

Guide de configuration de Sun StorEdge™ Fast Write Cache 2.0



THE NETWORK IS THE COMPUTER™

Sun Microsystems, Inc.
901 San Antonio Road
Palo Alto, CA 94303-4900 USA
650 960-1300 Fax (650) 969-9131

Référence 806-5482-10
Août 2000, Révision A

Envoyez tout commentaire relatif à ce document à docfeedback@sun.com

Copyright 2000 Sun Microsystems, Inc., 901 San Antonio Road • Palo Alto, CA 94303 Etats-Unis. Tous droits réservés.

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

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

Sun, Sun Microsystems, le Sun logo, NFS, The Network Is The Computer, Sun StorEdge, Fast Write Cache, SunVTS, Sun Enterprise Volume Manager, Solstice DiskSuite, Sun Enterprise, et Solaris sont des marques de fabrique ou des marques déposées, ou marques de service, de Sun Microsystems, Inc. aux Etats-Unis et dans d'autres pays.

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

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

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



Guide de configuration de Sun StorEdge Fast Write Cache

Introduction

Sun StorEdge™ Fast Write Cache est un accélérateur d'écriture résidant sur l'hôte et destiné aux systèmes de stockage A5X00. Il accélère le traitement des transactions et assure des temps de réponse plus rapides aux requêtes de données des utilisateurs en réduisant la fréquence des opérations de lecture et d'écriture sur disques. Les écritures sont stockées dans une mémoire cache rémanente, puis elles sont reportées sur le disque. Fast Write Cache s'installe sur des serveurs Solaris™ et comprend des cartes de mémoire vive rémanente (NVRAM), qui font office de mémoire cache, et un logiciel de gestion de cache de stockage.

Remarque - Fast Write Cache est mis en oeuvre sur le serveur Solaris à l'aide d'une paire de cartes NVRAM SBus ou PCI et d'un logiciel de gestion de cache de stockage. Chaque système supporte une paire de cartes, sauf le système Enterprise 10000, qui supporte une paire de cartes par domaine.

Sommaire

Le présent document contient l'information suivante :

- Documentation connexe
- Applications

- Architecture
- Considérations sur la configuration
- Performances
- Exemples de configuration

Renseignements sur l'installation

Pour des informations sur les aspects suivants de l'installation, reportez-vous au *Guide d'installation de Sun StorEdge Fast Write Cache 2.0*

- Configuration système requise
- Plateformes compatibles
- Compatibilité logicielle
- Restrictions

Documentation connexe

Application	Titre	Référence
Pages du manuel	fwcadm(1FWC) svadm(1SV)	S.O.
Notes de version	<i>Notes de version de Sun StorEdge Fast Write Cache 2.0</i>	806-5497-10
Installation	<i>Guide d'installation de Sun StorEdge Fast Write Cache 2.0</i>	806-5492-10
Utilisateur	<i>Guide de l'administrateur système de Sun StorEdge Fast Write Cache 2.0</i>	806-5487-10
Livres blancs	<i>Sun StorEdge Fast Write Cache Technical White Paper</i>	
Guides des meilleures pratiques	<i>Improving OLTP Performance on the Sun StorEdge A5000</i>	

Applications

Le modèle A5X00 est parmi les plus performants qui soient pour tout un éventail d'applications, notamment l'entreposage de données, l'aide à la décision et le traitement des images. Sun StorEdge Fast Write Cache permet à Sun Microsystems™ d'étendre les performances du A5X00 à des applications comme OLTP et les services de fichiers.

En effet, des performances jusqu'à 40 fois supérieures ont été mesurées selon la taille des transferts, le nombre d'unités d'exécution et d'autres facteurs. Les meilleures améliorations observées concernent les écritures séquentielles, lorsque la mémoire cache fusionne les petites écritures en plus grosses. Sous RAID 5, ces écritures occupent toute la bande, ce qui élimine l'opération lecture-modification-écriture.

Grâce à Fast Write Cache, l'administrateur système peut sélectionner les volumes qui sont mis en mémoire cache et ceux qui ne le sont pas. L'administrateur système devrait mettre en cache :

- les données NFS™,
- les données utilisées par les applications OLTP.

L'administrateur système ne devrait pas mettre en cache :

- les données utilisées par les applications d'aide à la décision,
- les volumes en lecture seule.

Architecture

Sun StorEdge Fast Write Cache sert de pilote de périphérique UNIX® utilisant une mémoire rémanente pour mettre en cache les requêtes d'écriture. Au fur et à mesure que la mémoire cache se remplit, les données les plus anciennes sont écrites de façon asynchrone sur le disque réel. Fast Write Cache fonctionne comme une couche entre les autres pilotes de disque et le reste du noyau UNIX. Des éléments de remplacement se substituent aux points d'entrée du pilote d'origine dans les tables de commutation de périphérique. Lorsque Fast Write Cache effectue des opérations d'E-S réelles (par exemple lorsque le contenu de la mémoire cache doit être reporté sur le disque), il utilise les sous-programmes du véritable pilote.

Fast Write Cache figure au-dessus de Volume Manager, ce qui lui permet de prendre en charge tout volume RAID 5, RAID 1 ou RAID 0+1.

Composants matériels

Fast Write Cache 2.0 comprend deux cartes NVRAM PCI ou SBus redondantes d'une capacité de 32 Mo. Elles sont dotées de fonctions de détection et de correction d'erreurs à un bit ainsi que d'une fonction de détection d'erreurs à deux bits. Si une erreur à deux bits survient, le logiciel met la carte hors ligne, reporte le contenu de la cache sur le disque et fait passer celle-ci en mode d'écriture immédiate. Durant les opérations normales d'E-S, tous les accès sont des écritures. Durant les récupérations (après un arrêt anormal), les blocs sont lus de la NVRAM.

Carte NVRAM SBus

Chaque carte est de largeur simple et comporte deux cartes mezzanines et trois piles au lithium. La durée de vie des piles est de 10 ans. Les piles peuvent conserver des données pendant un maximum de 18 mois. Chaque pile est activée par un cavalier placé sur la carte.

Les performances mesurées de la carte en écriture sont de 108 Mo/s.

La carte NVRAM est dotée d'une DEL alimentée par le système et non par les piles.

La DEL s'allume dans les cas suivants :

- quand la mémoire cache est activée,
- quand un arrêt anormal survient pendant que la mémoire cache est activée.

La DEL est éteinte dans les cas suivants :

- quand la mémoire cache est désactivée,
- quand la carte n'est pas installée,
- quand la carte est installée mais que le système est hors tension.

Carte NVRAM PCI

La carte NVRAM PCI est une carte pleine longueur d'une capacité de 64 Mo et qui comporte quatre piles. La durée de vie des piles est de 10 ans. Elles peuvent conserver les données pendant environ 34 mois. Les piles sont activées par un commutateur placé sur la carte. Le débit de transfert de données peut atteindre 90 Mo par seconde.

Composants logiciels

Fast Write Cache comprend les composants logiciels suivants :

- Pilote Storage Volume (SUNW`spsv`),

- Gestionnaire Storage Cache (SUNWscm),
- Pilote de NVRAM (SUNWnvm),
- Diagnostic (SUNWvt.snp),
- Compte rendu d'état uniforme (SUNWspuni).

Pilote Storage Volume (SUNWspsv)

Le pilote Storage Volume est un pilote hiérarchisé qui intercepte les appels d'E-S envoyés aux périphériques dotés d'une mémoire cache. Il copie ensuite les données depuis et vers les tampons de l'environnement d'exploitation Solaris dans les tampons intégrés à la cache.

Gestionnaire Storage Cache (SUNWscm)

Le gestionnaire Storage Cache contient le logiciel de cache et exécute les fonctions suivantes :

- fusion,
- report sur le disque,
- activation et désactivation du mode d'écriture directe,
- données marquées par le gestionnaire, soit les données qui se trouvent dans la mémoire cache et qui ne peuvent pas être reportées sur le disque (dans le cas d'une panne de disque par exemple).

Pilote de NVRAM (SUNWnvm)

SUNWnvm contient le logiciel requis pour prendre en charge les cartes NVRAM et gère l'écriture miroir entre les deux cartes NVRAM redondantes. Il intègre également le logiciel qui contrôle l'état de la mémoire, les piles et l'état général des cartes. S'il détecte une défaillance d'une carte, il avertit le gestionnaire Storage Cache, qui commande alors le passage au mode d'écriture directe.

Diagnostic (SUNWvt.snp)

SUNWvt.snp comporte des fonctions de diagnostic pour les cartes NVRAM.

Compte rendu d'état uniforme (SUNWspuni)

Ce module fournit des rapports d'erreurs aux services de données.

Considérations sur la configuration

Décision quant à l'utilisation de la mémoire cache

Avec Fast Write Cache, l'administrateur système peut sélectionner les volumes qui sont mis en mémoire cache et ceux qui ne le sont pas. Il s'agit de la formule idéale : accès direct aux disques haute vitesse lorsque nécessaire et mise en mémoire cache au besoin.

Puisqu'il est en mesure d'insérer et de retirer dynamiquement des volumes de la mémoire cache, l'administrateur peut mettre les volumes en mémoire cache lorsque cette méthode convient et accéder directement au disque si nécessaire.

Mise en mémoire cache

- Mettez en mémoire cache les données des applications qui effectuent de petites écritures séquentielles :
 - Journaux de base de données,
 - Données OLTP,
 - Données du système de fichiers.
- Mettez en mémoire cache les volumes RAID 5.

Utilisation déconseillée de la mémoire cache

Ne mettez pas en mémoire cache les éléments suivants :

- Périphériques de gestion de volume (par exemple, périphériques de consignment VxVM), car les résultats seraient alors imprévisibles.
- Systèmes de fichiers racine (/) et /usr (ou tout système de fichiers placé avant Fast Write Cache) ; pour pouvoir récupérer les données de disque stockées sur une carte NVRAM après une panne système, ces données doivent être restaurées sur le disque avant le montage des systèmes de fichiers ou l'enregistrement des applications sur les volumes bruts.
- Partition de mémoire virtuelle.
- Données d'applications qui effectuent d'importantes écritures asynchrones.
 - Exemple : applications d'aide à la décision.
- Volumes en lecture seule.

Volumes bruts ou mis en mémoire cache

Les volumes peuvent être mis en cache selon le mode *brut* ou *cache*. En mode *cache*, les E-S se terminent lorsque les données sont stockées dans la mémoire NVRAM. En mode *brut*, l'accusé de réception de fin des E-S n'est pas envoyé tant que les données ne se trouvent pas sur le support physique, ce qui permet aux volumes d'utiliser l'interface de la cache sans être placés en mémoire cache. Cette condition est obligatoire si l'administrateur souhaite mettre le volume à la disposition des services de données (comme Instant Image) sans le mettre en cache ou, encore, s'il n'a pas installé les cartes NVRAM. Si les cartes NVRAM ne sont pas installées ou ne fonctionnent pas, la mémoire cache fonctionne en mode d'écriture directe. Autrement dit, l'accusé de réception de fin des E-S n'est pas transmis à l'application tant que les données n'ont pas été stockées sur les disques matériels.

Performances

Ce sont les applications effectuant des écritures séquentielles qui bénéficient du meilleur gain de performances avec Fast Write Cache. Le tableau ci-après indique le gain de performances des écritures séquentielles à une unité d'exécution. Il montre les différences entre deux lecteurs A5000 à 7 200 tr/min, dont un est mis en mémoire cache.

TABLEAU 1 Performances comparées des écritures séquentielles

E-S séquentielles	Débit	IOPS	Temps de réponse
2 ko sans cache	187,93 ko/s	94,04	11 ms
2 ko avec cache	9 684,09 ko/s	4 843,98	0 ms ⁽¹⁾
8 ko sans cache	663,06 ko/s	82,92	12 ms
8 ko avec cache	10 550,14 ko/s	1 319,34	1 ms
16 ko sans cache	1 169,21 ko/s	73,11	14 ms
16 ko avec cache	10 567,28 ko/s	660,79	1 ms
256 ko sans cache	6 320,23 ko/s	24,71	40 ms
256 ko avec cache	10 822,59 ko/s	43,30	24 ms

(1) Le temps de réponse est inférieur à 500 microsecondes et arrondi à 0,000 seconde, ce qui donne 0 après la conversion en millisecondes.

Pour en savoir davantage sur les performances de ce produit, reportez-vous au Sun StorEdge Fast Write Cache Technical White Paper.

Exemples de configuration

Application	Configuration matérielle	Charge de travail	Taille des E-S	Configuration RAID	Cache
OLTP	E3000	75 % de lectures	2 ko	RAID 1	Journaux d'annulation
	A5000	25 % d'écritures			
	E4500	60 % de lectures	8 ko	RAID 1	Volumes sur lesquels les écritures ont lieu
	A5200	40 % d'écritures			
Aide à la décision	E6500	90 % de lectures	64 ko	RAID 5	Ne pas mettre en cache
	A5200	10 % d'écritures			