



Contribution ID: 356

Type: oral presentation

Evaluation of containers as a virtualisation alternative for HEP workloads

Monday, April 13, 2015 5:15 PM (15 minutes)

Cloud computing enables ubiquitous, convenient and on-demand access to a shared pool of configurable computing resources that can be rapidly provisioned with minimal management effort. The flexible and scalable nature of the cloud computing model is attractive to both industry and academia. In HEP, the use of the “cloud” has become more prevalent with LHC experiments making use of standard Cloud technologies to take advantages of elastic resources in both private and commercial computing environments.

A key software technology that has eased transition to a cloud environment is the Virtual Machine (VM). VM's can be dynamically provisioned, managed and run a variety of Operating Systems tailored to user requirements. From a resource utilisation perspective however, VM's are considered a heavyweight solution. Upon instantiation a VM will contain a complete copy of an operating system and all associated services leading to an increase in resource consumption when compared to standard “bare metal” deployment. This level of virtualisation is not required by the majority of workloads processed in HEP and can lead to increases in execution time on workloads that performs intensive I/O file operations such as LHC data analysis.

An alternative solution which is gaining rapid traction within industry is containerisation. Here the Linux Kernel itself can virtualise and isolate a user-land level instance of the operating system in which applications can run. Less resources are needed compared to a VM because only shared system libraries and files needed by the application are virtualised. Additionally, as the virtualisation takes place via namespaces (a mechanism provided by the Linux Kernel giving an isolated view of a global system resource) performance is close to that of the physical hardware with minimal tuning.

In this study the use of containers will be assessed as a mechanism for running LHC experiment application payloads. Using currently available software tools (Docker) deployment strategies will be investigated by the use of a distributed WLCG Tier-2 facility as an example computing site. The relative performance of Containers and VM's when compared with native execution will be detailed using hardware benchmarking tools such as the HEPSPEC suite and Bonnie++. System-level resource profiling will also be performed on a selection of representative LHC workloads in order to determine the relative efficiency of hardware utilisation in each scenario.

As part of this deployment alternative methods of providing resources to WLCG similar to those found in “Cloud” solutions will be investigated. The integration of Containers in existing Cloud platforms (OpenStack) will be explored as well as new and emerging platforms (CoreOS, Kubernetes). Additionally, the possibility of hosting Grid Services in containers to ease Middleware deployment will also be assessed.

Primary authors: WASHBROOK, Andrew John (University of Edinburgh (GB)); ROY, Gareth Douglas (University of Glasgow (GB))

Co-authors: Prof. BRITTON, David (University of Glasgow (GB)); CROOKS, David (University of Glasgow (GB)); QIN, Gang (University of Glasgow (GB)); Dr SKIPSEY, Samuel Cadellin

Presenter: WASHBROOK, Andrew John (University of Edinburgh (GB))

Session Classification: Track 7 Session

Track Classification: Track7: Clouds and virtualization