White Paper

SGI on the Grid

Visual Area Networking in Grid Environments

1.0 Introduction

1.1 What Is Grid Computing?

Grid computing evolved to provide scientists and engineers quick and transparent access to critical computing resources and advanced instrumentation. A grid user can directly access computers, software, data, and other critical resources with little concern for the physical location of those resources. A grid enhances the ability to solve difficult problems, enables collaboration across organizations, and increases the utilization of expensive resources.

While distributed computation—using multiple computer systems to solve a single large problem—is an important aspect of grid computing, distributed access to unique capabilities is the driving force behind much of the ongoing grid development. Particularly in Europe, a great deal of effort has gone into creating grids for resource sharing. Connection to a grid increases system utilization, ensuring maximum return on investment for advanced computing resources.

1.2 Visualization in a Grid Environment

Until SGI developed Visual Area Networking (VAN), one of the most powerful tools for increasing understanding and insight into difficult problems—advanced visualization—was not readily sharable on a grid. The output of advanced visualization systems could only be displayed locally. Scientists and engineers could efficiently share computer systems and instrumentation without regard for location, but to visualize results they had to either purchase visualization systems for every location where visualization was required or travel to a location that already had them. Likewise, to share visual data for collaboration, either all collaborators had to travel to the same location or each collaborator had to have a local copy of the data set and access to a visualization system. Of course, having multiple copies of data increased the difficulty of ensuring data integrity, data security, and, in some cases, protection of intellectual property. All these things greatly restricted the use of visualization.

Through the development of VAN, SGI has removed these limitations, resulting in greatly improved access to advanced visualization. With VAN, grid users can access advanced visualization systems in a manner analogous to other grid resources. Using the flagship SGI® VAN product, OpenGL Vizserver™, the rendered output of SGI® Onyx® family visualization systems can be delivered over the grid for display on virtually any client system.

This paper explores the benefits of OpenGL Vizserver for remote access to visualization resources and improved visual collaboration on grids. For more information on grid computing and other unique capabilities that SGI provides to grid users, see the companion SGI white paper titled; “SGI on the Grid: Unique Capabilities for Grid Computing.”

2.0 Visual Area Networking with OpenGL Vizserver

2.1 How OpenGL Vizserver Works

With OpenGL Vizserver, engineers and scientists can remotely access visualization resources on a grid, increasing resource utilization and improving collaboration and decision making. OpenGL Vizserver delivers high-resolution graphics to client systems on the grid. No application changes are required to use OpenGL Vizserver with OpenGL® applications. A visualization server, such as an SGI® Onyx® 3000 series or SGI® Onyx® 300 visualization system, runs the application, processes the data, and renders the graphics. Graphical images are then read back from the frame buffer, compressed (if necessary), and sent across the grid to the user. Since the client system receives rendered graphics, it requires no accelerated graphics

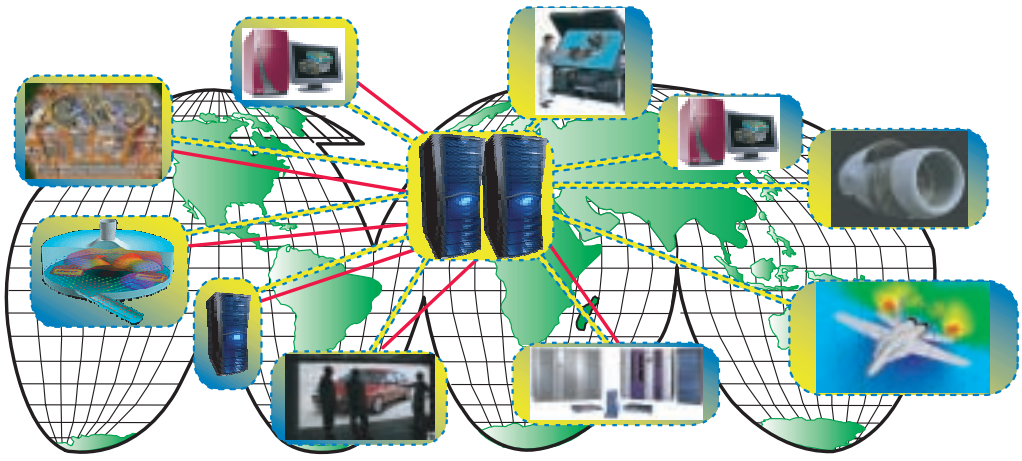

Fig. 1. Grid computing creates an infrastructure of shared resources that can be accessed transparently from any point.

capabilities. With centralized rendering, less-capable client systems make full use of the power of SGI advanced visualization systems to manipulate and visualize large data sets. An OpenGL Vizserver visualization grid consists of:

- **Visualization server(s):** The application and data reside on the server. All rendering is performed using the capabilities of the server. Rendered images are compressed before delivery to the client. The server software for OpenGL Vizserver runs on Onyx family platforms with InfiniteReality® or InfinitePerformance™ graphics. (See table 1.) Advanced visualization systems combine high-performance graphics, multiple CPUs, large amounts of memory, and exceptional I/O bandwidth to compute and visualize problems far beyond the capabilities of typical desktop systems.
- **Visualization client(s):** The client receives and displays rendered images from the application running on the server. Because the client simply displays pre-rendered images, accelerated graphics capabilities are not required. The client has full control of the application running on the server (or shared control during a collaborative session). OpenGL Vizserver supports a variety of UNIX®, Linux®, and Windows® OS clients. (See table 1.)
- **Existing grid infrastructure:** OpenGL Vizserver is designed to operate over existing grid infrastructure. Grid backbones such as NSFnet, superJANET, and CANARIE provide enough bandwidth to support full visualization capabilities. Grid connections with less than 40Mb-per-second bandwidth may impact performance, depending on the application. (Over slower connections, the bandwidth required by OpenGL Vizserver can be reduced by increasing the compression ratio or decreasing the frame size of the image being viewed.)

A single visualization system can support as many simultaneous sessions as it has graphics pipes, or a single session can utilize multiple pipes. Collaborative sessions can be set up between local and remote users in which each user sees on his or her screen the identical image and has shared control of that image. Any change to the image made by one user is immediately reflected in the review at the other location.

Table 1. OpenGL Vizserver Features

Supported Visual Servers	Silicon Graphics® Onyx2® with InfiniteReality graphics, SGI Onyx 300 with InfiniteReality3™ graphics, Onyx 3000 series with InfiniteReality3 or InfinitePerformance graphics
Supported Clients	Silicon Graphics® workstations (IRIX® 6.5.5+), Sun™ workstations (Solaris™ 2.6+), workstations with Intel® Pentium® III or better and Red Hat® Linux® 6.2, Microsoft® Windows NT® 4.0, Windows® 2000, or Windows® XP
Compression	4:1, 8:1, 16:1, 32:1, or compression API
Frame Spoiling	
Authentication API	Allows integration in a wide variety of authentication environments
Reservation System/ Reservation API	Users can reserve time on visualization server; API allows integration with existing calendar systems
Dynamic Pipe Allocation	Waiting sessions are initiated whenever a pipe becomes dormant
Accounting	A full per user usage log for planning, billing, etc.

Fig. 2. Grid users can access the advanced visualization abilities of SGI Onyx family systems from anywhere on the grid using OpenGL Vizserver. A single Onyx family system can serve the visualization needs of several grid users simultaneously. Collaborative sessions can be established in which multiple users can see and manipulate the same visualization.

2.2 Grid-Enabling Features

OpenGL Vizserver incorporates a variety of features that make it ideal for use in a grid environment. (See table 1.) Many of these features, such as compression, authentication, and reservations, include their own APIs, making it simple to integrate OpenGL Vizserver into existing grid environments.

2.2.1 Compression and Frame Spoiling

Compression reduces the amount of data that is required to transmit an image, trading off some image quality for greater interactive capability. The compression rate is chosen to facilitate the desired frame rate based on network conditions. Compression can be particularly valuable when grid traffic is high and results in congestion, or when backbone bandwidth is below 40Mb per second.

The four compressors that come with OpenGL Vizserver are based on Color Cell Compression (CCC) and Interpolated Color Cell Compression (ICC) techniques. CCC has a compression ratio of 8:1 and ICC is 4:1; there are also modified versions that deliver 32:1 and 16:1. Both algorithms reduce the size of the image by reducing the color space of small regions within the image and therefore result in a lossy compression scheme. All the compression mechanisms supplied with OpenGL Vizserver have fixed latency. Only the size of the frame—not the contents—must be known to calculate compression time. The algorithms are also asymmetric—the computational burden is greater for the compressor. This facilitates decompression on thin clients. Compression can be disabled for situations where network bandwidth is not an issue and maximum image quality is a necessity, or where clients lack the necessary compute capability to handle decompression.

ICC: Interpolated Color Cell	4:1 compression
CCC: Color Cell Compression	8:1 compression
SICC: Scaled Interpolated Color Cell	16:1 compression
SCC: Scaled Color Cell	32:1 compression
No Compression	Compression disabled for maximum image quality

Table 2. OpenGL Vizserver Compression Options

OpenGL Vizserver also provides a compression API. Using this API, a user/developer can integrate his or her own compression algorithm into OpenGL Vizserver. These algorithms can be based on frame-by-frame compression methods like CCC or ICC, or they can use interframe compression techniques like MPEG.

Frame spoiling is another feature that can help accommodate slower networks. With spoiling turned off, the server will not generate a new frame until the current frame is displayed on the client. When spoiling is enabled, the server generates frames at the maximum rate. This feature is helpful on slower grids in situations where some frame loss is acceptable.

2.2.2 Authentication

Because the use of grid resources may often cross organizational boundaries, security on grids is of particular importance. By default, OpenGL Vizserver supports standard UNIX authentication, but includes an authentication API to allow it to support Kerberos,

DCE, or other forms of authentication. This API ensures that OpenGL Vizserver can integrate seamlessly into the authentication environment of any grid where it is to be deployed. OpenGL Vizserver also supports the ability to traverse packet-filtering firewalls.

2.2.3 Reservations

Grid users can reserve time on a visualization server running OpenGL Vizserver using a simple graphical interface provided with the software. Alternatively, a reservation API is provided so that OpenGL Vizserver can integrate with existing grid-based reservation systems. When the reservation system is enabled on the server, users without reservations are blocked from starting OpenGL Vizserver sessions, to ensure fairness.

2.2.4 Dynamic Pipe Allocation

Dynamic pipe allocation works with the reservation system to ensure maximum utilization of graphics resources. When all pipes are utilized and someone is waiting to initiate a session, dynamic pipe allocation will move that user onto a pipe as soon as it becomes dormant.

2.2.5 Accounting

OpenGL Vizserver provides a per-user usage log, which can be used for accounting and billing purposes. This feature makes it possible for one organization to charge back another for system usage, and can facilitate the sale of visualization time as a service on commercial grids.

3.0 Using OpenGL Vizserver in a Grid Environment

OpenGL Vizserver can greatly increase the use of visualization for problem solving and collaboration in a grid environment.

3.1 Remote Visualization

The most straightforward use of OpenGL Vizserver is for remote access to advanced visualization, analogous to the way in which grid users access other grid resources. With OpenGL Vizserver, a user anywhere on a grid can initiate a visualization session on an available visualization server and have the output displayed on a local client system, leaving both the data and the application on the server. SGI Onyx family visualization systems provide a balance of storage, compute, memory, I/O, and visualization resources. A grid user can select a visualization server with the most appropriate characteristics in each of these areas to get optimum performance from the visualization application.

It is also important to note that, although the term "remote" may imply significant distance, this need not be the case. Users may choose to access visualization resources that are available in their local environments using OpenGL Vizserver, so that they don't have to disrupt their workflow by leaving their offices. OpenGL Vizserver can serve the visualization needs of a number of local and remote visualization users, taking the place of expensive desktop graphics workstations while providing considerably greater visualization capability.

In a novel application of remote visualization on a grid, SGI is collaborating with the Manchester Visualisation Centre (MVC), at the University of Manchester in the U.K., to bring interactive visualization to the operating room over a visualization grid. Most operating rooms don't have the visualization equipment necessary to view the results of CT scans or MRI in 3D, although the ability to do so would make the surgeon's job much simpler. Visualization equipment would be both too bulky and too expensive for the operating rooms at most hospitals, so instead surgeons continue to rely on traditional 2D x-ray films.

MVC has developed a specialized OpenGL visualization application for surgeons, called Op3D, and tested the use of OpenGL Vizserver to bring visualization to the OR. An Onyx 300 system acts as a visualization server, providing interactive 3D graphics to a laptop system in the OR. This laptop system is connected to a video projector—allowing the surgical team a larger view of the image. The surgeon controls the application with a joystick. The joystick is covered with a sterile bag, allowing the surgeon to manipulate the image during the surgery without risking contamination.

The team is working to enhance the Op3D application for voice-controlled, hands-free operation. Visualization access will be made available to other hospitals in Manchester, creating a visualization grid. Ultimately, this functionality can be replicated to bring visualization to operating rooms around the world.

3.2 Visual Collaboration

Another unique capability that VAN brings to the grid is visual collaboration. Collaborators in different locations can work together much more effectively and solve problems more quickly with shared access to the same visual information. OpenGL Vizserver allows the same visualization session to be directed to multiple locations with shared control. This means that not only do collaborators in different locations see the same image; any collaborator can manipulate the image during the discussion, and the other collaborators will automatically see the change. Many possible scenarios exist for the use of visual collaboration. In the simplest scenario, grid users consult with each other one-on-one while viewing and manipulating the same image with OpenGL Vizserver. This provides immediate access to the expertise needed to solve problems quickly.

More elaborate scenarios involve the use of SGI® Reality Center™ facilities. Reality Center facilities combine the advanced visualization capabilities of Onyx2, Onyx 300, or Onyx 3000 series visualization systems with immersive display technologies including desk displays, cylindrical or spherically curved walls, and rooms in which outputs are projected on three to six surfaces. Reality Center facilities provide an immersive environment where individuals or teams can gather to absorb themselves in problems or review designs and results of simulations. Because of the proven benefits of immersive visualization, grid-connected Reality Center facilities are increasingly being deployed. For example, Cardiff University is building a grid to provide academic and commercial access to visualization resources and HPC. Among their facilities are a Fakespace™ Portico WorkWall™ display and two Fakespace™ ImmersaDesk® R2 workbenches to provide immersive environments. Visual Area Networking makes Cardiff resources available to end users in engineering, physics, earth science, bioscience, and chemistry.

A team working at a Reality Center facility might include remote personnel or outside suppliers in a session using OpenGL Vizserver. More advanced forms of collaboration facilitate the cooperation of teams working in separate visualization centers. Such a collaborative effort might include data integration. Teams working on different aspects of a problem each make their data available for visualization, viewing their own data locally while sending the image to a collaborating site with OpenGL Vizserver. Thus each team can view visual representations of all data simultaneously, giving everyone an integrated view of the complete problem without data duplication.

4.0 Conclusion

Grid computing provides the means to efficiently share critical resources to solve difficult problems. With Visual Area Networking, SGI brings visualization to the grid, making it possible for grid users to share advanced visualization resources in the same way they now share compute resources and scientific instrumentation. Because of grid-enabling features such as compression, and APIs for authentication and reservations, OpenGL Vizserver can integrate quickly into any grid environment, making it simple to access and use the integrated compute, I/O, and visualization power of SGI Onyx family systems.

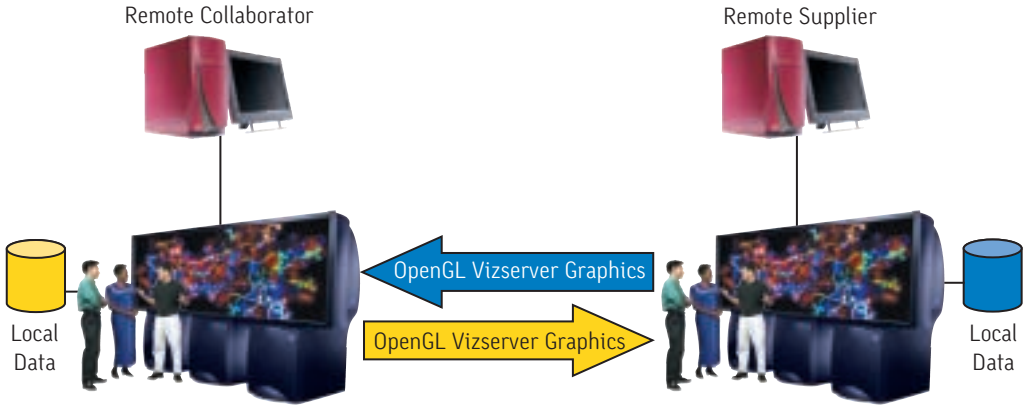

Fig. 3. Collaboration and data integration using OpenGL Vizserver: Teams working in separate visualization facilities view their own data set locally and receive data from the other facility via OpenGL Vizserver. Thus, both teams have the same integrated view. Remote collaborators or suppliers can be included as needed.

Corporate Office
1600 Amphitheatre Pkwy.
Mountain View, CA 94043
(650) 960-1980
www.sgi.com

North America 1(800) 800-7441
Latin America (52) 5267-1387
Europe (44) 118.925.75.00
Japan (81) 3.5488.1811
Asia Pacific (65) 771.0290

3308 (07/18/2002) JI4014