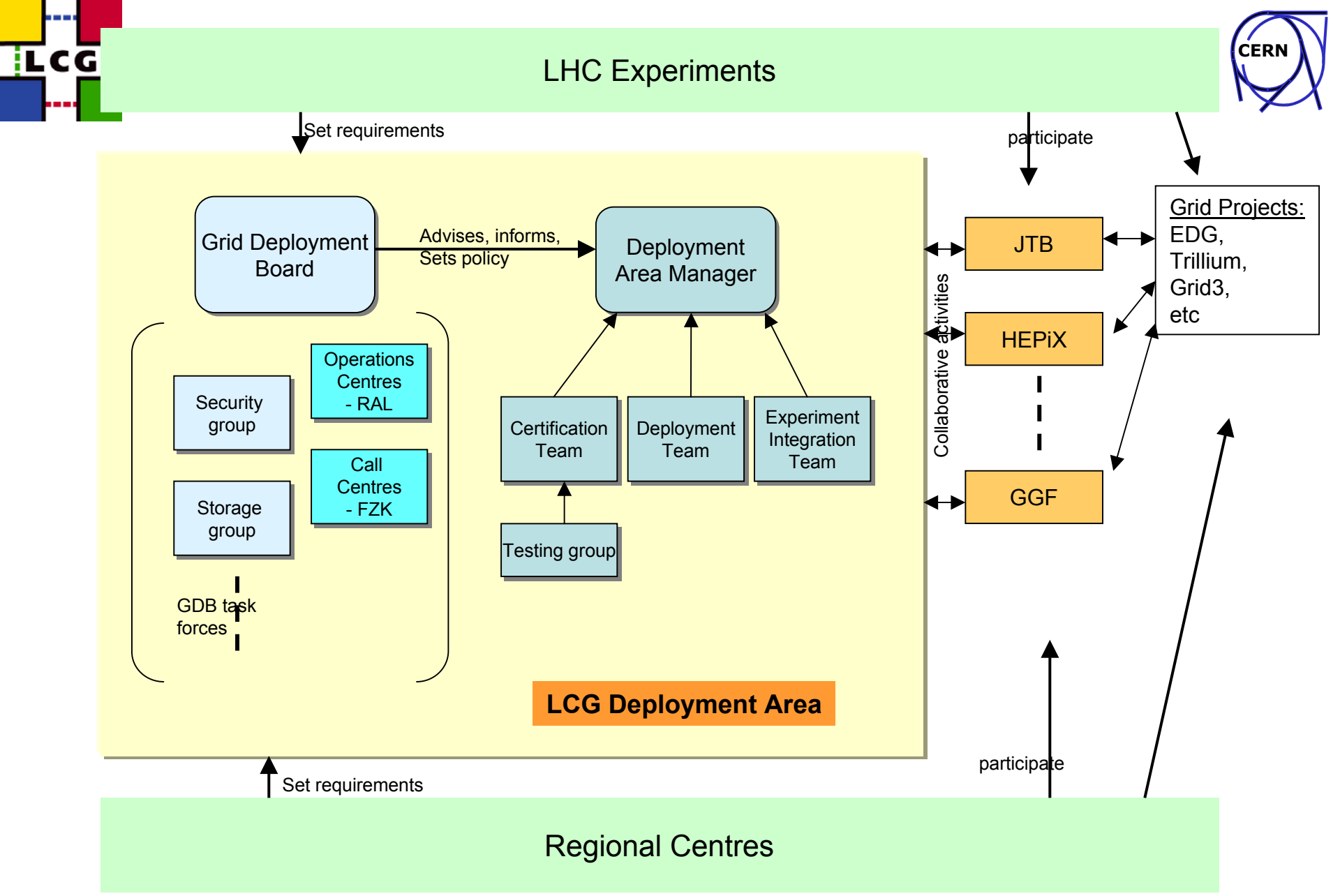
LCG Deployment Overview

Ian Bird
CERN IT/GD

LHCC Comprehensive Review
24-25 November 2003

LCG Grid Deployment Area

- Goal: - deploy & operate a prototype LHC computing environment
- Scope:
  - Integrate a set of middleware and coordinate and support its deployment to the regional centres
  - Provide operational services to enable running as a production-quality service
  - Provide assistance to the experiments in integrating their software and deploying in LCG; Provide direct user support
- Deployment Goals for LCG-1
  - Production service for Data Challenges in 2004
    - Initially focused on batch production work
  - Experience in close collaboration between the Regional Centres
  - Learn how to maintain and operate a global grid
  - Focus on building a production-quality service
    - Focus on robustness, fault-tolerance, predictability, and supportability
  - Understand how LCG can be integrated into the sites' physics computing services
    - Move away from dedicated testbeds



LCG Deployment Organisation and Collaborations

Deployment Activities: Human Resources

Activity	CERN/LCG	External	
Integration & Certification	6	External tb sites	In collaboration
Debugging/development/mw support	3		
Testing	3	2 + VDT testers group	Team of 3 Russians have 1 at CERN at a given time (3 months)
Experiment Integration & Support	5		
Deployment & Infrastructure Support	5.5	RC system managers	
Security/VO Management	2	Security Task Force	Refer to Security talk
Operations Centres		RAL + GOC Task Force	Refer to Operations Centre talk
Grid User Support		FZK + US Task Force	
Management	1		
Totals	**25.5**		

- The GDA team has been very understaffed – only now has this improved with 7 new fellows
- There are many opportunities for more collaborative involvement in operational and infrastructure activities

Implementation – 1

- A core team at CERN – Grid Deployment group
- Collaboration of the regional centres – through the Grid Deployment Board
- Partners take responsibility for specific tasks (e.g. GOC)
- Focussed task forces as needed
- Collaborative joint projects – via JTB, grid projects, etc.

- CERN deployment group, includes LCG funded staff, fellows, etc
  - Core preparation, deployment, and support activities
  - Integration, packaging, debugging, development of missing tools,
  - Deployment coordination & support, security & VO management,
  - Experiment integration and support

- GDB: Country representatives for regional centres
  - Address policy, operational issues that require general agreement
  - Brokered agreements on:
    - Initial shape of LCG-1 via 5 working groups
    - Security
    - What is deployed

Implementation – 2

- Several long-term groups, set up by the GDB:

➢ Security group
  - Members: site security contacts, experiments
  - Proposes policies, usage rules, registration, etc, operational issues – audit, incident response

➢ Grid Operations Centre – RAL
  - Includes a GOC steering group

➢ Call Centre – FZK
  - Works together with GOC groups

❖ These will be discussed in later talks

Task forces

- Limited time task forces set up to address specific issues:

➢ Mass storage access
  - Working to agree and implement a common strategy to provide access to mass storage (tape and disk) at LCG sites

➢ Packaging, installation, configuration tools
  - To address the problems of simplifying m/w installation and config

- …

Collaborative activities

- Via the HICB-JTB (HEP Inter-Grid Coordination Board – Joint Technical Board)
  - Members from US and EU grid projects
  - Address common issues of interoperability
    - GLUE schema, Testing, Replica location services
    - Installation tools, Monitoring, Workload management
  - Demonstrate interoperability …
- Through Global Grid Forum
  - PNPA research area – bring HEP experiences and requirements to GGF
  - Several relevant areas: production grid management, security, user support, SRM, etc.
- Other collaborative activities include:
  - Russian participation in testing group
  - Collaboration on monitoring with INFN/DataTag
  - Collaboration with Indian group
  - US-LHC
    - VO management tools

2003 Milestones

Project Level 1 Deployment milestones for 2003:

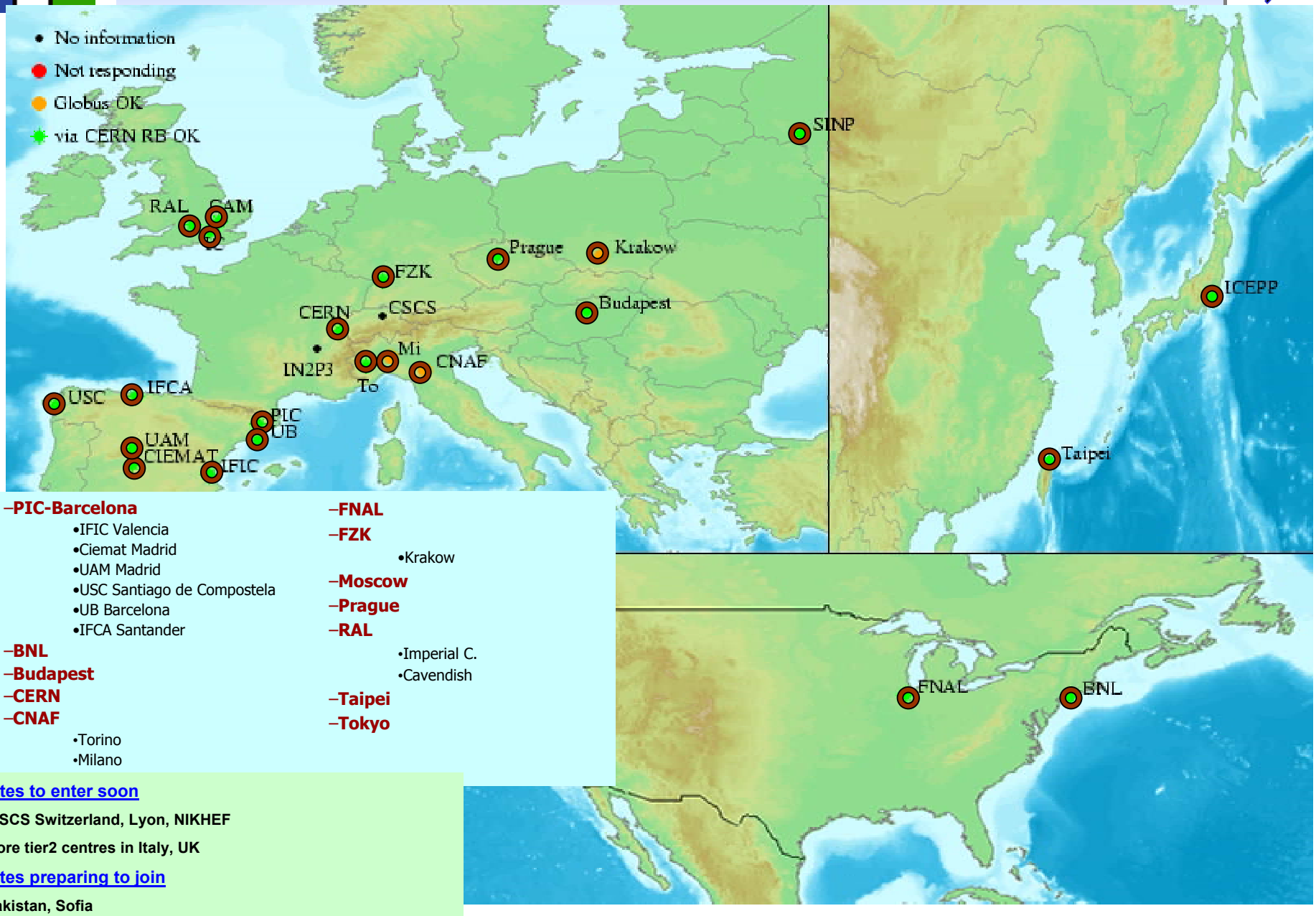
- **July:** Introduce the initial publicly available LCG-1 global grid service
  - With 10 Tier 1 centres in 3 continents
- **November:** Expanded LCG-1 service with resources and functionality sufficient for the 2004 Computing Data Challenges
  - Additional Tier 1 centres, several Tier 2 centres – more countries
  - Expanded resources at Tier 1s
  - Agreed performance and reliability targets

- The idea was:
  - Deploy a service in July
    - Several months to gain experience (operations, experiments)
  - By November
    - Meet performance targets (30 days running) – experiment verification
    - Expand resources – regional centres and compute resources
    - Upgrade functionality

Milestone Status

- July milestone was 3 months late
  - Late middleware, slow takeup in regional centres
- November milestone will be partially met
  - LCG-2 will be a service for the Data Challenges
  - Regional Centres added to the level anticipated (23), including several Tier 2 (Italy, Spain, UK)
  - But:
    - lack of operational experience
    - Experiments have only just begun serious testing
- LCG-2 will be deployed in December
  - Functionality required for DCs
  - Meet verification part of milestone with LCG-2 – early next year
  - Experiments must drive addition of resources into the service
  - Address operational and functional issues in parallel with operations
    - Stability, adding resources
  - This will be a service for the Data Challenges

Sites in LCG-1 – 21 Nov



- **PIC-Barcelona**
  - IFIC Valencia
  - Ciemat Madrid
  - UAM Madrid
  - USC Santiago de Compostela
  - UB Barcelona
  - IFCA Santander
- **BNL**
- **Budapest**
- **CERN**
- **CNAF**
  - Torino
  - Milano
- **FNAL**
- **FZK**
  - Krakow
- **Moscow**
- **Prague**
- **RAL**
  - Imperial C.
  - Cavendish
- **Taipei**
- **Tokyo**

Sites to enter soon

CSCS Switzerland, Lyon, NIKHEF

More tier2 centres in Italy, UK

Sites preparing to join

Pakistan, Sofia

Achievements – 2003

- Put in place the Integration and Certification process:
  - Essential to prepare m/w for deployment – the key tool in trying to build a stable environment
  - Used seriously since January for LCG-0,1,2 – also provided crucial input to EDG
  - ➢ LCG is more stable than earlier systems
- Set up the deployment process:
  - Tiered deployment and support system is working
  - Currently support load on small team is high, must devolve to GOC
- Support experiment deployment on LCG-0,1
  - User support load high – must move into support infrastructure (FZK)
  - CMS use of LCG-0 in production
  - Produced a comprehensive User Guide
- Put in place security policies and agreements
  - Particularly important agreements on registration requirements
- Basic Operations Centre and Call Centre frameworks in place

- Expect to be ready for the 2004 DCs
  - Essential infrastructures are ready, but have not yet been production tested
  - And, improvements will happen in parallel with operating the system

Issues

- Middleware is not yet production quality
  - Although a lot has been improved, still unstable, unreliable
  - Some essential functionality was not delivered – LCG had to address
- Deployment tools not adequate for many sites
  - Hard to integrate into existing computing infrastructures
  - Too complex, hard to maintain and use
- Middleware limits a site's ability to participate in multiple grids
  - Something that is now required for many large sites supporting many experiments, and other applications
- We are only now beginning to try and run LCG as a service
  - Beginning to understand and address missing tools, etc for operation
- Delays have meant that we could not yet address these fundamental issues as we had hoped to this year

Changing landscape

- The view of grid environments has changed in the past year
- From
  - A view where all LHC sites would run a consistent and identical set of middleware,
- To
  - A view where large sites must support many experiments each of which have grid requirements
  - National grid infrastructures are coming – catering to many applications, and not necessarily driven by HEP requirements
- We have to focus on interoperating between potentially diverse infrastructures ("grid federations")
  - At the moment these have underlying same m/w
  - But modes of use and policies are different
- Need to have agreed services, interfaces, protocols
- The situation is now more complex than anticipated

Expected Developments in 2004

- General:
  - LCG-2 will be the service run in 2004 – aim to evolve incrementally
  - Goal is to run a **stable** service
- Some functional improvements:
  - Extend access to MSS – tape systems, and managed disk pools
  - Distributed replica catalogs – with Oracle back-ends
    - To avoid reliance on single service instances
- Operational improvements:
  - Monitoring systems – move towards proactive problem finding, ability to take sites on/offline
  - Continual effort to improve reliability and robustness
  - Develop accounting and reporting
- Address integration issues:
  - With large clusters, with storage systems
  - Ensure that large clusters can be accessed via LCG
  - Issue of integrating with other experiments and apps

Services in 2004 – EGEE

- LCG-2 will be the production service during 2004
  - Will also form basis of EGEE initial production service
  - EGEE will take over operations during 2004 – but same teams
  - Will be maintained as a stable service
  - Peering with other grid infrastructures
  - Will continue to be developed
- Expect in parallel a development service – Q204
  - Based on EGEE middleware prototypes
  - Run as a service on a subset of EGEE/LCG production sites
  - Demonstrate new architecture and functionality to applications
- Additional resources to achieve this come from EGEE
- Development service must demonstrate superiority
  - All aspects: functionality, operational, maintainability, etc.
- End 2004 – hope that development service becomes the production service

Summary – 1

- Initial milestone was late –
  - middleware came late, functionality less than hoped for
- Many issues with deployment …
  - Packaging, dependencies, incompatibility of m/w installs with others, requirement to control full machine environment, etc…
- … and with lack of functionality for experiments …
- … were due to the legacies LCG inherited
  - Working hard to get away from them – but it is a complex problem
  - Lack of time to adapt research products to a production service environment
- Very little time to run this as a service – we are still resolving basic operational issues
  - This will continue during 2004

Summary - 2

- LCG-1 is deployed to 23+ sites
  - Despite the complexities and problems
- The LCG-2 functionality
  - Will support production activities for the Data Challenges
  - Will allow basic tests of analysis activities
  - And the infrastructure is there to provide all aspects of support
- Staffing situation at CERN is now better
  - Most of what was done so far was with limited staffing
  - Need to clarify situation in the regional centres
- We are now in a position to address the complexities