Measurement, Modeling and Analysis of the Internet

March, 2004

THE UNIVERSITY of WISCONSIN MADISON

Paul Barford
Assistant Professor
Computer Science
University of Wisconsin

Motivation – there's plenty!

- The Internet is a HUGE network of networks
  - Scientists love to study/model complex systems
    - Emergent characteristics
- Wide area network behavior is unpredictable
  - Many factors are pushing and pulling the infrastructure
  - Constant change is normal
- Many applications have minimum performance requirements
  - Reliability, predictability, …
- Network managers adjust systems to conditions

Tutorial goals

1. Present overview of the when, where, how, and why of Internet measurement and monitoring techniques
   – Particular focus on tools and infrastructure
2. Present overview of Internet measurement data analysis
   – Where we’ve been and where we are headed
3. Provide citations and pointers to Internet measurement resources
4. Stimulate discussion!

Tutorial outline

1. Network measurement overview
   – Challenges, tools and techniques
2. Internet measurement infrastructures
   – Today’s and tomorrow’s
3. Overview of methods of Internet data analysis
   – A lead into to part 2
4. Problems with network measurement work today
   – It’s a little grim…

Tutorial themes

- Measurement has been the basis for critical improvements
  - Without measurement, what do you know?
- Measurement capability in the Internet is limited
  - The systems not designed to support measurement
- Measurement tools and infrastructures are few and limited
  - Size, diversity, complexity and change
- Measurement data presents many challenges
  - Networking researchers need better connections with experts in other domains

Part 1: Network Measurement

Challenges

Successes

Tools and methods

Internet measurement challenges

- Size of the Internet
  - O(100M) hosts, O(1M) routers, O(10K) networks
- Complexity of the Internet
  - Components, protocols, applications, users
- Constant change is the norm
  - Web, e-commerce, peer-to-peer, wireless, next?
- The Internet was not developed with measurement as a fundamental feature
  - Nearly every network operator would like to keep most data on their network private
- Floyd and Paxson, “Difficulties in Simulating the Internet”, *IEEE/ACM Transactions on Networking, 2000.*

A small selection of past successes

- Leland *et al.*, "On the Self-Similar Nature of Ethernet Traffic", IEEE/ACM Transactions on Networking, 1994.
  - Thorough analysis of Bellcore LAN traces established *self-similar* properties of packet arrival process
- Cunha *et al.*, "Characteristics of WWW Client-based Traces", BU-TR 95-010, 1995.
  - Modeled a variety of WWW client use characteristics
- V. Paxson, "Measurement and Analysis of End-to-end Internet Dynamics", PhD. Thesis, 1997
  - Characterized routing and packet behavior in wide area

Past successes contd.

- Govindan and Reddy, "An Analysis of Inter-domain Topology and Route Stability", INFOCOM, 1997.
  – Establishes basic properties of inter-domain connectivity
- Caceres *et al.*, "Multicast-based Inference of Network-internal Loss Characteristics", IEEE Trans. On Info. Theory, 1999.
  – First use of *tomographic inference* to isolate packet loss
- Moore *et al*., "Inferring Internet Denial of Service Activity", USENIX Security Symposium, 2001.
  – First analysis of the extent of denial of service activity

Why do we measure the Internet?

- Some reasons have been presented already…
  - Basic component of scientific method
  - Operation
- We cannot improve the Internet if we don’t understand it’s structure and behavior
  - We cannot understand it if we don’t measure
  - We cannot build effective models or simulators if we don’t measure
- A long term objective – “a day in the life of the Internet”
  - NRC report: “Looking over the Fence at Networks, A Neighbors View of Networking Research”, 2001.

What can we measure in the Internet?

- Structure
  - Topology, routing, proxies, wireless, etc.
- Traffic
  - Transport, end-to-end performance, etc.
- Users and Applications
  - WWW, peer-to-peer, streaming, DNS, etc.
- Failures
  - In all areas
- Nefarious behavior
  - Pattern attacks, port scans

Where can we measure the Internet?

- For some measurements, this is obvious
  - For example, if you are studying the Web servers, Web logs are a good starting point
  - The goal for other measurements is to be "representative"
  - Various "Internet weather reports"
- Placement of measurement nodes is not a well understood problem
  - More is better??
- Where we *can't* measure is in commercial networks

How can we measure the Internet?

- Active methods
  - Probes, application simulation
- Passive methods
  - Application monitors (logs), system monitors, packet monitors
- Surveys (almost never)
- Significant infrastructure is **always** required
- All methods present difficulties
- Resources
  - SLAC: www.slac.stanford.edu/xorg/nmtf/nmtf-tools.html
  - CAIDA: www.caida.org/tools/taxonomy

When should we measure the Internet?

- Diurnal traffic cycle
- Time scales depend on “what” and “how”
- Passive measurements are typically continuous
  - Can generate **huge** data sets
  - Many people will not allow access to their logs
- Active measurements are typically discrete
  - Important characteristics can be missed
  - Probes can be filtered and/or detected

Who is measuring the Internet?

- Businesses do a great deal of measurement
  - What measurements are they taking and what do they do with their data?
- Instrumentation for measurement-based research is relatively new
  - Developments over the past 12 years have been slow
  - 10's of current studies
    - SLAC: www.slac.stanford.edu/comp/net/wan-mon/netmon.html
    - CAIDA: www.caida.org/analysis/performance/measinfra/
- Most studies are not coordinated and relatively narrowly focused

Active probes to study path properties

- Active probe tools send stimulus (packets) into the network and then measure the response
  - Network (IP), transport (UDP/TCP), application layer probes
- Active probes can measure many things
  - Delay/loss
  - Topology/routing behavior
  - Bandwidth/throughput
- Oldest examples of probe tools use Internet Control Message Protocol (ICMP)
  - Network layer probe

Simple delay/loss probing with *ping*

Simplest request/response probe tool using ICMP Echo capability

C:\WINDOWS\Desktop>ping www.soi.wide.ad.jp

Pinging asari.soi.wide.ad.jp [203.178.137.88] with 32 bytes of data:

Reply from 203.178.137.88: bytes=32 time=253ms TTL=240
Reply from 203.178.137.88: bytes=32 time=231ms TTL=240
Reply from 203.178.137.88: bytes=32 time=225ms TTL=240
Reply from 203.178.137.88: bytes=32 time=214ms TTL=240

Ping statistics for 203.178.137.88:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 214ms, Maximum = 253ms, Average = 230ms

Routing behavior with *traceroute*

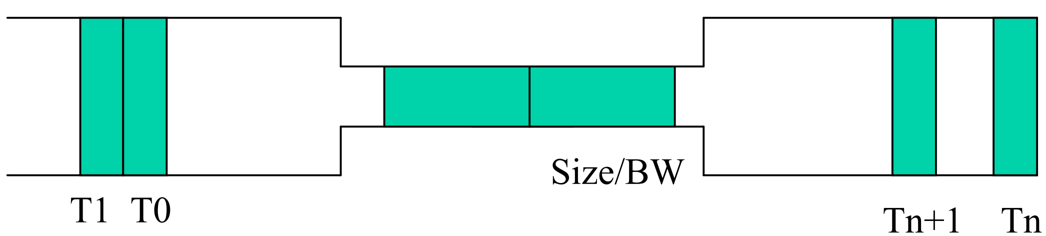
- Standard utility for assessing the route between hosts
- traceroute sends a series of probes to successive nodes along a route to an intended destination and records the source address of the message returned by each
- Operation
  - Routers decrement the "time to live" (TTL) field in IP pkts.
  - Router sends ICMP Time Exceeded message back to source if the TTL field is decremented to 0
  - If TTL starts at 5, source host will receive Time Exceeded message from router that is 5 hopes away
  - traceroute typically sends three probes to each hop and reports source address information and RTT for each probe

traceroute example

```
C:\windows\desktop> tracert www.soi.wide.ad.jp
Tracing route to asari.soi.wide.ad.jp [203.178.137.88]
over a maximum of 30 hops:
  1    19 ms    27 ms    23 ms  208.166.201.1
  2    17 ms    13 ms    14 ms  204.189.71.9
  3    25 ms    29 ms    29 ms  aar1-serial4-1-0-0.Minneapolismpn.cw.net [208.174.7.5]
  4    24 ms    27 ms    24 ms  acr1.Minneapolismpn.cw.net [208.174.2.61]
  5    26 ms    22 ms    23 ms  acr2-loopback.Chicagochd.cw.net [208.172.2.62]
  6    29 ms    29 ms    27 ms  cand-w-private-peering.Chicagochd.cw.net [208.172.1.222]
  7    28 ms    24 ms    28 ms  0.so-5-2-0.XL2.CHI2.ALTER.NET [152.63.68.6]
  8    26 ms    27 ms    28 ms  0.so-7-0-0.XR2.CHI2.ALTER.NET [152.63.67.134]
  9    25 ms    24 ms    26 ms  292.at-2-0-0.TR2.CHI4.ALTER.NET [152.63.64.234]
 10    73 ms    74 ms    73 ms  106.ATM7-0.TR2.LAX2.ALTER.NET [146.188.136.142]
 11    74 ms    76 ms    76 ms  198.ATM7-0.XR2.LAX4.ALTER.NET [146.188.249.5]
 12    73 ms    75 ms    77 ms  192.ATM5-0.GW9.LAX4.ALTER.NET [152.63.115.77]
 13    80 ms    73 ms    76 ms  kdd-gw.customer.ALTER.NET [157.130.226.14]
 14    84 ms    84 ms    91 ms  202.239.170.236
 15    97 ms    81 ms    86 ms  cisco1-eth-2-0.LosAngeles.wide.ad.jp [209.137.144.98]
 16   174 ms   174 ms   178 ms  cisco5.otemachi.wide.ad.jp [203.178.136.238]
 17   201 ms   196 ms   194 ms  cisco2.otemachi.wide.ad.jp [203.178.137.34]
 18   183 ms   182 ms   196 ms  foundry2.otemachi.wide.ad.jp [203.178.140.216]
 19   183 ms   185 ms   178 ms  gsr1.fujisawa.wide.ad.jp [203.178.138.252]
 20   213 ms   205 ms   201 ms  asari.soi.wide.ad.jp [203.178.137.88]
Trace complete.
```

Probing for link characteristics

- *Packet dispersion* techniques (Jacobson) can be used to infer characteristics of each link along an Internet path
  - Latency, bandwidth, and queue delays
  - Cross traffic causes problems
- Tools available: bprobe [CC97], clink [D99], nettimer [LB99], pathchar [J97], pchar [M00], pathrate [DRM01]
- A. Downey, "Using pathchar to Estimate Internet Link Characteristics," SIGCOMM, 1999.



Tn+1 - Tn = max(S/BW, T1 – T0)

clink output example

```
Probe for link bandwidths between Boston Univ. and Univ. Wisconsin
>clink pluto.cs.wisc.edu
```

8 probes at each of 93 sizes (28 to 1500 by 16)

0 localhost l n= 744 lat= 0.210 ms bw= (6.414, 6.411, 6.611) Mb/s
1 CS4NET-GW.BU.EDU (128.197.14.1) l n= 744 lat= -0.026 ms bw= (-387.989, -140.840, -136.152) Mb/s
2 crc-ext-gw.bu.edu (128.197.254.60) l n= 744 lat= 0.148 ms bw= (260.224, 346.367, 320.380) Mb/s
3 ATM10-410-OC12-GIGAPOPNE.NOX.ORG (192.5.89.13) l n= 744 lat= 2.556 ms bw= (493.574, 639.542, 23568.176) Mb/s
4 ABILENE-GIGAPOPNE.NOX.ORG (192.5.89.102) l n= 744 lat= 6.095 ms bw= (-1440.365, 705.495, 1438.433) Mb/s
5 clev-nycm.abilene.ucaid.edu (198.32.8.29) l n= 744 lat= 3.113 ms bw= (-748.522, 1502.420, 780.744) Mb/s
6 ipls-clev.abilene.ucaid.edu (198.32.8.25) l n= 744 lat= 4.243 ms bw= (-8.827, 29.665, 12998.206) Mb/s
7 r-peer-at-0-1-0-14.net.wisc.edu (144.92.20.137)
8 144.92.128.226 (144.92.128.226) l n= 744 lat= 0.449 ms bw= (-34.186, 23.717, 40.601) Mb/s
9 144.92.128.196 (144.92.128.196) l n= 744 lat= 0.626 ms bw= (-248.625, -37.351, -7.664) Mb/s
10 e1-2.foundry2.cs.wisc.edu (128.105.1.6) l n= 744 lat= -0.742 ms bw= (7.680, 18.018, 23.147) Mb/s
11 pluto.cs.wisc.edu (128.105.167.50)

n = number of probes, lat = latency (ms), bw = (low,best,high) bandwidth

Passive packet measurement

- Capture packet data as it passes by
  - Packet capture applications (tcpdump [JLM89]) on hosts use packet capture filter (bpf [MJ93], libpcap [MLJ94])
    - Requires access to the wire
      - Promiscuous mode network ports to see other traffic
  - Hardware-bases solutions
    - Endace, Inc.'s DAG cards – OC12/48/192 (0.622/2.5/10Gbps)
- Problems
  - Adds, deletes, reordering, timestamping
  - 10 Gbps Ethernet today, 40Gbps soon - LOTS of data!
  - Privacy issues

Example of *tcpdump* output

04:47:00.410393 sunlight.cs.du.edu.4882 > newbury.bu.edu.http: S 1616942532:1616942532(0) win 512 (ttl 64, id 47959) 04:47:03.409692 sunlight.cs.du.edu.4882 > newbury.bu.edu.http: S 1616942532:1616942532(0) win 32120 (ttl 64, id 47963) 04:47:03.489652 newbury.bu.edu.http > sunlight.cs.du.edu.4882: S 3389387880:3389387880(0) ack 1616942533 win 31744 (ttl 52, id 27319)

04:47:03.489652 sunlight.cs.du.edu.4882 > newbury.bu.edu.http: . ack 1 win 32120 (DF) (ttl 64, id 47964)

04:47:03.489652 sunlight.cs.du.edu.4882 > newbury.bu.edu.http: P 1:67(66) ack 1 win 32120 (DF) (ttl 64, id 47965) 04:47:03.579607 newbury.bu.edu.http > sunlight.cs.du.edu.4882: . ack 67 win 31744 (DF) (ttl 52, id 27469)

04:47:04.249539 newbury.bu.edu.http > sunlight.cs.du.edu.4882: . 1:1461(1460) ack 67 win 31744 (DF) (ttl 52, id 28879) 04:47:04.249539 newbury.bu.edu.http > sunlight.cs.du.edu.4882: . 1461:2921(1460) ack 67 win 31744 (DF) (ttl 52, id 28880)

04:47:04.259534 sunlight.cs.du.edu.4882 > newbury.bu.edu.http: . ack 2921 win 32120 (DF) (ttl 64, id 47968) 04:47:04.349489 newbury.bu.edu.http > sunlight.cs.du.edu.4882: P 2921:4097(1176) ack 67 win 31744 (DF) (ttl 52, id 29032)

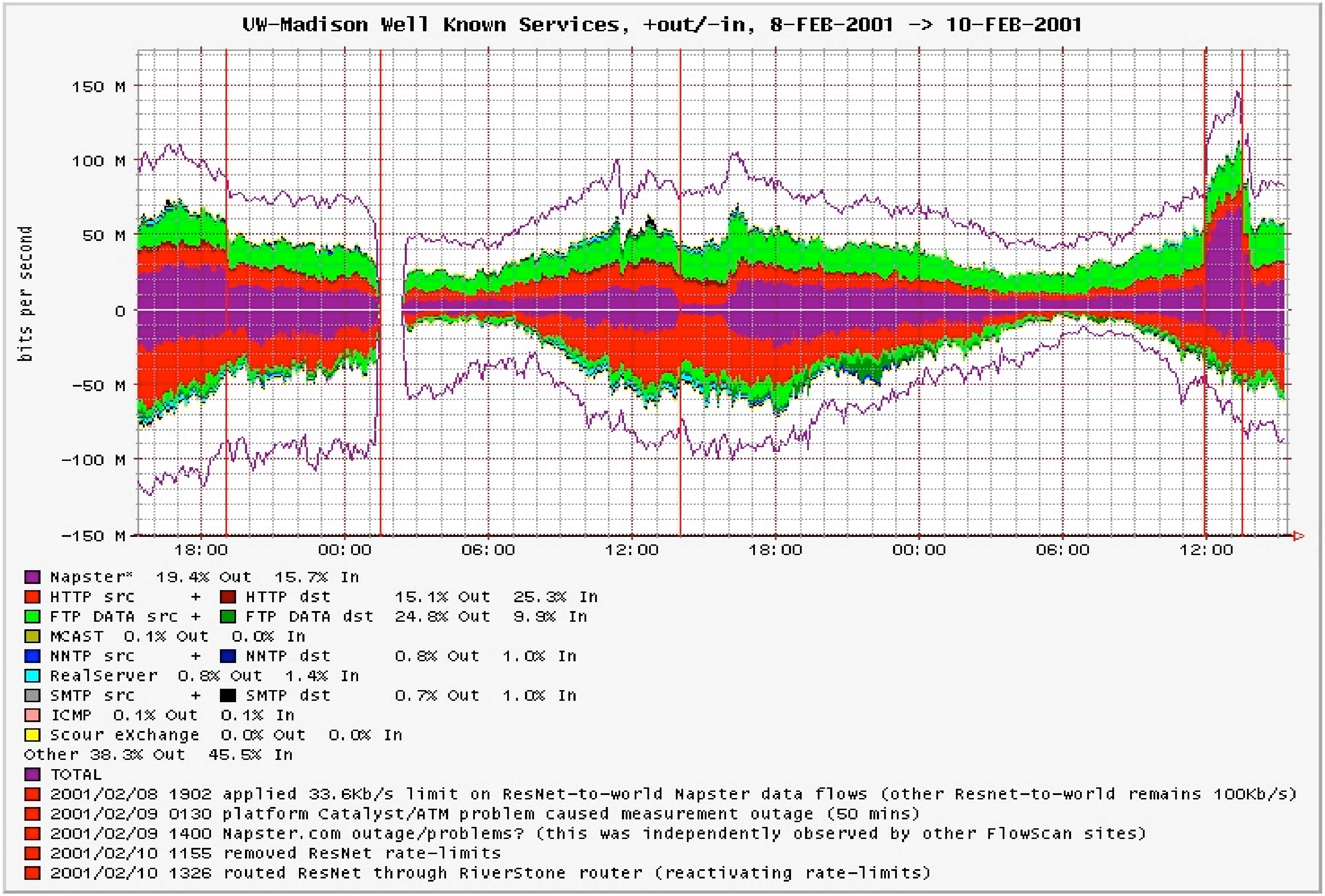
04:47:04.349489 newbury.bu.edu.http > sunlight.cs.du.edu.4882: . 4097:5557(1460) ack 67 win 31744 (ttl 52, id 29033)

Passive IP flow measurement

- An IP Flow is defined as “a unidirectional series of packets between source/dest IP/port pair over a period of time”

```
{SRC_IP/Port,DST_IP/Port,Pkts,Bytes,Start/End Time,TCP Flags,IP Prot …}
```

  – Exported by applications such as Cisco’s NetFlow
- We use FlowScan [Plonka00] to collect and process Netflow data
  – Combines flow collection engine, database, visulaization tool
  – Provides a near real-time visualization of network traffic
  – Breaks down traffic into well known service or application


UW-Madison Well Known Services, +out/-in, 8-FEB-2001 -> 10-FEB-2001
bits per second
150 M
100 M
50 M
0
-50 M
-100 M
-150 M
18:00
00:00
06:00
12:00
18:00
00:00
06:00
12:00
Napster* 19.4% Out 15.7% In
HTTP src + HTTP dst 15.1% Out 25.3% In
FTP DATA src + FTP DATA dst 24.8% Out 9.9% In
MCAST 0.1% Out 0.0% In
NNTP src + NNTP dst 0.8% Out 1.0% In
RealServer 0.8% Out 1.4% In
SMTP src + SMTP dst 0.7% Out 1.0% In
ICMP 0.1% Out 0.1% In
Scour eXchange 0.0% Out 0.0% In
Other 38.3% Out 45.5% In
TOTAL
2001/02/08 1902 applied 33.6Kb/s limit on ResNet-to-world Napster data flows (other Resnet-to-world remains 100Kb/s)
2001/02/09 0130 platform Catalyst/ATM problem caused measurement outage (50 mins)
2001/02/09 1400 Napster.com outage/problems? (this was independently observed by other FlowScan sites)
2001/02/10 1155 removed ResNet rate-limits
2001/02/10 1326 routed ResNet through RiverStone router (reactivating rate-limits)

Passive monitoring for intrusions

- There are plenty of bad guys out there
  - Cracking tools are readily available
- Detecting attacks and nefarious behavior (*eg*. port scans) is critical for protecting networks
- Passive measurements of packet traffic can be used to reconstruct higher level behavior
  - Most traffic is unencrypted
- Network Intrusion Detection Systems (NIDS) use packet filters to observe network traffic
  - V. Paxson, "Bro: A System for Detecting Network Intruders in Real-time", Computer Networks, 1999.
  - M. Roesch, "Snort: Lightweight Intrusion Detection for Networks", LISA, 1999.

Network intrusion detection systems

- Signature-based NIDS
  - Generates alerts based on observations with known attacks
- Anomaly-based NIDS
  - Generates alerts based on observed deviations from established profile of normal behavior
- Activity-based NIDS
  - Generates alerts based on observed deviations from a site's security policy
- There are many commercial systems
- All systems suffer from false positives and negatives

Example of *Snort* output

Snort portscan log:

Mar 11 19:37:00 130.253.192.2:51217 -> 130.253.192.200:79 SYN ******S*
Mar 11 19:37:00 130.253.192.2:51218 -> 130.253.192.198:79 SYN ******S*
Mar 11 19:37:01 130.253.192.2:51219 -> 130.253.192.207:79 SYN ******S*
Mar 11 19:37:01 130.253.192.2:51220 -> 130.253.192.195:79 SYN ******S*
Mar 12 16:14:11 130.253.192.7:1023 -> 130.253.192.219:32825 UDP
Mar 12 16:14:13 130.253.192.7:2049 -> 130.253.192.198:57741 UDP
Mar 12 16:14:13 130.253.192.7:1023 -> 130.253.192.215:63715 UDP
Mar 12 16:14:13 130.253.192.7:1023 -> 130.253.192.218:61213 UDP

Snort alert log:

[**] [100:1:1] spp_portscan: PORTSCAN DETECTED from 130.253.192.2 (THRESHOLD 4 connections exceeded in 1 seconds) [**] 03/11-19:37:00.874491

[**] [1:1243:2] WEB-IIS ISAPI .ida attempt [**] [Classification: Web Application Attack]
[Priority: 1] 03/12-01:32:56.468227 211.172.179.3:1245 -> 130.253.192.161:80
TCP TTL:240 TOS:0x10 ID:0 IpLen:20 DgmLen:1420
AP Seq: 0x264615F1 Ack: 0xBFDF7245 Win: 0x7BFC TcpLen: 20
[Xref => http://www.whitehats.com/info/IDS552]
[Xref => http://cve.mitre.org/cgi-bin/cvename.cgi?name=CAN-2000-0071]

iSinks: an intrusion data gold mine

- Observations
  - A great deal of IPv4 address space is allocated (routed) but not used (no system connected)
  - the bad guys don't know which addresses in a network are used
  - Monitoring allocated but unused addresses (network telescopes) is a great source of intrusion data
- An *iSink* is a scalable tool for both active and passive measurement on unused address space
  - Includes active responders used to gather details of exploit
  - http://www.cs.wisc.edu/~pb/publications.html

Part 2: Measurement Infrastructures

Today

Future

Design of measurement infrastructures

- Architecture consists of physical systems and management/operation environment
- Physical systems
  - Measurement method will inform system selection
    - Extra hardware (eg. GPS) could be necessary
  - Deployment is typically based on what you can get more than what you would like to have
  - Even deploying a small number of systems is difficult
  - Maintaining systems is always underestimated

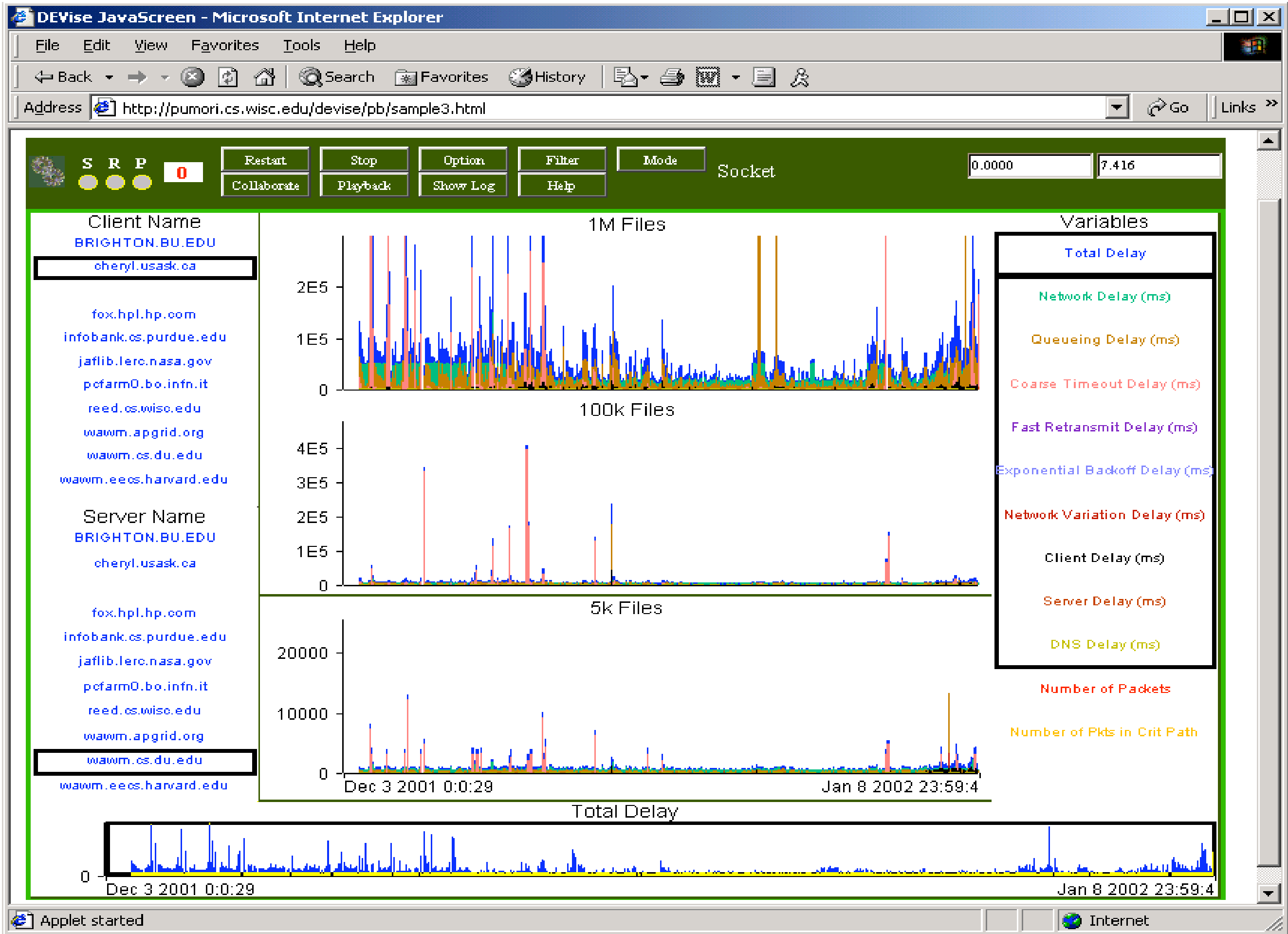
Operation and management systems

- Security is critical!!
  - Strong authentication is more important than encryption
  - Assume bad guys will break in and design system for quick reinstall
- Measurement scheduling system and method
  - Automated environment for scheduling and sometimes synchronizing measurements
  - Methods must consider things like synchronization in traffic
    - Use Poisson probing methods
- Data collection and archival system
  - Automated environment for collecting and storing results
  - Careful work in this area ALWAYS pays off in the end

Operation/management systems contd.

- Analysis and visualization systems
  - Standard scripts for evaluating data
  - Visualization of time series data is critical
  - Web front end
- Software deployment and maintenance
  - Standard distributions and management methods
    - Documentation and archives of configurations
  - In very large systems a PULL environment works better than PUSH

Rapid prototype visualization

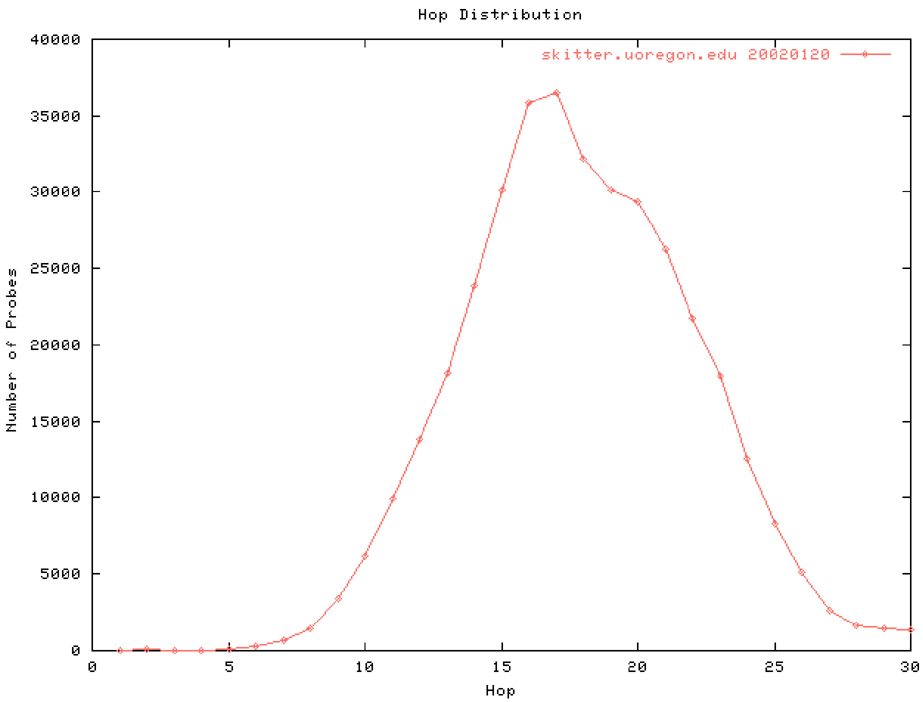


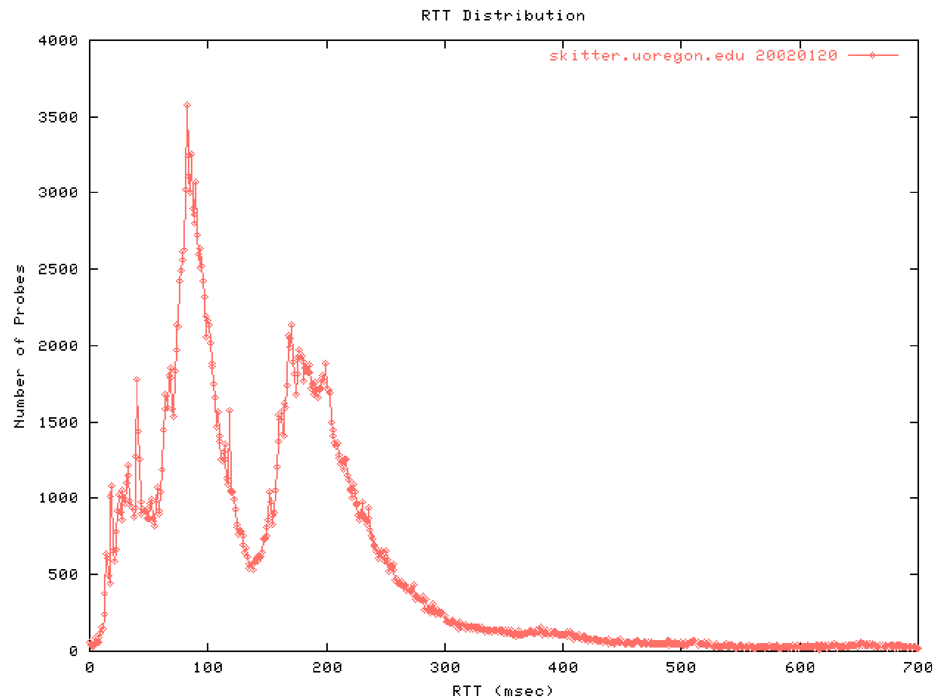
Yao et al., “DEVise and the Javascreen: Visualization on the Web”, SPIE, 2000.

The *Skitter* infrastructure

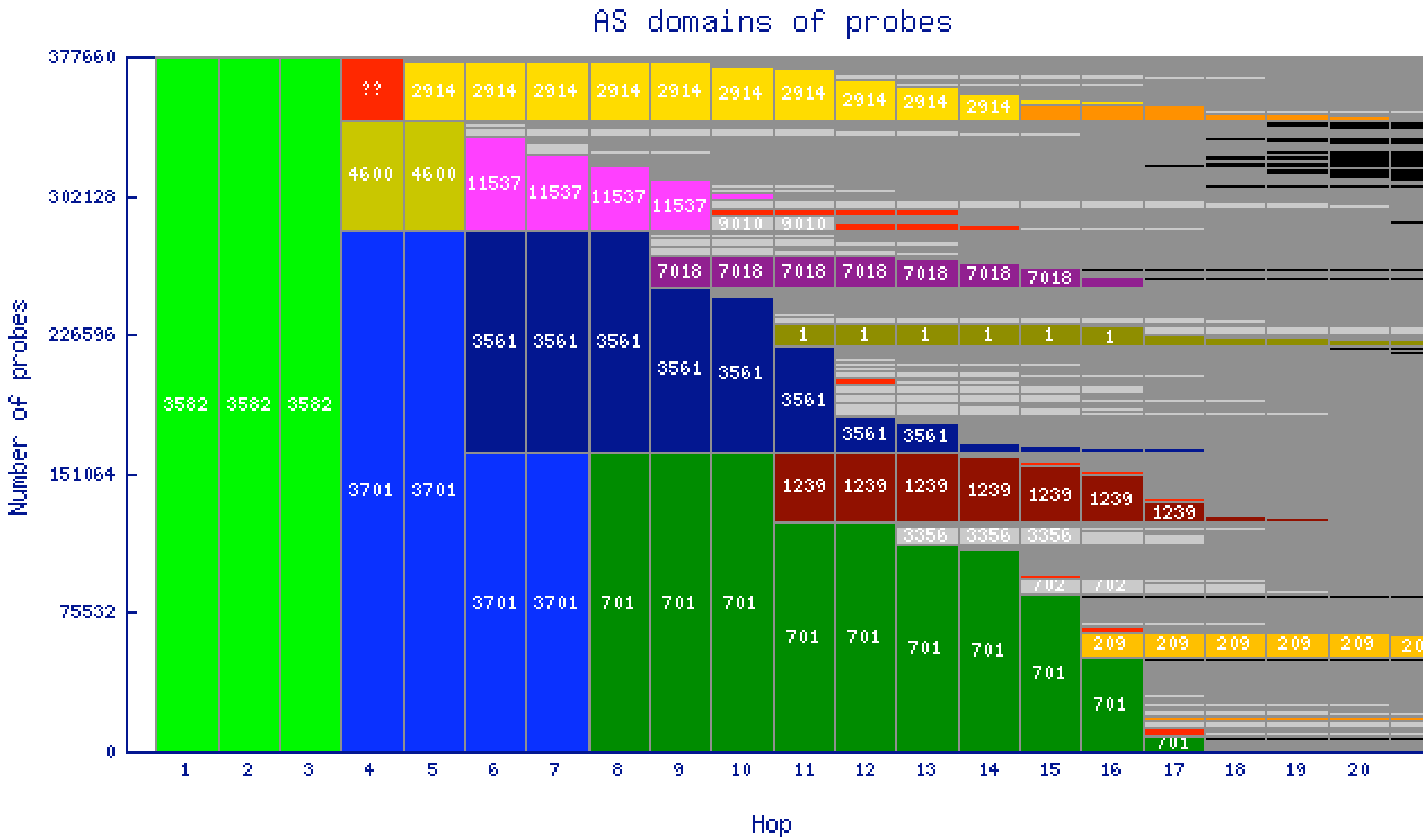
- CAIDA's Internet routing and topology measurement infrastructure
  - *traceroute* study focused on router/link discovery
    - Raw data available on request
    - Also trace packet delay and loss
- Infrastructure: 21 sources, ~500K destinations
  - World wide deployment
  - Not all destinations are reached by sources
- Methodology: sources *traceroute* to destinations 24x7
- Visualization is also a component of the project
- www.caida.org

Skitter daily summary

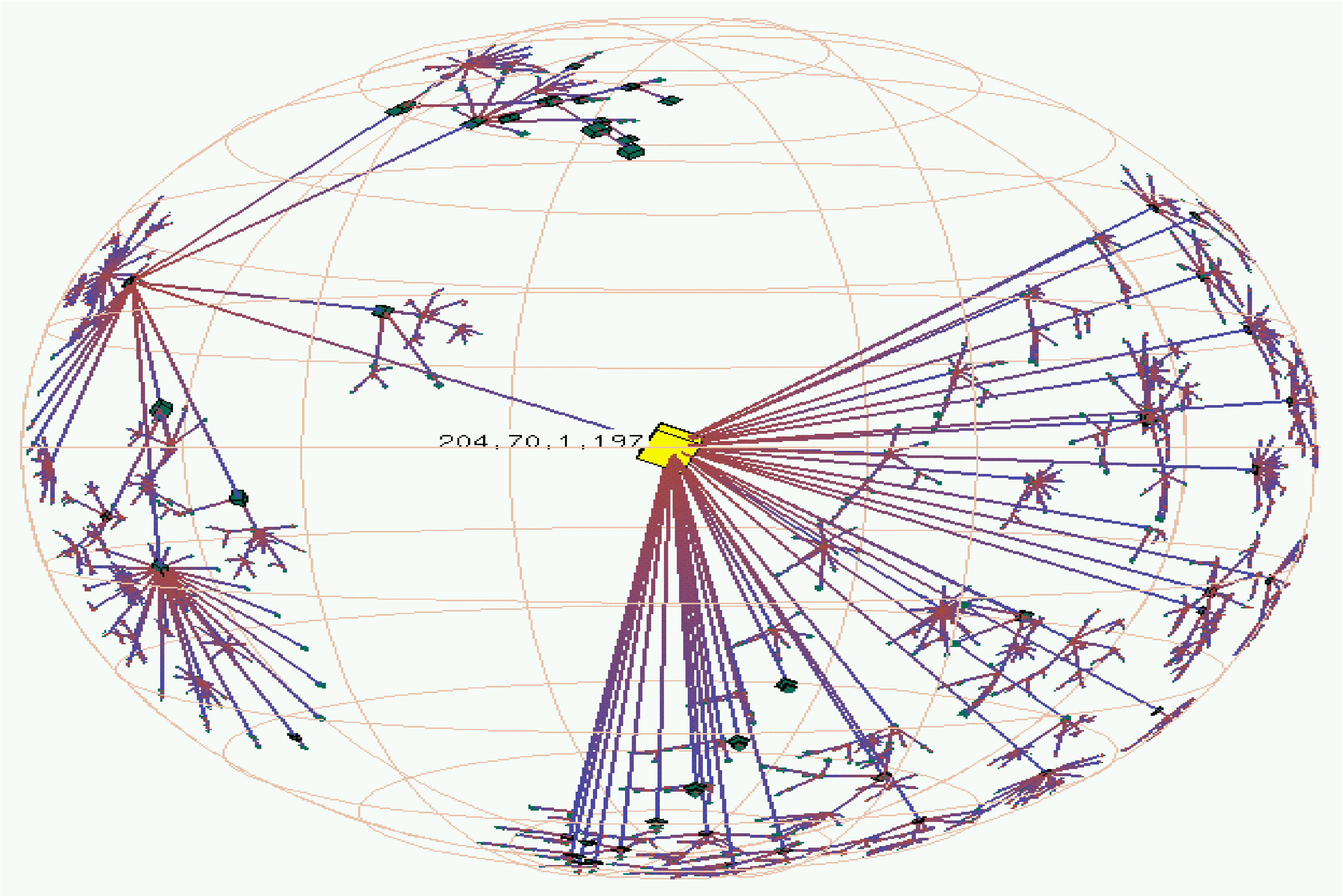




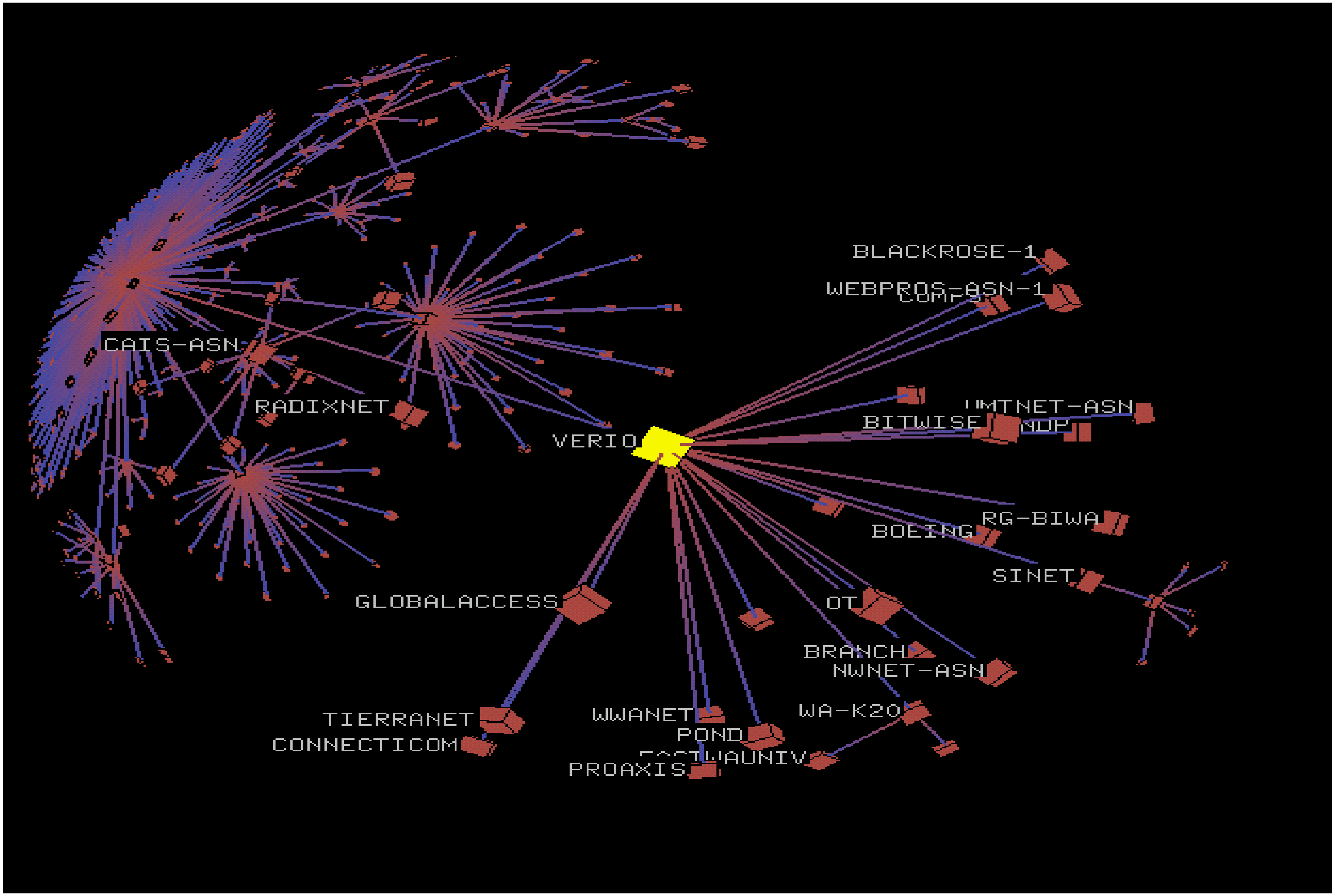
Skitter daily summary contd.



Skitter visualization – IP paths



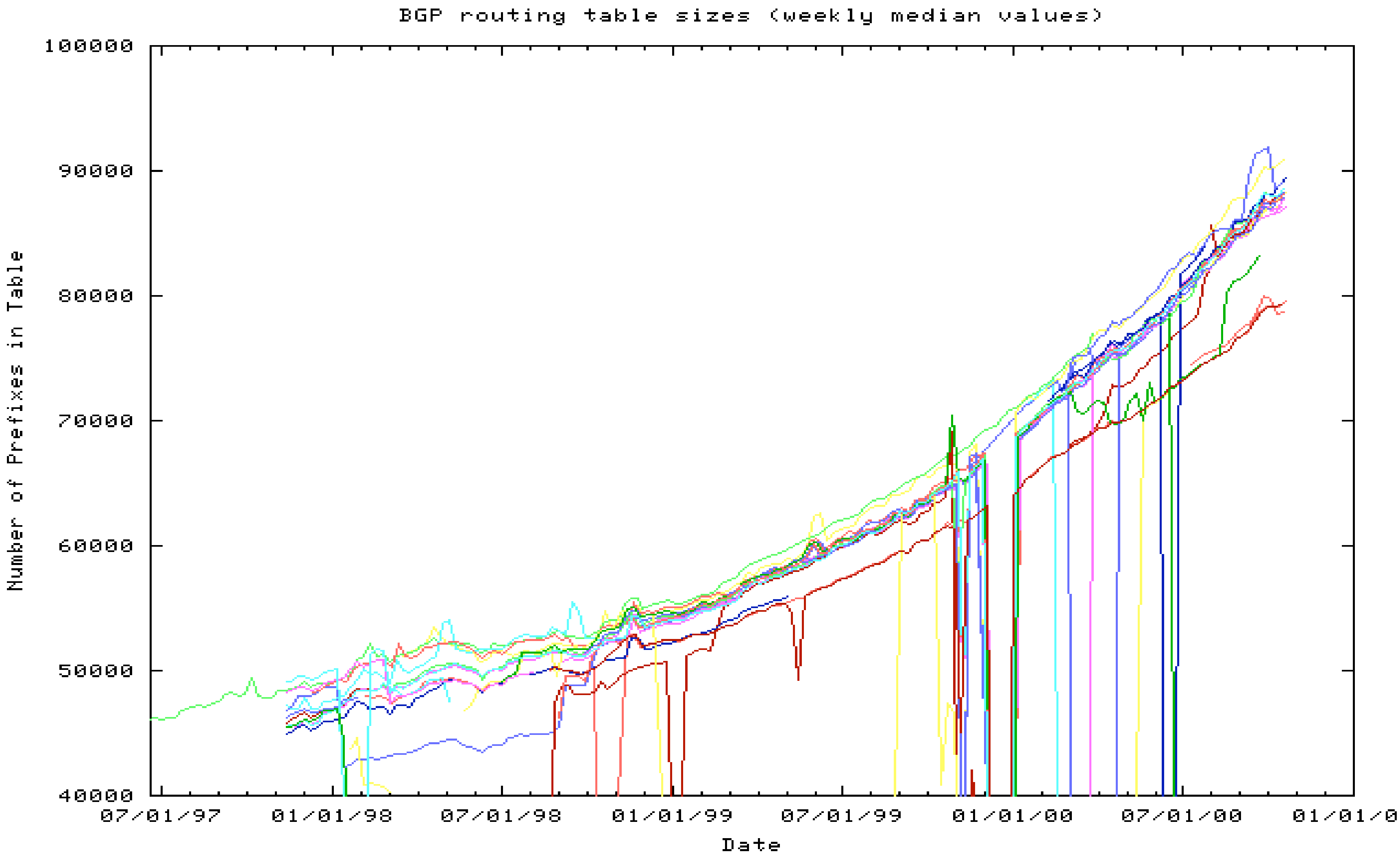
Skitter visualization – BGP paths



The *Route Views* infrastructure

- University of Oregon's inter-domain (autonomous system) routing measurement infrastructure
  - Passive collection system for Boarder Gateway Protocol (BGP - a path vector protocol) routing updates
  - Data used also used to understand size of routing tables
- Infrastructure: Looking glass router that receives BGP peering feeds from 41 networks world wide
- Methodology: database of both routing table snapshots and updates is made available in pseudo real-time
- www.antc.uoregon.edu/route-views
- www.ripe.net/ripencc/pub-services/np/ris-index.html

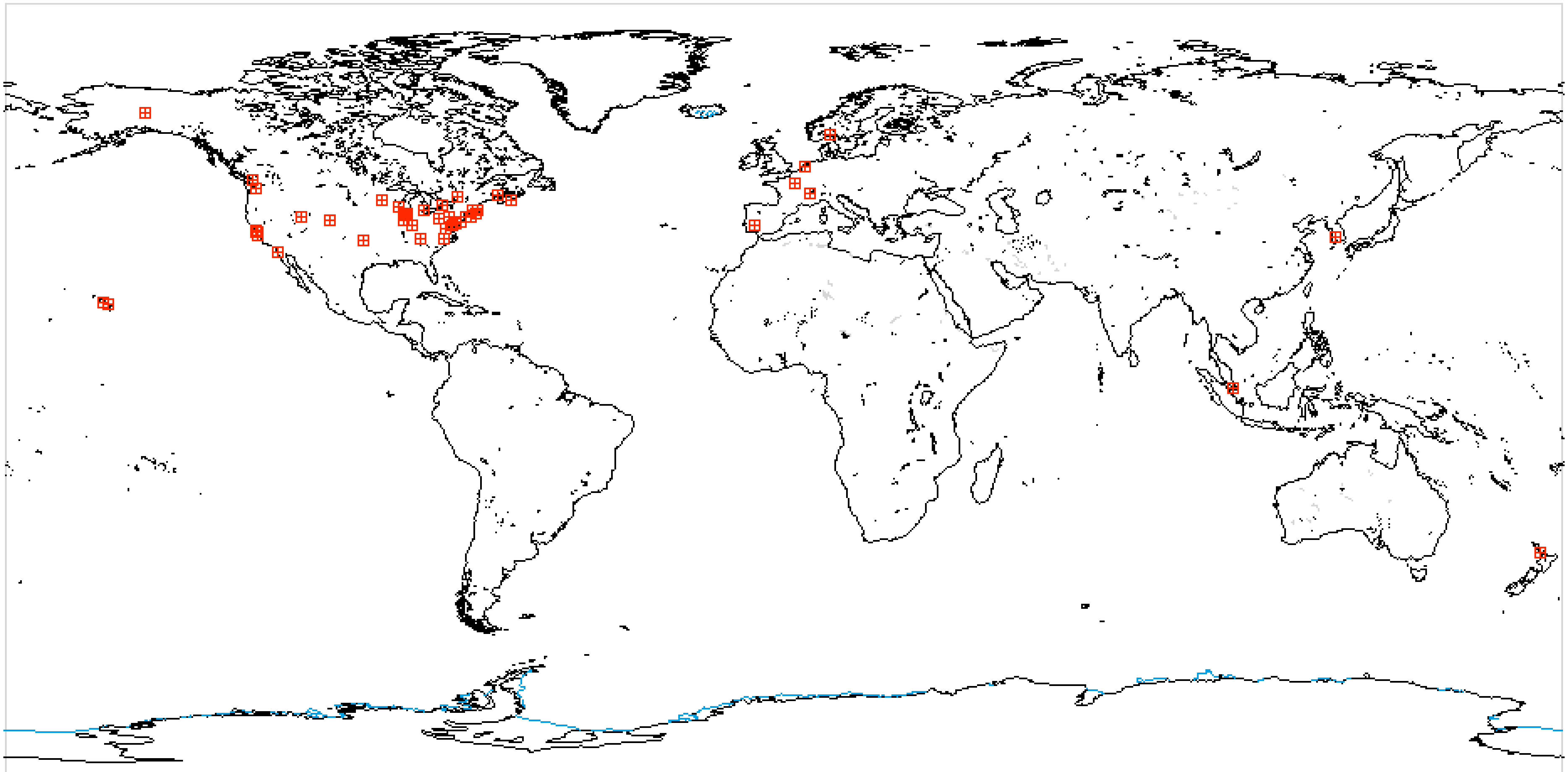
Route Views example



The *Surveyor* infrastructure

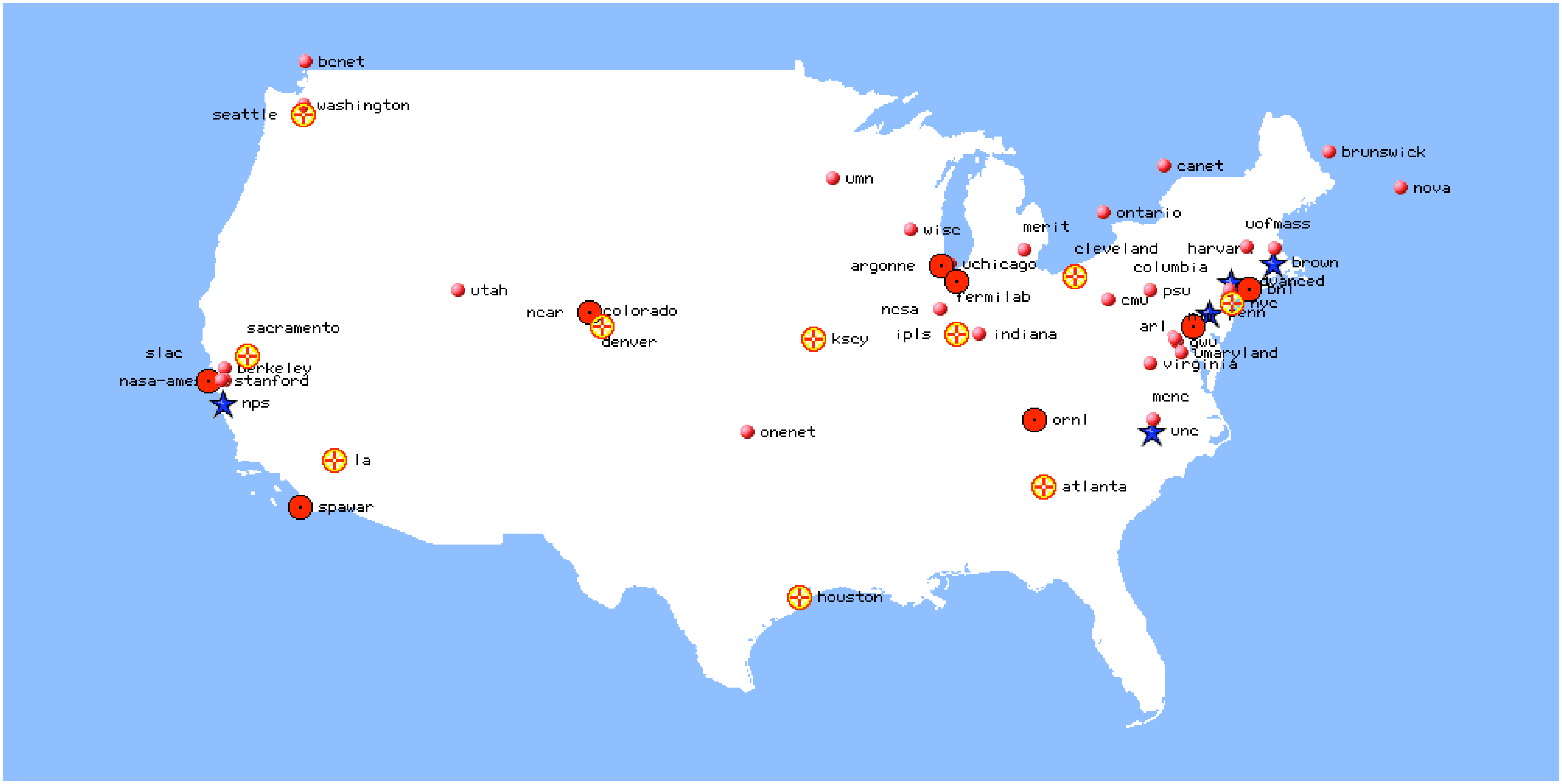
- Advanced Network Systems deployed this infrastructure for measuring Internet performance and reliability
  - Provides data on routing, latency and loss
- Infrastructure: 71 PC measurement systems deployed world wide
  - GPS enabled
  - Centralized database
  - Some analysis and visualization tools
- Methodology: One-way active probe measurement in Poisson intervals (2 Hz avg.) in full mesh 24x7. Traceroutes every 10 min.
- www.advanced.org

Surveyor node deployment



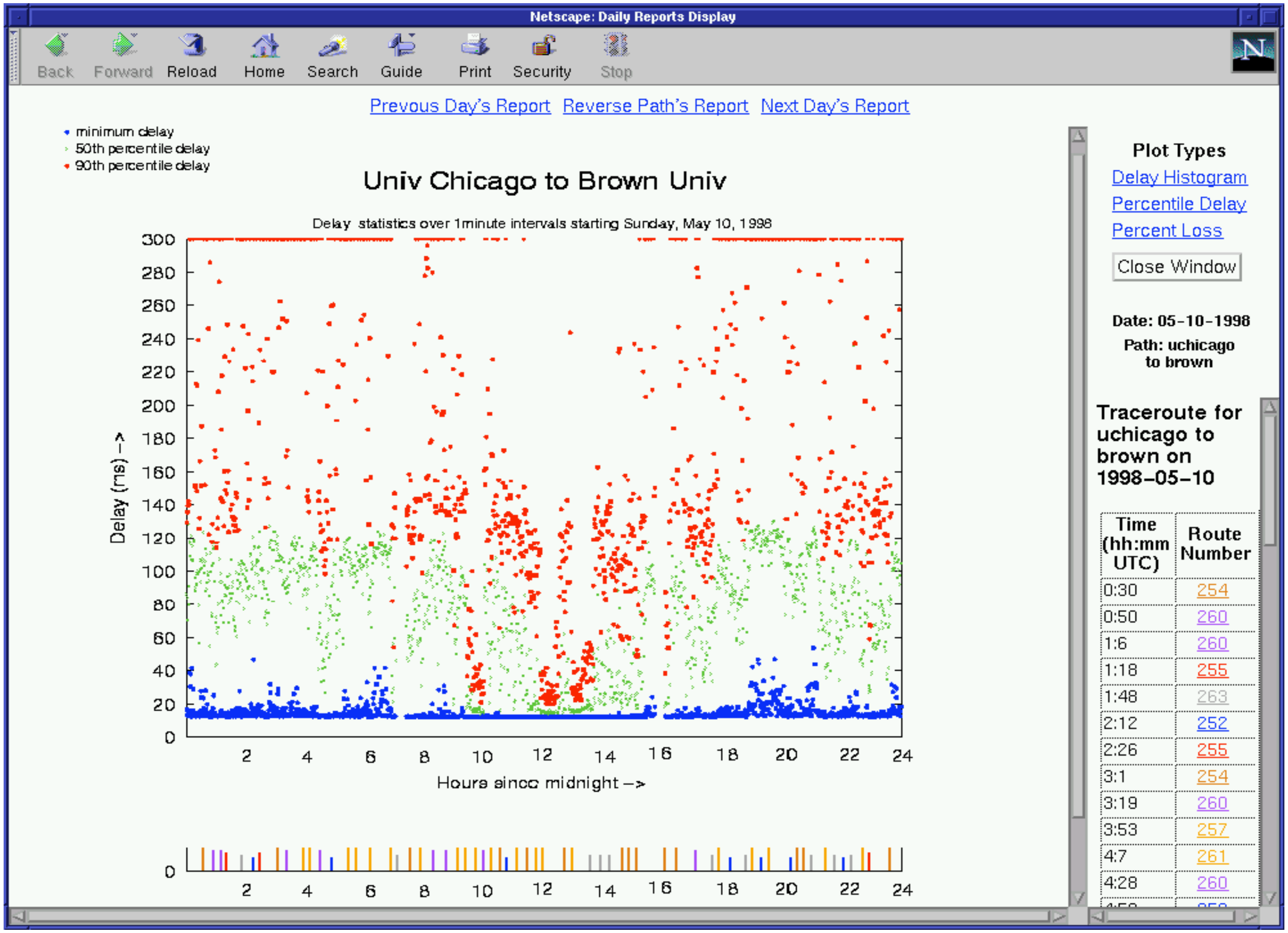
Courtesy Matt Zekauskas, Advanced Systems

Surveyor node deployment in US



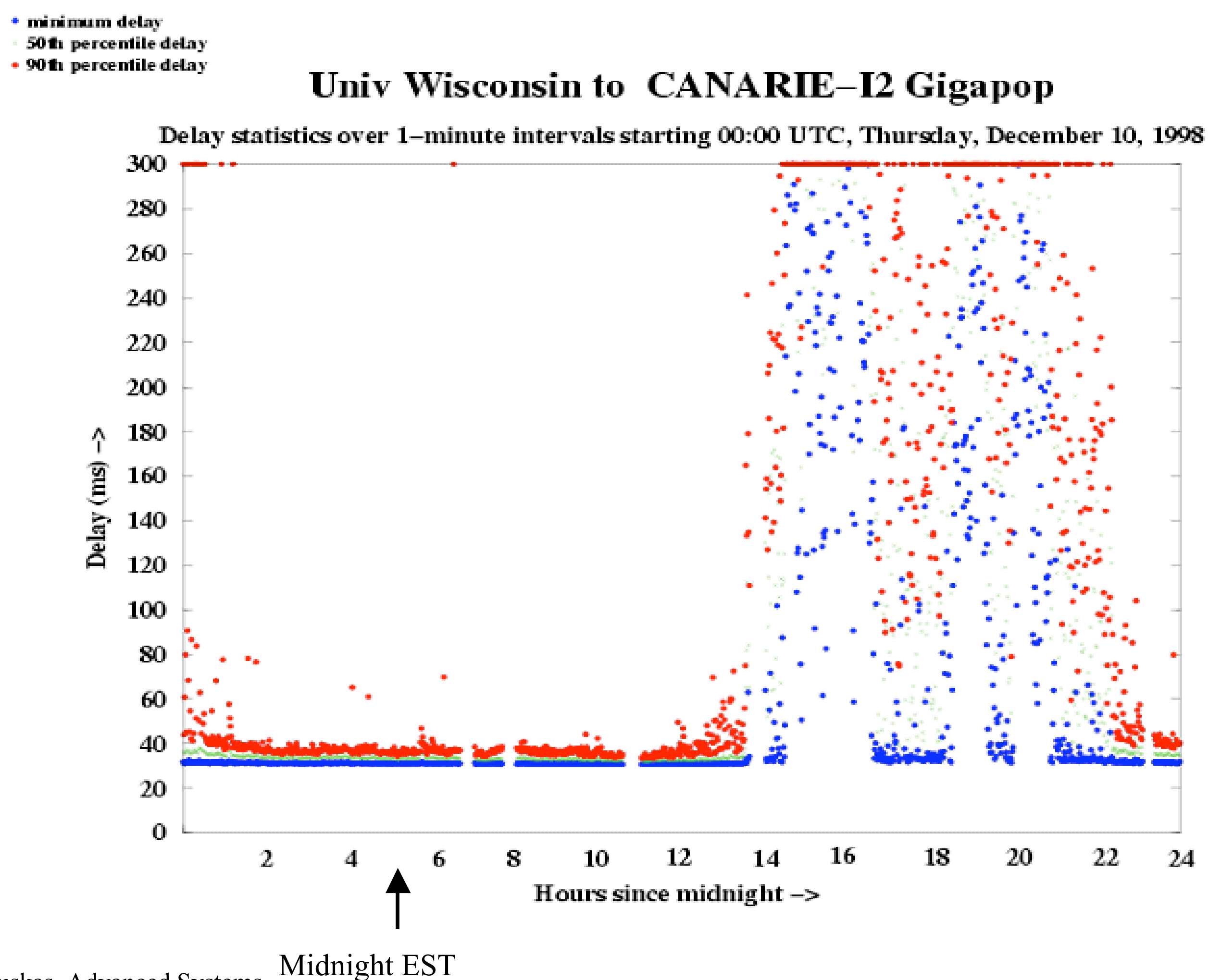
Courtesy Matt Zekauskas, Advanced Systems

Surveyor daily analysis example



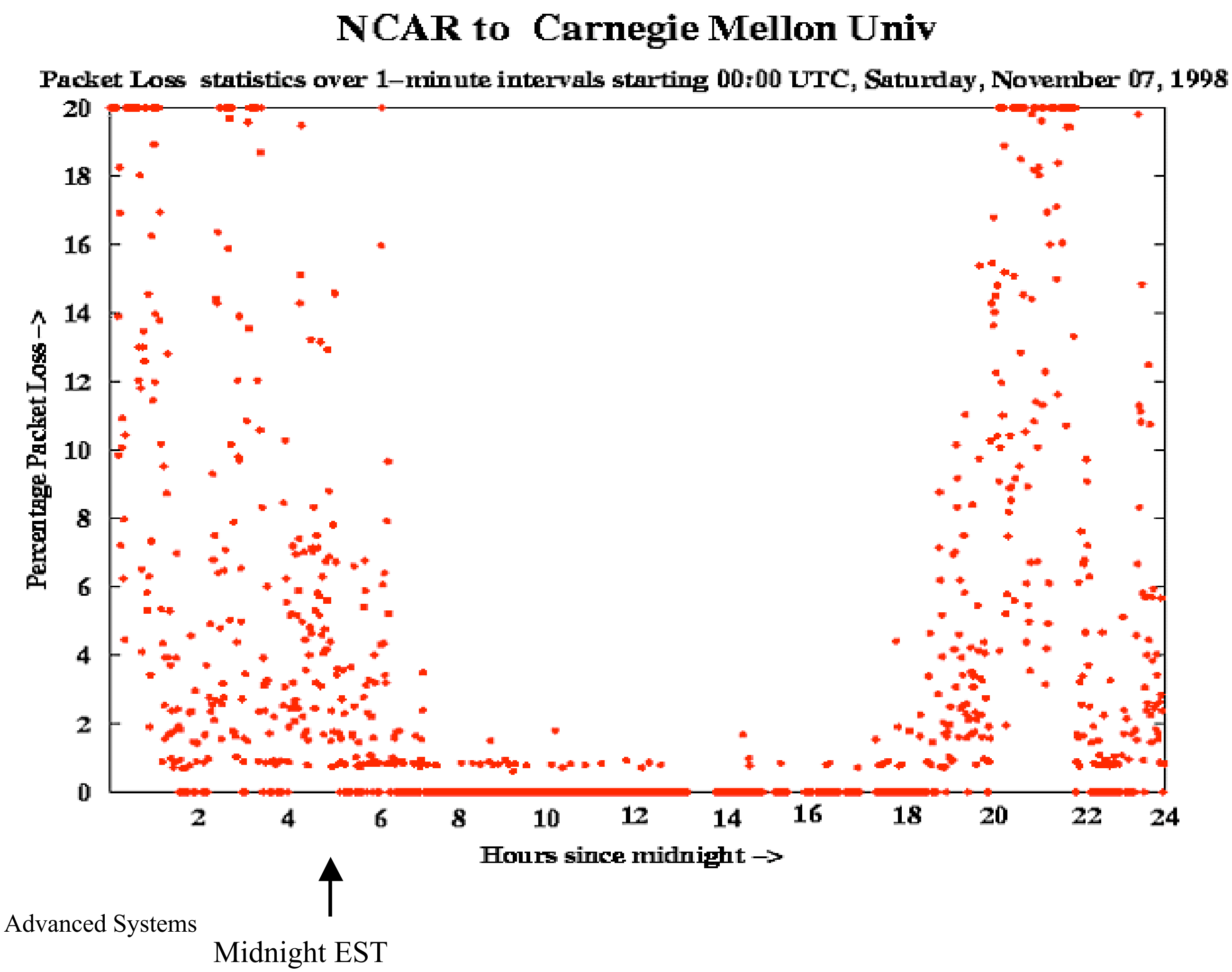
Courtesy Matt Zekauskas, Advanced Systems

Surveyor example - delay



Courtesy Matt Zekauskas, Advanced Systems

Surveyor example - loss



Courtesy Matt Zekauskas, Advanced Systems

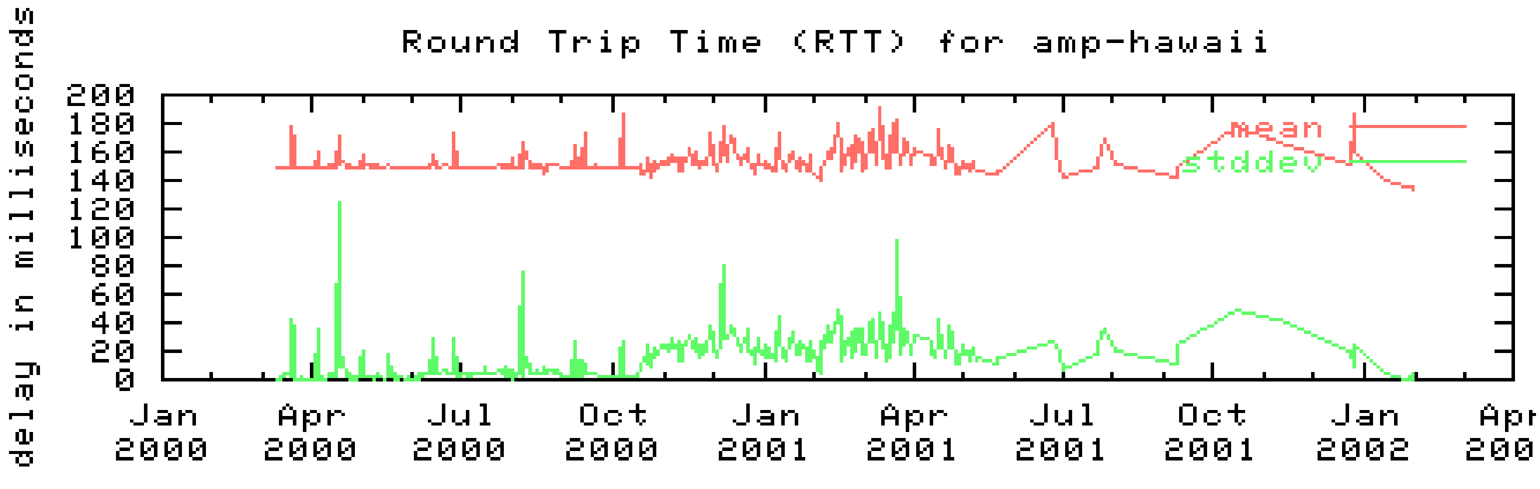
Surveyor status

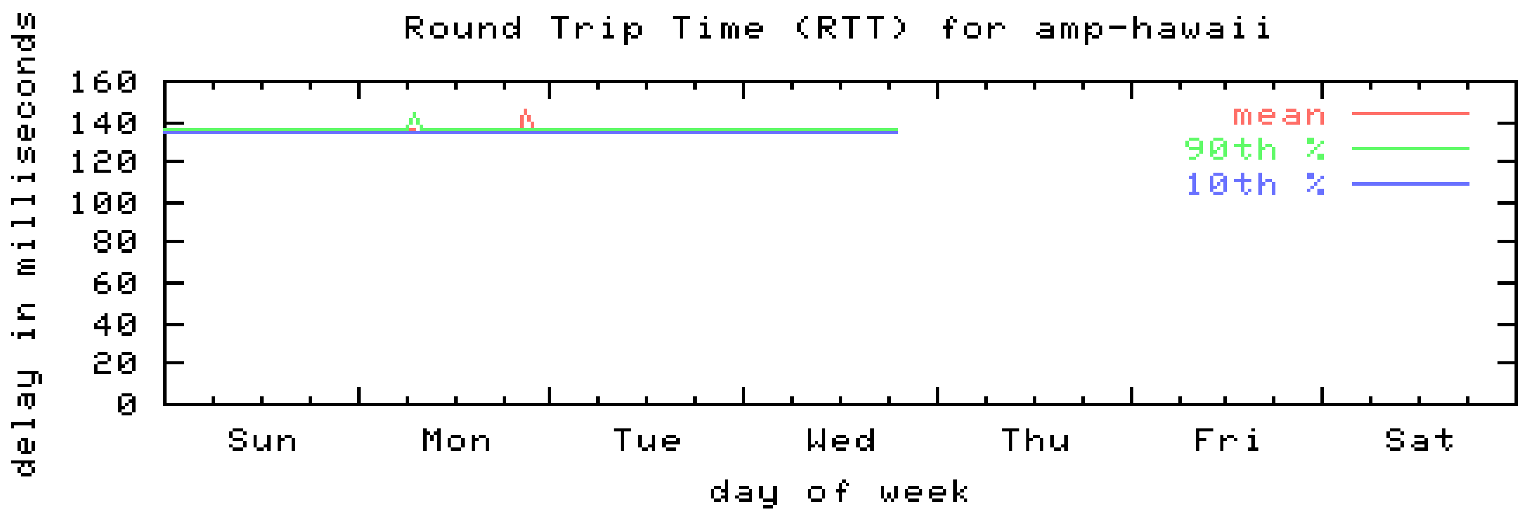
- Systems are currently being overhauled
  - OS transition from BSDi to Linux to enhance manageability
  - Open access to the community via *Scriptroute*
    - Monitoring activity will continue
  - Target: May, ‘04
- Database is currently being overhauled
  - New Web interface and *Netdb* backend
    - *Netdb* developed by DeWitt and Gray
  - 3 years worth of archived data plus new when Surveyor spins up again
  - Target: Summer, ‘04
- http://wail.cs.wisc.edu

The Network Analysis Infrastructure

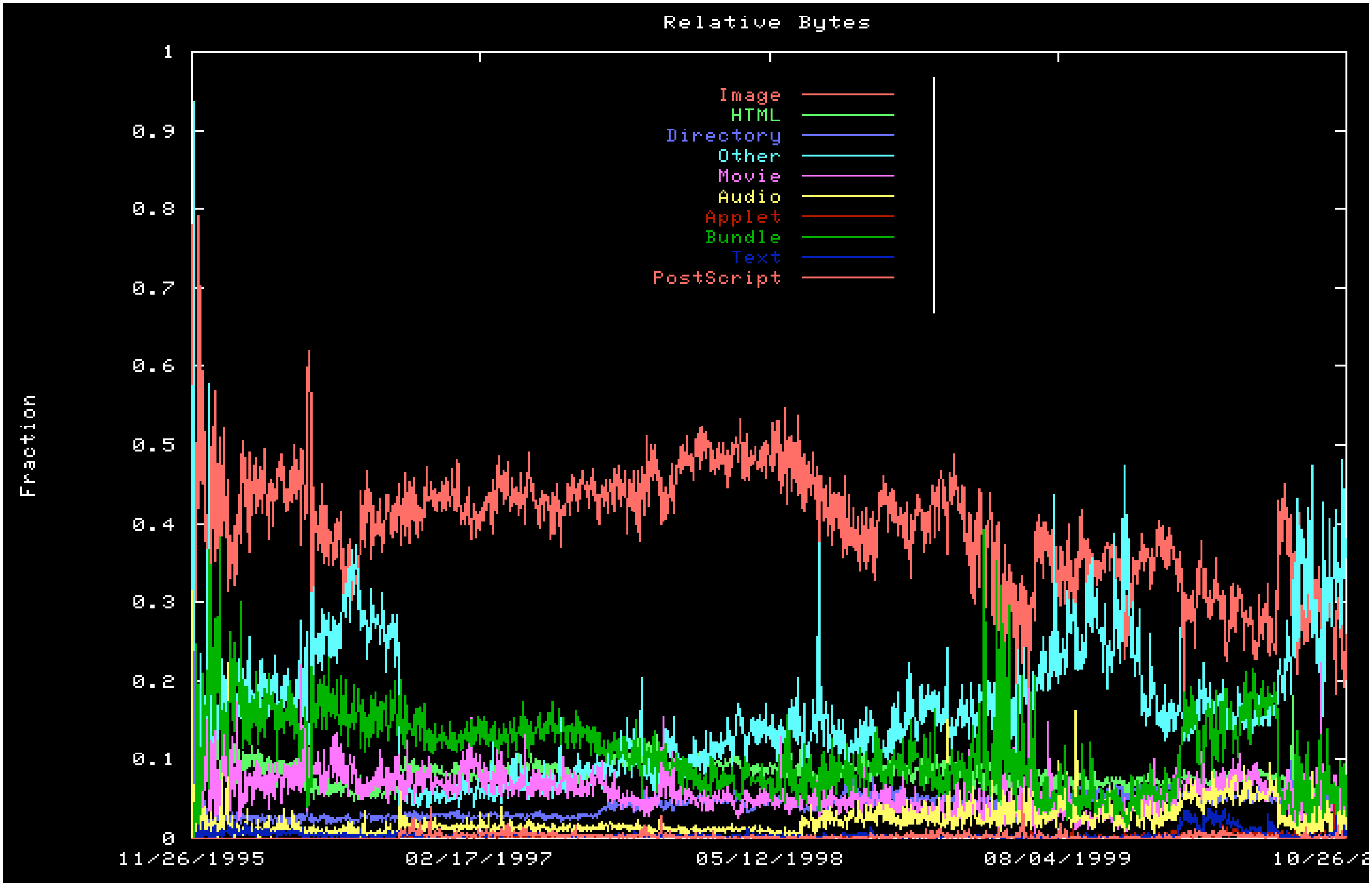
- National Laboratory for Applied Network Research (NLANR)
- Infrastructure for active (AMP) measurements
  - Confederation of universities (approx. 130)
  - RTT, loss and topology measurements
- Infrastructure for passive (PMA) measurements
  - High speed packet monitors across US (approx. 20)
  - Throughput, packet and flow analysis
- Squid cache hierarchy
  - Publicly available cache logs from 10 NLANR caches
- There is a TON of data at these sites!!
- amp.near.net, pma.nlanr.net, ircache.nlanr.net

AMP summary data example





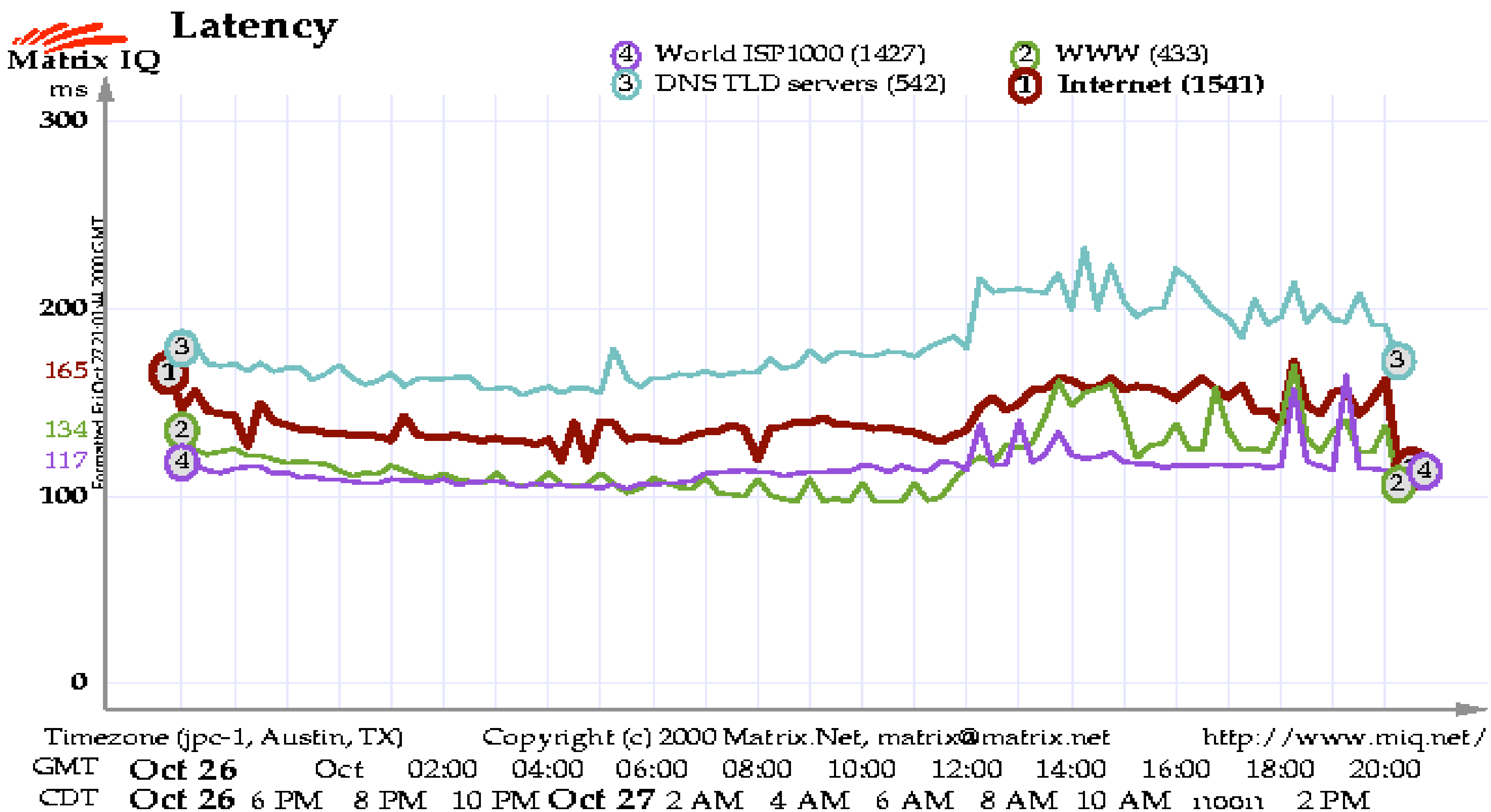
Squid Cache log example



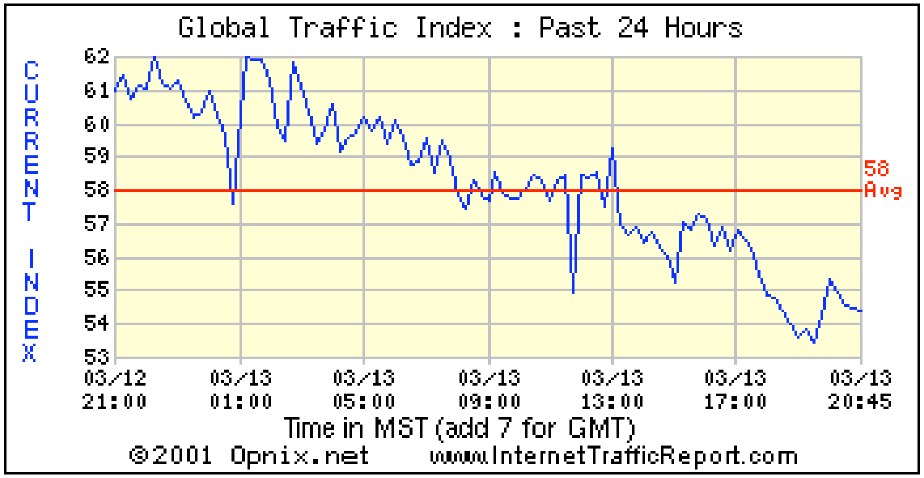
Internet weather/traffic reports

- Andover News, MIDS, others…
- Infrastructures meant to provide high level global Internet traffic statistics
  - Periodic pings to routers, DNS servers and WWW servers all over the world
  - Break down by provider and geographically
  - Commercial focus
- www.internettrafficreport.com, www.mids.org

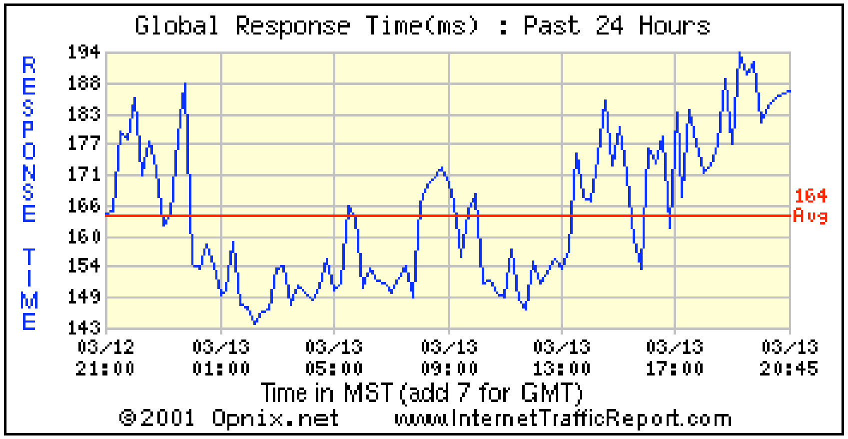
Internet traffic report example

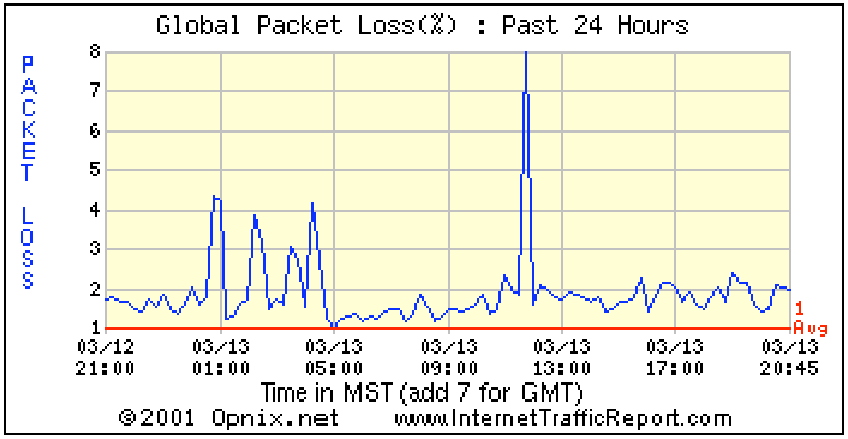


Global Internet traffic summary



Ping response time





"Chunk of data" response time

No response to several pings

The PlanetLab infrastructure

- "An open, globally distributed platform for developing, deploying and accessing planetary-scale network services."
  - An Intel-funded, distributed PC infrastructure
  - Systems deployed at 137 sites world wide
  - A systems development project with Internet measurement as a primary application
  - http://www.planet-lab.org

Managing experiments with *Scriptroute*

- A "measurement tool management system"
- "The goal is to allow any user to connect to any server and execute any safe network measurement"
  - Basic tools include *ping, traceroute*
  - Enables system resource restriction
- Deployed widely, including Planetlab
- http://www.cs.washington.edu/research/networking/scriptroute/

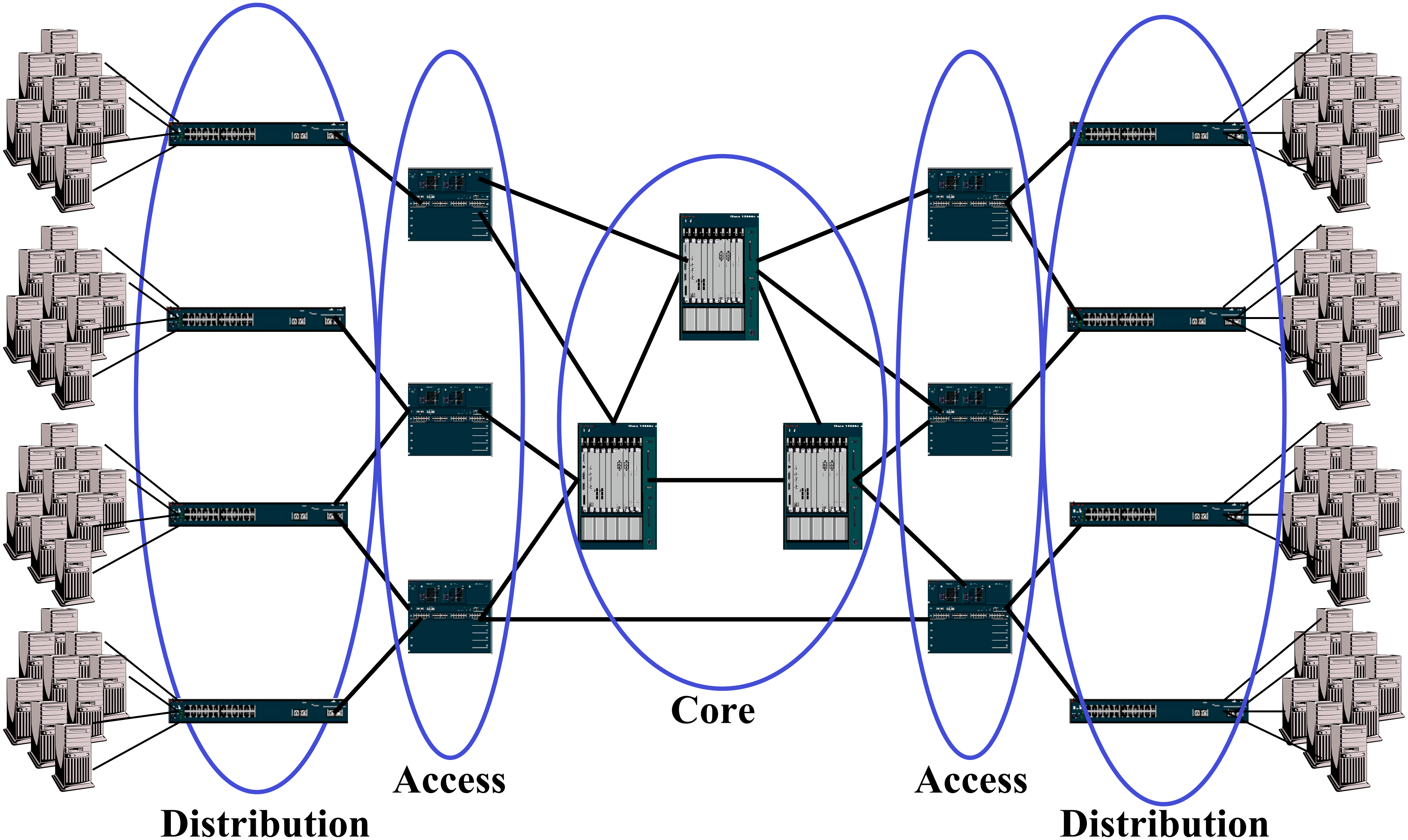
traceroute.org

- Routing and topology study is greatly influenced by *location* of measurement hosts
- There are hundreds of resources all over the Internet openly available for executing *traceroute*
  - Capabilities vary
- www.traceroute.org is a huge index of these resources
  - Be careful using these!

Another approach - The Wisconsin Advanced Internet Lab

- Why do we need an "internal" lab?
  - Enables instrumentation and measurement of entire end-to-end system
  - Enables new systems and protocols to be implemented in places where access is not possible in wide area
- Complement to external facilities
- Hands-on test bed which creates paths identical to those in the Internet from end-to-end-through-core
  - Variety of highly configurable equipment
- Vision of internal lab: New means for doing network research
- Status: Systems operational and open access by summer '04
- http://wail.cs.wisc.edu

WAIL Conceptual Design



Part 3: Measurement data analysis

Standard approaches to data analysis

- Summary statistics
  - Are these meaningful considering size, and complexity?
- Histograms and curve fitting
  - There is a danger that we are spending too much time here
- Assessment of upper tails
  - Many properties exhibit heavy tails
- Assessment of scaling properties
  - Self-similarity is one of the true success stories

Modeling and simulation

- A variety of models for Internet traffic have been proposed
  - Willinger et al. "Self-similarity Through High-variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level", *IEEE/ACM Transactions on Networking*, 1997.
    - Proposed ON/OFF model for network traffic
  - Many rely on a mathematical construction to generate data that can be shown to approximate Internet traffic behavior
    - These provide no insight into Internet mechanisms
- Simulations have been successful but are highly simplified
  - www.isi.edu/nsnam/ns, www.ssfnet.org

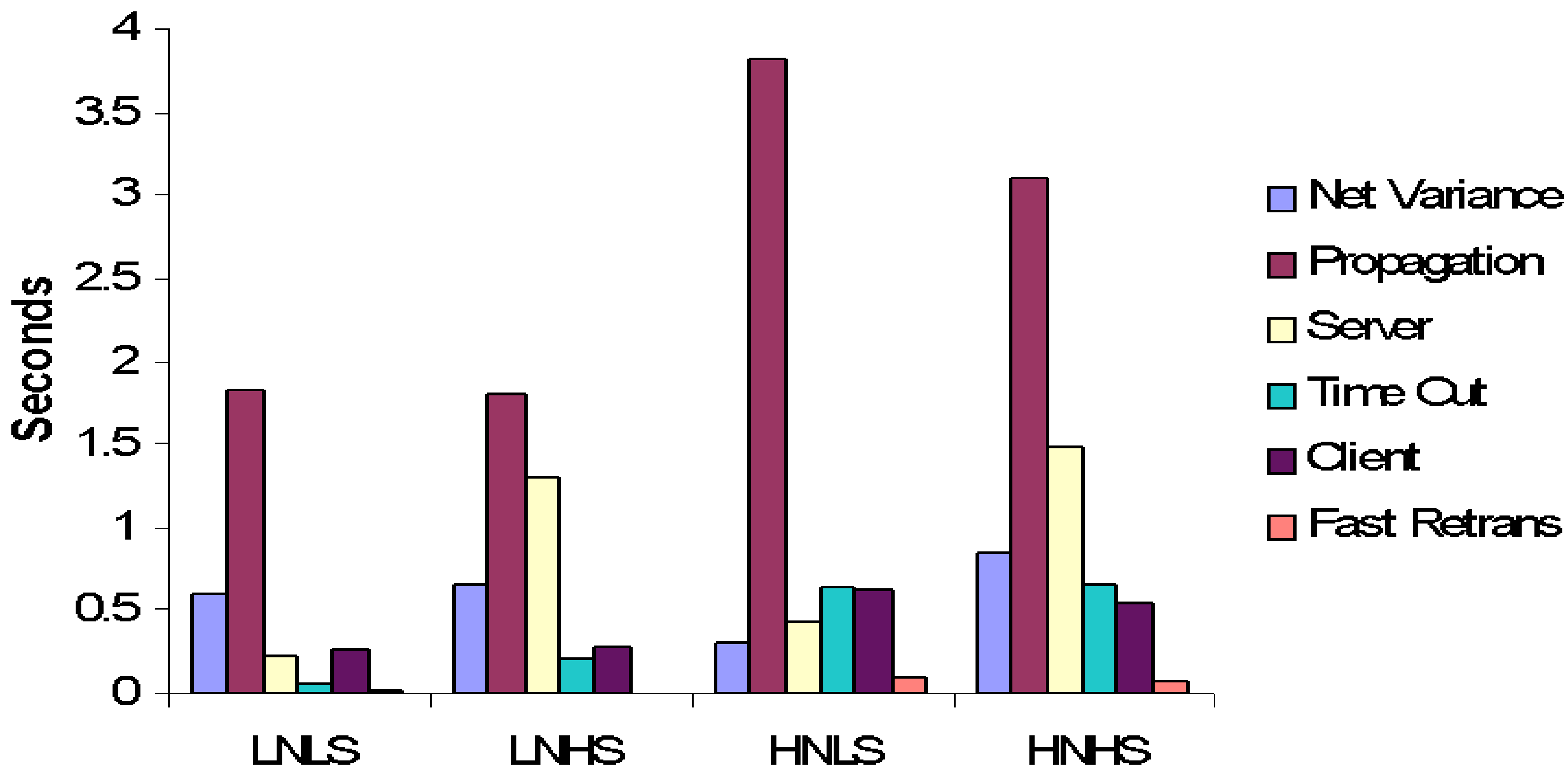
New analysis methods are necessary

- Innovative methods for extracting more information from existing data
  - Critical path analysis, fluid analysis, Rocketfuel
- Mechanistic models which explain behavior of the Internet at a variety of levels
- Application of new mathematical and statistical methods
  - Fractals, wavelets, non-linear dynamics, etc…
- Innovative visualization techniques are necessary
  - Dimensionality and magnitude must be addressed

Extracting more information from TCP packet traces

- Barford and Crovella "Critical Path Analysis of TCP Transactions", *IEEE/ACM Transactions on Networking*, 2001.
- CPA identifies the precise set of events that determine execution time of a distributed application
- Applying CPA to TCP transactions enables accurate assignment of delays to:
  - Server delay
  - Client delay
  - Network delay (propagation, network variation and drops)

CPA example



File transfer delay for 500KB file between Denver and Boston

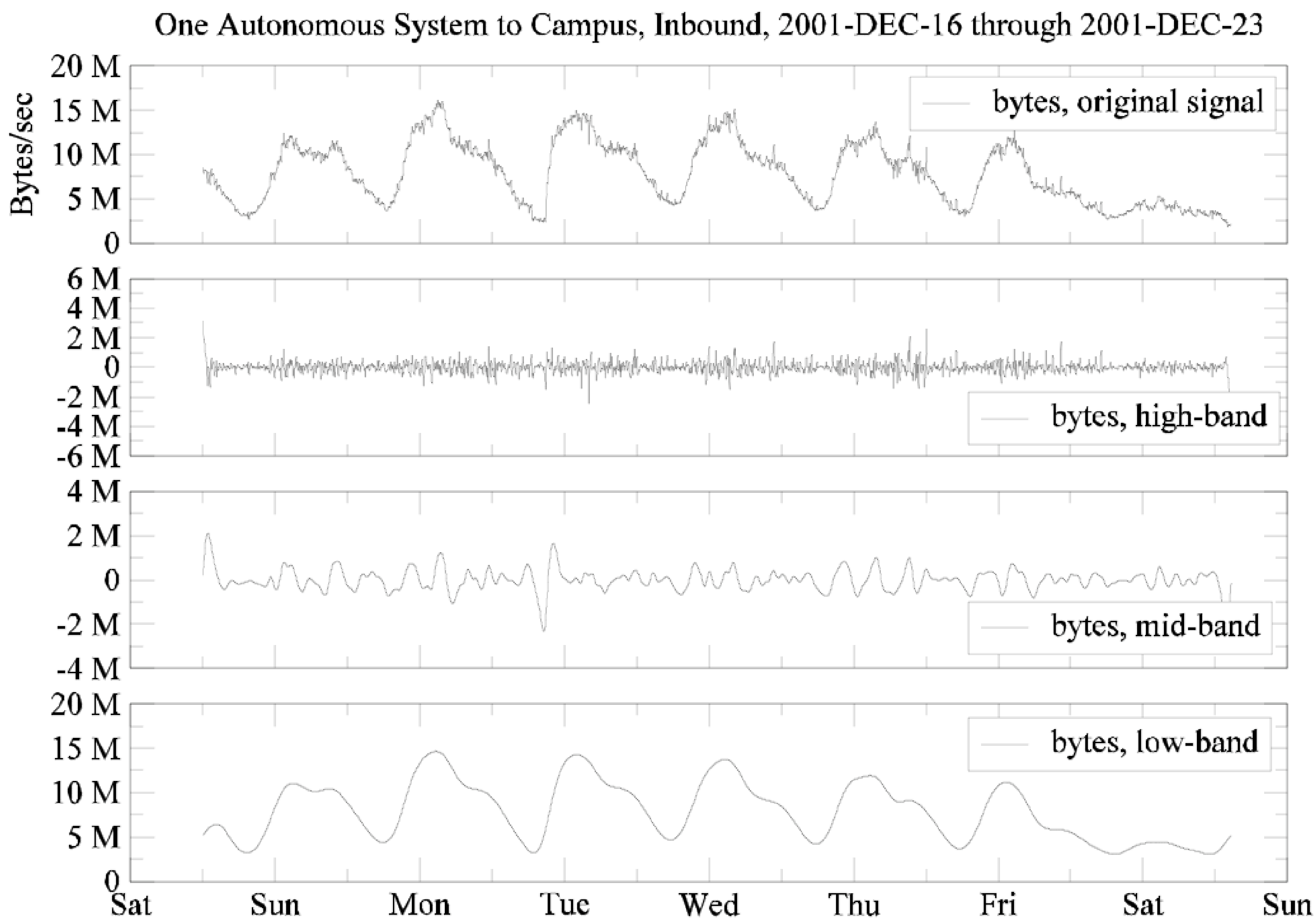
Analysis of traffic anomalies

- **Motivation:** Anomaly detection remains difficult
- **Objective**: Improve understanding of traffic anomalies
- **Approach**: Multiresolution analysis of data set that includes IP flow, SNMP and an anomaly catalog
- **Method**: Integrated Measurement Analysis Platform for Internet Traffic (IMAPIT)
- **Results**: Identify anomaly characteristics using wavelets and develop new method for exposing short-lived events
  - Barford *et al*., "A Signal Analysis of Network Traffic Anomalies", Internet Measurement Workshop, 2002.

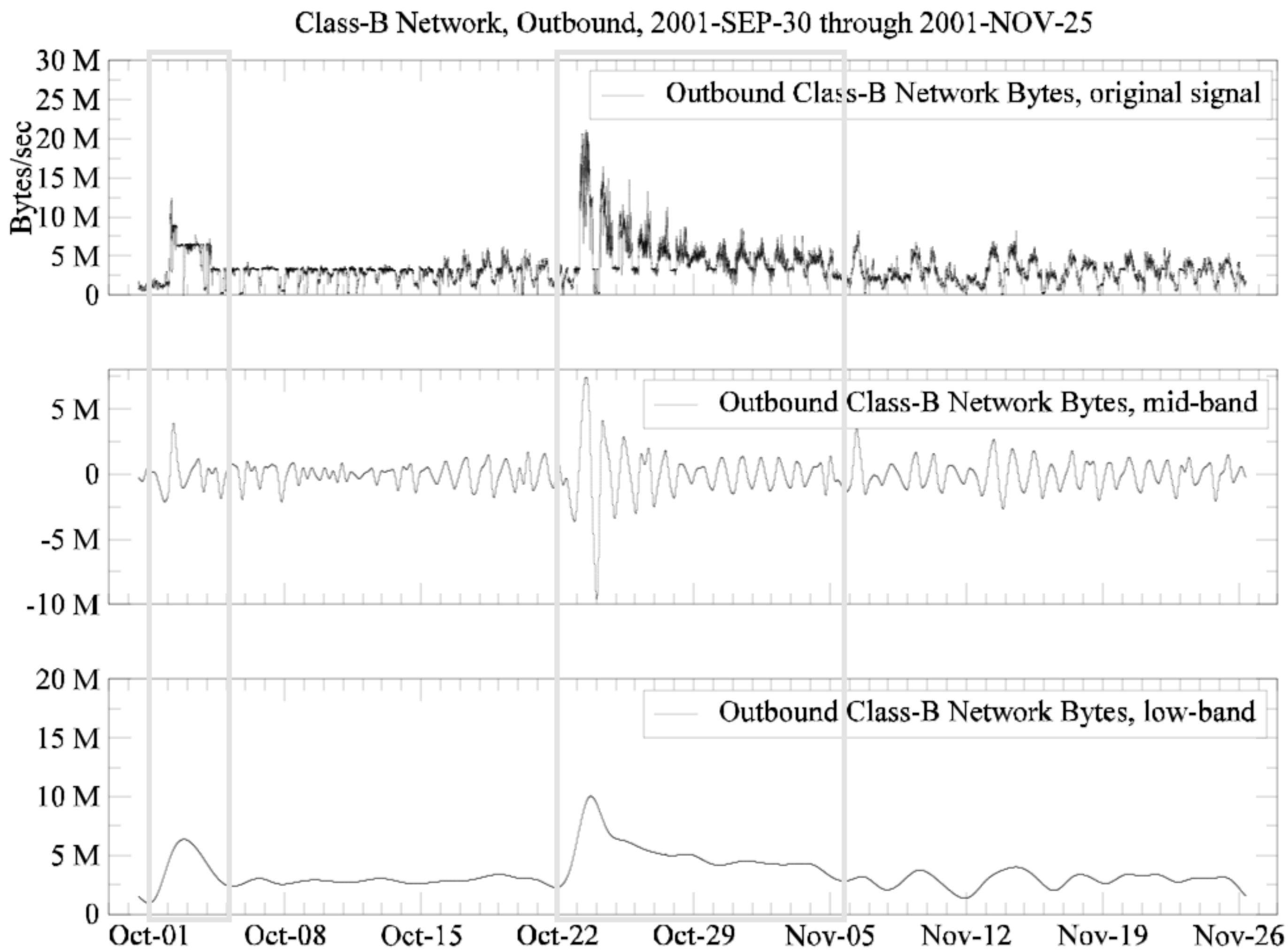
Multiresolution analysis

- Wavelets provide a means for describing time series data that considers both *frequency* and *time*
  - Powerful means for characterizing data with sharp spikes and discontinuities
  - Using wavelets can be quite tricky
- We use tools developed at UW which together make up IMAPIT
  - FlowScan software
  - The IDR Framenet software

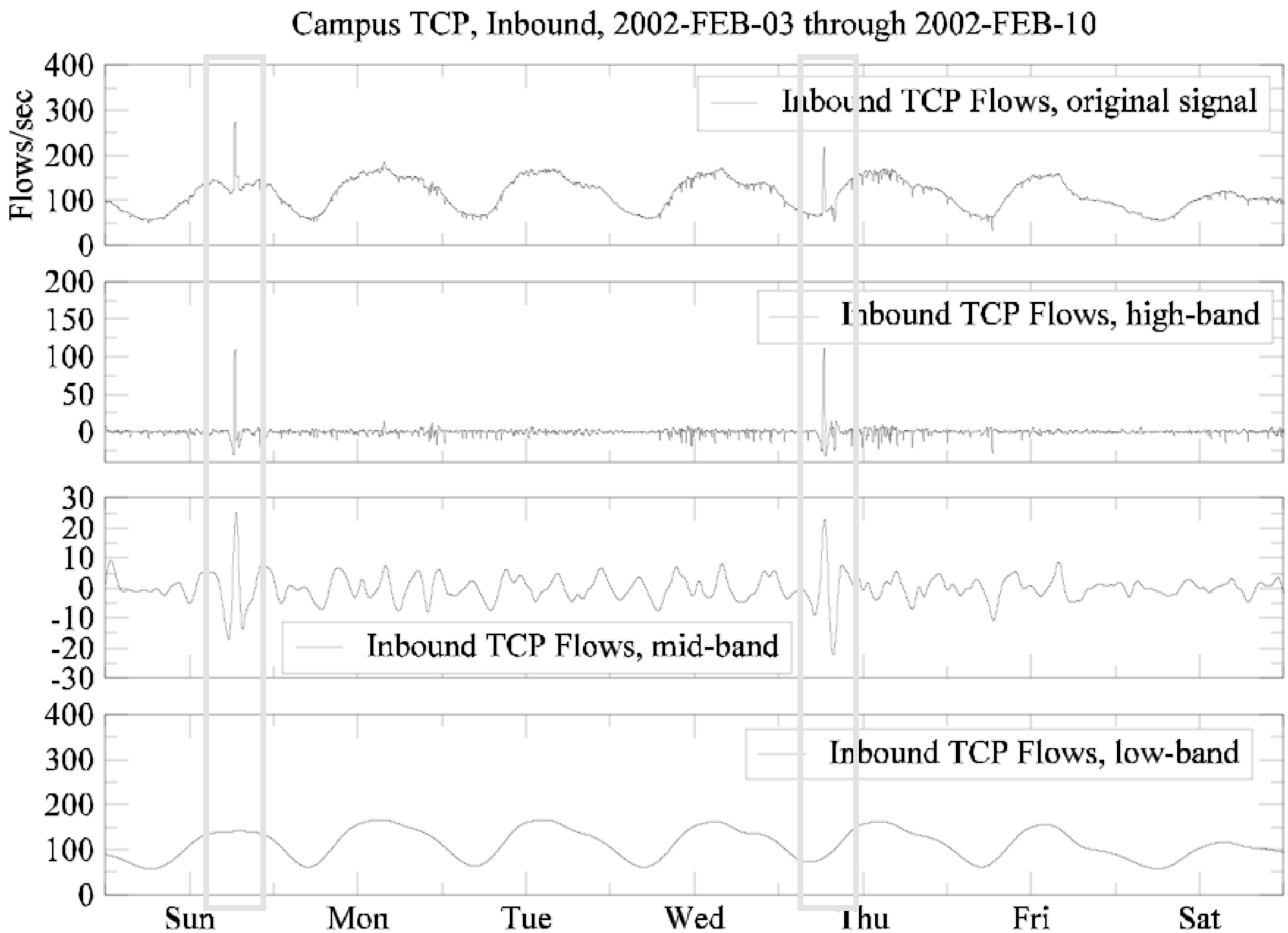
Ambient IP flow traffic



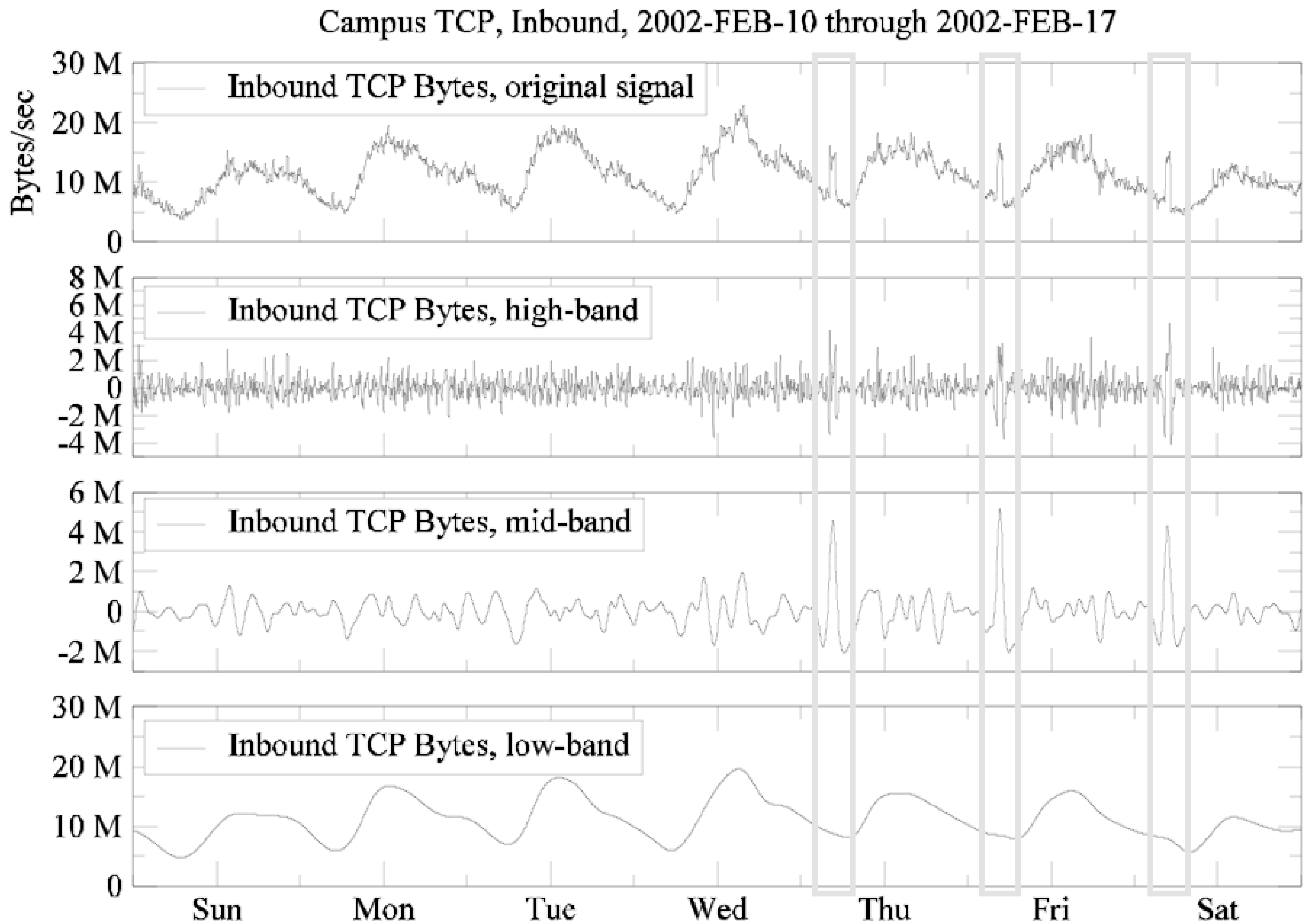
Byte traffic for flash crowd



Flow traffic during DoS attacks



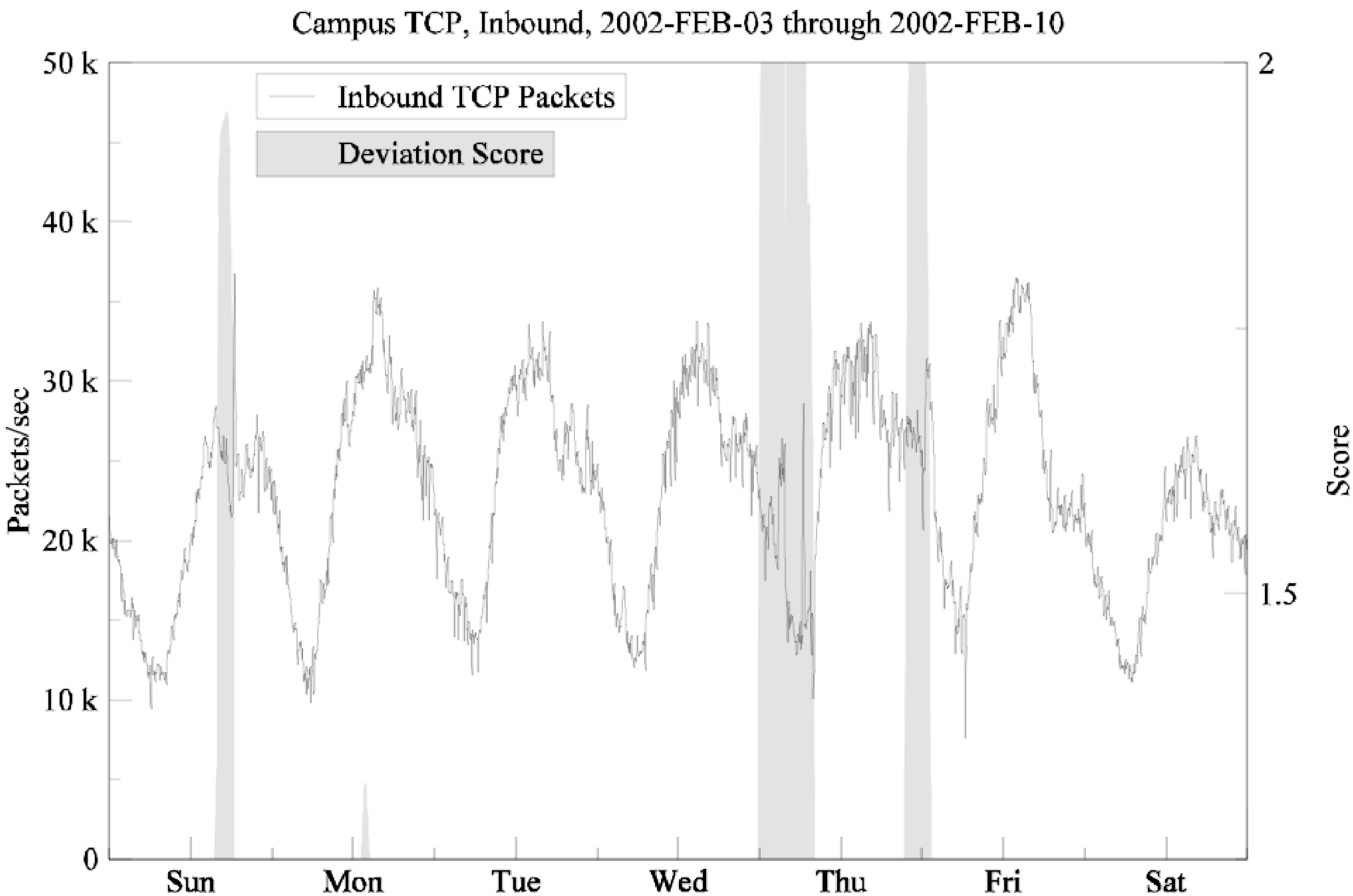
Byte traffic during measurement anomalies



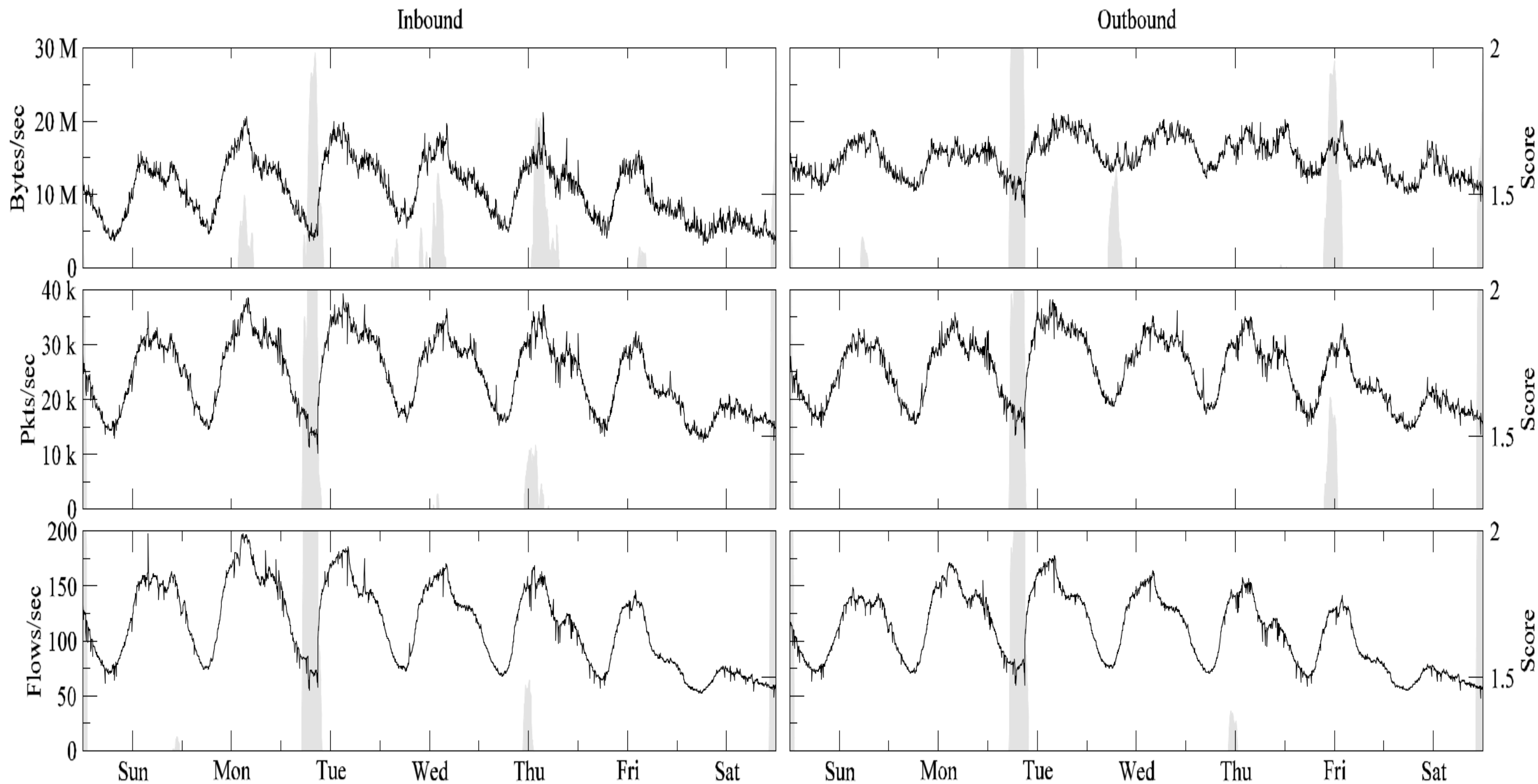
Anomaly detection via *deviation score*

- Short-lived anomalies can be identified automatically based on variability in H and M signals
  1. Compute local variability (using specified window) of H and M parts of signal
  2. Combine local variability of H and M signals (using a weighted sum) and normalize by total variability to get deviation score V
  3. Apply threshold to V then measure peaks
- Analysis shows that V peaks over 2.0 indicate short-lived anomalies with high confidence
  – We threshold at V = 1.25 and set window size to 3 hours

Deviation score for three anomalies



Deviation score for network outage



Deviation score evaluation

- How effective is deviation score at detecting anomalies?
  - Compare versus set of 39 anomalies
    - Set is unlikely to be complete so we don't treat false-positives
  - Compare versus ***Holt-Winters Forecasting***
    - Time series technique
    - Requires some configuration
- Holt-Winters reported many more positives and sometimes oscillated between values

Total Candidate Anomalies	Candidates detected by Deviation Score	Candidates detected by Holt-Winters
39	38	37

Part 4: Difficulties with Network Measurement

Poor data consistency

- Data collected for the same characteristic using different methods does not always agree
  - Packet delays in TCP versus probes
  - Logs from different sources do not always have the same information
- Perspective matters
  - Results for BGP by Chang *et al*. from UMich
- Clock synchronization is always difficult

Inaccurate tools

- Active measurement tools can be blocked
- Passive measurement tools can behave in all sorts of ways
- Systems operating at lower levels in the network are not visible
- Privacy issues limit the ability to validate
- Calibration is difficult

Representativeness

- Size, heterogeneity, constant and radical change
- Just what does “representative” mean?
  - There may be no such thing
  - There do seem to be some invariant properties
    - Self-similarity
    - Heavy-tails
- Infrastructures available to the community are important

Reproducing results

- The networking community does not have a culture of reproducing results
- There are very few instances of public repositories of data
  - research.cs.vt.edu/nrg/dbase/nrgsearch.html
  - ita.ee.lbl.edu
  - www.internet2.org
  - Infrastructures mentioned in this talk
- There is little sharing of analysis tools

Explosion of data!!

- Current state of the art is OC192 (10 Gbps)
- Many popular web sites get over 1B hits per day
- Understanding all aspects of the Internet require measurements across many layers
- No standard databases for Internet measurement data
- Datasets today overwhelm statistical methods and statistical tools

Conclusions

- Measurements are necessary for understanding and improving Internet structure and behavior
- Tools and methods for taking Internet measurements give a limited view of Internet characteristics and behavior
- Current measurement infrastructures can provide a great deal of data but fall short of the GIMI goal
- There is a significant need to expand the analysis methods employed to evaluate Internet data
- Internet measurements are easy to do poorly and difficult to do well

Acknowledgements

Thanks to the following people for their help and support:

Mark Crovella, Jeff Kline, Larry Landweber, Vern Paxson, David Plonka, Amos Ron, Walter Willinger, Vinod Yegneswaran, Matt Zekauskas