University of Huddersfield Repository

Kureshi, Ibad, Holmes, Violeta, Liang, Shuo, Gubb, D. and Cooke, D.

High performance distributed computing resources to enable e-science research

Original Citation

Kureshi, Ibad, Holmes, Violeta, Liang, Shuo, Gubb, D. and Cooke, D. (2012) High performance distributed computing resources to enable e-science research. In: Proceedings of The Queen's Diamond Jubilee Computing and Engineering Annual Researchers' Conference 2012: CEARC'12. University of Huddersfield, Huddersfield, p. 153. ISBN 978-1-86218-106-9

This version is available at http://eprints.hud.ac.uk/id/eprint/13489/

http://eprints.hud.ac.uk/

HIGH PERFORMANCE DISTRIBUTED COMPUTING RESOURCES TO ENABLE e-SCIENCE RESEARCH

I.Kureshi, V. Holmes, S. Liang, D. Gubb and D. Cooke
University of Huddersfield, Queensgate, Huddersfield HD1 3DH, UK

ABSTRACT

With pressure on Higher Educational Institutions to increase publication output, research using computational models is now being preferred to practical research. Utilizing computational models reduces researchers time to and cuts the infrastructure costs to the University (i.e. each researcher will not need expensive laboratory space and equipment). To be able to replace the bulk of research with computational models, these models would need to be able to mimic real life scenarios. This requires computational systems with either large memory or many processors or both! Standard desktops and even workstations cannot load, let alone processes such real life models as car crash testing, nuclear fuel performance modelling or weather prediction modelling. Models that would run on desktops may take months as the small repetitive tasks add up, while on a many core system these tasks can be completed in an order of hours (e.g. creating a molecular lattice by calculating bond strengths or 3D rendering).

High Performance Computing, the parent term for super computing, cluster computing and GPU computing, is the where a single monolithic system with many cores and a large quantity of RAM divides a single problem to each core in an effort to solve the model. High Throughput computing is where a problem with many discrete parts is divided across many machines which are usually loosely coupled to help reduce the processing time. In most organizations many multicore systems are purchased and allocated to staff members. These systems are usually sitting idle or being used for non intensive applications. By deploying software and utilizing these systems when idle is a cost effective way to create a HTC system. Linking these systems together and providing users with a single login interface (e.g. command line, graphical or web based), a single file transfer mechanism, a resource discovery system, a single resource selection system and a job management system, is a computational grid.

This poster will outline the results of a project undertaken at the University of Huddersfield, since October 2009, to setup a High Performance Computing (HPC) Grid. This resource serves the University's research community by providing a robust computing solution. The system which has almost reached 2 years of run time has to its credit 50+ publications and 6 PhD successes. This system is also a part of the regional North West Grid creating ties with universities like Liverpool, Lancaster, Manchester, UCLAN and research organisations like the STFC Daresbury. The HPC Grid known as the Queensgate Grid also forms a part of the National Grid Infrastructure which allows local users to scale to larger systems own by other universities.

Keywords campus grid, research computing, HPC, HTC, high performance throughput parallel processing, distributed system middleware