

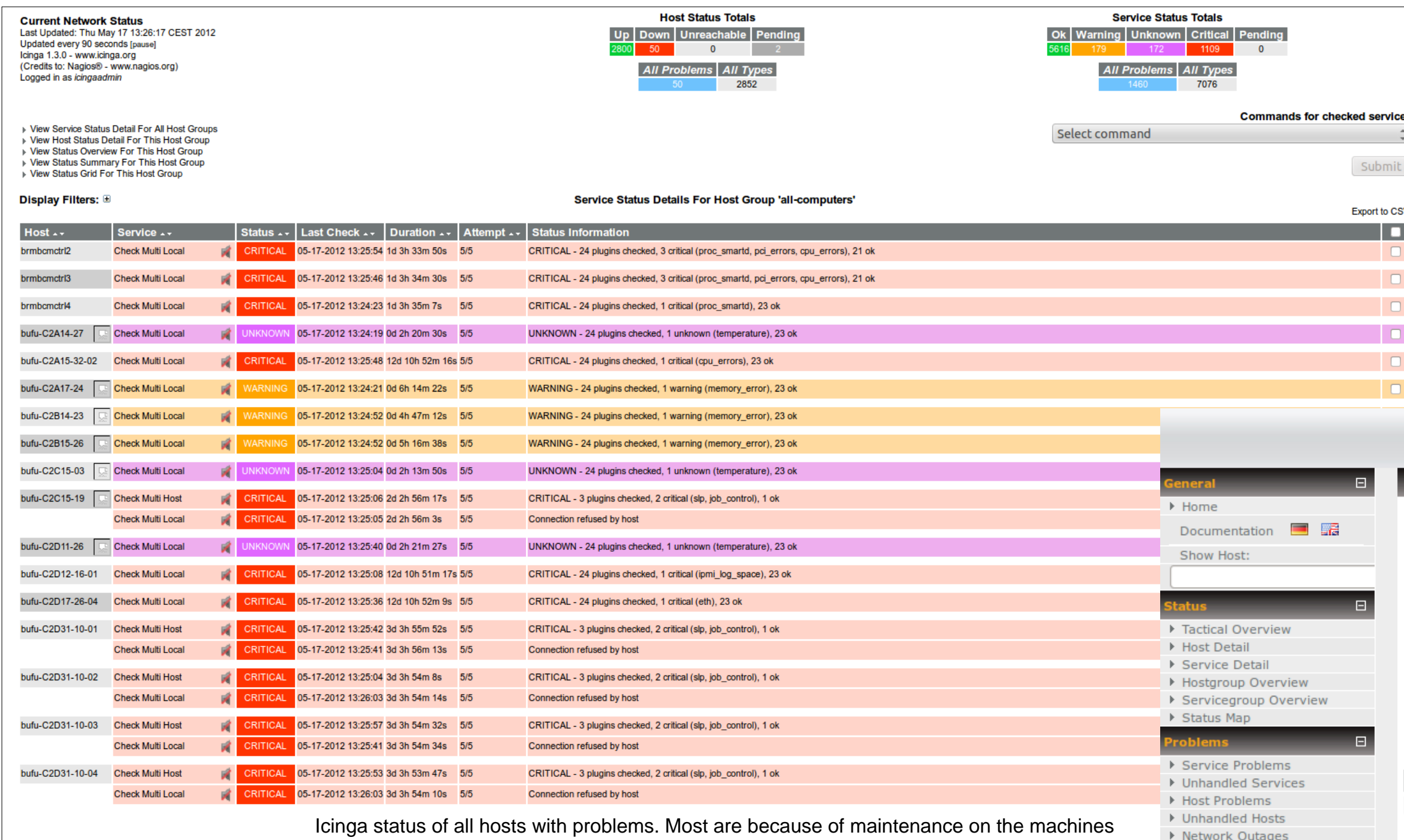


Health And Performance Monitoring Of The Online Computer Cluster Of CMS

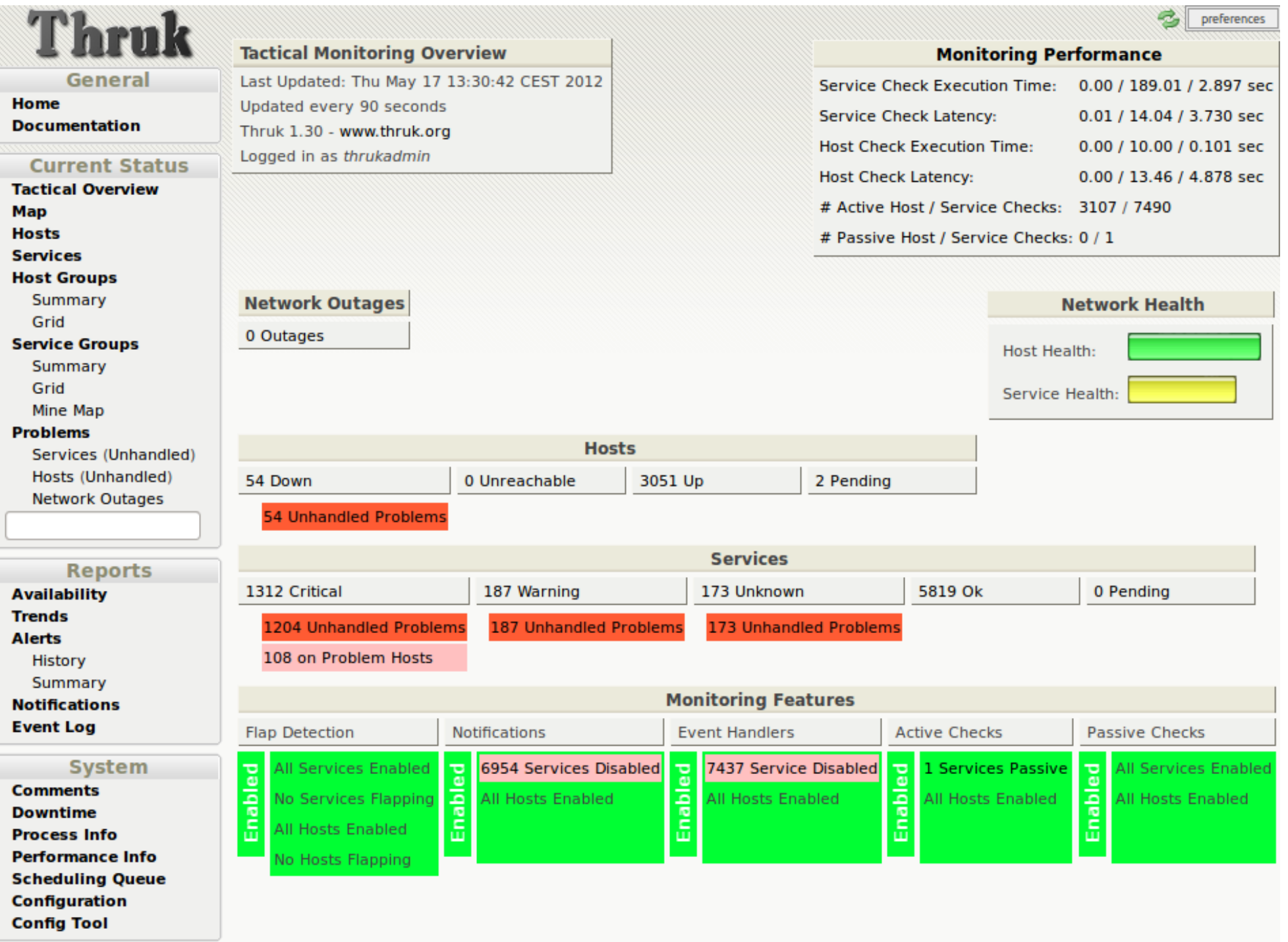
Olivier Raginel, on behalf of the CMS DAQ group



Massachusetts
Institute of
Technology



The CMS experiment's online cluster needs to be monitored constantly in order to react as quickly as possible to any problems that arise. When migrating our 4 Nagios servers from SLC4 to SLC5, it was decided to go with Icinga and some plugins to fix the issues we experienced in our growing cluster.



Thruk technical overview. Very similar to icinga's, but a bit nicer and easier to extend

Issues and solutions:

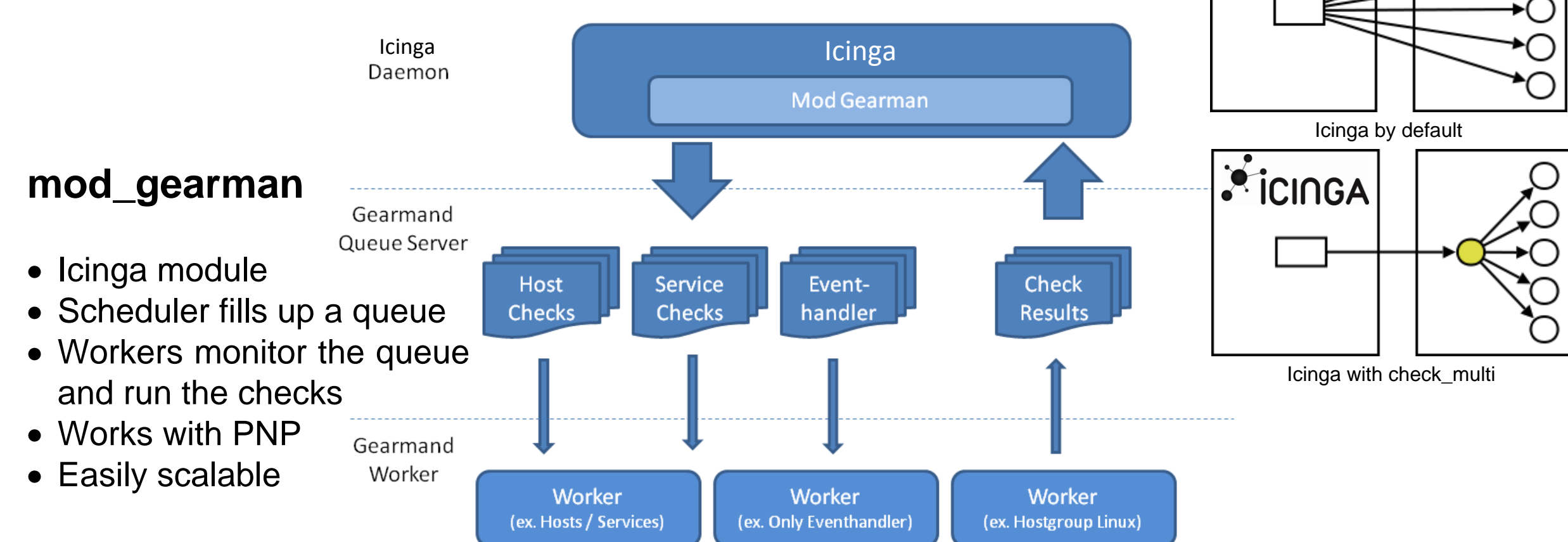
- nagios => icinga: many bug fixes
- hardcoded CGI interface => thruk: easily extendable
- PNP4nagios / RRDtool => same: generate plots
 - RRD host has 4x300 GB 15k SATA disks
- scalability issue => check_multi: channel service tests
 - 1 service, 25 checks in our configuration
- manual break of configuration per server => gearmand + mod_gearman: queue checks and performance data
- 4x8 cores machine => 1x4 cores (master + RRD) + 1x8 core (worker)
- 20'000 checks every 5 minutes => over 65'000 checks every 2 minutes

Monitoring System Summary:

- 2600 computers, 200 switches
- Configuration generated from Quattor
- 25 standard checks run through check_multi
- Additional checks added on a case-by-case basis
- Alerting based on groups, by sub-detector
- Email and SMS alerts

Icinga Configuration

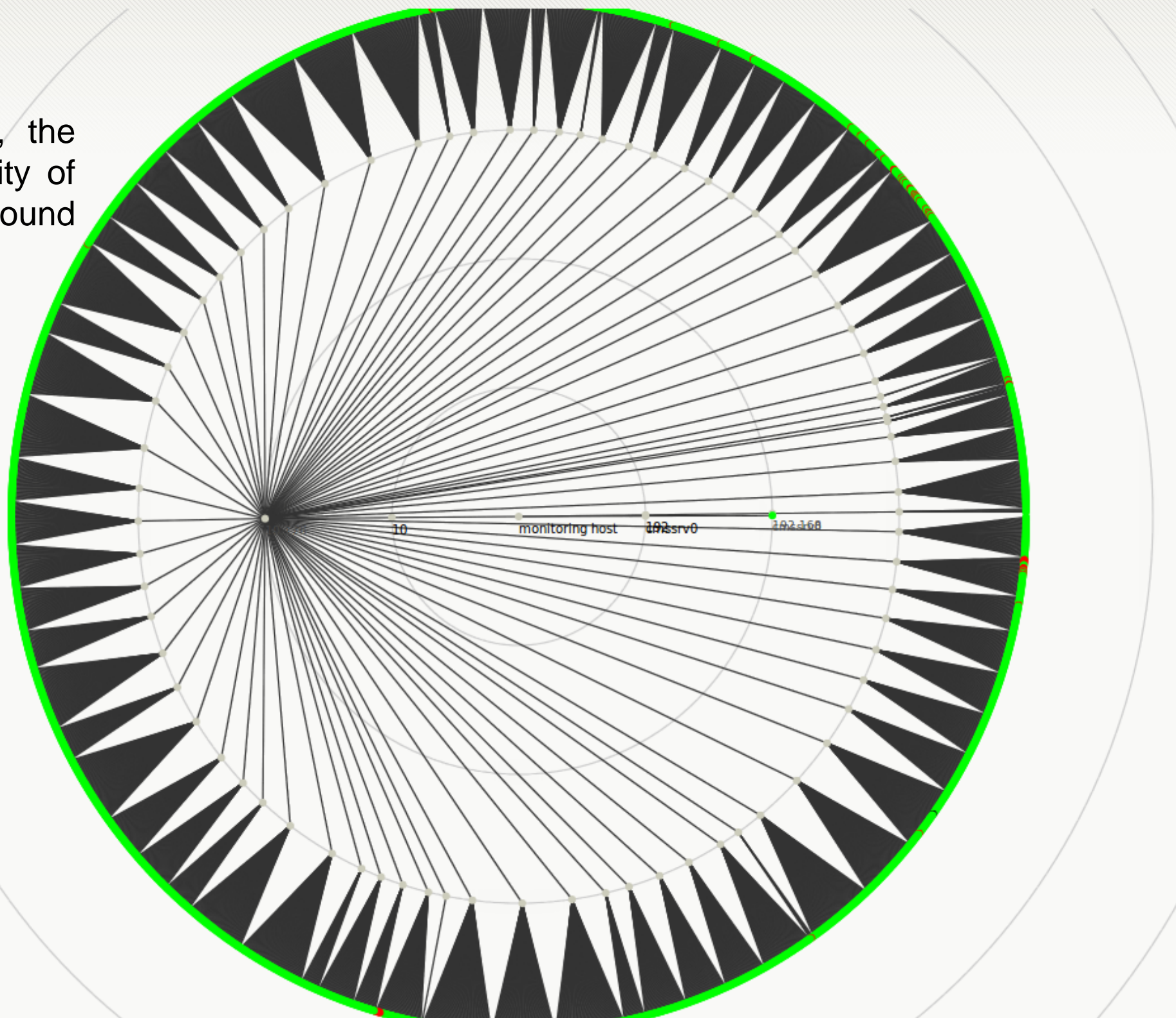
- Configuration settings automatically generated from a wiki page
- Processed by a perl script
- Generates hosts, based on list or Quattor templates (configuration management system)
- Also generates parents (network topology), contacts and groups



Thruk Map

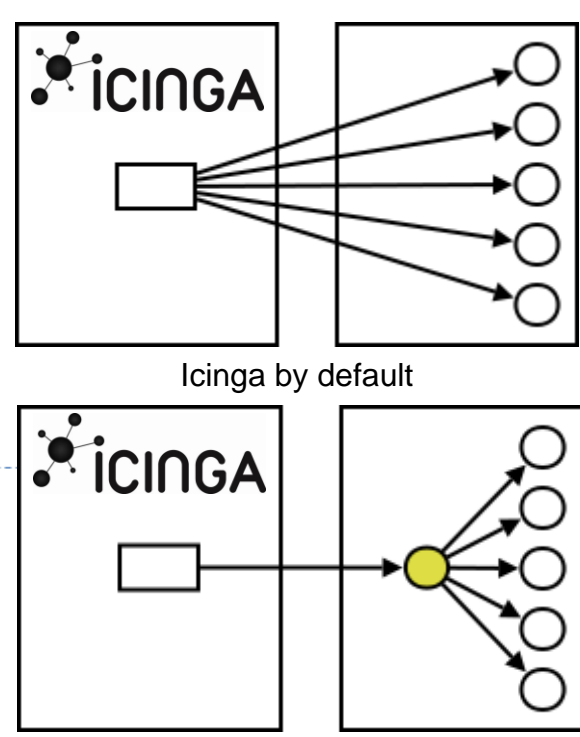
Parents being generated, the map reflects the complexity of the CMS online cluster: around 100 different networks

- Center point is the monitoring host
- First left point is the main class A network (10)
- Second point is the class B (10.176)
- All the class C derives from it
- Hosts are the final green / red points
- Right points are artefacts of the configuration



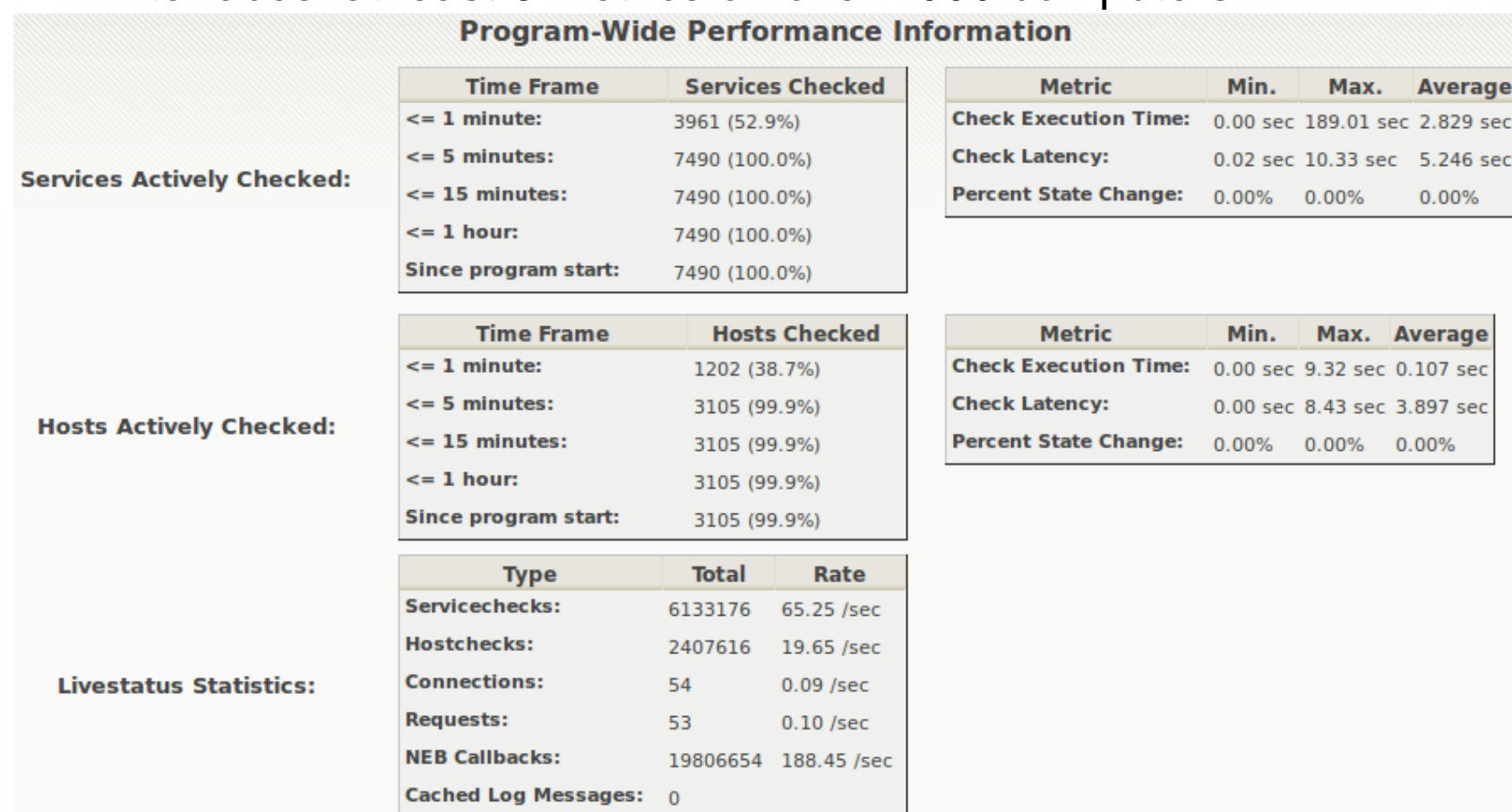
Check_multi

- Combines 25 tests into a single NRPE call
- Less load on monitoring server and worker



Thruk Performance Summary

- With only 12 cores, 4 years old CPUs, performance is impressive:
- 2800 devices are checked every 2 minutes
- Over 25 services are checked on each device
- Half the checks return performance data
- The network check return performance data for all the network interfaces: at least 6 metrics on over 2000 computers



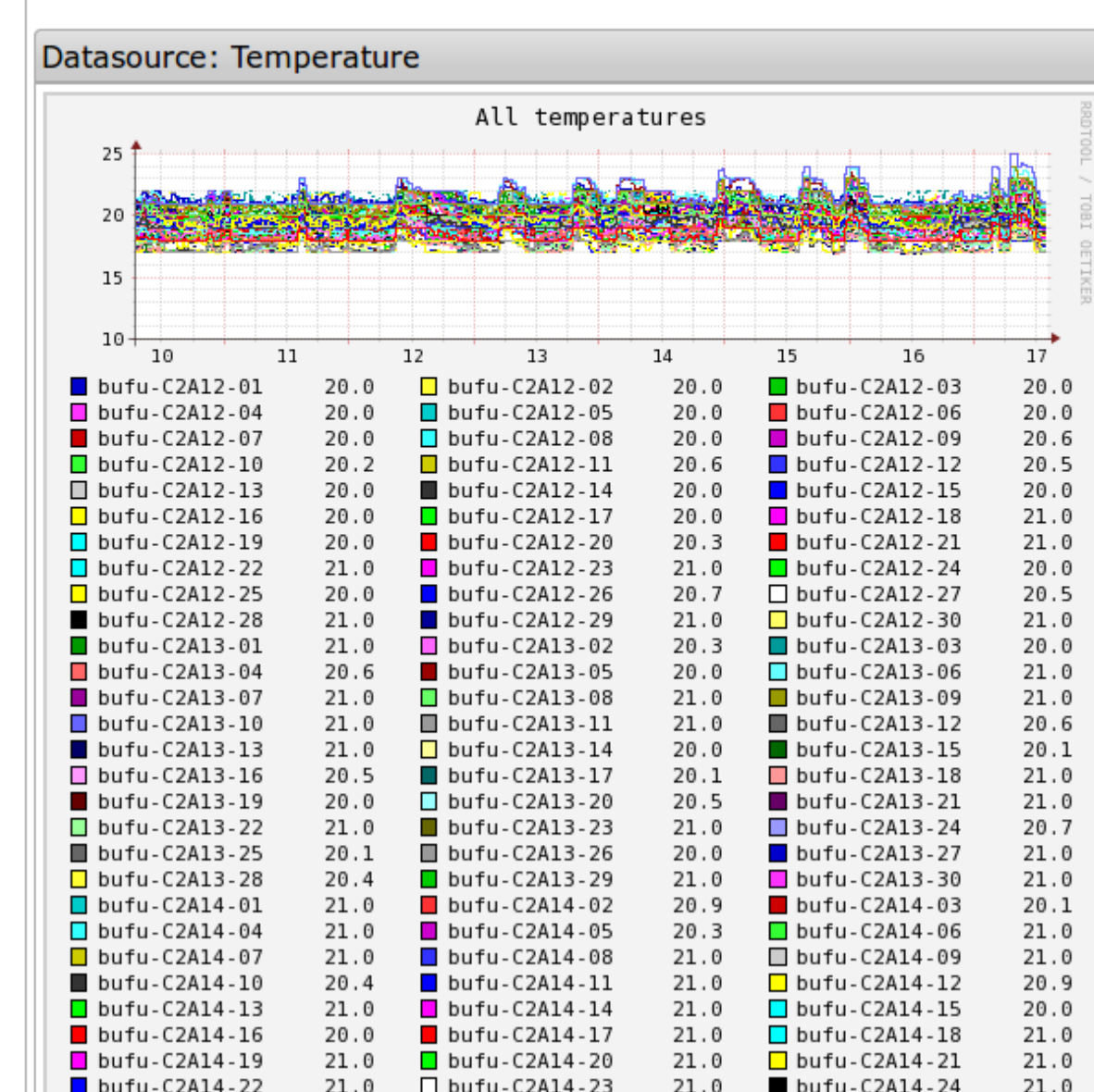
Temperature Plots

PNP collects all performance data generated by the checks, and graphs them, automatically. It can then be used to generate complete reports so any anomaly can be quickly spotted. One can correlate the slight raise in overall farm temperature with CMS being in data taking mode.

Temperature on hosts from BUFU racks

Temperature of the 1028 BUFU machines from racks C2A12 to C2F:

One Week 10.05.12 6:46 - 17.05.12 13:46



Plot Merging

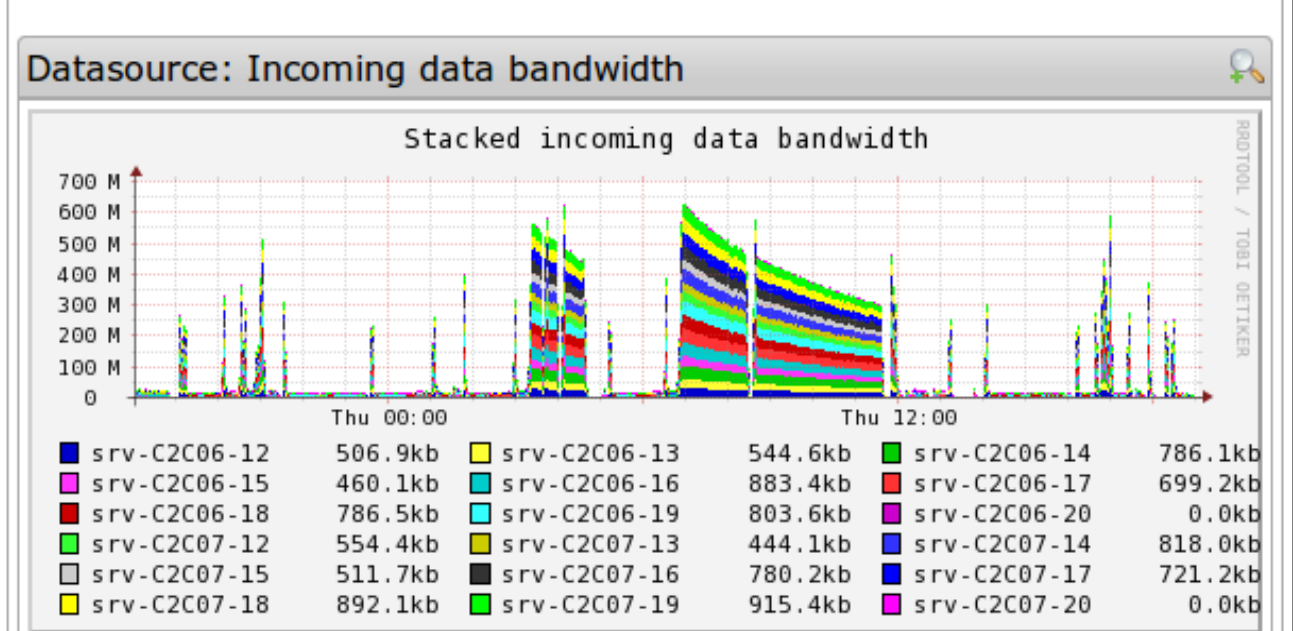
PNP is used to generate complete reports, and put together several metrics from multiple hosts. All the network interfaces of the machine can be stacked together to reflect the overall flows:

- 4 data interfaces
- 2 Tier0 interfaces bound together
- 1 service interface
- All 18 machines are easily viewable on one plot.
- The plots reflect the luminosity decrease

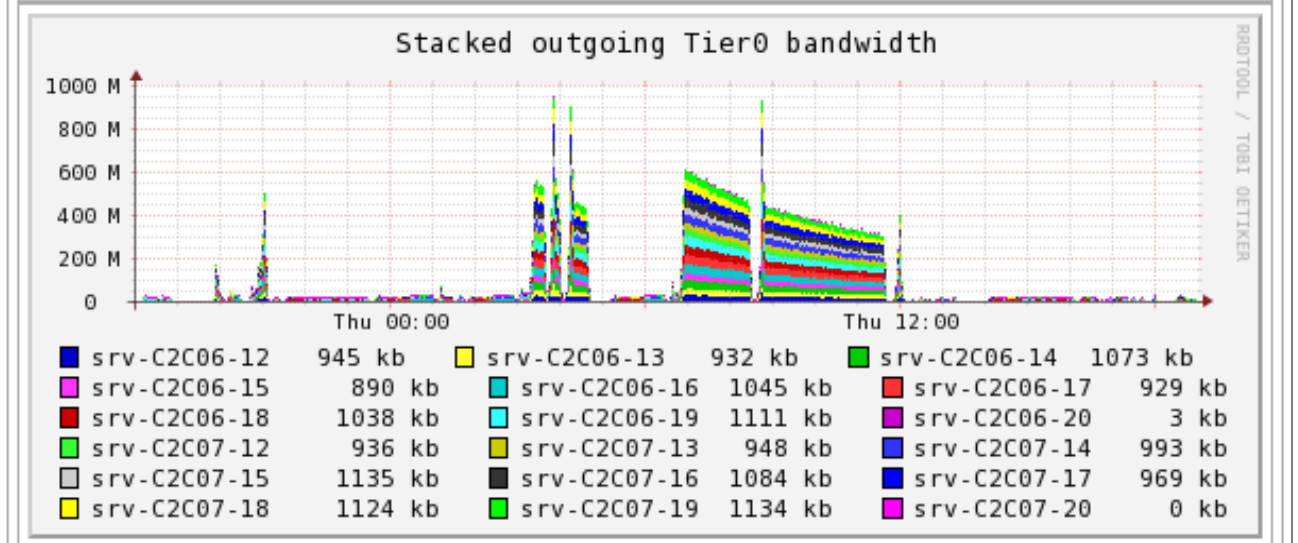
Data bandwidth on all SM hosts

Statistics of bandwidth on the data network per node

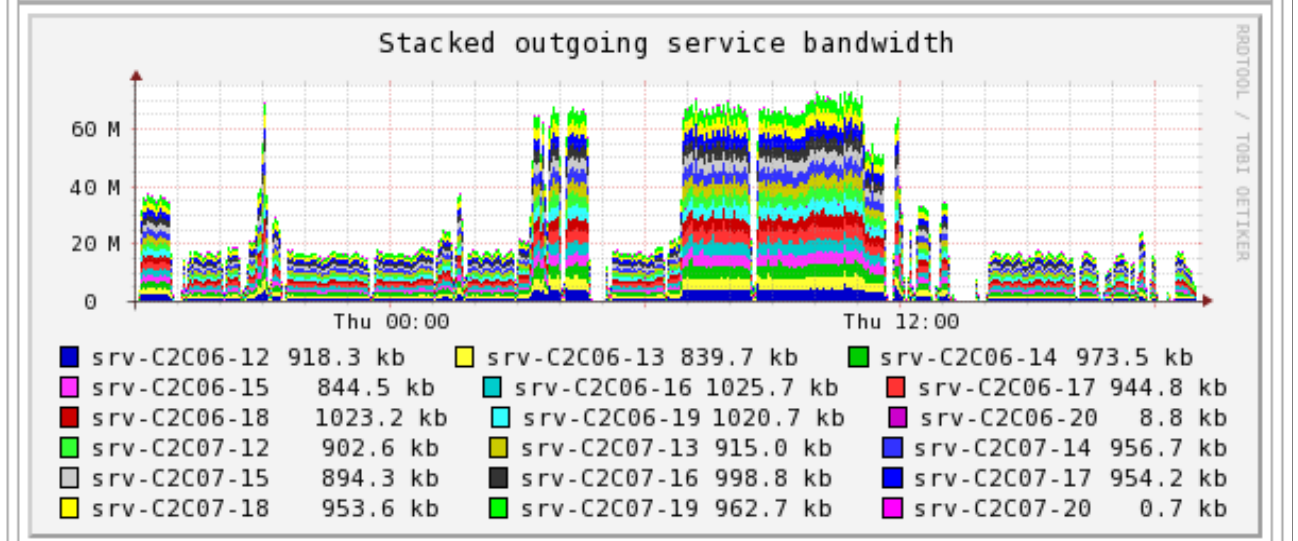
25 Hours 16.05.12 18:00 - 17.05.12 19:00



Datasource: Outgoing Tier0 bandwidth



Datasource: Outgoing service bandwidth



CHEP 2012