Network Analysis Times

Vol. 1 (1) - premiere issue *January, 2000*

Welcome to the Network Analysis Times

To meet the challenges posed by the rapid growth and expansion of the Internet, NLANR's measurement and network analysis activities are building and operating a network analysis infrastructure, with a primary focus on the NSF HPC sites. This is not to give all-encompassing answers, but rather, to gain more insight and knowledge of the inner workings of the Internet, thereby improving performance, and maximizing it as a resource.

We would like to use the *Network Analysis Times* to communicate our work, encourage the use of our data, and to solicit even more collaboration from the community. We believe that there are tremendous opportunities for more insight into issues such as a better understanding of Internet workload profiles, service characteristics and performance, models, metrics, and simulation.

We look forward to your feedback, ideas, and participation.

Hans-Werner Braun,
Principal Investigator,
NLANR Measurement and
Network Analysis activities

Thank you for your comments, questions and suggestions:
natimes@nlanr.net

In this issue:

This electronic newsletter is available on-line at: http://moat.nlanr.net/NATimes

We are happy to consider submissions for articles relating to network analysis research; for more information, please email: natimes@nlanr.net

The NLANR Measurement and Network Analysis Group is located at the San Diego Supercomputer Center (SDSC), University of California, San Diego (UCSD). This work is supported by the National Science Foundation (NSF) (grant no. ANI-9807479). Any opinions, findings and conclusions or recommendations expressed in this publication are those of the author(s), and do not necessarily reflect the views of the NSF.

NLANR
NATIONAL LABORATORY FOR APPLIED NETWORK RESEARCH

Introduction: The Role of Network Analysis

by Hans-Werner Braun, Principal Investigator, NLANR Measurement and Network Analysis activities

In today's environment, while the scope of the Internet is dramatically expanding in terms of global ubiquity, performance and degree of interconnections, the understanding of this increasing complexity and degree of manageability at the systemic level is diminishing. The services of the network today are assessed with perceptions and probabilities, from the fairly pure mathematical science of assessing earlier and simpler networks, we are moving closer and closer to the unpredictabilities and complexity of a biologic organism. As a result, more and more Internet service providers fight public information about user perceptions that put their services into a bad light regarding reported performances - defending themselves that their component network is in good shape, and that the problem must be somewhere else. In the same sense, participants in the global Internet are often inconsiderate of a common fate-sharing model where the union of ISPs has to provide and prove predictable and verifiable performances.

A quick verification of global routing tables in http://moat.nlanr.net/ASPL/ shows that commonly the distance between BGP-advertised Internet entities traverses 4 or 5 (or more) Autonomous System numbers, roughly mapable to service providers. This illustrates that from a user's point of view, the singular service parameters of ISPs are important, but typically not reflective of the end-to-end performance the user can expect.

In response to these issues, NLANR's measurement and network analysis activity is building and operating a network analysis infrastructure, with a primary focus on the NSF HPC sites. An analysis infrastructure has the objective to operate beyond the means of network measurements, as it is not good enough to just create large data silos containing measurement data that is hardly ever examined or studied. The definitions of analysis objectives have to be critically evaluated on an ongoing basis, and have to be flexible enough to adapt to new and extended requirements.

We focus on four areas for measurement data, which include: passive header trace data, active performance measurements, data collected from network entities via SNMP, and data related to stabilities and status of Internet routing. Currently, our primary focus is on our Passive Measurements and Analysis (PMA), and the Active Measurement Project (AMP), although we have some work going on related to Internet routing, as well.

We make all data, that we reasonably can, available on a regular basis, via the moat.nlanr.net Web and ftp server, resulting in gigabytes of data per day. We would like to use the Network Analysis Times to communicate our work, encourage the use of our data, and to solicit even more collaboration from the community. We believe that there are tremendous opportunities for more insight into issues such as, a better understanding of Internet workload profiles, service characteristics and performance, models, metrics, and simulation. Please feel free to provide feedback via our Web interface at http://moat.nlanr.net/.

We are starting a network analysis research "Idea Bank" to hold ideas for potential research topics using our data. There is a tremendous amount of data and nearly as many possibilities as to what can be done with it. We look forward to sharing ideas.

We focus on four areas for measurement data, which include: passive header trace data, active performance measurements, data collected from network entities via SNMP, and data related to stabilities and status of Internet routing.

We make all data, that we reasonably can, available on a regular basis, via the moat.nlanr.net Web and ftp server, resulting in gigabytes of data per day.

We believe that there are tremendous opportunities for more insight into issues such as, a better understanding of Internet workload profiles, service characteristics and performance, models, metrics, and simulation.

A Brief Overview of the Active Measurement Project (AMP)

Researcher: A. J. (Tony) McGregor, AMP Manager, NLANR Measurement and Network Analysis Group

The Active Measurement Project (AMP) is an important part of the Network Analysis Infrastructure (NAI) that is being developed by the NLANR Measurement and Network Analysis Group. The focus of AMP is site to site performance measurements taken across the National Science Foundation (NSF) approved, High Performance Connection (HPC) networks; these measurements include: round trip time (RTT), packet loss, topology and throughput. The first AMP monitors were deployed in December, 1998. In just over a year, this project has grown into a large network of remote machines (approximately 100) across the United States; all machines are owned by NLANR, and maintained and operated by the Measurement and Network Analysis Group.

The project began early in 1998, with a suggestion by Tony McGregor, a Department of Computer Science faculty member at the University of Waikato (Hamilton, New Zealand) to student, Stele Martin. Stele wrote the original code and a few Web pages that did analysis; modified versions of some of Stele's original code are still in use. They began with one monitor that pinged (contacted) other universities' networks.

In July, 1998, Tony began a six-month sabbatical as a visiting researcher with the NLANR program at SDSC, where he worked with Hans-Werner Braun and the rest of the Measurement and Analysis Group. Tony got the performance measurements going; he wrote the code for *am_master* and *am_slave* (an "active mirror" program which distributes changes in a directory at a master site (or sites) to slave sites as they happen). He further developed AMP by adding support for multiple monitors, a separate data server, real-time transfer of the data from the monitors to the server, and more sophisticated data analysis.

Although he returned to New Zealand and resumed teaching at Waikato University in January, 1999, Tony continues to work with NLANR, taking the lead for the AMP project. He attends weekly staff meetings (by phone) and works closely with the Measurement and Analysis Group on all AMP activities.

Some of this work has been developed by students working with NLANR, including Todd Hansen, Ryan Kassel and Neil Cotofana. In the following article in this newsletter, Todd Hansen details how AMP processes the data collected.

For more information,

Tony McGregor: http://www.moat.nlanr.net/~tonym/

AMP: www.moat.nlanr.net/AMP

AMP Data Delivery, Collection, and Archiving

Researcher: Todd Hansen, NLANR Measurement and Network Analysis Group

Approximately 100 Active Measurement Project (AMP) monitors are located across the U.S.; each measurement box (known as an "AMPlet") probes the network. AMPlets collect measurement data and store it locally for five days. An independent program (called am_master) running on the AMPlets, is responsible for sending the data back to the data collectors (central servers located at SDSC). Am_master connects to the am_slave program on a central data collector (either AMP, or VOLT); the connection is open for the life of the program, or until the system reboots. Every five minutes during normal processing, am_master sends updates of the files. Only the changes of each file are sent (so that the whole file does not have to be resent each time).

When the program first starts up, or when it reconnects to am_slave after a network failure, it goes through an initialization process, during which it synchronizes its knowledge of the files on the data collector. That way it knows what changes need to be sent. This synchronization involves several steps:

1. Am_master sends the file length and a 32-bit cyclic redundancy check (CRC) for every file which has been modified in the previous three days.

2. Am_slave receives this information and compares it with the file it already has. If the file size on the central data collector (am_slave) is less than the size sent by am_master (on the AMPlet), then it will acknowledge the smaller file size. If the file size is bigger than the file size sent by am_master, it will delete the local file and acknowledge it as having 0 file size (which will result in the file being sent from scratch because something must have changed).

3. If the file sizes match, then the system will do a local CRC calculation on the file, and if they match again, it will acknowledge the length of the file. If they do not match, it will acknowledge 0 and the file will be resent.

4. After am_master has verified every file from the previous three days, it can start sending the updates to the files on the central data collector. Therefore, no files will be resent unnecessarily.

This initialization process is repeated every 23 hours (+0-60 minutes, random) to verify the integrity of the files sent. This also allows us to catch files that may have changed internally but not externally, such as when a new vBNS.list file of the same size is installed. By redoing this initialization every 24 hours, we ensure that a checksum occurs on each stagnant file twice before it is ignored and subsequently deleted. This allows us to verify that we have all of the correct information from the AMPlet.

In the event that the network connection goes down, or we have to turn off a data collector (central server), up to 3 days worth of data can be automatically recovered. Using a slight modification, we can recover 5 days of data. (We can automatically recover a weekend's worth of data in an outage.)

The two data collectors (AMP and VOLT) are configured to collect the same data from all of the amp boxes; each of the collectors does this independent of everything else (and each other). The algorithm being used ensures that the two data collectors remain in sync with each other and the entire AMP constellation (all of the AMPlets). Therefore, we can recover lost data in the event of a data collector failure, (such as a bad hard disk). It would take about 2 days to copy the data from one system to the other, but it is feasible to do so within the time-frame before the AMPlets reach their maximum storage capability.

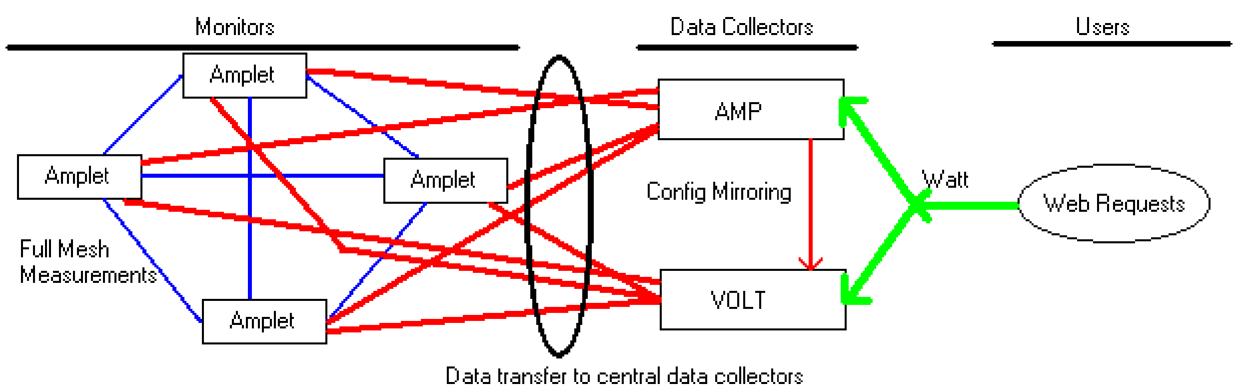
One data collector at a time serves as our main system, (where we make all of our changes and software updates). This system is then mirrored to the alternate system every night. Although we have two systems, they have the exact same content and Web page cgi's, so you can access them the same way. This allows us to put one domain name (watt.nlanr.net) which points to both AMP and VOLT. If one machine were to fail, users would automatically be shifted to the other system, without even noticing. This allows for minor system maintenance without any perceived downtime. In the event of a major outage, we might have to delay receiving new data for up to 2.5 days simply to resynchronize the systems. Up-to-the-minute access to the data would not be available at that time; however, no data would be lost in the long term. The end result is that under normal conditions, the measurement data on the two collectors is available within five minutes of its collection.

For more information,

Todd Hansen: http://www.moat.nlanr.net/~tshansen/

AMP information and current server:

http://www.moat.nlanr.net/AMP/



Data flow schematic by Todd Hansen

Pings from around the world -

notes on some of NLANR's collaborative projects in Network Analysis

Research for the Defense Research and Engineering Network (DREN)

Researchers: Phil Dykstra and Cindy Dykstra, WareOnEarth

WareOnEarth opened a new office in San Diego, CA, in October 1999. Led by Phil Dykstra, Chief Scientist, and Cindy Dykstra, Technical Director, they are focused on performance analysis, technology assessment, and security for the Defense Research and Engineering Network (DREN).

The DREN is part of the Department of Defense's (DOD's) High Performance Computing Modernization Program (HPCMP). It connects approximately 60 DOD R&D facilities in the continental United States, Alaska, and Hawaii. The DREN includes Internet Protocol (IP) and Asynchronous Transfer Mode (ATM) services from 10 Mbps through OC3, with OC12 expected in early 2000. The DREN peers directly with other federal, academic, and commercial high performance networks to facilitate collaboration with scientists at non-DOD facilities.

WareOnEarth plans to deploy a mesh of AMP measurement machines on the DREN and will be collaborating with the NLANR group on the issues of hierarchy and interface between multiple measurement clouds. They are particularly concerned about throughput for high-end applications and will be attempting to use AMP and/or OCxmon machines to test network-throughput capability. They have been evaluating the ability of those machines to make such measurements.

GigE Jumbo Frame - why you should care: Given the explosive growth of Gigabit Ethernet in campuses, gigapops, and network access points (NAPs), and because large maximum transmission unit (MTU) is so important to wide-area TCP performance, Phil has created a document about the importance of large MTU (specifically GigE jumbo frames) in high performance wide-area networks (WANs) (please see below for the URL). As of early 2000, there have been over 300 downloads. Phil wants to help get the word out "before we all put in 1500 byte bottlenecks."

Phil Dykstra came to WareOnEarth from the Army Research Laboratory where he co-chaired the Joint Engineering Team (JET) that focused on collaborations between Federal Networks, the Next Generation Internet, and Internet2. Cindy Dykstra was the Software Engineering Manager at Northrup Grumman's California Microwave Systems Division before joining WareOnEarth.

For more information,

Phil Dykstra: phil@wareonearth.com)

Cindy Dykstra: cindyd@wareonearth.com

WareOnEarth, San Diego: http://sd.wareonearth.com/

Joint Engineering Team (JET): www.ccic.gov/jet

DOD's High Performance Computing Modernization Program (HPCMP): www.hpcmo.hpc.mil

GigE Jumbo Frame: http://sd.wareonearth.com/~phil/jumbo.html

Inter-Autonomous Systems' (AS) Traffic Patterns and Their Implications

Researcher: Wenjia Fang, Computer Science Dept., Princeton University

In early 1999, Wenjia Fang, a doctoral candidate in the Computer Science Department at Princeton University, used NLANR Measurement and Analysis Group's passive monitoring infrastructure in her research on inter-autonomous systems' (AS) traffic patterns. She presented this work at Globecom '99, in Rio, Brazil.

For more information,

Wenjia Fang:

www.cs.princeton.edu/~wfang

full text of article:

www.cs.princeton.edu/~wfang/papers.html

(The abstract of this paper is available in the publications section of this newsletter, p. 13.)

How large round trip times (RTTs) affect Internet-2 applications

Researchers: Hank Nussbacher and Oded Comay,
Israel InterUniversity Computation Center

The InterUniversity Computation Center (IUCC) is a consortium of all eight Israeli universities. Part of its facilities are located inside the campus of Tel Aviv University. The IUCC operates Internet-2 in Israel; they maintain a T3 satellite link from Tel-Aviv University to StarTap and an E3 fiber link to QUANTUM in Europe.

The main purpose of their data collection and analysis is to see how large round trip times (RTTs), > 560ms, affect Internet-2 applications. They are planning on installing a set of TCP-spoofing boxes to allow faster than 936kb TCP streams to travel via satellites. In December, they installed an NLANR Measurement and Analysis Group passive monitor, which will be used to determine what effect these black boxes have on the total throughput of Internet-2.

For more information,

Hank Nussbacher: hank@interall.co.il

Internet-2 in Israel: http://www.internet-2.org.il

Satellite testing: http://www.internet-2.org.il/satellite-testing.html

Analysis of vBNS performance

Researcher: Ronn Ritke, Computer Science Department
University of California Los Angeles (UCLA)

Ronn Ritke is currently working with NLANR/SDSC and NSF to help provide answers on vBNS performance for the Presidential Information Technology Advisory Committee (PITAC). "What does a typical researcher get from the vBNS?" is an example of the kind of questions he hopes to answer.

To approximate end-to-end performance, one metric he is using is the bits/sec throughput for TCP connections using the vBNS. He is also looking at which well-known port number applications (http, FTP, etc.) are present in the vBNS traces, and what percentage of the total packets and bytes they represent. The approximation of end-to-end vBNS performance is a collaborative effort between Ronn and Hans-Werner Braun (NLANR/SDSC), it is anticipated that a paper will be written on the results, and submitted for publication.

At UCLA, Ronn coauthored a proposal entitled "High-Speed Network Measurements: Techniques and Tools"; NSF awarded them a one-year, $342,000, grant. The grant, which was based on Ronn's research and contacts, started January 1, 2000; if successful, the grant could be extended for an additional year. This NSF project will involve collaboration between UCLA, NLANR, and SDSC.

For more information,

Ronn Ritke: http://www.cs.ucla.edu/~ritke/

Presidential Information Technology Advisory Committee

(PITAC): http://www.ccic.gov/ac/

Passive Measurements and Monitoring of Internet Traffic

Researcher: Brynjar Aage Viken, Department of Telematics, Norwegian University of Science and Technology

Brynjar Aage Viken, a doctoral candidate in the Department of Telematics at the Norwegian University of Science and Technology, was a visiting researcher with the NLANR Measurement and Analysis Group from September 98 to February 99. His work on passive monitoring and passive measurements collected during Supercomputing '98, was presented at EUNICE '99 (Barcelona, Spain). Later in the fall, he enjoyed the opportunity to present this work to a Norwegian audience at 'The Norwegian Computer Science Conference.' Brynjar also plans to use passive measurements (packet traces) in parts of his doctoral thesis. (The abstract of this paper is available in the publications section of this newsletter, p. 13.)

For more information,

Brynjar Aage Viken: http://www.item.ntnu.no/~brynjar/

Norwegian Computer Science Conference:

http://www.ifi.ntnu.no/nik99/

“Passive measurement and analysis provide a better understanding of workload profiles. This in turn, allows for better network engineering and the creation of advanced Internet technology.”

Hans-Werner Braun

Passive Header Trace Measurements and Analysis (PMA): an overview

Researcher: Hans-Werner Braun, Principal Investigator, NLANR Measurement and Network Analysis activities

The Internet Protocol (IP) packet header trace collection and analysis is based on work done by Hans-Werner Braun and Kimberly Claffy in the early 1990’s, (eventually published as *"A Parameterizable Methodology for Internet Traffic Flow Profiling"* in the March 1995 IEEE JSAC Special Issue on the Global Internet, which examined traffic flows and workload profiles). Later, MCI vBNS staff augmented the work by evolving the code further and adding an OC3 ATM link interface, which resulted in the original OC3mon machines. Jon Dugan of NCSA did initial work to adapt the OC3mon software to a FreeBSD environment, an updated version of which we still use. Initially, we looked into using OC12 interface cards from a company in Chicago, but are now using cards, both ATM and Packet over Sonet capable, from the University of Waikato in New Zealand.

Passive measurements are performed in a non-invasive manner relative to the observed networking environment. The measurements do not impact the performance of the network while they are being done. The monitors tap into the light of a fiber interconnection by means of optical splitters, and collect packet header traces. The Web interface to our environment, data, and analysis results can be found at the Web site listed below.

Packet header traces can generate an immense amount of data, and, unless the desire is to keep the whole data set, it is usually critical to abstract the data locally, as quickly and as thoroughly as possible; however, this process loses information contained in the full traces. Abstractions can occur at a central data collection location, or, if possible, at the location where the data is being collected (to avoid transferring large volumes of data). Sometimes it is necessary to abstract already abstracted data further, such processing can take place at either the data generating/local site, or the central collection site.

If data abstraction happens at the location where the data is being collected, there are two ways to do it: in real time, or non-real time. Real-time abstractions allow for a continuous operation, but are limited by the capabilities of local CPU and memory resources to accommodate the analyses at the speed at which new data is arriving. This becomes an especially difficult issue when complex analyses for in-depth understanding of the networking environment are desired, or in high traffic load situations. In such cases the local collection-analysis machine turns into a bottleneck.

A non-real time analysis assumes staggered data collection and data analysis phases, presumably followed by a phase where the analyzed data is transferred to a central location, and the data collection resumes. However, even in this case, highly complex analyses can take a very long time. This methodology reflects a reality sampled in time, as the measurement phases are interleaved with the analysis phases. This is the mode we are using for the PMA data collection and analysis.

We are currently operating 15 data collectors (14 in the U.S., and one in Israel). An agreement between our research group and every collaborator is that we pay attention only to the headers and disregard packet contents. We have adopted the term “PMA monitors” to encompass the superset of OC3, OC12, DS3, FDDI, and Ethernet, etc. measurement machines. We expect to deploy about 25 more machines, in collaboration with the Internet2 community, over the coming months.

For more information,

PMA Web interface, data, analysis results:

http://www.moat.nlanr.net/PMA/

Raw packet header trace data repository: Access statistics for HTTP and FTP

Raw data access is provided for both HTTP and FTP interfaces. Based on our logs, during the 9/1998 to 12/1999 time frame, about 7500 trace files with 125 GB of data were accessed via the FTP interface; during the 1/1998 to 1/2000 time frame about 43000 trace files with 284 GB of data were accessed via the Web interface. Details are below.

HTTP:

	accesses:	volume:
1998-Jan	64	907 MB
1998-Feb	117	3445 MB
1998-Mar	373	8750 MB
1998-Apr	186	1705 MB
1998-May	2205	3256 MB
1998-Jun	2062	3473 MB
1998-Jul	4768	6540 MB
1998-Aug	941	2505 MB
1998-Sep	2130	3441 MB
1998-Oct	1030	2164 MB
1998-Nov	9336	90575 MB
1998-Dec	2352	12013 MB
1999-Jan	4510	1512 MB
1999-Feb	651	5189 MB
1999-Mar	861	5109 MB
1999-Apr	2583	23689 MB
1999-May	4840	56170 MB
1999-Jun	648	1382 MB
1999-Jul	624	4631 MB
1999-Aug	235	2465 MB
1999-Sep	243	898 MB
1999-Oct	804	9520 MB
1999-Nov	818	16943 MB
1999-Dec	692	7839 MB
2000-Jan	784	10846 MB
Total	**43857**	**284 GB**

FTP:

	accesses:	volume:
1998-Oct	17	1024 MB
1998-Nov	75	2517 MB
1998-Dec	796	10015 MB
1999-Jan	1740	11083 MB
1999-Feb	331	6601 MB
1999-Mar	822	25405 MB
1999-Apr	1177	17840 MB
1999-May	488	5374 MB
1999-Jun	149	4251 MB
1999-Jul	190	2400 MB
1999-Aug	265	4968 MB
1999-Sep	259	4381 MB
1999-Oct	307	4706 MB
1999-Nov	441	9551 MB
1999-Dec	427	11087 MB
Total	**7527**	**125 GB**

NLANR
NATIONAL LABORATORY FOR APPLIED NETWORK RESEARCH

Data Archiving for Passive Monitors

Researcher: Hans-Werner Braun, Principal Investigator, NLANR Measurement and Network Analysis activities

In the Passive Measurement and Analysis (PMA) program, the monitors collect traces onto one or more dedicated data hard drives, which, in the case of the initially deployed OC3 monitors is 4GB, and eventually transfer their data to a central server located at the San Diego Supercomputer Center. The machines do multiple kinds of analyses that are also sent to San Diego; previous day's traces get posted overnight. The data and analyses are available at the Web sites listed below.

Currently, about 2GB worth of compressed traces (there is an approximate 1:3 compression ratio) are posted daily. This means that we run short of disk space after a while. We use 12GB DDS-3 tapes for long-term storage and the UNIX tar command to write traces directly to the tape. A single tape takes hours to write or read as it holds about five days worth of trace data.

We have the capability to write data to a compact disk, but those can hold only around 650MB each. We have tried DVD RAM, but the media incompatibility between a DVD RAM drive and a common DVD reader does not make it an attractive solution even though the DVD RAM we tried allows for about 2.5GB per media surface.

For more information,

PMA data and analysis:

http://moat.nlanr.net/Datacube

PMA overnight traces:

http://moat.nlanr.net/Traces

Sample data: Sat., 6 Nov 1999, 23:55:06 -0800
Experimental Daily Max Throughput Summary from Packet Header Traces

Rank	Mbps Throughput	Site	Application Triplet	Bytes	Duration
001	17.956114	TXS	6:1351:2855	198344924	88.368752
002	13.967173	12SDC	17:47550:9875	153920	0.088161
...					
099	3.000765	OSU	6:20:34773	3083105	8.219515
100	2.998889	ODU	6:80:2221	114137	0.304478

The Cichlid 3-D Visualization System

Researcher: Jeff Brown, NLANR Measurement and Network Analysis Group

Originally intended as a single demonstration program, the Cichlid visualization system has been developed into its current state as a general purpose visualization tool that can transform raw data into great images and animation. Jeff Brown created Cichlid in response to a request from PI Hans-Werner Braun for something to "visually demonstrate" at Supercomputing Conference (SC) 1998.

The OpenGL based Cichlid visualization package consists of two closely linked sets of programs: servers and clients.

Cichlid-servers gather and analyze data and listen for requests from Cichlid-clients; Cichlid-clients request data from these servers and display them graphically for the user. Currently, the user interface (UI) is very rudimentary; improvements are forthcoming.

Cichlid visualizes data it receives from servers. In order to visualize data using Cichlid, server code must be written to make the data available. Usually, this involves writing code that produces and abstracts the data in which you're interested, then linking it with Cichlid's server library, (OpenGL is not necessary for the server). The code processes and produces the data, then the Cichlid server library code passes it out to a connected Cichlid-client. Unlike traditional tools which visualize local, static data sets (e.g., Excel, gnuplot), Cichlid requires writing code to do so. Cichlid was not created for visualizing static data, but specifically for continuous data import by the visualization client from the data server.

The power of Cichlid comes from the ability of a local 3-D manipulation of result data, while continuously importing additional data, shown as an animation, in real time. Rotating the animating data set, zooming, panning, as well as displaying multiple real-time animated data sets, all at once, is easily done in Cichlid. Cichlid adds minimal additional complexity for the visualization developer, based on the need to create a Cichlid server, and no additional effort on the client site, while producing pleasing and compelling visualizations. Cichlid visualization is portable, and will work with virtually any UNIX or Windows workstation that supports OpenGL.

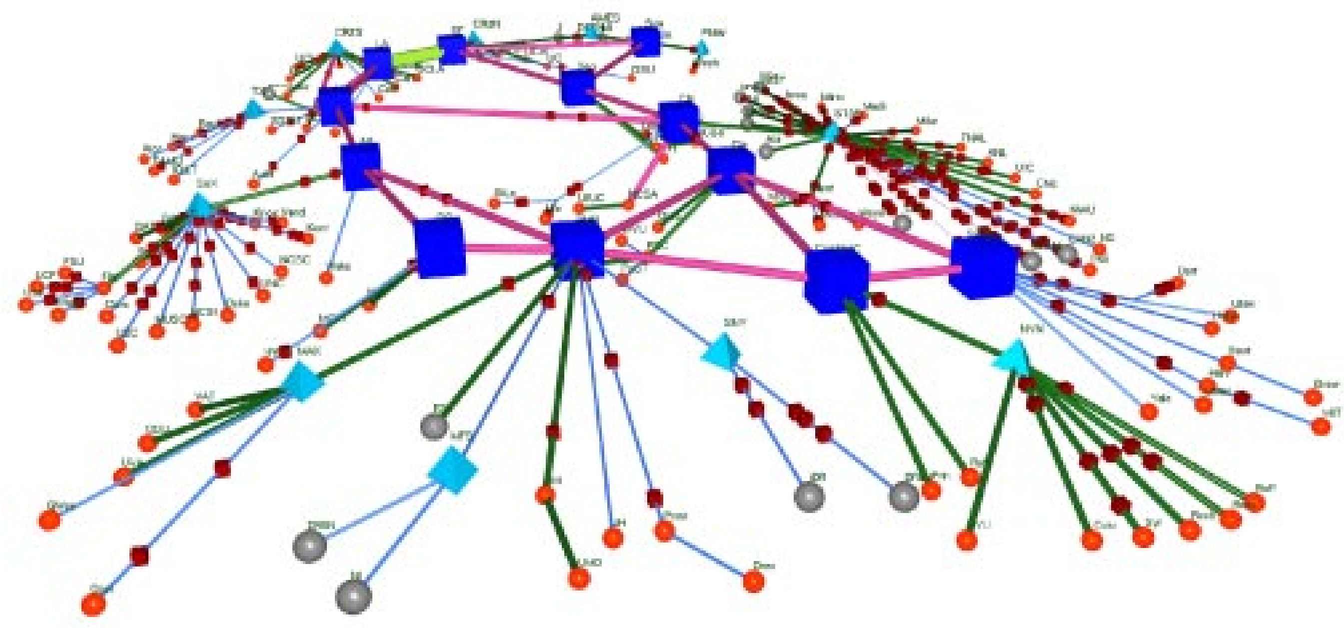
For more information,

Jeff Brown: jabrown@nlanr.net

Cichlid:

http://moat.nlanr.net/Software/Cichlid/

http://moat.nlanr.net/Software/Cichlid/gallery/gallery.html



logical vBNS, Cichlid screenshot, Jeff Brown.

BGP Network Routing Analysis

Researcher: Neil Cotofana, NLANR Measurement and Network Analysis Group

Routing instability over the Internet contributes to traffic congestion and, in some cases, impairs connectivity. BGP4 is the most widely used exterior gateway protocol (BGP) and the method by which the Internet knows which networks can and cannot be reached. Neil Cotofana researches data derived from BGP peering sessions in order to answer the following questions:

1) why is there routing instability,
2) how does it happen, and
3) how much of it is related to how BGP4 itself was designed?

The University of Oregon's Route Views project and the BGP peering session between moat.nlanr.net and a router on the vBNS, provide the raw data for Neil's work. He creates 2-D graphs using Perl scripts to analyze the data, and 3-D visualizations using Cichlid for analysis. His focus is the size of the Internet Routing Table (IRT) and the distribution of prefix lengths within the routing tables of different Internet Service Providers (ISPs). From these graphs, Neil has found that the /24 prefixes, which make up the largest group of prefixes in the IRT (approximately 30,000), offer reachability to a relatively small number of hosts. Neil has also experimented with graphs depicting the redundancy of /24 prefixes as seen across the routing tables of the core routers participating in the Route Views project.

For more information,

Neil Cotofana: http://www.moat.nlanr.net/~neil/

2-D graphs, using Perl:

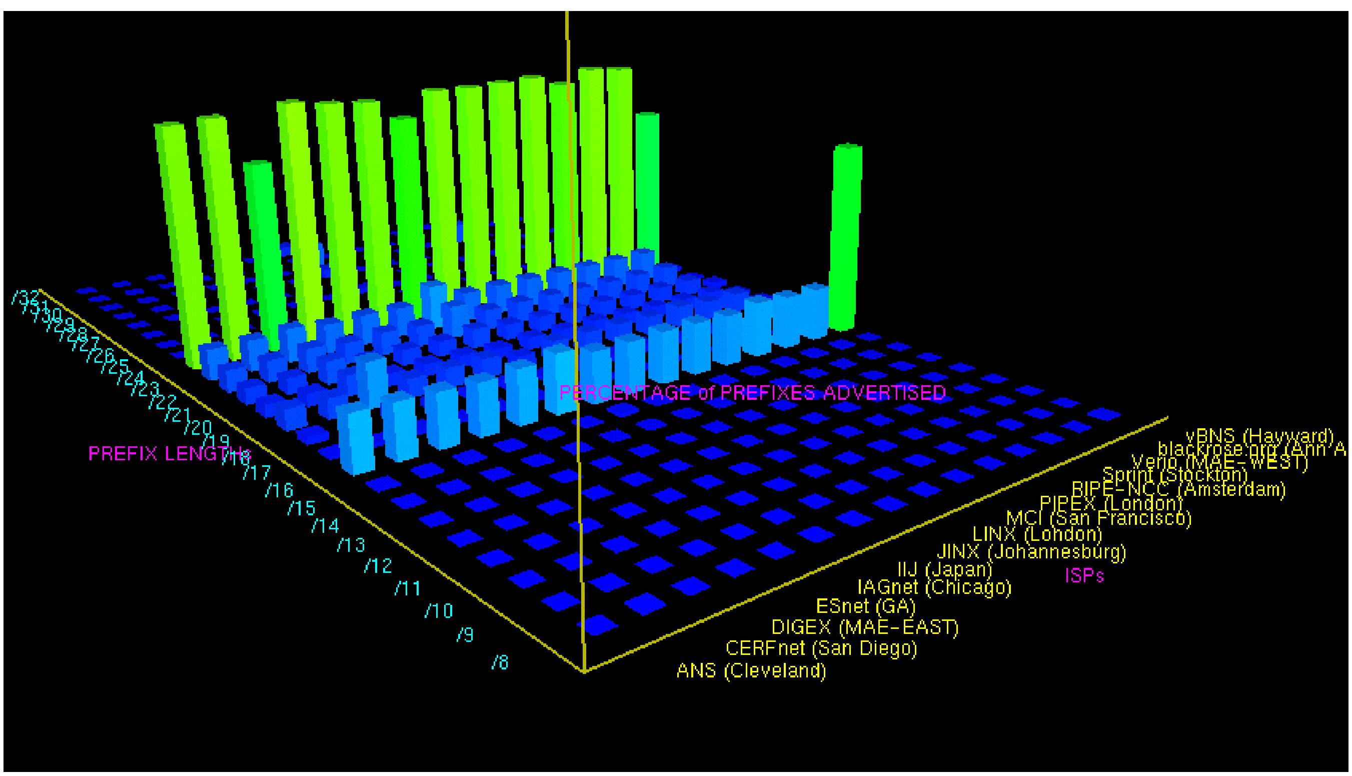
http://moat.nlanr.net/~neil/bars.html

3-D visualizations, using Cichlid:

http://moat.nlanr.net/~neil/cichlid.html

/24 prefix redundancy graphs:

http://moat.nlanr.net/~neil/24.html



Screenshot of Cichlid interactive 3-D graph (prefix percentages); data from Oregon Route Views project.

“The real synthesis of measurement information into an integrated services suite and user-driven interactive tools has just begun”

An Interview with Don Mitchell, National Science Foundation(NSF) Project Manager for NLANR

Meiling Guentzel, NLANR Measurement and Network Analysis Group

How do the NLANR Measurement and Network Analysis activities differ from other NSF funded projects?

“The NLANR measurement and network analysis effort is not a stand-alone project but an integrated part of a service suite (including engineering and user services) which helps to ensure that institutions and researchers using the advanced networking infrastructure that ANIR supports are able to optimize that use.”

Why is the measurement aspect of the program important?

“Measurement is really a fundamental building block upon which service is built. The effectiveness of network and applications ‘tuning’ efforts ultimately depends on being able to identify what problems exist in the network, where they are, and observing the effect of changes. That's what network traffic measurement and analysis is all about.”

Given a rock collection, a live turkey, and a chocolate éclair, which item do you think would most directly relate to your experiences with the Measurement and Network Analysis Group, and why?

“I would have to go with two of the three choices: the rock collection and the chocolate éclair. The rock collection because the OCx-monitors (measurement machines) have been rock-solid tools that have continued to percolate along since their introduction into the measurement and analysis program, and the chocolate éclair because sweet things have turned up during our work together.”

Have there been any surprises along the way? If so, how do you feel they will affect the future of measurement and network analysis?

“The Active Measurements Project (AMP) work was not originally anticipated. It may well add a new and better understanding of the larger high performance network system and how well it functions by allowing us to inject variants into the traffic and measure ‘end-to-end’ stress. The emphasis on ‘end-to-end’ implies a ‘user-centric’ rather than ‘network centric’ perspective. We're beginning to take the feedback which network measurement can provide and incorporate it into tools and processes to improve engineering and applications-related service. The real synthesis of measurement information into an integrated services suite and user-driven interactive tools has just begun.”

"In today's Internet, with the phenomenal growth that is being experienced, the ‘practice’ of networking has outrun the ‘theories’ on the basis of which advanced networks are designed, built, and operated.

Network traffic monitoring and measurement is the critical technology needed now to allow us to see and understand what is actually happening within and among networks.

This, in turn, will help both network researchers and operators to identify and mitigate nascent problems before they become infrastructure failures. It may even, in the longer run, allow us to develop reasonable predictive models to anticipate problems before they arise."

Aubrey Bush,
National Science Foundation Division Director for Advanced Networking Infrastructure and Research (ANIR)

News Briefs

New Utility - TCPtune:

Todd Hansen has created a new utility to tune the Transmission Control Protocol (TCP) stack in the Windows (95, 98, and NT) operating system. He recently finished a new version, 1.1.1, which supports all of the Windows 32-bit operating systems (OS). It also has some new features that make it easier to use, including that it automatically detects the current OS version, and modifies the user interface (UI) for that particular OS.
For more information:
http://moat.nlanr.net/Software/TCPtune/

Analyzing the gated.log File:

Neil Cotofana continued work on scripts analyzing data from the gated.log file which is resulting from the BGP session with a vBNS node. The file, which logs all of the BGP transactions, is being made publicly available.
For more information: http://moat.nlanr.net/BGP/VBNS/

Quarterly and Annual Reports for 1999 are available for viewing: http://moat.nlanr.net/Reports/

Instructional Web Sites:

CD-ROM Burning:

To support the burning of CD-ROMs for uses such as distributing sets of measurement data, Jeff Brown has created a Web site, complete with software to make the process intuitive and easy.
For more information: http://moat.nlanr.net/CDBurning/

Passive Monitor Assembly:

David Cheney has documented the process of building a passive monitor, step by step. The instructions are such that anyone can build a similar machine with little or no knowledge of computer hardware and maintenance. To complement this site, Todd Hansen has produced a rescue/install floppy which ensures any new passive monitor will be installed with the current up to date image and software.
For more information:
http://moat.nlanr.net/~dgcheney/makePMA.htm

On Deck: Short- and Long-term Goals

- We are in the process of creating a database to track the details of the large number of machines that we operate; the database will also provide site status reports. Plans are underway to provide these reports on-line to detail causes for data outages and planned corrective actions. We also hope to maximize the automation of problem warnings, as we need to spend more time on our own research.
- Several more AMP and PMA machines should be deployed by the end of the quarter.
- We hope to get the Research Manager position filled. If you know someone with the necessary high level skills and experience, who would enjoy the San Diego area and the SDSC/UCSD environment, please direct them to: http://joblink.ucsd.edu/bulletin/prog_jobs.html#C4851

Publications & Presentations Using NLANR Measurement and Network Analysis Data

Note: citations followed by (*abs*) have abstracts following the publications list.

Inter-AS Traffic Patterns and Their Implications. Wenjia Fang and Larry Peterson (Princeton University). *Proceedings of Global Internet 99,* Rio de Janeiro, Brazil. (*abs*)

The NLANR Network Analysis Infrastructure. A.J. McGregor, H-W Braun, J.A. Brown. Accepted for publication by *IEEE Communications Magazine* for publication in the feature topic issue on "Network traffic measurements and experiments." (*abs*)

Network Performance - Visualization: Insight through animation. J.A. Brown and A.J. McGregor. PAM2000: Passive and Active Measurement Workshop. Hamilton New Zealand. Accepted, to be published April 2000.

Passive Measurements of Internet Traffic from SCinet '98. Brynjar Aage Viken. Norsk Informatikkonferanse, Trondheim, Norway, 16-17 November 1999.

Passive Monitoring of Internet Traffic at Supercomputing '98. Brynjar Aage Viken. EUNICE '99, Barcelona, Spain, 1-3 September 1999. (*abs*)

Proceedings of the "Challenges and Opportunities for Measurement and Analysis in a High Performance Computing Environment" Workshop. 1 July 1999; Mike Gannis and Todd Hansen.

Proceedings of the "Measurement and Analysis Collaborations" Workshop. 29-30 June 1999; Mike Gannis and Todd Hansen. Several printed copies of both proceedings have been made, and are available upon request.

The NLANR Network Analysis Infrastructure.

A.J. McGregor, H-W Braun, J.A. Brown.

ABSTRACT

The Measurement and Network Analysis group within NLANR is developing a Network Analysis Infrastructure (NAI). It is intended that this infrastructure will provide both engineering and research support for the HPC community. Specifically the goal of the NAI project is to create an infrastructure that will support measurements and analysis through the collection and publication of raw data, visualization and analysis of network measurement. This paper details the foundation and various aspects of the NAI, as well as discussing its history and current status.

Passive Monitoring of Internet Traffic at Supercomputing '98.

Brynjar Aage Viken

ABSTRACT

Passive measurement data is collected without any impact on the network by the measurement itself. OCxmon/Coral is a passive stand-alone monitor that collects packet traces from a high-capacity link by tapping a mall fraction of the lights with an optical splitter. SCinet '98 was the show floor network for Supercomputing '98 where the latest in systems, applications, and services for all area of high performance networking and computing were demonstrated. NLANR/MOAT monitored the traffic on the OC3 link connecting the conference network to the vBNS. The measurement data collected consist of 79.7 million packets, 34.3 million packets sent from SCinet '98 to the vBNS and 45.5 million packets sent from the vBNS to the SCinet '98. Based on these measurement data, the traffic composition for IP protocols, packet length distributions, and interarrival times are presented. Further, the NLANR/MOAT demonstration of distributed real-time 3-D visualization of abstracted OC3mon/Coral data, presented at Supercomputing '98, is briefly described.

Inter-AS Traffic Patterns and Their Implications.

Wenjia Fang and Larry Peterson

ABSTRACT

This paper reports on a study of traffic patterns among Autonomous Systems (ASes), based on traces taken at various points on the Internet. The traces display a highly non-uniform distribution of traffic on flows between pairs of hosts, networks, and ASes. Aggregation along coarser granularities, such as networks or Ases, accentuates this non-uniform distribution. In one typical trace, for example, the top 9% of flows between Ases accounts for 86.7% of the packets and 90.7% of the bytes transmitted. A highly non-uniform traffic pattern suggests that routers need to maintain only limited Quality of Service (QoS) flow state. The paper discusses the implications of this phenomenon on different proposed QoS mechanisms.

full text: http://www.cs.princeton.edu/~wfang/papers.html

Meet the NLANR Measurement and Network Analysis Group

The group conducts performance and flow measurements for High Performance Connection sites; we are located at the San Diego Supercomputer Center (SDSC), University of California, San Diego (UCSD).

All members of the group have contributed significantly to this issue of *Network Analysis Times*. The members are listed here, followed by their area of primary responsibility/interest.

Project lead:	Hans-Werner Braun	Principal Investigator, NLANR Measurement and Network Analysis activities
Manager:	Tony McGregor	AMP (Active Measurements Project) Manager
Full time staff:	Maureen C. Curran	Technical writing; Editor, *Network Analysis Times*
	Bud Hale	Systems Administrator
UCSD Students:		
	Jeff Brown	Cichlid development
	David Cheney	Passive measurements
	Neil Cotofana	BGP network routing analysis
	Meiling Guentzel	Technical writing; researched, wrote, and edited newsletter prototype, which became the *Network Analysis Times*
	Todd Hansen	Active measurements, coordination, TCPtune, Windows development
	Ryan Kassel	Active measurements
	José Otero	Analysis tasks; working with Todd on papers; in addition to handling technical aspects, contributed significantly to the formatting and design of the *Network Analysis Times*
SDSC staff:		
	Mike Gannis	External Relations staff; Public Relations and outreach
	Marjorie Hardy	Business Manager, SDSC; support
	Edna Nerona	External Relations staff; Web design/development; assisted in the formatting and design of the *Network Analysis Times*

NLANR
NATIONAL LABORATORY FOR APPLIED NETWORK RESEARCH