# Grid Operations
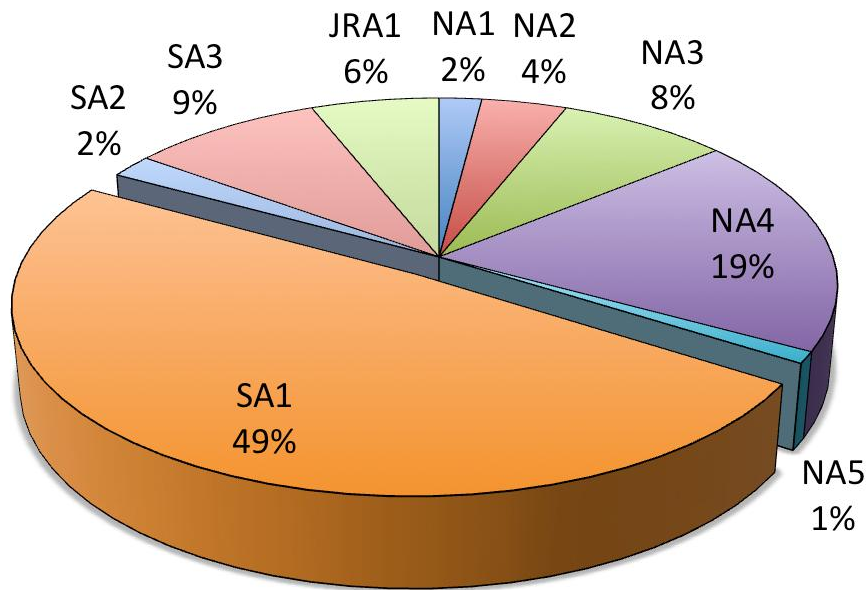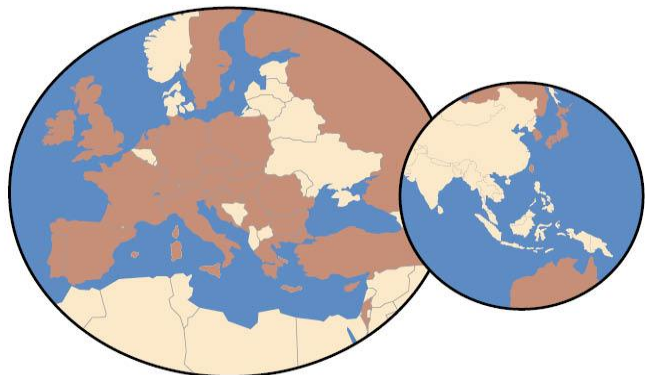# SA1 Status Report

*Steven Newhouse*

*(substituting for Maite Barroso Lopez)*

*EGEE-III Final Review, 23-24 June, 2010*

**www.eu-egee.org**

e-infrastructure

**Enabling Grids for E-sciencE**

## SA1 Partners

## 28 countries, 175 FTE

| Country | Total PM planned at M24 *(1)* | Total FTE |
|---|---|---|
| Austria | 37 | 1.5 |
| Belgium | | |
| Bulgaria | 60 | 2.5 |
| CERN | 420 | 17.5 |
| Croatia | 47 | 2.0 |
| Cyprus | 47 | 2.0 |
| Czech Republic | 58 | 2.4 |
| Finland | 24 | 1.0 |
| France | 450 | 18.8 |
| Germany | 392 | 16.3 |
| Greece | 131 | 5.5 |
| Hungary | 38 | 1.6 |
| Ireland | 36 | 1.5 |
| Israel | 52 | 2.2 |
| Italy | 468 | 19.5 |
| Netherlands | 204 | 8.5 |
| Norway | | |
| Poland | 152 | 6.3 |
| Portugal | 100 | 4.2 |
| Romania | 57 | 2.4 |
| Russia | 424 | 17.7 |
| Serbia | 55 | 2.3 |
| Slovakia | 33 | 1.4 |
| Slovenia | 16 | 0.7 |
| Spain | 317 | 13.2 |
| Sweden | 120 | 5.0 |
| Switzerland | 24 | 1.0 |
| Turkey | 66 | 2.8 |
| UK | 372 | 15.5 |
| **Total PM planned at M24** | **4200** | |
| **Total FTE** | | **175.0** |

SA3 9%, SA2 2%, JRA1 6%, NA1 2%, NA2 4%, NA3 8%, NA4 19%, SA1 49%, NA5 1%

**Enabling Grids for E-sciencE**

## Operate the EGEE production infrastructure, providing a high quality service to the application groups

- This is made possible by a number of support structures including
  - Regional Operation Centres
  - Global Grid User Support (GGUS)
  - Release and deployment management
  - Grid security coordination
  - Operational procedures and tools
- Structural changes needed for the transition to a sustainable model

**Number of EGEE-III certified sites (April 2010): 317 (in 52 countries)**
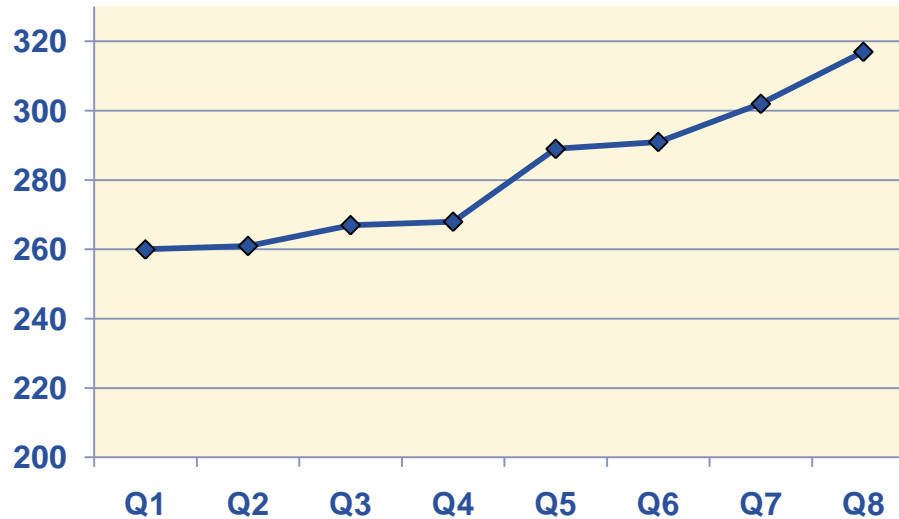


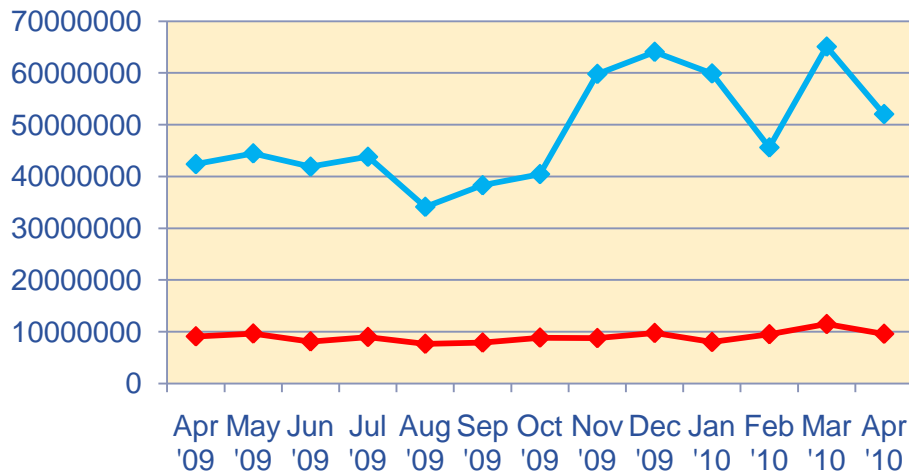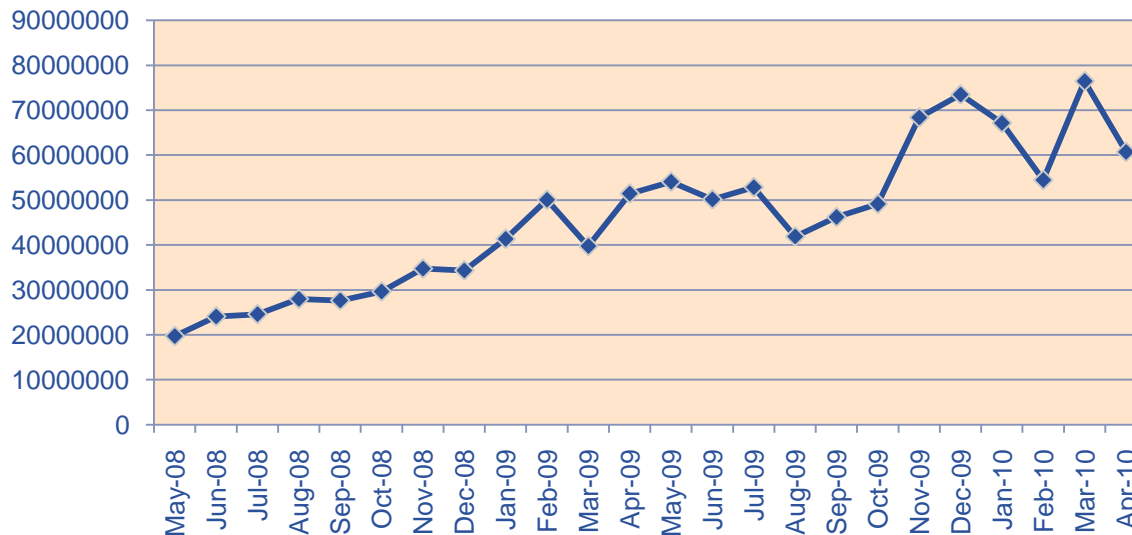**Computing resources, doubled since last year:**

- **243,020 cores** (139,000 last year)
- **500 MSI2k** (212 MSI2k last year)

(MSI2k = Normalised CPU time to a reference value of Mega SPECint 2000)

**Storage resources:**

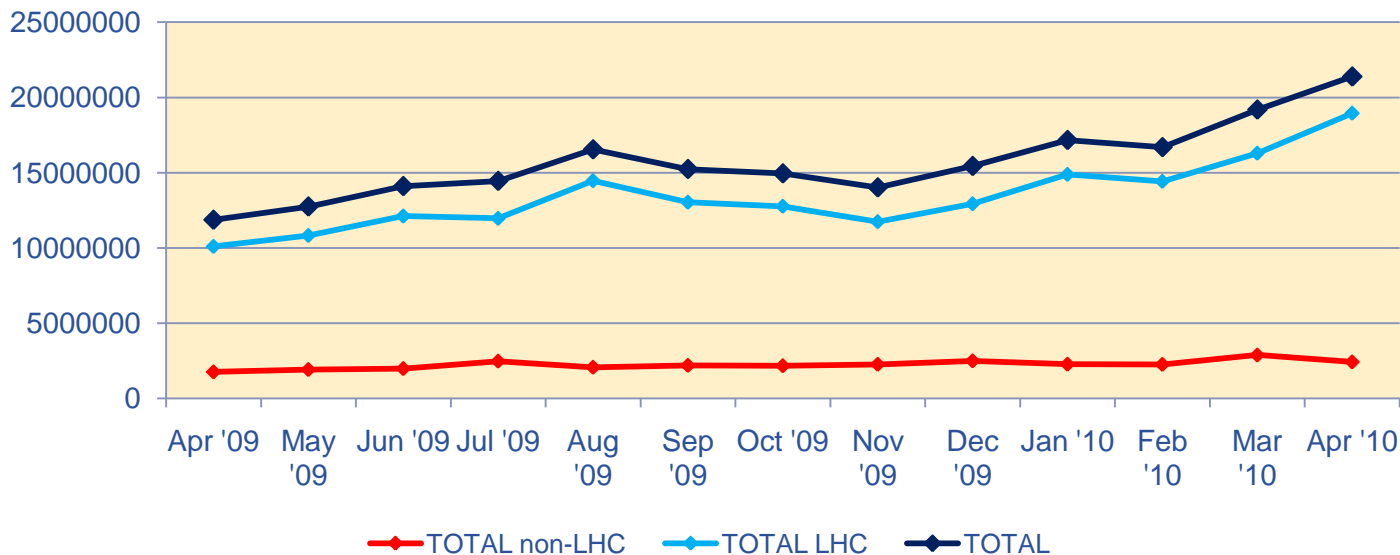- **40 Pb disk, 61 Pb tape** (25 Pb disk, 38 Pb tape last year)

**Enabling Grids for E-sciencE**

**Submitted**

**Jobs**



• usage from non-LHC communities stable

• LHC usage growing, real activity: Real data processing, simulations, analysis



TOTAL non-LHC    TOTAL LHC

# Usage of the infrastructure (II)

## Number of jobs



- **Average of 15 million jobs per month** (14M last year)
- **525.000 jobs/day** (480.000 jobs/day last year)
- Non-LHC VOs with stable workload
- LHC running increasingly high workloads
  - anticipate millions / day soon

# Site availability and reliability (I)

- **No changes in algorithm regarding last year**
  - Service available if any of the service instances are available (logical OR)
  - Site available if all services are available (logical AND of all service types)
- **EGEE League Tables (monthly reports) published and followed up with ROCs**
  - Also available for top 10 VOs
- **Underperforming sites requested to report justifications, collected in a wiki page**
  - https://twiki.cern.ch/twiki/bin/view/EGEE/MonthlyAvailability

| Italy | | |
|---|---|---|
| GRISU-CYBERSAR-CAGLIARI | 57/57 | Problem on local cluster and storage |
| GRISU-CYBERSAR-PORTOCONTE | 45/45 | problem on SE due to the lcg-voms certificate expired, problem with a bdii set by mistake on WNs |
| INFN-CAGLIARI | 63/63 | Maintenance on CE and some problems on few worker nodes on which we are investigating |
| INFN-GENOVA | 66/93 | Change CC cooling system and rearrange racks. Operations on UPS |
| INFN-PADOVA-CMS | 64/64 | problem related to a bug in the new STORM version, jobs submission problems (reinstallation of CE) and authentication problems due to few sgmops pool accounts |
| INFN-PERUGIA | 14/24 | Hardware upgrade & issues on SE |
| TRIGRID-INFN-CATANIA | 62/70 | UPS maintenance & Hardware problems - DCP |

**Enabling Grids for E-sciencE**

## April 2010

| Region | Avail-ability | Reli-ability |
|---|---|---|
| AsiaPacific | 92 % | 93 % |
| CERN | 98 % | 99 % |
| CentralEurope | 88 % | 88 % |
| France | 97 % | 97 % |
| GermanySwitzerland | 93 % | 93 % |
| Italy | 94 % | 94 % |
| NGI_GRNET | 96 % | 97 % |
| NGI_PL | 94 % | 95 % |
| NorthernEurope | 91 % | 91 % |
| ROC_Canada | 92 % | 93 % |
| ROC_IGALC | 79 % | 90 % |
| ROC_LA | 89 % | 89 % |
| Russia | 86 % | 89 % |
| SouthEasternEurope | 77 % | 77 % |
| SouthWesternEurope | 91 % | 91 % |
| UKI | 93 % | 95 % |

• Most regions have a good performance well above targets (70% availability and 75% reliability)

• April had highest average availability in the last year: 91%

• New ROCs started with lower A/R levels than experienced ones, took a few months to achieve targets

**eGee**

- **Sites with availability <50% for three consecutive months are removed from the Production infrastructure**
  - 7 sites removed, from AsiaPacific and Russia
- **Simplification of downtime declaration rules**
  - To enforce scheduled downtimes
  - Unscheduled downtimes has a big impact in site availability! And to users!
  - Ratio of scheduled/unscheduled downtimes **doubled** between May-09 and Apr-10
    - Sites are doing a better job following the procedures and pro-actively informing their user communities of planned maintenances
- **New ROCs needed a few months to get stable and achieve similar availabilities than the rest of the infrastructure**
- **Most regions have a good performance well above targets (70% availability and 75% reliability)**
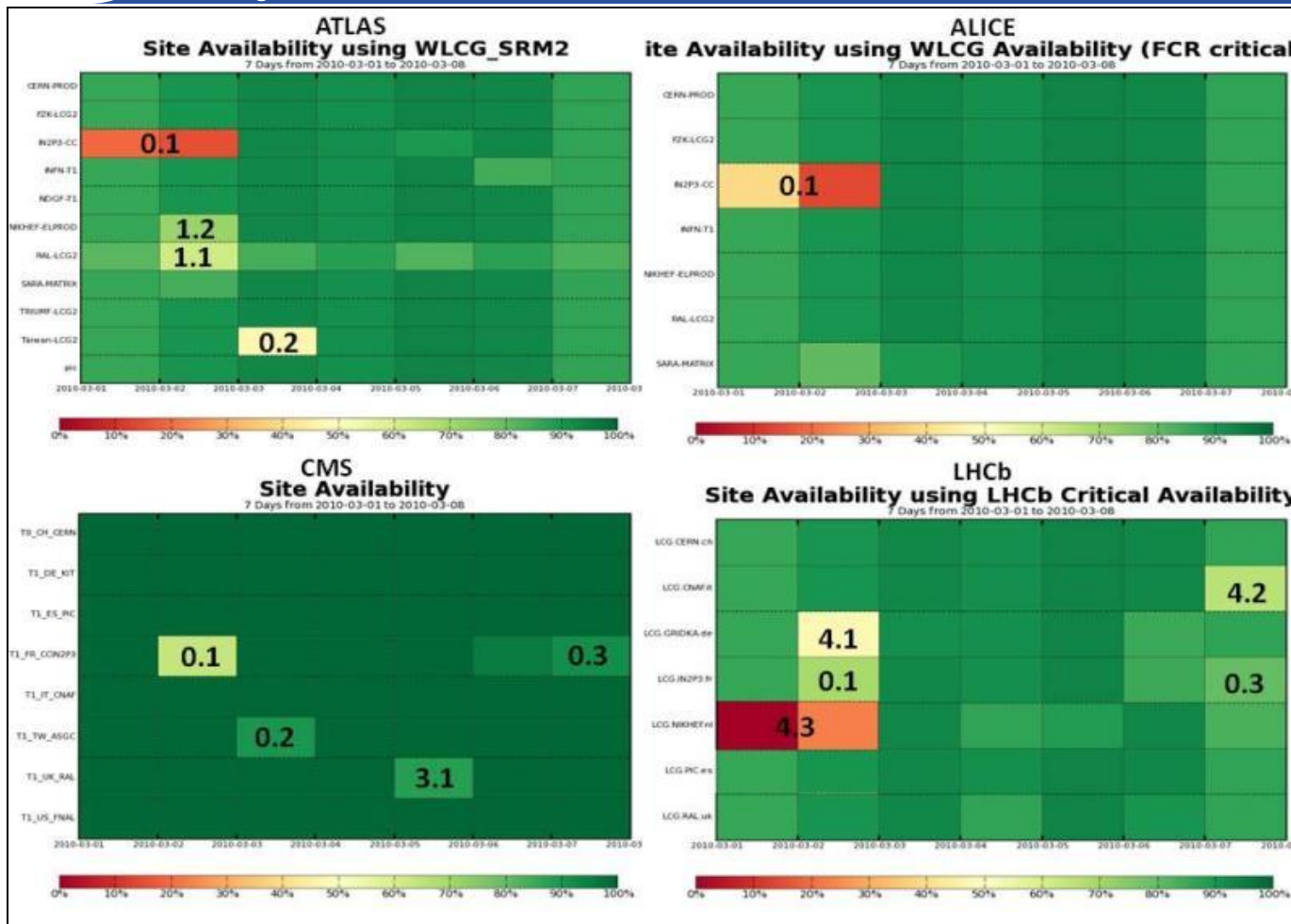
- **Views from WLCG dashboards**
- **Built on top of VO specific grid monitoring probe results**

- **Move to Staged rollout as transition to EGI model:**
  - Deploy middleware updates in a controlled way on the production infrastructure to get early feedback on the release, reducing the duplication of work at multiple sites
  - All ROCs agree with the concept, few sites volunteer to take part on it
    - Requires commitment to upgrade quickly in production
    - As it is in production, might affect availability and reliability
- **Process for retiring old versions of middleware services**
  - Prototyped with Resource Brokers
  - Trying with gLite 3.1 services (replaced by gLite 3.2)
    - Big/medium sites able to follow new versions
    - Small ones slow, different issues, e.g. Obsolete H/W not running SL5
    - Sites per Software version: 48% gLite 3.2, 52% gLite 3.1
- **List of gLite service versions supported by SA1 Operations**
  https://twiki.cern.ch/twiki/bin/view/EGEE/SupportedServiceVersions

**Enabling Grids for E-sciencE**

- **Distributed security operations**

- **Strong expertise in incident handling, including time deadlines for suspension of sites.**

  - Failure to apply security patches is one of the main sources of incidents
  - Two EGEE wide security vulnerabilities last year:
    - First one in August, EGEE completely patched after 2.5 months
    - First time a security patch was enforced at the sites, at all levels, including PMB
    - 60+ sites warned about suspension within 7 days, only 4 suspended
    - Second one in November, EGEE patched after 1 month
    - 30+ sites warned about suspension, none suspended

- **Effective security patching monitoring: Security monitoring tools developed and deployed to prevent security incidents and identify security weaknesses into sites**

- **Extensive connections outside the project**

  - Channels between National Research and Education Networks (NRENs) and EGEE to ease the communication between each ROC security contact and the respective NREN

**Enabling Grids for E-sciencE**

- **Grid Security Vulnerability Group**
  - Aims to help identify security weaknesses inorder to prevent incidents in deployed Grid middleware.
  - 33 issues handled during the last year (60 since the start of EGEE-III)
  - Transition to EGI Software Vulnerabilities Group (SVG) in progress, covering all the software distributed by EGI
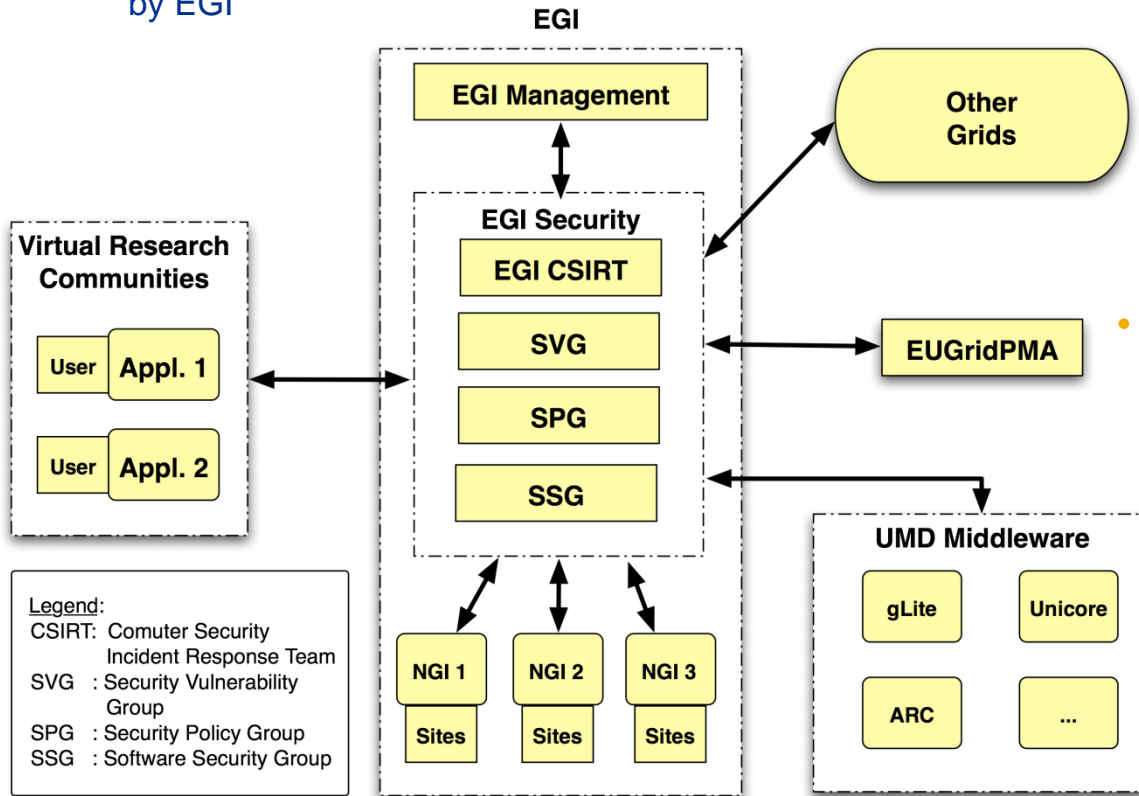


- **Joint Security Policy Group**
  - 5 old policies revised and 2 new polices:
    - VO Registration, VO Membership Management, Security Incident Response, Grid AUP, Site Registration
    - User-level Job Accounting, VO Portals
  - Security Policy Framework developed to present the current policies and to move forward in EGI

- **International Grid Trust Federation**
  - Bridged Research & Education federated ID and grid certificates
  - New guidelines for key management and the use of automated clients published

**Enabling Grids for E-sciencE**

**The EGEE helpdesk is a distributed system based on regional ticketing systems with central coordination and a central entry point for user support (GGUS).**

- **It is the only front-end exposed to end users**

  – Only developers/testers submit bugs directly to Savannah

  – GGUS tickets leading to Savannah bugs are set to "in *progress*" state, and automatically closed if the related Savannah bug reaches status "Ready for review".

  - 85 tickets ended in Savannah bugs since January 2010

- **Changes:**

  – Direct routing to sites (bypass 1st line support)

  - 21 % reduction in tickets handled by TPM in last year (2023 versus 3912)

  - ~80% tickets used direct routing, only 20% went through TPM

  - Additional automation analyzed and implemented as much as possible (cost of implementation versus ticket reduction):

    - *Direct assignment to some VOs and SUs, ALARM and TEAM tickets, Ticket escalation by submitter*

## User support tickets created per VO in the last year
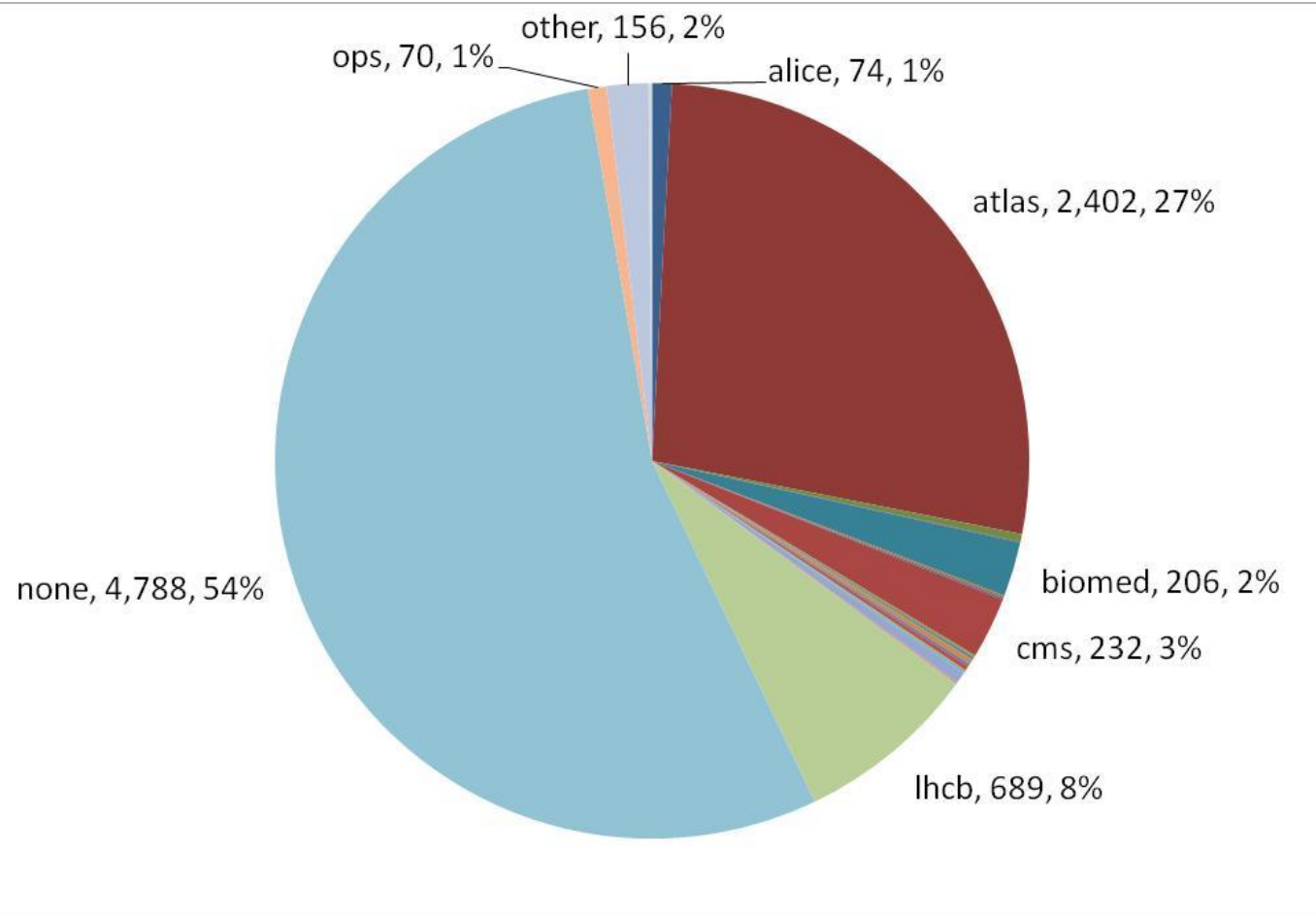


- **None:** tickets where the submitting user does not explicitly define the associated VO

- **Atlas** is the major user (27%), followed by **LHCb** (8%), **CMS** (3%) and **biomed** (2%)

- Each VO has different support model

**User support tickets created in the last year: per solution support unit**
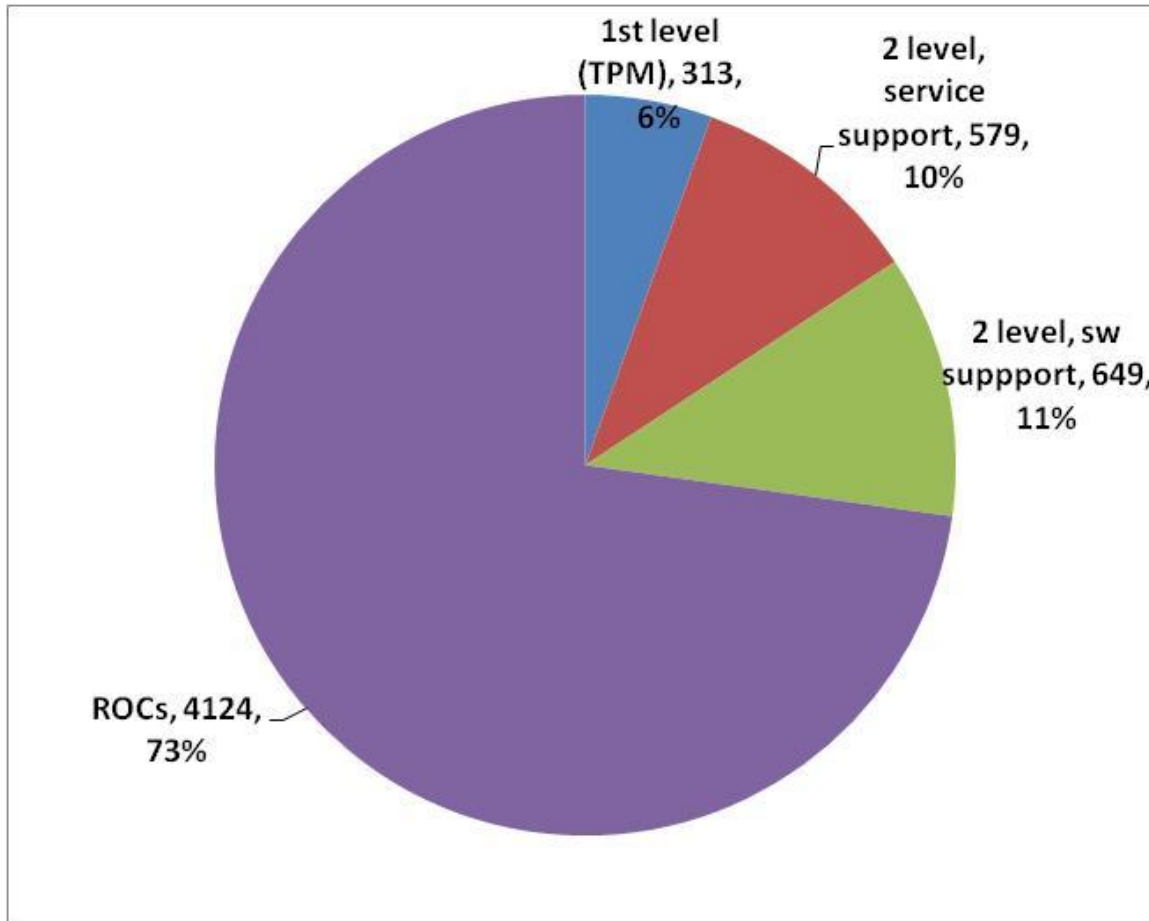


- **Most tickets solved at ROCs/sites, 73%**
  - **Savannah connection, problem types?**
- **1st line support (TPM) also solving tickets, 6%**
- **Specialized support units (2 level) solving 21% of the tickets**

## Average solution time per VO (NEED log scale to show low Vos?)

**Enabling Grids for E-sciencE**
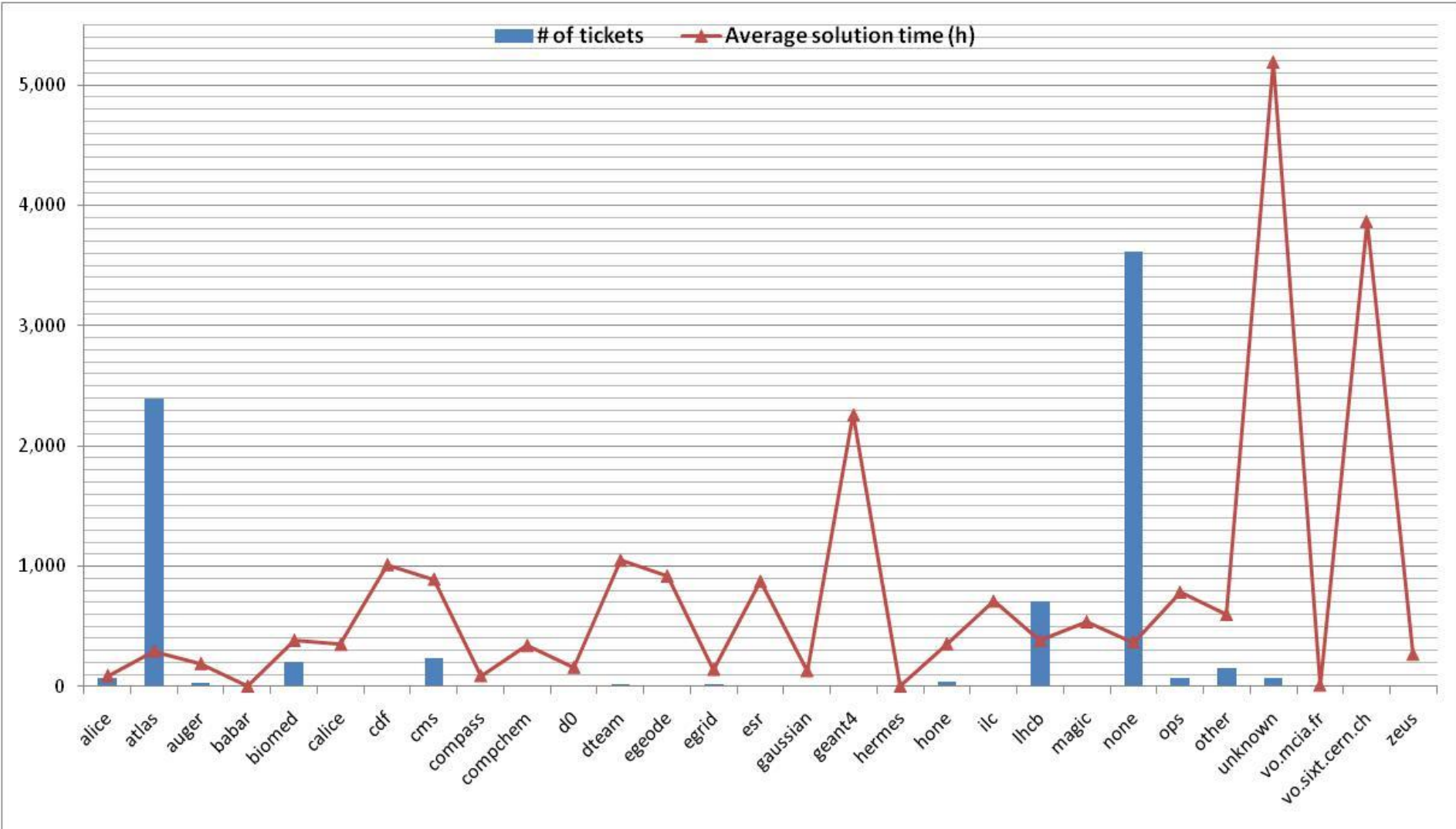
- **Role of oversight and 1$^{st}$ level support for grid production infrastructure**
  - teams of operators on duty (COD) opening tickets to sites in case of grid monitoring alarms

- **New model, based on the devolution to regions, fully implemented since summer '09**
  - First-line support done by each region, plus common layer for procedures, tools, escalation

- **Improvements:**
  - Regional teams closer to the sites
    - Aware of different site deployment scenarios
    - More incentivised to prevent incidents
  - Shorter response time: 5.75 days
    - Before regionalisation average response time 8 days
  - Distribution of responsibilities to ROCs → NGIs
    - Skills at ROC level will get partitioned when ROCs become NGIs

**Enabling Grids for E-sciencE**

## Tickets opened in the last year
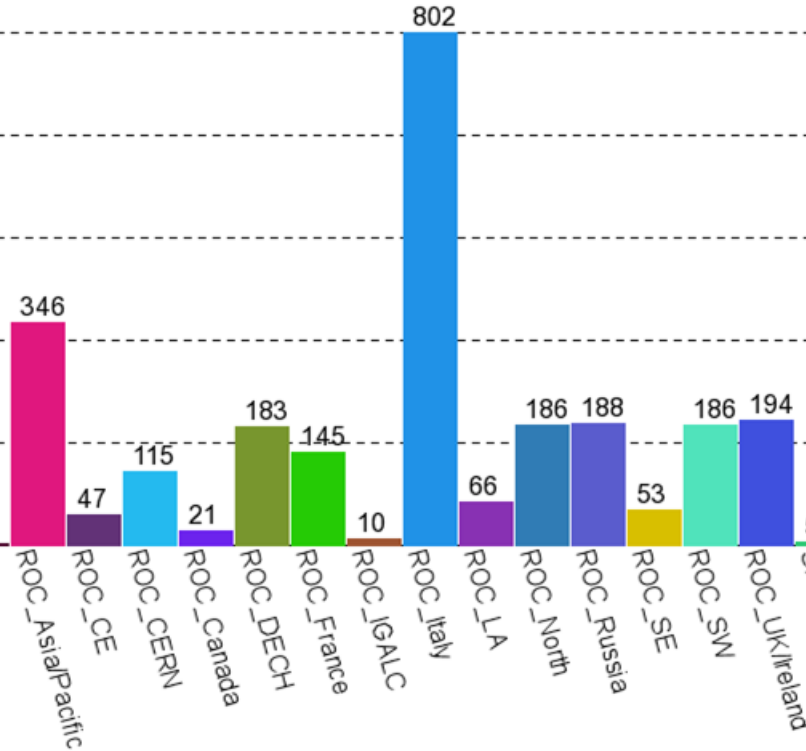
## Tickets opened per month
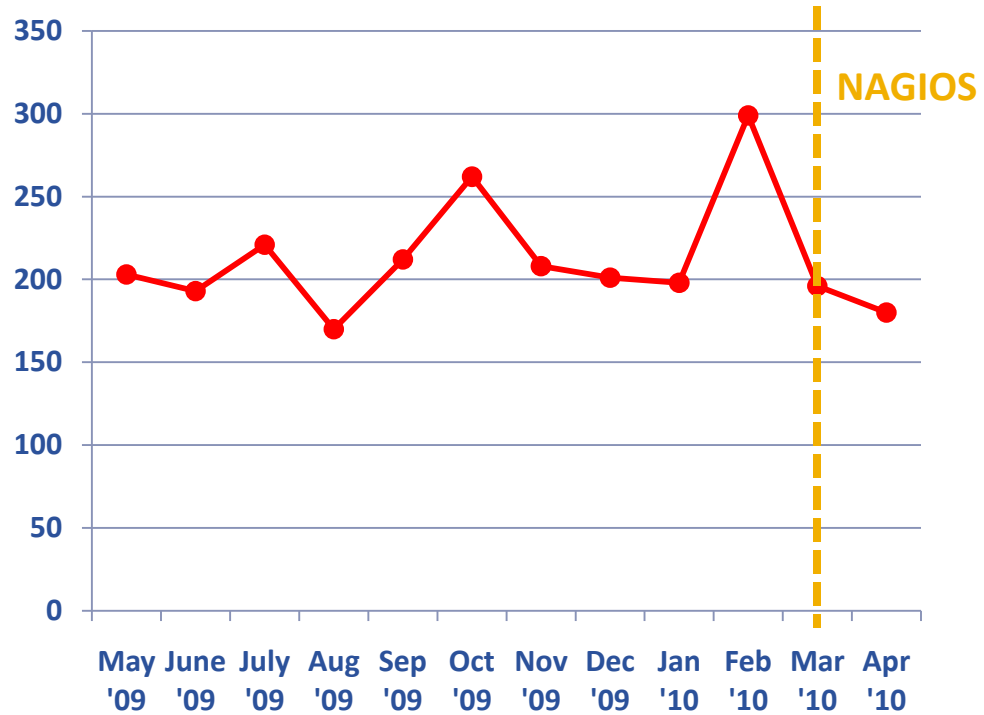
- 2570 tickets opened in the last year with average response time 5.75 days
- Stable number of tickets opened by month, also after migration from SAM to Nagios
  - Move to new technology did not impact operations
  - Expect reduction in regional alarms as many sites deploying Nagios
    - Sites see issues before they are reported to them by RODs

**Enabling Grids for E-sciencE**

- **Aims**
  - Improve reliability and availability of sites via operational tools
  - Prepare operational tools for use in an EGI/NGI structure

- **Achievements:**
  - Site and Regional Monitoring via Nagios, a commodity open-source monitoring framework
  - Better tools for site admins → faster problem detection & solution
  - Operations dashboard with regional views, to raise alarms and GGUS tickets, used by R-COD
  - Integration of operational tools via ActiveMQ, an open-source enterprise messaging system
  - MyEGEE visualization portal, collaboration with OSG
  - Availability/Reliability now calculated from regional Nagios results

- **Last 2 years:**



- **Now:**

**Enabling Grids for E-sciencE**
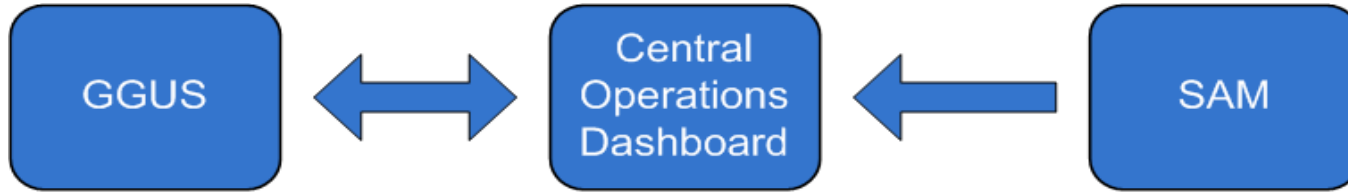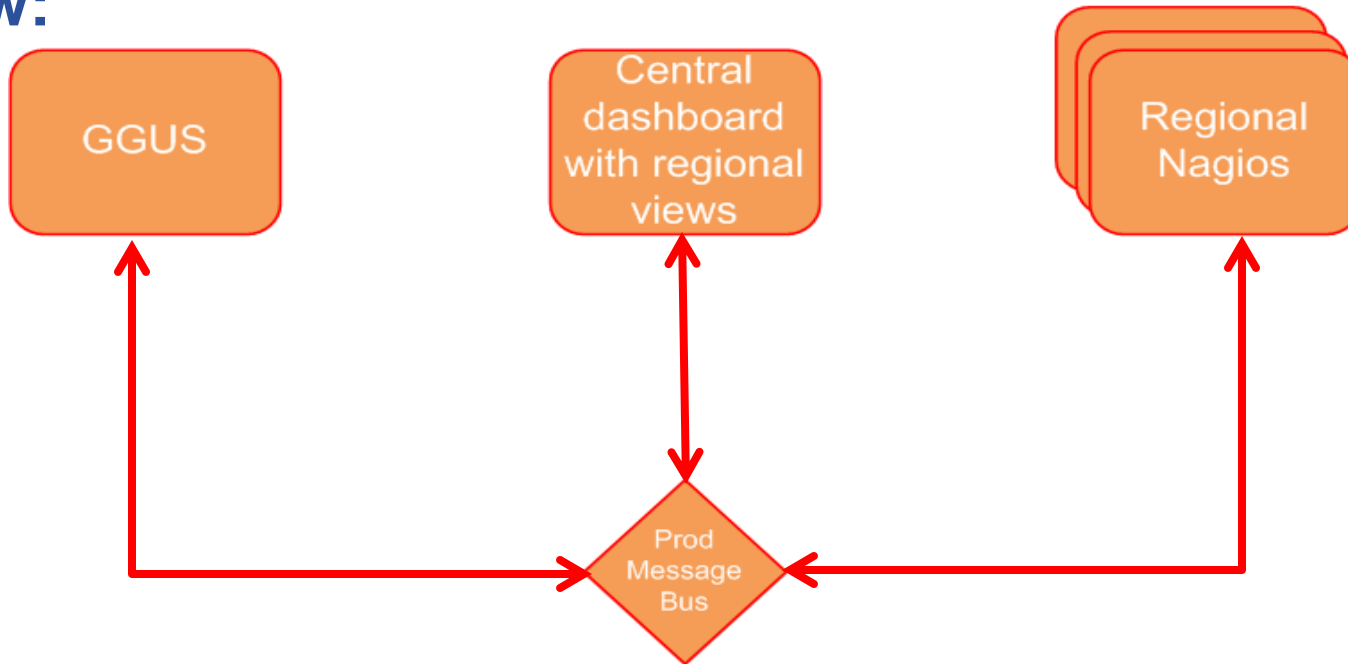
- **The process to start integrating NGIs into the grid operations structure started in the last months of the project**
  - 2 NGIs from 2 different ROCs (Poland from CE and Greece from SEE) volunteered to start and prototype the process
  - Included registration to managerial and operational structures, tool deployment, process adaptation when needed
  - Result: process prototyped, known and documented. Being followed in parallel by many other NGIs
  - Main difficulties: tight coordination needed between tools and teams to enable the new NGI
  - In both cases NGI staff overlapped with ROC staff with cumulated knowledge/experience. This is not the case for many NGIs, which will needed close support and training.

- **Regionalization is expensive due to extra layers**
  - Move to national based infrastructures has 'cheaper' core

- **Operation tools**
  - Importance of agreeing on formats for information exchange between tools
  - Rely where possible on widely adopted public domain solutions (e.g. Nagios)
  - Need to have a more integrated development activity

- **Provided a stable and reliable infrastructure**
  - While doubling (approx.) in capacity
  - Significant operational changes
- **Need to improve operational tools**
  - Home-grown tools & infrastructures costly to maintain and support
- **Good transition to EGI, no major issues left**
  - Full Implementation of ROC → NGI ongoing
- **Global scientific collaborations relying on it**
  - WLCG main user group, solid base for ESFRI engagement
- **Collaboration was essential, important not to break this with regionalization/NGIs**