

Table of Contents

Release notes for Patch #1841.....	1
------------------------------------	---

Release notes for Patch #1841

Release 08_98 of the WMS for gLite3.1/SL4. Changes with respect to the current production version (patch #1726):

- Enabled submission to CREAM CE. A newly introduced component in the WMS internal architecture, called ICE, implements the job submission service to CREAM. Its functionality can be compared to what the three components JC, LM and CondorG do for the submission to LCG CE
- Added recovery procedure for the WM. This feature is enabled by the option `EnableRecovery = true` in the `WorkloadManager` section of the configuration file <https://twiki.cnaf.infn.it/cgi-bin/twiki/view/EgeeJra1It/WMSConfFile>. It basically works in this way: upon restart, old requests are reconsidered and the LB is queried to know exactly where to resume their processing in such a way that no operation is performed twice or more.

Important: 1) **if the recovery is not enabled, simply starting and stopping the `glite-wms-workload_manager` process (and of course restarting after whatever kind of interruption) might cause duplicating requests.** 2) the recovery **only works with "JobDir"** (see below)

- "JobDir" is a mailbox-based persistent communication mechanism, for the moment adopted between the WM proxy and the WM. In the present release it is enabled by default. A tool is available for converting from the former mechanism based on filelist (conversion in the opposite way is also supported). At the moment this not done automatically. Of course, another option to handle this transition will consist in putting the WMS in drain and wait for the filelist to be empty.
- The output sandbox can be limited: How the OSB limit works
- LDAP queries to fetch information in the Information Supermarket from the BDII can now be pre-filtered. This can be very helpful whenever a WMS instance is dedicated to only one VO. Typically, using a production BDII, the ISM reaches a size of 6-7000 entries, with the consequence that the match-making for a job can take a time of the order of ten seconds * note: only about one second with the subsequent patches *. Using the filter on the VO name, as for the aforementioned use-case, significantly reduces the MM time. The filtering expression has to be set by assigning the relevant parameter in the `WorkloadManager` section of the configuration file, as shown in the following example:
 - ◆ `IsmIiLDAPCEFilterExt="(|(GlueCEAccessControlBaseRule=VO:cms)(GlueCEAccessControlBase`
- LDAP queries to the BDII can now be done asynchronously (attribute `IsmIiLDAPSearchAsync = true` in the WM section). This mode is typically faster than the usual synchronous one.
- Purchasing from CEMon has been temporarily disabled
- Purchasing from R-GMA has been removed
- Added support for MPI jobs according to the latest specifications from the MPI working group. The value "MPICH" for the JDL attribute `JobType` becomes deprecated from now on, just set it to "Normal" and follow the new guideline instead
- Support for interactive jobs has been dismissed. However, the functionality is not compromised because it can be achieved using a tool called `i2glogin` (formerly known as `glogin`). This different approach is actually more flexible, the user being totally in charge, and it follows the trend set by the new handling for MPI jobs).
- Known issues:

- ◆ Performance problems in the newly introduced ICE component when it has to deal with several CEs
- ◆ Very often, especially under high loads, the virtual memory occupation for the glite-wms-workload_manager process may reach very high values, such as one Gigabyte and more. This is not about a memory leak, but simply the effect of a well-known problem with the allocator which comes with the glibc (the so called ptmalloc2). See tcmalloc for a more detailed explanation. This problem can be avoided using run-time redirection to whatever lock-free, optimized alternative allocator, to avoid excessive swap activity. It is highly suggested doing so wherever RAM is less than or equal to 4Gb. Here is our recipe which makes use of the Tcmalloc, such an alternate allocator distributed by Google under BSD license:
 - ◇ install the two rpms, google-perftools-devel-???.rpm and google-perftools-???.rpm (just pick up the latest version, older versions should work anyway, just in case),
 - ◇ enable the malloc redirection for the WM by editing the glite-wms-wm script. It is just a matter of removing the comment in the following line:

```
#use_google_perf_tools=1
```

- ◆ Bug #35244 : Can't submit jobs using voms proxies with roles due to a mapping problem *fixed by patch 2055*
- ◆ Bug #40951 : Cleared event is not logged for nodes
- ◆ Bug #40982 : When a collection is aborted the "Abort" event should be logged for the sub-nodes as well /2
- ◆ Bug #42587 : Error processing DAG dependencies while generating the ISB for final node
- ◆ Bug #42590 : the WM terminates unexpectedly handling a cancel request.
- ◆ Bug #43368 : Long Nordugrid ARC Jobs go into the HELD state and get resubmitted

-- AlessioGianelle - 27 May 2008

This topic: EgeeJra1It > ReleaseNotes1841

Topic revision: r32 - 2011-02-10 - MassimoSgaravatto



Copyright © 2008-2020 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback