

# **Solaris 10 Update 4: New Features, featuring features of the IP Instances feature. (oh - and iSCSI + .02 zfs)**

**Nathan Kroenert  
NOT Sun Microsystems :) Solaris Fanboy  
nathan@sun.com**

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

- I'm a hardware guy. Really... Don't hassle me about SW. :)
- 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 still eat Steak.

# What went into S10u4?

- Officially -
  - > <http://docs.sun.com/app/docs/doc/817-0547>
  - > Points to what's new for EACH release.
  - > Worth a read, even if you don't think you have time.

# What went into S10u4?

- Where do we start? At the cool and funky for Admins:
  - > fsstat - An iostat style utility for filesystems.
  - > nscd performance improvements, AD compatibility
  - > iostat improvements for MPxIO devices
  - > Many ZFS improvements. Speed, stability
  - > Niagara 2 - T5120 and T5220 support
  - > Improved Diagnostics for AMD Rev F CPU's (FMA)
  - > x86: Predictive Self-Healing for PCI Express on x64 Systems
  - > stmsboot - help managing x86 MPxIO boot control
  - > SATA improvements for Marvel88sx driver. NCQ / concurrent read/write queued operations

# What went into S10u4?

- > NFSv4 Domain ID can be specified in sysid (yay!)
- > Solaris Live Upgrade enhancements (though zfs still tricky)
- > Upgrade / LU possible even with zones present
- > Deferred activation patching (feature? :)
- > Greatly improved IPSEC compatibility
- > Routing improvements, routeadm and quagga
- > DHCPv6 client
- > Single hosts file
- > LSO hardware large send offload capability
- > Solaris Key Management Framework - ease manageability
- > Solaris Data Encryption Supplement (**HOORAY!**)

# continued...

- > Support for iSCSI Target Devices
  - > You can now do both target and initiator.
- > Brandz support for Linux apps
- > Zone-wide resource controls improved
- > IP Instances: LAN and VLAN Separation for Non-Global Zones
- > Solaris Zones Boot Enhancements
- > System V Resource Controls for Zones
- > DTrace in a Non-Global Zones
- > For x86 users - Xvideo support in realplayer

## continued...

- > SunVTS - hardware test suite
- > Multi-level CMT Scheduling Optimizations
- > Process Count Scalability
- > MPSS Extended to Shared Memory

# Oh GAWD - 10pm here we come...

- > Fear not! We'll only be covering two of the big ticket features...
- > We'll focus on iSCSI and IP Instances
- > Also - Just a few 'cool' things here and there

# Not from U4, but cool - fsstat

- iostat style utility for filesystems
- Shows # of file creations, removal, attribute changes, lookups, readdirs, reads, writes etc.
- Very quickly zero in on certain meta-activity that might be impacting the performance of the disk
- Less manipulation required than iostat

# fsstat example output

```

huron:/root # fsstat /tmp / 5
new name name attr attr lookup rddir read read write write
file remov chng get set ops ops ops bytes ops bytes
3.21K 3.20K 3 3.47K 11 3.53K 32 20 33.3K 59 50.6K /tmp
461 107 11 113K 263 476K 1.51K 86.3K 209M 21.8K 106M /
0 0 0 0 0 0 0 0 0 0 0 /tmp
0 0 0 58 0 400 4 20 1.38K 28 2.96K /
489 0 0 491 0 2 0 0 0 0 0 /tmp
1 0 0 8.17K 0 24.0K 0 2.94K 235K 27.3K 218M /
847 0 0 867 0 20 0 0 0 0 0 /tmp
0 0 0 14.5K 0 42.7K 4 5.12K 408K 30.8K 246M /
824 0 0 865 0 41 0 0 0 0 0 /tmp
0 0 0 14.4K 0 42.3K 0 5.14K 408K 32.1K 256M /
798 0 0 866 0 68 0 0 0 0 0 /tmp
0 0 0 14.4K 0 42.7K 4 5.16K 408K 31.3K 249M /
766 0 0 862 0 96 0 0 0 0 0 /tmp
0 0 0 14.3K 0 42.1K 0 5.15K 406K 31.4K 250M /

```

# Separate IP instances per Zone

- Hoopy Froods get excited. It's here...
- Ability to assign a zone a complete ethernet device or VLAN.
- Yes - You can make a zone a DHCP client.
- VNICS not yet available. (boo! :)
- <http://www.opensolaris.org/os/project/crossbow/snap>

# Separate IP instances per Zone

- <http://www.opensolaris.org/os/project/crossbow/>
- <http://www.opensolaris.org/os/project/crossbow/faq/>
- Requires a GLDv3 driver based cards
- Use `dladm show-link` to determine if you have GLDv3 devices. (anything not 'Legacy' is GLDv3.)
- E1000g is a cheap, GLDv3 alternative with SPARC drivers, BGE is a cheap Vlanable one.

# Separate IP instances per Zone

- What services can now run in a zone that could not before?
  - > DHCP client
  - > DHCP server
  - > Routing daemon
  - > IPsec
  - > IPfilter
  - > IP Multipathing (IPMP)
  - > ndd commands
  - > ifconfig (with set or modify capabilities)

# Separate IP instances per Zone

- Still CANNOT run an NFS server from within a non-global zone. See 4964859
- Get in on the bug if this annoys you.
- If it would cost deals, where you would go with someone else, definitely let your rep's know.
- Will not be fixed without a really good business case.
- iscsi not working in zones even with separate IP
- Vnics a little ways off...

# What does it look like?

- Just like a standard interface, within a zone
- At some point we need to set ip-type=exclusive in the zonecfg
- Interface then behaves just like a bought one!
- Find all interfaces available - ifconfig -a plumb

# Separate IP instances per Zone

```
zonecfg:zone1> info
  zonename: zone1

  zonepath: /zfs/zones/zone1
  brand: native
  autoboot: false
  bootargs:
  pool:
  limitpriv:
  scheduling-class:
  ip-type: exclusive
  inherit-pkg-dir:
    dir: /lib
  inherit-pkg-dir:
    dir: /platform
  inherit-pkg-dir:
    dir: /sbin
  inherit-pkg-dir:
    dir: /usr

  net:
    address not specified
    physical: e1000g0
```

# Separate IP instances per Zone

```
# ifconfig -a

lo0: flags=2001000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4,VIRTUAL> mtu
8232 index 1

    inet 127.0.0.1 netmask ff000000

e1000g0: flags=201004843<UP,BROADCAST,RUNNING,MULTICAST,DHCP,IPv4,CoS>
mtu 1500 index 2

    inet 192.168.1.227 netmask ffffffff00 broadcast 192.168.1.255

    ether 0:e:c:6b:fa:52

lo0: flags=2002000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv6,VIRTUAL> mtu
8252 index 1

    inet6 ::1/128

# ping 192.168.1.1

192.168.1.1 is alive
```

# Separate IP instances per Zone

```
beaker:/root # zoneadm list -cv
```

ID	NAME	STATUS	PATH	BRAND	IP
0	global	running	/	native	shared
4	zone1	running	/zfs/zones/zone1	native	excl

# iSCSI - Target + Initiator

- Provide iSCSI targets to all sorts of OS's
- Solaris ZFS backed iSCSI is gold.
- Windows / Linux / Solaris iSCSI clients
- For systems, such as the Niagara 2 based boxes with 10GbE, high throughput access to iSCSI targets.
- Reduced need for massive internal storage
- Reduced dependency on SAN / FC switches etc
- Cheaper than the 'real thing'.

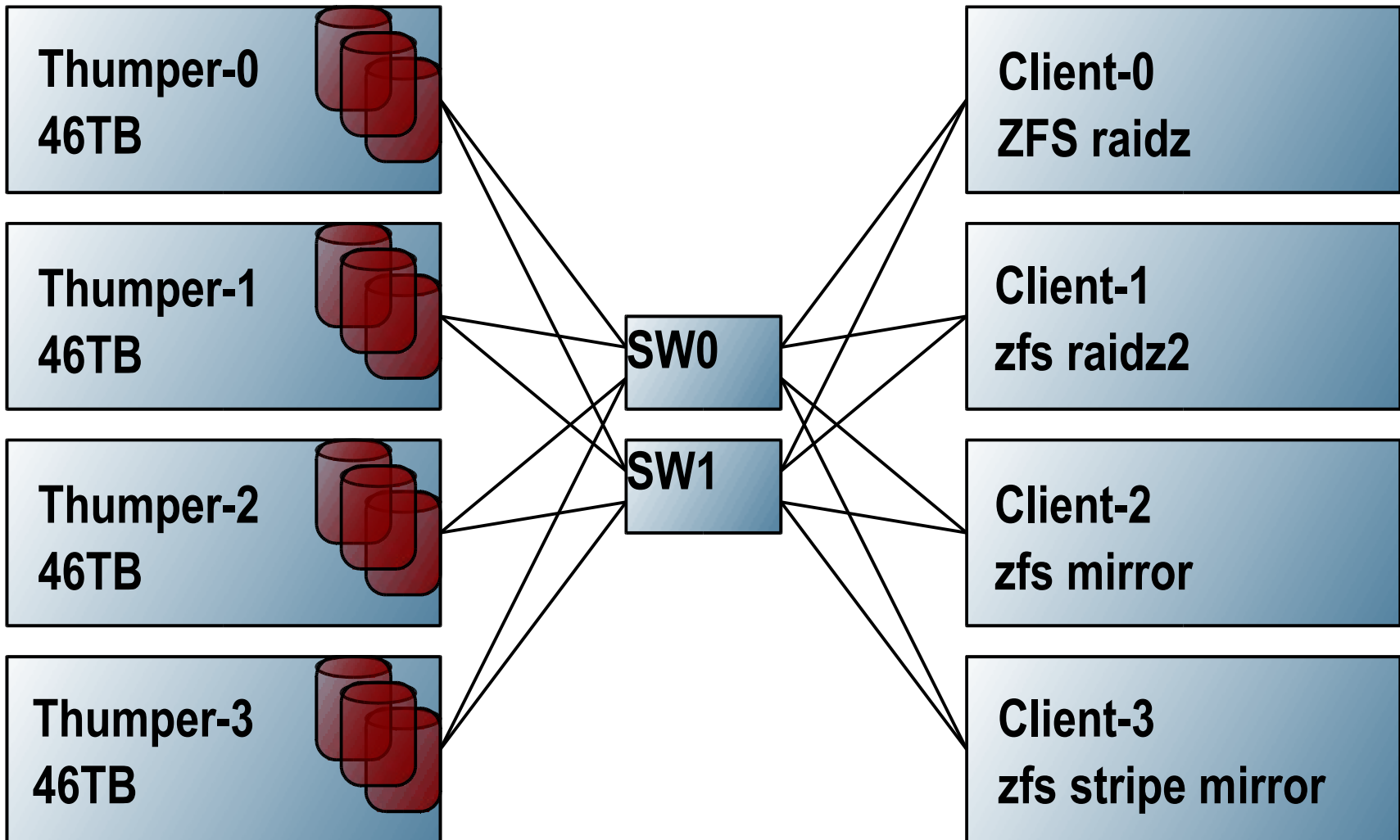
# iSCSI - For the geeky

- Redundant storage over either trunked or multi-homed server + client
- Redundant storage, mirrored / striped over **servers**
- ZFS backed, any FS you like on the iSCSI volume
- All your storage are belong to solaris
- Consider Thumper... Presenting iSCSI targets

# Caveats:

- Security security security...
- If using non-redundant storage, target consumers can pause waiting for I/O or LUN to time out
- If you make a zpool out of non-redundant storage, on the lan, a network interruption or target server reboot will most likely panic your system.  
(ZFS updates coming that will help to resolve some of this, with an onerror= style properly.)
- Has been in use internally and in SXDE for a while, but still relatively young.

# Possible config:



# ISCSI - Target (The Server!)

```
$ zfs create -s -V 200mb zfs/iSCSI_target_0
$ zfs create -s -V 200mb zfs/iSCSI_target_1
$ zfs create -s -V 200mb zfs/iSCSI_target_2
$ zfs create -s -V 200mb zfs/iSCSI_target_3
$ zfs create -s -V 200mb zfs/iSCSI_target_4
$ zfs set shareiscsi=on zfs/iSCSI_target_0
$ zfs set shareiscsi=on zfs/iSCSI_target_1
$ zfs set shareiscsi=on zfs/iSCSI_target_2
$ zfs set shareiscsi=on zfs/iSCSI_target_3
$ zfs set shareiscsi=on zfs/iSCSI_target_4
```

# ISCSI - Initiator (The Client!)

```
$ iscsiadm modify discovery --sendtargets enable
```

```
$ iscsiadm add discovery-address 192.168.1.15
```

```
$ svcadm enable iscsi-initiator
```

```
$ format (find out what disks are there)
```

```
$ zpool create tank cXtCdX
```

done!

# ISCSI - Initiator teardown

```
$ zpool destroy tank  
$ iscsiadm remove discovery-address 192.168.1.15  
$ iscsiadm modify discovery --sendtargets disable  
$ svcadm disable iscsi-initiator  
$ format (check that disks are gone)
```

done!

# References

- Zones + Dtrace: <http://docs.sun.com/app/docs/doc/819-2450/gelyb?a=view>
- Benr's blog on zfs+iSCSI - <http://www.cuddletech.com/blog/pivot/entry.php?id=775>
- S10U4 update notes - <http://docs.sun.com/app/docs/doc/817-0547>
- Crossbow opensolaris site:<http://www.opensolaris.org/os/project/crossbow/>

**And He said unto them:  
Let there be steak.**