Tackling Large Volume Visualization Challenges with Real-Time Ray Tracing

Table of Contents

January 2006

Executive Summary

In high-performance computing (HPC) markets, such as manufacturing, oil and gas exploration, and health care, users are struggling to visualize a rapidly growing body of volumetric information. The latest scanning and measurement devices capture far more detail than earlier versions – with some achieving up to 1,000 times the resolution of the previous generation. This puts enormous pressures on the graphics hardware and software used to visualize that data in 3D. It forces users to view information in piecemeal, with system components swapping out one portion of an overall model with another. This is a time-consuming, cumbersome approach that hurts productivity and can hobble an engineer’s or researcher’s ability to derive insights quickly and efficiently.

Optimized graphics processing technology has helped, though many acceleration techniques sacrifice rendering quality in the name of performance. But in certain HPC environments – where automobile or aircraft quality and safety are on the line, or where medical researchers attempt to detect how cancer spreads – compromises in data accuracy and detail can be especially costly.

Fortunately for these users, a proven and popular graphics rendering technique called “ray tracing” has moved beyond its roots as a computationally expensive approach to displaying data. With the advent of high-performance, lower-cost processing technology, and sophisticated software ray tracing engines, ray tracing is now an attractive option for real-time rendering of ever larger volumes of data. Key to this transformation is the availability of cost-efficient, 64-bit HPC processors that leverage shared-memory architectures, making the entire system’s memory available to render the entire data set.

Ray tracing solutions have been proven to scale on a sub-linear curve as a rendered scene grows more complicated, and can scale from two to hundreds of 64-bit processors. Real-time ray tracing

Sixty thousand (200 micron) 2D slices were captured by a very high resolution CT scanner at Stanford University. Real-time Ray Tracing technology is used to re-create a photo-realistic image of "Sherit" (a 2000 year old Egyptian child Mummy of San Jose's Rosicrucian Museum). A detailed view of the skeleton and outer casting is revealed, while leaving the mummy intact.

offers a massively scalable solution for on-demand visualization of volume data that vastly exceeds what can be viewed on today's typical desktop system. Ray-traced renderings present real-world 3D volume data with photo-realistic accuracy – a breakthrough for medical and manufacturing users whose work leaves little room for error.

For HPC users facing an escalation in volume data, real-time ray tracing offers a cost-effective visualization solution based on industry-standard technologies and can scale powerfully as data set sizes continue to increase.

The Road to Real-time Ray Tracing

In the past two decades, industry, science and medicine have grown increasingly reliant on visualization to save time and improve accuracy. Whether analyzing an engine block for microscopic fissures, determining the most efficient path for an oil well, or assessing brain tissue for abnormalities, the act of seeing and interacting with data has proven to be an indispensable part of high-performance computing (HPC) environments since Ivan Sutherland introduced Sketchpad in 1963.

Early visualization systems were built from loosely integrated graphics and computational components. They were often based on mainframe computers facing strict limitations in memory, processor power, and disk speeds. To improve performance, HPC graphics pioneers focused on more tightly integrating system resources. Key to this effort was the specialized graphics processing unit (GPU), a hardware processor developed to render polygons. GPU-based workstations became the platform of choice for driving most demanding visualization applications.

An essential shift occurred in the late 1990s when GPUs became increasingly commoditized. But over the past decade, while the performance of GPUs has soared, commodity system interconnect speeds have failed to keep pace.

The resulting bottleneck between CPUs and GPUs hasn't presented much of a problem for most applications. But as advances in scanning and measurement technologies create a data glut in HPC environments like manufacturing, oil and gas exploration, and health care, new problems are arising. Data sets are growing fast. And today's graphics acceleration hardware needs help carrying the load.

This white paper examines some of those problems. And it looks at how a long-popular rendering technique allows HPC users to leverage major technology trends to augment the contribution of GPUs to visualize even their largest data sets in their entirety, and then scale those capabilities as their needs evolve.

What Is Volume Visualization?

Today's scanning and measurement devices can record information on multiple 2D planes, capturing information about specific volumes to gain an understanding about the structure within the data. Volume visualization allows users to see that data in 3D, so it can be rotated, zoomed, and interactively clipped for detailed inspection.

Two dimensional CT scans are combined to create a 3D visualization of a human spine.

Because volumetric data can be combined from multiple sources, such as CT and MRI scans, volume visualization software spatially aligns or registers the data so that it can be represented in a single 3D visualization. This task requires powerful graphics and computing resources, particularly as data sets grow larger.

RAY TRACING TAKES HOLD

As GPUs came into their own, scientists were busy refining a rendering technique developed in the 1960s called "ray casting," which was famously used to simulate nuclear penetration effects and to create television ads and segments of the 1981 computer animated film Tron. The body of refinements to ray casting eventually came to be known as "ray tracing."

The ray tracing algorithm works in direct concert with its name: it projects a ray through every pixel on the screen and traces it to the 3D scene image. The algorithm then calculates how the ray and the light source intersect with the objects in the scene. As the rays land on various surfaces, three fundamental behaviors occur: the ray will be reflected, refracted or absorbed. If an object in the scene has a reflective surface, for instance, then the traced ray will generate another ray, which in turn may intersect with other objects in the scene. The algorithm stops when the ray stops bouncing and the actual color of the pixel is stored in the computer's frame buffer.

Ray tracing can achieve shadows and illumination at a level of precision that traditional GPU techniques alone cannot match. Indeed, ray traced scenes can exhibit photorealistic detail, whether it is derived from captured volumetric data, CAD models, or animated scenes. Only a few years ago, ray tracing was viewed as a computationally expensive way to attain high-quality imagery – the sole domain, it seemed, of animators and others with boundless budgets.

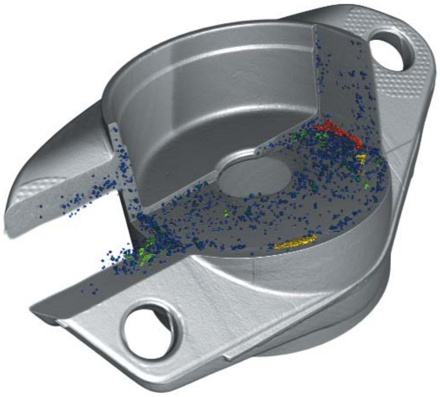
GOING MAINSTREAM

A convergence of key industry trends has changed all that. Today, ray tracing is used in the manufacturing, oil and gas, and animation industries. Its adoption is the result of plummeting hardware costs and soaring system performance, along with a broad range of increasingly sophisticated software renderers that efficiently leverage computational hardware and, to a lesser extent, GPUs.

The commoditization of hardware components has triggered massive economies of scale. In the first few years of this decade, for example, computer server performance alone leaped by 150 percent annually, and price/performance increased at an even faster rate[1]. Aggressive competition at both the component and system levels has ensured that CPUs, GPUs, memory, storage and complete solutions push price/ performance to new levels. As a result, increasingly powerful computing and visualization systems are proliferating.

The availability of commercial ray tracing software has also been essential. Since the 1990s, several companies have supplied users in targeted vertical markets with packaged solutions that allow them to interactively view data, either as traditional pixels or as volume pixels, also known as voxels. So today, many users already rely on ray tracing in its classic form.

In manufacturing environments, for instance, quality engineers routinely examine ray traced volumetric data. For this, they use computer tomography (CT) scanning devices combined with tactile coordinate measuring machines to gather vital information on new product prototypes or production samples off the assembly line. This practice boosts productivity by allowing them to drastically reduce the need for destructive testing – or the process of literally sawing and otherwise disassembling newly constructed products to examine and compare their properties against engineering specifications. By leveraging volume visualization enabled by affordable ray tracing technology, they can literally study the interior of an engine block wall or examine imperfections in a machined surface more quickly and efficiently than in years past.



Imperfections are detected in an aluminum casting without the need for destructive testing. Ray tracing technology includes image processing and analysis features that enable a virtual slice through an aluminum casting and identify defects.

Manufacturers such as Ford Motor Company, DaimlerChrysler, Toyota Motor Corporation, BMW, and The Goodyear Tire & Rubber Company rely on ray tracing in this way to streamline their manufacturing, testing and quality assurance practices.

Until now, these solutions have ably handled the rendering and visualization requirements of users. But these same users, accustomed to visualizing data sets on PCs and laptops, are grappling with the challenge of maintaining satisfactory rendering performance in the face of ever-increasing data sets. The problem of rendering these visualizations in real time is bigger than any single component working on its own can solve.

An Onslaught of 3D Data

The challenges of ever-growing data sets present issues for GPU-based visualization. Today's GPUs are severely memory-constrained, with each graphics processor offering no more than 512MB of memory. Even deploying four GPUs in a single system would apply little more than 2GB of available memory toward visualizing a data set.

While this may seem sufficient to many users, HPC industry trends point to an immediate future in which data sets of 2GB – or even 20GB – will appear almost quaint. Nowhere is this more evident than in two fields: manufacturing quality assurance and medical imaging. These industries are facing an explosion in the volume of data captured by Computer Tomography (CT) scans, Positron Emission Tomography (PET), fluoroscopy devices, Magnetic Resonance Imaging (MRI) scans, and two-photon microscopes. Consider these recent developments:

QUALITY AND PROCESS CONTROL

With CT scanning and other detection devices commonplace in quality assurance operations among automobile manufacturers, the ability to quickly and accurately visualize scan results helps automakers produce safer, higher quality, and more reliable products. Industry observers report that automakers, for instance, are working to scan to increasingly greater levels of detail, some resolving down to 1 micron. That level of detail allows manufacturers to detect tiny imperfections, cracks or fissures in an engine head or aluminum casting – problems that may measure a 100th of a millimeter. These studies can generate multiple gigabytes of data, and the ability to quickly and accurately visualize that information streamlines the quality control process and ultimately shortens the company's time to market.

One challenge is that the detection devices themselves are getting larger and more advanced, capable of capturing more and more information with each generation. Manufacturers worked for years with flat panel detectors that captured 512 pixels in a square. That soon doubled, and the number continues to increase.

Working with 3D data also tends to increase data sets geometrically. Doubling the number of pixels in the X and Y axes produces eight times more data. A scan from a 1K pixel scanner, for instance, can measure roughly 2GB. Doubling that capture rate to 2K increases the 3D data set to about 16GB. With a host of manufacturers, including Bosch, Volkswagen AG and DaimlerChrysler moving to 2K scanners, single scans measuring up to 16GB are fast becoming the standard.

At Stanford University School of Medicine, researchers use an AXIOM Siemens scanner to produce very high resolution CT scans. High resolution scanning capabilities combined with the latest visualization technology are transforming medical procedures and improving outcomes.

CT scans are also used as, or in conjunction with, geometry analysis devices, which measure the surface of finished products so manufacturers can ensure those parts meet original specifications. Some parts have greater tolerances than others, although the ability to understand where geometry errors are occurring can save companies significant time and money if they can catch manufacturing problems sooner rather than later.

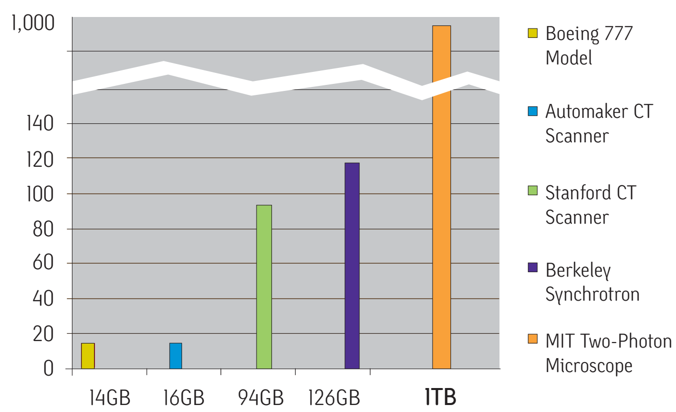
"It's easy to see a convergence of different types of measurement devices – a CT scan for X-ray images and a tactile coordinate measuring machine for geometry analysis – into a single device," says Christof Reinhart of Volume Graphics GmbH. "You'll have a higher accuracy than you have today, and you'll get information on a wider variety of attributes, such as porosity, wall thickness, fissures and geometry errors." At the same time, Reinhart predicts detection and visualization will move closer to the production line. Today, for instance, manufacturers test a representative sampling of parts for defects. But soon, that will change. "You'll see CT scanners right on the production line, capable of capturing information on every product. The goal will be to achieve a scan not just every half hour, but every half minute."

RESEARCH PUSHES THE ENVELOPE

The quest for higher-resolution CT scans in medical and biological research has produced new devices that can generate up to 100 times the detail of traditional CT scanners. At the **Lawrence Berkeley Laboratory** (LBL), a "micron CT" scanner known as a Synchrotron records data at a resolution of 10 microns (compared to typical hospital CT scanners with 500 microns resolution). For instance, Synchrotron scans can fill a 1mm square region with 500 points of data, compared to just four points for traditional scans. LBL's Synchrotron, one of some 40 such scanners located worldwide, recently produced a single data set totaling 126GB in size, or the memory of more than 250 of today's highest-end GPUs.

At the **Massachusetts Institute of Technology** (MIT), research undertaken using an advanced two-photon microscope is helping medical imaging pioneers create the most detailed 3D organ visualizations in history. The MIT device is approximately 1,000 times faster than commercially available two- photon microscopes – and generates far more data. The result is an enormous gain in productivity. "A typical two-photon microscope would take months to image a mouse heart at a resolution of one micron," notes Dr. Tim Ragan, biomedical research associate at MIT. "We're working on accomplishing that in half an hour. It works much faster, and generates much more data, so you can look at a larger sample or section and draw conclusions much more quickly." The mouse heart imaging study resulted in a 3D visualization of 1TB in size. To view the data on his current visualization system – a laptop PC – Ragan had to down-sample the data by a factor of 16 in each of the three dimensions. In other words, he was only able to visualize 1/16 of the data collected.

HPC Visualization Data is Exploding

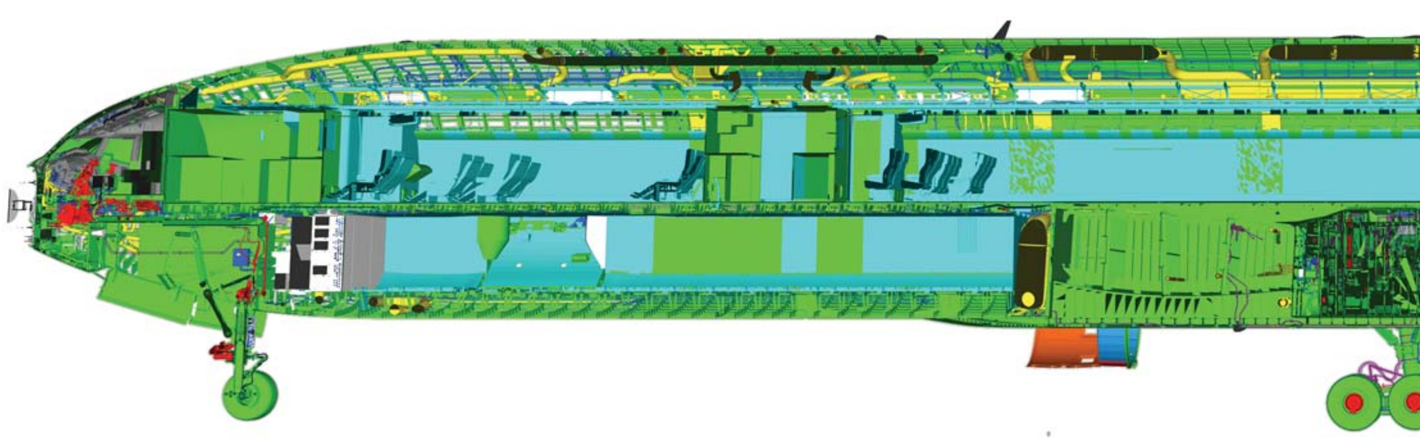


Boeing: Pioneering CAD Visualization

Aerospace giant The Boeing Company, as part of its ongoing effort to find cost-effective ways to maintain quality control, recently collaborated with SGI and Intel Corporation to demonstrate how real-time ray tracing technology could be useful in streamlining in-factory quality assurance processes. The demonstration used real-time ray tracing technology to interactively display the entire original CAD model of the first Boeing 777® aircraft, a 14GB data set made up of 350 million polygons and representing 16,000 different models. The demonstration showed that a Boeing Quality Assurance analyst on the factory floor, equipped with a portable computing device, could compare a finished part or assembly with the original design to check for variances. Using a Visual Area Network and video-conferencing technology, the demonstration also allowed the Quality Assurance analyst to remotely confer with a Boeing Liaison Engineer, who could interactively view the same model to validate any potential workarounds.

"The interesting aspect of looking at the entire model is two-fold," observes David Kasik, technical fellow and enterprise visualization architect at The Boeing Company. "First, you can look at any other aspect of the airplane that might be impacted by the problem or the fix. And second, you are guaranteeing that you will have the entire engineering definition at your fingertips, and virtually eliminate the risk that you forgot to load something."

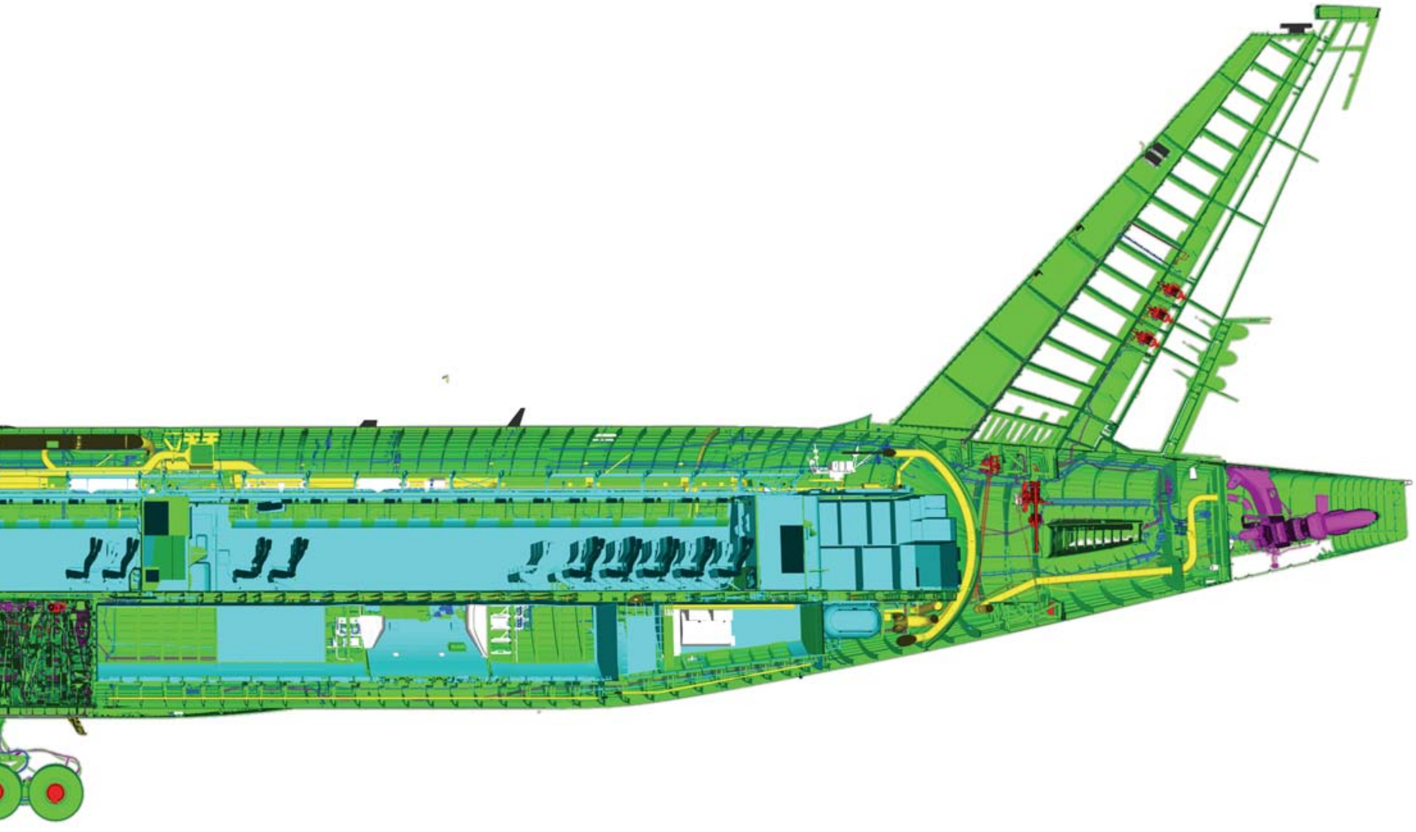
The collaborative demonstration used real-time ray tracing to render the model on demand, while exploiting such ray tracing capabilities as shadows, synthetic transparency and coloring. While multiple Intel® Itanium® 2 CPUs drive rendering, the demonstration also leverages GPU hardware to handle such tasks as frame buffer management and pixel memory.



The 350 million triangle Boeing 777 dataset is rendered in real-time using Manta open source ray-tracing software, Meanwhile, the sophisticated two-photon microscope at MIT, the only one of its kind in the world, will soon have company: MIT's Ragan expects that 20 to 30 research sites worldwide will have similar devices. For groundbreaking medical research, Ragan says, these advances are essential: "When you're studying cancer, for instance, blood vessel growth near a tumor is a very important parameter. You want to look at the largest portion of tissue possible to get a global sense of the how cells interact with the greater tissue mass," he says. "Being able to interactively view the heart as a whole, while at the same time being able to see zoom in on portions in 3D with sufficient detail to identify individual cancer cells, is very important. Five years ago, this wasn't even a possibility because it would be simply too expensive to store and process that data. Now we're doing it."

3D imaging has become an essential technology in clinical use providing detailed views of the patient's body (here, the kidneys) with higher and higher quality. The amount of data which has to be processed and viewed is continuously increasing. Ray-tracing technology exploits the power of both the CPUs and GPUs to enable interactive visualization of these large data sets.

The pace of change is also accelerating beyond research laboratories. In 2003, Siemens AG Medical Solutions unveiled the Siemens SOMATOM® Sensation 64, which makes it easier for practitioners to gather more CT data from faster scans. For many hospitals and clinics accustomed to single-slice scanners, the new product represented a revolutionary leap forward. In fact, not long ago, a CT scanner with these capabilities would have been considered utterly leading edge with limited real-world use. But Siemens now counts more than 350 installations for its 64-slice scanner, and Stephen Mohan of consulting firm Frost &



which includes large-model visualization capabilities contributed by SGI.

Sullivan's North American Healthcare Practice says that "health-care professionals now consider it an industry standard." Other manufacturers, such as GE Healthcare, Toshiba and Phillips Medical Imaging, have introduced 64-slice CT scanners as well, signaling that the march toward faster, higher-resolution scanning devices will continue.

LARGER BURDENS ON VISUALIZATION

Advances like these put exponentially larger burdens on the systems used to combine thousands of 2D slices into 3D volume visualizations.

"Each time you get a new technology, it handles more data," says Robert Cheng, a medical imaging research assistant working at LBL on behalf of Brown & Herbranson Imaging and Stanford University, agrees that the data explosion will only continue. "Because we're trying to explore how far we can stretch the technology, we have been trying to get the largest data sets possible."

Clearly, medical imaging data sets are exploding faster than ever. But the memory limitations inherent in GPU technology force users to discard significant portions of collected data or view large visualizations in small sections. This slows the process of analysis and potentially hinders insights that would come from interacting with the entire visualization.

Enabling Trends

Fortunately for HPC users, the technologies required to efficiently render this data in real time already is in place. Better still, they will evolve on a curve that will prove helpful for years to come.

PROCESSORS: 64-BIT AND THE MOVE TO MULTI-CORE

David Kasik, technical fellow and enterprise visualization architect at The Boeing Company, notes that today's high-performance 64-bit processors are key to enabling real-time ray tracing. "Most desktop architectures today are limited by 32-bit processors," he says. "Say you have 2GB of memory on your 32-bit desktop, and you have 10 GB of data that you want to display in real time. In that situation, you have to feed two beasts: the CPU RAM from disk, and the GPU VRAM from the CPU RAM. You have to do both tasks fast enough to compute a new frame with enough detail to be compelling in a tenth of a second or less. The advantage of 64-bit CPUs is that you can add enough RAM to contain the entire 3D data set. So the disk-to-CPU RAM problem ceases."

Can GPUs Alone Drive Ray Tracing?

Because the benefits of ray tracing are well-known, efforts are underway to create GPU-accelerated ray tracing solutions. However, many of these solutions require complicated programming techniques. Chief among these is compressing source data and textures so they "fit" within the memory constraints of GPUs. But compression leads to data loss, which presents grave problems to environments that require accurate analysis of data.

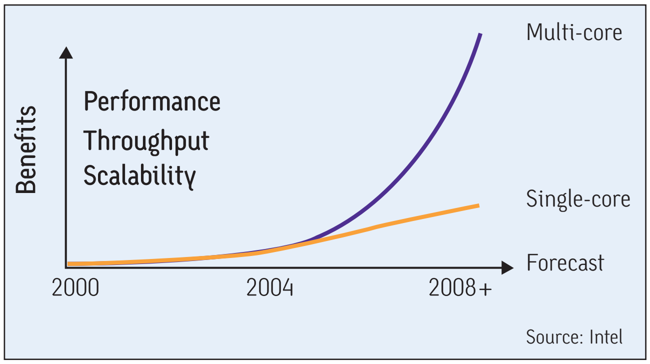
"When you're looking for faint details as you try to make a medical diagnosis or determine the production quality of a new engine block, you can't have the losses that come with compression," notes Christof Reinhart, of Volume Graphics GmbH, which provides ray tracing solutions to the medical and manufacturing markets. "If the user has to reduce his volume size to look only at a region of interest, go back and forth and load a portion of data and throw out other data, it doesn't work on a day-to-day basis. And the problem only gets worse as the data sets increase."

Kasik and others point out that GPUs and CPUs can co-exist well in a real-time ray tracing environment. Depending on the software ray tracing renderer used, a GPU can be used to handle shading, frame buffer management or pixel memory, while the CPU handles the bulk of the rendering workload.

And despite popular myth, HPC users will continue to benefit from Moore's Law, the dictum that the number of transistors on a die will double every 18 months. The reason is that processors from the Intel Corporation will soon feature two execution engines – or cores – on a single processor across their entire product line.

So-called "dual-core processors" represent the next major advance in CPU architectures, delivering a cost-effective way to escalate CPU performance via fully parallel execution of multiple software threads. "It's reasonable to apply Moore's

Multi-core Processors on the Rise



Law to the phenomenon of multi-core computing," advises Jim Hurley, principal engineer in the Application Research Lab of the Corporate Technology Labs at Intel. "In other words, one might imagine that the number of cores on a die could reasonably double every 18 months."

Hurley says this is particularly promising for HPC users who rely on ray tracing for visualization. "Ray tracing is ideal for multi-core computing, because each ray is independent from one another," notes Hurley, who says initial tests show strong scalability of ray tracing applications. "We've experimented on eight cores with hyperthreading turned on and off, and we get perfectly linear scalability on ray tracing."

SHARED-MEMORY OVER COMMODITY CLUSTERS

More and more, HPC environments are moving to shared-memory systems designed to hold entire data sets in memory for maximum access and interactivity. This is an enormous advantage over commodity Linux® clusters, whose distributed architectures operate their small nodes under individual instances of Linux. To ray tracing software, these cluster nodes may as well be separate systems, leading to latencies that make attempts to interact with large data sets completely unsatisfactory. Such clusters are not suited to visualization via software-enabled ray tracing when the visual dataset exceeds a 32-bit address space.

In contrast, open architecture, standards-based Linux shared-memory systems like SGI® Altix® servers are ideally suited to serve as platforms for real-time ray tracing visualization. Such systems offer massive scalability within each node, from two to hundreds of processors. SGI Altix systems are already popular as computation engines in many HPC markets throughout the world. The addition of real-time ray tracing software adds a visualization component to the systems, and extends the functionality of existing computer investments.

OPEN SOURCE AND STANDARDS-BASED COMPUTING

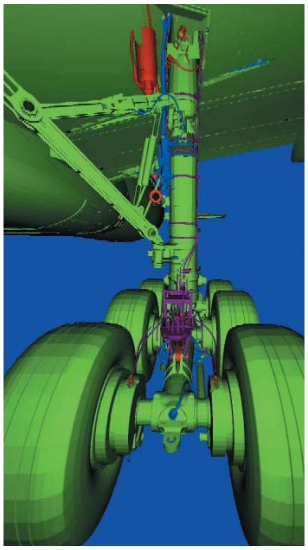
The move toward adopting Linux and standards-based computer systems has in large part been led by HPC markets. The pace of innovation afforded by open source technologies and the cost advantages enabled by standard components attract many performance-minded users, particularly researchers.

While commercial, proven, certified solutions are available from ISVs, the open source community has its own ray tracing platform that is rapidly gaining acceptance. MANTA, a framework for building ray tracers, was developed by engineers at the University of Utah. With contributions from SGI, MANTA now has the ability to render large data visualizations. In tests conducted by the University of Utah, MANTA scaled strongly on a 1,024-processor SGI® shared-memory server[2].

High-fidelity Interactivity

But as a growing number of HPC users learn about the advantages of real-time ray tracing, the approach is finding ever-broader acceptance as a resource-efficient and cost-effective technique for interactive visualization of large data. Among its advantages:

- **Efficient Rendering.** Ray tracing's inherent efficiency lies in a simple idea: It renders only what the light allows the user to see. And as the complexity and amount of information required for a particular scene increases, the efficiencies improve as well. For instance, traditional raster-based graphics will take twice as long to render twice the number of polygons in a scene. In contrast, ray tracing scales logarithmically. Tests conducted by Intel Corporation showed that engineers were able to increase the complexity of a viewed scene by 10 times before doubling the rendering time of a ray traced scene. So the more complicated the visual scene, the greater the benefits of ray tracing[3].
- **Greater Accuracy**. A common trick for traditional GPU-based visualization is to reduce polygon count and use textures to improve rendering performance. But polygon reduction often leads to loss of detail, which is unacceptable in many medical, research and manufacturing environments. Ray-traced visualization achieves photorealistic interactivity without relying on compression, making it ideal for environments in which visual fidelity is at a premium. "GPUs can play games with things like textures," notes Boeing's Kasik. "But they tend to break down when closely scrutinized."



This detail of the Boeing 777 nose landing gear is part of a much larger 350 million triangle Boeing 777 dataset. Manta ray-tracing software enables real-time, detailed and interactive visualization of each part of the dataset.

- **Vast Scalability**. Crucial to users' ability to visualize large data sets is ray tracing's reliance on CPU-based algorithms or routines. As a result, ray tracing leverages processors that can scale into the hundreds, as well as new multi-core processors from industry innovators such as Intel. Ray tracing also is especially well-suited to accessing a single memory-copy of the data.
- **Tighter Integration**. Because ray tracing algorithms are executed on CPUs, graphics and computational tasks are far more easily integrated than with GPU-centric visualization. For applications like medical imaging, the system faces a substantial computational problem in that it must pull information from volumetric source data to identify and treat differently the structures in the data as the user interacts with the 3D visualization. The general-purpose nature of the ray tracing algorithm eases this level of integration. At the same time, ray tracing algorithms can efficiently make use of both GPU hardware and CPU hardware, leveraging each to its best ability.
- **Progressive Rendering**. The computational "cost" of rendering is always directly proportional to the image size, so the higher the resolution, the more resources required for rendering. But using a technique called "progressive rendering," ray tracing algorithms can easily render an image at any desirable resolution – more easily than traditional GPU-based raster graphics algorithms.

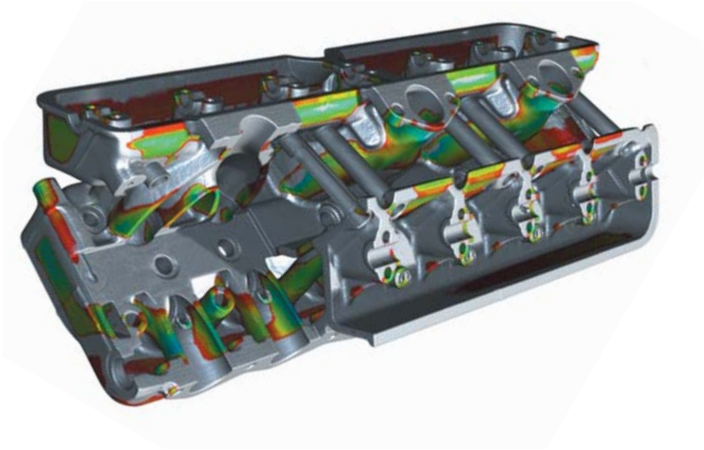
VISUALIZING VOLUMES WITH REAL-TIME RAY TRACING

Ray tracing software pioneers such as inTrace GmbH, Volume Graphics GmbH, mental images GmbH, and VRcontext s.a./n.v. have helped establish a thriving commercial marketplace for real-time ray tracing in life sciences and manufacturing. As CPUs grow more powerful and cost-efficient, and with shared-memory systems growing more ubiquitous and affordable, these companies and their customers are fomenting the development of a new standard for visualizing large volume data.

RAY TRACING IN MANUFACTURING

Manufacturing applications of real-time ray tracing include quality control processes, rapid prototyping and reverse engineering. With commercial and open source ray tracing tools, manufacturers can perform a variety of visualization functions, such as:

- **Wall thickness analysis**, including analyzing CT data sets for areas within a preset wall thickness interval
- **Porosity analysis**, which detects and analyzes internal defects
- **Isosurface extraction**, which generates polygonal models of CT extractions



Wall thickness analysis is visualized based on a CT scan of a BMW 328 roadster cylinder head, built in 1937. Volume Graphics's software enables interactive visualization of the 1GB voxel data set, allowing engineers to rapidly detect inclusions or material defects, measure internal wall thicknesses, and extract the geometry of a complex internal cavity.

RAY TRACING IN HEALTH CARE

In medical research, real-time ray tracing will play an increasingly important role as a resource-efficient way to visualize the growing body of patient data gathered from multiple scanning and sensing sources. Today, bodies are often scanned and visualized in sections, but there are efforts underway, including multiple virtual autopsy initiatives, to capture CT scans of entire bodies.

Who Uses Ray Tracing to Visualize Volume Data?

Medical Research and Clinical Practice
- Lawrence Livermore Laboratory
- Massachusetts Institute of Technology
- Stanford University
- Stryker Leibinger
- University of Utah

Automotive Industry
- BMW
- Bosch
- Daimler Chrysler
- Ford
- Goodyear
- Toyota
- Volkswagen

Aerospace Industry
- The Boeing Company
- EADS N.V.
- Eurocopter

Source: Volume Graphics GmbH and SGI

In this market, ray tracing enables a wide range of capabilities, such as:

- **3D data classification, segmentation, and manipulation**, which allows surgeons to, among other things, visualize key portions of patient data to enable less invasive procedures and treatments
- **Patient data and CAD model integration**, which combines patient-specific information with implant models to assist in preoperative planning of implant procedures
- **Sub-cellular analysis**, giving researchers a microscopic and macroscopic understanding of the behavior of cancer or other cells both at the cellular level and within the context of tissue sections

Capabilities like these will someday help speed day-to-day diagnoses and procedures in hospitals and clinics while enabling laboratories and institutions to break important new ground in medical research. At **Massachusetts General Hospital** (MGH), researchers are gearing up to leverage the power of ray tracing to study the process of metastasis, in which solitary cells from a cancerous tumor travel through blood vessels to take hold in another part of the body. Approximately 90 percent of cancer deaths are the result of metastasis[4]. Yet metastasis is difficult to study, because its root actions take place at the cellular level, and a single cell could travel anywhere in the body. To better understand the process, MGH researchers inject cancer cells into laboratory mice. The cells are labeled with a dye. Later, when the mice show signs of tumor growth, the animals are scanned to see where the cancer cells have lodged.

"The problem is that it's extraordinarily difficult to know where the cancer cells have ended up," notes MIT's Tim Ragan, whose team is working with MGH to investigate how highly detailed scans of mouse organs can be visualized via ray tracing to reveal clues to the onset and progression of metastasis. "You can search through entire organs or regions of tissue, and those few cells could be hiding anywhere. Using current tools and typical visualization techniques, you just can't do it; it's too labor-intensive, too time-intensive and too costly. But with real-time ray tracing on a multiprocessor server with significant shared memory, it's entirely possible to find those few cells."

OTHER MARKET SEGMENTS

Real-time ray tracing also has been making an impact in other market segment shares, such as oil and gas reservoir analysis. Here, energy exploration and production companies have analyzed seismic data sets totaling more than 10GB on a 64-bit workstation. But data sets here are growing as energy companies strive to learn as much as possible about the characteristics of oil and gas deposits located miles below earth's surface and beneath the depths of the world's oceans. With that knowledge, they can determine the most efficient well paths, pinpoint the safest and most productive locations for oil rigs, and maximize their return on billions of dollars of R&D investments.

An Emerging 'Killer App'

As data sets continue to grow, so will the challenge of presenting data so people can quickly and accurately understand it. A chief advantage of software-based ray tracing, however, is that it is ideally suited to scale along with the needs of HPC users.

As we have seen, every major industry trend is poised to elevate ray tracing volume visualization as the next "killer app" for big data HPC environments. A Salomon Smith Barney study predicts enterprise application data will increase at a compound annual growth rate of 30 to 40 percent[5], and most industry observers recognize that technical data is growing much faster than that.

HERE TODAY, AND TOMORROW

A confluence of tectonic industry trends – an explosion in the size of HPC, the fixed-memory limitations of GPUs, the rise of massively scalable and multi-core processor systems, and the availability of proven software-based ray tracing solutions – has thrust real-time ray tracing from the frontiers of 3D graphics to its very forefront.

In the next five years, users in key HPC markets will continue to realize the vast cost and performance benefits of real-time ray tracing as they work with some of the largest data sets ever visualized. Without investing in additional GPU hardware or corrupting their data through compression, they will interactively analyze and interpret information faster and more accurately than ever.

For those users, real-time ray tracing will deliver faster time to insights. For society, it will mean faster and more accurate diagnoses, a better understanding of deadly diseases, less invasive surgical procedures, safer and better automobiles and aircraft, and more efficient energy production. In so many ways, ray tracing is a "killer app" whose time has come.

SOURCES

[1] Giga Group, "IT Trends 2004: Server Systems," Richard Fichera, Jan. 29, 2004.
[2] David Weinstein, "The Star-Ray Interactive Ray-Tracer Debuts at SIGGRAPH 2002," University of Utah Scientific Computing and Imaging Institute, accessed at http://www.sci.utah.edu/stories/2002/sum_star-ray.html on Nov. 15, 2005.
[3] Jim Hurley, "Ray Tracing Goes Mainstream," Intel Technology Journal, Vol. 9, Issue 2, May 19, 2005, pp. 99-108.
[4] "Scientists identify gene responsible for tumor metastasis in cancer," MIT News Office, Aug. 9, 2000, accessed at http://web.mit.edu/newsoffice/2000/cancer-0809.html on Oct. 31, 2005.
[5] Salomon Smith Barney study cited in InformationWeek, "Every Little Gigabit Helps," Nov. 3, 2003.

3883 [January 2006] J15046