International Symposium on Grids and Clouds (ISGC) 2015

Contribution ID : 18

The "Cloud Area Padovana": an OpenStack based IaaS for the INFN User Community

Thursday 19 Mar 2015 at 16:00 (00h20')

Content :

At the end of 2013 INFN Padua division and Legnaro National Laboratories (LNL) jointly started a new project aiming at expanding their grid-based computing and storage facility, mainly targeted to the needs of the big LHC experiments, with a cloud-based offering best suited to address the needs of the smaller sized physics experiments carried out by the local teams. Despite the great success of the grid model in supporting large scale HEP experiments, its adoption within smaller experiments has been in fact quite limited due to the well known lack of flexibility of the grid, e.g. in terms of: authentication/authorisation mechanisms; few OS and execution environment supported; batch-like only access to resources (no interactivity); deep learning curve for deploying/using services. This lead to a proliferation of several small computing clusters, disconnected from the grid infrastructure, owned by each experiment and fully dedicated to it, often underutilised but not enough powerful to satisfy peak usage needs concentrated in short periods (typically close to a scientific conference deadline). This scenario clearly implies low efficiency and large waste of both human and hardware resources for the data centre. The new cloud-based infrastructure aims at merging these scattered computing and storage resources in a unique facility that can serve the different experimental teams on-demand with the maximum of flexibility and elasticity made possible by the cloud paradigm. Leveraging on the longstanding experience and collaboration as LHC Tier-2 of the Padua and LNL data centres, located 10 km away but connected with a dedicated 10Gbps optical link, the "Cloud Area Padovana" has been built and put in production at the end of October 2014, after six months of pre-production operations while a couple of pilot experiments tested the capabilities of the infrastructure with real use-cases. OpenStack was chosen as Cloud Management Framework for implementing a IaaS where computing and storage resources were shared between the two data centres. However, several customisations and innovative services were added to the standard OpenStack deployment in order to address the users' needs, ensure system reliability and implement an efficient resource allocation. These concern the integration with OpenStack of authentication protocols like SAML

and OpenID in order to enable user registration and access at first via INFN-AAI and later via Italian (IDEM) and possibly other international identity federations; the implementation of advanced functionalities for the management of users and projects by the IaaS administrators; the High Availability solution adopted to implement fault-tolerance of the cloud services; the development of a fair-share resource allocation mechanism analogous to the ones available in the batch system schedulers for maximizing the usage of shared resources among concurrent users/projects. An overall description of the cloud infrastructure and its operations will be given, together with the perspective of the main scientific applications running on it.

Primary authors : Dr. VERLATO, Marco (INFN)

Co-authors : Dr. AIFTIMIEI, Cristina (INFN) ; Dr. ANDREETTO, Paolo (INFN) ; Dr. BERTOCCO, Sara (INFN) ; Dr. BIASOTTO, Massimo (INFN) ; Mr. CRESCENTE, Alberto (INFN) ; Mrs. COSTA, Fulvia (INFN) ; Dr. FANTINEL, Sergio (INFN) ; Dr. FRIZZIERO, Eric (INFN) ; Dr. GULMINI, Michele (INFN) ; Dr. MICHELOTTO, Michele (INFN) ; Dr. SGARAVATTO, Massimo (INFN) ; Dr. TRALDI, Sergio (INFN) ; Dr. ZANGRANDO, Lisa (INFN) ; Dr. FANZAGO, Federica (INFN) ; Dr. VENARUZZO, Massimo (INFN) ; Dr. DORIGO, Alvise (INFN)

Presenter : Dr. VERLATO, Marco (INFN)

Session classification : Infrastructure Cloud I

Track classification : Infrastructure Clouds and Virtualisation

Type : Oral