

# Solaris Resource Manager™ 1.1 White Paper

---

*Consolidate Servers and Increase Resource  
Utilization*



Sun Microsystems, Inc.  
901 San Antonio Road  
Palo Alto, CA 94303  
1 (800) 786.7638  
1.512.434.1511

Copyright 1999 Sun Microsystems, Inc., 901 San Antonio Road, Palo Alto, California 94303 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 U.S. and other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the Sun logo, Solaris, Solaris Resource Manager, and Sun Enterprise are trademarks, registered trademarks, or service marks of Sun Microsystems, Inc. in the U.S. and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. 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.

**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).

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 1999 Sun Microsystems, Inc., 901 San Antonio Road, Palo Alto, Californie 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 logo Sun, Solaris, Solaris Resource Manager, et Sun Enterprise 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

# Contents

---

Introduction .....	1
Controlling System Resources .....	1
Complementary System and Resource Management Products .....	2
Product Availability .....	2
Solaris Resource Manager .....	3
Key Features and Benefits .....	3
Fair Resource Allocation .....	4
Hierarchical Resource Allocation Model .....	4
Policy-Based Resource Administration .....	5
Usage Reports .....	6
Key Environments for Resource Control .....	6
Server Consolidation .....	6
Web/Application Hosting .....	7
Sites with Large or Varied User Populations .....	7
Cluster Environments .....	7
Controlling Resources .....	8
Fine-Grained Control .....	8
Per-User Information .....	8
The Solaris Resource Manager Hierarchy .....	9
Configuration .....	9
How Solaris Resource Manager Controls Resources .....	9
Types of Resources .....	10

An Example .....	10
CPU Time .....	11
Virtual Memory .....	13
Number of Processes .....	14
Resource Parameters .....	14
Accounting .....	14
Managing Resources .....	15
Delegating Resource Control .....	15
Changing the Current Resource Policy .....	16
Tools .....	16
Summary .....	17
Additional Resources .....	18

## Introduction

---

Responding to fierce competition, businesses are requiring IT to control costs and guarantee service levels for their enterprise applications. IT managers need to reduce costs and simplify administration. They realize that today's system model of hosting one application per system for more predictable service availability is outmoded and expensive, and will only lead to a proliferation of systems that are expensive to administer.

IT managers are consolidating applications from smaller systems onto larger, more scalable servers. Servers are typically configured to run a single, mission-critical application at 70% of peak capacity. With average capacity utilization of 20%-30% at normal times, there are significant cost and administrative savings in increasing resource utilization to levels attained in the mainframe environment. However, if multiple mission-critical applications are running on a single server, they will compete for system resources with less important tasks. A single runaway process can potentially cause business-critical functions to starve for resources. It's not enough for key business applications to be up and running — they must also be assigned an appropriate share of system resources.

Essentially, IT managers want mainframe-class resource availability at a fraction of the cost. They need the ability to consolidate multiple applications on a single server while fully utilizing system resources. At the same time, they want every user and application to receive the appropriate level of resource.

## Controlling System Resources

Solaris Resource Manager™ software is a tool from Sun Microsystems for enabling resource availability for users, groups, and enterprise applications. While many system management tools provide the means to *monitor* the use of system resources, they typically lack the means to *control* system resources. Solaris Resource Manager offers the ability to control and allocate CPU time, processes, virtual memory, connect time, and logins. And it does so on a fine-grained basis.

# Complementary System and Resource Management Products

Solaris Resource Manager software complements a number of products that Sun offers for increasing resource availability:

*Dynamic Reconfiguration:* Solaris Resource Manager software complements this unique capability of Sun's mid-range and high-end servers. System resources, such as CPU, memory, or I/O boards, can be added to or removed from the server without any change to Solaris Resource Manager.

*Dynamic System Domains:* Solaris Resource Manager software augments Dynamic System Domains on the Sun Enterprise™ 10000 server. Domains offer the ability to have separate OS instances for each application or set of applications. The system administrator could increase or decrease domain capacity based on changes in processor load. With Solaris Resource Manager, resources can be even more finely tuned for each domain.

*Processor Sets:* Solaris Resource Manager is complementary to the “processor sets” feature introduced in the Solaris™ 2.6 Operating Environment. With processor sets, applications can be bound to specific processors, allowing the administrator to set aside specific CPUs for specific applications. Solaris Resource Manager can be configured to allocate resources on the remaining (“unspecified”) processors in a system.

*Solaris™ Bandwidth Manager 1.5:* By installing Solaris Bandwidth Manager 1.5 on the network's major links and known congestion points, bandwidth can be controlled throughout the network. Solaris Bandwidth Manager, previously known as Sun™ Bandwidth Allocator, allows service providers to manage network traffic flow, remotely monitor available resources in real-time, perform accounting based on traffic flow and network usage, and allocate network bandwidth to individual applications, users, and organizations.

*Sun™ Cluster 2.2:* Solaris Resource Manager enables IT managers to easily and reliably allocate and control major system resources on a highly available Sun cluster. Sun Cluster software provides high availability (HA) support for applications and parallel database access on a cluster of servers. Together, application access can be maintained with predictable performance — even in the unlikely event of a system failure.

## Product Availability

Solaris Resource Manager 1.1 software is available today on the Solaris 7 and 2.6 Operating Environment. It is supported on the full range of Sun's SPARC™ servers.

## Solaris Resource Manager

---

Solaris Resource Manager software enables increased resource availability for users, groups, and applications. It provides the ability to allocate and control major system resources such as CPU, virtual memory, and number of processes. After a resource policy is set, the system administrator can walk away, confident that mission-critical applications will get the resources they demand. Additionally, users and applications will receive a more consistent level of service.

### Key Features and Benefits

Solaris Resource Manager re-defines the traditional model of hosting an individual application on a single system. Now, server resources utilizing the Solaris operating system can be controlled using methods similar to mainframe-class systems. With this product, multiple applications and groups receive a consistent level of service on a single server. In fact, resources can be allocated to the individual user. Resource utilization can actually increase because unused capacity is dynamically allocated to active users and applications. This results in significant cost savings and greater administrative flexibility.

Solaris Resource Manager software provides a means for *proactively* controlling system resources. Other tools only allow a system administrator to see when valuable resources — processes, CPU, or memory — are about to be exhausted. At that time, the system administrator must then react to prevent mission-critical applications from failing. Solaris Resource Manager not only provides resource usage reports, it reduces the likelihood of resource exhaustion in the first place because it guarantees resources to key applications and users. The benefit of this is that it makes the performance of an application more predictable, and ensures that system response times are not adversely affected by other tasks on the system.

### *Fair Resource Allocation*

Solaris Resource Manager software allocates resources according to the administrator-defined policy, and prevents resources allocation to users and applications who are not entitled to them. Solaris Resource Manager services many users fairly, and can be used to control system resources for hundreds of users and applications.

Solaris Resource Manager prevents rogue processes from running wild and consuming all the available CPU power or virtual memory. When many applications are consolidated on a single large server, this protects key applications.

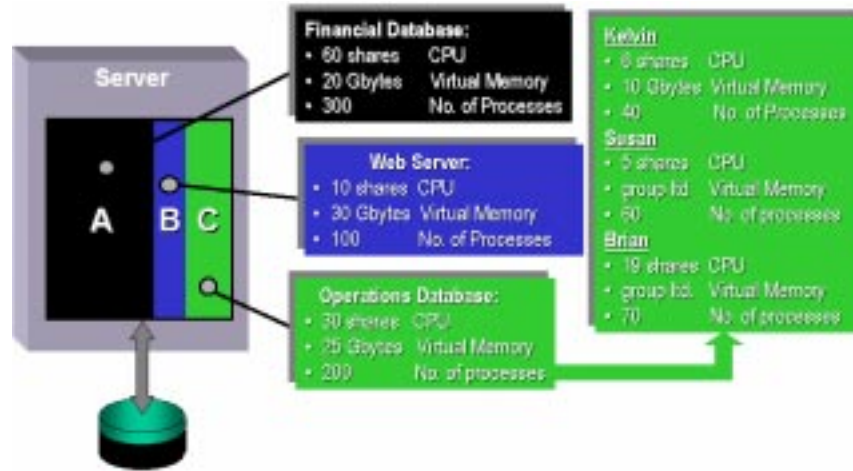
Available resources are never left unused. For example, CPU time is allocated based on an assigned number of *shares*, rather than on a flat percentage basis. When applications are idle or are consuming less than their designated CPU allotment, other applications or users benefit by taking advantage of additional processing power.

Solaris Resource Manager software also has the ability to vary CPU resources based on previous usage. It “remembers” how long ago a particular user logged on and what resources were used, and can allocate fewer resources to a more active user, ensuring that even low-priority users or applications are not starved of resources.

### *Hierarchical Resource Allocation Model*

One significant advantage of Solaris Resource Manager software is its hierarchical nature. The granularity of control can be as deep or flat as the system administrator desires. Shares can be allotted to a department, and then to individuals in the department.

Figure 2-1 provides an overview of how Solaris Resource Manager allocates resources in a hypothetical organization. In this example, the system administrator divides system resources between three groups of users: financial database (60 shares), Web server (10 shares), and operations database (30 shares). Resources allocated to the operations are then further distributed to Kelvin (6 shares), Susan (5 shares), and Brian (19 shares).



**FIGURE 1** Solaris Resource Manager allocates resources across groups and users

In Figure 1, the system administrator has allocated 30 shares for CPU resources for the operations. This represents a minimum of 30% (30 of 100 shares) of CPU resource if the financial database and Web server are also in operation. If the financial database was not running, Operations would receive 75% (30 of 40 shares).

Within operations, Kelvin receives 6 shares and Susan 5 shares. Without Brian on the system, Kelvin would receive 55% (6 of 11 shares) of the operations subset. If Brian were to log on, Kelvin's allocation would drop to 20% (6 of 30 shares) of the operations' CPU allocation. If operations' CPU allocation was increased, then Kelvin's allocation would increase accordingly, without the system administrator explicitly changing it. In this way, the Solaris Resource Manager hierarchical model greatly simplifies the system administrator's job of adding new resources and new users.

The Solaris Resource Manager hierarchy also allows the system administrator to delegate resource policy decisions without distributing superuser access. The system administrator can designate sub-policy administrators who can manage Solaris Resource Manager policy within their own workgroups.

### *Policy-Based Resource Administration*

Solaris Resource Manager software is particularly easy to configure and administer. The system administrator creates a configuration for virtual memory, number of processes, number of logins and connect time, and shares for CPU. Once the system administrator establishes this policy, Solaris Resource Manager enforces it — without requiring further operator intervention.

If the system administrator wants to update or fine-tune the policy, it's easy to make changes while the system is still running. The policy can also be changed to reflect different allocation policies at different times by running scripts with the changes via `cron`.

### *Usage Reports*

Solaris Resource Manager provides detailed usage data on individual and group usage. This can be input through user-defined tools or scripts into resource accounting or capacity-planning programs.

## Key Environments for Resource Control

Solaris Resource Manager software enables a Solaris server to act more like a mainframe because resources can be allocated for users, groups, and applications. It is useful in a variety of business situations including:

- Server consolidations
- Web hosting
- Sites with large or varied user populations
- Cluster environments

### *Server Consolidation*

Traditionally, UNIX® servers were configured so that there were dedicated servers for each application, designed for peak capacity. Solaris Resource Manager enables server consolidation and more effective system utilization because the main aspects of system resources can be controlled in a granular fashion. The cost and complexity of managing multiple servers is reduced as greater economies of scale are reached by hosting multiple applications on a single system.

For example, a single Sun Enterprise server (running the Solaris 2.6 Operating Environment) with Solaris Resource Manager could provide application, file, and print services for heterogeneous clients, messaging/mail service, Web service, and mission-critical database applications. Since Sun Enterprise servers scale to 64 processors, one server could be configured for several departments to share, or for an entire enterprise to use.

In other server consolidation efforts, the development, prototype, and production environments are combined on a single large machine like the Sun Enterprise 6500 server, rather than being hosted on three separate servers. Still other consolidation projects combine multiple database instances and application servers within a single machine.

## *Web/Application Hosting*

For Internet Service Providers (ISPs) as well as companies hosting Web servers, Solaris Resource Manager software can be a tremendous benefit. Instead of purchasing large numbers of small servers, multiple Web servers can be hosted on a single machine. ISPs do not have to worry about a faulty CGI script bringing the system down. To an ISP customer, it appears as if they have their own server dedicated to them when it's actually a large number of servers running on a single, highly reliable server. ISPs can easily add resources by changing allocation policies. In addition, e-commerce sites can divide resources allocated to different modes of operation. For example, the purchasing application can receive more resources over the information section of the site.

The same situation exists for application service providers. Both service providers and their customers gain economies of scale because multiple customer applications can be hosted on the same system.

## *Sites with Large or Varied User Populations*

Solaris Resource Manager software can help manage resources in any site that has a large number of users and a mixed workload.

These environments can include Sun Ray™ thin clients, servers, and Solaris™ PC NetLink settings. With Sun Ray, resources can be effectively portioned so that active users cannot starve resources from less-active users. With Solaris PC NetLink, multiple backup domain controllers, files servers, and/or application servers can be hosted on a single, reliable, available Solaris server.

Where there are mixed workloads, Solaris Resource Manager can be configured to favor certain users. For example, in large brokerage firms, traders intermittently require fast access to execute a query or perform a calculation. Other system users, however, have more consistent workloads. If the traders are granted a proportionately larger amount of processing power of shares, Solaris Resource Manager ensures that they will have the responsiveness they need.

## *Cluster Environments*

Before Solaris Resource Manager was available, the typical recommendation was to host one application per server for the maximum level of availability. However, in the event of a failure, the applications on a single system would contend for resources until the services were restored. With Solaris Resource Manager, this is no longer the case. Each application can receive a specified level of CPU resource — even in the event of a failure.

Moreover, with Solaris Resource Manager and Sun Cluster software, higher levels of availability can be cost-justified because more application can be hosted on each node.

# Controlling Resources

---

## Fine-Grained Control

Solaris Resource Manager software provides a hierarchical, user-based model of resource control that is easy to configure and administer for large groups of users. The system administrator defines a control policy for Solaris Resource Manager for users, groups, and applications, providing fine-grained control of system resources.

By default, the Solaris Operating Environment includes system parameters that impose limits on resources (e.g. the number of processes per user). However, these system parameters are defined within the kernel for *all* users and don't allow different values for different users. Solaris Resource Manager, on the other hand, defines a resource control policy for each individual user, or for a group (of users). With Solaris Resource Manager, a system administrator has a great deal of flexibility in defining the resource control policy for the system.

## Per-User Information

The fundamental building block in Solaris Resource Manager software is a per-user persistent object called an `lnode`, short for "limit node." An `lnode` exists for every user and group on the system. The `lnode` object contains system and user attributes that correspond to the resource control policy for that user. The kernel accesses `lnodes` to determine a user's specific resource limits and to track that user's resource usage. When a user starts a process, the user's `lnode` records the process and its activity. As the process runs and requests resources, the kernel decides whether or not the process is entitled to the requested resources based on the `lnode` information.

For fast access, a cache of `lnodes` resides in the system memory. If an `lnode` is not cached, it is automatically retrieved from disk, where `lnodes` are stored in a file. The caching of `lnodes` and kernel-based implementation means Solaris Resource Manager adds little overhead and has no noticeable impact on performance.

### *The Solaris Resource Manager Hierarchy*

To facilitate easy administration, Solaris Resource Manager software enables related users to be grouped together and administered as a whole. Within Solaris Resource Manager, an `lnode` can represent either a user or a group (of users). At the top of the Solaris Resource Manager hierarchy is an `lnode` for root, with each user or group `lnode` identifying its parent within the hierarchy or by default root. All resources are available at the root node, and are divided hierarchically among the `lnodes` below it in the tree.

This easy administration enables the system administrator to allocate resources to different departments and workgroups as well as fine-tune the allocations within those departments. Since *root* takes precedence over all other users, it can always get the required resources it needs to do any necessary tasks.

### *Configuration*

Through the command-line interface, the system administrator can use configuration commands to create the system's resource control policy. The configuration commands define both system and user attributes for the `lnode` objects. System attributes are used directly by the kernel, such as those used to control processes and virtual memory. User attributes specify the resource limits for users, groups, or applications.

Solaris Resource Manager software can be configured to delegate authority. This allows each group administrator to tune the configuration based on usage within a particular designated group.

## How Solaris Resource Manager Controls Resources

In general, Solaris Resource Manager software controls resources in two ways:

- *Limiting* – imposes limits on resource use
- *Accounting* – tracks a user's current and past resource usage

## *Types of Resources*

Generally, all system resources can be divided into one of two classes: *renewable* resources and *fixed* (or *non-renewable*) resources. Solaris Resource Manager manages these two types of resources differently.

- Renewable Resources

Renewable resources are those which are in continuous supply, such as CPU ticks which are replenished over time. Solaris Resource Manager monitors the rate at which these resources are used, and prioritizes allocations based on pre-set policies and competing demands. Over time, Solaris Resource Manager discounts past usage by applying a decay factor that gives recent usage more significance, so new or infrequent users cannot get “locked-out” of using CPU resources.

- Fixed Resources

Fixed or non-renewable resources are those which are available in a finite quantity, such as virtual memory, number of processes, and number of logins. Solaris Resource Manager tracks the usage of these resources and imposes limits to prevent over-consumption.

## An Example

Solaris Resource Manager provides a utility, `liminfo`, that lets a user view what resource controls are in effect. `liminfo` gives an overview of the information in an `lnode` object.

The sample `liminfo` output in Figure 1 shows information about the user and each resource utilized: CPU, virtual memory, and number of processes.

The `liminfo` output includes both the limit, usage, and accrual data for each resource.

```

$ liminfo
Login name:                michael      Uid:    223  R,Euid:  225,225
Sgroup (uid):              cpr (98)  Gid:    50  R,Egid:  50,50

CPU
  Shares:    5  Share:    3.57 %  Usage:    76000
  Myshares:  1  E-share:  2.37 %  Accrued usage:  6.46e+08

Memory
  Mem usage:    11.06 MB
  Mem limit:    50.00 MB
  Proc mem limit: 20.00 MB
  Mem accrue:   13.67 TB.s

Last used: Mon Feb 23 16:21:47 1998
Directory: /usr/people/michael
Name:      Michael Sebree
Shell:     /usr/local/bin/bash
Umask:     0000 (ugtrwxrwx)

Processes:    25
Process limit: 40
Current logins: 2

Flags: uselim+
Device      Limit      Usage      Decay/Interval      Accrued
Connect-time 16h      8h354s     16h/1w              1w1d17h3m

```

FIGURE 1 Sample liminfo output

All the information is stored in a database called srmDB, which can be accessed and changed with various Solaris Resource Manager commands such as `limreport`, `srmrates`, and `liminfo`. These commands enable users to access usage and limit information. System administrators can update the database at any time using `limadm`.

### CPU Time

CPU time is a renewable resource, rather than a fixed resource like virtual memory. Solaris Resource Manager software controls the rate of CPU usage through *fair-share scheduling*. Solaris Resource Manager introduces a new scheduling class into the Solaris Operating Environment — the “SHR” scheduling class. Instead of assigning processes to the timesharing scheduling class (“TS”), when Solaris Resource Manager is enabled, the Solaris Operating Environment assigns processes to the SHR class. The scheduler then allocates time slices to each SHR process based on the user’s *entitlement* and *usage*.

The system administrator assigns each group and user within the Solaris Resource Manager hierarchy a number of CPU shares, similar to shares in a company. The number of shares, recorded in the `lnode` object, is the user’s CPU *entitlement*.

In comparing entitlements, it is the ratio of shares between peers that is important, not the actual number of shares. For example, in Figure 2, the system administrator allocates 30 shares to R&D, 10 shares to administration, and 60 shares for mission-critical users. If all three groups are requesting their full amount of CPU time, their entitlements work out to 30 percent, 10 percent, and 60 percent, respectively.

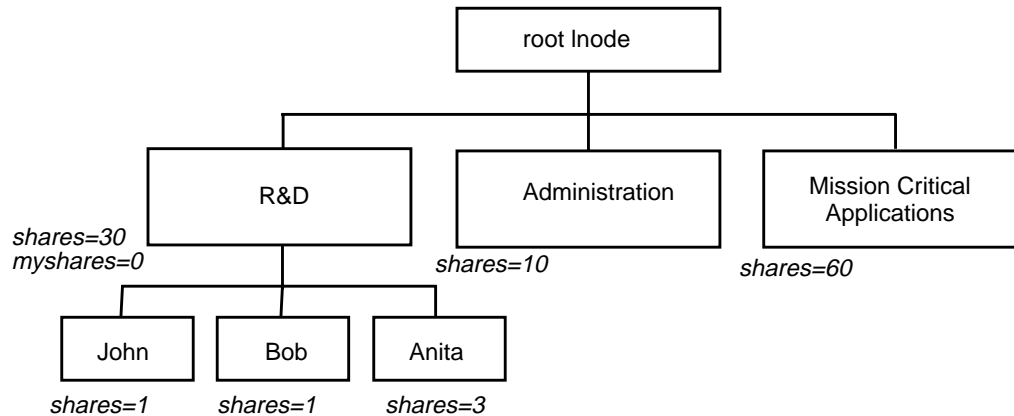


FIGURE 2 Comparing CPU entitlements

However, Solaris Resource Manager allocates CPU time according to the ratio of entitlements for *active* processes; no CPU resources are ever wasted. If users in the administration group are not requesting CPU resources, then the available resources are divided among the remaining two groups: 33 percent for R&D and 66 percent for mission-critical users. If a fourth department is added to the hierarchy, then it is not necessary to adjust the shares of the other three because it is the *ratio* of shares that matters.

The Solaris Resource Manager hierarchy makes it easy for the system administrator to establish CPU entitlements. In Figure 2, the users within the R&D department are subject to the group entitlement, which is split among the active group members according to the ratio of their shares. In this example, Anita has 3 out of a total of 5 shares, or 60 percent of the group's 30 share allotment. John and Bob each have 20 percent of the group's total. If the group's parent in the hierarchy is an actual user (for example, the R&D department head), then their portion of the group's shares is expressed in the lnode's *myshares* attribute. Here *myshares* is set to zero, which means that the R&D lnode represents a placeholder in the hierarchy rather than an actual user. Solaris Resource Manager offers the system administrator a great deal of flexibility in configuring the hierarchy and entitlements.

Solaris Resource Manager software internally tracks all CPU *usage* in the SHR scheduling class. As a process consumes CPU ticks, Solaris Resource Manager stores the costs of these actions in the corresponding `lnode`. To schedule processes fairly, Solaris Resource Manager takes into account both the user's entitlement and the pattern of their past CPU usage. Over time, users who consistently use more CPU time than their entitlement receive lower allocations than lighter users so that light users do not get "locked out."

Solaris Resource Manager progressively decays past usage so that only the most recent usage is significant. The Solaris Resource Manager system administrator sets a *half-life* parameter which controls the rate of decay. A long half-life favors a steady rate of consumption, typical of long-running batch jobs, while a shorter half-life favors interactive users by quickly "forgiving" high bursts of consumption. Using scripts to shorten decay rates each morning and lengthen decay rates each night results in a system that favors interactive use during the day and batch jobs at night.

One distinct advantage of the Solaris Resource Manager scheduler is that it schedules processes to achieve user outcomes, not process outcomes. With Solaris Resource Manager, the system administrator can assign a proportionally larger entitlement to users who run mission-critical applications, thus guaranteeing them a larger share of processing power when needed. During quiet periods, lower priority applications can still take advantage of the otherwise unused system capacity.

### *Virtual Memory*

System memory is a fixed resource, and Solaris Resource Manager imposes limits on virtual memory consumption. Solaris Resource Manager tracks memory use in two ways:

- Overall virtual memory used
- Per-process virtual memory used

Solaris Resource Manager software calculates and limits the amount of virtual memory used by all processes with the same `lnode` and at all levels in the `lnode` hierarchy. A limit is also imposed on per-process virtual memory (In Figure 1, the overall virtual memory limit is 50 Mbytes while the per-process limit is 20 Mbytes). Thus, a single user is prevented from unfairly monopolizing virtual memory. Consider an application with a memory leak that continuously requests memory without freeing it. Solaris Resource Manager allows system administrators to put a cap on the amount of virtual memory this application could receive, preventing it from starving other processes and users.

### *Number of Processes*

Solaris Resource Manager software imposes a limit on the number of processes that a user may run at any one time. The output of `liminfo` in Figure 1 shows the user “michaels” is currently running 25 processes from an allowable 40.

### *Resource Parameters*

System global parameters, such as decay rate and daemon options, define the overall performance of all `lnodes`. Simple one-line commands enable Solaris Resource Manager to be turned off if necessary, so that the system looks entirely as it did prior to installing Solaris Resource Manager.

## Accounting

Solaris Resource Manager software keeps track of CPU, virtual memory, number of processes, and number of logins used within the `lnode` object. (In the `liminfo` output in Figure 1, the `usage` and `accrue` fields display this information.) The `lnodes` continue to accumulate usage and accrual data until the system administrator re-initializes the counters.

Solaris Resource Manager software includes tools to develop a simple billing system. The system administrator must define costing functions for resource use. `lnode` information is stored in the `srmDB` and can be accessed by simple tools.

In addition, the system administrator can extract the raw accrual data, which can then be input into other reporting, capacity planning, and chargeback tools.

## Managing Resources

---

Solaris Resource Manager software includes the tools and interfaces that system administrators require.

### Delegating Resource Control

It is easy to configure the Solaris Resource Manager resource control policy. The primary reason for this is that Solaris Resource Manager is hierarchical, and all resource allocation works within that hierarchy. In most organizations, the system administrator configures the Solaris Resource Manager hierarchy to match that of the actual organization. In doing so, the system administrator can easily match organizational goals with the resource allocation policy. If the system administrator chooses, he or she may then delegate the decision-making for resource allocation to other administrators in the hierarchy. Delegating the Solaris Resource Manager administrative duties has a positive side effect — it passes responsibility and authority to deal with resource contention issues to the parts of the organization in which they occur.

To configure Solaris Resource Manager policy, a sub-administrator requires special privileges but not superuser access. The system administrator assigns these privileges via special flags in the `lnode` object. These privileged users can then make policy decisions within their workgroups, but not at any level outside of their group in the hierarchy. The privileges allow the system administrator to split up the task of resource policy decisions in a safe, secure fashion.

## Changing the Current Resource Policy

With Solaris Resource Manager software, it's easy for the system administrator to activate alternate resource policies. Once the resource policy is configured, Solaris Resource Manager controls resources automatically without any further intervention. However, system administrators typically want to do some additional fine-tuning of the Solaris Resource Manager policy. This can be done while the system is operating.

Some sites may implement new policies based on the time of day or time of year. For example, an organization's finance department may require additional resources at year end. Differing policies might be more appropriate outside normal working hours versus during normal hours. Using scripts policies used in Solaris Resource Manager can be automatically changed on a periodic basis.

## Tools

Solaris Resource Manager software provides a powerful set of commands for scripting. Included are utilities for:

- Configuring and monitoring the operation of Solaris Resource Manager
- Modifying and deleting `lnode`
- Manipulating processes attached to specific `lnodes`
- Extracting data from `lnodes`

In addition, Solaris Resource Manager provides a command to view their effective resource policy.

Within Solaris Resource Manager environments, system administrators could use the available commands to implement scripts for planning or billing purposes, or to invoke an alternate resource control policy.

## Summary

---

IT managers are responding to business needs by implementing service-level agreements to control costs and guarantee service levels for their enterprise applications. They realize that today's system model of hosting one application per system for more predictable service availability is outmoded and expensive, and leads to a proliferation of systems that are complex and inefficient to administer. As a result IT managers are consolidating a variety of applications on larger, more scalable servers to achieve mainframe-class RAS capabilities at a fraction of the cost.

Solaris Resource Manager software is a key enabler for server consolidation and increased resource utilization, because it enables system administrators to control and allocate major system resources to users, groups and applications.

Solaris Resource Manager is unique in its ability to control as well as monitor system resources. Its hierarchical model offers system administrators an easy and flexible way to implement corporate policies for users, groups and applications. Solaris Resource Manager software's innovative approach to CPU scheduling — allocating shares rather than setting a flat percentage — simplifies the task of adding new users and new resources.

Solaris Resource Manager provides fine-grained control of system resources, making it easy to configure an effective resource control policy. With Solaris Resource Manager software, mission-critical applications always have access to the crucial system resources they need.

## Additional Resources

---

Sun has a wealth of information available on how to configure and analyze Solaris Resource Manager in a wide variety of configurations and application environments. These include:

- Sun Blueprints: Resource Management

A practical guide that shows how to manage system resources to meet system and performance service levels. The theory of resource management is introduced and a detailed analysis of what is required to implement resource management in the Solaris Operating Environment is explored. The guide includes how Solaris Resource Manager facilities can be used with the following workloads:

- Databases
  - File service
  - Web servers
  - Batch tasks
  - ISP configurations
- Solaris Resource Manager on Sun Cluster 2.2 White Paper
  - Resource management articles (<http://www.sun.com/blueprints/>)
  - Solaris Resource Manager 1.1 Administration Guide (<http://docs.sun.com>)



Sun Microsystems, Inc.  
901 San Antonio Road  
Palo Alto, CA 94303

1 (800) 786.7638  
1.512.434.1511

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

September 1999