



# BigPanDA Status

**Kaushik De**

**Univ. of Texas at Arlington**

**Alexei Klimentov**

**Brookhaven National Laboratory**

**OSG AHM, Clemson University**

**March 14, 2016**



# Outline



- **BigPanDA Overview**
  - Goals and Plans
  - Accomplishments over the past 3 years
- **What is happening now**
  - And for the next 6 months
- **Future of BigPanDA**



# BigPanDA Project



## “Next Generation Workload Management and Analysis System for BigData”

- DOE ASCR and HEP funded project started in Sep 2012
  - Generalization of PanDA Workload Management System as meta application, providing location transparency of processing and data management, for HEP and other data-intensive sciences, and a wider exascale community
  - Program Managers : Rich Carlson (ASCR), Lali Chaterjee (HEP)
- Institutions :
  - Lead Institution : Brookhaven National Laboratory
    - Pis : Alexei Klimentov\*, Sergey Panitkin, Torre Wenaus and Dantong Yu
  - Argonne National Laboratory :
    - PI : Alexandre Vaniachine
  - University Texas at Arlington :
    - Pis : Kaushik De, Gergely Zaruba
- July 2015 : Project extended until Dec 31 2016



# BigPanDA Work Packages



- **WP1. Factorizing the code**
  - Factorizing the core components of PanDA to enable adoption by a wide range of exascale scientific communities
- **WP2. Extending the scope**
  - Evolving PanDA to support extreme scale computing clouds and Leadership Computing Facilities
- **WP3. Leveraging intelligent networks**
  - Integrating network services and real-time data access to the PanDA workflow
- **WP4. Usability and Monitoring**
  - Real time monitoring and visualization package
- **Schedule**
  - Year 1. Setting the collaboration, define algorithms and metrics
  - Year 2. Prototyping and implementation
  - Year 3. Production and operations



# WP1 Factorizing the core



- **New Code repository.**
  - Migration from CERN SVN to GitHub
  - New build system; distribution through RPMs (thanks OSG for collaborative effort)
- **PanDA server improvements.**
  - PanDA server consolidation: now **one** master branch that serves multiple experiments
    - Consolidated divergence into multiple branches over many years
    - Standardized installation thanks to collaboration with OSG
- **PanDA pilot improvements.**
  - Core pilot has been refactored to a generic (VO independent) version;
  - VO specifics are handled as plug-ins; execution backends are handled as plug-ins
- **Multiple database backends.**
  - Oracle database backend (ATLAS, AMS, COMPASS)
  - MySQL (running on EC2 PanDA server) – LSST, ALICE
  - New PanDA instance with MySQL backend deployed in Amazon EC2
    - Instance tuned for multi-VO support
- ~~Documented installation process for all components~~



# WP2 Extending the scope



- **Extend PanDA job management to Clouds and HPC**
  - Commercial and academic clouds
    - Google, Amazon EC2, NECTAR, etc...
  - Primary goal for HPC activity - Leadership Computing Facilities
    - Titan@OLCF – main target
    - Mira@ALCF, Aencelm, Ostrava (CZ), Archer (UK), Kurchatov (RF)
- See talk tomorrow by Frank/me on HPC's



# Titan Example



- 27 PFlops (Peak theoretical performance). Cray XK-7
- 18,688 compute nodes with GPUs
- 299,008 CPU cores
- AMD Opteron 6200 @2.2 GHz (16 cores per node)
- 32 GB RAM per node
- NVidia TESLA K20x GPU per node
- 32 PB disk storage (center-wide Luster file system)
- >1TB/s aggregate FS throughput
- 29 PB HPSS tape archive

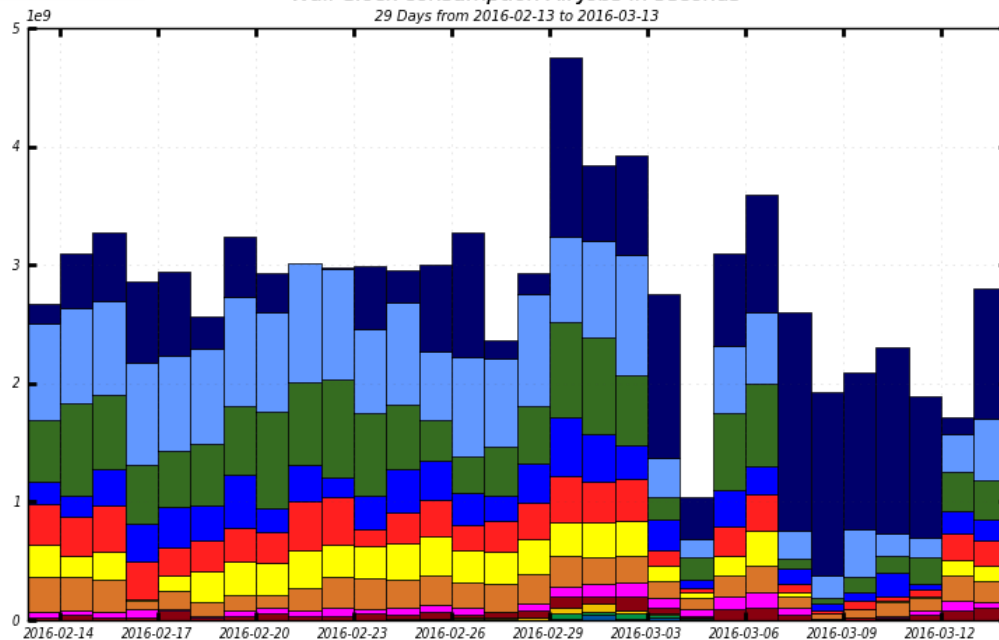


# Titan Usage Last Month

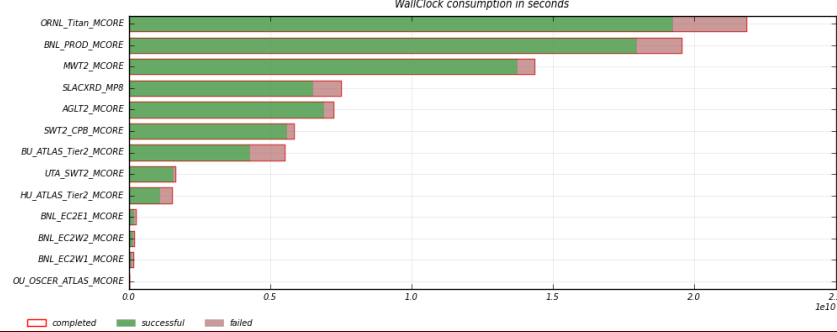


### Wall Clock consumption All Jobs in seconds

29 Days from 2016-02-13 to 2016-03-13



- ORNL Titan MCORE
- BNL\_PROD MCORE
- MWT2 MCORE
- SLACXRD MP8
- AGLT2 MCORE
- SWT2\_CPB MCORE
- BU\_ATLAS\_Tier2 MCORE
- UTA\_SWT2 MCORE
- HU\_ATLAS\_Tier2 MCORE
- BNL\_EC2E1 MCORE
- BNL\_EC2W1 MCORE
- OUI\_OSCER\_ATLAS MCORE







# WP3 PanDA and Networking



- **Goal for PanDA**
  - Direct integration of networking with PanDA workflow – never attempted before for large scale automated WMS systems
- **Pick a few use cases**
  - Cases which are important to PanDA users
  - Enhance workload management through use of network
- **Case 1: Improve User Analysis workflow**
- **Case 2: Improve Tier 1 to Tier 2 workflow**
- **Step by step approach**
  - Collect network information
  - Storage and access
  - Using network information



# Network Summary



- **All work completed in first 2 years**
  - Successful demonstrations – overflow and automatic MCP
  - Results shown at SC15 and other meetings
- **New implementation underway for “World Cloud”**
  - Details in Fernando’s talk earlier
  - Moving from demonstration to operation



# WP4 Usability and monitoring



- New PanDA monitoring web application was developed based on Django framework
  - All of you know this daily as BigPanDA monitoring
  - Allows rapid development and easy extension

ATLAS PanDA monitor Dashboards Jobs Tasks Errors Users Sites Incidents Search Prodsys Services VO Help

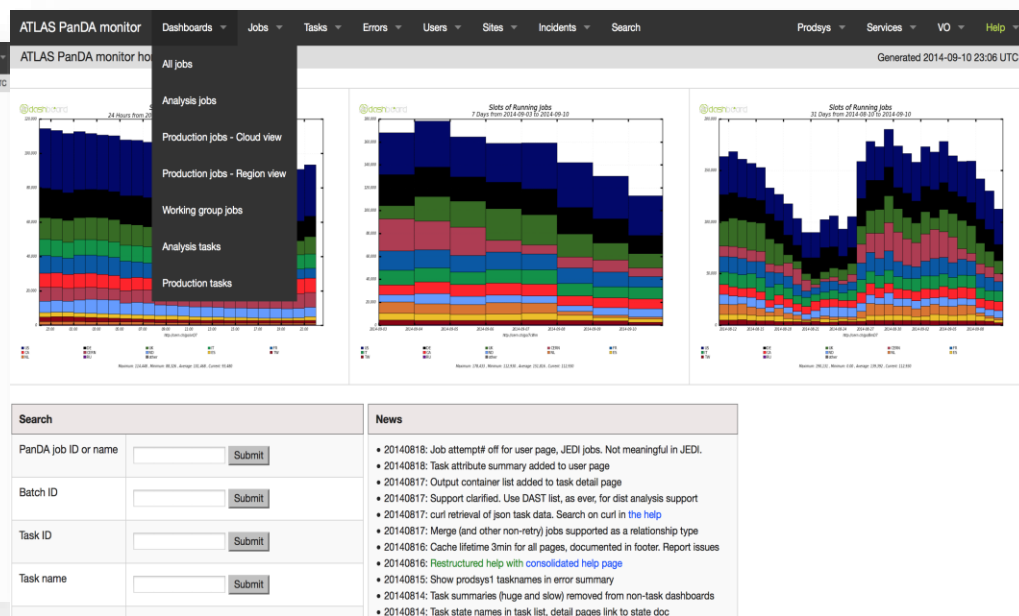
PanDA analysis dashboard, last 12 hours. Query params: Generated 2014-09-10 23:37 UTC

analysis task summary by cloud, last 7 days **Hover over state name to see full name. Task state documentation**

Cloud	nTask	req	def	assign	rtly	pend	scout	scld	run	prep	done	fail	finish	abrtg	abrdt	finshg	toprep	prepcr	tozbrk	broken	retry	inacc	refine
Cloud	1124	0	0	0	0	2	76	0	85	0	738	14	21	0	24	0	0	0	0	154	0	0	0
UK	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
US	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0

Cloud / Site summary of analysis jobs For a description see below

Cloud	Status	nJobs	nPlots	defined	waiting	assigned	throttled	activated	sent	starting	running	holding	transferring	finished	failed	cancelled	% failed
All clouds		524651	20713	19629	0	0	2008	63220	11	954	29176	4359	1	318070	54147	32858	14
Failed before site assigned		1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
CA ✓	online	28441	1199	333	0	0	0	3015	0	0	577	323	0	21241	1021	1931	4
CERN ✓	online	7195	422	414	0	0	0	3917	2	0	779	33	0	1324	493	233	27
DE ✓	online	103826	9150	1245	0	0	0	4090	1	0	4876	955	0	79948	11874	1935	13
ES ✓	online	16829	889	100	0	0	0	1226	0	0	1394	55	0	12694	1073	297	7
FR ✓	online	39002	1324	370	0	0	0	6211	0	0	3273	299	0	25143	2231	1515	8
IT ✓	online	32685	1084	991	0	0	0	1901	0	0	2557	235	0	15836	8948	2516	38
ND ✓	online	38829	42	12998	0	0	0	11038	0	951	264	7	1	917	966	11527	51
NL ✓	online	18327	1426	361	0	0	0	1983	0	0	1478	105	0	11635	1332	1433	10
RU ✓	brokeroff	486	36	1	0	0	0	27	0	0	305	7	0	79	81	16	39





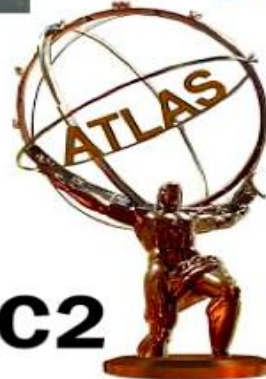
# “BigPanDA” Collaborations



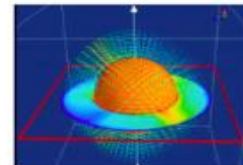
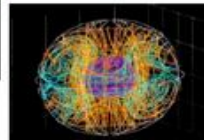
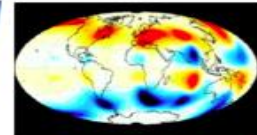
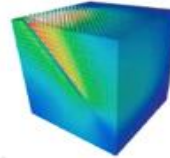
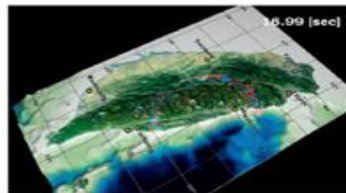
- Through ASCR project, PanDA has moved well beyond ATLAS
  - Collaborating with RADICAL group of Rutgers University and Oak Ridge LCF team (Shantenu Jha and Jack Wells)
    - Many joint papers and invited talks
- Collaboration between ATLAS, ALICE, nEDM experiments for efficient usage of opportunistic resources, especially HPC and LCF
- LSST evaluating PanDA for distributed data processing
- COMPASS and AMS02 installed their own PanDA instances
- Other communities getting involved



# BigPanDA - Generalizing PanDA Beyond Grid and ATLAS



Titan System (Cray XK7)			
Peak Performance	27.1 PF	18,608 compute nodes	24.5 PF GPU, 2.6 PF CPU
System memory	710 TB total memory		
Interconnect	Generic High Speed Interconnect		30 Teras
Storage	Local Filesystem		32 PB
Archive	High-Performance Storage System (HPS)		29 PB
I/O Nodes			512 Service and I/O nodes





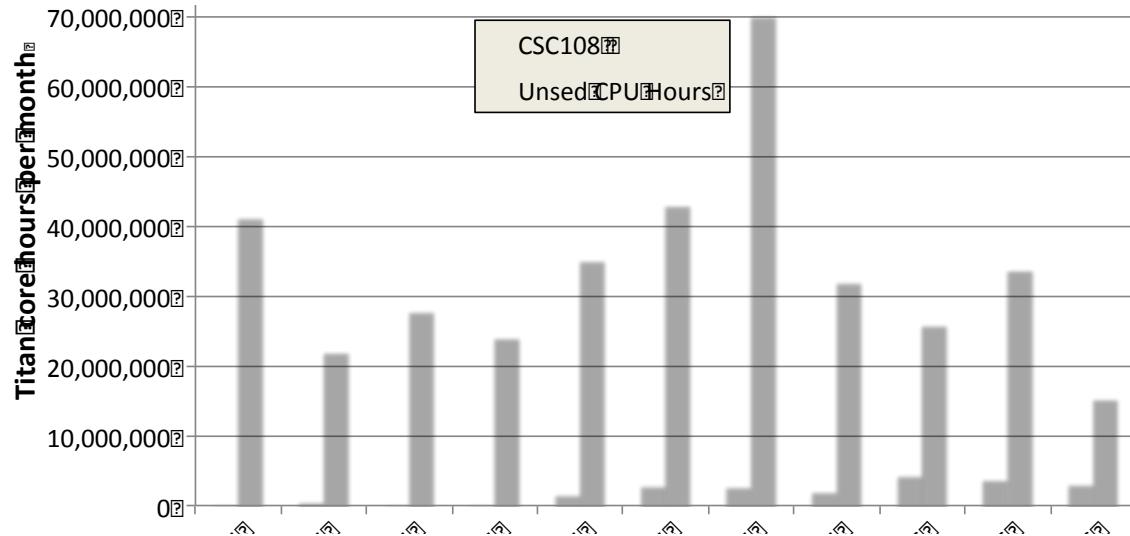
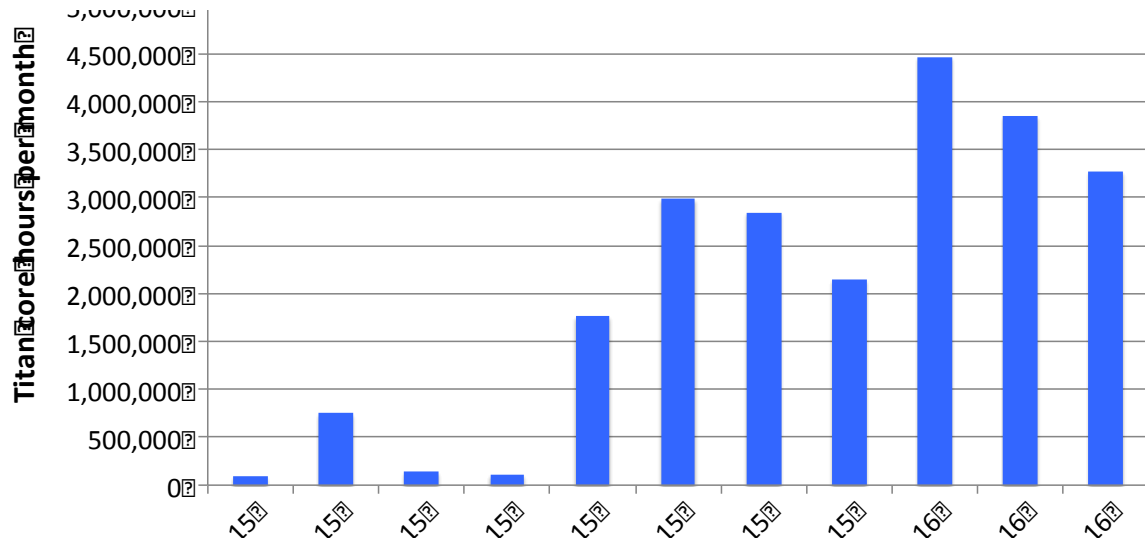
# What Are We Doing Now?



- **WP1 – continuing to improve packaging**
  - Bi-weekly meetings, focusing on non-ATLAS experiments
  - Work carried out by core PanDA and people from other experiments
- **WP3 – networking demonstration completed**
  - Ongoing work done by core PanDA and ADC teams
  - World Cloud, NWS, Configurator...
- **WP4 – BigPanDA monitoring big focus**
  - New features being added continuously
- **WP3 – increasing LCF usage**
  - Heavy focus of work
  - Collaborating with Rutgers CSE team and OLCF operations



# Titan Backfill





# Future of BigPanDA



- Submitted proposal to ASCR for continuation
  - Hope to hear back before summer
  - Focus on BNL+UTA+Rutgers+OLCF collaboration on Titan
  - Setup PanDA as a service on Titan for 'all' users
  - ATLAS (and ALICE) will benefit from big increase in CPU cycles