



# EGEE-III

## OPERATIONAL PROCEDURES MANUAL FOR ROCS AND SITE

Document identifier:	EGEE-III-SA1-840932- OperationalProceduresforROCsAndSites-v4.0.odt
Date:	9.04.2010
Activity:	SA1: Operations
Document status:	Released 4.0

Document link: https://edms.cern.ch/document/840932

#### OPERATIONAL PROCEDURES MANUAL FOR ROCS AND SITE

Date: 9.04.2010

#### Copyright notice:

Copyright  $\ensuremath{\mathbb{C}}$  Members of the EGEE-III Collaboration, 2008.

See <u>www.eu-egee.org</u> for details on the copyright holders.

EGEE-III ("Enabling Grids for E-science-III") is a project co-funded by the European Commission as an Integrated Infrastructure Initiative within the 7th Framework Programme. EGEE-III began in May 2008 and will run for 2 years.

For more information on EGEE-III, its partners and contributors please see <u>www.eu-egee.org</u>

You are permitted to copy and distribute, for non-profit purposes, verbatim copies of this document containing this copyright notice. This includes the right to copy this document in whole or in part, but without modification, into other documents if you attach the following reference to the copied elements: "Copyright © Members of the EGEE-III Collaboration 2008. See <u>www.eu-egee.org</u> for details".

Using this document in a way and/or for purposes not foreseen in the paragraph above, requires the prior written permission of the copyright holders.

The information contained in this document represents the views of the copyright holders as of the date such views are published.

THE INFORMATION CONTAINED IN THIS DOCUMENT IS PROVIDED BY THE COPYRIGHT HOLDERS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE MEMBERS OF THE EGEE-III COLLABORATION, INCLUDING THE COPYRIGHT HOLDERS, OR THE EUROPEAN COMMISSION BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THE INFORMATION CONTAINED IN THIS DOCUMENT, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Trademarks: EGEE and gLite are registered trademarks held by CERN on behalf of the EGEE collaboration. All rights reserved"

# **Table of Contents**

1 Operational Procedures Manual for ROCs and Sites	1
Revision history	1
1.1 Preface	1
2 Introduction	
2.1 Structure of this manual	
2.2 Role of Site Admin	
2.3 Role of 1st Line Support	
2.4 Role of Regional Operator (ROD)	
2.5 Role of Central COD (C-COD)	4
3 Getting Started	5
3.1 First time for ROC Managers	
3.2 First time for Sites and Site Administrators	
4 Sites and Site Administrators	6
4.1 Duties and Obligations.	
4.1.1 Communication Lines - contacting 1st Line Support, ROD	
4.1.2 Responding to a request to act on an incident	
4.1.3 Modifying/Updating tickets	
4.1.4 Providing Information in the Dashboard Notepad	
4.1.5 Communication Lines - contacting C-COD.	
4.1.5 Communication Lines - contacting C-COL	0
5 Current administrative operations	
5.1 GOC Database	
5.1.1 Site Status Flow	
5.1.2 Introducing a new site	
5.1.3 Site downtime scheduling	
5.1.4 Removing problematic sites	
5.1.4.1 Emergency suspension	
5.1.5 Removing unused sites	9
5.1.6 Removing resources	10
5.1.7 Emergency contacts	
5.2 Intervention Procedures	10
5.2.1 Types of Interventions	
5.3 Incident reporting	
5.4 Weekly Reporting	11
6 Monitoring Sites	12
6.1 Monitoring Tools	
6.2 NAGIOS Availability Monitoring (MyEGEE)	
6.2.1 ROC_OPERATORS profile	
6.2.1.1 Procedure to add a test to the ROC_OPERATORS profile	
6.2.2 Alarm generation in the Operations Dashboard	
6.3 GIIS Monitor	
7 Departing Problems	17
7 Reporting Problems 7.1 Problem detection	
7.1 Problem detection 7.2 Workflow and escalation procedure	
7.2 worknow and escalation procedure 7.3 Suspending a site	
7.4 Sites that fail NAGIOS tests but still continue to be operational	
7.5 Ticket handling during Weekends and Public Holidays	
7.5 TICKET Handling untilg weekeinds and rubile Holldays	1/

# **Table of Contents**

8 Security Items and Daily Operations	
8.1 Security matters	19
8.2 OSCT organization and OSCT-Duty Contact Role	19
8.3 Security incidents handling and Interaction with OSCT-DC	
8.4 Grid Security Vulnerability Handling	
Reason for a Vulnerability handling process	
Reporting a Vulnerability.	
Handling of issues reported	
Risk Categories	
On resolution or reaching the Target Date	
Further Details	
Other notes	
9 References	

# 1 Operational Procedures Manual for ROCs and Sites

## **Revision history**

Comment	Date	Version	Author
Complete revision accounting	March 2010	4.0	Vera Hansper, Malgorzata Krakowian, Peter
for new NAGIOS tools			Gronbech
Site status flow explainations added	23 Dec 2009	3.2	Michaela Lechner
Revision to reflect new downtime procedure	18 Nov 2009	3.1	Helene Cordier, Michaela Lechner
Procedures for ROCs and Sites in regional operations model	18 Sept 2009	3.0	Vera Hansper, Malgorzata Krakowian, Helene Cordier, Michaela Lechner,Ioannis Liabotis, Peter Gronbech
Changes to reflect demise of old COD model	01 July 2009	2.1	Vera Hansper
Procedures for Sites and ROCs	27/02/2009	2.0	Ioannis Liabotis
Various Changes	03 November 2008	1.8	David Bouvet, Helene Cordier, John Shade, Peter Gronbech, Ioannis Liabotis
Node removal	29 August 2008	1.71	David Bouvet
Site suspension on extensive downtimes	10 July 2008	1.70	Ioannis Liabotis
Updates Based on COD 15	09 May 2008	1.60	Ioannis Liabotis
Several additions and reorganization	20 July 2007	1.50	Ioannis Liabotis, Philippa Strange, Kai Neuffer
Update based on COD 12	16 Apr 2007	1.40	Ioannis Liabotis, Philippa Strange, Kai Neuffer
Updates based on COD 11	8 Nov 2006	1.30	Philippa Strange, Ioannis Liabotis
Added to SA1 twiki	24 Oct 2006	1.21	Ioannis Liabotis
Revisions to format and content	26 Sept 2006	1.20	Philippa Strange
Revisions to format and content	5 July 2006	1.10	Philippa Strange
Updates to First revision	19 April 2006	1.01	Helene Cordier&Philippa Strange
First Revision 2006	20 March 2006	1.00	Helene Cordier
Revisions to content – complete reformatting	20 November 2005	0.13	Helene Cordier
Minor revisions to content	20 October 2005	0.12	Helene Cordier
Minor revisions to content – complete reformatting	10 October 2005	0.11 0.2	Alistair Mills
Initial draft		0.01	Piotr Nyczyk&Helene Cordier

## 1.1 Preface

The EGEE Operational Procedures Manual (OPS Manual) defines the procedures and responsibilities of the various parties involved in the running of the EGEE infrastructure, namely the resource centers (also referred to as 'sites') consisting of local support and sites administrators, the staff of the Regional Operations Centres (ROCs) like the ROC Manager and the ROC support staff, the regional operations team consisting of the Regional Operator on Duty (ROD) and the 1st Line Support and the oversight grid monitoring operators (also

referred to as 'C-COD'). The OPS manual is currently structured into three separate documents, one covering the responsibilities of ROCs and Sites, one detailing the procedures and responsibilities of the regional operations team and its ROC, the last detailing those for the C-COD team members. To avoid dispersing the same information on multiple documents, and to allow for an easier update of the information, there is a fourth separate document containing the sections that are of relevance to the other manuals, the OPS 'common sections' manual. Sections of this document will be included in the relevant sections of the other three manuals using the twiki INCLUDE mechanism.

The OPS Manual for regional operations related to ROCs and Sites can be found at:

https://twiki.cern.ch/twiki/bin/view/EGEE/OperationalProceduresforROCsAndSitesObsolete09042010

The OPS Manual for the regional operations team can be found at:

https://twiki.cern.ch/twiki/bin/view/EGEE/OperationalProceduresforRODObsolete09042010

The OPS Manual for C-COD can be found at:

https://twiki.cern.ch/twiki/bin/view/EGEE/OperationalProceduresforCCODObsolete09042010

The above procedures can be also found in EDMS at: https://edms.cern.ch/document/840932

Readers of any one of the above manuals are encouraged to also read the other manuals in order to have a complete picture of daily operations within EGEE.

Please verify that the document you are using is the current release. The document can be found from:

- the CERN twiki
- EDMS repository

This document does not describe future procedures. It describes the tools and procedures used currently to operate the EGEE production service according to the operation model, defined at: https://edms.cern.ch/document/971628

Since operation of large scale production grids are not static we expect the document to be changed regularly. Major changes will be announced.

- Change requests come from users of the Operational Procedures Manual. Those are: Site managers, ROD, C-COD and SA1 Operational Teams
- To request a change in the manual any interested party should open a ticket via GGUS specifying the "Type of problem" as: "ROD and C-COD Operations";
- According to the type of change requested the following procedure is followed:
  - Significant changes in procedures and subsequent updates are discussed at C-COD/ARM meetings, which occur quarterly. These requests will be dealt with according to their priority and level of effort required to make the changes.
  - ♦ For urgent changes, proposals have to be discussed on the C-COD and ROC-managers mailing lists and agreed at the following Weekly Operations Meeting by a large majority of operational people.
  - When agreed and validated these changes are implemented and the procedure coordinator will release a new version of the document.
  - New versions of the Operations Manual that contain changes in the procedures will be BROADCASTED to all sites and ROCs via the broadcasting tool.
  - Small typographical and grammatical changes can be made without discussion or approval and published at short notice. No BROADCAST is sent for these types of changes.

# **2 Introduction**

Staff responsible for the operations of the EGEE grid, are broken up into the following areas:

- Operations and Co-ordination Centre OCC top level management responsible for all operations.
- Oversight Team C-COD (consisting of volunteer ROD representatives).
- Regional Operations Team(s) ROD and 1st Line Support (can be one team or two separate teams).
- Regional Operations Centre ROC Managers, ROC support staff.
- Resource Centres (sites) local support, site admins.

The Regional Operations team is responsible for detecting problems, coordinating the diagnosis, and monitoring the problems through to a resolution. This has to be done in co-operation with the Regional Operations Centres to allow for a hierarchical approach and overall management of tasks. ROCs decide themselves on how to manage Regional Operations and whether they wish to have ROD and 1st Line Support as one team or two separate teams.

Procedures have to be followed according to the formal descriptions to ensure predictable work flow and reduce effort duplication or no action at all.

The procedures described in this document refer to the responsibilities of ROCs and Sites that participate in the EGEE infrastructure.

## 2.1 Structure of this manual

This manual in several sections:

- Section 1 : Defines the purpose of the global operations manual, its structure and its process for upgrade ;
- Section 2 : Provides an introduction to the manual and a description of the duties and roles of the regional operations model ;
- Section 3 : Describes the actions needed for people to become operational within SA1 ;
- Section 4 : Describes the duties and obligations of site administrators ;
- Section 5 : Describes administrative tasks using the GOC Database such as adding new sites, changing details and scheduling downtime ;
- Section 6 : Describes the monitoring infrastructure and what the site administrators and ROCs should know about it ;
- Section 7 : Describes the effect to sites and ROCs of the ROD monitoring and problem reporting activities ;
- Section 8 : Describes Security Items and how to handle them ;
- Section 9 : Provides a table of references including web addresses.

## 2.2 Role of Site Admin

In the scope of Regional Operations, site administrators primarily receive and react on notification of one or more incidents. They should also provide information in the site notepad, which is available on the dashboard.

## 2.3 Role of 1st Line Support

A team responsible for supporting the site administrators to solve operational problems. The team is provided by each ROC and requires technical skills for their work. Organization and the presence of such a team is optional. However, if this team does not exist explicitly, duties of 1st Line Support must be absorbed by the Regional Operator.

## 2.4 Role of Regional Operator (ROD)

A team responsible for solving problems on the infrastructure according to agreed procedures. They ensure that problems are properly recorded and progress according to specified time lines. They ensure that necessary information is available to all parties. The team is provided by each ROC and requires procedural knowledge on the process (rather than technical skills) for their work. However, if the team also encompasses the role of 1st Line Support, then the necessary technical skills will be required.

## 2.5 Role of Central COD (C-COD)

A small team responsible for coordination of RODs, provided on a global layer. C-COD represents the whole ROD structure at the political level. Support tools developers should interact with C-COD about the tools, especially when issues arise.

# **3 Getting Started**

## 3.1 First time for ROC Managers

When a new federation joins EGEE, a ROC manager is nominated. This ROC manager has to

- 1. request subscription to project-egee-ROC-managers@cernSPAMNOTSPAMNOT.ch, and obtain the check-list to set-up their regional operations (ROD) teams. That is, set up a generic mailing list for their team so that there is a point of contact with EGEE.
- 2. send an e-mail containing the ROD email contact mailing list (as in 1., above) to cic-information@in2p3SPAMNOT.fr or to project-eu-egee-sa1-cic-on-duty@cernSPAMNOT.ch. This mailing list will be included in project-eu-egee-sa1-cic-on-duty@cernSPAMNOT.ch, which is the mailing list for general interaction of all regional teams. This is used for ROD members to interact with ROC managers.
- 3. ask their ROD teams to go to the GOCDB at https://goc.gridops.org/role/request to request the "regional staff" or "regional 1st line support" role. Both roles enable the use of the tool dedicated to the regional operations teams available at https://operations-portal.in2p3.fr/"

NOTE: Remarks and suggestions for best practices on an international scale, however, should be sent with a cc: to project-eu-egee-sal-c-cod-followup@cernSPAMNOTSPAMNOT.ch or put the information/request into a GGUS ticket assigned to the C-COD GGUS Support Unit.

## 3.2 First time for Sites and Site Administrators

Read and become familiar with this document, specifically the parts relevant to their role;

- Have a valid certificate delivered by a CA;
- Register into the dteam VO: https://lcg-voms.cern.ch:8443/vo/dteam/vomrs;
- Register into GGUS as support staff;
- Request GOCDB read-only access by completing the form at the following URL: https://goc.gridops.org/user/register (Note, you will not be able to access this URL if you are already registered.)
- Request a Site Administrator role in the GOCDB at https://goc.gridops.org/role/request for the site or sites that they administer.

# **4 Sites and Site Administrators**

Sites are responsible for providing site services at an agreed level, thus their main task is to make sure site problems are solved efficiently.

## 4.1 Duties and Obligations

In the scope of Regional Operations site administrators primarily receive and react on notification of one or more incidents. They should also provide information in the site notepad, available on the dashboard.

Site administrators can also view their site on the operations dashboard at the following URL: https://operations-portal.in2p3.fr/ .

#### 4.1.1 Communication Lines - contacting 1st Line Support, ROD

Site administrators can send a "request for help" to 1st Line Support and/or ROD through appropriate mailing lists. The contact lists for this should be provided by your ROC.

In general, communication with C-COD or ROD will be via the GGUS system when responding to tickets.

#### 4.1.2 Responding to a request to act on an incident

Sites should respond to requests to act on incidents in a suitable time frame. They will be notified of early alarms by 1st Line Support, or they can choose to monitor the dashboard directly for this information. (See https://operations-portal.in2p3.fr for the operations dashboard.) ROD will create a ticket for any alarm that is not acted on, or solved, within the first 24 hours.

### 4.1.3 Modifying/Updating tickets

Sites can respond to tickets directly, via email, or through the GGUS interface.

#### 4.1.4 Providing Information in the Dashboard Notepad

It is mandatory that sites provide information about their incidents, status of downtimes, etc. in the site notepad. Solutions to a problem can also be put into the site notepad. The site notepad is accessible from the dashboard: https://operations-portal.in2p3.fr/.

#### 4.1.5 Communication Lines - contacting C-COD

Sites can create tickets (GGUS) to C-COD for core or urgent matters. These can also be submitted by or through ROD for validation.

## **5 Current administrative operations**

There are a few administrative operations for ROC and site staff:

- GOCDB entries
- Weekly Reports
- EGEE Broadcasts

## 5.1 GOC Database

The GOC Database (GOCDB) is a core service within EGEE and is used by many tools for monitoring and accounting purposes. It also contains essential static information about the sites such as:

- site name;
- location (region/ country);
- list of responsible people and contact details (site administrators, security managers);
- list of all services running on the nodes (CE, SE, UI, BDII etc.);
- phone numbers.

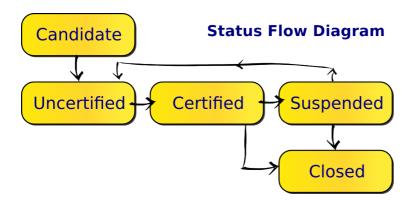
Site administrators have to enter all scheduled downtimes into the GOCDB. The information provided by GOCDB is an important information source during problem follow-up and escalation procedures.

#### 5.1.1 Site Status Flow

The GOCDB also provides information about a site's Certification Status. The possible statuses are as follows:

- Candidate: A site has just been declared to the GOCDB and information is still not complete.
- *Uncertified*: ROC has validated the site info. Clarification:
  - The *uncertified* status would generally indicate that a site is ready to start the certification procedure (again).
  - It can also be used as a timewise unlimited state for sites which have to keep an old version of the middleware as a requirement of an international VO or to flag a site coping with the ROC's requirements but not with EGEE availability/reliability thresholds.
- *Certified*: ROC has verified that the site has all middleware installed, passes the tests and appears stable.
- *Suspended*: A site temporarily does not conform to EGEE production requirements (e.g. EGEE SLAs, security matters) and requires ROC attention. Clarification:
  - Suspended is a temporary state. It is used to flag a site which is temporarily not meeting EGEE availability/reliability thresholds. If no change occurs in the site's availability it should be closed or uncertified by its ROC within 4 months. While suspended, sites can express their wish to pass certification again. The *suspended* status is useful to EGEE and to the ROCs themselves to flag the sites that require attention by the ROCs.
- *Closed*: A site is definitely no longer operated by EGEE and is only shown for historical reasons. Clarification:
  - The *closed* status is a terminal state. The site is no longer part of the infrastructure.

These definitions can also be found in the GOCDB user documentation. In order to avoid unclear state transitions only certain transitions are allowed by the GOCDB. A visualization of the guideline for a site's state transitions during normal operations is shown below.



Allowed examples of site state transition:

- suspended -> uncertified -> certified
- certified -> suspended -> closed
- *certified -> closed* (on site request)

Explicitly forbidden examples of site state transitions:

- *suspended* -> *certified*
- *candidate ->* something else but *uncertified* and *closed*
- *closed* -> anything else

An important note for the *suspended* status: ROC managers must regularly attend to all of their suspended sites, so that they are processed within the given maximum time of four months. Sites which are *suspended* should either be brought back in production via the *uncertified* status within that time frame **OR** set to *closed*.

#### 5.1.2 Introducing a new site

A new site must first contact the local ROC as detailed in the Site/ROC Association document [13]. As part of the process, the ROC registers the new site in the GOCDB under the status *candidate*. The global monitoring within GOCDB for the new site should be set to off. Monitoring for the site cannot be switched on for this status. When the site is established, the ROC switches the site status to *uncertified*.

Monitoring may now be switched on at the discretion of the site administrators for commissioning.

Once the site information within the GOCDB is completed and the site's specific tests set of the local NAGIOS instance have been in an OK state for a week the site status can be changed from *uncertified* to *certified* by its own ROC. Monitoring for the site and its nodes will now be switched on, and cannot be switched off.

The site is now classed as being in EGEE production: no rollout announcement is necessary.

#### 5.1.3 Site downtime scheduling

EGEE resources need to be switched off properly so that they do not disturb operations during maintenance periods.

When a site needs to upgrade, it should specify the downtime beforehand for any or all of the nodes involved. Instructions in such a case are:

• Set the downtime period in GOCDB and choose the criticality. A broadcast will be automatically sent to the relevant targets.

- The downtime procedure documentation can be found at: https://edms.cern.ch/document/1032984 .
- Ensure that the queues for the nodes affected are properly closed. It might be necessary to drain them first. At least, the GIIS should indicate that the following values are correct:
  - "GlueCEStateStatus: Draining" (while queues are draining.)
  - "GlueCEStateStatus: " is not available to gstat (while the nodes are in downtime.)

There is also help at the GOCWIKI:

http://goc.grid.sinica.edu.tw/gocwiki/GOCDB\_User\_Documentation#head-260f1818bcc28e952004775a41c1a526eac6

NOTE: For a downtime to be scheduled, it must be declared at least 24 hours in advance. Downtime declarations must follow specific rules given in the Intervention Procedures below. (§ 5.2).

#### 5.1.4 Removing problematic sites

The ROC must suspend one of their sites in either of the following two cases:

- For problems that are not addressed by either the site or the ROC: C-COD will apply an "escalation step procedure" described in the Workflow and Escalation Procedure section. The final step is suspension, and the site is taken out of the grid resources. For the site to be suspended, a given ROC would have to disregard answering several emails over a period of more than two weeks and not join the Weekly Operations Meeting when asked to. As soon as the site's status is modified, the ROC should be sent another notification mail.
- If a site is in downtime for more than one month: Before suspending the site the ROC should investigate the situation to exclude the possibility that the site will be up again in a short period of time. (i.e. within a few days.)

Before a site is suspended, the site's ROC should make contact with the senior people in the federation, the site and C-COD. If a site later indicates its readiness to rejoin the infrastructure, it will need to go through the certification procedure again before its status is set to *in production*.

#### 5.1.4.1 Emergency suspension

A site may be suspended directly in emergency cases, e.g. security incidents. The escalation procedure is then by-passed and either the ROC, ROD or the C-COD may set the site in suspension in agreement with the ROC's corresponding ROD. (In general, ROD can place a site in downtime if it is either requested by the site, or ROD sees an urgent need to put the site into downtime.)

Under exceptional circumstances, C-COD may suspend a site immediately without going through all the steps of the escalation procedure. The most important scenario for this is when a security incident occurs, and efforts by C-COD to contact the site's ROD or the site directly are unsuccessful. It is preferable that in such cases that the site and ROD are contactable, in which case the site can close down its resources and ROD can issue an immediate downtime. (A similar scenario is applicable for ROD where ROD is not able to contact the site in question.)

In all situations it is important that communication channels between all parties involved are active, and that C-COD (or ROD) inform the ROC and its site that the suspension has occured.

#### 5.1.5 Removing unused sites

There is a site status in GOCDB called *closed*. This status is for sites which are no longer in use because they have either been closed or replaced by another site. The status is essential for accounting to provide the correct information about the site. This lets the ROC get a historical reference of that site.

### 5.1.6 Removing resources

When a *resource or node* (not site) is to be removed from the GOCDB, it's also important that it is no longer published in the site-BDII. The SAM DB is still used for availability and reliability calculations, that is, it collects both static (GOCDB) and dynamic (BDII) information. However, currently, the Nagios tests will recognize the node as 'deleted', after it has disappeared from the GOCDB . It remains a good practice to remove the information from the BDII none-the-less as the current situation will change in the near future.

### 5.1.7 Emergency contacts

Currently the GOCDB also handles sites that are not committed to EGEE, but are part of the WLCG as far as EGEE/LCG operations are concerned. These sites are connected to ROCs outside of the EU: LA (Latin America), CA (Canada), IGALC (Latin America and Caribbean Grid Initiative).

Phone contacts for all staff connected to sites and ROCs are available within the GOCDB and VO contacts are available from the Operations Portal. Please see https://goc.gridops.org/user/security\_contacts .

NB: Since a recent revision of the GOCDB it is now possible to download these data on a regular basis through GOCDBPI. It is advised that the operations teams download and update this list once a month.

## **5.2 Intervention Procedures**

The following two sections are derived from from the EGEE intervention procedures as currently defined in those documents: EGEE-intervention procedures&Notification mechanism document.

This section will evolve as intervention procedures are updated.

### 5.2.1 Types of Interventions

Procedure:

- Declare interventions at least 24 hours in advance, specifying reason and duration.
- Any intervention declared less than 24 hours in advance will be considered unscheduled.

Best practice recommendations:

- For interventions that impact end users, declare downtime 5 working days in advance, specifying reason and duration.
- Sites shall declare unscheduled interventions when problems arise. They should be declared as soon as they are detected to inform the users. Unscheduled interventions can be declared up to 48 hours in the past in order to provide retroactive information for the user community. A post-mortem shall then be included in the next ROC report after the intervention.

## 5.3 Incident reporting

Grid site security contacts and site administrators should be aware of the Incident reporting procedure aimed at minimising the impact of security incidents by encouraging post-mortem analysis and promoting co-operation between the sites. This procedure is available at: http://osct.web.cern.ch/osct/incident-reporting.html.

## 5.4 Weekly Reporting

Each Regional Operation Centre (ROC) is obliged to fill in the Weekly ROC Report. This is used to keep track of operational issues and for these issues to be discussed in the Weekly Operations Meetings that occur at CERN every Monday at 16:00 CET.

#### Regional Operations Centre weekly reports are available for editing and viewing at

https://cic.gridops.org/index.php?section=roc&page=rocreport and can be completed from Saturday at 06:00 (GMT) until Monday 12:00(GMT). Note that submitting and approving a weekly report is compulsory for all ROCs regardless of whether they had problems or not.

Notifications that the weekly reports are due are sent to project-egee-roc-managers@cernSPAMNOTSPAMNOT.ch and advertised via the broadcast tool to all ROCs.

# **6 Monitoring Sites**

## 6.1 Monitoring Tools

The links below are for direct access to some of the monitoring tools.

- Nagios monitoring: https://sam-%INSERT\_YOUR\_ROC%-roc.cern.ch/myegee. There is a page for each ROC and the details of each ROC's URL can be found on https://twiki.cern.ch/twiki/bin/view/EGEE/NagiosROCURL ;
- GIIS Monitor: http://gstat-prod.cern.ch/gstat/summary/;
- GOC Database. https://goc.gridops.org/;
- GOC Real Time Monitor. http://gridportal.hep.ph.ic.ac.uk/rtm/;
- Google Earth. Add the link http://dashb-earth.cern.ch/dashboard/dashb-earth-all.kmz;
- GridView. http://gridview.cern.ch/GRIDVIEW/;
- GridMap. http://gridmap.cern.ch/gm/;

While the above tools are used on a daily basis there are many more monitoring systems available. For a current overview the GOC page on monitoring can be consulted:

• http://www.ukiroc.eu/content/view/117/265/.

## 6.2 NAGIOS Availability Monitoring (MyEGEE)

One of the most detailed information pages about possible problems is the MyEGEE Portal, https://sam-%INSERT\_YOUR\_ROC%-roc.cern.ch/myegee (1). On the front page it also gives two options, *Resource Summary* and *Status History*. There is also a link to the help pages for that portal on the right hand side of this page. The more useful option for operations is the status history.

#### **Status History:**

In MyEGEE you can view the status of a site and service according to different profiles by selecting the relevant profile in the 'Profile Selection' drop-down menu. For operational purposes, the ROC\_OPERATORS profile should be used. By careful manipulation of the parameters on the right side of the Status History page, you can select the sites and resources of interest to you for viewing. It will be useful to familiarise oneself with this tool by reading the relevant help pages at

https://sam-%INSERT\_YOUR\_ROC%-roc.cern.ch/myegee/help.html .

Two options are possible for further analysis:

- clicking on the history bar at or near a point of failure,
- selecting the table tab.

(Note, it is useful to select a shorter date range at this point, or select the *Current SAM Status* (2) option.) Any failure observed here should be recognised, understood, diagnosed and the follow-up process should be started. This is especially important for failures of the critical tests, which should be treated with the highest priority.

The two options mentioned above provide an abbreviated status of the test results from the probes. They also contain detailed information about each test results by selecting the *Show Detail* widget of a particular result. The detailed reports should be studied to diagnose the problems.

(1) Please replace the whole string %INSERT\_YOUR\_ROC% with your ROC name in capitals, i.e. UKI or SWE.

#### 6 Monitoring Sites

(2) Please note that SAM in this context does not refer to the old SAM tests.

## 6.2.1 ROC\_OPERATORS profile

In the Nagios system there is a set of profiles corresponding to groups of probes instead of a single list of critical tests. A profile is therefore a mapping from a service type to a list of metrics (tests) e.g. for a CE there are about 10 metrics gathered.

A profile may be used for different functionality requirements e.g. for raising notifications to operators (ROC\_OPERATORS profile), or for the calcuation of availability (ROC\_CRITICAL profile - note that this will be renamed to ROC\_AVAILABILITY). The list of tests in the ROC\_OPERATORS profile for Nagios can be found here: https://twiki.cern.ch/twiki/bin/view/LCG/SAMCriticalTestsForCODs .

Over time more profiles will be created for specific tests, or tests which will be applied to specific sets of resources e.g. MPI tests, which should only run on resources which support MPI.

Currently the *old* SAM results are still used for site availability calculations. More information can be obtained on the *(old)* SAM wiki pages dedicated to the *(old)* SAM critical tests for Availability Metrics calculations: https://twiki.cern.ch/twiki/bin/view/LCG/EGEEWLCGCriticalProbes

#### 6.2.1.1 Procedure to add a test to the ROC\_OPERATORS profile

Tests can be added to the ROC\_OPERATOR profile so that they generate an alarm in the Operations Dashboard when the tests return an error status.

This change to the ROC\_OPERATOR profile must be agreed upon by various management boards, and the procedures outlined below to define a test as critical for the OPS VO must be scrupulously followed.

Please note that in the scope of this document only OPS critical tests are considered.

"Candidate critical tests requests" i.e. new tests which will raise alarms on the dashboard upon failure and which operators on shift should handle are reviewed during the quarterly ROD Forum meetings. Urgent requests can come from the ROCs directly.

In face to face meetings, regular phone meetings or by mail, Operations Procedures (aka 'Pole 2') members (currently project-eu-egee-sa1-cod-pole2-proc@cernSPAMNOT.ch) will:

- gather input from people attending the ROD Forums meeting and make recommendations to the Nagios team,
- check that all the relevant activities and tools are ready to handle the new critical test.

This can be summarised in the following check-list:

- 1. Gather ROD feedback on the test implementation;
- 2. Gather ROD feedback on the impact of the general activity;
- 3. Record feedback and final recommendations in a GGUS ticket referenced in the wiki;
- 4. Check that the template mail in the dashboard accounts for the new critical test;
- 5. Check availability of appropriate links in the dashboard for the new critical test;
- 6. Check that appropriate instructions are defined in the Operations Procedure Manual.

Once ROC managers validate the new test to become critical for the OPS VO, it is announced to all the sites/ROCs/C-CODs via a BROADCAST during Week (W-2), and announced at the Weekly Operations Meetings one week later (W-1). Based on all feedback gathered in the process, and that they have been approved, these tests will be become critical on week W at the Weekly Operations Meetings time.

The above procedure is summarized in the following list:

- 1. A request for a new test comes from VO/ROC/Operations
- 2. The Nagios team evaluate and integrate the test. ROC managers are also officially informed of the new test.
- 3. The test put into a candidate status to be available on the dashboard as a non-critical test, but still with final "Error", "Warn" or "Ok" flags.
- 4. A report is submitted by VO/ROC/Operations to ROC managers on the progress of the test. This should include details of sites failing the test (by region).
- 5. ROCs review results with their sites until 75% of sites pass the new tests.
- 6. A standing item on the ROC managers meeting agenda is made to accept, reject or request more time until the test is validated. A target of two weeks should be put on this step.
- 7. Once the ROC managers accept the test, the acceptance is noted at the Weekly Operations Meeting. Wait one week.
- 8. A broadcast is sent to all sites one week before the test is set as critical. This is also mentioned at the Weekly Operations Meeting.
- 9. A Broadcast is sent to all sites one day before the test is set as critical.
- 10. Test is set as critical.

### 6.2.2 Alarm generation in the Operations Dashboard

Alarm generation in the operations dashboard relies on checks of service notifications. The ROC\_OPERATORS profile (see the section on ROC\_OPERATORS profile above) defines the services that are checked. The decision to send out notifications is made in the service check and host check logic. Host and service notifications occur in the following instances:

- When a hard state change occurs
- When a host or service remains in a hard non-OK state and the time specified by the </notification\_interval/> option in the host or service definition has passed since the last notification was sent.

Information on state types and hard state changes can be found at the following URLs:

- http://nagios.sourceforge.net/docs/2\_0/statetypes.html
- http://nagios.sourceforge.net/docs/2\_0/notifications.html

The actual alarm creation is made through a filter, and if the service state changes, this is registered by the filter, stored in a data base and displayed on the dashboard.

## 6.3 GIIS Monitor

There is a new GIIS monitor at CERN: the GStat 2.0 Monitor. It is used to display information about grid services, the grid information system itself and related metrics.

GStat 2.0 does not rely on any submitted job, but rather scans all site GIISes/BDIIs to gather the information and perform sanity checks to point out any potential problems with the information system of individual sites. The tool covers the following areas:

- Static information published by a site: site name, version number, contact details, runtime environment (installed software);
- Logical information about site's components: CE, number of processors, number of jobs, all SEs, summary storage space, etc;
- Information integrity: bindings between the CE and close SEs.
- Geographical Information

- LDAP query view
- Lists the top level bdiis status
- Useful statistics

The GStat 2.0 provides many different views of the grid; The Site View is similar to that provided by the old GSTAT tool. Filters can be applied to select by Country, EGEE\_ROC, Grid, VO or WLCG\_Tier, to select the sites of interest. A table is shown with summary information about the sites, including its status and stats for numbers of CPUs, Storage, and jobs. Each site is provided with an individual page with more detailed information about the results of tests and sanity checks that have been performed, and also several plots of historical values of critical parameters. These historical plots are useful for spotting intermittent problems with site's information systems.

Further information about the architecture of the new system.

# **7 Reporting Problems**

For local operations, problems tickets should be raised via the local ROC helpdesk. Each ROC helpdesk may be connected to the central EGEE helpdesk, known as GGUS. In some ROCs this integration is handled automatically to allow for the seamless transfer of and action on tickets. The ROCs can then decide whether to escalate the ticket to GGUS for EGEE-wide resolution. Tickets can, however, be created directly in GGUS and allocated to a choice of support units, including C-COD for grid-wide operational issues. Please see the tutorial on the GGUS website for detailed instructions on using GGUS.

For user problems please raise a ticket with GGUS according to the instructions to be found at http://www.ggus.org

## 7.1 Problem detection

There are currently a variety of monitoring tools in EGEE/LCG which are used to detect problems with sites. They also provide useful information about sites. The regional operations dashboard https://operations-portal.in2p3.fr/ provides links and utilises combined views of the most common monitoring tools.

## 7.2 Workflow and escalation procedure

This section introduces a critical part of operations in terms of sites' problems detection, identification and solving. The escalation procedure is a procedure that operators must follow whenever any problem related to a site is detected. The main goal of the procedure is to track the problem follow-up process as a whole and keep the process consistent from the time of detection until the time when the ultimate solution is reached.

Moreover, the procedure is supposed to introduce a hierarchical structure and responsibility distribution in problem solving which should lead to significant improvement in the quality of the production grid service. Consequently, minimizing the delay between the steps of the procedure is of utmost importance. The regular procedure the operators follow can be considered in four phases.

- submitting problems into the problem tracking tool after they are detected using monitoring tools or by a task created by a regional operations team (ROC);
- updating the task when a site state changes which can be detected either by a comparison of the monitoring information with the current state of the task in the problem tracking tool, or by input from a ROC;
- closing tickets or escalating outdated tickets when deadlines are reached in the problem tracking tool;
- initiate last escalation step and/or communication with site administrators and ROCs.

Below are the detailed steps of the "escalation procedure" if no response is received for the notification of a problem or the problem has been unattended for:

Step [#]	Max. Duration [work days]	Escalation procedure
1		When an alarm appears on the ROD dashboard (>24 hours old): 1st mail to site admin and ROC
2	3	2nd mail to site admin and ROC; At the end of this period escalate to C-COD
3		Ticket escalated to C-COD, C-COD should in that week, act on the ticket by sending email to the ROC, ROD and site for immediate action and stating that representation at the next weekly operations meeting is requested. The discussion may also include site suspension.
4		

		(IF no response is obtained from either the site or ROC) C-COD will discuss the ticket at the FIRST Weekly Operations Meeting and involve the the Operation and Coordination Center (OCC) in the ticket
5	5	Discuss at the SECOND weekly operations meeting and assign the ticket to OCC
6		Where applicable, C-COD will request OCC to approve site suspension
7		C-COD will ask ROC to suspend the site

NB : Theoretically, the whole process could be covered in a 2-3 week period. Most often a site is either suspended on the spot for security reasons, or the problem is solved (or the site and the ROC reacts) well before operators need to escalate the issue to C-COD, who then determines whether to bring it to the Weekly Operations Meetings.

NB: After the first 3 days, at the 2nd escalation step, if the site has not solved its problem, ROD should suggest to the site to declare downtime until they solve the problem and the ROC should be notified. If they do not accept the downtime then C-COD will proceed with the regular escalation procedure at the agreed deadlines.

## 7.3 Suspending a site

For normal course of operations, a site status would be in "production". The ROC must suspend one of their sites in the following two cases:

- For problems that are not addressed by either the site or the ROC, the ROD team will apply an "escalation step procedure" described in section 5.2 and also section 8.3. The final step is suspension, and the site is taken out of the grid resources. For the site to be suspended, a given ROC would have to disregard answering several emails over a period of more than two weeks and not join the weekly operations meeting when asked to. As soon as the site's status is modified, the ROC would get another mail of notification.
- If a site is in downtime for more than one month. Before suspending the site the ROC should make an investigation to exclude the possibility that the site will be up again in a short period of time.

Before this happens, the ROC should make contact with the senior people in the federation, the site and the ROD. After, the ROC would have to re-certify the site before its status is put into "production" again.

It is well understood, that such "suspending a site" action may directly apply in emergency cases, e.g. security incidents. The escalation procedure is then by-passed totally by either the ROC or the ROD.

For the site to be suspended by C-COD, a given ROC would have to disregard answering several emails over a period of more than two weeks and not join the weekly operations meeting when asked to." Sounds a bit complicated, but don't find a best sentence

# 7.4 Sites that fail NAGIOS tests but still continue to be operational

Sites that fail NAGIOS tests where an alarm gets raised on the dashboard, but continue to be operational can be handled as follows: Operators need to set the ticket to "unsolvable" or set a long term deadline. A mail must be sent to the C-COD mailing list and a note put in the dashboard "site notepad". These sites should not be candidates for suspension.

## 7.5 Ticket handling during Weekends and Public Holidays

Due to the fact that weekends are not considered working days it is noted that ROD teams do not have any responsibilities during weekends and that RODs should ensure that tickets do not expire during weekends.

Currently there is no automatic mechanism for handling ticket expiration over public holiday periods. As this can vary from site to site, RODs are encouraged to get their sites to announce their public holidays so that ticket expiration can be set accordingly. (Correspondingly, ROD operators also have no duties when they are on public holidays.)

# 8 Security Items and Daily Operations

## 8.1 Security matters

Specific security issues or questions can be sent to

project-egee-security-support@cernSPAMNOTSPAMNOT.ch. This list contains a list of generic security contacts in each ROC, who will provide a reply. As it is rather difficult to ensure that GGUS tickets are readable only by the affected parties, one should refrain from using GGUS to track operational security issues.

All sites are bound by the JSPG policies listed at http://cern.ch/osct/policies.html. The policies cover many areas, including:

- "You shall immediately report any known or suspected security breach or misuse of the GRID or GRID credentials to the incident reporting locations specified by the VO and to the relevant credential issuing authorities. The Resource Providers, the VOs and the GRID operators are entitled to regulate and terminate access for administrative, operational and security purposes and you shall immediately comply with their instructions." (Grid Acceptable Use Policy, https://edms.cern.ch/document/428036)
- "Sites accept the duty to co-operate with Grid Security Operations and others in investigating and resolving security incidents, and to take responsible action as necessary to safeguard resources during an incident in accordance with the Grid Security Incident Response Policy." (Grid Security Policy, https://edms.cern.ch/document/428008/)
- "You shall comply with the Grid incident response procedures and respond promptly to requests from Grid Security Operations. You shall inform users in cases where their access rights have changed." (Virtual Organisation Operations Policy, https://edms.cern.ch/document/853968/)

## 8.2 OSCT organization and OSCT-Duty Contact Role

The Operational Security Coordination Team (OSCT, http://cern.ch/osct) provides an operational response to security threats against the EGEE infrastructure. It focuses mainly on computer security incident handling, by providing reporting channels, pan-regional coordination and support. It also deals with security monitoring on the Grid and provides best practices and advice to Grid system administrators.

The OSCT is led by the EGEE Security Officer and includes security contacts from each EGEE region. The OSCT Duty Contact (OSCT-DC) role is assigned on a weekly basis to one of the ROC Security Contact persons who provide support for daily security operations, according to a schedule defined by the OSCT.

The OSCT-DC must perform the following actions:

1. All reported incidents **should** be coordinated by the ROC Security Contact of the site reporting the incident (although this responsibility MAY be delegated in the region), but the OSCT-DC **must** ensure a timely coordination of the incident effectively happens, which includes assuming responsibility to coordinate himself/herself in an appropriate time frame if needed. More details about the role of the incident coordinator are available in the EGEE incident response procedure.

2. Ensure that appropriate action is taken in a timely manner (often by the affected ROC) to solve ticket assigned to the GGUS Security Management unit or to any matter being raised on the team's mailing list.

3. If necessary, attend the weekly OPS meeting and report or follow-up as appropriate within the OSCT

4. Send a report to the OSCT list, before lunch time on the Monday following the duty week, containing a summary of the issues of the week and those that are carried over from a previous week. This handover should also be carried over the weekly security operations phone meeting to enable the current and next OSCT-DC to discuss open issues.

5. Write new (or update) 'best-practice' items for the security RSS feed, based on advisories from GSVG (items should be sent to project-egee-security-officer@cernSPAMNOTSPAMNOT.ch to be published) If the issue posted by GSVG requires urgent action, raise it at the weekly OPS meeting and send a specific EGEE broadcast after consulting the rest of the OSCT

6. Monitor CA updates and monitor the update process as described in (http://goc.grid.sinica.edu.tw/gocwiki/Procedure\_for\_new\_CA\_release).

A backup OSCT-DC is defined in the same manner whose role is expected to cover situations such as prolonged unexpected network outage with the Lead site.

# 8.3 Security incidents handling and Interaction with OSCT-DC

OSCT- Duty Contacts (OSCT-DC) are members of the OSCT group that provide security coordination at a given week in tandem with the C-COD team. OSCT-DC has to be ready to respond to tickets created on security matters that given week.

If a security incident is suspected, the EGEE incident response procedure, which is available and maintained at http://cern.ch/osct/incident-reporting.html, being an excerpt from the full procedure https://edms.cern.ch/file/867454/2/EGEE\_Incident\_Response\_Procedure.pdf, must be used.

When a security related ticket is created by a site, VO or user, the OSCT-DC should then assign the ticket to the Security Support unit. They then follow-up the issue according to the OSCT regulations summarized above, taking into consideration the restricted information and contact details they have access to.

Security incidents are handled via the communication channels described on the procedure and must not be discussed via GGUS. If a security incident is initially discovered via GGUS, communication with the affected parties must be switched to the appropriate channels and the existing ticket should be closed and only used later as a reference.

However, 1st Line Support may have to act rapidly according to OSCT-DC instructions depending on the severity of the incident. 1st Line Support should also inform ROD about the situation. ROD can then act as a follow up contact.

## 8.4 Grid Security Vulnerability Handling

### Reason for a Vulnerability handling process

A lot of care is taken to ensure that Grid software and its deployment is secure, however from time to time, Grid Security Vulnerabilities are found. Grid Security Vulnerabilities are problems in the software or deployment which may be exploited in order to create an incident. Such problems need to be resolved in a timely manner in order to prevent incidents, and, while this is happening, it is important that they are not advertised to potential hackers. Hence, within EGEE a process has been established for reporting and handling vulnerabilities which is run by the EGEE Grid Security Vulnerability Group (GSVG).

## **Reporting a Vulnerability**

If a possible vulnerability is found an e-mail should be sent to

grid-vulnerability-report@cernSPAMNOT.ch

Alternatively, the issue may be entered as a "bug" in the GSVG savannah at https://savannah.cern.ch/projects/grid-vul/ for obvious reasons these bugs are set to "private" so only a limited number of people can browse them.

Note that if a vulnerability has been exploited, it is an incident and the incident handling procedure should be followed.

#### Handling of issues reported.

The GSVG will investigate the issue, and inform the reporter of the findings. If the issue is found to be valid then the Risk Assessment Team will place the issue in one of 4 Risk categories, each which has a corresponding Target Date.

#### **Risk Categories**

- \* Extremely Critical Target Date = 2 days
- \* High Target Date = 3 weeks
- \* Moderate Target Date = 3 months

\* Low - Target Date = 6 months

It is then the responsibility of the appropriate development team to produce a patch. This issue will be kept private until either a patch is issued to resolve it, or the target date is reached.

#### On resolution or reaching the Target Date

An advisory will be released. Hopefully, in most cases a patch will be issued in time for the Target Date. Advisories are now placed at http://www.gridpp.ac.uk/gsvg/advisories/. Advisories should reference the release, and release notes reference the advisory.

#### **Further Details**

Details of the issue handling process is available in the document entitled 'The Grid Security Vulnerability Group - Process and Risk Assessments for Specific Issues' at https://edms.cern.ch/document/977396/1.

More information on the Grid Security Vulnerability Group is available at http://www.gridpp.ac.uk/gsvg/.

#### Other notes

If the issue is found to be operational, rather than being due to a software bug, then an advisory will be set to the OSCT. Please refrain from discussing vulnerabilities on open mailing lists, logged mailing lists, or reporting them as 'bugs' in open bug reporting systems.

The GSVG was setup to primarily handle vulnerabilities and improve the security in EGEE gLite Middleware, and does not handle vulnerabilities in operating systems, or in non-Grid software. However, if such issues are reported to the GSVG they will attempt to pass information onto the relevant party.

# 9 References

[1] CIC Portal
[2] Operations Dashboard
[3] GOCDB
[4] Global Grid User Support
[5] GIIS Monitoring pages
[6] Availability report page
[7] SAM Nagios Documentation page
[8] Generic Nagios Documentation (CGI)
[9] GOC Wiki page
[10] EGEE broadcast tool
[11] Dashboard HOWTO Guides
[12] Best Practices
[13] Intervention Procedures& Notification Mechanism
[14] Site/ROC Association document

This topic: EGEE > OperationalProceduresforROCsAndSites Topic revision: r57 - 12-Apr-2010 - 03:20:35 - VeraHansper

**TWiki** Copyright &© by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Ask a support question or Send feedback