Technical Status

Erwin Laure

4th EGEE Conference

Pisa, Italy

24th October 2005

Contents

- Major advances since Athens
- Overview of the conference programme
- Priorities through the end of the first phase

What have we done since Athens?

- ***Further increased the infrastructure and its usage***
  - *Sustained rate of ~**10.000 jobs/day***
  - *Biomed data challenge (WISDOM) – over **80 CPU years in ~6 weeks***
  - ***84 VOs** supported on the production infrastructure (**22 VOs** >1000 CPUh/m in average)*
  - *Number of NA4 **users doubled** over the past 9 months (~1000 in PM 18)*
- ***Improved middleware***
  - *LCG-2 stack further improved, gradually including gLite components*
- ***Extensive training***
  - *Summer schools & outreach to AP*
- ***Wide dissemination to the media and academic/scientific world***
  - *Over 300 conferences, 230 press cuttings, 9 TV and 5 Radio interviews*
- ***International cooperation and global relevance***
  - *IGTF (Int. Grid Trust Federation)*
  - *Interoperability efforts with OSG, NorduGrid intensified*
- ***Proposal for EGEE-II***
  - *EU hearing on November 7, 2005*

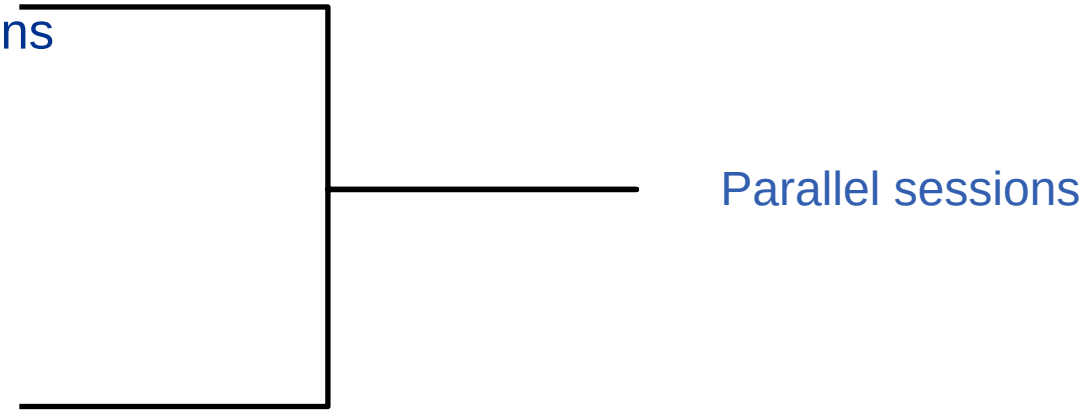
EGEE has clearly moved into production mode and is well prepared for the second phase

Thanks to everyone's hard work

Conference Programme

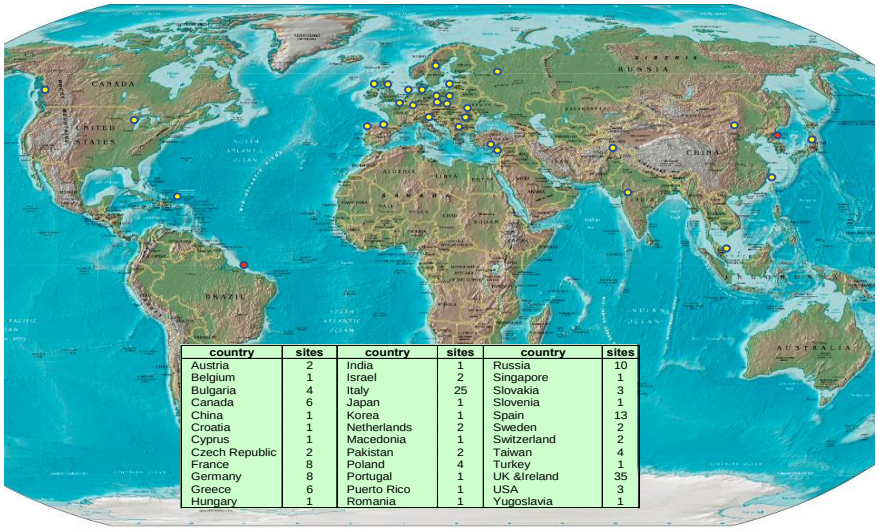
Prepare for the focused review in December and a smooth transition to EGEE-II

- **Monday (PM)**
  - Overview of project status and directions + related projects + second phase (EGEE-II)
- **Tuesday**
  - Related projects (SEEGRID, Diligent, DEISA, OSG)
  - Technical plenary panel – please post questions on the website!
  - Industry plenary session
  - Afternoon parallel sessions
- **Wednesday**
  - Parallel sessions
  - Demos
- **Thursday**
  - Parallel sessions
- **Friday (AM)**
  - Summaries of parallel sessions



SA1: Grid Operations, Mgmt & Support

- **Scale of the production service**
  - ***Den Haag***: ~8k CPUs/80 sites ***Athens***: ~14k CPUs/130 sites ***Pisa:*** ~16K CPUs/170 sites
    - **This greatly exceeds the no. sites planned for the end of EGEE**
    - **We now entered the consolidation phase after an impressive growth**
  - Continuous improvements to LCG-2 middleware
    - **LCG-2.6 recently released**
    - **Contains FTS, VOMS,**
    - **R-GMA from gLite**
- **Preproduction Service and Certification in place**
  - 14 sites (PPS), 4 virtual sites (cert)
- **Interoperability demonstrated with OSG (ongoing work with ARC)**



country	sites	country	sites	country	sites
Austria	2	India	1	Russia	10
Belgium	1	Israel	2	Singapore	1
Bulgaria	4	Italy	25	Slovakia	3
Canada	6	Japan	1	Slovenia	1
China	1	Korea	1	Spain	13
Croatia	1	Netherlands	2	Sweden	2
Cyprus	1	Macedonia	1	Switzerland	2
Czech Republic	2	Pakistan	2	Taiwan	4
France	8	Poland	4	Turkey	1
Germany	8	Portugal	1	UK &Ireland	35
Greece	6	Puerto Rico	1	USA	3
Hungary	1	Romania	1	Yugoslavia	1

- **Operational procedures are continuously improved as we go along**

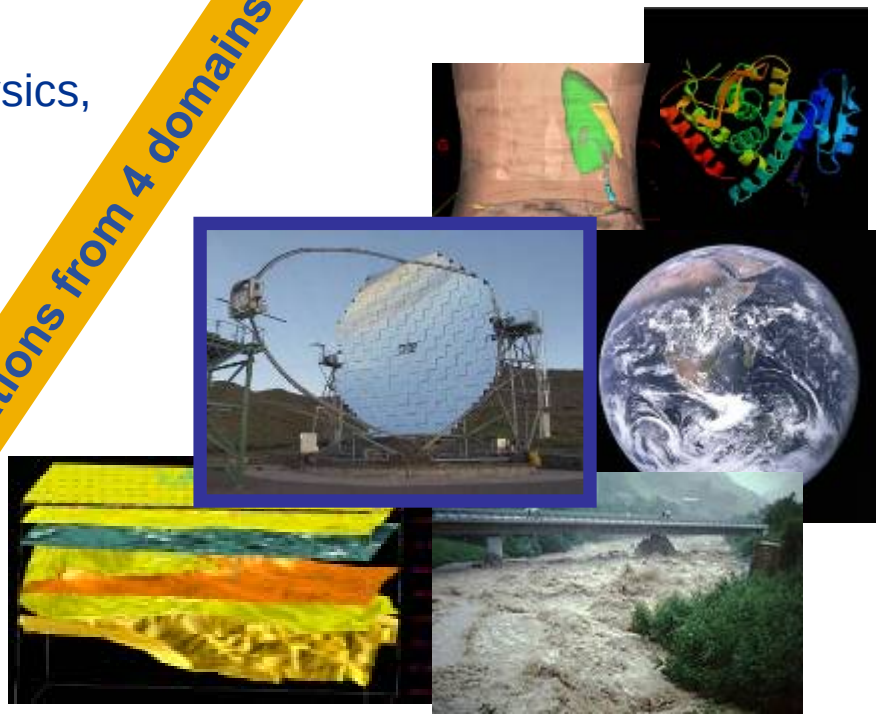
Network Resource Provision & Research (SA2 & JRA4)

- **Closer interactions between GGUS and NRENs (ENOC - EGEE Network Operation Center trial) during summer**
- **More work on network service provisioning model and bandwidth allocation and reservation specifications and prototype**
- **Network QoS experiment finalized**
- **Network performance diagnostic tool for ROCs and CICs produced**
- **Further work**
  - Further intensify the collaboration between GGUS and NRENs
  - Establish SLAs for allocation and reservation and integrate with GEANT2 service provisioning infrastructure
  - Deploy network monitoring and make it also available to middleware

NA4 Application Support

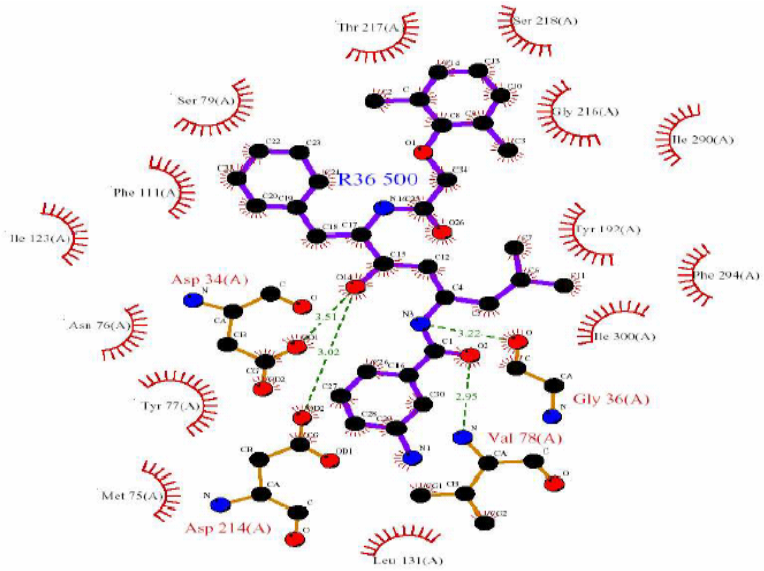
- **More than 20 applications from 7 domains**
  - High Energy Physics
    - 4 LHC experiments (ALICE, ATLAS, CMS, LHCb)
    - BaBar, CDF, DØ, ZEUS
  - Biomedicine
    - Bioinformatics (Drug Discovery, GPS@, Xmipp_M fine, etc.)
    - Medical imaging (GATE, CDSS, gPTM3D, SiM D, etc.)
  - Earth Sciences
    - Earth Observation, Solid Earth Physics, Hydrology, Climate
  - Computational Chemistry
  - Astronomy
    - MAGIC
    - Planck
  - Geo-Physics
    - EGEODE
  - Financial Simulation
    - E-GRID

Another 8 applications from 4 domains are in evaluation stage



NA4 Pilot: Biomed

- **Production: Biomed Data Challenge on molecular docking**
  - >80 CPU years in ~6 weeks
  - >40 Kjobs, 60 Kfiles produced (~1TB)
  - ~46M docked ligands
  - High cost in human resources



In progress

Medical files management (DICOM)
Metadata management (AMGA)
Security (file encryption)
Grid workflows processing

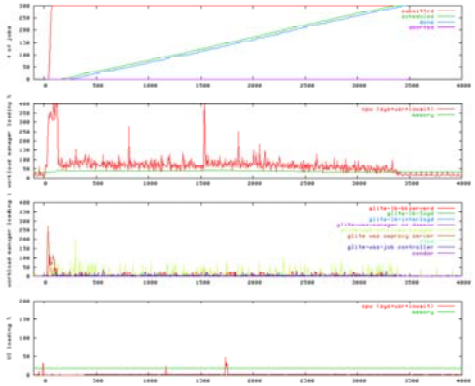
Next priorities

Efficient processing of short jobs
Fine grain access control (data and metadata)

NA4 Pilot: HEP

gLite test (e.g. WMS within the ATLAS Task Force)

First Measurement on gLite 1.4



- First gLite 1.4 WMS performance measurement in Milan
  - gLite 1.4 WMS (4CPU, 3GB memory)
  - 300 simple hello world jobs submitted by 3 parallel threads
  - submission rate ~ 4.1 jobs/sec
  - dispatching rate ~ 0.08 jobs/sec
  - All jobs in bulk are submitted to the same CE: atlasce.lnf.infn.it
- Thanks to Elisabetta Molinari for setting up the gLite WMS in Milan

30 Sept. 2005 LCG/EGEE Taskforce Meeting

LHCb production integrating record CPU usage

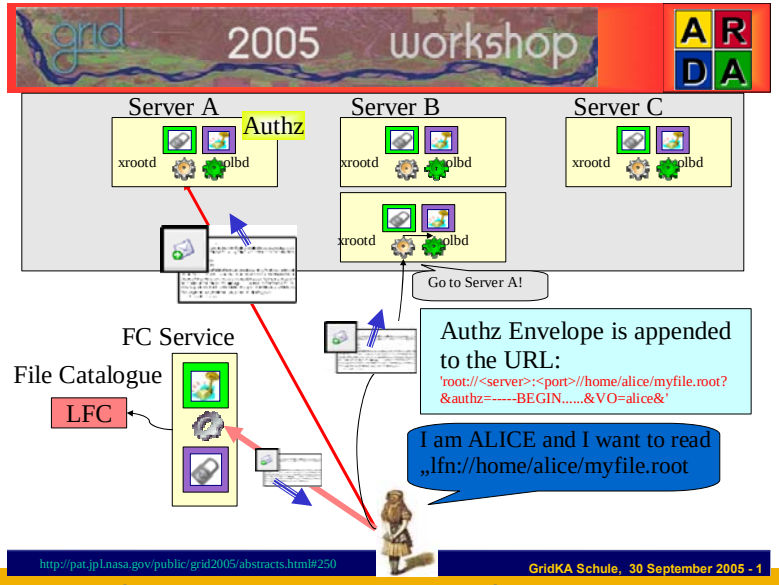
CPU used: 6,389,638 h
Data Output: 77 TB

DIRAC.Barcelona.es 0.214% | DIRAC.Bologna-T2.it 0.696%
DIRAC.CERN.ch 0.571% | DIRAC.Cambridge.uk 0.001%
DIRAC.CracowAgu.pl 0.001% | DIRAC.IF-UFRJ.br 0.175%
DIRAC.LHCBONLINE.ch 0.779% | DIRAC.Lyon.fr 2.552%
DIRAC.PNPI.ru 0.000% | DIRAC.Santiago.es 0.148%
DIRAC.ScotGrid.uk 3.068% | DIRAC.Zurich-spz.ch 0.003%
DIRAC.Zurich.ch 0.756% | LCG.ACAD.bg 0.106%
LCG.BHAM-HEP.uk 0.705% | LCG.Barcelona.es 0.281%
LCG.Bari.it 1.357% | LCG.Bologna.it 0.032%
LCG.CERN.ch 10.960% | LCG.CESGA.es 0.528%
LCG.CGG.fr 0.676% | LCG.CNAF-GRIDIT.it 0.012%
LCG.CNAF.it 13.196% | LCG.CNB.es 0.385%
LCG.CPPM.fr 0.242% | LCG.CSCS.ch 0.282%
LCG.CY01.cy 0.103% | LCG.Cagliari.it 0.515%
LCG.Cambridge.uk 0.010% | LCG.Catania.it 0.551%
LCG.Durham.uk 0.476% | LCG.Edinburgh.uk 0.031%
LCG.FZK.de 1.708% | LCG.Ferrara.it 0.073%
LCG.Firenze.it 1.047% | LCG.GR-01.gr 0.349%
LCG.GR-02.gr 0.226% | LCG.GR-03.gr 0.171%
LCG.GR-04.gr 0.056% | LCG.GRNET.gr 1.170%
LCG.HPC2N.se 0.001% | LCG.ICI.ro 0.088%
LCG.IFCA.es 0.022% | LCG.IHEP.su 1.245%
LCG.IN2P3.fr 4.143% | LCG.INTA.es 0.076%
LCG.IPP.bg 0.033% | LCG.ITEP.ru 0.792%
LCG.Imperial.uk 0.891% | LCG.Iowa.us 0.287%
LCG.JINR.ru 0.472% | LCG.KFKI.hu 1.436%
LCG.Lancashire.uk 6.796% | LCG.Legnaro.it 1.569%
LCG.Manchester.uk 0.285% | LCG.Milano.it 0.770%
LCG.Montreal.ca 0.069% | LCG.NIKHEF.nl 5.140%
LCG.NSC.se 0.465% | LCG.Napoli.it 0.175%
LCG.Oxford.uk 1.214% | LCG.PIC.es 2.366%
LCG.PNPI.ru 0.278% | LCG.Padova.it 2.041%
LCG.Pisa.it 0.121% | LCG.QMUL.uk 6.407%
LCG.RAL-HEP.uk 0.938% | LCG.RAL.uk 9.518%
LCG.RHUL.uk 2.168% | LCG.SARA.nl 0.675%
LCG.Sheffield.uk 0.094% | LCG.Torino.it 1.455%
LCG.Toronto.ca 0.343% | LCG.Triumf.ca 0.105%
LCG.UCL-CCC.uk 1.455% | LCG.USC.es 1.853%

4 Task Forces set up to boost this activities

CMS integration (CMS TF and SC3). Sizeable analysis activity (see also demo session)

ALICE integration (ALICE Task Force)



NA4 Continued

- **User survey conducted**
  - **User satisfaction, evaluation of grid added value – results to be published soon**
- **EGAAP reviewing new application requests**
  - **MoU for applications defined – take time to complete**
- **"Virtuous cycle" completely defined**
  - **Works with OAG (NA4/SA1) to ensure applications are well supported**
- **Technical fora**
  - **OAG (Operation Application Group)**
    - Discusses operation issues for applications (resource allocation)
  - **User forum planned for next spring**
- **Main issues**
  - **Achieving results is still labour intensive**
  - **User documentation needs improvements**
  - **Need to further demonstrate Grid added value**

JRA1: Middleware Reengineering

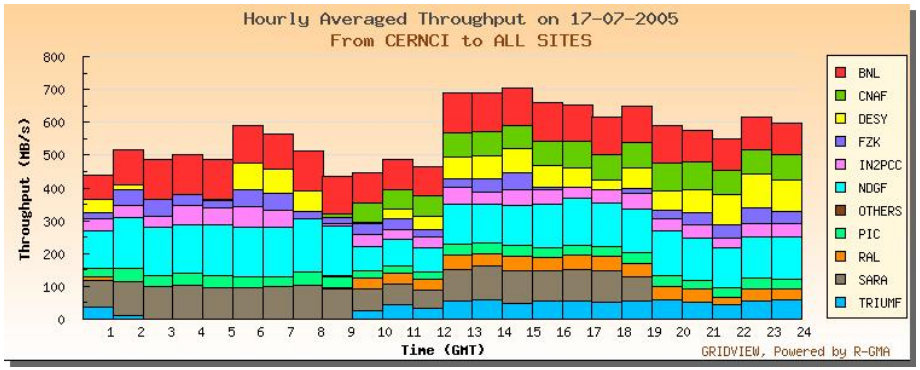
- A series of gLite releases have been produced (1.1, 1.2, 1.3, and 1.4)
  - *Driven by application and deployment needs*
  - *Focus on defect fixing*
- gLite deployed on PPS and made available to application use
  - *Independent evaluation outside EGEE by NGS and DILIGENT*
  - *gLite components also available via VDT*
- gLite components deployed on the infrastructure
  - *More scheduled by the end of the year*

- Emphasis is now to produce the second release deliverable (gLite 1.5)
  - *This is not the end of the story – maintenance and improvements will continue*
- Testing and integration tasks are still understaffed
- Need to preserve the gLite brand

gLite Lightweight Middleware for Grid Computing

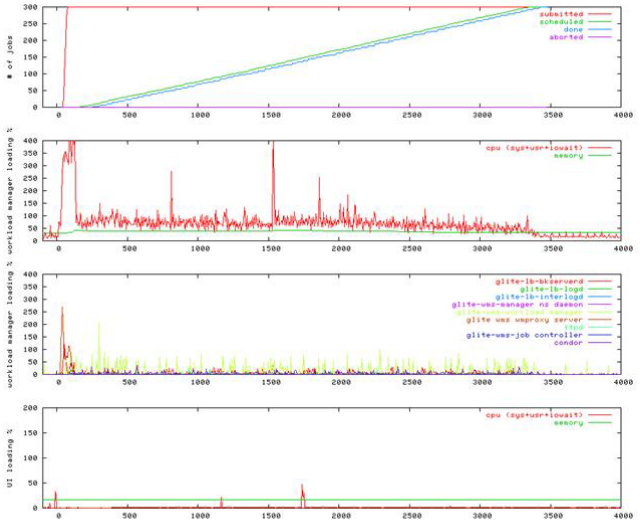
www.glite.org

Example gLite Usage



Service challenge throughput monitoring (FTS & R-GMA)

Atlas TF WMS measurements



- First gLite 1.4 WMS performance measurement in Milan
  - gLite 1.4 WMS (4CPU, 3GB memory)
  - 300 simple hello world jobs submitted by 3 parallel threads
  - submission rate ~ 4.1 jobs/sec
  - dispatching rate ~ 0.08 jobs/sec
  - All jobs in bulk are submitted to the same CE: atlasce.lnf.infn.it
- Thanks to Elisabetta Molinari for setting up the gLite WMS in Milan

PPS Layout (14 sites) all gLite services

Outreach & Training (NA2&NA3)

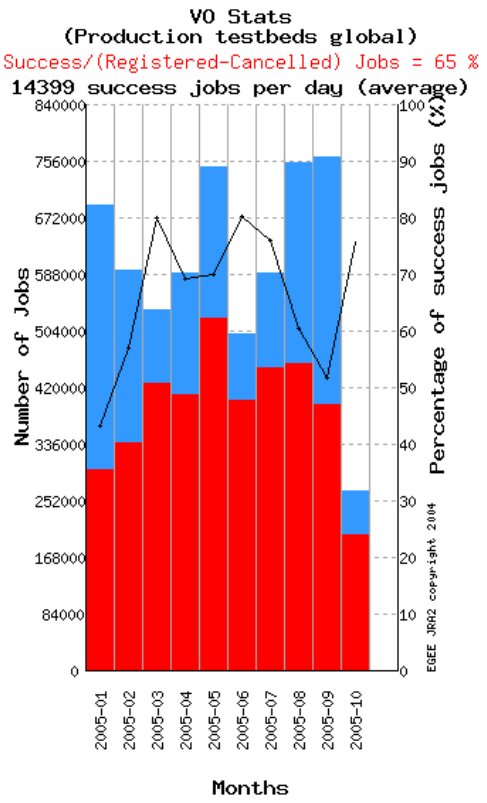
- Public, technical, and regional websites constantly evolving to expand information available and keep it up to date
  - Over 16,000 unique visits of an EGEE site each month
  - Many events attended and citations in the media
    - 92 news releases, 230 press cuttings etc.
  - New material to better address industry and decision makers as a target audience

- More than 170 training events (including ISSGC'05 and other summer schools) across many countries
  - More than 2000 people trained
  - Material archive with ~2000 presentations
  - Working closely with related projects (DILIGENT, EMBRACE, MAGIC, ICEAGE)
  - Developing eLearning framework started
  - t-Infrastructure developments (gLite integration and experience) started
  - Need to encourage more trainers
  - Improve organization/retrieval of training material

JRA2: Quality Assurance

- **Achievements**
  - Quality Group (QAG) has continued the work on metrics ensuring activities can collect relevant data
  - Improvements in job statistics collections, transition to gLite being planned
  - Program now in active use across activities
    - Collected statistics for Biomed data challenge
  - PPT in daily use and provided basis for period cost claims

Target	Current status (Q6)	End Year 2 target values
Number of Users	~ 1000	≥ 3000
Number of sites	120	50
Number of CPU	~12000	9500 at month 15
Numberof Disciplines	6	≥ 5
Multinational	24	≥ 15 countries



- **Issues**
  - Better understand the distribution for jobs submitted through RBs and other submission mechanisms
  - Define a process and associated tools to better measure the performance of project partners (more than 90 partners planned in EGEE-II)

JRA3: Security

- **Revised global security architecture. Secure credential storage procedures/recommendations document**
- **Middleware security group (MWSG) setting example for security interoperability between grid initiatives (EGEE, OSG, NAREGI)**
  - To be used for GGF work. Official MWSG meeting at GGF16
- **Actively contributing to the gLite middleware**
- **EUGridPMA continued work and was instrumental to**
- **IGTF launched,**
  - Chaired by David Groep (JRA3)
  - Coordinating European, Asian, and American GridPMAs
- **Vulnerability analysis database created**

- **For remaining 2005**
  - Reinforce middleware security component development and interoperability
  - Overview and recommendation document on accounting techniques
  - Second revision of security operational procedures document.
  - Assessment of security infrastructure – *Security Challenge*

NA5: Policy and International Collaboration

- Policy-related activities **:**
  - Coordinated support to the e-Infrastructures Reflection Group (e-IRG) defining grid policies across Europe and beyond
    - for the Luxembourg presidency white paper and roadmap
    - in the Luxembourg presidency workshop and meeting
  - Contributed to the planning of the UK and Austrian presidency events

- International cooperation activities**:**
  - "Concertation" with other EU projects (both from F2 and F3 Units)
    - Participated and contributed to many concertation-type events
    - Joint Working groups (security related) – more being discussed
    - MoUs, joint deliverables/events and LoS for other RI and IST projects
  - Reinforced cooperation with other geographical areas (in middleware, operations, training and dissemination)
    - USA, Japan, South Korea, China, Taiwan, India
  - Preparing an inventory of EGEE contributions to standardisation efforts

- **EGEE shall continue to contribute its experience with large-scale deployment of multi-science applications and knowledge of what is possible with the technologies today**

Priorities for the rest of the first phase

- **Focused technical EU review (Dec. 6/7 2005):**
  - No agenda received yet but focus will clearly be on NA4, SA1, and JRA1/JRA3
  - Must demonstrate
    - Availability and usage of the infrastructure (metrics not only showing the quantity but also quality of the service)
    - Usage of the infrastructure by applications other than the pilot ones (HEP and Biomed)
    - Grid added value in applications
    - Uptake of gLite components on the production service and convergence path
    - Impact of EGEE on the future development of Grids in Europe and worldwide
    - A clear path for the future development of Grids in Europe

- **Finish the program of work of phase I and prepare for a smooth transition to phase II**
  - Final deliverables have been moved forward
  - New EGEE-II structures (see Bob Jones' talk) need to be set up early
    - ***EGEE-II is a continuation not a new project!***
  - Final EU review probably in May 2006

Summary

- **Project Growth**
  - Ireland: Initial excitement
  - Den Haag: Reading the map
  - Athens: On the road
  - Pisa: **Building new roads**
- **High Points**
  - **Project buy in**, "Us" not "Them"
  - EGEE is ready for the **long term vision**

Dave Snelling, EGEE03 Athens

- **The focus of this conference**
  - Finalizing the program of work of phase 1
  - Prepare for a smooth transition to phase 2
  - Building new roads → **Upgrade existing roads**
    - Expand and intensify/deepening application uptake
    - Develop and start implementing the long term vision
      - *In collaboration with other Grid efforts world-wide – need to build the **World-Wide Grid***

Buon lavoro!