White Paper

The Silicon Graphics® 320 and Silicon Graphics® 540 Visual Workstation Graphics Subsystem

1.0 Introduction

The designers of the Silicon Graphics 320 and 540 visual workstations studied the latest, popular PC graphics and imaging applications to develop a new graphics architecture that optimizes the most commonly executed graphics and imaging tasks. At the heart of the system architecture, the graphics subsystem takes advantage of high-speed buses, a single, large, central pool of versatile memory, and a new SGI™ ASIC that incorporates both the geometry and the rendering pipelines—an industry first for ASIC design.

This paper discusses the technical features and architectural design philosophies of the Silicon Graphics 320 and 540 graphics subsystem as they apply to CAD/CAM, visual simulation, content creation, imaging, and other applications.

1.1 Visionary Graphics

The Silicon Graphics 320 and 540 visual workstation graphics subsystem extends the graphics capabilities of traditional PC-based systems by exploiting several features of the system architecture: multiple high-bandwidth buses, a highly accessible central memory pool, and high-speed multiprocessing. Within the graphics subsystem itself, a new graphics ASIC (the Cobalt™ chip) includes the necessary graphics functions to implement both the rasterization and geometry pipelines at the silicon level for maximum performance and throughput.

Because of its hardware-accelerated graphics and the system's efficient handling of large amounts of data, the Silicon Graphics 320 and 540 graphics subsystem particularly suits applications that require the creation, processing, and manipulation of complex, real-time visual data. Tight integration of system components provides both the bandwidth (see Table 1-1) and the processing power (see Table 1-2) to deliver workstation-class graphics while effectively competing in the PC price range.

Table 1-1. Critical Graphics Paths Performance (Peak Throughputs)

Graphics to Memory	3.2GB/sec
CPU to Memory	800MB/sec

1.2 Graphics Performance

The Silicon Graphics 320 and 540 graphics subsystem sets itself apart from PC graphics with several features:

- A dynamically allocated, high-speed memory pool (up to 1GB for the 320 and up to 2GB for the 540) makes it practical to simultaneously manage large amounts of framebuffer, texture, and z-buffer data
- A high-speed (3.2GB/second), low-latency memory bus eliminates the graphics-to-memory bottleneck and delivers three to six times the throughput available on PC systems
- A specialized, proprietary graphics ASIC that incorporates hardware accelerators for geometry acceleration, rendering, hardware-based texture mapping, dynamically allocated texture RAM, and 2D and image acceleration
- Direct motherboard connections among graphics, video, audio, network, and IEEE 1394 functions to optimize data availability, parallel processing, and content sharing among all subsystems

The result of this implementation is a visual computing system that executes graphics operations much more efficiently than any other system in its price class. While most PCs require the use of PCI and/or motherboard cards for graphics functionality, the 320 and 540 graphics ASIC puts essential graphics on the motherboard. Overall graphics performance is improved and system costs are lowered. Table 1-2 lists the essential graphics performance numbers for the Silicon Graphics 320 and 540 systems.

Table 1-2. Cobalt Sustained Graphics Performance Numbers 32-bit color, 24-bit z, 500 MHz PIII

Smooth-shaded, z-buffered pixels/sec	180M/sec
Bilinearly mip-mapped, textured, z-buffered pixels/sec	150M/sec
Trilinearly mip-mapped, textured, z-buffered pixels/sec	120M/sec
Pixel transfer (pixels/sec)	180M/sec
10-pixel non-anti-aliased vectors/sec	7M/sec
10-pixel anti-aliased vectors/sec	3.5M/sec
1-pixel lit, smooth-s shaded, 2 buf trils	7.4M/sec
25-pixel smooth-shaded, z-buffered triangles/sec	3.5M/sec
25-pixel trilinearly mip-mapped, textured, z-buffered triangles/sec	1.5M/sec
Clear rate	2.7GB/sec

2.0 Graphics Subsystem Description

2.1 Geometry and Rasterization Pipelines on a Chip

The Cobalt ASIC includes:

- 2D/3D rendering and geometry pipelines (minimizes CPU involvement in graphics)
- Memory controller
- Interfaces to the processor (P6) bus, memory, I/O subsystem, and the display subsystem

The incorporation of this much functionality and connectivity required the following package and chip technology:

- 1521 Ball flip-chip package for high pin count
- Five layers, 0.25 mm technology
- 13.6 mm x 13.6 mm die size
- 10M transistors

The rendering engine presents a high-level programming interface that offloads many graphics operations from the CPUs. Primitives are described by their window-space vertex coordinates and attributes such as vertex normal, color, texture coordinates, fog, and z. To reduce traffic between the processor bus and memory system, the interface also supports connected line and triangle mesh primitive descriptions. The Cobalt ASIC supports the following features and functions.

Rendering Engine (Geometry Operations)

- Vertex-attribute interface for point, line, line strip, triangle, triangle strip, and rectangle primitives
- Vertex transforms from object to window coordinates
- View frustum clip-check
- Normal transform and normalization
- Connected line and triangle mesh interface
- Rasterizer setup, attribute interpolation setup and anti-aliased line setup from primitive vertices and vertex attributes
- Front and back face culling
- Per-vertex lighting computation for up to four lights (and support for accumulating a host-supplied contribution from other light sources)

Other Rendering Engine Functions and Features (Rasterization Operations)

- Gouraud shading
- Depth buffering for 16-bit floating point and 24-bit fixed point z-buffer formats
- Stencil buffering for 8-bit stencil values
- Texture mapping with nearest, bilinear, and trilinear mip-mapped filtering
- Anti-aliased points and lines
- Fogging
- Scissored rendering
- Line and polygon stippling
- Alpha and chroma testing
- Alpha blending
- Logical operations
- Color plane masking
- Dithering for 5-bit RGB components
- Window clipping support through screen masks and window clip IDs
- Linear framebuffer addressing
- Occlusion testing
- Post-texture specular highlights
- 8, 16, 32-bit color formats and 16/16, 32/32 double-buffer formats
- Pixel transfers through rendering pipeline with format conversion
- Subsampled formats for YUV video
- 4x4 color matrix for color space conversion, scale/bias, and component swizzling
- Clear operations at full memory bandwidth
- Addressing up to 4k x 4k framebuffers and texture maps
- Memory management hardware for allocation of graphics buffers in system memory

Scalable Graphics Memory

- Up to 1GB for Silicon Graphics 320; 2GB for Silicon Graphics 540
- 32-bit, double-buffered
- 24-bit z buffer
- 8-bit overlay
- 8-bit stencil
- Texture memory, up to approximately 900MB/1.9GB (Silicon Graphics 320/540, respectively)

Resolution

- Up to 1920x1200 at 66 Hz in 32-bit RGBA (24-bit color plus 8-bit alpha)

2.2 System Memory

The main memory system, with a 288-bit bus (32-bit ECC), achieves data rates of up to 3.2GB per second from the synchronous DRAM (SDRAM) implementation. A single, central, dynamically allocated pool of memory offers several advantages for graphics performance and capabilities including simplified programming, high graphics processing rates, and lowered costs. Unlike typical PC architectures, where data must be transferred between main memory and graphics, video, and imaging memory located on separate interface cards, all of the 320/540 data resides in main memory where every subsystem has direct access to it.

Memory for the systemincluding, frame buffer, z-buffer, texturing, rendering, imaging, and video are all allocated by the Cobalt chipset. Instead of maintaining multiple memory systems throughout the computer (a design that forces data movement through limited bus architecture and therefore degrades performance), the Cobalt chipset allows the graphics memory to be dynamically allocated from system memory. Memory is allocated in tiled pages. To minimize page fragmentation and reduce DRAM page crossings incurred during rendering, the pages are organized linearly (without tiles).

3.0 Graphics Software

In addition to the hardware-accelerated functions, the Silicon Graphics 320 and 540 systems offer graphics performance enhancements in the form of library optimizations. Data and control flows are tuned for key Windows NT® graphics and imaging applications without sacrificing compatibilty.

The Silicon Graphics 320 and 540 systems support these standard API's:
- OpenGL® 1.2
- Microsoft® Direct3D®
- Microsoft® DirectDraw®

4.0 Application Areas

The designers of the 320/540 graphics subsystem optimized execution for a broad range of applications that require both geometry throughput and fast fill rates. The resulting graphics architecture delivers exceptional performance to these applications and presents new opportunities to expand the scope of many visual computing tasks.

4.1 CAD/CAM

CAD/CAM applications involve the assembly, texturization, and manipulation of large, complex models that are often comprised of more than 100,000 triangles. Sophisticated shading, lighting, and real-time capabilities are highly desirable. Graphics features that address these requirements include:
- High polygon and texture throughput: the system's high-speed, low-latency memory bus facilitates rapid transfer of geometries and textures in and out of memory
- High capacity: dynamically allocated memory supports the use of up to 1GB of system memory (up to 2GB on the Silicon Graphics 540 system) for model textures
- Geometry pipeline in hardware: unmatched performance for 3D geometric operations
- OpenGL extensions: optimized performance and quality including occlusion test and occlusion correct; post-production texture specular highlights

4.2 Imaging

Similar in requirements to CAD/CAM modeling tasks, imaging jobs involve the creation and manipulation of large, high-resolution files such as medical scan data, maps, photos, and GIS data files. For these applications, the Silicon Graphics 320 and 540 graphics subsystem provides:
- Rapid data movement: the high-speed memory bus optimizes storing and retrieving images to and from disk.
- Extremely high pixel fill and transfer rates
- High-speed interface to high-quality displays
- Support for large amounts of texture memory

4.3 Content Creation

Content creation applications require a broad range of capabilities. Complex models demand the geometry processing functions (similar to CAD/CAM) and the creation of texture-rich environments calls for advanced visual simulation functions. Content creation benefits from 320/540 features such as:
- A unique combination of vertex throughput and fast fill rates
- Large amounts of texture memory for rich environments

4.4 Visual Simulation

Visual simulation tasks require the ability to walk through complex 3D scenes, whether they be architectural structures, battlefield diagrams, or game environments. In order for a system to be used in this setting, it must be capable of real-time display and manipulation of realistic, highly textured spaces. The system offers:
- High-bandwidth, high-capacity memory for image storage and manipulation
- High-quality display capabilities
- Hardware-accelerated 3D graphics
- Support for large texture maps
- Support for fully textured, real-time simulations

4.5 New Application Possibilities

The tightly integrated system architecture and large, central memory pool introduce many possibilities for creatively combining graphics, imaging, and multimedia data. Unlike traditional PC architecture, the Silicon Graphics 320 and 540 workstations do not require that data be moved between graphics, imaging, and video subsystems. Once in memory, data can be shared among and accessed by all of the subsystems. This arrangement makes it practical to combine image, geometric, and video data.

The 320/540 I/O engine is capable of handling all of the bandwidth of two uncompressed streams of NTSC or PAL video. It also allows the use of video as a graphics component, not only supporting high-quality uncompressed video stream input, editing, and storage, but also allowing mixing of video with 2D and 3D graphics—all in real time.

Broadcasting, learning on demand, and content creation applications will benefit from the creativity unleashed by the new architecture.

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

U.S. 1(800) 800-7441
Europe (44) 118-925.75.00
Asia Pacific (81) 3-54.88.18.11
Latin America 1(650) 933.46.37

Canada 1(905) 625-4747
Australia/New Zealand (61) 2.9879.95.00
SAARC/India (91) 11.621.13.55
Sub-Saharan Africa (27) 11.884.41.47

2351 6/99 J10043