ATLAS Production System Monitoring

John Kennedy
LMU München
CHEP 07
Victoria BC

06/09/2007
Overview

- Production System – quick intro
- Monitoring Goals
- Design
- Functionality - examples
- Usage
- Conclusion + outlook
Prod-sys

• Central Database with
  • Tasks – Groups of jobs
  • Jobs – job definitions
  • Attempts – attempt info for jobs

• Supervisor(Eowyn)
  • Retrieves jobs from DB
  • Passes to GRID through Executor
  • Manages job attempt info

• Executor(per grid flavour)
  • Create wrappers etc for submission
  • Submit job
  • Manage Job + interpret output (failure reasons etc)
Prod-sys

O(50K) Jobs per Day

Jobs
Prod DB

Supervisor
Eowyn
Exe

Data
DDM

LCG

GridKa

Clouds
Production Database

- **Hierarchy of tables**
  - **Task Table** - 11,018
    - Contains task definition data
  - **JobDef Table** - 6,619,954
    - Contains definition of each job within a task
  - **JobExe Table** - 9,457,938
    - Contains record of each attempt at a specific job
Monitoring Goals

- Provide easy access to monitoring and accounting info
- **Avoid load on central database**
- Provide Views for
  - Recent jobs to allow problem spotting
  - Configurable time windows for # jobs/walltime and errors
  - Higher level views to give an understanding of job states and available jobs
- **Provide High and Low Level Views!**
Design

- Monitoring tables defined in specific monitoring account
- Oracle Stored procedures used to fill monitoring tables from Prodsys tables
- Monitoring tables used in combination with smaller proddb tables to provide web pages
- PHP pages used to interface to tables
- Cron jobs to give periodic views
Design

All Job Execution Records

Production Database

06/09/2007 - John Kennedy
Design

24hr Window of Execution Records

Summary Table
Filled from snapshot

Production Database

06/09/2007 - John Kennedy
Design

24hr Window of Execution Records

All Job Execution Records

10M

Snapshot filled with Last 24hrs

Monitoring Database

Summary Table
Filled from snapshot

Production Database

06/09/2007 - John Kennedy
Last 24hr View

• Snapshot table in the monitoring database contains 24hr sliding window of the job execution records.
• Pages generated from this table to give views of jobs/errors grouped by
  • Task
  • Site
  • Executor
• This 24hr view is aimed towards people on shift + spotting recent problems
Last 24hr View

Views group info based on Tasks – Sites - Executors
Aids problem solving
Last 24hr View

Views group info based on Tasks – Sites - Executors
Aids problem solving

Click to get breakdown

Details for sites giving pend/run/fin/fail jobs and colour coded efficiency
spot sinners!

06/09/2007 - John Kennedy
# Last 24hr View

<table>
<thead>
<tr>
<th>#</th>
<th>Site</th>
<th>PENDING</th>
<th>RUNNING</th>
<th>Finished</th>
<th>Failed</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>grid.uk.ac.uk</td>
<td>1</td>
<td>27</td>
<td>67</td>
<td>0</td>
<td>94</td>
</tr>
<tr>
<td>3</td>
<td>lcg-compute.hep.unimelb.edu.au</td>
<td>0</td>
<td>20</td>
<td>59</td>
<td>2</td>
<td>97</td>
</tr>
<tr>
<td>4</td>
<td>ce001.grid.unifr.ch</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>bigmac-log.ee.physics.utoronto.ca</td>
<td>1</td>
<td>74</td>
<td>205</td>
<td>15</td>
<td>93</td>
</tr>
<tr>
<td>6</td>
<td>lcg.ca.lps.umontreal.ca</td>
<td>0</td>
<td>10</td>
<td>46</td>
<td>95</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>lcg-ce.erdw.uveca</td>
<td>2</td>
<td>60</td>
<td>3</td>
<td>285</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>lcgpal1.triumf.ca</td>
<td>18</td>
<td>107</td>
<td>291</td>
<td>210</td>
<td>58</td>
</tr>
<tr>
<td>9</td>
<td>snowpatch.hep.sfu.ca</td>
<td>0</td>
<td>13</td>
<td>17</td>
<td>110</td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>ce101.cern.ch</td>
<td>0</td>
<td>68</td>
<td>111</td>
<td>30</td>
<td>79</td>
</tr>
<tr>
<td>11</td>
<td>ce102.cern.ch</td>
<td>1</td>
<td>34</td>
<td>89</td>
<td>0</td>
<td>92</td>
</tr>
<tr>
<td>12</td>
<td>ce107.cern.ch</td>
<td>0</td>
<td>26</td>
<td>93</td>
<td>10</td>
<td>90</td>
</tr>
</tbody>
</table>
## Last 24hr View

### Efficiency Breakdown

<table>
<thead>
<tr>
<th>Site</th>
<th>PENDING</th>
<th>RUNNING</th>
<th>Finished</th>
<th>Failed</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>grid-ubk.ac.at</td>
<td>1</td>
<td>27</td>
<td>67</td>
<td>0</td>
<td>94</td>
</tr>
<tr>
<td>lgc-compute.hpc.unimelb.edu.au</td>
<td>1</td>
<td>20</td>
<td>59</td>
<td>2</td>
<td>97</td>
</tr>
<tr>
<td>ces011.grid.unifr.ch</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>lgmc-ccp-nh.physics.utsa.edu</td>
<td>1</td>
<td>74</td>
<td>206</td>
<td>15</td>
<td>63</td>
</tr>
<tr>
<td>lgc-ccp.umontreal.ca</td>
<td>1</td>
<td>10</td>
<td>48</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>lgc-ccp.ubc.ca</td>
<td>1</td>
<td>60</td>
<td>3</td>
<td>285</td>
<td>3</td>
</tr>
<tr>
<td>lgcpr.famnité.ch</td>
<td>18</td>
<td>107</td>
<td>291</td>
<td>210</td>
<td>58</td>
</tr>
<tr>
<td>snowpatch.hpc.sfu.ca</td>
<td>1</td>
<td>13</td>
<td>17</td>
<td>110</td>
<td>53</td>
</tr>
</tbody>
</table>

Click on efficiency to get breakdown of errors
### Last 24hr View

<table>
<thead>
<tr>
<th>#</th>
<th>Site</th>
<th>PENDING</th>
<th>RUNNING</th>
<th>Finished</th>
<th>Failed</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>grid.nih.ec.at</td>
<td>1</td>
<td>27</td>
<td>87</td>
<td>0</td>
<td>94</td>
</tr>
<tr>
<td>3</td>
<td>kgcrp.computer.hpc.anl.edu.au</td>
<td>1</td>
<td>20</td>
<td>59</td>
<td>2</td>
<td>97</td>
</tr>
<tr>
<td>4</td>
<td>ces001.gwu.edu</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>bigmac-log.ee.physics.utoronto.ca</td>
<td>1</td>
<td>74</td>
<td>206</td>
<td>15</td>
<td>90</td>
</tr>
<tr>
<td>6</td>
<td>log-cc/cps.montreal.ca</td>
<td>0</td>
<td>10</td>
<td>46</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>log-cc/ref.uvic.ca</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>logppot1.triumf.ca</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>snowpatch.hpc.stlu.ca</td>
<td>0</td>
<td>13</td>
<td>17</td>
<td>110</td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>ces101.uncw.edu</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ces102.uncw.edu</td>
<td>1</td>
<td>34</td>
<td>69</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>ces107.uncw.edu</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Click on efficiency to get breakdown of errors**

**Click # to get logfiles (egee)**

**Click on error to get detailed info job execution info**
# Tasks Status

<table>
<thead>
<tr>
<th>Data Set View</th>
<th>How to Find Datasets</th>
<th>Help</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset Status: All</td>
<td>Grid: All</td>
<td></td>
</tr>
<tr>
<td>Match Dataset</td>
<td>Dataset name: PythonTH160WW</td>
<td></td>
</tr>
</tbody>
</table>

Select Task by Grid State and Name
**Tasks Status**

<table>
<thead>
<tr>
<th>Dataset/Task</th>
<th>TaskID</th>
<th>Done</th>
<th>Running</th>
<th>To be done</th>
<th>Wait/Inp</th>
<th>Aborted</th>
<th>Completed(%)</th>
<th>Grid</th>
<th>Atlas ver</th>
<th>Tier</th>
</tr>
</thead>
<tbody>
<tr>
<td>misall_mcel2.005345.PythiaH160/WW33v4j_neglep_H1T2_digt.v12003107.1.task</td>
<td>4696</td>
<td>99</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>99.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>ES</td>
</tr>
<tr>
<td>misall_mcel2.005347.PythiaH160/WW33v4j_neglep_H2T1_digt.v12003107.1.task</td>
<td>4698</td>
<td>99</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>99.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>UK</td>
</tr>
<tr>
<td>trigg1_misall_mcel2.005345.PythiaH160/WW33v4j_neglep_H1T2_recon.v12000601.task</td>
<td>6448</td>
<td>19</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>95.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.06</td>
<td>ES</td>
</tr>
<tr>
<td>trigg1_misall_mcel2.005346.PythiaH160/WW33v4j_poslep_H2T1_recon.v12000601.task</td>
<td>6449</td>
<td>18</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>90.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.06</td>
<td>DE</td>
</tr>
<tr>
<td>trigg1_misall_mcel2_V1.005346.PythiaH160/WW33v4j_poslep_H2T1_recon.v1200601.task</td>
<td>7855</td>
<td>19</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>95.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.06</td>
<td>IT</td>
</tr>
<tr>
<td>trigg1_pileos05_misall_mcel2.005346.PythiaH160/WW33v4j_poslep_H2T1_recon.v1200605.task</td>
<td>9861</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>95.45</td>
<td>LCG-DQ</td>
<td>Atlas-12.06</td>
<td>IT</td>
</tr>
<tr>
<td>pdc1sf05_misall_mcel2.005346.PythiaH160/WW33v4j_poslep_H2T1_digt.v1200605.task</td>
<td>9860</td>
<td>32</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>91.43</td>
<td>LCG-DQ</td>
<td>Atlas-12.06</td>
<td>IT</td>
</tr>
<tr>
<td>misall_mcel2.005345.PythiaH160/WW33v4j_neglep_H1T2_digt.v12003103.task</td>
<td>7838</td>
<td>99</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>99.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>ES</td>
</tr>
<tr>
<td>trigg1_misall_mcel2_V1.005345.PythiaH160/WW33v4j_neglep_H1T2_recon.v12000501.task</td>
<td>9740</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>86.67</td>
<td>LCG-DQ</td>
<td>Atlas-12.06</td>
<td>ES</td>
</tr>
<tr>
<td>trigg1_misall_mcel2_V1.005346.PythiaH160/WW33v4j_poslep_H2T1_recon.v12000601.task</td>
<td>9741</td>
<td>18</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>94.74</td>
<td>LCG-DQ</td>
<td>Atlas-12.06</td>
<td>IT</td>
</tr>
<tr>
<td>misall_mcel2.005345.PythiaH160/WW33v4j_neglep_H1T2_recon.v12000601.task</td>
<td>12144</td>
<td>145</td>
<td>55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>72.50</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>ES</td>
</tr>
<tr>
<td>trigg1_misall_mcel2_V1.005345.PythiaH160/WW33v4j_poslep_H2T1_recon.v12000501.task</td>
<td>12145</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.06</td>
<td>None</td>
</tr>
<tr>
<td>misall_mcel2.005346.PythiaH160/WW33v4j_poslep_H2T1_digt.v12003107.1.task</td>
<td>4697</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>DE</td>
</tr>
<tr>
<td>trigg1_misall_mcel2.005344.PythiaH160/WW33v4j_poslep_H2T1_recon.v12000601.task</td>
<td>6362</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>100.00</td>
<td>OSG</td>
<td>Atlas-12.06</td>
<td>None</td>
</tr>
<tr>
<td>mcel2.005344.PythiaH160/WW33v4j_poslep_H2T1_recon.v12003102.task</td>
<td>3706</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>100.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>DE</td>
</tr>
<tr>
<td>mcel2.005345.PythiaH160/WW33v4j_neglep_H1T2_recon.v12003102.task</td>
<td>3707</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>ES</td>
</tr>
<tr>
<td>misall_mcel2.005345.PythiaH160/WW33v4j_neglep_H1T2_digt.v12003107.1.task</td>
<td>6481</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>ES</td>
</tr>
<tr>
<td>misall_mcel2.005346.PythiaH160/WW33v4j_poslep_H2T1_digt.v12003107.1.task</td>
<td>6482</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>IT</td>
</tr>
<tr>
<td>misall_mcel2.005347.PythiaH160/WW33v4j_poslep_H2T1_recon.v12000601.task</td>
<td>6483</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>UK</td>
</tr>
<tr>
<td>mcel2.005346.PythiaH160/WW33v4j_poslep_H2T1_recon.v12003105.task</td>
<td>4228</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>IT</td>
</tr>
<tr>
<td>mcel2.005347.PythiaH160/WW33v4j_neglep_H2T1_recon.v12003105.task</td>
<td>4229</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.31</td>
<td>UK</td>
</tr>
<tr>
<td>misall_mcel2.005344.PythiaH160/WW33v4j_neglep_H2T1_recon.v12003108.task</td>
<td>5453</td>
<td>200</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>OSG</td>
<td>Atlas-12.06</td>
<td>None</td>
</tr>
<tr>
<td>trigg1_misall_mcel2.005347.PythiaH160/WW33v4j_neglep_H2T1_recon.v12000601.task</td>
<td>6450</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.06</td>
<td>UK</td>
</tr>
<tr>
<td>mcel2.005344.PythiaH160/WW33v4j_poslep_H2T1_recon.v12000401.task</td>
<td>4399</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00</td>
<td>OSG</td>
<td>Atlas-12.04</td>
<td>None</td>
</tr>
<tr>
<td>trigg1_misall_mcel2_V1.005347.PythiaH160/WW33v4j_poslep_H2T1_recon.v12000501.task</td>
<td>7856</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>100.00</td>
<td>LCG-DQ</td>
<td>Atlas-12.06</td>
<td>ES</td>
</tr>
</tbody>
</table>
OverViews

- Allow Views to be generated over defined time windows
- Uses Summary table
  - Ensures quick response
  - Low load on database
- Provide Views of
  - Finish/Failed Jobs/Walltime
  - Error reports, pie charts
  - Allocated Tasks/jobs – ensure grids have jobs to run
  - Allow split by grid flavour and sites
OverViews

Release Validation
Production Operation

High Failure rates during Validation Phase.
Savannah bug reports etc from operation teams

Successful jobs Oct-Dec 06
All Grids perform better after Validation Phase

06/09/2007 - John Kennedy
OverViews

Looking at Jobs to highlight problems and reduce load

Looking at Walltime to find where/why we waste resources

06/09/2007 - John Kennedy
OverViews

Views of available jobs split by SW version
Grid and State
- allows us to ensure we have jobs that can run

Definitions of Error Codes/Acronyms with Meaning
- helps shifters to interpret problems

06/09/2007 - John Kennedy
EGEE Shift System

• During the last 12 months an EGEE shift system has been put in place

• People responsible for
  • Assigning Tasks to Tiers
  • Monitoring Jobs
  • Monitoring Data movement
  • Ensuring Tasks are progressing
  • Reporting problems to sites/physicists

• Strong users of Monitoring + lots of feedback

• Pages Developed to help shifters
EGEE Shift Pages

Identify sites with ATLAS SW installations problems

Monitor Running Tasks
Identify old Tasks which need action
Identify nearly completed Tasks

06/09/2007 - John Kennedy
Monitoring Usage

- **Monitor the Monitoring**
- Monitor usage of main web pages
  - Total # users
  - Busy pages
- Log SQL query times of Cron Jobs
  - Spot increasing load and problems when crons run
- Monitoring account which can be tracked
Monitoring Usage

- Monitor the Monitoring
- Monitor usage of main web pages
  - Total # users
  - Busy pages
- Log SQL query times of Cron Jobs
  - Spot increasing load and problems when crons run
- Monitoring account which can be tracked

Currently O(40-50)
Individual users per day
Conclusion

- The model of using cooked down monitoring tables has proved successful
  - Quicker queries
  - Reduced load on database
- Both High and low level views have been implemented allowing for
  - Problem Identification
  - Accounting
- EGEE shift pages are well developed due to good communication with users – I am part of the shift team!
- Much more functionality provided than originally expected
- From a personal point of view - It's been fun and it's useful
Outlook

• Much has been learned during the project
  • By developers
    • Snapshot/summary table use
    • What users want
  • By users
    • What's useful
    • What's not
    • What would be nice
• Prod-sys monitoring being integrated into ARDA dashboard
  • Functionality based on experience gained here
  • Database snapshot tables etc will remain
  • PHP replaced by new framework
• Migration has started – expected to be complete by start of 2008
THANKS!

Thanks for your time and attention!