SOLARIS @STEP

DOAG 2014

Solaris 11.2

Neues aus der Zone

Martin Muschkiet, as-systeme

Solaris 11.2 Neue Zonen Features

- Kernel Zonen
- Immutable globale Zonen
- Zonen provisionieren und klonen mit Unified Archives
- "Live" Modifikation von Zonen Ressourcen
- Multiple Boot Environments für *solaris10* Zones
- Zonen Ressourcen für CMT-Systeme (T-Serie)
- IPS Pakete in der globalen Zone und in mehreren nicht globalen Zonen installieren

In diesem Vortrag

- Kernel Zonen
- Immutable globale Zonen
- Zonen provisionieren und klonen mit Unified Archives
- "Live" Modifikation von Zonen Ressourcen
- Multiple Boot Environments für *solaris10* Zones
- Zonen Ressourcen für CMT-Systeme (T-Serie)
- IPS Pakete in der globalen Zone und in mehreren nicht globalen Zonen installieren

Recap Was ist eine Zone?

- OS Virtualisierung ab Solaris 10
- Urspr. Sicherheitsfeature "*Project Kevlar*"
- Leichtgewichtig ≠ Hypervisor Lösungen
- technisch umgesetzt durch
  - :: Prozessgruppen, jeder Prozess hat eine *zoneid*
  - :: Chroot Umgebung
  - :: beschränkte Privilegien
- Alles läuft unter einem Kernel
  - :: Erweiterung durch Branded Zones

Zonen Typen

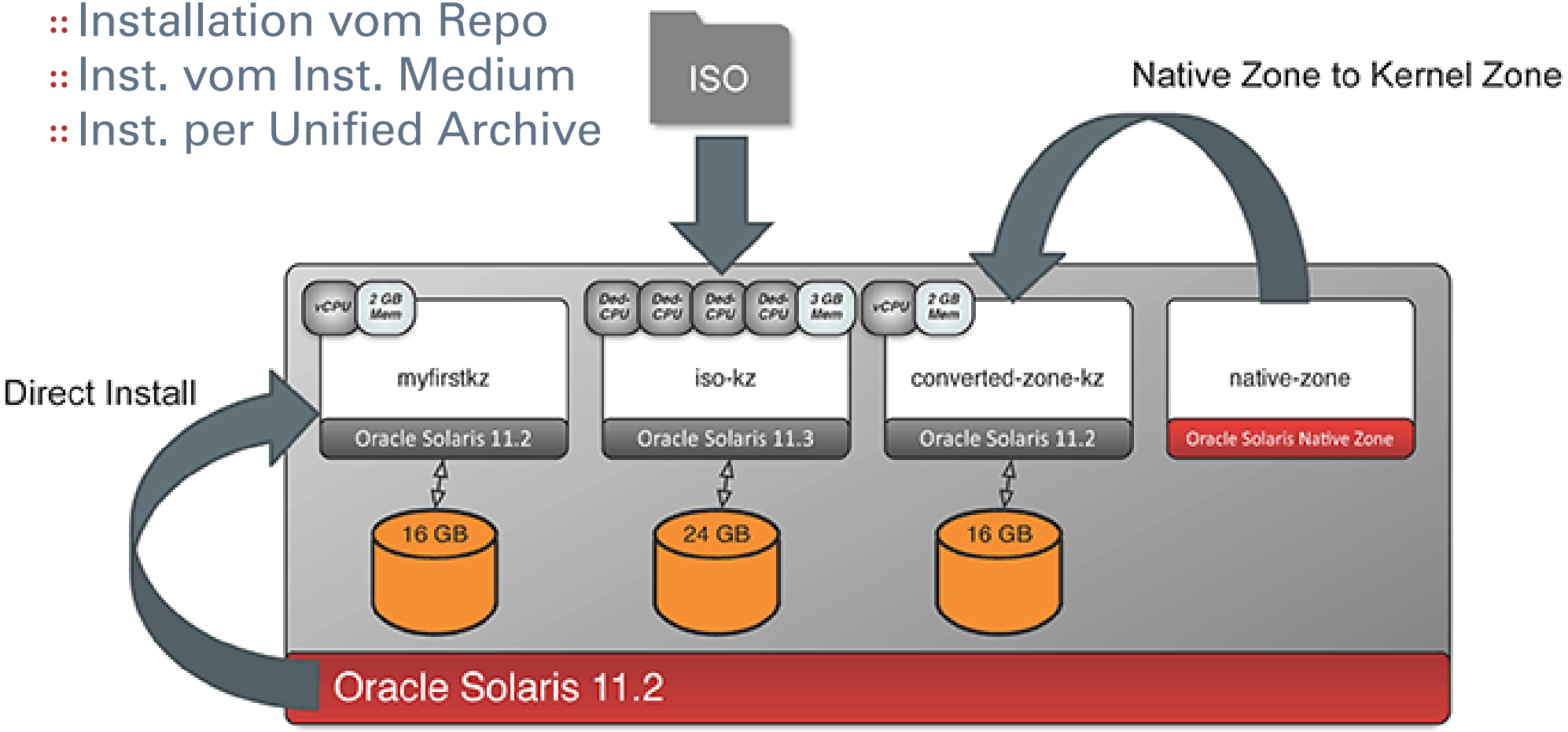
- Globale Zone (GZ)
- Nicht-globale Zone (NGZ)
  3 Brands
  - :: solaris
  - :: solaris10
  - :: solaris-kz (KeZ)

Warum Kernelzonen?

- Kernel Zonen ermöglichen NGZ mit unterschiedlichen Kernelversionen auf einem physischen System laufen zu lassen.
- NGZs mit speziellen Paket-Bedürfnissen können leichter realisiert werden
- Kernelzonen lassen sich individuell (früher/später) aktualisieren
- Kernelzone als Testumgebung für neue SRUs oder Releases
- Bei einer Konsolidierung kann ein *Unified Archive* auf ein Ziel mit einer anderen OS-Version migriert werden
  :: ohne Zwangsupdate
- Live Migration von Kernel Zonen (11.3 / 11.4)
  :: derzeit suspend/resume

3 Wege zur Kernelzone

- Installation vom Repo
- Inst. vom Inst. Medium
- Inst. per Unified Archive



Vorausetzungen

- Sparc
  - :: T4 mit FW min. 8.5.1,
  - :: T5, M5, M6 mit FW min. 9.2.1
- X86
  - :: Nehalem+(Intel) oder Barcelona+(AMD) basiertes X86 System mit aktivierter CPU Virtualization (VT-x)
  - :: läuft NICHT in einer Virtualisierungslösung wie z.B. VirtualBox
- 8 Gb RAM
  - :: ggf. muss ARC begrenzt werden
- pkg: *brand-solaris-kz*

```
# virtinfo -c supported list kernel-zone
NAME            CLASS
kernel-zone     supported
```

Kernel Zonen: Dezidierte Ressourcen

- RAM
  - kein Capping wie in regulären NGZ, sondern ein festes Quantum
  - derzeit keine dynamische Rekonfiguration
  - Std. 2 Gb
- Festplatte
  - Std. 16 Gb *zvol* für den rpool der KeZ
  - Optional: physische Platte oder Storage URIs
- CPU
  - Std: 1 virtuelle CPU
  - *virtual-cpu* (gemeinsam verwendet mit anderen Zonen)
  - *dedicated-cpu* (dezidiert), legt beim Start der KeZ einen temporären cpu-pool an

Kernel Zone konfigurieren

```
# zonecfg -z kzone
Use 'create' to begin configuring a new zone.
zonecfg:kzone> create -t SYSsolaris-kz
zonecfg:kzone> select device id=0
zonecfg:kzone:device> set storage=dev:/dev/zvol/dsk/pot/zones/kzone/disk0
zonecfg:kzone:device> end
zonecfg:kzone> add virtual-cpu
zonecfg:kzone:virtual-cpu> set ncpus=2
zonecfg:kzone:virtual-cpu> end
zonecfg:kzone> info
zonename: kzone
brand: solaris-kz
...
```

Kernel Zone installieren

```
# zoneadm -z kzone install -c /root/sc-kzone.xml

 AI Manifest: /tmp/zoneadm2966.ccayld/devel-ai-manifest.xml
  SC Profile: /root/sc-kzone.xml
Installation: Starting ...
              ...
DOWNLOAD                    PKGS       FILES    XFER (MB)   SPEED
Completed                 579/579   77133/77133  640.1/640.1    0B/s

PHASE                           ITEMS
Installing new actions            104101/104101
Updating package state database             Done
Updating package cache                   0/0
Updating image state                   Done
Creating fast lookup database              Done
Installation: Succeeded
        Done: Installation completed in 466.327 seconds.
```

Alternative: iso-Datei statt Repository

:: Mit 11.2-Text Iso-Image
:: Größe des *zvol* auf 24 Gb festgelegt

```
 # zoneadm -z proton install  -b /root/sol-11_2-text-x86.iso  -x install-size=24g
Progress being logged to /var/log/zones/zoneadm.20141118T174821Z.proton.install
[Connected to zone 'proton' console]
Boot device: cdrom1  File and args:

...normale interaktive Installation

[NOTICE: Zone rebooting]

[Connection to zone 'proton' console closed]
      Done: Installation completed in 1043.208 seconds.
/root #
```

Kernelzone Observability 1

- Der Kernel einer Solaris KeZ wird durch einen User Land Prozess in einer Solaris NGZ dargestellt
- In der GZ sind nur 2 Prozesse der KeZ sichtbar

```
# zoneadm list -cv
  ID  NAME      STATUS     PATH       BRAND     IP
   0  global    running    /          solaris   shared
   -  kzone     installed  -          solaris-kz excl
# zoneadm -z kzone boot
# ps -fZz kzone
   ZONE     UID   PID    PPID   C  STIME    TTY  TIME   CMD
   kzone    root  2458   1      0  12:10:57 ?    0:00     zsched
   kzone    root  2636   2458   0  12:10:58 ?    1:37     /usr/lib/kzhost
```

Kernelzone Observability 2

- Datalink und Volume Information, volle Privilegien
  - :: Anzahl der effektiven Privilegien ist gleich zwischen GZ und KeZ

```
# dladm show-link -z kzone
   LINK              CLASS     MTU    STATE      OVER
   kzone/net0        vnic      1500   up         e1000g1
# zfs list -o name,volsize,refreserv,used,avail,refer pot/zones/kzone/disk0
   NAME                      VOLSIZE  REFRESERV  USED   AVAIL  REFER
   pot/zones/kzone/disk0     16G      16.5G      16.5G  861G   1.44G
# ppriv -v $$ | grep E | tr , '\n' | wc -l
     84
# zlogin kzone
[Connected to zone 'kzone' pts/2]
Oracle Corporation      SunOS 5.11      11.2    April 2014
root@kzone:~# ppriv -v $$ | grep E | tr , '\n' | wc -l
     84
```

Tour de Kernelzone

- Alle Ressourcen wie konfiguriert
  - :: eigenständige IPS Verwaltung, k. System Publisher wie in regulären NGZ

```
 # zlogin kzone
[Connected to zone 'kzone' pts/2]
Oracle Corporation      SunOS 5.11      11.2    April 2014
root@kzone:~# psrinfo -p
2
root@kzone:~# prtconf | grep Mem
Memory size: 2048 Megabytes
root@kzone:~# zpool list
NAME  SIZE    ALLOC  FREE   CAP  DEDUP    HEALTH   ALTROOT
rpool  15.6G   4.46G  11.2G  28%  1.00x    ONLINE   -
root@kzone:~# pkg publisher
PUBLISHER      TYPE    STATUS    P  LOCATION
solaris        origin  online    F  http://nuklear.pool.bar:8888/
```

Suspend/Resume für Kernelzonen

- Statt die KeZ mit *zoneadm -z <zone> shutdown* runter zu fahren, kann der aktuelle Zustand mit *suspend* gesichert werden
- Liegen Dateisysteme und *Suspend* Abbild der KeZ auf shared Storage lassen sich KeZ mit minimaler downtime warm migrieren.

```
# zonecfg -z kzone info suspend
suspend:
    path: /system/zones/kzone/suspend
    storage not specified
# zoneadm -z kzone suspend
```

Suspend/Resume für Kernelzonen

- *zoneadm list -s|-p* zeigen den suspend Zustand an
- Das *Suspend* Abbild wird mit AES-128-CCM verschlüsselt. Schlüssel unter */etc/zones/keys*
- *zoneadm -z <zone> boot -R* verwirft das *suspend* Abbild

```
# zoneadm list -is
NAME        STATUS          AUXILIARY STATE
global      running
neutrino    running
kzone       installed       suspended
proton      running
# ls -lh /system/zones/kzone
-rw-------  1 root     root         350M      Jul 22 19:06   suspend
# zoneadm -z kzone boot
```

Zone in der Kernel-Zone

- Kernelzonen können weitere reguläre NGZ beherbergen
  - alle MAC-Adressen, die diese NGZ der KeZ verwenden, müssen in der GZ vorab festgelegt werden
  - Nach dem Neustart der KeZ stehen die MAC-Adr. zur Verfügung

```
root@global:~# zoneadm list -cv
  ID   NAME         STATUS       PATH                BRAND      IP
   0   global       running      /                   solaris    shared
   9   proton       running      -                   solaris-kz excl
root@global:~# zonecfg -z proton
zonecfg:proton> select anet id=0
zonecfg:proton:anet> add mac
zonecfg:proton:anet:mac> set id=1
zonecfg:proton:anet:mac> end
zonecfg:proton:anet> end
zonecfg:proton> exit
```

Zone in der Kernel-Zone

- durch *default* Einträge im *Zonentemplate* Konfiguration der NGZ als Einzeiler.
- Die *anet*-Ressource nutzt eine der zuvor festgelegten MAC-Adr.
- die KeZ fungiert dann als globale Zone

```
root@proton:~# virtinfo
NAME              CLASS
kernel-zone       current
non-global-zone supported
root@proton:~# zonecfg -z higgsboson create
root@proton:~# zoneadm -z higgsboson install
The following ZFS file system(s) have been created:
   rpool/VARSHARE/zones/higgsboson
...
root@proton:~# zoneadm -z higgsboson boot
root@proton:~# zlogin -C higgsboson
```

Und die Nachteile?

- erhöhter RAM Anforderungen
  - :: Bedarf des separaten Kernels
    NGZ im Leerlauf ca. 85Mb, KeZ phys. used ca. 450 Mb
  - :: feste Vergabe von RAM an die KeZ
  - :: bei Nutzung von *zvols* weniger effektives, doppeltes Caching
    ZFS-ARC in der GZ und der KeZ
- höherer Plattenplatzbedarf
  - :: Kez mit *solaris-large-server* incl. dump und swap 4,8Gb
  - :: NGZ mit *solaris-large-server* 850 Mb
- Administrativer Aufwand
  - :: Überwachung ausgehend von der globalen Zone eingeschränkt
  - :: KeZ muss separat aktualisiert werden

Vergleich mit Oracle VM für Sparc aka LDoms

Kernel Zone

- KeZ sind abhängig von der globalen Zone. Läuft diese nicht, kann auch die KeZ nicht laufen
- OS der KeZ ≥ Solaris 11.2

OVM für Sparc

- Fährt man die *Control domain* runter, können *Guest Domains* weiter laufen, solange I/O (per *io-domain*, *root-domain*)verfügbar ist.
- Live Migration verfügbar (nur reine *guest domains*)
- OBP in allen Domains
- dynamische Memory Rekonfiguration verfügbar
- Solaris 10, Solaris 11, Linux guests?

Immutable unveränderliche Zonen

- Rückschau: Solaris 10 sparse Zone
  - :: wichtige Systemverzeichnisse der NGZ sind durch *lofs readonly* mount vor Veränderung geschützt
- Solaris 11 immutable Zones
  - :: Konfiguration per *file-mac-profile* Eigenschaft mit *zonecfg*

```
# zonecfg -z neutrino set file-mac-profile=fixed-configuration
# zoneadm  -z neutrino reboot
# zoneadm  -z neutrino list -p
5:neutrino:running:/zones/neutrino:<uuid>:solaris:excl:R:fixed-configuration
# zlogin neutrino
[Connected to zone 'neutrino' pts/2]
Oracle Corporation     SunOS 5.11    11.2  April 2014
root@neutrino:~# touch /usr/xray
touch: cannot create /usr/xray: Read-only file system
```

Immutable unveränderliche Zonen

- Solaris 11 immutable Zones
  - :: Für Veränderungen kann die NGZ im Write-Modus gebootet werden
    # zoneadm -z <zone> boot -w
  - :: Solaris 11.2 führt *zlogin -T <zone>* ein; durch das Anmelden am *trusted path* werden Änderungen an sonst geschützten Bereichen möglich
  - :: für *flexible-configuration* Zonen ist *zlogin -U <zone>* nötig

```
# zlogin -T neutrino
[Connected to zone 'neutrino' pts/2]
Oracle Corporation      SunOS 5.11      11.2    April 2014
root@neutrino:~# touch /usr/xray
root@neutrino:~#
```

Immutable Zonen Konfiguration

- 4 Werte für *file-mac-profile*. Was ist schreibbar/möglich (W)
  - :: *strict*
  - :: *fixed-configuration*
  - :: *flexible-configuration*
  - :: *none* =Standard RW Zone

Verzeichnis / Aktion	strict	fixed-conf.	flexible-conf.	none
/usr	-	-	-	W
/etc	-	-	W	W
Teilbereiche /var, z.B. für Logs	-	W	W	W
gesamtes /var	-	-	-	W
/tmp	W	W	W	W
/root	-	-	W	W
$HOME normale User im rpool	-	-	-	W
per add dataset hinzugefügtes zfs	W	W	W	W
zfs create im rpool	-	-	-	W
svcadm enable|disable <fmri>	n. persistent	n. persistent	W	W

Globale Immutable Zonen

- Verfügbar ab Solaris 11.2
- Konfiguration analog der NGZ
- Für die GZ schreibbare Datasets können per *add-dataset* konfiguriert werden (wird vererbt)
- Zum Verändern der Immutable-GZ Login an der Console per *break*= Trusted Path

```
# zonecfg -z global
    zonecfg:global> set file-mac-profile=flexible-configuration
    zonecfg:global> add dataset
    zonecfg:global:dataset> set name=rpool/export
    zonecfg:global:dataset> end
    zonecfg:global> exit
updating /platform/sun4v/boot_archive
#  init 6
```

Links, Quellen

- The Zones Zone. Mike Gerdts Blog
  https://blogs.oracle.com/zoneszone/
- How to Get Started Creating Solaris Kernel Zones. Duncan Hardie
  http://www.oracle.com/technetwork/articles/servers-storage-admin/howto-create-kernal-zones-s11-2251331.html
- Immutable global Zones
  https://blogs.oracle.com/casper/entry/solaris_11_2_immutable_global
- Overview of Solaris Zones Security Models
  https://blogs.oracle.com/darren/entry/overview_of_solaris_kernel_zones
- man solaris-kz
  https://docs.oracle.com/cd/E36784_01/html/E36883/solaris-kz-5.html
- Memory Management Betw. ZFS and Applic. in Sol. 11.2 (Doc ID 1663862.1)
  https://support.oracle.com/epmos/faces/DocumentDisplay?id=1663862.1&_adf.ctrl-state=w3elcljm1_9&_afrLoop=245361974776441

SOLARIS @STEP

Danke

Martin Muschkiet, as-systeme