



Readiness of ATLAS Computing - A personal view

Michael Ernst

Brookhaven National Laboratory

WLCG Collaboration Meeting

September 1, 2007

ATLAS Production



- ❑ **Computer System Commissioning (CSC)**
 - ❑ Software integration and operations exercise
 - ❑ Started more than 1.5 years ago (since end of 2005)
 - ❑ Distributed (MC) production goals
 - Validation of Athena software using 'standard' samples
 - Continuous production of high statistics physics samples
 - To exercise widely distributed computing infrastructure
 - ❑ Distributed Data Management (DDM) goals
 - Worldwide data distribution and replication
 - Controlled flow of data T0 -> T1 -> T2 -> T1
- ❑ **Bring sites and services to steady operations mode**

Distributed Production Operations



- ❑ ATLAS utilizes three production grids
 - ❑ EGEE, NDGF and OSG (10 Tier 1's, ~40 Tier 2's, +++)
 - ❑ NDGF (Nordic) ~500 CPU's, ~60TB disk
 - ❑ OSG (U.S.) ~2500 CPU's, ~386 TB disk
 - ❑ EGEE ~3500 CPU's, ~358 TB disk
 - ❑ Common features
 - Task/Job definition, Job dispatch (supervisor), metadata (AMI)
 - Access to data through ATLAS DDM system
 - ❑ Differences
 - Production software (executors), grid middleware
 - Independent operations team for each grid
 - Service architecture, storage systems...

Production Service Infrastructure used in the U.S.



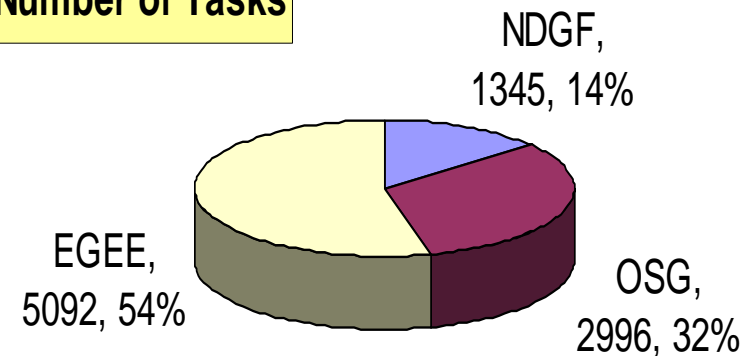
- ❑ PanDA is the distributed computing service infrastructure deployed in the U.S. and Canada (joined few weeks ago)
 - ❑ Includes all U.S. and Canadian Tier 1 and Tier 2's
 - ❑ Plus some opportunistic/shared sites
 - ❑ Works both with OSG (U.S.) and EGEE (Canada) middleware
- ❑ PanDA uses pilot jobs and provides a single task queue
 - ❑ Pilot jobs allow instant activation of highest priority tasks
 - ❑ Task queue brings familiar 'batch system' to distributed grids
 - ❑ Athena jobs transparently become 'pathena' jobs running on PanDA
- ❑ PanDA provides an integrated software and computing system for U.S./CA ATLAS sites, managing all production and user analysis activities

CSC Production Statistics

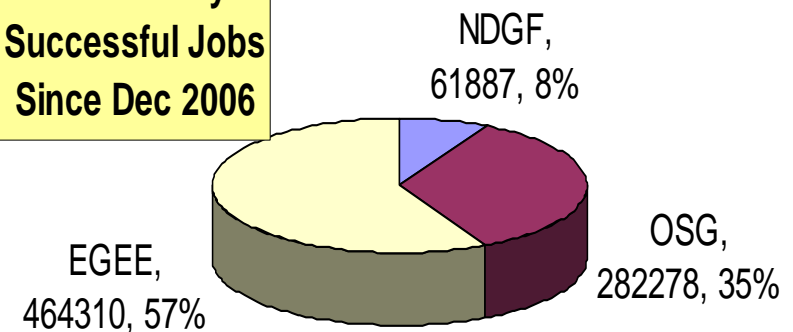


- Many hundreds of physics processes have been simulated
- Tens of thousands of tasks spanning two major releases
- Dozens of sub-releases (about every three weeks) have been tested and validated
- Thousands of 'bug reports' fed back to software and physics
- 50M+ events done from CSC12
- >300 TB of MC data on disk

Number of Tasks



Walltime Days Successful Jobs Since Dec 2006

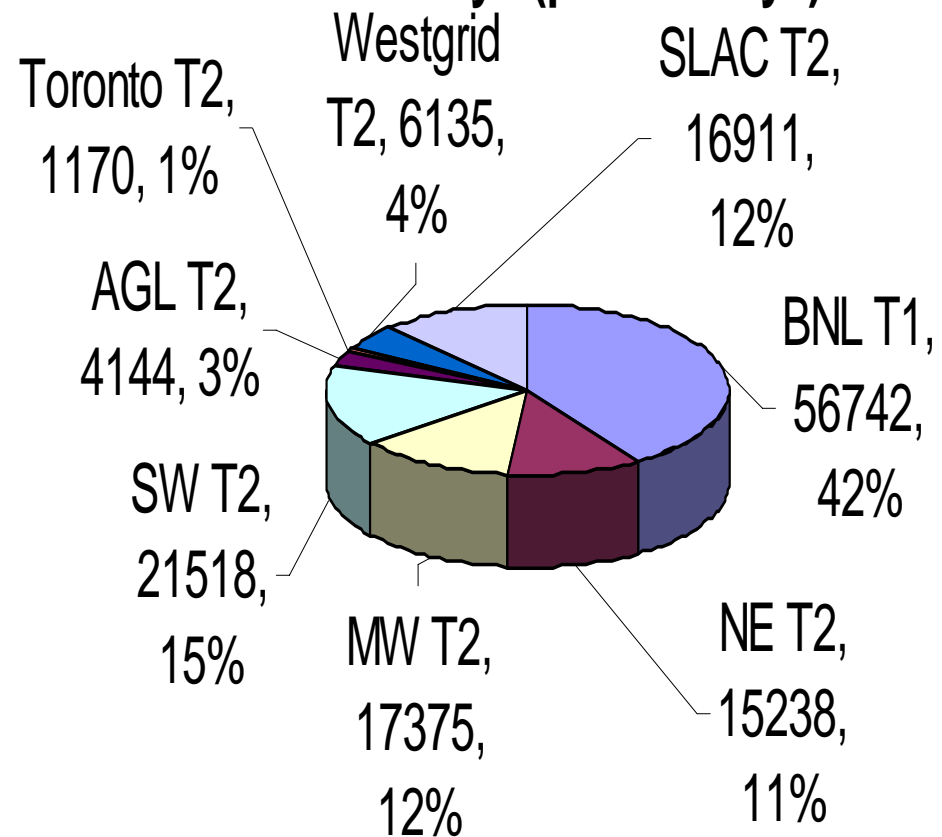


US / Canada Production Statistics (PanDA)



- ❑ PanDA has completed ~30M fully simulated physics events (simul+digit step), >30% of total central production
- ❑ Also successfully completed >15M single particle events
- ❑ Since November, all available CPU's occupied (ran out of jobs only for few days, plus few days of service outages)
- ❑ About 400 TB of original data stored at BNL T1 (includes data generated on other grids)
- ❑ Additional ~100 TB of replicas kept at U.S. ATLAS Tier 2 sites

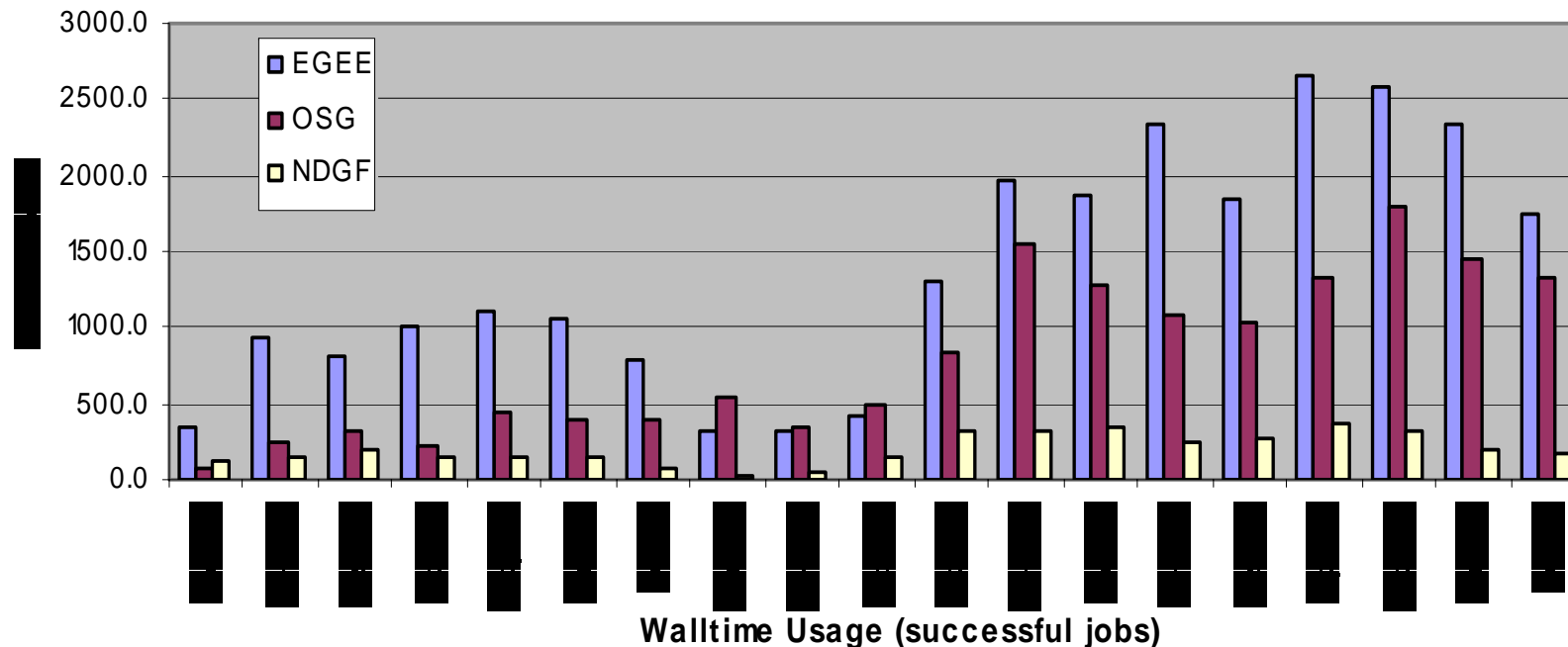
Walltime Usage by Successful Jobs in CPU days (past 14 days)



Resource Usage



- ❑ CPU and disk resources available to ATLAS rising steadily
- ❑ Production system efficiencies are steadily increasing
- ❑ But much more will be required for data taking
- ❑ Additional resources are coming online soon



DDM Status – a Personal View



- ❑ **ATLAS Distributed Data Management is a Critical Component**
 - ❑ To catalogue and manage flow of data
 - ❑ It has not performed at the level needed
- ❑ **Main issues**
 - ❑ Robustness of site services (transfer agents)
 - ❑ Quality of Service (completion of data movement tasks)
 - ❑ Too much manual intervention required
- ❑ **What can we do?**
 - ❑ Add more effort to DQ2 development
 - ❑ Develop backup plan for data movement (Catalogues are fine)
 - ❑ Giving up on ATLAS computing model is not an option!
 - ❑ We are discussing these topics actively and daily

The End Game



□ Goal: steady ramp up of operations till the end of 2007

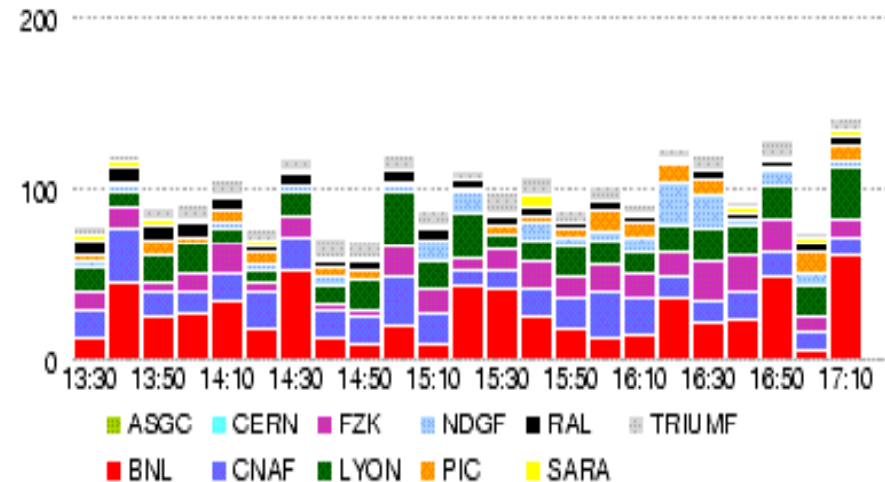
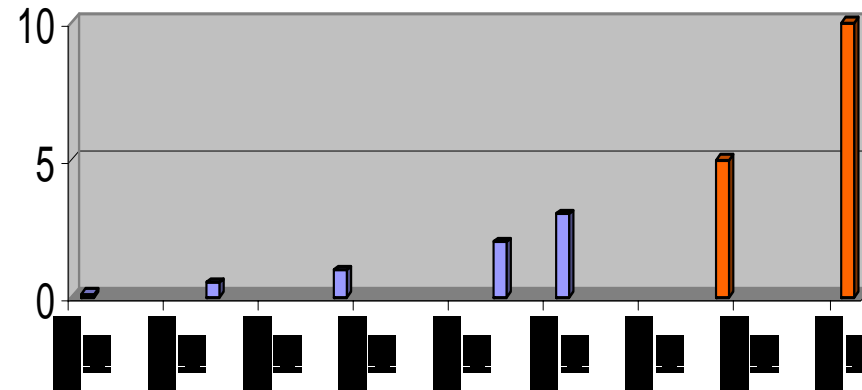
□ MC Production

- Achieve 10 Million evts/day by the end of 2007 (100 - 200k production jobs/day)
- All targets that were set in 2005 have been met – so far

□ DDM Operations

- Steady operations
- Automation and monitoring improvements
- Computing Model data flow implemented by end of 2007

Production Rate (Mevts/day)



Data Location (U.S.)



- ❑ Tier 1 – main repository of data (MC & Primary)
 - ❑ Store complete set of ESD, AOD, AANtuple & TAG's on disk
 - ❑ Fraction of RAW and all U.S. generated RDO data
- ❑ Tier 2 – repository of analysis data
 - ❑ Store complete set of AOD, AANtuple & TAG's on disk
 - ❑ Complete set of ESD data divided among 5 Tier 2's
- ❑ Data distribution to Tier 1 & Tier 2's will be managed
- ❑ Tier 3 – unmanaged data matching local interest
 - ❑ Data through locally initiated subscriptions
 - ❑ Mostly AANtuple's, some AOD's
 - ❑ Tier 3's will be associated with Tier 2 sites?
 - ❑ Tier 3 model is still not fully developed – need physicist's input

Conclusion



- ❑ Computing systems are in good shape – through long running commissioning (CSC) exercise
 - ❑ Current ATLAS Data Replication system needs significant improvements
- ❑ 2006 – successful transition into operations mode
- ❑ 2007 – high volume production and DDM operations
- ❑ Successful distributed MC production for CSC notes
- ❑ But many new challenges to come
- ❑ Need more participation by physicists – as users, developers, shifters... provide feedback to ATLAS Computing
- ❑ Lots of resources available for physicists
- ❑ Overall, progressing well towards full readiness for LHC data processing