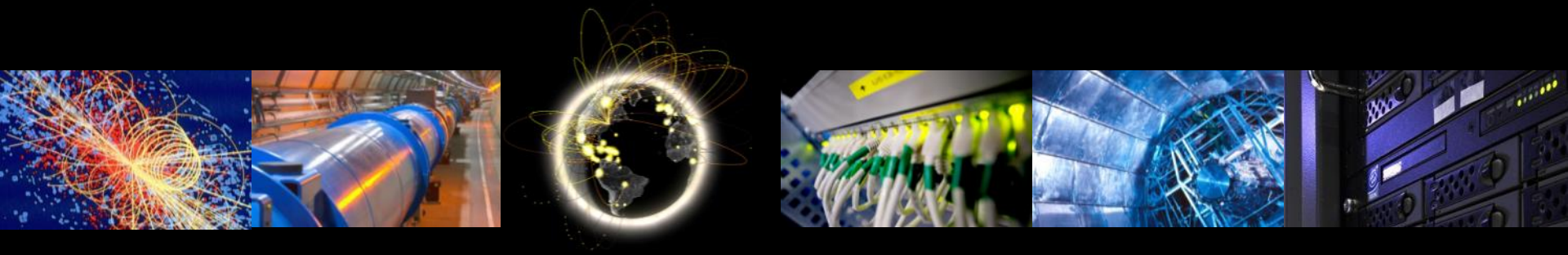
Integrating Network and Transfer Metrics to Optimize Transfer Efficiency and Experiment Workflows

Shawn McKee, Marian Babik for the
WLCG Network and Transfer Metrics Working Group

CHEP 2015, Okinawa, Japan
13th April 2015



WLCG
Worldwide LHC Computing Grid

Network Integration Motivations

- A crucial contributor to the success of the massively scaled global computing system that delivers the analysis needs of the LHC experiments is the **networking infrastructure** upon which the system is built.
  - The LHC experiments have been able to exploit excellent high-bandwidth networking in adapting their computing models for the most efficient utilization of resources.
  - However there are still challenges and opportunities to make this even more effective for our needs.

Evolving Network Capabilities

- We already have deployed End-to-end monitoring of our networks using perfSONAR.
  - perfSONAR, combined with data flow performance metrics, further allows our applications to adapt based on real time conditions.
- New, advanced networking technologies are slowly becoming available as production equipment is upgraded and replaced
  - Software Defined Networking(**SDN**) holds the potential to further leverage the network to optimize workflows and dataflows
- We foresee eventual proactive control of the network fabric on the part of high level applications such as experiment workload management and data management systems. We must prepare for this…

Network Monitoring in WLCG/OSG

- Goals:
  - Find and isolate “network” problems; alerting in time
  - Characterize network use (base-lining)
  - Provide a source of network metrics for higher level services
- Choice of a standard open source tool: perfSONAR
  - Benefiting from the R&E community consensus
- Tasks achieved by the perfSONAR TF:
  - Get monitoring in place to create a baseline of the current situation between sites
  - Continuous measurements to track the network, alerting on problems as they develop
  - Develop test coverage and make it possible to run “on-demand” tests to quickly isolate problems and identify problematic links

Network and Transfer Metrics WG

- Started in May 2014, bringing together network and transfer experts
- Follows up on the perfSONAR TF goals
- Mandate
  - Ensure all relevant **network** and **transfer metrics** are identified, collected and published
  - Ensure sites and experiments can better understand and fix networking issues
  - Enable use of network-aware tools to improve transfer efficiency and optimize experiment workflows
- Membership
  - WLCG perSONAR support unit (regional experts), WLCG experiments, FTS, Panda, PhEDEx, FAX, Network experts (ESNet, LHCOPN, LHCONE)

Network and Transfer Metrics WG

- Objectives
  - Coordinate commissioning and maintenance of WLCG network monitoring
    - Finalize, harden and maintain the perfSONAR deployment
    - Ensure all links continue to be monitored and sites stay correctly configured
    - Verify coverage and optimize test parameters
  - Identify and continuously make available relevant transfer and network metrics
  - Document metrics and their use
  - Facilitate their integration in the middleware and/or experiment tool chain
- Since inception, main focus was to finalize deployment and commissioning, extend the infrastructure, but also to jump start common projects with network and transfer metrics

perfSONAR Deployment

http://grid-monitoring.cern.ch/perfsonar_report.txt for stats

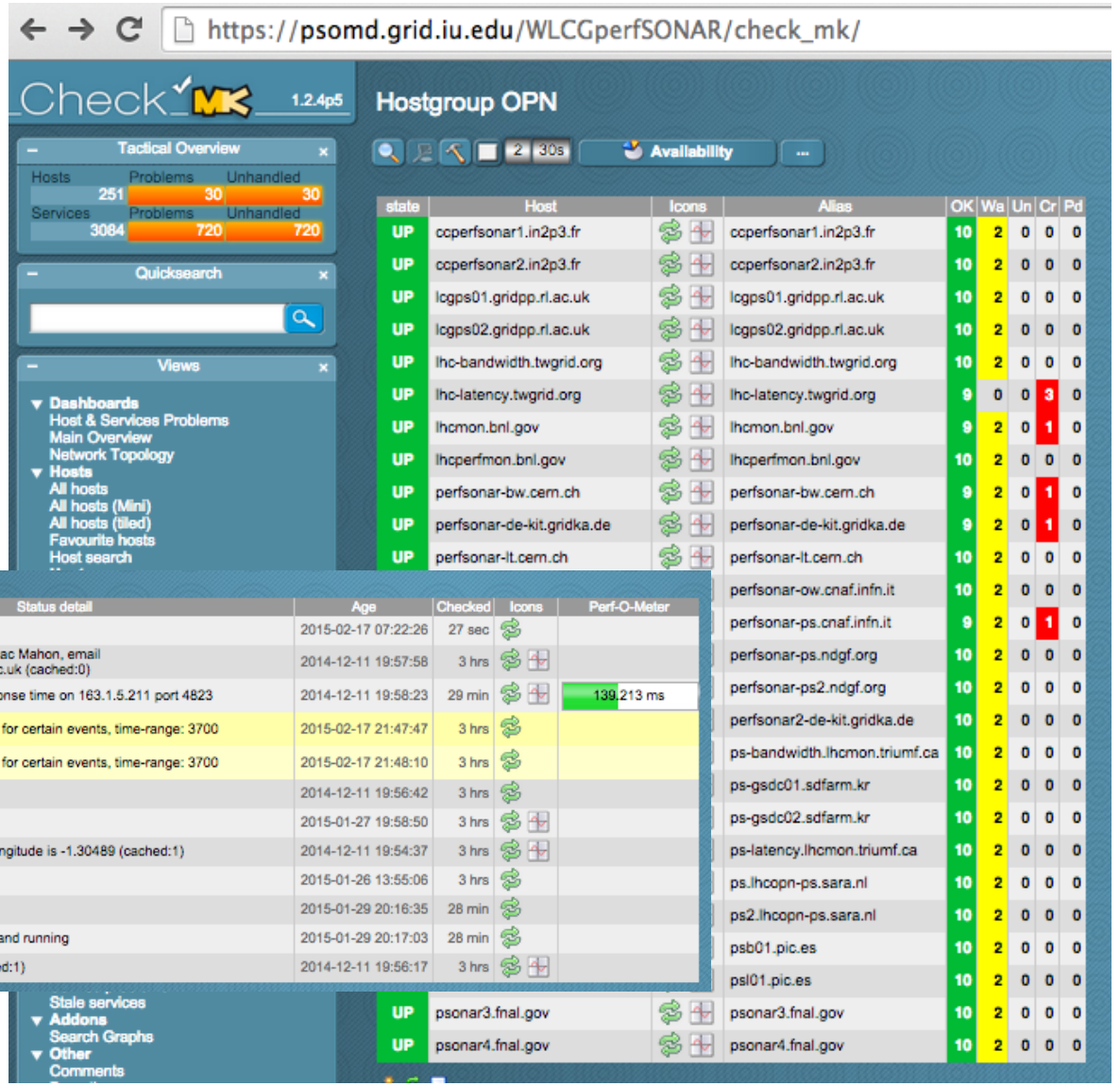
- Initial deployment coordinated by WLCG perfSONAR TF
- Commissioning of the network followed by WLCG Network and Transfer Metrics WG

Infrastructure Monitoring

- Based on OMD/check_mk and MadDash
- MadDash developed as part of perfSONAR – 1.2 version released recently
- OMD/check_mk extended to cover WLCG perfSONAR needs
- Developed bootstrapping and auto-configuration scripts
  - Synchronized with GOCDB/OIM and OSG configuration interface
- Packaged and deployed in OSG
- Developed new plugins - core functionality
  - Toolkit Version, Regular Testing, NTP, Mesh configuration, Esmond (MA), Homepage, Contacts
  - Updated to perfSONAR 3.4 information API/JSON
- High level functionality plugins
  - Esmond freshness – checks if perfSONAR node's local MA contains measurements it was configured to perform
  - Extremely useful during commissioning

Infrastructure Monitoring

- Auto-summaries are available per mesh
- Service summaries per metric type
- **GOAL:** To ensure we continue to reliably obtain ALL network metrics



OSG perfSONAR Datastore

- All perfSONAR metrics should be collected into the OSG network datastore
  - This is an Esmond datastore from perfSONAR (postgresql+cassandra backends)
  - Loaded via RSV probes; currently one probe per perfSONAR instance every 15 minutes.
- Validation and testing ongoing in OSG
  - Plan is to have it production ready by Q3 2015
- Datastore on psds.grid.iu.edu
  - JSON at http://psds.grid.iu.edu/esmond/perfsonar/archive/?format=json
  - Python API at http://software.es.net/esmond/perfsonar_client.html
  - Perl API at https://code.google.com/p/perfsonar-ps/wiki/MeasurementArchivePerlAPI

Integration Projects

- Goal
  - Provide platform to integrate network and transfer metrics
  - Enable network-aware tools (see ANSE http://cern.ch/go/M9Sj)
    - Network resource allocation along CPU and storage
    - Bandwidth reservation
    - Create custom topology
- Plan
  - Provide latency and trace routes and test how they can be integrated with throughput from transfer systems
  - Provide mapping between sites/storages and sonars
  - Uniform access to the network monitoring
- Pilot projects
  - FTS performance – adding latency and routing to the optimizer
  - Experiment's interface to datastore

Experiments Interface to Datastore

- Aim
  - Develop publish/subscribe interface to perfSONAR
  - Enable possibility to subscribe to different events (filter) and support different clients - integration via messaging, streaming data via topic/queue
  - Provide mapping/translation btw sonar infrastructure and experiment's topology
- Components
- esmond2mq – prototype already exists
  - Retrieves all data (meta+raw) from esmond depending on existing mesh configs
  - Publishes to a topic
- Proximity/topology service
  - Handle mapping/translation of services (service to service; storage to sonar), service to site (sonar to site)
  - Test different algorithms (site mapping, traceroutes, geoip)
  - Evaluate if existing tools can be reused for this purpose

Closing remarks

- perfSONAR widely deployed and already showing benefits in troubleshooting network issues
  - Additional deployments within R&E/overlay networks still needed
- Significant progress in **configuration** and **infrastructure monitoring**
  - Helping to reach full potential of the perfSONAR deployment
- **OSG datastore** – community network data store for all perfSONAR metrics – planned to enter production in Q3
- Integration projects aiming to aggregate network and transfer metrics
  - FTS Performance
  - Experiment's interface to perfSONAR
- ANSE project is providing "hooks" in PANDA and PhEDEx to utilize future SDN capabilities as they become available at our sites and in our networks.




Questions?