**5th EGEE User Forum** 



Contribution ID: 87

Type: Oral

# Towards an interoperable representation of the Grid Observatory information: an experiment with Common Base Event

Wednesday 14 April 2010 11:00 (20 minutes)

Reaping the full benefit of the Grid Observatory (GO) initiative requires providing end-users with convenient representation of the traces. The present lack of standardization creates considerable difficulties for developing automated analysis and situation handling solutions. We report an experiment on the representation of internal logs of the gLite WMS with the IBM Common Base Event (CBE) format, categorizing the information of hundred of thousand events to a few generic type situations, with potential applications of the open-source supporting technologies to job monitoring.

# **Detailed** analysis

The CBE XML schema defines the format of an event, which at its core is a 3-tuple (reporting component, impacted component, situation). Component may be hardware of software. Twelve generic situations are available, e.g. Start, Stop, and the very important Feature and Dependency describing component availability. As a test case, we considered the CondorG logs of GRIF-LAL site from 2008-09-16 to 2009-03-24, with 883,701 events amongst which 118,191 are associated to one or several identifiers. The main challenge is to type each logged event into a situation. The CondorG log format being un-documented, we developed a software suite for elucidating its syntax and to some extent semantics. The syntax is "class::fonction:message"; we identified 7 different classes, and 21 functions. We are currently in the process of segmenting the messages (finding and deleting url or process identifiers) in order to type them. The next step will be interpreting each type as a suitable CBE situation type. For instance, the event "ControllerLoop::run(): Aborting daemon..." will be classified as a CBE "StopSituation" with arguments "successDisposition=SUCCESSFUL" and "situationQualifier=ABORT INITIATED".

# **Conclusions and Future Work**

One year of GO experience shows that, despite documentation efforts, the simplest traces are by far the most used. The GO will exploit our work with CBE for the presentation of the traces broadening the scope of their utilization by the computer science community.

The a posteriori conversion of the operational logs is also an opportunity for interacting with system administrators about the possible usage of the advanced tools from IBM to improve the QoS to the end-users. These interactions will help assessing the interest of the technology for the future evolution of the EGI middleware.

#### Impact

The second priority of the GO, after data collection, is to provide parsimonious and informative representations of the traces. A first step is converting the traces towards user-friendly formats, in other words taking the burden of the "80% preprocessing" of data mining. While some traces are natively organized along standards (e.g. the information system –IS - is Glue compliant), other ones use proprietary formats (e.g. the Logging and

Bookkeeping –LB -), and are in the worst case fully undocumented (internal logs of the WMS). To maximize the added value of the costly conversion process, the target format should 1) be a (de-facto) standard and 2) come with an exploitation framework. We choose CBE, which as a format and associated technologies (automatic analysis engine, visualization tool) is the result of IBM's extensive experience with autonomic management. CBE is not suitable for all gLite logs, e.g. the IS, or in the WMS scope jobmap, are not event-oriented. However, CBE adequately covers many of them: in the WMS scope, wmproxy, jobcontroller, condorG, logmonitor and workload manager; outside, the LB. Our work will be a first step towards consolidating these disparate sources of information.

## Keywords

Monitoring, standardization, events

## URL for further information

www.grid-observatory.org

Authors: Prof. GERMAIN-RENAUD, Cecile (LRI); Dr GAURON, Philippe (LRI)

Presenter: Dr GAURON, Philippe (LRI)

Session Classification: Computer Science

Track Classification: End-user environments, scientific gateways and portal technologies