EGEE

SA1 – Operations Summary

Cristina Vistoli

INFN CNAF Italy

EGEE SA1

Regional Operation Centers coordinator

4th EGEE Conference

Pisa, 28th October 2005

Information Society

www.eu-egee.org

cristina.vistoli@cnaf.infn.it

Outline

- Metrics and job statistics
- Operations for users, VO managers, admins, users
- User support
- Deployment and release issue
- Experience from production sites
- OSCT, Intrusion detection, emergency planning
- Porting, access to other platform
- Storm
- EGEE-II: Sa1 Sa3 TCG
- Pre-production service
- Deployment tools

Metrics Categories

- Size + deployment metrics
- Operations
- Usage
- User support
- Services
- Defined list of metrics in all the categories
- Defined how to collect data
- Defined where to gather information
- Some examples:

Size + deployment

- Number of sites in production
- Number of job slots
- Total available kSpecInt: computing power available
- Storage (disk) + MSS
- Number of EGAP-approved VOs and apps
- Number of overall VOs: active
- Number of users active: current / future
- Total % of used resources

- M/w Security upgrade uptake

Operations

- **Site response to COD operational tickets**
  - (open/pending/closed within timeframe)
  - Average time of response
- **Site tests failed – or % time spent in the particular test-related state**
- **% availability of SE, CE – outside of downtime window**
- **ROC: Number of days down / site summed up?**

Usage

- Jobs per VO (submitted, completed, failed) (through RBs)
- Data transfer per VO
- CPU (wall clock time) and storage per VO
- % of sites blacklisted/whitelisted per timeframe
- Number of CE/SE available to a VO

User support

- Time to response to user tickets (+within different groups)
- Number of supporters
- Number of tickets escalated at various points + time related to it
- % of tickets assigned to wrong group + time related to it
- Above measures for ROCs: responsibility of ROCs

Services

- RB: % passing a job submission test
- BDII: Query time measurement
- MyProxy: % passing a test (register a proxy, access the proxy, delete the proxy)
- SRM-SE: % passing file movement and deletion tests
- Catalogue: % passing register, query, delete
- VOMS: % passing VOMS test
- RGMA: TBD

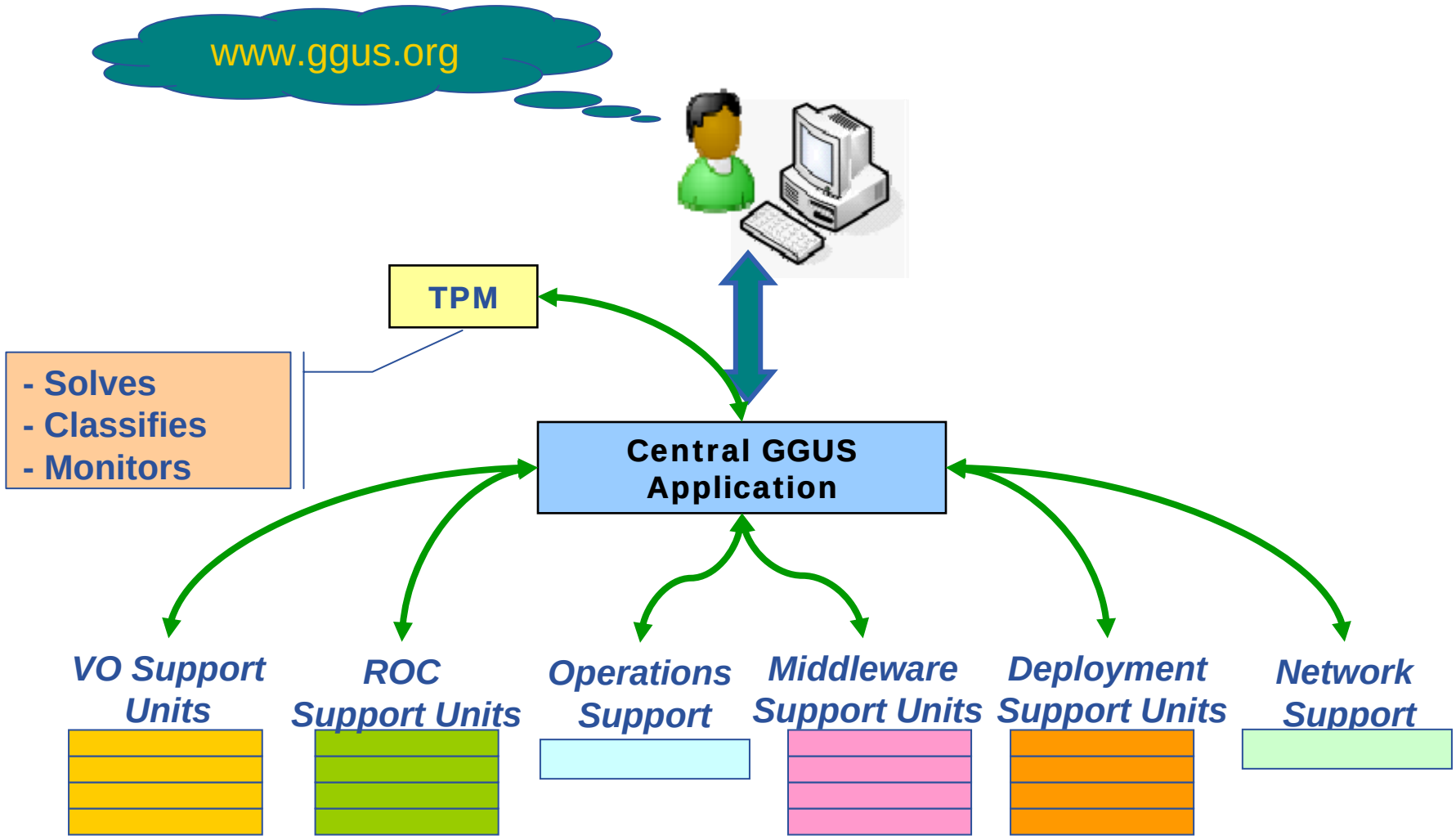
Metric discussion

- **We are going to publish these values**
- **Is it acceptable for all sites?**
- **Provide guidelines to publish correct specint and correlated benchmark**
- **Goal:**
  - provide understandable metrics with comparable values
  - Most of the metric are numbers with an associated scale
  - Other metrics are judgement of quality
  - Measure the quality of the infrastructure
- **Applications need to measure the quality of their application**
  - Basic numbers can be retrieved via with application SFT
- **Start publish some of these metrics soon as prototype,**
  - end of the year
  - Service metrics LCG requirement by December
  - Other metric sets to follow same timeline

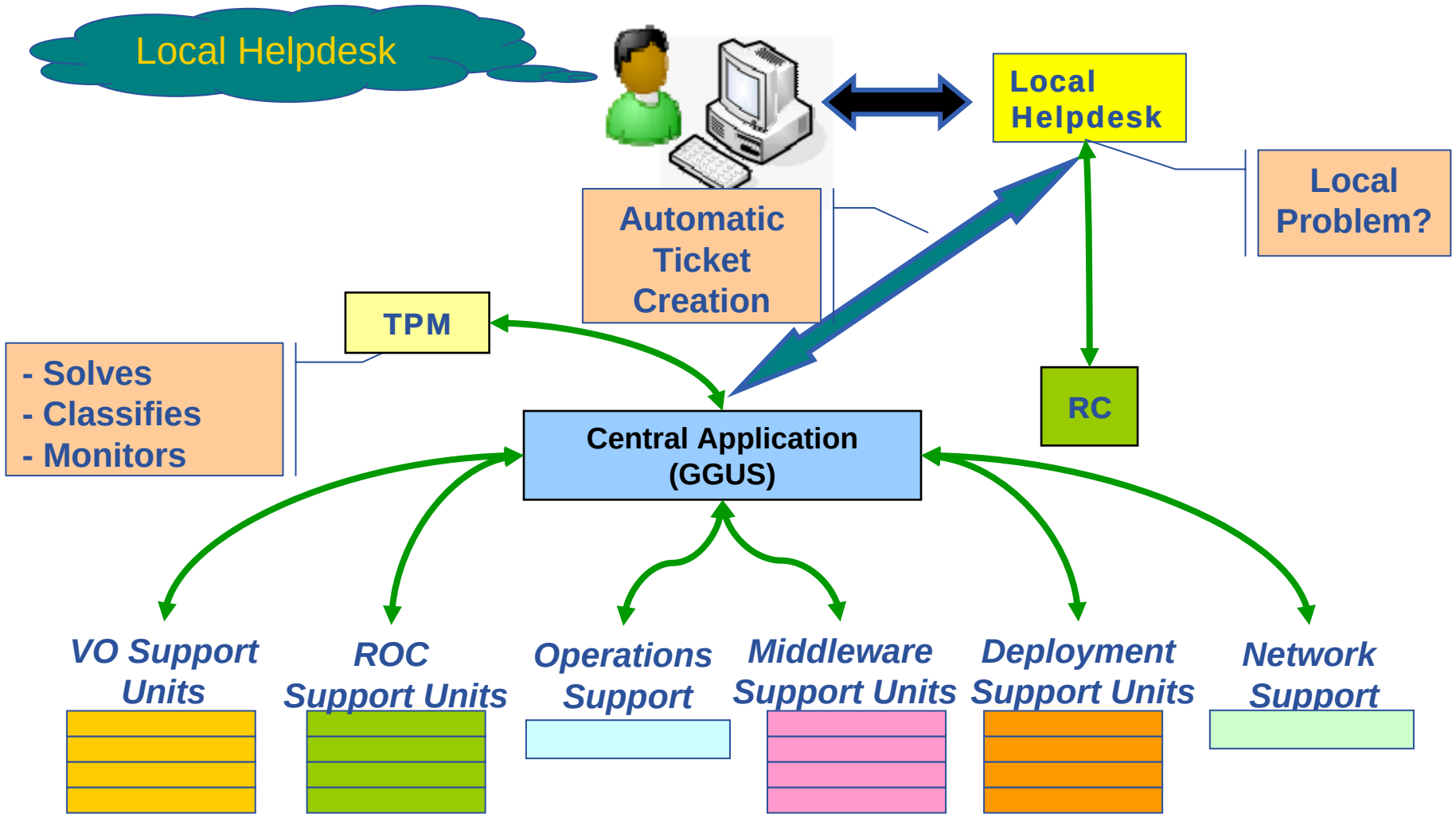
Operation Information in the CIC portal

- **Operation Information for end users, VO managers and Site administrators using the CIC portal**
  - "VO users" section
  - http://cic.in2p3.fr/index.php?id=endusers
  - "VO managers" section
  - http://cic.in2p3.fr/index.php?id=vo
  - "Site administrators" (RC) section
  - http://cic.in2p3.fr/index.php?id=rc
- **Request for Roc team to access CIC on Duty tools and dash boards**

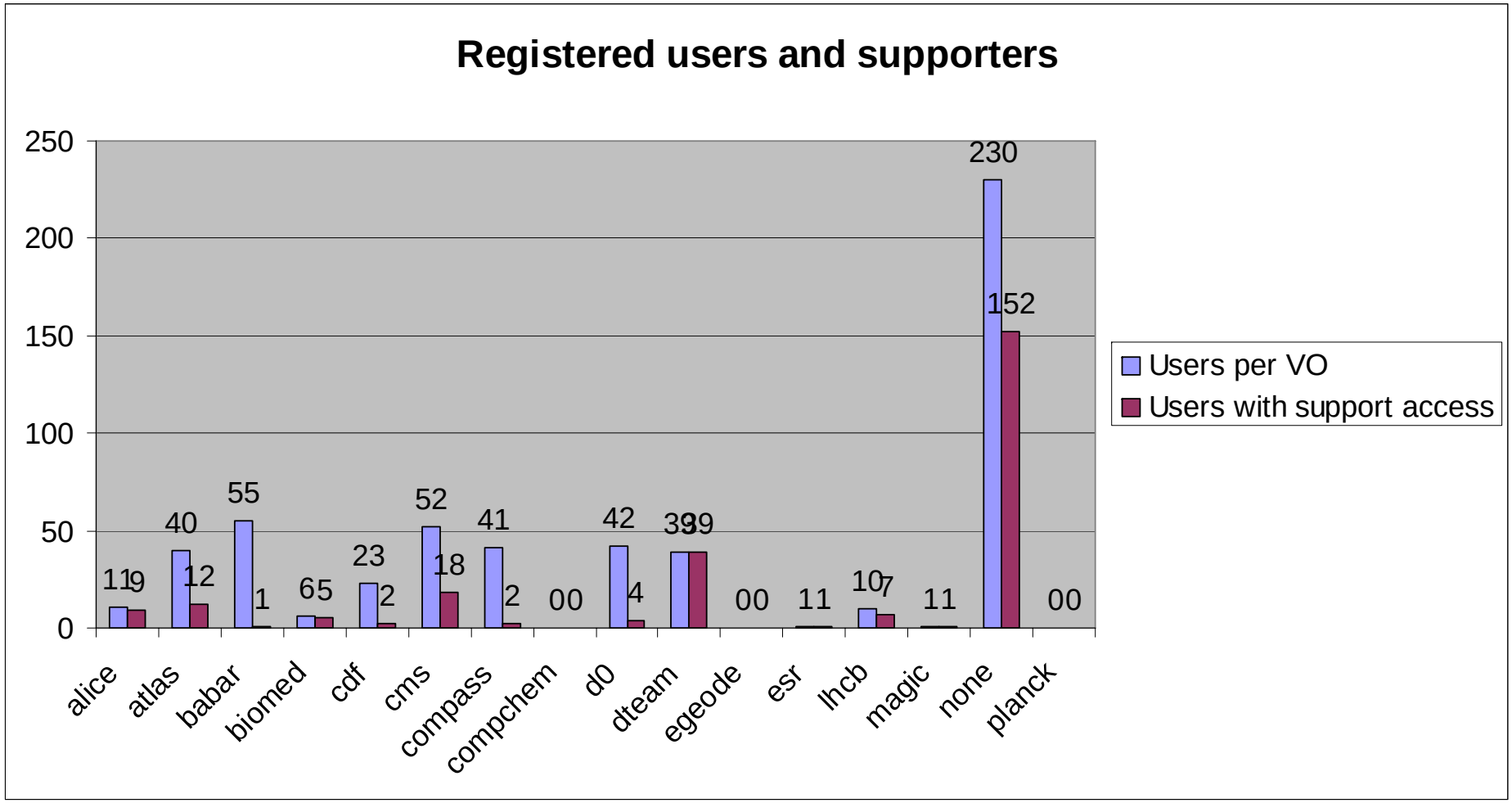
User Support Workflow



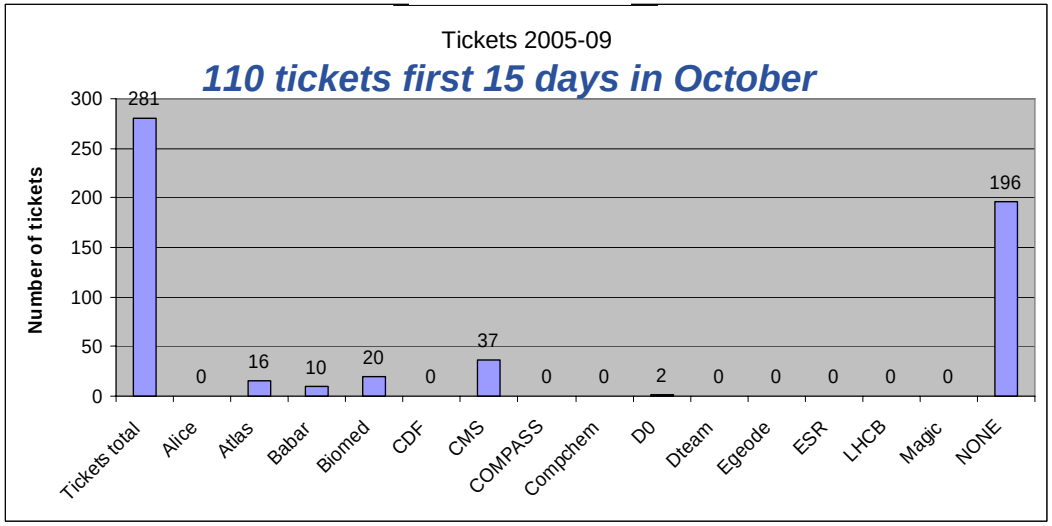
Support Workflow

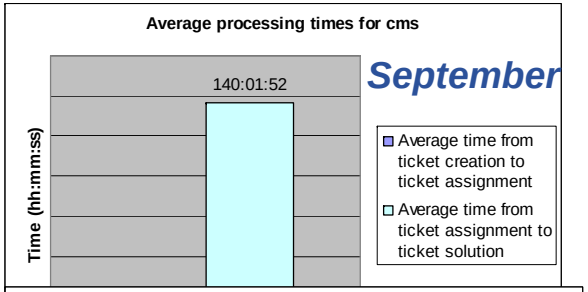


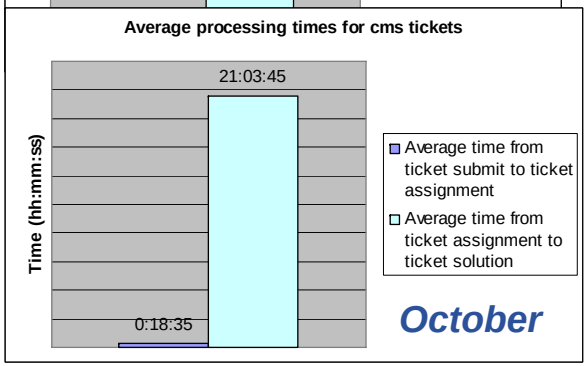
Some statistics: users per VO

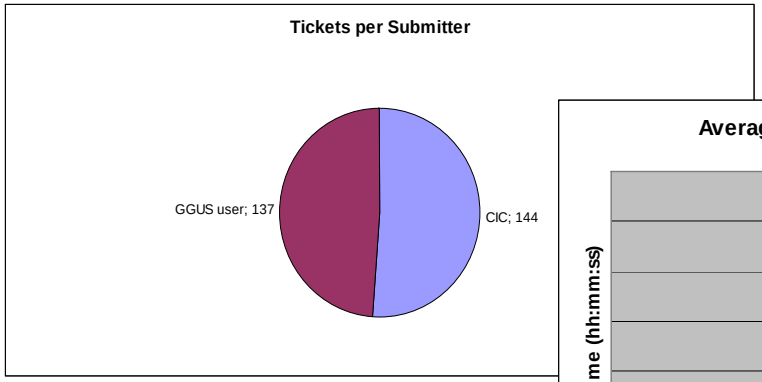


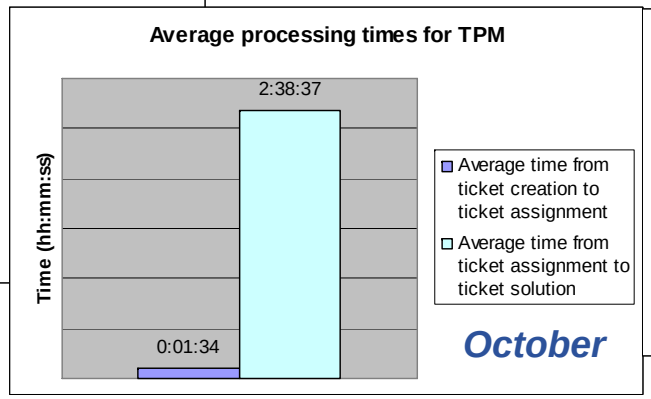
Performance statistics

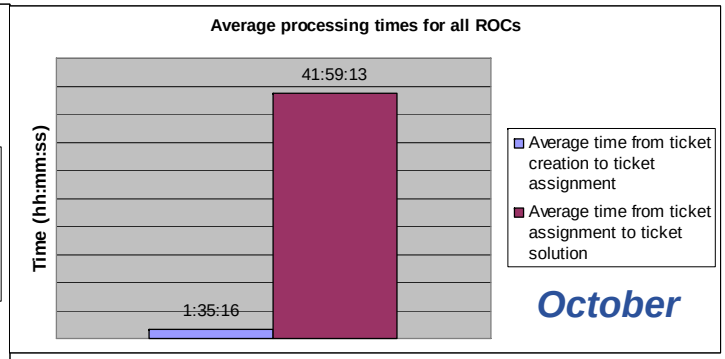












User support Conclusions

- The functionality and usability of the ***GGUS*** system ***has improved*** in the last months, thanks to the help of the ROCs (more tickets submitted, more customers and general appreciation of the service).
- ***GGUS/ESC*** coordinates the effort and operations: key body.
- The existent ***interfaces with the ROCs*** are quite practical and make the system function as one. Most ROCs have established functional interfaces with GGUS, the others are working on it.
- The ticket traffic is increasing. We still ***do not know*** what a ***real figure*** would be for the number of ticket to be expected. The system can be dimensioned appropriately with more TPMs and support units.
- A lot of ***metrics established*** to measure the performance of the system (performance of a supporter/support unit, tickets solved/week/VOs, # of tickets filed in Wiki pages, etc.). The measures refer only to the central system. Each ROC processes and solves also local requests. Measures for each ROC are also available.
- GGUS is working on a ***plan to offer resilience to system and network failures***.
- ***We need more specialized supporters*** in order to help the supporters at CERN who now are the main source of knowledge and help.

Deployment and Release Issues Straw Man I

- **Upgrade and introduce services as soon as ready (SC and gLite)**
  - Whenever upgrades are ready and tested
  - T1s experience can be used as additional deployment test
  - Handle these as upgrades to the last major release
  - Make client libs available in user space
  - Version tracking via the information system
- **With sufficient material accumulated**
  - Cut a new release
  - Starting point for sites to join
  - Reference point for external project
- **Security upgrades**
  - Upgrades + message to the site security contacts
- **Work continuously on tests and documentation**

Deployment and Release Issues Straw Man II

- **New services should be "integrable" with current CIC on duty**
  - Tick list produced at the operations workshop

Ticklist for new service

- User support procedures (GGUS)
  - Troubleshooting guides + FAQs
  - User guides
- Operations Team Training
  - Site admins
  - CIC personnel
  - GGUS personnel
- Monitoring
  - Service status reporting
  - Performance data
- Accounting
  - Usage data
- Service Parameters
  - Scope - Global/Local/Regional
  - SLAs
  - Impact of service outage
  - Security implications
- Contact Info
  - Developers
  - Support Contact
  - Escalation procedure to developers
- Interoperation
  - ???
- First level support procedures
  - How to start/stop/restart service
  - How to check it's up
  - Which logs are useful to send to CIC/Developers
    - and where they are
- SFT Tests
  - Client validation
  - Server validation
  - Procedure to analyse these
    - error messages and likely causes
- Tools for CIC to spot problems
  - GIIS monitor validation rules (e.g. only one "global" component)
  - Definition of normal behaviour
    - Metrics
- CIC Dashboard
  - Alarms
- Deployment Info
  - RPM list
  - Configuration details (for yaim)
  - Security audit

CERN IT-GD

Deployment and Release Issues Straw Man III

- **Integrated releases on the pre-production service**
  - Now:
    - production: lcg + gLite components
    - pre-production: gLite only
  - Better:
    - production: lcg + gLite components
    - pre-production: lcg + current gLite + new gLite components
- **New name for integrated production stack needed**
- **The winner is:**

– gLite X.Y

Deployment and Release Issues: conclusion

- **Rapid upgrade versus stability and site predictability: how to cope with these two requirements?**
  - continuous release upgrade cannot be operated
  - to operate a deployed middleware we need a reference version
- **Version number per service published in the information system**
- **Release: becomes a mix of services versions**
  - well defined rpm list for services and clear procedure for upgrade the single service
- **Is the backward compatibility guaranteed?**
  - test is possible to some extend completely regression test takes too time.
  - not ideal but service challenge is very demanding in new versions
- **gLite name is confusing for a complete release**
  - gLite release stops and start releasing separate components then the name is ok....
  - inside gLite services there are already diff. versions number
  - just continue the current numbering, add in the relevant LCG software
- **Integrating pre-production service with LCG and Production services**
- **Make PPS more similar to real next production version**
  - pure gLite version not more distributed to PPS
  - PPS is the last step of the certification for the application
  - additional partners in certification process for EGEE-2

Issues from Site/Grid Managers

- **Restrictions on OS**
  - Need at least few Linux distributions
  - List of required packages
  - Improved documentation of migrations
- **Few supported batch systems**
- **More Transparent Release management**
- **Insert middleware into the fabric management tools**
- **VO boxes**
- **Middleware Releases.**
  - **“continuous update” model will introduce problems**
  - **how will “versions” be tagged**
  - **support impact be if sites are running different software versions**
  - **YAIM or gLite installation tool?**

Issues by Site/Grid Managers

- Defining and measuring availability
- User/operational support homogeneous quality among federations
- Technical forum on storage issues
- More flexible SFT
  - Different subset of tests for regional SFT
  - Sites in sc3 have specific services need special SFT

OSCT, Intrusion detection

- **Security Service Challenge**
  - SSC_1 targeting ROC sites completed in June 2005
  - Transition to GGUS problem tracking tool t.b.d.
  - Procedure for SSC reponse should be included into the COD operation
  - Next SSC scheduled for Nov
- **Security Monitoring**
  - First test integrated into SFT framework
    - Monitoring last-download time for CRLs
- **Incident Response**
  - Proposed: *Incident Response Handbook*
- Grid-wide Intrusion Detection System

When are emergency procedures required?

- **Emergency procedures are required to cover the following cases:**
  - Incident response plans cannot be followed: critical parts of the infrastructure are unavailable (e.g. mailing lists)
  - Incident response plans are inappropriate: E.g. need to rapidly inform large parts of the community, beyond the security contacts, and incident communication channels are compromised

- **Examples**
  - Major power cut at Site A lasted several days
  - Cable cut network access to Site B
  - Major worm disrupted network access at Site C
  - Security incident blocks user access to accounts at Site D
  - Wide area exploit of the (homogeneous) security fabric

Emergency plan

- **Define a plan for disaster**
  - communication is important
  - out of band communication mechanisms are needed
  - coordinate with the Grid participating site how to handle disaster and define sites recovery plan
- **US sites are forced to have an emergency plan**
- **Grid changes the classical emergency disaster plan:**
  - the grid introduce dependability at global level

Need for porting

- **x86 clusters where SL3 cannot be installed**
- **Support for other architectures**
  - Grid-enable compute centres with existing resources
  - User applications optimised for a particular architecture
  - MacOS/X, AIX, IRIX, Solaris, IA64, …
- **Definite need to port (at least) the WN software**
- **Question: is there a requirement to port service nodes?**

Overview of porting projects

- **CERN Openlab & SPACI**
  - Itanium port available and tested for LCG 2.6.0 (all nodes)
- **CERN/UVienna/Apple**
  - MacOS X port available (focus on UI: WMS, …)
- **Grid-Ireland**
  - WN ports available for CentOS 4.1, Suse 9.3, RedHat 7.3/9
  - Work in progress on MacOS X, Solaris, EMT64, FC4, AIX, IRIX
- **GSI (Germany)**
  - Debian port (UI and WN?)
- **IRB (Croatia)**
  - Debian: tar fixes (UI), chroot (CE+WN), converting RPMs to DEBs (ongoing); FreeBSD: tar (UI)
- **HPC2N Umea (Sweden)**
  - Porting gLite to Ubuntu (Debian)
- **EGRID (Italy)**
  - LiveCD with all service nodes, UI-only relocatable installation

Issues on porting

- The future is to provide common infrastructure for all platform with no need to check out the code
- Installation of different platforms must not be an issue
- The code has to be 64 bit compatible
- Interoperability is also very important
- Significant interest to run Grid service in different Linux distribution non only SLC

StoRM objectives

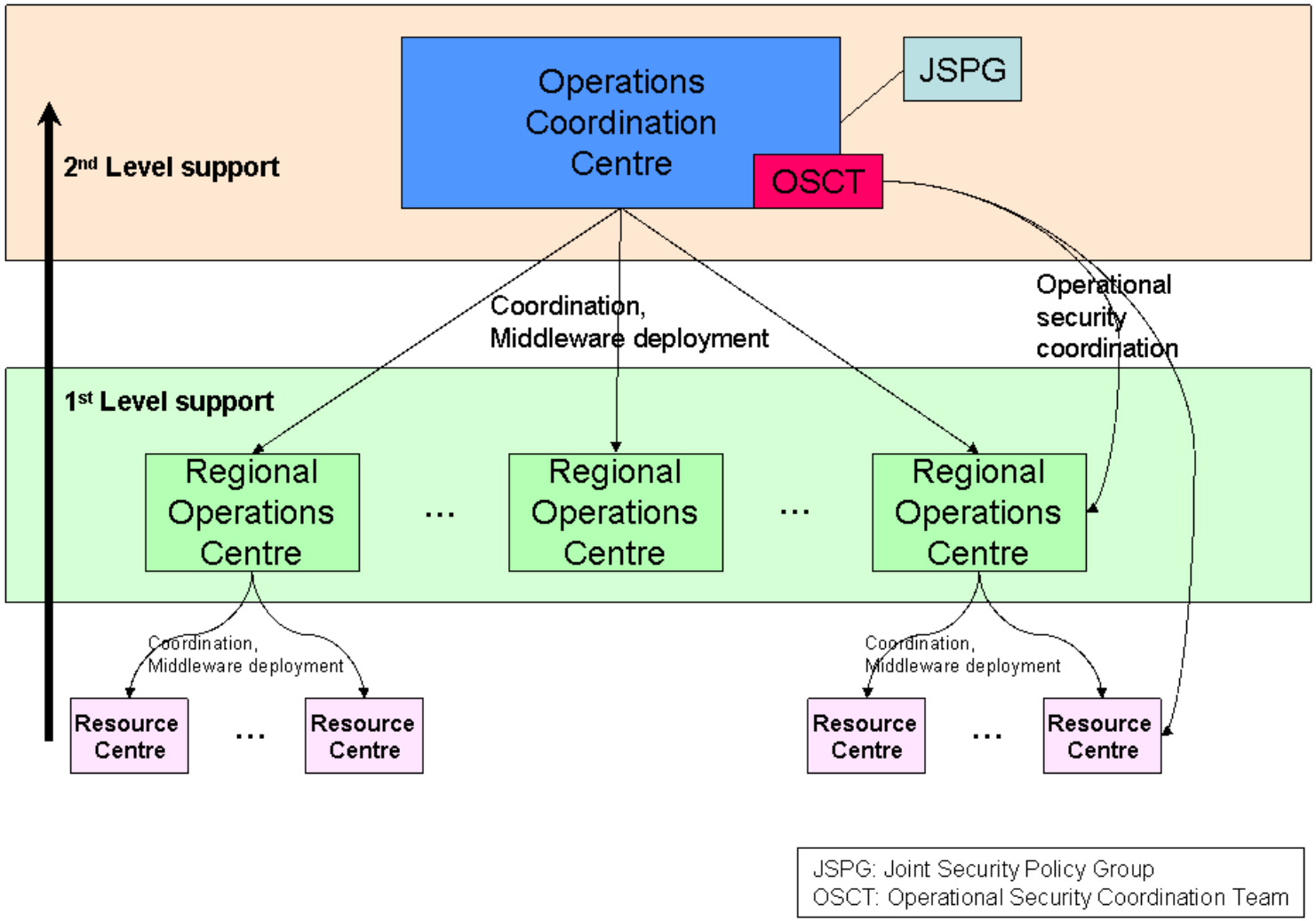
- **A disk based SRM 2.1.1 server**
- **StoRM's implementation of SRM 2.1.1 meant to meet three important requirements :**
  - **Space Reservation using** file system with **native support for space reservation**
  - **direct POSIX I/O call.** Performant parallel file system distributed over all WNs of farm
  - Security on data **VOMS proxies**.
- **Comments:**
  - Is interoperability tested with other SRM 2.1.1 implementation?
  - Is Raid5 supported?

- **Sa1 Sa3 Tcg**

SA1 in EGEE-II

- Consolidate and build on what has been done in EGEE
- Integration, certification, release preparation moved into SA3
- Maintain hierarchical structure, but simplify ROC/CIC → ROCs
- Strengthen (operational and user) support for many more applications and users
- Security – take operational tasks from JRA3; set up real response teams

Organisation



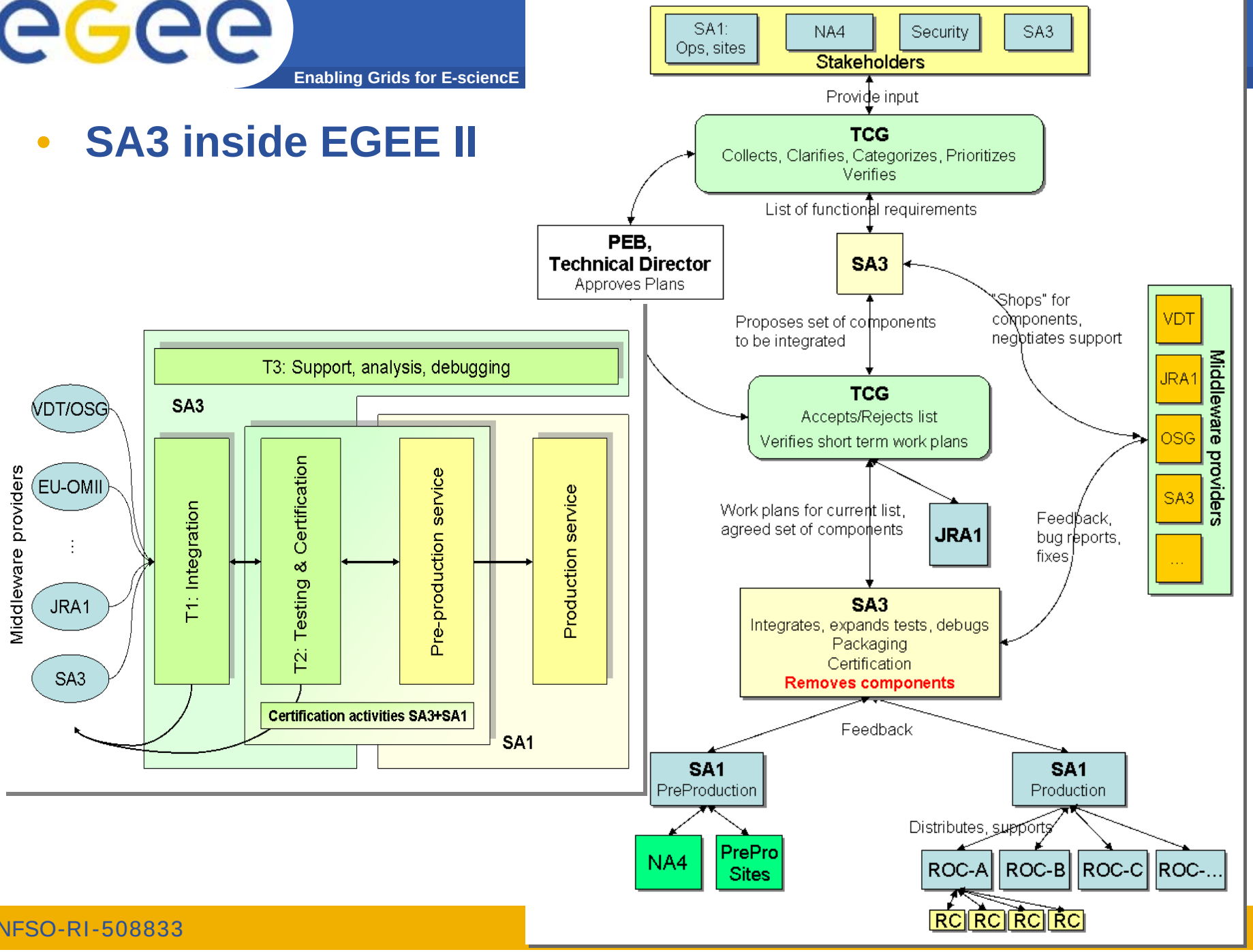
SA3 Tasks

- **T1: Integration and packaging**
- **T2: Testing and certification**
- **T3: Support, analysis, debugging, problem resolution**
- **T4: Interoperability**
- **T5: Capture Requirements**

- **Partners: CERN +**
  - PSNC,TCD, ▪ IMPERIAL, INFN,UKBH
  - UCY, GRNET, CSIC, PIC, CESGA, FZJ

  - Partners active in testbeds, tests, and specific interoperation tasks

- **SA3 inside EGEE II**



TCG Overview

- **The EGEE-II proposal defines a Technical Coordination Group (TCG):**

  *The TCG brings together the technical activities within the project in order to ensure the **oversight and coordination of the technical direction** of the project, and to **ensure that the technical work progresses according to plan**.*

- **Basically coordinating the work of SA1, SA2, SA3, NA4, and JRA1**
  - Membership from all these activities but still remain a “small” team
  - Additional experts will join based on the topic of discussion
  - Working groups will be spawn off to solve specific problems

- **Focus on practical short term solutions**
  - Long term projects will be sourced out to middleware providers

- **The group must have executive power**
  - Not just a discussion forum!
  - Decisions taken by the group must be honoured by the affected activities

Report from discussion

- The TCG should have control on the architecture/design team and have a solid requirement management process
- The most important members of the TCG should be users, sites and developers representatives with a distillation process to avoid TCG become too big
- User documentation is very important and has to be part of the certification process is to verify documentation
- The TCG picture is too complicated simplest approach taking into account all the actors
- Define strict acceptance criteria to avoid testing things that doesn't work.
- JRA1 is not in the stakeholders, because the list contains requirement providers
- Avoid confusion if debugging and fixing is in sa3 or in jra1, SA3 just minor fixes
- SA3 and SA1 are middleware providers only for small components, glue between services
- Make sure the contributions coming from SA3 follow the same criteria as the rest
- The deployment policies? is not a short term discussion and will be done in the TCG, certain decisions will also be taken by deployment managers depending on site needs

Proposal for PPS Evolution

- **PPS can hardly be defined the Pre-Production compared with the 'production services**
  - PPS includes only gLite services
  - PPS does not share Monitoring and Operation Procedures with the Production
  - PPS is far away to be a preview of what really will be in the next LCG release
- **How PPS should change**
  - Improve Operation Procedures
  - Same Support model as in production
  - To be included in Monitoring Infrastructure
  - Include all production services really in use: Catalogs/LFC, BDII etc
- **How Production should change**
  - Adding deployable gLite Services to existing production services

gLite 1.4 Certification Status

- **WMS**
  - push/pull: OK
  - Integration with the BDII: still waiting for the fix for the empty cache (it is not supposed to be released with 1.4.1, which may mean that likely the WMS will not be ready to be included in the next LCG release as proposed)
  - Bulk submission: not tested, test left to PPS users
  - Upgrade: OK
  - Status: Moved to PPS
- **IO, Fireman (Mysql), R-GMA :**
  - Upgrade: OK
  - Status: Deployable
- **FTS :**
  - Configuration instructions still not sufficient
  - It took several days to get it working
  - We could not get it working with dpm
  - Still does not work with catalogs
  - It works with url-copy
  - Upgrade: not tested
  - Status: Not deployable out-of-the-box. Expert support needed for installation, limited functionality, keep version 1.3

Deployment tools

- **Currentl situation: YAIM**
  - All changes concentrated in 1 file, easy to customize and extend it, appreciated by many site managers
  - Balance between simplicity and flexibility
- **Glite configuration approach is different**
  - For each glite module, there is a corresponding packages list + XML data + 1 configuration script
  - 1 XML document can describe the configuration of 1 site
- **There is an overlap between the two tools**
  - Discussion how to find a good compromise for site manager, integrators, developers
  - Everybody agree that the configuration for the sitemanager must be kept simple whenever possible
  - A. Di Meglio, O. Keeble, A. Forti will work for a temporary smooth transition

Thank you!