



Contribution ID: 50

Type: not specified

Using L&B to Monitor Torque Jobs across a National Grid

Wednesday, 28 March 2012 14:45 (45 minutes)

Description of the Work

The basics of the PBS job state diagram, identical with Torque, have already been present in L&B for some time. The set of states and transition rules has not been complete, though, since PBS was a closed-source application and information about jobs for L&B, up to that time, had to be generated by PBS log parsers. Some details were obviously not discernible from any of the logs.

With the transition to an open-source solution, the opportunity arose to actually insert the event logging calls into the batch system itself, making it possible to avoid log parsing altogether while also covering several additional sub-states and extracting values that cannot be usually read from log files. The instrumented Torque server contacts the L&B server directly using L&B's proven custom messaging infrastructure.

On submission, a jobID is returned to the user indicating not only the torque server they have contacted, but most importantly the address of the L&B server that holds job status information (same JobID structure as that of native gLite jobs). The user can then use all traditional L&B features (queries for job state, notifications...) regardless of whether the job has been forwarded to another Torque server or not. An L&B-oriented 'qstat' command alternative has even been implemented to mock the behavior of PBS/Torque's native qstat so that the end user sees no difference.

Another issue that had to be addressed related to authentication. The authentication infrastructure in MetaCentrum is not X.509-based but rather Kerberos-based, but this did not pose much of a problem in the end. L&B authentication layer, relying on GSS, provides very good abstraction from the actual authentication method used, and Kerberos is thus fully supported. L&B internals simply treat the user's principal name in the same way they would normally treat a DN.

Conclusions

gLite's Logging and Bookkeeping service, currently developed and released by EMI, has been successfully used as a solution for tracking grid jobs over multiple batch system instances. As it was already designed to handle different types of jobs at the same time, extending it to accept Torque job details was fairly straightforward and, combined with inserting appropriate logging calls into Torque's source, proved to be a one-person task. L&B has established itself further as a middleware-independent monitoring tool that performs well as a stand-alone monitoring solution for many kinds of processes. It helped to make Czech NGI MetaCentrum more redundant and scalable. Changes required in Torque to send appropriate events to L&B are available to the community.

Impact

In a distributed environment with multiple Torque instances, L&B serves as a common monitoring service that is aware of all jobs regardless of where they run or whether they have migrated between several Torque

worlds. It is the primary point of contact for users wishing to follow the status of their jobs. All standard L&B client tools are available, but –to simplify the transition –existing MetaCentrum users may also rely on the qstat command, which has been reimplemented to provide the same output as traditional qstat, but with information collected from the L&B server rather than individual Torque instances. Thus MetaCentrum users now have at their disposal all the functionality of L&B (notifications –pull model, user or application specific job tagging) with a backward compatible interface of qstat.

Undeniably, at this point, the L&B server constitutes a single point of failure (improbable since L&B is a proven and tested solution) but any such occurrence would only (temporarily) affect the users' ability to query job states since all message passing channels have been designed for reliability and delivery of event records is always assured. Short-term unavailability to users will be addressed in the future by making multiple L&B servers capable of handling each other's information.

Another issue that may be addressed in the future consists in pairing Kerberos principals with DNs so that users relying on Kerberos-based authentication in their UI may access their job information over HTTPs with nothing but standard X-509 certificates installed in their Web browsers.

Overview (For the conference guide)

gLite's Logging and Bookkeeping service (L&B) is designed to support other types of jobs than native gLite. Support for PBS jobs was first introduced at HPDC 2008, and a pilot deployment was performed in MetaCentrum, the Czech national grid that was relying on PBS as its central batch system at that time. Later, MetaCentrum commenced transition to Torque, an open-source alternative to PBS. To suit MetaCentrum's needs better, Torque was extended to allow for running several independent instances, which could nevertheless forward jobs to each other if one instance could not satisfy current resource requirements. This led to a new task: providing a tool to monitor grid jobs potentially migrating over several instances of Torque. L&B was invited to provide that functionality.

Primary authors: Dr KOURIL, Daniel (CESNET); Mr DVORAK, Frantisek (CESNET); Mr FILIPOVIC, Jiri (CESNET); Mr SITERA, Jiri (CESNET); Prof. MATYSKA, Ludek (CESNET); Mr POUL, Marcel (CESNET); Mr VOCU, Michal (CESNET); Dr RUDA, Miroslav (CESNET); Mr TOTH, Simon (CESNET); Mr SUSTR, Zdenek (CESNET)

Presenter: Mr VOCU, Michal (CESNET)

Session Classification: EMI: Job processing components