



# **OSL Storage Cluster 4.1**

## **Virtualization - Clustering**

OSL aktuell - Schöneiche, April 2015

## Das Prinzip:

- Sehen - Verstehen – Verbessern !
- Konsequenz systematisieren !
- Auf Dauer und erfolgreich nur mit eigener Technologie !

Wir sind mit dieser Sichtweise nicht allein!

## Das Prinzip:

- Sehen - Verstehen – Verbessern !
- Konsequenz systematisieren !
- Auf Dauer und erfolgreich nur mit eigener Technologie !

Wir sind mit dieser Sichtweise nicht allein!

## Die Idee:

- helle Köpfe zusammenführen
- Konzeptionen entwickeln
- Lösungen bauen

# OSL Storage Cluster

Was das Produkt ausmacht



**Symmetric Cluster Engine**

**Global  
Virtual  
Storage**

**Unified  
Virtual  
Networking**

**Application  
Virtual  
Environment**

**Unified  
Virtual  
Machines**

**Adaptive  
Resource  
Control / HA**

**Cross-Platform Capabilities**

# OSL Storage Cluster

Was das Produkt ausmacht



Symmetric Cluster Engine

Global  
Virtual  
Storage

Unified  
Virtual  
Networking

**It's just software!**

Application  
Virtual  
Environment

Unified  
Virtual  
Machines

Adaptive  
Resource  
Control / HA

Cross-Platform Capabilities

# Speichervirtualisierung

hostbasiert – clusterfähig - schnell

# Speichervirtualisierung

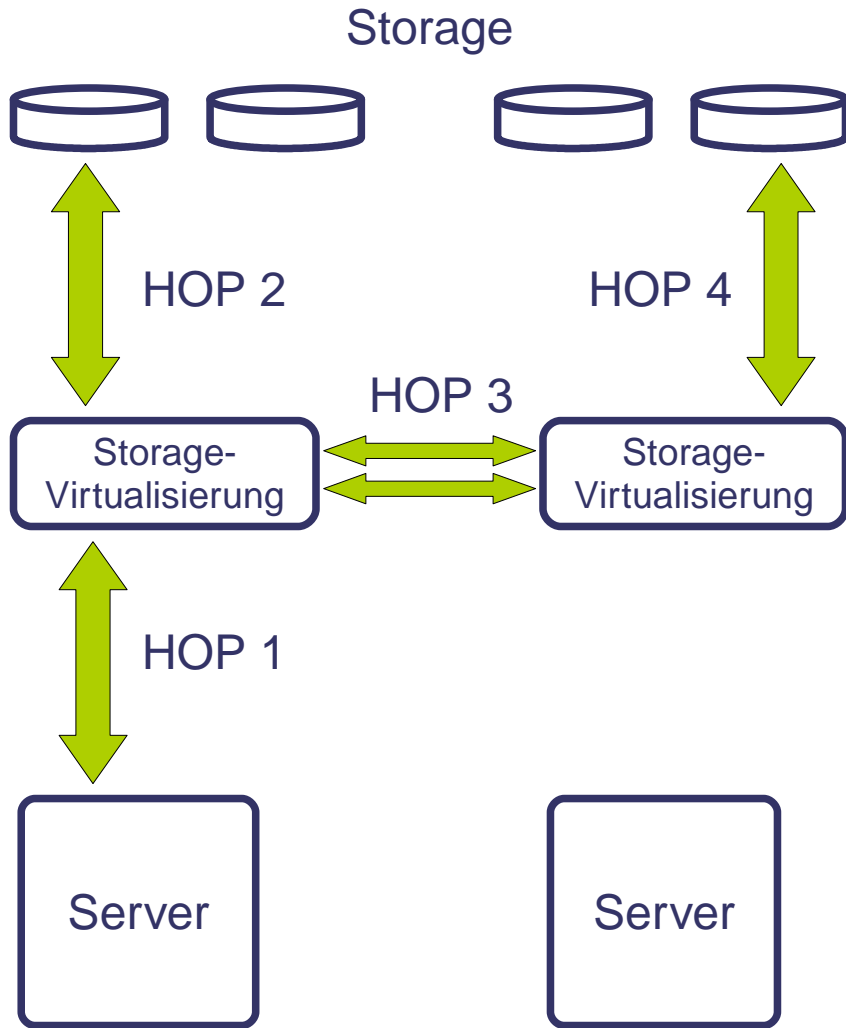
Worum geht es bei der Speichervirtualisierung?



- Hardwareabstraktion
- Verschieben physikalischer oder logischer Limits (Größen, Volumezahl)
- verbesserte I/O-Performance
- verbesserte Verfügbarkeit
- Herstellen einer Eignung für Disaster Recovery
- Daten online verschieben und reorganisieren
- Zusatzfunktionen wie
  - Bandbreitensteuerung
  - Backup to Disk
  - permanent Backup u./o. Snapshots
- entscheidend: intelligenter Cluster-Support (s. folgende Folien)

## Was davon kann Ihre Lösung?

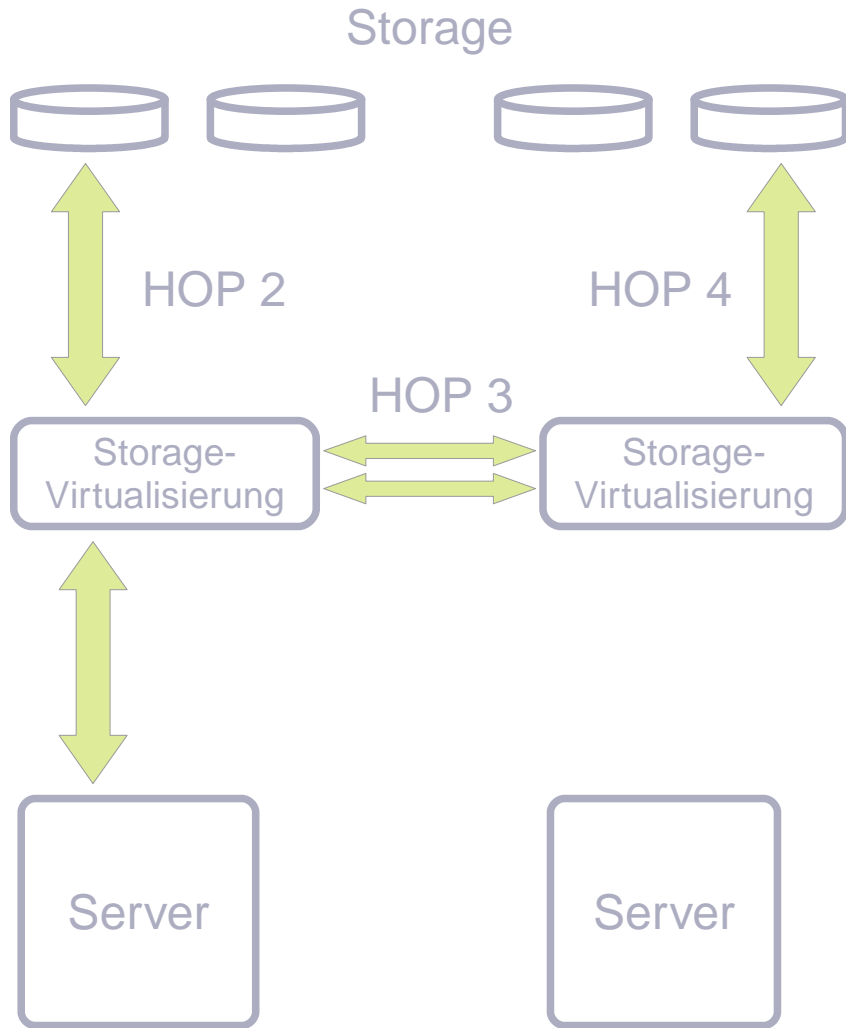
# Speichervirtualisierung - so lösen es andere...



Cluster ?



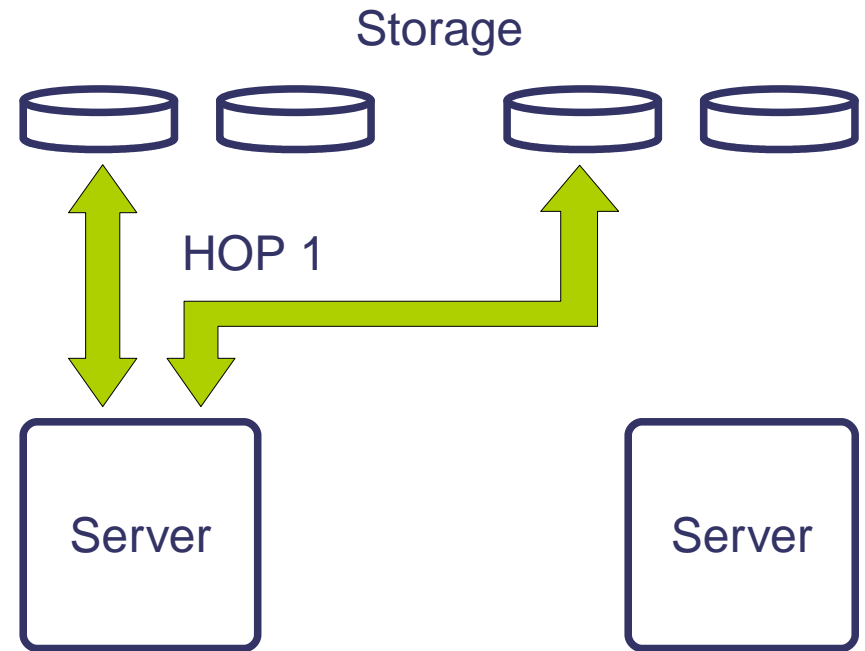
# Speichervirtualisierung - so lösen es andere...



Cluster ?

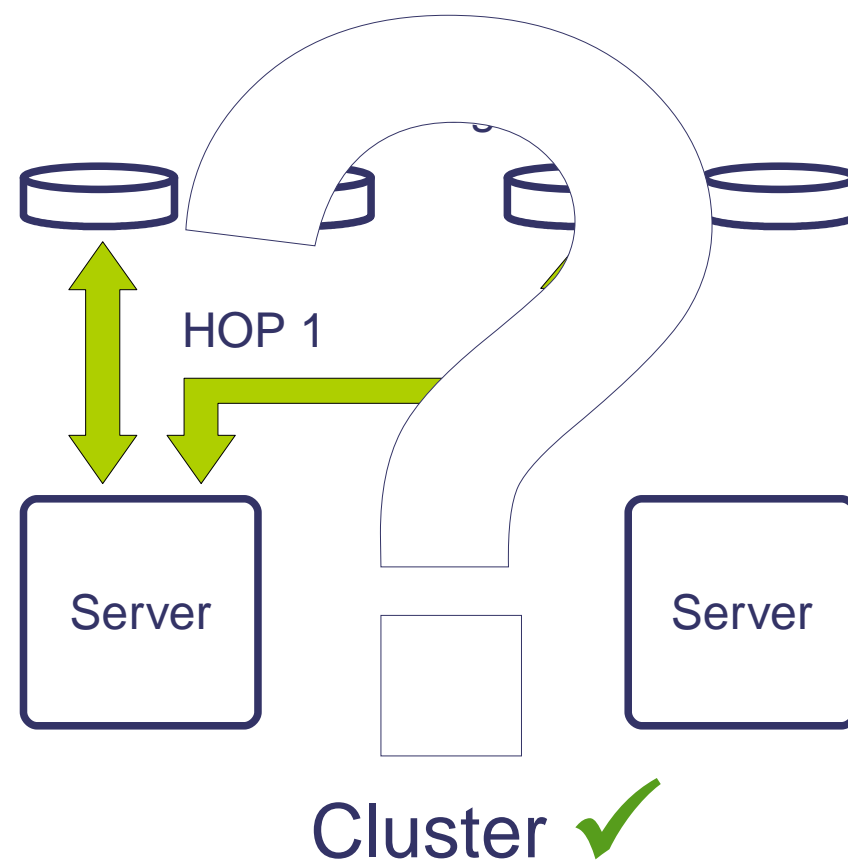
und ...

so löst es OSL:



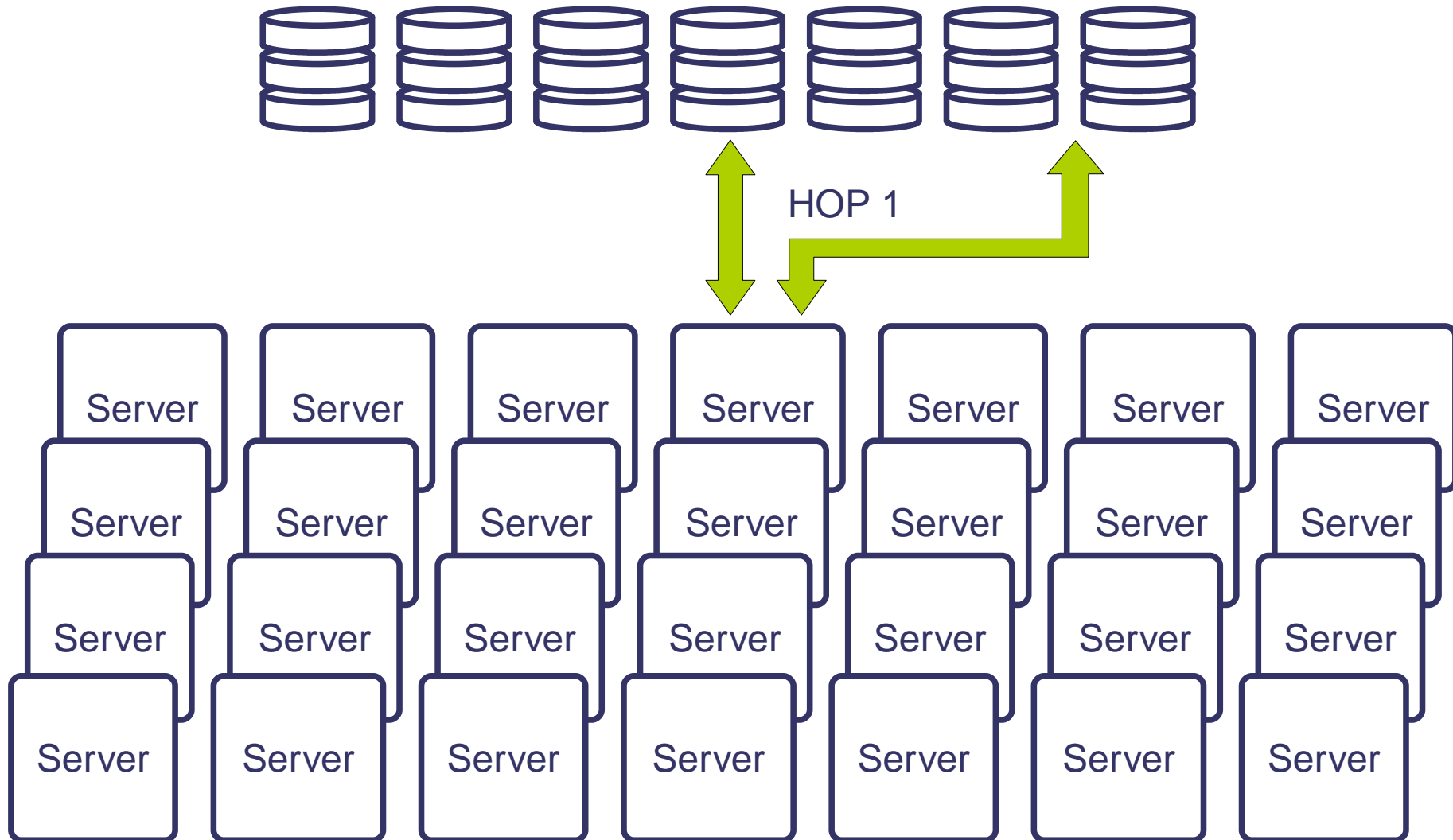
Cluster ✓

so löst es OSL:



# Was heißt Cluster im Kontext Speichervirtualisierung?

Massive Skalierbarkeit: Storage, SAN, Network, Throughput, Availability, Nodes

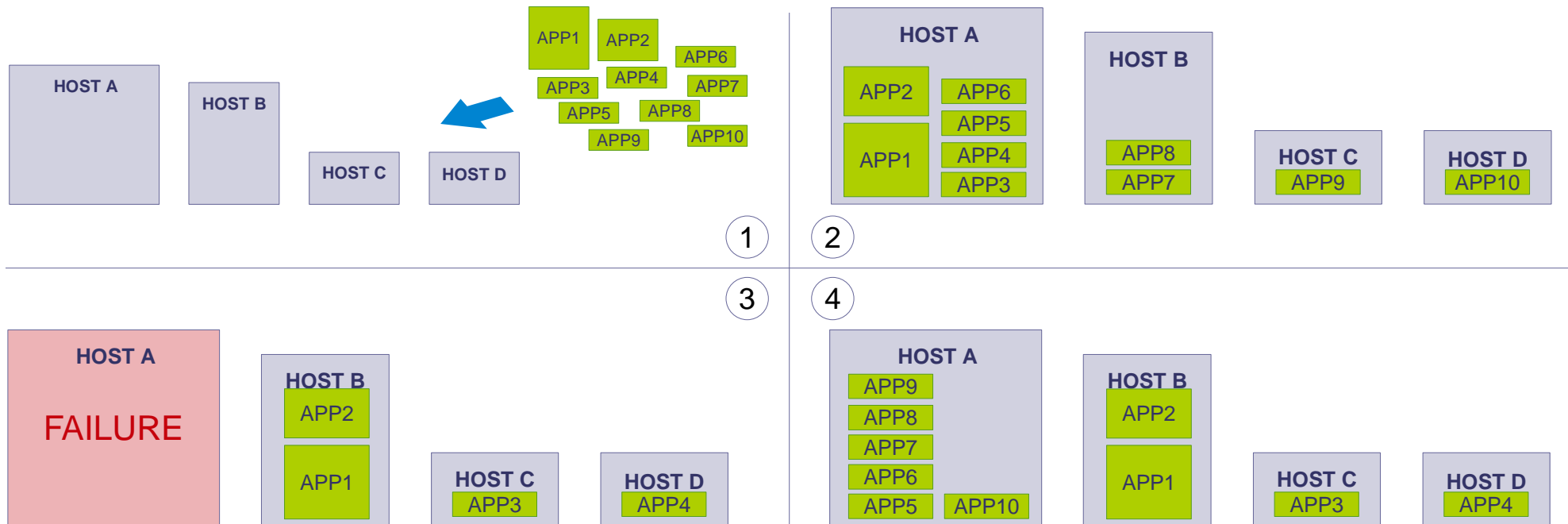


# Was der Cluster auch noch kann ...

Hochverfügbarkeit schon eingebaut



- selbstorganisierend – Berücksichtigung / Steuerung Ressourcen
- vollkommen symmetrisch
- herausragende Robustheit - kein Split Brain
- zentrale Administration von jedem Knoten aus
- Cross-Platform
- verknüpft mit Speichervirtualisierung (Applikationsbewußtsein, Backup & DR ...)

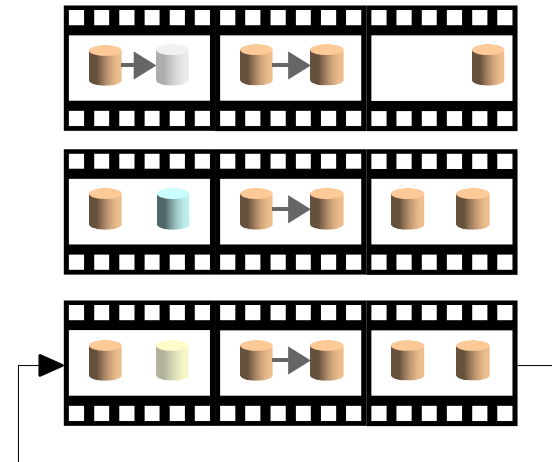
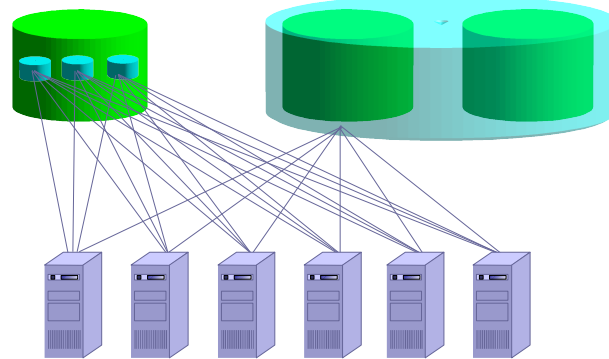


# Speichervirtualisierung als Software

Möglichkeiten – Freiräume - Unabhängigkeit



Speichervirtualisierung
clusterweit
globale Pools
Daten verschieben
Daten klonen
Daten spiegeln
Sonderfunktionen



Physical Volumes + Application Volumes  
linear oder integriert (simple, concat, stripe)  
Hardwareabstraktion und IO-Multipathing  
systemgestützte Speicherallokation  
Online-Konfig./Dekonfig./Vergrößerung

globale Geräte / globaler Namesraum  
vollautomatisiertes Zugriffsmanagement

globale Pools (hostübergreifend)  
globales Inventory (Verzeichnis)  
kein Verschnitt von Kapazitäten

Daten online verschieben / reorganisieren  
minimaler Einfluß auf laufenden Applikations-I/O

Online-Datenkopien auf wahlfreie Ziele  
atomare Operationen für mehrere Volumes

permanente Master-Image-Beziehungen  
mehrere Images + OSL-Universen  
inkrementelle Resynchronisation  
Überbrückung von Fehlern auf dem Master

XVC (Extended Volume Controls)  
z.B. Pause, Stop, Trigger, Aktionen  
Bandbreitensteuerung  
detaillierte Statistik

# Nutzen für RZ-Infrastrukturen

Vorteile mit der Speichervirtualisierung des OSL Storage Clusters u. a.:



- Speichervirtualisierung und Cluster in Einem -> überlegenes Design
- globaler Storage-Pool – Enterprise Storage Directory
- enorm vereinfachtes Gerätehandling
  - globale Geräte / globaler Namensraum / wahlfreie Gerätenamen
  - alle Anschlußtechniken von SCSI und iSCSI über FC und Infiniband nach gleichem Schema
  - integriertes, leicht verständliches Multipathing, Dynamic Hardware Reconfiguration Capabilities
  - identische Handhabung von Solaris 7 bis Solaris 11, Sparc, x86 und sogar Linux
  - **steht auch für Zonen und LDOMs zur Verfügung**
- automatisiertes Disk-Access-Management
  - allgemein enormer Sicherheitsgewinn in Clusterumgebungen bei NULL-Administration
  - ZFS sicher auf Shared Storage / im Cluster betreiben
- Applikations- / VM-Bewußtsein
  - applikationsbezogene, automatisierte Aktionen (Spiegeln, Backup-to-Disk, Backup-to-Tape, DR)
  - Übersicht Nutzung Storage Pool nach Applikationen / VMs
  - Allokation und Bandbreitensteuerung nach Applikationen
- herausragende Performance und Bandbreitensteuerung
  - keine Appliance, kein Flaschenhals – beliebige Verfügbarkeits- und Durchsatzskalierung
  - Fähigkeit zur Bandbreitensteuerung (per Volume und per Applikation/VM)

# OSL Storage Cluster

Speichervirtualisierung und Clustering in perfekter Symbiose



## Symmetric Cluster Engine

**Global  
Virtual  
Storage**

**Unified  
Virtual  
Networking**

**Application  
Virtual  
Environment**

**Unified  
Virtual  
Machines**

**Adaptive  
Resource  
Control / HA**

# OSL Storage Cluster 4.x – Unified Networking

Eine wesentliche Neuerung



Symmetric Cluster Engine

Global  
Virtual  
Storage

**Unified  
Virtual  
Networking**

Application  
Virtual  
Environment

Unified  
Virtual  
Machines

Adaptive  
Resource  
Control / HA

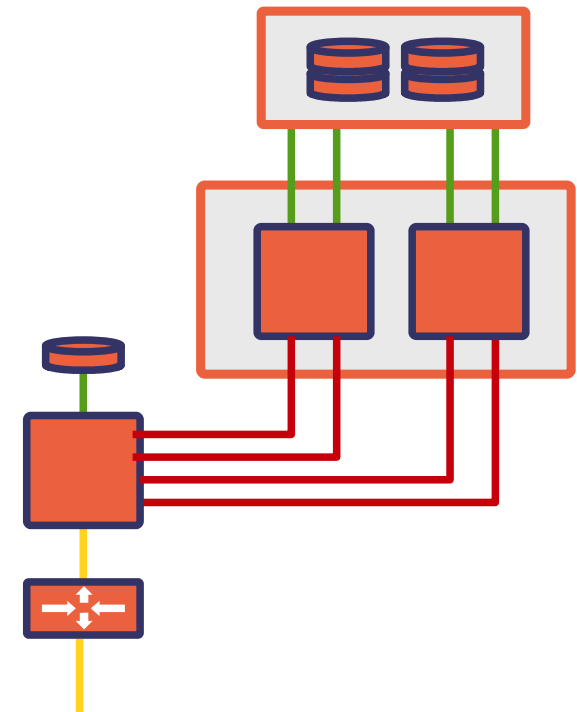
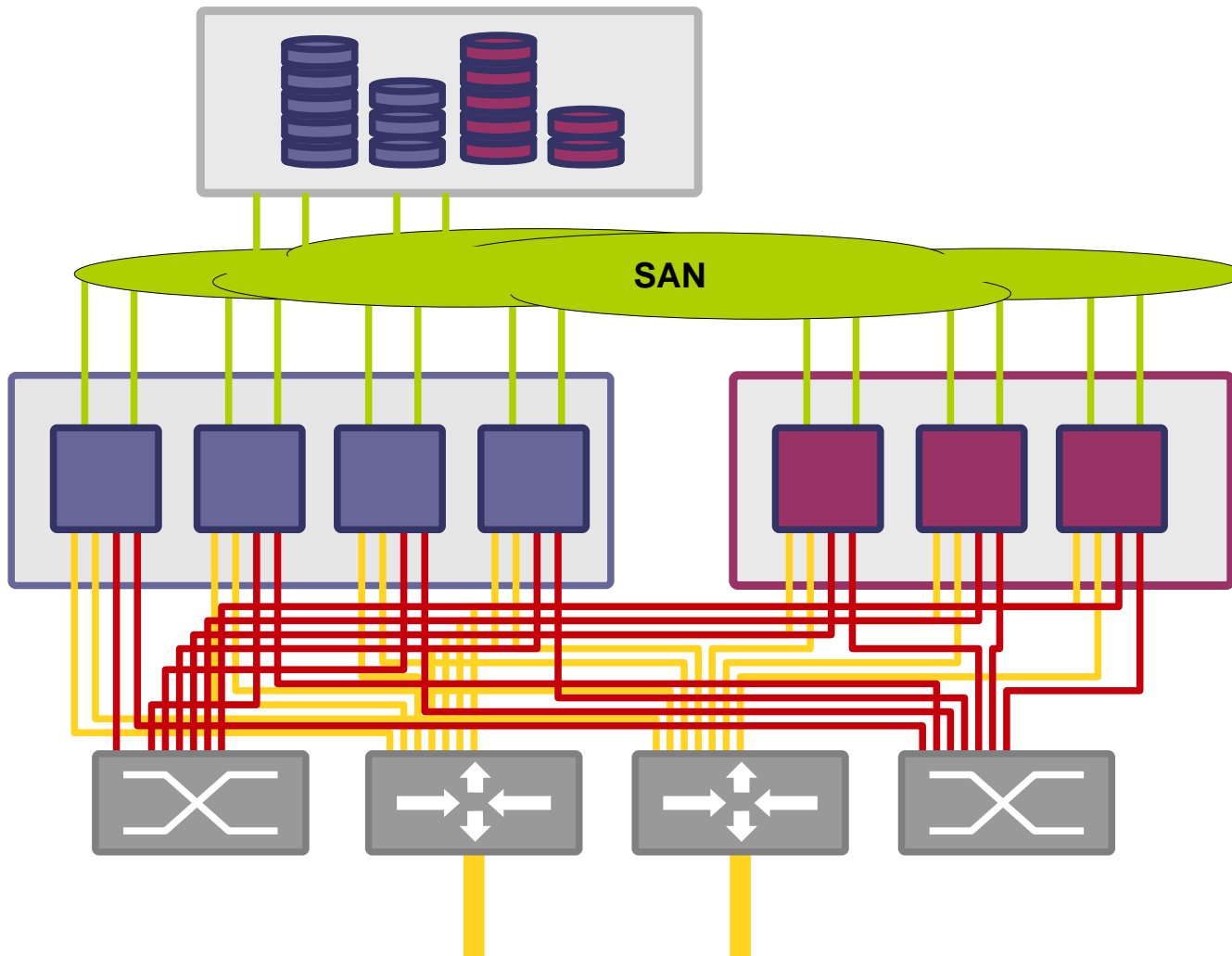
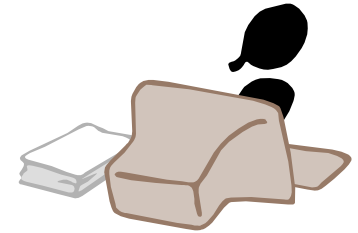


# Unified Networking

einfach – leistungsfähig - günstig

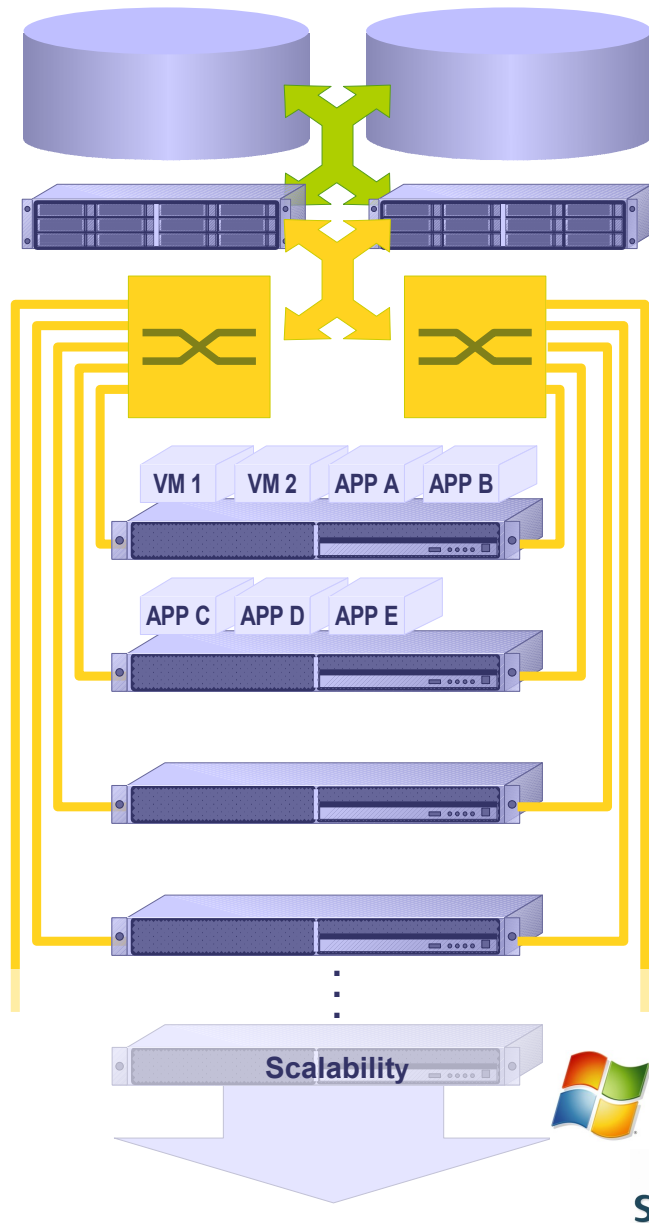
# Eine ganz normale Infrastruktur ...

mit klassischen RZ-Komponenten

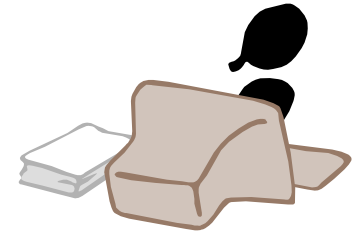


# ... kann auch ganz anders gedacht werden

radikal einfacher – sicherer, flexibler, leistungsfähiger



- Unified Networking
- weniger Komponenten
- niedrigste Kosten
- einfachste Administration
- aufgeräumte Client-Server-Architektur
- vorhandene Komponenten integrierbar
- beeindruckende Skalierbarkeit
- bewährte Technologie
- Mix: Platforms, Storage, Applications, VMs  
**Solaris, Linux, Hypervisors in einem Framework!**



# Unified Networking – die Bausteine

von Hardware über das Protokoll bis zur Administrationssoftware



- Unified Network      Ethernet 1G - 10 G - 40G  
                                 Infiniband
  
- Block-I/O over Network -> RSIO (Remote Storage I/O)
  - modernes Block-I/O-Protokoll mit herausragenden Skalierungs- und Verfügbarkeitseigenschaften
  - massive multithreaded Design
  - Verknüpfung mit modernen Netzarchitekturen und Speichervirtualisierung
  
- Virtual Networking Stack
  - löst Probleme virtualisierter Netzwerkkomponenten auf den Hosts
  - durchdachtes Konzept
  - identische Handhabung unter Solaris und Linux

# OSL Virtual Network Objects

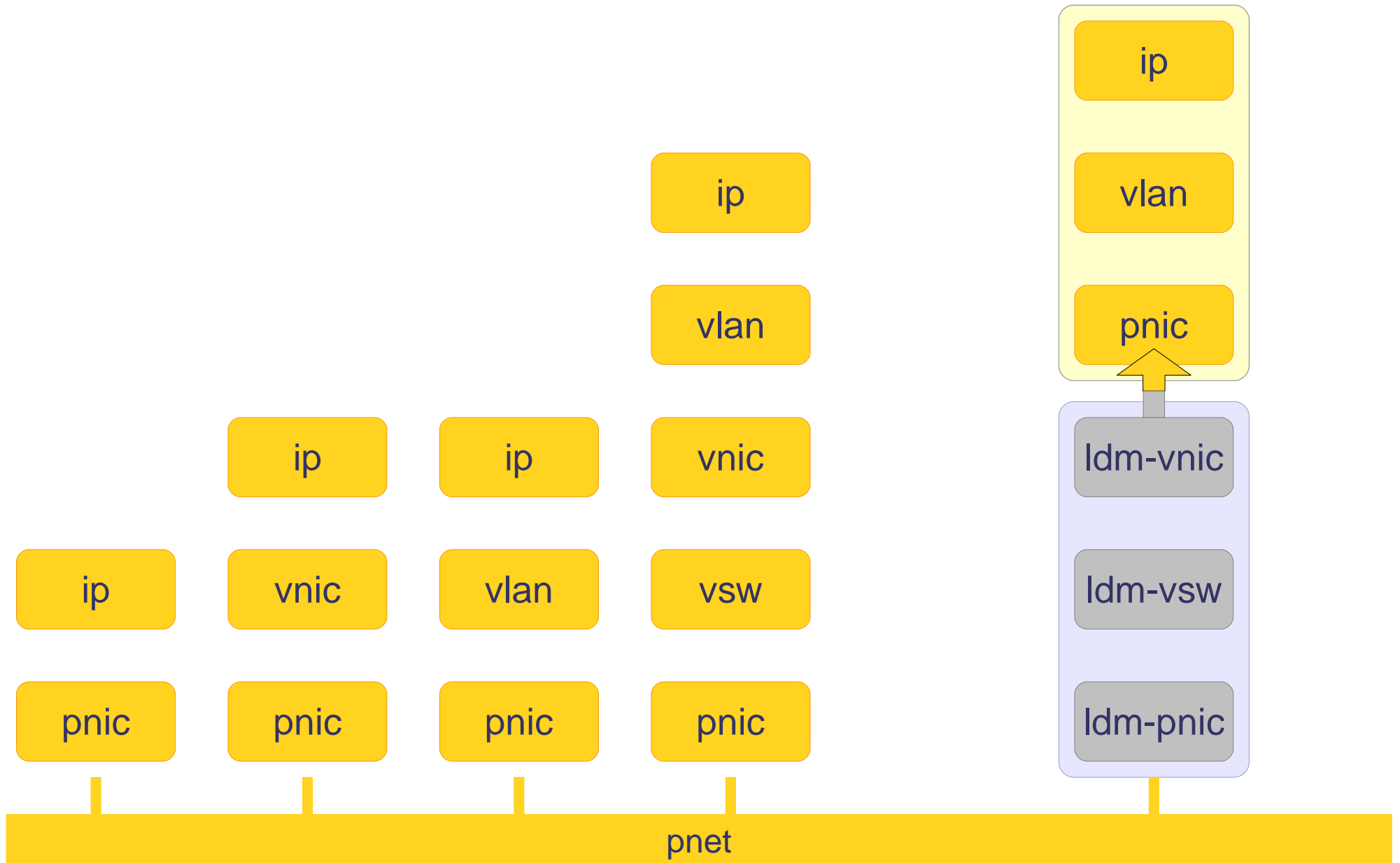
Die Antwort auf die Probleme aus In-Server-Consolidation/Virtualization



- Netzwerk-Setup auf heutigen Servern sehr komplex
  - Vielzahl verschiedener Netzwerk-Objekte, inzwischen bis hin zu Switches
  - Sharing von Controllern und virtuellen Objekten
  - Überlagerung durch VLANs
  - mehrere VMs oder Applikationen
- OSL Virtual Network Objects (OSL VNO)
  - Abstraktionslayer
  - einheitliche Bedienung unter Solaris und Linux, auch unter verschiedenen Versionen
- Jede Applikation / VM kann ihren eigenen Namespace aufbauen
  - Einstiegspunkt ist immer das Physical Network
  - komplett eigener Namespace für die Applikation bis zum obersten Layer (Interface/VLAN/IP)
  - System Network Objects werden bei Bedarf automatisch erzeugt oder entfernt
- OSL Virtual Network Objects ermöglichen applikationsweise Operationen
  - Device Trees
  - Abbau ganzer Trees mit einem Kommando
  - Schutz anderer Konfigurationen
- OSL VNOs werden auch für Systeminterfaces nutzbar sein
- erhebliche Steigerung der Bediensicherheit

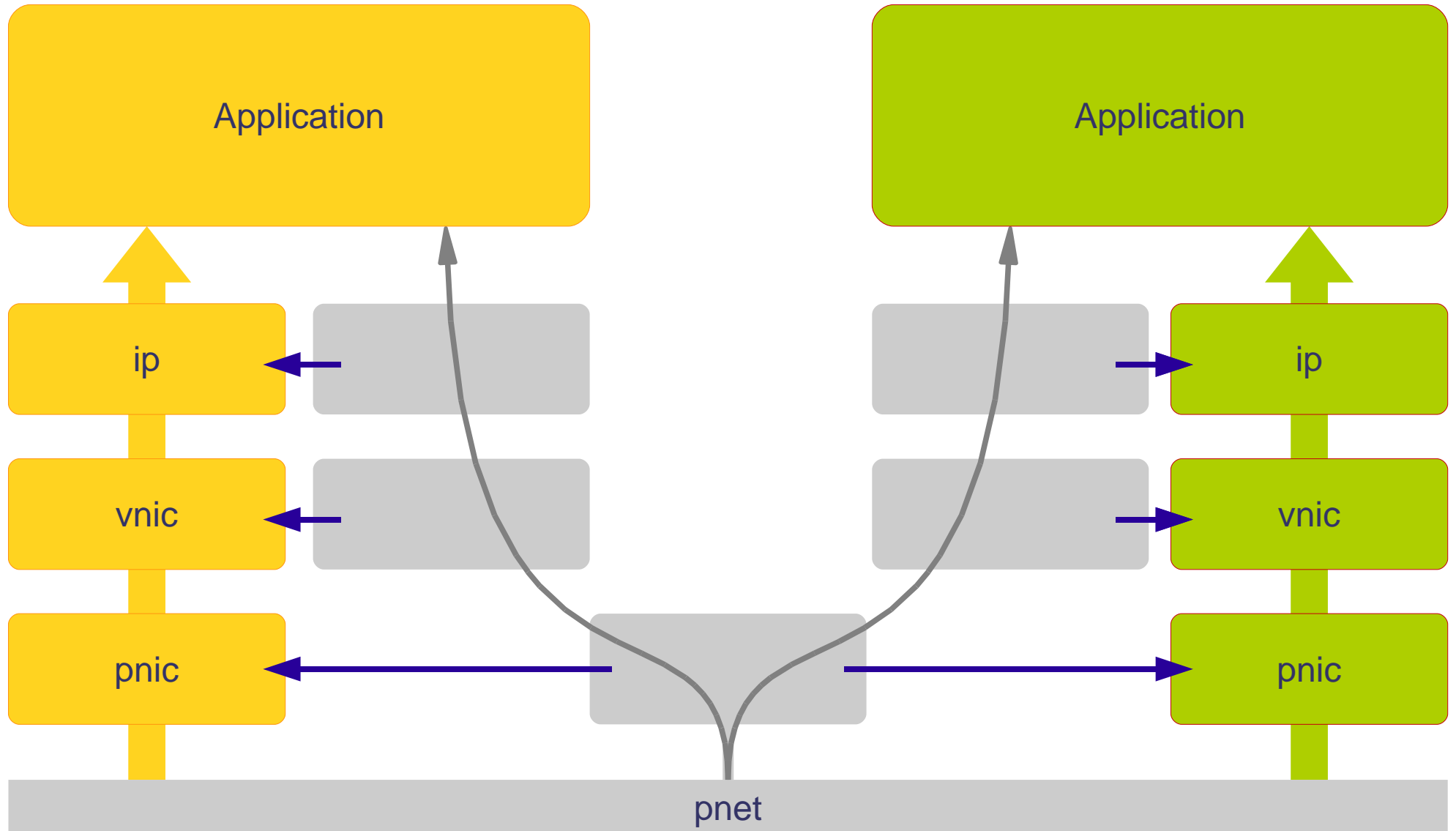
# OSL Virtual Network Objects

Konfigurationsvarianten (Auswahl)



# OSL Virtual Network Objects

Logical to Physical - System Network Object Sharing



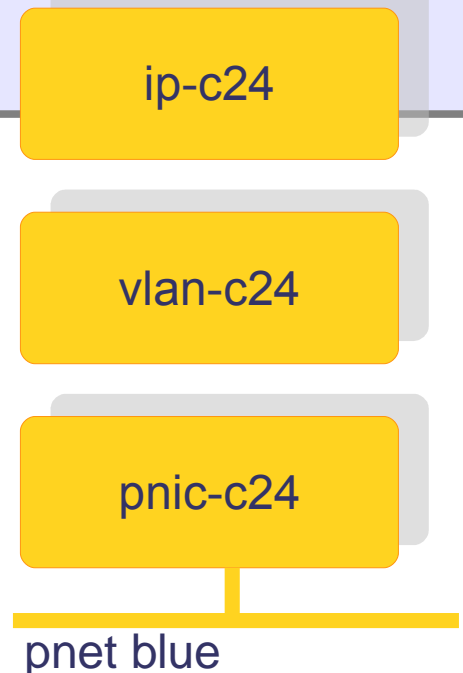
# OSL Virtual Network Objects

Kommandozeile



```
# pnadmin -q
red          eth      1g
blue        eth      1g
# pnadmin -l
e1000g1 at 1g-port of eth net red with 1g capability
e1000g2 at 1g-port of eth net blue with 1g capability
e1000g3 at 0m-port of eth net blue with 1g capability
# vnadmin -c pnic-c24 -t pnic -a c24 -d blue
# vnadmin -c vlan-c24 -t vlan -a c24 -o vid=123 -d pnic-c24
# vnadmin -c ip-c24 -t ip -a c24 -o ip=192.168.124.12/24 -d vlan-c24
# vnadmin -l ip-c24
ip-c24      192.168.124.12/24      on vlan-c24
```

- Aufbau einfach auf dem gewünschten pnet (hier "blue")
- System Network Objects (SNO) werden bei Bedarf erzeugt und ggf. geshared
- Konfiguration in der ARD der Applikation entsprechend



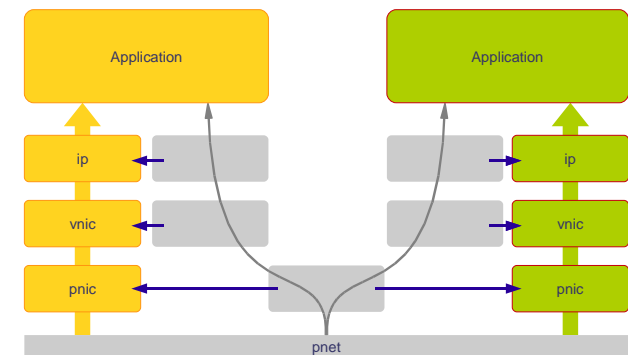


# OSL Virtual Network Objects

Nochmal zum "Warum"



- Netzwerk-Setup auf heutigen Servern sehr komplex
  - Vielzahl verschiedener Netzwerk-Objekte, inzwischen bis hin zu Switches
  - Sharing von Controllern und virtuellen Objekten
  - Überlagerung durch VLANs
  - mehrere VMs oder Applikationen
- OSL Virtual Network Objects (OSL VNO)
  - Abstraktionslayer
  - einheitliche Bedienung unter Solaris und Linux, auch unter verschiedenen Versionen
- Jede Applikation / VM kann ihren eigenen Namespace aufbauen
  - Einstiegspunkt ist immer das Physical Network
  - komplett eigener Namespace für die Applikation bis zum obersten Layer (Interface/VLAN/IP)
  - System Network Objects werden bei Bedarf automatisch erzeugt oder entfernt
- Applikations- / VM-weise Operationen
  - Device Trees
  - Abbau ganzer Trees mit einem Kommando
  - Schutz anderer Konfigurationen
- erhebliche Steigerung der Bediensicherheit

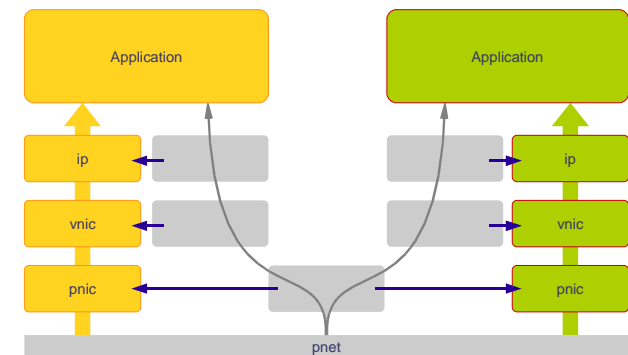


# OSL Virtual Network Objects

Nochmal zum "Warum"



- Netzwerk-Setup auf heutigen Servern sehr komplex
  - Vielzahl verschiedener Netzwerk-Objekte, inzwischen bis hin zu Switches
  - Sharing von Controllern und virtuellen Objekten
  - Überlagerung durch VLANs
  - mehrere **VMs** oder Applikationen
- OSL Virtual Network Objects (OSL VNO)
  - Abstraktionslayer
  - **einheitliche Bedienung** unter Solaris und Linux, auch unter verschiedenen Versionen
- Jede Applikation / **VM** kann ihren eigenen Namespace aufbauen
  - Einstiegspunkt ist immer das Physical Network
  - komplett eigener Namespace für die Applikation bis zum obersten Layer (Interface/VLAN/IP)
  - System Network Objects werden bei Bedarf automatisch erzeugt oder entfernt
- Applikations- / **VM**-weise Operationen
  - Device Trees
  - Abbau ganzer Trees mit einem Kommando
  - Schutz anderer Konfigurationen
- erhebliche Steigerung der Bediensicherheit



# VMs – einheitliche Bedienung & Vielfalt

Eine Säule des OSL Storage Clusters



Symmetric Cluster Engine

Global  
Virtual  
Storage

Unified  
Virtual  
Networking

Application  
Virtual  
Environment

Unified  
Virtual  
Machines

Adaptive  
Resource  
Control / HA

Cross-Platform Capabilities

# Unified Virtual Machines

einfach - flexibel – vielfältig

# Zur Systematik von Applikationen

Applications – Virtual Machine Applications – Virtual Nodes



## Applications

SAP

Informix

Oracle

Baan

Postgres

Backup

...

## Virtual Machine Applications

Solaris Zones

Solaris LDOMs

VirtualBox

KVM

XEN

andere

## Virtual Nodes

Solaris Zones

Solaris LDOMs

VirtualBox

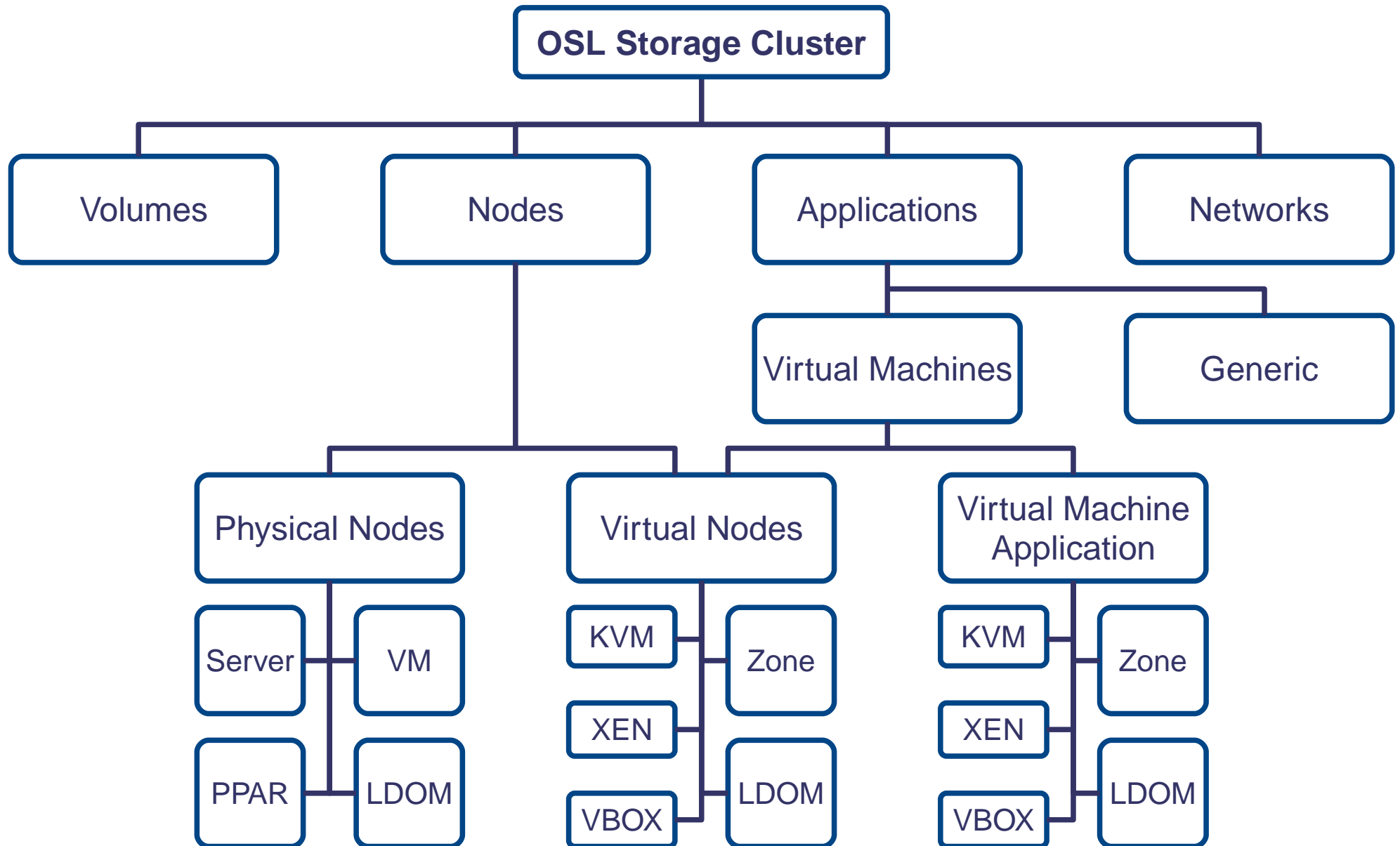
KVM

XEN

andere

# VMs und OSL Storage Cluster

Nahtlose Eingliederung in ein integriertes Konzept



# Linux als Hypervisor – integriert in OSL SC

Einbindung in ganzheitliches Konzept



- Anlegen / Ressourcensteuerung über vmadmin
  - # vmadmin -c vm\_name -F {lkvm | lfxen | vbox | sldm ...}
  - Zuweisung CPU / Memory / Storage
- Start / Stop / Failover über normale SC-Mechanismen
- automatisiertes Anlegen über Menüsystem oder vm\_install
- Detail-Konfiguration über Standard-Werkzeuge (virsh / virt-manager ...)
- Leichter Wechsel zwischen Technologien
- Backup to Disk / Tape integriert (dvam-Tools)
- selbstverständlich failoverfähig
- automatischer Aufbau der notwendigen Netzkonfiguration (hardwareabstrakt)

# Solaris als Hypervisor - Zonen/LDOMs als normale VM

Einbindung in ganzheitliches Konzept



- Anlegen / Ressourcensteuerung über vmadmin
  - # vmadmin -c vm\_name -F {lkvm | lfxen | **szone** | **sldm** ...}
  - Zuweisung CPU / Memory / Storage
- Start / Stop / Failover über normale SC-Mechanismen
- automatisiertes Anlegen über Menüsystem oder zone\_install
- Detail-Konfiguration über Solaris-Standard-Werkzeuge
- Solaris Zones waren schon immer Zonen auf Shared Storage und schon immer failoverfähig
- Backup to Disk / Tape integriert (dvam-Tools)
- selbstverständlich failoverfähig
- automatischer Aufbau der notwendigen Netzkonfiguration (hardwareabstrakt)
- Fähigkeit zur Migration von Solaris 10 nach Solaris 11



# Zonen und LDOMs als Virtual Node

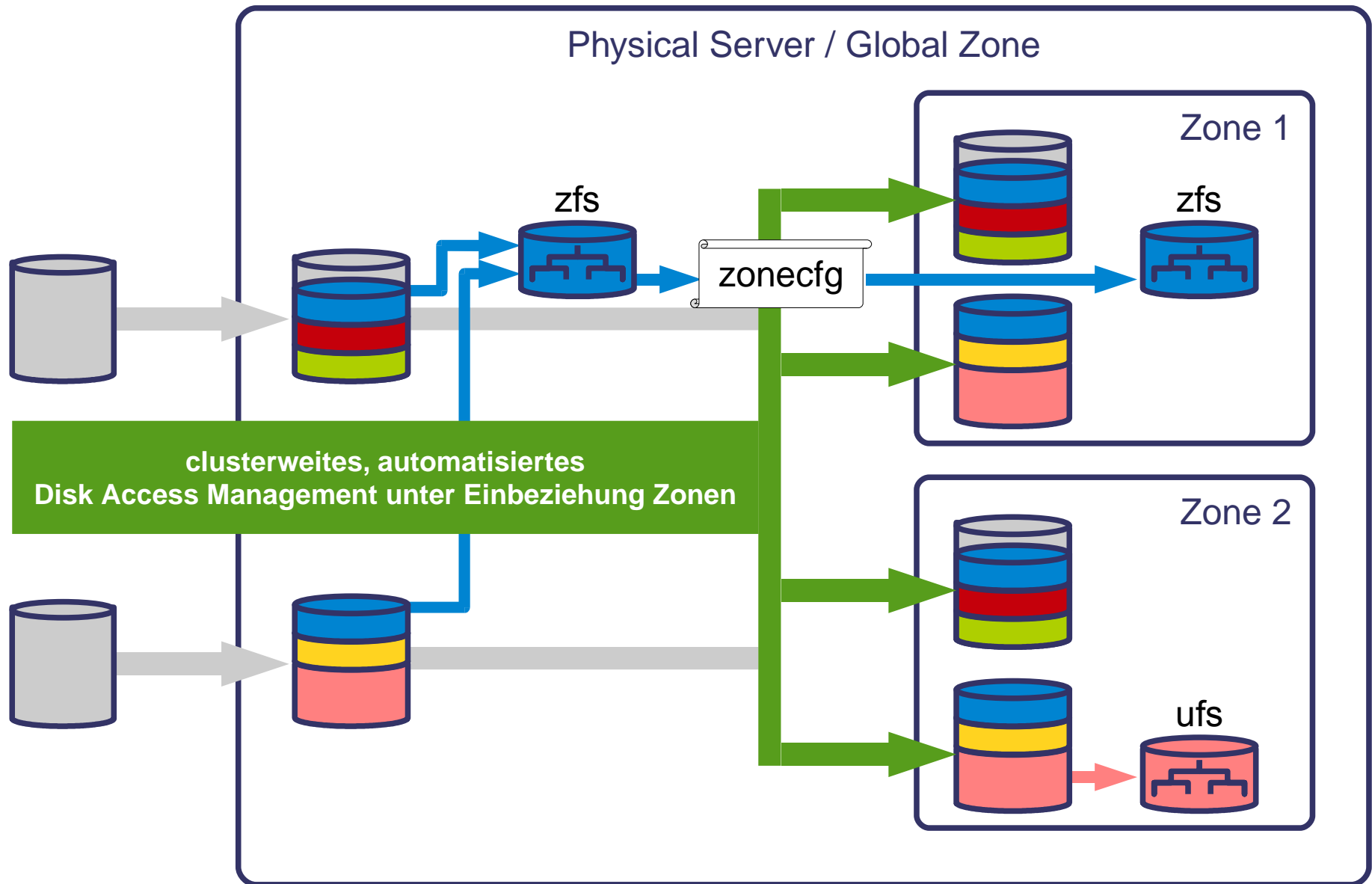
## Die VM als Cluster-Node



- OSL Storage Cluster ist in der Zone/LDOM installierbar und lauffähig
- Zone/LDOM wird zum Clusterknoten
- Anwendungssteuerung damit bis in die Zone/LDOM hinein
- Anwendungen können zwischen Hosts, Zonen und LDOMs beliebig migrieren
- Zone-Failover auf dem gleichen Host, zwischen Domainen oder Hosts
- voller Zugriff auf Speichervirtualisierung in der Zone
- Application Specific Devices + Disk Access Management für Zonen/LDOMs (ASM!)
- Applikation kann aus der Zone/LDOM heraus Spiegel steuern, Multipathing administrieren usw.
- Zone/LDOM kann sich selbst spiegeln
- Zone/LDOM kann (fast) wie ein normaler Host betrieben werden (vfstab ...)

# Zonen als Virtual Node

Vereinfachung und mehr Sicherheit nicht nur mit Raw Devices



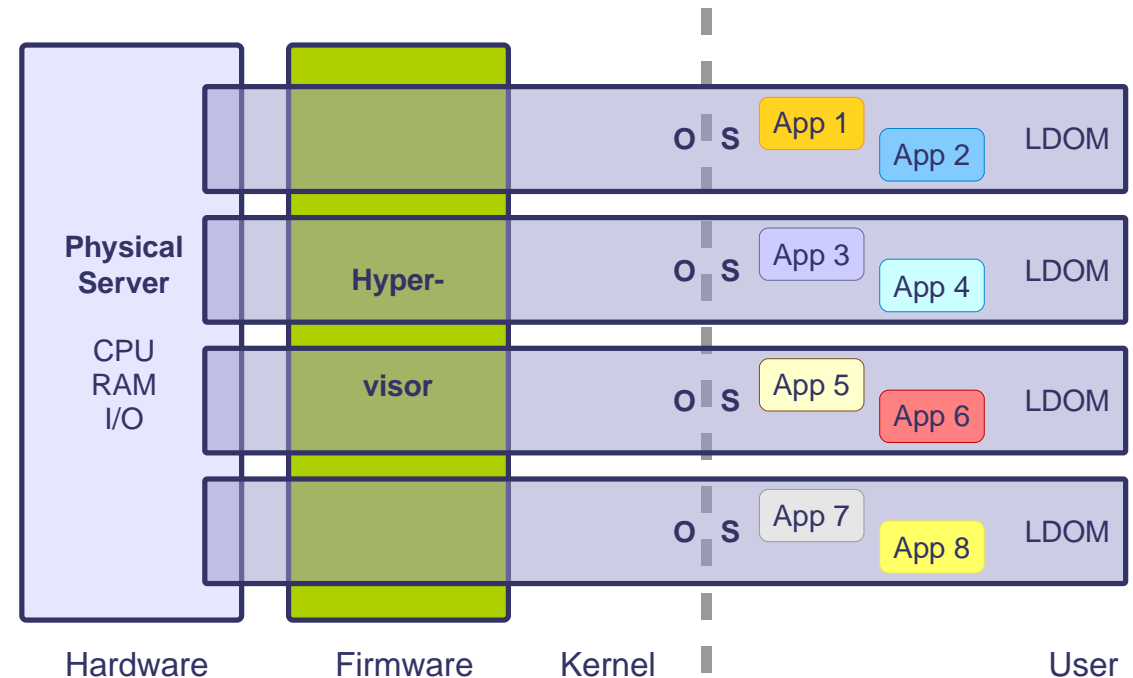
# **LDOMs & OSL Storage Cluster**

# Oracle VM for SPARC - Überblick

Das Pendant zu VMWare in der SPARC-Welt

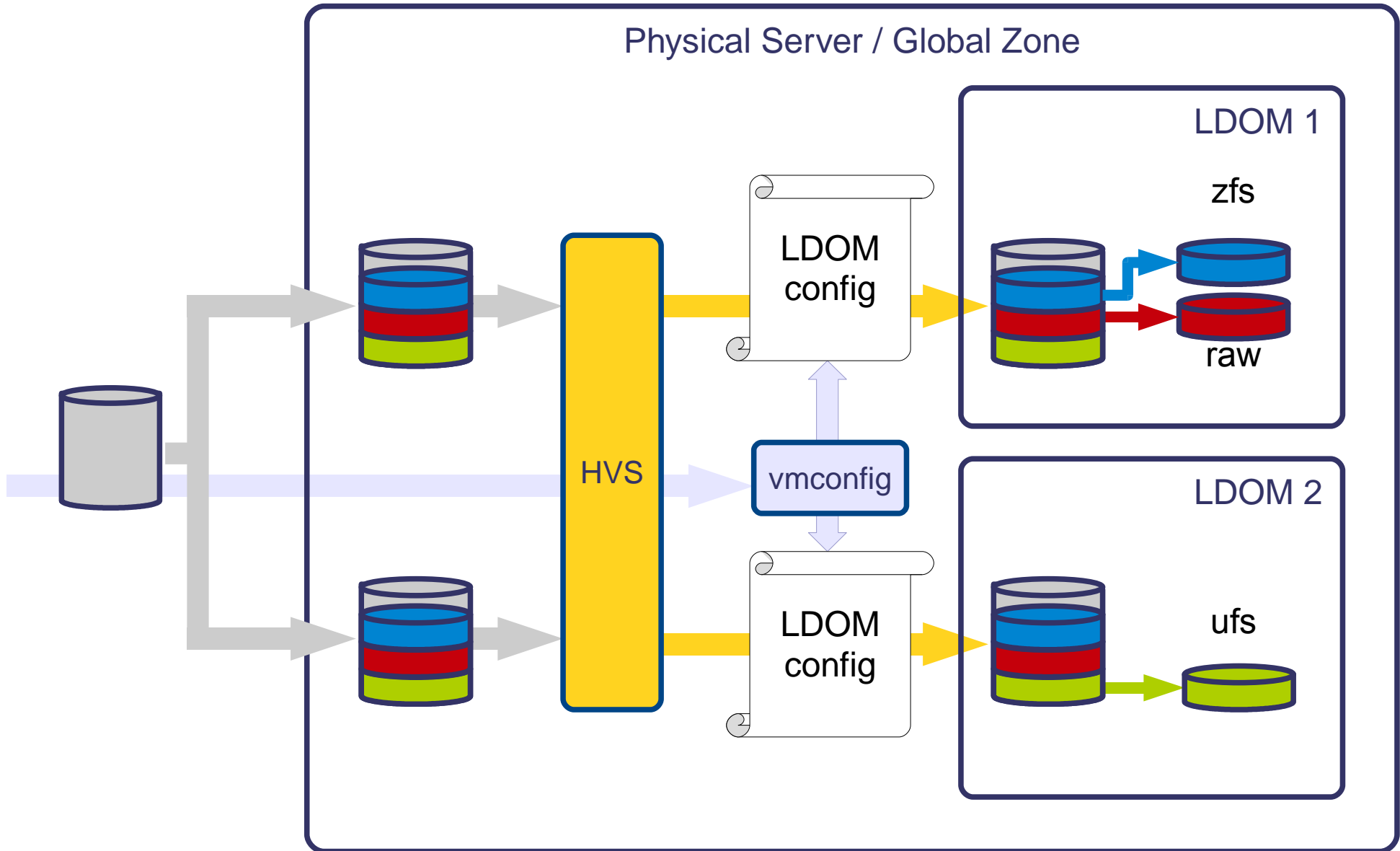


- Nicht wirklich neu
  - trat in Deutschland das erste Mal in Verbindung mit T2000-Systemen in Erscheinung (2007/2008)
  - wurde von Kunden lange Zeit in Produktionsumgebungen ignoriert
  - rückte mit T4-Systemen stärker ins Interesse
- Paravirtualisierung mit geringem Overhead
- Kostet nichts extra (Lizenz mit Hardware gebündelt)
- Verglichen mit Physischer Partitionierung (M-Serie) wesentlich feinere Granularität
- Dynamisch rekonfigurierbar
- Optional redundanter Virtual I/O
- Statischer direkter I/O
- Live-Migration
- Interessant für Optimierung Applikations-Lizenzen (Oracle)



# LDOMs als Virtual Nodes

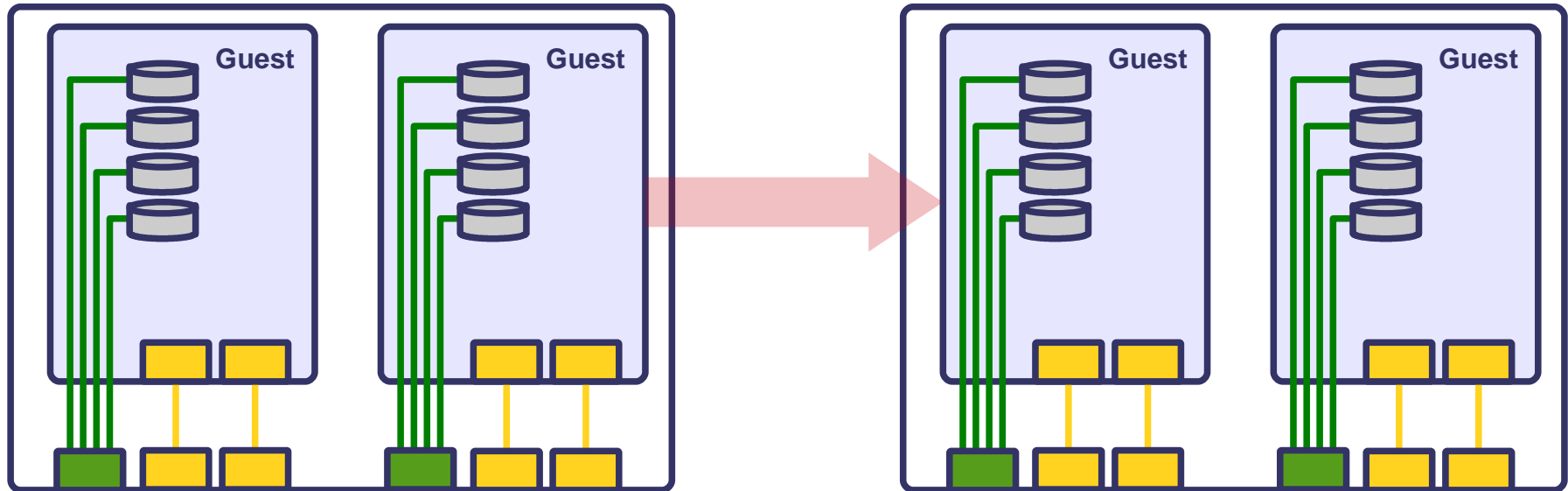
Virtual Storage, Clusterengine, Application Control ... in der LDOM



# Hintergrundinformation LDOMs und Live-Migration

# Das Problem mit Controllern und Live-Migration

VM-Technologie erfordert spezielle I/O-Lösungen

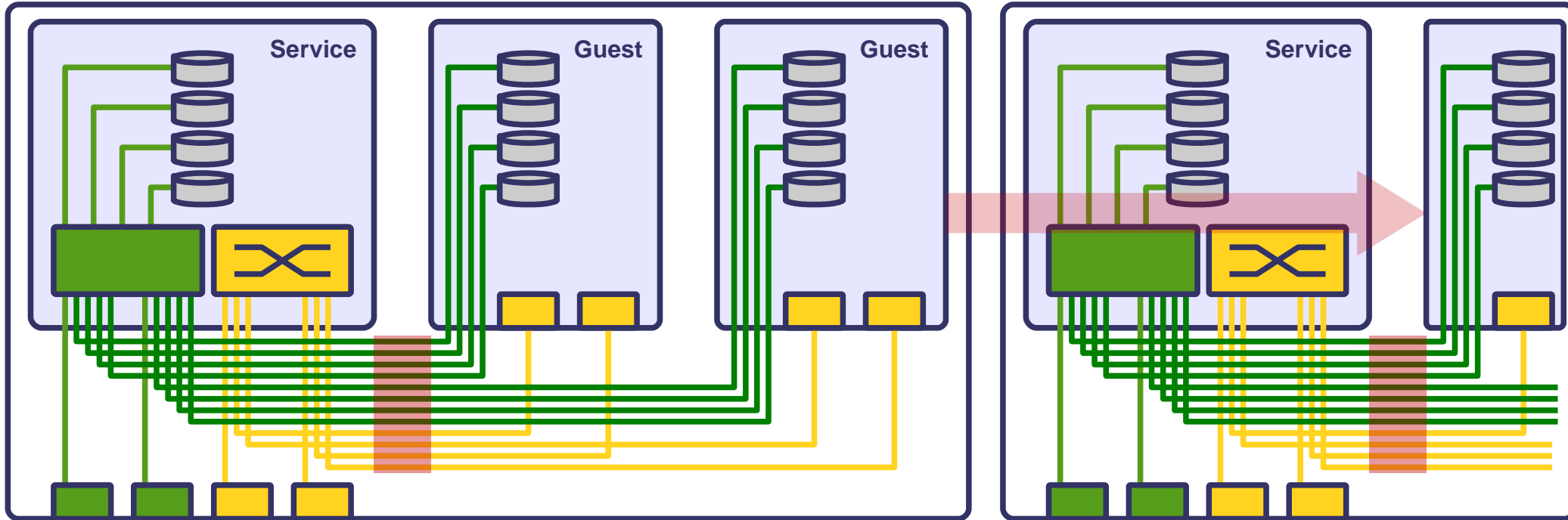


Diese Konfiguration hat nur einen Vorteil (Performance) aber einige handfeste Nachteile:

- Aufwendige Konfiguration
- Kein Controller-Sharing -> es werden viele teure Controller benötigt
- Zugriff Gast-OS auf Hardware schließt Live-Migration aus

# So löst Sun / Oracle das Problem

Virtuelle Device-Server und Virtuelle Netze



**Wir gewinnen mit dem VDS (Virtual Disk Service):**

- Controller-Sharing
- Anbindung der Gast-Geräte über "Logical Domain Channels"
- Fähigkeit zur Live-Migration

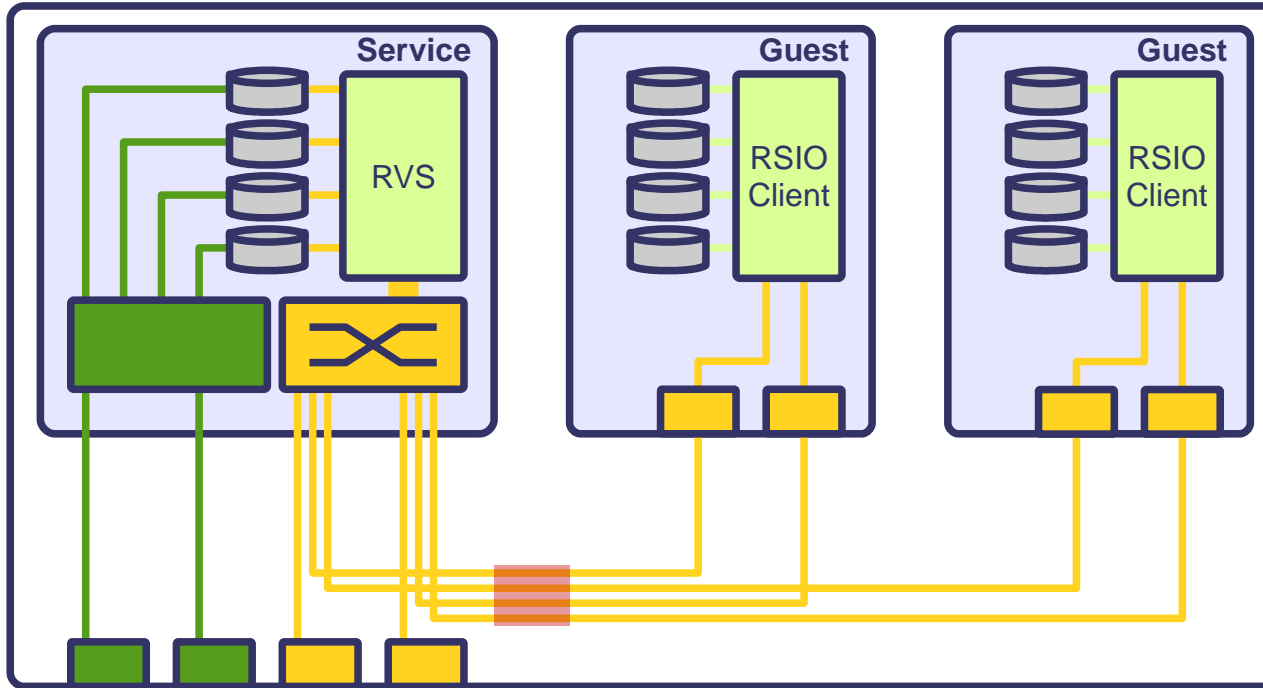
**Aber das Glück ist nicht perfekt:**

- Konfiguration ist aufwendig und muß auf allen Hosts konsistent gehalten werden
- Anzahl "Logical Domain Channels" ist limitiert
- Zu komplexer externer Connectivity kommt eine komplizierte hostinterne Infrastruktur



# Vielleicht hat OSL eine kühle Lösung

HVS: Virtueller Device-Server auf Basis der RSIO-Technologie



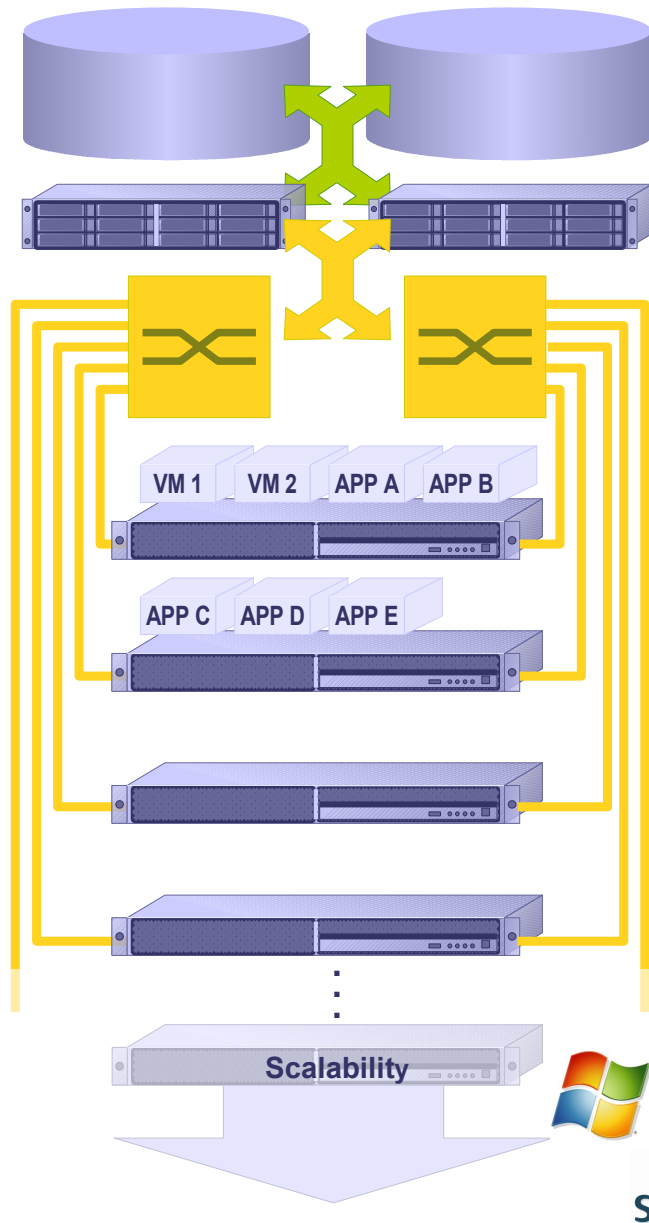
- Kann alles, was OVM VDS kann: - Controller-Sharing  
- Virtuelle Geräte  
- Fähigkeit zur Live-Migration

und es steckt noch deutlich mehr drin:

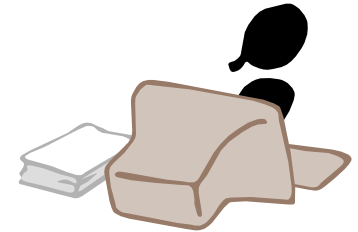
- nochmals vereinfachte Konfiguration (Unified Virtual Network)
- erheblich vereinfachte Administration
- erhebliche Einsparung LDCs
- Bündelung und Sharing von LDCs für I/O -> überlegene Durchsätze

# ... es kann auch ganz anders gedacht werden

radikal einfacher – sicherer, flexibler, leistungsfähiger



- Unified Networking
- weniger Komponenten
- niedrigste Kosten
- einfachste Administration
- aufgeräumte Client-Server-Architektur
- vorhandene Komponenten integrierbar
- beeindruckende Skalierbarkeit
- bewährte Technologie
- Mix: Platforms, Storage, Applications, VMs  
**Solaris, Linux, Hypervisors in einem Framework!**



# Radikale Vereinfachung der (Netzwerk-) Infrastruktur

Alles aus einem Guß



**Komplexität reduzieren**

**Integrieren**

**Kosten senken**

**Flexibilität, Geschwindigkeit, Verfügbarkeit steigern**

- Unified Networking
- weniger Komponenten
- niedrigste Kosten
- einfachste Administration
- aufgeräumte Client-Server-Architektur
- vorhandene Komponenten integrierbar
- beeindruckende Skalierbarkeit
- bewährte Technologie
- SDDC als Produkt



V3 = Virtual • (Storage + Server + Network)



virtualization and clustering – made simple