perfSONAR
Performance Monitoring Framework

Matt Zekauskas, matt@internet2.edu

GENI Measurement Workshop
June 26, 2009

Madison, Wisconsin

INTERNET®

Credits

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

INTERNET®

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

INTERNET®

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

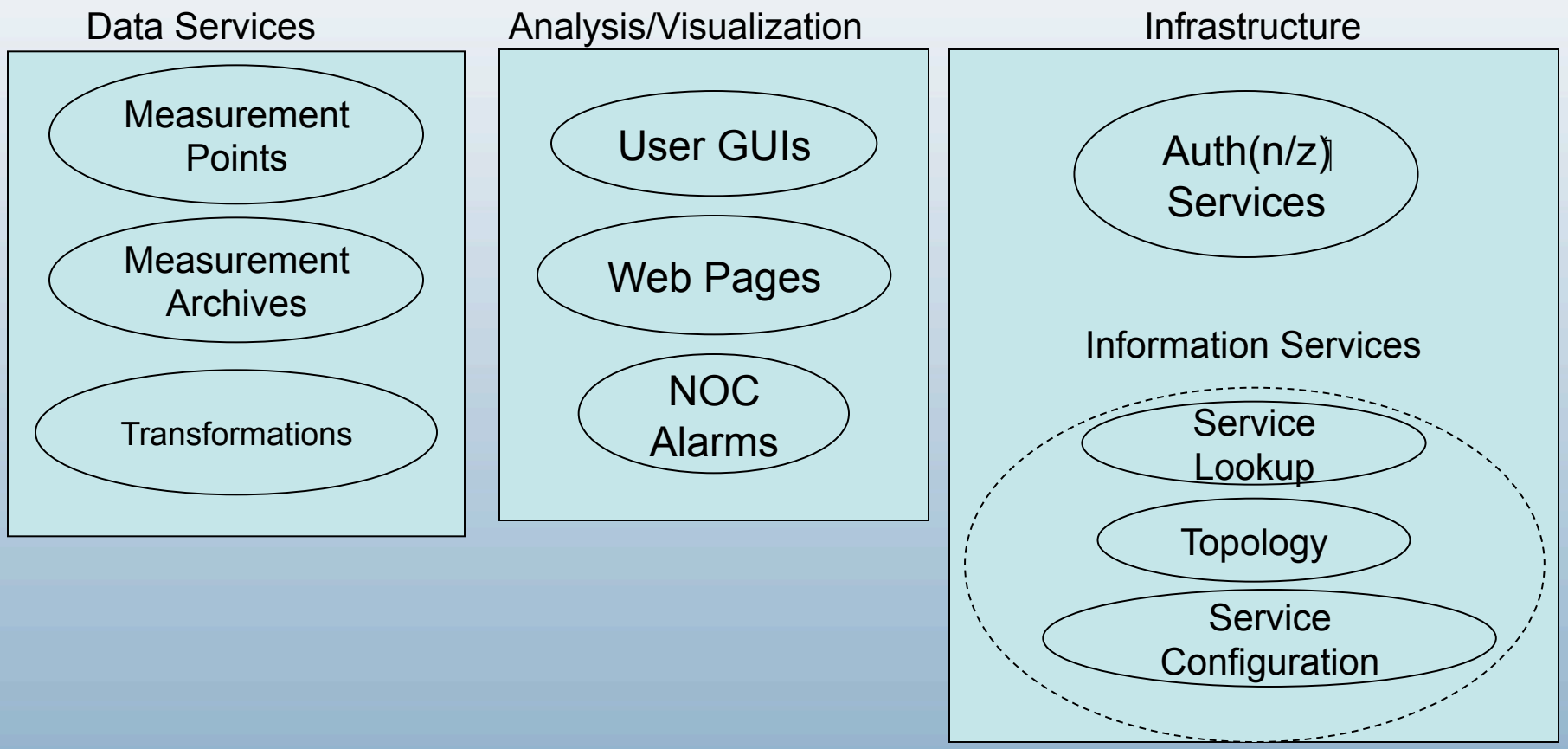
INTERNET®

perfSONAR Architecture

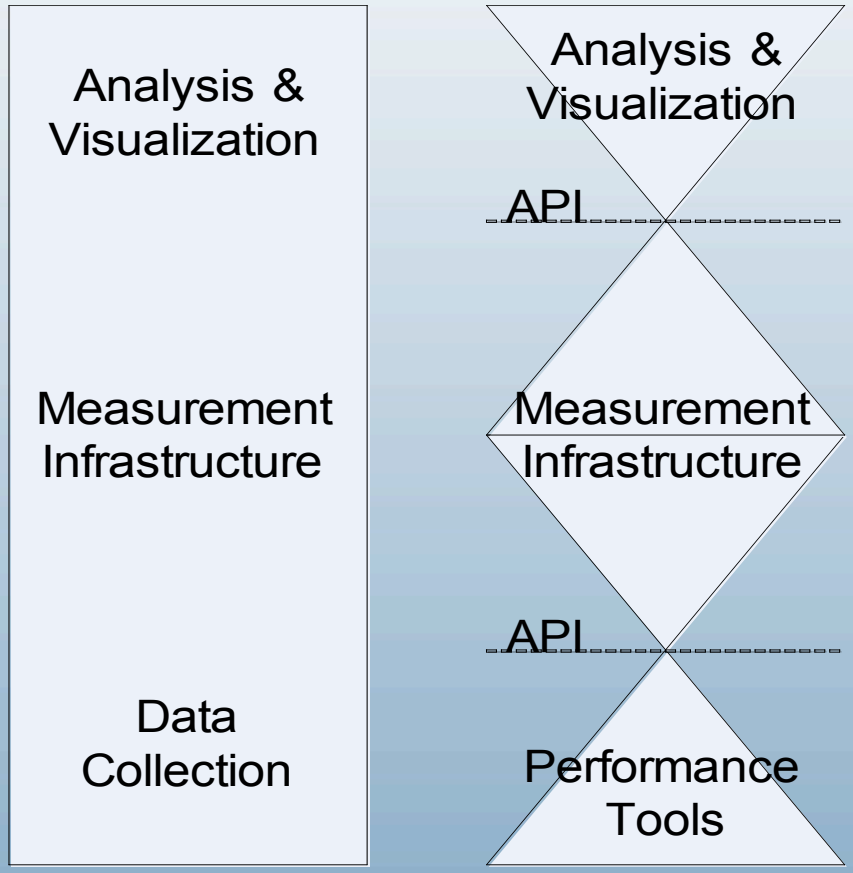
- Interoperable network measurement middleware (SOA):
  - Modular
  - Web services-based
  - Decentralized
  - Locally controlled
- Integrates:
  - Network measurement tools and archives
  - Data manipulation
  - Information Services
    - Discovery
    - Topology
    - Authentication and authorization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

INTERNET2®

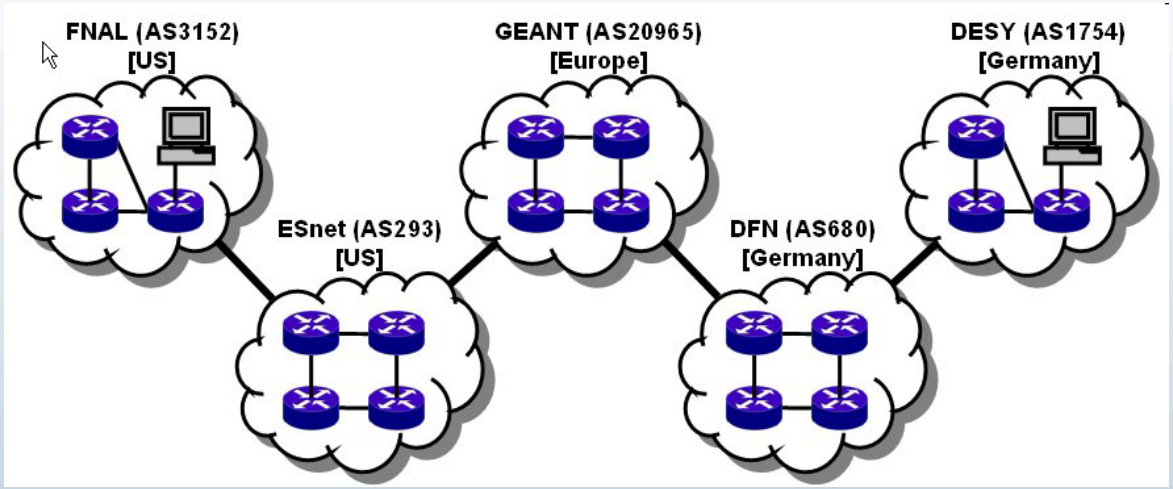
perfSONAR Architecture



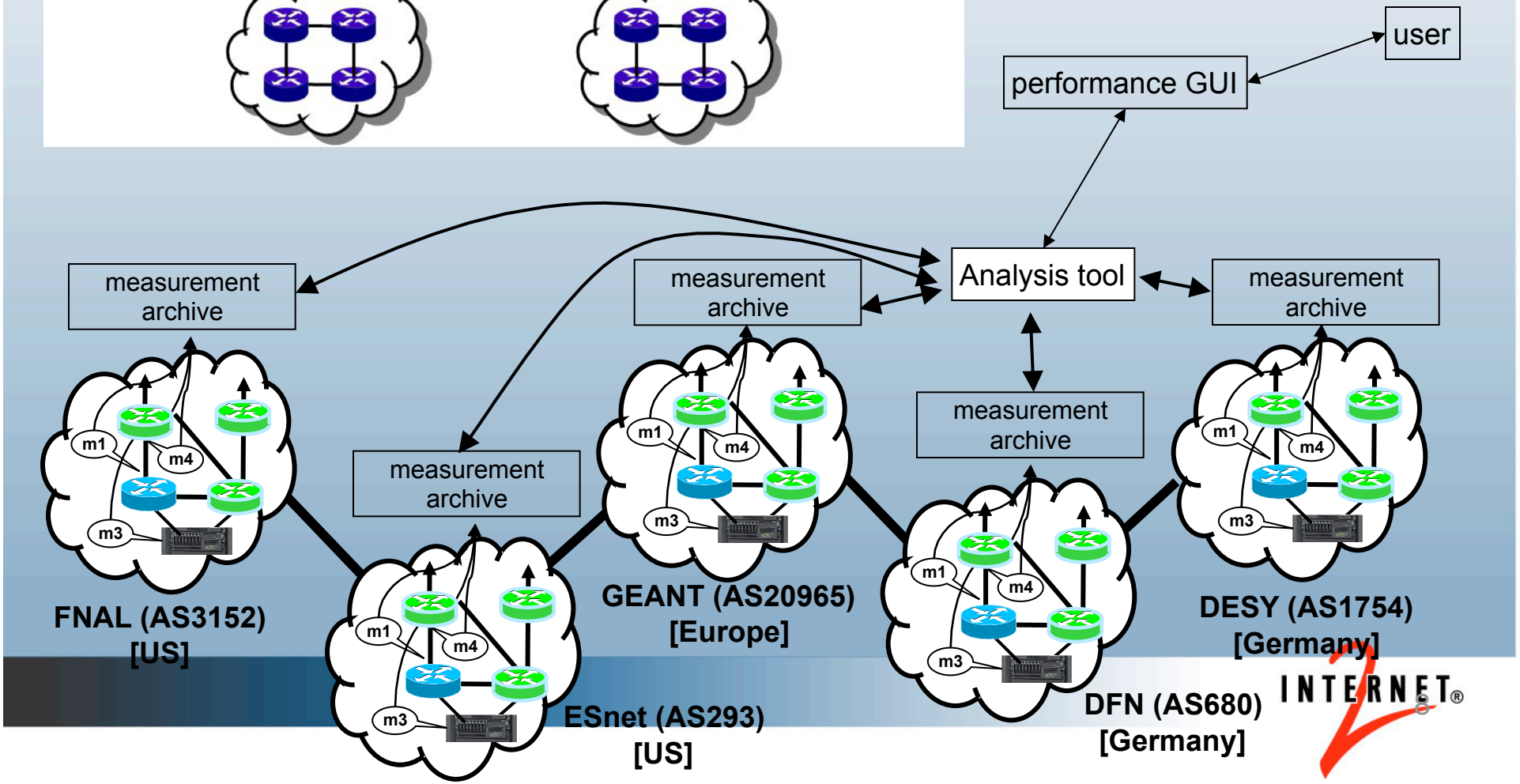
Decouple 3 phases of a Measurement Infrastructure



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

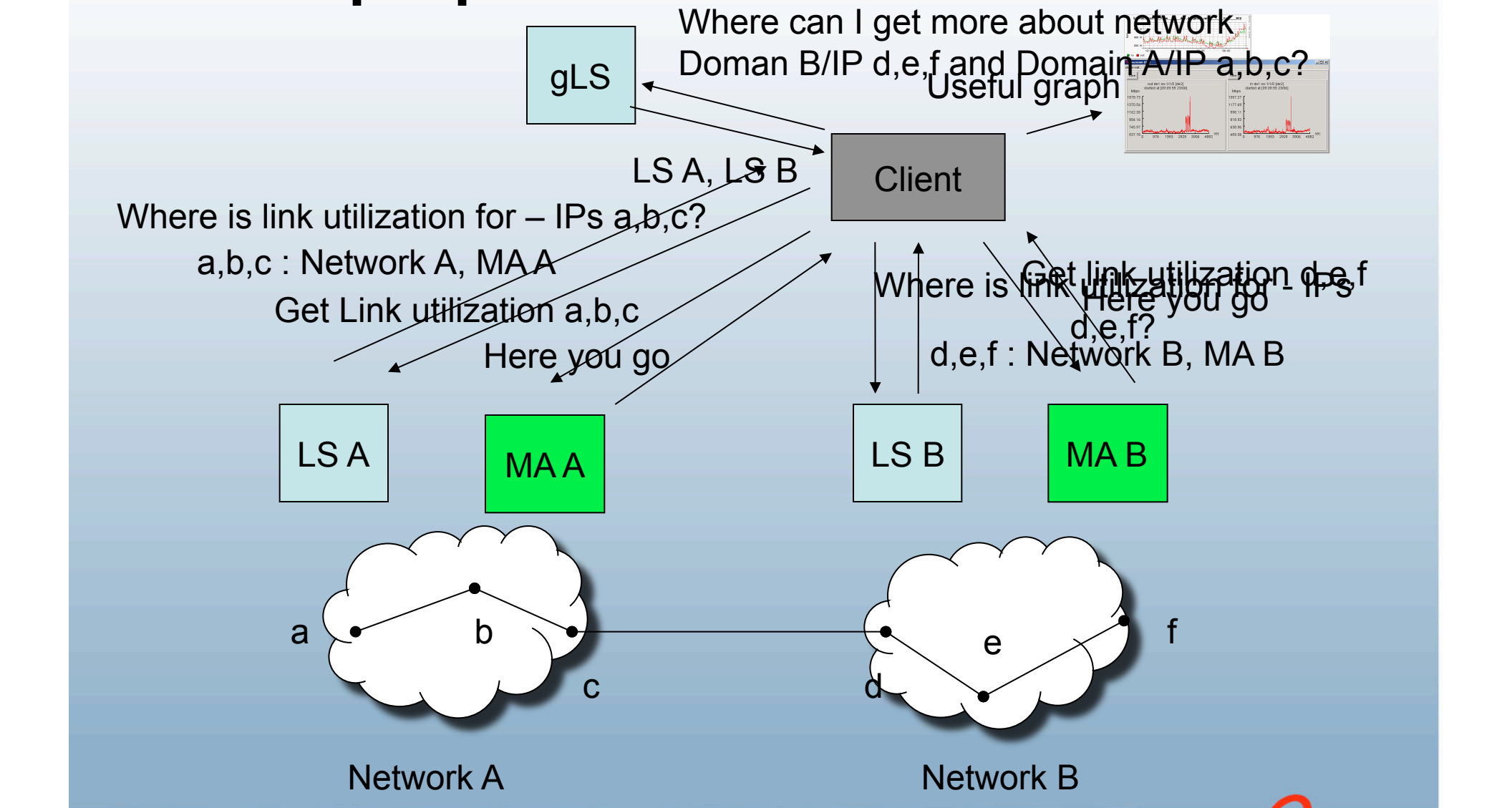
perfSONAR - Availability

- 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 configurion
  - 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)

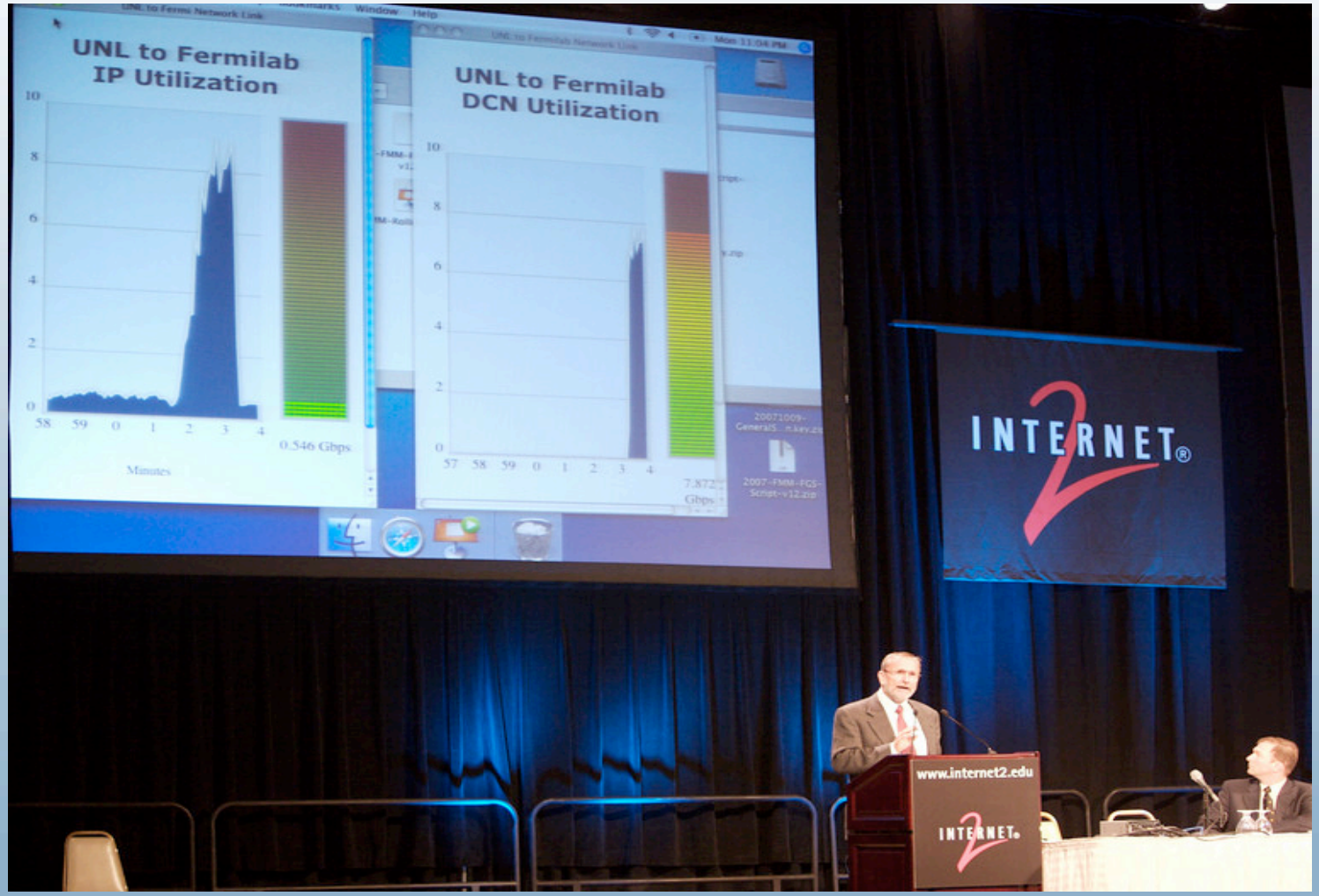
INTERNET®

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

INTERNET®

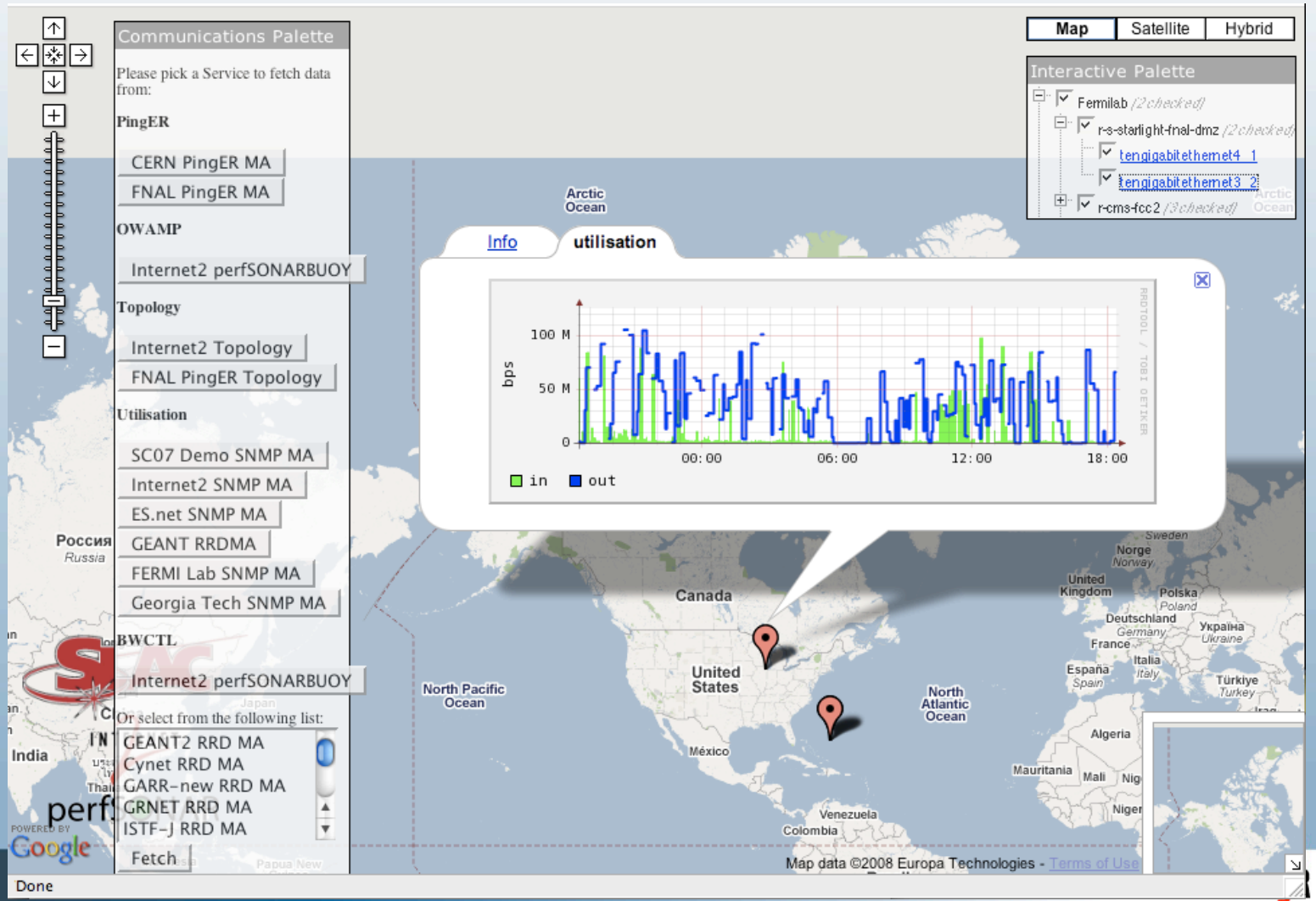
SNMP MA In Action - perfOMeter



INTERNET®

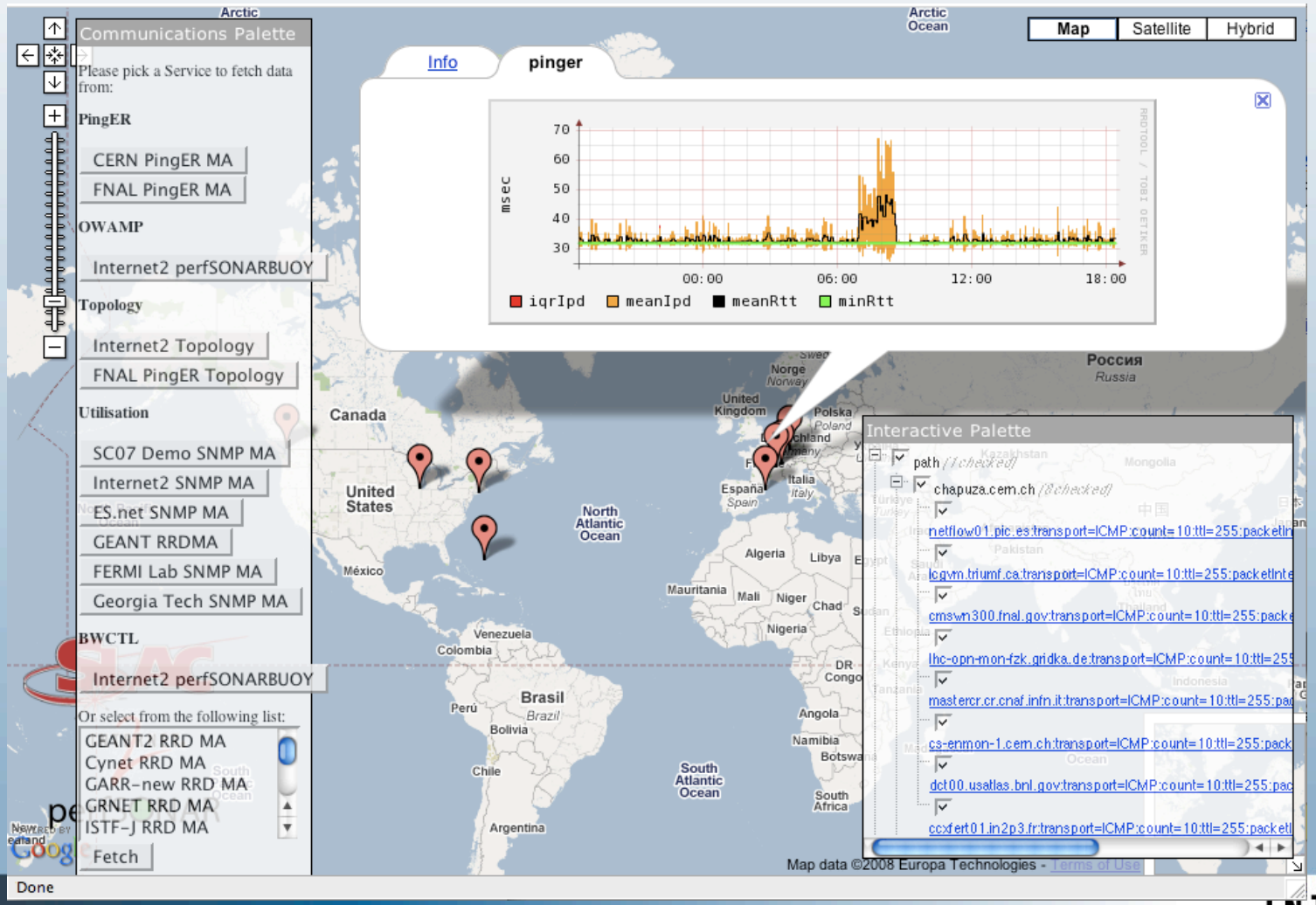
Gmaps (SNMP)

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



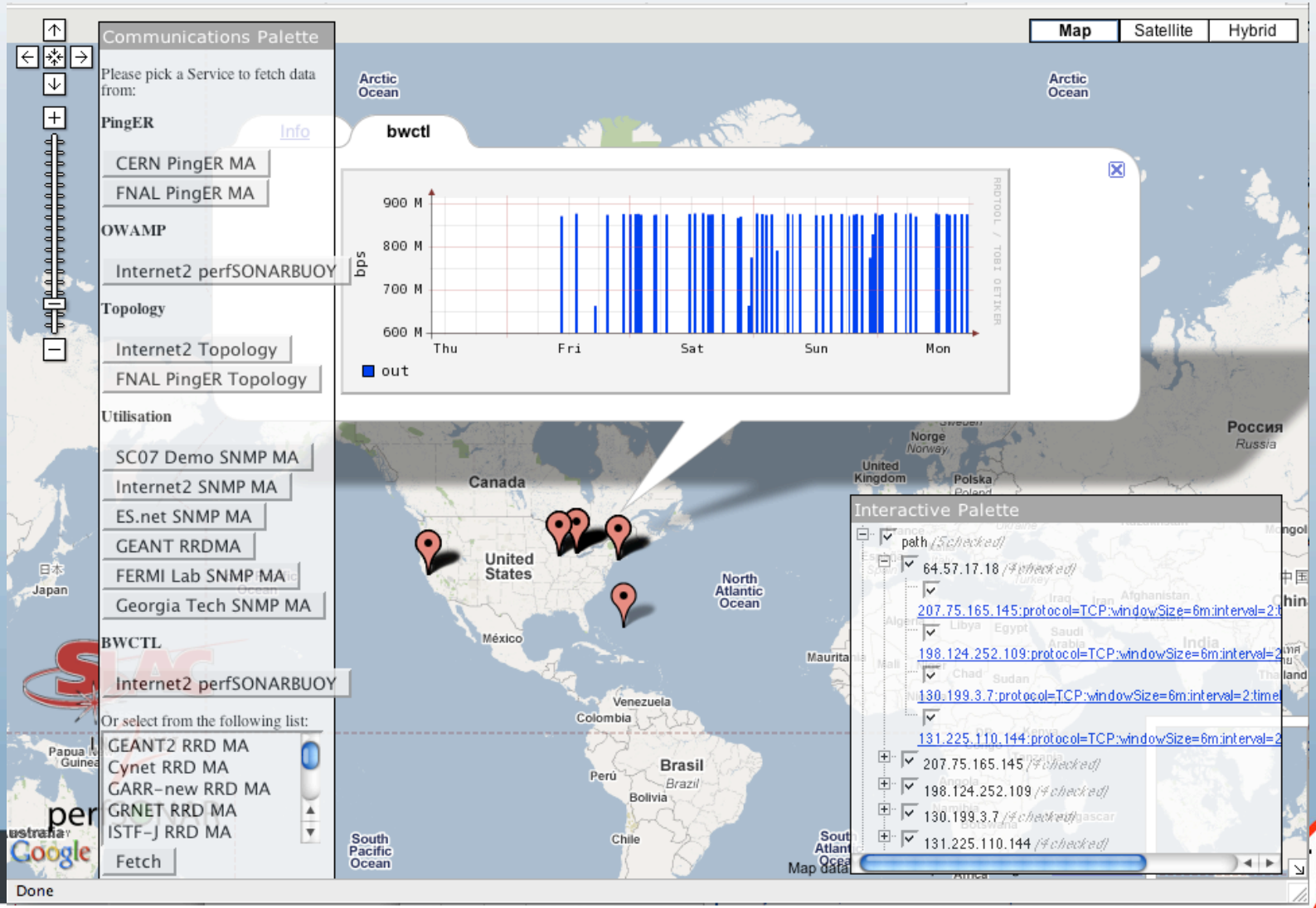
Gmaps (pingER)

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



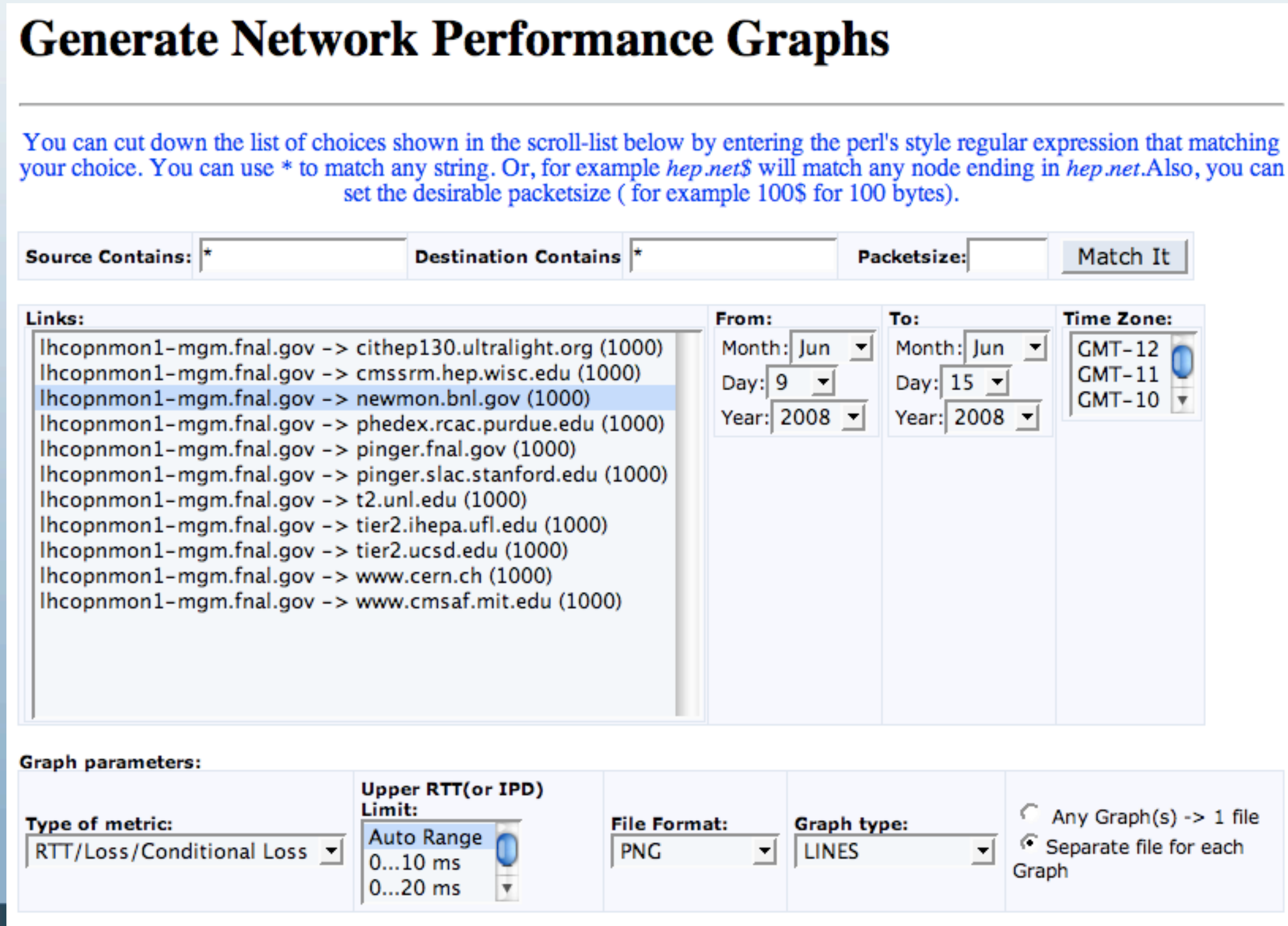
Gmaps (bwctl)

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



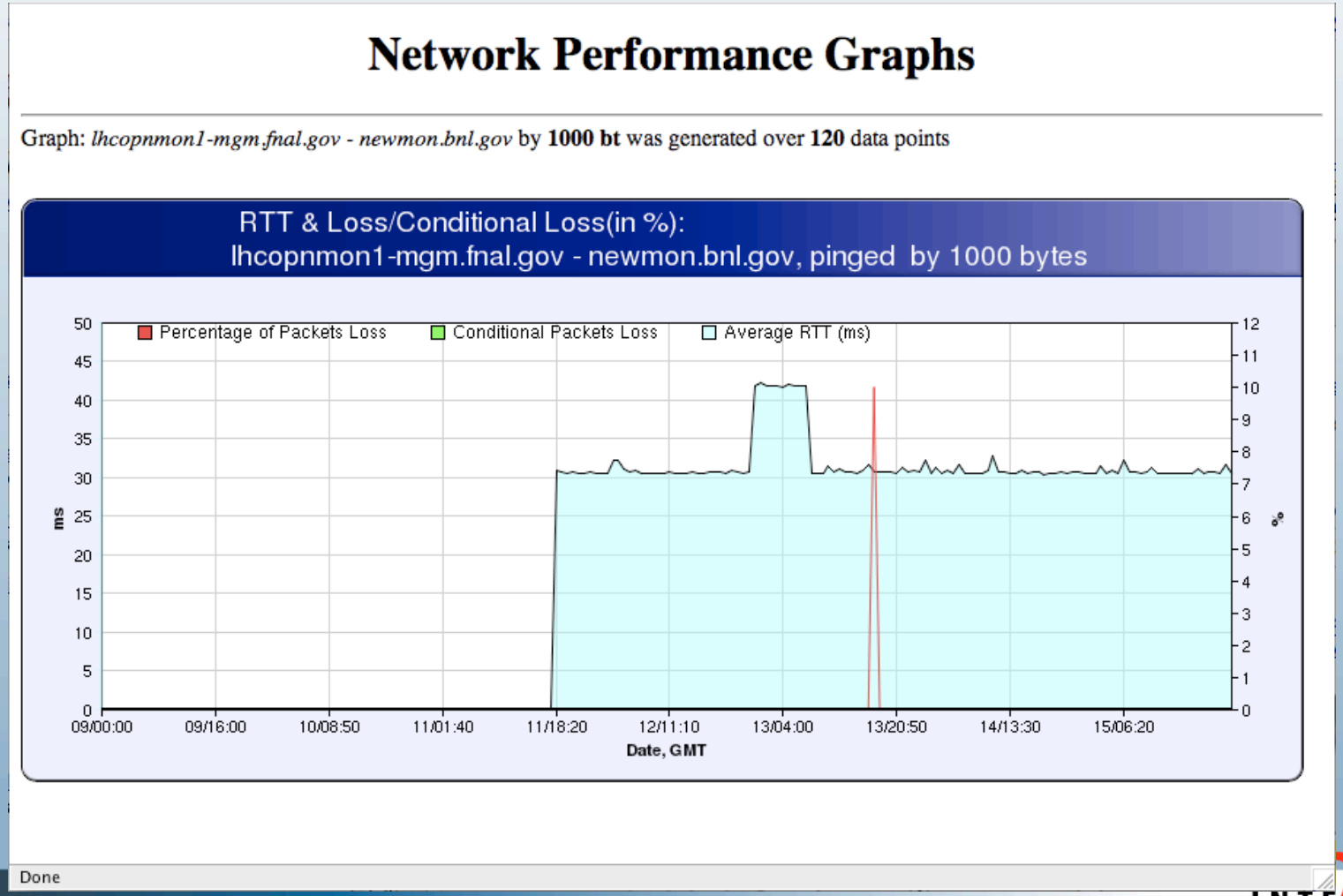
pingER Analysis

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



pingER Analysis

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



perfSONAR-BUOY (bwctl)

- http://ndb1.internet2.edu/cgi-bin/bwctl.cgi?name=OFFICEMESHBWTCP4

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

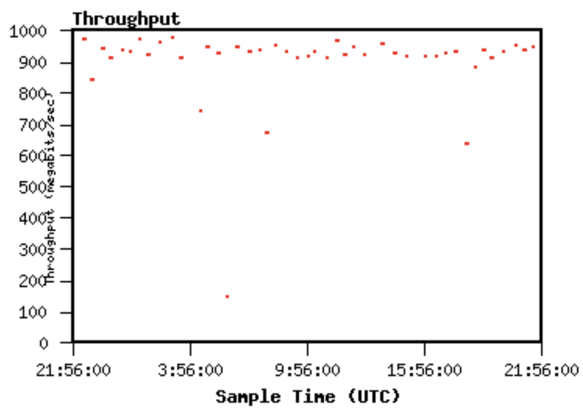
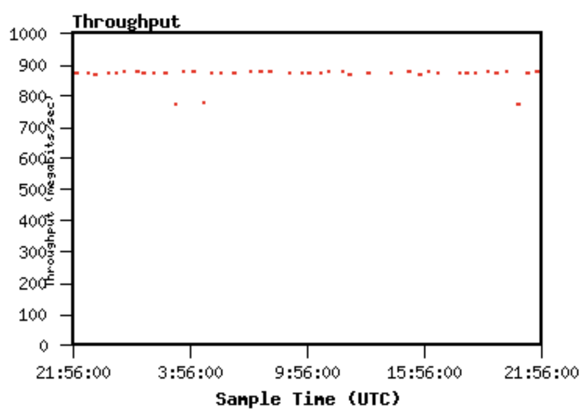
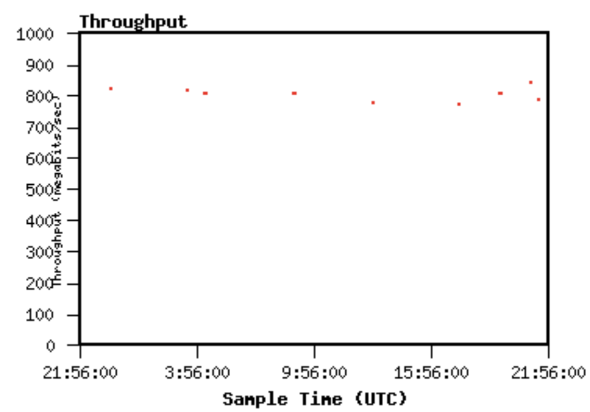
LHC Sites - IPv4 TCP Throughput

[Throughput(Mbps) / Date&Time]

BWTCP4	Senders					
Receivers		ann arbor	brookhaven national laboratory	chicago	fermi national accelerator laboratory	esnet
	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)

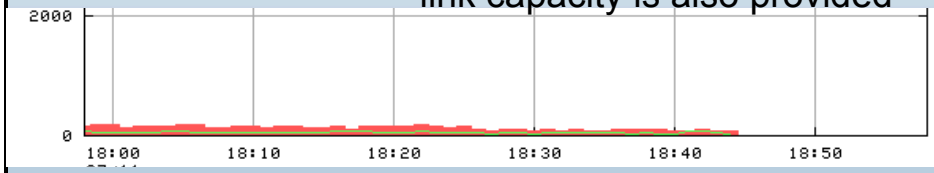
<table>
<tr><td>STARLIGHT <=> BNL</td><td>Sender: brookhaven national laboratory</td><td>Receiver: brookhaven national laboratory</td></tr>
<tr><td>ESNet</td><td> 
</td><td>No data available.</td></tr>
<tr><td>CHIC <=> BNL</td><td>Sender: brookhaven national laboratory</td><td>Receiver: brookhaven national laboratory</td></tr>
<tr><td>Chicago</td><td> 
</td><td> 
</td></tr>
</table>

Traceroute Visualizer

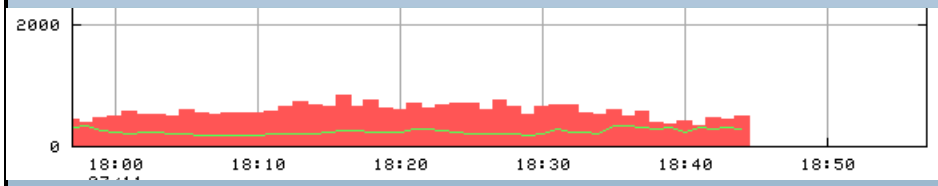
- Forward direction bandwidth utilization on application path from LBNL to INFN-Frascati (Italy)
  - traffic shown as bars on those network device interfaces that have an associated MP services (the first 4 graphs are normalized to 2000 Mb/s, the last to 500 Mb/s)

1 ir1000gw (131.243.2.1)
2 er1kgw
3 lbl2-ge-lbnl.es.net

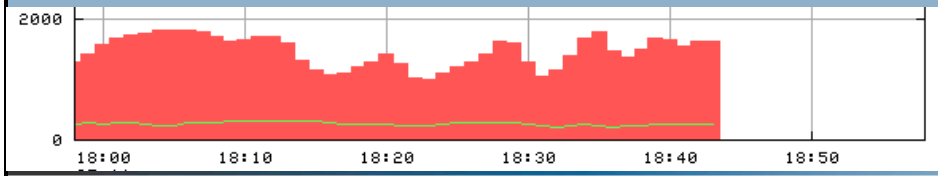
link capacity is also provided



4 slacmr1-sdn-lblmr1.es.net (GRAPH OMITTED)
5 snv2mr1-slacmr1.es.net (GRAPH OMITTED)
6 snv2sdn1-snv2mr1.es.net

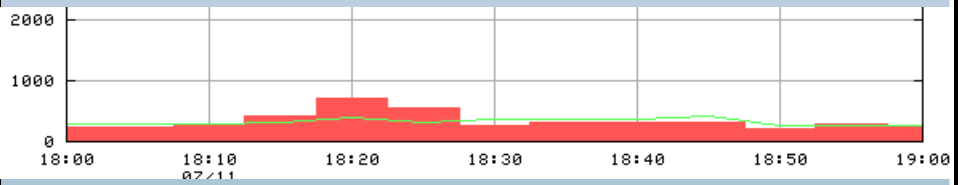


7 chislsdn1-oc192-snv2sdn1.es.net (GRAPH OMITTED)
8 chiccr1-chislsdn1.es.net

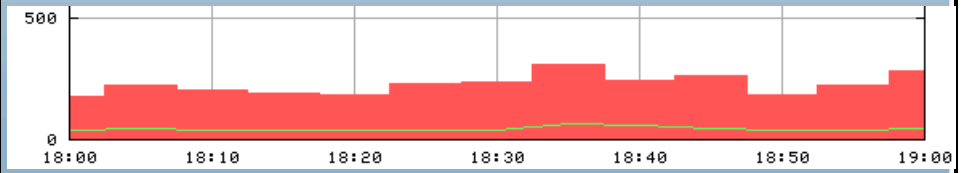


9 aofacr1-chicsdn1.es.net (GRAPH OMITTED)

10 esnet.rt1.nyc.us.geant2.net (NO DATA)
11 so-7-0-0.rt1.ams.nl.geant2.net (NO DATA)
12 so-6-2-0.rt1.fra.de.geant2.net (NO DATA)
13 so-6-2-0.rt1.gen.ch.geant2.net (NO DATA)
14 so-2-0-0.rt1.mil.it.geant2.net (NO DATA)
15 garr-gw.rt1.mil.it.geant2.net (NO DATA)
16 rt1-mi1-rt-mi2.mi2.garr.net



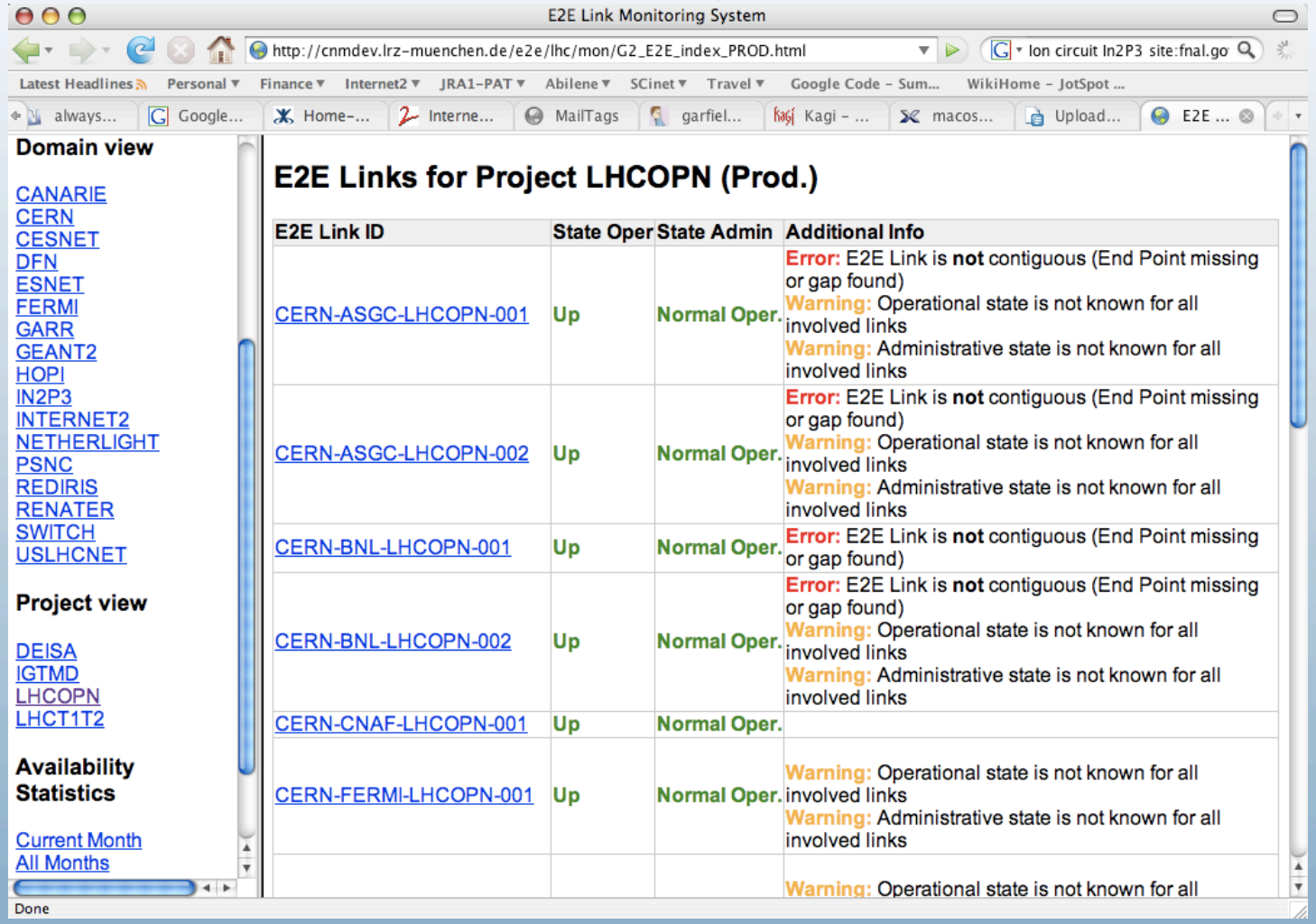
17 rt-mi2-rt-rm2.rm2.garr.net (GRAPH OMITTED)
18 rt-rm2-rc-fra.fra.garr.net (GRAPH OMITTED)
19 rc-fra-ru-lnf.fra.garr.net (GRAPH OMITTED)



20
21 www6.lnf.infn.it (193.206.84.223) 189.908 ms 189.596 ms 189.684 ms

INTERNET2®

E2Emon - Monitoring Circuits



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 generat

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	-

E2Emon generated view of the data for one OPN link [E2EMON]

INTERNET®

perfSONAR GUI List

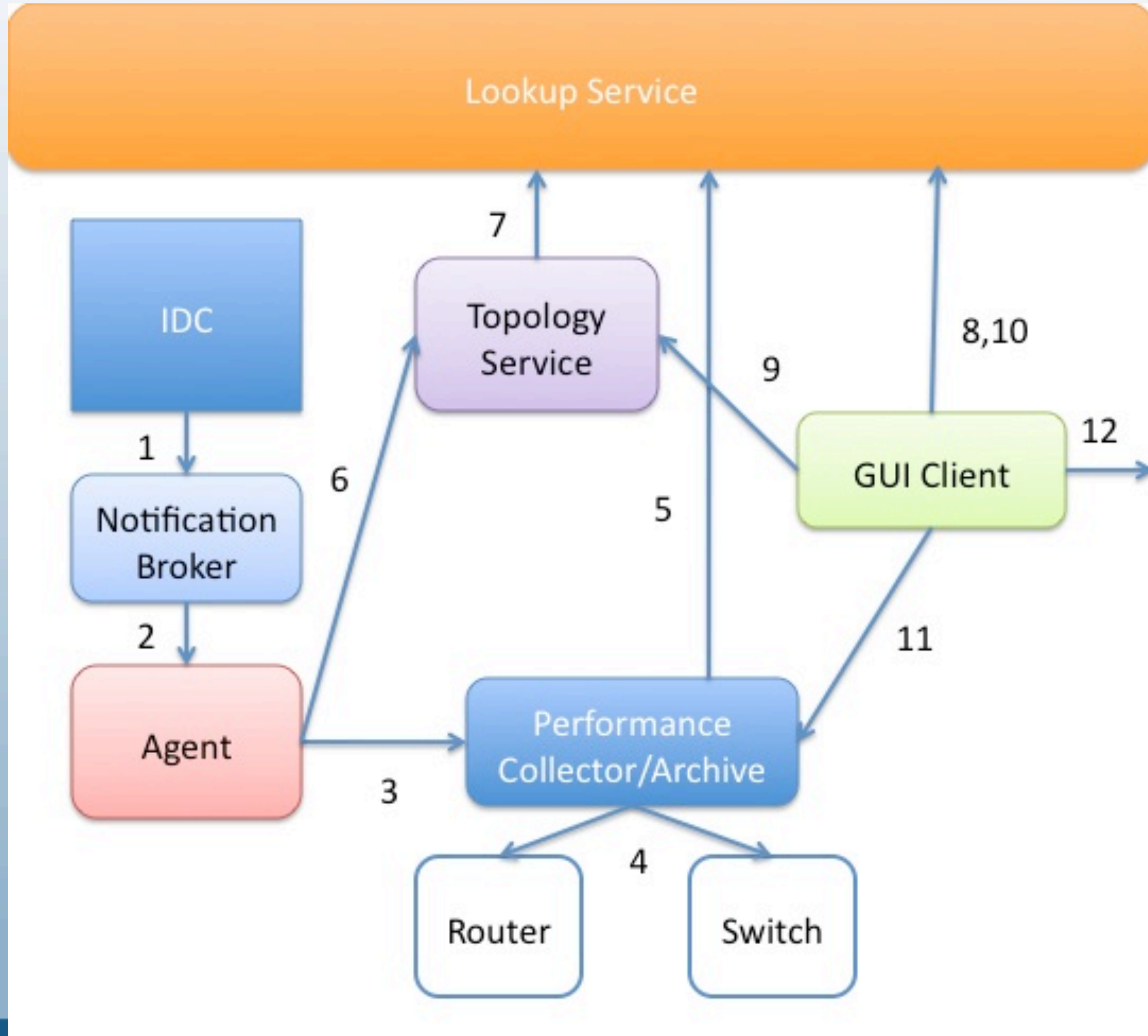
- http://www.perfsonar.net/activeServices
  - Listing of many types of active services
  - Graphs
  - Live Tests Points
- http://psvis0.internet2.edu:8008/
  - Gmaps interface
- http://perfsonar.acad.bg/
  - Java based visualization tool
- http://cnmdev.lrz-muenchen.de/e2e/lhc/mon/G2_E2E_index_ALL.html
  - 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

INTERNET®

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

INTERNET®

www.internet2.edu