



Socially-aware Management of New Overlay Application Traffic with Energy Efficiency in the Internet

European Seventh Framework Project FP7-2012-ICT- 317846-STREP

Deliverable D2.5 Report on Definition of Use-cases and Parameters (Final Version)

The SmartenIT Consortium

Universität Zürich, UZH, Switzerland
Athens University of Economics and Business - Research Center, AUEB, Greece
Julius-Maximilians Universität Würzburg, UniWue, Germany
Technische Universität Darmstadt, TUD, Germany
Akademia Gorniczo-Hutnicza im. Stanisława Staszica w Krakowie, AGH, Poland
Intracom SA Telecom Solutions, ICOM, Greece
Alcatel Lucent Bell Labs, ALBLF, France
Instytut Chemii Bioorganicznej PAN, PSNC, Poland
Interoute S.P.A, IRT, Italy
Telekom Deutschland GmbH, TDG, Germany

© Copyright 2015, the Members of the SmartenIT Consortium

For more information on this document or the SmartenIT project, please contact:

Prof. Dr. Burkhard Stiller
Universität Zürich, CSG@IFI
Binzmühlestrasse 14
CH—8050 Zürich
Switzerland

Phone: +41 44 635 4331
Fax: +41 44 635 6809
E-mail: info@smartenit.eu

Document Control

Title: Report on Definition of Use-cases and Parameters (Final Version)
Type: Public
Editor: Sylvaine Kerboeuf
E-mail: Sylvaine.Kerboeuf@alcatel-lucent.com
Authors: Valentin Burger, George Darzanos, Manos Dramitinos, Zbigniew Dulinski, Frédéric Faucheux, Gerhard Hasslinger, Fabian Kaup, Sylvaine Kerboeuf, Roman Lapacz, Ioanna Papafili, Patrick Poullie, Sabine Randriamasy, Michael Seufert, George D. Stamoulis, Rafal Stankiewicz, Matthias Wichtlhuber, Piotr Wydrych, Grzegorz Rzym, Krzysztof Wajda
Doc ID: D2.5-v1.0.

AMENDMENT HISTORY

Version	Date	Author	Description/Comments
V0.1	2014/09/24	S.Kerboeuf	ToC creation
V0.2	2014/11/13	S.Kerboeuf	ToC update
V0.3	2015/01/05	S.Kerboeuf et al	Description of use cases in chapter 3; inputs in chapter 4
V0.4	2015/03/27	S.Kerboeuf	sub-sections restructured in chapter 4 and 5
V0.5	2015/04/20	Valentin Burger, George Darzanos, Manos Dramitinos, Zbigniew Dulinski, Gerhard Hasslinger Fabian Kaup, Sylvaine Kerboeuf, Roman Lapacz, Ioanna Papafili, Patrick Poullie, Michael Seufert, Rafal Stankiewicz, Matthias Wichtlhuber	Input in chapters 4, 5 and 6
V0.6	2015/05/11 b	S.Kerboeuf	Introduction; summary of conclusion in chapter 6 Editing and comments of the document
V0.7	2015/05/20	Valentin Burger, George Darzanos, Manos Dramitinos, Zbigniew Dulinski, Gerhard Hasslinger, Fabian Kaup, Sylvaine Kerboeuf, Roman Lapacz, Ioanna Papafili, Patrick Poullie, Michael Seufert, Rafal Stankiewicz, Matthias Wichtlhuber	Inputs and addressed comments in section 3, 4 and 5
V0.8	2015/06/15	Manos Dramitinos, Gerhard Hasslinger, George Stamoulis	First review of document (preliminary version does not include all results of the work done during the extension period)
V0.9	2015/29/06	Valentin Burger, George Darzanos, Manos Dramitinos, Zbigniew Dulinski, Gerhard Hasslinger, Fabian Kaup, Sylvaine Kerboeuf, Roman Lapacz, Ioanna Papafili, Patrick Poullie, Michael Seufert, Rafal Stankiewicz, Matthias Wichtlhuber	Partial comments addressed
V0.10	2015/29/06	Valentin Burger, George Darzanos, Manos Dramitinos, Zbigniew Dulinski,	Editing work and input for section 4.4

		Gerhard Hasslinger, Fabian Kaup, Sylvaine Kerboeuf, Roman Lapacz, Ioanna Papafili, Patrick Poullie, Michael Seufert, Rafal Stankiewicz, Matthias Wichtlhuber	
V0.11	2015/07/10	Valentin Burger, George Darzanos, Manos Dramitinos, Zbigniew Dulinski, Gerhard Hasslinger, Fabian Kaup, Sylvaine Kerboeuf, Roman Lapacz, Ioanna Papafili, Patrick Poullie, Michael Seufert, Rafal Stankiewicz, Matthias Wichtlhuber	Completed version ready for second review
V0.12	2015/07/20	Manos Dramitinos, Sylvaine Kerboeuf, Ioanna Papafili, George Stamoulis	Address general comments of internal reviewers in all sections
V0.13	2015/07/20	Frédéric Fauchaux, Gerhard Hasslinger, Sabine Randriamasy	Completion of results for MUCAPS and for caching strategies
V1.0	2015/07/27	Valentin Burger, George Darzanos, Manos Dramitinos, Zbigniew Dulinski, Gerhard Hasslinger, Fabian Kaup, Sylvaine Kerboeuf, Roman Lapacz, Ioanna Papafili, Patrick Poullie, Michael Seufert, Rafal Stankiewicz, Matthias Wichtlhuber	Address comments of internal reviewers – contribution finalization

Legal Notices

The information in this document is subject to change without notice.

The Members of the SmartenIT Consortium make no warranty of any kind with regard to this document, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. The Members of the SmartenIT Consortium shall not be held liable for errors contained herein or direct, indirect, special, incidental or consequential damages in connection with the furnishing, performance, or use of this material.

Table of Contents

Table of Contents	4
List of Figures	7
List of Tables	11
1 Executive Summary	12
2 Introduction	17
2.1 Brief summary of past achievements of the project	17
2.2 Purpose of the Document	18
2.3 Document Outline	19
3 SmartenIT Use Cases	21
3.1 Methodology	21
3.2 Stakeholders	22
3.3 SmartenIT Use Cases	23
3.3.1 <i>Bulk data transfers for cloud operators</i>	23
3.3.2 <i>Host resource allocation in cloud federations</i>	25
3.3.3 <i>Video content transfer between storages of independent clouds</i>	27
3.3.4 <i>IoT data transfer for cloud operators</i>	31
3.3.5 <i>Service and content placement for users</i>	32
3.3.6 <i>Exploiting content locality</i>	34
3.3.7 <i>Social-Aware mobile data offloading</i>	35
3.3.8 <i>Access Technology Selection for Users</i>	37
3.3.9 <i>SDN based DC server selection</i>	39
3.4 Use cases summary	41
4 Parameters, Metrics and Evaluation of SmartenIT mechanisms	43
4.1 Parameters, Metrics and Evaluation Results of SmartenIT mechanisms for OFS43	
4.1.1 <i>DTM</i>	44
4.1.2 <i>Inter-Cloud Communication (ICC)</i>	56
4.1.3 <i>Multi-Resource Allocation (MRA) dependencies</i>	61
4.1.4 <i>DTM++</i>	67
4.1.5 <i>Model for Cloud Federation: Investigation of Pricing Aspects</i>	78
4.2 Parameters, Metrics and Evaluation of SmartenIT mechanisms for EFS	84
4.2.1 <i>RB-HORST</i>	85
4.2.2 <i>SEConD</i>	89
4.2.3 <i>vINCENT</i>	92
4.2.4 <i>MoNA</i>	94
4.2.5 <i>RB-HORST ++</i>	97
4.2.6 <i>MUCAPS</i>	105

4.3	Summary of theoretical models supporting TM mechanisms' evaluation	111
4.4	Main outcome of TM mechanisms evaluation	113
4.5	Key metrics and parameters of SmartenIT TM mechanisms	118
4.6	Key design goals of TM mechanisms for SmartenIT use cases	122
5	Tussle analysis of SmartenIT ecosystem	126
5.1	Tussles in the Operator Focused Scenario	127
5.1.1	<i>Stakeholders</i>	127
5.1.2	<i>Tussles identified and resolved by SmartenIT</i>	128
5.1.3	<i>Relation with SmartenIT Business Models</i>	132
5.2	Tussle in the End-user Focused Scenario	132
5.2.1	<i>Stakeholders</i>	133
5.2.2	<i>Tussles identified and resolved by SmartenIT</i>	134
5.2.3	<i>Relation with SmartenIT Business Models</i>	137
5.3	Tussles across Operator Focused and End-user Focused scenarios	137
5.3.1	<i>Synergy of DTM and RBH</i>	137
5.3.2	<i>Synergy of ICC and SEConD (AUeB)</i>	139
5.3.3	<i>Tussles identified and resolved by SmartenIT</i>	144
5.3.4	<i>Relation with SmartenIT Business Models</i>	146
5.4	Tussles in Cloud Federations	146
5.4.1	<i>Stakeholders</i>	147
5.4.2	<i>Tussles identified and resolved by SmartenIT</i>	147
5.4.3	<i>Federation policies and rules</i>	148
5.4.4	<i>Relation with SmartenIT Business Models</i>	148
5.5	Summary of tussle analysis	149
6	Summary and Conclusions (ALBLF)	150
6.1	Coverage of SmartenIT aspects	150
6.2	Lessons learnt	154
6.3	Future work	158
6.3.1	<i>Further evaluations of TM mechanisms</i>	158
6.3.2	<i>Applicability of SmartenIT solutions in Future Internet and 5G</i>	159
7	SMART Objectives Addressed	161
8	References	165
9	Abbreviations (ALL)	169
10	Acknowledgements	171
11	Appendices	172
11.1	Appendix A: DTM++ detailed specification	172
11.2	Appendix B: Model for QoE of Mobile Video Streaming	179
11.3	Appendix C: Cloud federation- Model and Pricing	180
11.3.1	<i>Cloud Service Provider Modeling</i>	180
11.3.2	<i>Federation Policies</i>	181
11.4	Appendix D: Evaluation of the energy efficiency of Multipath TCP	187
11.4.1	<i>Background and related work</i>	187
11.4.2	<i>Measurement setup:</i>	189

11.4.3 <i>Measurement results</i>	192
11.4.4 <i>Power model</i>	198
11.4.5 <i>Conclusion</i>	202
11.5 Appendix E - Options and Performance Evaluation of Caching	204
11.5.1 <i>Access Pattern for Web Content and Efficient Caching Strategies</i>	204
11.5.2 <i>Content delivery simulation framework</i>	204
11.5.3 <i>Efficiency Study of Web Caching Strategies Combining LFU and LRU</i>	206
11.5.4 <i>Content Popularity Dynamics</i>	212
Extended Request Model for Popularity Dynamics Including New Items	212
Conclusions	214
11.6 Appendix F: Pricing	216
11.6.1 <i>Pricing Layers</i>	216

List of Figures

Figure 2-1: Prioritization of UCs to be considered.	18
Figure 2-2: Prioritization of TM mechanisms to be considered.	19
Figure 3-1: Taxonomy of cloud business model according to Forrester Research Inc.	23
Figure 3-2: The spectrum of Cloud computing services. (Source: Gene Phifer, Managing Vice President at Gartner).....	25
Figure 3-3: The Cloud systems taxonomy and features [20].....	26
Figure 3-4: Appirio focus on brokering and building differentiated services.	27
Figure 3-5: Cloud brokers' services and importance [32].....	28
Figure 3-6: From low-value commodity services to high-value business solutions [29].	29
Figure 3-7: Cloud business and technical evolution [33].....	29
Figure 3-8: The Amazon repeatable business model [36].....	32
Figure 3-9: The TMforum transformation of the Amazon BM for cloud services [36].	33
Figure 3-10: The TM Forum Ecosystem Enablement Platform and sample service instantiations on top that ISPs could provision – but also OTTs [36].....	33
Figure 3-11: The Arzuna Agility high-level approach [39].	35
Figure 3-12: The Arzuna Agility approach materialized via SLAs [39].	36
Figure 3-13: The Deutsche Borse Cloud Exchange [44].....	37
Figure 3-14: Mapping of SmartenIT mechanisms to business models.	39
Figure 3-15: Mapping of SmartenIT theoretical investigations (presented in Chapter 4) to the business models that motivate them.	40
Figure 4-1: A single Cloud Provider as an M/M/1 queueing system	42
Figure 4-2: The two Cloud Providers operating separately.	43
Figure 4-3: One way federation of two cloud providers.....	44
Figure 4-4: Values of α^* in d, for $\lambda_1 = 2\lambda_2$ and $\mu_1 = \mu_2$	45
Figure 4-5: Full federation of two cloud providers.....	46
Figure 4-6: Values of α and β that minimize the average delay of total system, for $d=0$, $\lambda_1=2\lambda_2$ and $\mu_1=\mu_2$,.....	47
Figure 4-7: Global optimal (blue line) vs. P_1 optimal (red line) (α^* in d, $\lambda_1=2\lambda_2$ and $\mu_1=\mu_2$)....	48
Figure 4-8: Function for the dimension age.	51
Figure 4-9: 16bit AS address room.	54
Figure 4-10: 32bit AS address room.	54
Figure 4-11: Zipf-law of the number of active IP-addresses in autonomous systems.	55
Figure 4-12: Probability distribution of the number of active IP-addresses per AS.	56
Figure 4-13: Hotspots and Intersections of London.	56
Figure 4-14: Distribution of Hotspots with Respect to the Distance to the City Center.	57
Figure 4-15: London Angular Distribution.	58
Figure 4-16: Measurement setup and its wiring for the Raspberry Pi.	60
Figure 4-17: Power Consumption vs. CPU utilization.	61
Figure 4-18: Ethernet power consumption during download (left) and upload (right) vs. used bandwidth.....	62
Figure 4-19: WiFi power consumption during download (left) and upload (right) vs. used bandwidth.....	62
Figure 4-20: Virtual node concept.....	69

Figure 4-21: Different scheduling strategies. Lower numbers indicate pieces with higher priority. The base layer is at the bottom. 71

Figure 4-22: Workload [87]. 73

Figure 4-23: Main evaluation results (5 experiments, 95 % confidence intervals). 74

Figure 4-24: Evaluation of token trading cost..... 76

Figure 5-1: Example ISP and cloud/DC link topology. 79

Figure 5-2: Network prefixes announced by BGP for DTM via selected inter-domain links. BGP routing policy selectively puts prefixes in the routing tables for the DTM purposes 80

Figure 5-3: Tunnels used by DTM. End-points of each tunnel are addressed with different network prefix this results in different inter-domain link traversing by each tunnel. . 80

Figure 5-4: The network resources used by the DTM mechanism..... 81

Figure 5-5: Distribution of observed 95th percentile sample pairs, measured during the billing period; values of achieved traffic vector, reference vector, and related inter-domain traffic costs for the experiment without DTM (red) and with DTM (blue):(a) success rate 40%; (b) success rate 30%. 84

Figure 5-6: Sorted 5-minute samples with (blue) and without (red) DTM compensation and corresponding 95th percentile thresholds, link 1 and link 2, success rate 40%..... 85

Figure 5-7: ICC system architecture. 87

Figure 5-8: The Best Effort traffic pattern and 95th percentile. 92

Figure 5-9: The ICC_FAP traffic pattern, the 95th percentile and the extra delays incurred. 92

Figure 5-10: The ICC_FA traffic pattern, the 95th percentile and the extra delays incurred. 93

Figure 5-11: Comparison of traffic patterns attained under Best Effort, ICC_FA, ICC_FAP. 93

Figure 5-12: The Best Effort Internet traffic pattern (without ICC)..... 93

Figure 5-13: The ICC_FA traffic pattern, the 95th percentile and the extra delays incurred. 93

Figure 5-14: Comparison of traffic patterns without ICC and with ICC_FAP..... 94

Figure 5-15: Comparison of traffic patterns attained under ICC_FA and ICC_FAP. 94

Figure 5-16: Reminder on DTM operation over the inter-DC bulk data transfer scenario. 101

Figure 5-17: The rationale of the combined operation of DTM and ICC. 102

Figure 5-18: Traffic shaping in Link 1 performed by ICC after DTM is run for a trace of 1 week duration..... 104

Figure 5-19: Traffic shaping in Link 1 performed by ICC after DTM is run for a trace of 1 week duration..... 104

Figure 5-20: Distribution of observed 95th percentile sample pairs, measured during the billing period of 7 days: experiment with DTM (blue) and with DTM++ (green)... 105

Figure 6-1: Process-diagram of a WATCH-event. 113

Figure 6-2: City area of Darmstadt with WiFi access point locations (red) and end-user locations (blue)..... 114

Figure 6-3: Total ISP cache contribution and mean ISP cache hit rate dependent on home router sharing probability for different ISP cache capacities. 116

Figure 6-4: Total amount of requests served intra-AS dependent on home router sharing probability for different ISP cache capacities. 117

Figure 6-5: Total ISP cache contribution and mean ISP cache hit rate dependent on home router sharing probability for different content popularity distributions with Zipf-exponent α 117

Figure 6-6: Total amount of requests served intra-AS dependent on home router sharing probability for different content popularity distributions with Zipf-exponent α	118
Figure 6-7: ISP cache hit rate for the five autonomous systems with most end-users dependent on home router sharing probability for different ISP cache capacities.	119
Figure 6-8: WiFi offloading potential for three WiFi access point seding ranges dependent on the WiFi sharing probability.	120
Figure 6-9: The messaging overlay constructed for an interest category.	122
Figure 6-10: An example of the prefetching algorithm – The source node shares a video hosted in a third-party owned server. (Sequence numbers arrows.).....	123
Figure 6-11: Total inter-AS traffic during a day.	127
Figure 6-12: HORST-vINCENT privacy layer and incentive layer.	129
Figure 6-13: Basic OMNET++ setup.	132
Figure 6-14: Visualization of analyzed traces.	133
Figure 6-15: Schematic overview of the MoNA mechanism.....	134
Figure 6-16: Power Consumption of the Nexus S splitting a fixed rate data stream to multiple interfaces (3G, WiFi).	137
Figure 6-17: The MACAO Request Service: function, interface and client.....	138
Figure 6-18: Deployment of MUCAPS in the ISP network.	139
Figure 6-19: Steps involved in MUCAPS and details on the AMWS (Application Metrics and Weights Selection) process.....	140
Figure 6-20: MUCAPS demonstration and functional evaluation setup	143
Figure 6-21: Chunk-based representation of a video.....	148
Figure 7-1: Topological view of architecture with added scenario overlay (based on component map taken from D3.1[7]).....	151
Figure 13-1: Inter cloud communication - multi cloud transfer to single cloud.....	173
Figure 13-2: Inter cloud communication - single cloud transfer to multi cloud.....	175
Figure 13-3: Available Internet path determined by BGP routing policy (example 1).....	177
Figure 13-4: Available tunnels resulting from routing policies (refers to Figure 13-3).	177
Figure 13-5: Available Internet path determined by BGP routing policy (example 2).....	178
Figure 13-6: Available tunnels resulting from routing policies (refers to Figure 13-5).	178
Figure 13-7: Available Internet path determined by BGP routing policy (example 3).....	179
Figure 13-8: Available tunnels resulting from routing policies (refers to Figure 13-7).	179
Figure 13-9: The network nodes taking part in the DTM communication and data transfer between partner ISPs.....	183
Figure 13-10: The simplified DTM for inter cloud communication - network topology, network nodes and tunnels (system release 1.0).	186
Figure 13-11: Cost functions.	187
Figure 13-12: Initial communication between SmartenIT components in ISP-A domain..	189
Figure 13-13: Initial communication between SmartenIT components in ISP-B domain..	190
Figure 13-14: The DTM operation sequence diagram with only NetFlow link monitoring.	193
Figure 13-15: The DTM operation sequence diagram with SNMP and NetFlow link monitoring.	194
Figure 13-16: The areas used in the reference vector optimization procedure. Vectors $S1$ and $S2$ determine the range in which optimal solution may be found.....	196
Figure 13-17: The environment for the DTM tests including connection points for traffic generators and receivers.	201

Figure 13-18: The operation of the ICC mechanism when datacenters are non-federated.203

Figure 13-19: The ICC mechanism cloud layer when datacenters are federated. 205

Figure 13-20: SmartenIT System Architecture as shown in[9]. 211

Figure 13-21: Sequence Diagramm of the MRA Mechanism. 211

Figure13-22: SMT components (in orange) involved in RBH. 215

Figure 13-23: The Facebook OAuth process. 218

Figure 13-24 Process of creating or joining the overlay network. 219

Figure 13-25: Maintenance of overlay network executed in a fixed time interval t. 220

Figure 13-26: The tracking process is executed periodically for each content shared. 221

Figure 13-27The Overlay Based prediction is initiated by the CTM and executed in the OM. 221

Figure 13-28: Facebook information retrieval. 222

Figure 13-29: Social prediction uses information stored in the Social Monitor. 223

Figure 13-30: To find the closest UNaDa a traceroute is requested from all providers. ... 223

Figure 13-31: Before the app can be used the user has to be authenticated with Facebook and the user ID will be stored. 224

Figure 13-32: The mobile App has the Facebook ID of the user stored and uses this to authenticate with the foreign UNaDa. 225

Figure 13-33: The messaging overlay constructed for an interest category. 227

Figure 13-34: Steps of the SEConD’s prefetching algorithm. 228

Figure 13-35: The sequence diagram demonstrates the communication between a user and the local SPS during initialization time. 234

Figure 13-36: Sequence diagram of the prefetching algorithm. 235

Figure 13-37: PING messages ensure that the FacebookAPP always has an updated view of the uNadas online, their location, and their social relations. 236

Figure 13-38: Tunnel setup procedure sequence diagram. 237

Figure 13-39: Assumptions on regressors and regressands. 241

Figure 13-40: $I(r,t)$ determines the height of the incentive when deviating from the average that can be expected from the regression model. 242

Figure 13-41: Mockup of page design. 243

Figure 13-42: Overall structure of the Mobile Network Assistant (MoNA) and external interfaces. 244

Figure 13-43: Details of the Network Optimizer. 245

Figure 13-44: Sequence diagram of the NetworkOptimizer. 246

Figure 13-45: Components of the Network Optimizer. 247

Figure 13-46: Basic Scheduling decision. 249

Figure 13-47: Execution of scheduled connections. 250

Figure 13-48: MACAO with MARC embedded in an AEPG located in the ISP network. .253

Figure 13-49: Details on the functions involved in MUCAPS and operations sequencing.255

Figure 13-50: MARC integrated in a network located AEPG, here a DNS server. 258

Figure 13-51: Interaction between entities involved in the MACAO AEP ranking 258

List of Tables

Table 2-1: Selection criteria for TM mechanisms to be implemented.	18
Table 2-2: Additional properties (functionalities) for the extension of the DTM and RB-HORST TM mechanisms leading to the formation of synergetic solutions.....	19
Table 4-1: Parameters of content demand model with temporal dynamics	49
Table 4-2: Functions for the considered dimensions.	51
Table 4-3: Top five autonomous systems with most active IPs.....	55
Table 4-4: Goodness of Exponential Fits of the Distance Distribution.	57
Table 4-5: The power models of the Raspberry Pi. Here, u is the CPU utilization in the range 0 to 1 and r the traffic rate in Mb/s.	63
Table 4-6: Bandwidth distribution of network model based on [94] (OECD) and [101] (VNI).71	
Table 4-7: SVC bandwidth per layer [87].	73
Table 4-8: Evaluated parameters; optimal settings are underlined.	73
Table 5-1: Parameters and metrics associated to the DTM mechanism.....	83
Table 5-2: Key parameters for ICC.	89
Table 6-1: Cahing strategies employed by RB-HORST.....	110
Table 6-2: Demand models employed for the evaluation of RB-HORST.	111
Table 6-3: Default parameters of the content delivery simulation framework.....	114
Table 6-4: Proxy server and SPS contribution w.r.t. the AS size.	127
Table 6-5: For each example AEP: utility vector of normalized relative performances and value for 2 utility functions and optimal AEP (= AEP1).....	144
Table 6-6: Demonstration scenario with 3 cases.	145
Table 7-1: Mapping of TM mechanism and single-scenario synergies to the layers and objectives of the SmartenIT field.	152
Table 9-1: Overall SmartenIT SMART objectives addressed [1].....	160
Table 9-2: Specific SmartenIT SMART objectives addressed; excerpt from the set of Tables of [1] with all SMART objectives of the project.	161
Table 13-1: Integration of the implemented functions of RB-HORST in the SmartenIT architecture components.....	216

1 Executive Summary

Deliverable D2.5 “Report on Definition of Use cases and Parameters” provides the final results of Work Package 2 “Theory and Modelling” within the ICT SmartenIT Project 317846. According to the Description of Work of the project, the main objectives of this deliverable and the related achievements reported cover both, the Operator Focused Scenario (OFS) and the End-user Focused Scenarios (EFS) of the project:

Objective 1: to provide the final specification of relevant use-cases that will highlight the efficiency of SmartenIT traffic management mechanisms within the framework of the two scenarios OFS (Operator-focused) and EFS (End user-focused).

The use cases were inspired by actual needs and business cases of the Internet-Cloud services ecosystem and built on the two complementary business perspectives of the ecosystem addressed by the OFS and EFS scenarios. A template for the description of use cases was defined to enable a well-organized and structured presentation of the different use cases. For each use case, this template contains the goal, a value network configuration diagram that highlights relations between stakeholders and their roles, the interests of the various stakeholders, and incentives, the SmartenIT Traffic Management (TM) mechanisms applicable and the overall innovation introduced by the use case and its associated TM mechanisms. The use cases developed and studied aim to provide proper incentives to all stakeholders involved and together with the applicable TM mechanisms lead them to win-win situation.

In particular, the use cases related to the OFS are mainly driven on one hand by the collaboration among Cloud Service Providers (CSPs), possibly in the form of a federation, aiming at improving performance and/or cost-effectiveness by means of resource sharing among them, and on the other hand by ISPs offering the underlying traffic management as a service to the CSPs. Thus, the goal of the stakeholders of the OFS use cases is to minimize their costs, which may include inter-connection charges for ISPs due to traffic generated by cloud services and applications, and operating cost in terms of connectivity charges and energy cost for Cloud Service Providers/Data Center Operators. Different ways to achieve this goal rely on inter-cloud communication, cloud federation and data replication and migration, while taking into account the often competing interests of CSPs, CDNs and the ISPs, especially in terms of transit traffic and the associated cost. Moreover, for certain OFS use cases, QoE improvement is among their goals, or it is an indirect effect arising.

The use cases related to EFS are mainly driven by a collaborative traffic management approach with a direct involvement of the end-users and their resources. The goal of the EFS use cases is to provide improved QoE and energy efficiency for the end-user, by intelligent placement and movement of content and services in an energy and socially aware manner, whilst the interest of the ISP on traffic management and respective costs is also considered. In particular, ISPs aim at saving network resources and at achieving a reduction of their inter domain traffic. These EFS use cases focus on user’s needs and employ a broad set of services to better serve them.

The study of these diverse use cases reveals that SmartenIT TM solutions are broadly applicable. They can be adopted in a wide variety of cases, and can lead to significant

benefits for all the stakeholders involved. Also, since the use cases are closely related to the SmartenIT TM mechanisms, they also serve as the basis for the definition of experiments of WP4 “Evaluation and Prototypes”, to adequately reflect all major aspects of SmartenIT, namely, incentive compatibility among stakeholders, energy efficiency, social- and QoE-awareness in all WP4 validation activities.

Objective 2: to provide the final evaluation of the mature specifications of the incentive-based cloud traffic management mechanisms and synergetic solutions, that will complete the results from D2.4 on Report on Final Specifications of Traffic Management Mechanisms and Evaluation Results.

Theoretical and simulation models are used to evaluate aspects of Traffic Management (TM) mechanisms and synergetic solutions not studied so far in the project and thus attain a more complete assessment. The completion of the evaluation of the TM mechanisms covering the OFS scope leads to the following new major outcomes reported in this deliverable:

Simulation experiments show that Dynamic Traffic Management (DTM) is able to manage properly the traffic and distribute it among the transit links of a multi-homed ISP as desired to optimize traffic driven costs for volume based tariff and for 95th percentile tariff. In the case where the 95th percentile tariff applies, however the DTM algorithm is sensitive to the traffic profiles. The means to provide DTM with scalability, reliability and security have also been discussed.

Further evaluation of the ICC (Inter Cloud Communication) mechanism for the incentive-compatible cost-efficient use of ISP’s transit links shows that ICC using statistics to predict traffic patterns performs well (max +/-10% deviation from target goal), often achieving for the ISP transit charge a higher discount than the one original sought. Also the way of incentive-compatible sharing of profits among the stakeholders involved in ICC is studied, and alternative pricing schemes are proposed and evaluated.

Regarding the MRA (Multi Resource Allocation) mechanism, a study of the interdependency of multiple heterogeneous resources (such as CPU, RAM, disk space, and bandwidth) in a cloud (or cloud federation) was carried out. This study has shown that CPU and RAM utilization does not depend in static ratios. Furthermore, also VM performance may suffer, when the VM is equipped with additional VCPUs, which is counter intuitive. Also other findings contract what is currently assumed in literature, when reasoning about the fair multi-resource allocation in clouds. Thus, these results justify the need for a fair cloud resource allocation policy as was presented by SmartenIT in form of the greediness metric. DTM++ comprises a synergetic solution integrating DTM with ICC. It combines the so-called traffic “shift in space” of DTM and traffic “shift in time” of ICC to achieve further optimization of traffic distribution across multiple transit links while delaying delay-tolerant traffic when transit links are heavily loaded, so as to ultimately achieve even lower transit charges for the ISP than DTM alone.

Innovative economic models of the federated environment of CSPs are defined, considering both the case where CSPs act cooperatively and non-cooperatively. Cloud federation increases revenues from QoS-based pricing on customers and reduces the energy consumption cost, thus the total profits of federated CSP is increased. Moreover,

cloud federation improves the global QoS in the federated environment, i.e., the average delay of served requests is decreased and throughput is increased. Moreover, proper mechanisms for *pricing and compensation* among CSPs are introduced and investigated, in order to ensure that all CSPs actually benefit from the formation of a federation.

The completion of the evaluation of TM mechanisms covering the EFS leads to the following new major outcome reported in this deliverable:

RB-HORST (Replicating Balanced- tracker and Home Router Sharing based on Trust) evaluation and refinement of caching strategies including *flash crowd scenario* are focused on performance evaluation when demand for content exhibits fast temporal dynamics. The results on WiFi offloading in an urban environment show that 66% of the connections can be offloaded on the average, if only 10% of WiFi access points are shared, assuming a sending range of 50m, thus showing the high potential of this approach.

Further evaluation of SEConD (Socially-aware mechanism for Efficient Content Delivery), as reported in the present deliverable, shows that, in the experimental set-up considered, SEConD can attain a significant reduction (even up to ~87%) of the total inter-AS traffic compared to inter AS traffic generated when applying the client-server paradigm in this set-up, thus advocating that it is a promising TM mechanism. Evaluation of its joint application with RB-HORST shows the potential of QoS-based user-assisted video delivery as a means to boost users' QoE.

Further evaluation of the vINCENT (virtual incentive) extension for RB-HORST is also performed, addressing mobile offloading from cellular communication to WiFi, leveraging social data to do so in an ISP friendly way. Results address the asymmetries between *rural areas vs. city areas* of an offloading scheme to derive fair incentives for all users. Despite different user densities, the proposed scheme is proven fair to all users.

Further evaluation of MoNA (Mobile Network Assistant) is performed, considering traffic on the cell interface and WiFi access points with data rates derived from a measurement study in the order of 1-10 Mbps. It shows that aggregation of traffic or deferring transmissions until a more energy efficient connectivity option (such as WiFi) is available may reduce power consumption: between 34% and 85% of the otherwise consumed energy can be saved. Joint evaluation with RB-HORST++ investigates the energy consumption for mobile video streaming sessions and shows that the minimum energy consumption for the video download compared over all available network technologies is attained when all connections are offloaded to WiFi, which can only be achieved for high WiFi data rates, which are not currently available in the WiFi access point deployed in streets.

Further investigations of MUCAPS (MULTi-Criteria Application endPoint Selection) show that for a reliable evaluation of its impact, it is necessary to take into account contextual elements such as at least the access type and capabilities of the User End Point (UEP) and the network conditions, or simply user expectations on QoE. In addition, the shortest AS-paths are not necessarily the best ones in terms of delay, resources and ISP costs: these three criteria may even be conflicting and a safe way for efficient layer cooperation is to consider them jointly in application resource endpoint selection, e.g. in the choice of server to get an application from.

The most important metrics employed for the evaluation of each TM mechanism as well as the most important/influential parameters are identified from the evaluation already carried out, and they will be used in the evaluation of test cases in WP 4. Values of parameters have been investigated and tuned by means of simulations. TM mechanisms that address OFS scenario by definition operate on large traffic aggregates and not on individual flows, thus they are mainly assessed based on metrics expressing inter-domain traffic and transit cost reduction.

On the other hand, TM mechanisms addressing EFS demonstrate a larger variety including metrics associated with caching, metrics related to energy consumption/energy efficiency, and metrics reflecting QoE perceived by the end users. It has also been observed that the set of key parameters of TM mechanisms demonstrates an even larger variety, as different mechanisms employ different methods and specifically designed algorithms to achieve the targets of SmartenIT per use case.

For instance, OFS TM mechanisms of ICC, DTM and their combination DTM++ are addressing inter-cloud communication focusing on the network level, and take into account in their decisions important parameters such as traffic patterns, cost functions and time slotting, while MRA considers CPU stress and workload type, as it focuses on the cloud level.

On the other hand, EFS mechanisms employ important parameters and features of caching strategies and user interests concerning caching of content, sharing probability and cache contribution/participation regarding resource sharing, device type, bitrate and network availability in order to achieve energy efficiency.

For each TM mechanism, proper values of the corresponding parameters that depend on the circumstances and ensure effective operation of the mechanism, are given.

Finally, an indicative *tussle analysis* of the SmartenIT ecosystem is performed and exhibits how conflicts of interest of stakeholders in SmartenIT scenarios and use cases can be mitigated by means of the incentive compatible SmartenIT TM mechanisms according also to the project's theoretical investigations and analysis of models. This analysis has also been explicitly mapped to the *business models* that are relevant for SmartenIT by identifying the business models related to each case of tussle analyzed. In practice, the synergy of OFS and EFS mechanisms may inherently resolve many (if not most) of the tussles identified in the individual scenarios, but it may also generate new tussles (that are not mitigated with the present TM mechanisms). Such tussles arise due to the potential interaction among stakeholders that were not considered to co-exist in the individual scenarios, e.g., a CSP and an end-user that owns a user controlled Nano Data center (uNaDa). These new tussles can be possibly mitigated by means of the careful parametrization of the scope and operations of each of the SmartenIT mechanisms constituting the synergy, applying appropriate business models that allow fair competition throughout the respective value chain segments. This relation is also evident in the SmartenIT pricing classification framework, provided as Appendix, where we present the different layers and granularities of pricing mechanisms related to SmartenIT, their scope and operation and how they work together. An example of a SmartenIT pricing mechanism for the OFS scenario, applicable to the respective inter-cloud communication use cases is

provided, and it is also related to the Repeatable DSP, Federation and ISP Managed Services business models which were presented in Deliverable D2.4.

In the conclusion of this deliverable, we summarize the outcome of the aforementioned results and we justify the complete coverage of the SmartenIT targets and objectives. All TM mechanisms addressing the EFS employ modules operating in the cloud layer, while DTM and ICC also operate in the network layer and both address inter-domain traffic. Additionally, MRA is a collaborative mechanism, which also pursues incentive-compatibility, while DTM, ICC and their combination DTM++ are mechanisms whose impact is observable in larger time-scales, i.e., monthly period to derive 95th percentile values. Mechanisms addressing the EFS expectedly operate in the end-user layer. Among EFS mechanisms, RB-HORST, SEConD, and vINCENT are collaborative and incentive-based, while the two first employ social awareness and consider inter-domain traffic. RB-HORST, SEConD and MUCAPS are aware of QoE, while MoNA and vINCENT address needs of mobile users, with MoNA being the mechanism also addressing energy efficiency at the end-user layer, considering the heterogeneous availability and performance of mobile data access.

Overall, as revealed from this deliverable, as well as from the entire work of WP2, the SmartenIT TM mechanisms are effective, incentive-compatible, and in line with recent evolutions in business models and networking and cloud technologies and services as well as with best practices and in Internet services markets, and thus can prove very good solutions leading to significant benefits even in practical cases involving commercial networks. Their final evaluation will be carried out in WP4, input for which is provided in the present deliverable regarding metrics to be employed and values of parameters that are important for effective operation of the mechanisms.

2 Introduction

SmartenIT (Socially-aware Management of New Overlay Application Traffic combined with Energy Efficiency in the Internet) targets the *incentive-compatible cross-layer* resource management for network and cloud operators, Cloud Service Providers, and end-users [1]. The project aims to address this challenge by means of novel mechanisms for resource and traffic management of cloud and Internet services pertaining to use cases of high market value and impact, including yet not limited to inter-cloud communication and applications requiring end user *Quality-of-Experience (QoE)-awareness*. One key aspect of the investigated cross-layer traffic management is *social awareness*, i.e. the exploitation of users' social relationships and interests, which represents an extra source of information to characterize the end-users of the cloud services, and to predict future demand. Another key aspect of the SmartenIT advanced traffic management is *energy efficiency* with respect to both networking and cloud infrastructure, as well as wireless and wired end-user devices, where the key is to find a balance between QoE and lower energy usage both in the Operator-/Provider side and the End-User side.

This deliverable provides the final specification of relevant use cases that highlight the applicability, importance and efficiency of the SmartenIT traffic management mechanisms within the framework of the two key scenarios of the project, namely the Operator Focused Scenario (OFS) and the End-user Focused Scenario (EFS). It also finalizes the evaluation along with the complete specifications of the incentive-based cloud traffic management mechanisms and synergetic solutions that complete the results from D2.4 on Report on Final Specifications of Traffic Management Mechanisms and Evaluation Results [7]

2.1 Brief summary of past achievements of the project

An initial set of use cases (UCs), based on the scenarios definition in D1.2 [3] was specified in the deliverable D2.3 [6] to illustrate situations in which incentive-based cross-layer traffic management (TM) mechanisms are used. The goal, the relations between stakeholders, and their role and interests in the use case were identified and accompanied by a business analysis and investigation of incentives for each stakeholder, followed by the identification of success indicators of the use case. An initial selection and prioritization of the UCs and mechanisms to be prototyped in WP3 was performed on a per scenario basis. Furthermore, a methodology was proposed to describe parameters and metrics to be used to describe experiments for the validation of the performance of SmartenIT solutions deployed across all use cases. Parameters have been divided into two classes: *global parameters* that are high-level and not tailored to a specific traffic mechanism, and *specific parameters* of the relevant traffic management algorithms. Parameters and metrics have been further classified to the layer they pertain to, i.e. the network, the cloud or the end-user/application layer. Preliminary definitions of metrics included QoE metrics at end-user and energy efficiency metrics either in the operator network and devices or in the end-user devices, depending on the mechanisms.

In parallel, the existing traffic management approaches in the literature was overviewed and assessed in the deliverable D2.1 [4]. Design and specification of management mechanisms that handle inter-cloud traffic (cost-) effectively, some of which employ social-

and QoE-awareness, while possibly achieving energy efficiency were introduced into deliverable D2.2 [5] and pursued in D2.4 [7]. The mapping of each single mechanism to the components and interfaces of the SmartenIT architecture defined in D3.1 [8], as well as an example instantiation of its operation have been provided. Investigations on business models of cloud services were carried out to bring a deeper understanding of the technical and economic dependencies among the stakeholders of the SmartenIT playfield. TM mechanisms were assessed by means of both theoretical evaluations and simulations covering a wide variety of aspects of the traffic management mechanisms. Moreover, combinations (synergies) of multiple TM mechanisms were initiated so as to cover the most of the SmartenIT targets and objectives.

2.2 Purpose of the Document

D2.5 deliverable concludes the above studies by providing results within the following aspects:

It identifies and describes the most relevant use cases where tangible benefits can be attained by means of SmartenIT, within the framework of OFS and EFS scenario. In particular,

- It provides a public version of the most relevant use cases within the framework of OFS and EFS scenarios where tangible benefits can be attained by means of SmartenIT, concluding the initial investigation of use cases provided in the project-internal deliverable D2.3 [6].
- It documents and assesses how SmartenIT solutions can handle new use cases that arise from the evolution of networks (e.g. with the introduction of Software Defined Network technology) and growth of new type of traffic (e.g. development of the Internet of Things). Those two new use cases are respectively presented in sections 3.3.9 and 3.3.4.

It completes the analysis of theoretical and simulation models so as to evaluate new variations of and aspects of TM mechanisms and synergetic solutions not studied so far and thus attain a more complete assessment thereof. In particular, it provides the extension of works presented in D2.4 [7] as described below:

- Further evaluations of the single TM mechanisms as follows:
 - Mechanisms related to OFS: DTM (Dynamic Traffic Management), ICC (Inter Cloud Communication), and MRA (Multi-resource Allocation), see sections 4.1.1, 4.1.2 and 4.1.3 respectively,
 - Mechanisms related to EFS: RB-HORST (Replicating Balanced- tracker and Home Router Sharing based on Trust) in section 4.2.1 including flash crowd scenario and refinement of caching strategy in section 11.5, SEConD (Socially-aware mechanism for Efficient Content Delivery) in section 4.2.2, vINCENT (virtual incentive) extension for RB-HORST in section 4.2.3, MoNA (Mobile Network Assistant) in section 4.2.4, and MUCAPS (MULTi-Criteria Application endPoint Selection) in section 4.2.6

- Final specification and evaluation results of synergetic solutions made of the combination of complementary TM mechanisms, as follows:
 - Further evaluation of DTM++ (in section 4.1.4.3), a synergetic solution combining the so-called traffic “shift in space” of DTM and traffic “shift in time” of ICC to achieve further optimization of traffic distribution across multiple transit links while delaying delay-tolerant traffic when transit links are heavily loaded, so as to ultimately achieve even lower transit charges for the ISP than original DTM/ICC standalone.
 - Evaluation of RB-HORST++ (in section 4.2.5) capitalizing on two synergy combinations: one combining trust-based home router sharing based on social observations by RB-HORST, and content caching, prefetching, and chunk-based dissemination by SEConD to address the content placement use case; and one combining trust-based home router sharing based on social observations by RB-HORST, and incentive-based reciprocation and energy efficiency by MONA to address the WiFi offloading use case.
- Results on new theoretical models developed for the evaluation of SmartenIT mechanisms. In particular:
 - Model of the formation of an economically sustainable *cloud federation*, with design of incentive compatible pricing mechanisms that achieve mutual benefits for CSPs (see section 11.3).
 - Model addressing the MoNA mechanism for both optimization of energy consumption and QoE at the mobile device side based on a selection algorithm of access network using MPTCP (see section 11.4).
 - Refinement of optimal caching strategies including *flash crowd scenario* (see section 11.5).
 - Model of pricing that could be supported by SmartenIT also in lieu of the SmartenIT business models presented in Deliverable 2.4 [7], including a SmartenIT pricing proposal for ICC mechanisms (see section 11.6).

D2.5 also completes the definition of parameters and metrics associated with SmartenIT solutions initiated in D2.3 [6], by providing the definition of metrics and refined range of parameters values based on results from the mechanisms evaluation. The applicability of SmartenIT solutions to Future Internet and 5G scenarios is also briefly discussed. Finally, D2.5 conducts tussle analysis of the SmartenIT ecosystem, for the OFS and EFS scenarios, for the synergy of both OFS and EFS, as well as for a SmartenIT model.

2.3 Document Outline

This document is organized as follows. **Section 3** provides a structured description of the use cases that both are relevant to and motivate the SmartenIT traffic management mechanisms with respect to the scenarios. **Section 4** presents the evaluation results of the SmartenIT mechanisms for the Operator Focused and End-user Focused Scenario, as well as results for the synergetic solutions and a summary of theoretical models which are detailed in appendices in **Section 11**. The parameters and metrics associated to the

SmartenIT mechanisms are also presented. **Section 5** provides a tussle analysis of the SmartenIT ecosystem, also demonstrating the correlation of tussles, use cases, mechanisms and business models. **Section 6** discusses the coverage of the SmartenIT objectives by the designed mechanism and draws major conclusions on the identified use cases and the evaluation of specified TM mechanisms. Finally, **Section 7** reports which SMART objectives, as described in SmartenIT Description of Work [1], have been addressed by the work performed in WP2 and reported in D2.5.

3 SmartenIT Use Cases

3.1 Methodology

In order to describe the use cases in a consistent manner, proper template has been defined and used throughout. The template provides a clear and structured presentation of all the details needed to understand how the SmartenIT solutions may be deployed to bring a tangible benefit (traffic management optimization) related to the overall project challenges. The form of the template is based on the experience of project partners and the focus of the project, while the main guidance on how to build the information structure has been taken from [11].

The template contains the following fields:

- Use case name,
- Goal – a short description what shall be achieved,
- Scenario – End-user Focused Scenario or Operator Focused Scenario,
- Stakeholders roles and their interests,
- Incentives for stakeholders with an assigned role,
- Multiple roles of stakeholders – an explanation of possible assignment of multiple roles to one stakeholder, e.g., a company can offer connectivity services (ISP role) and CSP services (Cloud Service Provider role),
- Value Network Configuration (VNC) – business and technical relationships among stakeholders ,Precondition – conditions that must be satisfied in order to make the main success scenario of use case possible to happen,
- Trigger – an action or an event after which the interactions described in the use case happen,
- Main success scenario – the sequence of interactions meeting the goal of the use case,
- Traffic management solution – one or more solutions proposed within the SmartenIT deliverable D2.4 [7] which are suitable for the use case,
- Innovation –what is innovative in the use case,
- Related information – additional information, which does not fit to the above fields but is considered as important to provide a complete view of the use case.

It is very important for the members of SmartenIT to analyze the applicability of the SmartenIT solutions in the real business market and emphasize their novelty. That is why the use cases contain the information regarding incentives for stakeholders, diagrams of Value Network Configuration (VNC) and the innovation aspect. Moreover, the use case template is directly linked to the traffic management mechanisms developed in the SmartenIT project. The field “Traffic management solutions” includes one or more mechanism names which are suitable for the use case.

3.2 Stakeholders

The main actors and their roles in the SmartenIT ecosystem, as they have been initially identified in Deliverable D1.1 [2], are as follows:

- Cloud Service Providers: Offer application services to users, basically following the SaaS paradigm, satisfying their QoS/QoE requirements.
- Data Center Operators: Merged role of data center operator and cloud provider for IaaS, i.e., own and manage one or multiple data centers
- Internet Service Provider: (Access and backbone network) offering connectivity service over both public and private networks. To be further distinguished into:
 - Transit network providers: Providers selling global Internet connectivity (wholesale market).
 - Edge network providers: Providers serving end-users of a region (retail market), customers of, or to be combined with transit network providers
- End-users (or just users) can be divided in two categories:
 - Business users: companies mainly interested in purchasing IaaS service (Computing and Storage) , connectivity of their locations via VPNs, Internet presence etc., often under specific SLA conditions
 - Residential users/Consumers: private users who purchase Internet connectivity and eventually cloud service for entertainment, gaming and content sharing purposes etc.
- Social Information Providers: Gathers social information and meta-information either in collaboration with an Online Social Network (OSN), or by means of crawling, and provides the derived information to whom may request it for some compensation. It may be identical to the OSN provider. This is part of the big data evaluation and derived services on large Internet platforms.

Two more actors considered in [2] are the Energy Provider and the Energy Broker in the context of cloud federations for energy efficiency. We do not further refer to them here, because they are not of central interest for SmartenIT. Nonetheless, energy efficiency is a goal of SmartenIT mechanisms.

The interactions among groups of actors in the SmartenIT ecosystem are investigated in two main areas which pertain to the wholesale and retail market of Internet services and represent the main SmartenIT scenarios defined in WP1:

- Operator Focused Scenario describes the interactions in the wholesale market, i.e. among SmartenIT operators, as mainly driven by the ultimate goal of achieving highest operating efficiency in terms of high Quality-of-Experience (QoE), low energy consumption, low operating cost and maximized revenues where applicable.
- End-user Focused Scenario describes the interactions in the retail market where end-users are active along with other SmartenIT actors, mainly driven by the ultimate goal of providing increased QoE and energy efficiency for end-user.

3.3 SmartenIT Use Cases

3.3.1 Bulk data transfers for cloud operators

Use case name	Bulk data transfers for cloud operators
<p>Goal</p>	<p>We consider the case of N clouds/data centers, which are in-general located in different ISPs. Their traffic is handled by their home ISPs, which are typically Tier-2 or Tier-3 ISPs, so for global Internet connectivity they rely on purchasing transit from Tier-1 ISP. Therefore the inter-domain traffic is typically charged under the 95-th percentile rule. The goal of this use case is to meet the inter-cloud communication needs of clouds/DCs in the most efficient way in terms of bandwidth and ISP transit costs.</p> <p>We assume that bulk data of the cloud are periodically replicated to a backup facility, i.e. another cluster located in a different physical location, for redundancy. The traffic itself can consist of static content of a Content Provider (e.g., CNN), or online personal storage of end-users (e.g., Dropbox) or content from an IaaS Provider (Virtual Machine snapshots). Such data transfer can be offered by the DC operator as an additional service to its customers, both IaaS and SaaS Providers who in turn bundle this service to their service offerings to the end-user, so as to guarantee a certain degree of geographical redundancy for improved resiliency. Such an additional service must be offered with a certain degree of granularity and be regulated by proper SLA parameters (Recovery Point Objective, RPO). In this context the inter-domain traffic that is generated can also be periodic with a frequency dictated by the contractual SLA. For example DC Operator could sell the service with RPO of 6hrs meaning that in case of primary site failure, only data older than 6 hrs can be recovered; in such a case the base frequency for the data replica is obviously every 6 hrs.</p> <p>An alternative instantiation of this use case would be that the periodic massive bulk data transfers are performed due to the need (and respective agreement) for sharing data, due to business or scientific purposes. For instance, astronomic observation data could be exchanged periodically among space agencies clouds so as to collaborate on scientific projects on deep space exploration.</p> <p>This use case also applies to a situation when an owner of storages (IaaS Provider) builds distributed architecture of its storage resources. Faster and thus more expensive storage is provided to have instant user access. Slower and cheaper storages keep data which are older or less important (less active) data. This data management approach is known as storage tiering (storage hierarchy). The automated migration of data between different types of disk does not have to exist in a single DC/IaaS cloud. Additional storage (groups of them – tiers) may be allocated in other DCs/clouds within an appropriate business agreement (e.g. federation of clouds).. This allows part of the data to reside on slower, larger, cheaper disks (Tier 2 or 3) whilst the most active data resides on the more efficient and expensive drives (Tier 1, e.g., Fibre Channel and flash SSD). The goal of the use case with storage tiers is to make optimized connections (in terms of bandwidth and cost) between storages located in different network domains to transfer less active data (the migration can be in both directions, which means the “hot” data can be moved again to Tier 1 from Tier 2). Such data migration between tiers results in decreasing the cost of data storage and better utilization of network resources. Data transfer does not have to be done in real time. This can adapt to a workload profile changes (migrations may be postponed).</p> <p>The data transfers envisioned in this use case can be the result of business</p>

	<p>agreements among federated clouds/data centers, which perform data transfers amongst their members in accordance with the federation business policy and rules, such as to perform load balancing, energy savings, and push services close to their consumers.</p>
<p>Scenario</p>	<p>Operator Focused Scenario</p>
<p>Figure</p>	<p>The diagram illustrates a network architecture where three clouds (Cloud 1, Cloud 2, and Cloud 3) are connected to three ISPs (ISP1, ISP2, and ISP3). Each cloud is connected to its respective ISP. The ISPs are interconnected and connected to two Points of Presence (PoP1 and PoP2), which are part of an Upper Tier ISP. Green dashed lines represent inter-cloud communication, and green solid lines represent inter-cloud bulk data replication between the clouds and the Upper Tier ISP.</p>
<p>Value Network Configuration</p>	<p>The Value Network Configuration diagram is divided into two layers: the Cloud Layer and the Network (Interconnection) Layer. The Cloud Layer contains two boxes: 'Cloud Provider [Dest]' (Similar to Source Accessible via Transit (and Home NSP')) and 'Cloud Provider [Source]' (Datacenters, Storage and VM, Business agreements). The Network Layer contains two boxes: 'Transit Network' (Transit to lower tier with 95th percentile, Peer with same Tier, Global connectivity) and 'Home ISP' (Residential users, Business users, Transit from upstream). Arrows indicate 'reachable' between the Cloud Provider [Dest] and Transit Network, 'Connectivity' between the Cloud Provider [Source] and Home ISP, and 'SLA' between the Home ISP and Cloud Provider [Source]. A 'Transit' arrow points from the Transit Network to the Home ISP.</p>

Main success scenario	<p>The main success scenario includes the following steps:</p> <p>Step 0: SLA is negotiated between DC Operator and ISP for the connectivity and bandwidth required and the delay tolerance (RPO value in SLA) of the bulk data transfers. The ISP may offer discount to Cloud for traffic which can be shaped rather than be instantaneously transferred.</p> <p>Step 1: DC operator decide a certain granularity of the back-up service to be offered and generates the SLA for its customers (IaaS and SaaS Providers)</p> <p>Step 2: A certain service (e.g. storage for content) is purchased by an end-user whom is also offered an additional back-up service to guarantee the availability and integrity of its personal data</p> <p>Step 3: Initiation of bulk data transfer on a periodical basis</p> <p>Step 4: ISP shapes the traffic, so as to reduce the transit cost it is facing.</p> <p>Step 5: Destination receives the data sent with some delay.</p> <p>Step 6: DC operators verify the coherency and the integrity of exchanged data</p> <p>Step 7: Assessment of incurred delay w.r.t to the RPO value of SLA versus ISP transit cost savings. Assessment should drive a feedback which potentially should involve the SLAs in all-layers.</p>
Traffic Management Solutions	Mechanism for Inter-Cloud Communication (ICC), as well as DTM (Dynamic Traffic Management), employed either individually or together as DTM++.
Innovation	<p>The main innovation is the shaping of the traffic taking into account the 95th percentile rule and the fact that all the value chain stakeholders' interests are taken into account (including that of the ISP). The network is efficiently used by moving delay insensitive transfers to intervals where traffic is low and well below the expected charge for ICC and choosing optimally the outgoing transit link for DTM; DTM++ combines both benefits for multi-homed ISPs. This is a <i>win-win situation</i> since the other ISP customers will also face less congestion at peak hours, since this traffic is removed from the network and moved to slots (ICC) and/or links (DTM) where the network is under-utilized. An additional innovation can be the seamless support of cloud federations under ICC and DTM++. In particular, the data transfers envisioned in this use case can be the result of business agreements among federated clouds/data centers. This possibility is not explicitly mentioned in this use case template since it does not affect the TM mechanism applied.</p> <p>The importance of this use case is also highlighted by the similar vision of related projects, such as Horizon2020 SSICLOPS [57], FP7 FELIX [58] and FED4FIRE [59] that also advocates remote data transfers and remote Infrastructure /Platform/Software access, also possibly under a federation scenario.</p>

3.3.2 Host resource allocation in cloud federations

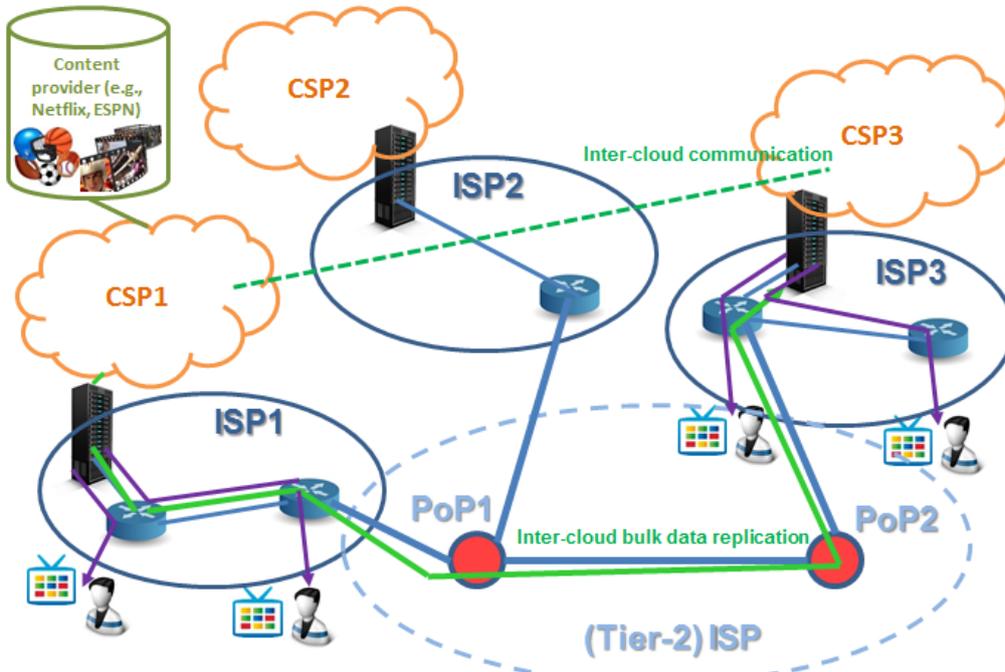
Use case name	Host resource allocation in cloud federations
Goal	<p>The goal is allowing DCOs to offer their services jointly, thus enabling CSPs to freely choose the set of resources they want from each DC in a cloud federation. In particular, CSPs will probably not use the exact amount of resources they announced, that is, CSPs will pick one of the flavours offered by the DCO for a VM, but the VM's resource usage will likely differ from this flavour during runtime.</p>

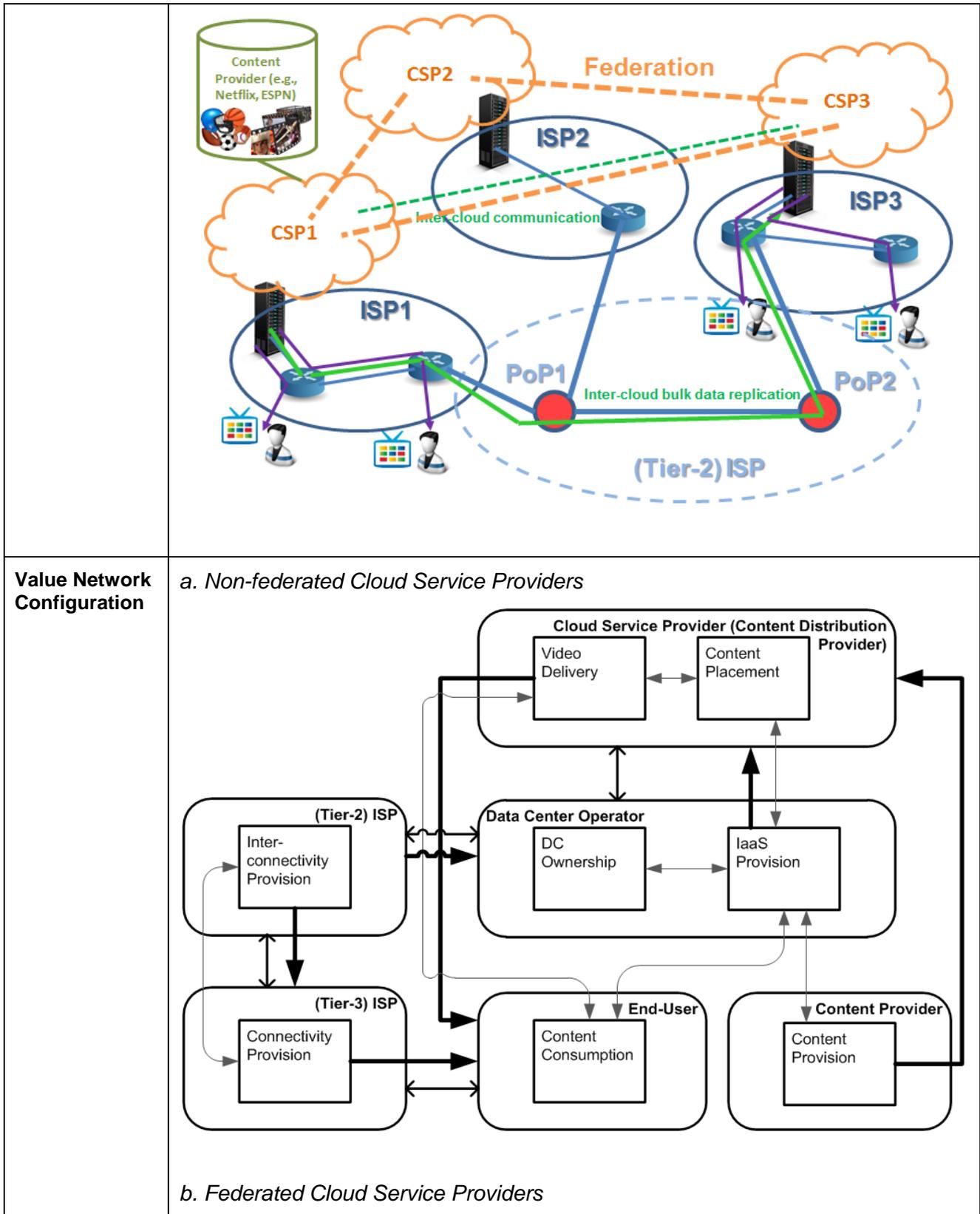
	<p>Therefore, unused and overused best-effort resources (i.e., resources that the CSPs have not paid for guarantees to receive them) are offset for each CSP to redistribute scarce resources in a way that is not only fair within a DC (and in particular guarantees each CSP the resources he demanded initially) but also when the entire federation is considered. This will make the services offered by DCOs more attractive since it gives CSPs flexibility with respect to where they consume resources. Since the offsetting may identify CSPs that overuse one DC, it may indicate to migrate some of their VMs to other DCs; this transfer will rely on SmartenIT's traffic management mechanisms. The resource consumption information is gathered constantly in the cloud or cloud federation for every VM and aggregated to and offset for each CSP. If need be (if scarcity occurs on hosts), individual hosts can access the aggregated information which is globally available, to make local decisions on how reallocate resources between VMs.</p>
<p>Scenario</p>	<p>Operator Focused Scenario</p>
<p>Figure</p>	<p>The diagram illustrates resource allocation across two Data Centers (DCOs). DCO 1 contains two physical machines. The first machine hosts two CSPs (CSP1, CSP2) and two VMs (VM1, VM2). The second machine hosts two CSPs (CSP3, CSP4) and two VMs (VM3, VM4). A grey box labeled 'Resource Conflict' is positioned between CSP3 and CSP4 in the second machine. A curved arrow labeled 'Migrate VM 4' points from VM4 in the second machine to a large empty box representing DCO 2. DCO 2 is currently empty.</p>
<p>Value Network Configuration</p>	<p>The Value Network Configuration diagram shows the flow of information and services. On the left, a dashed box labeled 'Cloud Federation' contains two DCOs. Each DCO has 'Greediness Calculation' and 'IaaS Provision' components. Bidirectional arrows connect the DCOs to each other and to a central 'CSP' box. The CSP box contains 'Resource Consumption' and 'Service Provision' components. Bidirectional arrows connect the CSP to the DCOs. Below the CSP, an 'End-User' box contains 'Service Consumption'. A downward arrow points from 'Service Provision' to 'Service Consumption', and an upward arrow points from 'Service Consumption' back to 'Service Provision'.</p>
<p>Main success scenario</p>	<p>Step 1: Resource scarcity occurs in DC1 of the federation Step 2: DCO1 identifies CSPs with immoderate resource use (in the DC or the entire federation) to reallocate resources to moderate CSPs, who are currently</p>

	<p>deploying DC1.</p> <p>Step 3: VMs of immoderate CSPs may be moved to other DCOs in the federation. In case migrating VMs is not possible (for example when bandwidth is the scarce resource on the congested host) VMs of the immoderate CSP may also be just constrained in their resource usage on this host. Therefore, moderate CSPs are not disrupted by immoderate ones and load is balanced more evenly throughout the DCOs in the federation.</p>
Traffic Management Solutions	Multi-Resource Allocation (MRA) Mechanism
Innovation	<p>Step 1: Resource scarcity occurs in DC1 of the federation</p> <p>Step 2: DCO1 identifies CSPs with immoderate resource use (in the DC or the entire federation) to constrain the immoderate CSP (or live migrate his VMs) and therefore be able to reallocate resources to moderate CSPs, who are currently deploying DC1.</p> <p>Step 3: VMs of immoderate CSPs may be moved to other DCOs in the federation. Therefore, moderate CSPs are not disrupted by immoderate ones and load is balanced more evenly throughout the DCOs in the federation. In case migrating VMs is not possible (for example when bandwidth is the scarce resource on the congested host) VMs of the immoderate CSP may also be just constrained in their resource usage on this host.</p>

3.3.3 Video content transfer between storages of independent clouds

Use case name	Video content transfer between storages of independent clouds
Goal	<p>We consider the case of N Cloud Service Providers, whose traffic is handled by their home ISPs from which IP connectivity was purchased; their home ISPs are typically Tier 2 or 3 ISPs so in order to deliver global Internet connectivity they rely on purchasing transit from Tier-2 (or Tier-1) ISP. Therefore, the inter-domain traffic of the Cloud Service Providers is delivered to Tier-3 ISP(s) through transit links; the respective traffic is typically charged under the 95-th percentile rule. The primary goal of this use case is to reflect the important aspects of video content delivery on inter-cloud communication, as it may include both delay-tolerant traffic, e.g. Video-on-Demand, and delay-sensitive traffic, e.g. live TV/live video streaming.</p> <p>Below, we distinguish two sub-cases depending on the formation of a federation between the interconnected Cloud Service Providers:</p> <p>A. <u>Non-federated Cloud Service Providers</u></p> <p>In this case, we assume that video content is replicated by a Cloud Service Provider to another data center/cloud facility belonging to a collaborative Cloud Service Provider in order to <u>improve QoS/QoE and achieve footprint expansion at low operational cost.</u></p> <p>B. <u>Federated Cloud Service Providers</u></p> <p>In this case, we assume that video content is replicated by a Cloud Service Provider to another data center/cloud facility, i.e. another cluster of the federation, located in a different physical location, in order to <u>improve QoS/QoE at low operational cost.</u></p> <p>In particular, we assume that service provision is performed over the federation of Cloud</p>

	<p>Service Providers, where the resources of the individual Cloud Service Providers are combined according to business policy rules so as to create a large virtual pool of resources, at multiple network locations to the mutual benefit of all participants.</p> <p>In this use case, we assume that the cloud federation will allow the creation of large “virtual” Cloud Service Providers that can efficiently provision their services over large geographical regions and across multiple networks. The service model for this use case is push-based, while a traffic management mechanism would be necessary in order to perform:</p> <ol style="list-style-type: none"> a. destination selection, i.e. in which cloud (data center) to replicate the data, and scheduling, i.e. when to perform the replication, b. taking possibly into account: <ol style="list-style-type: none"> i. the cost of interconnection between ISPs, ii. the cost of energy consumption by the Cloud Service Providers, and iii. predominantly, the QoE experienced by the end-users. <p><i>For simplicity reasons and without loss of generality, we assume that for each cloud involved in the content delivery service there is a single known Point of Interconnect (PoI) where the respective traffic either originates from or must be delivered to.</i></p>
<p>Scenario</p>	<p>Operator Focused Scenario</p>
<p>Figure</p>	<p><i>a. Non-federated Cloud Service Providers</i></p>  <p><i>b. Federated Cloud Service Providers</i></p>



<p>Main success scenario</p>	<p>The scenario is triggered by a request by a source Cloud Service Provider that aims to transfer video content for to another CSP within the cloud federation.</p> <p>The cloud federation is orchestrated by an Orchestrator, namely an entity (either third-party or one of the CSPs) responsible to allocate jobs w.r.t. to available resources announced by the various Cloud Service Providers. This entity, the Orchestrator would be responsible to reply to the request of the source CSP by sending the ratings of the various destinations to him taking into account the cloud information by other Cloud Service Providers.</p> <p>The home ISP provides to the source Cloud Service Provider ratings of the set of destinations taking into account network information, transit costs, and his QoS requirements.</p> <p>The source Cloud Service Provider makes decision and routes traffic generated by moving video content based on ratings replied by the Orchestrator and its home ISP.</p>
<p>Traffic Management Solutions</p>	<ul style="list-style-type: none"> • ICC • MRA
<p>Innovation</p>	<p>The main innovation to be demonstrated in this use case includes:</p> <ul style="list-style-type: none"> • The intelligence of the source Cloud Service Provider, i.e., the decision making on the destination Cloud Service Provider based on both network and cloud metrics, and the intelligence of the decision making of the ISP: i) assessment of destinations, local and remote ones, based on local information and information from neighboring ISPs, ii)

	<p>traffic shaping based cloud QoS requirements and transit cost mitigation,</p> <ul style="list-style-type: none"> • The communication protocols: between ISP and Cloud Service Provider, which could be based on the ALTO protocol, the inter-ISP communication protocol, which could be based on the inter-ALTO protocol, and the inter-Cloud Service Provider protocol, a protocol for communication between CSPs in the case of non-federation, or the communication protocol between Cloud Service Provider and Cloud Orchestrator, in the case of federation. • Model on incentives for the formation of a cloud federation.
--	---

3.3.4 IoT data transfer for cloud operators

Use case name	Integration of IoT with the core network
Goal	<p>The goal is to manage the network traffic of data generated by network(s) of sensors (IoT).</p> <p>A single sensor does not produce considerable amount of data but large-scale networks of them (tens of thousands or hundreds of thousands) may need some special handling. This is true for multiple business cases that for instance involve end users/customers who want to have access to the information about available parking places in the city, companies who conduct some business/socio analysis, or CSPs providing services that need some statistical/aggregated data collected in some region. In the following we consider that the sensor related applications/services are running in CSP domain.</p> <p>Cloud Service Providers (CSP) offering such services/applications which require such input from sensors may be interested in certain characteristics of transfer of such data (e.g. security, low delay, etc.).</p> <p>In this use case one can consider two types of data to be transferred: 1) delay tolerant and 2) delay sensitive (real time data).</p> <p>In case of the first one the CSP accepts the use of traffic management mechanisms in the ISP infrastructure which may postpone data delivery. This way the ISP may efficiently allocate available network resources to meet demanding requirements of different applications/customers. An example of delay tolerant data is meteorological information.</p> <p>On the other side the real time data collected by sensors should be immediately transferred to the CSP. The ISP should prioritize such traffic using suitable traffic management approaches to achieve expected QoS. An example of such real time application is video streaming from a network of cameras or health monitoring (recording patient health parameters).</p> <p>To prioritize the traffic of real time data the edge/access ISP aggregates the respective flows of sensor data at a certain aggregation point inside his domain and forward them to cloud based services /applications which are installed in a remote network domain.</p> <p>To guarantee the selected parameters of aggregated flows through multiple domains the cooperation between domains may be needed. ISPs should cooperate with each other using available or new mechanisms to maintain agreed level of multi-domain network service. Moreover, the CSP should be involved in this dialogue to automatically provide the information about expected network service parameters.</p>

Scenario	Operator Focused Scenario
Value Network Configuration	<p>The diagram illustrates the Value Network Configuration. It features five main components: CSP (Cloud Service Provider), Owner/Operator of sensors, DCO (Data Center Operator), ISP (transit), and ISP (Edge). CSP and Owner/Operator of sensors are connected by a bidirectional arrow labeled 'Access to sensordata'. DCO and CSP are connected by a bidirectional arrow labeled 'Resource allocation'. DCO and ISP (transit) are connected by a bidirectional arrow labeled 'Transit'. ISP (transit) and ISP (Edge) are connected by a bidirectional arrow labeled 'Transit'. Owner/Operator of sensors and ISP (Edge) are connected by a bidirectional arrow labeled 'Marking sensor's packets'. Inside the Owner/Operator of sensors box is a sub-box 'Sensor Data Provision'. Inside the ISP (Edge) box are sub-boxes 'Connectivity Provision' and 'Sensor Data Aggregation'. A thick arrow points from the Owner/Operator of sensors to the DCO, indicating data flow.</p>
Main success scenario	<p>The main success scenario includes the following steps:</p> <ul style="list-style-type: none"> • The CSP collects data from the owner of sensors, which initiate the respective data transfers. • The edge (access) ISP aggregates the respective traffic from the network(s) of sensors (it is agreed among the ISP and the DCO/CSP which sensor networks/data sources are taken into account). • The aggregated traffic is forwarded from the ISP aggregation point to a remote network domain. Each involved ISP has to guarantee agreed traffic parameters. • Data from sensors are received in a remote network domain, stored in DCO's storage resources and processed by a service of CSP.
Traffic Management Solutions	<p>DTM++</p> <ul style="list-style-type: none"> • to split the traffic and thus decrease the cost of inter-domain transfer • with the use of different optimization functions it is possible to have control over characteristics other than the cost • to delay some time-shiftable sensor traffic per transit link
Innovation	<p>The innovation is the smart management of IoT (sensor) data transfer in the multi-domain network environment in terms of cost and delay (use of DTM++). The CSP may require certain characteristics of such transfer inside a domain (this is a simpler sub-case of the scenario as a single ISP is responsible for network traffic management in its infrastructure) and between domains (this sub-case of the scenario requires cooperation between domains on policy and technical levels).</p>

3.3.5 Service and content placement for users

Use-case name	Service and content placement
Goal	In this use case decisions are taken, based on social information, when and where to place content or services. By analyzing social information made

	<p>available by OSNs, a user’s content/service consumption can be predicted. Based on this prediction the respective content or service can be placed close to the user, e.g. the user’s or a friend’s UNaDa. This placement can be done in times of low load and the content/service can be moved from a close resource. This is how this use case reduces load on the Cloud and the ISPs’ core networks and improves the QoE of end-users, thus meeting the goal of this use case.</p>
<p>Scenario</p>	<p>End-user Focused Scenario</p>
<p>Value Network Configuration</p>	
<p>Main success scenario</p>	<p>The main success scenario includes the following steps:</p> <p>Step 1: Infrastructure information from Internet Service Provider and Cloud Operator/Application Provider, and social information from Social Information Provider is collected</p> <p>Step 2: Location and volume of service/content demand is predicted based on social information</p> <p>Step 3: The most appropriate resources are determined based on demand and infrastructure information</p> <p>Step 4: The Cloud Operator/Application Provider is assisted to place the service/content to the most appropriate resources</p> <p>Step 5: The service/content can be consumed with lower resource demands (energy, operating cost) and higher end-user QoE</p>
<p>Traffic Management Solutions</p>	<p>The following traffic management solutions are suitable to be used in this use case:</p> <ul style="list-style-type: none"> • HORST • SEConD • RB-Tracker

	<ul style="list-style-type: none"> vINCENT
Innovation	<p>The most innovative aspect of this use case is the exploitation of social information to predict location and mix and volume of service demand and to derive an effective service/content placement. Effective proactive placement leads to considerable savings of cloud resources, energy consumption, and inter-AS traffic.</p>

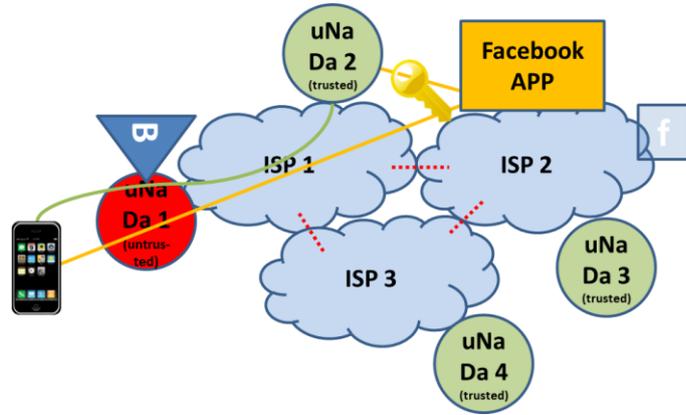
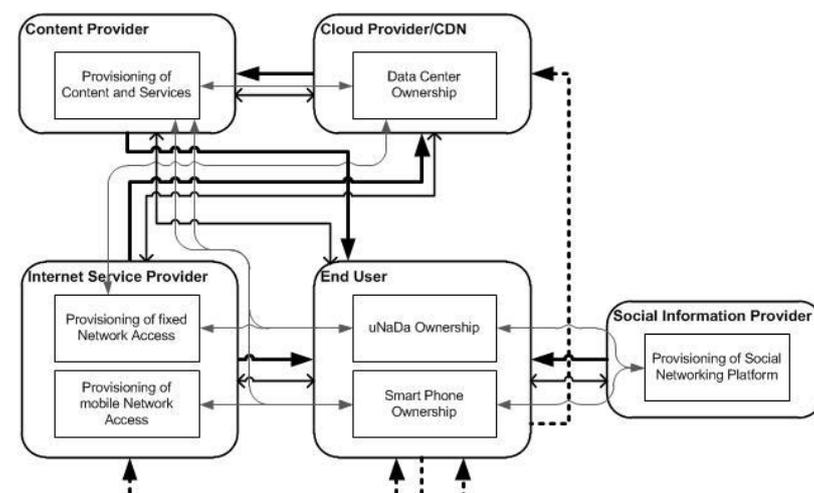
3.3.6 Exploiting content locality

Use case name	Exploiting content locality
Goal	<p>This use case has the goal of exploiting the location of UNaDas (User controlled Nano Data centers) to reduce inter-domain traffic and to improve QoE.</p> <p>Downloading or streaming a resource from a server to a client usually involves inter-domain links, which are expensive for ISPs. This traffic can be reduced if the resource is available in a neighbor AS or it can be completely eliminated if the resource is available within the AS.</p> <p>Therefore, this use case aims to find providers for a resource which is located within the same AS, or a neighboring AS, where the request originates from. Neighboring ASes are typically other access providers since content is stored on home routers of end users. Therefore, AS hops are resolved by using traceroute in the direction of the intended download. At the same time the QoE for the user can be improved since less links are traversed by the traffic, which means less chance of congestion.</p> <p>Furthermore, if multiple sources are found, then chunks can be downloaded from multiple sources and bandwidth can be increased making higher quality streaming possible. The QoE aspect of this use case is especially relevant during peak traffic hours when inter-domain links are typically at their capacity limits. Therefore, serving content from the edge of the network, ideally from the same AS, reduces the load on other links.</p>
Scenario	End-user Focused Scenario
Value Network Configuration	<p>The diagram illustrates the Value Network Configuration. It features three main components: an End User, an Edge ISP, and a Transit ISP. The End User contains a Customer. The Edge ISP contains a Customer and an ISP. The Transit ISP contains an ISP. Interactions are as follows: Traffic flows from the End User's Customer to the Edge ISP's Customer. A Peering Agreement connects the Edge ISP's Customer and ISP. The Edge ISP's Customer and ISP are connected via Transit. The Edge ISP's Customer and ISP are connected to the Transit ISP's ISP via IC Contracts. The End User's Customer is connected to the Edge ISP's ISP via Internet Access. The End User's Customer is connected to the Edge ISP's ISP via Subscription. The Edge ISP's ISP is connected to the Transit ISP's ISP via Access to End Users.</p>
Precondition	For this use case to work, the content requested by the user has to be available on one or multiple UNaDas.
Trigger	End-user requests a content, e.g., clicks play on YouTube video or starts a

	torrent
Main success scenario	<p>The main success scenario includes the following steps:</p> <p>Step 1: End-user requests content.</p> <p>Step 2: UNaDa receives or intercepts the request.</p> <p>Step 3: UNaDa finds that the content is available on at least one UNaDa.</p> <p>Step 4: UNaDa finds the closest UNaDa, <i>i.e.</i> maximum 1 AS Hop away.</p> <p>Step 5: UNaDa serves the content from the neighbor to the End-user.</p>
Extensions	<p>Step 3.1: UNaDa finds no neighbor containing the content (go to Step 5.1).</p> <p>Step 4.1: UNaDa finds that original source is closer (less AS hops) than any UNaDa (go to Step 5.1).</p> <p>Step 5.1: UNaDa serves the content from the original source to the End-user.</p>
Traffic Management Solutions	<p>The following traffic management solutions are suitable to be used in this use case:</p> <ul style="list-style-type: none"> • RB-Tracker • HORST
Innovation	<p>Using UNaDas for caching content and retrieving it from close neighbor UNaDas. Locality is considered for content serving from the edge of the network without the need of a centralized cache inside an ISP. While this is of similar spirit with p2p (ie. contribution of resources that are distributed among the users), using UNaDas also ensures the retrieval of content from a close neighbor via a short path. Compared to a hybrid P2P-CDN this approach does not require new infrastructure, it uses edge network resources (e.g. UNaDas) to keep traffic away from the core network and transit links.</p>

3.3.7 Social-Aware mobile data offloading

Use case name	Social-Aware Mobile data offloading
Goal	<p>The aim of this use case is to provide access to all access points within the range of a user depending on a socially-aware incentive scheme. For that purpose, each access point as well as each mobile device is made socially aware of the friendship connections of his owner. For that purpose, each device is equipped with a Facebook client. Assuming, the access point owner is a direct friend of the offloading mobile device's owner, unlimited access is granted. In case of an unknown user, the offloading mobile device's owner is only granted a limited service (B) coupled to the service he provided himself with his very own access point. The service limitation (B) is calculated by the Facebook App. In order to satisfy privacy constraints, a tunneling scheme (green) is applied and terminated by a trusted endpoint (e.g., the offloading mobile user's own uNaDa).</p>
Scenario	End user Focused Scenario

<p>Figure</p>	 <p><i>Authentication (orange), tunnel setup (green), and bandwidth limitation (blue triangle) to assure fairness of mobile offloading</i></p>
<p>Value Network Configuration</p>	 <p>(dashed lines depict possible potential for business relations enabled by this use case)</p>
<p>Main success scenario</p>	<p>The main success scenario includes the following steps:</p> <p>Step 1: The user's mobile device (A) requests access to a UNaDa nearby, thus providing offloading services to the mobile user.</p> <p>Step 2: The owner of the UNaDa (B) decides in an automated process (based on social data, A's contribution history, and geographical data), which maximum quality of service in terms of data rate the requesting node (A) will be allowed to retrieve from his UNaDa.</p> <p>Step 3: In case of granted permission, the requesting node (A) uses the UNaDa to access content. The amount of service provided by the UNaDa is registered in the system.</p> <p>Step 4: The content accessed by the user is delivered in a cost-effective (monetary, energetic) way while providing a high QoE, as the mobile offloading</p>

	scheme is based on a type of reciprocity
Traffic Management Solutions	The following traffic management solutions are suitable to be used in this use case: <ul style="list-style-type: none"> • RB-HORST • HORST-VINCENT
Innovation	This use case is innovative in creating incentives for providing offloading capabilities to end-users by other end-users. The trust for delivering the service is provided by a combination of social, geographical data and a reciprocal incentive mechanism. Moreover, mobile offloading increases the energy efficiency of network access for energy-limited mobile devices.

3.3.8 Access Technology Selection for Users

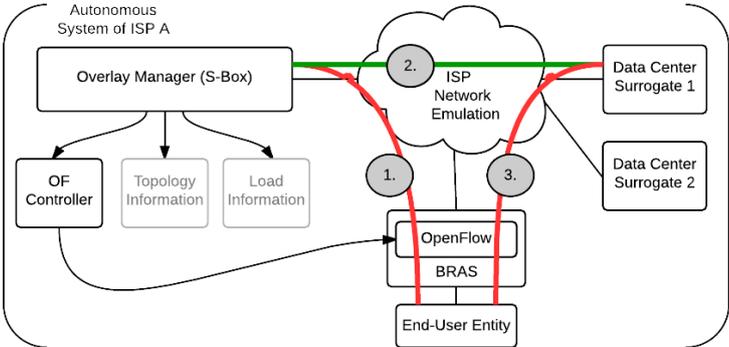
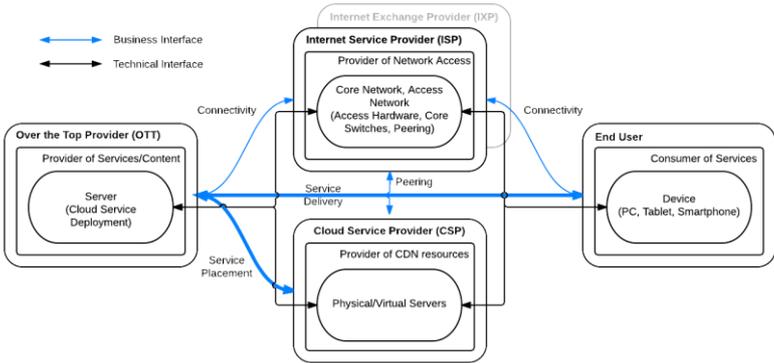
Use case name	Access Technology Selection for Users
Goal	<p>In Mobile Heterogeneous Networks (HetNets) multiple types of access nodes are used in a wireless network. This results in the challenge to decide which user should be connected to which kind of access node (3G, 4G, WiFi, indoor base station, pico base station, macro base station, etc.). The decision can be made on the basis of different criteria.</p> <p>By analyzing social information made available by OSNs, a user's content/service consumption can be predicted. The goal of this use case is to select the most appropriate access node or transmission technology based on this prediction. Furthermore, offloading strategies can be applied in a heterogeneous network. If the load is unevenly distributed among different cell types and access technologies, traffic or even users can be offloaded according to social information, energy efficiency criteria etc..</p> <p>This is how this use case balances the load in the network and improves the QoE of end-users.</p>
Scenario	End-user Focused Scenario

<p>Figure</p>	
<p>Value Network Configuration</p>	
<p>Main success scenario</p>	<p>The main success scenario includes the following steps:</p> <p>Step 1: Network information and social information from Social Information Provider are collected</p> <p>Step 2: Location of user and volume of user demand is predicted based on social information</p> <p>Step 3: The most appropriate access technology is determined based on demand and infrastructure information</p> <p>Step 4: The respective network elements (access nodes) are instructed to connect the user to the most appropriate access technology</p> <p>Step 5: The service/content can be consumed by the users at the new access</p>

	nodes with high or equal end-user QoE compared to the previously used. The new user allocation is also beneficial for the network with respect to operating costs or energy.
Traffic Management Solutions	The following traffic management solutions are suitable to be used in this use case: <ul style="list-style-type: none"> • RB-HORST • vINCENT • MONA
Innovation	The most innovative aspect of this use case is the exploitation of social information to predict user demands and to compute an effective user assignment to access nodes. Effective assignment also leads to considerable savings of network resources and energy consumption.

3.3.9 SDN based DC server selection

Use case name	SDN based DC server selection
Goal	<p>The collaboration of Internet Service Providers (ISPs) and Data Center Operator (DCOs) is beneficial for both parties. Influencing the selection of the delivery location server (surrogate) allows the ISP to manage the rising amount of traffic generated from DCOs to reduce the Operational Expenditures (OPEX) of his infrastructure, e.g., by preventing traffic over peered links/costly upstream traffic. At the same time, including the ISP's hidden network knowledge in the surrogate selection process influences the Quality of Service (QoS) a DCO can deliver positively.</p> <p>As a large amount of DCO traffic is Video on Demand (VoD) traffic, this use case advocates the DCO/ISP collaboration in order to address high-volume, long living flows. This type of flows is hardly manageable with state-of-the-art Dynamic Name Service (DNS) based redirection, as a reassignment of flows during the session is difficult to achieve. Consequently, varying load of surrogates caused by flash crowds and congestion events in the ISP's network are hard to compensate.</p> <p>Thus, a novel approach is proposed promoting ISP and DCO collaboration based on a minimal deployment of Software Defined Networking (SDN) switches in the ISP's network. The approach complements standard DNS based redirection by allowing for a migration of high-volume flows between surrogates in the backend even if the communication has state information, such as Hyper Text Transfer Protocol (HTTP) sessions. The interface between DCO and ISP can be implemented using the ALTO protocol.</p>
Scenario	Operator Focused scenario and End-user Focused Scenario

<p>Figure</p>	 <p>The diagram illustrates the architecture of an Autonomous System of ISP A. It includes an Overlay Manager (S-Box) which manages an OF Controller, Topology Information, and Load Information. The system connects to an End-User Entity through a BRAS (OpenFlow) and an ISP Network Emulation cloud. Traffic is redirected from the BRAS to two Data Center Surrogates (1 and 2) via the ISP Network Emulation cloud.</p> <p>The ISP owned Overlay Manager (S-Box) influences the surrogate selection process by transparently redirecting traffic at the edge (BRAS) using OpenFlow packet rewriting. The DCO provider is allowed to express preferences on redirection using an interface to the Overlay manager.</p>
<p>Value Network Configuration</p>	 <p>The Value Network Configuration diagram shows the relationships between different providers and the end user. It includes an Over the Top Provider (OTT) (Provider of Services/Content), an Internet Service Provider (ISP) (Provider of Network Access), a Cloud Service Provider (CSP) (Provider of CDN resources), and an End User (Consumer of Services). The diagram shows connectivity, service delivery, and peering between these entities.</p>
<p>Main success scenario</p>	<p>Step1: The ISP or the DCO detects a non-optimal placement of a flow, i.e., a flow traversing a peered link or a flow ending in a data center that is non-optimal due to, e.g., sub-optimal energy consumption.</p> <p>Step 2: Each party (DCO and ISP) can contact the other party and ask for a migration of the flow.</p> <p>Step 3: If both partners consent, the DC owner migrates the VM or content delivery process to a different data center; at the same time, the ISP redirects the flow in the network to point to the new destination. An additional state transfer mechanism guarantees that sessions do not break during the transmission.</p> <p>Step 4: ISP and DCO have reached a better position in terms of OPEX, if the scenario was successful.</p>
<p>Traffic</p>	<p>DTM and MUCAPS can serve as a base for this use case. Moreover, the</p>

Management Solutions	system can interface with RB-HORST, to be included in the selection of locations, i.e., RB-HORST UNadas can be considered as a cloud which can be utilized for redirecting sessions.
Innovation	ISP/DC interaction, migration of flows in the network using SDN while not breaking sessions, e.g., running TCP sessions.

3.4 Use cases summary

The use cases scoped in the **Operator Focused Scenario** deal with **business interactions of stakeholders at the wholesale market**. They share common stakeholders, namely **cloud/data center operators, Cloud Service Providers and ISP**, who are present in most of the use cases, with the exception of the use case on resource allocation, which does not involve directly the ISP. These use cases call for the use of **complementary functionalities** that will improve the operations and/or services of cloud/data centers and ISPs and will impact indirectly end users in terms of improved quality of experience (QoE). Those functionalities cover the optimization of data center resource (CPU, storage) through the creation of a large virtual pool of resources among cloud/data centers as well as the optimization of data transfer among cloud/data centers by taking into account the specific features of the data (delay tolerant and delay sensitive video data, IoT sensor data). OFS use cases meet compatible incentives of the stakeholders such as energy and traffic cost reduction for the providers and indirectly improved QoE for the customers of ISP or Cloud Service Provider. Some use cases belonging to the Operator Focused Scenario (OFS) either advocate or at least consider the use of **cloud federations**, which aim at resource sharing between different cloud players, giving opportunity to develop new business models for cloud/data center operators. Moreover, due to incentive compatibility OFS use cases emphasize on **win-win situations between Cloud Operators and ISPs** where ISPs can optimize traffic management solution and monetize them through SLA with cloud players. **Social awareness may** play also an **implicit part** in several use cases (e.g., to optimize data transfer for (aggregated) contents, which are expected to be demanded by the users in the near future, as predicted by the users' OSN information). **Inter-cloud communication** and **dynamic traffic management solutions** developed by SmartenIT are appropriate to instantiate effectively the OFS use cases.

On the other hand, the use cases scoped in the **End-user Focused Scenario** deal with **business interactions of stakeholders at the retail sale market**. They consider a **direct involvement of the end user and its resources in the service delivery chain**. They are capitalizing on functionalities and services offered by **user devices and owned nano data centers (UNaDa)**, such as the placement of static content and service in UNaDa in order to store content/service closer to the end user, in a neighbor UNaDa. Exploiting content locality attains the benefit of inter-domain traffic reduction for the ISP by delivering content/service from local UNaDa. Socially aware mobile offloading capabilities to end-users are provided by other end-users where the trust for delivering the service is provided by a combination of social, geographical data and a reciprocal incentive mechanism. Access technology selection for users takes advantage of UNaDas to provide WiFi

offloading opportunity bringing traffic reduction in the mobile network of the ISP, while providing still access to the network and contents for the end-user. The offloading relies on the exploitation of social information to predict users' demand and to compute an effective user assignment to access nodes. All EFS use cases rely on incentives motivating user contributions for e.g. sharing trusted WiFi access. Moreover they capitalize on **social information such as users' relations and users' interests** to serve users by providing **improved QoE** compared to traditional traffic management solutions and improved **energy efficiency for end users as well as for cloud providers. Collaborative traffic management** solutions taking into account the end user contribution developed by SmartenIT, such as RB-HORST, vINCENT and SEConD, are appropriate to instantiate effectively the EFS use cases.

Finally, the **SDN based Data Center server selection** covers **both Operator Focused Scenario and End user Focused Scenario**. This use case highlights again a **win-win situation of collaboration between Cloud Operators and Internet Service Provider**, especially in case of management of high volumes of long living flows. Such a use case aims at a reduction in terms of OPEX for ISP, by preventing traffic over peered links/costly upstream traffic. At the same time, for the Cloud Operator, the Quality of Service (QoS) it can deliver is positively influenced by including the ISP's hidden network knowledge in the surrogate selection process. The use case is also covering the End-user Focused Scenario in case the DC server is hosted in UNaDas. The use case illustrates the potential benefit of involving end user collaboration through DC sharing in UNaDas. **This use case serves as an example to show benefit of combination of OFS and EFS SmartenIT TM solutions** such as DTM and RB-HORST **which may nicely be complementary** to each other and beneficial for all stakeholders of the chain.

To sum up, the studied use cases pertain to real business models and needs, which can be efficiently met by means of providing proper **incentives and rising to win-win situation among stakeholders**. The use cases and associated SmartenIT mechanisms target to minimize costs in terms of inter-connection charges due to traffic generated by cloud services and applications for ISPs, operating cost in terms of connectivity charges and energy cost for Cloud Service Providers/Data Center Operators, and connectivity cost (WiFi vs. mobile) for end-users. Moreover, these use cases comprise benefits brought by **QoE- and social awareness**, as well as **energy efficiency**. In addition, from the end-user perspective, these use cases offer a large set of services and **focus on user's needs**. The analysis of these diverse use cases reveals that **SmartenIT solutions** are very broadly applicable. They can be adopted in a wide variety of cases, they are in line with recent evolutions in business models and networking technologies, and can lead to significant benefits for all the stakeholders involved particularly when complemented by the SmartenIT TM (Traffic Management) mechanisms.

4 Parameters, Metrics and Evaluation of SmartenIT mechanisms

This section documents the completion of TM evaluation studies initially reported in deliverable D2.4 [7]. Main parameters are fine tuned in order to achieve best performance of the SmartenIT traffic management mechanisms in terms of the metrics of interest. Additional results were integrated to address key design goals of TM mechanisms belonging to OFS and EFS scope that were still open in D2.4. In addition, it provides the final specification and the evaluation of synergetic solutions such as DTM++ and RB-HORST++ together with the identification of the main success scenarios for which they are applicable.

The rest of this section is structured as follows. Parameters, metrics and the evaluation of mechanisms for the operator focused scenario DTM, ICC, MRA, DTM++ and cloud federation are presented in section 4.1. The evaluation of mechanisms for the end-user focused scenario, RB-HORST, vINCENT, SEConD, MONA, MUCAPS and RB-HORST++, as well as the corresponding parameters and metrics are documented in section 4.2. Main outcomes of the performance evaluation of the mechanisms are summarized in section 4.3. Finally, key metrics for SmartenIT use cases are identified in section 4.4 to assess if design goals are met by the traffic management mechanisms.

Concluding work presented in Deliverable D2.4 and complementing work presented in this chapter, we also present in Appendix 11 the following material: in Appendix A (11.1), the complete specification of DTM++ is described as a follow up on the initial specification in D2.4, in Appendix B (11.2), a model employed for measuring QoE for mobile video streaming is described to complement the set of models described in D2.4, in Appendix C (11.3), the full version of the model on cloud federation is presented, in Appendix D (11.4), the evaluation of the energy efficiency of Multipath TCP is presented, which builds the basis for QoE aware offloading and handover decisions in future energy efficient mobile data access, in Appendix E (11.5), complementary material discussing further details on the evaluation of caching in small caches is presented, and finally, in Appendix F (11.6), we discuss economic aspects of ICC, especially focusing on pricing issues.

4.1 Parameters, Metrics and Evaluation Results of SmartenIT mechanisms for OFS

This section documents the important parameters, evaluation metrics and results for the SmartenIT traffic management mechanisms belonging to the Operator-Focused Scenario. Depending on the mechanisms, results are either a completion of evaluations presented into D2.4, or first report of full description and evaluation. Those mechanisms are briefly listed below with their evaluation goals:

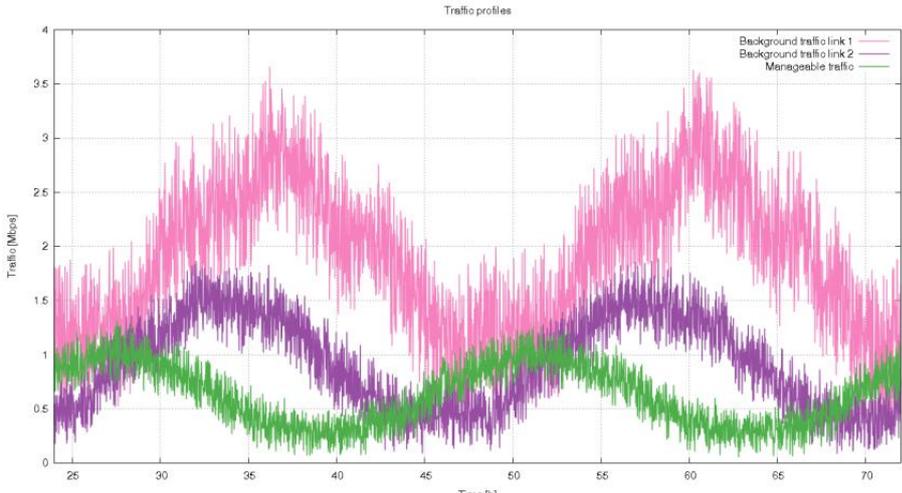
- **Dynamic Traffic Management (DTM)**, which minimizes the inter-domain traffic cost in multi-homed AS by influencing the distribution of the traffic among links. Results present DTM performance with 95th percentile based tariff and volume based tariff.

- **Inter-Cloud Communication (ICC)**, which attains a reduced 95th percentile transit charge by controlling the rate of delay-tolerant traffic, marked a priori accordingly by the ISP’s business customer (e.g. cloud/datacenter), and shifting its transmission at off-peak intervals. Results completion focuses on the assessment of traffic learning features with respect to the prediction of the expected real time traffic.
- **Multi-Resource Allocation (MRA)**, which aims to ensure a fair resource allocation among federated Cloud Service Providers. Results completion is focused on the interdependency of multiple heterogeneous resources (CPU, RAM, etc.).
- **DTM++** employing features of DTM and ICC to perform scheduling of data flows in space (transit links) and time (5-min intervals) in order to further improve the 95-th percentile inter-connection charge (compared to the individual mechanisms) and to perform improved load balancing. Results provide full description of the mechanism and evaluation of the concept.

4.1.1 DTM

This section presents selected simulation results for DTM. Two groups of results of simulation experiments, for volume based tariff and for 95th percentile tariff, are presented below in separate tables. This section is concluded with considerations on scalability, security and reliability of DTM.

Use case name	Bulk data transfer for cloud operators, Inter-Cloud Communication
Scenario	OFS
Goal	To evaluate DTM performance. Volume based tariff is used.
Figure	<p>Figure 4-1: Logical network topology for simulations of DTM operations.</p>

<p>Parameters</p>	<p>Tariff: volume</p> <p>Cost functions for link L1 and L2:</p> $f_1(x) = \begin{cases} 2 * 10^{-8} * x + 200 & 0 \leq x \leq 24 * 10^9 \\ 4 * 10^{-8} * x - 1000 & x > 24 * 10^9 \end{cases}$ $f_2(x) = \begin{cases} 4 * 10^{-8} * x & 0 \leq x \leq 15 * 10^9 \\ 6 * 10^{-8} * x - 300 & x > 15 * 10^9 \end{cases}$ <p>where x is the total traffic volume at the end of billing period.</p> <p>Billing period: 1 day Simulation time: 4 days SDN controller mode: proactive with reference vector (per packet traffic management) Compensation period: 30 s Traffic profiles: see Figure 4-2</p>  <p>Figure 4-2: Traffic profiles.</p> <p>The traffic profiles were selected intentionally in such way that the peak of manageable traffic (green line) is not at the same period as peaks of background traffic on links L1 and L2 (pink and violet lines, respectively). We wanted to check whether DTM will manage to compensate the traffic before the end of billing period. Traffic profiles on Figure 4-2 are presented for two days (two billing periods).</p>
<p>Metrics</p>	<p>Total amount of traffic at the end of the billing period; Total cost (expected, achieved, predicted for non DTM scenario) KPIs: $\xi^{(1)}$, $\xi^{(2)}$, $\Delta D^{(1)}$, $\Delta D^{(2)}$, ρ (refer to D4.2 [7])</p>
<p>Traffic Management Solutions</p>	<p>DTM</p>

<p>Evaluation framework</p>	<p>DTM, implementation of DTM in ns-3 simulator</p>
<p>Evaluation results</p>	<div style="text-align: center;"> <p>60 second samples in time order Link L1</p> <p>60 second samples in time order Link L2</p> </div> <p style="text-align: center;">Figure 4-3: Total and background traffic observed on links 1 and 2</p>

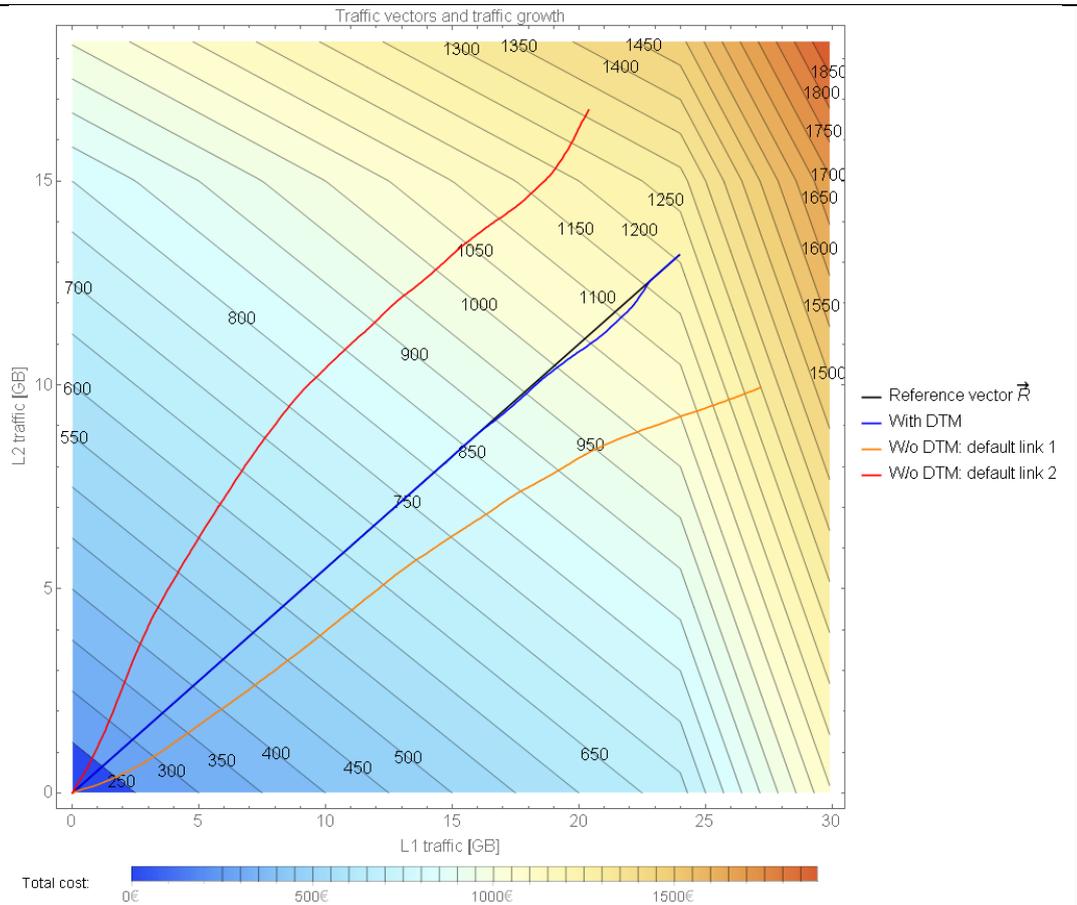


Figure 4-4: Traffic growth on a cost map

The traffic profiles in this experiment were as presented in Figure 4-2. It can be noticed that during almost half of the billing period there were not enough manageable traffic to compensate too much background traffic on link 1. Figure 4-3 shows that in that period all manageable traffic was sent over link L2. Just before the end of billing period the amount of manageable traffic increased. The undesired distribution of the traffic among links where compensated and during the last hour of the billing period that manageable traffic was again sent interchangeably over both links. The period in which there was not enough manageable traffic for compensation is also clearly visible at Figure 4-4. The actual traffic vector was not closely following the reference vector during the whole billing period by deviation from it for some period is observed. Figure 4-4 shows also the expected traffic growth if DTM was switched off and all manageable traffic was sent over a default BGP path, over link 1 or link 2. It can be read from the underlying cost map that the total cost would be higher in both cases.

Traffic costs and KPIs are presented in Table 4-1. It shows cost savings around 8% and very good compensation accuracy. Since actual traffic vector met the reference vector very closely, achieved and expected costs are very close and KPI ρ (defined as a ratio of achieved cost to the expected cost, see D4.2 [7]) is close to 1.

Table 4-1: Traffic costs and KPIs for selected billing period.

Costs		KPIs	
Achieved with DTM	1205.95	(1)	0.9276

		Optimal (w.r.t new \vec{R})	1204.81	⁽²⁾	0.9192
		Expected (w.r.t \vec{R})	1206.11	$\Delta D^{(1)}$	94.18
		w/o DTM: default link 1	1300.12	$\Delta D^{(2)}$	106.00
		w/o DTM: default link 2	1311.95	ρ	0.99986
Innovation	<p>We have presented only a single simulation experiment for volume base tariff for brevity reasons. Several simulations were implemented with multiple traffic profiles and cost functions. They all show that DTM is able to compensate the traffic and distribute it among links as desired. The reference vector is met with a good accuracy so the total traffic cost achieved is minimized, as desired by the ISP. If the case of volume based tariff, the total traffic volume sent over the billing period is calculated and used for billing. In fact there is a lot of time to compensate the traffic. Even traffic bursts or adverse traffic distribution close to the end of billing period could be compensated. Some deviation from the reference vector may happen but is usually small since DTM strives to follow \vec{R} during the whole billing period. The achieved cost is still lower than that expected without DTM.</p>				

Use case name	Bulk data transfer for cloud operators, Inter-Cloud Communication
Scenario	OFS
Goal	To evaluate DTM performance. 95th percentile based tariff is used.
Figure	The logical topology used for simulation experiments with 95th percentile is exactly the same as for volume base tariff presented in Figure 4-1.
Parameters	<p>Tariff: 95th percentile</p> <p>Cost functions for link L1 and L2:</p> $f_1(x) = \begin{cases} 2 * 10^{-6} * x + 300 & 0 \leq x \leq 120 * 10^6 \\ 4 * 10^{-6} * x + 60 & x > 120 * 10^6 \end{cases}$ $f_2(x) = \begin{cases} 4 * 10^{-6} * x + 100 & 0 \leq x \leq 100 * 10^6 \\ 7 * 10^{-6} * x - 200 & x > 100 * 10^6 \end{cases}$ <p>where x is the size (in Bytes) of 5-minute sample used for billing in the billing period.</p> <p>Billing period: 3 days</p> <p>Simulation time: 9 days</p> <p>SDN controller mode: proactive with reference vector (per packet traffic management)</p> <p>Compensation period: 30 s</p> <p>Sampling period: 5 minutes</p> <p>Traffic profiles: see Figure 4-5 and Figure 4-6.</p>

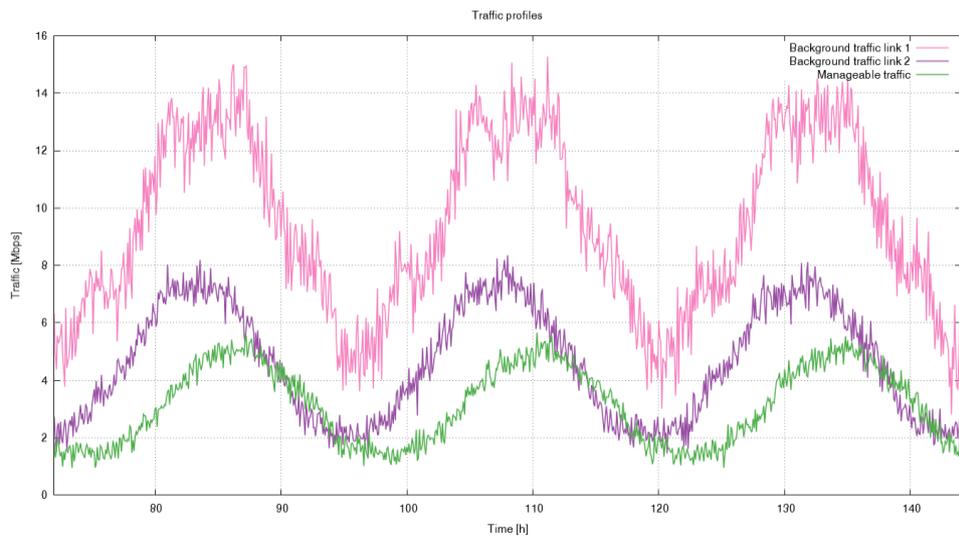


Figure 4-5: Traffic profiles - case 1.

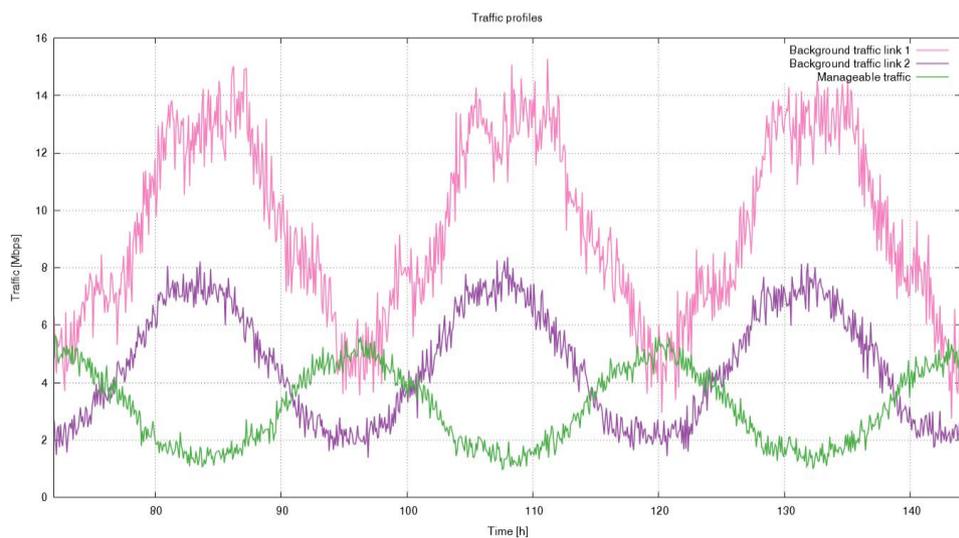
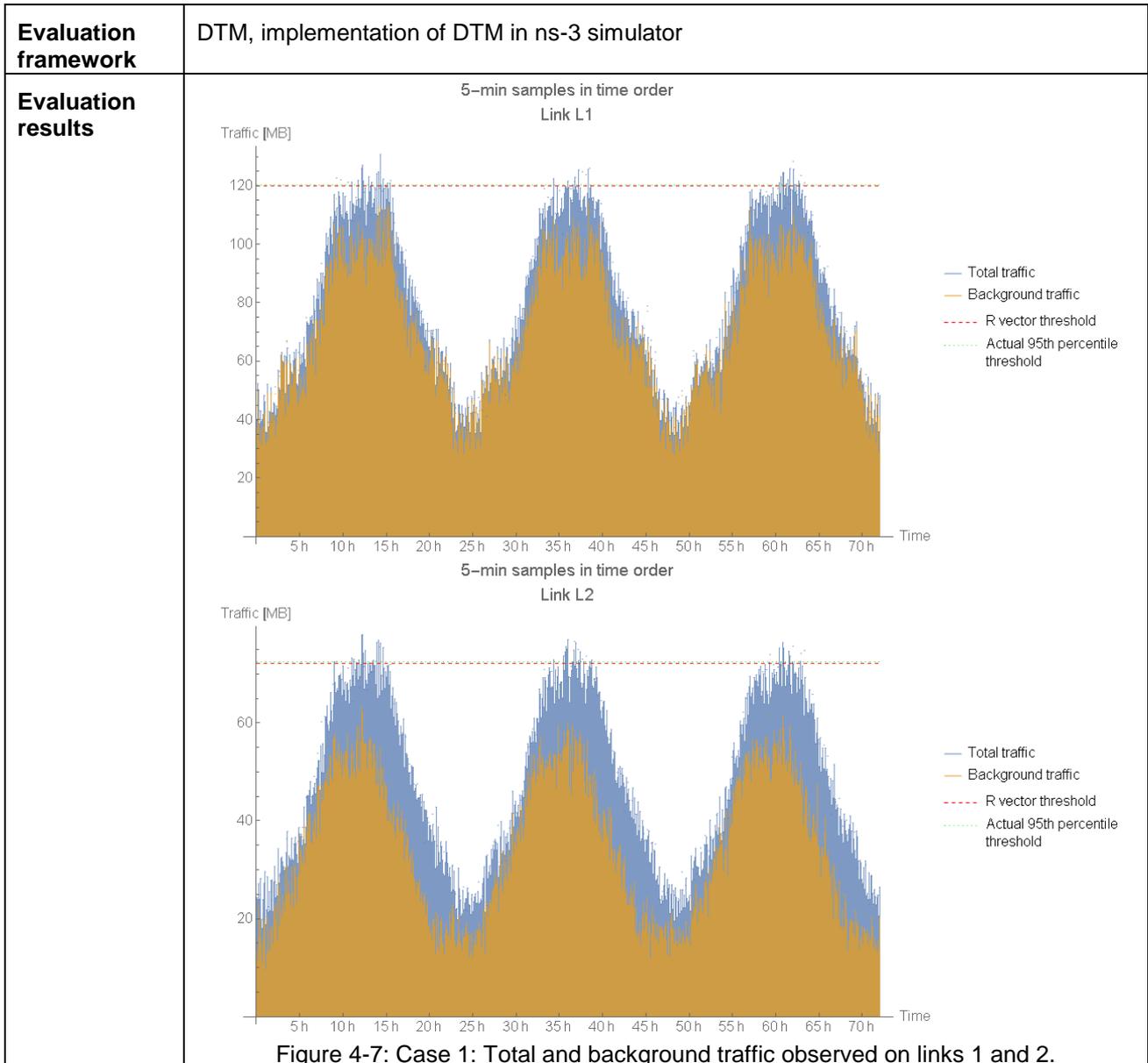


Figure 4-6: Traffic profiles - case 2.

We simulated two cases that differ in manageable traffic profiles. Background traffic profiles for both cases are exactly the same. The traffic envelope for manageable traffic is similar in both case, the only difference is a shift in time. In case 1 (Figure 4-5) the daily peak of manageable traffic is almost at the same time as for background traffic. In case 2 (Figure 4-6) there is a few manageable traffic during peak of the background traffic. It is in fact the worst case for DTM operating on 95th percentile. The intention was to evaluate DTM under such adverse conditions.

<p>Metrics</p>	<p>5-minute samples</p> <p>Total cost (expected, achieved, predicted for non DTM scenario)</p> <p>KPIs: $\xi^{(1)}$, $\xi^{(2)}$, $\Delta D^{(1)}$, $\Delta D^{(2)}$, ρ (refer to D4.2 [7])</p>
<p>Traffic Management Solutions</p>	<p>DTM</p>



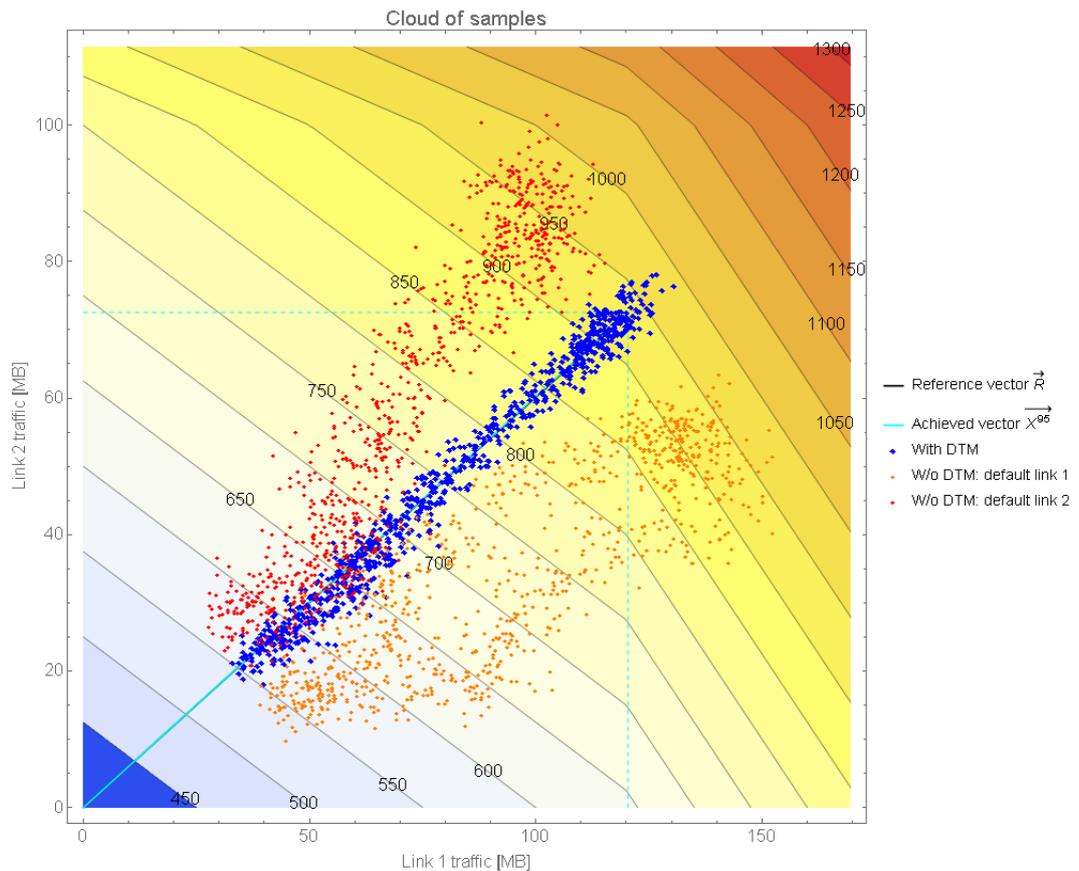


Figure 4-8: Case 1: Distribution of 5-minute sample pairs on a cost map.

Case 1

As shown in Figure 4-7 and Figure 4-8 the reference vector was achieved. The high of 5-min sample for which ISP is charged (95th percentile threshold in Figure 4-7) is approximately the same as stemming from reference vector component value. As presented in Figure 4-8: Case 1: Distribution of 5-minute sample pairs on a cost map the pairs 5-min samples obtained with DTM are condensed around reference vector. For comparison we present also traffic sample pairs that would be obtained inter the same traffic traces but if DTM were not used. Traffic costs and KPIs for this experiment are shown in Table 4-2.

Table 4-2: Traffic costs and KPIs for selected billing period

Costs		KPIs	
Achieved with DTM	932.35	(1)	0.9414
Optimal (w.r.t new \vec{R})	931.276	(2)	0.9544
Expected (w.r.t \vec{R})	928.719	$\Delta D^{(1)}$	58.05
w/o DTM: default link 1	990.399	$\Delta D^{(2)}$	44.58

	<table border="1" data-bbox="491 315 1289 376"> <tr> <td data-bbox="491 315 815 376">w/o DTM: default link 2</td> <td data-bbox="815 315 948 376">976.932</td> <td data-bbox="948 315 1118 376">ρ</td> <td data-bbox="1118 315 1289 376">1.004</td> </tr> </table> <p data-bbox="352 443 443 472">Case 2</p> <p data-bbox="352 488 1430 1283"> In this experiment the average amount of manageable traffic is approximately the same as for case 1, also the traffic envelope is similar, but it is shifted in time in such a way that there is few manageable traffic during daily peak periods of background traffic. The highest 5-min samples are obviously collected during peak periods. Thus peak periods are crucial from DTM traffic management point of view. As shown in Figure 4-9 there is almost no manageable traffic send over link 2 (all is sent over link 1) but a lot of 5-min samples are higher than the threshold stemming from reference vector. Thus the large amount of background traffic was not compensated and the actual level of 5-min sample for which ISP is charged is significantly higher than desired. The resulting cost of traffic on link 2 is higher than expected. At the same time the traffic and cost on link 1 is a bit lower than predicted by reference vector. Total cost of inter-domain traffic is higher than desired (Table 4-3: Traffic costs and KPIs for selected billing period). Reference vector would be met on both links if there were more manageable traffic on link 2 and less background traffic. The observed situation stems from the fact that in the current implementation of economic analyser the amount of manageable traffic that is used for the calculation of a new reference vector (the freedom for searching optimal solution) is estimated as an average amount of manageable traffic share in all 5 min samples collected in the billing period. As a result, in such an adverse profile of manageable traffic the amount of available manageable traffic during peak period is overestimated. The algorithm assumes too high amount of manageable traffic during peak periods and underestimates the amount of background traffic. The optimal cost found by economic analyzer and the calculated reference vector are not achievable. It also clear from Figure 4-10. Comparing the distribution of sample pairs for the situation with DTM enabled and without DTM and link 1 used as a default path one can notice that only distribution of lower samples is affected while the highest sample pairs that are crucial for charging are almost the same. The achieved cost is approximately equal to the cost achievable w/o DTM at default link 1. </p> <p data-bbox="352 1301 1134 1330">Traffic costs and KPIs for this experiment are shown in Table 4-3.</p>	w/o DTM: default link 2	976.932	ρ	1.004	
w/o DTM: default link 2	976.932	ρ	1.004			

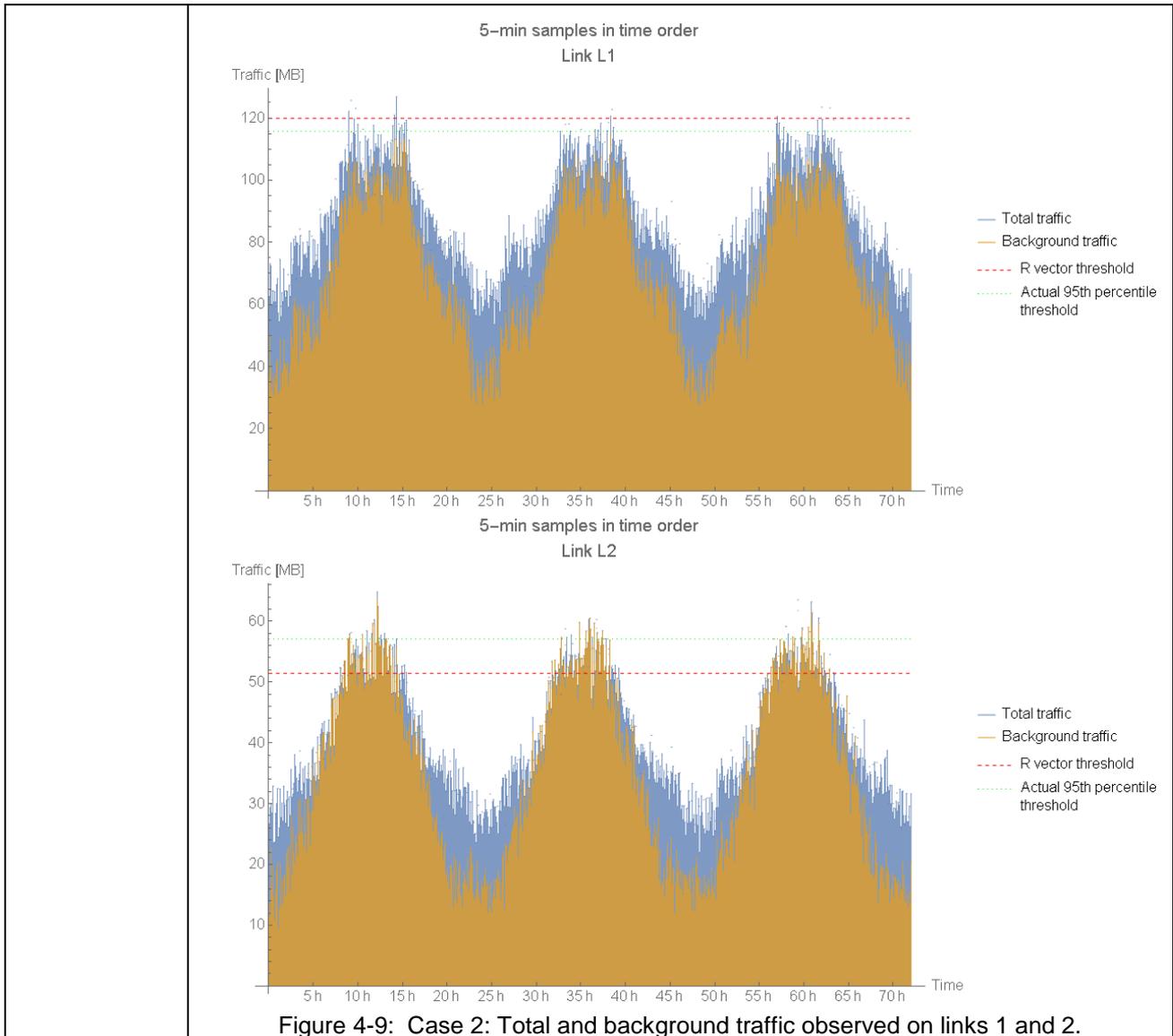


Figure 4-9: Case 2: Total and background traffic observed on links 1 and 2.

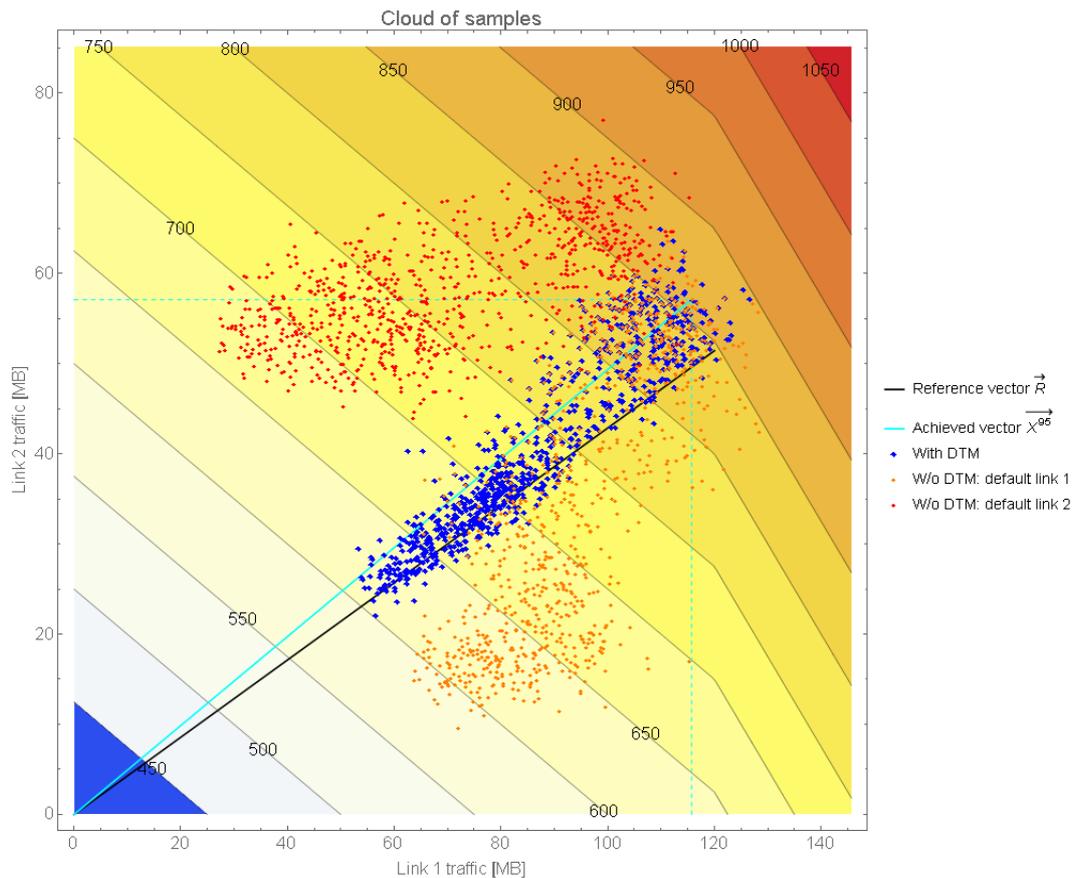


Figure 4-10: Case 2: Distribution of 5-minute sample pairs on a cost map.

Table 4-3: Traffic costs and KPIs for selected billing period.

Costs		KPIs	
Achieved with DTM	859.9	(1)	0.9969
Optimal (w.r.t new \vec{R})	849.2	(2)	0.9698
Expected (w.r.t \vec{R})	845.8	$\Delta D^{(1)}$	2.67
w/o DTM: default link 1	862.5	$\Delta D^{(2)}$	26.73
w/o DTM: default link 2	886.6	ρ	1.016

Innovation

The main conclusions from the above set of simulation experiments are as follows. In general, DTM is able to optimize traffic cost for 95th percentile tariff. However the algorithm is sensitive to the traffic profiles. The traffic must be compensated on a very short time scale of 5 minutes. It is clearly more sensitive than in the case of volume based tariff (in which the traffic is averaged over the whole billing period). The method for calculation of reference vector for 95th percentile tariff needs to be modified. The estimated amount of manageable traffic cannot be average over all 5-min samples. An improved algorithm should be defined; it should take into account amount of both types of traffic during daily peak periods.

4.1.1.1 Scalability Considerations

OpenFlow scaling: From version DTM 3.0 SDN controller can work with 4 operational modes. It can use reactive with reference, reactive without reference, proactive with reference and proactive without reference modes. Proactive modes offer very good scalability regarding a number of flow entries in a flow table. In this case only a single entry in a flow table is used for particular destination, a single flow with wildcard represents all flows going via a tunnel. Also multiple SDN controllers can be used for scaling which can perform load sharing. Different SDN controllers can serve separate DC/clouds in the same domain (not implemented).

Communication scaling: Communication scaling pertains to the inherent to DTM compensation and reference vector announcement to many DC/clouds. Currently S-Box uses unicast communication. S-Box in particular domain announces the same reference and compensation vectors to all cooperating partner domains. In order to achieve better scaling, the multicast communication can be used instead of unicast (not implemented).

Serving many compensation and reference vectors by SDN controller: in a domain where generating traffic DC/clouds are located, S-Box and SDN controller will receive many reference and compensation vectors. Each pair of these vectors is related to selected destination DC/cloud. In order to allow SDN controller process flows related with selected compensation vector we plan schedule delivery of new compensation vectors. Time slot scheduling for compensation vector delivery from different S-Boxes can be used. Compensation vectors are calculated periodically. A remote S-Box can synchronize and schedule delivery of compensation vectors from different S-Boxes (not implemented). From version DTM 3.0 S-Box can send compensation vector update only when the sign of compensation vector is changed. This significantly limits the number of sent compensation vector updates from many S-Boxes and also limits processing effort related to flow table rearrangement performed by SDN controller.

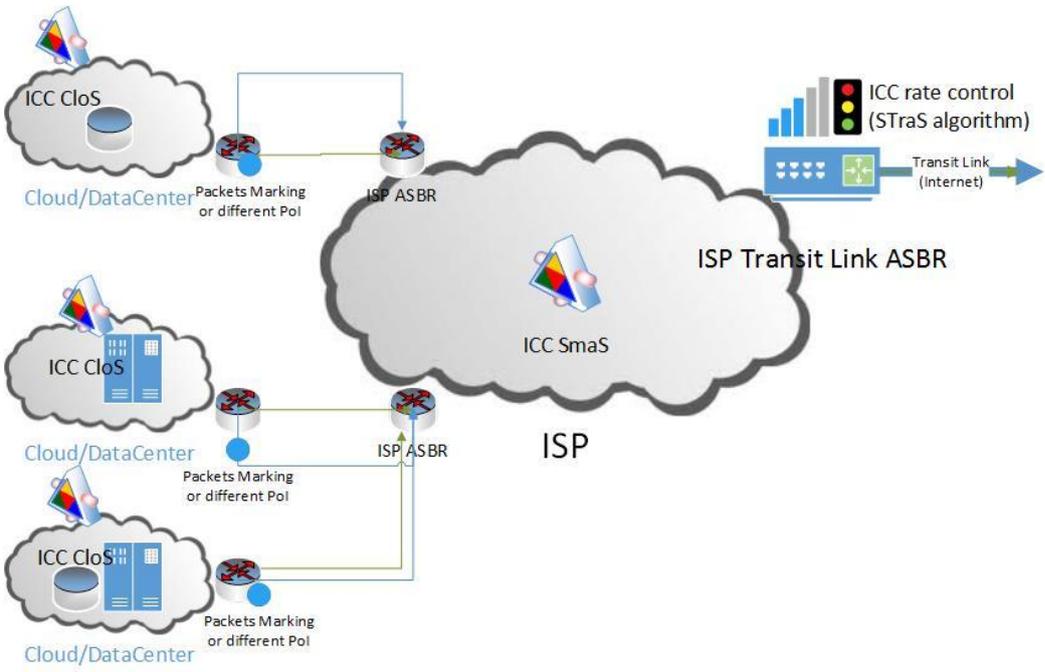
4.1.1.2 Security Considerations

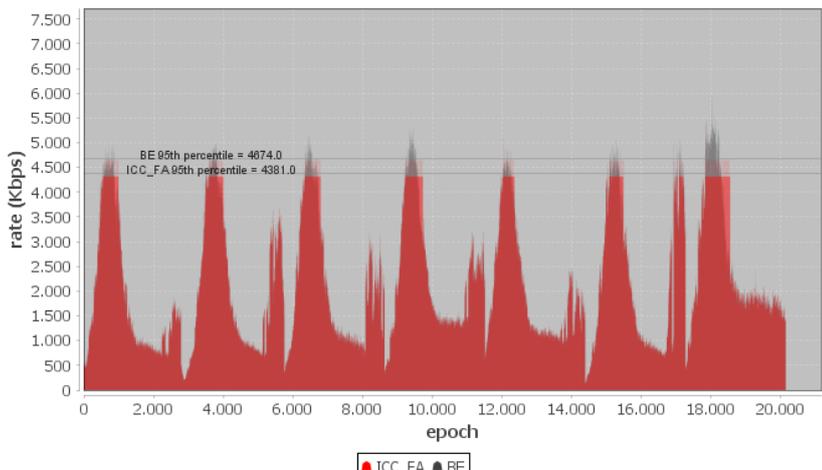
Communication between S-Boxes (currently using http messages) can be done via VPN tunnels set up between communicating parties. Another option is replacement of http communication by https (https not implemented).

4.1.1.3 Reliability Considerations

For reliability purposes redundant S-Boxes and SDN controllers can be used. They can operate in standby mode. S-Boxes can collect the same data and calculate compensation and reference vectors but they do not send any announcement to S-Boxes in partner domains (not implemented). In particular domain can operate two S-Boxes, each one process collects the same data from links and tunnel. They perform the same calculation. One is a master S-Box and the second one is a backup S-Box. They maintain keep-alive communication. Only master sends reference and compensation vector to a remote domains. When the backup S-Box detects master failure it starts sending compensation and reference vectors. In a similar way S-Boxes can work in a remote domain. Both master and backup accept reference and compensation updates but only master communicates with SDN controller.

4.1.2 Inter-Cloud Communication (ICC)

Use case name	Inter-Cloud Communication
Scenario	Operator Focused Scenario
Goal	<p>Applies to any operator data transfer, thus on the whole OFS UC collection of section 3. Performance evaluation of ICC when traffic is unknown, i.e. a variation of ICC (named ICC_STATS) which is additional to those evaluated in D2.4 and for which future traffic is derived from statistics and traffic patterns over a smaller training set. In particular the following questions are answered:</p> <ul style="list-style-type: none"> • How is performance of the ICC learning features when the traffic exhibits time-of-day patterns? • What is the impact on the target 95th percentile due to noisy patterns and inherently erroneous predictions? • How does the performance of this variation of ICC compare to the case where ICC is not used? • How can the logic of ICC be modified/fine-tuned so that the negative impact of erroneous expectations regarding real time traffic on the 95th percentile attained is minimized? • How does the performance of this variation of ICC compare to the previously assessed versions of ICC with perfect knowledge on the traffic patterns?
Overview	 <p>Figure 4-11: ICC TM operates on the ISP transit link and is agnostic to the specific UC. This figure depicts the set-up for inter-cloud/DC communication UC.</p>
Parameters	The key parameters over which sensitivity analysis is performed (but not reported here for brevity reasons) are the following:

	<ul style="list-style-type: none"> • Transit link capacity C • Target 95th percentile C_{target} • Number of epochs y • Threshold parameters $tholds[*]$ • Traffic traces (and their inherent features such as periodicity, variance etc.)
Metrics	<ul style="list-style-type: none"> • Traffic patterns when ICC is applied with the variation that does not assume knowledge of the traffic patterns but uses the 1st day as training set to estimate CAGR and predict traffic in the coming 6 days • 95th percentile attained • Delays for shiftable traffic
Traffic Management Solutions	The ICC mechanism and in particular its network layer logic, i.e. the Shiftable Traffic Scheduling algorithm (STraS).
Evaluation framework	ICC Simulation Framework as documented in D2.4 with extensions to accommodate the latest experiment needs regarding traffic traces and ICC logic for using statistics to predict expected real time traffic in each epoch.
Evaluation results	<p>The results provided comprise the worst case for ICC_STATS since the $tholds[*]$ value for the last epoch of the 5-min interval is set to 1. This means that we are perfectly confident on the predictions and the periodicity of traffic patterns, thus maximizing the impact of erroneous predictions. Better results are always obtained when this threshold value is lowered even by little (e.g. set to 0.985). ICC performance is still good (maximum deviation is 10%), results can vary depending on the periodicity of the traffic patterns and the thresholds values used by ICC for the traffic shaping (the $tholds[]$ values) but ICC always manages to attain better results compared to Best Effort transport without ICC (henceforth mentioned as Best Effort or BE for short) and close to the C_{target} value set. Some indicative plots and result are given below:</p> <p style="text-align: center;">ICC FA vs BE</p>  <p style="text-align: center;">Figure 4-12: The week traffic pattern of no-ICC Best Effort (BE) and how perfect knowledge allows ICC (ICC_FA variation) to perfectly meet the target 95th percentile.</p>

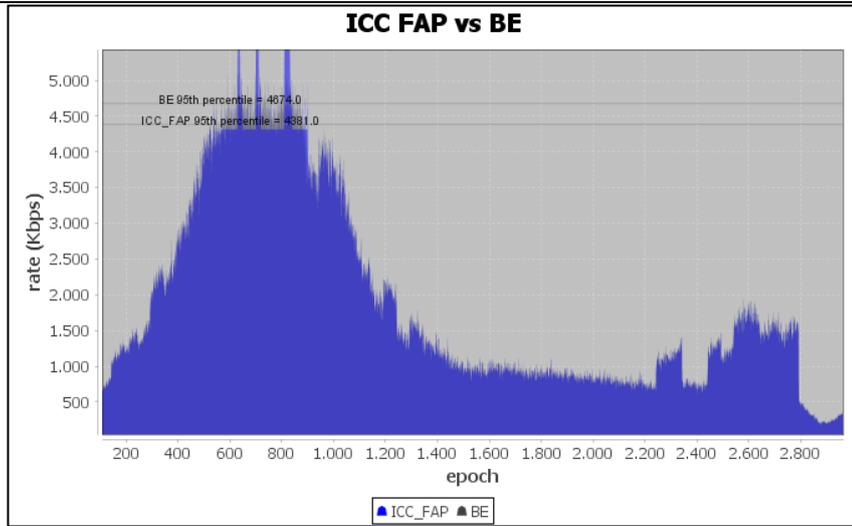


Figure 4-13: ICC_STATS training set: The first day of the week traffic pattern.

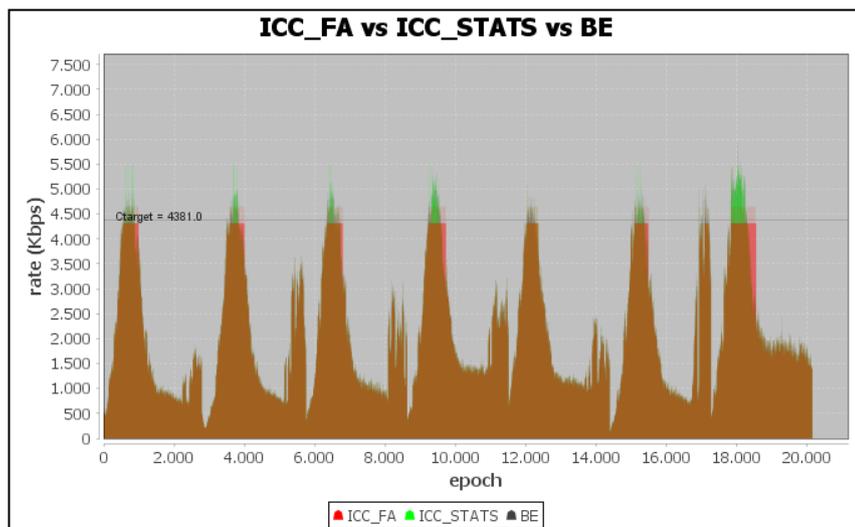


Figure 4-14: Comparison of the traffic patterns attained under ICC_FA, ICC_STATS and Best Effort.

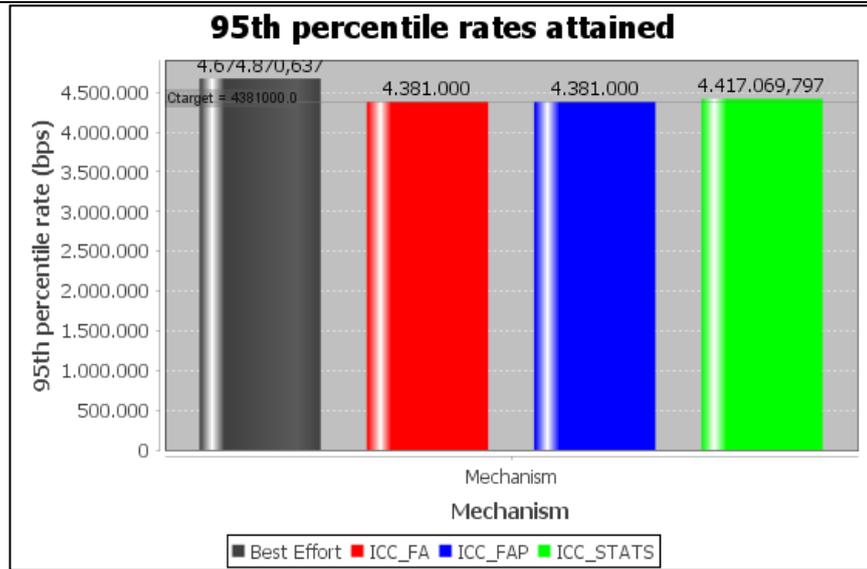


Figure 4-15: The 95th percentile rates attained by the three variations of ICC evaluated and cross-comparison with the Best Effort 95th percentile and the ICC target rate.

Note that even without perfect knowledge ICC performance is close to the target value set and always better than Best Effort without ICC, which by definition has the same 95th percentile for all the ICC simulation runs and is repeated for comparison reasons in terms of the 95th percentile attained. The maximum deviation of ICC_STATS attained 95th percentile from the *Ctarget* value goal observed in the experiments has been approximately 10% for the worst case of ICC_STATS (*tholds[*]* set to 1).

Setting a different value for the ICC *tholds[*]* values results in better performance of the mechanism, since mistakes in expectations regarding the future traffic to be handled has smaller impact and results in lower deviation from *Ctarget*. Below we provide some indicative results from the sensitivity analysis performed over the *tholds[*]* values; for simplicity all threshold values that pertain to the ten 30-sec epochs of the 5-min interval. Note that the deviation margin can drop to below 2%; negative deviation of the cost reduction attained from the *Ctarget* set indicate that a larger discount was attained by ICC_STATS. In particular, below we depict the actual and percentage distance of the 95th percentile attained under ICC_STATS compared to that of Best Effort without ICC for various target values of *Ctarget* and under different parametrization of ICC_STATS, i.e. different value for the parameter *tholds[*]* which defines the confidence of ICC_STATS traffic prediction algorithm and thus the aggressiveness of pursuing the desirable *Ctarget* value: Negative deviations from the *Ctarget* value sought indicates that ICC_STATS actually managed to attain a higher discount of 95th percentile than the one expected initially.

[ICC STATS Performance Sensitivity Analysis over *Ctarget* and *tholds\[*\]* and Comparison with Best Effort](#)

tholds[*] = 0.85

BE 95th percentile	Ctarget	ICC	STATS	ICC_STATS
---------------------------	----------------	------------	--------------	------------------

		95th perc.	deviation %
4674870	3285000	3268679	-0.496834094
4674870	3559000	3437469	-3.414751335
4674870	3833000	3627181	-5.369658231
4674870	4107000	3834504	-6.634915997
4674870	4381000	4036428	-7.865144944
4674870	4654000	4212442	-9.487709497
		AVG dev (%):	-5.544835683
tholds[*] = 0.9			
BE 95th percentile	Ctarget	ICC 95th perc.	STATS ICC_STATS deviation %
4674870	3285000	3381226	2.929254186
4674870	3559000	3582225	0.652570947
4674870	3833000	3797732	-0.920114793
4674870	4107000	4011633	-2.322059898
4674870	4381000	4201133	-4.105615156
4674870	4654000	4356671	-6.388676407
		AVG dev (%):	-1.692440187
tholds[*] = 0.95			
BE 95th percentile	Ctarget	ICC 95th perc.	STATS ICC_STATS deviation %
4674870	3285000	3510412	6.861856925
4674870	3559000	3735947	4.971817926
4674870	3833000	3966478	3.482337595
4674870	4107000	4174374	1.640467495
4674870	4381000	4338118	-0.978817622
4674870	4654000	4464214	-4.077911474
		AVG dev (%):	1.983291808

	Figure 4-16: ICC_STATS sensitivity analysis.
Innovation	The mechanism operates in very small time scales, in the order of seconds, and certainly less than the 5-min interval where most other mechanisms operate (e.g. NetStitcher). This allows for a finer granularity in decision making and control over the rate of the traffic that is to be sent, further empowering the ISP to attain significant savings even in cases where his traffic expectations may be wrong. The mechanism is applicable even in cases of single-homed ISPs, i.e. when there is only one outgoing transit link. ICC has built-in support for both federated and non-federated DCs/clouds

4.1.3 Multi-Resource Allocation (MRA) dependencies

Use case name	Multi-Resource Allocation
Scenario	Operator Focused Scenario
Goal	Determining basic dependencies between different resources, such as CPU, RAM, bandwidth, disk I/O, when consumed by virtual machines (VMs)
Figure	<p>The diagram shows a central blue box labeled 'Virtual Machine'. Above it is a horizontal bar containing logos for 'python', 'php', '7zip', 'NGINX', 'Apache', and 'other'. A magnifying glass labeled 'Performance score' is positioned over the VM. To the left, four RAM chips are shown with a magnifying glass labeled 'RAM utilization' and a double-headed arrow indicating interaction. Below the VM, five CPU chips are shown with a lightning bolt labeled 'Stress tests' and a magnifying glass labeled 'CPU time utilization'. To the right, a disk icon has a magnifying glass labeled 'Disk utilization', and a network switch icon has a magnifying glass labeled 'Network utilization'. Blue arrows and wrench icons indicate the flow and adjustment of resources between the VM and these components.</p>
Parameters	Since the dependencies and requirements of resources of a VM depend heavily on the workload it executes, the first parameter is the cloud workload. In particular, this

Figure 4-17: MRA operation principle.

	<p>parameter/workload is chosen from:</p> <ul style="list-style-type: none"> • Apache: the most used webserver on the Internet. We not only ran tests on how many requests can be served continuously with different VM configuration but also how long it takes to build the apache server. • nginx: is a webserver software that is currently used by 45,5 % of the top 10k websites. • Python scripts: python is one of the most widespread scripting languages. • 7zip: is one of the most widely spread file compression software • PHP: is a widely spread scripting language that is interpreted on the server side • aiostress: an a-synchronous I/O benchmark created by SuSE, to model applications that stress the disk. <p>In order to see how different resources influence the consumption of other resources and performance, i.e., to gain insights on resource dependencies, workloads were executed in VMs with varying configurations. The configurations were determined by varying the following resources:</p> <ul style="list-style-type: none"> • VCPUs: the number of virtual CPU cores (VCPUs) the virtual machine has • RAM: the amount of RAM the VM has • CPU quota: the CPU quota is also known as RAM bandwidth and specifies the data rate with that the CPU can exchange data with the main memory <p>Since resources in clouds are over provisioned, resource contention may occur on hosts, which means that at least some VMs do not receive the resources they are configured with. This resource scarcity was simulated by different technical measures depending on the resource. These “stress tests” constituted the last parameter set and included the following:</p> <ul style="list-style-type: none"> • CPU stress: the standard linux CPU stress test were workers spin on sqrt() • RAM ballooning: deducts RAM from the VM during runtime, by a balloon process that runs in the VM and allocates RAM. • CPU quota reduction: the reduction of the rate with which CPU and memory can exchange data
Metrics	<p>The metrics of the experiments were the following:</p> <ul style="list-style-type: none"> • Performance score: each workload was represented by a benchmark that resulted in a score, which is the performance indicator. • RAM: the amount of RAM the VM used over time. This amount was measured every second. • CPU: the CPU time that was used every second and the overall CPU time consumed for executing the entire workload. • Disk I/O: the bytes written to and read from disk at every second • Network: the bytes send and received via any virtual network interface of the VM
Traffic Management Solutions	<p>The results presented in this section confirm claims of SmartenIT that assumptions made about resource dependencies in literature are generally simplifying. Therefore the results directly justify the need for a new fairness metric, which was also presented by SmartenIT in form of the greediness metric. This metric can be implemented as a resource allocation policy in a cloud federations to allocate multiple heterogeneous resources between CSPs.</p>
Evaluation framework	<p>The evaluation framework is a python script that allows automatic configuration of VMs with the parameters discussed above. Then, the script starts the mentioned workloads on these VMs measure the listed metrics.</p> <p>The resource most complicated to measure is RAM, because many monitoring tools only display how much RAM a VM has configured with but not how much it actually utilizes.</p>

	<p>Therefore, the script infers the amount of RAM that is actually utilized by the VM via smem, which displays memory usage of processes in terms of resident set size and shared set size.</p>
<p>Evaluation results</p>	<p>Our results confirm a Leontief dependency between CPU cycles, disk I/O and network, that is, the consumption of these resources often changes in the same ratios. Leontief utility functions were introduced to computer science in [60] to describe that different resources are needed in static ratios, i.e., increasing the amount a consumer (in this case a VM) receives of resource by X%, does not increase its utility/performance, unless it also receive X% more of the other resources that are Leontief dependent. For example, consider Figure 4-18 shows the utilization of CPU time and network downstream of a VM hosting an apache server. For the different “blocks” in this figure from left to right, the same number of requests is issued over a smaller time frame. As can be seen, this results in an equal increase of CPU time and network downstream.</p>  <p>Figure 4-18: CPU time (green) and network downstream (yellow) consumption of a VM hosting an apache server, for an increasing number of requests per second.</p> <p>While this confirms assumptions made in the literature [60][61], we made an important finding, which largely contradicts them. That is, <i>the utilized RAM does almost never increase in the same ratio as CPU time</i>. This is an important observation, because the standard assumption in literature is that consumers in computing systems have Leontief utility functions. However, when a workload is executed in a VM on an unstressed and stressed host, the difference usually looks as in Figure 4-19. Figure 4-19 depicts that if a host only receives 50% CPU time (left) RAM utilization is decelerated but not decreased. In particular, at the end of the workload, the same amount of RAM is utilized even though one VM only received only 50% of CPU time per second compared to the other. This dependency between RAM and CPU time per second basically occurred for every workload, which strongly disproves the common assumption of Leontief utility functions.</p>  <p>Figure 4-19: Execution of the same apache workload in a VM on an unstressed host (left) and stressed host (right). The green line depicts the RAM utilization of the VM and the yellow line the CPU time received.</p> <p>While a reduction of CPU time per second did not decrease the amount of RAM that is</p>

utilized, also the number of CPU cores did not influence the amount of RAM utilized in general. In particular, we only found one workload, where the number of VCPUs influenced the amount of utilized RAM. This occurred for the 7zip program, when compressing a file. As can be seen in Figure 4-19, the amount of RAM utilized by the VM increases every two cores (except for the first three cores, which utilize the same amount of RAM).

However, this case where VCPUs and RAM show some Leontief dependency is rather the exception than the norm. This is particularly interesting, because CPU and RAM are the first two resources that are assumed to have a Leontief dependency in literature.

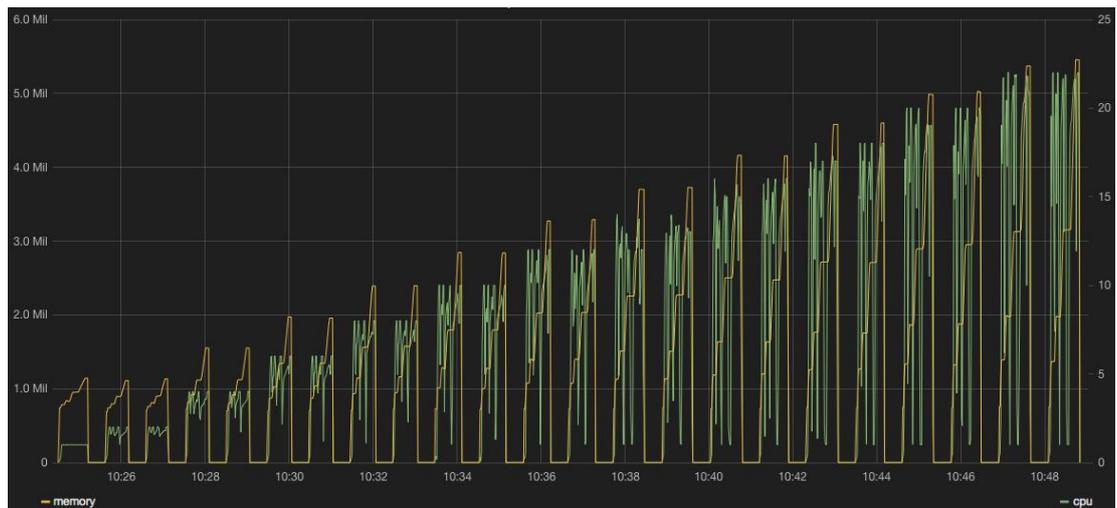


Figure 4-20: Execution of the same 7zip workload in a VM with 1 to 23 VCPUs (from left to right). The green line depicts the RAM utilization of the VM and the yellow line the cpu time received.

Another finding is that the performance of VMs may get worse with additional VCPUs. In particular, for several multi-core workloads we observed that adding additional VCPUs harmed performance. What makes this result even more interesting is that this effect may only occur after a certain number of VCPUs, while adding VCPUs initially increases performance. This effect can for example be seen in Figure 4-10. Here we can see that, when a VM executes the benchmark on an unstressed host, the score peaks when the VM has 3 or 4 VCPUs. However, for 5 and more cores the performance linearly decreases.

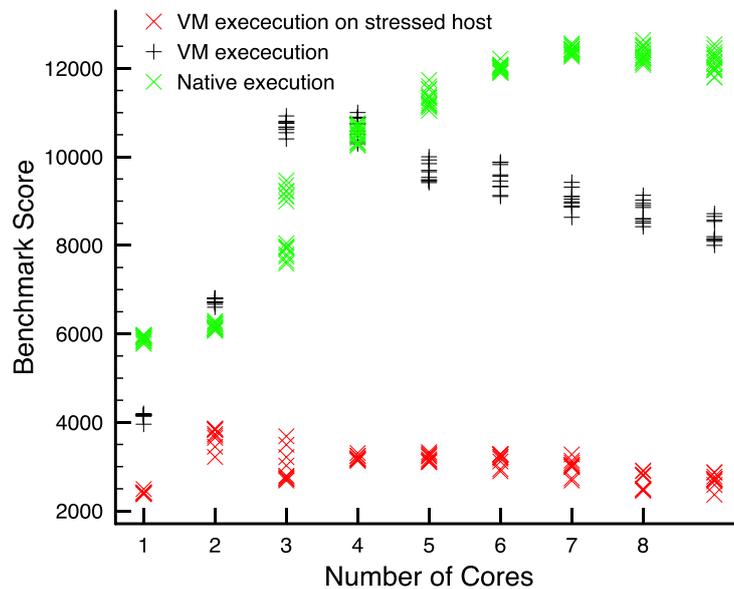


Figure 4-21: the score achieved for an apache benchmark in dependence on the number of VCPUs of the VM.

While this clearly shows that multi-core/parallel programs can suffer, when the VM has more VCPUs, we also found that programs, which can only utilize a certain number of cores, e.g. single-core programs, can profit, when the VM has more VCPUs. This is the case, when the host is stressed. As Figure 4-11 shows, the nginx benchmark cannot utilize more than 2 cores effectively. This can be seen on scores achieved for the unstressed host. However, when the host is stressed the performance keeps increasing for additional cores, although it never reaches the performance of an unstressed system. This phenomenon can be explained as follows: with a rising number of VCPs the VM can compete on more physical cores for CPU cycles, that is, it has higher entitlement to CPU cycles. By rescheduling, the operating system is able to grant these cycles on the same physical core. Therefore, with an increasing number of VCPUs, the VM can get a higher number of cycles on the same core (if it does not utilize the other cores). However, investigation of effects of blocking and high communication overhead is future work.

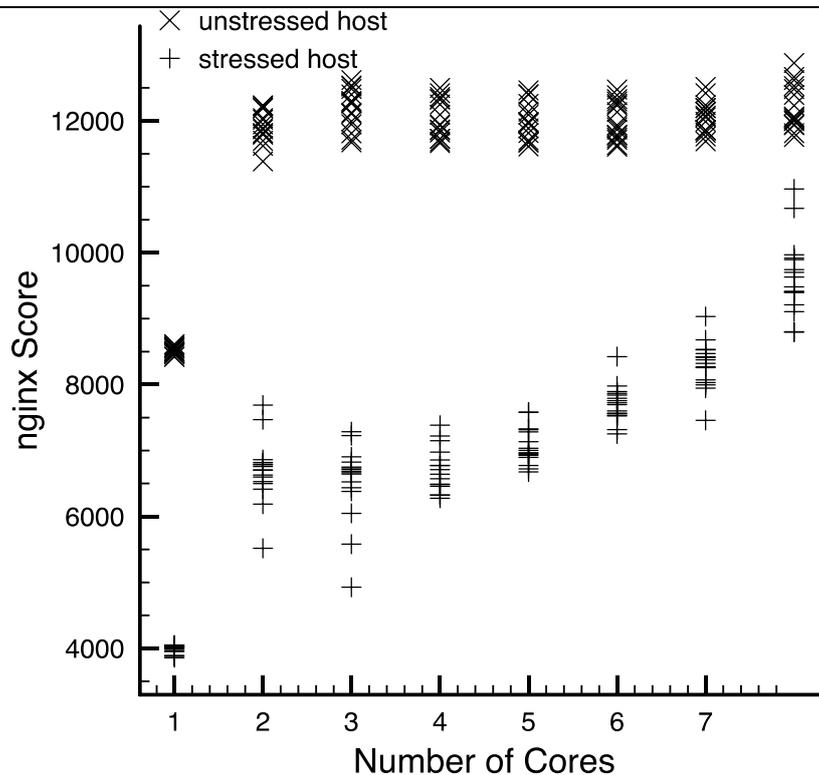


Figure 4-22: Performance score for the nginx benchmark executed by a VM with different numbers of VCPUs and on a stressed and unstressed host.

We have shown that contrary to popular belief the utilization of RAM almost never increases with an increase in CPU, i.e., these resources do not exhibit a Leontief dependency. On the other hand, CPU time, network requests, and disk I/O often show a Leontief relationship. We showed that, due to virtualization, more VCPUs for a VM can actually harm performance, especially, when it executes a multi-core workload. Contrary, when a host is under stress, a workload executed on a VM may perform better, when the VM has more VCPUs than workload can normally utilize.

Innovation

We have presented a fine-grained investigation of the dependency of CPU, RAM, disk I/O, and bandwidth. In particular, the RAM utilization required customized tools, because out-of-the-box VM monitoring tools only display how much RAM the VM has configured, but not how much of this RAM it actually utilized.

We found the commonly assumed Leontief relationship between CPU and RAM does not hold but that contrary CPU, disk I/O and network often show a Leontief dependency. Interestingly, the performance of VMs may suffer from additional VCPUs depending on the workload. On the other hand, when the host is stressed, additional VCPUs may be beneficial, even if the executed workload is single-core.

These findings will be implemented as an extension of the CloudSim simulator, which currently only simulates resource allocation based on CPU. However, the results clearly show that many host resource influence VM performance in unpredictable ways. The implications of these results are that fairness in clouds or cloud resource allocation policies cannot be defined based on resource dependencies, as these are too diverse. Instead, fairness and policies should be defined based on plain VM consumption patterns. Since SmartenIT already expected diversities in resource dependencies, the greediness metric

	defines fairness in clouds solely based on what their VMs consume, that is, the greediness metric does not need to assume any resource dependency to define fairness, contrary to many other definitions.
--	---

4.1.4 DTM++

4.1.4.1 DTM++ – Integration of DTM and ICC: description of the concept and specification

Theoretical introduction

The main rationale of the approach is to build incrementally on top of the already implemented DTM mechanism, thus integrating the minimum set of additional features and components of ICC that provide extra functionality to DTM. The DTM++ concept integrates DTM with ICC, thus allowing for optimizations over both the space (among multiple links – DTM) and time (shifting in time – ICC) axes. The rationale and role of both mechanisms can be summarized as follows:

- DTM is responsible for distributing manageable traffic among tunnels and what follows selects the inter-domain link used. The traffic should be distributed according to a reference vector R . Given the amount of background and manageable traffic in a previous billing period DTM finds such a traffic distribution that the cost of the (inbound) traffic is optimized for the ISP of the receiving domain. DTM does not delay the traffic or drops packets. Offered traffic and served traffic are equal at each point of time.
- ICC additionally protects the traffic on a given link from exceeding the optimal amount represented by a respective component of R vector. ICC influences manageable traffic (delay tolerant component) on each link independently, performing traffic shaping so that the resulting aggregate rate is limited trying to keep it below a threshold, which is set for DTM++ to the reference vector component related to that link.

ICC may operate only using 95th percentile based tariff, so the DTM part of DTM++ uses the same tariff. The DTM changes the distribution of the traffic among links and influences the distribution of 95th percentile samples on each link.

The distribution of samples is further influenced by ICC. It is expected that some high samples exceeding the reference vector value will be lowered. In turn, some lower samples get increased due to the “shifting” of time-shiftable manageable traffic (delay tolerant) from time epochs where the aggregate rate exceeds the reference vector constrain value to those where the aggregate traffic rate is lower. As a result, it is possible to further lower ISP’s cost of inter-domain traffic by using ICC. Even if DTM manages to achieve a reference vector, the further lowering of samples (and resulting costs) is possible with ICC. Therefore, by integrating both mechanisms it is possible to perform more aggressive optimization of reference vector (not limited to only changing the distribution of the traffic amongst links).

The currently implemented version of DTM optimizes the cost of the inbound traffic. Therefore, the traffic cost optimization is done in a domain (henceforth also referred as “local domain”) where the DC receiving the traffic is located. In the local domain, the reference vector and the compensation vector are calculated and sent to the remote domain where the traffic source (i.e. the DC sending the traffic) is located. The compensation vector and reference vector reflects the traffic distribution desired by the local domain. The traffic decision on tunnel selection is done in the local domain but traffic distribution is performed in the remote domain where traffic is generated and injected into the tunnels.

Building incrementally on top of this approach, we have decided to implement (and integrate into DTM++) the ICC mechanism for inbound traffic only. This implies that the ICC operations will be performed in the local domain only and independently per inter-domain link. Instead of buffering – and in order to minimize the implementation/integration overheads we decided to shape delay-tolerant traffic as prescribed in the ICC specification. To this end, the link capacity made available for delay-tolerant traffic will be limited. The limit will be changed dynamically according to the current traffic measurement on each link. The goal is to maintain the throughput below the level stemming from a reference vector component value; the latter is the equivalent of the *Ctarget* value in ICC specification terms.

In DTM the whole traffic is classified as background (non-manageable) and manageable (overlay traffic). The overlay traffic considered is e.g., inter-DC/inter-cloud traffic that has a well-defined origin and destination.

Additionally, for the purposes of ICC the manageable traffic may be classified as delay-sensitive and delay-tolerant. Only the latter can be influenced by ICC. The goal is to keep the information on the type of traffic for each packet. Delay-tolerant and delay-sensitive traffic are marked differently with DSCP and sent via the same tunnel. The receiving side distinguishes the type of traffic using DSCP after the traffic leaves the tunnel. Being a source of traffic the DC or has to mark selected packets with proper DSCP representing delay tolerant or delay sensitive traffic. This distinction can only be made by the DC which is the originator of the traffic.

Limiting delay-tolerant traffic

As mentioned before, the rationale of the DTM++ integration approach is to slow down the rate of the delay-tolerant manageable traffic when necessary. Slowing down the traffic is performed by limiting the link capacity available for delay-tolerant traffic. To do so, we assume that the DC-to-DC delay-tolerant traffic is constituted by TCP flows; this is a straightforward assumption given by the nature of inter-DC communication.

The adopted ICC implementation requires usage of hardware routers with hierarchical policers. We can shape the traffic by limiting the link capacity using hardware hierarchical policers. The current traffic throughput is measured by the router which applies hierarchical policers. Whenever the traffic throughput tends to exceed the limit, then some packets are dropped. This forces TCP sources to slow down their sending rate due to the unacknowledged packets which trigger the TCP rate control algorithm. In particular, the

TCP source (multiplicatively) decreases the size of transmission window. In this way, we are able to limit the amount of delay-tolerant traffic that enters the local domain.

Hierarchical policer operation

In our implementation we use Juniper routers so the terminology is taken from the Juniper configuration.

The hierarchical policer operates with two types of policers: aggregate and premium.

Below we provide the policer description from the Juniper documentation [55]

A hierarchical policer configuration defines two policers—one for EF traffic only and another for non-EF traffic—that function in a hierarchical manner:

Premium policer—You configure the premium policer with traffic limits for high-priority EF traffic only: a guaranteed bandwidth and a corresponding burst-size limit. EF traffic is categorized as nonconforming when its average arrival rate exceeds the guaranteed bandwidth and its average packet size exceeds the premium burst-size limit. For a premium policer, the only configurable action for nonconforming traffic is to discard the packets.

Aggregate policer—You configure the aggregate policer with an aggregate bandwidth (to accommodate both high-priority EF traffic up to the guaranteed bandwidth and normal-priority non-EF traffic) and a burst-size limit for non-EF traffic only. Non-EF traffic is categorized as nonconforming when its average arrival rate exceeds the amount of aggregate bandwidth not currently consumed by EF traffic and its average packet size exceeds the burst-size limit defined in the aggregate policer. For an aggregate policer, the configurable actions for nonconforming traffic are to discard the packets, assign a forwarding class, or assign a packet loss priority (PLP) level.

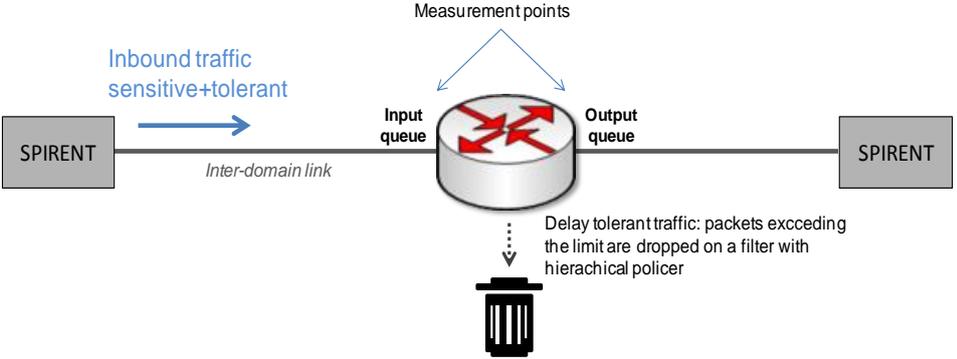
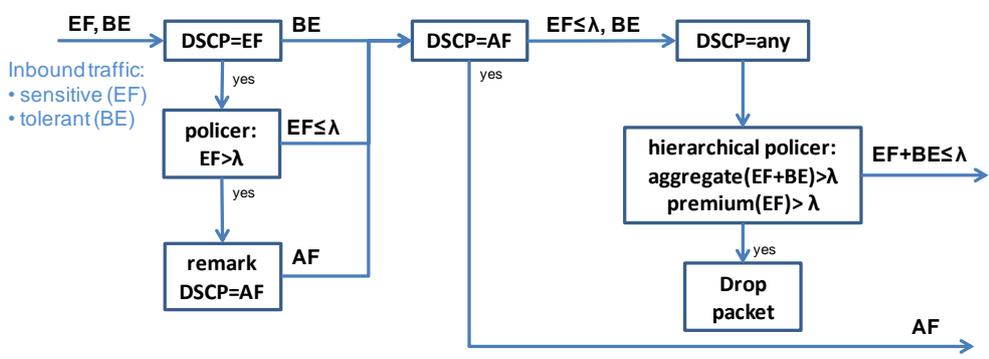
Figure 4-23: Policer description from Juniper documentation [55].

The delay-sensitive traffic is marked with EF and delay-tolerant with BE; other marking would be feasible too. In our approach delay-sensitive traffic must not be dropped at all, so the whole ICC management procedure requires remarking of delay-sensitive traffic. We use the same bandwidth limit and burst size for the aggregate and the premium policer. Before the application of the hierarchical policer the traffic can be remarked. If the delay-sensitive traffic exceeds the predefined limit, it is remarked to AF and it is sent without dropping and the hierarchical policer is not applied. Only when the delay-sensitive traffic throughput is lower or equal to the limit, the hierarchical policer is applied. This way only BE traffic may be dropped. The detailed description is presented in the Appendix in section 11.1. The bandwidth limits follows from reference vector components – reference vector component divided by sampling period (5 min.).

When the billing period expires a new reference vector is calculated which is used by DTM and new values for bandwidth limits for ICC are established. These new bandwidth limits are sent to the BG routers performing the policing procedures.

The detailed description is presented in the Appendix in section 11.1

4.1.4.2 ICC functionality implementation using hierarchical policer (step towards DTM++)

<p>Use case name</p>	<p>Bulk data transfer for cloud operators, Inter-Cloud Communication</p>
<p>Scenario</p>	<p>OFS</p>
<p>Goal</p>	<p>To evaluate the implementation of ICC functionality using hierarchical policer at hardware router. This is an intermediate step towards DTM++. The evaluation of DTM++ functionality will be done within WP4 and reported in D4.3.</p>
<p>Figure</p>	 <p>Figure 4-24: Logical network topology for experiment for testing ICC implementation on hardware router.</p>  <p>Figure 4-25: Configuration of traffic filter on Juniper MX240.</p> <p>Figure 4-24 shows the topology used in the experiment while Figure 4-25 shows the configuration of a traffic filter on the router. For more details please refer to the specification presented in Appendix in section 11.1. In this experiment packets are marked at the source (Spirent). Note that in DTM++, when the above ICC implementation is integrated with DTM, the inter-DC traffic is marked as delay tolerant or sensitive at the source DC and then put into tunnels. On the receiving side there is an additional marker that changes the DSCP codes of all the inbound traffic before forwarding it to the policer.</p>
<p>Parameters</p>	<p>Throughput limit $\lambda=10\text{Mbps}$ Simulation time 3 hours Sampling period 5 minutes</p>

	<p>Sensitive traffic: UDP flow, envelope, SPIRENT generated Tolerant traffic: 3 TCP flows, quasi CBR, aggregated throughput: 3,3 Mbps</p>
<p>Metrics</p>	<p>Attained 5-minute samples Traffic patterns on input and output interfaces of a router with and without ICC</p>
<p>Traffic Management Solutions</p>	<p>ICC: implementation for integration with DTM, hierarchical policer configured on hardware router, Juniper MX240</p>
<p>Evaluation framework</p>	<p>ICC implementation based on hierarchical policer; hardware router used due to lack of simulator ready for such experiment; traffic generated by Spirent</p>
<p>Evaluation results</p>	<div style="display: flex; justify-content: space-around;"> <div data-bbox="389 936 831 1301"> </div> <div data-bbox="887 936 1329 1301"> </div> </div> <p>Figure 4-26: Results for a reference case: ICC not enabled — traffic samples on input and output interfaces of the router.</p> <div style="display: flex; justify-content: space-around;"> <div data-bbox="389 1402 831 1767"> </div> <div data-bbox="887 1402 1329 1767"> </div> </div> <p>Figure 4-27: ICC enabled — traffic samples on input and output interfaces of the router.</p> <p>Experiment 1: Evaluation functionality of ICC implementation: ability to limit the traffic. Comparison of effect of ICC operation to the case without ICC</p> <p>The same traffic patterns were generated for scenarios with and without ICC mechanism. Figure 4-26 shows results for a reference scenario without ICC. Traffic</p>

on input and output interfaces is exactly the same, no packets are dropped. Traffic samples (5-minute samples) exceed the limit. After turning on the ICC mechanism we observe that the aggregated traffic throughput is limited (Figure 4-27). Only the sensitive traffic is affected. The traffic observed on the input interface is a bit higher than that on the output one. At the input interface we measure all incoming traffic while on the output we see only the traffic that passed the filter. Some tolerant packets were dropped. As a result, the tolerant traffic TCP sources were forced to slow down. They send the remaining traffic later (traffic is delayed) when more link capacity is allowed for tolerant traffic. Therefore, we observe higher 5-min samples until around 2:30 hour.

Experiment 2: ICC enabled, sensitive traffic exceeds the limit

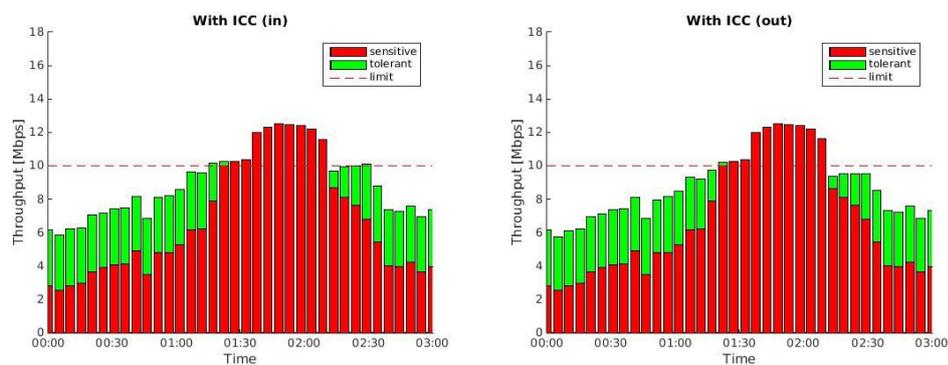


Figure 4-28: ICC enabled — traffic samples on input and output interfaces of the router, sensitive traffic exceeds the limit

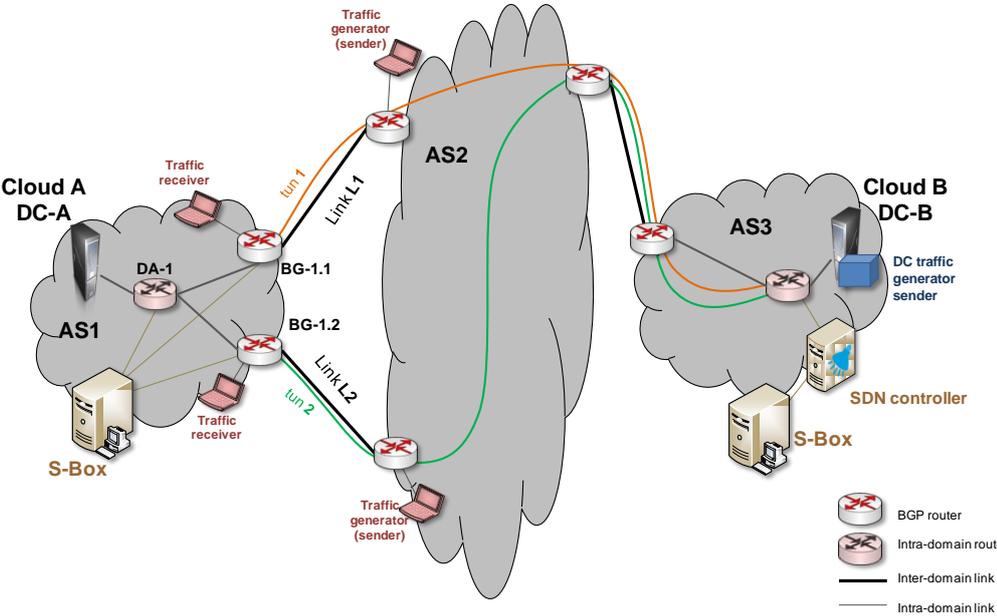
In this experiment we tested the system behaviour for the case where sensitive traffic exceeds the limit. As expected, all sensitive traffic passed the router and packets were dropped. At the same time all tolerant packets were dropped since no link capacity was available for this type of traffic (see Figure 4-28).

The above two experiments show that the ICC implementation on hierarchical policer operates as expected, hence it limits delay tolerant traffic to the predefined limit without affecting sensitive traffic. The intended functionality of ICC is thus realized.

Innovation	ICC implementation using hierarchical policer, limiting delay tolerant traffic to the predefined limit without affecting sensitive traffic. It is innovative, simple and does not require additional software. It is based on policers available in various hardware routers and relies on TCP self-regulating mechanism. The only extension needed is the automated configuration of the limit. So an ISP may control the 5-min samples size and reduce cost of inter-domain traffic on a link with ICC mechanism being enabled.
-------------------	---

4.1.4.3 Results for DTM++

Use case name	Bulk data transfer for cloud operators, Inter-Cloud Communication
Scenario	OFS

<p>Goal</p>	<p>To assess if DTM++ implementation works as expected. This is a test of functionality of DTM++. Results for DTM++ are compared with results obtained for DTM (without ICC enabled) as well as for scenario without traffic management.</p>
<p>Figure</p>	<p>The logical topology used for the DTM++ simulation experiments is presented in Figure 4-29. It is very similar to the topology presented in Figure 4-24. The main difference is that tunnels are terminated at border gateway routers BG-1.1 and BG-1.2 instead of DA-1 router. Additionally, the BG routers implement traffic filters built on hierarchical policers (as presented in Figure 4-23 and described in experiments for ICC implementation and in DTM++ specification in Appendix C).</p> <p>Note that the experiment presented below is not a simulation but it was done using the testbed environment. The preliminary results for DTM++ are presented here as a proof of concept and for completeness of D2.5.</p>  <p style="text-align: center;">Figure 4-29: Logical network topology for experiment for testing DTM++.</p>
<p>Parameters</p>	<p>Tariff: 95th percentile</p> <p>Cost functions for link L1 and L2:</p> $f_1(x) = \begin{cases} 200 & 0 \leq x \leq 83.3 * 10^6 \\ 6 * 10^{-6} * x - 300 & 83.3 * 10^6 < x \leq 416.7 * 10^6 \\ 18 * 10^{-6} * x - 5300 & x > 416.7 * 10^6 \end{cases}$ $f_2(x) = \begin{cases} 2.4 * 10^{-6} * x & 0 \leq x \leq 83.3 * 10^6 \\ 7.8 * 10^{-6} * x - 990 & 83.3 * 10^6 < x \leq 416.7 * 10^6 \\ 24 * 10^{-6} * x - 5310 & x > 416.7 * 10^6 \end{cases}$ <p>where x is the size (in Bytes) of 5-minute sample used for billing in the billing period.</p> <p>Billing period: 500 minutes Compensation period: 10 s</p>

	<p>Sampling period: 5 minutes</p> <p>Throughput limit λ calculated from reference vector</p> <p>SDN controller mode: proactive without reference</p> <p>Manageable traffic, SPIRENT generated:</p> <ul style="list-style-type: none"> • Tolerant traffic: single TCP flow, average throughput 1Mbps • Sensitive traffic: UDP, quasi CBR, aggregated throughput: 2,1 Mbps <p>Background traffic: UDP traffic from proprietary generator, quasi CBR, average throughput:</p> <ul style="list-style-type: none"> • link 1: 9.75 Mbps • link 2: 4.8 Mbps <p>Traffic envelopes: flat</p>
Metrics	<p>5-minute samples</p> <p>Total cost (expected, achieved, predicted for non DTM scenario)</p> <p>Traffic patterns on input interfaces of both BG routers with DTM++, with DTM (ICC switched off) and without traffic management.</p>
Traffic Management Solutions	<p>DTM and DTM++</p>
Evaluation results	<p>Figure 4-30 and Figure 4-31 show the results for DTM++ and DTM (without ICC) respectively. It can be easily noticed that when ICC is enabled the traffic is smoothed. The result of applying a throughput limit for delay tolerant traffic is visible on the traffic traces. Traffic traces for DTM (without ICC) is more bursty.</p> <p>Comparison of costs achieved with DTM++, DTM and without traffic management are shown in Table 4-4. Traffic costs for DTM++, DTM and no traffic management.. It can be noticed that enabling ICC may offer an ISP a further reduction of costs of the inter-domain traffic.</p> <p>In Figure 4-32, we graphically depict the effect of the operation of DTM and ICC on the distribution of sample pairs.</p>

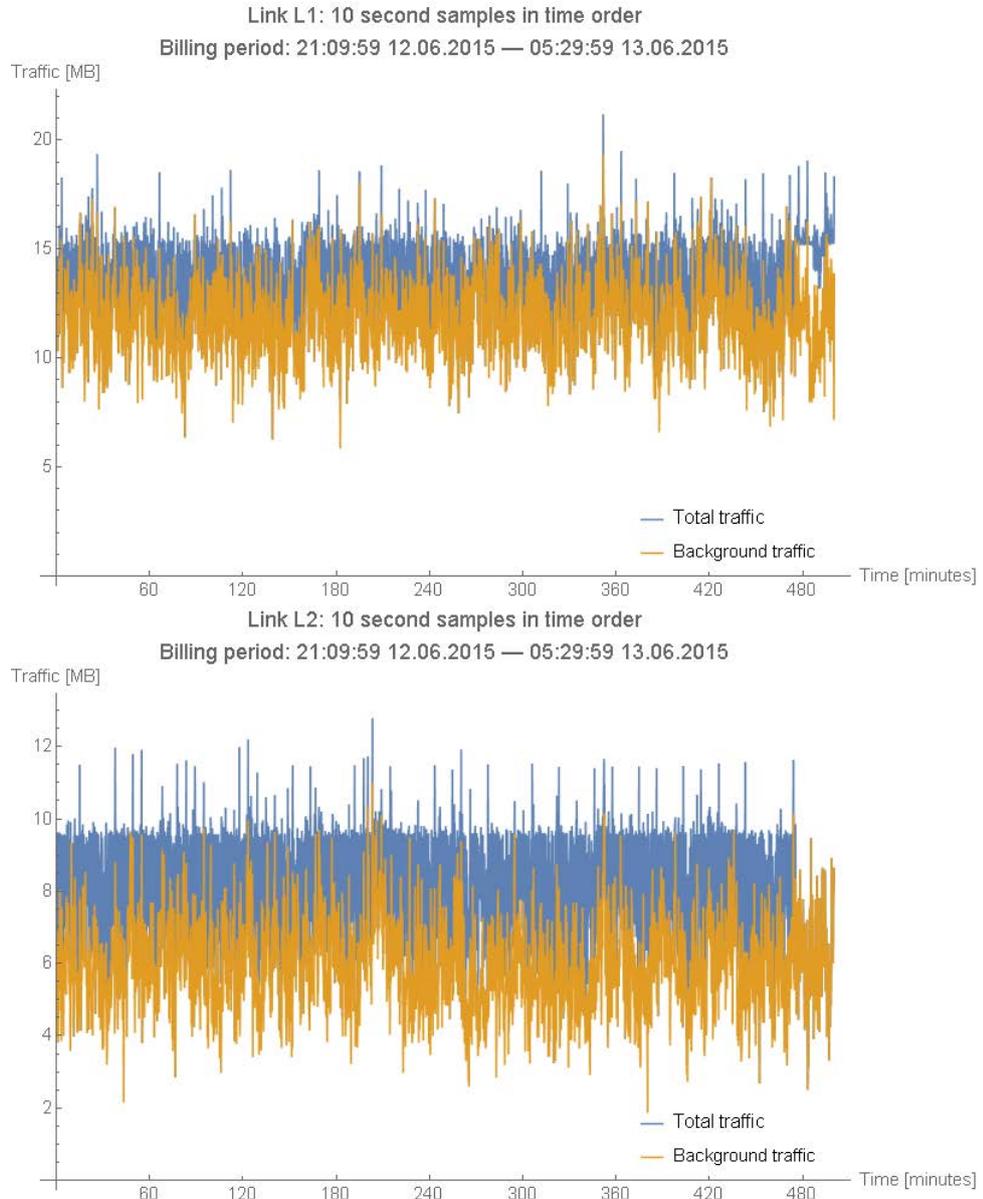


Figure 4-30: Traffic traces on links 1 and 2 when DTM++ is used.

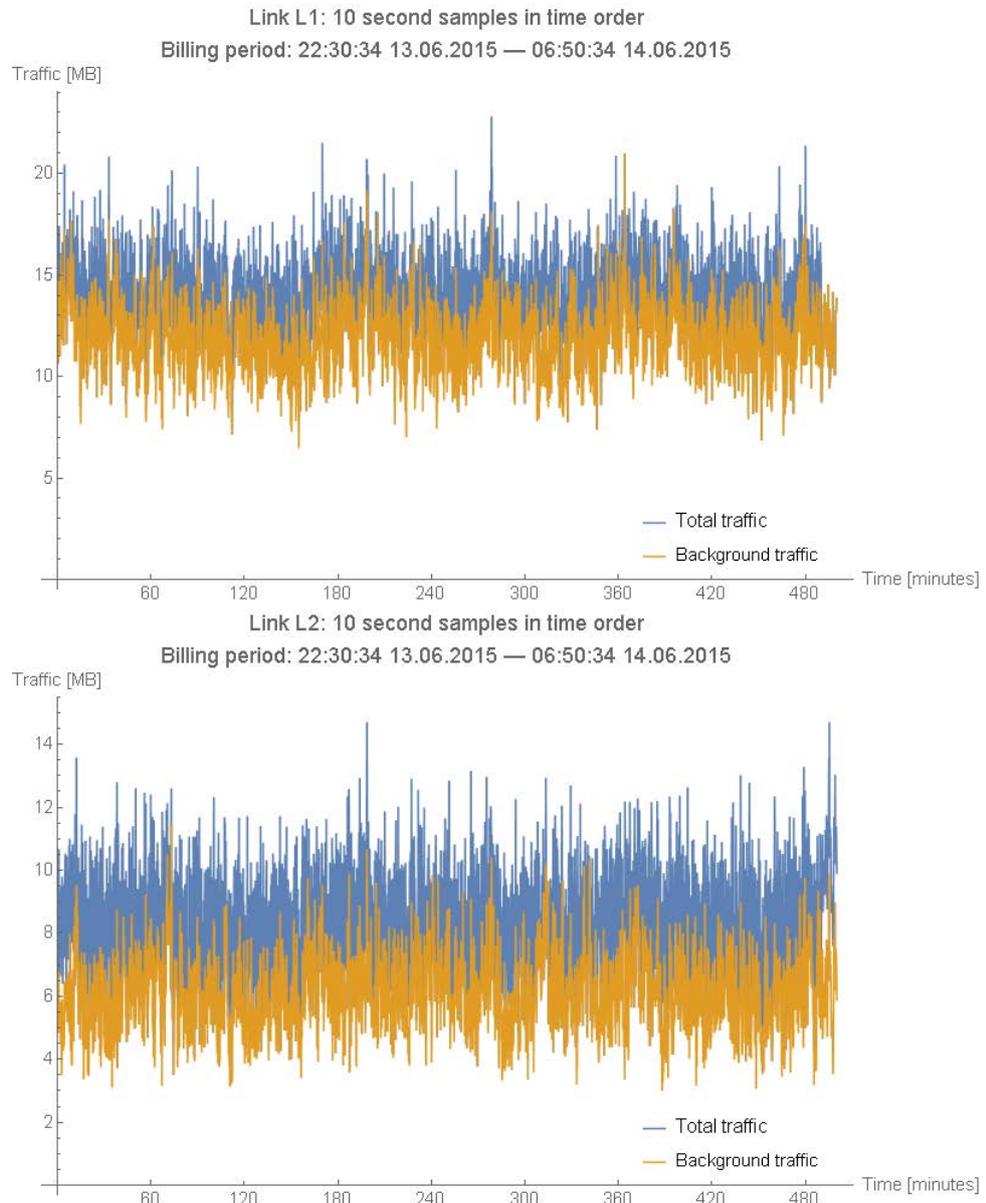
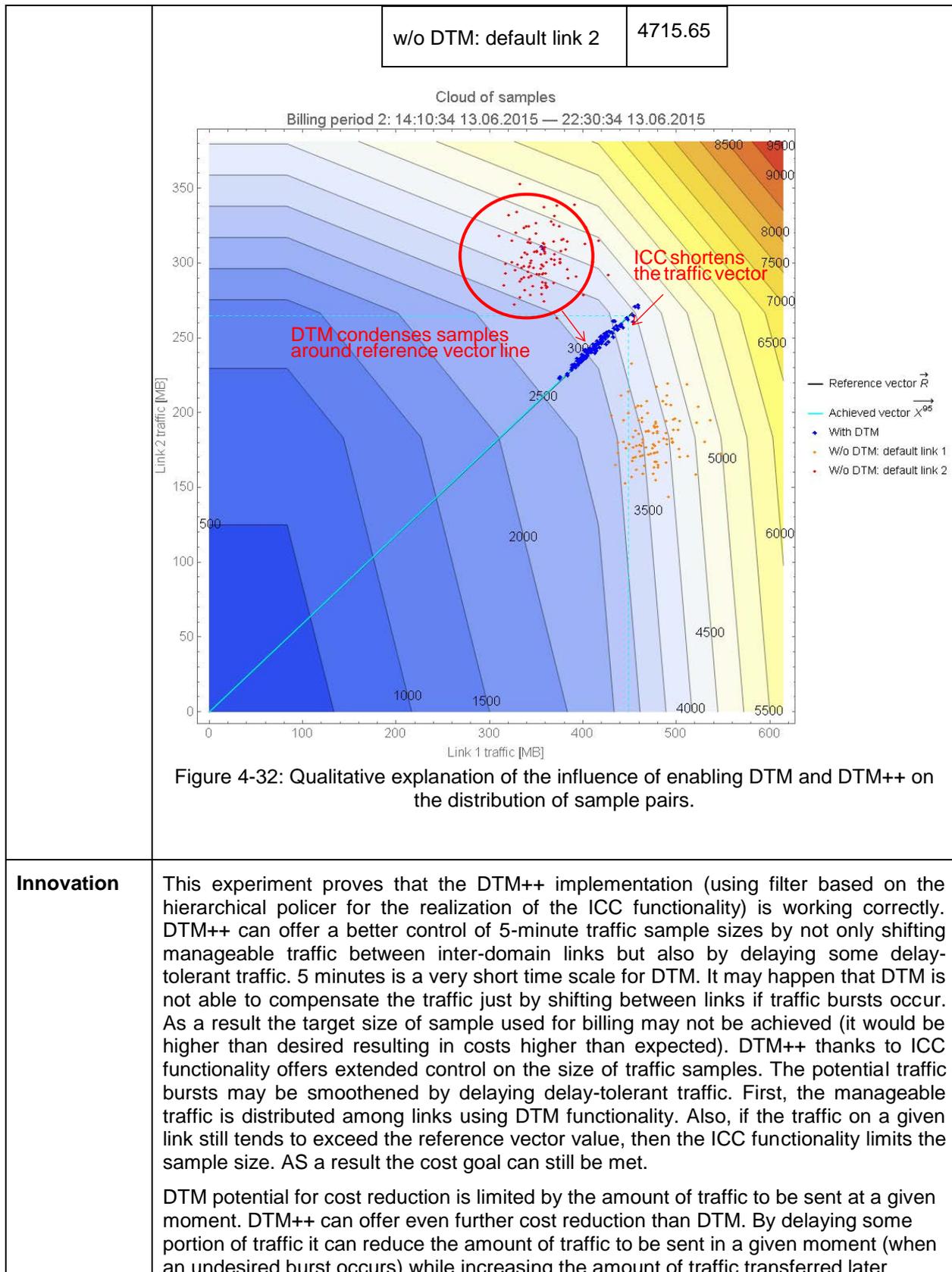


Figure 4-31: Traffic trace on links 1 and 2 when only DTM is used (ICC functionality switched off).

Table 4-4. Traffic costs for DTM++, DTM and no traffic management.

Costs	
Achieved with DTM ++	3514.31
Achieved with DTM	3764.09
w/o DTM: default link 1	4759.12



	(when there is no or limited burst). Thus, in the case of DTM++ a more conservative (and at the same time more optimistic) algorithm for calculating a reference vector can be considered and the target cost can be lower than in the case of DTM.
--	---

4.1.5 Model for Cloud Federation: Investigation of Pricing Aspects

Use case name	Bulk data transfer for cloud operators
Scenario	Operator Focused Scenario
Goal	<p>We considered the problem of the formation of an economically sustainable federation of computational resources among Cloud Service Providers (CSPs).</p> <p>We aim to model the service that a CSP offers to its customers and to define certain policies for the formation of a federation that:</p> <ul style="list-style-type: none"> • Guarantee an increased profit for the CSPs joining the federation. In the worst case a CSP’s profit should be equal to the profit in the standalone operation of this CSP. • Achieve an improved QoS in terms of average total delay per job for customers served by the federated CSPs. <p>We investigate the problem of federation formation both when the CSPs are cooperate or non-cooperative and we design incentive compatible pricing mechanisms that achieve mutual benefits for CSPs in both cases.</p>
Overview	<ul style="list-style-type: none"> • We develop an abstraction model for the CSP’s infrastructure and service as an M/M/1 queueing system. • We model the salient factors that determine the net benefit of a CSP, i.e. a pricing function that each CSP uses to charge its clients, and the cost from energy consumption at servers. • We model a federation policy among CSPs as the transfer of a portion of jobs’ requests from one CSP to others in order to be served through their server infrastructure. • We formulate the problem of finding the federation policy that maximizes the total profit for the CSPs, we find the optimal federation policy as the solution of nonlinear optimization problem and we provide a rule for the sharing of the generated profit of the federation. <p>Note: for simplicity, we henceforth deal with the case of two CSPs, which though is adequately indicative.</p>

	<p>Figure 4-33: Federation scenario for two CSPs, each modelled as an M/M/1 queue. The amount of job requests that is transferred to the other CSP undergoes a given average delay D over the Internet.</p>
<p>Parameters</p>	<p>The key parameters for the experiments conducted:</p> <ul style="list-style-type: none"> • Arrival rate of incoming requests in both CSP queues, expressed in requests per unit of time λ_1, λ_2 • Average load per job (flops): L • Computational capacity of CSPs' infrastructure, flops per unit of time : C_1, C_2 • Maximum price that a customer of each CSP should pay per job, for simplicity applying to the ideal case when the average delay per job tends to zero: x_1, x_2 • Pair of idle and total power consumption of the infrastructure of each CSP: $[W_{i,1}, W_{t,1}]$ and $[W_{i,2}, W_{t,2}]$ • Energy cost, expressed as price per Watt*sec that each CSP pays to its electricity provider: q_1, q_2 • Average network delay introduced by the transfer process over the Internet between servers of the two CSPs: D • Sensitivity of the CSPs pricing to QoS degradation: a^* • Portion of the stream of outsourced requested for each CSP: a_1, a_2
<p>Metrics</p>	<ul style="list-style-type: none"> • Individual profit of each CSP: $P_1(a_1, a_2), P_2(a_1, a_2)$ • Total profit of the federation: $P_{tot}(a_1, a_2) = P_1(a_1, a_2) + P_2(a_1, a_2)$ • Average delay per job, as metric for the customers' QoS
<p>Traffic Management Solutions</p>	<p>Economically sustainable cloud federation through service delegation.</p>
<p>Evaluation framework</p>	<p>We simulate an environment of two CSPs in standalone, in weak federation and in strong federation modes of operation. By standalone we mean that</p>

	<p>CSPs act in isolation from each other and serve only their own clients.</p> <p>We assume that the average number of processor flops L that a job requires in order to be completed is the same for both CSPs and equals to 2. (This value is selected arbitrarily; any other value would just serve as a normalizing factor.) We define the network delay D that models the average intermediate delay experienced by a request that is transferred from one CSP to the other. We assume that this delay is a small fraction of the delays d_1, d_2, since this is expected to be the case in reality. Thus, we set D to be an order of magnitude lower than d_1 and d_2. In our experiments we generate the value of the maximum prices x taking as input the electricity price q; given the price q we find the value of x_1, x_2 for which the profit of CSP becomes zero when the utilization factor approaches 1 and finally we choose x_1, x_2 at least one order of magnitude higher. This guarantees that both CSPs in standalone operation will not have negative profit under any value of utilization ρ. Next, we assume that the CSP₂ has fixed rate of incoming requests λ_2 and we set values for λ_1 in the range of values that does not make the queue of CSP unstable (should $\frac{C_i}{L} > \lambda_i$), from 1 to 9.9 with a step of 0.1. We run this type of experiment for different fixed values of λ_2 for 1 to 9.9.</p> <p>We ran experiments to compare the weak and strong federation with the standalone operation, considering either symmetric or asymmetric CSPs.</p> <ul style="list-style-type: none"> • Symmetric CSPs: In the first set of experiments we assume that CSP₁ and CSP₂ are symmetric with respect to their infrastructure $C_1 = C_2 = 20 \text{ flops/sec}$. For the power consumption of the servers we take $W_0 = 300 \text{ KWatt}$ and $W_1 = 1000 \text{ KWatt}$. We also assume that both CSPs pay the same price, $q = 15 \text{ \\$/KWatt} \cdot \text{sec}$ to their electricity provider, while they charge their clients according to the same pricing function, with the same maximum price $x \text{ \\$/job}$ when $d_i \rightarrow 0$ and sensitivity parameter $d^* = 1 \text{ sec}$. • Asymmetric infrastructure-symmetric pricing: We run the same type of experiments for asymmetric CSPs with respect to their infrastructure, i.e. $C_1 \neq C_2$. Since we assume that the power consumption of servers is related to their processing power, the CSPs are also asymmetric with respect to their power consumption. Therefore, we consider three different values of CSP dimensioning and the corresponding power consumption, $C = \{10; 20; 40\}$ and $(W_0; W_1) = \{(300; 1000), (600; 2000), (1200; 4000)\}$. Then we try all possible combinations of elements in the sets above for CSP₁ and CSP₂. • Asymmetric infrastructure and pricing: We also run a set of experiments for symmetric infrastructure, but now we assume that the energy price q_i and the maximum price per job x_i are asymmetric among CSPs. In particular we consider three different values of electricity price $q_i = \{5, 10, 15\}$ and we assign the x_i's produced from them to two identical CSPs.
Evaluation results	<p>Next, we provide results for the benefits of the strong federation vs standalone operation.</p> <p>Assuming that the CSPs are symmetric, the results Figure 4-34 show that</p>

for $\lambda_2 = 9$ and for low to medium load λ_1 , the federation leads to 50-100% more total profit compared to the case where each CSP serves its own clients. The benefits of the federation diminish as λ_1 tends to λ_2 . In the case where both CSP input loads are equal ($\lambda_1 = \lambda_2 = 9$) the benefit of the federation is zero due to the overall symmetry. These results provide valid guidelines regarding in which cases a federation is most profitable. That is, the more diverse the loads are, the more pronounced the benefit of the federation is. This can be explained by the fact that optimal federation balances the loads in the two servers appropriately.

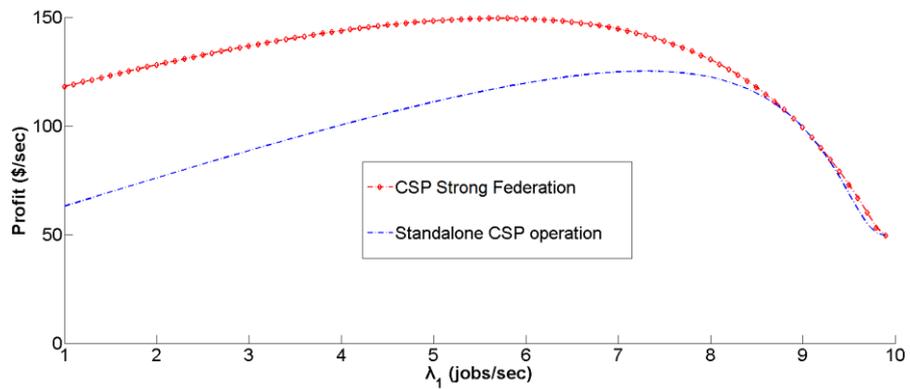


Figure 4-34: Maximum total profit of strong federation for $\lambda_2 = 9$ and $\lambda_1 \in [1, 9.9]$.

Figure 4-35 shows the individual profit for each CSP is higher than this in the standalone operation. Note that this is the profit of each CSP after the application of the cooperative profit sharing rule.

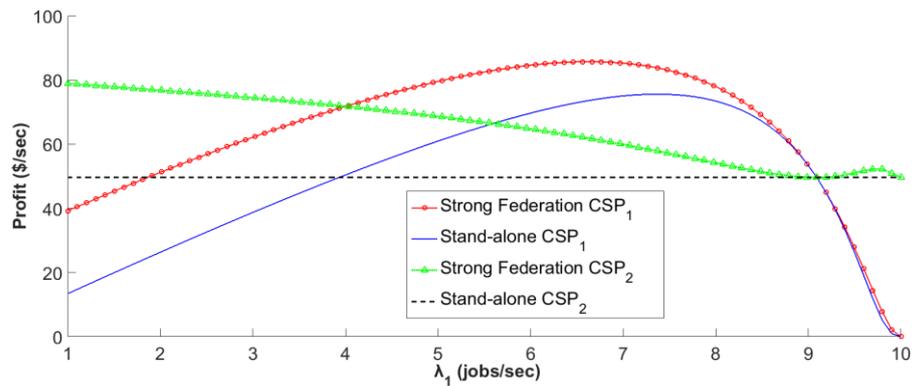


Figure 4-35: Individual CSP profit strong federation vs standalone, for $\lambda_2 = 9$ and $\lambda_1 \in [1, 9.9]$.

In Figure 4-36 we provide results for the average delay. For the average delay of clients of federated CSPs, it is shown that the optimal federation policy achieves an average delay that coincides or is close to the optimal average delay with respect to a QoS-based federation policy, namely a policy when the objective of the optimization problem includes only the delay. Therefore, the formation of a federation is incentive compatible for both the CSPs and the users.

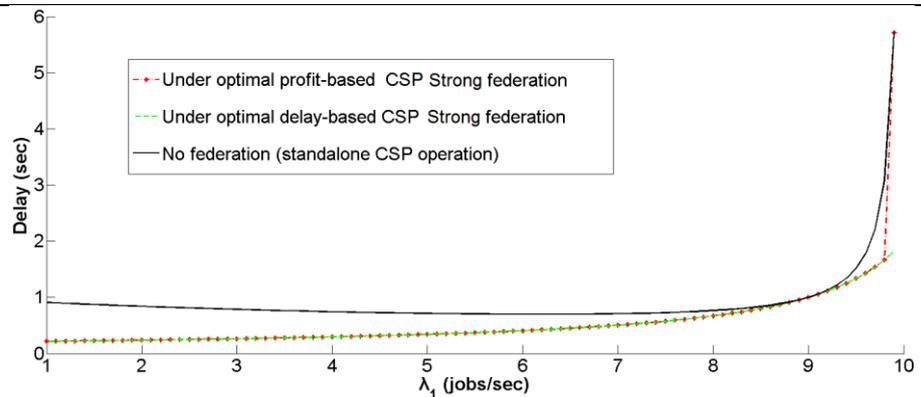


Figure 4-36: Average provisioned delay of all customers in the environment of two CSPs under different policies.

In Figure 4-37, we can observe that in the optimal solution of the optimization problem that gives the optimal federation policy, at least one of a_1 and a_w equals zero, while the non-zero value always corresponds to the most utilized CSP. When $\lambda_1 = \lambda_2$, then both a_1 and $a_2 = 0$, and the optimal federation coincides with standalone operation.

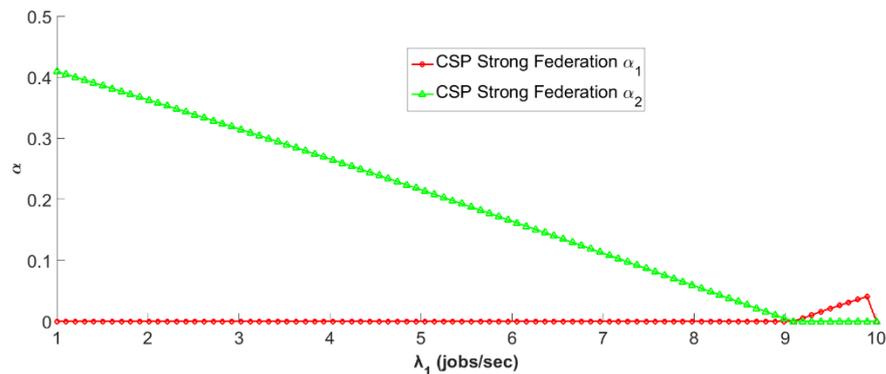


Figure 4-37: Optimal pairs (a_1^*, a_2^*) that denote the portions of request traffic transferred from one CSP to the other, for $\lambda_2 = 9$ and $\lambda_1 = [1, 9.9]$.

Moreover, the results in Figure 4-38 show that as the intermediate average delay D increases, the CSPs follow a more conservative traffic transfer strategy, and when D exceeds a certain high value, both a_1 and a_2 becomes zero. Consequently, as D increases the effectiveness of federation decreases, and thus the maximum total profit decreases and after a certain value it is best in terms of the total profit not to federate.

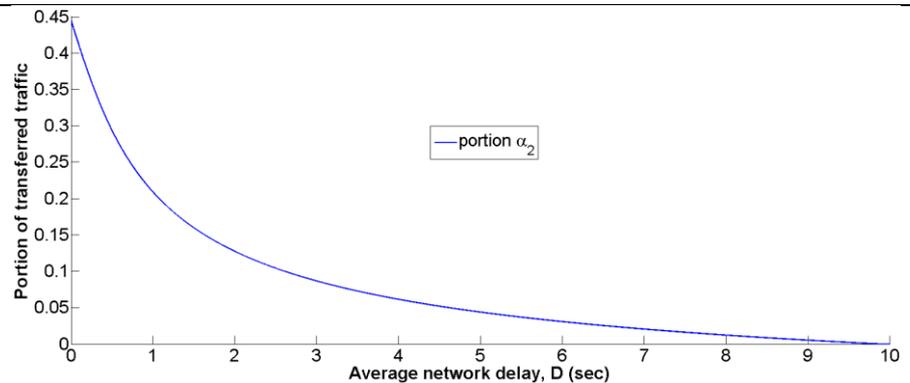


Figure 4-38: Optimal pairs (a_1, a_2) that denote the portions of request traffic transferred from one CSP to the other as a function of D , for $\lambda_1 = 1$ and $\lambda_2 = 9$. Note that $\alpha_2 = 0$.

For the asymmetric infrastructure case, the results reveal that the parameter that affects more the effectiveness of federation is again the utilization factor of the infrastructure of each CSP. In addition, when the largest of the two CSPs has a high utilization factor and the other has a low to medium utilization factor, then the federation can achieve higher benefit than in the opposite case of asymmetric CSPs, but it is also better than the case of identical CSPs.

For the asymmetric pricing scenario, the results show that in the case where utilization factors of the CSPs differ significantly, if the highly utilized CSP is the one with the highest value of x_i , then the benefit of the federation is higher and the portion of requests that are outsourced (i.e. a_1 or a_2) increase by up to 25% compared to the case of symmetric pricing. On the other hand, when the CSP with the lowest utilization has the highest x_i , then the benefits and the portion of outsourced requests decrease by up to 25% compared to that of symmetric pricing. The effect of price asymmetry is less pronounced when the CSPs have the same utilization level.

Next, we provide some results from the ongoing investigation and evaluation of the weak federation.

We assume that the CSPs adopt the scheme for aggregate compensation, i.e. an aggregate payment for the total stream of outsourced requests. In this case, we assume that each CSP compensates the other by paying half of the extra profit he manages to obtain from the requested outsourcing action. Figure 4-39 shows that the total profit of the weak federation is exactly the same with that of the strong federation presented in Figure 2. This happens because our aggregate compensation rule achieves an indirect maximization of the global profit, since the pair (α_1^*, α_2^*) in the equilibrium point is same with the optimal pair in the strong federation.

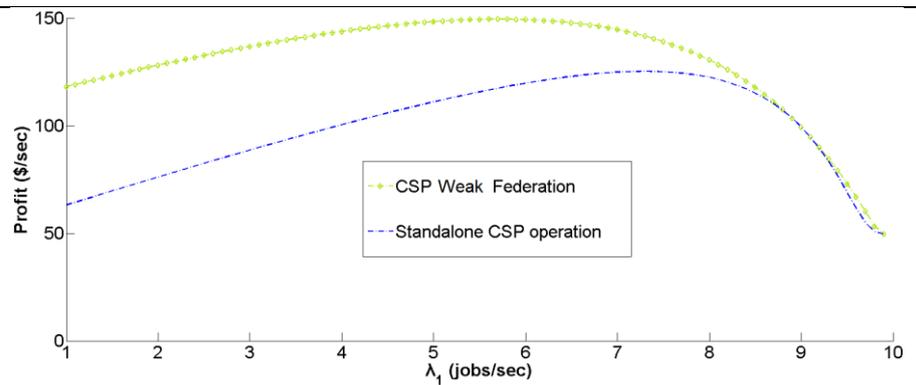


Figure 4-39: Maximum total profit of weak federation for $\lambda_2 = 9$ and $\lambda_1 \in [1,9.9]$.

The individual profit of each CSP in weak federation is depicted in Figure 4-40. Comparing this with Figure 4-35, it follows that this profit close but not the same as in the strong federation, since the compensation rule performs a different split than the profit sharing rule of compensation. However, the individual profit of each CSP is again greater than the profit of the standalone operation, thus ensuring incentive compatibility.

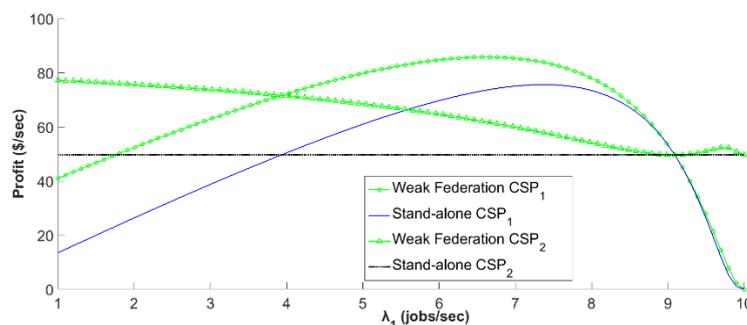


Figure 4-40: Individual CSP profit weak federation vs standalone, for $\lambda_2 = 9$ and $\lambda_1 \in [1,9.9]$.

<p>Innovation</p>	<p>We provide a performance and economic model of the federated environment of CSPs and we investigate utility-based optimal federation formation policies. Furthermore, we take into account the QoS offered to CSPs' clients in their optimization approach. In our work, the federation policy is optimal with respect to total CSPs' profit, but it is also beneficial and caters for client benefit, since a larger profit for the CSP is accompanied with better QoE for clients. Finally, we introduce incentive compatible profit sharing and pricing schemes.</p>
--------------------------	--

4.2 Parameters, Metrics and Evaluation of SmartenIT mechanisms for EFS

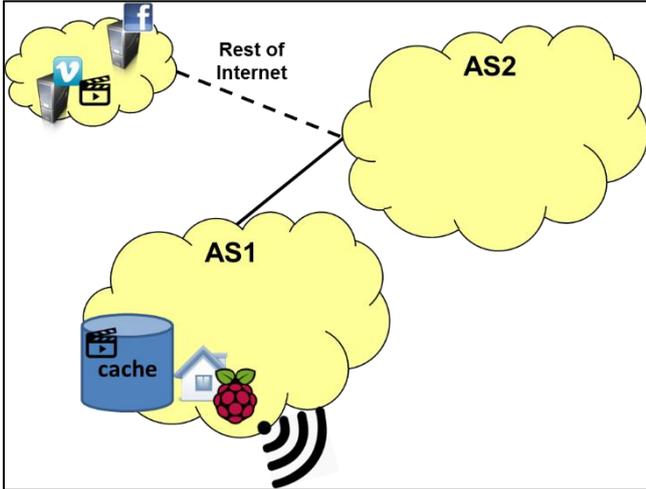
This section documents the results for the SmartenIT traffic management mechanisms belonging to the End user-Focused Scenario. Depending on the mechanisms, results are

either a completion of evaluations presented in D2.4, or full description and evaluation reports. Those mechanisms are briefly explained below as well as their evaluation goals:

- **Replicating Balanced Tracker and Home Router Sharing based on Trust (RB-HORST)**, which provides WiFi access to trusted users to offload their mobile traffic from 3/4G to WiFi. RB-HORST also enables home routers of trusted users identified through a social network to be organized in a content-centric overlay network supporting content prefetching. Results are focused on 1) refinement of caching strategy including flash crowd situation handling, and 2) on performance when demand exhibits temporal dynamics. Moreover it integrates consideration on the performance brought by social awareness in caching strategy.
- **Socially-aware Efficient Content Delivery (SEConD)**, which employs social information, AS-locality awareness, chunk-based P2P content delivery and prefetching. A centralized node is acting as cache and as P2P tracker with a proxy to improve the Quality of Experience of video streaming for users of OSNs to reduce inter-AS traffic. Results are focused on the inter-domain traffic reduction and the investigation of exploitability of peering links toward this direction.
- **Virtual Incentives (vINCENT)**, which aims to leverage unused wireless resources among users, while it ensures security by employing tunneling among trusted users. vINCENT exploits social relationships derived by the OSN, as well as interest similarities and locality of exchange of OSN content, and addresses the asymmetries between rural and urban participants of an offloading scheme to derive fair incentives for resource sharing among all users.
- **Mobile Network Assistant (MONA)**, which schedules wireless data transmissions to reduce the energy expenses on the air-interfaces. Results consider traffic on the cell interface and WiFi access points with realistic data rates.
- **RB-HORST++** employing features of RB-HORST, SEConD, vINCENT and MONA, in order to perform content prefetching and mobile to WiFi offloading in an energy efficient manner, while ultimately further improving end-users' Quality of Experience. Evaluation results are provided capitalizing on two synergy combinations: one combining trust RB-HORST, and SEConD to address the content placement use-case and one combining RB-HORST and MONA to address the WiFi offloading use-case.
- **Multi-Criteria Application Endpoint Selection (MUCAPS)**, which improves users' Quality of Experience by performing selection of communication endpoints, by employing basic ALTO functionality but also involving awareness on the underlying network topology. Evaluation results are provided for video streaming services delivery.

4.2.1 RB-HORST

Use case name	Exploiting content locality, Service and content placement for users
Scenario	End-user Focused Scenario
Goal	Performance evaluation of RB-HORST overlay <ul style="list-style-type: none"> • What is performance of the RB-HORST cache overlay in terms of hit

	<p>rate?</p> <ul style="list-style-type: none"> • How much inter-domain traffic can be saved? • How much requests can be served by the RB-HORST overlay? • How much load can be taken off the ISP cache?
<p>Figure</p>	 <p>The diagram illustrates the topology of RB-HORST. It features three main components: a 'Rest of Internet' cloud at the top left containing icons for Facebook, Twitter, and YouTube; a central cloud labeled 'AS1' which contains a blue 'cache' cylinder, a house icon, a Raspberry Pi, and a WiFi signal icon; and a cloud labeled 'AS2' at the top right. A dashed line connects the 'Rest of Internet' cloud to AS2, and a solid line connects AS2 to AS1.</p> <p>Figure 4-41: Topology of RB-HORST. Local UNaDas are used for caching and WiFi offloading.</p>
<p>Parameters</p>	<ul style="list-style-type: none"> • Cache replacement strategy LRU • Content demand model Zipf • Content popularity distribution • Catalogue size (number of objects provided by content provider) • Autonomous system size • Cache capacity • Sharing probability (pshare)
<p>Metrics</p>	<ul style="list-style-type: none"> • Cache hit rate • ISP cache contribution • Inter-domain traffic
<p>Traffic Management Solutions</p>	<p>RB-HORSTORST, SECOND, MONA</p>
<p>Evaluation framework</p>	<p>Content Delivery Simulation Framework, c.f. D2.4, Caching Evaluation Framework [7].</p>
<p>Evaluation results</p>	<p>To evaluate the performance of the overlay, an autonomous system is</p>

considered given the autonomous system size in terms of the number of end-users in the autonomous system. The probability that an end-user has RB-HORST installed and shares its home router (HR) for content is given by p_{share} . The probability that a user requests certain content items depends on the content’s popularity distribution, which is specified by a Zipf distribution with exponent α . The impact of the Zipf parameter on the cache efficiency has been shown in various studies. As we aim to assess the performance of the overlay we dependent on sharing probability and AS size, we fix α to 0.8. To evaluate the performance of the overlay, two cases are considered, (a) the tree case and (b) the overlay case. In the tree case (a), each user is assigned to one shared home router (HR) in its AS which runs RB-HORST. If a user shares its HR, it is assigned only to its HR. We consider a tiered caching architecture with three tiers. In tier-3 are caches deployed on shared HRs. Tier-2 cache is the cache hosted by the ISP. Tier-1 is the data-center of the content provider which can serve all requests. A requested item is looked up in the assigned HR initially, i.e. in the tier-3 cache. If the requested item is not found, the request is forwarded to the next tier. The hierarchical caching strategy is leave-copy-everywhere, which means that the object is cached in each cache on the look up path. In the overlay case (b), a requested item is looked up in the HR of the user. If it is not found, it is looked up in shared HRs in the same autonomous system using the overlay. If no tier-3 cache in the AS contains the item it is looked up in tier-2 caches and finally in the data center of the content provider. The hierarchic caching strategy is leave-copy-everywhere, with the constraint, that the item is cached in the tier-3 cache only, which was looked up first.

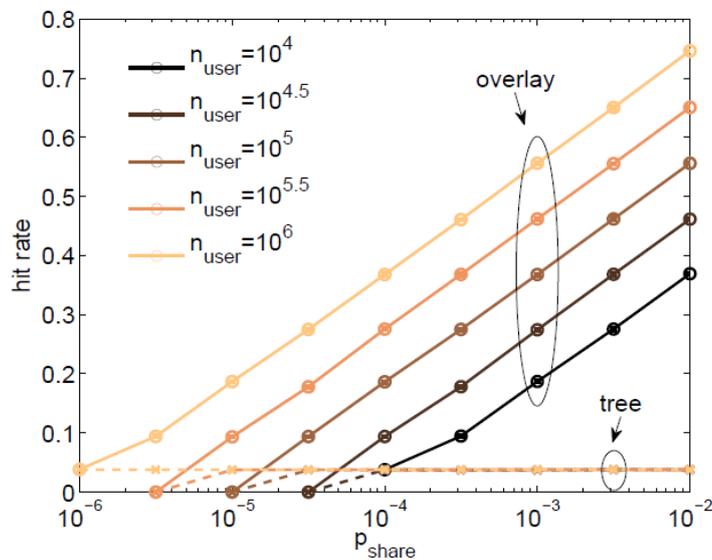


Figure 4-42: Hit rate of the overlay dependent on sharing probability.

Figure 4-42 shows how the hit rate of the overlay depends on the sharing probability. If the home routers are organized in an overlay their hit rate and contribution increases since an item is looked up in each home router in the same AS.

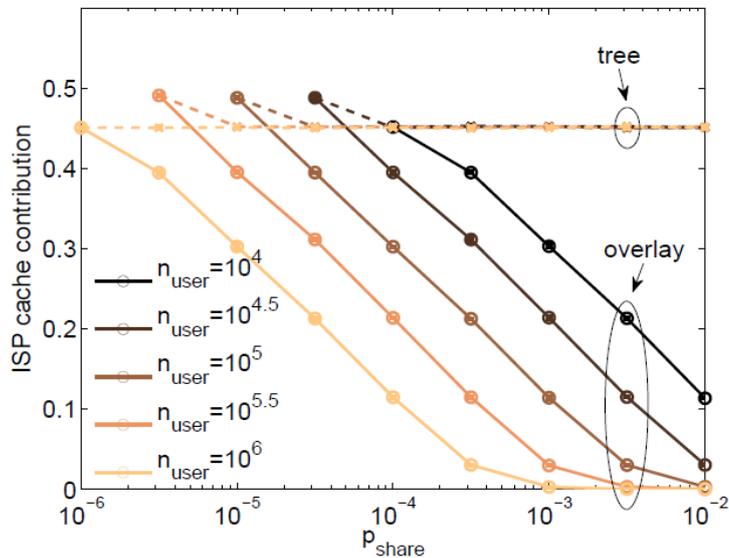


Figure 4-43: ISP cache contribution dependent on sharing probability.

Figure 4-43 shows how the ISP cache contribution depends on the sharing probability. In a tree structure the ISP cache contribution is independent of the home router sharing probability. If the home routers are organized in an overlay the ISP cache contribution decreases, because of the lower hit rate of the ISP cache.

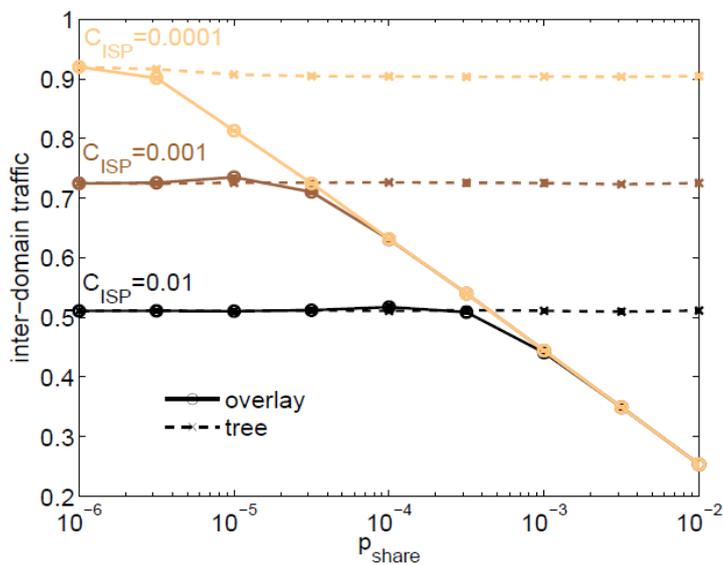


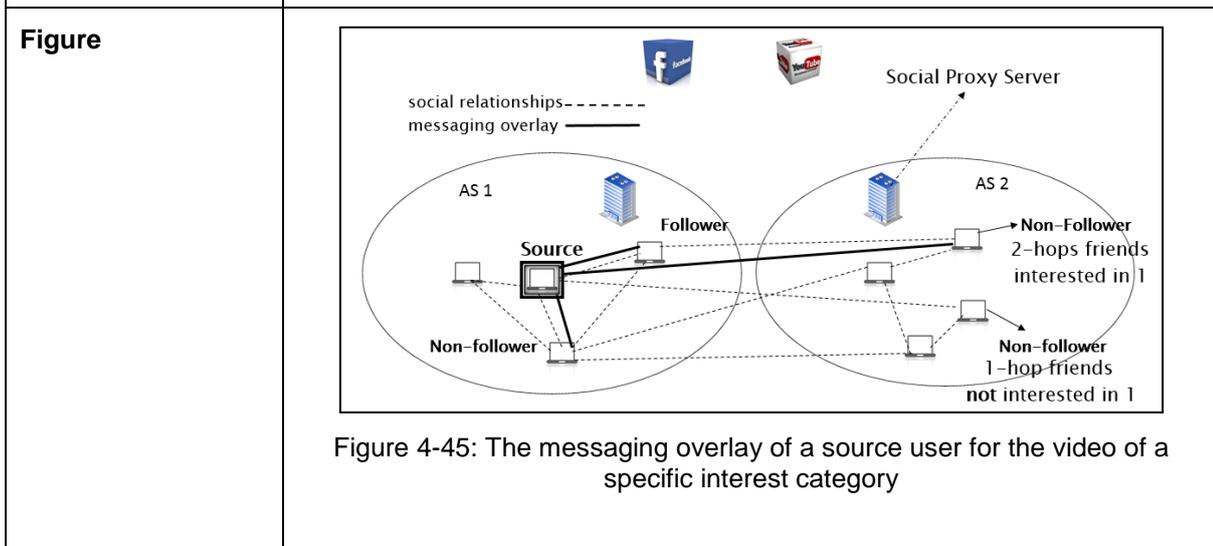
Figure 4-44: Inter-domain traffic dependent on sharing probability.

Figure 4-44 shows how the inter-domain traffic depends on the sharing probability. In the overlay case the number of requests served locally increases with the sharing probability which reduces the inter-domain traffic. If no overlay is present the amount request served locally is independent of the sharing probability. This shows the importance of an overlay for a

	mechanism like RB-HORST.
Innovation	RB-HORST Overlay increases the share of requests served locally and reduces costly inter-domain traffic. The load on the ISP cache can be reduced depending on the home router sharing rate. The ISP cache can be dimensioned accordingly to save energy. If the amount of shared home routers deploying caches increases less ISP cache servers have to be deployed to achieve the same cache efficiency within the AS.

4.2.2 SEConD

Use case name	Service and content placement for users
Scenario	End-user Focused Scenario
Goal	SEConD aims to: <ul style="list-style-type: none"> • Enhance the OSN users’ QoE by eliminating stall events during video viewing. The means to this improvement is the effective video prefetching and the peer-assisted and QoE-oriented video delivery. • Reduce the contribution of origin server in video delivery, and thus its relevant operational costs. • Reduce the inter-domain transit traffic and exploit peering links, which may lead to the reduction of transit inter-connection cost. • Restrict redundant traffic, i.e., traffic generated by prefetching the same content item several times, in order to mitigate intra-domain congestion



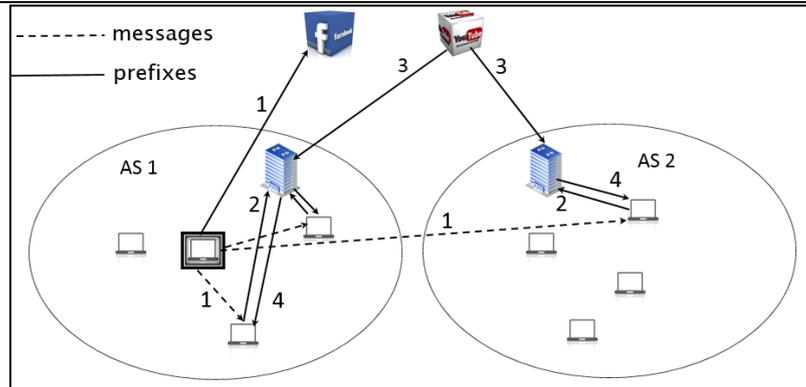


Figure 4-46: An example presenting the sequence of steps performed by the prefetching algorithm – The source node shares a video hosted in a third-party owned server

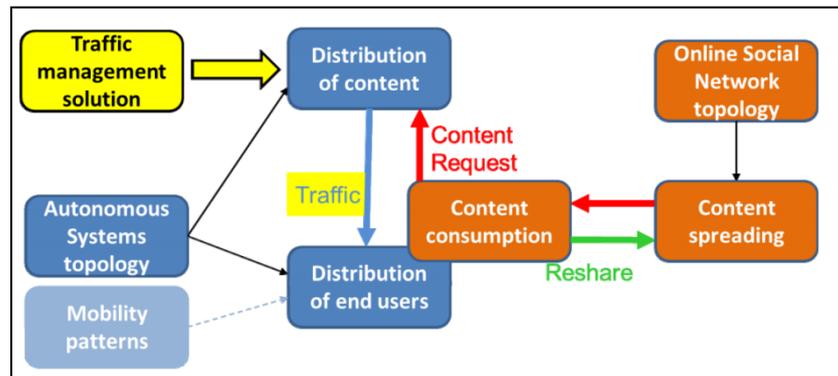


Figure 4-47: Parameters for service and content placement use case.

<p>Parameters</p>	<ul style="list-style-type: none"> • SPS cache size: The storage size of the SPS cache. • User-owned cache size: The storage size of user-owned caches. • Users’ distribution per AS: Distribution of user into ASes (Zipf). • Total upload bandwidth within a swarm: The total upload bandwidth within the swarm as a threshold for SPS participation in content-based P2P overlay. • Social Tie thresholds: Thresholds that decides the distribution of the audience of a user into viewer categories. The choice is based on the percentage of the videos of the uploader the viewer has watched and on his interests. • Users’ online time: The time that the users are online in order to support in video delivery. • Interest categories and their distribution over users • Social Graph • Content characteristics: demand and response
<p>Metrics</p>	<ul style="list-style-type: none"> • Inter/intra AS traffic • Contribution of server hosting the video • Caching accuracy • Prefetching accuracy • Useless and redundant prefetching

<p>Traffic Management Solutions</p>	<p>SEConD</p>																														
<p>Evaluation framework</p>	<p>Documented in D2.4 – Subsection 6.2.3 [7] and D2.2 – Subsection 3.2.1.1 [5]</p>																														
<p>Evaluation results</p>	<ul style="list-style-type: none"> • 87% average reduction of the total inter-AS traffic, compared to the inter AS traffic created from the traditional client-server architecture. • 88% average reduction to the contribution of the origin server where the video is hosted. • 80% average prefetching accuracy. • Minimizes redundant traffic in inter-AS links both for prefixes and videos. • Eliminates redundant prefetching, i.e., traffic originating from multiple sources with the same prefix to the same destination. <div data-bbox="528 880 1283 1290" data-label="Figure"> </div> <p>Figure 4-48: Total inter-AS traffic during the day compared to Client server and SocialTube [17].</p> <table border="1" data-bbox="523 1384 1289 1756"> <thead> <tr> <th></th> <th>AS1</th> <th>AS2</th> <th>AS 3</th> <th>AS4</th> </tr> </thead> <tbody> <tr> <td>Number of Users</td> <td>1925</td> <td>928</td> <td>634</td> <td>476</td> </tr> <tr> <td>Proxy Contribution</td> <td>48%</td> <td>61%</td> <td>75%</td> <td>79%</td> </tr> <tr> <td>Proxy Cache Size</td> <td>63GB</td> <td>30GB</td> <td>21GB</td> <td>15.7GB</td> </tr> <tr> <td>Proxy Cache Hit Accuracy (Videos/prefixes)</td> <td>90/94%</td> <td>78/86%</td> <td>68/80%</td> <td>60/74%</td> </tr> <tr> <td>External Server Contribution</td> <td>5%</td> <td>14%</td> <td>24%</td> <td>31%</td> </tr> </tbody> </table> <p>Figure 4-49: Proxy server and SPS contribution w.r.t. the AS size.</p>		AS1	AS2	AS 3	AS4	Number of Users	1925	928	634	476	Proxy Contribution	48%	61%	75%	79%	Proxy Cache Size	63GB	30GB	21GB	15.7GB	Proxy Cache Hit Accuracy (Videos/prefixes)	90/94%	78/86%	68/80%	60/74%	External Server Contribution	5%	14%	24%	31%
	AS1	AS2	AS 3	AS4																											
Number of Users	1925	928	634	476																											
Proxy Contribution	48%	61%	75%	79%																											
Proxy Cache Size	63GB	30GB	21GB	15.7GB																											
Proxy Cache Hit Accuracy (Videos/prefixes)	90/94%	78/86%	68/80%	60/74%																											
External Server Contribution	5%	14%	24%	31%																											
<p>Innovation</p>	<ul style="list-style-type: none"> • Multiple socially-aware messaging overlays are created per user, in each of which he can send demand indications to his friends for specific content items, in order to proactively store the prefix (first chunk) of those items. • Socially-aware Proxy Server (SPS) is employed in each Autonomous 																														

	<p>System (AS) in order to: orchestrate the formation of messaging overlays, to operate as P2P tracker for local content-based P2P overlays, and to achieve high traffic localization, by caching content to assist in sharing, when the local P2P is not adequate.</p> <ul style="list-style-type: none"> • SEConD employs a novel caching strategy, based on the demand patterns of OSNs rather than on general popularity of content.
--	---

4.2.3 vINCENT

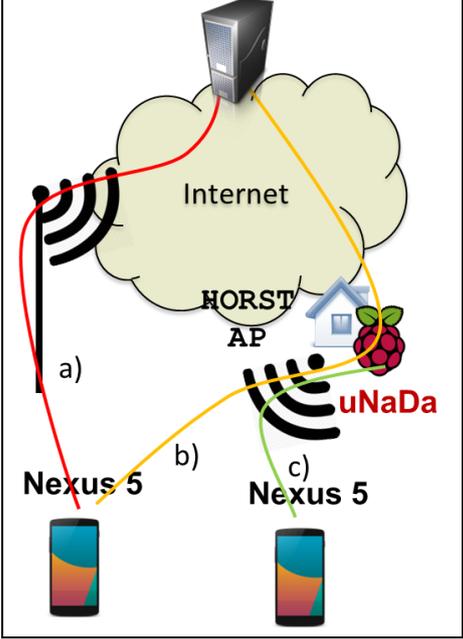
Use case name	<i>Socia- aware Mobile Data Offloading</i>
Scenario	End-user Focused Scenario
Goal	<p>vINCENT aims at creating incentives to users to provide WiFi network access to even unknown parties. In the scheme, each user participates with his mobile phone and his uNaDa. Service is granted by allowing other mobile users offloading at the uNaDa. Despite different user densities (e.g. rural areas vs. city areas), the scheme should be fair to all users and be able to isolate free riding nodes, i.e., mobile user consuming offloading resources while not contributing any service with their uNaDa. The height of service each mobile may receive is calculated by a Facebook App that has a global view on the contribution granted at all uNaDas.</p> <p>In a typical usage scenario, a mobile user offloads traffic to an access point whenever possible using an app on his smartphone. The more offloading capabilities are provided by the mobile user's own uNaDa, the higher the QoS that can be received.</p>
Figure	<p>Figure 4-50: vINCENT system overview. The traffic management mechanism aims at allowing for offloading at untrusted uNaDas. In this case, untrusted refers to peers not being friends in the social network.</p>
Parameters	<ul style="list-style-type: none"> • Accounting Method: Method used for accounting contributed service to the system (e.g., provided offloaded traffic volume) • Parameterization Regression Model: Parameterization of input data for regression model (e.g., size of the radius for access point density calculations) • The scenario parameters, e.g., user densities, probability of offloading

	<p>were taken from real-world mobility traces as described in D2.4.</p>																							
<p>Metrics</p>	<ul style="list-style-type: none"> • Percentage of offloaded traffic volume: The overall traffic volume that could be offloaded to WiFi during the simulation run. • Percentage of offloaded data by free riding nodes: Traffic volume that could be offloaded by free riding peers during the simulation run. 																							
<p>Traffic Management Solutions</p>	<p>RB-HORST, vINCENT</p>																							
<p>Evaluation framework</p>	<p>OMNeT++, INETMANET, BonnMotion, Python scripts as described in D2.4.</p>																							
<p>Evaluation results</p>	<div style="text-align: center;"> <p>Mean Offloaded Volume Total</p> <table border="1"> <thead> <tr> <th>Scheme</th> <th>Mean Offloaded Volume</th> </tr> </thead> <tbody> <tr> <td>vINCENT</td> <td>~0.82</td> </tr> <tr> <td>Unrestricted (no incentive)</td> <td>1.0</td> </tr> <tr> <td>Tit-for-Tat</td> <td>~0.75</td> </tr> </tbody> </table> </div> <p>Figure 4-51: Mean Offloaded volume.</p> <p>Figure 4-51 shows the mean offloaded volume for three different cases, The green bar shows the vINCENT scheme allowing for an increase of 5% overall offloaded volume as opposed to a Tit-for-Tat scheme (3). This increase in efficiency is driven by the possibility to level out contribution asymmetry caused by population differences of rural and municipal areas. Moreover, vINCENT is closer to the unrestricted case, in which all peers always have unrestricted access to all access points nearby.</p> <div style="text-align: center;"> <p>Mean Offloaded Volume per User/Free Rider</p> <table border="1"> <thead> <tr> <th>Fraction of Free Riders</th> <th>Contributing User (Mean)</th> <th>Free Rider (Mean)</th> </tr> </thead> <tbody> <tr> <td>0.1</td> <td>~0.55</td> <td>~0.75</td> </tr> <tr> <td>0.2</td> <td>~0.50</td> <td>~0.75</td> </tr> <tr> <td>0.5</td> <td>~0.48</td> <td>~0.75</td> </tr> <tr> <td>0.8</td> <td>~0.40</td> <td>~0.65</td> </tr> </tbody> </table> </div> <p>Figure 4-52: Mean offloaded volume free riders vs. contributing users.</p>	Scheme	Mean Offloaded Volume	vINCENT	~0.82	Unrestricted (no incentive)	1.0	Tit-for-Tat	~0.75	Fraction of Free Riders	Contributing User (Mean)	Free Rider (Mean)	0.1	~0.55	~0.75	0.2	~0.50	~0.75	0.5	~0.48	~0.75	0.8	~0.40	~0.65
Scheme	Mean Offloaded Volume																							
vINCENT	~0.82																							
Unrestricted (no incentive)	1.0																							
Tit-for-Tat	~0.75																							
Fraction of Free Riders	Contributing User (Mean)	Free Rider (Mean)																						
0.1	~0.55	~0.75																						
0.2	~0.50	~0.75																						
0.5	~0.48	~0.75																						
0.8	~0.40	~0.65																						

	Figure 4-52 shows the performance drop free riding nodes, i.e., nodes only consuming service from the system while not providing any service with their own uNaDas. The difference in performance is as high as 20%, regardless of the fraction of free riders in the system.
Innovation	Current incentive schemes for mobile offloading don't address user's locality, leading to unfairness in the offloading process. The proposed TM mechanism identifies the asymmetries between rural and urban participants of an offloading scheme to derive fair incentives for all users.

4.2.4 MoNA

Use case name	Social-aware Mobile Data Offloading
Scenario	End-user Focused Scenario
Goal	Mobile data consumption has seen tremendous growth in the recent past and studies predict an ongoing increase in the coming years. The sudden rise in traffic demand is mainly caused by mobile video consumption (e.g. YouTube). A promising solution to this problem is offloading mobile cellular traffic to WiFi. Knowing the state of the network (i.e. estimated RTT, estimated throughput based on past throughput measurements and utilization patterns) before initiating data transfers, and offloading data on WiFi networks, allows increasing the user experience. Combining this a-priori knowledge of the network with power models of the available interfaces allows predicting the expected energy consumption of each transfer mode, and such modifying scheduling decisions to reduce the energy expense while keeping the user experience stable. Exploiting the locality of content in RB-Horst, the energy efficiency can further be increased by utilizing the fact that locally available content transferred via WiFi (eliminating the possible bottleneck on the uplink) consumes less energy than regular offloading without cached content.
Figure	

	 <p>Figure 4-53: Selecting the most energy efficient transmission technology. Options: a) 3G b) conventional WiFi offloading c) social aware offloading.</p>
<p>Parameters</p>	<p>Dependent on access technology:</p> <ul style="list-style-type: none"> • Location based bandwidth distribution • Location based latency distribution • Location based energy consumption <p>WiFi sharing probability WiFi access point signal range WiFi access point coordinates End user location distribution</p>
<p>Metrics</p>	<ul style="list-style-type: none"> • Data volume offloaded from cellular network to WiFi • Available bandwidth to fulfill request • Received latency to fulfill request • Consumed energy on mobile device during data transfer
<p>Traffic Management Solutions</p>	<p>RB-HORSTORST, MONA</p>
<p>Evaluation framework</p>	<p>Mobile Offloading Simulation Framework, c.f. D2.4.</p>

Evaluation results

To evaluate the potential of MONA, both the throughput and the energy efficiency of the mobile connections were determined depending on the access technology. The throughput of mobile connections is evaluated by deterministic network measurements, executed on a variety of hardware devices. The data was gathered using the NetworkCoverage App I.) in a crowd sourcing based approach and II.) during dedicated measuring studies targeted at particular network metrics. The measurements were mostly executed in and around Darmstadt, Germany, representing a medium sized urban center.

To assure the quality of the measurements, the data was thoroughly filtered. Data points with invalid fields were removed from the data set as well as measurements with a velocity of more than 15 m/s (i.e., 4.6 km/h ~ walking speed). To eliminate effects of different network structures in the backbone, only the data of one large network provider was selected.

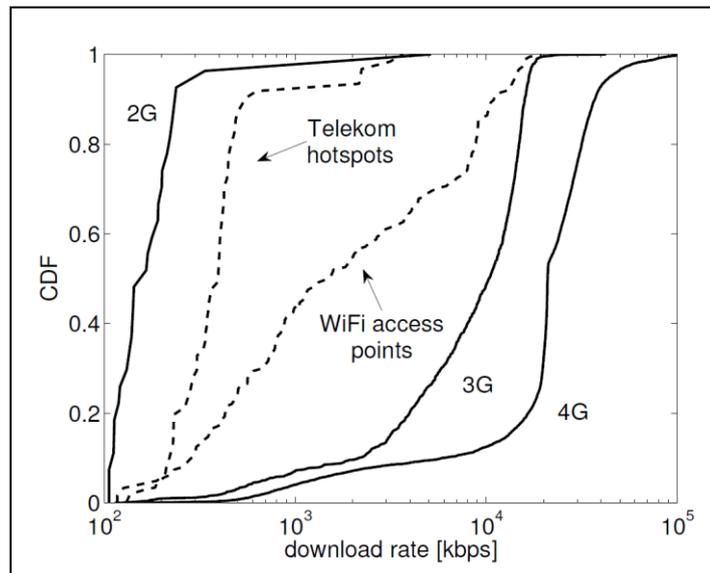


Figure 4-54: Throughput of mobile access technologies.

Figure 4-54 shows the cumulative distribution function of the download rate for various access technologies on logarithmic scale. The 2G connections have a maximum download rate of 5100 kbps, but 90% of the connections achieve a throughput of less than 230 kbps. The 3G connections show a maximum throughput of 42000 kbps, with 40% of the connections approaching this maximum value. Only 15% of the 3G connections have a lower download rate than 3000 kbps. 80% of the 4G connections show a higher throughput than the maximum speed of 3G, ranging up to the maximum download rate of 117000 kbps.

In case of WiFi access, we distinguish between hotspots by a major German provider (i.e., Deutsche Telekom) and other WiFi access points. It can be seen that the maximum speed of the Telekom hotspots is comparable to the maximum of 2G, although they usually offer a higher download rate than 2G. The maximum throughput of the other WiFi access points comes close to the 3G maximum, and the throughputs are usually higher than the Telekom hotspots but lower than the throughputs of 3G connections.

Based on these measurements, the power consumption of multipath TCP (MPTCP) is analyzed. MPTCP allows using multiple interfaces in parallel

with the goal of increasing the QoE for the end-user. Still, the power consumption, in particular for constant bitrate streaming, is not well analyzed. From the above measurements a general decision tree for energy-efficient mobile network access can be derived. This is given in Figure 4-54.

Based on the measurements, also energy efficiency improvements compared to the additive power consumption of both were determined. The Nexus S shows no synergies, while the power consumption on the Nexus 5 was 20% lower, reducing the absolute cost of using MPTCP. Details on this study can be found in Appendix D.

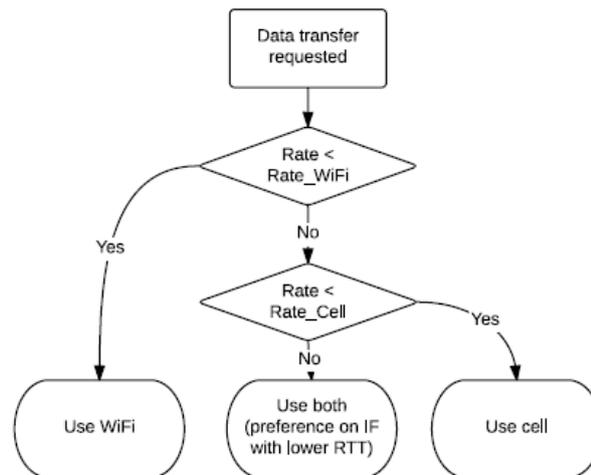


Figure 4-55: Decision tree for the most energy efficient mobile network access.

Innovation	Currently, WiFi offloading does not consider the quality of the underlying network technologies when switching between different interfaces. Neither is the energy consumption of the mobile device considered when selecting access technologies. By including both in the decision making process, the optimal trade-off between energy consumption, monetary cost and QoE can be adjusted based on user preferences and the type of network request (e.g. live streaming, web browsing, downloads, etc.). Such, the user experience can be improved while saving energy and/or monetary expenses.
-------------------	--

4.2.5 RB-HORST ++

4.2.5.1 RB-HORST + SEConD

Use case name	Service and content placement for users
Scenario	End-user Focused Scenario
Goal	The goal of RB-HORST + SEConD is to leverage unutilized resources in home user equipment like router or set top boxes for content delivery while keeping the QoE high in video sessions. If the aggregated upload bandwidth provided by the contributing uNaDa's is not sufficient to maintain a good QoE for the end-user in a video session, then the Socially-aware Proxy Server (SPS) participates in the content delivery. The trade-off between QoE and

	<p>SPS contribution is evaluated by studying the support threshold, which decides on the participation of the SPS.</p>
<p>Figure</p>	<pre> sequenceDiagram participant SPS participant uNaDa participant End-user Note over End-user: User requests a content and joins the swarm for this content_id Note over End-user: QoS monitoring Note over End-user: low_QoE End-user->>uNaDa: Request_SPS_participation uNaDa->>SPS: SPS_participation_response Note over SPS: Join swarm SPS->>uNaDa: Note over End-user: high_QoE End-user->>uNaDa: Disengage_SPS uNaDa->>SPS: Leave swarm of content providers Note over SPS: Leave swarm SPS->>uNaDa: </pre> <p>Figure 4-56: RB-HORST+SEConD communication.</p>
<p>Parameters</p>	<ul style="list-style-type: none"> • UNaDas upload bandwidth • Video bit rate • Support threshold for ISP cache contribution • Available resources when a UNaDa requests a video
<p>Metrics</p>	<ul style="list-style-type: none"> • Share requests served by the SPS / ISP cache contribution • Inter-AS traffic • Reduction of DC contribution • End-user QoS
<p>Traffic Management</p>	<p>RB-HORST+ SEConD</p>

<p>Solutions</p>	<p><u><i>RB-HORST+ SEConD algorithm for SPS participation</i></u></p> <ul style="list-style-type: none"> • The SPS should maintain a table with participation counters. Each counter belongs to a specific content that is stored in the SPS cache. • When this counter is greater than zero the SPS participates. • This counter is initially zero is incremented each time a UNaDa requests support and decremented when a UNaDa disengages the SPS. • During streaming, the UNaDa monitors the downloading rate and the video's bit rate: <ul style="list-style-type: none"> ○ If the downloading rate is lower than the support threshold, then the respective counter is increased. ○ If the downloading rate is higher than the support threshold, then the respective counter is decreased. 																												
<p>Evaluation framework</p>	<p>Content Delivery Simulation Framework</p> <p>The content delivery simulation framework, simulates a tiered caching architecture to evaluate the performance of the RB-HORST + SEConD mechanism. Therefore we model the home router upload bandwidth with a normal distribution with mean μ and standard deviation σ. We define the support threshold θ, which determines the SPS participation. We refer to the SPS as ISP cache in the following. To assess the QoE of end users of a video session, we implement the QoE model for mobile video streaming defined in the Appendix 11.1.</p> <p>As parameters we consider the mean home router upload bandwidth μ and the support threshold θ.</p> <p>We consider the ISP cache / home router contribution as a performance metrics to evaluate how much load the mechanism takes of the ISP cache and the inter-domain traffic saved by the mechanism. The quality perceived by end users is evaluated by the amount of video sessions having a good QoE.</p>																												
<p>Evaluation results</p>	<div style="text-align: center;"> <table border="1"> <caption>Data for Figure 4-57: Contribution to content delivery dependent on mean home router upload bandwidth for support threshold $\theta = 500$ kbps.</caption> <thead> <tr> <th>mean home router upload bandwidth [kbps]</th> <th>home router contribution</th> <th>ISP cache contribution</th> <th>content provider contribution</th> </tr> </thead> <tbody> <tr> <td>200</td> <td>~0.05</td> <td>~0.75</td> <td>~0.20</td> </tr> <tr> <td>400</td> <td>~0.48</td> <td>~0.32</td> <td>~0.20</td> </tr> <tr> <td>600</td> <td>~0.65</td> <td>~0.15</td> <td>~0.20</td> </tr> <tr> <td>800</td> <td>~0.70</td> <td>~0.10</td> <td>~0.20</td> </tr> <tr> <td>1000</td> <td>~0.73</td> <td>~0.07</td> <td>~0.20</td> </tr> <tr> <td>unlimited</td> <td>~0.78</td> <td>~0.02</td> <td>~0.20</td> </tr> </tbody> </table> </div> <p>Figure 4-57: Contribution to content delivery dependent on mean home router upload bandwidth for support threshold $\theta = 500$ kbps.</p>	mean home router upload bandwidth [kbps]	home router contribution	ISP cache contribution	content provider contribution	200	~0.05	~0.75	~0.20	400	~0.48	~0.32	~0.20	600	~0.65	~0.15	~0.20	800	~0.70	~0.10	~0.20	1000	~0.73	~0.07	~0.20	unlimited	~0.78	~0.02	~0.20
mean home router upload bandwidth [kbps]	home router contribution	ISP cache contribution	content provider contribution																										
200	~0.05	~0.75	~0.20																										
400	~0.48	~0.32	~0.20																										
600	~0.65	~0.15	~0.20																										
800	~0.70	~0.10	~0.20																										
1000	~0.73	~0.07	~0.20																										
unlimited	~0.78	~0.02	~0.20																										

ISP cache contribution highly depends on home router upload bandwidth. At this point we consider constant content size and home router cache capacity. The impact of content size will be investigated in a later stage.

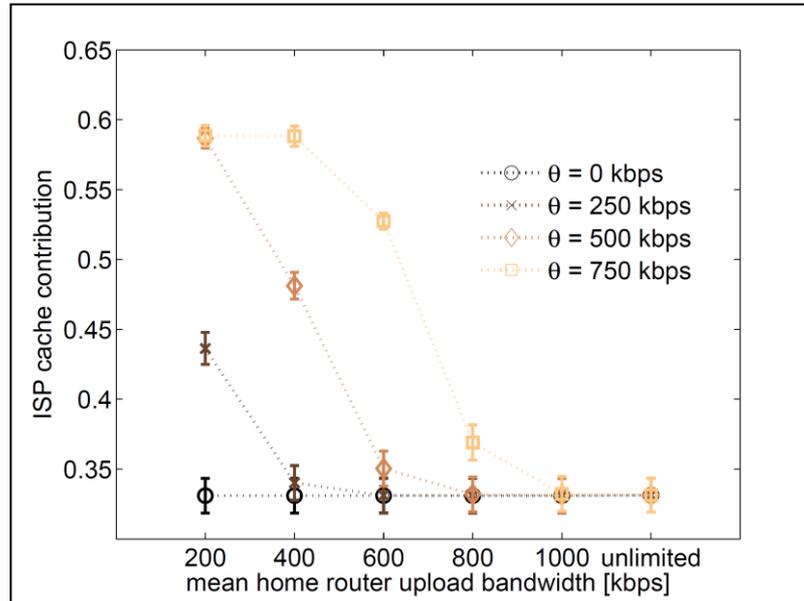


Figure 4-58: ISP cache contribution dependent on mean home router upload bandwidth for different support thresholds θ .

The ISP cache contribution depends on support threshold θ if the home router upload bandwidth is low.

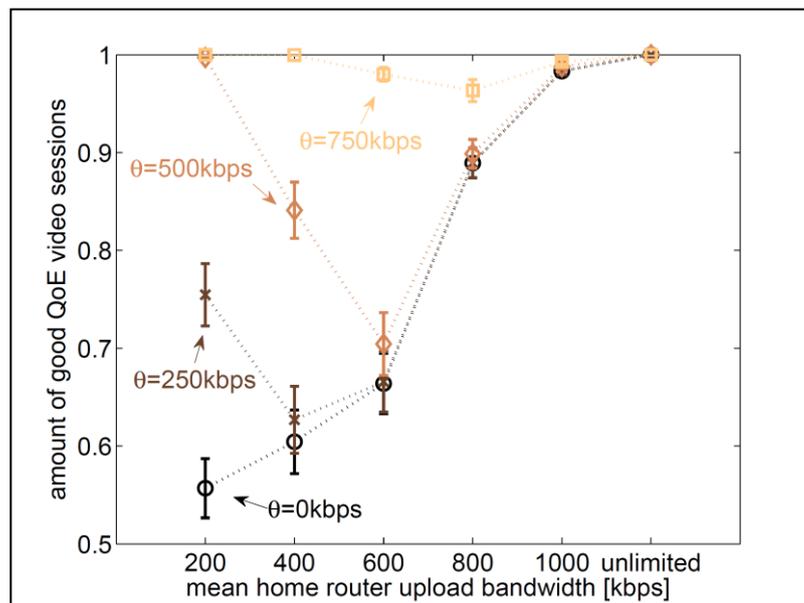


Figure 4-59: Amount of good QoE video sessions dependent on mean home router upload bandwidth for different support threshold θ .

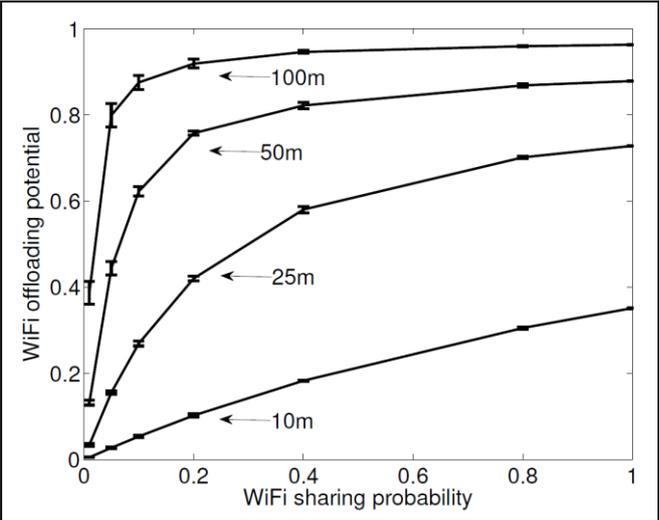
A low support threshold results in low QoE, except in the case of unlimited home router upload bandwidth. The QoE of video sessions increases with the support threshold θ . If the home router upload bandwidth is low there is a

	tradeoff between good QoE of video sessions and the ISP cache contribution. If home routers have a mean upload bandwidth of at least 800 kbps 9 of 10 video sessions receive a good QoE independent of the support threshold. In this case only one third of the requests are served by the ISP cache.
Innovation	RB-HORST + SEConD allows leveraging unutilized resources for content delivery while keeping the QoE high in video sessions. This is achieved by exploiting the upload bandwidth of local caches, which is to the best of our knowledge not the case in other existing approaches in the literature. The trade-off between QoE and ISP cache contribution can be set by a support threshold.

4.2.5.2 RB-HORST + MONA

Use case name	Social-Aware Mobile Data Offloading
Scenario	End-user Focused Scenario
Goal	The RB-HORST + MONA mechanism covers the EFS use cases: Social-Aware Mobile Data Offloading, and Access Technology Selection for Users. The mechanism combines the increased WiFi coverage provided by trusted RB-HORST access points with the offloading functionality of MONA. The combined mechanism aims to offload as much traffic as possible to shared WiFi access points, while maintaining a good QoE, e.g. for video sessions. In the following we use the QoE model described in section 11.1 to evaluate the WiFi offloading potential depending on the WiFi sharing probability and the impact of WiFi offloading on the QoE perceived by end-users in video sessions.
Figure	<p>Figure 4-60: End-users offloading mobile video sessions to shared WiFi networks of RB-HORST enabled home routers.</p>
Parameters	<ul style="list-style-type: none"> • Transmission range r • Sharing probability p

<p>Metrics</p>	<ul style="list-style-type: none"> • Fraction of offloaded requests • Fraction of good QoE sessions
<p>Traffic Management Solutions</p>	<p>RB-HORST + MONA</p>
<p>Evaluation framework</p>	<p>Data Sets</p> <p>For the evaluation of RB-HORST + MONA on different data sets are used to model the bandwidth of mobile access links for different technologies, the distribution of WiFi access points in a city area, and the location of end-users. The data sets are described in the following.</p> <p>The network performance data set is provided by the measurements conducted for evaluation of the MONA mechanism, c.f. section 4.2.4. The data set provides the throughput for mobile connections on street level for 2G, 3G, 4G and WiFi access technology.</p> <p>Access Point Location Data Set</p> <p>The WiFi access point location data set was measured by Panitzek et al. and is described in [16]. The data set consists of 1527 AP locations in an area of approximately 1.5km², covering the inner city of Darmstadt, Germany. From this data set, we use the interpolated locations of the access points as derived from the observed WiFi beacons at street level.</p> <p>OpenStreetMaps Data Set</p> <p>The probability of end-users to be at a specific location in the Darmstadt city area is derived by a street map of Darmstadt from OpenStreetMap [17]. The street map contains way points that are interconnected to define streets, or that describe buildings, facilities, local businesses or sights. The way points are all set up by users contributing to the OpenStreetMap platform. In that way, the way point locations provide a good model for end-user location probabilities.</p> <p>Simulation Model</p> <p>In the simulation we consider an area with a set of way points W and a set of access points A. The location of the way points and access points is specified by longitude and latitude. Each access point $\alpha \in A$ has a fixed transmission range r and is shared with probability p.</p> <p>For a given transmission range r we define a function $\chi_r: A \times W \rightarrow \{0,1\}$, where χ_r returns 1, only if a way point $w \in W$ is in transmission range of an access point $a \in A$, else 0.</p> <p>As set of way points W we use the way points from OpenStreetMap in the inner city area of Darmstadt. As set of access points A we use the access point location data set.</p> <p>The procedure of one run simulating n mobile requests is described in the following. A subset $A_s \subset A$ of shared access points is randomly chosen according to the sharing probability p. For each mobile request $1 \leq i \leq n$ a random way point $w_i \in W$ is determined. The mobile request i can be offloaded, if a shared WiFi access point is in range, i.e. $\exists a \in A_s \mid \chi_r(w_i, a) = 1$. The WiFi offloading potential</p>

	<p>is then calculated by the amount of offloaded requests.</p> <p>If the mobile request can be offloaded, WiFi is used as access technology. If the request cannot be offloaded the request is served by the cellular network which uses 2G, 3G or 4G access technology. The throughput ρ_i received for request i is determined randomly according to the access technology and its cumulative distribution function derived from the network performance data set.</p> <p>Finally the bit rate b_i of the requested video is determined randomly according to the encoding rate of itag36 videos.</p> <p>We determine if request i received a good QoE, if $\rho_i \geq 2 \cdot b_i$.</p> <p>The amount of good QoE sessions is determined by the number of requests that received a good QoE.</p>
<p>Evaluation results</p>	<p>In the following we describe simulation results to show the WiFi offloading potential in an urban environment dependent on the WiFi sharing probability. We further show the QoE benefits and degradations of WiFi offloading for different mobile access technologies. The results can be used by operators to assess the feasibility of establishing WiFi offloading according to their cellular network coverage, or to estimate the amount of users that share their access point, which is necessary to get a good WiFi coverage.</p> <p>The results show mean values with 95% confidence intervals of 10 runs with different random number seeds and $n=100000$ mobile requests in each run.</p> <p>We investigate the impact of the WiFi sharing probability on the WiFi offloading potential. As the transmission range of WiFi access points depends on the environment, the number of active connections and its configuration, we show results for different transmission ranges. Figure 4-8 shows the amount of mobile connections offloaded to WiFi dependent on the WiFi sharing probability p. The WiFi offloading potential is depicted for different transmission ranges r.</p>  <p>Figure 4-61: Amount of mobile connections offloaded to WiFi dependent on sharing probability for different access point transmission ranges.</p> <p>If a transmission range of only 10m is assumed, the WiFi sharing potential is rather low and increases almost linearly with the WiFi sharing probability. Roughly every second mobile connection can be offloaded for a transmission range of 25</p>

meters if 40% of the access points are shared and 3 of 4 connections can be offloaded if every access point is shared. If a WiFi transmission range of 50m is assumed, a decent WiFi offloading potential is obtained if only 10% percent of WiFi access points in an inner city area are shared. Hence, to obtain a good WiFi coverage, incentive mechanisms have to be designed, such that at least 10% of WiFi access points are shared.

In the following we set the WiFi transmission range to 50m and investigate the impact of WiFi offloading on video streaming QoE.

Figure 4-61 shows the amount of good QoE video sessions dependent on the WiFi sharing probability p for a WiFi transmission range r of 50m. The QoE of a video session is considered to be good, if the received throughput is at least twice the video bitrate. The dashed line depicts the probability that a video session receives a good QoE ifz WiFi is used in any case, which is about 80%.

The amount of good QoE sessions is depicted for three alternative access technologies, which are used if the connection cannot be offloaded to WiFi. If the alternative to WiFi is 2G, the number of sessions with good QoE increases with the WiFi sharing probability. This depends on the fact that the throughput of the WiFi connection is higher than the throughput of 2G in most cases. If more WiFi access points are shared, the probability to offload the connection and to receive more bandwidth increases. With a higher throughput the amount of good QoE sessions also increases.

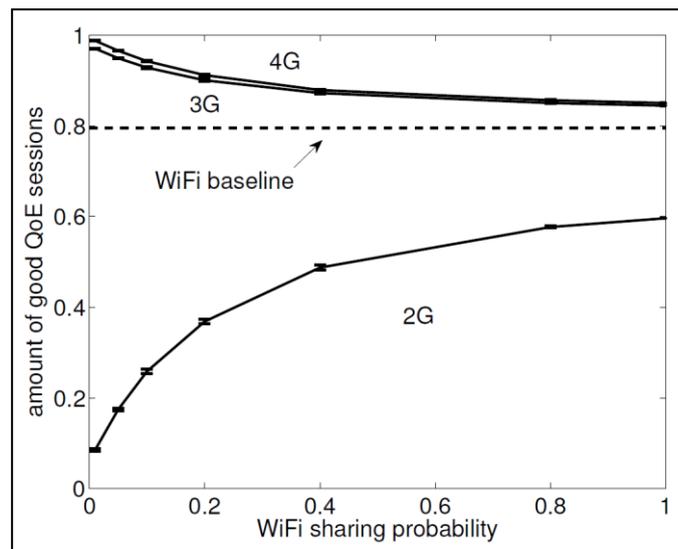
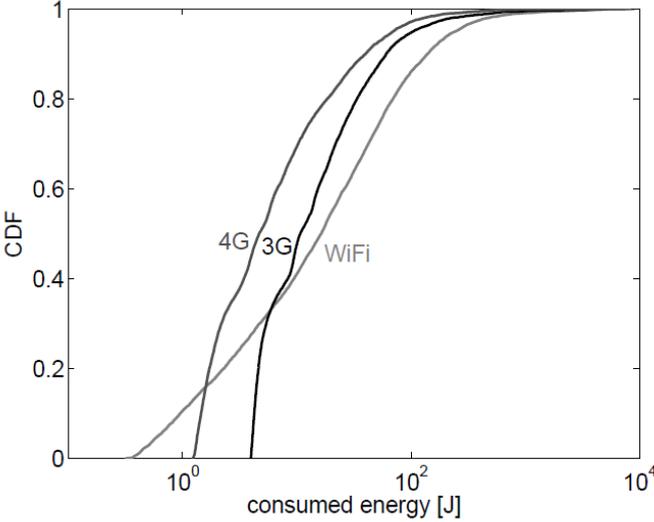


Figure 4-62: Probability to perceive good QoE during video session.

In our measurements the throughput of WiFi is lower than the throughput of 3G and 4G connections with high probability. The amount of good QoE sessions decreases with the WiFi sharing probability.

3G and 4G meet the requirements of mobile video streaming. Hence, if no WiFi is shared and every video session has to be streamed over 4G, the QoE is good in 100% of the sessions. The amount of good QoE sessions in 3G is slightly lower than in 4G, but is still close to 100%.

In the worst case from QoE perspective, hence if 4G is available and the WiFi sharing probability is 100% and mobile connections are offloaded if possible, still, in more than 86% of video sessions good QoE is perceived. In this case the load

	<p>on the cellular network is mitigated to only 12%.</p>  <p>Figure 4-63: Energy consumptions for different access technologies.</p> <p>A burst traffic model for video on demand was applied to estimate the energy consumed by mobile devices based on the energy model provided in section 4.2.4. Figure 4-63 shows the energy consumption for WiFi, 3G and 4G. The minimum energy consumption is achieved using WiFi access technology. However, WiFi consumes more energy than 3G in more than 60% of the requests and WiFi consumes more energy than 4G in more than 80% of the requests. This depends on the fact that the energy consumption decreases exponentially with the data rates and that the data rates obtained for WiFi in the measurements are very poor compared to 3G and 4G. Hence, although WiFi is more energy efficient than 3G and 4G for equal data rates it consumes more energy in this case due to lower data rates.</p>
<p>Innovation</p>	<p>RB-HORST + MONA mitigate the load on cellular networks while maintaining good QoE in video sessions. This is achieved by leveraging WiFi access points shared via RB-HORST. The energy consumption of video transmissions can be reduced by using WiFi access if high data rate is achieved.</p>

4.2.6 MUCAPS

<p>Use case name</p>	<p><i>Service and content placement selection for users</i></p>
<p>Scenario</p>	<p>End-user focused</p>
<p>Goal</p>	<p>Performance evaluation of the MUCAPS in-network TM prototype is illustrated in Figure 4-64. The set-up comprises:</p> <ul style="list-style-type: none"> • User Endpoint (UEP) receiving a video flow, • video servers AEP1, AEP2, AEP3, ... AEPk, supporting at least 3 video bit-rates (4, 1 and 0.640 Mbits/sec), connected to a UEP with different e2e path bit-rates (8, 1.6, 0.8 and 0.4 Mbits/sec) depending on path bandwidth and connection types, • the MUCAPS components: the MARC (MACAO Request handling Client) hooked to a DNS Server, the MACAO (Multi-Access and Cost AltO) decision block with its

	<p>MARS (MACAO Request handling Server) interface, an ALTO Server, all in the same ISP network than the UEP,</p> <ul style="list-style-type: none"> The QoE measurement and analysis components: VITALU tool (Video Inspector Tool of Alcatel LUcent) processing PCAP files captured at the UEP level. <p>The evaluation of the prototype answers the following questions by the following means:</p> <ul style="list-style-type: none"> What is the gain in terms of ISP-defined routing costs and ISP-defined path BW availability score? Assessed by analytical evaluation as described in D2.4. What is the impact on user QoE? measured and analyzed by VITALU. How are user conditions (access, device) considered by MUCAPS? Assessed by monitoring AEP changes in AEP; this study will be completed in WP4.
<p>Figure</p>	<p>The diagram illustrates the MUCAPS base prototype architecture. At the top, three Video Servers are shown with IP addresses 10.133.10.1, 10.133.10.2, and 10.133.10.53, each labeled with AEP1, AEP2, and AEP3 respectively. These servers are connected to a central network. Below them, a DNS Server (172.25.70.103) and a MACAO Block (172.25.70.153) are connected. To the right, an ALTO Server (172.25.70.153) is also connected. An End-User UEP (represented by a laptop) is connected to the network. A QoE Analyzer Vitalu is connected to the UEP, and a Network Capture component is shown capturing data from the UEP in pcap format. A Results Metrics box is connected to the Vitalu analyzer. A central orange box labeled 'Prototype: integration of VITALU functions' encompasses the DNS Server, MACAO Block, and ALTO Server. A vertical 'MARS' label is positioned between the DNS Server and MACAO Block.</p> <p>Figure 4-64: MUCAPS base prototype with QoE measurement and analysis.</p>
<p>Parameters</p>	<ul style="list-style-type: none"> $RC = \text{ISP defined Routing Cost} = \text{unitless metric}$ $BWscore = \text{unitless ISP score on e2e path BW availability with values in:}$ <ul style="list-style-type: none"> $[0,20]$ if pure score $[0, 10MBpF]$ MBpF with MBpF = mean MB allocation per flow, if to reflect real path BW UEP (User End Point) access type
<p>Metrics</p>	<ul style="list-style-type: none"> Cross Layer Utility w.r.t. ISP defined routing costs and path BW score VQM = video quality score Start Time delay = date of media play – date of media request NFrz, DFrz: Number of freezes, Duration of freezes (statistics)

	<ul style="list-style-type: none"> • Media Bit Rate: relative value wrt device conditions • % of corrupted frames when applicable transport protocol 																																																
<p>Traffic Management Solutions</p>	<p>MUCAPS</p>																																																
<p>Evaluation framework</p>	<p>MUCAPS has been implemented and demonstrated as a prototype. The prototype assessment has used an analytical and a measurement based evaluation framework. In both frameworks results are compared in 3 MUCAPS cases: direct selection (MUCAPS <i>off</i>), MUCAPS <i>on</i> and using Routing Costs (RC) only, MUCAPS <i>on</i> and using RC and Bandwidth Score (BWS). The MUCAPS evaluation framework is detailed in D2.4 [7]. The analytical evaluation framework: computes the Utility of the selected AEPs w.r.t. the ISP defined BWS and RC. It compares the utilities obtained in 3 MUCAPS cases. D2.4 shows preliminary results for UEPs connected via a LAN. In this section we complete this assessment with WiFi and Cellular connections.</p> <p>The measurement-based framework: performed 100 of video session measurements and analysis, for the 4 classes of path bit-rate and 3 classes of video bit-rate listed above. Measurements were done on devices with high, standard and low definition screens (HD, SD, LD). 5 videos were tested.</p> <p>The measurements and analysis results are provided in terms of the VQS metric. VQS stands for Video Quality Score and is obtained as follows:</p> <p>The measurements have been statistically aggregated over the Videos and the number of measurements. The VQS have been differentiated w.r.t. the screen quality to illustrate their impact.</p>																																																
<p>Evaluation results</p>	<p>1 - Analytical evaluation</p> <p>Table 4-5, Table 4-6 and Table 4-7 recap the ALTO values and metric weights for the 3 classes of AEPs: AEP1: moderate cost + good bandwidth, AEP2: low cost + low bandwidth, AEP3: high cost + high bandwidth). They also provide their utilities computed by the Multi-Criteria EP Evaluation function, with the method described in D2.4, basically the weighed proximity to the ideal vector, which is (7, 20) in this case. L1-Utility is the weighted sum of the normalized utility vector components. A table is given for each UEP access type, for which utility vectors values vary, while the ALTO values remain unchanged.</p> <p>Table 4-5: Utility value of the 3 AEP classes for UEPs connected via LAN</p> <table border="1" data-bbox="355 1550 1430 1771"> <thead> <tr> <th></th> <th>AEP1</th> <th>AEP2</th> <th>AEP3</th> <th>Metric weight</th> <th>Best EP = MAX Utility</th> </tr> </thead> <tbody> <tr> <td>ALTO RC</td> <td>11</td> <td>7</td> <td>60</td> <td>1</td> <td></td> </tr> <tr> <td>ALTO BWS</td> <td>18</td> <td>8</td> <td>20</td> <td>1</td> <td></td> </tr> <tr> <td>Utility vector</td> <td>(0.63, 0.9)</td> <td>(1, 0.4)</td> <td>(0.116, 1)</td> <td></td> <td></td> </tr> <tr> <td>L1-utility</td> <td>0.765</td> <td>0.7</td> <td>0.558</td> <td></td> <td>AEP1</td> </tr> </tbody> </table> <p>Table 4-6 : Utility value of the 3 AEP classes for UEPs connected via WiFi</p> <table border="1" data-bbox="355 1863 1430 1986"> <thead> <tr> <th></th> <th>AEP1</th> <th>AEP2</th> <th>AEP3</th> <th>Metric weight</th> <th>Best EP = MAX Utility</th> </tr> </thead> <tbody> <tr> <td>ALTO RC</td> <td>11</td> <td>7</td> <td>60</td> <td>2.3</td> <td></td> </tr> <tr> <td>ALTO BWS</td> <td>18</td> <td>8</td> <td>20</td> <td>4</td> <td></td> </tr> </tbody> </table>		AEP1	AEP2	AEP3	Metric weight	Best EP = MAX Utility	ALTO RC	11	7	60	1		ALTO BWS	18	8	20	1		Utility vector	(0.63, 0.9)	(1, 0.4)	(0.116, 1)			L1-utility	0.765	0.7	0.558		AEP1		AEP1	AEP2	AEP3	Metric weight	Best EP = MAX Utility	ALTO RC	11	7	60	2.3		ALTO BWS	18	8	20	4	
	AEP1	AEP2	AEP3	Metric weight	Best EP = MAX Utility																																												
ALTO RC	11	7	60	1																																													
ALTO BWS	18	8	20	1																																													
Utility vector	(0.63, 0.9)	(1, 0.4)	(0.116, 1)																																														
L1-utility	0.765	0.7	0.558		AEP1																																												
	AEP1	AEP2	AEP3	Metric weight	Best EP = MAX Utility																																												
ALTO RC	11	7	60	2.3																																													
ALTO BWS	18	8	20	4																																													

weighted Utility vector	(1.45, 3.6)	(2.3, 1.6)	(0.266, 4)		
L1-utility	0.789	0.609	0.666		AEP1

Table 4-7: Utility value of the 3 AEP classes for UEPs connected via Cellular

	AEP1	AEP2	AEP3	Metric weight	Best EP = MAX Utility
ALTO RC	11	7	60	3	
ALTO BWS	28	8	20	2	
weighted Utility vector	(1.89, 1.8)	(3, 0.8)	(0.348, 2)		
L1-utility	0.738	0.76	0.469		AEP2

Next, we comment on the weight selection among RC and BWS. In LAN, RC and BWS are equally important. In WiFi, the cost of BWS is more important than RS, as many users may have low performance otherwise. In cellular access, there is no need of high bandwidth (e.g. 4 to 8 megabytes) if access connection and devices do not support them, so the cost of RS is more important than the cost of BWS. The calculation of the weights is generally decided by the operator and depends on the estimated average bit-rate necessary for each connection.

Table 1-4 summarizes the ordering and selection of AEPs in all 3 MUCAPS activation cases, for all three types of access. The symbol “*”, means that any other AEP may be selected in this position of the ordered list.

Table 4-8 : AEP ordering w.r.t. MUCAPS status and UEP access type

MUCAPS	LAN	WiFi	Cellular
OFF	AEP3, *, *	AEP3, *, *	AEP3, *, *
RC	AEP2, AEP1, AEP3	AEP2, AEP1, AEP3	AEP2, AEP1, AEP3
(RC, BW)	AEP1, AEP2, AEP3	AEP1, AEP3, AEP2	AEP2, AEP1, AEP3

Note that with MUCAPS, the application does not systematically skip the high ISP cost server. Although the Best EP selection usually discards the highest cost AEP, one may note that for the (RC, BW) case, for a UEP connected via WiFi, MUCAPS prefers as a second choice AEP3, a high cost Server ensuring a high bandwidth, given the importance of BW availability in WiFi access. For a cellular connection, due to the lower needs for bandwidth, the BWScore weight is less important and MUCAPS prefers AEP2.

2 – analysis of the MUCAPS impact on QoE

The Table 4-9 and Table 4-10 below provide the average Video Quality Scores (VQS) obtained for the 3 types of UEP access and with 4 classes of path bandwidth associated to the BW Score.

The assumed video bit-rate is equal to 4 Mbits/sec for LAN and 1 Mbit/s for WiFi access. For Cellular connections it is assumed to be 640 Kbits/sec. The value range of a VQS is [1, 5].

Table 4-9: Average VQS when Optimal resolution requested

	AEP1	AEP2	AEP3

Target Equipment		Network Bitrate (kbps)	VQS	Network Bitrate (kbps)	VQS	Network Bitrate (kbps)	VQS
Mobile Phone (LD)	LAN	8000	4.5	1600	4.5	8000	4.5
	Wi-Fi	1600	4.5	800	4.5	1600	4.5
	Cellular	800	4.5	800	4.5	800	4.5
Tablet (SD)	LAN	8000	4.5	1600	4.5	8000	4.5
	Wi-Fi	1600	4.5	800	2.1	1600	4.5
	Cellular	800	2.1	800	2.1	800	2.1
TV (HD)	LAN	8000	4.5	1600	2.0	8000	4.5
	Wi-Fi	1600	2.0	800	1.4	1600	2.0
	Cellular	800	1.4	800	1.4	800	1.4

These evaluations were obtained under the hypothesis that each target equipment requests the video with the optimal resolution. The results show without surprise that when the underlying network bit-rate is lower than the video bit-rate for the selected video, the quality of experience will drop below a comfortable level because of freezes events.

It can be seen that selecting the AEP2 results in lower network bit-rates and hence in lower quality of experience for the user. In particular it is impossible in the case of the HD-TV equipment to reach an acceptable quality when AEP2 is selected. As stated earlier, the use of AEP2 should be limited to users using a cellular network. Table 4 showed that when MUCAPS operates considering only RC, AEP2 is selected in priority even on LAN and Wi-Fi. This means that the choice should be made considering RC as well as BW in order to be able to obtain a good QoE.

Overall what can be seen from Table 1-4 and Table 1-5 is that the “nominal mode” of the solution always results in a very good QoE for the user. What we call the nominal mode is the use of equipment adapted to the network in use: a Mobile phone on a cellular network, a tablet on Wi-Fi or a HD-TV over a LAN, and connecting to the first AEP among the list of AEPs returned by MUCAPS. This is summarized in Table 4-10

Table 4-10: Results in nominal mode

Target Equipment		First Selected AEP	Network Bitrate (kbps)	VQS
Mobile Phone (LD)	Cellular	AEP2	800	4.5
Tablet (SD)	Wi-Fi	AEP1	1600	4.5
TV (HD)	LAN	AEP1	8000	4.5

As a final conclusion, it should be noted that even under the hardest hypothesis that the optimal resolution is requested for the considered equipment and network, MUCAPS has permitted to select an AEP matching the conditions and resulting in a good quality of experience for the user.

Table 4-11: Results with adapted resolution

Target Equipment		AEP1		AEP2		AEP3	
		Network Bitrate (kbps)	VQS	Network Bitrate (kbps)	VQS	Network Bitrate (kbps)	VQS
Mobile Phone (LD)	LAN	8000	4.5	1600	4.5	8000	4.5
	Wi-Fi	1600	4.5	800	4.5	1600	4.5
	Cellular	800	4.5	800	4.5	800	4.5
Tablet (SD)	LAN	8000	4.5	1600	4.5	8000	4.5
	Wi-Fi	1600	4.5	800	3.9 (LD)	1600	4.5
	Cellular	800	3.9 (LD)	800	3.9 (LD)	800	4.5
TV (HD)	LAN	8000	4.5	1600	4.1 (SD)	8000	4.5
	Wi-Fi	1600	4.1 (SD)	800	3.6 (LD)	1600	4.1 (SD)
	Cellular	800	3.6 (LD)	800	3.6 (LD)	800	3.6 (LD)

Table 4-11 shows the highest result for the video quality score (maximizing the quality of experience of the user) considering that the client requests a video resolution compatible with the actual network bit-rate. Under this hypothesis, even with a suboptimal video resolution, the overall quality of experience stays acceptable. When a suboptimal resolution is used, this is noted in the table after the VQS value. To be more explicit, if we consider the case of a HD-TV equipment over a cellular network, the choice presented in Table 4-11 is to display the HD video, but because of the slow network, a lot of freezes happen during the playout and the overall VQS is only 1.4 (a lot worse than the value of 4.5 when no freeze occurred). Here the player chooses to download the video encoded in LD resolution, then up-scales it to HD resolution and finally displays it on the HD-TV. As it needs a lower network bit-rate, this up-scaled video can be played smoothly without any freeze event and results in a VQS of 3.6, lower than the theoretical optimal of 4.5 for HD without freezes, but a lot better than the score of 1.4 for HD with freezes. This mechanism corresponds to what happens with adaptive streaming.

In this case as well, AEP2 shows a better adaptation to cellular networks.

Innovation SmartenIT investigates incentive-based mechanisms using a cross-layer approach for the efficient management of traffic generated by overlay applications. MUCAPS provides a mutual incentive for the application layer and the network layer to cooperate, by involving on one hand selection criteria on transport network costs such as routing costs that meet ISP interests, and on the other hand, criteria impacting QoE that meet end-users interests such as end to end path bandwidth. The QoE criteria used in MUCAPS abstract end-to-end performances on application paths. This way, applications take the ISP costs and constraints into consideration, provided that applications get reliable performance indicators from the ISPs. The selection made by MUCAPS appears to represent a good trade-off wrt costs and QoE and to be in-line with the available access technology in the UEP. MUCAPS uses an extension of the IETF ALTO protocol that is being already standardized, see [62].

4.3 Summary of theoretical models supporting TM mechanisms' evaluation

We present in Appendix 11 a set of theoretical models and simulation models developed along the TM mechanisms to characterize:

- energy consumption in mobile and wireless data transmission,
- content demand in online social networks,
- topological distribution of home routers and WiFi hotspots,
- caching effectiveness,
- QoE assessment for mobile video streaming, and finally
- new pricing models for SmartenIT system.

The major outcomes from these studies are summarized below:

The **energy efficiency** of the SmartenIT solutions is addressed by modeling the power consumption of mobile/wireless data transmissions. Multipath TCP (MPTCP) permits to transmit seamlessly data over multiple paths between different network technologies. The energy cost of constant bitrate streaming using MPTCP on smartphones is comparable to the added cost of using both interfaces simultaneously. The reduction in energy cost with simultaneous use of multiple interfaces, compared to individual interfaces depends on the device. Here, the cost of MPTCP is 20% lower than the theoretical cost of both interfaces individually, whereas the Nexus 5 shows power consumption proportional to the added cost of both interfaces. The most energy efficient configuration to use MPTCP on mobile devices is obtained when using the most energy efficient interface only. Only if the data rate cannot be supported by a single interface, both interfaces should be used. The most energy efficient setting in this configuration can be achieved by transferring most traffic on the interface with the lower RTT, which in our case is the WiFi interface. This stands in contrast with the current MPTCP implementation, which uses both interfaces with equal data rates. Hence, it is suggested to add an energy efficiency mode to the MPTCP implementation, considering these observations to conserve the available energy on the mobile device.

Performance of **caching mechanisms** is very essential to the overall effective performance of the TM solution as deployed in RB-HORST. Evaluation of basic caching mechanisms have been studied and their efficiency has shown

- that the performance gain of statistics based caching strategies compared to LRU (least recently used) is considerably high going up to twice the hit rate especially for small caches,
- that variants of LRU, i.e. the investigated score-gated SG-LRU scheme can fully exploit this gain at only slightly higher control overhead and even lower cache update effort per request,
- that the hit rate of the LFU (least frequently used) strategy determines the maximum caching efficiency not only in case of the independent request model (IRM) with

static content popularity, but the LFU limit is closely approached also for moderate dynamics in content popularity as typically observed for web caching.

In practice, LFU type strategies regarding popularity statistics are usual in advanced systems, where cache updates are often done once per day during low network load phases, but LRU caching is also applied as a simple first implementation option.

The performance of LRU caches in terms of cache hit rate can be analytically calculated by the Che approximation, which uses a fixed point iteration to determine the characteristic time an item resides in the cache before being replaced. Although our results confirm fairly high accuracy of the Che approximation for the LRU hit rate, we recommend taking the LFU hit rate as the basis for caching efficiency in techno-economic studies of cache infrastructures. Last not least, score-based cache strategies like SG-LRU are flexible to include not only the popularity but also other cost and benefit criteria into account for selecting the most relevant set of items in the cache.

Crucial for the performance evaluation of caching systems as deployed in RB-HORST is also the study of proper **models of content demand**. A model is defined to generate requests with lifespan and popularity that exhibits a Zipf law (includes flash crowds). This model is used in simulation to evaluate the efficiency of content delivery in RB-HORST with demand exhibiting temporal locality. Such a model permits to simulate and investigate situation of *flash crowd* generated by demands with high popularity and short lifespan. A second model is proposed for the prediction of individual content requests based on social network traces (Twitter dataset). Such a model is used in RB-HORST for improving the performance of the mechanism by social prediction of content request. The lack of demand model that considers social and temporal locality and results in Zipf-law prevents an accurate evaluation of the prediction performance. It has also been noted that Facebook content requests may be different from Twitter. Thus the methodology adopted in the model was kept, but adaptive (self-learning) prediction was implemented.

In order to estimate the potential of home router sharing approaches like RB-HORST, **global distribution of home routers on AS is modeled**. This model is used in simulative performance evaluation of RB-HORST. A model for geographic distribution of WiFi hotspots achieves the complete modeling of the topology characteristics in SmartenIT system in terms of deployed WiFi hotspots and home routers (UNaDas). In particular, it was observed that the distribution of home routers on AS is heterogeneous. In addition, the performance of mechanisms like RB-HORST highly depends on the size of AS, as, for instance, the total cache capacity increases with the number of available home routers. The model developed for geographic distribution of WiFi hotspots is simple by using only two parameters and approximates nicely different characteristics of different cities.

We have specified a generic **pricing framework**, defining the layers of charging involved in pricing cloud services, also in lieu of the SmartenIT business models presented in Deliverable 2.4 [7]. Furthermore and in-line with this pricing classification framework, an original SmartenIT **pricing model** related to the Operator Focused Scenario and in particular the ICC mechanism is proposed. The key feature of ICC is to incentivize the key ISP business customers that generate the time-shiftable traffic (e.g. some inter-cloud/-datacenter traffic) to allow ICC to manage the time-shiftable portion of their traffic through proper pricing. Hence, each customer will receive a discount proportional to his fair share

on the total volume of traffic managed by ICC. The pricing scheme associated with the ICC mechanism is a good compromise between simplicity and incentives, also respecting current ISP operational and business practices on the one hand and fairness and proper incentives on the other. The justification of the pricing model proposed also includes a detailed discussion of potential alternatives and their pros and cons, which can be applied to pricing cloud and network services at the bulk-data layer in general, thus its scope is wider than ICC.

4.4 Main outcome of TM mechanisms evaluation

The TM mechanisms and models developed to address the aspects of the OFS are DTM, ICC, MRA and cloud federation. They cover the use cases of bulk data transfer among federated clouds (i.e., DTM, ICC, DTM++ TM mechanisms, and the model for cloud federation), video transfer between storage of independent clouds (ICC,MRA), host resource allocation in cloud federations (MRA) and IoT data transfer for cloud operators (DTM++). In addition, we specified and evaluated DTM++ a synergetic solution combining the so-called traffic “shift in space” of DTM and traffic “shift in time” of ICC to achieve further optimization of traffic distribution across multiple transit links while delaying delay-tolerant traffic when transit links are heavily loaded, so as to ultimately achieve even lower transit charges for the ISP than DTM alone. The major outcomes and lesson learnt from all the studies conducted in WP2 for the evaluation of the **TM mechanisms for OFS** are summarized below.

DTM is a traffic management mechanism that minimizes the inter-domain traffic cost in multi-homed AS by influencing the distribution of the traffic among links. DTM works for tariffs based on traffic volume and 95th percentile. Simulation experiments show that DTM is able to compensate the traffic and distribute it among links as desired to optimize traffic cost for volume based tariff and for 95th percentile tariff. The achieved cost is still lower than that expected without DTM. In the 95th percentile tariff case, however the algorithm is sensitive to the traffic profiles. The traffic must be compensated on a very short time scale: 5 minutes. It is clearly more sensitive than in the case of volume based tariff (in which the traffic is averaged over the whole billing period).The method for the calculation of the reference vector for the 95th percentile tariff needs to be modified. The estimated amount of manageable traffic used for \vec{s} vector (freedom for looking for optimal traffic distribution w.r.t cost) cannot be calculated as an average of the amount of manageable traffic observed in 5-min samples (mean over all 5-min samples).. An improved algorithm should be defined. It should take into account amount of both types of traffic during daily peak periods. Furthermore, considering the aspect of DTM scalability, usage of multicast instead of unicast will scale the communication of compensation and reference announcement to many DC/clouds, while using time slot scheduling for compensation vector delivery from different S-Boxes served by multiple SDN controllers sharing load. Security and reliability can be ensured through the usage of VPN tunnels or HTTPS and redundancy of S-Boxes and SDN controllers respectively.

ICC is a traffic management mechanism deployable by an ISP for utilizing its transit link(s) optimally in terms of cost and performance. ICC attains a reduction in the ISP’s 95th percentile transit charge by means of controlling the rate of a portion of this traffic (e.g.

delay-tolerant cloud traffic) that has been marked accordingly by the ISP's business customer (e.g. cloud/datacenter), and shifting its transmission at off-peak intervals, while real-time traffic is still forwarded a la Best Effort. Quantitative results show performance gains of ICC under both full information and partial information (traffic expectation from statistics). ICC always attains the transit charge reduction goal under full information regarding traffic. ICC using statistics to predict traffic patterns also performs well (max +/- 10% deviation from target goal), often achieving higher discount than the one original sought (impact of epoch *tholds[]* parameter). ICC is always better than current status quo in terms of 95th percentile and traffic shaping. It gives often impressive gains if there is significant volume of manageable traffic and the average rate within the billing period is much less than the max (5%) rate values that determine the transit link charge. Performance depends on the link traffic patterns: the traffic variability and periodicity over time, the percentage of manageable traffic, and the difference of avg and max (5%) rates. Fine-tuning ICC requires a trial and error process for optimal values (for *Ctarget*, *tholds[]*), but sub-optimal values also meet the goal of attaining significant cost savings. Achieving a slightly higher or lower discount than the one set is more than fine since the difference of the 95th percentile attained with and without ICC is always significant for typical ISP transit links, where transit traffic patterns inherently exhibit significant rate variations over time.

MRA is a mechanism that helps in allocating multiple heterogeneous resources, such as CPU, RAM, disk space, and bandwidth, in a fair manner between customers of a cloud or a cloud federation. Fairness takes the greediness of customers into account to reallocate the overloaded resource, i.e., VMs of customers who are rated as "moderate" (i.e., have a lower greediness) receive more of the scarce resource. Fairness in a cloud (federation), efficiency of VM scheduling and live migration can only be evaluated significantly, when the interdependency of multiple heterogeneous resources (CPU, RAM, etc.) is considered. Such interdependencies are not straightforward to determine and usually not considered in the literature. As the investigations of resource dependencies revealed, there is no clear correlation between resources usage. Although a Leontief relationship for CPU time, disk I/O, and bandwidth could be confirmed, there are certainly workloads that are more flexible with respect to consuming these resources. For example a VM could receive requests (consume bandwidth), process these afterwards, when CPU is available (consume CPU) and only then write the results to disks, when the disk is not congested (consume disk I/O). In the meantime data could be stored in RAM. Investigating such more flexible cases is left for future work. Also, it should be investigated, why additional VCPUs cores decrease performance of multi-threaded workloads in many cases. While these findings are interesting, their main purpose was to prove that assumptions about resource dependencies in literature are too simplistic. These simplistic assumptions have in particular been made in fairness-research. Therefore, the results achieved prove many proposed fairness metrics as being insufficient and thereby certify the need for a new fairness metric and mechanism to ensure fairness in clouds. Such metric and an according mechanism has been proposed by SmartenIT as the greediness metric/MRA mechanism. The MRA mechanism is designed to ensure fair resource allocation in clouds and or cloud federations among customers based on multiple resources, which are share in a best effort manner and not prescribed by SLAs. **DTM++** is a synergetic solution integrating DTM with ICC. It combines the so-called traffic "shift in space" of DTM and traffic "shift in time" of ICC to achieve further optimization of traffic distribution across multiple transit links

while delaying delay-tolerant traffic when transit links are heavily loaded, so as to ultimately achieve even lower transit charges for the ISP than DTM alone. A completely new approach to realization of ICC functionality has been proposed relying on hierarchical policers. Several experiments have been performed to validate the implementation which proved that the selected solution supports ICC functionality as desired. Experiments prove that the DTM++ implementation (using filter based on hierarchical policer for the realization of ICC functionality) is working correctly. DTM++ may offer a better control of traffic sample sizes and what follows, also the cost of inter-domain the traffic. It has even a potential for further cost reduction.

Regarding **Cloud federation**, we provide an economic model of the federated environment of CSPs, considering both the case where CSPs act cooperatively and non-cooperatively. Cloud federation increases revenues from QoS-based pricing on customers and reduces the energy consumption cost, thus the profit of federated CSP is increased. Moreover Cloud federation improves the global QoS in the federated environment, i.e. the average delay of served requests is decreased. We investigate utility-based optimal federation formation policies. Furthermore, we take into account the QoS offered to CSPs' clients in their optimization approach. In our work, the federation policy is optimal with respect to total CSPs' profit, but it is also beneficial and caters for client utility, since a larger revenue for the CSP is mapped to better QoS for clients. The jobs transferred from one CSP to the other through the Internet, thus as the communication delay increases the benefit of federation is mitigated. In the cooperative federation, the optimal policy for jobs' outsourcing maximizes the total profit of the federation, but the global optimum may not be straightforward beneficial for both CSPs. Therefore, a fair profit sharing mechanism is also required. The benefit of federation seems to diminish when both CSPs have the same utilization ρ , where ρ denotes the ratio of the average arrival rate of requests to the average CSP service rate. In a federation of two CSPs and under the assumption that all the jobs have the same QoS-requirements, the service delegation is always unilateral, i.e. at most one of two CSP outsources request to the other. On the other hand. in the non-cooperative federation, the selfish behavior of CSPs do not guarantee that the individual profit of both CSPs will be higher than their profit in their standalone operation. Thus, there is the need for a certain pricing/compensation function that guarantees participation of both CSPs in the federation due to their mutual benefits

The TM mechanisms developed to address the aspects of the EFS are RB-HORST, vINCENT, MONA and MUCAPS. They cover the use cases of Service and content placement for users (RB-HORST, SEConD, vINCENT, MUCAPS), Exploiting Content Locality (RB-HORST), Social-Aware mobile data offloading (RB-HORST), and Access Technology Selection for Users (RB_HORST, vINCENT, MONA). Moreover, we specified two synergetic solutions named RB-HORST++ for the EFS: a) one combining trust-based home router sharing based on social observations by RB-HORST, and content caching, prefetching, and chunk-based dissemination by SEConD to address the service and content placement use case, and b) one combining trust-based home router sharing based on social observations by RB-HORST, and incentive-based reciprocation and energy efficiency by MONA, so as to address the social-aware mobile data offloading use case. The major outcomes and lesson learnt from the all studies conducted in WP2 for the evaluation of the **TM mechanisms for EFS** are summarized below.

RB-HORST evaluation shows that an overlay is essential for efficient content delivery in systems with a large number of resources with little capacity each. The overlay highly increases the share of requests served locally and reduces costly inter-domain traffic. It takes load off the ISP cache, which allows the provider to save costs for cache infrastructure and to save energy. The ISP cache can be dimensioned accordingly to save energy. Load reduction on ISP cache is dependent on home router sharing rate. Furthermore, RB-HORST provides WiFi access to trusted users to offload their mobile traffic from 3/4G to WiFi. The results on WiFi offloading in a urban environment show, that 66% of the connections can be offloaded on the average, if only 10% of WiFi access points are shared, assuming a sending range of 50m. LRU caching policy is simple but effective under non-stationary demand. Evaluation of prefetching accuracy is really challenging because of lack of ground truth. Still lack of demand model that considers both social and temporal locality and results in Zipf-law prevents accurate evaluation of RB-HORST performance. Inclusion of local demand in the model would help to evaluate home router cache efficiency

SEConD is a TM mechanism employing social-awareness, AS-locality awareness, chunk-based P2P content delivery and prefetching, and a centralized node acting as cache, P2P tracker and proxy to improve the QoE of video streaming for users of OSNs. In the experimental set-up considered, SEConD can attain a significant reduction (even up to ~87%) of the total inter-AS traffic compared to inter AS traffic generated when applying the client-server paradigm in this set-up, thus advocating that it is a promising TM mechanism. It may achieves up to ~88% reduction to the contribution of the origin server where the video is hosted. Results show ~80% in prefetching accuracy. In addition, it minimizes redundant traffic in inter-AS links both for prefixes and videos. It also eliminates redundant prefetching (same prefix from multiple sources to same destination). As the number of customers within the AS increases, relatively less SPS cache size is required in order to achieve the desired QoS. However, SEConD does not manage the intra-AS traffic, thus an extension of the mechanism is needed in order to handle unfavorable events (e.g. congestion) in backhaul network.

Joint evaluation of SEConD with RB-HORST++ shows the potential of QoS-based user-assisted video delivery as a means to boost users' QoE. The obtained results meet the expectations with the only reservation being that QoE is addressed in an indirect way. In order to solve this problem, a more direct user-side mechanism for the QoE quantification could be implemented, to prove the applicability of the current results.

vINCENT provides an incentive scheme for users to offer WiFi network access to even unknown parties. In the scheme, each user participates with his mobile phone and his uNaDa. It grants their owners an appropriate incentive by allowing them to consume a fair share of the offloading capabilities of other access points. The main lesson learnt is that the incentive cannot be granted without taking an access point's geographical position into account, as differing user densities have a major influence. Despite different user densities (e.g. rural areas vs. city areas), the proposed scheme is fair to all users. The vINCENT scheme allows for an increase of 5% overall offloaded volume as opposed to a Tit-for-Tat scheme. Moreover, vINCENT is closer to the unrestricted case, in which all peers always have unrestricted access to all access points nearby. vINCENT should be able to isolate free riding nodes, i.e., mobile users consuming offloading resources while not contributing

any service with their uNaDa. The difference in performance when dropping nodes only consuming service from the system while not providing any service with their own uNaDas, is as high as 20%, regardless the fraction of free riders in the system.

MONA addresses the energy consumption of smartphones. The available energy on mobile devices is inherently limited, while at the same time mobile data connectivity becomes more power demanding. Still, optimizations are possible by always using the most energy-efficient interface, and coordinating network requests, if QoE requirements permit. Aggregation of traffic or deferring transmissions until a more energy efficient connectivity option is available may reduce power consumption: between 34% and 85% of the otherwise consumed energy can be saved. The proposed MPTCP connection selection algorithm optimizes the energy efficiency of the mobile device. At the same time this algorithm considers the QoE of the end user by adjusting the selection of the network interfaces to the traffic demands of the mobile user. This works is in close connection with RB-Horst, and vINCENT, as by increasing the availability of local content, and thus increasing the available data rates, or in the case of vINCENT providing additional offloading opportunities, the mobile connectivity selection attains its full potential. Hence, both the energy efficiency aspect and the QoE aspect are addressed by the MONA mechanism.

Joint evaluation of MONA with RB-HORST++ investigates the energy consumption for mobile video streaming sessions and shows that minimum energy is consumed for connections offloaded to WiFi, which can only be achieved for high data rates that are not available in streets with current WiFi access point deployment.

MUCAPS adds network layer awareness to decisions on application resource endpoint selection. It is a pro-active mechanism based on ISP specified non real-time network state considerations and cost policy. Such a process can complement adaptive decisions on social and energy awareness derived by other mechanisms proposed in SmartenIT. MUCAPS addresses the EFS, since it selects the best end-to-end path between application and user endpoints. It uses the IETF ALTO protocol and actually may also address the OFS if the utilized ALTO service is the “filtered cost map service” which evaluates paths between network regions rather than endpoints. Last, MUCAPS basically performs the automatic decision making for ALTO Clients, by enabling them to decide what ALTO Service to request w.r.t. the application and the User Equipment receiving application data and what to do with the received information. ALTO is already part of the SmartenIT architecture as the benefits of using ALTO are considered to be granted since a couple of years. So MUCAPS in SmartenIT is halfway between TM mechanisms research and “a new architectural component”. It ensures a mutual winning situation for applications, users and ISPs: ISPs costs will benefit from network aware decisions by applications; users and apps will not choose the least ISP cost path if it has poor bandwidth resources; but they will also adapt the AEP selection to the user conditions and needs. Initial MUCAPS results’ have shown that a reliable evaluation requires to involve User Endpoint UEP related parameters such as at least the access type of the User Endpoint UEP, or simply user expectations on QoE. The present ones in addition show the importance of device capabilities. They also motivate the utility of integrating QoE in the AEP cost, as proposed in AQAS. Indeed, the current path cost definition in ALTO cannot

assess the overall quality of an access path such as a WiFi or has no clue of user perceived or expected QoE.

4.5 Key metrics and parameters of SmartentIT TM mechanisms

This section summarizes the most important metrics and parameters (including some indicative values of parameters) from all mechanisms for both OFS and EFS.

The most important metrics employed for the evaluation of each TM mechanism are summarized in Table 4-5. As expected, TM mechanisms that address OFS mainly are assessed based on metrics expressing inter-domain traffic and transit cost reduction. On the other hand, TM mechanism addressing EFS demonstrate a larger variety including both metrics associated to caching mechanisms, metrics related to energy consumption/energy efficiency, and QoE-awareness.

Table 4-8: Most significant metrics of TM mechanisms for OFS and EFS.

Metrics	M1	M2	M3	M4	M5
Operator focused scenario					
DTM	Total cost of inter-domain traffic Cost on each inter-domain link may be based on volume based tariff or 95th percentile based tariff	Goodness of compensation ρ It is defined as a ratio between achieved cost to the predicted one (expected according to reference vector)	Relative traffic cost reduction (in comparison to non-DTM scenario) $\Delta D^{(i)}$ where $i=1..N$ - number of inter-domain links	Absolute traffic cost reduction $\Delta \xi(i)$ where $i=1..N$ - number of inter-domain links	Traffic prediction accuracy, accuracy of reference vector calculation
ICC	95 th percentile-based cost of inter-domain traffic 95th percentile attained at the ISP transit link under ICC; this is also compared with the 95th percentile attained when ICC is not run.	Shiftable traffic delays To estimate the delay of the non-manageable traffic, we increase a counter initially set to 0 every time one bit is delayed per one epoch of 300/y secs	Volume of manageable traffic sent The total volume of time-shiftable traffic sent within the experiment	Traffic patterns: 5% peak rates For both cases (with and w/o ICC) we compare the average rate to the 5% peak rate value for the traffic pattern for both cases so as to estimate the potential savings with traffic shaping	Traffic patterns: average vs peak rate For both cases (with and w/o ICC) we compare the average to peak rate for the traffic pattern so as to estimate the potential savings with traffic shaping
MRA	Degree of fairness in the cloud/federation. The degree of fairness can be quantified by mapping the (global) greediness	Cost and performance loss due to resource reallocation. Each reallocation operation should be associated to cost and	Performance Isolation The degree of performance isolation should be measured by the performance decrease a		

	vector to a fairness index with various single-resource fairness metrics, such as Jain's index	performance loss. Subsequently it should be counted how often each of these operations is carried out	moderate VM exhibits, when heavy VMs are started on its PM		
End user focused scenario					
RB-HORST	<p>Cache hit ratio</p> <p>We determine the cache efficiency by calculating the ratio of cache hits and cache misses.</p>	<p>ISP cache contribution</p> <p>We determine the number of requests served by the ISP cache to evaluate the load taken off the ISP cache.</p>	<p>Offloaded traffic</p> <p>We determine the number of connections that can be offloaded to WiFi to evaluate the potential to take load of the cellular network.</p>	<p>Inter-domain traffic</p> <p>We estimate the traffic generated by video distribution in inter-domain links, since inter-domain traffic may lead to transit traffic charges.</p>	<p>Prefetching accuracy</p> <p>The prefetching accuracy is calculated by the fraction of prefetched items that is actually downloaded and is used to evaluate the efficiency of social and overlay prediction.</p>
SEConD	<p>Inter/Intra-AS traffic</p> <p>We estimate the traffic generated by video distribution in inter-AS links, since inter-AS traffic may lead to transit traffic charges.</p>	<p>Prefetching accuracy</p> <p>We estimate the percentage of video prefixes that have been proactively stored in users' equipment and the user eventually used them</p>	<p>Reduction of contribution of Content Provider</p> <p>We define as contribution of the Content provider the ratio of the (load of content shared by the Content Provider)/(total load of content downloaded by all end-users).</p>	<p>Caching accuracy of Social Proxy Server</p> <p>We estimate the percentage of video prefixes or videos that had already been stored in the cache of the SPS when a user requested it.</p>	<p>SPS vs P2P contribution</p> <p>Social Proxy Server vs local P2P contribution.</p>
VINCENT	<p>Percentage of offloaded traffic volume</p> <p>The overall traffic volume that could be offloaded to WiFi during the simulation run.</p>	<p>Percentage of offloaded data by free riding nodes</p> <p>Traffic volume that could be offloaded by free riding peers during the simulation run.</p>	<p>Price of Anarchy</p> <p>Comparison of performance gap between the applied incentive scheme and the no incentive case assuming altruistic nodes, only.</p>		
MONA	<p>Energy consumption</p> <p>Consumed energy during the transfer of content of a given size.</p>	<p>Throughput</p> <p>Average application level bitrate (TCP goodput) over the duration of a file transfer.</p>	<p>RTT</p> <p>Round-trip time to the content provider (e.g. cloud, uNaDa)</p>	<p>Available network technology</p> <p>Round-trip time to the content provider (e.g. cloud, uNaDa))</p>	<p>Signal Strength</p> <p>Received signal power at the mobile device. The signal strength affects the energy consumption by limiting the available throughput, RTT, but also defines</p>

					the power consumed by power amplifier.
MUCAPS	VQM Video Quality Metric (VITALU specific MOS)	Start-time delay date of media (video/audio) play – date of media request	NFrz, DFrz Number of freezes, Duration of freezes (statistics)	Media Bit Rate Media = video, audio..., measured values + relative values w.r.t. device conditions	XL-Utility of the selected AEP, w.r.t. weighted decision vector (w1.RC, w2.BWS) with : RC = routing cost, BWS = Bandwidth Score, those 2 values being associated to the AEP.

Furthermore, the key parameters used in the evaluation of performance of each TM mechanisms are summarized in Table 4-6. Their value ranges have been assessed by means of simulations. It can be observed that the set of key parameters demonstrates an even larger variety, as different mechanisms employ different methods and a wide variety of specifically designed algorithms to achieve the targets of SmartenIT per use case (as defined in section 3.3). For instance, DTM/DTM++ and ICC, OFS TM mechanisms that address inter-cloud communication focusing on the network level, consider traffic patterns, cost functions and time slotting, while MRA considers CPU stress and workload type, as it focuses on the cloud level. On the other hand, EFS mechanisms employ parameters and features, such a caching strategy and user interests concerning caching of content, sharing probability and cache contribution/participation regarding resource sharing, device type, bitrate and network availability in the case of energy efficiency.

Table 4-9 : Most important parameters of TM mechanism for OFS and EFS.

Parameters	P1	P2	P3	Values
Operator focused scenario				
DTM	Input parameters for DTM: Cost functions per link and per domain Type of tariff 95th percentile	DTM configuration parameters: report period DTM, compensation period [D2.4, D3.4]	SDN controller mode	Report period DTM: 10s, 30s Compensation period: 10s or 30s (95th perc. tariff), 30s or 5min (volume tariff) SDN controller modes: <ul style="list-style-type: none"> reactive without reference reactive with reference proactive without

	or volume Billing period length (same for each domain)			reference • proactive with reference
ICC	Ctarget target transit cost	Number of epochs The value of the number of epochs per 5-min interval <i>y</i> defines how many times the ICC control rate algorithm is applied	Threshold values tholds[1,...,y] The <i>threshold values</i> per epoch determine how close to the <i>Ctarget</i> will be the attempted rate decided by the ICC algorithm within each <i>y</i> -epoch: this is needed due to the a priori unknown values of the real time traffic	P1:Ctarget in [0.7,0.95] * Capacity,P2: <i>y</i> =10,P3: tholds[1,...,y] in [0.85,0.95]
MRA	CPU Stress: A standard Linux cpu stress test, which spins on sqrt(), is pinned to every physical core that a VCPU of the measured VM is mapped to.	Number VCPUs The number of VCPUs that the VM is configured with. This value is not changed mid-measurement.	Workload type The workload that is executed inside the VM. It is chosen from a set of workloads that often run in clouds.	P1:True/False, P2: 1-23, P3 apache/compression/aio-stress, nginx, python, php
End user focused scenario				
RB-HORST	sharing probability The probability that a home router is shared for content delivery and offloading using RB-HORST.	caching strategy The caching strategy used by home routers using RB-HORST.	support threshold The threshold determining the available upload bandwidth of a home router to take part in content delivery.	P1: 1, P2: LRU, P3: 750kbps (dependent on upload bandwidth)
SEConD	SPS cache size The storage size of the SPS cache	SPS QoS-based participation rule The total upload bandwidth within the swarm as a threshold for SPS participation in content-based P2P	Social Tie thresholds Thresholds that decide the distribution of the audience of a user into viewer	P1= ln (# users * (video_size + prefix_size)), P2 > video bit-rate, (≈ 2 * video bit-rate), P3=(80%,30%,20%)

		overlay.	categories. The choice is based on the percentage of the videos of the uploader the viewer has watched and on his interests.	
VINCENT	Accounting Method Method used for accounting contributed service to the system (e.g., provided offloaded volume)	Parameterization Regression Model Parameterization of input data for regression model (e.g., size of the radius for access point density calculations)		Varying algorithms for the accounting method, radius size in meters
MONA	Available. network performance Maximum throughput, minimum RTT and jitter available per technology	Device type Defines the power consumption of the individual components, in particular available access network technologies and their power requirements	Network availability Defines, which network type (e.g. EDGE, UMTS, HSDPA+, LTE, 802.11g, 802.11n, etc) is available at the time of the content request.	Determined using measurements, from which CDFs are derived as input to the simulations
MUCAPS	RC ISP defined Routing Cost = unitless metric	BWScore unitless ISP score on e2e path bandwidth availability	Access Type	P1 >= 0 e.g. in [0, 100] P2= [0,20] if pure score [0, 10MBpF] – MBpF MBpF = mean MB allocation per flow

4.6 Key design goals of TM mechanisms for SmartenIT use cases

This section documents the key design goals addressed by TM mechanism to cover the SmartenIT use cases. Table 4-7 documents the key design goals of the TM mechanisms. For each goal it is stated by which mechanisms it is met and how. Details on the evaluation of the goals can be found in the respective subsection of section 4.

Table 4-10: Key design goals met by SmartenIT TM mechanisms.

Key Design Goal	Goal Met
Inter-domain traffic reduction	- RB-HORST serves requests from local

	<p>shared UNaDas</p> <ul style="list-style-type: none"> - SEConD serves requests from SPS and local P2P connections
Inter-domain traffic cost reduction	<ul style="list-style-type: none"> - DTM is designed to reduce transit costs on inter-domain links and considers absolute and relative cost reduction - ICC reduces the 95th percentile of inter-domain traffic by shifting traffic in peak periods
Inter-cloud traffic management	<ul style="list-style-type: none"> - ICC reduces shifts inter-cloud traffic in peak periods - DTM distributes inter-cloud traffic efficiently among inter-domain links
Energy efficiency of user equipment	<ul style="list-style-type: none"> - MONA uses the most energy efficient technology for mobile access - RB-HORST offloads mobile traffic to close WiFi access points
Energy efficiency of cloud resources	<ul style="list-style-type: none"> - ICC is designed to consider load and energy cost of data-center resources for inter-cloud/-DC bulk data transfers - MRA reduces resource consumption and execution time of jobs executed in cloud
Take load off cellular networks	<ul style="list-style-type: none"> - RB-HORST takes load off cellular networks by offloading to shared local WiFi access points
Social Awareness	<ul style="list-style-type: none"> - SEConD prefetches content based on social information for efficient content delivery - RB-HORST uses information from social networks, as well as neighboring home routers for efficient prefetching in the overlay

Improve QoE	<ul style="list-style-type: none">- SEConD prefetches chunks important for fluent video playback- Content distribution is supported by SPS in SEConD to ensure good QoE- RB-HORST+MONA increase the throughput of mobile access to improve video streaming QoE- MUCAPS optimizes QoE by selecting appropriate communication endpoints and providing cross layer incentives- ICC (and DTM++) has positive effect on QoE by means of removing load from the transit links at peak times and prioritizing QoE sensitive flows while delaying time-shiftable bulk data transfers.
Fairness	<ul style="list-style-type: none">- vINCENT implements an incentive scheme to provide fairness among end-users- MRA fairly distributes delay experienced by VMs of different customer relative to the amount of resource requests
Incentive Compatibility	<ul style="list-style-type: none">- ICC uses a pricing scheme for incentive compatible marking of traffic and sharing of profit- DTM implements a pricing model to provide incentives for inter-domain traffic cost reduction- MUCAPS provides cross layer incentives to optimize traffic and QoE
Scalability	<ul style="list-style-type: none">- RB-HORST leverages edge resources for scalable content distribution and WiFi offloading- Content delivery in SEConD is supported by P2P connections

5 Tussle analysis of SmartenIT ecosystem

SmartenIT considers an ecosystem of stakeholders such as ISPs, Cloud Service Providers, Content Providers, end-users, as presented in the collection of use cases presented in section 3.2 and section 3.3. Stakeholders are entities, individuals or organizations, supervising or making decisions that affect how the Internet ecosystem operates and evolves. They are considered to be rational, self-interested players, which following their interests perform certain actions towards the satisfaction of them. In complex, multi-stakeholder environments, it can be expected that different stakeholders involved in a specific activity may result having conflicting interests. To describe the competitive behaviour due to conflicting interests of stakeholders in the Internet, the term tussle was introduced by Clark et al [49]. A tussle is a process in which stakeholders interactions usually lead to contention. Reasons for tussles to arise in our scenarios of interest are manifold, e.g., overlay / cloud traffic management and routing decisions between autonomous systems constitute a typical example for tussle space. Thus, with the ongoing success of the Internet and with the assumption of a future Internet being a competitive marketplace with a growing number of stakeholders such as users, service providers, network providers, and cloud providers, tussle analysis becomes an important approach to assess the impact of stakeholder behaviour.

To address an up-rising tussle among Internet stakeholders, incentives, as an economic mechanism to motivate an entity to perform certain actions or to pursue certain goals, can be the tool to motivate stakeholders towards an incentive-compatible behaviour. In SmartenIT, we tried to design mechanisms that address incentives of the various involved stakeholders, so as to avoid potential tussles among them. We focused on two types of incentives: monetary incentives, e.g., revenues for providing a specific service, and performance incentives, e.g., enjoying high(er) QoE or experiencing low(er) congestion.

In the next sections, we describe three tussle scenarios corresponding to an OFS use case, an EFS use case and a cross-scenario one; the latter is considered to be addressed by two pairs of OFS and EFS mechanisms.

A tussle scenario is an overall outline of real world actions and events occurring during an interaction with the underlying SmartenIT system. It describes the general behaviour of all system mechanisms and enables the identification of challenges and problems to be solved in order to achieve desired functionality. Then, we perform tussle analysis with an eye on the SmartenIT mechanisms that have been proposed by the SmartenIT consortium to address the targets and objectives of each scenario. Tussle analysis is also employed as a means to qualitative access the incentive compatibility achieved by those TM mechanisms in the context of the reference scenarios.

The tussle analysis comprises also a stakeholder analysis which includes the identification and assessment of the major entities related to an activity based on their influence and importance, and a stakeholder relationship analysis, which identifies the types of interdependence between stakeholders and how these dependencies can be affected towards some specific outcome.

5.1 Tussles in the Operator Focused Scenario

The use case " Bulk Data Transfers for Cloud Operators " of the operator focused scenario is motivated by the increasing communication among datacenters (DCs) and cloud service providers (CSPs), as well as the web back office traffic, i.e. the backbone traffic supporting the back office operations of web applications – such as the traffic exchanged among caches of the same CDN – efficiently in terms of delay and throughput.

The idea is that datacenters or small cloud service providers need to exchange data in order to either support their services. In particular, CSPs/DCs push some content or service instance towards the users that are expected to need it (resp. invoke it), or to perform operations that are inherent in their business, such as forwarding and replicating content to remote datacenters for backup or redundancy reasons. Since the clouds/datacenters are in general hosted by different Internet Service Providers (ISPs) and their respective traffic are constantly increasing in volume, the impact of the inter-cloud communication data transfers on the network performance of the ISP, its inter-domain transit link traffic and thus its respective transit charge are becoming increasingly more apparent and important. This motivates the need to properly manage the inter-cloud communication traffic in a way that is beneficial for all the stakeholders involved and resolve the respective tussles involved.

This section contains the tussle analysis of the "Bulk Data Transfers for Cloud Operators" use case. The major stakeholders are identified, as well as their respective interests, actions and concerns. The most prominent tussles are overviewed and the way these are handled by the relevant SmartenIT mechanisms is briefly discussed. In particular, the SmartenIT mechanisms attempt to resolve the respective tussles by means of adopting the "design-for-tussle" principles in the mechanism design and providing proper interfaces and incentive schemes for stakeholders to communicate their interests and resolve or mitigate potential tussles and conflicts in an incentive-compatible way. In order to perform the tussle analysis, the stakeholders involved are identified in the next subsection, followed by the identification of the respective tussles and their analysis.

5.1.1 Stakeholders

This subsection contains the major stakeholders of the "Bulk Data Transfers for Cloud Operators" use case for completeness reasons so as to make the section self-contained. Note that some grouping has been performed in cases where a detailed breakdown of stakeholders (such as Application Service Providers) does not affect the tussle analysis of this use case.

5.1.1.1 Cloud Service Providers and Datacenters

Cloud Service Providers owning infrastructure and Datacenters offer storage, computational infrastructure and content delivery and application hosting to businesses (e.g. Application Service Providers, Content Delivery Networks) and end users. They offer a rich set of managed services over different market segments and typically buy Internet connectivity from an ISP. They generate substantial amounts of traffic both towards the users that access their services and among components that need to communicate in the background so as to efficiently provide services, such as communication for data transfers

among the caches of a CDN. A portion of this traffic is delay-sensitive (e.g. real-time/“zero-tolerance” cloud services) but there is also traffic that applies to data that need to be transferred over potentially large time window without Quality of Service (QoS) constraints, e.g. bulk data transfers for performing content replication or backups.

5.1.1.2 Internet Service Providers

An Internet Service Provider (ISP) sells Internet connectivity to users and forms also business relationships with other ISPs. An ISP is a Tier-1, 2, 3 network operator that would normally own and operate its network, and is responsible for the provisioning of connectivity services and respective functionalities. Tier-1 ISPs are large ISPs with a big customer base and network footprint: Tier-1 ISPs only sell and do not purchase Internet connectivity contracts (transit interconnection) from any other ISP. All Tier-1 ISPs maintain settlement-free peering agreements among them, realized in their networks via multiple interconnection points (in prominent Internet exchange points) allowing the mutual termination of the traffic of their users within their networks.

Note that for transit agreement there is a buyer-seller business relationship between two ISPs where the seller has the obligation to terminate the packets of the seller for any Internet destination, hence also outside its network. Transit interconnection contracts offer a statistical guarantee of uptime and bandwidth over a transit link and the buyer is also charged according to the traffic crossing the link via the 95th percentile rule. On the other hand, there is no exchange of money or buyer-seller relationship in settlement-free peering, which is a non-transitive agreement between two peer ISPs.

There are also interesting variations of peering involving multiple ISPs, such as “donut peering”, or exchange of money, such as “paid peering”, which are outside the scope of this analysis.

5.1.1.3 Users

Users may be distinguished to either corporate (ISP business customers) or residential (end users). End users enjoy the services of their home ISP, i.e. Internet connectivity, which in turn allows them to access a rich set of Over-The-Top services and applications such as gaming, VoIP, video conferencing, infotainment, email, cloud services. The end users access services that are either sensitive to Quality of Service (QoS) parameters such as throughput and delay (e.g. video conferencing) or not (e.g. email). Corporate users include CSPs and DCs, typically have some ICT infrastructure and purchase Internet connectivity from the ISPs typically via dedicated leased lines of a certain capacity.

5.1.2 Tussles identified and resolved by SmartenIT

After the involved stakeholders and their interests are identified, conflicts among them, i.e., the tussle space, can be identified too. In each of the paragraphs below, we identify the reason for such a tussle and the associated trade-offs.

5.1.2.1 Optimal Traffic Destination Selection

For various cases of the Inter-Cloud Communication use case there may be multiple destinations (DCs and CSPs) where the respective traffic can be forwarded to. For instance, this is a case for backups where multiple destinations can serve the needs of replications and redundancy. The selection of the optimal destination is currently done by the sending CSP/DC without any knowledge on the underlying network load which may affect both the time within which the data transfer will terminate as well as the underlying network load. It is possible that the selection of the source CSP/DC, solely based on local information, is inefficient for the data transfer and causes extra overhead and congestion to the underlying network links from which the traffic is forwarded to. A different destination selection could be more beneficial to the service and to the network, provided that there is accurate information sharing between the cloud and network layer stakeholders and both layers' constraints are taken into account in the decision making. Thus there is contention over the control of the decision of the flow destination and the amount of information made available to the respective stakeholder to make this decision. The goal would be to resolve this tussle in a win-win fashion for both the CSP/DC and the ISP.

5.1.2.2 Intra-Domain Inter-Cloud Communication Traffic Management

There is a tussle between the ISP and the CSP/DC regarding the effects of traffic management that the ISP can apply to the Inter-Cloud Communication traffic, once the traffic destination has been selected and the respective traffic flow is generated. In particular, the higher the degree of intervention of the home ISP to the way this traffic is managed – thus changing its traffic profile by means of exogenous to the CSP/DC interests “reasonable traffic management” – is, then the higher are the cost savings/gains that the ISP can attain in terms of load balancing its network and prioritizing other QoS-sensitive instead over its backbone. On the other hand, the CSP/DC is primarily concerned on getting his traffic managed in a way that results in the best QoS for its customers, regardless of the home ISP network load and links' utilization. The contention of the ISP and the CSP/DC regarding the scope, goals and means of traffic management of the Inter-Cloud Communication traffic within the home ISP domain comprises an additional tussle to be resolved.

5.1.2.3 Inter-Domain Inter-Cloud Communication Traffic Management

Though the ISPs' backbone is typically over-provisioned, the inter-domain links are typically highly utilized and thus there is higher resource utilization and competition among flows, while congestion may also appear at peak hours. In fact, peering links are typically under-provisioned mostly due to business issues such as backbone free-riding and business stealing, while inter-cloud communication traffic is typically forwarded via the transit inter-domain link(s). The latter is also charged based on its actual utilization, based on the 95th percentile rule. To this end, the shaping of the Inter-Cloud Communication traffic at the transit link can be beneficial for the ISP since reducing the peak traffic by removing the Inter-Cloud Communication traffic from peak hours can reduce its transit charge. On the other hand, this will result in more delay for the sending CSP/DC traffic, thus resulting in dis-satisfaction for the CSP/DC and its customers. These two conflicting effects need to be mutually addressed and balanced in an incentive-compatible way.

5.1.2.4 Net Neutrality

The traffic management of (a portion of) the Inter-Cloud Communication in a fashion that ISP can have problems to cope with the fast changing and growing traffic demands. Then transmissions can be subject to delays, throughput can be limited when the transit link is heavily utilized.

The European Commission is in the process to regulate such net neutrality aspects and has recently published their perspective on ISP traffic engineering [63] as part of the digital agenda:

Blocking and throttling measures are estimated as unfair traffic management practices with the effect of weakening the competition and decreasing innovation.

In order to deal with network congestion, ISPs have to implement certain quality parameters in their network. On the other hand, fast broadband coverage should be provided with to all European citizens at speeds equal to or higher than 30 Mbps.

Most ISPs nowadays have QoS-enabled networks, meaning that the active elements of the network can support multiple QoS classes, thus different portions of traffic can experience different quality and performance. Though these traffic classes are typically used for better than Best Effort performance and assured QoS, the ISP may be tempted to assign the time-shiftable traffic to a lower priority class inside his network, similarly to the way P2P traffic has been traditionally managed by many ISPs. This raises a net neutrality tussle, regarding the extent of traffic management that can be employed by the ISP for the Inter-Cloud communication use case, which needs to be resolved by means of proper incentive mechanisms.

5.1.2.5 Information Asymmetry and Hidden (Network) Effort

The current network and interconnection contracts among the ISPs as well as between ISPs and their client CSPs/DCs do not provide any kind of statistical QoS guarantees; there are solely guarantees regarding maximum possible rate (theoretical), uptime and/or connection recovery. Hence, the CSP/DC traffic is typically handled a la Best Effort, without allowing the CSP/DC to verify whether his traffic could be handled in a better way by its home ISP. Hence, it is not clear whether for instance the delay exhibited over a data transfer is a result of the actual network conditions over the respective network path or this performance has been degraded due to lack of effort by the respective ISP(s). Since there is no way for the traffic source to observe the actual effort of the network stakeholders, there is a tussle over the delay and throughput of the respective traffic, since the underlying ISP may favor other flows (e.g. video flows of its CDN/video platform) in the expense of the CSP/DC. This tussle could be resolved if there was a connection between the statistical QoS attained and the resulting ISP monetary charge (or compensation), which is currently lacking for Best Effort Internet connectivity contracts [34].

5.1.2.6 Discussion

The SmartenIT ICC Traffic Management mechanism for which we apply the tussle analysis for this use case – and thus also DTM++ where ICC has been integrated for multi-homed ISPs – have been designed adopting the design-for-tussle principle without

favoring any kind of stakeholder in the expense of the other. Moreover, they provide the necessary interfaces and communication protocols for the stakeholders to exchange their preferences and thus resolve the respective tussles via incentives.

In particular, the ICC Cloud-Network layer interface allows for the cloud and network load to be communicated to the source CSP/DC (or to the federation CSP/DC) so that the optimal destination for bulk data transfers of this use case can be selected. This mitigates the potential negative impact to the network of a bad selection of traffic destination. However the control remains at the cloud layer stakeholder (sending or federation CSP/DC) in order to respect their respective business processes and decisions. For further details on this, as well as the ICC mechanism's features mentioned below the reader may refer to section 5.2 and section 13.3 of Deliverable D2.4 for the detailed presentation and specification respectively of the ICC mechanism.

The tussles pertaining to the intra- and inter-domain traffic management are resolved by means of:

- a) prescribing that the sender CSP/DC is to mark the traffic that can be delayed/time-shifted, as opposed to allowing the ISP to choose this traffic e.g. via Deep Packet Inspection; this ensures that critical delay-sensitive traffic will be served in the best possible way – and in fact better due to the “removal” of some “time-shiftable” traffic from the transit link when it is highly utilized thus resulting in better delays – thus causing no harm to CSP/DC critical traffic flows and contributing to load balancing the ISP network and reducing the respective transit link charge.
- b) prescribing that the ISP will provide a discount for the additional delay incurred to the CSP/DC traffic flow in the form of a discount for the respective traffic charge; thus a cut of the ISP transit link cost savings attained due to the delay of the Inter-Cloud Communication traffic flows is returned to the CSPs/DCs that willingly marked a portion of their traffic as “time-shiftable”/manageable thus enabling those gains. This provides an incentive loop that allows the ICC mechanism to work, also in a net neutral way, since a “win-win” solution is sought in the traffic management, instead of imposing the network traffic management to the CSP/DC flows.

Furthermore, the fact that the performance of the network for the Inter-Cloud Communication use case is linked with the amount of charge of the client CSP/DC by the ISP provides an incentive:

- a) for the ISP to mitigate delays since these must be compensated, and
- b) for the CSP/DC to mark as time-shiftable all the portion of the traffic that is indeed time-shiftable, since this will lead to higher cost savings for the ISP and thus a lower respective charge for the sending CSP/DC.

Finally, the incentive schemes built into ICC (and thus DTM++) do apply between the CSP/DC and its home ISP but cannot apply among ISPs (e.g. between the home ISP and its transit ISP) that may be involved in carrying and terminating the traffic to its destination. Note that this is an inherent problem to the Best Effort Internet, which can be surpassed only in paradigms where incentive-based interconnection contracts and connections can be provided [35]. Though the SmartenIT mechanisms are intentionally designed for the

Best Effort Internet so as to enhance their applicability and impact, they can be extended to such paradigms, thus enabling tussle resolution also in beyond Best Effort contexts.

5.1.3 Relation with SmartenIT Business Models

This tussle analysis pertains to the SmartenIT mechanisms, namely ICC, DTM, DTM++ and business models at the wholesale (operator) level for inter-cloud and inter-network communication and transport networks over the Best Effort Internet and beyond. In particular, the tussles of subsection 5.1.2 apply to Best Effort network but also in the context of Future Internet where some form of quality of service could be provisioned even at inter-domain level. In the latter context the tussles presented earlier could also be resolved via structuring proper Service Level Agreements, interconnections contracts and respective incentive-compatible pricing schemes, thus the tussle analysis of this section could serve as a helpful guide to specifying them when needed.

The business models outlined in Chapter 3 of SmartenIT Deliverable D2.4, for which this tussle analysis is relevant are the Repeatable DSP, Federation and ISP Managed Services business models. We remind to the reader that the Repeatable DSP model is a TM Forum attempt to adapt the Amazon repeatable retail business model to cloud and Internet services based on an Ecosystem Enablement Platform (EEP) providing rich functionality to digital service providers; the ICC mechanism, and also DTM and DTM++, could be a module of this platform, to be invoked when needed by CSPs (additional details are provided in subsection 3.3.3 of Deliverable D2.4). The Federation model (additional details are provided in subsection 3.4.2 of Deliverable D2.4) is also a direct match since ICC has by design a cloud layer facilitating optimal destination selection of bulk data transfers which explicitly considers federations of clouds. Finally, the ISP Managed Services model is also an excellent match since ICC, DTM and DTM++ could be an integral part of this paradigm as a whole and also for particular managed services such as Virtual Private Cloud (VPC) and Managed Storage and Backup (additional details are provided in section 3.4 of Deliverable D2.4); in the latter cases ICC intelligence could be integrated and used for the bulk data transfers among the VPC instances and to perform scheduled managed back-ups. Therefore, given the close relations of ICC and DTM mechanisms with the aforementioned business models, all the tussles presented in this section are relevant to these contexts. Also, additional tussles may arise due to the different forms of customer ownership, intermediation and control. In particular for the cooperative/competitive business model of Federation additional tussles arise regarding the role and boundaries of Federation as well as the competing forms of conducting business – individually or via the Federation. These issues are revisited and investigated in the last section of this chapter.

5.2 Tussle in the End-user Focused Scenario

The use case "Exploit Content Locality" has the goal of exploiting UNaDas proximity to end-users to reduce inter-domain traffic and to improve QoE. The idea is to cache or pre-fetch content in end-user hosted Nano Data centers (uNaDas) and deliver it from there to other end-users nearby who request it. Since uNaDas are located in the edge network, RTTs and overall traffic will decrease significantly. However, arbitrary content cannot be

cached, because the chance of cache hits is too low for it. Instead, content to be cached has to be popular for a large user community or it has high relevance for sharing within a regional user group. Therefore, content from OSNs (e.g., pictures and movies shared via Facebook) may qualify for efficient caching as well as content from multimedia platforms such as YouTube. Content and accounts on multimedia platforms are often also linked from OSN profiles.

This section contains a preliminary tussle analysis of the "Exploit Content Locality" use case. The major stakeholders are identified, as well as their respective interests, actions and concerns. The most prominent tussles are overviewed and the way these are handled by the relevant SmartenIT mechanisms is briefly discussed. Concluding, the SmartenIT Traffic Management mechanisms for the "Exploit Content Locality" use case attempt to resolve the respective tussles by means of adopting the "design-for-tussle" principles in the mechanism design and providing proper interfaces and incentive schemes for stakeholders to communicate their interests and resolve or mitigate potential tussles and conflicts in an incentive-compatible way.

5.2.1 Stakeholders

5.2.1.1 End-users

End-users setup uNaDas and request content to be cached or pre-fetched by uNaDas. Their interest is to receive high QoE and to have their personal and usage data transmitted and stored confidentially. While end-users profit, if their uNaDa has content pre-fetched for them or is served by close-by uNaDas, they have no inherent interest in serving requests of other users with their uNaDa and even face costs for serving others (energy, bandwidth, deterioration of their other flows QoS/QoE).

5.2.1.2 ISPs

ISPs are interested in reducing traffic inside their network, because it decreases infrastructure and energy costs. Even more important to an ISP than decreasing traffic inside his network, is reducing inter-domain traffic, because this reduces transit costs and/or costs for expensive interconnection links such as long-distance or submarine cables. Transit costs can be high for small ISPs whereas large ISPs, e.g. Tier-1, and interconnection providers offer and make profit from transit contracts. NaDas may be deployed and controlled by an edge ISP, e.g. in home gateways, to reduce intra and inter AS traffic.

5.2.1.3 Cloud provider

This stakeholder group develops and deploys the software that is necessary to build OSNs and multimedia platforms. For deployment they install or rent data centers or use the services of CDN providers (introduced next). Users then upload or request content from the cloud provider. Cloud providers make not only revenue from ads or service fees but also from gathering and processing end-user behavior information, i.e., data mining. Therefore, these stakeholders seek ways to improve ad placements. They often track and analyze precisely, how end-users view content, e.g., how long they remain on a certain page, where their mouse pointer moves, and how/where they scroll.

The interest of this stakeholder in uNaDa caching may highly depend on whether they serve content without business constraints, i.e., by users and non-profit organizations e.g., founded wikis or civil services. Caching of such content is beneficial by improving download performance and/or saving resources for all stakeholders. This even includes part of content from commercial cloud providers, e.g. software updates, bug fixes for free or other information to support efficient access and proper usage for their services, which they want to be distributed as efficient as possible. On the other hand, if a cloud providers primary business model is data mining and provisioning of paid content, he will likely prefer serving this content from centralized/non-end-user-controlled infrastructures, as this gives him a higher degree of control.

Since the social or interest graph is critical to perform efficient caching and prefetching and the cloud provider is the only stakeholder, who has access to it, it is critical, that cloud providers support uNaDa caching and prefetching. Furthermore, nowadays most connections to OSNs or multimedia platforms are established via encrypted connections, making it impossible to cache content, without support of cloud providers.

5.2.1.4 CDN Provider

CDN Providers support cloud providers via distributed overlays including cache servers to host their applications and deliver their content to end-users. Because CDN providers are not directly connected to the end-user, they may pay ISPs to forward the traffic that originates from their servers. Therefore, they are interested in reducing transport costs to or from their data centers. While they are interested in reducing traffic they are also interested in securing their business, i.e., ensuring that they are needed to deliver data. Therefore, while the deployment of uNaDas can reduce traffic to their infrastructure it also threatens their business, as cloud providers may primarily rely on uNaDas to deliver their content and therefore buy less services from cloud providers. On the other hand, they could cooperate with uNaDas e.g. by facilitating transparent caching for their content in order to reduce transport costs.

5.2.2 Tussles identified and resolved by SmartenIT

After the involved stakeholders and their interests are identified, conflicts between these, i.e., the tussle space, can be identified. The first tussle is also the main motivation to introduce the uNaDa caching technology and is therefore solved as discussed. The other tussles are open, however, for each tussle at least one possible solution is proposed.

5.2.2.1 Reduction of cloud provider traffic traversing ISPs

As pointed out recently [50], it is important for all players involved in the video content delivery process to cooperate to make the content delivery more efficient. In particular, streaming video and popular content distribution platforms dramatically increased the volume and character of traffic, where much of the traffic originates from cloud providers and contracted CDN providers. In particular, "content originators" (cloud provider and CDN providers) argue with ISP over who is responsible for delivering the cloud provider content to end-users. More precisely, ISPs argue they should be paid, because they enable cloud providers to reach their end-users. On the other hand, content originators often argue, that ISPs are then also responsible alone for upgrading connection points. While there is a

point to both arguments, it recently has been shown, that this tussle is often not resolved, with neither party upgrading capacities, negatively affecting not only the traffic under discussion but due to caused congestion, the entire Internet. When considering the problem in more detail it is evident that both parties need to support a possible solution. It was also proposed to position the content under discussion in the ISPs network (thereby avoiding congested inter-AS links). This is exactly the approach uNaDa caching offers: the content will be positioned/cached in the edge ISP network. If this solution is supported by ISPs as well as content originators, it will greatly reduce the traffic while not affecting or even improving user experience. Therefore, the tussle about who has to pay for the delivery of cloud provider content is resolved. Particularly, the proposed solution solves the tussle in a sustainable manner, because not judicial or economical means are taken, but the root cause of the tussle is resolved (too much traffic nobody is willing to pay for).

5.2.2.2 Traceability of User Behavior

If content is provided from caches and not via a controlled connections to cloud provider servers (or infrastructure they rent and is therefore fully controlled by a business partner of them, e.g., a CDN provider), data mining and optimal ad placement is harder to achieve for cloud providers. Since data mining and ad placement is the primary business model of a considerable number of cloud providers and it is mandatory to have them on board to deploy uNaDas, this tussles deserves much attention in the future.

SmartenIT recommends solving this tussle by technical means. In particular, economical means would be to compensate the cloud provider for information he loses, however, this might not always be possible, for example, when completely bypassing the cloud provider. This compensation would have to happen by the ISP, because he sees a reduction in traffic. However, as ISPs and cloud providers are even unable to reach an agreement on who should cover the costs for the delivery of the traffic originating from an cloud provider, it is unlikely that an agreement about an appropriate compensation could be found. Also judicial means, i.e., legally forcing the cloud provider to integrate uNaDas in the content delivery process for the greater good of overall traffic reduction, seems unlikely to be effective. In particular, the problem is that laws that cover this highly technical and fast evolving environment are missing to reach a grounded verdict. Therefore, the only possible solution to this tussle, seems to be of technical nature: this means that cloud providers need to be enabled to trace end-user behavior, even if their requests are served from uNaDas.

SmartenIT proposes to solve this tussle by encryption: in particular, cloud provider could distribute the content to the uNaDas encrypted (only with unencrypted headers). The keys to decrypt the content, could be handled by centralized servers just as the entire content was before. Therefore, if a user requests content, that can be served by a uNaDa it is provided by it (the headers are unencrypted). However, to decrypt the content he will have to contact the “centralized” cloud provider infrastructure in order to receive the key. Therefore, the cloud provider will still have a full picture of which user consumes which content and can also mine data and place adds accordingly.

This solution was presented at a recent ITU-T workshop, where many security experts were attending [65]. The feedback at this workshop pointed out that since uNaDas will provide content rather unselectively, there are many similarities to TV broadcasts via

satellite (satellites also distribute information unselected and it needs, therefore, to be protected by encryption to only be viewed by authorized users). Since the latter determines a mature industry, it was suggested to look for refinements of the SmartenIT proposed solution.

5.2.2.3 *Tit-for-tat*

This tussle is similar to a problem encountered in P2P networks many years ago: while end-users profit if their uNaDa performs prefetching and receives content from other uNaDas, they have no intrinsic interest that their uNaDas serve other end-users. In particular, this will cost them energy and bandwidth. Since they have physical access to the uNaDa they might be able to reconfigure it to not serve other end-users. However, as P2P file sharing systems showed, end-users can show altruistic behavior. Furthermore, not many users are expected to have the expertise or motivation to modify the behavior of their uNaDa. For these two reasons, it can be assumed that this will not be a predominant tussle. Nonetheless, there are two means ISPs (who probably profit most from uNaDa deployment) can take, to motivate even selfish end-users with technical knowledge to provide other end-users with content.

The first approach would be that the ISP resolves this tussle on a technical level by increasing the bandwidth of end-users, who's uNaDas serve other end-users. Analogously, the ISP could solve the tussle on an economic level by granting discounts to end-users who serve other end-users.

5.2.2.4 *Inter-domain traffic changes*

This tussle arises when geographically co-located users are customers of different access ISPs and one user's uNaDa provides content to the other. While the users do not care about this, on a large scale this mechanism could affect the peering link load, thus changing peering ratios to one side's expense. Also if peering links are well provisioned then the uNaDas could serve as a "business stealing" trojan horse: a well-provisioned ISP is actually used for free by his national competitors ISPs via the uNaDas. This is due to the fact that the other ISPs refrain from investing in upgrading their own transit connections and backbone links infrastructure and their customers are still served in high QoS through the well-provisioned ISP's uNaDas. This results in a situation where the ISP that invests most in the market sees his infrastructure being utilized for free by his direct competitors who actually face lower costs due to their lower investments.

However, in the SmartenIT solution uNaDas can be configured to only serve other uNaDas that are a certain number of autonomous systems (AS) "away". By default, this number is 0, which means that uNaDas only serve uNaDas located in the same AS, i.e., which implies that they are hosted by the same ISP. To also allow serving other AS of the same ISP, uNaDas could be equipped with black or white lists of IP addresses they serve. Then an access ISP could update these lists to make the uNaDas in his network only serve other uNaDas in his network. While this would be the technical approach, the economical approach would be that interconnection contracts are changed and the ISP who serves customers of other ISPs with his uNaDas actually is financially rewarded. While this would allow to use the full caching potential of uNaDas (because, contrary to the technical

solution, every uNaDa can still serve every other) it will be hard to negotiate these new agreements.

5.2.3 Relation with SmartenIT Business Models

This tussle analysis pertains to the SmartenIT mechanisms, namely RBH, SEConD, vINCENT, MONA, MUCAPS and RBH++ and business models at the end-user level for resource sharing. In particular, the tussles of this section apply to Best Effort network but also in the context of Future Internet where some form of improved QoE would be provided to the end-users, accompanied by reduction of inter-domain traffic and energy efficiency for ISPs. In the latter context the tussles presented earlier could also be resolved via structuring proper incentive-compatible pricing schemes that will enable efficiency value and cost sharing between the ISP, the CSP and the end-users, thus the tussle analysis of this section could serve as a helpful guide to specifying them when needed. The business models, among those outlined in Chapter 3 of SmartenIT Deliverable D2.4, for which this tussle analysis is more relevant is the Terminal and Service business model. In particular, the “terminal+service” and “HaaS+Internet” trends and respective business models also indicate an increasing interest on innovation at the end user devices side. Thus, the SmartenIT EFS mechanisms, namely RBH, SEConD, vINCENT, MONA, MUCAPS and RBH++ that can deliver energy efficiency, advanced services and resource management, increased agility and performance, can be very useful means for new differentiated services and terminal operations. A prominent example is the cloudlets business case: Cloudlets are decentralized and widely-dispersed Internet infrastructure whose compute cycles and storage resources can be leveraged by nearby mobile computers, i.e. a “data center in a box”. The aforementioned SmartenIT EFS mechanisms are well positioned within this emerging business context, with which all the tussles presented in this section are relevant.

5.3 Tussles across Operator Focused and End-user Focused scenarios

This section focuses on tussles across the Operator Focused and End-user Focused scenarios. Concrete examples of synergies when combining SmartenIT OFS and EFS mechanisms are provided and it is demonstrated that super-additive gains can be attained. Then, tussle analysis is performed and its findings are also related to the business models that are relevant for SmartenIT.

5.3.1 Synergy of DTM and RBH

The DTM-RBH synergy covers both the EFS and OFS scenarios. RBH reduces the amount of inter-domain traffic by content caching and social-aware prefetching thus making the content available locally for users located in a given domain. From DTM perspective, the amount of manageable traffic is reduced but still there is a need to optimize traffic distribution among inter-domain links with DTM such that the costs for operators are minimized. Also, thanks to caching and prefetching RBH reduces load of data-centers since the content is downloaded by end-users from local caches (uNaDas) instead of data-centers.

Thus, the objectives met in this synergy are inter-domain cost mitigation, energy efficiency, and social awareness. QoE is enhanced by reducing latency with local caches.

RBH contributes to reduction of inter-domain traffic. RBH prefetches the content making socially-aware decisions, makes the content available locally, or increases local availability so end users need to download the content directly from remote DCs less frequently. There is a trade-off between availability and inter-domain traffic which can be tuned by the amount of content to prefetch. If the primary goal is to save inter-domain the prefetching accuracy targeted by the mechanism needs to be high enough to produce less inter-domain traffic than in the case prefetching is not used. DTM performs cost-aware decision on choosing the inter-domain link (tunnel) for the manageable traffic so as to minimize the cost of inter-domain traffic. DTM does not decrease the inter-domain traffic itself. The total traffic remains the same, only the distribution of the traffic among links is changed.

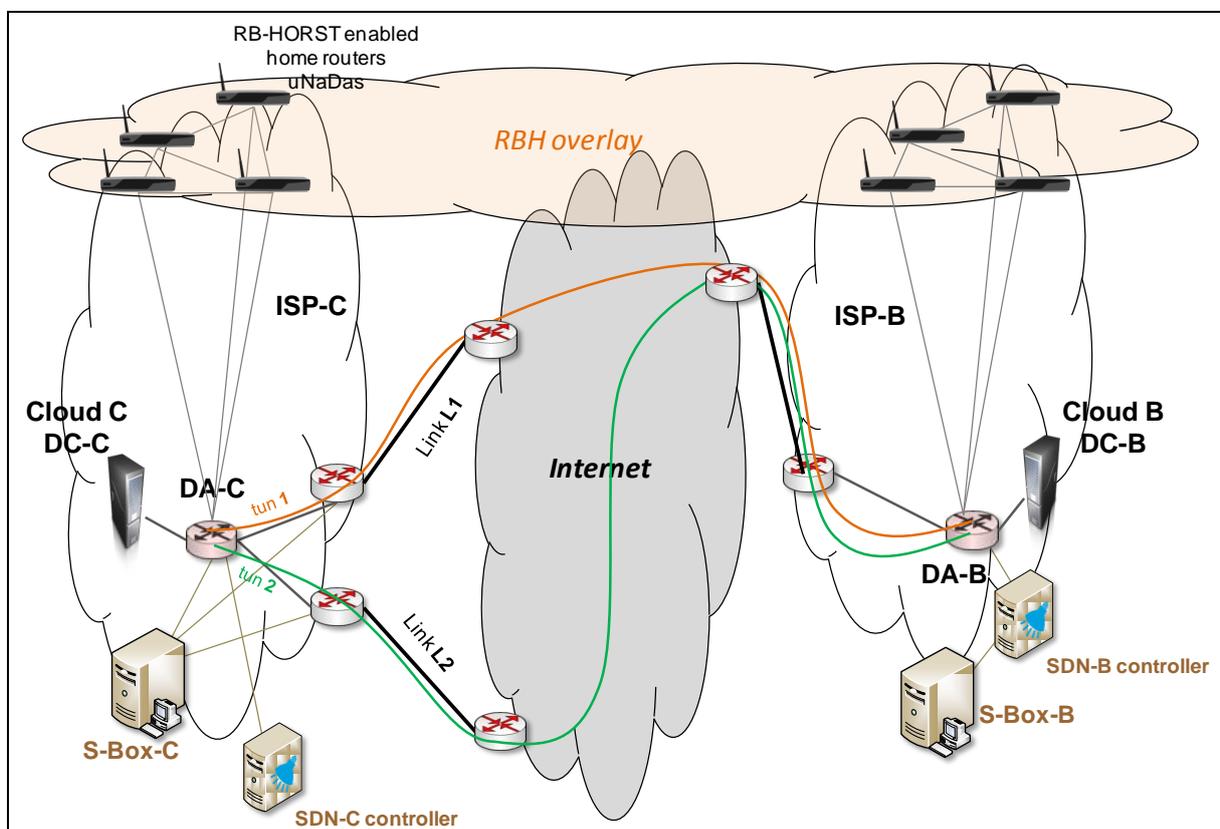


Figure 5-1: DTM-RBH architecture.

In a success scenario RBH and DTM would operate independently yet complementarily. A result of RBH operations is overall decrease of inter-domain traffic thanks to localization of the content while DTM focuses on inter-DC, inter-cloud manageable traffic and distributes it among inter-domain links in a cost effective manner. Additionally, the traffic from DC to uNaDas may be also managed by DTM similarly to inter-DC traffic. In this way ISP may decide on the inter-domain link used by RBH to prefetch the content from DC. However, this would require cooperation between RBH and DTM.

To sum up:

- no cooperation/integration of RBH and DTM is necessary but desirable
- DTM does not influence RBH traffic (inter-domain traffic exchanged between local and remote uNaDas or between uNaDa and DC is treated by DTM as non-manageable).
- ISP benefits from RBH and DTM
- end-users have benefits from RBH but DTM is transparent for them

However, in case of RB-HORST + DTM, the cooperation of operators and end-users can add more benefits for different stakeholders. End-users benefit from information provided by operators, e.g., when prefetching content to an RB-HORST cache, the inter-domain link for the content download is chosen by the ISP (using DTM). Operators benefit from information provided by end-users, e.g., information offered by RBH about social activity of end users can be utilized by ISPs to predict huge flows. This information may help in more accurate prediction of amount manageable inter-domain traffic. A more sophisticated algorithm for DTM may be considered. Traffic prediction based on social information may be used in calculation of compensation vector or even the reference vector may be recalculated on demand during billing period

Table 5-1 summarizes the design objectives address the synergy across scenarios of DTM and RBH. As expected, the set of addressed objectives is vastly broader than the respective set of objectives met by each of the mechanisms individually. The quantitative assessment of such synergies is left as future work.

Table 5-1: Design objectives addressed.

cross-layer design	Yes:
cloud layer,	RBH
network layer,	DTM/RBH
user layer,	RBH
applicability in the inter-domain,	Yes
incentive compatibility,	<i>Yes Incentives for ISP are clear, incentives for end-user to cooperate with ISP (in scenarios 2, 3, and 4) need investigation/discussion</i>
pursue collaboration,	Yes: ISP-ISP (DTM), cloud-ISP (DTM), end-user - ISP (RBH/DTM), end-user-cloud (RBH) <i>As above: collaboration of end-user with ISP needs investigation</i>
efficient network management,	Yes (DTM/RBH)
energy efficiency,	-
social awareness,	Yes (RBH)
QoE awareness	Yes (RBH)

5.3.2 Synergy of ICC and SEConD (AUEB)

The ICC-SEConD synergy covers both the EFS and OFS scenarios in a complementary fashion. In particular, SEConD can be used to mitigate the fact that ICC is solely focused on inter-provider data transfers without being able to estimate the impact on the end user

QoE. In particular, this decision is left to the business logic of the cloud service provider (CSP) or datacenter (DC), who in turn may not have enough information to decide on the timing of the bulk data transfers in an efficient and optimal in terms of user QoE way.

Thus, the objectives met in this synergy are inter-domain cost mitigation, end user QoE enhancement, and social awareness.

The idea is thus to inject into ICC social awareness and QoE aspects by means of exploiting social information to predict demand and thus make educated decisions regarding the scheduling of the data transfers. The social awareness logic of SEConD will be used in this synergy to decide on what needs to be moved, where and when, also in an ISP-friendly way that is enforced by ICC. This synergy can be implemented by means of introducing social awareness in the Cloud Scheduler (CloS) input to the SmartenIT Information Service (SmaS). In particular, this input is the amount of data, the priority level (urgent or delay-tolerant) and the candidate destinations that will be selected via the social prediction module of the SPS component of SEConD. Apart from the social awareness, the adoption of SPS gives an extra boost in the QoE of the ISP’s customers. This is achieved by enabling local and QoE-aware content dissemination techniques within the domain of ISP. This is depicted in Figure 5-2, for the case of federated cloud service providers/datacenters; the same rationale can be applied to the non-federated case as well.

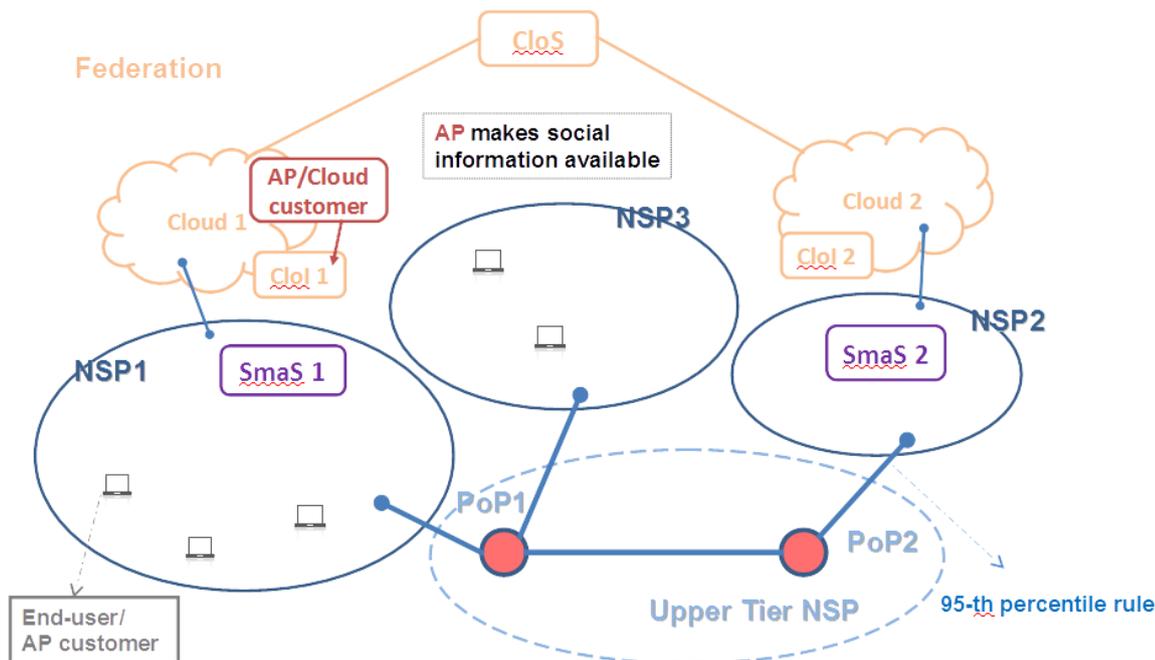


Figure 5-2: ICC and SEConD synergistic deployment in the case of federated clouds/datacenters.

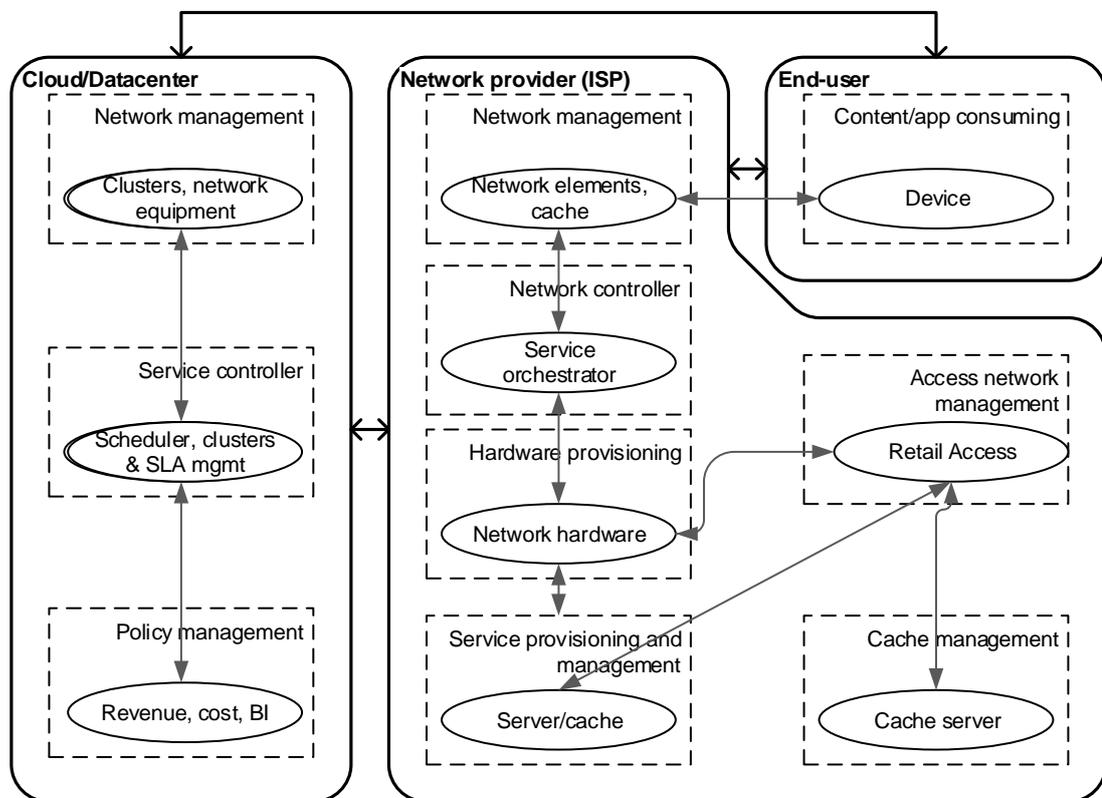


Figure 5-3: VNC for ICC-SEConD synergy scenario

A success scenario highlights the interworking and complementarity of two mechanisms addressing both EFS and OFS and includes the following steps:

Step 0: SLA between DC Operator and ISP for the connectivity required and the delay tolerance of the bulk data transfers. The ISP will offer a discount to Cloud for traffic that can be shaped rather than be instantaneously transferred.

Step 1: The DC/CSP decides a Service Level Specification of the back-up service to be offered and generates the SLA for its customers (IaaS and SaaS Providers).

Step 2: A certain service (e.g. storage of content) with predefined attributes and statistical guarantees provided in the respective Service Level Specification is purchased by an end-user whom is also offered an additional back-up service to guarantee the availability and integrity of its personal data (e.g. remote Datacenter locations where the data can be replicated periodically).

Step 3: Clol creates clusters and messaging overlay based on social information.

Step 4: Overlay assesses friends' actions and content consumption to trigger demand indications.

Step 5: CloS decides on bulk data transfer of aggregates given these demand indications.

Step 6: ISP shapes the time-shiftable traffic, so as to reduce the transit cost it is facing.

Step 7: Destination receives the data sent with some delay.

Step 8: CSP/DCs verify the coherency and the integrity of exchanged data.

Step 9: Assessment of QoE impact, incurred delay and ISP transit cost savings. Assessment should involve the verification of the SLAs conformance in all-layers (network and cloud).

The design objectives addressed by the synergy of ICC and SEConD comprise end-user QoE, social awareness and cost reduction for ISPs are handled by ICC-SEConD. Moreover, energy efficiency could also be considered, if such logic is incorporated into the CSP/DC business logic for the cloud service providers and datacenters.

5.3.2.1 Case A: Single SPS instance in the topology

The first case is the ISP-owned SPS case where there is a single instance of SPS per ISP network domain. All the customers of the ISP should be aware of the local SPS and exchange all the required information. Specifically, each SPS gathers social information from the local users' actions in social networks and thus is able to predict the demand for content within the ISP. Moreover, SPS operates as tracker in local content-specific p2p swarms and also cache content items in order to participate in swarms and assist in content distribution locally. In this ISP-owned SPS scenario, the ISP has the advantage and the ability to build a complete social profile for each customer, based on customers' actions in a broad range of different applications. This information could be retrieved from the end user actions (HTTP/application requests) that are logged by the ISP. Thus, under the assumption that the ISP can have access to different applications of user, the ISP is able to predict future demand with high accuracy.

Therefore, in the synergetic solution the intelligence of SEConD is used to enable the social awareness in the destination selection of a data transfer and for the intra-ISP content dissemination. In the initial specification of ICC, we assume that the source CSP of a data transfer creates a list of candidate destinations and then the cloud source or the federation in the federated CSPs/DCs scenario, decides on the actual destination CSP. **In this synergy, the SPS's functionality for prediction of demand can be used as input to Cloud layer in order to induce the transfer of data toward specific locations.** The data transfers can pertain to multiple applications such as video content or Web pages, documents accessed by the users and predicted to be of value to their friends residing in other parts of the Internet.

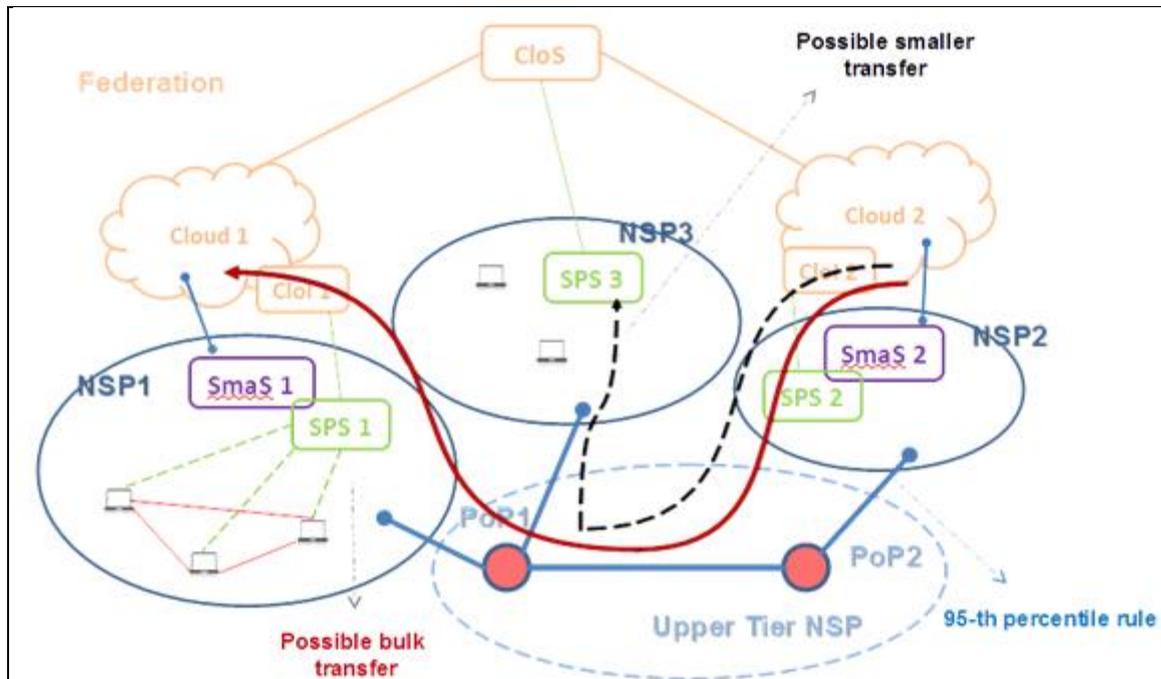


Figure 5-4: The ISP-owned SPS case, i.e. one instance of SPS per ISP.

5.3.2.2 Case B: Presence of SPS instance in each CSL/DC

The second case is the CSP/DC-owned SPS case where there is a single instance of SPS per CSP/DC. From the technical side, the SPS operation and functionalities remain the same as the previous scenario, but there is limited information for the SPS compared to the first scenario. In particular, the CSP cannot obtain the full information for his customers' actions, since he can only have access to application-specific information, i.e. information related to the services that the CSP offers to each customer. Once again though, **both the CSP and the ISP have incentive to collaborate and exchange information in order to achieve mutual benefits**. Therefore the CSP can follow an ISP-friendly policy in the SPS operation, while the ISP can provide extra social information to CSP so that the resulting inter-CSP/DC data transfers are beneficial for both in terms of inter domain cost and user quality of experience. On the other hand, since the SPS is under the control of CSPs the social information can directly become available in the Cloud layer and specifically in CloS.

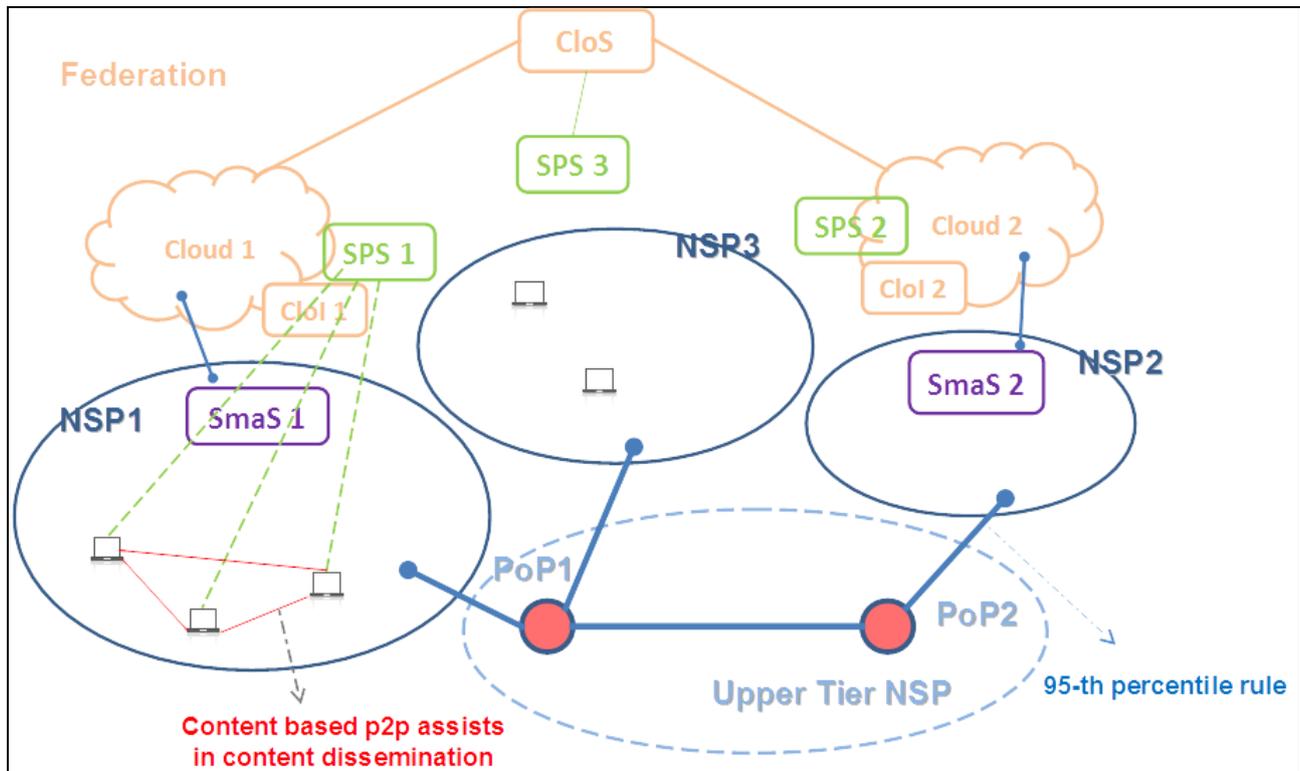


Figure 5-5: The CSP/DC-owned SPS case, i.e. the cloud/datacenter owns the SPS instance.

5.3.3 Tussles identified and resolved by SmartenIT

In practice, the synergy of OFS and EFS mechanisms may inherently resolve many (if not most) of the tussles identified in the individual scenarios, but it also may generate new tussles that arise due to the interaction among stakeholders that were not considered to co-exist in the individual scenarios, e.g. a Cloud Service Provider and an End-user that owns a uNaDa. In this section, we will focus on the tussles that arise because of the synergy of OFS and EFS mechanisms, while we will describe how these could be resolved taking into account a high-level synergy across the mechanisms, DTM and RBH, or ICC and SEConD, as defined in section 5.3.1 and section 5.3.2.

5.3.3.1 Optimal Traffic Destination Selection

The context is similar to that of the tussle described in section 5.1.2.1, with the important difference that the feasible destinations may not be DCs and CSPs, but also the edge caches, i.e. uNaDas that have been deployed by end-users.

For instance, in the case of video content dissemination across multiple network domains, with the support of both centralized facilities as well as distributed mini-caches is a case where multiple destinations can serve the needs of replications and redundancy. The selection of the optimal destination would then be done by the sending entity, CSP/DC or uNaDa, without any knowledge on the underlying network load which may affect both the time within which the data transfer will terminate as well as the underlying network load. It is possible that the selection of the source entity, solely based on local information, is

inefficient for the data transfer and causes extra overhead and congestion to the underlying network links from which the traffic is forwarded to.

A destination selection based on network conditions, economic agreements (e.g. transit or peering) and availability or adjacency (physical or social) of uNaDas could be much more beneficial to the service and to the network, provided that there is accurate information sharing between the cloud, network and end-user layer stakeholders and all three layers' constraints would be taken into account in the decision making. Thus there is contention over the control of the decision of the flow destination and the amount of information made available to the respective stakeholder to make this decision. The goal would be to resolve this tussle in a win-win-win fashion for all involved stakeholders, e.g. end-user, CSP/DC and ISP.

5.3.3.2 Net Neutrality

The differentiated traffic management of (a portion of) the video content traffic, e.g. delay-tolerant video traffic, in a worse than Best Effort fashion, can be beneficial for the ISP, however it deteriorates the performance of this traffic. Thus, the ISP may be tempted to assign this traffic to a lower priority class inside his network, similarly to the way P2P traffic has been traditionally managed by many ISPs. This raises a net neutrality tussle, regarding the extent of traffic management that can be employed by the ISP for the video content transfer, which needs to be resolved by means of proper incentive mechanisms, i.e. an appropriate pricing mechanism that will convince source entities to send their traffic either in different timeslots, e.g. low peak hours, and thus be compensated by a lower price; in fact such a pricing mechanism is in progress in the context of ICC mechanism.

5.3.3.3 Information Asymmetry

Current network and interconnection contracts among the ISPs as well as between ISPs and their client CSPs/DCs do not provide any kind of statistical QoS guarantees; there are solely guarantees regarding maximum possible rate (theoretical), uptime and/or connection recovery. Hence, the CSP/DC traffic is typically handled a la Best Effort, without allowing the CSP/DC to verify whether his traffic could be handled in a better way by its home ISP. Moreover, the inclusion of uNaDas or other edge storage may further burden ISPs, as apart from generating potentially expensive inter-domain traffic, they would also generate a significant increase in intra-domain traffic, i.e. traffic within the administrative domain of an ISP. To avoid OPEX increase for ISPs and allow improved performance for end-users, i.e. the end-customers of CSPs/DCs, the uNaDas could be provided to the end-users by the ISP and thus, be fed with (abstracted) network information, so as to perform prefetching and content distribution. ICC in combination with SEConD could assure an optimal transmission of traffic across administrative domains, along with an optimal distribution within the domain of each ISP.

5.3.3.4 Inter-domain traffic changes

The inclusion of edge storage devices and the support of video content dissemination has been considered to bring traffic changes in the pattern and volume of inter-domain traffic which may have significant impact on one ISP's inter-connection costs. While the users do not care about this, on a large scale this mechanism could affect the peering link load, thus

changing peering ratios to one side's expense. Also if peering links are well provisioned then the uNaDas could serve as a "business stealing" trojan horse: a well-provisioned ISP is actually used for free by his national competitors ISPs via the uNaDas. This is due to the fact that the other ISPs refrain from investing in upgrading their own transit connections and backbone links infrastructure and their customers are still served in high QoS through the well-provisioned ISP's uNaDas. This results in a situation where the ISP that invests most in the market sees his infrastructure being utilized for free by his direct competitors who actually face lower costs due to their lower investments. However, the combination of an EFS solution with DTM or ICC would enable the selection of the inter-domain link or the time epoch in such a way that the video content traffic would be transferred across inter-domain links in a cost-efficient way.

5.3.4 Relation with SmartenIT Business Models

This tussle analysis across the Operator-Focused and End-user Focused Scenario inevitably implies that the tussles identified are relevant for all the SmartenIT business models presented in Chapter 3 of SmartenIT Deliverable D2.4. In particular, the SmartenIT EFS mechanisms serve as a smart network connectivity services' layer. This layer allows for new differentiated network services that could be provisioned under the Network as a Service (NaaS) paradigm. On top of these services, the Federation business model (for instance) can be established among multiple clouds offering e.g. a video service on pan-European scale, using the SmartenIT EFS mechanisms to optimize inter-domain traffic costs. This layer of TM mechanisms is not directly visible to the end users and besides reduction of inter-domain traffic can also lead to content and services being properly managed and pushed close to the user. This functionality is complemented with the synergetic EFS mechanisms, which are most suited to the "smart terminal+service" model plus end-user side ISP services under the Cloudlets model. Hence, complementary SmartenIT mechanisms are simultaneously applicable under and business models jointly addressing the wholesale and retail market for cloud and Internet services, each operating on a different layer, scope, granularity of flows and time, also involving different stakeholders and relationships amongst them. This co-existence creates value in terms of synergies among the mechanisms of the OFS and EFS scenarios as well as additional potential for new value propositions and respective business models, e.g. by means of vertical integration: For instance a large ISP may offer advanced EFS caching, offloading and energy efficiency capabilities to his subscribers combined with managed services and virtual private clouds relying on SmartenIT OFS traffic management for ubiquitous Internet and Cloud service connectivity. This way, substantial competitive advantage can be attained and the customer base of the retail market can be additionally exploited for attracting business customers at the wholesale market, e.g. CSPs or CDNs wishing to reach the ISP's end users via NaaS contracts where SmartenIT OFS functionality is in place.

5.4 Tussles in Cloud Fedearations

The SmartenIT theoretical investigations and models also comprise an area where interesting tussles arise and whose analysis allows a deeper understanding of the aspects that the SmartenIT mechanisms and business models come up against. A prominent case

is the model of cloud federation. Cloud federation has been considered for multiple SmartenIT use cases and mechanisms. Therefore, a brief tussle analysis of the cloud federation issues is of high importance.

5.4.1 Stakeholders

The federation model considers the problem of the formation of an economically sustainable computational resource federation by Cloud Service Providers (CSPs). In this context, a federation policy is specified by the portions of workload transferred from each CSP towards other CSPs and the problem of finding the federation policy that maximizes the total profit (revenue minus cost) of CSPs is solved; finally, the model defines a rule for the fair sharing of the profit that is induced by the federation. Therefore, the stakeholders are the following:

5.4.1.1 Cloud Service Providers (CSPs)

Cloud Service Providers (CSPs) have geographically dispersed servers in order to satisfy client requests through their storage and/or computational resources. Client requests are themselves arising in different locations and the stream of requests has time-varying characteristics. As a result of the time-varying request load of clients, the load at the servers of a CSP is time-varying, and thus the quality of provisioned service (e.g. the job execution delay) is also time dependent and unpredictable. In order to alleviate the temporal variation of request load, a solution would be to invest more in resources (e.g. servers and computational capacity) at the expense of increased costs. A natural means to refrain from this investment is to respond to such load variations by forming cloud federation. Their main interest is to provide high QoS to their customers and CAPEX/OPEX reduction regarding infrastructure deployment, management and operation.

5.4.1.2 Federation

In a cloud federation, CSPs cooperate and pool together their resources in order to improve the QoS of their client requests in a seamless manner. The Federation as a whole can be seen a super-CSP controlling a large amount of resources.

5.4.1.3 End users

End users enjoy the cloud services and wish to receive them at high quality for a reasonable price.

5.4.2 Tussles identified and resolved by SmartenIT

The major tussles that arise in this model are the following:

5.4.2.1 Cooperation versus myopic competition among CSPs and Federation

CSPs need to decide whether they want to join a federation and whether once admitted to it they will indeed cooperate smoothly with the other Federation members, even trying to displace or reject admittance of some of them. Economics indicate that in cooperative environments such games among providers are inevitable and larger and more powerful operators are less eager to cooperate with smaller ones. The reason is that business stealing may occur from their Federation partners who can take advantage of the fact that

their limited infrastructure is less costly and can be dynamically scaled with competitors' resources. Federation policies regarding how jobs should be allocated and managed as well as pricing are the typical ways to resolve these tussles. SmartenIT has successfully proposed such solutions that mitigate this tussle.

Furthermore, the Federation as a whole can be seen as a direct competitor for each CSP. However, the cost savings attained by the Federation, combined with the inability of CSPs to maintain and operate world-wide infrastructure in a cost-efficient way mitigate the cases where CSPs would prefer not to join a Federation coalition and form a cooperative business relation but instead directly compete.

5.4.2.2 Quality discrimination

Cooperative CSPs tend to prioritize their own tasks and provide inferior performance to other CSPs, even when better quality is indeed possible. Once again, the model's policy and pricing approach contributes to providing incentives to the Federation members for proper tasks management.

5.4.2.3 Information Asymmetry

The end user or even a CSP member of the federation cannot deduce whether for instance the delay exhibited over a task execution is a result of the actual load of the cloud resources or this performance has been degraded due to lack of effort by the respective ISP(s). Since there is no way for the task dispatcher to observe the actual effort of the hosting CSP, there is a tussle over the performance delivered and also for the information revealed and mapped to SLAs within the Federation. This tussle could be resolved by means of proper contracts and tamper-proof policies such as the models.

5.4.3 Federation policies and rules

Federation policies and rules are of prominent importance since they determine customer ownership, intermediation, revenue flows and revenue sharing schemes. The actual definition of these policies and rules is a major control tussle for the CSPs. This tussle is orthogonal to the SmartenIT Federation model, which mitigates this threat by setting as objective the total profit maximization and pre-specified policies. Thus, the admittance to the Federation also results in explicit acceptance of these predefined policies and rules.

5.4.4 Relation with SmartenIT Business Models

This model is directly mapped to the Federation business model, presented in Chapter 3 of SmartenIT Deliverable D2.4.[7] It is interesting to note that the business model of Federation can take many forms in the context of services. In particular, a federation may be homogeneous or heterogeneous depending on the nature of its members. The SmartenIT federation model is mapped to homogenous federations, since all the Federation members are CSPs. The federation model could serve as the means to determine the collaboration and information model of such federations.

5.5 Summary of tussle analysis

This section has focused on tussle analysis of the SmartenIT ecosystem. This analysis covers the two SmartenIT scenarios, namely OFS and EFS, the hybrid combination of the two including synergies for combinations of SmartenIT mechanisms, as well as the SmartenIT models. Since an exhaustive analysis of the entire SmartenIT solutions space would not be feasible, we have opted to provide indicative analysis over specific instances for each major SmartenIT solution “type”, namely mechanisms, scenarios and models. This analysis has also been explicitly mapped to the business models that are relevant for SmartenIT and have been presented in the Chapter 3 of the SmartenIT deliverable D2.4. Furthermore, despite the fact that project consciously focuses on Best Effort Internet, the impact of beyond Best Effort Internet connectivity has been briefly discussed where some quality of service assurance can be provided even in inter-domain level. Thus, the tussle analysis of this section exhibits the close relation of SmartenIT theoretical investigations/models, scenarios, use cases, synergies, mechanisms and business models and how inevitably correlated these are.

6 Summary and Conclusions (ALBLF)

Deliverable D2.5 has identified and analysed the most relevant use cases that highlight the efficiency and scope of the traffic management mechanisms proposed in SmartenIT. The use cases involving Cloud Service Providers, Data Center Operators and Network Operators (or ISPs) were elaborated on SmartenIT scenarios resulting from WP1 and refined in WP2 over concrete use cases. They were built on the two complementary business points of view of the ecosystem addressed by those scenarios of the project: the End-user Focused Scenario and the Operator Focused Scenario. A template for the description of use cases was defined to allow a well-organized and structured development and presentation of the different use cases. The template was created to guide the identification of use cases that solve real and concrete problems of incentive based traffic management for overlay networks and cloud-based applications driven by social awareness, QoE awareness, and energy efficiency. For each use case, a value network configuration figure highlights the relationships between stakeholders. Success indicators for achieving the use case goal are presented. Afterwards, for each use case, it has been considered which traffic management mechanisms of SmartenIT are the most appropriate for the use case.

Furthermore, D2.5 has provided final evaluation of the mature specifications of the incentive-based cloud traffic management mechanisms and synergetic solutions completing the results from D2.4 [7]. We maintained the mapping of the TM mechanisms to the two main scenarios of SmartenIT, which were introduced in deliverable D1.2 [3], in order to present in a structured way the mechanism and their potential synergies. Synergy between mechanisms consists in a combination of mechanisms that allows a broader coverage of the SmartenIT playfield and higher gains for the ecosystem than when applying individual mechanism. We have also evaluated aspects of TM mechanisms and synergetic solutions not studied so far in the project. Analysis results of each mechanism or synergetic solution are mapped into a structured template. In the template, the goal of each conducted analysis on mechanisms and synergetic solutions is firstly described. Furthermore, the associated parameters and metrics, and the evaluation framework for each mechanism or synergetic solution are defined, while the evaluation results obtained by means of theoretical evaluations, and simulations are presented. Tussle analysis has been conducted on single mechanisms, as well as on a SmartenIT model and synergetic solution covering the two complementary business points of view of the ecosystem addressed by SmartenIT (end-user focused and Operator Focused Scenario as well as cross scenario) also in lieu of the business models presented in D2.4.

Finally, D2.5 has achieved results of work items from the requested extension period detailed in section 2.2 and thus provided a deeper understanding of the technical, economic and business aspects of the developed solutions.

6.1 Coverage of SmartenIT aspects

Below, we discuss the coverage of the SmartenIT targets and objectives by the specified TM mechanisms and their synergies to address an individual scenario, either OFS or EFS, or both in a cross-scenario manner. To do so, we remind how the SmartenIT playfield is

formed based on SmartenIT key objectives and targets. Figure 6-1 presents the mapping of the SmartenIT architecture on the cloud, network and end-user domain [8]. This topological diagram can be used as an overview of the TM mechanism and synergies with respect to the SmartenIT architecture. As apparent from Figure 6-1, three domains of entities are present:

- **Data Center/Cloud Layer:** This layer comprises data centers and their virtual interconnections over (Best Effort) Internet.
- **(Core and Access) Network Layer:** The core and access layer network contains components in the ISP network and the private networks of data center operators, as well as access network infrastructure.
- **End User Layer:** The end-user layer covers the end-users’ devices. In particular, this layer contains the user’s terminal devices and the home router.

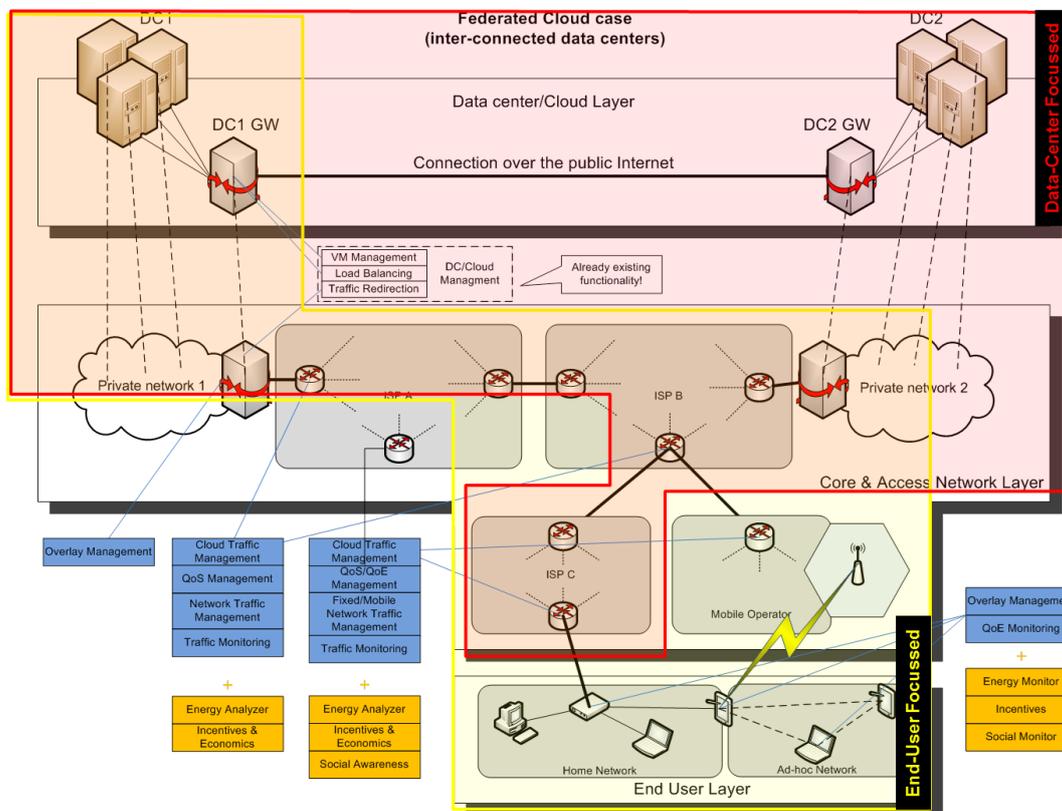


Figure 6-1: Topological view of architecture with added scenario overlay (based on the component map taken from D3.1 [8]).

Furthermore, in order to assess the coverage of the SmartenIT field by the developed TM mechanisms, we employ the categorization methodology developed in Deliverable D2.1 [4]. In order to assess the relevance, applicability and potential benefit for SmartenIT for each of these approaches found in literature, we performed a categorization based on the following criteria (briefly described here – a more extensive description is available in [4]):

- **Cloud layer** (mentioned as application layer in [4]): We distinguish TM mechanisms whose modules or functionalities are executed in centralized cloud/datacenter infrastructures controlled by Cloud Service Providers,
- **Network layer** (mentioned as lower layer in [4]): TM mechanisms whose modules or functionalities involve the core or access part of ISPs,
- **End-user layer** (mentioned as application layer in [4]): We distinguish TM mechanisms whose modules or functionalities involve the devices at the end-users premises, e.g. a mobile application, or the access router (CPE),
- **Inter-domain**: TM mechanisms that are applicable within a single network's administrative domain or it can be generally applied in multiple domains,
- **Collaborative**: TM mechanisms that address the incentives of several entities, either in the same (e.g. cloud federation) or in different layers, so as to reach a goal by means of collaboration,
- **Energy efficiency**: TM mechanisms that consider and/or pursue energy efficiency,
- **Social awareness**: TM mechanisms which employ information derived by OSNs so as to perform some decision making,
- **Resource enhancing**: TM mechanisms that insert extra capacity, in terms of storage, bandwidth or computation, in the considered setup,
- **Long-term time scale**: TM mechanisms whose effect is demonstrable in the course of weeks or months,
- **Incentive-based**: TM mechanisms that pursue incentive-compatibility, rather than being applied compulsorily.

Note that we omit here the “End-user related” criterion, as it is addressed jointly with the “End-user layer” one, while we include two more criteria referring to the “Cloud layer” and “QoE-awareness”. Thus, mapping the TM mechanisms and their synergies, single scenario ones such as DTM++ and RB-HORST++, and cross-scenario ones such as DTM&RB-HORST and ICC&SEConD, we derive the following table (cf. Table 6-1).

Table 6-1: Mapping of TM mechanism and single-scenario synergies to the layers and objectives of the SmartenIT project.

	Cloud layer	Network layer	End-user layer	Inter-domain	Collaborative	Energy efficiency	Social awareness	QoS/QoE awareness	Resource enhancing	Long-term time scale	Incentive-based
DTM		x		x	x					x	x
ICC	x	x		x	x					x	x
MRA	x				x						x
RB-HORST		x	x	x	x		x				x
SEConD		x	x	x	x		x	x	x		x
vINCENT			x		x						x
MONA			x			x					
MUCAPS		x	x		x			x			
DTM++	x	x		x	x					x	x
RB-HORST++		x	x	x	x	x	x	x	x		x
DTM&RB-HORST		x	x	x	x		x			x	x
ICC&SEConD	x	x	x	x	x		x	x	x	x	x

As it can be observed, DTM and ICC also operate in the network layer while collaborating with the cloud layer and both address inter-domain traffic, whilst ICC and more MRA TM mechanisms employ modules that operate in the cloud layer. Additionally, MRA is a collaborative mechanism, which also pursues incentive-compatibility, while DTM and ICC are mechanisms whose impact is observable in the larger time-scales of the wholesale Internet services market, i.e. monthly period to derive 95th percentile values.

On the other hand, all mechanisms addressing the EFS expectedly operate in the end-user layer. Out of the EFS mechanism, RB-HORST, SEConD, and vINCENT are collaborative and incentive-based, while the first two employ social awareness and consider inter-domain traffic. RB-HORST, SEConD and MUCAPS, are QoE-aware while MONA is the one addressing energy efficiency at the end-user layer. RB-HORST provides WiFi access for mobile users and supports mobility in content delivery by prefetching content to potential locations based on overlay and social prediction. Since content is delivered by an overlay formed by shared home routers the capacity of cache resources scales with the number of active users and has the ability to deal with flash crowds.

Regarding the two single-scenario synergetic solutions, we observe that they “inherit” the characteristics of the mechanisms (or their modules) that they employ. Thus, DTM++ operates in both the cloud and network layers, it addresses inter-domain traffic and is

applicable in a long time-scale, inheriting the attributes of DTM and ICC. On the other hand, RB-HORST++ practically, addresses all criteria/objectives except the fact that it does not operate in the cloud layer and in large time scales. Thus, we conclude that indeed the integration of the mechanisms into synergetic solutions leads to a broader coverage of the SmartenIT objectives and targets.

Finally, the two cross-scenario synergies, i.e. DTM-RB-HORST and ICC-SECoND, apparently address most of the targets of the SmartenIT project, which is reasonable taking into account that the OFS and EFS approaches are complementary in the Internet services market; thus the operation of the one affects the operation of the other only in an indirect way, since different TM issues are resolved at the traffic aggregate and the per-flow layer, while their combination and synchronization may result in even higher benefits for all stakeholders as discussed in section 5.3.3. The quantitative evaluation of these approaches remains as future work.

6.2 Lessons learnt

We summarize the lessons learnt on the evaluation of a set of TM mechanisms and synergetic solutions addressing the OFS and EFS, and highlight the most important metrics and their key parameters. The examination of various use cases addressing EFS and OFS reveals the high degree of applicability, high potential for adoption, and the diverse and positive impact of SmartenIT solutions. Those use cases show benefits of incentive-compatibility brought by SmartenIT mechanisms. Incentives for the stakeholders are to minimize costs in terms of inter-connection charges due to traffic generated by cloud services and applications for ISPs, operating cost in terms of connectivity charges and energy cost for Cloud Service Providers/Data Center Operators, and connectivity cost (WiFi vs. mobile) for end-users. Moreover these use case cover benefits brought by QoE- and social awareness, as well as energy efficiency. In addition, from the end-user perspective, the use cases focus on user's needs and comprise a large set of services enabled for end users.

- The use cases related to the Operator Focused Scenario (OFS) are mainly driven on one hand by collaboration of cloud providers (possibly in the form of a federation) aiming at better performance, more revenue and load balancing by means of resource sharing between different cloud players, and on the other hand by ISPs offering the underlying traffic management as a service to the operators of the clouds. The incentive of the stakeholders of the OFS use cases is to minimize their costs in terms of inter-connection charges due to traffic generated by cloud services and applications for ISPs, operating cost in terms of connectivity charges and energy cost for Cloud Service Providers/Data Center Operators, and improve (indirectly) QoE of their customers for Cloud Service Providers and ISPs. Different ways to achieve these goals rely on inter-cloud communication, cloud federation and data replication/migration, while taking into account the often competing interests of Cloud Service Providers, Cloud Operators and the ISPs, especially in terms of transit traffic and the associated cost.
- The use cases related to the End-user Focused Scenario (EFS) are mainly driven by a collaborative traffic management approach with a direct involvement of the end-users

and their resources. The goal of the EFS use cases is to provide improved QoE and energy efficiency for the end-user, by intelligent placement and movement of content and services in an energy and socially aware manner, whilst the interest of the ISP on traffic management and respective costs is specially considered. In particular, ISPs benefits from saving network resources, also achieving a reduction of their inter domain traffic. These EFS use cases focus on user's needs and comprise a large set of services enabled for end users.

Next, D2.5 presented the finalization of the TM mechanisms and of their analysis of evaluation (further to the material in D2.4 [7]) as well as synergetic solutions not studied so far. We summarized the simulation framework employed for their evaluation and the evaluation results derived by the relevant simulations.

Completion of TM mechanisms evaluation addressing the aspects of the OFS provides the major outcomes as follows:

- Simulation experiments show that DTM is able to compensate the traffic and distribute it among links as desired to optimize traffic cost for volume based tariff and for 95th percentile tariff. In the 95th percentile tariff case, however the algorithm is sensitive to the traffic profiles. Considerations to provide DTM scalability, reliability and security have also been discussed.
- Further evaluation of the ICC (Inter Cloud Communication) mechanism for the incentive-compatible cost-efficient use of ISP's transit link(s) show that ICC using statistics to predict traffic patterns also performs well (max +/-10% deviation from target goal), often achieving higher discount than the one original sought
- Studying the interdependency of multiple heterogeneous resources (such as CPU, RAM, disk space, and bandwidth) in a cloud (federation) has proven that many fairness metrics for resource allocation, which were proposed in literature, are insufficient and thereby certify the relevance of the greediness metric (which does not depend on interdependency of multiple heterogeneous resources), and MRA mechanism proposed by SmartenIT mechanism.
- DTM++ specification and experiment document a synergetic solution integrating DTM with ICC. It combines the so-called traffic "shift in space" of DTM and traffic "shift in time" of ICC to achieve further optimization of traffic distribution across multiple transit links while delaying delay-tolerant traffic when transit links are heavily loaded, so as to ultimately achieve even lower transit charges for the ISP than DTM alone.
- Economics model of the federated environment of CSPs are defined, considering both the case where CSPs act cooperatively and non-cooperatively. Cloud federation increases revenues from QoS-based pricing on customers and reduces the energy consumption cost, thus the profit of federated CSP is increased. Moreover Cloud federation improves the global QoS in the federated environment, i.e. the average delay of served requests is decreased.

Completion of TM mechanisms evaluation addressing the aspects of the EFS provides the major outcomes as follows:

- RB-HORST (Replicating Balanced- tracker and Home Router Sharing based on Trust) evaluation and refinement of caching strategy including flash crowd scenario are focused on performance when demand exhibits temporal dynamics. LRU caching policy is simple but effective under non-stationary demand. Furthermore the results on WiFi offloading in an urban environment show, that 66% of the connections can be offloaded on the average, if only 10% of WiFi access points are shared, assuming a sending range of 50m.
- Further evaluation of SEConD (Socially-aware mechanism for Efficient Content Delivery), as reported in the present deliverable, shows that, in the experimental set-up considered, SEConD can attain a significant reduction (even up to ~87%) of the total inter-AS traffic compared to inter AS traffic generated when applying the client-server paradigm in this set-up, thus advocating that it is a promising TM mechanism. Evaluation of its joint application with RB-HORST shows the potential of QoS-based user-assisted video delivery as a means to boost users' QoE.
- Further evaluation of vINCENT (virtual incentive) extension for RB-HORST addressing mobile offloading from cellular communication to WiFi, leveraging social data to do so in an ISP friendly way. Results address the asymmetries between rural areas vs. city areas of an offloading scheme to derive fair incentives for all users. Despite different user densities (e.g.), the proposed scheme is proven fair to all users.
- Further evaluation of MONA (Mobile Network Assistant), consider traffic on the cell interface and WiFi access points with realistic data rates. It shows that aggregation of traffic or deferring transmissions until a more energy efficient connectivity option is available may reduce power consumption: between 34% and 85% of the otherwise consumed energy can be saved; Joint evaluation with RB-HORST++ investigates the energy consumption for mobile video streaming sessions and shows that minimum energy is consumed for connections offloaded to WiFi, which can only be achieved for high data rates that are not available in streets with current WiFi access point deployment
- Further evaluations of MUCAPS (MULTI-Criteria Application endPoint Selection) have shown that for a reliable evaluation, it is necessary to introduce contextual elements such as at least the access type and capabilities of the UEP and the network conditions, or simply user expectations on QoE. In addition, the shortest AS-paths are not necessarily the best ones in terms of delay, resources and ISP costs: these three criteria may even be conflicting and a safe way for efficient layer cooperation is to consider them jointly.

The most important metrics employed for the evaluation of each TM mechanism are summarized both OFS and EFS. TM mechanisms that address OFS mainly are assessed based on metrics expressing inter-domain traffic and transit cost reduction. On the other hand, TM mechanism addressing EFS demonstrate a larger variety including both metrics associated to caching mechanisms, metrics related to energy consumption/energy efficiency, and metrics reflecting QoE perceived by the end users. It has also been observed that the set of key parameters of TM mechanisms demonstrates an even larger variety, as different mechanisms employ different methods and specifically designed algorithms to achieve the targets of SmartenIT per use case (as defined in section 3.3).

For instance, DTM/DTM++ and ICC, OFS TM mechanisms that address inter-cloud communication focusing on the network level, consider traffic patterns, cost functions and time slotting, while MRA considers CPU stress and workload type, as it focuses on the cloud level. On the other hand, EFS mechanisms employ parameters such a caching strategy and user interests concerning caching of content, sharing probability and cache contribution/participation regarding resource sharing, device type, bitrate and network availability in the case of energy efficiency.

Furthermore, we presented and discussed the tussle analysis of SmartenIT ecosystem. In SmartenIT, we designed mechanisms that address incentives of the various involved stakeholders, so as to avoid potential tussles among them. We focused on two types of incentives: monetary incentives, e.g., revenues for providing a specific service, and performance incentives, e.g., enjoying high(er) QoE or experiencing low(er) congestion.

A basic tussle analysis of the Inter-Cloud Communication and Exploiting Content Locally use cases was conducted. During the tussle analysis, DTM++ where ICC has been integrated, and RB-HORST++ including components by SEConD, vINCENT and MONA have been considered. The major stakeholders are identified, as well as their respective interests, actions and concerns. The most prominent tussles are overviewed, their relationship to the business models that are applicable in the SmartenIT case was highlighted, and the way these are handled by the SmartenIT mechanisms is briefly discussed. The conclusion from this analysis is that the SmartenIT Traffic Engineering mechanisms for the Inter-Cloud Communication and Exploiting Content Locality use cases do resolve (or at least mitigate) the respective tussles by means of adopting the “design-for-tussle” principles in the mechanism design without favoring any kind of stakeholder in the expense of the other. Moreover they provide proper interfaces and incentive schemes for stakeholders to communicate their interests and resolve or mitigate potential tussles and conflicts in an incentive-compatible way. For instance, to avoid increase of inter-connection charges, the use of uNaDas within the domain of an ISP can be restricted to serve only users within that AS, resulting in minimizing egress traffic towards other domains, and thus, avoiding a potential increase of associated costs.

Then, we have argued that the various TM mechanisms address different objectives of the SmartenIT landscape, i.e. the coverage of the cloud layer, the network layer and the end-user layer, inter-domain cost mitigation, collaborative optimization, energy efficiency, social awareness, QoE improvement, resource usage enhancement, long-term time scale and finally incentive based. SmartenIT TM mechanisms are mainly complementary and in fewer cases overlapping, while the synergetic solutions addressing only EFS or OFS, as well as cross scenario solutions addressing both EFS and OFS, achieve to tackle a broader area of the SmartenIT landscape.

Since DTM and RB-HORST can be run independently, yet complementarily, they cover both the EFS and OFS scenarios. RB-HORST reduces the amount of inter-domain traffic and takes load of data-centers by caching and social-aware prefetching. The remaining inter-domain traffic is optimized by DTM such that the costs for operators are minimized. The objectives met in this synergy are inter-domain cost mitigation, energy efficiency, and social awareness. QoE is enhanced by reducing latency with local caches.

The ICC-SEConD synergy covers both the EFS and OFS scenarios. The two mechanisms can be run in a coordinated and complementary fashion. In particular, SEConD can be used to mitigate the fact that ICC is mainly focused on inter-provider data transfers and may not have enough information to decide on the timing of the bulk data transfers in an efficient and optimal in terms of user QoE. The social awareness logic of SEConD can be used inject into ICC social awareness and QoE aspects by means of exploiting social information to predict demand and thus make educated decisions regarding the scheduling of the data transfers. The objectives met in this synergy are inter-domain cost mitigation, end user QoE enhancement, and social awareness.

Overall, the SmartenIT use cases and Traffic Management mechanisms are in line with recent evolutions in business models and networking technologies and their adoption can lead to significant benefits for the respective stakeholders even in practical cases involving commercial networks in an incentive-compatible manner, thus fulfilling the objectives and challenges of WP2 and of the project as a whole. These results will be completed by the demonstration of some real test cases that will be carried out in WP4.

6.3 Future work

While WP2 has completed its work and met its objectives, several additional interesting research directions on the rich set of TM mechanisms can still be identified and pursued. In particular, we envision that the SmartenIT mechanisms could be further extended as follows:

6.3.1 Further evaluations of TM mechanisms

Further evaluations of TM mechanisms developed in SmartenIT can be carried out. In particular

- Quantitative assessment by means of evaluations for cross-scenario synergies, as described in section 5.3.
- The modeling and evaluation of cache efficiency including dynamic popularity can be further consolidated. Results obtained on caching efficiency in the Appendix 11.5 and in sections 4.2.1 and 4.2.5 can be used for techno-economic analysis of caching infrastructures at ISPs, CSPs and data center providers. Last but not least, the interworking of caching in different domains under different administration remains a difficult task waiting for solution approaches that are viable for all involved parties and enable better inter-domain optimization.
- The developed power models for mobile connectivity can be used for developing an algorithm scheduling the user generated traffic in future versions of MONA, considering QoE requirements of the end user, while minimizing the power consumption of the mobile device. This algorithm is also to consider the network energy consumption and location of content in its decision process, reducing the energy cost generated in the network.

- ICC is inspired by the increasing need to manage inter-cloud and back-office Web traffic and a performance evaluation using real traces has been made. More traces and traffic models will be used for additional evaluations in the near future, while larger traces when available will allow a deeper understanding of the fine-tuning and learning capabilities of the ICC network layer. Assessing the cloud layer of the ICC mechanism, taking into account both the network load and the energy cost of the IT infrastructure of the datacenters, will enable further optimizations on the management of inter-domain traffic for the specific case of cloud services.
- The evaluation of the MRA mechanism in simulations and experiments and comparison to other metrics. For these experiments, a CloudSim extension is developed and mimics the observed resource dependencies.
- MUCAPS and ICC/DTM++ will benefit from further investigation on cross layer cooperation between users, network providers and cloud/application providers and by the development of cross-layer decision making algorithms.

6.3.2 Applicability of SmartenIT solutions in Future Internet and 5G

6.3.2.1 Applicability of SmartenIT mechanisms in the scope of OFS

As service differentiation and premium connectivity services over the Internet gain acceptance and momentum, the SmartenIT mechanisms could be extended accordingly so as to take advantage of such premium connections in the medium term future. In particular, regarding the OFS TM mechanisms:

a) Evaluating the potential of ICC to operate with multiple delay classes and its potential adoption over inter-domain network paths of assured quality, as well as coupling the ICC pricing model with that of DiffServ for multiple quality layers comprises additional promising future research steps. This also fits nicely in the context of Future Internet and assured quality interconnection in order to overcome the inefficiencies of Best Effort Internet and make Internet the preferred way of delivering content and services, as opposed to competing extranet solutions that currently fill-in this gap, such as IPX and its related business initiatives

b) DTM could be extended to use multiple tunnel connections, each pertaining to a different service for the same source-destination pair. DTM rationale could optimize the usage of these tunnels while the tunnel features could result in differentiated performance plus bypassing the BGP single source-destination route. Clearly those extensions could also be transferred to DTM++ which is the superposition of ICC and DTM.

5G is considered as a technology that enables machine-to-machine (M2M), human-to-human (H2H) and human-to-machine (H2M) communication without limit (infinite bandwidth, infinite capacity). Many 5G services and Future Internet Applications accessed through 5G will require huge data processing in DC/Cloud. Then, DTM and DTM++ may help in managing traffic crossing the borders of mobile operator domains, limiting transfer costs. In addition, as 5G network will rely on SDN technology, DTM can be used for managing traffic transfer in different delivery regions of the mobile network where more capacity is present. DTM uses only very simple features of SDN technology, one of the

possible extensions is usage of MPLS tunnels managed via SDN. The whole physical network infrastructure can be shared by a few virtual mobile operators. DTM can also apply in 5G, where virtualization will lead to the conception and deployment of network slice (virtual 5G network which is managed by a single virtual operator). The DTM can offer virtual DA routers and they can be deployed in separate slices belonging to different virtual operators. Given that the 5G architecture is still undefined it is impossible to derive a solution but the applicability of these SmartenIT mechanisms remains to be investigated in the future.

6.3.2.2 *Applicability of SmartenIT mechanisms in the scope of EFS*

From the EFS point of view, the MONA approach is expected to be applicable with the development of the future internet and of 5G, especially because 5G will integrate seamless multi radio access technologies. In addition to QoE criteria, energy efficiency as well as device power consumption will stay fundamental to the 5G ecosystem. As MONA supports vertical as well as horizontal handovers, it will have potential for adapting to the future mobile network evolution. Furthermore, the MUCAPS approach is compatible with the network evolution toward SDN based architecture. Further investigation can aim to allow coherent coexistence of application layer traffic optimization (ALTO) with software defined networks (SDN).

7 SMART Objectives Addressed

Throughout this document, nine SmartenIT SMART objectives defined in section B1.1.2.4 of the SmartenIT DoW [1] have been partly addressed. Namely, two overall objectives as reported in Table 7-1, and four specific ones as presented in Table 7-2.

Overall Objectives 2 and 3 are addressed in D2.5. They are defined in DoW as following:

- Objective 2** SmartenIT will develop theory for new pricing and business models, applicable in example use cases, combined with the design and prototyping of appropriate traffic management mechanism that can be combined with today's and IETF's proposed relevant communication protocols.
- Objective 3** SmartenIT will investigate economically, QoE, and energy-wise such models in theory by simulations to guide the respective prototyping work in terms of a successful, viable, and efficient architectural integration, framework and mechanisms engineering, and its subsequent performance evaluation.

The work reported in this deliverable covers to a significant extent both of the above objectives. In section 3, eleven relevant use cases have been identified in the playfield of SmartenIT. A well structured template for use case description has been built to prove that the solutions proposed in SmartenIT are applicable to solve real and concrete problems of incentive based traffic management for overlay networks and cloud-based applications driven by social awareness, QoE awareness and energy efficiency. In majority, the studied use cases belongs to one the two scenarios addressed by the project, namely the Operator Focused Scenario and the End-user Focused Scenario. They cover the whole ecosystem in two complementary points of view (or focuses): the retail (or end-user) and the wholesale (or operator) business focuses. Those use cases motivate the deployment of SmartenIT technology, which benefits to the opposite sides of operators and end-user. The last use case though covers the field of cross scenario, thus proving that SmartenIT solutions can simultaneously be beneficial to both side of operators and end-user perspectives, e.g. SEConD has been proven to be beneficial to both ISPs and end-users (see section 4.2.2).

Moreover, the business models covered in D2.4 are also considered in the tussle analysis section (see section 5) and their relation with use cases and mechanisms have been explicitly documents. Tussle analysis has been performed for single OFS and EFS framework, respectively in section 5.1 and in section 5.2. The synergy of OFS and EFS mechanisms may inherently resolve many (if not most) of the tussles identified in the individual scenarios as discussed in section 5.3.

Furthermore the work done in D2.5 provides a SmartenIT pricing classification framework (in section 11.6), detailing where the different layers and granularities of pricing mechanisms related to SmartenIT operate and how they work together. This work is completed by an example of a SmartenIT pricing mechanism provided in the case of ICC (in section 11.6.2).

Model and pricing of Cloud federation was studied in section 11.3. We considered two options for a federation based on the CSPs' behavior, the strong and the weak federation.

In the strong federation the CSPs are fully cooperative and each of them follows a jobs' outsourcing strategy that optimizes the global profit of the federation. On the other hand, in the weak federation each CSP is non-cooperative following the outsourcing strategy that maximizes its individual profit. Both strong and weak federation policies aim to an increased individual profit for all federated CSPs compared to their profit in the stand-alone operation.

Next, in section 4, we completed the evaluation of TM mechanisms developed to address the aspects of the OFS and the EFS initiated in the deliverable D2.4. It describes the parameters, metrics, simulation framework employed for evaluation and the evaluation results derived by the simulation are provided. Moreover, some TM mechanisms are combined and their performance evaluated for both EFS and OFS in order to demonstrate the complementarity that exists among certain mechanisms. Then, in section 6.1, we analyzed the coverage of the SmartenIT playfield (as defined by the SmartenIT objectives and key targets described in[1], by the TM mechanisms specified and evaluated in this deliverable, and we investigated potential cross-scenario synergies that constitute a unified SmartenIT approach to address selected UCs of interest.

Table 7-1: Overall SmartenIT SMART objectives addressed [1].

Objective No.	Specific	Measurable	Achievable	Relevant	Timely
		Deliverable Number			Mile Stone Number
O2	Theory design for traffic management mechanisms	D2.4	Design, simulation	Advanced	MS2.4
	Prototyping of traffic management mechanisms	D2.4	Design, simulation, implementation	Complex	MS2.4
O3	Simulative investigations	D2.4	Analysis, simulation, evaluation	Complex	MS2.2

Table 7-2: Specific SmartenIT SMART objectives addressed; excerpt from the set of Tables of [1] with all SMART objectives of the project.

Objective ID	Specific	Measurable	Achievable	Relevant	Timely
		Metric			Project Month
O1.1	How to align real ISP networks, while optimizing overlay service/cloud requirements?	Savings in inter-domain traffic (inMbit/s) and possible energy savings due to optimized management mechanisms	Design, simulation T2.2	Major output of relevance for provider and in turn users	M36

Objective ID	Specific	Measurable	Achievable	Relevant	Timely
		Metric			Project Month
O1.3	Which incentive schemes will be needed (for application overlays and clouds) to adapt to the existing physical network structure? For what optimization criteria (revenue, energy efficiency) can this be attained more effectively?	Number of identified incentive schemes and their cost reduction in terms of money, traffic, and energy footprint	Design, simulation T2.3, T2.4	Extremely relevant output of relevance for providers and users	M24
O1.4	Are decentralized traffic management schemes superior to traditional schemes; if yes, why? If not, what is the associated efficiency loss?	Number of identified and tested scenarios for the comparison of the different traffic management schemes	Design, simulation T2.5	Highly relevant output of relevance for providers	M24
O2.1	Which key design goals (incentive compatibility, user expectations, etc.) are met by the designed mechanisms in terms of their performance?	Number of identified design goals, number of met design goals (evaluated separately for each mechanism)	Design, simulation, prototyping T2.5	Major output of entire project	M36

Deliverable D2.5 contributes to the answering of four specific theoretic questions:

Objective O1.1: How to align real ISP networks, while optimizing overlay service/cloud requirements?

The ISP-friendliness of the TM mechanism was a prerequisite during the design and development phases. Thus, the alignment of the interests of a real ISP network with the TM mechanisms can be achieved due to the transit cost minimization/reduction attained by means of most of the OFS TM mechanisms, i.e. DTM or ICC and their combination, namely DTM++, and selected EFS TM mechanisms as well, such as SEConD.

Objective O1.3: Which incentive schemes will be needed (for application overlays and clouds) to adapt to the existing physical network structure? For what optimization criteria (revenue, energy efficiency) can this be attained more effectively?

Incentives for resource sharing and load/VM migration among clouds (with or without federation) or load sharing among home routers have been investigated by dedicated models in section 4. The optimization criteria served are minimization of transit charges for the ISP, cloud metrics, QoS/QoE and energy metrics. Moreover, an indicative *tussle analysis* of the SmartenIT ecosystem has been performed and exhibits how conflicts of interest of stakeholders in SmartenIT scenarios and use cases can be mitigated by means of the incentive compatible SmartenIT TM mechanisms according also to the project's theoretical investigations and analysis of models. These new tussles can be possibly mitigated by means of the careful parametrization of the scope and operations of each of the SmartenIT mechanisms constituting the synergy, applying appropriate business models that allow fair competition throughout the respective value chain segments. This relation is also evident in the SmartenIT pricing classification framework, provided as Appendix (in section 11.6).

Objective O1.4: Are decentralized traffic management schemes superior to traditional schemes; if yes, why? If not, what is the associated efficiency loss?

Decentralized TM mechanisms have been compared against traditional (centralized) ones and have been found to be more efficient in certain cases (see section 4). For example, new caching strategies for small caches have been compared against traditional caching schemes in section 11.5, while SEConD was compared with SocialTube [36] and traditional client-server dissemination in section 4.2.2. In the latter case, the evaluation of SEConD in ASes of varying size (i.e. number of subscribers) proved that the contribution of the P2P overlay is bigger in larger ASes while the contribution of the centralized node is higher in smaller ASes, where the distributed components and resource of the mechanisms do not suffice.

Objective O2.1: Which key design goals (incentive compatibility, user expectations, etc.) are met by the designed mechanisms in terms of their performance?

For each TM mechanism a set of key metrics has been defined which addressed the key design goals for all involved stakeholders including end-users. Important examples of key metrics related to the end-user are: QoE, e.g. in terms of video bit-rate or video stalling time, accessibility, energy consumption in user device, improvement of the performance of non-shiftable traffic etc. A list of 5 key metrics per TM mechanism, which has also been considered by WP4 for the test-bed trials, has been summarized in Table 4-5.

8 References

- [1] The SmartenIT project: Grant Agreement for STREP: Annex I 'Description of Work (DoW)'; 2012.
- [2] The SmartenIT project: Deliverable D1.1: Report on Stakeholders Characterization and Traffic Characteristics; April 2013.
- [3] The SmartenIT project: Deliverable D1.2: Report on Cloud Service Classifications and Scenarios; October 2013.
- [4] The SmartenIT project: Deliverable D2.1: Report on Overview of Overlay Traffic Management Solutions; April 2013.
- [5] The SmartenIT project: Deliverable D2.2: Report on Definitions of Traffic Management Mechanisms and Initial Evaluation Results; October 2013.
- [6] The Smarten IT project: Deliverable D2.3: Report on Definition of Use-Cases and Parameters (Initial Version); April 2014.
- [7] The SmartenIT project: Deliverable D2.4: Report on Final specifications of Traffic Management Mechanisms and Evaluation Results; October 2014.
- [8] The Smarten IT project: Deliverable D3.1: Report on Initial System Architecture; April 2013.
- [9] The SmartenIT project: Deliverable D3.2: Technologies, Implementation Framework and Initial Prototype (Initial Version), April 2014.
- [10] The SmartenIT project: Deliverable D3.3: Final Report on System Architecture; October 2014.
- [11] Alistair Cockburn: Writing Effective Use Cases, Addison-Wesley Longman, 2000.
- [12] J. J. Ramos-Munoz, J. Prados-Garzon, P. Ameigeiras, J. Navarro-Ortiz, and J. M. Lopez-Soler, "Characteristics of Mobile YouTube Traffic," IEEE Wireless Communications, vol. 21, no. 1, pp. 18–25, 2014.
- [13] B. Wang, J. Kurose, P. Shenoy, and D. Towsley, "Multimedia Streaming via TCP: An Analytic Performance Study," ACM Transactions on Multimedia Computing, Communications and Applications, vol. 4, no. 2, p. 16:116:22, 2008.
- [14] T. Hoßfeld, R. Schatz, M. Seufert, M. Hirth, T. Zinner, and P. Tran-Gia, "Quantification of YouTube QoE via Crowdsourcing," in Proceedings of the IEEE International Workshop on Multimedia Quality of Experience- Modeling, Evaluation, and Directions (MQoE 2011), Dana Point, CA, USA, 2011.
- [15] T. Hoßfeld, S. Egger, R. Schatz, M. Fiedler, K. Masuch, and C. Lorentzen, "Initial Delay vs. Interruptions: Between the Devil and the Deep Blue Sea," in Proceedings of the 4th International Workshop on Quality of Multimedia Experience (QoMEX 2012), Yarra Valley, Australia, 2012.

- [16] K. Panitzek, I. Schweizer, T. Bönning, G. Seipel, and M. Mühlhäuser, "First Responder Communication in Urban Environments," *International Journal of Mobile Network Design and Innovation*, vol. 4, no. 2, pp. 109–118, 2012.
- [17] "OpenStreetMap." <http://www.openstreetmap.org/export#map=14/49.8788/8.6628>
- [18] Traverso, Stefano, Mohamed Ahmed, Michele Garetto, Paolo Giaccone, Emilio Leonardi, and Saverio Niccolini. "Temporal locality in today's content caching: why it matters and how to model it." *ACM SIGCOMM Computer Communication Review* 43, no. 5 (2013): 5-12.
- [19] L. Breslau et al., Web caching and Zipf-like distributions: Evidence and implications, *Proc. IEEE Infocom* (1999)
- [20] R. Bolla et al., A measurement study supporting P2P file-sharing community models. Special Issue on Content Distribution Infrastructures for Community Networks, *Computer Networks* 53 (2009) 485-500
- [21] M. Cha et al., I tube, you tube, everybody tubes: Analyzing the world's largest user generated content video system, *Internet measurement conference IMC'07*, San Diego, USA (2007)
- [22] J. Charzinski, Traffic properties, client side cacheability and CDN usage of popular web sites, *Proc. 15th MMB conference*, Essen, Germany, Springer LNCS 5987 (2010) 182-194
- [23] R. Fielding, M. Nottingham and J. Reschke, Hypertext transfer protocol HTTP/1.1: Caching, IETF standardization, RFC 7234 (2014)
- [24] C. Fricker, P. Robert and J. Roberts, A versatile and accurate approximation for LRU cache performance, *IEEE Proc. 24th International Teletraffic Congress*, Kraków, Poland (2012)
- [25] R.G. Garroppo et al., The greening potential of content delivery in residential community networks *Computer Networks*, <http://www.sciencedirect.com/science/journal/13891286/73/supp/C>", (2014) 256-267
- [26] G. Hasslinger, Efficiency of caching and content delivery in broadband access networks, Chapter 4 in *Advanced Content Delivery, Streaming, and Cloud Services*, M. Pathan et al. (Ed.), Wiley (2014) 71-90
- [27] G. Hasslinger and K. Ntougias: Evaluation of Caching Strategies based on Access Statistics on Past Requests, *Proc. MMB & DFT 2014*, Bamberg, Germany, Springer LNCS 8376 (2014) 120-135
- [28] Internet Engineering Task Force, CDN interconnection working group <tools.ietf.org/wg/cdni/charters>
- [29] D. Lee et al., LRFU: A spectrum of policies that subsumes the least recently used and least frequently used policies, *IEEE Transactions on Computers* 50/12 (2001) 1352-1361
- [30] N. Megiddo and S. Modha, Outperforming LRU with an adaptive replacement cache algorithm, *IEEE Computer* (Apr. 2004) 4-11

-
- [31] T. Qiu et al., Modeling channel popularity dynamics in a large IPTV system, Proc. 11th ACM SIGMETRICS, Seattle, WA, USA (2009)
- [32] R.K. Sitaraman et al., Overlay networks: An Akamai perspective, Chapter 16 in Advanced Content Delivery, Streaming, and Cloud Services, M. Pathan et al. (Ed.), Wiley (2014) 307-328
- [33] S. Zhao, D. Stutzbach and R. Rejaie, Characterizing files in the modern Gnutella network: A measurement study, SPIE/ACM Proc. Multimedia Computing and Networking (2006)
- [34] Constantiou, I.D., Courcoubetis, C.A.: "Information asymmetry models in the Internet connectivity market". In Proceedings of 4th Internet Economics Workshop, 2001.
- [35] ETICS: Economics and Technologies for Inter-Carrier Services, www.ict-etics.eu.
- [36] Haiying Shen; Ze Li; Yuhua Lin; Jin Li, "SocialTube: P2P-Assisted Video Sharing in Online Social Networks," Parallel and Distributed Systems, IEEE Transactions on , vol.25, no.9, pp.2428,2440, Sept. 2014
- [37] S. Barré, C. Paasch, and O. Bonaventure, "MultiPath TCP: From Theory to Practice," in IFIP Networking, 2011.
- [38] C. Pluntke, L. Eggert, and N. Kiukkonen, "Saving Mobile Device Energy with Multipath TCP," in ACM MobiArch, 2011.
- [39] T. A. Le, C. S. Hong, M. A. Razzaque, S. Lee, and H. Jung, "ecMTCP: An Energy-Aware Congestion Control Algorithm for Multipath TCP," IEEE Communications Letters, vol. 16, no. 2, pp. 275–277, 2012.
- [40] S. Chen, Z. Yuan, and G.-M. Muntean, "An Energy-Aware Multipath-TCP-Based Content Delivery Scheme in Heterogeneous Wireless Networks," in IEEE WCNC, 2013.
- [41] Y.-C. Chen, Y.-S. Lim, R. J. Gibbens, E. M. Nahum, R. Khalili, and D. Towsley, "A Measurement-based Study of MultiPath TCP Performance over Wireless Networks," in IEEE IMC, 2013.
- [42] Y.-S. Lim, Y.-C. Chen, E. M. Nahum, D. Towsley, and R. J. Gibbens, "Improving Energy Efficiency of MPTCP for Mobile Devices," in ACM CoNEXT, 2014.
- [43] Hampel, A. Rana, and T. Klein, "Seamless TCP Mobility Using Lightweight MPTCP Proxy," in ACM MOBIWAC, 2013.
- [44] C. Paasch, G. Detal, F. Duchene, C. Raiciu, and O. Bonaventure, "Exploring Mobile/WiFi Handover with Multipath TCP," in ACM CellNet, 2012.
- [45] S. Chen, Z. Yuan, and G.-M. Muntean, "A Traffic Burstiness-based Offload Scheme for Energy Efficiency Deliveries in Heterogeneous Wireless Networks," in IEEE GLOBECOM-CTEMD, 2013.
- [46] S. Rado, "Optimizing the Energy Efficiency of Multipath TCP for Mobile Devices," Bachelor's Thesis, Technische Universität Darmstadt, 2014.

- [47] C. Gross, F. Kaup, D. Stingl, B. Richerzhagen, D. Hausheer, and R. Steinmetz, “EnerSim: An Energy Consumption Model for Large- Scale Overlay Simulators,” in IEEE LCN, 2013
- [48] R. D. Cook, “Detection of Influential Observation in Linear Regression,” *Technometrics*, vol. 19, no. 1, pp. 15-18, 1977.
- [49] Clark, David D., et al. "Tussle in cyberspace: defining tomorrow's internet." *ACM SIGCOMM Computer Communication Review*. Vol. 32. No. 4. ACM, 2002
- [50] D. Clark, S. Bauer, K. Claffy, A. Dhamdhere, B. Huffaker, W. Lehr, and M. Luckie, “Measurement and Analysis of Internet Interconnection and Congestion,” in Telecommunications Policy Research Conference (TPRC), Sep 2014
- [51] <http://aws.amazon.com/ec2/pricing/>
- [52] <https://cloud.google.com/storage/#pricing>
- [53] Briscoe, B. and Rudkin, S.: “Commercial Models for IP Quality of Service Interconnect”. In BTTJ Special Edition on IP Quality of Service, 23(2) (Apr 2005).
- [54] Varian, H. R.: “Estimating the demand for bandwidth”. Available at: <http://people.ischool.berkeley.edu/~hal/Papers/wtp/wtp.html>
- [55] <http://www.statslab.cam.ac.uk/~frank/eb.html>
- [56] http://www.juniper.net/documentation/en_US/junos13.2/topics/reference/statement-hierarchy/policer-edit-firewall-hierarchy.html
- [57] Horizon2020 SSICLOPS, available at www.ssiclops.eu
- [58] FP7 FELIX, available at <http://www.ict-felix.eu/>
- [59] FED4FIRE available at <http://www.fed4fire.eu/>
- [60] Ali Ghodsi, Matei Zaharia, Benjamin Hindman, Andy Konwinski, Scott Shenker, and Ion Stoica. Dominant Resource Fairness: Fair Allocation of Multiple Resource Types. 8th USENIX Conference on Networked Systems Design and Implementation
- [61] Danny Dolev, Dror G. Feitelson, Joseph Y. Halpern, Raz Kupferman, and Nathan Linial. No Justified Complaints: On Fair Sharing of Multiple Resources. 3rd Innovations in Theoretical Computer Science Conference (ITCS'12), pp 68–75, Cambridge, MA, USA, January 2012.
- [62] draft-ietf-alto-multi-cost-00: “Multi-Cost ALTO”, (work in progress) S. Randriamasy and W. Roome and N. Schwan, May 22nd 2015, <http://tools.ietf.org/html/draft-randriamasy-alto-multi-cost-10.txt>
- [63] <https://ec.europa.eu/digital-agenda/en/net-neutrality-challenges>
- [64] <http://www.capacitymagazine.com/Article/3397337/Telenor-joins-the-Chicago-agreement.html>
- [65] Patrick Poullie: Tussles for Edge Network Caching; ITU Workshop on "Future Trust and Knowledge Infrastructure", Geneva, Switzerland, Geneva, Switzerland, April 2015

9 Abbreviations (ALL)

3GPP	3rd Generation Partnership Project
AEP	Application Endpoint
ALTO	Application-Layer Traffic Optimization
AP	Access Point
AS	Autonomous System
BE	Best Effort
BGP	Border Gateway Protocol
BR	Border Router
CAGR	Compound Annual Growth Rate
CAPEX	Capital Expenditure
CDN	Content Delivery/Distribution Network
CDNI	Content Distribution Network Interconnection
CPE	Customer Premises Equipment
CSP	Cloud Service Provider
DC	Data Center
DCO	Data Center Operator
DSCP	Differentiated Services Code Point
DoW	Description of Work
DTM	Dynamic Traffic Management
EFS	End-user Focused Scenario
GRE	Generic Routing Encapsulation
HTTP	Hyper Text Transfer Protocol
IaaS	Infrastructure-as-a-Service
IANA	Internet Assigned Numbers Authority
ICC	Inter-Cloud Communication
ICMP	Internet Control Message Protocol
IETF	Internet Engineering Task Force
IP	Internet Protocol
ISP	Internet Service Provider
IT	Information Technology
MONA	Mobile Network Assistant
MPLS	Multi-Protocol Label Switching
MRA	Multi Resource Allocation
MUCAPS	Multi-Criteria Application Endpoint Selection

OFS	Operator Focused Scenario
OPEX	Operational Expenditures
OSN	Online Social Network
OTT	Over-The-Top
P2P	Peer-to-Peer
Pol	Point of Interconnect
PoP	Point-of-Presence
QoE	Quality.Of-Experience
QoS	Quality-of-Service
RB-HORST	Replicating Balanced Tracker and Home Router Sharing based on Trust
RPO	Recovery Point Object
RTT	Round Trip Time
SaaS	Software-as-a-Service
SDN	Software Defined Network
SEConD	Socially-aware Efficient Content Delivery
SLA	Service Level Agreement
SmartenIT	Socially-aware Management of New Overlay Application Traffic with Energy Efficiency in the Internet
TCP	Transport Control Protocol
TM	Traffic Management
TTL	Time-To-Live
UDP	User Datagram Protocol
UNaDa	User-owned Nano Data Center
vINCENT	Virtual Incentives
VM	Virtual Machine
VNC	Value Network Configuration
VoD	Video-on-demand
WiFi	Wireless Radio Transmission
WP	Work Package

10 Acknowledgements

Besides the authors, this deliverable was made possible due to the large and open help of the WP2 team of the SmartenIT team. Many thanks to all of them! Special thanks go to the internal reviewers Manos Dramitinos (AUEB), Gerhard Hasslinger (DT), Paolo Cruschelli (IRT) and George Stamoulis (AUEB) for their detailed revision and valuable comments.

11 Appendices

11.1 Appendix A: DTM++ detailed specification

This section presents updates to release 3.0 (Deliverable D2.4 section 13) necessary to implement DTM++. The description encompasses the realization of ICC functionality on hardware routers and integration with DTM. The ICC implementation is based on hierarchical policers.

For the ICC operation the Border Gateway (BG) routers need to distinguish a DSCP field and perform traffic measurement for traffic policing. Traffic measurements performed for the S-Box operation are almost the same like in the case of pure DTM; we have only changed the traffic measurement points: The tunnels in the receiving domain have to be terminated on BG routers because these routers have to be able to recognize the DSCP marking done by the sending cloud/DC which are hidden in tunnels. The tunnel traffic has to be measured before the decapsulation, so in comparison to the previous DTM implementations the measurement point is moved from DA-C router to BG-C-1 and BG-C2 routers (Figure 11 1). In this figure one can see that the lines representing tunnels are terminated on the BG routers. In the present implementation approach SDN controllers are used only for DTM and they operate on the remote domains. First we summarize the changes required for the operation of DTM in the scope of the DTM++ framework (the DTM can operate separately without ICC in this configuration):

- tunnels have to be terminated on BG-C-1 and BG-C-2 routers
- inbound tunnel traffic measurements have to be performed on BG-C-1 and BG-C-2 routers,
- QoS/QoE Analyzer acquires information about the tunnel traffic from BG-C-1 and BG-C-2 routers via SNMP. In Figure 11-2 one can see that the S-Box gets the information about total traffic X_1 on link 1 and the tunnel traffic Z_1 on the same link.

Now we present the ICC operation when hierarchical policers are used. We assume that the delay tolerant traffic is TCP traffic. The hierarchical policer will discard the delay tolerant traffic when it exceeds the predefined limit and we expect that sources of this traffic will decrease the transmission rate due to the TCP rate control mechanism. We present our consideration only for link 1, it is identical for link 2.

The traffic from the remote cloud/DC comes to link 1 in a tunnel. The traffic inside the tunnel is distinguished by the DSCP field. The delay tolerant traffic is marked as BE and the delay sensitive traffic is marked as EF. We do not use the tunnel DSC marking for DTM++. The tunnel DSCP can have any marking and we do not respect this marking because on the path toward ISP-C it can be changed by any operator. The DSCP marking schema is presented in Figure 11-3. The BG router marks all traffic going via link 1 with DSCP=EF. This procedure does not change the tunnel internal traffic marking, the delay tolerant and delay sensitive traffic stay with the DSCP marked by the sending cloud/DC (Figure 11-3). After the marking procedure, the tunnel DSCP becomes EF and also the background traffic is marked with EF. In the next step the tunnel traffic is decapsulated and after this procedure the background traffic stays marked EF but now we can see that the delay tolerant traffic is marked with BE and the delay sensitive traffic is marked as EF

(Figure 11-3). The whole traffic is directed to the hierarchical policer, which is a part of the traffic filter represented in Figure 11-3. This policer discards the BE traffic which exceeds the predefined bandwidth limit. Only BE traffic is policed, the EF traffic goes without limit, in accordance to our aim of influencing only delay tolerant traffic.

In Figure 11-4 we present the logical functionalities of the traffic filter. After the tunnel decapsulation, the packets DSCP value is either EF or BE. These packets are sent to the firewall filter which is composed of a chain of matching rules: DSCP=AF, DSCP=any. Below each matching rule one can see actions. These actions are terminating actions, meaning that if a match rule is true no other rule in the chain is triggered. The terminating action is the default behavior but we can redirect the traffic to the next match rule after performing the action. This is done in case of DSCP=EF after policing and remarking. The match rule DSCP=AF is used to extract the EF traffic which is less than predefined bandwidth limit λ . In our case, the AF traffic represents the amount of traffic which exceeds the limit λ . The EF traffic, reaching match rule DSCP=any, is always less or equal to limit λ . In case of the last match rule there is used a hierarchical policer. The hierarchical policer introduces two classes of traffic: premium and aggregate. Premium traffic is recognized by DSCP field marked with EF. Aggregate traffic means all traffic (including premium).

In general one can put separate limits for premium and aggregate traffic. The premium traffic is not policed if it does not exceed predefined limit for this traffic. When it exceeds limit it is discarded. When the aggregate traffic (whole traffic) is higher than the limit for this type of traffic, the traffic different from premium (EF) is discarded, leaving premium unaffected.

For our purposes we use the same limit λ for premium and aggregate traffic. It discards the premium traffic only when it exceeds the limit λ . In our case, EF traffic coming to this hierarchical policer never exceeds this limit (fig.11.4), so all EF traffic is transferred. The second condition imposed by $aggregate \leq \lambda$ allows the sum of EF and BE traffic to be less or equal than limit λ . So if aggregate is greater than λ , then only BE traffic is discarded.

As a whole, the described traffic filter does not limit EF traffic at all but it may limit BE traffic trying to prevent the whole traffic from exceeding the λ value.

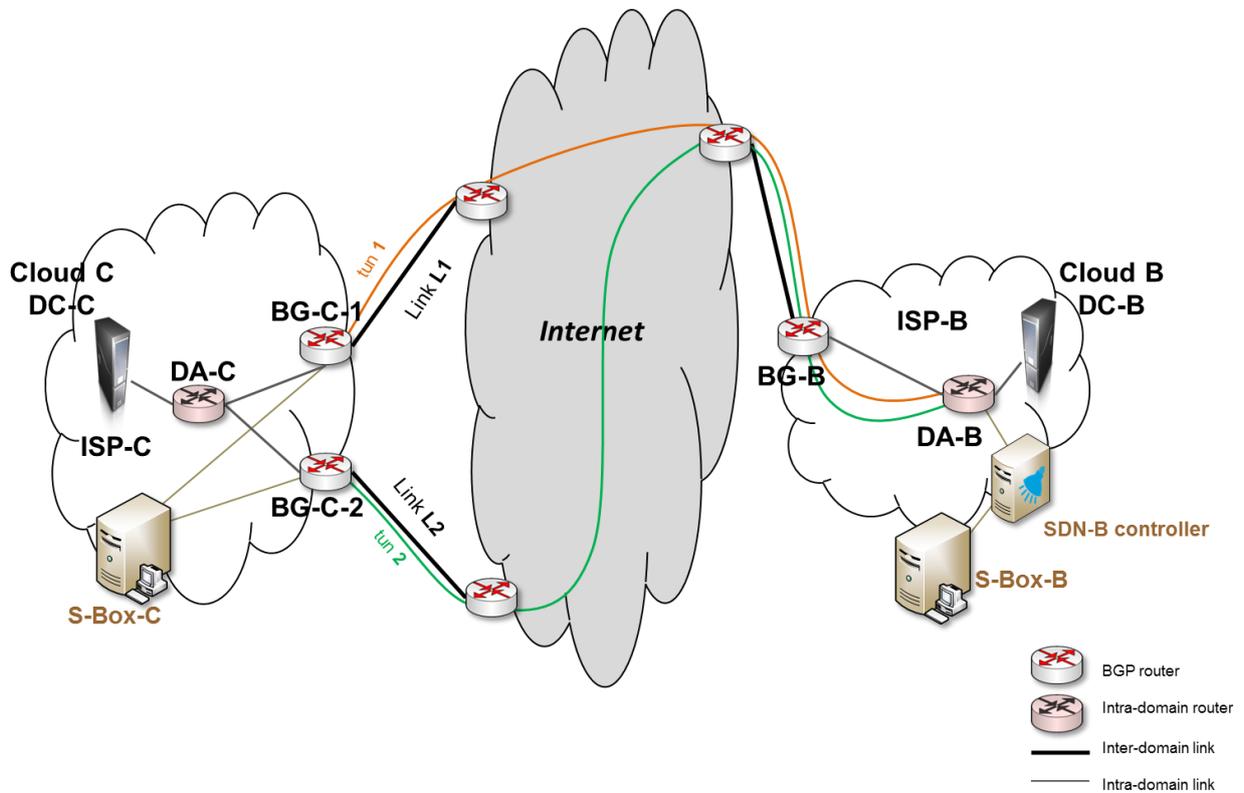


Figure 11-1: The network configuration for the DTM++ operation.

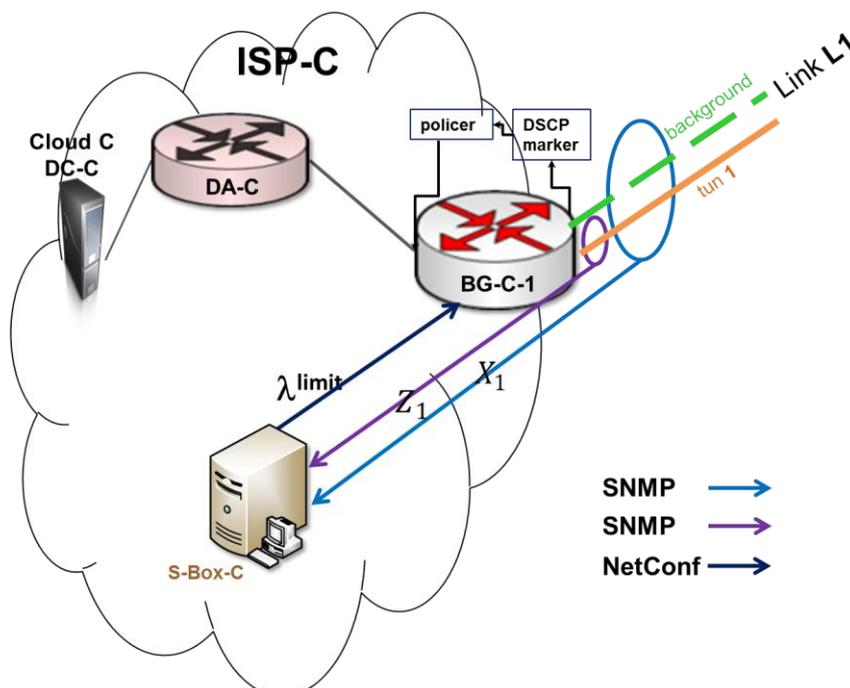


Figure 11-2: Information flow between the BG router and the S-Box; the same applies to all the BG routers in the ISP-C domain.

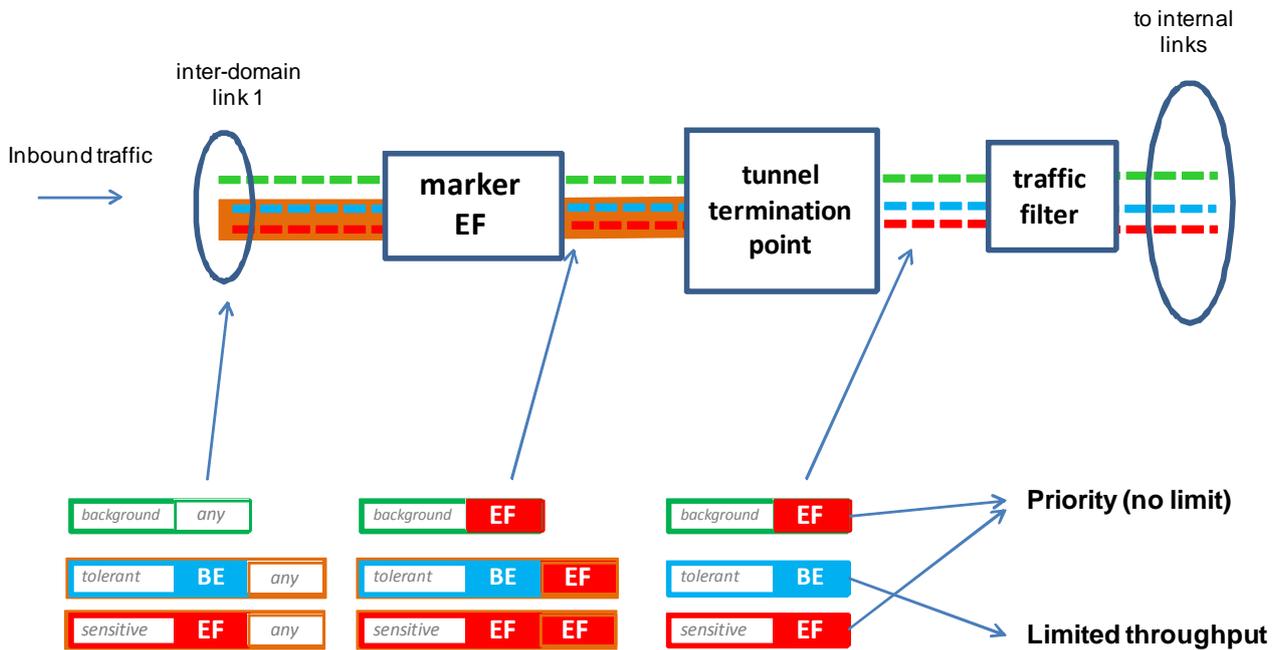


Figure 11-3 The DSCP marking procedures performed by the BG routers.

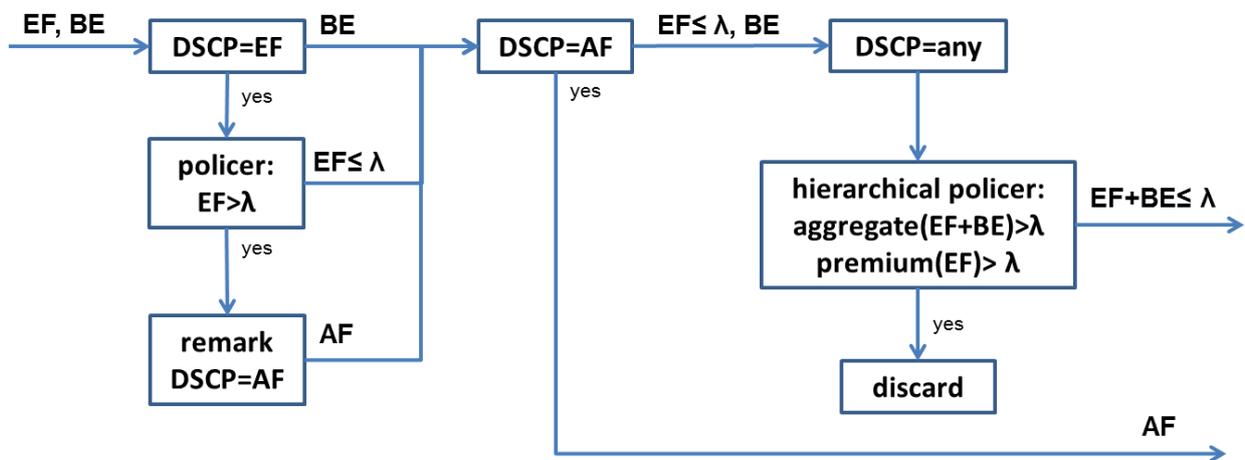


Figure 11-4 Logical operation of the traffic filter from fig.11-3

Economic Analyzer modification

This modification is not an integral part of ICC. It can be used also by pure DTM. In green and purple we presented the changes in the pseudo-code in comparison to release 3.0. The green color refers to changes required only by the ICC extension in comparison to pure DTM. The ICC extension requires modifications marked by both green and purple color. This modification takes into account the fact that it can happen that before the end of billing period the number of 95th percentile samples which are below the supposed limit may be 95% of total number of samples to be collected during billing period. In such a case, we neither need to compensate traffic in DTM nor discard traffic in ICC.

...

```

int currentSampleNumber = 0;
int numberOfSamplesBelowR[2];
reset(numberOfSamplesBelowR);
int flagSamplesBelowR[2];
reset(flagSamplesBelowR);
int onICC;
. . .
currentReportNumberEA++;
if(currentReportNumberEA==reportNumberPerSampleEA) {
    ...
    if(flagSamplesBelowR[link]==0 && linkTraffic[link] <= rVector[link]){
        numberOfSamplesBelowR[link]++;
        if(numberOfSamplesBelowR[link] == sample95Percentile-1){
            flagSamplesBelowR[link]=1;
            updateFlagSamplesBelowR(flagSamplesBelowR);
        }
    }
    ...
currentSampleNumber=0;
reset(numberOfSamplesBelowR);
reset(flagSamplesBelowR);
reset(linkTraffic95Samples);
. . .

```

Network Traffic Manager modification

Every time that the `flagSamplesBelowR` is updated by the Economic Analyzer, the Network Traffic Manager (NTM) sends by NetConf a command deactivating the filter on a link specified by a flag `flagSamplesBelowR`. Also the compensation vector has to be changed in order to redirect traffic to the proper tunnel. The NetConf communication between S-Box and BG router is presented in fig. 11-2.

Below we present the part of pseudo code exemplifying NTM behavior:

```

//new variable indicating compensation switch-off, must be defined when S-box is
activated
int onICC;
int flagSamplesBelowRActivated[2];
reset(flagSamplesBelowRActivated)
. . .
currentReportNumberDTM++;
if(flagSamplesBelowRActivated[1]==0 && flagSamplesBelowR[1]==1)
{
    cVector(1)=100*(rVector(1)+rVector(2));
    cVector(2)=-100*(rVector(1)+rVector(2));
    flagSamplesBelowRActivated[1]=1;
    send(cVector, dstIspId);
    if(onICC)sendNetConf(1,deactivateHierahicalFilter)
        // first parameter in method above indicates link number
}
if(flagSamplesBelowRActivated[2]==0 && flagSamplesBelowR[2]==1)
{

```

```

cVector(1)=-100*(rVector(1)+rVector(2));
cVector(2)=+100*(rVector(1)-rVector(2));
flagSamplesBelowRActivated[2]=1;
send(cVector, dstIspId);
if(onICC)sendNetConf(2,deactivateHierahicalFilter)
    // first parameter in method above indicates link number
}
if(flagSamplesBelowRActivated[1]==0 && flagSamplesBelowRActivated[2]==0){
    . . .
}
else
    if(if(flagSamplesBelowRActivated[1]==1 && flagSamplesBelowRActivated[1]==1)
        if(currentReportNumberDTM % periodicCompensationNumber==0){
            cVector[1]=-cVector[1];
            cVector[2]=-cVector[2];
            send(cVector, dstIspId);
        }
}

```

When a new billing period starts, all previously deactivated filters must be activated:

```

If(onICC){
    if(flagSamplesBelowRActivated[1]==1)
        sendNetConf(1,activateHierahicalFilter);
        // first parameter in method above indicates link number
    if(flagSamplesBelowRActivated[2]==1)
        sendNetConf(2,activateHierahicalFilter);
        // first parameter in method above indicates link number
}

```

The modifications introduced till now may be used by pure DTM without ICC. They can improve DTM operation, that before the end of billing period the number of 95th percentile samples which are below the supposed limit may be 95% of total number of samples to be collected during billing period. In such a case there is no need in traffic compensation because we have obtained reference vector value on particular link and we can transfer whole traffic via this link. Also these changes can work in integrated version of DTM++ (DTM-ICC) to attain even better performance.

Below we present the DTM modifications which are strictly related to the ICC integration. When the billing period expires a new value of the reference vector is sent to the remote S-Box. The expiration of the billing period indicates also the moment when a new value of `bandwidthLimit` and `burstSizeLimit` are sent to the BG routers. These limits have to be applied to a hierarchical policer and policer performing remarking.

New variables have to be defined in NTM:

```

double toleranceICC[2];//default {1.0, 1.0}
long bandwidthLimit[2];
long burstSizeLimit[2];
reset(bandwidthLimit);
reset(burstSizeLimit);

```

When a new reference has been received from ECA, the NTM assigns values to `bandwidthLimit` and `burstSizeLimit`:

```
if (onICC) {  
    bandwidthLimit[1]=rVector[1]/300*toleranceICC[1];  
    bandwidthLimit[2]=rVector[2]/300*toleranceICC[2];  
    burstSizeLimit[1]=625000;  
    burstSizeLimit[1]=625000;  
}
```

BG router configuration using NetConf

Next these values are reconfigured on the BG routers using NetConf communication:

```
sendNetConf(1,setPolicerNewLimits(bandwidthLimit[1],burstSizeLimit[1]));  
sendNetConf(2,setPolicerNewLimits(bandwidthLimit[2],burstSizeLimit[2]));  
    // first parameter in method above indicates link number
```

11.2 Appendix B: Model for QoE of Mobile Video Streaming

To assess the QoE of mobile video streaming in simulation frameworks, a model is needed, which estimates the QoE of mobile video sessions given the bandwidth of the video transmission. The worst quality degradation of video streaming is stalling [14], i.e., playback interruption because of insufficient downloaded video data. The authors [14] found that users tolerate at most one stalling event of up to three seconds length for good QoE. In our work, a simplified QoE model is used inspired by the work in [13]. The authors used discrete-time Markov models for an analytic performance evaluation of video streaming over TCP. They found that a good streaming performance, which results in a low probability of stalling, can be achieved if the network throughput is roughly twice the video bit rate when allowing a few seconds of initial delay. [15] showed that the impact of initial delays on QoE is not severe, as users are already used to them and tolerate them. Therefore, our simplified QoE model only considers the received throughput of the video streaming connection:

$$QoE = \begin{cases} \text{good, if throughput} \leq 2 \cdot \text{video bit rate} \\ \text{bad, otherwise} \end{cases}$$

To derive the bit rate of videos streamed by mobile devices we use the results from [12] where the video formats in mobile networks were characterized by analyzing 2000 videos streamed from the video on demand platform YouTube.

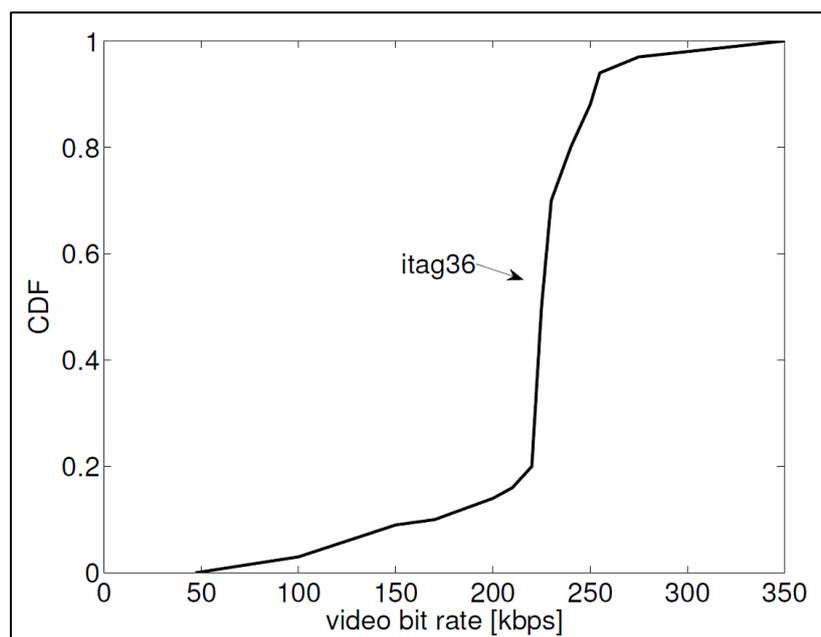


Figure 11-5: Bit rate of YouTube videos in *itag36* format [12].

The authors find that the format *itag36* is used in 80% of the streams.

Figure 11-5 shows the cumulative distribution of video bit rates for mobile videos in *itag36*. The majority of the videos have a bit rate between 220 and 250 kbps.

11.3 Appendix C: Cloud federation- Model and Pricing

In our approach, we investigate the cloud federation as service delegation. First, we model each Cloud Service Provider (CSP) (as an M/M/1 queueing system) and the federation and two CSPs, and we define the functions that determine the net benefit of a CSP, i.e. a pricing function that each CSP uses to charge its clients, and the cost from energy consumption at servers. Then, we propose a federation policy among CSPs as the transfer of a portion of jobs' requests from one CSP to others in order to be served through their server infrastructure. We consider two options for a federation based on the CSPs' behavior, the strong and the weak federation. In the strong federation the CSPs are fully cooperative and each of them follows a jobs' outsourcing strategy that optimizes the global profit of the federation. On the other hand, in the weak federation each CSP is non-cooperative following the outsourcing strategy that maximizes its individual profit. Note that both strong and weak federation policies aim to an increased individual profit for all federated CSPs compared to their profit in the stand-alone operation.

11.3.1 Cloud Service Provider Modeling

We model each CSP as an M/M/1 queueing system. Therefore for a CSP i , we assume that the job requests arrive according to a Poisson process of rate λ_i (jobs/sec). The size of the job in terms of the number of operations it entails is random. We assume that this size follows the exponential distribution with mean number of operations per job L (flops). Let C_i denote the total computational capacity of servers for CSP in (flops/sec). Hence, the average service rate (in jobs/sec) for CSP i is $\mu_i = C_i/L$. Finally, the service time of a job is exponentially distributed with mean $1/\mu_i$. We use as metric for the QoS estimation that a CSP offers to its customers the average delay (queueing and execution) of each job in the M/M/1 system. According to Little's Law for M/M/1 systems the average delay d_i is:

$$d_i = \frac{1}{\mu_i - \lambda_i}$$

Considering the infrastructure of each CSP as a single server we estimate its power consumption. The power consumption of a single server is linearly increasing in its utilization factor $\rho_i = \lambda_i/\mu_i$. In particular, the power consumed is the sum of idle power consumption and dynamic power consumption. The idle power $W_{0,i}$, is the power consumed when the server is powered on and does not serve any request. The dynamic power consumption depends on the utilization ρ_i . If we denote by $W_{1,i}$ the power of the server when it is fully utilized (namely, at $\rho_i = 1$), then the range of dynamic energy consumption is $[0, W_{1,i} - W_{0,i}]$. The total power consumption of the server as function of ρ_i is $W_i(\rho_i) = W_{0,i} + (W_{1,i} - W_{0,i}) \rho_i$. Given the estimated power consumption (Watts) and a price q_i that the CSP should pay to its electricity provider per Watt · sec, then the cost of energy consumption per unit of time is:

$$E_i = W_i q_i.$$

In our approach, each CSP charges its clients based on the offered QoS namely the average delay. The pricing function $p_i(\cdot)$ of a CSP i should be decreasing in the

average delay d_i of its clients. Furthermore, it should also be convex, because a marginal change of delay is perceived more by the client for smaller values of the delay, and hence the pricing should reflect that. A pricing function definition that satisfies the above conditions is

$$p_i(d_i) = x_i e^{-d_i/d^*},$$

where x_i denotes the price per job that CSP charges his customers for offering the service in the best possible level of QoS (for simplicity of the formula, this is taken as $d_i \rightarrow 0$ although this is clearly impossible), while d^* is a parameter that specifies the sensitivity of the price to QoS degradation. The revenue rate, in monetary units per unit of time for CSP i is

$$R_i = \lambda_i p(d_i).$$

Finally, considering both the revenue (cash inflow rate) and the energy cost (cash outflow rate), the profit rate for CSP i is:

$$P_i = R_i - E_i.$$

11.3.2 Federation Policies

11.3.2.1 Model

Our federation model enables the transfer of a portion of the incoming requests from a CSP to other CSPs within the federation. In order to demonstrate our approach, we fix our attention here to the case of two CSPs. For each CSP i , we define a variable α_i (with $0 < \alpha_i < 1$) that denotes the portion of the requests from clients of CSP i that are transferred to the other CSP. Consequently, a federation policy is specified by a pair (α_1, α_2) , where $\alpha_1 \lambda_1$ denotes the arrival rate of requests that will be transferred from CSP₁ to CSP₂ and $\alpha_2 \lambda_2$ the arrival rate of outsourced requests towards the opposite direction. We assume that the requests transferred from one CSP to the other experience a fixed average delay D . This models the delay introduced by the transfer of requests over Internet links between servers of the two CSPs. We assume that the jobs in both CSPs have the same mean size in flop requirements L , since all the requests belong to the same class of jobs. Figure 11-6 depicts the federation scenario for two CSPs.

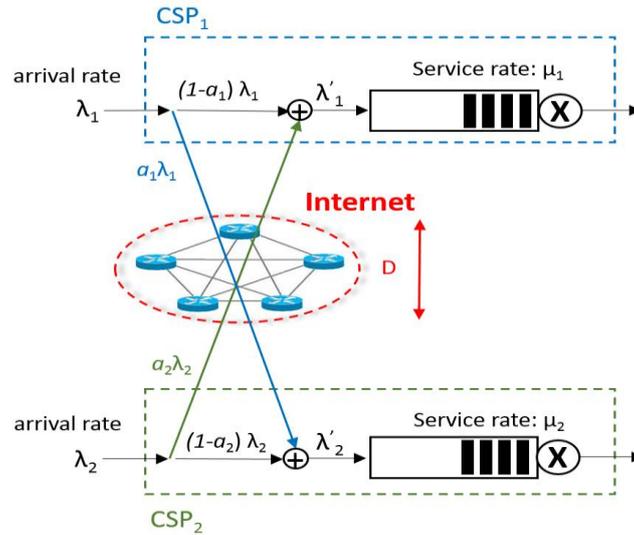


Figure 11-6: Federation scenario for two CSP, each modeled through an M/M/1 queue. The amount of request traffic that is transferred to the other CSP undergoes a fixed average delay D .

Under the federated operation, the input in two CSP queues depends on the values of α_1 and α_2 . The ultimate arrival rates in CSP₁ and CSP₂ queues are given by $\lambda'_1(\alpha_1, \alpha_2) = (1 - \alpha_1)\lambda_1 + \alpha_2\lambda_2$ and $\lambda'_2(\alpha_1, \alpha_2) = (1 - \alpha_2)\lambda_2 + \alpha_1\lambda_1$ respectively. Consequently, the average delay d_i of requests that are served by the queue of CSP i , is:

$$d_i = \frac{1}{\mu_i - \lambda'_i(\alpha_1, \alpha_2)}$$

Part of the arriving requests from each CSP's clients is served by that CSP's own infrastructure, while another part is served by the infrastructure of the other federated CSP. Therefore, the average delay experienced by the clients of each CSP depends on the average delays at both CSPs' queues, $d_1(a_1, a_2)$ and $d_2(a_1, a_2)$. The average $T_i(a_1, a_2)$ experienced by clients of CSP $i, i = 1, 2$, are as follows:

$$T_1(a_1, a_2) = (1 - a_1)d_1(a_1, a_2) + a_1[d_2(a_1, a_2) + D]$$

$$T_2(a_1, a_2) = (1 - a_2)d_2(a_1, a_2) + a_2[d_1(a_1, a_2) + D]$$

There is significant difference between $d_i(\cdot)$ and $T_i(\cdot)$. While $d_i(\cdot)$ denotes the average delay of any job served by the queue of CSP i regardless of whether it originated from clients of CSP₁ or CSP₂, $T_i(\cdot)$ denotes the average delay of jobs originated from clients of CSP i , regardless of the server they are actually served.

We now revisit the definitions of revenue and energy cost. In the presence of federation, we have re-define the function presented in section 1.1 in order to be applicable in the federation. Thus, the power consumption is

$$W_i(a_1, a_2) = W_{0,i} + (W_{1,i} - W_{0,i})\rho_i$$

where $\rho_i = \frac{\lambda'_i(a_1, a_2)}{\mu_i}$, since as explained above $\lambda'_i(a_1, a_2)$ denotes the total incoming requests rate at the queue of CSP i . Therefore, the energy cost per unit of time is

$$E_i(a_1, a_2) = W_i(a_1, a_2) q_i.$$

Since the pricing is based on delay $T_i(\cdot)$ and not on $d_i(\cdot)$ the pricing function is $p_i(a_1, a_2) = x_i e^{-T_i(a_1, a_2)/d^*}$, and thus the revenue rate is

$$R_i(a_1, a_2) = \lambda_i p_i(a_1, a_2).$$

Finally, the profit is

$$P_i(a_1, a_2) = R_i(a_1, a_2) - E_i(a_1, a_2).$$

11.3.2.2 Strong Federation – Cooperative CSPs: Profits and Compensation

In strong federation the CSPs who participate are fully cooperative. Therefore, the federated CSPs jointly decide on the best possible outsourcing policy that is beneficial for both of them. The determination of the optimal federation strategy reduces to solving an optimization problem. The output of this problem is the optimal pair (a_1^*, a_2^*) of the portions of the request traffic at the input of each CSP queue that are routed to the other CSP, such that the total profit of federated CSPs is maximized. The optimization problem is as follows:

$$\begin{aligned} & \max_{a_1, a_2} [P_1(a_1, a_2) + P_2(a_1, a_2)] \\ & s. t. \quad 0 \leq a_i \leq 1, \quad i = 1, 2 \\ & \quad \lambda'_i(a_1, a_2) < \mu_i, \quad i = 1, 2 \end{aligned}$$

The second constraint is due to stability in the queues of each CSP, so that the rate of the stream of incoming requests does not exceed the service rate of the CSP. If that constraint were not included in the formulation, the delay would grow unbounded.

For a given $D > 0$, there exists a unique pair that maximizes the objective above, and for this optimal pair it should be $a_1^* \cdot a_2^* = 0$. Therefore, we always have unilateral service delegation, i.e. at most one of the two CSPs transfers a portion of its request load to the other. This is because the optimal solution essentially entails an optimal load balancing, for which unilateral shift of load suffices. However, if $D = 0$, there exist in general multiple optimal solution pairs (a_1^*, a_2^*) since CSPs will be exchanging loads between them at no cost D . In particular, the solution can be succinctly described as is a pair $(z(a_2^*), a_2)$, where $z(\cdot)$ is an increasing function.

The solution of the optimization problem leads to a total profit $P_{tot}(a_1^*, a_2^*) = P_1(a_1^*, a_2^*) + P_2(a_1^*, a_2^*)$ that may exceed or be equal to the corresponding total profit of CSPs if these were not involved in a federation. The latter is attained for $(a_1, a_2) = (0, 0)$ and thus can be written as $P_{tot}(0, 0) = P_1(0, 0) + P_2(0, 0)$. If the federation is beneficial for CSPs as a whole, the issue arises how to share the profits incurred by the federation. By incurred profit we mean the difference $P_{tot}(a_1^*, a_2^*) - P_{tot}(0, 0)$.

It should be noted that the extra workload increases the energy consumption cost, due to the higher utilization and thus higher power consumption of its infrastructure. Therefore, the CSP to whom load is delegated by the other has reduced profits and may be unwilling to conform to the federation, unless some rule is applied for compensating it for these losses. Since the total profits of the federation exceeds that for the standalone case, the CSP that delegates part of its workload definitely has higher profit than before. This CSP should compensate the other for loss in profit and reach an agreement for the sharing of the additional profit that satisfies both of them. A cooperative sharing policy that serves the above objective is one where each CSP gets at least the profit it had in the no-federation case, while the extra profit generated from federation is shared according to some proportionality rule. If this rule concerns the served request load, then CSP i gets profit. Thus, the payoff that CSP i eventually obtains is given by

$$\frac{\lambda'_i(a_1^*, a_2^*)}{\lambda_1 + \lambda_2} (P_{tot}(a_1^*, a_2^*) - P_{tot}(0,0)) + P_i(0,0)$$

where the second term represents the profit of CSP in the standalone operation, while the first term is the share of the extra profit induced by the federation that is given to CSP i . This is an **incentive compatible profit-sharing rule**; that is, both CSPs have the incentive to participate in the cooperative federation knowing that this rule is applied.

11.3.2.3 Weak Federation – Non-cooperative CSPs

In weak federation the CSPs who participate are non-cooperative, meaning that each CSP aims to maximize its individual profit rather than the global profit of the federation. However, there still exists a cooperation in the sense that both CSPs are able to outsource part of the stream of their incoming requests to the other. The difference from the strong federation is that each CSP i independently decides on its own outsourcing policy (i.e. a_i^*) that maximizes its individual profit, without taking into account the impact of this action on the profit of the other CSP. The determination of the optimal outsourcing strategy for each CSP reduces to solving its individual profit maximization problem.

Each CSP independently determines its outsourcing strategy, but the decision affects the profit of both CSPs, and thus a non-cooperative game arises. The players in this non-cooperative game are $\langle P_1, P_2 \rangle$ and their strategies are their decisions on the part of the requests will outsource $\langle a_1, a_2 \rangle$. Our objective in this case is to define a **pricing mechanism**, according to which each CSP is charged by the other for the jobs that it outsources, so that the Nash equilibrium of this game still leads to a **mutually beneficial collaboration of the CSPs** even in this weak form of a federation. The payoff of each player in this game is given by the solution of its individual maximization problem. Therefore, in a situation of repetitive best-response dynamics, the game starts with the initial values of both be set in zero, and then in every step each CSP i takes as input the strategy of the other CSP and by solving its maximization problem decides on its new outsourcing strategy a_i^* . In every step of the game, the players use as input in their maximization problem the resulted strategy of the other CSP in the previous step. The games continues until the system reach a Nash equilibrium, where none of the players can increase its payoff by changing its strategy. Of course, the same equilibrium can be derived by solving a system of equations and constraints (see below).

Coming back to our objective for a proper definition of a pricing mechanism, it should be noted that the selfish behavior of CSPs can drive the game in equilibrium points that do not guarantee that the individual profit of both CSPs will be higher than their profit in their stand-alone operation. If this condition is violated for one of the CSPs, then this CSP does not have the incentive to participate. To this end, we have to define a certain mechanism that guarantees participation of both CSPs in the federation due to their mutual benefits. Thus, we define a pricing/compensation function for each CSP that determines the monetary amount a CSP should pay for the outsourcing requests to the other. Consequently, the objective function of the maximization problem includes the compensation amounts of both CSPs. For instance, the maximization problem for CSP₁ is as follows:

$$\begin{aligned} \max_{a_1} & P_1(a_1, a_2) - Q_1(a_1, a_2) + Q_2(a_1, a_2) \\ \text{s. t.} & 0 \leq a_1 \leq 1 \\ & \lambda'_i(a_1, a_2) < \mu_i, i = 1, 2 \end{aligned}$$

where $Q_1(a_1, a_2)$ denotes the compensation that CSP₁ will pay to CSP₂, while $Q_2(a_1, a_2)$ is the compensation in the opposite direction.

The compensation that a CSP i should pay for its requests transferred to the other CSP depends on the value of a_i . One approach is to assume that CSP i pays a price c_i per job (\$/job). Therefore, the monetary amount the CSP i pays per unit of time as compensation to the other CSP is given by

$$Q_i(a_1, a_2) = a_i \lambda_i c_i.$$

We assume that c_i is pricing function (or even a fixed price) that is mutually selected by both CSP and must satisfy the following conditions guaranteeing that the equilibrium of the game is mutually beneficial for both CSPs:

$$\begin{cases} P_1(a_1^*, a_2^*) \geq P_1(0, 0) \\ P_2(a_1^*, a_2^*) \geq P_2(0, 0) \end{cases}$$

In our analysis so far, we have not found a single c_i capable to satisfy the above pair of conditions for all the pairs of λ_1, λ_2 . An example of such a function for CSP₁ that satisfies the conditions for a broad range of pairs of λ_1, λ_2 is

$$c_1 = x_2 e^{-T_2(0,0)/T_1(0,0)}$$

where x_2 denotes the price per job that CSP₂ charges his customers for offering the service in the best possible level of QoS, and the exponent is the ratio of their customers delay in the stand alone operation. Thus, the CSP with the higher delay pays less for outsourcing a job.

Another way to define the compensation is to use it as an aggregate payment for the total stream of outsourced requests. In this case, we assume that each CSP compensates the other by paying half of the extra profit he manages to obtain from the requested outsourcing action:

$$Q_i(a_1, a_2) = R_i(a_1, a_2) - R_i(0,0) + E_i(a_1, a_2) - E_i(0,0).$$

where $R_i(a_1, a_2)$ and $E_i(a_1, a_2)$ are the revenues and the energy cost of CSP i in the outsourcing policy of (a_1, a_2) , while $R_0(0,0)$ and $E_0(0,0)$ are the revenues and energy cost in the standalone operation.

11.4 Appendix D: Evaluation of the energy efficiency of Multipath TCP

The energy efficiency of mobile connections is an important aspect of the traffic management mechanisms developed in the scope of the project. Here Multipath TCP (MPTCP) promises to increase the perceived network quality by allowing seamless handovers between different network technologies. However, besides the benefits in terms of service continuity, it remains unclear how the parallel usage of multiple interfaces via MPTCP influences the cost in terms of energy consumption on the mobile device. As battery lifetime is a critical factor for mobile users, it is important to quantify these effects. Consequently, this section analyses the energy consumption of MPTCP and determines energy-optimal configurations. In particular, this work aims to answer the following research questions:

- What is the energy cost of constant bitrate streaming using MPTCP on smartphones?
- Can the simultaneous use of multiple interfaces reduce the energy cost compared to individual interfaces?
- What is the most energy efficient configuration to use MPTCP on mobile devices?

To answer those questions, an MPTCP energy model is proposed. As opposed to simulation based optimization (cf. [38], [39], [40], [CLG13]), this model is based on real world hardware measurements, which are performed for two different types of smartphone. Contrary to regular downloads as analyzed in [42], the derived power model considers constant bitrate streaming. The resulting model allows optimizing for the most energy efficient interface or a combination of multiple interfaces, depending on the requirements of the application in use. By measuring power models for the individual and combined use of the network interfaces, the cost of MPTCP can be quantified, the optimal configuration is found, and the performance gain is quantified. Possible applications are video streaming, mobile cloud gaming, or any other real-time entertainment service. Furthermore, suggestions on improving the traffic control in MPTCP are proposed to further minimize the energy cost of network transmissions on mobile devices.

11.4.1 Background and related work

The basic features and mechanisms of the MPTCP transport protocol are described by the maintainers of the reference implementation in [37]. MPTCP's most important value proposition is the bundling of multiple TCP subflows over a single or multiple interfaces in order to provide parallel data transmission. The feature of parallel transmission may be used to achieve different goals. If the application's throughput demands exceed the capabilities of a single network interface, another network interface can be added to increase throughput (full MPTCP mode). Another mode of operation is the utilization of subflows for seamless handover (backup mode). In this mode, backup flows are established as soon as the endpoint can be reached over a different network interface. However, MPTCP only uses the cheaper interface (e.g., WiFi instead of 3G) for transmitting data. Finally, MPTCP can also be used as a replacement of TCP sending data over a single interface (single path mode), where it provides application transparent connection recovery.

A comparison of the body of related works investigating MPTCP is shown in [37]. All publications listed there consider the network performance of MPTCP (indicated as 'Throughput') in combination with regular downloads (column 'Regular Downl.'). The related works presented in [41] and [43] focus on throughput only, i.e. only the performance is measured while energy consumption is neglected. The authors of [41] examine the characteristics of MPTCP over wireless networks, concluding that the MPTCP performance is at least close or better than single-path performance in terms of download times. The work presented in [43] proposes a lightweight MPTCP proxy to be used for mobile users allowing incremental deployment of MPTCP. Even though these works do not consider energy efficiency, they provide valuable insights into MPTCP experiment setups and performance evaluation.

Other presented works on MPTCP investigate performance in combination energy efficiency (column 'Energy Cons.') [38], [39], [40], [42], [44], [45]. Most of these works are based on simulations and not on measurements (denoted as 'sim' in the measurement column): The authors of [38] propose a method to compute multipath schedules by solving Markov decision processes (MDP) offline. The evaluation shows that different interfaces are energy optimal depending on the throughput. A number of publications also focus on load balancing between interfaces (column 'Load Balancing'). In [39], an energy aware congestion control algorithm for MPTCP is proposed. The algorithm maintains MPTCP's throughput improvements, fairness to single-path flows and load-balancing properties. Chen et al. [40] develop an energy-aware data distribution scheme (eMTCP) that is built around MPTCP. The same authors later extended their proposal in [45] to take the traffic's level of burstiness into account. Bursty traffic causes the interfaces to frequently switch between active and idle states. The authors find a strong influence on energy efficiency.

Table 11-1: Comparison of Related Work

Reference	Throughput	Energy cons.	Load Balancing	Regular Downl.	Constant Bitrate	Measured
[41]	✓			✓		✓
[44]	✓	✓		✓		✓
[43]	✓			✓		✓
[38]	✓	✓		✓		Sim ¹
[39]	✓	✓	✓	✓		Sim ¹
[40]	✓	✓	✓	✓		Sim ¹
[45]	✓	✓	✓	✓		Sim ¹
[42]	✓	✓	✓	✓		✓
This work	✓	✓	✓		✓	✓

¹ Simulation only

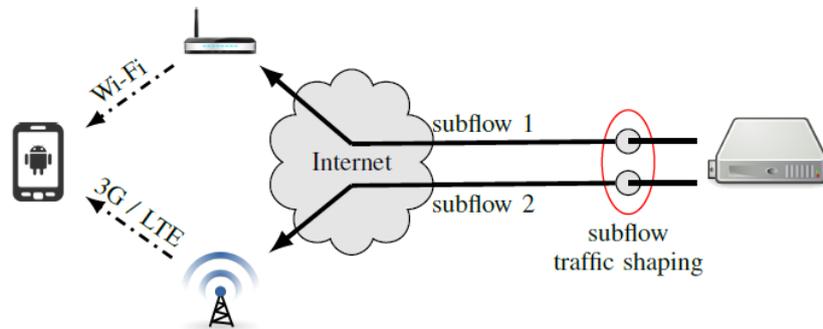


Figure 11-7: Experiment Setup: Traffic requested by the smartphone is split into two subflows transmitted via the WiFi and the cellular interface. Traffic shaping is applied to each subflow on the server.

However, for a realistic energy model, power measurements are necessary: Paasch et al. [44] experimentally analyze the energy consumption of WiFi/3G handover. Energy measurements using a Nokia N950 show that the 3G interface consumes about double the energy per bit compared to the WiFi interface. The authors of [42] measure a Galaxy S3 smartphone considering downloads and optimize the load balancing between interfaces. The optimized load balancing allows to save up to 15% of energy compared to standard MPTCP's.

The work presented here differs from the related work by considering constant bitrate streaming (column 'Constant Bitrate'), as is encountered in real-time entertainment (e.g. live video streaming, cloud gaming) or other time-sensitive (monitoring) applications. Moreover, this work is based on hardware measurements and not on simulation.

11.4.2 Measurement setup:

The measurement setup required for the evaluation of the energy consumption of MPTCP consists of a smartphone capable of connecting to 3G/4G and WiFi, and a server offering a constant bitrate data stream. The experiment setup is depicted in Figure 11-7. Both are equipped with an MPTCP capable kernel. The server runs two services: One providing a data stream to the mobile device, the other an interface to configure the test setup. On the mobile devices, the received traffic is monitored. The power consumption of the Nexus 5 is recorded on the device, while the power consumption of the Nexus S is measured externally. Combining both allows drawing conclusions on the energy efficiency of the different configurations. The setup for the individual components is detailed in the following sections and in [46].

Server Setup

On the server side, Debian 7.5 with a kernel version 3.11.10 was used. This was placed inside the department's server room and connected via a 1 Gbps link to the university network. This kernel was compiled with MPTCP in Version 0.88.11. This allows splitting the data stream requested by the mobile client to multiple subflows. The data for the mobile device is provided in the form of HTTP downloads by a PHP script, allowing configuring the size of the requested file, and optionally the download rate and bursting

intervals. Limiting the bandwidth of individual MPTCP subflows requires a bandwidth limiter on kernel level.

The download requests by the smartphone are handled by the download server, a PHP script running on the high performance HTTP server *nginx*. This ensures high performance and low delays. The downloads are configured by a number of parameters. This script reads data from */dev/urandom* according to the parameters of the HTTP GET request and sends these to the client. Possible parameters are:

- size Size (in bytes) of the download
- rate The requested data rate in kbyte/s
- burst_interval Interval between data bursts in seconds

These parameters allow simulating a wide range of requests patterns including web surfing, file downloads, and HTTP streaming. This is achieved by setting the average streaming rate in the rate parameter, but at the same time configuring burst intervals. This way, the maximum available data rate is used to send data to the client, but after a sufficient number of bytes for the configured interval are transmitted, the process pauses until the next interval is due. This is advantageous, as it allows the mobile device to switch the network interfaces to a power saving mode instead of keeping the device powered up to receive a continuously small data rate. This case can also be seen as a periodic HTTP download with the size of $r_{\text{req}} \cdot t_{\text{burst}}$, where r_{req} is the required data rate, and t_{burst} is the burst interval.

The traffic shaping is configured by the traffic control server. This is a set of python scripts, accepting commands via a telnet like interface to configure bandwidth limitations for the MPTCP subflows. This is achieved by acquiring the *inode* of the connection to the server using *fstat*. First, it verifies that MPTCP is used by reading */proc/net/mptcp* and searching for the respective *inode*. This is required, as MPTCP does not provide an API to control and monitor the functionality. After the remote IPs have been acquired, the traffic shaping is configured. For this, the Linux tool *tc* (traffic control) is used. Limiting individual subflows is achieved by installing queuing disciplines (short: *qdiscs*) for the individual TCP connections established by MPTCP. As *qdisc*, the Hierarchy Token Bucket (HTB) is used to limit the bandwidth, as it results in the highest accuracy. This is achieved by forwarding packets up to the configured rate, and queuing or discarding packets above. Configuring the respective data rate for each subflow, it is possible to limit the data rates on the cellular and WiFi interface individually.

Combining the features of both servers allows configuring the data rates in a high granularity. The download server limits the overall bandwidth of the data transfer to the mobile device, allowing to simulate a streaming connection. Via the traffic control server, the load on the individual interfaces is balanced, allowing to measure the energy efficiency of data transfers with differing interface utilization.

Smartphone Configuration

The smartphone in the test setup is used to receive the streaming data, control the test configuration, and record the experiment data. Therefore, different applications are used,

part of which was developed to allow recording the required data set or control the test setup.

The smartphones used for the measurement run a customized version of CyanogenMod² with the kernels 3.0.11 (Nexus S) and 3.4.0 (Nexus 5). These were patched with the MPTCP code in versions 0.86.6 and 0.86.7 respectively and compiled into system images. The functionality of the MPTCP implementation is verified by connecting to a test server³.

Still, this configuration does not allow using both network interfaces simultaneously. This can be activated by calling an undocumented function of Androids Connectivity Manager. After activating these, the policies and routes on the local device can be configured. These are changed by writing the respective values to the *proc* file system. This is done programmatically using an Android application.

As already noted in the previous sub-section, the configuration of the server is controlled by the mobile device. For this, different applications are used. The traffic control server is configured by connecting to the provided port via telnet, for which a number of applications is readily available. After the log-in, the shaping based on the connected IPs can be configured by executing simple commands. This is done for each test configuration, after which a series of tests with the respective configuration are run.

After the bandwidth is configured on the server, the data is requested from the download server. Depending on the test case, different file sizes and download rates are requested. This is done via an Android application, the NetworkEnergyMeter, written to simplify this task. Furthermore, it monitors the achieved throughput on the local device. For this, the time stamps and traffic received by the HTTP request are collected for later evaluation. As the load balancing for receiving data is controlled on the server, no additional configuration is required on the mobile phone.

Furthermore, the NetworkEnergyMeter measures the power consumption of the Nexus 5. Therefore, the download and power measurements run as a background service, allowing to minimize the error of the system state (i.e. display brightness) on the power measurement. Accurate power measurements are achieved by configuring idle periods before and after the tests. This allows the user to switch off the screen, and the system to hibernate tasks. All measurements are written to a log file after the tests have finished, such reducing the influence on the power measurement.

Power Measurements

The cost of data transmissions on the mobile device is calculated by allocating the consumed energy to the traffic received in the respective time frame. For this, the power consumption is monitored and integrated over the time of the transfer to calculate the consumed energy. Dividing this by the size of the transferred data volume, the cost of the transmission can be calculated.

² <http://www.cyanogenmod.org/> accessed 2014-11-20

³ <https://amiusingmptcp.com/> accessed 2014-12-05

The power consumption of the Nexus S can only be measured using external hardware as described in [47]. Therefore, the battery is removed from the device and placed in a modified charging cradle. This connects the battery via a high precision measurement shunt (error < 1 %) to a battery dummy, which is placed inside the smartphone. The battery voltage and voltage drop over the measurement shunt are measured using a 16 bit A/D converter (Measurement Computing USB-1608FSPlus). This allows measuring the power consumption with high accuracy and low error. The measurements are recorded on a connected PC and written to a file for later analysis.

Special care was taken to synchronize the clocks between both devices, to not affect the accuracy when later combining both measurements. For this purpose, the Network Time Protocol (NTP) with a common server was used on both devices. Compared to the measurements on the Nexus S, the measurements on the Nexus 5 are comparatively straightforward. Here, the built in current sensor (MAX17048) was read by a background service and written to a separate log file. This eliminates the need for synchronizing the clocks, but requires some thought to not affect the measurements due to changed system load. This effect is minimized by running reference measurements while the device is idle, which are later subtracted from the active measurements. The current sensing chip measures the current consumed by the smartphone with a resolution of 16 bit, resulting in a high accuracy of the measurements. These are available via the kernel file system. Tests have shown that contrary to the rest of the *proc* file system, this value changes for each request. Sample rates of up to 40 samples/s have been achieved. A sample rate of 20 samples/s was chosen, limiting the system load when idle to approximately 10 %. This leaves sufficient resources to not affect the throughput by limited CPU resources. Here, it is inevitable to keep the sample rate between the reference and the active measurements constant, as otherwise the results are biased. Combining the measurements from the built in current and voltage sensors allows calculating the instantaneous power consumption of the Nexus 5.

11.4.3 Measurement results

To measure the power consumption of MPTCP, a number of reference measurements are required. First, the idle power of the device must be determined and eliminated from the transmission cost. To reduce the effect of background processes, care was taken to limit their number. The noise in the measurement was further reduced by measuring only when the display was off. On top of the baseline consumption, the cost of each interface was determined. This is used as reference for the later comparison of the energy efficiency between individual connections and an MPTCP connection. Based on the power consumption of individual interfaces, the overall power when using multiple interfaces is determined.

The following sections describe the reference power measurements and models used to estimate the MPTCP power consumption, and the final MPTCP power measurements.

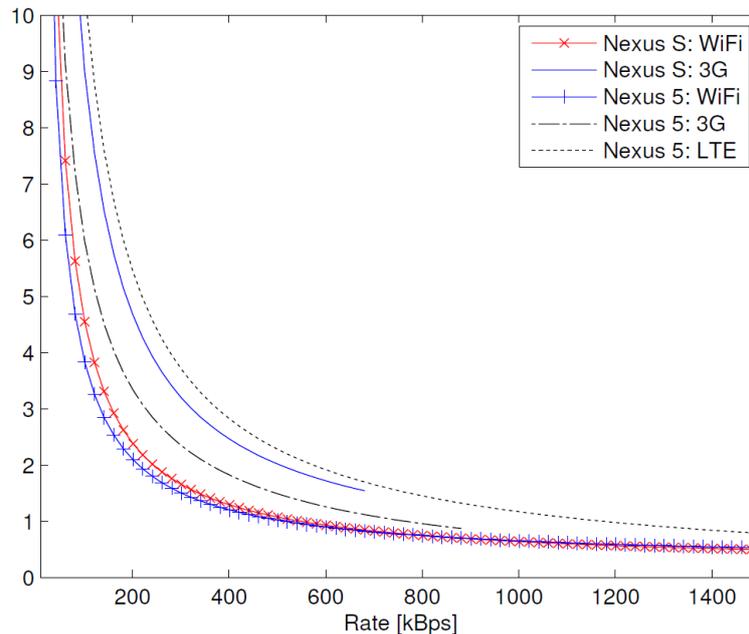


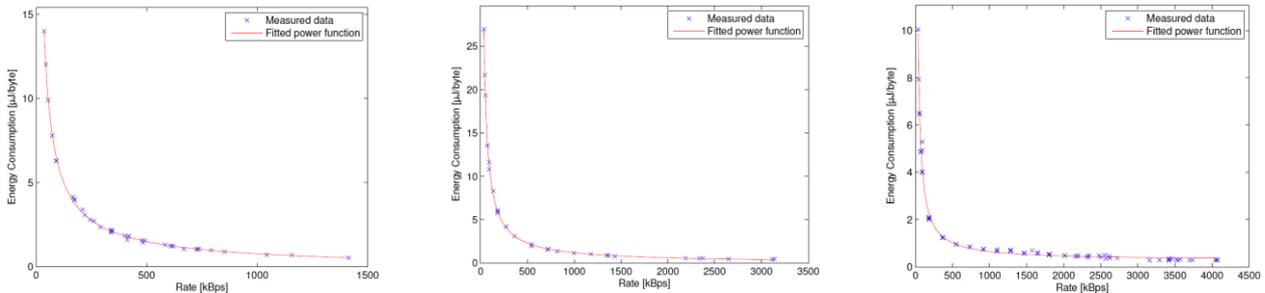
Figure 11-8: Cost in $\mu\text{J}/\text{B}$ for different access technologies and devices

Reference Power Models

The reference power consumption of the 3G/4G and WiFi interface is measured using the measurement setup described above. To allow a comparison of the measurements, the same location was used for all measurements. The effects of varying network performance were reduced by measuring at night time. Still, parallel tests were not executed to reduce interference of the different flows with each other. The measurements of the cost of transmitting data with varying rates over different interfaces of the Nexus 5 and the resulting cost models are shown in Figure 11-8. Similar measurements were also conducted for the Nexus S, but are omitted due to space limitations. As is visible from these, the cost for transmitting one byte at a given rate shows an exponential behavior. Hence, a power function in the form $C(R) = a \cdot R_b + c$ was fitted to the measurements. All plots show an accurate fit to the measurements, although small deviations are visible in the case of higher data rates using WiFi on the Nexus 5. The cost functions for the different interfaces of both devices are compared in Figure 11-9.

The exponential decay of the cost functions is expected, as the idle power of the interface is included in the transmission cost. The fraction of the overall energy required to transmit one byte is particularly high for low data rates, leading to the curves seen in Figure 11-10. For higher data rates, the constant part is shared between a larger number of packets, leading to a lower cost per byte. The cost per byte transmitted is lowest for WiFi on both devices, followed by 3G on the Nexus 5, 3G on the Nexus S, and LTE on the Nexus 5. From this, the general conclusion can be drawn that for low data rates (excluding ramp and tail energies) the use of 3G is generally more energy efficient than LTE. Only for data rates higher than the ones supported by 3G, LTE should be used in terms of energy efficiency. Still, one must consider the different Roundtrip Times (RTTs) of 3G and LTE, if

the service is highly interactive. The cost for transmitting one byte on the individual interfaces (in $\mu\text{J}/\text{B}$) is also given in Figure 11-9.



(a) Cost of streaming with various data rates via 3G

(b) Cost of streaming with various data rates via LTE

(c) Cost of streaming with various data rates via WiFi

Figure 11-9: Cost of receiving data on the Nexus 5 using different interfaces

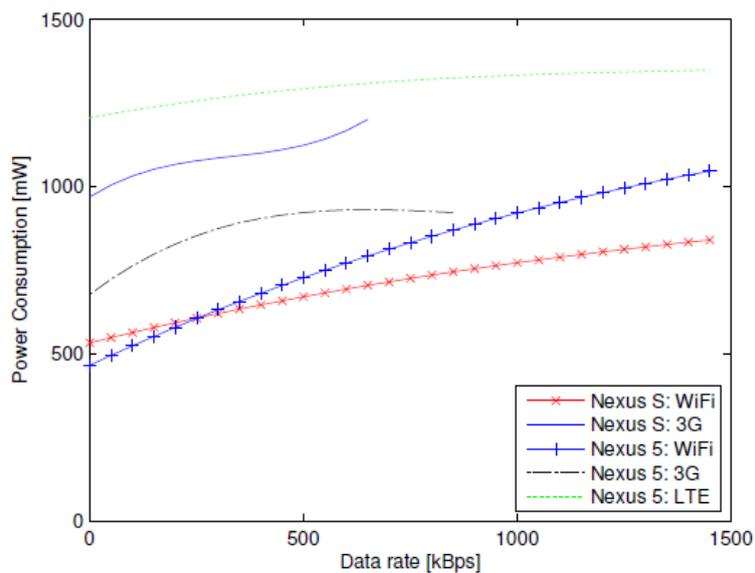


Figure 11-10: Derived power models for the Nexus S and Nexus 5 while using different network access technologies

Table 11-2: Power consumption of the Nexus S and Nexus 5 for different interfaces and data rates

Nexus S	
WiFi only	$C(r) = 449.6 \text{ mW} \cdot r^{-1.005} + 0.2063 \mu\text{J B}^{-1}$
3G only	$C(r) = 831.3 \text{ mW} \cdot r^{-0.9857} + 0.2024 \mu\text{J B}^{-1}$
Nexus 5	
WiFi only	$C(r) = 315.7 \text{ mW} \cdot r^{-0.9714} + 0.2670 \mu\text{J B}^{-1}$
3G only	$C(r) = 264.1 \text{ mW} \cdot r^{-0.8141} - 0.1863 \mu\text{J B}^{-1}$
4G only	$C(r) = 798.6 \text{ mW} \cdot r^{-0.9393} - 0.0447 \mu\text{J B}^{-1}$

Converting the cost to power models, the functions in Table 11-2 are derived. The graphs show the power consumption of the Nexus S and Nexus 5 while transferring data with the indicated bit-rate. For both devices the WiFi connections show the lowest power consumption, followed by 3G and LTE on the Nexus 5. It is remarkable to see the difference in power consumption between the different 3G implementations and chip sets of the Nexus S and Nexus 5, which exhibit a difference of 300mW for identical data rates. The 3G data rates on the Nexus 5 are higher than the ones on the Nexus S. This is caused by the availability of HSDPA+ on the Nexus 5. LTE on the Nexus 5 results in a comparatively high power consumption, but also data rates comparable to WiFi.

MPTCP Throughput

To determine the cost of using MPTCP, further traffic measurements were conducted. As already noted, the cellular network is the one aspect of the test, which was not controllable. Furthermore, the requested data rates are well under the maximum data rates available using the different network technologies at the given location. This eliminates the effect of external bandwidth limitations on the measured effects. To assert the accuracy of the generated model, the achieved throughput using MPTCP was also measured on the mobile device. In the following measurements, the traffic shaping limits the overall data rate of the individual subflows. The traffic measurements are conducted on the mobile device, logging the traffic received via the HTTP connection.

Figure 11-14 (a) shows the measured throughput for a configured data rate of 200 kBps. As is visible in the figure, the maximum data rate achieved is slightly above 180 kBps. This relates to an overhead of 10% when using MPTCP in combination with HTTP streaming. Still, downloads with a fraction of smaller 50% of the traffic on the WiFi interface result in a significant packet loss, and such reduced bandwidth. This is even more apparent in Figure 11-14(b), showing the maximum throughput of the Nexus 5 for a requested data rate of 500 kBps via 3G and WiFi. It is also interesting to observe the high variance of the measurements when using MPTCP mainly on the cellular interface.

The MPTCP throughput measurements of a 2 MBps stream on the Nexus 5 are shown in Figure 11-14(c). Here, WiFi in combination with LTE is used. The maximum net throughput achieved is approximately 1.8 MBps. The lowest throughput measured was 850 kBps. A similar behavior as in the previous measurements is visible, although the drop in

performance is smaller. Still, the higher variance for a lower fraction of traffic on the WiFi interface is visible, as well as a reduced performance for these configurations.

The effect of lower maximum throughput is also visible in other measurements not shown here. This is expected, as the traffic shaping is conducted on the network layer, while the traffic measurements on the mobile device are conducted on the application layer. These traffic measurements show that an overhead of 10% of the original traffic size is added when using HTTP streaming via MPTCP. The detailed cost of each component should be analyzed in detail in future work.

The higher variance for a lower fraction of traffic on the WiFi interface is thought to be caused by the traffic management algorithms of MPTCP. As is noted in [42], the primary interface when using multiple interfaces on the smartphone is the WiFi interface. Hence, if the MPTCP control messages are delayed by the congestion of the cellular interface, or queued before transmission on the WiFi interface, the overall performance drops. Further analysis of this effect in future work is recommended.

The traffic rates, as measured on the application layer, are then used to model the cost of receiving streaming data with various rates on the Nexus S and Nexus 5. For this, the above power and throughput models are combined to create an MPTCP power model.

MPTCP Power Measurements

The power consumption of MPTCP transmissions using multiple interfaces is measured, combining the MPTCP subflow shaping with the bandwidth limited HTTP server. The power measurements are conducted analog to the measurements described above. From the device power consumption the idle power is subtracted, resulting in the cost of data transmissions only. As these measurements are targeted at streaming solutions, the ramp and tail energies are neglected. In the case of downloads or video streaming using data bursts, these need to be added before and after each transmission.

The measurements of the different possible combinations (i.e. Nexus S Wifi/3G, Nexus 5 Wifi/3G, Nexus 5 Wifi/LTE) are given in Figures 7 to 9 as crosses. The horizontal axis shows the fraction of traffic received on the WiFi interface. Hence, 0% reflects a pure 3G/LTE transmission, while 100% represents a pure WiFi transfer. In between, the traffic is transferred using parallel subflows on both interfaces. As might be expected from the power measurements, the power consumption per byte is lowest when using WiFi alone. The second cheapest option is using the cell interface alone. In terms of energy consumption, the use of parallel subflows should be considered only if the required transfer rate cannot be supported by a single interface. When using MPTCP, it is visible that data transfers are cheapest, when the fraction of traffic on the cell interface is minimized.

From these observations, a general, energy efficient load balancing model can be derived. It can be derived, that first the interface with the lowest RTT should be used to capacity. If the available data rate on this interface is not sufficient, the next interface should be checked. If this provides a sufficient throughput, it should be used exclusively. Only if the requested data rate cannot be supported by either interface, the second interface should be added. Still, the majority of the traffic should be processed by the primary interface, as

this result in the highest performance and lowest cost. This decision logic is depicted in Figure 11-11. The detailed derivation of this can be found in [46].

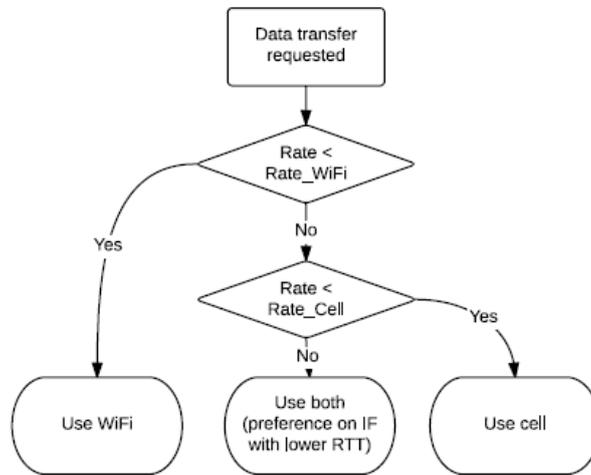


Figure 11-11: Decision tree for finding the most energy efficient transfer mode for a given data rate

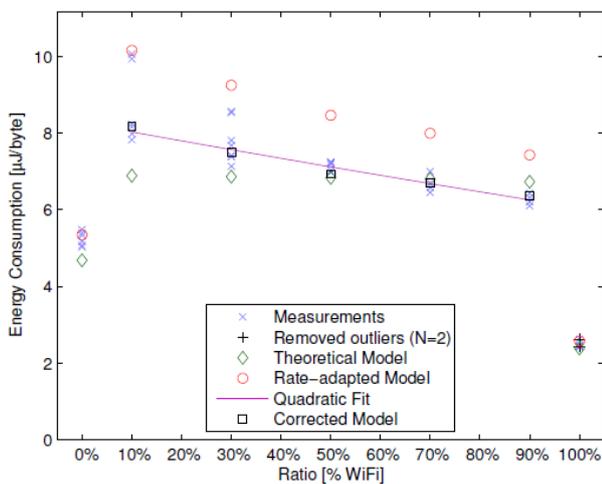


Figure 11-12: 200 kbps download via 3G and WiFi on the Nexus S

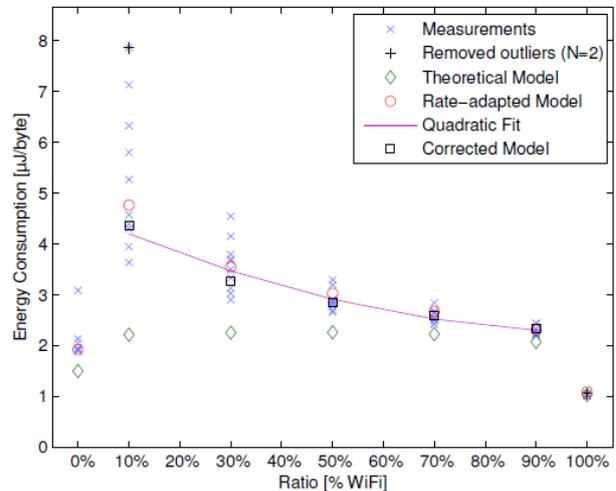
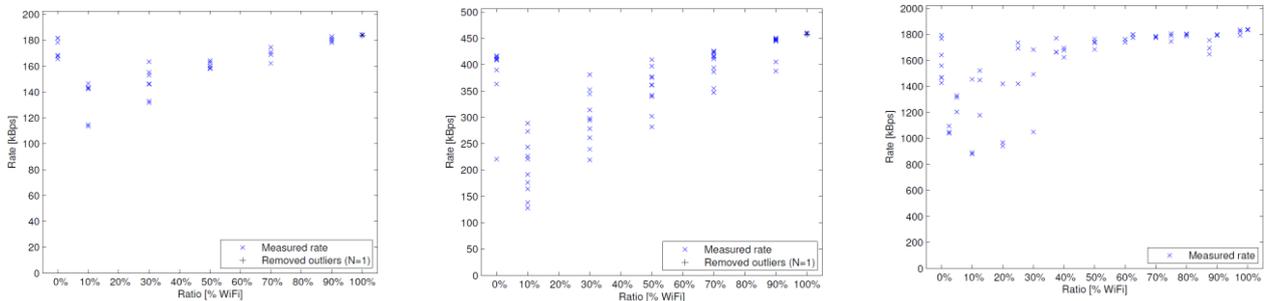


Figure 11-13: 500 kbps download via 3G and WiFi on the Nexus 5



(a) MPTCP throughput for streaming data limited to 200 kbps on the Nexus S via 3G and WiFi

(b) MPTCP throughput for streaming data limited to 500 kbps on the Nexus 5 via 3G and WiFi

(c) MPTCP throughput for streaming data limited to 2 MBps on the Nexus S via 3G and WiFi

Figure 11-14: Measured net throughput using MPTCP for varying interfaces and streaming rates

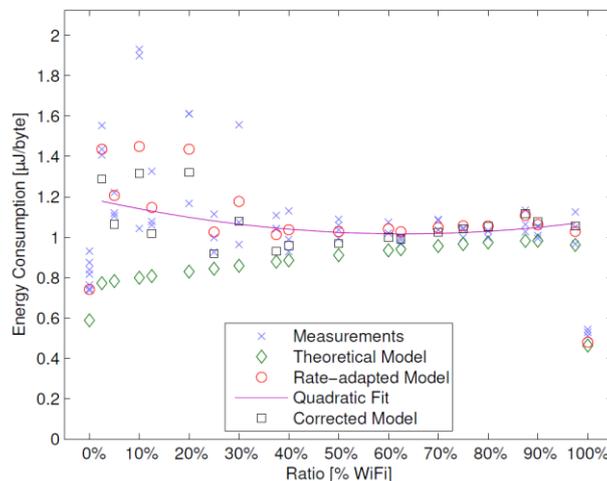


Figure 11-15: 2 MBps download via LTE and WiFi on the Nexus 5

Cleaning of the Measurements

The gathered data was analyzed to remove possible outliers. For this, Cook’s distance [48] was calculated for each configuration (i.e. WiFi, cellular, MPTCP). Values with a distance of over $4 = N_{\text{samples}}$ of the respective configuration were removed from the data set. The number of discarded data points is given in the legend of the respective plots. The location of the outliers is indicated in the form of plus signs.

11.4.4 Power model

There are two possible approaches to model the MPTCP power consumption: 1) Adding the power consumption of both interfaces, or 2) fitting a function to the measured values. The first approach is based on the assumption that components within the smartphone are independent, while the second aims to derive the simplest model based on the measurements. In this work a model is derived, which is both based on the physical preconditions and approximates the power consumption based on the measurements.

A. Additive Power Model

Using an additive power model is intuitive, as often separate hardware components within the smartphone (i.e. baseband chips, power amplifiers) are used for different technologies. Hence, the power consumption of the MPTCP connections is first modeled by calculating the weighted sum of the cost per byte for the individual interfaces according to the power models determined in section 11.4.3. Considering the cost C for transmitting data on an interface is defined as the energy E consumed while transmitting data with the size S :

$$C = \frac{E}{S} \quad (1)$$

As the energy E is the integral of the power P over the time t ($E = \int P dt$), and the transferred volume S is the data rate R integrated over the time ($S = \int R dt$), the above relation can be simplified to

$$\bar{C} = \frac{\bar{P}}{\bar{R}} \quad (2)$$

Therefore, the average cost using both interfaces is given by

$$\bar{C} = \frac{\bar{R}_1}{\bar{R}} \cdot C_1(\bar{R}_1) + \frac{\bar{R}_2}{\bar{R}} \cdot C_2(\bar{R}_2) \quad (3)$$

Here, R_1 and R_2 are the rates on the interfaces 1 and 2, and $C_{1/2}$ is the cost for receiving data at the given rates. The approximations based on the theoretical model with configured rates $R_{1/2}$ are given in Figure 11-12 and Figure 11-14 as circles. Using the measured data rates, R_{meas} results in the rate-adapted model in form of diamonds.

Figure 11-12 shows the cost per byte of the Nexus S while streaming data with 200 kbps. The rate-adapted power model in the shape of diamonds clearly does not fit the actual power consumption of the device. Contrary, the power model based on the shape rates (circles) leads to a smaller error, but does not reflect the cost of the different interfaces well. Figure 11-13 shows a 500 kbps stream via 3G and WiFi on the Nexus 5. A similar problem with the theoretical model (diamonds, based on the configured rates) is visible. Using the additive, rate-adapted model (circles) leads to an acceptable fit. Still, an error between the measurements and the fitted model is visible. Figure 11-15 shows the power models for a download stream of 2 MBps on the Nexus 5. Again, the theoretical model (diamonds), based on the shaping rates, performs worst. The rate-adapted model (circles) performs better, but still does not fit the quadratic model fitted to the measured data. Hence, an improved power model based on the interface power consumption is required.

B. Improved Power Model

Comparing the additive models with the measurements indicates that the power consumption of the MPTCP transmissions is lower than expected. Hence, some synergies must exist when using both interfaces. As this is not reflected in the additive models, these effects are analyzed in detail. Substituting the cost per interface $C_{1/2}$ with Equation (2), a simplified representation for the average cost for using MPTCP can be achieved:

$$\bar{C} = \frac{1}{\bar{R}} [P_1(\bar{R}_1) + P_2(\bar{R}_2)] \quad (4)$$

Hence, the cost function can be modeled as the additive power consumption of the used interfaces, depending on the individual data rates, divided by the overall data rate. Using the second order approximations, a general dependency for the MPTCP power model can be derived:

$$C = \frac{1}{R} [a_{1,0} + a_{1,1} \cdot R_1 + a_{1,2} \cdot R_1^2 + a_{2,0} + a_{2,1} \cdot R_2 + a_{2,2} \cdot R_2^2] \quad (5)$$

For a fixed data rate R , R_1 and R_2 are related via

$$R_1 = f_1 \cdot R \quad (6)$$

$$R_2 = f_2 \cdot R \quad (7)$$

The fraction $f_{1/2}$ is the traffic processed by the interface 1 and 2 respectively. f_1 and f_2 are related by $1 = f_1 + f_2$. Hence, the MPTCP power consumption can be modeled based on the overall rate R and the fraction of traffic on a single interface. Substituting R_1 and R_2 with equations 6 and 7, equation 5 can be simplified to

$$C = \frac{1}{R} [d_0(R) + d_1(R) \cdot f_1 + d_2(R) \cdot f_1^2] \quad (8)$$

where $d_0(R)$, $d_1(R)$, and $d_2(R)$ are defined as follows:

$$d_0(R) = a_{1,0} + a_{2,0} + a_{2,1} \cdot R + a_{2,2} \cdot R^2 \quad (9)$$

$$d_1(R) = (a_{1,2} - a_{2,1}) \cdot R - 2 \cdot a_{2,2} \cdot R^2 \quad (10)$$

$$d_2(R) = (a_{2,2} + a_{2,1}) \cdot R^2 \quad (11)$$

Hence, for each target rate R , a quadratic fit over the fraction of traffic on each interface is plausible. This quadratic fit to the MPTCP connections is given in Figure 11-12 and Figure 11-14 as a line. The function was fitted using the Matlab *robustfit* function, automatically weighting the samples according to their distance to the approximation, and such minimizing the error of the final fit.

Table 11-3: Second order fits to the MPTCP Measurements

Nexus S	
WiFi/3G (200 kbps)	$C(f_W)/\mu\text{W B}^{-1} = 8.261 - 2.382 \cdot f_W + 0.178 \cdot f_W^2$
Nexus 5	
WiFi/3G (500 kbps)	$C(f_W)/\mu\text{W B}^{-1} = 4.627 - 4.517 \cdot f_W + 2.141 \cdot f_W^2$
WiFi/LTE (1 MBps)	$C(f_W)/\mu\text{W B}^{-1} = 2.248 - 0.313 \cdot f_W - 0.231 \cdot f_W^2$
WiFi/LTE (2 MBps)	$C(f_W)/\mu\text{W B}^{-1} = 4.627 - 0.584 \cdot f_W + 0.454 \cdot f_W^2$

Table 11-4: Differences between the rate-adapted model and the quadratic fit and resulting correction functions

Nexus S		
Setting	Dev.	Function
WiFi/3G (200 kBps)	-21.2 %	$C_c(f_W)/\mu\text{J B}^{-1} = +1.144 \cdot f_W - 2.099$
Nexus 5		
Setting	Dev.	Correction Function
WiFi/3G (500 kBps)	-5.2 %	$C_c(f_W)/\mu\text{J B}^{-1} = +0.523 \cdot f_W - 0.446$
WiFi/LTE (1 MBps)	+2.5 %	$C_c(f_W)/\mu\text{J B}^{-1} = -0.198 \cdot f_W - 0.153$
WiFi/LTE (2 MBps)	-5.8 %	$C_c(f_W)/\mu\text{J B}^{-1} = +0.185 \cdot f_W - 0.152$

In the case of the Nexus S, the function is almost linear, while quadratic terms are visible for the Nexus 5. The models for the Nexus S overestimate the power consumption when using both interfaces in parallel, while the model for the Nexus 5 is quite accurate. The resulting cost functions for streaming over MPTCP are given in Table 11-2. The mean deviation between the quadratic fit to the measurements as given in Table 11-5 and the additive, rate-adapted power model as derived from the models in Table 11-3 is given in the second column of the Table 11-4. This can be interpreted as additional cost or energy savings when using MPTCP with the given setting compared to the theoretical cost of using both interfaces independently with the requested data rates. Negative values denote a gain in efficiency, as the energy expense is lower than the estimates, while positive values indicate additionally consumed energy. The energy savings on the Nexus S confirm measurements by Lim et al. [42], which have shown an increase of energy efficiency of up to 15% on the Galaxy S3, although for higher data rates. Generally, the power model derived for the Nexus 5 fits well to the measured cost per byte. This is true over the full range of measurements executed. From this, it can be concluded, that synergetic effects between the interfaces are negligible.

Considering the error as is visible in Figure 11-12, an additive correction is likely. The validity of this is checked by fitting a linear function to the deviation between the additive, rate-adapted model and the quadratic fit. The resulting correction functions C_c for the MPTCP cost are given in the third column of Table 11-4. From the coefficients, the hypothesis of an additive correction can be confirmed. The static term in the functions $C_c(f_W)$ gives a constant to be added to the model to eliminate the constant error term. This is considerable with a cost of 2:099 $\mu\text{J/B}$ for the Nexus S, while comparatively small for the Nexus 5, where it is in the range 0:1522 $\mu\text{J/B}$ to 0:4458 $\mu\text{J/B}$, which compared to the overall cost of the transmission is approximately 10% of the cost of receiving data on the interface.

By checking the slope of the function, the dependency on the fraction of traffic on the WiFi interface f_W is determined. These terms are given in $\mu\text{J s}$, which multiplied by a traffic rate r in B/s results in the cost in $\mu\text{J/B}$. These terms show a general dependency on the fraction of traffic on the WiFi interface. Still, the tendency is not consistent. Hence, no general conclusion on the dependency of the synergies when using MPTCP can be drawn.

Table 11-5: RMSE of the different models

Nexus S			
Setting	Static	Rate-adapted	Corrected
WiFi/3G (200 kbps)	0.9972	1.2665	0.5263
Nexus 5			
Setting	Static	Rate-adapted	Corrected
WiFi/3G (500 kbps)	1.3508	0.5148	0.5753
WiFi/LTE (1 Mbps)	0.4376	0.1834	0.1714
WiFi/LTE (2 Mbps)	0.3452	0.1365	0.1655

Adding these (empirical) correction functions to the original functions, an improved cost model for using MPTCP on mobile devices can be generated. Hence, eq. (8) can be extended by adding the correction functions. The resulting equation for the MPTCP operation is:

$$C = \frac{1}{R} [d_0(R) + c_0 + (d_1(R) + c_1) \cdot f_1 + d_2(R) \cdot f_1^2] \quad (12)$$

The comparison of the Root Mean Square Error (RMSE) for the different models is given in Table 11-5. The table shows that a considerable gain in accuracy can be achieved when applying the correction to the MPTCP cost model of the Nexus S. The accuracy of the rate-adapted model cannot be improved by the additional correction terms. To assess the best approach for reducing the final error, a larger sample size is required.

11.4.5 Conclusion

The main findings are summarized along the research questions posed in the beginning, for which the answers are outlined in the following.

What is the energy cost of constant bitrate streaming using MPTCP on smartphones? The cost of using MPTCP on smartphones is comparable to the added cost of using both interfaces simultaneously. In the case of the Nexus 5, this is the best approximation possible. The Nexus S exhibits a lower cost when streaming data using MPTCP. For a 200 kbps stream, the cost is 21.8% lower than the added cost of both interfaces.

Can the simultaneous use of multiple interfaces reduce the energy cost compared to individual interfaces? Synergies when using MPTCP have only been observed on the Nexus S. Here, the cost of MPTCP is 20% lower than the theoretical cost of both interfaces individually. The Nexus 5 shows power consumption proportional to the added cost of both interfaces.

What is the most energy efficient configuration to use MPTCP on mobile devices? According to the measurements, the cost of MPTCP is lowest when using the most energy efficient interface only. If the achieved data rate is not sufficient, and the data rate of the secondary interface (cellular) is higher, using only the cellular interface is most energy efficient. Only if the data rate cannot be supported by a single interface, both interfaces should be used. The most energy efficient setting in this configuration can be achieved by transferring most traffic on the interface with the lower RTT, which in our case

is the WiFi interface. This stands in contrast with the current MPTCP implementation, which uses both interfaces with equal data rates. Hence, it is suggested to add an energy efficiency mode to the MPTCP implementation, considering these observations to conserve the available energy on the mobile device.

As MPTCP is one of the lower layer technologies allowing seamless handover between different network technologies, it was selected to be used by MoNA as the technology for shifting loads between networks. As detailed in the previous sections, the current network selection mechanism is not optimal. Hence, the improved connection selection mechanism together with the power model derived for MPTCP connections is to be used by MoNA to select the best access technology depending on the current user requirements and network availability.

The developed MPTCP connection selection algorithm (cf. Figure 11-11) is to be used to optimize the energy efficiency of the mobile device. At the same time this algorithm considers the QoE of the end user by adjusting the selection of the network interfaces to the traffic demands of the mobile user. This works in close connection with RB-Horst, and vINCENT, as by increasing the availability of local content, and such increasing the available data rates, or in the case of vINCENT providing additional offloading opportunities, the mobile connectivity selection attains its full potential. Hence, both the energy efficiency aspect and the QoE aspect are addressed by the MoNA mechanism.

11.5 Appendix E - Options and Performance Evaluation of Caching

11.5.1 Access Pattern for Web Content and Efficient Caching Strategies

The efficiency of caching basically depends on the hit rate, i.e. the fraction of requests and data volume that can be served from the cache. Preferences and usage pattern are different in local environments or special networks for enterprises, universities, social networks or for IP-TV services. Caching for a single user already can save or prefetch a considerable amount of data transfer due to partly repetitive usage pattern [22]. For small user communities as involved in uNaDas, the efficiency is expected to be higher. Caches in user premises can completely avoid network load and delay, which is most valuable for congested access networks e.g. on air interfaces of wireless and mobile networks and for real time applications.

For single users or small user communities, request pattern are varying without general models known in literature, whereas Zipf laws seem to be uniquely valid for a large user population accessing a large number of objects [19]. A Zipf distribution assigns the object in popularity rank R an access probability of $Z(R) = \alpha R^{-\beta}$. The skewness of a Zipf law can be adapted to 70:30-, 80:20- or 90:10-rules by appropriate choice of the parameter β where an $x:y$ -rule means that the top $y\%$ of the objects attract $x\%$ of the user requests. More than two dozen measurement studies confirm Zipf-like access distributions for different Internet platforms and applications including video streaming, file downloads and IP-TV in client-server and P2P distribution schemes. Estimations of the Zipf parameter β are given in more than half a dozen studies, which all fall into the range $0.55 \leq \beta \leq 0.88$.

11.5.2 Content delivery simulation framework

In the content delivery simulation framework (see section 4.2.1), different caching strategies, like LRU, LFU or score gated LRU are implemented. Further on different document request processes are implemented, in particular a Zipf request process and the box request process described in deliverable D2.4[7], which exhibits temporal dynamics.

We simulate a simple topology that consists of only one cache with cache size C and one user that requests documents according to a Zipf process and the box process with parameter β . The box process uses two more parameters, which is the mean lifespan τ and the standard deviation of the lifespan σ . We simulate 10^5 document requests and show numeric examples for a catalogue size of 10^4 documents. We set the simulation time to one month and the mean lifespan to $\tau = 10.6$ days and the standard deviation to $\sigma = 3.7$ days, based on the results from [18].

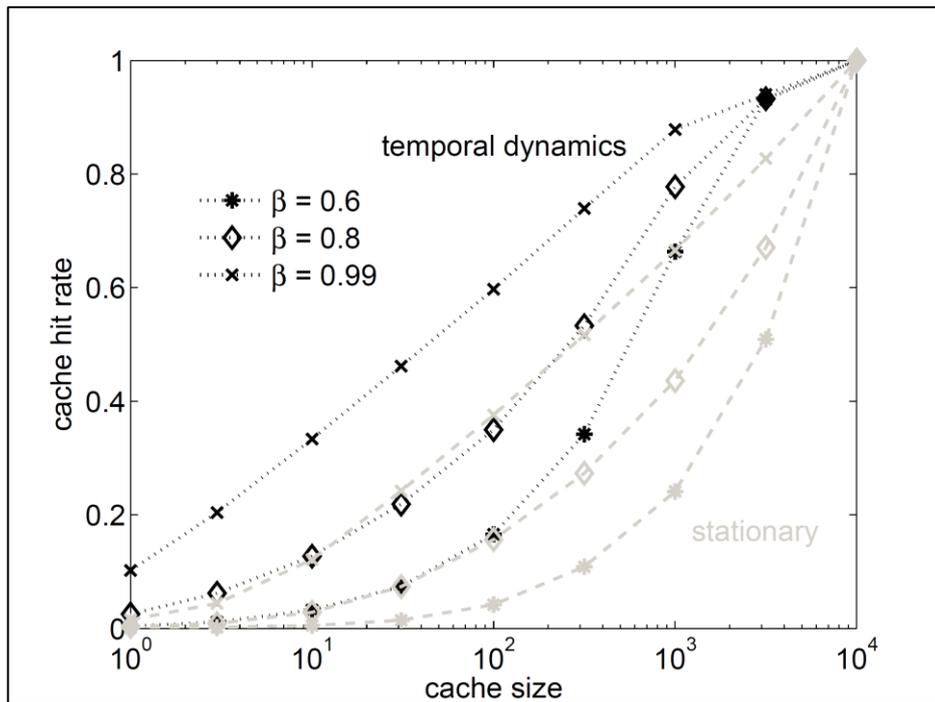


Figure 11-16: Hit rate of an LRU cache. Comparison of dynamic and stationary request processes.

Figure 4-13 shows the hit rate of the LRU cache. The cache hit rate increases with the cache size and approaches 1 as the cache size approaches the catalogue size. The cache is more efficient if the Zipf parameter β is high, which results in a more heterogeneous content popularity. In this case the LRU cache performs better with the dynamic request process compared to the stationary request process. This depends on the fact that documents with expired lifespan that would not generate more hits are replaced by more recent documents in the LRU scheme. However this depends on the lifespan distribution and the resulting dynamics. It is part of future work to investigate the impact of the lifespan distribution on cache performance.

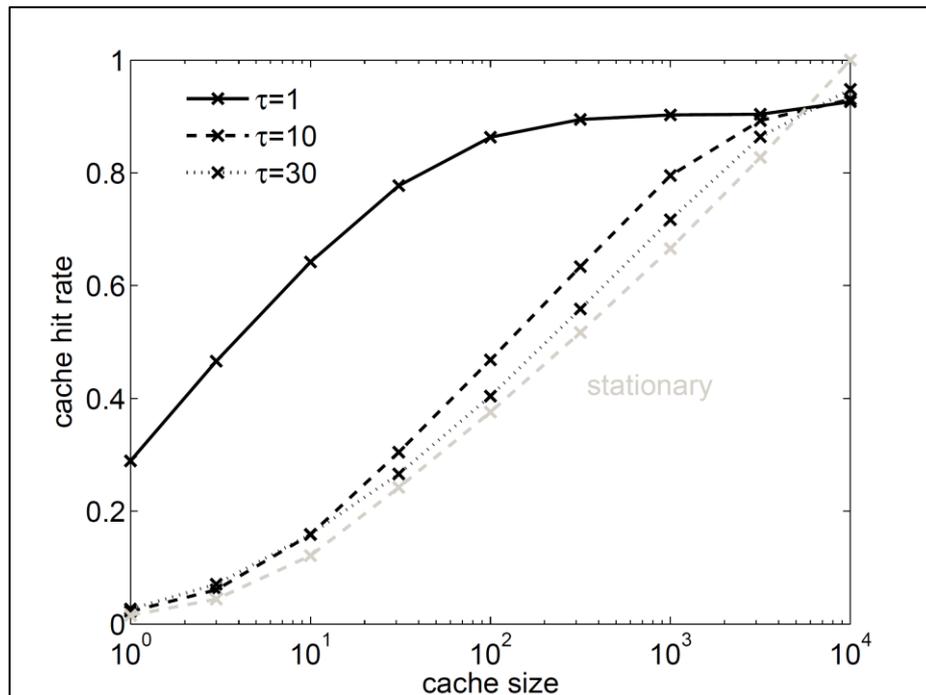


Figure 11-17: Hit rate of an LRU cache dependent on mean lifespan of documents

- *Access Pattern for Web Content and Efficient Caching Strategies*

The efficiency of caching basically depends on the hit rate, i.e. the fraction of requests and data volume that can be served from the cache. Preferences and usage pattern are different in local environments or special networks for enterprises, universities, social networks or for IP-TV services. Caching for a single user already can save or prefetch considerable amount of data transfers due to partly repetitive usage pattern [22]. For small user communities as involved in uNaDas, the efficiency is expected to be higher. Caches in user premises can completely avoid network load and delay, which is most valuable for congested access networks e.g. on air interfaces of wireless and mobile networks and for real time applications.

For single users or small user communities, request pattern are varying without general models known in literature, whereas Zipf laws seem to be uniquely valid for a large user population accessing a large number of objects [19]. A Zipf distribution assigns the object in popularity rank R an access probability of $Z(R) = \alpha R^{-\beta}$. The skewness of a Zipf law can be adapted to 70:30-, 80:20- or 90:10-rules by appropriate choice of the parameter β where an $x:y$ -rule means that the top $y\%$ of the objects attract $x\%$ of the user requests. More than two dozen measurement studies confirm Zipf-like access distributions for different Internet platforms and applications including video streaming, file downloads and IP-TV in client-server and P2P distribution schemes. Estimations of the Zipf parameter β are given in more than half a dozen studies, which all fall into the range $0.55 \leq \beta \leq 0.88$.

11.5.3 Efficiency Study of Web Caching Strategies Combining LFU and LRU

Demands for Efficient Web Caching

Efficient web caching strategies have to be implemented regarding the following demands:

- (A) Simple implementation with constant $O(1)$ update effort per request as for LRU;
- (B) High hit rate for the basic scenario of independent random requests (IRM);
- (C) Adaptive response to changing popularity of objects;
- (D) Low input traffic to the cache avoiding multiple loading of the same object.

The least recently used (LRU) caching strategy meets the demands (A) and (C), but has obvious deficits regarding (B) and (D). On the other hand, least frequently used (LFU) meets demands (A), (B) and (D), but fails on (C), making it inapplicable in practice.

Many variants of caching strategies have been proposed and evaluated in the literature which include additional information or statistics about past requests than LRU but also limit the memory backlog in contrast to LFU and therefore stay adaptive to changing popularity over time. The class of such strategies is attributed in [29] as LFRU spectrum. The caching schemes include LRU and LFU as extreme cases depending on a parameter.

Usual caching strategies are based on a score function which assigns a score value to each object and puts objects with highest scores into the cache. A straightforward implementation keeps an ordered list of the objects according to scores, but this makes updates per request often more complex than constant $O(1)$ effort.

In the sequel, we refer to a simpler updating scheme studied by the authors in recent work [27], which shows favourable properties with regard to all demands (A)-(D). The considered score-gated LRU (SG-LRU) scheme doesn't enforce strict ordering of objects in the cache according to highest scores, but extends LRU by requiring a higher score for an external object compared to the bottom object of the cache before entering the cache. This scheme combines simple LRU cache updates with preference for objects based on a score function, which proved to be sufficient when object scores are stable or only slowly varying over time. In our evaluations we focus on SG-LRU with a geometrically fading request count as a most promising caching strategy which combines LFU and LRU and includes both basic strategies as extreme cases of the fading factor ρ (SG-LRU equals LRU for $\rho \leq 0.5$; SG-LRU \rightarrow LFU for $\rho \rightarrow 1$).

Advanced Caching Simulations: Run Time versus Precision

We present simulation results for independent Zipf distributed requests (IRM) comparing the hit rates of LFU/LRU caching strategies. We start with examples for estimating the accuracy of results depending on the simulation run time, which is measured in number of requests. In a second part we provide an exhaustive study for Zipf distributed requests. A first simulation example demonstrates how deviations from a long term mean are decreasing with more run time.

We evaluate the hit rate h as the fraction of requests to items in the cache. Simulations start with an empty cache. While the cache is filling with objects, the score-gated and pure LRU caching strategies have the same stochastic behaviour. As soon as the cache is full, pure LRU already enters steady state regarding the set of items in the cache because the requests are independent. Consequently, the mean hit rate per simulated request with a full cache equals the long term LRU hit rate. Thus it is sufficient to exclude only the cache

filling phase as non-representative start phase for pure LRU simulations. Figure 11-18a shows simulation results for an example of Zipf distributed requests to $N = 10^6$ items with $\beta = 1$. Then a small cache size of $M = 200$ is sufficient to achieve 27.8% LRU hit rate.

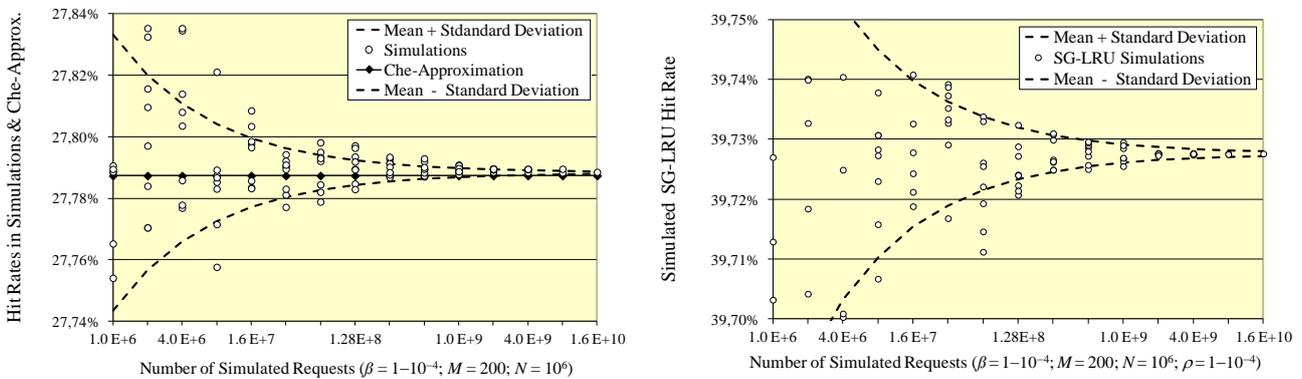


Figure 11-18: Simulation results for a) LRU and b) SG-LRU hit rates for different run times

Each dot in the figure refers to a hit rate simulation result with varying number R of included requests in a range $10^6 \leq R \leq 1.6 \cdot 10^{10}$. If each request would represent an independent experiment for a hit then a binomial distribution would be observed for the number of hits per simulation with mean $\mu_h = hR$ and standard deviation $\sigma_h = [h(1 - h)R]^{0.5}$. Figure 11-18 includes dotted curves showing $(\mu_h \pm \sigma_h)/R$, where the estimate μ_h for h is the hit rate obtained from the longest simulation run for $R = 1.6 \cdot 10^{10}$. 8 simulation runs are included for each number R of requests up to $R = 10^9$, where some results for short runs are beyond the shown range.

Results of an extended test series confirm that the long term mean hit rate is close to a 68-95-99.7 rule which corresponds to independent experiments, i.e. about 68% of the results fall into the range $(\mu_h \pm \sigma_h)/R$, about 95% into $(\mu_h \pm 2\sigma_h)/R$ and 99.7% into $(\mu_h \pm 3\sigma_h)/R$. In fact, hit rates are not independent for successive requests because the cache content can change only in one object per simulated request, but the hit rate variations due to changing cache content are usually small.

Results of a case study for score-gated LRU are shown in Figure 11-18b. Then the convergence of simulations depends on the development of scores for the items requiring much longer time than the cache filling phase after which LRU enters steady state. Therefore we exclude the first quarter of each simulation run from the evaluation of the hit rate. In Figure 11-18b, score-gated LRU is investigated for the same Zipf distributed IRM example as in Figure 11-18a. Geometrical fading scores are considered with a factor $\rho = 0.999$. Again, up to 8 simulation runs over R requests are shown with $10^6 \leq R \leq 1.6 \cdot 10^{10}$ in the evaluation phase.

In this case, deviations in the simulation results are also in the range $(\mu_h \pm \sigma_h)/R$ for independent requests because the distribution of scores shows sufficient convergence at least in order to prefer the $M = 200$ most popular objects for caching. The LRU simulation results in Figure 11-18 a are close to the Che approximation $h_{Che} \approx 27.787\%$ for the LRU

hit rate, whereas score-gated LRU exploits most of the LFU hit rate $h_{LFU} = z(1) + \dots + z(M) \approx 40.667\%$ as the maximum under IRM conditions which is achieved for $\rho \rightarrow 1$.

We consider another case of Zipf distributed requests with parameter $\beta = 0.6$, where requests to top items are less dominant and larger caches are required for a hit rate demand. Access to set of $N = 10^6$ items is supported by a cache of size $M = 200\ 000$. In this case the convergence of simulated LRU to steady state is again observed after the cache filling phase, whereas score-gated LRU needs more than 10^6 requests.

Figure 11-19 shows results for simulation runs of different length with R in the range $10^6 \leq R \leq 10^{10}$ and for different fading factors ρ . We observe that the hit rate is increasing with ρ from $h_{LRU} \approx 39.19\%$ to $h_{LFU} \approx 52.38\%$. SG-LRU stays close to h_{LRU} for $\rho \leq 1-10^{-5}$ and approaches h_{LFU} for $\rho \geq 1-10^{-8}$.

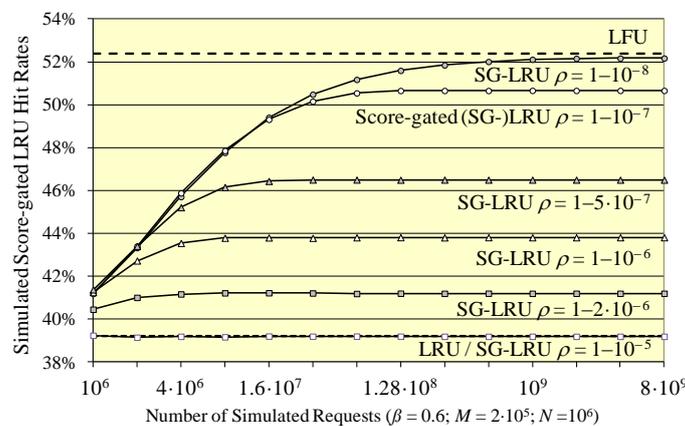


Figure 11-19: SG-LRU hit rate simulations with different run times

Exhaustive LFU / LRU Hit Rate Evaluations for Zipf Distributed IRM

We extend caching simulations over the basic range of Zipf distributed requests regarding

- the number N of objects,
- the cache size M ($M < N$) and
- the shaping parameter β of the Zipf distribution

as characteristic parameters for web caching. In particular, we evaluate the LFU and the LRU hit rate for each grid point in the relevant region covering

$$\beta = 0.4, 0.5, \dots, 1.1, N = 10^2, 10^3, \dots, 10^7 \text{ and } M = M_{1\%}, M_{2\%}, \dots, M_{99\%},$$

where $M_{x\%}$ is the minimum cache size required to obtain an LRU hit rate of $x\%$ depending on β and N . The Che approximation is applied to determine $M_{x\%}$ together with a final check through simulation.

On the whole, $8 \cdot 6 \cdot 99 = 4752$ simulation runs have been performed to cover this range. Each simulation includes at least 10^8 requests in the evaluation phase to keep the standard deviation below $5 \cdot 10^{-5}$. The simulation results are shown in Figure 11-20 for

object sets of size $N = 10^3$, and $N = 10^7$. Three graphs are shown for the results of both cases,

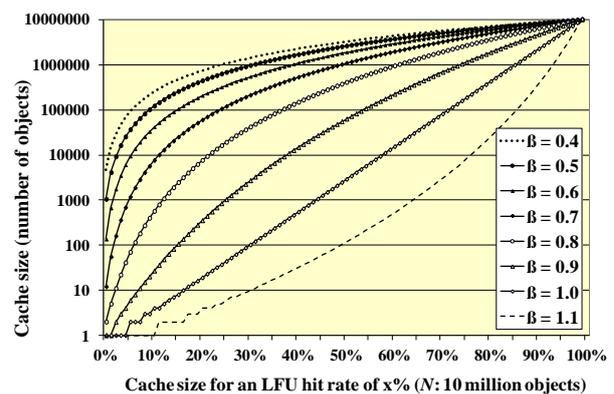
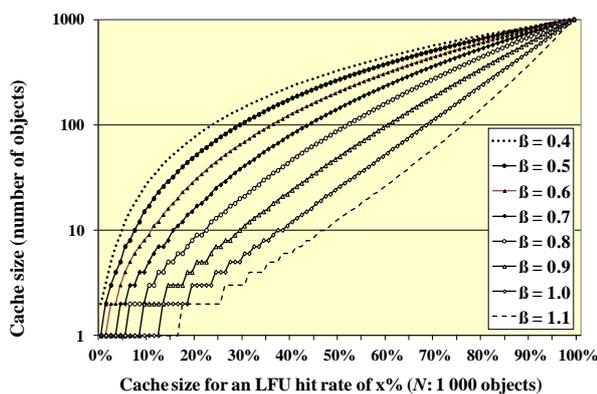
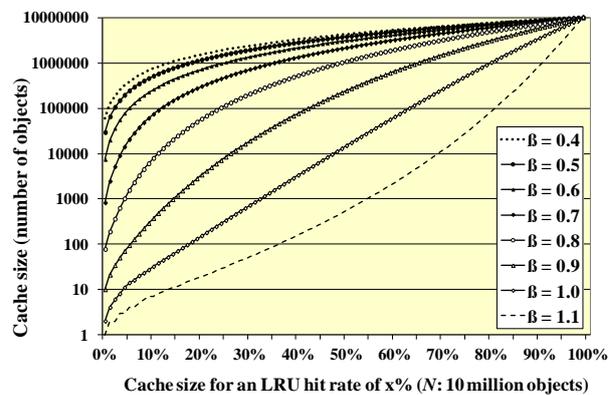
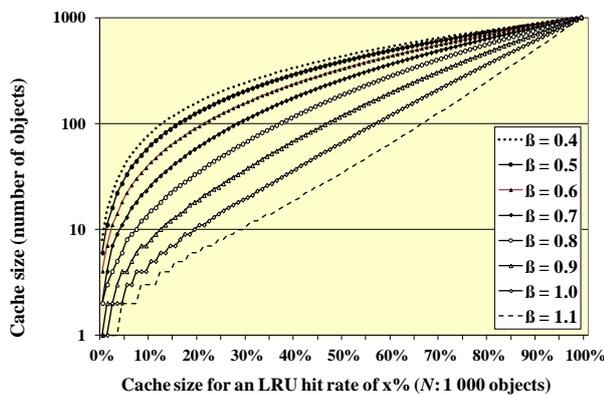
- for the cache size $M_{x\%}$ required to obtain $x\%$ LRU hit rate,
- for the cache size required to obtain $x\%$ LFU hit rate and
- for the gain $h_{LFU} - h_{LRU}$ of the LFU cache hit rate over LRU, which is again evaluated for cache sizes $M_{1\%}, M_{2\%}, \dots, M_{99\%}$.

We restrict the relevant range for web caching to

$$\beta = 0.5, \dots, 1.0, N = 10^2, 10^3, \dots, 10^7 \text{ and } M = M_{10\%}, M_{11\%}, \dots, M_{50\%},$$

assuming useful web caches to achieve hit rates of $>10\%$ as minimum efficiency and $<50\%$ due to cache size limitation. In this relevant range we obtain the gain $h_{LFU} - h_{LRU}$ of LFU over LRU:

$$9.6\% \leq h_{LFU} - h_{LRU} \leq 19.3\% \text{ and } h_{LFU} - h_{LRU} \approx 13.7\% \text{ on the average over all 6-6-41 cases.}$$



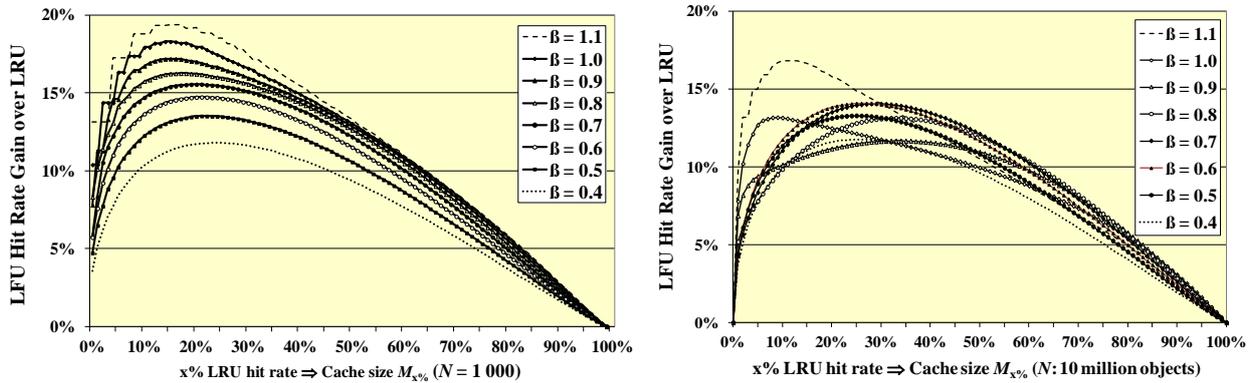


Figure 11-20: Comparing LFU/LRU hit rates for IRM Zipf requests

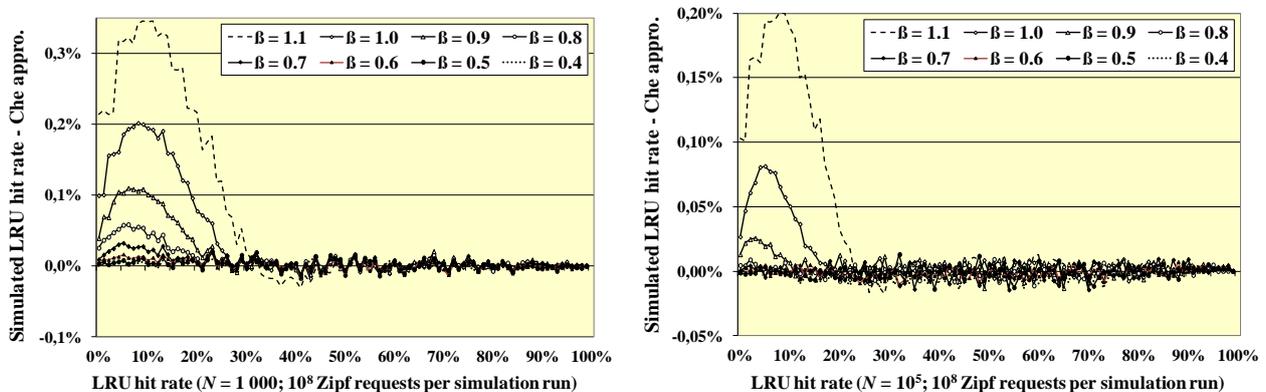
This roughly confirms a 10%-20% cache hit rate gain of LFU and statistics-based caching strategies like SG-LRU over pure LRU performance not only in some measurement based studies and extreme cases [30], but over the entire relevant range of Zipf distributed independent (IRM) requests for web caching.

The relative gain is obviously higher, most for small caches. When LRU hit rates are in the range 1 - 10% then LFU at least doubles the LRU hit rate achieving 10 - 25% for the same cache size M . Finally, we compare the Che approximation for LRU hit rates to the simulation and confirm a surprisingly high precision as experienced and investigated in [24].

Figure 11-21 shows the differences between simulation results and the Che approximation for the same cases as in Figure 11-20. In the relevant grid range

$$(\beta, N, M) \in \{0.5, 0.6, \dots, 1\} \times \{10^2, 10^3, \dots, 10^7\} \times \{M_{1\%}, M_{2\%}, \dots, M_{99\%}\}$$

for Zipf distributed requests we observe a maximum absolute difference of 0.4% and a majority of evaluated differences below 0.02%. Each simulation result is based on 10^8 or more requests and is expected to be subject to a standard deviation of less than 0.01% according to results for simulations with varying run time in Figure 11-19. Even longer simulations runs are partly necessary in order to evaluate the precision of the Che approximation. On the other hand, the largest deviations of the Che approximation are encountered for small cache sizes M , for a large number of items $N > 10^6$ and for $\beta \rightarrow 1$.



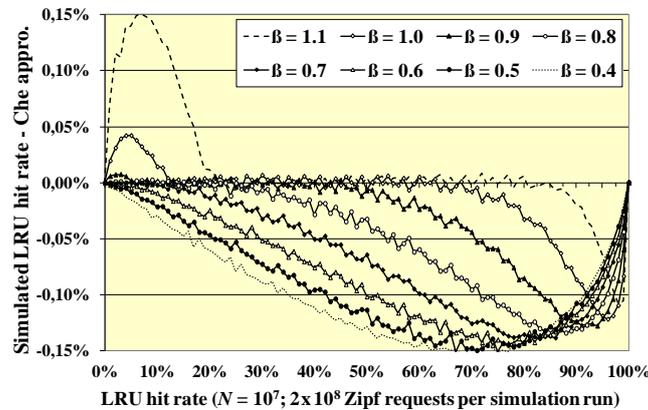


Figure 11-21: LRU hit rate for Zipf requests: Che approximation confirmed by simulation

11.5.4 Content Popularity Dynamics

- Numerous measurement studies have indicated that the popularity of Internet content varies over time. Most of these papers focus on user-generated content, especially videos, either uploaded on sharing platforms or embedded in online social networks [20][21][33]. Studies are also confirming this phenomenon in IPTV [31] as well as in P2P file-sharing systems. Some interesting conclusions can be drawn from these papers:
- Temporal popularity dynamics are studied on the scale of days, weeks, or months [20][21].
- The popularity evolution of each object is characterized by a rapid growth towards the maximum popularity value followed by a phase of slowly decreasing popularity [20][33].
- Therefore, when the uploaded content is young we notice significant rank changes starting from initially low popularity [21] until it becomes stable at the maximum.
- During the stable popularity phase, an oscillation of the popularity around a mean value in short time scales is observed even with bursts [20], especially for very popular content.
- Overall, most studies indicate low content popularity dynamics. For example, only 1-3% daily drift in the popularity of the top 100, 1 000 and 10 000 Gnutella files is reported [33].
- Another study focusing on Gnutella [20] confirms that more than 99% of the files in a dataset comprised from more than 6 million files have a relatively constant popularity profile. Similarly, in the measurement study [31] is reported that the cosine similarity value between the popularity of individual TV channels at day 0 and day 3 is about 0.97, thus indicating stable ranking of the content on the time scale of several days.

Extended Request Model for Popularity Dynamics Including New Items

Finally, we extend the Zipf distributed independent request model (IRM) to include new items over time with corresponding modifications in the ranking of items. It has been observed in many studies that new items appear with fast increasing phase towards maximum popularity followed by a long period of slowly decreasing popularity [31][33]. We

reflect this behaviour while preserving the Zipf distributed request pattern for requests to the current set of items.

In the extended model, each request to an item in the cache is modified with probability γ to address a new item. Consequently, the probability that a request refers to a current item in rank R is reducing to $(1 - \gamma) \cdot \alpha \cdot R^{-\beta}$, still following a Zipf distribution. Otherwise, a new item is assigned to an initial popularity rank R_{new} , which is uniformly chosen among the ranks of the currently cached objects. The insertion of the new object in the rank list is compensated by shifting the current item in this rank to the next rank $R_{new} + 1$ and consequently by shifting all items in a rank $R > R_{new}$ to the next rank $R + 1$ until the item in rank N , which is removed from the set of items.

In this way, the underlying Zipf distribution is preserved for each request to the set of current items, if no new object is introduced. Considering a single item, it starts at an initial Zipf rank R_{new} on its highest popularity level. Afterwards the rank is incremented each time when a new item is inserted at the same or higher popularity, thus reducing the popularity and request probability of the shifted items step by step. Sooner or later the item is downgraded to rank N and finally is removed when the next new item appears.

Compared to the usually observed popularity profile of web items over time, this model omits the fast ramp up phase for new items and thus is more disruptive regarding rank changes for increasing popularity but reflects the long phase of slowly decreasing popularity. The model combines the IRM for $\gamma = 0$ with the introduction of new items at a rate γ per request. It can be refined to also closely follow a predefined ramp up phase profile of growing popularity and it can be applied to arbitrary other request distributions instead of Zipf requests.

In our implementation, we do not insert new items during an initial cache filling phase and we again exclude the first quarter of requests from the evaluation of a simulation run in order to focus on steady state behaviour also for dynamically changing items.

The consequence of dynamic changes in the set of requested items on cache efficiency is a decrease in the hit rate, because a request to a new object is no hit. For the LRU strategy, the cache hit rate in this model is given by $h_{LRU} = h_{LRU,IRM} \cdot (1 - \gamma) \approx h_{Che} \cdot (1 - \gamma)$, i.e. the hit rate is decreasing with the factor $1 - \gamma$ starting from the hit rate obtained for independent requests $\gamma = 0$, which is closely approached by the Che approximation. Again, LRU is in steady state as soon as the cache is full and insertion of new items has started.

For score-gated LRU the decreasing effect of dynamically changing items on the hit rate is more severe, because new items gaining high request probability $p_{new} = \alpha \cdot (R_{new})^{-\beta}$ still will need several requests until they reach the score level of the bottom item of the cache and then can enter the cache. When we assume e.g. sliding window based scores and a new item has been present only for the last w requests with $w < W$ (W : window size) then the expected score $(1 - \gamma) \cdot p_{new} \cdot w$ is lower than $(1 - \gamma) \cdot p_{new} \cdot W$ for an "old" item.

In order to overcome this handicap for new items, we use the ratio of the request count and the age of an item as its score, where the age is measured by the number of requests to all items since the first request to the new item. The score function keeps the maximum IRM hit rate and leads to better performance for dynamically changing objects than sliding

window or geometric fading scores. The simulation results in Figure 11-21 confirm SG-LRU to be more efficient than LRU also for changing items but the advantage is decreasing with higher dynamics γ . LRU and SG-LRU are compared with the proposed score function for Zipf distributed requests with $\beta = 0.6, 0.8, 0.9999$ for $N = 10^5$ items. The cache size is $M = 2000$ for $\beta = 0.6$ and $M = 1000$ for both other cases.

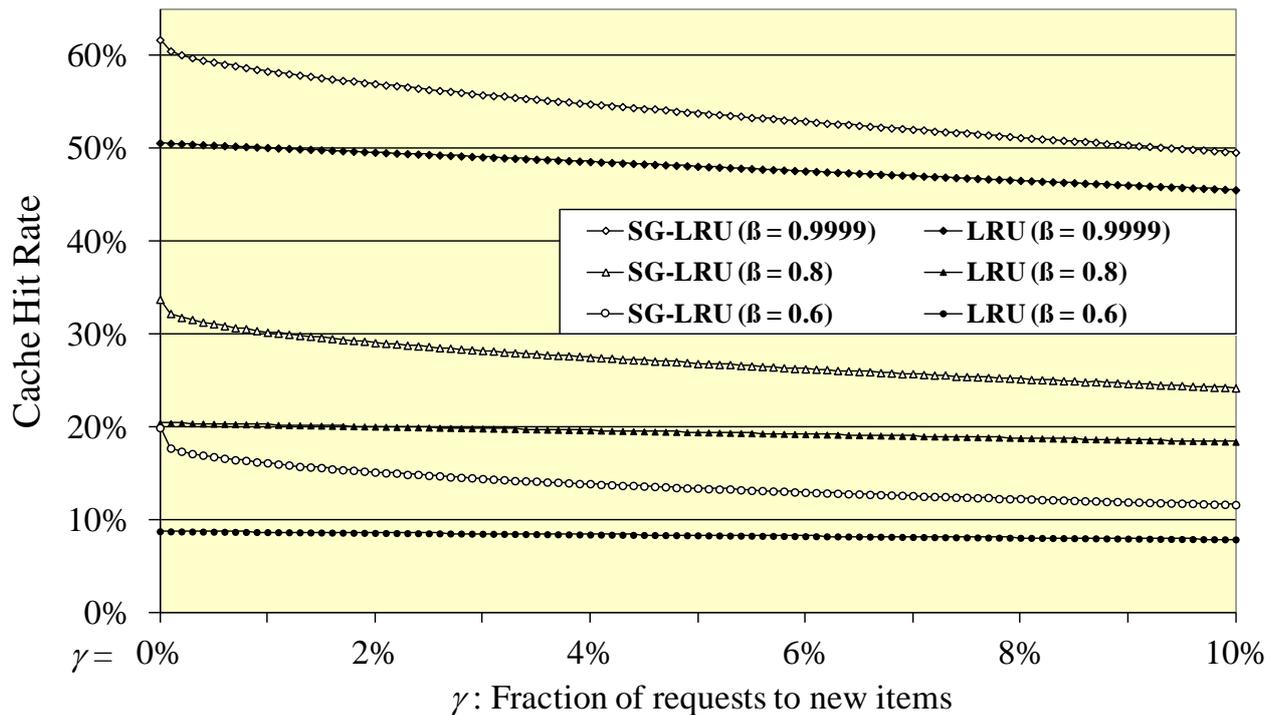


Figure 11-21: Impact of Dynamic Popularity Changes Including New Items on Hit Rates

The evaluation covers the range $0 \leq \gamma \leq 10\%$. To the author's knowledge, there are only few publications providing clear results on the dynamics of web items [31][33]. They indicate that only a small fraction of about 2% of the top-M most popular items are changing per day. When we assume that a cache is processing K requests per day then a dynamics rate $\gamma = 0.02 \cdot M / K$ produces a 2% daily exchange in the top M items. For a large user community processing a million requests per day and for a cache of size $M = 1000$, we obtain $\gamma = 2 \cdot 10^{-5}$ or for a smaller population generating 1 000 requests to cached items per day, we obtain $\gamma = 0.02$. For both cases, the SG-LRU and the LRU hit rates are still close to the IRM hit rate. We conclude that the dynamics rate for new items in web caching is usually small enough to be neglected and that the IRM model ($\gamma = 0$) mainly determines the caching efficiency. Therefore the IRM results for SG-LRU and LRU in the Appendix are confirmed to be characteristic for web caching.

Conclusions

In this work we evaluate a score-gated LRU (SG-LRU) caching strategy which

- has update effort comparable to pure LRU
- is capable of including count statistics and
- is flexibly adaptive based on one parameter between LFU and LRU as extreme cases.

We use efficient simulation to show that the absolute hit rate gain of LFU over LRU is in the range of 9.6%-19.3% over the entire relevant parameter set for independent (IRM) Zipf distributed requests. We conclude that SG-LRU essentially outperforms pure LRU going even beyond twice the hit rate for small caches by combining the best of LFU and LRU strategies. Our study also confirms high precision of the Che approximation for LRU hit rates, which has been shown analytically but without clear precision bounds in [24].

The simulations are extended to study the impact of new items being encountered at a given rate γ on the caching performance. They indicate that the rate of change in item sets for web caching is usually small enough to yield hit rates close to the presented IRM results.

11.6 Appendix F: Pricing

This Appendix addresses pricing that could be supported by SmartenIT also in lieu of the SmartenIT business models presented in Deliverable 2.4[7]. We begin by defining the layers of charging, namely the *cloud/application layer*, the *bulk-data layer* and the *per-flow layer*, thus providing a “map” for classifying concrete pricing proposals. Then we overview a SmartenIT pricing proposal related to the Operator Focused Scenario and in particular the ICC mechanism.

11.6.1 Pricing Layers

The *cloud/application layer* comprises charging for cloud resources and/or per-service charge (e.g. per movie on a Video on Demand application). Cloud resources are typically priced with a fixed fee per unit and per type of resource and customers pay only for what they actually use [50]. Elastic scaling of resources is supported so as to accommodate demand spikes with resource policies that allow the dynamic reservation of additional resources, such as Virtual Machines of a certain type. Cloud storage, such as Google Cloud Storage typically adopts a fixed fee per GB per month pricing model [51]. Regarding the network layer, charging mechanisms commonly adopted by ISPs today in the Internet access retail market are either flat-fee, with periodic charges for a speed tier, independent of usage, or usage-based pricing e.g. with volume caps, or sometimes a combination of these two models. Both models are normally independent of quality of service and only uptime guarantees are provided.

Pricing schemes need to overcome two main obstacles: *service coordination* and *business coordination* [52]. Service coordination involves ensuring technical capabilities to fulfil requests with the required characteristics. Business coordination, on the other hand, involves the identification of sustainable business models which for example enable value/revenue and cost apportionment among the stakeholders. *Service coordination* may be interpreted as spanning across two main control layers: *bulk- data* and *flows*. The *bulk-data layer* includes functions such as rate policing and path characterisation while the higher *per-flow layer* encompasses functions such as rate control and admission control.

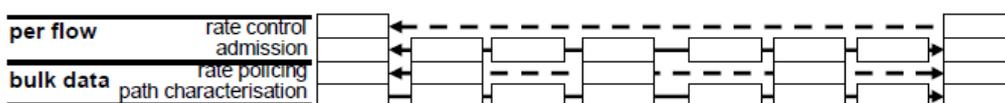


Figure 11-22: Pricing layers for the network [52]

To the **bulk-data layer** there can correspond two charging layers: 1) capacity charging independent of usage, and 2) capacity charging dependent on usage; and at the **per-flow layer** there corresponds per-session charging. Inter-cloud and inter-network data communication are classified to the **bulk-data layer** so as to ensure scalability and OPEX reduction for the network management and pricing functions. Flat-fee pricing offers advantages for both for the ISP and the end-user as it provides predictable revenues for the ISP and costs for the end-users. In general, end-users have strong preference for predictable fees and are prepared to pay a considerable premium for this [53]. The main

drawback with flat-fee pricing is that it does not motivate users to limit usage at times of congestion. The distribution of scarce network resources will therefore be determined by competition between network protocols, controlled by the cloud/content and service providers in general. Flat-fee pricing furthermore does not allocate marginal revenue incentives to the ISP when usage grows. The only up-sale opportunity will be to a higher speed tier, and this may require large investments in infrastructure.

To overcome these limitations without engaging in the less preferred general volume capping, ISPs have experimented with various alternatives. Some have also applied deep packet inspection to determine the nature of different traffic flows and applied this to limit flow rates in general or in combination with more specific volume or time limitations. At the *wholesale level*, pricing operates on aggregate usage from multiple individual flows. It can depend on different metrics such as:

- Provisioned link capacity or bandwidth (maximum number of bits per second)
- Consumed data volume over the billing period (in total number of bytes, or average bits per second)
- 95th percentile of 5 minutes traffic (in bits per second)
- Geographical scope of network connectivity (e.g. continents)
- External interconnection capacity with third-party networks
- Specific charging events, e.g. observed in user session signalling

A drawback with a simple volume-based pricing is that it does not account for the constantly varying nature of network utilization and ISPs are not encouraged to incentivize their end-users to use the network sensibly. *The 95th percentile peak-demand method and time-of-day bandwidth-pricing are alternative pricing regimes that give carriers (and potentially in turn their customers) incentives to smooth out their traffic and hence reduce peak-time network usage and congestion.*

11.6.1.1 A SmartenIT Pricing Proposal: ICC Pricing

SmartenIT has proposed a set of mechanisms for resource allocation and traffic management. The design goals of these mechanisms typically include OPEX reduction, efficiency, incentives, cost reduction or fairness. It is clear that the suitable pricing model for each of these mechanisms is correlated with the type of application and respective business model envisioned, as well as the respective pricing layer where it can be classified to. In this subsection we restrict attention to one of these mechanism, namely ICC. ICC – and also DTM and DTM++ -are tailored for the EFS scenario and thus can be used at the *bulk-data (wholesale) layer*. We now proceed to present the SmartenIT approach regarding ICC pricing pertains to inter-cloud and inter-network communication which is in line with the design-for-tussle principles and incentive compatibility and discuss interesting issues regarding also other candidate pricing options.

The key feature of ICC is to incentivize the key ISP business customers that generate the time-shiftable traffic (e.g. some inter-cloud/-datacenter traffic) to allow ICC to manage the time-shiftable portion of their traffic through proper pricing. The specification of the pricing scheme for ICC affects our mechanism's incentives for stakeholders, complexity and

scalability. In particular, we specify that the ISP will return a cut p , where $0 \leq p \leq 1$, of the total transit cost savings $ISP_{savings}$ attained due to ICC, to each one of his customers $i \in I$ that accepted ICC traffic management for a portion of their traffic of volume vol_i as a discount to their billing charge by the ISP. Hence, each customer will receive a discount proportional to his fair share on the total volume of traffic managed by ICC, as depicted in equation (1).

$$discount_i = p \cdot ISP_{savings} \cdot \frac{vol_i}{\sum_{j \in I} vol_j} \quad (1)$$

Firstly, we comment on whether the benefit sharing rule of equation (1) should be computed ex post or instead an estimate of its value should be a priori announced to the ISP business customers. Clearly both options are viable, as long as the expected or observed ex post discount per billing period (typically month) is substantial enough to motivate some datacenters/clouds to mark a portion of their traffic as time-shiftable and this discount is also economically viable for the ISP. We advocate in favour of the a priori announcement of an expectation of this value, rather than its ex post computation at the end of the billing period, so that customers can have a clear incentive to allow ICC operation and know beforehand the discount they will acquire. This a priori announcement also requires that the ISP can indeed estimate the expected $ISP_{savings}$. This is indeed feasible since the ISP is aware of the volume of traffic injected per business customer, the transit cost incurred when ICC was not applied and can compute the cumulative aggregate growth rate (*cagr*) of the traffic from that past point in time. Hence, the ISP could deduce that without the ICC operation the expected transit charge would increase proportionally to the increase of the total volume of traffic denoted by *cagr*.

Formula (1) has been chosen since it is increasing in the volume of data offered to ICC, it is simple, and it does not impose additional requirements for traffic metering or monitoring especially at per-flow level. This is an *absolute must* for aggregate flows and it prevents us from adopting more fine-grained pricing proposals with potential that could more accurately compute the load contribution/relative weight of the per-flow manageable traffic injected at the peak periods and affecting the 95th percentile rule. Therefore, sophisticated tools such as the effective bandwidths theory [54] cannot be applied in practice: Though it would be theoretically possible to use the effective bandwidth theory at the operating point of the transit link border router to perform essentially call admission control per interval for the flows of manageable traffic that would be allowed to cross it, this would result in tedious computations and would require inspecting the traffic at per-packet layer which is extremely cost for network operations and not scalable. Therefore, we have opted for simpler yet practically feasible solutions that respect ISPs' business models and network management best practices.

This fundamental requirement for simplicity and low overhead also affects how we envision that $ISP_{savings}$ could be computed. The main issue here is that computing accurately the value of the resulting transit link 95th percentile under both the presence and absence of ICC, denoted as $perc_{ICC}$ and $perc_{BE}$, would be infeasible in practice since the operation or absence of ICC inevitably alters the traffic profile and the resulting charge in a real ISP network. This is due to the fact that removing some time-shiftable traffic from a link at some time would inevitably alter the way the non-manageable by ICC flows crossing this

link would adapt their rates so as to adapt to the different conditions, e.g. due to TCP rate control. Since such information is inevitably lacking and in order to keep the computational complexity low we prescribe that $perc_{BE}$ can be computed through traffic statistics by multiplying the $perc_{BE}$ of a billing period of a previous year when ICC was not operating times the cumulative aggregate growth rate. This is a fair approximation, also due to the long-lived time-of-day and day-of-week traffic patterns that are inherent to the ISP traffic patterns.

Additional options would be to have a different definition for vol_i and consider it to be the volume of time-shiftable traffic that customer i has during the intervals and epochs where the ISP actually needs to shift traffic in order not to exceed the C_{target} value or even specify it to be the value of the traffic that was *actually shifted in time* by the ISP resulting in throughput degradation. There are two main reasons that such variations of the rule in equation (2) should be avoided: The first is *simplicity and scalability*, as also previously explained: the total volume of traffic injected in the ISP network is typically metered by the ISP also for billing and accounting reasons. The second main reason is *incentives*: Providing higher discounts for customers who inject more time-shiftable traffic in epochs where the transit link is heavily utilized - and thus ICC operates to shift traffic - would provide the incentive for bargain-hunter business customers to change their traffic patterns so as to inject the time-shiftable traffic when it is least desired by the ISP and most harmful for all users to do so: during peak periods. On the contrary, equation (1) does not provide such an incentive and thus is preferred.

Concluding, the pricing scheme associated with the ICC mechanism is a good compromise between simplicity and incentives, also respecting current ISP operational and business practices on the one hand and fairness and proper incentives on the other.