

Solaris Virtualization

Ryan Matteson

matty91@gmail.com

<http://prefetch.net>

What is virtualization?

- Virtualization is a way to abstract resources (e.g., CPUs, memory, NICs, HBAs, etc.) from the consumers using them
- Software (e.g., Solaris zones, VMWare) and hardware (e.g., AMD Pacifica, Intel VT, etc.) virtualization technologies are becoming more and more prevalent

Why would I use virtualization?

- Better utilization of hardware resources
- Reduced time to market
- High availability
- Disaster recovery
- Testing new technologies
- Migrating services to new platforms

Which virtualization technologies come with Solaris / OpenSolaris?

- Solaris zones
- Linux branded zones
- Xen (dom0 and domU)
- Crossbow (virtualized network stacks)
- More to come ...

What are Solaris zones?

- Virtual Solaris execution environments that look and feel identical to a native Solaris installation
- Zones operate inside a single instance of the Solaris kernel, and run with reduced privileges (which are configurable) to improve security
- There is little to no overhead introduced by zones (the sysbench OLTP tests indicate that zone overhead is less than 2%)

Types of zones

- Zones come in two flavors:
 - Global zone - Global zones manage hardware resources and are the administrative domain for local zones
 - Local zones - Virtualized Solaris execution environments that look and feel just like a normal standalone Solaris installation
- Local zones come in two types:
 - Sparse zones - Share binaries with the global zone
 - Non-sparse zones - Contain a complete copy of the Solaris binaries that are installed in the global zone

Installing Solaris zones

- Zones are configured by running the zonecfg command
- Administrative actions (e.g., halt, reboot, install, move, etc.) are performed on zones with the zoneadm command
- Installing and configuring a zone is a four step process:
 - Create the zone configuration
 - Install the zone
 - Boot the zone
 - Console into the zone and configure it just like any other Solaris installation

Creating a Solaris zone

- To create a new zone, the `zonecfg` utility can be run with the "-z" option, the name of the zone to create, and one or more configuration directives can be used to customize the zone:

```
$ zonecfg -z zone1
zone1: No such zone configured
Use 'create' to begin configuring a new zone.
zonecfg:zone1> create
zonecfg:zone1> set autoboot=true
zonecfg:zone1> set zonename=/zones/zone1
zonecfg:zone1> add net
zonecfg:zone1:net> set physical=ni0
zonecfg:zone1:net> set address=192.168.1.17
zonecfg:zone1:net> end
zonecfg:zone1> add fs
zonecfg:zone1:fs> set dir=/usr/local
zonecfg:zone1:fs> set type=lofs
zonecfg:zone1:fs> set options=ro
zonecfg:zone1:fs> set special=/usr/sfw
zonecfg:zone1:fs> end
zonecfg:zone1> verify
zonecfg:zone1> commit
```

Installing a Solaris zone

- To install the binaries and configure the Solaris package database, the zoneadm command can be run with the install option and the name of the zone to install:

```
$ zoneadm -z zone1 install
```

- You can also use the "clone" option to create a zone from an existing zone:

```
$ zoneadm -z zone1 clone template
```

- If you are using ZFS, zone creation is almost instantaneous!:

```
$ timex zoneadm -z new clone centos
```

```
Cloning snapshot zones/centos@SUNWzone1
```

```
Instead of copying, a ZFS clone has been created for this zone.
```

```
real      0.58  
user      0.09  
sys       0.12
```

Booting a Solaris zone

- Zones can be stopped, started and restarted with the zoneadm command
- The "boot" option can be used to boot a zone:
\$ zoneadm -z zone1 boot
- The "reboot" option can be used to reboot a zone:
\$ zoneadm -z zone1 reboot
- The "halt" option can be used to stop a zone:
\$ zoneadm -z zone1 halt

Consoling into a Solaris zone

- Each zone contains a virtual console, which can be accessed with the `zlogin` command:

```
$ zlogin -C zone1
```

- You can also run commands in a local zone from inside the global zone

```
$ zlogin zone1 "ls -l /usr/bin"
```

Solaris zones demonstration

What are Linux branded zones?

- Zones capable of running **unmodified** Linux ELF executables (i.e., no recompile is required to run Linux ELF executables in a Linux branded zone)
- Linux branded zones currently support Linux ELF executables built to run on RHEL 3 and CentOS 3
- Linux branded zones are currently only available in OpenSolaris, and will hopefully make an appearance in a Solaris 10 update next year

Installing Linux branded zones

- Linux branded zones are configured identically to native Solaris zones, but use a different type (SUNWlx) to indicate that they are Linux branded
- Administrative actions (e.g., halt, reboot, install, move, etc.) are performed on zones with the zoneadm command
- Installing a zone and accessing a zone is a four step process:
 - Create the zone configuration
 - Install the zone
 - Boot the zone
 - Console into the zone and configure it just like any other RHEL or CentOS Linux installation

Creating a Linux branded zone

- To create a new Linux branded zone, the `zonecfg` utility can be run with the "-z" option, the name of the zone to create, and one or more configuration directives can be used to customize the zone:

```
$ zonecfg -z centos
centos: No such zone configured
Use 'create' to begin configuring a new zone.
zonecfg:centos> create -t SUNWlx
zonecfg:centos> set autoboot=true
zonecfg:centos> set zonepath=/zones/centos
zonecfg:centos> add net
zonecfg:centos:net> set physical=ni0
zonecfg:centos:net> set address=192.168.1.17
zonecfg:centos:net> end
zonecfg:centos> verify
zonecfg:centos> commit
```

Installing a Linux branded zone

- To install the binaries and configure the Linux RPM package database, the zoneadm command can be run with the install option and the name of the zone to install:

```
$ zoneadm -z centos install -v -d /home/matty/CentOS
```

- You can also use the "clone" option to create a zone from an existing zone:

```
$ zoneadm -z centos clone template
```

- If you are using ZFS, zone creation is almost instantaneous!:

```
$ timex zoneadm -z centosnew clone centos
```

```
Cloning snapshot zones/centos@SUNWzone1
```

```
Instead of copying, a ZFS clone has been created for this zone.
```

```
real      0.58  
user      0.09  
sys       0.12
```

Booting a Linux branded zone

- Zones can be stopped, started and restarted with the zoneadm command
- The "boot" option can be used to boot a zone:
\$ zoneadm -z centos boot
- The "reboot" option can be used to reboot a zone:
\$ zoneadm -z centos reboot
- The "halt" option can be used to stop a zone:
\$ zoneadm -z centos halt

Consoling into the branded zone

- Each Linux branded zone contains a virtual console, which can be accessed with the `zlogin` command:

```
$ zlogin -C centos
```

- You can also run commands across branded zones from inside the global zone

```
$ zlogin centos "yum upgrade"
```

Linux branded zones demonstration

Network virtualization

- Project crossbow is a project within Sun to virtualize the Solaris network stack
- Crossbow will allow multiple virtual network stacks to be created on one or more physical NICs (Network Interface Cards)
- Each virtual network stack can be assigned a priority level and bandwidth limits can be applied to each stack
- To ensure that one virtual NIC doesn't impact other virtual NICs, each virtual NIC is dedicated a set of hardware resources (e.g., RX/TX rings, DMA channels, etc.) from the underlying NIC
- The crossbow reference implementation is available as a set of BFU archives on opensolaris.org, and currently has limited hardware support (this will be fixed when the clearview project integrates)

Creating virtual NICs

- Virtual NICs can be created and destroyed with the dladm utilities “create-vnic” and “delete-vnic” options
- To create a virtual NIC, the dladm utility can be run with the “create-vnic” option, a physical device to create the VNIC on, an optional bandwidth limit or priority, and a unique identifier:

```
$ dladm create-vnic -d bge0 -m 0:1:2:3:4:5 -b 10000
```

Assigning virtual NICs to zones

- Once the virtual NIC is created*, the zonecfg utility can be used to assign the vnic to a zone or Xen instance:

```
$ zonecfg -z zone1
zonecfg:zone1: No such zone configured
zonecfg:zone1> create
zonecfg:zone1> set zonepath=/export/zone1
zonecfg:zone1> create
zonecfg:zone1> add net
zonecfg:zone1> set physical=vnic1
zonecfg:zone1> set address=192.168.1.100
zonecfg:zone1> verify
zonecfg:zone1> commit
```

* You can also configure virtual network devices inside the zonecfg net configuration stanza

Resource Management

- Resource management is used to ensure that resources are available for a given application, zone or Xen instance
- Helps to ensure that service levels are met (you wouldn't want that pesky Oracle database to steal all the resources from your Quake server, would you?)
- Solaris provides resource controls to limit the amount of CPU, memory or network bandwidth an application, zone or Xen instance can consume
- Extended accounting is available to aggregate usage data by zone, project or process, which can then be used to charge customers or business units for the computing resources they used

References

- Brendan Gregg's zone reference:
<http://brendangregg.com/zones.html>
- Crossbow documentation
<http://opensolaris.org/os/project/crossbow/>
- Linux branded zone documentation
<http://opensolaris.org/os/community/brandz>
- Resource manager documentation:
<http://www.sun.com/blueprints/0505/819-2679.pdf>
- Solaris zones documentation:
<http://opensolaris.org/os/community/zones/>
- ZFS documentation:
<http://opensolaris.org/os/community/zfs/>

Questions?