



Managed by Fermi Research Alliance, LLC for the U.S. Department of Energy Office of Science

PanDA and ProdSys II overview

Fermilab

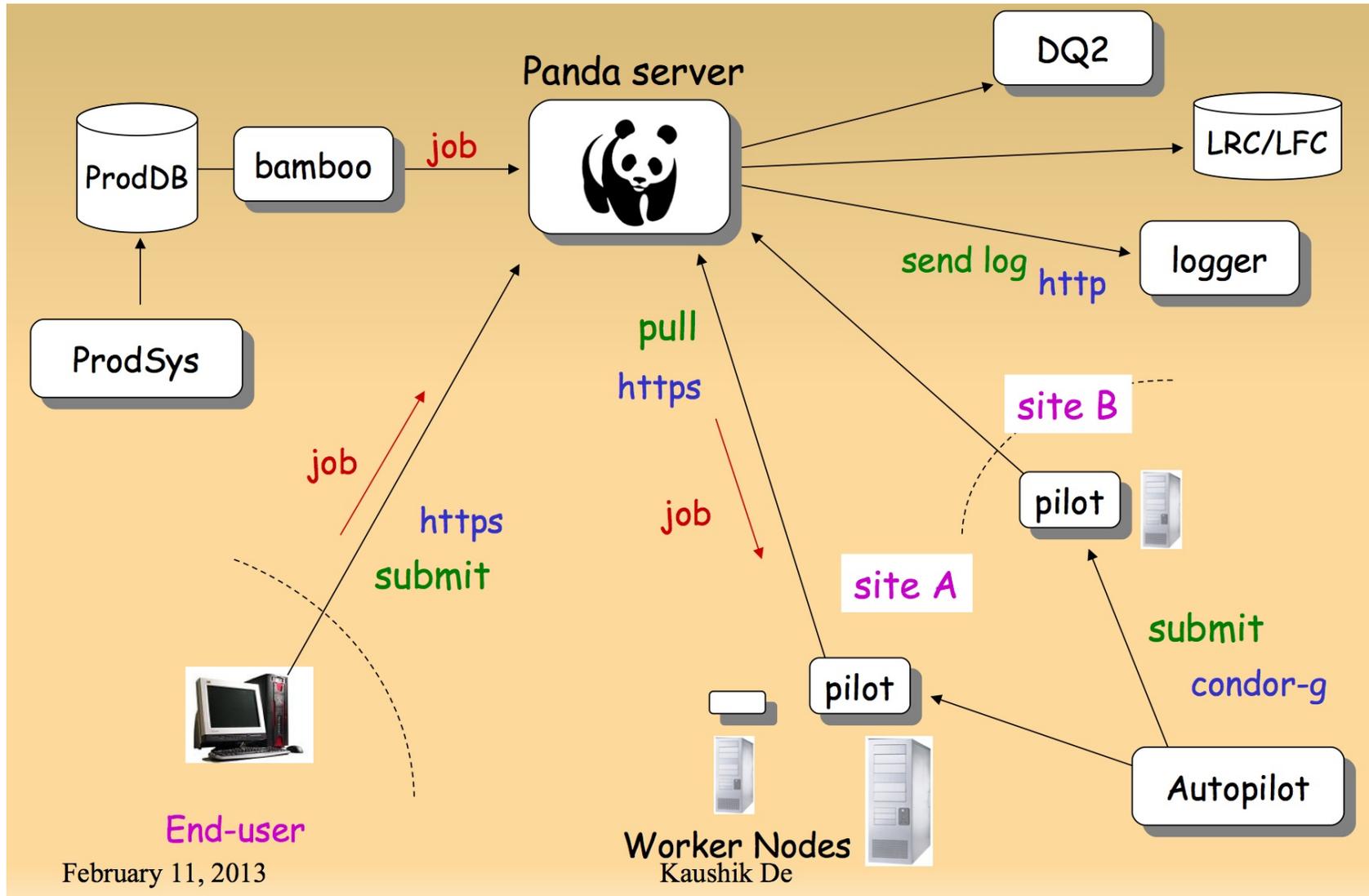
5/3/2016

Marco Mambelli

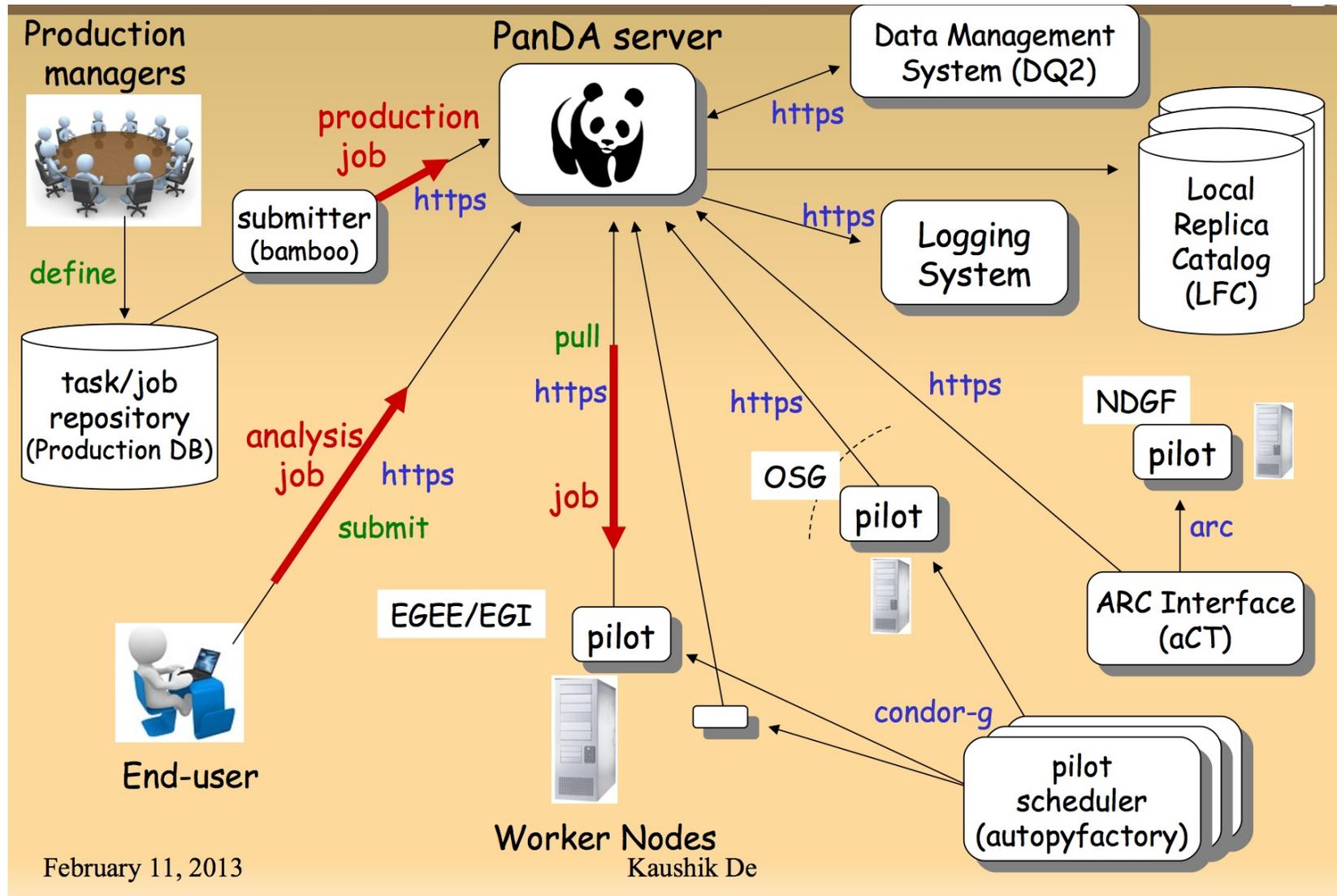
Panda ecosystem

- (Big) Panda
 - Panda server
 - Panda monitoring
 - Panda client (pathena, prun, psequencer, pbook)
- Bookkeeping – pbook
- Athena (AthenaMP, AthenaMT)
- AGIS – ATLAS Grid Information System
- Pilot
 - Plugins
 - Data movers
- DQ2 and Rucio – Data management/movement
- DEFT – Database Engine For Tasks
- JEDI – Job Execution and Definition Interface
- Yoda – ES for HPC

(Traditional) PanDA and ProdSys



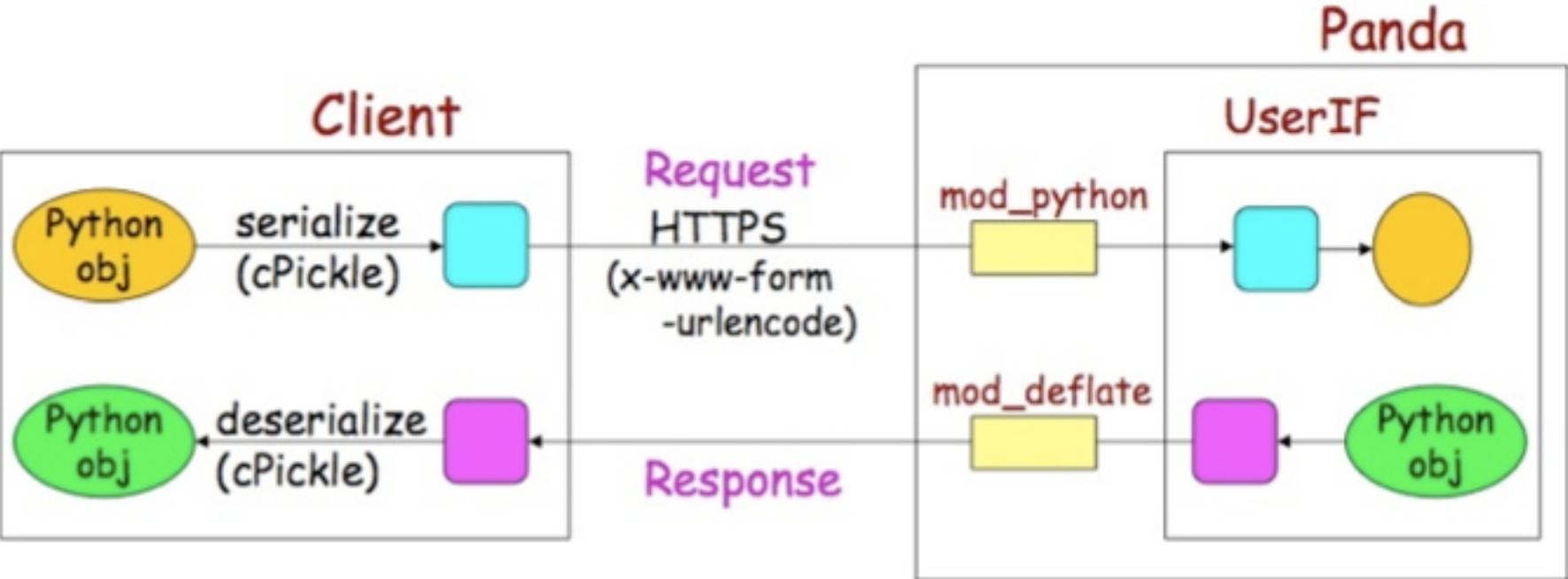
Panda Workflow management



Project status and sites

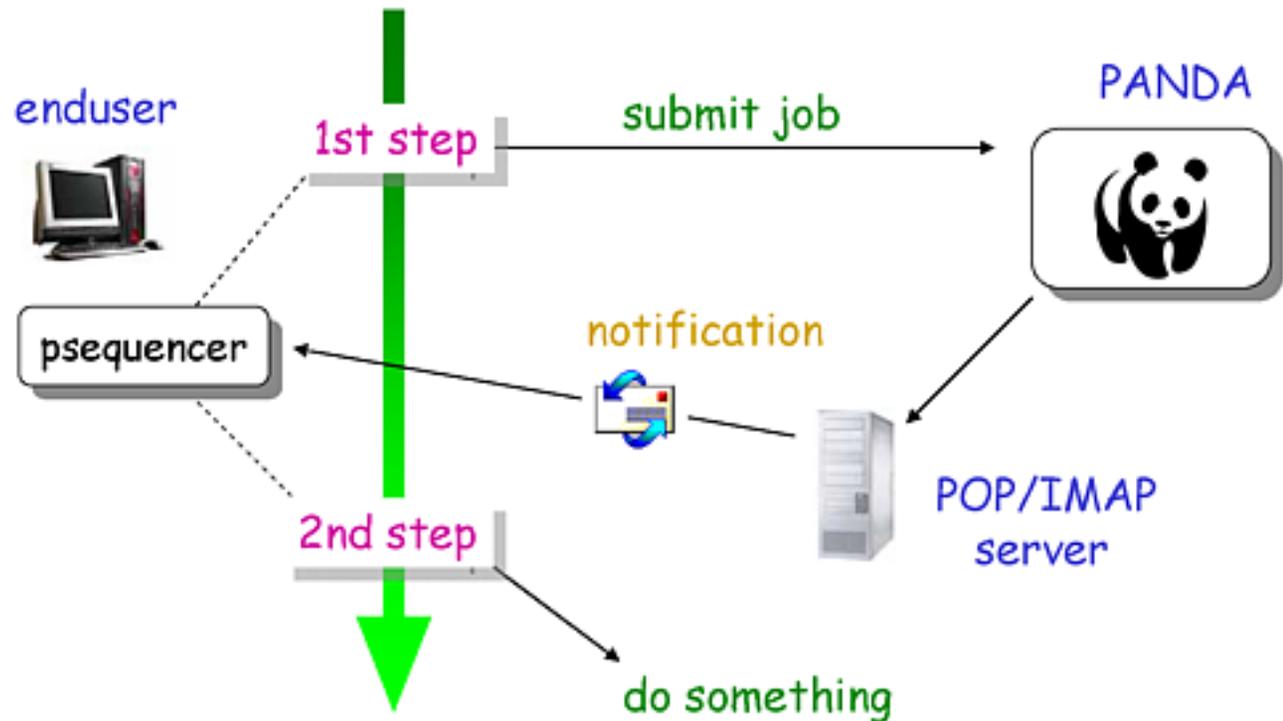
- Re-development for new ATLAS requirement and generalization using also ASCR (DOE Advanced Scientific Computing Research) grant
 - Used for ATLAS and (soon) LSST and has other experiments including CMS interested in it
 - Used in the past for CHARMS
 - <http://atlascloud.org:8080/pandawms/>
 - <http://prodsys.blogspot.com/>
- <https://twiki.cern.ch/twiki/bin/view/PanDA>
- <http://bigpanda.cern.ch/> (monitoring)
- <https://svnweb.cern.ch/trac/panda/browser/panda-server/current/pandaserver/> (old SW repository)
- <https://github.com/PanDAWMS> (SW repository)

HTTP communication



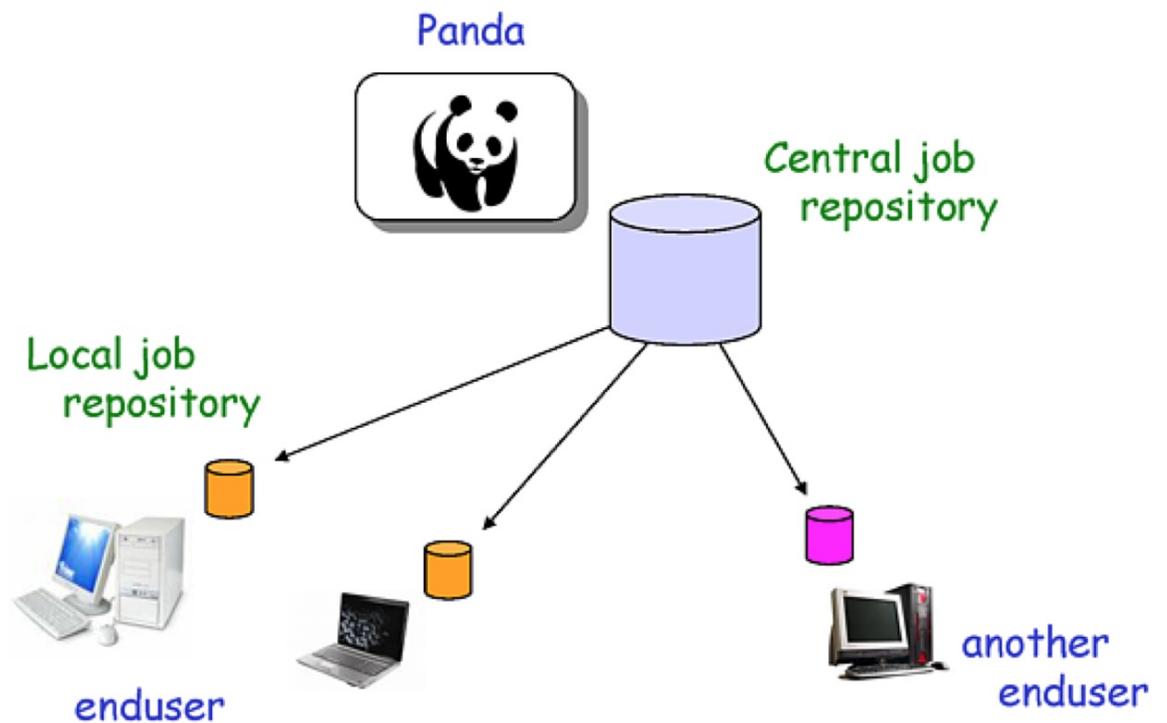
psequencer

- Chain activities using email notification



Pbook – PanDA bookkeeping

- To submit, monitor, kill jobs
- Sqlite based, client (can go away) server



AGIS – ATLAS Grid Information System

- Grid information system based on heterogeneous data sources
- https://atlas-agis.cern.ch/agis/pandaqueue/detail/CONNECT_ES/full/
- Web UI:
 - Panda object
 - Update log
 - Attributes

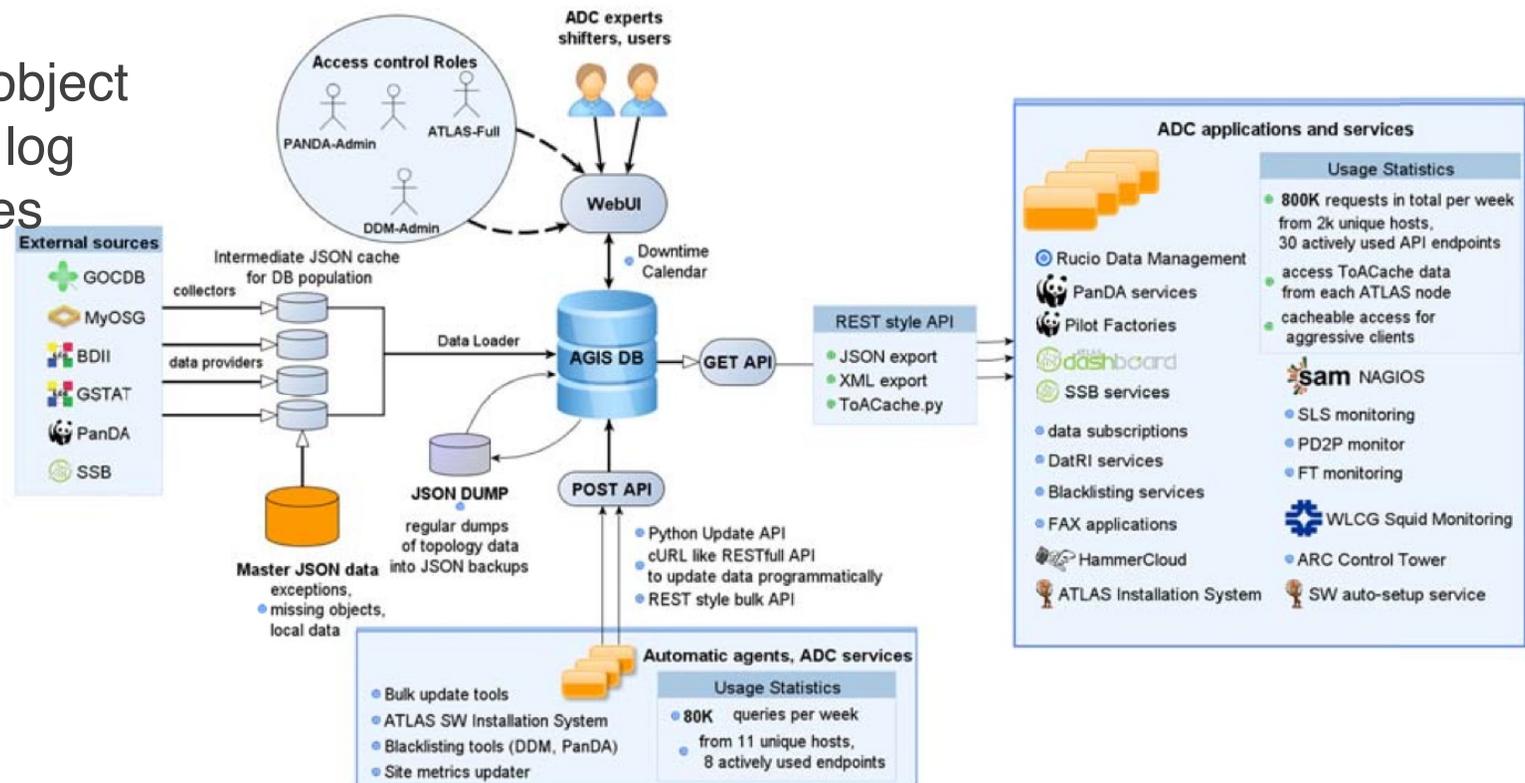


Figure 3. Schematic view of the AGIS client-server architecture.

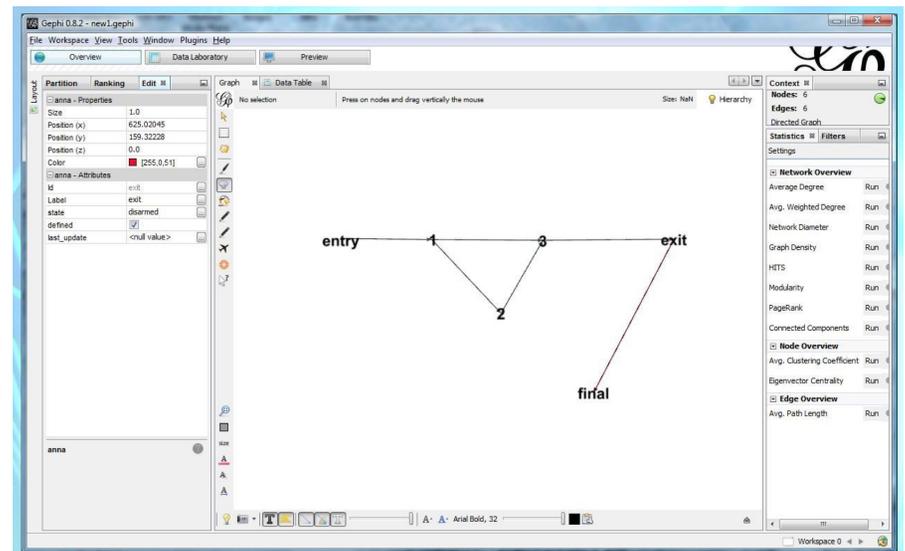
- <http://iopscience.iop.org/article/10.1088/1742-6596/664/6/062001/pdf>

Task, Job, Event

- Task: a collection of jobs. Is used as a main unit of computation
- Job: small data processing tasks. Splitting of task into jobs is similar to splitting in packets for networks.
 - Users do not care about transient job failures (packet drop)
 - Completion time and low overall loss ($O(10^{-8})$ event loss) are the metrics
- Event: smallest unit in ATLAS data and processing
- Meta-Task: task composed by multiple tasks (sequence/chain, parallel/bag, generic/graph)

DEFT – Database Engine For Tasks

- Platform-neutral database engine for bookkeeping of tasks
 - Recovery
 - Speedup & force finish
 - Forking & Extension
 - Transient & Intermediate
- (Meta-)Task definition
- Bookkeeping
- Tools to ease definition/editing



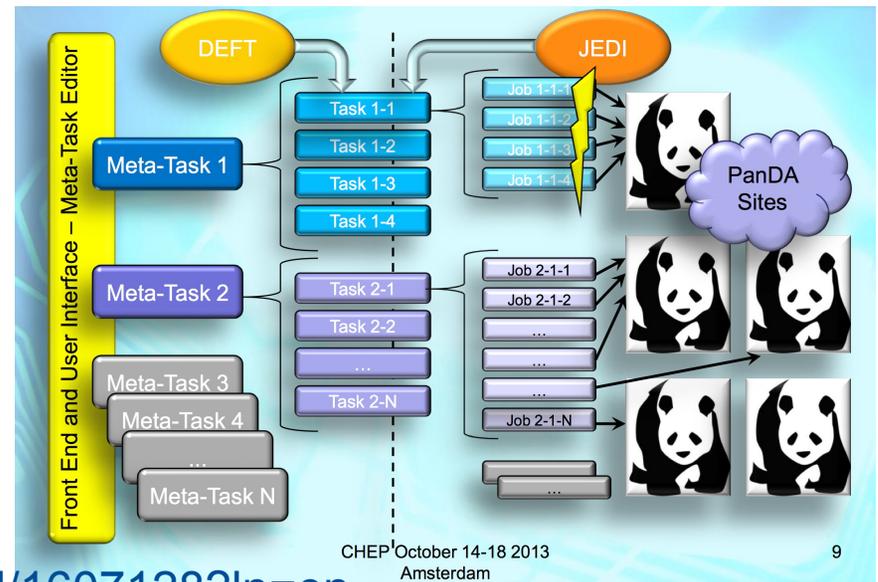
CHEP October 14-18 2013
Amsterdam

29

CHEP 2014 - <https://cds.cern.ch/record/1607128?ln=en>

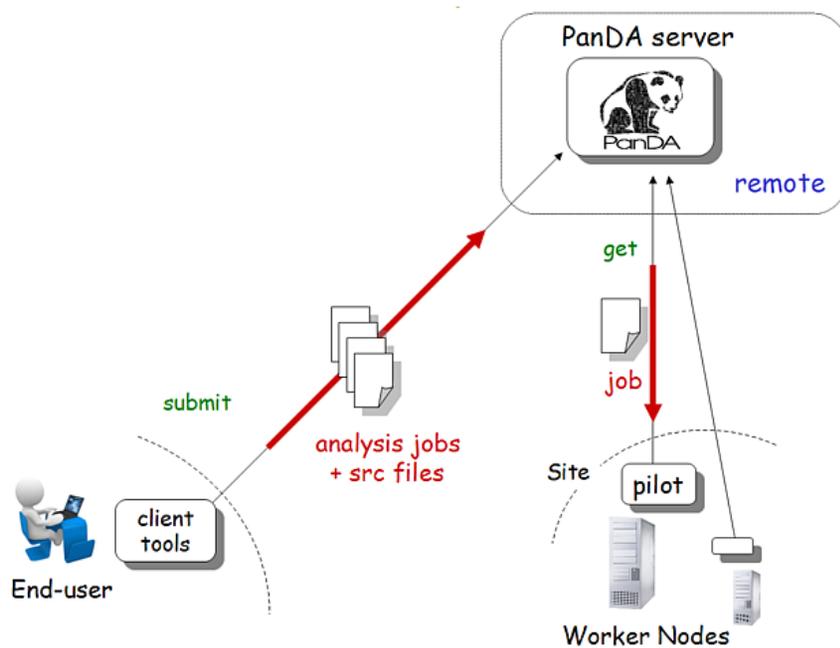
JEDI - Job Execution and Definition Interface

- Job Execution and Definition Interface
- Strongly integrated with all PanDA functions
- Owns the task database
- Defines jobs based on task definitions
- Dynamic job redefinition
- Jobs bookkeeping



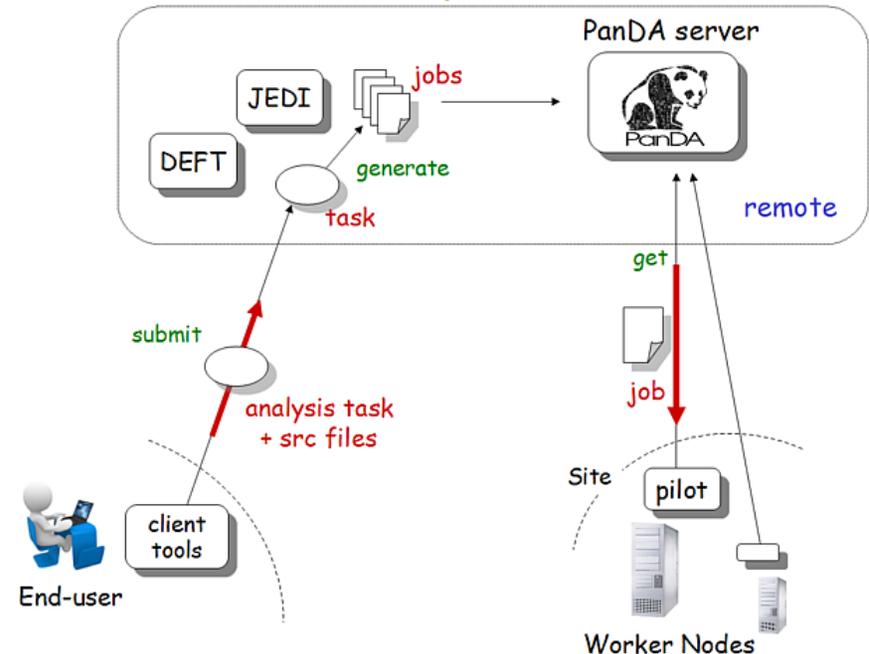
CHEP 2014 - <https://cds.cern.ch/record/1607128?ln=en>

From user to execution: ProdSys vs ProdSys II



ProdSys

End-user (production or analysis) defines jobs

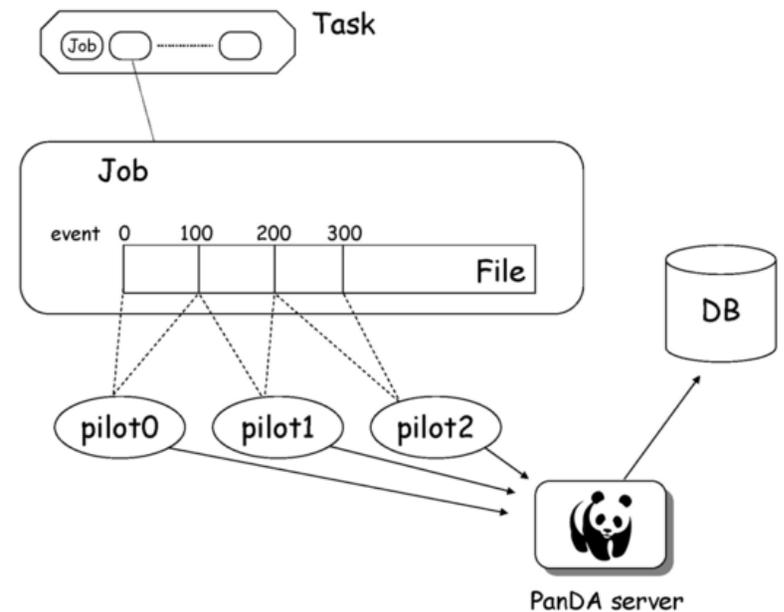


ProdSys II

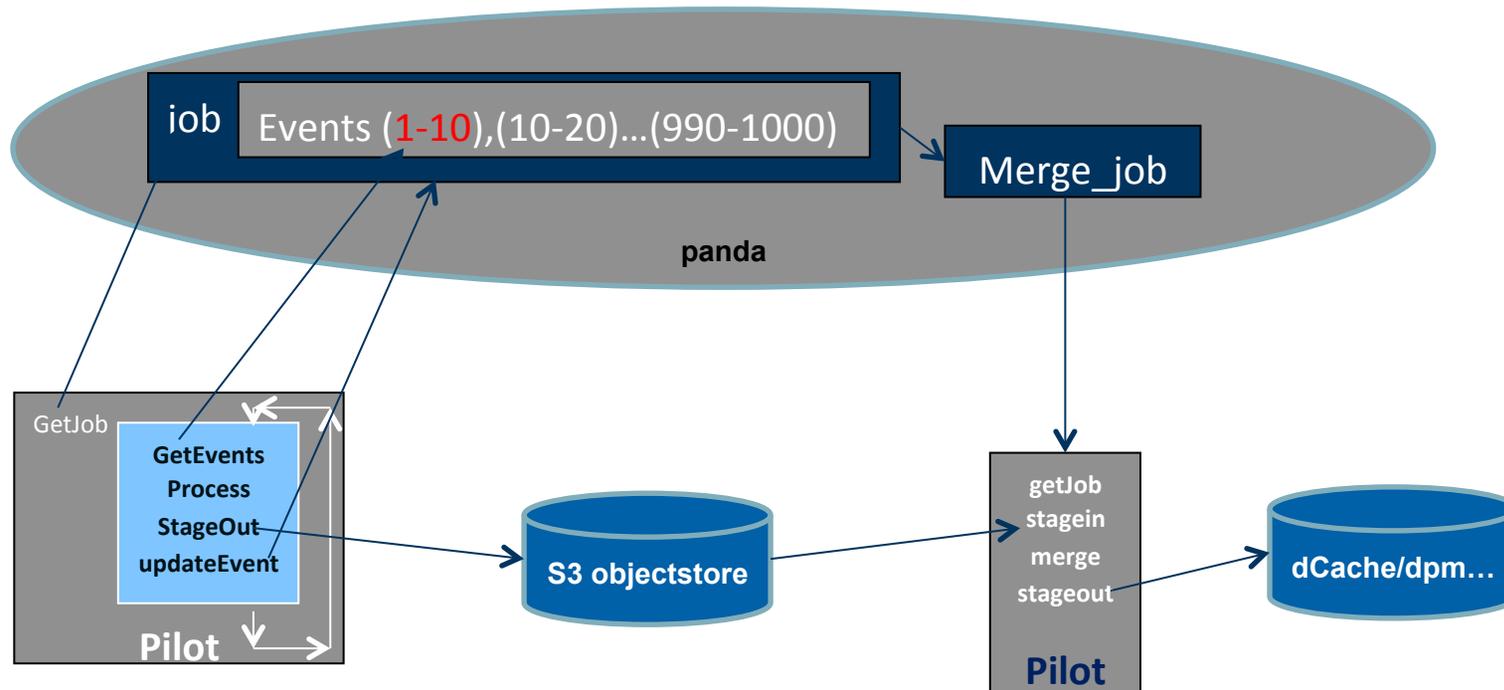
End-user defines tasks
JEDI splits the job to adapt to resources

Event Level Processing

- Job is split in many events chunks that can run separately
- Different resources have different abilities
- Pilot runs getJob to request work from Panda.
- A payload is returned from Panda which can be normal or ES (Event Service) work
- Pilot parses the payload.
- Pilot can run one or more jobs
- Pilot automatically selects different processes for different jobs.

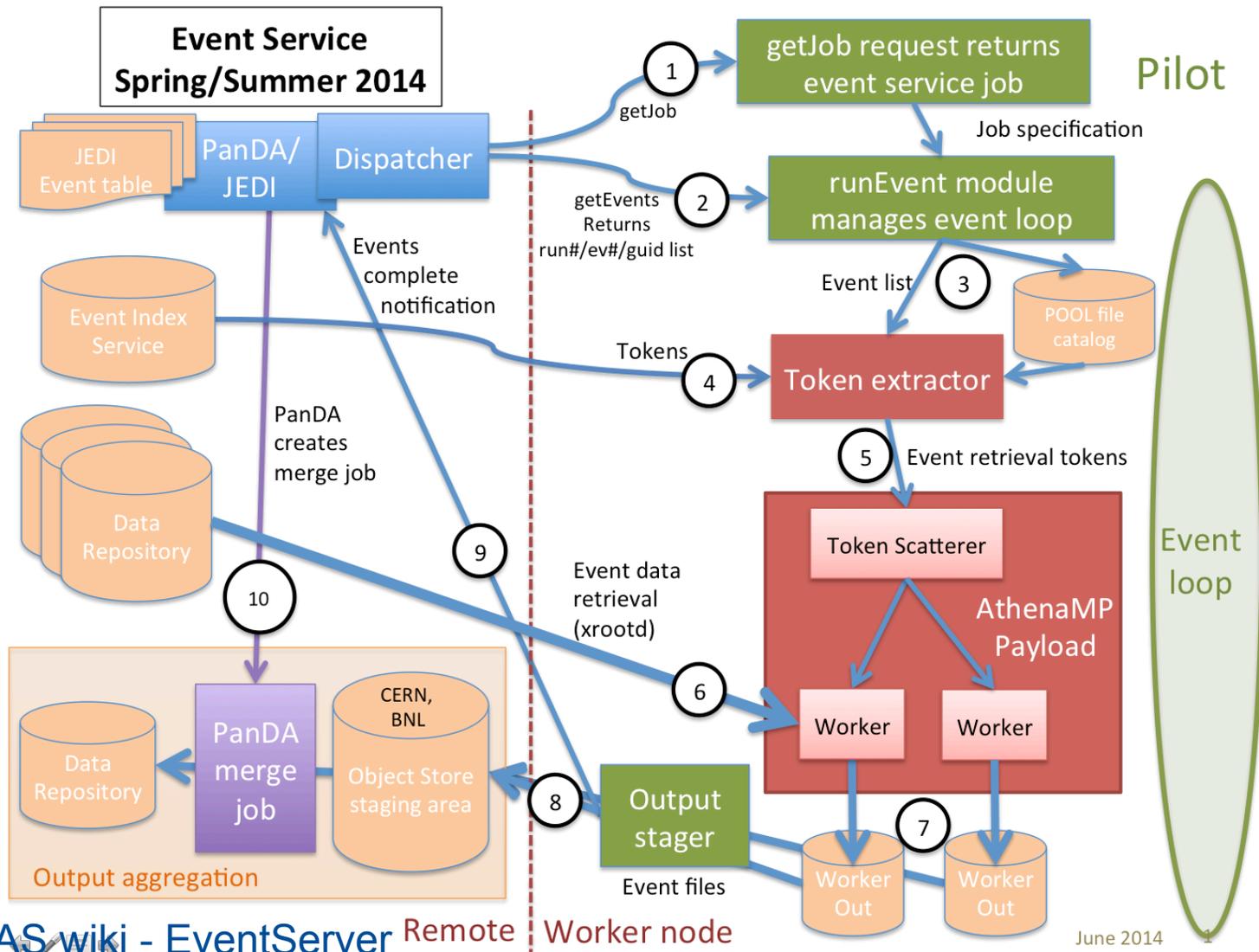


Event Service processing



- Special ES queue
- Requires Objectstore (OS, e.g. CEPH, S3, ...)

Event Service architecture



ATLAS wiki - EventServer Remote

June 2014

Yoda: Running on HPC

- Goal
 - Make use of HPC with many CPUs in one job.
 - No outbound internet connection
 - Prevents from conventional ES
- Yoda
 - Run ES as a single MPI job.
- Activity
 - Running on NERSC Edison (PBS, Slurm) for over a year
 - Running on NERSC Cori
 - Testing on ARC (Nordugrid CE)

Schematic view of YODA

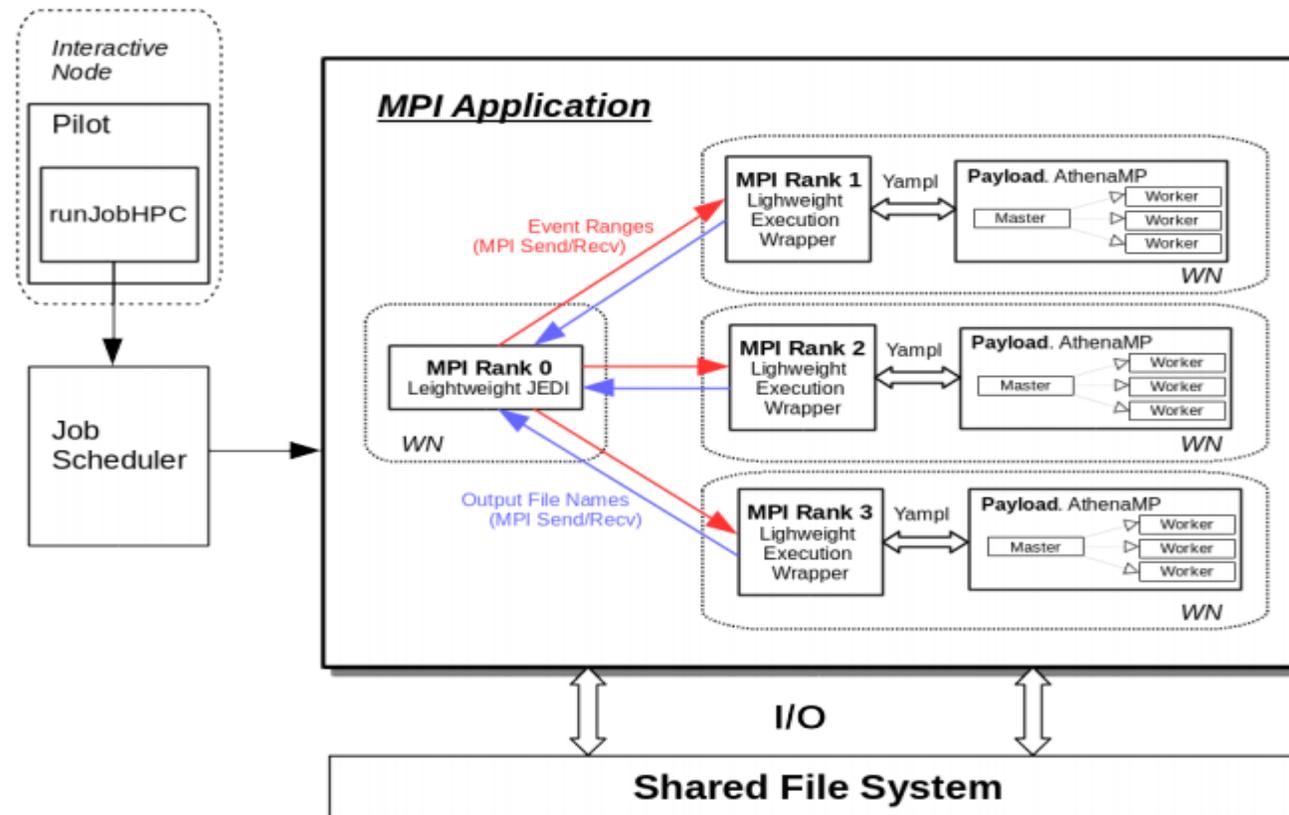


Figure 1. Schematic view of Yoda (MPI-based lightweight Event Service)

Wen Guen – ATLAS Jambore 2016