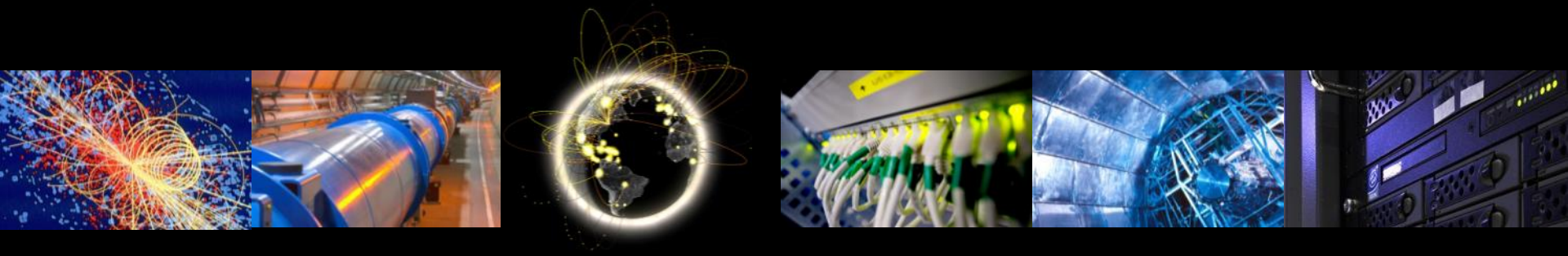
Network and Transfer Metrics WG Meeting

Shawn McKee, Marian Babik

Network and Transfer Metrics WG Meeting
18h March 2015



WLCG
Worldwide LHC Computing Grid

Outline

- Next Meetings/Events
  - 8 Apr (CHEP), 6 May, 3 June, 8 July, 2 Sept - all at 4pm CEST
  - HEPiX and CHEP
- Input to use cases document still missing !

Today

- perfSONAR status
  - Mesh configuration changes
  - Operations status
- esmond status/plans
- Network performance incidents/degradations
- Integration projects
  - Experiment's interface to perfSONAR

perfSONAR Status

Mesh Configuration changes

- Disabled all inter-cloud meshes and full mesh bandwidth
- Introduced full mesh latency mesh – with LHCOPN hosts for now
- Regional meshes unchanged
- Introduced Latin America mesh – lead by Renato and Pedro
- Introduced Dual-stack mesh – lead by Duncan
  - IPv6 tests for nodes that are dual stack
  - Plan is to progressively add nodes that support IPv6
  - Test frequency can be lowered if necessary
- Configuration interface at https://oim.grid.iu.edu/oim/meshconfig
  - You need to be registered in OIM and authorized to access
  - Please contact me or Shawn if you have issues

perfSONAR operations

- For LHCOPN/LHCONE started to investigate if sonars are consistently delivering metrics
  - Added new checks to psomd.grid.iu.edu to query freshness of local MAs and identify missing links
  - Non-critical tests that check high level functionality
- Results used to calculate overall completeness
- Detect links that are not tested
  - Plotting missing links to see the patterns - http://grid-monitoring.cern.ch/perfsonar-debug/
- Several issues were found thanks to this
  - Dual-stack sonars were not tested in any way if connection to IPv6 failed – resolved by forcing all tests to be ipv4 only
  - Latency nodes gradually removing tests until almost nothing is tested – regular testing almost never recovers (requires reboot)
  - Regular testing not running on some nodes
  - Bandwidth nodes in better shape, but prone to misconfigurations (NDGF)
  - Current status information from sonars not sufficient to understand if node is healthy – limited information on status of regular testing and esmond

perfSONAR operations

- But some of the issues still not understood
  - Very difficult to debug since everything has to be done remotely in LHCOPN/LHCONE
  - Attempted to get admin access to OPN perfSONARs - problematic for some sites due to policies
- Changed focus to 3.4.2rc – established perfSONAR testbed
  - Controlled environment
  - Participants: AGLT2, BNL, UNL, IU, MWT2
    - https://maddash.aglt2.org/maddash-webui/index.cgi?dashboard=perfSONAR%20Testbed
  - Focus on validation and testing of release candidates (3.4.2rc)
  - Following up via infrastructure monitoring - includes base sonar checks, freshness checks, additional os/host level checks
  - Confirmed issue with OWAMP nodes and regular testing not running
  - Issues found were fixed, validation still ongoing

esmond Status/Plans

esmond Status/Plans

- Regular meetings held in OSG - Shawn coordinates the activity
- Planning to make it production ready by Q3
- Performance was improved on VMs: configured RSV data-gathering probes to distribute starts
  - Original conflicted with other I/O intensive VMs
  - Gathering 100% of meshes (But many w/o data ☹)
- Completeness and accuracy
  - Plan is to re-use the freshness tests to check accuracy and completeness of the data

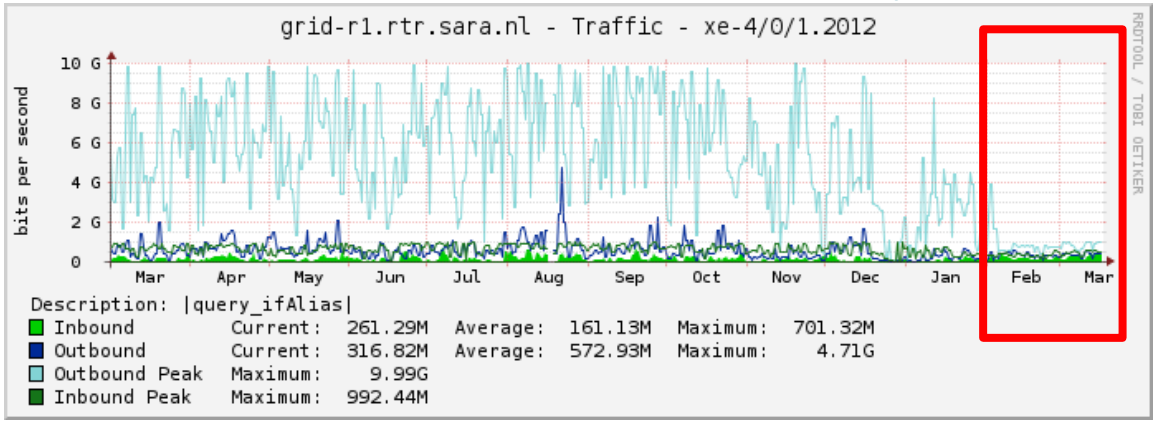
Network performance incidents follow up

Network Incidents Follow up

- Incidents not frequent but very difficult to resolve
- Complex tracking is needed since many parties are involved
- Difficult to identify the real problem
- Missing process to track and follow up on incidents
- Case in point - recent AGLT2 – SARA incident
  - Detected in FTS, confirmed by FAX and perfSONAR
  - Investigated and narrowed down with current perfSONAR network
  - Ticket opened with Internet2 – followed up with a coordinated ticket in GEANT
  - Issue still unresolved

perfSONAR in Troubleshooting

- Problems found SARA->AGLT2. FTS timed out because rate for large files < 2-300 Kbytes/sec
- perfSONAR tests confirmed similar results
- Setup “temp” debug mesh using mesh-config GUI
  - https://maddash.aglt2.org/maddash-webui/index.cgi?dashboard=Debug%20Mesh%20(temp)
- Opened ticket with I2, who opened ticket with GEANT
  - GEANT brought up LHCONE pS instance. Tests to AGLT2 showed 3 times BW vs (much closer) SARA
  - Problematic link identified. 0.2-0.5% packet loss



Impacts overall link throughput.
Many fixes tried. No solution yet.

Where should we track/document cases?

Network Incidents Follow up

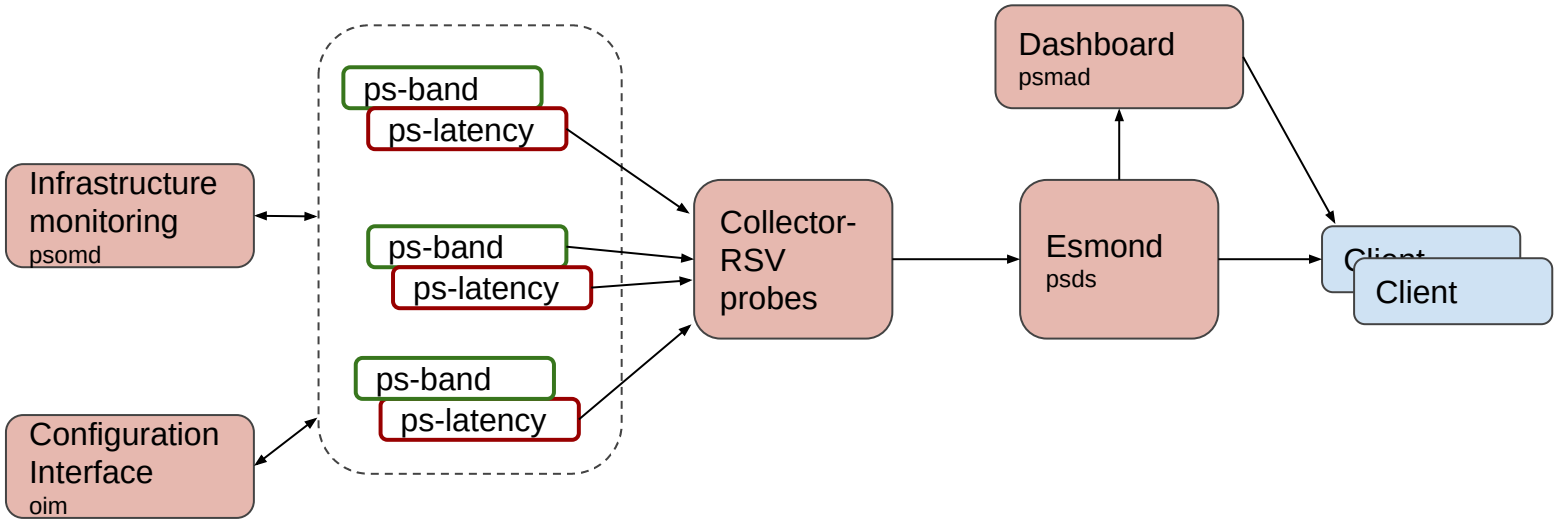
- Process strawman TBD
- Experiments
  - Report to mailing list (wlcg-network-throughput ?) – with experiments, wlcg perfsonar support unit (incl. lhcone/lhcopn experts)
  - WLCG perfSONAR support unit to confirm if this is network issue - contacts sites(?)
  - Concerned sites report to their NREN informing mailing list
  - List of ongoing incidents on WG page – detailed to be provided
- Sites
  - Report to their NREN, inform mailing list
  - WLCG perfSONAR support unit can assist if needed
  - Escalate to WLCG operations coordination to resolve non-technical (policy) inter-site issues

Integration projects

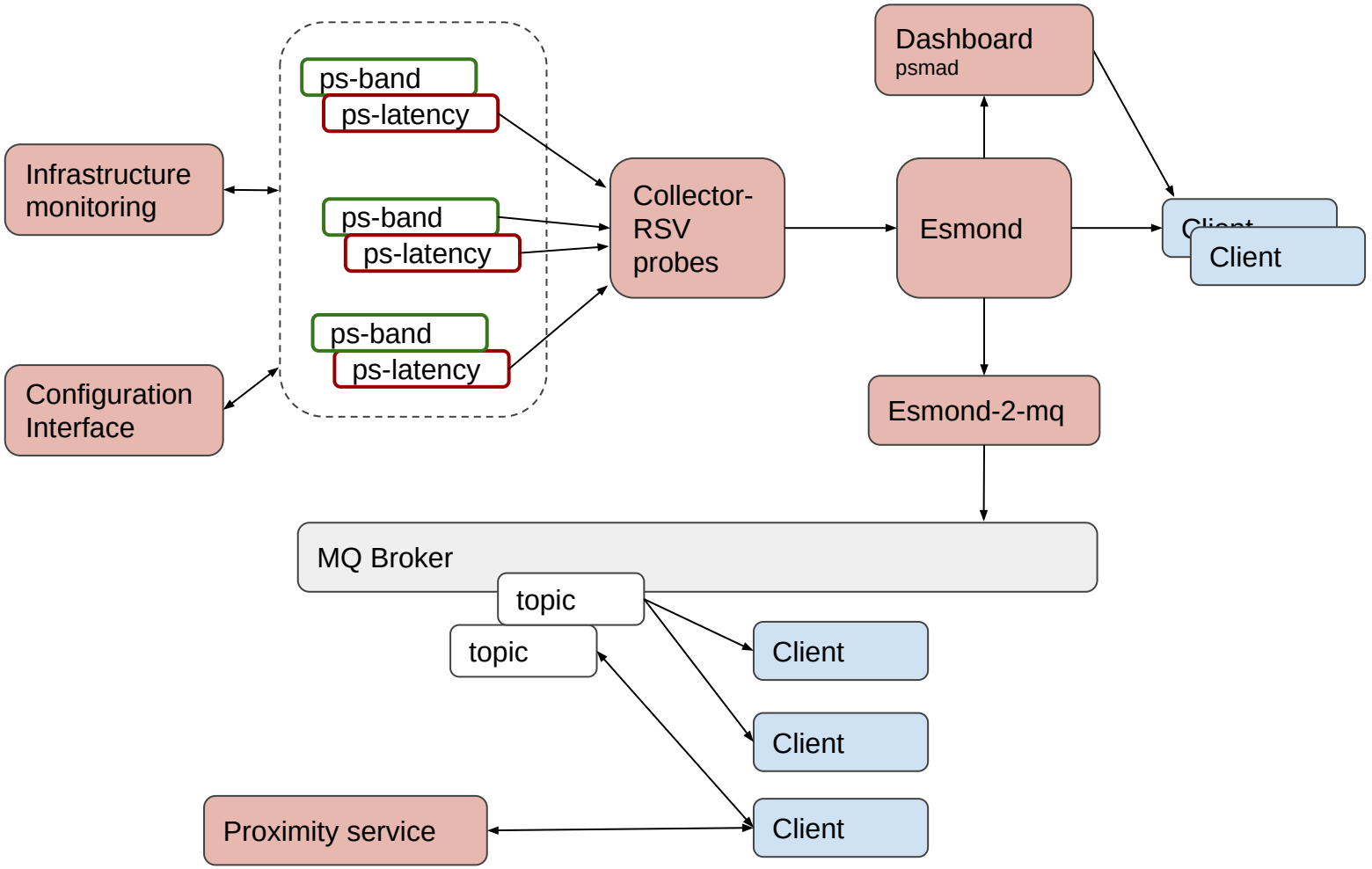
Experiments interface to esmond

- Meeting held 25th of February – revise the proposal taking into account comments received during the last meeting
- Aim
  - Develop experiment's 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
- Current event-types measured by sonar infrastructure
  - histogram-owdelay – one way delays over time period – in total 16k events/5 mins (currently 10%) - each event contains histogram
  - packet-loss-rate – number of packets lost/packets sent – in total 16k links/5 mins (currently 10%) – each event contains key/value pairs (ts, loss)
  - packet-count-sent – packets sent
  - packet-count-lost – packets lost
  - packet-trace – in total 16k events/hour (currently 60%) - each event contains tracepath
  - throughput – observer amount of data sent over period of time – in total 16k events/week – each event contains key/value pair (ts, throughput)

Architecture (Current)



Architecture (TBD)



Components

- RSV probes (OSG) – collecting metrics
- esmond (OSG) – datastore
- esmond2mq – Henryk developed a prototype in python
  - Retrieves all data (meta+raw) from esmond depending on existing mesh configs
  - Publishes to a topic
  - Lessons from prototype - esmond currently too slow to get raw data – requires optimization on both sides
- 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
- Clients
  - consume data from topic, optionally connect to proximity/topology service and re-publish extended metrics
  - but also connect directly to esmond if needed

AOB