



Infrastructure Technologies for Life Science Computing

A Summary of Sun's Recent White Papers & Use Cases

White Paper, September 2004

by Loralyn Mears, PhD

TABLE OF CONTENTS

INTRODUCTION	2
SOLARIS™ 10 (for SPARC™ & x86)	3
OPTERON™ SYSTEMS (including Starter Cluster for Bioinformatics)	6
N1™ Grid Engine 6 (Sun™ Grid Engine and Utility Computing Model for Storage)	8
JAVA™ ENTERPRISE SYSTEM	9
REFERENCES & KEY URLs	10

© 2004 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, CA 95054 USA

All rights reserved. This product or document is protected by copyright and distributed under licenses restricting its use, copying, distribution and decompilation. No part of this product or document may be reproduced in any form by any means without prior written authorization of Sun and its licensors, if any. Thirdparty software, including font technology, is copyrighted and licensed from Sun suppliers. Sun, Sun Microsystems, the Sun logo, Solaris, Java, Forte, Sun Fire and J2EE are all trademarks and logos that contain Sun, Solaris, or Java, and certain other trademarks and logos appearing on this website, are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and other countries. All other trademarks mentioned in this document are the property of their respective owners. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the United States and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc. RESTRICTED RIGHTS: Use, duplication, or disclosure by the U.S. Government is subject to restrictions of FAR 52.227-14(g)(2)(6/87) and FAR 52.227- 19(6/87), or DFAR 252.227-7015(b)(6/95) and DFAR 227.7202-3(a). DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS HELD TO BE LEGALLY INVALID.

INTRODUCTION

Enabled by Partner Power

Sun Microsystems has been a key technology provider to the life sciences industry for more than a decade. This white paper highlights new technologies that can and are already benefiting the life science community. For example, Sun-AMD Opteron servers have recently set benchmark records for price *and* performance running life science (and other) applications. The Solaris 10 Operating System has yet to be released but demand for early access by both life science customers and ISVs has been overwhelming, largely due to a few novel features that are outlined later and include one common set of compilers and developer tools that are designed for SPARC or x86 architectures. This white paper also focuses on existing technologies that offer cost saving and enhance productivity.

Sun's strength is not only in high performance computing with new leadership in high throughput computing, but in our numerous deep and broad partnerships that we have developed along the way. Sun has more than 14,000 software vendor partners where some of the larger ones include AMD, Fujitsu, Oracle, SAS, Documentum, OpenText and Microsoft, which support multiple industries but have a special focus on life science. In addition, we have more than 400 software partners focused solely on life science, along with 22 academic Centers of Excellence (COE) engaged in life science research projects ranging from systems biology to virtual surgery and bioinformatics on a super-computing scale. We have also donated more than \$100 million in hardware grants through our Academic Excellence Grant program (www.sun.com/edu) and continue to make substantial awards each quarter. Regular meetings with our COE and computational biology special interest group members are held around the globe to foster collaboration within our community (see References to learn how to join).

The life science community is very comfortable with open source codes, while Sun is equally comfortable with contributing to and supporting these efforts. For example, Sun has recently donated Java3D to the open source community which is expected to have profound impact on the life science community given the rich visualization needs of this industry. And, few are aware that we are the second largest contributor of Linux code, behind only University of California Berkeley, and continue to donate millions of lines of code, as well as financial assistance to efforts that increase commonality between Linux and Solaris.

Sun is focused on ensuring that both open source and commercial life science applications in greatest demand are optimized for deployment on our platform. Sun offers excellent engineering support and dedicated annual programs, including Tunathon, which was established five years ago with the purpose of exposing and removing performance bottlenecks at the level of the source code. Modified code is then posted on the public domain (for example, BLAST, HMMER, etc.) or incorporated into the next release of the application for much improved performance.

Our life science partners and customers have been driven to control costs – just as Sun has. Many are making gains by focusing their domain talent to develop life science applications on Solaris using Java™ Enterprise System (Java ES) to free them from the burden of developing and maintaining LDAP or other horizontal IT standards-based software. Java ES enables streamlined, cost-efficient management of combined high throughput and high performance computing environments. The need for single sign-on in R&D is growing as the number of boutique and home-grown applications explodes with emerging data types and lab instrumentation. Sun's partners who have adopted our identity management solution or knowledge sharing solution (both embedded within Java ES) have enjoyed competitive advantage.

In the past year, Sun has been delivering low-cost, x86 solutions that give customers choice, a choice of Solaris or Linux operating systems, a choice of 32- or 64-bit microprocessors, with full front-line support available for each. What's more, year-over-year and quarter-over-quarter, Sun outgrew the market in terms of units shipped. Sun achieved significant revenue and volume growth in almost every category reported in the 2Q04 Worldwide Server Database report by Gartner Dataquest. Sun outperformed all vendors in revenue market share growth on a quarter-to-quarter basis. Year-on-year, Sun outpaced all competitors in unit shipment growth at 38.4%. Additionally, according to an August 24, 2004 IDC report, Sun posted its 40th consecutive quarter as the undisputed leader in traditional workstation unit shipments while the top five competitors have all posted declines.

We invite you to read this white paper and to contact Sun to discuss how we can advance your efforts in life science computing.

SOLARIS 10

Screaming Performance at Compelling Prices

Solaris is the most trusted, robust, scalable and lowest cost operating system on the market today.

- **PERFORMANCE:** Sun's newest workstations run BLAST on Solaris x86 v9 (32-bit) 61% faster than equivalent systems from Dell running Red Hat Linux. Imagine what performance will be like when Solaris x86 v10 (64-bit) is released in January, 2005!
- **LOW COST:** Up to 80% below cost of Windows Server 2003 and 70% below cost of Red Hat Linux. Sun offers Solaris x86 right-to-use licences at very competitive prices: \$99 for single-CPU systems, \$250 for 2-CPU systems, \$1,500 for 4-CPU systems and \$2,499 for 8-CPU systems.
- **FLEXIBLE:** Optimized Sun and third-party hardware including x86 (Intel Xeon™ and AMD Opteron, plus SPARC-based architectures for 1- to 4-CPU servers).
- **SAFE:** There is no IP risk, indemnified.
- **CONSISTENT QUALITY DESIGN:** Forward and backward binary compatibility save you the hassle of patching and recompiling daily. No other operating system can offer all these features. NOTE: Non-commercial (educational, developer) licenses are available at no charge and can be download at <http://sun.com/software/download>. Attractive OEM license terms are also offered.

Solaris 10 will bring Solaris scalability and robustness to entry-level systems (AMD Opteron and Intel Pentium™ 4 SSE/SSE2) with its release January, 2005. Developers can get early access to the Solaris 10 OS soon (see References). Advanced functionality including hyperthreading, system call and library optimization, D-trace and containers (zones) will enable developers to identify bottlenecks in their code, readily correct them to optimize performance, and run beta and production code at the same time, with no risk of a Solaris *panic*. Several customers and partners have already begun working with Solaris 10 for SPARC and x86 (*pre-release* versions) while others have already adopted Solaris 9 for x86. Here are some use cases that describe their experiences:

- Use Case 1: Sun Java Application Server more scalable on Solaris 10/Opteron with JDK (Java Developer Kit) 1.5 beta vs. Red Hat Enterprise Linux 3/Opteron with 1.4.2_03 JDK. Near linear scalability reported with 99% more users and 6% more hits/second.
- Use Case 2: Solaris 9/Opteron scales better at 1 million entries than Red Hat Enterprise Linux 3/Opteron. Solaris outperforms in I/O intensive operations such as Modify and “Siteminder” Simulations.
- Use Case 3: Solaris 10/Opteron scales better to 4 CPUs than Red Hat Enterprise Linux 3/Opteron. Web server performance is 19% to 8% faster with Solaris 10/Opteron vs. Red Hat Enterprise Linux 3/Opteron. The following performance gains were recorded: Sun Java Enterprise System Web Server gained 19%, Apache™ Web Server gained 38%, and Solaris/Opteron runs select Java benchmarks 7% to 29% faster than Solaris/Xeon.
- Use Case 4: Two data-intensive applications were migrated to Solaris 10/Opteron and one gained 41% more throughput while the other gained 28% more. A third application (single-threaded, little data I/O) was 54% faster using *gcc* (although 65% faster is expected with optimal use of compilers).
- Use Case 5: BLAST: the Sun W1100z with 2.4 GHz Opterons running Solaris x86 v9 shows up to 61% better performance than Dell 650 workstation with a 3.2 Ghz Xeon running Linux. The Sun W1100z with 2.4 GHz Opterons running Solaris x86 v9 shows up to 48% better performance than Dell 360n workstation with a 3.4 GHz Pentium running Linux. The Sun W2100z with two 2.4 GHz Opterons running Solaris x86 v9 shows up to 34% better performance than Dell 650 workstation with two 3.2 GHz Xeon running Linux.
- Use Case 6: Sun Studio 9 running on Solaris for x86 generates code that is 65% to 70% faster than *gcc* running on Red Hat Linux.
- Use Case 7: In mid-2003, Developer 6.2 running on Solaris for SPARC was 80% slower than Intel tools running on Linux in mid-2003. Today, Solaris for SPARC is at parity on the same benchmark using Sun Studio 9 on Solaris for x86 v9.

The white paper, **Porting Guidelines for Solaris Operating System, x86 Platform Edition**, is an overview of what it entails to port to the Sun Solaris Operating System and describes actual case studies (see References). The white paper is especially useful if you are considering a port from Solaris OS SPARC Platform Edition, or from Linux, to Solaris OS x86. It provides a comprehensive idea about where to look for Solaris OS x86-

related modifications and indicates the effort required. When you are ready to port to Solaris OS x86, there are some areas you may want to look into to get additional insight into the porting process.

First, understand the Application Components: Are the components in native C, C++, assembly language and/or Java programming language? If all the components are Java technology-based, then it is a very *easy* port (as good as a “non-port”). Any Java technology-based components will work without any modifications. For example, client GUI components in the Java programming language will work as they are on Solaris OS x86 with no changes needed in the source files. Second, discover the source code: What are the platform-specific code segments? These segments may need to be modified to support Solaris OS x86. You can use *#define* statements (with *#ifdef*) to separate Solaris OS x86 platform-specific code segments. You can use *#define* statements for different possible hardware platforms (x86 or SPARC) or for different operating systems such as Solaris OS SPARC Platform Edition, Solaris OS x86 Platform Edition, Red Hat, Linux or Windows.

Third, include header files: If there are specific header files needed for Solaris OS x86, you may need to use *#include* statements in the source files in order to include them. Fourth, modify Driver Code: If you have a platform-specific driver code, then this may need to be modified to work on Solaris OS x86 Platform Edition. If the Solaris OS SPARC driver is developed using the standard APIs like DDI/DDK, then it takes just a recompile to work on Solaris OS x86. NOTE: A validation test should always be conducted to ensure scientific integrity. If a driver does not adhere to standard interfaces, then it may need to be modified to make it work on Solaris OS x86. Driver modifications can be minimized by adopting an x86-based Linux driver code for (if this application supports Linux) Solaris OS x86. Due to differences in the driver frameworks of Linux OS and Solaris OS, it may require some modifications in order to port Linux drivers to Solaris OS x86. NOTE: Modifying hardware-dependent code in assembly will need some modifications to work on Solaris OS x86 Platform Edition. For example, assembly code on SPARC hardware will need to be modified to work on the x86 hardware platform. This is due to hardware architectural differences. You can minimize the changes needed to port to Solaris OS x86 by adopting the assembly code from Linux (instead of Solaris OS SPARC Platform Edition).

BASIC STEPS TO PORT TO SOLARIS x86

1. Check the source code: If it is Java technology-based then it is a “non-port.”
2. Verify that third-party (and/or open source) libraries/dependencies are available on Solaris OS x86 platform.
3. Modify Makefiles and build process to generate Solaris OS x86-based application. *Autoconf*, *qmake* and *Jam* types of tools can easily generate Solaris x86-based Makefiles.
4. If needed, add Solaris OS x86 Platform Edition specific changes using *#define* statements (with *#ifdef*).
5. If needed, add *#include* statements to include Solaris OS x86-specific header files and/or other third-party header files.
6. If needed, modify platform-specific drivers for Solaris OS x86 platform. It makes for much easier porting if
7. adopted from Linux-based driver code.
8. If needed, modify compiler flags to generate optimized Solaris OS x86 technology-based applications.

The Porting Process:

- **Recompile the Application:** If porting from Solaris OS SPARC Platform Edition to Solaris OS x86, the application may just need to be recompiled. If the application is developed using published Solaris OS interfaces on SPARC, it does not need any changes. If you are porting from Linux to Solaris OS x86, you may need to discover the differences in a few system/function calls used in both these platforms. You will need to discover the equivalent system/function calls used in Solaris OS x86. You can look through Linux man pages for those system/function calls and review the “Conforming To” section. You can also use tools like *Autoconf* (<http://www.gnu.org/software/autoconf/autoconf.html>), which can indicate portability issues after scanning through the code, as well as help in creating portable build environments.
- **Changes in Makefile(s):** There may be some changes required in the Makefile(s) to generate Solaris OS x86 platform-specific object files, and linking these object files to get the Solaris OS x86 platform binaries. Makefiles may need to be modified to even include appropriate libraries specific to Solaris OS x86. Thus, porting to Solaris OS x86 may affect the build process that you use to generate platform-specific binaries. The tools like *Autoconf* (<http://www.gnu.org/software/autoconf/autoconf.html>) or *qmake* (which generates platform specific Makefiles) (<http://doc.trolltech.com/3.3/qmake-manual-1.html>) or *Jam* (<http://www.perforce.com/jam/jam.html>) can make this process much easier.

- **Third-Party Libraries:** You also need to discover what third-party libraries (and/or open source libraries) are needed to generate Solaris OS x86-based applications. If these libraries are not available yet on Solaris OS x86, you may want to consider alternate libraries, which may already be available on Solaris OS x86.
- **Modify Compiler Flags:** You may also need to modify the compiler flags used to create optimized executables for Solaris OS x86 Platform Edition. If you are already using Sun Studio or Sun Forte™ tools, then building the application on Solaris OS x86 is much easier. If you are using a different compiler for the Solaris OS x86 platform, you need to consider the appropriate command-line options to use for generating optimized applications for Solaris OS x86.
- **Byte-Order Dependent Code Segments:** If you are porting your application from Solaris OS SPARC Platform Edition to Solaris OS x86, you may want to look into code segments which are byte-order dependent when stored in the memory. For example, SPARC hardware supports *big-endian* (Most Significant Byte [MSB] is stored in the lowest memory address) whereas x86 hardware supports *little-endian* (Least Significant Byte [LSB] is stored in the lowest memory address).
- **Source Code and Version Control:** You may also investigate what source code and version control system is currently used. CVS (Concurrent Versions System) and SCCS are available on Solaris OS x86. If the source control system you are using now is not supported on Solaris OS x86, there are various simple workarounds you can use. For example, you can use the source control system on a supported OS, and check out to NFS or Samba (www.samba.org) which will make the files available to Solaris OS x86.

Reviewing all the case studies, we can conclude that in most of the cases, porting to Solaris OS x86 is a very easy process. This is especially true if you have an application that is written completely in the Java programming language. In this case, it is as good as a “non-port” because it is truly a “write once, run anywhere” application. You can expect to spend more time tweaking your Makefile(s) to include Solaris OS x86-related header files and libraries, as well as modifying certain compiler flags to get the best performance on the Solaris OS x86 platform. Similarly, you will spend some time modifying your existing build process to generate executables for the Solaris OS x86 platform. Most importantly, ascertain if the application has any dependencies. Find out the availability of these dependencies on Solaris OS x86.

Did you know...

Sun believes in collaborative development as a lifestyle. In fact, Sun has contributed more code to the open source community than any other commercial vendor. Sun also contributes money and engineering as part of its five year standing commitment to increasing commonality between the Solaris and Linux operating systems. Sun has been offering Linux-based systems since August 2002. Sun ships and fully supports Red Hat and SuSE Linux, which together make up more than 50% of the Linux market (Linux Development Survey, Spring 2004). Both of these Linux distributions are qualified on all Sun x86 hardware platforms (Intel Xeon and AMD Opteron). By buying their Linux systems from Sun, customers gain “one-stop” support for their heterogeneous IT environment. List prices for one-year Linux license starts at \$299 for Red Hat (WS Standard), \$799 (ES Standard) and \$799 for SuSE.

Did you know...

There are more than a half-million registered licenses for the Solaris 9 x86 OS. The Solaris OS also runs on third-party x86 systems, including AMD Opteron-based systems, further enabling flexibility and choice for customers. There are approximately 200 systems listed on the Solaris x86 hardware compatibility list today, plus support for hundreds of add-in cards and peripherals. And, the Solaris 10 OS, due out in January 2004, has driven more enquiries than any other Sun technology in recent months.

Did you know...

The Southwest Foundation for Biomedical Research, Department of Genetics, is using the Sun Solaris OS x86 Platform Edition running on a cluster of 700 dual-processor machines. To date, this is the world's largest cluster dedicated to statistical genetic analysis of data from human and animal families. The goal of this research is to identify genes influencing susceptibility to common diseases such as heart disease and diabetes. By using Solaris OS x86 Platform Edition, the development team was able to quickly and easily develop a cluster environment, saving valuable time, money, and resources.

OPTERON SYSTEMS

Extreme Computing Performance

The far-reaching Sun Microsystems-AMD alliance combines 20 years of enterprise expertise with the industry economics of x86 to give customers extreme computing performance at compelling prices and a growth path to 64-bit x86 computing. Sun and AMD are maximizing the performance of the Java platform and tools, optimizing Solaris OS performance on Opteron processors and undertaking joint technology initiatives. Sun is offering customers leadership and flexibility virtually unmatched by IBM, HP or Dell through industry-standard technologies that lead on price/performance, and Sun's complete spectrum of services and solutions. Sun's AMD Opteron processor-based solutions run the platinum standard for enterprise operating systems: Solaris, as well as Linux and Microsoft Windows through Windows Hardware Qualification List (WHQL) certification. Sun currently offers a growing range of x86-based hardware products, priced to compete fiercely with similar products from other vendors. Sun's x86 hardware products include the Sun Fire™ B100x and B200x blade servers; Sun Fire V60x, V65x, V20z, V40z servers; the Sun Fire V60x and V20z Compute Grid rack systems; and the V60x and V20x Starter Clusters for Bioinformatics.

- **The Sun Fire V40z server:** Available with the Solaris OS and Java Enterprise System, the new Sun Fire V40z **outperforms IBM and HP** in real-world and industry-standard benchmarks. The Sun Fire V40z server starts at \$8,495 U.S. list price and is optimal for usage in life science with its industry-standard 32- and 64-bit architectures for high performance computing and enterprise applications, Web services, data management, workgroup collaboration, grid-structured database management and more. The V40z is expanding the Sun Fire line to include both 2- and 4-way rack-mount AMD Opteron processor-based servers.
- **Sun Java Workstation W1100z and Sun Java Workstation W2100z:** These new workstations start at \$1,995 U.S. list price and \$4,695 U.S. list price, respectively, and build on Sun's number one position in the workstation market (based on IDC reports) and continue Sun's heritage of delivering some of the fastest 64-bit workstations on the market, with the ability to boot three operating systems from the same unit. The Sun Java Workstation W2100z demonstrated best overall performance running the OCUS (Olaf Corten's Utilities) benchmark V4 for PTC Pro/ENGINEER Wildfire, one of the world's most prominently used mechanical computer aided design applications. The Sun W1100z with 2.4 GHz Opterons running **BLAST on Solaris x86 shows up to 61% better performance than Dell 650** workstation with a 3.2 GHz Xeon running Linux. These new workstations also outperformed all other Xeon and Opteron processor-based systems on floating point operations based on the SPEC CPU2000 benchmark. More information can be found at: <http://www.sun.com/desktop/workstation>. Workstations are offered with the Java Desktop System running on either the Solaris x86 OS or Linux. More information on the Java Desktop System can be found at: <http://www.sun.com/javadesktopsystem>.
- **Sun Fire V20z Enterprise Essentials Promotions and new V20z server configurations:** In response to customer demand for better control over licensing, planning and deploying network computing applications such as large databases or grid computing, Sun is introducing the Sun Fire V20z Server Enterprise Essentials Promotion package. This includes a three-year subscription to the Solaris x86 OS and Sun Silver Spectrum services support for \$492/year U.S. list price plus a choice of entry-level or premium hardware configurations. The promotion creates dramatic reductions in customers' total cost of acquisition - nearly **40% lower** over three years than comparably configured systems from Dell running Red Hat Linux. The promotions provide customers with a highly secure, reliable, available and scalable architecture unmatched in the x86 market. Sun is also offering new configurations based on AMD's higher performing Opteron processor Model 250 and the Solaris x86 OS.
- **Sun Fire V20z and V60x Starter Cluster for Bioinformatics:** Starter Cluster is a powerful 4-node (8-CPU) x86 (32-bit Xeon or 64-bit Opteron) cluster in a small rack on wheels designed for operation in any space-constrained lab environment using standard room electrical power outlets running the Linux or Solaris x86 operating system. Resource management software (Sun Grid Engine or N1GE6) and MPI software enable the efficient full-time use of all available CPUs to execute batch jobs. It is pre-loaded, pre-tested and shipped assembled ready-to-use via our Customer Ready Systems center. By the same process that includes integration, pre-testing and shipping as a complete solution, life sciences software is integrated into this cluster and deployed as Sun Starter Cluster for Bioinformatics to yield "bioinformatics-in-a-box." Software includes: a) *BioBox*, a collection of >20 public algorithms optimized for Sun's x86 hardware running Linux, Solaris and Solaris x86 is

available through Sun's Asia Pacific Science & Technology Center; b) the INCOGEN *VIBE* (Visual Integrated Bioinformatics Environment) commercial application is built on 100% Java technology, grid-enabled and ships with a software development kit (SDK) to make it easy for the user to extend functionality without the aid of consultants. *VIBE* combines the most common bioinformatics tools into one unified and intuitive environment that facilitates simple, drag and drop pipelining capabilities; c) *GenomeQuest* by Gene-IT is also grid-enabled and provides users with the ability to compare bioinformatics results with patent data to streamline research focus; d) *iNquiry* by The BioTeam is a grid-enabled, ready-to-use cluster and Web portal optimized for use in life science informatics settings preconfigured with hundreds of open source scientific applications and can easily be extended to support additional commercial, open-source or internally developed applications. Finally, because of Sun's long-standing commitment to providing the most scalable and widely compatible architectures, this hardware/software combination can be readily integrated into an existing heterogeneous computational environment. It can be upgraded to include technologies for persistent and compliant storage, CPU provisioning and grown to thousands of nodes. And, because Sun solutions can be deployed via the Sun Customer Ready Systems center, the Starter Cluster can be customized to include any third party software, hardware or peripherals for a price and value that is hard to beat. Watch for our 2004-2005 **webinar series** that will showcase customer use cases!

Did you know...

Sales of the Sun Fire V60x server and Sun Fire V60x Compute Grid Rack System propelled Sun to the number one position in total server units shipped in the grid computing market (IDC Q3CY03 Worldwide Technical Server report). Sun's HPTC units grew 111% compared with 66% for the market as a whole over the same period in 2002, making Sun the only vendor to outgrow the market in units shipped both year-over-year and quarter-over-quarter figures (IDC Q3CY03 Worldwide Technical Server report).

Did you know...

The San Diego Supercomputer Center at the University of California at San Diego, in cooperation with San Diego State University, is hosting a Web site that runs on a 12-CPU Sun Fire B1600 blade server cluster, including Sun Fire B100x blade servers, to allow members of the public anywhere on the Internet to view and manipulate actual data beamed back to Earth from the NASA/JPL Mars rovers and do mission planning for the rovers with the same software that NASA uses.

N1 GRID ENGINE 6

The Ultimate Efficiency Engine

To date, Sun Grid Engine software has been downloaded tens of thousands of times, with widespread use across the life science community (see www.bio-itworld.com "Inside IT" March, April, May and, September 2004 articles). **N1 Grid Engine 6 (N1GE6)** is the next evolution and boasts unprecedented scalability – it can now handle more than 5 million jobs at a time distributed over more than 10,000 hosts – and was recognized with the Frost & Sullivan 2004 Excellence in Technology award. Key features are advanced scheduler functionality that includes resource reservation, urgency-based priorities, and throughput scheduling; cluster queue configuration (the most commonly requested feature by customers); and, conformance with grid standards. Monitoring and accounting tools are also now available. A feature that will be available in the September, 2004 release will include support of Windows. N1GE6 is a tool designed to drive efficiency and productivity and many have already realized its benefits, including Cognigen whose scientists began saving about 1 hour per day after the transparent introduction of N1GE6.

To dispatch jobs, the scheduler selects queues for the submitted jobs via attribute matching in an N1 Grid Engine 6 cluster as opposed to submitting to a queue, popular in other DRM products. However you can still submit to a specific queue if desired. Several enhancements have been made to the scheduler in N1 Grid Engine 6 that improve scalability, functionality, performance and overall ease of use through a common interface. The accounting system now includes an integration with relational databases and a reporting tool. To comply with standards, the N1 Grid Engine 6 software implements the DRMAA 1.0 specification and delivers everything needed to integrate applications using the DRMAA interface with N1GE6.

The N1 Grid Engine 6 *qmaster* daemon uses multiple threads in order to separate independent *qmaster* tasks. Dynamic events, such as message passing, and periodic tasks, such as job priority adjustment, run on separate threads. This multithreading improves efficiency, response time, and throughput by taking advantage of common usage patterns in N1GE6. The N1GE6 data subsystem is designed for scalability, programmability, and extensibility. The Grid Engine product was initially developed with only the concept of queues attached to a single host. The new version of the software also introduces the concept of cluster queues.

Utility computing certainly benefits from grid technologies, but application is not limited to servers alone – it can extend to storage. Sun offers a complete offering for storage through our own technology and alliance with Hitachi Data Systems. Storage products include software (ranging from availability, to business continuity and beyond), Tape (L180, L6000, etc.), SAN (3511, 6900 series, etc.), Direct Attached (3120, 3900 and 9900 series, etc.) and Network Attached Storage in a variety of configurations to satisfy a wide range of customer needs and budgets. With Sun utility computing offerings, using your IT staff to maintain hardware and software is a thing of the past – you simply access these resources as you need them, and pay only for what you use. One such offering, Sun Utility Computing for Sun StorEdge systems, is a pay-for-use (as low as \$2/Sun Power Unit/month), high-end storage infrastructure service with a 100% variable price. This service allows you to enjoy the benefits of pay-as-you-go storage capacity with predictable pricing, and reduced IT costs. You choose where your storage assets reside: either onsite or at one of our partner sites. To learn more about utility computing for high-end Sun StorEdge systems, visit: <http://www.sun.com/service/utility/utilitystoreedge.html>.

Did you know...

The Delaware Biotechnology Institute purchased multiple Sun Fire V60x systems as part of project *BioWolf*, the institute's new grid computing infrastructure. Sun created a 128 dual-node Compute Grid rack, which was recently ranked among the 30 most powerful academic supercomputers in the United States.

Did you know...

Today, Sun offers affordable online basic training (WE-1600) on how to install Sun Grid Engine software. Beginning October, 2004, Sun will also offer classroom-style workshop advanced training (ES-337) to instruct users on how to optimize the use of Sun's Grid Engine software. The new N1GE6 training modules were produced with The BioTeam (recognized as life science industry leaders for cluster configuration). Additionally, Sun has several white papers that describe the use of Sun Grid Engine and its optimization for life sciences (please see References) To learn more about these courses, please see: <https://slp.sun.com/sun> .

JAVA ENTERPRISE SYSTEM

Secure Collaboration for Knowledge-Sharing

The **Sun Java Enterprise System (Java ES)** changes the enterprise software industry with a radical, new approach to how enterprise software is designed, developed and operated – an approach that synchronizes IT investment with business priorities. The Java ES combines software and services as a single entity, for a single price, per employee, per year based on the total number of full-time employees. And by doing so, Sun solves the problem of integrating complex infrastructure software before the product even ships. The system is designed to simplify management and unify high throughput with high performance computing environments.

Life science is notorious for “communication silos,” having so much data its value is unknown, and for wasting domain resources on IT administration. Users need a simple way to **collaborate and communicate securely** and can do so with Sun's instant venue for real-time exchange of work, information, and ideas. Users need to create, publish and **manage** content, and those who have adopted Java ES have found that they save money and can readily enable secure collaborations to exchange knowledge. Portal enables users to search and organize knowledge assets, automate indexing of all file types and Web sites including discussion threads – both internal and external with full search capabilities: specific passage retrieval, dynamic summarization with assets categorized and ranked based on the rules you set using built-in knowledge management with taxonomy capabilities, and keyword highlighting. Only documents that the user has the access rights to view are shared. Users can subscribe to areas of interest, then receive alerts on the relevant new content. Users then need to share knowledge and can do so by **creating communities** (employees, partners and customers, or combinations thereof) while enforcing policy and ensuring that service delivery is both personalized and secure. Web single sign-on (SSO) readily supports multiple roles, personalized authentication methods, self-registration for these communities, and increased IT-independent usability. SSO can be readily customized to integrate and associate the applications, algorithms and databases you require to do your work. Users also need to streamline **resource management**. Having a common infrastructure of identity information helps eliminate data inconsistencies that can create multiple problems, security breaches in particular. It also greatly decreases redundancies, administrative effort and cost through the use of shared components. All of the above services are managed from the main Portal console for seamless administration. In summary, Sun has developed a solution for secure knowledge sharing that is particularly suited to the life science industry with its data volumes, complexity, compliance regulations, security, collaborative approach towards research and many other needs.

The **Sun Java System Portal Server** (“Portal”), is one of the primary components of Java Enterprise System, exists as IT's first identity-enabled portal server solution and the first to enable secure knowledge sharing (a patented Relaxation-Ranking Passage Retrieval algorithm was developed by Sun Microsystems Labs over the last decade at NASA's request). Portal (and Java ES) is supported on leading Java 2 Platform, Enterprise Edition (J2EE) technology-based application servers from Sun, IBM, and BEA and enables users to search content and receive only those results they are authorized to access through a secure, next-generation language- and data format-independent, **semantic search engine**. Sun leads in support of industry standards facilitating development and integration, including JSR 168, LSB (Linux), the Liberty Alliance, LDAP, SOAP, and others. Sun's advanced search technology helps enterprise users locate information up to **five times faster** than standard Boolean search engines. Portal also delivers secure, personalized content and access, applications and services to end users by dynamically aggregating information based on an end user's role(s). Users can customize the layout and interface to fit their needs simply and easily. Enterprises and service providers can offer new applications and services on demand without the need to restart the portal server. And finally, portal developers are able to create and generate easily deployable portlets with the click of a button.

Did you know...

Portal customers specifically focused on life science research include Applied Scientific Research, Caprion Pharmaceuticals, Government of Singapore, HPCVL, OSU (Ohio State University Super Computing Center), Ribotargets, SurroMed, Spotfire (and their customers), among others. For more details, please see <http://www.sun.com/DynPub/modules/casestudy/index.jsp> .

References & Key URLs

ARTICLES

- Porting Guidelines for Solaris Operating System, x86 Platform Edition: An Overview with Real Case Studies, White Paper April, 2004 by *Tushar Patel*, Sun Microsystems Inc.
- N1 Grid Engine 6 Features and Capabilities, White Paper May, 2004 by *Paulo Tibério Bulhões, Chansup Byun, Rick Castrapel & Omar Hassaine*, Sun Microsystems Inc.
- Searching Versus Finding: Why Systems Need Knowledge to Find What you Really Want, White Paper April, 2004 by *W.A. Woods*, Sun Microsystems Inc.
- The BioBox Initiative: BioClusterGrid, White Paper by *Simon See, Gilbert Thomas, Deepak Jeevan Kumar, Maddie Jing Theng Wong*, Sun Microsystems Asia-Pacific Science & Technology Center
- Integrating BLAST with Sun Grid Engine, White Paper July, 2003 by *Bogdan Vasiliu*, Sun Microsystems Inc. http://developers.sun.com/solaris/articles/integrating_blast.html
- Sun Briefing in Computational Biology – Volume 2, White Paper November, 2003 by *Stefan Unger & Ilya Sharapov*, Sun Microsystems Inc.

LIFE SCIENCES

Life Sciences

<http://www.sun.com/lifesciences>

Life Science Monthly Newsletter (subscribe to Sun Boardroom Minutes)

<https://see.sun.com/cgi-bin/gx.cgi/mcp?q=STHm2BTFCCH1in>

BioBox

<http://apstc.sun.com.sg/biobox.php>

Computational Biology (and info on Starter Cluster for Bioinformatics)

<http://www.sun.com/products-n-solutions/edu/commofinterest/compbio/>

To Join Sun's Computational Biology Special Interest Group, send an email to:

compbio-sig-info@sun.com

SOLARIS

Early Access To Solaris 10 for Sun Partners (iForce)

<http://iforce.sun.com/partners/solaris/>

Guide to Adding Support for Solaris OS, x86

http://developers.sun.com/solaris/articles/support_for_x86.html

Compiler Differences

http://developers.sun.com/solaris/articles/x86_compiler_diffs.html

Case Study: Porting Apache Web Server from Solaris OS SPARC to Solaris OS x86

http://developers.sun.com/solaris/articles/casestudy_port_x86.html

Migration to Sun Platforms for Developers

<http://iforce.sun.com/partners/migration/>

Solaris Operating System for x86 Platforms

<http://www.sun.com/software/solaris/x86/index.html> or <http://www.solaris-x86.org> or <http://sun.com/solaris/x86>

Solaris OS Forums

<http://forum.sun.com>

TUNING/ADMINISTRATION

Hardware Compatibility Lists

<http://www.sun.com/bigadmin/hcl/>

Tunathon

<http://sdc.sun.com/tunathon2004/>

System Administration Portal

<http://www.sun.com/bigadmin/>

contact: loralyn.mears@sun.com with any questions or requests for white papers