The WMS:

Recent problems and current status

Nick Thackray

Grid Operations

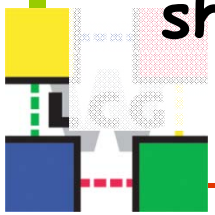
CERN

WLCG Grid Deployment Board,
CERN 11 March 2009

The problem – a recap

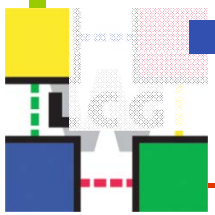
- Christmas period: ALICE seeing issues in all 3 of their dedicated WMSs at CERN, causing serious problems for production
- Symptoms:
  - High load on WMS machines
  - Effect: newly submitted requests stayed in status Waiting or Ready forever

  ➔ suicide mode! – new requests still coming and being accepted, thus worsening the situation…
- But at least one of the machines seeing the problem was 8 core / 8 GB, so unlikely to be a shortage of processing power



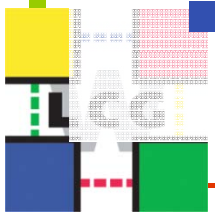
Possible causes?

- Destination queue not available or configuration problem at the destination site?
  - Submitted jobs are then kept for 2 hours (up to 3 retries per job)
- ALICE JDL construction?
  - Complicated JDL will slow down the matchmaking process ➔ The workload manager will not be able to keep up with the requests which are being sent by the WMProxy
  - Possible cause at CERN, but problem also seen at sites with many fewer ALICE queues than CERN
- BDII overloaded? Network problems? MyProxy server overload?
  - All discarded



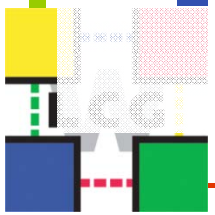
Some obvious questions…

- Why only *ALICE*?
  - The other experiments weren't complaining of problems, so what's the difference?
  - *ALICE* submit individual jobs rather than collections
- Why only *CERN*?
  - Turns out other sites *have* seen similar issues (e.g. GRIF)
- Why now (and not earlier)?
  - Appears that RB and gLite 3.0 WMS didn't suffer from these problems: Did we swap performance for stability
  - What changed in WMS from gLite 3.0 to 3.1?
- Why the disagreement on sustainable submission rate?
  - Jobs/day vs. Peak rate
  - Problems seemed to occur when some job submission rate was exceeded



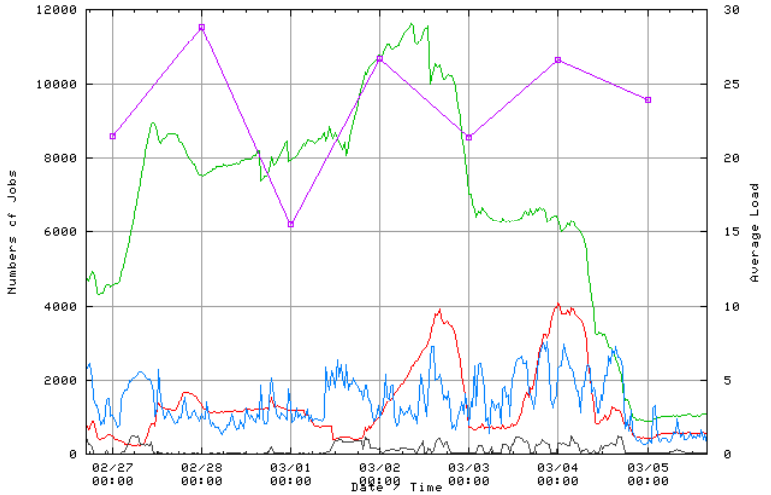
Getting to the bottom of it

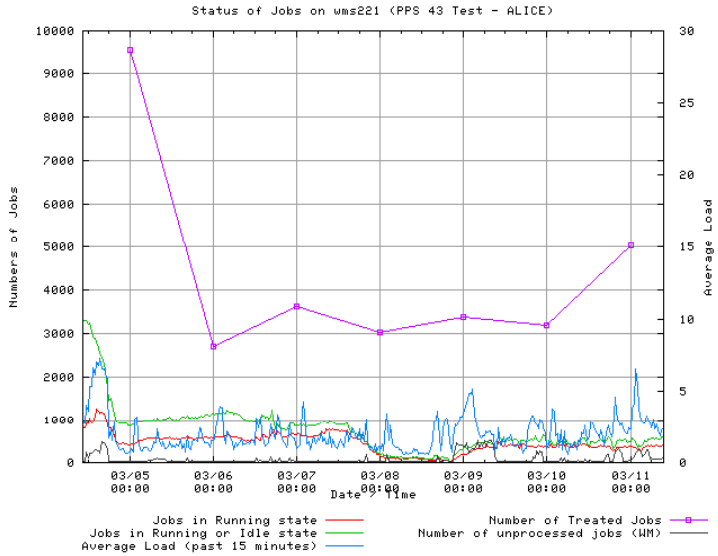
- Issue tracked at daily WLCG operations meeting but no resolution was found after > 1 month
- On 12 Feb a war council was called to decide how to tackle this issue and bring it to resolution
  - ALICE (Patricia)
  - Grid Operations (Nick T.)
  - WMS expertise (Maarten, Steve T.)
  - CERN WMS service manager (Ewan)
- Top priority: get ALICE operations up and running!
  - Reconfigure ALICE framework to use more than just CERN WMSs for production
  - Monitor WMSs at CERN and drain as soon as problem is seen



But how to fix the problem?

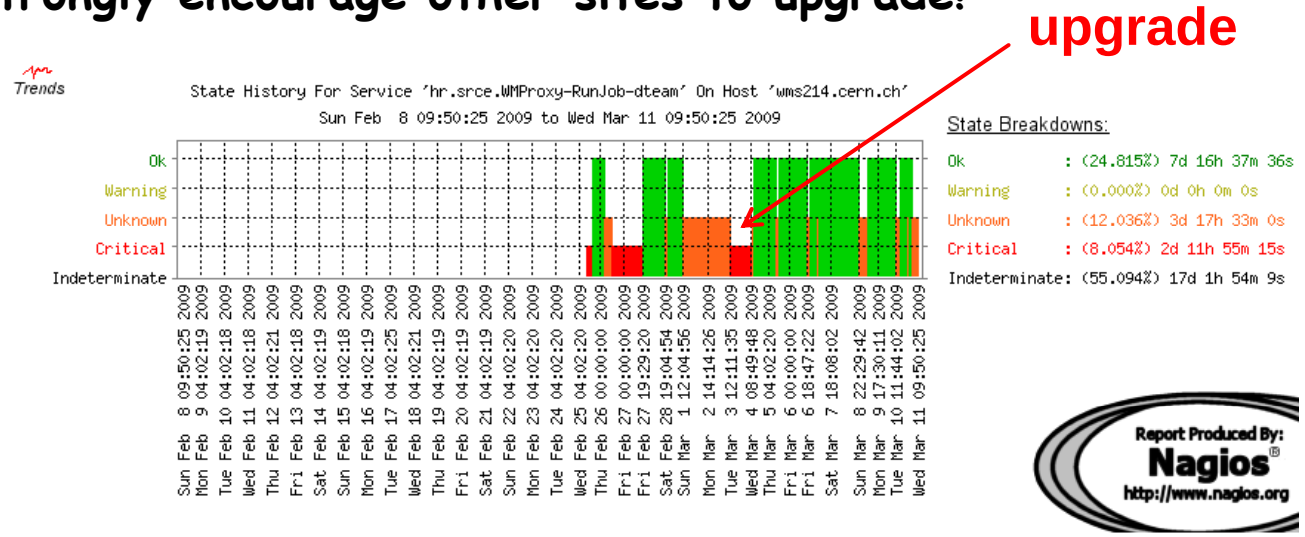
- Might the WMS "mega-patch" solve (or at least help) the problem?
  - Patch 2562 contained ~6 months of distilled wisdom, including a "limiter" based on number of unprocessed jobs
  - This was applied to WMS221.cern.ch (ALICE production WMS at CERN) on 18th Feb. (3 weeks ago)
- WMS221 progressively loaded up with jobs by ALICE and closely monitored by Ewan





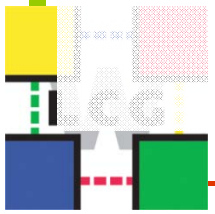
Results

- Results
  - Limiter greatly improved stability of WMS - but this doesn't improve throughput
  - Currently seeing 10,000 individually submitted jobs/day without problems. Will ramp this up next week
  - Came across 2 Known Issues (#47150, #47040).The workarounds are a little painful – these need to be fixed soon
- Once it was clear that the mega-patch helped it was pushed into production (26 Feb)
  - In the process of upgrading all CERN WMSs
  - Strongly encourage other sites to upgrade!



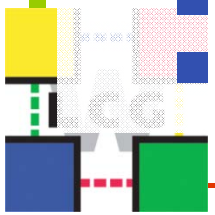
Next steps

- Patch 2562 allows pre-filtering of the information put into the ISM; i.e. can be made VO specific
  - Faster matchmaking due to fewer entries (linear relation)
  - This has already been tried at CNAF but needs more stress testing
- Matchmaking time window is now configurable with YAIM
- DNS load balancing with arbiter for WMS is in the pipeline
  - Suffers from bug 43633 which may cause jobs to be refused even when some nodes in the cluster are OK
  - Also bug 47677 : when a WMS is removed from the DNS, any UI config referring to that WMS must be updated otherwise the UI can no longer submit jobs for the affected VOs



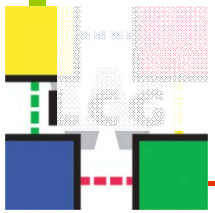
Summary

- 3.1 WMS could not handle a backlogs of jobs
- Several improvements introduced in the "mega-patch"
  - Limiter for WMS job acceptance
    - Can now handle around 10,000 individually submitted jobs/day without collapsing (many more with bulk submission)
    - Higher rates possible for single job submission? – further testing needed
  - Pre-filtering of information in ISM
    - Will probably help those WMSs dedicated to 1-2 VOs
    - Again, further testing needed
- Further improvements in the pipeline
  - DNS aliasing
    - Needs some bugs fixing before going to production
- But still some issues
  - Bug 47150, Bug 47040 (Known Issues)
  - Bug 46209 which prevents WMS being published correctly
    - Needs to be fixed as soon as possible



Conclusions

- We still don't really understand the causes
  - Or why this wasn't seen in the gLite 3.0 WMS
- We were lucky that the mega-patch improved things
  - If it hadn't, we could only now have started debugging
    - And who knows how long it might have taken
- But is the gLite 3.1 WMS good enough now?
  - Is job matching rate of 1 job per 6 seconds enough? (14,400 jobs/day)
- Need to stress test the ICE-WMS / CREAM CE combination well !



Questions…?