



System Management Services Software: An Inside Look

Thomas Chalfant, Enterprise Server Products

Sun BluePrints™ OnLine—January 2003



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

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95045 U.S.A.
650 960-1300

Part No. 817-1659-10
Revision 1.0, 1/13/03
Edition: January 2003

Copyright 2003 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, California 95045 U.S.A. All rights reserved.

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

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

Sun, Sun Microsystems, the Sun logo, Sun BluePrints, Sun Fire, Sun Enterprise, and Solaris are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, 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 licensees. Sun acknowledges the pioneering efforts of Xerox in researching and developing the concept of visual or graphical user interfaces for the computer industry. Sun holds a non-exclusive license from Xerox to the Xerox Graphical User Interface, which license also covers Sun's licensees who implement OPEN LOOK GUIs and otherwise comply with Sun's written license agreements.

U.S. Government Rights—Commercial use. Government users are subject to the Sun Microsystems, Inc. standard license agreement and applicable provisions of the Far and its supplements.

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

Copyright 2003 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, Californie 95045 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 enregistrée aux Etats-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company Ltd.

Sun, Sun Microsystems, le logo Sun, Sun BluePrints, Sun Fire, 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. Toutes les marques SPARC sont utilisées sous licence et sont des marques de fabrique ou des marques déposées de SPARC International, Inc. aux Etats-Unis et dans d'autres pays. Les produits portant les marques SPARC sont basés sur une architecture développée par Sun Microsystems, Inc.

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

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



Please
Recycle



Adobe PostScript

System Management Services Software: An Inside Look

The Sun Fire™ 15K server is monitored and controlled by a system controller (SC) that runs System Management Services (SMS) software in the Solaris™ Operating Environment (Solaris OE). SMS is highly integrated into the Solaris OE and uses several of its features while performing its role for the platform.

This Sun BluePrints™ OnLine article addresses some of the more advanced topics of SMS software including the Management Network (MAN) and SMS security. In addition, it provides insight to a new security feature that is available through a patch to SMS version 1.2 and that is integrated in follow-on releases such as the upcoming SMS version 1.3. This new feature provides the much requested capability to use secure shell for file synchronization between SCs.

This article contains the following sections:

- “Understanding the Management Network (MAN)” on page 2
- “Managing Complete Link Failure” on page 5
- “Understanding SMS Security Design” on page 7
- “Using `ssh` for SC to SC File Propagation” on page 11

Understanding the Management Network (MAN)

The Sun Fire 15K server has two SCs within the platform cabinet. One of the SCs is designated as the main SC, and the other is designated as the spare SC. These two SCs must be kept synchronized with one another and must be apprised of each others' statuses. The Sun Fire 15K server allows the hardware to be partitioned into one or more environments that are capable of running separate images of the Solaris OE, commonly referred to as domains. With this in mind, the main SC has the following additional functions:

- Controlling dynamic reconfiguration (DR)
- Providing consoles for each domain
- Recording message logs for each domain
- Assisting with the installation of Solaris OE domains
- Synchronizing time between the SC and domains

These functions are implemented over a network that is internal to the platform chassis called the Management Network or MAN. The MAN is not a general-purpose network and should only be used for its intended purpose. To that end, neither the domains themselves, nor the SC will route traffic to these networks by default, other than as mentioned above.

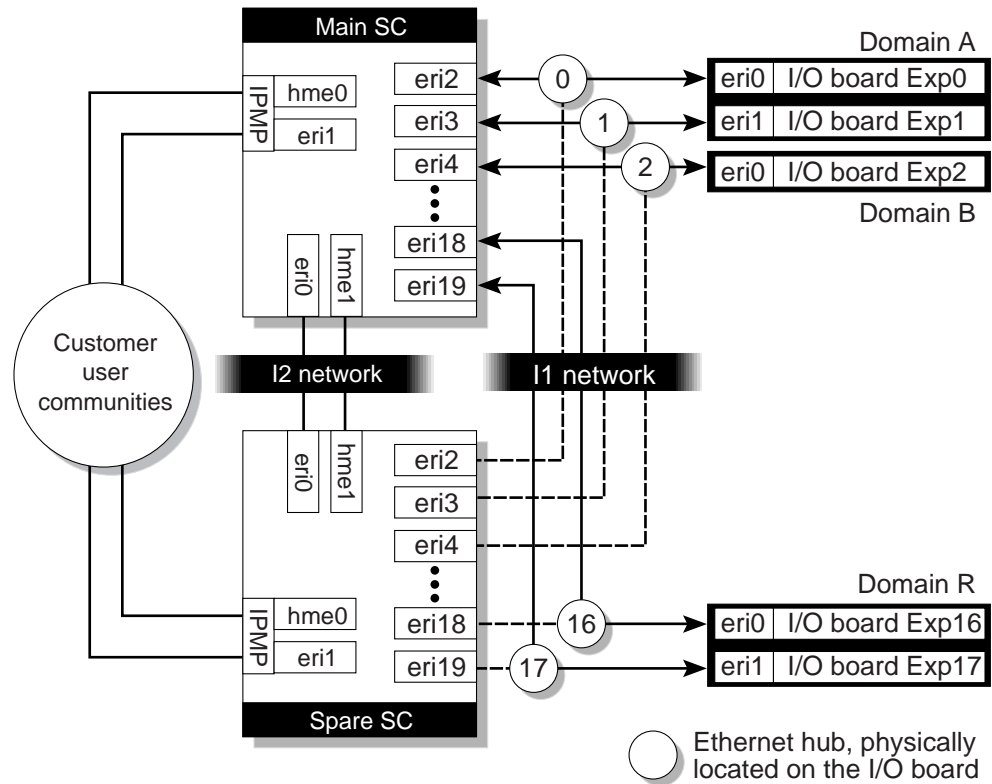


FIGURE 1 MAN Overview

The SC-to-domain network is referred to as the I1 network. It is constructed from 18 separate 100BASE-T network interface controllers (NICs) on each SC, which then connect to an Ethernet hub on each input/output (I/O) board, forming a point-point network between the SC and each I/O board. Because the Ethernet hub has connections from each SC, it is programmable so that only the NIC connecting to the main SC is active. These networks operate at half-duplex because the hub on the I/O board does not support full-duplex transfers.

Between the two SCs themselves, another network referred to as the I2 network exists. This network operates at 100-megabit full-duplex and does not involve the use of hubs.

I1 MAN Functions

- Domain consoles
- Message logging
- Dynamic reconfiguration
- Network boot/Solaris OE installation
- Time synchronization

I2 MAN Functions

- SC heartbeat
- File synchronization
- Failover

The MAN drivers on the SC create meta-interfaces for individual NICs to reduce complexity and administrative overhead. The drivers also implement the concept of a community network that can be monitored by the SC and that can be configured in a highly available manner by using both external NICs on the SC. A floating IP address is created and owned by the main SC, as an aid to reach the main SC for external clients that might not have specific knowledge of which SC is performing a certain role at a given time.

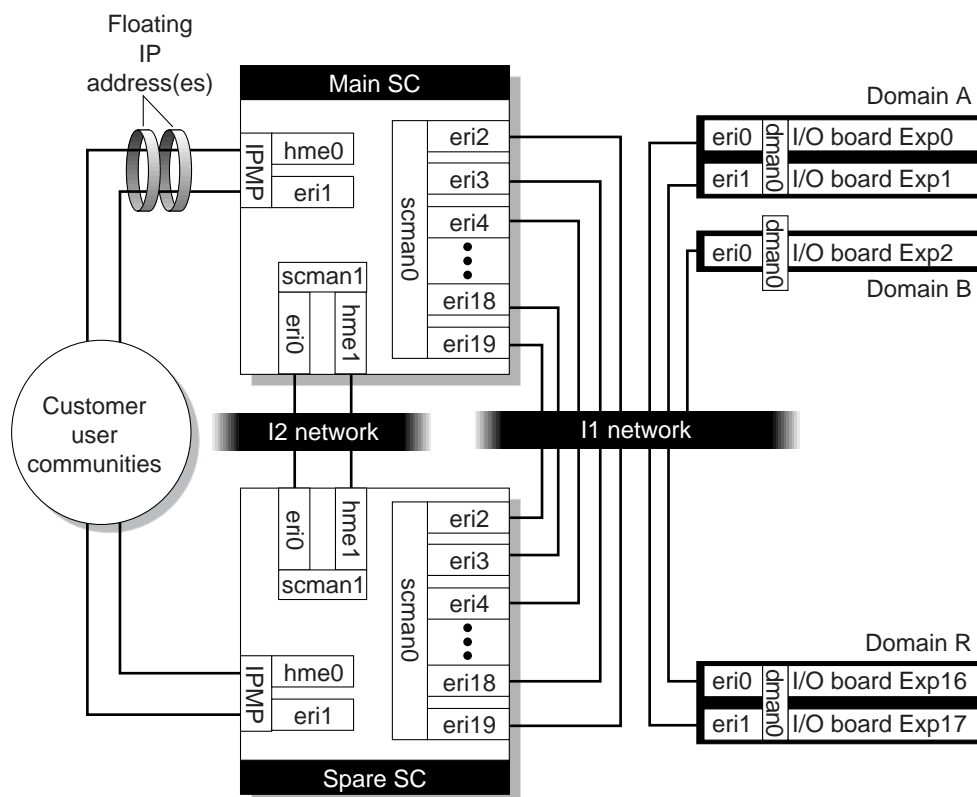


FIGURE 2 Simplified MAN Overview

Note that it is possible for there to be multiple interfaces on the domain side of the I1 network. A single meta-interface represented by `scman0` on the SC is created. It reacts to path failures and automatically switches the active network path, provided that one exists. It also enforces domain isolation, keeping domain traffic exclusively

between the domain and SC, while making various point-point links appear as a normal Ethernet network. On the domain side of the network, there is a corresponding meta-interface called `dman0`.

The I2 network has its own meta-interface called `scman1`. Unlike I1 network `scman0` interfaces, `scman1` interfaces are active on both SCs. The `scman1` network has two possible paths, and the network driver will detect path failures and switch paths automatically just as the `scman0` meta-interface does.

Automatic path switchover is handled as follows:

- **I1 network.** By default, a domain's active NIC is on the I/O board that contains the *golden* input/output static random access memory (IOSRAM). IOSRAMs are located on every I/O board, with the golden IOSRAM acting as the master IOSRAM, hence the term golden. This is typically the lowest numbered I/O board in the domain; however, this may not always be the case. `dman0` pings the SC through the active NIC every 10 seconds. Every 30 seconds, `dman0` checks the inbound packet count. If the packet count has increased, the connection is considered good. If not, a path switch is started and the next available path is selected.
- **I2 network.** By default, the active NIC is `eri0`, but this may change in the future. `scman1` on the *main* SC pings the *spare* SC through the active NIC every 10 seconds. Every 30 seconds, `scman1` on the main SC checks the inbound packet count. If the packet count has increased, the connection is considered good. If not, the active path is switched to the other NIC, provided it was not previously marked as failed.

Managing Complete Link Failure

If all available links for a domain on the I1 network fail, domain consoles and domain initiated DR can still successfully complete because they can operate over a backup mechanism called the IOSRAM. IOSRAM is not discussed here. Services that become unavailable when all links fail include time synchronization, SC initiated DR, message logging from the domain, and network booting from the SC.

If all links on the I2 network fail, file synchronization between SCs is the most affected function. Heartbeat and failover still function, because they do not solely rely on the I2 network.

When troubleshooting problems concerning the MAN network, it might be desirable to force a switch to another path. Perhaps intermittent problems on a given network path are causing a degradation in connectivity, but not enough to induce an automatic path switch. MAN provides a method for forcing a path change through `ndd`.

I1 Network Path Switch

For the I1 network, you can force a path switch from either the domain or the SC. Switchover from the domain is preferred because `dman` drives path recovery on the I1 network. Therefore, if the change is made to `dman` on the domain, it is immediately propagated to `scman` on the SC.

To perform the path switch, use the `ndd` commands:

```
domainA# ndd /dev/dman man_pathgroups_report
MAN Pathgroup report: (* == failed)
=====
Interface  Destination  Active Path  Alternate Paths
-----
dman0      Master SSC   eri0         eri0 exp 6, eri2 exp 12, eri3 exp 14
```

Now the first command is complete. We've listed available paths. Now let's set the path as we desire. See the new command. This command produces no output.

```
domainA# ndd -set /dev/dman man_set_active_path 0 0 2
```

The first parameter is the instance of the driver. The second parameter is the number of the domain. The third parameter is the `eri` instance to set active. For the domain side, the first and second values are always 0. The key parameter is the third number passed to `man_set_active_path`.

This indicates which `eri` instance to make active. In the prior example, `eri2` is made active. The path change is reflected immediately to `scman0`. Furthermore, the path that was previously active is marked failed (*). (If that NIC is in fact healthy, `dman` will clear the failed flag as part of its normal link policing.) Performing manual path failover on the `scman0` driver from the SC is similar but not discussed here.

I2 Network Path Switch

The I2 network has similar manual path switch commands.

To perform the path switch, use the `ndd` commands:

```
sc# ndd /dev/scman man_pathgroups_report
MAN Pathgroup report: (* == failed)
=====
Interface  Destination                Active Path  Alternate Paths
-----
scman      C 8:0:20:be:f4:ed          eri8        eri8 exp 6, eri14 exp 12, eri16 exp 14
scman0     A 8:0:20:b7:2f:20          eri4        eri4 exp 2
scman1     Other SSC                  eri0        exp 0, hme1 exp 0
```

The first command is complete and we listed the available paths. Now let's set the path as we desire. See the new command. This command produces no output.

```
sc# ndd -set /dev/scman man_set_active_path 1 0 1
```

The first parameter is the instance of the driver, the second is the number of the domain, and the third is the `eri` instance to set active. For the I2 network, the first parameter is always 1 and the second is always 0. The key parameter is the third number passed to `man_set_active_path`. It indicates which NIC to make active (0 = `eri0`, 1 = `hme1`).

Above, `hme1` has been made active. There is a small delay while the main and spare SCs `scman` drivers converge on the new active path. As with the I1 network, the previously active path is marked failed (*).

Typical network troubleshooting techniques such as using `ping`, `netstat`, and `snoop` are effective for debugging problems with the I1 and I2 MAN networks.

Understanding SMS Security Design

Along with the hardware design for allowing partitioning into dynamic system domains, SMS software enables strict domain separation on the system controller. For instance, the privileges for administrators of one domain can be configured to prevent them from viewing the console output of another domain.

In my experience, I have not seen many data centers actually use this capability. Some people state it is easier to have a single login with full permissions, as is the case with the model provided by the Sun Enterprise™ 10000 server system service processor (SSP). Certainly, this model can be implemented on the Sun Fire 15K server SC. Other people state an unawareness of the capability itself or a lack of

understanding of what it can do for them. In today's more security-conscious environment, many customers are reevaluating past decisions regarding security practices. The next several paragraphs discuss security features of SMS software, delve into how it works behind the scenes, and further clarify the rationale for use.

Administrative Privileges

SMS splits functionality into domain and platform administrative privileges, along with a subset of privileges for platform operators and domain configurators. For instance, platform administrators are provided the ability to configure the platform, including assigning boards to domains, getting environmental status, and other administrative tasks. Domain administrators can access the console of the respective domain and perform reconfiguration on the boards assigned to it. However, they cannot perform global configuration as performed by the platform administrator. For a complete description of the various roles and their capabilities, refer to the section on SMS security in the *System Management Services (SMS) 1.2 Administrator Guide*.

How does SMS split administrative privileges on the SC? Several Solaris OE features come into play to allow this to function properly. In particular, SMS uses Solaris OE groups and access control lists.

Example

Let's begin by taking a look at what user `jim` can do on a given SC. Logging into the SC as user `jim`, we can see what this user can do.

```
f15k-sc0:jim:2> groups
staff platadm platsvc dmnaadm
```

User `jim` is a member of several groups, including the platform administrator group, the platform service group, and the domain A administrator group. Because `jim` is part of the domain A administrator group, let's snoop around a bit on the SC, and see what we find.

```
f15k-sc0:jim:15> cd $SMSVAR/adm
f15k-sc0:jim:16> pwd
/var/opt/SUNWSMS/SMS1.2/adm
f15k-sc0:jim:17> ls -ld A
drwxrwx---+ 4 root bin 1024 Jul 10 14:08 A
```

Note the + at the end of the directory permissions. This indicates there is a Solaris OE access control list assigned to this directory. You can use the Solaris OE `getfacl` command to display the access control list for a directory or file (see the `getfacl` man page for more information).

```
f15k-sc0:jim:18> getfacl A
# file: A
# owner: root
# group: bin
user::rwx
user:jim:rwx #effective:rwx
group::rwx #effective:rwx
mask:rwx
other:---
```

You might also note that `jim` has an additional privilege assigned. Without an access control list, only the user `root` and the group `bin` would have permissions to this directory. In the preceding sample screen output, you can see that user level permissions have been extended for user `jim`, effectively making them the same as permissions for the directory owner, `root`. The following example clearly demonstrates the effect of this.

```
f15k-sc0:jim:36> cd /
f15k-sc0:jim:37> pwd
/
f15k-sc0:jim:38> ls -ld testfacl
drwxrwx--- 2 root bin 512 Aug 7 19:21 testfacl
f15k-sc0:jim:39> cd testfacl
testfacl: Permission denied
```

The previous example demonstrates another directory with the same permissions as before, but there is no additional access control list. Note that permission to access that directory is denied to user `jim`.

The following command shows the empty access control list.

```
f15k-sc0:jim:40> getfacl testfacl
# file: testfacl
# owner: root
# group: bin
user::rwx
group::rwx #effective:rwx
mask:rwx
other:---
```

To add an access control list, as user `root`, you would type the following command.

```
# setfacl -m user:jim:rwX testfacl
```

See the `setfacl` man page for more information about adding access control lists. Essentially, adding an access control list using the preceding command gives `jim` the same permissions that user `root` already has. Here's the result:

```
f15k-sc0:jim:51> ls -ld testfacl
drwxrwx---+ 2 root bin 512 Aug 7 19:34 testfacl

f15k-sc0:jim:52> getfacl testfacl
# file: testfacl
# owner: root
# group: bin
user::rwX
user:jim:rwX #effective:rwX
group::rwX #effective:rwX
mask:rwX
other:---

f15k-sc0:jim:53> cd testfacl
f15k-sc0:jim:54> ls
f15k-sc0:jim:55> touch t.t
f15k-sc0:jim:56> ls -l
total 0
-rw-r----- 1 jim staff 0 Aug 7 19:34 t.t
```

Even though the normal permissions of the directory would not allow `jim` to list directory contents or create a file, the access control list allows this capability. Does this mean you need to configure the access control lists for SMS software yourself? No. SMS software includes a utility called `smsconfig` that adds users to the appropriate groups and maintains appropriate access control list permissions.

Earlier, I mentioned that user `jim` was a member of several groups, but then segued into a discussion on access control lists. Let's go back to the group mechanism for just a bit. The Solaris OE groups are mainly used by SMS commands before the functionality of that command is executed. Essentially, a library call is made to determine if the user executing the command is in the appropriate Solaris OE group that is eligible to execute the particular command. If it is not in the appropriate group, an appropriate message is printed. If it is in the appropriate group, the command is executed.

Now you can see how SMS software uses the Solaris OE access control list and group mechanisms to implement this facet of SMS security. If you are trying to determine why a given user cannot issue a particular command or edit a file in a particular directory as expected, you are now armed with the tools to discover the

reason. Now, with a better understanding of this SMS software feature and the Solaris OE, you can use this feature to manage your Sun Fire 15K server platform in a more secure fashion.

If you are hosting multiple projects run by multiple departments with multiple administrators in multiple domains, utilizing these features provides several benefits. The separation of responsibility becomes obvious and helps with accountability. You no longer have a long list of possible suspects who may have reconfigured the critical domain. Instead, you have a limited few who possess the appropriate permissions.

Using `ssh` for SC to SC File Propagation

SMS software supports the concept of a main and spare system controller. This allows one system controller to take over for the other under certain conditions. While I won't go into the specifics of the conditions here, suffice it to say in order for this to work properly, there must be a way within SMS for configuration information created on the main system controller to be replicated on the spare. This is referred to as file propagation. Prior to patch 112481-05 for SMS 1.2 and the release of SMS 1.3, file propagation was performed using what is commonly referred to as `r*` commands such as `rcp`. Many sites dislike the use of `r*` commands even in the limited context of use within the MAN between the system controllers due to concerns over security. To address this, the listed patch for SMS 1.2, and SMS 1.3, will allow for the use of `ssh/scp` (secure shell/secure copy) to perform the functions of file propagation in a more secure fashion. This should help to mitigate the concerns brought about by the prior use of the `r*` commands. OpenSSH is an open source internet available implementation of `ssh` which can be used on Solaris 8 OE, but if you are using Solaris 9 OE a version of secure shell is included and its use is recommended. Third party commercial implementations are available as well. Refer to the Sun BluePrints OnLine article "Securing the Sun Fire 12K and 15K System Controllers" for more detailed information.

About the Author

Tom Chalfant received his bachelor's degree in Computer Science from the University of Kentucky, and received a master's degree in Computer Science from the University of Dayton. Before joining Sun Microsystems in 1996, Tom was a US Air Force Officer and later worked for the US Government as a civilian. Since joining Sun Microsystems, Tom has worked as a System Support Engineer, Regional System Support Engineer, Corporate Problem Resolution Engineer, and now as a Staff

Engineer in support of the Sun Enterprise 10000 server and Sun Fire 15K server product groups. His experience spans many years, many disciplines, and many data centers, in both mainframe and UNIX-based shops across the US and abroad.

Acknowledgements

I would like to thank my peer, Scott Davenport, for his work on MAN and his relentless quest to enlighten his coworkers. His efforts are greatly utilized in the MAN section of this paper.

References

System Management Services (SMS) 1.2 Administrator Guide
Part # 816-3267-10, available at <http://docs.sun.com>

Alex Noordergraaf and Dina K. Nimeh, "Securing the Sun Fire 12K and 15K System Controllers," available at <http://www.sun.com/blueprints>

Jason Reid and Keith Watson, "Building and Deploying OpenSSH on the Solaris Operating Environment," available at <http://www.sun.com/blueprints>

Ordering Sun Documents

The SunDocsSM program provides more than 250 manuals from Sun Microsystems, Inc. If you live in the United States, Canada, Europe, or Japan, you can purchase documentation sets or individual manuals through this program.

Accessing Sun Documentation Online

The docs.sun.com web site enables you to access Sun technical documentation online. You can browse the docs.sun.com archive or search for a specific book title or subject. The URL is <http://docs.sun.com/>

To reference Sun BluePrints OnLine articles, visit the Sun BluePrints OnLine Web site at: <http://www.sun.com/blueprints/online.html>