Update on workloads for HPC

Sergey Panitkin

BNL

Introduction

- I will talk mostly about current situation with workloads on Cray machines accessible for BigPanDA project

    - Titan at ORNL

    - Edison, Hopper at NERSC

- Taylor may comment on situation with Mira at ALCF

- Andrej may comment on the situation with European HPC

Some of the HPC Centers available to ATLAS

PowerPC

Minimum Job 8192 cores

MIRA

75% GPUs

TITAN

Intel Phis

STAMPEDE
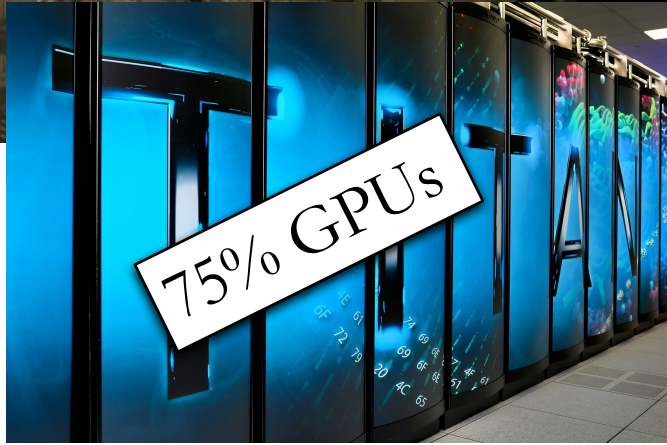
- 786,432 CPU cores
- 6.9B CPU-hours/year

- 299,008 CPU cores
- 2.6B CPU-hours/year

- 102,400 CPU cores
- 0.9B CPU-hours/year

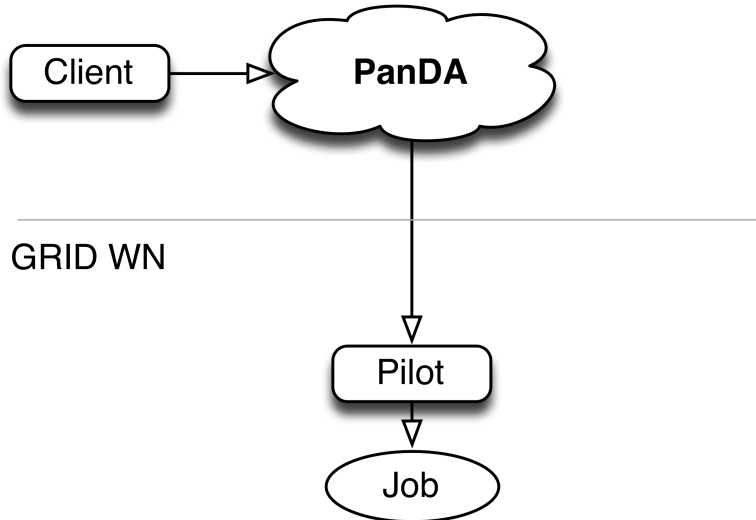These are not Grid Machines, No SLC here.
Sometimes not even x86 based !

Hydra

Xeon Phi

SuperMUC

Xeon,Phi,GPU

Xeon

archer
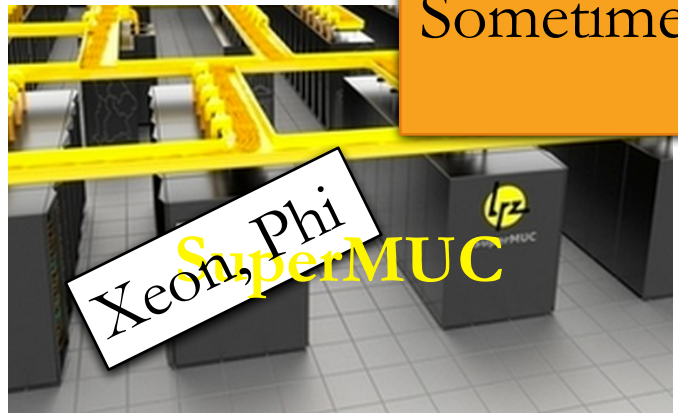
- 155,656 CPU cores
- 1.3B CPU-hours/year

- 83,000 CPU cores
- 0.7B CPU-hours/year

- 72,192 CPU cores
- 0.6B CPU-hours/year

Courtesy of Taylor Childers

† CPU-hours not made equal
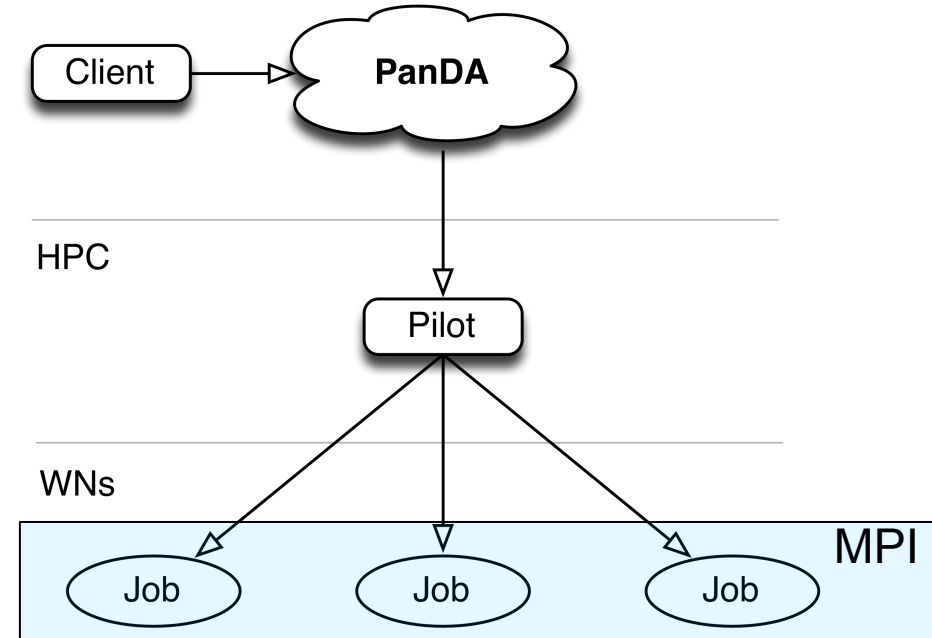
3

MPI wrapper for workloads

In BigPanDA we use MPI wrapper/overlay scripts that allow to run multiple "single node" workload instances in parallel on multiple HPC nodes. We do not modify workloads.

GRID Behavior

Client → PanDA

GRID WN

Pilot → Job

"One to One"

HPC Behavior

Client → PanDA

HPC

Pilot

WNs

MPI

Job   Job   Job

"One to Many"

Standalone workloads on Titan

- Several standalone workloads were ported to Titan

  - Good binary compatibility with Grid, system incompatibility needs to be worked around

- Root,etc

  - Root based ATLAS analysis

  - Limits setting code (aTGC)

- Event generators ported

  - SHERPA (v. 2.0.b2 and v. 1.4.3)

  - MadGraph 5 (v. 1.5.12)

  - ALPGEN v 1.4

  - Powheg

  - Simple examples and tutorials for EvGens run

- Geant 4, including multithreaded v4.10

- Full GEANT simulation chains for ALICE@LHC and EIC@RHIC tested

Situation with CVMFS on HPC

- On the Grid we use CVMFS with Fuse kernel module installed on a host machine

- On HPC machines installation of kernel module is often out of question

- CVMFS via Parrot works on Titan login nodes. How to expose them to WN?

  - Cvmfs copy to shared file system. Worked nicely for ALICE

  - Didn't work for full ATLAS repository (copy takes long time, lost sessions, etc)

  - Parrot instability . See Rob Gardner's talk from August workshop a Berkeley

  - Working with CVMFS team to understand and resolve the issue

- Since we position (at least now) HPCs as simulation platforms there is no real pressing need for CVMFS installation on supercomputers.

  - We use limited number of releases for simulations

  - Many simulations are standalone

ATLAS workloads on Titan I

- Needed other mechanism to expose ATLAS software to worker nodes (WN)

- Looked at individual releases installation using pacman

  - Worked very well !

  - Releases, 17.2.12, 17.2.11.15, 18.9.0, 19.2.0 installed on Titan shared file system

  - Many thanks to Grigory Rybkin for help with details of release installations

  - Many thanks to Vakho Tsulaia for help with missing libraries on HPC and DBRelease update procedure!

- Highly scriptable procedure. Can be executed by Pilot via special PanDA queue (see Danila's talk)

- Tested Athena "Hello World" on Titan's WNs (works at NERSC too)

  - Single threaded Athena on a single WN

  - AthenaMP on a single WN with 16 threads

  - MPI wrapped single and multi-threaded Athena on multiple WN

ATLAS workloads on Titan II

- We started collaboration with group of Prof. Rostislav Konoplich from Manhattan College and NYU

- They are interested in running several large scale simulations for their ATLAS analysis

- Ported, stand alone, custom MadGraph5 (MG5_aMC_v2.1.2) to Titan
    - Ran several processes with MG5 on small scale for validation purposes
        - pp_X2pmin_ZZ_4l_0j1j2j at 14 TeV
        - pp_X2pmin_ZZ_4l_0j1j at 14 TeV and 8 TeV
    - Pythia showering in Athena
    - Reco.trf in Athena

- Tested the same chain with Powheg

- Dimitriy Krasnopertsev, a PhD student from MEPhI, started physics validation under Prof. Konoplich guidance.

- Initial look at generated data from all steps looks very similar to samples generated on the Grid

ATLAS Workloads on Titan III

- Receive request from Prof. Konoplich to run large scale H->ZZ->4l sample for VBF analysis.
    - Carbon copy of their private simulation chain on the Grid with some optimizations.
    - Standalone Powheg, No Pythia showering yet, but were asked to implement this step too
    - Final goal is to generate ~100M events
    - As of now delivered ~16M events for initial evaluation/validation
- Had several discussions with FTK simulation group
    - Computationally intensive
    - Possibility to optimize for GPU. Very interesting possibility for Titan
- FTKSim Standalone was ported to Titan ;ast week. Waiting for further instructions from FTKSim group
- Talking with people from SUSY group about common project on Titan

Conclusions

- We are in pretty good shape with ATLAS specific workloads on current Cray machines

- Good integration with PanDA pilot for HPC

- Can run HEP workloads as MPI groups on HPC

- Can install any ATLAS release necessary

- Ran several EvGen simulation chains

- Started collaboration with several ATLAS groups
    - Helps with quick validation of results from HPC

- Would be interested in running more ATLAS workloads