Toolset zur Planung und Qualitätssicherung von verteilten virtuellen Netzwerkstrukturen

Martin Gründl, Susanne Naegele-Jackson

Friedrich-Alexander Universität Erlangen-Nürnberg
Martensstrasse 1, 91058 Erlangen
{martin.gruendl, susanne.naegele-jackson}@rrze.uni-erlangen.de

Zusammenfassung: Durch die zunehmende Verbreitung von leicht handhabbarer Software für anspruchsvolle Anwendungen wie Videoconferencing oder Server-Virtualisierung sind neue verteilte virtuelle Netzwerkstrukturen und zwischenbetriebliche Vernetzungen möglich. Bei solchen verteilten virtuellen Strukturen ist es für den Benutzer von großem Vorteil, wenn schon im Vorfeld abgeklärt werden kann, ob gewisse Applikationen über alle verteilten Standpunkte nutzbar sind und auch in der entsprechenden Dienstgüte gewährleistet werden können. Diese Arbeit beschreibt ein Toolset, welches dazu verwendet werden kann, Schwachstellen in einem Multi-Domain-Netz aufzuzeigen, Dienstgüteparameter zu messen und die Erfüllung der Anforderungen einer Applikation für eine verteilte virtuelle Organisationsstruktur zu testen. Weiterhin werden Mechanismen zur Überprüfung von Protokollparametern und deren prinzipieller Funktion sowie Werkzeuge zur Bewertung der IPv6-Konnektivität angeboten. Um den verteilten Einsatz des Systems zu erleichtern, liegt ein besonderer Fokus auf der Integration verschiedener Sensor-Implementierungen, um sowohl kostengünstige Messpunkte mit Hilfe von Plug Computern oder virtuellen Maschinen als auch hochpräzise Sensoren für hohe Bandbreiten-Anforderungen realisieren zu können.

1 Einleitung

Die informationstechnische Vernetzung bietet Firmen die Möglichkeit, ihre Betriebsprozesse über neue virtuelle Organisationsstrukturen zu verteilen. Bei einer solchen Verteilung kann es sich beispielsweise um eine Teilmenge von Firmenstandorten handeln, die alle an einem bestimmten Projekt arbeiten und zusätzlich auch externe Berater in diese projektabhängige Organisationsstruktur einbinden möchten [We06]. Das Netz als Verbindungsglied dieser Struktur stellt die Basis der gemeinsamen Projektarbeit dar und nimmt dadurch einen sehr hohen Stellenwert ein. Störungen im Netz oder ungenügende Dienstqualität für einzelne Standorte der virtuellen Netzwerkstruktur können zu einer massiven Beeinträchtigung der Arbeitsabläufe innerhalb des gesamten Projektes führen. Es ist daher von besonderem Vorteil, wenn Netzwerkzeuge zur Verfügung stehen, die zum einen die Fehlersuche bei Störungen im Netz vereinfachen, auf der anderen Seite aber auch Möglichkeiten bieten, Applikationen schon im Vorfeld auf ihre Einsetzbarkeit an den einzelnen Standorten zu testen. Ein

typischer Fall hierfür sind Videokonferenzen, die über alle Standorte hinweg in ausreichend guter Dienstqualität durchführbar sein sollten. Der Benutzerkreis möchte sich vor der Nutzung der Applikation einen Überblick darüber verschaffen, ob über die Teilbereiche der Netzinfrastruktur genügend Ressourcen zur Verfügung stehen, so dass die erforderliche Dienstgüte der Anwendung gewährleistet werden kann.

Diese Arbeit beschreibt TSS, einen Netzwerk Troubleshooting Service, der Funktionstests ermöglicht, welche für den Einsatz in verteilten virtuellen Strukturen von Interesse sind. Zum Funktionsumfang von TSS zählen unter anderem prinzipielle Netzfunktionstests wie `ping` und `traceroute`, darüber hinaus kann aber mit TSS auch die prinzipielle Funktionalität von Netzservices wie DHCP oder DNS überprüft werden. TSS erlaubt außerdem die Bestimmung von protokoll- und anwendungsspezifischen Parametern und bietet die Möglichkeit, IP Performance Metrics zur Qualitätssicherung einzusetzen.

2 Verwandte Arbeiten

Bei TSS handelt es sich um eine Weiterentwicklung von perfSONAR-Lite Troubleshooting Service (TSS) [NGH10], das im Rahmen des EGEE-III (Enabling Grids for E-sciencE) Projektes entwickelt wurde. Da EGEE als eines der führenden Grid-Projekte mehr als 280 Einrichtungen in Europa vernetzte, war es notwendig, eine Lösung zur Fehlersuche auf Netzebene zur Verfügung zu haben, um schnell bei akuten Problemen eine leicht handhabbare Untersuchungssoftware einsetzen zu können. Diese am Regionalen Rechenzentrum Erlangen entwickelte Software basierte auf Teilen des im GN2/GN3-Projekt entwickelten Monitoringsystems perfSONAR [BBD05,HKM08] und stellte Tools wie `ping`, `traceroute`, `dns lookup`, `nmap` [Nm11] sowie `bwctl` [Bw11] für Bandbreitenmessungen zur Verfügung. In dieser Arbeit wird nun die vom DFN-Verein finanzierte Weiterentwicklung von TSS vorgestellt, die außer den bereits vorhandenen für EGEE-III entwickelten Tools weitere Werkzeuge enthält, die nicht nur für eine Fehlerdiagnose entscheidend sind, sondern darüber hinaus zusätzliche Netzfunktionalitäten prüfen können und daher vor allem auch bei der Planung von verteilten virtuellen Organisationsstrukturen oder zur Überprüfung von Qualitätsanforderungen bei Anwendungen hilfreich sind.

Ein vergleichbares Schnelldiagnose-System findet man in der Literatur unter NDT (Network Diagnostic Tool [Nd11]): NDT bietet unter anderem Tests, um Ethernet-Duplex-Einstellungen zu überprüfen oder Links zu ermitteln, die nicht den zugesagten Datendurchsatz erreichen. Allerdings führt NDT diese Tests immer zwischen dem Desktop Computer des Benutzers und dem NDT-Server durch. TSS hingegen erlaubt die Schnelldiagnose losgelöst vom PC eines Benutzers zwischen zwei beliebigen Sensoren im Netz. Dadurch wird eine Schnelldiagnose im Fehlerfall flexibler und auch Qualitätsanforderungen für Anwendungen lassen sich präzise über den betroffenen Pfad bestimmen.

Im folgenden Abschnitt wird zunächst die Implementierung von TSS erläutert, während Abschnitt 4 die möglichen Anwendungsszenarien der TSS Werkzeuge im De-

taits beschreibt. Schließlich werden in Abschnitt 5 die Kernaussagen zusammengefasst und ein Ausblick auf mögliche weitere Entwicklungen gegeben.

3 Architektur

Die grundlegende Architektur des TSS-Systems setzt sich aus einem Server und den für die eigentlichen Messungen genutzten Sensoren zusammen.

3.1 Zentraler Server

Der Webserver als zentrale Komponente kommuniziert über einen abgesicherten Kanal mit den Sensoren und stellt die Schnittstelle für die Benutzer zur Verfügung. Der Ablauf einer entsprechenden Messung ist in Abbildung 1 dargestellt.

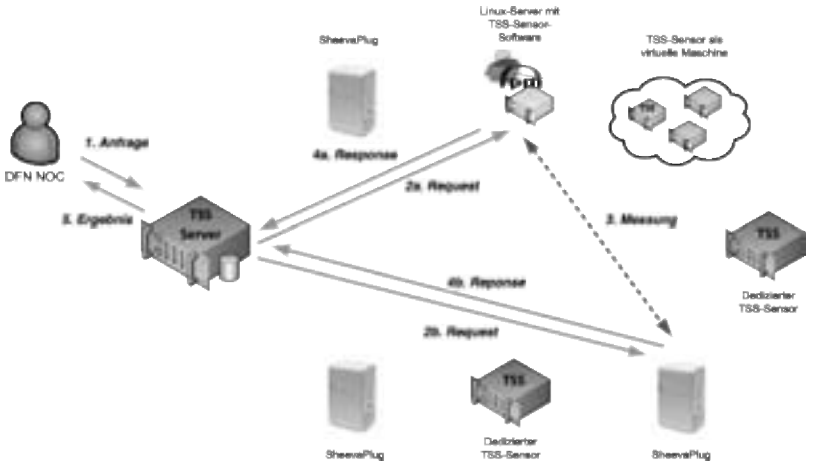

Abbildung 1: Ablauf einer Messung

Das Webserver Interface führt eine zugangsberechtigte Person durch die Auswahl der entsprechenden Messtypen und Parameter (1) und leitet schließlich die entsprechenden Mess-Aktionen über einen SSL-Kanal an die Sensoren des Systems weiter (2). Diese wiederum nehmen ausschließlich über diesen Kanal Befehle entgegen, führen sie aus (3) und liefern die Ergebnisse wieder an den Server zurück (4). Schließlich werden die Resultate der Messung dem Benutzer präsentiert (5).

3.2 Sensoren

Um einen möglichst breiten Bereich an Anwendungsfällen abdecken zu können, ist es ein zentrales Ziel des hier vorgestellten Systems, an verschiedene Einsatzszenarien angepasste Sensoren zur Verfügung zu stellen (Abb. 2). Grundsätzlich handelt es sich

bei einem solchen Sensor um einen speziellen Client zur Durchführung von Messungen und Funktionstests zwischen zwei Messpunkten; die dazugehörigen Ergebnisse werden dann über das Webserver Interface dem Anwender übermittelt.

Neben der für das EGEE-III Projekt entwickelten Software-Lösung, welche auf einem bestehenden Linux-Server in wenigen Schritten installiert werden kann und daher auch keine zentrale Administration des Systems erlaubt, sind für mit begrenzten Mitteln ausgestattete Vorhaben oder noch in der Planungsphase befindliche Projekte vor allem die Sensoren auf Basis der SheevaPlug-Familie [Sh11] interessant. Dieses Gerät, welches trotz seiner Kompaktheit und der sehr geringen Anschaffungskosten ein vollständiges Linux-System zur Verfügung stellt, benötigt lediglich eine Netzwerkverbindung sowie einen Stromanschluss, um seinen Betrieb aufnehmen zu können.

Da das Betriebssystem dieser Sensoren von einer SD-Karte gestartet wird, ist im Fehlerfall keine langwierige und kostenintensive Wiederherstellung erforderlich. Im für den Benutzer einfachsten Fall wird lediglich die SD-Karte getauscht, alternativ kann sie natürlich auch mit Hilfe eines für registrierte Nutzer zugänglichen Images auf einen definierten Zustand zurückgesetzt werden.

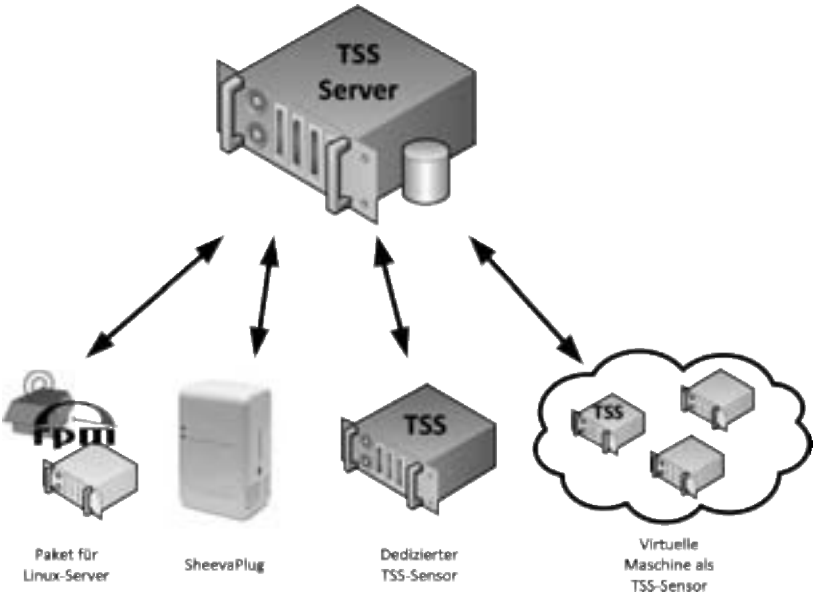

Abbildung 2: Verfügbare Sensoren

Natürlich besteht weiterhin die Möglichkeit, einen vollwertigen Server, ähnlich zu den erfolgreich in nationalen und internationalen Forschungsnetzen betriebenen HADES-Messrechnern [Ha11], in das System zu integrieren. Ein solcher HADES Messrechner erlaubt aufgrund seiner Bauform und seiner erweiterten technischen

Ausstattung vor allem im Bereich des Zeitverhaltens mit Hilfe einer GPS-Karte sehr viel präzisere Messungen und kann beispielsweise auch - ausgerüstet mit einer entsprechend leistungsfähigen Netzwerkkarte - Durchsatzmessungen bis in den Bereich von mehreren zehn Gigabit pro Sekunde durchführen.

Gerade im Rahmen der rasant zunehmenden Virtualisierung ist es darüber hinaus möglich, eine mit der entsprechenden Software ausgestattete virtuelle Maschine ebenfalls als Sensor zu integrieren, um auch innerhalb derartiger virtueller Umgebungen die diversen Möglichkeiten des TSS-Systems nutzen zu können.

3.3 Verfügbare Werkzeuge

Um die zahlreichen Anwendungsfälle des Systems abdecken zu können, wurde eine breit gefächerte Auswahl an Werkzeugen integriert, welche in der folgenden Tabelle 1, jeweils nach Funktionalität gegliedert, dargestellt ist.

Tabelle 1: Werkzeuge

Konnektivität	IPv6-Unterstützung	Protokoll- und anwendungs-spezifisch	IPPM-Metriken und Qualitäts-Indikatoren
ping	ping6	tracepath	OWAMP
traceroute	traceroute6	tracepath6	iperf
host / nslookup	Router Discovery	netcat	
DHCP-Diagnose		nmap	
NTP			

3.4 Verwaltung der Berechtigungen

Die Authentifizierung der Benutzer erfolgt über X.509-Zertifikate, um eine eindeutige und zuverlässige Identifikation sicherstellen zu können. Die Berechtigungen innerhalb des Systems können wiederum spezifisch auf Ebene der Sensoren zugewiesen werden. Einem nicht privilegierten Benutzer des Systems ist es somit möglich, jeweils Messungen ausgehend von einem für ihn freigegebenen Sensor zu allen übrigen Sensoren auszulösen, sowie in entgegengesetzter Richtung von allen übrigen Sensoren zu den für ihn freigegebenen Sensoren durchzuführen.

4 Anwendungsszenarien

Um die Arbeit der mit dem Aufbau und dem Betrieb eines Netzwerkes befassten Personen zu unterstützen, werden mit Hilfe der bereits beschriebenen Architektur zahlreiche Parameter, welche einen Netzwerkanschluss charakterisieren, ermittelt, um

sowohl bei der Planung eines Netzes als auch bei der späteren Behebung von Anomalien oder Fehlern möglichst effizient Hinweise auf Fehlerquellen zu erhalten. Im Folgenden werden einige Szenarien für den Einsatz des TSS-Systems skizziert.

4.1 Überprüfung der prinzipiellen Funktionsfähigkeit und Qualität der Netzanbindung

Im Rahmen der Planung neuer PoPs oder neuer Standorte von virtuellen Organisationsstrukturen ist zunächst eine uneingeschränkte Konnektivität der dortigen Netzanbindungen abzuklären. Neben den trivialen Mechanismen zur Überprüfung der Erreichbarkeit wie ping ist es hier beispielsweise von Interesse, ob alle bestehenden oder geplanten PoPs von dem jeweils untersuchten Standort aus erreicht werden können oder ob hier eventuell gewollte oder durch Störungen verursachte Einschränkungen vorliegen. Neben der prinzipiellen Erreichbarkeit der Gegenstellen ist natürlich auch die Nutzung gleichbleibender Wege innerhalb der Transit-Netzwerke erstrebenswert. Auch wenn dies prinzipiell keine Anforderung im Rahmen des IPv4-Standards ist, da dieser paketvermittelt arbeitet und somit explizit auf die variable Wegführung der IP-Pakete ausgelegt ist, sind aus Gründen des Zeitverhaltens und anderer potentiell störender Nebeneffekte konstant bleibende Routen, welche sich lediglich im Fehlerfall oder während Wartungsarbeiten verändern, ein relevantes Qualitätsmerkmal. Zu diesem Zweck kommen, abhängig von den konkreten Anforderungen und zu bestimmenden Parametern, Vertreter der unter dem Oberbegriff traceroute zusammengefassten Werkzeuge zum Einsatz.

Darüber hinaus ist natürlich auch die einwandfreie Funktionalität von Netzdiensten wie dem Domain Name System DNS und dem Dynamic Host Configuration Protocol DHCP (Abb. 3) eine Voraussetzung für eine zuverlässige und leistungsfähige Netzanbindung. Daher können hier ebenfalls einige grundlegende Funktionen geprüft werden, beispielsweise wie sich der lokal angebotene DNS-Server im Falle wiederholter identischer Anfragen verhält, um eventuelle Caching-Effekte erkennen und gegebenenfalls berücksichtigen zu können.

Abbildung 3: Screenshot DHCP-Diagnose

4.2 Qualität der IPv6-Konnektivität

Die bisher genannten Messungen bekommen, vor allem aufgrund der Erschöpfung des verfügbaren IPv4-Adressraumes, mit der rasch zunehmenden Verbreitung des Protokolls IPv6 eine neue Dimension, da in naher Zukunft sehr viele Einrichtungen ihre entsprechenden Netzstrukturen an das prinzipiell bereits seit mehr als 10 Jahren standardisierte und verfügbare Protokoll anpassen und den zuverlässigen Betrieb dieses Protokolls sicherstellen müssen. Da dieser Vorgang bisher keineswegs abgeschlossen ist, werden momentan noch diverse Mechanismen genutzt, um auch ohne native IPv6-Konnektivität bereits das entsprechende Protokoll nutzen zu können. Diese Technologien weisen durch ihre Aufgabe natürlich einen inhärenten Einfluss auf das Verhalten der darüber verwendeten Protokolle und im Speziellen auf das Zeitverhalten auf, welches beispielsweise für zeitkritische Projekte durchaus relevante Auswirkungen haben kann.

Daher ist an dieser Stelle ein wesentliches Qualitätsmerkmal der IPv6-Konnektivität, ob diese uneingeschränkt von der den Netzanschluss zur Verfügung stellenden Infrastruktur unterstützt wird oder vorerst der Einsatz der beschriebenen Mechanismen nötig ist. Durch die Überprüfung diverser Parameter am topologischen Standort des Senders und einige Messungen kann sowohl das prinzipielle Verfahren der IPv6-Anbindung als auch deren Qualität gemessen und überprüft werden.

4.3 Bestimmung von protokoll- und anwendungsspezifischen Parametern

Neben der prinzipiellen Funktionalität einer Netzanbindung, welche oben beschrieben wurde, sind weitere Parameter für die optimale Nutzung von Protokollen höherer Schichten und Anwendungen relevant. Um beispielsweise eine möglichst effiziente Übertragung von großen Datenmengen realisieren zu können, ist der Parameter Maximum Transmission Unit (MTU) von großer Bedeutung, da dieser Parameter die maximale Größe eines IP-Paketes beschreibt. Im optimalen Fall ist die Nutzung von Jumbo Frames mit einer Größe von bis zu 9000 Bytes möglich, wodurch sich sowohl volumenmäßig als auch auf technischer Ebene der Overhead gravierend reduzieren lässt. Daher sind Werkzeuge zur Bestimmung der MTU auf einem Pfad ebenfalls Bestandteil der hier beschriebenen Lösung.

Mit Hilfe des hier vorgestellten Werkzeugs kann somit eine Verbindung mit den entsprechenden Quell- und Zielports beziehungsweise den Protokollnummern simuliert werden, um einen prinzipiellen Verbindungsaufbau nachstellen zu können. Neben der Funktionsfähigkeit von Anwendungen, welche im späteren Betrieb eingesetzt werden sollen, und eventuell störenden Netzkomponenten wie beispielsweise Access-Listen oder Firewalls wird mit diesem Mechanismus zudem überprüft, ob die für die Administration der TSS-Sensoren nötigen Verbindungen sowie die für Messungen benötigte Werkzeuge und ihre jeweiligen Protokolle über den entsprechenden Netzanschluss nutzbar sind.

4.4 Bestimmung von IPPM-Metriken und weiterer Qualitätsindikatoren

Die Qualität einer Netzanbindung wird neben den für eine grundlegende Funktion nötigen Protokollen und Parametern im Wesentlichen durch Metriken, welche größtenteils unter dem Oberbegriff IP Performance Metrics zusammengefasst werden, erfasst und bewertet. Vor allem für die Übertragung von Multimediadaten zur Echtzeitkommunikation sowie für zeitkritische Steuerungsaufgaben haben die Parameter One-Way Delay und One-Way Delay Variation erhebliche Bedeutung, da sie das zeitliche Verhalten einer Verbindung im Bezug auf eventuelle Verzögerungen oder Schwankungen im Zeitverhalten charakterisieren.

Im Rahmen des Einsatzes von Voice over IP ist dies beispielsweise relevant, um die Größen der für den Ausgleich von Laufzeitschwankungen nötigen Pufferspeicher vorab oder auch dynamisch während des Betriebs festlegen zu können. Ein positiver Nebeneffekt einer derartigen Messung ist weiterhin eine Erkennung von gegebenenfalls auftretenden Paketverlusten.

Zu diesem Zweck wird das One-Way Active Measurement Protocol OWAMP eingesetzt (Abb. 4), welches beide Parameter auf der Grundlage entsprechender RFCs [STK06] bestimmt. Natürlich wäre prinzipiell auch der Einsatz des aktiven HADES-Messsystems denkbar, allerdings ist dies in seiner momentanen Form auf vorkonfigurierte periodische Messungen ausgerichtet und unterstützt keine Ad-hoc-Messungen.

Abbildung 4: Screenshot OWAMP-Messung

Neben dem Zeitverhalten des über die Internetverbindung fließenden Verkehrs ist natürlich gerade für wissenschaftliche Projekte und HPC-Anwendungen, wie beispielsweise die mehrstufige Datenverarbeitung der LHC-Experimente mit ihren enormen Datenmengen, der Durchsatz eines Netzanschlusses von hoher Wichtigkeit. Zur Bestimmung dieser Parameter kommt momentan das Werkzeug `iperf` [Ip11] zum Einsatz.

5 Fazit und Ausblick

Mit dem vorgestellten Werkzeug wurde eine Lösung geschaffen, welche eine Reihe von Anforderungen im Rahmen der Netzwerkplanung und der Erweiterung bestehender Netze sowie der Fehlerbehebung in entsprechenden Netzen abdeckt. Durch die verteilte Infrastruktur, welche durch ein zentrales System gesteuert wird und einen an die jeweiligen örtlichen Gegebenheiten und technischen Anforderungen angepassten Einsatz von Messpunkten verschiedener Komplexität ermöglicht, wird ein sehr breites Spektrum abgedeckt. Während bei sehr verteilten Projekten im Planungsstadium eine kostengünstige und portable Lösung mit Sicherheit vorgezogen wird, können für den Betrieb wichtiger Projekte mit hohen Qualitätsansprüchen bei Bedarf Messpunkte eingesetzt werden, welche deutlich komplexere und auch exaktere Messungen erlauben.
Um die Flexibilität und den Nutzwert des hier beschriebenen Systems weiter zu erhöhen, ist zukünftig die Integration weiterer Werkzeuge angedacht, wie beispielsweise eine anwendungsspezifische Simulation von Datenströmen mit geeigneten Werkzeugen.

Literatur

[We06] Weiss, P., Virtuelle Organisationsstrukturen, Konzept, Analyse, Lösungen und Praxisbeispiele, Öffentlicher Abschluss-Workshop Projekt ARBEIT@VU, FZKI Forschungszentrum Informatik und Institut AIFB Universität Karlsruhe, 02.06.2006, http://www.virtuelleunternehmen.com/arbeit_at_vu/infothek/praesentationen/20060602_workshop_abschluss.pdf

[NGH10] Naegele-Jackson S., Gründl M., Hanemann A., PerfSONAR-Lite TSS: Schnelldiagnose von Netzverbindungen im EGEE-III-Projekt, 3. DFN-Forum „Verteilte Systeme im Wissenschaftsbereich“, 26. Und 27. Mai 2010, Universität Konstanz, Lecture Notes in Informatics (LNI) – Proceedings Series of the Gesellschaft für Informatik (GI), Paul Müller, Bernhard Neumair, Gabi Dreo Rodosek (Hrsg.), Volume P-166, Bonn 2010, pp. 127-136, ISBN 978-3-88579-260-4, http://subs.emis.de/LNI/Proceedings/Proceedings166/P-166.pdf

[BBD05] 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

[HKM08] Hanemann, A., Kraft, S., Marcu, P., Reinwand, J., Reiser, H., Schmitz, D., Venus, V., perfSONAR: Performance Monitoring in europäischen Forschungsnetzen, In *Proceedings Erstes DFN–Forum Kommunikationstechnologien — Verteilte Systeme im Wissenschaftsbereich*, GI–Verlag/DFN, Kaiserslautern, Deutschland, Mai, 2008

[Nm11] NMAP, http://nmap.org/, 06.04.2011

[Bw11] Bandwidth Test Controller (BWCTL), http://e2epi.internet2.edu/bwctl/, 06.04.2011

[Nd11] Network Diagnostic Tool (NDT), http://www.internet2.edu/performance/ndt/, 06.04.2011

[Sh11] SheevaPlug, http://www.plugcomputer.org/, 06.04.2011

[Ha11] HADES Active Delay Evaluation System, http://www.win-labor.dfn.de/, 06.04.2011

[STK06] Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., Zekauskas, M., RFC 4656 – A One-Way Active Measurement Protocol (OWAMP), IETF Network Working Group, 2006

[Ip11] Iperf, http://iperf.sourceforge.net, 06.04.2011