Contribution ID: **419**   Type: **poster**

Distributed Tracking, Storage, and Re-use of Job State Information on the Grid

Wednesday, 29 September 2004 10:00 (1 minute)

The Logging and Bookkeeping service tracks job passing through the Grid. It collects important events generated by both the grid middleware components and applications, and processes them at a chosen L&B server to provide the job state. The events are transported through secure reliable channels. Job tracking is fully distributed and does not depend on a single information source, the robustness is achieved through speculative job state computation in case of reordered, delayed or lost events. The state computation is easily adaptable to modified job control flow.

The events are also passed to the related Job Provenance service. Its purpose is a long-term storage of information on job execution, environment, and the executable and input sandbox files. The data can be used for debugging, post-mortem analysis, or re-running jobs. The data are kept by the job-provenance storage service in a compressed format, accessible on per-job basis. A complementary index service is able to find particular jobs according to configurable criteria, e.g. submission time or "tags" assigned by the user. A user client to support job re-execution is planned.

Both the L&B and Job Provenance index server provide web-service interfaces for querying. Those interfaces comply with the On-demand producer specification of the R-GMA infrastructure. Hence R-GMA capabilities can be utilized to perform complex distributed queries across multiple servers. Also, aggregate information about job collections can be easily provided.

The L&B service was deployed in the EU DataGrid and Cern LCG projects, the Job Provenance will be deployed in the EGEE project.

Authors: GUARISE, A. (INFN Torino); GIANELLE, A. (INFN Padova, Italy); KRENEK, A. (CESNET, CZECH REPUBLIC); MARASCHINI, A. (DATAMAT); TERRACINA, A. (DATAMAT, Italy); Mr KOURIL, D. (CESNET, CZECH REPUBLIC); REBATTO, D. (INFN Milano, Italy); RONCHIERI, E. (INFN Cnaf, Italy); GIACOMINI, F. (INFN Cnaf); PACINI, F. (DATAMAT, Italy); PRELZ, F. (INFN Milano, Italy); AVELLINO, G. (DATAMAT, Italy); PATANIA, G. (INFN Torino, Italy); Mr POSPISIL, J. (CESNET, CZECH REPUBLIC); SITERA, J. (CESNET, CZECH REPUBLIC); SKRABAL, J. (CESNET, CZECH REPUBLIC); MATYSKA, L. (CESNET, CZECH REPUBLIC); ZANGRANDO, L. (INFN Padova, Italy); MARCHI, M. (INFN Milano); MEZZADRI, M. (INFN Milano); MORDACCHINI, M. (INFN Padova); MULAC, M. (CESNET); PAPPALARDO, M. (INFN Catania, Italy); Mr RUDA, M. (CESNET, CZECH REPUBLIC); SGARAVATTO, M. (INFN Padova, Italy); VOCU, M. (CESNET, CZECH REPUBLIC); ANDREETTO, P. (INFN Padova, Italy); ANDREOZZI, S. (INFN Cnaf, Italy); BECO, S. (DATAMAT, Italy); BORGIA, S. (INFN Padova, Italy); MONFORTE, S. (INFN Catania); CIASCHINI, V. (INFN Cnaf, Italy); Mr SALVET, Z. (CESNET, CZECH REPUBLIC)

Presenter: MATYSKA, L. (CESNET, CZECH REPUBLIC)

Session Classification: Poster Session 2

Track Classification: Track 4 - Distributed Computing Services