

## **perfSONAR: Performance Monitoring in europäischen Forschungsnetzen**

Andreas Hanemann<sup>#</sup>, Stephan Kraft<sup>\*</sup>, Patricia Marcu<sup>+</sup>, Jochen Reinwand<sup>\*</sup>, Helmut Reiser<sup>~</sup>, David Schmitz<sup>+</sup>, Verena Venus<sup>\*</sup>

<sup>+</sup> DFN-Verein, c/o Leibniz-Rechenzentrum

<sup>\*</sup> Friedrich-Alexander Universität Erlangen-Nürnberg

<sup>#</sup> DFN-Verein

<sup>~</sup> Leibniz-Rechenzentrum; Munich Network Management Team

**Zusammenfassung:** Zur QoS- und SLA-Überwachung sowie zur Vereinfachung der domänenübergreifenden Fehlersuche der Dienste im europäischen Netzwerkverbund GN2 wurde eine föderierte Multi-Domain Monitoring Architektur namens perfSONAR entwickelt. perfSONAR definiert geeignete Metriken und Messverfahren zum Performance Monitoring. Es realisiert eine Architektur, um verschiedenste unabhängig entwickelte Sensoren und Messverfahren in eine übergreifende und umfassende Architektur zu integrieren, einheitliche Messungen durchzuführen, zu analysieren und zu visualisieren. Von der Qualitätssicherung profitieren dabei sowohl große internationale Projekte sowie individuelle Nutzer der europäischen Wissenschaftsnetze.

### **1 Einführung**

Der über GEANT2 realisierte europäische Verbund von 30 nationalen Forschungsnetzen (NRENs) stellt Netzdienste im paneuropäischen Maßstab zur Verfügung. Ein Beispiel für einen solchen Dienst sind Optical Private Networks (OPNs), die z.B. im Umfeld des LHC-Grid verwendet werden, um das CERN mit den Tier1- und Tier2-Zentren zu verbinden. Die Dienste werden in einer multinationalen Kooperation unabhängiger NRENs und DANTE erbracht.

Für die Überwachung der Dienstqualität ist ein Grenz-, Domänen- und Provider-übergreifendes Performance Monitoringsystem unabdingbar. Eine solche föderierte Multi-Domain Monitoring Architektur wurde im Rahmen des GN2-Projektes JRA1 mit perfSONAR konzipiert und realisiert. Diese wird mittlerweile im europäischen Maßstab produktiv eingesetzt.

Im folgenden Abschnitt werden die wichtigsten von perfSONAR unterstützten Metriken und Messverfahren vorgestellt. Das in internationaler Kooperation entwickelte perfSONAR Framework wird in Abschnitt 3 beschrieben. Die an verschiedene Nutzergruppen angepassten Visualisierungstools, mit denen eine Analyse und Darstellung der Messergebnisse ermöglicht wird, fasst Abschnitt 4 zusammen. Der Artikel schließt mit einer Zusammenfassung und einem Ausblick, der insbesondere auf den Regelbetrieb von perfSONAR abzielt.

## 2 Performance Monitoring

perfSONAR bedient sich einer Reihe von Messeinrichtungen, die Netzwerkkenngrößen ermitteln. Einen Teil der möglichen Metriken – insbesondere die, die sowohl im X-WiN als auch im GEANT2-Projekt gemessen werden – sind Inhalt des folgenden Abschnitts. Die Internet Engineering Task Force (IETF) hat mehrere dieser Dienstgüte-Parameter als IP Performance Metrics (IPPM) standardisiert.

### 2.1 Metriken

Bei IPPM handelt es sich um ein umfassendes Rahmenwerk, das Definitionen von Metriken zur Messung der Internet-Performance beinhaltet [RFC 2330]. Prinzipiell verfolgt IPPM das Ziel, Messverfahren der Dienstgüte und deren Auswertung zu standardisieren und dadurch die Aussagekraft der erzielten Ergebnisse zu erhöhen. Die verschiedenen Metriken werfen dabei ihr Hauptaugenmerk auf die unterschiedlichen Transportprotokolle des Internets. Die wichtigsten der in perfSONAR verwendeten und von der IPPM-Arbeitsgruppe definierten Metriken werden im folgenden kurz zusammengefasst. Eine ausführliche Beschreibung der IPPM-Metriken sowie deren Ermittlung findet sich in [Holl2007].

Unter dem *One-Way Packet Delay* (OWD) [RFC 2679] versteht man diejenige Zeitverzögerung, die einem Paket auf dem Weg von einem Quell- zum Zielrechner widerfährt. Dabei gilt als (optimaler) Startzeitpunkt der Messung der Zeitpunkt, zu dem der Quellrechner das erste Bit des Pakets überträgt und als Endzeitpunkt derjenige, zu dem der Zielrechner das letzte Bit des Pakets empfängt. Der OWD im Bezug auf ein einzelnes Paket ergibt sich schließlich als Differenz von Start- und Endzeitpunkt.

*One-Way Packet Loss* (OWPL) [RFC 2680] gibt an, ob ein bestimmtes von einem Quell- an einen Zielrechner gesendetes Paket dort tatsächlich angekommen oder auf dem Weg zu diesem verloren gegangen ist. Wendet man diese Metrik auf eine Menge nacheinander versendeter Pakete an, lässt sich schließlich der Anteil der auf dem Weg von der Quelle zum Ziel verloren gegangenen Pakete bestimmen.

*IP Packet Delay Variation* (IPDV oder OWDV, One Way Delay Variation) [RFC 3393] bezieht sich jeweils auf zwei Pakete aus einer Menge von sequentiell versendeten Paketen. IPDV ist dabei definiert als die Differenz der Verzögerungen (OWD), die den beiden Paketen auf dem Weg vom Quell- zum Zielrechner widerfahren.

*Utilization* (Auslastung) bezeichnet das tatsächlich auf einem Pfad aufgelaufene Datenvolumen. Hierbei wird die übertragene Datenmenge pro Zeit zwischen zwei Endpunkten ermittelt und über einen Zeitraum dargestellt.

Neben diesen Metriken werden in perfSONAR auch andere ermittelt (passive Performance Daten, Interface Errors, Interface Drops), die in diesem Beitrag nur erwähnt werden sollen, im einem Projekt-Bericht aber genau beschreiben sind [hlms06].

## 2.2 Messmethoden

In Abhängigkeit von der Metrik erfolgt deren Erfassung über drei unterschiedliche Systeme: Hades, BWCTL und SNMP-Polling. Die Metriken OWD, OWPL und IPDV werden mit dem im DFN-Labor entwickelten Hades-Messsystem (Hades Active Delay Evaluation System [DFN-2]) ermittelt. Zur Ermittlung der verfügbaren Bandbreite (Available Bandwidth, AB) wird BWCTL [BWCTL] (Bandwidth Controller) benutzt, ein Wrapper für iperf [iperf], entwickelt von Internet2. Das SNMP-Polling wird zurzeit vom CNM [CNM] (Customer Network Management), das am LRZ München entwickelt wurde, um Auslastungsdaten in X-WIN zu erfassen und zu visualisieren, eingesetzt (s. Abschnitt 4).

Die *Hades*-Messungen werden durch Messstationen ausgeführt, die aktiv Testpakete generieren und diese in das Netzwerk einspeisen und empfangen. Die Messstationen bestehen in der Regel aus einem PC mit Linux-Betriebssystem. Zwischen allen Standorten können Messungen durchgeführt werden. Für die notwendige Uhrensynchronisation sind die Messrechner mit einer GPS-Karte ausgerüstet. Dadurch ist eine Genauigkeit von < 10 Mikrosekunden möglich [DFN-2]. Aus der Ermittlung der One-Way Delays werden die abgeleiteten Größen OWPL und IPDV bestimmt.

Die gleichen Messboxen, die zur Ermittlung der Delay-Daten herangezogen werden, dienen über ein zweites Netzwerkinterface zur Messung der verfügbaren Bandbreite. Hierbei wird zwischen zwei Endpunkten mittels BWCTL ein Verkehrsstrom generiert, der auf dem zu messenden Pfad ein möglichst großes Datenvolumen pro Zeit versendet. Im Normalfall werden hier TCP-Ströme benutzt, um den eigentlichen Nutzdatenverkehr nicht unnötig zu behindern. Regelmäßige Messungen werden derzeit im GEANT2-Netz durchgeführt.

## 3 perfSONAR-Schichtenmodell und Web Services

Für das Monitoring der im vorherigen Abschnitt vorgestellten Metriken ist es neben dem Messen der Daten notwendig, deren Management in geeigneter Weise zu organisieren sowie die Daten mit Hilfe von Visualisierungs- und Analysewerkzeugen zugänglich zu machen. Hieraus wurde im perfSONAR-Projekt [bbdh05] ein Modell mit drei Schichten entwickelt, das in Abbildung 1 dargestellt ist und im Abschnitt 3.1 erläutert wird. In Abschnitt 3.2 wird dann auf die Details der Service Layer eingegangen, d.h. die momentan zur Verfügung gestellten Web Services werden genauer betrachtet.

### 3.1 Die drei Schichten der Messarchitektur

Die Measurement Point Layer ist die unterste Schicht der Architektur. Mit dieser werden Messungen der im vorherigen Abschnitt vorgestellten Metriken durchgeführt, die entweder direkt über ein perfSONAR-Interface zugreifbar sind (sog. Measurement Points, MPs) oder nicht direkt zugreifbare Datensammlungen für Messarchive

(Measurement Archive, MA) realisieren, die über Dienste der Service Layer (s.u.) nutzbar sind. Hierbei werden vom Rahmenwerk keine festen Vorgaben gemacht, welche

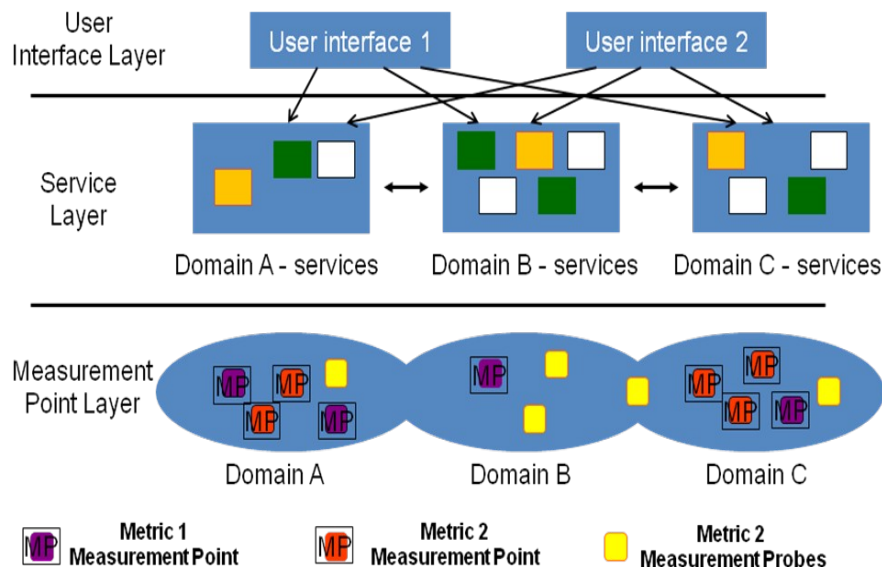


Abbildung 1: Das Drei-Schichten-Modell von perfSONAR

Arten von Messpunkten installiert sein müssen, sondern die Auswahl bleibt den einzelnen Domains überlassen.

Die Service Layer dient dem Management der Messungen innerhalb und zwischen Netzwerkdomänen. Der Name der Schicht bezieht sich auf deren Implementierung als Web Services, wodurch eine logische Aufteilung von unterschiedlichen Funktionen erreicht wird. Die Aufgabe eines Dienstes kann beispielsweise die Authentifizierung und Autorisierung von Nutzern, die Verwaltung von anderen Diensten in einem Dienstverzeichnis, der Schutz von Ressourcen zur Überlastvermeidung oder die Archivierung von Messdaten sein wie in 3.2. genauer dargestellt. Die Kommunikation der Dienste erfolgt über SOAP und verwendet dabei ein gemeinsames Protokoll der NMWG (Network Measurement Working Group) des OGF (Open Grid Forum), in dessen Standardisierung die Anforderungen von perfSONAR berücksichtigt werden.

Auf der obersten Schicht befinden sich Visualisierungswerkzeuge [hjkm06], die die gemessenen Kennzahlen in graphischer Form darstellen und deren Analyse ermöglichen.

### 3.2 perfSONAR Web Services

Die wichtigsten Dienste der perfSONAR Service Layer werden im folgenden kurz beschrieben.

Der Telnet/SSH MP ist ein Werkzeug, um mittels der Protokolle Telnet oder SSH auf Router zuzugreifen. Der Messpunkt ermöglicht derzeit den Zugriff auf Cisco und Juniper Router und übersetzt herstellerunabhängige Befehle in deren proprietäre Kommandos. Der MP ermöglicht damit (zusammen mit dem dazu entwickelten Visualisierungstool) eine vereinfachte Fehlersuche über Domaingrenzen hinweg.

Auf den Hades-Messboxen werden aktive Messungen von OWD, IPDV und OWPL durchgeführt. Diese Messungen sind nicht direkt zugreifbar, sondern werden zentral beim DFN-Labor gesammelt und in einer Datenbank verwaltet. Diese Datenbank ist mit einer perfSONAR-Schnittstelle ausgestattet und wird als HADES MA bezeichnet. Derzeit sind insgesamt etwa 100 Messstationen in sieben Domains im Einsatz, die kontinuierlich Hades Messungen durchführen.

Der BWCTL MP (Bandwidth Controller Measurement Point) ist ein Dienst zum Testen der verfügbaren Bandbreite. Dabei wird zwischen zwei MPs soviel Verkehr erzeugt, wie man erfolgreich übertragen kann. BWCTL Messungen sind ebenfalls auf allen Messboxen möglich, es wird jedoch nicht vollvermascht gemessen.

Der am weitesten verbreitete Service (Installationen in 15 Domains) ist das Round Robin Database Measurement Archive (RRD MA), das eine perfSONAR-Schnittstelle für den Zugriff auf Daten in RRD-Dateien dargestellt. Hierbei geht es vor allem um Auslastungsdaten, aber auch um Paketverwerfungen an Routerinterfaces. Alternativ zum RRD MA gibt es das SQL MA, das einen Zugriff auf SQL-Datenbanken wie z.B. MySQL oder PostgreSQL ermöglicht.

Der Lookup Service ist ein Verzeichnisdienst für perfSONAR-Services, bei dem sich Services registrieren und darüber gefunden werden. Er wird in den nächsten Monaten dahingehend erweitert, dass die Lookup Services eine Föderation bilden und damit Anfragen an andere Lookup Services weiterleiten können.

Der Authentication and Authorization Service dient dazu, dass einzelne Domains festlegen, wie Installationen von Diensten in der Domain verwendet werden können. Der Dienst erlaubt dabei die Festlegung einer Zuordnung zwischen Attributen eines Nutzers und dessen Rechten, z.B. kann man einem Mitarbeiter eines NOCs den Zugriff auf eine Telnet/SSH MP-Installation gewähren und dieses für andere Nutzer untersagen.

cNIS (common Network Information Service) ist eine Topologiedatenbasis, die statische Topologieinformationen über eine Domain zur Verfügung stellt. Dieses wird insbesondere für die Visualisierungs- und Analysetools gebraucht.

## **4 Visualisierung und Benutzerschnittstellen**

Die perfSONAR Web Services erlauben den Zugriff auf die Messungen über eine XML-Schnittstelle, auf die in JRA1 entwickelte Visualisierungstools zugreifen. Die Tools sind zielgruppenspezifisch entwickelt worden. Als Zielgruppen werden dabei Mitarbeiter von Network Operation Centers (NOC), Projektgruppen, einzelne Wissenschaftler und

administrative Mitarbeiter unterschieden (näheres hierzu in [hjkm06]). Den Schwerpunkt der Aktivitäten liegt im Moment auf den NOC-Mitarbeitern, denen genaue Daten zur Fehleranalyse bereitgestellt werden sollen, sowie bei der Unterstützung von Projekten, insbesondere des LHC-Projekts.

Das Tool perfsonarUI bietet einen direkten Zugriff auf perfSONAR-Dienste und stellt diese gruppiert nach Metriken dar. Das Tool setzt Grundkenntnisse über perfSONAR voraus und ist für die Fehleruntersuchung konzipiert. Abbildung 2 zeigt die Darstellung von Auslastungsdaten, bei der man entweder die gesamten Daten in einem Archiv anzeigen kann (im Beispiel vom Geant2 MA) oder eine Route nachverfolgen kann. Diese wird mit einem Traceroute beschrieben und die Daten aus den entsprechenden MAs werden so zusammengetragen, dass man Stellen mit hoher Auslastung entlang des Pfades, die möglicherweise zu Beeinträchtigungen führen, identifizieren kann.

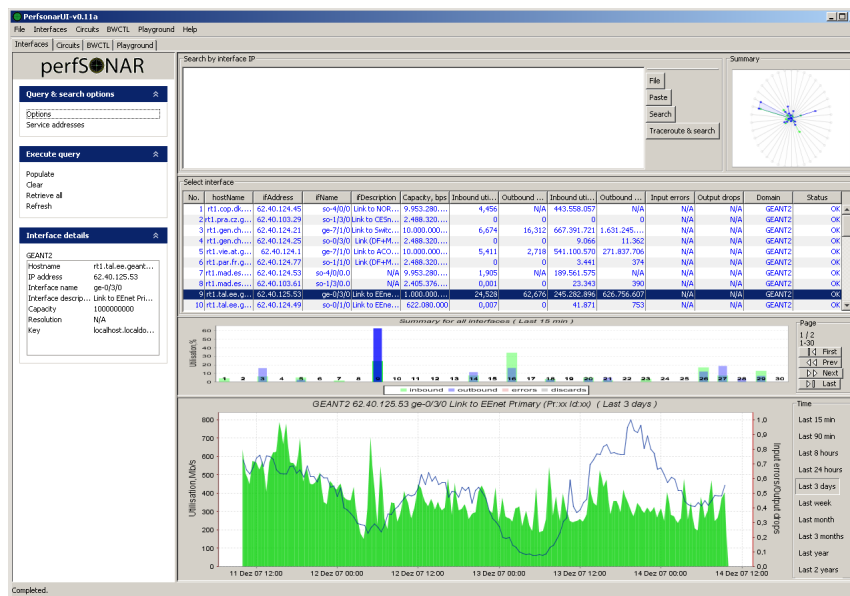


Abbildung 2: Darstellung von Auslastungsdaten im perfsonarUI

In perfsonarUI stehen außerdem weitere Oberflächen zur Darstellung von HADES-Daten und für BWCTL-Messungen zur Verfügung.

Das VisualperfSONAR-Tool läuft innerhalb einer Web-Oberfläche und ist auf die Analyse von Pfaden spezialisiert. Bei diesen werden die Daten von verschiedenen Metriken zusammengetragen und sollen in der Zukunft in einen automatischen Report zusammengestellt werden. Der Pfad wird zur Veranschaulichung in Google Maps mit den Koordinaten der Interfaces dargestellt.

Das CNM-Tool stellt die Topologie von Netzen mittels hierarchischer Karten dar. Die Karten beinhalten sowohl Netzknoten (Router) und Verbindungen (Links) zwischen

unterschiedlichen Knoten als auch Statusinformationen und Kennzahlen. Eine Baum- und eine graphische Darstellung stehen dabei zur Verfügung. In der Baumdarstellung kann man einzelne Unterbäume ausklappen und Sichten auswählen, die dann in der graphischen Darstellung auf der rechten Seite angezeigt werden. In Abbildung 3 wird die Topologie des bulgarischen Forschungsnetzes dargestellt und die auf den Links angezeigte Kennzahl ist der Durchsatz. Es ist zusätzlich möglich zwischen den unterschiedlichen Kennzahlen (Jitter, Delay, Paketverlust usw.) beliebig zu wechseln (links oben im Fenster). Diese Karten zeigen den aktuellen oder historischen Status eines Netzes an. Für einzelne Knoten oder Links stehen Tages-, Wochen-, Monats- und Jahresstatistiken zur Auswahl, um den Verlauf der Messungen zu betrachten.

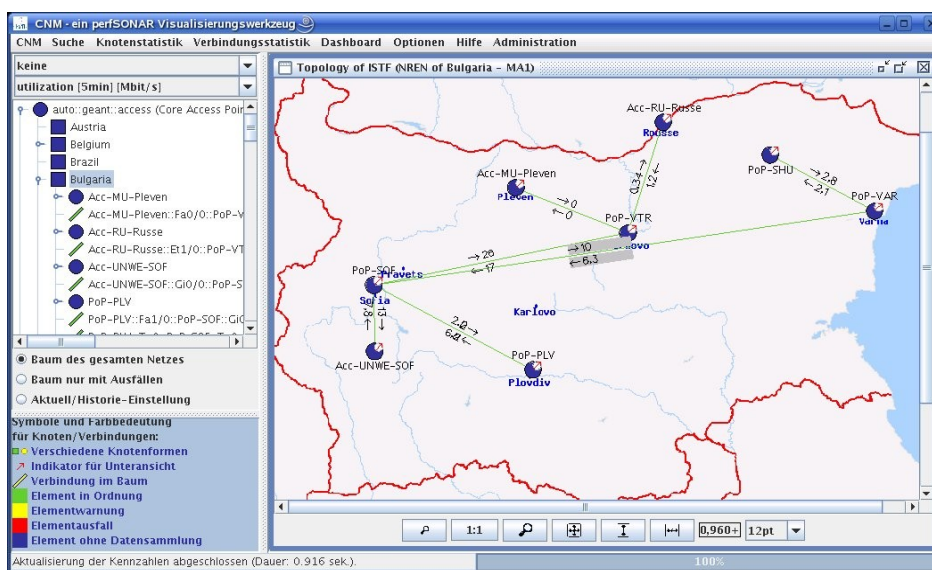


Abbildung 3: Netztopologie des bulgarischen Forschungsnetzes

Das CNM-Tool kann kundenspezifisch angepasst werden, indem man für verschiedene Kunden verschiedene Karten bzw. Kennzahlen freigibt. Das wird durch einen Authentisierungsmechanismus realisiert und ist in Projekten dieser Größenordnung notwendig.

Parallel zu perfonarUI wurden zur Darstellung der gemessenen Metriken OWD, OWPL und IPDV am DFN-Labor geeignete Visualisierungen entwickelt [HADES]. Als primäre Übersicht werden die Messdaten in Form einer Karte dargestellt. Entsprechend der Zugehörigkeit zu einer Messdomain erhält man so ein Bild der Delaymessungen entlang der Netzwerktopologie. Zusätzlich können von allen Standorten aus die vorhandenen Messungen (derzeit ca. 2400) zur zeitlichen Darstellung selektiert werden.

Auch BWCTL Messungen können in die Datendarstellung integriert werden. Ein Beispiel für die Darstellung der Messdaten findet sich in Abbildung 4.

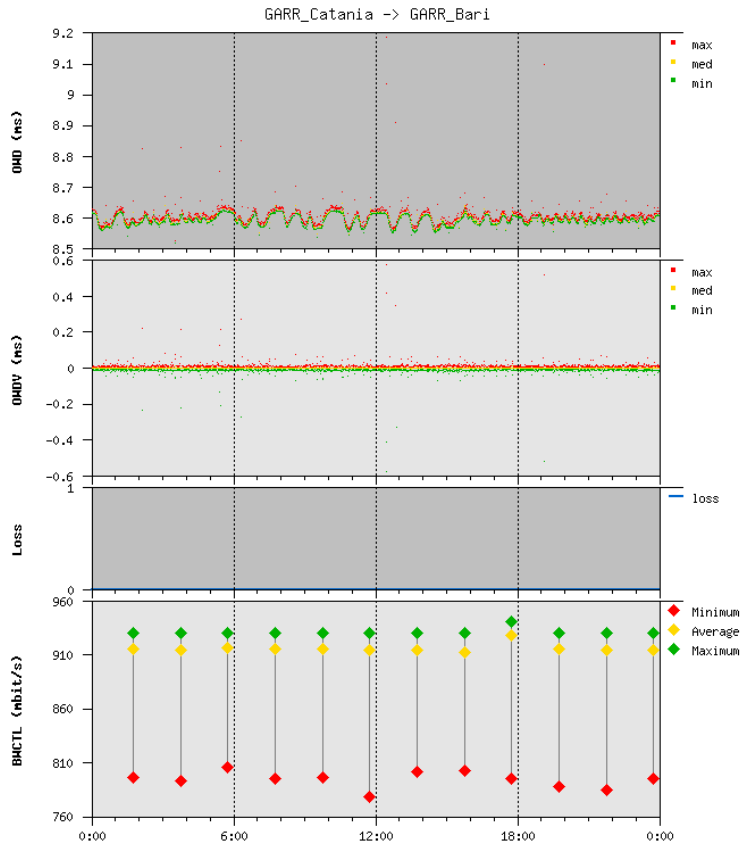


Abbildung 4: Delay und Bandbreitenmessungen

## 5 Zusammenfassung und Ausblick

Im europäischen Verbundprojekt GEANT2 entsteht unter Mitwirkung der angeschlossenen nationalen Forschungsnetze unter der Leitung von DANTE ein Rahmenwerk zur Leistungsüberwachung in nationalen und internationalen Weitverkehrsnetzwerken.

Der Ansatz perfSONAR stellt einen einzigartigen Weg dar, über Domänengrenzen hinweg Methoden zur Ermittlung von Performance-Metriken integrativ zur Verfügung zu stellen und Messergebnisse zusammenfassend über benutzerorientierte Oberflächen zu visualisieren (Multi Domain Monitoring).

In einem Drei-Schichten Modell werden, ausgehend von unabhängigen Messwerkzeugen, über eine Service-Oriented-Architecture (SOA) Schnittstellen



definiert, die maßgeschneidert für die Erfordernisse der Benutzer die gewünschten Informationen liefern.

perfSONAR soll es jedem Netzbetreiber ermöglichen, auf einfache und transparente Weise eine Monitoring-Infrastruktur aufbauen zu können, die entsprechend seiner Vorgaben die passenden Werkzeuge zur Verfügung stellt.

Neben dem eigentlichen Monitoring ist angestrebt, die erzeugten Daten zu analysieren und bedarfsgerechte Alarmsysteme zu kreieren, um die perfSONAR Infrastruktur effizient nutzen zu können.

Im Laufe des Jahres 2008 ist die Installation der Services in einer Reihe von europäischen Netzen (auch im X-WiN) geplant, um dann im Regelbetrieb die Überwachung der Netzqualität in der beschriebenen Weise zu ermöglichen.

## Literaturverzeichnis

- [Holl2007] T. Holleczeck: Redesign und Implementierung eines Softwarepakets zur Messung der IP Performance nach OWAMP-Standard (Studienarbeit), Erlangen 2007.
- [DANTE]  
[RFC 2330] [www.dante.net](http://www.dante.net)  
Paxson, V. , G. Almes, J . Mahdavi , M. Mathis: RFC 2330: Framework for IP Performance Metrics. RFC, IETF, Mai 1998, <http://www.rfc-editor.org/rfc/rfc2330.txt>
- [RFC 2679] Almes, G. , S. Kalindindi, M. Zekauskas: *RFC 2679: A One-way Delay Metric for IPPM*. IETF, September 1999, <http://www.rfc-editor.org/rfc/rfc2679.txt>
- [RFC 2680] Almes, G. , S. Kalindindi, M. Zekauskas: *RFC 2679: A One-way Packet Loss Metric for IPPM*. IETF, September 1999, <http://www.rfc-editor.org/rfc/rfc2680.txt>
- [RFC 3393] Demichelis, C., P. Chimento: *RFC 3393: IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)*. IETF, November 2002, <http://www.rfc-editor.org/rfc/rfc3393.txt>
- [RFC 3148] Mathis, M., M. Allman: *RFC 3148: A Framework for Defining Empirical Bulk Transfer Capacity Metrics*. IETF, Juli 2001, <http://www.rfc-editor.org/rfc/rfc3148.txt>
- [iperf] Iperf. <http://dast.nlanr.net/Projects/Iperf/>
- [BWCTL] Bandwidth Test Controler (BWCTL). <http://e2epi.internet2.edu/bwctl/>
- [CNM] Customer Network Management (CNM). <http://www.cnm.dfn.de/>
- [DFN-2] IPPM – Performancemessungen in Hochgeschwindigkeitsnetzen. <http://www-win.rze.uni-erlangen.de/ippm/messprogramm.html>
- [hjkm06] Hanemann, A., Jeliakov, V., Kvittem, O., Marta, L., Metzger, J., Velimirovic, I., Complementary Visualization of perfSONAR Network Performance Measurements, In Proceedings of the International

Conference on Internet Surveillance and Protection (ICISP), 2006, IARIA/IEEE, Cap Esterel, France, August, 2006.

[bbdh05]

Boote, J. W., Boyd, E. L., Durand, J., Hanemann, A., Kudarimoti, L., Lapacz, R., Swany, D. M., Zurawski, J., Trocha, S., PerfSONAR: A Service Oriented Architecture for Multi-Domain Network Monitoring, In Proceedings of the Third International Conference on Service Oriented Computing, 241–254, LNCS 3826, Springer Verlag, ACM Sigsoft, Sigweb, Amsterdam, The Netherlands, Dezember, 2005.

[hlms06]

Hanemann, A., Liakopoulos, A., Molina, M., Schmitz, D., Solberg, A., Swany, D. M., Velimirovic, I., van Maele, A., van den Berghe, S., Deliverable DJ1.2.3: Network Metric Report, GEANT2/JRA1 project report, Februar, 2006.