

WAN Monitoring

Prepared by Les Cottrell, SLAC, for the
Joint Engineering Taskforce Roadmap Workshop
JLab April 13-15, 2004

www.slac.stanford.edu/grp/scs/net/talk03/jet-apr04.ppt



Partially funded by DOE/MICS Field Work Proposal on
Internet End-to-end Performance Monitoring (IEPM), also
supported by IUPAP

Why

(Can't manage what you can't measure)

- Need measurements for both production networks & testbeds:
 - Planning, setting expectations, policy/funding
 - Trouble-shooting: reliability & performance
 - Problems may not be logical, e.g. most Internet problems caused by operator error (Sci Am Jun'03), most LAN problems are Ethernet duplex, host config, bugs
 - Made hard by transparency, size & rate of change of network
 - *A distributed system is one in which I can't get my work done because a computer I never heard of has failed.* Butler Lampson
 - Application steering (e.g. Grid data replication)
- E2E performance problem is THE critical user metric

E.g. Policy - trends

C. Asia, Russia, S.E. Europe,
L. America, M. East, China:
4-5 yrs behind

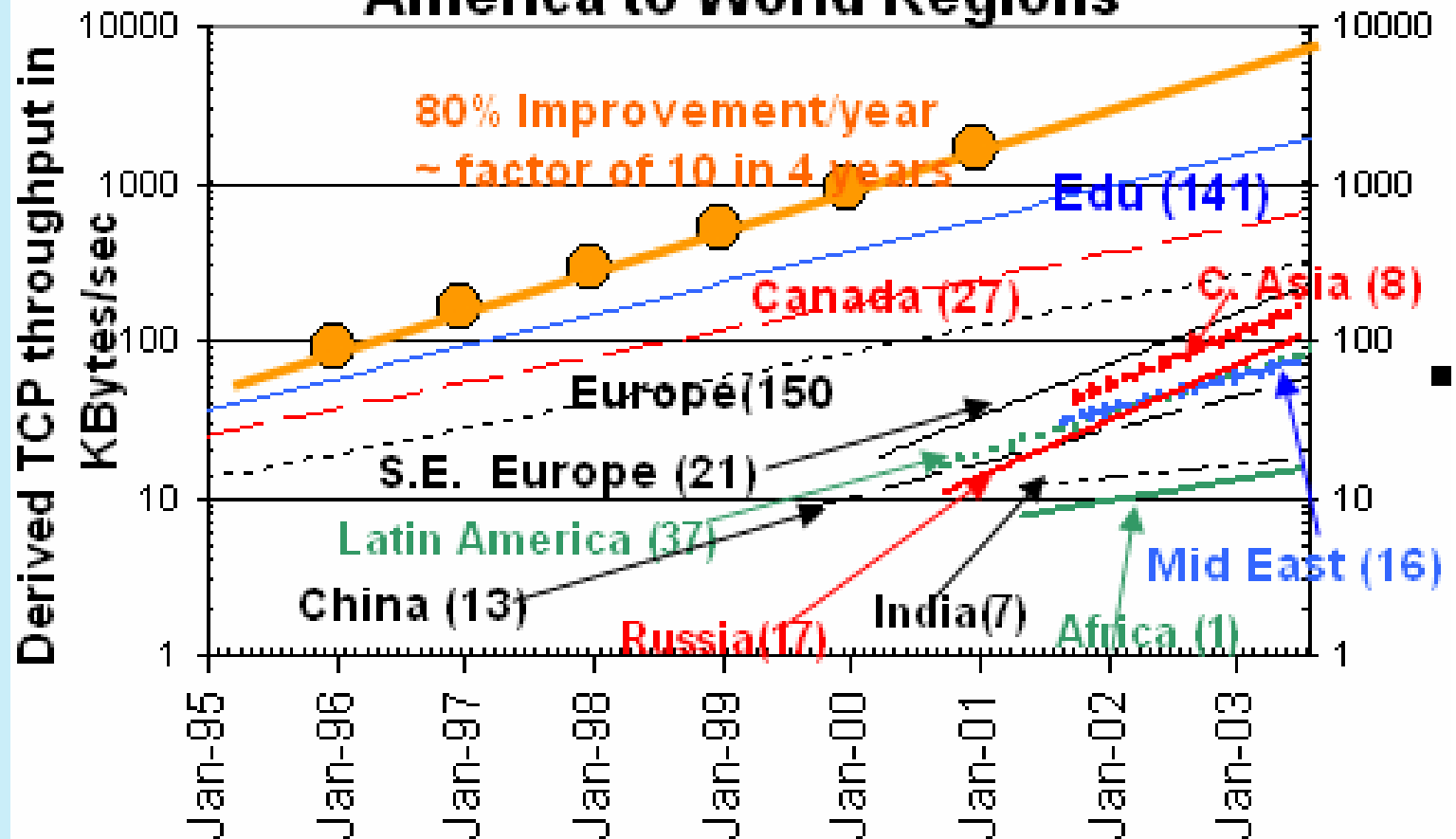
S.E. Europe, Russia: **catching up**

Latin Am., Mid East, China: **keeping up**

India, Africa: **falling behind**

Important
for policy
makers

TCP throughput measured from N. America to World Regions





E.g. Changes in network topology (BGP) result in dramatic change in performance

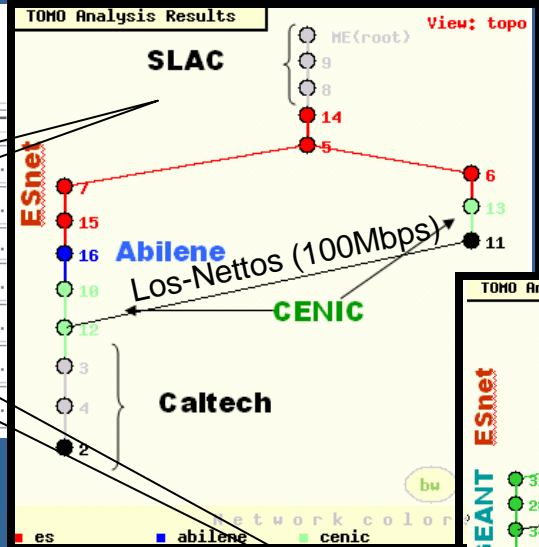
Traceroute Analysis for 10/09/2003

Yesterday's Summary | Reverse Traceroute Summary | Directory of Historical Traceroutes

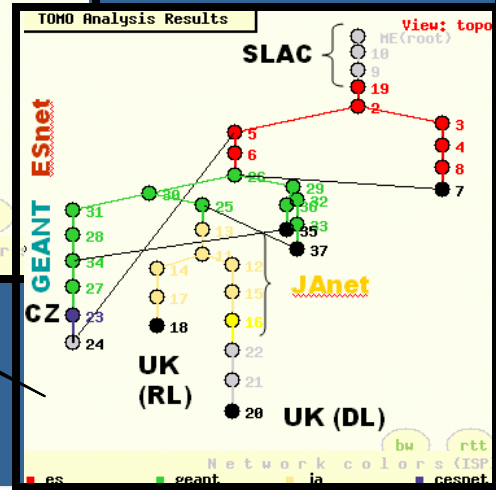
SUBMIT Topology request | RESET FIELDS

Hour (PST) →

NODE \ Hour =>	00	01	02	14	15	16	17
node1.cacr.caltech.edu* R Sum Log*	105	110	105
node1.cesnet.cz* R Sum Log*	35	58	35	39
node1.crc.ac.uk* R Sum Log*	67	74	67	67
node1.dl.ac.uk* R Sum Log*	97	102	97	97
node1.ece.rice.edu* R Sum Log*	104
node1.fnal.gov* R Sum Log*	8
node1.in2p3.fr* R Sum Log*	29	77	100	32



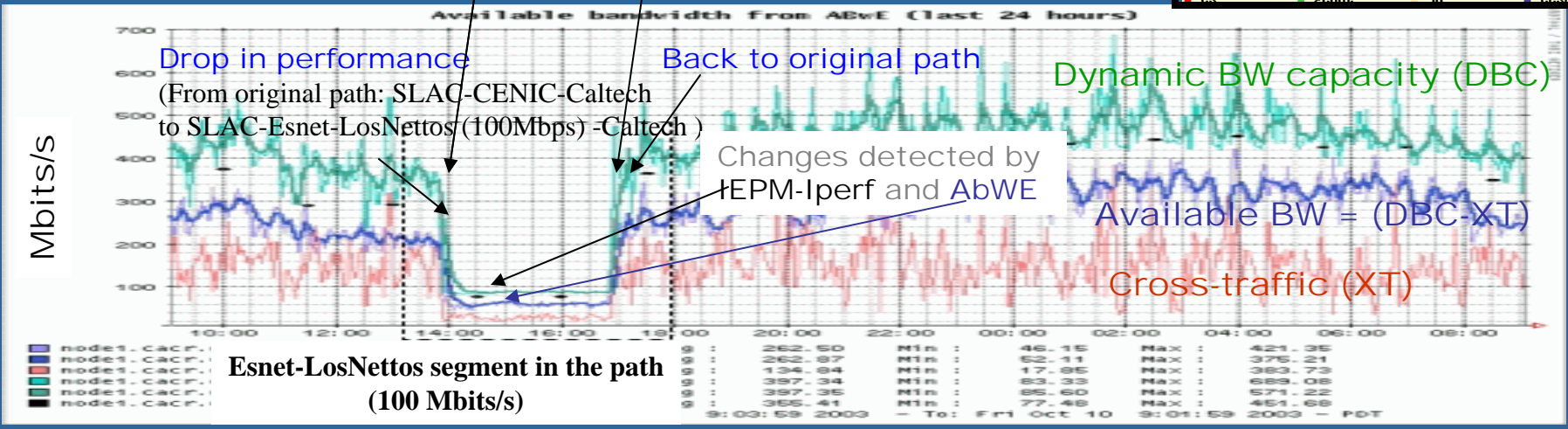
Samples of traceroute trees generated from the table



Snapshot of traceroute summary table

Notes:

1. Caltech misrouted via Los-Nettos 100Mbps commercial net 14:00-17:00
2. ESnet/GEANT working on routes from 2:00 to 14:00
3. A previous occurrence went un-noticed for 2 months
4. Next step is to auto detect and notify



ABwE measurement one/minute for 24 hours Thurs Oct 9 9:00am to Fri Oct 10 9:01am

- Active Measurement probes:
 - Include: Ping, traceroute, owamp, pathload/abwe, major apps (e.g. bbftp, bbcp, GridFTP...)
 - Typically used for end-to-end testing
 - Inject data into network
- Passive tools:
 - Include: SNMP, NetFlow, OCxMon, NetraMet, cflowd, SCNM
 - Typically used at border or inside backbones
 - SNMP heavily used for utilization, errors on LAN & backbones
 - Flows for traffic characterization and intrusion detection
 - Need access to network devices (e.g. routers, taps)
- Need to put together data from multiple sources
 - Different probes, different source & destinations, network-centric & end-to-end

Some Challenges for Active monitoring

- Bandwidth used, e.g. iperf etc. & apps
- For TCP tools: configuring windows at clients/servers and optimizing windows, streams
- Some lightweight tools (e.g. packet pairs) not effective at $\gg 1\text{Gbits/s}$
- Many tools tuned for shared TCP/IP nets not for dedicated circuits
- Simplifying use and understanding for end-user, automating problem detection & resolution, need close collaboration today

- Many measurement projects with different emphases, different communities
 - Passive (usually requires network control, used at borders and on backbones, e.g. MICSmon/Netflow, ISP/SNMP, SCNM)
 - Active
 - Lightweight (PingER, AMP, Surveyor, RIPE ...)
 - Medium weight (PiPES, NWS, IEPM-Lite ...)
 - Heavy weight/hi-perf (IEPM-BW, NTAF)
 - End-to-end vs net centric (skitter, macroscopic views)
 - Repetitive (PingER, AMP, IEPM, PiPES, NWS, NTAF, ...)
 - On demand, or non-production (NDT, NIMI, PiPES ...)
 - Dedicated hardware (AMP, RIPE, NDT, PlanetLab ...)
 - Hierarchical (e.g. AMP) vs Full mesh (e.g. PingER)
- For a table comparing 13 public domain infrastructures, see: www.slac.stanford.edu/grp/scs/net/proposals/infra-mon.html

- Sustaining deployment/operation in multi-agency / international world
- Scaling beyond hundreds of hosts very hard over the long term:
 - Hosts change, upgrade, new OS
 - No control over shared hosts
 - Depend on friendly admin contacts who may be busy, uninterested, have moved etc.
 - Policy/fears at remote site can make dedicated changes painful
 - web100 upgrades not coordinated with Linux upgrades
 - New TCP kernel upgrades not coordinated with OS upgrades
 - Hosts age, become measurement bottleneck
 - Need constant upgrades for dedicated hosts
 - Access policies change (pings & ports filtered)
 - Probes (iperf etc.) change: new features, patches
- Appropriate security

So Recognize

- Unrealistic to think multiple admin domains will all deploy one and the same infrastructure
 - Scaling and interests make unrealistic
- Multiple-domain, multi-infrastructures will be deployed
- Need to tie together heterogeneous collection of monitoring systems
 - Create a federation of existing NMs
 - Infrastructures work together
 - Share data with peer infrastructures and others using a common set of protocols for describing, exchanging & locating monitoring data (e.g. GGF NMWG)
 - Enables much improved overall view of network using multiple measurement types from multiple sources

- Measurement and Analysis for the Global Grid and Internet End-to-end performance
- Contribute to, utilize the GGF NMWG naming hierarchy and the schema definitions for network measurements
- Develop tools to allow sharing
 - Web services based
 - Integrate information from multiple sources
- Brings together several major infrastructure participants: LBNL (NTAP, SCNM), SLAC (IEPM-PingER/BW), Internet2 (PiPES, NDT), NCSC (NIMI), U Delaware, ESnet
- Will work with others, e.g. MonALISA, AMP, UltraLight, PPDG, StarLight, UltraScienceNet

Federation goals

- Appropriate security
- Interoperable
- Useful for applications, network engineers, scientists & end users
- Easy to deploy & configure
- As un-intrusive as possible
- As accurate & timely as possible
- Identify most useful features of each NMI to improve each NMI faster than working alone

NMI Challenges:

- Reduce “Wizard gap”
- Applications cross agency AND international funding boundaries (includes Digital Divide)
- Incent multi-disciplinary teams, including people close to scientists, operational teams
 - Make sure what is produced is used, tested in real environment, include deployment in proposals
- *Network management research historically underfunded, because it is difficult to get funding bodies to recognize as legitimate networking research, IAB*
- Without excellent trouble-shooting capabilities, the Grid vision will fail

More Information

- Some Measurement Infrastructures:
 - CAIDA list: www.caida.org/analysis/performance/measinfra/
 - AMP: amp.nlanr.net/, PMA <http://pma.nlanr.net>
 - IEPM/PingER home site: www-iepm.slac.stanford.edu/
 - IEPM-BW site: www-iepm.slac.stanford.edu/bw
 - NIMI: ncne.nlanr.net/nimi/
 - RIPE: www.ripe.net/test-traffic/
 - NWS: nws.cs.ucsb.edu/
 - Internet2 PiPES: e2epi.internet2.edu/
- Tools
 - CAIDA measurement taxonomy: www.caida.org/tools/
 - SLAC Network Tools: www.slac.stanford.edu/xorg/nmtf/nmtf-tools.html
- Internet research needs:
 - www.ietf.org/internet-drafts/draft-iab-research-funding-00.txt
 - www.slac.stanford.edu/grp/scs/net/talk03/lsn-jun03.ppt