

Tracking Tier-2 Resources

Presentation not intended to be complete, rather to trigger a discussion



Computing Resources Session
CMS Plenary Week, CERN
December 8th, 2008

THOMAS KRESS

Tier-2 Resources and Connections



- CPU
 - Storage <==== „Hard“ resources
 - (Network)
 - Computing Model Capability
 - Site Reliability & Performance <==== „Soft“ resources
 - Support
 - CMS Central Tasks
 - Detector & Physics Groups
 - National and Local CMS Users
 - Other CMS Users
 - Other VOs
- Consumers ==>
Competitors

SiteDB & Pledge Information



SiteDB :: RWTH Resource

Below is the current resource pledge for [RWTH](#). To edit the p

View pledge for Quarter 4 2008

Compute

Total processing power available to CMS.

Cpu	450.0 (kSI2k)
Job Slots	360.0 (#)

Storage

Total storage allocated to CMS.

Disk Store	100.0 (TB)
Tape Store	0.0 (TB)

Storage available for local users, and available for transfer (e.g. via PhEDEx).

Local Store	(TB)
Wan Store	(TB)

Network

Expected national and international bandwidth.

NREN connection speed	5.0 (Gbps)
OPN connection speed	0.0 (Gbps)

- Application, database, API
<https://twiki.cern.ch/twiki/bin/view/CMS/SiteDB>
- Nice tool with information & links, e.g.:
site contacts, installed software, Grid info, monitoring
- Sites' pledges:
 - stored quarterly
 - should be kept up-to-date by the sites
 - meaning of some entries might be a bit confusing
 - SpecInt & jobs slots fuzzy by at least 35%
 - completion discipline not overwhelming
- Proposals:
 - evaluate/reconsider pledge entries and explanations
 - **use actual/expected status instead of pledges !**
 - define technical notations (Spec, # slots, brutto/netto TB)
 - SiteDB has to be the only source for pledge information
 - force sites to keep pledge information for current quarter up-to-date, require to fill in an estimate for NQ and NNQ

User Home Space Resources



- We anticipate **0.5 – 1 TB per local user** storage space at his/her „home T2“
- To be expected that the range will be much broader since some sites have substantial additional local resources integrated transparently
- For bookkeeping and e.g. default CRAB stage-out we will have a single user – Tier-2 home location mapping in SiteDB
- `/store/user/<userID>` capacity is not (fully) traced in CMS DB systems but most sites could provide this information by e.g. `cron du -sh /store/user/(*)`
is it sufficient to monitor this locally at the Tier-2 or do we need indiv. / T2-integral tracking centrally ?
- Users with an official **CERN affiliation** will be provided with home storage **resources at CERN** (Tier2 part of the CAF), about 50-100 (?) people
- Some countries do not operate a Tier-2 site and do not have a T3 with sufficient CMS Grid capability – I expect $O(5\%)$ of our users are affected
we **have to find „T2 friends“** (using established relations) or provide CERN CMS resources

Storage Resources con't



- **Break-down of disk resources** to central/groups/individuals/(local) at most sites not trivial but feasible to some extent:
 - official data in /store/data/ and /store/mc/ registered in global DB
 - central vs. detector/physics groups data tagg becomes available for new data transfers this month
 - data in /store/user/<userID> not/not all in global/local-scope DB – high level of dynamics; site has to use local tools to monitor and to protect space

CPU and Soft Resources



- Jobs at T1 & T2 sites report **SPEC** wall time **consumptions** centrally (EGEE APEL), **based on VO**; no VOMS granularity; some O(25%) arbitrariness ((slow) progress by WLCG task force) with static SPECint normalization factor
- **Rather complicated** to correctly adjust and monitor site's **CPU batch fair share** for numerous VOs, VOMS groups, VOMS roles, local submission, ...
- In principle we have the possibility to monitor **I/O throughput** which measures combined CPU/disk/network/job intelligence performance
- Several sophisticated **tools to monitor general site performance** (quantity & quality) are in operation
- Individual Tier-2 analysis **user support** is going to enter the game now, however no clear policy is defined yet

An Attempt of Conclusions

- Resource verification is necessary not because of mistrust but to find and react on **performance bottlenecks** and to **optimize utilization** of Tier-2 sites
- With **SiteDB** we have a powerful tool to provide us with Tier-2 (mainly static & integral) **resource information** to be used for many aspects of the planning
- We should **introduce strict policies** to have more benefit from SiteDB:
 - the actual CPU & disk status of the current quarter has to be kept up-to-date by the sites
 - in addition, the planning would benefit if estimates for forthcoming CPU and disk resources could be provided at least for the following two quarters
- To get more **detailed/individual information is sometimes hard**
 - we have to evaluate what granularity is necessary where and which tools exists or have to be developed
- We have nice tools and excellent central support to **monitor and classify site performance** and reliability; quality of user support will be less easy to measure
- The real challenge is still in front of us – massive parallel chaotic user analysis
- **How dynamic/flexible is/have to be our computing model for resource re-allocations ?**
- What counts at the end of the day is (hard) **resources * efficiency !**