Towards Path Quality Metrics for Overlay Networks

Andri Lareida, Daniel Meier, Thomas Bocek, Burkhard Stiller
Communication Systems Group CSG@IfI, University of Zurich
Binzmühlestrasse 14, CH-8050 Zürich, Switzerland
Email: [lareida|bocek|stiller]@ifi.uzh.ch, dani@two-options.info

***Abstract*—Decentralized systems, such as overlay networks or content delivery networks, form a virtual network on top of a physical network leading to inefficiencies in handling network traffic as physical locations are not considered. Therefore, this work investigates, analyzes, and evaluates two feasible network locality metrics for decentralized systems to overcome such inefficiencies. Furthermore, those metrics to determine network locality have been implemented and evaluated within an international testbed. These metrics evaluated consider aspects of path stability and path symmetry. The evaluation performed did reveal that the stability and symmetry metrics provide accurate locality indication.**

I. Introduction

Today, there are different types of decentralized systems. Prominent examples are Peer-to-Peer (P2P) networks and Content Distribution Networks (CDN). The property they have in common is that they build a virtual network, termed overlay network, on top of an existing physical network, termed underlay network. Overlay networks are typically not aware of the underlay network topology leading to inefficient overlay network topologies or operations, such as long paths. As a consequence, this can lead to degraded service quality for overlay users. Depending on the purpose of the overlay network, this can be experienced in terms of increased latency and reduced bandwidth. Furthermore, involved Internet Service Providers (ISPs) have to handle unnecessary traffic, which leads to higher operating expense. Those inefficiencies are a result of overlay networks that select partner hosts without consideration of the underlay network.

The aforementioned inefficiencies can be reduced by introducing locality metrics, which represent the "distance" between two hosts. Feasible approaches tend to represent the characteristics of the underlay network between two hosts. Therefore, this paper investigates metrics that are generally available to overlay network nodes to determine the underlay distance. As a main contribution this work proposes two network locality metrics for decentralized systems and evaluates those based on traceroute data collected throughout Europe within the EMANICSLab [2] infrastructure.

II. Related Work

To optimize overlay network topology, information about the structure of the underlay network is required. The network topology itself can be described on three different levels [1]: link layer topology, internet layer topology, and overlay topology. To improve overlay network topologies, internet layer topology information can be used since it can be discovered across domains.

Traceroute is a widely used internet diagnostic tool which discovers hops between a source and a destination node [5]. Traceroute discovers the *ingress interface* of routers and, thus, does not provide any visibility into the egress interface as well as the return paths of the ICMP messages. Since Internet routing does not guarantee a symmetric return path, traceroute can only provide the forward path hops. IP alias resolution [4] is one way of overcoming this limitation by identifying the IP addresses that belong to the same router. However, those techniques require too much effort to be used in lightweight locality metrics.

The measured RTT for each hop comprises of the time the forward probe required, the time for ICMP generation on the router and finally the reverse path delay of the message. Because the reverse path is invisible to traceroute and can introduce latency very different to the forward path, it is necessary to examine the reverse path when in control of both source and destination hosts. Large additional latency jumps from one hop to the next might indicate a congested router and/or a suboptimal reverse path. A forward traceroute itself is insufficient when determining the location of networking issues. Asymmetric paths start often at network boundaries, e.g. on inter-AS links where administration policies change, *i.e.* "hot-potato" routing [7]. However, asymmetric path behavior can potentially occur at each and every hop along the path.

The measured RTT values include multiple latency sources. By definition, there are three main types of delay sources that can be observed in traceroute [6]: Serialization Delay, Queuing Delay, and Propagation Delay. *Serialization delay* is in the order of nano seconds and can be neglected. Queuing Delay depends on network utilization, thus, is a varying component of total delay. Propagation delay depends on the path chosen since the speed of light in fiber optical cables is constant. However, ICMP generation is not implemented in hardware and is considered to be one of the lowest priority functions within routers. Furthermore, it is typically rate-limited to 50-300 packets per second range [6].

III. Metric Design

This section introduces two metrics targetting different use cases, which are explained together with the metrics. The parameters used in the metrics can be derived from traceroute measurements and are summarized in Table I.

A. Metric 1: Path Stability

The first metric, termed *path stability* (PST), provides a measure of how stable a forward path is over time and, thus, indicating the quality of a path. The primary use case for this metric is based on applications that would rely on stable one-way communication, e.g. applications that use UDP to push latency and jitter sensitive information to other nodes, *e.g.,* (P2P) video streaming, selecting a Voice over IP (VoIP) relay. In such a context, the efficiency of the overlay network can be improved by selecting nodes or peers that are on stable paths. PST can be expressed by different variances in metrics such as route, number of hops, RTT, and packet loss. Considering the use case for the proposed PST metric the focus is laid on RTT and packet loss since these will have an effect on the quality of the service. Furthermore, if changes in route and hop count are relevant they have an effect on the RTT otherwise they are negligible.

The PST metric is defined in Equation 1 consists of two factors, $F_{packetloss}$ and $F_{congestion}$ presented in Equations 2 and 3 respectively. Table I lists the parameter definitions.

$$PST = F_{packetloss} \times F_{congestion} \tag{1}$$

$$F_{packetloss} = \frac{n_{resp}}{\sum_{i=1}^{n} \frac{probe_i}{total_i}} \tag{2}$$

$$F_{congestion} = \frac{\sum_{i=1}^{n} \frac{RTT_i}{n_{resp}}}{RTT_{last}} = \sum_{i=1}^{n} \frac{minRTT_i \times RTT_{last}}{n_{resp}} \tag{3}$$

$F_{packetloss}$ is based on the inverse value of the average number of all responsive probes that have been gathered on responsive hops. This results in a value that is an element of the half-open interval [1,∞). Therefore, it can be considered as a packetloss indicator where the value 1 represents no packetloss and values above indicate the amount of occurring packetloss.

The $F_{congestion}$ computes the average sum of the best RTT of all resolvable hops on the path in proportion to the best RTT on the target host. The results range within the interval (0,∞) and continuous, stable, and ascending forward paths are expected to range within the interval (0,1). Values pointing exactly to 1 are considered a special case and should only appear if the target host is one hop away. Values greater than 1 indicate a congestion along the path.

Multplying the two factors results in the PST metric. Optimal non-congested paths should lead to results in the range of (0,1). However, in reality the results are expected to lie within the interval (0,∞). For the purpose of identifying the best path, it is enough to only consider values within the range (0,1) since paths with a higher value are congested. The one hop distance special case is not considered to be relevant since this case would only appear within a local network and not typically in decentralized scenarios.

TABLE I
Summary parameters for all three metrics.

Parameter	Definition
n	Total hop count from source to destination host.
n_{resp}	Number of responding hops in path.
n'	Total hop count of the reverse path.
$probe_i$	Number of returned probes from hop i.
$total_i$	Total amount of probe sent to hop i.
RTT_i	Lowest RTT measured to hop i.
RTT_{last}	The lowest RTT measured to the last hop.
RTT'_{last}	Lowest RTT from destination to source host.

B. Metric 2: Path Symmetry

The second metric is called *path symmetry* (PSY) and provides a measure on the symmetry of the paths between two hosts. This metric targets use cases that benefit from symmetry of routing paths such as, distributed file systems, which rely on time sensitive file locking and versioning mechanisms. A key limitation of traceroute is that the reverse is invisible, but still part of the RTT measurement. In general, it is unlikely that the reverse path is the same as the forward path due to the behavior of Internet routing in general. There are two aspects that have an impact on the resulting asymmetry. First, a router determines the next hop for a packet based on its destination. Secondly, an AS can choose the next hop among its alternatives. Therefore, two adjacent ASes may use different peering points or apply intra-domain traffic engineering, which finally results in path asymmetry [3].

Considering the use cases CDN and P2P it can be assumed that source and destination hosts can actively participate in measurements. This allows gathering forward and reverse paths for two hosts, with a second traceroute from destination to source, rendering both paths visible. However, due to the limitations of traceroute it is not possible to simply compare the hops between forward and reverse path. Therefore, hop count and the RTT to the furthest hop of both paths are considered. Table I lists the element definitions that have been included to compute the metric. The metric is defined as follows:

$$PSY = \frac{n}{n'} \times \frac{RTT_{last}}{RTT'_{last}} \tag{4}$$

Based on the formula above, a perfectly symmetric path would result in a value of $PSY = 1$. For value ranges with $PSY > 1$, the reverse path is considered to be relatively inflated and vice versa $PSY > 1$ indicates relative inflation on the forward path.

The advantages of this metric are based on its simplicity and comparability in general. Values close to 1 imply that both paths have similar length in terms of hops and final

RTT, providing a good combination between hop count and end-to-end latency. In terms of path inflation, the metric might be biased by symmetric path inflation, but in use cases where the symmetry is of primary concern, this aspect is not of importance. Furthermore, a second trivial metric, such as hop count, can be used to identify inflated paths. On the other hand, the metric is intended to indicate relative path inflation that occurs on only one of the paths.

IV. Evaluation

This section shows how the metrics can be applied based on actual measurements. To execute the measurements a software was written and deployed in EmanicsLab [2] an international testbed for distributed systems. At the time of measurement execution 8 sites were available located in France, Germany Spain, Switzerland, the United Kingdom, and the Netherlands. The data collection was carried out in the last two weeks of July 2015. At the end, the evaluation data set was reduced to six days of measurements. Time frames longer than six days did not provide more measurement diversity.

The average path length for the evaluation data is 15.41 hops and has a median value of 14 hops. In relation, the statistics for each traceroute measurement indicate 2 hops for the third quartile. At first glance, this data indicates that the aspect of non-responding hops should not be that critical within the evaluation data set. Nevertheless, this depends on the amount of outliers on a path and their hop RTTs. Further conclusions in this area are not appropriate since the data set was collected for 8 nodes and therefore basically 56 paths. Nevertheless, the proposed path stability metric can always be interpreted as relative value over time. Routes that are considered stable will replicate similar results.

A. Metric 1: Path Stability (PST)

Section III-A proposed the PST metric that is computed by the product of a Packetloss Factor (PLF) and a Congestion Factor (CF) based on the data set, consisting of 161,272 entries. Those entries are based on the parsed traceroutes and therefore can include traceroute measurement anomalies from probing. Such anomalies can occur and have been considered as subject for pre-filtering input data [8]. Considering the packetloss factor, further outliers can be excluded. Assuming a packetloss of 33.33% the packetloss factor would be 1.5. Within the evaluation data set, there are ≈ 0.8% of measurements with a packetloss factor larger than 1.5. Indicating that packetloss was rarely observed.

The second component, congestion factor, can be biased much more. While some outliers can be explained by traceroute anomalies, others are influenced by the way the factor is calculated. The metric considers non-responding hops by discarding them and not by trying to estimate them. Nevertheless, in case of paths where groups and sequences of non-responding hops occur more frequently, an inaccuracy margin is implicitly introduced. For example, a path contains a high count of non-responding hops at the end. As a consequence, the congestion factor indicates lower congestion. The same characteristics apply for the reverse case. While this aspect might bias the distribution of the congestion factors, the primary source of outliers is based on intermediate hops reporting a RTT far higher than the RTT to the destination host. Considering that meaningful values of the congestion lie within the range (0,1), extreme outliers should be discarded. Excluding congestion factors greater than 1 results in the exclusion of ≈ 2.7% of the evaluation data set. The resulting statistics have been summarized in a box plot in Figure 1a based on the outlier corrected data set for all measured paths.

TABLE II
Path symmetry statistics for all paths.

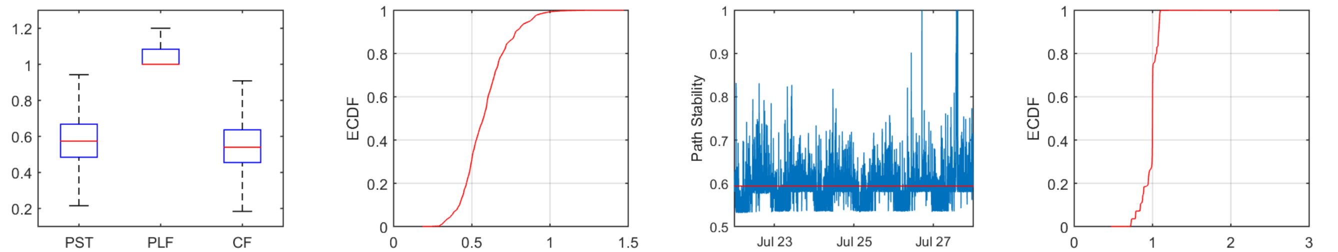
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.3824	0.9717	1.0000	1.0050	1.0290	2.6150

This leads to an adjusted evaluation data set that contains ≈ 96.5% of the original measurements. The resulting PST statistics for all measured paths can be found in Figure 1. Based on the adjusted data set, the Empirical Cumulative Distribution Function (ECDF) was plotted for the PST of all paths in Figure 1b. The ECDF depicts that approximately 75% of all PST computations lead to values below 0.6674.

The metric can be applied in two ways. First, the PST metric can be used to monitor a specific route over time. Figure 1c depicts the change of the PST during the evaluation period for the forward path between the sites Munich (DE) and Twente (NL). The figure shows that the PST metric follows diurnal patterns which show better results between midnight and ten o'clock in the morning, which reflect the traffic patterns observable on Internet exchanges. Second, absolute values of different paths in a network can be compared. This can be achieved when all hops on the paths are resolvable and responding, which was the case for 39% of all measurements. Otherwise, the routes need to be analyzed further to ensure that the metric performs as supposed.

B. Metric 2: Path Symmetry (PSY)

Section III-B presented the PSY metric, which was applied on the original data set. An inherent advantage of this metric is its simple design. While PST can be sensitive for RTT outliers on the intermediate path or non responding hops, this is not the case for PSY because it solely relies on the hop count and the final hop RTT for two corresponding hosts. Those are two measurement values that nearly always can be retrieved, but may require different probing protocols. The summary statistics of the PSY calculation is presented in Table II, which provides the minimum, maximum, 2nd and 3rd quartiles, mean,



(a) Path stability and its components.

(b) ECDF of PST.

(c) PST from UniBw to UT.

(d) ECDF plot of the PSY metric.

Fig. 1. PST calculation based on all measurements and PSY results. Outliers are removed in 1a and 1b but kept in 1c.

and the median of the PSY from all routes. To do so, the complete path symmetry data set has been analyzed and the reverse values have been computed. Table II lists the path symmetry statistics for all possible paths. Outliers are caused by inflated RTTs while the hop count has been consistent. The hop count for the paths on all routes has been very stable besides rare occurrences of alternate path lengths due to load balancing. Figure 1d depicts the ECDF for the path symmetry metric. It illustrates that all routes within the test data set can be considered highly symmetric.

Since PSY relies on initiating reverse traceroutes the metric provides a better approximation to the real path a packet could take compared to one-way metrics. This information is useful for applications that benefit from preferably symmetric path properties, such as VoIP. It is feasible to abstract the metric from hop to router level by using alias resolution techniques or even further by extracting the AS path.

V. Conclusions and Future Work

The novel metrics of path stability and path symmetry have been designed and applied to real world traces. These metrics assist in finding more efficient overlay network topologies while being simple to compute based on traceroute measurements acquired directly on nodes. These metrics have been evaluated for a time frame of six days, since longer time frames did not provide improved measurement diversity, indicating that weekly usage patterns have more influence than routing and latency fluctuations within the EMANICSLab research and experimentation network.

In conclusion, the path stability metric indicated that there is rarely any packetloss and congestion within the paths evaluated in EMANICSLab. This metric provides more profound indications on the forward path locality in comparison to path length and latency metrics. Therefore, this metric can be used in any case where traffic is flowing one way such as in P2P file sharing systems like BitTorrent or in CDN replication mechanisms where the optimal source to replicate from has to be selected. In a CDN context path stability is especially useful as nodes are very stable and paths can be observed over a long time.

The path symmetry metric has shown that even in the relatively stable EMANICSLab environment asymmetric paths exist. Since this metric indicates equal path qualities in both directions it is well suited for two way communication as it is required whenever P2P systems, *e.g.,* WebRTC, need to chose a relay for communication due to firewalls and Network Address Translation. Furthermore, both metrics can be used in conjunction to find the best path available.

Since both metrics have been evaluated within EMANICSLab, results may be biased due to the stable routing behavior within the research network. Therefore, further work includes the extension of the underlying measurements to include additional diverse networks, which include end-user service providers and oversea paths.

Acknowledgments

This work has been performed partially in the framework of the EU ICT NoE FLAMINGO (No. 318488) supported by the European Commission under its Seventh Framework Programme.

References

[1] B. Donnet and T. Friedman, “Internet Topology Discovery: a Survey,” *Commun. Surveys Tuts.*, vol. 9, no. 4, pp. 56–69, Oct. 2007.

[2] EMANICSLab Website, “EMANICSLab,” available online at http://www.emanicslab.org//. Last visited: June 8th, 2015.

[3] E. Katz-Bassett, H. V. Madhyastha, V. K. Adhikari, C. Scott, J. Sherry, P. Van Wesep, T. Anderson, and A. Krishnamurthy, “Reverse Traceroute,” 7th USENIX Conference on Networked Systems Design and Implementation 2010 (NSDI’10), San Jose, CA, USA, April 2010, pp. 15–15.

[4] K. Keys, “Internet-scale IP Alias Resolution Techniques,” *SIGCOMM Comput. Commun. Rev.*, vol. 40, no. 1, pp. 50–55, Jan. 2010.

[5] Lawrence Berkeley National Laboratory Website, “Lbnlś network research group,” available online at http://ee.lbl.gov/. Last visited: July 21st, 2015.

[6] R. A. Steenbergen, “Traceroute,” 2011, available online at http://cluepon.net/ras/traceroute.pdf. Last visited: July 28th, 2015.

[7] R. Teixeira, A. Shaikh, T. Griffin, and J. Rexford, “Dynamics of Hot-potato Routing in IP Networks,” ACM SIGMETRICS 2004, New York, NY, USA, June 2004, pp. 307–319.

[8] U. Weinsberg, Y. Shavitt, and Y. Schwartz, “On the Diversity, Stability and Symmetry of End-to-End Internet Routes,” IEEE INFOCOM 2010, San Diego, CA, USA, March 2010, pp. 1–6.