



# OSL Storage Cluster 4.0 & RSIO

Welten verbinden

**OSL Technologietage**  
Berlin • 12./13. September 2012

**Bert Miemietz**

OSL Gesellschaft für  
offene Systemlösungen mbH

# Die großen Themen in Storage Cluster 4.0

*Bewährtes beibehalten und mit Ergänzungen den Nutzen vervielfachen*



- *RSIO – Data Center Block I/O over Ethernet*
- *Einführung Virtual Nodes*
- *Zonen als Virtual Nodes*
- *Speichervirtualisierung in der Zone*
- *Integration Linux-Systeme*
- *Framework für verschiedene Virtualisierungstechnologien einschl. KVM*
- *Vergrößerung Knotenzahl auf 64*
- *Vergrößerung Zahl Application Volumes auf je 2048 in 4 Universen*

# Hier stehen wir heute

Der Ausgangspunkt für OSL Storage Cluster 4.0

<i>Application Awareness</i>	<b>Application Control Option</b>	<i>Applikationen spiegeln</i>		
	<i>clusterweite Steuerung von Applikationen</i>		<i>Applikationen klonen</i>	
	<i>virtualisierte (hardwareabstrakte) Ablaufumgebungen</i>			<i>B2D / DASI / DR -Werkzeuge</i>
	<i>Hochverfügbarkeit</i>			
	<i>ressourcenbasiertes Selbstmanagement</i>			
<b>Application Resource Description</b>				

<b>Cluster- Speichervirtualisierung</b>
<i>globale Storage-Pools (rechnerübergreifend)</i>
<i>globales Disk-Inventory</i>
<i>globale Geräte / globaler Namensraum</i>
<i>Clustervol-Manager/automatische Allokation</i>
<i>Disk Access Management</i>
<i>I/O-Multipathing</i>

<b>Extended Data Management</b>
<i>Integration RAID-basierter Datenkopien &amp; Snapshots</i>
<i>hostbasierte Spiegelung</i>
<i>online Datenmigration</i>
<i>Daten klonen</i>

# Hier stehen wir heute

Der Ausgangspunkt für OSL Storage Cluster 4.0





# Hier stehen wir heute

Der Ausgangspunkt für OSL Storage Cluster 4.0

<i>Application Awareness</i>	<b>Application Control Option</b>	<i>Applikationen spiegeln</i>		
	<i>clusterweite Steuerung von Applikationen</i>		<i>Applikationen klonen</i>	
	<i>virtualisierte (hardwareabstrakte) Ablaufumgebungen</i>			<i>B2D / DAS / DR -Werkzeuge</i>
	<i>Hochverfügbarkeit</i>			
	<i>ressourcenbasiertes Selbstmanagement</i>			
<b>Application Resource Description</b>				

<b>Cluster- Speichervirtualisierung</b>	<b>RSIO</b>	<b>Extended Data Management</b>
<i>globale Storage-Pools (rechnerübergreifend)</i>	<i>Datacenter Block-I/O z. B. via Ethernet</i>	<i>Integration RAID-basierter Datenkopien &amp; Snapshots</i>
<i>globales Disk-Inventory</i>	<i>Integration mit Speichervirtualisierung</i>	<i>hostbasierte Spiegelung</i>
<i>globale Geräte / globaler Namensraum</i>	<i>Multipathing + Trunking</i>	<i>online Datenmigration</i>
<i>Clustervol-Manager/automatische Allokation</i>	<i>Cross-Platform</i>	<i>Daten klonen</i>
<i>Disk Access Management</i>		
<i>I/O-Multipathing</i>		

# Hier stehen wir heute

Der Ausgangspunkt für OSL Storage Cluster 4.0



***RSIO***

***Data Centre Block I/O over Ethernet***

# RSIO - Remote Storage I/O

*Eckdaten der neuen Technologie für LAN-attached (Shared) Block Devices*



- *neues, von OSL entwickeltes Protokoll*
- *direkter Transport aller relevanten IO-Aufrufe (read, write, ioctl)*
- *integriert Verbindungsaufbau, Überwachung, Path-Multiplexing, Trunking*
- *fähig zu Selbstkonfiguration und Error Recovery*
- *kann alle modernen Storage-Szenarien abbilden:*
  - *einfache Server und Clients, ggf. mit Multipathing*
  - *Cluster von Storage-Servern (Targets)*
  - *Cluster von Storage Clients (Initiators)*
  - *integrierte Cluster von Servern und Clients*
  - *Storage Server Farms*
  - *Cloud-Konzepte*
- *besondere Eignung für Kombination mit Speichervirtualisierung*
  - *eingängige Namen*
  - *fdisk (Partitionierung) auf Clientseite entfällt*
  - *On-Demand-Allokation und Online-Rekonfiguration*
  - *viele weitere Sonderfunktionen*
  - *ermöglicht Administration vom Client aus*

OSL Gesellschaft für offene Systemlösungen mbH  
[www.osl.eu](http://www.osl.eu)



# RSIO – das Performance-Beast

Hohe Performance und beeindruckende Skalierbarkeit bereits mit 1GBit Ethernet



## Server-Performance bei Cache Read / 8k

<i>iSCSI</i>	<i>10 Clients</i>	<i>100 Threads</i>	<i>7,6 Cores</i>	<b>31.000 IOPS</b>
<i>iSCSI / comstar</i>	<i>10 Clients</i>	<i>100 Threads</i>	<i>10,0 Cores</i>	<b>85.000 IOPS</b>
<i>RSIO</i>	<i>4 Clients</i>	<i>64 Threads</i>	<i>5,6 Cores</i>	<b>98.000 IOPS</b>
<i>RSIO</i>	<i>4 Clients</i>	<i>128 Threads</i>	<i>6,3 Cores</i>	<b>102.000 IOPS</b>

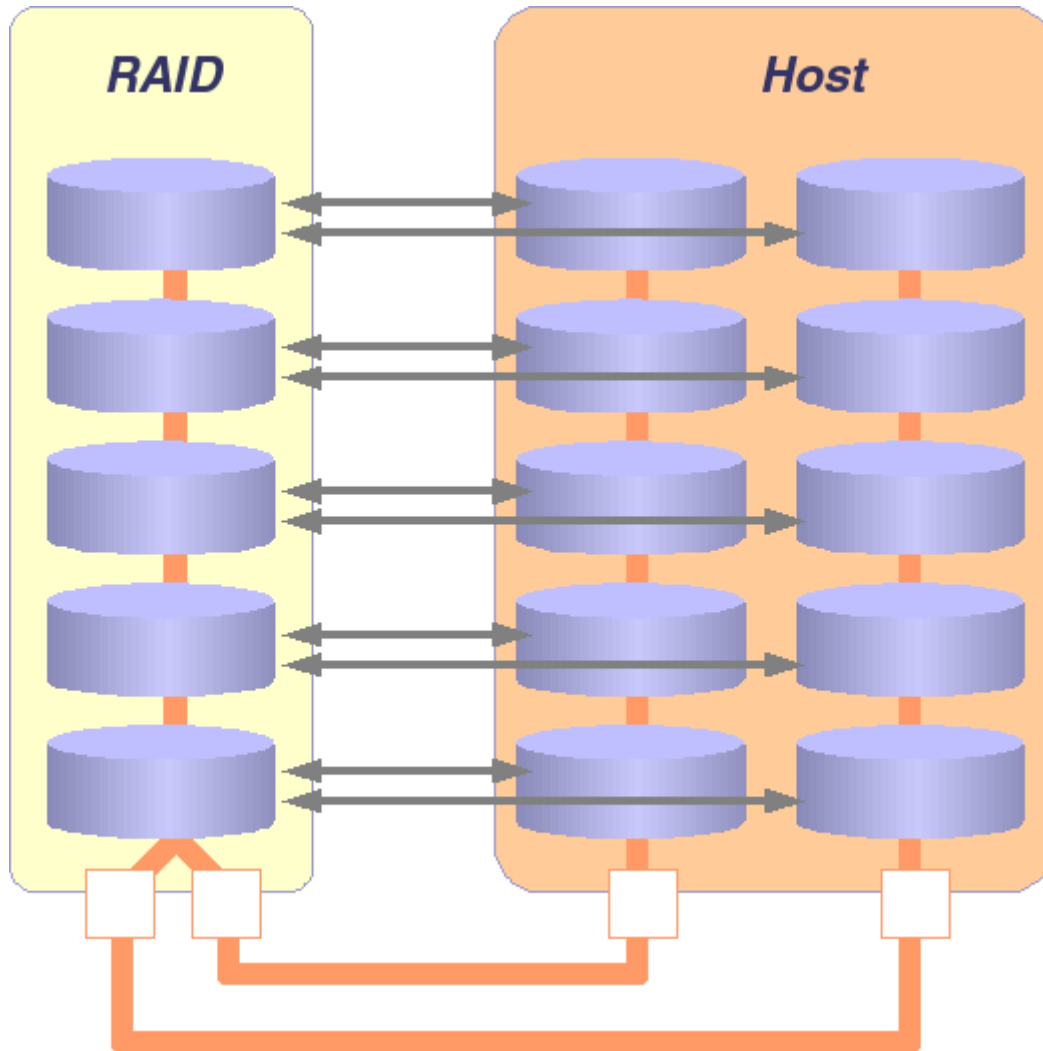
## Client-Performance Throughput

<i>RSIO</i>	<i>1 x 1 GBit</i>	<i>ca. 0,5 Cores</i>	<b>&gt; 110 MByte/s</b>
<i>RSIO</i>	<i>2 x 1 GBit</i>	<i>ca. 1,0 Cores</i>	<b>&gt; 220 MByte/s</b>
<i>RSIO</i>	<i>4 x 1 GBit</i>	<i>ca. 2,0 Cores</i>	<b>&gt; 440 MByte/s</b>
<i>RSIO</i>	<i>8 x 1 GBit</i>	<i>&gt; 4,0 Cores</i>	<b>bis &gt; 900 MByte/s</b>

OSL Gesellschaft für offene Systemlösungen mbH  
[www.osl.eu](http://www.osl.eu)

# Das Wichtigste: Ein völlig anderes Konzept

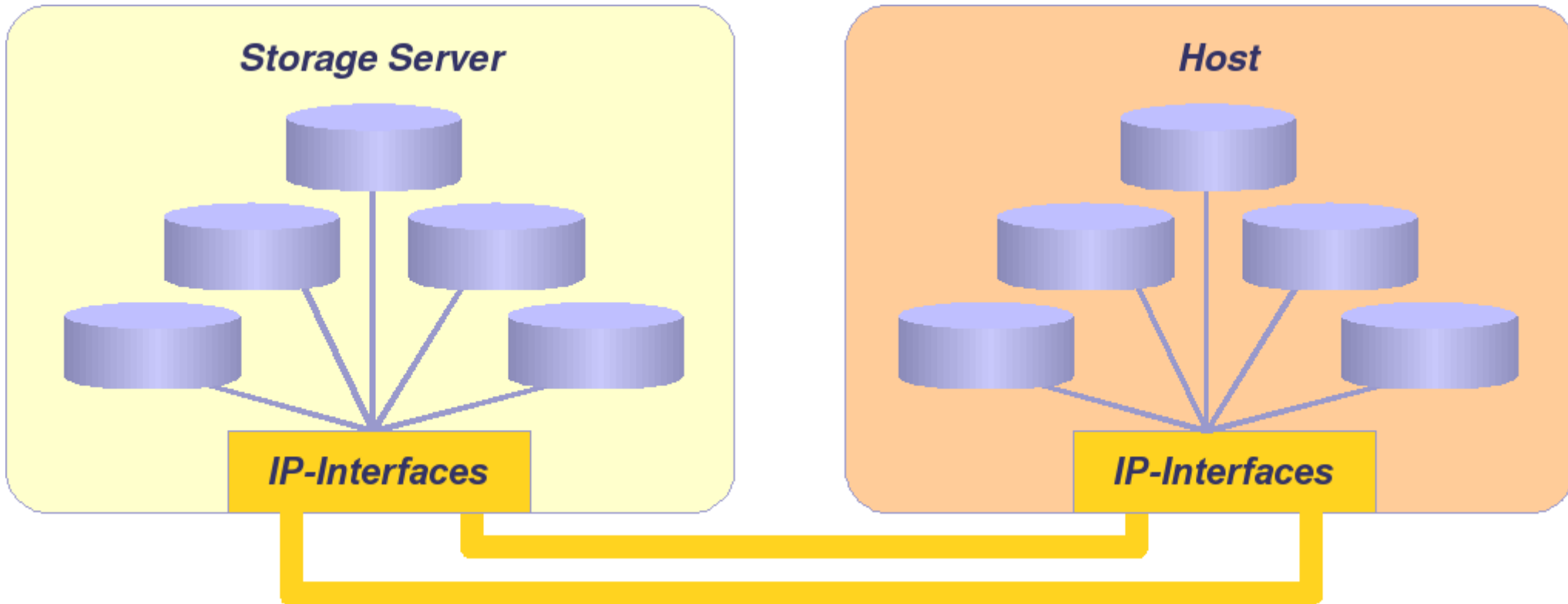
Warum die Welt mit SCSI / Fibre Channel so komplex sein kann



- SCSI über FC, iSCSI etc. emulieren die SCSI Daisy Chain
- Verbindungsstatus pro Lun, bei Multipath entspr. mehrfach -> Zuordnung zum Controller wichtig, daneben hunderte logischer Verbindungen
- SCSI-Layer hat kein Verständnis von der Netztopologie

# Das Wichtigste: Ein völlig anderes Konzept

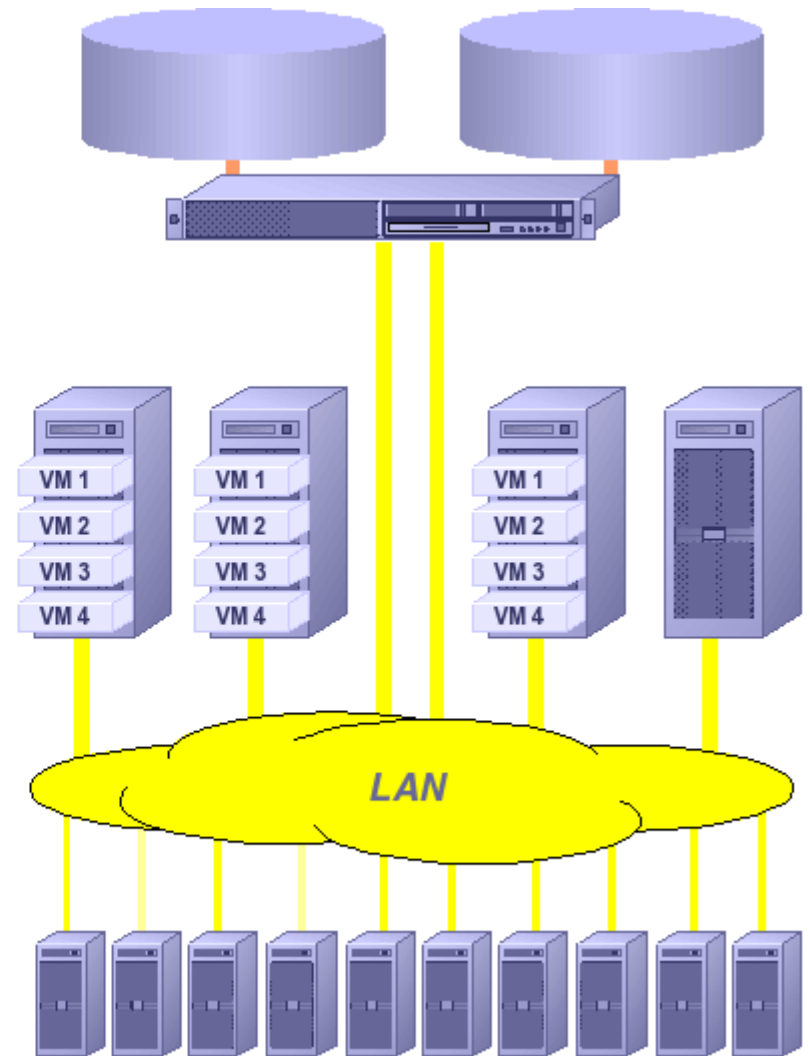
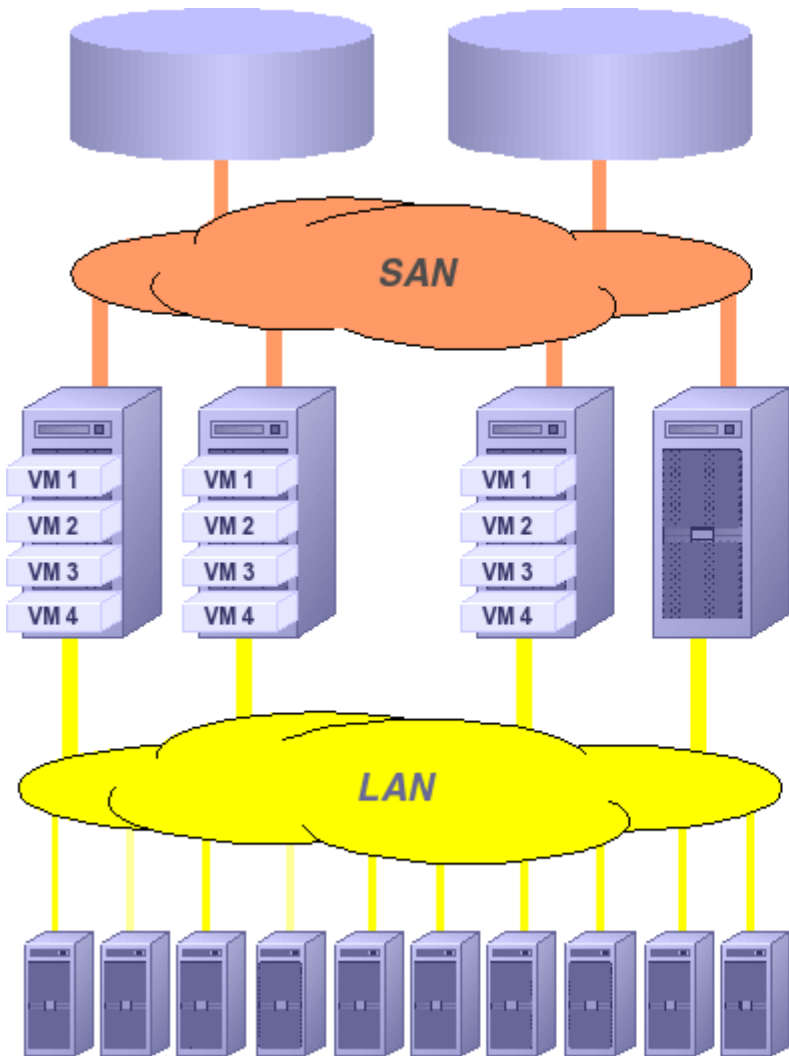
Mehr Übersichtlichkeit und völlig neue Möglichkeiten



- *Mit RSIO werden die Geräte über IP “getunnelt”*
- *Admin und Software interessieren sich nur noch für die logische Verbindung*
- *Die Geräte entstehen aus dem IP-Interface heraus und stehen damit überall zur Verfügung, wo ich ein IP-Interface habe.*

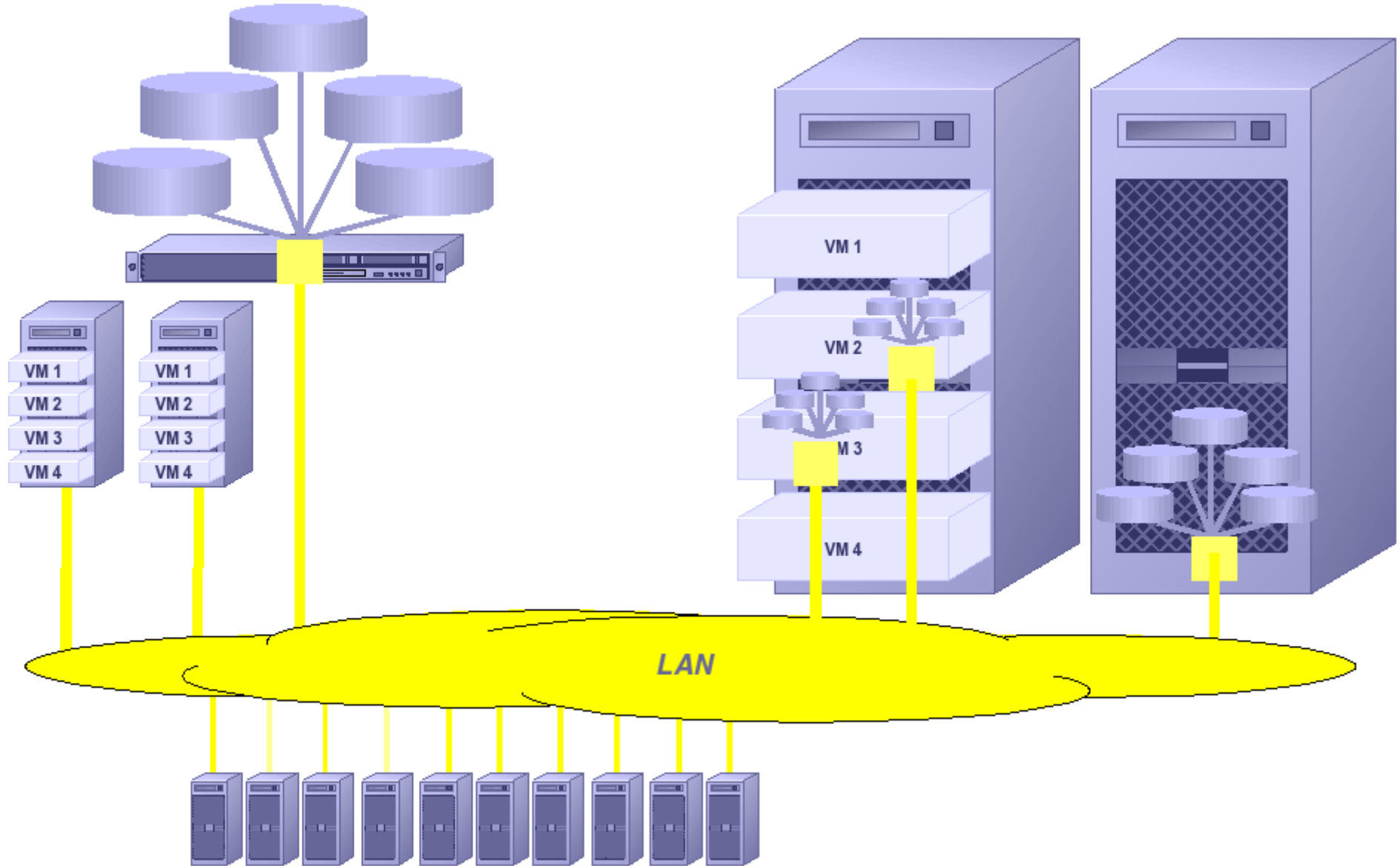
# Was folgt daraus? Einsatzkonzepte

Erste Möglichkeit: Eine Netzwerktechnologie statt zweier!



# Im Detail

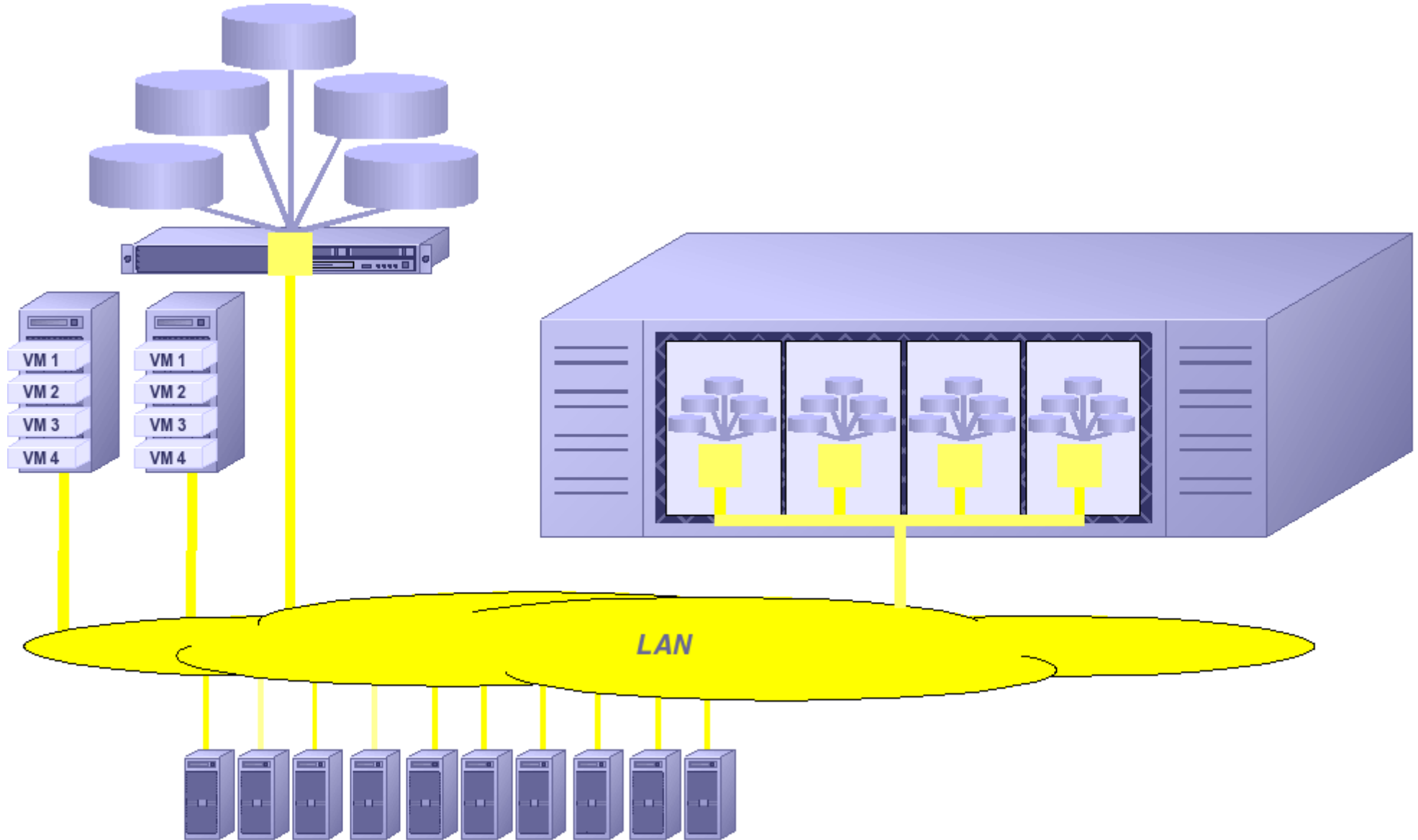
Überall wo IP ist: Storage in 5 Minuten!





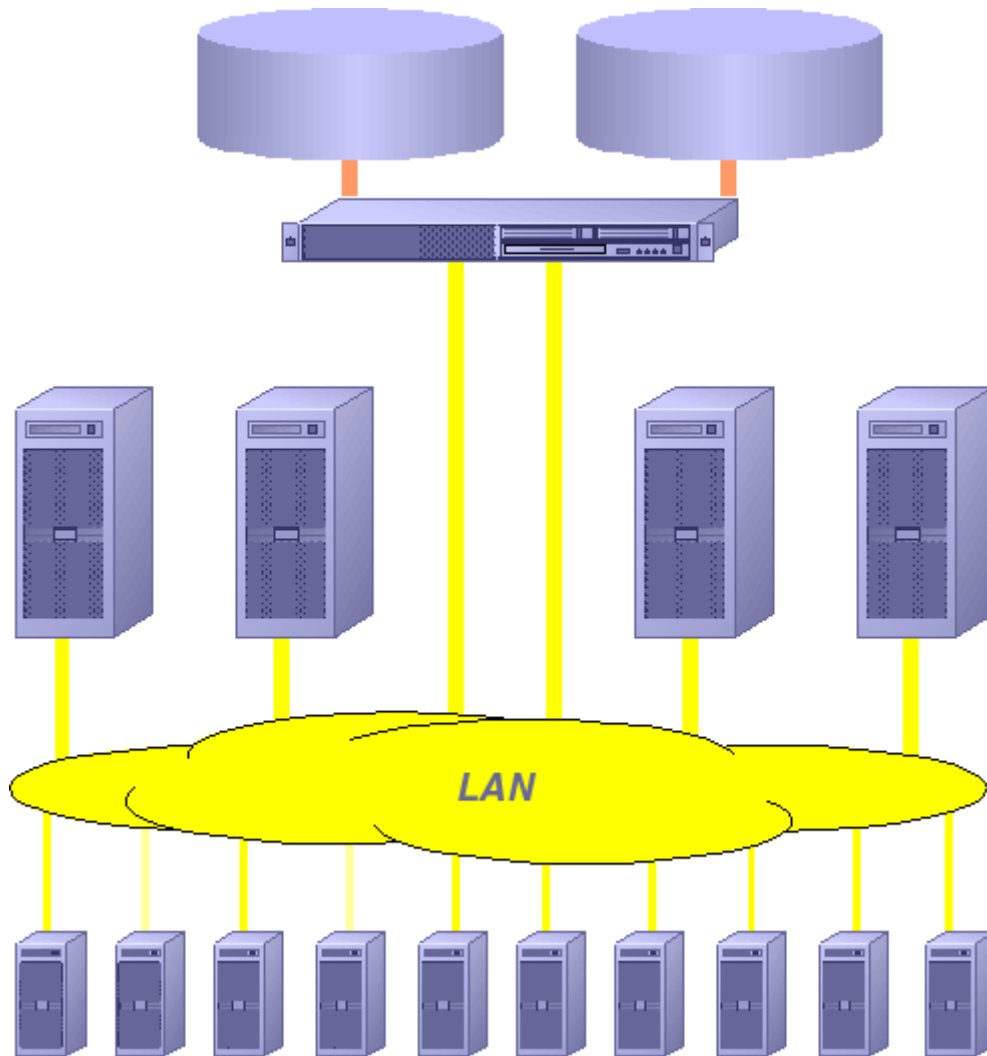
# Das gilt natürlich auch für Blade-Server

Flexible Softwarelösung anstelle proprietärer Hard- und Firmware



# Ein einfaches RZ-Szenario

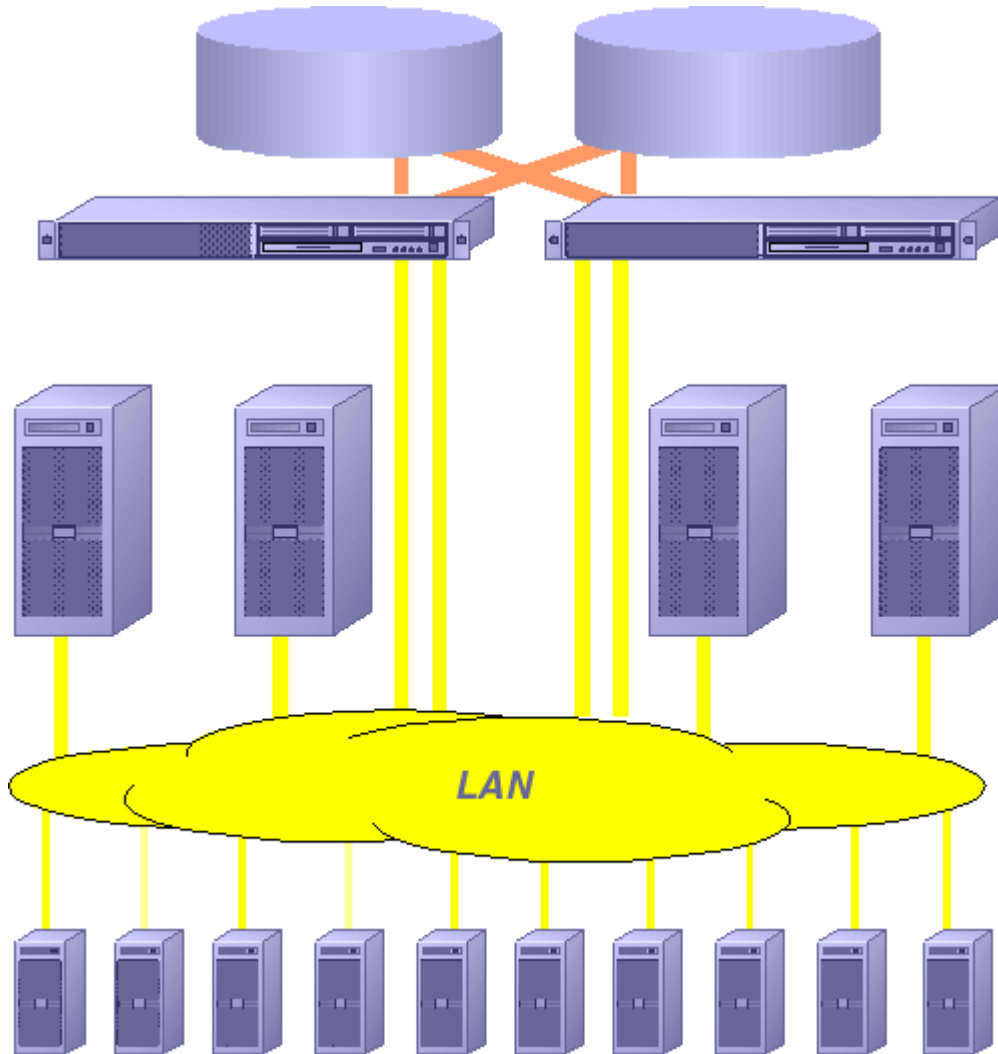
Einbeziehung von FC-Devices - RSIO-Server als LAN-to-FC-Gateway



- *bestmögliches Fanout-Verhältnis für die FC-Ports der RAID-Systeme*
- *preiswerteste und trotzdem performante und zuverlässige Storage-Anbindung*

# Für höhere Ansprüche

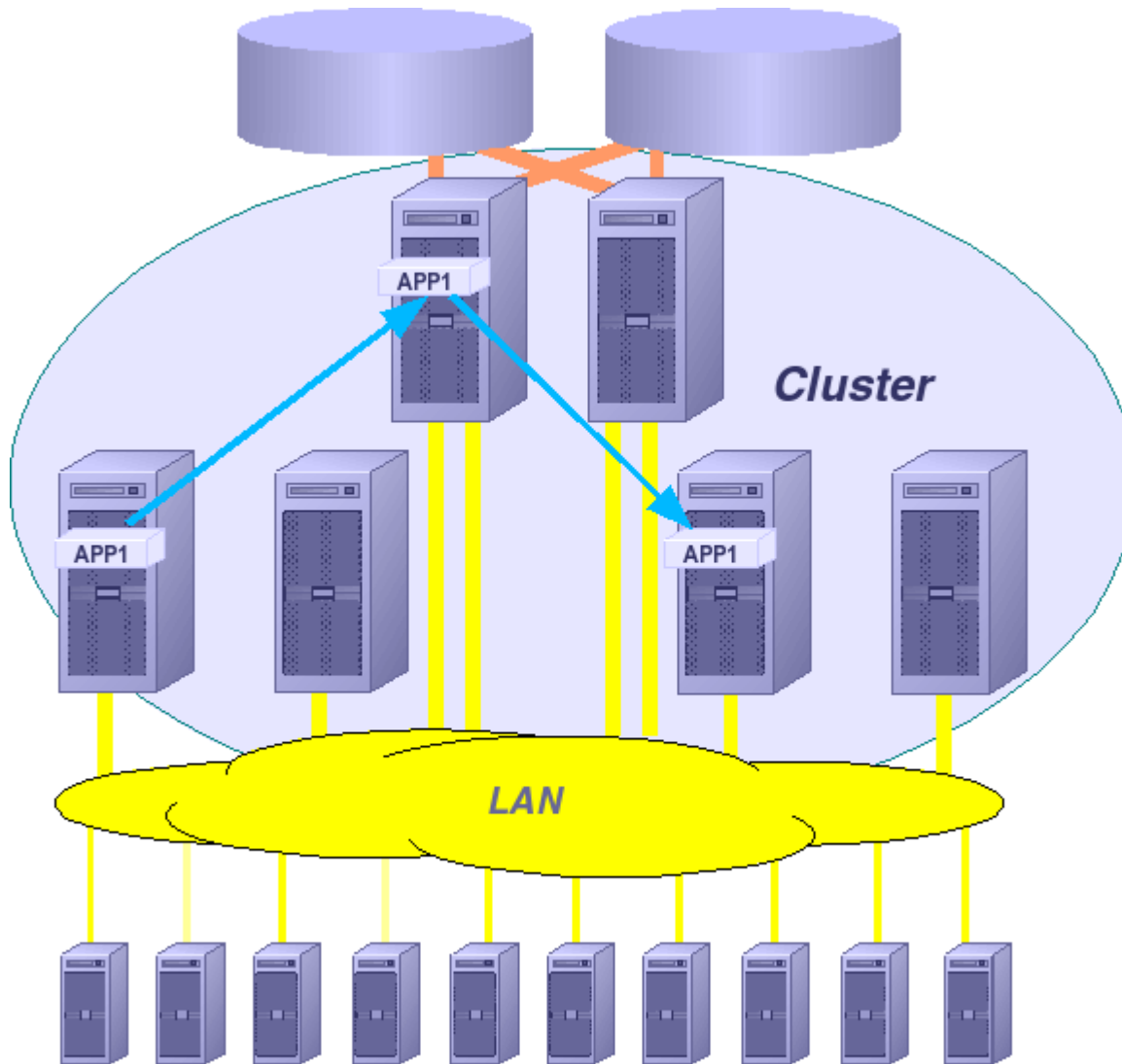
Geclusterte Storage-Server – Skalierbar in Redundanz und Durchsatz



- *Multipathing für die Clients über mehrere Storage-Server*
- *Bei nicht zu vielen Storage-Servern weiterhin Verzicht auf FC-Switches möglich*

# Ausbau zum multifunktionalen Storage-Cluster

Storage-Server und Storage-Clients in einem Cluster

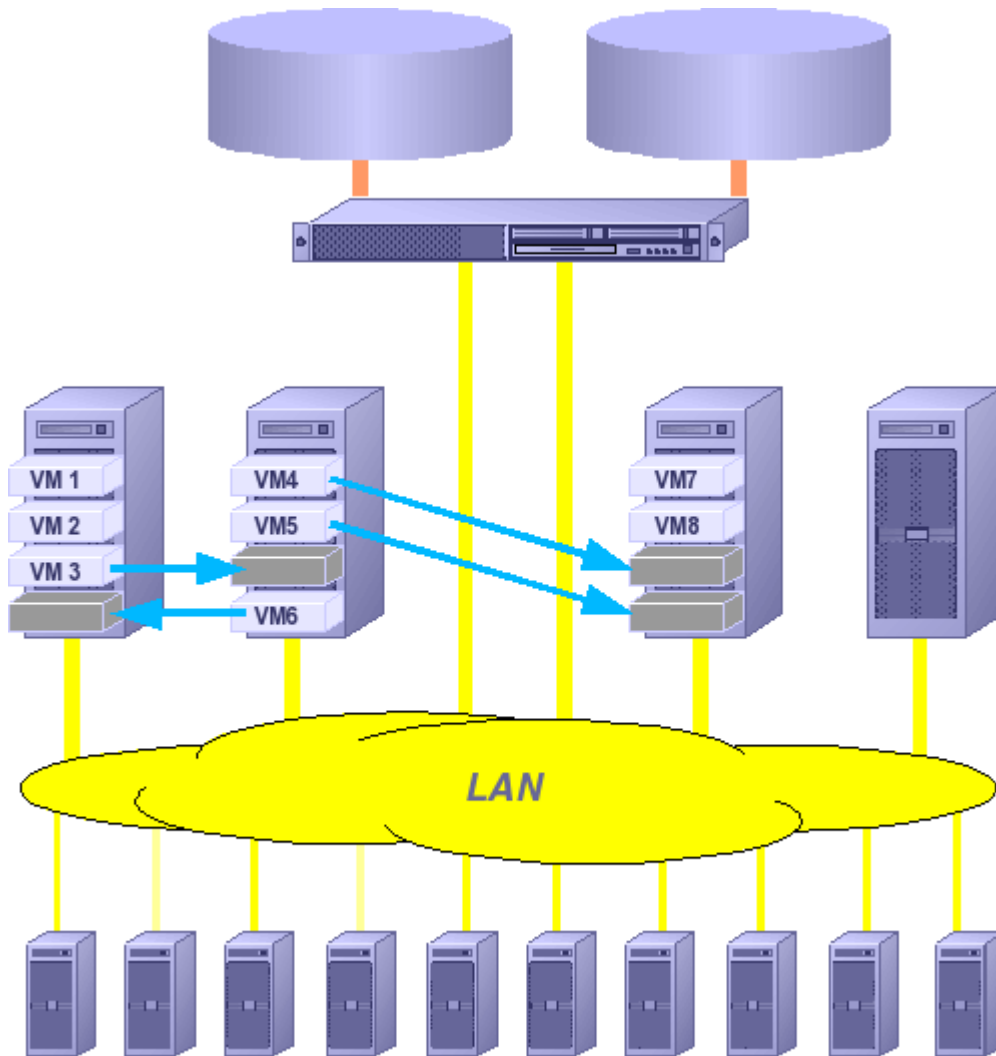


Mit Installation der Storage-Cluster-Pakete läßt sich ein RSIO-Client zum vollwertigen Cluster-Mitglied machen, d.h.:

- Storage-/Device-Allokation vom Client aus
- Spiegeloperationen vom Client aus
- automatisches Access-Management
- identische Gerätesicht auf Storage-Servern und -Clients
- optionale HV für Applikationen (Applikationen auf jedem Clusternode lauffähig, gleich ob Storage-Server oder -Client)

# Storage Cluster, RSIO und KVM

Einheitliche Netzwerk-Infrastruktur, einfaches Handling



In Verbindung z.B. mit KVM erhalte ich einen VM-Cluster:

- Storage, LAN, Live-Migration etc. über die gleiche Infrastruktur: Ethernet
- Hochverfügbarkeit
- kein Trennung von Storage- und Control-Verbindungen
- kein Split Brain
- zentrale Administration möglich
- Alle Funktionen der Speichervirtualisierung für VM-Abbilder nutzbar
- via IP Zugriff auf die gesamte Storage-Welt einschl. Storage-Virtualisierung aus der VM heraus installierbar / konfigurierbar



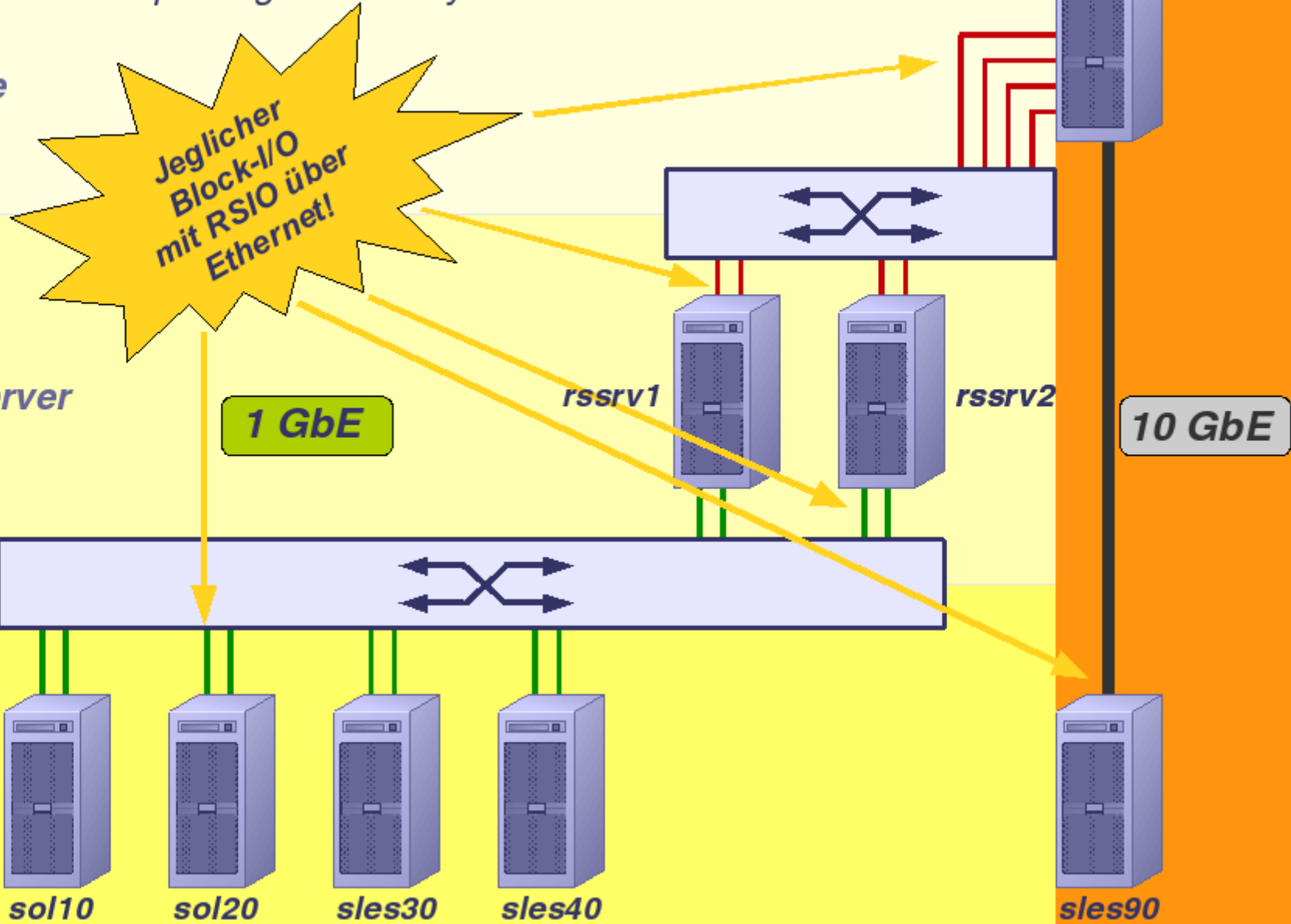
# Ein Blick auf die Praxis

Die Vorführkonfiguration in praxisgerechter Systematik

Backend Storage

RSIO Storage Server

RSIO Clients

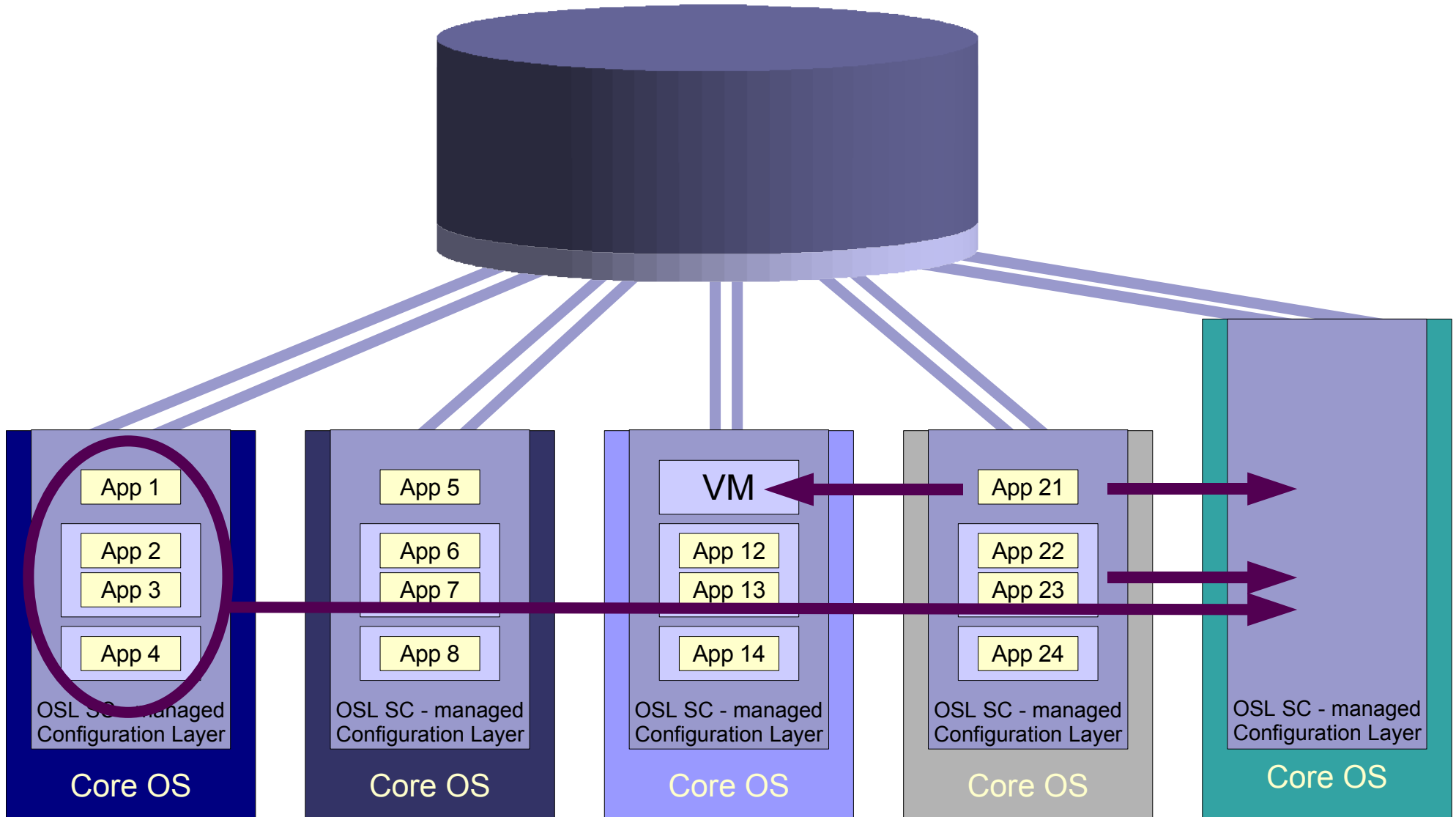


***Mittendrin!***

***Virtual Nodes***

# Was hat es mit Virtual Nodes auf sich?

Flexibilität nach Kundenwunsch: Virtual Nodes zwischen Physical Nodes verschieben, Applikationen zwischen Physical Nodes und Virtual Nodes verschieben



# Ein Blick auf's System

Ein wirklich integratives Clusterkonzept

Physical Node

Remote Storage Attachment

```
# ndadmin -lvvv
```

nodename	id	state	tf	os	cpu-isa	vcpu	clock	memory
rssrv1	1	ONLINE	p	SunOS 5.10	amd64	4	2833	7863
sles30	3	ONLINE	pR	Linux 3.0	amd64	4	3201	15811
sles40	4	OFFLINE	pR	Linux 3.0	amd64	4	3201	15811
sol10	5	OFFLINE	pR	SunOS 5.10	amd64	2	3000	7866
rssrv2	6	ONLINE	p	SunOS 5.10	amd64	4	2833	7868
z1	7	ONLINE	v	SunOS 5.10	amd64	4	2833	7863
slesvm	8	OFFLINE	vR	Linux 2.6	amd64	2	3192	998

```
# ndadmin -qo
```

nodename	id	os	release	brand	revision
rssrv1	1	SunOS	5.10	Solaris 10	10/09
sles30	3	Linux	3.0.13-0.27	SLES 11	11.2
sles40	4	Linux	3.0.13-0.27	SLES 11	11.2
sol10	5	SunOS	5.10	Solaris 10	10/09
rssrv2	6	SunOS	5.10	Solaris 10	10/09
z1	7	SunOS	5.10	Solaris 10	10/09
slesvm	8	Linux	2.6.32.12-0.7	SLES 11	11.1

Virtual Node

## Applications

### Virtual Machine Applications

Solaris Zones

VirtualBox

KVM

andere

### Virtual Nodes

Solaris Zones

VirtualBox

KVM

andere



# Solaris Zonen als Virtual Nodes

*Der tatsächlichen Nutzung bei SAP-Anwendern entsprechen*

- *Installation aller Pakete ab OSLscsys*
- *Knoten erscheint in der Knotenliste*
- *voller Zugriff auf Speichervirtualisierung  
(Volumes anlegen, vergrößern, löschen)*
- *voller Zugriff auf XDM-Operationen  
(Zone kann sich selbst spiegeln oder klonen)*
- *Zugriff auf volle Anwendungssteuerung*
- *Zone verhält sich weitestgehend wie ein Physical Node*
- *Zonen lassen sich clustern*
- *Applikationen können zwischen Zonen oder Zonen und Physical Nodes verschoben werden*
- *Zonen werden wie VMs administriert*

# KVM als Virtual Node

Die Open Source - Virtualisierung



- *Enorme Leistungsfähigkeit*
- *überlegenes Softwarekonzept – mit Linux-Kernel stets für aktuelle Hardware verfügbar*
- *voller Support durch SAP*
- *OSL Storage Cluster bietet umfangreiches Framework:*
  - *Provisionierung*
  - *Start / Stop / Überwachung / Live-Migration / Failover*
  - *Verflechtung mit Speichervirtualisierung (Spiegelung, Clonen, Datenverschiebung, Reorganisation, restorefreier Wiederanlauf)*
  - *Möglichkeit der Anwendungssteuerung in der VM*
  - *ausgeklügelte Möglichkeiten für Systemkopien, für Point-in-Time-Recovery, Offline-Redolog-Archivierung ...*
- *Unabhängigkeit von bestimmten Hardwarelösungen*
- *Offenheit für neue Konzepte und leistungsfähigste Hardware*
- *Für SAP-Anwender: Wenn meine neue strategische Plattform SLES sein sollte – was spricht dann für eine proprietäre OS-Virtualisierung im Stack?*

OSL Gesellschaft für offene Systemlösungen mbH  
[www.osl.eu](http://www.osl.eu)

# VirtualBox als Virtual Node

*Auch VirtualBox ist Open Source*



- *auch hier geht mehr, als man gemeinhin denkt*
- *verfügbar für Solaris und Linux, und das auch noch kompatibel!*
- *in der VBox kann auch Solaris laufen*
- *Support durch Oracle möglich*
- *OSL Storage Cluster bietet umfangreiches Framework:*
  - *Provisionierung*
  - *Start / Stop / Überwachung / Live-Migration / Failover*
  - *Verflechtung mit Speichervirtualisierung*  
(*Spiegelung, Clonen, Datenverschiebung, Reorganisation, restorefreier Wiederanlauf*)
  - *Möglichkeit der Anwendungssteuerung in der VM*
  - *ausgeklügelte Möglichkeiten für Systemkopien, für Point-in-Time-Recovery, Offline-Redolog-Archivierung ...*
- *Unabhängigkeit von speziellen Hardwarelösungen*
- *ist ebenfalls im OSL Storage Cluster implementiert*

# Einheitliche Bedienerchnittstellen für alle VMs

OSL bietet Abstraktionslayer und eigenes Framework

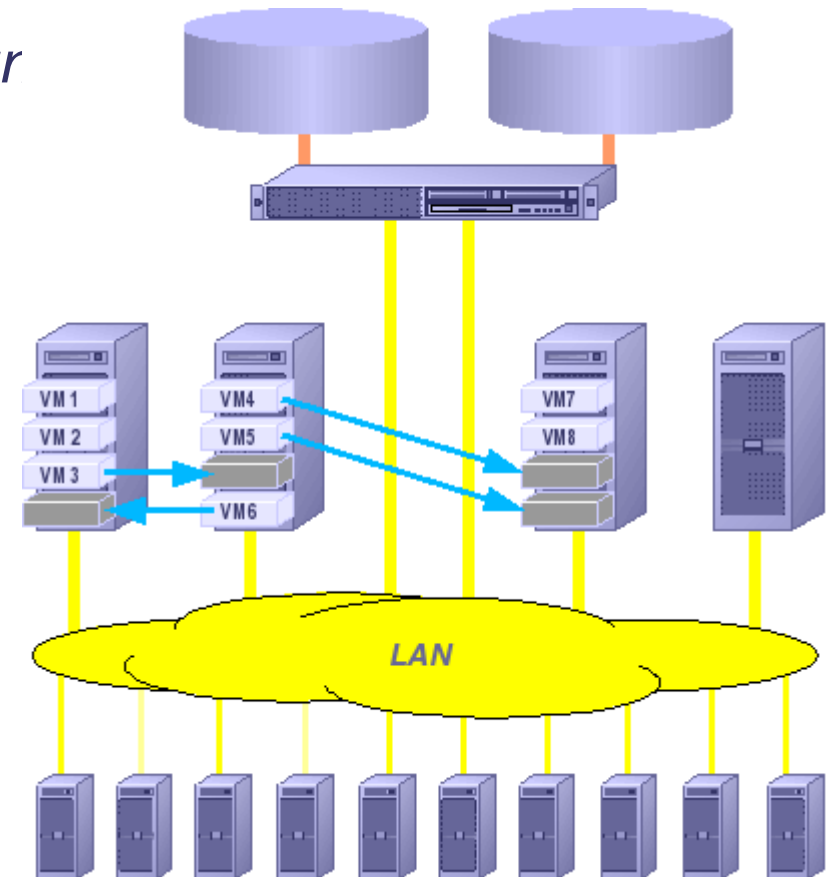
- abstraktes CLI: vmsash

## Clusterkommandos (nur in Verbindung mit ACO)

- vmsadmin einschl. Ressourcenbeschreibur
- vmstart
- vmstop
- vmbreak
- vmrestart

**Hochverfügbarkeit von der einfachen Applikation bis hin zur VM für**

**Solaris + Linux**



# Was ist noch neu?

Diverse “unscheinbare” Änderungen

- *bis zu 64 Knoten (Physical + Virtual)*
- *2048 AVs pro Universum (bei 4 Universen insgesamt 8192)*
- *differenziertere Applikationszustände (jetzt mehrere tausend möglich)*
- *verbesserte Funktionen in RSIO*

# OSL Storage Cluster 4.0 - Zusammenfassung

*Bewährtes + Neues = vielfacher Nutzen*



- *RSIO – Data Center Block I/O over Ethernet*
- *Einführung Virtual Nodes*
- *Zonen als Virtual Nodes*
- *Speichervirtualisierung in der Zone*
- *Integration Linux-Systeme einschließlich Hochverfügbarkeit*
- *Framework für verschiedene Virtualisierungstechnologien einschl. KVM*
- *Vergrößerung Knotenzahl auf 64*
- *Vergrößerung Zahl Application Volumes auf je 2048 in 4 Universen*

OSL Gesellschaft für offene Systemlösungen mbH  
[www.osl.eu](http://www.osl.eu)