

Performance monitoring of high-speed networks from the NREN perspective

Sven Ubik, Vladimír Smotlacha <ubik, vs@cesnet.cz>, CESNET, Prague, Czech Republic
 Nicholas Simar, <Nicolar.Simar@dante.org.uk>, DANTE, Cambridge, United Kingdom

Abstract—Performance monitoring evaluates network qualitative and quantitative characteristics in order to assist with network planning and network application performance tuning.

In this paper we summarize current state and propose future directions of performance monitoring from the NRENs (National Research and Educational Network) viewpoint. We describe performance monitoring requirements from the NREN viewpoint, what we already can monitor, what lessons we have learned, what are existing problems and what we need to do in near future. We will introduce Perfmonit, a multidomain network performance management infrastructure being developed by European NRENs.

Keywords: performance monitoring, end-to-end performance, bandwidth measurement

I. CURRENT NRENs AND THEIR EXPECTED DEVELOPMENT

Most backbone NREN circuits currently operate at speeds ranging from 1 Gb/s to 10 Gb/s. Typical long-term loads (daily averages) range from 50 Mb/s to 1 Gb/s. That is the networks are currently on the verge of overprovisioning, loosely defined as being loaded up to 10% of their installed bandwidth. This implies two things. First, we need to monitor network behaviour carefully and this monitoring must work at high speeds. Second, speed upgrade will be needed in near future even where we now have 10 Gb/s circuits.

II. PERFORMANCE PROBLEMS

We found that throughput between tuned PCs over Géant and connected NRENs is often lower than what we could expect from the installed bandwidth and current load. There are two likely causes: 1) short utilization peaks, which are not observable by conventional throughput measurements aggregating over longer periods of time, 2) intertwined congestion control details, especially in Linux [1], which make performance tuning difficult. Additionally, low-level problems such as faulty transceivers sometimes reduce throughput and are difficult to locate.

III. WHAT WE NEED TO MONITOR

We can identify *three types of users of performance monitoring*: 1) network administrators, who need to know development of aggregated traffic for network planning, 2) advanced application users, such as Grid people, who need to know how well they applications will run over the network, 3) PERT, who need to know detailed network behaviour in short timescales for performance debugging.

To satisfy needs of these three user groups, we need a *performance monitoring system* with the following properties:

- Monitor in order of importance: available bandwidth, ideally at network layer, packet loss rate, jitter and delay.
- Characteristics available in various timescales for each network link and network path (end-to-end), which implies inter-domain monitoring.
- Acceptable accuracy, timescales, aggregation and presentation should be further studied. The need for more detailed information increases in the order in which the three user groups were mentioned.

IV. WHAT WE CAN MONITOR AND LESSONS LEARNED

A. Active bandwidth measurement and estimation

Throughput achieved by *Bandwidth measurement* tools, such as iperf heavily depends on TCP buffer size, number of parallel streams, TCP implementation, dynamics and elasticity of existing traffic.

Available bandwidth estimation tools, such as pathload or ABWE, send just a few carefully scheduled packets and try to estimate the available bandwidth from the analysis of packet departure and arrival times.

We ran several tools together along two paths over Géant network, making one round of measurements and estimations per hour for the period of one month. Each path crossed more than 10 routers, all links were Gigabit Ethernet or OC-48. We will describe our observation in the full paper. Basically, ABWE accuracy and reliability was still insufficient to assist congestion control to set optimal sending rate and results from different tools generally did not match and it was not easy to conclude what was the real available bandwidth.

B. Bandwidth in small timescales

Link load can be reliably computed from 64-bit *byte counters* associated with a port on a router and accessible by SNMP. Achievable granularity depends on how often we read the counters. Reading just a few SNMP variables is a lightweight operation for the router and it can be done frequently, such as once per second.

When we sample byte counters at different rates, we find fluctuations of different magnitudes and frequencies. For example, statistical values of load on one of our backbone links are summarized in Tab. I. More frequent sampling revealed higher fluctuations. Unfortunately, we cannot take more than one sample per a few seconds. It takes some time for the router to propagate information about transferred packets from port adapters to the SNMP agent, which updates its MIB in some intervals. Short-term fluctuations are distorted by mea-

Interval	Avg Mb/s	Min Mb/s	Max Mb/s	Var (Mb/s) ²	Std dev Mb/s
60s	550	435	657	1474	38
30s	551	426	672	1753	41
10s	551	400	718	2693	51
5s	551	376	770	5736	75
1s	551	297	1562	47268	217

TABLE I

CHARACTERISTICS OF LOAD COMPUTED FROM SAMPLES TAKEN WITH DIFFERENT FREQUENCY

surement errors. We will illustrate this effect in the full paper.

C. Packet loss rate and delay

Active monitoring of packet loss rate is difficult because what we measure is packet loss rate of monitoring stream, which has little relation to packet loss rate of other traffic on the same network path, particularly when loss is low and measurement periods short [3].

Active monitoring of delay between two end stations is possible. However, we cannot monitor RTT to routers along a path by ICMP or UDP ping to assist congestion control. The point where delay fluctuates and losses occur is a likely bottleneck. However, as we will illustrate in the full paper, ICMP implementation in Cisco causes fluctuations that distort expected dependence of RTT on queue length.

Again, a reliable solution would be to use an independent monitoring device.

V. FUTURE OF PERFORMANCE MONITORING

There are two primary tasks to be completed:

- To develop an extensible inter-domain monitoring platform for end-to-end performance monitoring. Development of such platform has started within TF-NGN and will continue as part of the Géant2 project.
- To develop a programmable monitoring adapter needed for fine-grained monitoring at speeds higher than 1 Gb/s. The adapter of this kind is being developed as part of the SCAMPI project.

We will describe the performance monitoring system in this article, while the description of the programmable monitoring adapter is provided in [5].

VI. PERFMONIT - EUROPEAN MONITORING INFRASTRUCTURE

Initiated by several European NRENs, the performance monitoring activity started as part of Terena TF-NGN with the goal to build "perfmnit", a multi-domain network performance monitoring platform. The activity is now lead by Dante and will continue as a part of Géant2 project.

The infrastructure should allow to access information from a database of regular tests as well from tests started on demand with specific characteristics between certain measurement points under different administrative authorities.

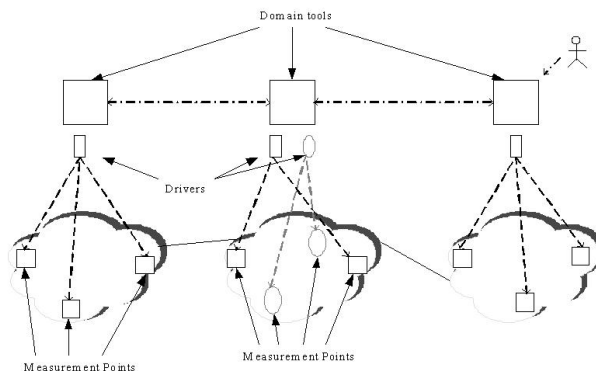


Fig. 1. Perfmnit architecture

Well-defined interfaces and modular structure should allow two things: extendable heterogeneous instrumentation (using different measurement devices, including RIPE TTM, SCAMPI, OWAMP devices, etc.) and multi-domain operation for end-to-end monitoring.

The overall perfmnit infrastructure is illustrated in Fig. 1. Each domain has one domain tool and several measurement points (MPs). Users or monitoring applications communicate their monitoring requests including required end points to the domain tool of their domain. If the request requires information from other domains, the domain tool contacts domain tools of other domains and properly concatenates obtained data. Each domain tool uses pathfinder to locate all measurement points that need to be asked. Communication with different measurement tools in measurement points takes place through drivers with well-defined interfaces. Information is either obtained from the repository or a new test is started on demand.

REFERENCES

- [1] Pavel Cimbal. *Incremental Development of Transmission Control Protocol and Related problems*, CESNET technical report, in preparation, 2003.
- [2] Sven Ubik, Pavel Cimbal. *Achieving Reliable High Performance in LFNs*, Terena Networking Conference, 19.-22. May 2003, Zagreb, Croatia.
- [3] Paul Barford, Joel Sommers. *A Comparison of Active and Passive Methods for Measuring Packet Loss*.
- [4] SCAMPI - A Scaleable Monitoring Infrastructure for the Internet, IST-2001-32404, <http://www.ist-scampi.org>.
- [5] Vladimir Smotlacha, Sven Ubik, Jiri Novotny. *High-speed programmable monitoring adapter*, submitted to TNC2004.

Biography

Sven Ubik received his MSc. and PhD. in computer science from the Czech Technical University in 1990 and 1998, respectively. He is currently with the Department of research and development of CESNET, operating the Czech NREN.

Vladimir Smotlacha received his MSc. degrees in Computer Science from the Czech Technical University in 1984 and in Theoretical Cybernetics from the Charles University of Prague in 1986. He is currently with the Department of research and development of CESNET.

Nicholas Simar is a project engineer in Dante, UK.