T2 Summary

Michel Jouvin
LAL, Orsay
jouvin@lal.in2p3.fr
http://grif.fr

July 9, 2009
STEP09 Postmortem, CERN

Methodology

- Sent an email to all T2 contacts asking for feedback: got 1 answer

  - Reliable, VO-independent contact with T2s is still a problem…

  - Not a lack of interest: I received several answers after posting first draft of this presentation…

- Atlas summary per cloud

  - https://twiki.cern.ch/twiki/bin/view/Atlas/Step09Feedback#CA_Cloud_T2s

- D. Bonacorsi private summary

- Experience of French T2s

  - http://indico.in2p3.fr/conferenceDisplay.py?confId=2049

- Not necessarily representative… but hopefully not too specific!

High-Level Summary…

- Useful exercise: first time a big focus was put on T2s for analysis

  - Competition with production

  - Multi-VO for T2s supported ATLAS/CMS(/ALICE)

- Not necessarily tried heroic efforts to solve everything during the exercise

  - Solve the problems in a « sustainable way »

  - Identify problems requiring more thinking or significant configuration changes

  - Avoid ad-hoc changes just to meet the target…

- Concentrate on post-mortem work for the most difficult issues

... High-Level Summary

- Overall successful for the VOs but hiding a lot of discrepancies between sites

  - ATLAS: 50% of analysis by 11 sites

  - Site may have been running well only part of the exercise

- Situation is easier/better for VO-dedicated T2s

  - But multi-VO choice is highly dependent of local conditions (eg. funding)

Storage

- Shortage of resources: not necessarily representative of a real problem
    - Several T2s delayed some procurements until September
    - Some T2s had on-going infrastructure work required before deployment of (available) new resources
        - Eg. GRIF
    - Sites cannot behave as if data were coming...
- Resource shortage may always happen: importance of « accurate » requirements by experiments
    - ATLAS was a problem : 50% above requirements for ATLASMCDISK
    - Need ability to quickly remove/reroute activity to another site
- STEP09 was the first time VOs used their share... requirements more credible in the future!

Transfers

- Many sites affected by transfer « instabilities »
    - ATLAS more affected than CMS ?
    - A well-behaving site suddenly failing a lot of transfers without any trivial reason, human error or misconfiguration

- Many sites (all in France) hit by lcg-cp timeouts
    - 30 minutes, probably not related to the load
    - Found evidence at several sites of hung gsiftp processes, 2 weeks after the end of STEP09
    - Same behavior observed at DPM and non-DPM sites (LIP with StoRM/LUSTRE, see last pre-GDB)
    - A Linux bug ? Very difficult to find a troubleshooting procedure to make progress
        - Try to concentrate on sites affected a lot by the problem
        - Compare with SL5 (as DPM now ready) ?

Data Access / CPU Efficiency

- A lot of work needed to assess the exact performance for each site

    - Hammerclould statistics very valuable

    - Data access performance is the issue to focus on for large T2s

- GRIF/LAL (and may be others) demonstrated good performance of 10 Gb/s connection for disk servers

    - Sustained 4 Gb/s on all disk servers

- Access to shared area hosting VO SW area may have a non-negligible impact on CPU efficiency too

    - Particularly true for Atlas where SW setup is putting a high load on SW area

    - Take into account several 100 of jobs can start at the same time…

Job Scheduling…

- ATLAS and CMS insisted for intra-VO fairshare to guarantee balanced resource access for all activities
  - Very different from giving priority to one activity

- Has not been a major problem so far because resources were under-used in average
  - More resources than pledged at many T2s, in particular large multi-VO T2s
  - Very low-level of VO concurrency in average

- Inter-VO fairshare working pretty well
  - Fairshare history may impact the access to resources when a VO takes advantage of under usage by others
  - Seen at GRIF/LAL with ATLAS which started at a very high level before CMS

... Job Scheduling

- Intra-VO scheduling is very dependent on batch scheduler features

  - Require hierarchical fairshares

  - Else can only mimics and probably need very frequent adjustments

  - Unfortunately MAUI doesn't support hierarchical fairshare

    - May boost priority for some users/accounts based on their fairshare target but not relative to the VO share

    - Analysis is difficult because this is not one user, not even a group of user

- Sharing experiences with different schedulers would be important

  - Pre-GDB ?

  - Some work started in France

Jobs and Data Placement

- From GRIF experience with CMS…

- Jobs sent where the data are… as expected

- Observed very large batch of jobs 1-2K) submitted by 2 or 3 users at the same time to the same CE

  - If CMS share is a few 100s slots, may take quite some times

  - Also impacted by the fairshare history : cannot take advantage of fairshare but refuse drawbacks…

- Very high number of waiting jobs may impact performance/stability of the batch scheduler

  - Particularly true for MAUI…

  - Saw little impact on job scheduling but may lead to inaccurate information into the BDII

  - Would be interesting to get feedback for other schedulers

- Solution is not in the hand of the site…

VO Tests Scheduling

- Also based on experience with CMS… but may have a wider scope

- CMS running analysis tests supposed to run with 12 hours

  - Failing to do it lead to the site being automatically blacklisted

  - Automatically removed from blacklist as soon as the « problem » disappear

- Because of the fairshare issues mentioned and the load spike from end-users, test job turnaround can be slow

  - Is it a problem to solve ? CMS tends to insist for a specific scheduling for these jobs to ensure a « fast » turnaround…

  - My personnal view is that this reflects the real state of the site and VO should be prepared to use another site in this situation : impact on data replication

ATLAS: WMS vs. Panda

- Seen in France... large job CPU efficiency impact depending on the submission method

  - No idea why?

  - ATLAS help needed...

Conclusions

- STEP09 was a very useful exercise and T2s are committed to solved the problems highlighted asap

  - First time we have this level of feedback and figures

- Storage resources sometimes undersized but probably not a long-term problem

- Data transfers showed strange timeout problems that may not be related to storageware

  - Need more investigation: how? Who?

- Intra-VO fairshare is the big issue as soon as there is competition for the resources

  - Not clear if MAUI can do the job in multi-VO context

- VO communications with T2s now effective and efficient but still need to establish VO-independent communication channels

  - Mailing lists seem not to be very effective. Cleanup ?