Solaris™ Resource Manager 1.x to Solaris 9 Resource Manager

Migration Guide
February 2003

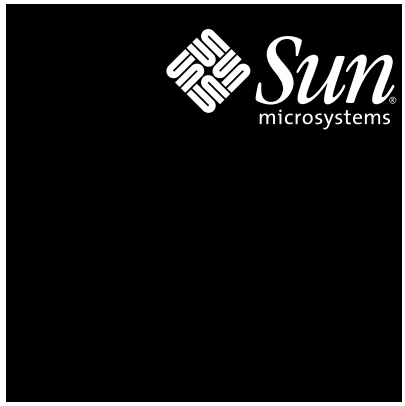


Table of Contents

Chapter 1

Migrating from Solaris™ Resource Manager 1.x to Solaris 9 Resource Manager Software

Abstract

Organizations running Solaris™ Resource Manager 1.x software on the Solaris 8, 7, or 2.6 Operating Systems and wishing to move to the Solaris 9 Operating System must migrate existing settings to Solaris 9 Resource Manager software.

This guide provides an overview of the architectural differences between Solaris Resource Manager 1.x and Solaris 9 Resource Manager, including new features as well as those no longer supported. The steps required to move from Solaris Resource Manager 1.x to Solaris 9 Resource Manager software are also discussed. A migration example is provided in which a system running Solaris Resource Manager 1.2 managing multiple workloads, including two Oracle databases, a batch job, and a Web server, is upgraded to the Solaris 9 Resource Manager.

Chapter 2

Introduction

At a time when budgets are tight and the demand for productivity and customer service excellence remains high, systems sprawl can generate excessive ownership costs with unsatisfactory returns on investment. This condition — which affects companies of all sizes and industries all over the world — *is* avoidable. Consolidating multiple applications and services onto a single system is the solution. As the first step toward designing a more rational, efficient, and flexible IT environment, consolidation delivers consistently high levels of service throughout the organization and helps bring total cost of ownership (TCO) under control.

Resource Management

Successful consolidation efforts require sophisticated resource management tools. Resource management software not only provides IT organizations with the fine-grained control they need over network-wide system resources, it also eases consolidation efforts by helping organizations control and allocate resources for maximum utilization, as well as meter and monitor resource usage for capacity planning, chargeback, or billing. By providing the ability to change service-level agreements or their policies dynamically, resource management software helps organizations ensure service-level agreements are met — even when demand rises.

Integration With the Solaris Operating System

While many system management tools provide the means to monitor the use of system resources, they typically lack the means to control resources. Initially, resource management existed in Solaris Resource Manager 1.x software. Because this was an unbundled product, it was not tightly integrated with the Solaris Operating System (OS), resulting in several disjoint schemas, including separate configuration and accounting files. Furthermore, the Solaris OS has historically had its own simple resource controls, such as *nice*, *rlimits*, and *processor sets*, which were unaware of the influence of Solaris Resource Manager 1.x on the system.

To address these issues, Sun integrated resource management functionality into the Solaris OS, enabling organizations to exert optimum control over network-wide system resources. The first release of the newly integrated software is Solaris 9 Resource Manager. A complete overview of the software can be found in the *Solaris 9 Resource Manager — Technical White Paper* listed in Chapter 7, *For More Information*.

By fully integrating resource management functionality into the Solaris 9 OS, organizations can use a single framework to configure, monitor, and control the system from the resource management point of view. Furthermore, integration enables the Solaris 9 Resource Manager to become a core foundation technology that Sun and its software partners can build on to create future innovations in resource sharing, instrumentation, and optimization. As a result, other key software components can take advantage of these powerful capabilities — at no additional cost. However, to fully integrate resource management features into the Solaris OS, the functionality, names, and commands have partially changed, though the principles have stayed the same. This means that if organizations running Solaris Resource Manager 1.x on the Solaris 8, 7, or 2.6 Operating Systems wish to move to the Solaris 9 Operating System, they must migrate existing settings to the Solaris 9 Resource Manager.

Chapter 3

Migrating to Solaris 9 Resource Manager

Prior to migrating existing configurations, it is important to understand the differences between the Solaris Resource Manager 1.x and Solaris 9 Resource Manager releases, as well as the steps involved and the tools provided to ease the migration effort.

Solaris Resource Manager 1.x

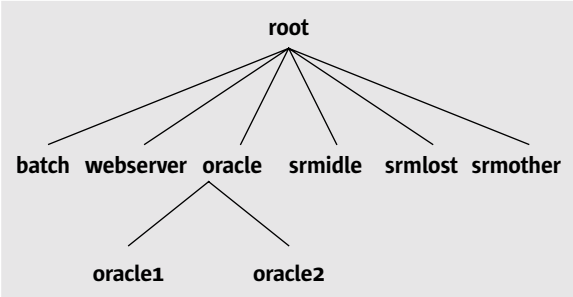
Features relevant to migration include[1]:

- **Operating system** — The Solaris 8, 7, and 2.6 Operating Systems are supported.
- **Lnodes (limit node) —** Lnodes are labels used to classify workloads. By default, they are linked to a userID. Resource limits are assigned to each lnode.
- **/var/srm/srmDB** — This file, which is the location of the configuration, can only be altered through Solaris Resource Manager 1.x commands and is unique to every system.
- **Tree structure** — Lnodes are stored in a tree structure that reflects their relative importance and resource assignment (Figure 3-1).
- **Default lnodes** — Solaris Resource Manager 1.x includes a default lnode tree that contains default lnodes, including `srmother`, `srmidle`, `srmlost`, and others.

1. Features relevant to migration can be found in the Solaris Resource Manager 1.2 Collection listed in Chapter 7, *For More Information*.

- **SHR**[2] — SHR is the classic fair share scheduler. SHR is unaware of processor sets and their configuration.
- **Other commands** — Other helpful commands are located under */usr/srm/...,* including `limreport`, `limadm`, `liminfo -v`, `schedtree`, and `limid -p`.

Figure 3-1: An example lnode tree



Solaris 9 Resource Manager

Features relevant to migration include[3]:

- **Operating system** — The software integrates resource management capabilities with the Solaris 9 OS. The result is a framework that can be expanded on in later releases.
- **Projects** — Project labels are used to classify workloads. Projects can be linked to userID, groupID, or any name. Resource limits can be assigned to each project.
- **/etc/project** — This file, which is the location of the configuration, can be altered with any text editor and placed into a directory, such as LDAP or NIS.
- **Flat structure** — Projects are not ranked in a tree structure, reducing complexity. This is particularly important when managing many systems with projects placed in a directory (Figure 3-2).
- **Default projects** — Every system has its own default */etc/project* file that includes the projects `system`, `user.root`, `noproject`, `default`, and `group.staff`.
- **FSS**[4] — FSS is the revised fair share scheduler. The FSS is aware of the processor set.
- **Other commands** — Other helpful commands are located in */usr/bin* and */usr/sbin*, including `projadd`, `projmod`, `projdel`, `prtctl`, `prstat -J`, `pgrep -J`, and `pkill -J`.

Figure 3-2: A flat project structure

```
system
user.root
noproject
default
group.staff
batch
webserver
oracle
oracle1
oracle2
```

2. SHR is the scheduler name, like TS (time share) and SYS (system). A easy way to see the scheduler/process relationship is to use the `ps -cafe` command.
3. Features relevant to migration can be found in the Solaris 9 Resource Manager — Technical White Paper listed in Chapter 7, *For More Information*.
4. FSS is the scheduler name, like TS (time share) and SYS (system). A easy way to see the scheduler/process relationship is to use the `ps -cafe` command.

Summary of Changes

Figure 3-3 summarizes the differences between Solaris Resource Manager 1.x and Solaris 9 Resource Manager software.

Figure 3-3: Summary of changes between Solaris Resource Manager 1.x and Solaris 9 Resource Manager software

	Solaris Resource Manager 1.x	**Solaris 9 Resource Manager**
Operating System	Solaris 8, 7, and 2.6	Solaris 9
Workload Classification	lnodes	projects
Configuration Location	/var/srm/srmDB	/etc/projects
Structure	Tree structure	Flat structure
Default Workload Classes	srmother, srmidle, srmlost	system, user.root, noproject, default, group.staff
Fair Share Scheduler	SHR	FSS
Other Commands	/usr/srm/... limreport, limadm, liminfo -v, schedtree, limid -p	/usr/bin and /usr/sbin projadd, projmod, projdel, prtctl, prstat -J, pgrep -J, pkill -J

Migration

Migrating from the Solaris 8, 7, or 2.6 OS with Solaris Resource Manager 1.x to the Solaris 9 OS with Solaris 9 Resource Manager requires the following steps:

- Convert lnodes into projects — Note that unlike lnodes, projects can be bound to a userID, although this is not required.
- Migrate information — Extracting information from the */var/srm/srmDB* file and moving it into the */etc/project* file.
- Change the tree structure to a flat structure — All relative values of the shares down in the tree structure must be translated into absolute share values for the flat structure.
- Change tools — Using different tools to monitor and maintain the resource management on the system.

Two scripts[5] are provided to assist in the migration to the Solaris 9 Resource Manager software.

- The *srm_backup* script backs up the existing database employed by Solaris Resource Manager 1.x software. Run prior to upgrading the host to the Solaris 9 OS, the *srm_backup* script creates a text file that describes the existing Solaris Resource Manager 1.x lnode database.
- The *srm_convert* script uses the description created by the *srm_backup* script to convert the backed-up lnode structure to a file with the same structure as the */etc/project* file. The *srm_convert* script uses the output of the *srm_backup* script to determine what share of CPU resources each user currently receives, and produces a project file entry for each user. The project entries created also include those lnodes with zero shares. The output of the *srm_convert* script may then be appended to the */etc/project* file after conflicting projects are resolved.

5. The scripts can be found in the Solaris Resource Manager 1.x to Solaris 9 Resource Manager Migration Tool listed in Chapter 7, *For More Information*.

During the conversion process, the *srm_convert* script extracts the information that makes a given configuration unique and converts it to a project structure. It is important to note that some lnodes created through the Solaris 8, 7, or 2.6 OS or Solaris Resource Manager 1.x installation are not needed in the new project structure and are ignored by the *srm_convert* script. These lnodes include: `adm`, `bin`, `daemon`, `listen`, `lp`, `noaccess`, `nobody`, `nuucp`, `root`, `srmidle`, `srmlost`, `srmother`, `sys`, and `uucp`. These lnodes have been replaced by the `system`, `user.root`, `noproject`, `default`, and `group.staff` projects, which exist in every system by default. If necessary, administrators can create projects for the missing lnodes. Also, certain limits found in Solaris Resource Manager 1.x, such as virtual memory, connection time, and number of logins, are not available in the Solaris 9 Resource Manager software. As a result, these limits are not converted by the migration tool.

Chapter 4

Converting Lnodes to Projects

Administrators can take the following steps to convert existing Solaris Resource Manager 1.x configurations:

1. Back up the existing lnode configuration using the *srm_backup* script. For example:

```
# srm_backup /var/srm/srmDB.backup
```

Figure 4-1 illustrates an *srmDB.backup* file containing the information that represents the lnode structure.

Figure 4-1: An example srmDB.backup file

/var/srm/srmDB.backup

```
0         4294967295    root            1       0
1         43            daemon          1       1
60001     43            nobody          1       1
1001      0             batch           10      10
2         43            bin             1       1
60002     43            noaccess        1       1
3         43            sys             1       1
4         43            adm             1       1
5         43            uucp            1       1
9         43            nuucp           1       1
37        43            listen          1       1
41        0             srmidle         0       1
42        0             srmlost         1       1
43        0             srmother        1       1
71        43            lp              1       1
1200      0             webserver       20      1
2002      43            dan             10      1
800       0             oracle          30      50
801       800           oracle1         50      1
802       800           oracle2         50      1
2003      43            bob             5       1
```

2. Run the *srm_convert* script to convert the lnode structure to a project structure. The result of the conversion is a list of projects that is equivalent to the original lnode structure. For example:

```
# srm_convert /var/srm/srmDB.backup > /var/tmp/newprojects
```

Figure 4-2 describes the transformed *newprojects* and */etc/project* files. This example illustrates how to use the ability of the Solaris 9 Resource Manager to link a project to a user, enabling the same result as that obtained when using Solaris Resource Manager 1.x software. The *newprojects* file contains all the information that made the lnode tree in Figure 4-1 unique.

Figure 4-2: The newprojects and /etc/project files

/var/tmp/newprojects

```
user.batch:1001::::project.cpu-shares=(privileged,50,none)
user.webserver:1200::::project.cpu-shares=(privileged,100,none)
user.dan:2002::::project.cpu-shares=(privileged,2,none)
user.oracle:800::::project.cpu-shares=(privilged,50,none)
user.oracle1:801::::project.cpu-shares=(privileged,50,none)
user.oracle2:802::::project.cpu-shares=(privileged,50,none)
user.bob:20023::::project.cpu-shares=(privileged,1,none)
```

/etc/project

```
system:0::::
user.root:1::::
noproject:2::::
default:3::::
group.staff:10::::
user.batch:1001::::project.cpu-shares=(privileged,50,none)
user.webserver:1200::::project.cpu-shares=(privileged,100,none)
user.dan:2002::::project.cpu-shares=(privileged,2,none)
user.oracle:800::::project.cpu-shares=(privileged,50,none)
user.oracle1:801::::project.cpu-shares=(privileged,50,none)
user.oracle2:802::::project.cpu-shares=(privileged,50,none)
user.bob:2003::::project.cpu-shares=(privileged,1,none)
```


3. If the Solaris 9 OS was installed as an initial install, skip this step and go directly to step 4. If this is an upgrade install and Solaris Resource Manager 1.x software is running, uninstall the existing software packages. For example:

   ```
   # pkgrm SUNWsrmb SUNWsrmm SUNWsrmr
   ```

4. Install the Solaris 9 OS through an upgrade or an initial install. Administrators should ensure the newly installed host has access to the *srm_convert* script and the *srmDB.backup* file, and that the `passwd` and `group` databases are substantively equivalent.
5. Merge the contents of the */tmp/newprojects* file into the */etc/project* file.
6. Instruct the system to use the FSS software. For example:

   ```
   # dispadmin -d FSS
   ```

   This change takes effect upon the next system reboot[6].
7. Test the system to verify that the conversion from an lnode structure to a project structure is correct. Administrators can compare the output of the `schedtree` utility with the new */etc/project* entries, checking for similarities in the ratio of shares of lnodes and projects to the total number of shares.

Figure 4-3 illustrates how the `schedtree` command can be used to display the current lnode structure.

Figure 4-3: Example `schedtree` output

```
# schedtree
 root(0) [1]
  batch(1001) [10]
  srmidle(41) [0]
  srmlost(42) [1]
  webserver(1200) [20]
  srmother(43) [1]
 daemon(1) [1]
 nobody(60001) [1]
 bin(2) [1]
 noaccess(60002) [1]
 sys(3) [1]
 adm(4) [1]
 uucp(5) [1]
 nuucp(9) [1]
 listen(37) [1]
 lp(71) [1]
 dan(2002) [10]
 bob(2003) [5]
  oracle(800) [30]
 oracle1(801) [50]
 oracle2(802) [50]
```

6. Enabling the FSS software can also be done without a reboot. See the FSS(7) man-page for more information.

Chapter 5

New Opportunities With Solaris 9 Resource Manager

Once a configuration is converted to the new project structure, organizations can begin to take advantage of new features in the Solaris 9 Resource Manager. The following sections describe the features often found to be initially interesting.

Resource Pools

With Solaris 9 Resource Manager software, administrators can control system resources through a resource pool and reserve them for exclusive use by an application or set of applications, represented by projects. Using resource pools, system administrators can partition a system into a set of smaller virtual environments, each providing resources for one or more projects. These partitions provide fixed boundaries between workloads, ensuring each has access to a consistent set of resources regardless of resource usage or contention on the rest of the system. As a result, administrators can separate workloads so that resource consumption does not overlap, helping achieve predictable application and system performance. See the `poolcfg(1M)`, `pooladm(1M)`, and `poolbind(1M)`man pages for more information.

Statistical Tools

Solaris 9 Resource Manager is integrated with the extensive statistical tools built into the Solaris 9 Operating System. With these tools, administrators can record resource consumption on a project, task, or process basis. Solaris 9 Resource Manager software is integrated with many Solaris statistical tools, including the `kstat`, `prstat`, and `ps` commands. Other commands like `pgrep` and `pkill` can also execute on projects as a whole.

Accounting Tools

The extended accounting feature in Solaris 9 software provides an extensible facility for recording information about system activity on a per-process or per-task basis. Using the accounting tools, administrators can collect process and task data to capture a detailed set of resource consumption statistics and label the usage records with the project for which the work was done. If resource chargeback, workload monitoring, and capacity planning are a concern, the extended accounting and reporting information can be integrated with third-party billing software. See the `acctadm(1M)` man page for more information.

Solaris Admin GUI

The Solaris 9 OS includes a graphical user interface (GUI) for administration that contains windows to set up and monitor projects. See the `smc(1M)` man page for more information.

Chapter 6

Using the Solaris 8 Operating System With Projects

Organizations may wish to continue using the Solaris 8 OS and Resource Manager 1.2 software on certain systems. Some functionality, such as projects and the statistical and accounting tools, are already available in the Solaris 8 OS[7].

To take advantage of this functionality, system administrators can use the migration tool to convert — or mirror — an existing Solaris Resource Manager 1.2 lnode configuration to the new project definition. As a result, system administrators can use Solaris Resource Manager 1.2 software to control resources and use the project feature for statistical and accounting information. To do so, administrators can:

1. Back up the existing Solaris Resource Manager 1.2 lnode configuration using the *srm_backup* script. For example:

   ```
   # srm_backup /var/srm/srmDB.backup
   ```

2. Run the *srm_convert* script to convert the lnode structure to a project structure and merge it with the */etc/project* file. For example:

   ```
   # srm_convert /var/srm/srmDB.backup > /tmp/newprojects
   ```

   This creates two different definitions that represent the workloads on the system; one representation in lnodes and the other in projects.

7. The minimum requirement is Solaris 8 06/00 (Update 1) software.

3. Use the built-in Solaris statistical tools to verify that the same results are viewable in both places. To do so, administrators can use the `srmstat` and `liminfo` commands on lnodes and the `prstat` and `ps` commands on projects.

If similar updates are always made to both the lnode database and the project database, organizations can use Solaris 9 Resource Manager accounting in concert with the Solaris Resource Manager 1.2 scheduler. This combination is most effective in static environments.

Chapter 7

For More Information

Organizations should contact a local Sun support representative to learn how Sun can help ease the transition from Solaris Resource Manager 1.x to the Solaris 9 Resource Manager software. Figure 7-1 identifies other sources of information related to the Solaris 9 Resource Manager software and other Sun products and technologies.

Figure 7-1: Web links for more information on the Solaris 9 Resource Manager software and other Sun products and technologies

	Description or Title and Web Site URL
[1]	Solaris 9 Resource Manager — Technical White Paper www.sun.com/software/solaris/whitepapers.html
[2]	Solaris Resource Manager 1.2 Collection docs.sun.com/db?p_coll/409.4
[3]	Solaris Containers — How Advancements in Server Virtualization Will Change the Future of Enterprise Computing www.sun.com/software/solaris/whitepapers.html
[4]	Solaris Resource Manager 1.x to Solaris 9 Resource Manager Migration Tool www.sun.com/software/download/sys_admin.html

Please
Recycle

Adobe PostScript

Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 USA **Phone** 800 786-7638 or +1 512 434-1577 **Web** sun.com

We make the net work.

Sun Worldwide Sales Offices: Africa (North, West and Central) +33-13-067-4680, Argentina +5411-4317-5600, Australia +61-2-9844-5000, Austria +43-1-60563-0, Belgium +32-2-704-8000, Brazil +55-11-5187-2100, Canada +905-477-6745, Chile +56-2-3724500, Colombia +571-629-2323, Commonwealth of Independent States +7-502-935-8411, Czech Republic +420-2-3300-9311, Denmark +45 4556 5000, Egypt +202-570-9442, Estonia +372-6-308-900, Finland +358-9-525-561, France +33-134-03-00-00, Germany +49-89-46008-0, Greece +30-1-618-8111, Hungary +36-1-489-8900, Iceland +354-563-3010, India–Bangalore +91-80-2298989/2295454; New Delhi +91-11-6106000; Mumbai +91-22-697-8111, Ireland +353-1-8055-666, Israel +972-9-9710500, Italy +39-02-641511, Japan +81-3-5717-5000, Kazakhstan +7-3272-466774, Korea +822-2193-5114, Latvia +371-750-3700, Lithuania +370-729-8468, Luxembourg +352-49 11 33 1, Malaysia +603-21161888, Mexico +52-5-258-6100, The Netherlands +00-31-33-45-15-000, New Zealand–Auckland +64-9-976-6800; Wellington +64-4-462-0780, Norway +47 23 36 96 00, People's Republic of China–Beijing +86-10-6803-5588; Chengdu +86-28-619-9333; Guangzhou +86-20-8755-5900; Shanghai +86-21-6466-1228; Hong Kong +852-2202-6688, Poland +48-22-8747800, Portugal +351-21-4134000, Russia +7-502-935-8411, Singapore +65-6438-1888, Slovak Republic +421-2-4342-94-85, South Africa +27 11 256-6300, Spain +34-91-596-9900, Sweden +46-8-631-10-00, Switzerland–German 41-1-908-90-00; French 41-22-999-0444, Taiwan +886-2-8732-9933, Thailand +662-344-6888, Turkey +90-212-335-22-00, United Arab Emirates +9714-3366333, United Kingdom +44 0 1252 420000, United States +1-800-555-9SUN or +1-650-960-1300, Venezuela +58-2-905-3800 FE1953-0