UCISA-Infrastructure Group Case Study

University of Reading Campus Grid

1. Introduction

The University of Reading has developed a Campus Grid to exploit the spare capacity available within the Library and Laboratory PCs, many of which are unavailable to students overnight, while providing a valuable resource for researchers who have embarrassingly parallel computation applications. The Campus Grid complements other research computing resources in the University: the THAMESBLUE supercomputer and school-based clusters.

The development of the Reading Campus Grid followed a smaller trial in the Systems Engineering department and has drawn on their research and experience.

2. Description of the work

The Condor workload management system from the University of Wisconsin is the de-facto standard for deploying cycle-savaging Grids for high-throughput computing, providing job scheduling, resource monitoring and resource monitoring and can be installed under both UNIX and Windows and can support both dedicated computational resources and desktops, where it monitors the machine usage at the terminal to run jobs when the machine is unused and moving the jobs on to other resources when the user returns.

Although Condor can be installed under both UNIX and Windows, the key issue for the Campus Grid deployment is that the IT Services department only provides Windows-based desktop PCs in its labs, while the majority of research computing is undertaken using Linux-based programs. The solution to this dilemma was developed by staff in the School of Systems Engineering - to deploy Condor within a Virtual Machine provided by the Open Source CoLinux. CoLinux is a very stable port of the Linux Kernel to Windows, with the whole Linux machine running as a Windows service and the Linux processes running unmodified directly on the host CPU.

Deployment

Once the file system image of the CoLinux system has been produced, it is packaged up into a zip file and passed on to the Windows administrators team which then push out the package and install the service using the Microsoft's Systems Management Server (SMS).

IT Services provisioned the Linux server that is to form the head-node of the Campus Grid and the login node. The Condor software was not installed directly, but as part of the Virtual Data Toolkit (VDT) from the Open Science Grid, which packages the core Grid components making upgrade and the installation and interfacing of the components straightforward.

Both the head nodes and worker nodes mount the user's home file system using NFS from the University's NetApp file server, providing the user with a common shared file system - although they are required to obtain a special account for the Campus Grid - thus limiting the use of the Campus Grid to registered users.

The CoLinux systems have their networking bridged to the host PC's network connection, with a second virtual LAN for the CoLinux hosts for communication to the head node and NetApp server, using 10.x.x.x IP addresses. These are based on the IP address of the Windows PC and so are straightforwardly configured by the CoLinux software.

Further Services

In addition to the Campus Grid's local job submission service, the Campus Grid also supports the National Grid Service's interface for the passing of jobs between Grid sites. This is based on the Globus Toolkit from the University of Chicago and available as part of the VDT package. This provides secure shell access, file transfer and job submission, authenticated by Digital Certificates issued by one of the National Academic Grid Certification Authorities, as well as an information service. This means that jobs can be submitted from the Campus Grid to other Grid systems throughout the world and jobs can be securely run from other sites or from web portals, subject to appropriate accounts.

3. Conclusion

The Campus Grid has 23 registered users and on average the Grid usage is equivalent to 55 CPUs, with peaks of up to 300 CPUs running in parallel on the same job. Without the Campus Grid making this free computing power available to researchers either much of this work would not happen or the researchers would have to buy their own dedicated resources. Providing the large, free Campus Grid server has meant that schools will not need to buy clusters for these projects which would be costly and probably still not large enough to solve the problems in a timely way.

While the capital costs of the service is low, just one relatively lightweight server, there was staff work as follows: UNIX Administrator: 2 staff days; and 1 Windows Administrator 1 day; School of Systems Engineering IT Support: 1 day and Network Administrator: 1 hour.

Future plans

The Campus Grid has been a success for many researchers in Science (e.g. Climate, Computers, Maths/Physics), who are relatively happy with the UNIX command line interface and have written their own code in a traditional programming language. On the other hand we have recognised the need to more effectively support the service for those who do not fit into this category and this means both staff effort to write detailed documentation, handhold users as they start on the Campus Grid and to deploy simpler user interfaces (for example web-portal software). In particular we wish to be able to support users or mathematical/statistics packages like Matlab and R.

4. Contact

Please direct any enquiries to: David Spence, d.r.spence@reading.ac.uk