Solaris™ 9 Resource Manager Quick Start Guide for System Administrators

Manage system resources more effectively.

Sun microsystems

Key feature highlights

- Improves resource utilization through fine-grained resource controls
- Helps ensure predictable service levels through resource partitions
- Increases accuracy of billing and capacity planning through detailed resource usage data
- Simplifies administrative tasks for resource management with a graphical user interface (GUI) tool

IT managers and system administrators are increasingly required to guarantee high service levels while reducing costs. Solaris 9 Resource Manager can help you address these important, yet often conflicting, goals.

The traditional way to guarantee service levels is to deploy a system or domain for each application. This model leads to a proliferation of expensive systems — expensive both in terms of physical maintenance and administration.

One way to avoid the proliferation model is to consolidate multiple applications onto a single system. In order to employ this solution, IT managers and system administrators must be confident that resource utilization will be greatly increased without a negative performance impact on their critical applications.

Sun's Solaris™ 9 Resource Manager technology, now integrated with the Solaris 9 Operating System, helps administrators manage system resources more effectively. With Solaris 9 Resource Manager, system administrators can define workloads, partition system resources, and monitor resource usage. By monitoring resource usage, administrators can identify users or applications that tend to use more resources than they should and compile more accurate data over time for capacity planning and billing. Once resource usage is linked to specific applications and users, various Solaris 9 Resource Manager mechanisms can be employed to control the amount of system resources an individual application can consume.

Solaris 9 Resource Manager provides the flexible solution for system administrators to perform server consolidation more effectively.

The Differences Between Solaris 9 Resource Manager and Solaris Resource Manager 1.x

Solaris Resource Manager 1.x is an unbundled product used for controlling system resources. It supports the Solaris 2.6, 7, and 8 Operating Systems. In the Solaris 9 Operating System, resource management is handled with Solaris 9 Resource Manager.

While Solaris 9 Resource Manager contains many features similar to Solaris Resource Manager 1.x, the code base and architectural design of Solaris 9 Resource Manager is completely different. Solaris 9 Resource Manager also contains a number of features that are not part of Solaris Resource Manager 1.x.

Solaris 9 Resource Manager provides the tools to assist system administrators in delivering predictable service levels, reducing service level costs, and increasing ease of management — today.

If you are already familiar with Solaris Resource Manager 1.x, you can expect a smooth transition to Solaris 9 Resource Manager. Sun has published a technical white paper (located at sun.com/software/whitepapers/solaris9/srm.pdf) that describes the steps to transition from Solaris Resource Manager 1.x to Solaris 9 Resource Manager. A migration script is also available for free download at sun.com/software/download/sys_admin.html.

Under the Hood: Core Technology

So how does Solaris 9 Resource Manager work? The building blocks of Solaris 9 Resource Manager are tasks, projects, and resource controls.

Tasks

A task is simply a collection of related processes; it is similar to a session or a process group. Tasks are assigned task IDs, which are similar to process IDs. For example, when a user logs into the system, the user's shell is placed into a newly created task. If the user logs in a second time, a second task is created. You can also invoke *newtask*(1) to create additional tasks.

Projects

A project is a tag that is used to classify a service, such as a database instance. A project may consist of one or more tasks that represent a workload. Projects and tasks also enable administrators to more effectively track resource usage through extended accounting.

Projects are defined in the *project*(4) database (which may be in */etc/project,* or in your NIS or LDAP name service). For example, you could define a project called *students* and associate that project with the UNIX® groups called *firstyear*, *sophomore*, *junior*, and *senior*. When any of the members of these groups log into the system, their processes will be associated with the *students* project.

The projects database is maintained using the following administrative commands:

- *projadd*(1M) to add entries to the database
- *projdel*(1M) to delete entries in the database
- *projmod*(1M) to modify existing database entries

Resource Controls

In Solaris 9 Resource Manager, you can establish resource limits on a per-process, per-task, and per-project basis.

Resource controls dictate how the Solaris kernel will manage the controlled entity, as well as how the system will react when the resource limit has been reached. For example, you could limit the number of lightweight processes (LWPs) in each task to 1000 for all tasks in the *student* project. This would help prevent students from launching "fork-bombs" because each process uses at least one LWP. Alternatively, you could configure the resource control so that a syslog message is issued when this limit is exceeded.

In addition to configuring resource controls in the projects database, you can make changes to workloads on a running system by using the *rctladm*(1M) administrative command. For resource control changes to running processes, you can use the *prctl*(1M) command.

Fair Share Scheduler

Solaris 9 Resource Manager incorporates the fair share scheduler, *FSS*(7), to provide you with the ability to specify that certain processes be given more CPU resources than the others. Administrators controls the allocation of available CPU resources among projects based on their importance. The relative importance of an application is expressed in terms of *shares*. The more shares assigned to a project, the more CPU resource the project receives.

The goal of FSS differs from the traditional time-sharing scheduling class (TS). In addition to scheduling individual LWPs, FSS balances work fairly between projects. This makes it impossible for any project to acquire more CPU cycles simply by running more processes concurrently.

You can set FSS as the default scheduling class for the entire system using the *dispadmin*(1M) command.

Resource Partitioning With Pools

In order to provide predictable service levels to applications on a shared system, system administrators need a way to help ensure that applications will always have access to a consistent set of resources — regardless of the resource usage on the rest of the systems.

Using the resource pool facility in Solaris 9 Resource Manager, you can partition system resources (such as CPUs) into named groups. Processes can be bound to these pools directly using *poolbind*(1), or use an assignment to the *project.pool* attribute in the *project*(4) database.

You can also define minimum and maximum resource values for each resource type in a pool to maintain resource allocations across dynamic reconfiguration operations and other events that change the available resources on the system.

While the resource pool facility provides the ability to partition system resources, it does not allow you to specify how resources are to be shared by applications within the resource pool. The FSS can be used within the resource pool to ensure higher priority applications receive more resources. Besides the FSS, the Solaris Operating System provides other schedulers, such as time share scheduler (TS) and the interactive scheduler (IA). With Solaris 9 Resource Manager, you can associate different schedulers with different resource pools in order to take advantage of the schedulers that best serve the needs of the applications running on them.

The various resource pool characteristics described are set in pool configuration files by the *poolcfg*(1M) command. The *pooladm*(1M) command is used to activate and deactivate pool configurations.

GUI Management Tools

Solaris 9 Resource Manager features can be managed from panels in Solaris Management Console 2.1. With this GUI, you can configure the project database as well as monitor system activity. You can obtain information for total number of processes, percent CPU usage, percent memory usage, etc. on a per-project or per-user basis.

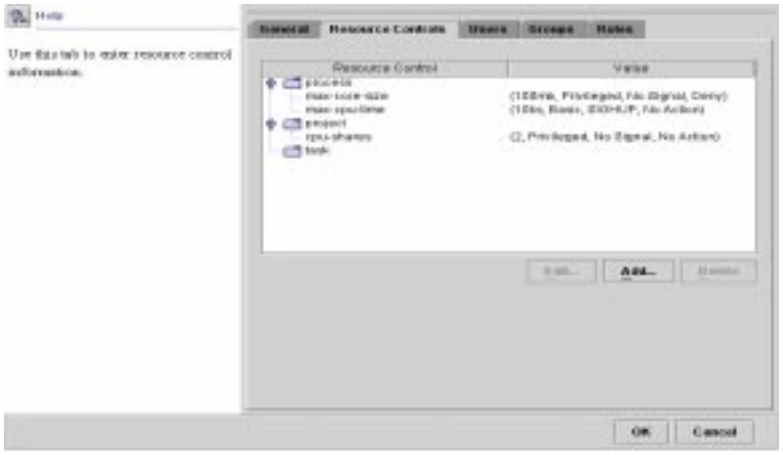

Figure 1: A view of the project manager tool user interface

Solaris™ 9 Resource Manager Quick Start Guide for System Administrators

IP Quality of Service (IPQoS)

IPQoS — an implementation of the Differentiated Services model defined in *RFC2475: An Architecture for Differentiated Services* — has been available in the Solaris 9 Operating System since the 9/02 release. You can use IPQoS to define and manage network flows using a sophisticated control language.

With the *ipgpc*(7IPP) packet classifier, you can classify packets based on IP header, user ID, project ID, and other parameters. Using *ipqosconf*(1M), you can subsequently meter flows, mark the differentiated services codepoint (DSCP) field in packets, or drop packets.

IPQoS includes a powerful flow accounting facility, which is fully integrated with the extended accounting framework. IPQoS can mark packets on a VLAN with a user priority as defined by the IEEE 801.D standard. An overview of IPQoS can be found in the *ipqos*(7IPP) manpage.

Extended Accounting

With Solaris 9 Resource Manager, you can obtain detailed resource usage information to track resource consumption. Solaris 9 Resource Manager takes advantage of the extended accounting (or *exacct*) feature, a redesign of the standard Solaris accounting facility. With extended accounting, you can aggregate accounting data using tasks and projects. This provides a comprehensive solution for workload accounting. Additionally, extended accounting is easy to administer using the *acctadm*(1M) command.

Exacct also includes a convenient programming API and an extensible file format; *libexacct*(3LIB) enables you to generate and read *exacct* files. Demonstration source code is included in */usr/demo/libexacct* as part of the *SUNWosdem* package. In the Solaris 9 Operating System 4/03 release, a Perl interface to *exacct* will be provided that will enable you to construct your own custom reports.

Upcoming Developments

Sun continues to invest in its resource management capabilities for the Solaris Operating System. New features planned for Solaris 9 software include:

- Physical memory control
- Dynamic resource pools

Summary of Main Benefits

The Solaris 9 Resource Manager is designed to help system administrators manage system resources more efficiently by:

- Delivering predictable application service levels
- Improving resource utilization
- Helping reduce service level costs

Platforms and Requirements

Solaris 9 Resource Manager is integrated into Solaris 9 Operating System and is supported on SPARC® and x86 processor-based systems running the Solaris 9 Operating System.

Learn More

Get the inside story on the trends and technologies shaping the future of computing by signing up for the Sun Inner Circle program. You'll receive a monthly newsletter packed with information on the latest innovations, plus access to a wealth of resources. Register today to join the Sun Inner Circle Program at sun.com/joinic.

To receive additional information on Sun software, products, programs, and solutions, visit sun.com/software.

Additional Information

- For more information on Solaris 9 Resource Manager, visit sun.com/solaris.
- Specific configuration examples can be found at sun.com/blueprints/0902/816-7753-10.pdf.
- Documentation for all Solaris 9 Resource Manager features can be found at docs.sun.com.
- Solaris Resource Manager 1.x to Solaris 9 Resource Manager migration information can be found at sun.com/software/whitepapers/solaris9/srm.pdf.
- A migration script is available for free download at sun.com/software/download.

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 +44-1-276-20444, United States +1-800-555-9SUN or +1-650-960-1300, Venezuela +58-2-905-3800

v3.0 3/03 DE1847-0