

Association Rule Mining on Grid Monitoring Data to Detect Error Sources

Monday 23 March 2009 08:00 (20 minutes)

Grid computing is associated with a complex, large scale, heterogeneous and distributed environment. The combination of different Grid infrastructures, middleware implementations, and job submission tools into one reliable production system is a challenging task. Given the impracticability to provide an absolutely fail-safe system, strong error reporting and handling is a crucial part of operating these infrastructures.

There are various monitoring systems in place, which are also able to deliver error codes of failed Grid jobs. Nevertheless, the error codes do not always denote the actual source of the error. Instead, a more sophisticated methodology is required to locate problematic Grid elements. In our contribution we propose to mine Grid monitoring data using association rules. With this approach we are able to produce additional knowledge about the Grid elements' behavior by taking correlations and dependencies between the characteristics of failed Grid jobs into account. This technique finds error patterns - expressed as rules - automatically and fast, which helps tracing back errors to their origin. Therewith a significant decrease in time for fault recovery and fault removal is achieved, yielding an improvement of a Grid's reliability. This work presents the results of investigations on association rule mining algorithms and evaluation methods to find the best rules with respect to monitoring data in a Grid infrastructure.

Authors: Ms MAIER, Gerhild (Johannes Kepler Universität Linz); Dr SCHIFFERS, Michael (Ludwig-Maximilians-Universität München)

Co-authors: Dr GAIDIOZ, Benjamin (CERN); Prof. KRANZLMÜLLER, Dieter (Ludwig-Maximilians-Universität München)

Presenter: Ms MAIER, Gerhild (Johannes Kepler Universität Linz)

Session Classification: Poster session

Track Classification: Distributed Processing and Analysis