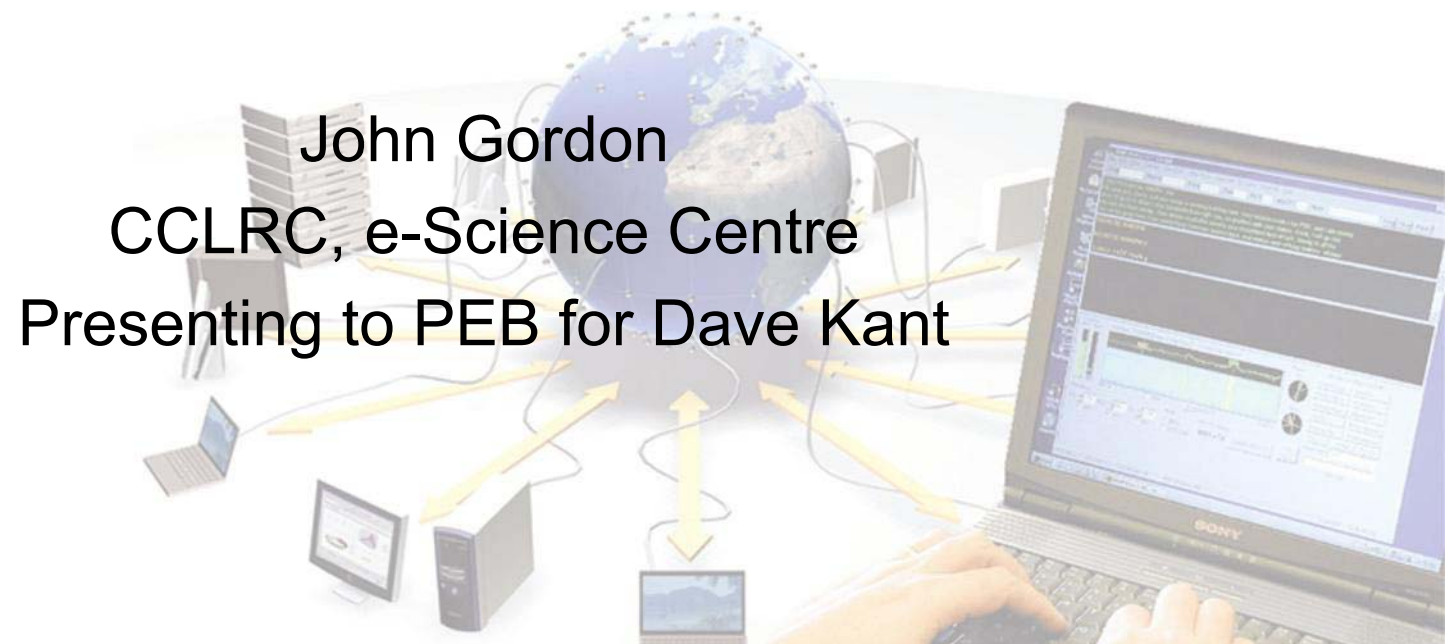




Enabling Grids for  
E-science in Europe

## DSA1.3

# Accounting and Reporting Web Site



# Overview

This deliverable is a web site, not a document.

The document describing the deliverable is

<https://edms.cern.ch/document/489455>

1. Requirements
2. Design
3. Description
4. Deployment
5. Issues
6. Future Plans

# Requirement Capture

- 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

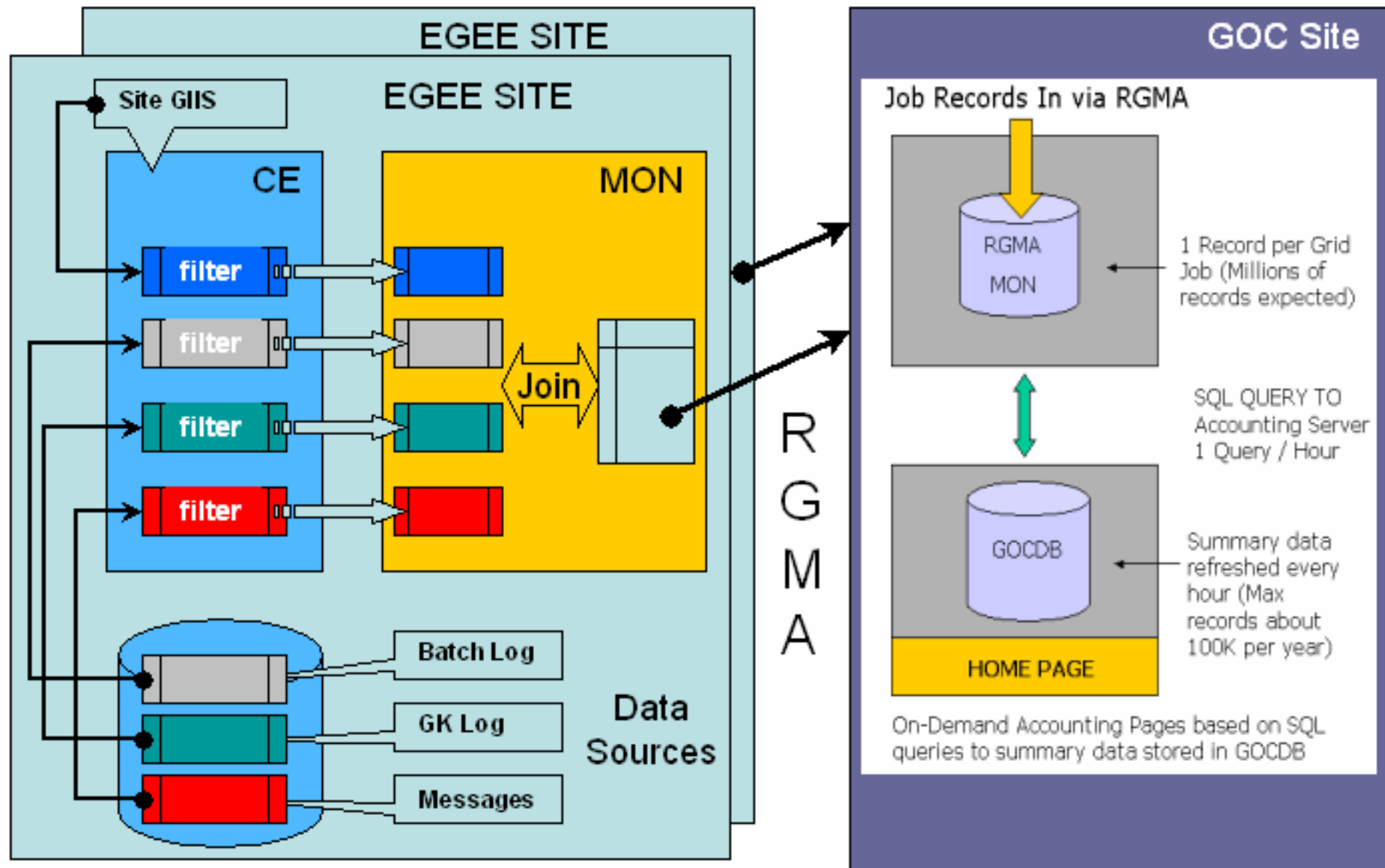
- A historical record of grid usage to identify the use of individual sites by VOs as a function of time
- To demonstrate the total delivery of resources by that site to the Grid
- Aggregated views of the collected data by:
  - VO
  - Country – a requirement of LCG which has a country-based structure
  - 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

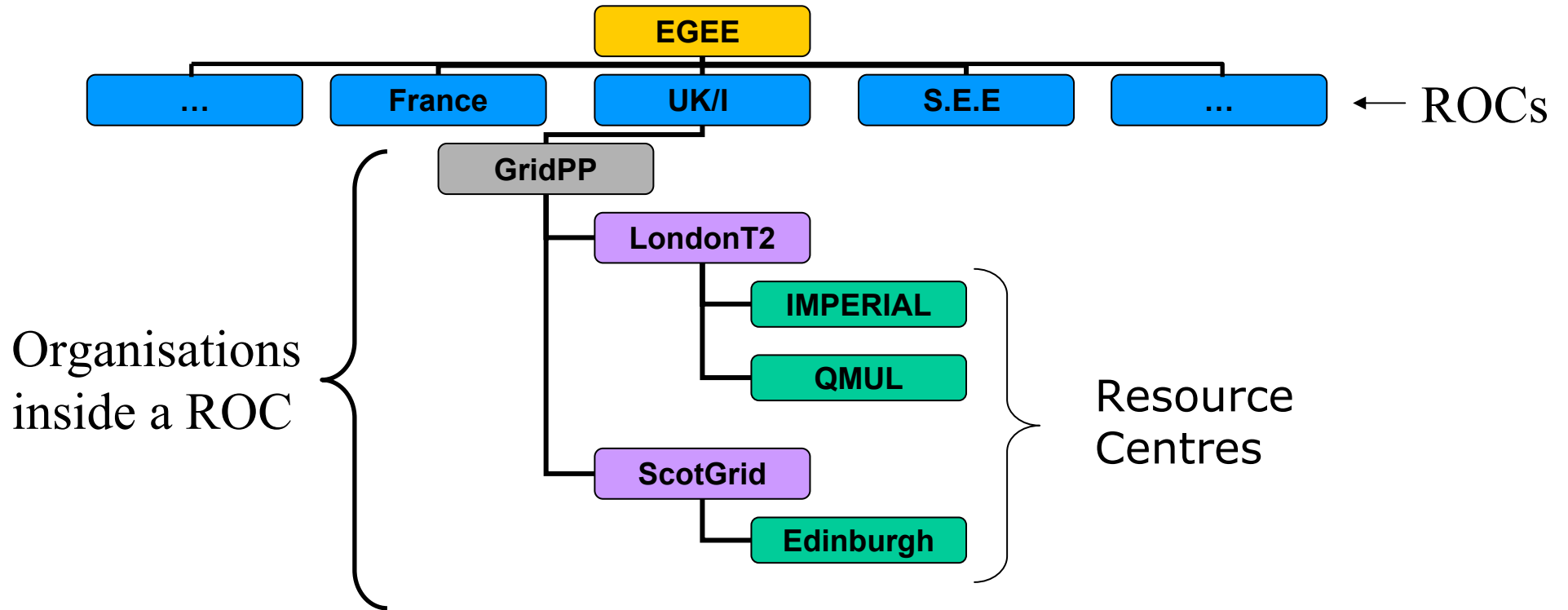
- It was not an explicit requirement that user information be captured but we included this in the design as we were sure this would be a secondary requirement
- This is a reporting system, **not** a charging mechanism.
- The information is under the control of the site, so it does not meet the requirement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

- 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
  
- Information schema from GGF Usage Record
- Structure of Grid taken from GOC DB – the grid configuration database.
- Only normalised cpu time collected

# Accounting Flow Diagram

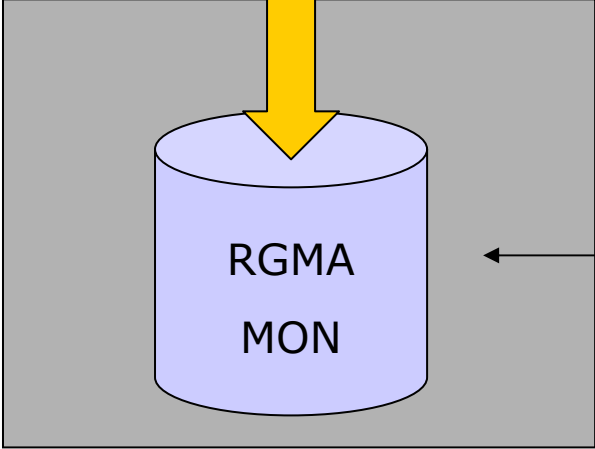


# EGEE Organisational Structure

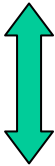




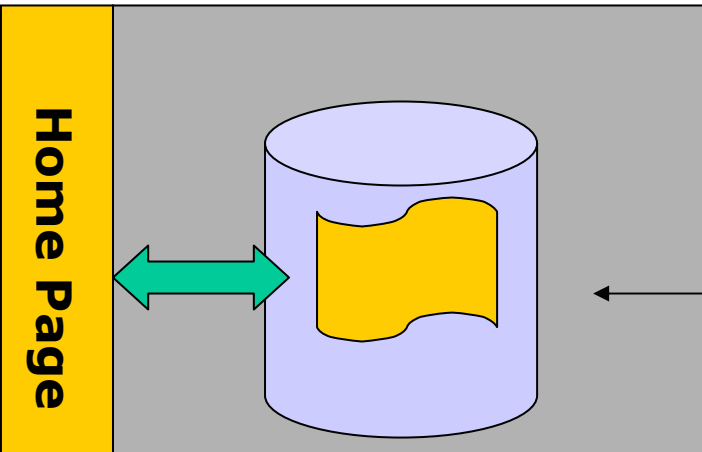
### Job Records In via RGMA



1 Record per Grid Job (Millions of records expected)



SQL QUERY TO Accounting Server  
1 Query / Hour



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



User queries

Graphs

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

# Description



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

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

**Accounting**

- > Accounting Home
- Accounting Plots**
- > CIC View
- > ROC View
- > Country View
- > Custom Query
- > Site View
- General**
- > Contact the GOC accounting team

**CIC Accounting Services**

CIC view provides a summary of resource usage (un-normalised cpu time in seconds) for each VO, across EGEE and the member ROC organisations.

Date range: Start year - 2004 Start month - 4 End year - 2004 End month - 9

VOs:  alice  atlas  babar  bfactory  cms  
 d0  desy  dteam  dzero  esr  
 ghep  gks  h1  hone  ilc  
 lhcb  zeus

Start point: NorthernEurope

Refresh

Select date range

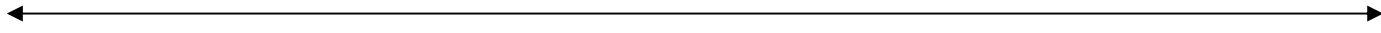
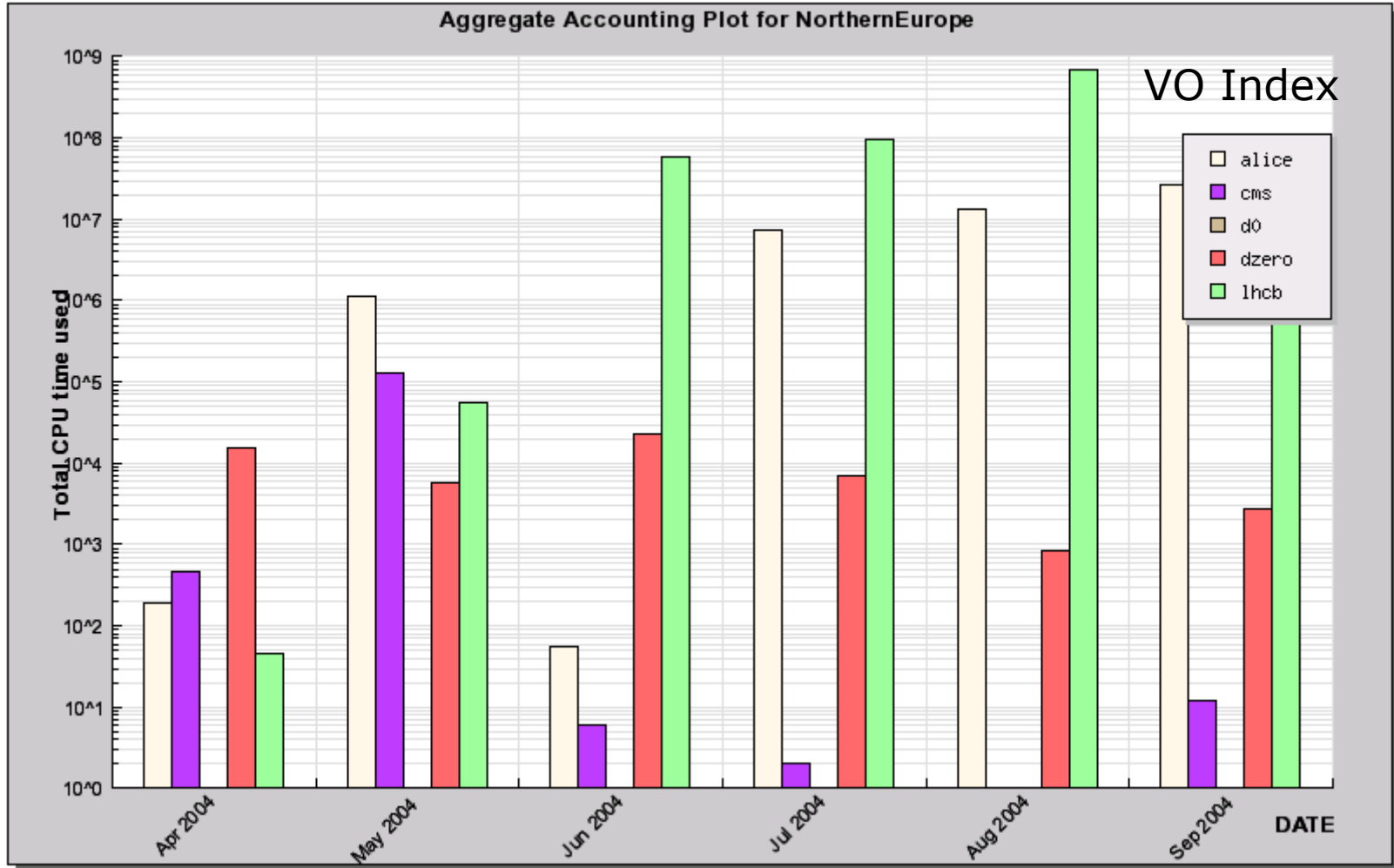
Select VOs (Default = All)

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	
	<b>alice</b>	0	0	0	0	0	0
	<b>cms</b>	0	0	0	0	0	0
	<b>d0</b>	0	0	0	0	0	0
	<b>dzero</b>	0	0	0	0	0	0
	<b>lhcb</b>	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	
	<b>alice</b>	196	1121109	55	7382501	13510882	26549930
	<b>cms</b>	474	128542	6	2	0	12
	<b>d0</b>	0	0	0	0	0	0
	<b>dzero</b>	15531	5685	23080	7014	854	2812
	<b>lhcb</b>	46	54798	58884356	95705295	689396955	64743962

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

# Deployment

1. Package was released to LCG in August 2004 and certified soon afterwards.
2. There was no LCG release after that until LCG2\_3\_0 on ??th December 2004
3. Sites successively running R-GMA in 2\_2\_0 were approached to install Accounting manually. Today there are still very few 2\_3\_0 sites. There are 22 sites producing accounting records today.
4. A few of them are historic (ie CE has been replaced and both old and new ones appear).

Accounting

> Accounting Home

Accounting Plots

> CIC View

> ROC View

> Country View

> Custom Query

> Site View

General

> Contact the GOC accounting team

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

GOC Accounting Services

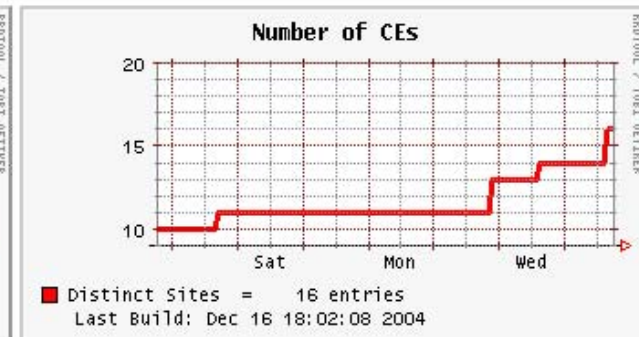
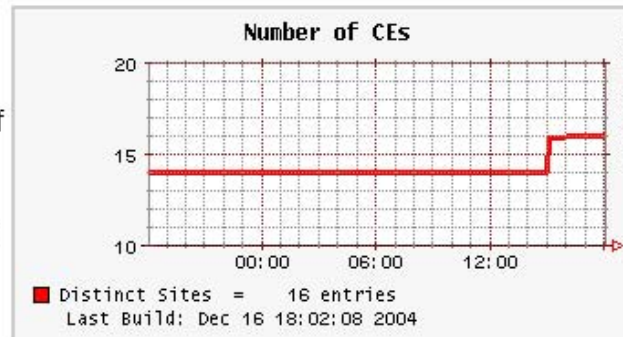
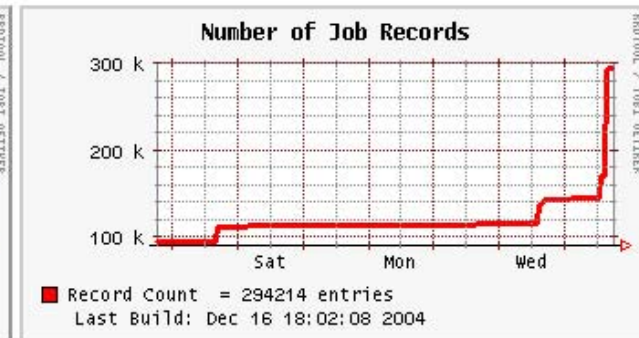
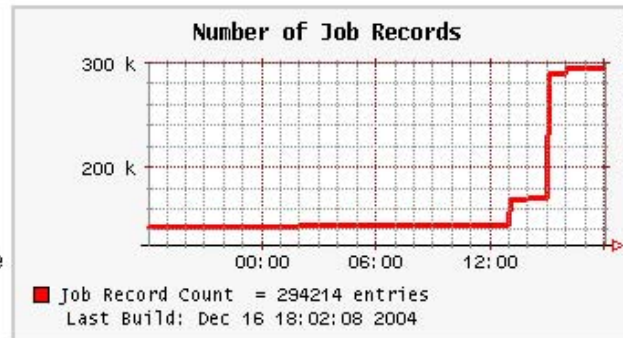
Latest News

- **Dec-08-2004:**  
The accounting RPMs packaged in LCG 2\_3\_0 (version 3.4.37) have been recently updated! The latest release (3.4.38) is found [here](#).
- **Dec-02-2004:**  
The Accounting Service is live.

Database Statistics

Total Number of Records in Database

Total number of distinct sites reporting to GOC



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

## 1. Scalability

- database can contain millions of records
- on-demand plots do not query this database but aggregated views which are updated hourly
- **Other Accounting Packages**
  - There are a variety of other packages in existence now
  - DGAS, TeraGrid, OMII(ComputationalMarkets), OSG(?)
  - All claim to use the GGF Schema so information can be aggregated/exchanged/merged (potential future project)



# Future Plans

- Support of the LSF batch system.
- More views of data
- Extend schema to include information about the worker node and the globalJobID.
- Investigate scalability and performance issues further.