Status of EGEE Production Service

Ian Bird, CERN
SA1 Activity Leader
EGEE 1st EU Review
9-11/02/2005

www.eu-egee.org

Information Society

Introduction

- **Overview of the Grid Operations Service activities (SA1, SA2) – structure, successes, issues, and plans**

- **Strategy has been to**
  - have a robust certification and testing activity,
  - simplify as far as possible what is deployed, and to make that robust and useable.
  - In parallel construct the essential infrastructure needed to operate and maintain a grid infrastructure in a sustainable way.

- **Current service based on work done in LCG – culminating in the current service ("LCG-2")**
  - Now at the point where in parallel we need to deploy and understand gLite – whilst maintaining a reliable production service.

SA1: Key points

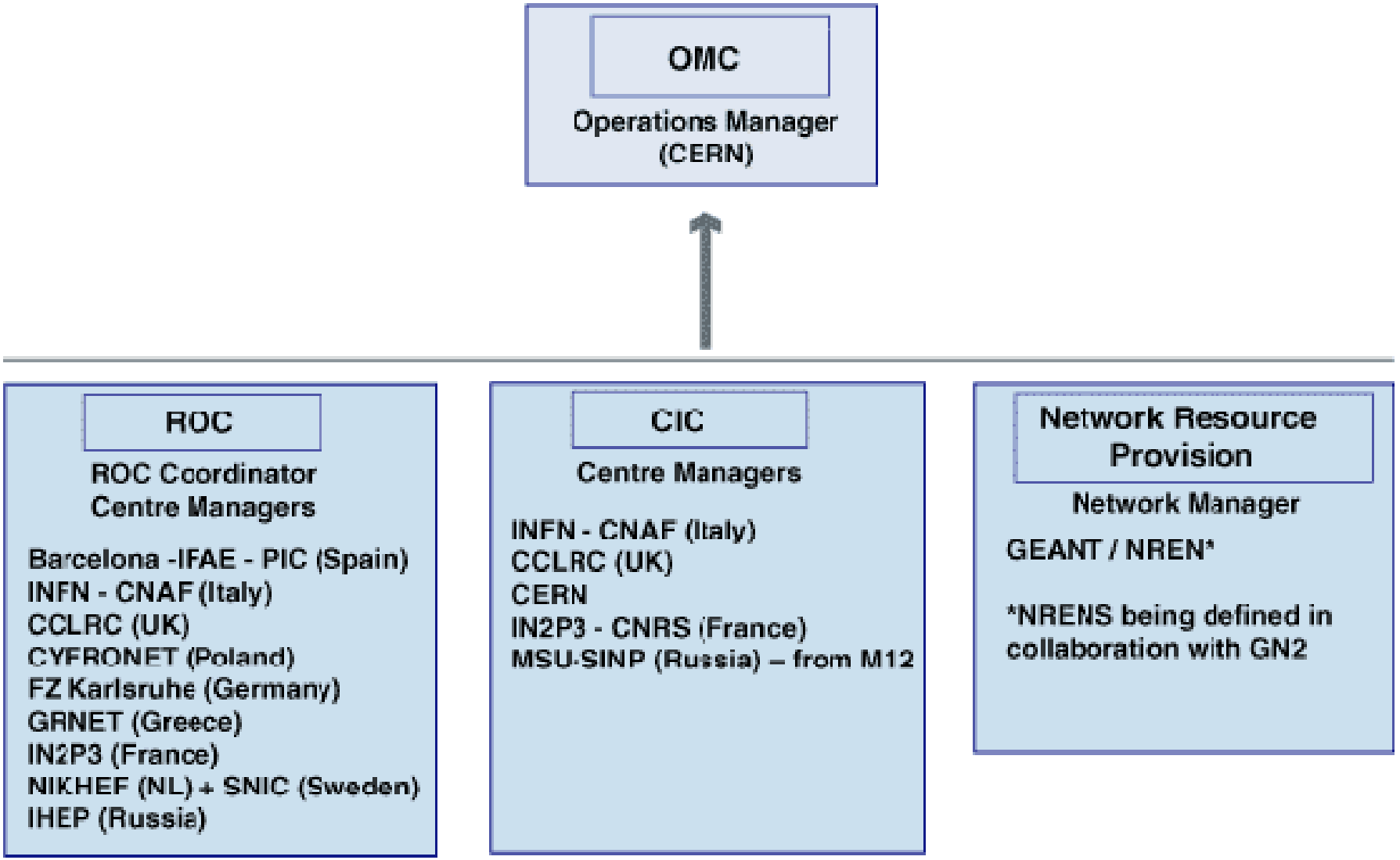
- **Successes:**
  - A large operational production grid infrastructure in place and in use
    - Managed certification and deployment process in place
      - ☞ *Markus Schulz - talk*
    - Managed grid operations process in place
      - ☞ *Hélène Cordier - demo*
  - Have supported extensive and intensive use by the LHC experiments during 2004 data challenges (10 months)
      - ☞ *NA4 talk*
  - Now has Bio-medical community using the infrastructure, and others close

- **Issues:**
  - Continue to improve the quality, reliability and efficiency of the operations
    - How to approach "24x7" global operations.
  - Continue to develop a trusted, reliable and usable user support infrastructure
  - Introducing and deploying new VOs is too heavy weight

SA1 Objectives

- **Grid Management**
  - Coordinate Regional Operations Centres (ROC) and Core Infrastructure Centres (CIC)
  - Manage the relationships with resource providers via service-level agreements
- **Grid Operation**
  - Operate core Infrastructure services:
  - Provide grid monitoring and control:
- **Support:**
  - Validate, deploy, and support middleware releases
  - Provide user support – helpdesk and call centres
  - Provide support to resource centres
- **International Collaboration**
  - Drive collaboration with peer organisations in the U.S. and in Asia-Pacific
  - Ensure interoperability of grid infrastructures and services for cross-domain VO's
  - Participate in liaison and standards bodies in wider grid community

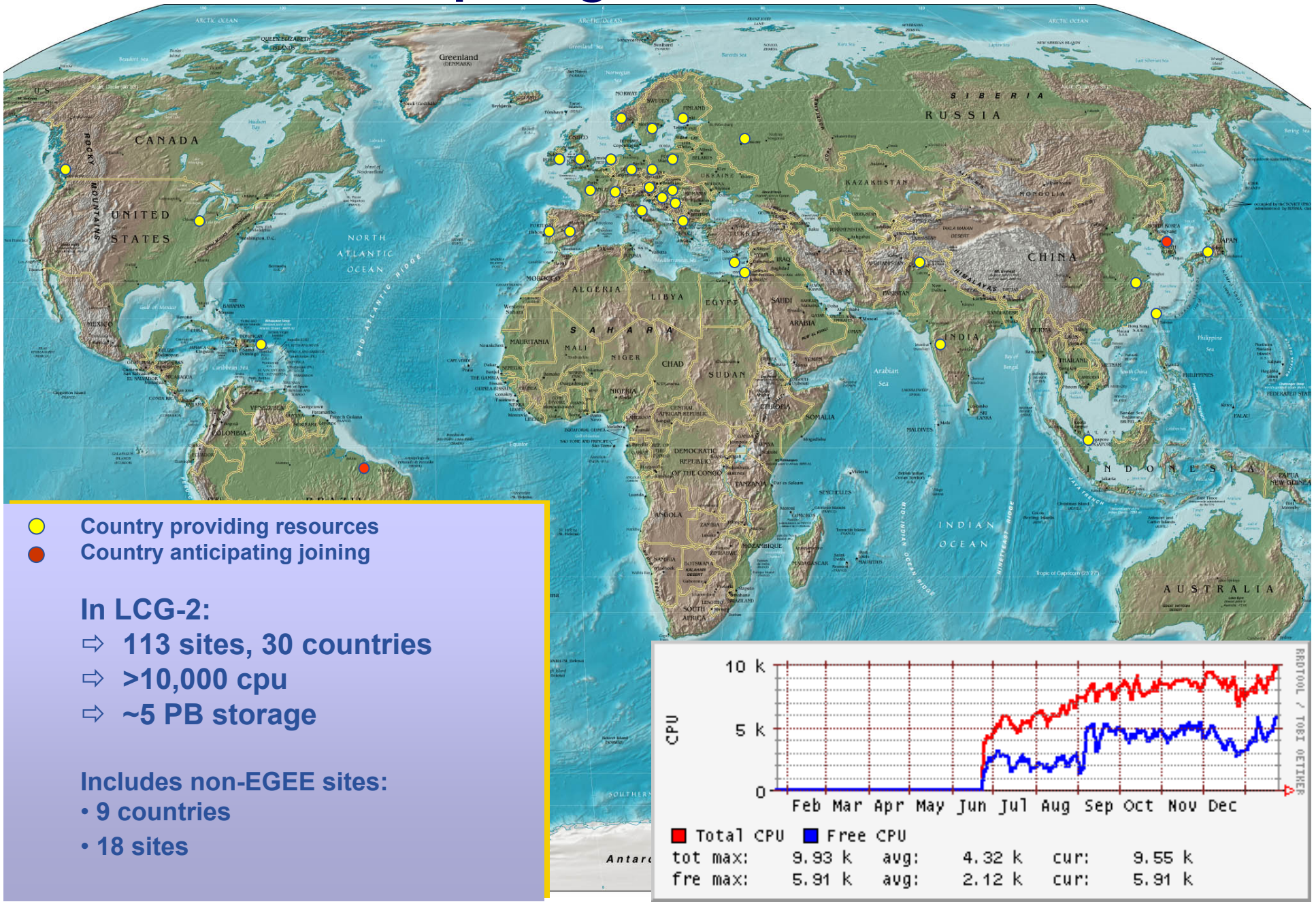
Operations (SA1, SA2) Management



Milestones & Deliverables

Month	Deliverable/ Milestone	Item	Lead
M03	DSA1.1	Detailed execution plan for first 15 months of infrastructure operation	CERN
M06	MSA1.1	Initial pilot production grid operational **10 sites**	
M06	DSA1.2	Release notes corresponding to the initial pilot Grid infrastructure operational	INFN
M09	DSA1.3	Accounting and reporting web site publicly available	CCLRC
M09	MSA1.2	First review	
M12	DSA1.4	Assessment of initial infrastructure operation and plan for next 12 months	IN2P3
M14	DSA1.5	First release of EGEE Infrastructure Planning Guide ("cook-book"),	CERN
M14	MSA1.3	Full production grid infrastructure operational **20 sites**	
M14	DSA1.6	Release notes corresponding to the full production Grid infrastructure operational	CCLRC
M18	MSA1.4	Second review	
M22	DSA1.7	Updated EGEE Infrastructure Planning Guide	CERN
M24	DSA1.8	Assessment of production infrastructure operation and outline of how sustained operation of EGEE might be addressed.	IN2P3
M24	MSA1.5	Third review and expanded production grid operational **50 sites**	
M24	DSA1.9	Release notes corresponding to expanded production Grid infrastructure operational	INFN

Computing Resources: Feb 2005



Infrastructure metrics

Region	coun-tries	sites	cpu M6 (TA)	cpu M15 (TA)	cpu actual
CERN	0	1	900	1800	942
UK/Ireland	2	19	100	2200	2398
France	1	8	400	895	886
Italy	1	20	553	679	1777
South East	5	7	146	322	133
South West	2	12	250	250	498
Central Europe	5	8	385	730	373
Northern Europe	2	4	200	2000	427
Germany/Switzerland	2	10	100	400	1207
Russia	1	6	50	152	238
EGEE-total	21	95	3084	9428	8879
USA	1	3	-	-	458
Canada	1	6	-	-	316
Asia-Pacific	6	8	-	-	394
Hewlett-Packard	1	1	-	-	100
Total other	9	18	-	-	1268
Grand Total	30	113	-	-	10147

Countries, sites, and CPU available in EGEE production service

EGEE partner regions

Other collaborating sites

Service Usage

- **VOs and users on the production service**
  - In addition to these there are many VO that are local to a region, supported by their ROCs, but not yet visible across EGEE
- **Active HEP experiments:**
  - 4 LHC, D0, CDF, Zeus, Babar
- **Active other VO:**
  - Biomed, ESR, Compchem, Magic, EGEODE
- **6 disciplines**
- **Registered users in these VO: 500**

- **Work performed:**
  - LHC Data challenges 2004:
    - >1 M SI2K years of cpu time (~1000 cpu years)
    - 400 TB of data generated, moved and stored
    - 1 VO achieved ~4000 simultaneous jobs (~4 times CERN grid capacity)

Introducing VOs

- **Mechanics:**
  - The recipe is straightforward and clear
  - But, this is a heavy weight process and must be improved
  - Requires a lot of configuration changes by a site
    - Often leads to problems
- **Policy:**
  - Joint group of SA1/NA4 (called OAG in the TA)
  - Members are the application representatives and the ROC managers; chaired by NA4
    - Mandate
      - *Understand application resource requirements*
      - *Negotiate those resources within the federations – the ROC manager is responsible to make the negotiation*
    - NB. A site is often funded for specific applications – it is by and large NOT the case that any application is entitled to run anywhere
      - *But let's demonstrate the value of being able to do that …*

LCG-2 software

- **Evolution through 2003/2004**
  - Focus has been on making these
    - rather than additional functionali
  - Respond to needs of users, adm
- **The software stack is the followi**
  - Virtual Data Toolkit
    - Globus (2.4.x), Condor, etc
  - EU DataGrid project developed higher-level components
    - Workload management (RB, L&B, etc)
    - Replica Location Service (single central catalog), replica management tools
    - R-GMA as accounting and monitoring framework
    - VOMS being deployed now
  - Operations team re-worked components:
    - Information system: MDS GRIS/GIIS → BDII
    - edg-rm tools replaced and augmented as lcg-utils
    - Developments on:
      - *Disk pool managers (dCache, DPM)*
        - Not addressed by JRA1
  - Other tools as required:
    - e.g. GridIce – EU DataTag project

- Maintenance agreements with:
  - VDT team (inc Globus support)
  - DESY/FNAL - dCache
  - JRA1 teams:
    - WLM, VOMS, Data Management

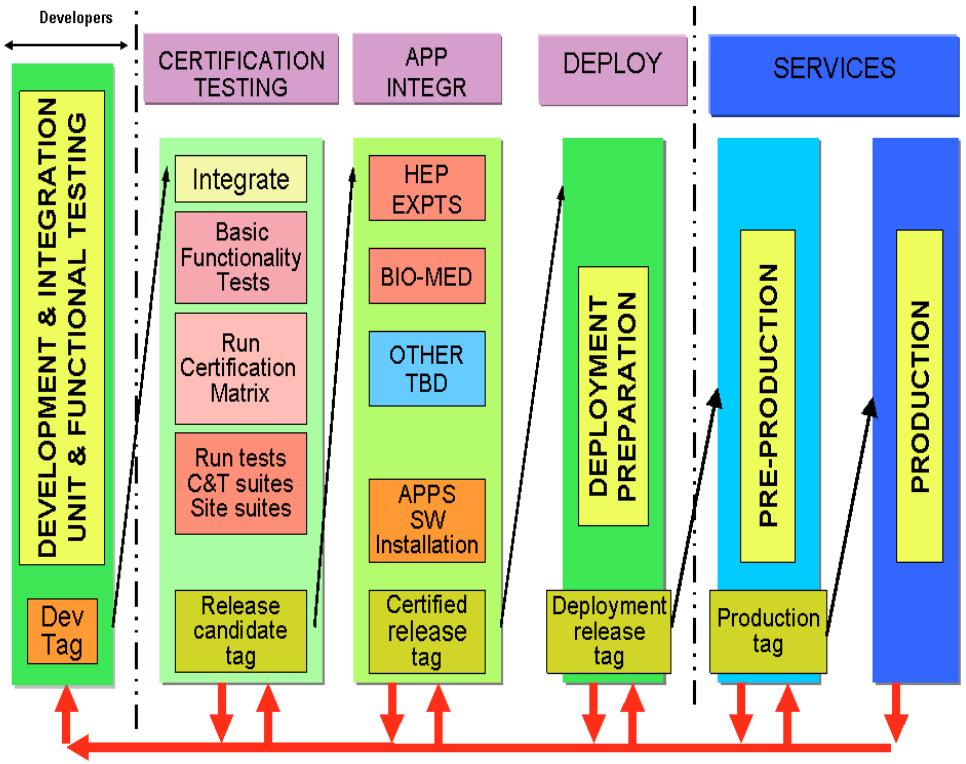
LCG-2 software

- **Platform support**
  - Was an issue – limited to RedHat 7.3
  - Now ported to: Scientific Linux (RHEL), Fedora, IA64, AIX, SGI

- **Another problem was heaviness of installation**
  - Now much improved and simpler – with simple installation tools, allow integration with existing fabric management tools
  - Very light installation on worker nodes – user level

The deployment process

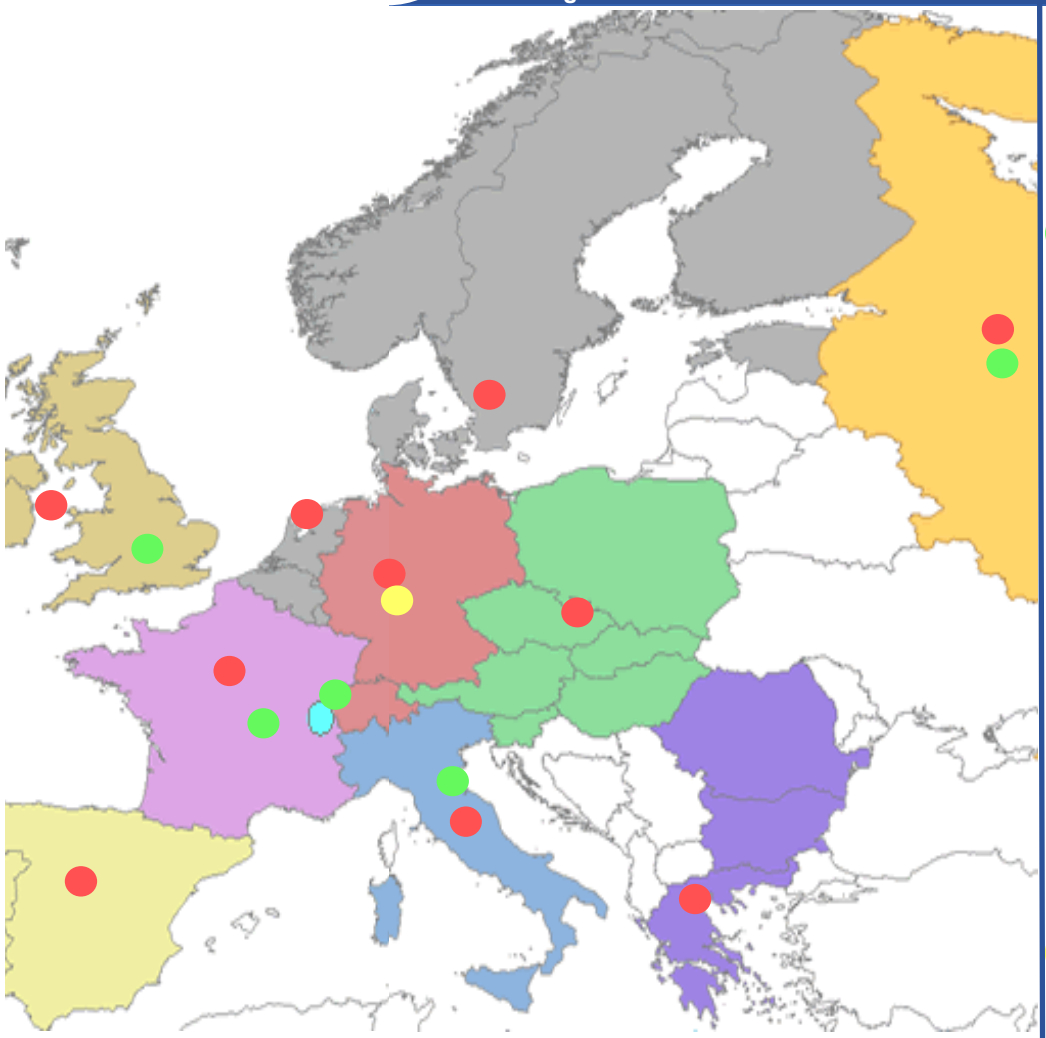
- **Key point – a certification process is essential**
  - However, it is expensive (people, resources, time)
  - But, this is the only way to deliver production quality services
  - LCG-2 was built from a wide variety of "research" quality code
    - Lots of good ideas, but little attention to the "mundane" needs of production …
  - Building a reliable distributed system is hard –
    - Must plan for failure, must provide fail-over of services, etc
  - Integrating components from different projects is also difficult
    - Lack of common standards for logging, error recovery, etc
- **→ Markus Schulz talk**



Overall status

- **The EGEE production grid service is quite stable**
  - The services are quite reliable
  - Remaining instabilities in the IS are being addressed
    - Sensitivity to site management
  - Problems in underlying services must be addressed
    - e.g. Globus gridftp → reliable file transfer service
- **The biggest problem is stability of sites**
  - Configuration problems due to complexity of the middleware
  - Fabric management at less experienced sites
- **Job efficiency is not high, unless**
  - Operations/Applications select stable sites (BDII allows a application-specific view)
- **Operations workshop last November to address this**
  - Fabric management working group – write fabric management cookbook
  - Tighten operations control of the grid – escalation procedures, removing bad sites

SA1 – Operations Structure



- **Operations Management Centre (OMC):**
  - At CERN – coordination etc
- **Core Infrastructure Centres (CIC)**
  - Manage daily grid operations – oversight, troubleshooting
  - Run essential infrastructure services
  - Provide 2nd level support to ROCs
  - UK/I, Fr, It, CERN, + Russia (M12)
  - Taipei also run a CIC
- **Regional Operations Centres (ROC)**
  - Act as front-line support for user and operations issues
  - Provide local knowledge and adaptations
  - One in each region – many distributed
- **User Support Centre (GGUS)**
  - In FZK – manage PTS – provide single point of contact (service desk)
  - Not foreseen as such in TA, but need is clear

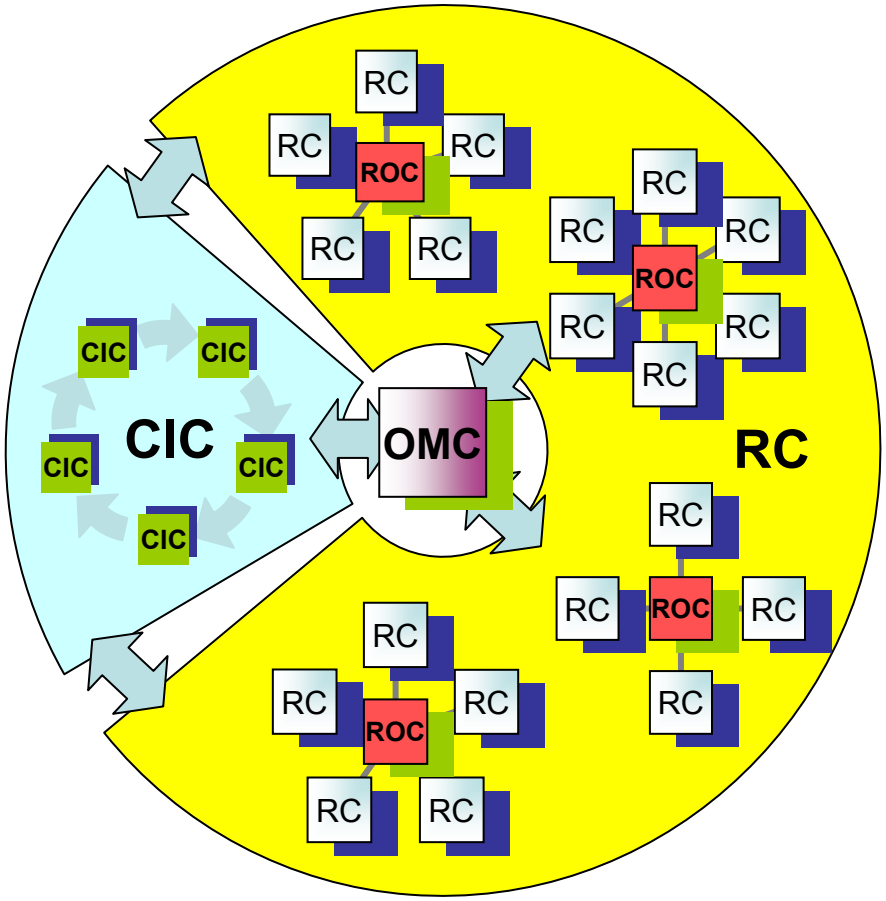
The Regional Operations Centres

Enabling Grids for E-sciencE

- **Represent a significant investment of the project**
- **Provide the front-line for**
  - User support → local helpdes
  - Deployment support in the re
  - Coordination, planning and s
  - Negotiate resource access fo
  - Local adaptations of middlew
  - Site certification and operatio
  - Negotiate and monitor SLAs in the region
- **The ROC is the first point of contact for all:**
  - New sites joining the grid and support for them
  - New users and user support
- **ROCs were set up early in the project and have been active in several areas**
  - Operations requirements for gLite
  - Provide incident security team members
  - Part of user support task force to build up GGUS and support infrastructure
- **This hierarchy of regional responsibility is essential**
  - Coordinate planning and reporting in the region,

- **Teams:**
  - Deployment team
  - Support team (answers user and rc problems)
  - Operations training at RC's
  - Organise tutorials for users

Grid Operations



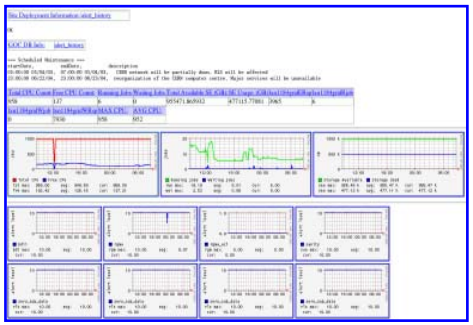
- **The *grid* is flat, but**
- ***Hierarchy* of responsibility**
  - Essential to scale the operation
- **CICs act as a single Operations Centre**
  - Operational oversight ***(grid operator)*** responsibility
  - rotates weekly between CICs
  - Report problems to ROC/RC
  - ROC is *responsible* for ensuring problem is resolved
  - ROC oversees regional RCs
- **ROCs responsible for organising the operations in a region**
  - Coordinate deployment of middleware, etc
- **CERN coordinates sites not associated with a ROC**

Grid Operations & Monitoring

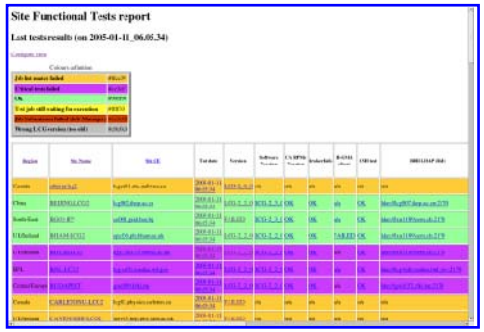
- **CIC-on-duty**
  - Responsibility rotates through CIC's – one week at a time
  - Manage daily operations – oversee and ensure
    - Problems from all sources are tracked (entered into PTS)
    - Problems are followed up
    - CIC-on-duty hands over responsibility for problems
  - Hand-over in weekly operations meeting
- **Daily operations:**
  - Checklist
  - Various problem sources: monitors, maps, direct problem reports
- **Next step:**
  - Continue to develop tools to generate automated alarms and actions

Operations Monitoring

GIIS Monitor



GIIS Monitor graphs



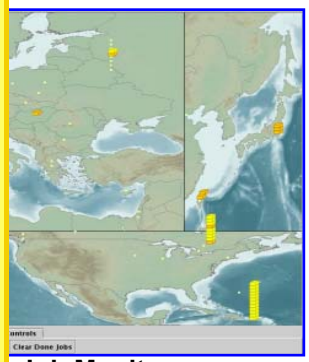
Sites Functional Tests

GOC Dat

Variety of monitoring tools are in daily use.

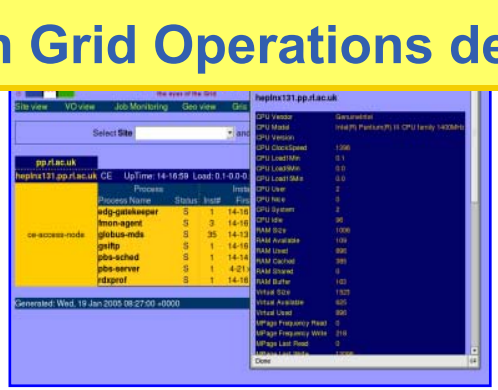
Some of these are shown on the live display monitors.

More details in Grid Operations demo.

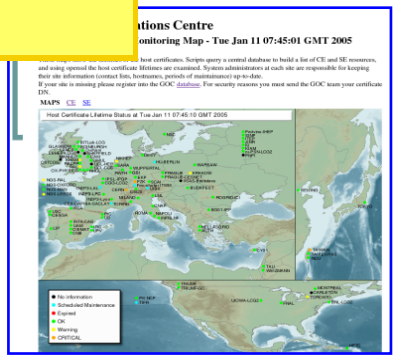


Job Monitor

Gridlce – VO view



Gridlce – fabric view



Certificate Lifetime Monitor

Note: Those thumbnails are links and are clickable.

Live displays

- **Running at the side of the auditorium**

1. **GIIS (Information system) monitoring**
   - Live state of information system
   - Check on information published
2. **Service test results**
   - Checks of services and installations – site validations
3. **Live job status applet**
   - Shows jobs queued, dispatching, executing, etc.
4. **GridIce – views of**
   - Remote site fabric management,
   - VO-specific views

Escalation procedures

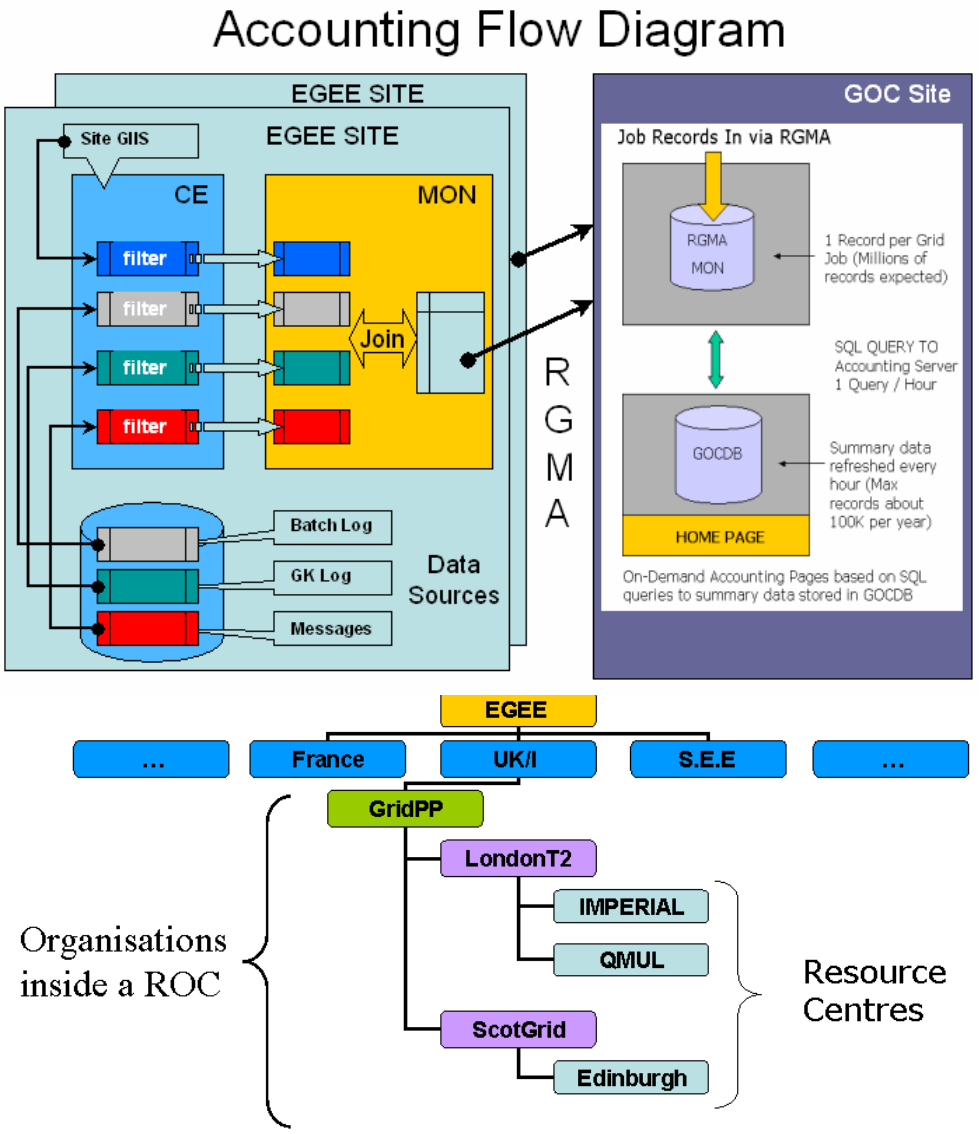
- **Need service level definitions (1st pass at SLA)**
  - What a site supports (apps, software, MPI, compilers, etc)
  - Levels of support (# admins, hrs/day, on-call, operators…)
  - Response time to problems
  - Agreement (or not) that remote control is possible (conditions?)
- **Sites sign-off on responsibilities**
- **Publish sites as bad in info system**
  - Based on unbiased checklist (written by CICs)
  - Consistently bad sites → escalate to political level (e.g. Project Management Board)
- **Small/bad sites**
  - Remote management of services
  - “Force” sites to follow upgrades
  - Remote fabric monitoring (GridICE etc)

24x7 extended support

- **How to move towards a 24x7-like global support?**
  - Separate security (urgent issues) from general support
  - Distributed CIC provides “24x7” by using EGEE, Taipei, (America/Canada?)
  - What is the requirement for 24x7?
    - Real 24x7 coverage only at CERN and large centres (CIC-centres)
    - Or other specific crucial services that justify cost
    - Loss of capacity – vs damage
    - Classify what are 24x7 problems
  - Direct user support not needed for 24x7
    - Massive failures should be picked by operations tools
- **Having an operating production infrastructure should not mean having staff on shift everywhere**
  - “best-effort” support
  - The infrastructure (and applications) must adapt to failures

Accounting in EGEE

- **Accounting at the moment is "after the fact"**
  - The most important way to determine how many resources were consumed by each VO (and potentially each user)
  - No attempt to establish or impose quotas
    - But of course, each site can and does do so
    - Not a trivial problem – jobs should not go to a site where they have no resource, but a modern batch system cannot give a definitive reply



Accounting views

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

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

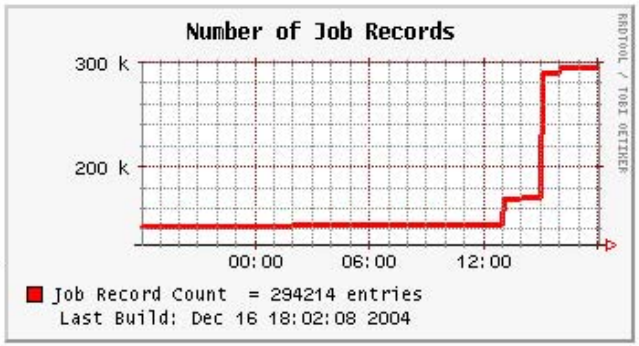
GOC Accounting Services

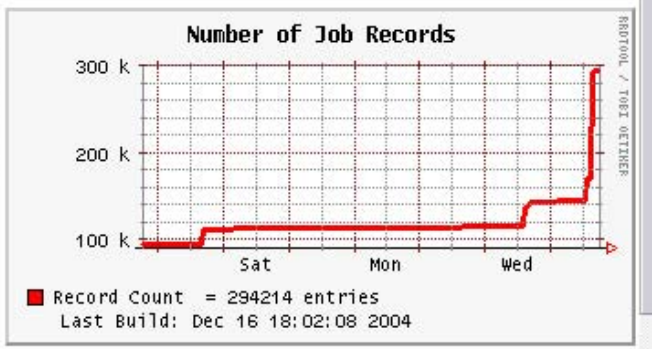
Latest News

- **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.

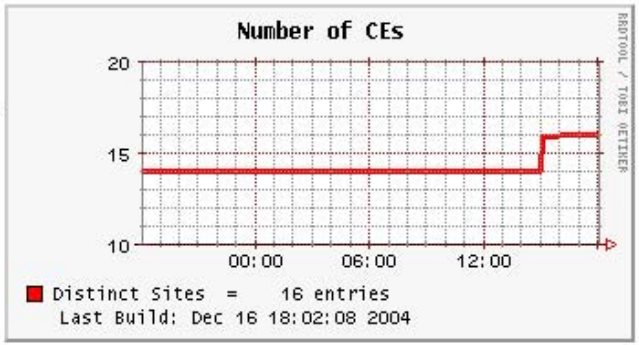
Database Statistics

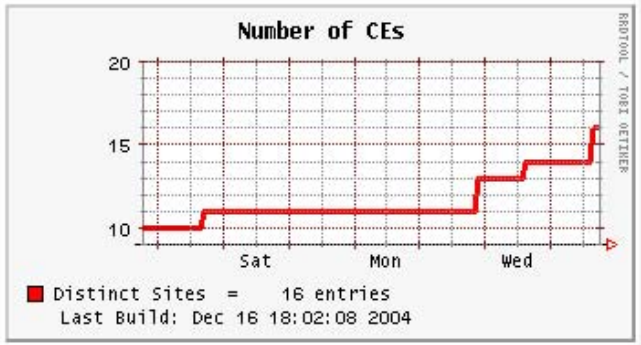
Total Number of Records in Database





Total number of distinct sites reporting to GOC

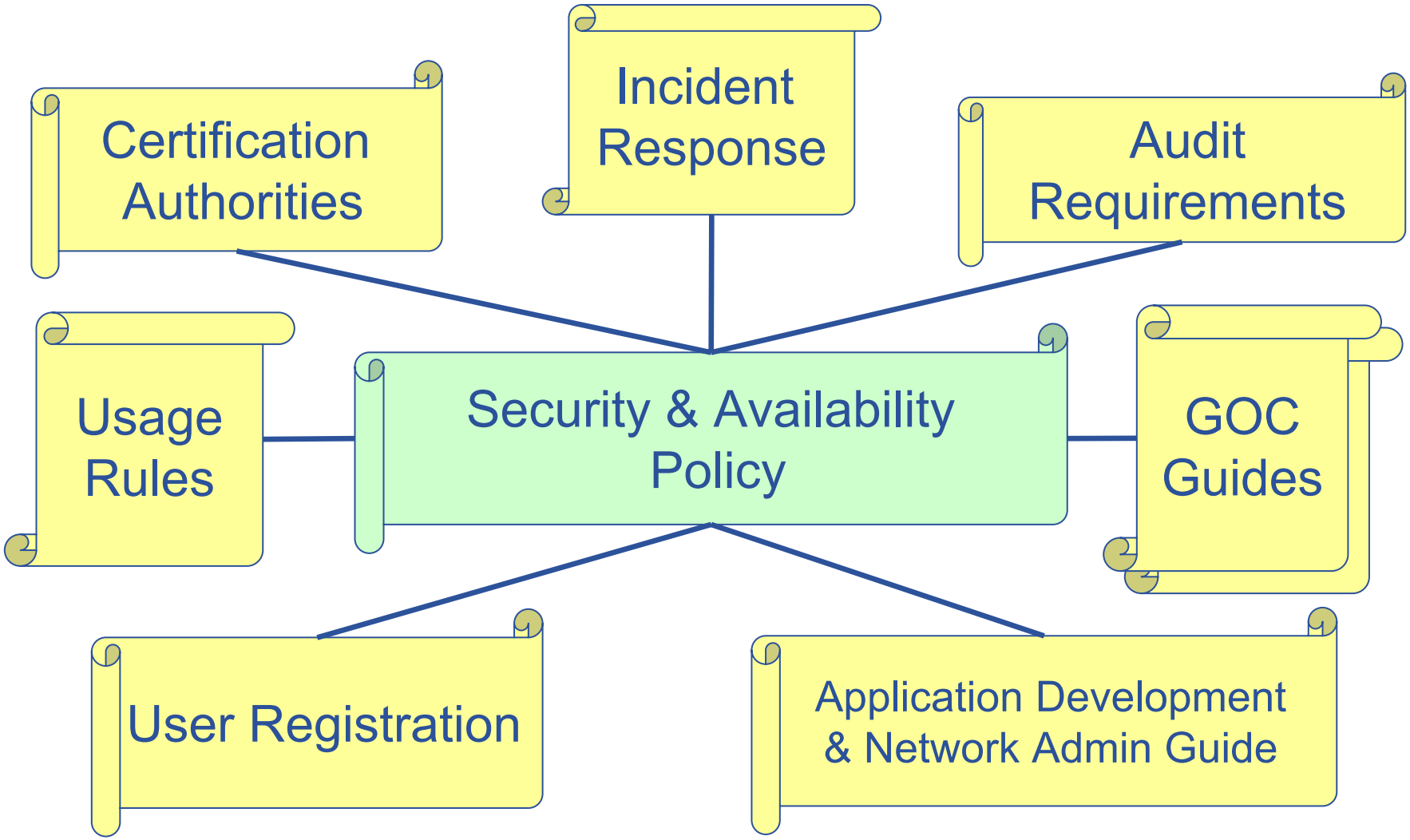




Operational Security

- **Operational Security team in place**
  - EGEE security officer, ROC security contacts
  - Concentrate on 3 activities:
    - Incident response
    - Best practice advice for Grid Admins, creating dedicated web
- **Security group and work was started in LCG – was from the start a cross-grid activity.**
- **Much already in place at start of EGEE: usage policy, registration process and infrastructure, etc.**
- **We regard it as crucial that this activity remains broader than just EGEE**
    - handling and response to cyber security incidents on Grids
  - Basic framework for incident definition and handling
- **Site registration process in draft**
  - Part of basic SLA
- **CA Operations**
  - EUGridPMA – best practice, minimum standards, etc.
  - More and more CAs appearing

Policy – Joint Security Group

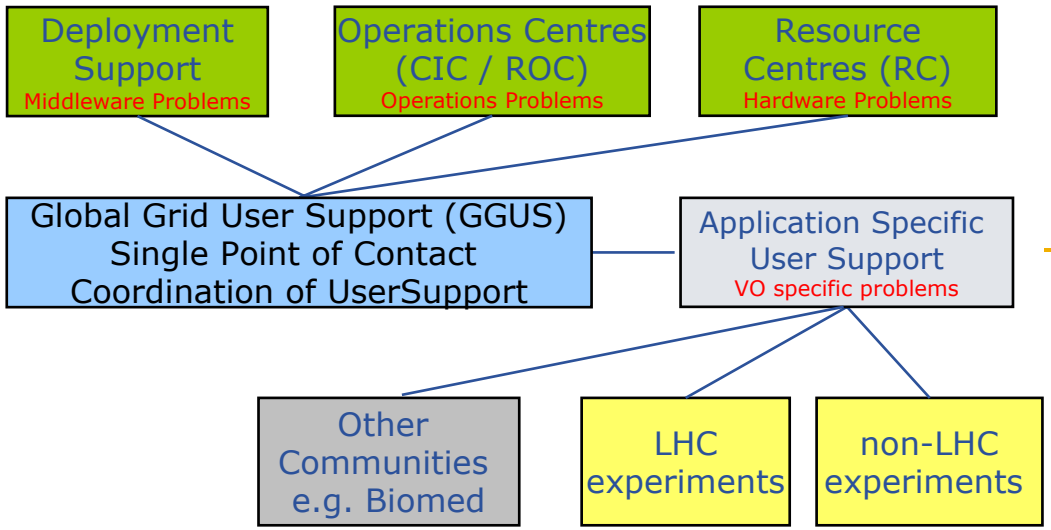


http://cern.ch/proj-lcg-security/documents.html

User Support

We have found that user support has 2 distinct aspects:

- **User support –**
  - Call centre/helpdesk
  - Coordinated through GGUS
  - ROCs as front-line
  - Task force in place to improve the service



- **VO Support**
  - Was an oversight in the project and is not really provisioned
  - In LCG we have a team (5 FTE):
    - Help apps integrate with m/w
    - Direct 1:1 support
    - Understanding of needs
    - Act as advocate for app
  - This is really missing for the other apps – adaptation to the grid environment takes expertise

Relationship to other grids

- **National Grids within EGEE**
  - The large national grid infrastructures in EGEE regions are becoming integrated into the overall service:
    - Italy – Grid.IT sites are part of EGEE
    - UK/I – National Grid Service sites are part of EGEE
    - Nordic countries – Some sites run EGEE in parallel with NorduGrid, others still to be integrated
    - SA1 supports SEE-Grid directly (1 FTE at CERN)
    - Looking at similar arrangement with Brazil (EELA project)
- **Strong relationship with Asia-Pacific**
  - Taipei acts as CIC and hopefully will become a ROC
- **External Grids**
  - Most important are Grid3 (→ Open Science Grid) in USA and the Canadian Grid efforts (WestGrid and GridCanada)
    - OSG and EGEE use same base sw stack – we have demonstrated job interoperability in both directions
      - *Operations and security teams have much in common – proposing specific joint activities*
    - Canada – at Triumf a gateway from EGEE to Canadian resources has been built and used in production
- **This momentum has to be maintained as we move to the next generation of middleware**

Plan for next 15 months

- **Milestones**
  - MSA1.3 (M14) Full production grid infrastructure operational
    - 20 sites, using re-engineered middleware
  - MSA1.4 (M18) Second project review
  - MSA1.5 (M24) Expanded production grid operational
    - 50 sites
- **Deliverables**
  - DSA1.4 (M12) Assessment of operation of 1st 12 months
  - DSA1.5 (M14) First release of “cook-book”
  - DSA1.6 (M14) Release notes corresponding to MSA1.3
  - DSA1.7 (M22) Second edition of “cook-book”
  - DSA1.8 (M24) Assessment of production operation
    - Include thoughts on how to make the infrastructure sustainable
  - DSA1.9 (M24) Release notes corresponding to MSA1.5
- **Changes wrt TA**
  - No significant change

Summary

- **Production grid is operational and in use**
  - Larger scale than foreseen, use in 2004 probably the first time such a set of large scale grid productions has been done
  - Leveraged work done in LCG over previous 18 months
  - Modest growth in resources foreseen over next year
- **Operational infrastructure in place and working**
  - Need to improve reliability of service
  - Need to improve user support
- **Support for applications and VOs**
  - VO deployment should become simpler and more routine
  - Application support needs more resources than foreseen
- **Strong collaboration with projects outside of Europe**
- **Deployment and migration to gLite is now a major focus**

⇨ We have come a long way in the last few months

⇨ There is still much to be done