

2nd Workshop on adapting applications and computing services to multi-core and virtualization

Wrap-up and Final Discussion
21-22 June 2010

Scheduling Multi-core Jobs

- ▶ One core/one batch slot model is put in question
- ▶ Test of LSF conclusive. Exponential wait time!
 - Probably due to the local configuration of LSF
- ▶ Experiments:
 - CMS proposal: moving to “whole node” scheduling
 - ATLAS: schedule jobs taking over a complete computing node (or at least know how many slots out of the computing node one can take over)
 - LHCb: Support needed from the batch systems
- ▶ Small setup already exists in IT (sftcms)
- ▶ Any disagreement?

Scheduling Multi-core Jobs: additional issues

- ▶ End-to-end implementation with Grid
 - One End: The batch scheduler should be prepared (e.g. special queue)
 - The other End: The Production manager/Physicist should be able to indicate that his job is using more than one core
 - Information needs to be passed to all layers...
- ▶ Accounting (memory, cpu)
- ▶ Handling bigger files (bigger jobs → bigger files)
 - Services like FTS needs to be adapted

Virtualization in General

- ▶ Good in general
 - CPU performance lost acceptable
- ▶ I/O performance problem
 - Potentially affecting everybody
 - Proposal
 - Put together some common effort to understand the issue and find solutions. All experiment participation.

Virtualization of Services

- ▶ ‘Physics servers’, VOBoxes, other services
 - Important also for Tier3 services (scale)
- ▶ So far very positive experience. Plan for July 2010 the first VOboxes.
- ▶ Any issue?
- ▶ Question: why the images need to be quattorized?

Virtualization of WN

- ▶ Virtualization of Batch farms
 - Driven by the expected flexibility and efficiency
 - Explicitly undesired by CMS
- ▶ Lxcloud starting to take shape. Very encouraging.
 - ‘virtual lxbatch’ on top of lxccloud
- ▶ Issues:
 - VM accounting. Incentive to shutdown images
- ▶ Extending local Batch with Cloud resources
 - Positive experience from KIT

VM Images

- ▶ HEPiX document prepared establishing obligations for people preparing images
 - Skeleton schema is being put in place. Contextualization
- ▶ Scale of the problem
 - How big are the images and what parts are really used?
 - How many images?
 - How often they will change?
- ▶ Experiment software part of the virtual image?
 - CVMFS is a possible solution. Can we agree?
- ▶ Why real clouds do not have these concerns?

Cluster Virtualization

- ▶ Experiments needs:
 - Alice: PROOF on Demand
 - ATLAS: Virtual Cluster Appliances
- ▶ Support from a Ixcloud-like computing services is required

Storage

- ▶ ROOT is well placed to optimize read and write on multi-cores (TTreeCache, Tbuffer merge, etc.)
- ▶ ROOT offers powerful caching capabilities
 - Issue: Cache management. Who does it?
- ▶ Storage Virtualization solution proposal
 - Xrootd + FUSE + extra pieces

CVMFS

- ▶ The CernVM FS should be officially supported service “long-term”
 - Endorsed by experiments. Can we agreed?
 - Work together to certify it security-wise
- ▶ Plan to move the “back-end infrastructure” to IT
- ▶ Piggyback on existing Frontier (or Content Delivery Network)

Data/Code Preservation

- ▶ Another use case for virtualization
- ▶ Long-term preservation of VM images?
 - Strategy: Test them often. Act when something stops working

Volunteering Computing

- ▶ Virtualization is the enabling technology that makes it possible
 - CernVM has a number of additional nice features (size, software updates, etc.)

GPU

- ▶ Should we start a common effort with the experiments to investigate/evaluate were this technology can be efficiently used for the LHC experiments?

AoB