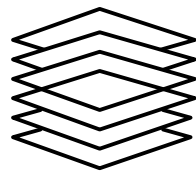


*Fujitsu PRIMEPOWER  
Server Family Offers  
Industry-Leading Capabilities*

*September 2002*



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

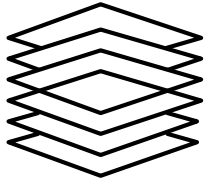
This document is copyrighted © by D.H. Brown Associates, Inc. (DHBA) and is protected by U.S. and international copyright laws and conventions. This document may not be copied, reproduced, stored in a retrieval system, transmitted in any form, posted on a public or private website or bulletin board, 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

<b>PRIMEPOWER NEW SERIES REVEALED</b> .....	<b>1</b>
<i>Table 1: PRIMEPOWER New Series Models Compared</i> .....	2
<i>Table 2: New Series Models 900, 1500, 2500 High-Availability Features Compared</i> .....	3
<b>OPEN PRIMEPOWER KEY COMPONENTS</b> .....	<b>3</b>
INDUSTRY-LEADING SOLARIS.....	3
A NEW SPARC.....	4
<b>PRIMEPOWER SYSTEM ARCHITECTURE</b> .....	<b>5</b>
<b>PRIMEPOWERS COMBINE IN A PRIMECLUSTER</b> .....	<b>7</b>
<i>Figure 1: PRIMECLUSTER Modular Software Architecture</i> .....	7
<i>Table 3: PRIMECLUSTER Module Requirements</i> .....	8
SCALABILITY SERVER.....	8
PARALLEL SERVER .....	8
HA SERVER .....	8
ENTERPRISE EDITION.....	9
<b>PRIMEPOWER SYSTEM MANAGEMENT</b> .....	<b>9</b>
RESOURCE MANAGEMENT.....	9
HIGH-AVAILABILITY MANAGEMENT .....	10
SECURITY MANAGEMENT.....	11
<b>REVIEW OF KEY POINTS</b> .....	<b>12</b>



## *Fujitsu PRIMEPOWER Server Family Offers Industry-Leading Capabilities*

Fujitsu's PRIMEPOWER new series is a state-of-the-art server family suitable for stand-alone server or cluster use. Based on the industry-standard Solaris Operating Environment and the Fujitsu SPARC64 V microprocessor as well as many novel system architecture features, the PRIMEPOWER family offers many benefits to CIOs, CTOs, and other IT team members responsible for selecting the optimum infrastructure. PRIMEPOWER's industry-leading capabilities in performance, scalability, manageability, high availability, and other benefits to the infrastructure place it in a strong position. As a result, the end user infrastructure customer will be able to enjoy new levels of service quality.

CIOs and CTOs need to understand these capabilities at a top functionality level so that the IT environment benefits are clear. (It is not necessary for the CIO or CTO to discern all of the details of all the new PRIMEPOWER's capabilities.) This white paper provides a suitable overview of these capabilities. It is one in a series of seven PRIMEPOWER white papers that provide information on PRIMECLUSTER, the Solaris Operating Environment, ARMTech software resource management, the PRIMEPOWER system architecture, the PRIMEPOWER SPARC64 V microprocessor, and PRIMEPOWER system management.

### **PRIMEPOWER NEW SERIES REVEALED**

With headquarters in Tokyo and \$38 billion in revenues for the fiscal year ending March 31, 2002, Fujitsu maintains 170,000 employees in 65 countries. Fujitsu is a leading global OEM of computer hardware and software.<sup>1</sup> Its range of platform products includes the Solaris-based PRIMEPOWER offerings and the Intel-based PRIMERGY offerings. The firm also offers disk arrays and other hardware.

Fujitsu is number four in worldwide UNIX market share and plans to double its UNIX market share in the next three years. Worldwide coverage is provided by Fujitsu in Japan and the APAC (Asia-Pacific) region, Fujitsu Technology Solutions in the United States, and Fujitsu Siemens Computers in the EMEA (Europe-Middle East-Africa) region. Fujitsu has shipped more than 20,000 PRIMEPOWER servers, which have an excellent industry benchmark record.<sup>2</sup>

---

<sup>1</sup> Fujitsu's Annual Report may be obtained through [www.fujitsu.com](http://www.fujitsu.com).

<sup>2</sup> These world-leading benchmarks are a matter of public record and include TPC-C, SAP, BAAN, Oracle, and more. These can be found on the appropriate industry or benchmark organization websites as well as at [www.primepower.fujitsu.com/en/benchmark.html](http://www.primepower.fujitsu.com/en/benchmark.html).

Fujitsu bases its PRIMEPOWER new series on the firm's mainframe heritage<sup>3</sup> to provide superior reliability, availability, and serviceability (RAS) characteristics for both the PRIMEPOWER server and the new SPARC64 V microprocessor. The company also relies on its supercomputer heritage to provide a superior server architecture (e.g., a high-speed crossbar interconnect).

The PRIMEPOWER new series offerings include the PRIMEPOWER 900, 1500, and 2500. These shared memory processing (SMP) machines all run on, and are shipped with, the Solaris operating environment. PRIMEPOWER servers use the 1.3 GHz Fujitsu-designed SPARC64 V microprocessor. This microprocessor complies with the SPARC International V9 Level 2 specification. Table 1 compares the main features of these PRIMEPOWER models.

PRIMEPOWER	Model 900	Model 1500	Model 2500
<b>Server Function</b>	Rack-mount midrange	Enterprise class business	High-end enterprise class business
<b>Maximum Number of System Boards</b>	2	4	16
<b>Number of CPUs</b>	4-16	4-32	8-128
<b>Amount of PC266 DIMM (with ECC) (GB)</b>	64	128	512
<b>Maximum PCI Slots Clock (MHz) (Maximum Number of Slots at Clock)</b>	36 33/66 (24,12)	72 33/66 (48, 24)	320 33/66 (224, 96)
<b>Maximum Ethernet Ports</b>	8	32	80
<b>Maximum Serial Ports</b>	8	32	80
<b>Maximum USCSI Storage Bays</b>	16	32	128
<b>Redundant Power Supply</b>	200 VAC	200 VAC	200 VAC
<b>Redundant Air Cooling</b>	Front-to-back	Bottom-to-back	Bottom-to-top
<b>Hard Partitions with Dynamic Reconfiguration</b>	8	15	15

**TABLE 1:**  
**PRIMEPOWER New Series Models Compared**

Table 2 compares the high-availability features of the three PRIMEPOWER new series models. When a capability is available, the technology it is based on is shown in parentheses (where necessary) and explained in the key below the table. Table 2 shows that PRIMEPOWER provides built-in high availability for its component parts that is second to none in the server industry.

<sup>3</sup> Amdahl's mainframe and Fujitsu Siemens Computers' mainframe are based on Fujitsu's mainframe hardware architecture.

**TABLE 2:**  
*New Series Models 900,  
1500, 2500 High-  
Availability Features  
Compared*

PRIMEPOWER Component		Redundancy	Hot Plug	900	1500	2500
System Board with Solaris Kernel	Processor	Y (DD)	Y (DR)	Y	Y	Y
	Memory	Y (CK)	Y (DR)	Y	Y	Y
	PCI Connection	Y (IOMP)	Y (DR)	Y	Y	Y
	DC-DC Converter	Y	Y (DR)	NA	NA	Y
PCI	PCI Board <sup>4</sup>	Y (IOMP)	Y (DR)	Y	Y	Y
	PCI Card	Y (IOMP)	Y (PHP)	Y	Y	Y
Crossbar-to- Backplane	L2 Crossbar <sup>5</sup>	Y	NA	NP	NA	Y
	Clock Distributor	Y	NA	NP	NA	Y
	DC-DC Converter	Y	NA	NP	Y	Y
Power Supply		Y	Y	Y	Y	Y
Cooling (Fan Tray)		Y	Y	Y	Y	Y
System Control	Control Board	Y	Y	NA	Y	Y
	Remote Cabinet Interface (RCI) <sup>6</sup>	Y	Y	NA	Y	Y
	Control LAN	Y	Y	NA	Y	Y
Network		Y (IOMP)	Y	Y	Y	Y
Storage (FC and SCSI)		Y (IOMP)	Y	Y	Y	Y
DD=Dynamic Degradation, DR=Dynamic Reconfiguration, IOMP=I/O Multipathing, CK=Chipkill-Like Technology, PHP=PCI Hot Plug, NA=Not Applicable, NP=Not Present, Y=Yes						

## OPEN PRIMEPOWER KEY COMPONENTS

The operating system and microprocessor are key components of any mission-critical server and the new PRIMEPOWER series models are no exceptions. Recognizing the need to provide an industry-leading operating system and microprocessor, earlier PRIMEPOWER servers used the best selling Solaris operating system and a SPARC-compliant microprocessor.

### INDUSTRY-LEADING SOLARIS

This Solaris tradition of industry-leading performance continues. Solaris from Sun Microsystems runs PRIMEPOWER new series. The PRIMEPOWER strategy is to use Solaris as a global standard for high independent hardware vendor (IHV) and independent software vendor (ISV) support as well as to allow Fujitsu to develop its own server components for high performance and high availability.

<sup>4</sup> A PCI board is considered part of a system board.

<sup>5</sup> PRIMEPOWER 1500 and 2500 both have a L1 and L2 crossbar. L1 is in the system board crossbar while L2 is in the inter-system board crossbar.

<sup>6</sup> RCI is an interface to manage the power supplies of multiple unites by simple operations. The RCI may be used for clustered systems or Model 2500 cabinets.

Solaris is a well-known, industry-accepted (first in UNIX market share) operating environment and is the subject of a separate white paper in this series. From the IT infrastructure point of view, it is important to remember that:

- Almost every UNIX application of interest to the IT community has been ported to Solaris (over 15,000).
- Solaris is based on, and is compatible with, the open SPARC V9 Level 2 architecture and fully supports 64 bits including a 64-bit virtual address space and 64-bit file systems. It is also compatible with 32-bit applications. For the vast majority of 32-bit applications, no recompilation is necessary.
- Due to Solaris' affinity to SPARC and its scalability, IT applications that run on the PRIMEPOWER new series will also run on other PRIMEPOWER machines, and, in fact, will run on a wide range of servers (from workstations to large SMP machines and clusters) from other Solaris-based server manufacturers.
- In addition to the Solaris media kit, Fujitsu ships additional Solaris-compliant software called the Enhanced Support Facility (ESF). This software covers system administration, automatic power control, and various drivers for PRIMEPOWER hardware such as the System Control Facility (SCF).

It is clear from the first three bullets above that Solaris provides an IT infrastructure with the power needed for today's advanced applications. For the infrastructure using PRIMEPOWER, the additional software mentioned in the fourth bullet, which does not interfere with Solaris in any way, provides still more value.

#### **A NEW SPARC**

A new Fujitsu-designed SPARC microprocessor lies at the heart of the PRIMEPOWER new series servers. It is Fujitsu-designed, which includes for example, mainframe RAS features to ensure high availability and unique internal caches and buses for performance. An overview of the SPARC64 V microprocessor is given in *Sidebar 1: SPARC64 V Microprocessor Revealed*. (This microprocessor is also the subject of a separate white paper in this series.)

While all of this is going on, and to further enhance RAS, hardware-based notification of error occurrences is handled by the previously mentioned SCF. Error information, both fatal and non-fatal, and error history are accumulated by the SCF for further analysis. This information is obtained from the SPARC64 V through its industry-standard, dedicated JTAG<sup>7</sup> interface pins, which in turn pass the error information to the SCF.

---

<sup>7</sup> Joint Technical Advisory Group (JTAG) is a specification for instrumenting chips so that internal chip errors may be monitored, detected, gathered, and distributed outside of the chip through dedicated chip pin connections.

#### Sidebar 1: SPARC64 V Microprocessor Revealed

The latest in a long line of Fujitsu-designed, SPARC International specification-compliant microprocessors, the SPARC64 V brings new power and mainframe high-availability features to the SPARC environment. The microprocessor has some 191 million transistors and is built on 0.13 micron copper technology. It consumes a low 50 Watts at a clock speed of 1.3 GHz. Since such high clock rates usually mean high heat production (up to twice as high as slower speeds), this low power consumption contributes to the PRIMEPOWER's reliability.

For speedy operation on PRIMEPOWER information, the SPARC64 V enjoys an on-chip, non-blocking Level 1 128 KB instruction and a 128 KB data cache, as well as an on-chip 2 MB Level 2 cache. These caches help ensure optimum PRIMEPOWER performance by having data and instructions ready for the microprocessor's use as soon as possible.

To further speed PRIMEPOWER operation, the SPARC64 V uses state-of-the-art branch prediction, speculative fetch, and parallel and out-of-order instruction execution for six arithmetic units (two for address generation, two for floating point, and two for fixed point). Up to six general or four floating point operations can run simultaneously.

#### Mainframe High Availability

A completely novel feature set of the SPARC64 V is its mainframe-like high-availability capabilities. As a result of the presence of this feature set, no other microprocessor for the UNIX domain is comparable to the SPARC64 V's ability to reduce errors and enhance server system availability in the IT infrastructure.

The SPARC64 V maintains parity protection on its arithmetic units and architecture registers. This guarantees data path integrity. Moreover it has Error Correcting Code (ECC)-based single-bit automatic correction and double-bit detection of errors on its built-in Level 2 cache for both data and TAG. It also provides ECC on its Level 1 cache for data and the Level 1 TAG is duplicated in the Level 2 cache. Even if there were a Level 1 TAG error, recovery is immediate by copying the TAG information from the Level 2 cache. Finally ECC is provided on the microprocessor's low-latency, high-throughput (e.g., 40 GB/sec. on a fetch) internal buses.

For automatic error recovery purposes, including mainframe operation, the SPARC 64 V microprocessor permits automatic, hardware-based instruction retry for intermittent errors. Moreover, while there is ECC- or duplication-based automatic correction for cache error, automatic degradation occurs when there is a fixed breakdown in these caches. When the number of correctable errors occurring in a processor exceeds a threshold, Solaris recognizes the condition as a prediction of uncorrectable error and can logically take the processor offline.

## PRIMEPOWER SYSTEM ARCHITECTURE

PRIMEPOWER's hardware architecture is based on Fujitsu's long experience with mainframe and supercomputer architectures. Much of this hardware remains unique. Consider, for example, the interconnect fabric, known as the crossbar, which links PRIMEPOWER system boards on the new series Models 1500 and 2500.<sup>8</sup> A crossbar is a point-to-point connection between backplanes, which is, in turn, attached to PRIMEPOWER system boards.

In a fully populated Model 2500, the crossbar connects two backplanes with a total of sixteen system boards, each with eight SPARC64 V microprocessors, and 64 GB of ECC memory. Fujitsu's crossbar runs at an unheard of 520 MHz clock rate<sup>9</sup> (Models 900 and 1500 use a 270 MHz clock) and provides an aggregate 133 GB/sec. bandwidth (provided through two 66.5 GB/sec. crossbars that may be degraded if necessary). (Throughput is 8.3 GB/sec. in or out for each system board.)

<sup>8</sup> The new series Model 900 uses direct connection and does not require a crossbar.

<sup>9</sup> This figure is achieved as a result of Fujitsu's supercomputer-like wave pipeline technique to raise the crossbar clock frequency while keeping data skew at a low level using a novel clock and data transmission synchronization method.

Crossbars speed up performance and increase system availability. They minimize the distance between system boards and reduce latency, lower the number of parts, and eliminate the use of cabling. In the event of a crossbar failure, the crossbar (actually two 66.5 GB/sec. crossbars) can be degraded by a failure isolation, crossbar reconfiguration, and restart to use its remaining resources (one 66.5 GB/sec. crossbar). With these features, the PRIMEPOWER server does not have to be removed from service.

Infrastructures that need to deal with the high-performance needs of high system or peak load situations in OLTP, data warehousing, or other channel-intensive applications will be best served by PRIMEPOWER with this state-of-the-art, industry-leading, high-clock-rate, high-bandwidth crossbar.

There are other architecture features of interest as well that enhance the IT infrastructure. For example, as Table 2 indicates, PRIMEPOWER offers a dual-clock configuration. At the failure of the operating clock, a reboot reconfigures PRIMEPOWER to the second clock.

PRIMEPOWER also features industry-leading resource dynamic reconfiguration, I/O Multipathing, and PCI board hot swap. Finally, hard partitioning for error containment and system security with software partition-like granularity (down to one CPU and one GB of memory<sup>10</sup> with PRIMEPOWER's new Extended Partitioning [XPAR]) provides the best features of both hardware and software partitions.<sup>11</sup> XPAR allows dynamic, real-time resource modification and redistribution to match varying loads. The result is a large decrease in cost of ownership as different workload situations can be accommodated with the same server.

Finally, the new PRIMEPOWER series offers architectural flexibility through the separation of its PCI slots for I/O from its system boards. Previously, a server had to add system boards to increase I/O. This is no longer necessary as the PCI expansion is physically separate in PRIMEPOWER's I/O (PCI/Disk) boxes. Each Model 2500 system board connects to two such PCI/Disk boxes. (Further details on all of these architectural features and functions and their value to the IT infrastructure are available in the PRIMEPOWER system architecture white paper in this series.)

---

<sup>10</sup> 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*.

<sup>11</sup> This is discussed later in the resource management section of this white paper.

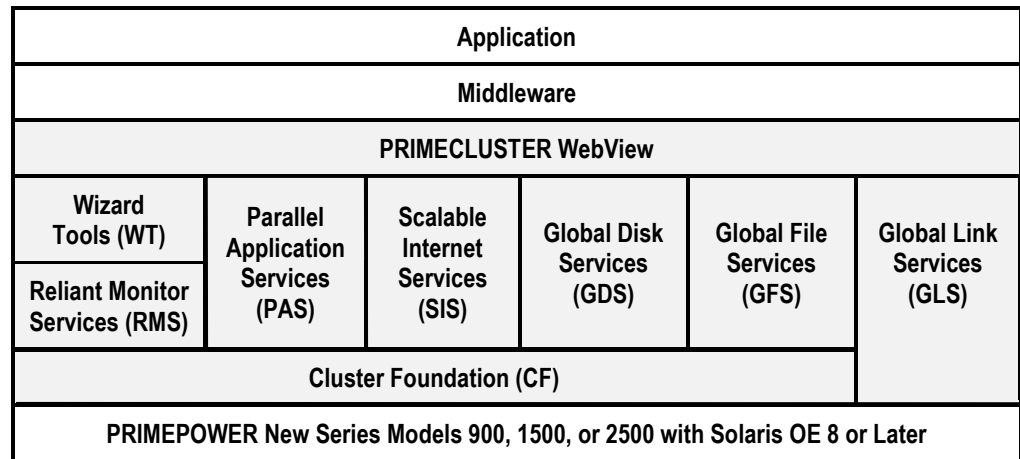
## PRIMEPOWERS COMBINE IN A PRIMECLUSTER

PRIMEPOWER servers are designed for the highest levels of availability and scalability. When PRIMEPOWER nodes are combined using Fujitsu's fourth generation PRIMECLUSTER with its state-of-the art, clustering software modules, the optimum IT infrastructure is possible.<sup>12</sup>

PRIMECLUSTER is a modular software-based product. Its modules can be mixed and matched to meet cluster node, storage, and network infrastructure needs for high availability, scalability, parallel database operation, and manageability. PRIMECLUSTER grows out of Fujitsu's long history of UNIX-based clusters. This experience spans clusters with mainframe-based technology (SynfinityCLUSTER) and a mature cluster system offered by Fujitsu Siemens Computers (Reliant UNIX), which is rooted in technology derived from Pyramid Technologies.

The modular software architecture of PRIMECLUSTER is shown in Figure 1. A WebView module provides a Java-based GUI administrative view so that PRIMECLUSTER is managed from a single console. The modular structure built on the Cluster Foundation (CF) module allows the easy addition of future modules to accommodate other cluster needs as they arise.

**FIGURE 1:**  
**PRIMECLUSTER**  
*Modular Software  
Architecture*



The separately purchasable, yet fully integrated modules required for the four PRIMECLUSTER products (Scalability Server, Parallel Server, HA Server, and Enterprise Edition) are shown in Table 3 along with a brief module description in the table key. (Further details on these modules and PRIMECLUSTER may be found in a separate white paper in this series.)

<sup>12</sup> While PRIMECLUSTER is discussed here with reference to Fujitsu's PRIMEPOWER computers and Solaris, PRIMECLUSTER software is not limited to this environment. Through its clever use of operating- system dependent and independent modules and interfaces, PRIMECLUSTER is operating system, hardware platform, and interconnect technology independent. For example, PRIMECLUSTER software modules also run on IA Linux and are InfiniBand-ready. The result of such extensive applicability is lower cost of ownership due, in part, to decreased administration and training costs.

**TABLE 3:**  
**PRIMECLUSTER**  
*Module Requirements*

PRIMECLUSTER Product	Module							
	CF	RMS	WT	PAS	SIS	GDS	GFS	GLS
Scalability Server	R				R			
Parallel Server	R			R		R	R	R
HA Server	R	R	R			R	R	R
Enterprise Edition	R	R	R	R	R	R	R	R

**Key:** R = Required; The CF module provides cluster architecture components; RMS provides failover management components; WT provides Wizard tools for applications; PAS provides parallel database components; SIS provides load balancing components; GDS provides storage mirroring and management components; GFS provides file sharing and scalability components; and GLS provides components for network or card failures.

**SCALABILITY SERVER**

The PRIMECLUSTER Scalability Server provides service to client requests (via the Internet or direct connections). Typically, it is suitable for a front-end web or proxy server or an application server in a tiered architecture. It is configured with a gateway node with one or more backup gateway nodes. Various other cluster nodes (service nodes) provide the responses to the user requests. In the Scalability Server configuration, PRIMECLUSTER provides dynamic load balancing for TCP/IP-based multi-instance applications. This load balancing provides scaling capabilities for these applications that are often CPU-intensive. To provide high availability for this scaling, especially if a node fails, the load can be redistributed. Persons responsible for the IT infrastructure use Scalability Server to accommodate unpredicted or time-varying loads in their business environment.

**PARALLEL SERVER**

PRIMECLUSTER Parallel Server provides the cluster operating environment for parallel databases such as Oracle9i RAC. It is suitable for the database tier of a tiered architecture. For optimum use of Oracle’s unique small message and lock caching scheme, Cache Fusion, PRIMECLUSTER employs a private, low-latency, message passing protocol. With Parallel Server, the database is available even if a node fails. Moreover, database parallel access is synchronized, and the preferred architectural concept of a single database design with a database instance on each cluster node can be maintained for the IT infrastructure.

**HA SERVER**

The HA Server is suitable typically as an application server in a tiered architecture. It delivers application high availability through resource monitoring (by means of “detectors” in the PRIMECLUSTER nodes, network, and applications), failure detection, local recovery where possible, and failover services. With the HA Server, PRIMECLUSTER administrators can choose the cluster architecture environment’s most suitable configuration. This includes the

commonly used multiple node backup, multiple node failover, cascading failover, selective failover, manual or automatic failover, and local recovery.

#### **ENTERPRISE EDITION**

The Enterprise Server is the high-end PRIMECLUSTER offering and performs all the required cluster operations for the business-critical IT infrastructure. It is possible, for example, to put together a two-node PRIMECLUSTER with the PRIMEPOWER Model 2500 servers. Each of these servers can have an unheard of 128 SPARC64 V microprocessors, 512 GB of PC266 ECC RAM, 320 PCI I/O slots, and 15 hard partitions with dynamic degradation and dynamic reconfiguration capability. Such a cluster exceeds the capability of anything else available in the industry.

### **PRIMEPOWER SYSTEM MANAGEMENT**

The optimum delivery of all of the PRIMEPOWER capabilities discussed in this white paper requires the highest level of server system management. This high level of system management means that the system must be managed so that all of its mission-critical attributes are available at the same time. This quality of service is critical. Without it the attributes lose most, if not all, of the value they provide to the information technology (IT) infrastructure user and the persons responsible for providing and maintaining this infrastructure.

Fujitsu provides PRIMEPOWER system software and dedicated hardware facilities to manage the three key components of the server family's operation. These components include computing resource management, high-availability management, and security management. All of these management capabilities are accommodated in an integrated manner through the use of PRIMEPOWER's GUI-based WebSysAdmin browser.

#### **RESOURCE MANAGEMENT**

The ability to divide a single physical system into multiple virtual systems has long been available in mainframes. Such hardware partitioning is available today for UNIX-based SMP machines. PRIMEPOWER's hardware partitioning allows virtual system performance and quality of service guarantees. It provides defective device isolation until repair can be accomplished, which occurs without taking the partition out of service. Also available is inter-partition security so that one virtual system does not become a security risk to another. Such hardware partitioning is often of greater value than firmware-controlled software or logical partitioning wherein faults can affect other partitions, that share joint resource ownership.

With the just-introduced PRIMEPOWER new series XPAR, it is possible to achieve partition granularity down to one CPU and one GB of memory (a part of a

system board).<sup>13</sup> This is similar to the granularity levels expected of software partitioning, but the advantages of hardware partitioning are maintained. Moreover, XPARs use dynamic reconfiguration to add or remove partition resources without

taking the partition down. The partition continues to be highly available for user service. This flexibility is also similar to what has been traditionally achieved with software partitioning.

Where even the flexible XPARs are not sufficient to meet resource allocation needs, PRIMEPOWER offers ARMTech ShareEnterprise resource management software for CPUs and memory allocated to processes in an XPAR (but not across XPARs). (See *Sidebar 2: ARMTech Technology* for an ARMTech overview.) Further details are available in a separate ARMTech white paper in this white paper series.

Resource management is not possible without knowing the status of the PRIMEPOWER resources. PRIMEPOWER administrators use GUI-based tools to track the state of their partition or server resources. With the aid of such monitoring, administrators can review workloads and performance, use dynamic reconfiguration to change resources to meet changing needs, and more.

## HIGH-AVAILABILITY MANAGEMENT

To ensure the best in high availability, PRIMEPOWER monitors itself during system boot-up and system operation so that defective subsystems and components can be isolated. In this manner, any effect on system availability is held to a minimum. In the long term, PRIMEPOWER's goal is to autonomically perform such detection and isolation without

the need for human intervention. Table 2 outlines PRIMEPOWER's industry-leading managed high-availability features and functions.

Monitoring PRIMEPOWER's processors, memory subsystem, operating system and disk storage is performed by PRIMEPOWER's redundant SCF. SCF depends on an independent processor for its system management capabilities. It operates independently of the system's SPARC64 V microprocessors.

### Sidebar 2: ARMTech Technology

To provide PRIMEPOWER with industry-leading resource management, Fujitsu provides Active Resource Management Technology (ARMTech). To provide the optimum in such management, Fujitsu offers ARMTech ShareEnterprise, the most comprehensive of the ARMTech products. By using ARMTech ShareEnterprise on Solaris, users achieve increased PRIMEPOWER stand-alone server and cluster efficiency. This also allows a superior quality of service, and reduced cost of ownership.

ARMTech allows administrators to share resources on a PRIMEPOWER server between users, groups, and applications. ARMTech software is tightly coupled to the Solaris kernel so that the resource use of individual processes can be controlled so they do not affect other users or processes. It is thereby possible to guarantee infrastructure service levels. ARMTech also monitors and reports resource usage. This is particularly important in a server consolidation environment.

ARMTech slices up available processor resources and shares using an administrator policy-determined, hierarchical-use policy that ranges across users, groups, and applications. For example, to eliminate resource conflicts, ARMTech enforces user virtual memory quotas or limits the number of processes a user can initiate. When needed, ARMTech assigns critical databases a higher proportion of resources. Alternatively, databases can be assigned a fixed resource proportion to ensure a continued quality of service to the database users regardless of database load.

ARMTech accommodates all of these needs by assigning either resource shares or resource reservations. This automated policy can be changed, on the fly, while resources are being consumed, and while system loads are changing. Shares entitle a resource consumer to a proportion of all the shares held by all contesting resource consumers. The more shares a resource consumer has, relative to its peers, the greater resource proportion it receives. In contrast, reservations are an allocation of an absolute percentage of the processor resources. This approach protects high-priority resource consumers.

<sup>13</sup> 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*.

During system boot-up, the SCF looks for the possibility of system “hangs” generated by hardware failures. If a defective, or potentially defective, device is detected, the system is restarted after the defect is isolated. Note that, when the SCF detects a CPU failure (even during boot-up), for example it reports this event to the System Management Console (SMC). In turn, the SMC keeps logs and reports to the Remote Customer Support System (REMCS).

All of this, including whether or not a partially failed device is still used, and when it is replaced, is subject to the particular operating policy of the IT environments. Such dynamic degradation is under the control of the system administrator. Even after the PRIMEPOWER has booted successfully, the SCF is monitored to detect undesired operation. Here too, a failing or failed device is isolated to minimize the need to bring the system down.

Finally, PRIMEPOWER minimizes system downtime when a defective system board is replaced. Its Flexible Memory Architecture (FMA) feature supports such hot swapping even if the system board is loaded with the Solaris kernel. This represents an advance over previous PRIMEPOWER servers in which a kernel-loaded system board could not be hot swapped.

#### **SECURITY MANAGEMENT**

PRIMEPOWER’s XPAR, as noted earlier, offers the advantages of hard partitioning together with the granularity capabilities of software partitions. The hard boundaries between XPARs prevent users and applications in one XPAR from causing security violations in another XPAR. A secure system for an enterprise environment requires additional elements. Fujitsu provides PRIMEPOWER and third-party based security management capabilities such as:

- Security for the SMC required for the PRIMEPOWER 900/1500/2500.
- Solaris Operating Environment security, which includes an integrated firewall and IPSec security functions for data integrity and protection.
- Third-party package support for a variety of scalable firewalls, trusted operating systems, VPNs, intrusion detection, enterprise security management packages, and real-time monitoring.

The combination of these security offerings, which are based on the security policies of an IT environment, provide PRIMEPOWER installations with the highest achievable security levels and ensure system uptime even in the face of overt attacks as well as software glitches.

## REVIEW OF KEY POINTS

This white paper provides an overview of Fujitsu's PRIMEPOWER new series servers and how they meet the needs of the business-critical IT environment for performance, high availability, use of industry standards, low cost of ownership, system and resource management, and more. All of these PRIMEPOWER features and attributes are the basis for PRIMEPOWER's ability to deliver a high quality of service to the IT environment.

This quality of service continues a long Fujitsu tradition in high-performance computing that began with supercomputers and highly available mainframes. Fujitsu's low latency, high-speed PRIMEPOWER interconnect grows out of its supercomputer experience. In addition, the Fujitsu-designed SPARC64 V microprocessor – at the heart of PRIMEPOWER – achieves a unique level of high availability based on the Fujitsu mainframe high-availability error detection, isolation, and recovery.

Fujitsu's goal of providing low cost of ownership and high performance in the same offerings extends to PRIMECLUSTER, the cluster available in PRIMEPOWER. Based on a software module approach wherein the modules can be mixed and matched to provide highly available, scalable, database or enterprise clusters, PRIMECLUSTER provides the highest level of cluster capability.

PRIMEPOWER's capabilities provide the IT infrastructure with numerous advantages in today's volatile business environment. These advantages include,

- PRIMEPOWER supports the market's demand for fast, cost-effective processing due to its industry-leading hardware and software architecture geared to handle the most challenging applications.
- PRIMEPOWER allows the quick change of business strategies thanks to its fast accommodation of new applications and changing application workloads and growth.
- PRIMEPOWER minimizes user downtime and maximizes quality of service for users because of its high-availability features and functions.
- PRIMECLUSTER reduces cluster cost of ownership (especially the "people cost" component) due to its extensive use of GUIs, Wizards, and automated operation and administration.
- PRIMEPOWER and PRIMECLUSTER offer world-record performance as published by results showing the PRIMEPOWER nodes and clusters continuing to capture industry-standard benchmarks on TPC-C, Oracle, SAP, BAAN, and SPEC.
- PRIMEPOWER and PRIMECLUSTER are ideal for server consolidation due to their ability to reduce costs and add redundancy with minimal effort.
- PRIMEPOWER and PRIMECLUSTER enjoy industry-leading scalability thanks to their use of state-of-the art components such as the SPARC64 V microprocessor and a supercomputer-based crossbar switch.

- PRIMEPOWER employs the industry-standard Solaris operating system and the SPARC International-compliant SPARC64 V microprocessor. As a result, the system offers maximum “openness” and consequent lower costs due to the availability of trained personnel and more than 15,000 applications.
- PRIMEPOWER and PRIMECLUSTER are tuned for optimum performance with industry-standard Oracle9i RAC including Cache Fusion support for speedy transmission of Oracle locks and data.
- PRIMECLUSTER offers tight integration with PRIMEPOWER hardware features such as partitioning, dynamic reconfiguration, error detection, and correction to ensure the highest availability cluster.

Simply put, PRIMEPOWER offers industry-leading hardware and software features in one package to ensure the highest performance, best availability, easiest management, and lowest cost of ownership. PRIMEPOWER is available worldwide along with local service so that multinational corporations using PRIMEPOWER may enjoy hardware, software, and personnel economies of scale at all of their locations.

Considering these advantages and the other technologies discussed in this white paper, it is clear that PRIMEPOWER should be a short-list candidate for single SMP server or clustered SMP server operation in the business-critical IT infrastructure. Further information concerning PRIMEPOWER’s attributes can be found in the other six white papers in this series, and on the Fujitsu websites.