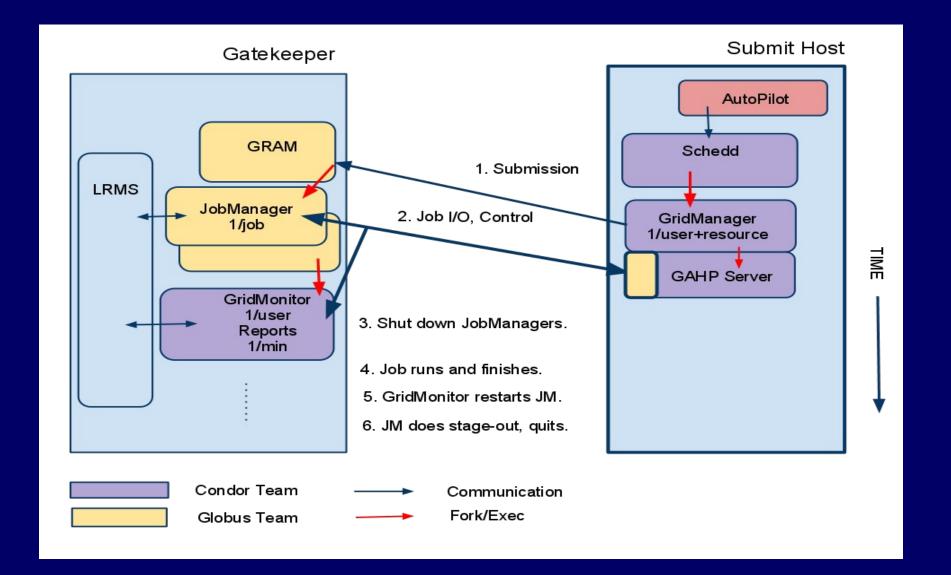
PanDA Pilot Submission using Condor-G: Experience and Improvements

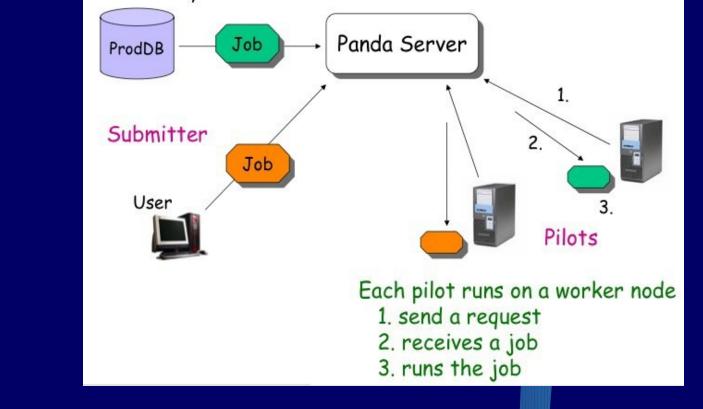
Xin Zhao, John Hover, Tomasz Wlodek, Torre Wenaus Brookhaven National Laboratory Jaime Frey, Todd Tannenbaum, Miron Livny University of Wisconsin-Madison



Production system

PanDA (Production and Distributed Analysis) is the workload management system of the ATLAS experiment, used to run production and user analysis jobs on the grid. As a late-binding, pilot-based system, to maintain a stable and scalable pilot submission system is critical for PanDA operation.





The ATLAS Computing Facility (ACF) at BNL, as the ATLAS Tier1 center in the US, operates the pilot submission systems for the US cloud. This is done using the PanDA "AutoPilot" scheduler components which submits pilot jobs via Condor-G, a grid job scheduling system developed at the University of Wisconsin-Madision.

Improving Scalability and Stability of Condor-G and AutoPilot

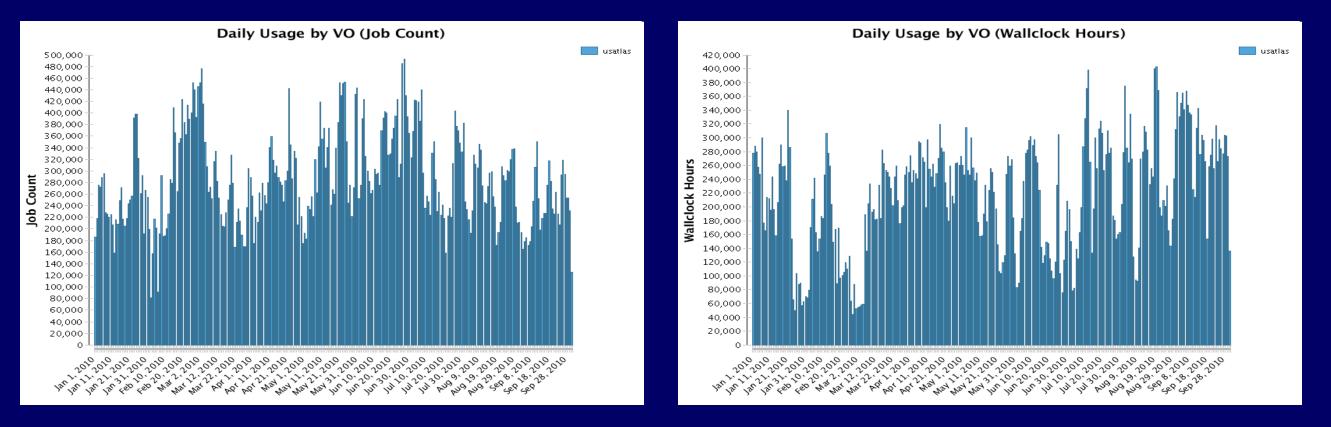
Isolation of site specific issues from affecting submission to other sites

PanDA AutoPilot Operation at BNL ACF

- Pilot submission in the US cloud
 - ⇒92 PanDA queues at 43 gatekeepers, with HEPSPEC-06 ~ 111,000
 - Full scale production has >10,000 real jobs running all the time in the US cloud, daily pilots peak reached ~ 480,000 pilots

- Separate GridManager process for each remote CE
- GRIDMANAGER_SELECTION_EXPR = GridResource
- Introduction of the "Nonessential" job attribute
 - Unconditional removal and cleanup of stale pilots to prevent clogging the whole submission process
- Role-based throttle on limiting jobmanager processes on remote CE
 - Half for job submission, half for job completion/cleanup
- Grid Monitor restart behavior adjustable
 - GRID_MONITOR_DISABLE_TIME
 - Refined GridManager error handling to avoid flooding sites with the Grid Monitor jobs
- Better scheduling in AutoPilot to avoid overloading remote CE
 - * Nqueue adjustable based on real job status
 - Multi-job pilots
- Bug fixes in Condor and underlying Globus libraries
 - Internal inefficient loops fixed
 - Integrated updtream Globus bug fixes into GAHP server

- →5 Condor-G submit hosts (3 primary) at BNL Tier1
- Stress test results: 50,000 jobs managed on one submit host; 30,000 jobs submitted to one remote CE (GT2 gatekeeper)

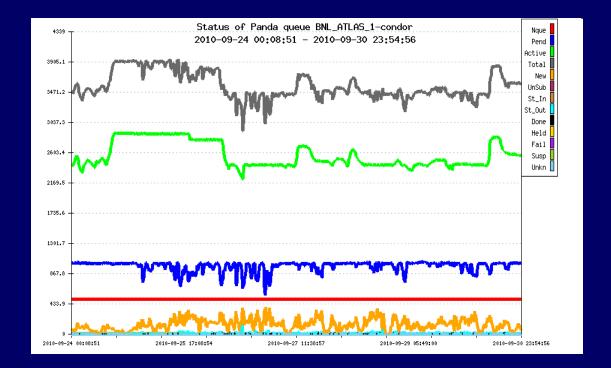


- Some best practices in Condor-G submission
 - Reduce frequency of proxy renewal
 - Avoid hard-kill (-forcex) jobs from client side
 - Optimize PanDA queues setup among several submit hosts to avoid starvation of pilots for any queues.
 - Multiple AutoPilot instances to increase injection rate for large PanDA queue
- Caching condor_q command results to reduce load of condor Schedd on the submit host

- Improved grid job status monitoring
 - condor_status -grid



- Monitoring of pilot submission
 - Dashboard & Nagios alerts
 - https://nagios.racf.bnl.gov/nagios/cgi-bin/prod/dashboard.php



Main Page Production System Overview Panda Production Pilots	Current Panda Jobs													
Panda Production Queues Current Panda Jobs	Queues on machine gridui07.usatlas.bnl.gov; status as of 2010-10-01 18:34:52													
Queues on Submit Hosts	Panda Queue name	Total	New	UnSub	St In	Pend	Active	St Out	Done	Held	Fail	Susp	Unkn	Nque
BNL gatekeepers	AGLT2-OSG-condor	150	0	0	10	107	32	1	0	91	0	Ō	0	50
BNL Gatekeepers	AGLT2-condor	563	8	0	0	110	453	0	0	0	0	0	0	110
BNL Grid Ftp Servers	ANALY_BNL_ATLAS_1-condor	451	0	0	0	309	142	0	0	0	0	0	0	250
	ANALY LONG BNL ATLAS	798	6	0	0	155	643	0	0	0	0	0	0	150
	ANALY_Tufts_ATLAS_Tier3-1sf	80	10	0	0	56	24	0	0	0	0	0	0	10
	BNL_ATLAS_1-condor	1726	91	0	0	464	1259	0	0	0	0	0	0	500
	BNL_ATLAS_2-condor	354	7	0	0	190	161	0	0	0	0	0	0	200
Other monitoring links:	BNL_ITB_ATLAS_TEST-condor	2	2	0	0	0	0	0	2	0	0	0	0	1
	BNL_ITB_Test1-condor	8	6	0	0	2	4	2	0	0	0	0	0	4
RACF Nagios Services Table	BNL_SITE_GR02-condor	350	0	0	0	350	0	0	0	0	0	0	0	50
Current RSV status	FIU-PG-condor	70	0	10	0	0	0	0	0	60	0	0	0	10
	Firefly_SBGRID-pbs	10	9	1	0	0	0	0	0	0	0	0	0	10
	Harvard-East_SBGRID-condor	12	12	0	0	2	10	0	0	0	0	0	0	10
	IU_OSG-pbs	1	1	1	0	0	0	0	0	0	0	0	0	1
	LBNL_DSD_ITB-condor	15	12	0	0	0	14	0	1	0	0	0	0	10
	MWT2_IU-pbs	1153	17	0	0	512	640	1	0	0	0	0	0	512
	OSG_LIGO_PSU-pbs	130	0	130	0	0	0	0	0	0	0	0	0	10
	OUHEP_ITB-condor	19 10	16 0	10	0	6 N	10	1	0	0	0	0	0	10
	SBGrid-Harvard-Bast-condor	90	0	90	0	0	0	0	0	0	0	0	0	10
	SBGrid-Harvard-Exp-condor SWT2 CPB-pbs	120	0	90	0	20	100	0	0	0	0	0	0	20
	TTU TESTWULF ITB	1956	0	0	0	20	1902	0	0	0	0	0	0	20
	TTU_TESTWOLF_ITB Tufts ATLAS Tier3-1sf	1956	20	0	2	34 46	1902	0	0	0	0	0	0	20
	UCITB EDGE7-pbs	16	10	0	2 0	11	4	1	0	0	0	0	0	10
	UConn=08G=condor	2100	10	0	1971	11	72	23	0	2028	0	0	0	50
	UFlorida-IHEPA-condor	130	10	130	19/1	0	10	23 N	0	0202	0	0	0	10
	UFlorida-IHEFA-hg-atlas-condor	130	0	130	0	0	0	0	0	0	0	0	0	10
	UFlorida-PG-condor	130	ň	130	ň	0	0	ő	0	0	0	ŏ	0	10
	UFlorida-PG-pg-atlas-condor	12	ŏ	100	n	10	0	0	ň	ő	n	2	0	10

