OPTIMAL SCHEDULING OF SCIENTIFIC APPLICATION WORKFLOWS FOR CLOUD-AUGMENTED GRID INFRASTRUCTURES

Funding Body: EPSRC/JISC
Programme: Pilot Projects in Cloud Computing
Start Date: 01/01/2011
End Date: 30/06/2011

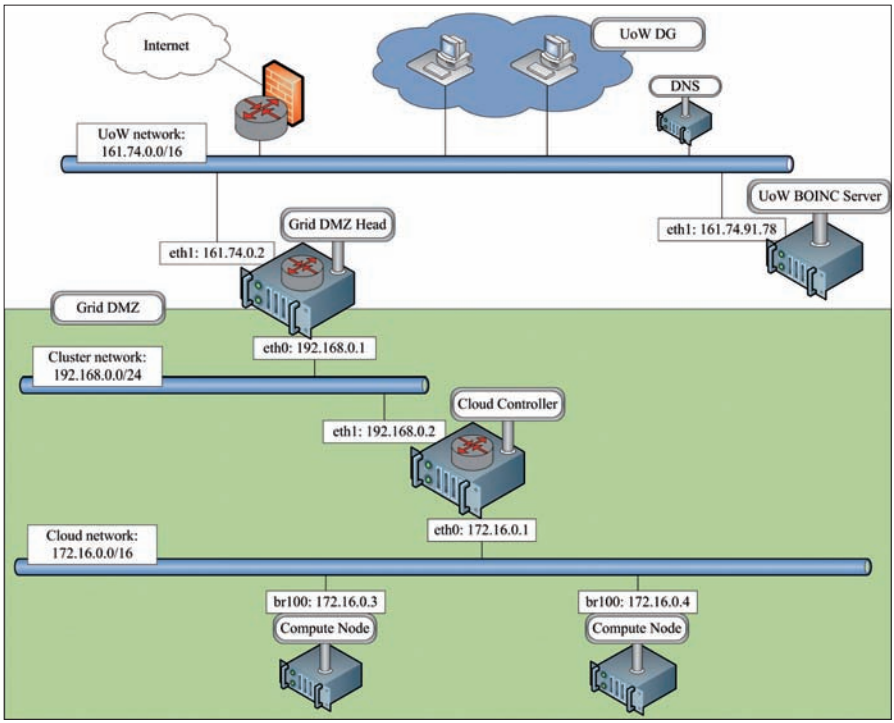
Synopsis

Research scientists need cloud computing to flexibly support their computational requirements but also require user-friendly tools in order to engage. Demanding computational requirements for science research have been addressed through grid computing, which until recently has been based on a fixed physical wide-area infrastructure. Physical grid resources can become overloaded however – scientists may find this restricts their needs where the performance required at a particular moment in time cannot be addressed in real-time. User-friendly tools and environments are required to support grid computing in science communities. P-GRADE is a web-based portal, co-developed and by the investigators, for designing, submitting and monitoring workflows on the grid. It is highly graphical, and thus very appealing to a wide range of non-computing specialists.

Familiarity is important to users. In order to develop the potential for cloud computing for the same community of science researchers, it is highly desirable, from their perspective, to retain continuity of computational support environment, as much as possible. The P-GRADE environment has been influenced by user requirements and aspirations over many years, and has already achieved a high degree of acceptance within the research science community, in view of its usability. In any case, to redesign a user interface from scratch is a very costly development activity.

Interoperability between service grids and cluster grids was achieved by the investigators in the EU FP7-funded EDGeS project, so that users of each community may now access the resources of the other. Grid computing can also be extended onto cloud resources. The investigators have recently developed solutions for extending both service and cluster grids with virtual cloud resources.

The usability of the tools and environment developed for the grid has been experimentally evaluated by the investigators in collaboration with biological research groups at the University of Westminster and Imperial College. These collaborative experiments have been very successful, in that the research biologists have found them very acceptable working environments,



and have actively adopted them for their live research programmes. The collaborations have also confirmed that the computational performance of the grid solutions was significantly enhanced. The combined effect of a high degree of usability and enhanced computational performance has been to facilitate significant shifts in biological experimental methodology, leading to increased research productivity.

Nevertheless, there are limits to the performance achievable on fixed grid implemented on physical infrastructure, due to the physical limitations of the infrastructure itself. Cloudbursting, ie. the ability to provide additional compute capacity that may be required on demand at specific times to cope with unpredictable peaks of research computing, is an attractive aspect of cloud computing which has the potential to break through the current performance constraints achievable on fixed grids. A cloud infrastructure will be attached to existing fixed grids to achieve a mixed grid infrastructure comprising both existing physical and virtualised cloud resources. The tools and environment will be re-engineered to exploit cloudbursting in the mixed grid to achieve even greater performance. Finally, the performance of the mixed physical/virtual grid will be evaluated, to include statistics on resource utilisation, to inform the development of a costed performance model.

Brief USP

Enhancing the performance of scientific DCI workflows by efficient use of cloud resources (cloudbursting).

Contact:

Professor Stephen Winter
E: s.c.winter@westminster.ac.uk