

OSL Aktuell

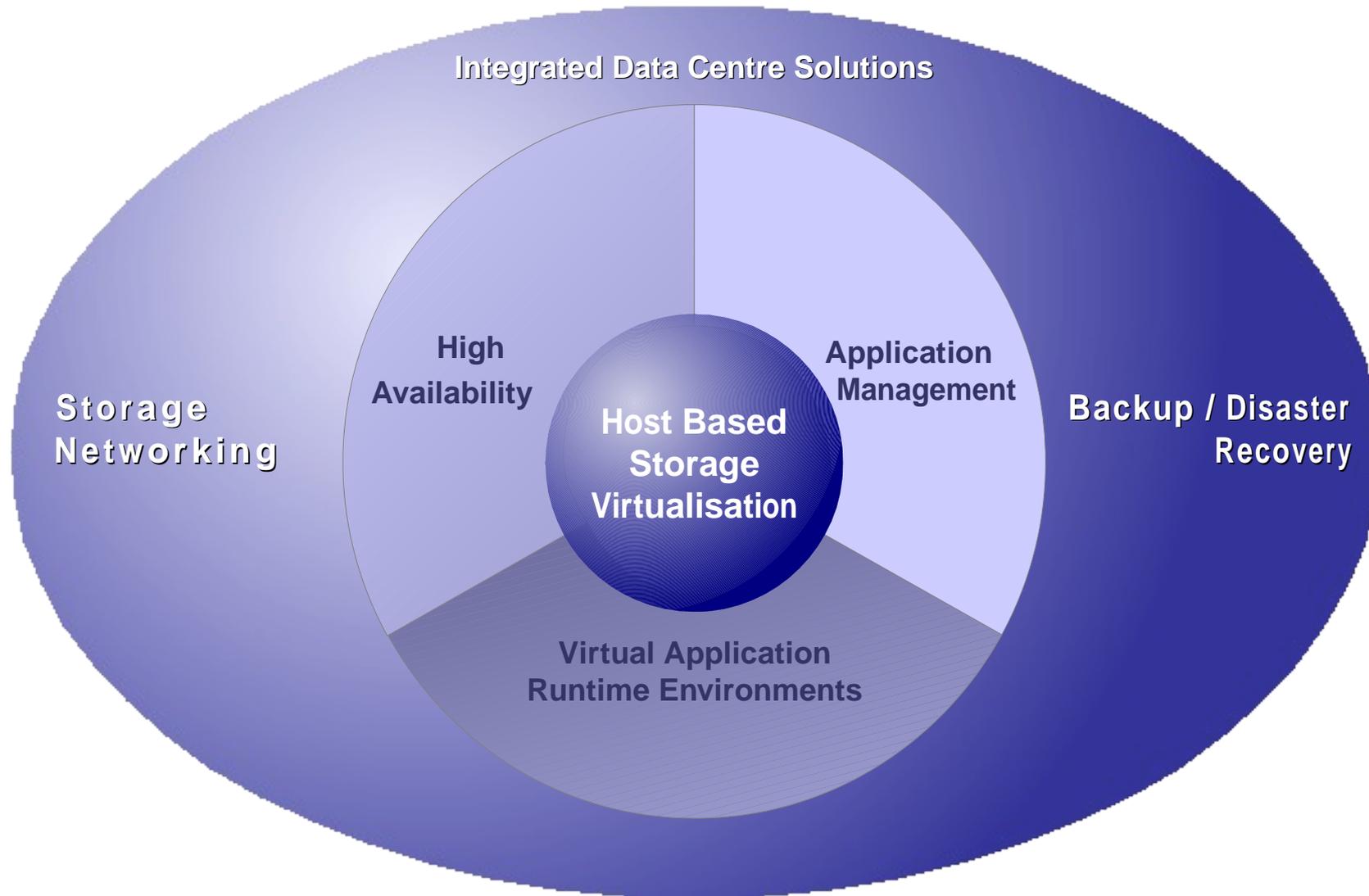
Optimierte RZ-Infrastrukturen

24. April 2013
Schöneiche / Berlin

Wer ist OSL?

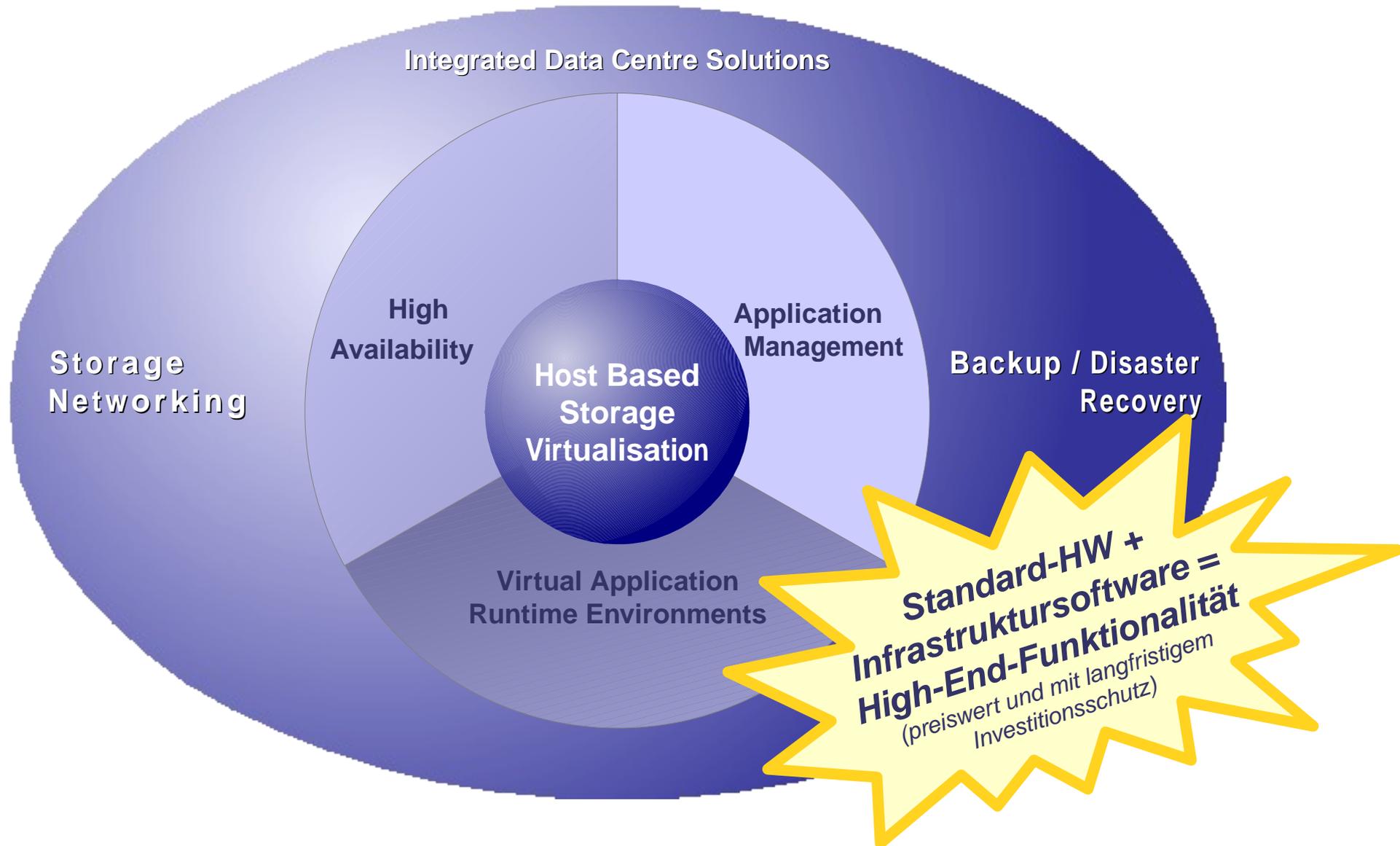
OSL entwickelt Infrastruktur-Software

Storage Virtualisation • Volume Management • Converged Networking
Virtual Machines • Clustering • High Availability • Disaster Protection



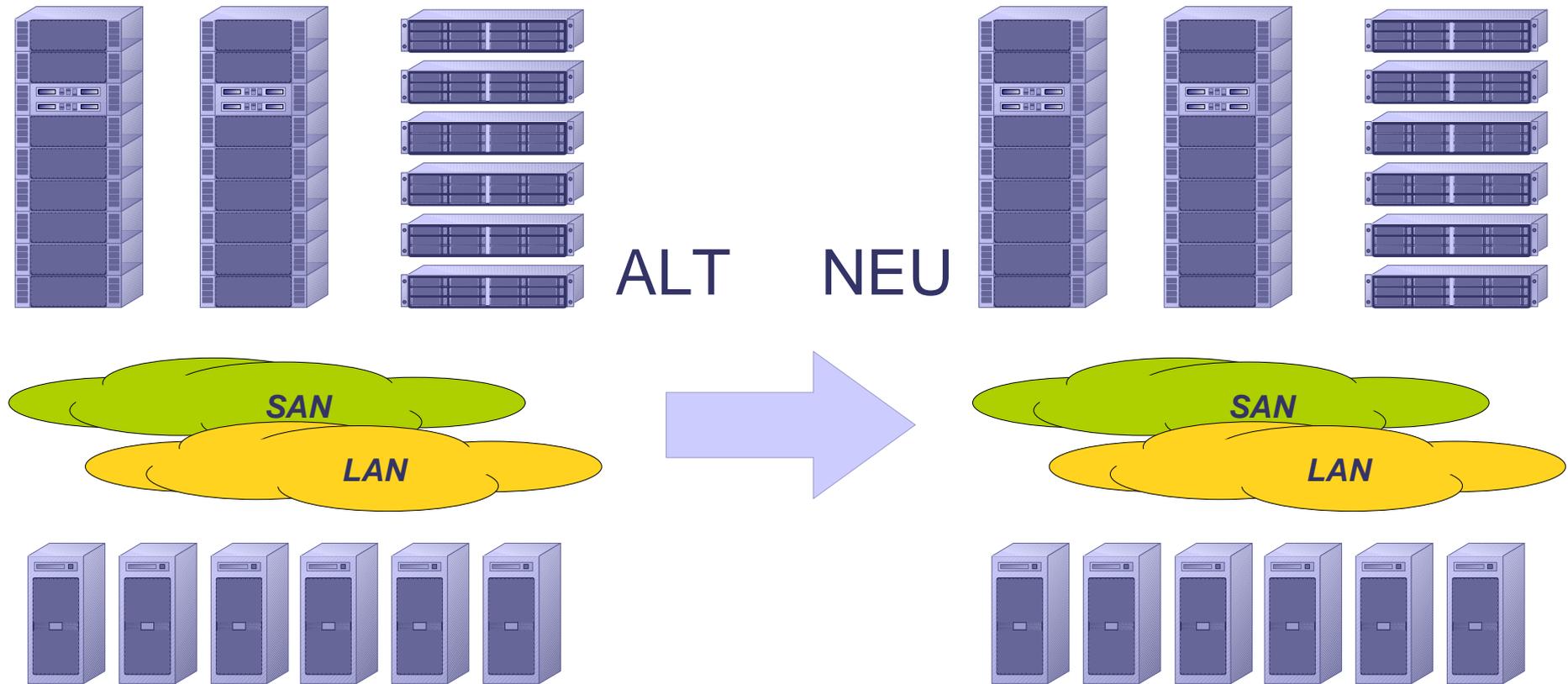
OSL entwickelt Infrastruktur-Software

Storage Virtualisation • Volume Management • Converged Networking
Virtual Machines • Clustering • High Availability • Disaster Protection



Warum sollte ich so etwas einsetzen?

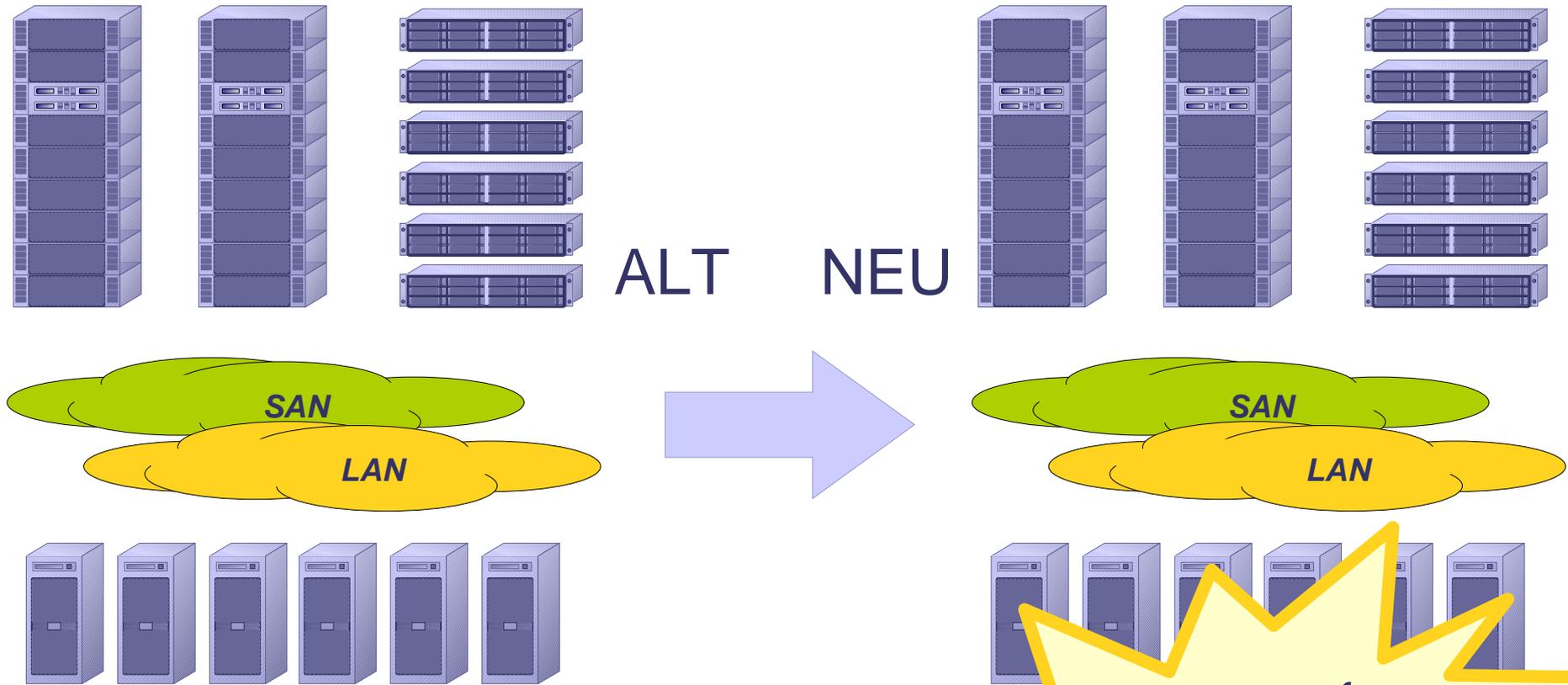
Ein Beispiel



- einfacheres, sicheres Management, schlankere Prozesse
- mehr Verfügbarkeit und Flexibilität
- Hardwareabstraktion \Rightarrow keine Angst vor Systemwechseln
- bessere Auslastung der Systemressourcen \Rightarrow mit weniger mehr erreichen

Warum sollte ich so etwas einsetzen?

Ein Beispiel

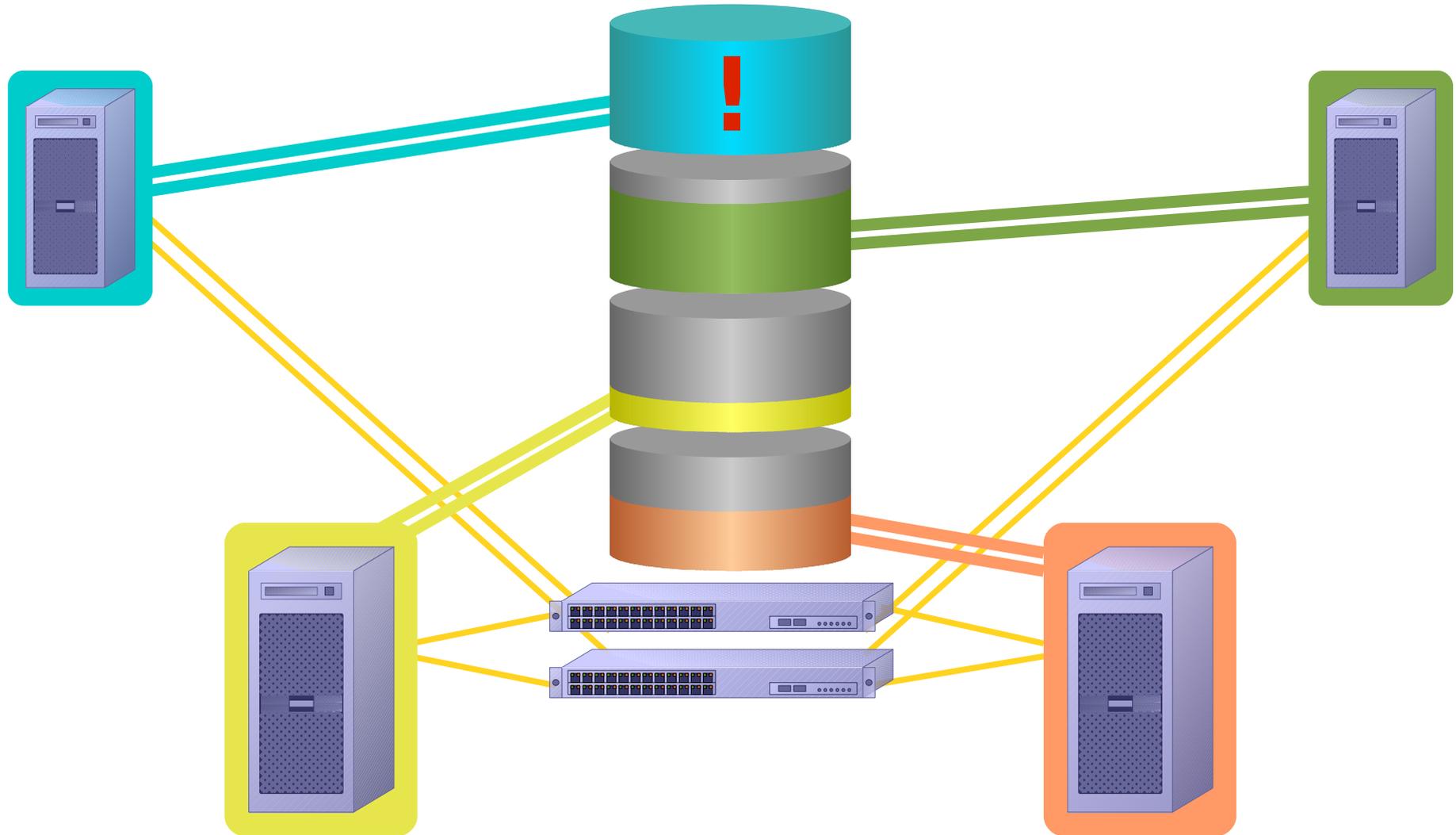


- einfacheres, sicheres Management, schlankere Hardware
- mehr Verfügbarkeit und Flexibilität
- Hardwareabstraktion \Rightarrow keine Angst vor Systemwechsel
- bessere Auslastung der Systemressourcen \Rightarrow mit weniger mehr erreichen

Funktionsübersicht

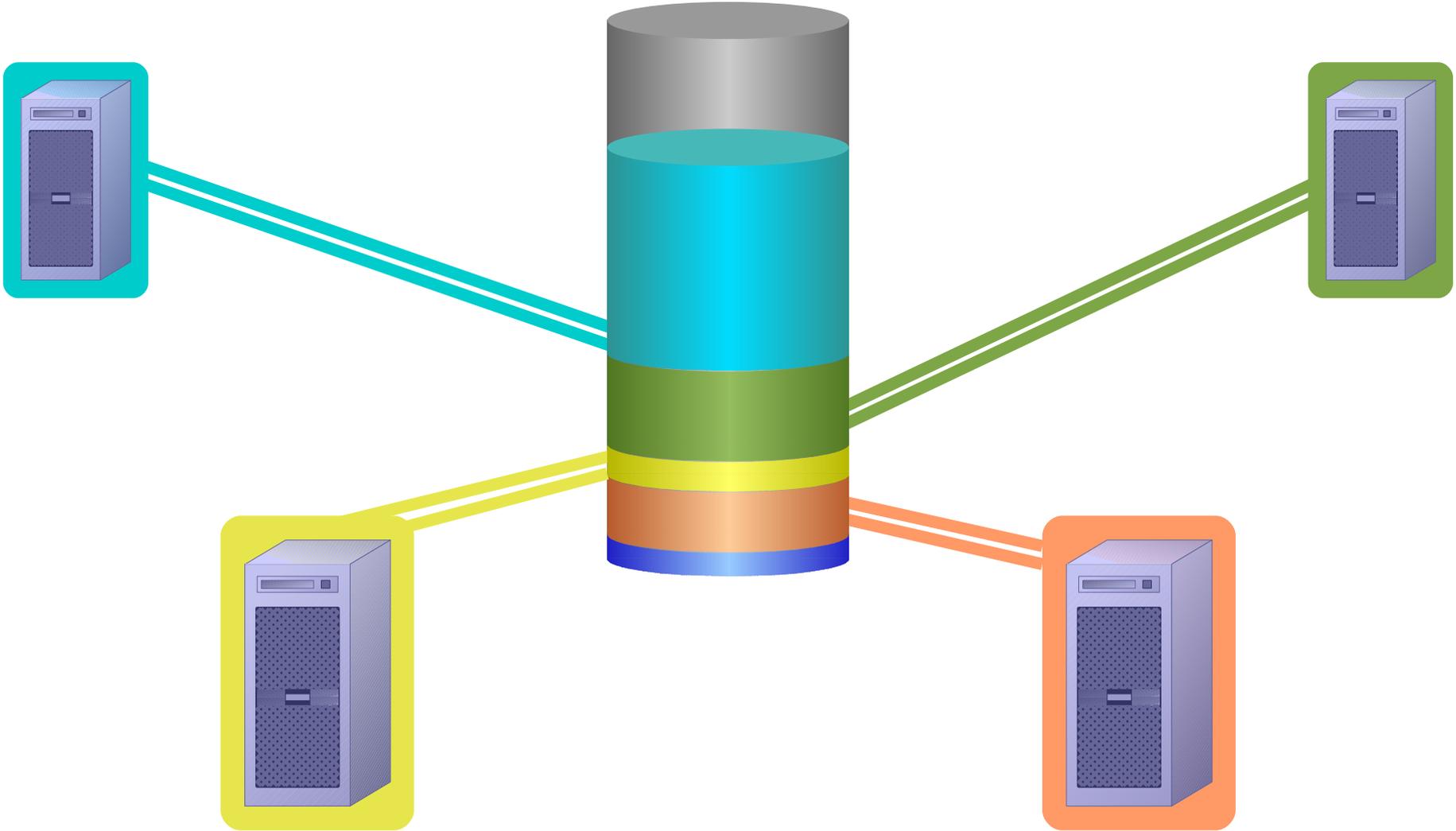
Speichervirtualisierung

Von der komplexen Ansammlung von Einzelbausteinen ...



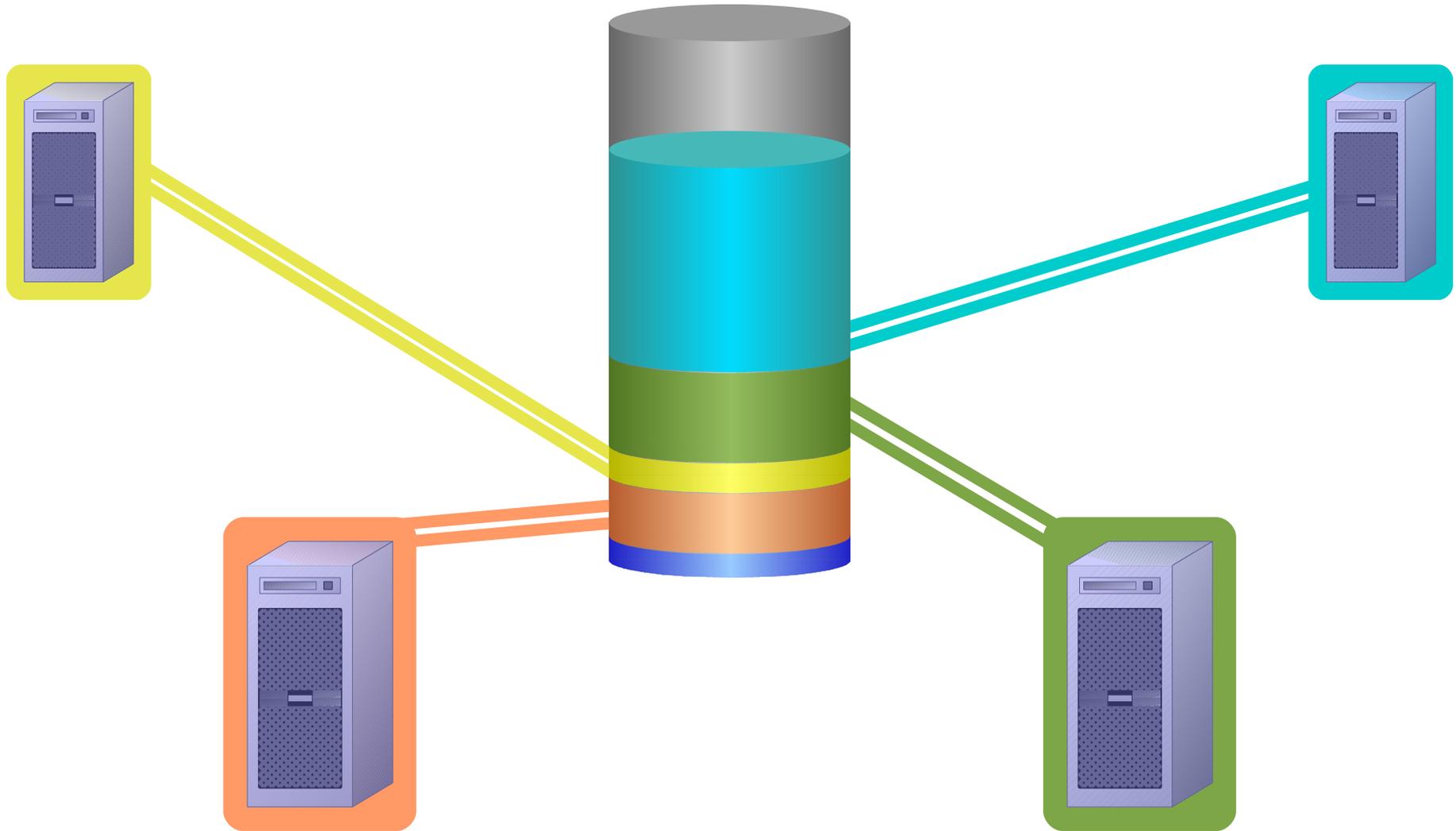
Speichervirtualisierung

... zur integrierten Sichtweise



Speichervirtualisierung

hostbasiert, hardwareabstrakt, flexibel, skalierbar, austauschbare Komponenten

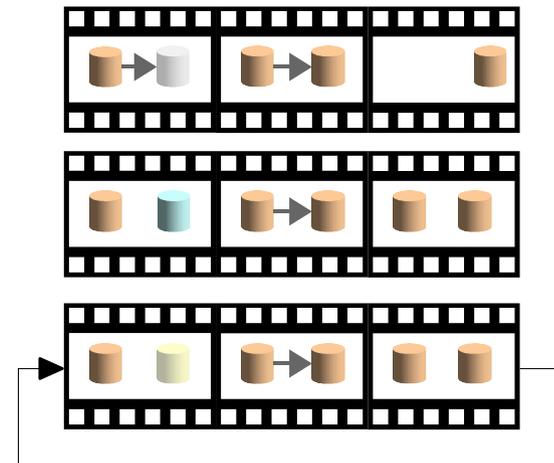
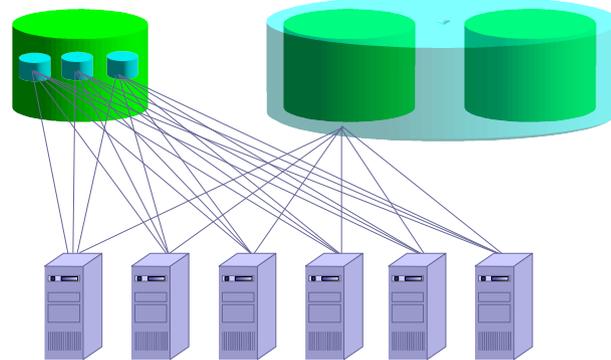


Speichervirtualisierung

Funktionen im Überblick



Basis-Virtualisierung
clusterweit
Globale Pools
Daten verschieben
Daten klonen
Daten spiegeln
Sonderfunktionen



keine
spezielle Hardware
erforderlich !

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

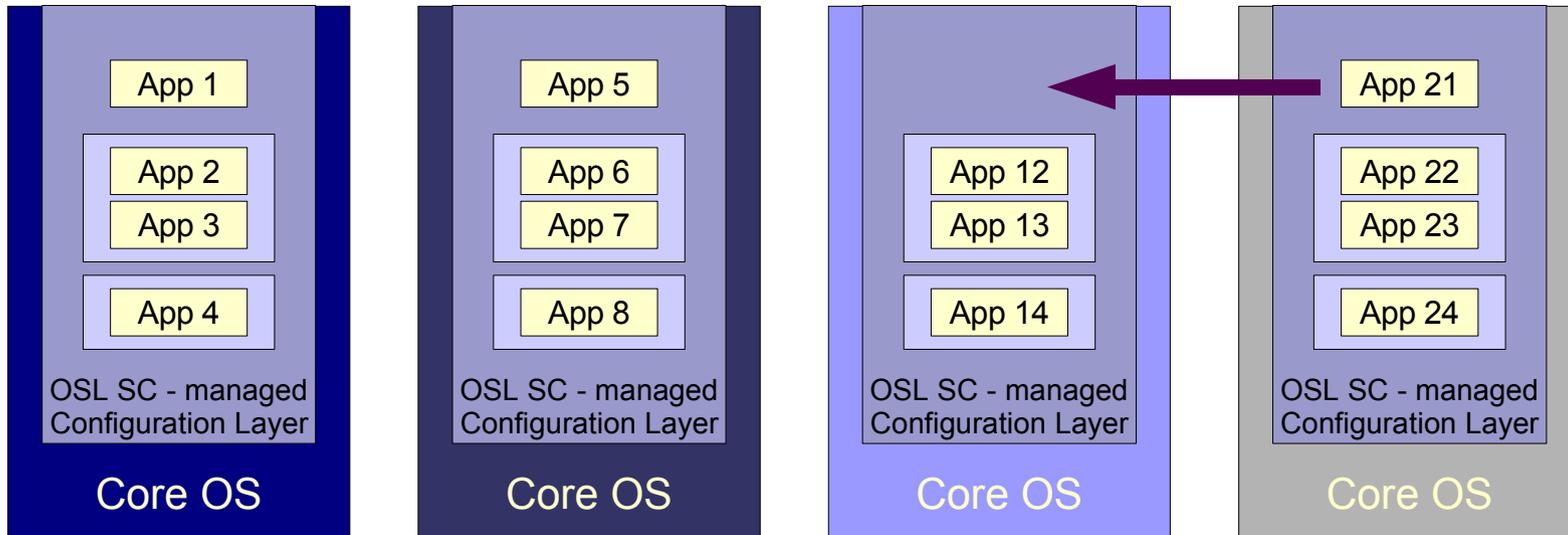
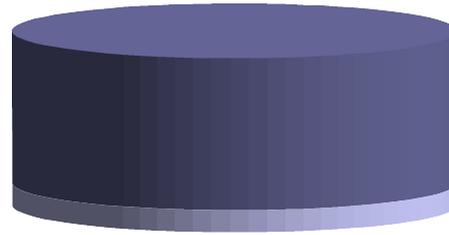
Online-Datenkopien auf wahlfreie Ziele
atomare Operationen für mehrere Volumes

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

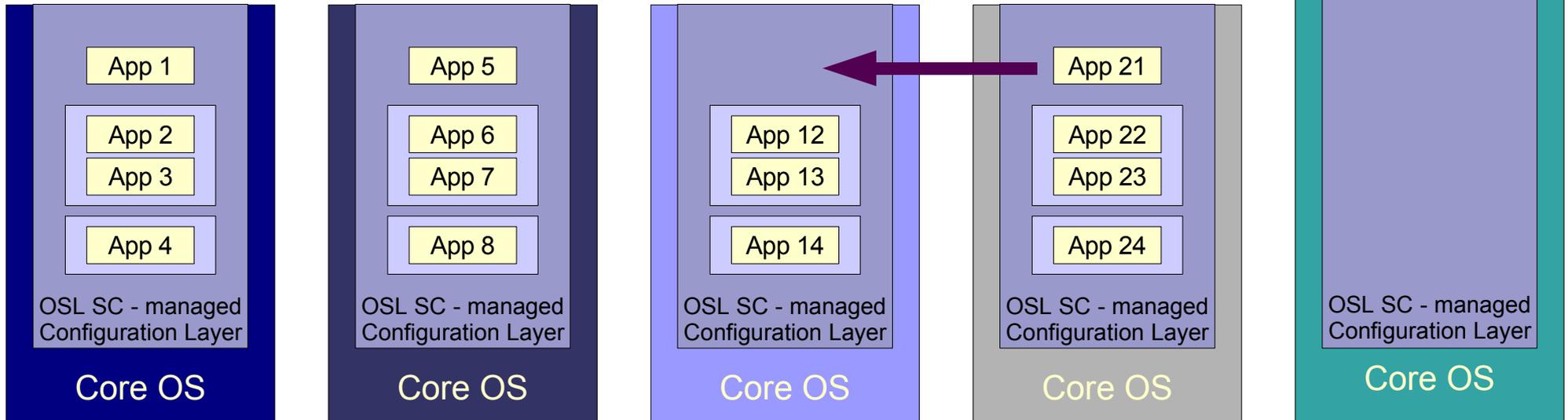
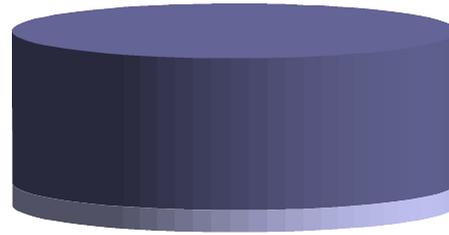
Hochverfügbarkeit

flexibel



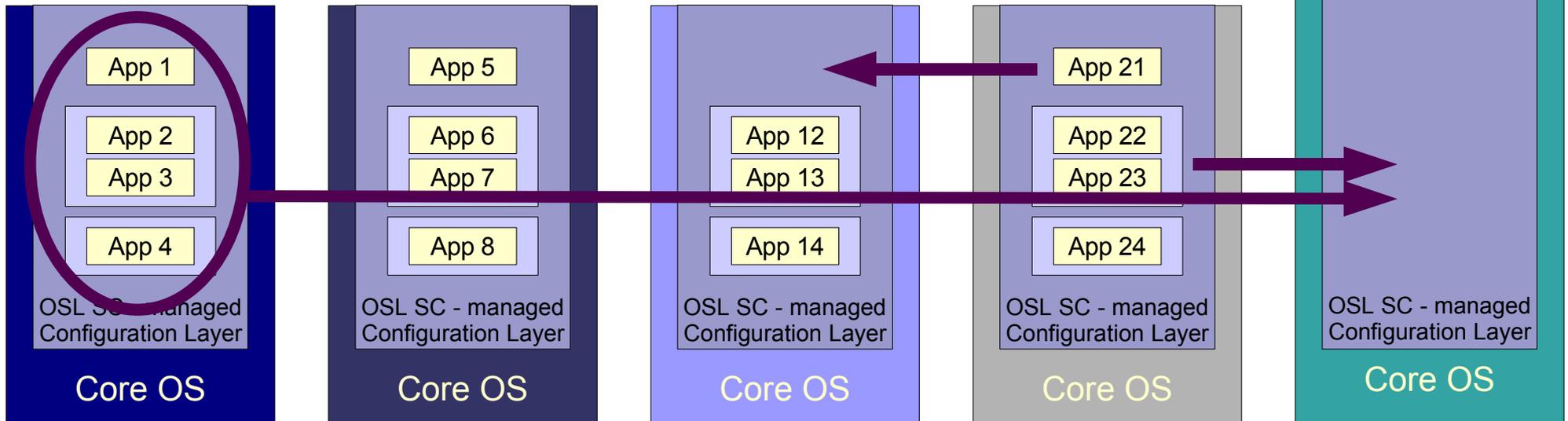
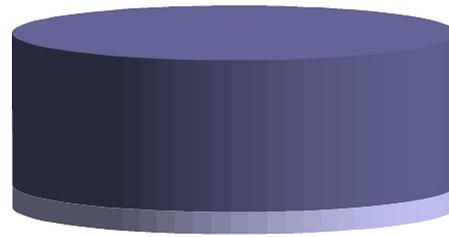
Hochverfügbarkeit

flexibel



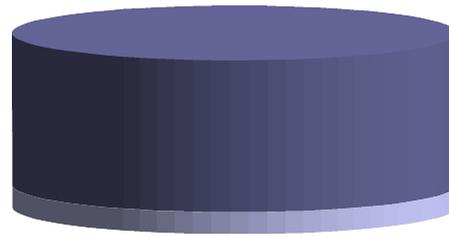
Hochverfügbarkeit

flexibel

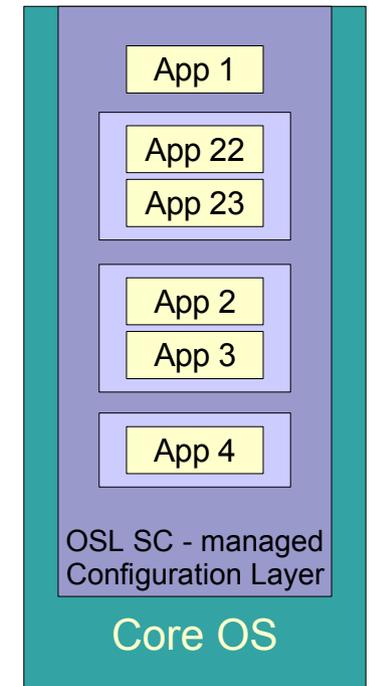
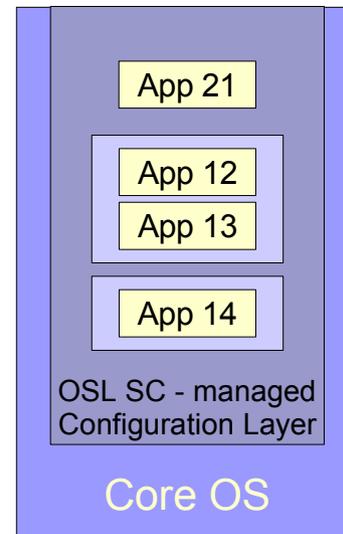
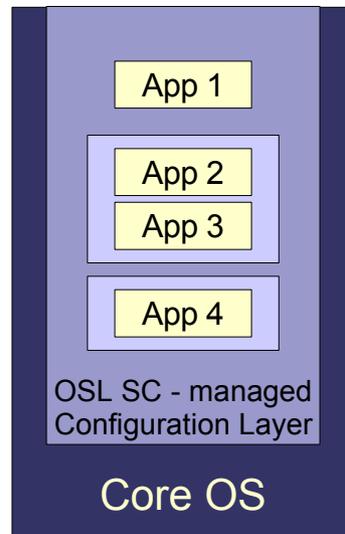


Hochverfügbarkeit

hardwareabstrakt, einfach zu bedienen



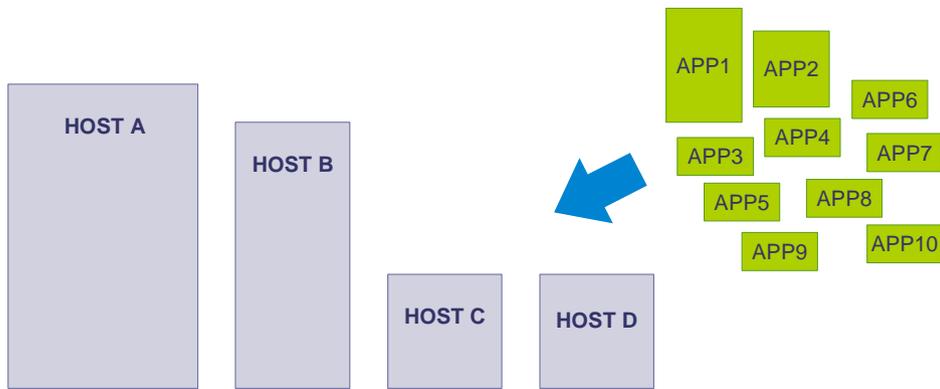
keine einzige
Änderung der
Cluster-Config



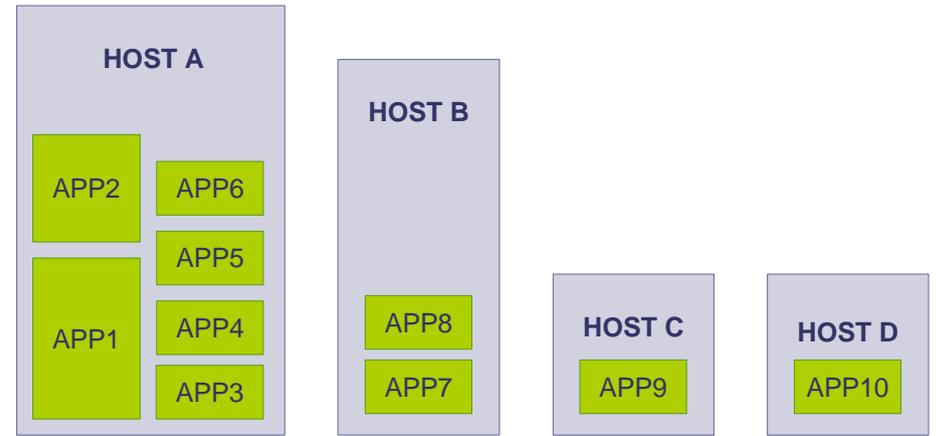
Core OS

Hochverfügbarkeit

selbstorganisierend

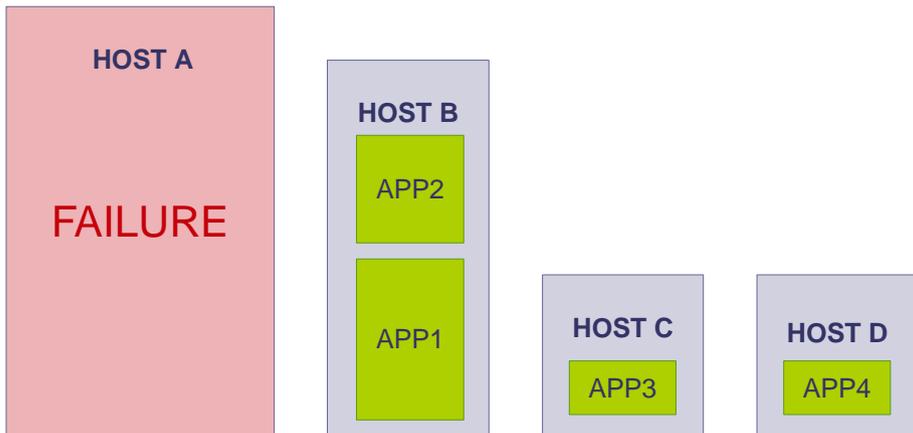


1

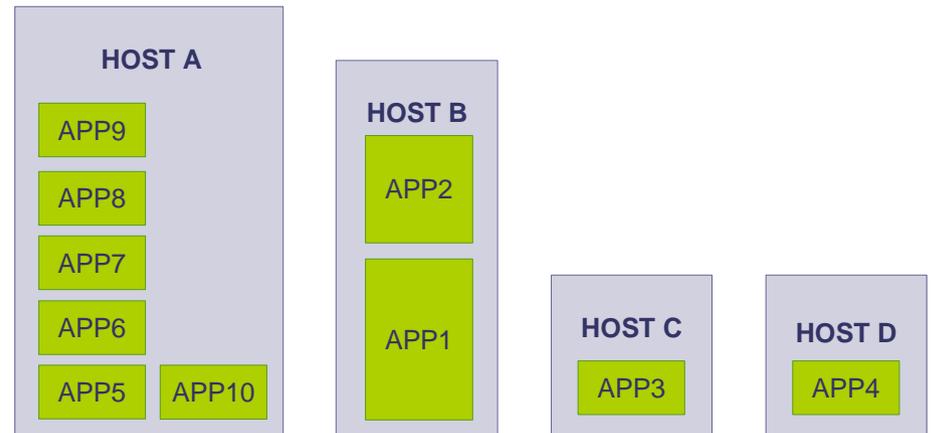


2

3



4



Was steckt dahinter?

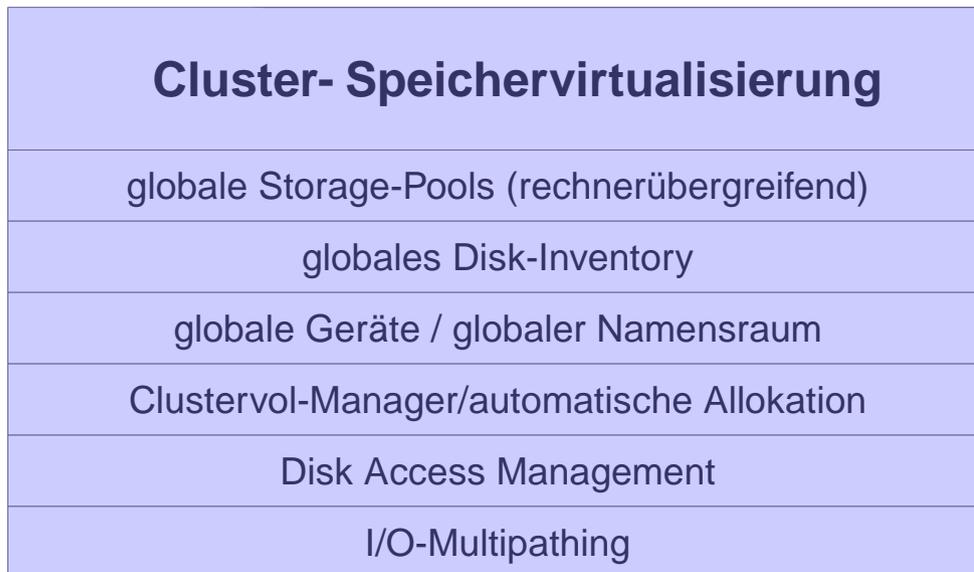
Ausgeklügelte, integrierte Gesamtarchitektur



- Anwendungsbeschreibung für alle Funktionen sichtbar machen
 - Speichervirtualisierung mit "Applikationsbewußtsein"
 - ⇒ enormer Funktionsgewinn
 - Selbstkonfiguration (z. B. Backup für komplexe Applikationen mit 1 Kommando)
- Trennung OS - Applikationen
 - wichtige Informationen leben unabhängig von Host und Speichersystem weiter
 - Eliminierung von Migrationsaufwänden
- Plattformunabhängigkeit
 - Mischung verschiedener Rechnerarchitekturen möglich
 - Mischung verschiedener Betriebssysteme möglich
 - Mischung verschiedener Connectivity-Lösungen möglich
- Intelligenz aus der Infrastruktur zum Host zurückholen
 - einfache Handhabung (Schulungsaufwände!)
 - höhere Verfügbarkeit
 - leichtere Migration

Ausgereiftes Portfolio

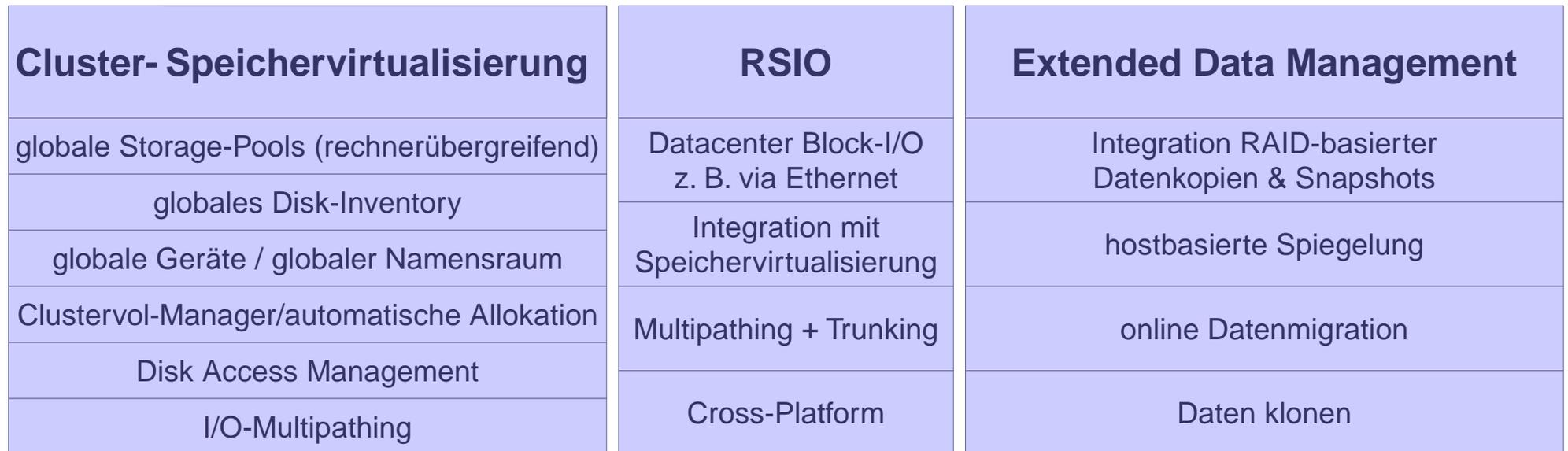
Widerspiegelung der Gesamtarchitektur im Produktpaket



Neuheiten

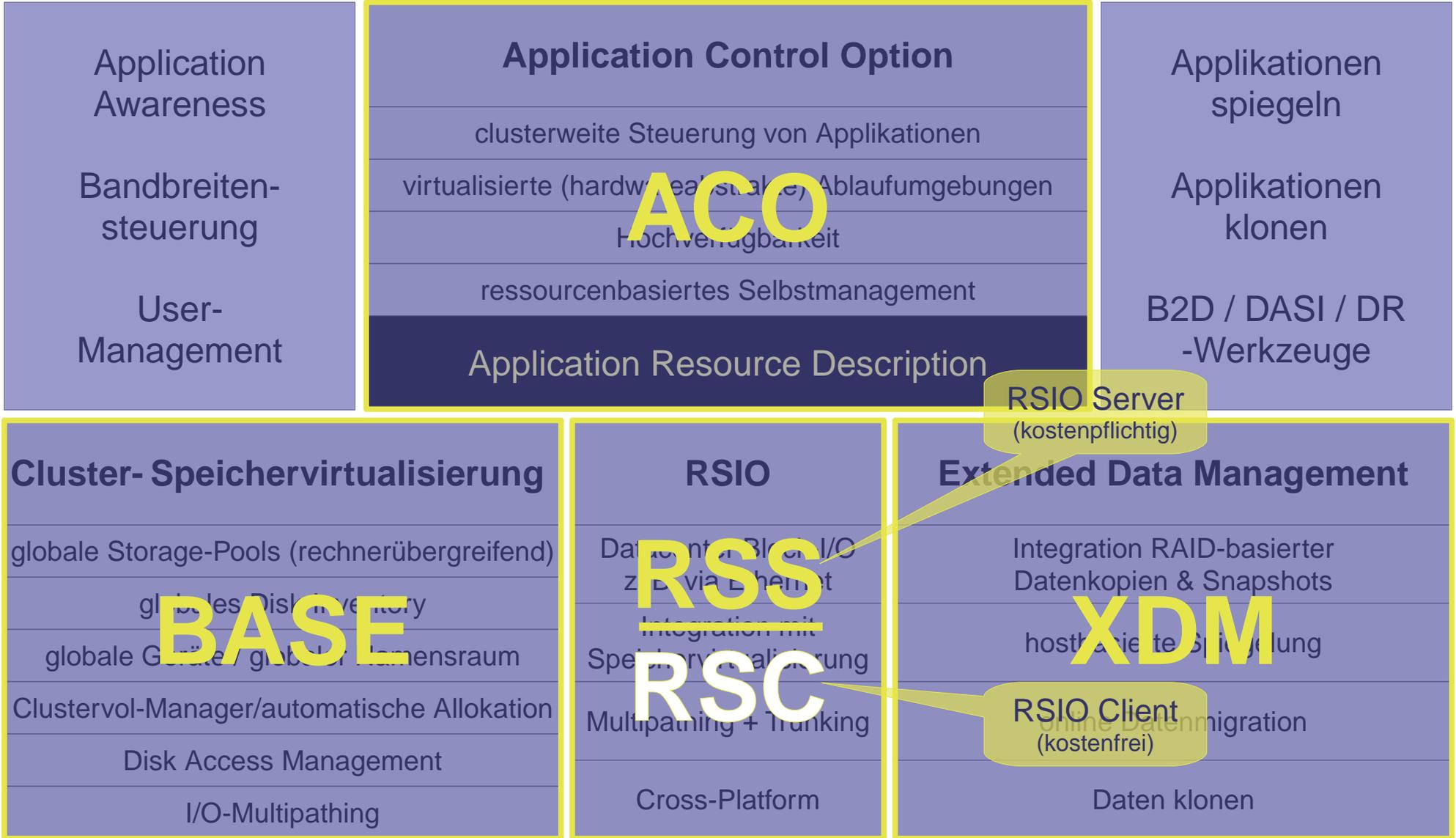
OSL Storage Cluster 4.0

Erweiterung des Portfolios



OSL Storage Cluster 4.0

Erweiterung des Portfolios



OSL Storage Cluster 4.0

Erweiterung des Portfolios – auch für andere Plattformen, etwa SLES



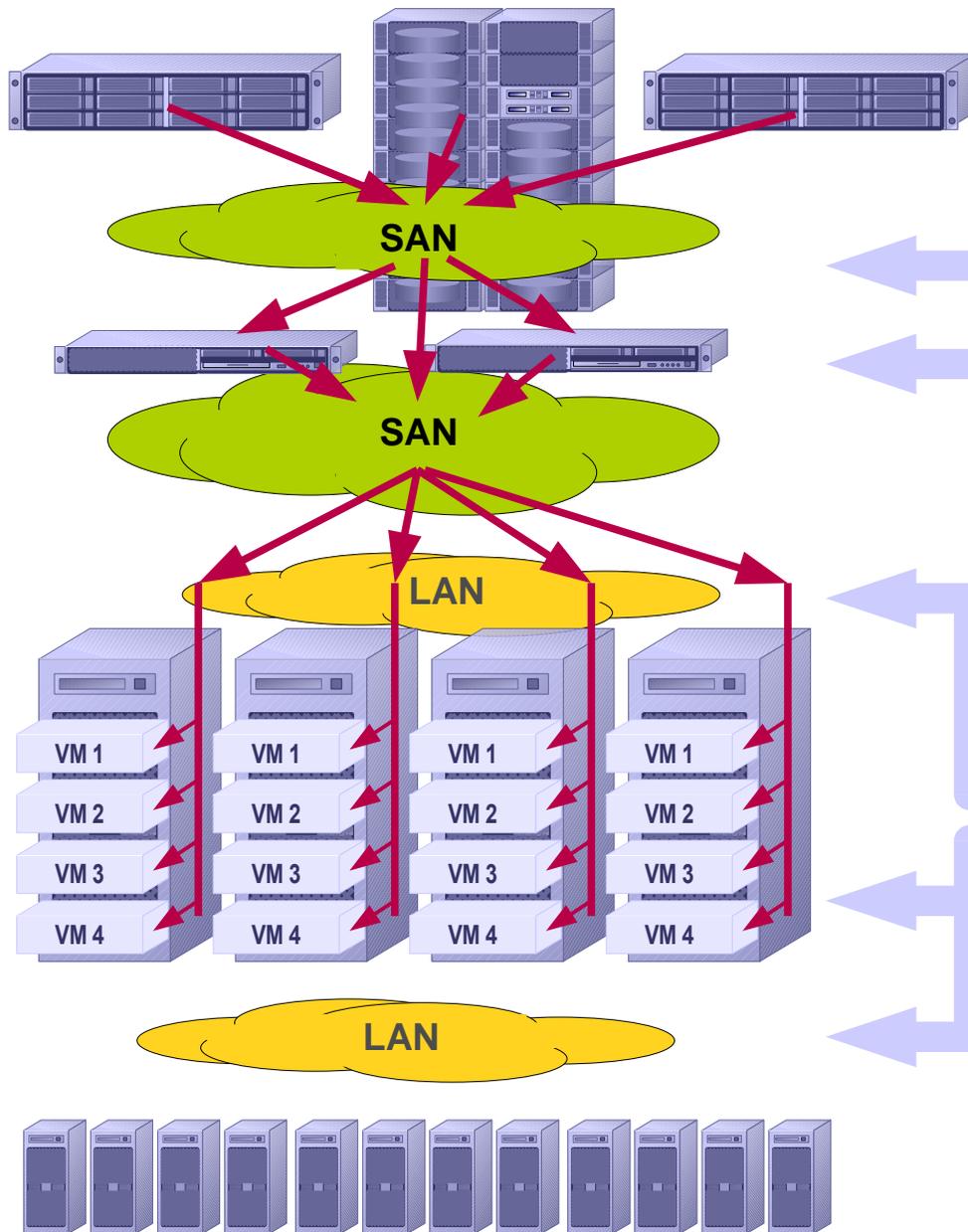
OSL Storage Cluster 4.0

Erweiterung des Portfolios – auch für andere Plattformen, etwa RHEL



Storage Virtualisierung und Netzwerke

Situation im heutigen Rechenzentrum



Storage Management

- Provisionierung
- Performance
- Bandbreite
- Verfügbarkeit

Speichervirtualisierung

- mit preiswerter HW mehr erreichen
- Verbess. Verfügbarkeit, Performance
- mehr Flexibilität, neue Möglichkeiten

SAN-Management

- Provisionierung
- Performance
- Verfügbarkeit

LAN-Management

- Performance
- Verfügbarkeit

OS-Virtualisierung

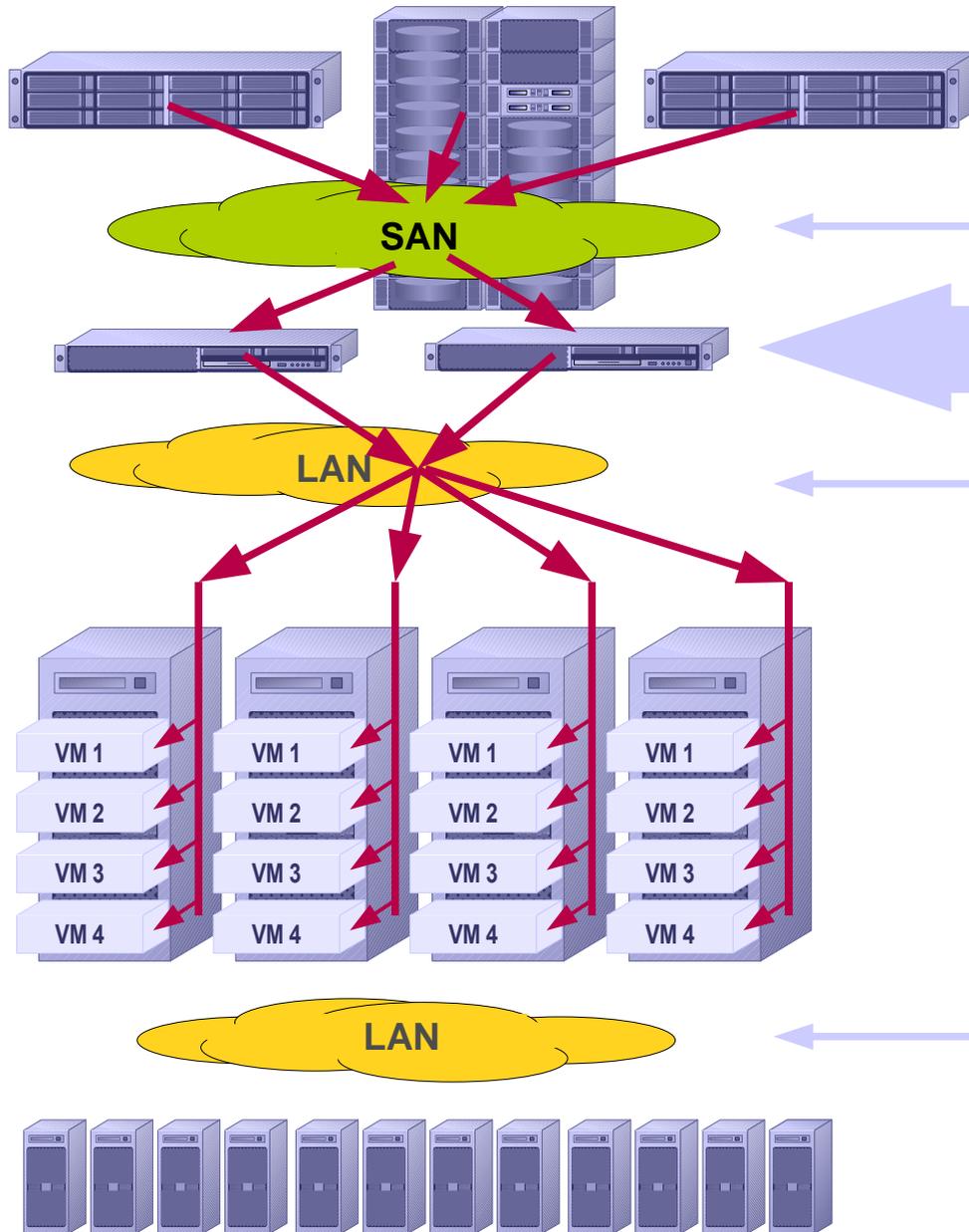
- SLAs
- Flexibilität
- Verfügbarkeit

Integration

- alles zusammen am Laufen halten
- Backup, Disaster Recovery ...

Storage Virtualisierung und Netzwerke

Gleiche Hardware – anderes Betriebskonzept mit anderer Software



Storage Management

- Provisionierung
- Performance
- Bandbreite
- Verfügbarkeit

Speichervirtualisierung

- mit preiswerter HW mehr erreichen
- Verbess. Verfügbarkeit, Performance
- mehr Flexibilität, neue Möglichkeiten

SAN-Management

- Provisionierung
- Performance
- Verfügbarkeit

LAN-Management

- Performance
- Verfügbarkeit

OS-Virtualisierung

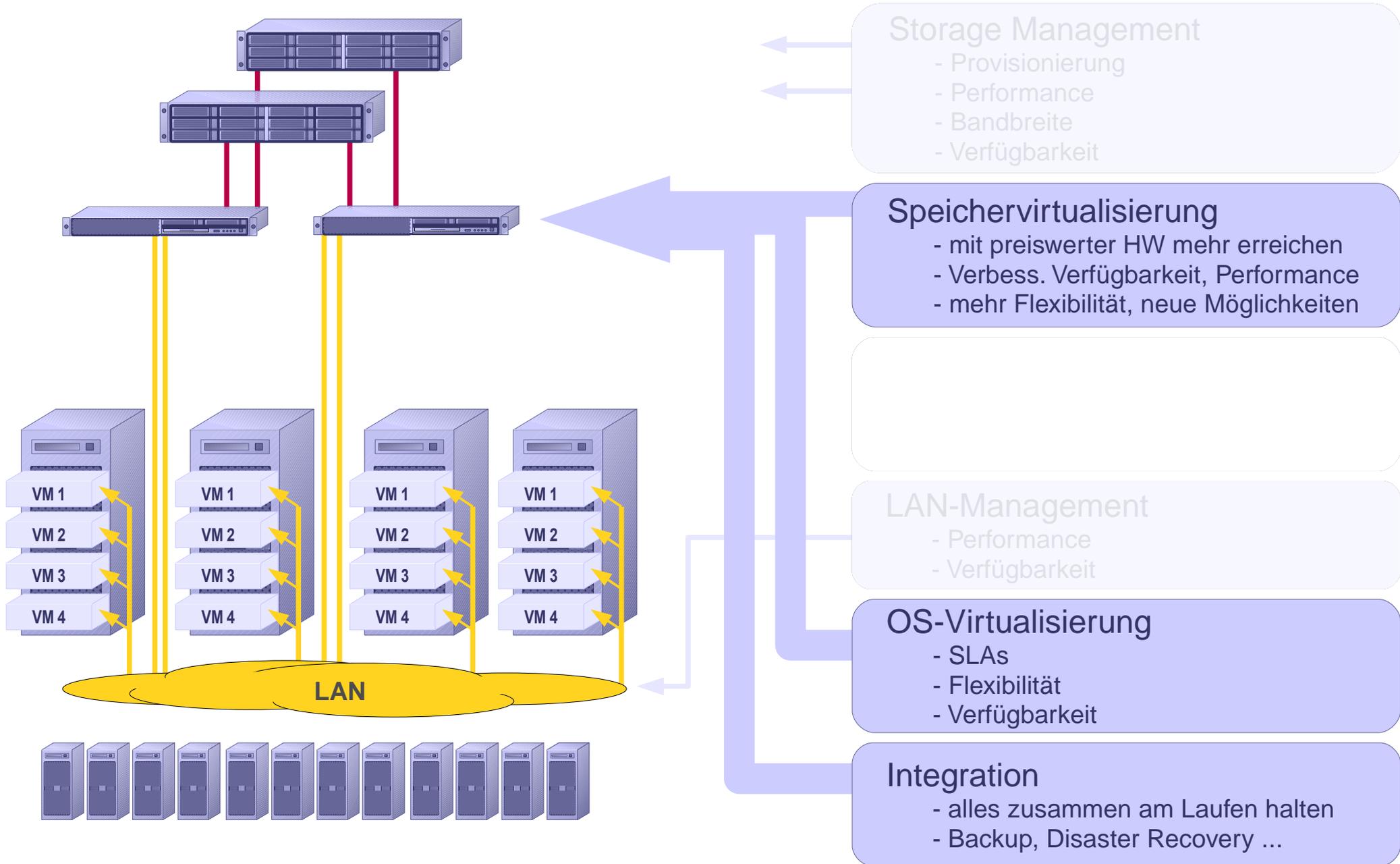
- SLAs
- Flexibilität
- Verfügbarkeit

Integration

- alles zusammen am Laufen halten
- Backup, Disaster Recovery ...

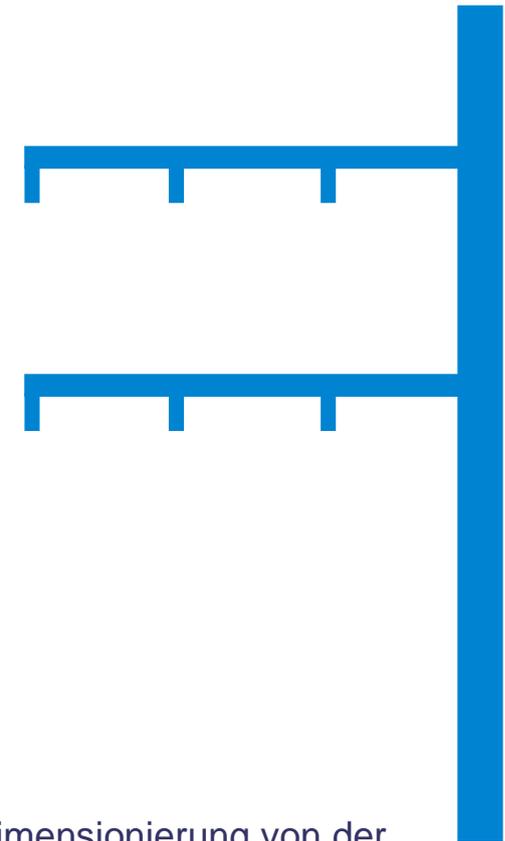
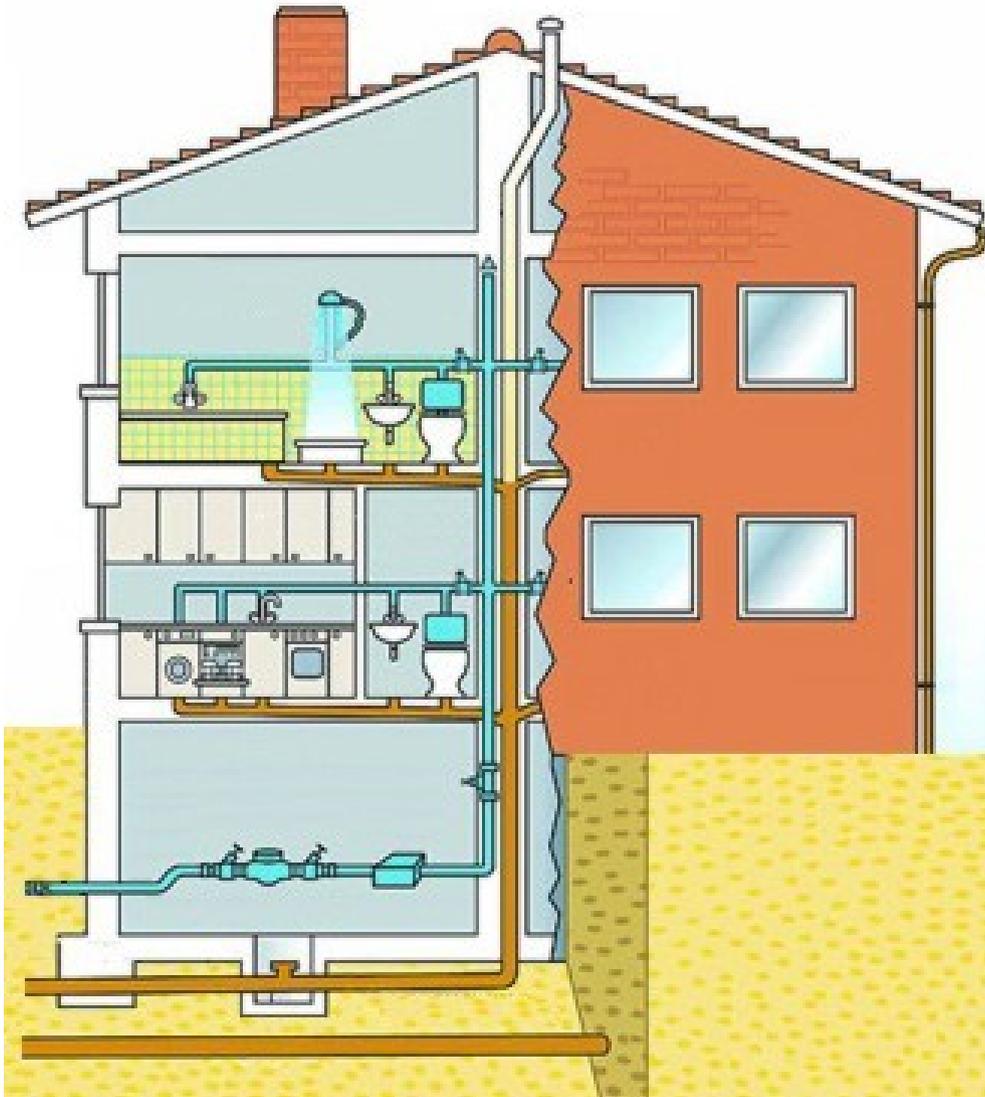
Die gleiche Idee – nur konsequenter: RSIO

Vereinfachung von Infrastruktur bis Management



Kann ich das denn machen?

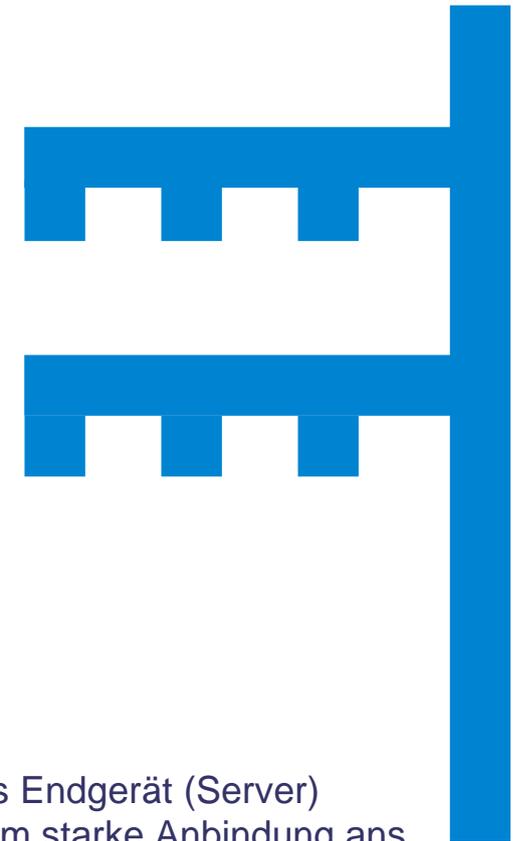
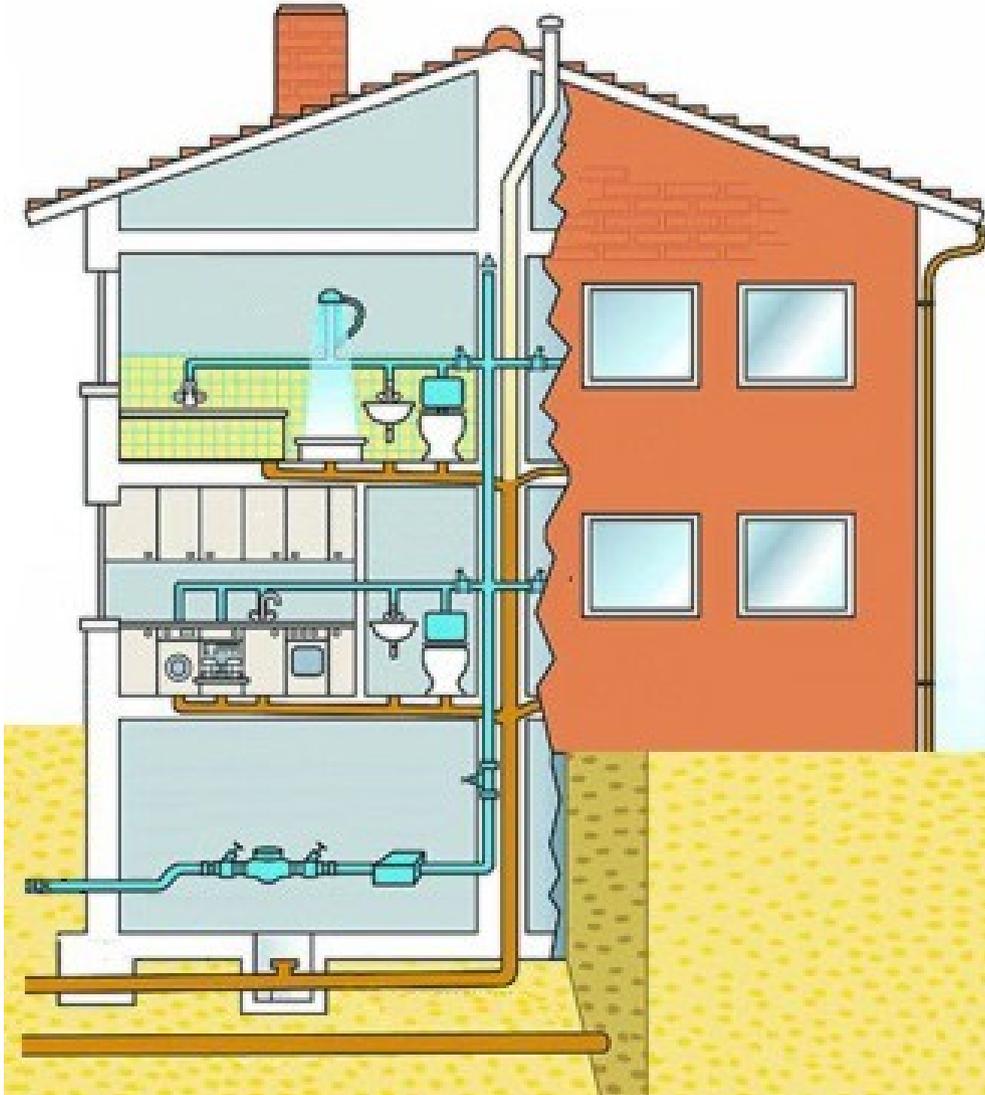
Sind heutige RZ-Infrastrukturen richtig dimensioniert ? Ein Vergleich ...



abgestufte Dimensionierung von der Zentralen Versorgung bis hin zu den Abnehmern an der Peripherie

Würden Sie sich so etwas bauen lassen?

Sind heutige RZ-Infrastrukturen richtig dimensioniert ? Ein Vergleich ...



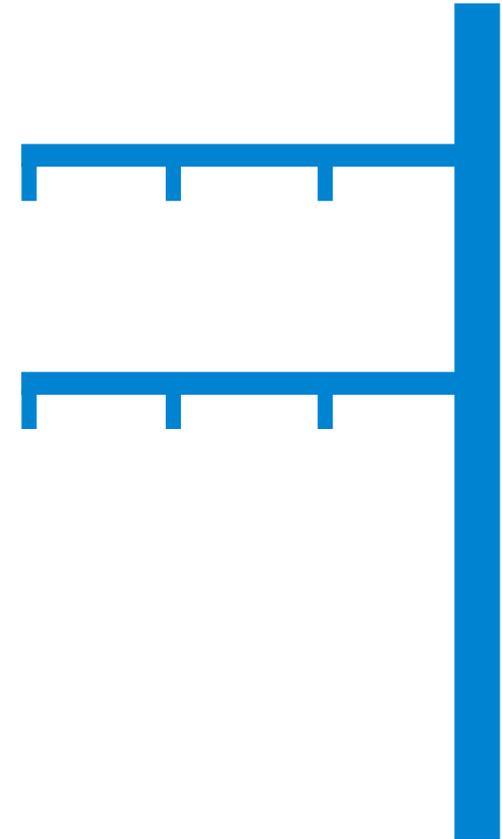
Im RZ für das Endgerät (Server) oft eine extrem starke Anbindung ans SAN im Vergleich zur Anbindung des zentralen Storage systems.

Natürlich ist Rightsizing vernünftig!

Sind heutige RZ-Infrastrukturen richtig dimensioniert ? Ein Vergleich ...



- absolut ausreichende Performance
 - 2 x 1Gbit LAN ⇒ bis zu 230 MiB/s **pro Node**
 - 4 x 1Gbit LAN ⇒ bis zu 460 MiB/s **pro Node**
 - 2 x 10Gbit LAN ⇒ über 500 MiB/s **pro Node**
 - 2 x 40Gbit LAN ⇒ >> 1000 MiB/s **pro Node**
 - Backup muß heute nicht mehr über's Netz
- keine Hotspots / Überlastungen
- höhere Portdichten, ggf. Senkung Energieverbrauch
- niedrigere Ausfallraten
- einfachere Handhabung
- alles über ein Netz!
- enorme Kostensenkungen in der Infrastruktur

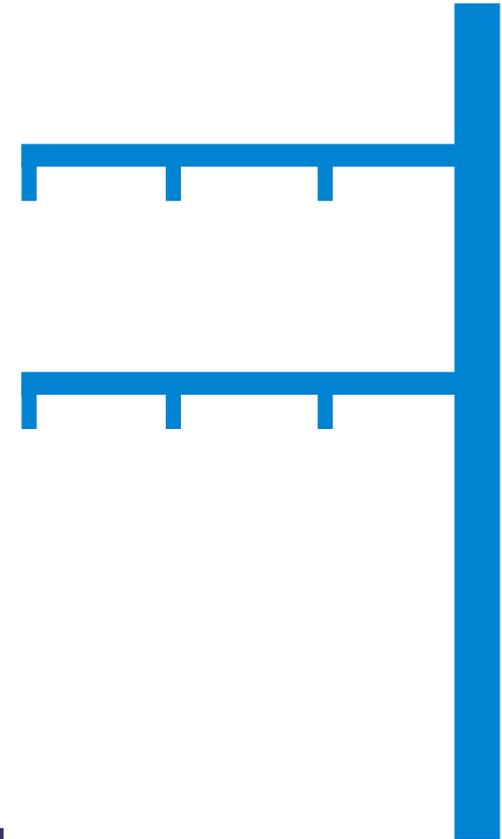


Natürlich ist Rightsizing vernünftig!

Sind heutige RZ-Infrastrukturen richtig dimensioniert ? Ein Vergleich ...



- absolut ausreichende Performance
 - 2 x 1Gbit LAN ⇒ bis zu 230 MiB/s **pro Node**
 - 4 x 1Gbit LAN ⇒ bis zu 460 MiB/s **pro Node**
 - 2 x 10Gbit LAN ⇒ über 500 MiB/s **pro Node**
 - 2 x 40Gbit LAN ⇒ >> 1000 MiB/s **pro Node**
 - Backup muß heute nicht mehr über's Netz
- keine Hotspots / Überlastungen
- höhere Portdichten, ggf. Senkung Energieverbrauch
- niedrigere Ausfallraten
- einfachere Handhabung
- alles über ein Netz!
- enorme Kostensenkungen in der Infrastruktur



... wenn ich das richtige Protokoll habe!

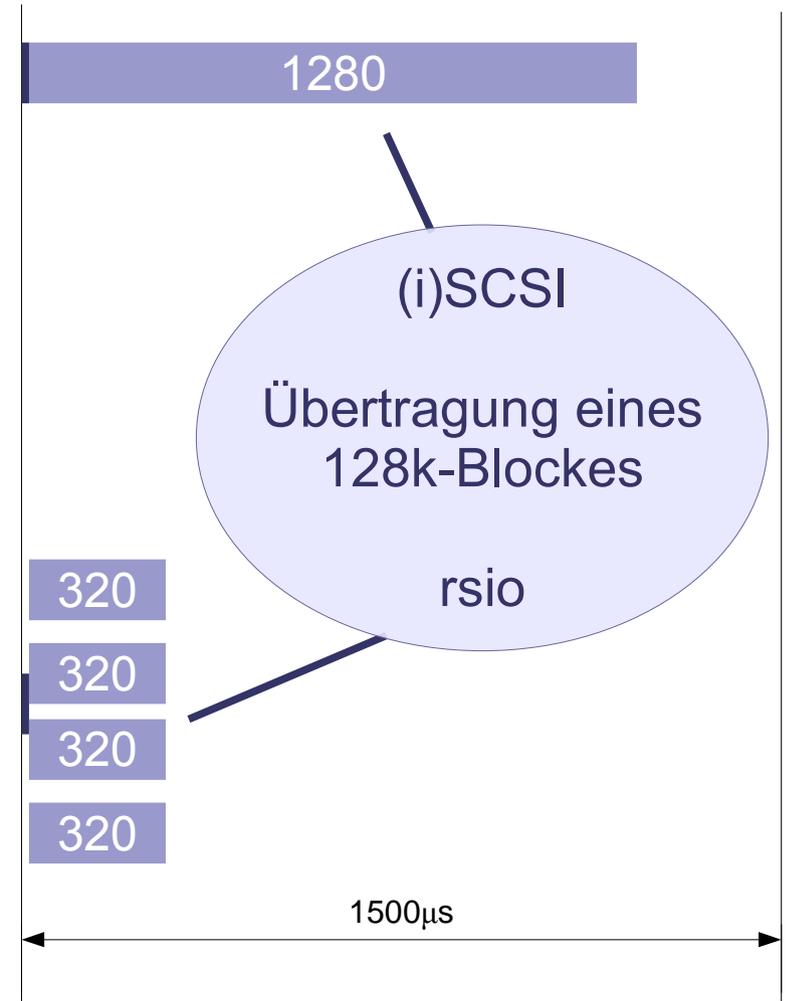
Block-I/O neu definiert

RSIO – Remote Storage I/O



Vorteile der RSIO - Technologie

- Definition eigener Frames
 - Unabhängigkeit von TCP
 - ermöglicht erst **Scatter-Gather-Multipathing**
 - ermöglicht Zusatzfunktionen wie Checksum / Encryption
 - Frames mit variabler Größe
 - Overhead per Frame nur 16 Byte
- Trennung von Treiber und Transport
 - größerer Funktionsumfang bei hoher Portabilität
 - besseres Error-Handling
 - Performance offensichtlich kein Problem
 - hochflexibler Multithreading-Support
 - bessere Abschirmung des Kerns
- Integriertes Multipathing und Trunking
- Selbstkonfiguration und Error Recovery
- Unterstützung geclusterter Server
- vollständige Abbildung relevanter Schnittstellen (dkio)
- Performance-"Nachteile" (Latenz $250\mu\text{s}$ 10GE vs. $130\mu\text{s}$ FC) werden je nach Kontext kompensiert oder überkompensiert



Nochmal Performance

Was RSIO bereits mit 1 GBit-LAN leistet



Server-Performance bei Cache Read / 8k

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

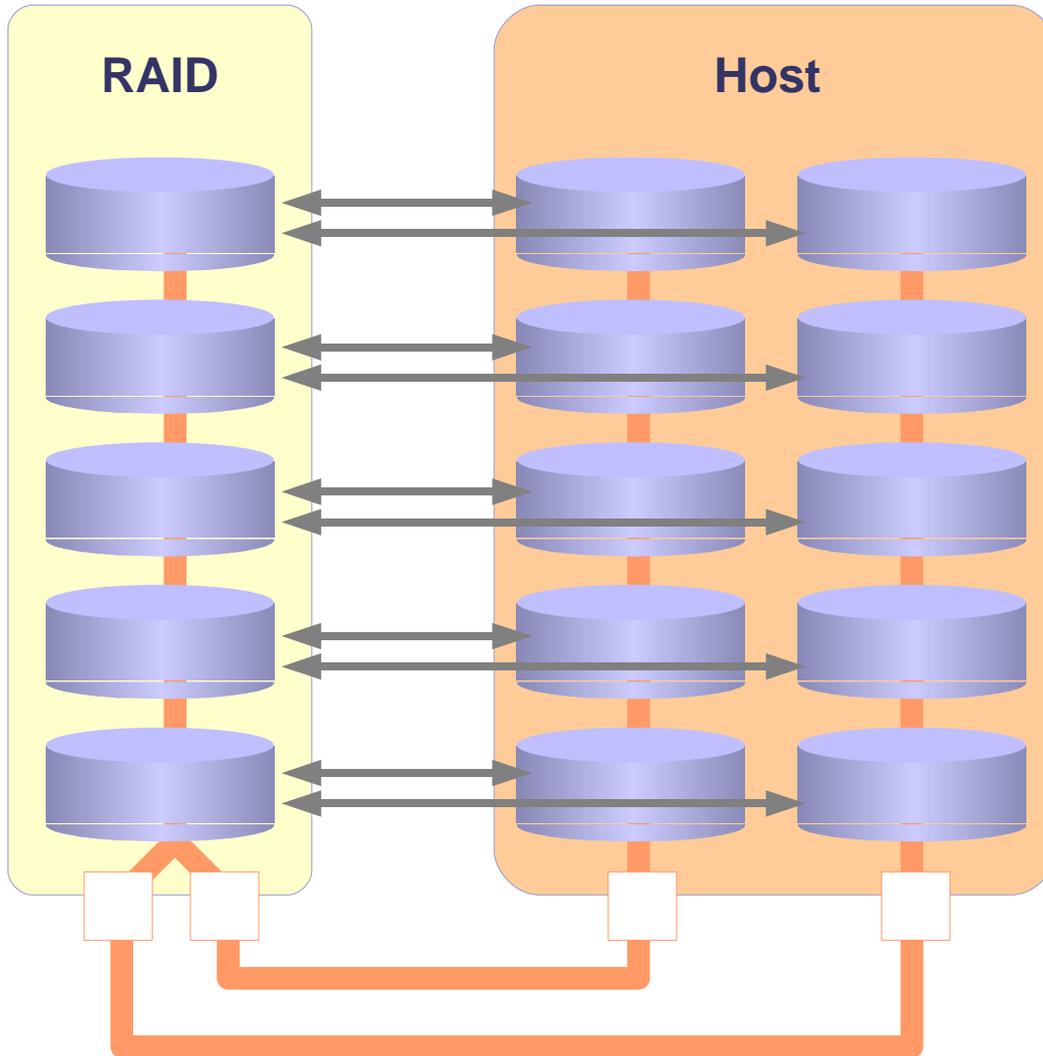
Client-Performance Throughput

RSIO	1 x 1 GBit	ca. 0,5 Cores	> 110 MByte/s
RSIO	2 x 1 GBit	ca. 1,0 Cores	> 220 MByte/s
RSIO	4 x 1 GBit	ca. 2,0 Cores	> 440 MByte/s
RSIO	8 x 1 GBit	> 4,0 Cores	bis > 900 MByte/s

Mehr als 10.000 IOPS je Client sind in heutigen Speichernetzen die absolute Ausnahme!

Das Wichtigste: Ein völlig anderes Konzept

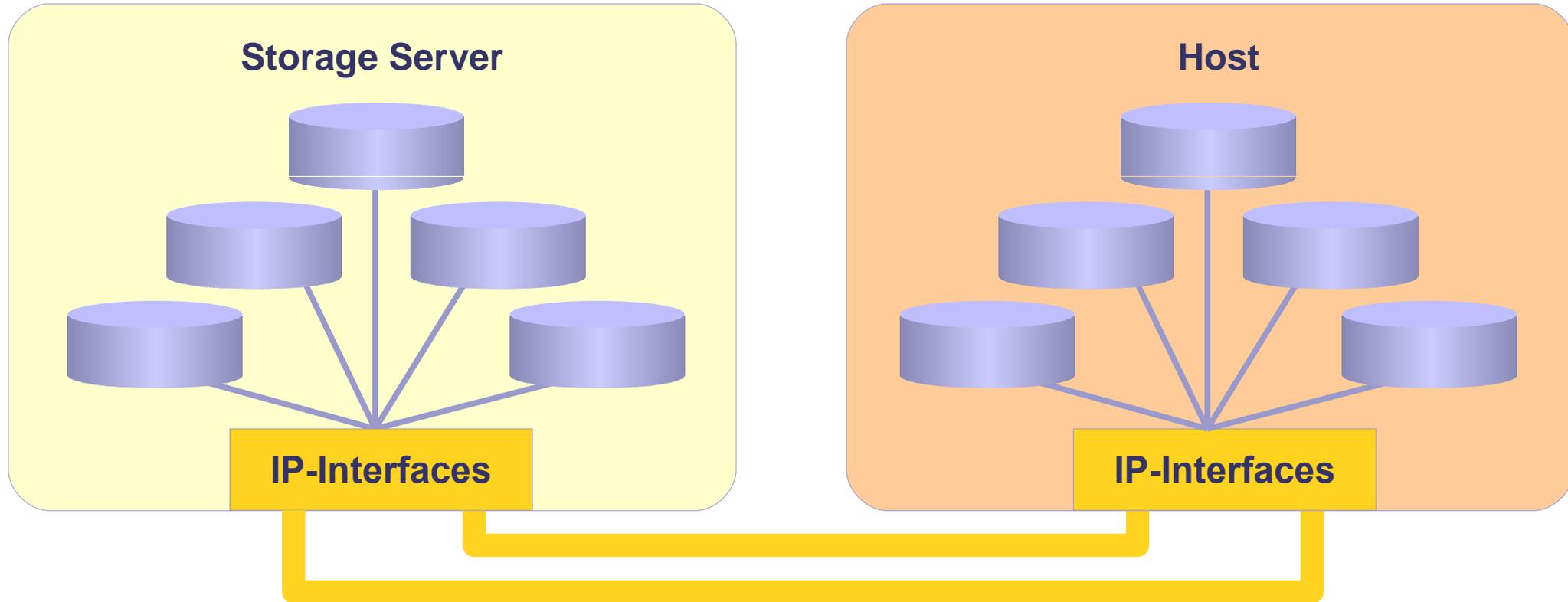
Warum die Welt mit SCSI / Fibre Channel so komplex sein kann



- 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

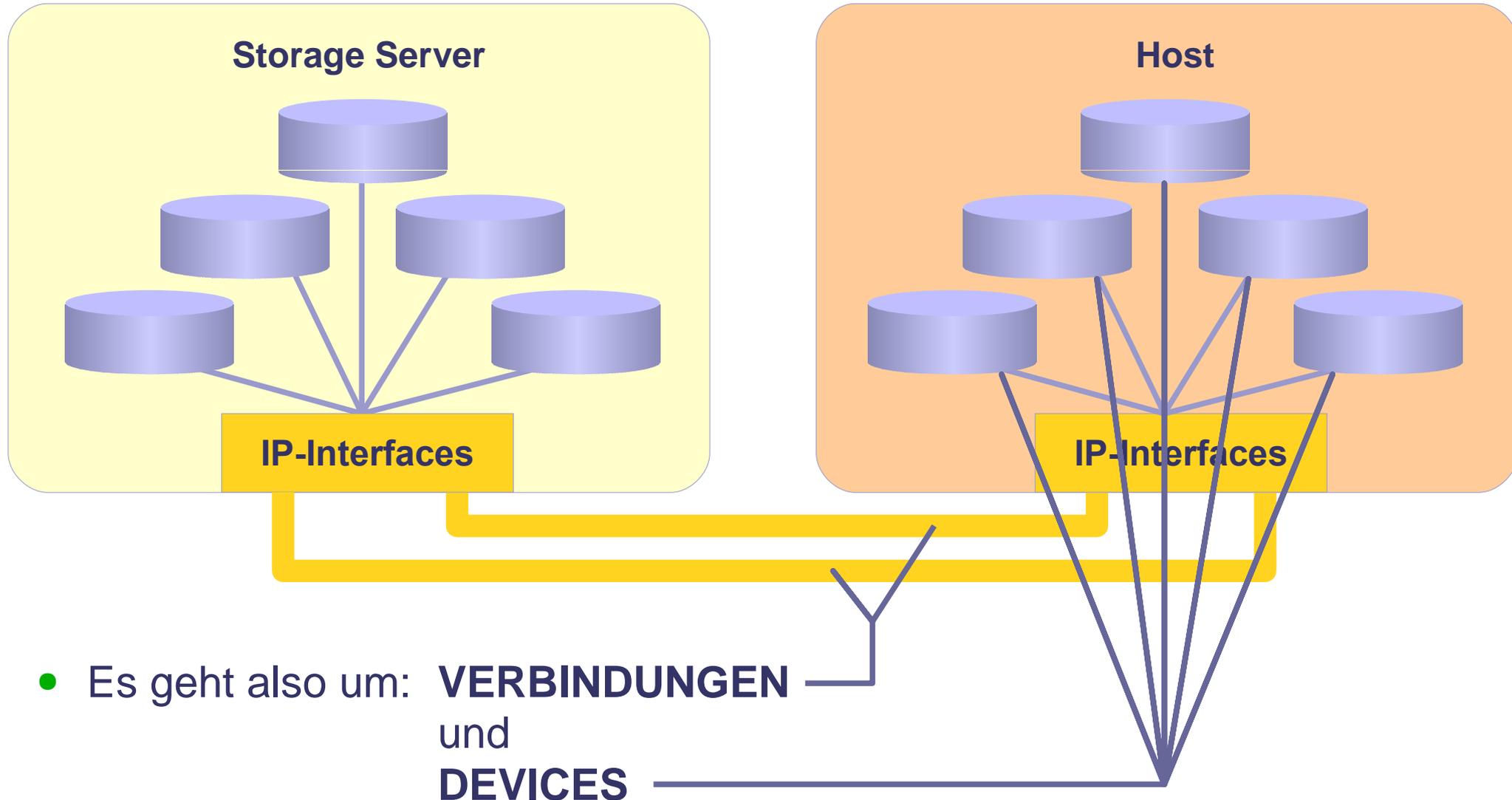
Mehr Übersichtlichkeit und völlig neue Möglichkeiten



- 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.

Das Wichtigste: Ein völlig anderes Konzept

Mehr Übersichtlichkeit und völlig neue Möglichkeiten

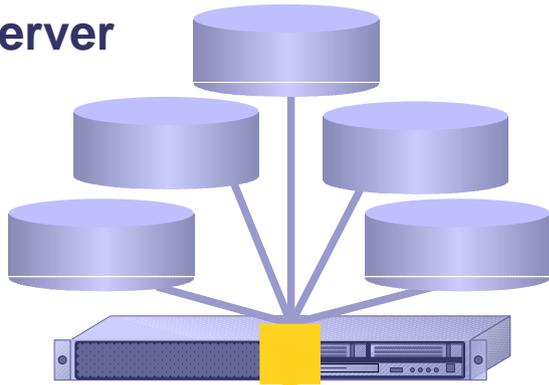


Im Detail

Überall wo IP ist: Storage in 5 Minuten!



RSIO Server

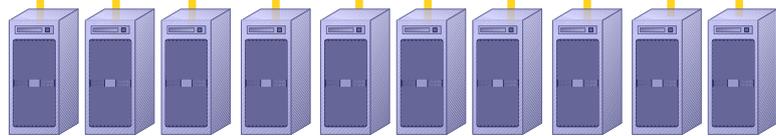
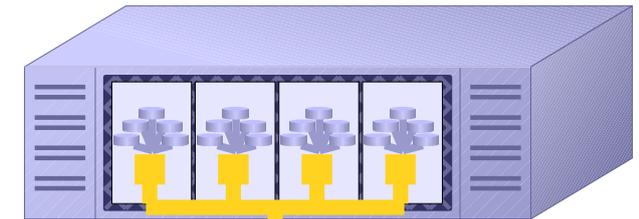
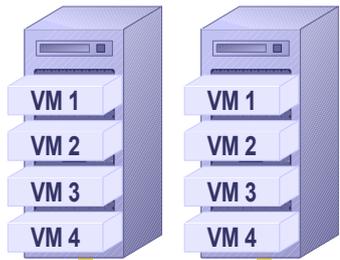
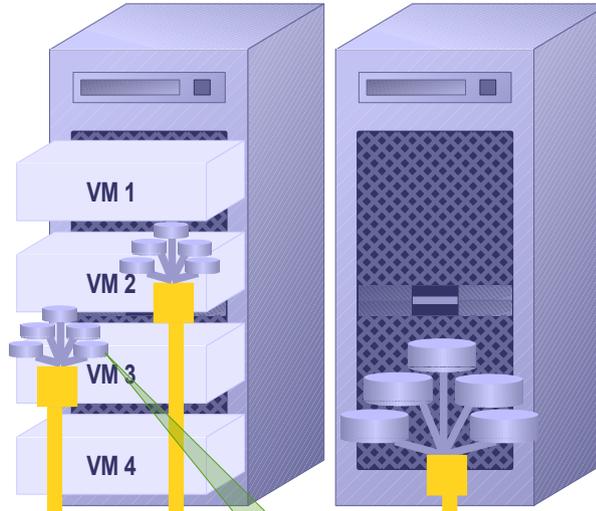


RSIO Clients

Virtual Machines

Physical Machines

Blade Servers



Auf Wunsch tausende Devices in der VM - sofort!

Was bringt RSIO noch?

Mehr als Block-I/O



- Verknüpfung mit Speichervirtualisierung
 - nie mehr fdisk ...
 - ready-to-use Devices
 - Spiegelung, Clones, DR-Konfigurationen, Backup-to-Disk
 - Clusterfähigkeiten
 - vergessen Sie umständliche Gerätenamen -> Sie als Admin sind Chef!
- Netzwerkkonvergenz mit heutigen, preiswerten Komponenten
- Lösung für spezielle Probleme (z. B. komplexe Konfigurationen in Linux-VMs)
- große Zahl von Servern – große Zahl von Clients beherrschbar
- viele Konfigurationsmöglichkeiten
- direkte Anbindung für Solaris und Linux
- über OS-Virtualisierung beliebige Betriebssysteme erreichbar

Weitere wichtige Neuerungen SC 4.0

Bewährtes weiter verbessern



- RSIO Server und Simple Client für Solaris und Linux
- volle Unterstützung für Linux
 - Virtual Storage via RSIO
 - Clusterframework
 - Base, XDM, ACO (vom Admin-Cluster über HA bis zu Backup und DR)
- besonderer Support für große Cluster
 - von bisher 32 auf 64 Knoten erweitert
 - 2048 AVs pro Universum (bei 4 Universen insgesamt 8192)
 - verbesserter Disk-Access-Manager
- differenziertere Applikationszustände und Application Variable Data
- verbesserte Intergration Multi-Volume-Filesysteme
- zahlreiche Verbesserungen für Solaris Zones
- einheitliches Interface für virtuelle Maschinen
- Virtual Nodes

Virtual Machines und Virtual Nodes

Der Cluster in der VM



Applications

Virtual Machine Applications

Solaris Zones

VirtualBox

KVM

andere

Virtual Nodes

Solaris Zones

VirtualBox

KVM

andere
(z. B. XEN)

Virtual Machines und Virtual Nodes

Der Cluster in der VM



Applications

Virtual Machine Applications

Einheitliche
Bediener-
schnittstelle

Solaris Zones

VirtualBox

KVM

andere

Virtual Nodes

Solaris Zones

VirtualBox

KVM

andere
(z. B. XEN)

Der Virtual Node hat Zugriff auf alle Clusterfunktionen einschl. Speichervirtualisierung und Anwendungssteuerung – damit Speichervirtualisierung auch in der VM / Solaris Zone!

V3

OSL

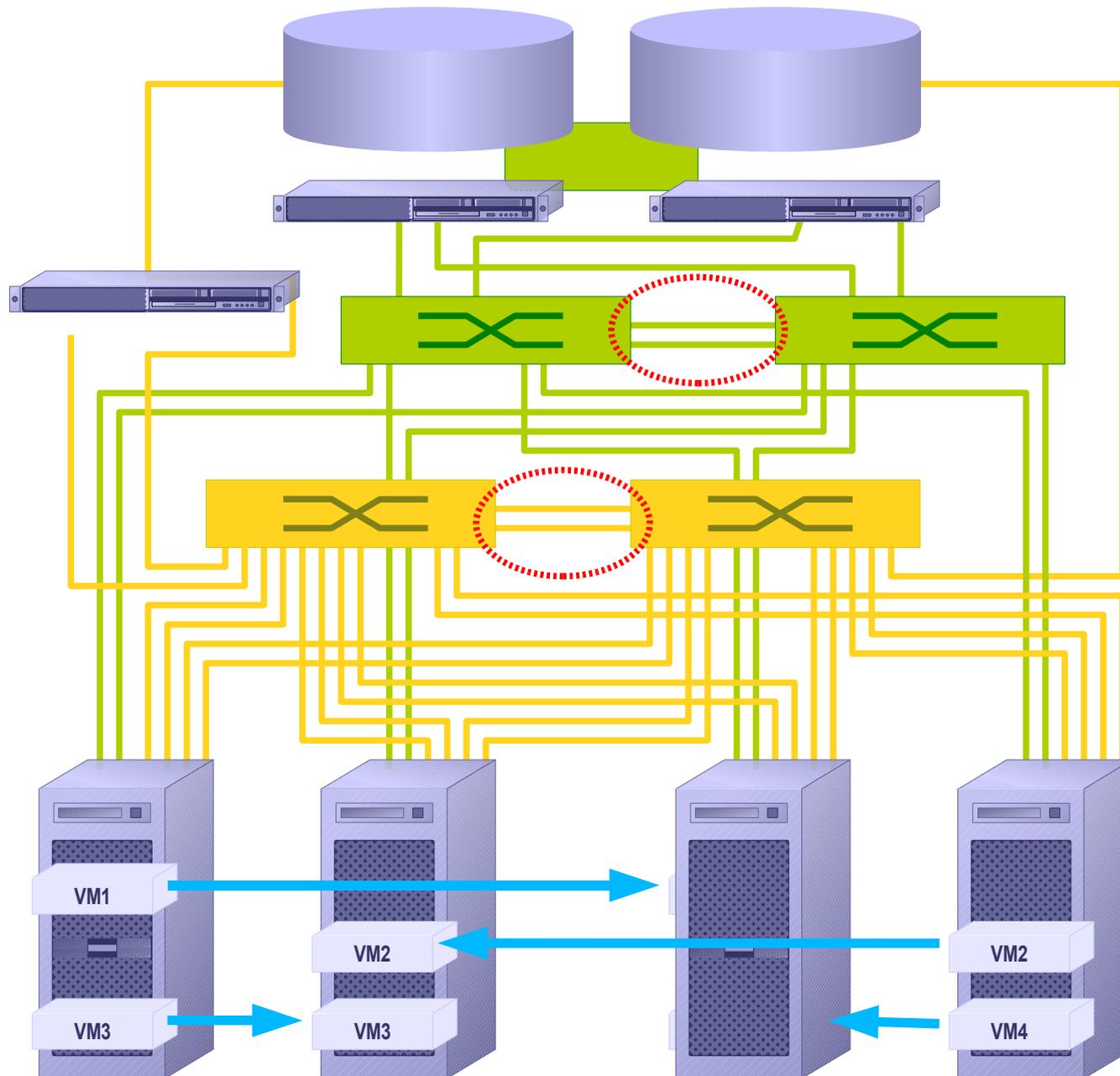
Unified Virtualisation Server

V3

**Virtual Storage
Virtual Machines
Virtual Networking
Over A
Single Server**

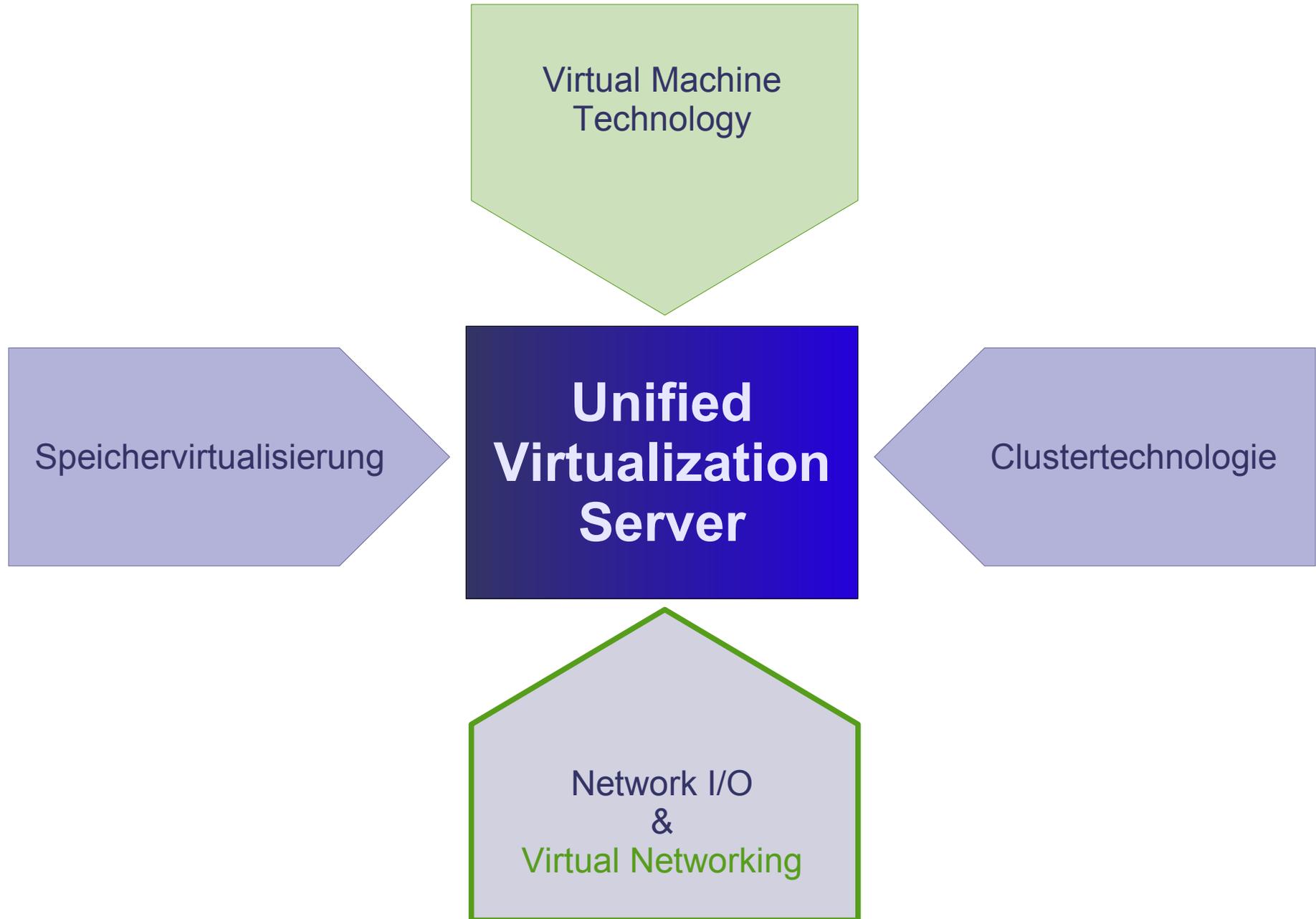
Referenzarchitektur üblicher VM-Cluster

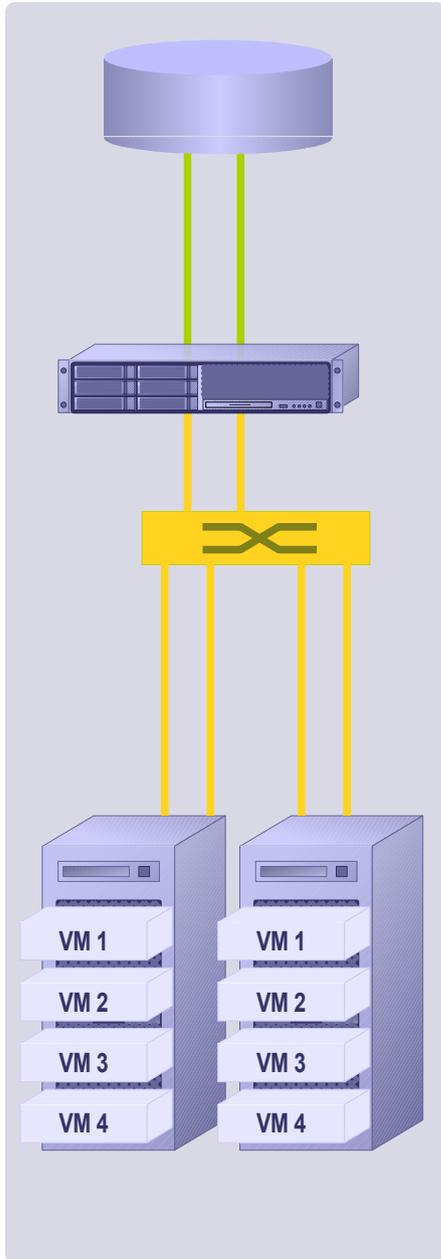
Die ganze Wahrheit



Da muß doch was zu machen sein!

Mit Bündelung der besten Technologien eine neue Qualität erreichen



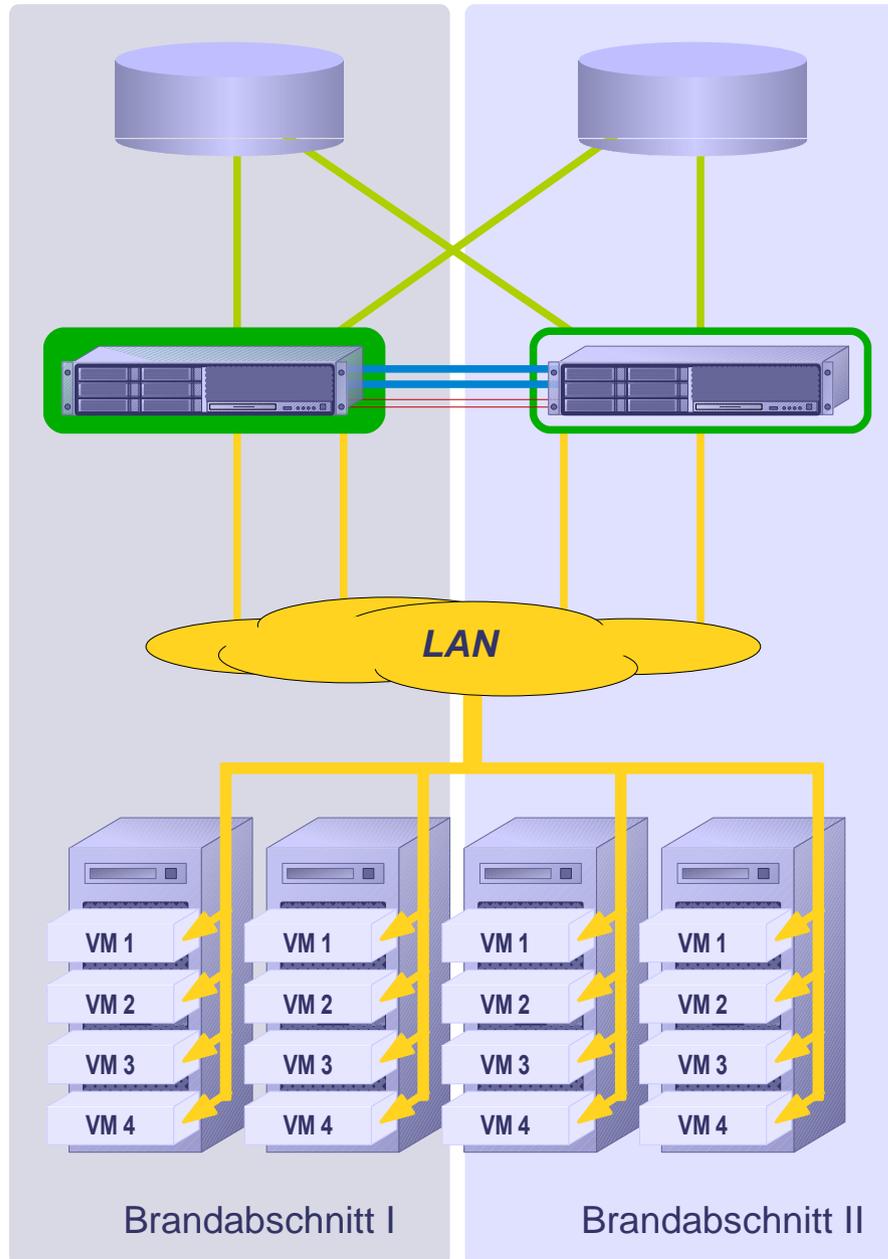


Merkmale

- Collapsed SAN – FC-Fabric verzichtbar
- Typische Serveranbindung via Ethernet (1G/10G/40G) oder ggf. IB
- I/O und Networking über die gleiche physikalische Infrastruktur ⇒ prinzipiell nur ein Kabel je Server
- keine Intelligenz, keine Administration im Netzwerk
- keine Administration auf dem Speichersystem
- Steuerung aller Funktionen vom UVS-Server aus (Speichervirtualisierung, Netzwerk, VMs)

UVS – Das Prinzip

Virtual Storage – Virtual Machines – Virtual Networking – alles integriert

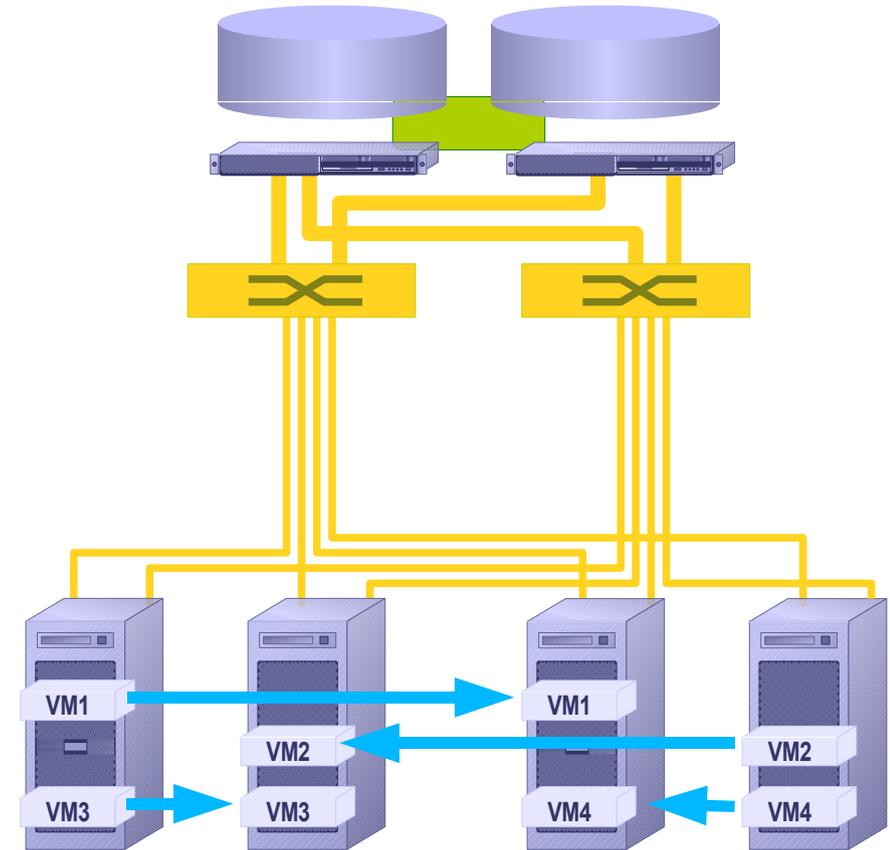
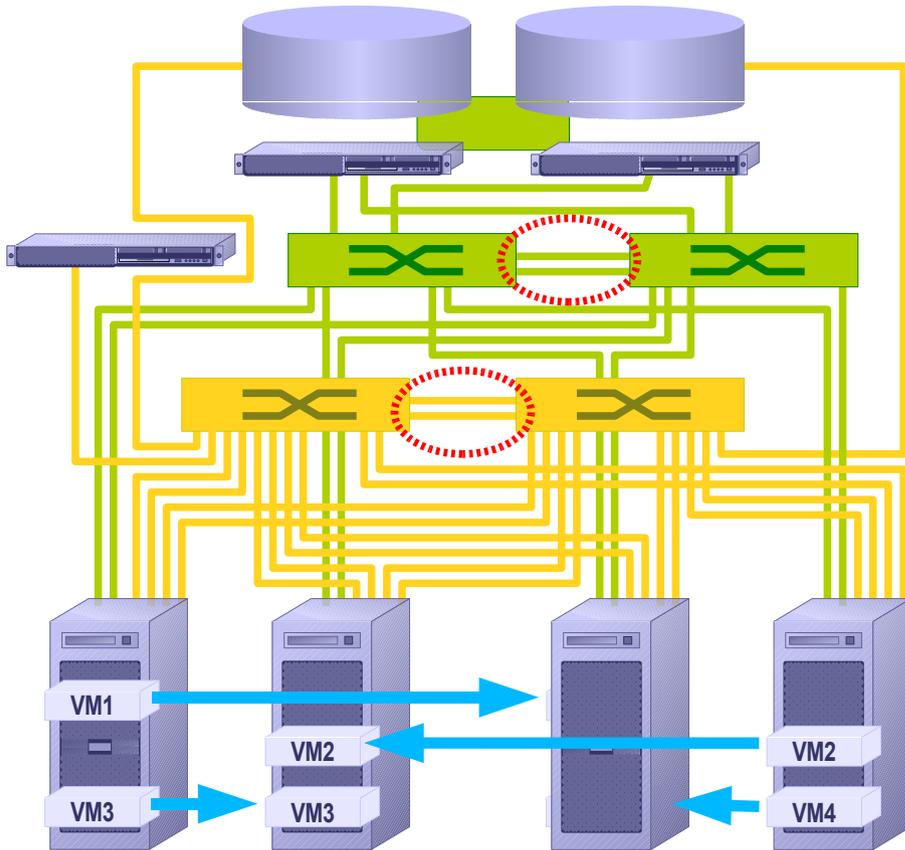


Merkmale

- Collapsed SAN – FC-Fabric verzichtbar
- Typische Serveranbindung via Ethernet (1G/10G/40G) oder ggf. IB
- I/O und Networking über die gleiche physikalische Infrastruktur ⇒ prinzipiell nur ein Kabel je Server
- keine Intelligenz, keine Administration im Netzwerk
- keine Administration auf dem Storage
- Steuerung aller Funktionen vom UVS-Server aus (Speichervirtualisierung, Netzwerk, VMs)
- Natürlich in redundanter Auslegung möglich, dann:
 - Redundanz auf allen Ebenen
 - automatische Spiegelung über zwei RAID-Systeme
 - Austausch **aller** Komponenten im laufenden Betrieb
 - weiterhin zentrale Administration auf dem UVS-Server
- Volle Steuerung der VM-Funktionen (Anlegen, Start, Stop, Provisionierung, HA, Ressourcen ...)
- Volle Integration Netzwerkfunktionen einschl. HA
- Netzwerke der VMs auf Wunsch getrennt
- Backup to Disk / restorefreier Wiederanlauf integriert
- DR-Funktionalität an Bord

Um fair zu sein

Wie sieht die praxistaugliche Verkabelung im Vergleich aus?



Und so integriert man UVS in das Datacentre

Eine skalierbare Serviceplattform mit definierter Übergabeschnittstelle



Nach außen sind nur die gewünschten Services sichtbar:

- kein Storage
- kein Netzwerk
- keine physischen Server

Intern ist alles redundant und skalierbar:
Die private Cloud-Lösung!

