

A Dynamic System for ATLAS Software Installation on OSG Sites

Xin Zhao, Tadashi Maeno, Torre Wenaus
(Brookhaven National Laboratory, USA)

Frederick Luehring (Indiana University, USA)

Saul Youssef, John Brunelle (Boston University, USA)

Alessandro De Salvo (Istituto Nazionale di Fisica Nucleare, Italy)

A.S. Thompson (University of Glasgow, United Kingdom)



Outline

- Motivation
 - ATLAS Software deployment on Grid sites
 - Problems with old approach
- New Installation System
 - Panda
 - Pacball via DQ2
 - Integration with EGEE ATLAS Installation Portal
- Implementation
- New system in action
- Issues and future plans



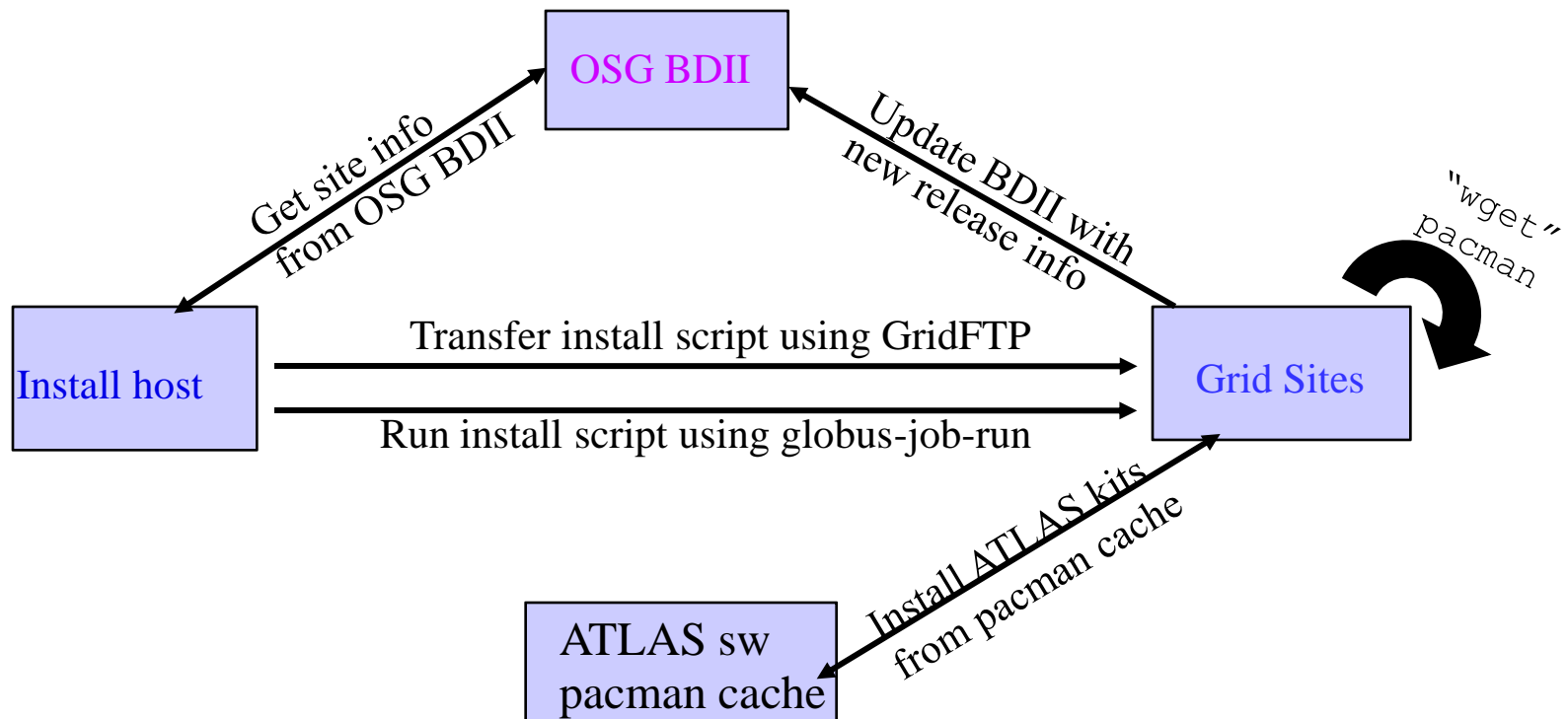
Motivation

- ATLAS software deployment on *Grid* sites
 - ATLAS *Grid* production, like many other VO applications, requires the software packages to be installed on *Grid* sites in advance.
 - A dynamic and reliable installation system is crucial for the timely start of ATLAS production and to reduce job failure rate.
 - OSG and EGEE each have their own installation system.



Motivation

- Previous ATLAS installation system for OSG



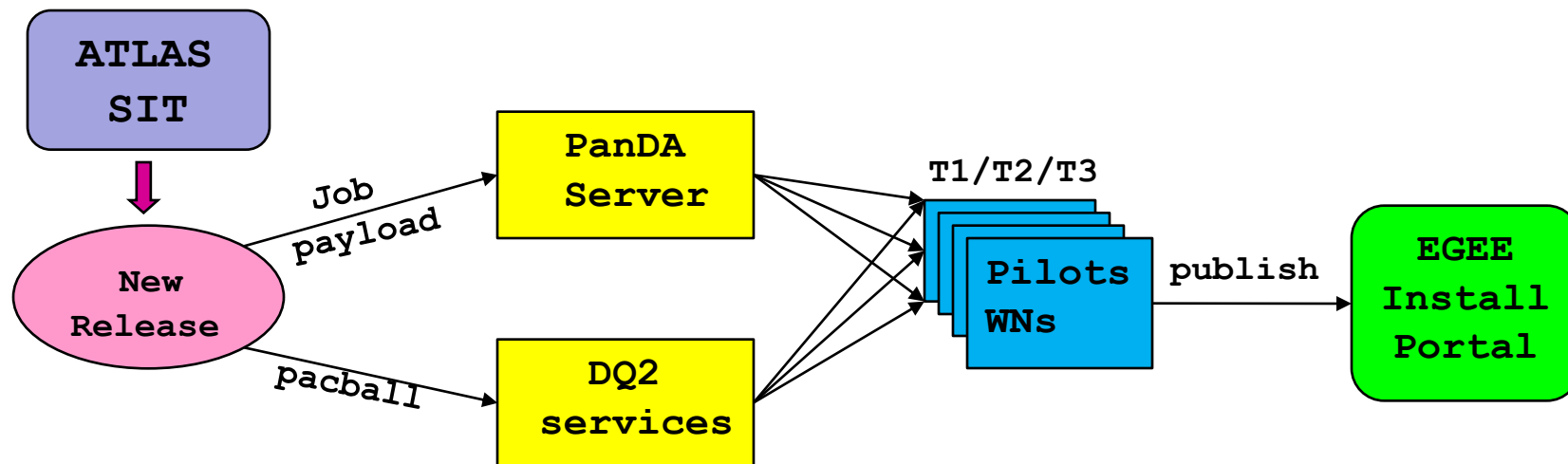


Motivation

- Previous ATLAS installation system for OSG (continued)
 - Problems
 - Installation done on gatekeeper
 - Not the real job running environment
 - Gatekeeper could run different OS than the worker nodes
 - Increase loads on the gatekeeper
 - Accessing directly CERN pacman cache
 - Increase load on CERN web server
 - Installation is slow with thousands of small files to transfer across the Atlantic ocean.
 - OSG and EGEE each has its own install scripts, monitoring tools etc. Unification is desirable.



New Installation System



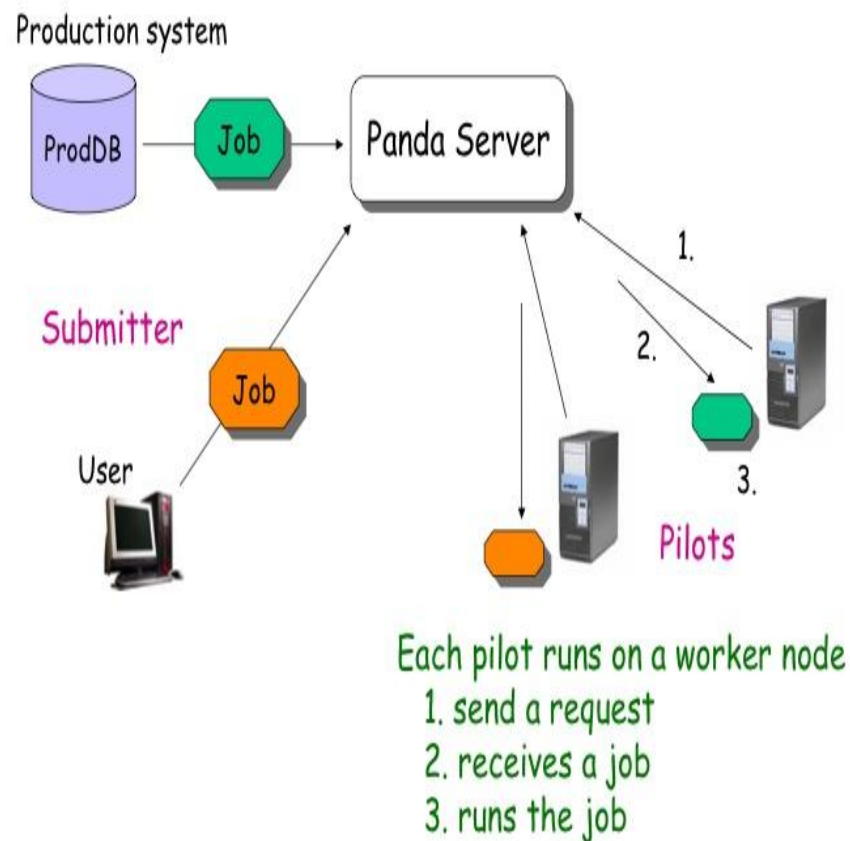
- Installation is done locally on the worker nodes
 - Panda pilots
- Release cache localization
 - Pacball via DQ2
- Unification
 - EGEE ATLAS installation portal and installation scripts



New Installation System

Job flow in PanDA

- Installation on WNs
 - PanDA provides a steady pilots stream to all WNs
 - PanDA job flow
 - Availability: wherever runs production, there are pilots
 - Installation task is a job payload to PanDA server, just as normal production and user analysis jobs, then dispatched to next available pilot.





New Installation System

- Pacman cache localization
 - Pacball - local pacman cache wrapped in a script
 - Machine independent, self-contained, self-installing
 - Easy to transfer and run locally, without accessing other pacman tools and internet
 - Un-updateable --- md5 checksum attached
 - Uniquely identifiable and reproducible software environment
 - In ATLAS, starting with 14.X.X series, every new ATLAS base kit and production cache will have a corresponding pacball created
 - Transfer pacballs to worker nodes
 - DQ2 treats each pacball as a data file, puts them in data container, and subscribes to all T1s
 - On OSG sites, pandamover jobs transfer pacballs from T1(BNL) to all T2 sites, just like the standard procedure of staging in job input files
 - Pilots retrieve pacballs from local SE and put them on WNs local disk



New Installation System

- Unification
 - Integration with EGEE
 - Job payload uses the same installation scripts as used by EGEE sites
 - Installation status published to EGEE ATLAS installation portal
 - Easily accessible across OSG and EGEE for ATLAS users
 - Information is also available on OSG and EGEE Grid BDII
 - Panda and DQ2 are used in ATLAS worldwide, as the main workflow and dataflow management systems
 - A uniform platform for the new installation system



Implementation

- New installation queues are defined in PanDA server
 - Each production site has a corresponding installation queue. Usually the installation queue is defined in PanDA to share the same batch job slots with the production queue on a site.
 - PanDA scheduler sends installation pilots to all installation queues, at a low but steady rate (queue depth is 1)
 - Installation pilots use voms proxy with /atlas/usatlas/Role=software role, mapped to usatlas2 account on OSG sites (production role is mapped to usatlas1).
 - Pilot adds a new job type to distinguish installation jobs from production and user analysis jobs



Implementation

- Pacballs are defined and put into DQ2
 - For each new ATLAS release or cache, create its pacball, put into AFS area, which is mirrored to an INFN web server
 - Use dq2-put to create a new dataset and register to DQ2 catalog
 - DQ2 replicates pacball datasets from T0 to T1s
 - On DQ2 site service, LFC ACL is changed to grant "software" role voms proxy permission to register output (log) files of the installation pilots
 - SRM space token "PRODDISK" is used to store output files of installation pilots



Implementation

- Job payload
 - EGEE installation scripts (Alessandro) is adopted as the job payload, with some modifications to cover OSG environment and use cases
 - It installs, validates, tags and cleans up ATLAS base kits and caches
 - Connects and publishes installation status to EGEE installation portal
 - EGEE installation DB is extended to be multi-grid enabled
 - Manually entered the existing releases' information on OSG sites to EGEE installation DB.
- Installation job definition
 - PanDA team provides a standard API to define and insert installation jobs to PanDA server



New system in action

- Pacballs at BNL T1
 - `$> dq2-list-dataset-site -n "sit09*" BNL-OSG2_DATADISK`
sit09.000001.AtlasSWRelease.PAC.v14022507
sit09.000001.AtlasSWRelease.PAC.v14050200
sit09.000001.AtlasSWRelease.PAC.v14050201
sit09.000001.AtlasSWRelease.PAC.v14050202
- Job definition
 - `$> python installSW.py --sitename BU_Atlas_Tier2o_Install --pacball sit08.000001.AtlasSWRelease.PAC.v14020001:AtlasProduction_14_2_0_1_i686_slc4_gcc34_opt.time-2008-07-07-14-19-25.md5-238cc278d67104657de35b5dfffa87bf8.sh --gridname OSG --release 14.2.0.1 --project AtlasProduction --arch _i686_slc4_gcc34 --reqtype installation --resource BU_Atlas_Tier2o_Install --cename BU_Atlas_Tier2o_Install --sitename BU_Atlas_Tier2o_Install --physpath /atlasgrid/Grid3-app/installtest/atlas_app/atlas_rel/14.2.0 --reltype patch --pacman 3.26 --jobinfo 14.2.0.1-validation.xml --tag VO-atlas-AtlasProduction-14.2.0.1-i686-slc4-gcc34-opt --requires 14.2.0`



New system in action

PanDA monitor for job status

Panda monitor and browser - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://panda.cern.ch:25880/server/pandamon/query?job=~&prodSourceLabel=software

Red Hat, Inc. Red Hat Network Support Shop Products Training

Panda monitor
Times are in UTC

Panda job information

Panda info and help [Click for help](#)

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

Quick search
Job
Dataset
Task request
Task status
File

Summaries
Blocks: 0 days
Errors: 0 days
Nodes: 0 days
Daily usage

Tasks - search
Generic Task Req
EvGen Task Req
CTBsim Task Req
Task list
New Tag
Bug Report

Datasets - search
Aborted MC datasets
Panda subscriptions

Datasets Distribution
DDM Req
Req list
AODs
EVNTs
Conditions DS
DB Releases
SIT pacballs
Validation Samples
Functional Tests
ATLAS Data
EDR_Datasets
Processed Datasets

Summary of all jobs for the last 1 days in any state at any site

22 jobs. Click job number to see details.
States: defined:0 assigned:0 waiting:0 activated:4 running:0 transferring:7 holding:0 finished:11 failed:0
Users: Xin Zhao:22
Sites: AGLT2_Install:2 BNL_ATLAS_Install:2 BU_Atlas_Tier2o_Install:2 HU_ATLAS_Tier2_Install:2 IU_OSG_Install:2 MWT2_IU_Install:2 MWT2_UC_Install:2 OU_OCHEP_SWT2_Install:2 SWT2_CPB_Install:2 UC_ATLAS_MWT2_Install:2 WT2_Install:2

Showing 22 jobs modified from 2009-03-13 17:01:48 to 2009-03-13 15:59:55

PandaID, Owner	Job	Status	Created	Time to start	Duration	Ended/Modified	Cloud/Site, Type	Priority
28698488 Xin Zhao	pathena jobID=1963 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.BU_Atlas_Tier2o_Install_2cb87c8b-0233-49b7-86b4-c1ba486bc479	finished	03-13 16:32	0:07:43	0:21:22	03-13 17:01	US/BU_Atlas_Tier2o_Install, software	10000
28698487 Xin Zhao	pathena jobID=1962 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.OU_OCHEP_SWT2_Install_17f31812-3086-4458-b7de-abc503bc22a8	finished	03-13 16:32	0:37:47	0:27:31	03-13 17:38	US/OU_OCHEP_SWT2_Install, software	10000
28698486 Xin Zhao	pathena jobID=1962 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.UC_ATLAS_MWT2_Install_71a12aed-3494-4d96-893a-07decc5f2814	activated	03-13 16:32	0:05:09		03-13 16:37	US/UC_ATLAS_MWT2_Install, software	10000
28698485 Xin Zhao	pathena jobID=1956 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.SWT2_CPB_Install_24187500-1287-4756-a0b9-43abb56104f9	finished	03-13 16:32	0:07:48	0:04:13	03-13 16:44	US/SWT2_CPB_Install, software	10000
28698483 Xin Zhao	pathena jobID=1951 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.MWT2_IU_Install_51017d41-73f7-4b7c-9749-5fa2c7108e9	transferring	03-13 16:32	3:38:21	0:04:10	03-13 20:27	US/MWT2_IU_Install, software	10000
28698481 Xin Zhao	pathena jobID=1945 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.MWT2_UC_Install_5f000fb8-2934-4890-ae90-e3a8521bb6ea	finished	03-13 16:32	0:08:15	0:06:13	03-13 16:46	US/MWT2_UC_Install, software	10000
28698478 Xin Zhao	pathena jobID=1939 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.IU_OSG_Install_47f616f7-013f-4418-bb32-b7d9bb0c12b4	transferring	03-13 16:32	4:21:00	0:16:24	03-13 21:21	US/IU_OSG_Install, software	10000
28698477 Xin Zhao	pathena jobID=1934 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.HU_ATLAS_Tier2_Install_7dbccb15-0378-49ea-a149-a93f70da8331	activated	03-13 16:32	0:05:19		03-13 16:37	US/HU_ATLAS_Tier2_Install, software	10000
28698476 Xin Zhao	pathena jobID=1928 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.BNL_ATLAS_Install_84df406c-2852-488e-b74b-0ca73425e49	finished	03-13 16:32	0:12:23	0:04:14	03-13 16:48	US/BNL_ATLAS_Install, software	10000
28698475 Xin Zhao	pathena jobID=1923 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.WT2_Install_2fcc5a30-11fa-4100-9663-29eab127226a	finished	03-13 16:32	1:20:25	0:14:22	03-13 18:06	US/WT2_Install, software	10000
28698474 Xin Zhao	pathena jobID=1917 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.AGLT2_Install_a383d4f3-36e4-497d-b35d-f72fa850d17d	finished	03-13 16:32	0:09:39	0:11:00	03-13 16:52	US/AGLT2_Install, software	10000
28692909 Xin Zhao	pathena jobID=9607 installAtlasSW In: sit09.000001.AtlasSWRelease.PAC.v14050203 Out: testpanda.BU_Atlas_Tier2o_Install_1b284b36-4714-45e0-6092-3c4b65510e5	transferring	03-13 15:53	0:06:58	5:34:32	03-13 21:44	US/BU_Atlas_Tier2o_Install, software	10000

Done

Termin Pand

Fri Mar 13 7:41 PM



New system in action

- Release installation monitoring

- [EGEE installation portal](#) shows release installation status

- One can also query the installation DB using command line tools:

- `$> curl -k -E $X509_USER_PROXY https://atlas-install.roma1.infn.it/atlas_install/exec/showtags.php -F "grid=OSG" -F "rel=14.5.2.1" -F "showsite=y"`

VO-atlas-production-14.5.2.1-i686-slc4-gcc34-opt,SWT2_CPB_Install

VO-atlas-production-14.5.2.1-i686-slc4-gcc34-opt,WT2_Install

VO-atlas-production-14.5.2.1-i686-slc4-gcc34-opt,MWT2_IU_Install

VO-atlas-production-14.5.2.1-i686-slc4-gcc34-opt,IU_OSG_Install

VO-atlas-production-14.5.2.1-i686-slc4-gcc34-opt,MWT2_UC_Install

VO-atlas-production-14.5.2.1-i686-slc4-gcc34-opt,BNL_ATLAS_Install

VO-atlas-production-14.5.2.1-i686-slc4-gcc34-opt,OU_OCHEP_SWT2_Install

VO-atlas-production-14.5.2.1-i686-slc4-gcc34-opt,AGLT2_Install

VO-atlas-production-14.5.2.1-i686-slc4-gcc34-opt,BU_Atlas_Tier2o_Install



Issues and future plans

- Integrity check
 - Due to site problems or installation job errors, there are always some sites that fall behind on installation schedule, which causes job failures.
 - It's error prone and time consuming to browse the webpage for missing installations and failed installation jobs.
 - Plan --- automation in job payload scripts
 - Every time it lands on a site, always scan the existing releases against the target list. If any software is missing, go ahead and do the installation
 - Target list could be retrieved from DQ2 central catalog by querying for the available pacball datasets
 - This way a new pacball creation will trigger the whole installation procedure for all sites and guarantee integrity to a certain level, reduce the time of daily troubleshooting



Issues and future plans

- Latency in pacball datasets replication to T1s
 - ATLAS releases are the first thing sites need to run production and analysis jobs. Production manager usually requires a new release to be installed on Grid sites in a couple of hours.
 - In the early days of running the new installation system in production mode, one major issue is DQ2 can't transfer the new pacballs to T1s promptly. Sometimes we have to use the old method to start the deployment
 - Plan
 - Make pacball creation and registration part of the standard SIT release procedure
 - Work with DDM team to increase the priority of pacball replication to T1s in DQ2



Issues and future plans

- Procedure of adding new sites to PanDA
 - Some T3 sites want the releases to be installed in the standard way as other T1/T2 sites do.
 - PanDA has its own procedure of testing and adding sites to its server
 - Installation queue in panda is always attached to an existing site in production
 - Plan
 - Work with production manager and PanDA team to work out the operational details and come up with a standard procedure
- More installation job payloads
 - Besides ATLAS releases, OSG USATLAS sites also need to have some standard supporting software packages installed and available for worker nodes, e.g. DQ2 client.
 - Plan
 - Define more job payloads for the adhoc ATLAS software packages