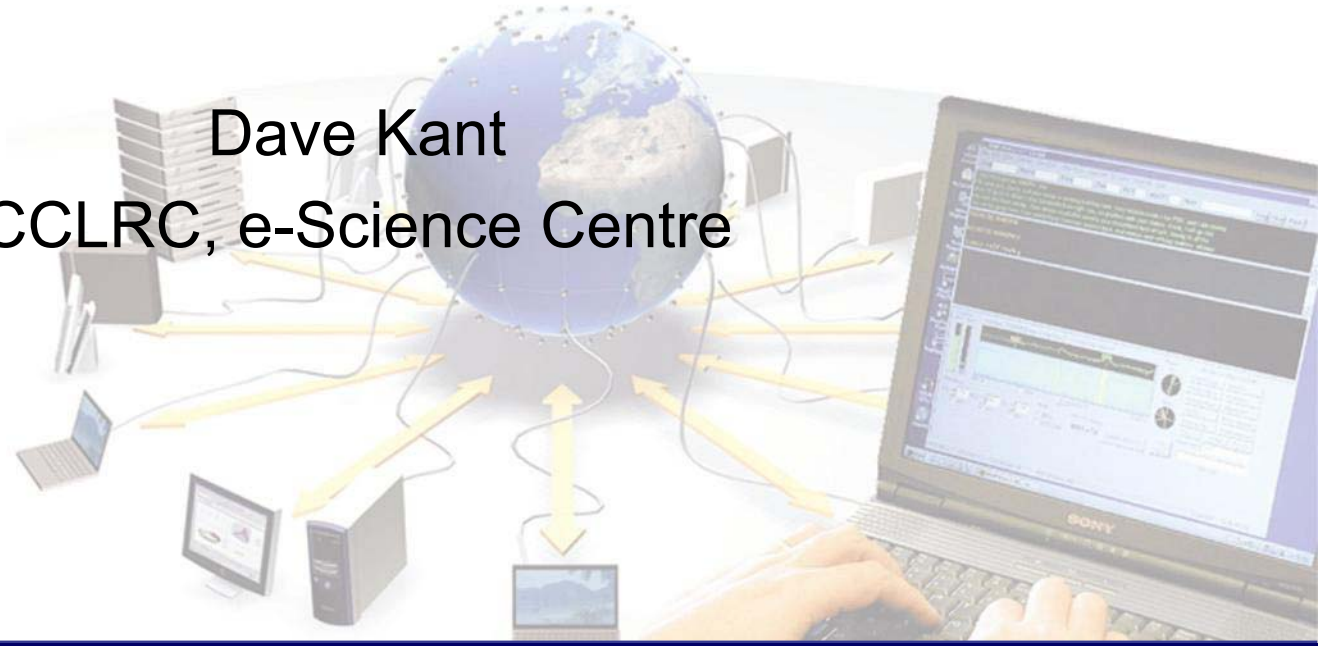
Enabling Grids for E-science in Europe

DSA1.3
Accounting Update

Dave Kant
CCLRC, e-Science Centre



Overview

Enabling Grids for E-science in Europe

This is a summary of the status of DSA1.3 Accounting & Reporting following its deployment in LCG2_3_0

1. Overview
2. Installation Problems
3. Issues
4. Future Plans

Requirement Capture

Enabling Grids for E-science in Europe

- Originally a requirement of the LHC Computing Grid project.
- Requirements were originally captured through presentations to
  - LCG's Grid Deployment Board
  - Deployment Team.
  - LHC experiments and the Tier1 centres are represented on the GDB.

Requirements

- A historical record of grid usage to identify the use of individual sites by VOs as a function of time
- To demonstrate the total delivery of resources by that site to the Grid
- Aggregated views of the collected data by:
  - VO
  - Country – a requirement of LCG which has a country-based structure
  - EGEE Region – for use by EGEE Regional Operations Centre (ROC)
- A presentation front-end to the data to allow the selection on-demand of the views described above for different VOs and periods of time.
- To present the data as
  - A graphical view for interpretation
  - A tabular view for precision
- To support sites that already had their own methods of data collection by allowing arbitrary data collection techniques and insertion of the data in the standard schema into the central database.

Requirements

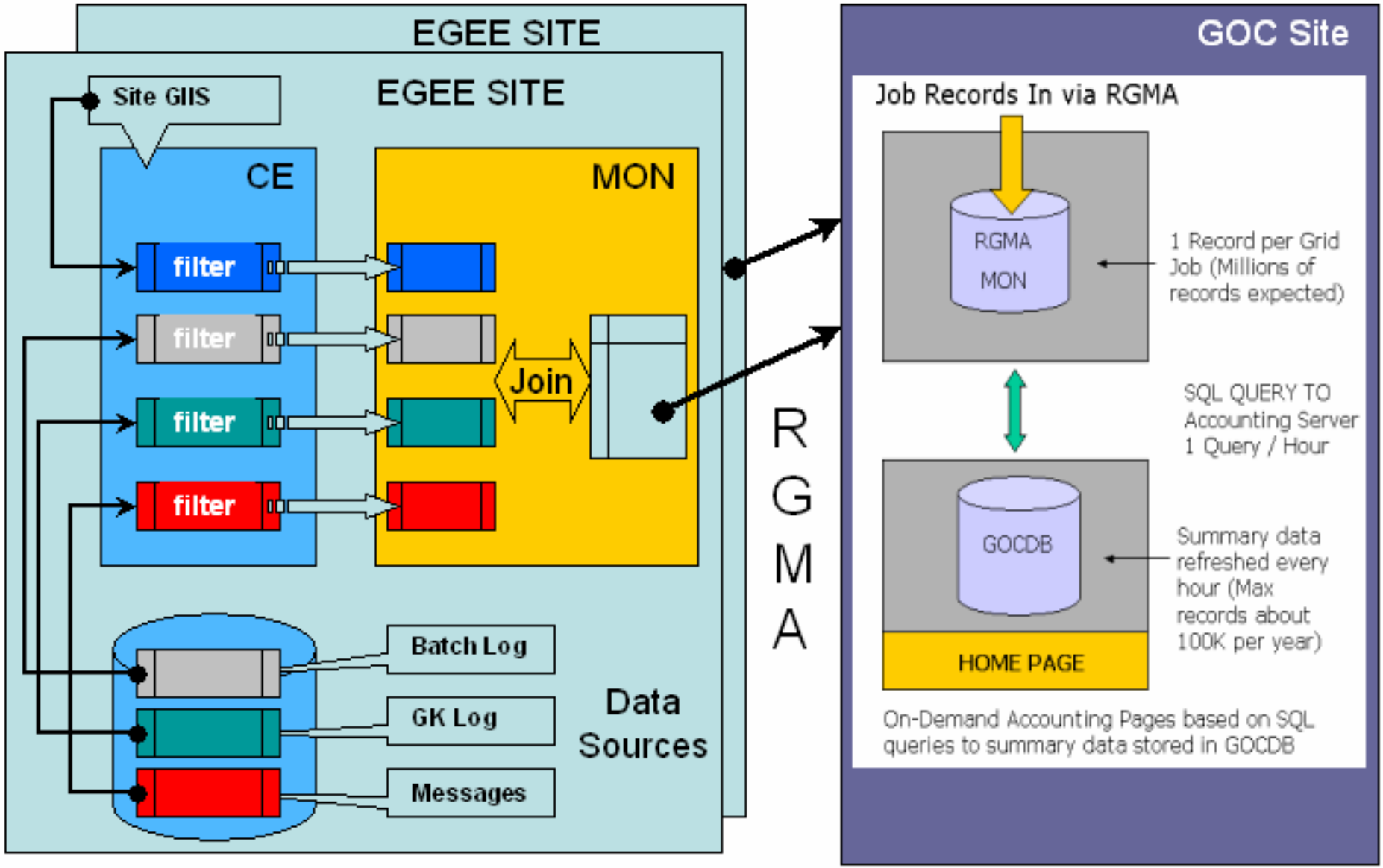
Enabling Grids for E-science in Europe

- It was not an explicit requirement that user information be captured but we included this in the design as we were sure this would be a secondary requirement
- This is a reporting system, **not** a charging mechanism.
- The information is under the control of the site, so it does not meet the requirement of a charging system to be digitally signed and irrefutable.
- Information is gathered centrally, not under the control of the VO

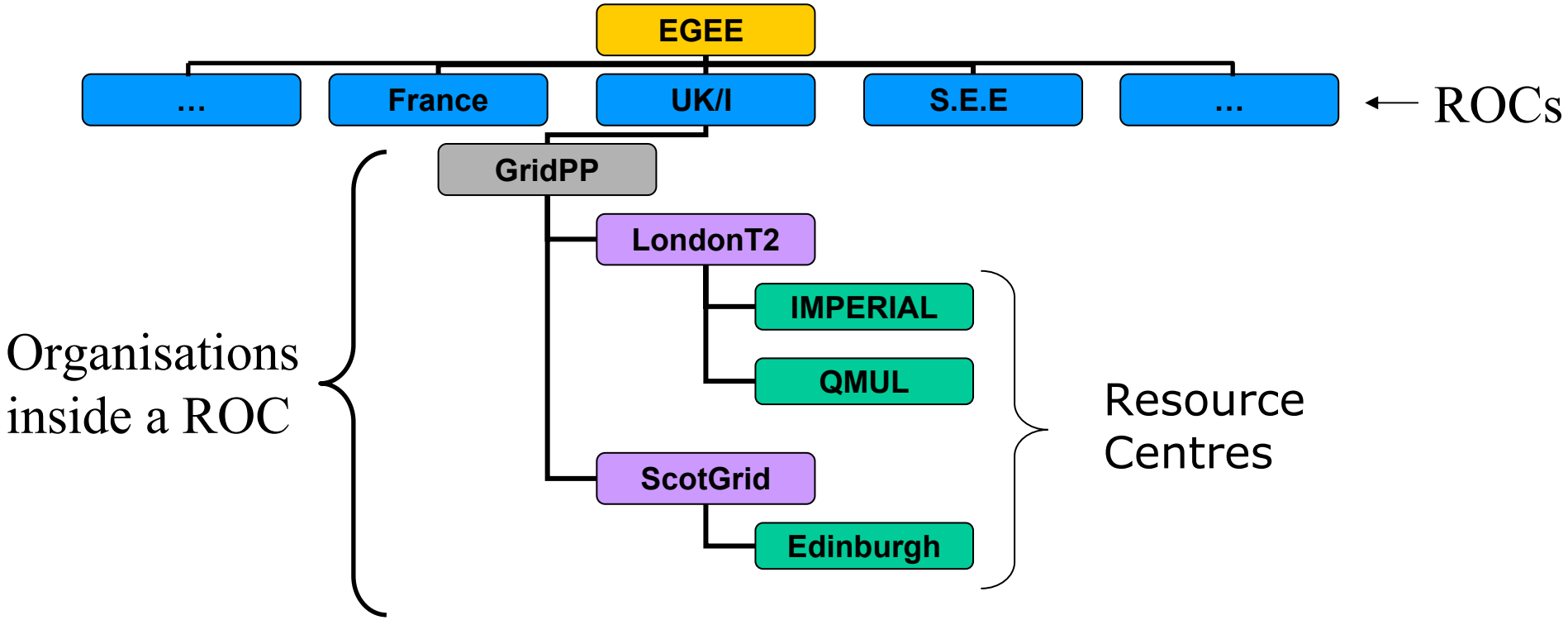
Design

- Information collected at each site from batch logs, gatekeeper logs etc
- Information joined at site level to select grid jobs and stored in database on R-GMA MON box at site.
- Information published through R-GMA and collected centrally in an R-GMA archive at GOC
- Web site presents various views of this data for presentation
- Structure of Grid taken from GOC DB – the grid configuration database.
- Only normalised cpu time collected

Accounting Flow Diagram



EGEE Organisational Structure



GOC

Job Records In via RGMA

RGMA

MON

1 Record per Grid Job (Millions of records expected)

SQL QUERY TO Accounting Server
1 Query / Hour

User queries

Graphs

Home Page

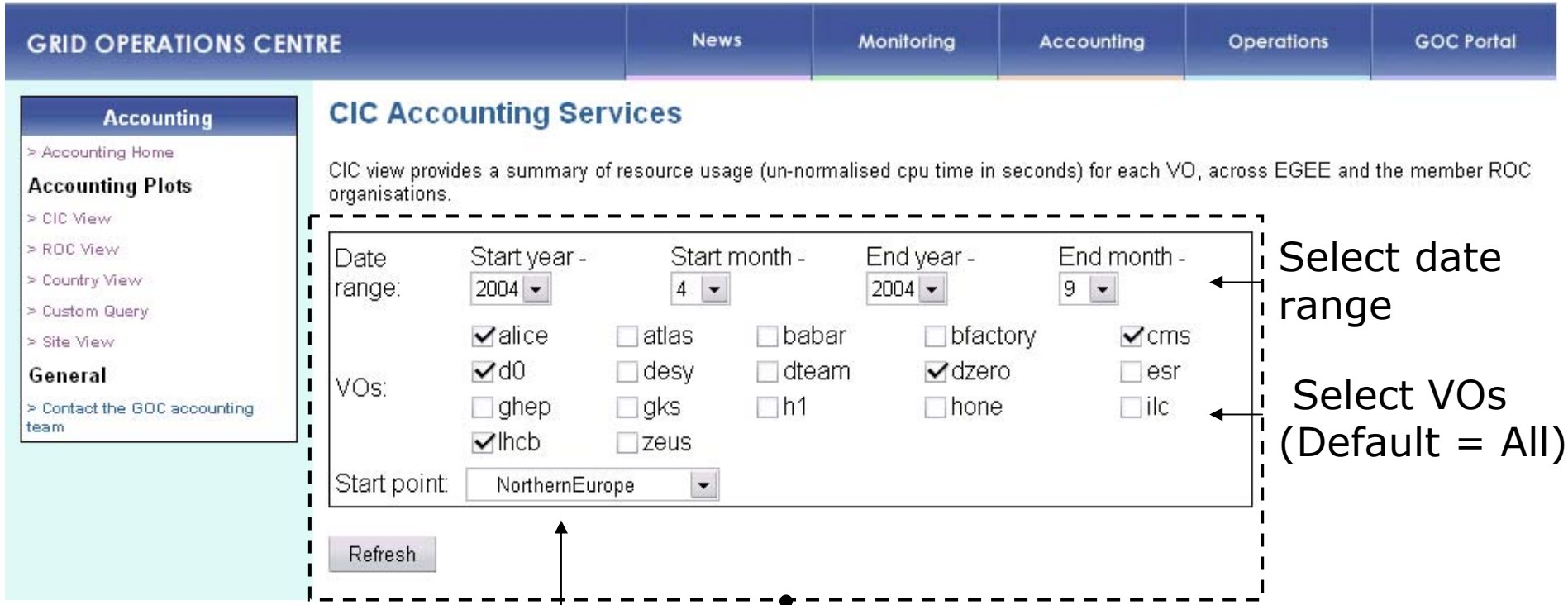
Summary data refreshed every hour (Max records about 100K per year)

On-Demand Accounting Pages based on SQL queries to summary data

Description

Enabling Grids for E-science in Europe

- Web allows information to be selected by
  - VO, time range, (Whole Grid, Country, EGEE Region, site)
- Also shows information on data collected



Select date range

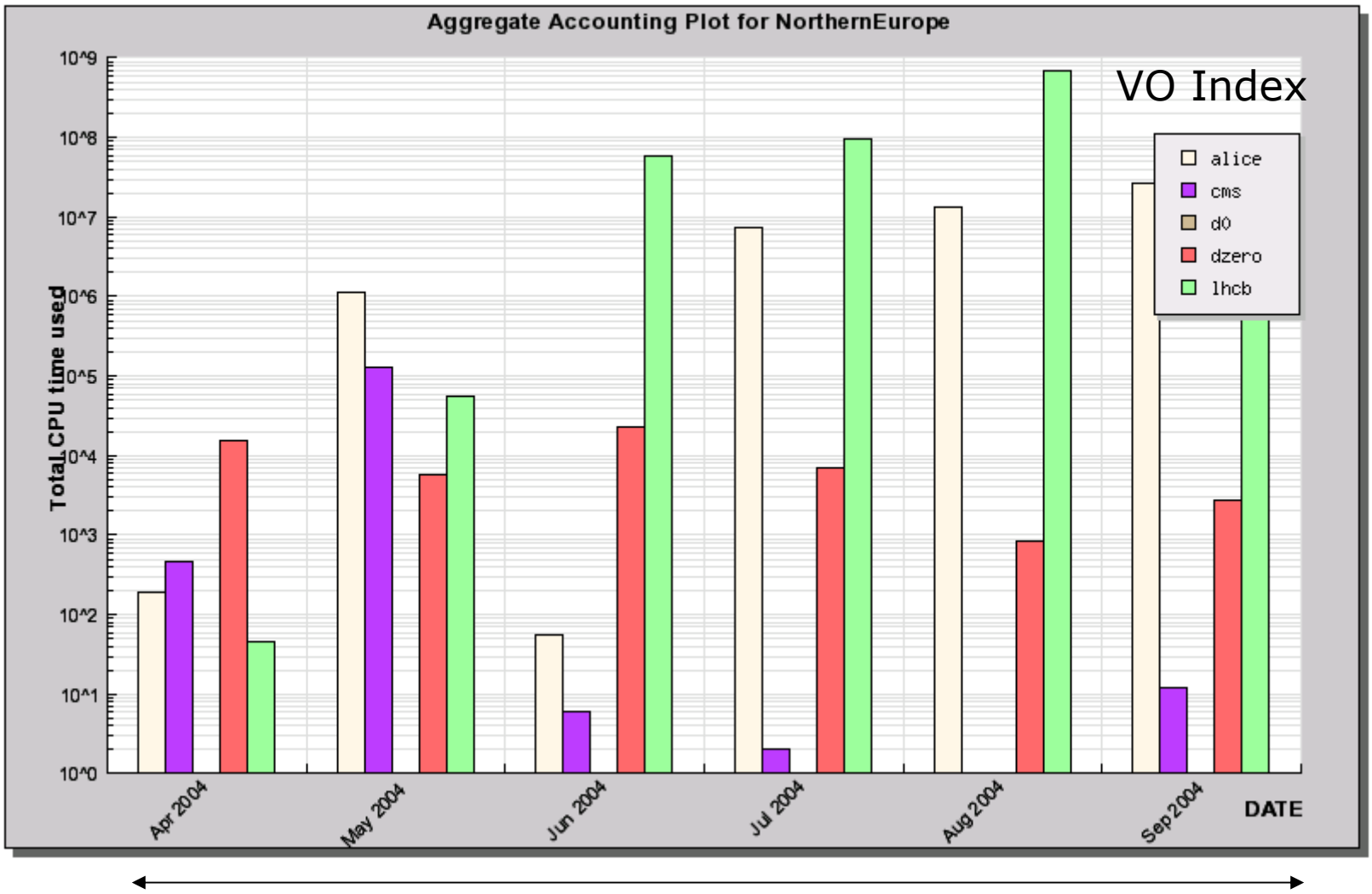
Select VOs (Default = All)

Aggregate data across an organisation structure (Default= All ROCs)

Web form to apply selection criteria on the data

- **Summed CPU (Seconds) consumed by resources in selected Region**

Accounting plot:



Selected Date Range

- List of Sites Belonging to the Selected ROC

Results for HPC2N

Click to zoom to HPC2N

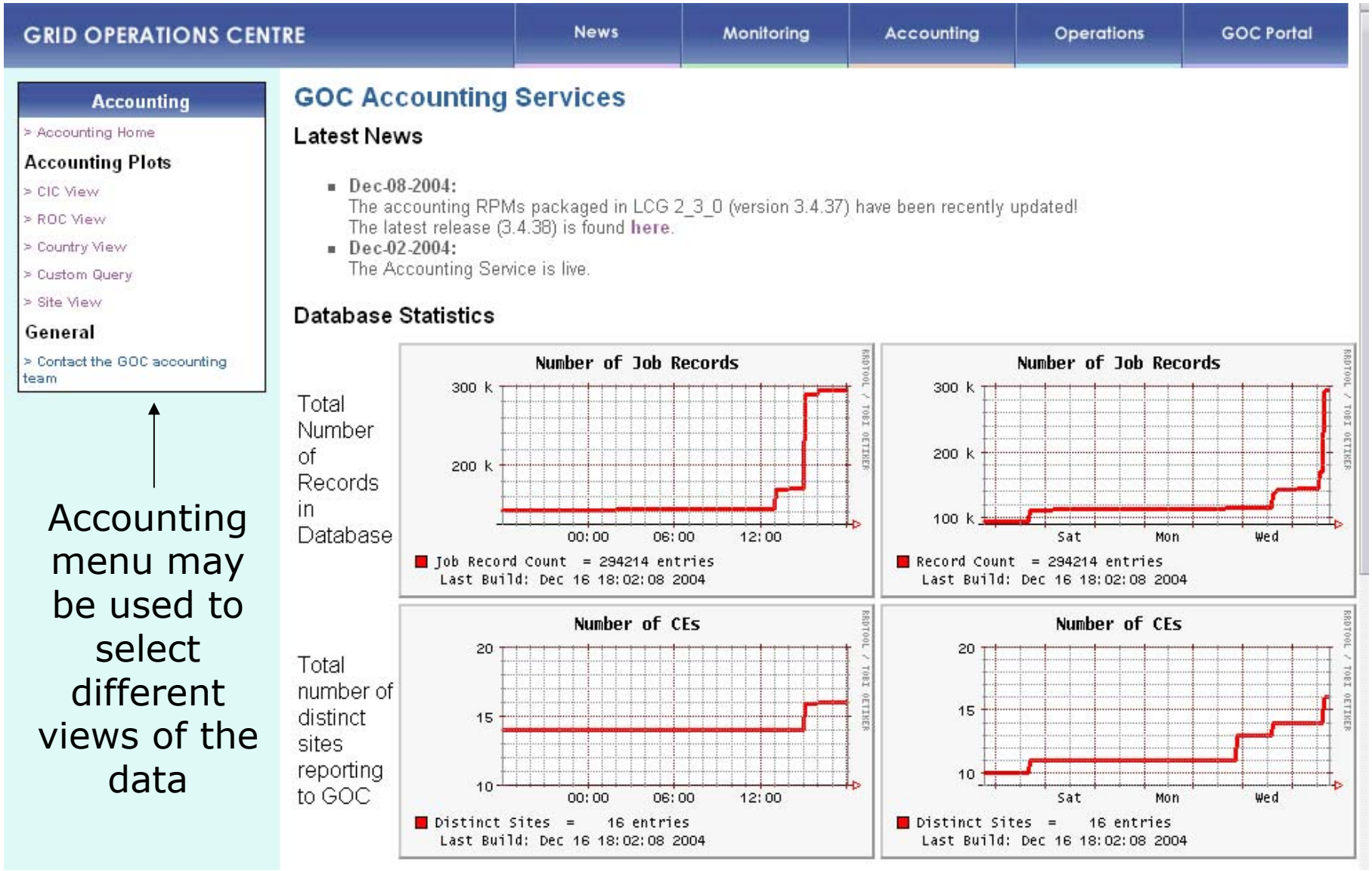
	Apr 2004	May 2004	Jun 2004	Jul 2004	Aug 2004	Sep 2004
alice	0	0	0	0	0	0
cms	0	0	0	0	0	0
d0	0	0	0	0	0	0
dzero	0	0	0	0	0	0
lhcb	0	0	0	0	0	0

Results for NIKHEF.NL

Click to zoom to NIKHEF.NL

	Apr 2004	May 2004	Jun 2004	Jul 2004	Aug 2004	Sep 2004
alice	196	1121109	55	7382501	13510882	26549930
cms	474	128542	6	2	0	12
d0	0	0	0	0	0	0
dzero	15531	5685	23080	7014	854	2812
lhcb	46	54798	58884356	95705295	689396955	64743962

A breakdown of the resource usage per Site, per VO, per Month



Accounting menu may be used to select different views of the data

Accounting Home Page displays latest news and global statistics of the accounting database

GRID OPERATIONS CENTRE | News | Monitoring | Accounting | Operations | GOC Portal

Accounting

- > Accounting Home

Accounting Plots

- >Normalised CPU
- > CIC View
- > ROC View
- > Country View
- > Custom Query
- > Site View
- > Privacy Statement

General

GOC Accounting Services

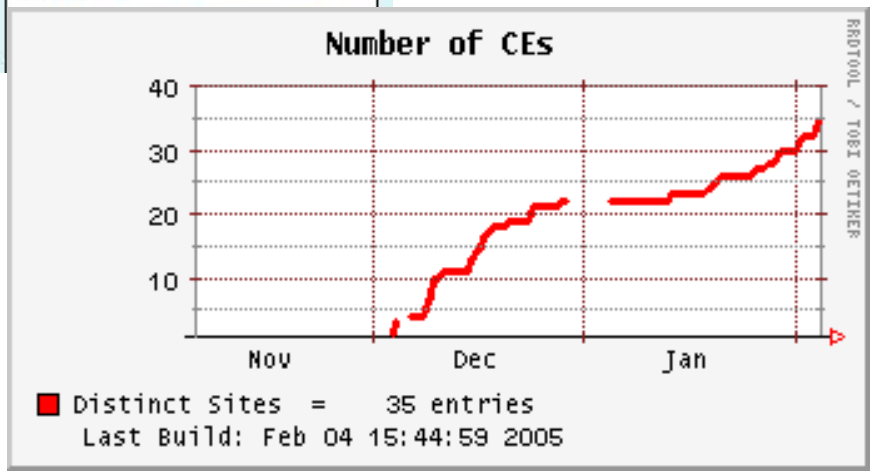
Latest News

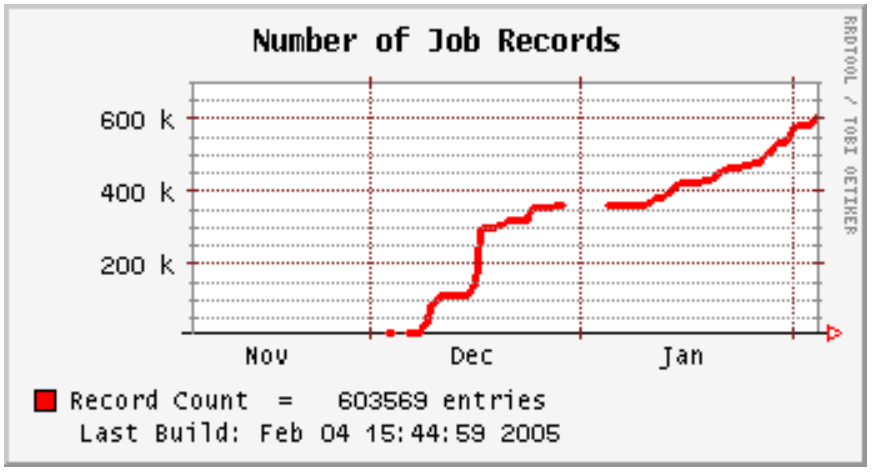
http://goc.grid-support.ac.uk/

- **Tues-18-2005:**
  We have released a new version (3-4-40) of the accounting processor which minor fixes to the derivation of SpecInt
  New Accounting RPMs Download: RPM-3.4.41 Release-Notes Installation Guide FAQ
- **Dec-02-2004:**
  The Accounting Service is live today.

Database Statistics

Total Number of Records in Database



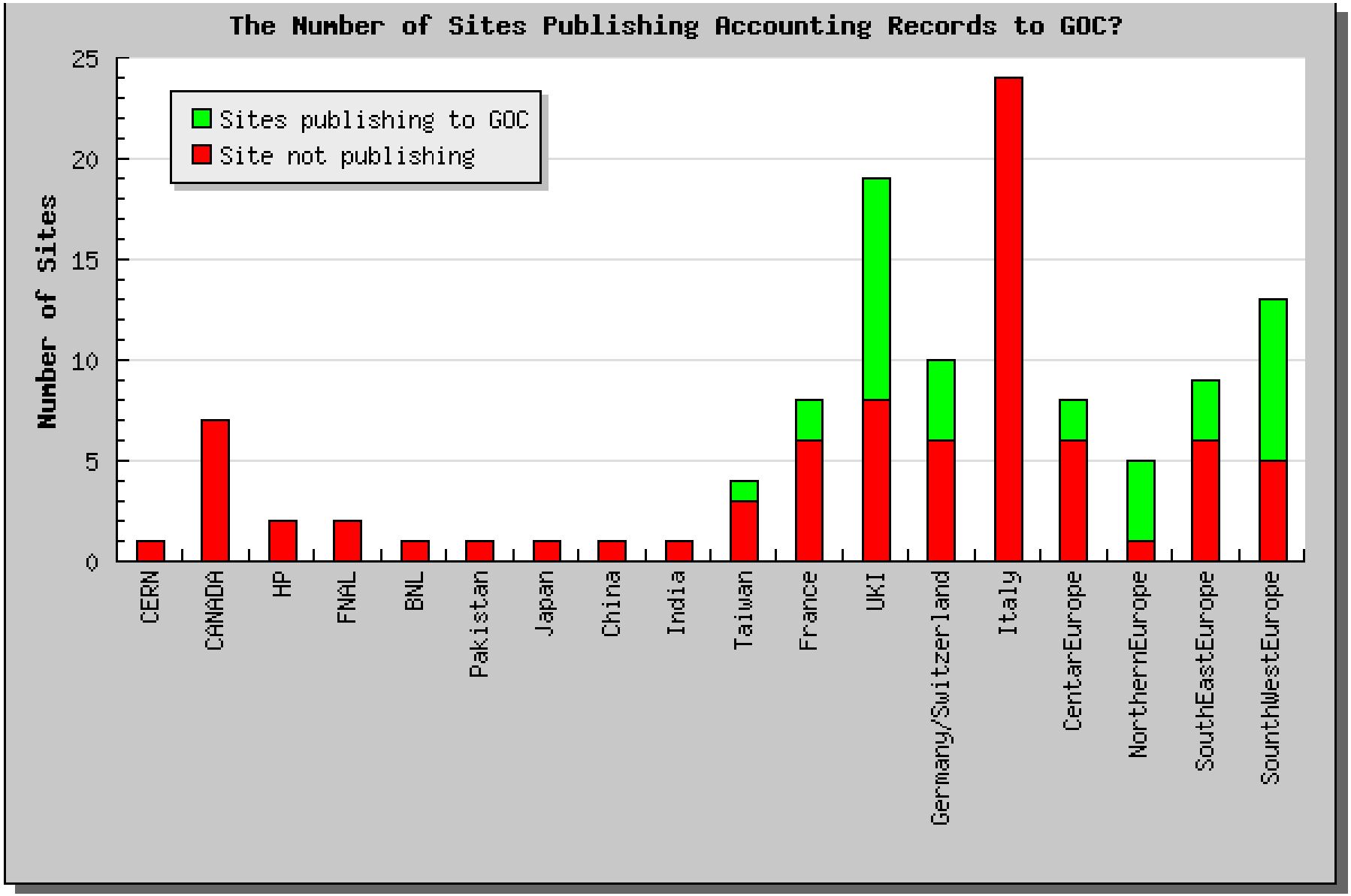


Since LCG2_3_0 Release: 35 Sites publishing data to GOC (Feb 7th 2005)

Over 0.6 Million Job records

~ 50K records per week

Number of Sites per ROC Publishing Accounting Records to GOC



Installation Problems

Enabling Grids for E-science in Europe

1. What problems did site administrators have?
    1. Configuration script to setup the accounting database was not executed.
    2. Didn't run the tests on their local system to check APEL
    3. Unclear about administering a MySQL database.
    4. Some had no problems with the configuration process.

2. What did we do about it?
    1. Provided FAQ to help with the setup procedure.
    2. Talk to site administrators, walk them through the procedure.
    3. Update FAQ, improve documentation.

Issues

Enabling Grids for E-science in Europe

1. Which RPM Version?
   - Apel 3.4.37 included in 2_3_0 release but configuration problems with the database setup script.
   - Testing phase missed this
     - What can we do to improve quality testing?
     - Provide guidelines on “fresh-install” and “upgrade” scenarios
   - Latest version 3.4.41 submitted as a patch via Savanna (# 317)
   - Change Log 3.4.37 to 3.4.41
     - Apel 3.4.41 (Feb 2nd) Minor fix to SQL script
     - Apel 3.4.40 (Jan 17th)
       - Normalisation issue (see later)
       - CatchAll specInt/specFloat set to value in GIIS rather than 0
     - Apel 3.4.39 (Dec 16th) Current PBS log excluded from archive
     - Apel 3.4.38 (Nov 19th)
       - Bug in “reprocess” option during Join
       - Added “cleanAll” option
     - Apel 3.4.37 (Oct 14th)
       - grant mechanism to allow GK and CE to connect to MySQL database

Issues

Enabling Grids for E-science in Europe

2. VO Filtering
   - National Grid VOs activities running on same infrastructure as EGEE/LCG
   - Privacy reasons why sites don't want to publish National VO data to GOC
   - APEL does not discriminate between the VOs
   - Develop a solution? What can we suggest today?
   - GOCDB can hide resources
   - APEL made the requirement to exclude Local work not published but non LCG work does come through.
   - Whats the model 1 CE per VO…what do people do?
     - Don't need to install Apel on non-LCG VO CEs
   - SARA-LCG2, IISAS-Bratislava

Issues

3. Development of Tests to Check the Accounting Service
   - Is site accounting working?
   - Is the GOC listening for new data?
   - Is the RGMA Registry working?

- GSTAT
  - GOC Flexible archiver service listens for accounting producers
  - If the service is down, no data can be sent to the GOC!
  - Use the service every 5 minutes to update a timestamp in a test record in the accounting database. GSTAT can query table, look at the timestamp and compare with the current date/time.
  - 3rd party to use the flexi service.
  - Use RGMA to compare records in the site database and GOC
- Site Functional Tests
  - Can check the RPM version installed on the CE
- Testing the Whole Thing instead of the Pieces
  - Investigate an Apel “heart-beat”
  - Site cron writes a test record every hour and publishes to GOC

Issues

Enabling Grids for E-science in Europe

4. Which Log Files Should Site Administrators Backup?
   - To build accounting records, we need to process data from **THREE** log file sources. This is a **mandatory requirement** in order to reconstruct what has been done during the 2004 period.
   - **/var/log/globus-gatekeeper***
     - Match between grid-user dn to GramScriptJobId
   - **/var/pbs/accounting/server_priv/***
     - Local jobID and details of resources consumed
     - No distinction between grid jobs and non-grid jobs.
   - **/var/log/messages***
     - Map GramScriptJobID to local JobID
   - This is how we separate grid jobs from local user jobs which run on the local fabric.
   - If the site has deleted its messages files, we may be able to work around this by matching local unix groups in the batch logs. Accounting records formed in this way will not contain the dn of the grid-user.

ISSUES

Enabling Grids for E-science in Europe

5. siteName Changes
   - Recent problem with presenting data from the French ROC
   - CCIN2P3 renamed to IN2P3-CC
   - All records associated with the site have to be updated in order for SQL queries to match the new siteName.
   - Need to implement an automatic procedure to deal with this in the GOCDB.

6. Namespace Convention?
   - Naming scheme to identify data belonging to large sites which provide services for different communities etc.
   - NIKHEF: lcgprod.nikhef.nl , lcg2prod.nikhef.nl, edgapptb.nikhef.nl
   - *SiteName* is a bad choice because we get multiple hits
     - *IC-LCG2* gives multiple matches PIC-LCG2 and IFIC-LCG2
   - Request sites stick to the convention *.SiteName
     - h1.desy.de, zeus.desy.de

ISSUES

Enabling Grids for E-science in Europe

5. siteName Changes
   - Recent problem with presenting data from the French ROC
   - All records associated with the site have to be updated in order for SQL queries to match the new siteName.
   - Need to implement an automatic procedure to deal with this.

6. Normalisation
   - We want to perform a reasonably sensible first order estimate to account for the differences in worker node performance.
   - Homogeneous vs Heterogeneous
   - PBS Job Records don't have any information about the worker node benchmarks => insert one manually
   - PBS Farms setup in different ways; can lead to an error in the normalisation calculation (Blindman vs internal normalisation)
   - Histories - What SpecInts do we use in order to process archived Job Records?
   - LSF Job Records have a CPU_FACTOR (1 - 4) in the Job Record.
     - What does a value of 1 correspond to?
     - Different "calibration" value at each site
     - Conversion table?

Future Plans

Enabling Grids for E-science in Europe

1. Support for the LSF batch system. We see this as the next major priority
   - Already support for 5.1 has been developed;
   - Version 6 to be supported.
2. More views of data and improve UserInterface
   - Motivated by Users
   - Histograms; Pie; Gantt
   - Real/virtual memory usage per VO
   - Total CPU usage as function of time as a stacked bar (stacked according to VO)
3. Extend schema to include information about the worker node and the globalJobID.
4. Job Efficiency at Batch Level
   - Exit status of jobs
   - What happens if the RB dies and you loose the L&B?
   - PBS job continues, writes data to SE; Job not lost
5. Storage equivalent for APEL.
   - Cron runs, pubblishes into GOC
   - SE snapshot e.g “df” of filesystem
     - Use of disk and tape
     - Cron runs on SE which is a script; but script tailored for different SE e.g. Dcache, tapestore etc
6. How do we show users?
   - Not what cpu used across the whole grid
   - But within a site, provide a breakdown of what the users done on a weekly basis.
7. VO Filtering
   - Something simple that can be included in next major release.