

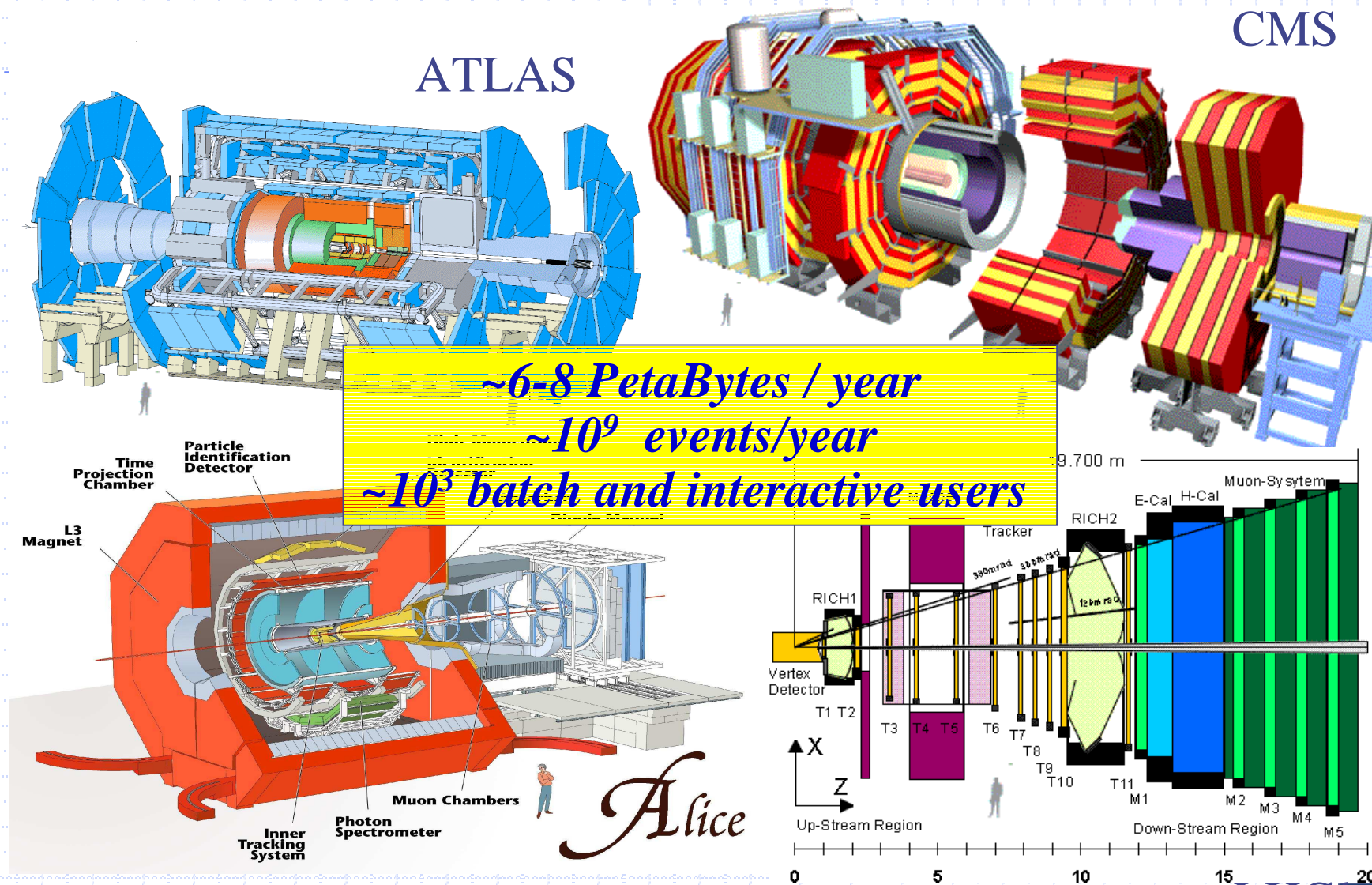
Experiments needs of Grid technology and infrastructure

LHC experiments

Geneva, February 22, 2003



A reminder of the problem



February 22, 2003

EGEE town meeting

LHC

2



Objectives of Phase I (02-05)

- ◆ Prepare the computing environment for the analysis of LHC data
 - Including applications, GRID MW and infrastructure
- ◆ Deploy and coordinate a global grid service
 - Acquire, deploy and operate robust and maintainable middleware
 - Requirements in HEPCAL
www.cern.ch/lcg/sc2/rtag4/finalreport.doc
- ◆ 2H03 service ramp-up
 - batch service, data management, reliability, operability, scaling & performance
- ◆ 1H04 experiment data challenges
- ◆ 2H04 full LCG Pilot
 - Fulfill HEPCAL + prototype interactive analysis
- ◆ 1H05 Technical Design Report for Phase 2



Centres taking part in LCG-1



around the world → around the clock

February 22, 2003

EGEE town meeting

4



Current status

- ◆ GRID middleware exists and is not vapourware
- ◆ Experiments [have used | are using it] for realistic productions, however
 - Basic reliability & functionality problems exist
 - Some HEPCAL “simple” requirements are far from being satisfied
 - Several system level issues are not yet addressed
- ◆ We still have a long way to go to get to a solid service
 - LCG-1 plan looks now ambitious, and this is worrying
 - Some of the advanced functionality we need is addressed only now

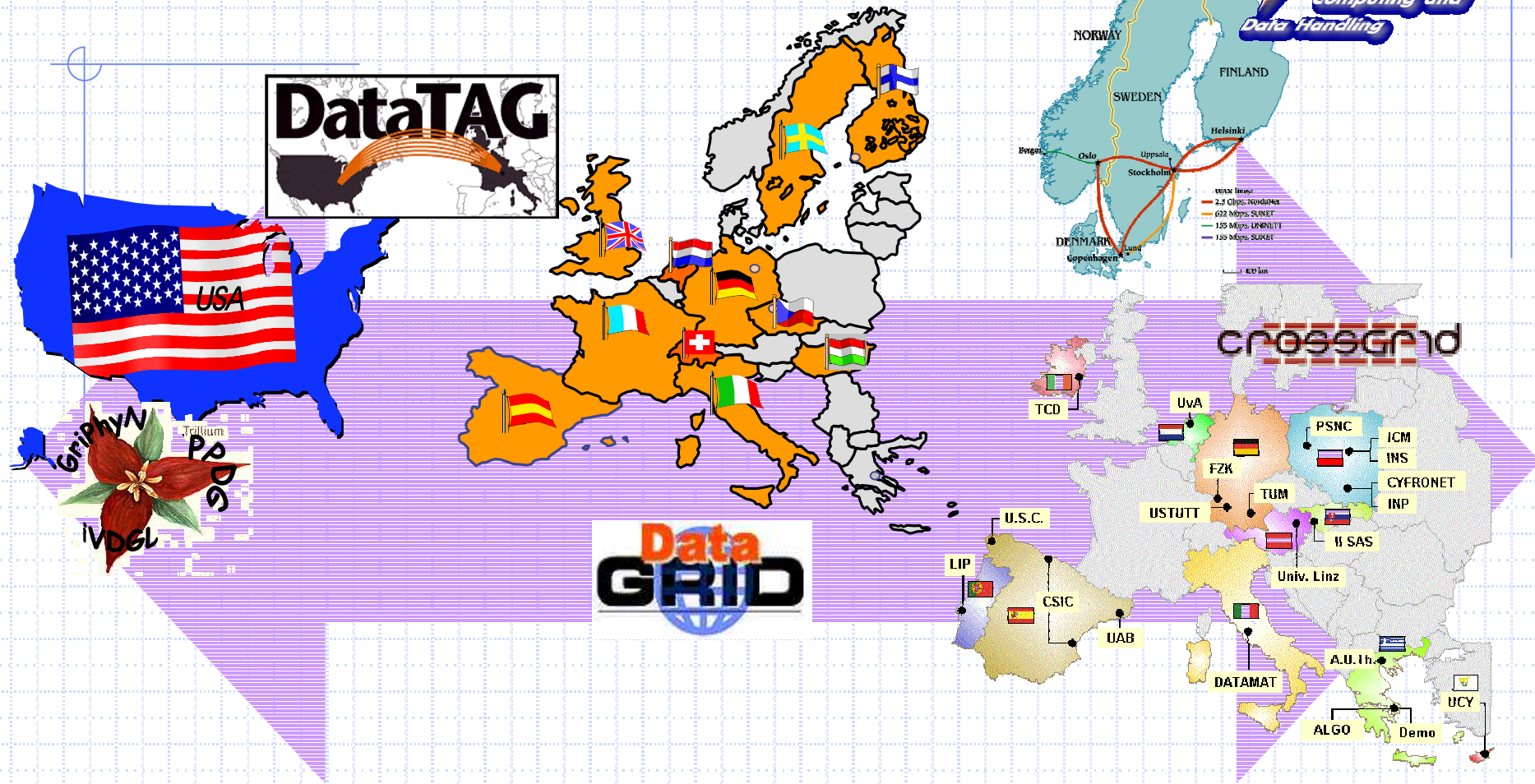


Current status (cont.)

- ◆ Middleware projects are collaborating only partially
 - Substantial duplication is present (could be beneficial initially, but is becoming wasteful...)
 - Most of them are short-lived
- ◆ Experiments have developed “higher functionality” middleware
 - AliEn, Dirac, Ganga, Grappa, Magda, Boss, Impala
 - Some solutions are surprisingly (or may be not) similar
- ◆ We need robustness and simplicity
 - Remember? GRID is about “seamless and reliable access to high-end resources”
 - It is easy to expand scope, much harder to contract it ! (D.Foster)
- ◆ How do we achieve this?



Related Grid projects



The GRID empire is developing... but it might strike back

Current status (cont.)

- ◆ A tremendous experience has been gained
 - See for instance the EDG testbed
 - From all the parties, i.e. middleware, users and sysadmins
 - Arguably some of the mistakes made were unavoidable

An expert is a man who has made all the possible mistakes in a restricted field
- ◆ We have learnt a lot from experience
 - But we also understand that it is very hard!
- ◆ Time is now short, and we cannot afford more “faux pas”
 - Duplication must be avoided (e.g. EDG vs. LCG-1 testbed)
 - ◆ Too much emphasis on interoperability may constrain evolution
 - But also too ambitious programmes and functionality
 - ◆ The EDG EU reviewers “Congratulates the project management for taking the risk of concentrating on quality”



LCG-1 priorities

- ◆ Develop a stable infrastructure providing basic functionality
 - Reference to HEPCAL + operational needs
 - From batch production to batch and then interactive analysis
- ◆ Converge on a set of MW tools that could evolve into Phase-2
 - Or be easily replaced if necessary
- ◆ Limit / avoid duplication of efforts
- ◆ Continue understanding experiment needs
 - HEPCAL II may be necessary here
- ◆ Understand experiment need for “higher level” or “HEP specific” middleware
 - See why it was developed and try to coordinate it
 - Or “push” some of it into MW development
- ◆ Focus on well understood / documented functionality, stability and simplicity



EGEE Enabling Grids for E-science in Europe

(how it has been presented to us)

◆ Goal

- create a general European Grid production quality infrastructure on top of present and future EU RN infrastructure

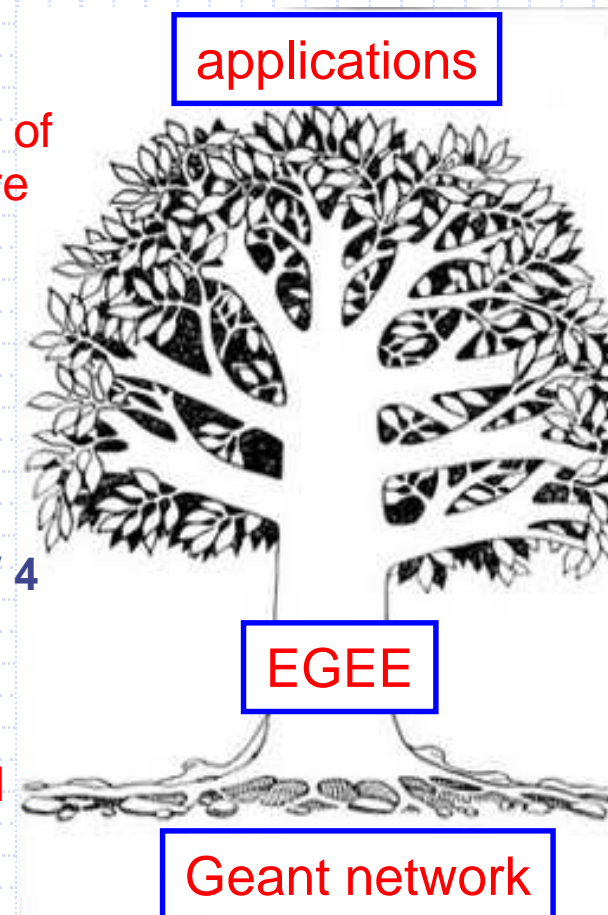
◆ Build on

- EU and EU member states major investment in Grid Technology
- Several pioneering prototype results
- Largest Grid development team in the world

◆ Goal can be achieved for about €100m/ 4 years on top of the national and regional initiatives

◆ Approach

- Leverage current and planned national and regional Grid programmes (e.g. LCG)
- Work closely with relevant industrial Grid developers, NRNs and US



Integrated Infrastructure Initiative (I3)

- ◆ Three lines of funding supported (with possible budget breakdown)
 - Networking activities (nothing to do with networks.):
 - ◆ This is the overhead: management, coordination, dissemination and outreach (7-10% of the total funding)
 - Specific service activities:
 - ◆ Provision and procurement of Grid services (60% of total funding)
 - Joint research activity
 - ◆ Engineering development to improve the services provided by the Grid infrastructure (20% of total funding)
 - ◆ Application support and focused R&D (10% of total funding)



EGEE proposal timeline

- ◆ Tentative Schedule (continued)
- ◆ EU call out on Dec 17th
- ◆ Draft 1: overall project structure end of February 2003
- ◆ Draft 2: with detailed workpackages end of March 2003
- ◆ Final proposal including admin and management end of April 2003
- ◆ Submission by May 6th 2003
- ◆ First feedback from EU in June-July
- ◆ Contract negotiation late summer, fall '03
- ◆ Contract signature by the end of '03
- ◆ Start of project Q1-Q2 '04



EGEE & LCG -- opportunities

- ◆ EGEE can be an unique opportunity to
 - Build a well operated testbed
 - Provide the necessary personnel to harden (simplify?) the existing MW
 - Provide quality MW using the existing middleware R&D
- ◆ The EGEE operated LCG testbed has the potential to provide a focal point for convergence
 - Of different user communities within LHC
 - Of different middleware projects working with LHC
- ◆ However such a large project has potential dangers
 - Which are proportionally big!



EGEE & LCG -- caveats

◆ Divergence

- US MW projects and US experiment groups MUST “buy in”
 - ◆ Cooperate with EGEE / LCG and ensure complementarity
- Application requirements have to be coordinated and *controlled*
 - ◆ All sciences involved have to get their requirements into the Program of Work
 - ◆ But this should not lead to too much divergence
- All efforts should converge on the same testbed
 - ◆ We are seeing with EDG-LCG-1 that the operational interference of several testbeds is destructive on the LHC Regional Centre staff
 - ◆ EGEE and LCG-x testbed must coincide (software and hardware) in all sites that belong to both



EGEE & LCG -- caveats

- ◆ Rewriting may not be the only option
 - Review current middleware packages with respect to LCG requirements
→ need repackage, simplify, interoperate, eliminate duplicates ??
 - An architecture would be extremely helpful (components, functionality, API's, protocols)
 - However remember that scrapping software is not failure providing you retain the knowledge -- that would be a good start
- ◆ Overhead & timing
 - Planning has to be carefully done as we cannot afford the overhead of running two large projects, supporting two planning/reporting/review processes
 - EGEE timing should be largely in line with LCG timing
- ◆ Resources
 - All this has significant costs, EGEE can probably cover it, but only if things are done right from the start
 - If CERN becomes a e-science competence centre this should not be to the detriment of LHC!



EGEE & LCG -- caveats

◆ Requirements

- Experiments must make sure that their requirements make it into the Programme of Work
- But we should be realistic
 - ◆ *Asking for the moon* will not work, no matter how much manpower is there
- LHC experiments must be involved in the definition of the workpackages and of their goals
- But to stand a chance to be heard LHC experiments should speak with a single voice (GAG has been setup for this)



Experiment participation

- ◆ Even in the best of all worlds EGEE will draw on experiment resources
 - Installation of software, testing and evaluation of EGEE
 - Necessary participation into the project bodies
 - Collaboration with the different components of the project, in particular MW
- ◆ This is not an overhead imposed by EGEE
- ◆ It is necessary manpower that we need to build LCG-x
- ◆ But EGEE can and must *compensate* for this
 - Failure to secure this manpower would make the participation of experiments into EGEE impossible
 - And therefore would reduce / eliminate the interest of the whole project for LHC



Experiment participation

- ◆ We need a “WP8” inside EGEE
- ◆ A HEP application work package
 - Some Experiment Independent People (~ 4) and some additional personnel into the experiments (1-2 people per experiment)
 - Build on the knowledgeable, experienced team within EDG – “loose” cannons
 - Provide support to experiments on the EGEE testbed for installation, evaluation, problem reporting, liaison with the other workpackages
- ◆ The EDG experience shows that this is essential
 - Only with such a body the experiments will be able to make the most out of the testbed, properly evaluating it and providing qualified feedback



Relation with GAG

- ◆ GAG will continue its work in parallel to EGEE
 - Requirement definition and refinement at a more “abstract” level without getting directly involved with the testbed
 - Look for commonalities in experiment “high level middleware”
 - Official representation in EGEE for LHC requirements
 - Involved in all the phases of the preparation of the EGEE workplan
- ◆ It is important that experiments are represented
 - But they must have a common representation
- ◆ GAG has been formed to create a common viewpoint of the experiments on GRID



Conclusions

- ◆ To avoid dispersion and divergence experiments will have to interact in a highly coherent way with EGEE
 - GAG will act as the LCG forum for developing & monitoring the common requirements → strong i/p to EGEE (& ITR)
- ◆ Experiments need extra support to evaluate the testbed and provide qualified feedback
 - EDG has shown that a WP8-like structure is necessary and must be properly manned
 - EDG type “loose cannons” essential for a coherent implementation & evaluation
- ◆ EGEE has the potential to be a great success, as we have the expertise and the experience
 - EDG has shown its necessary to have an upfront architecture
 - Essential to have well-described, comprehensive set of use cases

