

Table of Contents

CMS Tests with the gLite Workload Management System.....	1
11 October, 2006.....	1
Memory usage.....	1
Performances.....	1
13 October, 2006.....	2
30 October, 2006.....	2
Summary table.....	2
Comments.....	3
1 November, 2006.....	3
Summary table.....	4
Comments by CE.....	4
ce01-lcg.cr.cnaf.infn.it.....	4
ce03-lcg.cr.cnaf.infn.it.....	4
ce04.pic.es.....	4
ce101.cern.ch.....	4
ce102.cern.ch.....	5

CMS Tests with the gLite Workload Management System

11 October, 2006

- Application: CMSSW_0_6_1
- WMS host: rb109.cern.ch
- RAM memory: 4 GB
- LB server: rb109.cern.ch (*)
- Number of submitted jobs: 25000
- Number of jobs/collection: 100
- Number of collections actually submitted: 234
- Number of CEs: 24
- Submission start time: 10/10/06, 18:45
- Submission end time: 10/12/06, 19:10
- Maximum number of planners/DAG: 2

Memory usage

During the submission, the swap memory usage increased linearly up to 40%, and decreased rapidly shortly after the job submission stopped. This means that, at some point, the total memory used was 5.8 GB.

The number of planners reached about 250, which accounted for about 1.4 GB. Other processes which used a lot of memory are the WMPProxy server (>1.5 GB), the WM (>0.5 GB) and Condor (>0.4 GB).

Concerning WMPProxy, the reason why it took so much memory is not clear, and it was suggested to decrease the number of server threads (30 being the current value). In detail, one could:

- reduce the maximum number of processes running simultaneously (-maxProcesses, -maxClassProcesses)
- make the "idle" processes killing policy more aggressive (-KillInterval: default is 300 secs))

This is done by changing the "FastCgiConfig" directive at the end of file /opt/glite/etc/glite_wms_wmproxy_httpd.conf as follows:

```
FastCgiConfig -restart -restart-delay 5 -idle-timeout 3600 *-KillInterval 150 \  
  -maxProcesses 25 -maxClassProcesses 10 -minProcesses 5* ..... (keep the rest as it is)
```

Concerning the WorkloadManager, it is a known issue that the Task Queue eats a lot of memory; this is not seen in gLite 3.1.

Performances

During the job submission, attempts to submit jobs from another UI took unreasonable amounts of time (~3' for a single job), probably due to the high level of swapping.

During the submission, the number of jobs in Submitted status kept increasing, meaning that the WMS could not keep up with the submission rate. After the end of the submission, it took about 10 hours to dispatch all the jobs. Again, this is probably due to a general slowness of the machine due to the swapping. The submission rate was also very close to the maximum dispatch rate, and if it was actually a bit higher, jobs would keep accumulating even without the swap memory effect. Therefore it is recommended to submit at a rate significantly lower (maybe 70%?).

It is also important to have as soon as possible the fix which limits the time for which the WM tries to match jobs in the task queue: the current limit of 24 h is too long, because a collection whose jobs cannot be matched is kept alive for a long time, even if it is clear that the jobs cannot ever be matched.

I noticed that, on a very busy RB, jobs in a collection may be matched even 24 hours after submission:

```
- JOBID: https://rb109.cern.ch:9000/nrn1NkJeABRP9u6f1FPfyA
Event      Time          Reason      Exit Src Result      Host
RegJob     10/10/06 20:52:12          NS          rb109.cern.ch
RegJob     10/10/06 20:52:15          NS          rb109.cern.ch
RegJob     10/10/06 20:53:22          NS          rb109.cern.ch
HelperCall 10/11/06 21:10:33          BH          rb109.cern.ch
Pending    10/11/06 21:29:02 NO_MATCH     BH          rb109.cern.ch
```

Note: I discovered that I never really used a separate LB server: the LBAddress attribute must be in the common section of the JDL for a collection, not in the node JDL. Another possibility is to configure the RB to use it in the RB configuration. In a recently released tag, an LB server can be specified also in the UI configuration.

13 October, 2006

* Application: CMSSW_0_6_1

- WMS host: rb109.cern.ch
- RAM memory: 4 GB
- LB server: lxb7026.cern.ch
- Number of submitted jobs: 14000
- Number of jobs/collection: 100
- Number of collections actually submitted: 140
- Number of CEs: 28
- Submission start time: 10/13/06, 11:10
- Submission end time: 10/14/06, 9:43
- Maximum number of planners/DAG: 2

30 October, 2006

* Application: CMSSW_0_6_1

- WMS host: lxb7283.cern.ch
- Flavour: gLite 3.1
- RAM memory: 4 GB
- LB server: lxb7283.cern.ch
- Number of submitted jobs: 2400
- Number of jobs/collection: 100
- Number of collections actually submitted: 24
- Number of CEs: 24
- Submission start time: 10/30/06, 12:30
- Submission end time: 10/30/06, 12:59
- Maximum number of planners/DAG: 10

Summary table

Site	Submit	Wait	Ready	Sched	Run	Done(S)	Done(F)	Abo	Clear	Canc
cclcgceli02.in2p3.fr	23	0	0	1	0	76	0	0	0	0
ce01-lcg.cr.cnaif.infn.it	0	0	0	0	0	100	0	0	0	0
ce01-lcg.projects.cscs.ch	0	0	0	0	0	100	0	0	0	0

GLiteWMSCMSTests < LCG < TWiki

ce03-lcg.cr.cnaf.infn.it	0	0	0	0	0	100	0	0	0	0
ce04.pic.es	0	0	0	0	0	100	0	0	0	0
ce106.cern.ch	0	0	0	0	0	100	0	0	0	0
ceitep.itep.ru	0	0	0	0	0	100	0	0	0	0
cmslscge.fnal.gov	0	0	0	0	0	100	0	0	0	0
cmsrm-ce01.roma1.infn.it	0	0	0	0	0	100	0	0	0	0
dgc-grid-40.brunel.ac.uk	0	0	0	0	0	100	0	0	0	0
egeece.ifca.org.es	80	20	0	0	0	0	0	0	0	0
grid-ce0.desy.de	0	0	0	0	0	100	0	0	0	0
grid-ce1.desy.de	0	0	0	0	0	100	0	0	0	0
grid-ce2.desy.de	0	0	0	0	0	100	0	0	0	0
grid10.lal.in2p3.fr	0	0	0	0	0	100	0	0	0	0
grid109.kfki.hu	0	0	0	0	0	100	0	0	0	0
gridba2.ba.infn.it	0	0	0	0	0	100	0	0	0	0
gridce.iihe.ac.be	0	0	0	0	0	97	3	0	0	0
gw39.hep.ph.ic.ac.uk	0	0	0	49	0	3	0	48	0	0
lcg00125.grid.sinica.edu.tw	0	0	0	9	9	73	0	9	0	0
lcg06.sinp.msu.ru	0	0	0	100	0	0	0	0	0	0
oberon.hep.kbfi.ee	0	0	0	100	0	0	0	0	0	0
polgrid1.in2p3.fr	0	0	0	0	0	100	0	0	0	0
t2-ce-02.lnl.infn.it	0	0	0	0	0	100	0	0	0	0

Comments

The Submitted jobs at cclcgceli02.in2p3.fr have indeed finished, but in the logging info the last event is a RegJob, whose timestamp, however, is close to the other RegJob events. In addition, for those jobs glite-job-status -v 3 reports timestamps only for Submitted and Waiting. This is linked to the fact that the sequence code of the logged event is wrong: the last RegJob event had a sequence code

```
UI=000000:NS=0000000001:WM=000000:BH=0000000000:JSS=000000:LM=000000:LRMS=000000:APP=000000
```

while the first event from WM/BH had

```
UI=000000:NS=0000000000:WM=000000:BH=0000000001:JSS=000000:LM=000000:LRMS=000000:APP=000000
```

instead of

```
UI=000000:NS=0000000001:WM=000000:BH=0000000001:JSS=000000:LM=000000:LRMS=000000:APP=000000
```

which causes all subsequent events to be considered prior to the last RegJob. The reason of this behaviour is not yet understood.

The aborted jobs at gw39.hep.ph.ic.ac.uk and lcg00125.grid.sinica.edu.tw had the "unspecified gridmanager error".

The 3 failed jobs at gridce.iihe.ac.be have the "Got a job held event, reason: Globus error 124: old job manager is still alive" error.

The jobs at egeece.ifca.org.es are either Submitted or Waiting because the no CE can be matched. The Waiting jobs are 20, which is strange because the maximum number of planners/DAG is 10.

1 November, 2006

- VOMS proxy duration: 48 hours
- Application: CMSSW_0_6_1
- WMS host: lxb7283.cern.ch
- Flavour: gLite 3.1
- RAM memory: 4 GB

- LB server: lxb7026.cern.ch
- Number of submitted jobs: 21000
- Number of jobs/collection: 200
- Number of collections actually submitted: 105
- Number of CEs: 21
- Submission start time: 11/01/06, 12:03
- Submission end time: 11/02/06, 11:27
- Maximum number of planners/DAG: 10

Summary table

Site	Submit	Wait	Ready	Sched	Run	Done (S)	Done (F)	Abo	Clear	Canc
ce01-lcg.cr.cnaf.infn.it	20	0	0	0	0	980	0	0	0	0
ce03-lcg.cr.cnaf.infn.it	107	0	0	0	0	893	0	0	0	0
ce04.pic.es	37	0	0	0	2	961	0	0	0	0
ce101.cern.ch	40	1	0	0	1	934	0	24	0	0
ce102.cern.ch	215	12	0	0	0	0	0	773	0	0
ce105.cern.ch	178	13	0	0	0	0	0	809	0	0
ce106.cern.ch	236	0	0	0	0	764	0	0	0	0
ce107.cern.ch	288	25	0	0	0	117	0	570	0	0
ceitep.itep.ru	110	0	0	0	0	890	0	0	0	0
cmslsgce.fnal.gov	49	0	0	0	0	951	0	0	0	0
cmsrm-ce01.roma1.infn.it	259	0	0	0	0	741	0	0	0	0
dgc-grid-40.brunel.ac.uk	26	0	0	0	0	974	0	0	0	0
grid-ce0.desy.de	228	1	0	0	0	771	0	0	0	0
grid10.lal.in2p3.fr	69	0	0	0	0	931	0	0	0	0
grid109.kfki.hu	50	0	0	0	0	950	0	0	0	0
gridce.ihe.ac.be	244	0	0	0	0	745	11	0	0	0
gw39.hep.ph.ic.ac.uk	0	0	0	0	0	442	168	383	0	7
lcg00125.grid.sinica.edu.tw	7	57	0	0	0	773	17	146	0	0
lcg02.ciemat.es	16	0	0	0	0	980	0	4	0	0
oberon.hep.kbfi.ee	206	0	0	0	0	392	359	43	0	0
t2-ce-02.lnl.infn.it	375	0	0	0	0	625	0	0	0	0

Comments by CE

ce01-lcg.cr.cnaf.infn.it

20 jobs apparently Submitted but finished, with a RegJob event at the end of the logging info.

ce03-lcg.cr.cnaf.infn.it

38 jobs stuck in Submitted status (only 3 RegJob events in logging info): error message "cannot create LB context". For the other Submitted jobs, see above.

ce04.pic.es

37 jobs apparently Submitted.

ce101.cern.ch

24 jobs are Submitted with reason (!) "no matching resources found"; they have a Pending event before the third RegJob event. The other 16 Submitted jobs have no reason and the third RegJob is before the first Pending. 13 Aborted with "X509 proxy expired" because no CE could be matched before the proxy expired. 11 Aborted with "request expired" because no CE could be matched in 24 hours. 1 Running, no termination events were received. 1 Waiting, no Abort event was logged.

ce102.cern.ch

-- AndreaSciaba - 30 Oct 2006

This topic: LCG > GLiteWMSCMSTests

Topic revision: r6 - 2006-11-06 - AndreaSciaba



Copyright &© 2008-2021 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.
or Ideas, requests, problems regarding TWiki? use [Discourse](#) or [Send feedback](#)