

**The ALICE Christmas Production**  
L. Betev, S. Lemaitre, M. Litmaath, P.  
Mendez, E. Roche

WLCG LCG Meeting  
14th January 2009

# Outlook

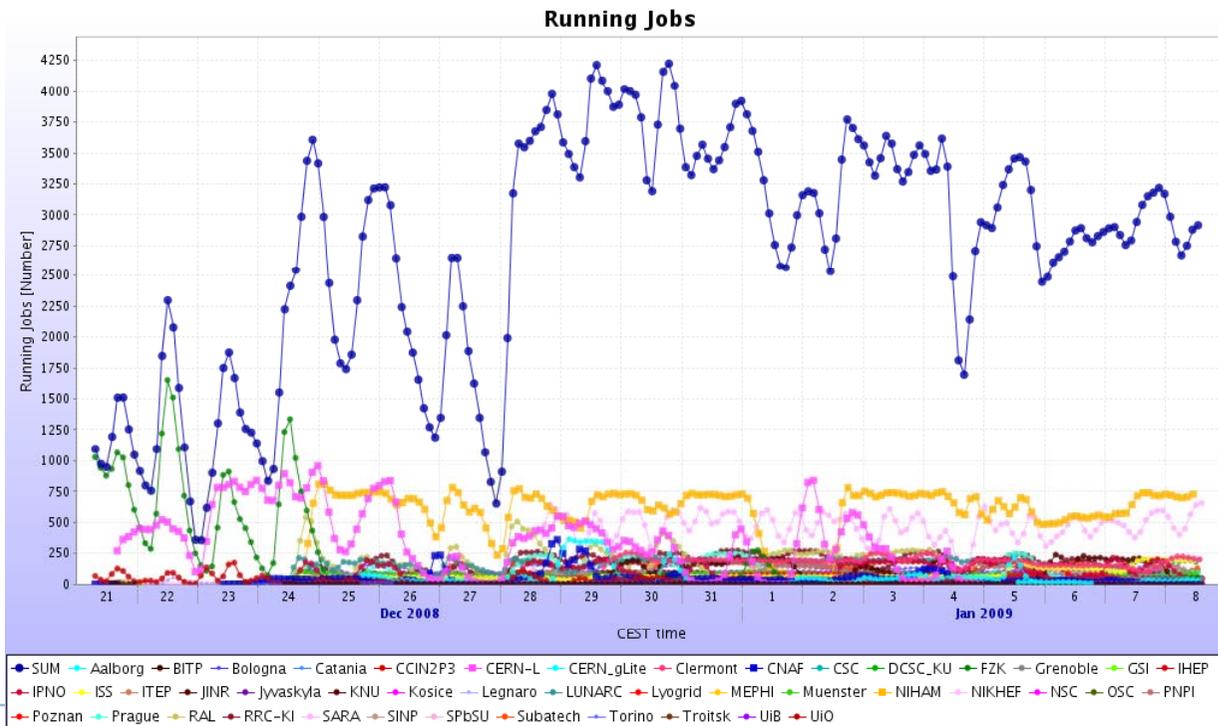
---

- ▶ The last ALICE production run began the 21st of December
  - ▶ This new run includes the latest AliEnv2.16 previously deployed at all sites
  - ▶ **This is the 1st AliEn version which fully deprecates the RB usage**
    - ▶ **HOWEVER:ALICE has been using the WMS submission mode in most of the sites since months in previous AliEn versions**
  - ▶ It implements also a CREAM-CE module for CREAM submissions
  - ▶ **3500 jobs daily run in average**
  - ▶ The following slides collect the experiences of the experiment during the Christmas period and also the conclusions of the post-mortem meeting between IT/GS members and WMS experts



# General summary of the production

- ▶ 36 sites have participated in the Christmas production (50%), all TI sites OK
- ▶ WMS reconfiguration needed during the vacations period
  - ▶ This will be the focus of the Alice report in the following slides



# The Services: WMS (I)

---

- ▶ **ALICE counts currently on 3 dedicated WMS nodes at CERN**
  - ▶ wms103 and wms109 (gLite3.0)
  - ▶ wms204 (8 cores machine, gLite3.1)
  - ▶ These nodes play a central role in the whole ALICE production
- ▶ **The 29th of December Maarten announces to ALICE a huge backlog in wms204**
  - ▶ A large number of jobs were being submitted through this node
  - ▶ As result the further job submission processing were slower and slower



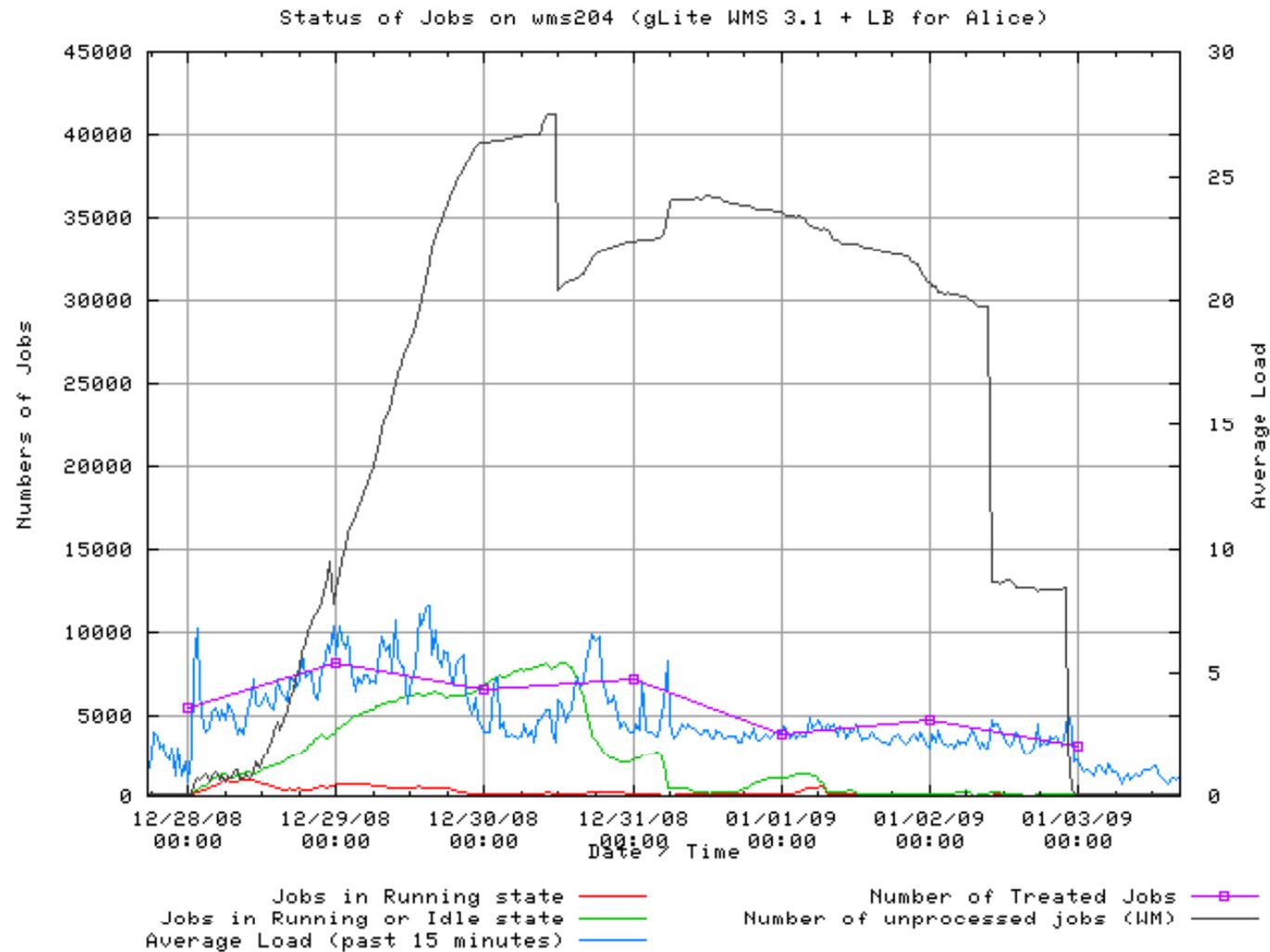
# Some results seen during Christmas

---

- ▶ i.fl: it represents the number of unprocessed entries in the input.fl file, the input queue for the Workload Manager service
- ▶ q.fl: it represents the number of unprocessed entries in the queue.fl file, the input queue for the Job Controller service
- ▶ <http://wmsmon.cern.ch/monitoring/monitoring.html>

						i.fl	q.fl
<a href="#">wms103</a>	01/08-11:30	112	1357	11913	9.13	4282	37682
<a href="#">wms105</a>	01/08-11:30	325	942	2317	2.06	4	0
<a href="#">wms106</a>	01/08-11:30	304	877	2500	1.93	7	0
<a href="#">wms107</a>	01/08-11:30	815	1322	2355	3.96	1	0
<a href="#">wms108</a>	01/08-11:30	3	3	0	0.25	0	0
<a href="#">wms109</a>	01/08-11:30	159	365	3940	10.80	9264	64474

# Example: wms204 backlog



# The Services: WMS (II)

---

- ▶ Where this was hapenning?
  - ▶ Basically 2 T2 sites were catching a huge number of jobs: MEPHI in Russia and the T2 is Prague
- ▶ Why this was hapenning?
  - ▶ Normally several reasons can drive to this situation:
    - ▶ **The destination queue is not available**
      - The submitted jobs are then kept for a further retry: (up to 2 retries: unmatched requests are discarded after 2 hours)
      - But ALICE has set the Shallow resubmission to zero and explicitly asked the WMS experts to set the nodes avoiding any possible resubmission
    - ▶ **Any configuration problem at the site keeps on submitting jobs**
      - Since these jobs are visible nowhere, they do not exist for ALICE and therefore, the system keeps submitting and submitting
    - ▶ **In any case the submission regime of ALICE is not so high to provoque such a huge backlogs in nodes as wms204**
      - The previous reasons can be ingredientes to the problem, but cannot be the only reason for such a load
      - On wms204 the matchmaking became very slow due to unknown causes; the developers have been involved



# Effects of this high load

---

- ▶ ALICE was seeing jobs in status READY and WAITING for a long time
  - ▶ The experiment still does not consider READY and WAITING as problematic status so it keeps on submitting and submitting... SNOWBALL: creating huge backlogs
  - ▶ **Request: Could the WMS be configured to avoid new submissions once it gets in such a state?**
    - ▶ **Proposed during the post-mortem meeting with the WMS experts, it could be in place for the end of February 2009 (earliest)**
- ▶ Why these job status?
  - ▶ the Workload Manager has to deal with all preceding requests in the "i.fl" queue at least once (a submission that does not match will be retried later if it has not expired by that time);
  - ▶ the Job Controller has to deal with all preceding requests in the "q.fl" queue.



# The WMS: ALICE Procedures

---

- ▶ ALICE stopped immediately the submission through wms204 at all sites putting the highest weight on wms103 and wms109
  - ▶ The situation was solved in wms204 but appeared in wms103 and wms109
    - ▶ wms103 and wms109 (gLite 3.0) had a different problem that could not be explained satisfactorily either
  - ▶ In addition access to wms117 was also ensured to ALICE for this period
    - ▶ The node developed the same symptoms as wms204
- ▶ As result a continuous care of the WMS has been followed during this period changing the wms in production when needed



# Possible source of problems

---

- ▶ **ALICE jdl construction?**
  - ▶ The experiment has always defined simple jdl files for their agents
- ▶ **BDII overloaded?**
  - ▶ It should be then affecting all VOs while performing the matchmaking
  - ▶ In addition several tests were made while quering the BDII and obtaining positive results
- ▶ **Network problems?**
  - ▶ During several days?...And affecting ALICE only?
- ▶ **Overloading myproxy server**
  - ▶ Indeed it was found a high load of myproxy by ALICE
  - ▶ However this seems to be uncorrelated with the WMS issue
    - ▶ **Although an overload on myproxy server can slow down the WMS processing, this should then be visible for all WMS of all VOs**



# How to solve the myproxy server issue

---

- ▶ Faster machines have been already requested to replaced the current nodes of myproxy server
  - ▶ Proposed during the Christmas period the request has been already done
- ▶ In addition ALICE is currently changing the submission procedure to ensure a proxy delegation request once per hour
  - ▶ In case of any problem at a VOBOX, this procedure can ensure a 'frugal' myproxy server usage
  - ▶ The new submission procedure will have a beta version this week at Subatech (France)



# Conclusions

---

- ▶ Still pending the issue with the WMS: We still cannot conclude why such a big backlogs have been created during this vacation period
- ▶ Two new WMS@CERN have been already announced: wms214 and wms215 in addition to wms204
  - ▶ All of them with independent LB
  - ▶ 8 core machines
  - ▶ Glite3.1
  - ▶ wms103 and wms109 will be fully deprecated end of February
- ▶ At this moment and due to an AliRoot update ALICE is not in full production
  - ▶ **As soon as the experiment restarts production we will follow carefully the evolution of the 3 nodes reporting any further issue to the developers**



# Final remarks

---

- ▶ **ALICE has a lack of WMS**
  - ▶ France still is not providing any WMS which can be put in production
  - ▶ WMS provided at RDIG, Italy, NL-TI, FZK and RAL
  - ▶ CERN WMS play a central role for many ALICE sites and are always a failover for the sites, even if a local WMS is available
  - ▶ ALICE wishes to thank the IT/GS (Maarten and Patricia in particular) for the efficient support during the Christmas running

