HTCondor Integration

Laurence Field
CERN IT
Computing Workflow

Offline Simulation w/GEANT4

Online trigger and filtering

Event simulation

Offline Reconstruction

Selection & reconstruction

Event reprocessing

Offline Analysis

Batch physics analysis

Processed Data (Active tapes)

Analysis objects (extracted by physics topic)

Interactive analysis

Event summary data
WLCG Computing Infrastructure

- Global data movement: 15 GB/s
- 1.5 PB/week recorded
- 2-3 GB/s from CERN
- 170 sites, ~8000 users
- nearly 40 countries

- 2 M jobs / day
- 250 000 CPU days/day
- 200PB Storage

* Slide from 2014
The Challenges

- Running HEP Software on Windows
  - ~85% of the volunteers
- Seamlessly integrating with the existing workflow
  - Experiment specific infrastructure and services
  - Trusted and untrusted environments
  - Low prioritization
    - Data taking and analysis is always higher
    - Focus on resources at stake rather than unfulfilled potential
- Reduce the overall operations cost
  - Build upon existing tooling and support structures
CernVM and CVMFS

- Small image size but need to “bake” the images to reduce unnecessary downloads
HTCondor

- Open Source batch system from the University of Wisconsin
  - Focus on High Throughput Computing
- Symmetric matching of job requests to resources
  - Using ClassAds of job requirements and machine resources
- Long history in HEP and elsewhere
  - Used extensively in OSG
  - Also for the CMS global pool (160K+ cores)
  - CERN currently migrating from LSF to HTCondor
    - Build upon existing expertise and operational support infrastructure
- Use HTCondor with BOINC to implement the Vacuum model
  - Overlay a batch system upon autonomous elastic resources
HTCondor Model

- Test4Theory
- Schedd
- Collector
- Negotiator
- Broker
- VMS
- Startd
- AccountingGroup
- Submit side
- Send jobs to reserved pslot
- Pull list of idle jobs
- Send machine properties (ClassAds)
- Execute Side
Sandboxing and Authentication

**Trusted Domain**

- IGTF CA
- Request Certificate
- Grid Service
- Submit Job
- Data Transfer

**Untrusted Domain**

- Online CA
- VCCS*
- GET Proxy
- VM
- Get Job
- Data Transfer

**Grid**

*Volunteer Computing Credential Service*
The DataBridge

http://svnweb.cern.ch/trac/lcgdm/wiki/Dynafeds
The Solution

Application Server

Job Manager
- condor_submit
- gfal-copy

Common Infrastructure

Condor

DataBridge
- DynaFed
- S3

Volunteer’s machine

Volunteer
- Instant Glidein
- VM
- VBoxwrapper

GET Proxy
Join Pool
Data IO

Grid

FTS
Direct Submission For Sixtrack

- Simplify for the project scientists
  - HTCondor interface for both batch and BOINC
  - Reduce BOINC specific knowledge
  - DAGs to seamless use both together
- All jobs going through the same system
  - Common monitoring and accounting
- Simplify for the service managers
  - Same skills set required
FifeMon
What Do We Care About?

- BOINC submission RPC
  - Surprised this is not the default approach
- HTCondor integration
  - Both approaches
  - RPC and GAHP scalability
- A platform rather than a project
  - Not all scientists are computer scientists
    - Not all computer scientists are computer engineers
- boinc_gahp
- vboxwrapper
Summary

• HTCondor plays a key role in HEP
  • Seamlessly integration
  • Common skill set
• Provided a few complementary solutions
  • CernVM and CVMFS
    • The original vboxwrapper
  • VCCS
    • Create an x509 credential from a BOINC credential
  • DataBridge
    • Authentication overlay for CEPH
• Using two different approaches
  • An elastic batch system
  • Job routing