

perfSONAR

Performance Monitoring Framework

Matt Zekauskas, matt@internet2.edu

GENI Measurement Workshop

June 26, 2009

Madison, Wisconsin



Credits

- Jeff Boote, Jason Zurawski, Aaron Brown, Eric Boyd at Internet2
- Lots of others at UDel, DANTE/GEANT, ESnet, RNP...

perfSONAR Introduction

- Most organizations perform monitoring and diagnostics of their own network
- Networking is increasingly a cross-domain effort
- Monitoring and diagnostics must also become a cross-domain effort

What is perfSONAR

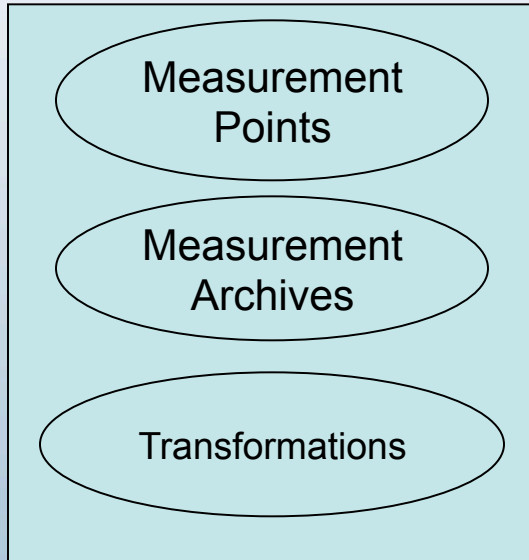
- An architecture & a set of protocols
 - Services Oriented Architecture (SOA)
 - Web Services Interfaces
 - Protocols being standardized in the OGF NMC-WG
- Also
 - A collaboration
 - Production network operators focused on designing and building tools that they will deploy and use on their networks to provide monitoring and diagnostic capabilities to themselves and their user communities.
 - Several interoperable software implementations
 - Java & Perl
 - A Federated set of Deployed Measurement Infrastructures

perfSONAR Architecture

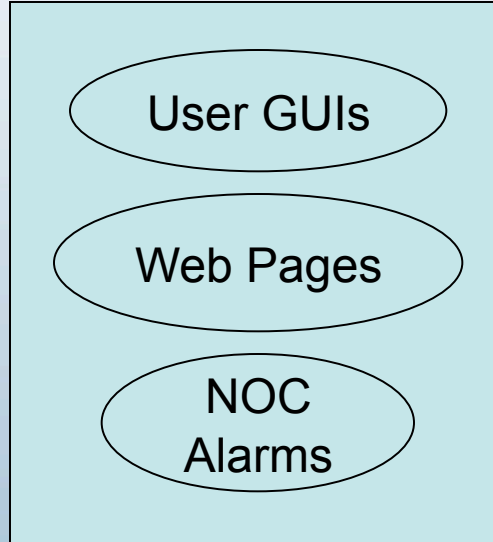
- Interoperable network measurement middleware (SOA):
 - Modular
 - Web services-based
 - Decentralized
 - Locally controlled
- Integrates:
 - Network measurement tools and archives
 - Data manipulation
 - Information Services
 - Discovery
 - Topology
 - Authentication and authorization
- Based on:
 - Open Grid Forum Network Measurement Working Group schema
 - Currently attempting to formalize specification of perfSONAR protocols in a new OGF WG (NMC)

perfSONAR Architecture

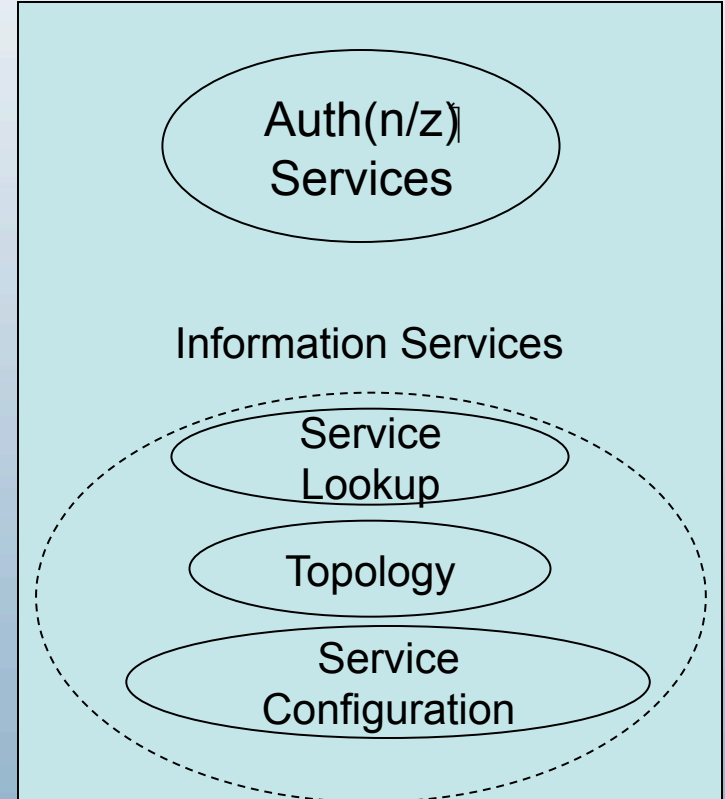
Data Services



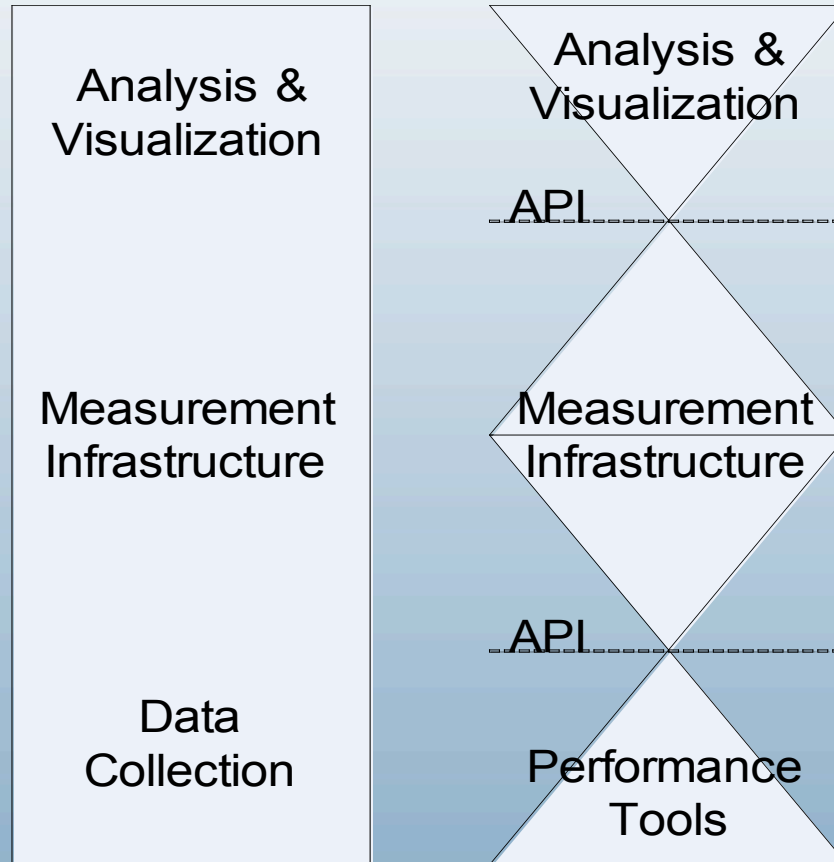
Analysis/Visualization



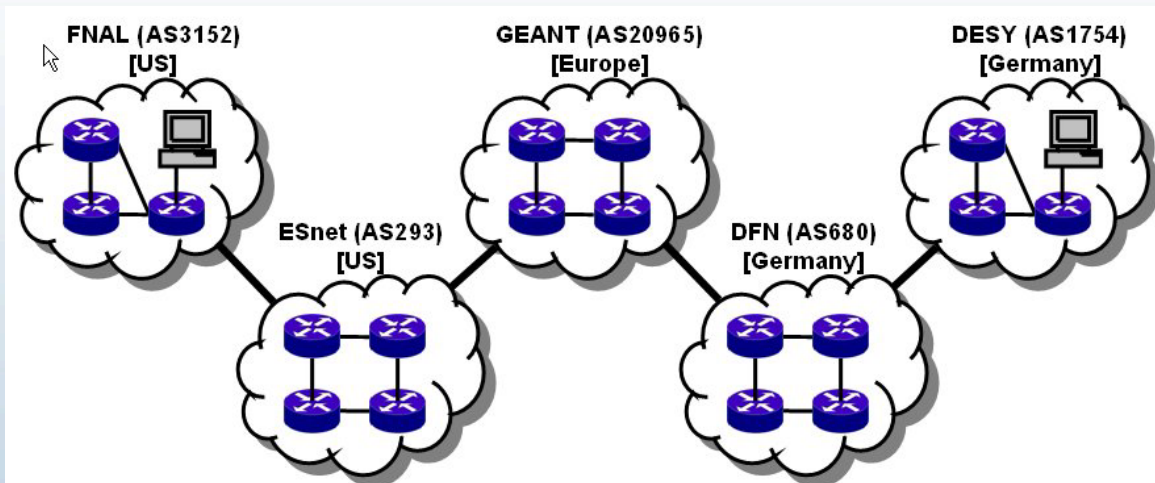
Infrastructure



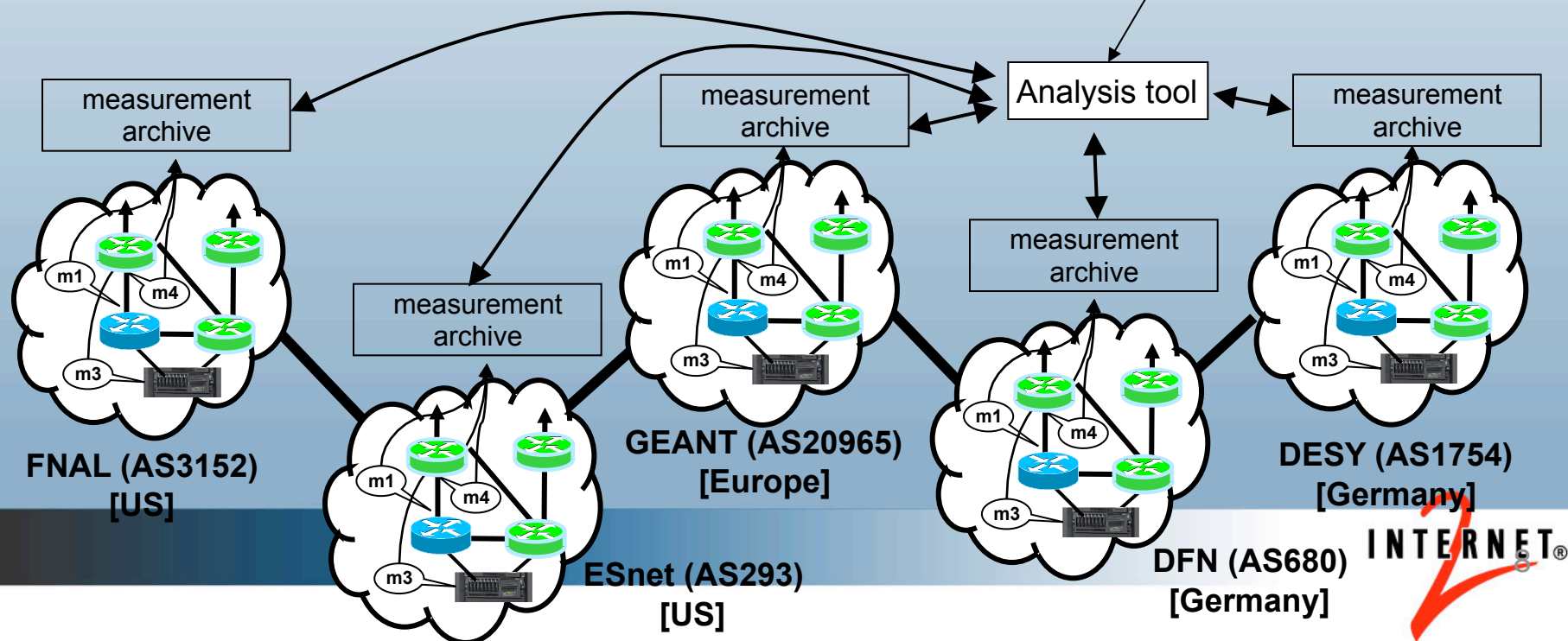
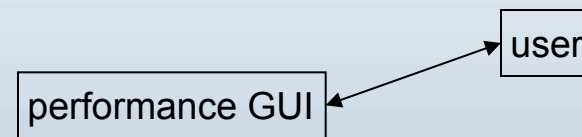
Decouple 3 phases of a Measurement Infrastructure



perfSONAR works E2E when All Networks Participate



Many collaborations are inherently multi-domain, so for an end-to-end monitoring tool to work everyone must participate in the monitoring infrastructure



perfSONAR Services

- Measurement Point Service
 - Enables the initiation of performance tests
- Measurement Archive Service
 - Stores and publishes performance monitoring results
- Transformation Service
 - Transform the data (aggregation, concatenation, correlation, translation, etc)

These services are specifically concerned with the job of network performance measurement and analysis

perfSONAR Services - MP

- **BWCTL (HADES)**
 - Wraps the BWCTL tool to perform regular throughput tests
- **BWCTL/OWAMP (perfSONAR-BUOY)**
 - Wraps the OWAMP (latency) and BWCTL (throughput) tools to perform tests
 - Packaged with an Archive (SQL based)
- **PingER**
 - Wraps the PingER latency tool
 - Packaged with an Archive (SQL based)
- **CMP (Command Line MP)**
 - Tool that wraps several tools (ping, iperf, traceroute, etc.)
 - Stores results to an SQL based MA
- **SSH/Telnet**
 - Using these protocols, attach to capable routing equipment to perform proxy commands

perfSONAR Services - MA

- **SNMP Measurement Archive**
 - Store SNMP Observations
 - Integrates directly with Cacti/Cricket/MRTG
 - Used in GUI production (Weathermap)
- **perfSONAR-BUOY Measurement Archive**
 - Store latency and throughput test results
 - Packaged with related MP to perform scheduled tests
- **PingER Measurement Archive**
 - Store results of PingER latency tests
 - Packaged with related MP to perform scheduled tests
- **RRD Measurement Archive**
 - Read results directly from RRD type databases (counters, gauges, etc)
- **SQL Measurement Archive**
 - Read results directly from SQL type databases (Postgres/MySQL)
- **Circuit Status Measurement Archive**
 - Collector polls for up/down with SNMP or TL1

Information Services

- Lookup Service
 - Allows the client to discover the existing services and other LS services.
 - Dynamic: services registration themselves to the LS and mention their capabilities, they can also leave or be removed if a service goes down.
- Topology Service
 - Make the network topology information available to the framework.
 - Find the closest MP, provide topology information for visualisation tools
- Authentication Service
 - Based on Existing efforts: Internet2 Shib, GN2-JRA5
 - Authentication & Authorization functionality for the framework
 - Users can have several roles, the authorization is done based on the user role.
 - Trust relationship between networks

These services are the infrastructure of the architecture concerned with the job of federating the available network measurement and diagnostic tools

perfSONAR Services - LS

- Stores both the location information and a directory of contained measurement information for each perfSONAR service
- Currently two layers
 - hLS – Home LS that accepts ‘registrations’ from MA/MP type services
 - gLS – Global LS that facilitates discovery of remote resources across domains
- Accepts direct queries about data or specific services
- Helps users locate remote resources on paths of interest

perfSONAR Services - TS

- Similar to LS – stores location and nature of various ‘topology’ elements (e.g. network resources)
- perfSONAR services register topology elements
 - Interfaces
 - Links
 - Nodes
- Related tools (DCN) are also starting to register

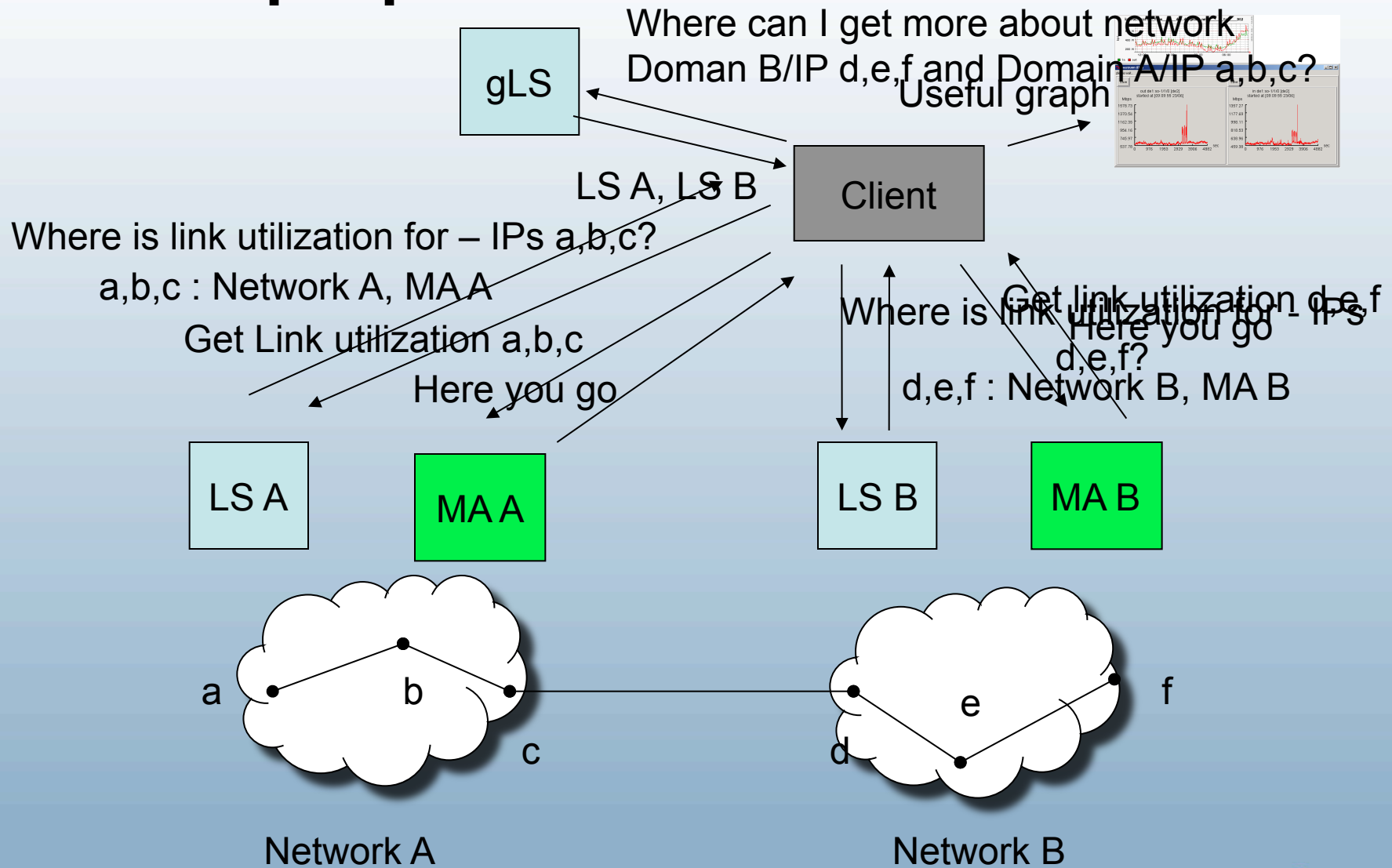
perfSONAR - Availability

- Java Tools
 - SVN Repository
 - RPM / Debian Packages
 - Use Java Build Tools (Ant, Maven)
- Perl Tools
 - SVN Repository
 - Source Packages
 - CPAN Perl System
 - Download services and all pre-reqs
 - RPM Packages
 - YUM/Up2date Coming Soon (helps manage pre-reqs)
 - Debian Packages (Coming soon)
 - NPToolkit

perfSONAR – Performance Node Live CD

- <http://code.google.com/p/perfsonar-ps/wiki/NPToolkit>
- Modification of disk you will use in this class
- Knoppix Based
- Requires additional configuration
 - Step by step process to configure perfSONAR
- Future Plans
 - Fedora Based (RPMs)
 - Inclusion of additional tools
 - Periodic OWAMP/BWCTL testing to remote sites
 - DCN Monitoring
 - Inclusion of additional GUIs
 - Client Focused
 - Domain Focused (Participation in Global Information Services)

Example perfSonar client interaction



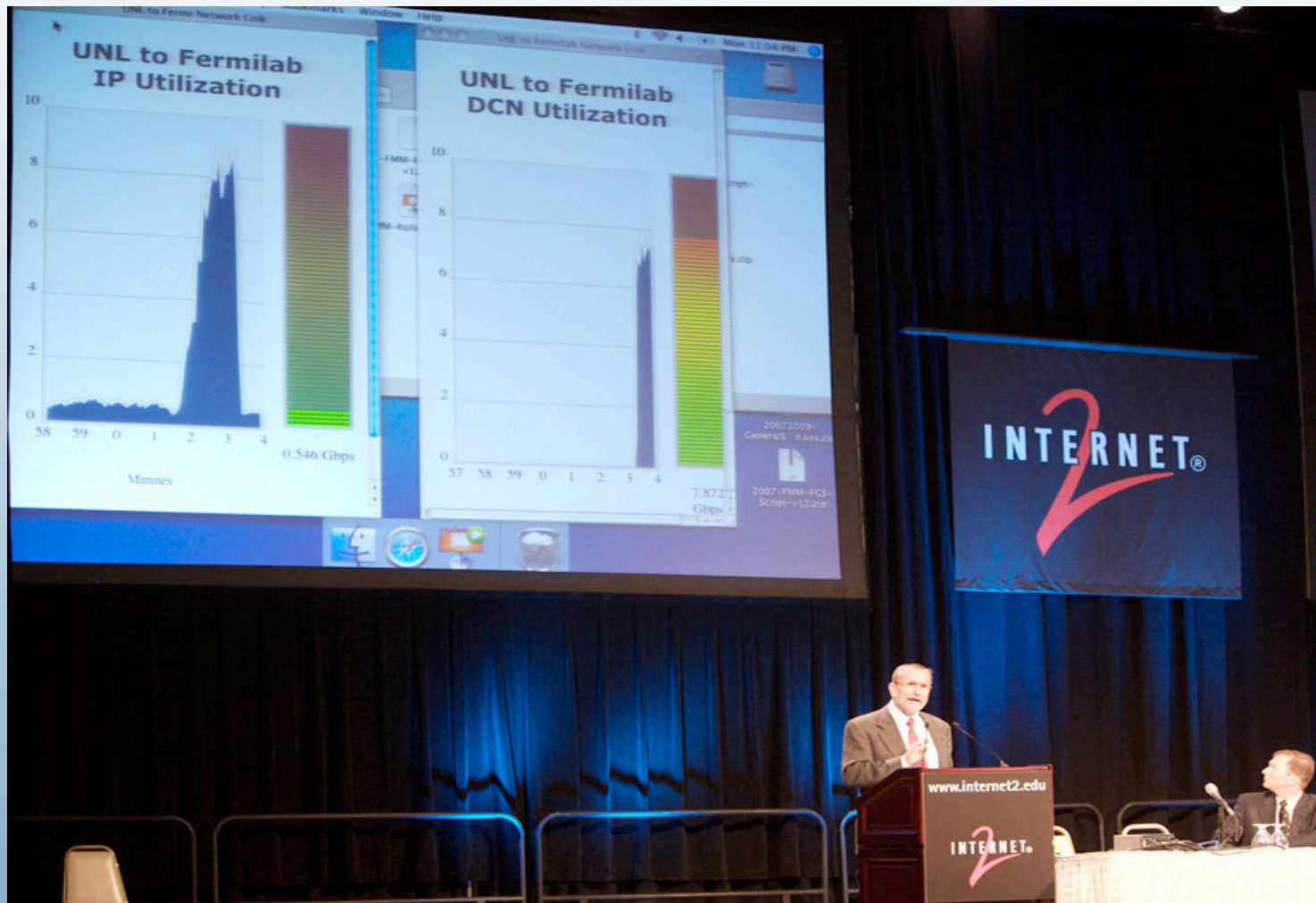
perfSONAR Client Developments

- Most tuned to specific services currently
- Different user focus (micro vs macro view)
- Client applications
 - perfSONAR-UI (acad.bg)
- Web Based
 - GMAPS (SLAC)
 - Domain Utilization Browser (ESnet)
 - pS-PS Weathermap (Internet2)
 - pingER Analysis (FNAL)
 - perfAdmin (Internet2)
 - E2EMon (DFN)

Data Views

- These are very quick GUI's we set up to show what is 'possible'
- The fact that the data is available (and known due to the LS) means that analysis clients are more easy to write and tune for specific user communities

SNMP MA In Action - perfOMeter



Gmaps (SNMP)

- <http://packrat.internet2.edu:8008/>

The screenshot displays the Gmaps (SNMP) web interface. On the left, a 'Communications Palette' lists various services for data fetching, including PingER (CERN, FNAL), OWAMP (Internet2), Topology (Internet2, FNAL), and Utilisation (SC07 Demo, Internet2, ES.net, GEANT, FERMI Lab, Georgia Tech, BWCTL). A search bar and a 'Fetch' button are at the bottom of the palette. The main area shows a map of the United States with two red location pins. An 'Info' window is open over the map, displaying a line graph titled 'utilisation' showing 'in' (green) and 'out' (blue) traffic in bps over a 24-hour period. The graph shows significant fluctuations, with peaks reaching 100 M bps. On the right, a 'Map' control bar includes 'Map', 'Satellite', and 'Hybrid' options, and an 'Interactive Palette' with a tree view of network components like 'Femilab', 'r-3-starlight-fnal-dmz', 'tengigabithemet4_1', 'tengigabithemet3_2', and 'r-cms-fcc2'. The bottom right corner features a 'NET' logo and a copyright notice: 'Map data ©2008 Europa Technologies - Terms of Use'.

Gmaps (pingER)

- <http://packrat.internet2.edu:8008/>

The screenshot displays the Gmaps (pingER) web application interface. On the left, the 'Communications Palette' allows users to select services for data fetching, including PingER (CERN, FNAL), OWAMP (Internet2 perfSONARBUOY), Topology (Internet2, FNAL), and Utilisation (SC07 Demo, Internet2, ES.net, GEANT, FERMI Lab, Georgia Tech). Below this, the 'BWCTL' section offers a list of services like GEANT2 RRD, Cynet RRD, GARR-new RRD, GRNET RRD, and ISTF-J RRD. A 'Fetch' button is at the bottom of the palette.

The main map shows the Atlantic Ocean with several red location pins. An 'Info' window for 'pinger' is open, displaying a line graph of network performance metrics over a 24-hour period. The graph plots four metrics: iqrIpd (red), meanIpd (orange), meanRtt (black), and minRtt (green). The y-axis represents milliseconds (msec) from 30 to 70. The x-axis shows time from 00:00 to 18:00. The graph shows a significant spike in meanIpd around 06:00.

On the right, the 'Interactive Palette' lists various network paths and their status, such as 'path (checked)', 'chapuza.cem.ch (checked)', and several paths involving 'netflow01.pic.es', 'ic.gvm.triumf.ca', 'cmswn300.fnal.gov', 'lhc-opn-mon-fzk.gridka.de', 'mastercr.crnaf.infn.it', 'cs-enmon-1.cem.ch', 'dct00.usatlas.bnl.gov', and 'ccxfert01.in2p3.fr'. Each path has a checkmark indicating its status.

Gmaps (bwctl)

- <http://packrat.internet2.edu:8008/>

The screenshot displays the Gmaps (bwctl) web application interface. The main map shows the United States with several red location pins. The 'Communications Palette' on the left offers various service selection options, including PingER, OWAMP, Topology, and Utilisation. The 'bwctl' window shows a bar chart of bandwidth usage in bps over time, with a legend indicating 'out' traffic. The 'Interactive Palette' on the right lists IP addresses and their checked status.

Communications Palette

Please pick a Service to fetch data from:

PingER

- CERN PingER MA
- FNAL PingER MA

OWAMP

- Internet2 perfSONARBUOY

Topology

- Internet2 Topology
- FNAL PingER Topology

Utilisation

- SC07 Demo SNMP MA
- Internet2 SNMP MA
- ES.net SNMP MA
- GEANT RRDMA
- FERMI Lab SNMP MA
- Georgia Tech SNMP MA

BWCTL

- Internet2 perfSONARBUOY

Or select from the following list:

- GEANT2 RRD MA
- Cynet RRD MA
- GARR-new RRD MA
- GRNET RRD MA
- ISTF-J RRD MA

Fetch

bwctl

900 M
800 M
700 M
600 M

bps

Thu Fri Sat Sun Mon

■ out

Interactive Palette

- path [5 checked]
- 64.57.17.18 [4 checked]
- 207.75.165.145:protocol=TCP>windowSize=6m:interval=2
- 198.124.252.109:protocol=TCP>windowSize=6m:interval=2
- 130.199.3.7:protocol=TCP>windowSize=6m:interval=2
- 131.225.110.144:protocol=TCP>windowSize=6m:interval=2
- 207.75.165.145 [4 checked]
- 198.124.252.109 [4 checked]
- 130.199.3.7 [4 checked]
- 131.225.110.144 [4 checked]



pingER Analysis

- <http://lhcopnmon1-mgm.fnal.gov:9090/pinger/gui>

Generate Network Performance Graphs

You can cut down the list of choices shown in the scroll-list below by entering the perl's style regular expression that matching your choice. You can use * to match any string. Or, for example *hep.net\$* will match any node ending in *hep.net*. Also, you can set the desirable packetsize (for example 100\$ for 100 bytes).

Source Contains: *	Destination Contains: *	Packetsize:	Match It
--------------------	-------------------------	-------------	----------

Links: lhcopnmon1-mgm.fnal.gov -> cithep130.ultralight.org (1000) lhcopnmon1-mgm.fnal.gov -> cmssrm.hep.wisc.edu (1000) lhcopnmon1-mgm.fnal.gov -> newmon.bnl.gov (1000) lhcopnmon1-mgm.fnal.gov -> phedex.rcac.purdue.edu (1000) lhcopnmon1-mgm.fnal.gov -> pinger.fnal.gov (1000) lhcopnmon1-mgm.fnal.gov -> pinger.slac.stanford.edu (1000) lhcopnmon1-mgm.fnal.gov -> t2.unl.edu (1000) lhcopnmon1-mgm.fnal.gov -> tier2.ihepa.ufl.edu (1000) lhcopnmon1-mgm.fnal.gov -> tier2.ucsd.edu (1000) lhcopnmon1-mgm.fnal.gov -> www.cern.ch (1000) lhcopnmon1-mgm.fnal.gov -> www.cmsaf.mit.edu (1000)	From: Month: Jun Day: 9 Year: 2008	To: Month: Jun Day: 15 Year: 2008	Time Zone: GMT-12 GMT-11 GMT-10
--	--	---	---

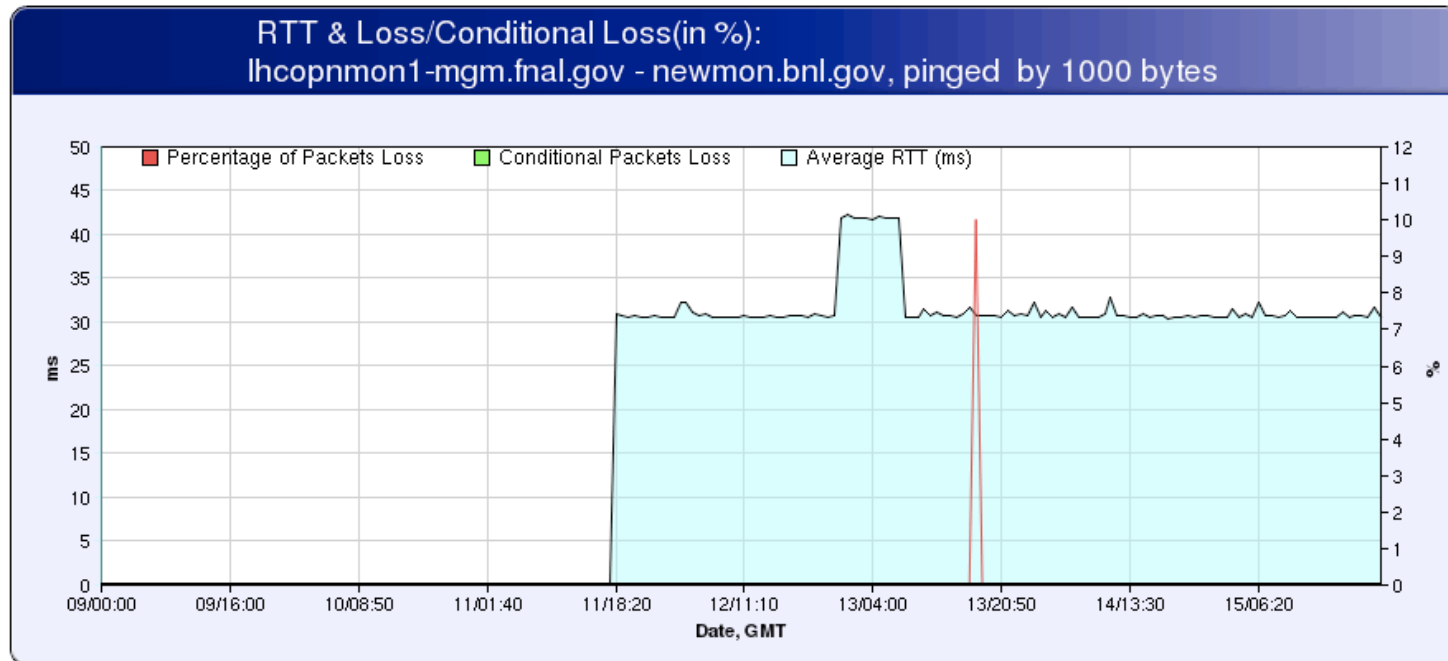
Graph parameters:	Upper RTT(or IPD) Limit: Auto Range 0...10 ms 0...20 ms	File Format: PNG	Graph type: LINES	<input type="checkbox"/> Any Graph(s) -> 1 file <input checked="" type="checkbox"/> Separate file for each Graph
Type of metric: RTT/Loss/Conditional Loss				

pingER Analysis

- <http://lhcopnmon1-mgm.fnal.gov:9090/pinger/gui>

Network Performance Graphs

Graph: *lhcopnmon1-mgm.fnal.gov - newmon.bnl.gov* by 1000 bt was generated over 120 data points



Done

INTERNET®

perfSONAR-BUOY (bwctl)

- <http://ndb1.internet2.edu/cgi-bin/bwctl.cgi?name=OFFICEMESHBTCP4>

Sun Jun 15 22:02:55 UTC 2008 --- Mon Jun 16 22:02:55 UTC 2008

LHC Sites - IPv4 TCP Throughput

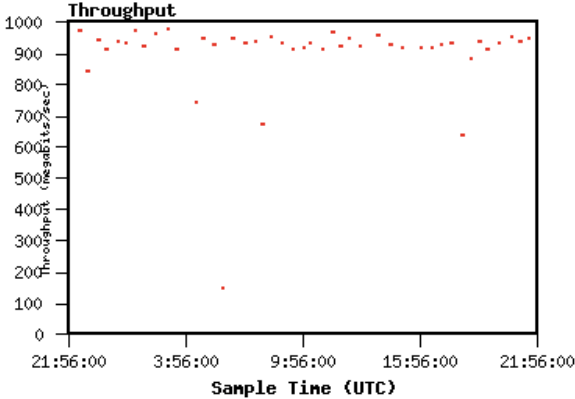
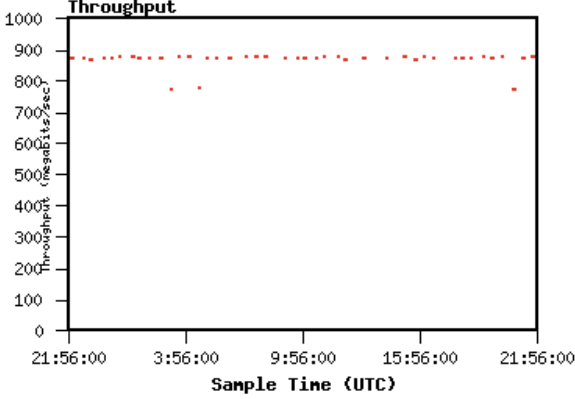
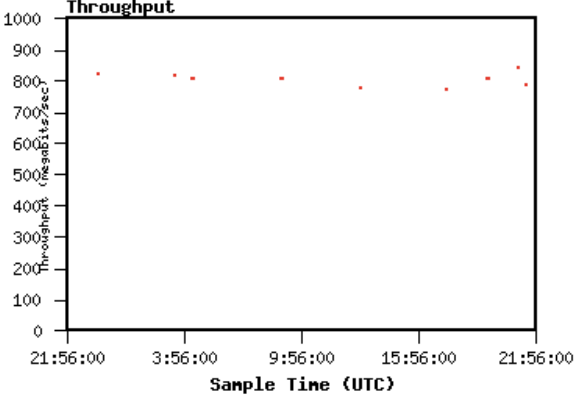
[Throughput(Mbps) / Date&Time]

BWTCP4	Senders				
	ann arbor	brookhaven national laboratory	chicago	fermi national accelerator laboratory	esnet
Receivers	ann arbor	*/ *	180.1 / Mon Jun 16 21:00:48 UTC 2008	159.0 / Mon Jun 16 21:58:28 UTC 2008	*/ *
	brookhaven national laboratory	117.4 / Mon Jun 16 21:55:42 UTC 2008	785.8 / Mon Jun 16 21:29:37 UTC 2008	535.6 / Mon Jun 16 20:52:57 UTC 2008	*/ *
	chicago	315.4 / Mon Jun 16 21:05:08 UTC 2008	876.9 / Mon Jun 16 21:46:45 UTC 2008	623.6 / Mon Jun 16 21:42:12 UTC 2008	*/ *
	fermi national accelerator laboratory	334.7 / Mon Jun 16 20:40:26 UTC 2008	389.2 / Mon Jun 16 21:41:07 UTC 2008	870.7 / Mon Jun 16 20:00:35 UTC 2008	*/ *
	esnet	370.8 / Mon Jun 16 20:56:34 UTC 2008	914.3 / Mon Jun 16 22:01:37 UTC 2008	947.9 / Mon Jun 16 21:58:49 UTC 2008	540.1 / Mon Jun 16 22:00:05 UTC 2008

Done

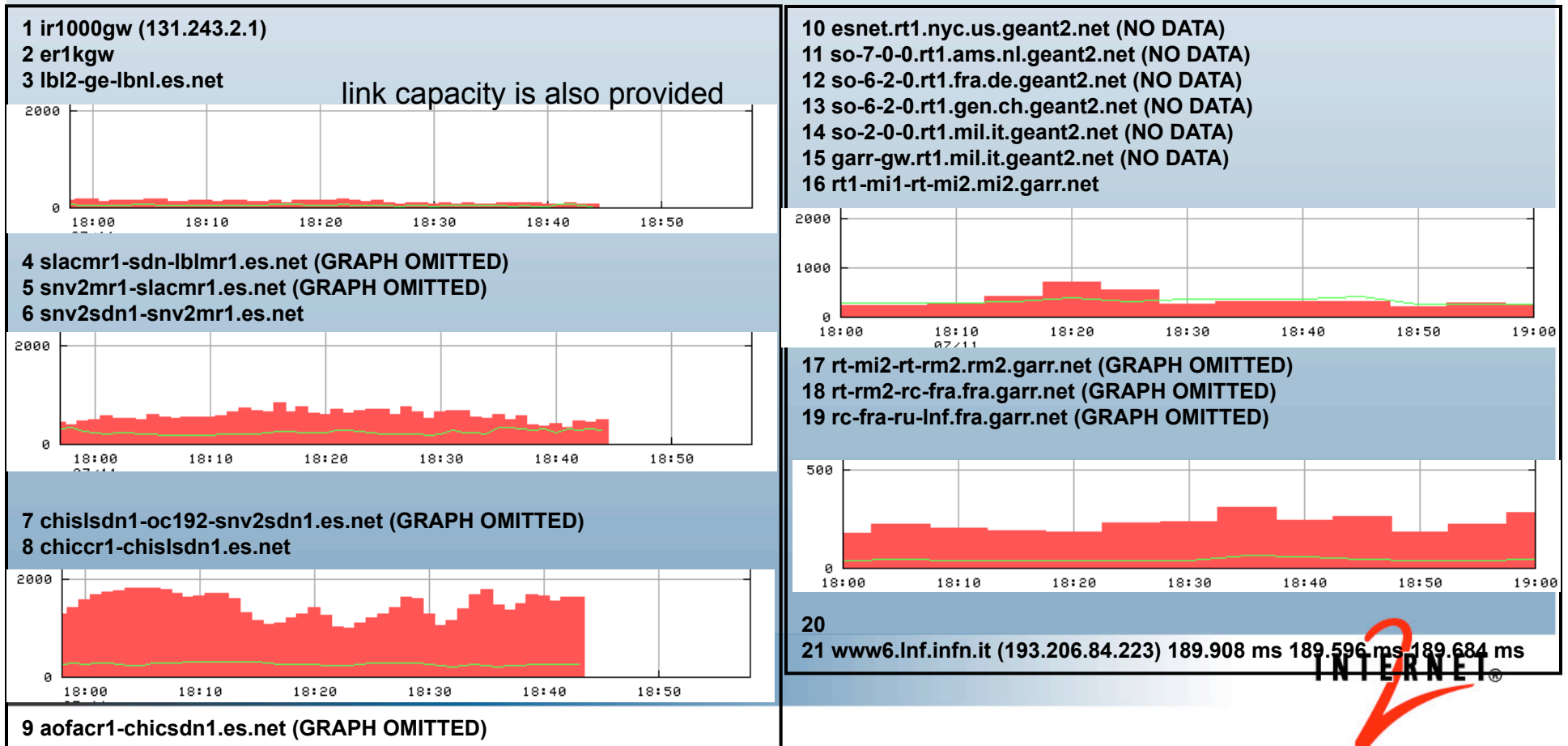


perfSONAR-BUOY (bwctl)

STARLIGHT ↔ BNL	Sender: brookhaven national laboratory	Receiver: brookhaven national laboratory
ESNet		No data available.
CHIC ↔ BNL	Sender: brookhaven national laboratory	Receiver: brookhaven national laboratory
Chicago		

Traceroute Visualizer

- Forward direction bandwidth utilization on application path from LBNL to INFN-Frascati (Italy)
 - traffic shown as bars on those network device interfaces that have an associated MP services (the first 4 graphs are normalized to 2000 Mb/s, the last to 500 Mb/s)



E2Emon - Monitoring Circuits

E2E Link Monitoring System

http://cnmdev.lrz-muenchen.de/e2e/lhc/mon/G2_E2E_index_PROD.html

Latest Headlines Personal Finance Internet2 JRA1-PAT Abilene SCinet Travel Google Code - Sum... WikiHome - JotSpot ...

always... Google... Home... Interne... MailTags garfiel... Kagi - ... macos... Upload... E2E ...

Domain view

- [CANARIE](#)
- [CERN](#)
- [CESNET](#)
- [DFN](#)
- [ESNET](#)
- [FERMI](#)
- [GARR](#)
- [GEANT2](#)
- [HOPI](#)
- [IN2P3](#)
- [INTERNET2](#)
- [NETHERLIGHT](#)
- [PSNC](#)
- [REDIRIS](#)
- [RENATER](#)
- [SWITCH](#)
- [USLHCNET](#)

Project view

- [DEISA](#)
- [IGTMD](#)
- [LHCOPN](#)
- [LHCT1T2](#)

Availability Statistics

[Current Month](#)

[All Months](#)

E2E Links for Project LHCOPN (Prod.)

E2E Link ID	State Oper	State Admin	Additional Info
CERN-ASGC-LHCOPN-001	Up	Normal Oper.	Error: E2E Link is not contiguous (End Point missing or gap found) Warning: Operational state is not known for all involved links Warning: Administrative state is not known for all involved links
CERN-ASGC-LHCOPN-002	Up	Normal Oper.	Error: E2E Link is not contiguous (End Point missing or gap found) Warning: Operational state is not known for all involved links Warning: Administrative state is not known for all involved links
CERN-BNL-LHCOPN-001	Up	Normal Oper.	Error: E2E Link is not contiguous (End Point missing or gap found) Error: E2E Link is not contiguous (End Point missing or gap found)
CERN-BNL-LHCOPN-002	Up	Normal Oper.	Warning: Operational state is not known for all involved links Warning: Administrative state is not known for all involved links
CERN-CNAF-LHCOPN-001	Up	Normal Oper.	
CERN-FERMI-LHCOPN-001	Up	Normal Oper.	Warning: Operational state is not known for all involved links Warning: Administrative state is not known for all involved links
			Warning: Operational state is not known for all involved links

Done

Oper. State: **Up**
 Admin. State: **Normal Oper.**

Domain	CERN			USLHCNET			
Link Structure	EP	←.....→	DP	↔	DP	←.....
Type	EndPoint	ID Part.Info	ID Part.Info	Demarc	Domain Link	Demarc	ID Part.Info
Local Name	CERN-T0	S513-C-BE1	CERN-FERMI-LHCOPN-001-GVA-CERN	USLHCNET-GEN	CERN-FERMI-LHCOPN-001-GVA-CHI	USLHCNET-CHI	CERN-FERMI-LHCOPN-001-CHI-ESNET
State Oper.	-	Up	Up	-	Up	-	Up
State Admin.	-	Normal Oper.	Normal Oper.	-	Normal Oper.	-	Normal Oper.
Timestamp	-	2007-04-08 T05:04:08+02:00	2007-04-08 T05:04:11+02:00	-	2007-04-08 T05:04:53+02:00	-	2007-04-08 T05:03:59+02:00

Page generated

ESNET				FERMI				
.....→	DP	↔	DP	←.....→	DP	↔	EP
ID Part.Info	Demarc	Domain Link	Demarc	ID Part.Info	ID Part.Info	Demarc	Domain Link	EndPoint
CERN-FERMI-LHCOPN-001-STARLIGHT-Tail	ESNET-STARLIGHT	CERN-FERMI-LHCOPN-001-FERMI-STARLIGHT	ESNET-FERMI	CERN-FERMI-LHCOPN-001-Site-Tail	md8	FERMI-ESNET	md2	FERMI-T1
Up	-	Up	-	Up	Up	-	Up	-
Normal Oper.	-	Normal Oper.	-	Normal Oper.	Normal Oper.	-	Normal Oper.	-
2007-04-08 T01:40:37.0	-	2007-04-08T01:40:37.0	-	2007-04-08 T01:40:37.0	2007-04-08 T01:40:01.0-6:00	-	2007-04-08 T01:40:01.0-6:00	-



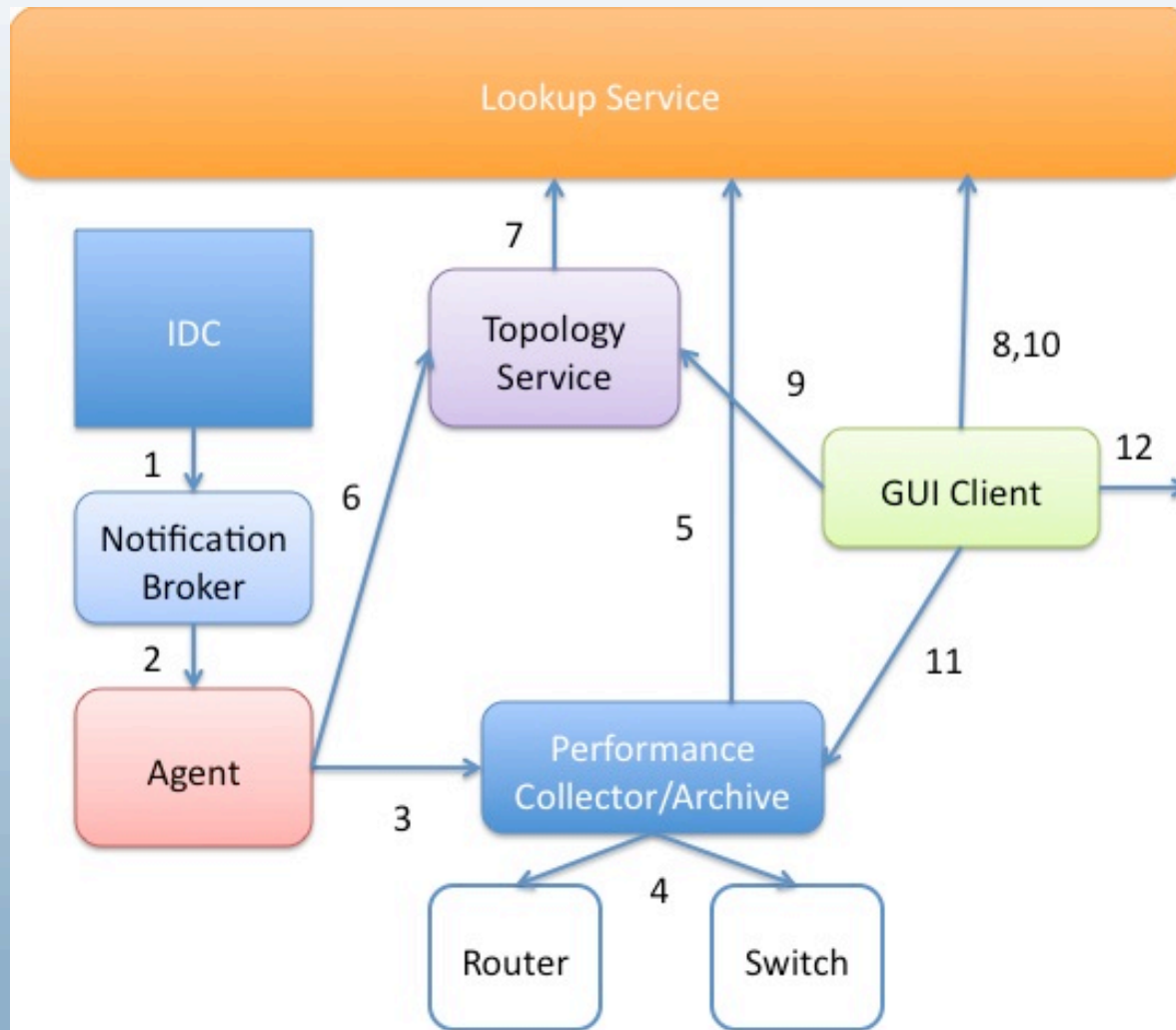
perfSONAR GUI List

- <http://www.perfsonar.net/activeServices>
 - Listing of many types of active services
 - Graphs
 - Live Tests Points
- <http://psvis0.internet2.edu:8008/>
 - Gmaps interface
- <http://perfsonar.acad.bg/>
 - Java based visualization tool
- http://cnmdev.lrz-muenchen.de/e2e/lhc/mon/G2_E2E_index_ALL.html
 - E2emon link monitoring
- <https://performance.es.net/cgi-bin/level0/perfsonar-trace.cgi>
 - Traceroute visualization

Near-term Futures

- Harden packages
- Circuit monitoring (more than up/down)
<http://code.google.com/p/perfsonar-ps/wiki/CircuitMonitoring.../perfsonar-ps/wiki/CircuitMonitoringMoreDetails>
- Adding federated authentication
- Exposing more Internet2 datasets

Circuit Monitoring



Conclusions

Summary

- Open Source licenses and development model
- Interfaces for any application to consume the data
- Partners are committed to supporting these tools
- Potential GENI base, or provide supplemental data

More information

<http://www.internet2.edu/performance/pS-PS>

Internet2 Community
Performance WG

<https://mail.internet2.edu/wws/info/performance-announce>

www.internet2.edu

