



Contribution ID: 125

Type: **Oral Presentation**

A new integrated monitoring tool for grid site administrator

Wednesday, 13 April 2011 11:00 (30 minutes)

Overview

In this work we will show the development and the work carried on in order to build a monitoring tool that gives an aggregate view of all the users activities on a given grid site.

The tool is able to show the job submitted by each user together with information about the file accessed on the storage system. Also in a farm with posix like parallel file-system, the tool is able to track down both SRM standard operation and the local "posix" file access.

We will put a particular attention in highlight how this monitoring system works in a mixed environment like a farm used both via grid and with local job submission. Moreover it could easily work with different type of computing elements and batch system, as it is highly modular and customizable.

This monitoring system will help the sys-admin to have a complete and detailed view of what is happening with the computing center.

Impact

All information gathered in the central database, could be aggregated and presented by users, host, processes etc. in order to have a much clearer view of what is happening on the computing nodes. Using this monitoring facility indeed makes far easier to track down misuse of the computing facilities from a given user. The database was designed in order to speed up the procedure of building web pages. This was obtained by means of an accurate data movement between on-line and near-line tables and a deep study on how to speed up the query needed to build the interfaces. Using this monitoring tool it is also possible to keep under control the disk space used by each VO or user. This monitoring infrastructure was already used with success at INFN-Bari site in order to have statistics on CMS dataset usage: this is particularly useful in order to take decisions about the dataset deletion. It is also easy to monitoring the usage of the Wide Area Network bandwidth as it is possible to have a clear view of the users that are transferring file into (or out of) the farm, and it is also possible know which files are transferred. Looking to the job submission, using the web interface, it is possible to find out several information: how many jobs are submitted from a given DN, or from a given VOMS group. Which executables they are running on the farm, and which node they are using. It was already evident, while using it at a medium-large site made by 1000 CPU and 700TB of storage, that the system is able to handle a huge amount of data without decreasing the responsiveness of the web interface. In the monitoring system are also integrated several status views for each service that should be up and running in order to have normal operational condition on the farm. This monitoring system is capable (differently from other system already developed) to give a complete and aggregate view of the status of the farm together with the capability to have historical information on each observed metric

Description of the work

The monitoring system was built starting from different agents and monitoring services. There is a central database system that take care of storing, aggregating and presenting the information gathered from each monitored node.

In particular each Computing Element has its own agent in order to sent to the information about the jobs, indeed this agent provide information about: user DN, FQAN, grid-jobid, local-jobid, queue, local user, VO.

Also the StoRM and the gridftp servers provide information about the file accessed both from the farm itself and/or from remote sites. Also in this case the monitoring agents provides: DN, FQAN, name and path of the file, VO. For each file accessed locally through lustre file-system, the local user that access the file, the node from which the file is accessed, the pid of the process accessing the file, the name of the process accessing the file.

Thanks to sensors installed in all the nodes of the farm it is possible for the site admin to know each accessed file over a Lustre/GPFS parallel file-system.

All the monitoring agent are as lightweight as possible in order to run it every one or few minutes.

By design the monitoring system allow the sys-admin to change, add, or switch off each plug-in used to find out the data.

The Database schema is built in order to keep track of the dependency between several observed values: for example it is easy to match between the job running on a given node, and the files accessed by the user on that machine.

In order to gather as much information as possible and in order to be easily adapted to new batch system, we uses pre and post exec scripts.

We have already developed sensors for several services: LCG-CE, CREAM-CE, StoRM, Gridftp servers, Xrootd servers, Torque/Maui, Lustre.

The web interface that allow the site admin to look at the status of the farm exploits the new available web technologies in order to give to the final user an advanced user experience.

Conclusions

In every grid computing farm, there are several services and each of those has different logging or monitoring system, it is important to have a unique point in which the site admin could aggregate and analyze all the information in order to have a clear view of what is going on.

This is much more important in those computing farms that are shared among several VOs in order to guarantee that activities of a VO is not blocked or affected by others users.

Using this tool the site administrator in the INFN-BARI farm were able to better fulfill the user requirements in terms of performance and reliability and also detect proactively problems and failures that could stop or affect the users activities.

Primary author: Dr DONVITO, Giacinto (INFN)

Co-authors: Prof. MAGGI, Giorgio Pietro (Politecnico di Bari + INFN); SPINOSO, Vincenzo (Università degli studi di Bari)

Presenter: Dr DONVITO, Giacinto (INFN)

Session Classification: User Support Services

Track Classification: User Support Services - Infrastructure