Data Staging Developments

David Cameron
University of Oslo
Reminder of Purpose

Problems with the down/uploaders:

- Per-job
- Waiting blocks other jobs
- Cached files blocked by others
- No priority system
- Tightly-coupled transfer preparation and transfer

Solution:

- Single queue of files with priorities
- Disconnect physical transfer
- Asynchronous requests to e.g. SRM
Differences

- No waiting to go to PREPARING and FINISHING
- Priorities in xrsl (priority=“80”)
- gm-jobs -s

<table>
<thead>
<tr>
<th>Preparing/Pending files</th>
<th>Transfer share</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/159</td>
<td>atlas:pilot</td>
</tr>
<tr>
<td>3/456</td>
<td>atlas:production</td>
</tr>
<tr>
<td>1/0</td>
<td>ops:null</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Finishing/Pending files</th>
<th>Transfer share</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/48</td>
<td>atlas:pilot</td>
</tr>
<tr>
<td>3/12</td>
<td>atlas:production</td>
</tr>
</tbody>
</table>
Current Status

- In svn trunk since Feb 2011
- Enabled by newdatastaging=yes
- Main tester is Andrej with pikolit.ijs.si
  - Running real ATLAS jobs
  - A lot of close collaboration to iron out bugs and performance issues
    - Seg faults, queue logic, etc..
  - Conclusion: new system much better at using bandwidth
    - Cached files processed instantly
- Other sites encouraged to try
Monitoring of Data Staging States

Gangliarc (more in a later talk)
Current Issues

• All data staging runs as threads of arched
  – A seg fault kills the entire process
• Memory leaks (bug 2573)
  – Main source of leaks (OpenSSL) was fixed
  – Some unavoidable e.g. setenv
• Queueing logic is hard to tune right
• P>>T not possible, now P=T
Plans

- Multi-host data staging
  - To replace multi-GM
  - Split between hosts is done at lowest level
  - Remote hosts run HED service which accepts transfer requests
  - Prototype available now

http://wiki.nordugrid.org/index.php/Data_Staging/Multi-host
Plans

- Better configuration
- Investigating algorithms for Scheduler logic
- Quantifying gain from new system (CHEP paper)
- Gaining more experience from more sites
- Monitoring requirements

- How does this fit to EMI?
- When to turn on by default?
Thank you!

EMI is partially funded by the European Commission under Grant Agreement RI-261611