



OSL Storage Cluster 4.0

Neue Perspektiven

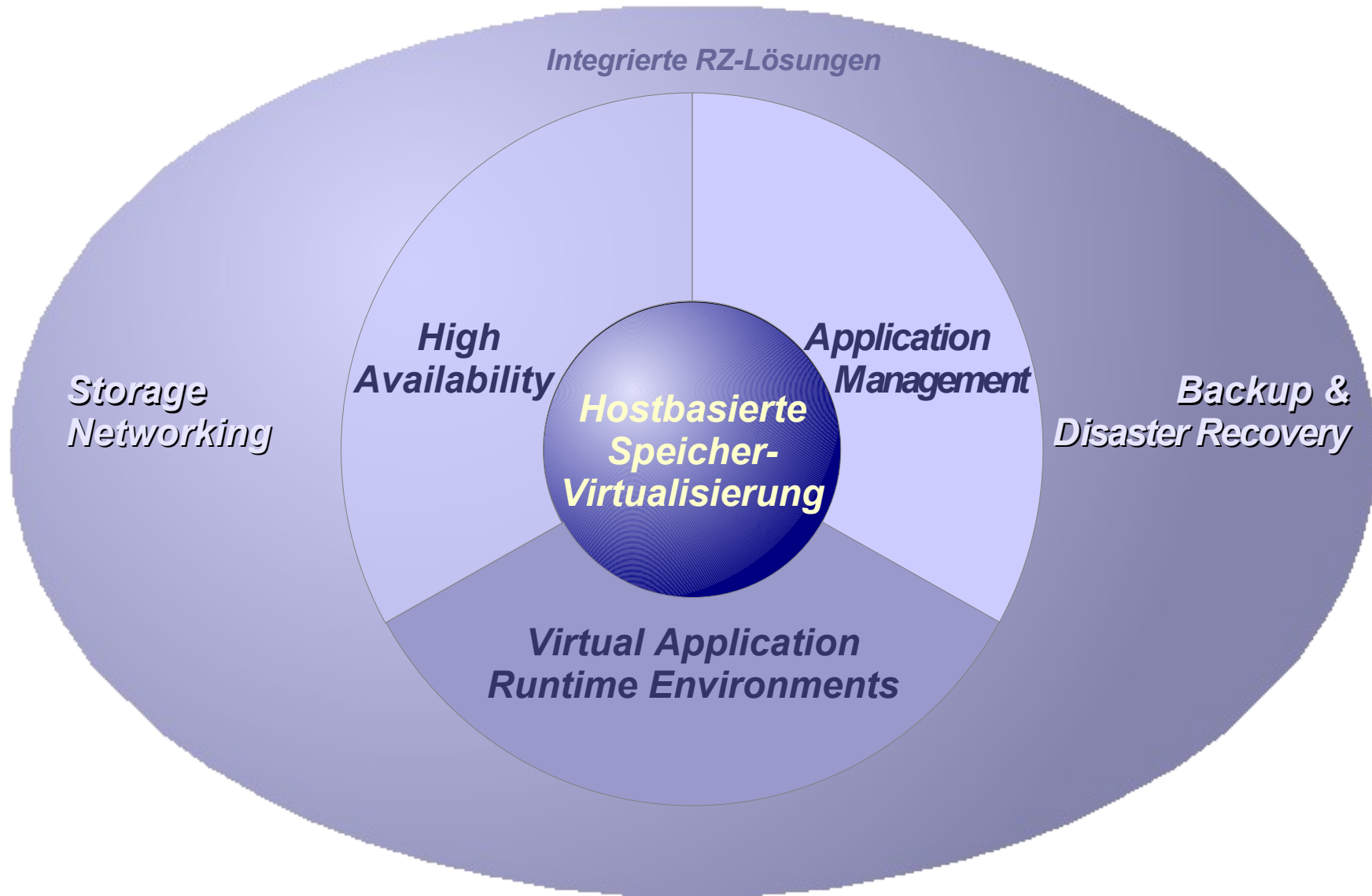
OSL Technologietage 2010
Innovation, Zweckorientierung und Kostendruck
im Rechenzentrum

Bert Miemietz

OSL Gesellschaft für
offene Systemlösungen mbH

Was bietet OSL eigentlich an?

Wir entwickeln Software für Open Systems



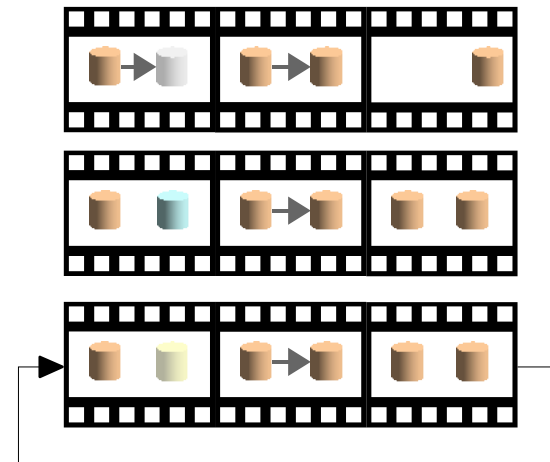
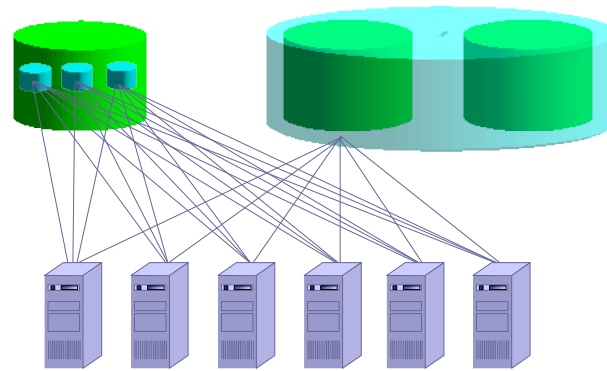
OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Wo stehen wir heute?

Hostbasierte Speichervirtualisierung

Spitzentechnologie von OSL – Funktionsübersicht

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

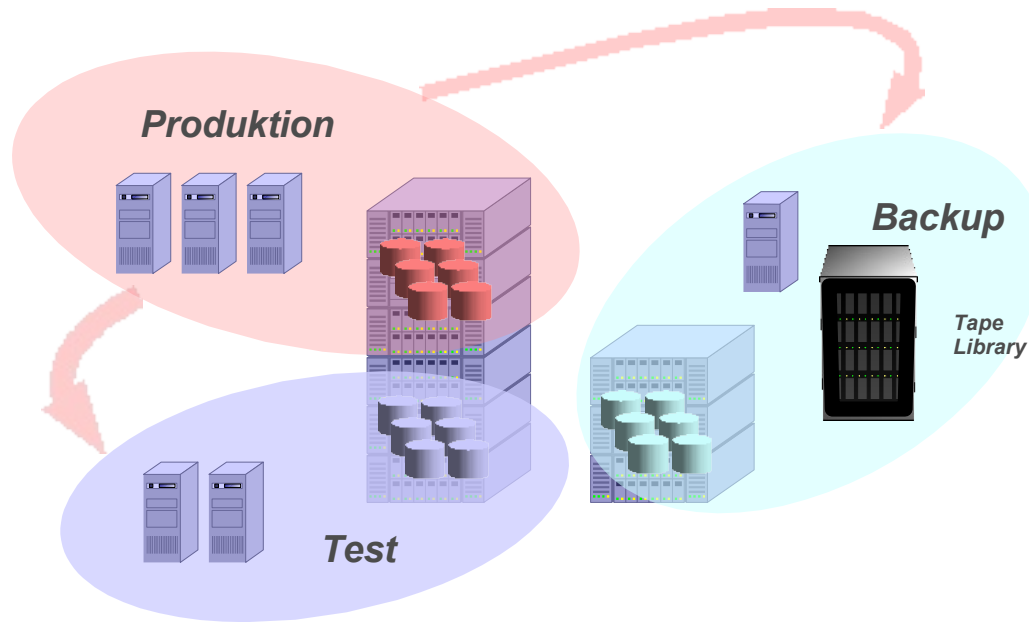
Online-Datenkopien auf wahlfreie Ziele
atomare Operationen für mehrere Volumes

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

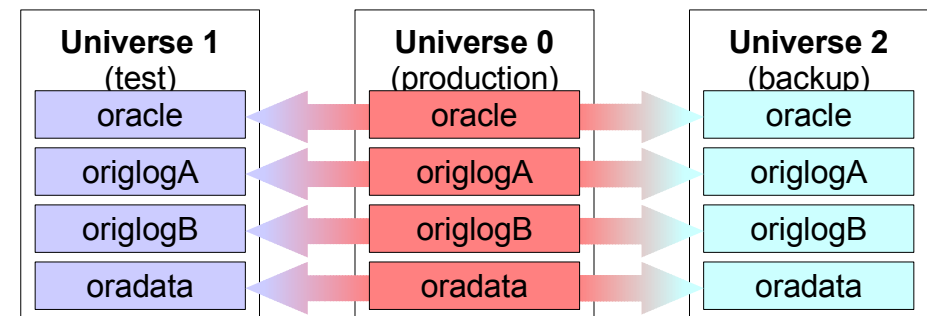
Hostbasierte Speichervirtualisierung

Ein entscheidendes Differenzierungsmerkmal - OSL Universen



- OSL Storage Universen bilden die Aufteilung der Ressourcen nach der Art der Nutzung ab
- Kopien eines Originals können jederzeit erstellt werden, auch unter Beibehaltung des Namens
- OSL Storage Cluster besitzt Informationen über die logischen Beziehungen zwischen den Universen

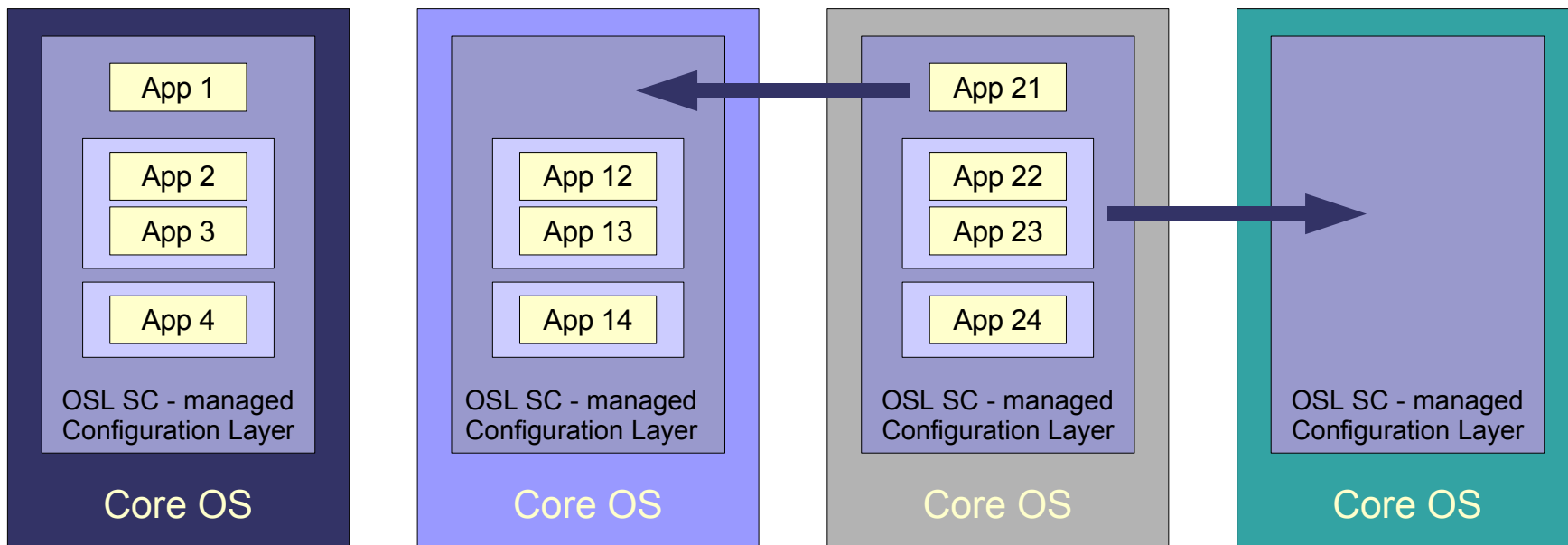
- Volle Integration der Universen in das Betriebssystem
- leichte Identifikation anhand des Namens
- Zugriff auf jede Instanz
 - jederzeit
 - von jedem Host aus



Virtualisierte Ablaufumgebungen für Applikationen

Mehr als nur Speichervirtualisierung und HA

- *Global Devices und Application Resource Description*
 - Raw- und Blockdevices + Dateisysteme
 - ZFS
 - IP-Adressen und NFS
- *Globales Nutzer- und Gruppenmanagement*
- *Automatische Adaption ASCII-Konfigurationsdateien*
- *Globales Management und Migrationsdienste für Zonen*

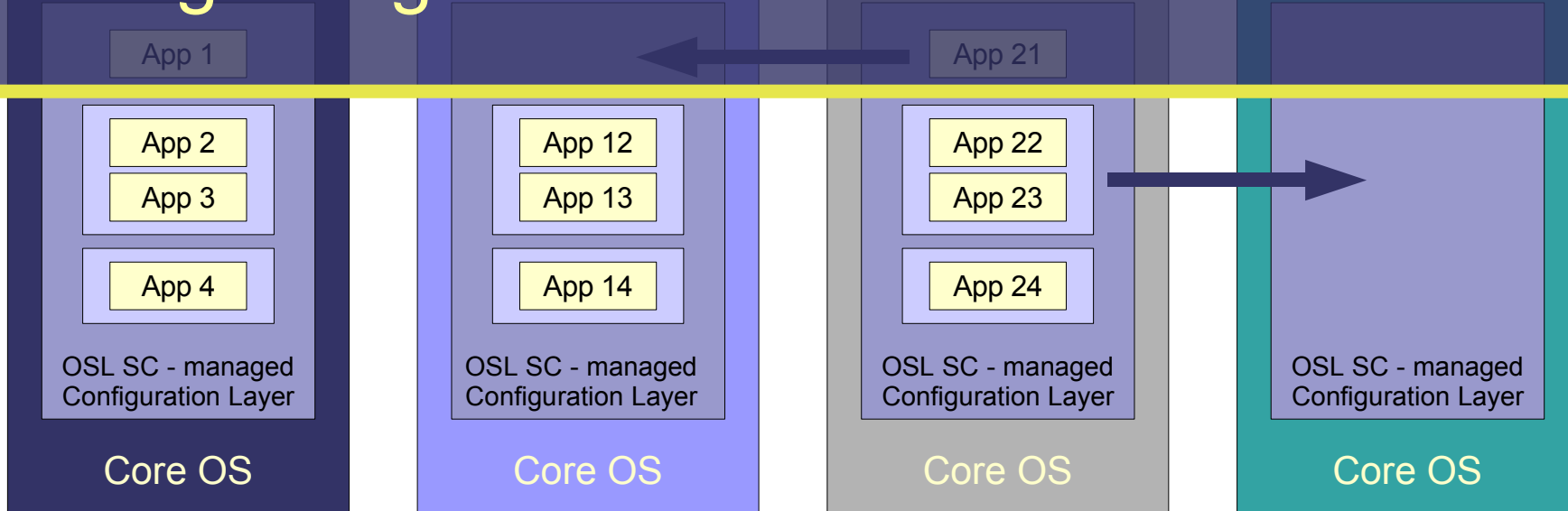


Virtualisierte Ablaufumgebungen für Applikationen

Mehr als nur Speichervirtualisierung und HA

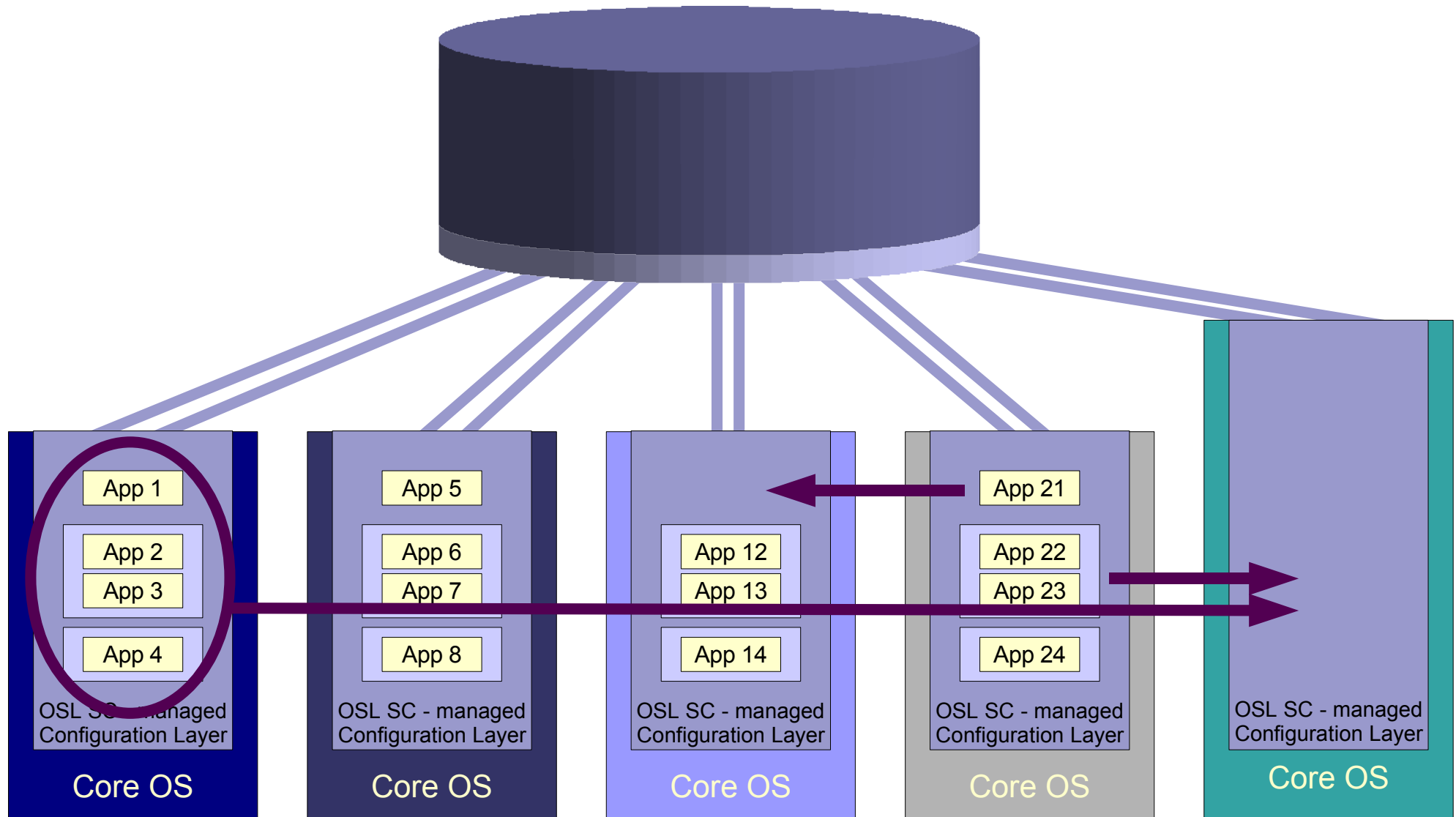
- *Global Devices und Application Resource Description*
 - Raw- und Blockdevices + Dateisysteme
 - ZFS
 - IP-Adressen und NFS
- *Globales Nutzer- und Gruppenmanagement*
- *Automatische Adaption ASCII-Konfigurationsdateien*
- *Globales Management und Migrationsdienste für Zonen*

Umgebung lebt außerhalb der Maschine weiter



Der nächste Schritt: Application Control und HA

Ein adaptiver Cluster im dynamischen Rechenzentrum

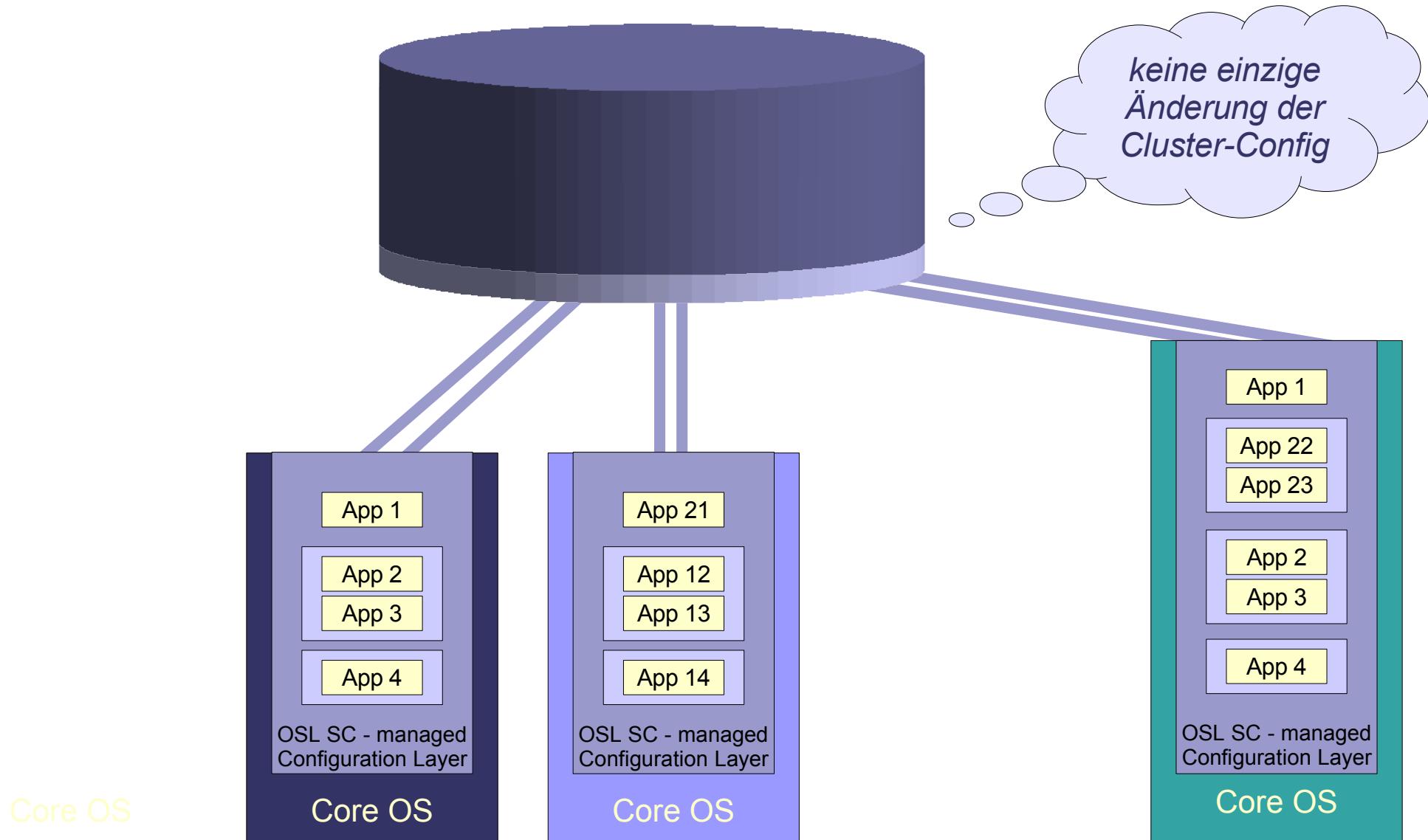


OSL Gesellschaft für offene Systemlösungen mbH

www.osl.eu

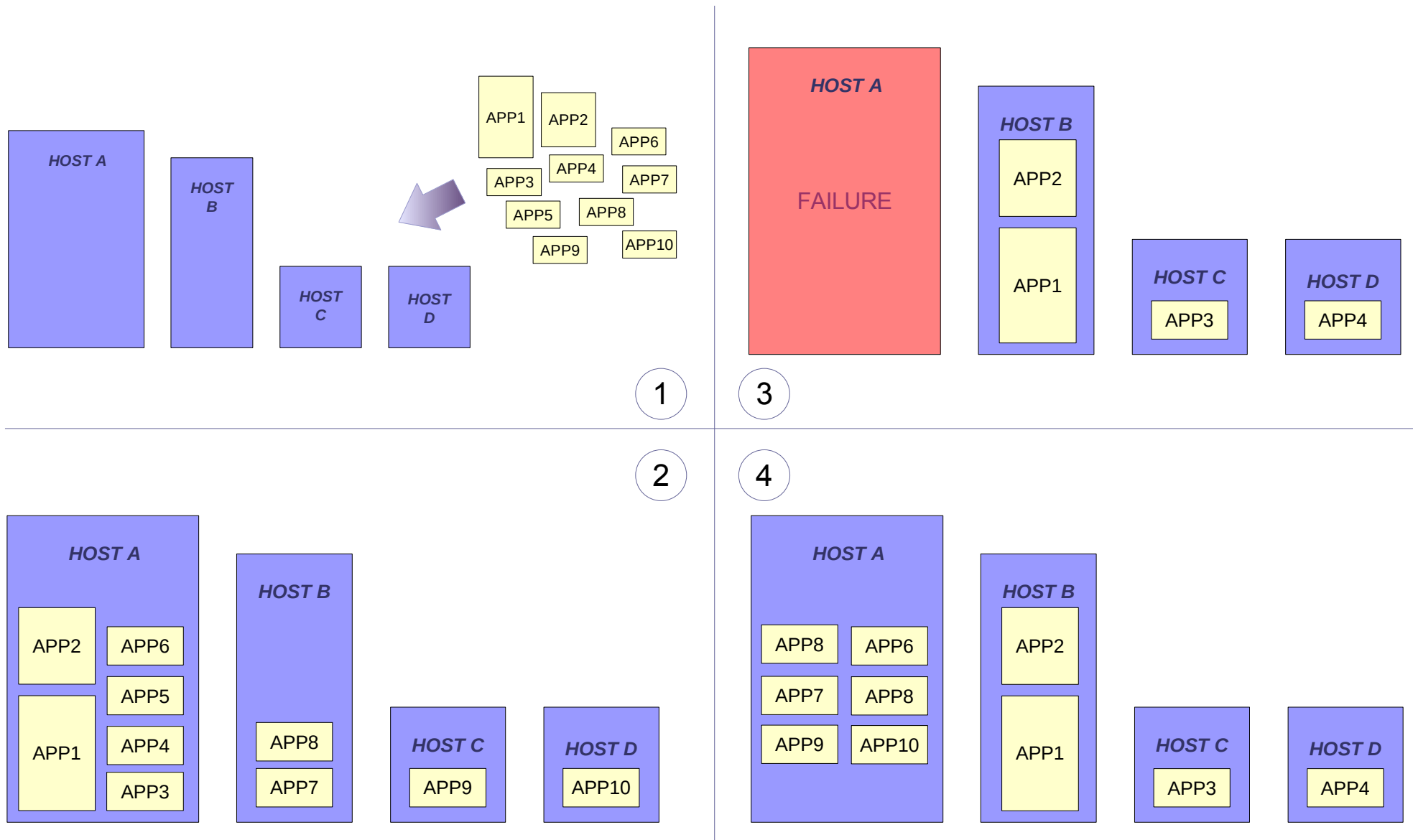
Der nächste Schritt: Application Control und HA

Ein adaptiver Cluster im dynamischen Rechenzentrum



OSL ACO (Application Control) ist mehr als HA

Ressourcenbasiertes Selbstmanagement



Alles zusammen ...

Die aktuelle Produktstruktur

*Application
Awareness*

*Bandbreiten-
steuerung*

*User-
Management*

Application Control Option

clusterweite Steuerung von Applikationen

virtualisierte (hardwareabstrakte) Ablaufumgebungen

Hochverfügbarkeit

ressourcenbasiertes Selbstmanagement

Application Resource Description

*Applikationen
spiegeln*

*Applikationen
klonen*

*B2D / DASI / DR
-Werkzeuge*

clusterfähige Speichervirtualisierung

globale Storage-Pools (rechnerübergreifend)

globales Disk-Inventory

globale Geräte / globaler Namensraum

Cluster-Volumenmanager mit automatisierter Allokation

Disk Access Management

I/O-Multipathing

Extended Data Management

Integration von RAID-basierten Datenkopien / Snapshots

hostbasierte Spiegelung

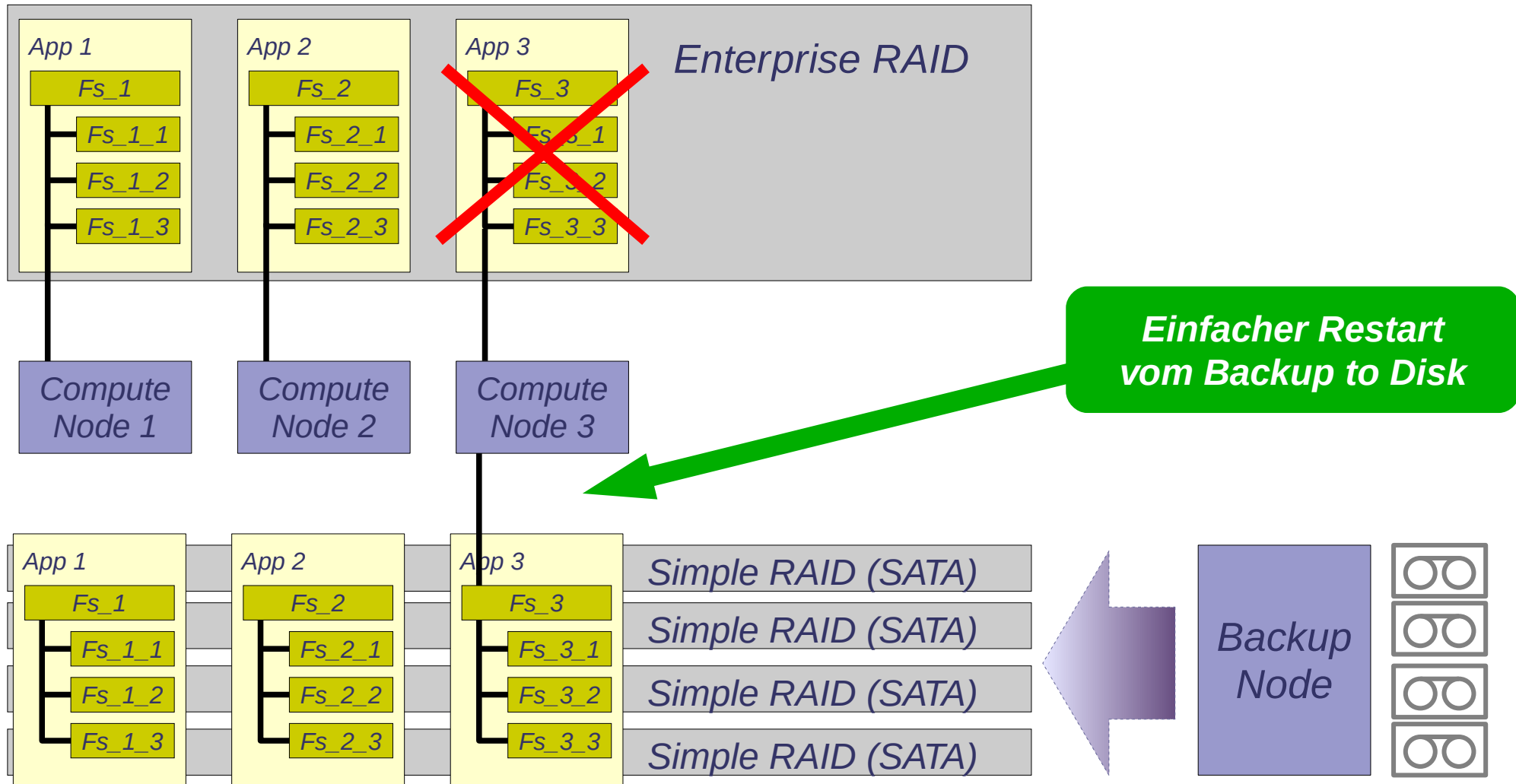
online Datenmigration

Daten klonen

Was hat sich in den letzten zwei Jahren getan?

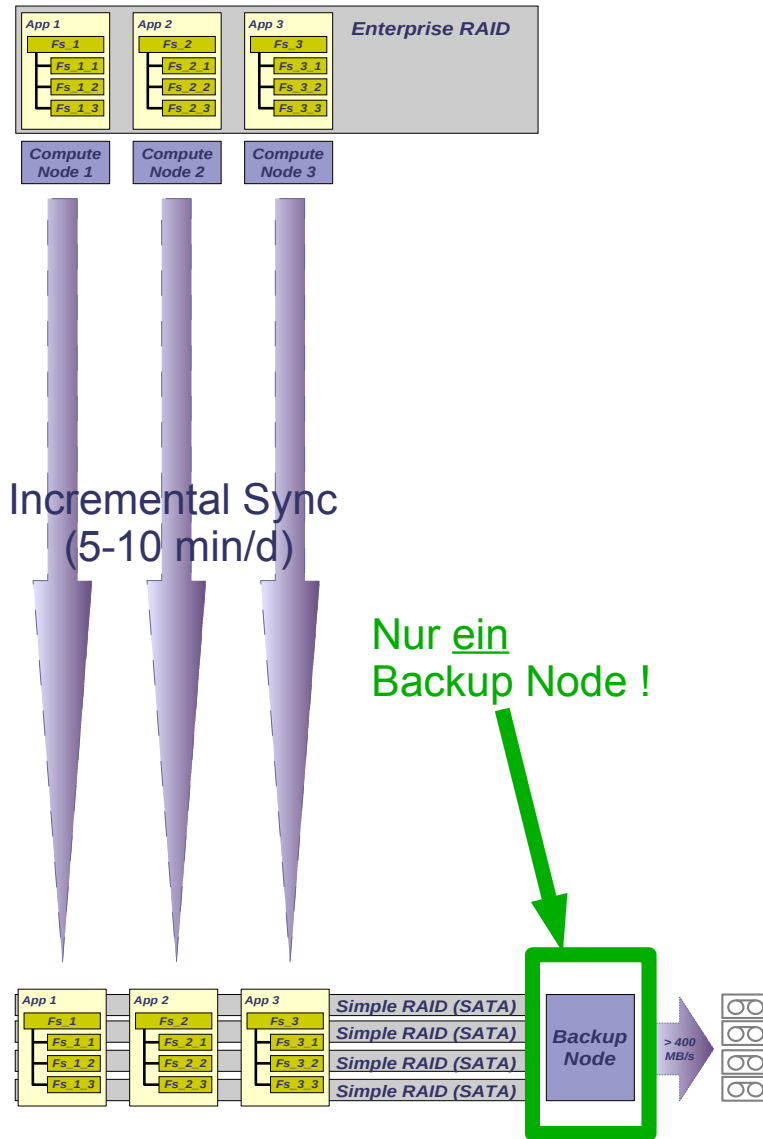
und Instant Recovery

Restore vom Band wird zumeist nicht benötigt



Tiefer in's Detail: Networker-Integrationslösung

Was andere Lösungen so nicht bieten

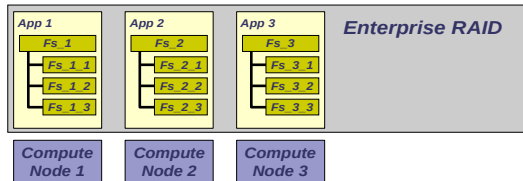


Was bietet die Integrationslösung von OSL?

- kombiniert B2D mit Bandsicherung
- applikationsorientiertes Verfahren
- sofortiger Neustart von Backup-Disk (kein Tape-Restore)
- Steuerung der Sicherungen über Networker-GUI oder CLI
- integriertes Pre- und Postprocessing
- mehrere Sicherungen pro Tag möglich
- differenzierte Erfolgskontrolle über die Networker-Indizes
- Aufzeichnung von Dateisystem-Informationen für Restore
- eindeutige Identifikation kompletter Sicherungen
- einfacher Restore kompletter Sicherungen
- weitere Funktionen für Oracle
 - Tool für Logrestore und Roll Forward
 - Archivierungslösung
- LAN-free Backup
- „cluster aware“, d.h. kein Eingriff bei Umschaltungen nötig
- automatische Gleichverteilung der Plattenlast
- hohe Durchsätze – VTL unter diesem Aspekt entbehrlich

Tiefer in's Detail: Networker-Integrationslösung

Was andere Lösungen so nicht bieten



Was bietet die Integrationslösung von OSL?

- kombiniert B2D mit Bandsicherung
- applikationsorientiertes Verfahren

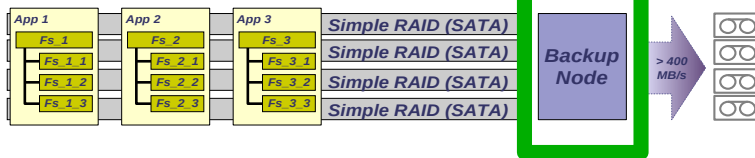
analoge Funktion für:

**Oracle, Informix, SAP
generische Applikationen**

Networker, Hypertape, NetVault

Incremental Sync
(5-10 min/d)

Nur ein
Backup Node !



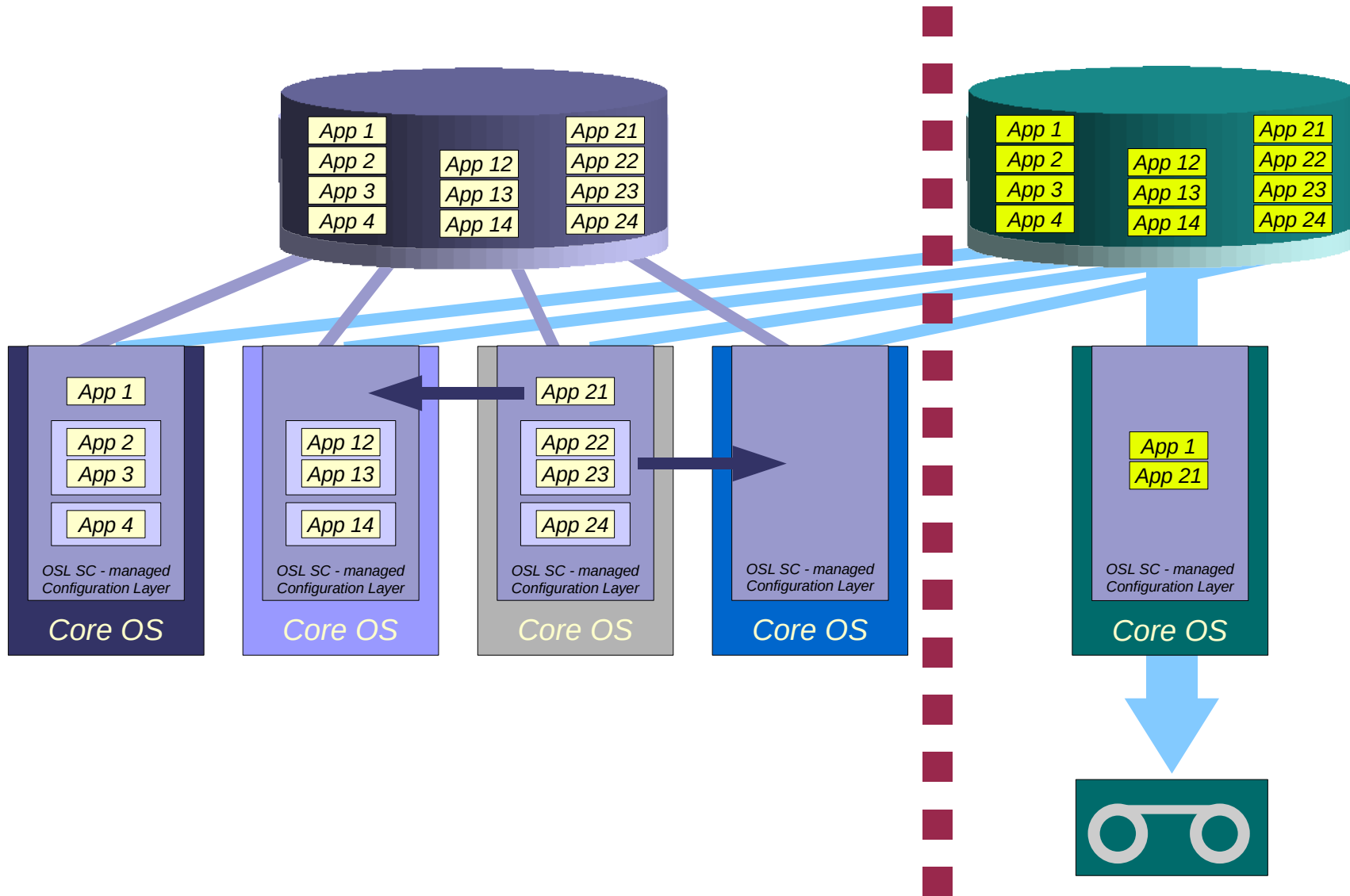
- sofortiger Neustart von Backup-Disk (kein Tape-Restore)
- Steuerung der Sicherungen über Networker-GUI oder CL
- integriertes Pre- und Postprocessing
- mehrere Sicherungen pro Tag möglich
- differenzierte Erfolgskontrollen über die Networker-Indizes
- Aufzeichnung von Dateisystem-Informationen für Restore
- eindeutige Identifikation kompletter Sicherungen
- einfacher Restore kompletter Sicherungen
- weitere Funktionen für Oracle
 - Tool für Logrestore und Roll Forward
 - Archivierungslösung
- LAN-free Backup
- „cluster aware“, d.h. kein Eingriff bei Umschaltungen nötig
- automatische Gleichverteilung der Plattenlast
- hohe Durchsätze – VTL unter diesem Aspekt entbehrlich

OSL Gesellschaft für offene Systemlösungen mbH

www.osl.eu

Alles zusammen

Dynamische, flexible und hochverfügbare Infrastrukturen

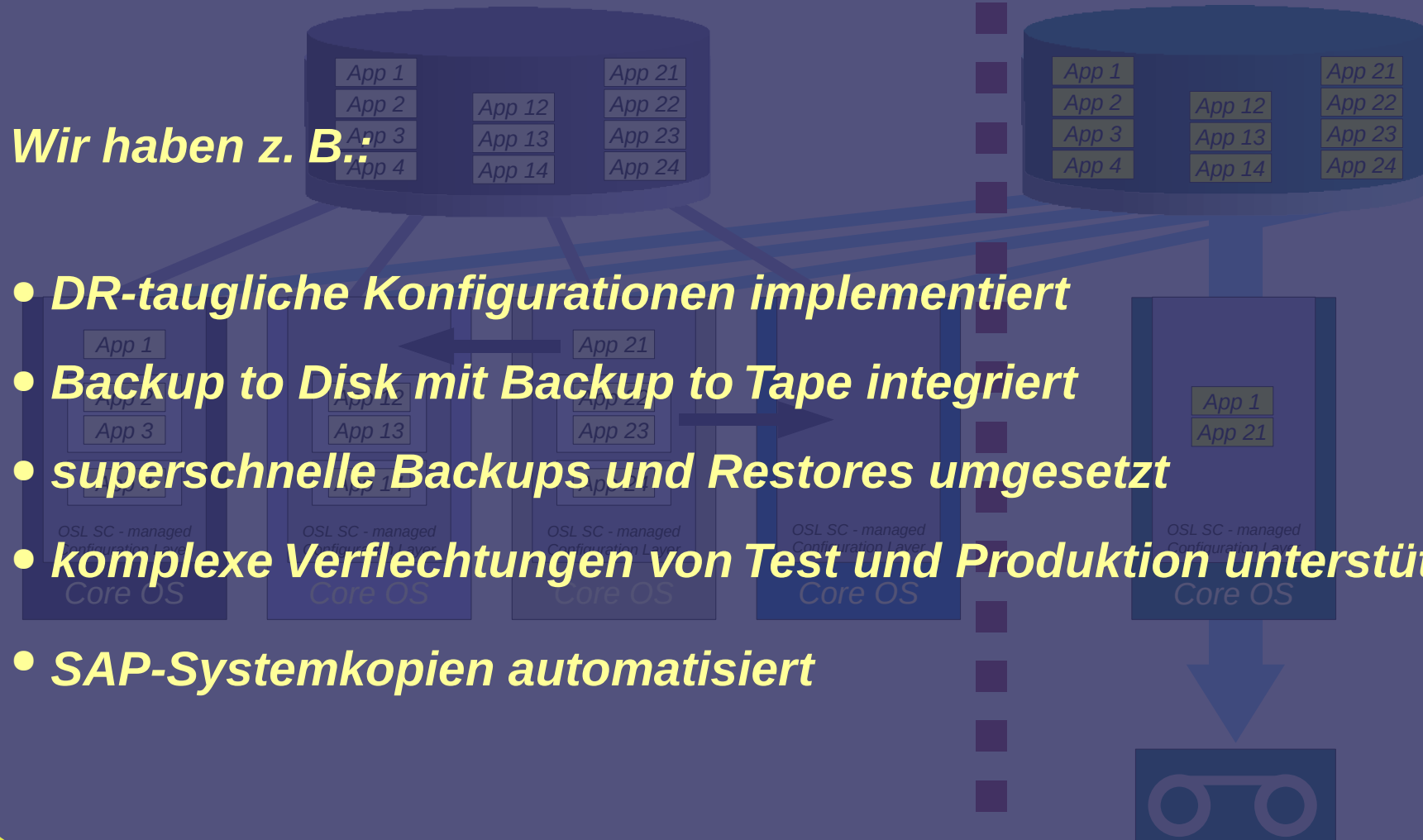


Alles zusammen

Dynamische, flexible und hochverfügbare Infrastrukturen

Wir haben z. B.:

- **DR-taugliche Konfigurationen implementiert**
- **Backup to Disk mit Backup to Tape integriert**
- **superschnelle Backups und Restores umgesetzt**
- **komplexe Verflechtungen von Test und Produktion unterstützt**
- **SAP-Systemkopien automatisiert**



Welten verbinden:
OSL Storage Cluster 4.0

OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

OSL SC 4.0 – Integration von OS -Virtualisierung

Nahtlose Integration virtueller Maschinen in clusterweites Management



- **Einführung von Virtual Nodes**
 - können per Software “erzeugt” werden
 - Modifikationen in Node-Administration erforderlich
 - Administration wie auf einem Physical Node
 - HA-Konfigurationen auf einem Physical Node möglich
 - voller Zugriff auf Speichervirtualisierung etc.
 - Integration von Ressourcenmanagement
- **Maximale Anzahl von Nodes mdst. 64**
 - VM's steigern die Zahl von Knoten deutlich
 - durchschnittliches Unternehmen soll alles in einem einzigen Cluster fahren können
- **Zonen sind besondere Herausforderung**
 - keine eigene OS-Instanz -> keine eigenen Treiber
 - neuer Stand soll aber vollen Zugriff auf SC-Dienste ermöglichen, z. B. applikationsbezogenes, automatisches Backup
 - keine besonderen Programme mehr für Applikationen in Zonen
- **Verbessertes Applikationsmanagement**
 - virtualisierte Applikationen sollen sich nahtlos zwischen Physical Nodes und Virtual Nodes verschieben lassen

RSIO – Remote Storage I/O

Warum müssen es eigentlich immer spezialisierte Netzwerke sein?

Fibre Channel

- Spezialprotokoll
- Block I/O
- Kanaleigenschaften
- Niedrige Latenz
- Hoher Durchsatz
- Niedrige CPU-
Belastung

Ethernet / IP

- Universalprotokoll
 - Primär von Applikationen getrieben
 - Zweck: Kommunikation
 - Client/Server-Applikationen
 - Implementierung wesentlich im OS
-> höhere CPU-Belastung
 - Seit langem auch für Storage
genutzt (NFS, Backup ...)
- NFS, SMB ...
 - Backup
 - IP over FC
 - FC over IP

RSIO – Remote Storage I/O

Warum müssen es eigentlich immer spezialisierte Netzwerke sein?

Fibre Channel

- Spezialprotokoll

Aus Sicht des Anwenders sind sich beide

- Kanaleigenschaften

- Niedrige Latenz

- Hoher Durchsatz

- Niedrige CPU-
Belastung

Ethernet / IP

- Universalprotokoll

- Zweck: Kommunikation

- Client/Server-Applikationen

- Implementierung wesentlich im OS
→ höherer CPU-Lastung

- Seit langem auch für Storage
genutzt (NFS, Backup ...)

**Netzwerktechnologien heute im Verhalten
oftmals ähnlicher als erwartet**

Warum Storage über Ethernet?

Anforderungen und Möglichkeiten

- **Anforderungen und Erwartungen**

- Erfordernisse der Anwendungen und Protokolle (Kommunikation, Filesharing etc.)
- Preisliche Motivationen
- Einheitliche Infrastruktur, weniger Ports ?
- Einfachheit, Flexibilität ?
- Virtualisierungstechnologien, Verfügbarkeit von Treibern
- Zusatzfunktionen (Konvertierungen, Filesystemsnapshots ...)

- **Möglichkeiten**

- Gigabit-LAN heute vergleichsweise preiswert
- Gigabit-LAN heute in vernünftiger Relation zur Geschwindigkeit einer Festplatte bzw. eines RAID-Systems
- Gigabit-LAN heute in günstiger Relation zu Durchsatz-Anforderungen der Applikationen
- Mehrere Gigabit-Ports je Server
- Ethernet ist eigentlich (fast) kein Ethernet mehr -> Switching-Technologie
- RAID-Systeme / Filer sprechen direkt die erforderlichen Protokolle
- Neue Performance-Erwartungen an 10Gbit-Ethernet

Warum Storage über Ethernet?

Anforderungen und Möglichkeiten

- **Anforderungen und Erwartungen**

- Erfordernisse der Anwendungen und Protokolle (Kommunikation, Filesharing etc.)
- Preisliche Motivationen
- Einheitliche Infrastruktur, weniger Ports ?
- Einfachheit, Flexibilität ?

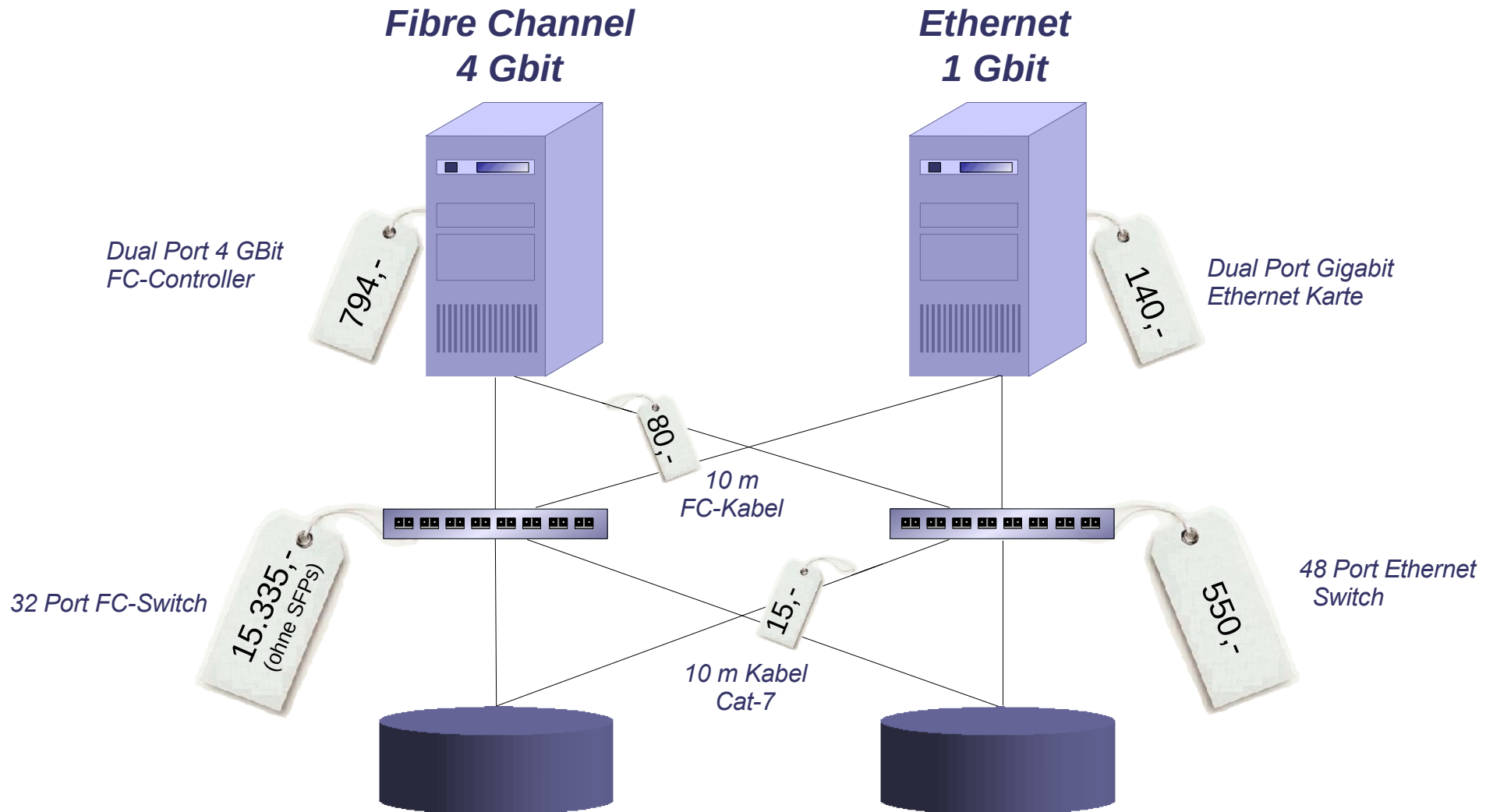
**Vielfalt an Virtualisierungstechnologien, Plattformen ...
erhöht Uniformierungsdruck bei Connectivity**

- **Möglichkeiten**

- Gigabit-LAN heute vergleichsweise preiswert
- Gigabit-LAN heute in vernünftiger Relation zur Geschwindigkeit einer Festplatte bzw. eines RAID-Systems
- Gigabit-LAN heute in günstiger Relation zu Durchsatz-Anforderungen der Applikationen
- Mehrere Gigabit-Ports je Server
- Ethernet ist eigentlich (fast) kein Ethernet mehr -> Switching-Technologie
- RAID-Systeme / Filer sprechen direkt die erforderlichen Protokolle
- Neue Performance-Erwartungen an 10Gbit-Ethernet

Das Dauerthema: Kostenvorteile?

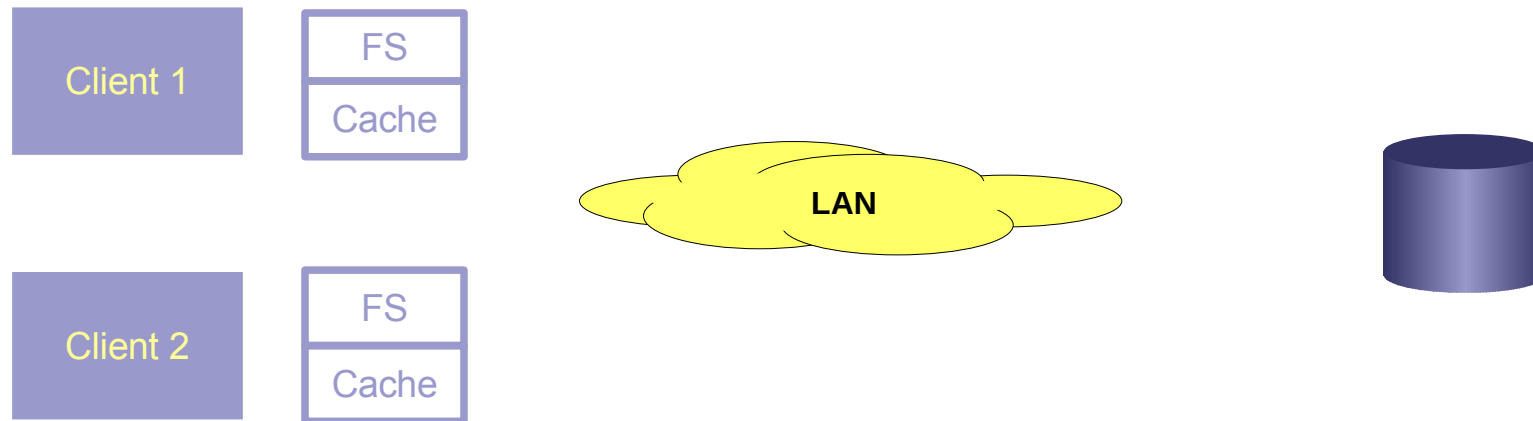
Ein Vergleich mit Fibre Channel (Stand 10/2009)



Starke Argumente für Block-I/O im RZ

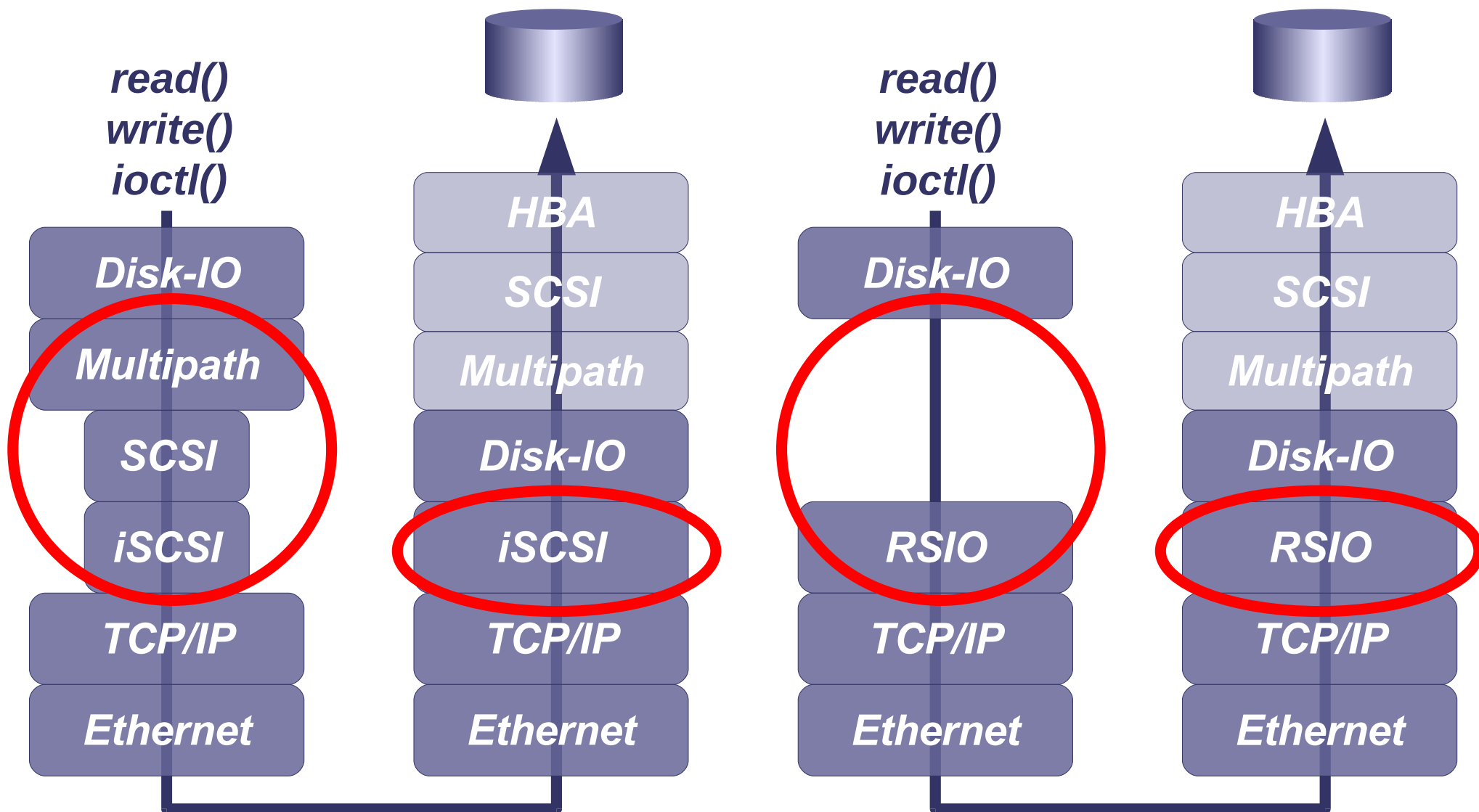
Jenseits von Filesharing (NFS, SMB) überwiegen die Vorteile

- Volle Kontrolle des Client-OS über das Storage-Device
- Nutzbar für beliebige Filesysteme
- Keine Kopplung an Server-OS (Isolation, privates Identity Management)
- Nur Übertragung von I/O, nicht von Cache-Inhalten
- Cache liegt beim Client -> schnellster Zugriff, Client-Caches summieren sich auf
- Einfache Administration, schlankes Protokoll, hohe Geschwindigkeit



Block-I/O über Ethernet/IP – ohne SCSI

Ein anderer Ansatz



RSIO - Remote Storage IO

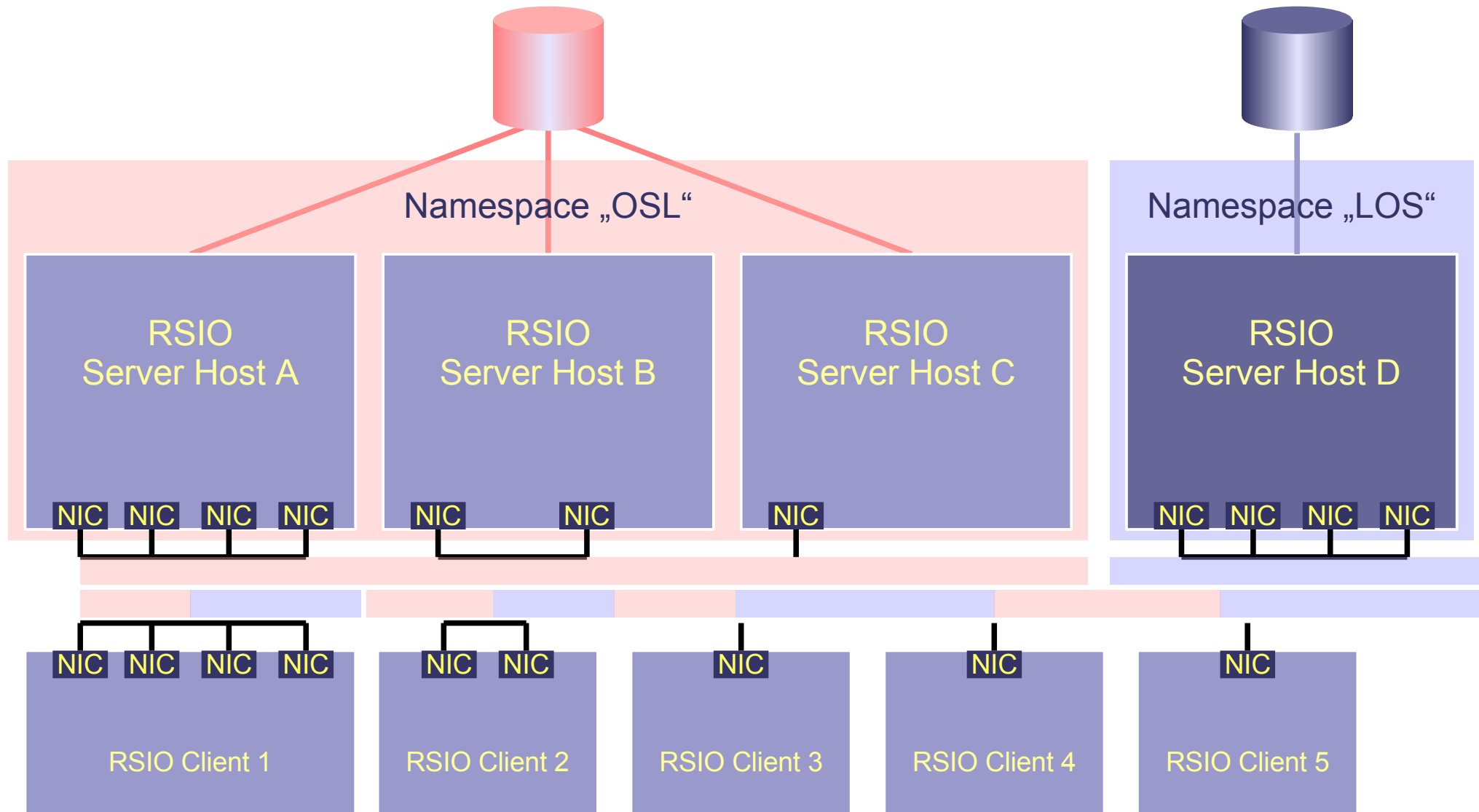
Eckdaten der neuen Technologie für LAN-attached (shared) Block Devices



- *eigenes, von OSL entwickeltes Protokoll*
- *natürliche Erweiterung der Speichervirtualisierung auf LAN (Ethernet)*
- *voller Clustersupport möglich (Client und Server), vorrangig natürlich:*
- *Einbindung in OSL-Clustertechnologie*
- *alle relevanten IO-Aufrufe (read, write, ioctl ...) abbildbar*
- *hochportable Implementierung*
- *internes Layout berücksichtigt moderne CPU- und Serverkonzeptionen*
- *guter Durchsatz / gute Verfügbarkeit mit heutiger preiswerter Standardtechnik*
- *integrierte und bequeme Administration vom Host aus*
- *Erweiterbarkeit, Raum für intelligente IO-Lösungen*

RSIO - Architektur ganz praktisch

Klar gegliedertes und flexibles administratives Konzept



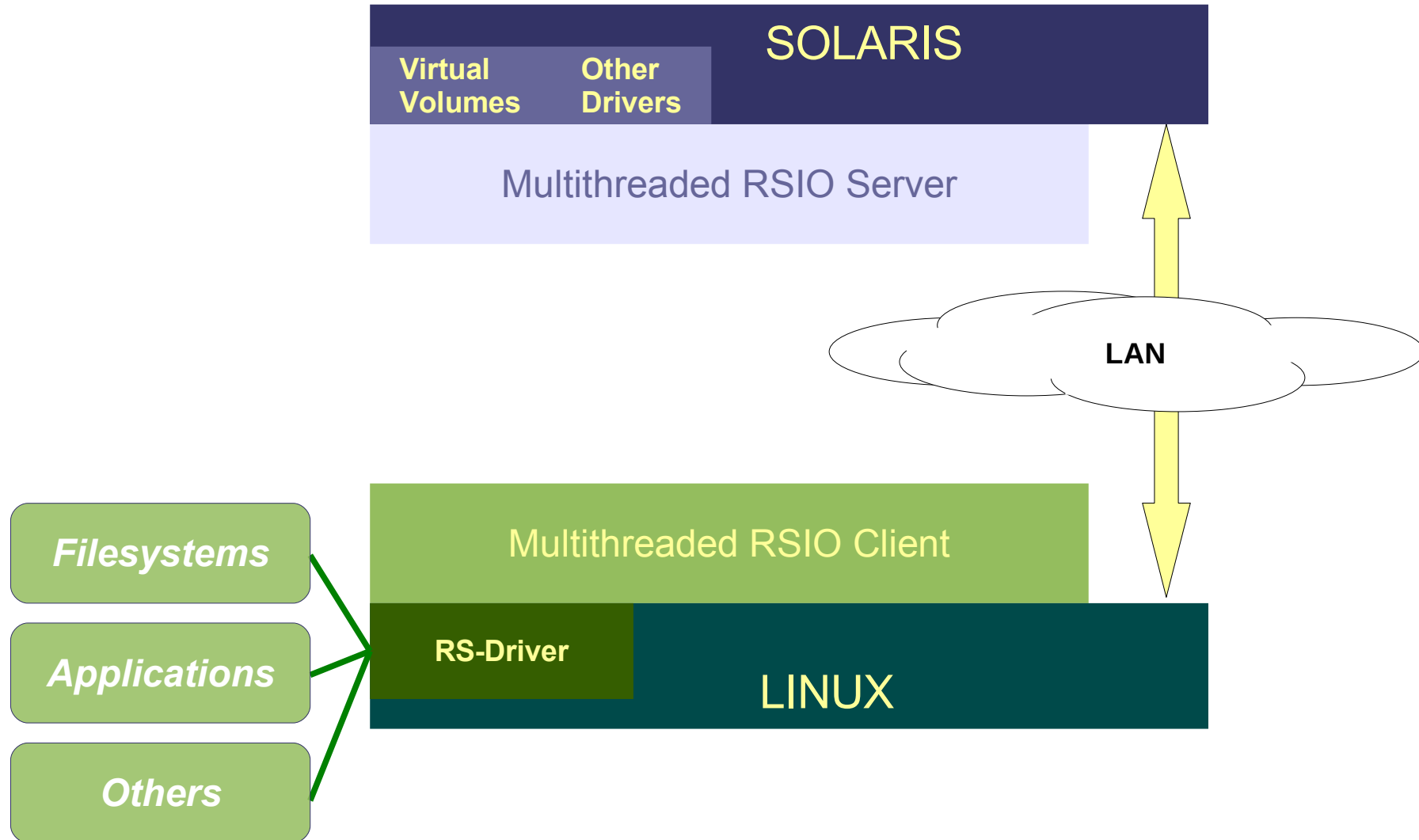
Parameter der RSIO-Architektur

Flexible Client-Server-Implementierung

- *Ein Namespace definiert Server (und Clients) mit Zugriff auf dieselben Storage-Ressourcen*
- *Auf einem Serverhost können (nahezu) beliebig viele Server(prozesse) laufen*
- *Jeder Serverhost kann (nahezu) beliebig viele Clients bedienen*
- *jeder Client unterstützt den Zugriff auf bis zu 256 Server*
- *jede Maschine (Client und Server) unterstützt bis zu 8 Interfaces*
- *jeder Client hat simultan Zugriff auf verschiedene Namespaces*
- *Auto-Explorer*
 - *Ermitteln verfügbarer Verbindungen*
 - *Ermitteln der Schnittstelleneigenschaften*
 - *Test der Parameter auf der Übertragungsstrecke*

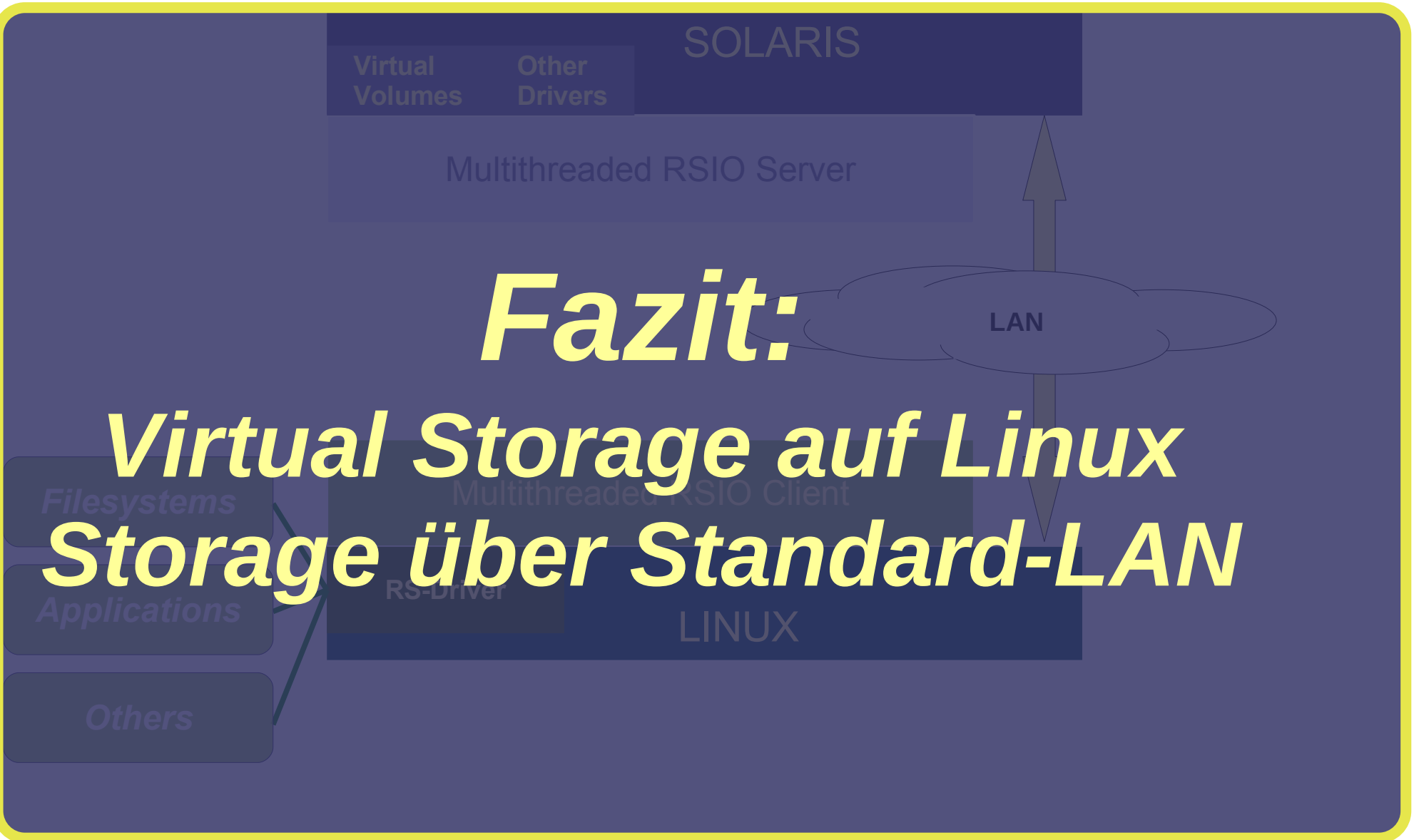
RSIO – unsere Lösung für Linux

Ein Blick hinter die Kulissen...



RSIO – unsere Lösung für Linux

Ein Blick hinter die Kulissen...



Aber ist das nicht zu langsam?

Performance bei Cache-Reads -> der Netzwerkprotokoll-Test

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

Hinweise zur Positionierung

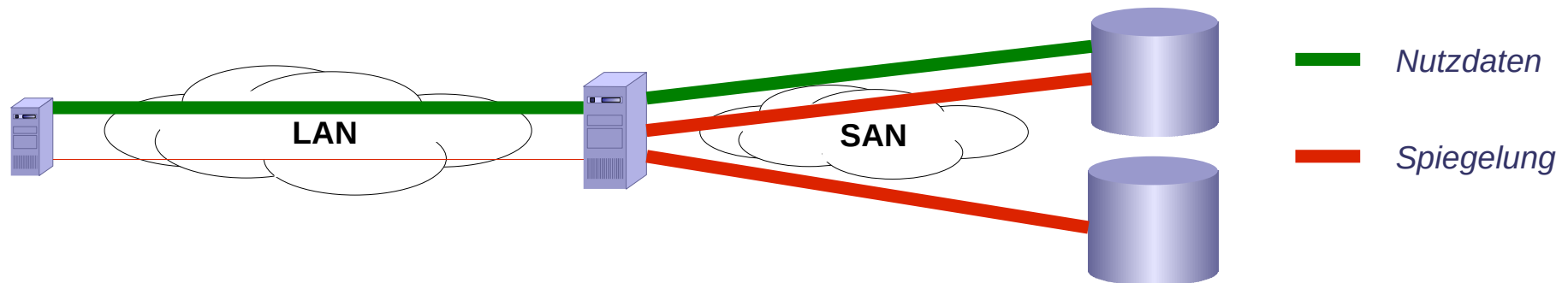
Performanceaspekte

- *exzellente Performance*
 - *keine Spezialsettings für TCP/IP -> Performance "out of the Box"*
- *aber: LAN ist kein SAN*
 - *brauche ich extreme Performance bei niedrigster CPU-Last: FC*
- *RSIO wird oft schneller sein als lokale Platten*
 - *es wird Virtual Storage aus spezialisierten Systemen genutzt*
 - *Ausnutzung von Cache-Algorithmen der Speichersysteme*
- *Standard-Server hat mit nur 2x1Gigabit > 200MByte/s zur Verfügung*
- *Intelligente RSIO-Server -> nicht alle Daten müssen durch's LAN*

RSIO - was heißt intelligenter Server?

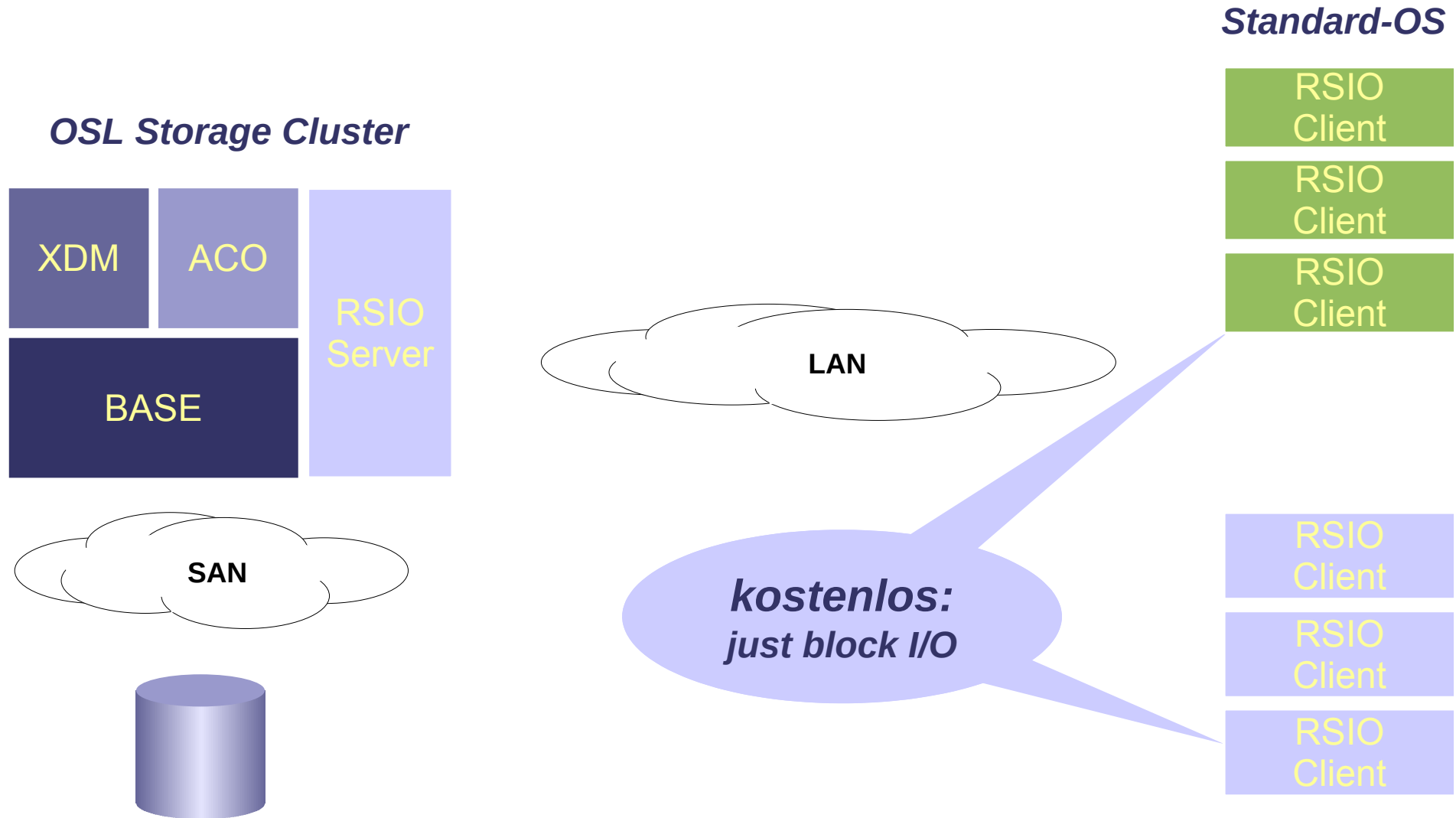
Virtual Storage und Transportvermeidung

- automatische Nutzung des I/O-Caches des OS
 - Daten werden effizienter transportiert
- in Verbindung mit OSL Storage Cluster: Virtual Storage
 - kein “format”, “fdisk” auf dem Client
 - global Namespace
 - Disk-Striping etc.
- Server kann weitere Optimierungen anbieten
 - Optimierung der I/O-Aufträge
 - zusätzliche Cache-Algorithmen
- Datenspiegelung etc. ohne Transport der Daten via LAN
 - volle Steuerung vom Client aus
 - nur Control-Daten laufen über LAN, Nutzdaten bleiben im Backend



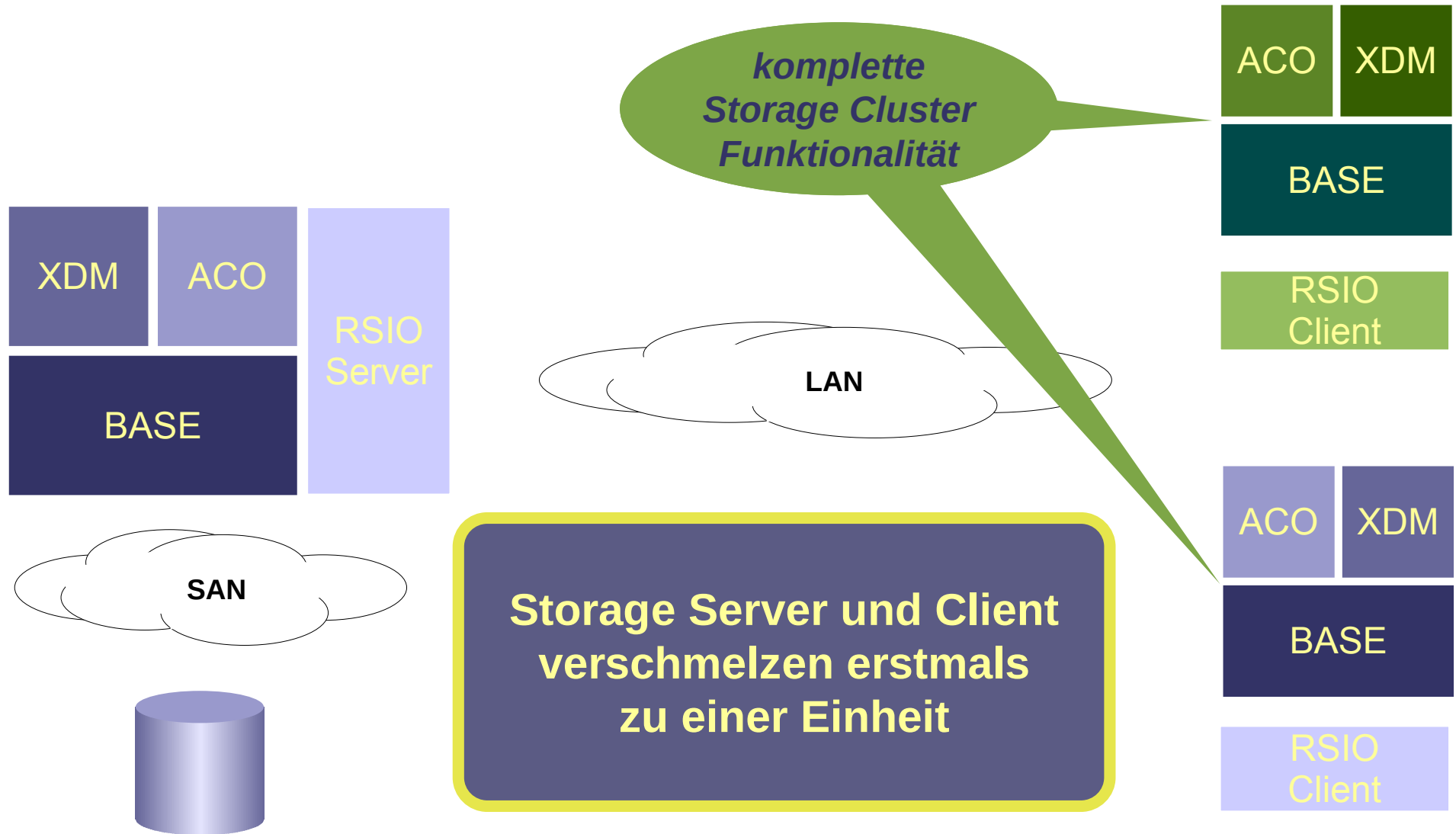
RSIO - was kann ich damit anfangen?

Ein einfaches Szenario in Verbindung mit OSL Storage Cluster



RSIO - was kann ich damit anfangen?

Volle Clusterfunktionalität für Server und Client



RSIO – Entdecke die Möglichkeiten

SAN-LAN-Konvergenz



RSIO-Server

SAN

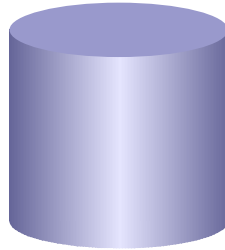
LAN

LAN

Anwender

RSIO – Entdecke die Möglichkeiten

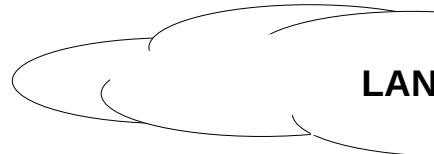
SAN-LAN-Konvergenz



RSIO-Server



SAN



LAN



LAN

Anwender

Komplexität reduzieren

SAN-Ports einsparen

Portauslastung verbessern

Hosts preiswert und einfach anbinden

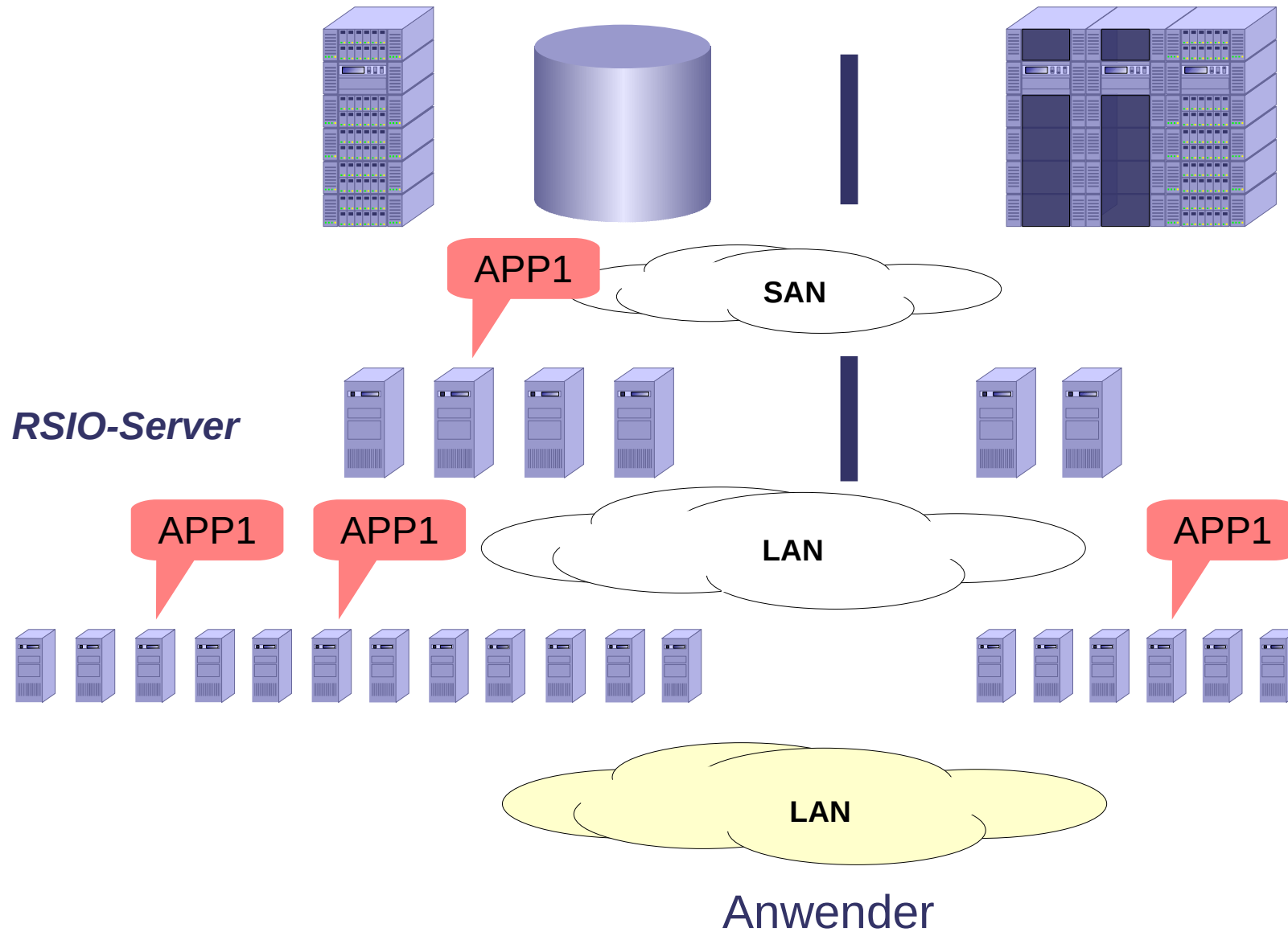
Performance-Rightsizing

erleichterte OS-Virtualisierung

Standardkomponenten

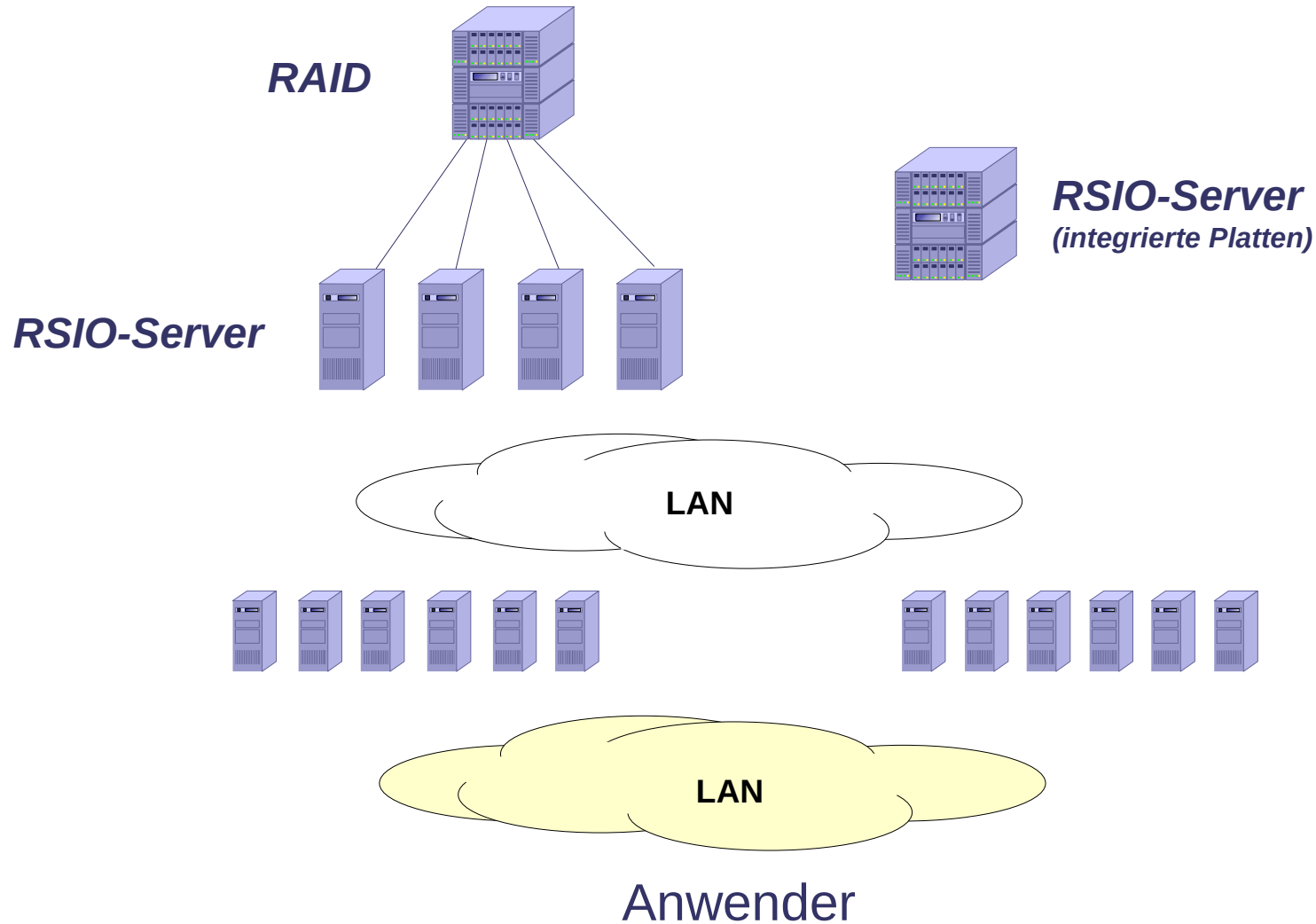
RSIO – Entdecke die Möglichkeiten

Es geht nicht nur um I/O: Run Applications Anywhere



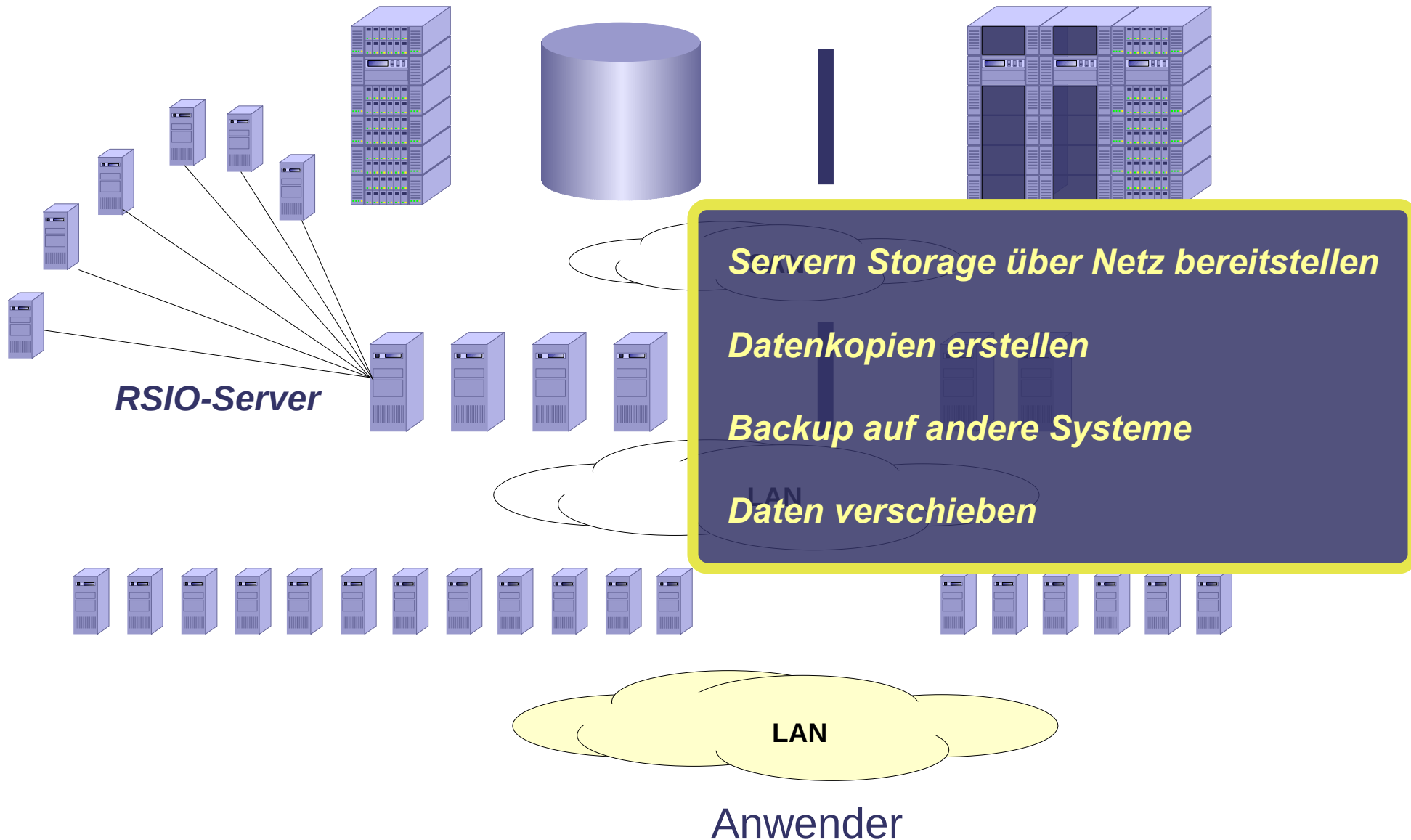
RSIO – Entdecke die Möglichkeiten

Für nicht ganz so große Anwender: Gar kein SAN mehr



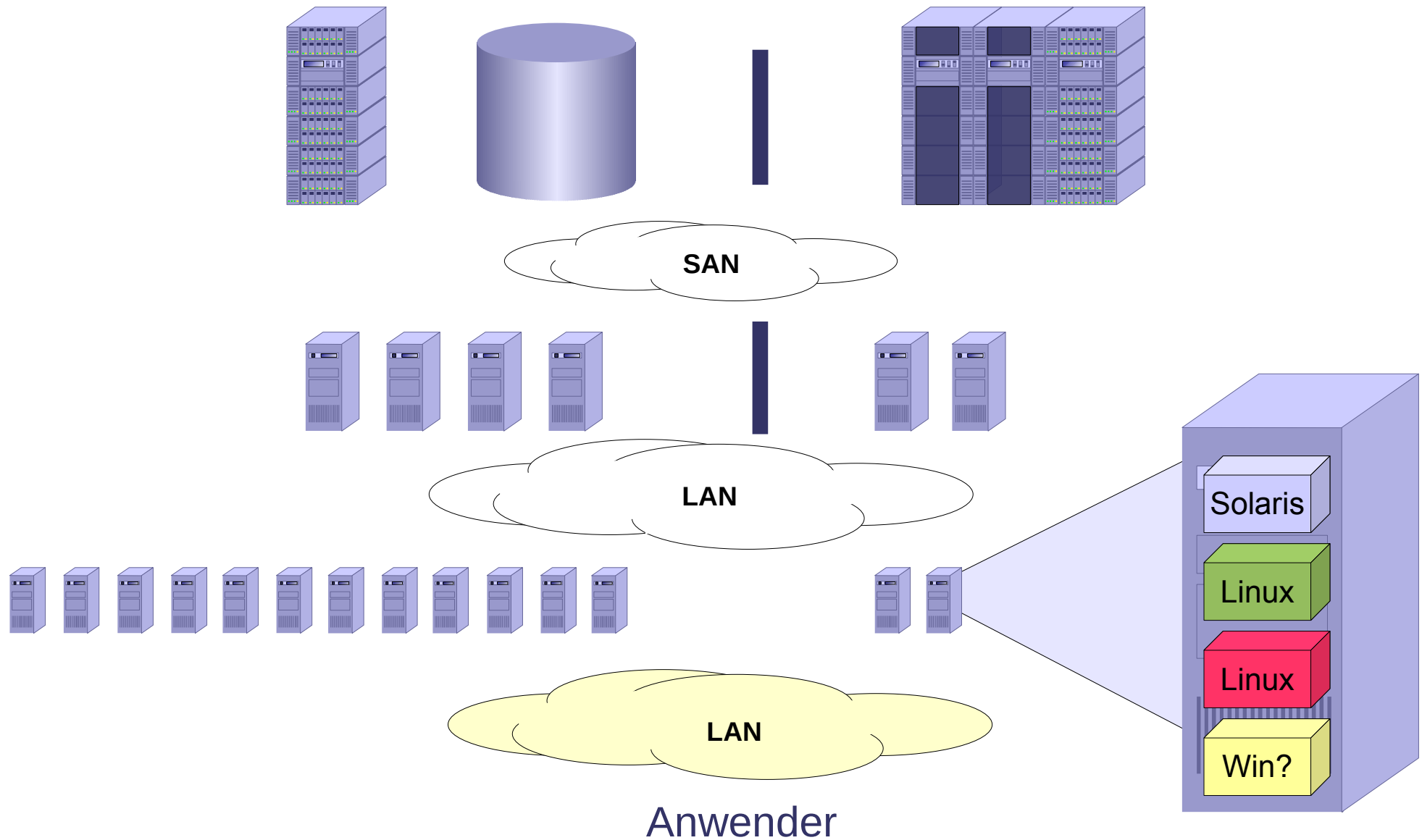
RSIO – Entdecke die Möglichkeiten

Für Freidenker: Storage-Engpässe überbrücken / Storage einsammeln



RSIO - Passend zu Virtuellen Maschinen

Drastische Vereinfachung der Storage-Anbindung für virtuelle Maschinen



RSIO - Passend zu Virtuellen Maschinen

Drastische Vereinfachung der Storage-Anbindung für virtuelle Maschinen

- kaum eine VM hat native SAN-Devices

- LAN kein Problem

- mit RSIO individuelle Devices

- einfacher Zugriff

- geeignet für Konsolidierung (MT)

- Vorteile mit 10GB-Boards z. B. durch VMDq

Anwender

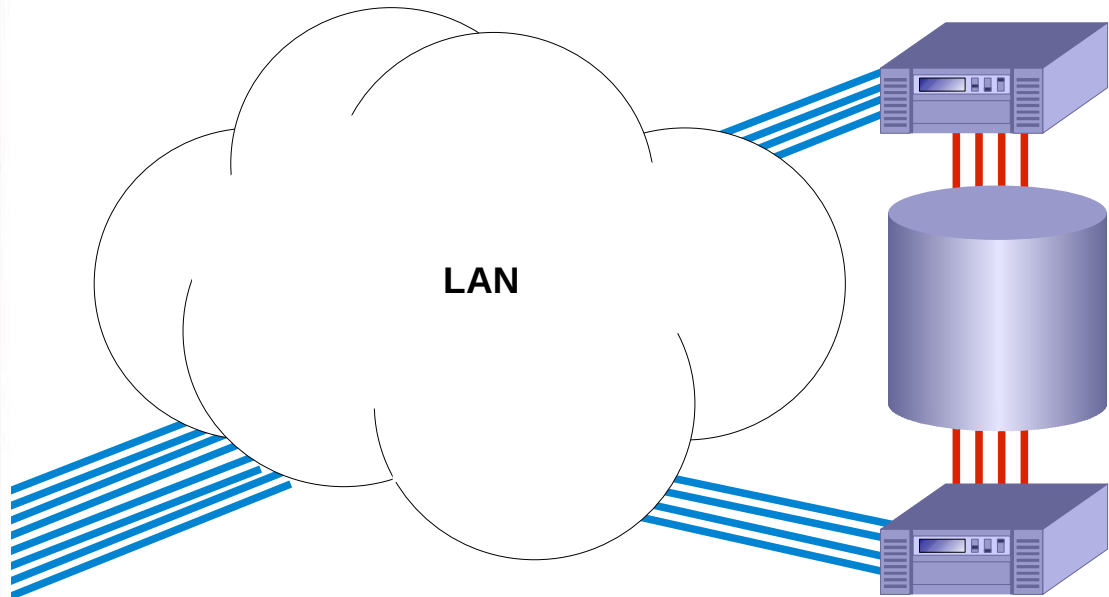
RSIO + SC - Passend zu Realen Maschinen

Storage, Management und HA für Cheap Server Farms*



CX-1000: 38 x ca. 300 RIP -> ca. 11.500 RIP*

zum Vergleich: M9000 32x SPARC64 VI 2400MHz ca. 1200 RIP



**RSIO bringt Storage heran
OSL SC liefert geeignetes Framework
für Management, Backup/DR, HA***

* Konzept-Idee

OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Zusammenfassung RSIO und OSL SC 4.0

- *noch nie war es so einfach, Standard-Server in leistungsfähige Storage-Server zu verwandeln*
- *noch nie war es so einfach und so preiswert, Server performant an Storage anzubinden*
- *noch nie bildeten Storage-Server und Storage-Clients einen so hochintegrierten, leistungsfähigen Cluster*
- *noch nie gab es alles über ein Kabel (Virtual Storage, HA-Cluster, Admin)*
- *noch nie gab es diesen Grad an Applikationsbewußtsein und diese Verflechtung mit HA/DR in der Storage-Administration*
- *noch nie gab es so eine weitreichende Integration von Speichervirtualisierung und Backup to Disk/Tape*
- *noch nie gab es diese Verflechtung von Solaris und Linux*
- *noch nie bot eine Clusterlösung diese Offenheit für verschiedene VMs*
- *noch nie gab es so vielfältige Kombinationsmöglichkeiten*
- *noch nie gab es diesen Komfort auf mehreren Plattformen*



OSL RSIO

Remote Storage I/O



Storage-Networking der nächsten Generation

Wir freuen uns auf Projekte und Gespräche:

**SNW Europe
iX Solaris Day**

**Frankfurt/M.
Stuttgart**

**26./27.Oktober
30.9. - 1.10.**

Bert Miemietz

OSL Gesellschaft für
offene Systemlösungen mbH