Open Science Grid

PanDA
Multi-User Pilot Jobs

Maxim Potekhin
Brookhaven National Laboratory
Open Science Grid

Panda Intro

Open Science Grid
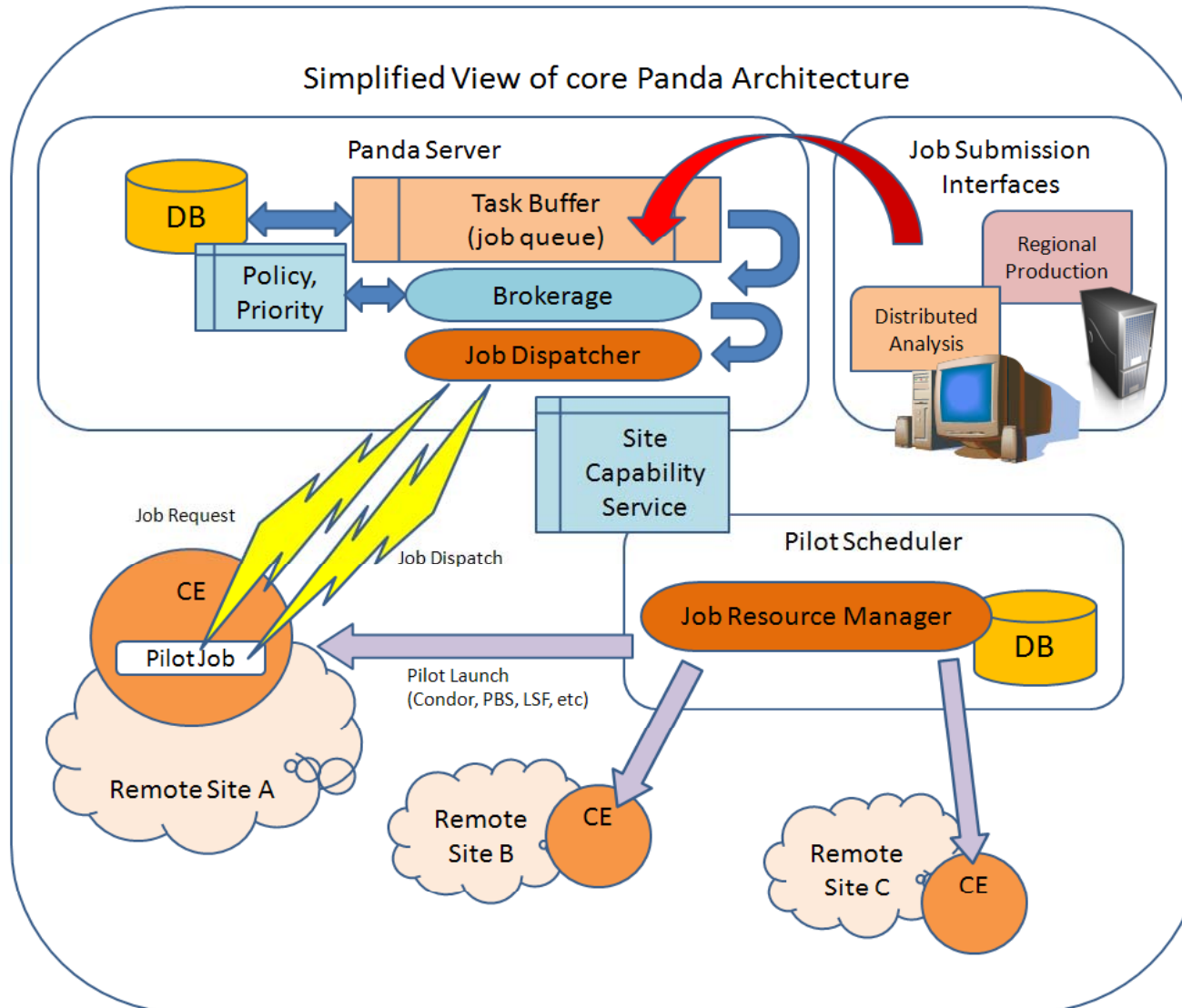
In the next slides, we present a brief overview of Panda and the current status of its Pilot-based framework. Project pages:

https://twiki.cern.ch/twiki/bin/view/Atlas/Panda

A Quick Look at Panda Architecture

Panda Multi-User Pilots: Work Areas

- ATLAS-specific development of pilot code (Paul Nilsson, Graeme Stewart, others)

- Infrastructure work related to DB migration from MySQL to Oracle and central services migration from BNL to CERN

- Monitoring (A.Thor)

- Ongoing extensions of production workflow automation (T.Maeno)

- Security enhancements according to adopted OSG/WCLG guidelines and documents (J.Caballero, M.Potekhin)
    - cf recent SCAS activity

- Support of generic data transport and other features of Panda as a workload management system for OSG (J.Caballero, M.Potekhin): using the functionality originally created for Atlas, but in simpler form and without Atlas-specific components, for use by non-HEP Virtual Organizations in OSG

Panda Multi-User Pilots: Security Overview

Open Science Grid

- PanDA services use the standard GSI grid security model of authentication and authorization based on X509 grid certificates

- Interactions with PanDA require secure https

- Proxy's VOMS attributes are checked to ensure user is a member of a VO authorized for PanDA use

- Authenticated users' DN is part of the metadata of a PanDA job, so the identity of the user is known and tracked throughout PanDA operations

- Production job execution and file management relies on production certificates, with the pilot carrying a specific 'pilot' VOMS role to control its rights

- Analysis jobs also run with a production proxy unless gLExec is employed in identity switching mode (next slide)

Panda Multi-User Pilots: Current Work on Security

Open Science Grid

•Optional gLExec based identity change on WN for Panda user jobs has been tested on US sites (proxy management done by MyProxy)

• Document produced by Atlas, titled "Answers to questions from WLCG pilot jobs security review", underwent a few revisions and can be found at:

https://twiki.cern.ch/twiki/bin/view/Atlas/PandaSecurity

• The focus of the current development and integration work is to ensure that the policies set forth in that document are implemented and fully tested:

- • since identity switch relies on proxies that can be retrieved from MyProxy server(s) by client processes, we are taking measures to secure that retrieval by using a system of tokens (keys), as explained in the document

• With identity switch taking place, there are implications for the Panda workflow automation and data movement mechanisms, which also require additional development

• Recently, the SCAS platform became available for testing of OSG/EGEE interoperability – work in progress (see previous talk by A.Retico)

- • action item – we in OSG need to maintain a testbed with same version of gLExec and other necessary elements of the software stack, to be able to efficiently test and debug the pilot software

Summary

- Panda continues to operate smoothly
- There has been significant amount of work done in the area of workflow automation and recently, in aligning security mechanisms with policies of OSG/EGEE/WLCG
- Focus now is in continued testing and interoperability effort

7