

GridPP 19 Collaboration Meeting, University of Cumbria, Ambleside

GridPP Discussion Session 2 ~ F2F ~ 30 August 2007

Chair: Dave Britton

Panel A: Pete Clarke, Catalin Condurache, Santanu Das, Andrew Elwell, Robert Fay, Barney Garrett

DB put up his issues list which had resulted from the Tier-2 Review, and invited panel members to bring forward concerns.

- | | |
|--------------------------|-------------------------------|
| 1) SGE support | 13) local users |
| 2) SL4 readiness | 14) middleware support |
| 3) 64 bit-ness | 15) upgrade process |
| 4) GGUS performance | 16) Tier-2 hardware purchases |
| 5) Tier-2 service levels | 17) site tests |
| 6) Security policies | 18) job requirements |
| 7) distributed disk | 19) virtualisation |
| 8) monitoring/Nagios | 20) Experiment requirements |
| 9) installing VOs | 21) site problems |
| 10) inefficient jobs | 22) definition of 'available' |
| 11) software licences | 23) MoUs/SLAs |
| 12) travel funds | |

Regarding SGE support (1), it was reported that IC, UCL, QMUL, and Bristol were using SGE, yet SGE was not supported. It was asked why this had not been discussed at LCG Management level? Why was the UK the only one to be going this way? It was noted that an issues summary document was needed, like Condor. The more general problem related to being tied to one version of everything. There also weren't enough developers.

CC brought-up the issue of the Tier-1 review. Problems with VO installation (9) related to VO boxes and the MoU. There had been no feedback from the Experiments. There had been a problem with installing the TOTAL VO catalogue and a long delay had ensued. The general issue was that it was possibly too difficult to install VOs and it was noted that YAIM was too heavyweight for this. They had never removed a VO and it could break everything, yet there seemed to be no preferable solution. Had there been a PMB action on this (SB)? This would need checked. There was a question about the length of time people should remain in the GridPP VO - six months was suggested and agreed as reasonable.

RF raised the issue of site problems (21), which fell into different categories: 1. air conditioning/watercooling, 2. the ones which showed on the SAM tests, where often the problem was not known about at all until output was received from an Experiment member. RF noted that SL's tests don't show up everything. It was difficult to know what to monitor, and the biggest problem was not being notified when there was a problem. DB commented that the SAM tests, and SL's tests, identify certain things, but the real users running real jobs were paramount and Experiments should be encouraged to feedback more consistently; a user will simply try somewhere else rather than raising the issue. GS noted that the Grid paradigm was that users don't see sites. DB asserted that we need to engage local user communities in feedback discussion.

RF asked what happens when there aren't local users? How else where they to pick up on problems? GS noted that the DTeam has a role and if sites have an issue it can therefore be fixed

across the UK - most commonly for worker nodes failing - but we still lack knowledge-sharing. RF noted that a balance needed to be struck between investigating an issue and doing other things. DB pointed to local users and the DTeam - two routes to work out what was going wrong, but timelag was an issue. TD commented that Experiment-specific SAM tests should also help sites. GS countered that the SAM interface was difficult - to check a site he has to check around 23 webpages every day. JJ also noted that the tape was not monitored, 'green' lights can be there yet lots of problems can exist. AE advised that SL's tests were there for other VOs - the availability tests were in a non-particle format - however if you were blacklisted by VOs you generally weren't informed, the jobs just stop. DB concurred that there were loads of places you have to look for monitoring, and asked whether they couldn't be put together, but this would need a GridPP dashboard where information can be centrally available. Dave Colling agreed that a GridPP dashboard was a good idea, to include a list of sites that were blacklisted. BG noted that what was required was an at-a-glance status page showing availability.

Paul Miller commented that the Grid is supposed to both monitor and raise alarms at sites, Nagios plugins are good if they are running Nagios, but if they're not then there are problems. DB suggested a monitoring workshop; TD suggested GridView world. PM noted that LCG had three monitoring groups set up, and also used GridView and dashboard. TD advised that we could organise a workshop in the UK to map-out what should be there, and review monitoring design. DB noted that a workshop on how to use Nagios and Ganga would be better, and a GridPP dashboard would be ideal. PM noted that you can publish graphs out of Ganglia into Google, and reported that a monitoring workshop was planned after Hepsysman at Imperial, the proposed date was 31st October, TBC.

JJ raised another issue, definition of 'available' (22). In the Quarterly Reports the graphs are there but for existing storage, what does it mean that a resource is available? DB advised that available = usable. JJ noted that if the right interfaces are there then that is OK but it is complicated and difficult to penetrate - for LCG what does 'reserve' space mean? LCG decided that 'available' meant only disk when JJ was trying to calculate tape. In future this could get worse and we need agreed standards, plus a sufficient level of support within the community.

DB advised that 'availability' was an evolving recipe but we need to measure availability in a common way. JJ asked what is 'usable' space? He noted that the JANET network SLA had just been revised, they had discussed what 'available' means for both provider and user, including network accessibility and network availability etc. DB asked whether we are at the stage where we can agree a recipe for this across the project? It was suggested that two definitions were required: availability for disk and CPU. DB noted that the ultimate test was how much a site is used. Greig Cowan asked if this could be measured by using job exit codes? Dave Colling advised that even if a job exits successfully, it might not mean that it could run. There was a discussion of the SAM tests, SL's tests, and job environment - R-GMA was an issue.

JJ noted that the definition of 'available' changes according to updates at sites, therefore something that was running 6 months ago might not be possible now. DC noted that the DTeam meetings should address this - if issues are raised with the Tier-2 Co-ordinator, the meeting will react accordingly. JJ noted that a definition of available/usable was still required. SL proposed that there was no answer to this. DB agreed. JJ countered that if the project has a milestone and we are measured by that then we need a definition of 'available' and tests need to be run. DB advised that the definition changes over time, is discussed at the DTeam and run by consensus - we have a definition of 'available': it is something that evolves

GS raised the issue of Tier-2 service levels (5) - currently these are 2 hours and 95%. He asked why we have these targets at all for numbers and levels - sites might have broken setups and don't

run jobs. The objective of tests is to provide services that users need - for a working site this means, does the site run my job? If it shows red there could be many reasons for this - wading through the causes of 'red' is difficult every time they fail a test, but they need to know why they fail, and what can be done to prevent it happening again. Initial targets were 85% and sites were still failing that.

PC noted that when JANET says available, it is aggregate availability independent of sites. We need a figure for external consumption as well as a figure for individual site availability. DB advised that the 2 hours and 95% came from the LCG MoU, and it was undefined - they tried to inject commonality across the globe, and they are aspirational numbers. Within the two-hour response time, the Tier-2 must acknowledge they have a problem - this must be done within two hours during normal working time, whether this is 7.30 am - 3.00 pm or 9.00 am - 5.00 pm. The complication comes with holidays/sickness. We have distributed Tier-2s, and all the Tier-2 have to do is acknowledge a problem within two hours - this is considered possible most of the time. However, 95% ultimately applies to the Tier-2 site. It was noted that it was difficult to measure responses. SB advised that the information is there but we need a footprints database, and this is difficult. If the user doesn't respond to the site for two days then the site can't be blamed. RJ noted that the MoU was a 'gentlemen's agreement'. BG suggested that we need an automated response even if this is from an individual - this would give the user some assurance that the problem doesn't disappear. It was agreed that this was not too difficult to do, most of the time. Serious issues were brought-up at DTeam meetings. There are other routes to people if normal service is not available.

Panel B: Pete Gronbech, Jens Jensen, Dave Kant, Peter Love, Giuseppe Mazza, Paul Millar, Gidon Moont, Duncan Rand, Robin Tasker, Graeme Stewart

PG raised the issue of SL4 readiness (2). This was an issue at many sites - initially a couple of sites volunteered to have a look at it, and discovered problems from the experimental end. This was still an issue and we were almost at SL5. TD advised that this had been raised with the experiments, the CMS issue was a concern. CMS sites aren't generally running DPM and there was a lack of awareness of DPM. DB noted that we have metric-driven objects, therefore need a mechanism for penalising. Some sites have moved their nodes to SL4 and they are not being used, therefore they are reluctant to move any more in case the VOs aren't going to use them. This has resulted in usage figures being lower - sites are being penalised for doing what they were asked to do.

Raja Nandakumar commented that from the experiments point of view, when a job fails it can be due to SL4 but not exclusively - it is only one of several issues. DB noted that he understood the sites' situation, and the message to the experiments was that it would be preferable if they were to test on the PPS beforehand. Sites should not be penalised for doing pioneering work. DB advised that the pledge numbers were being revisited at present by STFC at high-level. TD noted that we would be going through the whole exercise again with a new regime.

Duncan Rand raised the issue of disk usage/distributed disk (7). At one stage they had a full disk and couldn't run CMS jobs - whose responsibility was that? It was a VO problem? Yes - if they had filled-up the disk then they couldn't use it. It was suggested that we consider a system of lifetimes - if a file is not accessed for 1 year then it should be deleted. DR noted that if old files are there, it reduces the value of the disk. They tried partitioning the data but this was problematic and DPM wasn't flexible enough at the time.

TD noted that the key point was when can you delete in order to be effective overall - a draft policy should be defined. SB asked why should a VO not be allowed to store data? Storage is storage, and if the VO chooses to put data there then why shouldn't they? It was suggested that the VO should state the kind of storage they require. RJ noted that storage was under the experiments' management - storage was supposed to be durable, files are pinned, a large chunk was for users and

should be explicitly pinned by VOs, and there should be another area marked-out as volatile. Tools were needed to enforce such a policy, but we need to have the ability to pin data.

PG noted that lots of sites have full disk. Sites were now buying a 2nd tranche of disk - they can partition per VO and this is less likely to have problems in the future than it was in the past. RJ noted that there will always be a problem, especially with students' storage. PG noted that if more storage was required, quotas should be put in place. DB advised that some partition has always been used for staged disk as a minimum so that CPUs can be used. GC noted that this should come with SRM2.2 ok and space can be reserved with a lifetime on it. SB noted that this would then need GLUE schema to publish it. RJ noted that in ATLAS, contacts are well defined.

DR advised that with respect to CPU efficiencies, the dashboard tells if a site is successful - a comparison of site and users would be useful, as long as jobs are integrated then you can build-up the idea of users whose jobs are failing and give an overview of the VO. The structure is there and could be used more.

GM raised the issue of communication with the external world, and gaining new users. He suggested that the procedure be easier for new customers to engage with GridPP. SP and NO could point new users to the London Tier-2 Co-ordinator for analysis, and he in turn could point to the Tier-2s generally. At the Tier-2 level they have a Co-ordinator and a VO that a new user can use to get in the Grid. Regarding communication between us as Grid people, and people doing physics, is it possible to have templates of JDL that do standard jobs? Some tasks are standard and a wiki page with templates would be useful - new users could look on there and see an example.

DB reported that we had tried to compile information on a 'people's page' on the GridPP website, showing GridPP PMB people and roles, there was also the ATLAS workbook. TD noted that the list of 'ranked users' also functions as a dissemination tool, and we might publish that kind of information too.

The issue of virtualisation was raised (19). DB noted that this was on the list because sites had asked what GridPP's plans were for virtualisation. Dave Colling noted that for compatibility, Enterprise3 had to be running in order to be compatible with the current release of gLite. It was asked whether anyone was using virtualisation? It was noted that it might be used if someone was using the PPS to RAL and testing UI, but it was not planned to be used for the CE or WMS. SL noted that STFC had organised a workshop on virtualisation for 21st Sept. The people in OpenLab have a new tool for building virtual machines and have published a link. DC noted that this was a growing area, running virtual machines on servers, and it can help to buy RAM.

The next issue raised was that of security policy (6). The new Security Officer, Mingchao Ma, asked how many policies existed and where can they be found? He noted that security was the responsibility of everyone and all should read the appropriate security policy as users. MM noted that some information was available on the website but that he intended to add a page with links to the up-to-date policies. SL pointed out that not all of the policies had been approved by either GridPP or the Tier-2 Board. MM noted that the Joint Policy Group were dealing with security at present. TD advised that all 8 policies need to be discussed and re-approved. It was agreed that MM send documents to SL for discussion at the Tier-2 Board.

DB noted that all users should be aware of security and acceptable use policy. The site admin response to security issues should be: do they have the right procedures and are the sites generally aware of them? MM reported that all sites would be receiving a questionnaire at end Nov/beginning Dec, that deals with current issues. DB drew the meeting to a close and thanked the panels for their contributions.

Action: Mingchao Ma to send extant security policies to SL for discussion at the Tier-2 Board. These need to be re-approved prior to being available as links for users on the website.