CSIRO Remote Visualisation Capability

Justin Baker[1], Peter Tyson[2]
[1]CSIRO, Melbourne, Australia, justin.baker@csiro.au
[2]CSIRO, Clayton, Australia, peter.tyson@csiro.au

BACKGROUND

CSIRO is a large, geographically dispersed research organisation with over six-thousand staff at fifty-six sites. Like all modern research organisations, CSIRO researchers are dealing with ever larger, more complex scientific datasets.

In response to these two significant and disparate issues, geographical separation of staff and increasing data complexity, an enterprise-wide remote visualisation capability has been developed. Scientific visualisation is generally recognised as a key to addressing many data complexity issues [1]. As an adjunct to that, remote visualisation provides researchers with virtualised access to advanced visualisation hardware and software, regardless of their location.

Remote visualisation is defined here as a means of providing a local client – usually a desktop PC - with efficient virtualized access to dedicated 3D graphics hardware. Like other virtualization technologies, remote visualisation systems are based on commodity hardware and the virtualization is performed through a software layer. Different remote visualisation libraries are generally tied to specific operating systems.

REMOTE VISUALISATION SOLUTION AT CSIRO

CSIRO researchers primarily use either desktop Windows or Linux based systems for their visualisation work. For this reason, the chosen remote visualisation solution needed to support native applications for these two operating systems. A variety of remote visualisation technologies were investigated before we settled on an architecture based on the open source projects VirtualGL [2], VizStack [3] and VirtualBox [4].

VirtualGL allows most Linux based OpenGL applications to run with full 3D hardware acceleration. Once configured, the remote visualisation layer is completely transparent to running applications.

VizStack is the resource and scheduling tool used to allocate GPU and CPU resources to visualisation applications. The GPU can be located on one or more nodes and is used to support the shared remote visualisation cluster.

VirtualBox is a cross-platform virtualisation application. It allows any one of several operating systems to run as virtualized "guests" on a choice of host operating systems. VirtualBox is itself an OpenGL application and so a combination - of Windows guest on VirtualBox and Linux host - enables users to run Windows 3D applications remotely.

While a number of VNC clients are available, TurboVNC is the prefered VNC client. It has been optimised to work as a fast X proxy with VirtualGL.

REMOTE VISUALISATION ADVANTAGES

There are several significant advantages to a having a remote visualisation capability in place.

1. The remote visualisation service enables scientists to access high-end graphics hardware directly from their desktop PCs, mitigating the need to purchase dedicated visualisation workstations. CSIRO also offer dedicated visualisation workstations to researchers, and it is likely that some of these will always be needed in certain instances. However, it is anticipated that the need for these dedicated systems will diminish over time as the remote visualisation capability becomes more widely adopted, and is progressively enhanced.
2. The TurboVNC based client provides a shared/remote collaborative viewing capability enabling researchers at different locations to interact with one another's data in real-time. There are other means of sharing desktops between users, but the TurboVNC client is one of few that allow true multi-party interaction with the underlying data. This capability is likely to prove very useful in disciplines that are heavily focused on volumetric datasets.
3. The virtualised environment provides access to a wider range of applications on different platforms rather than being limited to the user's local operating system. In addition, the same applications often perform better on one operating system than another for various reasons, such as better memory management.
4. Lastly, the compute hardware is centrally managed in a data centre, and makes use of an integrated cluster based storage system minimizing the need to transfer large volumes of data between different systems. For very large datasets, it may prove completely impractical to transfer data elsewhere for visualisation anyway, in which case an integrated storage solution is the only viable option. Also, as with multi-user HPC systems, a shared

visualisation system is likely to result in more efficient use of hardware resources with lower power consumption per user.

VIRTUALGL PERFORMANCE

If researchers are encouraged to make use of VirtualGL based remote visualisation, instead of dedicated local hardware, it is understanderable they should be concerned about performance. This is especially true given the interactive nature of most visualisation applications. The overall performance of any virtualised visualisation application will be related to its patterns of usage. An application displaying high-resolution data in a large rapidly changing window, for example, will obviously make greater network demands than an application displaying lower resolution data in a small window.

Performance tests undertaken by the VirtualGL development team [5] indicated that no more than 50mbit/s of network bandwidth was required to achieve full performance at the highest quality setting, and acceptable performance could be achieved with 20 Mbit/sec bandwidth. When the quality was reduced somewhat, say by tuning the VNC client, then acceptable performance were achieved with only 10Mbit/s of bandwidth. Similar tests conducted by CSIRO verified these figures as accurate. In one internal test, a two GB volumetric dataset was rendered and the application provided quite good interactive performance over a WAN link using a 10 Mbit/sec broadband connection.

TECHNICAL ISSUES

While the combincation of VirtualGL, VizStack and VirtualBox have proven to be very effective, there are a few technical issues.

1. Not all OpenGL Windows based applications function correctly under VirtualBox. In particular, texture mapped data displays well on some applications and incorrectly on others.
2. OpenGL is a cross platform visualisation API, whereas DirectX is generally only available under Windows. While VirtualGL has proven to be very effective, it is only suited to OpenGL based applications. It provides no remote visualisation support for DirectX.

UPTAKE AND FUTURE DIRECTIONS

Following the success of the pilot, the remote visualisation system has been operational for approximately twelve months. The remote visualisation service itself now forms the core of CSIRO's Visualisation and Collaboration platform which includes very high-resolution tiled displays and stereoscopic display systems.

The user base is progressively growing and includes researchers from various domains including materials science, environmental modeling and earth sciences. An active outreach program is underway to make the capability more widely known across the organization and further increase uptake. Coupled with that, it is also hoped to broaden the remote visualisation service by providing a support function that targets visualisation applications and issues.

To provide support for Windows DirectX applications – as outlined above - we are investigating various Windows specific remote visualisation technologies such as Citrix and XenDesktop.

New hardware has recently been ordered that will be more closely integrated with some high-performance computing clusters, including the GPU cluster. This should allow a tighter workflow aligned with compute intensive jobs – such as long running simulations – as well as access to an integrated data storage system. Further down the track, it is envisaged that a kind of "computational steering" capability will enable scientists to visualize data from their long running jobs and make appropriate adjustments along the way.

REFERENCES

1. *Data Visualization is NOT useful. It's Indispensable.* Available from http://fellinlovewithdata.com/reflections/indispensable-visualization accessed June 17 2011
2. *VirtualGL,* Available from http://www.virtualgl.org/ accessed 20 June 2011
3. *VizStack,* Avaliable from http://vizstack.sourceforge.net/ accessed 20 June 2011
4. *VirtualBox,* Available from http://www.virtualbox.org/ accessed 20 June 2011
5. *A Study of the Performance of VirtualGL 2.1 and TurboVNC 0.4,* Available from http://www.virtualgl.org/pmwiki/uploads/About/vglperf21.pdf accessed June 20 2011
6. .

ABOUT THE AUTHOR(S)

Justin Baker is the manager of CSIRO's Visualisation and Collaboration Team as a part of the eResearch Program.
Peter Tyson is the principal architect of CSIRO's remote visualisation system.