Contribution ID: **31**   Type: **Oral presentation to parallel session**

Integration of Cloud resources in the LHCb Distributed Computing

Tuesday 15 October 2013 13:30 (22 minutes)

This contribution describes how Cloud resources have been integrated in the LHCb Distributed Computing. LHCb is using Dirac and its LHCb-specific extension LHCbDirac as an interware for its Distributed Computing. So far it was seamlessly integrating Grid resources and Computer clusters. The cloud extension of Dirac (VMDIRAC) extends it to the integration of Cloud computing infrastructures.

Several computing resource providers in the eScience environment are planning to deploy IaaS in production during 2013. VMDIRAC is able to interface to multiple types of infrastructures in commercial and institutional clouds, supported by multiple interfaces (Amazon EC2, OpenNebula, OpenStack and CloudStack). It instantiates, monitors and manages Virtual Machines running on this aggregation of Cloud resources. These VMs then create an overlay of the computing resources in the same way as pilot jobs do on the Grid: jobs submitted to the LHCbDirac infrastructure can be executed seamlessly either on standard Grid resources or on Cloud resources.

This work addresses the specifications for institutional Cloud resources proposed by HEPIX and WLCG. The WLCG Cloud approach defines an instance framework on service level basis similar to the spot instances of commercial clouds. VMDIRAC is in particular able to deal with the agreed constraints on the VM lifetime: based on default limits defined by the resources provider as well as short notice on demand requests for shutdown. It also allows to instantiate VMs running on multiple cores, under control of the VO. In a first instance, the WLCG scenario considers the static assignment of cloud slots to each VO, which is good enough for a starting point of WLCG cloud deployment, but the VMDIRAC implementation also makes provision for more dynamic, job-driven, VM management.

We describe the solution implemented by LHCb and VMDIRAC for the contextualisation of the VMs, and how job agents are instantiated on these VMs. We report on operational experience of using in production several institutional Cloud resources that are thus becoming integral part of the Distributed Computing resources used by LHCb. We present a comparison of the performance of those Cloud resources with Grid traditional resources, in particular for what concerns data access and memory footprint. An outlook is also given on optimizing the memory footprint of VMs running on multiple cores by using parallel processing applications based on GaudiMP.

Primary authors: UBEDA GARCIA, Mario (CERN); MENDEZ MUNOZ, Victor (PIC)

Co-authors: CASAJUS RAMO, Adrian (University of Barcelona (ES)); Dr TSAREGORODTSEV, Andrei (Centre National de la Recherche Scientifique (FR)); Mr STAGNI, Federico (INFN Ferrara); CLOSIER, Joel (CERN); CHARPENTIER, Philippe (CERN); GRACIANI DIAZ, Ricardo (University of Barcelona (ES)); Dr ROISER, Stefan (CERN); FERNANDEZ ALBOR, Victor Manuel (Universidade de Santiago de Compostela (ES))

Presenters: UBEDA GARCIA, Mario (CERN); MENDEZ MUNOZ, Victor (PIC)

Session Classification: Distributed Processing and Data Handling A: Infrastructure, Sites, and Virtualization

Track Classification: Distributed Processing and Data Handling A: Infrastructure, Sites, and Virtualization