Assessing Scalability of the Sun Ray™ 1 Enterprise Appliance

Technical White Paper

Sun microsystems

Printed in the United States of America.
901 San Antonio Road, Palo Alto, California 94303 U.S.A

Contents

1. Introduction 1

The Hot Desk Technology and the Sun Ray 1 Enterprise Appliance 2

Sun Ray Hot Desk Architecture 3

2. Sun Ray Enterprise System 8

Traditional Desktop Architectures 8

Hot Desk Technology 10

Sun Ray Enterprise Server Execution Environment 12

3. Verifying Scalability 15

Application Profiles 16

Components of Scalability 17

Typical Configurations 25

References 29

Introduction 1

Organizations today face the conflicting challenges of increasing productivity while controlling costs. Higher productivity often entails providing faster application performance to larger and larger groups of users. To some, the traditional PC model has seemed like a solution to this problem since it has provided steady increases in performance at decreasing hardware prices. Unfortunately, constantly replacing user's hardware with faster systems along with the significant software and support costs of the PC model have made large deployments difficult and expensive to manage.

Customers are demanding technology that not only scales to meet increasing computing needs, but also presents a predictable cost model. The *Sun Ray™ 1 enterprise appliance* and associated *Hot Desk technology* represent a fundamentally new approach that anticipates growth and can scale to address small and large workgroups without driving Total Cost of Ownership (TCO) to unacceptable levels.

This document is intended to provide background on the essential components that affect scalability in the Sun Ray enterprise system and presents the results of initial scalability-testing done by Sun.

Hot Desk Technology and the Sun Ray 1 Enterprise Appliance

The Sun Ray 1 enterprise appliance is a stateless, zero administration "plug-and-work" device that is centrally managed from a Solaris server running Sun Ray enterprise server software. Underlying Hot Desk technology provides users with the ability to access their desktop environment instantly from any Hot Desk enabled appliance in the workgroup — access to applications running on both Solaris™ and Microsoft Windows NT servers is provided. By sharing resources and leveraging the inherent reliability, availability, and scalability of Sun servers, Hot Desk technology delivers enterprise computing performance to the desktop with economics that dedicated "fat clients" generally cannot match.

The stateless Sun Ray 1 enterprise appliance and integral smart card technology give users a high degree of mobility within the workgroup and prevent data loss in the event an appliance fails. In case of appliance hardware failure, a new appliance can simply be installed, or the user can move to another desktop appliance, insert their smart card, and resume work immediately. Users are returned to their desktop computing environment exactly as they left it, reducing downtime increasing productivity.

The Sun Ray 1 enterprise appliance also makes sharing desktop hardware a practical and secure reality. Multiple users (agents working different shifts or job sharing employees) can share the same appliance, each finding their customized desktop environment in the same state as when they left, with no danger of corrupting another users files.

Access to the Solaris™, Operating Environment, other UNIX®, and Microsoft Windows NT applications within a single integrated desktop environment further boosts user productivity. The appliance's fanless, noise-free enclosure and sharp 24-bit color images lets users enjoy a more comfortable and productive work environment.

The Sun Ray 1 enterprise appliance's user benefits are accompanied by a low cost of ownership that is appealing to IT managers. Existing multi-user applications run without modification, eliminating application porting expense and user retraining. System administration is simplified because operating system and application software resides only on server systems — upkeep costs are eliminated, as the desktop appliance requires no maintenance and no upgrades. Finally, software installation and distribution are centralized and efficient bringing software licensing costs into line with actual application usage. Hot Desk technology delivers what enterprises demand — a productive user environment coupled with a dramatically lower total cost of ownership.

Sun Ray Hot Desk Architecture

The Sun Ray Hot Desk Architecture is the first step towards a model of computing where client sessions are maintained on the server and instantly available from any device, anytime, anywhere. Currently targeted at the workgroup, the Sun Ray Hot Desk Architecture will eventually be capable of extending beyond the workgroup across the enterprise.

The Sun Ray Hot Desk Architecture succeeds by combining key advantages of existing architectures with today's inexpensive hardware components and high-speed networking technology. As shown in Figure 1-1, the Sun Ray Hot Desk Architecture is comprised of three components: the *Sun Ray 1 enterprise appliance*, a *Sun*™ *server* with the Solaris (SPARC™ Platform Edition) Operating Environment, running *Sun Ray enterprise server software*, and a *dedicated interconnect*.

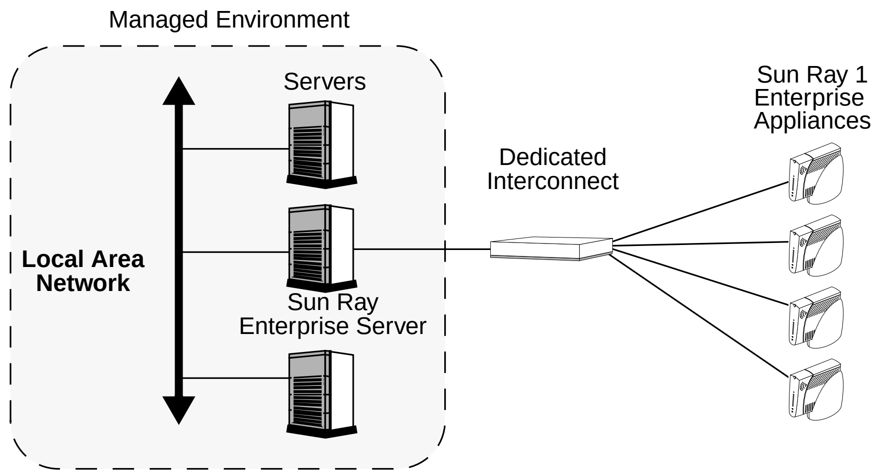

Figure 1-1 High-level view of the Sun Ray enterprise system

The Sun Ray 1 enterprise appliances contain video and audio I/O connectors, a 10/100 Mbit Ethernet connection, an integrated smart card reader, and four USB ports. The appliance connects to a display device, keyboard and mouse, and will support other USB peripherals (such as a printer or scanner) with later releases of the Sun Ray enterprise server software. Sun Ray enterprise server software is used to manage, administer and provide the screen display for any Sun Ray 1 enterprise appliance attached to the dedicated interconnect and includes an Authentication Manager, a Session Manager, and an Administration tool. The dedicated *interconnect* is a 10 or 100 Mbit switched Ethernet communications channel that connects the enterprise appliances with the Sun Ray enterprise server.

No client software is stored or executed on the appliance. Unlike X-terminals and similar devices, all user applications (including X11 server and the Graphical User Interface (GUI)) run on one or more centralized server systems. The enterprise appliance contains only the resources necessary for the *human interface* — input devices such as microphone, keyboard and mouse, and output devices such as the display and audio.

Graphical output from applications is transmitted to an X11 server process on the Sun Ray enterprise server which renders the results to a *virtual device driver*, and then transmits them via the dedicated interconnect to an attached Sun Ray 1 enterprise appliance. Similarly, all user input (keystrokes, mouse clicks, etc.) is transmitted from the appliance through the interconnect and on to the appropriate client application. The user sees a fully-functional desktop environment and window system which provides a composite view of the user's currently active applications.

Because applications execute independently of the location of their input and output, this design creates a *virtual session* for the user. A user can be *switched* from one desktop appliance to another instantaneously by redirecting the input and output for the session. Session switching is done based on authentication information obtained from the desktop appliance's smart card interface.

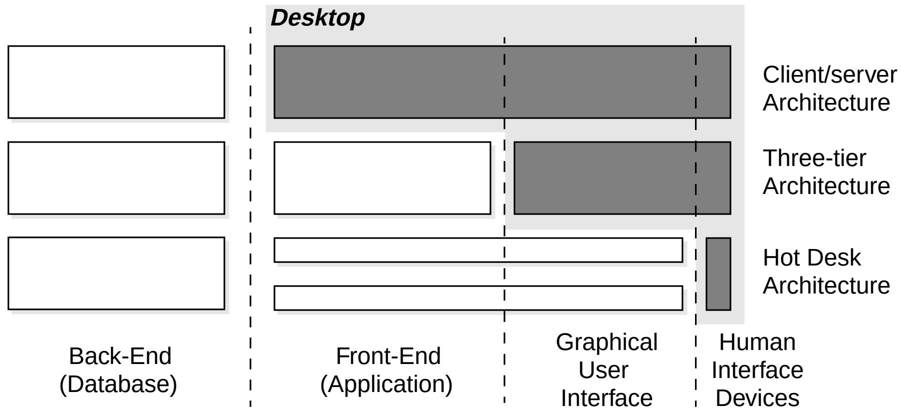

Figure 1-2 Sun's Hot Desk architecture leaves only the human interface devices on the client desktop.

Figure 1-2 illustrates the partitioning of functionality for traditional Client/server architectures, three-tier thin-client architectures (JavaStation™ NCs) and the Hot Desk architecture. By moving the GUI to the server and leaving only the human interface devices on the client desktop, the Sun Ray enterprise system retains many of the desirable features of previous approaches without their attendant drawbacks. For example, the cost-of-ownership and administrative benefits of three-tier, thin-client computing are retained without constraining the applications that can be delivered to the desktop. In addition, the Sun Ray enterprise system inherits the applications, performance, and resources of distributed (client/server) architectures while significantly reducing desktop acquisition costs, administration, and maintenance. By emphasizing shared resources, Hot Desk technology delivers server-class performance to a responsive desktop system which never needs upgrades or maintenance.

Sun Enterprise™ Servers and Solaris Operating Environment

Application performance in a Sun Ray enterprise system environment is directly dependent on the servers that provide computational resources. Sun Enterprise servers lead the industry in offering some of the most powerful and reliable systems available today. Sun's family of servers provide scalable, symmetric multiprocessing capabilities. From one to 64 high-performance UltraSPARC™ processors can be configured along with up to 64 GB of physical memory and up to 20 TB of disk storage, providing the necessary performance for peak demands as well as virtually unlimited growth. For the highest levels of availability, Sun servers also support clustering technology that can produce high levels of system availability.

The power of Sun's servers is further enhanced by the Solaris Operating Environment — the premiere environment for enterprise network computing. Designed with the needs of the enterprise in mind, the Solaris 7 Operating Environment features full 64-bit processing, mainframe-class reliability, superior scalability, and unprecedented performance. The Operating Environment has significant functionality that enhances multi-user environments, and is uniquely suited to hosting the Sun Ray enterprise system.

Microsoft Windows NT Support

In many organizations, users require full and easy access to popular office productivity applications running on Windows NT servers. The Sun Ray enterprise system supports the display of applications running on Windows NT 4.0 Terminal Server Edition (TSE) servers with the addition of MetaFrame software from Citrix Systems. After installing Citrix MetaFrame software on an available Windows NT 4.0 TSE server(s) and corresponding Citrix ICA client software on the Sun Ray enterprise server running the Solaris Operating Environment, the Windows NT server can redirect output from Windows applications to Citrix client applications running on the Solaris server as shown in Figure 1-3.

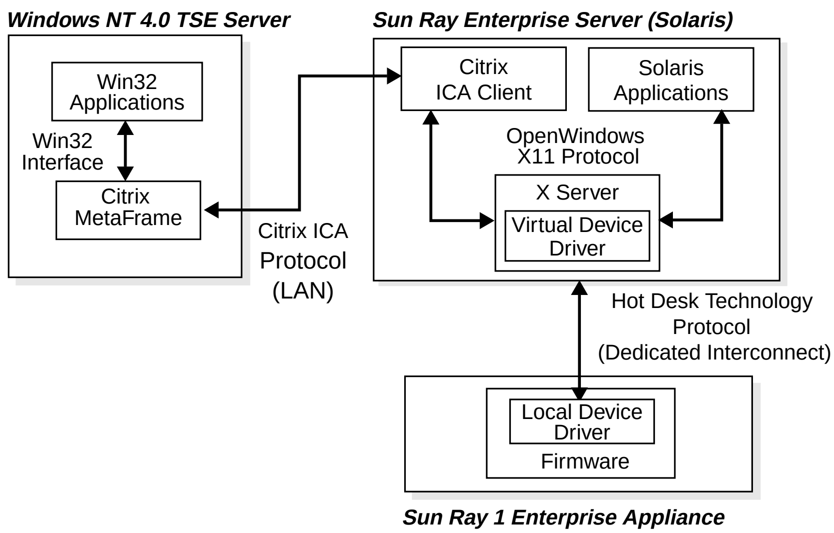

Figure 1-3 Optional MetaFrame software from Citrix Systems displays applications running on Windows NT servers onto enterprise appliance desktops.

Like other UNIX applications, the Citrix client application is displayed onto the Sun Ray 1 enterprise appliance's screen through a virtual device driver within the X11 / OpenWindows server. This approach combines high performance native execution for Windows applications with the ease and ease of administration of Solaris applications.

The Sun Ray Enterprise System 2

The Sun Ray enterprise system represents the currently deployed components of the Sun Ray Hot Desk Architecture and includes the Sun Ray 1 enterprise appliance, a SPARC Solaris server, Sun Ray enterprise server software and the components of the interconnect. As a prelude to site planning, it is important to understand how the Sun Ray enterprise system differs from traditional approaches and how these differences impact end-user performance.

Traditional Desktop Architectures

Traditional desktop systems (PCs and workstations) deliver a certain minimum level of performance and responsiveness by dedicating hardware resources to an individual user. Unfortunately dedicated desktop systems also assert a limit to performance which can become all too real with increased application demands or new operating systems. Even powerful desktop systems can be limited in some key resource (CPU, memory, or I/O). Essential applications which perform poorly can blunt user productivity — usually justifying an entire desktop system replacement.

Additionally, desktop systems are often some of the most under-utilized systems in the enterprise. Even when an active user is manipulating the system (moving windows, typing, pushing buttons), most desktop systems are in a near-idle state. Desktop clients that are alternately performance-limited, or virtually idle represent a poor deployment of computing resources. Figure 2-1 illustrates a desktop system under typical usage patterns; mostly idle but occasionally operating at maximum throughput.

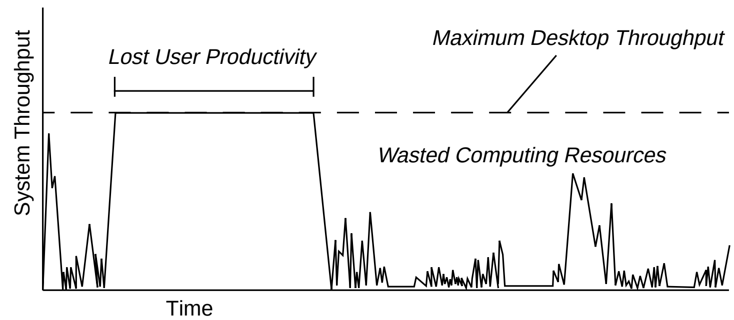

Figure 2-1 Typical desktop throughput — alternately idle or performance limited

For the purposes of this discussion, maximum throughput is defined by an application running out of any one resource (CPU, memory, I/O). When a system is operating at its maximum throughput, some resource is exhausted and the system will generally take longer to complete a given task. Users lose valuable productivity as they wait for tasks to complete or struggle to get them to run at all. Jobs which are CPU-bound take longer to run. Even worse are jobs that exhaust memory resources, and cause the system to “thrash” — rendering both the system and the application useless.

As seen in Figure 2-1, typical desktop usage consists of multiple bursts of activity in an environment of relative idleness causing valuable computing resources to be wasted. This is particularly true in single-user PC environments because there is no mechanism for other users to take advantage of idle computing resources on remote desktops. Networked multi-user Sun workstations solve this problem in part since they allow remote usage and can be equipped with a network queuing system that allows jobs to be distributed to idle systems. For example, Platform Computing’s Load Sharing Facility (LSF) workload management software can be used to unite large numbers of networked Sun systems into a single, powerful computing resource.

Hot Desk Technology

Hot Desk technology provides the underlying technology for the Sun Ray Hot Desk Architecture — including a high quality of service (QOS) interface, smart card technology, and server software which instantly maps users' computing sessions to appliances. Hot Desk technology differs from the traditional desktop implementation in that it is based on an understanding that the desktop system only needs to deliver services at a performance level set by "human bandwidth". Once the input/output requirements set by the limits of human perception are met, a faster processor or more memory will not provide any additional benefit to the user. With this fundamental understanding, Sun system architects focused resources where they could best improve application performance and administration.

Shared Resource Model

With Hot Desk technology, computational resources are combined, centralized, and shared rather than being secreted away on personal desktop systems. As a result, Sun Ray 1 enterprise appliance users enjoy a responsive and highly-interactive computing experience as if they were sitting at the console of a large server. In addition their applications have a higher performance ceiling and may actually run faster because they are hosted on more powerful platforms. The "bursty" load produced by multiple users is aggregated across a more powerful server or set of servers (Figure 2-2)).

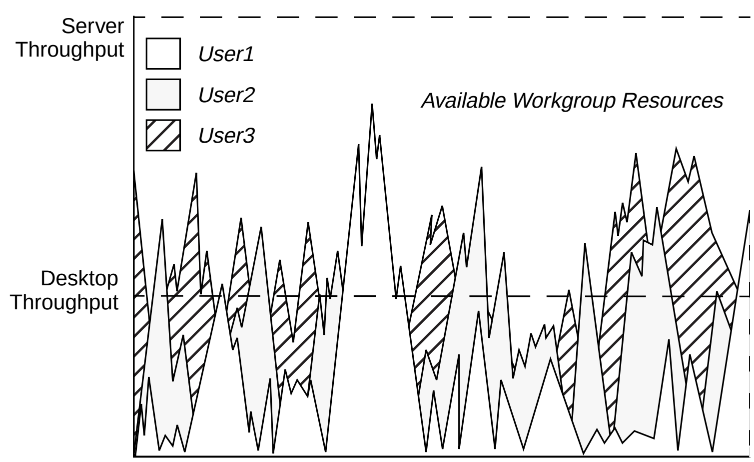

Figure 2-2 Hot Desk technology enables efficient use of resource

Far from being wasted, unused resources are better utilized as a result of being shared and are now automatically available for use at any time by anyone in the workgroup. Users benefit directly and automatically when resources (software, memory, additional processors) are added to Sun Ray enterprise servers.

Resource sharing is automatic within a Sun Ray enterprise server, and no special load sharing software is required though tools are available to reserve capacity for a given user or application. Solaris Resource Manager™ can be installed on a Sun Ray enterprise server to define system resources for groups of users sharing a server. Solaris processor sets (Solaris 2.6 or later) allow jobs to be bound to single processors or groups of processors.

Reliability and availability in a Sun Ray enterprise server environment can also greatly exceed that of dedicated desktop systems. Redundancy can be added to Sun Ray enterprise servers through disk mirroring, and redundant hot-standby hardware components — techniques which are seldom cost-effective to obtain or manage at the desktop level.

Hot Desk technology removes many traditional boundaries to growth and scalability. In the event that resources do become constrained, administrators can add memory or processors to existing servers as needed, or can move applications to new servers. Hardware and software investments in Sun Ray enterprise servers enhance the productivity of the workgroup as a whole, and can be applied where the need is greatest.

Dedicated Interconnect

Commodity Fast Ethernet and Gigabit Ethernet technology is used to provide the dedicated interconnect between Sun Ray enterprise servers and the Sun Ray 1 enterprise appliances, making modular changes and upgrades simple. Though standard networking hardware is used, it is important to understand that the interconnect is designed to be fully dedicated to the Sun Ray enterprise server and the Sun Ray 1 enterprise appliances.

With CPU, memory, and I/O concentrated in the Sun Ray enterprise server, only 2D pixels (and occasional audio/video data and peripheral I/O) are sent over the interconnect to the appliances while keystrokes and mouse events are sent back to the applications running on the Sun Ray enterprise server and other servers (Figure 2-3).

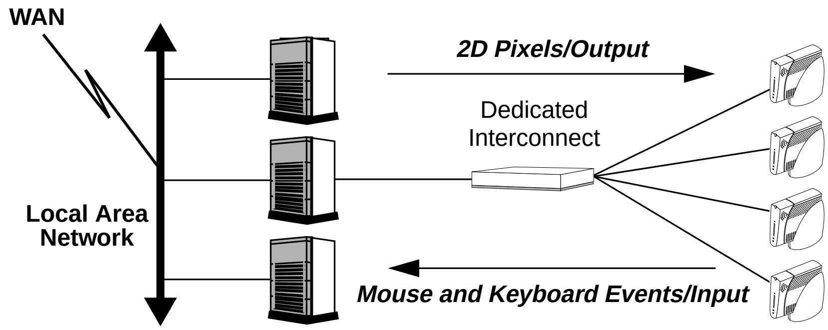

Figure 2-3 The flow of information between the appliances and applications running on the Sun Ray enterprise server and other servers

Sun Ray Enterprise Server Execution Environment

Because Sun Ray enterprise servers provide all of the computational resources for basic desktop application operation, understanding their operating environment is essential to assessing scalability. Sun Ray enterprise servers provide the following:

- *Solaris Operating System*

  Solaris 2.6 or 2.7 systems provide the operating environment for Sun Ray enterprise servers and the infrastructure for other Sun Ray enterprise server software. Solaris also provides highly-tuned low-level networking support to accelerate the dedicated interconnect.

- *Sun Ray Enterprise Server Software*

  Sun Ray enterprise server software provides the screen display for the Sun Ray 1 enterprise appliance and includes the Authentication Manager, Session Manager, virtual device drivers, administration tools and peripheral device support for the desktop appliance.

- *User Desktop Environment*

  For each active appliance user, resource load on the server minimally includes the window system (X11 server and CDE window manager) in combination with end-user desktop applications.

- *Other Applications*

  Sun Ray enterprise server systems need not be dedicated only to serving enterprise appliances and can be shared for other purposes if capacity allows. For example, a server running Sun Ray enterprise server software might also provide database or Web services.

Verifying Scalability 3

For users to remain productive, they must experience responsive interactive performance, even when the system is operating under peak load conditions. Any effective computing platform must maintain a stable and responsive environment as it scales to support larger and larger numbers of users and their applications.

Exploring scalability for any platform is a matter of examining how different functional components of the architecture drive scalability and application performance. To assure that Hot Desk technology scales effectively to meet the demands of both small and large workgroups, Sun has conducted studies that focus on real users working within their everyday application environments on Sun Ray 1 enterprise appliances. In addition to verifying basic scalability of the architecture, these studies also serve as a basis to approach capacity planning.

Application Profiles

The number of users that can be supported on a given platform is directly dependent on the nature and mixture of applications they use. To begin to study how different applications scale in a Sun Ray enterprise system environment, Sun profiled users working with a variety of applications on Sun Ray 1 enterprise appliances. Applications with very different computational and display characteristics were chosen for study.

- Netscape — Netscape provides a general web browsing environment which is bursty and interactive in nature. Web browsers display a wide range of content and generally cause large display updates as users change web pages.
- Adobe FrameMaker — FrameMaker represents a desktop publishing environment with a high degree of interactivity. Frame Maker operates mostly on formatted bi-color text but includes some bit-mapped graphics.
- Adobe Photoshop — Photoshop can make significant computational demands for image processing and uses large amounts of memory for image storage and manipulation. Photoshop usage involves a moderate amount of interactivity with occasional large, full-screen updates.
- Personal information management tools (PIM) — Most computer users use these tools to operate on e-mail, calendar, and text information. Usage for personal information management tools is typically bursty and interactive with frequent window manipulation and relatively low resource demands.
- Database font end (DB FE) — Administration tools analyzed in this study are implemented as lightweight Java™ technology-based applications. The Java-based front-end (running on the Sun Ray enterprise server system) interfaces to back-end database software running on separate database systems.
- Full database system — This application represents a scenario where both the front end and database itself run on the Sun Ray enterprise server.

Components of Scalability

The principal variables that can affect scalability in a Sun Ray enterprise system include memory and CPU resources in Sun Ray enterprise servers and the throughput of the dedicated interconnect. To understand the impact of these variables in a Sun Ray enterprise system, Sun analyzed the sample applications and used the generated application profiles to study capacity and performance.

It is important to note that all of the results given in this document represent *simultaneously active* users and are thus highly conservative. The application profiles were used directly in all simulations with no provision for varying user load. In practice, only a fraction of all potential users are generally active at any one time. In addition, the profiles portray dedicated, active users engaged with an application for a fixed period of time. Real users typically work in bursts with many breaks for think-time and other activities.

In the real world multiple servers might host applications to enterprise appliance user's. To simplify things somewhat, this document will examine a bounded case consisting of a single Sun server equipped with theSun Ray enterprise server software. Scalability metrics for a system with multiple servers can be extrapolated from the single server case in a straightforward manner.

Sun Ray Enterprise Server Memory Requirements

In order to provide a responsive environment, Sun Ray enterprise server systems must be equipped with sufficient main memory to support intended applications. Applications which run out of memory can cause the system to "thrash" — constantly moving pages between memory and disk. When systems become I/O bound in this fashion, disk speed becomes the limiting factor and performance is significantly degraded. To avoid I/O bound applications, Sun Ray enterprise server systems are typically equipped with large amounts of system memory. Large-memory Sun Ray enterprise server systems are cost effective since the low cost of RAM is aggregated across all the appliances served by that system.

To understand the memory impact of an application on a Sun Ray enterprise server, the Resident Set Size (RSS) of the application must be considered. The RSS, or “memory footprint”, is defined as the amount of physical memory an application needs in order to avoid being paged to disk. On a Sun Ray enterprise server, the RSS of the X11 server must also be considered since one X11 server process runs for each active enterprise appliance user. Additionally, applications cause the RSS of the X11 server to grow by an amount which is dependent on the application. This growth in RSS in the X11 server is not cumulative with additional applications — The X11 server RSS grows and shrinks as users switch applications.

Sun systems and the Solaris Operating Environment are designed to enable extremely large memory configurations (up to 64 GB) allowing them to potentially support thousands of Sun Ray 1 enterprise appliance users. Further, the Solaris Operating Environment features an efficient shared memory system which can reduce the effective RSS for multiple instances of an application. Multiple instances of applications that are written to use shared memory can share memory segments (particularly code segments) providing a very positive effect on multi-user scalability. The cost for a shared page of memory is only incurred once for all instances of an application, whereas applications’ private memory needs scale linearly. The memory impact for applications that do not use shared memory also scales linearly with each additional instance of the application.

Configuring Sun Ray enterprise server memory should be based on an understanding of the peak-load needs of the application environment including not only application use, but an indication of how many users are likely to be active at one time. The total number of instances of the application that might be run concurrently must be taken into account along with both the private and shared memory footprints of the applications.

Figure 3-1 illustrates the relative amounts of private and shared memory which are needed by the sample applications. Memory resources (both private and shared) needed by the X11 server are also shown

.

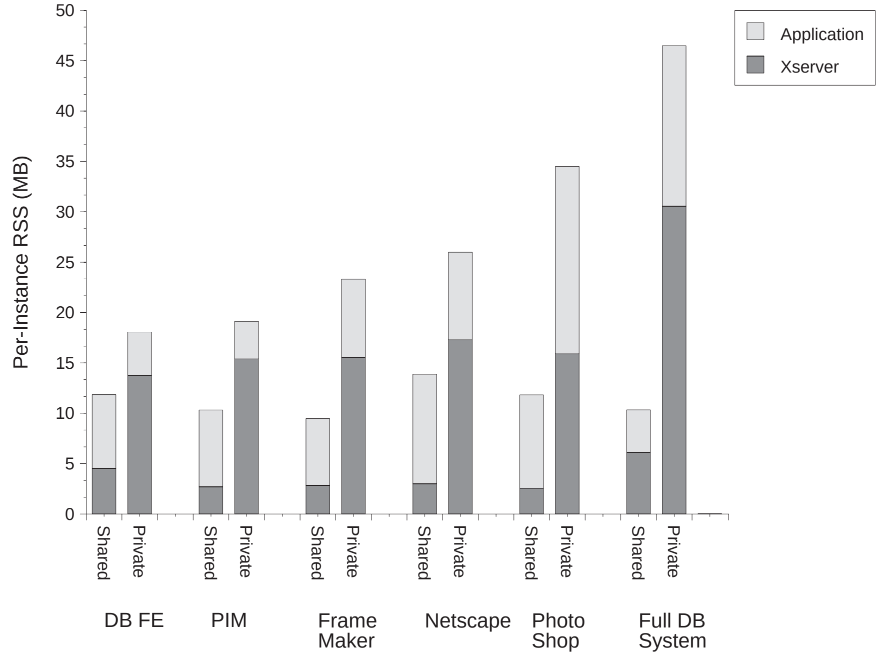

Figure 3-1 Per-instance Resident Set Size (RSS) for sample applications

Sun Ray enterprise server CPU Requirements

Quantifying processor scalability in the Sun Ray enterprise system requires understanding the computational demand placed on Sun Ray enterprise server systems. In the studies that Sun conducted, a single system was used to run the Sun Ray enterprise server software, the GUI (X11 server, CDE), and the sample applications. A single server was used to allow engineers to study the effect that CPU loading had on responsiveness and interactivity for users of enterprise appliances. In real-world deployments, applications might be provided by additional servers as needs dictated to support a given user base.

Though users will occasionally forgive substandard performance suffered by an individual application, poor system responsiveness (screen refresh, mouse movements, key clicks) is a true limiting factor for productivity and a considerable source of aggravation. Like memory, server systems' processing capability should be sized to ensure responsiveness during peak loading situations and should provide capacity for inevitable growth.

To gauge how increasing numbers of users and their applications impact responsiveness, Sun had users run their applications on Sun Ray 1 enterprise appliances while the Sun Ray enterprise server system was subjected to increasing load. Through testing, a strong correlation was found between the size of the run-queue (the list of jobs waiting for each individual CPU) and the users' perceived quality of performance. Users were asked to rate performance and responsiveness for various run-queue lengths. Though responses varied widely, in general responsiveness was considered "annoying" once the run queue depth reached forty five jobs and was considered "unusable" once the run queue depth exceeded fifty five. Engineers then used these run-queue "high water-marks" to measure different system's effectiveness in serving multiple users running the sample applications.

To determine maximum loading conditions where the system still provided good interactive response, a load generator was used along with the individual sample application profiles to emulate increasing numbers of simultaneous users. The run queues on several server systems were monitored as increasing numbers of individual application profiles were played back. Server systems with 1, 2, 4, and 8 300 Mhz UltraSPARC II processors were evaluated. Table 3-1 lists the configurations of the servers used as Sun Ray enterprise servers in the scalability testing.

System	Number of 300 Mhz Processors	System Memory (Megabytes)	Swap Space (Megabytes)
Sun Enterprise 2	1	512	1141
Sun Enterprise 2	2	512	1141
Sun Enterprise 450	4	4096	6699
Sun Enterprise 450	8	4096	6699

Table 3-1 Tested Sun Ray enterprise server, memory, and swap configurations

Figure 3-2 illustrates the results from a single-processor system and demonstrates the wide range of supportable users depending on the applications in use.

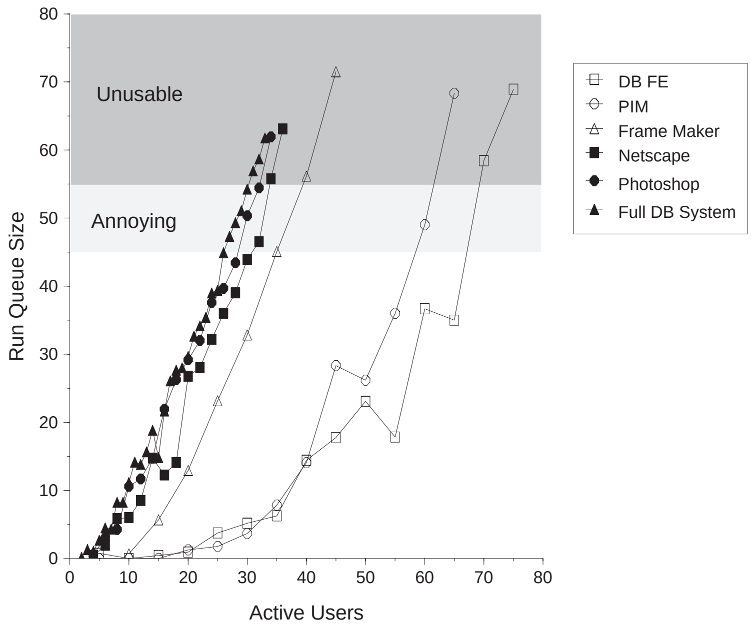

Figure 3-2 Individual application scaling for a single processor system

It should be noted that these experiments modeled multiple users all running the same application and do not approximate a mixed application environment. Though much can be inferred from these results, work is in progress to explore loading and server sizing in mixed environments.

Figure 3-3 illustrates the complete results of the testing and demonstrates that Sun's symmetric multiprocessing servers scale gracefully to serve increasing numbers of enterprise appliance users.

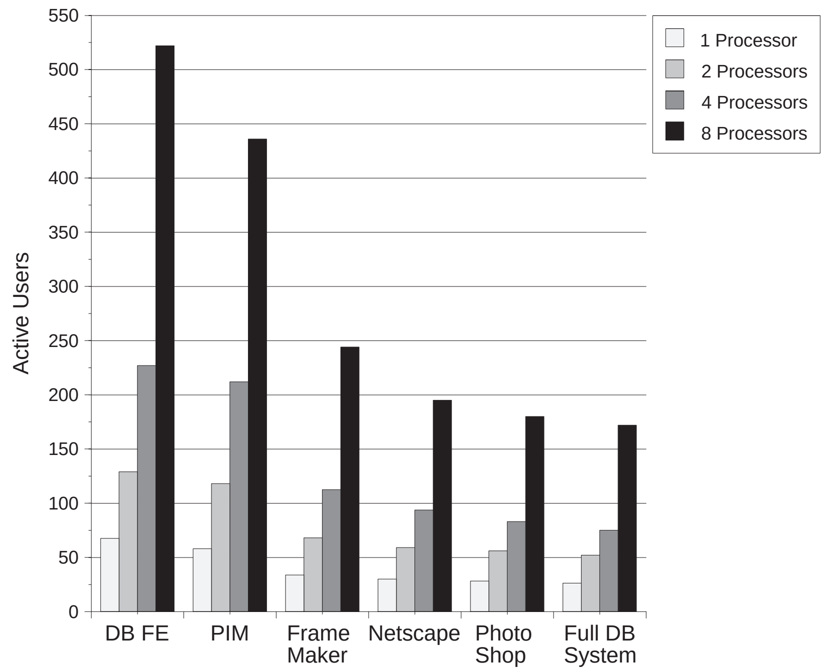

Figure 3-3 Scalability of simultaneous active users on various systems

Application Impact on the Dedicated Interconnect

A low-latency interconnect is essential to deliver performance and responsiveness to the enterprise appliance which literally relies on the interconnect for each pixel and keystroke. To provide low latency, the interconnect should be an entirely dedicated interconnect and not considered as a part of a broader-purpose network. High level network protocols like NIS, NFS, LDAP, or SMTP should not be run on the dedicated interconnect — routing of the Hot Desk technology protocol is also not supported.

By using standard Fast Ethernet and Gigabit Ethernet technology for the dedicated interconnect, organizations can leverage standard low-cost, unmanaged switches and can often utilize existing building wiring to deploy

Sun Ray 1 enterprise appliances. This approach gives customers the reliability and cost benefits of commodity off-the-shelf network components along with Solaris' high-performance low-level networking support.

Because the interconnect is strictly dedicated to traffic between Sun Ray enterprise servers and the desktop appliances, and not used like a real network, full-scale network management is unnecessary. Customers save considerably on administration. By deploying Hot Desk technology and replacing existing fully-networked desktop systems with Sun Ray 1 enterprise appliances, organizations can dramatically reduce the number of true network nodes they support. In fact, the Sun Ray enterprise system can significantly reduce the amount of network address management, system administration, troubleshooting and general maintenance that is required to support a given number of desktop users.

Though standard low-level networking protocols are used, the dedicated interconnect presents very different traffic patterns than a conventional network. Sun Ray enterprise servers use a very efficient protocol to send 2D pixels to the appliances only when they need to be updated on the screen. This approach provides a predictable impact on the amount of interconnect bandwidth that is required by different applications. Those applications which are visually dynamic (video and animation, for example) require more bandwidth to display because they tend to update larger numbers of pixels more frequently.

To preserve interactive end-user responsiveness, saturation of the interconnect must be avoided. In order to verify the scalability of the interconnect and to understand the impact of different applications, Sun once again used the profiles gathered from users running applications in the Sun Ray enterprise system. The profiles were used in conjunction with BEST/1 resource planning software from BMC software which uses queueing theory to predict load. Given the different application profiles, the BEST/1 software was used to predict saturation points for multiple instances of each application and for various classes of network technology.

Figure 3-4 provides the results of the saturation modeling for both 10 Mbps Ethernet and 100 Mbps Fast Ethernet in both half duplex (HDX) and full duplex (FDX) configurations. On hubs and switches (and interfaces) that support FDX, devices at each end of the link can negotiate to send and receive data simultaneously — providing for additional bandwidth. FDX support assumes that each end of the link connects to a single device.

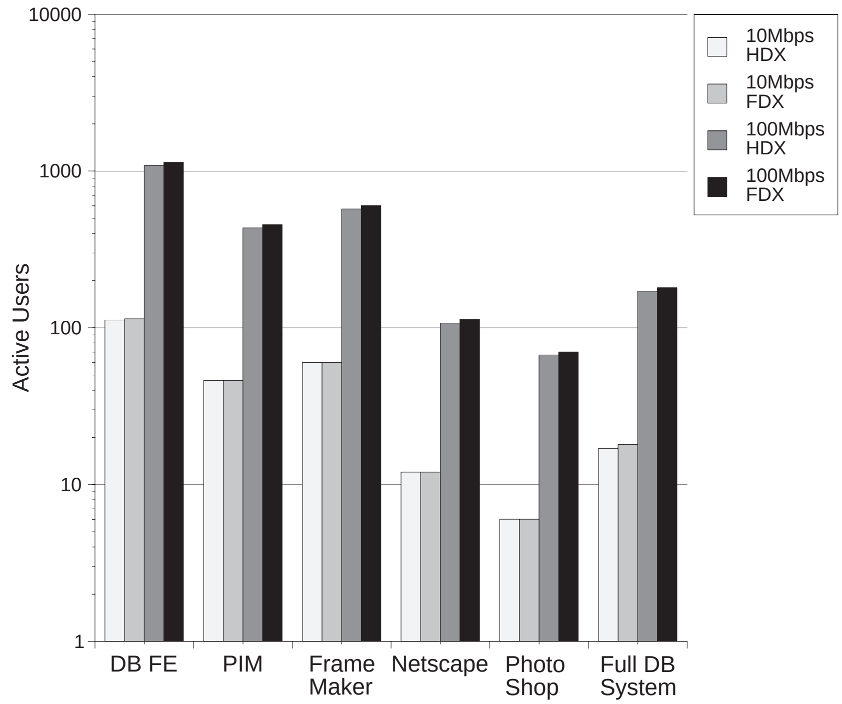

Figure 3-4 Estimated network saturation points from sample application profiles

Again, it is worth noting that Figure 3-4 illustrates interconnect saturation for multiple copies of individual applications and does not represent a mixed-application environment. Also as this data represents simultaneously active users, real-world situations should be able to support considerably larger numbers of users. These results demonstrate that 100 Mbps technology scales well enough to support a significant number of users in the Sun Ray enterprise system.

Typical Configurations

While detailed site planning is beyond the scope of this document, some example scenarios are presented here to illustrate typical configurations. Work is ongoing to provide true capacity planning for Sun Ray enterprise server systems based on additional testing and real-world application usage. More information on interconnect design and server sizing can be found respectively in the Sun white papers *Deploying Hot Desk Technology* and *Sizing Sun Ray Enterprise Servers*.

General Guidelines

Ultimately, Sun Ray enterprise server requirements are entirely dependent on user activity levels, applications deployed to the desktop appliances, and other roles performed by the server. In general, an Ultra 10S server is the smallest system recommended by Sun to serve as a Sun Ray enterprise server. A multiprocessor server is recommended for all but the smallest workgroups (15 or fewer simultaneously active users). It is also a good idea to over-configure Sun Ray enterprise servers by one CPU and a memory increment equivalent to a memory SIMM. Over-configuring in this fashion enables a system to maintain its expected performance level in the event of component (CPU or memory SIMM) failure since SPARC / Solaris systems have the ability to map out a failed component upon reboot.

With Sun's scalable server product line, memory and CPUs can be steadily added as demand increases until a server's maximum configuration is reached. Maximally configured servers can be upgraded to a larger binary compatible model if necessary or multiple servers can be configured to add application performance and higher availability.

As discussed, network scalability is also important to the performance of enterprise appliances. Switched Ethernet is ideal for the interconnect because additional users can be added without incurring network conflicts or congestion. Additional Fast Ethernet or Gigabit Ethernet ports can be added to Sun Ray enterprise servers and additional switches can be added to raise the total interconnect bandwidth.

Small Workgroup Configuration

Small workgroups consisting of approximately 5 to 50 Sun Ray 1 enterprise appliances can be satisfactorily served by using a a low-cost Fast Ethernet (100BaseT) switched network as the dedicated interconnect. Using commodity layer-2 switches keeps both purchase and administration costs low and gives each Sun Ray 1 enterprise appliance its own dedicated 100 Mbps of bandwidth.

Figure 3-5 shows an entry-level Sun Enterprise Ultra™ 10S server with two Fast Ethernet ports dedicated to the interconnect while an additional Fast Ethernet port connects the Sun Ray enterprise server to the LAN. Each of the 100 Mbps interconnect links are connected to one of the 100BaseT switches — each handling the traffic from six desktop appliances. Because user traffic is generally bursty in nature and not all users are typically active at any one time, this configuration should supply ample bandwidth for most applications.

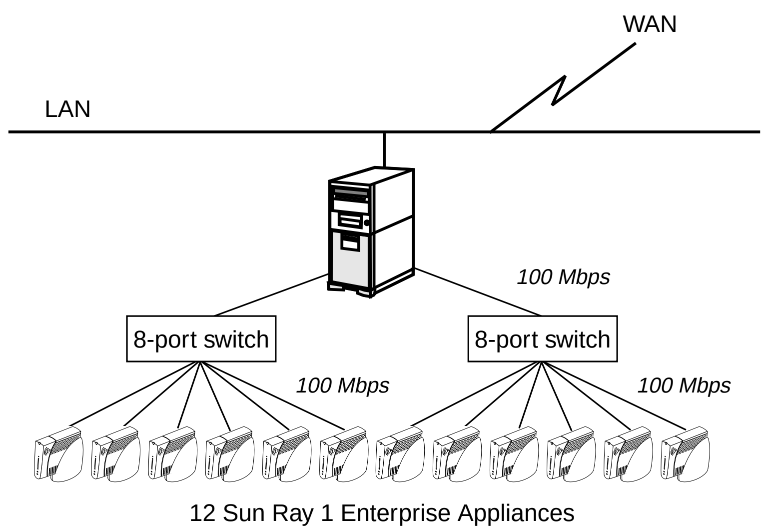

Figure 3-5 Inexpensive 100BaseT switches are used as the interconnect for small workgroup configurations

The Enterprise 10S server pictured is configured with one 270 MHz CPU, 1 GB of memory, a single 1 GB disk, and two Fast Ethernet interfaces in addition to the system's primary Fast Ethernet interface.

Departmental Configuration

For departmental groups consisting of approximately 100 to 1000 desktops, a large-scale Gigabit Ethernet/Fast Ethernet switched network is more suitable for the interconnect. In this configuration, a Sun Ray enterprise server with multiple Gigabit Ethernet cards connects to a set of Gigabit/ Fast Ethernet switches, which in turn connect to the Sun Ray 1 enterprise appliances (Figure 3-6). The switches in this example must be able to perform layer-2 switching, and must include IGMP multicast snooping to reduce the flow of multicast traffic to unwanted ports.

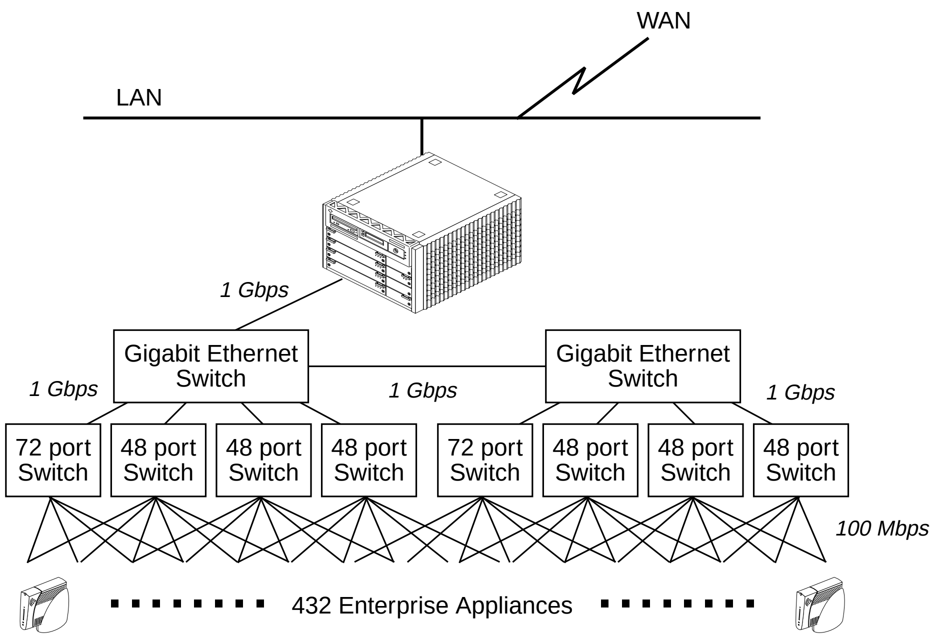

Figure 3-6 Larger, departmental configurations utilize a switched Gigabit Ethernet/Fast Ethernet interconnect.

The Enterprise 4500 depicted in Figure 3-6 is assumed to have eight 333 MHz CPUs, 8 GB of system memory, and seven Gigabit Ethernet interfaces enabling support of up to 500 Sun Ray enterprise appliance.

References

Sun Microsystems Computer Company posts product information in the form of data sheets, specifications, and white papers on its Internet World Wide Web Home page at: *http://www.sun.com*.

Look for the these and other Sun technology white papers:

Sun Ray 1 Enterprise Appliance Overview and Technical Brief, White Paper, Sun Microsystems.

Deploying the Sun Ray Hot Desk Architecture, White Paper, Sun Microsystems.

Sizing Sun Ray Enterprise Servers, White Paper, Sun Microsystems.

Integrating Sun Ray 1 Enterprise Appliances and Microsoft Windows NT, White Paper, Sun Microsystems.

Sun Microsystems Incorporated
901 San Antonio Road
Palo Alto, CA 94303 USA
650 960-1300
FAX 650 969-9131
http://www.sun.com

Sales Offices

Africa (North, West and Central):
+33 1 30674680
Argentina: +54-11-4317-5600
Australia: +61-2-9844-5000
Austria: +43-1-60563-0
Belgium: +32-2-716 79 11
Brazil: +55-11-5181-8988
Canada: +905-477-6745
Chile: +56-2-3724500
Colombia: +571-629-2323
Commonwealth of Independent States:
+7-502-935-8411
Czech Republic: +420-2-33 00 93 11
Denmark: +45 4556 5000
Estonia: +372-6-308-900
Finland: +358-9-525-561
France: +33-01-30-67-50-00
Germany: +49-89-46008-0
Greece: +30-1-6188111
Hungary: +36-1-202-4415
Iceland: +354-563-3010
India: +91-80-5599595
Ireland: +353-1-8055-666
Israel: +972-9-9513465
Italy: +39-039-60551
Japan: +81-3-5717-5000
Kazakhstan: +7-3272-466774
Korea: +822-3469-0114
Latvia: +371-750-3700
Lithuania: +370-729-8468
Luxembourg: +352-49 11 33 1
Malaysia: +603-264-9988
Mexico: +52-5-258-6100
The Netherlands: +31-33-450-1234
New Zealand: +64-4-499-2344
Norway: +47-2202-3900
People's Republic of China:
Beijing: +86-10-6803-5588
Chengdu: +86-28-619-9333
Guangzhou: +86-20-8777-9913
Shanghai: +86-21-6466-1228
Hong Kong: +852-2802-4188
Poland: +48-22-8747800
Portugal: +351-1-412-7710
Russia: +7-502-935-8411
Singapore: +65-438-1888
Slovak Republic: +421-7-522 94 85
South Africa: +2711-805-4305
Spain: +34-91-596-9900
Sweden: +46-8-623-90-00
Switzerland: +41-1-825-7111
Taiwan: +886-2-2514-0567
Thailand: +662-636-1555
Turkey: +90-212-236 3300
United Arab Emirates: +971-4-366-333
United Kingdom: +44-1-276-20444
United States: +1-800-555-9SUN OR +1-650-960-1300
Venezuela: +58-2-905-3800
Worldwide Headquarters:
650-960-1300 or 800-555-9SUN
Internet: www.sun.com

Printed in USA