

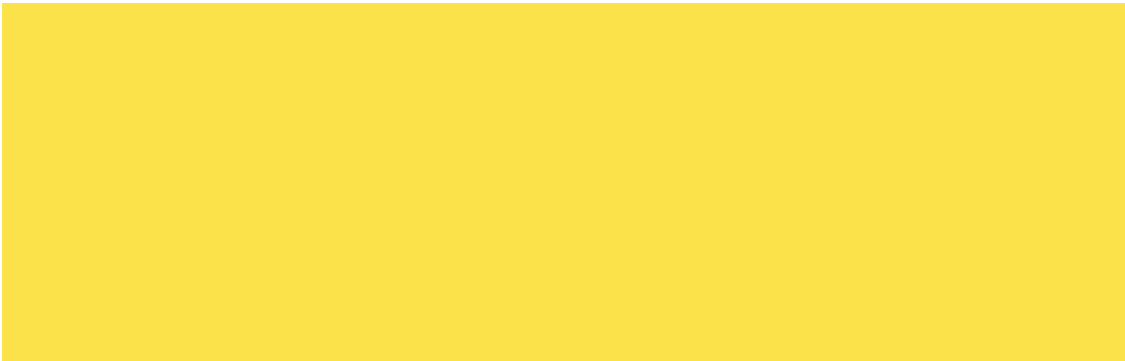


Customer
Engineering
Conference
2003

February 28 - March 3
Colorado Convention Center
Denver, Colorado

Solaris Containers

• Server Virtualization using Zones



All customers must sign the Confidential Disclosure Agreement

This presentation contains information related to projects that are in the planning and/or development stages. All aspects of this project, including but not limited to funding, availability, shipping dates, configurations, capacities, performance, and all other characteristics are subject to change and/or cancellation without notice.

This material is proprietary and confidential to Sun Microsystems, Inc. and may only be disclosed to non-employees of Sun Microsystems, Inc. under the terms of a signed confidential-disclosure agreement (CDA).

Under no circumstances will a hard copy or electronic copy of this presentation be left with, or forwarded to, non-employees or unauthorized contractors of Sun Microsystems, Inc.



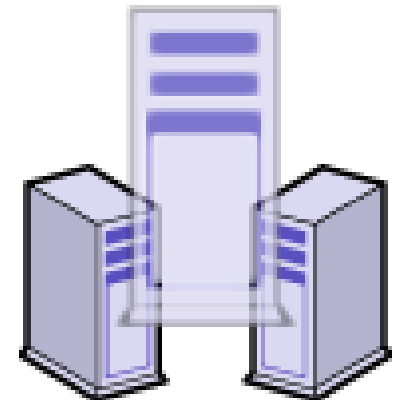
“What we're working on now is enabling customers to virtualize more of their resources into a single, easy-to-manage pool.”

Scott McNealy,

InformationWeek *Jan 27, 2003*

Agenda

- Containers & Virtualization
- Introduction to Zones
- Technology Comparison
 - LPAR, Domain, Zone
- Solaris Zone Features
- Example Architectures
- Further Information



Container Virtualization Tools

RBAC Sun Cluster 3.x

Dynamic System Domains

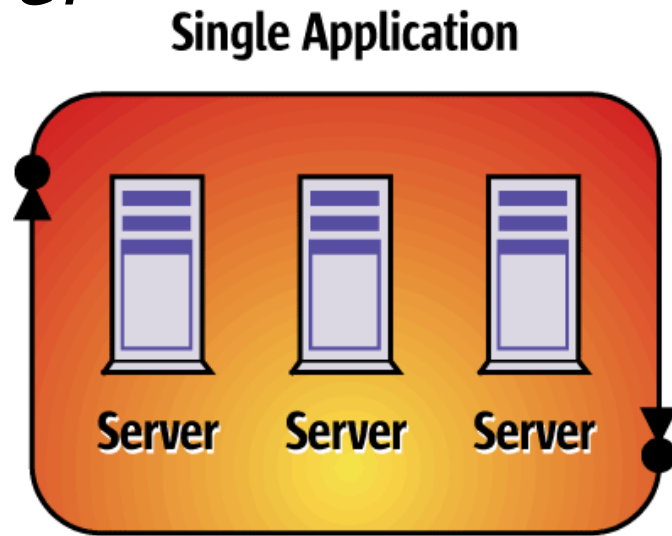
IPQoS

Sun ONE Grid Engine

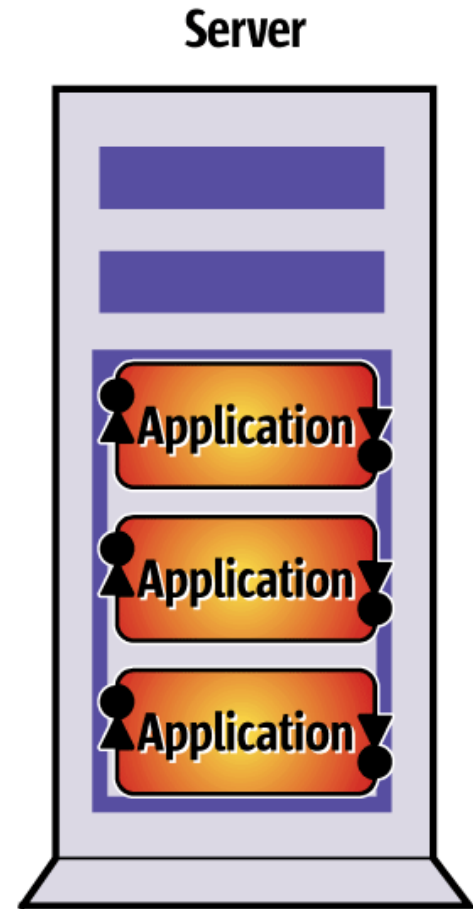
S9RM Solaris Zones

Server Virtualization

Removes the 1:1 ratio of application to server



Grid = 1:N



Zone = N:1

Solaris Virtualization

Solaris Zones

A **Zone** is a virtualized operating system environment within an instance of Solaris.

A.K.A., Software Partition
Similar to BSD Jails



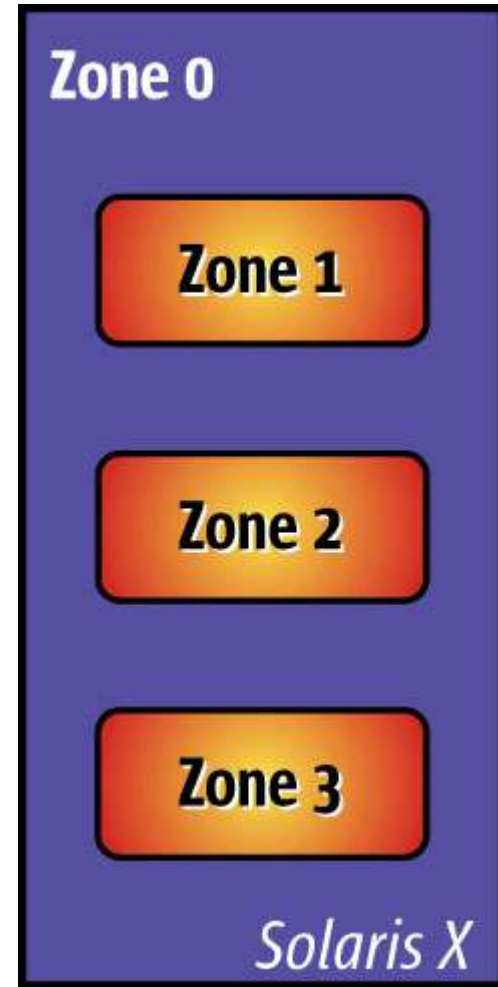
Sun Server

Solaris Virtualization

Zones allow 1 or more processes to run in isolation from other system activities.

Each *Zone* has access to:

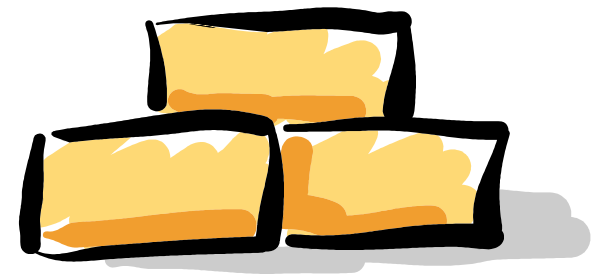
- Network Interface(s)
- Storage
- Solaris OE



Sun Server

Isolation Technologies

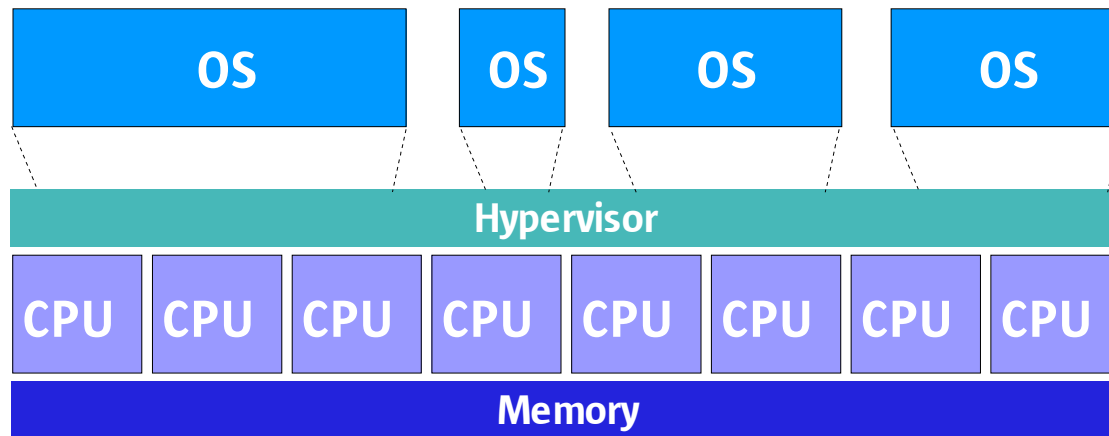
- IBM LPARS
- Sun Fire Domains
- Solaris Zones



Server Virtualization

IBM LPAR

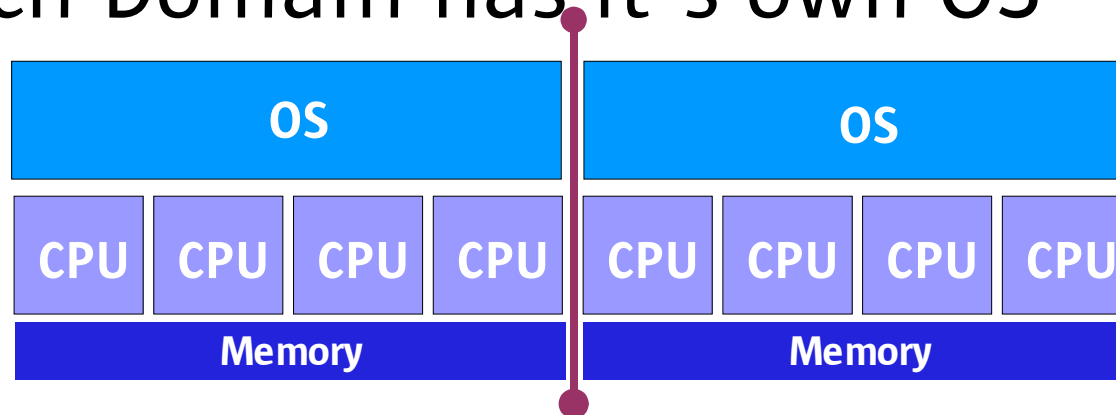
- Virtualize just above the hardware
 - Sub-CPU granularity
 - Each Partition has it's own OS
 - Overhead of multiple kernels & memory



Server Virtualization

Sun Domains

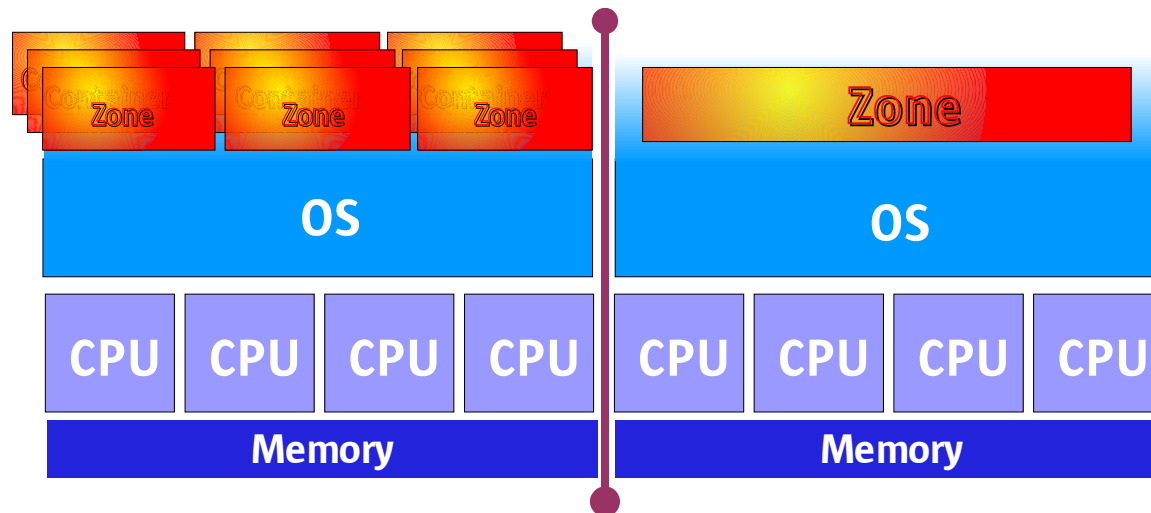
- Physical hardware separation
 - Complete Isolation and fault containment
 - 2 CPU granularity
- Dynamic resource control
- Each Domain has it's own OS



Server Virtualization

Solaris Zones

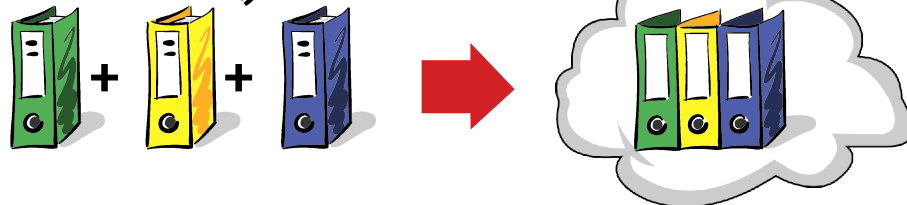
- Single Solaris instance
 - Sub-CPU granularity
 - Appearance of many OS instances
 - Minimal performance impact



Solaris Zones for Consolidation

- Software partitioning
- High resource utilization
- Repository for many small apps
- Entire Lifecycle on 1 Domain
 - Dev, Test, Stage, Production

Consolidation. Not, Availability



Features



Features of a Zone

Security

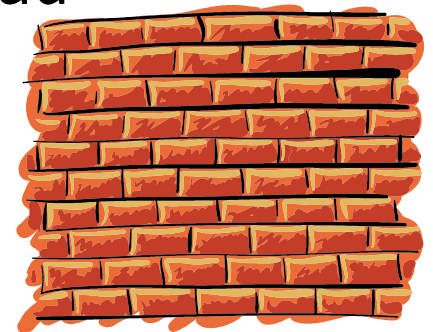
- No access to other Zones
- Restricted root access
- Functions not allowed include;
 - Reboot or shutdown of the entire system
 - Kernel memory through /dev/kmem
 - Physical devices from other Zones



Features of a Zone

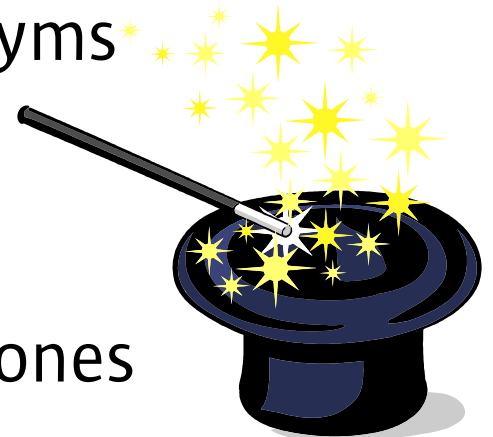
Isolation

- FS restriction similar to chroot
- Shared network port
 - No view of other traffic
- No access to objects in other Zones
 - No control, modify, monitor, or read



Features of a Zone Virtualization

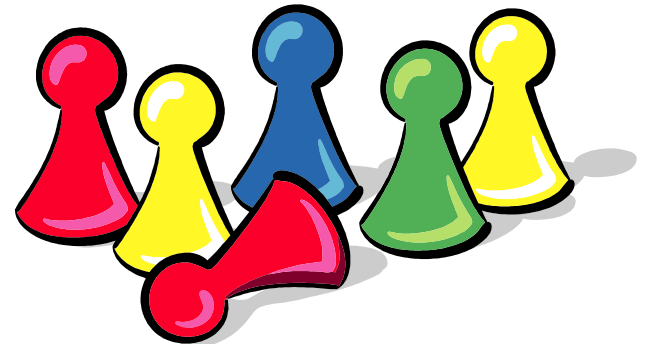
- Complete Solaris Environment
 - In appearance at least
 - Separate Zone “root” passwd
- Restricted global system state
 - kmem, lockstat, trapstst, cpc, ksyms
- Hides:
 - Physical devices
 - IP address/hostnames of other Zones



Features of a Zone

Granularity

- No dedicated physical devices
- Multiplexed resources
- Arbitrary granularity
- 100+ Zones on a 1 CPU system
 - Throttle: Disk space for unique Zone files



Features of a Zone

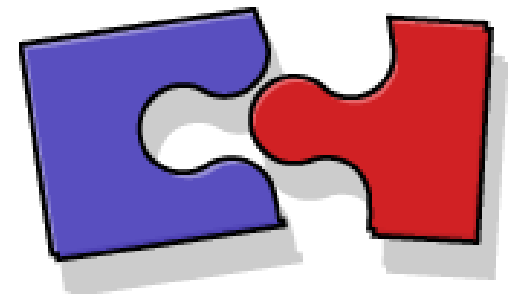
Transparency

- Standard Solaris Interfaces
 - SysV IPC shared memory segment
- Seems like a separate server
- *ps -ef* shows only current Zone
- Restrictions on low-level hardware operations

Features of a Zone

Compatibility

- Global Zone
 - Apps will run without modification
- Local Zone
 - Apps will run without modification, unless they:
 - Load custom kernel modules
 - Use physical network interfaces
 - Tested root apps include;
 - iAS, iDS, Apache, Oracle, sendmail, DNS



Features of a Zone

Resource Management

- Not aligned to HW boundaries
- Controlled by the Global Zone
- Local Zones cannot:
 - assign procs to RT scheduling class
 - create processor sets
 - lock down memory



Features of a Zone

Resource Management (Cont'd)

- Likely presented by SunMC
- SgRM Control by the Global Zone
 - Project shares honored within Local Zones



Sample Zone Usage

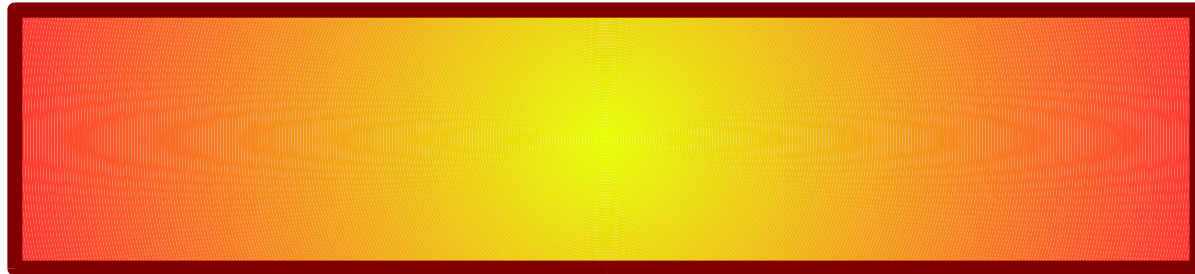


3-Tier ERP Example

- Single Sun Fire 6800
 - 24 CPUs
 - 2 Domains
 - 8 CPU Oracle license
- Web, App, DB, DNS, FTP, Development Environments

Sample Zone Usage

SF6800 with 2 Domains



Domain A
8CPU, 16G



Domain B
16CPU, 32G

Sample Zone Usage

SF6800 with 2 Domains



Zone 0: Global, nodename: construction

Oracle

Domain A
8CPU, 16G

Zone 0: Global, nodename: twilight

Domain B
16CPU, 32G

Sample Zone Usage

SF6800 with 2 Domains



Zone 0: Global, nodename: construction

Oracle

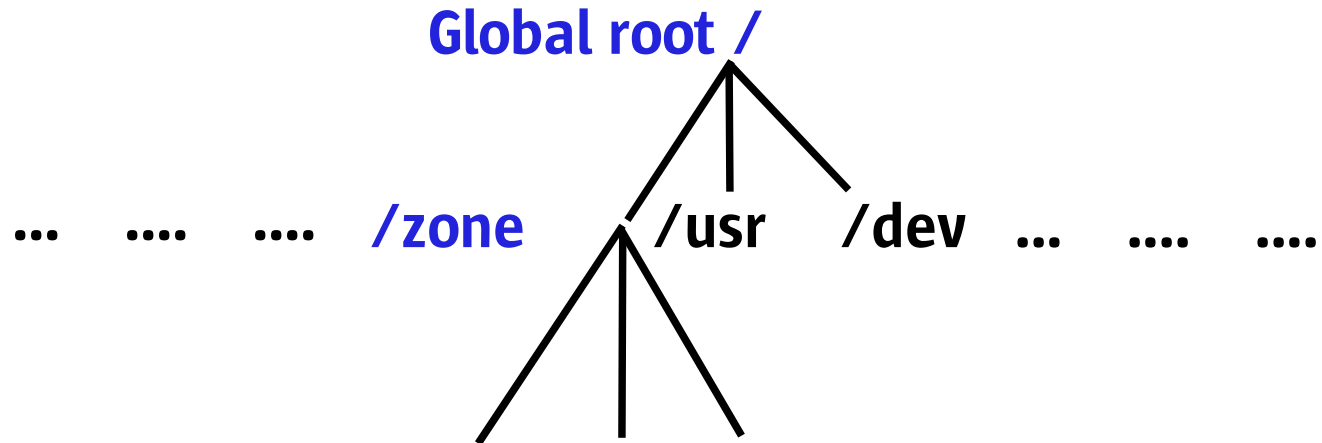
Domain A
8CPU, 16G

Zone 0: Global, nodename: twilight

**Each Global Zone has a
Supreme root login.**

Domain B
16CPU, 32G

Zone Filesystem



Zone Creation


Fresh Solaris load, Domain B

```
Terminal
Window Edit Options Help
# zoneadm info -v
ZID ZONENAME NODENAME ROOT
0 global twilight / Create
#
# zonecfg school
# zoneadm -i school Install
```

```
#
```

Zone Creation (cont'd)

```
Terminal
Window Edit Options Help
# zoneadm boot school
#
# zoneadm info -v
ZID  ZONENAME  NODENAME  ROOT
0    global    twilight  /
1    school    school    /zone/1
#
```



Sample Zone Usage

SF6800 with 2 Domains



Zone 0: Global, nodename: construction

Oracle

Domain A
8CPU, 16G

Zone 0: Global, nodename: twilight

Zone 1: school
Web Server

Domain B
16CPU, 32G

Sample Zone Usage

SF6800 with 2 Domains



Zone 0: Global, nodename: construction

Oracle

Domain A
8CPU, 16G

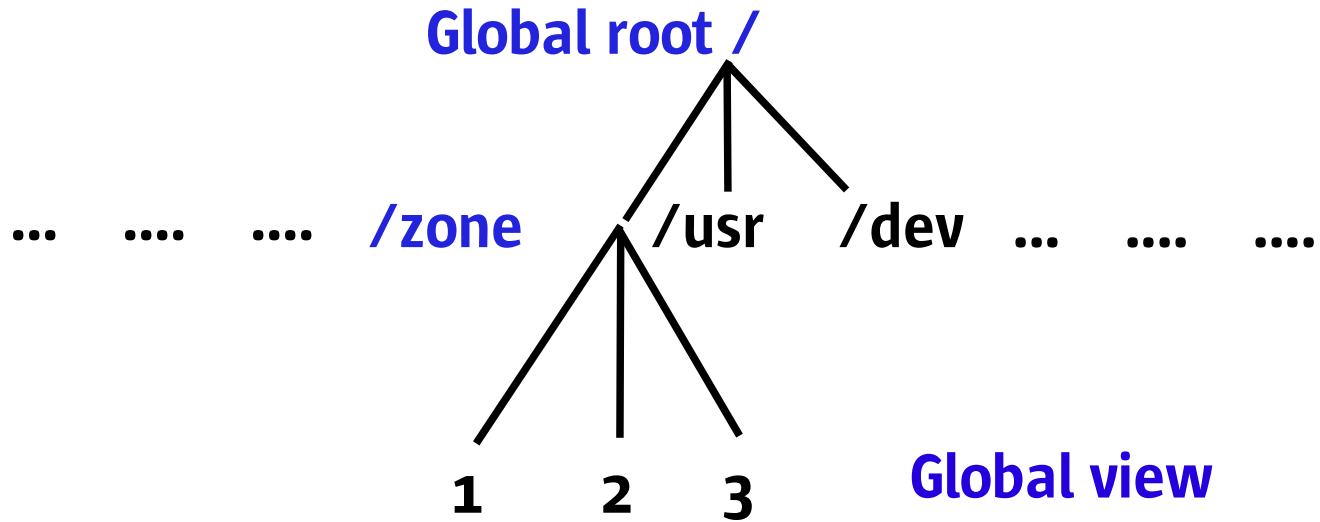
Each local Zone also has a root login for it's realm.

Zone 0: Global, nodename: twilio

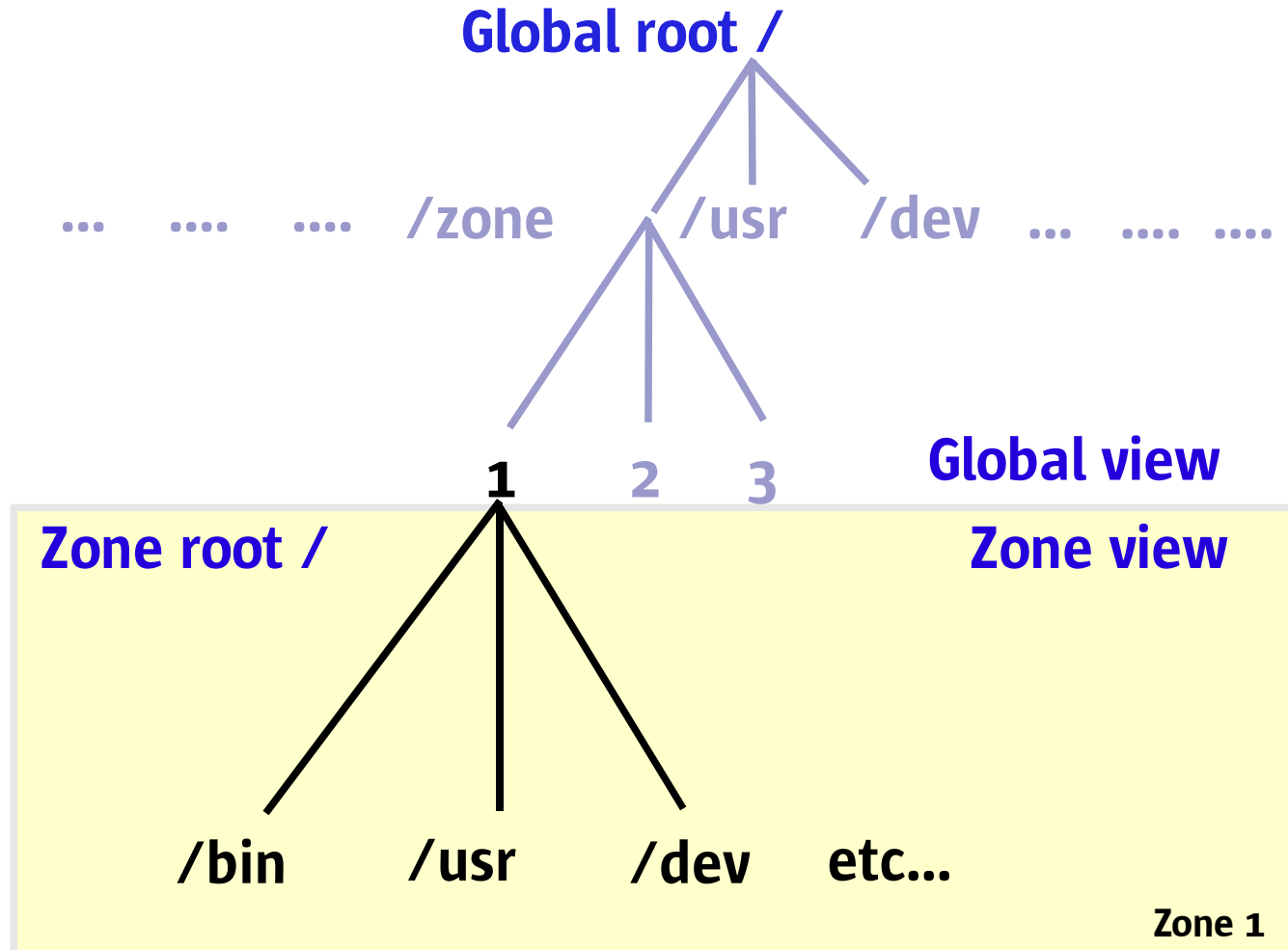
**Zone 1: school
Web Server**

Domain B
16CPU, 32G

Zone Filesystem

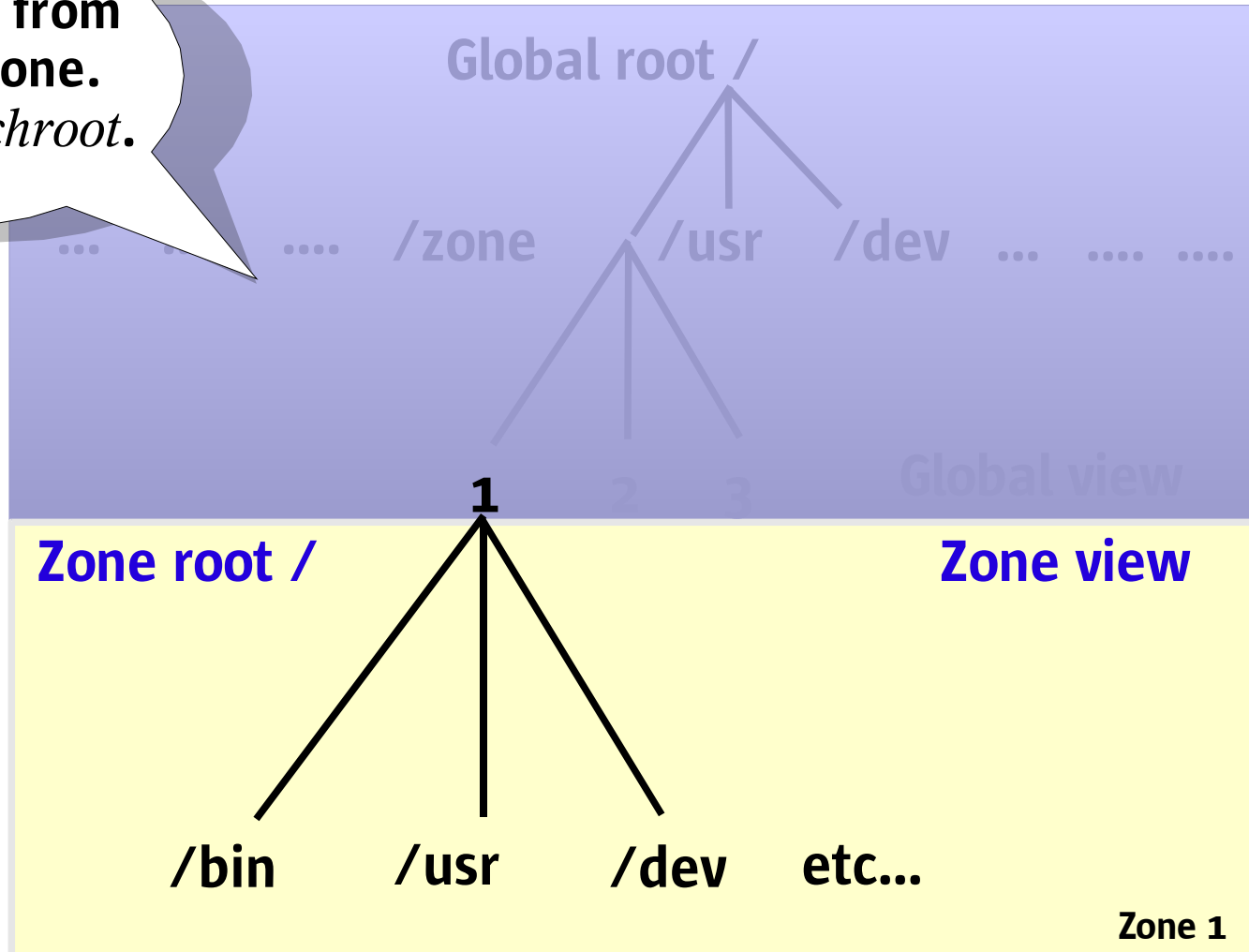


Zone Filesystem



Zone Filesystem

Not visible from the new Zone.
Similar to *chroot*.



Zone Login

From the Global Zone

```
Terminal
Window Edit Options Help
# zone
global
# zlogin school
Sun Microsystems Inc. SunOS 5.10 kevlar:Mar 2, 2003
# zone
school
#
```

Network Interfaces

Terminal

Window Edit Options

Help

```
# ifconfig -az
lo0: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4>
    mtu 8232 index 1 inet 127.0.0.1 netmask ff000000
lo0:1: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4>
    mtu 8232 index 1 zone school inet 127.0.0.1 netmask ff000000
hme0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4>
    mtu 1500 index 2 inet 129.146.86.89 netmask ffffffff00
    broadcast 129.146.86.255 ether 8:0:20:b9:37:ff
hme0:1: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4>
    mtu 1500 index 2 zone school inet 129.146.86.77 netmask ...

# zlogin school
# ifconfig -a
lo0:1: flags=1000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4> mtu 8232
    index 1 inet 127.0.0.1 netmask ff000000
```

Sample Zone Usage

SF6800 with 2 Domains



Zone 0: Global, nodename: construction

Oracle

Domain A
8CPU, 16G

Zone 0: Global, nodename: twilight

Zone 1: school
Web Server

Zone 3: comfort
DNS Server

Zone 2: towaway
App Server

Zone 4: loading
FTP Server

Domain B
16CPU, 32G

Sample Zone Usage

SF6800 with 2 Domains



Zone 0: Global, nodename: construction

Oracle

Domain A
8CPU, 16G

Zone 0: Global, nodename:

<p>Zone 1: school Web Server</p> <p>20%</p>	<p>Zone 3: comfort DNS Server</p> <p>30%</p>
<p>Zone 2: towaway App Server</p> <p>40%</p>	<p>Zone 4: loading FTP Server</p> <p>10%</p>

Domain B
16CPU, 32G

Sample Zone Usage

SF6800 with 2 Domains



Zone 0: Global, nodename: construction

Oracle

Domain A
8CPU, 16G

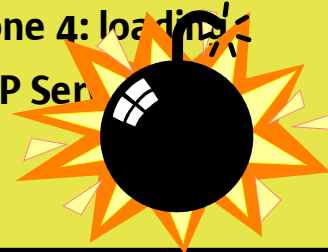
Zone 0: Global, nodename: twilight

Zone 1: school
Web Server

Zone 3: comfort
DNS Server

Zone 2: towaway
App Server

Zone 4: loading
FTP Server



Domain B
16CPU, 32G

Sample Zone Usage

SF6800 with 2 Domains



Zone 0: Global, nodename: construction

Oracle

Domain A
8CPU, 16G

Zone 0: Global, nodename: twilight

Production

Development

Zone 1: school
Web Server

Zone 3: comfort
DNS Server

Zone 5: school
Web Server

Zone 7: comfort
DNS Server

Zone 2:
towaway
App Server

Zone 4: loading
FTP Server

Zone 6:
towaway
App Server

Zone 8: loading
FTP Server

Domain B
16CPU, 32G

Sample Zone Usage

SF6800 with 2 Domains



Zone 0: Global, nodename: construction

Oracle

Domain A
8CPU, 16G

Zone 0: Global, nodename: twilight

Zone 1: school Web Server	Zone 3: comfort DNS Server	Zone 5: school Web Server	Zone 7: comfort DNS Server
Zone 2: towaway App Server	Zone 4: loading FTP Server	Zone 6: towaway App Server	Zone 8: loading FTP Server
Zone 9: school Web Server	Zone 11: comfort DNS Server	Zone 13: school Web Server	Zone 15: comfort DNS Server
Zone 10: towaway App Server	Zone 12: loading FTP Server	Zone 14: towaway App Server	Zone 16: loading FTP Server

Domain B
16CPU, 32G

Sample Zone Usage

SF6800 with 2 Domains

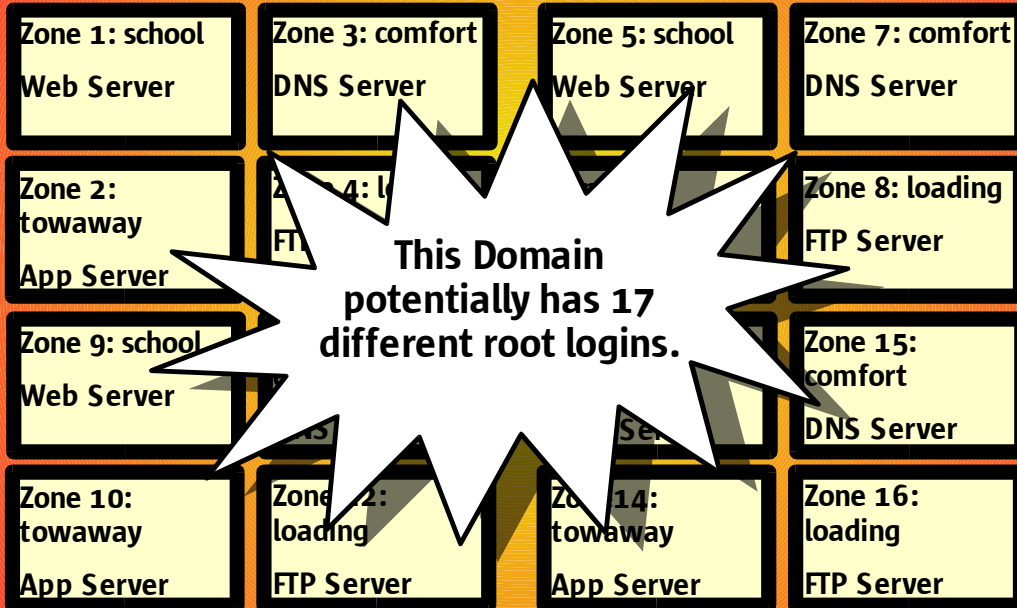


Zone 0: Global, nodename: construction

Oracle

Domain A
8CPU, 16G

Zone 0: Global, nodename: twilight

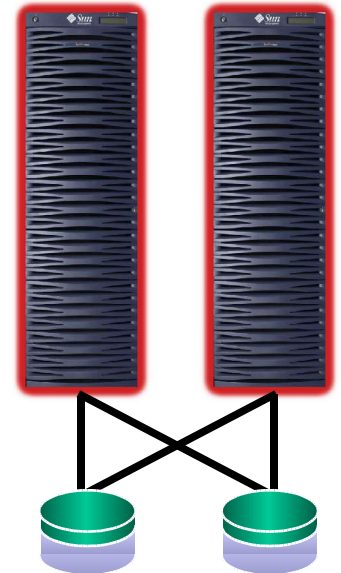


Domain B
16CPU, 32G

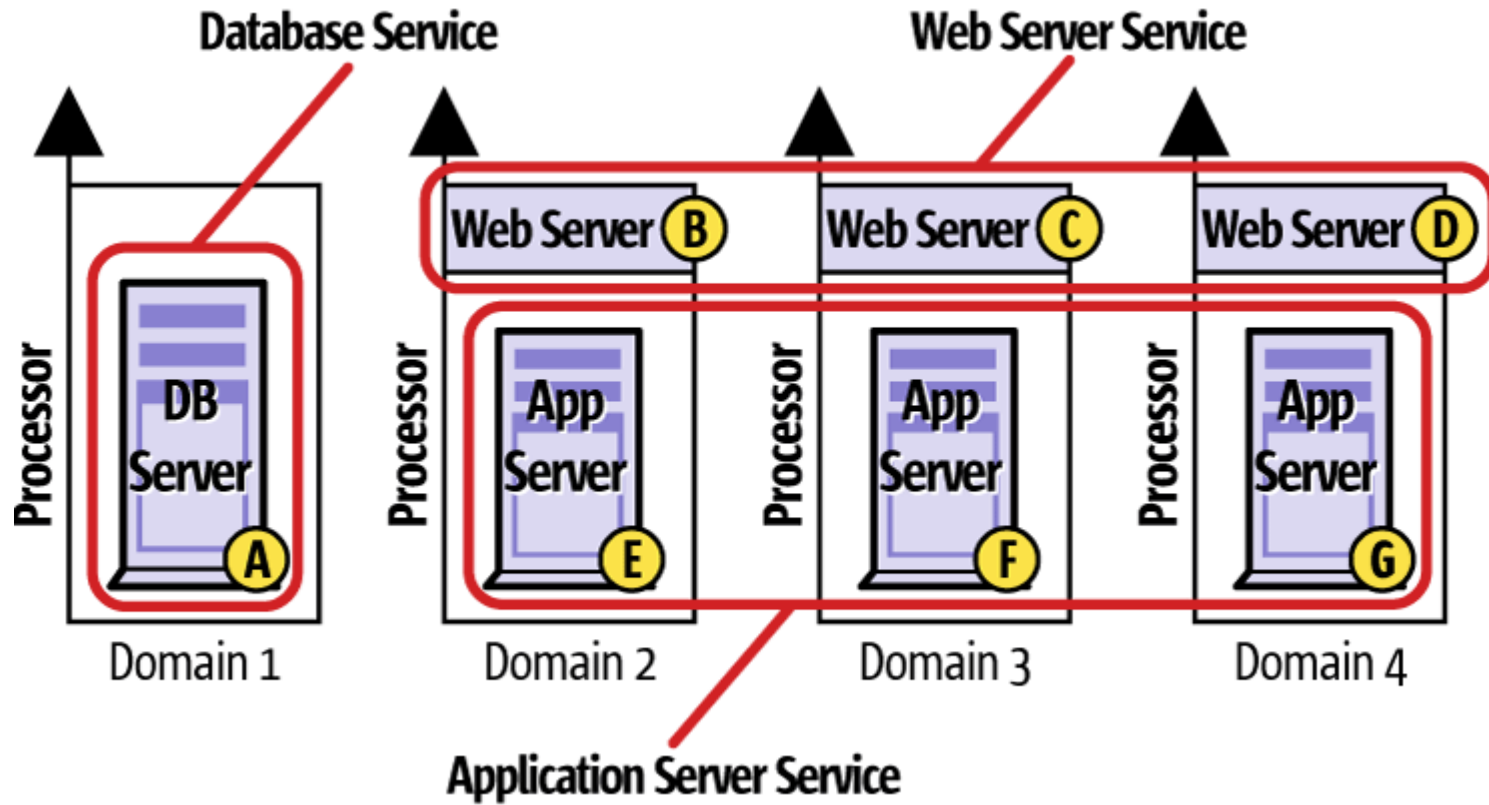
Cluster Example

3-Tier HA ERP Example

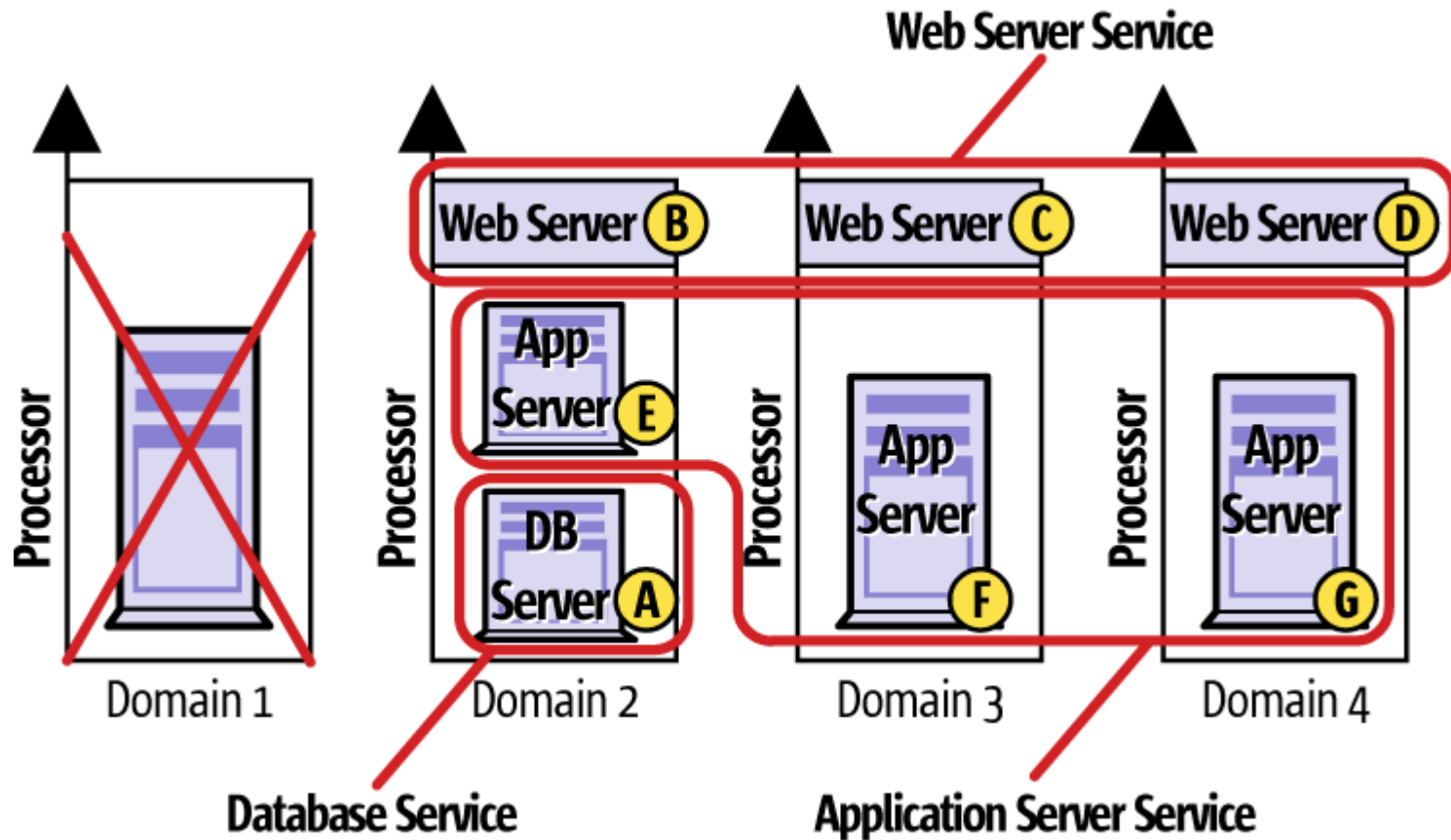
- Two *Sun Fire 6800s*
 - 4 Domains (or separate servers)
- Primary DB Server
- 3 Web Servers
- 3 App Servers



Cluster Example

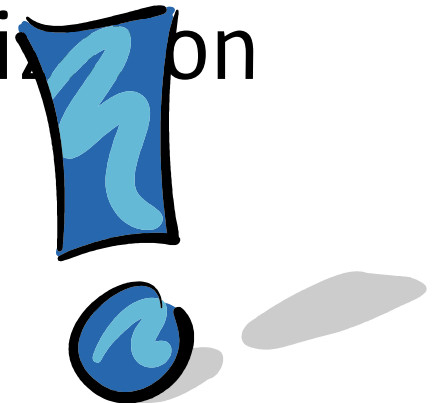


Cluster Failover



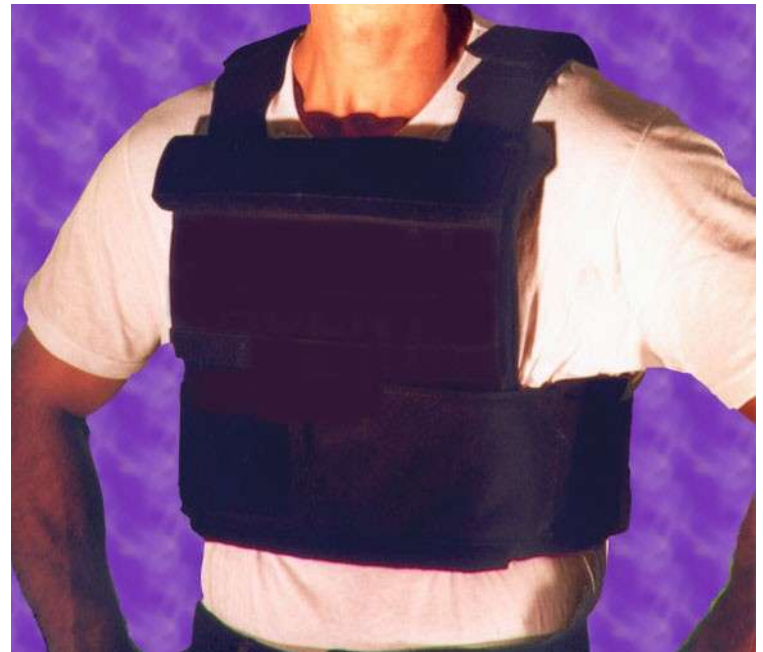
Zone Summary

- Part of Sun's *N1 Strategy*
- Isolates applications
 - providing secure containers on 1 Solaris instance
- Tool for server consolidation
- Drives efficient resource utilization
- Feature of Solaris 10
 - Beta available this Fall
 - Checked in at build 24
 - Updated at build 27 (Feb 1999)



Further Information

- Zones Project
 - Zones.eng – Design Document rev 1.3
 - Public – wp-svccont.pdf
 - Engineers
 - Andy Tucker (lead)
 - John Beck
 - Jim Carlson
 - David Comay
 - Andrew Gabriel
 - Ozgur Leonard
 - Dan Price



Zone Q&A

