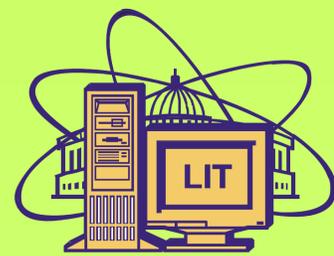


Tier 3 Monitoring Software Suit (T3MON)

S. Belov, I. Kadochnikov, N. Kutovskiy, D. Oleynik, A. Petrosyan, L. Valova
Laboratory for Information Technologies, Joint Institute for Nuclear Research, Dubna, Russia



The ATLAS Distributed Computing activities were concentrated so far in the "central" part of the computing system of the experiment, namely the first 3 tiers (CERN Tier0, 10 Tier1s centers and 60+ Tier2s). This is a coherent system to perform data processing and management on a global scale and host (re)processing, simulation activities down to group and user analysis.

Many ATLAS Institutes and National Communities already built or have plans to build Tier3 facilities. The definition of Tier3 concept has been outlined [1,2]. Tier3 centers consist of non-pledged resources mostly dedicated for the data analysis by the geographically close or local scientific groups.

Tier3 sites comprise a range of architectures and many do not possess Grid middleware, which would render application of Tier2 monitoring systems useless. T3MON software suite enables local monitoring of the Tier3 sites and the global view of the computing activities of the LHC virtual organizations at the Tier3 sites.

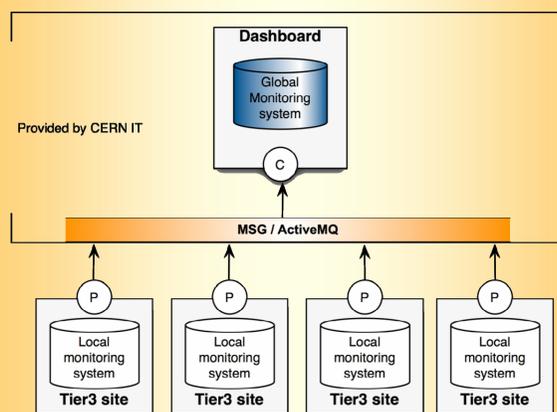
Aims of the project

- Provide a reasonable monitoring solution for 'off grid' sites (unplugged geographically close computing resources)
- Monitoring of computing facility of local groups with a collocated storage system (Tier1+Tier3, Tier2+Tier3)
- Present Tier3 sites activity on global level
- Data transfer monitoring across XRootD federation

Tier3 sites monitoring levels

There are two levels of Tier3 sites monitoring:

- Monitoring of the local infrastructure for site administration.
 - Central system for monitoring of the VO activities at Tier3 sites.
- This system should collect, aggregate and display information from the local Tier3 monitoring systems.



The objectives of the local monitoring system for a Tier3 site:

- Detailed monitoring of the local fabric (overall cluster or clusters monitoring, monitoring each individual node in the cluster, network utilization)
- Monitoring of the batch system (distribution of tasks on nodes)
- Monitoring of job processing (PROOF, PBS, OGE, Condor).
- Monitoring of the mass storage system (total and available space, number of connections, I/O performance)
- Monitoring of VO computing activities at a local site

The objectives of the global Tier3 monitoring:

- Monitoring of the VO usage of the Tier3 resources in terms of data transfer, data access and job processing and the quality of the provided service based on the job processing and data transfer monitoring metrics.

Site monitoring

Based on Ganglia monitoring system

- Collect basic metrics
- Plugin system for monitoring specific metrics
- Job processing systems, storage solutions, specifics protocols
- Special solution for XRootD and PROOF monitoring

XRootD

- Special representation of monitoring data: summary, detailed stream
 - Summary stream provides a lot of information for understanding the "health" of the system
 - A lot of parameters (see <http://tinyurl.com/bn2or4l> for details)
- Detailed monitoring streams: provides information about file access
- Special parser with different backends.

PROOF

- Set of job related parameters to collect: wall time, CPU time, memory consumption, number of read bytes, number of workers, exit code, number of processed events...
 - Information about data access from job: list of files, list of datasets
- Data collected in local database (PostgreSQL) presented by Ganglia
- PROOF add-on: stores Proof report to PostgreSQL
 - Metrics:
 - User
 - Start/end time
 - CPU time, Wall time
 - Dataset name
 - Number of files in the dataset
 - Number of events, Number of workers

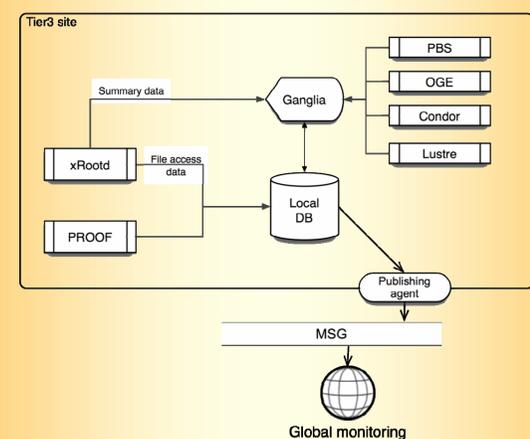
Global monitoring

The central Tier3 monitoring system is based on monitoring data published by Tier3 sites and will provide a global picture of how ATLAS uses Tier3 resources. The necessary condition for the development of the central Tier3 monitoring system is a consistent registration of the Tier3 sites in the ATLAS Grid Information System (AGIS). Another important factor is encouraging the Tier3 user community to use data transfer and job submission systems which are instrumented for reporting the monitoring data, like for example Ganga, Pathena, ATLAS DQ2.

Main components:

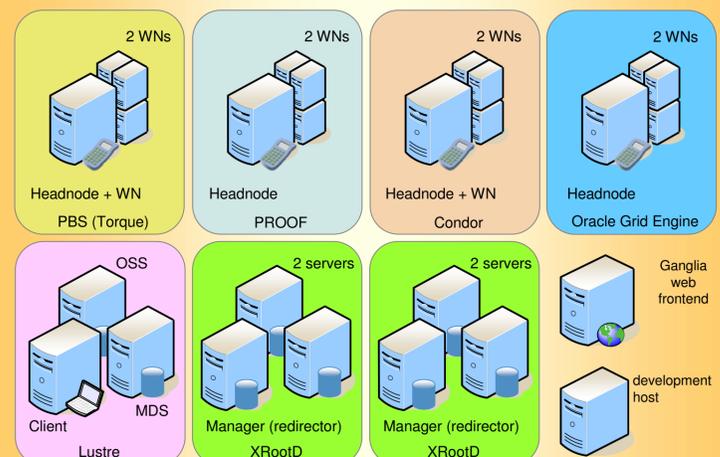
- Publisher on local site: intercommunicate with local DB and send information to MSG system (prototype is ready)
- MSG as a transmitting system: provided and supported by CERN IT
- Backend: consumer(s) of messages and integration with Dashboard system

Data flows in T3Mon



Tier3 testbed

- Created and supported by JINR
- 6 sites; virtual servers on multicore nodes
- Deployed most popular facilities (in accordance with ATLAS Tier3 sites survey):
 - XRootD
 - PROOF
 - Lustre
 - Oracle Grid engine (OGE, former Sun Grid Engine)
 - PBS (Torque)
 - Condor
- Set of applications for simulation load from real users



Project's framework and used technologies

- Monitoring tools distributions are prepared for Scientific Linux 5
- Python 2.6 used as a base programming language
- Data Base: PostgreSQL
- Ganglia monitoring system
 - RRD and PostgreSQL for storage
 - Web UI for visualization
- Build and packaging system adopted from ARDA Dashboard
- YUM repository for installation of T3Mon RPM packages on sites

References

[1] Tier3 Workshop. 25-26 January 2010 CERN, <https://indico.cern.ch/conferenceDisplay.py?confId=77057>

[2] U.S. ATLAS Tier 3 Task Force (March 27, 2009) by R. Brock et al.

[3] J. Andreeva, D. Benjamin, A. Klimentov, V. Korenkov, D. Oleynik, S. Panitkin, A. Petrosyan, Tier-3 Monitoring Software Suite (T3MON) proposal, ATL-SOFT-PUB-2011-001, <http://cdsweb.cern.ch/record/1336119>