A DOCTORAL THESIS PROPOSAL:

—

IMPLEMENTATION, EVALUATION AND MODELING OF IP NETWORK BANDWIDTH MEASUREMENT METHODS

Andreas Johnsson
The Department of Computer Science and Electronics
Mälardalen University
Sweden
andreas.johnsson@mdh.se

1. Project description

A team of researchers from Ericsson Research (Svante Ekelin, Bob Melander, Jan-Erik Mångs), Linköping University (Erik Hartikainen), Mälardalen University (Mats Björkman and Andreas Johnsson) and the Swedish Institute of Computer Science (Martin Nilsson) jointly work on a project that aims at development of novel methods for active end-to-end bandwidth estimation in IP networks as well as development of the theoretical understanding of the underlying concepts.

The joint project is currently, and has previously been, funded by several sponsors; TFR (later VR) through the Traffic measurement and analysis project, VINNOVA through the Evalunet project, KKS through the Evalunet II project and SSF through the ARTES program. Funding also comes from Ericsson Research in Kista and Ericsson in Katrineholm.

The goal of the joint project is to address challenges within bandwidth estimation research, both practical and theoretical. Some of the challenges are listed below.

- Find new and lightweight active end-to-end methods to obtain the bandwidth estimates in reasonable time without consuming too much resources of the computer.
  - Can bandwidth estimation methods fit inside small devices such as cellular phones?
- Active bandwidth estimation methods affect other traffic flows on the network. How can the impact be minimized while at the same time keep statistically sound measurement samples?
  - The injection of probe packets must be scalable, otherwise the benefits of active measurements can not be spread to a wider application area.
  - Is it possible to use application traffic instead of probe packets to obtain measurement samples?
  - Are there ways of integrating bandwidth measurements and TCP?
- Minimize the time between the actual measurement of the path and the presentation of the bandwidth estimate. This is especially important when applications need estimates of the available bandwidth to operate properly. Old estimates are outdated since the network conditions may have changed.
- Further develop the theoretical understanding of the challenges with bandwidth estimation.

- It is important to describe and model how old and new methods are affected when network technologies, such as 802.11 and 3G, are combined with the ordinary wired Internet.

My personal goal in the project is to complete my doctoral thesis by addressing some of the challenges above. This article aims at describing the completed work as well as the research to be carried out within the next 9 to 12 months. Since this is a doctoral thesis proposal the work and papers listed are my own or at least where I'm a co-author. Other papers within the joint project are left out.

The rest of this doctoral thesis proposal is organized as follows. In Section 2 a background to the active bandwidth measurement research area is given. In Section 3 a brief introduction to bandwidth measurement methods as well as theoretical literature is given. Section 4 points out my contribution within the joint project. Further, in Section 5 a list of papers intended to be part of my doctoral thesis is reviewed while other published papers are listed in Section 6. The article ends with an outline of the doctoral thesis in Section 8 and a time plan to its completion in Section 9.

2. Background: Why active end-to-end bandwidth measurements?

In large-scale networks, such as the Internet, it is impossible to keep a updated centralized view of network conditions for the entire network. This is partly because the amount of information needed would be so large that no single computer would be able to process it. Another problem is that the network is divided into subnetworks that are managed by different corporations, and the status of their network is considered a business secret. However, to obtain the network conditions and characteristics of an end-to-end path is a feasible task by using active measurement methods. Many of these methods do not require previous knowledge of the network at all.

Active end-to-end measurements of the network conditions and path characteristics, such as available bandwidth and link capacity, are important in best-effort IP networks where bandwidth resource reservations can not be made. The result from such active measurements is a way of describing the current status of the network path. Ideally, a forecast of the network conditions can also be made.

There exist many examples on applications that would benefit from having an updated view of the available bandwidth of a network path. Suppose a user wants to download a large file from a website on the Internet. By deploying active end-to-end measurements he or she (or the web browser itself) can decide which mirror site to use by estimating the available bandwidth to each of the mirror sites and thus estimate the minimum download time. Other examples of applications are network error diagnosis and load balancing in routers but also as a part of the adaptive machinery of network applications such as streaming audio and video.

3. Introduction to active end-to-end bandwidth measurements in IP networks

This section gives an introduction to the field of active bandwidth measurements.

The available bandwidth on a single link can be defined, in simple terms, as the unused portion of the link capacity during some time period. Depending on the time period the available bandwidth varies. The end-to-end available bandwidth over an entire network path is the minimum available bandwidth for each link constituting the path. That link defines the bottleneck link.

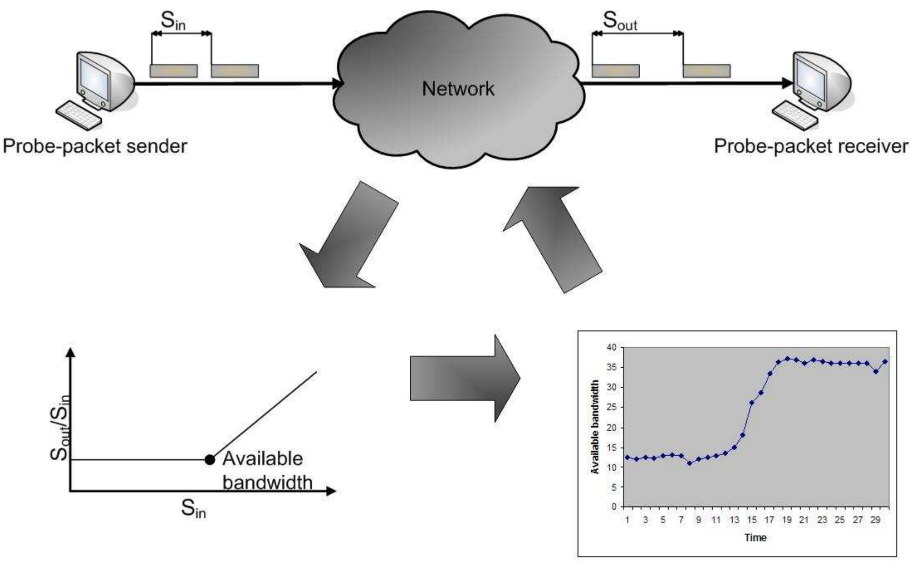

Figure 1. The process of bandwidth estimation.

In Figure 1 the essence of active end-to-end bandwidth measurements is shown. A probe-packet sender injects probe packets into the path at some input separation S_{in}. After the packets have traversed the network path the separation may have changed due to the network congestion, from S_{in} to an output separation S_{out}. If the network was congested, S_{out} is greater than S_{in}. Otherwise the input and output separations are the same. The relation between the two separations is seen in the left diagram. When the input separation is high the probe packets does not cause congestion and thus S_{out}/S_{in} equals 1. When the input separation exceeds the threshold for causing congestion the quotient will deviate from 1 as seen in the diagram. The two curves define the rate response curve of the path between the two communicating computers. The available bandwidth is obtained by calculating the intersection of the two lines; the intersection corresponds to an input rate defined as the available bandwidth. To obtain values corresponding to the sloping segment of the rate response curve the probe packets must cause congestion for a short time interval.

The estimate is then used in an application. In this case the estimate is plotted in a time diagram to the right in Figure 1. If the view of the network path is to be monitored over a time interval new probe packets are injected and the process is repeated.

3.1 Methods and tools

Well known active end-to-end bandwidth measurement methods are for example Pathchirp [Ribeiro et al., 2003], Pathload [Jain and Dovrolis, 2002b], PTR [Hu and Steenkiste, 2003], Spruce [Strauss et al., 2003] and TOPP [Melander et al., 2002]. Newly developed methods are ABGET [Antoniades et al., 2006] and BART [Ekelin et al., 2006]. The above methods can be classified in two categories: *direct probing* and *iterative probing* [Jain and Dovrolis, 2004]. Independent on the classification of the above methods they all have one thing in common: they all rely on self-induced congestion. That is, the methods must cause network congestion for a short time interval in order to estimate the available bandwidth as described above.

In direct probing each injection of probe packets gives a sample of the available bandwidth. *Spruce* is one such method. As other methods that relies on direct probing Spruce needs knowledge of the bottleneck link capacity. By comparing the input rate and the output rate the cross-traffic rate can be deducted. The available bandwidth is then the link capacity minus the cross-traffic rate.

In iterative probing no knowledge of end-to-end path is needed. Instead, probe packets are injected into the network at different rates to investigate whether the probe-packet rate is greater than or less than the available bandwidth.

TOPP uses linear regression in order to find the equation of the sloping segment of the rate response curve shown in Figure 3. Then it is a simple task to find the intersection of the two segments and thus the available bandwidth.

Pathload is based on observations of the one-way delay in probe-packet trains. If the input rate of the probe packets is above the available bandwidth the one-way delay will show an increasing trend within the train. If not, the one-way delay is more or less constant. To locate the available bandwidth Pathload performs a binary search.

Pathchirp injects probe packets in chirps instead of probe-packet trains. A chirp is a sequence of probe packets with exponentially decreasing time intervals between the probe packets. It is then possible to investigate a whole range of input rates in one chirp. The analysis is then performed by investigating the relation between delayed and non-delayed probe packets. One chirp gives an estimate of available bandwidth.

BART uses a Kalman filter in its analysis. A state vector describes the sloping segment, similar to the TOPP model. In BART the state vector is updated for each measurement sample. This way an estimate of the available bandwidth is obtained in real time. The Kalman filter tells how much to trust the new sample compared to the old estimate of the available bandwidth.

ABget relies on using a fake TCP client which instruments a TCP server to send packets according to a scheme similar to a probing session. This way it is not necessary to have one probe-packet sender and one receiver. The actual analysis of the received packets on the client side is done by using the ideas from Pathload.

A more in-depth review of available bandwidth measurement methods and tools can be found in [Prasad et al., 2003].

There also exist methods for estimating the link capacity (not just the available bandwidth) of the bottleneck link. Pathrate [Dovrolis et al., 2001] is one example and TOPP is another. Pathrate looks for a capacity mode obtained from packet-pair probing while TOPP produces the link capacity as a side effect when estimating the available bandwidth.

Pathneck [Hu et al., 2004] is a tool that uses ICMP packets and ordinary probe packets to find the actual location of a bottleneck link in an end-to-end path. However, it cannot estimate the link capacity nor the available bandwidth.

There exist a variety of older tools that uses ICMP and variable packet sizes in order to estimate the link capacity, but these methods fails in many cases [Prasad et al., 2002].

3.2 Theoretical work

Much theoretical research has been produced throughout the years. In this subsection a short review of the literature is given.

In [Dovrolis et al., 2001] it was studied how probe-packet pairs and trains were affected by cross traffic. Depending on the scheme used, different modes were visible in histograms of probe-packet time separations. In this study they defined one of the modes as the asymptotic dispersion rate, which was called the proportional share in [Melander et al., 2000]. This value corresponds to the proportion of the link capacity in use by the probe packets. In [Dovrolis et al., 2001] a study of the impact of probe packet size was also investigated. The findings lead forward to the link capacity estimation method Pathrate.

In [Melander et al., 2000][Melander et al., 2002] discussion about the proportional share, among other things, lead forward to the TOPP model to measure the end-to-end available bandwidth and link capacity.

In [Pásztor and Veitch, 2002] a delay-variation model for packet-pair like methods was developed. Using this model several histogram signatures could be identified. The modes in the histograms corresponded to bottleneck link and secondary bottleneck link capacities. The impact of the probe-packet size to link capacity estimation methods was also discussed.

The effect of layer-2 store-and-forward devices on per-hop capacity estimation was studied in [Prasad et al., 2002]. In this paper it was shown why many older methods for estimating bottleneck link capacity reported erroneous estimates.

A description on how the one-way delay within a probe-packet train varies depending on different parameters such as cross traffic and train length was presented in [Jain and Dovrolis, 2002a]. By investigating the one-way delay the authors could conclude that the one-way delay increased within a probe-packet train if the rate was above the available bandwidth. Otherwise the one-way delay remained constant. A method for detecting increasing trends was developed and resulted in the tool Pathload for estimating available bandwidth.

In [Liu et al., 2004][Liu et al., 2005][Liu et al., 2006] the rate response curve was investigated in-depth. In this work a more comprehensive mathematical model for describing the cross-traffic effect on the rate response curve was presented. However, it has been shown by measurements that the fluid model for describing the rate response curve gives good estimates of the available bandwidth [Ekelin et al., 2006][Johnsson et al., 2004].

Sometimes the available bandwidth itself is not sufficient for an application. In [Jain and Dovrolis, 2005] a study of how to measure the variation range was presented. The variation range is defined using second order statistics such as the variance. That method was implemented in a tool called Pathvar.

Initial work on how broadband access links, such as 802.11 and ADSL, affect bandwidth measurements obtained from different measurement methods was presented in [Lakshminarayanan et al., 2004]. Here they showed how the probe-packet size affect the available bandwidth estimates obtained from a set of tools.

There exists much more work in the area of bandwidth estimation research due to many active researchers in recent years. However, the above literature survey, together with the short presentation of state-of-the-art tools, covers much of the important research made in this area.

4. My contributions in existing papers

This section presents my contributions of the completed papers that are intended to be part of my doctoral thesis. Each bullet has a reference to its main paper(s). A brief description of the papers themselves can be found in Section 5.

- Development of a packet-interaction framework. This framework describes, at the discrete packet level, how probe-packet trains and cross-traffic packets interact with each other when traversing a common network bottleneck link. [Paper A]

- Investigation on how the properties of wireless bottleneck links affect bandwidth estimates obtained from state-of-the-art methods. [Paper B]

- Development of an extended TOPP model in order to describe why bandwidth estimates obtained from methods that rely on self-induced congestion varies with probe-packet size and cross-traffic intensity in wireless networks. [Paper B]

- Evaluation of a new available bandwidth measurement method called BART in wired testbeds as well as on Internet paths. The BART method uses Kalman filtering in order to obtain accurate bandwidth estimates in real time. [Paper C]

- A simulation study on how probe packets affect the performance of TCP flows sharing the same bottleneck link has been made. In this work it was shown that the flight pattern of the probe packets as well as the amount of probe packets per time unit is crucial to the performance of the TCP flows. [Paper D]

- Implementation of two available bandwidth measurement methods. A version of the TOPP model, called DietTopp, has been implemented. Further, a version of BART has also been implemented. Both tools are written in C/C++ and runs under UNIX systems. [Paper B][Paper C]

5. Completed papers to be part of the doctoral thesis

This section presents a list of completed research papers that are intended to be included in the doctoral thesis. A short description of the contest of the papers is given along with comments on my participation.

Paper A

Andreas Johnsson, Bob Melander and Mats Björkman
On the Analysis of Packet-Train Probing Schemes
In proceedings to the International Conference on Communication in Computing, Special Session on Network Simulation and Performance Analysis, Las Vegas, USA, June 2004

This paper describes the interaction, at the discrete packet level, between probe-packet trains and cross-traffic packets when they traverse a common network bottleneck link. Within the paper three different packet-interaction patterns are defined. Their impact on the the rate response curve is discussed.

I have written most of the paper, performed analysis together with the co-authors, conducted all measurements and co-authored the measurement tool.

Paper B

Andreas Johnsson, Mats Björkman and Bob Melander
An Analysis of Active End-to-end Bandwidth Measurements in Wireless Networks
In proceedings to the 4th IEEE/IFIP End-to-end Monitoring Techniques and Services workshop, Vancouver, Canada, April 2006

In this paper the results from a study of how wireless bottlenecks affect bandwidth estimates are discussed. It is shown that the probe-packet size is crucial to the estimates of link capacity and available bandwidth. A smaller probe-packet size results in a lower estimate. Further, the estimated link capacity is dependent on the cross-traffic intensity. The variations of the link capacity and available bandwidth estimates above are not visible in wired networks. Hence, this is due to the properties of wireless link. In the paper the TOPP model is extended in order to explain why the probe-packet size and the cross-traffic intensity has impact on the bandwidth estimates, both link capacity and available bandwidth. However, recent results have shown that this extended model is applicable only in special cases.

I have written most of the paper, performed most of the analysis, constructed the testbed, conducted all measurements and implemented the measurement tool.

Paper C

Svante Ekelin, Martin Nilsson, Erik Hartikainen, Andreas Johnsson, Jan-Erik Mångs, Bob Melander and Mats Björkman **Real-time Measurement of End-to-End Available Bandwidth Using Kalman Filtering**
In proceedings to the 10th IEEE/IFIP Network Operations and Management Symposium, Vancouver, Canada, April 2006

In this paper a new measurement method called *Bandwidth Available in Real Time*, BART, is presented along with an evaluation of the method in testbed and Internet scenarios. BART relies on the TOPP model and deploys a Kalman filter to obtain estimates in real time. The evaluation of BART shows that the accuracy and speed is very good.

I have co-authored the part on evaluation of BART, I have participated in discussions concerning the evaluation process, co-designed test scenarios and participated in conducting measurements. I have also implemented the measurement method into a tool used in the evaluation process.

Paper D

Andreas Johnsson and Mats Björkman
Measuring the Impact of Active Probing on TCP
In proceedings to the International Symposium on Performance Evaluation of Computer and Telecommunication Systems, July 2006

This paper discusses the impact of active probing, such as available bandwidth measurements, on TCP. In the paper it is shown, by analyzing results from simulations in NS-2, that the decrease of TCP performance depends on how the probe packets are injected. The TCP performance decreases due to changed estimates of round-trip time and increased loss rate caused by the probe packets. However, the overall impact is rather low.

I have written most of the paper, co-analyzed the simulation results, constructed the simulation setup and performed all simulations.

6. Proposed work to be part of the doctoral thesis

To complete my doctoral thesis I propose a list of important research problems to study. The topics are related to each other and focus on BART measurements in 802.11 networks. They are summarized below:

- Is the proportional share assumption valid in 802.11 wireless networks as it is when sending packets through wired link with a FIFO queue? This was the assumption when extending the TOPP model in paper B, however recent results give hints that this is not the case. The bandwidth measurement methods described in Section 3 all assume that the probe packets get a proportional share of the bottleneck link when the probe packets cause congestion. What happens if the 802.11 MAC layer uses RTS/CTS?
- Tuning of BART parameters when performing measurements in 802.11 networks. The BART method uses a Kalman filter in its estimation process. In the Kalman filter there are parameters that can be tuned in order to optimize the performance. When the bottleneck is a wireless link the choice of parameter settings may differ compared to the parameters used in wired networks. Other BART parameter settings, such as the probe-packet train length and the probe-packet size, will also be investigated.
- Is it possible to estimate the link capacity of an 802.11 bottleneck link at the link layer? As previous research has shown the IP-layer link capacity varies with the probe-packet size and the cross-traffic intensity.

The findings made when studying these research problems will be published as Paper E (and possibly Paper F).

7. Additional publications

This section lists additional publications that are not intended to be part of the doctoral thesis.

Andreas Johnsson
How does Available Bandwidth Measurement Methods Affect TCP?
In Proceedings to the Swedish National Computer Networking Workshop, Halmstad, November 2005.

Andreas Johnsson, Bob Melander and Mats Björkman
Bandwidth Measurement in Wireless Networks
In Proceedings to the Fourth Annual Mediterranean Ad Hoc Networking Workshop, Porquerolles, France, June, 2005.

Andreas Johnsson
Bandwidth Measurements in Wired and Wireless Networks
Technology Licentiate thesis, Mälardalen University Press, April 2005.

Andreas Johnsson, Bob Melander and Mats Björkman
DietTopp: A First Implementation and Evaluation of a Simplified Bandwidth Measurement Method
In proceedings to the Second Swedish National Computer Networking Workshop, Karlstad, November 2004.

Andreas Johnsson, Mats Björkman and Bob Melander
A Study of Dispersion-based Measurement Methods in IEEE 802.11 Ad-hoc Networks
In proceedings to the International Conference on Communication in Computing, Special Session on Network Simulation and Performance Analysis, Las Vegas, USA, June 2004.

Mats Björkman, Andreas Johnsson and Bob Melander
Bandwidth Measurements from a Consumer Perspective - A Measurement Infrastructure in Sweden
Presented at the Bandwidth Estimation (BEst) workshop, San Diego USA, December 2003.

Andreas Johnsson
On the Comparison of Packet-Pair and Packet-Train Measurements
In proceedings to the First Swedish National Computer Networking Workshop, Arlandastad, September 2003.

8. Thesis outline

The doctoral thesis will be a collection of the five papers described in Section 5 and 6 (Paper A - E).

The doctoral thesis will also contain an introduction to the field of bandwidth measurement methods, a problem statement, related work and the main contributions of the thesis.

9. Time plan

The research that has to be completed before the defence of the doctoral thesis was described in Section 6. More precisely, a set of measurements has to be performed using BART in a wireless testbed (already existing). First, the proportional share assumption will be studied. Thereafter an analysis is made where the parameters of the Kalman filter in BART is tuned to optimize the estimation process with respect to accuracy, adaptivity and other aspects. It will also be studied whether it is possible to estimate the link capacity at the link layer. The time plan is shown in the activity table below.

Activity	Deadline
Ph.D. proposal defence	September 2006
Measurements completed for paper E	October 2006
Analysis of measurement results	November 2006
Completion and submission of paper E	December 2006
Completion of doctoral thesis	February 2007
Doctoral thesis defense	March 2007

In parallel to the research activities a set of graduate courses must be completed. Completed, ongoing and future graduate courses are shown in the course table below.

Course	Credits	Status
Distributed systems MN1	5p	Completed
Computer communications MN2	5p	Completed
Computer architecture MN2	5p	Completed
Secure computer systems	5p	Completed
Case-study methodology	3p	Completed
Reading course in computer communications	4p	Completed
Simulations using NS-2	3p	Completed
Markov processes	5p	Completed
Research methodology and planning	10p	Ongoing
Leadership and organization	5p	Sep - Oct 2006

References

Antoniades, Demetres, Athanatos, Manos, Padadogiannakis, Antonis, Markatos, Evangelos P., and Dovrolis, Constantine (2006). Available bandwidth measurement as simple as running wget. In *Passive and active measurement (PAM) workshop*, Adelaide, Australia.

Dovrolis, Constantinos, Ramanathan, Parameswaran, and Moore, David (2001). What do packet dispersion techniques measure? In *Proceedings of IEEE INFOCOM*, pages 905–914, Anchorage, AK, USA.

Ekelin, S., Nilsson, M., Hartikainen, E., Johnsson, A., Mångs, J., Melander, B., and Björkman, M. (2006). Real-time measurement of end-to-end available bandwidth using kalman filtering. In *Proceedings to the IEEE/IFIP Network Operations and Management Symposium*, Vancouver, Canada.

Hu, Ningning, Li, Li (Erran), Mao, Zhuoqing Morley, Steenkiste, Peter, and Wang, Jia (2004). Locating internet bottlenecks: algorithms, measurements, and implications. In *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*.

Hu, Ningning and Steenkiste, Peter (2003). Evaluation and characterization of available bandwidth probing techniques. *IEEE Journal on Selected Areas in Communication*.

Jain, Manish and Dovrolis, Constantinos (2002a). End-to-end available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput. In *Proceedings of ACM SIGCOMM*, Pittsburg, PA, USA.

Jain, Manish and Dovrolis, Constantinos (2002b). Pathload: a measurement tool for end-to-end available bandwidth. In *Passive and Active Measurement (PAM) Workshop 2002 Proceedings*, pages 14–25, Ft Collins, Co, USA.

Jain, Manish and Dovrolis, Constantinos (2004). Ten fallacies and pitfalls in end-to-end available bandwidth estimation. In *Internet Measurement Conference (IMC)*, Sicily Italy.

Jain, Manish and Dovrolis, Constantinos (2005). End-to-end estimation of the available bandwidth variation range. In *SIGMETRICS*, Canada.

Johnsson, Andreas, Melander, Bob, and Björkman, Mats (2004). Diettopp: A first implementation and evaluation of a new bandwidth measurement tool. In *Swedish National Computer Networking Workshop*.

Lakshminarayanan, Karthik, Padmanabhan, Venkata N., and padhye, Jitendra (2004). Bandwidth estimation in broadband access networks. In *In Proceedings to the Internet Measurement Conference*.

Liu, Xiliang, Ravindran, Kaliappa, and Loguinov, Dmitri (2004). Single-hop probing asymptotics in available bandwidth estimation: Sample-path analysis. In *Proceedings of ACM SIGCOMM Internet Measurement Conference*.

Liu, Xiliang, Ravindran, Kaliappa, and Loguinov, Dmitri (2005). Multi-hop probing asymptotics in available bandwidth estimation: Stochastic analysis. In *Proceedings of ACM SIGCOMM Internet Measurement Conference*.

Liu, Xiliang, Ravindran, Kaliappa, and Loguinov, Dmitri (2006). Measuring probing response curves over the ron testbed. In *Passive and Active Measurement (PAM) workshop*.

Melander, Bob, Björkman, Mats, and Gunningberg, Per (2000). A new end-to-end probing and analysis method for estimating bandwidth bottlenecks. In *Proceedings of IEEE Globcom*, San Francisco.

Melander, Bob, Björkman, Mats, and Gunningberg, Per (2002). Regression-based available bandwidth measurements. In *Proceedings of the 2002 International Symposium on Performance Evaluation of Computer and Telecommunications Systems*, San Diego, CA, USA.

Pásztor, Attila and Veitch, Darryl (2002). The packet size dependence of packet pair like methods. In *Tenth International Workshop on Quality of Service (IWQoS 2002)*, Miami Beach, USA.

Prasad, Ravi S., Dovrolis, Constantinos, and Mah, Bruce A. (2002). The effect of layer-2 store-and-forward devices on per-hop capacity estimation. In *IEEE Infocom*.

Prasad, R.S., Murray, M., Dovrolis, C., and Claffy, K. (2003). Bandwidth estimation: metrics, measurement techniques, and tools. IEEE Network Magazine.

Ribeiro, Riedi, Baraniuk, Navratil, and Cottrel (2003). pathchirp: Efficient available bandwidth estimation for network paths. In *Passive and Active Measurement Workshop*.

Strauss, Katabi, and Kaashoek (2003). A measurement study of available bandwidth estimation tools. In *ACM SIGCOMM Internet Measurement Workshop*.