Performance Measurement Platform (PMP) service

Szymon Trocha (Poznań Supercomputing and Networking Center)
WP6T3, PMP subtask

TNC19, Tallin, June 16, 2019

Public

www.geant.org

Agenda

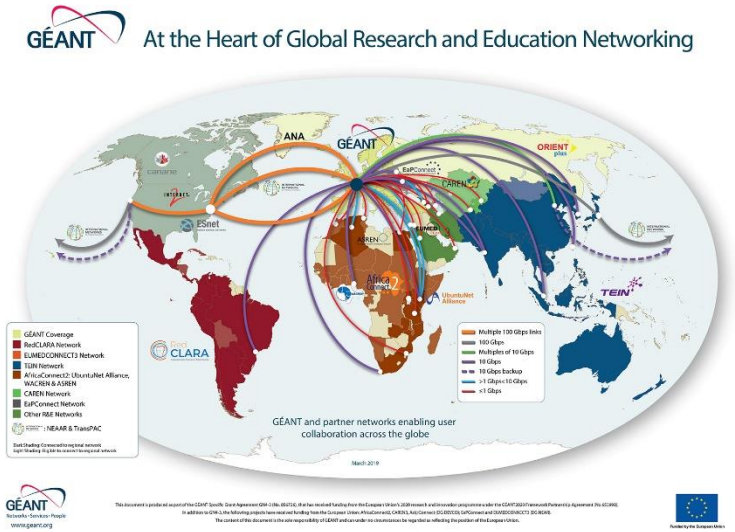
- What is PMP?
- From Small Nodes to PMP
- Architecture
- Topology
- Measurements
- Dashboard
- perfSONAR Toolkit GUI

This document partially uses results of work by the perfSONAR Project (http://www.perfsonar.net). Full set of training materials is available in https://www.perfsonar.net/resources/training-materials/#perfSONAR_Training_Materials

What is PMP?

Heterogeneous world

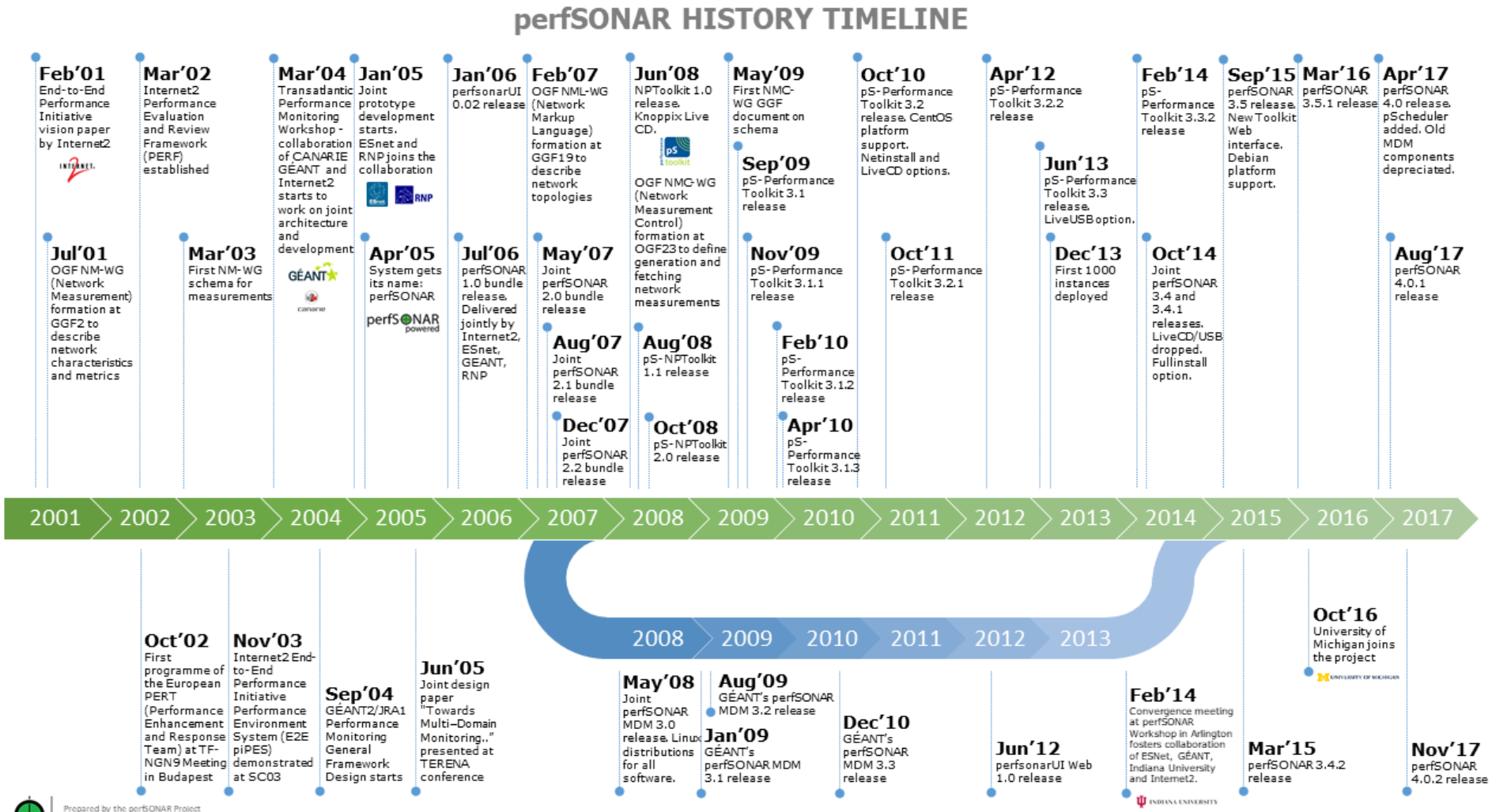
- The global Research & Education network ecosystem is comprised of hundreds of international, national, regional and local-scale networks
- While these networks all interconnect, each network is owned and operated by separate organizations (called “domains”) with different policies, customers, funding models, hardware, bandwidth and configurations
- This complex, heterogeneous set of networks must operate seamlessly from “end to end” to support **your** science and research collaborations that are distributed globally



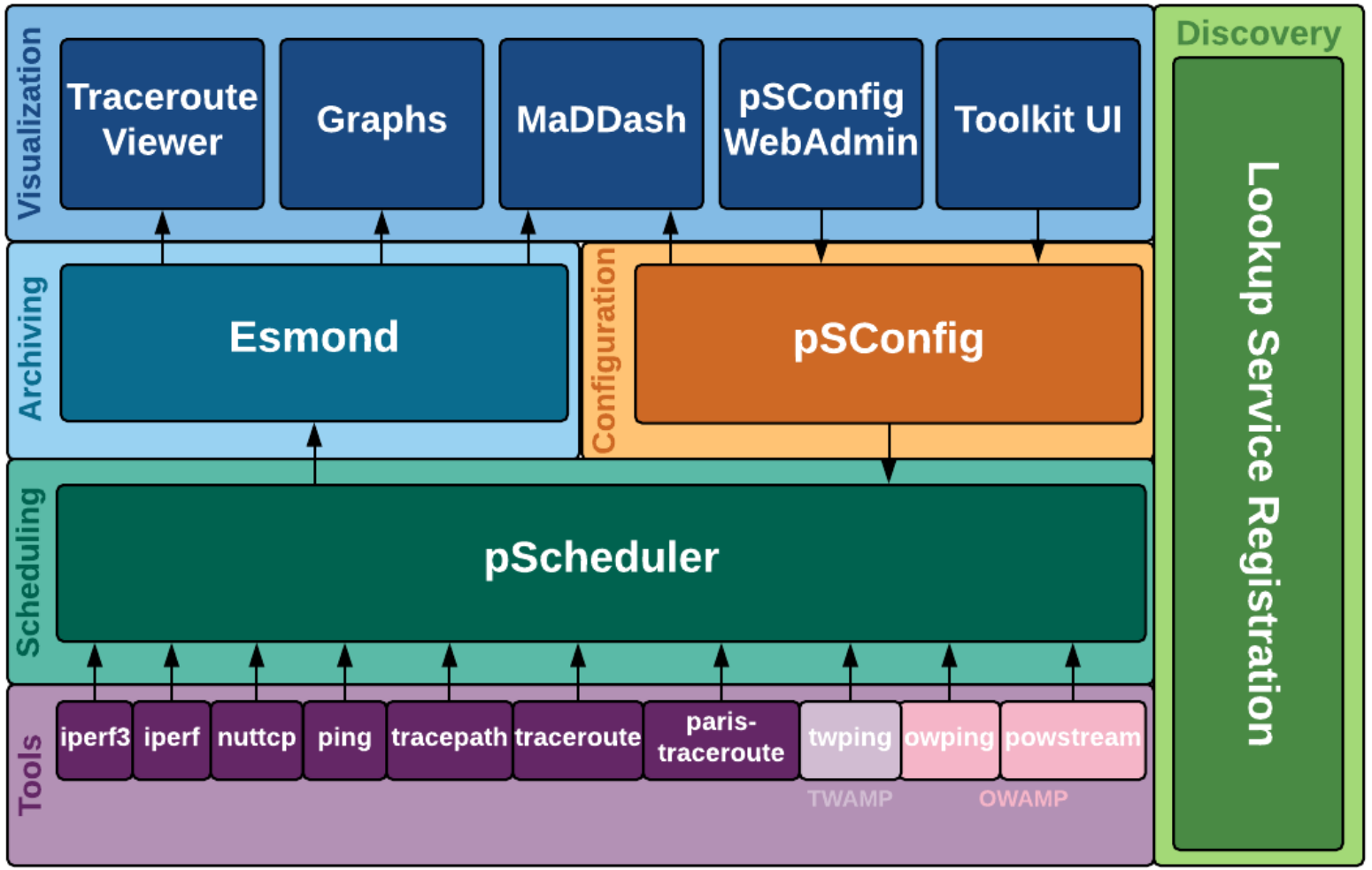
What is perfSONAR?

- It’s infeasible to perform at-scale data movement all the time – as we see in other forms of science, we need to rely on simulations
- perfSONAR is a tool to:
  - Set network performance expectations
  - Find network problems (“soft failures”)
  - Help fix these problems
  - All in multi-domain environments
- These problems are all harder when multiple networks are involved
- perfSONAR is provides a standard way to publish monitoring data
- This data is interesting to network researchers as well as network operators

Years of experience and competence



Architecture



perfSONAR Toolkit

- Network performance comes down to a couple of key metrics:
  - Throughput (e.g. “how much can I get out of the network”)
  - Latency (time it takes to get to/from a destination)
  - Packet loss/duplication/ordering (for some sampling of packets, do they all make it to the other side without serious abnormalities occurring?)
- We can get many of these from a selection of measurement tools – the perfSONAR Toolkit

- The “perfSONAR Toolkit” is an open source implementation and packaging of the perfSONAR measurement infrastructure and protocols
- All components are available as RPMs, DEBs, and bundled as a CentOS ISO
- Very easy to install and configure (usually takes less than 30 minutes for default install)

Performance Measurement Platform (1/2)

- Consists of a set of low-cost hardware nodes with preinstalled perfSONAR software
- Coupled with GÉANT MPs to create a partial mesh for NRENs
- The central components that manage the platform elements, gather, store and represent the performance data, are operated and maintained by the GÉANT project

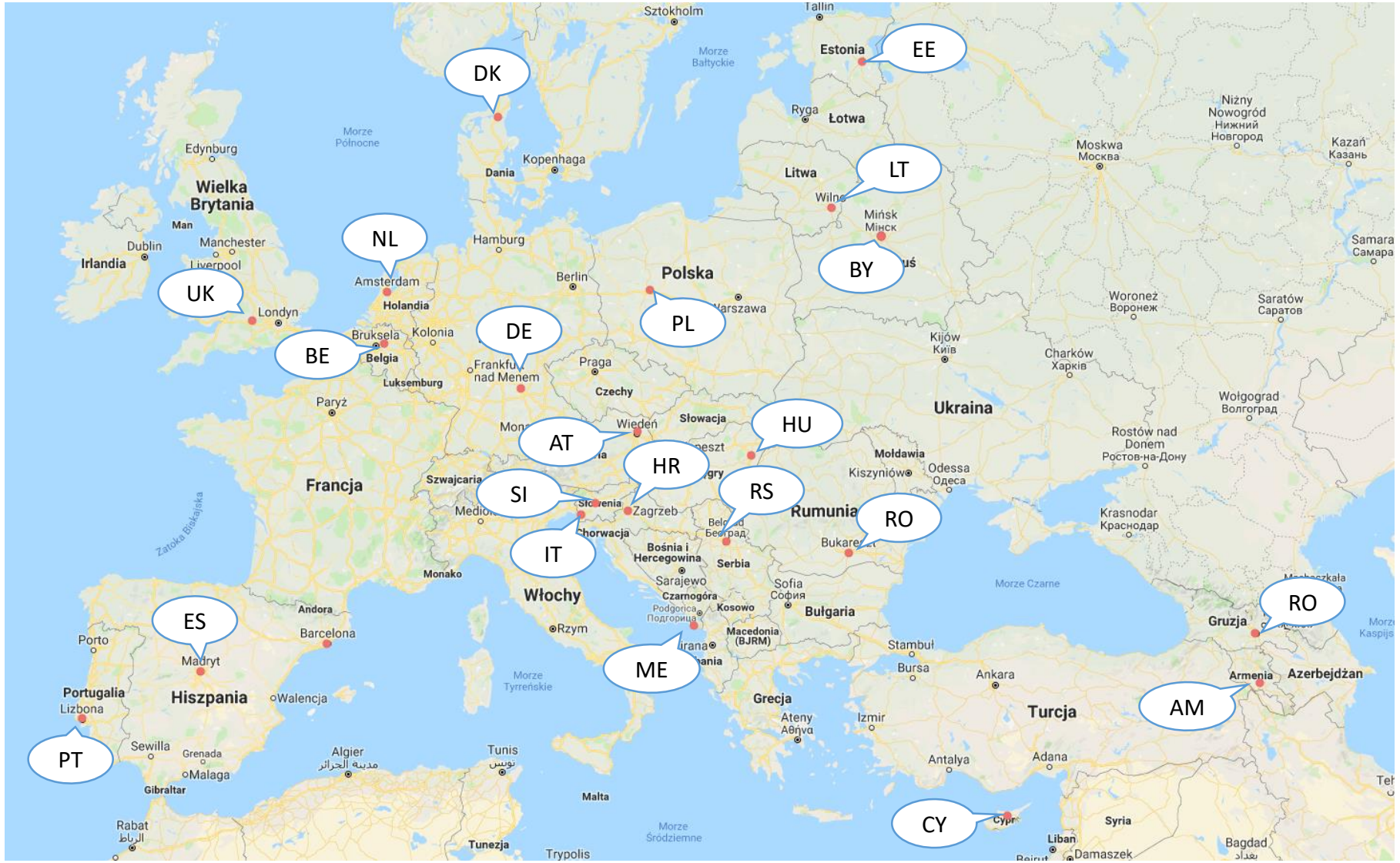
Performance Measurement Platform (2/2)

- Small nodes users can shape the predefined setup and configure additional measurements to their needs and get more familiar with the platform
  - Can become example measurement experimentation and training platform about network measurement, network management, network performance
  - Can provide an easy way to setup a new perfSONAR small nodes on new small devices through providing ways for image creation and guidelines

Current coverage

http://stats.es.net/ServicesDirectory/

- +3 in African NRENs
- Some contries have >1



From small nodes to PMP

It all began at TNC16

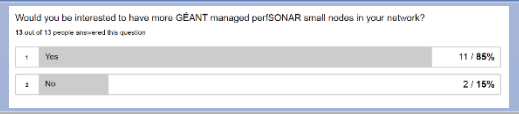
A pilot transformed into GÉANT service

pSmall – Small Nodes Project

Enable **you** to have first hand experience with perfSONAR

Inexpensive but with known capabilities

Good community feedback



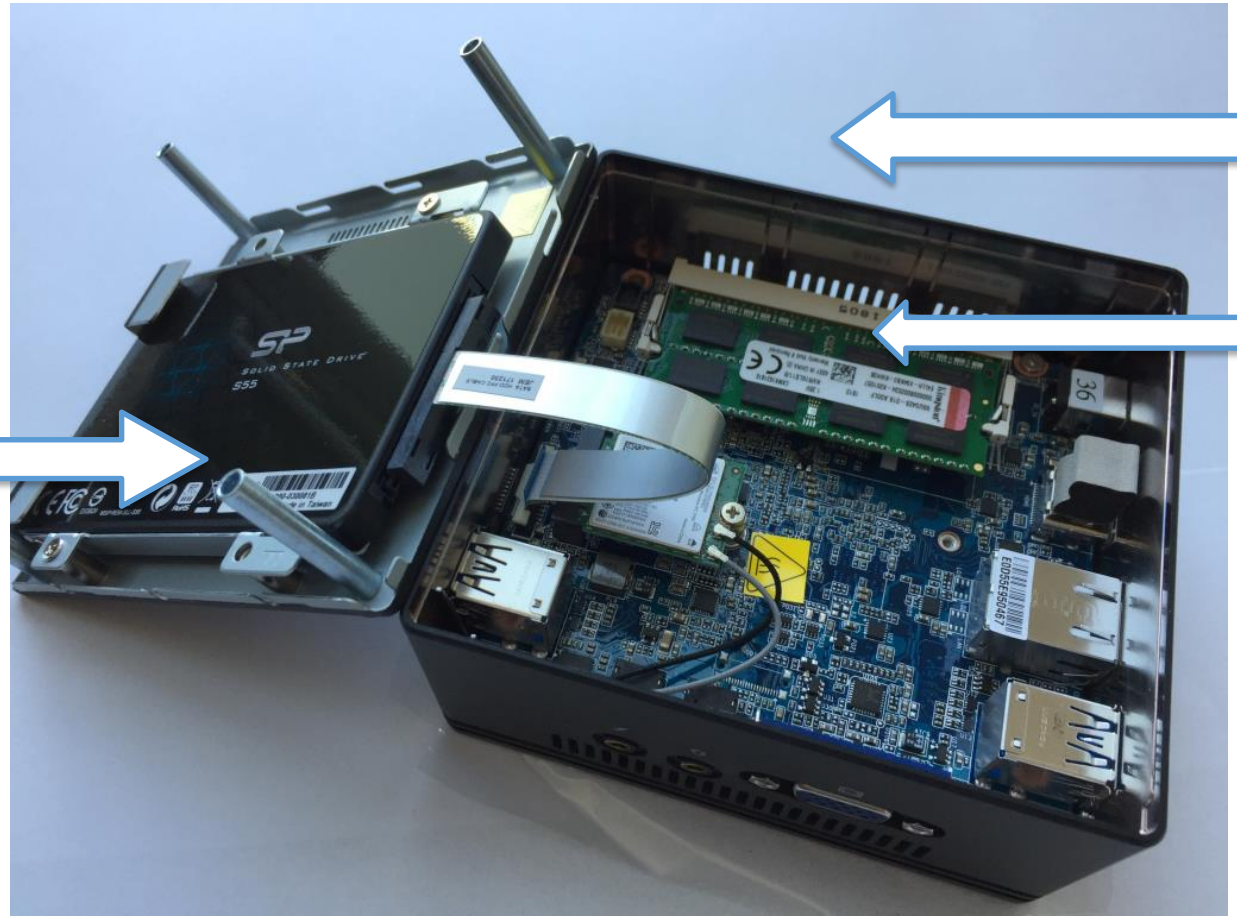
GÉANT service

WP6T3

Infrastructure and Ops team

Small technology choice

56 x 107 x 114 mm



Intel Celeron J3160
1.6GHz (up to 2.24 GHz)
4 cores

8GB RAM
Kingston

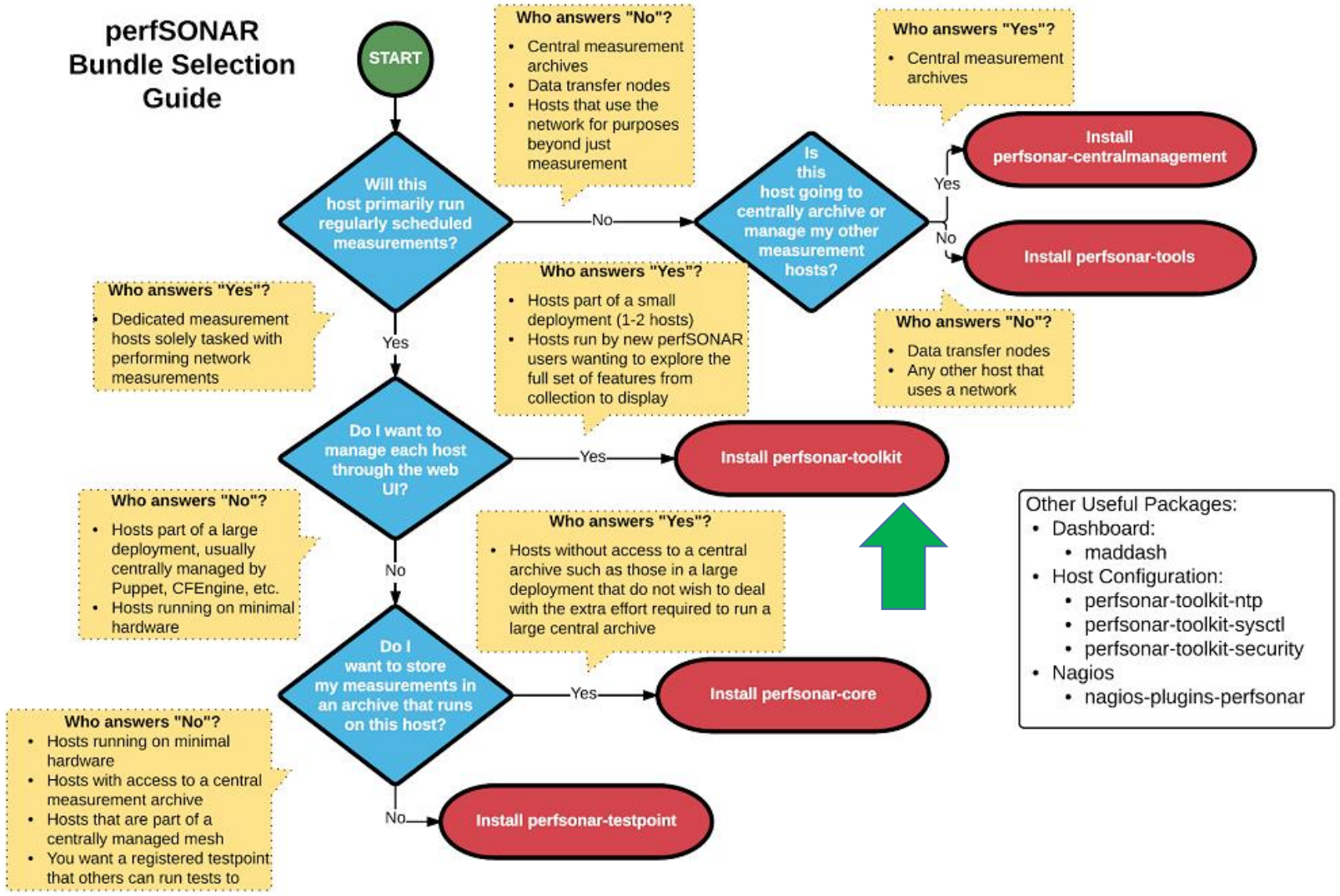
120GB
Patriot SSD

/	20 GB
/var/lib	86 GB
Swap	8 GB

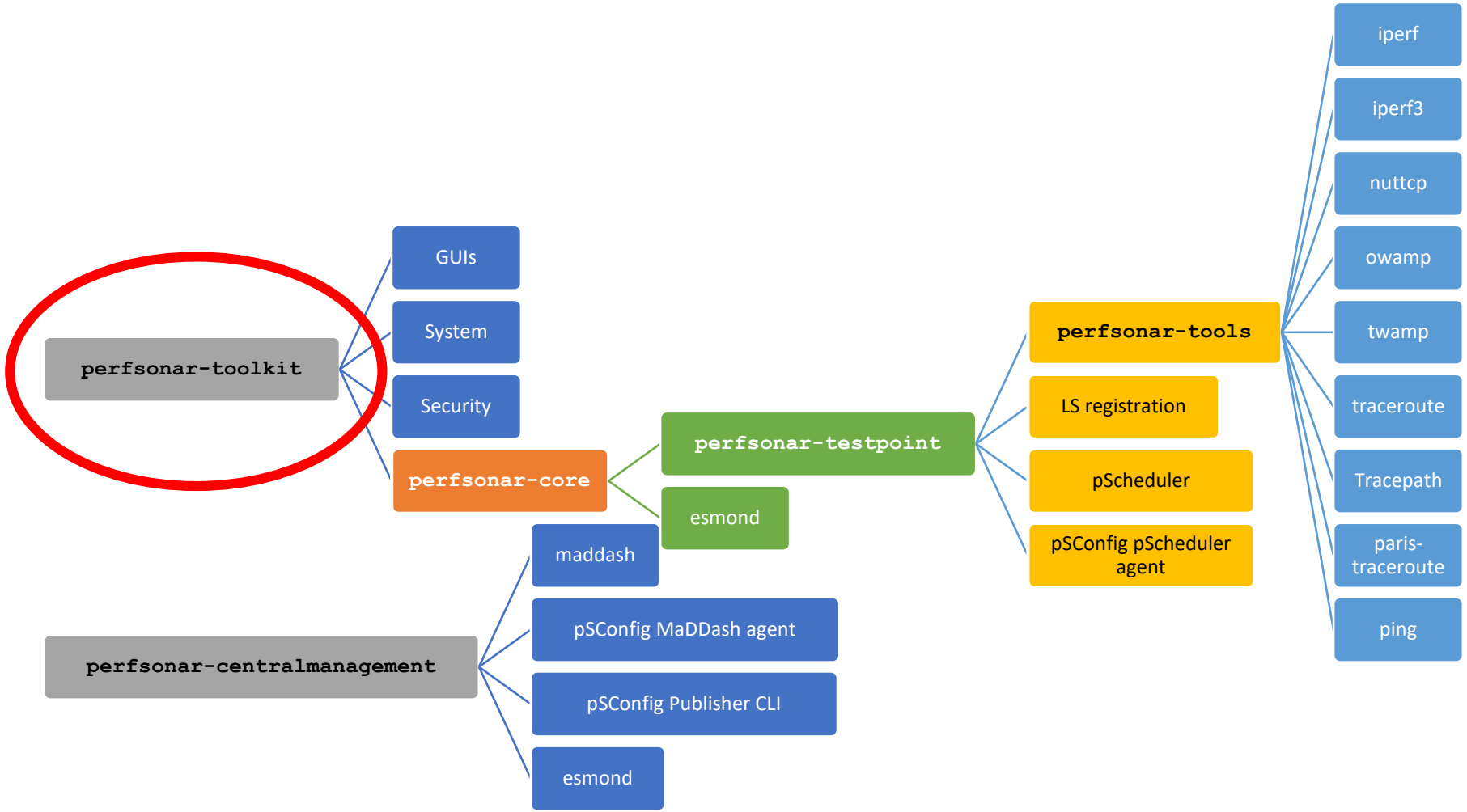
- 1Gb/s RJ45 LAN Realtek
- WiFi + BT
- VGA, HDMI
- 4 x USB 3.0
- Micro SD
- External power supply

GIGABYTE BRIX GB-BACE-3160

What did we choose?



Bundle solutions



Preconfigured distribution

- pre-install of the perfSONAR toolkit on CentOS7
  - Auto-update turned on
- 1 master node, rest cloned
  - Procedure available
- Tagged with communities : GEANT, pSmall-GEANT, PMP-GEANT
- SNMP agent

Central parts

Central server

Hosts measurement results

Hosts Ansible configuration

Measurements dashboard

Visualization of test mesh

Publicely available

Infrastructure monitoring

Grafana and Prometheus based

Central server and nodes' health

Alerting

Automation & management

Via Ansible

Mesh configuration

User configuration

Operation team

Maintains central parts

Reacts to monitoring

Supports deployment

Disseminates information

Topology and testing

Importance of regular testing

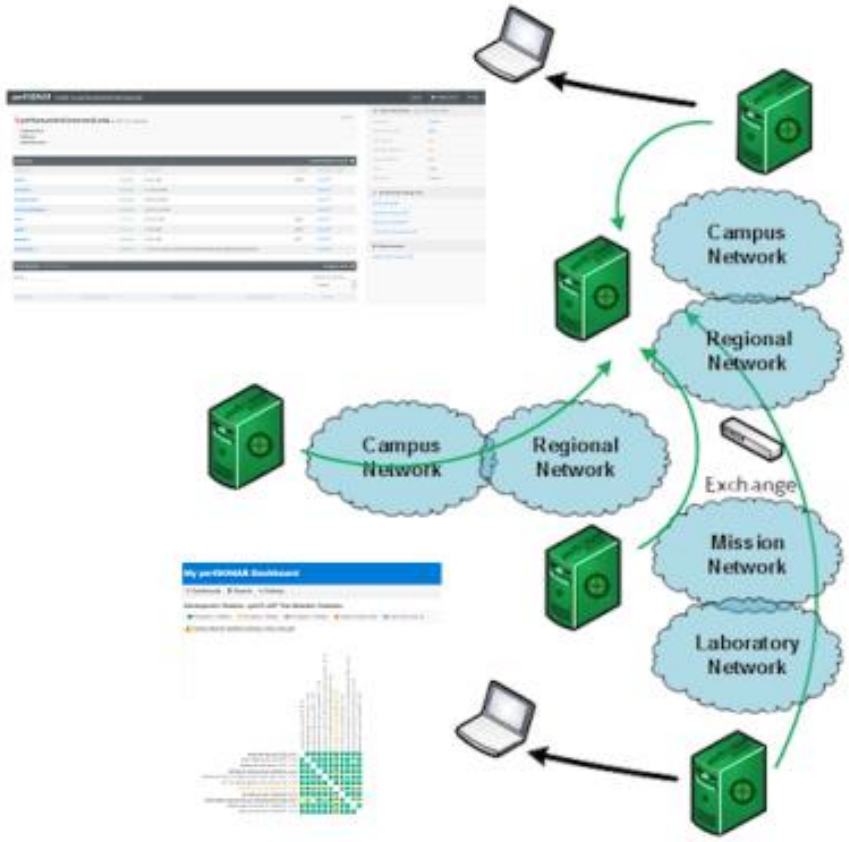
- We can’t wait for users to report problems and then fix them (soft failures can go unreported for years!)
- Things just break sometimes
  - Failing optics
  - Somebody messed around in a patch panel and kinked a fiber
  - Hardware goes bad
- Problems that get fixed have a way of coming back
  - System defaults come back after hardware/software upgrades
  - New employees may not know why the previous employee set things up a certain way and back out fixes
- Important to continually collect, archive, and alert on active test results

Regular testing

- There are a couple of ways to do this:
  - Beacon: Let others test to you (e.g. no regular configuration is needed)
  - Island: Pick some hosts to test to – you store the data locally. No coordination with others is needed
  - Mesh: full coordination between you and others (e.g. consume a testing configuration that includes tests to everyone, and incorporate into a visualization)

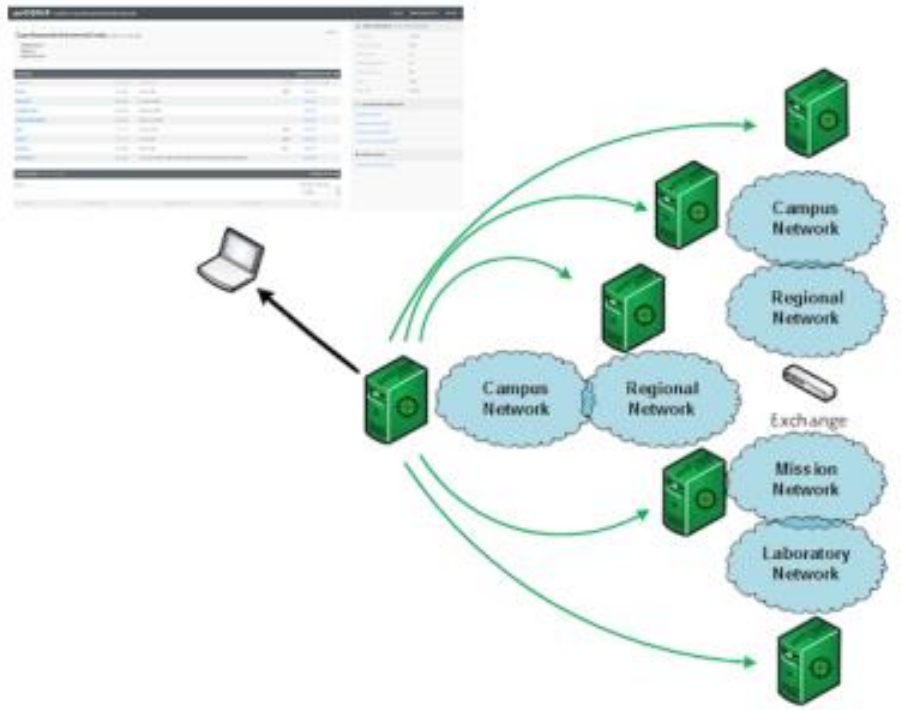
Regular testing - Beacon

- The beacon setup is typically employed by a network provider (regional, backbone, exchange point)
  - A service to the users (allows people to test into the network)
  - If no regular tests are scheduled, minimum requirements for local storage.
  - Makes the most sense to enable all services (bandwidth and latency)



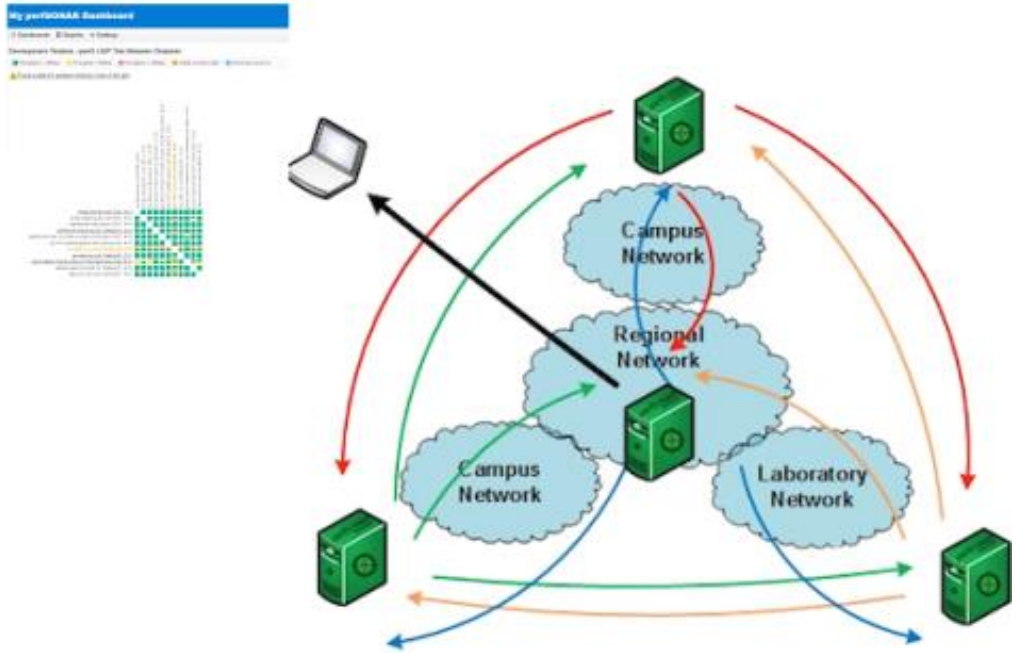
Regular testing - Island

- The island setup allows a site to test against any number of the 2000+ perfSONAR nodes around the world, and store the data locally.
  - No coordination required with other sites
  - Allows a view of near horizon testing (e.g. short latency – campus, regional) and far horizon (backbone network, remote collaborators).
  - OWAMP is particularly useful for determining packet loss in the previous cases.
  - Throughput will not be as valuable when the latency is small



Regular testing - Mesh

- A full mesh requires more coordination:
  - A full mesh means all hosts involved are running the same test configuration
  - A disjoint mesh could mean only a small number of related hosts are running a testing configuration
- In either case – bandwidth and latency will be valuable test cases



Regular Measurements

MaDDash & pSConfig

- Measurement results are more useful when they can be “seen”, because this implies they will be acted on
- MaDDash is a software package that can be used to visualize the results of many perfSONAR tests
- The pSConfig is a template framework for describing and configuring a **topology** of **tasks**
  - E.g. this is in contrast to the other method of configuration
    - the “Island” model
  - Changes node from ‘testing as an island’ to being a part of a larger testing strategy
- More info: http://docs.perfsonar.net/psconfig_intro.html

pSconfig basic concepts

*A **template** is a description of the task topology in a machine readable format*

- The pSConfig templates are formatted in JSON. The files containing this JSON data are referred to as ***pSConfig templates***.

*A **task** is a job to do consisting of a test to be carried out, scheduling information and other options.*

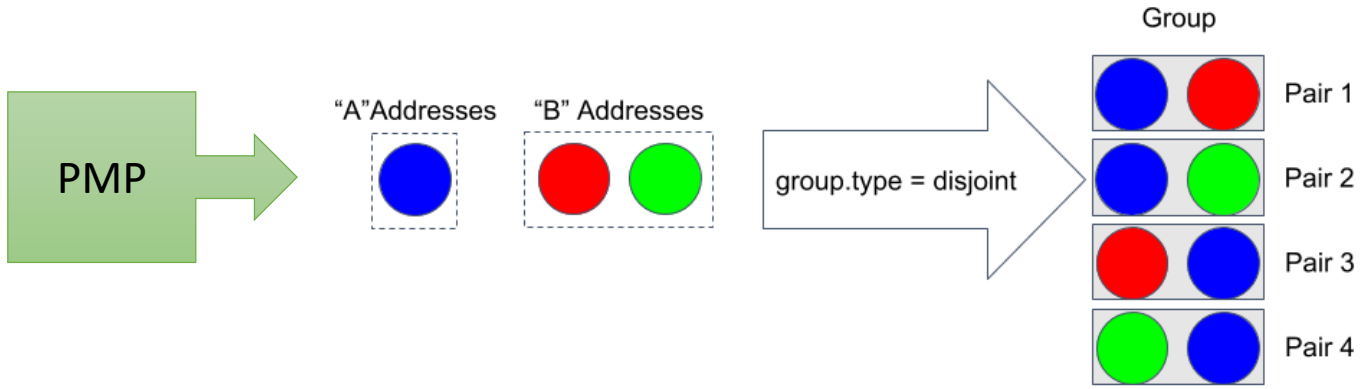
- In pScheduler a single task consists of a number of components, and these elements carry-over to pSConfig: Tests, Tools, Schedules, Archivers, Contexts

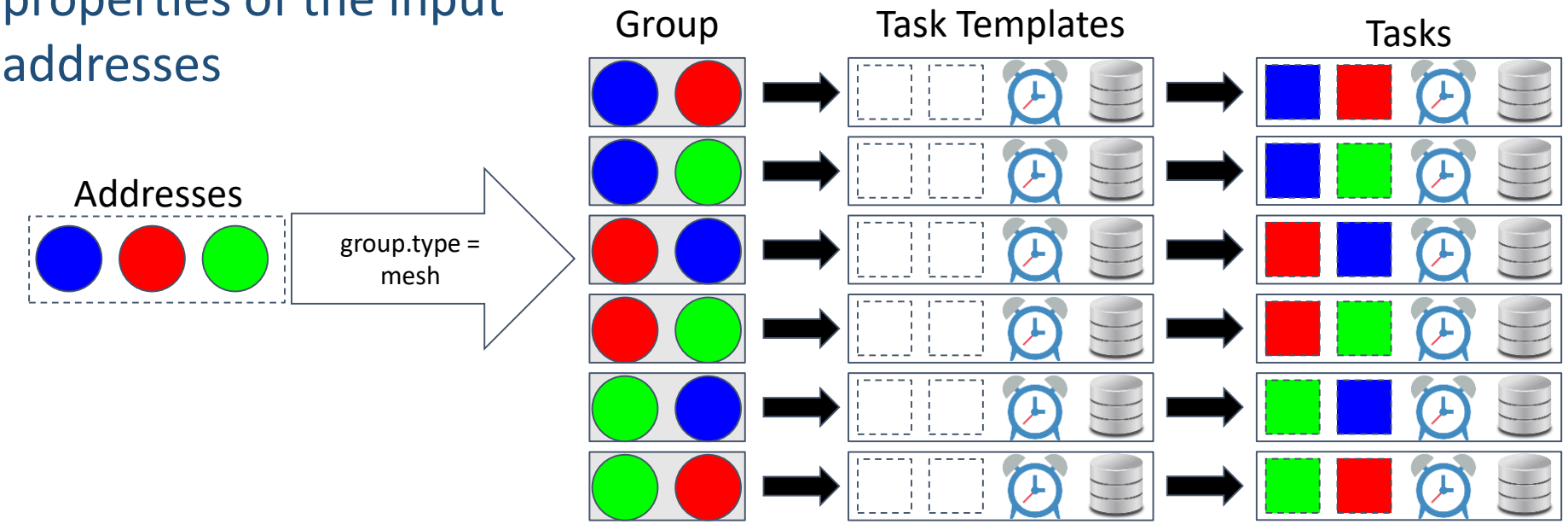
*A **topology** is the way in which tasks are interrelated and arranged*

- Ultimately we want a list of tasks to be performed
- Many of these tasks have common components. These common components often represent relationships which we care about when looking at multiple tasks

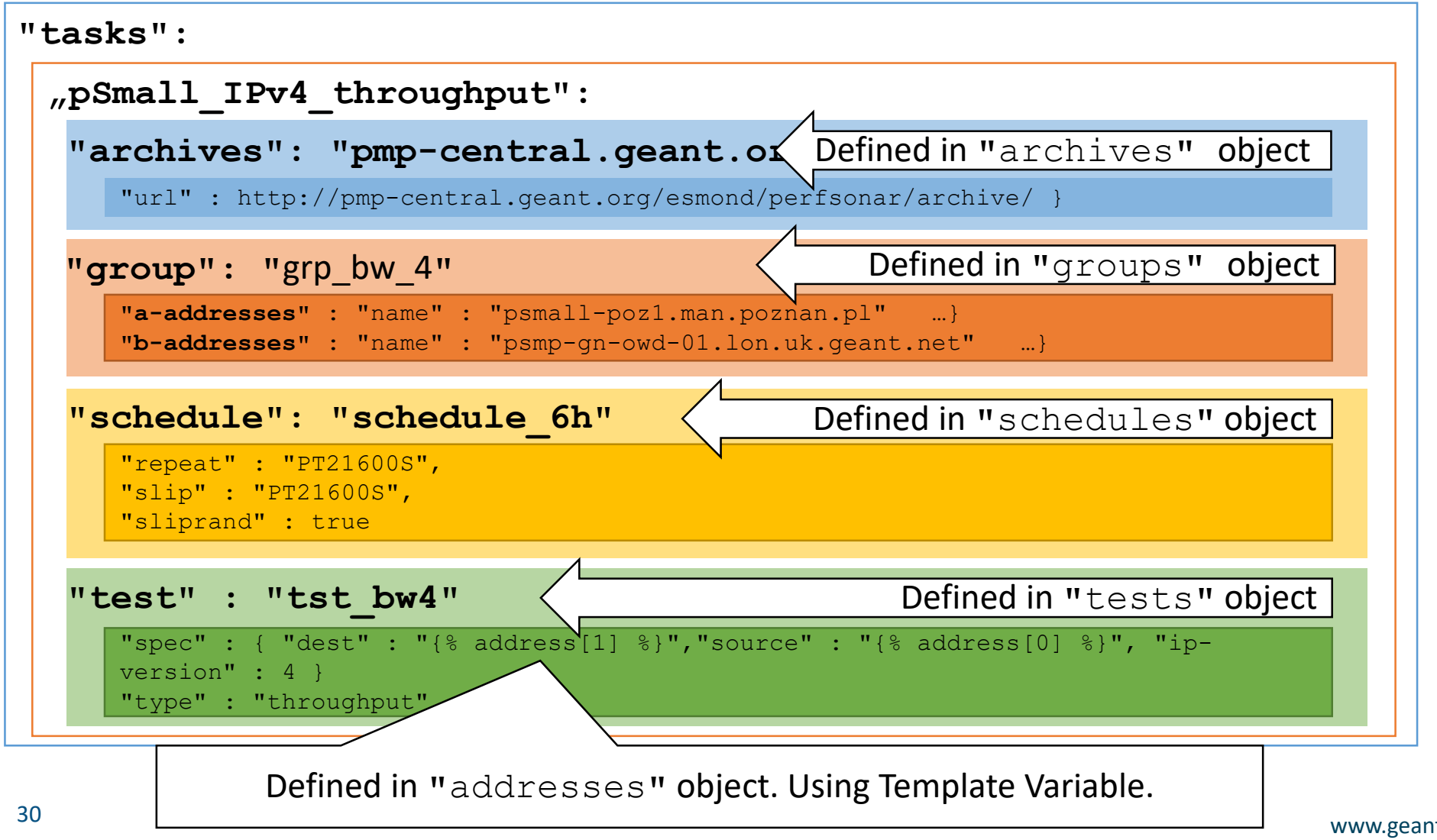
Creating tasks

For each pair in the group, we generate a task to be run using properties of the input addresses





Creating tasks



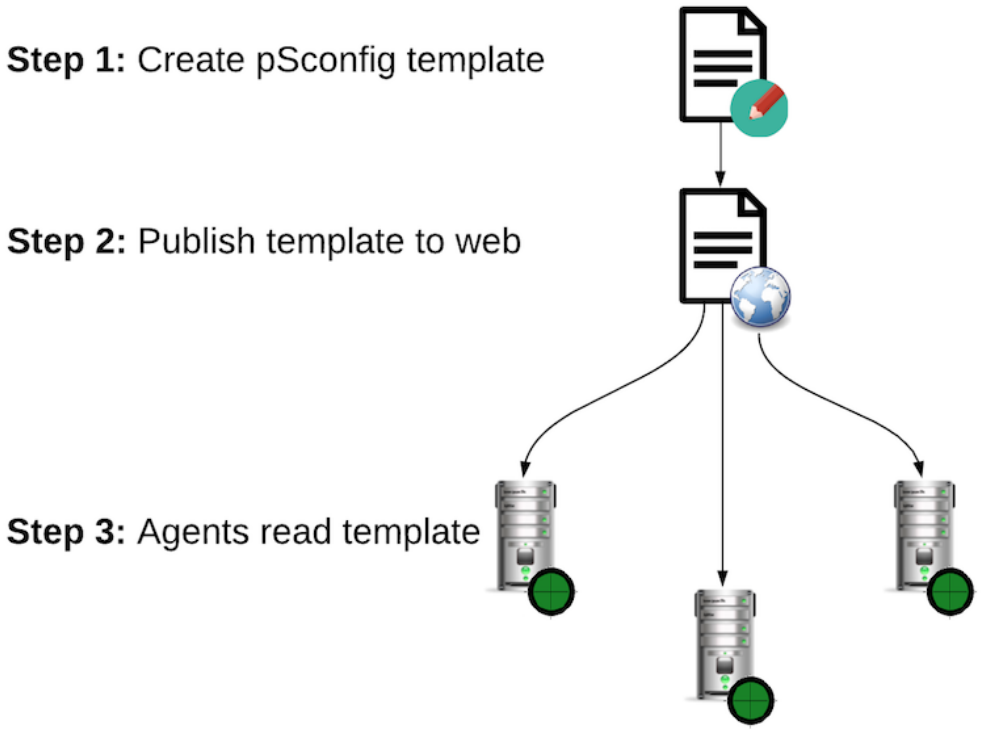
PMP Test plan

- What are we going to measure?
  - IPv4 and IPv6 (subject to availability)
  - Achievable bandwidth
    - ***5 selected GEANT MP destinations***
    - 4 times per day to each destination
    - 20 second tests across continent
  - Loss/Availability/Latency
    - OWAMP
  - Traceroute
- What are you going to do with the results?
  - MaDDash

Agents

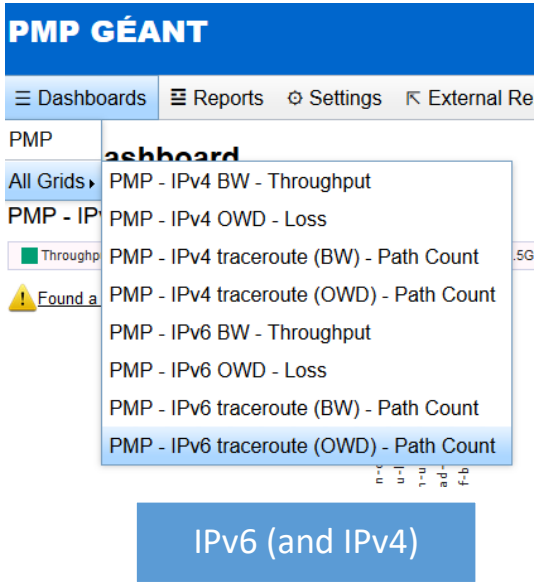
*An **agent** is software that reads one or more pSConfig templates and uses the information to perform a specific function.*

- We currently have two agents:
  - **psconfig-pscheduler-agent**: It reads the template file(s) and generates pScheduler tasks
  - **psconfig-maddash-agent:** It reads the the template file(s) and generates a maddash.yaml file

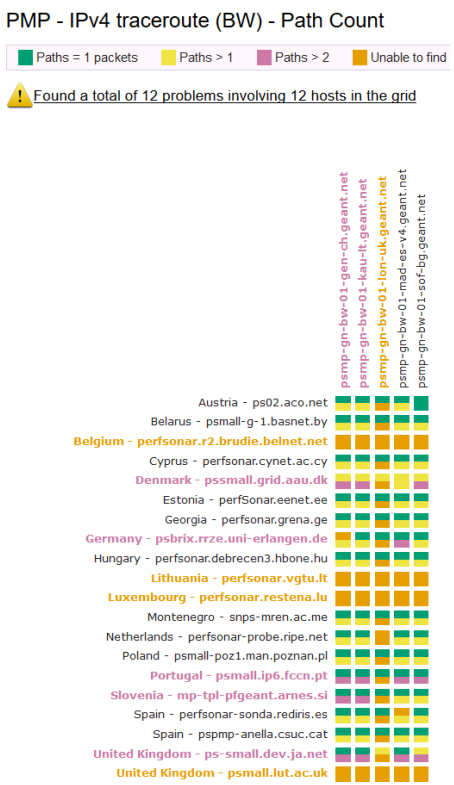


Measurements

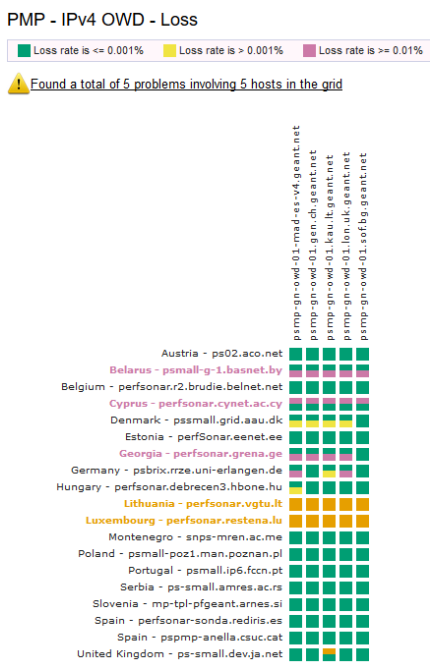
Throughput



IPv6 (and IPv4)



Traceroute / hop count



Latency / packet loss

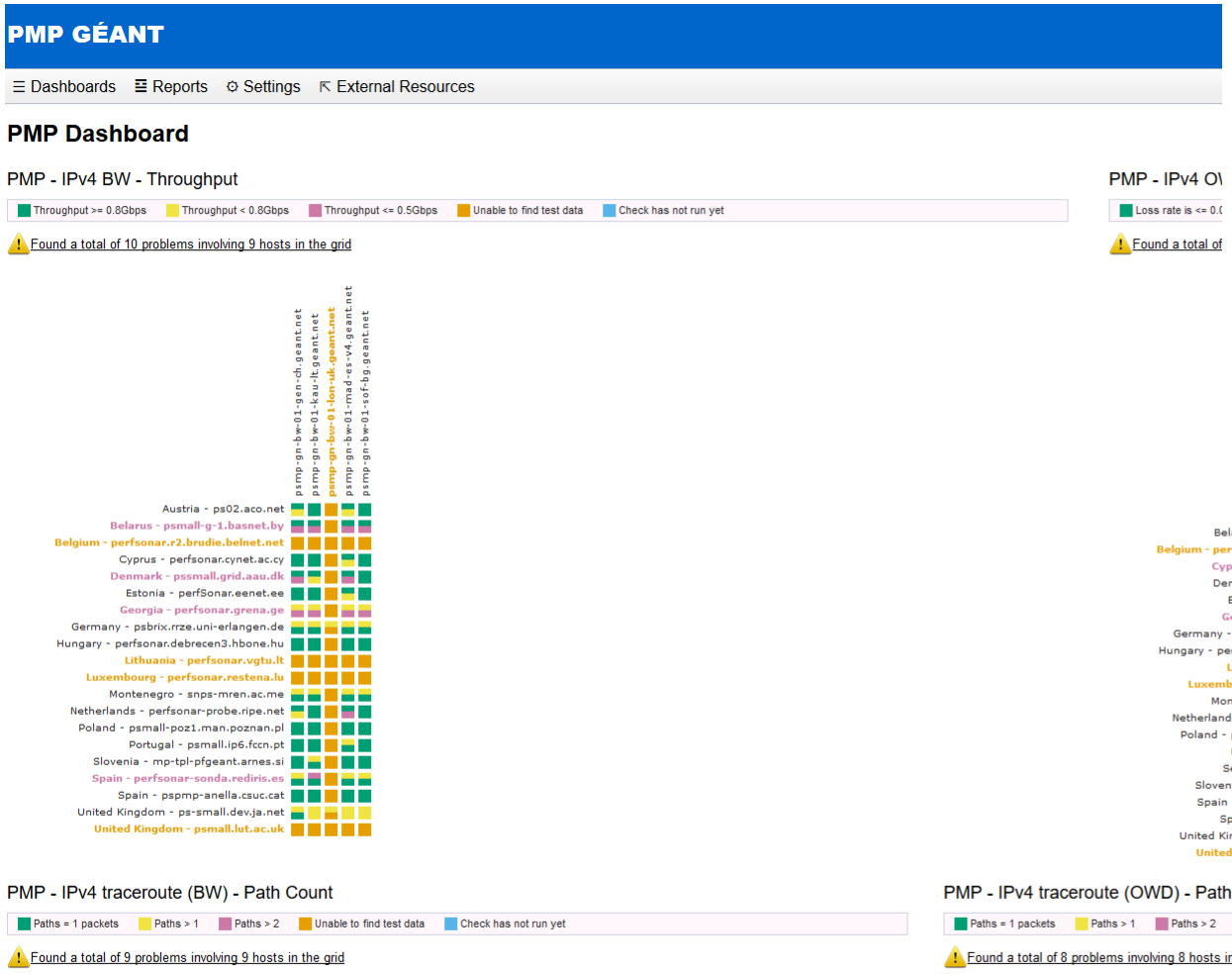
PMP disjoint mesh walkthrough

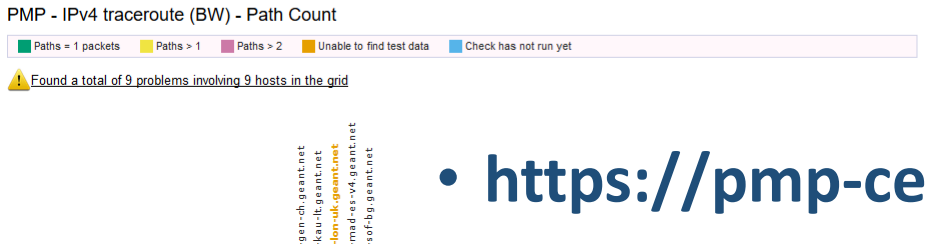
- https://pmp-central.geant.org/pscfg-psmall.json

- Have complex mesh?
  - Use psConfig Web Admin
    - Relies on Lookup Service

Dashboard

PMP MaDDash





- **https://pmp-central.geant.org/maddash-webui/**

What is MaDDash?

- **M**onitoring **a**nd **D**ebugging **Dash**board
- Set of grids (dashboard)
  - Two-dimensional monitoring data
  - Point to point network measurements between 2 hosts
- Set of jobs that report on metrics
  - Usually Nagios checks
  - Lots of configurable options (but defaults are usually fine)
  - Getting data out of perfSONAR backend (esmond)
- REST API
  - Makes it possible to integrate with other platforms

Dashboard and visualization

- Integrated with psConfig templates
  - You define your measurement mesh
  - Then MaDDash grids come up automatically
- Each square represent current status
  - Gives access to graphs of measurement history
  - Split in 2:
    - Upper: from line host to column host
    - Lower: from column host to line host (reverse)
- Link gives access to 2 different pages with both direct and reverse
  - Same graphs as on toolkit

PMP MaDDash customization

- GUI
  - Dashboards and grid names, host labels
  - Additional menu items
- Thresholds
  - Fully customisable: values and descriptions
  - You define your performance expectations
- Alerts and notifications
  - Nagios
  - Email (native or through Nagios)

Measurement archives

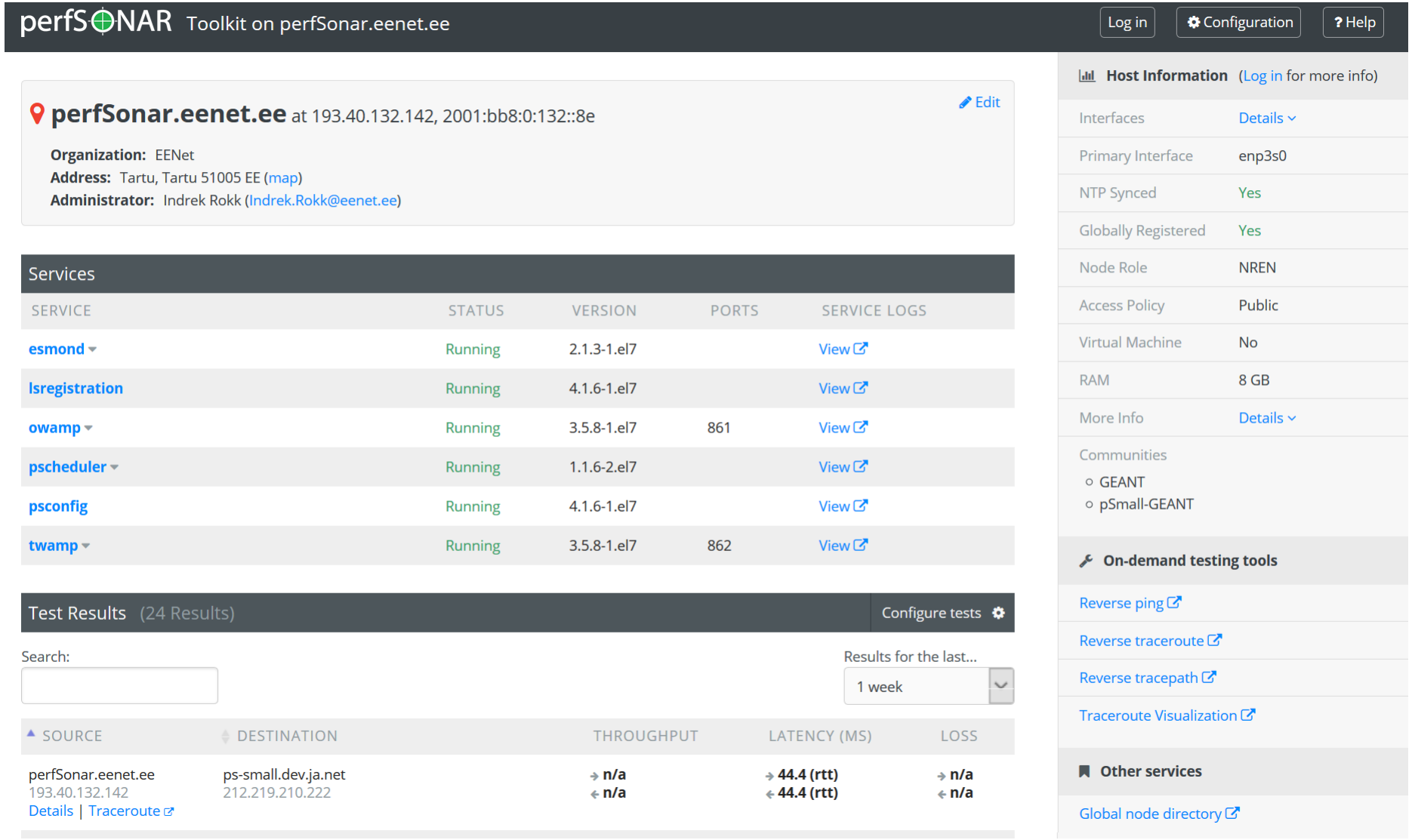
- Centralised big archive
  - All nodes post their data in a single place
  - Pros: one central storage, one storage policy
  - Cons: access management
- Distributed archiving on each node
  - Each node keeps data localy
  - Pros: reduce a bit network traffic
  - Cons: If node offline, data offline too
- Multiple archives possible

MA access management

- IP based ACL
  - Easy, though a bit less secure…
- API key
  - Finer grained controled
  - But need to distribute key
  - Used locally on toolkit nodes
- Mix and match at will

Toolkit GUI

Toolkit GUI


perfSONAR Toolkit on perfSonar.eenet.ee
Log in
Configuration
Help
perfSonar.eenet.ee at 193.40.132.142, 2001:bb8:0:132::8e
Edit
Organization: EENet
Address: Tartu, Tartu 51005 EE (map)
Administrator: Indrek Rokk (Indrek.Rokk@eenet.ee)
Services
SERVICE STATUS VERSION PORTS SERVICE LOGS
esmond Running 2.1.3-1.el7 View
lsregistration Running 4.1.6-1.el7 View
owamp Running 3.5.8-1.el7 861 View
pscheduler Running 1.1.6-2.el7 View
psconfig Running 4.1.6-1.el7 View
twamp Running 3.5.8-1.el7 862 View
Test Results (24 Results)
Configure tests
Search:
Results for the last...
1 week
SOURCE DESTINATION THROUGHPUT LATENCY (MS) LOSS
perfSonar.eenet.ee
193.40.132.142
Details | Traceroute
ps-small.dev.ja.net
212.219.210.222
n/a
n/a
44.4 (rtt)
44.4 (rtt)
n/a
n/a
Host Information (Log in for more info)
Interfaces Details
Primary Interface enp3s0
NTP Synced Yes
Globally Registered Yes
Node Role NREN
Access Policy Public
Virtual Machine No
RAM 8 GB
More Info Details
Communities
GEANT
pSmall-GEANT
On-demand testing tools
Reverse ping
Reverse traceroute
Reverse tracepath
Traceroute Visualization
Other services
Global node directory

Participation and resources

- Want to participate?
  - If you are an NREN or institution connected to an NREN
  - If you belong to the R & E community
    - Being part of a networking team, service or research group
- Identify responsible person and sign in

- www.perfsonar.net
- pmp-central.geant.org/maddash-webui/

Thank you

Any questions?

szymon.trocha@psnc.pl

www.geant.org

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 856726 (GN4-3).

Performance Measurement Platform (PMP) service

Szymon Trocha (Poznań Supercomputing and Networking Center)
WP6T3, PMP subtask

TNC19, Tallin, June 16, 2019

Public

www.geant.org

The scientific/academic work is financed from financial resources for science in the years 2019 - 2022 granted for the realization of the international project co-financed by Polish Ministry of Science and Higher Education.