©2011 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The definitive version of this paper has been published in Proceedings of the 7th Conference on Next Generation Internet Networks (NGI), 2011.

Performance of Concurrent Multipath Transmissions - Measurements and Model Validation

Thomas Zinner*, Dominik Klein*, Kurt Tutschku†, Tanja Zseby‡, Phuoc Tran-Gia*, Yuval Shavitt§
*University of Wuerzburg, Institute of Computer Science,
Wuerzburg, Germany
†University of Vienna, Professur "Future Communication",
Universitätsstrasse 10/T11, 1090 Wien, Austria
‡ FOKUS - Fraunhofer Institute for Open Communication Systems,
Kaiserin-Augusta-Allee 31, 10589 Berlin, Germany
§ Tel Aviv University, School of Electrical Engineering,
Tel Aviv, Israel

***Abstract*—Concurrent multipath transport layer mechanisms have gained recently increasing interest in research and standardization because of the potential for bandwidth aggregation, load balancing and increased reliability. Multihomed end devices may benefit from IP-based multipath protocols like Multipath TCP or Concurrent Multipath Transmission via SCTP. In the future, concurrent multipath transport might be transparent to network and transport layer protocols as proposed by the concept of Transport Virtualization. This mechanism enables the pooling of heterogeneous transmission technologies or physical paths. However, the selection and application of multiple paths and its impact on the transmission is non-intuitive. Our prior work on transport virtualization discussed the impact of delay diversity of pooled paths on concurrent data transmissions. A mathematical model was introduced enabling the analysis of packet re-ordering that occurs due to different path delays. In this paper we extend our investigations by validating the previously presented analytical and simulative models with measurements performed in Planetlab Europe and the Etomic testbed.**

I. Introduction

Multipath transmission techniques and protocols on transport layer have gained increasing interest in standardization and research. This trend is mostly driven by multihomed end user devices like smartphones, tablets or netbooks. The most popular protocols currently under development are Multipath-TCP (MPTCP) [1] and Concurrent Multipath Transmission-SCTP (CMT-SCTP) [2]. The main benefits of these state-of-the-art transport protocols are their potential for bandwidth aggregation, increased reliability, load balancing, and joint congestion control [3].

All these multipath transport concepts have one thing in common. They logically build a new, pooled transport resource out of several transport resources, typically with different resource characteristics like capacity or transmission delay. In Future Networks, such concepts might be transparent to network or transport layer protocols as proposed by the concept of Transport Virtualization (TV) [4] or possible extensions of MPLS [5].

However, the impact of the dissimilarity of the different used resources on the pooled transport resource is a fundamental problem and motivates the design of appropriate functions for selecting the best set of available resources. Furthermore, different scheduling mechanisms on the utilized paths might influence the system performance.

These problems are typically addressed with analytical or simulative models. For the case of Transport Virtualization, such models were introduced enabling the analysis of packet re-ordering that occurs due to different path delays. Although first results on the impact of different varying path delays on the re-ordering probability have been attained, these theoretical models still lack validation and their limitations are still unknown.

The validation of theoretical models in local laboratories is possible but those laboratories suffer from scalability limits since physical distances and the number of resources are limited. Also the acquisition of specific measurement equipment for performing highly accurate measurements is often difficult due to the high costs of such hardware. A possible solution is the utilization of experimental facilities as provided by the European FIRE program. It provides the required scalability features and allows the use of special equipment like the ETOMIC nodes [6].

The aim of this paper is the validation of the theoretical models of Transport Virtualization with measurements conducted in experimental facilities. These models allow a computation of the re-sequencing buffer and the end-to-end delay out of the path capacity and the path delay distribution. For that the TV mechanisms were implemented and an adequate amount of measurements were conducted.

The reminder of this paper is structured as follows. Section II discusses different multipath transport mechanisms, summarizes common problems and introduces experimental facilities used in the following. In Section III, we detail the investigated system and describe the performance metrics. After that we focus on the validation methodology in Section IV. The results are presented in Section V and the paper is concluded in Section VI.

II. Background and Related Work

In this section we shortly summarize different multipath transport mechanisms and introduce experimental platforms.

A. Multipath Transport Mechanisms

This section presents a short overview of different multipath transmission techniques and highlights similarities between the different approaches.

1) Multipath TCP: Multipath TCP is an extension of TCP for the usage over multiple available paths and is currently discussed within the IETF working group MPTCP [1]. It enables the usage of multiple paths and ensures fairness in case of shared bottlenecks. This is achieved by a coupled congestion control of the single TCP connections over each path, cf. [7]. Currently different implementations of Multipath TCP are discussed within the IETF, for instance Multipath TCP [8] or Multi-Connection TCP (MCTCP) [9]. However, they mainly discuss implementation details on kernel level and do not change the coupled congestion control schema as presented in [10]. Performance questions how different path characteristics like bandwidth and packet end-to-end delay influence the overall performance are not discussed in detail yet.

2) SCTP-CMT: The idea to use SCTP-multihoming for transmitting data concurrently via different paths was published in 2004 by Iyengar et al. [11]. An evaluation of the presented SCTP extension in [12] revealed problems occurring in case of multipath transmissions and introduced additional mechanisms to solve these problems. However, for the case of dissimilar path characteristics like latencies and available bandwidth, buffer blocking at the sending and receiving buffer occurs and has to be solved, as discussed in [13], [14]. Currently, the standardization of SCTP-CMT is discussed within the IETF Transport Area WG, cf. [2].

3) Transport Virtualization: The concept of Transport Virtualization (TV) enhances the capabilities of future networks. TV can be considered as an alternative mode of Network Virtualization (NV) [15]. While NV typically facilitates the sharing of resources, TV creates virtual resources (e.g. virtual links) based on the aggregation of resources [15]. To cope with out-of-order arrivals which occur inevitably in case of concurrent multipath transmissions, a re-sequencing buffer is introduced. In [4], a theoretical model is introduced which allows the computation of the buffer occupancy with regard to the path delay distributions of the involved paths.

4) Similarities: The mentioned multipath approaches share the fact that different path capacities or latencies may lead to performance degradations due to buffer blocking or additional delays due to re-sequencing mechanisms. The impact of this performance degradation might be reduced by appropriate selection strategies in case of different path characteristics or appropriate scheduling mechanisms on the different paths. However, this issues are not yet addressed as can be seen in [2], where the section about scheduling is missing. An initial discussion in this direction and a generic mathematical modeling can be seen in a previous work [4] where appropriate path selection mechanisms were discussed.

B. Experimental Platforms

Large research programs in the US, Europe and Asia work on the establishment of global experimental facilities for future Internet research. Examples are the GENI program in the US and the FIRE program in Europe. Many nations also have own national programs to support experimental research (G-Lab in Germany, F-Lab in France, etc.). PlanetLab Europe is the European part of the worldwide experimental platform PlanetLab. PlanetLab forms an overlay of machines which are spread around the world and can be reserved by PlanetLab users to perform global-scale experiments. PlanetLab Europe is administered by the PlanetLab Europe Office in Paris and is supported by the European Project OneLab2. PlanetLab Europe integrates further technologies like wireless testbeds and platforms for disruptive Future Internet research like autonomic communication, delay tolerant networks and packet switched networks in the platform to make such technologies available for scientists. The PlanetLab Europe testbed is federated with PlanetLab Central, PlanetLab Japan, and other experimental facilities (e.g. wireless testbed NITOS, measurement infrastructure ETOMIC [6], and others). PlanetLab Europe provides highly sophisticated control and monitoring functions to support the needs of experimental-driven research. The OneLab2 project has developed a measurement solution that integrates passive and active measurements, the Advanced Network Monitoring Equipment (ANME), which is specified in [16]. Sophisticated active measurements are supported by the federation of PlanetLab with the world-wide measurement infrastructure ETOMIC. PlanetLab Europe also supports passive multipoint measurements to follow the path that packets take in the network (multi-hop packet tracking) and allows resource controlled measurements enabled by sampling technologies.

III. System Description and Performance Metrics

In this section, we briefly present the system setup for the multipath experiments and detail the performance metrics.The investigated multipath system is based on Transport Virtualization as described in [17]. For the validation of the models we decided to setup a concurrent multipath transmission via two paths as shown in Figure 1. Due to different path characteristics, the packets experience a different delay. Furthermore, this delay may vary due to volatile network congestion. Thus, the packets arrive unordered at the destination. In order to resolve this issue, a buffer is introduced enabling packet re-sequencing. This however leads to an additional waiting time which increases the experienced end-to-end delay of the packets.

With respect to this system, we have the following input parameters:

- **capacity** describes how many packets are transmitted per time unit over a path. In our experiment one packet is transmitted every 10 milliseconds.

- **scheduling** describes how the packets are transmitted over the different paths. In this paper, we investigate a scheduling which transmits a packet per time unit over both paths. At time i packet $2i+1$ is transmitted via the first path and packet $2i+2$ via the second path.
- **path selection** depicts the influence of different path selection strategies on the system. To design appropriate selection mechanisms, a deep analysis of the system behavior has to be performed. Since the purpose of this paper is to validate the theoretical models and investigate the end-to-end delay this input parameter is omitted.
- **path delay distribution** depicts the distribution of the delays each packet experiences. The delays may vary and can be combined to a path delay distribution / histogram. In this paper, we investigate different transmission paths from source to destination leading to different delays.
- **buffer size** illustrates the size of the re-sequencing buffer. For the course of this work we assume this buffer to be infinite, i.e., packets are not lost.

The above mentioned input parameters have an impact on the following performance metrics:

- **buffer occupancy** describes the number of packets which are stored in the buffer. The buffer occupancy is investigated every time unit.
- **waiting time** describes the additional waiting time introduced by the buffering. Due to the high correlation of this metric with the buffer occupancy, we omit this metric during the course of this work.
- **end-to-end delay** depicts the perceived delay for each packet combing its path delay and its waiting delay.

In the following, we present an evaluation of this system by means of measurements, simulation [18], and analysis [4] in order to validate our prior theoretical investigations. In the next sections we outline the different evaluation methods.

IV. Methodology

The methodology to cross validate our simulator and the analytical model relies on measurements conducted in the

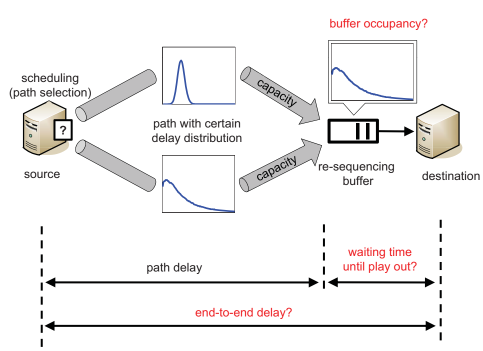


Fig. 1. System description

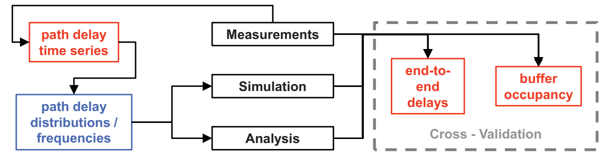


Fig. 2. Validation methodology

PlanetLab Europe testbed. It is illustrated in Figure 2 and described in detail in the following.

A. Analysis

Generic relationships between input and output parameters of a system are usually investigated with analytical methods since they enable a broader investigation than simulations on an abstract level. For the investigated multipath system we adapted the analytical model of Nebat and Sidi [19] and performed an evaluation on the influence of different path characteristics on the buffer occupancy, [4].

The model assumes a continuous data stream for the multipath transmissions over m concurrent paths. The delays on the paths are independent and described by discrete delay distributions with a resolution of one time unit. We further consider paths with equal capacity and that the transmission rate on each path is equal to one packet per time unit. A detailed explanation of the mathematical model can be found in [19], [20]. The used model ensures that no packet reordering on a single path can occur. Packets send over one path can not overtake each other. To facilitate the explanation, the following notations are used:

The packets transmitted at time 0 over path $1, 2, ..., m$ are packets $1, 2, ..., m$ respectively. After transmitting the first m packets, the packets are appointed to the sources in a round-robin manner. Thus at time t, packets $1+mt, 2+mt, ..., m+mt$ are transmitted. We further use the term *minimum valued packet (mvp)*, as introduced in [19], denoting the lowest indexed packet at time t that has not arrived at the destination by time t. For instance, if packets 1 through 5 arrived, but packet 6 did not, packet 6 is the mvp. Thus, the re-sequencing buffer occupancy at time t is equal to the number of packets with higher index than the mvp that have arrived by time t. Since packets transmitted on every path arrive in transmission order, no packet in the re-sequencing buffer was transmitted via the path of the mvp.

We denote the index of the path of the mvp by s_n and $\delta_{X,t}$ as the time passed since the last packet, received via path X, was transmitted at time t. For brevity we refer to $\delta_{X,t}$ as δ_X in the following. With this notation the re-sequencing buffer occupancy can be computed as:

$$P(B=k) \quad = \quad \sum_{i=1}^{m} \sum_{x=0}^{\infty} P(B=k, s_n=i, \delta_i=x). \quad (1)$$

The right hand side of the equation denotes the buffer occupancy probability for each path transmitting the mvp and each possible value for the time passed since the last packet transmitted over this path was received. As discussed in [20], this yields to:

$$
\begin{aligned}
&P(B=k, s_n=i, \delta_i=x) \\
&= P(\sum_{j=1, j\neq i}^{m} \delta_j = (m-1)x + 1 - i - k, \\
&\qquad\qquad \delta_j < x \forall j < i, \delta_j \leq x \forall j > i) \\
&= \sum_{S_{i,x,k}} P_i(x) \prod_{j=1}^{i-1} P_j(l_j) \prod_{j=i+1}^{m} P_j(l_j)
\end{aligned}
\tag{2}
$$

where $S_{i,x,k}$ defines the delay configuration on path before the arrival of the mvp:

$$
\begin{aligned}
&S_{i,x,k} = \{l_1, ..., l_{i-1}, l_{i+1}, ..., l_m : l_1 < x, ..., l_{i-1} < x, \\
&l_{i+1} \leq x, ..., l_m \leq x, \sum_{j=1, j\neq i}^{m} l_j = (m-1)x + 1 - i - k\}.
\end{aligned}
$$

With this formula, we can compute the re-sequencing buffer occupancy in case of transmission over m paths with equal transmission rate. However, the analytical approach lacks the possibility to investigate the perceived end-to-end delay and, as we will see later, correlated path delays.

B. Simulation

In order enable a broader investigation of different input parameters on the system behavior, a prior work investigates a simulative approach, cf. [18]. This simulation framework was ported to the OMNeT++ simulator and provides the possibility to be used with arbitrary path delay distributions. Further, we enhanced it to be used with delay time series as measured within the ETOMIC experiments and to investigate the end-to-end delay on packet level. With these enhancements, we can easily validate our simulation with the performed measurements in the PLE/ETOMIC testbed and investigate the influence of predefined scenarios on the buffer occupancy and the end-to-end delay.

C. Measurements

Our measurement setup consists of a source, a destination, and different paths between source and destination. The packet forwarding was realized by application layer packet forwarding. For measuring the one way delays between source and destination we relied on ETOMIC nodes with DAG cards. These nodes provide high precision GPS synchronized timestamps with an accuracy of nanoseconds, cf. [6]. As source and destination node we used the ETOMIC nodes located in Pamplona and Budapest. The application layer packet forwarding was realized on PlanetLab Europe hosts located in Ireland, China, Canada and Brazil. We investigate three different multipath transmission scenarios each with two paths:

- **Brazil-Ireland** respectively via nodes located at RNP - Para and Waterford Institute of Technology. The purpose of this experiment was to investigate the system behavior for a setup with a low and a high delay path, i.e. the path characteristics are very different.
- **China-Canada** respectively via nodes located at PLA University of Science and Technology and University of Waterloo. We choose this setup for evaluating the system with two paths with similar delays, i.e. a very homogeneous scenario.
- **Brazil-Brazil** respectively via nodes located at RNP - Para and RNP - Rio de Janeiro. In this scenario we also expect very similar delays for both paths. However, since we sent packets from Europe via two nodes in Brazil the packets are sent partly via the same link. If this is a bottleneck link we can observe the influence of a shared bottleneck on the multi path transmission mechanism.

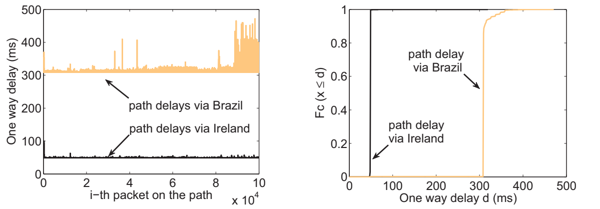
For each scenario, we conducted 14 experiments and transmitted 100.000 packets per path, that means 200.000 packets in total, with constant inter departure times of 10 ms, as described in Section III. Each packet contains a unique id. Thus, packets with odd ids were sent via the first path and packets with even ids via the second path. In case of lost packets on a path, e.g. packet 7 arrives after packet 3 on the first path and hence packet 5 is missing, we play out all packets in the buffer including the last successfully arrived packet, e.g. packet 7. This is a valid approach since we can confirm with our measurements that no packet reordering on a single path occurred. It should be noted that the experimental setup was very complicated since the different paths had to be set up manually.

With measurements, the true system behavior can be investigated which implies a high reliability of the gained results. However, the number of possible scenarios are limited. For our experiment, we cannot affect the network conditions on the different paths. We rather have to check that different configurations involve different path delays and thus a different system behavior.

V. Results

A. Investigated One Way Path Delays

We start with illustrating typical time series and typical path delay frequencies of the first scenario, a multipath transmission via Brazil and Ireland. After that we discuss the path delay distributions for the other scenarios. Results for the first scenario, Brazil-Ireland, are depicted in Figure 3. The experienced one way delay for each packet, i.e. the delay time series, is depicted in Figure 3(a). It can be seen that the one way delay of packets routed via Ireland is significantly lower than the one way delay of packets routed via Brazil. Further, the delay variation on the path via Brazil is higher since packets experience delays between approximately 310 ms and 420 ms. The results are also depicted as cumulative frequencies in Figure 4(b). We see that about 90% of the packets on the path via Brazil experience a similar one way delay of around



(a) One way delay time series for each packet (b) One way delay cumulative frequencies

Fig. 3. One way delay measurements for Brazil - Ireland

310 ms and that the rest of the packets experience higher one way delays. In contrast, the path via Ireland is rather constant and the delay gap between both paths is around 260 ms. Due to the constant inter departure time of 10 ms of packets on each path, we expect that the buffer is always filled with more than 20 packets for this scenario.

The cumulative one way delay frequencies for the other scenarios are depicted in Figure 4. As explained in Section IV, we chose the multipath scenario via China - Canada because of the expected similar one way delays. The measurements are illustrated in Figure 4(a). It can be seen that the gap between the path delay frequencies is around 100 ms, smaller as in the Brazil - Ireland scenario. Thus, we expect a smaller minimum number of packets which are always stored in the re-sequencing buffer. Further, it can be seen that the delay variation is small for both paths. The one way delay frequencies for the third scenario with both paths via Brazil are depicted in Figure 4(b). It can be seen that the one way delays for the path via Rio are shorter than for the path via Para. Further, the delay variation on each of the utilized paths is higher than in the other scenarios.

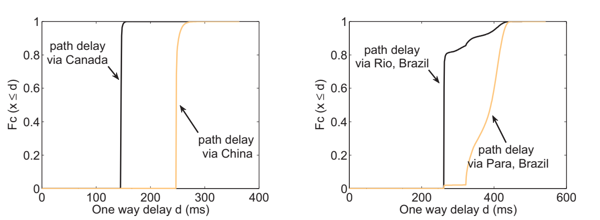
B. Validation of the Buffer Occupancy

In this subsection, we investigate the re-sequencing buffer occupancy for the different scenarios. Due to the number of measurements we show for each scenario an experiment with a close match of measurements, simulations and analysis and the scenario with the worst match. For describing the goodness of fit between measurement / analysis and measurement / simulation we computed the mean squared errors (MSE) between the curves. The re-sequencing buffer occupancies, depicted on the x-axis, are illustrated as cumulative relative frequencies Fc, depicted on the y-axis. The MSE's for all experiments will be investigated in V-B4

1) Brazil - Ireland: The first scenario to discuss is for a path via Brazil and one path via Ireland. The results for two experiments, one with a good match of the models and one with the worst match, are depicted in Figure 5. It can be seen that due to the high delay difference of the used paths the re-sequencing buffer is always filled with at least 20 packets. The closest match between models and measurements according to MSE is depicted in Figure 5(a). The computed MSE between measurements and simulation $mse_{sim} \approx 0.0004$ and between measurements and analysis $mse_{ana} \approx 0.0007$. The differences between measurements and simulations can be explained by the fact that for the measurements the inter departure times for the packets on a specific path are not constant but varying slightly. This is not reflected by the simulation. Another effect only happening in the measurements is that a transmission time is skipped and at the next transmission time two packets are sent. The analysis uses the histogram as input parameter and not the time series which constitutes a strong abstraction and thus explains the gap between measurements and analysis.

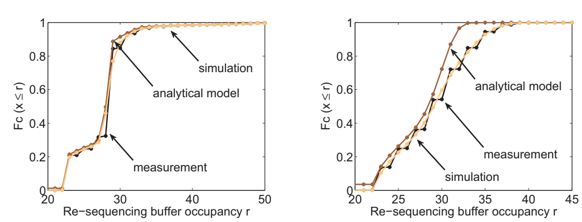
The scenario yielding to the highest MSE and thus to the highest difference between measurements and the models is depicted in Figure 5(b). However, it can be seen that the simulation still matches the measurements pretty good, and that the analysis is also a good match for the measurements. The computed MSE between measurements and simulation $mse_{sim} \approx 0.003$ and between measurements and analysis $mse_{ana} \approx 0.004$.

2) China - Canada: In the second scenario, we investigate a path via China and a path via Canada. Again, we show the results for a scenario with a small MSE, i.e., a good match between measurements, simulation and analysis, and for a scenario with a high MSE, i.e., a worse match. The results are depicted in Figure 6. For this scenario, the buffer occupancy is lower as compared to the previous scenario. That is due to the lower difference in path delays between both paths. The computed MSE between measurements and simulation $mse_{sim} \approx 0.0002$ and between measurements and analysis $mse_{ana} \approx 0.0002$ are illustrated in Figure 6(a). In contrast, the computed MSE between measurements and analysis or simulations is higher for the experiment depicted in Figure



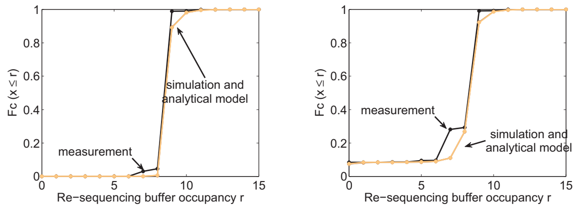
(a) China - Canada (b) Brazil - Brazil

Fig. 4. Measured cumulative one way delay frequencies for the other scenarios



(a) good match between measurements and models (b) worst match between measurements and models

Fig. 5. Re-sequencing buffer occupancy for measurements, simulation and analysis

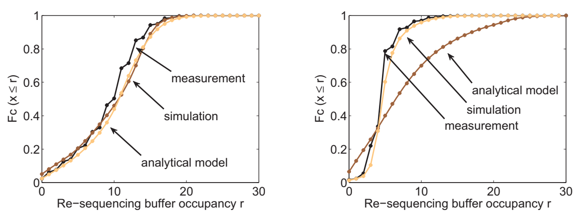


(a) good match between measurements and models (b) worst match between measurements and models

Fig. 6. Re-sequencing buffer occupancy for measurements, simulation and analysis

6(b), respectively $mse_{sim} \approx 0.001$ and $mse_{ana} \approx 0.001$. This can again be explained by the effects occurring in measurements which are not reflected by the simulation and the analysis. However, it can be seen that the theoretical models are still very close to the measurements.

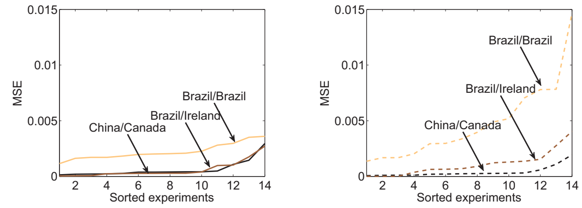
3) Brazil - Brazil: In the third scenario, we examined two paths in Brazil: one via Rio, the other via Para. The results we present in detail are depicted in Figure 7. As can be seen



(a) good match between measurements and models (b) worst match between measurements and models

Fig. 7. Re-sequencing buffer occupancy for measurements, simulation and analysis

in both subfigures, the buffer occupancy is mostly very small which is due to the similar delay magnitudes on the paths. The experiment with the lowest MSE is depicted in Figure 7(a), with $mse_{sim} \approx 0.002$ and $mse_{ana} \approx 0.001$. The MSE is higher as in the previous scenarios, but the models still fit the measurements very well. The experiment with the worst match for this scenario is depicted in Figure 7(b). Although the simulation is very close to the measurements, as also indicated by the $mse_{sim} \approx 0.002$, the difference between measurements and analysis is very high. This is also reflected by the MSE which is much higher, $mse_{ana} \approx 0.015$. Since this difference cannot be explained by minor modeling inaccuracies, we investigated the results in more detail. It turned out that for this experiment the delay correlation between both paths. In case of a high delay value on path one it is very likely that the delay value on path two is also very high and vice versa. In case of uncorrelated path delays, a high delay value on path one does not allow any prediction about the delay value on the second path. The analytical model represents the uncorrelated case which becomes very imprecise in case of path delay correlations. Due to the high delay correlation, packets on different paths experience similar delays. This means that



(a) MSE between measurements and simulations (b) MSE between measurements and analysis

Fig. 8. Mean squared error (MSE) between measurements and analysis or simulation for the different scenarios

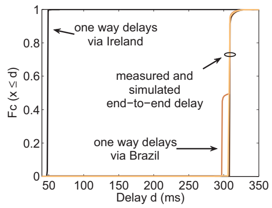
it is very likely that missing packets arrive soon after their predecessor on the other path. Thus, the buffer occupancy keeps much lower.

The experiment shows the limitations of the analytical model; path delay correlations are not considered and thus the model becomes imprecise for high correlations. On the other side, the values computed with the simulation are still very close to the measurements. This scenario shows the effect of a multipath transmission via a shared link on the re-sequencing buffer occupancy and the limitations of the analytical model.

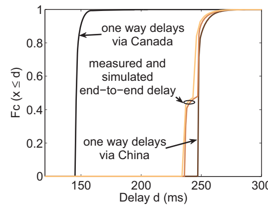
4) Mean Squared Errors: In the last part of this subsection, we want to discuss the MSE values for all experiments and show that analysis and simulation match the measurements mostly. The results for both models are depicted in Figure 8. Figure 8(a) details the results for the comparison of the measurements with the simulations. As can be seen, the MSE is always very low which means that the simulation is always very close to the measurement results. The experiments for the scenario with both paths via Brazil are worst in the comparison between the values. However, as can be seen in Figure 7, the simulation model is still very accurate. The comparison between analytical model and measurements are depicted in Figure 8(b). Concerning the scenarios via China/Canada and Brazil/Ireland, the results of the model are always very close to the measurements. The scenario with a concurrent multipath transmission via two nodes in Brazil, yields to five experiments with a MSE higher than 0.005, which is worse than the MSE of all other experiments. As discussed before this is mainly due to the path delay correlation which is $\sigma_{path1,path2} > 0.1$ for these experiments. Thus, we can conclude that we can not use the analytical model to predict the behavior of the re-sequencing buffer in case of a significant delay correlation between the involved paths.

C. Impact on the End-to-End Delay

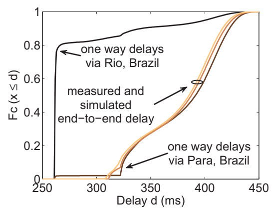
In this subsection, we investigate the impact of different path delays on the perceived end-to-end delay which comprises the means of the path delay of a packet and the corresponding waiting time in the re-sequencing buffer. For that we check whether the simulated results match the measurements, and we investigate the influence of the path delays on the end-to-end delay. The results are depicted again as cumulative frequencies in Figure 9. A typical result for the first scenario,



(a) Delays for the scenario Brazil - Ireland



(b) Delays for the scenario Canada - China



(c) Delays for the scenario Brazil-Brazil

Fig. 9. Measured one way delays and end-to-end delays and simulated end-to-end delays for the investigated scenarios

a multipath transmission via Brazil and Ireland, is depicted in Figure 9(a). As discussed previously, the one way delay via Ireland is much smaller than via Brazil leading to a buffer occupation of 20 packets and more, mostly packets sent via Ireland. These packets have to be stored in the re-sequencing buffer and wait for missing packets sent via Brazil. Thus, the packets sent via Brazil are mostly in sequence and thus can be played out immediately. For these packets the end-to-end delay is mostly the experienced one way delay. Accordingly, a packet sent via Ireland can be played out after the previous packet sent via Brazil arrived. Hence, the packet experiences a similar end-to-end delay which is dominated by the higher one way delay path. However, the end-to-end delay is compound of the shorter one way delay and the waiting delay. Thus, the overall experienced end-to-end delay is dominated by the path via Brazil. As a result, the cumulative end-to-end delay frequencies are very similar to the cumulative delay frequencies of the path via Brazil. Further, it can be seen that simulated end-to-end delay frequencies closely match the measured end-to-end delay. For the multipath transmission via Canada-China, depicted in Figure 9(b), and via Brazil-Brazil, illustrated in Figure 9(c) both investigations hold also. Thus, we can conclude that the simulation model predicts the perceived end-to-end delay very accurate and that the perceived end-to-end delay of each packet can be approximated fairly with the higher one way delay frequency.

VI. Conclusion

In this paper we validated the theoretical models designed for Transport Virtualization by measurements in experimental facilities. For that we implemented the multipath mechanism for the usage in the Etomic / Planetlab Europe testbed and performed measurements for investigating the influence of dissimilar paths, i.e. paths with different one way delays, on the multipath transmission mechanisms. For each of the three setups we performed 14 experiments where we measured the packet delays, the re-sequencing buffer occupancy and the packet end-to-end delay. With the measured packet delays we also computed the buffer occupancy and the end-to-end delay with theoretical models.

It turned out that the theoretical models provide a good approximation of the re-sequencing buffer occupancy as long as the delay correlation between the utilized paths is lower than $\sigma_{path1,path2} < 0.1$. This holds for the presented simulative and analytical models. Additionally, it turned out that the analytical model becomes inaccurate for higher positive delay correlations. However, it can still be used as upper bound approximation since it overestimates the buffer occupancy significantly. In case of a proper dimensioning of the system it is necessary to integrate the path delay correlation in the theoretical models in order to avoid over provisioning and thus a waste of resources.

Further, we investigated the impact of different path delays on the perceived end-to-end delay with measurements and simulations. It turned out, that the simulated end-to-end delays match the measured end-to-end delays very well. Further we have seen, that for the investigated dissimilar path combinations, the path with the higher delay dominates the end-to-end delay, i.e. the delay frequencies of this path are nearly similar to the end-to-end delay. Thus, the end-to-end delay frequencies can be approximated by the delay frequencies of the path with the higher delay.

We have seen that the previously presented theoretical models can be used to compute the performance metrics buffer occupancy and end-to-end delay with measured one way delay. This increases the confidence of the theoretical models and enables the design of an appropriate resource selection for a pooled multipath transmission. Further, the impact of the dissimilar paths on the transport can be estimated before its conducted.

Network Virtualization and Network Federation enable the safe sharing and the temporal leasing of variable resources. The transmission concepts presented in this paper enable the combination of these resources dynamically. The performance models, presented in this contribution, provide performance envelops for these mechanisms. Thus, they provide foundations for the expected performance capabilities of possible Future Internet concepts.

Future work will have to deal with a deep simulative and analytical investigation of the system. At first, the influence of positive and negative delay correlations between the utilized paths has to be analyzed in more detail. Then we have to investigate the trade-off between the impact on the end-to-end delay and the re-sequencing buffer occupancy in detail. What happens, for instance, if the utilized paths have similar one way delays. Does this have a significant impact on the end-to-end delay? Or can we still approximate the end-to-end delay with one of the delay frequencies? Last but not least it needs to be investigated if the system can be optimized by the usage of appropriate scheduling mechanisms. These investigations have to be performed with the help of the validated theoretical models, an investigation with measurements would be too demanding.

ACKNOWLEDGEMENTS

The authors would like to thank Tobias Hoßfeld for the fruitful discussions and his support during the course of this work. A special thanks goes to Sven Scheuring for implementing the software and performing the measurements within the testbeds. The sponsorship of this research by the European FP7 Network of Excellence "Euro-NF" through the Specific Joint Research Project "Multi-Next" is thankfully acknowledged.

REFERENCES

[1] Yoshifumi Nishida and Philip Eardley, "Charter of the Multipath TCP (MPTCP) working group," .

[2] Martin Becke, Thorsten Dreibholz, Janardhan Iyengar, Preethi Natarajan, and Michael Tuexen, "Load Sharing for the Stream Control Transmission Protocol (SCTP) draft-tuexen-tsvwg-sctp-multipath-01.txt," Jan 2011, Information available at http://tools.ietf.org/html/draft-tuexen-tsvwg-sctp-multipath-01.

[3] Damon Wischik, Mark Handley, and Marcelo Bagnulo Braun, "The resource pooling principle," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 5, pp. 47–52, 2008.

[4] Thomas Zinner, Kurt Tutschku, Akihiro Nakao, and Phuoc Tran-Gia, "Using Concurrent Multipath Transmission for Transport Virtualization: Analyzing Path Selection," in *Proceedings of the 22nd International Teletraffic Congress (ITC)*, Amsterdam, Netherlands, Sept. 2010.

[5] Xavier Hesselbach, Ramon Fabregat, and Christos Kolias, "The impact over the packets sequence at the output interface in load balancing strategies," in *Transparent Optical Networks, 2006 International Conference on*. IEEE, 2006, vol. 3, pp. 263–266.

[6] D. Morato, E. Magana, M. Izal, J. Aracil, F. Naranjo, F. Astiz, U. Alonso, I. Csabai, P. Haga, G. Simon, J. Steger, and G. Vattay, "The european traffic observatory measurement infraestructure (etomic): A testbed for universal active and passive measurements," in *Proceedings of the First International Conference on Testbeds and Research Infrastructures for the DEvelopment of NeTworks and COMmunities*, Washington, DC, USA, 2005, pp. 283–289, IEEE Computer Society.

[7] Damon Wischik, Mark Handley, and Costin Raiciu, "Control of Multipath TCP and Optimization of Multipath Routing in the Internet," in *Proceedings of the 3rd Euro-NF Conference on Network Control and Optimization*, Berlin, Heidelberg, 2009, NET-COOP '09, pp. 204–218, Springer-Verlag.

[8] Alan Ford, Costin Raiciu, and Mark Handley, "TCP Extensions for Multipath Operation with Multiple Addresses draft-ietf-mptcp-multiaddressed-02," Jan 2011, Information available at http://tools.ietf.org/html/draft-ietf-mptcp-multiaddressed-02.

[9] Michael Scharf, "Multi-Connection TCP (MCTCP) Transport, draft-scharf-mptcp-mctcp-01," Jan 2011, Information available at http://tools.ietf.org/html/draft-scharf-mptcp-mctcp-01.

[10] Costin Raiciu, Mark Handley, and Damon Wischik, "Coupled Congestion Control for Multipath Transport Protocolsdraft-ietf-mptcp-congestion-01," Jan 2011, Information available at http://tools.ietf.org/html/draft-ietf-mptcp-congestion-01.

[11] Janardhan Iyengar, Paul Amer, and Randall Stewart, "Concurrent Multipath Transfer Using SCTP Multihoming," in *SPECTS'04*, 2004.

[12] Janardhan Iyengar, Paul Amer, and Randall Stewart, "Concurrent Multipath Transfer using SCTP Multihoming over Independent End-to-End Paths," *IEEE/ACM Trans. Netw.*, vol. 14, no. 5, pp. 951–964, 2006.

[13] Hakim Adhari, Thomas Dreibholz, Martin Becke, and Erwin Rathgeb, "Evaluation of concurrent multipath transfer over dissimilar paths," in *1st International Workshop on Protocols and Applications with Multi-Homing Support (PAMS)*, 2011.

[14] T. Dreibholz, M. Becke, E. P. Rathgeb, and M. Tüxen, "On the Use of Concurrent Multipath Transfer over Asymmetric Paths," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, Miami, Florida/U.S.A., Dec. 2010, ISBN 978-1-4244-5637-6.

[15] Kurt Tutschku, Thomas Zinner, Akihiro Nakao, and Phuoc Tran-Gia, "Re-sequencing Buffer Occupancy of a Concurrent Multipath Transmission Mechanism for Transport System Virtualization," in *Proc. of the 16. KiVS 2009*, Kassel, Germany, Mar. 2009.

[16] Gábor Vattay, Attila Fekete, Péter Hága, József Stéger, Javier Aracil, and Giuseppe Iannaccone, "Specifications for advanced monitoring boxes, onelab deliverable d4.2," 2008, available at http://www.onelab.eu/index.php/results/deliverables.html.

[17] Kurt Tutschku, Thomas Zinner, Akihiro Nakao, and Phuoc Tran-Gia, "Network Virtualization: Implementation Steps Towards the Future Internet," in *Electronic Communications of the EASST, Volume 17: Kommunikation in Verteilten Systemen 2009*, Kassel, Germany, Mar. 2009.

[18] Thomas Zinner, Phuoc Tran-Gia, Kurt Tutschku, and Akihiro Nakao, "Performance Evaluation of Packet Reordering on Concurrent Multipath Transmissions for Transport Virtualisation," *Int. J. Communication Networks and Distributed Systems*, vol. 6, no. 3, pp. 322–340, 2011.

[19] Yoav Nebat and Moshe Sidi., "Resequencing Considerations in Parallel Downloads," in *Proc. 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, 2002.

[20] Y. Nebat et al., "Parallel Downloads for Streaming Applications – a Resequencing Analysis," *Performance Evaluation*, vol. 63, no. 1, pp. 15–35, Jan. 2006.