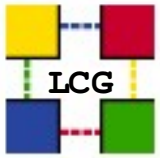


LHC Computing Grid Project – LCG

The LHC Computing Grid Service and EGEE

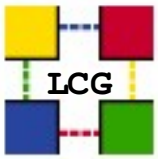
EGEE-LHC town meeting

CERN – 22 February 2003

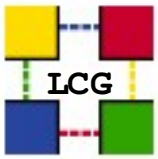
Les Robertson
Information Technology Division
CERN

les.robertson@cern.ch

last update: 22/02/2003 11:32



- General issues
- Deploying a service
- Acquiring Grid middleware
- LCG & EGEE



What is the LHC Computing Grid Project?

Mission of Phase 1 (2003-2005) -
prepare the computing environment for the analysis of LHC data

- applications software - environment, tools, frameworks, common developments
- deploy and coordinate a global grid service
  - acquire and organise support for **robust, maintainable** middleware
  - set up and learn how to operate a **service**
    - encompassing the LHC Regional Centres not a service apart
    - strategy and policy for resource allocation
    - authentication, authorisation, accounting
    - user support, operation
    - performance and optimisation
    - evolution

LCG service development timeline

- 1H03 – development of the basic service
  - service definition
  - establish certification, distribution & support processes
- 2H03 – LCG-1 service ramp-up
  - batch service – used for simulated event productions
  - basic data management
  - work on reliability, availability, operability
  - develop operations centre, call centre
  - improve scaling, data movement performance
- 2004 – LCG-2 the principal service for data challenges
  → computing model TDRs due end 2004
- 2005 – LCG-3 the full LCG Pilot
  - fulfill basic requirements
  - initial experience with interactive analysis

  → Technical Design Report for Phase 2 – mid-2005

Timeline for the LCG computing service

VDT, EDG tools building up to basic functionality

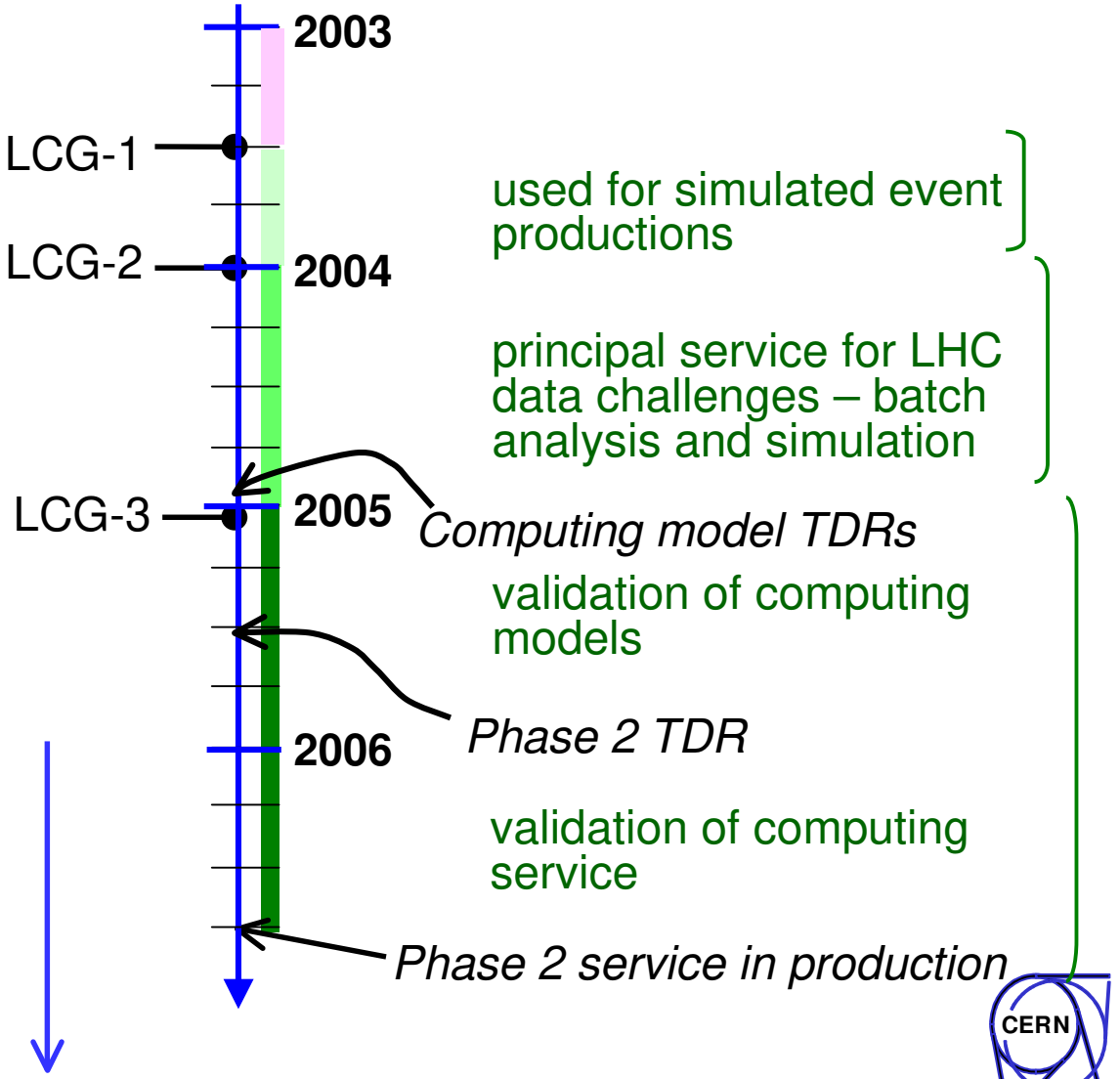
Stable 1st generation middleware

Developing management, operations tools

More stable 2nd generation middleware

Very stable full function middleware

Acquisition, installation, commissioning of Phase 2 service (for LHC startup)



Observations

- The middleware tools that we can identify today must be hardened this year and supported into 2005

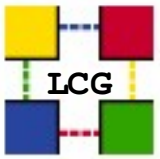
- It becomes increasingly difficult to introduce new middleware into a stable, production service

- The LHC analysis facility will be fully distributed from day 1
  - no site will have more than ~15% of the analysis capacity

- Our minds and mails have been concentrating on the middleware

  But there are many (equally difficult) problems to solve in deploying a coherent, productive service for data analysis

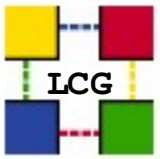
- The middleware that will run the service when LHC starts in April 2007, must be deployed at least one year before



Situation Today

- We are still solving basic reliability & functionality problems
  - We still have a long way to go to get to a solid service
  - A solid service in mid-2003 looks ambitious
- We have not yet addressed system level issues
  - How to manage and maintain the Grid as a system providing a high-quality reliable service.
  - Few tools and treatment in current developments of problem determination, error recovery, fault tolerance etc.
- Some of the advanced functionality we will need is only being thought about now
  - Comprehensive data management, SLA's, reservation schemes, interactive use.
- Many many initiatives are underway and more coming

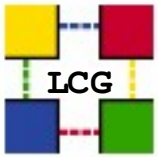
How do we manage the complexity of all this ?



Establishing Priorities

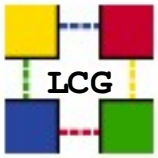
- We need to focus on a model that we can easily explain and understand
- The basic requirements of the 2004 Data Challenges are a good starting point
  - Focus on robust job scheduling, data handling.
- We must make the simple things work well before we expand the scope
- This is not to say that we should not be working on advanced requirements –

  But we must recognise the difference between R&D and providing a service

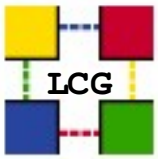


LCG Guidelines

- Focus on a **Service for Physics**
  - simulation first
  - then batch analysis
  - and later interactive analysis
- Keeping everything else **as simple as possible**

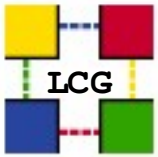


Deploying a service



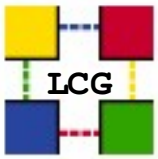
Deploying the LCG Service

- Middleware:
  - Testing and certification
  - Packaging, configuration, distribution and site validation
  - Support – problem determination and resolution; feedback to middleware developers
- Operations:
  - Grid infrastructure services
  - Site fabrics run as production services
  - Operations centres – trouble and performance monitoring, problem resolution – 24x7 globally
- Support:
  - Experiment integration – ensure optimal use of system
  - User support – call centres/helpdesk – global coverage; documentation; training



Certification and Testing

- Will be an ongoing major activity of LCG
  - Part of what will make LCG a production-level service
- Goals:
  - Certify/validate that middleware behaves as advertised and provides the required functionality (HEPCAL)
  - Stabilise and harden middleware
  - Provide debugging, problem resolution *and feedback to developers*
- Testing activities at all levels
  - Component/unit tests
  - Basic functional tests, including tests of distributed (grid) services
  - Application level tests – based on HEPCAL use-cases
    - Driven/implemented by the experiments – GAG
  - Experiment beta-testing before release
  - Site configuration verification
- JTB collaborative project – LCG, Trillium, EDG
  - Gather existing tests
  - Write/obtain missing tests

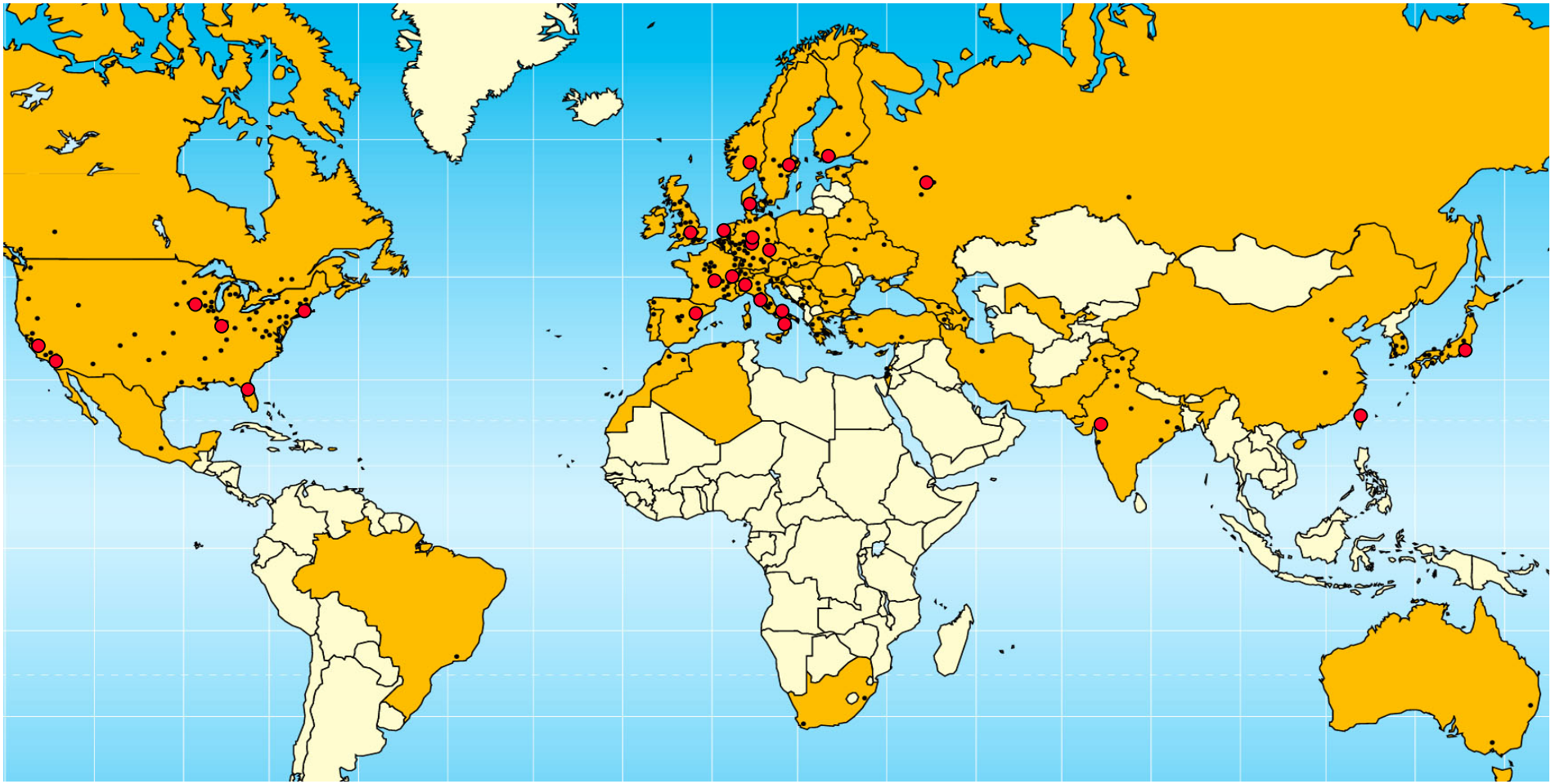


Packaging and distribution

- Obviously a major issue for a deployment project
- Joint activity started –
  - Discussions LCG, EDG, VDT, EDT, iVDGL, etc.
  - Have produced a draft discussion document
  - Will soon lead to a JTB joint project

- Want to provide a tool that satisfies needs of the participating sites,
  - Interoperate with existing tools where appropriate and necessary
  - Does not force solution on sites with established infrastructure
  - Solution for sites with nothing
- Configuration is essential component
  - Essential to understand and validate correct site configuration
  - Effort will be devoted to providing configuration tools
  - Verification of correct configuration will be required before sites join LCG

Centres taking part in the LCG prototype service (2003-05)



around the world → around the clock

Centres taking part in the LCG prototype service – 2003-05

Tier 0

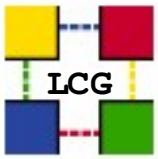
- CERN

Tier 1 Centres

- Brookhaven National Lab
- CNAF Bologna
- Fermilab
- FZK Karlsruhe
- IN2P3 Lyon
- Rutherford Appleton Lab (UK)
- University of Tokyo
- CERN

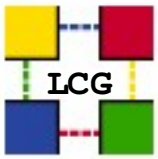
Other Centres

- Academica Sinica (Taipei)
- Barcelona
- Caltech
- GSI Darmstadt
- Italian Tier 2s (Torino, Milano, Legnaro)
- Manno (Switzerland)
- Moscow State University
- NIKHEF Amsterdam
- Ohio Supercomputing Centre
- Sweden (NorduGrid)
- Tata Institute (India)
- Triumf (Canada)
- UCSD
- UK Tier 2s
- University of Florida-Gainesville
- University of Prague
- ……



LCG-1 Ramp-up Schedule

	Date	RegionalCenter	Experiment
Pilot 1 start – Feb 1			
0	15/2/03	CERN	All
1	28/2/03	CNAF, RAL	All
Pilot 2 start – March 15			
2	30/3/03	FNAL	CMS
3	15/4/03	Tokyo	Atlas
4	30/4/03	Karlsruhe	All
5	7/5/03	IN2P3	All
6	15/5/03	BNL	Atlas
7	21/5/03	Russia(Moscow), Taiwan	All
LCG-1 InitialPublic Service Start – July 1			



Setting up the LCG Service

Grid Deployment Board

- senior computing service management, experiment production managers, ..
- policies, strategy, scheduling, standards, recommendations

Grid Resource Coordinator

- facilitating scheduling (data challenges, etc.) between experiments, regional centres within limits set by funding agencies

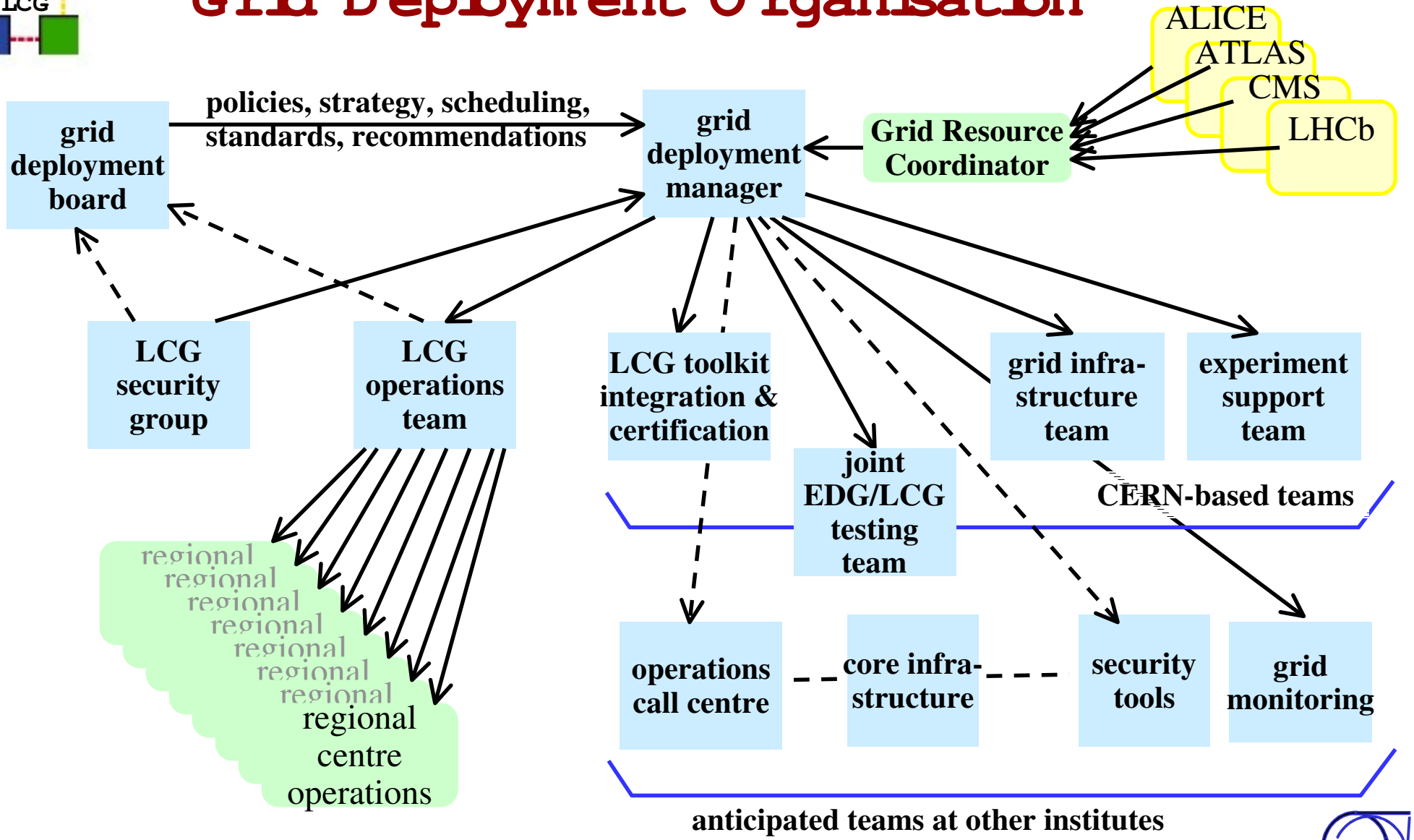
LCG Operations Team

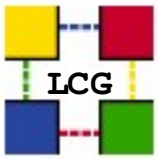
- service managers
- detailed scheduling, day-to-day operations

LCG Security Group

- Security Policy
- Close ties to site security officers

Grid Deployment Organisation

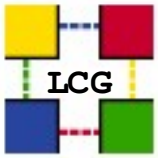




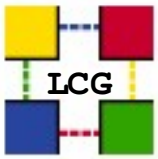
Grid Operations Infrastructure

- Few well staffed centres to provide:
  - Operations and monitoring service – performance monitoring, troubleshooting, problem resolution
  - Call Centre – providing user support, documentation, training
  - Basic grid infrastructure services
  - Centre of expertise in grid tools – provide debugging, problem resolution, interaction with developers
  - Grid configuration management
  - VO management
  - Resource management
- Would like to see a few well-staffed and managed centres rather than a lot of centres with 1 or 2 people.

- LCG is setting up centres such as this – specifically for LHC
  - but exactly how they will operate is not yet clear



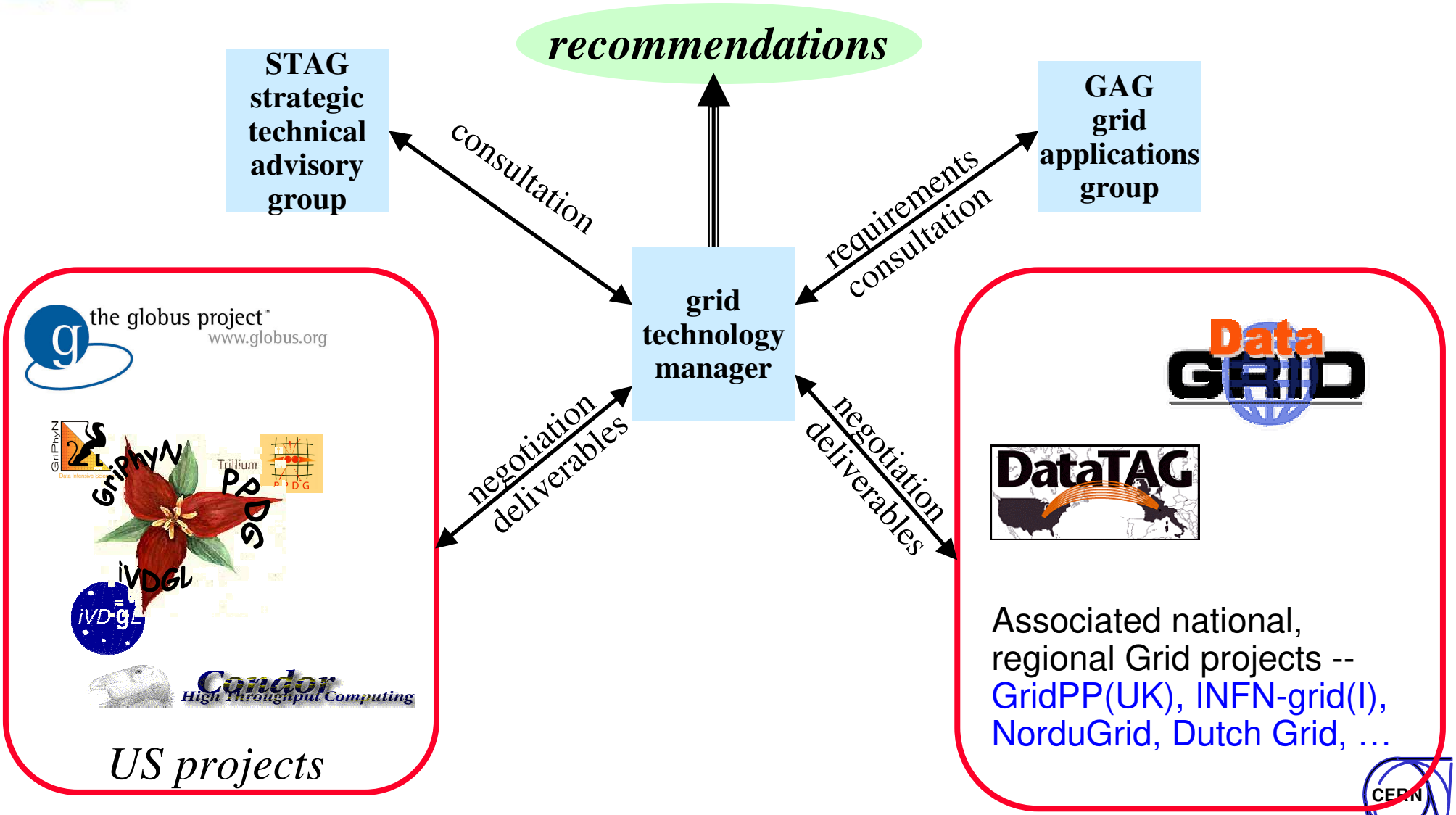
Acquiring Grid Middleware

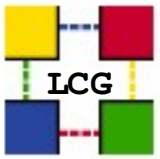


LCG Grid Middleware Challenges

- Have identified the starting technologies to be deployed
  - Driven pragmatically through the GDB/WG1
  - Initial suppliers – VDT and EDG
- Identify the medium term supply & support strategies
  - Requirements from GAG, advice from STAG
  - Short life-time projects (EDG, Trillium, NMI ….) with unclear continuations
  - We need to see credible supplier projects that focus on product quality, maintainability, support, end-user service
  - What about industrial products
- Work towards future middleware solutions that are coherent, acceptable and supportable
  - Inter-working, standards, …
    - — essential for evolution
    - — LCG must be mainline, not HEP-special
  - OGSA helps – but will standards emerge on the LCG timescale?
  - or will we need to apply some Super-GLUE?

LCG Grid Technology Organisation

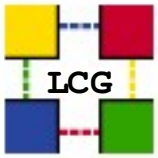




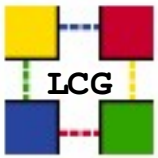
Middleware Support

- Is a process that involves key stakeholders
  - The development community
  - The system s administrators
  - The user support team s
- Is an infrastructure
  - That allow s problem s to be isolated
  - Perm its fixes (short and long term) to be created and deployed
- Is a know ledge system
  - That provides direct input to development
  - That provides feedback for design
  - That gives confidence to users

Middleware support is not an isolated activity

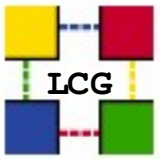


LCG & EGEE



EGEE as a Solution for LCG

- Middleware
  - EGEE could provide the critical mass
    - to fund **hardened & supported** middleware serving a wider community
    - and form a **global** middleware partnership
  - LCG cannot do this on its own
- Operation
  - funds the operation of a core grid in European countries
  - assuming that –
    - the EGEE infrastructure is integrated with national infrastructures
    - and with US, Asian infrastructures
    - the LCG regional centres are smoothly integrated from the start
    - the aim is a long-term infrastructure
- Simplification
  - LCG (and HEP) can concentrate on physics services and leave middleware and grid operation to someone else
- The long term model is provision of a core Grid service
  - GEANT + NRENs provide a model for this
  - but they had a more tangible deliverable
  - and the model took some time to mature



LCG as a Catalyst for EGEE

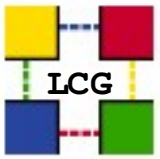
- LCG has a real need
  and a reasonable scale
  and has mature global collaborations of scientists

- LCG must acquire basic middleware for a global science grid

- Solutions for LCG are quite general - readily applicable to other sciences

- LCG will deploy a core grid infrastructure in Europe, America, Asia

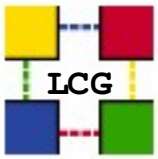
  EGEE can build on this
  - to learn how to provide a "production quality" service during LHC preparation (data challenges)
  - exploit the LHC core infrastructure as the foundation for a general science grid



Timing is Critical

- The two-year first phase could deliver just in time (but only just) for LCG
- No time for a slow start-up
- Optimistic approach essential
- Must complete first round – requirements, planning, identify teams, design, ... recruitment .....

  before funding starts to flow



EGEE Priorities (as seen from LCG)

1. Robust industrial-strength middleware, with complete support structure.
   - Functionality at a "basic" level – providing data management, resource scheduling, and information services
   - Part of a Global middleware programme

2. Basic grid operational and support infrastructure – managed as a production service
   - Operations centres, call centres (user support), centres of expertise, system-level support