



Contribution ID: 174

Type: **Poster**

Health and performance monitoring of the large and diverse online computing cluster of CMS

Tuesday, May 22, 2012 1:30 PM (4h 45m)

The CMS experiment online cluster consists of 2300 computers and 170 switches or routers operating on a 24 hour basis. This huge infrastructure must be monitored in a way that the administrators are proactively warned of any failures or degradation in the system, in order to avoid or minimize downtime of the system which can lead to loss of data taking. The number of metrics monitored per host varies from 20 to 40 and covers basic host checks (disk, network, load) to application specific checks (service running) in addition to hardware monitoring (through IPMI). The sheer number of hosts and checks per host in the system stretches the limits of many monitoring tools and requires careful usage of various configuration optimizations in order to work reliably. The initial monitoring system used in the CMS online cluster was based on Nagios, but suffered from various drawbacks and did not work reliably in the recently expanded cluster. The CMS cluster administrators investigated the different open source tools available and chose to use a fork of Nagios called Icinga, with several plugin modules to enhance its scalability. The Gearman module provides a queuing system for all checks and their results allowing easy load balancing across worker nodes. Supported modules allow the grouping of checks in one single request thereby significantly reducing the network overhead for doing a set of checks on a group of nodes. The PNP4nagios module provides the graphing capability to Icinga, which uses files as round robin databases (RRD). Additional software (rrdcached) optimizes access to the RRD files and is vital in order to achieve the required number of operations. Furthermore, to make best use of the monitoring information to notify the appropriate communities of any issues with their systems, much work was put into the grouping of the checks according to, for example, the function of the machine, the services running, the sub-detectors they belong to, and the criticality of the computer. An automated system to generate the configuration of the monitoring system has been produced to facilitate its evolution and maintenance. The use of these performance enhancing modules and the work on grouping the checks has yielded impressive performance improvements over the previous Nagios infrastructure allowing for the monitoring of X metrics per second (compared to Y on the previous system). Furthermore the design allows the easy growth of the infrastructure without the need to rethink the monitoring system as a whole.

Primary author: RAGINEL, Olivier (Massachusetts Inst. of Technology (US))

Co-authors: Mr FLOSSDORF, Alexander (DESY); HOLZNER, Andre Georg (Univ. of California San Diego (US)); PETRUCCI, Andrea (CERN); SPATARU, Andrei Cristian (CERN); Dr RACZ, Attila (CERN); DUPONT, Aymeric Arnaud (CERN); DELDICQUE, Christian (CERN); HARTL, Christian (CERN); PAUS, Christoph (Massachusetts Inst. of Technology (US)); SCHWICK, Christoph (CERN); SHPAKOV, Dennis (Fermi National Accelerator Lab. (US)); GIGI, Dominique (CERN); MESCHL, Emilio (CERN); GLEGE, Frank (CERN); MEIJERS, Frans (CERN); BAUER, Gerry (Massachusetts Inst. of Technology (US)); Dr POLESE, Giovanni (CERN); SAKULIN, Hannes (CERN); BRANSON, James (Univ. of California San Diego (US)); Dr HEGEMAN, Jeroen (CERN); Dr COARASA PEREZ, Jose Antonio (CERN); SUMOROK, Konstanty (Massachusetts Inst. of Technology (US)); MASETTI, Lorenzo (CERN); ORSINI, Luciano (CERN); Dr DOBSON, Marc (CERN); PIERI, Marco (Univ. of California San Diego (US)); CIGANEK, Marek (CERN); SANI, Matteo (Univ. of California San Diego (US)); BOWEN, Matthew (University of the West of England); SIMON, Michal; BOUFFET, Olivier (CERN); MOMMSEN, Remi (Fermi

National Accelerator Lab. (US)); GOMEZ-REINO GARRIDO, Robert (CERN); ERHAN, Samim (Univ. of California Los Angeles (US)); BUKOWIEC, Sebastian (CERN); CITTOLIN, Sergio (Univ. of California San Diego (US)); BEHRENS, Ulf (Deutsches Elektronen-Synchrotron (DE)); O'DELL, Vivian (Fermi National Accelerator Laboratory (FNAL)); HWONG, Yi Ling (CERN)

Presenter: RAGINEL, Olivier (Massachusetts Inst. of Technology (US))

Session Classification: Poster Session

Track Classification: Computer Facilities, Production Grids and Networking (track 4)