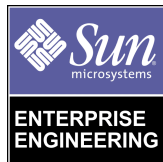




Dynamic Reconfiguration

By Enrique Vargas - Enterprise Engineering

Sun BluePrints™ OnLine - April 1999



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

Sun Microsystems, Inc.
901 San Antonio Road
Palo Alto, CA 94303 USA
650 960-1300 fax 650 969-9131

Part No.: 806-3837-10
Revision 02, April 1999

Copyright 1999 Sun Microsystems, Inc. 901 San Antonio Road, Palo Alto, California 94303 U.S.A. All rights reserved.

This product or document is protected by copyright and distributed under licenses restricting its use, copying, distribution, and decompilation. No part of this product or document may be reproduced in any form by any means without prior written authorization of Sun and its licensors, if any. Third-party software, including font technology, is copyrighted and licensed from Sun suppliers.

Parts of the product may be derived from Berkeley BSD systems, licensed from the University of California. UNIX is a registered trademark in the U.S. and other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the Sun logo, The Network Is The Computer, Sun Enterprise, Starfire, Gigaplane-XB, Sun Enterprise SyMON, and Solaris are trademarks, registered trademarks, or service marks of Sun Microsystems, Inc. in the U.S. and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.

The OPEN LOOK and Sun™ Graphical User Interface was developed by Sun Microsystems, Inc. for its users and licensees. Sun acknowledges the pioneering efforts of Xerox in researching and developing the concept of visual or graphical user interfaces for the computer industry. Sun holds a non-exclusive license from Xerox to the Xerox Graphical User Interface, which license also covers Sun's licensees who implement OPEN LOOK GUIs and otherwise comply with Sun's written license agreements.

RESTRICTED RIGHTS: Use, duplication, or disclosure by the U.S. Government is subject to restrictions of FAR 52.227-14(g)(2)(6/87) and FAR 52.227-19(6/87), or DFAR 252.227-7015(b)(6/95) and DFAR 227.7202-3(a).

DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID.

Copyright 1999 Sun Microsystems, Inc., 901 San Antonio Road, Palo Alto, California 94303 Etats-Unis. Tous droits réservés.

Ce produit ou document est protégé par un copyright et distribué avec des licences qui en restreignent l'utilisation, la copie, la distribution, et la décompilation. Aucune partie de ce produit ou document ne peut être reproduite sous aucune forme, par quelque moyen que ce soit, sans l'autorisation préalable et écrite de Sun et de ses bailleurs de licence, s'il y en a. Le logiciel détenu par des tiers, et qui comprend la technologie relative aux polices de caractères, est protégé par un copyright et licencié par des fournisseurs de Sun.

Des parties de ce produit pourront être dérivées des systèmes Berkeley BSD licenciés par l'Université de Californie. UNIX est une marque déposée aux Etats-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company, Ltd.

Sun, Sun Microsystems, le logo Sun, The Network Is The Computer, Sun Enterprise, Starfire, Gigaplane-XB, Sun Enterprise SyMON, et Solaris[®] sont des marques de fabrique ou des marques déposées, ou marques de service, de Sun Microsystems, Inc. aux Etats-Unis et dans d'autres pays. Toutes les marques SPARC sont utilisées sous licence et sont des marques de fabrique ou des marques déposées de SPARC International, Inc. aux Etats-Unis et dans d'autres pays. Les produits portant les marques SPARC sont basés sur une architecture développée par Sun Microsystems, Inc.

L'interface d'utilisation graphique OPEN LOOK et Sun™ a été développée par Sun Microsystems, Inc. pour ses utilisateurs et licenciés. Sun reconnaît les efforts de pionniers de Xerox pour la recherche et le développement du concept des interfaces d'utilisation visuelle ou graphique pour l'industrie de l'informatique. Sun détient une licence non exclusive de Xerox sur l'interface d'utilisation graphique Xerox, cette licence couvrant également les licenciés de Sun qui mettent en place l'interface d'utilisation graphique OPEN LOOK et qui en outre se conforment aux licences écrites de Sun.

CETTE PUBLICATION EST FOURNIE "EN L'ETAT" ET AUCUNE GARANTIE, EXPRESSE OU IMPLICITE, N'EST ACCORDEE, Y COMPRIS DES GARANTIES CONCERNANT LA VALEUR MARCHANDE, L'APTITUDE DE LA PUBLICATION A REpondre A UNE UTILISATION PARTICULIERE, OU LE FAIT QU'ELLE NE SOIT PAS CONTREFAISANTE DE PRODUIT DE TIERS. CE DENI DE GARANTIE NE S'APPLIQUERAIT PAS, DANS LA MESURE OU IL SERAIT TENU JURIDIQUEMENT NUL ET NON AVENU.



Please
Recycle



Adobe PostScript

Dynamic Reconfiguration

This is the first article in an on-going series about Dynamic Reconfiguration. This introductory article focuses on DR architecture and Sun Enterprise 3000-6500 and Starfire DR Technology Differences.

Dynamic Reconfiguration (DR) is a software mechanism that allows resources to be attached (logically added) or detached (logically removed) from the Solaris^(TM) operating environment control without incurring any system downtime. The current second generation DR technology was originally introduced in 1995 with the Cray CS6400 machine (it later became the Sun CS6400 server).

DR was an exclusive feature of the Sun EnterpriseTM 10000 Server (StarfireTM) platform, but it now supports the entire Sun Enterprise 3000-6500 Server line. The Starfire DR allows the relocation of resources within a single system, while the Sun Enterprise 3000-6500 DR allows the relocation of resources between separate systems.

DR works in conjunction with the hot swap hardware capability to allow the physical insertion or removal of system boards. Additionally, DR increases the level of availability and serviceability by providing a way to execute online hardware upgrades and maintenance.

The Starfire DR provides full functionality (CPU, memory and I/O interfaces) starting with the Solaris 2.5.1 release. In contrast, the Sun Enterprise 3000-6500 DR provides full functionality after the Solaris 7 5/99 release and limited support for I/O interfaces in the Solaris 2.6. 5/98 release (see TABLE 1).

TABLE 1 Dynamic Reconfiguration Functionality Matrix

Solaris Version	Starfire	Sun Enterprise 3000-6500
Solaris 2.5.1	Yes	No
Solaris 2.6	Yes	Limited support for I/O interfaces in 5/98 release
Solaris 7	Yes	Full support for I/O interfaces in 5/99 release

If required, Alternate Pathing (AP) provides an additional software mechanism in support of DR to provide redundant disk and network controllers (and their respective physical links) to sustain continuous network and disk I/O when resources are removed (detached) from a DSD. While AP is supported on both the Starfire and the Sun Enterprise 3000-6500 servers, it is *not* a prerequisite for DR.

Sun Enterprise™ 3000-6500 and Starfire™ DR Technology Differences

To have DR best serve your application needs, it is imperative to understand the differences between the DR capabilities applicable to the Starfire and those applicable to the Sun Enterprise 3000-6500 servers.

The Sun Enterprise 3000-6500 DR is simpler in nature since it doesn't support *Dynamic System Domains* (DSDs), and its functionality is limited to bringing brand new system board resources into the Solaris operating environment or detaching system boards to be serviced or shared with other Sun Enterprise 3000-6500 servers.

In contrast, the Starfire DR is more complex in nature because it includes all of the Sun Enterprise 3000-6500 DR functionality as well as DSD support to provide resource sharing within the same platform.

Resource Sharing

The Sun Enterprise 3000-6500 platforms are limited to *physically* sharing system board resources with other Sun Enterprise 3000-6500 platforms. Physical sharing involves the powering down and physical removal of a system board from a system (after having it removed from Solaris operating environment control). After the system board has been removed from the original system, it can then be inserted into an alternate system, powered up and registered with the active Solaris operating environment instance.

Both the Starfire and Sun Enterprise 3000-6500 platforms share the same risk of having bent pins temporarily or permanently damage the backplane interface and bring down the whole Solaris operating environment instance when exercising physical sharing. The Starfire DR removes the danger of having bent pins or worn out connectors by allowing the logical relocation of system boards between DSDs.

Note – Physical removal or installation of system boards might invalidate service contracts in place if attempted by non-qualified service personnel.

Memory Interleaving

The Starfire and Sun Enterprise 3000-6500 platforms *require* that memory interleaving be disabled in order to support DR. Memory interleaving helps increase memory subsystem performance by reducing the probability of *hot spots* or contention in a few memory banks by spreading access to multiple memory banks.

The active Gigaplane-XB™ interconnect on the Starfire server provides main memory access through a point-to-point data router which isolates data traffic between system boards and minimizes any performance degradation when memory interleaving is disabled. The passive Gigaplane bus interconnect on the Sun Enterprise 3000-6500 platforms is shared by all system boards and may introduce performance degradation with memory intensive workloads when memory interleaving is disabled.

System Board Diagnostics

The Sun Enterprise 3000-6500 DR executes POST as soon as the system board is attached to the system. The Sun Enterprise 3000-6500 system boards do not have hardware isolation applicable to DSDs and will expose the server to the possibility of bringing its active Solaris operating environment instance down by registering faulty on-board components. The Starfire DR, in contrast, controls the POST operation from the SSP over the JTAG interface and does not activate a newly attached system board until its full functionality is guaranteed.

Note – The Starfire DR allows an increased test coverage on all system board components by temporarily increasing the POST level through modification of the SSP's `blacklist(4)` file. User Interfaces

User Interfaces

Both the Sun Enterprise 3000-6500 DR and Starfire DR functions have their own private user interfaces that require manual intervention from a system administrator. The Sun Enterprise 3000-6500 DR uses the Sun Enterprise SyMON™ 2.0 framework and the `cfgadm(1M)` command interface to invoke the required DR functions.

The `cfgadm(1M)` command provides an interface to build automated DR scripts. Script building for unattended DR operations should be handled with caution when detaching a system board because removing system resources may have a negative impact on applications executing in the machine.

The SSP manages all the resources that are allocated to DSDs and keeps track of their boundaries; thus, it has to initiate all DR operations. The Starfire DR is performed through the following SSP applications:

- The `drview(1M)` command, which produces a Graphical User Interface to be invoked only through the `hostview(1M)` interface, see FIGURE 1.
- The `dr(1M)` command, a Tool Command Language (TCL) application that produces a command line interface monitor that supports a set of commands that integrate all DR functionality, see FIGURE 2.

Currently, the Starfire DR user interface does not provide mechanisms to automate DR operations since the current user interfaces are exclusively interactive.

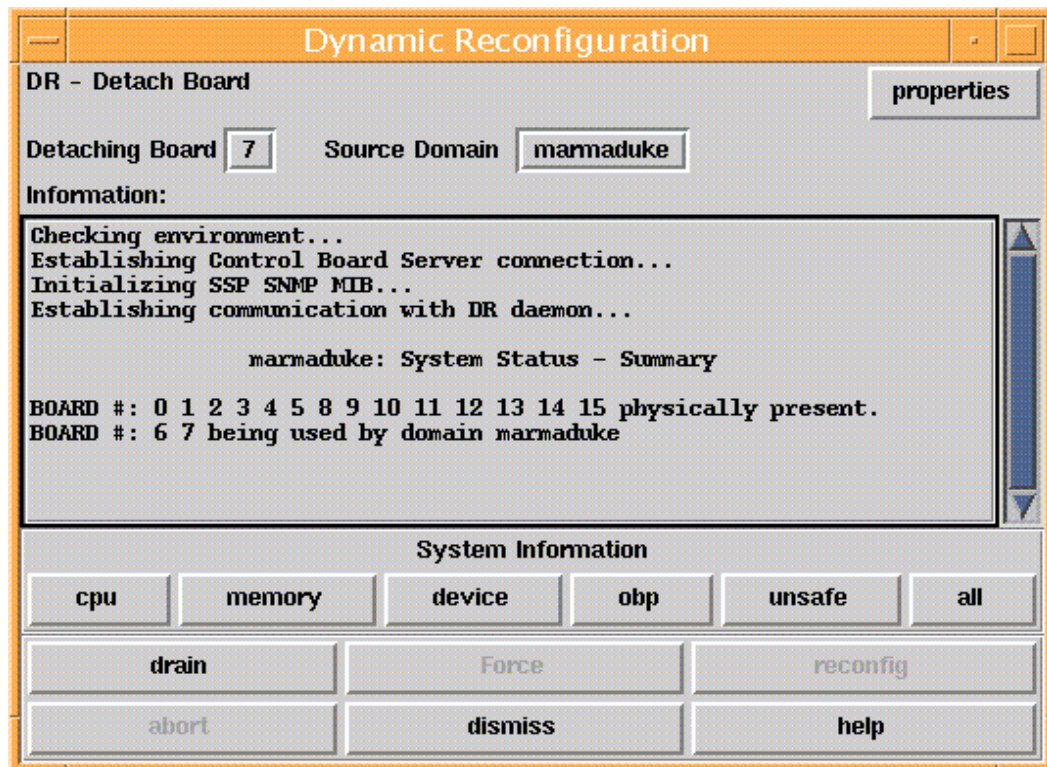


FIGURE 1 The SSP Hostview `drview` Application

```

cmdtool - COMICS-SSP
comics-ssp:flinstones% /opt/SUNWssp/bin/dr
Checking environment...
Establishing Control Board Server connection...
Initializing SSP SNMP MIB...
Establishing communication with DR daemon...

flinstones: System Status - Summary

BOARD #: 0 1 2 3 4 6 8 9 10 11 12 13 14 15 physically present.
BOARD #: 5 7 being used by domain flinstones
dr> drshow UNSAFE

Unsafe Devices

No unsafe devices currently open.
dr> drshow 7

Cpu Configuration for Board 7

cpu  status  | Processor Set | Bound Threads
---  -
id  count  | user  sys  procs
---  -
28  online  | None    0  |    0    0
29  online  | None    0  |    0    0
30  online  | None    0  |    0    0
31  online  | None    0  |    0    0

Cpu 20 is the bootproc.
dr>

```

FIGURE 2 The SSP dr Monitor Application

Solaris 7 Operating Environment™ DR Architecture

The DR framework was merged in the Solaris 7 operating environment to provide some module commonality between the Starfire and the Sun Enterprise 3000-6500 platforms. FIGURE 3 shows all software modules involved in the Solaris 7 Starfire DR for the SSP and DSD as well as the Sun Enterprise 3000-6500 server sides.

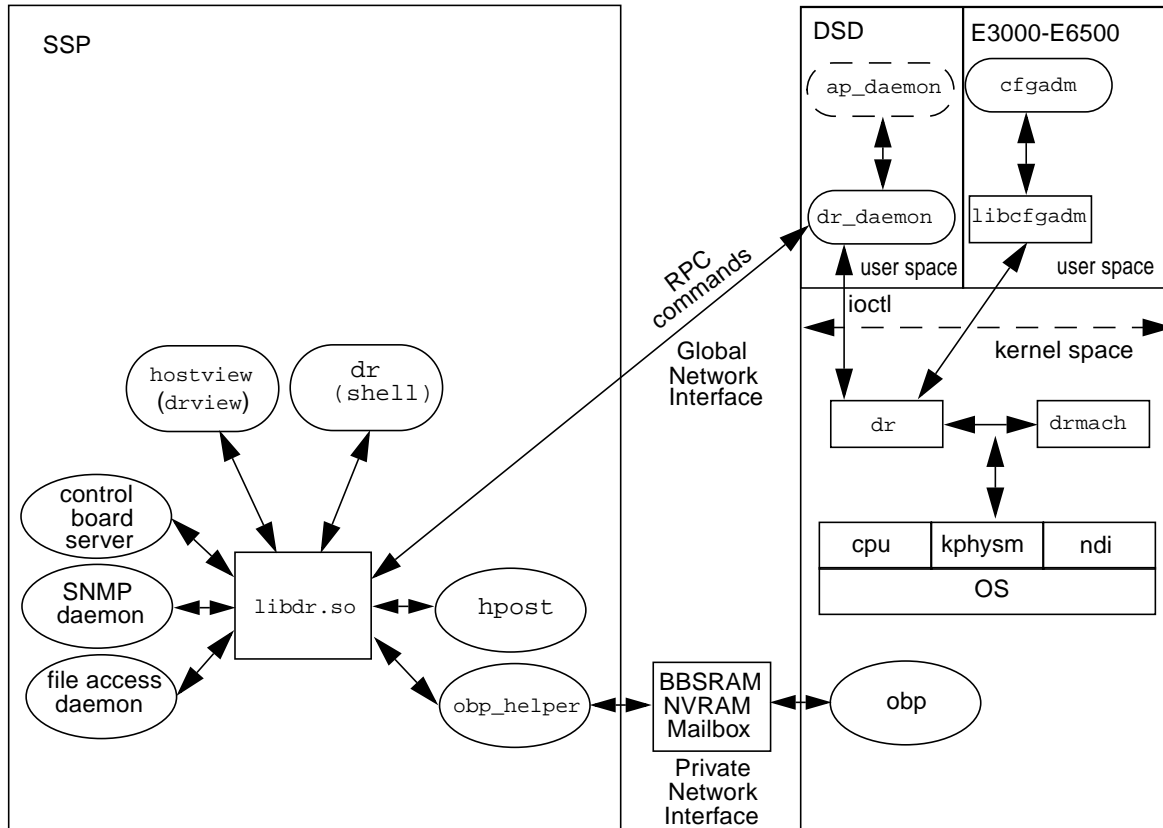


FIGURE 3 Solaris 7 5/99 Starfire DR Software Architecture

On the Sun Enterprise 3000-6500 platforms, all DR operations are invoked through the `cfgadm(1M)` user interface dynamically linked to the `libcfgadm(1M)` library. On the Starfire platform, all DR operations are initiated by the SSP, and only one outstanding operation is allowed at a time. Both the SSP and the DSD manage their own states to synchronize and guarantee the integrity of the DR operations.

The `hostview(1M)` and `dr(1M)` applications on the SSP provide user interfaces for executing DR operations. Both applications are dynamically linked to the `libdr.so` shared-object library, which provides access to the `dr_daemon` process on the DSD side through TCP/IP Remote Procedure Calls (RPC). The DR daemon process acts as a translator and intermediary between the SSP and the DSD.

The `libdr.so` library invokes both the `hpost(1M)` and `obp_helper(1M)` processes to register resource changes in the OBP kernel device tree and to execute Power-On Self-Test (POST) on added system boards.

If the `ap_daemon` was present on the server, the `dr_daemon` process would communicate appropriate device information to invoke the rerouting of I/O calls through registered alternate devices. The `dr_daemon` invokes the `dr` driver module, which is a common, platform-independent module used for both the Starfire and Sun Enterprise 3000-6500 servers. The `dr` driver communicates with the loadable `drmach` platform-specific module to invoke hardware functions specific to either the Starfire or Sun Enterprise 3000-6500 platforms.

The `ndi` (nexus device interface) layer provides the framework to manipulate the I/O device tree in support of the DR-Attach and DR-Detach operations. Both the `dr` and `drmach` modules will register changes in the CPU, memory and I/O resources while the operating system is manipulated to support the addition or removal of system resources.

Closing Statements

DR allows the logical sharing of resources between DSDs within a single Starfire platform to achieve committed application service levels. Physical sharing involves the introduction of brand new resources into the Starfire platform and the entire Sun Microsystems Sun Enterprise E3500-E6500 server line. Physical sharing may require involvement from Sun service organization to avoid non-compliance with the service agreements.

Author's Bio: Enrique Vargas

Enrique brings a wealth of large systems experience to Sun and specializes in high end UNIX offerings including the Enterprise 10000. Enrique came to Sun from Amdahl where he also focused on the high end Solaris systems.