SLACC
SLA Support System for Cloud Computing

Guilherme Sperb Machado, Burkhard Stiller
Department of Informatics IFI, Communication Systems Group CSG,
University of Zürich UZH
machado | stiller @ifi.uzh.ch

Motivation and Problem
Use Cases
SLACC Process
System Architecture

Motivation

- ❑ "Companies struggling with Cloud performance" [1]
  - – Survey from Compuware, by Vanson Bourne:

Reference [1]: http://www.computerworlduk.com/news/cloud-computing/ 3239390/companies-struggling-with-cloud-performance/

Motivation

❑ "Companies struggling with Cloud performance" [1]

  – Survey from Compuware, by Vanson Bourne:

**57% European businesses were stopping
further investments on Cloud Computing until
they provide more specific guarantees**

Reference [1]:   http://www.computerworlduk.com/news/cloud-computing/
3239390/companies-struggling-with-cloud-performance/

Motivation

- "Companies struggling with Cloud performance" [1]
  - Survey from Compuware, by Vanson Bourne:

> **57% European businesses were stopping further investments on Cloud Computing until they provide more specific guarantees**

> **72% of businesses: cloud platform was hampering their ability to maintain set levels of service, also affecting company's revenue**

Reference [1]:   http://www.computerworlduk.com/news/cloud-computing/3239390/companies-struggling-with-cloud-performance/

Motivation

- "Companies struggling with Cloud performance" [1]
  - Survey from Compuware, by Vanson Bourne:

57% European businesses were stopping further investments on Cloud Computing until they provide more specific guarantees

72% of businesses: cloud platform was hampering their ability to maintain set levels of service, also affecting company's revenue

For revenue-generating websites: performance would mean revenue

Reference [1]: http://www.computerworlduk.com/news/cloud-computing/3239390/companies-struggling-with-cloud-performance/

Cloud Provider	Service
Amazon	S3
	EC2
	SimpleDB
SalesForce	CRM
Google	Google Apps (inc. Gmail business, Google Docs, etc.)
Rackspace Cloud	Cloud Server
	Cloud Sites
	Cloud Files

Cloud Provider	Service	SLA Parameters
Amazon	S3	**Availability** (99.9%) with the following definitions: Error Rate, Monthly Uptime Percentage, Service Credit
	EC2	**Availability** (99.95%) with the following definitions: Service Year: 365 days of the year, Annual Percentage Uptime, Region Unavailable/Unavailability, Unavailable: no external connectivity during a five minute period, Eligible Credit Period, Service Credit
	SimpleDB	Subject to the Amazon Web Services Customer Agreement, since no specific SLA is defined. **Such agreement does not guarantee availability**.
SalesForce	CRM	The company's Web site does not contain information regarding SLAs for this specific service.
Google	Google Apps (inc. Gmail business, Google Docs, etc.)	**Availability** (99.9%) with the following definitions: Downtime, Downtime Period: 10 consecutive minutes downtime, Google Apps Covered Services, Monthly Uptime Percentage, Scheduled Downtime, Service, Service Credit.
Rackspace Cloud	Cloud Server	**Availability** regarding the following: Internal Network: 100%, Data Center Infrastructure: 100% **Performance** related to service degradation: Server Migration in case of performance problems: migration is notified 24 hours in advance, and is completed in 3 hours (maximum). **Recovery Time**: In case of failure, guarantee the restoration/recovery in 1 hour after the problem is identified.
	Cloud Sites	**Availability**, Unplanned Maintenance: 0%, Service Credit.
	Cloud Files	**Availability**: 99.9%, Service Credit.

Problem

- ❑ Cloud Providers do not offer/guarantee
  - – *SLA specification tailored to Cloud Users' interests*
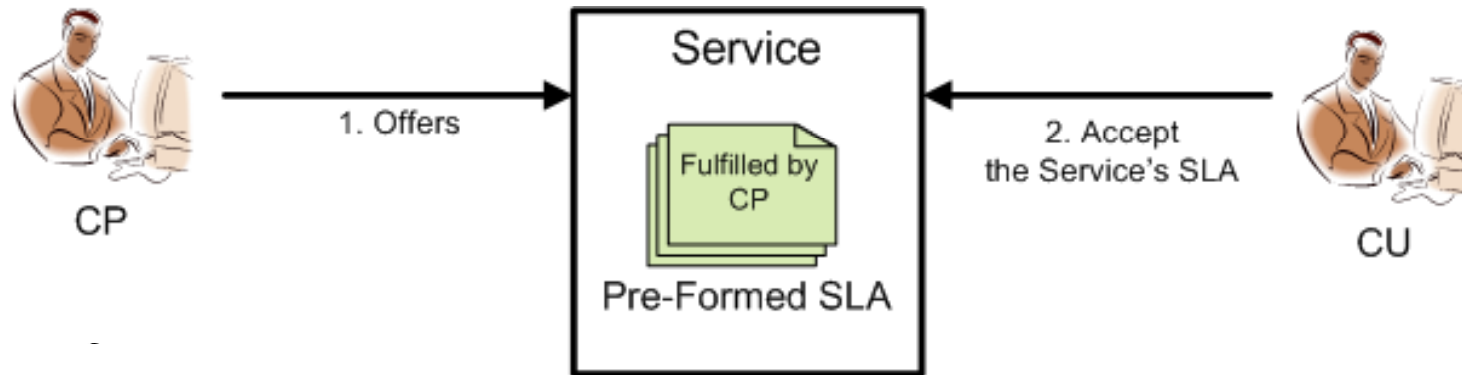   - • Mostly, "Service Availability"

Problem

❑ Cloud Providers do not offer/guarantee
  – *SLA specification tailored to Cloud Users' interests*
    • Mostly, "Service Availability"

❑ *Cloud Providers offering performance parameters*

Problem

- ❏ Cloud Providers do not offer/guarantee
  - – *SLA specification tailored to Cloud Users' interests*
    - • Mostly, "Service Availability"

- ❏ *Cloud Providers offering performance parameters*
  - – The solution is not obvious
    - • Huge size of Providers' IT Infrastructure
    - • High complexity with multiple inter-dependencies of resources (physical or virtual)
    - • Diversity of performance parameters

Solution Approach

- ❑ SLACC: SLA Supporting System for Cloud Computing
  - – Estimate SLA parameters (KPIs and SLOs) in a formalized methodology based on
    - • Historical data (and the lack of data, as well)
    - • IT infrastructure information (dependency between components)
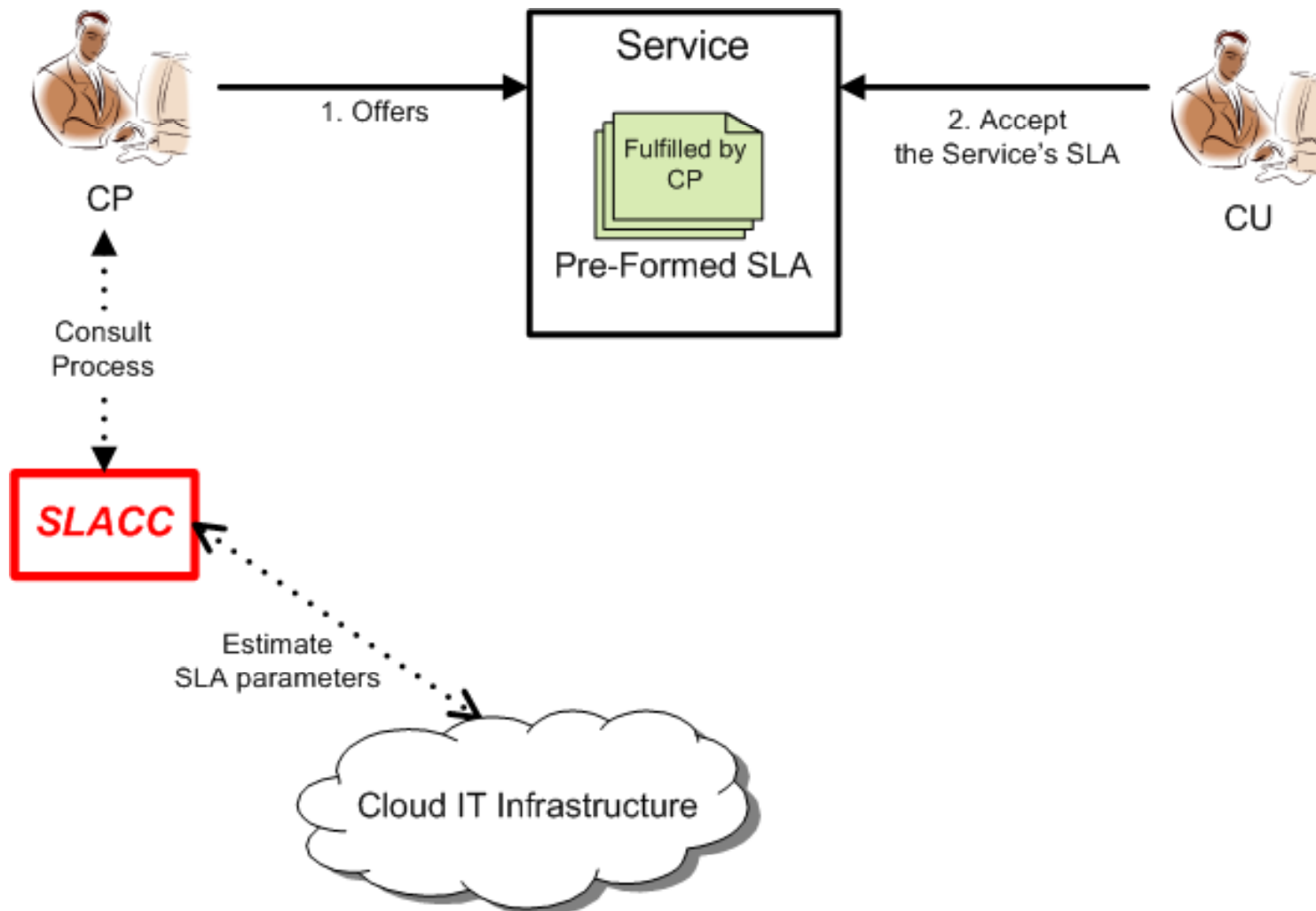 - – Focusing on performance parameters

- ❑ The benefits:
  - – Enhance the level of SLA specificity
  - – Decision support in SLA negotiation processes (CPs)
  - – Better knowledge of IT infrastructures' capabilities
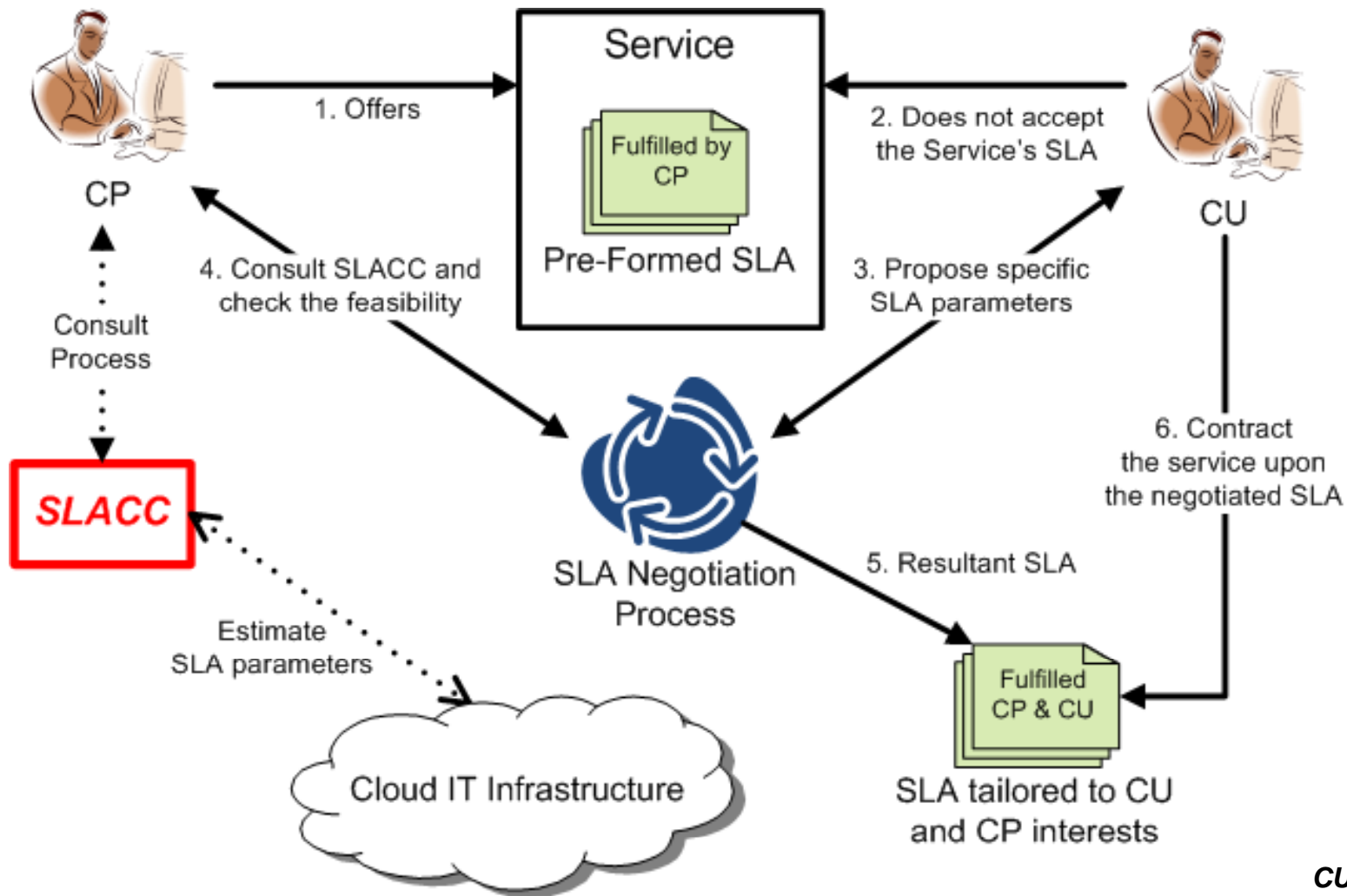
Use Cases (1)

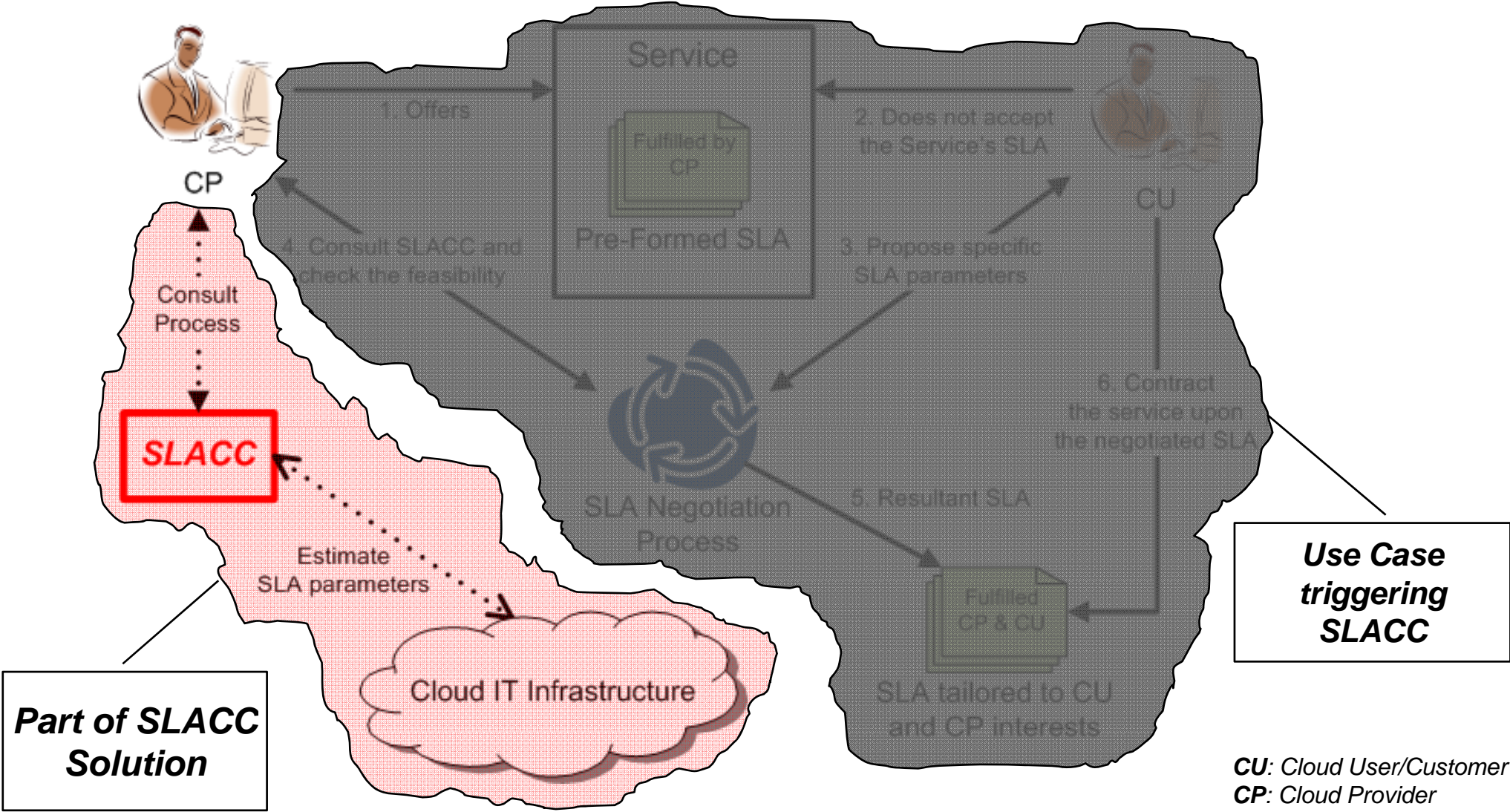*CU*: Cloud User/Customer
CP: Cloud Provider

Use Cases (1)

CU: Cloud User/Customer
CP: Cloud Provider

Use Cases (1)

CU: Cloud User/Customer
CP: Cloud Provider

Use Cases (1)

Part of SLACC Solution

Use Case triggering SLACC

CU: Cloud User/Customer
CP: Cloud Provider

Use Cases (2)

❑ SLACC handles typical Cloud Computing estimation cases in different levels (IaaS, PaaS, SaaS)

  – <u>Response time </u>of an operation (*e.g.*, query data, insert new customers) from a CRM application (Customer Relationship Management)

  – <u>Deployment time</u> of a specific Virtual Machine template provided by the Cloud provider

  – <u>Backup time completion </u>of several VM instances

  – <u>Minimal bandwidth</u> between VM instances (in different geographical localities)

 – <u>Minimal CPU processing </u>capacity for a given VM

Use Cases (2)

❑ SLACC handles typical Cloud Computing estimation cases in different levels (IaaS, PaaS, SaaS)

- <u>Response time </u>of an operation (*e.g.*, query data, insert new customers) from a CRM application (Customer Relationship Management)

- <u>Deployment time</u> of a specific Virtual Machine template provided by the Cloud provider

- <u>Backup time completion </u>of several VM instances

- <u>Minimal bandwidth</u> between VM instances (in different geographical localities)

- <u>Minimal CPU processing </u>capacity for a given VM

Use Cases (2)

- Response time of an operation (*e.g.*, query data, insert new customers) from a CRM application (Customer Relationship Management)

  - (example) Cloud Customer requires the information retrieval in less than 1 second, having 50.000 clients at the database

  - Composed of measurements:
    - time of distributing HTTP requests (load balancing distribution)
    - time that the application (CRM) can process the request
    - time of establishing a database connection
    - time to perform the SELECT on the "users table" (learned from populated databases)

SLACC Process

SLACC Process

SLACC Decision Support System
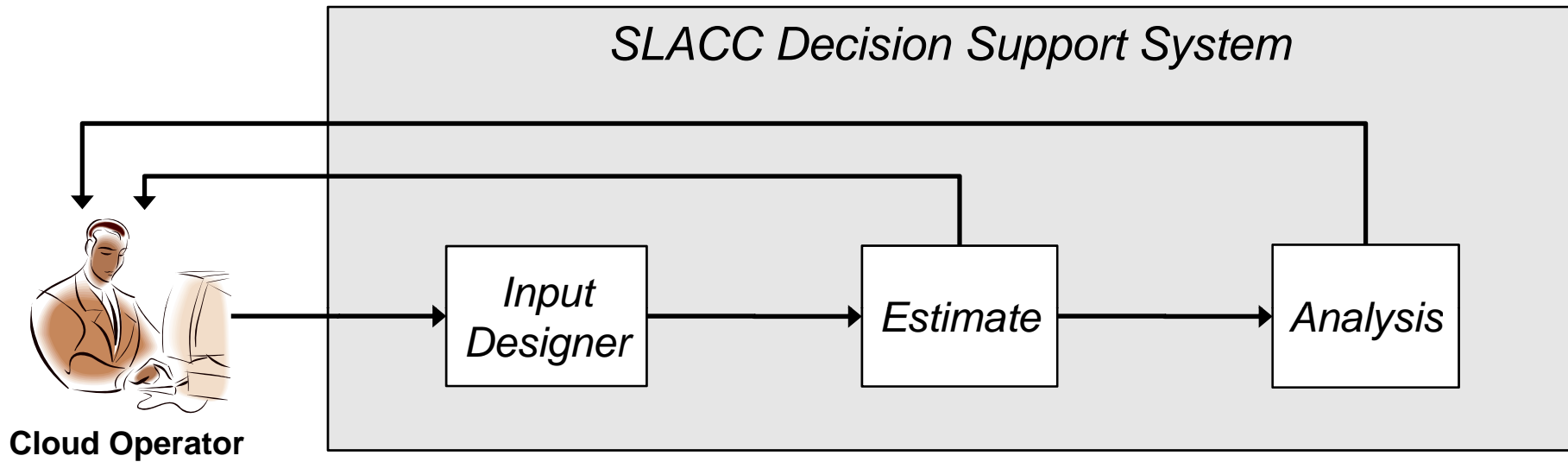
Cloud Operator → Input Designer → Estimate → Analysis
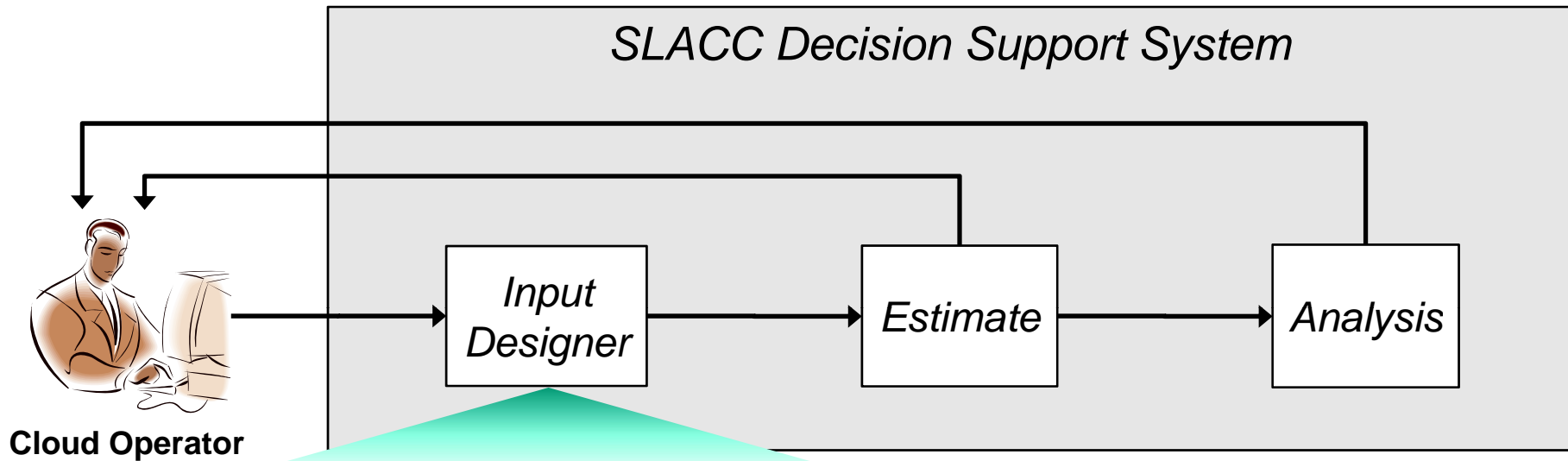
R	Time that the Load balancing distributes the R	Time that Application process R	Time to Est. DB Conn	Time to perform SELECT on Users table
1	0.012	0.123	0.050	1.150
2	0.056	0.100	0.073	1.012
3	0.023	0.223	0.098	1.344
4	0.028	0.145	0.012	0.983
5	0.043	0.245	0.033	0.974
…	….	…	…	…

SLACC Process
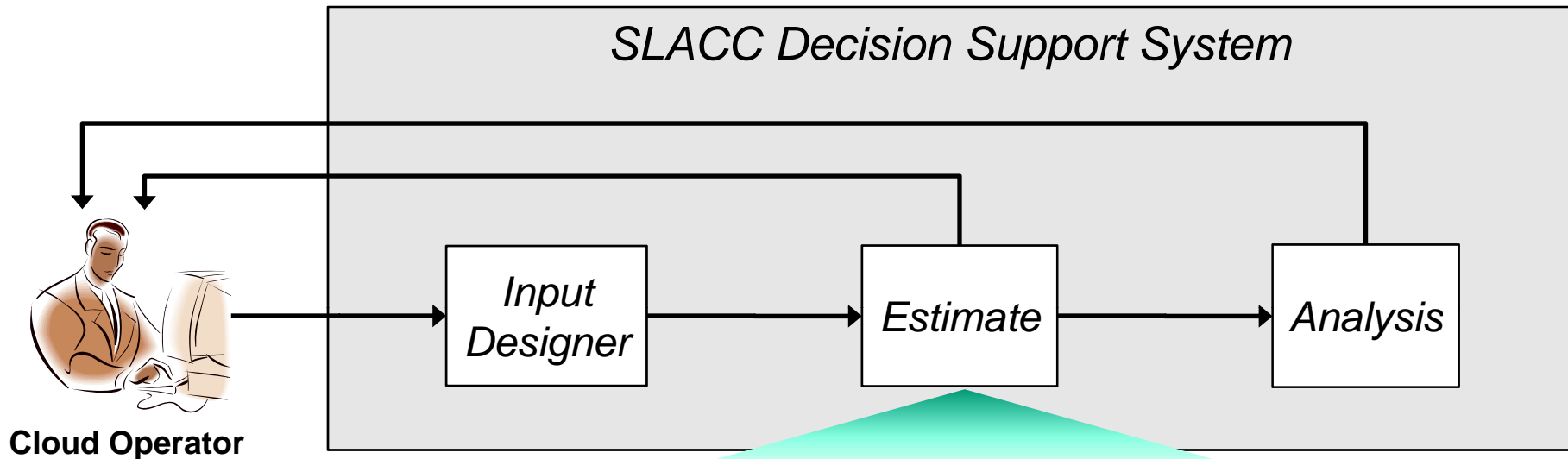
SLACC Decision Support System

Cloud Operator

Input Designer → Estimate → Analysis

Correlation:
- Correlate some variables to check how significant they are for the estimate
- E.g., the time of load balancing of HTTP Requests has an influence of X%

Regression:
- Come up with a function that also gives point interval values

Hypothesis testing

SLACC Process

Cloud Operator

SLACC Decision Support System

Input Designer → Estimate → Analysis

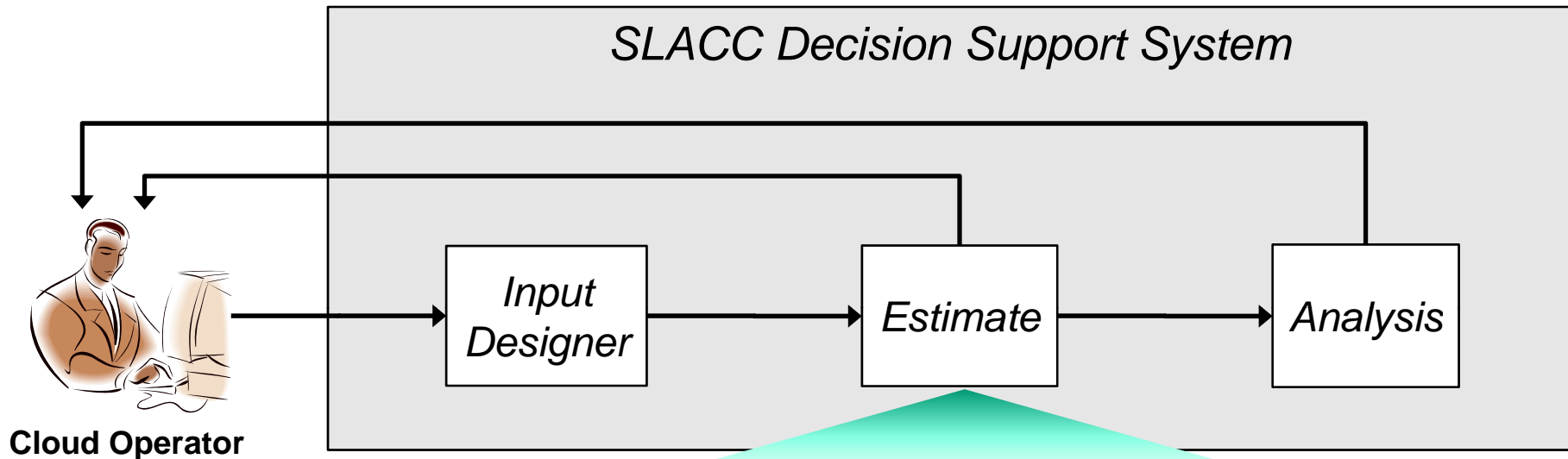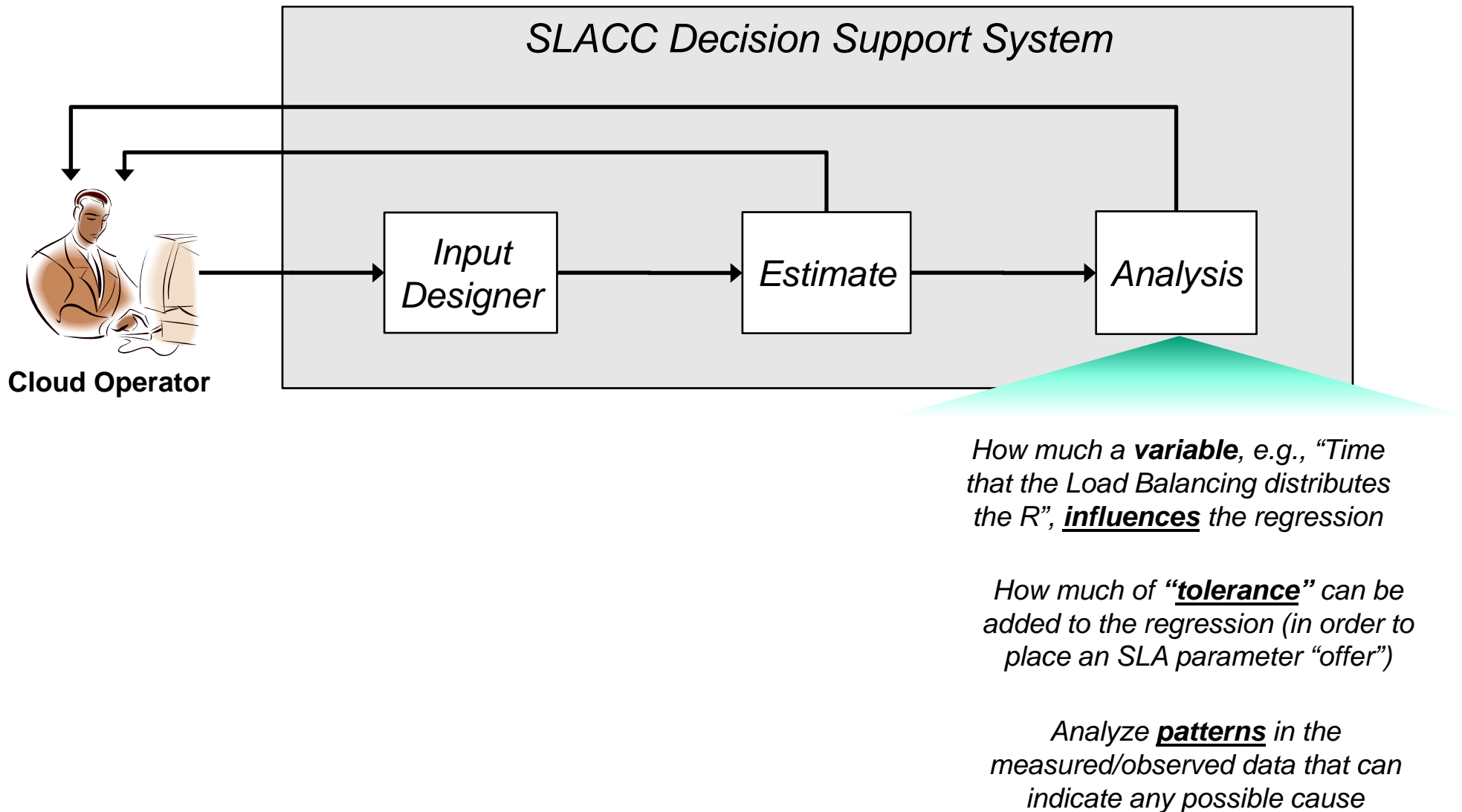*on-going work*

Correlation:
- Correlate some variables to check how significant they are for the estimate
- E.g., the time of load balancing of HTTP Requests has an influence of X%

Regression:
- Come up with a function that also gives point interval values

Hypothesis testing

SLACC Process

SLACC Decision Support System

Cloud Operator

Input Designer → Estimate → Analysis

*How much a **variable**, e.g., "Time that the Load Balancing distributes the R", **influences** the regression*

*How much of **"tolerance"** can be added to the regression (in order to place an SLA parameter "offer")*

*Analyze **patterns** in the measured/observed data that can indicate any possible cause*

System Architecture

Legend:
- A ←→ B : A and B exchange information
- A → B : A provide information to B
- A ⤳ B : A reflect B

CP Operator

SLA Designer

SLACC
- Estimates Repository
- SLA Decision Support System

Accounting Records Repository

SLAs Repository

Accounting and Monitoring System

SLA Monitoring System

Infrastructure Model

Cloud IT Infrastructure

CP: Cloud Provider

Summary

- ❑ Estimate SLA parameters in order to evaluate **what Cloud Providers will be able to offer/accept as SLOs or KPIs**
  - – Analyzing *historical data*, *current information about IT infrastructure, and considering possibly changes*

- ❑ SLACC, Decision Support System
  - – It aims to be part of the system without interfering in the current Cloud IT architecture
  - – Work with typical Cloud Computing performance parameters
  - – Service-Oriented