BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS
DEPARTMENT OF TELECOMMUNICATIONS

Some Performance and Design Aspects of Overlay Networks

Ph.D. Thesis summary

by

Csaba Király

Advisors:

Dr. Tien Van Do, BUTE
Dr. Renato Lo Cigno, UNITN

Budapest, Hungary
2012

1 Introduction

In recent years, services built using overlay network technologies have seen a tremendous increase both in their number and in the sheer amount of bytes transferred. Proliferation of broadband access, stronger CPUs and advances in high-level programming languages (such as simplified networking APIs and libraries) have led to system designs that handle the network with a high level of abstraction, masquerading network connections as simple pipes, making the application almost independent, but also unaware of the underlying networking technologies. This trend has led to systems where several networking functionalities are implemented in application layer and the interaction between these high level algorithms and the network itself is a decisive factor in system performance.

Two prominent examples of the power of such overlay networks are Anonymous Routing Overlays and Peer-to-Peer Streaming Overlays. In both cases, the use of application layer overlay techniques allowed the quick development and introduction of large scale distributed systems, boosting vast areas of research, experimentation, as well as real services.

Anonymous Routing Overlays provide privacy enhancing services, protecting the identity of communicating parties both from each other and from eavesdroppers or other malicious attackers. When we contrast this to the amount of information revealed in a single IP packet on the Internet, it becomes evident how fundamental this service is in certain use cases, most notably to enable freedom of speech over the Internet. On the Internet, when node A sends an IP packet to node B, the packet will contain both A's and B's IP address in clear form all along the path. This means that A's identity is revealed to B, therefore services that would require anonymity such as anonymous voting or anonymous payment are difficult to implement. Even if anonymity is granted at higher levels, the IP layer reveals real identities, potentially breaking all the protection provided in higher layers. Sending A's and B's address also means that B's identity is revealed to A's ISP, and indirectly to A's government, hindering privacy and allowing the detailed profiling of individuals. Finally, anyone eavesdropping on any single point of the path becomes aware of the communication between the two parties.

Anonymous Routing Overlays provide techniques to hide these identifiers and thus anonymize communication. The most well known example of such systems is Tor [12], used by hundreds of thousands of people all over the world. Although Tor is already deployed as a service, it is known to suffer from serious performance bottlenecks [11]: available bandwidth is relatively low and end-to-end delays can be an order of magnitude higher compared to normal IP, effectively limiting Tor's applicability to only a few applications.

In the case of live Peer-to-Peer streaming, the bottleneck it overcomes is again to be found in the IP layer. IP multicast, which works well in a single operator's managed domain, is simply not applicable on the Internet scale due to administrative boundaries and interoperability problems [17]. Since changing the behavior of the IP layer is almost impossible, overlay technologies have been developed to broadcast live TV channels over the Internet to large audiences with optimized resource usage [15, 26]. Services have already been deployed in this case as well (PPlive, SopCast, PPStream, TV Ants, to name just a few). However, there are still numerous performance issues to analyze and solve in order to optimize streaming performance and resource utilization.

2 Research objectives

The general objective of this thesis is to identify key performance issues and bottlenecks of overlay networks, analyze their impact on state of the art systems, and to provide new techniques to improve performance. The thesis is divided into two main parts.

The goal of the first part is to analyze performance characteristics of Anonymous Routing Overlays, and to overcome their most important performance problem, i.e. high end-to-end delays.

First, we provide deeper insight into the transport mechanisms hidden behind some of these abstract overlay technologies by showing an analysis of TCP dynamics in the case of short-lived connections, deriving detailed performance characteristics (such as completion time distribution) through an open multiclass queuing network model of TCP [J1].

Performance differences between TCP based and datagram based overlay networks are then studied through the example of Anonymous Routing Overlays. First, onion routing networks and their most prevalent example, the Tor application is described. We present an analysis of some of the networking technologies applied and their consequences. Then we propose a novel system design for anonymous routing, called IPpriv, an overlay network similar in its scope to Tor, but different in its design choices and performance characteristics [C3]. The peculiarity of IPpriv is not just its datagram based design approach, but that it entirely relies on IPsec, therefore providing improved performance due to kernel level or router based operation.

In the second part of the thesis another widely used example of overlay networks will be considered: peer-to-peer (P2P) networks. More specifically, our goal is to analyze and improve the performance of live P2P Streaming Overlays.

In P2P streaming, scheduling is the decision of what to send and whom to send it to. First, we present some performance bounds valid for idealized conditions, and propose a novel scheduling algorithm that achieves this bound under these simplified assumptions [B1]. We extend this study by lifting some of the assumptions, namely, bandwidth homogeneity, describing a network-aware variant of this scheduling algorithm that considers the bandwidth of other peers and outperforms other algorithms known from the literature [C2, C1].

Finally, we provide an example of how overlay networks allow the application of some networking concepts even if the underlying network does not provide support for it. Namely, packet priority support is mostly missing on the large-scale Internet, still, applications serving as ”overlay routers” allow us to implement effective prioritization techniques. We introduce the design of a media-aware scheduler and show how it influences system performance, including a study of the effective end-user perceived video quality improvements [C7, C8].

3 Research methodology

We tackle two largely different areas of overlay networks in the thesis, namely Anonymous Routing and live P2P Streaming.

In both cases, we first analyze the performance of existing systems in detail through analytical techniques. We do not analyze the system as a whole, but we concentrate on some key areas hindering performance.

Then we introduce new algorithms and/or system designs, and evaluate the performance compared to state of the art solutions. Depending on the complexity of the problem, we use analytical techniques, simulations, or experiments using real implementations. Analytical results are also validated through simulation and/or controlled experiments.

TCP related results are validated using the well-known ns-2 simulator (widely used to verify results on TCP protocol behavior [6–8, 14]). IPpriv is directly evaluated through controlled experiments. Care has been taken to isolate various factors contributing to performance by running experiments both in our lab and over the Internet.

In the case of P2P Streaming, no simulators were readily available for the task. Therefore, two simulators were developed together with the algorithms to evaluate performance. First, SSSim (the Simple and Scalable Simulator) was developed as a round-based simulator that provides a one-to-one mapping to the idealized model used in the analytical part of the work [C12]. Later, the more complex event-based P2PTVsim simulator was developed to evaluate performance also in scenarios where some of the assumptions of the mathematical model have been relaxed [1].

4 New results

New results are organized in two main groups: results related to Anonymous Routing (including some generic results on TCP as well) and results related to P2P Streaming.

4.1 Performance of Anonymous Routing Overlays

Claim 1.1. *[J1, C4] I have proposed a method for the calculation of the completion time distribution of short-lived TCP connections based on OMQN (open multiclass queuing network) models of protocol behavior. (see Section 2.4 of the thesis)*

As mentioned earlier, several overlay designs use TCP as their underlying transport mechanism. While a lot is known about TCP's average throughput for long-lived connections [2, 3, 20], the completion time of short-lived connections is less explored.

Several papers [8, 19, 23, 24] proposed methods to obtain the *average* completion time of short-lived TCP connections. However, depending on the occurrence of losses, the life of an N-segment long TCP connection may exhibit many different behaviors, and may correspondingly have largely different durations. The variability of QoS requirements for different types of services, hence different types of TCP connections, requires the computation of more sophisticated performance metrics, such as completion time *distributions*, and *quantiles* of the completion time for TCP connections that need to transfer a given number of segments.

I have developed a new technique for the analytical evaluation of distributions (and quantiles) of the completion time of short-lived TCP connections. The proposed technique derives from known open multiclass queuing network (OMQN) models of the TCP protocol and computes a discrete approximation, with arbitrary accuracy, of the distribution of sojourn times of customers in the OMQN, which corresponds to the distribution of completion times of the modeled TCP connections.

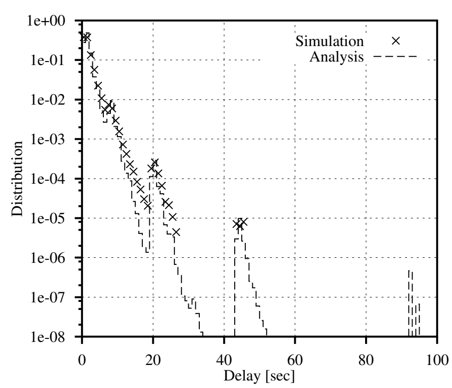

Figure 1: Probability density function of the completion time of an $N_s = 100$ segment TCP transfer. Precision with Analysis and Simulation

The method allows the computation of completion time distributions with high precision, as shown on the example of Fig. 1, comparing results obtained through simulation and OMQN model based analysis.

We simulated 1.125 million flows in order to obtain the simulation curves (we plot a point if we have at least 5 samples), and it took 3 hours. The analytical results were computed in less than 2 s.

Claim 1.2. *[J1, C4] I have shown — through ns-2 simulations — that the model predicts TCP connection completion times with high precision in various networking scenarios. The proposed technique is computationally efficient, and its asymptotic complexity does not depend on the network topology, on the number of concurrent flows, and on other network parameters. (see Section 2.5 of the thesis)*

The method has been evaluated and compared to simulation results in various realistic scenarios, including single- and multiple-bottleneck scenarios, as well as UDP cross-traffic.

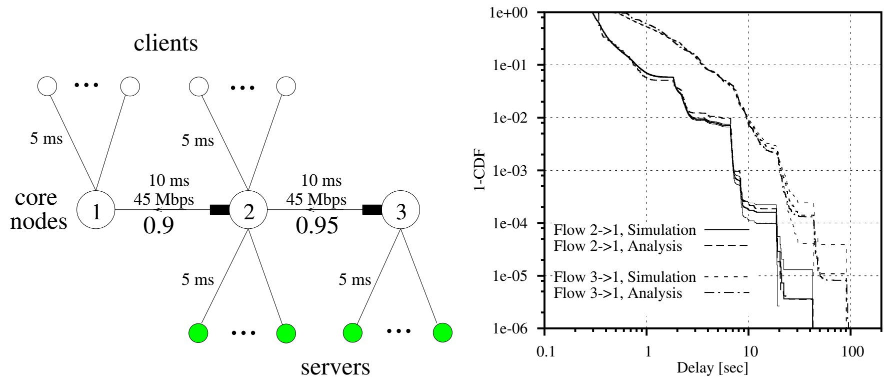

Figure 2: Two-bottleneck scenario. Setup (left) and results (right). Performance of flows from servers behind router 2 to clients behind router 1 and from servers behind router 3 to clients behind router 1

Fig. 2 shows the multiple-bottleneck scenario and the completion time distributions for various flows passing through the network, obtained with different methods. Both the results from simulation (thinner lines show 99% confidence intervals) and those obtained using the analytical approach are shown.

Several interesting features of the completion time distribution can also be highlighted with the proposed method: e.g. it allows the exploration of quantiles over a large parameter space in a reasonable time, as shown on Fig. 3. This would otherwise be computationally too complex to derive through simulation.

Claim 1.3. *[C3, C5, C6] I have proposed a new system design for anonymous routing overlays, based entirely on standard IPsec functionality, thus avoiding bottlenecks found in other systems due to the use of TCP tunnels. (see Section 3.3 of the thesis)*

Based on previous experience in performance issues rooted in the use of TCP as overlay tunnel technology [C5], I have developed an anonymous overlay network design to overcome these problems.

Privacy on the "Internet" was first discussed by David Chaum [9], who introduced the concept of anonymous communications. After this initial discussion,

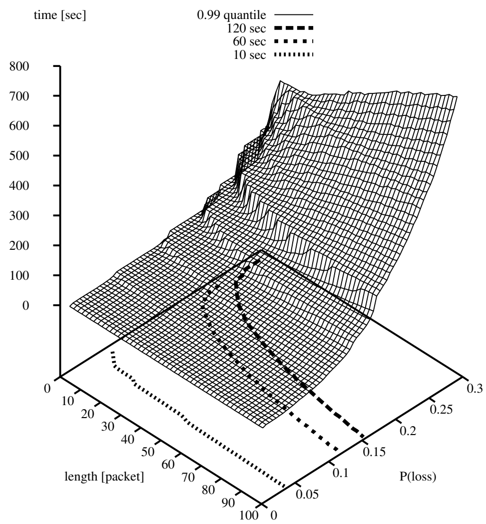

Figure 3: Analytical results: 0.99 completion time quantiles as a function of the connection length N_s and TCP segment loss probability P_L.

however, the topic remained confined to small communities for a long time. The concept that was originally intended for anonymous e-mail only, was later extended to low-latency communications by application level overlay systems such as Tor [12], Tarzan [13] or Freedom [5]. Low-latency in this context means that the communication is not based on storing and forwarding *entire application level messages* in intermediate nodes, but rather small information units are forwarded, reducing end-to-end delays to seconds.

Our new design, called *IPpriv*, not just improves performance by a datagram-based design, but it builds entirely on standard IPsec features, allowing a novel operator-centric deployment model for Anonymous Routing Overlays. We show that IP address protection can be provided *within the network layer itself*, without the need of building cumbersome and hardly-scalable application-level overlays as other Anonymous Routing solutions do.

Claim 1.4. *[C3, C5] I have implemented the design of Claim 1.3 and confronted its performance with a TCP based implementation in both emulated and real environments, demonstrating that end-to-end delays can be largely improved in anonymous routing systems. (see Section 3.4 of the thesis)*

In order to gain a better understanding of the impact of the overlay networking technology and of the 'price' of anonymization, performance measurements have been conducted both in local controlled testbeds and over the Internet. We have used two sets of overlay nodes for our tests: a circuit of three nodes in our lab (*local testbed*, lower part of Fig. 4), and a circuit of three nodes in different European universities (*Internet testbed*, upper left part of Fig. 4). In both circuits we had Tor as well as IPpriv installed, thus having four different cases. We used direct, non-anonymous download as the fifth reference case.

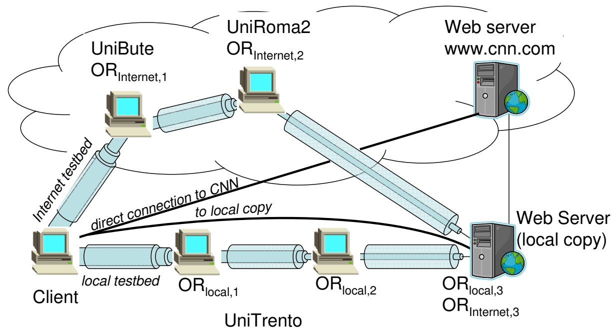

Figure 4: Testbed configuration and experiment setup

A performance comparison with Tor showed that on the one hand the overhead (in terms of additional transmitted bytes) of IPpriv is almost identical to application level solutions, while on the other hand the performance in terms of download delay is 3–4 times faster with the IPsec solution.

4.2 Performance of Peer-to-Peer Streaming Overlays

The second group of new results is centered on the performance of live P2P Streaming systems, more specifically on the problem of chunk and peer scheduling.

Chunk and peer selection strategies (or scheduling) are among the main drivers of live P2P Streaming system performance (other factors being overlay topology optimizations [C9], chunk exchange and signaling protocol design, aggregate and P2P transmission rate control [C10, C11], etc.).

Claim 2.1. *[B1] I have proposed the new Earliest-Latest peer scheduling algorithm (ELp) for P2P streaming systems, and proved that — in combination with the well-known Latest Useful chunk scheduler (LUc) — it achieves strict optimality in idealized conditions. (see Section 4.4 of the thesis)*

It is well known that the lower bound on the dissemination delay of any piece of information, given that nodes have exactly the bandwidth necessary for the streaming itself, is $\delta_{lb} = (\lceil \log_2(N) \rceil + 1)T$ where N is the number of nodes in the system and T is the transmission time of a single chunk. The bound comes from the fact that each node can transmit the information only after receiving it, and the number of nodes owning the chunk at most doubles every T. It is also known [18] that centralized schedulers can distribute every chunk of a stream in exactly δ_{lb}. Also, in [4] it was proved that a bound holds for several *distributed* schedulers if $N \to \infty$ and $M_c \to \infty$ (M_c is the number of chunks). However, when real-time distribution systems are considered, such an asymptotic bound is not equivalent to δ_{lb}.

A key observation when looking for a scheduler that provides strict optimality can be derived by looking at schedulers from the *redistribution perspective*: to achieve optimal diffusion delay, it must be ensured that the number of peers owning a given chunk doubles in each cycle. In other words, a peer that receives a chunk should immediately start redistributing it, and it should continue the distribution of the same chunk till it gets diffused in the whole system.

Based on this observation, I have designed the Earliest-Latest peer scheduler and formally proved its optimality in the framework of the idealized diffusion model considered also in [4].

Claim 2.2. *[B1] I have shown that the ELp algorithm still achieves strict optimality in combination with a whole class of Deadline-based chunk scheduling algorithms (DLc) in the same idealized conditions. This combination is also robust to impairments when relaxing some idealistic assumptions of the mathematical model. (see Section 4.5 of the thesis)*

While ELp/LUc provides optimal performance in an idealized diffusion model, it is known that LUc itself is fragile to a number of realistic impairments due to its greedy nature. Restrictions to the topology or differences in chunk sizes can easily create situations when some chunk overtakes a later one, effectively blocking its diffusion in the system.

Taking inspiration from "Deadline Driven Scheduling" [16], we have designed the novel Deadline-based chunk scheduler (DLc) that overcomes many of these shortcomings of LUc.

We present an analytical proof that — when combined with ELp — an entire class of Deadline-based chunk schedulers provide strict optimality in idealized conditions.

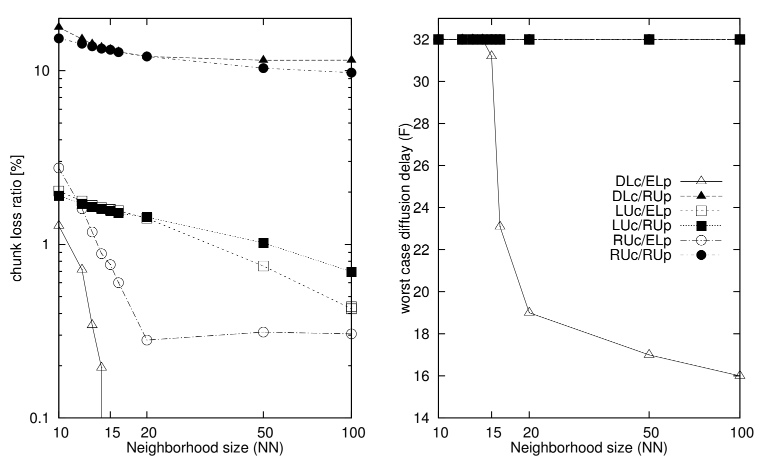

Figure 5: Chunk loss and worst-case diffusion delay as a function of the neighborhood size (10000 peers, buffer size of 32); note that worst-case delay is 32 for all schedulers with losses

We also demonstrate by simulations that the DLc/ELp scheduler is resilient to the reduction of the neighborhood size down to values as small as $\log_2(N)$. Figure 5 illustrates some of the simulation results.

Claim 2.3. *[C2, C1] I have proposed a network-aware extension of the algorithm of Claim 2.1, and analyzed its performance in heterogeneous network scenarios. The new BAELp algorithm is equivalent to ELp in the homogeneous case, and it outperforms other known algorithms in the heterogeneous case. (see Chapter 5 of the thesis)*

Contrary to the assumptions of the model used in Claims 2.1 and 2.2, networks are neither ideal nor homogeneous. The choice of the destination peer will affect the performance in transferring the chunk to it, thus indirectly it will also affect how this specific chunk will be diffused in the future: peers more endowed with resources will diffuse the chunk more efficiently than peers with scarce resources.

Previous studies on this topic either focus on asymptotic properties of distributed gossiping algorithms [22], or probabilistic bounds for specific well known algorithms [4], or do not distinguish clearly the scheduling problem from the protocol to exchange the information ([21, 25, 26].

A peer scheduler, in order to be robust to different bandwidth distribution scenarios, should be both bandwidth aware [10] and select peers that are in the best conditions to redistribute the chunk as ELp does [B1].

A first way to integrate ELp with some kind of bandwidth awareness is to use hierarchical scheduling. Two possible hierarchical combinations can be implemented: $\mathrm{EL_{BA}p}$ and $\mathrm{BA_{EL}p}$. $\mathrm{EL_{BA}p}$ uses ELp to select a set of peers having the earliest latest chunk, and then uses a bandwidth aware strategy to select the peer having the highest output bandwidth among them. $\mathrm{BA_{EL}p}$ first uses a bandwidth aware scheduler to select the set of peers having the highest output bandwidths and then applies ELp scheduling to this set.

Although hierarchical scheduling can be very effective in some situations, a better integration of the two scheduling algorithms can improve the system's performance. The new BAELp (Bandwidth Aware Earliest Latest) peer scheduler changes ELp to consider both the peer's redistribution power and the foreseen arrival time of the chunk to the peer.

The scheduling performance of BAELp has been confronted with $\mathrm{BA_{EL}p}$, $\mathrm{EL_{BA}p}$, and several other peer schedulers from the literature in various heterogeneous bandwidth scenarios (Figure 6 shows an example), and we have found that BAELp outperforms all the other scheduling algorithms in a variety of different conditions. This peer scheduler is also able to cope with the presence of free riders, namely peers that do not contribute to uploading chunks and sustaining the stream.

Claim 2.4. *[C7, C8] Based on the algorithm of Claim 2.2, I have proposed a novel media-aware scheduling algorithm that assigns priorities to chunks based on their media content. I have shown that the new algorithm improves overall system performance. (see Chapters 6 and 7 of the thesis)*

Up to this point, we have dealt with an abstract notion of chunk (simply treating it as the unit block of data) and measured performance only in terms

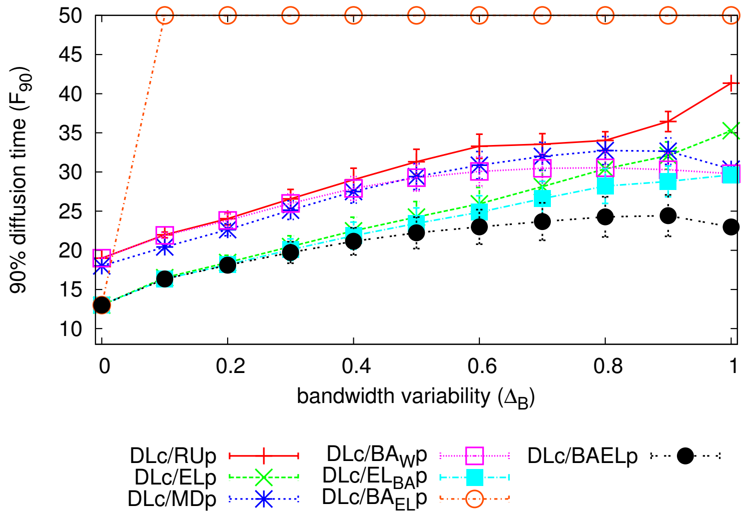

Figure 6: Scheduling performance as a function of bandwidth heterogeneity (neighborhood size 20; playout delay 50; $N = 600$ peers, $M_c = 2000$ chunks.

of chunk diffusion delay distribution, or — in another view — as percentage of chunks received before a given deadline. However, the quality perceived by the user also depends on the way chunks are generated from the media stream, on the encoding bitrate, on the audio and video codecs used, etc.

Traditional P2P distribution systems generally broadcast a file to multiple peers by splitting it into fixed-size chunks. Many P2P streaming systems adapt this approach to media streaming as well, dividing the encoded media into fixed size chunks, independently from the structure of the underlying media stream.

In this chapter, a more direct measure of the received media quality is used to analyze the relationship between media encoding, chunkisation policies (i.e. the way of splitting the media stream into chunks), playout delay, chunk loss, and the final received media quality.

Priorities are assigned to chunks based on the analysis of their media content, and an extension of the DLc chunk scheduling algorithm is proposed to handle these priorities in the chunk diffusion process.

We show how the overlay network allows for the application of prioritization even if the underlying network does not provide support for it. Detailed results on end-user perceived video quality enable the evaluation of the effect of system parameters that otherwise — based on chunk loss and delay metrics — would be impossible to evaluate. The effect of encoding rate on perceived quality in different user groups is illustrated in Figure 7.

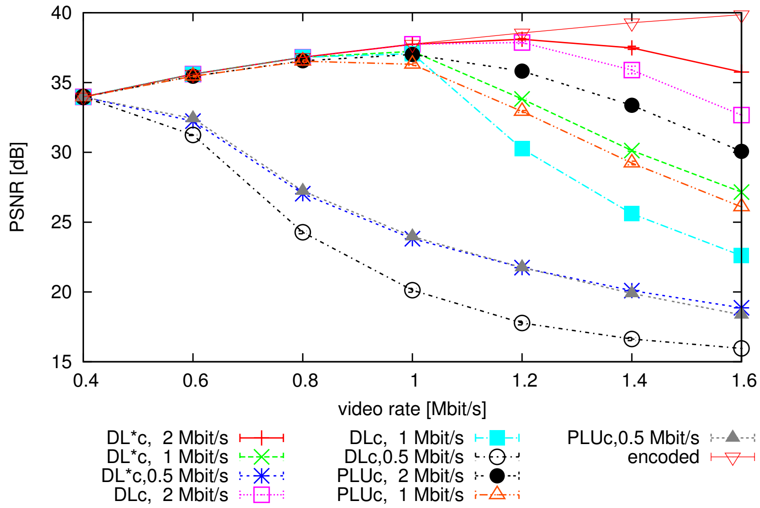

Figure 7: PSNR with different chunk schedulers (DL*c: priority-aware DLc; PLUc: LUc scheduler with strict priorities), at different peer classes (heterogeneous network with 3 peer classes: 2 Mbit/s; 1 Mbit/s and 0.5 Mbit/s peers), as a function of video encoding rate

5 Applicability of results

Results on TCP performance are applicable in the performance analysis of a wide range of systems. In fact, TCP connections are being used in a large variety of ways, often as a consequence of using higher level programming libraries, without even knowing what is underneath. It is enough to mention the fast opening and closing of TCP connections in some P2P file sharing applications, the use of several parallel connections in browsers, or various HTTP based IPC mechanisms often used in composite web services to demonstrate that detailed characterization of TCP's short-term behavior is essential for understanding system behavior in some of today's complex systems.

IPpriv presents a fast and efficient datagram based anonymous system design. On the one hand, it has some applicability limitations: namely, it does require system and network administrators to enable IPsec policies in some selected routers. On the other hand it also conveys two important messages: first, many of the performance bottlenecks attributed to "routing around the world" are not in fact due to the increased length of the transmission path, but rather to the misguided selection of underlying overlay techniques. Second, it presents a realistic alternative to end-user based anonymous routing overlays, providing techniques that can anonymize traffic in the network itself with standard carrier grade routing equipment, without introducing application level routing components.

Results on live P2P Streaming are not only applicable, but they are already being applied in the open-source PeerStreamer streaming framework. The results presented and their improvements are also being discussed in the IETF PPSP working group, and will hopefully influence the design of the future peer-to-peer streaming protocol standards.

6 Acknowledgement

The author would like to express his gratitude for the financial support received from the following organizations: Budapest University of Technology and Economics, University of Trento, Consorzio Nazionale Interuniversitario per le Telecomunicazioni, the FP6 project DISCREET and the FP7 project NAPA-WINE.

References

[1] The P2P-TVSim simulator. URL `http://www.napa-wine.eu/cgi-bin/twiki/view/Public/P2PTVSim`.

[2] E. Altman, K. Avrachenkov, and C. Barakat. A stochastic model of tcp/ip with stationary random losses. In *Proceedings of the conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, SIGCOMM '00, pages 231–242, New York, NY, USA, 2000. ACM. ISBN 1-58113-223-9. doi: http://doi.acm.org/10.1145/347059.347549. URL `http://doi.acm.org/10.1145/347059.347549`.

[3] A. Berger and Y. Kogan. Dimensioning bandwidth for elastic traffic in high-speed data networks. *Networking, IEEE/ACM Transactions on*, 8(5):643 –654, oct 2000. ISSN 1063-6692. doi: 10.1109/90.879350.

[4] T. Bonald, L. Massoulié, F. Mathieu, D. Perino, and A. Twigg. Epidemic live streaming: optimal performance trade-offs. In *Proc. of SIGMETRICS*, pages 325–336, Annapolis, Maryland, USA, June 2008. ACM. ISBN 978-1-60558-005-0.

[5] P. Boucher, A. Shostack, and I. Goldberg. Freedom systems 2.0 architecture. White paper, Zero Knowledge Systems, Inc., December 2000.

[6] G. Buchholcz, A. Gricser, T. Ziegler, and T. V. Do. Explicit loss notification to improve tcp performance over wireless networks. In *HSNMC'03*, pages 481–492, 2003.

[7] G. Buchholcz, T. Ziegler, and T. V. Do. Tcp-eln: On the protocol aspects and performance of explicit loss notification for tcp over wireless networks. In *Proceedings of the First International Conference on Wireless Internet*, pages 172–179, Washington, DC, USA, 2005. IEEE Computer Society. ISBN

0-7695-2382-X. doi: 10.1109/WICON.2005.31. URL `http://dl.acm.org/citation.cfm?id=1092359.1092623`.

[8] N. Cardwell, S. Savage, and T. Anderson. Modeling tcp latency. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 3, pages 1742 –1751 vol.3, mar 2000. doi: 10.1109/INFCOM.2000.832574.

[9] D. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 4(2), February 1981.

[10] A. Couto da Silva, E. Leonardi, M. Mellia, and M. Meo. A bandwidth-aware scheduling strategy for p2p-tv systems. In *Proceedings of the 8th International Conference on Peer-to-Peer Computing 2008 (P2P'08)*, Aachen, September 2008.

[11] R. Dingledine and S. J. Murdoch. Why tor is slow and what we're going to do about it? Technical report, Tor project, March 2009.

[12] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. In *Proceedings of the 13th USENIX Security Symposium*, August 2004.

[13] M. J. Freedman and R. Morris. Tarzan: A peer-to-peer anonymizing network layer. In *Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS 2002)*, Washington, DC, November 2002.

[14] M. Garetto, R. Lo Cigno, M. Meo, and M. Marsan. Closed queueing network models of interacting long-lived tcp flows. *Networking, IEEE/ACM Transactions on*, 12(2):300 – 311, april 2004. ISSN 1063-6692. doi: 10.1109/TNET.2004.826297.

[15] B. Li, Y. Qu, Y. Keung, S. Xie, C. Lin, J. Liu, and X. Zhang. Inside the new coolstreaming: Principles, measurements and performance implications. In *Proc. of IEEE INFOCOM 2008*, Phoenix, AZ, USA, Apr. 15–17 2008.

[16] C. L. Liu and J. Layland. Scheduling alghorithms for multiprogramming in a hard real-time environment. *Journal of the ACM*, 20(1), 1973.

[17] J. Liu, S. G. Rao, B. Li, and H. Zhang. Opportunities and challenges of peer-to-peer internet video broadcast. *Proceedings of the IEEE*, 96(1):11–24, 2008. doi: http://dx.doi.org/10.1109/JPROC.2007.909921. URL `http://dx.doi.org/10.1109/JPROC.2007.909921`.

[18] Y. Liu. On the minimum delay peer-to-peer video streaming: how realtime can it be? In *MULTIMEDIA '07: Proceedings of the 15th international conference on Multimedia*, pages 127–136, Augsburg, Germany, September 2007.

ACM. ISBN 978-1-59593-702-5. doi: http://doi.acm.org/10.1145/1291233.1291259.

[19] M. Mellia and H. Zhang. Tcp model for short lived flows. *Communications Letters, IEEE*, 6(2):85 –87, feb 2002. ISSN 1089-7798. doi: 10.1109/4234.984705.

[20] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose. Modeling tcp reno performance: a simple model and its empirical validation. *Networking, IEEE/ACM Transactions on*, 8(2):133 –145, apr 2000. ISSN 1063-6692. doi: 10.1109/90.842137.

[21] D. Ren, Y. T. H. Li, and S. H. G. Chan. On reducing mesh delay for peer-to-peer live streaming. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pages 1058–1066, 2008. doi: http://dx.doi.org/10.1109/INFOCOM.2008.160. URL `http://dx.doi.org/10.1109/INFOCOM.2008.160`.

[22] S. Sanghavi, B. Hajek, and L. Massoulié. Gossiping with multiple messages. In *Proceedings of IEEE INFOCOM 2007*, pages 2135–2143, Anchorage, Alaska, USA, May 2007.

[23] B. Sikdar, S. Kalyanaraman, and K. Vastola. Analytic models and comparative study of the latency and steady-state throughput of tcp tahoe, reno and sack. In *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, volume 3, pages 1781 –1787 vol.3, 2001. doi: 10.1109/GLOCOM.2001.965883.

[24] B. Sikdar, S. Kalyanaraman, and K. S. Vastola. An integrated model for the latency and steady-state throughput of tcp connections. *Perform. Eval.*, 46:139–154, October 2001. ISSN 0166-5316. doi: 10.1016/S0166-5316(01)00048-7. URL `http://dl.acm.org/citation.cfm?id=570289.570294`.

[25] M. Zhang, Q. Zhang, L. Sun, and S. Yang. Understanding the power of pull-based streaming protocol: Can we do better? In *IEEE JSAC, special issue on Advances in Peer-to-Peer Streaming Systems*, 2007.

[26] X. Zhang, J. Liu, and T. shing Peter Yum. Coolstreaming/donet: A data-driven overlay network for peer-to-peer live media streaming. In *in IEEE Infocom*, 2005.

Publications

Journal papers

[J1] C. Kiraly, M. Garetto, M. Meo, M. Ajmone Marsan, and R. Lo Cigno. Analytical computation of completion time distributions of short-lived tcp connections. *Performance Evaluation*, 59:179 – 197, 02/2005 2005. ISSN 01665316. doi: 10.1016/j.peva.2004.07.002.

[J2] R. Birke, E. Leonardi, M. Mellia, A. Bakay, T. Szemethy, C. Kiraly, R. Lo Cigno, F. Mathieu, L. Muscariello, S. Niccolini, J. Seedorf, and G. Tropea. Architecture of a network-aware p2p-tv application: the napa-wine approach. *Communications Magazine, IEEE*, 49(6):154 –163, june 2011. ISSN 0163-6804. doi: 10.1109/MCOM.2011.5784001.

[J3] C. Kiraly, T. V. Do. Sztratoszferikus platformok és szolgáltatásaik. (in Hungarian) *Magyar Távk ozlés* (Hungarian Telecommunications Journal) 2004/3 :23 – 26, 2004.

Book chapters

[B1] L. Abeni, C. Kiraly, and R. Lo Cigno. *On the Optimal Scheduling of Streaming Applications in Unstructured Meshes*, volume 5550 of *Lecture Notes in Computer Science*, pages 117–130. Springer Berlin / Heidelberg, 2009.

Conference and workshop papers

[C1] L. Abeni, C. Kiraly, and R. Lo Cigno. Scheduling P2P multimedia streams: Can we achieve performance and robustness? In *2009 3rd International Conference on Internet Multimedia Services Architecture and Application (IMSAA)*, pages 1 – 6, Bangalore, India, 2009. IEEE. ISBN 978-1-4244-4792-3. doi: 10.1109/IMSAA.2009.5439486.

[C2] L. Abeni, C. Kiraly, and R. Lo Cigno. Robust scheduling of video streams in network-aware P2P applications. In *ICC 2010 - 2010 IEEE International Conference on Communications*, pages 1 – 5, Cape Town, South Africa, 2010. IEEE. ISBN 978-1-4244-6402-9. doi: 10.1109/ICC.2010.5502542.

[C3] C. Kiraly and R. Lo Cigno. IPsec-based anonymous networking: A working implementation. In *ICC 2009 - 2009 IEEE International Conference on Communications*, pages 1 – 5, Dresden, Germany, 2009. IEEE. doi: 10.1109/ICC.2009.5199102.

[C4] C. Kiraly, M. Garetto, R. Lo Cigno, M. Meo, and M. Ajmone Marsan. Computation of the completion time distribution of short-lived TCP connections. In *Performance Modelling and Evaluation of Heterogeneous Networks (HET-NETs '03)*, Ilkley, England, 2003.

[C5] C. Kiraly, G. Bianchi, and R. Lo Cigno. Solving performance issues in anonymization overlays with a L3 approach. Technical Report TR-DISI-08-041, University of Trento, September 2008.

[C6] C. Kiraly, S. Teofili, G. Bianchi, R. Lo Cigno, M. Nardelli, and E. Delzeri. *Traffic Flow Confidentiality in IPsec: protocol and implementation*, volume 262 of *IFIP International Federation for Information Processing*, pages 311 – 324. Springer US, Boston, MA, 2008. ISBN 978-0-387-79026-8.

[C7] C. Kiraly, L. Abeni, and R. Lo Cigno. Effects of P2P streaming on video quality. In *ICC 2010 - 2010 IEEE International Conference on Communications*, pages 1 – 5, Cape Town, South Africa, 2010. IEEE. ISBN 978-1-4244-6402-9. doi: 10.1109/ICC.2010.5502003.

[C8] C. Kiraly, R. Lo Cigno, and L. Abeni. Deadline-based differentiation in P2P streaming. In *GLOBECOM 2010 - 2010 IEEE Global Communications Conference*, pages 1 – 6, Miami, FL, USA, 2010. IEEE. ISBN 978-1-4244-5636-9. doi: 10.1109/GLOCOM.2010.5683450.

[C9] C. Kiraly and R. Lo Cigno. On the effects of overlay localization on P2P networks. In *IEEE INFOCOM Workshops 2009 - IEEE Conference on Computer Communications Workshops*, pages 1 – 2, Rio de Janeiro, Brazil, 2009. IEEE. ISBN 978-1-4244-3968-3. doi: 10.1109/INFCOMW.2009.5072170.

[C10] R. Birke, C. Kiraly, E. Leonardi, M. Mellia, M. Meo, and S. Traverso. Hose rate control for P2P-TV streaming systems. In *Peer-to-Peer Computing (P2P), 2011 IEEE International Conference on*, pages 202 –205, 31 2011-sept. 2 2011. doi: 10.1109/P2P.2011.6038736.

[C11] R. Birke, C. Kiraly, E. Leonardi, M. Mellia, M. Meo, and S. Traverso. Aggregate flow control for P2P-TV streaming systems. In *NEM Summit 2011*, Sept. 2011.

[C12] L. Abeni, C. Kiraly, and R. Lo Cigno. SSSim: a simple and scalable simulator for P2P streaming systems. In *2009 IEEE 14th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pages 1 – 6, Pisa, Italy, 2009. IEEE. ISBN 978-1-4244-3532-6. doi: 10.1109/CAMAD.2009.5161473.