

EGEE TAB: draft report

February 22, 2003

Matthias Kaseman

TAB chair

TAB charter

- *The main role of the Technical Advisory Board is to advise to the EGEE Executive Committee on a project structure which is technically sound and in accordance with all requirements coming from the European Commission (call for proposal, work programme, etc...)*
- *This project structure shall consist of a series of well identified, self contained workpackages and shall include if necessary, transverse structures*
- *The workpackage contents and boundaries should be such that small task forces can be launched to write the corresponding section of the proposal, in the required timeframe*

TAB charter

It is expected that the following topics will be covered by the Technical Advisory Board:

- *-Grid Middleware re-engineering*
- *-Grid production support and operation*
- *-Interface with applications.*

Non-technical Work packages such as project management, industry relationship, dissemination, relationship with other projects and networks of excellence will be handled directly by the EC

The Technical Advisory Board can solicit the help of outside experts to accomplish this task

The role of the TAB during a second phase will be to monitor the progress of the various task forces and to guarantee the technical coherence of the whole project proposal

Task Forces will be appointed by the Executive Committee

The TAB will be formally ~~dissolved~~^{fixed} on submission of the project proposal

TAB composition and work

- [Matthias Kasemann](#) (Chair)
CERN/DESY
- [Marian Bubak](#) CYFRONET
Krakow
- [Bob Jones](#) CERN
- [Miron Livny](#) Wisconsin Univ.
- [Malcolm Atkinson](#) UK National
e-Science
- [Thierry Priol](#) IRISA – France
- [Francesco Prelz](#) INFN
- [Dave Pearson](#) ORACLE UK
- [Ian Bird](#) LCG
- The board had a three day meeting on February 13-15, 2003 at CERN.
 - It heard presentations from the HEP and Bio-Medical research communities,
 - reviewed the experience of Grid projects and
 - worked out a draft report.
- The board plans to have the next meeting on March 30 and 31, 2003, where it will review and give advice to the EGEE proposal writing procedure.

**In the following I will present the TAB proposal
for the EGEE work structure**

High level EGEE goals and scope

- Operate a reliable and dependable set of Grid services to a broad and diverse set of users that have a significant impact on the computing and data access capabilities
- Maintain and support a suite of software tools capable of providing these services that are based on proven Grid technologies.
- Facilitate by means of guidance and facilities advances in Grid technology and the development of new Grid middleware
- Establish the vision, framework (structure and procedures) and tools needed to guarantee the longevity of a Grid infrastructure for research.

We see EGEE as a service organization: „customer satisfaction” is major consideration in any decision

Build on existing solutions(1)

- build on the assets of EDG and related projects.
 - a large-scale grid testbed;
 - innovative middleware;
 - enthusiastic, experienced, well-connected and competent teams across Europe within a well functioning organisation;
 - active participation of application groups.
- Primary goal of EGEE:
 - deploy and operate a long-term, large-scale grid facility across Europe that will deliver powerful and reliable computing services to the European research community at large.
 - in an incremental process where communities with established Grid experience like the HEP and Bio-medical communities joining at the early stages
 - others coming on board when their infrastructure and applications are ready to benefit from the services provided by the EGEE.

Build on existing solutions(2)

- It is necessary to improve the reliability, robustness and availability of the existing services and
- to increase the total amount of computing resources at their disposal.

This requires a change of emphasis from middleware development and exploration towards the operational aspects of deploying and operating a world-class grid utility. (EGEE ≠ EDG')

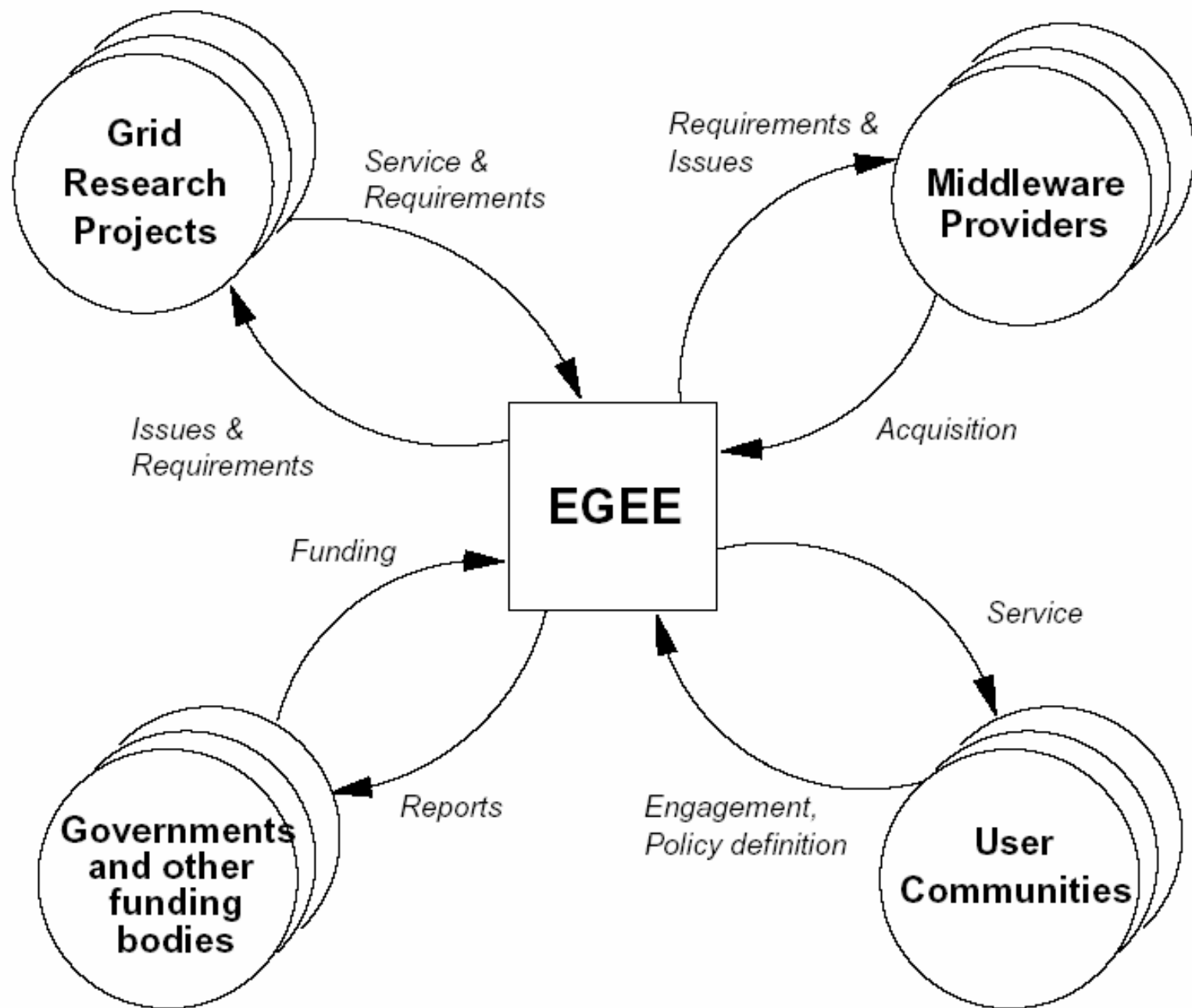
- As EGEE must operate with current technology it will require a **strong operations team** and **investment in improving the software** infrastructure.
- A long-term goal of this improvement is to reduce the deployment, maintenance and running costs associated with large-scale grid utilities.
 - EGEE will do this by applying professional software engineering, quality assurance and management techniques and processes to existing grid technology

Relationships with other entities(1)

- **Research communities (i.e. LCG):**
 - EGEE will turn the resources *owned* by the community into a *seamless* computing environment that is governed by *policies* provided by the community.
 - the resources of the community will be transformed in to a production grid that **operates according to rules defined by the community**.
 - EGEE has to **record and provide the information** needed to verify that the resources are actually managed according to these policies.
 - The community will not relinquish control and will **remain the entity responsible** to its members.
 - Requests for resources generated by members of such a community in one *domain* may be *routed* by the EGEE to resources in another domain that belongs to the community (i.e. US-grids, NorduGrid, Asia-Grid)
 - EGEE will enable the formation of inter-community grids, managed according to policies mutually defined by member communities.
 - Note that the EGEE not need be aware of the internal structure of each of the domains it services.
 - Each of these domains (and in some cases when the entire community appears as a single domain) may internally operate as grids.

Relationships with other entities(2)

- **Middleware providers** –will deliver software components, documentation, training materials and support that meet requirements and address bug reports provided by the EGEE. EGEE may also **facilitate testing and evaluation** of pre-release software versions.
- **Grid Research Projects** – EGEE will **share practice and experience information as well as needs for new technologies** with such projects and will seek information on their R&D directions and timelines
- **Funding Agencies** – as an entity established by the EU commission with a charter to serve the research activities funded by the commission as well as other funding bodies, the EU funding for EGEE may be augmented by funds from other sources. Each of these sources is likely to have **different review, reporting and auditing procedures.**



The EGEE Workforce must be flexible to adapt to each cycle of the project (no static allocation of all resources)

- **Management Team:**

- Responsible for overall management of the project and its resources,
- coordination and oversight of all teams.
- *All* contention and escalation procedures will eventually lead to a final decision by a mechanism run by this group.
- This group must retain substantial financial resource, e.g. funds for commissioning software from external providers, so that it can resolve or head off problems.

- **Requirements Team:**

Responsible for assimilating, tracking and specifying accurately the requirements arising from all stakeholders (operations team, user support team, end-user communities, provider sites, etc.).

- A combination of technical competence and interpersonal skills are required.
- They must steer the requirements capture to take due consideration of existing investment, working practices and feasibility.

- **Quality Assurance Team:**

Scrutiny of all activities of all teams to ensure that EGEE

- responds to stake holder requirements,
- sustains dependable services and
- delivers dependable software.

This includes ensuring that

- security is considered at all stages,
- designs and plans are technically and logistically feasible,
- testing and certification of releases is conducted properly
- all aspects of RAMPS (Reliability, Availability, Maintainability, Performance and Scale) are considered by all teams at appropriate stages.
- Consider balance with timely delivery

The normal resolution of a conflict is to reduce targets in the current cycle rather than run risks.

This team will be the final arbiter of risk assessment.

They require technically skilled staff and compute resources for inspection and testing.

- **Technology Acquisition Team:**

Responsible for all technology acquisition.

- this team is responsible for tracking and assessing the availability and quality of required components.
- This will include developing and running relevant stress tests.
- Their advice on the suitability of software to the strategy group must be based on firm evidence that they can vouch for.
- This group may also commission external software production and oversee the contracts before conducting their standard acquisition procedures.
- This team will also be responsible for developing and maintaining cooperative relations with other Grid projects and relevant standards organisations.

- **Strategy Team:**

This team will take as input the results of the requirements team, and the technology acquisition team, with appropriate scrutiny for quality.

- be aware of the current operational infrastructure and experiences of EGEE
- be responsible for steering work commitments so that they best serve EGEE's specified goals, balancing political, managerial and technical issues.
- It will **produce a prioritised and approximately sized** (in terms of effort required) **list of tasks** that are considered important candidates **for the next cycle**.
- It will inform the requirements team of its views on their requirements and alert the technology acquisition team to missing technology that prevented tasks being added to the priority list.

- **Design and Planning Team:**

- take the prioritised list from the strategy group, re-evaluate the effort required for tasks, taking into account skill and knowledge requirements, and **develop a safe and feasible plan** leading to the next release cycle.
- They will provide feedback to the strategy team and engage with implementation, integration and deployment to validate their assessment of feasibility.

- This plan will be agreed by review at a meeting representing all the stake-holders.

- are responsible for having **contingency plans** and being prepared with quantitative information about **alternatives**.
- produces a high-level design to guide implementation.
 - It therefore needs excellent technical skills, relevant experience (spanning all aspects of EGEE infrastructure and operations) and well defined software engineering processes.
- It will remain available for design consultations, planning revisions and component review, throughout the development to release cycle.
 - In particular, it is the final arbiter with respect to change management,
 - all proposals and impacts of changes must be referred to this group, who will adjudicate as to whether they can be incorporated in the current cycle.

- **Implementation Teams:**

These are teams of well-trained and well-managed software developers, with sufficient relevant experience in software engineering procedures.

- They will require a well-found development environment.
- They will all work according to standards and procedures defined by the Design and Planning team, e.g. with respect to software repository use, testing procedures and code review procedures.
- There may be from one to three such teams, each with a well-defined centre.
 - They may “bid” for the work identified by the design and planning team.
- The majority of work undertaken by these teams will be **incremental changes to existing parts of the software infrastructure** and improvements to the security and RAMPS properties of imported software to prepare it for deployment.

- **Integration Team:**

This team will take the output from implementation teams, combine them with the unchanging parts of the software infrastructure, and test the integrated whole.

- They must then package this ready for deployment and work with the deployment team, who will be a nominated part of the operations team, to assess whether the result is ready for deployment.
- The team will consist of well-managed software engineers with appropriate skills.
- This team will need appropriate computing resources for integration and testing.

- **Operations Team:**

This team has the full responsibility

- for the entire running software infrastructure,
 - for delivering a dependable service,
 - for diagnosing the causes of deficiencies in the service,
 - for supporting the user-support team,
 - for dealing directly with the resource providers (typically site managers) and
 - for feeding its priorities into the planning cycle.
- It will deploy new versions of software incrementally only when it judges it is appropriate to do so.
 - That is, the operations team **has final say over which software they use** to deliver the EGEE grid service.
 - It will also **have a major say in the addition of new sites and new user communities**, who are expected to increase the demands on this team, and may ask that participation of sites, user communities, virtual organisations or users be suspended while issues are addressed.
- This team has to establish a service that is trusted by all of the stakeholders.
 - This will include all aspects of security.
- To achieve this, the operations team must be empowered to take immediate remedial action when they judge a problem to be significant. **This is a major team that must be well resourced.**

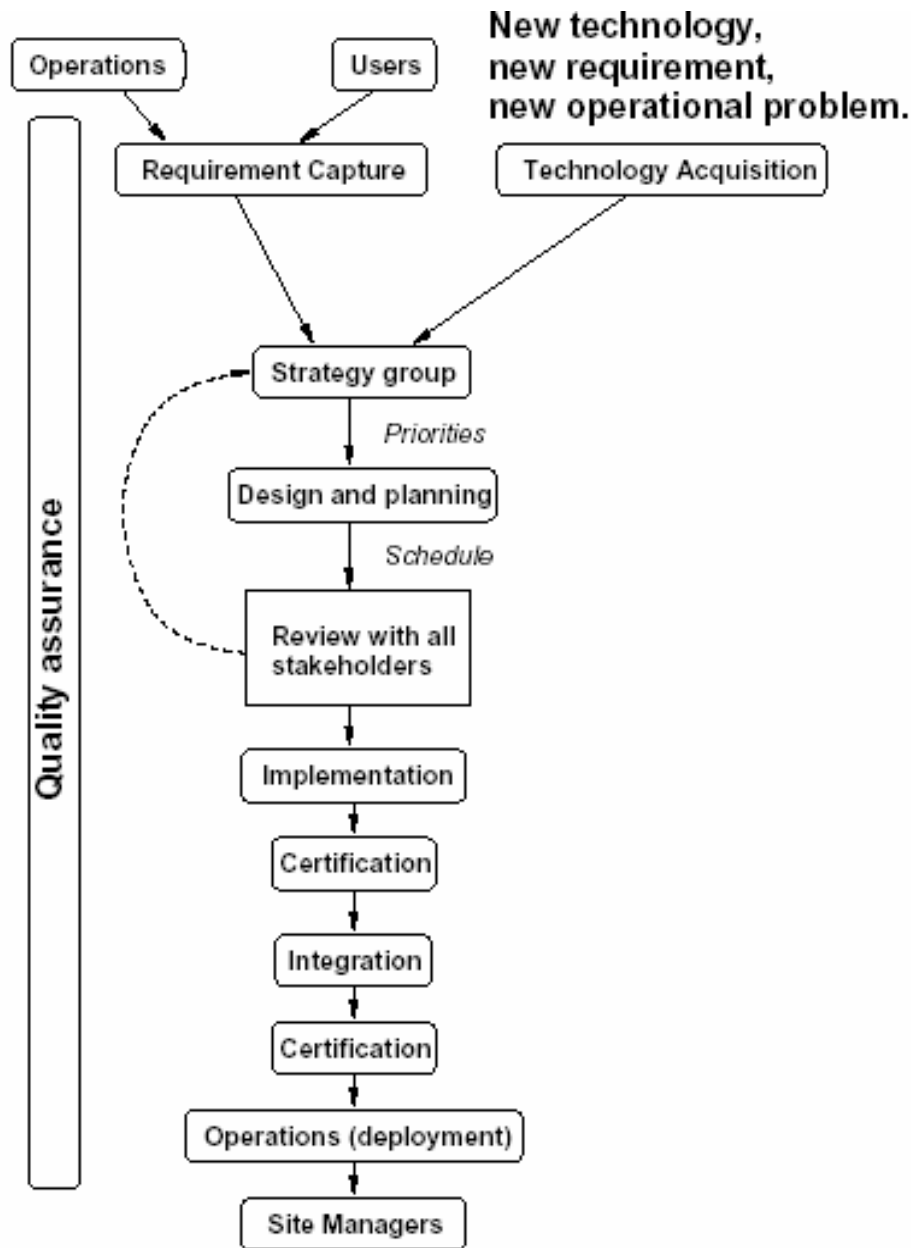
- **User Support, Training and Outreach Team:**
This team is responsible for the induction of new users, new communities and new virtual organisations, and of new members to all the teams above.
 - it must support appropriate, multi-lingual identification, registration, information and training facilities
 - in the case of new members of EGEE teams, also help in developing the appropriate quality-of-service-and-products ethos.
- It is also responsible for outreach;
 - that is identifying potential groups who may join as users, communities, industrial users, industrial partners, or industrial providers.
 - It should develop and disseminate appropriate information to these groups, including proactively addressing them.
- It will run a help desk and consultancy service primarily **dealing with expert representatives of communities.**
- It will assimilate and evaluate records of this work and provide information to the requirements and planning activities.

Work-package definition

- Each work package should identify **how it will deliver** its responsibilities and **how it will interact** with other work packages, and where relevant, with users, user communities, resource providers, external software providers and external service providers.

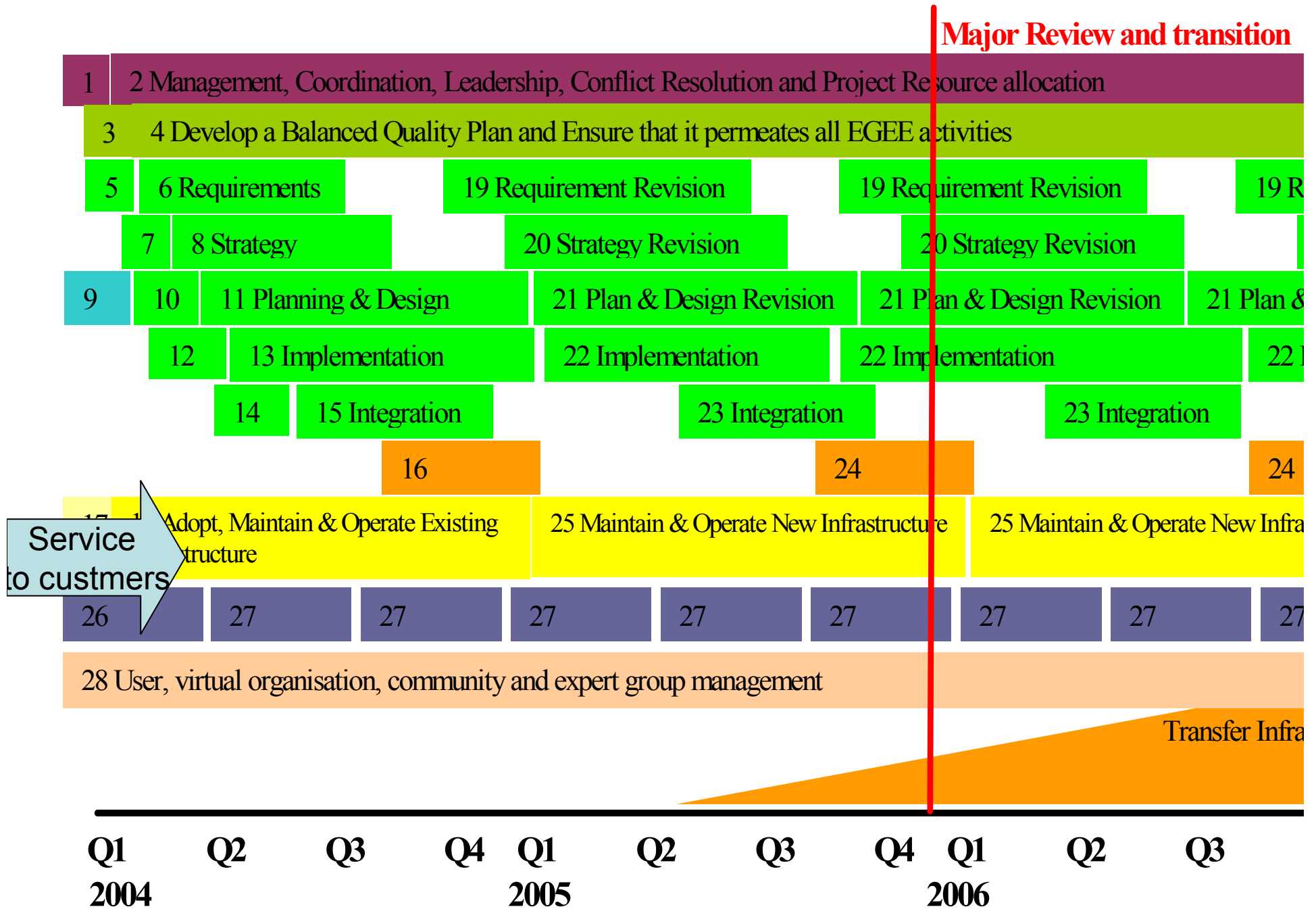
Work Package EGEE Team(s)

- | | |
|---|---|
| 1 | Quality Assurance |
| 2 | Requirements, Technology Acquisition and Strategy |
| 3 | Design and Planning |
| 4 | Implementation and Integration |
| 5 | Operations, Infrastructure Call Centre, Site Managers |
| 6 | User Support, Training and Outreach |

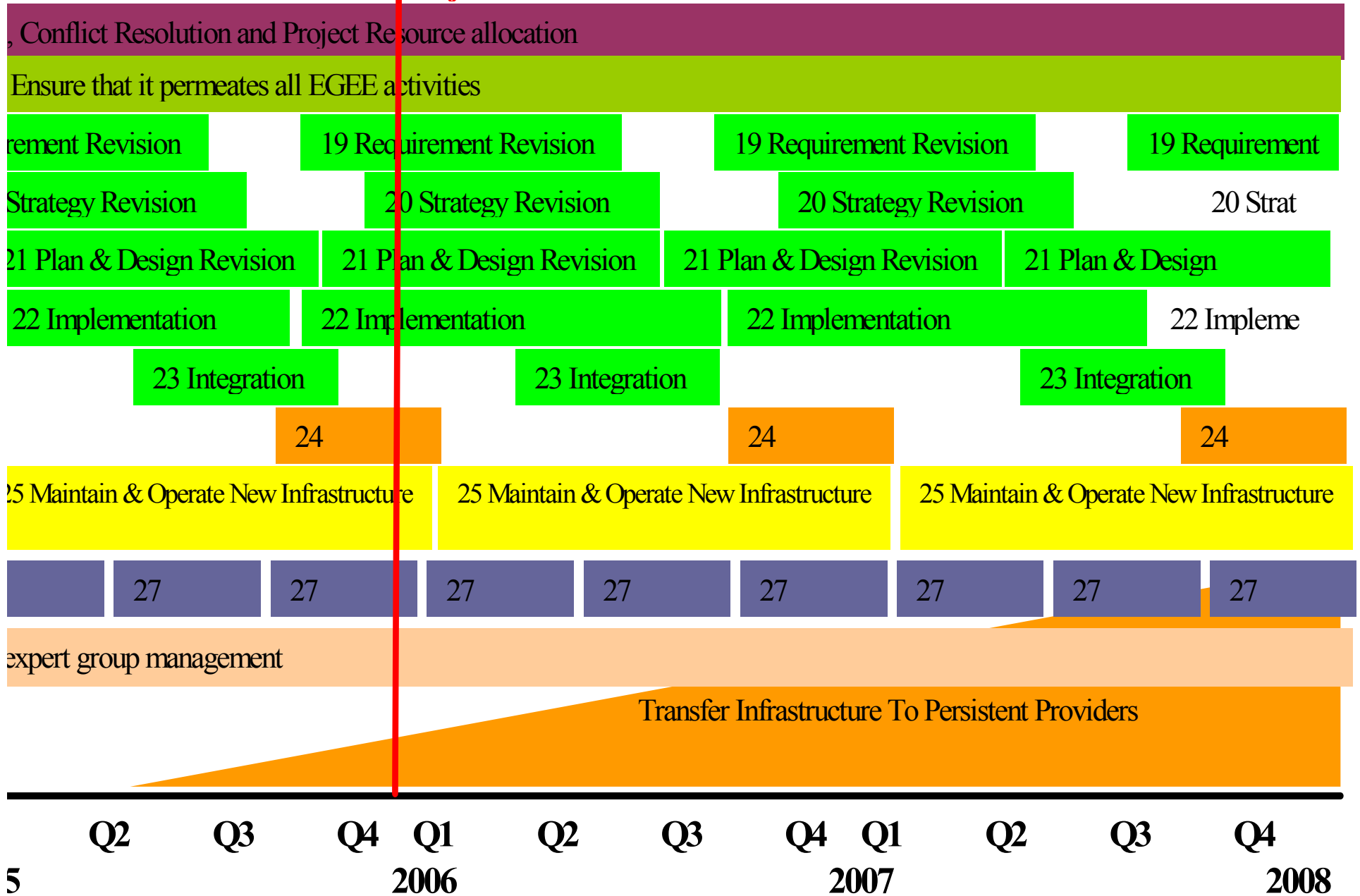


EGEE Cycle

- Goal:
 - Deliver quality of service
 - Deliver stable service
 - Plan changes carefully
 - Changes at predefined release times
 - Apart from bug-fixes
 - Maintain (rolling) 5 year plan
 - Even beyond EGEE lifetime



Major Review and transition



1. Establish the management team, including full-time managers
2. Manage all aspects of EGEE
3. Establish the QA team
4. QA team prepares a quality plan and interacts with all EGEE project processes to ensure quality
5. Establish the requirements team
6. Perform Cycle 1 Requirements capture and analysis (with input from previous projects)
7. Establish Strategy Team
8. Develop strategy, priority list a putative five-year plan for Cycle 1 (input from technology acquisition not shown)
9. Establish the computational infrastructure for development, e.g. build and test servers and CVS repository
10. Establish Design & Planning Team
11. Perform Cycle 1 of Design & Planning (feedback loop to Req Team not shown)
12. Establish the Implementation Team(s)
13. Implementation Team implements according to Cycle 1 Design & Plan (feedback to design & planning and QA certification of components not shown).
14. Establish Integration Team
15. Integration Team prepare Version 1 of Production Platform for Deployment (feedback to planning team and certification of quality of platform not shown)
16. Operations team conduct handover and prepare to deploy "New Infrastructure"
17. Establish the operations team
18. Adopt a viable Grid software infrastructure and policies from previous projects and develop a service using it.
19. Conduct requirements capture for Cycles 2, 3, 4, ...
20. Revise priority list and long-term plan for Cycles 2, 3, 4, ... (work of technology acquisition team not shown).
21. Revise detailed plans and designs progressively for Cycles 2, 3, 4, ...
22. Develop software infrastructure, including ruggedising imported components, incrementally improving existing components and thorough component testing for Cycles 2, 3, 4, ...
23. Integrate, test and prepare for deployment new software infrastructure for Cycles 2, 3, 4, ...
24. Operations team manage handover of software infrastructure produced by Cycle 2 to be operated in Cycle 3, and so on.
25. Operations team manage the use of the current infrastructure and use it to provide services during Cycles 2, 3 & 4.
26. Establish the User Support, Training and Outreach Team. It should build an EGEE team spirit and establish inter-team communication by holding a series of events to train members of EGEE.
27. Run various training events, induction events, industrial awareness events, researcher awareness events, supporting these with tutorial material and including illustrative simple applications using EGEE.
28. The operations team and the user support, training and outreach team, set up and operate combined mechanisms to support the process of disciplines, expert communities, users, and providers joining EGEE. Expert user communities should be encouraged and supported via formal arrangements with each community's own user-support team.

Metrics for success: quantitative

- Goal: deliver substantial and long-term benefits to a broad spectrum of EU researchers: this can be counted!!
- EGEE should not be evaluated primarily by documents, papers etc...
- We propose an evaluation after 2 and 4 years,
 - passing the 2 year evaluation makes a convincing case for continuation.

Q4 2005

- Users
 - > 3000 Certified users
 - > 15% Peer-Approved Allocation
 - > 5 Disciplines
 - > 15 Countries
- Providers
 - > 10K Processors
 - > 5 PB
 - > 10 Countries, > 50 sites
- Dependability
 - MTBF > 5 Days
 - Availability > 95% of 24x7 last Q
 - > 10% of operations guaranteed > 5 years
- Usability – Training + Platforms
 - Scientific Application Developers
 - Induction of > 200 / year
 - EGEE new Operators > 25 / year
 - EGEE new Infrastructure > 20 / year

Q4 2007

- Users
 - > 5000 Certified users
 - > 50% Peer-Approved Allocation
 - > 10 Disciplines
 - > 30 Countries
- Providers
 - > 30K Processors
 - > 15 PB
 - > 15 Countries, > 100 sites
- Dependability
 - MTBF > 30 Days
 - Availability > 99% of 24x7 last Q
 - > 50% of operations guaranteed > 5 years
- Usability – Training + Platforms
 - Scientific Application Developers
 - Induction of > 300 / year
 - EGEE Operations > 20 / year
 - EGEE new Infrastructure > 20 / year

Proposed technical components in production quality

1. **Resource Access Service**

Provides an access channel to a given resource for properly authorized work requests. Possible features include advance reservation/schedule publishing, pre-emption capabilities, etc.

- Related examples: GRAM, CondorG

2. **Data/Application Repository and Access/Deploy Service**

Provides a way to uniquely identify, store and retrieve data. This can apply both to the application execution sandbox (executables, configuration, etc.) and/or to input and output data.

- Related examples: GDMP, EDG WP2+5

Proposed technical components in production quality

3. Information Collection Service

Provides a way to collect and access the description and status of a set of grid objects (applications, computing, storage, network elements).

- Related examples: MDS, R-GMA, EDG WP3

4. User ID and Authentication Service

Provides grid-wide identification of users, along with information about their membership and capabilities within various organizations or groups.

- Related examples: GSI, EDG VOMS

Proposed technical components in production quality

5. Resource Discovery and Brokering Service

Provides the selection of resources matching a given request, based on information from (3) and (4).

- Related examples: Condor central manager, EDG WP1 Resource Broker, Eurogrid services

6. Accounting Service

Collects resource usage records, organizing them into user/organization accounts that are made accessible for queries.

- Related examples: NIMROD-G, EDG WP1 DGAS

A series of properties need to be watched and assured across these services, towards the production quality goal of EGEE:

- Fault tolerance
- Service authorization policy
- Generalized adoption of authentication/privacy techniques that are acceptable to all stakeholders
- Consistent public interface
- Consistent and ample status/exception reporting

- **Grid Support Centres:**

These should be well staffed and operated in order to provide a 24x7 service.

- These major centres should be located at sites that already have expertise and experience in deploying and operating grids, and supporting large or multi-disciplinary user communities.
- Their functions would include providing the following services:
 - Operate basic grid infrastructure services – for example, those described in the previous section
 - Performance monitoring, troubleshooting, problem resolution
 - Monitoring tools need development
 - Grid configuration management
 - Resource management

- **Call centres.**

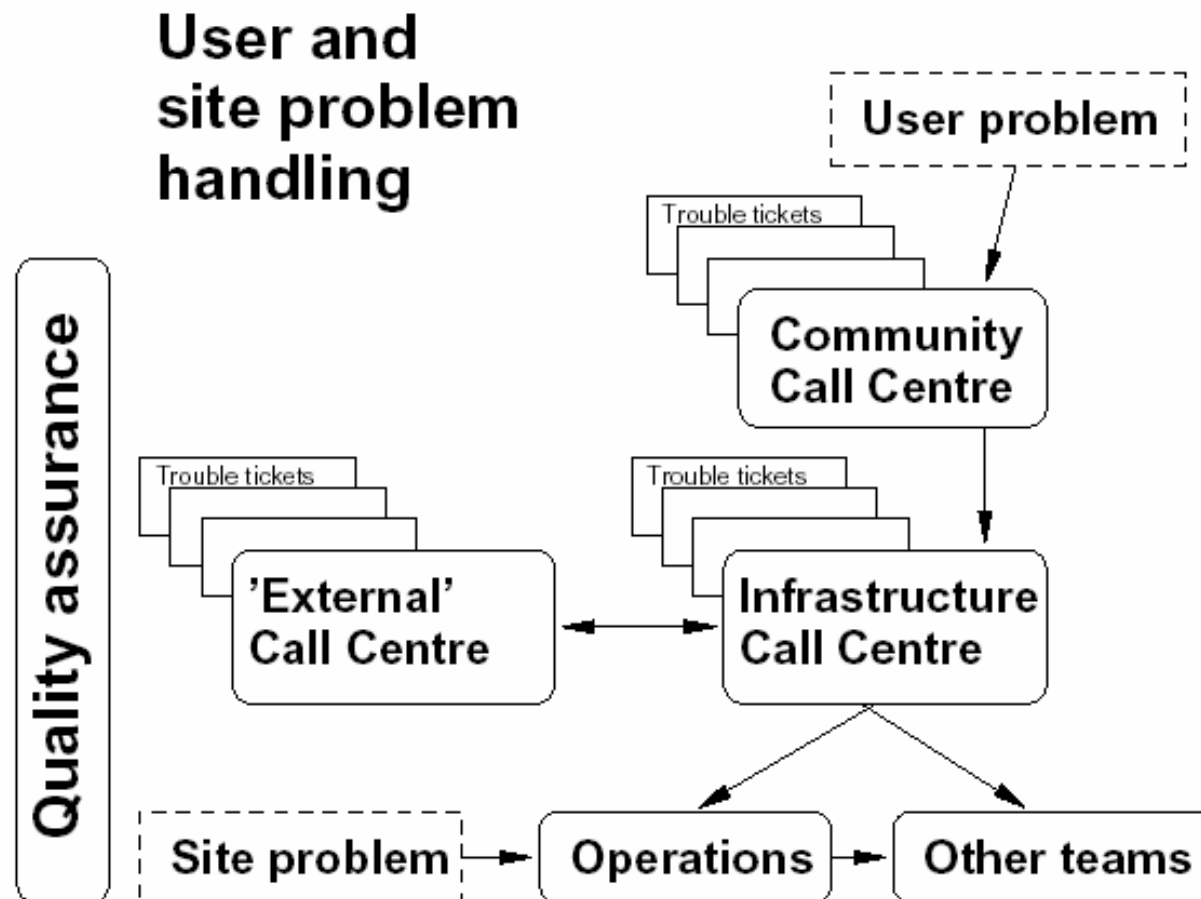
The call centres should provide the following basic set of services:

- User support and helpdesk.
 - must tie in to the local user support structures at participating sites
 - but must provide a central layer of coordination.
 - The local user support services can be seen as gateways into the system.
 - The user support process must provide the user with a single point of contact.
 - enable the tracking of problems across multiple domains of responsibility.
 - The user communities will frequently require access to resources that extend across multiple resource domains (of which EGEE will be one).
 - A user community should provide a support centre for its users and makes a decision on responsibility for the problem and forwards it to the infrastructure (EGEE or other) support organisation.
 - There must be well defined processes for EGEE to collaborate with its peers, with the site operators, with network operators, etc.
- Provision of documentation and training for end users, developers, service providers or site managers, and to the project work packages.
- Provide support interfaces to providers and application developers
- Act as centres of expertise in grid middleware to provide facilities and resources to assist in debugging, middleware/application problem solving, etc.

- **Points of Presence:**

These might be staffed with 1-2 people and provide a business-hours user support and technical advice service

- in geographical regions not served by one of the main Grid Support Centres or Call Centres.



The actual production operation model of a grid must be an essential component of the EGEE

- model of operations should be allowed to evolve.
- requires a flexible and agile management policy.

Incremental change management while maintaining a sustained service

- Upgrade process for deployment of middleware releases.
 - Must be possible at the granularity of individual services.
 - It is essential that the overall grid service be available continuously even during upgrades.

Middleware deployment process:

- Testing and certification process of the middleware and of the participating sites.
- The deployment process will use a sequence of testbeds for development, integration, and certification; before the middleware is put into production.
- The deployment process must provide tools and facilities to control the packaging, distribution, and configuration of the middleware services:
 - The middleware should be deployable using site-specific (and existing) installation tools
 - Provision must be made for the separate distribution of middleware and applications, and must include the process of publishing what is available

Process for new sites joining EGEE and bringing new resources – site deployment and integration

- It is essential in providing a stable service that each site is certified as it joins the grid.
- This process should include verifying the configuration and ensuring conformance to the operational standards.

Process for bringing in new user scientific disciplines,

- Describes the mechanism by which new user communities join the EGEE,
- should include consideration of contribution of resources, agreeing policies and so on.

Process for including new applications (or new groups) within a discipline,

- Describes how a new application area within a community can become part of EGEE and what the requirements are.

Service quality assurance and monitoring process

- This will require the development of the appropriate tools and metrics.

Resource usage accounting

- is an essential activity that ensures that the policies of the customer applications or domains were implemented by EGEE
- This is an area where much development is needed.

Operational security procedures are essential.

- Security policies must be developed from the outset
- provision made for adequate security operations and management
 - including such things as **incident handling**.

Executive Summary: Goals and Scope(1)

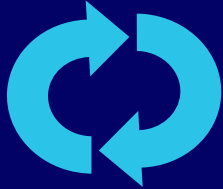
- Set up and operate a multi-disciplinary science Grid infrastructure for scientific research.
 - Initial applications will come from High Energy Physics and from Biomedical sciences.
 - The Grid production infrastructure should be deployed on hardware which is planned for big scientific projects and dispersed laboratory systems.
 - The EGEE operations and support teams should be based on (=leveraged on) plans and resources foreseen in the established scientific communities.
- Facilitate the application of Grid technologies to scientific and industrial groups
 - disseminate experience,
 - support a suite of high quality grid middleware
 - providing start-up resources.
 - It should work with user communities through managed alliances;
 - new communities and applications should be attracted.
 - EGEE should arrange the creation of industrial outreach organizations.
 - To encourage the search for novel Grid applications it is suggested to **establish an annual award** for the most advanced, most challenging and most creative new Grid applications.

Executive Summary: Goals and Scope(2)

- Deliver and support a high quality production Grid software suite with a consistent and stable API for user applications.
 - by building on existing technology through a well defined technology acquisition, support and maintenance process.
 - The emphasis should be on the provision of a stable operating environment and a carefully planned extension of functionality.
- The EGEE project- and management structure must be set up to maintain flexibility throughout the project to accomplish the most economic/efficient/appropriate assignments to reach this goal.

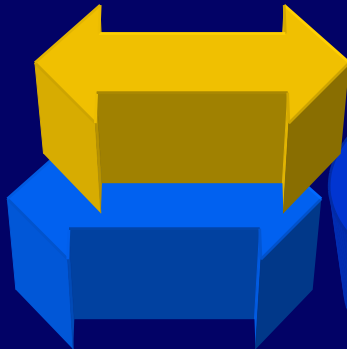
GÉANT and GRIDs: The model

GRIDs use
GÉANT infrastructure



GÉANT profits from
technological innovation

GRIDs employ
GÉANT infrastructure



International
dimension

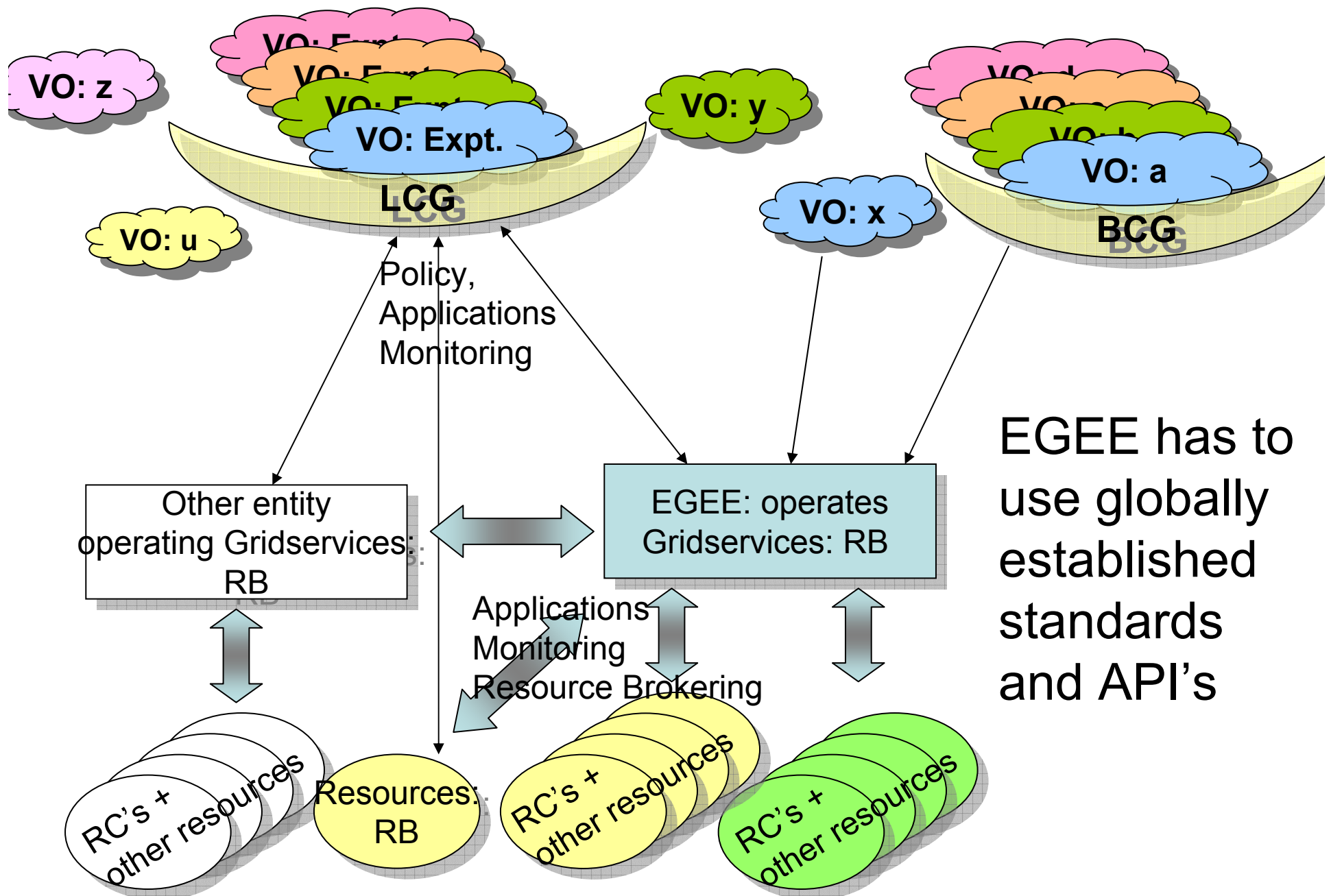
Managed alliances:
i.e. LCG

Application areas

GRIDs platforms

GÉANT network

R&D on GRIDs



EGEE has to use globally established standards and API's

LCG will be able to contract out to EGEE part of the operation of Grid infrastructure