**New entries to the TCB Catalogue of User Requirements (CoR)**

**Prepared for the 11th TCB meeting**

**Version 1.0, 05/04/2012**

This document was prepared by the User Community Support Team (UCST; ucst@egi.eu) and the EGI Operations team (operations@egi.eu) for the EGI Technical Coordination Board (TCB) from the requirements that recently reached UCST and Operations and that these teams are unable to address without the assistance of the TCB.

## User requirements (3)

|  |  |
| --- | --- |
| Unique ID | #3404 |
| Title/Topic | CE/batch information discovery about fair shares policy for the VO/user |
| Reference | <https://rt.egi.eu/guest/Ticket/Display.html?id=3404> |
| Priority | 5 |
| Submitter(s) | VO COMPCHEM |
| Status at TCB | Submitted |
| Description | VO COMPCHEM would like to query information about the capabilities of endpoint per USER, and not per VO or per QUEUE.e.g. information is needed by VO COMPCHEM per User- Number of running Jobs per user (current state / limit)- Number of waiting Jobs per user (current state / limit)- Number of Processors/Cores per user (current state / limit)- Number of Computing Nodes per user (current state / limit)According to Glue Schema 1.3 specification document such information is not available.But the information which is available is misleading and brings even more confusions.The descriptions about the attributes are different between what is in Glue Schema at running Top/Site BDIIs and what is in Glue Schema 1.3 spec. document (<http://glueschema.forge.cnaf.infn.it/Spec/V13>)Descriptions from Top/Site BDIIs Glue Schema:TotalCPUs - GlueCEInfoTotalCPUs (GlueCE) - Number of CPUs available to the queue. NB: this number should not be used to total available resources as more then one queue may be pointed to the same physical resourcesPhysicalCPU - GlueSubClusterPhysicalCPUs (GlueCE) - The total number of real CPUs in the subclusterLogicalCPU - GlueSubClusterLogicalCPUs (GlueCE) - The effective number of CPUs in the subcluster, including the e®ect of hyperthreadingSMPSize - GlueHostArchitectureSMPSize (GlueCE) - number of CPUs in an SMP nodeMaxRunningJobs - GlueCEPolicyMaxRunningJobs (GlueCE) - The maximum number of jobs allowed to be runningMaxCPUTime - GlueCEPolicyMaxCPUTime - The maximum CPU time allowed for jobs submitted to the CE in minsMaxTotalJobs - GlueCEPolicyMaxTotalJobs (GlueCE) - The maximum allowed number of jobs in the queueAssignedJobSlots - GlueCEPolicyAssignedJobSlots (GlueCE) - number of slots for jobs to be in running state (it represents the maximum number of single-processor jobs that can be running at a given timePriority - GlueCEPolicyPriority (GlueCE) - Info about the Queue PriorityMaxWCTime - GlueCEPolicyMaxWallClockTime (GlueCE) - The maximum wall clock time allowed for jobs submitted to the CE in minsTotalJobs - GlueCEStateTotalJobs (GlueCE) - Number of jobs in the CEVOTotalJobs - GlueCEStateTotalJobs (GlueVOView) - Number of jobs in the CEVOFreeJobSlots - GlueCEStateFreeJobSlots (GlueVOView) - Number of free job slots, i.e. number of single-processor jobs which could be started if no other job are submitted and no jobs finish in the interimIt means that user already must keep in mind Glue Schema 1.3 specifications document in order to understand what is meant by Glue Schema in Top/Site BDII:GlueCEStateFreeJobSlots (GlueVOView) - Number of free job slots, i.e. number of single-processor jobs which could be started if no other job are submitted and no jobs finish in the interimWhile in Glue Schema 1.3 spec. document it is said:Number of free job slots (sometimes called logical CPUs), i.e., number of single-processor jobs which could be started if no other jobs are submitted and no jobsfinish in the interim. This could be the size of the queue less the used resources,but a policy could influence this number (N)Investigation/proof about having limits per user pool account which cause the information present in information system not to be valid at supported VO COMPCHEM clusters.1) the list of all clusters supporting VO COMPCHEM$ ./lcg-info --list-ce --bdii ldap://[lcg-bdii.cern.ch:2170](http://lcg-bdii.cern.ch:2170) --vo compchem --attrs 'Cluster' |egrep Cluster|awk '{print $3}'|sort|uniq > Clusters\_VO\_COMPCHEM$ wc -l Clusters\_VO\_COMPCHEM$ 50 Clusters\_VO\_COMPCHEM2) the list of jobs limits defined by MAXJOBPERUSER, MAXPROCPERUSER (per pool account) in the batch system (targeting PBS)$ for i in `cat Clusters\_VO\_COMPCHEM`; do uberftp $i 'cat /var/spool/maui/maui.cfg' > $i.maui.cfg; done$ grep -iR compchem \*ce01.ariagni.hellasgrid.gr.maui.cfg:CLASSCFG[compchem]      MAXJOBPERUSER=60,100 MAXPROCPERUSER=60,100 FSTARGET=5 QDEF=egeece01.marie.hellasgrid.gr.maui.cfg:CLASSCFG[compchem]      MAXJOBPERUSER=70,100 MAXPROCPERUSER=70,100 FSTARGET=5 QDEF=egeece02.marie.hellasgrid.gr.maui.cfg:CLASSCFG[compchem]      MAXJOBPERUSER=12,20 MAXPROCPERUSER=12,20 FSTARGET=5 QDEF=egeece0.m3pec.u-bordeaux1.fr.maui.cfg:GROUPCFG[compchem] MAXPROC=365 PLIST=pubce.reef.man.poznan.pl.maui.cfg:GROUPCFG[compchem]                      MAXJOB=50ce.reef.man.poznan.pl.maui.cfg:# GROUPCFG[compchem]             PRIORITY=10cream01.athena.hellasgrid.gr.maui.cfg:CLASSCFG[compchem]      MAXJOBPERUSER=114,190 MAXPROCPERUSER=114,190 FSTARGET=10 QDEF=egeecream01.grid.uoi.gr.maui.cfg:GROUPCFG[compchem]     MAXPROC=64,118cream01.grid.uoi.gr.maui.cfg:CLASSCFG[compchem]     FSTARGET=10 QDEF=total QLIST=total,egee  MAXPROCPERUSER=64,118 MAXJOBPERUSER=64,118cream01.kallisto.hellasgrid.gr.maui.cfg:CLASSCFG[compchem]      MAXJOBPERUSER=60,90 MAXPROCPERUSER=60,90 FSTARGET=10 QDEF=egeecream02.athena.hellasgrid.gr.maui.cfg:CLASSCFG[compchem]      MAXJOBPERUSER=200,333 MAXPROCPERUSER=200,333 FSTARGET=10 QDEF=egeegridgate.cs.tcd.ie.maui.cfg:GROUPCFG[compchem]      MAXPROC=400,500gridgate.cs.tcd.ie.maui.cfg:SRCFG[r410] GROUPLIST=compchem,biomed,cosmo,solovomario.univ-lille1.fr.maui.cfg:GROUPCFG[compchem] FSTARGET=5+ MAXJOB=300 MAXIJOB=15 PLIST=WN3) User from VO COMPCHEM finds that he has VOFreeJobSlots=123, but after he sends 123 Jobs, only 60 are running, because MAXJOBPERUSER=60 is defined in a local batch system.$ ./lcg-info --list-ce --bdii ldap://[lcg-bdii.cern.ch:2170](http://lcg-bdii.cern.ch:2170) --vo 'compchem' --query 'Cluster=[ce01.ariagni.hellasgrid.gr](http://ce01.ariagni.hellasgrid.gr)' --attrs 'TotalCPUs,PhysicalCPU,LogicalCPU,SMPSize,VMemory,Memory,MaxRunningJobs,MaxTotalJobs,AssignedJobSlots,TotalJobs,MaxWCTime,MaxCPUTime,VOTotalJobs,VOFreeJobSlots,VOEstRespTime'- CE: [ce01.ariagni.hellasgrid.gr:2119/jobmanager-pbs-compchem](http://ce01.ariagni.hellasgrid.gr:2119/jobmanager-pbs-compchem)  - TotalCPUs           212  - PhysicalCPU         120  - LogicalCPU          120  - SMPSize             2  - VMemory             4096  - Memory              2048  - MaxRunningJobs      119  - MaxTotalJobs        999999999  - AssignedJobSlots    212  - TotalJobs           0  - MaxWCTime           4320  - MaxCPUTime          2880  - VOTotalJobs         0  - VOFreeJobSlots      123  - VOEstRespTime       0Documentation used:1) GlueSchema 1.3 REF: [www.italiangrid.it/sites/default/files/**GLUESchema**1-3.pdf](file:///C%3A%5CUsers%5Cgergely.sipos%5CAppData%5CLocal%5CTemp%5Cwww.italiangrid.it%5Csites%5Cdefault%5Cfiles%5CGLUESchema1-3.pdf)2) MAUI scheduler REF: <http://www.adaptivecomputing.com/resources/docs/maui/12.3nodepolicies.php> |
| EGI Helpdesk | Ticket at GGUS, opinions from IS experts: <https://ggus.eu/tech/ticket_show.php?ticket=79309> |
| Goal | Ability to more efficiently use resources |
| Impact | Users get wrong information about how many jobs they can run from that time they see the free slots, after they try to do submission of let’s say 10 jobs, only 5 of them are running, because local batch system has a limit for 5 jobs per pool account. |
| Affected services | Top-BDII, Site-BDII, local batch systems. |

|  |  |
| --- | --- |
| Unique ID | #3406 |
| Title/Topic | Implementation of realistic cyclic and dynamic Workflows |
| Reference | <https://rt.egi.eu/guest/Ticket/Display.html?id=3406> |
| Priority | 5 |
| Submitter(s) | VO COMPCHEM |
| Status at TCB | Submitted |
| Description | According to experts in GGUS, current DAG or collection jobs can’t help in using these approaches below.VO COMPCHEM provides 2 examples suggesting to implement them..1. Implement a first JobType called a-Workflow (acyclic Workflow) enabling static Workflows formed by ordered-dependent Job Collections to be run on both the Grid and a HPC platform (in this respect see also the so called HPC JobType), according to the following simplified schema:JobType = a-Workflow;Dependencies = [ {hpc,htc} || {htc,hpc} ];Nodes = [hpc = [description = [JobType = "HPC";<I/O, addressing common specs>Nodes = {[<I/O specs>],[<I/O specs>…]}];];htc = [description = [JobType = "Collection";<I/O common specs>Nodes = {[<I/O specs>],[<I/O specs>…]} ];];];

An extension of DAG where Nodes can be Collections: this allows users to combine parameter study approaches on different job subclasses (platforms). We call 'a-Workflows' this type of 'enhanced DAG' (yet static and acyclic)1. Implement a second JobType called c-Workflow (cyclic Workflow) enabling dynamic Workflows according to the following example:JobType = “c-Workflow”;// Nodes definitionA = [<I/O, addressing and other specs>];B = [<I/O, addressing and other specs>];C = [<I/O, addressing and other specs>];…StartingNodes = {A};// Static and/or Dynamic dependenciesDependencies = {{A, if(var1>1, A, B)},{A, if(var2>1, C, []}}…where A, B and C have to be considered as Single Jobs which run transparentlyon both the Grid and one or more HPC platforms, according to their preliminarydefinition. In this respect, c-Workflows comprise a-Workflows even if various datamanagement issues have to be taken into account for this more generalapproach.

An extension of DAG where Dependencies can be dynamic. As reported on the simple workflow attached below, it can be cyclic when the HP block needs to be repeated more and more times. We call 'c-Workflows' this type of 'enhanced DAG' (now dynamic and cyclic).Which 'c-Workflows' would support 'Collections as Nodes', the final type could be simply defined as 'Workflows' allowing users to run dynamic and cyclic workflows supporting parameter study strategies.cyclicity |
| EGI Helpdesk | Ticket at GGUS:: <https://ggus.eu/tech/ticket_show.php?ticket=79361>Marco Cecchi commented that it can be implemented, but we need to go through the TCB. |
| Goal | To make a better usage not only of the various gLite computational resources (in particular the implementation of realistic cyclic and dynamic Workflows) but also of the possibility to bridge well-defined HPC clustered infrastructures. |
| Impact | REF: https://rt.egi.eu/rt/Ticket/Attachment/131852/47090/COMPCHEM%20Middleware%20Requirements.pdf |
| Affected services | WMS, CE. |

|  |  |
| --- | --- |
| Unique ID | #3563 |
| Title/Topic | Full support for PCKS8 private keys  |
| Reference | <https://rt.egi.eu/guest/Ticket/Display.html?id=3563> |
| Priority | 5 |
| Submitter(s) | Community WLCG (VRC) - <http://www.egi.eu/about/news/news_0116_MoU_with_wLCG.html> |
| Status at TCB | Submitted |
| Description | The users raise again the issue of slightly older version of Globusused in UMD, which does not support PKCS#8 format of private key.OpenSSL 1.0 uses this format by default, therefore virtually all themiddleware functionality breaks if those keys are not coverted.A fix is already available in Globus, it should be include in UMD releases. |
| EGI Helpdesk | Ticket at GGUS:: https://ggus.eu/tech/ticket\_show.php?ticket=80382https://ggus.eu/tech/ticket\_show.php?ticket=77148 |
| Goal | Enable support for PCKS8 private keys |
| Impact | REF: https://ggus.eu/tech/ticket\_show.php?ticket=80382 |
| Affected services | dCache openssl globus |

|  |  |
| --- | --- |
| Unique ID | #3424 |
| Title/Topic | avoid double counting of URs |
| Reference | https://rt.egi.eu/guest/Ticket/Display.html?id=3424 |
| Priority | 4 |
| Submitter(s) | Operations Office |
| Status at TCB | Submitted |
| Description | Some sites provide access to the same set of batch queues through different CE implementations. In order to prevent double counting, it is important that usage of a job is submitted only once (regardless of the CE interface through which grid submission happended). This requirement requires to assess the logic implemented by the non-glite sensors, since the gLite parser already matches the job with the CE that run it (avoiding double reporting). |
| EGI Helpdesk |  |
| Affected services | Apel parser |

## Operations Requirements (3)

This section contains the requirements approved by OMB for the 11th meeting of the EGI Technical Coordination Board (TCB).

|  |  |
| --- | --- |
| Unique ID | #3546 |
| Title/Topic | static default ordering for FQANs |
| Reference | https://rt.egi.eu/guest/Ticket/Display.html?id=3546 |
| Priority | 3 |
| Submitter(s) | NGI\_IT |
| Status at TCB | Submitted |
| Description | VOMS currently does not allow to define a static default ordering for FQANs issued to users when a specific ordering is not requested at voms-proxy-init time.Having the possibility to define such ordering would be very useful for catch-all VOs that divide users in subgroups according to their geographic location and institution affiliation and would leverage the ordering to properly account the usage of resources. |
| EGI Helpdesk | ***To be submitted directly to PT through a GGUS ticket.*** |
| Affected services | VOMS |

|  |  |
| --- | --- |
| Unique ID | #3413 |
| Title/Topic | Implement CREAM-CE WN Tests |
| Reference | https://rt.egi.eu/rt/Ticket/Display.html?id=3413 |
| Priority | 3 |
| Submitter(s) | NGI\_CH |
| Status at TCB | Submitted |
| Description | Currently jobs for WN testing are submitted through WMS/LB. These tests should rely on direct CREAM-CE submission, so that they are not affected by WMS unavailability. |
| EGI Helpdesk | ***To be submitted directly to PT through a GGUS ticket.*** |
| Affected services | CREAM-CE |