

APEL: An implementation of Grid accounting using R-GMA

Rob Byrom^a, Roney Cordenonsi^b, Linda Cornwall^a, Martin Craig^a,
Abdeslem Djaoui^a, Alastair Duncan^a, Steve Fisher^a, John Gordon^a, Steve Hicks^a, Dave Kant^a,
Jason Leake^c, Robin Middleton^a, Matt Thorpe^a, John Walk^a, Antony Wilson^a

^a Rutherford Appleton Laboratory, UK

^b Queen Mary, University of London

^c Objective Engineering Limited

Abstract

This article describes the implementation of an accounting tool in the LHC Computing Grid (LCG): a distributed computing grid project consisting of over 100 resource centers and more than 10,000 CPUs. APEL (Accounting Processor for Event Logs) parses batch, system and gatekeeper logs generated by a site and builds accounting records, which provide a summary of the resources consumed based on attributes such as CPU time, Wall Clock Time, Memory and grid user DN. The accounting data is published into the R-GMA information and monitoring system, and archived for processing by the graphical front-end utilised by the accounting web tool.

1. LCG Accounting Overview

In the LCG [1] grid environment, the computing resources, the application data and the grid users belonging to Virtual Organisations (VO) are distributed. Jobs submitted by these users may be sent to computing resources where the data is located nearby, or may go to remote resources where there are available job slots in order to reduce queue times. Jobs that run on LCG resources must be properly accounted for, so that the resources consumed by VOs and the resources provided by sites (resource centres), as a function of time can be determined.

In this paper we distinguish between two flavours of grid accounting: “Grid Job accounting” in which a job usage record provides a complete description of resource consumption and, “Real-time Grid accounting”, which is based on an incremental determination of resource value while the job is being executed. In the former, the cost of computing is determined on the basis of what was done (after job execution), whilst in the latter, data feeds into a pricing model on which a cost for consumption is determined in real time.

An implementation of real-time accounting (DGAS [2]) is under development within the gLite [3] middleware framework of the EGEE [4] grid-computing project.

Below we summarise the implementation of a web-based job accounting tool that collects usage records from distributed resources using R-GMA [5] to a central point, and consolidates the information for presentation on the web.

2. Data Collection Mechanism and R-GMA

The collection of accounting usage records is done through R-GMA, an implementation of the Grid Monitoring Architecture (GMA) proposed by the Global Grid Forum (GGF). GMA models the information and monitoring system infrastructure of a Grid as a set of Consumers (which request information), Producers (which provides information) and, a registry, which mediates the communication between the producers and consumers.

In LCG accounting, each site publishes its own accounting data using an R-GMA primary producer using its locally assigned R-GMA server. The published data are automatically streamed to a centralised database via a secondary producer, in order to consolidate the data from all the participating sites.

The central database is located on the Grid Operations Centre (GOC) and is used to provide a web front end, providing a summary of resource usage across the LCG grid network.

3. APEL

Apel is a log processing application, which is used to scan batch system logs, such as PBS and LSF, to build accounting records. It is able to piece together (by means of a join) the exact resource usage with a particular grid user from finding the necessary details stored within system and gatekeeper logs.

Apel is then used to publish the generated accounting records into R-GMA whereby they are collated at the GOC using an R-GMA secondary producer, as shown in Figure 1.

Apel provides support for republishing the complete local copy of accounting records to R-GMA (in cases when the GOC was offline). It also provides a mechanism for reliable delivery using a basic integrity check to compute the number of records that were last published compared with the actual count stored on the GOC.

Each accounting record is unique and there is only one record per grid job. The records may be consolidated in different ways to provide high-level views [6] of accounting data, such as the total CPU time consumed for each LHC VO at the tier-1 computing centres, as shown in Figure 2.

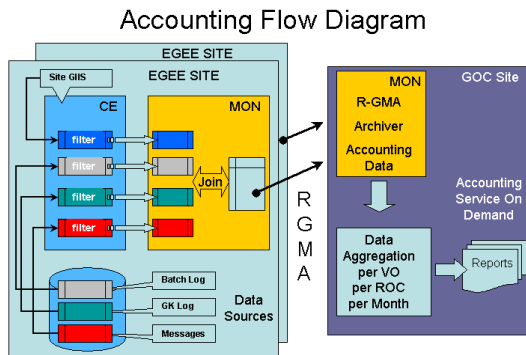


Figure 1: Accounting flow diagram providing a global overview of the data collection process (APEL), and the web reporting service.

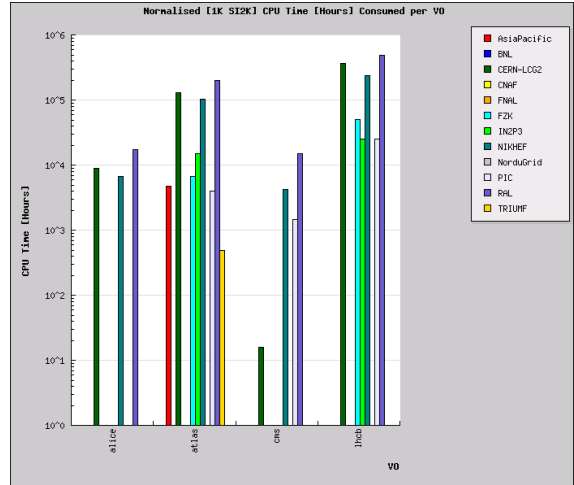


Figure 2: The total CPU time (Hours) provided by Tier-1 computing centres supporting LHC VOs during the latter part of 2004. Accounting data is gathered from sites running the LCG-2 middleware suite which currently excludes a few sites (e.g. BNL, FNAL and NorduGrid).

References



- [1] <http://lcg.web.cern.ch/LCG/>
- [2] <http://egee-jra1-wm.mi.infn.it/egee-jra1-wm/org.glite.dgas.shtml>
- [3] <http://glite.web.cern.ch/glite/>
- [4] <http://public.eu-egee.org/>
- [5] <http://www.r-gma.org/index.html>
- [6] <http://goc.grid-support.ac.uk/gridsite/accounting/>