

# Solaris and Xen: Where's the Beef?

**Nathan Kroenert**

nathan@sun.com

Heavily plagiarized from **Stuart Maybee's**

FROSUG talk. Thanks, Stu. :)

## Caveats / disclaimer / get out of jail free...

- XEN is un-released, though available in SX
- What is presented here is from the public domain
- I might be wrong!
- Public Domain might be wrong
- Sun definitely makes no warranties of correctness...
- I continue to eat Steak. It's tasty.

# Overview

- Hypervisors 101
- Paravirtualization vs unmodified OS
- About Xen
- Performance
- Migration
- Solaris on Xen
- Xen and Zones

# Hypervisors 101

- Provides a “Virtual Machine” for a guest
- Not new – VM/370 over 30 yrs ago
- Controls hardware - memory/cpu/io devices
- Schedules cpus/memory/io
- May emulate real devices
- For x86/x86-64 multiple choices available:
  - > Xen, Vmware, MSFT Virtual Server
  - > others

# Para vs Full virtualization

- Full virtualization:
  - > Runs binary image of “metal” OS
  - > Must emulate real i/o devices
  - > Can be slow/need help from hardware
  - > May use trap and emulate or rewriting
- Para-virtualization:
  - > Runs OS ported to special arch
  - > Uses generic “virtual” device drivers
  - > Can be more efficient since it is hypervisor aware

# Solaris Xen Port

- “Platform” rather than “arch” port
- Wrote generic virtual device drivers
  - Net
  - Disk
  - Console
- Support Xen “event” interrupt model
- HAT layer changes for TLB/page table management
- Replace privileged instructions with hypervisor calls

# Xen and Sun - Opportunities

- “Deep-frozen” solution stacks
  - Preconfigured, suspended, ready-to-resume
- Embedded Stacks
  - Driver domains
  - Network switch
  - domU JavaOS
- Grid computing
  - “Internet Suspend/Resume”
- Massive Vertical Scale
  - Reliability

# Xen and Sun – How to Play?

- Participate in Xen community
  - Fix bugs, improve stability and portability
- Innovate in Solaris-on-Xen
  - Fault Management, Observability
  - Performance, Scaling
  - Networking, Security, Manageability
- Manage Xen-based domains
  - N1P1

# Solaris on Xen - Status

December, 2007

- Solaris Nevada build 79 on Xen 3.x
  - > 64-bit and 32 bit, domU, multiprocessor
  - > Limited Audio
  - > Network
  - > Video
  - > Some parav Video drivers (nvidia)
  - > Available now for download from OpenSolaris

# Solaris Zones and Xen

- Zones
  - Scalable, fast, virtual platform, sparc & x64
  - Emphasis on sharing, simpler admin
  - Improved fault isolation over “single system.”
  - Alternate brands
- Xen (hardware virtualization)
  - Familiar admin model
  - Emphasis on separation
  - Better performance isolation
  - Better fault isolation, SPOFs remain

# Solaris Zones and Xen (continued)

- Complementary, not alternatives
  - Use one, the other, or both
- Key differentiators for HW virtualization
  - Different OSes on same machine
  - Different patch levels on same machine
  - Suspend/Resume, and Live Migration
- Technology Lifecycle
  - Zones – nearly two years old via Solaris Express, mission-critical deployment >today<
  - Xen – still being improved, will take time to mature and be released in production

# Summary

- Xen with Solaris...
  - provides new capabilities from a new virtualization layer
  - Provides ALL the goodies Solaris has to offer to the platform
    - FMA, hardware fault isolation, IPMP, MPXIO, ZFS, DTRACE etc!
  - Will further commoditize HW virtualization
    - Everyone wants to 'control the metal'. This will be the first class alternative.

# Let's use some Xen!

- Get version of SX that has Xen integrated. Latest would be best!
- Install
- Boot non-xen first
- Reboot into Xen-boot
- Start Xen services
- Create config files
- Go the hack!

# Downloading:

- Must be based on nv\_75 or later.
- SXDE or SXCE are good choices. At present SXCE would be the best choice.

# Installing

- Nothing special here – Standard install.
- Note that lots of VM's will require lots of disk
- ZFS is ideal for managing backing stores for Xen
  - > Pooled Storage
  - > Zvols – can even be sparse volumes
  - > Snapshot, clone, promote etc
- Make sure there is adequate disk, memory and a relatively fast CPU

# First Boot

- I have encountered some interesting issues when booting into Xen for the first boot after an install.
- Might not actually be required, but I have had better results (at least for now)
- Second boot, into Xen requires we start some services

# Services:

```
# svcs -a | grep xvm
```

```
online      17:41:08 svc:/system/xvm/store:default
```

```
online      17:41:15 svc:/system/xvm/xend:default
```

```
online      17:41:16 svc:/system/xvm/console:default
```

```
online      17:41:19 svc:/system/xvm/domains:default
```

```
# svcadm enable
```

```
^^^^^^^^^^
```

# Create Config files

- See the Xen community page for examples.
- Nominally, need to describe the resources for the DomU; CPU(s), Memory, Disk devices, CD devices, Audio etc.
- .hvm files for HVM domains and .py for Para-virtualized DomU's.

# Example Para-virtualized DomU

```
beaker:/export/xen/solaris-pv-0-0 # more *  
name = "solaris-0-0"  
memory = "640"  
disk = [ 'phy:/dev/zvol/dsk/zfs/solaris-pv-0-0,0,w' ]  
vif = [ " ]  
on_shutdown = 'destroy'  
on_reboot = 'destroy'  
on_crash = 'destroy'  
boot='c'
```

# Example HVM DomU

```
beaker:/export/xen/win # more *hvm | grep -v '^#' | grep -v "^$"
```

```
import os, re
```

```
arch = os.uname()[4]
```

```
if re.search('64', arch):
```

```
    arch_libdir = 'lib64'
```

```
else:
```

```
    arch_libdir = 'lib'
```

```
kernel = "/usr/lib/xen/boot/hvmloader"
```

```
builder='hvm'
```

# HVM continued

memory = 512

shadow\_memory = 16

name = "windows-0"

vcpus=1

vif = [ 'type=ioemu' ]

disk = [ 'phy:/dev/dsk/c1t0d0s4,hdc,w',  
          'file:/export/install/winxp.iso,hda:cdrom,r' ]

on\_poweroff = 'destroy'

on\_reboot = 'restart'

on\_crash = 'preserve'

# HVM continued

```
device_model = '/usr/' + arch_libdir + '/xen/bin/qemu-dm'
```

```
boot='d'
```

```
sdl=1
```

```
vnc=0
```

```
vnclisten="0.0.0.0"
```

```
vncconsole=1
```

```
vncpasswd=""
```

```
nographic=0
```

```
stdvga=0
```

```
serial='null'
```

```
soundhw='es1370'
```

```
fullscreen=1
```

# Go the Hack!

- Your world to Xen management is xm
- xm top, create, destroy, console most common.
- xm create mydomain.hvm
- xm destroy mydomain.hvm
- xm create mypara.py
- xm console mypara
- Xm destroy mypara

# Questions?

- Except for when... ;)

# OpenSolaris on Xen Community

- Email, blogs, FAQs

- Find it at:

<http://www.opensolaris.org/os/community/xen>