



Solaris Containers & Solaris™ 10 Operating System

Teresa Majer

Data Center Practice

Sun Microsystems, Inc.



Solaris 10: A Generation Ahead



Unparalleled Security



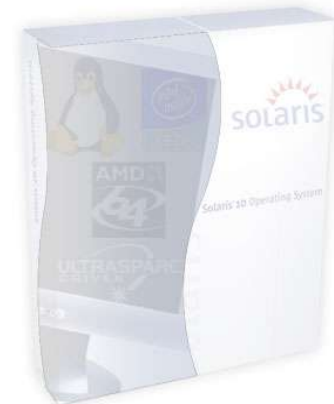
Relentless Availability



Optimal Utilization



Extreme Performance



Platform Choice

Solaris Containers

Consolidation Made Simple, Safe and Secure

- Breakthrough approach to virtualization
 - Virtualized application environment
 - Host 1,000s of applications/services on one system, with a single OS instance
 - Hardware independent
- Superior resource utilization
 - Dynamically adjust to business goals
 - Less than 1% system overhead
 - Easy to create, replicate and maintain
- Significant increase in uptime and security
 - Each service fault- and intrusion- isolated
 - Instant Restart: containers start in seconds
- Reduced costs
 - Simplifies and accelerates consolidation



Solaris Containers

Definition



Container Principals

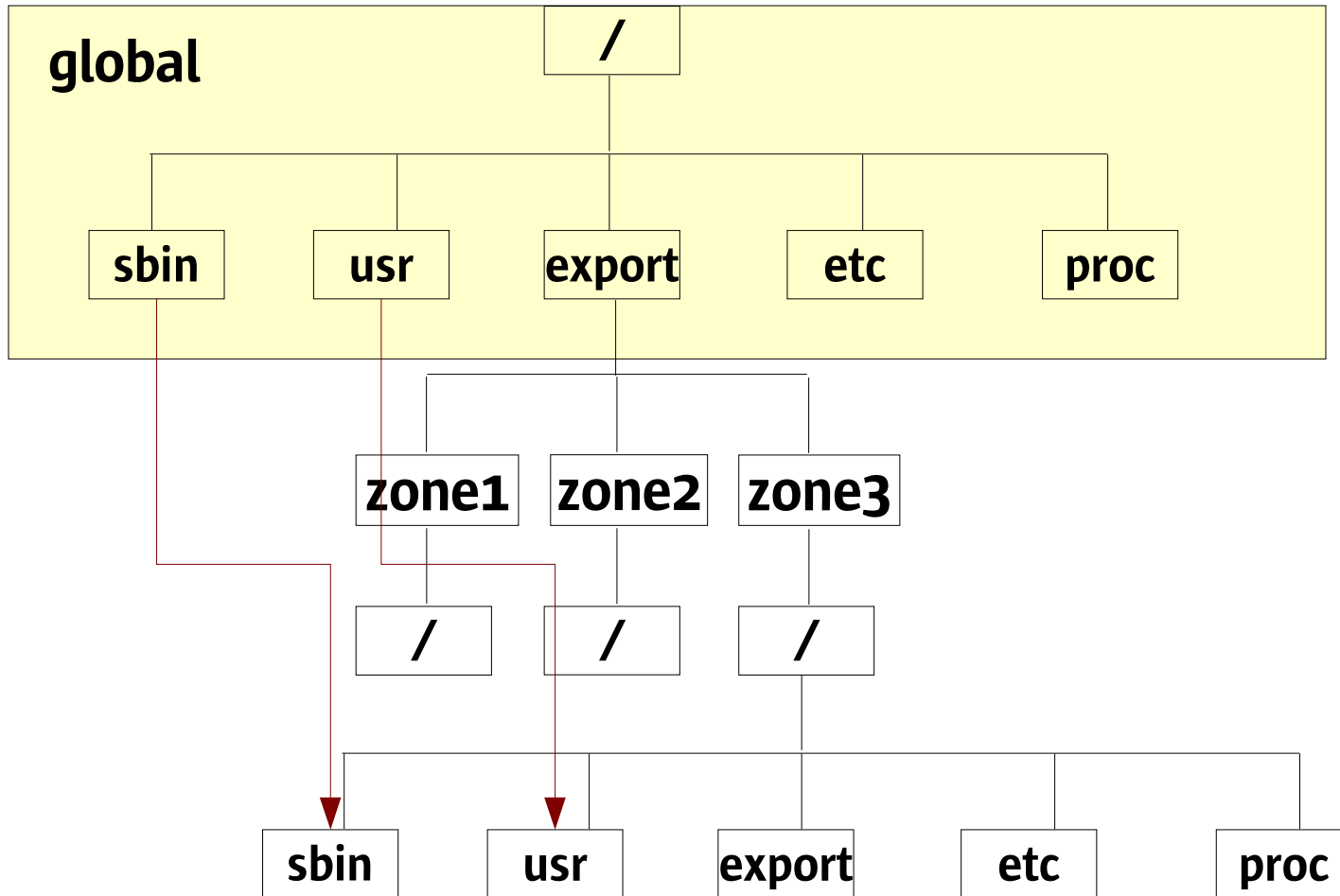
- Isolation
- Virtualization
- Granularity
- Transparency

Container Principals

Isolation

- Everything in a local Zone is isolated from things in other local zones.
- A zone contains :
 - processes, files, interfaces, and System V IPC elements.
- Includes a separate root file system
- Includes separate users (including root)
- Faults and Security

Isolation: root filesystem



Container Principals

Virtualization

- A zone appears as an instance of Solaris

An instance within an instance
zched and init

- Zones can be administered separately

Separate name services
Separate users, patches, packages

Container Principals

Granularity

- Partitions at sub-CPU level (FSS)
- Projects, Tasks, Pools & Privileges
- Only file system space is dedicated
- HW resources multiplexed the zones (IPQos)
- Resource allocation such as cpu, memory, bandwidth can be used

Container Principals

Transparency

- Applications run as is
- Standard Solaris interfaces are provided
- Restrictions limited to privileged operations
 - No global reboots, access to `/dev/kmem`, physical devices

Fair Share Scheduler

“CPU shares are not equivalent to percentages of CPU resources. They are used to define the relative importance of one workload in relation to other workloads.”

project1 -> 10 shares -> no less than 40% if needed

project2 -> 10 shares -> no less than 40% if needed

project3 -> 5 share -> no less than 20% if needed

25 total shares

project1 % = project1 shares / (\sum all **active** shares)

If project2 is inactive, then project1 could get 67% [10/15 = 67%]

Fair Share Scheduler

- Shares limit usage of CPU resources only when there is competition
- CPU cycles are never wasted
- More fine grained control than processor sets
- Projects are the workload containers
- Shares are allocated in `/etc/project`
 - `project.cpu-shares`
 - `zones.cpu-shares`
- `dispadm -d FSS`
- `prctl` for temporary modifications
- System project gets what it needs despite share allocation

Memory Resource Control

- `rcapd` is the daemon that controls memory usage
- Max values per project are defined in `/etc/project`
- `rcapadm` to configure `rcapd`
- `rcapstat` to monitor usage

Zone Administration

The Zone Commands

<code>zonecfg</code>	Used to create and modify the configuration of a zone.
<code>zoneadm</code>	The <code>zoneadm</code> utility is used to administer system zones.
<code>zlogin</code>	Used to enter a zone.
<code>zonename</code>	Can be used to give the name of the zone after <code>zlogin</code>

Solaris Containers



1) Zone Creation

```
bash-2.05b# zonecfg -z zone1
```

```
zone1: No such zone configured
Use 'create' to begin configuring a new zone.
zonecfg:zone1> create
zonecfg:zone1> set zonepath=/zone/1
zonecfg:zone1> set autoboot=true
zonecfg:zone1> add net
zonecfg:zone1:net> set address=192.168.35.210
zonecfg:zone1:net> set physical=hme1
zonecfg:zone1:net> end
zonecfg:zone1> commit
```

2) Zone List

```
bash-2.05b# zoneadm list -vc
```

ID NAME	STATUS	PATH
0 global	running	/
- zone1	configured	/zone/1

3) Zone Install

```
bash-2.05b# zoneadm -z zone1 install
```

```
Preparing to install zone <zone1>.
Creating list of files to copy from the global
zone.
Copying <2521> files to the zone.
Initializing zone product registry.
Determining zone package initialization order.
Preparing to initialize <808> packages on the zone.
Initializing package <7> of <808>: persen complete:
0%
Successfully initialized zone <zone1>.
```

4) Zone Boot

```
bash-2.05b# zoneadm list -vc
```

ID NAME	STATUS	PATH
0 global	running	/
- zone1	installed	/zone/1

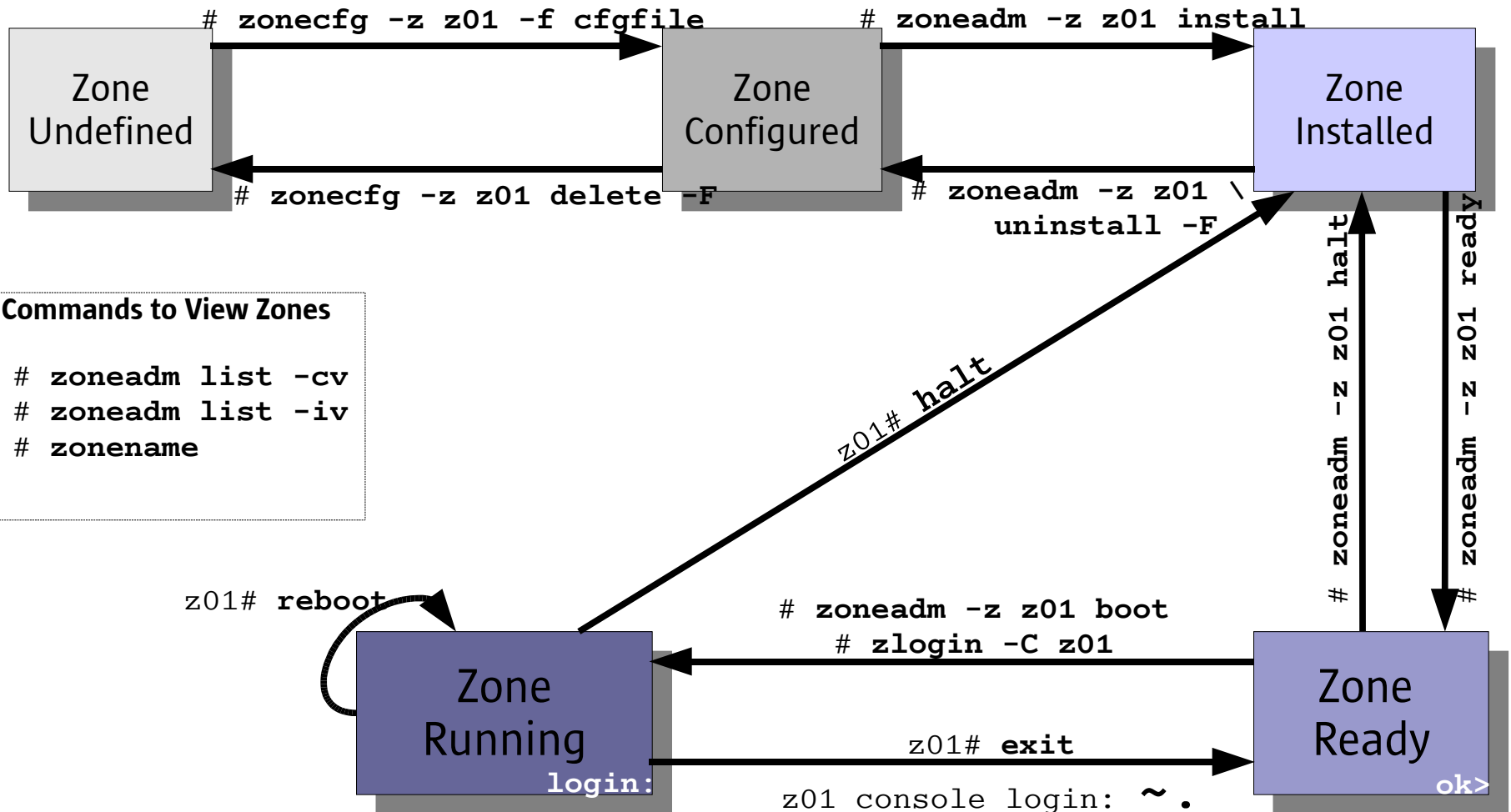
```
bash-2.05b# zoneadm -z zone1 boot
```

```
bash-2.05b# zoneadm list -vc
```

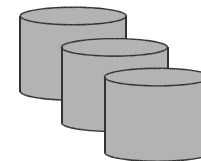
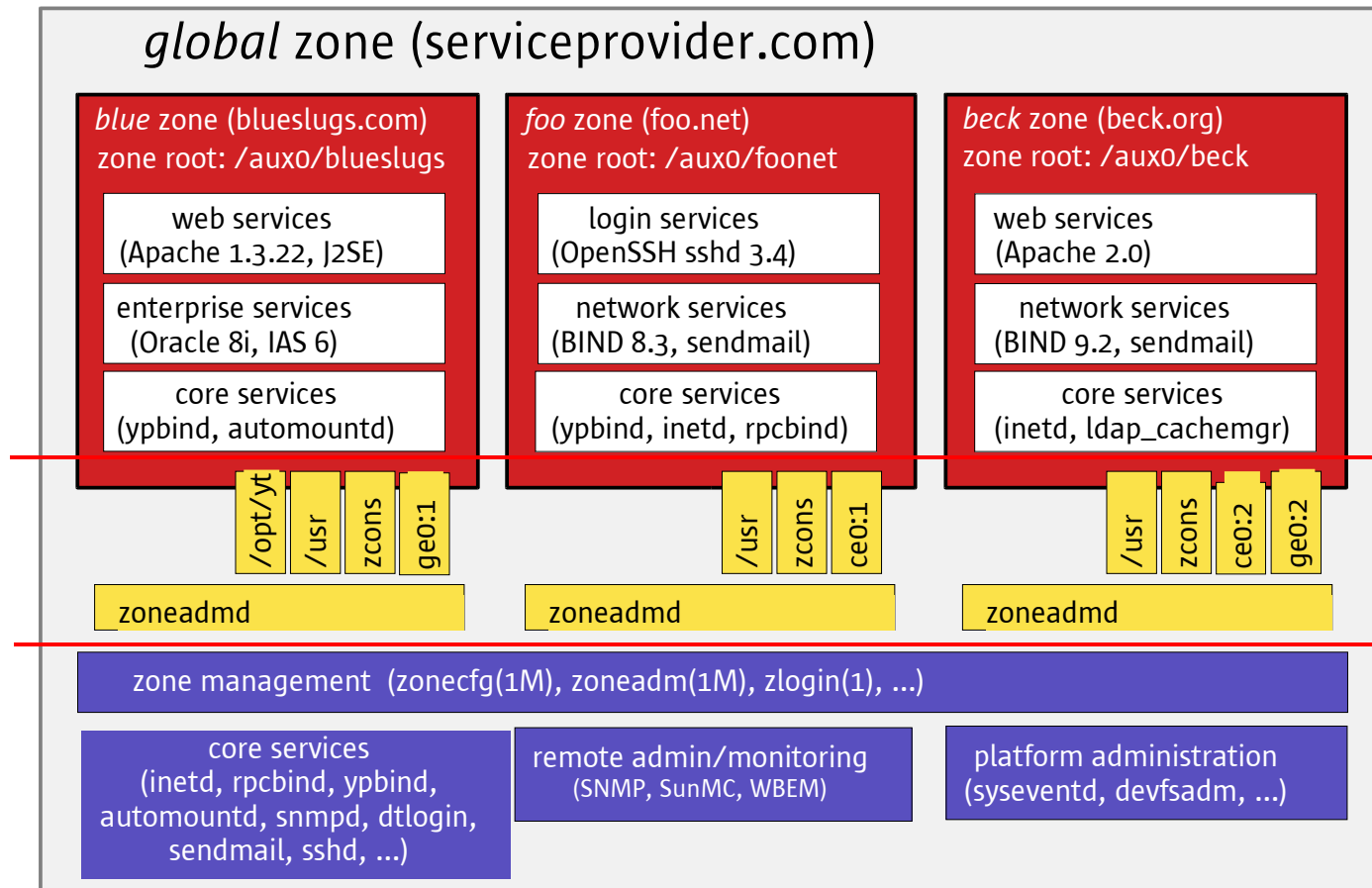
ID NAME	STATUS	PATH
0 global	running	/
2 zone1	running	/zone/1

Zone States

Commands to Create and Administer Zones



Zones Block Diagram



Zones – global vs. local

Global

- Patch kernel
- Run Dtrace
- Change zone config (ip addr)
- IPQoS

Local

- Patch applications
- Change own root password
- Add users
- Different name service

Not allowed:

- Dtrace
- Access to physical devices unless given by globalzone
- NFS server
- Dynamic ISM (Oracle)
- Jumpstart server (NFS)
- IPfilter
- prtdiag
- /etc/system
- DHCP

Solaris Container Manager



The screenshot shows the N1 Grid Console - Container Manager web interface. The browser window title is "N1 Grid Console - Container Manager - Netscape". The address bar shows "https://my-server:6789/containers/common/SCMIndex". The interface includes a navigation menu with "Hosts" and "Containers" tabs. The "Containers" tab is active, showing a list of containers for "host1". The list includes columns for "Container Name", "Status", "Resource Pool Name", and "CPU Reservation (CPUs)".

host1 - Containers

To view a container's properties, click the container name. To activate, deactivate, or delete a container, select the checkbox next to the name and click the appropriate button. [More on Host - Containers](#)

Containers (1 - 9 of 9)

<input type="checkbox"/>	<input type="checkbox"/>	Container Name ▲	Status ▲	Resource Pool Name ▲	CPU Reservation (CPUs)
<input type="checkbox"/>	<input type="checkbox"/>	Default (Read Only)	active	pool_default	0.01
<input type="checkbox"/>	<input type="checkbox"/>	music	active	pool_default	0.2
<input type="checkbox"/>	<input type="checkbox"/>	payroll	active	pool_default	0.3
<input type="checkbox"/>	<input type="checkbox"/>	Processes with No Project (Read Only)	active	pool_default	0.01
<input type="checkbox"/>	<input type="checkbox"/>	Root User (Read Only)	active	pool_default	0.01
<input type="checkbox"/>	<input type="checkbox"/>	SunMC	active	pool_default	0.3
<input type="checkbox"/>	<input type="checkbox"/>	System Processes (Read Only)	active	pool_default	0.01
<input type="checkbox"/>	<input type="checkbox"/>	Users with Group Staff (Read Only)	active	pool_default	0.01
<input type="checkbox"/>	<input type="checkbox"/>	web-sales	active	pool_default	0.2

Page: 1 of 1 Go

Solaris Container Manager

- Add-on to Sun Management Center (SunMC 3.5U1)
- Simplify configuration of Solaris Resource Manager 1.3
- Resource allocation, organization and control
- Resource usage monitoring
- Gathering of extended accounting information

Solaris 10 Resources

Free Solaris 10 New Features Web Training:

<http://training.sun.com/US/catalog/courses/WS-9000-S10-90.html>

Sun Web Site:

<http://www.sun.com/software/solaris/10/index.jsp>

Documentation Web Site:

<http://docs.sun.com/db/prod/solaris.10>

Solaris Zones on Big Admin:

<http://www.sun.com/bigadmin/content/zones/>

Solaris Express (download site):

<http://www.sun.com/software/solaris/solaris-express/>



Thank You

Teresa.Majer@sun.com
Sun Microsystems, Inc.

