Computing Operations and around

Christoph Wissing (DESY) & Christoph Paus (MIT)
For CMS Computing Operations

Operations Intelligence Forum
April 2019

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

Operations related Entities

This is **not** the official Computing organization chart!

Computing Operations

- Tier-0 Ops
- P&R
  All central processing
- Data Management
  Phedex Ops
  DDM Ops
- Integration & Tools

Distributed Analysis

- CRAB Ops &Dev
- User Support

Workflow & Data Management Dev

- WM Dev
- Monitoring Dev

CMS Rucio Team

Data Management

Facility Services

- Support for Grid Sites
- VO box infrastructure
- Frontier/Squid

Dynamic Resource Provisioning

- Access to HPC
- Volunteer resources (BOINC)
- Inclusion of Clouds

Submission Infrastructure

- CMS Global Pool
- Factory Ops
- Interface to HTCondor Dev

General Remarks

Many areas involved in Operations related tasks – O(100) individuals

Large spectrum of resources

- CERN
  - Tier-0 for PromptRECO and RAW archiving, but also general processing
  - CMS HLT for processing, dynamically as data taking allows
- Grid sites
  - Over 50 classical sites: 7 Tier-1s, other sites varying a lot in size and reliability
- Dynamic resources
  - Diverse kind of resources: academic and commercial Clouds, various HPC sites

Communication

- Meetings
  - CompOps: 1h (sharp!) weekly meeting, weekly team chats, weekly team leader/coordinator chat
  - Most other areas: weekly meeting and/or weekly team meetings
- Many tools in regular use
  - Ticket systems: GGUS (main for site communication), JIRA (a lot for P&R)
  - Slack, ELOG, HyperNews, e-groups, individual IM, e-mail, F2F….

Distributed Analysis

Main tool: CRAB3

- Allows users to run jobs on the Grid
- Wide spectrum of applications
  - Private MC production
  - Production of thin trees/n-tuples for analysis
- Tool is rather mature and receives mainly fixes and very little feature enhancements

Support

- Only little personpower available
  - Basically 2 operators for operations + small improvements supervised & supported by part time seniors
- Community support via HyperNews
  - Large spectrum of issues
    - From: users really finding a problem
    - To: All jobs crash while message clearly saying killed by memory watch dog

Submission Infrastructure

Main focus: CMS Global Pool

- All CMS resources should be reachable via one HTCondor pool
- Typical size ~250k+ cores
- Actually the Global Pool is a federation of a few pools with HTCondor flocking
- Identify (and overcome, if possible) limitations in scale, number of jobs, match making attributes …

Factory Operations

- Rather hand-crafted maintenance of configuration

Interface to HTCondor development

- Regular meeting with HTCondor developers
- Discuss CMS needs and feature requests

Dynamic Resource Provisioning

Make non-Grid resources usable

Volunteer resources

- Enable BOINC resources as part of the Global Pool infrastructure
- Solve challenges to allow data in- and output to/from a non-trusted infrastructure

HPC resources

- Very diverse conditions
  - Ranging from almost easy (Grid interface provided) to almost impossible (no outbound networking)
  - Integration into the data management

Cloud resources

- HLT Cloud: Dynamic scaling of CPU for processing depending on beam conditions
- HEPCloud: Extension of FNAL Grid site by commercially provided Cloud resources
- Test/integration of other academic or commercial clouds

Vacuum approach

- Pilots launched by site to join the Global Pool - DODAS

Organize or provide support for those non-Grid resources

Transparent integration into production system

Facility Services

Support the classical Grid sites

- Consult with site contacts
- Configuration or upgrade campaigns
- Large deviations regarding know-how at the sites

Readiness metrics and tools

- SAM tests
- Hammercloud
- Enable/disable sites for Production or Analysis

Chase problems

- Help to distinguish site issues from workflow issues
- Followup fixes at the sites

VOBOX infrastructure

Frontier/Squid

Computing Operations Area: P&R, T&I, Data Management

Production and Reprocessing

- All centrally managed production
- Team of ~5 handling thousands of workflow per week
- Invested in automation: Unified tool
  - Placing of input data
  - Composition of site whitelists
  - Injection into DDM
  - Various checks in the end
  - Highly configurable
- Hugh potential for further automation

Integration & Tools

- Utilities for operations
- Remote site content lister
  - Key ingredient for consistency
  - On its way into Rucio
- Operator console
  - Track operator action
  - Basis for DL network
- ML based suggestion for operator
  - Considers so far only small set of options
  - Extension of considered inputs
  - Still proof of concept

Data Management

- Ops team focused on present transfer tools
  - Phedex
  - Dynamo (DDM on top of phedex)
- Mainly debugging transfer issues
- Management of DDM rules
- AAA data federation
- Planning of larger staging campaigns
- Preparation of tape deletion campaigns
- Slowly starting with Rucio

CMS Tier-0 is also part of CompOps, but omitted here– rather CMS specific anyway

DMWM Development – Operations related items

Data Management (DM)

- Phedex & Dynamo development is frozen
- Data Bookkeeping System (DBS) and Data Aggregation Service (DAS)
- CMS Rucio team presently mainly connected to development, growing towards operations

Monitoring

- Consolidate monitoring towards MONIT infrastructure
- Build new sophisticated tools employing analytix

Workflow Management (WM)

- WMCore
  - Main library for all workflow management
  - Main product: WMAgent
- Ever changing requirements for new/modified workflows

Support from tool developers is crucial for successful operations

Computing Shifts

Used to have 24x7 shifter coverage in Run1 and Run2 (reduced coverage in LS1)

- Shift personnel was usually remote, acting at day time in their home time zone
- Shift person going to various monitoring pages a few times per shift
- Observed issues should be reported to relevant operation teams, sites or experts

Abandoned shifts in LS2

- Supporting, training and supervising shifters found higher than return
- Large variation in commitment and reliability: from almost doing nothing to motivated
- Attempt to get a few of the good ones to join a operations team – moderate success only

Computing Run Coordinator (CRC)

- Experienced person on duty for one week
- Shift frequently not covered with little impact
  - Most issues covered by operators and/or relevant coordinators
- Looking into adjustment of this role during LS2

Summary and What next….

Infrastructure and requirements keep changing

Operations effort needs to adapt to changing conditions

Operations involves many entities of CMS Computing (actually even beyond) and resource providers

Support from developers & experts required for efficient operations

Steady effort to automatize

Opportunities for collaboration to do operations better and more efficient

- Areas of common tools
  - Rucio and FTS
  - MONIT infrastructure
- Distributed infrastructure
- Exchange of DL networks for error detection
- Possibly much more: No restrictions of interesting topics, but of people to address them

Closing Remark

Note from the CMS Operations room in building 8 at CERN