A Study on Network Performance Metrics and their Composition

Andreas Hanemann[1], Athanassios Liakopoulos[2], Maurizio Molina[3], D. Martin Swany[4]

[1] German Research Network, c/o Leibniz Supercomputing Center, Barer Str. 21, D-80333 Munich, Germany
[2] Greek Research & Technology Network S.A., 56 Mesogion Ave., 11574, Athens, Greece
[3] DANTE, 126-130 Hills Road, Cambridge CB2 1PG, United Kingdom
[4] Dep. of Computer and Information Sciences, University of Delaware, Newark, DE 19716, USA

Email: hanemann@dfn.de, aliako@grnet.gr, Maurizio.molina@dante.org.uk, swany@cis.udel.edu

Keywords: network performance, network monitoring, network metrics, multi-domain, post processing

Abstract: Research backbone networks like GÉANT2 and the National Research and Education Networks are used by a variety of scientists and research projects. These users and the network engineers operating the networks would like to get access to network performance metrics to optimise their use of the network and to troubleshoot performance degradations, when they happen. A variety of tools for performing network measurements already exist, and the perfSONAR architecture developed within the Joint Research Activity 1 (JRA1) of GÉANT2 aims at integrating them in a coherent framework. However, a harmonised definition of which metrics are mostly interesting and how measurements must be carried out is still lacking. In this paper we suggest the set of elementary metrics which are more relevant, along with indication about how to post process (or "transform", or "compose") them in order to obtain derived summary values that can quickly and intuitively give an indication of network performance. Methods to perform the composition are presented, together with constraints which have to be taken into account to get accurate results. In particular, delay measurements are the most delicate ones to compose. We carried out a series of experiments for proofing the validity of composition of delay metrics, and we briefly present some preliminary results.

1 Introduction

Many modern networking applications can benefit from improved Quality of Service (QoS) supported across multiple administrative domains. GÉANT2, the Gigabit core pan-European research network, for example, supports the Premium IP service to the European National Research & Education Networks (NRENs). Provisioning of end to end advanced transport services requires methods for verifying the established Service Level Agreements (SLAs) between the service provider and its customers. Even in a well engineered network, however, occasional equipment fault or misconfiguration can cause severe service performance degradation.
Therefore, GÉANT2 is committed to constantly assess the QoS in the network and verify that the performance guarantees agreed upon with the NRENs are met. Moreover, end users should be able to access the measurement infrastructure or the archived measurement data, even if with lower privileges than the GÉANT2/NRENs Network Engineers. This requires the deployment of an appropriate monitoring infrastructure in the GÉANT2/NRENs networks, and coordination in the performance metric collection and exchange.
The Joint Research Activity 1 (JRA1) [1] in the GÉANT2 project, in cooperation with the Internet2's End-to-End piPEs [2] initiative and the US Department of Energy's ESnet [3], defined a general framework for a multi-domain network measurement infrastructure. Currently, a prototype implementation, called "Performance focused Service Oriented Network monitoring Architecture" (perfSONAR), is under development and testing. In this context, it is fundamental to harmonise the type of collected measurements so that they are useful also in a multi-domain context, and to define common procedures to post-process (or "compose") them.
This paper presents the work in progress in GN2-JRA1 related with network performance metric composition. In section 2, a brief description of the JRA1 monitoring architecture (perfSONAR) is provided. Section 3 presents our selection of the most significant metrics and their classification into categories. Section 4, explains the main reasons for post processing monitoring data. This operation is called "metric composition". Section 5 presents, as an example, experimental results of a composition of One Way Delay data. Finally, section 6 references related work on metric composition and our conclusions as well as future plans are discussed in the last section.

2 GN2-JRA1 Monitoring Architecture (perfSONAR)

The perfSONAR system is a framework that enables network performance information to be gathered and exchanged in a multi-domain, federated manner. The goal of perfSONAR is to enable ubiquitous gathering and sharing of this performance information in order to ease management of advanced networks, facilitate cross-domain troubleshooting and to allow next-generation applications to tailor their execution to the state of the network. This system has been designed to accommodate easy extensibility for new network metrics and to facilitate the automatic processing of these metrics as much as possible.

The perfSONAR architecture is composed of three different layers, as shown in Figure 1. The Measurement Point Layer is responsible for performing active or passive measurement tests via multiple Measurement Points (MPs), i.e. existing network monitoring tools. The MP is wrapped into a higher level abstraction called Measurement Point Service, belonging to the Service Layer, which hides the implementation details of the MP. The Service Layer is composed of multiple services that control the monitoring infrastructure, receive, store and exchange measurement and network topology data. Services interact with each other without human intervention (e.g. measurement data retrieved by a Measurement Point Service is fed into a Measurement Archive Service and manipulated by a Transformation Service) and with the upper User Interface Layer. The end users interact via the visualisation tools at the User Interface Layer only.

The whole architecture is based on Web Services (WS) technology, which allows defining the interaction between services through well defined, language independent interfaces. Web Services are closely tied to the eXtensible Markup Language (XML). perfSONAR uses and extends a schema defined by the Global Grid Forum's Network Measurement Working Group [9]. This schema defines an extensible message and storage format for network measurements. The perfSONAR approach removes any dependencies from the lower networking technologies and permits new services to be easily added. The following services have been defined in the perfSONAR framework:

- Measurement Point (MP) service: performs the measurements and forwards data to other services
- Measurement Archive (MA) service: stores the measurement data
- Lookup service (LS): registers information regarding active services and their capabilities
- Topology service (TS): stores network topology information
- Authentication service (AS): provides authentication and authorisation services required in users - services interactions
- Transformation service (TrS): performs manipulation (aggregation, statistics) on available data sets
- Resource Protector (RP) service: arbitrates the use of limited measurement resources

Currently, perfSONAR is focusing on IP level metrics, as this is the main service provided by the NRENs. The framework has been build flexible enough to cater for new metrics and for different type of technologies.

In order for the described architecture to be truly useful in a multi-domain environment, there is the need to harmonise the type of collected measurements and the procedures for their composition in the TrS. This is the main focus of the study described in this paper. A more extended description is available in [5].

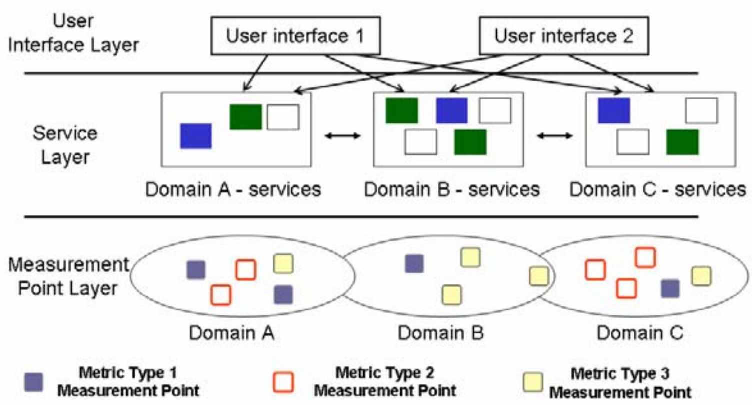


Figure 1 – The PerfSONAR Service Oriented Architecture for multi-domain network monitoring

3 Network Metric Selection and Classification

We surveyed several network metrics, both the ones defined in standards ([6], [11]) and non-standard ones which are commonly collected by network operation centres. Furthermore, we analysed the replies to a questionnaire circulated by the NRENs among potential users of perfSONAR in the first phase of the JRA1 project. As a result, we selected the metrics of greatest relevance for network performance, i.e. useful for assessing the service level offered to IP traffic forwarded though a network. They can be divided into four main groups:

- availability
- loss & error
- delay
- bandwidth

Availability metrics assess how robust the network is, i.e. the percentage of time the network is running without any problem impacting the availability of services. It can also be referred to specific network elements (e.g. a link or a node), and in that case it will measure the percentage of time they are running without failure. Loss and error metrics are indicative of the network congestion conditions and/or transmission errors and/or equipment malfunctioning. They usually measure the fraction of packets lost in a network due to buffer overflows or other reasons, or the fraction of errored bits or packets. Delay metrics also assess the network congestion conditions or effect of routing changes. They measure the delay (One Way Delay-OWD and Round Trip Time-RTT) and Delay Variation (IPDV, or "jitter") of the packets transferred by a network. Finally, bandwidth metrics assess the amount of data that a user can transfer through the network in a time unit, both dependent and independent from the existing network traffic.

Besides of the performance-related metrics, several additional metrics are often useful to explain the causes of performance degradations. Examples are the CPU load, memory consumption, or even chassis temperature of network devices. The monitoring infrastructure may observe these additional metrics to ease troubleshooting when their values indicate degradation in service levels, or to prevent degradation by upgrading equipments before they reach critical conditions. These metrics are further divided to device specific, flow monitoring and routing metric groups.

For each metric relevant in the context of perfSONAR, in [5] a definition was given (referencing standards when possible), and a procedure for its measurement was described, along with accuracy considerations. This effort tried to reconcile the variety of metric definitions and measurement methods that are often possible.

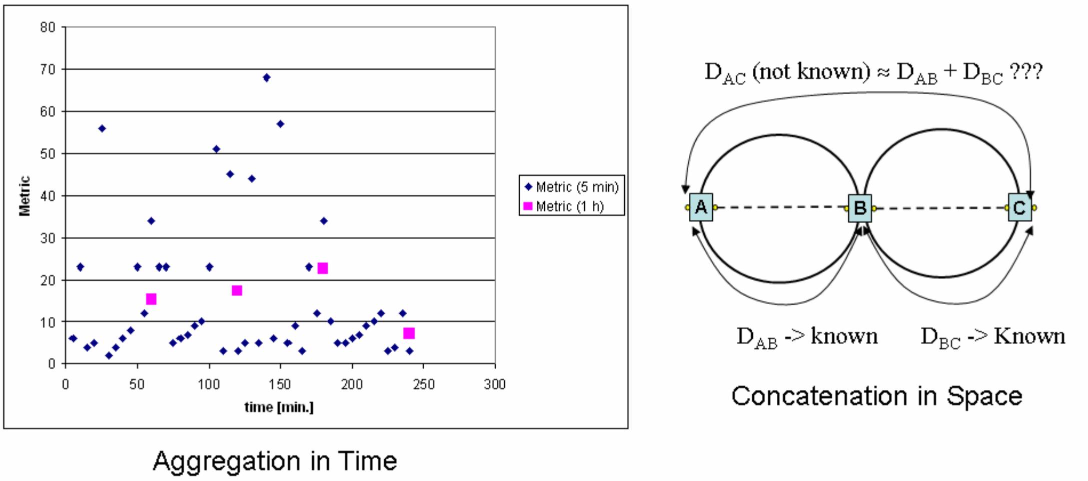
4 Network Metric Composition (Transformation Service)

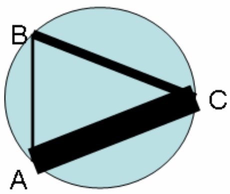
Setting a common understanding of network metric definitions, their measurement methodologies and their accuracy is, unfortunately, not enough. In general, network measurements need to be post-processed (composed) to be useful for the several tasks of network engineering, management and planning. This becomes fundamental in a multi-domain environment such as the one targeted by perfSONAR.

There are several reasons for composing network metrics: The first one is data reduction. Consider for example a network domain in which delay measurements are performed on all links. A network manager might ask whether there is a general problem with the network delay. Therefore, it would be desirable to obtain a single summary value calculated from the delay measurements on single domain's links. We call this composition "aggregation in space". In the example in Figure 2, a weight proportional to the carried traffic on the links is applied, to produce a summary OWD value for this single domain. Other rules to produce a summary value may be used (e.g. the maximum of the average OWD on single links, without any weight), depending on the foreseen usage.

Another important reason for composing network metrics is to perform trend analysis. For doing so, a single value for an hour, a day or, a month is computed from the basic measurements which are scheduled with finer granularity, e.g. every five minutes (see Figure 2). In this way, trends can be more easily witnessed, like an increasing usage of a backbone link which might require the installation of alternative links or the rerouting of some network flows. This type of composition is called "aggregation in time". This method reduces the amount of monitoring data at the expense of data resolution. Aggregation in time is widely used by visualisation tools, such as MRTG [10], that present various network performance parameters in different time scales.

Finally, composition may be performed for scalability. Due to the number of network elements in large networks like GÉANT and the connected NRENs, it is impossible to perform a full mesh of measurements between all the equipment, neither regularly nor on demand. However, if regular measurements are scheduled between selected Measurement Point pairs, say A to B and B to C, we can try to infer the value of a network performance metric (e.g. the OWD) on a path, say A to C, even in absence of a direct measurement among the end points of that path. This type of composition is called "concatenation in space" (see Figure 2).





	Delay	Load
A-B	24.5 ms	1 Gbit/s
B-C	7.8 ms	3 Gbit/s
A-C	4 ms	9 Gbit/s
Domain	1/13*24.5+3/13*7.8+9/13*4 = 6.4 ms	13 Gbit/s

Aggregation in Space

Figure 2 – The three possible metric composition types

For each selected network metric, we examined which composition operation can reasonably be applied and for what purpose, which statistical operations are more useful (e.g. average, max, min, X-percentiles, median) and when it is useful to perform in sequence two or more compositions. A summary of the results is shown in Table 1, and some details are explained in the following. The full work is contained in [5].

The lower part of the table shows which composition operations are useful with respect to metrics, and the specific statistical operation done during the composition. For example, it is reasonable to aggregate OWD measurements in time by computing a value for a longer time interval taking the average of measured values. Always regarding aggregation in time of OWD, it is better to use a 97.5% - 2.5% percentile aggregation to avoid extreme values resulting from measurement inaccuracies, instead of simply calculating a maximum/minimum value.

The remark “ToS effect” for aggregation in space of packet loss measurements means that it can be interesting to track whether the prioritisation of packets in router waiting queues leads to different packet loss for different ToS values.

The concatenation in space for available bandwidth and capacity is only reasonable for minimum, because the minimum bandwidth of a link is the bottleneck when transferring data through a concatenation of links.

The achievable bandwidth on a path requires sending a lot of test packets which will likely disturb other traffic. Therefore, these measurements will be carried out in specific situations only so that no data for composition will be available (no longer periods of time, only along the specific path). Moreover, since the end to end Round Trip Time plays a role in these kinds of tests, concatenating results on consecutive path portions makes little sense.

For the concatenation in space of OWD, RTT, and IPDV special constraints arise so that an experimental analysis has been conducted. It is described in the next section.

Table 1 – Summary of the composition study

	Aggregation in time	Aggregation in space	Concatenation in space
Definition	Aggregate measurements of the same scope and type performed in different time windows or time instants.	Aggregate measurements of the same type but of different (physical or logical) scope.	Concatenate measurements of the same type performed on consecutive paths
Usability	Reduce the amount of collected data, observe trends.	Provide a summary metric value for a group of network elements or links in a domain.	Combine the results from multiple measurements in order to estimate the e2e performances for a longer path.
Requirements	NA	NA	Measurements should be taken in consecutive links.
	NA	Measurements should be performed in the same timeframe	
	Measurements should be performed with the same type-packets, e.g. size, ToS, etc. (For space aggregation, this applies to *physical* space aggregation only. *Locical* space aggregation is by definition over packet properties!)		
	Measurements should to be collected during all the time widow. Otherwise, measurements have to be weighted.	Measurements should be weighted according to the link characteristics (e.g. capacity, utilisation) and their significance	
	NA	Measurements should have comparable accuracy.	
	Operations should be performed over an adequate data set.		
Most relevant Operations			
OWD, RTT	Average, percentiles	Average, maximum, percentiles	Average Percentile (but difficult to compute exacty)
IPDV	Average, percentiles	Average, maximum, percentiles	Average Percentile (but difficult to compute exacty)
Packet Loss	Average, median, percentiles	Average minimum,maximum (ToS effect)	Average
Available Bandwidth	Average, minimum, maximum	Average, minimum, maximum, percentiles	Minimum
Utilisation	Average, median	Average, minimum, maximum, percentiles	NA
Capacity	NA (capacity is a slowly varying "metric")	Average, minimum, maximum	Minimum
Achievable bandwidth	NA (Not likely that tests are performed regularly)	NA	NA
Availability	Average	Average	Average

5 Experimental Assessment of Concatenation in Space

It is simple to apply concatenation in space operation using mean OWD values. For example, the mean OWD value along the path from host[1] A to host C via host B ($<OWD_{AC}>$), knowing the corresponding mean OWD values along the path from host A to host B ($<OWD_{AB}>$) and B to C ($<OWD_{BC}$,>), is given by the following (intuitive) formula:

$$<OWD_{AC}> = <OWD_{AB}> + <OWD_{BC}>$$

However, the mean OWD value is not a sufficient metric to assess the performance along a path, especially when delay-sensitive applications are deployed over the network. In such cases, it is highly recommended to

[1] Measurement points are typically located "close" to the host/router (e.g. connected through a high speed LAN switch) so that the time the measurement packet is received and timestamped by the host is a good approximation of the time it transits through a host's interface.

estimate a high quantile[2] for the OWD along a path. This introduces challenges from a statistical point of view as the quantile of OWD along the path AC can be computed only knowing the corresponding OWD distribution along the path AC. If the latter is unknown, the only possibility is to compute it from the distributions of OWD_{AB} and OWD_{BC} by performing the convolution of these two distributions. The convolution [8] of two distributions f and g is a mathematical operation involving an integration and is denoted as "f*g" or "f⊗g", or "f(x)g". It is important to note that we assume that the OWD variables along the two path portions, i.e. from A to B and from B to C, are independent.

To prove the applicability of the convolution approach to infer the high quantiles of OWD_{AC} given two independent OWD_{AB} OWD_{BC} datasets, we collected OWD measurement data among three different IPPM measurement boxes [7]; in Erlangen (Germany), Frankfurt (Germany) and Rome (Italy). We collected measurements on both A to B, B to C and A to C, to be able to compare the values given by composition formulas with a direct measurement. In our experiments, we were sure that the routing of packets from A to C was through B, which is of course a requirement for such an analysis to make sense. Each IPPM measurement box transmitted 5 packets per second. We observed that the OWD distributions (of both AB, and BC and AC) are very concentrated around the typical, minimum value for each path. As shown in Figure 3, most of the singleton delays along the path A-B-C (Erlangen-Frankfurt-Rome) are around 15 ms, as "represented" by the thick line at the bottom of the graph. However, there are quite regularly high values, appearing throughout the 12-hour measuring period. These high values represent less that 0.6 % of the total collected values, but they are exactly what needs to be kept under control to ensure the good functioning of the network.

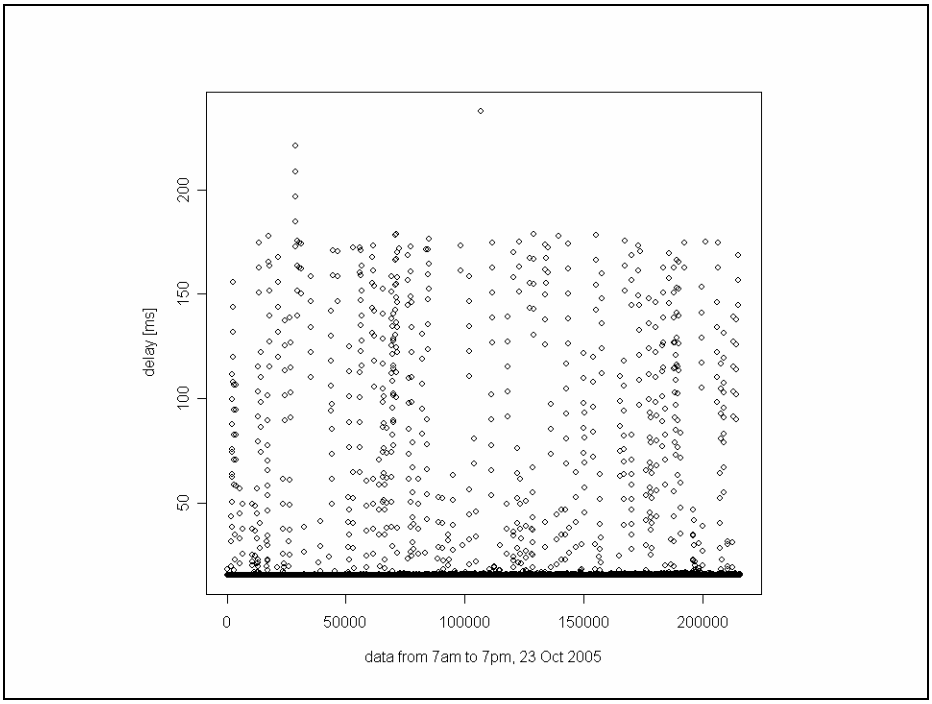


Figure 3 – Snapshot of 12-hour measurement of OWD along the path "Erlangen to Rome via Frankfurt".

In Figure 4, we use a *Quantile-Quantile* plot to assess the difference among the convolution of distributions AB and BC with the "real" distribution of AC. If the convolution AB⊗BC was exactly like the distribution of AC, all the points would marked on the diagonal .We used the following range of quantile values: 0.1, 0.2, …, 0.9, 0.91, 0.92, …, 0.99, 0.991, 0.992, …0,999. The bottom-right graph in Figure 4 zooms into the quantiles up to 0.993, which appear concentrated on a single point on the upper-left graph. We can note that for the highest quantiles (0.994 and over) there is a significant drift. This corresponds to the points in the graph in Figure 3 that have a value above 20ms (thus is the upper "sparse" region of the graph). Although the drift is significant, we note first that it does not tend to diverge as the quantile approaches 1 (on the contrary, it tends to get closer again to the diagonal), and second that the prediction given by AB⊗BC is below the diagonal (i.e. the prediction of the OWD is higher), and thus can be used as a conservative value for an SLA. For the lowest quantiles (up to 0.993 - see bottom part of Figure 4) we see that the prediction is good, as it remains within 1.3% of the real value (consider that the scale of the graph is enlarged).

[2] A q-quantile of a random variable X is any value x such that $Pr(X \le x) = q$. While quantiles can take whatever value (e.g. 0.999) percentiles are specializations of quantiles constrained to take values with only two significant decimal digits, (e.g. 0.90, or 0.95). The terminology percentile exists for historical reasons.

The presented results are preliminary, but to our knowledge a similar validation test for the convolution approach has not been performed yet.
Further analysis is needed to confirm the obtained results, and understand the reason of the big differenced for the highest quantiles in the upper graph (it may be due e.g. to the fact that the delays on AB and BC are not temporally independent) as well as the reason of the small, fixed offset that appears in the lower graph.

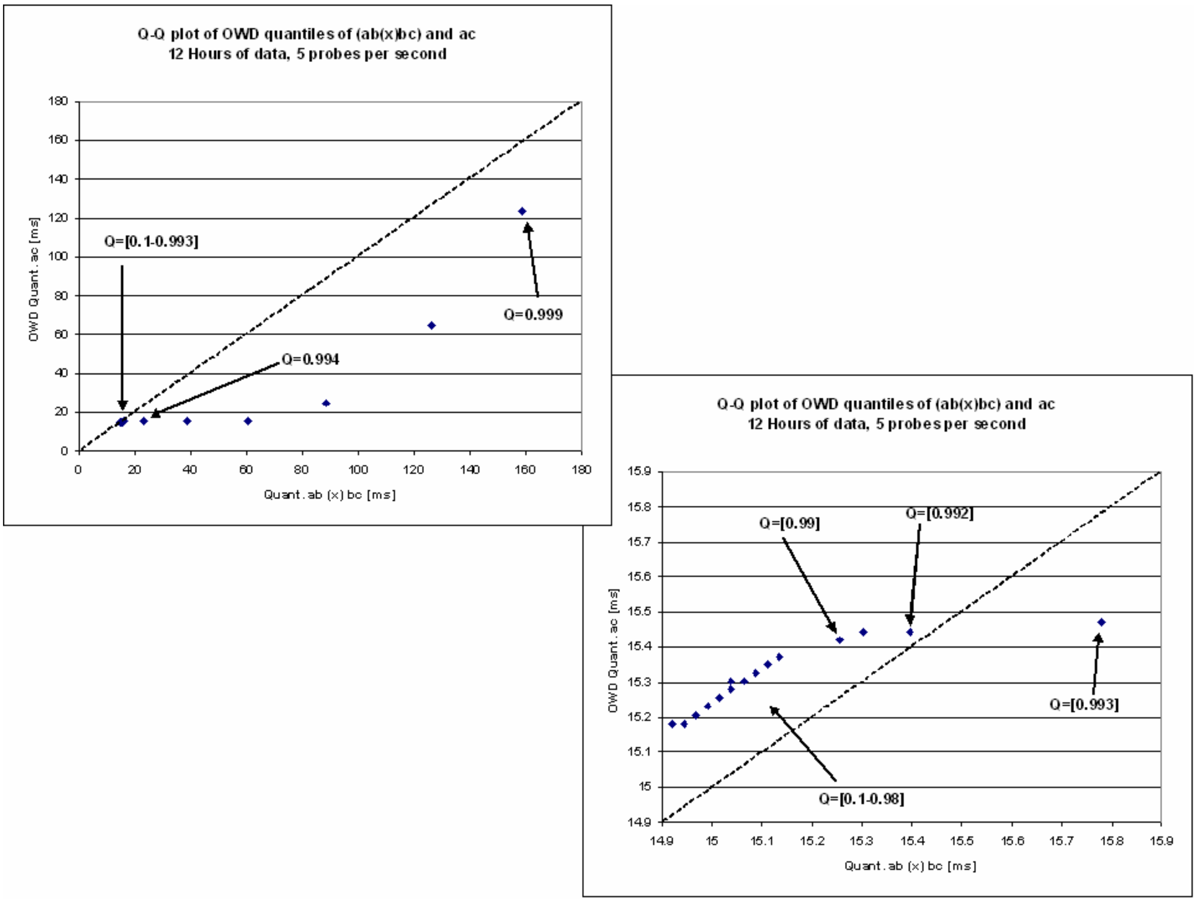


Figure 4 – Comparison of the OWD quantiles of a direct measure from A to C and of the quantiles of the convolution of AB (x) BC. The lower graph is a zoom of the initial part of the upper one

6 Related work

In the standardization arena, both the ITU-T and the IETF produced several recommendations (respectively, RFCs) about performance metrics for IP networks. The more relevant ITU-T recommendations are Y.1540 [11], defining the performance metrics, and Y.1541 [12], defining 6 different classes of service, and specifying the performance bounds for network delays, losses and errors that define these classes. The IETF IPPM WG published several RFCs about performance metrics for IP networks as well ([6]), but without defining any service classes on the basis of their values. The IETF also specifies in RFC 3763 **[19**] requirements for a One-Way Active Measurement Protocol (OWAMP), similar to the IPPM protocol [7] used in this work.
The issue of metric composition has been addressed by the ITU-T in an amendment to Y.1541 [13], limited to the case of concatenation in space of losses, errors and delays. In a forthcoming revision, [14] some approximate formulas (not based on the full convolution) to compose OWD quantiles are proposed, but these formulas are based on a modelling approach, without the support of real data observations.
In the IETF, no significant work in the metric composition has been performed yet, but recently the group decided to undertake this activity, and draft documents for time and space composition are planned by the end of 2006 in the IPPM Working Group [6]. A Framework document [15] has already been published, containing some of the concepts developed by JRA1 in [5].
The perfSONAR project is related to, and influenced by a variety of other network monitoring efforts. The Network Weather Service [16] uses statistical forecasting techniques to predict network performance from a time-series of network measurements. Internet2's piPEs infrastructure [2] includes an implementation of the OWAMP protocol as well as support for the scheduling of recurring active measurements of available bandwidth. The Intermon [17] project focuses on many aspects of inter-domain network monitoring and analysis

with a particular emphasis on Quality of Service (QoS) issues. MonALISA [18] is a distributed network montoring system with a wide deployment in Grid computing systems.

7 Conclusion and Future Work

The JRA1 [1] in the GÉANT2 project, in cooperation with the Internet2 [2] and ESnet [3], developed perfSONAR [4], a Service Oriented monitoring architecture for retrieving, storing, processing and presenting network performance metric measurement data in a multi-domain network environment. The first development phase for perfSONAR prototype has been completed and a few demonstrative services are already available for testing. We presented in this paper some results of a study for harmonising the choice of network performance metrics in a multi-domain environment, and for defining common procedures for metric post-processing (or composition). In particular, we classified the possible compositions in three main categories: aggregation in time, aggregation in space, concatenation in space, and explained the utility and challenges associated with each of them. For concatenation in space, we also presented the analysis of some One Way Delay data we collected, the analysis being finalised to validate a procedure for getting a picture of the performances on an end-to-end path given the availability of performance measurements only on disjoint sections of the path. This operation has an important practical utility, since in large network it is unpractical to setup dedicated measurements among each possible couple of end points of interest.
Future work will extend this experimental validation of concatenation in space, possibly applying it also to the IPDV (jitter) metric. Moreover, we wish to address the detailed study of the error propagation when compositions are applied and the experimental verification in real networks of all the proposed composition strategies.

Acknowledgements

We would like to thank Stephan Kraft and Roland Karch for collecting the OWD measurement data, and all the contributors to GN2 deliverable DJ1.2.3

References

[1] "GÉANT2 Joint Research Activity 1 (GN2-JRA1): Performance Measurements and Monitoring", http://www.geant2.net/server/show/nav.754.
[2] E2E piPEs: End-to-End Performance Improvement Performance Environment System http://e2epi.internet2.edu/e2epipes/
[3] US Department of Energy - ESnet http://www.es.net/
[4] perfSONAR: Performance focused Service Oriented Network monitoring Architecture, http://www.perfsonar.net
[5] DJ1.2.3 Network Metric Report, GN2-JRA1 deliverable, Febuary 2006.
[6]RFCs of the IETF IPPM Working Group: http://www.ietf.org/html.charters/ippm-charter.html
[7] IPPM Measurement tool: http://www-win.rrze.uni-erlangen.de/ippm/messprogramm.html
[8] Convolution operation, http://mathworld.wolfram.com/Convolution.html.
[9] Global Grid Forum's Network Measurement Working Group: https://forge.gridforum.org/projects/nm-wg
[10] Multi Router Traffic Grapher (MRTG): http://mrtg.hdl.com/mrtg.html
[11] ITU-T rec. Y.1540 - Internet protocol data communication service – IP packet transfer and availability performance parameters – 2002
[12] ITU-T rec. Y.1541 - Network performance objectives for IP-based services - 2002
[13] Y.1541 (2002) Amendment 2 New Appendix XI – Concatenating QoS values
[14] A. Morton et al - Revised Version of Rec. Y.1541, Network Performance Objectives for IP-based Services – work in progress
[15] A. Morton, S. Van den Berghe - Framework for Metric Composition: http://www.ietf.org/internet-drafts/draft-ietf-ippm-framework-compagg-00.txt - work in progress
[16] R. Wolski, L. Miller, G. Obertelli, and M. Swany, Performance Information Services for Computational Grids, In Resource Management for Grid Computing, Nabrzyski, J., Schopf, J., and Weglarz, J., editors, Kluwer Publishers, Fall, 2003.
[17] INTERMON - Advanced architecture for INTER-domain quality of service MONitoring, modelling and visualisation, http://www.ist-intermon.org
[18] MonALISA – MONitoring Agents using a Large Integrated Services Architecture, http://monalisa.cacr.caltech.edu
[19] S. Shalunov and B. Teitelbaum, One-way Active Measurement Protocol (OWAMP) Requirements, RFC 3763, http://www.ietf.org/rfc/rfc3763.txt

Vitae

Andreas Hanemann received a diploma degree (M.Sc.) in Computer Science from the University of Karlsruhe (TH), Germany. He is involved in the JRA1 performance visualisation starting from the project launch in 2004. Since October 2002 he is responsible for the Customer Network Management project at the Leibniz Supercomputing Center, which provides a network performance visualisation tool for the German Research Network.

Athanassios Liakopoulos received the Dipl.-Ing. degree in Electrical and Computer Engineering from the Nat. Technical Univ. of Athens (NTUA) in 1996, MSc with Distinction in Telematics from the Electrical Engineering Dept. in Univ. of Surrey (UniS) in 1998 and PhD in Electrical and Computer Engineering from the Nat. Technical Univ. of Athens (NTUA) in 2005. Since 2000, he joined the GRNET S.A. and participated in national and European research projects, such as SEQUIN, 6NET, SEEREN, GN2. He has awarded for his performance during his academic studies and has published articles in recognized technical journals.

Maurizio Molina graduated in Electronic Engineering (Italian Laurea) from the Polytechnic of Turin in 1993. Since then he has worked in the telecommunications industry, including Telecom Italia Labs (Turin, Italy) and the NEC Network Laboratories (Heidelberg, Germany). He published several papers about IP and ATM traffic modelling and network measurements. He contributed to IETF, ITU-T and 3GPP standards. He joined DANTE's Systems group in November 2004, working on performance monitoring, security and authentication.

Martin Swany is an Assistant Professor in the Department of Computer and Information Sciences at the University of Delaware. He completed his Ph.D. at the University of California, Santa Barbara in 2003. He is a 2004 recipient of the US Department of Energy Early Career Principal Investigator award and a 2005 Internet2 Faculty Fellow. His research interests include high-performance networking and distributed computing.