LAMP: Leveraging and Abstracting Measurements with perfSONAR

Guilherme Fernandes

fernande@cis.udel.edu

perfSONAR – Short Intro

- perfSONAR is a multi-domain performance monitoring framework, which defines a set of protocol standards for ***sharing data*** between measurement and monitoring systems
  - End-to-end performance problems on paths crossing several networks.
- Design Goals:
  - Standards-based, Decentralized, Locally Controlled
  - Open Source, Modular, Extensible
    - Applicable to multiple generations of network monitoring systems
    - Grows “beyond our control”
    - Customized for individual science disciplines

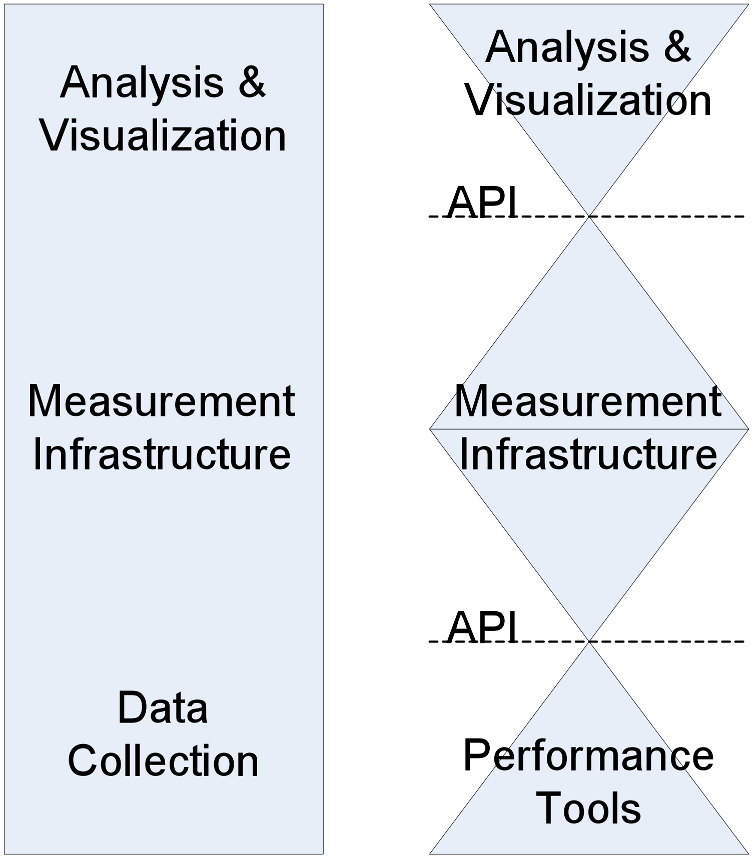
perfSONAR - Architecture

- Interoperable network measurement middleware designed as a Service Oriented Architecture (SOA):
  - Components are Web Services (WS) based

- The framework is made up of several unique components and design considerations, all of which operate in a cooperative yet independent manner
  - Each functionality is separated into a specific function
  - Clients and servers interact through scripted, XML Based protocols
  - Measurement data is encoded in expressive XML formats

- perf*SONAR* Integrates:
  - Network measurement tools and archives (e.g. stored measurement results)
  - Data manipulation
  - Information Services
    - Discovery
    - Topology
  - Authentication and authorization

perfSONAR – as a middleware



Analysis & Visualization

Measurement Infrastructure

Data Collection

Analysis & Visualization

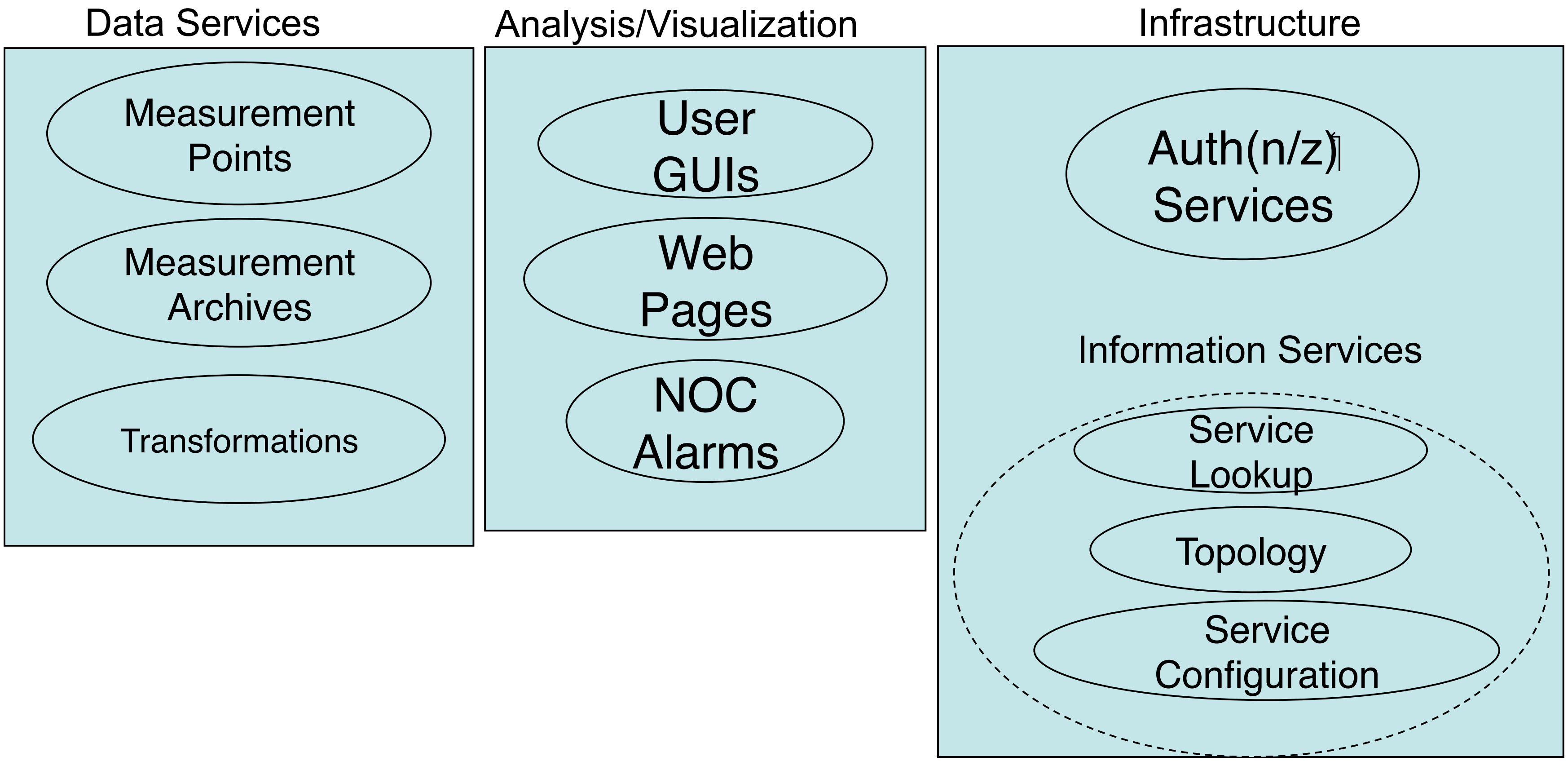
API

Measurement Infrastructure

API

Performance Tools

perfSONAR - Components



perfSONAR - Services

- Measurement Point (MP) Service
  - Enables the initiation of performance tests
- Measurement Archive (MA) Service
  - Stores and publishes performance monitoring results
- Transformation Service
  - Transform the data (aggregation, concatenation, correlation, translation, etc)
- Resource protector
  - Arbitrate the consumption of limited resources
  - Other services delegate a limited portion of the authorization decision here

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

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 MAT, 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 concerned with discovering and federating the available network services

perfSONAR - Open Protocols & Schemata

- Base network measurement schema
  - OGF Network Measurement Working Group
- Topology Schema
  - OGF Network Markup Language (NML-)WG
  - Includes Topology Network ID
- perfSONAR Protocol Documents
  - OGF Network Measurement and Control (NMC-)WG

Measurement Schema

- Key Goals: Extensibility, Normalization, Readability
  - Extensibility achieved through XML namespaces – can represent basically any measurement

- Break representation of performance measurements down into basic elements
  - Measurement Data
    - A set of measurement events that have some value or values at a particular time
  - Measurement Metadata
    - The details about the set of measurement data

Schema Basic Elements - Metadata

- Subject (Noun)
  - The measured/tested entity (who)
  - E.g. A pair of hosts (end-point-pair), or a Layer 3 interface
- EventType (Verb)
  - What type of measurement, value, or event occurred
  - Characteristic, tool output, or generic event
  - E.g. *latency*, *bandwidth*, *utilization*, or simply *iperf*
- Parameters (Adjectives and Adverbs)
  - How, or under what conditions, did this event occur?
  - E.g. buffer sizes used, *TCP* vs *ICMP* packets
- Key
  - Shortcut substituted in place of previous three items
  - No predefined format

Schema Basic Elements - Data

- Datum: The actual result (values) of measurement.
  - Can contain time (e.g. a Time element or attribute).
  - Existence of an event might point to the case where there no additional value
    - As in “Link up/down” or threshold events

- Time: Representation of a time stamp or time range in a specified format.
  - Must be extensible since even agreement about the right structure is not easy
    - E.g. UNIX timestamp vs NTP time

All measurements have some sort of Data and Time

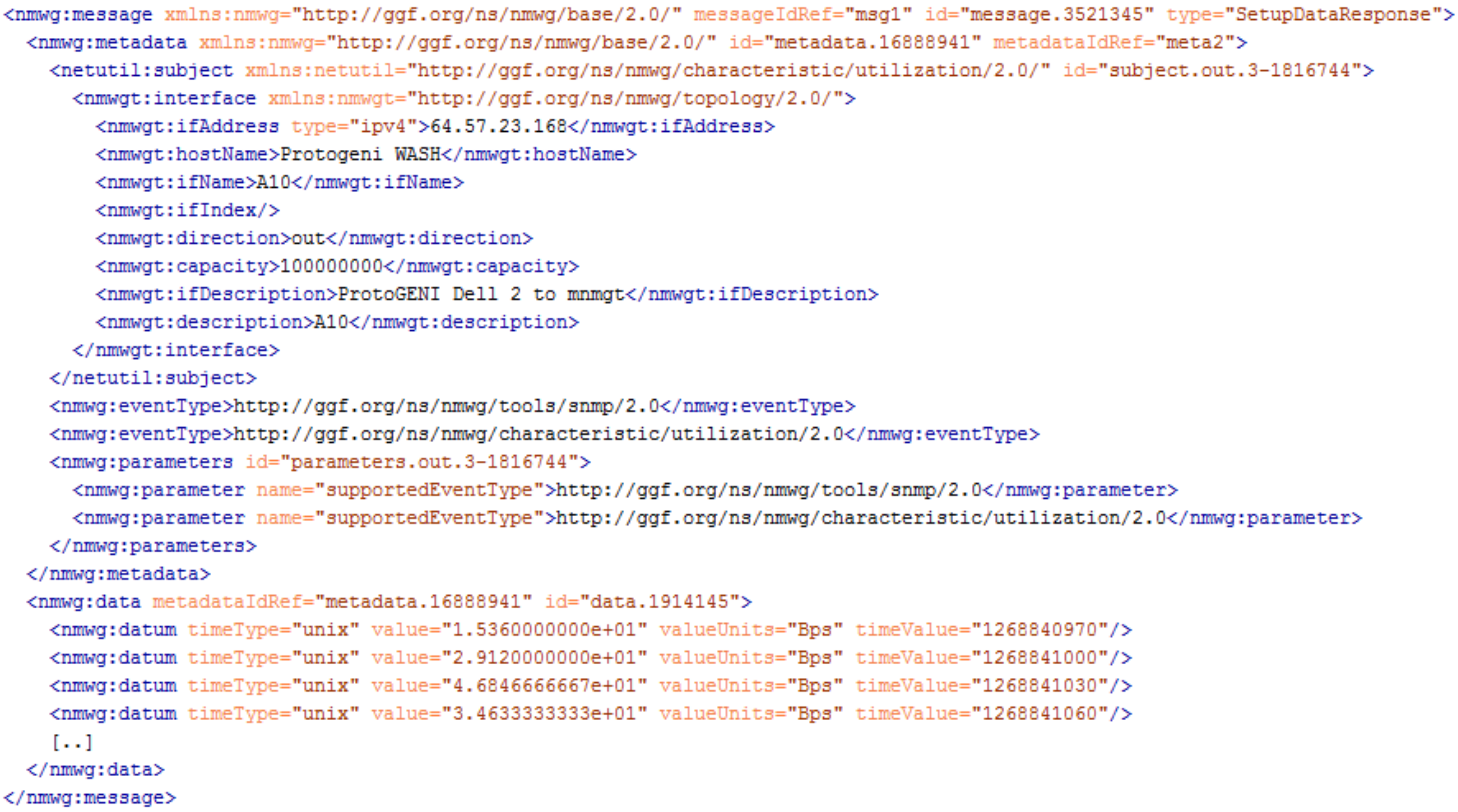
Schema Namespaces & Extensibility

- A namespace:
  - http://ggf.org/ns/nmwg/base/2.0/
  - ***MAY NOT*** be a URL
- All measurements can be described by the Metadata identifying who/what/how
- Each measurement might have specific needs regarding data/metadata elements
  - Approach: Use Data and Metadata elements and vary the namespaces of the specific elements
- We encode the measurement/event type in the namespace (and as a standalone element)

Schema Namespaces & Extensibility

- Extensibility achieved through hierarchy with delegation
  - Similar to OIDs in the IETF management world
- The NM-WG has a hierarchy of network characteristics
  - Good starting point
  - E.g. http://ggf.org/ns/nmwg/characteristic/utilization/2.0, http://ggf.org/ns/nmwg/characteristic/bandwidth/achievable/2.0
- However, not all tools are cleanly mapped onto the Characteristic space
  - Often a matter of some debate
- Organization-rooted tools namespace addresses this
  - Easy to add new tools in organization-specific namespaces
  - E.g. http://ggf.org/ns/nmwg/tools/nuttcp/2.0

ProtoGENI SNMP MA Example



```
<nmwg:message xmlns:nmwg="http://ggf.org/ns/nmwg/base/2.0/" messageIdRef="msg1" id="message.3521345" type="SetupDataResponse">
  <nmwg:metadata xmlns:nmwg="http://ggf.org/ns/nmwg/base/2.0/" id="metadata.16888941" metadataIdRef="meta2">
    <netutil:subject xmlns:netutil="http://ggf.org/ns/nmwg/characteristic/utilization/2.0/" id="subject.out.3-1816744">
      <nmwgt:interface xmlns:nmwgt="http://ggf.org/ns/nmwg/topology/2.0/">
        <nmwgt:ifAddress type="ipv4">64.57.23.168</nmwgt:ifAddress>
        <nmwgt:hostName>Protogeni WASH</nmwgt:hostName>
        <nmwgt:ifName>A10</nmwgt:ifName>
        <nmwgt:ifIndex/>
        <nmwgt:direction>out</nmwgt:direction>
        <nmwgt:capacity>100000000</nmwgt:capacity>
        <nmwgt:ifDescription>ProtoGENI Dell 2 to mnmgt</nmwgt:ifDescription>
        <nmwgt:description>A10</nmwgt:description>
      </nmwgt:interface>
    </netutil:subject>
    <nmwg:eventType>http://ggf.org/ns/nmwg/tools/snmp/2.0</nmwg:eventType>
    <nmwg:eventType>http://ggf.org/ns/nmwg/characteristic/utilization/2.0</nmwg:eventType>
    <nmwg:parameters id="parameters.out.3-1816744">
      <nmwg:parameter name="supportedEventType">http://ggf.org/ns/nmwg/tools/snmp/2.0</nmwg:parameter>
      <nmwg:parameter name="supportedEventType">http://ggf.org/ns/nmwg/characteristic/utilization/2.0</nmwg:parameter>
    </nmwg:parameters>
  </nmwg:metadata>
  <nmwg:data metadataIdRef="metadata.16888941" id="data.1914145">
    <nmwg:datum timeType="unix" value="1.5360000000e+01" valueUnits="Bps" timeValue="1268840970"/>
    <nmwg:datum timeType="unix" value="2.9120000000e+01" valueUnits="Bps" timeValue="1268841000"/>
    <nmwg:datum timeType="unix" value="4.6846666667e+01" valueUnits="Bps" timeValue="1268841030"/>
    <nmwg:datum timeType="unix" value="3.4633333333e+01" valueUnits="Bps" timeValue="1268841060"/>
    [..]
  </nmwg:data>
</nmwg:message>
```

Topology Schema

- Topology schema grew from network measurement description
  - Reusable "Subject" elements for common cases
    - Also reduces redundancy
  - Relationships between measurement Subjects
- Structured by layers and the same elements recurring there (Base, L2, L3, L4)
  - networks as *graphs*
- Elements:
  - Domain
  - Node
  - Port
  - Link
  - Network
  - Path
  - Service

Topology Schema

- Varied by namespaces (extensibility)
  - Reuse visualization logic, etc.
  - Validate layer- or technology-specific attributes
- Used by perfSONAR, IDC Protocol (ION, OSCARS, AutoBAHN), Phoebus
  - Currently calling it the UNIS Topology Schema
- OGF NML-WG to unify NDL and UNIS Topology schema
  - Happening as we speak at OGF28

LAMP Objectives

- Collaborate on defining a common but extensible format for data storage and exchange for GENI I&M systems
  - Use perfSONAR NM-WG schema as starting point
  - Identify new characteristics/tools namespaces
- Develop a representation of GENI topology to be used to describe measurements and experiment configuration
  - UNIS topology schema can be easily extended
- Collaborate with related GENI measurement and security projects on a common GENI I&M architecture
  - The new GENI I&M Arch. Draft defines very similar services (MP, MC, MDA, MAP), and new ones (MO)
  - perfSONAR is a good starting point, not currently a final solution (for GENI); Use cases have been different, but much can be reused and the framework can be extended

Questions?

THANK YOU!

Credits: Jeff Boote, Jason Zurawski, and many more from the perfSONAR community.

Migrating ProtoGENI RSpec Topology

- Benefits:
  - Standardized representation used by other big projects (perfSONAR, Internet2 ION, Phoebus, etc)
  - Different GENI facilities require topology information, but at different layers of abstraction
    - O&M, I&M, all should use same schema for integration
    - RSpec would need to grow and increase in complexity
    - NML-WG focuses on arbitrarily complex, multi-layer network representation
  - Interaction with external network control and monitoring (perfSONAR/IDC) will be possible
    - E.g. perfSONAR data for GENI resources can be correlated with perfSONAR data taken for external items. Similarly it would be possible to control Dynamic Circuit networks and GENI resources using the same Control Plane calls.

Migrating ProtoGENI RSpec Topology

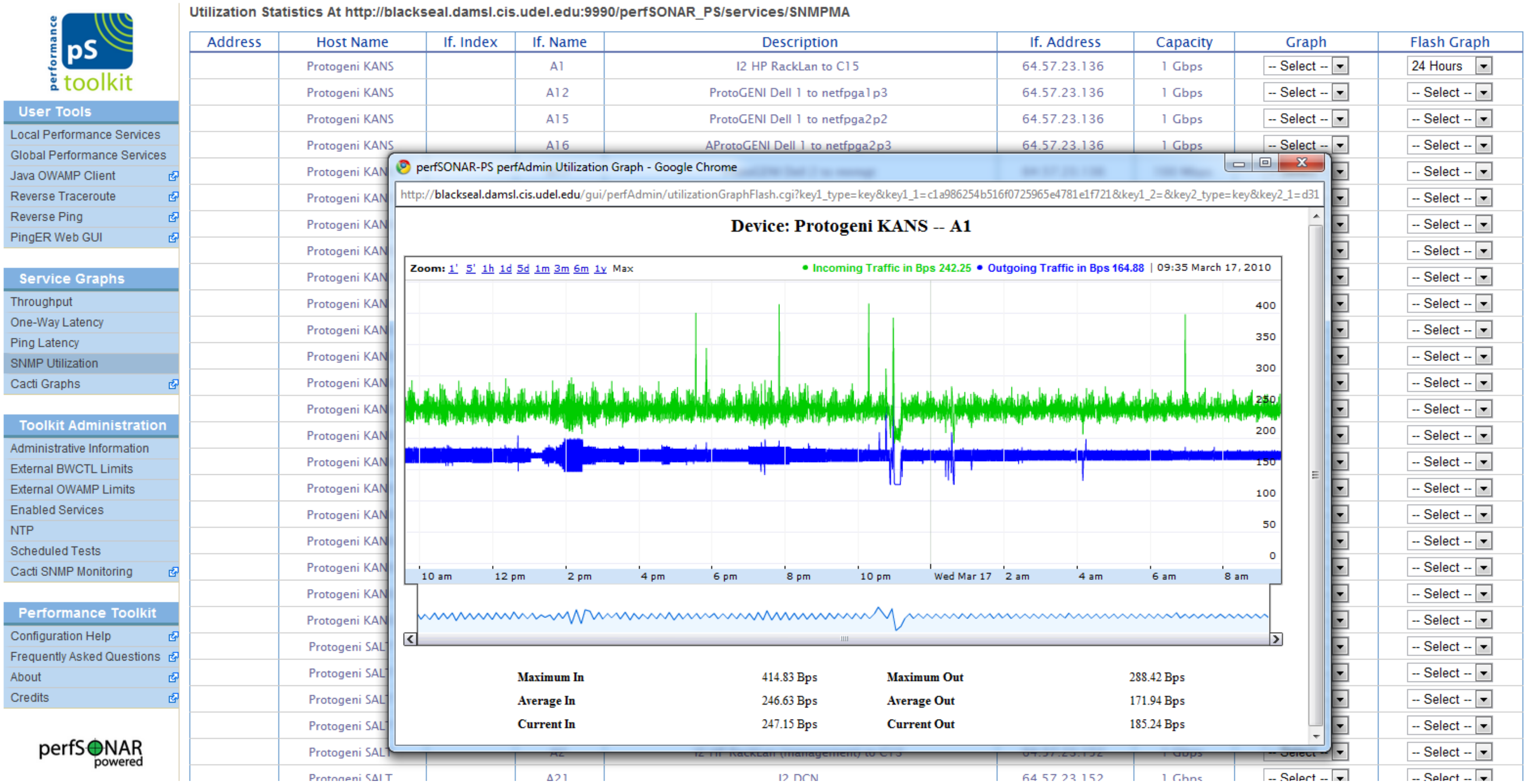
- Cons:
  - UNIS is not yet a standard, (minimal) changes are always possible as a standard is approached
  - Current RSpec is not completely compatible with NML descriptions, key constructs may be missing or form an incompatible mapping.
    - New version of Rspec won’t be backwards compatible
  - Mechanical translation will become a burden in a realtime system, particularly when trying to invoke on-demand measurement and recording through information services.

perfSONAR-based GENI I&M

- Operations & Management
  - Federated GENI is perfect match to perfSONAR
  - perfSONAR SNMP MA setup to query ProtoGENI switches on Internet2

- Experimenters
  - perfSONAR architecture is modular and extensible
  - Integration with GENI will need
    - Measurement Portal: allows users to augment their slice with measurement infrastructure (mix and matched to their needs); visualization of measurement data
    - Measurement Orchestration?: Bootstrapping the necessary resources, (re)configure underlying tools/sensors

perfSONAR on GENI Backbone



User Tools

- Local Performance Services
- Global Performance Services
- Java OWAMP Client
- Reverse Traceroute
- Reverse Ping
- PingER Web GUI

Service Graphs

- Throughput
- One-Way Latency
- Ping Latency
- SNMP Utilization
- Cacti Graphs

Toolkit Administration

- Administrative Information
- External BWCTL Limits
- External OWAMP Limits
- Enabled Services
- NTP
- Scheduled Tests
- Cacti SNMP Monitoring

Performance Toolkit

- Configuration Help
- Frequently Asked Questions
- About
- Credits

Utilization Statistics At http://blackseal.damsl.cis.udel.edu:9990/perfSONAR_PS/services/SNMPMA

Address	Host Name	If. Index	If. Name	Description	If. Address	Capacity	Graph	Flash Graph
	Protogeni KANS		A1	I2 HP RackLan to C15	64.57.23.136	1 Gbps	-- Select --	24 Hours
	Protogeni KANS		A12	ProtoGENI Dell 1 to netfpga1p3	64.57.23.136	1 Gbps	-- Select --	-- Select --
	Protogeni KANS		A15	ProtoGENI Dell 1 to netfpga2p2	64.57.23.136	1 Gbps	-- Select --	-- Select --
	Protogeni KANS		A16	AProtoGENI Dell 1 to netfpga2p3	64.57.23.136	1 Gbps	-- Select --	-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni KAN							-- Select --
	Protogeni SAL							-- Select --
	Protogeni SAL							-- Select --
	Protogeni SAL							-- Select --
	Protogeni SAL							-- Select --
	Protogeni SALT		A2	I2 HP RackLan (management) to C15	64.57.23.152	1 Gbps	-- Select --	-- Select --
	Protogeni SALT		A21	I2 DCN	64.57.23.152	1 Gbps	-- Select --	-- Select --

perfSONAR-PS perfAdmin Utilization Graph - Google Chrome

http://blackseal.damsl.cis.udel.edu/gui/perfAdmin/utilizationGraphFlash.cgi?key1_type=key&key1_1=c1a986254b516f0725965e4781e1f721&key1_2=&key2_type=key&key2_1=d31

Device: Protogeni KANS -- A1

Zoom: 1' 5' 1h 1d 5d 1m 3m 6m 1y Max
● Incoming Traffic in Bps 242.25 ● Outgoing Traffic in Bps 164.88 | 09:35 March 17, 2010
400
350
300
250
200
150
100
50
0
10 am
12 pm
2 pm
4 pm
6 pm
8 pm
10 pm
Wed Mar 17
2 am
4 am
6 am
8 am

Maximum In	414.83 Bps	**Maximum Out**	288.42 Bps
Average In	246.63 Bps	**Average Out**	171.94 Bps
Current In	247.15 Bps	**Current Out**	185.24 Bps