The PanDA Distributed Production and Analysis System

Torre Wenaus
Brookhaven National Laboratory, USA

ISGC 2008
Taipei, Taiwan
April 9, 2008
Panda Overview

Workload management system for Production AND Distributed Analysis
Core team @ BNL, UT Arlington/CERN. 4 developers, 1 full time

- Launched 8/05 by US ATLAS to achieve scalable data-driven WMS
- Designed for analysis as well as production
- Insulate users from distributed computing complexity
  - Lower entry threshold
- US ATLAS production since end ‘05
- Analysis since Spring ’06
- ATLAS-wide production since early ‘08
- OSG WMS program since 9/06
  - VO-neutrality
  - Condor integration
  - Security enhancements
PanDA Attributes

• Pilots for ‘just in time’ workload management
  • Efficient ‘CPU harvesting’ prior to workload job release
  • Insulation from grid submission latencies, failure modes and inhomogeneities
• Tight integration with data management and data flow
  • Designed/developed in concert with the ATLAS DDM system
• Highly automated, extensive monitoring, low ops manpower
• Based on well proven, highly scalable, robust web technologies
• Can use any job submission service (CondorG, local batch, EGEE, Condor glide-ins, ...) to deliver pilots
• Global central job queue and management
• Fast, fully controllable brokerage from the job queue
  • Based on data locality, resource availability, priority, quotas, ...
• Supports multiple system instances for regional partitioning
Job Submission

- Clients submit jobs to PANDA server with
  - Grid proxy
    - generated using LCG-UI, VDT, NorduGrid-UI, stand-alone Globus
  - HTTP client
    - e.g., curl or urllib
- Communication protocol is independent from Grid-middleware
  - Jobs can be submitted from any Grid flavor (OSG, EGEE, NorduGrid)

End-user

Scheduler

Worker Nodes

T. Maeno
PanDA Server

- Server manages all job information centrally
  - Priority control
  - Resource allocation
  - Job scheduling

- LAMP stack
  - RHEL3 / SLC4
  - Apache 2.0.59
  - MySQL 5.0.27 - InnoDB
  - Python 2.4.4

- Multi-processing
  - Apache child-process

- Multi-threading
  - Python threading

T. Maeno
Pilots are prescheduled to batch system and grid sites, and run jobs as soon as CPU becomes available.

Use resources efficiently:
- Exit immediately if no job available
- Rate is regulated according to workload

Multi-tasking:
- Job execution
- Status reporting
- Zombie detection
- Error recovery
Pilot Scheduling Management

Scheduling runs with a grid cert appropriate to and managed by the region and usage mode.
PanDA Data Flow

Data management uses the ATLAS DDM system Don Quixote 2 (DQ2)

Torre Wenaus, BNL
Analysis with PanDA: pathena

User

PANDA

buildJob x 1
runAthena x N

DDM

source.tgz

outputs

outputs

libraries.tgz

Storage

inputs

inputs

outputs

outputs

Running the ATLAS software:
Locally: athena <job opts>
PanDA: pathena --inDS --outDS <job opts>

Outputs can be sent to xrootd/PROOF farm, directly accessible for PROOF analysis
Production vs. Analysis

- Use the same PanDA and underlying grid infrastructure
  - Same software, monitoring system, facilities
  - No duplicated operations manpower
  - Analysis benefits from problems found/resolved by shifters
- Separate computing resources
  - Different queues mapping to different clusters
  - Relative usages can be controlled without competition
- Different policies for data management
  - Data delivery to sites is in production workflow, but not in analysis
  - Analysis jobs go to where the data currently resides
    - User can specify the site, or let the system decide
  - Analysis outputs reside on the local SE until retrieved by user
- PanDA user-level accounting/quota system supports analysis; being extended to production (at physics working group and project levels)
  - Priorities also supported; important basis for brokerage decisions
PanDA in Operation

• 2006 - 2007: PanDA responsible for US ATLAS production and analysis, performed extremely well
  • ~2500 CPUs, ~35% of total ATLAS production, well in excess of US share of ~20% through efficient resource use
  • Event gen, simu, digitization, pile-up, mixing, reco
  • Many hundreds of different physics processes simulated
  • ~350 analysis users, 125 with >1k jobs, >2M total jobs
  • 1 shifter, ~.1 FTE to manage PanDA ops itself

• 2008: PanDA responsible for ATLAS-wide production
  • Roll-out smooth, and almost complete; no scaling issues seen
    • But scaling studies ongoing, on web server and DB levels, to keep ahead of the growth curve
    • Expansion includes a second full PanDA instance at CERN
      • For ‘human’ more than technical reasons
      • But a valuable further scaling knob for the future
US ATLAS PanDA Production ‘06–’07

CSC Production - Jobs finished in 2007

Distribution among US Tier 1 (BNL) and 5 Tier 2s

Torre Wenaus, BNL

K. De
Job Error Rates Apr-May 07

Errors from all sources. PanDA errors mainly application errors
• Roll-out across ATLAS began fall ‘07
• Today PanDA production spans all 10 ATLAS regions
  • almost 200 sites/queues
  • 10k-18k concurrent jobs
• Nordic grid (NDGF) just coming online
• Italy ramping up; served by distinct (CERN) PanDA system instance
Panda Production Operations Dashboard

Panda monitor

Shift log Wiki

Jobs - search
Recent running, activated, waiting, assigned, defined, finished, failed jobs
Select analysis, prod, install, test jobs

Quick search
Job
Dataset
Task req
Task status
File

Summaries
Blocks: days
Errors: days
Nodes: days
Daily usage

Tasks - search
Generic Task Req
EvGen Task Req
CTBs, sim Task Req
Task list
Task browser
Bug Report

Datasets - search
Dataset browser
Aborted MC datasets
Panda subscriptions

Datasets
Distribution
DDM Req
Req list
AODs
EVENTs
RDOs

Conditions DS
DB Releases
Validation Samples
Functional Tests
CosmicRuns
FOR Datasets
GECO Datasets

Sites - see all

Servers: BNL:OK BNLdev:OK CERN:OK Logger:OK


Bamboo submissions, status over last 12 hours

Jobs updated >12 hrs ago: activated:6526 running:0
Jobs updated >36 hrs ago: transferring:366

Ganglia World Wide Summary

Space available at sites:

<table>
<thead>
<tr>
<th>Site</th>
<th>GB</th>
<th>As of (UTC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANALY_MWT2-condor</td>
<td>53743</td>
<td>04-08 22:01</td>
</tr>
<tr>
<td>ANALY_SLAC-Ist</td>
<td>25241</td>
<td>04-08 21:10</td>
</tr>
<tr>
<td>BU_ATLAS_Tier2</td>
<td>23777</td>
<td>04-08 21:28</td>
</tr>
<tr>
<td>BU_ATLAS_Tier20</td>
<td>23774</td>
<td>04-08 21:56</td>
</tr>
<tr>
<td>IU_OSG</td>
<td>58622</td>
<td>04-08 22:01</td>
</tr>
<tr>
<td>MWT2 IU</td>
<td>58622</td>
<td>04-08 22:02</td>
</tr>
<tr>
<td>MWT2 UC</td>
<td>53743</td>
<td>04-08 22:02</td>
</tr>
<tr>
<td>OU_OCHEP_SWT2-condor</td>
<td>5839</td>
<td>04-08 21:47</td>
</tr>
<tr>
<td>SLACRFD-Ist</td>
<td>25240</td>
<td>04-08 21:29</td>
</tr>
<tr>
<td>UC_ATLAS_MWT2</td>
<td>53744</td>
<td>04-08 21:58</td>
</tr>
<tr>
<td>UTAT_SWT2</td>
<td>2830</td>
<td>04-08 22:01</td>
</tr>
</tbody>
</table>

Other SEs reporting in last 3 days

Production job summary, last 12 hours (Details: errors, nodes)

<table>
<thead>
<tr>
<th>Cloud Information</th>
<th>Nodes</th>
<th>Jobs</th>
<th>Latest</th>
<th>Pilots (3hrs)</th>
<th>defined</th>
<th>assigned</th>
<th>waiting</th>
<th>activated</th>
<th>running</th>
<th>holding</th>
<th>transferring</th>
<th>finished</th>
<th>failed</th>
<th>tot</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Production</td>
<td>6541</td>
<td>67253</td>
<td>04:08 18:03</td>
<td>4834</td>
<td>0 / 0</td>
<td>1829 / 0</td>
<td>1 / 0</td>
<td>22564 / 0</td>
<td>10960 / 0</td>
<td>6047 / 0</td>
<td>8348 / 366</td>
<td>15486 / 0</td>
<td>1617 / 0</td>
<td>9%</td>
<td>0%</td>
</tr>
<tr>
<td>CA</td>
<td>779</td>
<td>8021</td>
<td>04:08 18:03</td>
<td>399</td>
<td>0</td>
<td>402</td>
<td>0</td>
<td>3227</td>
<td>1559</td>
<td>147</td>
<td>800 / 0</td>
<td>1603</td>
<td>283</td>
<td>15%</td>
<td>0%</td>
</tr>
<tr>
<td>DE</td>
<td>1103</td>
<td>13400</td>
<td>04:08 18:03</td>
<td>634</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>5360</td>
<td>2464</td>
<td>64</td>
<td>2924 / 0</td>
<td>2531</td>
<td>638</td>
<td>20%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Torre Wenaus, BNL
# Production Snapshot Apr 8 ’08

## Production Job Summary, Last 12 Hours

<table>
<thead>
<tr>
<th>Cloud Information</th>
<th>Nodes</th>
<th>Jobs</th>
<th>Latest</th>
<th>Pilots (3hrs)</th>
<th>defined</th>
<th>assigned</th>
<th>waiting</th>
<th>activated</th>
<th>running</th>
<th>holding</th>
<th>transferring</th>
<th>finished</th>
<th>failed tot tf other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Production</td>
<td>6541</td>
<td>67253</td>
<td>04-08 18:03</td>
<td>4864</td>
<td>0 / 0</td>
<td>1829 / 0</td>
<td>1 / 0</td>
<td>22964 / 0</td>
<td>10960 / 0</td>
<td>6047 / 0</td>
<td>8347 / 366</td>
<td>15489 / 0</td>
<td>1617 / 0</td>
</tr>
<tr>
<td>CA</td>
<td>779</td>
<td>8021</td>
<td>04-08 18:03</td>
<td>399</td>
<td>0</td>
<td>402</td>
<td>0</td>
<td>3227</td>
<td>1559</td>
<td>147</td>
<td>800 / 0</td>
<td>1603</td>
<td>283</td>
</tr>
<tr>
<td>DE</td>
<td>1103</td>
<td>13400</td>
<td>04-08 18:03</td>
<td>634</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>5350</td>
<td>2464</td>
<td>64</td>
<td>2324 / 0</td>
<td>2531</td>
<td>636</td>
</tr>
<tr>
<td>ES</td>
<td>225</td>
<td>1561</td>
<td>04-08 18:02</td>
<td>198</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>681</td>
<td>204</td>
<td>81</td>
<td>216 / 0</td>
<td>359</td>
<td>10</td>
</tr>
<tr>
<td>FR</td>
<td>1201</td>
<td>6082</td>
<td>04-08 18:03</td>
<td>1994</td>
<td>0</td>
<td>806</td>
<td>0</td>
<td>1573</td>
<td>1299</td>
<td>157</td>
<td>396 / 1</td>
<td>1797</td>
<td>54</td>
</tr>
<tr>
<td>IT</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0 / 0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NL</td>
<td>231</td>
<td>2161</td>
<td>04-08 17:55</td>
<td>223</td>
<td>0</td>
<td>241</td>
<td>0</td>
<td>519</td>
<td>39</td>
<td>0</td>
<td>663 / 365</td>
<td>329</td>
<td>370</td>
</tr>
<tr>
<td>UK</td>
<td>1649</td>
<td>17404</td>
<td>04-08 18:03</td>
<td>862</td>
<td>0</td>
<td>318</td>
<td>1</td>
<td>3784</td>
<td>2079</td>
<td>4887</td>
<td>1856 / 0</td>
<td>4327</td>
<td>153</td>
</tr>
<tr>
<td>US</td>
<td>1170</td>
<td>17055</td>
<td>04-08 18:03</td>
<td>516</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>6951</td>
<td>3064</td>
<td>709</td>
<td>2036 / 0</td>
<td>4172</td>
<td>111</td>
</tr>
<tr>
<td>TW</td>
<td>183</td>
<td>1569</td>
<td>04-08 18:03</td>
<td>38</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>879</td>
<td>252</td>
<td>2</td>
<td>56 / 0</td>
<td>371</td>
<td>0</td>
</tr>
</tbody>
</table>
Non-ATLAS OSG Usage
Currently CHARMM protein folding application. Paper based on results

User/VO does
- job submission, using simple http-based Python client
- pilot submission, such that pilots carry their DN identity
- queue selection: tagging queues for inclusion in a logical ‘CHARMM’ queue

BNL/ATLAS/OSG provides
- PanDA service/DB infrastructure; same as used by ATLAS
- Panda monitoring, VO customized
- Machine(s) for VO pilot submission (@Madison)
- Support from ~3 FTE pool at BNL
- Future: Data mgmt and data-driven workflow
Security in PanDA

- Uses grid certificate based security for the PanDA server and its client communications (https)
- OSG workload management effort at BNL in collaboration with BNL Tier 1 and OSG very active in security extensions
- Currently integrating glExec to switch pilot identity to that of the downloaded user job
  - Prototype implemented and working; integration in production underway
  - Fermilab acting as US/OSG test site; finding an EGEE site is an ongoing challenge!
- Implementing limited proxies (pilot role) to restrict pilot capabilities
  - Prevent users from gaining access to extensive rights via the pilot proxy
  - Pilot proxy also concealed from the application code (removed from disk)
- Client<->Server validation, payload validation to come
  - Drawing on CMS/FNAL work
  - Will prevent hijacking PanDA pilots to run unauthorized jobs
- Authorizing pilot submitters and pilots themselves (beyond simply holding valid certificate) with time-limited token also to come
Near Term PanDA Activities, Plans

- **Scaling.** Have not hit limits yet and want to keep it that way.
  - MySQL performance, load sharing, failover, partitioning
  - Apache performance, tuning, redundancy, multi-host
  - System partitioning (time, region, activity, ...)
- **Security** -- as discussed
- **ATLAS-wide deployment**
  - Largely complete but tuning and extensions based on operations experience continues -- principal developer activity
- **Code organization, configuration management, packaging**
  - Housecleaning after 2.5 years! Maintainability, ease of component and service installation, easier pickup outside ATLAS
- **OSG.** PanDA is part of the OSG WMS program, managed from BNL
  - Security, Condor pilot glideins (with CMS), Pilot Factory (Condor scheduler glideins), production OSG WMS service, OSG infrastructure services (site attribute discovery/publishing, resource usage management)
Conclusion

• PanDA has delivered its objectives well
  • Efficient resource utilization (pilots and data management)
  • Ease of use and operation (automation, monitoring, global queue/broker)
  • Makes both production and analysis users happy
• Current focus is on broadened deployment and supporting scale-up
  • ATLAS-wide production deployment
  • Expanding analysis deployment (glExec dependent in some regions)
• Also support for ever-lengthening list of production, analysis use cases
  • Reconstruction reprocessing just implemented
  • PROOF/xrootd integration implemented and expanding
• Data management/handling always the greatest challenge
  • Extending automated data handling now that PanDA operates ATLAS-wide
• Leveraging OSG effort and expertise, especially in security and delivering services
• On track to provide stable and robust service for physics datataking
  • Ready in terms of functionality, performance and scaling behavior
  • Further development is incremental
  • Starting to turn our scalability knobs, in advance of operational need