ESnet Network Measurement
Network Research Needs

Joe Metzger

Energy Sciences Network
Lawrence Berkeley National Laboratory

Feb 7th 2008

Networking for the Future of Science
Outline

• Measurement Challenges for ESnet
• Measurement Challenges for ESnet Users
• Research and Development Directions
• Moving from Research to Production
• Collaborations
Measurement Challenges for ESnet

• Network Complexity is Increasing Significantly
  – More services
  – More devices and circuits
  – More complex routing plan
  – More dynamic
    • Physical circuits: average 1 new physical circuit every ~2 weeks for next couple years
    • Virtual cross domain circuits
  – How do we ensure that all of the components are working properly?

• How do we identify & resolve problems with multi-domain data flows that are running to slow (or to fast)
  – Ratio of well provisioned end station capacity to backbone link size is approaching 1.
    • Larger numbers of users may be able to saturate backbone links at will (or even accidentally without knowing it) in the near future.
    • May lead to transient queue saturation on the ms timescales.

• How should work-flow systems interact with the network?
  – Scheduling current & future resources
  – Determine if work-flow systems are receiving what they asked for?
  – Determine resources consumed
  – Exchange info about & diagnose problems
Measurement Challenges for ESnet Users

• Establishing Expectations
  – What are the characteristics of the network between A and B?
    • Now, last week & next week
  – How much of the resources in the path are available to a particular user?

• Verifying circuit functionality
  – Are dynamic circuits available
    • What path did a circuit actually take
    • Did it meet the requirements
  – What happens when a part of a circuit fails
    • Dynamic re-routing or resizing
    • Notifications?

• Circuit effectiveness tools
  – Provide quantitative measures about how effectively they are utilizing dynamic circuits
    • IE, users will want to know if they are burning up their virtual circuit allocations when the storage system is down.

• Debugging Problems
  – Need end-user tools that assist in differentiating between Network and Application problems
  – Tracking problem sources down to a domain
Research & Development Directions

• perfSONAR
  – AAA
    • Basic framework
      – Use a SAML shim in a standard web services architecture
    • How can we get it to work across multiple authentication infrastructures
    • How can we get it integrated into the different tools.
  – Dynamic Circuits
    • Tighter integration of OSCARS & perfSONAR PS
      – Measurement Collection & Archiving
        » Circuit Up & Down Status
        » Circuit utilization
      – Using the Measurements
        » Automated Analysis tools
        » Visualization Tools
  – User Tools
    • What tools are needed for NOC’s PERT’s and end users?
  – Automated Alerting and Alarming Tools looking at:
    • Latency
    • Achievable Bandwidth
    • Utilization
    • Circuit Status
    • Hotspots (past & future)
  – Middleware API
    • Allow workflow systems to manage their network interactions
Moving from Research to Production

- The current perfSONAR deployment strategies
  - Commit that ESnet, Internet2 and GEANT will deploy perfSONAR services and protocols. This is a defacto majority of the R&E community in Europe & US.
  - Engage the target user community and have them participate in the development so they have a vested interest in deployment in their community.
  - Focus on global projects that have critical network dependencies and a hierarchal structure. The core perfSONAR collaboration assists the top level with deployment. Then allow the projects top level support team to encourage & support the lower levels when they deploy the tools.

- Issues
  - GEANT support AUPs
  - Internet2, SLAC, & FERMI funding issues
  - Current focus on developing the portions of the framework necessary for the backbone networks to collect, manage and export backbone measurement data
    - Development effort on user interface and analysis tools is lagging behind.
    - “Need a NOC quality GUI” for general adoption
External Collaborations

• ESnet Sites
  – FERMI
  – SLAC

• US R&E Community
  – Internet2
  – U Delaware
  – Georgia Tech

• International
  – GEANT & European NRENS
  – RNP (Brazil)
  – LHC Community
Additional Backup Slides
What is perfSONAR

• 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.

• An architecture framework & a set of protocols
  – Web Services Architecture
  – Protocols based on the Open Grid Forum Network Measurement Working Group Schemas

• Several interoperable software implementations
  – Java & Perl

• A growing deployed measurement infrastructure
perfSONAR Collaborators

- RNP
- ARNES
- BELNET
- CARNET
- CESNET
- CYNET
- DANTE
- DFN
- ESnet
- FCCN
- FERMI
- GARR
- GEANT
- GRNET
- HEAnet
- Internet2
- ISTF
- POZnan
- UNINETT
- University of Delaware
- Renater
- RedIRIS
- SLAC
- SWITCH
- SURFnet

And anybody else I missed
## Widely Deployed perfSONAR Services

<table>
<thead>
<tr>
<th>Utilization MA:</th>
<th>Layer 2 Circuit Status for LHCOPN Circuits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet2</td>
<td>ASCC</td>
</tr>
<tr>
<td>ISTF/ACAD</td>
<td>CERN</td>
</tr>
<tr>
<td>CARNet</td>
<td>ESnet</td>
</tr>
<tr>
<td>Cesnet</td>
<td>GARR</td>
</tr>
<tr>
<td>CYnet</td>
<td>GEANT2</td>
</tr>
<tr>
<td>ESnet</td>
<td>NORDUnet</td>
</tr>
<tr>
<td>FERMI</td>
<td>Renater</td>
</tr>
<tr>
<td>GARR</td>
<td>REDIris</td>
</tr>
<tr>
<td>GEANT2</td>
<td>Sara/SURFnet</td>
</tr>
<tr>
<td>GRNET</td>
<td>SWITCH</td>
</tr>
<tr>
<td>MREN</td>
<td>CANARIE</td>
</tr>
<tr>
<td>PIONIER</td>
<td>USLHCPnet</td>
</tr>
<tr>
<td>RedCLARA</td>
<td>UKERNA</td>
</tr>
<tr>
<td>RENATER</td>
<td></td>
</tr>
<tr>
<td>RNP</td>
<td></td>
</tr>
<tr>
<td>SEEREN</td>
<td></td>
</tr>
<tr>
<td>SWITCH</td>
<td></td>
</tr>
<tr>
<td>Uninett</td>
<td></td>
</tr>
</tbody>
</table>
perfSONAR Architecture

- Interoperable network measurement middleware:
  - Modular
  - Web services-based
  - Decentralized
  - Locally controlled

- Integrates:
  - Network measurement tools
  - Network measurement archives
  - Discovery
  - Authentication and authorization
  - Data manipulation
  - Resource protection
  - Topology

- Based on:

- Each Domain
  - Collects measurement data internal to their domain
  - May collect cross-domain measurement data
  - **Controls publication of the information they collect and manage**
perfSONAR Services (1)

- **Measurement Point Service**
  - Enables the initiation of performance tests
  - Currently in use: bandwidth, latency, looking glass, circuit status

- **Measurement Archive Service**
  - Stores and publishes performance monitoring results
  - Currently in use: interface utilization and errors, bandwidth, latency, circuit status

- **Lookup Service**
  - Allows the client to discover the existing services and other LS services.
  - Dynamic: services register themselves to the LS and describe their capabilities, they can also leave or be removed if a service goes down.
  - Currently available: service discovery in a single domain. Global integration getting close.

- **Authentication Service**
  - Based on Existing efforts: Internet2 MAT, GN2-JRA5
  - Authentication & Authorization functionality for the framework
  - Trust relationship between networks
  - Currently status: Some services work with eduGAIN identities
perfSONAR Services (2)

• Transformation Service
  – Transform the data (aggregation, concatenation, correlation, translation, etc).
  – Current status: perfSONAR_PS includes a transform service that determines circuit status based on info from several other services

• Topology Service
  – Publish the network topology information via the framework.
  – Currently status: perfSONAR_PS includes a useful server. CNIS exports a perfSONAR topology interface

• Resource protector
  – Arbitrate the consumption of limited resources.
  – Current status: not in perfSONAR, but some capabilities in underlying tools.
Gmaps from SLAC
CNM from DFN
CNM from DFN
perfSONARUI from acad.bg
Oscars Circuit plugin - Internet2
Oscars circuit plugin
E2Emon - Monitoring Circuits

![Image of E2Emon monitoring interface](image)

### E2E Links for Project LHCOPN (Prod.)

<table>
<thead>
<tr>
<th>E2E Link ID</th>
<th>State Oper</th>
<th>State Admin</th>
<th>Additional Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>CERN-ASGC-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td><strong>Error:</strong> E2E Link is not contiguous (End Point missing or gap found)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Warning:</strong> Operational state is not known for all involved links</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Warning:</strong> Administrative state is not known for all involved links</td>
</tr>
<tr>
<td>CERN-ASGC-LHCOPN-002</td>
<td>Up</td>
<td>Normal Oper</td>
<td><strong>Error:</strong> E2E Link is not contiguous (End Point missing or gap found)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Warning:</strong> Operational state is not known for all involved links</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Warning:</strong> Administrative state is not known for all involved links</td>
</tr>
<tr>
<td>CERN-BNL-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERN-BNL-LHCOPN-002</td>
<td>Up</td>
<td>Normal Oper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERN-CNAF-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERN-FERMI-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERN-FERMI-LHCOPN-002</td>
<td>Up</td>
<td>Normal Oper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERN-GRIDKA-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERN-IN2P3-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERN-NDGF-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td><strong>Warning:</strong> Operational state is not known for all involved links</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Warning:</strong> Administrative state is not known for all involved links</td>
</tr>
<tr>
<td>CERN-PIC-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td><strong>Error:</strong> E2E Link is not contiguous (End Point missing or gap found)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Warning:</strong> Operational state is not known for all involved links</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Warning:</strong> Administrative state is not known for all involved links</td>
</tr>
<tr>
<td>CERN-RAL-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERN-SARA-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERN-TRIUMF-LHCOPN-001</td>
<td>Up</td>
<td>Normal Oper</td>
<td><strong>Error:</strong> E2E Link is not contiguous (End Point missing or gap found)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Warning:</strong> Operational state is not known for all involved links</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Warning:</strong> Administrative state is not known for all involved links</td>
</tr>
</tbody>
</table>
E2Emon: Status of E2E link CERN-LHCOPN-FNAL-001

<table>
<thead>
<tr>
<th>Domain</th>
<th>CERN</th>
<th>USLHCNET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link Structure</td>
<td>EP</td>
<td>DP</td>
</tr>
<tr>
<td>Type</td>
<td>EndPoint</td>
<td>ID Part.Info</td>
</tr>
<tr>
<td>Local Name</td>
<td>CERN-TO</td>
<td>SS13-C-BE1</td>
</tr>
<tr>
<td>State Oper.</td>
<td>-</td>
<td>Up</td>
</tr>
<tr>
<td>Timestamp</td>
<td>-</td>
<td>2007-04-08 T05:04:08+02:00</td>
</tr>
</tbody>
</table>

E2Emon generated view of the data for one OPN link [E2EMON]
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
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.mil2.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
```

link capacity is also provided
In Summary, perfSONAR is:

- A global collaboration
- A framework, and a set of protocols that implements parts of the framework in a web services environment
- Several interoperable software implementations
- A measurement infrastructure that is deployed in over a dozen different networks
Research Project Ideas

Application community specific cross domain network weathermaps

• Performance analysis dashboard

• Analyze the current network using current and historical data to forecast future hotspots.

• Automated dynamic circuit based flow accelerator
  – Identify, characterize and rank real-time flows that are candidates for dynamic circuits.
  – Dynamically re-route them using OSCARS

• Identify & build the tools necessary for workflow systems to manage network resources.