Der Aufbau von Cloudumgebungen mit Oracle Solaris 11

Detlef Drewanz

Oracle Deutschland B.V. & Co. KG

Potsdam

Schlüsselworte:
Oracle Solaris 11, Solaris Zonen, iSCSI, ZFS

Einleitung

Dieser Vortrag erläutert die Möglichkeiten von Oracle Solaris 11 zum Aufbau von Laufzeitumgebungen, für die im Umfeld von Clouds breite Einsatzmöglichkeiten bestehen. So zum Beispiel für den Betrieb von Oracle Datenbanken, Application Servern oder Oracle Applications.
Den Hauptbestandteil bilden Oracle Solaris Zones als sichere und performante Virtualisierungstechnologie. Für eine günstige Datenablage wird iSCSI verwendet. Die Konfiguration und Verwaltung von Solaris Zonen auf iSCSI wird seit Solaris 11.1 von dem Feature "Zones on Shared Storage(ZOSS)" unterstützt und stark vereinfacht. Die Kombination von Solaris Zonen mit ZOSS und verschlüsselten ZFS-Datasets auf Standard iSCSI führt zu einer performanten und sicheren Laufzeitumgebung. In diesem Vortrag werden die einzelnen Schritte zum Aufbau der Umgebung vorgestellt und diskutiert.

Cloud-Computing

Der Begriff des Cloud-Computing wird vielfach für Verfahren und Ansätze benutzt, in denen IT-Infrastrukturen über ein Netzwerk Konsumenten zur Verfügung gestellt wird. Diese Infrastrukturen können Speicherkapazitäten oder auch Speicherdienste, Rechenkapazitäten, Laufzeitplattformen oder komplette Dienste sein, wobei die bereitgestellten Kapazitäten an den Bedarf dynamisch angepasst werden können. Da die Struktur und Architektur der bereitgestellten Infrastruktur für den Konsumenten nicht sichtbar ist oder abstrahiert wird, erscheint sie für den Benutzer undurchsichtig oder wie aus einer "Wolke".

Je nachdem, wo und wie die Infrastrukturen betrieben werden und wie die Leistung durch den Konsumenten abgenommen wird, ergeben sich unterschiedliche Liefermodelle für Clouds:

- Public Cloud: Für die Öffentlichkeit über das Internet
- Private Cloud: Innerhalb einer Organisation für eine geschlossene Nutzergruppe
- Hybrid Cloud: Kombination aus Public und Private Cloud
- Community Cloud: Für eine geschlossene Nutzergruppe im Internet, die die Kosten teilt

Zusätzlich ergeben sich unterschiedliche Servicemodelle für Clouds. Diese richten sich danach, welche Art von Service bereitgestellt wird. Daraus ergeben sich Konsequenzen für die Art und den Umfang der verwendeten Architektur:

- Infrastructure as a Service (IaaS): Bereitstellung von OS-Laufzeitumgebung für Software oder Bereitstellung von Speicher
- Platform as a Service (PaaS): Bereitstellung von Softwareumgebungen zur Ausführung von optimierter Anwendersoftware
- Software as a Service (SaaS): Bereitstellung von Diensten

Zum Aufbau von Iaas-Clouds, werden im Weiteren die folgenden Technologien betrachtet:

- Virtualisierung zur Bereitstellung von Laufzeitumgebungen
- Vernetzung zum Zugriff auf die Laufzeitumgebungen
- Storage für die Laufzeitumgebungen
- Storage zur Nutzdatenspeicherung

Gerade in Cloudumgebungen sind für alle genannten Punkte noch die folgenden Eigenschaften wichtig, die die eingesetzen Implementierungsformen der Technologien auszeichnen müssen:

- Flexibilität, um schnell auf veränderte Anforderungen reagieren zu können
- Skalierbarkeit, um auch zukünftige wachsende Anforderungen bedienen zu können
- Stabilität und Verfügbarkeit, um die Akzeptanz der Cloud bei den Nutzern zu sichern
- Security, zum Schutz der Laufzeitumgebung und der Daten vor dem Zugriff Dritter

Die Zielsetzung für dieses Papier ist zu zeigen, wie IaaS-Cloudumgebungen mit Oracle Solaris 11 und Solaris Zonen unter Nutzung dem Oracle ZFS Storage aufgebaut werden können und wie solche Umgebungen die Basis für PaaS-Clouds bilden können.

Virtualisierung mit Oracle Solaris 11

Oracle Solaris Zonen sind die Standardmethode, um Anwendungen in einer Solaris Umgebung zu virtualisieren und zu konsolidieren. Die Möglichkeit, auf einfache Weise Ressourcen zu teilen, zu kontrollieren und administrative Rechte zu delegieren und damit sichere und gekapselte Ablaufumgebungen zu implementieren, haben zu einer breiten Akzeptanz bei Solaris Nutzern geführt.

Zonen werden in nicht-globale Zonen und die globale Zone unterschieden. Nicht-globale Zonen sind voneinander isoliert, teilen sich jedoch gemeinsam die globale Zone. Diese enthält neben dem Betriebssystem-Kernel die Device-Treiber, die Devices, das Memory Management System, die Filesystem-Treiber und in vielen Fällen auch den Netzwerk-Stack. Zur Begrenzung des Ressourcenverbrauches der einzelnen Zonen, können Funktionalitäten des Ressourcenmanagements (RM) genutzt werden. Lediglich die Ressourcen, die durch die globale Zone für eine nicht-globale Zone bereit gestellt werden, können durch diese genutzt werden.

- Die globale Zone "sieht" alle physikalischen Ressourcen und stellt sie den Prozessen in den nicht-globalen Zonen zur Verfügung.
- Je Solaris System können bis zu 8192 nicht-globale Zonen erzeugt werden. Die tatsächliche Anzahl nutzbarer Zonen ist von den spezifischen Leistungsanforderungen der Zonen und der benutzten Hardware abhängig.
- Eine nicht-globale Zone erscheint dem Kernel wie eine Gruppe von Prozessen mit besonderen Eigenschaften. Diese verfügen über einen gemeinsamen Identifier (zoneid) und zeichnen sich durch Security Credentials aus.
- Einer Anwendung oder einem Benutzer erscheint eine Solaris Zone wie eine eigene Oracle Solaris Installation.
- Durch die Bereitstellung eigener root-Verzeichnisse für jede Zone, können durch die lokalen Namensumgebungen in den Zonen eigene Festlegungen zu den Sicherheitseinstellungen getroffen werden (RBAC - Role Based Access Control, passwd-Datenbank, ...) und eigene Nutzerdatenbanken bereitgestellt werden.
- In jeder nicht-globalen Zone kann sich ein anderer Administrator mit seinem root-Passwort anmelden. Andere Zonen oder ihre genutzten Ressourcen sind aus einer nicht-globalen Zone nicht "sichtbar".
- Die globale Zone hat eine transparente Sicht auf alle Zonen.
- Die nicht-globalen Zonen und die globale Zone benutzen für eine begrenzte Anzahl von Solaris-Paketen einen identischen Release-Stand. Eine Ausnahme bilden hierbei die Solaris 10 Branded Zones.

Auf Grund der Einfachheit bei der Installation und dem Betrieb von Zonen, eignen sie sich hervorragend zur schnellen und effizienten Bereitstellung von Laufzeitumgebungen.

- Zonen können sehr schnell erzeugt werden - z.B. aus Templates oder durch Cloning vorhandener Zonen.
- Zonen können sehr schnell gestartet, gestoppt und restartet werden, da die Initialisierung von Hardware, Kernel, Treibern und Filesystemen entfällt.
- Zonen sind durch die Veränderung von Ressource Controls skalierbar und können in den Ressourcegrenzen der globalen Zone vergößert oder verkleinert werden.

- Zonen beziehen ihre Stabilität und Verfügbarkeit aus der globalen Zone durch die Nutzung der Fault Management Architektur (FMA) und des Service Management Facility (SMF).
- Die Sicherheit von Solaris Zonen wird durch Privilegien-Sets realisiert. Diese legen fest, über welche Privilegien der Administrator einer Zone verfügt, welche Aktionen er ausführen kann oder ob auf bestimmte Funktionen von Geräten oder des Betriebssystems zugegriffen werden darf oder nicht. Durch Securitykonzepte wird ebenfalls die Sichtbarkeit von Geräten in der Zone gesteuert.

Zur Konfiguration einer Zone sind die folgenden Schritte notwendig:

- Festlegung eines Zonenamens
- Festlegung des Root-Filesystems einer Zone
- Zuweisung des benutzen Netzwerkinterfaces durch die Zone

Für eine Zone mit dem Namen *keetonga* auf einem separaten zpool *keetonga_rpool* ist das die folgende vereinfachte Konfiguration:

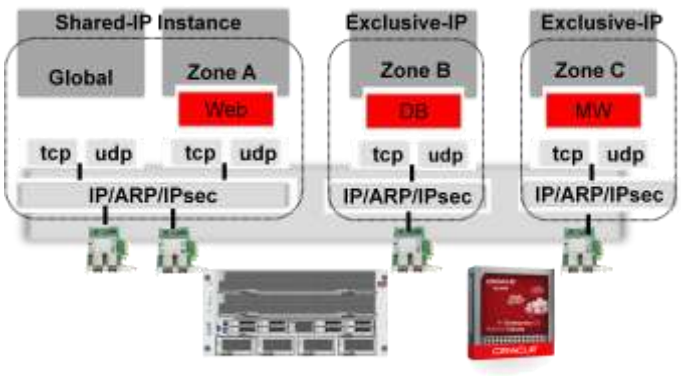
```
# zpool create keetonga_rpool
# zfs create keetonga_rpool/zoneroot
# zonecfg -z keetonga
zonecfg:keetonga> create
zonecfg:keetonga> set zonepath=/keetonga_rpool/zoneroot
zonecfg:keetonga> commit
zonecfg:keetonga> exit
#
```

In Solaris 11 ist damit eine Zone konfiguriert und kann installiert werden.

Vernetzung mit Oracle Solaris 11 und Zonen

Eine Zuweisung und Konfiguration eines Netzwerkinterfaces zu einer Zone ist möglich, aber nicht notwendig. Zur Vereinfachung wird beim Start einer Zone standardmäßig ein virtuelles Netzwerkinterface (VNIC) auf einem Netzwerkinterface der globalen Zone erzeugt und der Zone zugewiesen. Diese Einstellung kann verändert werden. So kann z.B. festgelegt werden:

- Ob der Zone kein VNIC, sondern ein physikalisches Interface zugewiesen werden soll
  `zonecfg: set physical=`
- Auf welchem Interface der globalen Zone das VNIC erzeugt werden soll
  `zonecfg: select anet linkname=net0; set lower-link=`
- Ob weitere Interfaces oder VNIC der Zone zugewiesen werden sollen
  `zonecfg: add net`
  `zonecfg: add anet`
- Welche Eigenschaften die Interfaces haben sollen (Mac-Adressen, MAC- und IP-Adressbeschränkungen, Netzwerk-Ressourcenmanagement)
  `zonecfg: select anet linkname=net0; set allowed-address= ....`
- Ob die Zone ihren eigenen TCP/IP-Stack erhält oder den durch die globale Zone bereitgestellten TCP/IP-Stack mitbenutzt.
  `zonecfg:set ip-type=`



Ressourceneinstellungen für Zonen

Das Management und die Begrenzung des Ressourcenverbrauches durch Zonen steuert die kontrollierte Zuteilung von gemeinsam genutzten Ressourcen der globalen Zone, um z.B. unterschiedlichen Anforderungen in Hinblick auf Durchsatz, Antwortzeit oder Verfügbarkeit gerecht zu werden. Dabei kommen unterschiedliche Prinzipien der Ressourcenzuteilung wie die Partitionierung, die Kappung oder das Scheduling von Ressourcen in Hinblick auf CPU, Hauptspeicher, Virtuellem Hauptspeicher und I/O zum Einsatz. Ressourceneinstellungen werden in den Konfigurationen der Zonen vorgenommen und werden vor allem duch die Anwendungen beeinflusst, die in den Zone ablaufen werden.

Installation von Zonen

Nachdem alle Konfigurationseinstellungen vorgenommen wurden, kann die Zone installiert werden. Dabei werden die ZFS Datasets für die Zone erzeugt und aus dem in der globalen Zone konfigurierten Paket-Repository die Pakete zur Installation der Zone geladen und in die Zone installiert.

```
# zoneadm -z keetonga install
```

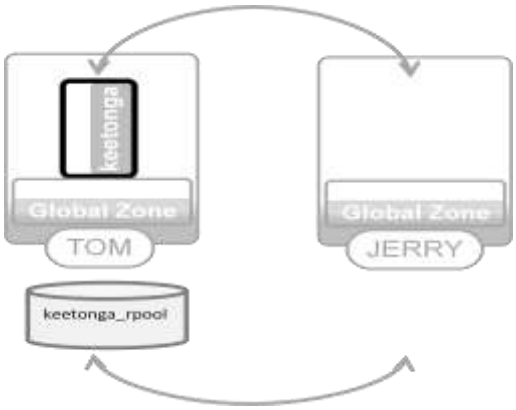
Betrieb, Update und Verschiebung von Zonen

Die installierte Zone kann nun mit *zoneadm -z keetonga boot* gestartet werden und stellt sich aus Sicht eines Anwenders wie eine unabhängige Solaris Instanz dar. Nun kann z.B. die Installation einer Datenbank oder eines Application-Servers erfolgen.

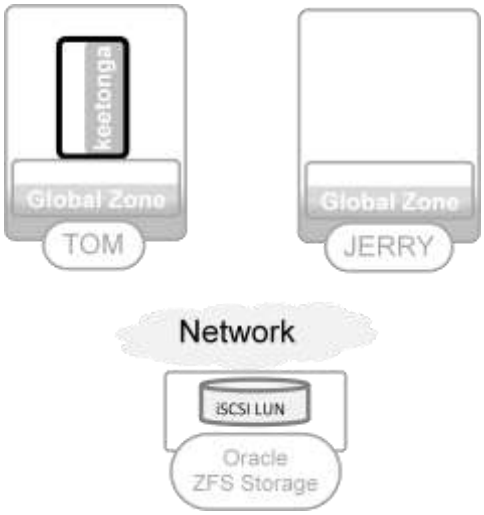
Sind Paket-Updates einer Zone notwendig, so erfolgt dies für systemnahe Pakete während des Update-Prozesses der globalen Zone (`pkg update`). So wird ein konsistenter Softwarestand zwischen der globalen Zone und den nicht-globalen Zonen sichergestellt.

Für einen flexiblen Betrieb von Zonen ist zwar auf Grund der benutzten Virtualisierungstechnologie der Zonen keine Live-Migration zwischen zwei Systemen möglich, aber Zonen können "kalt" migriert werden. Dazu werden sie auf einem System `detached` und auf einem anderen System `attached`. Dabei werden beim `attach` ggf. bestehende Unterschiede in System-Paketen der Zone an die der globalen Zone des Zielsystems während des `attach` angeglichen. Während der Verschiebung der Zone muss sichergestellt werden, dass das Root-Filesystem der Zone als auch ggf. vorhandene Anwenderdaten mit verschoben werden, damit sie von dem zweiten System nutzbar werden.

```
tom # zoneadm -z keetonga detach
tom # zpool export keetonga_rpool
jerry # zpool import keetonga_rpool
... Zonenkonfiguration von tom nach jerry kopieren ...
jerry # zoneadm -z keetonga attach
```



Zones on Shared Storage (ZOSS) mit Oracle ZFS Storage



Für eine weitere Vereinfachung des Aufbaus ist die Benutzung von iSCSI-Storage in der Form des Oracle ZFS Storage denkbar. Hierbei entfällt der Aufbau von SAN-Infrastrukturen zur Bereitstellung von Shared-Storage. Storage wird als in LUNs über iSCSI Targets über das Netzwerk bereitgestellt und kann von iSCSI Initiators genutzt werden.

Mit Solaris 11.2 wurden Funktionalitäten zur Verwaltung des Zones Storage mit in das Zones-Framework integriert, um die Benutzung von iSCSI-LUNs, aber auch FC- oder SAS-LUNs einfacher zu gestalten.. So genügt die Erzeugung und Bereitstellung der iSCSI-Targets und LUNs am ZFS Storage. Die LUNs werden in der Konfiguration der Zone angegeben. Damit entfällt die Erzeugung und das Import des zpool vor der Konfiguration und der Erzeugung der Zone. Das Zones-Framework übernimmt diese Aufgabe nun während der Erzeugung der Zone.

Auch die Verschiebung von Zonen wird durch diese Funktionalität unterstützt, denn das Zones-Framework exportiert beim `zoneadm detach` und importiert beim `zoneadm attach` den oder die entsprechenden zpools.

Hier ein Beispiel:
An einem ZFS Storage wurde ein iSCSI Target und eine LUN erzeugt.

Die LUN hat die Nummer `600144F0F69444CB000052299DE9000A`. Storage Komponenten werden in der Zonekonfiguration durch Storage Unique Resource Identifier (SURI) angegeben, um die sich von Gerätebezeichnungen wie /dev/dsk/... zu lösen und unabhängig zu werden. Das ist wichtig, da ja Zonen auch auf andere Server bewegt werden sollen und ein iSCSI-Gerät auf einem anderen Server als "Festplatte" mit einer anderen Device-Bezeichnung gemapped werden können. Das ist im übrigen auch für FC-Geräte zu beachten. Deshalb werden `suri` benutzt, die iSCSI-Target Server und LUN-IDs oder bei FC WWN Nummern enthalten.

Da wir hier iSCSI benutzen, setzt sich eine suri wie folgt zusammen:
`iscsi://<Name oder IP des iSCSI-Target>/luname.naa.<LUN-ID>`

Diese LUN soll als Storage für den zpool der Zone keetonga genutzt werden.

```
# zonecfg -z keetonga
zonecfg:keetonga> create
zonecfg:keetonga> set zonepath=/keetonga
zonecfg:keetonga> add rootzpool
zonecfg:keetonga:rootzpool> add storage \
iscsi://192.168.175.10/luname.naa.600144F0F69444CB000052299DE9000A
zonecfg:keetonga:rootzpool> end
zonecfg:keetonga > commit
zonecfg:keetonga > exit
#
```

Fertig. Alle weiteren Schritte folgen wie oben. Das Zones-Framework übernimmt die Verwaltung des Storage bei der Installation und Verschiebung der Zonen. Dabei werden die iSCSI-LUNs importiert, Device-Einträge erzeugt, der zpool `keetonga_rpool` (Standardname gebildet aus <zonename>_rpool) erzeugt, importiert und die Zone installiert.

Ein `zoneadm detach` exportiert den zpool und setzt die Zone in den Zustand `configured`. Um eine ZOSS-Zone zu verschieben genügt es, die Zonenkonfiguration auf das Zielsystem zu übertragen und ein `zoneadm attach` auszuführen. Entsprechende /dev/dsk/... Einträge werden durch das Zones-Framework erzeugt, die iSCSI-LUNs importiert, der zpool importiert und die Zone attached.

iSCSI Security

Ausgehend von dem obigen Beispiel ergibt sich eine einfache Möglichkeit, Zonen auf Shared Storage zu erzeugen:

- Der Storage Administrator erzeugt iSCSI Targets und LUN's.
- Die Targets und LUN's werden dem Administrator der Server mitgeteilt.
- Auf Basis der LUN's werden Zonen erzeugt.
- Zonen können sehr einfach verschoben werden.

Bei der Benutzung von iSCSI-Storage sind jedoch einige Dinge zu beachten, um den Zugriff zu begrenzen.

- iSCSI-Daten zwischen Server und Storage sollten über ein separiertes Netzwerk übertragen werden.
- Nicht jeder iSCSI-Initiator sollte alle iSCSI-Targets und LUNs sehen dürfen. Deshalb sollte eine Form der Maskings auf iSCSI-Ebene genutzt werden.z.B. Welche iSCSI-Initiatoren dürfen welche iSCSU-Targets sehen.
- Zur weiteren Erhöhung der Sicherheit sollte CHAP verwendet werden. Dabei authentisiert bei unidirectionalem CHAP das Ziel (Target) den Initiator, aber nicht der Initiator das Target. Beim der bidirectionalen CHAP wird zusätzlich das Target durch den Initiator authentisiert. Eine weitere Möglichkeit zu Athentisierung ist die Verwendung eines RADIUS-Servers.
- Zusätzlich kann bei Solaris der in den iSCSI-LUN's abgelegte zpool encrypted Datasets verwenden. Das führt zu einem weiteren Schutz der abgelegten Datenblöcke im iSCSI-Storage, da die Blöcke verschlüsselt abgelegt werden. Auch der Transportweg ist, da die ZFS-Blöcke bereits verschlüsselt zum iSCSI-Storage übertragen werden.

Über die verschiedenen Möglichkeiten lässt sich so sicherstellen, dass jeweils nur berechtigte Initiatoren auf die iSCSI-LUN's zugreifen.

Zusammenfassung

Solaris Zones bieten eine leichtgewichtige und kostengünstige Form der Virtualisierung zur Erzeugung von Laufzeitumgebungen für verschiedene Arten von Diensten. Werden für das root-Filesystem der Zone und die Daten Shared Storage-Technologien verwendet, so ergeben sich Vereinfachungen beim Zones- und Storage-Handling. iSCSI-Storage kann hier eine preiswerte Storage-Plattform bilden, wenn Sicherheits- und Performanceaspekte beachtet werden.

Kontaktadressen:

Detlef Drewanz
Oracle Deutschland B.V. & Co. KG
Schiffbauergasse 14
D-14467 Potsdam
Telefon: +49 (0) 331 200 7341
E-Mail Detlef.Drewanz@oracle.com
Internet: http://www.oracle.com