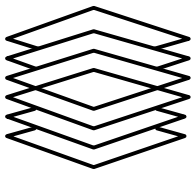
Solaris Operating Environment Use on PRIMEPOWER Provides Multiple IT Environment Benefits

September 2002



A D.H. Brown Associates, Inc. White Paper Prepared for

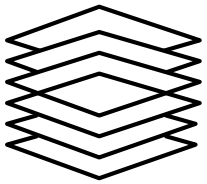
Fujitsu

This document is copyrighted © by D.H. Brown Associates, Inc. (DHBA) and is protected by U.S. and international copyright laws and conventions. This document may not be copied, reproduced, stored in a retrieval system, transmitted in any form, posted on a public or private website or bulletin board, or sublicensed to a third party without the written consent of DHBA. No copyright may be obscured or removed from the paper. D.H. Brown Associates, Inc. and DHBA are trademarks of D.H. Brown Associates, Inc. All trademarks and registered marks of products and companies referred to in this paper are protected.

This document was developed on the basis of information and sources believed to be reliable. This document is to be used "as is." DHBA makes no guarantees or representations regarding, and shall have no liability for the accuracy of, data, subject matter, quality, or timeliness of the content. The data contained in this document are subject to change. DHBA accepts no responsibility to inform the reader of changes in the data. In addition, DHBA may change its view of the products, services, and companies described in this document.

DHBA accepts no responsibility for decisions made on the basis of information contained herein, nor from the reader's attempts to duplicate performance results or other outcomes. Nor can the paper be used to predict future values or performance levels. This document may not be used to create an endorsement for products and services discussed in the paper or for other products and services offered by the vendors discussed.

TABLE OF CONTENTS



Solaris Operating Environment Use on PRIMEPOWER Provides Multiple IT Environment Benefits

Fujitsu's use of the SPARC processor in the PRIMEPOWER server allows Fujitsu to take advantage of the leading edge Solaris Operating Environment (OE) to support its advanced scalability and reliability functions. Solaris OE benefits from exceptionally strong functional capabilities, and now resides at the forefront of providing high-end scalability and RAS (reliability, availability, and serviceability) capabilities for UNIX systems.

The use of Solaris OE on PRIMEPOWER servers offers a variety of other benefits, including the strong market-share and mind-share leadership that Solaris OE has established over the years. For example, Solaris OE now enjoys a very strong base of available software, with more than 12,000 applications running on the platform.

Solaris OE also boasts a pool of worldwide administrative expertise matched by very few alternative system software environments. When complemented with key value-added functions from Fujitsu, Solaris OE provides an extraordinarily rich system software infrastructure that matches the powerful underlying capabilities in PRIMEPOWER hardware.

This white paper provides an overview of the Solaris OE's capabilities. It is one in a series of seven PRIMEPOWER white papers that provide an overview of PRIMEPOWER, information on PRIMECLUSTER, the ARMTech ShareEnterprise resource management software, the PRIMEPOWER system architecture, the PRIMEPOWER SPARC64 V microprocessor, and PRIMEPOWER system management.

UNIX IN THE DATACENTER

The most demanding commercial computing applications were historically reserved for proprietary mainframe systems. However, during the past few years, UNIX systems have matured sufficiently to be able to meet the functional requirements of the highest-end environments. Although most of the time IT managers cannot simply drop UNIX systems into traditional datacenters as a replacement for mainframes, the gap between their functional capabilities and those of mainframes has narrowed sufficiently so that they can be realistically considered for handling similar classes of workloads.

At the same time, the functional requirements for enterprise-class systems have evolved to reflect changes in the business climate. In today's environment, managers increasingly focus on tactics that control costs, such as server consolidation. Simultaneously, in spite of readjusted expectations for

Copyright 2002 © D.H. Brown Associates, Inc.

opportunities on the web, high-end web applications in business-critical roles continue to demand very high levels of scalability and uptime.

As a result, systems targeting modern datacenters require a flexible portfolio of functions that can adapt to a variety of scalability and reliability conditions rather than focusing on the monolithic, high-throughput workloads required by batch processing on mainframes. These conditions may include,

- Large shared memory processing (SMP) ranges for supporting applications that scale best by sharing a single address space across multiple processors. Because SMP incurs fewer penalties related to management and processing overhead than other multiprocessing techniques such as clustering, and is also relatively easy for application developers to exploit in their code, it remains one of the most effective ways to increase system performance for many key business applications. These include database servers and online transaction processing (OLTP) functions. Today, sufficiently tuned SMP systems are capable of supporting more than one hundred processors in SMP configurations.
- Partitioning support, which allows administrators to run multiple instances of an operating system within a single server. Each instance behaves as if it were running on a stand-alone machine, and reliable barriers between the different environments maintain overall system robustness, so that even the most extreme application failure or operating system crash in one partition leaves the others unaffected. The entire environment, i.e., all partitions, can be managed from a single point. This allows administrators to deploy multiple discrete system images within a single server for purposes of server consolidation or for supporting applications that can scale horizontally.
- Resource management tools, which work within a single operating system instance to effectively manage massive, constantly changing workloads, allow multiple dominant applications to coexist in a single environment. The tools work by efficiently allocating system resources such as CPU, memory, and I/O to different applications with flexible scheduling policies. These resource management functions effectively override the operations of the default UNIX scheduler, taking into consideration customized policies instead.
- Online system maintenance capabilities, which minimize planned downtime by reducing the number of administrative tasks that require a system restart. As IT infrastructures become increasingly web-based and globally oriented, servers must truly be able to respond to requests 24 hours a day and 365 days a year. Also, reliability becomes an increasingly important concern when users take advantage of partitioning capabilities to perform server consolidation, since the system risks becoming a single point of failure that affects more workloads.
- System management tools that are user-friendly, yet powerful, and provide administrators with precise control over system functions. The earliest UNIX implementations had notoriously poor tools for system management. Traditional UNIX system management procedures required an experienced

 Copyright 2002 © D.H. Brown Associates, Inc.

administrator to hand-edit a large and dispersed set of cryptic configuration files stored in the */etc* directory – a crude and error-prone process. However, as UNIX systems take on ever more roles in enterprise and datacenter environments, the need has grown for graphical user interface (GUI) tools that allow less-seasoned administrators to employ the easier-to-learn approach of "recognize and point" rather than "remember and type." Also, remote manageability is becoming increasingly important to manage multiple servers in geographically distributed organizations and large enterprises, as well as to enable easy access to multiple discrete system images in server consolidation projects.

FUJITSU AND SOLARIS OE

Complete support for advanced scalability and reliability functions such as high-end SMP ranges, server partitioning, and failure resiliency requires not only advanced hardware functions, but also a variety of complementary functions in operating system software. In addition, the trend in system software packaging has been to offer an increasingly broad set of capabilities in the base operating system that might have required separate purchase and installation in the past.

This trend has resulted in the availability of complete "operating environments" that go far beyond the basic functions needed to run hardware. These functions were traditionally provided by UNIX systems, and minimize the burden for users to acquire and install the necessary components for various layered software functions.

Fujitsu's use of the SPARC processor in the PRIMEPOWER server allows it to meet both the requirements for advanced scalability and reliability software support, as well as rich integrated functions, by using Solaris OE. Solaris OE was designed from the ground up for the SPARC architecture, and its design resided at the forefront of the charge to drive mainframe-grade scalability and RAS functions into UNIX systems. Recent versions of the system such as Solaris OE 8 substantially step up the level of integration of layered services as well, resulting in a complete operating environment that provides the necessary functions for deploying a variety of services.

The use of Solaris OE on PRIMEPOWER servers offers other benefits, including the strong market-share and mind-share leadership that Solaris OE has established over the years. Solaris OE and its predecessor operating systems originally made UNIX fashionable. Indeed, UNIX remains synonymous with Solaris OE in the minds of many users.

As a result, Solaris OE now enjoys a very strong base of available software, with more than 12,000 applications running on the platform. Because of the broad range of binary-compatible platforms that Solaris OE and SPARC support, the hurdles to developing Solaris OE applications are relatively low, compared with traditional mainframes and other UNIX platforms.

Copyright 2002 © D.H. Brown Associates, Inc.

Independent software vendors (ISVs) and end users can begin development on sub-$1,000 SPARC workstations and servers, and then scale those same applications to high-end mainframe-class systems. The availability of a Solaris OE version for the Intel platform also contributes to the flow of applications, since moving applications from PC desktops, laptops, and low-end servers to high-end SPARC systems is a straightforward process.

For the same reasons, Solaris OE enjoys a strong mind share with system administrators. Again, many administrators first learned UNIX on Solaris OE systems, and they have evolved their skills as new high-end functions were introduced into Solaris OE. As a result, Solaris OE boasts a pool of worldwide expertise that is matched by very few alternative system software environments.

Finally, much of the Internet was built with UNIX systems such as Solaris OE. Solaris OE's strong affinity with key web protocols and technologies, coupled with its rapidly evolving scalability and reliability capabilities, enabled Solaris OE to establish a strong position on back-end servers and other parts of critical infrastructure underlying today's web.

Furthermore, as web browsers become the primary entry point for a growing number of day-to-day computing activities, application developers have increasingly begun to explore possibilities for segmenting application designs along web boundaries, i.e., shifting application logic from clients to web servers and implementing user interfaces with HTML-based presentation layers. The incumbency of Solaris OE in terms of hosting basic web infrastructure extends to a solid position with developers of web applications, as shown by its strong support for virtually all of the key Java-based web application server platforms.

SOLARIS OE TECHNICAL LEADERSHIP

Solaris OE benefits from exceptionally strong functional capabilities, delivering a considerable advantage for Fujitsu PRIMEPOWER systems. After becoming established as the *de-facto* UNIX standard on high-performance workstations during the 1980s and early 1990s, Solaris OE developers increasingly adopted functions that optimized it for use on high-end servers. They copied as many scalability and enterprise RAS features as possible from the mainframe into the Solaris OE design.

As a result, Solaris OE now resides at the forefront of providing high-end scalability and RAS capabilities for UNIX systems. When complemented with key value-added functions from Fujitsu, Solaris OE provides an extraordinarily rich system software infrastructure that matches the powerful underlying capabilities in PRIMEPOWER hardware. What follows are several key scalability, RAS, and manageability highlights in Fujitsu's version of Solaris OE 8, the operating environment currently supported on Fujitsu PRIMEPOWER servers.

Copyright 2002 © D.H. Brown Associates, Inc.

SCALABILITY

SMP boosts system performance by harnessing multiple processors within a single server in which all share the same processors, memory, and I/O resources. SMP incurs fewer penalties related to management and processing overhead than other multiprocessing techniques, such as clustering, and is also relatively easy for application developers to exploit in their code. For this reason, it remains one of the most effective ways to increase system performance for many key business applications, including database servers and OLTP.

Enabling a kernel to effectively manage large numbers of processors has traditionally presented an extraordinary and tedious challenge for operating system developers. Typically, hardware and software design teams must cooperate closely and consider a variety of factors throughout the system, including memory-bus bandwidth and cache sizes, to optimize their implementation.

Moreover, they must gather huge amounts of empirical data to determine how well their design works with a variety of applications, a process that can last for years before good results are achieved. Solaris OE has been undergoing this process for years, resulting in an environment that is highly tuned for large SMP servers. One of the benefits of running Solaris OE on Fujitsu's high-end SMP servers is that applications can take full advantage of these optimizations.

While there is no generally applicable way to predict exactly how well an application will scale on an SMP system under real-world conditions, certain benchmark tests provide a reasonable approximation. In particular, the Transaction Processing Performance Council's TPC-C benchmark remains the most widely accepted method to assess the SMP range of server systems.

In addition to being truly vendor-neutral and subject to rigorous auditing procedures, the TPC-C test stresses a number of system components that are frequently exercised in commercial server applications, including I/O, inter-processor communication, and cache management. Fujitsu's PRIMEPOWER has not only been tested with more processors than any other SMP server (at 128 CPUs), but it also holds the world record for TPC-C results achieved on a single server.[1]

RELIABILITY, AVAILABILITY, AND SERVICEABILITY (RAS)

Solaris OE 8 offers a very strong set of RAS features, including several that are not yet available from any competitors and that can dramatically reduce downtime for planned or unplanned reasons. For example, Solaris OE 8 can arbitrarily add and remove processors online, and is the only UNIX system that can add and remove memory online. The hardware partitions function in Solaris OE is the only available UNIX partitioning tool that allows the partitions between operating system instances to be adjusted while keeping them online. All

[1] For further information, see the TPC website at www.tpc.org.

Copyright 2002 © D.H. Brown Associates, Inc.

of the other UNIX hardware partition solutions currently require the affected partitions to be rebooted after reconfiguration.

On Fujitsu's servers, Solaris OE hardware partitioning support attains an advantage even compared to Dynamic Domains on Sun's own Sun Fire servers with PRIMEPOWER new series Extended Partitioning (XPAR) capabilities. With XPARs, it is possible to achieve partition granularity down to one CPU and one GB of memory[2] (a part of a system board). By contrast, Sun's Dynamic Domains have a partition granularity of a minimum of four CPUs.

For managing multiple workloads that do not require "bullet-proof" isolation between instances, resource management tools provide a critical complement to hardware partitions. Resource management tools work within a single operating system instance to effectively manage massive, constantly changing workloads so that multiple dominant applications can coexist in a single environment. The tools work by efficiently allocating system resources such as CPU, memory, and I/O to different applications with flexible scheduling policies.

These resource management functions effectively override the operations of the default UNIX scheduler, taking into consideration customized policies instead. Although UNIX has always offered simple resource management tools such as disk quotas and the *nice* command to set application priorities, UNIX system developers have recently invested in far more sophisticated capabilities, modeled in part after long-standing mainframe functions.

On PRIMEPOWER servers, Fujitsu provides the ARMTech ShareEnterprise resource management software to provision CPU resources to processes in an XPAR. ARMTech guarantees processing stability through its ability to provide CPU "shares" or "reservations."

Shares are designed to improve resource use when the system's resources are contested. By allocating shares, a resource consumer gains entitlement to a portion of the sum of the shares held by all contesting resource consumers. In other words, a share guarantees CPU or memory based on a relative ratio to the total system CPU or memory capability. By contrast, a reservation provides a minimum resource guarantee when there is a resource conflict. This means that a reservation is an absolute percentage of a resource and is not demand-dynamic. Reservations are commonly used to ensure a proper level of resources to address mission-critical tasks.

On PRIMEPOWER servers, Fujitsu's SPARC64 V processor supports a function called Error ID injection that enables the operating system to detect error conditions such as defective CPU, memory, or I/O modules, or multiple

[2] XPAR can achieve single processor partitions on the PRIMEPOWER 900 and 1500 models. The minimum PRIMEPOWER 2500 XPAR configuration contains two processors. For a fuller explanation, please refer to the companion white paper *PRIMEPOWER Server Architecture Excels in Scalability and Flexibility.*

 Copyright 2002 © D.H. Brown Associates, Inc.

Error Correcting Code (ECC) errors in memory. The processor can identify a defective module by CPU ID, memory slot number, or PCI slot number.

The SPARC64 V processor also offers register parity so that the system can detect register corruption or other unrecoverable errors that may severely affect the system continuity. Solaris OE on PRIMEPOWER has the ability to respond to these errors and continue running the system until a "panic" procedure can shut it down in a safe and orderly manner.

SYSTEM MANAGEMENT

As enterprises depend ever more on networks, the IT infrastructure becomes more distributed, dramatically increasing the number of servers that need to be deployed. Large enterprises routinely disperse servers geographically, in some cases across different continents and time zones.

Server farms and blade architectures involve deploying large numbers of servers in parallel, each of which has to be managed over the network. Thus, it becomes increasingly important to have the capability to effectively manage operating systems remotely from a single point of control. If an enterprise depends on a thousand servers, for example, it is simply not feasible to maintain a thousand system administrators locally.

Strong remote manageability is also a key requirement for server consolidation because it minimizes user dependence on the IT organization for intervention, even after a server has been physically moved or otherwise redeployed. Web-based system management tools allow users to maintain their servers from anywhere, using ordinary web browsers as entry-points.

Fujitsu offers a GUI-based system management facility for its PRIMEPOWER servers called WebSysAdmin, which can be used to manage Solaris OE-based servers remotely over the web and enterprise networks from either Windows or Solaris OE clients. Using SNMP-based agents, WebSysAdmin allows connections to Computer Associates' Unicenter, IBM's Tivoli, and Fujitsu's SystemWalker enterprise system management frameworks and applications.

WebSysAdmin is capable of managing multiple PRIMEPOWER Solaris OE-based systems from a single point, and allows the monitoring and configuration of a variety of Solaris OE and PRIMEPOWER system parameters including,

- hierarchical domains and management of nodes and their interfaces;
- hardware status (e.g., processors, memory) and configuration allocation;
- system processes (with associated logs for individual or all domains);
- user status (e.g., user definitions, group definitions, and passwords);
- software management (e.g., packages and versions);
- task management (e.g., task definition);

Copyright 2002 © D.H. Brown Associates, Inc.

- SNMP management (e.g., monitors and controls what is passed to Unicenter or Tivoli); and
- performance status (e.g., display of attributes of system performance).

REVIEW OF KEY POINTS

This white paper provides an overview of Fujitsu's PRIMEPOWER Solaris OE. This describes how it contributes to PRIMEPOWER's ability to meet the demands of the business-critical IT environment for an operating environment that offers industry-leading scalability, RAS, system management, and other capabilities. The Solaris OE is yet another part of the PRIMEPOWER design and architecture infrastructure that is the basis for PRIMEPOWER's ability to deliver a high quality of service to the IT environment.

As a result, the Solaris OE and the other technologies discussed in the other six white papers in this series make it clear that PRIMEPOWER is a short-list candidate for single SMP server or clustered SMP server operation in the business critical IT infrastructure. (Additional information concerning PRIMEPOWER's Solaris OE operating environment and other PRIMEPOWER advantages can be found on the Fujitsu websites.)

 Copyright 2002 © D.H. Brown Associates, Inc.