

# Meso-scale Monitoring: Introduction











Sarah Edwards, GPO  
Chaos Golubitsky, GPO

## Two goals for today's meeting:

- #1 Bottom-up
  - Continue to address day-to-day issues & pains
- #2 Top-down
  - Look ahead 6-36 months to where GENI is headed
  - Ensure:
    - We don't ignore any important pieces
    - Architecture decisions reflect needs of monitoring so that GENI Clearinghouse, I&M, etc serve our needs
    - Where possible, we build tools which can be adapted to new software when it becomes available

- Introduction [5 min]
  - Sarah Edwards, GPO
- Bottom-Up
  - FOAM: New OpenFlow Aggregate Manager [10 min]
    - Josh Smift, GPO
  - Using SNAPP to Find and Visualize GENI Monitoring Data [20 min]
    - Camilo Viecco, Indiana University & John Meylor, Indiana University
  - Topology [15 min]
    - Chaos Golubitsky, GPO
- Top-Down
  - Monitoring/Mgmt Requirements & Discussion [1 hour]
    - Sarah Edwards, GPO

# GEC12 Status & Change since GEC11

Status		Requirement/Pain Point
	(1)	Standard monitoring stats: cpu, mem usages; i/f stats
✓ 	(2)	Topology information. (a) End-to-end testing/Network reachability (b) Integrity of topology.
	(3)	OpenFlow Stats
<b>NEW</b>  @ GEC11	(4)	Make monitoring software easier to deploy (ie. plastic slices monitoring software)
✓ <b>NEW</b>  @ GEC11	(5)	Administering OpenFlow AM (OF Opt-In) and tie-together GENI things for OF (ie slivernames hard to determine from slice name)
	(6)	Aggregate/campus/slice view of resources and their availability
✓ 	(7)	Sharing data (a) SNMP data (b) Sharing of non-SNMP data
	(8)	Privacy
	(9)	Collecting and sharing of GENI usage/overall experiment stats
	(10)	Accountability report: How to prove if this is not my fault?
	(11)	Others?

- Introduction [5 min]
  - Sarah Edwards, GPO
- Bottom-Up
  - FOAM: New OpenFlow Aggregate Manager [10 min]
    - Josh Smift, GPO
  - Using SNAPP to Find and Visualize GENI Monitoring Data [20 min]
    - Camilo Viecco, Indiana University & John Meylor, Indiana University
  - Topology [15 min]
    - Chaos Golubitsky, GPO
- Top-Down
  - Monitoring/Mgmt Requirements & Discussion [1 hour]
    - Sarah Edwards, GPO

# GENI

## Meso-scale Monitoring - FOAM

**Josh Smift, GPO**  
**Kansas City, Missouri November 3, 2011**  
**[www.geni.net](http://www.geni.net)**

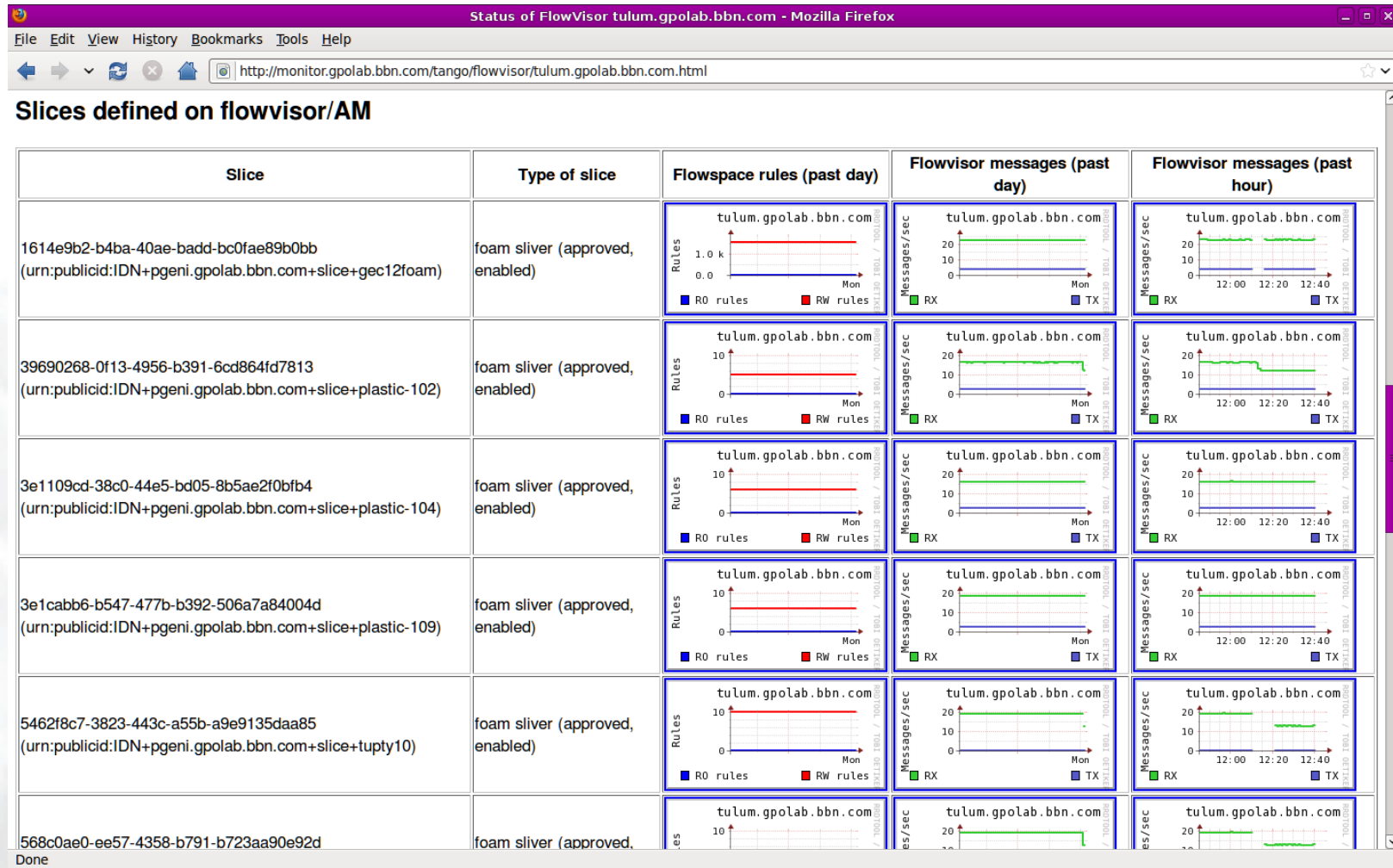
- **FlowVisor OpenFlow Aggregate Manager**
  - Replaces the AM and OIM parts of Expedient
  - Only an AM so far – no end-user opt-in (yet)
  - GENI AM API interface for experimenters
  - JSON API interface for administration
  - foamctl command-line interface to the JSON API
  - Lightweight nginx front-end provides HTTPS
  - Runs on Ubuntu 10.04 LTS now, Red Hat 6 soon
  - 0.4 running at BBN now (for the past few weeks)
  - 0.6 release Nov 21<sup>st</sup> for widespread deployment

<https://openflow.stanford.edu/display/DOCS/FOAM>

- Pain point: Hard to link FlowVisor slices, GENI slivers, GENI slices, and experimenter identities
- Foam helps: `show-sliver` shows lots of info

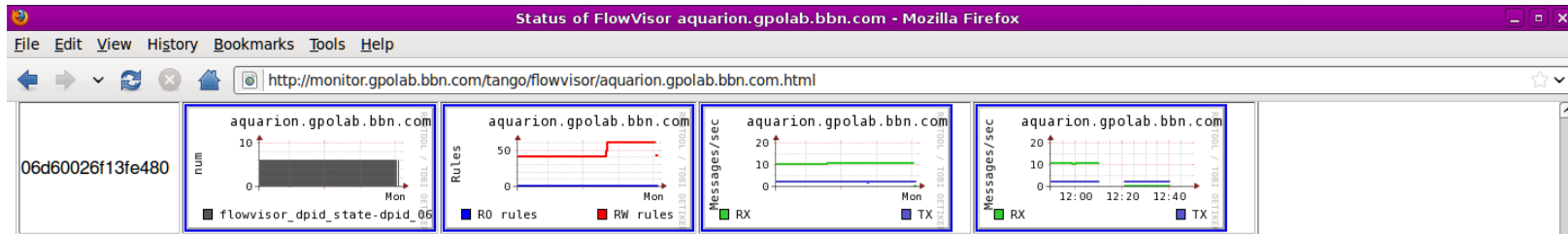
```
+$ foamctl show-sliver -u $sliver_urn --passwd-file=/opt/foam/etc/foampasswd
{
  "sliver": {
    "status": "Approved",
    "deleted": "False",
    "slice_urn": "urn:publicid:IDN+pgeni.gpolab.bbn.com+slice+plastic-104",
    "enabled": true,
    "email": "jbs@bbn.com",
    "controllers": [
      {
        "url": "tcp:naxos.gpolab.bbn.com:33104",
        "type": "primary"
      }
    ],
    "expiration": "2011-11-04 16:20:15.204288",
    "desc": "plastic-104 OpenFlow resources at BBN.",
    "sliver_urn": "urn:publicid:IDN+pgeni.gpolab.bbn.com+slice+plastic-104:3e1109cd-38c0-44e5-
      bd05-8b5ae2f0bfb4",
    "id": 31,
    "flowvisor_slice": "3e1109cd-38c0-44e5-bd05-8b5ae2f0bfb4"
  }
}
```



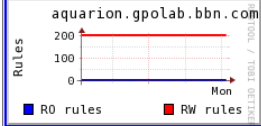
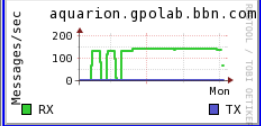
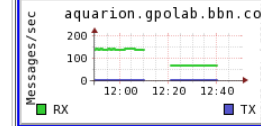
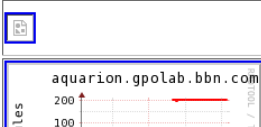
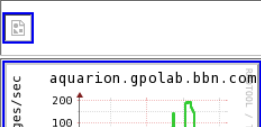
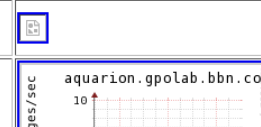
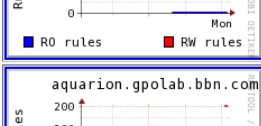
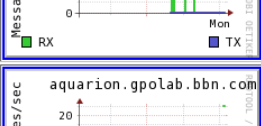
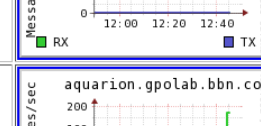
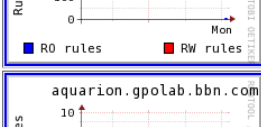
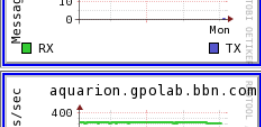
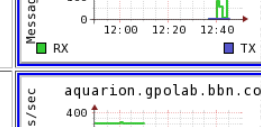
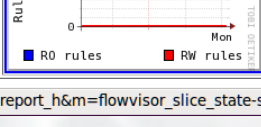
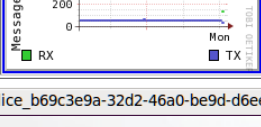
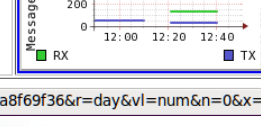


(our production FOAM instance)

<http://monitor.gpolab.bbn.com/tango/flowvisor/tulum.gpolab.bbn.com.html>



## Slices defined on flowvisor/AM

Slice	Type of slice	Flowspace rules (past day)	Flowvisor messages (past day)	Flowvisor messages (past hour)
0ec14a7d-e7b6-4018-9640-f6b85663296f (urn:publicid:1DN+pgeni.gpolab.bbn.com+slice+tut02)	foam sliver (approved, enabled)			
1938ee9f-4515-41f8-bec8-b41f42b0ab33 (urn:publicid:1DN+pgeni.gpolab.bbn.com+slice+slice01)	foam sliver (pending)			
b69c3e9a-32d2-46a0-be9d-d6eea8f69f36 (urn:publicid:1DN+pgeni.gpolab.bbn.com+slice+testdef)	foam sliver (approved, enabled)			
e0b807bd-e575-425d-999f-f0e79bf1771d (urn:publicid:1DN+pgeni.gpolab.bbn.com+slice+tut01)	foam sliver (approved, enabled)			
fvadmin	native slice			

[http://monitor.gpolab.bbn.com/ganglia/graph.php?c=Core&h=aquarion.gpolab.bbn.com&g=flowrule\\_report\\_h&m=flowvisor\\_slice\\_state-slice\\_b69c3e9a-32d2-46a0-be9d-d6eea8f69f36&r=day&v1=num&n=0&x=10&z...](http://monitor.gpolab.bbn.com/ganglia/graph.php?c=Core&h=aquarion.gpolab.bbn.com&g=flowrule_report_h&m=flowvisor_slice_state-slice_b69c3e9a-32d2-46a0-be9d-d6eea8f69f36&r=day&v1=num&n=0&x=10&z...)

(our tutorial FOAM instance – note Pending sliver)

<http://monitor.gpolab.bbn.com/tango/flowvisor/aquarion.gpolab.bbn.com.html>

Status of GENI slice urn:publicid:IDN+pgeni.gpolab.bbn.com+slice+plastic-101 - Mozilla Firefox

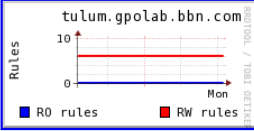
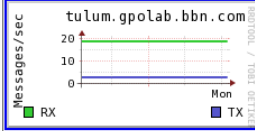
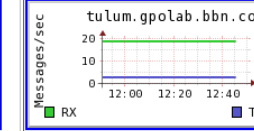
File Edit View History Bookmarks Tools Help

http://monitor.gpolab.bbn.com/tango/slice/urn:publicid:IDN+pgeni.gpolab.bbn.com+slice+plastic-101.html

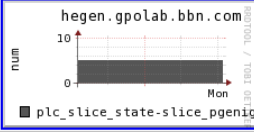
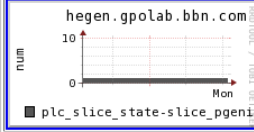
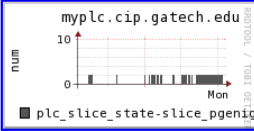
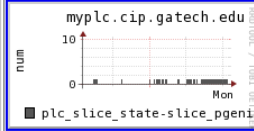
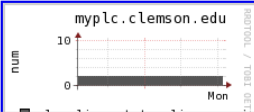
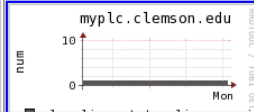
## Status of GENI slice urn:publicid:IDN+pgeni.gpolab.bbn.com+slice+plastic-101

Last update: 2011-10-31 12:42:57.566778

### FlowVisors containing FOAM sliver in slice

Flowvisor	FV slice name	Flowspace rules (past day)	Flowvisor messages (past day)	Flowvisor messages (past hour)
<a href="http://tulum.gpolab.bbn.com">tulum.gpolab.bbn.com</a>	c0b17abe-7274-451f-859e-ef8aa3762f45			

### MyPLC servers containing sliver in slice

MyPLC	Nodes in sliver (past day)	Users in sliver (past day)
<a href="http://hegen.gpolab.bbn.com">hegen.gpolab.bbn.com</a>		
<a href="http://myplc.cip.gatech.edu">myplc.cip.gatech.edu</a>		
<a href="http://myplc.clemson.edu">myplc.clemson.edu</a>		

Done

(things in the plastic-101 slice, at multiple aggregates)

<http://monitor.gpolab.bbn.com/tango/slice/urn:publicid:IDN+pgeni.gpolab.bbn.com+slice+plastic-101.html>

- Pain Point: Approving (“opting in”) slivers is hard
- FOAM helps: An engine for automating approval decisions
- 0.4 only has a default policy setting
- 0.6 will have hooks for other policies, based on:
  - Slice owner
  - Flowspace characteristics (overlaps, or doesn’t, etc)
  - Other fields in the sliver record (e.g. controller URL)
  - Fuzzier things, e.g. “a sliver like this one I approved before”
- Some of those may not make it by the Nov 21<sup>st</sup> release, but will be added soon after

<https://openflow.stanford.edu/bugs/browse/FOAM-152>

- Pain point: Expedient's algorithm for turning rspecs into flowspace was very verbose
- FOAM helps: Its algorithm is much more succinct
- ...and you can see the flowspace, and the number of flowspace rules, with show-sliver
- Our production FV at BBN went from 6000 rules to 200 (higher at the moment due to a legitimately large sliver for a GEC 12 tutorial)

- Introduction [5 min]
  - Sarah Edwards, GPO
- Bottom-Up
  - FOAM: New OpenFlow Aggregate Manager [10 min]
    - Josh Smift, GPO
  - Using SNAPP to Find and Visualize GENI Monitoring Data [20 min]
    - Camilo Viecco, Indiana University & John Meylor, Indiana University
  - Topology [15 min]
    - Chaos Golubitsky, GPO
- Top-Down
  - Monitoring/Mgmt Requirements & Discussion [1 hour]
    - Sarah Edwards, GPO

# Using SNAPP to Find and Visualize GENI Monitoring Data

Camilo H. Viecco  
GMOC- Indiana University



# SNAPP

- A suite of tools initially designed to do SNMP collection and display.
  - Time series based data
  - Expanded to do also NON SNMP data collection
    - Data API
    - Planetlab Comon
- Has a very nice UI and API for programmatic access.
  - You can embed it on your own front/end



# Terminology

- Collections:

Groups of closely related Time series based data. Such as:

- Interface data (in/out pps, in/out bytes).
- Sliver state

Each collection is of one type. (specifies what is measured and its units).

- Categories

Groups of Collections that are related in some manner.

These can be hierarchical. (think iTunes by: artist, by album)

# Using SNAPP

- How to get where you want?
  - Using the browsing (Good for deep categories)
  - Using search (Good for wide categories)
  - Using the portal
- In the GENI world most of we have very wide categories so search is the best option IF you have a keyword(s) for what you are looking for
  - IE. Slice name
  - Resource name

# Using SNAPP(2)

- Currently GMOC hosts we have  $\approx$  4000 time series data (most of it is sliver data).
  - when the unification and SNMP of planetlab is done we will be at around 50000.
- Further most data is not hierarchically tagged (only location + aggregate).
  - Browsing thus is not very effective.
  - We will focus on searching

<http://gmoc-db.grnoc.iu.edu/measurement/>

# Search

- We use a 'google' like syntax

We support only 2 logic operators: and and not

- We can do group intersection: for example:
  - 'category:atla category:openflow' would match all collections that are both in atla and in openflow.

# Use Cases

- Researcher using SNAPP to check his/her experiment
- Campus NOC checking the state of their resources.
- ??? Something else

# Demo

# Demo

Portal:

[http://gmoc-db.grnoc.iu.edu/measurement/  
portal.cgi](http://gmoc-db.grnoc.iu.edu/measurement/portal.cgi)

Regular access:

<http://gmoc-db.grnoc.iu.edu/measurement/>

# Scenario: User (slice creator)

- The search parameter -slice is the key here

The screenshot displays the SNAPP (SNMP Network Analysis and Presentation Package) web interface. At the top, there are logos for GlobalNOC (Global Research Network Operations Center), SNAPP, and GENI MOC (Meta-Operations Center). The interface is divided into a left navigation pane and a main content area.

**Navigation Pane:**

- Buttons: Search, Browse, Saved, Advanced
- Search input: -silver plastic 109
- Buttons: Aggregate Graph, All Graphs Side by Side, Save Current View
- Table with 5 results:

Name	Description
hegen.gpolab.bbn.com-plc_s	hegen.gpolab.bbn.com-plc_slice_state-slice_p
myplc.clemson.edu-plc_slice	myplc.clemson.edu-plc_slice_state-slice_pger
myplc.stanford.edu-plc_slice	myplc.stanford.edu-plc_slice_state-slice_pgen
of.cs.washington.edu-plc_slic	of.cs.washington.edu-plc_slice_state-slice_pg
wings-openflow-1.wail.wisc.e	wings-openflow-1.wail.wisc.edu-plc_slice_stati

5 total results

**Main Content Area:**

- Header: `plc_slice_state-15sec` Parameters
- Text: *Click on a graph to see the more detailed version.*
- Section: Aggregate
- Time Range: Mon 17 Oct 2011 13:15:57 EDT to Mon 24 Oct 2011 13:15:57 EDT
- Graph: A line graph showing two data series over time. The y-axis ranges from 2.5 to 15. The x-axis shows dates from Oct 17 to Oct 23. The orange line (auto-tree-slice\_nodes) fluctuates around 15, while the blue line (auto-tree-slice\_users) fluctuates around 5. Both lines show a sharp drop on Oct 23.
- Legend: auto-tree-slice\_nodes (Crosshair: 15), auto-tree-slice\_users (Crosshair: 5)
- Footer: User: Public, Developed by Global Research NOC Systems Engineering Copyright 2011, The Trustees of Indiana University, Help



# Scenario: Campus NOC

- “clemson -slice -sliver”

GlobalNOC Global Research Network Operations Center

SNAPP SNMP Network Analysis and Presentation Package

gMOC GENI META-OPERATIONS CENTER

Navigation

Search Browse Saved Advanced

-sliver clemson -slice

Aggregate Graph All Graphs Side by Side Save Current View

Name	Description
expedient.clemson.edu-flowvisor_dpид_state-d	expedient.clemson.edu-flowvisor_
expedient.clemson.edu-flowvisor_dpид_state-d	expedient.clemson.edu-flowvisor_
expedient.clemson.edu-flowvisor_dpид_state-d	expedient.clemson.edu-flowvisor_
expedient.clemson.edu-flowvisor_dpид_state-d	expedient.clemson.edu-flowvisor_
expedient.clemson.edu-flowvisor_dpид_state-d	expedient.clemson.edu-flowvisor_
expedient.clemson.edu-flowvisor_dpид_state-d	expedient.clemson.edu-flowvisor_
expedient.clemson.edu-flowvisor_dpид_state-d	expedient.clemson.edu-flowvisor_
expedient.clemson.edu-flowvisor_dpид_state-d	expedient.clemson.edu-flowvisor_
expedient.clemson.edu-flowvisor_dpид_state-d	expedient.clemson.edu-flowvisor_
expedient.clemson.edu-flowvisor_dpид_state-d	expedient.clemson.edu-flowvisor_
myplc.clemson.edu-plc_node_count	myplc.clemson.edu-plc_node_co
myplc.clemson.edu-plc_node_state-node_plar	myplc.clemson.edu-plc_node_sta
myplc.clemson.edu-plc_node_state-node_plar	myplc.clemson.edu-plc_node_sta
myplc.clemson.edu-plc_session_count	myplc.clemson.edu-plc_session_
myplc.clemson.edu-plc_site_count	myplc.clemson.edu-plc_site_cour
myplc.clemson.edu-plc_site_state-site_cu	myplc.clemson.edu-plc_site_state
myplc.clemson.edu-plc_site_state-site_pgenig	myplc.clemson.edu-plc_site_state
planetlab4.clemson.edu-phnode_state	planetlab4.clemson.edu-phnode_s
planetlab5.clemson.edu-phnode_state	planetlab5.clemson.edu-phnode_s

15 total results

<< first < prev 1 next > last >>

Click on a graph to see the more detailed version.

plc\_node\_count-15sec Parameters

myplc.clemson.edu-plc\_node\_count  
myplc.clemson.edu-plc\_node\_count-auto  
Mon 24 Oct 2011 13:24:58 EDT to Mon 24 Oct 2011 14:24:58 EDT

Legend:

- auto-tree-plc\_boot\_nodes Crosshair: 2
- auto-tree-plc\_disabled\_nodes Crosshair: 0
- auto-tree-plc\_fallboot\_nodes Crosshair: 0
- auto-tree-plc\_install\_nodes Crosshair: 0
- auto-tree-plc\_nodes Crosshair: 2
- auto-tree-plc\_reinstall\_nodes Crosshair: 0
- auto-tree-plc\_safeboot\_nodes Crosshair: 0

plc\_node\_state-15sec Parameters

User: Public

Developed by Global Research NOC Systems Engineering Copyright 2011, The Trustees of Indiana University

Help

# Graph Controls

Parameters:

- Series (what do you want to see)
- Time range. (from when to where)
- For aggregate graphs you also have the options of stacked/overlay graphs.

# Getting help

There is documentation on the page:

See the bottom of the page

Campuses:

What do you want on your portal?

Do you want another portal?

# SNAPP- Monitoring GMOC Roadmap

- Allow a single snapp frontend to also query data from remote backends (mid 2012)
- Time machine (fall 2012)
- Reporting (fall 2012)
- Auth integration
- Integration with the current monitoring by GPO

END

# Nagios

- Public Face of the data reported by our internal collections
  - Currently limited to OF devices on the mesoscale effort

# SNAPP UI

- Two mechanisms to get data
  - Browsing
  - Search
- Two types of data
  - Collections: Sets or related time series data
  - Categories: groups of collections with hierarcchy

- Introduction [5 min]
  - Sarah Edwards, GPO
- Bottom-Up
  - FOAM: New OpenFlow Aggregate Manager [10 min]
    - Josh Smift, GPO
  - Using SNAPP to Find and Visualize GENI Monitoring Data [20 min]
    - Camilo Viecco, Indiana University & John Meylor, Indiana University
  - Topology [15 min]
    - Chaos Golubitsky, GPO
- Top-Down
  - Monitoring/Mgmt Requirements & Discussion [1 hour]
    - Sarah Edwards, GPO





# GENI Monitoring

## Topology in the Mesoscale

Chaos Golubitsky, GPO  
GEC12 Nov 3, 2011  
[www.geni.net](http://www.geni.net)



- Layer 2 mesoscale network is different from commercial networks:
  - Topology contains intentional loops, for use by experimenters in path selection experiments
  - Experimenters know how packets are switched and routed through the multisite network
  - Experimenters modify packet switching and routing, independently from other experiments
- Mesoscale has been operational for a year now
  - Examples of network problems seen so far in the wild
  - Implications for future monitoring and debugging

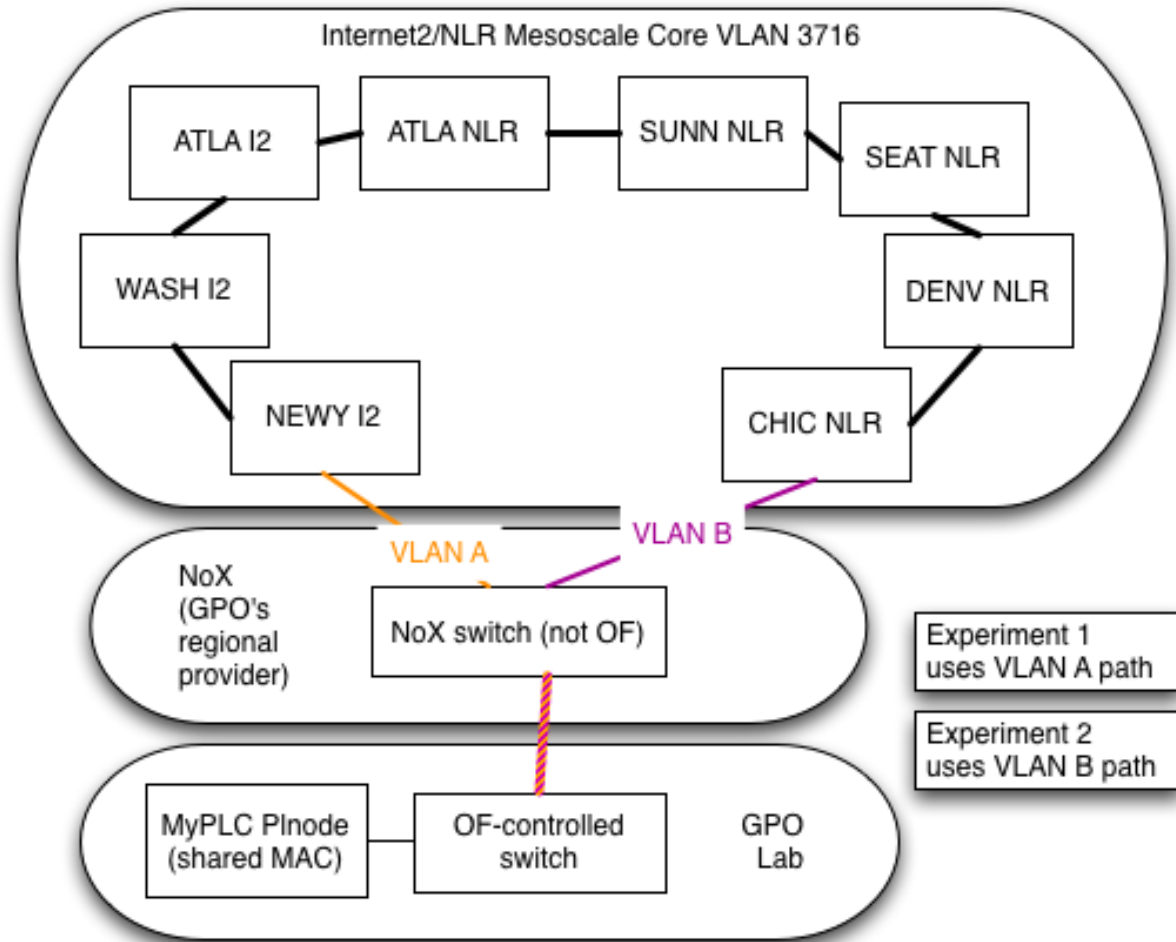
- What can happen?
  - Mesoscale core consists of two VLANs
  - VLANs are bridged at each campus by an OpenFlow-controlled VLAN
  - If bridge fails open, traffic leaks between core VLANs
- Why does it happen?
  - Interface misconfiguration
  - OpenFlow VLAN fails open (default config on HP)
  - VLAN loses its OF configuration entirely and reverts to learning switch
  - Experiment misconfiguration

# Detecting uncontrolled VLAN bridging

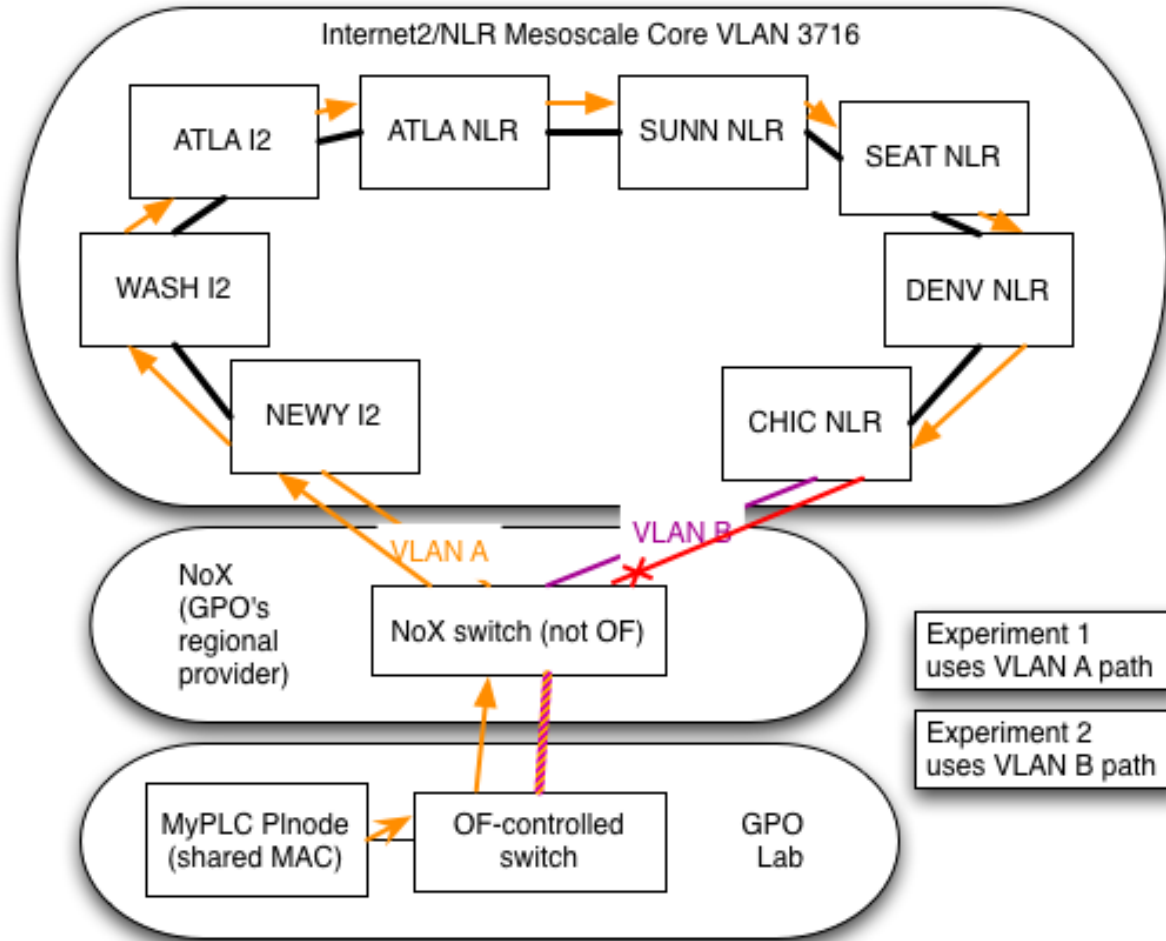
- Detection of severe failures is easy:
  - Define traffic which should never be seen on each VLAN (IP monitoring subnet of other VLAN)
  - Alert if you see that traffic
  - <http://groups.geni.net/geni/NetworkCore/TrafficLeaks>
- Tracking down failures so they can be resolved:
  - GMOC and GPO have tracked down many severe failures using suspicious MAC address table patterns
  - But: the leaking traffic is not the traffic causing the leak
  - Intermittent failures are much harder to find
- How do we find leaks affecting only a subset of (non-monitoring) traffic?

- Why MACs get shared between experiments:
  - Physical interfaces are expensive
  - Subinterfacing puts one interface on many networks
  - Default (Linux) subinterfaces share MAC address
  - Examples: mesoscale MyPLC, monitoring nodes
- When this causes problems:
  - Different topologies using the same MACs on the same devices
  - Depending on bridging, maybe even the same VLANs
  - Complicated by non-OF devices in an OF path

# Shared MAC Problems: GPO Example (1)



# Shared MAC Problems: GPO Example (2)





# Shared MAC Problems: Detection

- **NoX detected this problem via syslog:**

```
Sep  8 11:46:18 EDT: %SW_MATM-4-MACFLAP_NOTIF: Host 001b.215f.8fe1  
in vlan 3745 is flapping between port Gi0/9 and port Te0/1
```

```
Sep  8 11:47:16 EDT: %SW_MATM-4-MACFLAP_NOTIF: Host 001b.214b.3fad  
in vlan 3745 is flapping between port Gi0/9 and port Te0/1
```

- **General solutions feasible in the short term:**

- Track MAC addresses of known hosts, and alert if you see them somewhere unexpected
- Collect problematic log messages which might be seen on non-OpenFlow devices connecting OpenFlow networks; ask regionals to watch for them



## What these problems have in common

- They are ways the topology can fail to match experimenter expectations
  - ...which could lead to experimental results mismatching expectations
- These problems have occurred “in the wild” in the early mesoscale deployment
- We want to accurately detect problems like these
  - GENI software will help in the long term
  - Some simple diagnostics have been useful so far
  - We can do more in the near term

- **Stitching:**
  - Design a custom end-to-end per-experiment topology
  - Virtualize resources which we manually provision, and often share, right now (e.g. VLANs, MACs)
  - Programmatic verification of topology may be part of topology reservation
- **Instrumentation and Measurement:**
  - Richer capture of activity on intermediate networks to help experimenters validate the topology they expect
  - Within-slice on-demand measurement to verify complex experimental configurations

- Define something you expect not to see, and alert if you see it. Examples:
  - Bridged VLAN detection (“Monitoring subnets A and B are distinct. If they see each other’s traffic, something is wrong”)
  - Shared MAC address (“MAC address M belongs to site B. If it is learned on an interface facing away from site B, something is wrong”)
- These checks are easy to do
- With reasonable definitions, false positives are low

- Tools to improve detection and response:
  - “Known bad symptom” tests which can run at sites, regionals, and backbones
    - Then share code for those tests, and run them
    - Simple things for regionals to be aware of are especially useful
  - Monitor MAC addresses learned by experimental switches, and alert on unexpected paths when feasible
- Other suggestions?
- Thank you for helping operate the mesoscale!

- Introduction [5 min]
  - Sarah Edwards, GPO
- Bottom-Up
  - FOAM: New OpenFlow Aggregate Manager [10 min]
    - Josh Smift, GPO
  - Using SNAPP to Find and Visualize GENI Monitoring Data [20 min]
    - Camilo Viecco, Indiana University & John Meylor, Indiana University
  - Topology [15 min]
    - Chaos Golubitsky, GPO
- Top-Down
  - Monitoring/Mgmt Requirements & Discussion [1 hour]
    - Sarah Edwards, GPO

# Monitoring & Management Requirements

Sarah Edwards, GPO



- Introduction
- Breakdown of Top Level Requirements
- Data Requirements
- Conclusions
  
- There is time built in to discuss as we go.

# INTRODUCTION



# Why are we here?

- Would like to ensure:
  - We don't ignore any important pieces
  - Architecture decisions reflect needs of monitoring so that GENI Clearinghouse, I&M, etc serve our needs
  - Where possible, we build tools which can be adapted to new software when it becomes available
- Therefore we need to answer the following questions:
  - 1) Are there any holes in our understanding of our requirements?
  - 2) Do we have agreement on what we need in a way that can be communicated to other groups working on topics of interest to ops?
  - 3) Do we know how to build tools in an adaptable way?
- And also...
  - 4) What do we work on next?

We'll come back to these questions at the end

- 10.2-3 Visible operational status
  - The GENI system shall make sufficient data available that researchers and maintainers will be able to evaluate the availability and **operational** status of the system.
- 10.2-5 Federated event escalation
  - The GENI system shall provide operations and **management** support for event **management** and escalation, including security events, within GENI and with those organizations that interconnect with GENI.
- 10.2-6 Federated operations data exchange
  - The GENI system shall support operational and **management** data exchange according to [TBS] GENI O&M Policy between GENI and operators/owners of federated components, aggregates, and networks.

- Meta-operations (a.k.a. GMOC)
- Aggregates
  - Examples: ProtoGENI, PlanetLab, Orca
- Campuses
  - Which host & run aggregates
  - Which only host aggregates
- Backbone & Regional Networks
  - Networks which are GENI participants: I2, NLR
  - Networks which carry GENI traffic: some regionals (eg NoX)
- Experimenters

# Aggregates, Racks and the GENI CH

- Each GENI rack is a **SINGLE** aggregate
  - Therefore requirements are the same as for aggregates
- Aggregates can outsource (some of) their responsibilities to the GENI Clearinghouse

- Monitoring
  - Act of collecting data and measuring what is happening
- Management
  - Act of fixing problems and responding to requests
- What does monitoring & management involve?
  - Observe unexpected events
    - THEN fix what's wrong
  - Observe expected events
    - THEN develop policy for fixing what's wrong
    - THEN fix what's wrong (by responding to monitoring)
  - Plan for the future
    - Monitor long-term trends in resource usage
    - THEN provision resources to meet forecasted needs

# What makes GENI different?

- Federated entities managed by different institutions with different policies
- People and information needed to troubleshoot and resolve problems are spread across several physical locations
- Users (end user and experimenters), managers and hosts of a given piece of equipment may all be different.
- Interactions between groups are governed by GENI federation agreements (e.g. aggregate provider agreement) and mutual understanding.

- We are not covering:
  - Monitoring and management which fits entirely within the purview of aggregates, campuses, etc
- For example, we will do (but not discuss here)
  - Keeping logs
  - Obeying local laws and policy
  - Answering the phone when someone has a concern
  - ... and tie your shoes and everything else.
- These things do **NOT** make GENI different



# Top-level Aspects of GENI Monitoring

- Information must be shareable
- Information must be collected
- Information must be available when needed
- Cross-GENI operational statistics collected and synthesized to indicate GENI as a whole is working
- Preserve privacy of users (opt-in, experimenters, other users of resources)



# Top-level Aspects of GENI Management

- For both debugging and security problems:
  - Must be possible to escalate events
  - Meta-operations and aggregate operators must work together to resolve problems in a timely manner
- Must be possible to do an emergency stop in case of a problem
- Orgs must manage GENI resources consistent with local policy and best practices
  - e.g security procedures, logging, backups, etc
- Develop policies for monitoring
- All parties should implement agreed upon policies
- Security of GENI as a whole and its pieces

# BREAKDOWN OF TOP LEVEL REQUIREMENTS

- *GENI monitoring is more than the sum of the monitoring at GENI's parts. In order to know if GENI is working properly, additional monitoring is required beyond that done by each of its constituent pieces.*
- Collect and synthesize additional operational statistics which indicate whether GENI is working
  - e.g. meso-scale ping tests, topology
  - Collect cross-GENI stats
  - Make cross-GENI stats available when needed

- Preserve privacy of users (opt-in, experimenters, other users of resources)
  - → TBD – This is an area needing major discussion

# Requirement: Troubleshooting & Event Escalation

- For both debugging and security problems:
  - Meta-operations and aggregate operators must work together to resolve problems
    - Aggregates must advertise resources accurately
      - (threshold) statically → **Fill out aggregate page**
      - (objective) dynamically → **Advertise resources via AM API**
    - Aggregates notify meta-operations when resources are unavailable → **via e-mail (doing SOME of the time)**
    - Aggregates cooperate with meta-operations on the resolution of security events
    - Aggregates cooperate with LLR on the resolution of security events
  - Must be possible to escalate events

- Must be possible to do an emergency stop in case of a problem
  - Must maintain POC information at meta-operations
    - Aggregate → send contact info to GMOC
    - Campus → send contact info to GMOC
    - Experimenter → slice e-mail
    - Other infrastructure → contacted by relevant campus
  - Aggregates & Meta-operations must each have policies and procedures in place to support an emergency stop
    - → Has been dry run

- Orgs must manage GENI resources consistent with local policy and best practices (e.g security procedures, logging, backups, etc)
  - In general, follow local policy and procedures
  - Follow best practices which if not followed would affect other members of the GENI community
- Develop policies for monitoring
- All parties should implement agreed upon policies
  - Follow Aggregate Provider Agreement
  - Follow LLR
  - Follow other GENI policies as they come into effect



- Security of GENI as a whole and its pieces
  - Two things we want to prevent:
    - Compromise of GENI resources
    - Use of GENI resources to compromise other entities
  - Two things we can do about this:
    - Follow best practices to hinder compromise
    - Detect and respond to compromise
  - Allow interesting research for which experimenters and operations have to coordinate for security and management reasons
    - Security experimentation BOF tonight and session at GEC13!!!
  - → **TBD – This is an area needing major discussion**



# Info must be shareable/collected/available

- Information must be shareable
  - Consistent definitions of data
  - Consistent data exchange format
  - Consistent data collection mechanisms
  - Data sharing mechanisms
  - The following benefit from shared common processes:
    - Accessing data, finding data, visualizing data
- Information must be collected
  - Verify continued successful data collection
  - Debug collection and reliability outages
- Information must be available when needed
  - Privacy of data must be maintained

# DATA REQUIREMENTS

- Consistent definition of data
  - Relational data
    - Resources (incl. connectivity)
    - List of aggregates
    - List of slices
    - List of users
    - Aggregate contact information
  - Timeseries data
    - Examples: Host and network statistics
  - Events
    - Examples: SNMP Traps

# Data Storage & Collection Methods

- Data collection methods
  - Relational data → store in relational DB
    - Resources → **Rspecs available via AM API**
    - List of aggregates → **ctrl framework clearinghouse & GENI wiki**
    - List of slices → **control framework slice authority**
    - List of users → **TBD**
    - Aggregate contact information → **aggregate page and GMOC**
  - Timeseries data → **store in RRD**
    - → **collect via SNMP (ie host and network stats)**
    - → **by asking the aggregate (ie custom OpenFlow API)**
  - Events → **store in relational DB (?)**
    - → **TBD**

- Sharing Data
  - → publish to central DB at GMOC
  - → publish locally via webpage or local API
  - → TBD: publish via a distributed mechanism
- Accessing, Finding and Visualizing Data
  - → GMOC Portals
  - → GMOC SNAPP Interface (with search)
  - → GMOC data available to interested consumers via API
  - → TBD: More to do here

## Other people who need data

- Troubleshooting info from aggregates, campuses, meta-operations
- Accountability report: How to prove if this is not my fault?

# CONCLUSIONS

# Why are we here? [repeat]

- Would like to ensure:
  - We don't ignore any important pieces
  - Architecture decisions reflect needs of monitoring so that GENI Clearinghouse, I&M, etc serve our needs
  - Where possible, we build tools which can be adapted to new software when it becomes available
- Therefore we need to answer the following questions:
  - 1) Are there any holes in our understanding of our requirements?
  - 2) Do we have agreement on what we need in a way that can be communicated to other groups working on topics of interest to ops?
  - 3) Do we know how to build tools in an adaptable way?
- And also...
  - 4) What do we work on next?



# 1) Are there any holes?

- From above:
  - Security
  - Privacy
  - Topology
    - Can stitching or other SW efforts help?
  - Accessing, finding and visualizing data
- From yesterday:
  - Topology
  - Event notification system
  - How does a third party help troubleshoot a slice?
    - Slice traceroute?
  - Universal names (eg for circuits)
    - need to coordinate with the software group

## 2) Agreement on requirements?

## 3) Building adaptable tools?

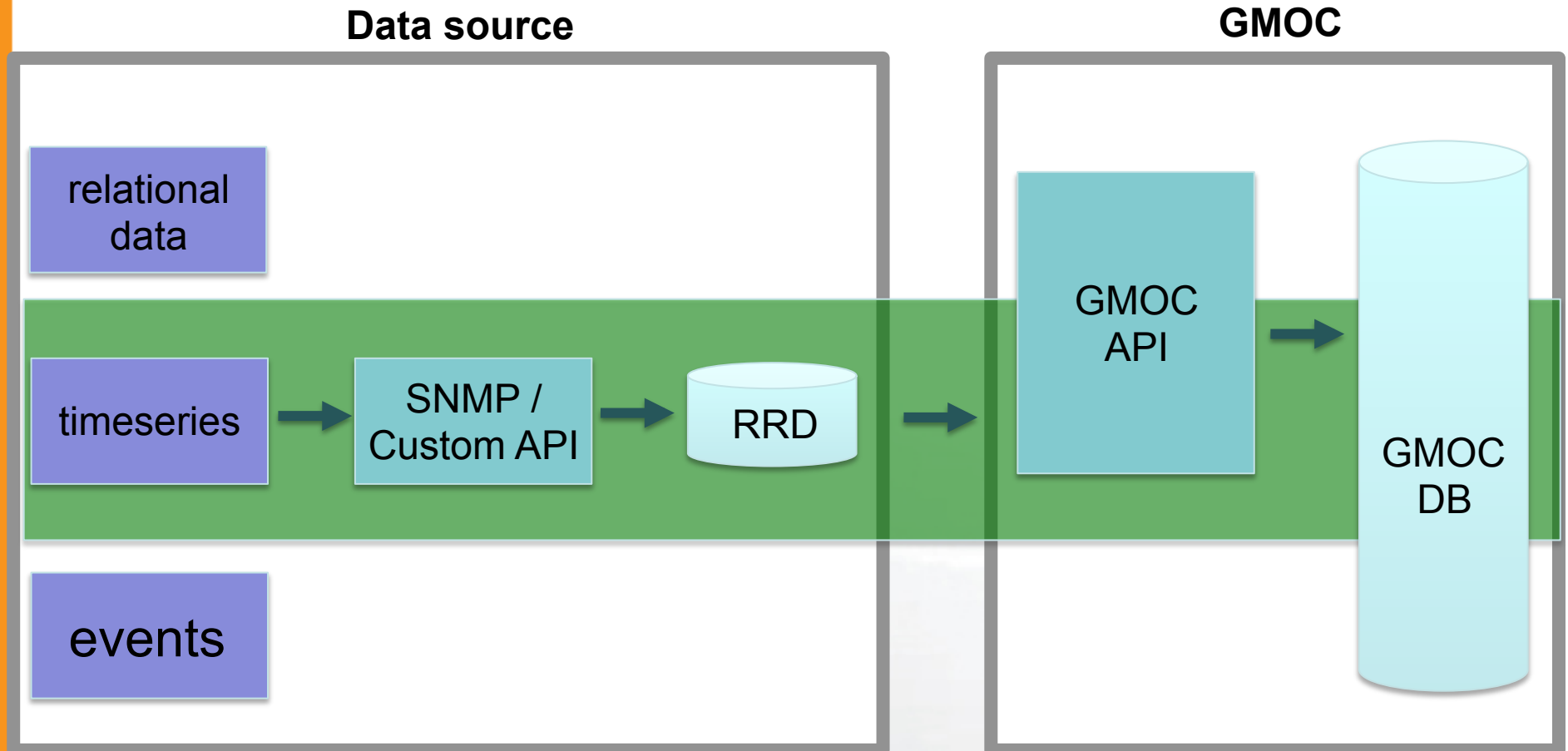
relational  
data

timeseries

events

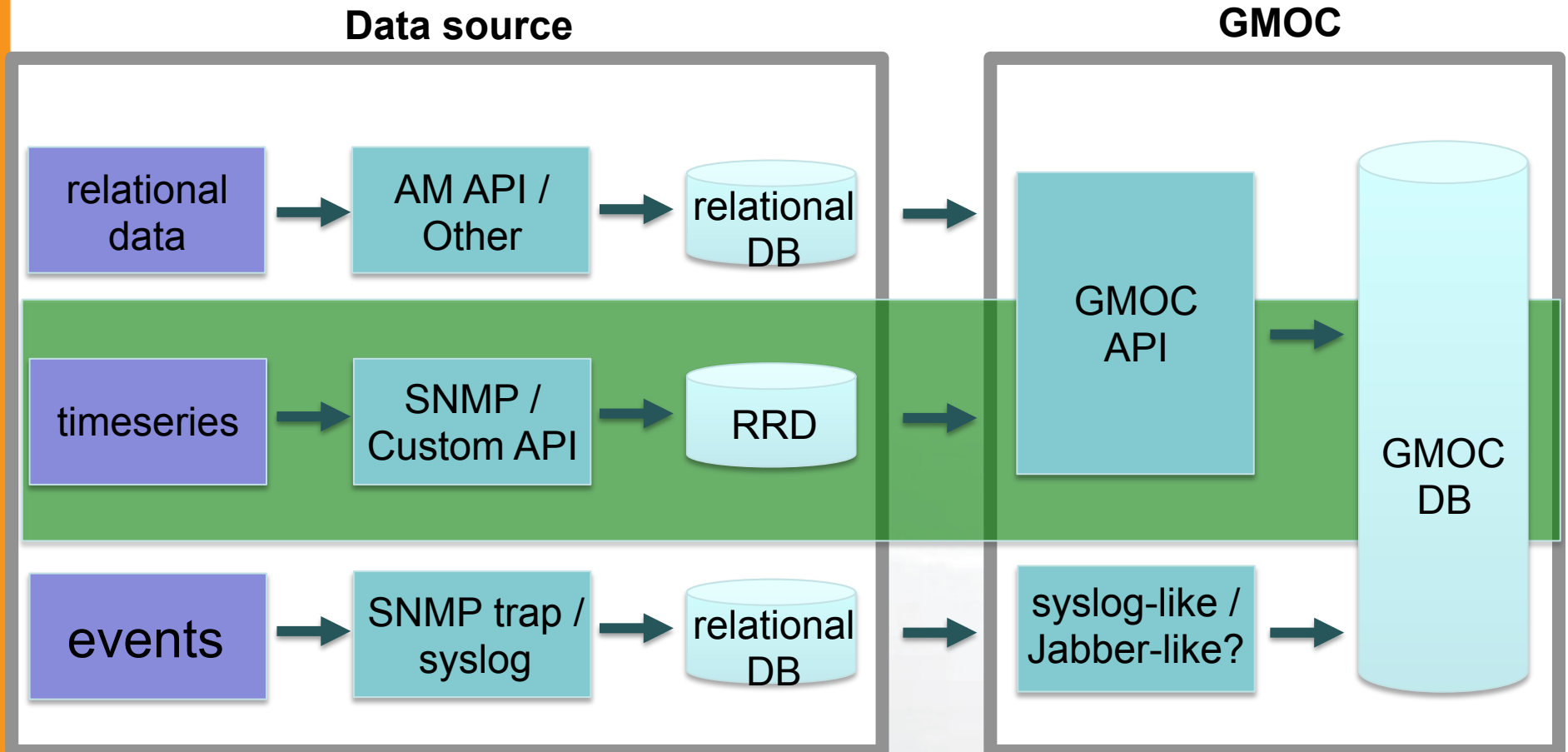
# 3) Building adaptable tools?

## Publishing Data



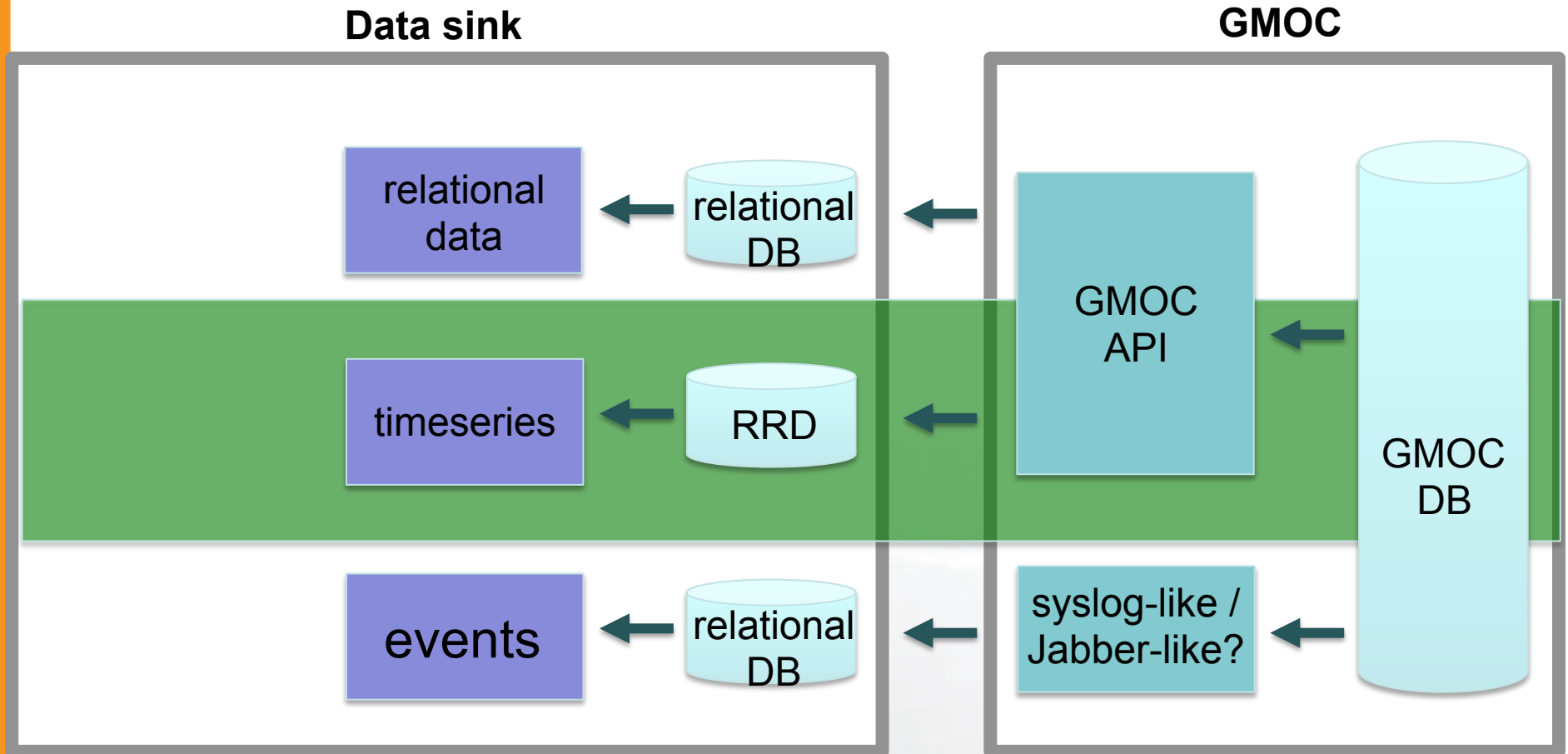
# 3) Building adaptable tools?

## Publishing Data






### 3) Building adaptable tools?

## Accessing Data



## 4) What's next?

Status		Requirement/Pain Point
	(11)	
	(12)	
	(13)	
	(14)	
	(15)	
	(16)	
	(17)	
	(18)	
	(19)	
	(20)	
	(21)	Others?