

PRIMEPOWER, SPARC, and the Advanced Product Line

In June 2004 Sun Microsystems and Fujitsu ended months of speculation by announcing a partnership and consolidated roadmap for their products in an effort to cut research/development and manufacturing costs of the next generation product line.

It is currently anticipated that Sun will introduce the UltraSPARC IV+ processor in late 2005, to be followed by "APL" (Advanced Product Line) the first line of machines developed in partnership with Fujitsu.

Interestingly, this agreement currently only covers the SPARC-based servers: Sun and Fujitsu will continue to develop and market their own individual x86 based servers.

Over the last eighteen months, I have been in a privileged position to be able to work with many of the current Sun and Fujitsu products. This article is intended to introduce the current PRIMEPOWER range to those who are not already familiar with them and to consider the future of the APL product.

PRIMEPOWER

As a seasoned bigot of Sun Microsystems servers, I was deeply suspicious when we started to receive the Fujitsu-Siemens branded hardware. However, my initial scepticism has since proved to be unfounded as to this day, the equipment continues to be installed, and it is clear that the product range is well engineered and dependable.

The machines are constructed with rugged metal casings rather like the NEBS compliant Netra servers and finished with a smart grey paintjob. As the servers are all locked away in the datacentre, aesthetics are not a primary concern, but it does inspire confidence in the product to have a "solid" feel about them. A small sector of the market may be disappointed that the PRIMEPOWER range are not NEBS certified, but this is of no concern to the vast majority of customers.

SPARC64

Whilst Sun has a strategic alliance with Texas Instruments to fabricate their processors, Fujitsu manufacture their own SPARC64 processor. Any compatibility worries are addressed by SPARC V9 level 2 compliance, as certified by SPARC international— it will run any SPARC Solaris application that will run on the equivalent Sun hardware.

The performance from the Fujitsu CPU is a significant improvement over currently available processors for Sun hardware. A bonus prize is also awarded to Fujitsu for having a 1.89GHz processor where Sun can only support 1.2GHz (at the time of writing).

Sun Vs Fujitsu Single Processor CPU Performance (as reported by SPEC)

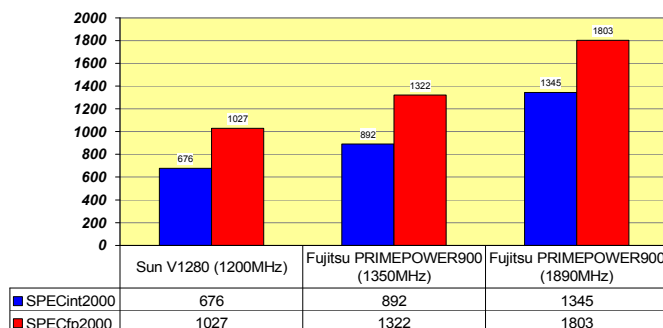


Figure 1 – Single Processor Performance

The SPEC CPU2000 speed benchmarks in Figure 1 are geared towards the performance of a single-processor system, which disadvantages the facilities available to Sun's dual-core UltraSPARC IV. However, even when considering multicore and multiprocessor configuration Fujitsu still appears to hold a marginal advantage. Figure 2 shows the comparison SPEC CPU2000 rate benchmarks between a multiprocessor PRIMEPOWER 900 against a multiprocessor E4900.

Sun Vs Fujitsu Multiprocessor CPU Performance (as reported by SPEC)

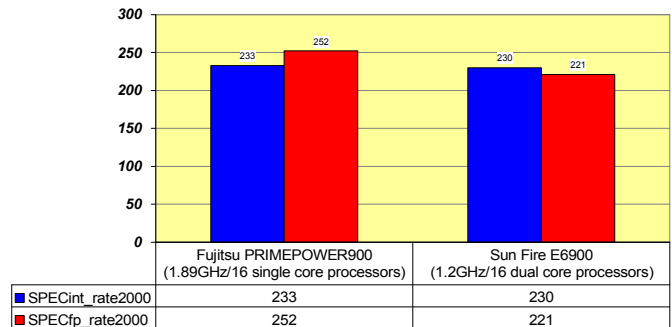


Figure 2 – Multiprocessor CPU performance

With my formative years as a SysAdmin spent working with IBM servers, I've always felt that having a display panel on the front of a machine is a worthwhile addition. I was pleased to see that Fujitsu have provided exactly this on their midrange and enterprise servers. As the machine boots, status messages from the POST (power on self test) are displayed on the two-line LCD display, giving a constantly updated status of the component being tested.

How many times have we all seen unlabelled (or mislabelled) equipment in an unfamiliar datacenter? When the system finishes booting, the LCD display usefully shows the hostname of the system, thus giving a valuable verification of system identity.

Additionally, by using the two control buttons beside the LCD panel a menu system can be navigated that allows the operator to health check the hardware (AC input and Temperature, for example), and also force a reset or even a crash dump (similar to dropping to OBP and issuing the "sync" command).

PRIMEPOWER 2500

The top-of the range PP2500 is the notional equivalent of Sun's E15k/E25k servers. The system cabinet capacity is 16 system boards, each containing up to eight of Fujitsu's SPARC64 processors – giving a maximum of 128 processors. The machine can be divided into domains, much like the technology that Sun has had in its' product lines since the early days of the E10k.

What is different here though is that a system board can be logically divided into two, giving 32 "logical" system boards that can be distributed between a maximum of 15 domains. This implies that a single physical system board can potentially be a member of two domains: this isn't a problem – it just takes a little more thought and care when designing your system and performing any dynamic reconfiguration work.

We have been managing two of these beasts for approximately a year now, and they have thus far proven themselves to be very dependable, and computationally very quick.

The PP2500 is accompanied by a Systems Management Console (SMC) much like the E10k System Service Processor (SSP), which provides management and monitoring facilities. Much like Sun's SSP, Fujitsu's SMC is standalone workstation that runs some specialised software for managing and monitoring the PP2500. Although, unlike the E10k the PP2500 does not *depend* on the SMC either to boot or for DR capabilities.

One peculiar feature of the PP2500 (and its' predecessor PP2000) range is the lack of a flexible console connection like the Sun equivalent. Taking the example of the E10k where console services are provided from the SSP over Ethernet to the control board. The control board then communicates with the boot processor of any particular domain via the backplane.

Unfortunately, this appears to be a disappointing omission from the Fujitsu range at present – there is of course a console connection, but is implemented as clumsy, thick RS/232 serial cables connected to the lowest numbered system board in each domain.

On the face of it, this doesn't appear to be a big problem, but one of the benefits of the dynamic system domains is the ability to reconfigure the number of boards/CPU's/Memory dynamically without having to visit a potentially remote or difficult to access site.

Many organisations running servers of this size also have strict "secure access" policies regarding entry to datacentres, and in this situation, reducing the need for physical access to the server can only be a good thing. Reconfiguring or creating a domain could potentially require a cable to be moved on the server.

That aside, remote console services are provided via an arbitration daemon running on the SMC that controls access to the Console Connection Unit. The CCU is essentially just a Network Terminal Server, connecting the serial cables of the domains to the network. In comparison to the feature rich "netcon" of the E10k, this set-up appears to be somewhat Heath Robinson.

The essentials are taken care of – a single session can have read-write access and there can be multiple simultaneous read-only sessions (switching of a session between RW and RO is unfortunately not possible). Read-write sessions can be forcibly terminated for those situations when your colleague has disappeared off to lunch with the console locked.

Software

The entire PRIMEPOWER range runs Solaris, just like any other Sun/SPARC product. If you examine a normal Solaris 8 media kit, you'll notice that the label describes it as "FOR SUN™ COMPUTER SYSTEMS" (I transcribe this from a media kit that is conveniently parked on my desk at the moment). This label is the only externally noticeable difference between the Sun and Fujitsu media kits.

Solaris 9 08/03 is a step in the right direction: on this release, Sun have provided the necessary OEM code to support PRIMEPOWER.

Disappointingly, although we have OEM support on the Sun media kit it is still necessary to use Fujitsu qualified patches, downloaded from Fujitsu's website, rather than from the usual source at SunSolve.

When using the operating system, there are very few clues to tell you that you're not using a Sun machine – "uname -i" shows that the Fujitsu hardware is identified by a prefix of "FJSV", followed by a code to identify the model of the machine: here is the output from a PP2500 domain:

```
# uname -i
FJSV, GPUZC-L
```

Fujitsu also bundles a plethora of extra packages to cover diverse functions such as crash dump analysis, performance monitoring and management. Included in the bundled software is very flashy looking Java based web-based management GUI.

I however, am not a particular fan of graphical interfaces and found it intolerably slow. Additionally, I hear on the grapevine that Sun is bundling "Webmin" with Solaris 10, which sounds like a much more comprehensive solution for those who crave the ability to reboot the wrong machine with a slip of the mouse.

```
# pkginfo |grep FJSV
application FJSVana      Auto Dump Analyze Tool
application FJSVapcs     Auto Power Control System
```

application	FJSVapcsw	APCS Administration View
application	FJSVbse	Enhanced Support Facility Information Management
application	FJSVcmnd	Machine Administration
application	FJSVcmrm2	Machine administration REMCS Interface for High-end Server
system	FJSVcpc	CPU Performance Counter package
system	FJSVcpcx	CPU Performance Counter package
application	FJSVcpcupd	CPU patrol diagnosis
system	FJSVdef	Server Default Configuration package
application	FJSVdmp	Crash Dump Assistant
application	FJSVdrcmd	Dynamic Configuration Core Command
system	FJSVdrd	Fujitsu PRIMEPOWER DR Daemon
system	FJSVdrdr	Fujitsu PRIMEPOWER DR Daemon init scripts (Root)
system	FJSVged	Fujitsu Gigabit Ethernet Adapter Driver
system	FJSVgid	Fujitsu Gigabit Interface Adapter
system	FJSVgidx	Fujitsu Gigabit Interface Adapter (64-bit)
system	FJSVhea	SunOS Header Files
application	FJSVIomp	IOMP Common Framework Level 2 package
application	FJSVmadm	Machine administration common package
application	FJSVmaom	Machine Administration Online Manual
application	FJSVmasv	Machine Administration Package (server)
system	FJSVmdb	Fujitsu Platform Modular Debugger
system	FJSVmdbx	Fujitsu Platform Modular Debugger (64-bit)
system	FJSVpiclu	Fujitsu PICL Libraries, and Plugin Modules (Usr)
system	FJSVpnl	Panel switch driver package
system	FJSVpnlx	Panel switch driver package (64-bit)
application	FJSVracm	REMCS Advanced Agent Installer
application	FJSVrrda	SystemWalker SS-Agent Resource Delivery Agent
system	FJSVscd3	System Control Driver 3 (SCF3 Driver)
system	FJSVscr3	System Control 3 (root)
system	FJSVscu3	System Control 3 (user)
system	FJSVscuj3	System Control 3 - japanese
system	FJSVsec	Fujitsu Serial Driver for FJSV, GPUZC-L and FJSV, GPUZC-L
application	FJSVsirms	SIRMS
application	FJSVsnap	System Information Output Tool
system	FJSVssf	Server Support Feature
system	FJSVulsa	FUJITSU ULTRA LVD SCSI Host Bus Adapter Driver
system	FJSVulsax	FUJITSU ULTRA LVD SCSI Host Bus Adapter Driver (64-bit)
system	FJSVvplr	Fujitsu platform links
system	FJSVvplu	Fujitsu usr/platform links
application	FJSVvts	Fujitsu/PFU Validation and Test Suite
system	FJSVwarn	System Warning Package
application	FJSVwvbs	Web-Based Admin View
application	FJSVwvcnf	WWW Server for Admin View
root	#	

Resource Management

Sun has for a number of years touted their "Solaris Resource Manager" (SRM) product. This product allows more focussed management of resources within the operating system, for example, it allows an operator to assign a maximum percentage of CPU or memory that an application or user can occupy.

Since its initial introduction SRM has been integrated into Solaris 9 and is now part of the standard install. In fact, SRM is based on technology that was developed by a company called Aurema – Sun established a relationship and licensed the Aurema "Sharell" technology for integration into the operating system.

Fujitsu appear to have taken this one step further – they have maintained a relationship with Aurema, and distribute the latest of its ARMTech software with the PRIMEPOWER servers.

Fundamentally ARMTech is a dynamic resource manager whereas SRM is a static tool. ARMTech introduces the concept of a Resource Consumer that may be a Solaris User, Group of Users, or an Application (and any hierarchy including all of these).

An Application Resource Consumer is a defined executable which may be qualified with command line switches or environmental parameters that are specified at invocation. This makes it easy for ARMTech to resource manage (for example, multiple Oracle instances in a single Solaris Operating Environment).

ARMTech allows for resource reservation (a guarantee of a minimum amount of resource), hard and soft caps (a limit on the amount of resource consumed) and sharing (a relative allocation of resource dependant on active resource consumers). The value of ARMTech is that these settings can be changed dynamically without the need to restart either ArmTech or the Solaris Operating System.

teleservice

Teleservice is available as an optional component to the SMC. This allows a modem to be connected to the SMC workstation that is capable of dialling home in case of a reported failure. This has the potential to be an invaluable service for organisations running mission critical servers, but don't have the luxury of 24x7 operations staff.

One neat piece of functionality that teleservice also provides is the ability to enable remote access for remote vendor support. The obvious security problems associated with such a configuration have been addressed by making the dial-in account password expire 30 minutes after it has been set.

When a Fujitsu engineer requires access, he would telephone the customer and ask for the password to be reset. The SysAdmin performs this with the "teleadm" menu interface, which takes care of the detail of generating and setting a valid password which can then be communicated to the engineer. However, it would appear that token-card authentication systems such as RSA SecurID is not yet supported for extra security.

Of course, for obvious reasons with a single-modem setup like this the teleservice is unable to dial-home in the event of a problem for the duration that the engineer is connected.

```

1  ----- TELESERVICE administration -----
a  (A) Connect to Center Information Service (CIS)
>b (A) Operate Teleservice session
c  (A) Common adjustments
d  (A) Install Teleservice channel
e  (A) Install and administer Teleservice center
f  (C) Remote operating with SINIX TE
g  (A) Analyze log files
h  (A) Manage database
i  (A) Backup/restore database
j  (A) Informations

2  ----- Operate Teleservice session ----- (A)
a  (A) Activate shadow
b  (A) Activate shadow for extension window
c  (B) Diagnosis session by a center
>d (B) Activate temporary entry
e  (A) Show state of Teleservice-Channel

3  Activate temporary entry ----- (B)
>0 (B) Bracknell Teleservice centre
6  (B) an: 4 ----- Temporary entry for center 0 -----

Password:      xen46
Modem mode:    9
Telephone Number: xxxxxxxxxxxx

Return to previous frame by pressing <F6>.

HELP  PREV PAGE  NEXT PAGE          CANCEL  CMD-MENU  EXIT

```

Figure x, teleadm interface

APL

By mid-2006, APL will replace the existing Sun Fire and PRIMEPOWER product lines. According to information from Fujitsu, the design of the low-end hardware (workstations and entry level servers) is to be entirely provided by Sun, using Sun's forthcoming Niagara processor. The mid-ranger servers are to couple Fujitsu processors (multicore SPARC64 VI due in early 2006) and Sun system boards.

The engineering of the enterprise class server line will be provided entirely by Fujitsu (also using the SPARC64) – suggesting the high-end servers will be the next generation of the current PP2500.

When the servers are launched, it is currently anticipated that the product will continue to be branded to the vendor, regardless of the production facility where it was manufactured. Fujitsu and Sun will each receive revenue based on the engineering that each partner has contributed to the product. Brian Sutphin, Senior VP of Corporate Development at Sun illustrates: "If a system is sold by Sun that contains components developed by Fujitsu, there's going to be revenue to Fujitsu on that sale".

We hear that APL support can be provided by either company. In fact, even today, Fujitsu are authorised to sell and support the entire Sun product range.

Production Facilities

Sun currently have two manufacturing facilities: one here in Scotland and one in Oregon (alas, the Newark plant has closed its' manufacturing facilities in recent times). Current speculation about the future of APL suggests that manufacturing will be shared between Sun's existing manufacturing facilities and also a Fujitsu plant near Tokyo.

It would make sense for the facilities to be capable of production to accommodate the needs of their geographic responsibilities (Tokyo for APAC, Linlithgow for EMEA and Hillsboro for the Americas), however recent press releases suggest that the high-end APL products (presumably the successor to the PP2500), will be manufactured exclusively in Fujitsu's Tokyo plant.

This appears to be a single point of failure: Sun were in a similar position when the E10k product was being produced exclusively in Oregon. An issue that Sun overcame by building a \$57m extension to the Scottish production facility and ramping up production of the F15k range at that site also.

It is possible that this plan is designed to reduce the learning curve when they roll out production to other manufacturing sites, however it remains to be seen how Fujitsu will address the risk of disaster recoverability on this site.

Conclusion

Overall, I believe that the partnership will bring together the very best of the two organisations. Fujitsu has some innovative engineering to help keep SPARC/Solaris competitive, and Sun has a long and successful history of developing and supporting the Solaris environment.

The current PRIMEPOWER range has proved to us to be a solid and dependable line of systems with an exceptional price/performance.

Mike Scott is the director of Hindsight IT Ltd, a small Solaris consultancy based in Central Scotland. He has been working in the North East and the central belt for the last ten years, specialising in systems management with a keen interest in security and performance management. He can be contacted at sysadmin@hindsight.it

References

- [1] <http://www.spec.org/> Standard Performance Evaluation Corporation
- [2] <http://www.aurema.com/> Developers of ARMTech software