1 Current operational procedures documentation

1.1 Revision history

Comment	Date	Version	Author
Minor revisions to content – complete reformatting	10 October 2005	0.11	Alistair Mills
Initial draft	8 September 2005	0.01	Helene Cordier

Please verify that the document you are using is the current release. The document can be found on the CIC portal URL:
https://cic.in2p3.fr

This document does not describe future procedures. The purpose is to document the procedures used currently to operate the LCG2/EGEE production service and to describe the tools that are in use to implement the procedures.

Since there is not sufficient experience with the operation of large scale production grids we expect the document to be changed significantly and frequently.

2 Introduction

One of the main roles of the operations team is to take care of operational problems at Resource Centers (sites), Core Infrastructure Centers (CICs - Core Services) and grid wide problems.

The operations team is responsible for detecting problems, coordinating their diagnosis, and starting the follow-up procedure (the so-called escalation procedure) to track issues and document the solutions. This has to be done in coordination with the Regional Operations Centers to allow for a hierarchical approach.

Procedures have to be followed according to the formal descriptions to allow sharing the work. However, good judgment and common sense cannot be replaced by these procedures.

This document is loosely based on a description of the operational procedures that have been in use by the operations team at CERN during October 2004 and reflects the changes to the structure of EGEE and the outcome of the previous LCG/EGEE Operations workshops.

- Section 3 describes the CIC-on-duty operator's task;
- Section 4 describes most current administrative tasks such as handling scheduled downtimes and introducing new sites or contacting ROC managers are dealt with in conjunction with the GOC database;
- Section 5 describes the tools to detect and diagnose problems within the day-to day grid management such as problems at sites and services. Problems are dealt with in conjunction with the SFT2 tests;
- Section 6 describes the follow-up process, the ticket handling and escalation procedure are dealt with in conjunction with Global Grid User Support and described in section 6;
- Section 7 provides practical guidelines;
- Section 8 provides information on the current scripts;
- Section 9 provides a table of references.

3 Operators' tasks - CIC portal

This section describes the regular tasks an operator has to carry out and how the operation team should prepare for the task.

3.1 First time

The document of reference is to be found on the CIC portal {ref 11}. New CIC-on-duty team members have to do the following before taking their first shift:

- Have a valid certificate delivered by a recognized CA;
- Register into the dteam VO (also known as the LCG deployment team) at the following URL: https://lcg-register.cern.ch/cgi-bin/register/account.pl ;
- Refer to the current Operations Procedure document at the following URL: https://cic.in2p3.fr/index.php?id=cic&subid=cic_dash2 ;
- Refer to the history log from the previous CIC-on-duty shift available at the following URL: https://cic.in2p3.fr/index.php?id=cic&subid=cic_dash&view=v4 ;
- Register into GGUS as support staff and read the current manual at the following URL: https://gus.fzk.de/pages/home.php ;
- Request a GOC DB read-only access by completing the form at the following URL: https://goc.grid-support.ac.uk/gridsite/gocdb/request.php. The generic GOC DB mail list is the following: lcg-goc@listserv.rl.ac.uk ;
- Ask to be entitled with the CIC role in the GOC DB;
- Subscribe to the LCG-ROLLOUT mailing list (follow the Join link): http://listserv.cclrc.ac.uk/lists/LCG-ROLLOUT.html .

Here are a few recommendations. Before you do grid operations for the first time you can do a few things to prepare:

- Go to the CIC-COD dashboard manual and get used to the operations dashboard on the CIC portal {ref 12};
- Go to the FAQ pages and read the FAQs {ref 8}. Here the goal is not to know the material by heart, but to get an idea of the symptoms that you will be confronted with;
- Work a bit on the grid and read the User Guide. This will give you valuable background information and with some hands on experience you can tell user errors from GRID service errors more easily;
- Ask questions. Find people who have done it before or even better join an experienced team for a few days and look over their shoulders while they do their work. Not everything can be written up;
- Do not panic!

3.2 Activities

The operations team is expected to watch a few things continuously. Here are the most important ones:

1. The very first action on a shift is related to the weekly report available on the CIC page https://cic.in2p3.fr.
   There you find under the CIC-View the link to the CIC on duty log file;
2. Read the entries made on the previous week and add your own summary at the end of the day. You can find instructions on this here:
   https://cic.in2p3.fr/index.php?id=cic&subid=cic_dash&view=v4 ;
3. Operate the CIC-on-duty dashboard on the following link:
   https://cic.in2p3.fr/index.php?id=cic&subid=cic_dash.;
4. At the beginning of the shift, have a look at the tasks that need to be escalated. To get a standard problem tracking procedure throughout the different shifts, you should deal with troubles from the top down. The list is already sorted according to a priority criterion based on the problem severity and on the number of CPUs the affected site offers to the grid;
5. Check the Certificate Lifetime monitor and trigger an alarm if you spot sites with host certificates that will expire in less than a week;
6. Watch the LCGROLLOUT mailing list and react to operational problems mentioned there. The list can have a high traffic and not every thread is relevant to operations. However, the operations team is expected to skim through the list's messages;
7. After taking over from the previous operation team you should have a look at the new FAQs and troubleshooting guides that have been created.

By now you might have already realized that this is more than a single person can realistically get done in a day. From experience we know that between 1 and 2 persons can handle the workload. It is advisable not to split this between more than 3 individuals. The workload shows large variations.

4 Current administrative operations - GOC database

The GOC Database is obviously not a monitoring tool, but it 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 visible nodes (CE, SE, UI, RB, BDII etc.);
- phone numbers.

Site administrators have to enter all scheduled downtimes into the GOC-DB. The information provided by GOCDB is not used directly in problem detection and diagnosis but is an important information source during problem follow-up and Escalation procedure see {section 5.4}.

4.1 Introducing a new site

At first a new site has to contact his ROC for negotiation. The ROC can register this site in the GOC DB {ref 2} under the status candidate. Its global monitoring within GOC DB is then set to off.

When the site is established, the ROC switches the site status to uncertified.

When the site form within the GOCDB {ref 2} is completed and when the ROC's specific tests set or the local SFT instance is ok for a week - the site gets its status changed to certified by its own ROC.

At this stage, a site is registered as certified in the GOC DB, and it gets into regular EGEE production and no rollout announcement is necessary. The monitoring of its nodes must be switched on in the GOC DB. It is automatically registered in the top-level BDII and in the Freedom of Choice for Resources tool, to which VOs can refer in order to build their own specific BDII instance {ref 13}.

The last optional status is suspended. The site is then removed from the top-level BDII and its monitoring is turned off. It is then actually out of the grid. This status has to be cautiously used by ROCs, as this status is specific to sites for

which tickets have been raised and not solved on time, i.e. it is the last escalation step in the procedure of daily operations that the CIC-on-duty apples as described in {section 5.4}.

CIC-on-duty operators may also suspend a site without going through all the steps of the escalation procedure. For example, if a security hazard occurs, CIC-on-duty people must suspend a site on the spot in case of such an emergency.

Consequently, for the normal course of operations, when a ROC needs to take one of its sites out of production, it should go through the site downtime scheduling procedure described below. For example, a ROC may want to use the status suspended when one of its sites has to be permanently or indefinitely taken out of production. Note that when a site modifies any field of its site form within the GOC DB, its ROC shall get a notification mail.

4.2 Site downtime scheduling

EGEE resources need to be switched off properly in order not to disturb operations.

When a site needs to switch off for upgrade, then the site should specify the appropriate downtime beforehand for the nodes involved. Instructions in such a case are:

- Set downtime period in GOC;
- Announce the downtime through the EGEE broadcast tool {ref 14} to all VO managers (and only to them);
- Change the status of the site to closed in your information system, i.e. GlueCEStateStatus: Closed.

At the time of the writing of this document, downtimes are set at the site level. It is going to be specified at the node level in the very next future. E-mails notifications for ROCs will be enabled at the same time so that when a site form in the GOC GB undergoes any type of changes, ROCs will be notified.

4.3 Emergency contacts

Currently the GOC-DB handles sites that are not committed to EGEE, but are considered dependent from CERN-CIC as far as LCG/EGEE operations are concerned. CERN-CIC acts as a catch-all ROC for non-EGEE sites. Examples of this are generally outside of Europe such as India, Pakistan and USA.

Phone contacts for ROC managers are available within the GOC DB and also from the CIC web site.

5 Unscheduled troubles - SFT2 tests set

5.1 Problem detection

Currently there are a variety of monitoring tools in EGEE/LCG2 which can be used to detect problems with sites. They also provide useful information about sites.
The list of tools that is recommended to look at by the operations team is the following {table 0}:

- Sites Functional Tests (a.k.a. Piotr's Tool). Update period: every 3 hours (or more frequent on demand). http://lcgtestzonereports.web.cern.ch/lcgtestzonereports/cgibin/listreports.cgi ;
- GIIS Monitor. Update period: 5 minutes. http://goc.grid.sinica.edu.tw/gstat/ ;
- GOC Database. Update period: non regular. https://goc.grid-support.ac.uk/gridsite/db/ ;
- GOC Job Monitor. Update period: several times a day. http://goc.grid-support.ac.uk/gppmonWorld/gppmon_maps/lcg2.html ;
- GOC Certificate Lifetime. Update period: 1 day. http://goc.grid-support.ac.uk/gppmonWorld/cert_maps/CE.html , http://goc.grid-support.ac.uk/gppmonWorld/cert_maps/SE.html ;
- GOC Live Job Monitor. Update period: frequent. http://www.hep.ph.ic.ac.uk/e-science/projects/demo/index.html .

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://goc.grid-support.ac.uk/gridsite/gocmain/monitoring/.

A framework that includes for many sites information down to the fabric level is GridIce.

5.1.1 Sites Functional Tests

One of the most detailed information pages about possible problems is the Sites Functional Test page (known also as: SFT tests). It displays the results of a site testing script which is executed at every site on a worker node. The test script contains a number of grid functional tests as well as several basic tests that check the general worker node set-up. As a result, a large table called the result matrix is produced. Each row of the table represents a single site and each column gives the information about the result of a single test performed on a worker node at a given site. Some of the tests are considered to be critical for grid operations and some of them are not. Any failure observed in this table should be recognized, 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 list of the tests and their purpose are provided below:

- General tests – critical:
  - Middleware Version published in GIIS must be the current release. Sites are asked to update within 3 weeks of a given release;
  - Certificate Authority RPMs of all CA certificates on Worker Nodes are checked;
  - CSH test checks if CSH shell works correctly by running a small CSH script;
  - BrokerInfo checks edgbrokerinfo command - information about CE and close SEs.
- Replica management tests using edg-rm command – critical:
  - CopyAndReg. WN -> defaultSE copy and register a small text file from the WN to the default SE;
  - Replicate defaultSE to central SE replicate the file from last test to a chosen central SE;
  - Copy defaultSE -> WN get the file replicated in previous steps back to WN;
  - 3rd Party Rep. from central SE to default SE replicate a well known, small text file from a central SE to the default SE;
  - 3rd Party cp to WN get the file from last step to the WN;
  - Delete Replica from default SE delete the replica of file from previous steps from the default SE.
- Information system test (BDII used for replica management) critical:
  - GFAL infosys checks if LCG_GFAL_INFOSYS variable links to a valid BDII - contact and query.
- Replica management tests using lcgutils (the same sequence as for edgrm based tests) - critical:
  - lcgcr -> defaultSE;
  - lgrep defaultSE -> castorgrid;
  - lcgcp defaultSE -> WN;
  - 3rd party lcgrep castorgrid to defaultSE;
  - lcgcp castorgrid to WN;
  - lcgdel from defaultSE.
- Tests performed by SFT - critical for operations:
  - RGMA client software test using test scripts provided by RGMA RPMs;
  - Software Version on WN software version on WN from the version of GFAL RPM.

The test script is submitted to all sites automatically every 3 hours. However, during the interaction with site administrators, the need for test resubmission may arise. In that case, the operator should manually resubmit the test job to the selected sites (on-demand resubmission).

The main results table contains just summary information about the test results (mostly OK/FAILED values). However, by clicking on a particular result it is possible to see a fragment of a detailed report which corresponds to the test result. Additionally there are links to the GOC Database (Site Name column) and to the GIIS Monitor (Version column). The detailed reports should be studied to diagnose the problems.

In some cases, the same functionality fails on many sites at the same time. If this happens it is useful to look at the sites' history. If sites that have been stable for a significant amount of time (weeks) start to fail, it might be a problem with one of the central services used to conduct the tests. Since the regions have developed their own specific set of tests, it may be very helpful to contact the ROCs during the process of problem identification.

5.1.2 GIIS Monitor

The GIIS Monitor page provides the results of information system tests which are run every 5 minutes. The test does not rely on any submitted job, but rather scans all site GIISes/BDIIs to gather the information and perform so called

sanity checks to point out any potential problems with the information system of individual sites. The test 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.

The GIIS Monitor provides an overall view of the grid; there is one large table with summary information about the sites, and color codes to mark potential problems. Additionally for each site it provides a detailed page with textual information about the results of tests and sanity checks that have been performed, and also several plots of historical values of critical parameters like jobs, CPUs, storage space. These historical plots are useful for spotting intermittent problems with site's information systems.

5.2 Diagnosis

After a problem is detected using the tools from {section 5.1}. The operation team should not only spot the problem and report it, but they should first of all find the possible reason and understand the nature of the problem. This is especially critical as far as all central problems are concerned like problems with the RB, BDII, Replica Manager central services and RGMA registry. Additionally the operations' team should be able to give sufficient explanation to the ROCs and site administrators and also contribute to building a central knowledge database of typical problems to allow for quick and smooth problem detection and resolution in future.

The diagnosis process is described in the following subsections. However there are three types of information sources that are recommended during this stage:

1. The monitoring and information tools like the Sites functional tests results page;
2. The GIIS monitor and GOC Database are described in {section 5.1};
3. The knowledge database for CICs is currently implemented as a WIKI page and contains information about typical problems. The page is available at the following URL:
   http://goc.grid.sinica.edu.tw/gocwiki/SiteProblemsFollowUpFaq
   It contains links to more complete write-ups for troubleshooting certain services; While the aim is to have a central FAQ list at some time in the future it can be useful to have a look at FAQs maintained by other sites. Here is a selection:
   http://gridit.cnaf.infn.it/index.php?knowledgebase,
   http://www.gridpp.ac.uk/tbsupport/faq/index.html.
   Since these FAQs are all provided on an as is basis you should report errors or inconsistencies that you spot;
4. The mailing list for CIC members to discuss the problems that occur in grid operation and to evolve the knowledge database:
   project-egee-sa1-followup@cern.ch .

Spotting detailed error messages

There are several ways to find the root of a problem, depending of the problem's nature. The very first step after a problem is detected is to spot a particular error message that corresponds to the failure. In the simplest case the error message can be easily found out by clicking on the failure in the main results matrix. This gives a detailed report which usually contains the command line that was invoked by test script, the output and the error message itself.

For example if the 3rd party replication test were to fail, then when clicking on FAILED link, the following output may show up:

```
Checking 3rd party replication from lxn1183.cern.ch to the default SE
++ edgrm vo dteam lr lfn:TheUniversalFilewn040701a.cr.cnaf.infn.it.txt
++ grep lxn1183.cern.ch
+ ufilesfn=sfn://lxn1183.cern.ch/storage/dteam/generated/20041 [...] 66304a417a909baaca139049c2
+ edgreplicamanager v vo dteam replicateFile sfn://lxn1183 [...]
/dteam/generated/20041022/file233e2e66304a417a909baaca139049c2 edgreplicamanager starting..
Issuing command: replicateFile
Parameters: sfn://lxn1183.cern.ch/storage/dteam/generated/20041022/file23 [...]
909baaca139049c2
Call replica manager replicatefile function
Unknown Storage Resource: castorftp.cnaf.infn.it
+ result=255
+ set +x
```

In this case the error message is the following:
Unknown Storage Resource : castorftp.cnaf.infn.it

Unfortunately, sometimes the failure of a single test can result in many other failures (because of inter-test dependencies, for example in the Replica Manager tests). That is why the operations team should always be very careful and by understanding the nature of individual tests they should try to avoid mixing consequences and reasons for problems.

Finding the reason and the solution
Unfortunately there is no simple way to find the reason for a problem not to find the solution for the problem in a general case. This is in fact the most difficult stage in the process which requires a comprehensive knowledge about individual grid components and the interactions between them. To acquire the required knowledge, the development of the knowledge database has been started. The database, which currently is based on the WIKI mechanism, contains the most frequent problems that have been encountered in grid sites operation during past several months. However, the database is a subject to extension in the future as soon as new problems are detected and more partners are involved in operations.

The main use case for the knowledge database follows after a problem with a site is detected, and the relevant error message is spotted out. The operations team should try to find the description of the problem in the database. It can be done either by using the categorized structure of the database or by using the WIKI search engine. Once the description of the problem is found and the problem itself is clearly understood, the follow up process starts according to the Escalation procedure {see section 6}. Since the ROCs are actively spotting problems it is possible that problems are spotted by them at the same time that the operation team finds them. Until the regional problem tracking tools are integrated with the one used for the CIC operations, the operations team should try to contact the ROC in question and synchronize their activities. However, it is better to have a problem being in two tracking systems than in none. Unfortunately in many cases the operator can discover a new problem, which is not yet described in the database. In that case he or she has to relay on his own experience and knowledge but also should exchange his or her ideas with other people involved in follow up process.

Extending the knowledge database
To extend the current FAQ DB the members of the operations team have to register with the WIKI page. New contributions should be announced on the CIC mailing list and reviewed by partners experienced in the relevant field. Here again, it is better to add a FAQ that might be slightly questionable than to add nothing at all. At least the question is useful.

5.3 Problem tracking tool

In order to keep track of the follow-up process, all operators have to submit each detected problem to a problem tracking tool. The current problem tracking tool is Global Grid User Support based on Remedy run by FZK and has been in use for CIC operations since mid-April 2005 {ref 4}. GGUS ticketing system has two available interfaces:

- CIConduty dashboard on CIC portal should be used only by CIC people to report new problems and assign them to ROCs;
- Generic GGUS interface should be used by ROCs once tickets are assigned to them.

Problem categorization
In EGEE/LCG operations a single problem in each individual site and the associated follow up process is represented by a single ticket in GGUS. In order to organize and categorize the tickets the following structure has been put in place:

- Ticket; each ticket represents a problem in a single site;
- Category; identifies a site by the site name;
- Priority; used internally to mark sites with higher operational importance according to the number of provided resources (big sites - high priority);
- Item group; represents a problem type, eg. LCG Version.

The task fields which are introduced above describe the individual problems in terms of location (site), type, and importance.

Task field usage

Apart from the information provided by the fields which were introduced above, there are a number of fields which describe the details of the problem and current status in terms of Escalation procedure (see {section 6}). These fields should be utilized as follows:

- Should be Finished on - the deadline for the current escalation step;
- Assigned to - ROC which currently takes care of followup for the particular problem;
- Last action taken - last action which was taken according to Escalation procedure;
- Person contacted - the name, email address and possibly phone number of the person who was contacted in the last action;
- Response - a summary of communication with the person responsible for the site in the last action;
- Summary - a summary of the problem, it is highly recommended to put affected host's name as first part of the summary;
- Original submission - the original error message plus any comments that may be useful for problem identification and solving.

6 Escalation procedure and tickets handling - GGUS

This section introduces the most critical part of operations in terms of sites' problems detection, identification and solving. The Escalation procedure is a procedure to be followed by operators whenever any problem related to a site, CIC or region is detected. The main goal of the procedure is to keep track of the whole problem follow-up process and keep the process consistent from the moment of detection until 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 of the quality of the production grid service. For this, the aspect of minimizing the delay is significant.

The procedure can be considered in five phases. The following figure (*Figure 1:*) shows the decision diagram for phases I to V by the source of information. The expression phase means here not a step in a linear sequence, but rather a part of the procedure which is triggered by an observation or information coming from various sources:

- phase I - problems are detected using monitoring tools or by a task created by a regional operations team (ROC);
- phase II - site changes can be detected either by a comparison of monitoring information with the current state of the task in the problem tracking tool, or by input from a ROC;
- phase III, IV - deadlines can be reached in problem tracking tools;
- phase V - communication is initiated with site administrators and ROCs.

These phases are discussed in more detail in the following sections of this document.

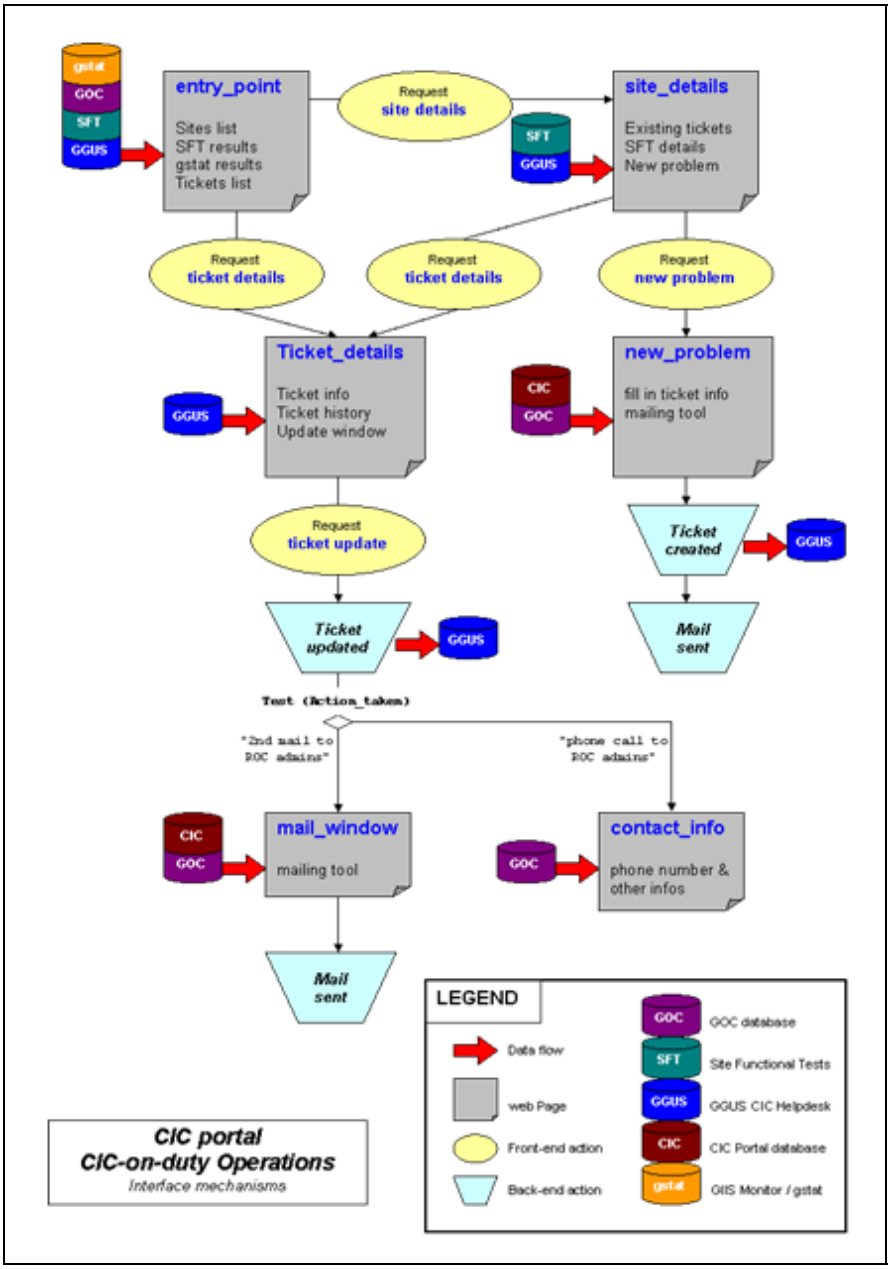

Figure 1: Decision diagram with phases

6.1 Phase I - Submitting problems into GGUS

Table 1: Operational Priorities

CPUs	priority
>= 100	high
>= 20	normal
< 20	low

1. Assign operational priority to sites using history graph and averaged values from GIIS monitor according to the values in the table (*Table 1: Operational Priorities*). The currently used priority is based on the number of CPUs that a site provides. It is understood that this is not the only critical aspect. This classification is currently used, as there is a strong correlation between the computing and the storage capacity that a site provides. The operations team has some freedom in increasing the priority of a site if it receives input from the VOs that this site is especially critical for ongoing production work;

2. Go to the GGUS interface on CIC Portal, select the site from the list and click View button to see the details of the site;
3. For each problem check if there is an open task for the site for this particular problem in GGUS. If yes proceed to Phase II, - Updating task when a site state changes;
4. Select problem type from the list and corresponding sub problem. This results in an integrated form which allows submitting a ticket and sending notification email in one step;
5. Enter the details of new problem and edit the body of notification email which will be sent to the site administrator and ROC:
    1. Priority: high, normal, low depending on importance;
    2. Should be finished on - transition to next step (depending on importance) according to the following table (*Table 2: Escalation deadlines*);
    3. Summary: very short description of the problem, if possible a listing of the VOs affected i.e.:

```
job submission failed;
replication failed;
lxn1181.cern.ch is still LCG2_0_0.
Original Submission: details of failure with error message if available, example:
copyAndRegisterFile and replicateFile failed with the following message: java:
globus_fifo.c:159: globus_fifo_peek: Assertion [...]

Body: edit the body of notification email and is possible give more details and suggested
reason/ solution (include an indication about the expected time needed to address the
problem).

for the rest use default values (or leave empty).
```

    4. Click Submit&Send button. Review the details and if it is correct confirm using Create&Send button.

Table 2: Escalation deadlines

priority	deadline
high	set to 1 day later
normal	3 days later
low	3 days

6.2 Phase II - Updating task when a sites state changes

1. For each site with an open task, for which the state changed in a new report from the failure to OK, update associated ticket in the problem tracking tool. However when a new failure for the site is detected the existing ticket should not be modified (only the deadline should be extended) but rather a new one should be submitted (Phase I). Problems that have been assigned to a ROC can be closed by the ROC. A problem that has been assigned to a ROC is handled by the ROC, the operations team only acts when the time limits are exceeded.
2. If the problem was trivial and required just a simple action (for example GIIS service restart) without any configuration changes close the ticket (see Phase III).
3. If the problem required longer treatment and involved configuration changes, start the quarantine period by changing the following attributes:
    - Change expiration date to: set to: now + 3 days (quarantine period, to see if site is stable);
    - Escalate: set to: Quarantine;
    - Add a comment to diary: feel free to give any details about the fixed problem;
    - Please make sure to select the check boxes for any fields you set. The quarantine has been introduced based on experience gained during the DC04 and is subject to future refinement.

6.3 Phase III/ IV - Closing tickets or escalating outdated tickets

Table 3: Escalation steps

Action taken	low	normal	high

1st mail to site admin and ROC	3	3	1
2nd mail to ROC	3	3	1
phone call to ROC	3	3	1
Site suspension	—	—	—

1. Use ticket browser in GGUS interface on CIC portal to find all expired tickets by selecting Ticket Status: Only expired tickets radio button, and clicking View button.
2. For all tickets, click on the ticket ID (first column) to go to GGUS Ticket Update page.
3. If Last Action taken for the ticket is Quarantine or Site OK, check if problem is really not there according to the monitoring tools and close ticket by updating the following attributes:
   - Escalate: set to Site OK;
   - Close ticket: set to Yes;
4. For all expired deadline problems for which no progress in communication with responsible people has been made (either at the site or at the ROC level) the operator should move to the next escalation step:
   - Change expiration date to: set to now + deadline depending on new action and site importance, according to values in the table (*Table 3: Escalation steps*);
   - Escalate: set change to next value on the list of actions (site suspension is the last action. Steps' names will be changed according to EGEE project structure);
   - Click Update button which allows you to perform selected action (mail tool for 2nd email and information page with phone numbers for phone call).

6.4 Phase V - Communication with sites, CICs and ROCs

Communication with sites, CICs and ROCs has to be described in a somewhat formal way because we experienced that the person in charge did not respond to the notification mails created by the problem tracking tool (GGUS). This was not a problem with related to the tool, but to the state of the project. We expect the need for this complex procedure to become obsolete in the future. This escalated communication with partners that are not living up to their responsibilities will be removed from the document as soon as we experience reliable reactions.

If an email was received as a response to Phase IV modify the task:

- Change expiration date to: set according to agreed final date, when site should become operational;
- Person contacted: update if got a response to general email, or responsible person changed;
- Response: short description of received response.

If during Phase IV a telephone call was SUCCESSFULLY made, modify the task:

- Change expiration date to: set according to agreed final date, when site should become operational again;
- Response: short description of received response.

For all problems that should have been finished at the current date, but still there, and for which there is no progress in communicating with the responsible people (either at the site, ROC level, move to the next escalation step. The escalation steps are described in the following figure (*Figure 2: Escalation steps*).

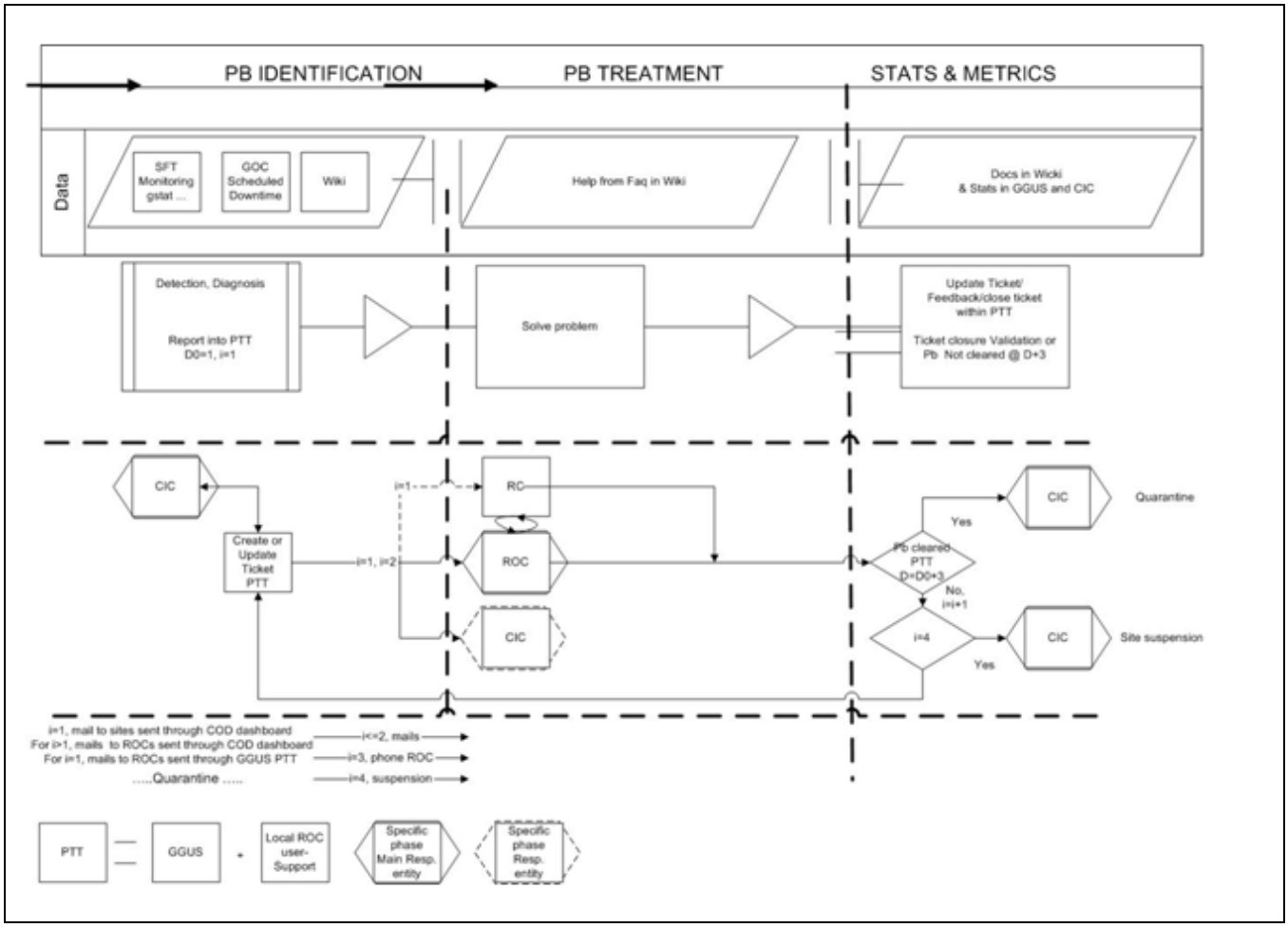

Figure 2: Escalation steps

The last step is suspending a site from Grid which implies no monitoring of the site. Moreover, no more jobs will be submitted through the grid. The site is set to suspended in GOC DB by COD after the current operator brings the problem up to the attention of the next weekly operations meeting. The Grid Deployment Area phone meetings are held on Mondays at 14:00 CET. SA1 managers are informed of the status of the site and a given ROC must ask the current CIC-on-duty person for the site to be reintegrated into the status production.

The current question underway is whether a given site should undergo certification tests after a scheduled downtime (for upgrade and not otherwise) and after a suspension time (in every case). It is well understood, that such an action may directly apply in emergency cases, e.g. security incidents. The escalation procedure is then by-passed directly to the last step of site suspension by either ROCs or CODs.

7 Practical notes and guidelines for CIC on duty operators

7.1 Introduction

CIC administrators need to investigate problems encountered, so they could give instructions, suggestions to the sites, about what and how to fix. The purpose of these guidelines is to briefly show, how this is done in practice, with our current tools.

7.2 Preparation

Before starting the actual task, the new CIC administrator has to get familiar with Sites Functional Tests and GIIS Monitoring pages. They should try to understand errors that appear. Finding ones that they could not follow up if they were on duty, they might ask the CIConduty mailing list. The GOC Wiki pages are also useful. They should also read some emails which have been sent on the followup-list.

7.3 The ten rules for identifying a problem

1. Error appears on Sites Functional Tests pages {ref 16}.
2. Check error message (starting from Replica Manager tests as those have the highest priority). If that concerns a brand new site, then it should get a mail describing the testing tools. A good example has already been sent to the followup list.
3. Depending on error message, check GIIS Monitoring page for the site. This is linked directly to the Version column:
   Have a look at Sanity Check and the first 3 big graphs (CPUs, jobs, storage).
   - Job Submission failed:
     Check if GIIS was up, and published CE information when jobs were submitted;
   - Replica Manager tests failed, problem with the site's SE ?
     Check if GIIS was up, and published information about the SE and CE:
     Note that GlueCESEBindCEAccesspoint, the base directory for SE access (usually /storage) is published by the CE! Check the LCG_GFAL_INFOSYS test;
   - Site GIIS down:
     When all three graphs show a valley, at the same time, probably GIIS was down at that point;
   - GIIS Monitoring pages could also show that GIIS or just CE/ SE info was unstable. Reason could be, for instance a busy CE with long response time (see Troubleshooting Guide/Information System/Value for GlueCEUniqueID not published);
   - GlueSAStateAvailableSpace not published. Please see below about this error.
4. Might need to check other tests and/or Test History (linked to Test date field) or if the problem needs further investigation:
   SE not found, with the following message for instance: Could not find a Storage Element to copy file to.
   Suppose that GIIS Monitoring did not show any problem. Check what BDII the site is using. This is displayed on Sites Functional Tests page. If the site has its own top level BDII, Troubleshooting Guide/Site BDII/No data published by top level BDII should be referred to.
   Typical problem was that SE dynamic info about available storage space (GlueSAStateAvailableSpace) was not published, due to an error in /opt/lcg/var/lcginfogeneric.conf.
   (Recipe on this is given at Troubleshooting Guide/Information System/Value for GlueSAStateAvailableSpace is not published). This could be verified by copying and checking the file:
   ```
   globusurlcopy gsiftp://<SEname>/opt/lcg/var/lcginfogeneric.conf \
   file:/tmp/infogeneric<site>.conf
   ```
5. Check if the problem at the site is reported in GGUS, and if it is, then what's the next step in the Escalation Procedure (described in Current Operational Procedures Draft). Suppose that it is Sending an email.
6. See if there was any relevant mail on followup list about the error detected (Section {B5}. Describes Sites Followup Archive access.), as you may have the right instruction for the site.
7. If you are doubtful about a problem, ask others about it (on CIConduty mailing list or directly to someone expert).
8. Error detected
   Check if there is an entry about it on the GOC Wiki page. If not, you could set it up.
9. Sending the mail
   If you are unsure, please double check and send a mail first to the CIConduty mailing list! We must not confuse the sites with inappropriate suggestions!
10. Add new entry to GGUS, or update the existing one.

More hints:

- Manual job submission;
- edg-grid-ftpls, ldapsearch and similar checks;
- check the accessibility of a port.

7.4 Mailing lists and follow-up archive

- followup list: <project-egee-sa1-followup@cern.ch> mails sent to the sites
- CIC on duty mailing list: < project-eu-egee-sa1-ciconduty@cern.ch > internal communication between CIC administrators.
- Sites Followup Archive
  http://simba2.cern.ch/archive/project-egee-sa1-followup
  login: <your full email>, password: <password>
  Getting password:
  http://weba5.cern.ch/externalsweb/loginexternal.aspx?param=newpasswindow
  Type in your email address registered for the mailing list. Press Send New Password button. In a few minutes you will receive the password.

7.5 Email contents

The email is addressed to the corresponding ROC, together with the site. ROC mail addresses are listed on CIC Portal (https://cic.in2p3.fr) under ROC views.

```
Forming address: To: <sitecontactemail>, <roccontactemail>
Cc: SA1 Followup Mailing List projectegeesa1followup@cern.ch
Subject: <problem> at <sitename>
ReplyTo: SA1 Followup Mailing List <projectegeesa1followup@cern.ch>
```

The mail has to contain:

- Name and URL of the page, where the error was detected;
- Category of the error and the error message;
- Recipe from the GOC Wiki page (Preferably, mainly for new site administrators provide access path from the main page, as in the example, instead of direct URL, so they would get used to the page, and see other entries also.).

Of course, obvious errors (like an MDS restart) do not require all these steps as a one line mail will do.

```
To: <sitecontactemail>
Cc: SA1 Followup Mailing List <projectegeesa1followup@cern.ch>
Subject: GIIS at <sitename>
ReplyTo: SA1 Followup Mailing List <projectegeesa1followup@cern.ch>

Dear Site Administrators,

According to the site's monitoring pages
(http://goc.grid.sinica.edu.tw/gstat/<site_page> )
site does not publish dynamic value about storage space (variable GlueSAStateAvailableSpace) for
VO biomed on SE <SEname>.  Having a look at the GRIS configuration file
(/opt/lcg/var/lcginfogeneric.conf) on the SE, there are 2 problems with line

dynamic_script= [...]

Un substituted LCFG variable appears there.
VO biomed does not appear there

You will find a fix about both of these errors at GOC wiki page
(http://goc.grid.sinica.edu.tw/gocwiki/FrontPage) Troubleshooting Guide, (Information System
section),

"Value for "GlueSAStateAvailableSpace" is not published" entry, together with the description of the
syntax for this line.

Please follow instructions described there.
Do not hesitate to contact us, if you have questions, or need help.

Thanks,
Judit Novak
CERN, ITGDGIS
```

7.6 Modifying the GOC Wiki page

Not yet recorded errors, updates on already existing entries should be added to the wiki page. Before editing, please refer to How to contribute to this page on the top of the Troubleshooting Guide, where you will also find an example entry in order to keep consistent structure and layout. Or, if you prefer not to bother with formatting the text, please send what should be added to <Judit.Novak@cern.ch> in an email. If you choose to modify the page yourself, please send a mail about it to ciconduty mailing list, so CIC administrators would know about new entries and updates.

8 Test Scripts

8.1 Test script installation

The whole package which contains the submission environment, the test script itself and a set of CGI scripts is available in LCG2 CVS repository {ref 17}. To install Site Functional Tests (SFT) you need to do the following get the newest version of the scripts from the CVS repository. This can be done either by using command line for LCG CVS Repository:

```
export CVSROOT=:ext:<YOUR AFS ID HERE>@isscvs.cern.ch:/local/reps/lcgware
export CVS_RSH=ssh
cvs cod sft lcg2/sft
```

Or by using a CVS web interface:
http://isscvs.cern.ch:8180/cgi-bin/cvsweb.cgi/sft2/?cvsroot=lcgware#dirlist

Then, please have a look at the files in sft/ doc directory and follow the installation and configuration instructions.

8.2 CGI script installation

This step is necessary only if you are not using CERN to publish test results, but your own storage and web server for reporting page. To be able to generate web reports for tests results you have to install and configure a set of CGI scripts which are provided in the main package. The instructions are provided in sft/doc/CGI QuickStart.txt file {ref 17}.

8.3 On-demand test resubmission and report generation

After the test environment is successfully installed you can submit a test job to selected sites, retrieve and publish the results and generate the reports. A quick guide for on demand job submission is the following:

- submission of jobs to all sites or just to selected ones:
  `sft/sftests submit [<filter>]`
  where filter can refer to either CE hostname, site name (according to GOC DB) or region name (also according to GOC DB), e.g.:
  `sft/bin/sftests submit France cern.ch`
- checking the status of submitted jobs:
  `sft/sftests status`
- getting the jobs' output:
  `sft/sftests getoutputs`
- publishing the results on the web:
  `sft/sftests publish`
- cancellation of one or more jobs:
  `sft/sftests cancel [<filter>]`
  where filter can refer only to CE host name at this point
- if everything fails and you want to clean it up and start from scratch:
  `sft/sftests clean`

8.4 CRON job installation

To enable automatic test jobs submission and report generation while you are a CIC-on-duty on call, you can install cron jobs that submit test jobs at definite times and periodically regenerate the report. It is recommended to submit test jobs once per day in the morning (for example at 6:05am) and regenerate the report every 30 minutes. To avoid overlapping, it is also recommended to submit new jobs about 5 minutes after periodical regeneration of the report (that is why 6:05am is suggested rather than 6:00am).

Warning : in order to submit cron jobs using an AFS account, you need an AFS enabled CRON to allow access to your home directory !

Proxy certificate renewal

In order to submit grid jobs and retrieve outputs, each cron job must have access to valid user's proxy certificate. This is the most critical and dangerous part of setting up the cron job. Currently the script is regenerating the proxy each time it is needed, based on the passphrase given in text file stored on user's account. Although this is really dangerous to store passphrase in a file, using myproxy service is not yet fully supported.

To allow the test script to regenerate the proxy you just need to put your Globus certificate's passphrase in a file using the following command:

```
echo '<passphrase>' > $HOME/sft/conf/grid-cert-passphrase
```

Afterwards do not forget to make this file private by setting proper access rights:

```
chmod 600 $HOME/sft/conf/grid-cert-passphrase
```

Do not forget AFS rights (use fs setacl).

Setting up the cron Jobs

If you are not using AFS to publish the results in a central place you can easily set up the cron jobs by adding the following lines to your crontab (crontab -e command):

```
0,30 * * * * $HOME/sft/sftests crongen    > /dev/null 2> /dev/null
5    6 * * * $HOME/sft/sftests cronsubmit > /dev/null 2> /dev/null
```

This will setup automatic job submission to all sites every day at 6:05am and automatic output retrieval and report regeneration every 30 minutes. However, usually the results are published using AFS accounts. In that case instead of using classic Cron deamon, AFS/Kerberos enabled version of CRON must be used. This is because each cron job must be executed with a valid Kerberos ticket in place. Therefore you have to use a UI machine which has acrontab installed and properly configured. Certainly installation and configuration of acrontab is outside the scope of this document and is usually site specific.

To setup the cron jobs with a crontab just add the following lines using acrontab e command:

```
0,30 * * * * <UIhostname> $HOME/sft/bin/sftests crongen    > /dev/null 2>/dev/null
5    6 * * * <UIhostname> $HOME/sft/bin/sftests cronsubmit > /dev/null 2>/dev/null
```

9 References

[1] CIC Portal:
https://cic.in2p3.fr

[2] GOC DB:
http://goc.grid-support.ac.uk/gridsite/gocdb/

[3] Sites Functional Tests - SFT2:
https://lcg-sft.cern.ch:9443/sft/lastreport.cgi

[4] Global Grid User Support:
https://gus.fzk.de/pages/home.php

[5] GIIS Monitoring pages:
http://goc.grid.sinica.edu.tw/gstat

[6] Metrics report page:
https://lcg-sft.cern.ch:9443/sft/metrics.html

[7] SFT2 Documentation page:
http://goc.grid.sinica.edu.tw/gocwiki/Site_Functional_Tests

[8] GOC Wiki page:
http://goc.grid.sinica.edu.tw/gocwiki/

[9] Sites Followup Archive:
http://simba2.cern.ch/archive/project-egee-sa1-followup

[10] Current Operational Procedures Document:
https://cic.in2p3.fr/include/pdf/opMan.pdf

[11] COD team management:
http://egee-docs.web.cern.ch/egee-docs/list.php?dir=.\cic_managers\&/CIC-COD_management-v1_1.pdf

[12] COD dashboard how-to:
http://egee-docs.web.cern.ch/egee-docs/list.php?dir=.\cic_managers\&/CIC-COD_dashboard-manaual-v1_0.pdf

[13] Freedom of Choice for Resources:
https://goc.grid-support.ac.uk/gridsite/bdii/site-apps/FCR-cgi/fcr.cgi

[14] EGEE broadcast tool:
https://cic.in2p3.fr/index.php?id=rc&subid=rc_publish

[15] SFT suite list:
http://lxb2001.cern.ch:8083/sft/sftestcases.html

[16] SFT web interface:
http://lxb2001.cern.ch:8083/sft/

[17] SFT Installation guide, quick start, adding new testshttp://isscvs.cern.ch:8180/cgi-bin/cvsweb.cgi/sft2/?cvsroot=lcgware

10 Comments

[FS1] Maybe this has to be agreed/ discussed? We are not meant to solve all problems, but to guide people to solutions.

[FS2] This diagram is not clear - I asked Helen to design a new one.

[FS3] I never do it - it is boring and useless AND should be automatic when the ticket is closed (asked GILLES to implement this).

[FS4] Is this really necessary? I think it is not the CIC duty to decide who is working but only to decide who is NOT working. I think this should be the SA1 leaders who decide.