



ulm university universität
uulm



Virtuelle Welten Netzwerk Virtualisierung in Solaris 11

Thomas Nau, kiz (Thomas.Nau@uni-ulm.de)

Kommunikations- und Informationszentrum (kiz)

- Die Aufgaben der "Abteilung Infrastruktur" umfassen
 - Cluster basierende universitätsweite Mail-, LDAP-, Portal-, Datenbank- und File-Services, ...
 - Betreuung von ca. 600 Desktop und Laptop Arbeitsplätzen
 - 25% Linux, 75% Windows
 - Backup Service für mehrere Universitäten in Baden-Württemberg
 - Landes HPC Cluster mit Schwerpunkt "Theoretische Chemie"
 - 4 lokale Netzwerke plus flächendeckendes Campus WLAN und MAN im Ulmer Stadtbereich
 - Telefonanlage mit ca. 14.000 Anschlüssen unter Einsatz von VoIP und 2-Draht Technik
 - Azubi Ausbildung

Es geht um die Basics

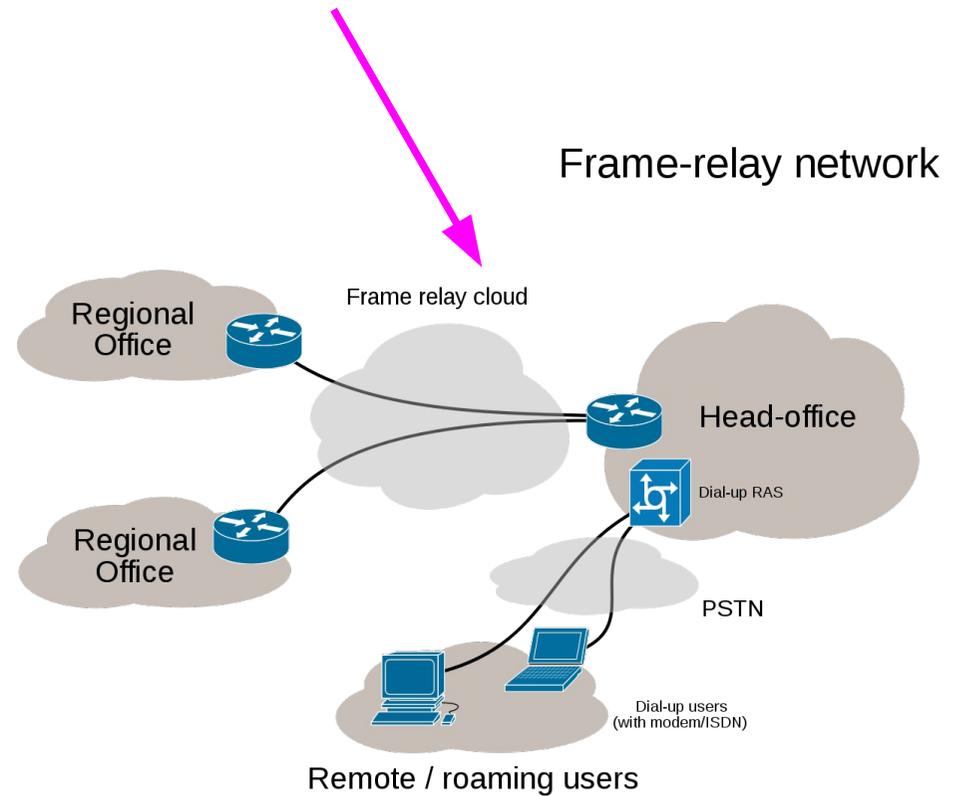
**Netzwerk Virtualisierung
umfasst viel mehr**

Solaris 11 kann viel mehr

Clouds existieren schon lange ...



Source: Wikipedia/Flickr, Autor: secretlondon123



Source: Wikipedia, Autor: Ludovic.ferre

Warum Netzwerk Virtualisierung?

- verfügbare lokale aber auch externe Bandbreite steigt
 - 10G Ethernet Standard bei Servern
 - derzeitiger sweet spot im Backbone noch bei 10GE
 - macht QoS notwendig
 - Uni Ulm WAN Anbindung bald 100G
- Konvergenz von HPC-, Daten- und Speicher Netzen
 - Ethernet everywhere iSCSI, FcoE; aber auch Infiniband iSER, SRP
 - neue IEEE Protokolle etwa *Data Center Bridging* (DCB)
- Systeme werden seit Jahren immer leistungsfähiger
 - Server Virtualisierung
 - hunderte von cores und multi-Terabyte Speicherausbau

Warum Netzwerk Virtualisierung?

- potentiell hunderte von Gast-Systemen stellen neue Herausforderung dar
 - Stichwort "multi-tenant" Fähigkeit
- Total Cost of Ownership (TCO)
 - bessere Auslastung **aller** vorhandenen Ressourcen
- Weg in die Cloud
 - "To boldly go where no one has gone before": überschreiten der nächsten Grenze nach "System", "Raum", "Standort", ...
- essentielle Voraussetzung
 - integriertes Konfigurations- und Ressourcen Management für den gesamten Stack einschließlich Netz- und Storage

Pros und Cons von Solaris 11

Pro

- "leichtgewichtige" Virtualisierung mit Solaris Zonen
 - **ein** Kernel
- einfaches Ressourcen- und Konfigurations-Management
- Virtualisierung war und ist Schwerpunkt bei Planung bzw. Weiterentwicklung des Netzwerkstacks in Solaris 11 (Crossbow)

Cons

- nur Solaris
- nur ein Solaris Kernel (derzeit)

Stand der Netzwerktechnik in Solaris 11.1

- Virtualisierung
 - für die Anwendung transparente virtuelle Interfaces
 - VNICs
 - Etherstubs
 - pseudo Geräte auf deren Basis sich auch VNICs erzeugen lassen
 - Grundlage der internen virtuellen Netzwerk Infrastruktur
 - "network in a box"
 - virtuelle Router, Switches, load-balancer, firewalls
- Skalierung
 - parallele und optimierte Nutzung von NIC und CPU Ressourcen
 - Wechsel zwischen polling und Interrupt
 - Hardware off-loading für Prüfsummen, Crypto, ...

Stand der Netzwerktechnik in Solaris 11.1

- Ressourcen Kontrolle und Management
 - Zuweisung von CPUs, NICs bzw. deren processing units
 - QoS auf Basis von flows
- Monitoring
 - DTrace provider für IP, TCP, UDP, SRP, iSCSI, NFSv3, NFSv4
 - real-time Daten für VNICs und flows
 - erweitertes *kstat(1m)* Interface
- Erweiterte und verbesserte Funktionalität
 - *Virtual Router Redundancy Protocol* (VRRP)
 - sniffing auch auf lokalen Interfaces, VNICs, ...
 - freie Wahl der Interface Namen

Stand der Netzwerktechnik in Solaris 11.1

- bye-bye *ifconfig(1m)*

- *dladm(1m)*, *dlstat(1m)*

Verwaltung von data-links also NICs, virtuelle NICs (VNICs), VLANs, link- aggregations, virtuellen bridges und etherstubs

Statistik Daten von data-links, etwa übertragene Bytes

- *ipadm(1m)*

konfiguriert IP Interfaces sowie TCP/IP Parameter einschließlich IP multipathing

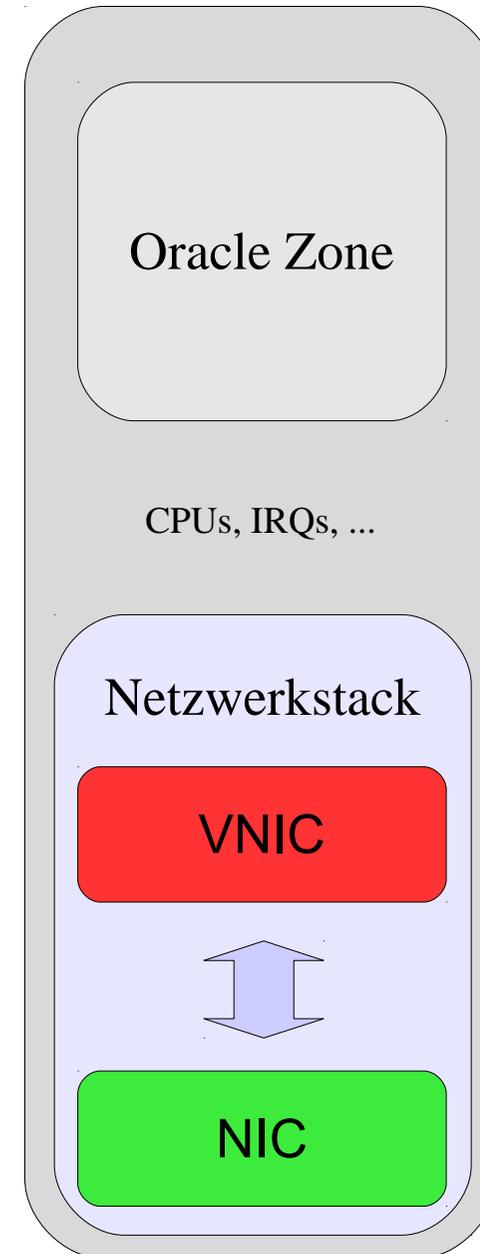
- *flowadm(1m)*, *flowstat(1m)*

Bandbreiten Verwaltung auf Basis von flows – definiert durch Attribute der Layer 3 und 4 sowie von Zonen

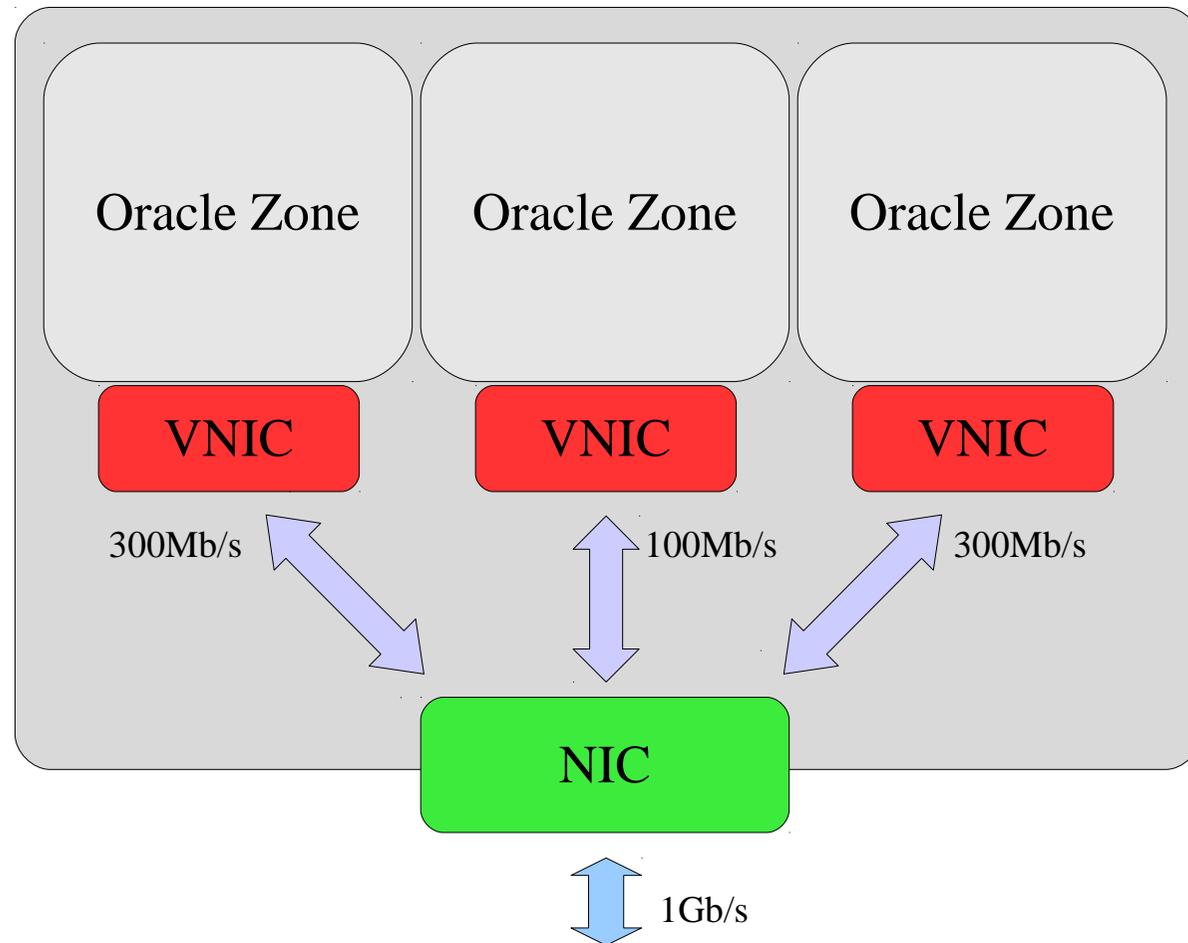
Statistik Daten von flows, etwa übertragene Bytes

Baukasten System

- Solaris Zone
 - VNIC
 - Memory
 - CPUs
 - ZFS
- Netzwerk Stack
 - flows
 - QoS
 - NIC / Etherstub

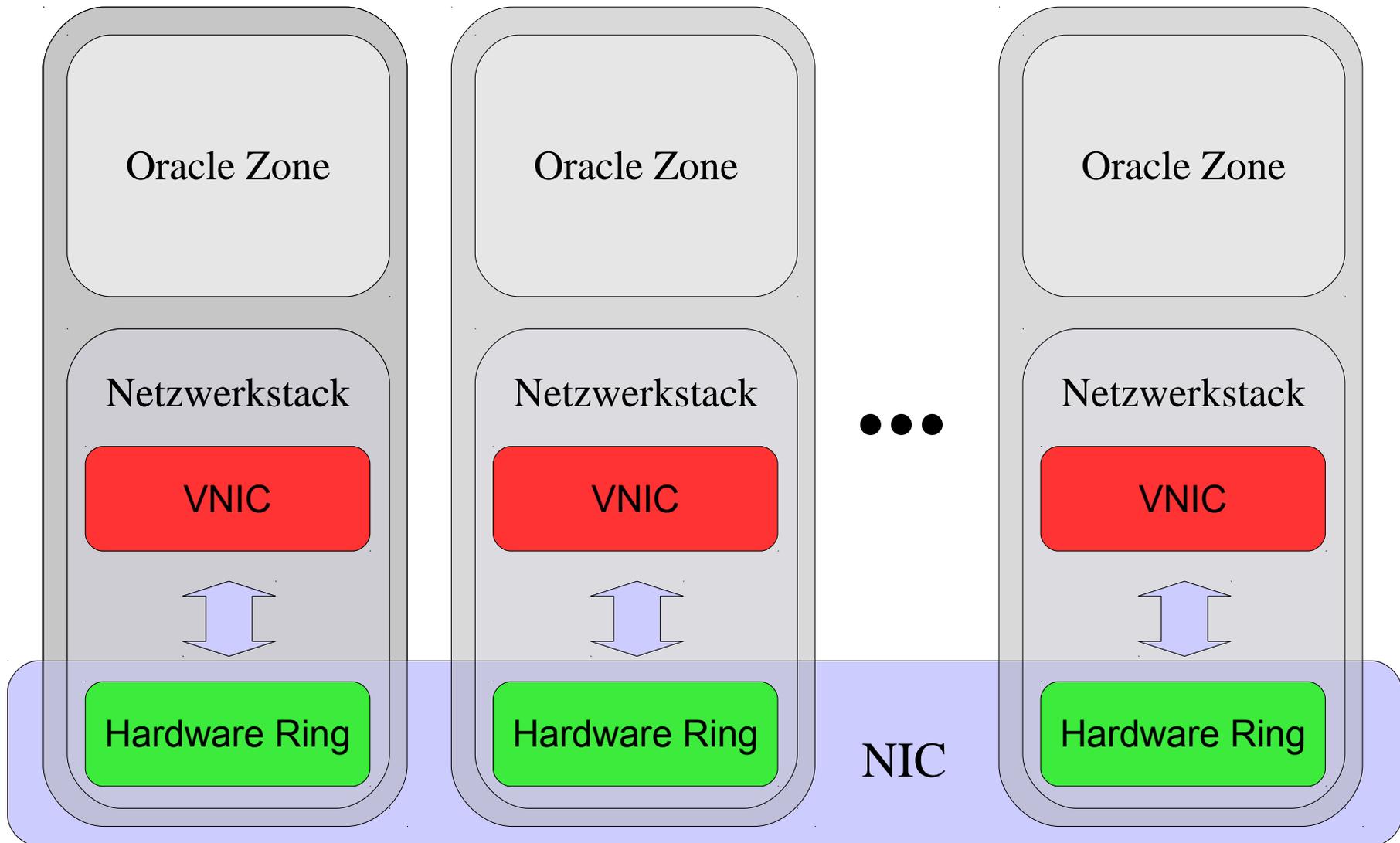


Baukasten System – Ressourcen



```
jedi# dladm create-vnic -l net0 -p maxbw=300M myvnic0
```

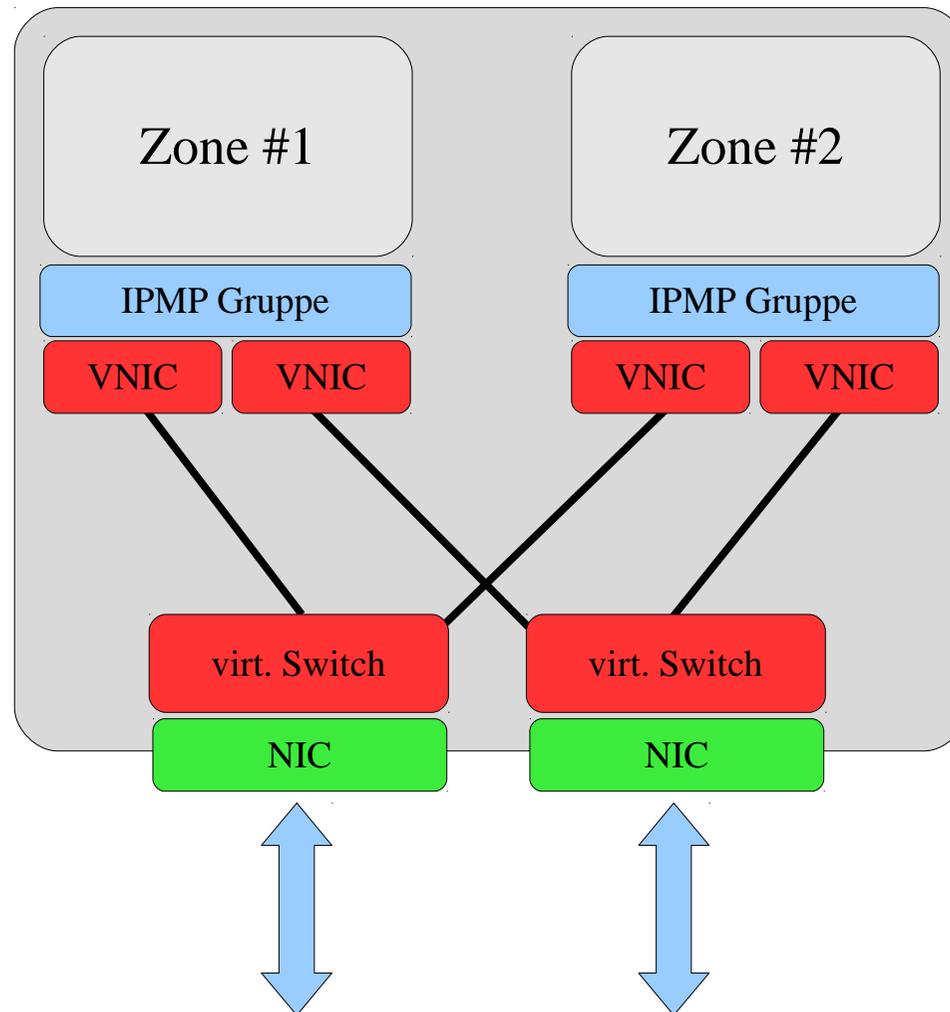
Solaris 11 Netzwerk Virtualisierungs-Infrastruktur



Redundanzen: IP Multipathing (IPMP)

Pro	Contra
failover über mehrere Switches hinweg ohne Nutzung proprietärer Protokolle	erfordert individuelle Konfiguration für jede Zone oder virtuelle Maschine
Switch Konfiguration nicht notwendig	

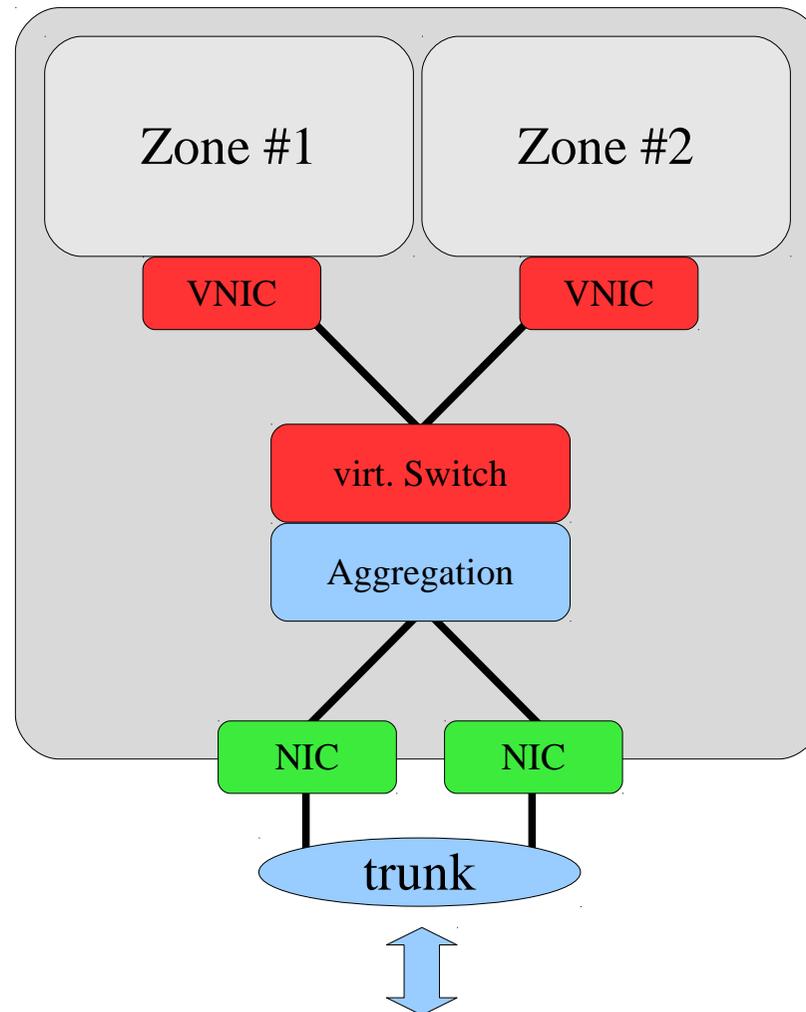
Redundanzen: IP Multipathing (IPMP)



Redundanzen: Link Aggregation / Trunking

Pro	Contra
transparent für Zonen und virtuelle Maschinen; einfache Administration	erfordert spezielle Konfiguration der Switches
erhöht die verfügbare Bandbreite	Beschränkt die Konnektivität auf einen einzelnen Switch oder verlangt proprietäre Protokolle
automatisches failover/fallback	alle verwendeten Interfaces müssen identische Duplex Modi und Geschwindigkeiten haben

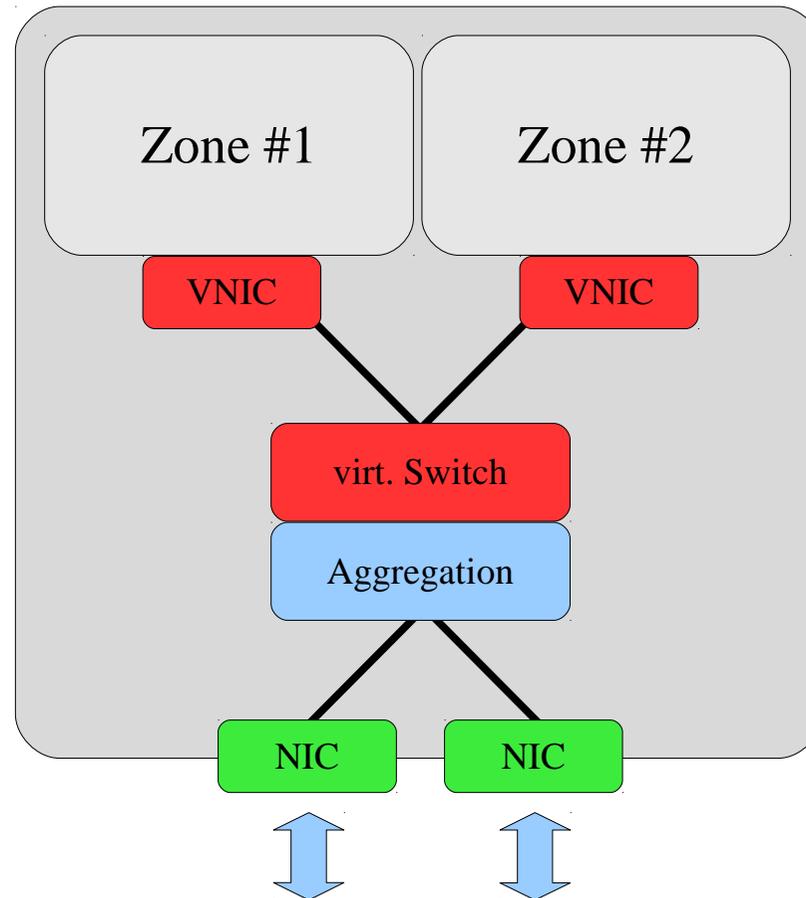
Redundanzen: Link Aggregation / Trunking



Redundanzen: Data-link Multipathing (DLMP)

- neu in Solaris 11.1
- das beste aus beiden Welten für viele Anwendungen
 - Hochverfügbarkeit ohne proprietäre Protokolle auch über Switch Grenzen hinweg
 - derzeit nur begrenztes load-balancing durch Verteilung der VNICs
- fail-over auf Basis des Link Status
 - keine aktive Überprüfung wie bei IPMP
 - Raum für Verbesserungen in Solaris 11.x

Redundanzen: Data-link Multipathing (DLMP)



Beispiel: Virtualisierung eines Rechenzentrums

- Ziel Virtualisierung einer vorhandenen Umgebung aus
 - Web-Servern
 - Mail-Servern
- Abschottung in mehrere Bereiche
- Trennung von Netzwerken
 - Internet Verkehr
 - Storage
 - Management

Beispiel: Virtualisierung eines Rechenzentrums



Beispiel: Virtualisierung eines Rechenzentrums

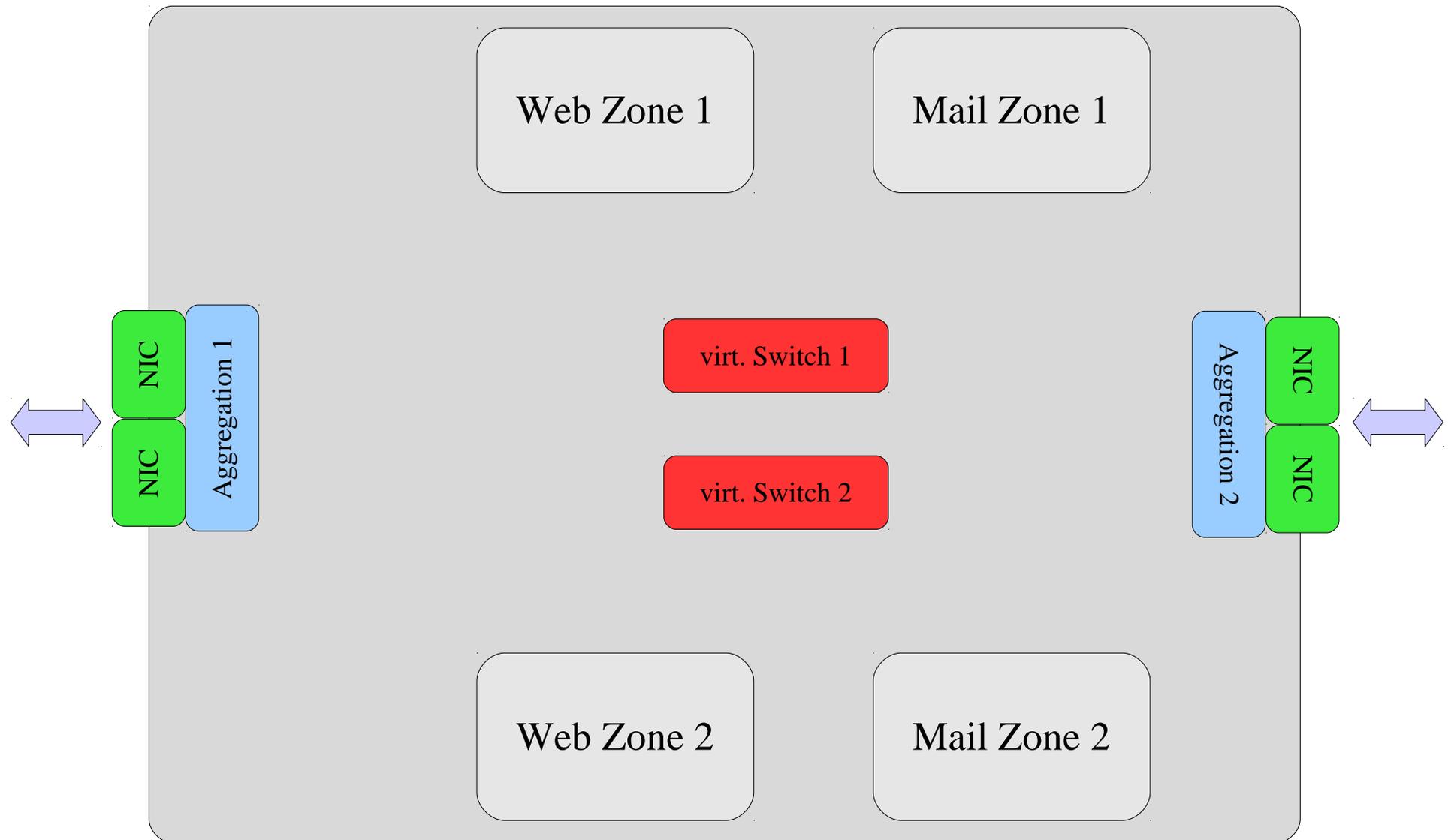
- DLMP aggregations anlegen

```
# dladm create-aggr -m dlmp -l net2 -l net3 aggr1  
# dladm create-aggr -m dlmp -l net4 -l net5 aggr2
```

- virtuelle Switches (etherstubs) anlegen

```
# dladm create-etherstub vsw1  
# dladm create-etherstub vsw2
```

Beispiel: Virtualisierung eines Rechenzentrums



Beispiel: Virtualisierung eines Rechenzentrums

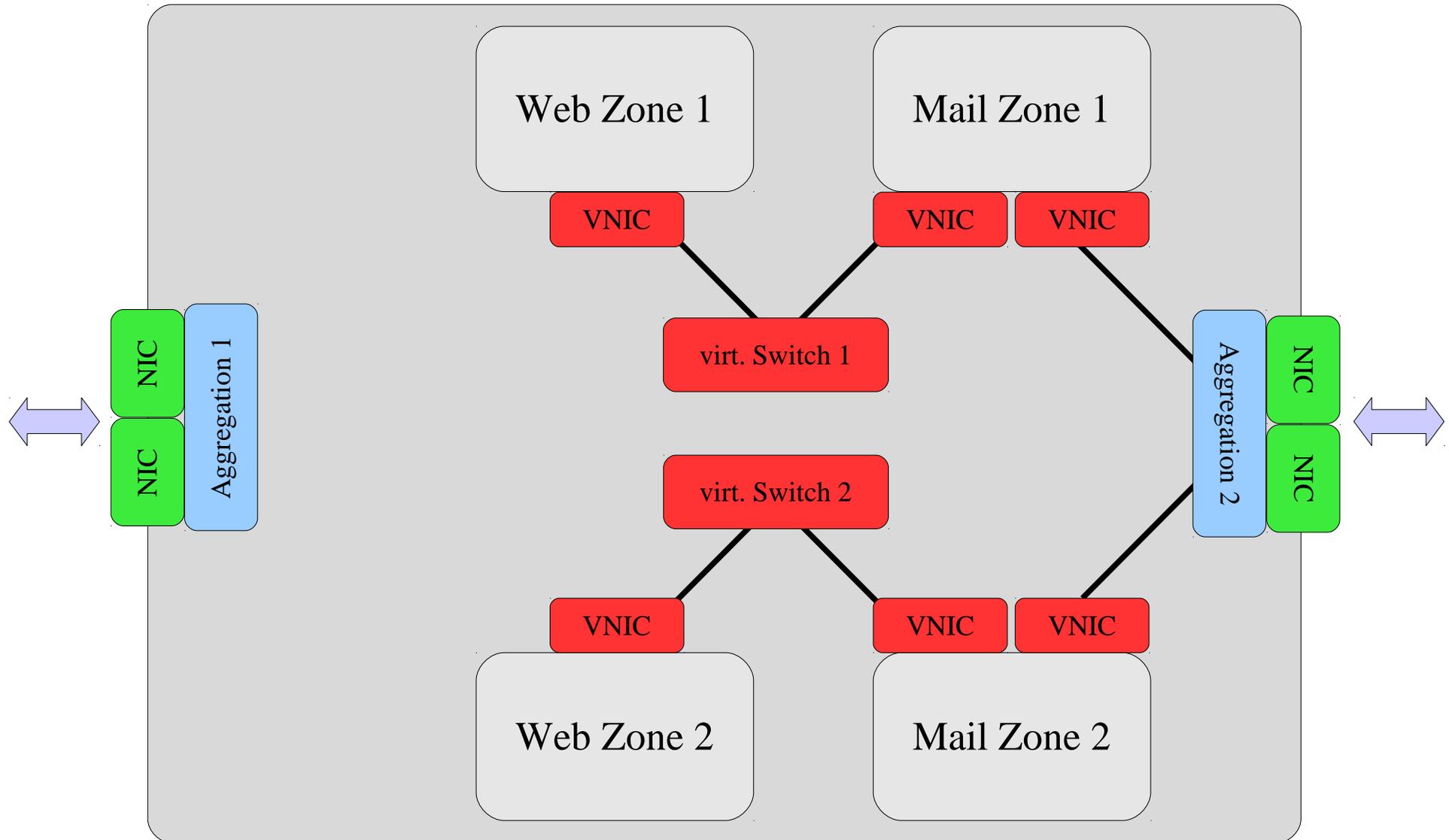
- VNICs anlegen und mit Switches verbinden

```
dladm create-vnic -l vsw1 vnic_wz1_0
dladm create-vnic -l vsw1 vnic_mz1_0
dladm create-vnic -l vsw2 vnic_wz2_0
dladm create-vnic -l vsw2 vnic_mz2_0
```

- VNICs für weitere Netze anlegen

```
# dladm create-vnic -l aggr2 vnic_mz1_1
# dladm create-vnic -l aggr2 vnic_mz2_1
```

Beispiel: Virtualisierung eines Rechenzentrums

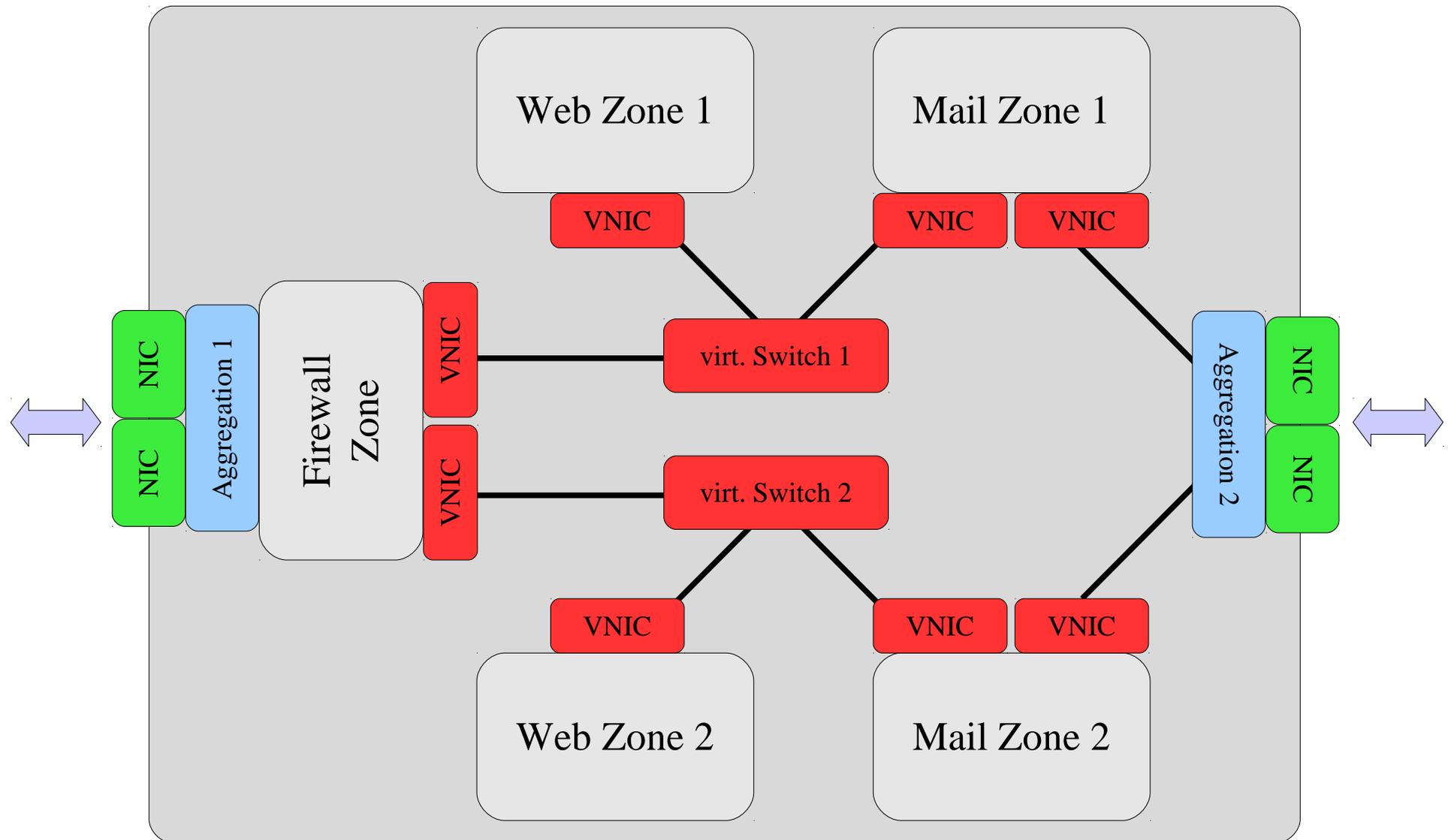


Beispiel: Virtualisierung eines Rechenzentrums

- Firewall-Zone einfügen

```
# dladm create-vnic -l vsw1 vnic_fw_0  
# dladm create-vnic -l vsw2 vnic_fw_1
```

Endprodukt: virtuelles Rechenzentrum



Das Netzwerk aus Kommandozeilen Sicht

```
obi-wan# dladm
LINK          CLASS      MTU      STATE    OVER
net0         phys      1500    up       --
net1         phys      1500    down     --
net2         phys      1500    down     --
net3         phys      1500    down     --
net4         phys      1500    down     --
net5         phys      1500    down     --
vsw1         etherstub 9000    unknown --
vsw2         etherstub 9000    unknown --
aggr1        aggr      1500    unknown net2 net3
aggr2        aggr      1500    unknown net5 net4
vnic_wz1_0   vnic      9000    up       vsw1
vnic_mz1_0   vnic      9000    up       vsw1
vnic_wz2_0   vnic      9000    up       vsw2
vnic_mz2_0   vnic      9000    up       vsw2
vnic_fw_0    vnic      9000    up       vsw1
vnic_fw_1    vnic      9000    up       vsw2
```

Weitere technische Trends

- Einsatz von Infiniband (IB) für Storage Netzwerke
 - geringere Latenz durch DMA fähige Transportmechanismen
 - 2x Bandbreite im Vergleich zu Fibre-Channel und günstigere Komponenten
 - hervorragend in Solaris eingebunden
- Virtual extensible LAN (VXLAN)
 - overlay Netzwerk
 - tunnelt Layer 2 über Layer 3 Protokoll (IP)
 - "scale Ethernet to the Cloud"
 - nicht Bestandteil von Solaris 11.1

Die Zukunft

- Transparente Verflechtung von Virtualisierung und Netzwerk Management- und Konfigurations-Tools
 - erlaubt Umsetzung von service level agreements (SLA)
- *Software Defined Networks (SDN)*
- IEEE Standard *Edge Virtual Bridging (EVB)*
 - regelt Zusammenspiel zwischen Hypervisor und physikalischer Netzwerk Hardware
- *Data Center Bridging (DCB)*
 - stellt isolierte Pfade auf Ethernet Basis zur Verfügung
 - gekennzeichnet durch Bandbreite, Flusskontrolle, ...
 - verhindert Überlastung sensibler Pfade (storage)

Danke für's Zuhören