
Contexte documentaire

Dossier de Pilotage E25K



Sun Microsystems France S.A.

13, Avenue Morane-Saulnier
BP 53
78142 Vélizy Cedex
Tél. : 01 34 03 00 00
Fax : 01 34 03 04 73

Avertissement

This product or document is 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 copyright 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 US. and other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the logo Sun, AnswerBook, SunDocs, Solaris, OpenBoot, OpenWindows, Solstice AdminSuite, Solstice Backup, Solstice DiskSuite, Solstice SyMON, SunSwift, SunVTS, Ultra and Enterprise are trademarks, registered trademarks, or service mark of Sun Microsystems, Inc. in the US. 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 licencees. Sun acknowledges the pioneering efforts of Xerox in researching and developing the concept of 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 licencees who implement OPEN LOOK GUIs and otherwise comply with Sun's written license agreements.

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.

Ce produit ou document est distribué dans le cadre d'accords qui en restreignent l'utilisation, la copie, la distribution et la décompilation. Aucune partie de ce produit ou document ne peut être reproduit 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 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 États-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company, Ltd.

Sun, Sun Microsystems, le logo Sun, AnswerBook, SunDocs, Solaris, OpenBoot, OpenWindows, Solstice AdminSuite, Solstice Backup, Solstice DiskSuite, Solstice SyMON, SunSwift, SunVTS, Ultra et Enterprise sont des marques de fabrique ou des marques déposées ou marques de service de Sun Microsystems, Inc. aux États-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 États-Unis et dans d'autres pays. Les produits portant les marques SPARC sont basés sur une architecture développée par Sun Microsystems, Inc.

L'interface graphique utilisateur OPEN LOOK et Sun™ a été développée par Sun Microsystems, Inc. pour ses utilisateurs et licenciés. Sun reconnaît les efforts de pionnier 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 DOCUMENTATION 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 PRODUITS DE TIERS. CE DENI DE GARANTIE NE S'APPLIQUERAIT PAS, DANS LA MESURE OU IL SERAIT TENU JURIDIQUEMENT NUL ET NON AVENU

1 PRÉSENTATION DES ÉQUIPEMENTS.....	7
1.1 ELÉMENTS DU SUN FIRE 25K.....	7
1.2 LE SERVEUR SUNFIRE 25K.....	8
1.2.1 Identification des sous-ensembles du serveur.....	9
1.2.2 Présentation des sous-ensembles du serveur.....	10
1.2.3 Alimentations 220V ac.....	13
1.2.4 Disjoncteurs associés aux sous-ensembles.....	14
1.2.5 Tiroirs de ventilation.....	14
1.2.6 System Boards.....	17
1.2.7 IO Boards (IB - Cartes d'entrées/sorties).....	18
1.2.8 System Controller.....	20
1.2.9 Centerplane board.....	22
1.2.10 Expander board.....	23
1.2.11 Domaines et réseaux internes.....	24
1.2.12 Disques systèmes (Se3120) des domaines.....	28
2 CONFIGURATION DE LA PLATE-FORME.....	29
2.1 PARAMÈTRES NÉCESSAIRES DU SUN FIRE 25K.....	29
2.1.1 Nom de la plate-forme.....	29
2.1.2 Configuration des System Controllers.....	29
2.2 FONCTION DES SYSTEM CONTROLLERS (SC).....	32
2.2.1 Administration matérielle du serveur.....	32
2.2.2 Administration des domaines.....	32
2.2.3 Fonctions et limites des SC lors du boot de domaines.....	32
2.2.4 Fonctionnalités portées par les SC aux domaines.....	33
2.3 LES DISPOSITIFS PROPRES AU SERVEUR SUN FIRE 25K.....	34
2.3.1 Basculement de System Controller.....	34
2.3.2 Reconfiguration dynamique.....	34
3 DESCRIPTION DE LA CONFIGURATION DES PLATES-FORMES SF25K.....	35
3.1 SUNFIRE25K D'ARCUEIL.....	35
3.2 SUNFIRE25K DE ROISSY.....	40
PROCÉDURES SYSTÈME.....	46
3.3 INTRODUCTION.....	46
3.4 TABLEAU DES MOTS CLÉS.....	47
3.5 CONNEXION AUX SC.....	48
3.6 IDENTIFIER LE SC PRINCIPAL (MAIN SC).....	49
3.7 ACTIVER LE MODE FAILOVER.....	50
3.8 BASCULEMENT DU SC SECONDAIRE EN SC PRINCIPAL.....	52
3.9 LES PROCESS ACTIFS SUR CHAQUE SC.....	53
3.10 ARRÊT D'UN DOMAINE.....	55
3.11 MISE HORS TENSION DU SUNFIRE 25K.....	56
3.12 MISE SOUS TENSION DU SF25K.....	57
3.13 ACTIVATION D'UN DOMAINE.....	59
3.14 BOOT D'UN DOMAINE.....	63
3.15 RÉCUPÉRATION DU CRASH DUMP.....	64
3.16 GESTION DES JETONS COD (CAPACITY ON DEMAND).....	65
3.17 RECONFIGURATION DYNAMIQUE (DR): DÉTACHEMENT D'UNE SYSTEM BOARD D'UN DOMAINE ACTIF.....	67
3.18 RECONFIGURATION DYNAMIQUE (DR): ATTACHEMENT D'UNE SYSTEM BOARD À UN DOMAINE ACTIF.....	70
3.19 RECONFIGURATION DYNAMIQUE (DR): DÉTACHEMENT D'UNE CARTE FC GÉRÉE PAR STMS.....	73
3.20 RECONFIGURATION DYNAMIQUE (DR): ATTACHEMENT D'UNE CARTE FC GÉRÉE PAR STMS.....	77
3.21 RECONFIGURATION DYNAMIQUE (DR): DÉTACHEMENT D'UNE CARTE SCSI/ETHERNET GÉRÉE PAR SVM/IPMP OU VXVM/IPMP.....	81
3.22 RECONFIGURATION DYNAMIQUE (DR): ATTACHEMENT D'UNE CARTE SCSI/ETHERNET GÉRÉE PAR SVM/IPMP OU	

VXVM/IPMP.....	85
3.23 CONTRÔLER L'ÉTAT DE LA PLATE-FORME.....	89
3.24 PROCÉDURE DE RESTAURATION D'UN SYSTEM CONTROLLER.....	90
3.24.1 Contexte:.....	90
3.24.2 Boot cdrom du system controller.....	90
3.24.3 Création du partitionnement du disque système.....	90
3.24.4 Création des systèmes de fichiers.....	91
3.24.5 Restauration de / et opérations manuelles à effectuer.....	91
3.24.6 Restauration de /export/install.....	93
3.24.7 Reboot en single-user.....	93
3.24.8 Réencapsulation dans Solaris Volume Manager du disque de boot.....	93
3.24.9 Mise en miroir du disque système avec SVM.....	94
4 TRAITEMENTS DES DYSFONCTIONNEMENTS.....	95
4.1 HISTORIQUE STANDARD DES MESSAGES DE DOMAINE.....	96
4.2 LISTE NOIRE (BLACKLIST FILE).....	96
4.3 OUTILS DE DIAGNOSTIC.....	97
4.3.1 hpost.....	97
4.3.2 prtdiag.....	97
4.4 ERREURS DE MÉMOIRE CORRIGÉES.....	98
4.5 DÉFAILLANCES SYSTÈME.....	98
4.5.1 Demande de redémarrage (reboot).....	99
4.5.2 Panic.....	99
4.5.3 Réinitialisations par watchdog, redmode ou xir.....	100
4.5.4 Hôte bloqué (heartbeat failure).....	101
4.5.5 Arbstop.....	102
4.5.6 Domainstop et recordstop.....	102
4.6 DOMAINE BLOQUÉ.....	103
4.7 INFORMATIONS TECHNIQUES POUR LE SUPPORT.....	105
4.7.1 Informations générales.....	105
4.7.2 Informations propres à un problème donné.....	105

Utilisation du document

Comment utiliser ce document ?

Ce document est principalement destiné à valider le bon fonctionnement de la plate-forme. Il est destiné au Chef de Projet Client et au représentant du Client habilité à signer le procès-verbal de recette, au Chef de Projet et Ingénieurs Projet Sun Microsystems et au *Technical Account Manager* Sun Microsystems.

Conventions typographiques

Les conventions typographiques utilisées dans ce document sont les suivantes.

Police ou symbole	Signification	Exemple
AaBbCc12	Les noms de commandes, fichiers, répertoires et affichages écran sont présentés en caractères non-proportionnels.	Editer votre fichier <code>.login</code> Utiliser <code>ls -a</code> pour lister tous les fichiers. Affichage écran : <code>nom_machine% you have mail</code>
AaBbCc12	Les commandes entrées sont présentées en caractères gras non-proportionnels pour les distinguer des affichages écran.	<code>nom_machine% mkdir /repertoire</code>
<i>AaBbCc12</i>	Les variables de la ligne de commande à remplacer par une valeur ou un nom sont présentées en caractères italiques non-proportionnels	Pour supprimer un fichier, entrer <code>rm nom_de_fichier</code>
<i>AaBbCc12</i>	Les titres de documents, éléments devant être mis en valeur et anglicismes usuels sont identifiés en italiques simples.	Faire un <i>dump</i> sur le lecteur de bande de la station d'administration.

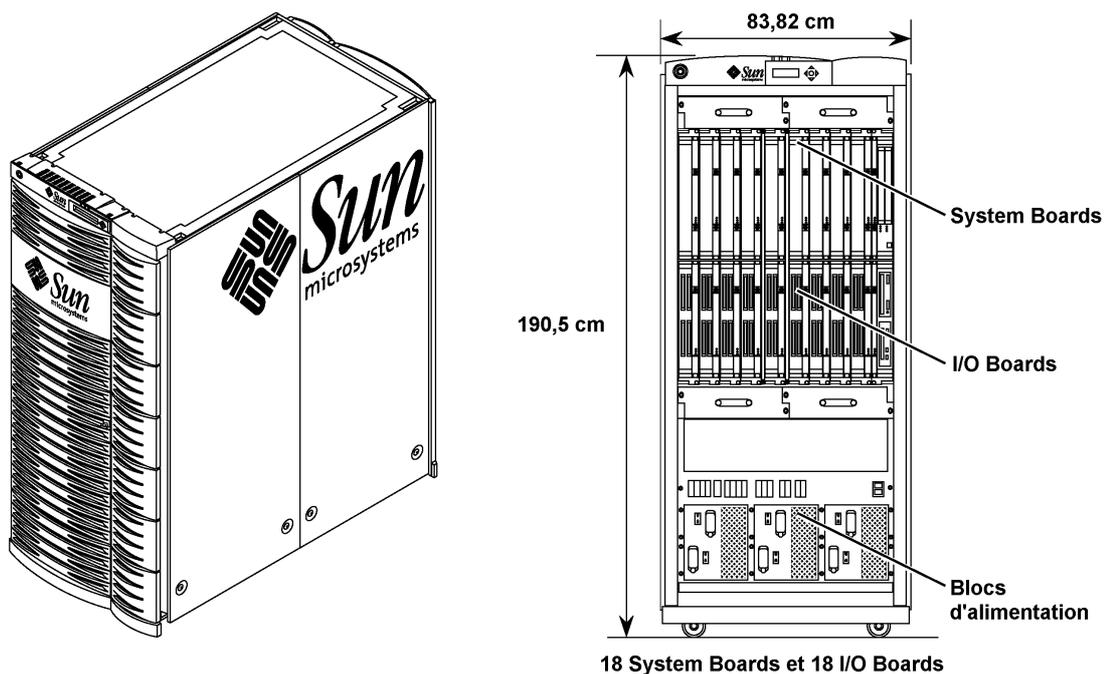
1 Présentation des équipements

1.1 Eléments du Sun Fire 25K

La plate-forme Sun Fire 25K comprend les éléments suivants:

- un serveur SunFire 25K,
- ses armoires d'extension pour les disques de boot.

Figure 1: Vue d'ensemble du SunFire 25K



1.2 Le serveur SunFire 25K

Le serveur SunFire 25K comprend trois ensembles fonctionnels :

- Les équipements attachés au fonctionnement système : alimentations et dispositifs de ventilation,
- Les cartes sur lesquelles sont réparties les fonctions logiques de la machine,
- Les équipements d'interconnexion associés.

Tableau 1: Récapitulatif configuration maximum d'un Sun Fire 25K

Équipement	Configuration
System boards (CPU/Memory)	18
CPUs	72
Nombre de barrettes mémoire DIMMs	576
Capacité mémoire maximum (avec 2-GByte DIMMs)	1152 GB
Sun Fireplane Interconnect	Active
Repeater boards	NA
Expander boards	18
Domaines	18
I/O boards (assemblies)	18
PCI assembly types	HsPCI+
PCI slots per assembly	4
Maximum PCI slots	72
Bulk power supplies	6
Power requirements	24 kW
System Control boards	2

Cinq types de cartes équipent le serveur par insertion sur les faces avant et arrière :

- Les *System Boards* qui regroupent les processeurs et mémoires,
- Les *I/O boards* qui permettent d'insérer jusqu'à 4 cartes d'entrée/sortie de type PCI,
- Les *expanders boards* qui assurent l'interface entre des SB et le CenterPlane,
- Les *Centerplane Support Boards* qui fournissent les ressources de chaque moitié de la carte fond de panier,
- Les *system controller* qui assurent l'interface à la console et autorise le *monitoring* des domaines configurés et de la plate-forme dans sa globalité (alimentations, système de ventilation, environnement). Les *system controller* sont chacun équipés d'un lecteur CDROM/DVD et d'une unité de sauvegarde de type DAT72 ainsi que de 2 disques embarquant Solaris et le produit SMS.

La carte d'interconnexion en fond de panier (*Centerplane board*) assure l'interconnexion des cartes du système. Elle présente une configuration physique symétrique sur chacune de ses faces autorisant ainsi l'insertion, à l'avant comme à l'arrière, de dix-huit *System Boards* et d'un *system controller* sur un slot (à droite) dédié par face.

1.2.1 Identification des sous-ensembles du serveur

Figure 2: Présentation face avant du serveur SF25K



1.2.2 Présentation des sous-ensembles du serveur

Figure 3: Vue générale du Sun Fire 25K

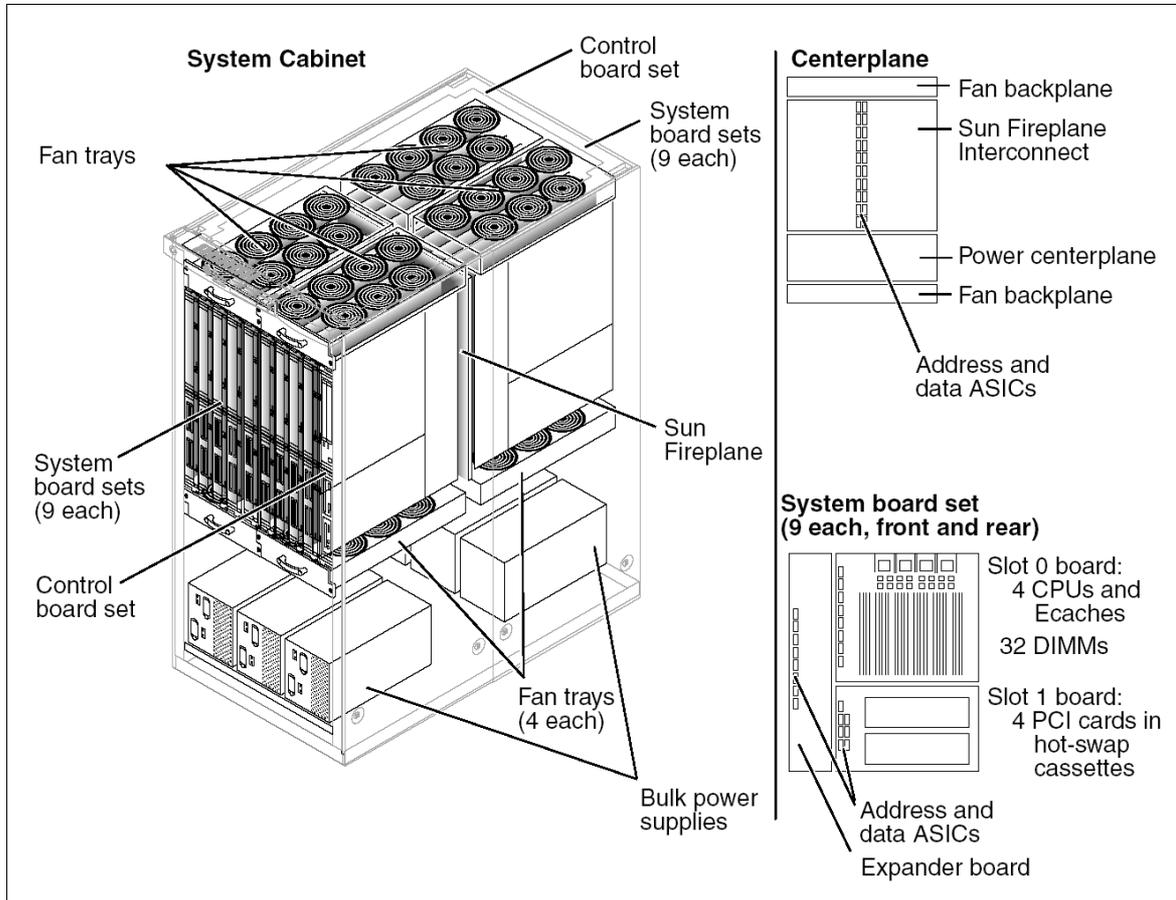
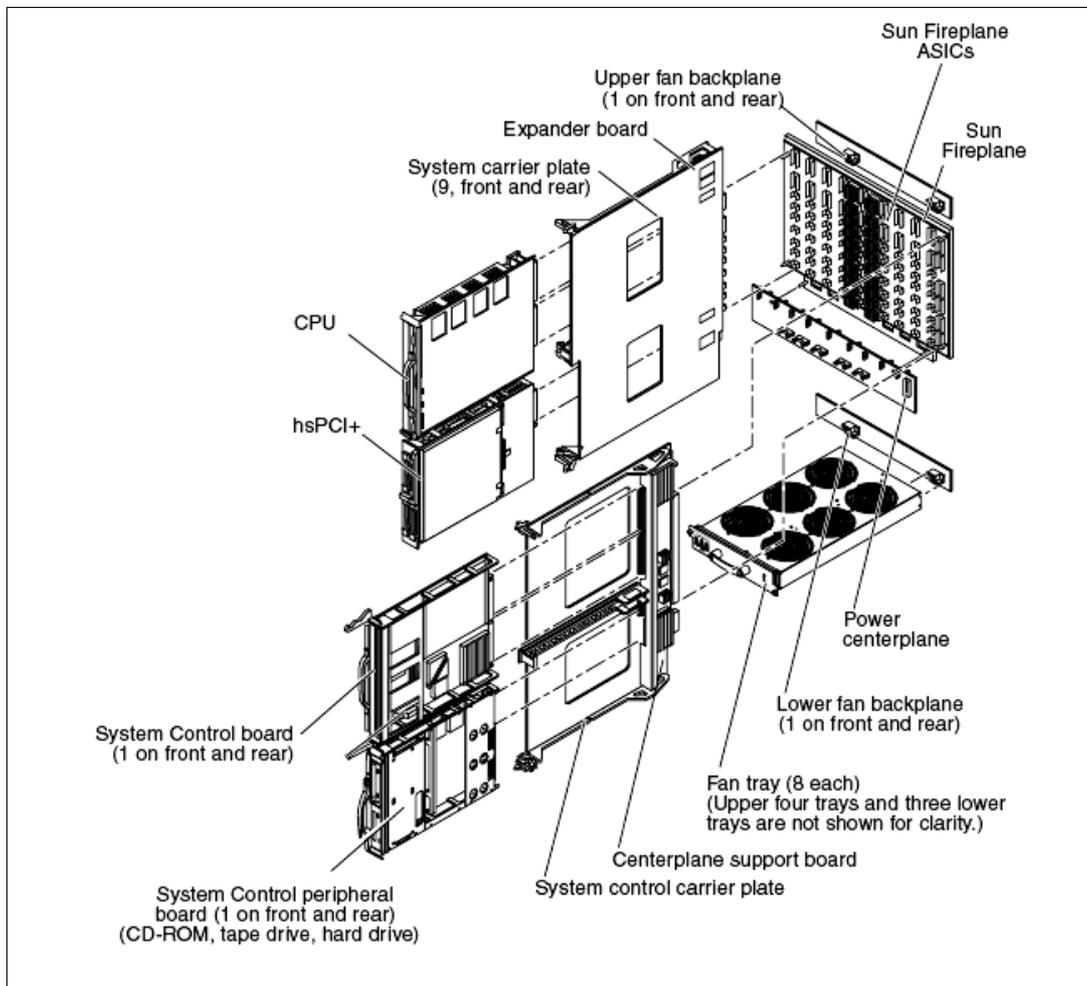


Figure 4: Vue éclatée du Sun Fire 25K



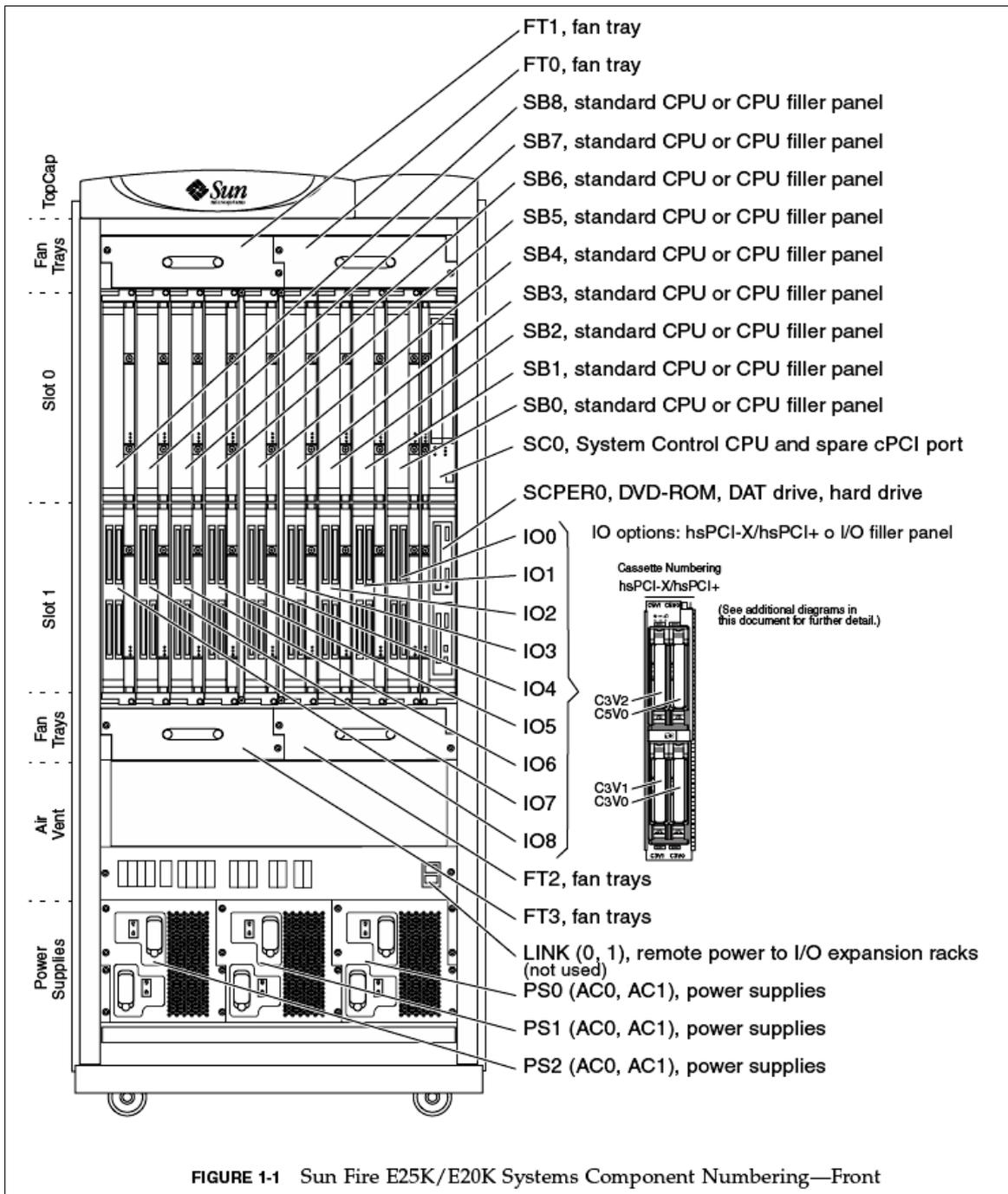


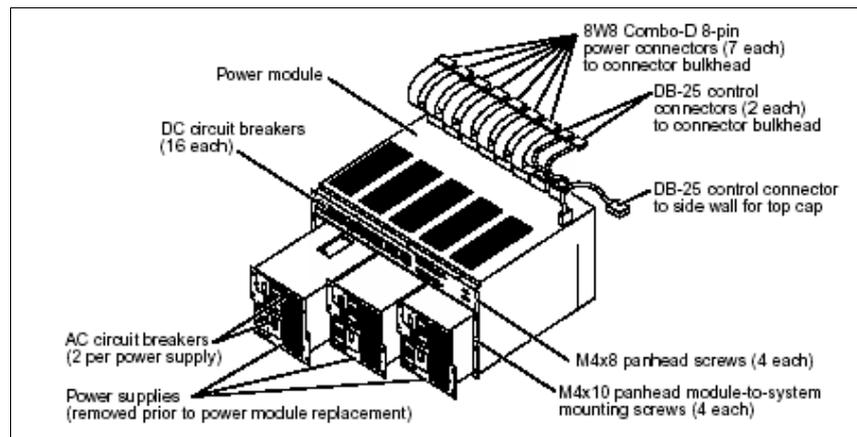
Figure 5: Identification des sous-ensembles



1.2.3 Alimentations 220V ac

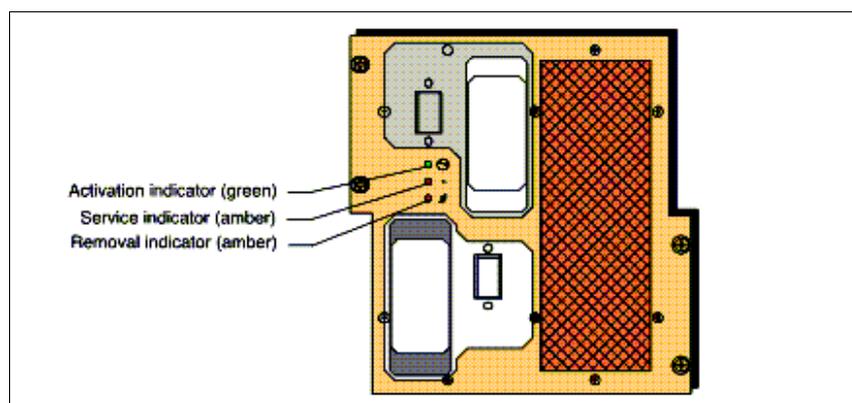
Elles comprennent les modules d'entrée 220 Vac (4000 Watt) Hot swappable.

Figure 6: Vue d'ensemble des blocs d'alimentations



Les modules d'entrée 220V_{AC} assurent le contrôle de l'énergie délivrée et fournissent à la plate-forme les tensions nécessaires à son bon fonctionnement; ils comportent en face avant : les prises d'entrée, les interrupteurs - disjoncteurs de chacun des blocs associés et les signalisations lumineuses de contrôle.

Figure 7: Vue d'ensemble d'un module d'entrée



Le pré-câblage interne de l'armoire permet le raccordement à l'énergie des modules d'entrée équipant le serveur. Il est recommandé de connecter les câbles sur des axes de distributions 220V_{AC} distincts.

Le Sun Fire 25K possède six blocs d'alimentation secteur (3 en face avant et 3 en face arrière). Chaque bloc d'alimentation est connecté à deux sources électriques distinctes et est capable de basculer automatiquement d'une source à l'autre en cas de défaillance de l'alimentation.

1.2.4 Disjoncteurs associés aux sous-ensembles

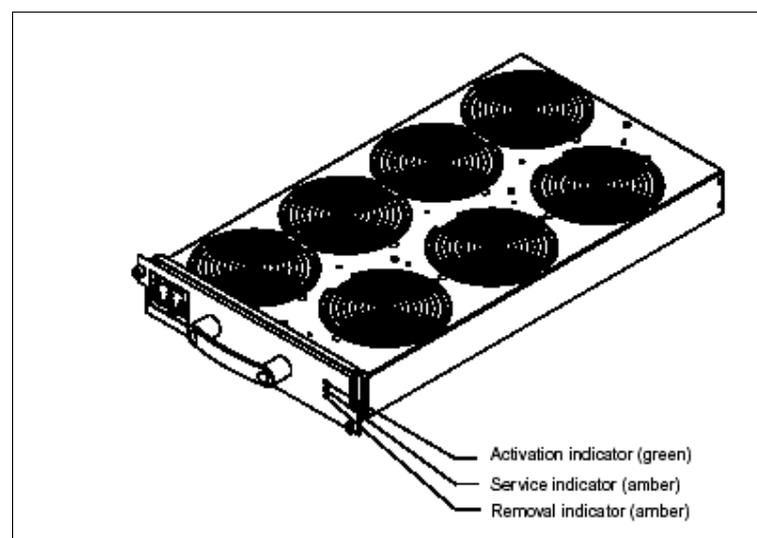
Situés sous les alimentations, ils commandent la distribution du 48V_{DC} aux sous-ensembles du système. Ils se déclenchent automatiquement lors d'un court-circuit. Ils sont repérés de la manière suivante :

- « FT x - x » (x compris de 0 à 7 par face) pour les quatre disjoncteurs associés aux tiroirs de ventilation « *Fan Trays* »,
- « EXx » (x compris de 0 à 9) pour les 9 disjoncteurs commandant la distribution aux *expander Boards*,
- « CSx » (x = 0 ou 1) pour les disjoncteurs associés aux *centerplane support board*,
- « SCx » (x = 0 ou 1) pour les disjoncteurs associés aux *system controllers*,
- « SCPERx » (x = 0 ou 1) pour les disjoncteurs associés aux *system controllers peripheral board*.

1.2.5 Tiroirs de ventilation

L'armoire serveur est refroidie par un flux d'air pressurisé ascendant produit par un dispositif de ventilation redondant comprenant huit tiroirs de ventilation.

Figure 8: Vue d'ensemble d'un tiroir de ventilation

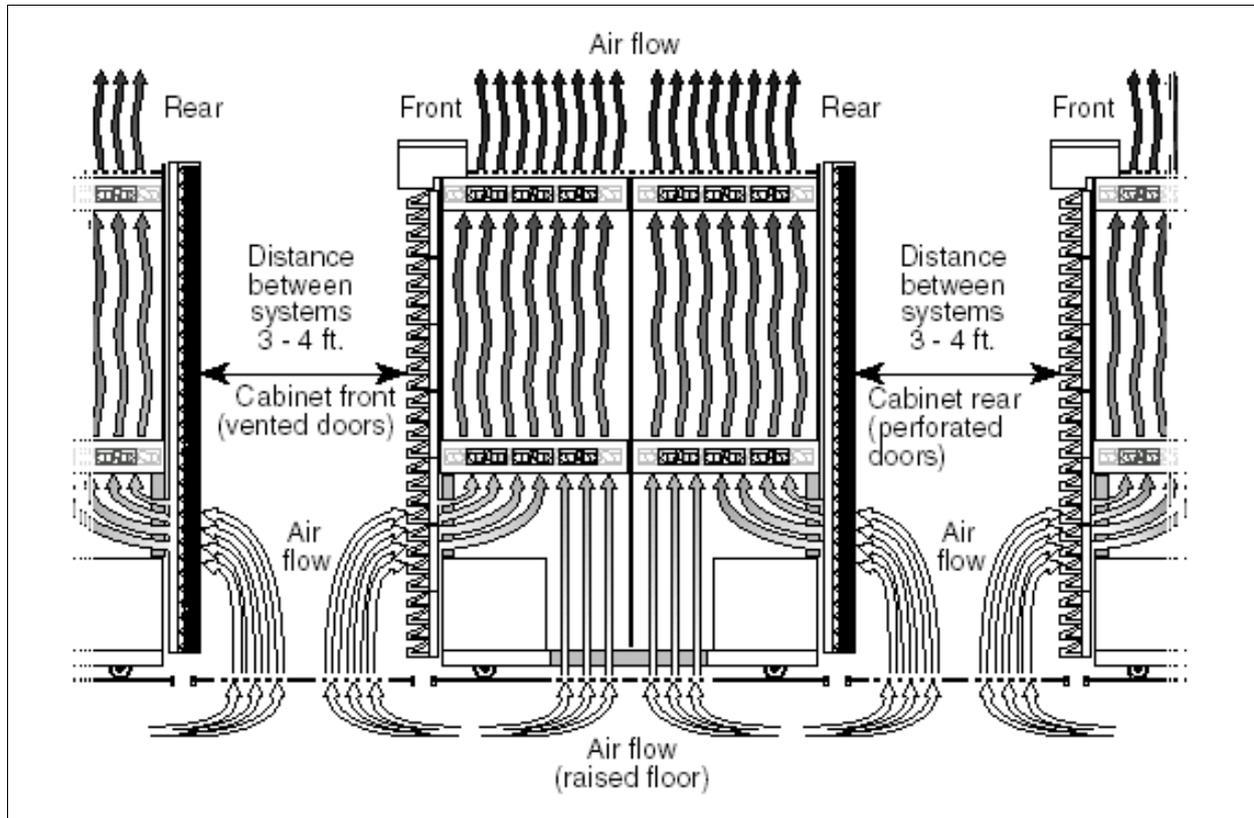


Huit disjoncteurs commandent la distribution de l'énergie à ces tiroirs.

Lors d'une procédure manuelle de mise sous tension du serveur, ils doivent être positionnés sur « *On* » (vers le haut) préalablement au basculement de tous les autres disjoncteurs (interrupteurs des modules d'entrée 220V_{AC} exceptés).

A l'inverse, lors de la mise hors tension du serveur, ils doivent être les derniers à être actionnés (modules d'entrée 220V_{AC} exceptés). Les signalisations en face avant des tiroirs permettent de vérifier leur mise sous tension.

Figure 9: Schéma des flux d'air des circuits de ventilation



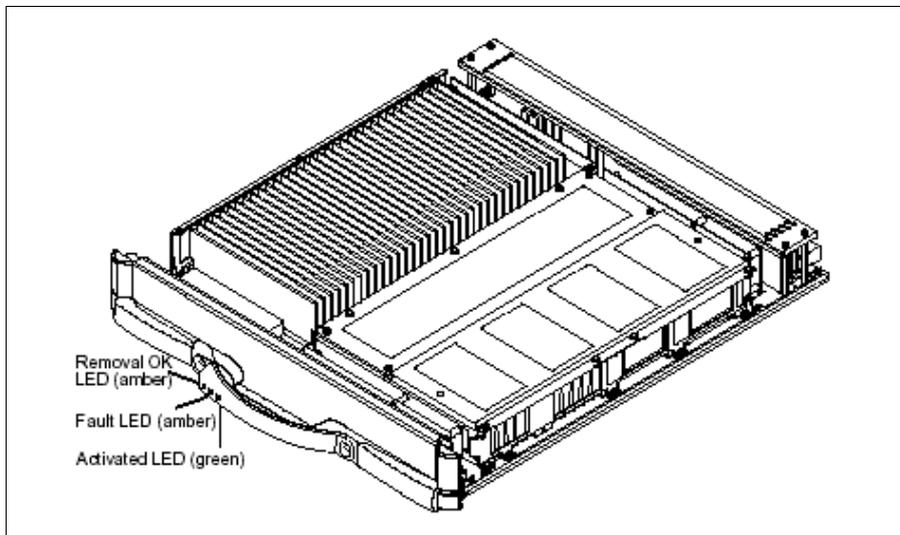
1.2.6 System Boards

Chaque *System Board* (SB) est munie de capteurs de température et de flux permettant la supervision du dispositif de ventilation. Les données correspondantes sont remontées au *system controller* actif.

De une à dix-huit *System Boards* (cartes système) peuvent équiper le serveur Sun Fire™ SF25K. Elles regroupent les processeurs et la mémoire. Chaque *System Board* peut être individuellement configurée par variation du nombre et de la nature de ces éléments :

- Un à quatre processeurs UltraSparcIV+™ (activation par licence Capacity On Demand)
- De 4 Go à 32 Go de mémoire réparti sur 4 *bank* de 8 slots.

Figure 10: Vue d'ensemble d'une system board de Sun Fire 25K



La carte système est une carte multicouches assurant l'interconnexion des processeurs et de la mémoire principale à la carte fond de panier. Les fonctions de contrôle, d'arbitrage, de tampons et de multiplexage sont assurées par des ASICs.

Chaque carte comporte ses convertisseurs 48V_{DC}/Tensions de service.
La distribution du 48V_{DC} à la carte est commandée par un disjoncteur individuel.

Les *System Boards* prennent place dans les dix-huit logements situés en partie milieu de l'armoire serveur tant à l'avant qu'à l'arrière (SF25K uniquement). Cette dernière regroupe également les signalisations lumineuses de maintenance.

1.2.7 I/O Boards (IB - Cartes d'entrées/sorties)

Le Sun Fire 25K peut accueillir jusqu'à 18 I/O Boards. Chaque I/O Board possède quatre emplacements PCI hot-swap.

Figure 11: Présentation de la face avant des IO Boards

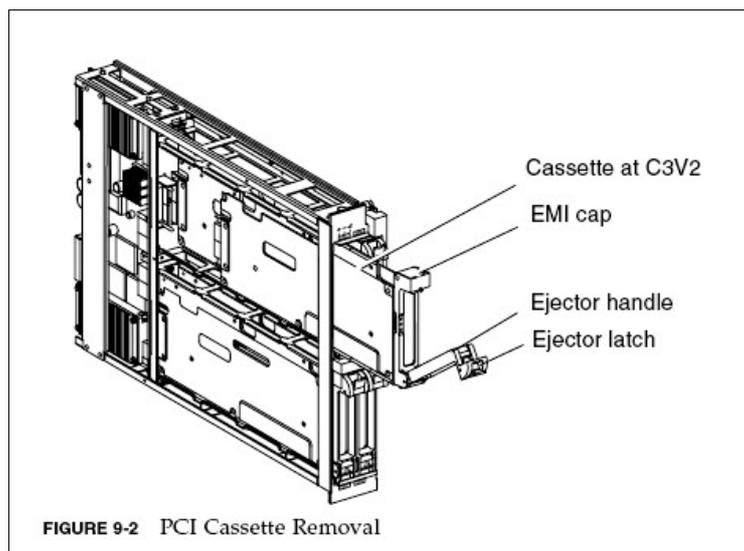
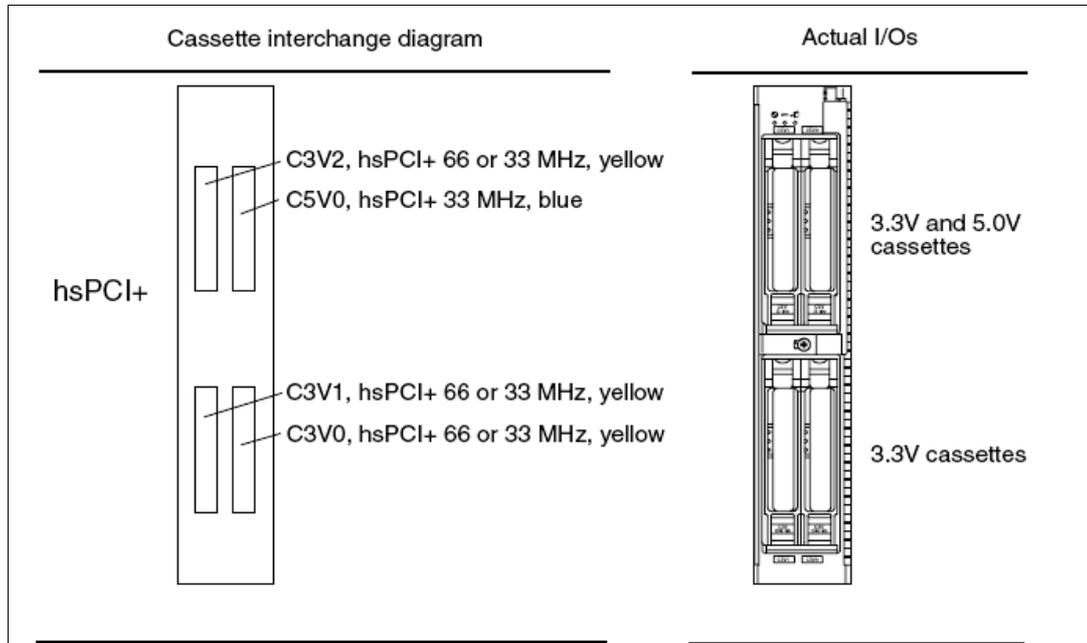


Figure 12: Emplacement et fonction des LEDs

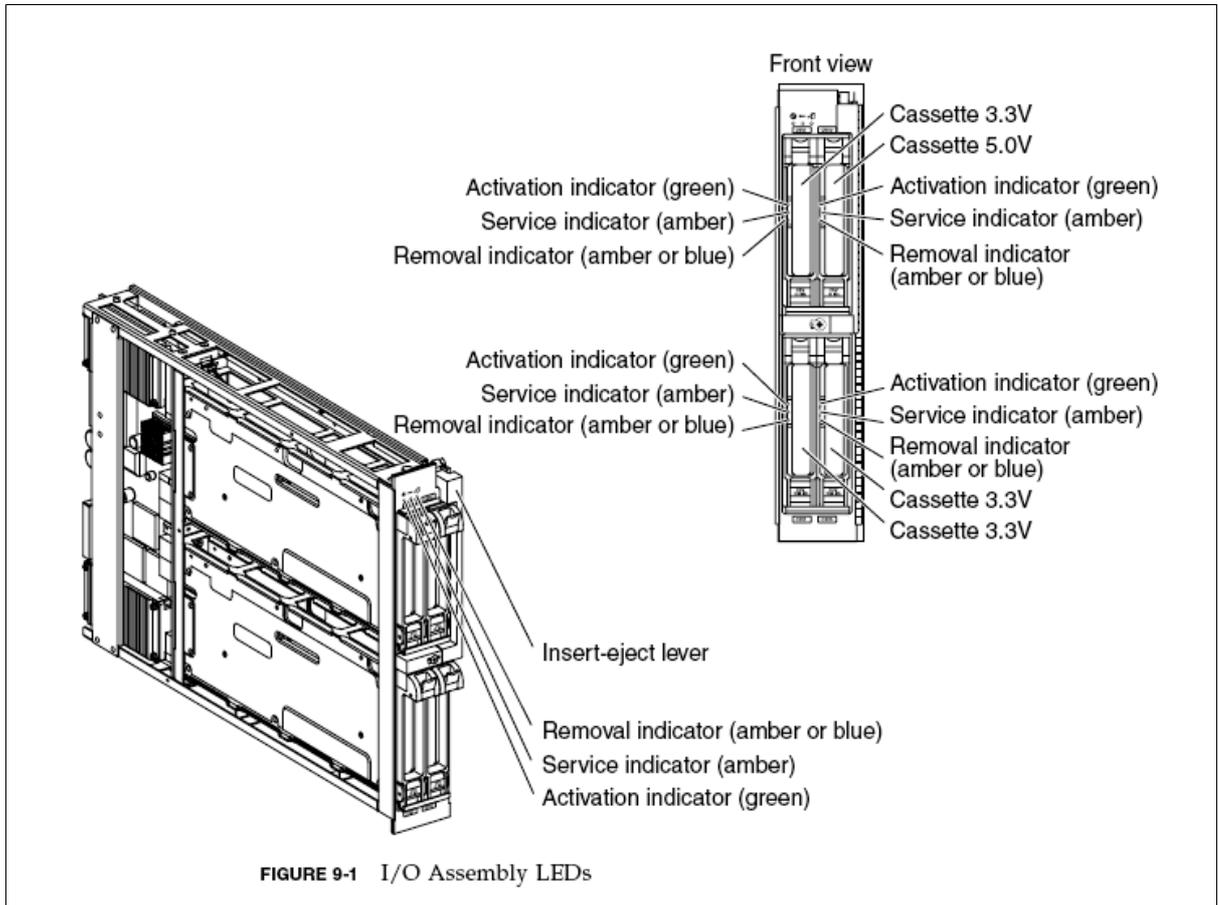


FIGURE 9-1 I/O Assembly LEDs

1.2.8 System Controller

Les *System Controllers* (SC) assurent la configuration et la surveillance de la plate-forme. Ils gèrent le démarrage des domaines, ainsi que la détection de pannes.

Le Sun Fire 25K possède deux cartes *System Controllers* redondantes permettant d'assurer la disponibilité et la reprise sur erreur (1 en face avant et 1 en face arrière). Les *System Controllers* sont nommés SC0 et SC1.

Les cartes *System Controllers* sont des stations embarquées faisant tourner leur propre système d'exploitation Solaris et le logiciel spécifique de gestion de la plate-forme Sun Fire 25K : SMS. (*System Management Services*). Elles permettent en outre d'accéder à la console des domaines du Sun Fire 25K.

A un instant donné, un seul des *System Controllers* assure les fonctions de gestion de la plate-forme. On le nomme *System Controller* principal (*Main*). L'autre est alors le *System Controller* secondaire (*Spare*).

Le *system controller* peut être assimilée à une station autonome constitué d'un processeur UltraSparc Ili (comme dans les Ultra 10) équipé de 1 Gb de mémoire, de 6 bus PCI, d'interfaces série et Ethernet, d'un lecteur CDRom/DVD et d'un DAT DDS4, ainsi que de la logique de commande et de réinitialisation.

Un Solaris 9 standard est embarqué sur 2 disques systèmes configurés en raid1 (miroir) via la produit SVM (Solaris Volume Manager).

Un *system controller* est divisé en quatre parties fonctionnelles:

- Le *controller board*,
- Le *CPU board* (CP1500),
- Le *system controller peripheral board* (il contient les disques, DVD et DAT),
- Le *centerplane support board*.

Liste des principales ressources embarquées:

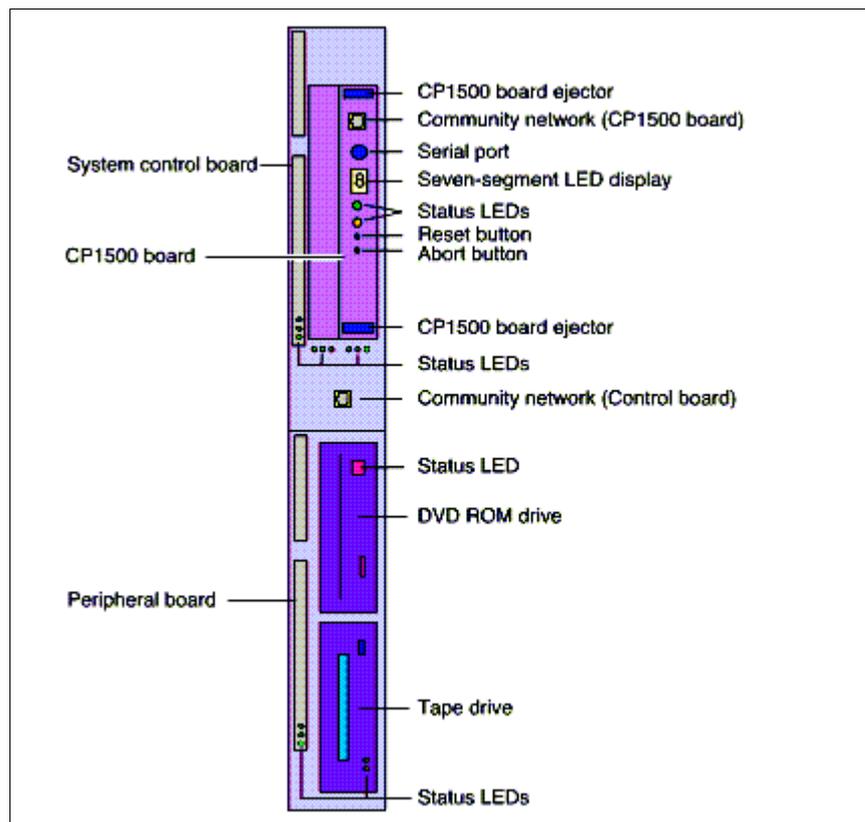
- 13 Interfaces réseau 100BASE TX d'administration,
 - 2 interfaces (réseau C) pour les connexions aux réseaux externes,
 - 9 interfaces (réseau I1) pour les connexions internes SC/domaines (console),
 - 2 interfaces (réseau I2) dédié aux connexions entre les 2 SC.
- Bus console vers chaque system board,
- Port serie,
- Interface I²C vers
 - Modules d'alimentation et de ventilation,
 - Contrôle et gestion des alimentations,
 - Contrôle et gestion de la temperature,
 - Contrôle des *leds* sur chaque élément,
- Routage des consoles de chaque domaine,

- Centralisation des erreurs,
- Contrôle *system controller* opposé,
- Gestion de la *clock*.

Principe de fonctionnement du « *system clock* » sur SF25K:

- Le SunFire 25K utilise un mécanisme de *clock* redondant : la *clock line*.
- La *clock* est générée par chaque SC, si le SC Main se retrouve stoppé, la ligne de *clock* reste valide en prenant comme source le second SC jusqu'au redémarrage du primaire.
- Chaque *system controller* genere une *clock* qui est distribuée séparément aux boards (SB et IO) et au second SC. Chaque board contient la logique nécessaire pour recevoir les informations provenant de la *clock line* et la redistribuer aux *clocks* locales.

Figure 13: System Controller – face avant



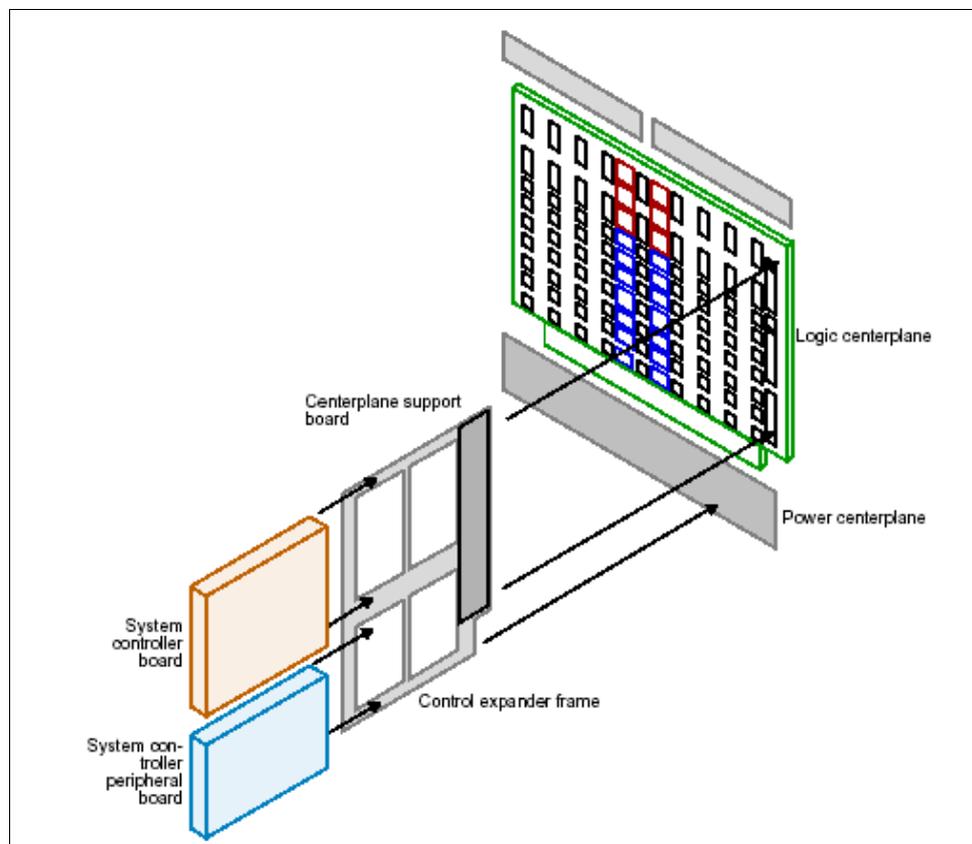
1.2.9 Centerplane board

Les *System Boards* du serveur Sun Fire 25K communiquent entre elles via la *Centerplane board* assurant la fonctionnalité de carte fond de panier.

A contrario des précédentes générations de serveurs architecturées autour de bus passifs, le *Centerplane board* serveur Sun Fire 25K exploite les interconnexions *Gigaplane-XB* combinant adresses actives et routeurs de données pour l'obtention d'une bande passante élevée sur les bus d'adresses et de données.

Le mode de communication inter-cartes obtenu est de type point à point (modèle *Crossbar*). La commutation des signaux est réalisée par les ASICs équipant la carte. Cette technologie autorise le partitionnement en domaines physiquement indépendants les uns des autres.

Figure 14: Vue d'ensemble du centerplane d'un Sun Fire 25K



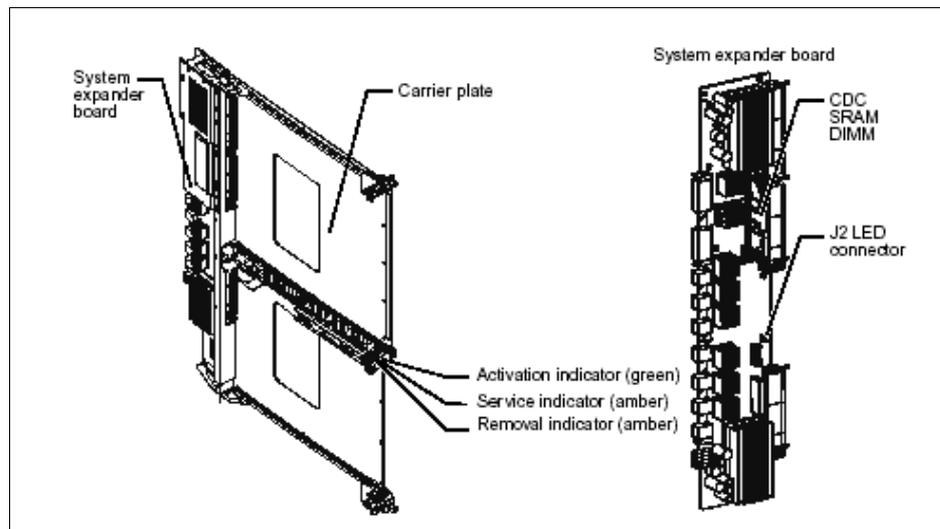
1.2.10 Expander board

Les *Expander boards* du serveur Sun Fire 25K permettent la communication entre le *Centerplane board* et les *System boards* (slot 0) et *IO boards* (slot 1).

Les *expander boards* permettent d'isoler les *system boards* du *centerplane*, augmentant ainsi la sécurité lors d'opération d'insertion ou de retrait dynamique.

Les *expander boards* contiennent un hub ethernet afin d'assurer la redondance du réseau I1 vers les 2 SC.

Figure 15: Vue d'ensemble du système d'expander board



1.2.11 Domaines et réseaux internes

Un domaine correspond à un ensemble de ressources matérielles (*System Boards* et *I/O Boards*) qui sont, à un instant donné, sous le contrôle d'une seule et unique instance du système d'exploitation Solaris.

Un serveur SunFire 25K contient au maximum dix-huit domaines, nommés de A à R. Le SunFire 25K est pourvu de deux réseaux internes précablés permettant les communications entre, d'une part, les *System Controllers* et les domaines et d'autre part entre les *System Controllers* eux-mêmes.

Tableau 2: Nomenclature des réseaux du Sun Fire 25K

Nom du réseau	Description
I1	Réseau accueillant les communications entre les <i>System Controllers</i> et les domaines. Ce réseau est interne et précablé. Le routage est désactivé par défaut sur ce réseau et ne doit pas être activé.
I2	Réseau entre les deux <i>System Controllers</i> . Ce réseau est interne et précablé. Le routage est désactivé par défaut sur ce réseau et ne doit pas être activé.
Cx	Cx est le nom employé pour désigner les réseaux client afin de connecter les deux <i>System Controllers</i> et les domaines.

1.2.11.1 Interfaces réseau

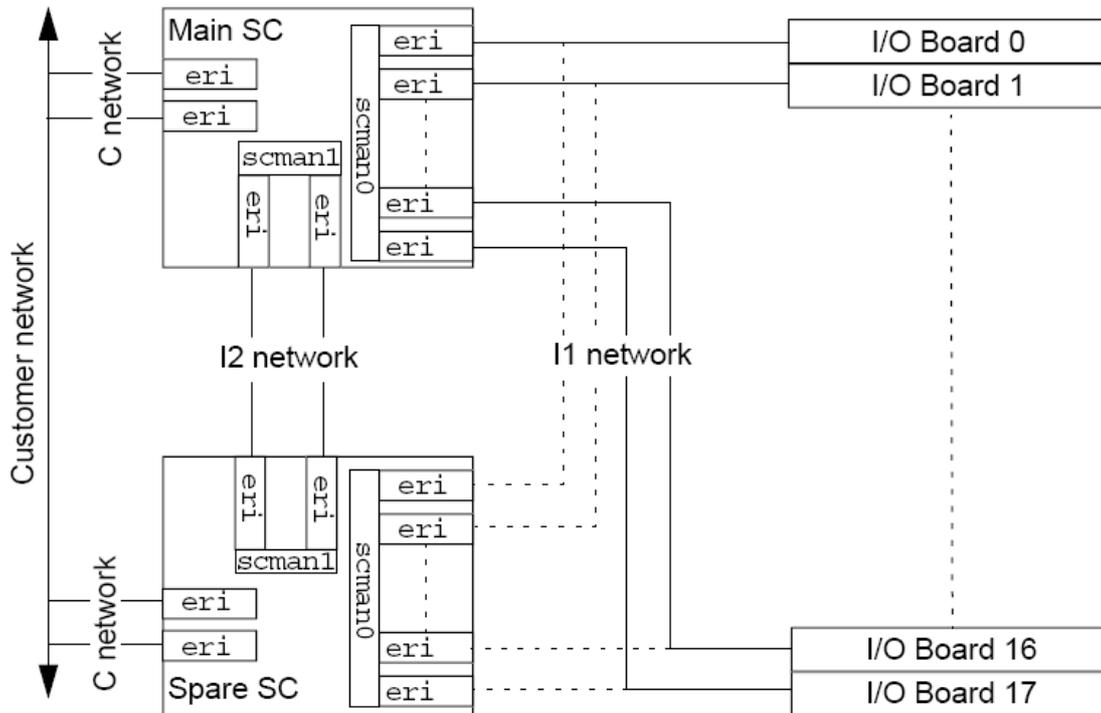
Le tableau suivant présente la nomenclature utilisée pour les différentes interfaces réseau du Sun Fire 25K dans la suite du document.

Tableau 3: Nomenclature des interfaces réseau du Sun Fire 25K

Nom de l'interface	Description
SC-I1	Connexion du <i>System Controller</i> principal sur le réseau I1. Une seule adresse IP est utilisée : seul le SC principal communique sur le réseau I1.
D[A-I]-I1	Connexion des domaines A à I au réseau I1.
SC[0,1]-I2	Connexion des SC0 et SC1 au réseau interne I2.
SC[0,1]-Cx	Connexion des SC0 et SC1 au réseau client.
SC[0,1]-PCx	Adresses « flottantes » pour un groupe d'interfaces redondantes.
L-Cx	Interface logique pour chaque réseau client sur le SC principal.
D[A-I]-Cx	Connexion des domaines A à I au réseau client.

1.2.11.2 Topologie des réseaux interne

Figure 16: Topologie des réseaux internes

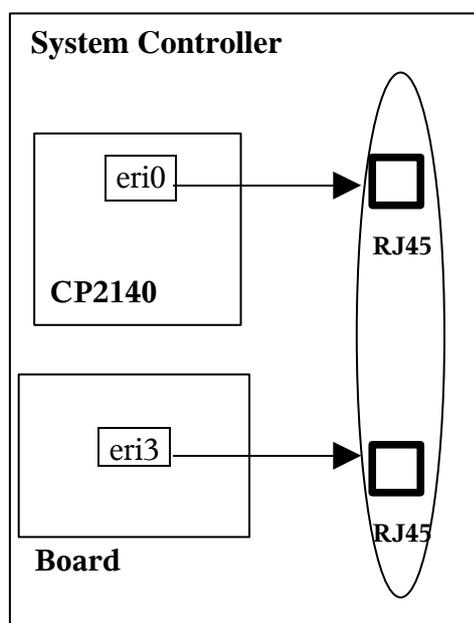


1.1.1.1.1. *Le réseau C*

Les 2 interfaces du réseau C (COMMUNITY) sont dédiées aux communications externes, et sont disponibles pour être intégrées à un réseau client. Ces 2 interfaces sont de type RJ45 et sont nommées *eri0* et *eri3*.

Figure 17: Diagramme du réseau C

Sun Fire 25



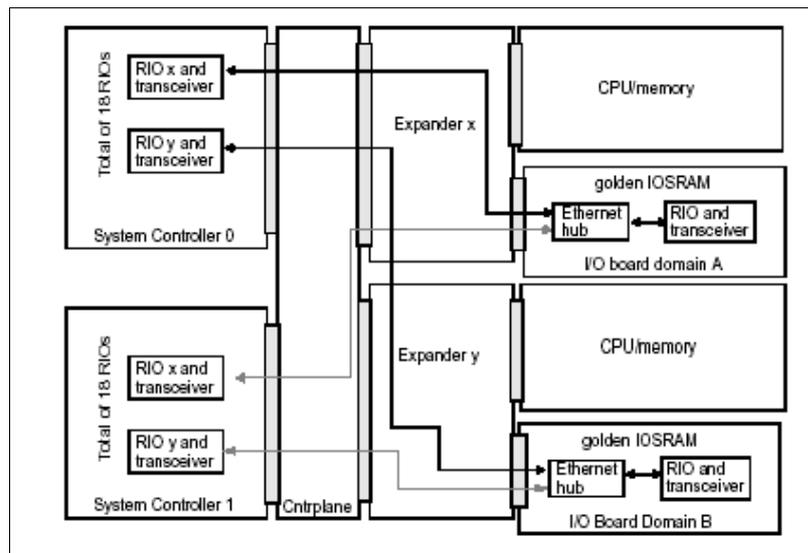
1.1.1.1.2. Le réseau I1

Le réseau I1 est le premier réseau interne au SunFire 25K. Il est utilisé pour les communications et le contrôle des domaines depuis le SC main.

La console de chaque domaine utilise ce réseau.

Les messages syslog et les opérations de DR transitent également par ce réseau. La redondance du réseau I1 vers les *system controllers* est assurée par les hubs embarqués sur chaque *expander board*.

Figure 18: Diagramme du réseau I1



1.1.1.1.3. Le réseau I2

Le réseau I2 comporte les 2 interfaces utilisées par les *system controllers* afin de gérer les mécanismes de failover.

Ce réseau est utilisé pour la synchronisation entre les 2 *system controllers*.

1.2.12 Disques systèmes (Se3120) des domaines.

Chaque domaine dispose de ses propres disques pour embarquer le système.

Les domaines SF25K ont leur système installé sur des disques Ultra SCSI intégrés dans un tiroir rackable de type Se3120.

Un tiroir Se3120 contient 3 disques de 73 Gb. Il est connecté au domaine via deux interfaces Ultra SCSI.

Figure 19: Vue d'ensemble tiroir Se3120



2 Configuration de la plate-forme

2.1 Paramètres nécessaires du Sun Fire 25K

Cette section décrit l'ensemble des paramètres nécessaires à l'installation de la plate-forme Sun Fire 25K hors installation des domaines proprement dits.

Ces paramètres sont les suivants :

- Nom de la plate-forme
- *hostnames*, réseaux, adresses IP et masques de sous-réseaux pour les *System Controllers*
- *hostnames* des domaines

2.1.1 Nom de la plate-forme

Un nom de plate-forme est nécessaire sur le Sun Fire 25K. Ce nom ne doit correspondre à aucun domaine de la plate-forme et à aucune autre machine du réseau.

2.1.2 Configuration des System Controllers

Ce chapitre présente l'ensemble des paramètres utilisés pour configurer les *System Controllers* (ou SC) de votre plate-forme Sun Fire 25K.

2.1.2.1 Configurations réseau

Deux configurations réseau sont possibles : la configuration normale et la configuration haute-disponibilité.

Dans les deux cas :

- Les domaines peuvent être connectés à autant de réseaux différents que nécessaire.
- Les *System Controllers* ne peuvent être connectés qu'à deux réseaux client différents maximum.
- Il n'y a pas de contrainte de connexion des domaines au même réseau que celui ou ceux des *System Controllers*.
- Les deux *System Controllers* doivent avoir des câblages réseau identiques.
- Si le SC principal perd la connexion à un des réseaux client, une bascule a lieu (*SC failover*), affectant tous les réseaux client du SC.
- L'accès à la plate-forme elle-même est hautement disponible. En cas de problème sur le SC principal, une bascule sur le SC secondaire a lieu.

2.1.2.2 Configuration normale

Les adresses IP notées F-CA et F-CB dans ce schéma correspondent aux adresses IP flottantes du *System Controller* principal et sont donc susceptibles de basculer d'un SC à l'autre.

Les adresses IP notées SC0-CA, SC0-CB, SC1-CA et SC1-CB correspondent aux adresses IP fixes des *System Controllers* permettant d'accéder à chacun des réseaux A et B. Elles ne basculent pas avec le SC principal.

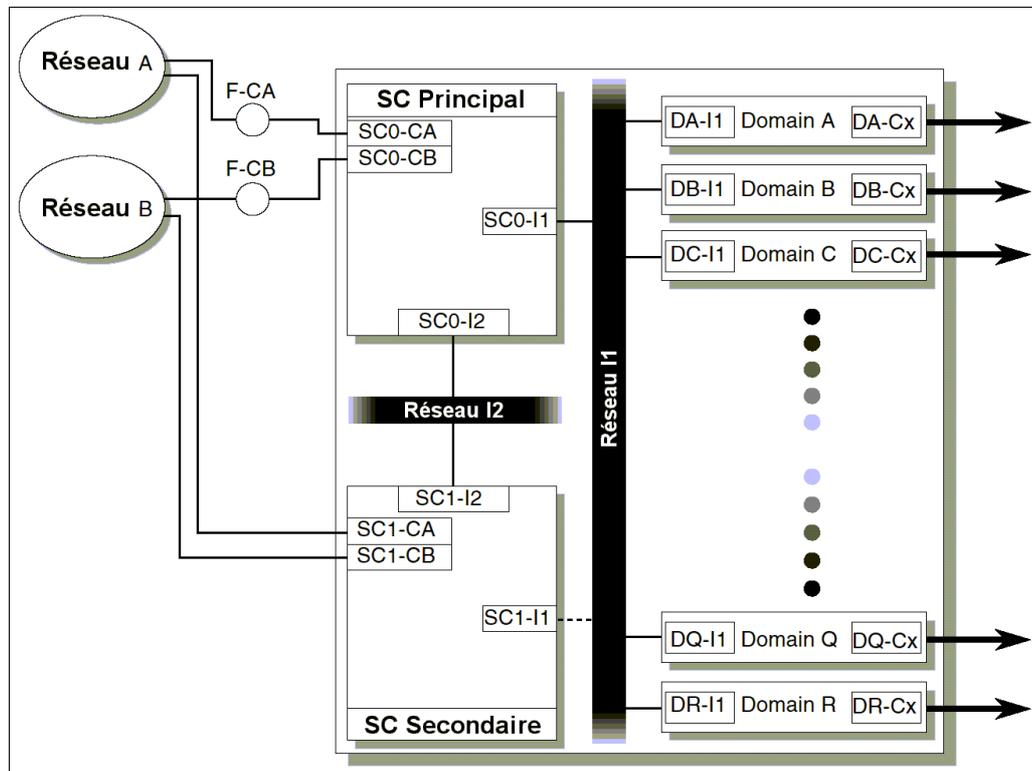
Avantages:

Chacun des SC peut être connecté à deux réseaux différents.

Inconvénients:

Une panne sur un des réseaux provoque un basculement du SC principal.

Figure 18 : Schéma de Configuration Réseau normal



2.1.2.3 Configuration haute-disponibilité

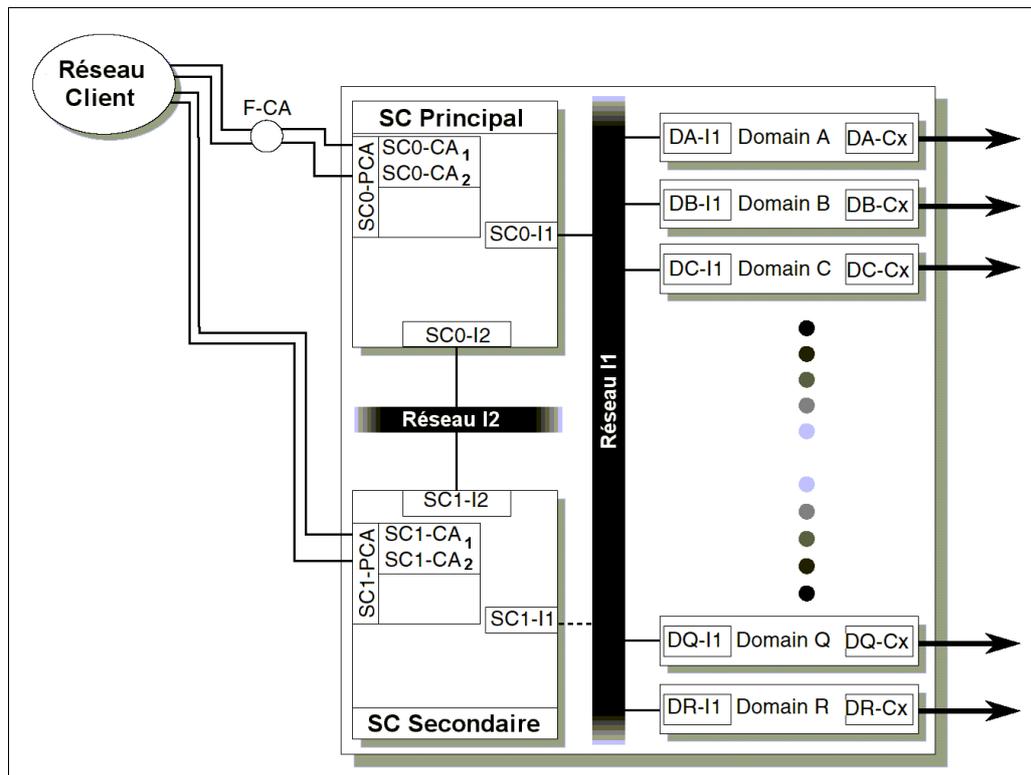
Avantages:

Chacun des SC possède deux connexions redondantes au réseau client.
La panne d'une connexion réseau sur un *System Controller* ne provoque pas la bascule du SC principal.

Inconvénients:

Un seul réseau client peut être connecté aux SC.

Figure 19 : Schéma de Configuration Réseau haute-disponibilité



2.2 Fonction des system controllers (SC)

2.2.1 Administration matérielle du serveur

Toute l'administration matérielle du serveur est réalisée par les *system controllers*. Les SC commandent de façon logicielle la mise sous tension des composants hors fonctionnalités « *House Keeping* », et leur éventuel arrêt électrique. Elle assure la supervision de la température et des alimentations via les données de monitoring. La séquence logique de mise sous tension du serveur est donc la suivante :

- Basculer sur « *On* » les interrupteurs des modules d'entrée 220V_{AC},
- Mettre sous tension et vérifier le *boot* du SC principal,
- Basculer sur « *On* » les disjoncteurs selon l'ordre suivant :
 - *Fan Trays*,
 - *Centerplane Support Boards*,
 - *System Boards*,
 - *Expander boards*.
- Mettre sous tension le système par les commandes en lignes entrées sur le SC principal.

La séquence inverse doit être conduite à l'arrêt du système.

2.2.2 Administration des domaines

Le serveur SunFire 25K associé au logiciel SMS peut supporter jusqu'à dix-huit domaines indépendants, chacun étant une instance Solaris™ différente. Les dix-huit *System Boards* pouvant équiper le serveur doivent alors être réparties entre les domaines. Ces derniers comportent, ad minima, une *System Board* et une *IO board*; celles-ci constituant un ensemble autonome de ressources (microprocesseurs, mémoire, I/O).

L'architecture *Crossbar* du serveur garantit l'indépendance de chaque domaine qui ne « voit physiquement » que ses propres ressources. La notion de domaine dynamique (variable en ressources par attachement ou détachement de *System Boards*) induite par la mise en place du dispositif de reconfiguration dynamique est donc applicable sur ce type d'architecture.

Les SC assurent la gestion des domaines: création, démarrage, arrêt, destruction ou reconfiguration,

2.2.3 Fonctions et limites des SC lors du boot de domaines

Lors du démarrage d'un domaine la présence de la SC *main* est indispensable pour les actions suivantes :

- Mise sous tension des *System Boards*,
- Exécution du test matériel hpost déterminant l'architecture effectivement utilisable pour le domaine,
- Téléchargement de la PROM.

A partir du *boot* Solaris™ le domaine peut fonctionner indépendamment des SC.

2.2.4 Fonctionnalités portées par les SC aux domaines

Les SC font office de :

- Console du domaine :
 - après le *boot* Solaris™ du domaine, via le réseau privé le (mode Ethernet).
 - avant le *boot* ou en cas de perte du réseau, via le réseau privé (mode JTAG).
- *Syslog* : archivage des messages système des domaines.

2.3 Les dispositifs propres au serveur Sun Fire 25K

2.3.1 Basculement de System Controller

Il s'agit de la procédure de passage de l'équipement défaillant sur celui de secours. Le basculement d'un SC est sans impact sur les domaines déjà démarrés.

Un basculement de SC peut être déclenché automatiquement par les SC eux-mêmes en cas de détection d'anomalie ou peut être forcé manuellement.

2.3.2 Reconfiguration dynamique

La reconfiguration dynamique (*Dynamic Reconfiguration - DR*) est une fonction propre au Sun Fire 25K qui offre la possibilité d'ajouter ou de retirer une *System Board* ou *I/O Board* à la volée sans interrompre l'exploitation. Un certain nombre de conditions doivent être respectées pour la mise en œuvre de cette fonction :

- **Conditions système au détachement d'une *System Board* :**
 - la capacité mémoire de la carte détachée doit pouvoir être répartie sur les autres cartes et les zones de *swap*,
 - les drivers des I/O sont classifiés *DR - Safe*.

Conditions générales de mise en œuvre:

- les contraintes de temps des applications sont compatibles avec les opérations de reconfiguration.

3 Procédures système

3.1 Introduction

Ce chapitre est destiné aux administrateurs et pupitreurs de la plate-forme Enterprise SF25K. Il est présumé que les utilisateurs disposent des connaissances minimales requises à l'exploitation de systèmes UNIX[®] sous l'environnement opérationnel Solaris, et notamment savent :

- Se *logger* sous l'identité d'un utilisateur quelconque.
- Ouvrir une fenêtre « terminal » sous CDE.
- Tirer parti des commandes de base de l'éditeur vi
- Identifier si l'on se trouve sous l'environnement Solaris ou sous l'OBP (*OpenBoot™ PROM*).

REMARQUES :

Il est indispensable de savoir distinguer si la commande entrée s'exécute sur le SC, ou sur un domaine sur lequel on est *loggé* via le SC (utilisation de la commande *uname*).

Il est indispensable, pour appliquer ces procédures de disposer des mots de passe de *root* et d'utilisateur *sms-svc* sur chaque domaine et sur les SC.

3.2 Tableau des mots clés

Le tableau ci-dessous dresse la liste des mots clés utilisés dans les procédures. Chaque élément en italique de la première colonne est à remplacer, lors de l'exécution de la procédure, par une valeur choisie dans la troisième colonne.

Tableau 8: Liste des mots clés utilisés dans les procédures

Mots clés	Signification	Valeurs possibles
<i>nom_plate-forme</i>	Nom de la plate-forme au sens du logiciel SMS	
<i>nom_domaine</i>	Nom d'un domaine au sens du logiciel SMS	
<i>liste_carte_systeme</i>	Numéros des <i>System Boards</i> ** composant un domaine	
<i>SC</i>	Nom du <i>system controller</i> principal du serveur SF25K (main SC)	
<i>SC0</i>	Nom (<i>hostname</i>) du premier <i>system controller</i>	
<i>SC1</i>	Nom (<i>hostname</i>) du second <i>system controller</i>	
<i>nom_device_bande</i>	Périphérique de sauvegarde	

3.3 Connexion aux SC.

Objectif: Pouvoir se connecter aux System Controllers

Opérations:

Deux modes de connexions sont possible: en mode console ou en mode telnet

- Connexion à la console de chacun des system controller via la connexion de type série RS232.

Le paramétrage de cette connexion est le suivant: 9600 bps, 8 bits, 1 stop bit, XON/XOFF.

La connexion en mode console n'est véritablement nécessaire que pour visualiser le boot d'un system controller.

Nota: il est possible, depuis un SC, de passer en mode console sur le SC opposé via la commande: *smsconnectsc*. La déconnexion à cette session se faisant par «~.».

```
sms-svc:1> smsconnectsc
Attempting to determine remote SC power state. This may take
some time.
Attempting to get TTY connection state. This may take some
time.
TTY connection is OFF.
About to connect to other SC. Do you want to continue
(yes/no)? yes
Attempting to set TTY connection state. This may take some
time.
connected

console login: root
Password: password

Sourcing //.profile-EIS.....
root@su1234SC1#
```

3.4 Identifier le SC principal (Main SC)

Objectif: Pouvoir identifier le rôle des SC.

Opérations:

Le rôle de chaque SC est donné via la commande *showfailover(1m)*,

- Sur le System Controller principal:

```
su1234SC0:sms-svc:> showfailover -r  
MAIN
```

- Sur le System Controller secondaire:

```
su1234SC1:sms-svc:> showfailover -r  
SPARE
```

3.5 Activer le mode failover

Objectif: Savoir activer la reprise automatique de la fonction System controller

Opérations: Sur le System Controller principal

- Visualiser l'état du mécanisme de *failover*.

```
sms-svc:> showfailover -v
SC Failover Status:      DISABLED
sms-svc:>
```

- Activer le mécanisme de *failover*.

```
sms-svc:> setfailover on
```

- Visualiser à nouveau l'état du mécanisme de *failover*

```
su1234SC0:sms-svc:> showfailover -v
SC Failover Status:      ACTIVE
Status of Shared Memory:
HASRAM (CSB at CS0):
.....Good
HASRAM (CSB at CS1):
.....Good

Status of su1234SC0:
Role: .....MAIN
SMS Daemons: .....Good
System Clock: .....Good
Private I2 Network: .....Good
Private HASRAM Network: .....Good
Public Network:
  Group "C1":
.....Up
  eri0:
.....Up
  eri3:
.....Up
  Logical IP Addr. -
C1:.....Up
System Memory:
.....4.9%
Disk Status:
  /:
.....2.3%
Console Bus Status:
  EXB at EX0:
.....Good
  EXB at EX1:
.....Good
  EXB at EX2:
.....Good
```

```

EXB at EX3:
.....Good
EXB at EX4:
.....Good
EXB at EX5:
.....Good
EXB at EX6:
.....Good
EXB at EX7:
.....Good
EXB at EX8:
.....Good

Status of sul234SC1:
Role: .....SPARE
SMS Daemons:
.....Good
System Clock:
.....Good
Private I2 Network:
.....Good
Private HASRAM Network:
.....Good
Public Network:
Group "C1":
.....Up
eri0:
.....Up
eri3:
.....Up
LogicalIPAddr.C1:.....Inactiv
e
System Memory:
.....4.6%
Disk Status:
/:
.....2.2%
Console Bus Status:
EXB at EX0:
.....Good
EXB at EX1:
.....Good
EXB at EX2:
.....Good
EXB at EX3:
.....Good
EXB at EX4:
.....Good
EXB at EX5:
.....Good
EXB at EX6:
.....Good
EXB at EX7:
.....Good
EXB at EX8:
.....Good
sul234SC0:sms-svc:>

```

3.6 Basculement du SC secondaire en SC principal

Objectif: Utiliser le SC secondaire en SC principal si ce dernier est indisponible ou si un arrêt est planifié.

- Appliquer la procédure «Activer le mode failover»
- Forcer le basculement de SC à l'aide de la commande suivante:

```
sms-svc> setfailover force
```

- Attendre 1minute. Le *System Controller* principal redémarre lors du *failover*
- Se connecter sur le *System Controller* qui était principal avant le *failover* comme administrateur ou opérateur de plate-forme et vérifier qu'il est bien secondaire (*SPARE*) en exécutant la commande suivante :

```
sms-svc> showfailover -r
```

- Le résultat de la commande précédente doit être **SPARE**
- Se connecter au *System Controller* principal (précédemment secondaire) comme administrateur ou opérateur de plate-forme et exécuter la commande suivante :

```
sms-svc> showfailover -r
```

- Le résultat de la commande précédente doit être **MAIN**.
- Se connecter sur chacun des domaines et exécuter la commande *uptime* pour vérifier que les domaines tournent toujours et qu'aucun n'a redémarré.
- Se connecter sur le *System Controller* principal comme administrateur ou opérateur de plate-forme et exécuter la commande :

```
sms-svc> showfailover -v
```

- Vérifier l'état de la clock (la ligne *system clock* doit être à *good*)
- Si le failover n'est pas activé, exécuter la commande suivante :

```
sms-svc> setfailover on
```

3.7 Les Process actifs sur chaque SC.

Objectif: identifier les démons spécifiques aux SC lancés par SMS lors du boot ou en cas de basculement.

- Liste des processus spécifiques sur le System Controller Spare:

```

root@sul234SC1# ps -ef
      UID    PID  PPID  C   STIME TTY      TIME CMD
      root    651     1   0 12:04:15 ?          0:00 ssd -i SMS 1.5
start-up initiated -iSC POST results: 'CP2140 cPOST Passed;
SSC
      root    628     1   0 12:04:12 ?          0:03 mld
      root   3872    512   0 11:56:55 ?          0:00 sleep 3
      root    668     1   0 12:04:15 ?          0:48 hwad
      root    675     1   0 12:04:16 ?          0:00 mand
sms-frad    684     1   0 12:04:16 ?          0:00 frad
      root    687     1   0 12:04:17 ?          9:56 fomd

root@sul234SC1#

```

- Liste des processus spécifiques sur le System Controller Main:

```

root@sul234SC0 # ps -ef
      UID    PID  PPID  C   STIME TTY      TIME CMD
 sms-tmd   1041     1   0   Sep 07 ?          0:44 tmd -t 12
 sms-dca   3865    650   0   Sep 13 ?          0:00 dca -d A
      root    650     1   0   Sep 07 ?          0:04 ssd -i SMS 1.5
start-up initiated -iSC POST results: 'CP2140 cPOST Passed;
SSC
      root    686     1   0   Sep 07 ?        198:54 fomd
sms-esmd   1055     1   0   Sep 07 ?          83:45 esmd
      root    667     1   0   Sep 07 ?          420:54 hwad
      root    670     1   0   Sep 07 ?           1:25 mand
sms-dsmd   1052     1   0   Sep 07 ?          12:19 dsmd
 sms-pcd   1029     1   0   Sep 07 ?           5:43 pcd
 sms-frad   683     1   0   Sep 07 ?           6:18 frad
 sms-svc   29368  29257   0 14:36:36 pts/7    0:00 console -d E
 sms-osd   1064     1   0   Sep 07 ?           0:01 osd
 sms-codd   1058     1   0   Sep 07 ?           0:01 codd
 sms-dca   27275   650   0   Sep 15 ?           0:00 dca -d E
      root   3610    160   0 14:54:12 ?           0:00 in.telnetd
 sms-efe   1226     1   0   Sep 07 ?           0:01 efe
 sms-efhd  1399     1   0   Sep 07 ?           1:44 efhd
      root   1202     1   0   Sep 07 ?           1:35 kmd
      root  26137  26131   0 14:23:43 pts/10    0:00 -sh
 sms-svc   29367  29258   0 14:36:34 pts/9    0:00 console -d D
      root  17021     1   0   Sep 08 ?           0:00 devfsadmd
 sms-svc   29340  29261   0 14:36:29 pts/10    0:00 console -d A

```

```

sms-elad 1410 1 0 Sep 07 ? 0:00 elad
sms-dxs 3294 650 0 Sep 13 ? 1:50 dxs -d B
sms-dxs 1567 650 0 Sep 07 ? 3:23 dxs -d D
sms-dca 1566 650 0 Sep 07 ? 0:00 dca -d D
sms-svc 9673 9557 0 09:19:50 pts/4 0:00 -csh
sms-erd 1403 1 0 Sep 07 ? 0:00 erd
sms-dca 1568 650 0 Sep 07 ? 0:00 dca -d C
 sms-svc 26333 26285 0 11:20:47 ? 0:00
/usr/lib/ssh/sshd
sms-dxs 1570 650 0 Sep 07 ? 3:53 dxs -d C
sms-dca 27285 650 0 Sep 15 ? 0:00 dca -d F
sms-dca 3296 650 0 Sep 13 ? 0:00 dca -d B
sms-svc 29354 29259 0 14:36:32 pts/11 0:00 console -d C
sms-dxs 27272 650 0 Sep 15 ? 1:22 dxs -d E
sms-dxs 27283 650 0 Sep 15 ? 1:18 dxs -d F
sms-svc 29381 29256 0 14:36:37 pts/12 0:00 console -d F
sms-dxs 3863 650 0 Sep 13 ? 2:04 dxs -d A
sms-svc 29341 29260 0 14:36:30 pts/5 0:00 console -d B
root@sul234SC0 #

```

3.8 Arrêt d'un domaine

Objectif: Arrêter logiquement un domaine.

Opérations:

- Se logger sur le main SC

```
telnet su1234SC
Trying 10.151.22.13...
Connected to su1234SC.
Escape character is '^]'.

SunOS 5.9
```

```
login: root
```

```
Password: password
```

- Exécuter les commandes suivantes pour se *logger root* sur le domaine :

```
root@su1234SC0 # su - sms-svc
su1234SC0:sms-svc:1> console -d nom_domaine
nom_domaine console login: root
password: password
```

- Saisir la commande suivante :

```
# init 0
```

- Attendre l'apparition du prompt *ok*

```
{0} ok
```

3.9 Mise hors tension du SunFire 25K

Objectif: Arrêter électriquement le serveur.

Opérations:

- S'assurer que la procédure «Arrêt d'un domaine» a été appliquée tour à tour sur chaque domaine puis saisir les commandes suivantes sur le SC principal pour stopper électriquement chaque domaine :

```
sms-svc:> setkeyswitch -d <domain_ID> off
Current virtual key switch position is "ON".
Are you sure you want to change to the "OFF" position
(yes/no)? y
Domain is down.
Waiting on exclusive access to EXB(s): 00010.
Powering off: HPCI at IO0
Resetting and deconfiguring: CPU at SB0
Powering off: CPU at SB1
Powering off: HPCI at IO1
Powering off: EXB at EX1
sms-svc:>
```

- L'étape suivante consiste à stopper les 2 System Controller en débutant par le Spare. Se logger root sur les SC:

```
SCconsolelogin:root
password : password
```

- Saisir la commande suivante:

```
# init 0
```

- En connexion sur le port série des system controlers, attendre l'apparition du *prompt* <#>ok
- Positionner les interrupteurs des 12 modules d'entrée 220V_{AC} PS[0-5] AC[0-1] sur "off"

3.10 Mise sous tension du SF25K

Objectif: Démarrer électriquement le serveur.

Opérations:

- Positionner les interrupteurs des 12 modules d'entrée 220V_{AC} vers le haut en position «On». Cette première phase de la mise sous tension initialise automatiquement le boot des 2 SC.
- Le cas échéant, positionner les disjoncteurs sur «on», selon l'ordre suivant: FT, SC, SCPER, CS et EX
- Attendre que les 2 SC aient correctement démarrés.
- Vérifier que les systèmes d'alimentation et les systèmes de ventilation soient opérationnels en saisissant les commandes suivantes sur le SC principal:

```
sms-svc:> showenvironment -p fans
FANTRAY  POWER  SPEED  FAN0  FAN1  FAN2  FAN3  FAN4  FAN5
FAN6
-----  -
FT0      ON      NORMAL OK    OK    OK    OK    OK    OK
FT1      ON      NORMAL OK    OK    OK    OK    OK    OK
FT2      ON      NORMAL OK    OK    OK    OK    OK    OK
FT3      ON      NORMAL OK    OK    OK    OK    OK    OK
FT4      ON      NORMAL OK    OK    OK    OK    OK    OK
FT5      ON      NORMAL OK    OK    OK    OK    OK    OK
FT6      ON      NORMAL OK    OK    OK    OK    OK    OK
FT7      ON      NORMAL OK    OK    OK    OK    OK    OK
sms-svc:19> showenvironment -p powers
POWER          UNIT  AC0  AC1  DC0  DC1  FAN0
FAN1
-----  -
PS-A196 at PS0 OK    OK    OK    ON    ON    OK
OK
PS-A196 at PS1 OK    OK    OK    ON    ON    OK
OK
PS-A196 at PS2 OK    OK    OK    ON    ON    OK
OK
PS-A196 at PS3 OK    OK    OK    ON    ON    OK
OK
PS-A196 at PS4 OK    OK    OK    ON    ON    OK
OK
PS-A196 at PS5 OK    OK    OK    ON    ON    OK
OK

POWER          VALUE  UNIT  STATUS
-----  -
PS-A196 at PS0
Current0      23.90  A     N/A
Current1      23.90  A     N/A
48VDC         49.40  V     N/A
Power         2361.32 W     N/A
PS-A196 at PS1
Current0      25.10  A     N/A
Current1      25.10  A     N/A
48VDC         49.40  V     N/A
```

<i>Power</i>	2479.88	W	N/A
<i>PS-A196 at PS2</i>			
<i>Current0</i>	24.70	A	N/A
<i>Current1</i>	24.30	A	N/A
<i>48VDC</i>			
<i>Power</i>	49.20	V	N/A
<i>Power</i>	2410.80	W	N/A
<i>PS-A196 at PS3</i>			
<i>Current0</i>	24.30	A	N/A
<i>Current1</i>	23.90	A	N/A
<i>48VDC</i>			
<i>Power</i>	49.00	V	N/A
<i>Power</i>	2361.80	W	N/A
<i>PS-A196 at PS4</i>			
<i>Current0</i>	23.10	A	N/A
<i>Current1</i>	23.90	A	N/A
<i>48VDC</i>			
<i>Power</i>	49.00	V	N/A
<i>Power</i>	2303.00	W	N/A
<i>PS-A196 at PS5</i>			
<i>Current0</i>	24.30	A	N/A
<i>Current1</i>	25.10	A	N/A
<i>48VDC</i>			
<i>Power</i>	49.20	V	N/A
<i>Power</i>	2430.48	W	N/A
<i>Total Power</i>	14347.28	W	N/A
<i>sms-svc:></i>			

3.11 Activation d'un domaine

Objectif: Exécuter un `setkeyswitch` afin de réaliser un auto-test à la mise sous tension (POST) garantissant que le matériel d'un domaine est prêt à être utilisé; charger l'*OpenBoot™ PROM* (OBP).

Opérations:

Chaque domaine est testé par défaut avec un `hpost` au niveau 16.

- La modification du niveau de `hpost` se réalise via le fichier `$/SMSETC/config/domain_id/.postrc` en positionnant la variable "level" au niveau souhaité.

Lors de la modification du fichier `.postrc` s'assurer que les droits et permissions de ce fichier soient bien préservés.

```
-rw-r--r-- 1 sms-svc staff    9 Aug 9 07:46 /etc/opt/SUNWSMS/config/A/.postrc
```

- Les `hpost` sont exécutés lors du démarrage à froid du domaine via la commande:

```
sms-svc:> setkeyswitch -d <domain_ID> on
Powering on: CSB at CS0
  Already powered on: CSB at CS0
Powering on: CSB at CS1
  Already powered on: CSB at CS1
Powering on: EXB at EX1
  Already powered on: EXB at EX1
Powering on: HPCI at IO1
  Already powered on: HPCI at IO1
Powering on: CPU at SB1
  Already powered on: CPU at SB1

Significant contents of .postrc (platform)
  /etc/opt/SUNWSMS/SMS1.5/config/platform/.postrc:
# ident      "@(#)postrc 1.1  01/04/02 SMI"

Reading      domain          blacklist          file
/etc/opt/SUNWSMS/config/A/blacklist ...
# ident      "@(#)blacklist  1.1  01/04/02 SMI"
Reading      platform          blacklist          file
/etc/opt/SUNWSMS/config/platform/blacklist ...
# ident      "@(#)blacklist  1.1  01/04/02 SMI"
port        6.0.3             # No COD
port        6.0.1             # No COD
port        4.0.1             # No COD
port        4.0.3             # No COD
port        5.0.1             # No COD
port        5.0.3             # No COD
SEEPROM probe took 0 seconds.
Reading Component Health Status (CHS) information ...
stage_lport_reset: Assert reset to IOC ports in -Q mode...
stage_lport_reset(): Not -Q mode; Skipping Stage_lport_reset
stage_bus_probe: Check in-use bus configurations...
stage_asic_probe: ASIC probe and JTAG/CBus integrity test...
WARNING: Component ID is up-version: Actual: 426980FD   for
Man Ether Hub IO00

Assuming equivalent to: 326980FD
```

```
WARNING: Component ID is up-version: Actual: 426980FD for
Man Ether Hub IO01
          Assuming equivalent to: 326980FD
stage brd_rev_eval: Board Revision Evaluation and
Compliance...
stage cpu_probe: CPU Module probe...
stage cdc_probe: CDC DIMM probe...
stage mem_probe: Memory dimm probe...
stage adapter_probe: I/O adapter probe...
stage cp_shorts: Centerplane Shorts...
stage lbist: Logic BIST...
stage ibist: Interconnect BIST...
stage field_ict: Field Interconnect Tests...
stage mbist1: Internal memory BIST...
stage mbist2: External memory BIST...
stage domain_sync: Domain sync test...
stage cbus_bbsram: Console Bus test of bootbus sram...
stage sc_interrupt: DARB to SC interrupt...
stage cdc_clear: CDC DIMM clear...
stage cpu_lpost: Test all L1 CPU boards...
Performing ASIC config with bus config a/d/r = 333...
    Slot0 in domain: 00003
    Slot1 in domain: 00003
    EXBs in use: 0007C
sgcpu.flash file: NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
Fprom SB0/F0: NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
Fprom SB0/F1: NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
Fprom SB1/F0: NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
Fprom SB1/F1: NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
stage nmb_cpu_lpost: Non-Mem Board Proc tests...
Performing ASIC config with bus config a/d/r = 333...
    Slot0 in domain: 00003
    Slot1 in domain: 00003
    EXBs in use: 0007C
stage_cpu_lpost(): No NMB Boards in config. Skipping Stage
nmb_cpu_lpost.
Acquiring licenses for all good processors...
Proc SB01/P2 deconfigured: no license available.
Proc SB01/P3 deconfigured: no license available.
Proc SB01/P1 deconfigured: no license available.
stage wib_lpost: Wildcat interface board tests...
stage_wib_lpost(): No good Wcis; Skipping Stage wib_lpost
stage pci_lpost: Test all L1 I/O boards...
Performing ASIC config with bus config a/d/r = 333...
    Slot0 in domain: 00003
    Slot1 in domain: 00003
    EXBs in use: 0007C
```

```

pcilpost.elf NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
stage exp_lpost: Domain-level board and system tests...
explpost.elf NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
stage cpu_lpost_II: CPU L1 domain/system tests...
sgcpu.flash file: NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
Fprom SB0/F0: NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
Fprom SB0/F1: NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
Fprom SB1/F0: NOTE: lpost_vercheck(): Using up-rev LPOST
version 5.19.6 Build 1.0 I/F 12 (from: 5.19.0-5 Build 0 I/F
12)
stage pci_lpost_Q: Init all L1 I/O boards under -Q...
stage cpu_lpost_II_Q: CPU L1 domain/system init under -Q...
stage final_config: Final configuration...
Creating CPU SRAM handoff structures...
Creating GCD IOSRAM handoff structures in Slot IO0...
Writing domain information to PCD...

Key to resource status value codes:
?=Unknown      p=Present      c=Crunched      _=Undefined
m=Missing      o=FailedOBP    f=Failed        b=Blacklisted
i=Misconfig    r=Redlisted
x=NotInDomain  u=G,unconfig   P=Passed        ==G,lockstep
l=NoLicense
e=EmptyCasstt

CPU_Brds:  PortCore
           3 2 1 0  Mem P/B: 3/1 3/0  2/1 2/0  1/1 1/0  0/1
0/0
Slot  Gen  10101010  /L: 10  10  10  10  10  10  10
10    CDC
SB00:  P  P P P P P P P  PP  PP  PP  PP  PP  PP  PP
PP    P
SB01:  P  1 1 1 1 1 1 P P  CC  CC  CC  CC  CC  CC  PP
PP    P

I/O_Brds:      IOC  P1/Bus/Adapt  IOC  P0/Bus/Adapt
Slot  Gen  Type  P1  B1/10 B0/10  P0  B1/eb10 B0/10
(e=ENet, b=BBC)
IO00:  P  hsPCI+  P  p _p p _p  P  p PP_e p _p
IO01:  P  hsPCI+  P  p _p p _e  P  p PP_e p _p

Configured in 333 with 5 procs, 20.000 GBytes, 5 IO adapters.
Interconnect frequency is 149.996 MHz, Measured.
Golden sram is on Slot IO0.
POST (level=16, verbose=20) execution time 8:37

sms-svc:>

```

Le cas échéant, une fois le domaine testé, repositionner le niveau de hpost à sa valeur par défaut de 16.

- Les messages générés par hpost sont loggés dans le répertoire `$SMSVAR/adm/<domaine-name>`. Pour le domaine A:

```
su1234SC0:sms-svc:10> cd $SMSVAR/Adm/A
su1234SC0:sms-svc:10> more post060912.1956.20.log
```

où 030403.1956 signifie 12 Septembre 2006 à 19:56.

3.12 Boot d'un domaine

Objectif: Obtenir le *prompt* de *login* d'un domaine.

Opérations:

- S'assurer que la procédure «*activation d'un domaine*» a été appliquée.
- Saisir les commandes suivantes sur le SC principal:

```
sms-svc> console -d <domain_ID>
```

- A l'apparition du *prompt* ok saisir la commande:

```
Ok boot
```

3.13 Récupération du crash dump

Objectif: Transmettre à la *hotline* SUN les images résultant du *panic* d'un domaine.

Opérations:

- Se *logger root* sur le domaine (à partir du SC principal) :

```
% rlogin nom_domaine
```

- Aller dans le répertoire */var/crash/nom_domaine* :

```
# cd /var/crash/nom_domaine
```

- Transférer le contenu du répertoire sur le portail de la hotline Sun Microsystems en suivant les instructions fournies par l'ingénieur du support.

3.14 Gestion des jetons COD (Capacity On Demand)

Objectif : Etre capable à tout instant de connaître l'état des jetons ou RTU (Right to Use) sur la plate-forme et de les affecter aux différents domaines

Opérations:

- Relever l'état des RTU à l'aide de la commande *codusage*.

Dans le relevé suivant:

- 10 RTU ont été affectés à la plate-forme,
- 8 RTU sont actuellement utilisés,
- 2 RTU sont disponible (le domaine G est à l'arrêt).

```
E25KLIX-SC0:sms-svc> showcodusage
Resource:
=====
Resource      In Use   Installed  Licensed  Status
-----
PROC          8        28         10      OK: 2 available

Domains:
=====
Domain/Resource  In Use  Installed  Reserved
-----
A - PROC         1        4          1
B - PROC         1        4          1
C - PROC         1        4          1
D - PROC         2        4          2
E - PROC         2        4          2
F - PROC         1        4          1
G - PROC         0        4          2
H - PROC         0        0          0
I - PROC         0        0          0
J - PROC         0        0          0
K - PROC         0        0          0
L - PROC         0        0          0
M - PROC         0        0          0
N - PROC         0        0          0
O - PROC         0        0          0
P - PROC         0        0          0
Q - PROC         0        0          0
R - PROC         0        0          0
Unused - PROC    0        0          0

E25KLIX-SC0:sms-svc>
```

- Modifier l'affectation des RTU par domaine à l'aide de la commande *setupplatform*.

Dans le relevé suivant:

- 1 RTU est retiré au domaine G (BOMBYXI)

```
E25KLIX-SC0:sms-svc> setupplatform -p cod
PROC RTUs installed : 10
PROC Headroom Quantity (0 to disable, 8 MAX) [0]?
```

```

PROC RTUs reserved for domain SIDEFPP (1 MAX) [1]?
PROC RTUs reserved for domain ILOGPP (1 MAX) [1]?
PROC RTUs reserved for domain ANAXOSPP (1 MAX) [1]?
PROC RTUs reserved for domain BOMBYXPP (2 MAX) [2]?
PROC RTUs reserved for domain TIMRATPP (2 MAX) [2]?
PROC RTUs reserved for domain ANAXOSI (1 MAX) [1]?
PROC RTUs reserved for domain BOMBYXI (2 MAX) [2]? 1
PROC RTUs reserved for domain H (1 MAX) [0]?
PROC RTUs reserved for domain I (1 MAX) [0]?
PROC RTUs reserved for domain J (1 MAX) [0]?
PROC RTUs reserved for domain K (1 MAX) [0]?
PROC RTUs reserved for domain L (1 MAX) [0]?
PROC RTUs reserved for domain M (1 MAX) [0]?
PROC RTUs reserved for domain N (1 MAX) [0]?
PROC RTUs reserved for domain O (1 MAX) [0]?
PROC RTUs reserved for domain P (1 MAX) [0]?
PROC RTUs reserved for domain Q (1 MAX) [0]?
PROC RTUs reserved for domain R (1 MAX) [0]?
E25KLIX-SC0:sms-svc:9>

```

Dans le relevé suivant:

- 1 RTU est ajouté au domaine B (ILOGPP)

```

E25KLIX-SC0:sms-svc> setupplatform -p cod
PROC RTUs installed : 10
PROC Headroom Quantity (0 to disable, 8 MAX) [0]?
PROC RTUs reserved for domain SIDEFPP (2 MAX) [1]?
PROC RTUs reserved for domain ILOGPP (2 MAX) [1]? 2
PROC RTUs reserved for domain ANAXOSPP (1 MAX) [1]?
PROC RTUs reserved for domain BOMBYXPP (2 MAX) [2]?
PROC RTUs reserved for domain TIMRATPP (2 MAX) [2]?
PROC RTUs reserved for domain ANAXOSI (1 MAX) [1]?
PROC RTUs reserved for domain BOMBYXI (1 MAX) [1]?
PROC RTUs reserved for domain H (0 MAX) [0]?
PROC RTUs reserved for domain I (0 MAX) [0]?
PROC RTUs reserved for domain J (0 MAX) [0]?
PROC RTUs reserved for domain K (0 MAX) [0]?
PROC RTUs reserved for domain L (0 MAX) [0]?
PROC RTUs reserved for domain M (0 MAX) [0]?
PROC RTUs reserved for domain N (0 MAX) [0]?
PROC RTUs reserved for domain O (0 MAX) [0]?
PROC RTUs reserved for domain P (0 MAX) [0]?
PROC RTUs reserved for domain Q (0 MAX) [0]?
PROC RTUs reserved for domain R (0 MAX) [0]?
E25KLIX-SC0:sms-svc>

```

Remarque:

Pour activer une nouvelle CPU gérée par COD sur un domaine actif et après affectation du RTU supplémentaire à l'aide de la commande *setupplatform -p*, un arrêt/redémarrage complet du domaine (*setkeyswitch off/on*) est nécessaire.

3.15 Reconfiguration Dynamique (DR): Détachement d'une System Board d'un domaine actif

Objectif: Retirer dynamiquement une ressource de type SB (system board) d'un domaine actif.

Opérations:

Nota sur les opérations de détachement: la quantité de ressource (CPU et mémoire) et le type de ressource (mémoire système, ...) peut influencer de façon considérable sur les temps de détachement d'une SB. De même certains processus peuvent être associé (*bind*) à un processeur. Il convient alors d'identifier le ou les processus impliqués et de les stopper ou de couper leur association au processeur impactés par la DR (voir la commande *pbind*). C'est le cas du mécanisme de synchronisation du temps *ntp*.

- Utiliser la commande suivante pour savoir si le board choisi héberge la mémoire permanente.

```

sul234SC0:sms-svc> rcfgadm -d a -av |grep permanent
SB1::memory    connected    configured    ok    base address
0x2000000000, 4194304 KBytes total, 1736904 KBytes permanent
sul234SC0:sms-svc>

```

- La commande équivalente `cfgadm -av |grep memory`` peut être exécutée depuis le domaine concerné sous l'utilisateur *root* et donne le même résultat.
- L'exemple présenté ci-après montre une opération de détachement sur la SB1. **Attention: le détachement de la carte système sur laquelle réside la mémoire permanente entraîne un blocage temporaire du domaine.**
- Visualiser les CPUs présentes sur le domaine (5 CPUs dans l'exemple suivant):

```

root@E25K-SIDEF01 # prtdiag
System Configuration: Sun Microsystems sun4u Sun Fire E25K
System clock frequency: 150 MHz
Memory size: 20480 Megabytes

===== CPUs =====

Slot ID      CPU      Run      E$      CPU      CPU
-----      -
/SB00/P0     0, 4    1500    32.0    US-IV+   2.2
/SB00/P1     1, 5    1500    32.0    US-IV+   2.2
/SB00/P2     2, 6    1500    32.0    US-IV+   2.2
/SB00/P3     3, 7    1500    32.0    US-IV+   2.2
/SB01/P0    32, 36   1500    32.0    US-IV+   2.2

root@E25K-SIDEF01 #

```

- Retirer la SB du domaine depuis le system controller main:

```

sul234SC0:sms-svc> deleteboard sb1
System may be temporarily suspended, proceed
(yes/no)? yes
request suspend SUNW_OS
request suspend SUNW_OS done

```

```
request delete capacity (2 cpus)
request delete capacity (524288 pages)
request delete capacity SB1 done
request offline SUNW_cpu/cpu32
request offline SUNW_cpu/cpu36
request offline SUNW_cpu/cpu32 done
request offline SUNW_cpu/cpu36 done
unconfigure SB1
unconfigure SB1 done
notify remove SUNW_cpu/cpu32
notify remove SUNW_cpu/cpu36
notify remove SUNW_cpu/cpu32 done
notify remove SUNW_cpu/cpu36 done
notify capacity change (2 cpus)
notify capacity change (524288 pages)
notify capacity change SB1 done
disconnect SB1
disconnect SB1 done
poweroff SB1
poweroff SB1 done
notify resume SUNW_OS
notify resume SUNW_OS done
SB1 successfully unassigned.
sul234SC0:sms-svc>
```

- Vérifier la suppression des éléments matériels par Solaris :

```
root # prtdiag
System Configuration: Sun Microsystems sun4u Sun Fire E25K
System clock frequency: 150 MHz
Memory size: 16384 Megabytes

===== CPUs =====

Slot ID      CPU      Run      E$      CPU      CPU
-----      -
/SB00/P0     0, 4    1500    32.0    US-IV+   2.2
/SB00/P1     1, 5    1500    32.0    US-IV+   2.2
/SB00/P2     2, 6    1500    32.0    US-IV+   2.2
/SB00/P3     3, 7    1500    32.0    US-IV+   2.2

root@E25K-SIDEF01 #
```

- Vérifier que la SB est bien disponible pour un attachement ultérieur:

```
sul234SC0:sms-svc:2> showboards
Retrieving board information. Please wait.
.....
Location Pwr   Type of Board   Board Status   Test Status
Domain
-----
SB0      On    V3CPU           Active         Passed
SIDEF
SB1      Off   V3CPU           Available      Unknown
Isolated
SB2      On    V3CPU           Active         Passed         ILOG
SB3      On    V3CPU           Active         Passed
ANAXOS
```

<i>SB4</i>	<i>On</i>	<i>V3CPU</i>	<i>Active</i>	<i>Passed</i>
<i>BOMBYX</i>				
<i>SB5</i>	<i>On</i>	<i>V3CPU</i>	<i>Active</i>	<i>Passed</i>
<i>RATUTI</i>				
<i>SB6</i>	<i>On</i>	<i>V3CPU</i>	<i>Active</i>	<i>Passed</i>
<i>SIDEFINT</i>				
<i>SB7</i>	<i>-</i>	<i>Empty Slot</i>	<i>Available</i>	<i>-</i>
<i>Isolated SB8</i>	<i>-</i>	<i>Empty Slot</i>	<i>Available</i>	<i>-</i>
<i>Isolated SB9</i>	<i>-</i>	<i>Empty Slot</i>	<i>Available</i>	<i>-</i>
<i>Isolated</i>				
<i>.....</i>				

3.16 Reconfiguration Dynamique (DR): Attachement d'une System Board à un domaine actif

Objectif: Affecter dynamiquement une ressource de type SB (system board) à un domaine opérationnel.

Nota: avant une opération de DR, vérifier que les ACL (access control list) soient correctement configurés.

Opérations:

- L'exemple présenté ci-après montre une opération d'attachement de la SB1 sur le domaine B.
- Vérifier la disponibilité de la SB1 (*Isolated*)

```

sul234SC0:sms-svc> showboards
Retrieving board information. Please wait.
.....
Location Pwr   Type of Board  Board Status  Test Status
Domain
-----
SB0      On    V3CPU         Active        Passed
SIDEF
SB1      Off    V3CPU         Available     Unknown
Isolated
SB2      On    V3CPU         Active        Passed      ILOG
SB3      On    V3CPU         Active        Passed
ANAXOS
SB4      On    V3CPU         Active        Passed
BOMBYX
SB5      On    V3CPU         Active        Passed
RATUTI
SB6      On    V3CPU         Active        Passed
SIDEFINT
SB7      -     Empty Slot    Available     -
Isolated
SB8      -     Empty Slot    Available     -
Isolated
SB9      -     Empty Slot    Available     -
Isolated
SB10     -     Empty Slot    Available     -
Isolated
SB11     -     Empty Slot    Available     -
Isolated
SB12     -     Empty Slot    Available     -
Isolated
SB13     -     Empty Slot    Available     -
Isolated
SB14     -     Empty Slot    Available     -
Isolated
SB15     -     Empty Slot    Available     -
Isolated
SB16     -     Empty Slot    Available     -
Isolated

```

```
SB17      -      Empty Slot      Available      -
Isolated
...
su1234SC0:sms-svc>
```

- S'assurer que les RTU COD sont correctement distribués en vue de l'opération d'attachement (cf §4.14 Gestion des jetons COD):

```
su1234SC0:sms-svc> showcodusage
Resource:
=====
Resource      In Use      Installed      Licensed      Status
-----
PROC           8           24             9      OK: 1 available

Domains:
=====
Domain/Resource  In Use      Installed      Reserved
-----
A - PROC         0           0              0
B - PROC         1           4              2
C - PROC         1           4              1
D - PROC         2           4              2
E - PROC         2           4              2
F - PROC         2           4              2
G - PROC         0           0              0
H - PROC         0           0              0
I - PROC         0           0              0
J - PROC         0           0              0
K - PROC         0           0              0
L - PROC         0           0              0
M - PROC         0           0              0
N - PROC         0           0              0
O - PROC         0           0              0
P - PROC         0           0              0
Q - PROC         0           0              0
R - PROC         0           0              0
Unused - PROC    0           4              0

su1234SC0:sms-svc>
```

- Visualiser les CPUs actives sur le domaine cible (1 seule CPU dans l'exemple suivant):

```
root@E25K-ILOG01 # prtdiag
System Configuration: Sun Microsystems sun4u Sun Fire E25K
System clock frequency: 150 MHz
Memory size: 4096 Megabytes

===== CPUs =====

Slot ID      CPU      Run      E$      CPU      CPU
ID           ID       MHz     MB     Impl.    Mask
-----
/SB02/P0     64, 68  1500    32.0   US-IV+   2.2

root@E25K-ILOG01 #
```

- Attacher la carte libre (**SB1**) au domaine:

```
su1234SC0:sms-svc> addboard -d b sb1
```

```

assign SB1
.
assign SB1 done
poweron SB1
poweron SB1 done
test SB1
.....
test SB1 done
connect SB1
...
connect SB1 done
configure SB1
.
configure SB1 done
notify online SUNW_cpu/cpu32
notify online SUNW_cpu/cpu36
notify add capacity (2 cpus)
notify add capacity (524288 pages)
notify add capacity SB1 done
.
sul234SC0:sms-svc>

```

- Vérifier la prise en compte des nouveaux éléments matériels par Solaris :

```

root@E25K-ILOG01 # prtdiag
System Configuration: Sun Microsystems sun4u Sun Fire E25K
System clock frequency: 150 MHz
Memory size: 8192 Megabytes

===== CPUs =====

Slot ID   CPU      Run    E$    CPU    CPU
         ID      MHz   MB    Impl.  Mask
-----
/SB01/P0  32, 36  1500  32.0  US-IV+  2.2
/SB02/P0  64, 68  1500  32.0  US-IV+  2.2

root@E25K-ILOG01 #

```

Remarque:

La commande *moveboard* permet de réaliser le détachement d'une SB depuis un domaine source suivi de l'attachement vers un domaine cible en une seule commande.

3.17 Reconfiguration Dynamique (DR): Détachement d'une carte FC gérée par STMS

Objectif: Retirer dynamiquement une ressource de type IO board - carte PCI FC de connexion au SAN gérée par STMS (Sun Storage Management System) - d'un domaine actif.

Procédure applicable dans le cadre d'une opération de maintenance sur la carte PCI.

Opérations:

Ce test n'est applicable qu'aux domaines mettant en œuvre du multipathing (STMS) sur 2 cartes PCI distinctes. En l'occurrence, ce test n'est applicable qu'aux domaines A des plates-formes.

Dans l'exemple suivant, il s'agit de retirer la carte du type **SG-PCI2FC-QF2** se trouvant dans le slot **C3V2** de l'I/O Board **0**.

- Vérifier que les 2 paths vers le SAN sont bien actifs, à l'aide de la commande *luxadm* exécutée sur l'une quelconque des Luns:

```
root # luxadm display /dev/rdisk/c8t6006048000028789081953594D333836d0s2
DEVICE PROPERTIES for disk:
/dev/rdisk/c8t6006048000028789081953594D333836d0s2
Vendor:                EMC
Product ID:            SYMMETRIX
Revision:              5670
Serial Num:            890819386000
Unformatted capacity: 15356.250 MBytes
Read Cache:           Enabled
  Minimum prefetch:   0x0
  Maximum prefetch:   0xffff
Device Type:           Disk device
Path(s):

/dev/rdisk/c8t6006048000028789081953594D333836d0s2
/devices/scsi_vhci/ssd@g6006048000028789081953594d333836:c,r
aw
Controller
/devices/pci@1d,700000/SUNW,qlc@1/fp@0,0
Device Address          500604844a3760e8,1f
Host controller port WWN 210000e08b8ff373
Class                   primary
State                   ONLINE
Controller
/devices/pci@3d,700000/SUNW,qlc@1/fp@0,0
Device Address          500604844a3760e7,1f
Host controller port WWN 210000e08b8fe874
Class                   primary
State                   ONLINE
root#
```

- Identifier l'emplacement de la cible de l'opération de DR comme suit :

Déterminer le numéro de l'*Expander Board* et l'emplacement de la carte PCI concernée :

- Pour le slot 0.0 (C3V2), emplacement PCI = slot3,
- Pour le slot 0.1 (C5V0), emplacement PCI = slot1,
- Pour le slot 1.0 (C3V1), emplacement PCI = slot2,
- Pour le slot 1.1 (C3V0), emplacement PCI = slot0.

L'emplacement PCI s'écrit **exxb1sloty**, ou **xx** est le numéro de l'*Expander Board* et **sloty** est l'emplacement PCI décrit ci-dessus.

- Identifier le point d'attachement de la cible de l'opération de DR (C3V2 *Expander 0* dans notre exemple):

```
root# cfmadm -v | grep C3V2 |grep e00
pcisch2:e00b1slot3          connected          configured      ok
C3V2
root#
```

- Identifier le path hardware de la carte PCI C3V2 de l'*Expander Board 0*

```
root# cfmadm -lv pcisch2:e00b1slot3
Ap_Id      Receptacle  Occupant    Condition  Information
When      Type      Busy      Phys_Id
pcisch2:e00b1slot3          connected          configured      ok
C3V2
Sep 22 14:21 mult/hp      n  /devices/pci@1d,700000:e00b1slot3
root#
```

- Identifier le numéro de contrôleur Solaris correspondant au path hardware de la carte PCI C3V2 de l'*Expander Board 0*

```
root# cfmadm -alv
Ap_Id      Receptacle  Occupant    Condition  Information
When      Type      Busy      Phys_Id
....
c4::500604844a3760e8          connected          configured
unknown
unavailable          disk          n
/devices/pci@1d,700000/SUNW,qlc@1/fp@0,0:fc::500604844a3760e8
....
root#
```

- Déconfigurer le contrôleur à l'aide de la commande *cfmadm* puis vérifier la désactivation du path à l'aide de la commande *luxadm*:

```
root# cfmadm -c unconfigure c4
root# luxadm display /dev/rdisk/c8t6006048000028789081953594D333835d0s2
DEVICE PROPERTIES for disk: /dev/rdisk/c8t6006048000028789081953594D333835d0s2
Vendor:          EMC
Product ID:      SYMMETRIX
Revision:        5670
Serial Num:      890819385000
Unformatted capacity: 15356.250 MBytes
Read Cache:      Enabled
Minimum prefetch: 0x0
```

```

Maximum prefetch: 0xffff
Device Type:      Disk device
Path(s) :

/dev/rdisk/c8t6006048000028789081953594D333835d0s2
/devices/scsi_vhci/ssd@g6006048000028789081953594d333835:c,r
aw
Controller
/devices/pci@3d,700000/SUNW,qlc@1/fp@0,0
Device Address    500604844a3760e7,1e
Host controller port WWN 210000e08b8fe874
Class             primary
State             ONLINE

root#

```

- Déconfigurer et déconnecter la carte PCI SG-PCI2FC-QF2 se trouvant en slot C3V2 de l'Expander Board 0, à l'aide de la commande `cfgadm` et en utilisant le point d'attachement relevé précédemment :

```

root# cfgadm
Ap_Id      Type          Receptacle  Occupant    Condition
IO0        HPCI+        connected   configured  ok
IO1        HPCI+        connected   configured  ok
SB0        V3CPU        connected   configured  ok
SB1        V3CPU        connected   configured  ok
c0         scsi-bus     connected   configured  unknown
c1         scsi-bus     connected   configured  unknown
c2         scsi-bus     connected   unconfigured unknown
c3         scsi-bus     connected   unconfigured unknown
c4         fc-fabric    connected   unconfigured unknown
c5         fc           connected   unconfigured unknown
c6         fc-fabric    connected   configured  unknown
c7         fc           connected   unconfigured unknown
pci_pci0:e00b1slot1  unknown     connected   unconfigured
unknown
pci_pci3:e01b1slot1  unknown     connected   unconfigured
unknown
pcisch1:e00b1slot0   pci-pci/hp  connected   configured  ok
pcisch2:e00b1slot3   mult/hp     connected   configured  ok
pcisch3:e00b1slot2   pci-pci/hp  connected   configured  ok
pcisch5:e01b1slot0   pci-pci/hp  connected   configured  ok
pcisch6:e01b1slot3   mult/hp     connected   configured  ok
pcisch7:e01b1slot2   unknown     connected   unconfigured
unknown
root# cfgadm -v -c disconnect pcisch2:e00b1slot3
root# cfgadm
Ap_Id      Type          Receptacle  Occupant    Condition
IO0        HPCI+        connected   configured  ok
IO1        HPCI+        connected   configured  ok
SB0        V3CPU        connected   configured  ok
SB1        V3CPU        connected   configured  ok
c0         scsi-bus     connected   configured  unknown
c1         scsi-bus     connected   configured  unknown
c2         scsi-bus     connected   unconfigured unknown
c3         scsi-bus     connected   unconfigured unknown
c6         fc-fabric    connected   configured  unknown
c7         fc           connected   unconfigured unknown

```

```
pci_pci0:e00b1slot1  unknown      connected   unconfigured
unknown
pci_pci3:e01b1slot1  unknown      connected   unconfigured
unknown
pcisch1:e00b1slot0   pci-pci/hp   connected   configured  ok
pcisch2:e00b1slot3   unknown      disconnected unconfigured
unknown
pcisch3:e00b1slot2   pci-pci/hp   connected   configured  ok
pcisch5:e01b1slot0   pci-pci/hp   connected   configured  ok
pcisch6:e01b1slot3   mult/hp      connected   configured  ok
pcisch7:e01b1slot2   unknown      connected   unconfigured
unknown
root#
```

Remarque:

La contrôleur c5 a également disparu de la liste des contrôleurs disponibles car la carte déconfigurée est une carte Dual FC et gère en l'occurrence les contrôleurs c4 et c5 sur cette instance de Solaris.

- La carte PCI peut désormais être physiquement retirée du slot C3V2 de l'Expander Board 0.

3.18 Reconfiguration Dynamique (DR): Attachement d'une carte FC gérée par STMS

Objectif: Attacher dynamiquement une ressource de type IO board - carte PCI FC de connexion au SAN gérée par STMS (Sun Storage Management System) - sur domaine actif.

Procédure applicable dans le cadre d'une opération de maintenance sur la carte PCI.

Opérations:

Ce test n'est applicable qu'aux domaines mettant en œuvre du multipathing (STMS) sur 2 cartes PCI distinctes. En l'occurrence, ce test n'est applicable qu'aux domaines A des plates-formes.

Dans l'exemple suivant, il s'agit de réintroduire la carte du type **SG-PCI2FC-QF2** se trouvant initialement dans le slot **C3V2** de l'*Expander Board 0*.

Etat initial:

Le domaine est considéré simplement attaché au SAN, le slot C3V2 de l'expander board 0 est libre (cf §4.17 Détachement d'une carte FC gérée par STMS).

- Identifier l'emplacement de la cible de l'opération de DR comme suit :
Déterminer le numéro de l'*Expander Board* et l'emplacement de la carte PCI concernée :
 - Pour le slot 0.0 (C3V2), emplacement PCI = slot3,
 - Pour le slot 0.1 (C5V0), emplacement PCI = slot1,
 - Pour le slot 1.0 (C3V1), emplacement PCI = slot2,
 - Pour le slot 1.1 (C3V0), emplacement PCI = slot0.

L'emplacement PCI s'écrit **exxb1sloty**, où **xx** est le numéro de l'*Expander Board* et **sloty** est l'emplacement PCI décrit ci-dessus.

- Vérifier que le point d'attachement de la cible de l'opération de DR est libre (C3V2 *Expander 0* dans notre exemple):

```
root# cfmadm -v | grep C3V2 |grep e00
pcisch2:e00b1slot3          empty          unconfigured      unknown
C3V2
root#
```

- Installer la carte PCI dans le slot 3V2 *Expander 0* et connecter la fibre d'attachement au SAN dans le slot identifié *Port 1*.
- Configurer la carte nouvellement installée et vérifier l'apparition des nouveaux contrôleurs (carte Dual FC) au niveau de Solaris, à l'aide de la commande *cfadm*:

```
root# cfmadm
Ap_Id          Type          Receptacle      Occupant      Condition
IO0            HPCI+        connected       configured    ok
IO1            HPCI+        connected       configured    ok
```

```

SB0          V3CPU          connected    configured  ok
SB1          V3CPU          connected    configured  ok
c0           scsi-bus       connected    configured  unknown
c1           scsi-bus       connected    configured  unknown
c2           scsi-bus       connected    unconfigured unknown
c3           scsi-bus       connected    unconfigured unknown
c6           fc-fabric     connected    configured  unknown
c7           fc            connected    unconfigured unknown
pci_pci0:e00b1slot1 unknown    connected    unconfigured
unknown
pci_pci3:e01b1slot1 unknown    connected    unconfigured
unknown
pcisch1:e00b1slot0 pci-pci/hp connected    configured  ok
pcisch2:e00b1slot3 unknown    disconnected unconfigured
unknown
pcisch3:e00b1slot2 pci-pci/hp connected    configured  ok
pcisch5:e01b1slot0 pci-pci/hp connected    configured  ok
pcisch6:e01b1slot3 mult/hp     connected    configured  ok
pcisch7:e01b1slot2 unknown    connected    unconfigured
unknown
root# cfgadm -v -c configure pcisch2:e00b1slot3
root# cfgadm
Ap_Id        Type           Receptacle    Occupant      Condition
IO0          HPCI+         connected     configured    ok
IO1          HPCI+         connected     configured    ok
SB0          V3CPU         connected     configured    ok
SB1          V3CPU         connected     configured    ok
c0           scsi-bus      connected     configured    unknown
c1           scsi-bus      connected     configured    unknown
c2           scsi-bus      connected     unconfigured  unknown
c3           scsi-bus      connected     unconfigured  unknown
c4           fc-fabric     connected     unconfigured  unknown
c5           fc            connected     unconfigured  unknown
c6           fc-fabric     connected     configured    unknown
c            fc            connected     unconfigured  unknown
pci_pci0:e00b1slot1 unknown    connected     unconfigured
unknown
pci_pci3:e01b1slot1 unknown    connected     unconfigured
unknown
pcisch1:e00b1slot0 pci-pci/hp    connected     configured    ok
pcisch2:e00b1slot3 mult/hp       connected     configured    ok
pcisch3:e00b1slot2 pci-pci/hp    connected     configured    ok
pcisch5:e01b1slot0 pci-pci/hp    connected     configured    ok
pcisch6:e01b1slot3 mult/hp       connected     configured    ok
pcisch7:e01b1slot2 unknown       connected     unconfigured
unknown
root#

```

- Configurer le contrôleur au niveau de Solaris:

```

root# cfgadm -c configure c4
root# cfgadm
Ap_Id        Type           Receptacle    Occupant      Condition
IO0          HPCI+         connected     configured    ok
IO1          HPCI+         connected     configured    ok
SB0          V3CPU         connected     configured    ok
SB1          V3CPU         connected     configured    ok
c0           scsi-bus      connected     configured    unknown
c1           scsi-bus      connected     configured    unknown
c2           scsi-bus      connected     unconfigured  unknown

```

```

c3          scsi-bus      connected   unconfigured unknown
c4          fc-fabric     connected   configured  unknown
c5          fc            connected   unconfigured unknown
c6          fc-fabric     connected   configured  unknown
c7          fc            connected   unconfigured unknown
pci_pci0:e00b1slot1 unknown    connected   unconfigured
unknown
pci_pci3:e01b1slot1 unknown    connected   unconfigured
unknown
pcisch1:e00b1slot0 pci-pci/hp connected   configured  ok
pcisch2:e00b1slot3 mult/hp     connected   configured  ok
pcisch3:e00b1slot2 pci-pci/hp connected   configured  ok
pcisch5:e01b1slot0 pci-pci/hp connected   configured  ok
pcisch6:e01b1slot3 mult/hp     connected   configured  ok
pcisch7:e01b1slot2 unknown    connected   unconfigured
unknown
root#

```

- Identifier le path hardware de la carte PCI C3V2 de l'Expander Board 0 correspondant au numéro de contrôleur Solaris:

```

root# cfgadm -alv
Ap_Id          Receptacle  Occupant    Condition  Information
When          Type        Busy        Phys_Id
....
c4::500604844a3760e8          connected   configured
unknown
unavailable          disk          n
/devices/pci@1d,700000/SUNW,q1c@1/fp@0,0:fc::500604844a3760e8
....
• root#

```

- Vérifier que ce path est désormais ONLINE pour accéder au SAN:

```

root# luxadm display /dev/rdisk/c8t6006048000028789081953594D333836d0s2
DEVICE PROPERTIES for disk: /dev/rdisk/c8t6006048000028789081953594D333836d0s2
Vendor:          EMC
Product ID:      SYMMETRIX
Revision:        5670
Serial Num:      890819386000
Unformatted capacity: 15356.250 MBytes
Read Cache:      Enabled
  Minimum prefetch: 0x0
  Maximum prefetch: 0xffff
Device Type:     Disk device
Path(s):

  /dev/rdisk/c8t6006048000028789081953594D333836d0s2
  /devices/scsi_vhci/ssd@g6006048000028789081953594d333836:c,r
aw
Controller
/devices/pci@3d,700000/SUNW,q1c@1/fp@0,0
  Device Address      500604844a3760e7,1f
  Host controller port WWN 210000e08b8fe874
  Class               primary
  State               ONLINE
Controller
/devices/pci@1d,700000/SUNW,q1c@1/fp@0,0
  Device Address      500604844a3760e8,1f
  Host controller port WWN 210000e08b8ff373
  Class               primary

```

<i>State</i> ONLINE
<code>root#</code>

3.19 Reconfiguration Dynamique (DR): Détachement d'une carte SCSI/Ethernet gérée par SVM/IPMP ou VXVM/IPMP

Objectif: Retirer dynamiquement une ressource de type IO board - carte PCI Dual SCSI/Ethernet (4422A) de connexion aux disques de boot et au réseau gérée par IPMP - d'un domaine actif.

Procédure applicable dans le cadre d'une opération de maintenance sur la carte PCI.

Opérations:

Les interfaces réseaux de la carte 4422A doivent être gérées par IPMP (*IP MultiPathing*). Chaque interface réseau IPMP de la carte 4422A doit disposer d'un lien IPMP de secours connecté à une autre carte PCI.

Les disques présents sur les interfaces SCSI de la carte 4422A doivent être gérés par Solaris Volume Manager et doivent disposer de miroirs présents sur une autre carte PCI.

Dans l'exemple suivant, il s'agit de retirer la carte du type **X4422A** se trouvant dans le slot **C3V1** de l'*Expander Board 5*.

- identifier l'emplacement de la cible de l'opération de DR comme suit :

Déterminer le numéro de l'*Expander Board* et l'emplacement de la carte PCI concernée :

- Pour le slot 0.0 (C3V2), emplacement PCI = slot3,
- Pour le slot 0.1 (C5V0), emplacement PCI = slot1,
- Pour le slot 1.0 (C3V1), emplacement PCI = slot2,
- Pour le slot 1.1 (C3V0), emplacement PCI = slot0.

L'emplacement PCI s'écrit **exxb1sloty**, où **xx** est le numéro de l'*Expander Board* et **sloty** est l'emplacement PCI décrit ci-dessus.

- Identifier le point d'attachement de la cible de l'opération de DR (C3V0 *Expander 1* dans notre exemple):

```
root# cfgadm -v | grep C3V1 |grep e05
pcisch3:e05b1slot2    connected    configured    ok           C3V1
root#
```

- Tenter la déconfiguration de la carte PCI 4422A se trouvant en slot C3V1 de l'*Expander Board 5*, à l'aide de la commande *cfgadm*, en utilisant le point d'attachement relevé précédemment :

```
root# cfgadm -v -c disconnect pcisch3:e05b1slot2
cfgadm: Component system is busy, try again:
Resource                               Information
-----
/dev/dsk/c1t10d0s0 component of concat "/dev/md/dsk/d20"
```

```
/dev/md/dsk/d20    submirror of "/dev/md/dsk/d0"  
/dev/md/dsk/d0    mounted filesystem "/"  
/dev/dsk/c1t10d0s1 component of concat "/dev/md/dsk/d21"  
/dev/md/dsk/d21    submirror of "/dev/md/dsk/d1"  
/dev/md/dsk/d1    swap area  
/dev/md/dsk/d1    dump device (swap)  
/dev/dsk/c1t10d0s3 component of concat "/dev/md/dsk/d23"  
/dev/md/dsk/d23    submirror of "/dev/md/dsk/d3"  
/dev/md/dsk/d3    mounted filesystem "/usr"  
/dev/dsk/c1t10d0s4 component of concat "/dev/md/dsk/d24"  
/dev/md/dsk/d24    submirror of "/dev/md/dsk/d4"  
/dev/md/dsk/d4    mounted filesystem "/var"  
/dev/dsk/c1t10d0s7 contains metadb(s)  
root#
```

Remarque:

Le résultat de la commande précédente nous informe que l'opération de DR ne peut pas être effectuée. En effet l'interface SCSI c1 est utilisée sur cette carte (disque système miroir). Les accès au disques c1 doivent être désactivés à l'aide des commandes Volume Manager.

Cas Solaris Volume Manager

- Désactiver les accès au contrôleur c1 en détachant le disque miroir système puis en détruisant les réplicats présents sur ce disque:

```
root# metadetach d0 d20  
d0: submirror d20 is detached  
root# metadetach d1 d21  
d1: submirror d21 is detached  
root# metadetach d3 d23  
d3: submirror d23 is detached  
root# metadetach d4 d24  
d4: submirror d24 is detached  
root# metadb -d /dev/dsk/c1t10d0s7  
root#
```

Cas Veritas Volume Manager

- Désactiver les accès au contrôleur c1 en détachant le disque miroir système. Utiliser la commande vxdiskadm

```
# vxdiskadm  
  
Volume Manager Support Operations  
Menu: VolumeManager/Disk  
  
1 Add or initialize one or more disks  
2 Encapsulate one or more disks  
3 Remove a disk  
4 Remove a disk for replacement  
5 Replace a failed or removed disk  
6 Mirror volumes on a disk  
7 Move volumes from a disk  
8 Enable access to (import) a disk group  
9 Remove access to (deport) a disk group
```

```

.../...
16 Make a disk available for hot-relocation use
17 Prevent multipathing/Suppress devices from VxVM's view
18 Allow multipathing/Unsuppress devices from VxVM's view
19 List currently suppressed/non-multipathed devices
list List disk information

? Display help about menu
?? Display help about the menuing system
q Exit from menus

Select an operation to perform: 4

```

Choisir l'option 4, « Remove a disk for replacement »

- Visualiser la configuration IPMP de l'interface ce2 active sur la carte à détacher

```

root# ifconfig -a
lo0: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4> mtu 8232 index 1
    inet 127.0.0.1 netmask ffffffff
ce0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
    inet 10.151.231.10 netmask ffff0000 broadcast 10.151.231.255
    groupname fret
    ether 0:14:4f:23:37:a9
ce0:1:
flags=9040843<UP,BROADCAST,RUNNING,MULTICAST,DEPRECATED,IPv4,NOFAILOVER> mtu
1500 index 2
    inet 10.151.231.11 netmask ffff0000 broadcast 10.151.231.255
ce1: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 3
    inet 10.40.34.184 netmask ffff0000 broadcast 10.40.35.255
    ether 0:14:4f:23:37:aa
ce2: flags=9040843<UP,BROADCAST,RUNNING,MULTICAST,DEPRECATED,IPv4,NOFAILOVER>
mtu 1500 index 4
    inet 10.151.231.12 netmask ffff0000 broadcast 10.151.231.255
    groupname fret
    ether 0:14:4f:23:37:7b
dman0: flags=1008843<UP,BROADCAST,RUNNING,MULTICAST,PRIVATE,IPv4> mtu 1500
index 5
    inet 192.168.1.6 netmask ffff0000 broadcast 192.168.1.31
    ether 0:0:be:a9:8d:60
root#

```

- Déconfigurer la carte PCI 4422A:

```

root# cfgadm -v -c disconnect pcisch3:e05b1slot2
root#

```

- Constater la disparition au niveau de Solaris du contrôleur c1 ainsi que de l'interface réseau ce2:

```

root# cfgadm

```

Ap_Id	Type	Receptacle	Occupant
Condition			
I05	HPCI+	connected	configured ok
SB5	V3CPU	connected	configured ok
c0	scsi-bus	connected	configured unknown
c2	scsi-bus	connected	unconfigured unknown
c4	fc	connected	unconfigured unknown
c5	fc	connected	unconfigured unknown
pci_pci0:e05b1slot1	unknown	connected	unconfigured unknown

```
pcisch1:e05b1slot0 pci-pci/hp connected configured ok
pcisch2:e05b1slot3 mult/hp connected configured ok
pcisch3:e05b1slot2 unknown disconnected unconfigured
unknown
root# ifconfig -a
lo0: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4> mtu 8232 index 1
    inet 127.0.0.1 netmask ff000000
ce0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
    inet 10.151.231.10 netmask fffffc00 broadcast 10.151.231.255
    groupname fret
    ether 0:14:4f:23:37:a9
ce0:1:
flags=9040843<UP,BROADCAST,RUNNING,MULTICAST,DEPRECATED,IPv4,NOFAILOVER> mtu
1500 index 2
    inet 10.151.231.11 netmask fffffc00 broadcast 10.151.231.255
cel: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 3
    inet 10.40.34.184 netmask fffffc00 broadcast 10.40.35.255
    ether 0:14:4f:23:37:aa
dman0: flags=1008843<UP,BROADCAST,RUNNING,MULTICAST,PRIVATE,IPv4> mtu 1500
index 5
    inet 192.168.1.6 netmask fffffe0 broadcast 192.168.1.31
    ether 0:0:be:a9:8d:60
root#
```

- La carte PCI peut désormais être physiquement retirée du slot C3V1 de l'Expander Board 5.

3.20 Reconfiguration Dynamique (DR): Attachement d'une carte SCSI/Ethernet gérée par SVM/IPMP ou VXVM/IPMP

Objectif: Attacher dynamiquement une ressource de type IO board - carte PCI Dual SCSI/Ethernet (4422A) de connexion aux disques de boot et au réseau gérée par IPMP - à un domaine actif.

Procédure applicable dans le cadre d'une opération de maintenance sur la carte PCI.

Opérations:

Les interfaces réseaux de la carte 4422A doivent être gérées par IPMP (*IP MultiPathing*). Chaque interface réseau IPMP de la carte 4422A doit disposer d'un lien IPMP de secours connecté à une autre carte PCI.

Les disques présents sur les interfaces SCSI de la carte 4422A doivent être gérés par Solaris Volume Manager et doivent disposer de miroirs présents sur une autre carte PCI.

Dans l'exemple suivant, il s'agit d'attacher une carte du type **X4422A** dans le slot **C3V1** de l'*Expander Board 5*.

Etat initial:

Le domaine est considéré simplement attaché au disque de boot, le slot C3V1 de l'expander board 5 est libre (cf §4.17 Détachement d'une carte Dual SCSI/Ethernet gérée avec SVM et IPMP).

- Identifier l'emplacement de la cible de l'opération de DR comme suit :

Déterminer le numéro de l'*Expander Board* et l'emplacement de la carte PCI concernée :

- Pour le slot 0.0 (C3V2), emplacement PCI = slot3,
- Pour le slot 0.1 (C5V0), emplacement PCI = slot1,
- Pour le slot 1.0 (C3V1), emplacement PCI = slot2,
- Pour le slot 1.1 (C3V0), emplacement PCI = slot0.

L'emplacement PCI s'écrit **exxb1sloty**, où **xx** est le numéro de l'*Expander Board* et **sloty** est l'emplacement PCI décrit ci-dessus.

- Vérifier que le point d'attachement de la cible de l'opération de DR est libre (C3V2 Expander 0 dans notre exemple):

```
root# cfdm -v | grep C3V1 |grep e05
pcisch3:e05b1slot2 empty unconfigured unknown C3V1
root#
```

- Installer la carte PCI dans le slot *3V1 Expander 5* et connecter:

- le bus SCSI0 au disque miroir système,
- l'interface Ethernet ENET0 au réseau.

- Configurer la carte nouvellement installée et vérifier l'apparition du contrôleur c1 et de l'interface Ethernet ce2 au niveau de Solaris, à l'aide de la commande `cfgadm`:

```

root# cfgadm
Ap_Id          Type          Receptacle    Occupant
Condition
IO5            HPCI+        connected     configured    ok
SB5            V3CPU        connected     configured    ok
c0             scsi-bus     connected     configured
unknown
c2             scsi-bus     connected     unconfigured
unknown
c4             fc           connected     unconfigured
unknown
c5             fc           connected     unconfigured
unknown
pci_pci0:e05b1slot1 unknown     connected     unconfigured
unknown
pcisch1:e05b1slot0 pci-pci/hp  connected     configured    ok
pcisch2:e05b1slot3 mult/hp     connected     configured    ok
pcisch3:e05b1slot2 unknown     disconnected   unconfigured
unknown
root# ifconfig -a
lo0: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4> mtu 8232 index 1
      inet 127.0.0.1 netmask ff000000
ce0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
      inet 10.151.231.10 netmask fffffc00 broadcast 10.151.231.255
      groupname fret
      ether 0:14:4f:23:37:a9
ce0:1:
flags=9040843<UP,BROADCAST,RUNNING,MULTICAST,DEPRECATED,IPv4,NOFAILOVER> mtu
1500 index 2
      inet 10.151.231.11 netmask fffffc00 broadcast 10.151.231.255
cel1: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 3
      inet 10.40.34.184 netmask fffffc00 broadcast 10.40.35.255
      ether 0:14:4f:23:37:aa
dman0: flags=1008843<UP,BROADCAST,RUNNING,MULTICAST,PRIVATE,IPv4> mtu 1500
index 5
      inet 192.168.1.6 netmask fffffe0 broadcast 192.168.1.31
      ether 0:0:be:a9:8d:60
root# cfgadm -v -c configure pcisch3:e05b1slot2
root# cfgadm
Ap_Id          Type          Receptacle    Occupant    Condition
IO5            HPCI+        connected     configured  ok
SB5            V3CPU        connected     configured  ok
c0             scsi-bus     connected     configured  unknown
c1             scsi-bus     connected     configured  unknown
c2             scsi-bus     connected     unconfigured unknown
c3             scsi-bus     connected     unconfigured unknown
c4             fc           connected     unconfigured unknown
c5             fc           connected     unconfigured unknown
pci_pci0:e05b1slot1 unknown     connected     unconfigured
unknown
pcisch1:e05b1slot0 pci-pci/hp  connected     configured  ok
pcisch2:e05b1slot3 mult/hp     connected     configured  ok
pcisch3:e05b1slot2 pci-pci/hp  connected     configured  ok
root# ifconfig -a
lo0: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4> mtu 8232 index 1
      inet 127.0.0.1 netmask ff000000
ce0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
      inet 10.151.231.10 netmask fffffc00 broadcast 10.151.231.255
      groupname fret
      ether 0:14:4f:23:37:a9

```

```
ce0:1:
flags=9040843<UP,BROADCAST,RUNNING,MULTICAST,DEPRECATED,IPv4,NOFAILOVER> mtu
1500 index 2
    inet 10.151.231.11 netmask fffffc00 broadcast 10.151.231.255
cel: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 3
    inet 10.40.34.184 netmask fffffc00 broadcast 10.40.35.255
    ether 0:14:4f:23:37:aa
dman0: flags=1008843<UP,BROADCAST,RUNNING,MULTICAST,PRIVATE,IPv4> mtu 1500
index 5
    inet 192.168.1.6 netmask fffffe0 broadcast 192.168.1.31
    ether 0:0:be:a9:8d:60
ce2: flags=9040843<UP,BROADCAST,RUNNING,MULTICAST,DEPRECATED,IPv4,NOFAILOVER>
mtu 1500 index 6
    inet 10.151.231.12 netmask fffffc00 broadcast 10.151.231.255
    groupname fret
    ether 0:14:4f:23:37:7b
root#
```

Cas Solaris Volume Manager

- Recréer les réplicats SVM sur le disque miroir (contrôleur c1) puis le resynchroniser avec le disque système primaire:

```
root# metadb -a -c3 /dev/dsk/c1t10d0s7
root# metattach d0 d20
d0: submirror d20 is attached
root# metattach d1 d21
d1: submirror d21 is attached
root# metattach d3 d23
d3: submirror d23 is attached
root# metattach d4 d24
d4: submirror d24 is attached
root#
```

Cas Veritas Volume Manager

- Afin que VxVM prenne en compte ce nouveau disque dans sa configuration, exécuter la commande :

```
# vxdctl enable
```

Utiliser la commande vxdiskadm

```
# vxdiskadm

Volume Manager Support Operations
Menu: VolumeManager/Disk

1 Add or initialize one or more disks
2 Encapsulate one or more disks
3 Remove a disk
4 Remove a disk for replacement
5 Replace a failed or removed disk
6 Mirror volumes on a disk
7 Move volumes from a disk
8 Enable access to (import) a disk group
9 Remove access to (deport) a disk group
```

```
.../...
16 Make a disk available for hot-relocation use
17 Prevent multipathing/Suppress devices from VxVM's view
18 Allow multipathing/Unsuppress devices from VxVM's view
19 List currently suppressed/non-multipathed devices
list List disk information

? Display help about menu
?? Display help about the menuing system
q Exit from menus

Select an operation to perform: 5
```

- Choisir l'option 5, « Replace a failed or removed disk ». Puis vérifier avec la commande `vxtask list` l'avancement de la resynchronisation des volumes.

3.21 Contrôler l'état de la plate-forme.

Objectif : S'assurer de la disponibilité et de l'état des sous ensembles de la plate-forme.

Les *system controllers* vérifient en permanence l'état des composants de la plate-forme. Il est possible à tout instant de visualiser cet état.

Deux grandes familles sont identifiables:

- La disponibilité et l'état des domaines configurés,
- Le monitoring et l'état du Hardware présent.

Opérations:

- Vérifier la cohérence et le niveau de disponibilité des domaines via la commande *showboards* et interprétation des résultats :

Colonne Pwr:

- ◆ *ON* signifie sous tension
- ◆ *OFF* signifie hors tension

Colonne Board Status:

- ◆ *Active* signifie sous contrôle d'un domaine booté (Solaris)
- ◆ *Assigned* signifie que le composant est affecté à un domaine non démarré.
- ◆ *Available* signifie que le slot correspondant est vide ou le composant non affecté.

Colonne Tests Status :

- ◆ *Passed* signifie qu'un *hpost* a été déroulé sur ce composant (lors d'un *setkeyswitch on* ou *setkeyswitch standby* par exemple).
- ◆ *Failed* signifie que le composant est testé mais qu'une erreur est détectée.
- ◆ - ou les autres états signifient que le composant n'est pas testé et donc pas utilisable par un domaine

- Monitoring de la plate-forme :

Le monitoring de la plate-forme se repose principalement sur la commande *showenvironment* qui permet d'avoir une vision globale ou plus ciblée des sous-ensembles du SunFire 25K.

```
Showenvironment -p domain_id -p temps|volts|currents|fans|powers|faults  
[-v]
```

- ◆ Temps: affiche uniquement les relevés de température,
- ◆ Volts: affiche les informations sur le voltage,
- ◆ Currents: affiche les informations sur le courant (alimentation uniquement),
- ◆ Fans: affiche les informations relatives au système de ventilation,
- ◆ Powers: affiche, par bloc d'alimentation, un état général des alimentations,
- ◆ Faults: Affiche uniquement les mesures relevées hors norme (niveau trop bas ou trop haut).

3.22 Procédure de restauration d'un System Controller.

Objectif: restaurer un SC.

La restauration d'un SC est identique à celui d'un domaine.

Il s'agit de restaurer à partir d'une bande via la commande *ufsrestore*, ou de la commande appropriée en fonction du type de backup (tar, cpio).

Les configurations et les informations liées à la plate-forme seront mise à jour, si nécessaire, par le mécanisme de «*datasync*» embarqué sur les SC et géré par le SC MAIN.

Un point est à noter, les disques systèmes des SC sont mirrorés via Solaris Volume Manager et nécessite donc quelques opérations spécifiques:

3.22.1 Contexte:

La politique de sauvegarde appliquée aux *system controllers* de la plate-forme SunFire 25K s'articule ordinairement autour d'un mécanisme de sauvegarde du système sur une cartouche DAT indépendante, au format *ufsdump Of*.

La cartouche DAT contenant le backup système doit être préalablement introduite dans le lecteur local du SC à restaurer.

Par convention, on appelle ci-après "disque système" le(s) disque(s) qui contien(nen)t le système d'exploitation, c'est-à-dire les systèmes de fichiers dont les points de montage sont / et /var ainsi qu'une partition de swap. Il s'agit des systèmes de fichiers requis, pour faire booter un *system controller*.

3.22.2 Boot cdrom du system controller

Introduire une distribution Solaris 9 dans le lecteur de DVD local, puis depuis le port série console du SC à restaurer, saisir:

```
ok> boot cdrom - s
```

3.22.3 Création du partitionnement du disque système

Il est possible que le partitionnement du disque système *c0t2d0*, ne soit plus correcte. Pour vérifier ou recréer le partitionnement du disque système, du SC à restaurer, on utilise la commande interactive *format*:

```
# format
Searching for disks...done
AVAILABLE DISK SELECTIONS:
    0. c0t2d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
       /pci@1f,0/pci@1,1/scsi@2/sd@2,0
    1. c0t3d0 <SUN72G cyl 14087 alt 2 hd 24 sec 424>
       /pci@1f,0/pci@1,1/scsi@2/sd@3,0
Specify disk (enter its number): 0 <- Selectionner c0t2d0
format> partition <- Selectionner le menu partition
format> print <- Afficher le partitionnement
Current partition table (original):
Total disk cylinders available: 14087 + 2 (reserved cylinders)
Part  Tag   Flag  Cylinders          Size              Blocks
  0    root   wm     0 - 1648           8.00GB            (1649/0/0) 16780224
  1    swap   wu   1649 - 2473         4.00GB            (825/0/0)  8395200
  2    backup  wm     0 - 14086          68.35GB           (14087/0/0) 143349312
  3 unassigned wm     0                   0                 (0/0/0)      0
  4 unassigned wm   2474 - 2480         34.78MB           (7/0/0)         71232
  5 unassigned wm   2481 - 2487         34.78MB           (7/0/0)         71232
  6 unassigned wm     0                   0                 (0/0/0)      0
  7 unassigned wm   2488 - 14086        56.28GB           (11599/0/0) 118031424
partition> quit
```

3.22.4 Création des systèmes de fichiers .

Remarque : Par convention, dans la suite du texte, la partition 0 correspond à /, la partition 1 à la swap, et la partition 7 à /export/install du SC à restaurer.

- Sur le port série console du SC à restaurer, taper:

```
# newfs /dev/rdisk/c0t2d0s0 (création de /)
# newfs /dev/rdisk/c0t2d0s7 (création de /export/install)
```

3.22.5 Restauration de / et opérations manuelles à effectuer.

- Restauration du système de fichier / :

- Sur le port série console du SC à restaurer, taper :

```
# mount /dev/rdisk/c0t2d0s0 /mnt
# cd /mnt
# ufsrestore rf /dev/rmt/0n
# cd /mnt/etc
```

- Positionnement de la variable TERM est requis pour avoir une émulation correcte sous l'éditeur vi:

- Sur le port série console du SC à restaurer, taper :

```
# TERM=vt100
# export TERM
```

- Sauvegarde des fichier /mnt/etc/system et /mnt/etc/vfstab originaux

- Sur le port série console du SC à restaurer, taper :

```
# cd /mnt/etc
# cp system system.org
# cp vfstab vfstab.org
```

- Modification de vfstab

- Sur le port série console du SC à restaurer, taper :

```
# vi /mnt/etc/vfstab
```

- ✓ Remplacer le *metadevice* d10 de la partition / par c0t2d0s0
- ✓ Remplacer le *metadevice* d20 de la partition / par c0t2d0s1
- ✓ Remplacer le *metadevice* d30 de la partition /export/install par c0t2d0s7

- Fichier vfstab restauré (avant modification)

#device	device	mount	FS	fsck	mount	mount
#to mount	to fsck	point	type	pass	at boot	options
#						
#/dev/dsk/c1d0s2	/dev/rdisk/c1d0s2	/usr	ufs	1	yes	-
fd	-	/dev/fd	fd	-	no	-
/proc	-	/proc	proc	-	no	-
/dev/md/dsk/d20	-	-	swap	-	no	-
/dev/md/dsk/d10	/dev/md/rdisk/d10	/	ufs	1	no	logging
/dev/md/dsk/d30	/dev/md/rdisk/d30	/export/install	ufs	1	no	logging
swap	-	/tmp	tmpfs	-	yes	-

- Fichier vfstab modifié

#device	device	mount	FS	fsck	mount	mount
#to mount	to fsck	point	type	pass	at boot	options
#						
#/dev/dsk/c1d0s2	/dev/rdisk/c1d0s2	/usr	ufs	1	yes	-
fd	-	/dev/fd	fd	-	no	-
/proc	-	/proc	proc	-	no	-
/dev/dsk/c0t2d0s1	-	-	swap	-	no	-
/dev/dsk/c0t2d0s0	/dev/rdisk/c0t2d0s0	/	ufs	1	no	logging
/dev/dsk/c0t2d0s7	/dev/rdisk/c0t2d0s7	/export/install	ufs	1	no	logging
swap	-	/tmp	tmpfs	-	yes	-

- Modification du fichier system

Elle s'effectue par suppression des références à l'encapsulation SDS du disque de boot

- Supprimer les lignes suivantes :

Lignes du fichier /mnt/etc/system à supprimer

```

rootdev:/pseudo/md@0:0,10,blk
* End MDD root info (do not edit)
* Begin MDD database info (do not edit)
set md:mddb_bootlist1="sd:20:16 sd:20:1050 sd:20:2084 sd:21:16
sd:21:1050"
set md:mddb_bootlist2="sd:21:2084 sd:28:16 sd:28:1050 sd:28:2084
sd:29:16"
set md:mddb_bootlist3="sd:29:1050 sd:29:2084"
* End MDD database info (do not edit)

```

- Supprimer le fichier mddb.cf
 - Sur le port série console du SC à restaurer tapez :

```
# rm /mnt/etc/lvm/mddb.cf
```

- Installation du bootblock.
 - Sur le port série console du SC à restaurer tapez :

```
# installboot /usr/platform/sun4u/lib/fs/ufs/bootblk \  
/dev/rdisk/c0t2d0s0  
# cd /  
# umount /mnt  
# fsck /dev/rdisk/c0t2d0s0
```

3.22.6 Restauration de /export/install

- Sur le port série console du SC à restaurer tapez :

```
# mount /dev/rdisk/c0t2d0s7 /mnt  
# cd /mnt  
# ufsrestore rf /dev/rmt/0n  
# cd /  
# umount /mnt
```

3.22.7 Reboot en single-user

- Sur le port série console du SC à restaurer tapez :

```
# init 0  
ok boot -s
```

- Se logger root.

3.22.8 Réencapsulation dans Solaris Volume Manager du disque de boot

- Recréer les replicaDB de SDS:
 - Sur la console du SC à restaurer tapez:

```
# metadb -fac 3 c0t2d0s4  
# metadb -fac 3 c0t2d0s5  
# metadb -fac 3 c0t3d0s4  
# metadb -fac 3 c0t3d0s5
```

- Vérifier le fichier `/etc/lvm/md.tab`
 - Le fichier doit être renseigné avec des miroirs à une voie sur le disque `c0t2d0`:

```
d10 -m d11 1
d11 1 1 c0t2d0s0
d20 -m d21 1
d21 1 1 c0t2d0s1
d30 -m d31 1
d31 1 1 c0t2d0s7
d32 1 1 c0t3d0s7
d12 1 1 c0t3d0s0
d22 1 1 c0t3d0s1
```

- Recréer tous les metadevices
 - Sur la console du SC à restaurer tapez :

```
# metainit -fa
# metaroot d0
```

- Restaurer le fichier `vfstab` correspondant à SVM.
 - Sur la console du SC à restaurer tapez :

```
# cd /etc
# cp vfstab.org vfstab
```

- Rebooter
 - Sur le port série console du SC à restaurer tapez :

```
# init 0
ok boot
```

3.22.9 Mise en miroir du disque système avec SVM

- Mettre `/`, `swap` et `/export/install`

```
# metattach d10 d12
# metattach d20 d22
# metattach d30 d32
```

- Attendre la fin des synchronisations.

```
# metastat
```

4 Traitements des dysfonctionnements

Ce chapitre décrit en détail les diverses procédures utilisées pour le traitement des dysfonctionnements éventuels de la plate-forme SunFire 25K. Il détaille également les traitements d'échec et de récupération d'une plate-forme Enterprise™ SF25K. Les autres éléments abordés dans ce chapitre sont les suivants :

- Comprendre le rôle des SC dans l'historique des défaillances,
- Identifier les différents types de défaillances,
- Localiser les enregistrements des informations de défaillance,
- Interpréter les informations de défaillance,
- Comprendre les exigences du support en termes d'informations.

4.1 Historique standard des messages de domaine

Les historiques standards des messages de domaine contiennent des informations importantes relatives aux problèmes affectant un seul domaine (ou tout le serveur) et pouvant être dus à une défaillance matérielle ou logicielle quelconque.

Ces messages peuvent traduire tant un *panic* du noyau (arrêt dû à une condition exceptionnelle) qu'une indication de surchauffe dans le matériel. De plus, des erreurs logicielles utilisateur peuvent être consignées, notamment les informations relatives aux actions exécutées, les refus de démarrage et les erreurs sémantiques des commandes incorrectes.

Chaque domaine comprend son fichier de messages SC spécifique :

```
$SMSVAR/adm/nom_domaine/messages
```

Consultez prioritairement ce fichier dès l'apparition d'un problème dans un domaine.

Le fichier de messages général à la plate-forme SunFire 25K se nomme :

```
$SMSVAR/adm/messages
```

4.2 Liste noire (*blacklist file*)

Au démarrage, *hpost* lit le fichier *blacklist*. Celui-ci définit les ressources que *hpost* doit considérer comme défaillantes ; *hpost* peut les réinitialiser ou les suspendre, mais elles ne seront pas configurées dans le système.

Vous pouvez faciliter la résolution d'un problème en plaçant les composants susceptibles d'être défaillants en *blacklist*. Ceci permet aussi de faire fonctionner un domaine en isolant un composant préalablement identifié comme défaillant.

Utilisez le manuel en ligne sur le SC pour obtenir le détail des mots clés du fichier *blacklist*, ou exécuter la commande suivante:

```
sms-svc> hpost -?blacklist
```

ATTENTION : Ne placez jamais un composant en *blacklist* sans instructions spécifiques du personnel de support Sun.

Les fichiers *blacklist* (`$SMSETC/config/domain_name/blacklist`) ne doivent pas être modifiés manuellement mais via les jeux de commande `disablecomponent/enablecomponent`. `Showcomponent` permet de vérifier l'état des *blacklists* (les fichiers *blacklist* restant visualisables).

Trois niveaux sont possibles :

- ◆ Niveau domaine : `$SMSETC/config/domain_id` : le composant est blacklisté uniquement s'il appartient au domaine spécifié,
- ◆ Niveau plate-forme : `$SMSETC/config/platform` : le composant est blacklisté quelque soit le domaine propriétaire,
- ◆ Niveau ASR : `$SMSETC/config/asr` : le composant est « signalé » comme defectueux et est configuré dans la liste ASR.

En cas de multiple *blacklist*, les directives dans tous les fichiers sont combinées et exécutées.

La prise en compte des fichiers *blacklist* est réalisée lors du `setkeyswitch -d domaine_id on` du domaine concerné.

Il est fortement recommandé de commenter la raison du blacklistage. Ce commentaire pouvant être consulté à tout moment via la commande `showcomponent`.

- Exemple de retrait de la SB1 du domaine A :

```
sms-svc> disablecomponent -d A SB1 -c commentaire
```

- Vérification par :

```
sms-svc> showcomponent -d A
```

- Réactivation par :

```
sms-svc> enablecomponent -d A SB2
```

4.3 Outils de diagnostic

Plusieurs programmes fournissent des informations de diagnostic relatives au serveur SunFire 25K. Ces outils sont :

- hpost,
- prtdiag.

4.3.1 hpost

La commande `hpost` peut être exécutée directement à partir de la ligne de commande `SC` *bien qu'il soit préférable de la lancer par la commande `setkeyswitch`* ; les options citées ci-après sont transposables d'une commande sur l'autre.

AVERTISSEMENT: `hpost` provoquera le *crash* d'un domaine en fonctionnement. Il ne vérifie pas l'état du domaine avant son démarrage.

La commande `hpost` peut être exécutée sur un domaine pendant le fonctionnement d'autres domaines .

Veillez à spécifier le nom de domaine approprié au moyen de la commande `setkeyswitch -d domain_id` .

La commande `hpost` comprend de nombreuses options ; veuillez vous référer à la page de manuel en ligne de `hpost` pour de plus amples informations.

AVERTISSEMENT : Si vous ne comprenez pas l'action d'une option, ne l'utilisez pas. Vous pourriez engendrer des dommages importants au système.

La commande `hpost` peut être exécutée selon plusieurs niveaux par les deux options suivantes :

- `-l n`, où `n` est le niveau de précision des tests, compris entre 7 et 127. La valeur par défaut est 16. Le niveau 64 est le maximum recommandé pour les opérations de dépannage sur le terrain ; l'exécution de la commande requiert alors de une heure trente minutes à deux heures (temps moyens).
- `-v m`, où `m` est le niveau de verbosité est compris entre 0 et 255. La valeur par défaut est 20. L'exécution au niveau 255 est utile (une seule fois) pour comprendre le fonctionnement de `hpost` ; cependant, une sortie volumineuse peut masquer les rapports d'erreurs. Le niveau de verbosité par défaut signale toutes les défaillances.

REMARQUE: Le format et la syntaxe du fichier `.postrc` sont accessibles via la commande suivante:

```
sms-svc:> hpost -?postrc
```

4.3.2 prtdiag

Lorsque `prtdiag` est exécutée à partir du domaine, elle génère l'affichage des éléments suivants :

- La configuration système,

- La fréquence de l'horloge système,
- Les vitesses des processeurs et les tailles de cache,
- Les tailles de chaque banque de mémoire,
- Une liste de toutes les cartes E/S et leur fréquence.

Pour exécuter `prtdiag` en tant qu'utilisateur `root` ou régulier, entrez la ligne suivante :

```
# /usr/platform/sun4u/sbin/prtdiag
```

- Pour des informations de diagnostic, voir sur le SC le fichier :

```
/var/opt/SUNWSMS/adm/nom_domaine/messages.
```

4.4 Erreurs de mémoire corrigées

Les erreurs de mémoire corrigées, tout en n'ayant aucun effet sur l'exécution des logiciels (autre qu'une dégradation mineure des performances) peuvent révéler un dysfonctionnement de mémoire DIMM susceptible d'occasionner des défaillances ultérieures plus sévères.

Le serveur Sun Fire 25K comprend une fonction décodant l'adresse du DIMM défectueux à partir de l'emplacement mémoire signalé, vous permettant ainsi de localiser et de remplacer le DIMM défectueux (en utilisant, si possible, la commande de reconfiguration dynamique DR). Cette fonction permet au noyau d'afficher le numéro de carte système, le numéro de banque mémoire et l'emplacement du DIMM sur la carte. L'affichage de sortie construit par l'OBP apparaît sous la forme suivante :

Le message d'erreur du système d'exploitation apparaît sous la forme suivante :

```
Softerror: Intermittent ECC Memory Error SIMM
Board# 3 Bank# 0 P# MM 0_3
ECC Data Bit 63 was corrected
```

dans lequel P# et MM 0_0 sont les repères sérigraphiés sur la carte support du DIMM affecté par les erreurs de mémoire.

4.5 Défaillances système

Plusieurs événements peuvent se produire sur le serveur SunFire 25K nécessitant une manipulation spécifique ou l'enregistrement d'informations d'état. Ces événements sont les suivants :

- Demande de redémarrage (*reboot*),
- Arrêt dû à une condition exceptionnelle (*panic*),
- Réinitialisation par *watchdog*, *redmode* ou *xir*,
- Hôte bloqué (*heartbeat failure*),
- *Arbstop*.

4.5.1 Demande de redémarrage (reboot)

Un *reboot* se produit lorsqu'une commande de redémarrage est émise à partir de la console de domaine de la manière suivante :

```
nom_domaine# reboot
```

Lorsqu'une telle commande est exécutée, le traitement suivant est effectué :

1. Le système d'exploitation s'arrête.
2. L'OBP envoie une demande de redémarrage au SC.
3. Le démon de détection d'événements (dsmd ou esmd) détecte la demande.
4. le démon fofd consigne l'événement.
5. fofd redémarre le domaine La demande de redémarrage peut être lancée avec les commandes reboot, shutdown ou init.

4.5.2 Panic

Un arrêt dû à une condition exceptionnelle (*panic*) se produit lorsque le système d'exploitation détecte un problème irréparable. Il peut également être forcé depuis le SC. L'arrêt peut être causé par un problème matériel ou logiciel. Analysez l'historique des messages de *panic* pour en déterminer la cause.

Lorsqu'une telle condition est détectée, le traitement automatique suivant est réalisé :

1. Le noyau enregistre un *dump* de lui-même dans la partition de *swap* principale.
2. dsmd ou esmd détecte le *panic* (notification envoyée à partir du démon hwad).
3. dsmd ou esmd consigne le *panic*.
4. fod redémarre le domaine avec hpost.
5. Les fichiers vmcore.n et unix.n sont créés par savecore dans le fichier `/var/crash/nom_domaine` du domaine

Les *dumps* de *panic* peuvent être analysés avec crash et kadb. Vous pouvez également utiliser le script *ISDA* de SunSolve®.

Remarques :

- La commande `savecore` doit être activée dans `/etc/init.d/sysetup` pour enregistrer le *crash dump*. Par défaut, les *dumps* de *panic* ne sont pas enregistrés.
- La partition de *swap* principale doit être dimensionnée de manière à contenir le *dump* brut.
- Le fichier `/var` doit être suffisamment grand pour contenir le *dump* transféré à partir de la zone de *swap* par `savecore`. Rappelez-vous que `/var` peut contenir plusieurs *dumps* simultanément.

4.5.3 Réinitialisations par *watchdog*, *redmode* ou *xir*

Ces conditions sont définies dans le *Manuel de Référence d'Architecture SPARC™* et font parties des fonctions du processeur SPARC.

Elles sont invoquées quand le processeur détecte certaines conditions d'erreur.

- Une réinitialisation par *watchdog* se produit lorsqu'une interruption est rencontrée et que le noyau est déjà dans une section critique du code gestionnaire des interruptions. Bien que l'architecture Ultra rarefie les réinitialisations par *watchdog*, elles peuvent encore survenir.
- Un *redmode* est un événement se produisant lorsque le système a reçu autant d'interruptions simultanées qu'il peut en gérer.
- Une *xir* est une réinitialisation lancée de manière externe. Elle se produit lorsqu'un signal spécifique est envoyé au processeur SPARC™. Ce signal peut être envoyé à partir du SC par la commande `reset -x`.

Ces trois conditions entraînent l'arrêt du domaine. Lorsque l'une d'entre elles est détectée, le traitement automatique suivant est déclenché :

1. `osd` sauvegarde les registres des processeurs dans un emplacement matériel spécial.
2. `dsmd` ou `esmd` consigne l'événement dans le SC.
3. `dsmd` lit les informations du registre et crée un fichier ASCII `resetinfo` dans `$SMSVAR/adm/nom_domaine/resetinfo.timestamp`
4. `edd` exécute ensuite un `hpost` et redémarre le domaine.

Les fichiers `resetinfo` sont en ASCII et leur affichage ne nécessite pas l'emploi du programme de *debuggage* `redx`. Ils contiennent les registres des processeurs tels qu'ils étaient au moment de la défaillance.

4.5.4 Hôte bloqué (*heartbeat failure*)

Un *heartbeat failure* se produit en l'absence de détection, par le démon `cbd`, de la pulsation des processeurs d'un domaine. Ceux-ci mettent à jour régulièrement le registre spécifique (*heartbeat register*), indiquant qu'ils sont actifs. Si ce registre n'est pas en cours de mise à jour, le domaine est considéré bloqué et un traitement de récupération est démarré.

Lorsqu'un *heartbeat failure* est détecté, le traitement automatique suivant est déclenché :

1. Le démon de détection d'événements (`dsmd`) détecte et consigne le blocage (une notification est envoyée à partir du démon `hwad`).
2. `dsmd` force le *panic* du domaine.
3. `dsmd` redémarre le domaine .
4. Intervention manuelle sur un hôte bloqué :

Un événement logiciel, ou autre, peut simuler un *heartbeat failure* alors que le registre de pulsation est en cours de mise à jour. Si ceci se produit, vous devrez intervenir manuellement pour lancer le traitement de récupération suite au problème.

La liste suivante présente les actions de récupération possibles ; elle a été définie en fonction de l'ensemble des informations d'état enregistrées. Essayez les actions séquentiellement, ne pas poursuivre après l'obtention d'un résultat.

- Tentez une nouvelle ouverture de session pour déterminer si le domaine répond.
- Tentez de rentrer de force dans l'OBP du domaine à partir de votre session *console*, et demandez un *dump* de *panic* :

```
~# (Forçage vers l'OBP)
ok sync
```

- Forcez un *panic* à partir du SC. Un *dump* de *panic* sera créé lorsque vous entrez la commande suivante :

```
sms-svc> reset -d domain_id
```

- Forcez une réinitialisation *xir* à partir du SC. Un fichier `resetinfo` sera créé lorsque vous entrez la commande suivante :

```
sms-svc> reset -d domain_id -x
```

- En dernier recours, forcez un *setkeyswitch* à partir du SC. Aucune information de défaillance ne sera enregistrée. Entrez la commande suivante :

```
sms-svc> setkeyswitch -d domain_id standby
sms-svc> setkeyswitch -d domain_id on
```

4.5.5 Arbstop

Un *arbstop* se produit lorsque l'équipement d'interconnexion de la carte centrale du système (*Centerplane board*) détecte une erreur. Celle-ci peut être liée au Code Correcteur d'Erreur (ECC), à la parité, à un dépassement de délai d'octroi, à un dépassement de file ou à tout autre problème. Généralement, les *arbstops* et les *recordstops* associés affectent uniquement le domaine dans lequel ils se produisent, à l'exception des erreurs se produisant dans la carte centrale qui affectent tous les domaines.

Si une erreur se produit localement sur une carte système (*System Board*), la demande d'arrêt d'arbitrage est diffusée à l'ensemble des cartes système du domaine, qui s'arrêtent. Tous les autres domaines de la plate-forme continuent à tourner. Si le problème n'est pas propre à une carte système mais survient, par exemple, sur la *Centerplane board*, tous les domaines s'arrêtent. Lorsqu'un *arbstop* se développe, le traitement suivant s'exécute :

1. dsmd consigne l'événement *arbstop*.
2. dsmd capture les informations matérielles en exécutant :
3. hpost -D
4. dsmd lance ensuite un redémarrage du domaine en exécutant un hpost :

Les fichiers binaires *arbstop dump* nécessitent l'utilisation du programme de *debuggage* redx pour leur interprétation.

4.5.6 Domainstop et recordstop

Un *domainstop* (*dstop*) se produit lorsque le démon de détection des événements hardware (hwad) détecte une erreur non recouvrable sur un domaine.

Un *recordstop* (*rstop*) se produit lorsque le démon de détection des événements hardware (hwad) détecte une erreur recouverte (corrigé). Il peut s'agir d'une erreur mémoire ECC, d'une *data interconnect ecc error* ou de façon plus générale d'une erreur corrigée par les mécanismes ECC.

Les erreurs *dstop* et *rstop* n'impactent normalement que le domaine propriétaire de l'élément défectueux.

En cas de détection d'une *dstop* une requête *dstop* est « broadcastée » aux *system boards* du domaine qui sont alors stoppés. Les autres domaines ne sont pas impactés. Le system controller actif log alors l'erreur.

- Les dumps *dstop* sont créés dans le répertoire suivant :
\$SMSLOGGER/domain_name/dsmd.dstop.yymmdd.hhmmss
- Les dumps *rstop* sont créés dans le répertoire suivant :
\$SMSLOGGER/domain_name/dsmd.rstop.yymmdd.hhmmss
- Les erreurs détectées durant les *hpost* sont créés dans le répertoire suivant :
\$SMSLOGGER/domain_name/xcstat.yymmdd.hhmmss

nota: la date indiquée est la date du SC.

Les fichiers de dump sont visualisable via la commande *redx*.

La taille de dump d'une SB est d'environ 4-5 Kbytes.

4.6 Domaine bloqué

Ce chapitre décrit la démarche à adopter lorsqu'un domaine est bloqué.

Prérequis

- Etre *loggé* sur le SC en tant qu'utilisateur *sms-svc* avec deux fenêtres « terminal ».
- Utiliser les commandes *showboards* et *showplatform* pour pointer sur le bon domaine.
- Utiliser l'une des deux sessions comme console système sur ce domaine au moyen de la commande *console*.
- Utiliser l'autre session pour toutes les commandes SC.
- Utiliser *ps -elf* pour rechercher les processus gros consommateurs de *cpu*, *df -lk* pour vérifier les taux d'utilisation des systèmes de fichiers et *who* pour déterminer qui sont les utilisateurs actuellement *loggés* et quels processus ils exécutent.

Exécutez la procédure

Vous devez engager cette procédure :

- Si vous êtes incapable de joindre un domaine à partir du SC au moyen de la commande *ping*,
- Si vous êtes incapable de générer une session *telnet* sur le domaine.

Le tableau suivant décrit les actions à réaliser.

Tableau 9 : Procédure de restauration d'un domaine bloqué

N°	Actions	Description
1	% showboards % showplatform	Entrez les commandes SC ci-contre et sauvegardez le résultat.
2	% reset -d domain_id -x	Essayez de forcer le domaine à retourner sous l'OBP en entrant <i>reset</i> . Observez l'activité de la session <i>console</i> . La commande <i>reset</i> aboutit si le prompt <i>ok</i> de l'OBP apparaît.
3	<i>Action à effectuer sous l'interface graphique.</i>	Si la commande <i>reset</i> a abouti avec succès, exécutez la séquence de commandes 4 à 7. Quand c'est terminé, sauvegardez le contenu du tampon de la fenêtre dans un fichier ou effectuez un copier/coller vers un fichier. Ces données sont utiles pour l'analyse de la cause du blocage du domaine.
4	ok ctrace	Cette commande vous délivrera une trace avant le lancement de l'OBP.
5	ok .registers	Cette commande génère un <i>dump</i> des registres globaux à l'instant du lancement de l'OBP.
6	ok .local	Cette commande génère un <i>dump</i> des registres locaux à l'instant du lancement de l'OBP.

-
- | | |
|--|--|
| 7 ok sync | Sync émet un appel en retour vers le noyau afin d'obtenir un <i>dump</i> du noyau. Le système devrait effectuer ce <i>dump</i> puis redémarrer (<i>reboot</i>) après l'émission de cette commande OBP. |
| 8 <i>Laissé en blanc intentionnellement.</i> | Si la commande <i>reset</i> n'a pas fonctionné, essayer de forcer le <i>panic</i> du système au moyen de la commande <i>reset -x</i> sur le SC.

Si tout a échoué, émettre une commande <i>setkeyswitch standby</i> suivi d'une commande <i>setkeyswitch on</i> pour remettre le domaine en opération. |
-

4.7 Informations techniques pour le support

Ces informations, qui doivent être collectées dès la constatation du problème, avant l'ouverture d'un appel support, se scindent en informations générales et informations propres à un problème donné.

4.7.1 Informations générales

Pour les informations de configuration système, récupérez les sorties obtenues, pour chaque domaine, par les commandes `prtdiag`, `sysdef` et `showrev -p`.

4.7.2 Informations propres à un problème donné

- Indiquez si possible le type de problème, tel que *panic*, *arbstop*, *domainstop*, réinitialisation par *watchdog*, blocage, erreur disque...
- Indiquez la sévérité du problème, tel que : « système arrêté » ou « système partiellement déconfiguré ».
- Répertoriez toutes les modifications système récentes qu'elles soient matérielles, logicielles ou liées à l'adjonction de patchs et de nouvelles applications ; utilisez à cette fin les feuillets présentés en *Annexe B – Cahier de Suivi*
- Lancez un *explorer* sur les domaines et les SC en se *loggant root* puis en entrant la commande : `# /opt/SUNWexplo/explorer` et tenir à disposition les résultats ou , à défaut, observez les recommandations ci-dessous.

- tenez à disposition toutes les informations pertinentes sur les fichiers résidents du SC, duplicata des fichiers `/var/adm/messages` des domaines :

```
/var/opt/SUNWSMS/adm/nom_domaine/messages  
/var/opt/SUNWSMS/adm/messages, et  
/var/opt/SUNWSMS/var/nom_domaine/post
```

- récupérez l'historique des consoles SC, et, lorsque cela est possible, les fichiers :
`.postrc` et `/var/opt/SUNWSMS/etc/nom_plate-forme/blacklist`.
- récupérez les informations du fichier `/etc/system` ainsi que toutes informations pertinentes du fichier `/var/adm/messages`.
- pour les *arbstop*, tenez à disposition le contenu du fichier :
`/var/opt/SUNWSMS/adm/nom_domaine/arbstopdump` sur le SC.
- pour les réinitialisations dues à un *watchdog*, mettez à disposition le contenu du fichier :
`/var/opt/SUNWSMS/adm/nom_domaine/resetinfo` sur le SC.

pour un *panic*, tenez à disposition le contenu du *crash dump* du domaine hôte SunFire 25K récupéré dans `/var/crash/nom_domaine`.