GENI Operations: Monitoring

Monitoring with distributed sources and collection points

Ryan Irwin, GPO

GEC19: March 18, 2014

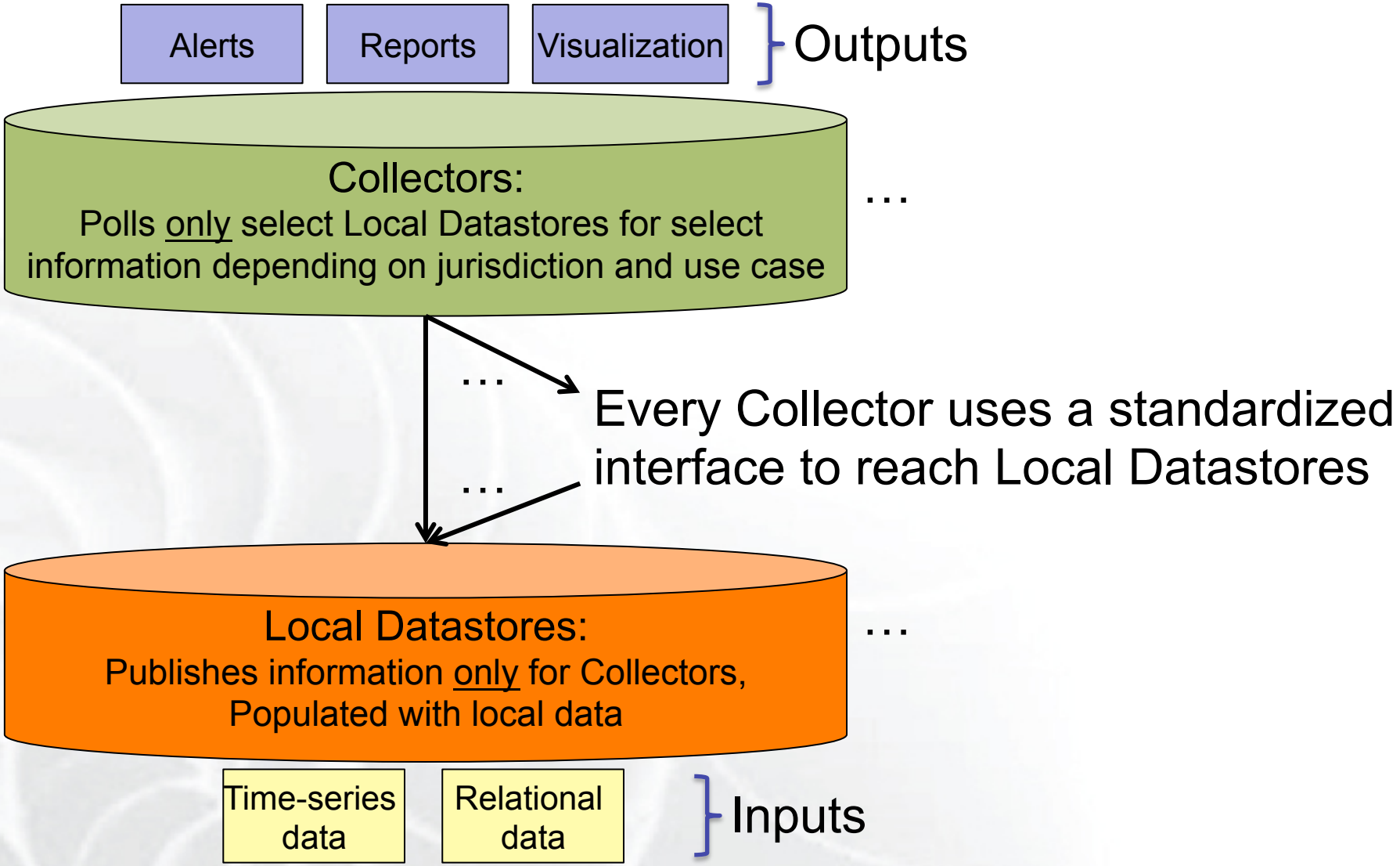
Motivations

- Number of teams participating in GENI is growing
  - Teams have interest in collecting operational data from different sources for different purposes
  - Teams need a unified way for publishing data and for others to have access to it
- Need an approach that scales with additional aggregates and has fewer bottlenecks
- Previous design did not specify use of common tools in a collaboratively monitored system

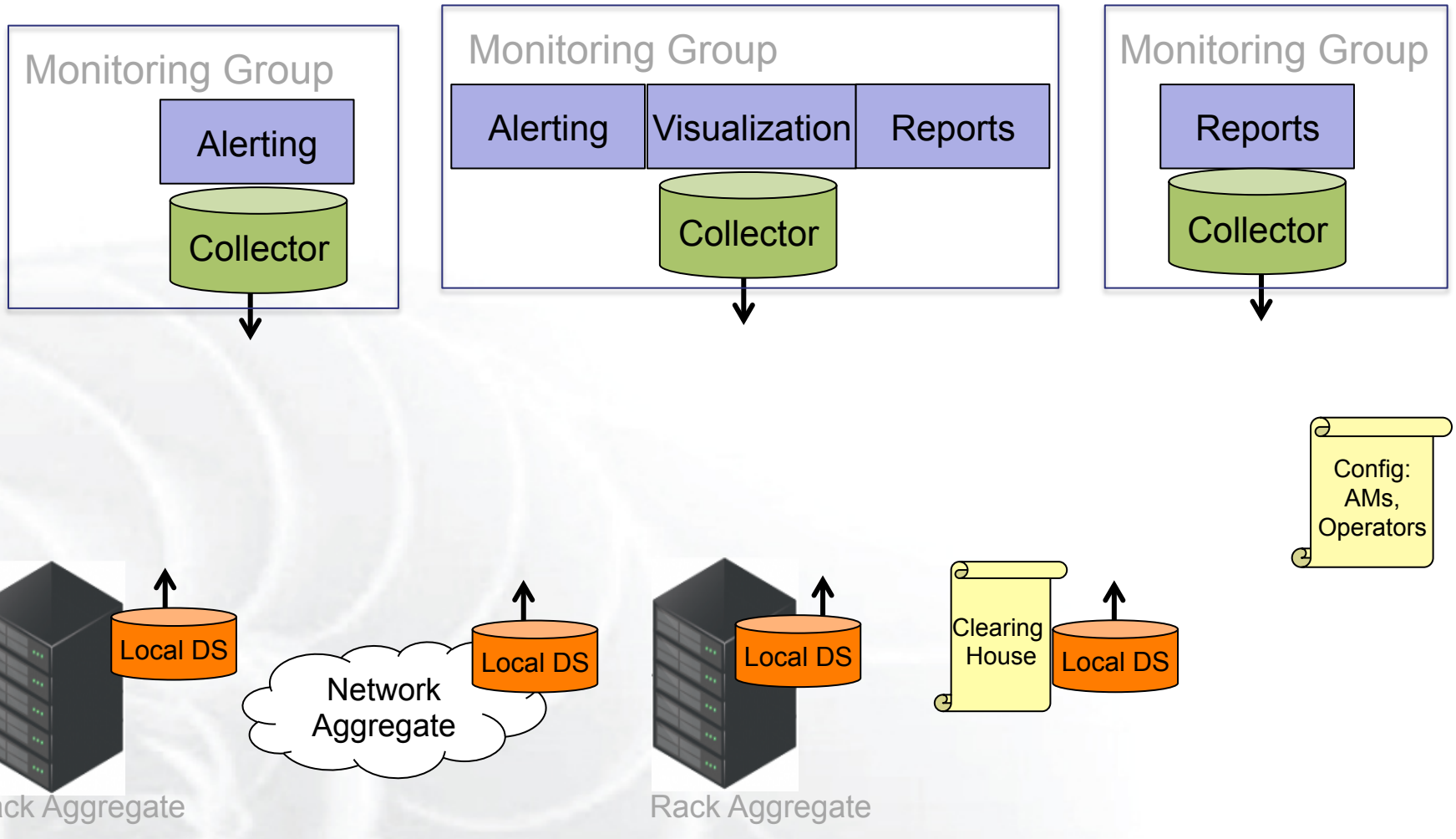
Overview of Approach

- Data sources and collection points are distributed
  - Main requirement: aggregates have data about their aggregate accessible through an interface
  - Collection points access data sources pertaining to their supported use case(s) and monitoring domain
- Enables a diverse set of operator-preferred tools to be used in a coordinated fashion
- Extensible to new aggregates and monitoring metrics
- Current implementation will transition to the GENI community

Overview of Architecture



Implementation in GENI



REST and JSON Based Framework

- Two types of REST calls to the local datastores
- Info about local datastore collections
  - Pointers to objects with in the local datastore through object IDs
  - Implemented examples: ID's of nodes, interfaces, vlans
  - Examples of things to come: slice, sliver, user
- Data about measurements or time-series data
  - Query a set of object IDs and measurements
  - Implemented examples: Utilization statistics (cpu_util, rx_bps)
  - Example of things to come: AM availability, OF statistics
- Both types of data support alerting, reporting, and visualizations

Info REST Calls

- A collector gets information about the contents of the local datastore
  - Response contains references to get info on the aggregate's resources and slivers
  - Response contains a reference to get time-series data

Ex: http://datastore.utah.geniracks.net:5001/info/aggregate/utah-ig

- A collector makes successive info calls to gather all the info of interest at a local datastore
  - Like a web-crawler that gathers info
  - Saves object IDs for making time-series data calls

Ex: http://datastore.utah.geniracks.net:5001/info/node/utah.geniracks.net_node_pc3

Data REST Calls

- Then, the collector makes periodic data queries for object id's gathered from info queries

  http://datastore.utah.geniracks.net:5001/?q=<*filters_dictionary*>

- Filters dictionary contains:
  - Event types (i.e., cpu_util, rx_bps)
  - Timestamp window
  - Object IDs of a single type (i.e., node or interface)
    - Object IDs match those used in info calls
  - JSON response is an array of length num_event_type x num_object_id
- Time window is an implementation design point

Implement to a standard

- Aggregates required to have local datastores
- Local datastores are validated
  - Required to follow the api and have consistent data (i.e., consistent object IDs)
- Validation tools
  - json-schema validation similar to UNIS (Martin Swany)
  - Custom validation tools

Local DS and Collector Reference Implementation

- Standard PostgreSQL and MySQL DBs
- Python programs for:
  - Transforming REST calls to SQL queries and JSON response creation
  - *Crawling* local datastores for info (called by CRON)
  - Fetching time-series data (CRON)
- Apache, Python-Flask, and WSGI for hardened, multi-threaded, webfacing servers
- Code: http://trac.gpolab.bbn.com/ops-monitoring/wiki
- Docs: http://groups.geni.net/geni/wiki/OperationalMonitoring

Deployment Status

- Local Datastores:
  - Rack teams:
    - IG @ UTAH
    - EG @ RENCI,WVNET
  - Network aggregates:
    - AL2S
    - Internet2 (mesoscale)
    - ION
    - MAX (regional access to ION)
- Collectors:
  - UK is visualizing networking data from networking aggregate and Utah IG local datastores (see demo)
  - GPO is alerting on multiple resources using Nagios
  - Reference implementation enables new collectors

ExoGENI Info Call

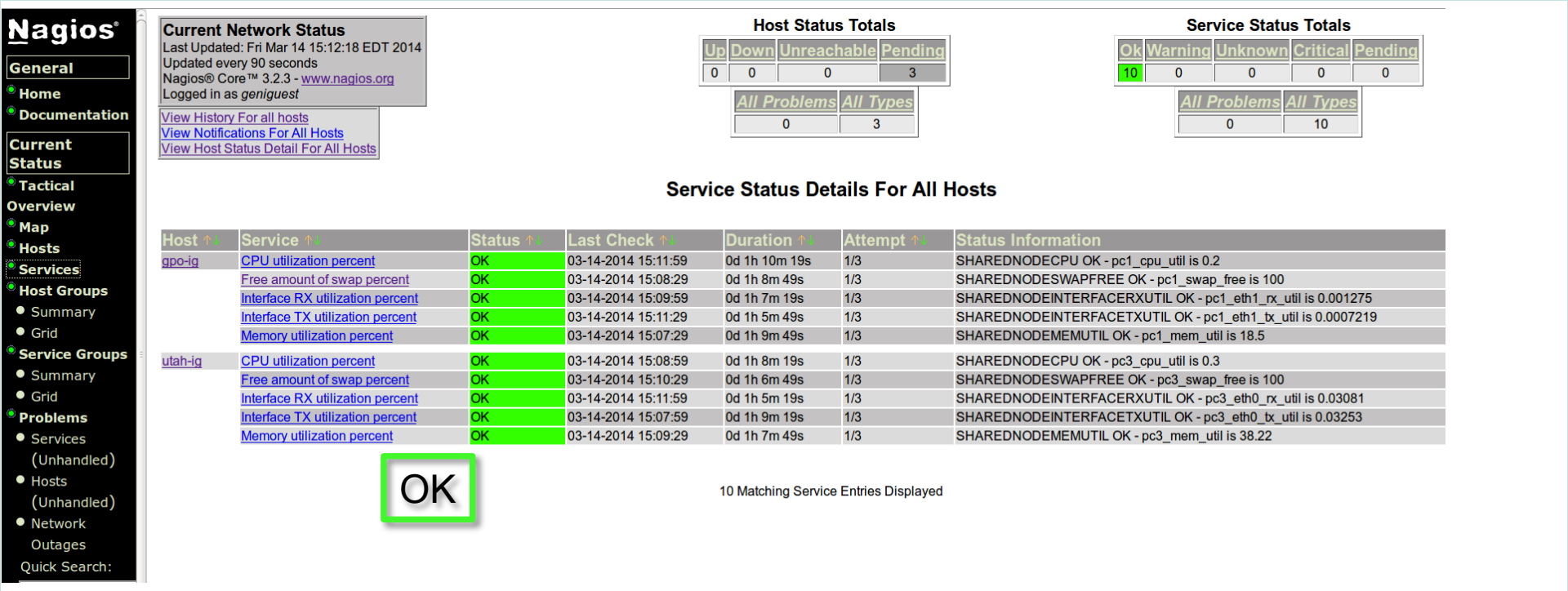
https://wvn-hn.exogeni.net/ops-monitoring/info/node/exogeni.net:wvnvmsite+node+wvn-w7

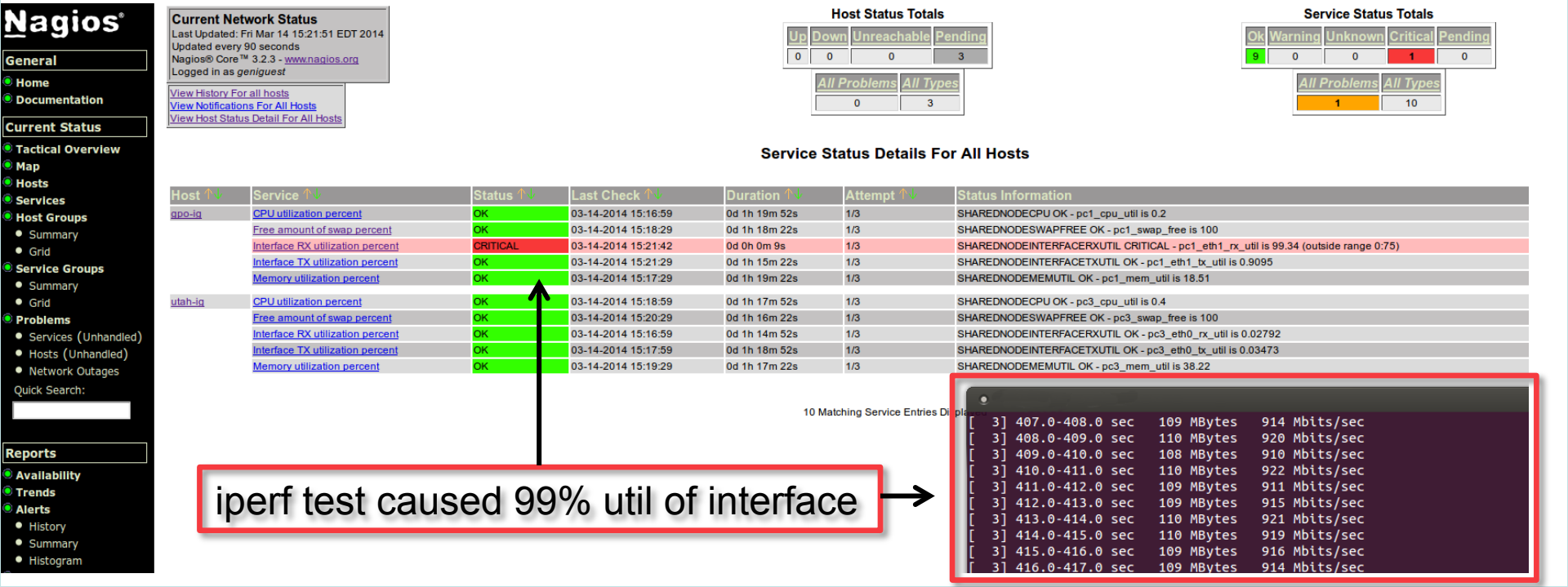
{"selfRef": "https://wvn-hn.exogeni.net/ops-monitoring/info/node/exogeni.net:wvnvmsite+node+wvn-w7",
"urn": "urn:publicid:IDN+exogeni.net:wvnvmsite+node+wvn-w7",
"ts": 1395163151000000,
"id": "exogeni.net:wvnvmsite+node+wvn-w7",
"$schema": "http://unis.incntre.iu.edu/schema/20120709/node#",
"properties": {"ops_monitoring": {"mem_total_kb": 1586307072}},
"ports": [{"urn": "urn:publicid:IDN+exogeni.net:wvnvmsite+interface+wvn-w7:eth3",
"href": "https://wvn-hn.exogeni.net/ops-monitoring/info/interface/exogeni.net:wvnvmsite+interface+wvn-w7:eth3"},
…
{"urn": "urn:publicid:IDN+exogeni.net:wvnvmsite+interface+wvn-w7:eth0",
"href": "https://wvn- hn.exogeni.net/ops-monitoring/info/interface/exogeni.net:wvnvmsite+interface+wvn-w7:eth0"}]}

AL2S Data Call

http://aj-dev6.grnoc.iu.edu/geni-local-datastore/data/al2s.net.internet2.edu/?q={ "filters":
{"eventType":["ops_monitoring:tx_bps","ops_monitoring:rx_bps"],
"ts":{"gte":1391192225475202,"lt":1391192326480000},
"obj":{"type":"interface","id":["sdn-sw.atla.net.interne2.edu:100GigabitEthernet1/1","sdn-sw.atla.net.internet2.edu:
100GigabitEthernet1/2"] } } }

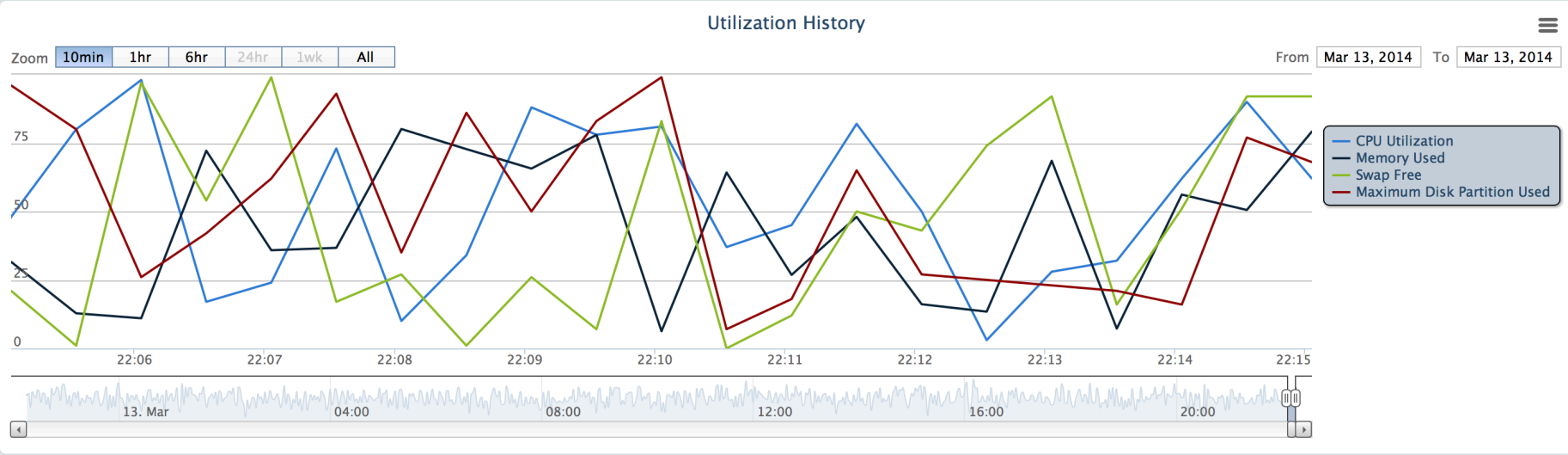
[
{"eventType":"ops_monitoring:rx_bps",
"subject":"http://aj-dev6.grnoc.iu.edu/geni-local-datastore/info/interface/al2s.net.internet2.edu/sdn-sw.atla.net.internet2.edu/100GigabitEthernet1/2",
"id":"rx_bps:al2s.net.internet2.edu+interface+sdn-sw.atla.net.internet2.edu:100GigabitEthernet1/2",
"description":"bits per second received on this interface",
"units":"float",
"$schema":"http://www.gpolab.bbn.com/monitoring/schema/20140131/data#",
"tsdata":[{"ts":1391192220000000,"v":47856109814.4}, ... ,{"ts":1391192330000000,"v":49359791877.12}]}*condensed*
,
{"eventType":"ops_monitoring:tx_bps",
"subject":"http://aj-dev6.grnoc.iu.edu/geni-local-datastore/info/interface/al2s.net.internet2.edu/sdn-sw.atla.net.internet2.edu/100GigabitEthernet1/2",
"id":"tx_bps:al2s.net.internet2.edu+interface+sdn-sw.atla.net.internet2.edu:100GigabitEthernet1/2",
"description":"bits per second sent on this interface",
"units":"float",
"$schema":"http://www.gpolab.bbn.com/monitoring/schema/20140131/data#",
"tsdata":[{"ts":1391192220000000,"v":35551686048.64}, ... ,{"ts":1391192230000000,"v":36441987075.2}]}*condensed*
]


Nagios
General
Home
Documentation
Current Status
Tactical Overview
Map
Hosts
Services
Host Groups
Summary
Grid
Service Groups
Summary
Grid
Problems
Services (Unhandled)
Hosts (Unhandled)
Network Outages
Quick Search:
Current Network Status
Last Updated: Fri Mar 14 15:12:18 EDT 2014
Updated every 90 seconds
Nagios® Core™ 3.2.3 - www.nagios.org
Logged in as geniguest
View History For all hosts
View Notifications For All Hosts
View Host Status Detail For All Hosts
Host Status Totals
Up Down Unreachable Pending
0 0 0 3
All Problems All Types
0 3
Service Status Totals
Ok Warning Unknown Critical Pending
10 0 0 0 0
All Problems All Types
0 10
Service Status Details For All Hosts
Host Service Status Last Check Duration Attempt Status Information
gpo-ig CPU utilization percent OK 03-14-2014 15:11:59 0d 1h 10m 19s 1/3 SHAREDNODECPU OK - pc1_cpu_util is 0.2
Free amount of swap percent OK 03-14-2014 15:08:29 0d 1h 8m 49s 1/3 SHAREDNODESWAPFREE OK - pc1_swap_free is 100
Interface RX utilization percent OK 03-14-2014 15:09:59 0d 1h 7m 19s 1/3 SHAREDNODEINTERFACERXUTIL OK - pc1_eth1_rx_util is 0.001275
Interface TX utilization percent OK 03-14-2014 15:11:29 0d 1h 5m 49s 1/3 SHAREDNODEINTERFACETXUTIL OK - pc1_eth1_tx_util is 0.0007219
Memory utilization percent OK 03-14-2014 15:07:29 0d 1h 9m 49s 1/3 SHAREDNODEMEMUTIL OK - pc1_mem_util is 18.5
utah-ig CPU utilization percent OK 03-14-2014 15:08:59 0d 1h 8m 19s 1/3 SHAREDNODECPU OK - pc3_cpu_util is 0.3
Free amount of swap percent OK 03-14-2014 15:10:29 0d 1h 6m 49s 1/3 SHAREDNODESWAPFREE OK - pc3_swap_free is 100
Interface RX utilization percent OK 03-14-2014 15:11:59 0d 1h 5m 19s 1/3 SHAREDNODEINTERFACERXUTIL OK - pc3_eth0_rx_util is 0.03081
Interface TX utilization percent OK 03-14-2014 15:07:59 0d 1h 9m 19s 1/3 SHAREDNODEINTERFACETXUTIL OK - pc3_eth0_tx_util is 0.03253
Memory utilization percent OK 03-14-2014 15:09:29 0d 1h 7m 49s 1/3 SHAREDNODEMEMUTIL OK - pc3_mem_util is 38.22
OK
10 Matching Service Entries Displayed


Nagios
General
Home
Documentation
Current Status
Tactical Overview
Map
Hosts
Services
Host Groups
Summary
Grid
Service Groups
Summary
Grid
Problems
Services (Unhandled)
Hosts (Unhandled)
Network Outages
Quick Search:
Reports
Availability
Trends
Alerts
History
Summary
Histogram
Current Network Status
Last Updated: Fri Mar 14 15:21:51 EDT 2014
Updated every 90 seconds
Nagios® Core™ 3.2.3 - www.nagios.org
Logged in as geniguest
View History For all hosts
View Notifications For All Hosts
View Host Status Detail For All Hosts
Host Status Totals
Up Down Unreachable Pending
0 0 0 3
All Problems All Types
0 3
Service Status Totals
Ok Warning Unknown Critical Pending
9 0 0 1 0
All Problems All Types
1 10
Service Status Details For All Hosts
Host Service Status Last Check Duration Attempt Status Information
gpo-ig CPU utilization percent OK 03-14-2014 15:16:59 0d 1h 19m 52s 1/3 SHAREDNODECPU OK - pc1_cpu_util is 0.2
Free amount of swap percent OK 03-14-2014 15:18:29 0d 1h 18m 22s 1/3 SHAREDNODESWAPFREE OK - pc1_swap_free is 100
Interface RX utilization percent CRITICAL 03-14-2014 15:21:42 0d 0h 0m 9s 1/3 SHAREDNODEINTERFACERXUTIL CRITICAL - pc1_eth1_rx_util is 99.34 (outside range 0:75)
Interface TX utilization percent OK 03-14-2014 15:21:29 0d 1h 15m 22s 1/3 SHAREDNODEINTERFACETXUTIL OK - pc1_eth1_tx_util is 0.9095
Memory utilization percent OK 03-14-2014 15:17:29 0d 1h 19m 22s 1/3 SHAREDNODEMEMUTIL OK - pc1_mem_util is 18.51
utah-ig CPU utilization percent OK 03-14-2014 15:18:59 0d 1h 17m 52s 1/3 SHAREDNODECPU OK - pc3_cpu_util is 0.4
Free amount of swap percent OK 03-14-2014 15:20:29 0d 1h 16m 22s 1/3 SHAREDNODESWAPFREE OK - pc3_swap_free is 100
Interface RX utilization percent OK 03-14-2014 15:16:59 0d 1h 14m 52s 1/3 SHAREDNODEINTERFACERXUTIL OK - pc3_eth0_rx_util is 0.02792
Interface TX utilization percent OK 03-14-2014 15:17:59 0d 1h 18m 52s 1/3 SHAREDNODEINTERFACETXUTIL OK - pc3_eth0_tx_util is 0.03473
Memory utilization percent OK 03-14-2014 15:19:29 0d 1h 17m 22s 1/3 SHAREDNODEMEMUTIL OK - pc3_mem_util is 38.22
10 Matching Service Entries Di
iperf test caused 99% util of interface
[3] 407.0-408.0 sec 109 MBytes 914 Mbits/sec
[3] 408.0-409.0 sec 110 MBytes 920 Mbits/sec
[3] 409.0-410.0 sec 108 MBytes 910 Mbits/sec
[3] 410.0-411.0 sec 110 MBytes 922 Mbits/sec
[3] 411.0-412.0 sec 109 MBytes 911 Mbits/sec
[3] 412.0-413.0 sec 109 MBytes 915 Mbits/sec
[3] 413.0-414.0 sec 110 MBytes 921 Mbits/sec
[3] 414.0-415.0 sec 110 MBytes 919 Mbits/sec
[3] 415.0-416.0 sec 109 MBytes 916 Mbits/sec
[3] 416.0-417.0 sec 109 MBytes 914 Mbits/sec

University of Kentucky Monitoring

History



Current Status

CPU Utilization: 0.9%

Polled: Mar 13 2014 18:15:59

Memory Used: 1044537344/2435476 KB

Swap Free: 99.4%

Disk Partition Max Used: 35.6%

Ports

utah.geniracks.net_interface_pc3:eth0

Updated: Mar 12 2014 05:29:48

Slivers

No slivers associated with this Resource

Dashboard of Collector

Thanks!

Many thanks for taking this from design to deployment between GEC18 and GEC19

- ExoGENI: Jonathan Mills
- GRNOC: AJ Ragusa, Mitch McCracken
- InstaGENI: Gary Wong
- MAX: Tom Lehman and Xi Yang
- UK: Cody Bumgardner, Caylin Hickey, and Ray Hyatt
- GPO: Chaos Golubitsky, Tim Upthegrove, and Heidi Picher Dempsey

- And others who have participated on the design decision-making process

Join Us at the Coding Sprint

Wednesday at 1:30

Resources -

Head wikipage:
http://groups.geni.net/geni/wiki/OperationalMonitoring

Use cases:
http://www.gpolab.bbn.com/monitoring/use_cases.html

Code:
http://trac.gpolab.bbn.com/ops-monitoring/wiki

Monitoring email/IRC:
monitoring@geni.net
#geni-monitoring on irc.freenode.net.