# Neue Technologien zur Netzwerk-Virtualisierung in Solaris 11

**Detlef Drewanz**
**Oracle Deutschland B.V. & Co. KG**
**Komturstraße 18a, 12099 Berlin**

**Schlüsselworte:**

Oracle Solaris Container, Solaris Zonen, Netzwerk, Solaris, Virtualisierung, VirtualBox

## Einleitung

Mit der wachsenden Nutzung virtualisierter Betriebssystem-Instanzen wächst auch die Bedeutung der Vernetzung dieser Instanzen - einerseits im Hinblick auf eine performante und zugleich flexible Nutzung der physischen Schnittstellen und andererseits im Hinblick auf die Vernetzung virtualisierter Instanzen untereinander - innerhalb eines Systems.

Solaris 11[1] stellt eine Reihe neuer Technologien bereit, die mit einfachen Mitteln das Bandbreiten-Management und die Virtualisierung des Netzwerkstacks von Solaris 11 ermöglichen und die auf dem OpenSolaris Projekt „Crossbow“ basieren. So können beispielsweise physische Netzwerk-Interfaces in virtuelle Interfaces mit Bandbreitenbegrenzungen unterteilt werden, virtuelle Ethernet-Switches in Software erzeugt werden oder virtuelle Softwarestacks für spezielle Netzwerk-Services erzeugt werden. Daraus ergeben sich neue Möglichkeiten der Server- und Netzwerk-Konsolidierung oder für die Bereitstellung von Services in Clouds.

Im Folgenden werden die grundlegenden Funktionsweisen und die Anwendung dieser Technologien erläutert, sowie Möglichkeiten dargestellt, wie daraus neue Architekturen für vernetzte virtualisierte Anwendungen entstehen können.

## Konsolidierung und Virtualisierung im Rechenzentrum

In den vergangenen Jahren ist die Leistungsfähigkeit von Prozessoren und entsprechenden Systemen immens gewachsen. Dieser Leistungszuwachs ist nicht nur auf die Steigerungen der Taktfrequenz der Prozessoren zurückzuführen, sondern auch durch die ständig größer werdende Anzahl von Prozessorkernen, die per CPU-Socket bereitgestellt werden.

Allerdings wächst mit diesen entwickelten Systemen auch der Stromverbrauch und die Wärmeleistung je System. Ein Mehrverbrauch an Strom und Kühlungsressourcen erhöht die Kosten für den Betrieb neuer Systeme. Andererseits können neue Systeme oft ein Vielfaches der Arbeitslast älterer Systeme leisten.

---

1 Oracle plant die Verfügbarkeit von Oracle Solaris 11 für das Kalenderjahr 2011. Das erste Oracle Solaris 11 Express Release - zum Ende des Kalenderjahres 2010 erwartet - ermöglicht Kunden Zugang zu neuesten Oracle Solaris 11 Features; über einen optionalen Vertrag auch mit Support. Solaris 11 Express basiert auf Technologien früherer Solaris- und OpenSolaris-Releases. Für weitere Details siehe: http://www.oracle.com/us/corporate/press/173478 und http://www.oracle.com/technetwork/server-storage/solaris11

Zur besseren Ausnutzung dieser leistungsfähigen Systeme werden u.a. verschiedene Virtualisierungstechnologien eingesetzt. Das ermöglicht die Arbeitslast vieler kleiner, älterer, wenig leistungsfähigerer Systeme, auf leistungsfähige Systemen zu verlagern. Im Ergebnis reduziert sich die Anzahl der zu betreibenden Systeme und dadurch die Kosten für Strom, Kühlung und Stellfläche.

Virtualisierung ermöglicht aber auch die weitere Flexibilisierung von Computing-Ressourcen und ein schnelleres Ausrollen von Anwendungen, da die Installation der physischen Hardware entfallen kann. Virtualisierte Anwendungen können oft bereits von der benutzten Hardware unabhängig betrachtet werden und gestatten so eine von der Hardware abstrahierte Betrachtungsweise der Services. Die Bereitstellung und das Management von Services, sowie das Upgrade von benötigter Hardware vereinfachen sich so drastisch. Services in virtualisierten Umgebungen bilden heute z.B. eine Grundlage für das Cloud Computing.

Virtualisierung im Bereich von Prozessoren, Systemen oder Betriebssystemen ist heute in jedem Rechenzentrum in der einen oder der anderen Art im Einsatz. Jedoch sind die Bereitstellung von Services und die Virtualisierung ohne Vernetzung undenkbar. Die Vernetzungskomponente des gesamten Virtualisierungsstacks gewinnt sogar noch in dem Maße an Bedeutung, in dem die Architekturen zur Virtualisierung immer komplexer werden, die einzelnen virtualisierten Services immer leistungsfähiger werden und die räumliche Verteilung der einzelnen Virtualisierungskomponenten in immer komplexeren Netzwerkarchitekturen abgebildet werden müssen.

Die Betrachtung von Netzwerktechnologien in Betriebssystemen bzgl. deren Eignung für die Virtualisierung wird so immer wichtiger. So werden IT-Architekten vor neue Herausforderungen gestellt, wenn z.B. 20 Virtuelle Maschinen oder Oracle Solaris Zonen auf einem System betrieben werden sollen, aber jeder Einheit ein eigenes unabhängiges Netzwerkinterface zur Verfügung gestellt werden soll. Oftmals reicht die Anzahl verfügbarer I/O-Steckplätze bereits nicht aus, um die benötigte Anzahl Netzkarten aufzunehmen. Die Nutzung einer Form der Netzwerkvirtualisierung kann hier helfen.

**Limitierungen vorhandener Netzwerktechnologien bei der Virtualisierung**

- Vorhandene Formen der Netzwerkvirtualisierung basieren oft auf VLAN's, die in den Netzwerkswitches konfiguriert werden. Wird so eine VLAN-ID an eine virtuelle Maschine (VM) gebunden, ist diese ebenfalls zu migrieren oder am Switch umzukonfigurieren, wenn die VM migriert wird.
- Vorhandene Lösungen für virtuelle Interfaces können selten als vollwertiger Ersatz für physische Interface benutzt werden. Oft können darauf keine Tools wie Netzwerkmonitore und Firewalls genutzt werden oder die virtuellen Interfaces verfügen nicht über eigene Statistikzähler.
- Ein Netzwerk-Accounting, bezogen auf virtuelle Interfaces und damit bezogen auf virtuelle Maschinen oder virtualisierte Services, ist nahezu unmöglich oder nur mit sehr großem Aufwand realisierbar.
- Teilen sich mehrere virtuelle Maschinen ein Netzwerkinterface, so ist eine definierte Bandbreitenzuordnung nur mit komplexen Konfigurationen möglich (z.B. IP-Quality of Service=IPQoS).
- IPQoS-Konfigurationen sind oft von zusätzlichen Funktionalitäten der Netzwerkinfrastuktur abhängig (Router, Switches). Sie sind an den physischen Host gebunden und nicht an die VM, für die sie konfiguriert wurden.

**Netzwerkvirtualisierung in Oracle Solaris 11**

In OpenSolaris sind im Rahmen des Crossbow-Projektes Technologien zur Netzwerkvirtualisierung und zum Ressourcenmanagement für Netzwerke entwickelt worden. Mit der Verfügbarkeit von Oracle Solaris 11 und Oracle Solaris 11 Express werden diese Funktionalitäten auch im Enterprise-Bereich nutzbar.

Bei der Entwicklung des Crossbow-Projektes wurde besonderer Wert auf die Einfachheit in der Benutzung, die Unabhängigkeit von der verwendeten Netzwerkinfrastuktur und die Nutzbarkeit mit anderen Virtualisierungstechnologien gelegt. Es wurde ein ganzheitlicher Ansatz verfolgt, der - ausgehend von den Netzwerkschnittstellen - auch die CPU, die Netzwerktreiber, den Betriebssystemkern als auch den TCP/IP-Stack mit seinen verschiedenen Komponenten erfasst. So werden z.B. Möglichkeiten bereitgestellt, für jeden Netzwerkservice einen eigenen virtuellen Netzwerkstack zur Verfügung zu stellen.

Unter anderem sind die folgenden Komponenten im Crossbow-Projekt implementiert worden:

- **VNIC**: Virtualisierung der physischen Netzwerkinterfaces als virtuelle Interfaces

  VNIC gestatten die Erzeugung von beliebig vielen virtuellen Netzwerkinterfaces und so deren transparente Benutzung beim Aufbau von Architekturen mit Virtuellen Maschinen und Oracle Solaris Zonen.

- **Etherstubs:** Erzeugung von VNIC auf pseudo-Netzwerkinterfaces

  Komplette Netzwerkarchitekturen können so in Software realisiert werden. Verschiedene Broadcast-Domains werden erzeugt, unabhängig von dem Vorhandensein phys. Netzwerkinterfaces. Diese werden per VNIC verbunden.

- **Flows**: Klassifizierung von Netzwerkverkehr im Kommunikationsstack nach Adressen und Protokollen

  Diese Klassifizierung ermöglicht eine definierte Behandlung von Netzwerkverkehr bei der gemeinsamen Benutzung von Kommunikationsstacks und Netzwerkkarten.

- **Ressourcenmanagement**: Bandbreitenbegrenzung, CPU-Zuteilung und Prioritäten für VNIC und Flows

  Das Ressourcenmanagement steuert und begrenzt den übermäßigen Ressourcenverbrauch einzelner Komponenten und garantiert Antwortzeiten bzw. verhindert einen Serviceausfall durch Überlastung.

- **Einfaches Management** für VNIC, Virtuelle Switches und Flows

  Alle Funktionalitäten sind sehr einfach und nach kurzem Studium der Handbücher möglich. Dazu wurden einheitliche Konfigurationswerkzeuge geschaffen.

**VNIC**

Durch virtuelle Netzwerkinterfaces (VNIC) werden physische Netzwerkinterfaces in virtuelle Interfaces aufgeteilt und können so z.B. von mehreren virtuellen Maschinen (VM) oder Oracle Solaris Zonen gleichzeitig genutzt werden. Jede Zone oder VM kann so ihr eigenes VNIC erhalten oder die VNICs können vom Solaris im TCP/IP-Stack der globalen Zone genutzt werden.

Jedes VNIC verhält sich wie ein physisches Interface, mit eigenen Netzwerkstatistiken und eigener MAC-Adresse. Bei der Erzeugung der VNIC kann angegeben werden, ob eine der vordefinierten MAC-Adressen der Netzkarte (wenn vorhanden), eine zufällige oder eine vom Administrator vorgegebene MAC-Adresse benutzt werden soll. So ist aus Sicht der Anwendung oder der virtuellen Umgebung nicht mehr unterscheidbar, ob ein physisches Netzwerkinterface oder ein VNIC benutzt wird, ist also transparent von Virtuellen Maschinen, Zonen oder Anwendungen nutzbar.

Die VNIC auf einem phys. Interface sind durch einen virtuellen Netzwerkswitch verbunden, befinden sich also in einer sog. Broadcast-Domain.  Der Datenverkehr der VNIC untereinander bleibt innerhalb des virtuellen Switch.

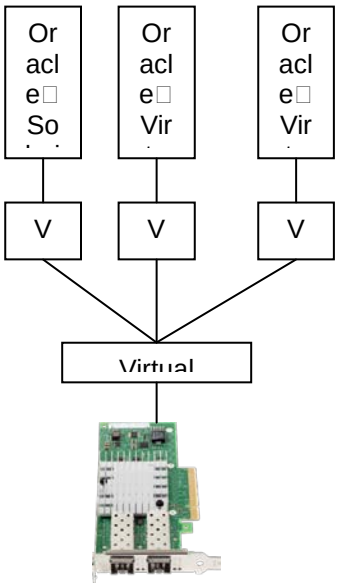


*Abb. 1: VNICs auf einem Interface, mehreren Zonen und VM zugeordnet*

Moderne Netzwerkkarten verfügen zusätzlich über Klassifizierungsmöglichkeiten des Netzwerkverkehr, die in Hardware Rings oder separaten DMA-Kanälen abgebildet werden. Bei der Erzeugung von VNICs werden die Hardware Rings von den VNICs benutzt, d.h. durch die Nutzung der Hardwarefunktionalitäten bei modernen Netzwerkkarten kann der Mehraufwand für die Virtualisierung erheblich reduziert werden. Sind keine Hardware Rings verfügbar, wird die Funktionalität in Software abgebildet.

**Etherstubs**

Wie in Abb. 1 gezeigt, werden VNIC durch virtuelle Ethernetswitches verbunden, die oberhalb der physischen Netzwerkkarten erzeugt werden. Virtuelle Ethernetswitches für VNIC können jedoch auch ohne vorhandene physische Interfaces erzeugt werden. Dazu werden Etherstubs als pseudo Ethernetadapter erzeugt und auf ihnen die VNIC erzeugt.

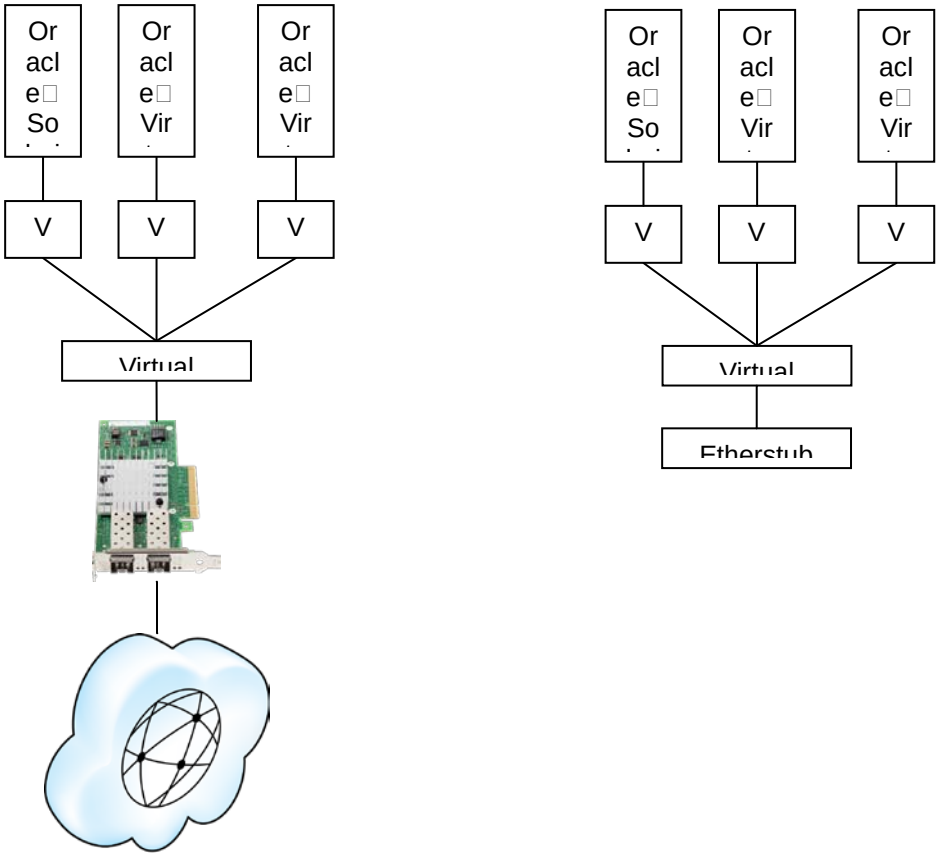


*Abb. 2: VNIC mit physischem Interface und mit Etherstub*

Etherstubs bieten so die Möglichkeit, komplexe Netzwerkarchitekturen komplett in Software abzubilden, ohne sie mit einem physischen Netzwerk zu verbinden.

**Flows**

Flows bezeichnen eine Teilmenge des Netzwerkverkehrs, der den gesamten Netzwerkstack umfassen kann - vom Netzwerkinterface über die Sockets bis hin zum Kernelthread, der diese Daten bearbeitet. Flows werden z.B. durch IP-Adressen, Portnummern oder Protokolltypen beschrieben. Dadurch wird eine eindeutige Zuordnung von Flows zu Datenströmen im Netzwerkstack möglich. Soll also ein bestimmter Protokolltyp oder eine bestimmte Kommunikationsverbindung besonders behandelt werden, kann ihr ein Flow zugeordnet werden. Später können diesem Flow Eigenschaften zur Ressourcenlimitierung zugeordnet werden. Auf diese Weise ist sehr einfach ein Ressourcenmanagement auf Verbindungs- oder Protokollebene im Solaris möglich, unabhängig von den von der Verbindung genutzten Netzwerkadaptern, Switchen oder Routern. Die Ressourcenkontrolle erfolgt am Kommunikationsendpunkt.

**Ressourcenmanagement**

Erzeugte VNIC, Etherstubs oder Flows benutzen ohne Limitierung die maximal verfügbare Bandbreite der Netzwerkverbindung bzw. die maximal verfügbaren CPU- und TCP/IP-Stack Ressourcen zur Realisierung der Kommunikation.

Im Umfeld der Virtualisierung ist jedoch oft die Anzahl der Virtuellen Maschinen oder Zonen erheblich größer, als die Anzahl der verfügbaren physischen Interfaces. In diesem Fall müssen sich mehrere virtuelle Umgebungen ein physisches Interface teilen. Dazu werden VNIC erzeugt und einzelnen virtuellen Umgebungen zugeordnet. Hierbei haben unterschiedliche virtuelle Umgebungen jedoch oft unterschiedliche Anforderungen an die Bandbreite und Verfügbarkeit von Netzwerkverbindungen. Wird nun jedoch nicht der Ressourcenverbrauch einzelner VNIC begrenzt, kann es dazu führen, das einzelne Verbindungen die gesamte Bandbreite einer Netzkarte oder eines gemeinsam genutzten Stacks „verbrauchen“ und anderen Verbindungen keine Bandbreite mehr zur Verfügung stehen.

Zusätzlich werden auch CPU-Ressourcen benötigt, um z.B. die Daten von Netzwerkkarten „abzuholen“ bzw. Daten von Verbindungen im TCP/IP-Stack zu verarbeiten. Auch hier kann es dazu führen, das einzelne Verbindungen zu Lasten anderer zu viele Ressourcen belegen und so u.U. die Arbeit anderer Verbindungen behindern bzw. ihnen nicht ausreichend Ressourcen zur Verfügung stehen.

Ein Beispiel sind hier die sogenannten „Denial-of-Service“ Attacken, bei denen einzelne Dienste mit hohen, teilweise unsinnigen Anforderungen überflutet werden und so ganze Systeme lahmgelegt werden können. Die Ursache dafür liegen darin, das auch fehlerhafte oder unsinnige Pakete zunächst empfangen und verarbeitet werden müssen um festzustellen, dass sie unsinnig sind. Allein für diese Verarbeitung werden aber auch bereits eine Menge CPU- und Bandbreitenressourcen benötigt. Bisher half hier oft als Gegenmaßnahme nur das Abschalten oder Sperren der Dienste bzw. bestimmter Adressen, womit Angreifer Ihr Ziel erreicht hätten.

Um solche Szenarien zu verhindern wäre es wünschenswert, den Ressourcenverbrauch von VNIC, Etherstubs und Flows zu begrenzen. So könnte man die Verarbeitungspriorität und damit den CPU-Verbrauch einzelner Verbindungen sehr weit herabsetzen, kann sie so also weiterhin aktiv halten, den Einfluss auf das Gesamtsystem jedoch absenken.

Das Ressourcenmanagement für Netzwerkverbindungen und Interfaces kann auf verschiedene Bereiche angewendet werden:

- Bandbreite
- CPU
- Priorität

**Accounting und Statistiken**

Je nach Einsatzfall für VNIC und Flows ist es erforderlich, eine Aufstellung über die Datenmenge zu erstellen, die von einem VNIC, einem virtuellen Netzwerkstack oder einem Flow verarbeitet wurde. Dazu sind entsprechende Statistikzähler und Tools in Oracle Solaris 11 integriert worden. So können zur Laufzeit aktuelle Statistiken für VNIC und Flow erzeugt werden oder die Daten aufgezeichnet und später analysiert werden.

**Managementinterface**

Das Managementinterface für die Netzwerkvirtualisierung und das Ressourcenmanagement in Oracle Solaris 11 setzt sich sehr einfach aus 2 Kommandos mit verschiedenen Optionen zusammen:

- dladm(1M) erzeugt, konfiguriert und löscht unter anderem VNIC und Etherstub, wird aber noch für eine Vielzahl anderer Funktionen im Netzwerk-Umfeld benutzt
- flowadm(1M) erzeugt, konfiguriert und löscht Flows

Zur Erzeugung von Statistiken für Interfaces, VNICs, Links und Flows stehen die folgenden Kommandos zur Verfügung:

- dlstat(1M) zeigt Laufzeitstatistiken über Datenlinks an
- flowstat(1M) zeigt Laufzeitstatistiken über definierte Flows an
- acctadm(1M) ist das Solaris extended Accounting Facility und kann u.a. zur Aufzeichnung von Netzwerkstatistiken benutzt werden, die später analysiert werden sollen

**VNIC Management**

dladm (1M) ist das administrative Kommando für VNIC und Etherstub. Der Name für VNIC und Etherstub kann frei gewählt werden, muss jedoch auf eine Ziffer enden. Das folgende Kommando erzeugt ein VNIC auf dem physischen Interface e1000g0 und weist eine zufällige MAC-Adresse zu.

```
# dladm create-vnic -l e1000g0 vnic1
```

VNIC können auch mit vorgegebener MAC-Adresse, VLAN-ID (hier 17) und beschränkter Bandbreite (hier 2 Mbps) erzeugt werden.

```
# dladm create-vnic -l e1000g0 -v 17 \
  -m 0:1:2:3:4:5 -p maxbw=2M vnic1_in_vlan17
```

Eine Anzeige der beiden vorhandenen VNIC sieht danach wie folgt aus:

```
# dladm show-vnic
LINK            OVER    SPEED  MACADDRESS      MACADDRTYPE
VID
vnic1           e1000g0  1000  2:8:20:a9:7c:d2 random
0
vnic1_in_vlan17 e1000g0   2    0:1:2:3:4:5     fixed
17
```

**Etherstub Management**

Etherstubs können zur Erzeugung virtueller Switches benutzt werden, die nicht an physische Interfaces gebunden sind.

```
# dladm create-etherstub luftnummer1
```

Auf diesem Etherstub kann ein VNIC erzeugt werden.

```
# dladm create-vnic -l luftnummer1 luftnummer_vnic1
```

Danach zeigt die Aufstellung der vnic:

```
# dladm show-vnic
LINK            OVER    SPEED  MACADDRESS      MACADDRTYPE
VID
vnic1           e1000g0  1000  2:8:20:a9:7c:d2 random
0
vnic1_in_vlan17 e1000g0   2    0:1:2:3:4:5     fixed
17
luftnummer_vnic1 luftnummer1 0 2:8:20:13:c7:17 random
0
```

**Flow Management**

Flows können mit flowadm(1M) erzeugt werden und zur fein granulierten Steuerung von Datenverkehr im Netzwerkstack genutzt werden. Dazu werden Flows einem Datentyp zugeordnet und dann einem Interface Priorität und Bandbreite zugewiesen.

```
# flowadm add-flow -l e1000g0 -a transport=icmp -p
maxbw=1200K,priority=low icmp-flow

# flowadm show-flow
FLOW      LINK             IPADDR PROTO  LPORT   RPORT   DSFLD
icmp-flow e1000g0          --     icmp   --      --      --
```

**Anzeige von Statistiken**

Flows und VNIC verfügen über eigene Statistiken, die direkt abgerufen werden können oder aufgezeichnet und zu einem späteren Zeitpunkt angezeigt werden können. Statistiken der VNIC werden mit dlstat(1M) ermittelt, die der Flows mit flowstat(1M). Wird nach obiger Konfiguration ein `ping -s` über Interface e1000g0 ausgeführt, ergibt sich die folgende direkte Statistik.

```
# flowstat -i 5
FLOW      IPKTS   RBYTES    IERRS    OPKTS   OBYTES    OERRS
icmp-flow   11    1,08K        0       11    1,08K        0
icmp-flow    5      490        0        5      490        0

# dlstat e1000g0 -i 5
LINK       IPKTS   RBYTES    OPKTS   OBYTES
e1000g0   41,51K    2,93M    2,27K  195,97K
e1000g0        1       60        1       42
```

Aufzeichnungen für eine spätere Analyse können so aktiviert werden:

```
# acctadm -e extended -f /var/tmp/netlog.acct net
```

und so später ausgewertet werden:

```
# dlstat show-link -h -f netlog.acct e1000g0
```

```
LINK     START    END          RBYTES   OBYTES   BANDWIDTH
e1000g0  08:11:24 08:11:44    0        42       0 Mbps
e1000g0  08:11:44 08:12:04    326      42       0 Mbps
```

```
# flowstat  show-link -h  -f /tmp/netlog icmp-flow
FLOW       START      END        RBYTES   OBYTES   BANDWIDTH
icmp-flow  08:11:24   08:11:44   1960     1960     0.001 Mbps
icmp-flow  08:11:44   08:12:04   1960     1960     0.001 Mbps
```

Sind die Aufzeichnungen der Statistik nicht mehr erforderlich, sollte das Accounting abgeschaltet werden:

```
# acctadm -x net
```

**Anwendungsbeispiele: VNIC in VirtualBox-Instanzen**

VNIC können z.B. separaten Oracle VirtualBox-Instanzen (VBox) zugewiesen werden, um die genutzte Bandbreite vom Host aus zu begrenzen. Oder es können Setups entworfen werden, in denen zwei VBox über VNIC und Etherstub miteinander kommunizieren, ohne das ein physisches Netzwerkinterface benötigt wird.

Das folgende Beispiel zeigt die Verbindung von zwei VBox zu Testzwecken. Das Testsetup kommt ohne physische Netzwerkports aus und basiert auf einem Etherstub mit zwei VNIC.

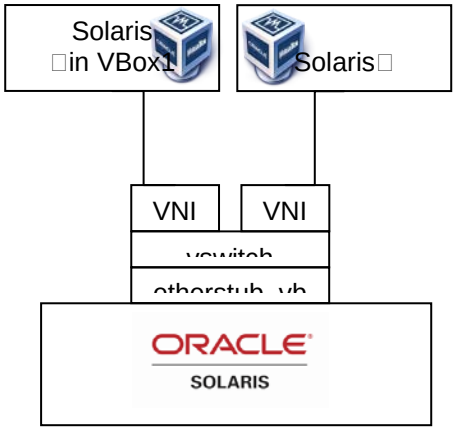


*Abb. 3: VNIC und VirtualBox*

1. Erzeugung des Etherstub und der VNIC im Solaris Hostsystem

```
host # dladm create-etherstub etherstub_vbox1

host # dladm create-vnic -l etherstub_vnic1 vnic1

host # dladm create-vnic -l etherstub_vnic1 vnic2

host # dladm show-vnic
```

```
LINK   OVER              SPEED  MACADDRESS      MACADDRTYPE
VID
vnic1  etherstub_vnox1  0      2:8:20:b6:f5:92 random         0
vnic2  etherstub_vbox1  0      2:8:20:8e:29:5f random         0
```

2. Erzeugung der VBox mit dem jeweils gewünschten Betriebssystem.

3. Zuweisung der VNIC zur VBox
Zur Realisierung der Funktionsfähigkeit in den VBox mit VNIC, wird der Modus „Netzwerkbrücke“ verwendet. Die MAC-Adressen der VNIC müssen den VBox bekannt gemacht werden.

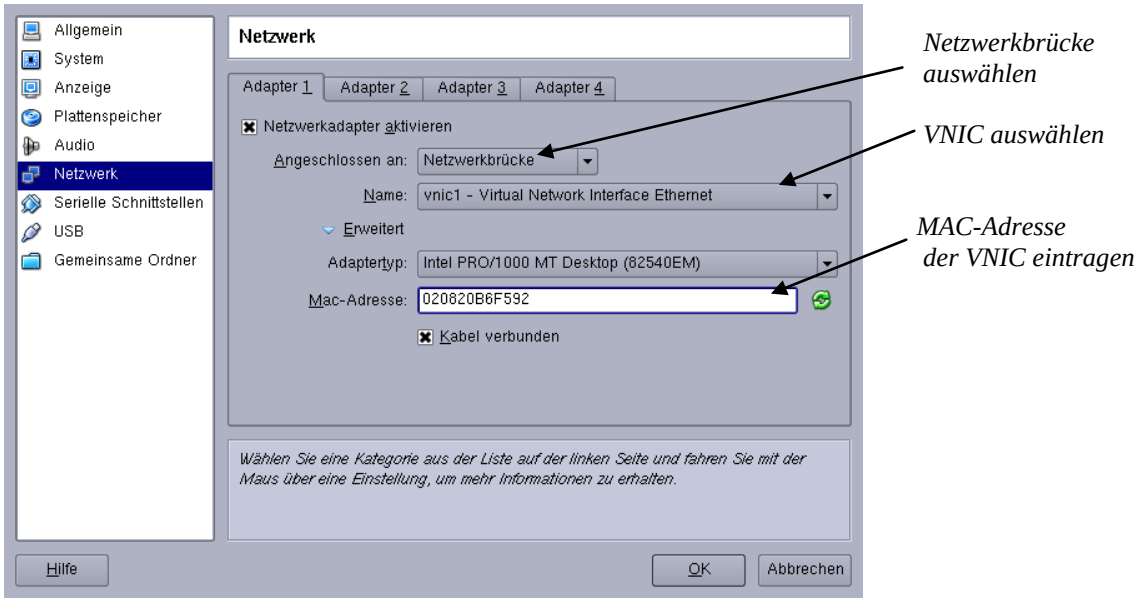

*Abb. 4: VNIC, Netzwerkbrücke und MAC-Adresse für VirtualBox-Instanzen setzen*

4. Innerhalb der VBOX sind die VNIC wie physische Interface mit den MAC-Adressen sichtbar und können entsprechend genutzt werden. Im Folgenden werden die Einstellungen beispielhaft für die vbox1-Instanz mit Solaris gezeigt.

```
vbox1 # ifconfig e1000g0 plumb

vbox1 # ifconfig e1000g0 192.168.100.1 up

vbox1 # ifconfig e1000g0
e1000g0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu
1500 index 3  inet 192.168.100.1 netmask ffffff00 broadcast
192.168.100.255  ether 2:8:20:b6:f5:92
```

5. Die Verbindung ist hergestellt und kann benutzt werden.

**VNIC und Oracle Solaris Zonen**

Oracle Solaris Zonen ist die Bezeichnung für eine virtualisierte Ausführungsumgebung in Solaris - eine Virtualisierung auf Betriebssystem-Ebene. Alle Zonen teilen sich gemeinsam den Solaris Betriebssystemkern, teilweise auch den TCP/IP-Stack und Netzwerkinterfaces.

In verschiedenen Installationen ist jedoch eine Trennung des TCP/IP-Stack erforderlich, so dass jeder Zone ihr eigener TCP/IP-Stack zugewiesen werden kann. Dafür wurden in Solaris 10 die Exclusive-IP Instanzen eingeführt. Eine Zone mit einer Exklusiven-IP Instanz hat ihre eigene Kopie von Variablen und Tabellen, die vom TCP/IP-Stack benutzt werden. Diese Zone hat im Ergebnis eine eigene IP-Routingtabelle, ARP-Tabelle, IPsec-Policies, IP-Filter Regeln oder ndd-Einstellungen.

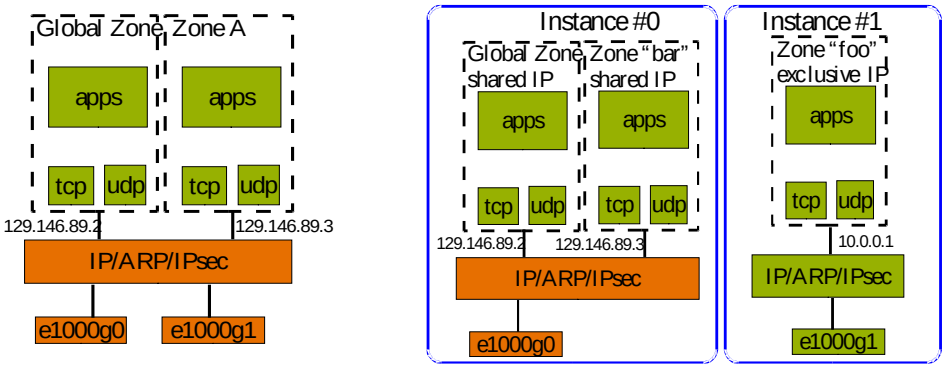


*Abb. 5: Vergleich Shared-IP und Exclusive-IP Instanz in Solaris Zonen*

Zur Benutzung einer Exclusive-IP Instanz pro Zone ist es jedoch notwendig, der Zone mindestens ein physisches Interface oder ein tagged-vlan Interface exklusive zuzuweisen. Werden jedoch viele Zonen in einem System benutzt, können oft nicht so viele Interfaces wie benötigt bereitgestellt werden. Die Gründe dafür liegen oft in der limitierten Anzahl von verfügbaren I/O-Steckplätzen oder den Bildungsregeln für tagged-vlan Interfaces im Betriebssystem. VNICs zeigen hier einen Ausweg auf und erlauben eine erhebliche flexiblere Nutzung von Netzwerk-Interfaces. So können nun genügend VNIC pro physischem Interface erzeugt werden und den Zonen exklusive zugeordnet werden.

Das Beispiel aus Abschnitt 6 kann jetzt um eine Zone als Kommunikationspartner erweitert werden.

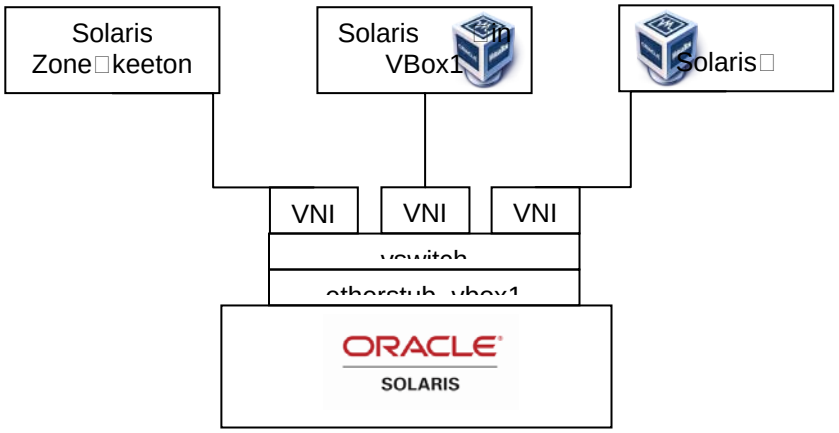


*Abb. 6: VNIC mit Solaris Zonen und VirtualBox*

1. VNIC erzeugen

```
host # dladm create-vnic -l etherstub_vbox1 vnic3

host # dladm show-vnic
LINK   OVER            SPEED  MACADDRESS        MACADDRTYPE
VID
vnic2  etherstub_vbox1 0      2:8:20:8e:29:5f   random
0
vnic1  etherstub_vbox1 0      2:8:20:b6:f5:92   random
0
vnic3  etherstub_vbox1 0      2:8:20:2b:f5:8e   random
0
```

2. Zone konfigurieren und erzeugen (Konfiguration der Zone keetonga mit einer Exclusive-IP Instanz)

```
host # zonecfg -z keetonga info
zonename: keetonga
zonepath: /zones/keetonga
brand: ipkg
autoboot: false
bootargs:
pool:
limitpriv:
scheduling-class:
ip-type: exclusive
hostid:
fs-allowed:
net:
        address not specified
        allowed-address not specified
        physical: vnic3
        defrouter not specified
```

3. Konfiguration des Interfaces in der Zone

Die Konfiguration des Interfaces in der Zone erfolgt wie die von physischen Interfaces. Im Gegensatz zu VirtualBox sind die VNIC auch in den Zonen unter ihrem VNIC Namen bekannt. Dadurch ist also auch eine eingängige Benennung von VNIC möglich.

```
keetonga #  ifconfig vnic3 plumb

keetonga# ifconfig vnic3 192.168.100.3 up

keetonga # ifconfig vnic3
vnic3: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu
9000 index 2  inet 192.168.100.3 netmask ffffff00 broadcast
192.168.100.255  ether 2:8:20:2b:f5:8e
```

4. Die Kommunikation kann nun zwischen den beiden VBox und der Zone erfolgen, ohne das ein physisches Interface vorhanden sein muss. Je nach Erfordernis können zusätzlich noch Funktionalitäten zur Ressourcenlimitierung eingesetzt werden.

## Zusammenfassung

Die Funktionalitäten zur Netzwerkvirtualisierung und zum Ressourcenmanagement in Oracle Solaris 11 bieten vielfältige Möglichkeiten beim Aufbau von Netzwerkarchitekturen. Dieses Manuskript erläutert die grundlegenden Funktionsweisen und zeigt an kurzen Beispielen die Konfigurationsmöglichkeiten auf. Den konkreten Anwendungsmöglichkeiten sind jedoch kaum Grenzen gesetzt. Weitere vielfältige Einsatzmöglichkeiten sind z.B.:

- Konsolidierung von Architekturen verschiedener Netze in einem System
- Konsolidierung von Multi-Tier Architekturen
- Erzeugung von virtuellen Netzwerkgeräten
  - Firewall in einer Zone
  - Router in einer Zone
  - Loadbalancer in einer Zone
- Testaufbau von Netzwerkarchitekturen vor dem Roll-out in der Praxis
- Untersuchung und Analyse von Architekturproblemen in Netzwerken durch Abbildung der Strukturen in Software und Nutzung der Analysefunktionalitäten
- Ausbildungsplattform für Netzwerkarchitekturen

## Literaturverzeichnis

[1]S. Tripathi, N. Droux, K. Belgaeid, S. Khare: Crossbow Virtual Wire: Network in a Box, Usenix LISA 09, http://www.usenix.org/events/lisa09/tech/full_papers/tripathi.pdf

[2] OpenSolaris Projekt Crossbow: http://hub.opensolaris.org/bin/view/Project+crossbow/

[3] http://www.virtualbox.org

[4] D. Drewanz, U. Gräf: Solaris 10 Container Leitfaden - Version 3.1, 2009

## Kontaktadresse:

**Detlef Drewanz**
Oracle Deutschland B.V. & Co. KG
Komturstraße 18a
D-12099 Berlin

Telefon: +49 (0) 30-747096 856
Fax: +49 (0) 30-747096 878
E-Mail Detlef.Drewanz@oracle.com
Internet: www.oracle.de