Evolution of the ATLAS PanDA Production and Distributed Analysis System

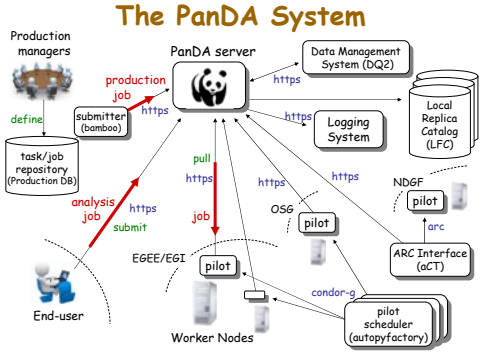
T. Maeno[1], K. De[2], T. Wenaus[1], P. Nilsson[2], R. Walker[3], A. Stradling[2], V. Fine[1], M. Potekhin[1], S. Panitkin[1], G. Compostella[4], for the ATLAS Collaboration

[1] Brookhaven National Laboratory, USA, [2] University of Texas at Arlington, USA,
[3] Ludwig-Maximilians-Universitat Munchen, Germany, [4] Max-Planck-Institut für Physik, Germany

Introduction

The PanDA (Production and Distributed Analysis System) is the ATLAS workload management system for processing all Monte-Carlo simulation and data reprocessing jobs in addition to user and group analysis jobs. PanDA has performed well with high reliability and robustness during the two years of LHC data-taking, while being actively evolved to meet the rapidly changing requirements for analysis use cases. We will present a brief overview of the PanDA system, an overview of system evolution, results from the analysis of two years of PanDA usage statistics, and plans for the future.

The PanDA System



- The PanDA server [1]
  - the main component for managing a central task queue for all jobs
  - a brokerage module operates to prioritize and assign work on the basis of job type, priority, input data and its locality, and available CPU resources
- The pilot [2]
  - retrieves a job based on its priority from the PanDA server
  - runs the jobs as soon as CPU slots become available
- The autopyfactory [3]
  - pre-schedules pilots to OSG and EGEE/EGI grid sites using Condor-G
- The PanDA monitor [4]
  - Web-based monitoring system both for production and analysis

Improvements

Rebrokerage for Analysis jobs

✓ The brokerage assigns jobs to sites based on input data locality, workload distribution, software and CPU availability, site downtime and status, exactly when each job is submitted
✓ The situation may change while jobs are waiting at sites to be picked up by pilots
  ✓ More replicas of input may become available
  ✓ Workload distribution may change
✓ Rebrokerage has been implemented
  ✓ Waiting jobs are periodically reassigned to other sites
  ✓ When sites are blacklisted by HammerCloud [5] or Site Status Board [6], waiting jobs are immediately reassigned to other sites

Automatic reattempts for analysis jobs

✓ The pilot investigates the cause of failure for each failed job and reports it back to the PanDA server
✓ Automatic reattempt mechanism has been implemented to do as follows:
  ✓ For recoverable problems, retries failed jobs at most three times at the same site
  ✓ For site problems, sends failed jobs to another site

Adaptation for the CernVM File System

✓ Many ATLAS sites have been configured to use software on CernVM Files System [7]
✓ The brokerage still checks validity of software for those sites

Improvements (cntd)

Support for the multi-cloud model

✓ A cloud was composed of one Tier-1 site and regional Tier-2 sites
  ✓ Production tasks are assigned to clouds
  ✓ Data for production were transferred in each cloud → simplification of data transfers
✓ The constraint has been relaxed after accumulation of operational experience
✓ A Tier-2 site can be associated to multiple Tier-1 sites, i.e., can belong to multiple clouds
  ✓ Better usage of Tier-2 CPU resources
  ✓ Tier-2 sites can survive even if a Tier-1 site is down

Beyond-pledge Resource Management and Preferential Brokerage

✓ Some sites have regional CPU/storage resources by using budgets beyond ATLAS MoU share in addition to official resources
✓ Each site can be configured to use additional resources only for the users who belong to a particular country group when their jobs are waiting in the queue, otherwise, use those resources for other general users
✓ When the users who belong to a country group submit jobs, the brokerage preferentially assigns them to sites that provide additional resources for the country users

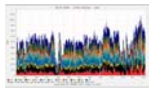
Fairshare for production activities

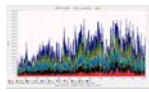
✓ ATLAS can set a fairshare policy at each site to define how CPU resources are allocated to production activities and/or physics working groups

Outbound network connection monitoring

✓ To block and/or kill problematic analysis jobs which naively do DoS attacks by connecting remote hosts
✓ connect and execve system calls are trapped on the worker node and all outbound network connections are monitored

Current Status





✓ The PanDA system has performed very well during the LHC data-taking year
✓ Concurrently running over 100k production jobs with 8% of single job failure rate
✓ Steady increase of analysis activities

Future Plans

✓ Adoption of new technologies
  ✓ NoSQL, MQ, cloud service …
✓ Development of Job Execution and Definition Interface (JEDI)
  ✓ Dynamic job generation in PanDA
  ✓ Improvement of automation and efficiency
✓ Moving client functionalities to the server side
✓ Utilization of multi-core queues and AthenaMP
✓ Extension of PanDA as a generic high level workload manager usable by the wider distributed scientific computing community

References

1. Maeno T., *Overview of ATLAS PanDA Workload Management*, J. Phys. Conf. Ser. 331 (2011)
2. Nilsson P., *The ATLAS PanDA Pilot in Operation*, J. Phys. Conf. Ser. 331 (2011)
3. Caballero J., *AutoPyFactory: A Scalable Flexible Pilot Factory Implementation*, in Proc. of CHEP2012
4. Potekhin M., *The ATLAS PanDA Monitoring System and its Evolution*, J. Phys. Conf. Ser. 331 (2011)
5. Legger F., *Improving ATLAS grid site reliability with functional tests using HammerCloud*, in Proc. of CHEP2012
6. Iglesias C. B., *Automating ATLAS Computing Operations using the Site Status Board*, in Proc. of CHEP2012
7. Predrag B., *Status and Future Perspectives of CernVM-FS*, in Proc. of CHEP2012