GRIDPP
UK
Particle
Physics

Fabric/Software WG

LCG Workshop 4th November 2004

Estimated Traversal Time

- Summary
  - "If I submit a job now how long will it take to run once it lands on the CE."
- Currently there is one published ETT per queue so this is meaningless when fairsharing is in place.
- The solution of a queue per VO while easy to implement now but is not a good or long term solution.
- The algorithm was questioned and other proposals were given.

ETT per VO

- We can work towards ETT per VOs but not users at this time.
- The following steps needed.
  1. VO specific ETT published within each queue.
  2. Any new algorithm must be deployable, ie portable across PBS, Maui, LSF, SGE, BQS, Condor,….
  3. Resource broker must be able to use these values.
     - It was unsure if the current resource broker can now rank on any value presented to it.
- This is something site admins wanted soon, the ability of one VO to block another is currently in the system.

GRIDPP UK Particle Physics

Algorithm

- A few solutions suggested.
  - Using Maui's test mode to predict what will happen to the next job.
    - Not feasible since it takes minutes to calculate.
  - Using statistical based method from NIKHEF.
  - Keep the current system.
    - ((walltime/2) * jobs queued)/(total number of CPUs).

GRIDPP UK Particle Physics

Changing the Schema for ETT

- We must change the schema anyway for adding the ETT for each VO.
- Should we publish the algorithm used to calculate ETT.
  - Advantages , multiple algorithms can be tested and used.
  - Disadvantage , just more complexity and confusion and people were sceptical of its value.
- No conclusion was reached and it should be discussed with Glue schema maintainers.

Avoiding ETT Completely

- Suggestion that we could publish
  - "Willingness to run jobs for a VO"
- Did not reach any conclusions since difficult to give a metric for this. Would a site ever not want more jobs? More queued jobs improves efficiency.
- One outcome in this area was RB should respect max number of jobs queued – just a bug fix really.

Submission with Job Attributes

- There were definite use cases for this.
  - Some sites, Scotgrid, Lyon, CERN and RAL would rather run with one routing queue into the site.
    - The alternatives is queues for every permutation of resource.
  - With federated grids the RB may just be submitting to another grid, eg LCG->Grid Canada.
    - The encapsulated grid must have these values.
- It was commented the RB already passes over multiple CPU requirements.

Installation of App SW on WNs.

- Current solution.
  - An extra VO exists of software managers exists who are able to write to NFS mounted area.
- Issues with this.
  - With 4000 WNs this expected to be problem if millions of jobs are reading the same files.
- The opposite solution to install everything on every WN.
  - Not being used anywhere and difficult to manage by the VOs, tank and spark may be a solution.

GRIDPP UK Particle Physics

Software Managers

- Software managers also run jobs.
- Sites can easily limit managers to one job and maximum priority for their benefit.
- But this requires end users to get a second certificate so they can also runs jobs.
- It was decided we wanted to push the managers this way until some role based solution is deployed, i.e. VOMS and LCMAPS.

GRIDPP UK Particle Physics

Some Requirements From Sites

- This appears to be different at most sites.
- Large sites wanted full control on when software would and would not be mounted vs. copied to the WNs.
- Smaller sites were happy with the NFS solution.
- The current environment variable solution for locating software was good.
  - A possibility of different environment variables (locations) for installing and using software was suggested.
  - This would allow different protocols to be used.
  - Can be tested within one site.

GRIDPP UK Particle Physics

Tank and Spark

- Some concern that sites not consulted and created some immediate concerns.
  - CE may not have space.
  - Space management of WNs.
- Tank and spark is just one option of many for this that can be used under the environment variable.
- There will be lots of solutions within the fabric for making software available.