

Passive Netzwerkmessungen am Internet-Zugang der Universität Leipzig

Klaus Mochalski *
Rechnernetze und Verteilte Systeme
Institut für Informatik, Universität Leipzig
Augustusplatz 10-11, 04109 Leipzig
E-Mail: mochalski@informatik.uni-leipzig.de

Zusammenfassung

Passive Netzwerkmessungen erlauben einen präzisen Einblick in ein operationales Netzwerk. Mit unserem passiven Messsystem greifen wir den Internet-Verkehr der Universität Leipzig vor und hinter dem zentralen Zugangs-Router ab und zeichnen ihn zu Analyse-zwecken auf. Wir beschreiben die dabei angewandten Methoden, die Analyseverfahren und präsentieren erste Ergebnisse unserer Auswertungen.

1 Einführung

Vor ca. einem Jahr begannen wir an der Universität Leipzig mit der Machbarkeitsstudie für passive Internetmessungen am Zugang zum Gigabit-Wissenschaftsnetz (G-WiN). Bei einer solchen Messung werden die Header aller ein- und ausgehenden Datenpakete aufgezeichnet und in einer Trace-Datei für spätere Auswertungen gespeichert. Vielversprechende Ergebnisse von Messungen an anderen Punkten im Internet [5], an denen wir beteiligt waren, bestärkten uns in der Überzeugung, dass aus solchen Aufzeichnungen wertvolle Erkenntnis-

se über das komplexe Zusammenspiel von Verkehrsströmen im Internet gewonnen werden können.

Außerdem erlauben sie einen ungefilterten Blick auf den Netzwerkverkehr, wie ihn Log-Mechanismen von Standardnetzwerk-komponenten wie Router und Switches nicht bieten können. Damit konnten wir die Unterstützung des Universitätsrechenzentrums gewinnen, das nicht zuletzt wegen des dramatischen Anstiegs des Volumens an Peer-to-Peer(P2P)-Verkehr ein starkes Interesse an solchen Auswertungen hat. Wir möchten an dieser Stelle für die Unterstützung und Kooperation danken.

Die Implementierung und der Betrieb eines Systems zur Aufzeichnung von Paket-Header-Traces ist mit einer Reihe von Schwierigkeiten verbunden. Auf diese Schwierigkeiten und deren Lösung werden wir im Abschnitt 3 genauer eingehen, nachdem wir im Abschnitt 2 die Konfiguration des Messpunktes Leipzig beschrieben haben.

Aufgrund des riesigen Volumens der anfallenden Daten ergeben sich besondere Anforderungen an die verwendeten Auswertungsmethoden. Diese werden im Abschnitt 4 beschrieben. In den Abschnitten 5 und 6 stellen wir zwei in Leipzig aufgezeichnete Traces vor und präsentieren erste Auswertungen.

*Das Projekt wird an der Universität Leipzig unter Leitung von Prof. K. Irscher in Kooperation mit dem DFN-Verein und dem Universitätsrechenzentrum realisiert.

2 Der Messpunkt Leipzig

Der G-WiN-Zugang der Universität Leipzig ist, wie in Abb.1 dargestellt, zur Zeit durch einen Cisco 7505-Router über eine SONET-OC3c-Verbindung (155,53 MBit/s) mittels Packet-over-Sonet (PoS) realisiert. Die Anbindung an des Campus-Netz erfolgt über eine 1000BaseSX-Verbindung zu einen Catalyst 6509. In beiden Verbindungen wurde ein messtechnischer Abgriff installiert.

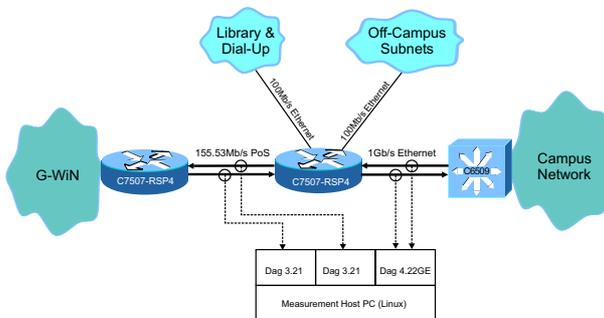


Abbildung 1: Konfiguration des Messpunktes Leipzig

Durch eine Messung an einem der beiden Abgriffpunkte lassen sich Aussagen über die Charakteristika des Datenstroms treffen. Mittels einer Zweipunkt-Messung können darüber hinaus durch den Router verursachte Veränderungen des Datenstroms registriert werden.

Bei einer solchen Zweipunkt-Messung ist zu beachten, dass nur ein Teil des gesamten Internet-Verkehrs über die instrumentierte Gigabit-Ethernet-Verbindung zum zentralen Campus-Switch fließt. Die Bibliothek, die Wählzugänge und einige externe Einrichtungen sind direkt an den Router angeschlossen und entziehen sich damit der Analyse einer Zweipunkt-Messung.

3 Methodologie von passiven Netzwerkmessungen

Bei einer passiven Messung wird für jedes Paket, das über die instrumentierte Netzwerkverbindung fließt, ein Datensatz erzeugt, der aus einem hochpräzisen Zeitstempel und dem Header des Pakets besteht. Die Länge des aufgezeichneten Headers kann dabei konfiguriert werden. Für unsere Messungen werden in der Regel die ersten 40 Byte eines IP-Pakets aufgezeichnet, um neben Analysen auf IP-Ebene auch solche auf TCP- und UDP-Ebene zu ermöglichen.

Der Zeitstempel ist wichtig, um eine Aussage über den genauen Ankunftszeitpunkt eines Pakets treffen zu können. Bei schnellen Netzwerktechnologien dringt man dabei in Bereiche vor, die nur noch mittels spezieller Hardware beherrschbar sind.

Zentraler Bestandteil unseres passiven Messsystems sind die Dag-Messkarten der Firma Endace Measurement Systems [7]. Dabei handelt es sich um PCI-Karten, die in einem Standard-PC unter Linux lauffähig sind. Diese Karten sind für verschiedene Netzwerktechnologien verfügbar, u.a. 10/100BaseT, 1000BaseSX, SONET OC3c/OC12c, OC48c und OC192c.

Die Zeitsynchronisation der Karten erfolgt mittels GPS. Dazu wurde eine GPS-Antenne vom Typ "Trimble Acutime 2000"[8] installiert und direkt mit den Messkarten verbunden. Mehrere Messkarten können dabei in Reihe geschaltet werden. Durch diesen Mechanismus ist auch eine Synchronisation mehrerer geographisch entfernter Messsysteme möglich. Die Synchronisationsgenauigkeit liegt im Bereich um 100 ns [2].

Für den Eingriff in die zu messende Netzwerkverbindung gibt es verschiedene Möglichkeiten. Bei einer Glasfaserverbindung kommen optische Splitter zum Einsatz. Bei Kupfer-

verbindungen können spezielle Pass-Through-Messkarten oder elektronische Leitungssplitter verwendet werden. Alternativ kann der Verkehr auch direkt von Hubs oder Spiegelports abgegriffen werden, wobei allerdings die spezifischen Einschränkungen, wie z.B. Paketverluste am Spiegelport durch fehlende Bandbreite, berücksichtigt werden müssen.

4 Analyseverfahren von passiven Netzwerkmessungen

Während einer Messung mit dem im vorangegangenen Abschnitt beschriebenen System werden Trace-Dateien erzeugt und auf Festplatten aufgezeichnet. Dort stehen sie für spätere Analysen zur Verfügung. Da für jedes einzelne Datenpaket ein Datensatz bestehend aus Zeitstempel und Paket-Header erzeugt wird, entsteht je nach Bandbreite der Netzwerkverbindung ein sehr großes Datenvolumen, das schnell mehrere 100 GByte erreicht.

Daraus ergeben sich spezifische Anforderungen an die Analyse. Sie muss mit hoher Effizienz und zumindest teilautomatisch erfolgen. Zu diesem Zweck haben wir zusammen mit der WAND-Gruppe an der Waikato University in Hamilton, Neuseeland, [4] eine Sammlung von Analysewerkzeugen geschaffen. Dabei handelt es sich um C-Programme, die für eine Trace-Datei verschiedene statistische Auswertungen vornehmen und diese graphisch aufbereiten.

Gegenwärtig gibt es Werkzeuge zur Analyse von Bandbreite und Paketzahl für den gesamten Datenstrom und einzelne Applikationen (basierend auf Port-Nummern). Weiterhin können Flow-Statistiken generiert werden (Anzahl aktive Flows, neue Flows pro Sekunde, Dauer und Volumen von Flows).

Bei den bereits erwähnten Zweipunkt-Messungen entstehen zwei Trace-Dateien simultan – eines für jeden Abgriffpunkt. Wir haben ein spezielles Programm entwickelt, dass

die zusammengehörigen Pakete aus beiden Traces herausucht und dann anhand des Zeitstempels die Verzögerungszeit zwischen den Messpunkten ermittelt.

5 Leipzig-I

Für die erste Messung am Internetzugang der Universität Leipzig wurde nur der Messpunkt in der SONET-Verbindung zum G-WiN benutzt. Dort wurden in der Zeit vom 21.11. bis 26.11.2002 ca. 3,8 Mrd. Paket-Header aufgezeichnet und in Trace-Dateien von insgesamt 226 GByte Volumen gespeichert. Durch Kompression mittels `gzip` konnte das Volumen auf 100 GByte reduziert werden.

Aus diesen Daten werden mit Hilfe der im Abschnitt 4 beschriebenen Werkzeuge verschiedene Grafiken erzeugt, die einen ersten Einblick in das Verhalten und die Zusammensetzung des Datenstroms erlauben. Um die Weiterverarbeitung der Traces zu vereinfachen, werden sie in 6-Stunden-Abschnitte zerlegt. Abb.2 zeigt einen solchen Abschnitt beginnend am 4.11.2002 um 18:00 Uhr. Dargestellt ist die durchschnittliche Bandbreite für 1-Minuten-Intervalle aufgeschlüsselt nach TCP/UDP-Portnummern für die volumenstärksten Ports. Ein- und ausgehender Verkehr werden summiert. Zur besseren Übersicht werden hier lediglich die Kurven der ersten drei Ports gezeichnet. Eine komplette Sammlung von detaillierteren Graphen sind auf unseren Internetseiten [6] zu finden.

Abb.2 veranschaulicht den gegenwärtigen Trend im Internet hin zu einer rapiden Zunahme von P2P-Verkehr generiert durch Tauschbörsen. Port 1214 wird von Kazaa, Port 4662 von E-Donkey verwendet. Allein das Datenvolumen dieser beiden Ports, die nicht einmal den gesamten Verkehr der genannten Tauschbörsen erfassen, liegt oft über dem Volumen von HTTP. Vorläufige Auswertungen einer späte-

Leipzig-20021104-180000 Applications - Bandwidth

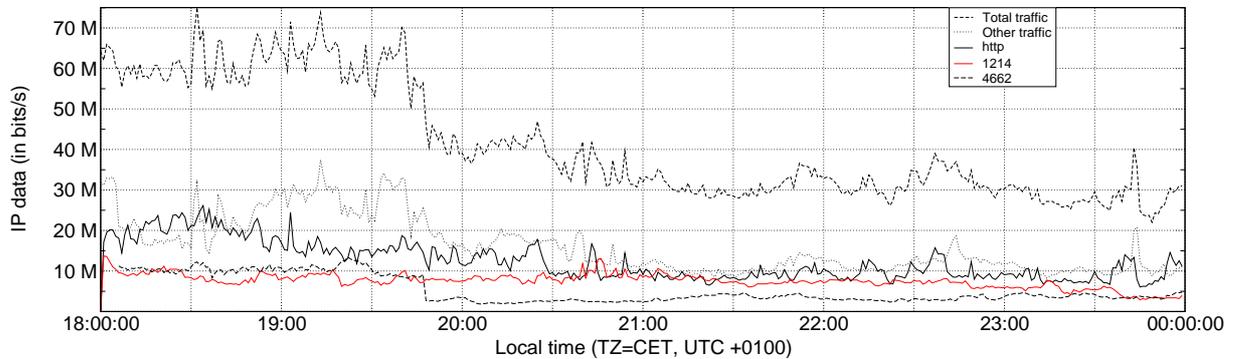


Abbildung 2: 1-Minuten-Mittelwerte der Bandbreite für die drei aufkommenstärksten Applikationen

ren Messung, die im nächsten Abschnitt beschrieben wird, unterstreichen diesen Trend.

6 Leipzig-II

Der Leipzig-II-Datensatz ist eine Zweipunkt-Messung durchgeführt vom 21.2. bis 22.2.2003. Es wurden ca. 2,3 Mrd. Paket-Header in Trace-Dateien von 165 GByte (69 GByte komprimiert) Volumen aufgezeichnet.

Wie bereits erwähnt, lässt sich zusätzlich zu den im Abschnitt 5 vorgestellten Analysen auch die Verzögerung der Pakete durch den Router ermitteln. Alles, was einem Paket im Router widerfahren kann, ist Verzögerung oder Verlust. Beides wird von unserer Zweipunkt-Messung registriert. Es ist also möglich, das genaue Verhalten des Routers auf Paketebene zu beobachten.

Abb.3 zeigt ein Diagramm mit Verzögerungszeiten auf der x-Achse (negative Werte repräsentieren ausgehende, positive Werte eingehende Pakete) und Paketlängen auf der y-Achse für ca. 2,5 Mio. Pakete. Aus dieser Darstellung lassen sich interessante Erkenntnisse über das Verhalten des Routers gewinnen.

Auffälligstes Merkmal dieser Darstellung ist die ausgeprägte V-Form. Die Innenseiten der

Schenkel repräsentieren die minimal beobachteten Verzögerungszeiten für jede Paketlänge. Zu einem Teil resultiert der Anstieg dieser Schenkel aus den zur Paketlänge proportionalen Serialisierungszeiten. Um dies zu veranschaulichen, wurden die Linien für die Serialisierungszeiten entsprechend der Bandbreiten auf der Außen- bzw. Innenseite des Routers (155,53 MBit/s bzw. 1000 MBit/s) eingezeichnet. Die 155,53 MBit/s-Linie muss dabei auf der Seite mit den positiven Verzögerungswerten

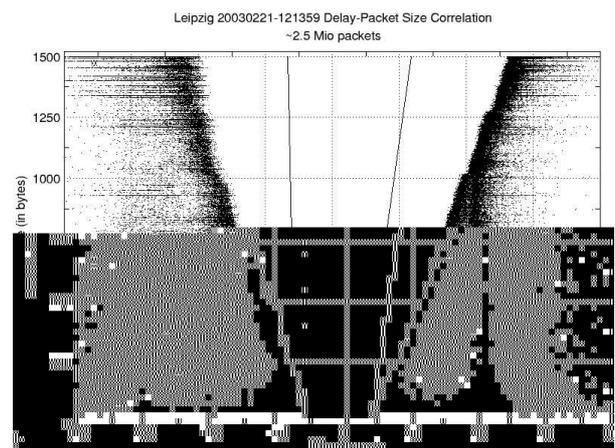


Abbildung 3: Paketverzögerung durch den Router im Verhältnis zur Paketlänge

ten (eingehender Verkehr) gezeichnet werden. Grund dafür ist, dass die Dag-Zeitstempel den Beginn eines Pakets markieren und die Serialisierungszeit demzufolge am Eingangsport des Routers beobachtet wird. Entsprechend wurde die 1000 MBit/s-Linie auf der negativen Seite mit den die Universität verlassenden Paketen gezeichnet.

Der Betrag des Anstiegs dieser Linien ist deutlich größer als der der V-Schenkel. Das bedeutet, dass zusätzlich zur Serialisierungszeit eine weitere paketlängenabhängige Verzögerungskomponente vom Router hinzugefügt wird. Das bedeutet, dass der Router nicht in der Lage ist, den Verkehr mit der vollen Bandbreite der jeweiligen Verbindung weiterzuleiten. Für eingehenden Verkehr beträgt die tatsächlich erzielte Bandbreite des Routers approximiert durch den Anstieg des rechten Schenkels ca. 54 MBit/s. Für ausgehenden Verkehr liegt der Wert bei ca. 82 MBit/s. Die enorme Differenz zwischen theoretischen und tatsächlich erzielten Werten macht deutlich, wie stark der Router überlastet ist.

Als weiteres Merkmal der Grafik fällt die Stufenform der Schenkel auf. Die Verzögerungszeit erhöht sich nicht gleichmäßig mit der Paketlänge, sondern springt bei allen ganzzahligen Vielfachen der Paketlänge von 512 Byte. Dieses Phänomen ist für beide Übertragungsrichtungen erkennbar. Vermutlich wird dieser Effekt durch die interne Organisation des Cisco 7505-Routers hervorgerufen.

7 Ausblick

Die beiden bisher in Leipzig durchgeführten Messungen verdeutlichen den Anstieg im Verkehrsaufkommen von Internet-Tauschbörsen. Basierend auf den uns vorliegenden Messdaten werden wir eine genaue Analyse dieses Verkehrs vornehmen – welches genaue Volumen er hat und wie dadurch klassische Applikationen

wie HTTP oder SMTP aber auch Echtzeitanwendungen zur Audio- und Videoübertragung beeinflusst werden.

Weiterhin arbeiten wir gegenwärtig an einem System zum intelligenten Management von Peer-to-Peer-Verkehr. Damit wollen wir Netzwerkadministratoren ein Werkzeug zur Verfügung stellen, das es ermöglicht, steuernd in den P2P-Verkehr einzugreifen. Dabei soll es nicht um ein vollständiges Blockieren sondern vielmehr um eine ressourcen- und bedarfsgerechte Zuteilung von Bandbreite gehen.

Im Laufe dieses Jahres wird der G-WiN-Zugang der Universität Leipzig von OC3c auf OC12c (622 MBit/s) umgestellt werden. Im Rahmen dieser Umstellung wird auch ein leistungsfähigerer Router zum Einsatz kommen. Mit Vergleichsmessungen werden wir ermitteln, welchen Veränderungen im Verhalten des Datenstroms sich ergeben. Besonderes Augenmerk gilt dabei dem Router, dessen Leistungsfähigkeit in Bezug auf die im Abschnitt 6 diskutierten Aspekte wir genauestens analysieren werden.

Alle bisher beschriebenen Analysen funktionieren nach dem gleichen Schema. Zuerst werden Daten in Form von Trace-Dateien gesammelt. Im Anschluss an eine Aufzeichnung beginnt die eigentliche Analyse. Bedingt durch die dabei anfallenden großen Datenmengen ergibt sich ein hoher Aufwand bei den Auswertungen. Um die in den vorigen Abschnitten präsentierten Grafiken für einen kompletten Trace von ca. 100 GByte Datenvolumen zu erzeugen, ist mit Laufzeiten der Analysetools von mehreren Stunden zu rechnen. Daraus ergeben sich signifikante Einschränkungen für potenzielle Einsatzmöglichkeiten eines solchen passiven Messsystems.

Ein wichtiges Ziel unserer weiteren Arbeit ist daher die Evaluation der Echtzeitfähigkeiten passiver Netzwerkanalysen. Die Analysen sollen nicht wie bisher nach dem Aufzeichnen eines Traces erfolgen, sondern unmittel-

bar während der Aufzeichnung. Je nach Bandbreite der beobachteten Netzwerkverbindung muss dabei eine Filterung bzw. Aggregation der anfallenden Daten vorgenommen werden. Daraus sollen dann in Echtzeit Informationen über das Verhalten des eventuell vorgefilterten Datenstroms bereitgestellt werden. Dieses System soll zur Netzwerkdimensionierung eingesetzt werden und dabei Netzwerkadministratoren bei ihren Entscheidungen unterstützen. Darüber hinaus ist auch der Einsatz in der Anomalie-Erkennung denkbar.

Literatur

- [1] Klaus Mochalski, Klaus Irmischer: On the Use of Passive Network Measurements for Modeling the Internet, in Irmischer, K., Fähnrich, K.-P. (Hrsg.), Tagungsband zur 13. Fachtagung Kommunikation in Verteilten Systemen KiVS 2003, Leipzig, 25.-28.2.2002
- [2] Jörg Micheel, Ian Graham and Stephen Donnelly: Precision Timestamping of Network Packets, Proceedings of the ACM SIGCOMM Internet Measurement Workshop, San Francisco, California, USA, November 1st/2nd 2001
- [3] Klaus Mochalski, Jörg Micheel and Stephen Donnelly: Packet Delay and Loss at the Auckland Internet Access Path, Proceedings of the PAM2002 Passive and Active Measurement Workshop, Fort Collins, Colorado, USA, March, 25-26th, 2002
- [4] WAND Group, Waikato University, Hamilton, Neuseeland: <http://wand.cs.waikato.ac.nz/wand/wits/>
- [5] Waikato Internet Traffic Storage: <http://wand.cs.waikato.ac.nz/wand/wits/>
- [6] Leipzig Trace Archive: <http://rnvs.informatik.uni-leipzig.de/traces/>
- [7] Web sites of Dag development at the University of Waikato and Endace Measurement Systems: <http://dag.cs.waikato.ac.nz> and <http://www.endace.com>
- [8] Trimble Navigation Limited: <http://www.trimble.com>