



OSL Storage Cluster 3.1

Just another Cluster Framework?

Sascha Viehweger
Vertriebs- und Partnermanager

Gliederung

Ordnung ist das halbe Leben

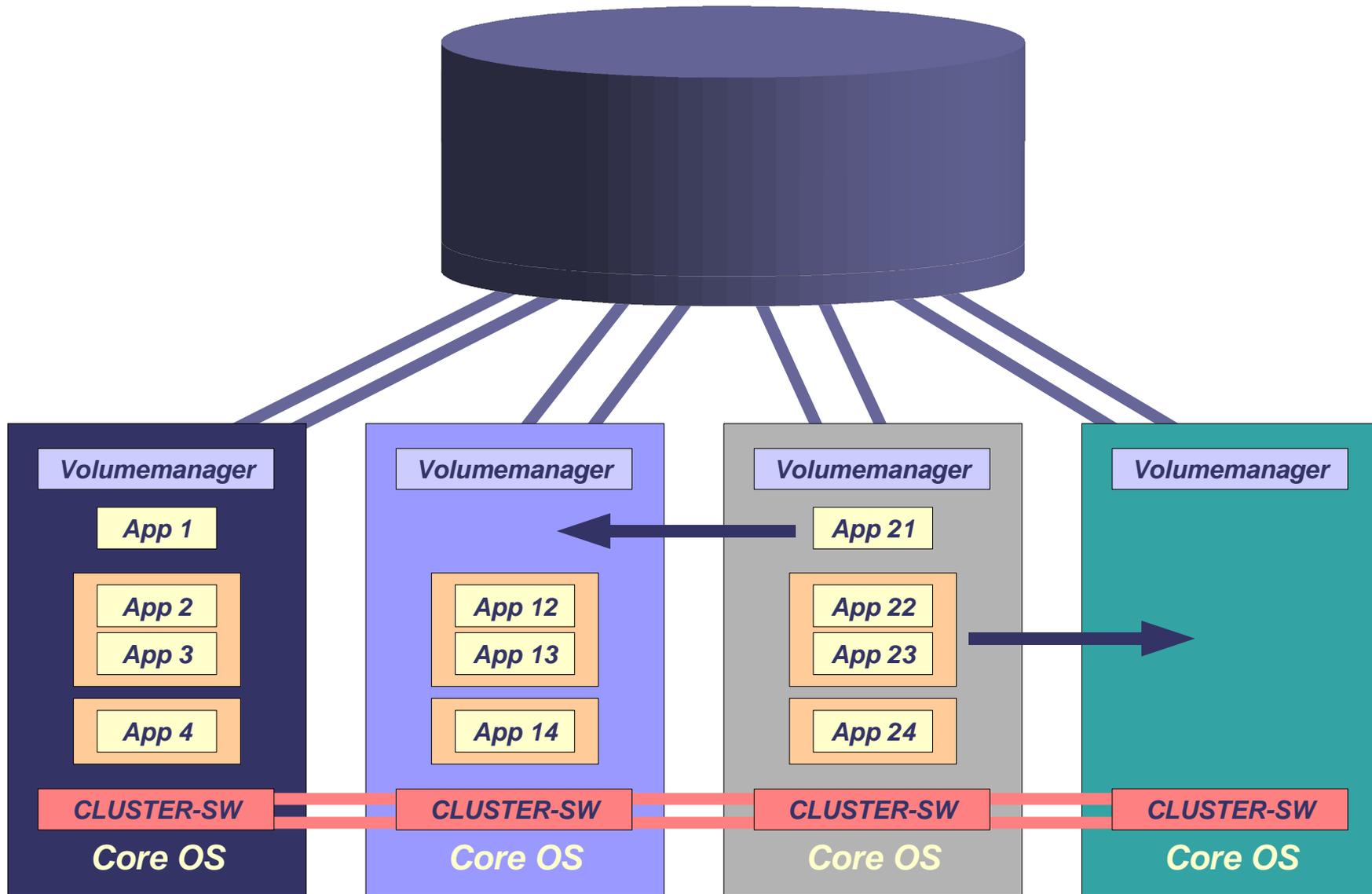


- *Cluster-Konzept*
- *Hochverfügbarkeit*
- *Servervirtualisierung*
- *Speichervirtualisierung*
- *Datenspiegelung*
- *Systemkopien*
- *Backup-to-Disk*
- *Bandbreitensteuerung und Nutzermanagement*
- *File- und Storage-Server*
- *Support & Kundendienst*

OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Klassische Cluster-Architektur

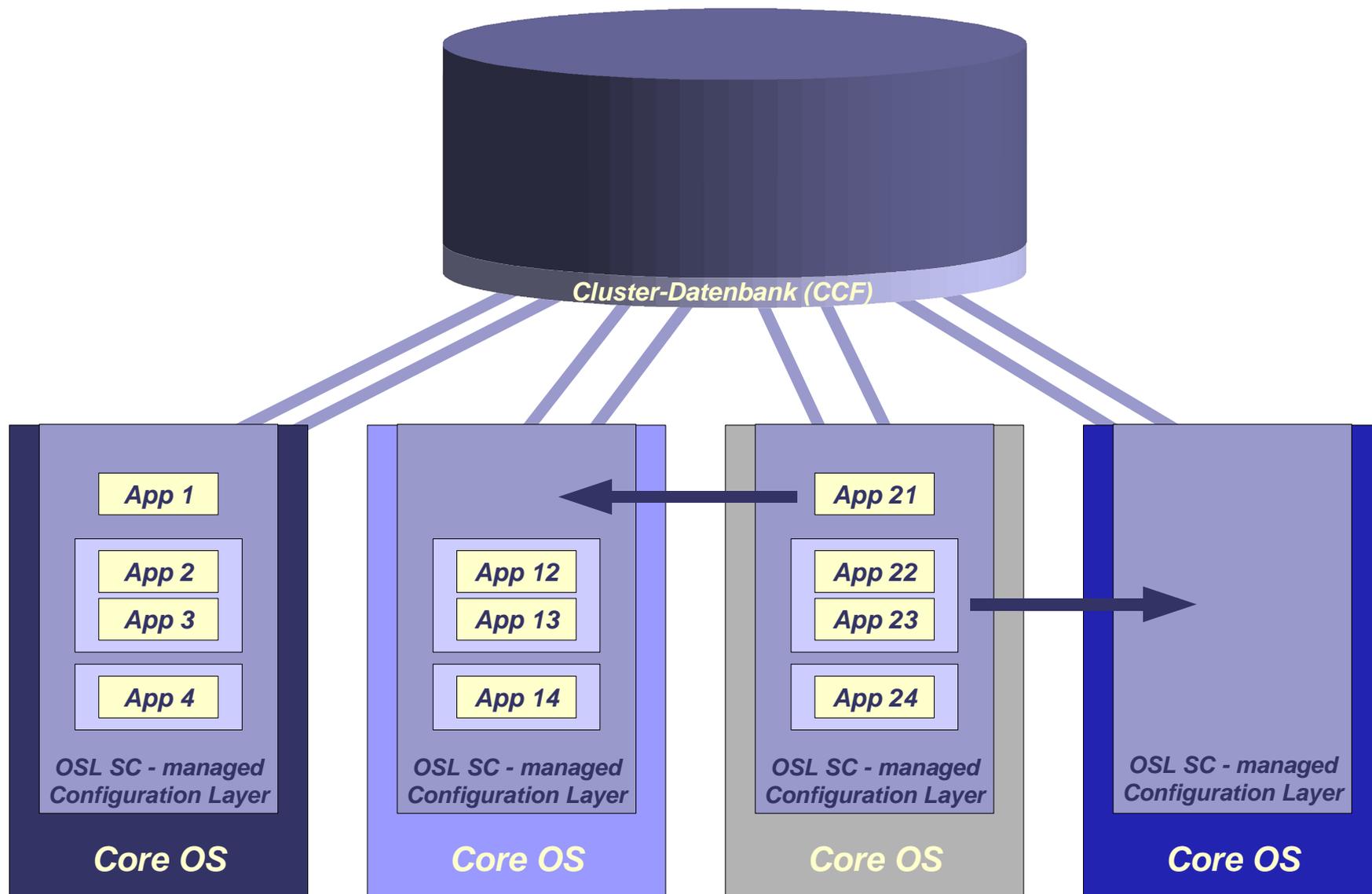
Peer-to-Peer oder Client-Server



OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Das Organisationsprinzip im OSL Storage Cluster

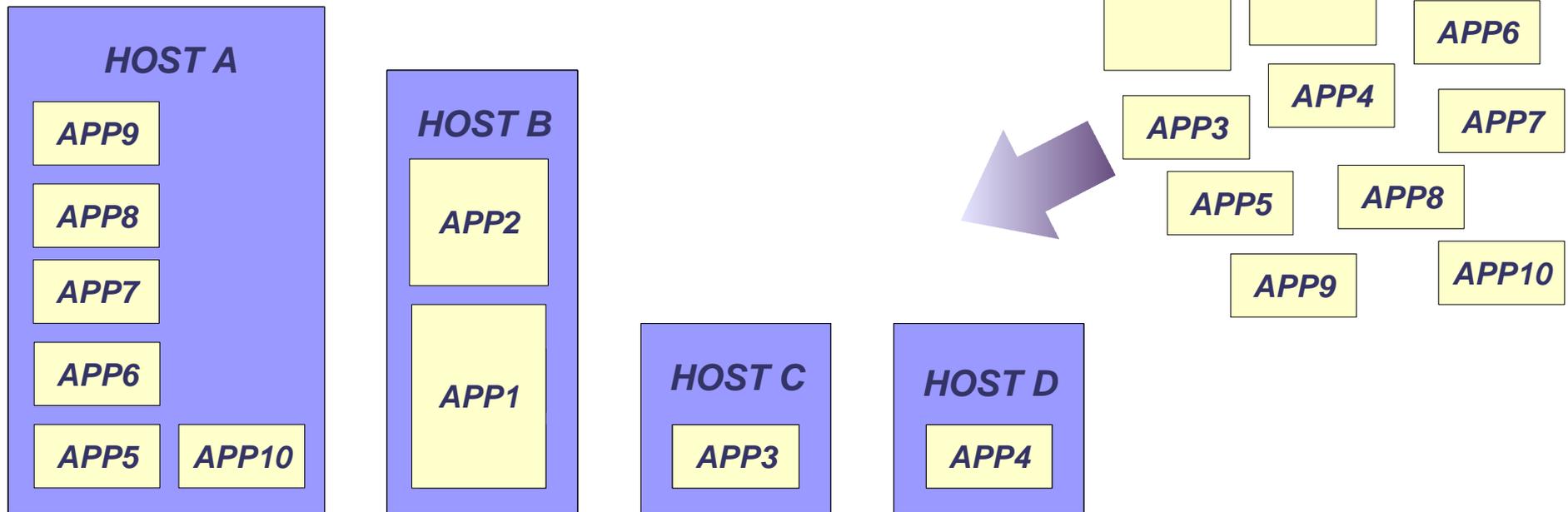
Zentralisation ersetzt ein Peer-to-Peer-Modell



OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Application Control Option

Hochverfügbarkeit – von ganz einfach bis optimal strukturiert



Shared Storage Clustering

Was bringt's?



- *Verschiedenste Cluster-Szenarien*
 - *aktiv/aktiv*
 - *aktiv/passiv*
 - *bedarfsgerecht angepasste Konfigurationen*
- *Adaptive Computing per ressourcenbasiertem Selbstmanagement*
- *Unmöglichkeit sonst clustertypischer Fehlerszenarien (Split Brain)*
- *Einsparung dedizierter Interconnect-Hardware*
- *Zugriffsmöglichkeit auf Server ganz ohne Netzwerkverbindung*
- *Weniger Administrationsaufwand und Komplexität als bei VMs*
- *Hardwareunabhängigkeit*
- *Flexible und bedarfsgerecht skalierbare HV-Umgebungen*

ABER: Zentrales Speichersystem erforderlich

Servervirtualisierung – wie und warum

Ein Vergleich der Möglichkeiten



Ressourcen-Management	Betriebssystem-Virtualisierung	Weiche Partitionierung	Harte Partitionierung
Softwaretools OS-Tools (SRM)	(Branded) Zones Jails	Virtuelle Maschinen LDoms	Serverpartitionen
sehr flexible Ressourcenzuweisung online rekonfigurierbar	teilweise verschiedene OS/OS-Releases auf einem Server	unterschiedliche Betriebssysteme je Partition	unterschiedliche Betriebssysteme je Partition
sehr geringer Overhead	geringer Overhead	vergleichsweise flexible Partitionierbarkeit	maximale Separation kein Overhead
geringer Administrationsaufwand	relativ flexible Ressourcenzuweisung	relativ hohe Separation	sehr geringer Administrationsaufwand
geringe Fehleranfälligkeit	gute Separation HV-Integration möglich	relativ flexible Ressourcenzuweisung HV-Integration möglich	sehr gute Fehlertoleranz HV-Integration möglich
geringe Separation	relativ hoher Administrationsaufwand	spürbarer Overhead	recht starre Ressourcenzuweisung
nur ein OS mit einem Patchstand einsetzbar	Fehlerintoleranz	sehr hoher Administrationsaufwand	
HV-Integration teils nur eingeschränkt möglich		hohe Fehlerintoleranz	

Servervirtualisierung – wie und warum

Ein Vergleich der Möglichkeiten



Ressourcen-Management	Betriebssystem-Virtualisierung	Weiche Partitionierung	Harte Partitionierung
Software-Tools OS-Tools (SRM)	(Branded) Zones Solaris Zones	Virtuelle Maschinen LDoms	Serverpartitionen
sehr flexible Ressourcen	Ressourcenmanagement-Tools integriert in Ansätzen verschiedene OS in Zonen lauffähig	verschiedene Systeme je Partition	unterschiedliche Betriebssysteme je Partition
online rekombinierbar	<i>HV-Integration steckt noch in Kinderschuhen</i>		
sehr geringe Overhead	<i>hoher Administrationsaufwand</i>	sehr flexible Separierbarkeit	maximale Separation
geringer Administrationsaufwand	<i>bisher sehr mäßige Fehlertoleranz</i>	relativ hohe Separation	kein Overhead
OSL Storage Cluster Applikationskapselung recht deutliche Separation stabil und langfristig erprobt angeborene HV sehr geringer Administrationsbedarf <i>keine OS-Virtualisierung</i>	hohe Separation	relativ flexible Ressourcenzuweisung	sehr geringer Administrationsaufwand
	hohe Separation möglich	Rela	sehr gute Fehlertoleranz
	hoher Administrationsaufwand	HV-Integration möglich	HV-Integration möglich
	hohe Fehlertoleranz	spürbarer Overhead	recht starre Ressourcenzuweisung
		sehr hoher Administrationsaufwand	hohe Fehlerintoleranz

Speichersystem-basiert

*in aller Regel performant und einfach handhabbar
oft ohne zusätzliche Kosten im Speichersystem integriert*

Unabhängigkeit von OS-Plattformen

Virtualisierung nur innerhalb des Speichersystems

vernachlässigt restliche Speicherlandschaft

fehlender Applikationsbezug

SAN-basiert

Virtualisierung der kompletten Speicherlandschaft

Unabhängigkeit von OS-Plattformen

mit teils hohen Kosten verbunden

oftmals komplex (z.T. Administration von Appliances)

fehlender Applikationsbezug

Host-basiert

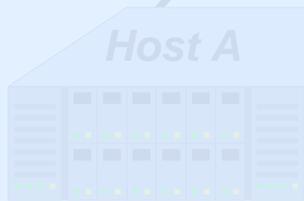
Virtualisierung der kompletten Speicherlandschaft

Sichtweise der Anwendung ermöglicht Applikationsbezug

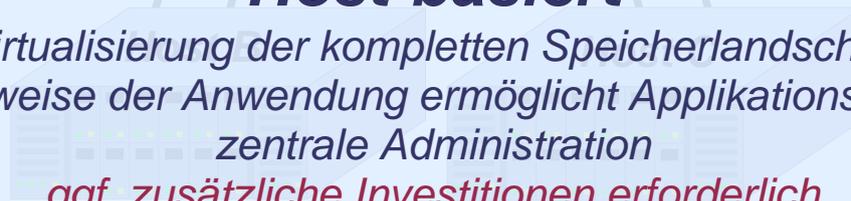
zentrale Administration

ggf. zusätzliche Investitionen erforderlich

Abhängigkeit von der OS-Plattform



Host A



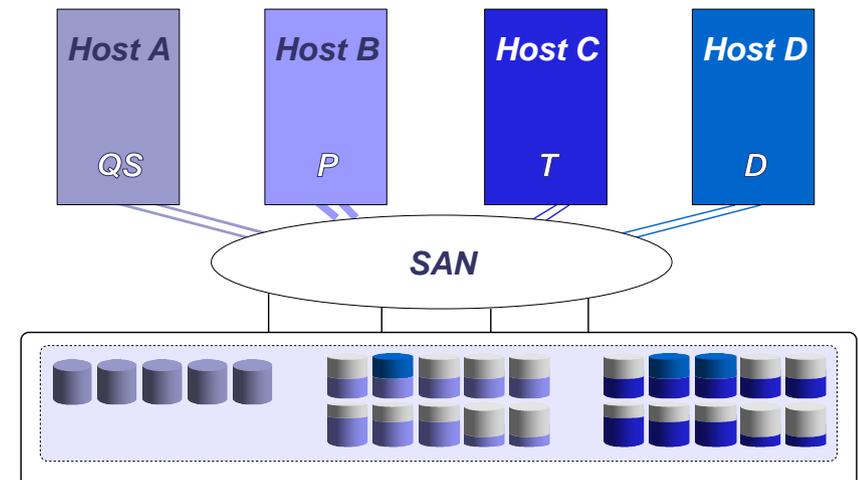
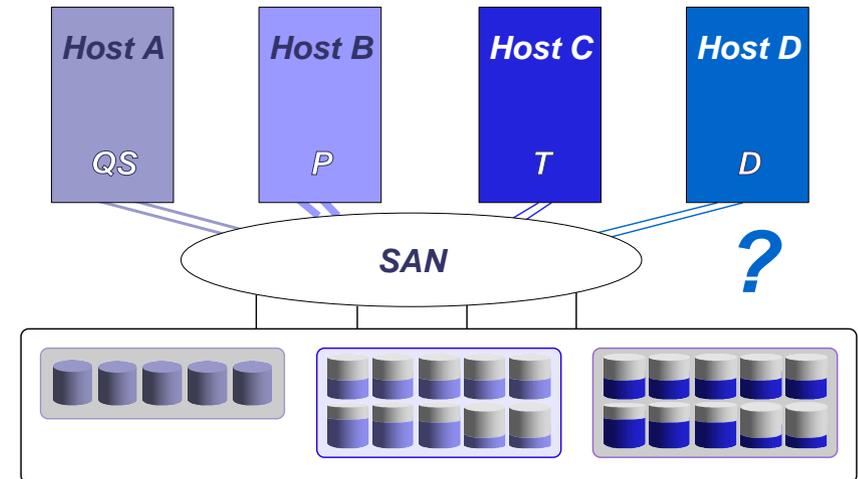
Host D

Storage Management per OSL SC Basispaket

Bedarfsgerechte und zeitgemäße Speichervirtualisierung



- Globale Storage Pools
- Speicherzuschnitt und -zuweisung je nach Bedarf
- Online Volumes vergrößern
- Aggregierte Durchsätze und Load Balancing mit integriertem Multipath-Treiber
- Performancegewinne durch Striping
- Bandbreitensteuerung für einzelne Volumes und ganze Applikationen
- Ganz ohne jeden Stress Daten verschieben, auch zwischen verschiedenen RAID-Systemen
- Verschnittfreie Ausnutzung der Speicherkapazitäten



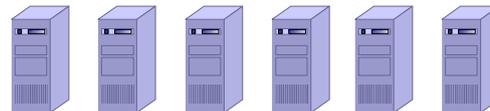
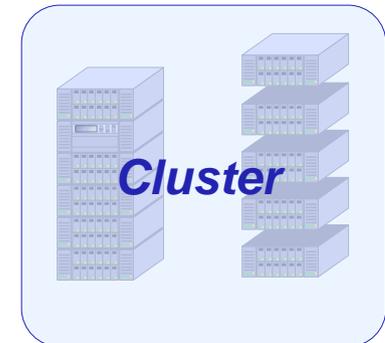
OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Warum hostbasierte Speichervirtualisierung?

Systematik, Vereinfachung und Gewinn an Funktionalität



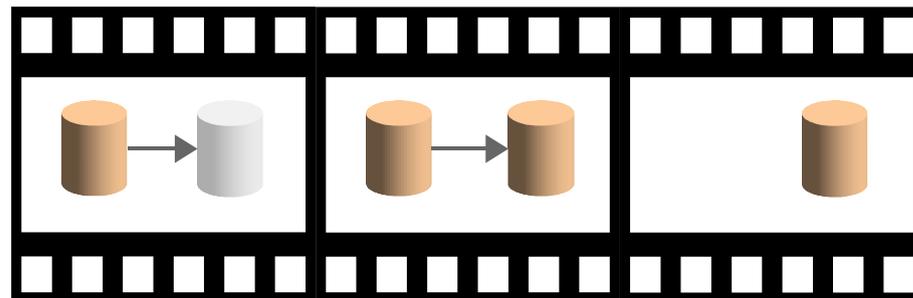
- Ermöglicht erst einheitliche Sichtweise auf den Storage
- Offenheit in der Auswahl von Speichersystemen
Beispiel Spiegeln und Verschieben von Daten
- Verbindet unterschiedlichste Speichersysteme
- Erstklassige Skalierbarkeit
- Performance-Vorteile
- Optimale Integration mit dem Betriebssystem
- Möglichkeit der Verknüpfung mit Anwendungen
- Integration mit Clustertechnologie



OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

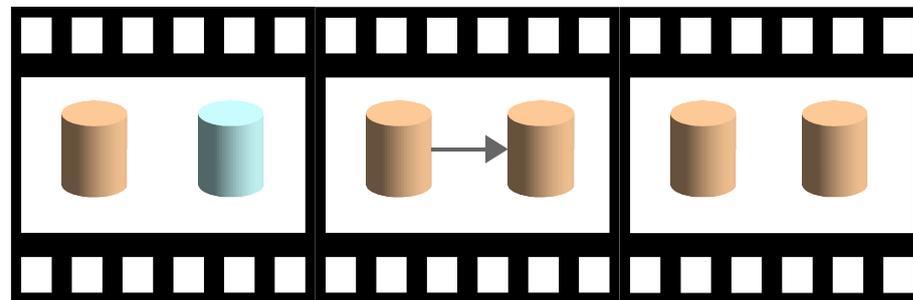
Daten verschieben

online Daten verschieben/reorganisieren
automatische Priorisierung von Anwendungs-IO



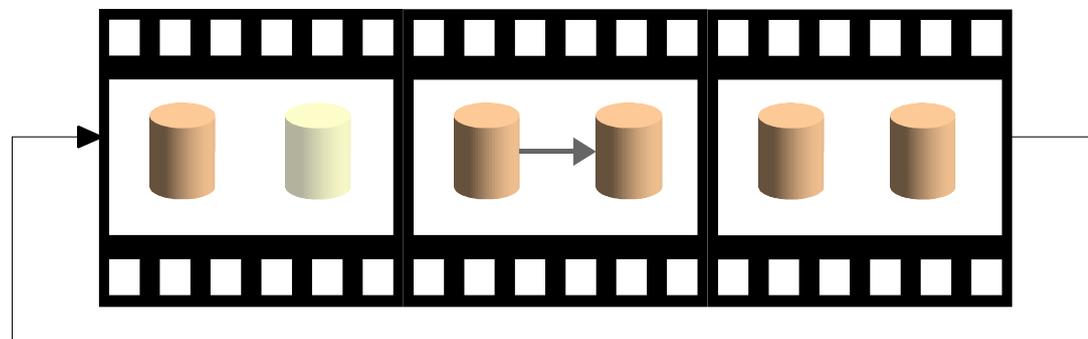
Daten clonen

einmalig online auf bestimmte Ziele kopieren
atomare Operationen für mehrere Volumes



Daten spiegeln

dauerhafte Beziehung Master -> Image
bis zu 3 Images
inkrementelle Resynchronisation
atomare Operationen für mehrere Volumes
Überbrückung von Fehlern auf dem Master

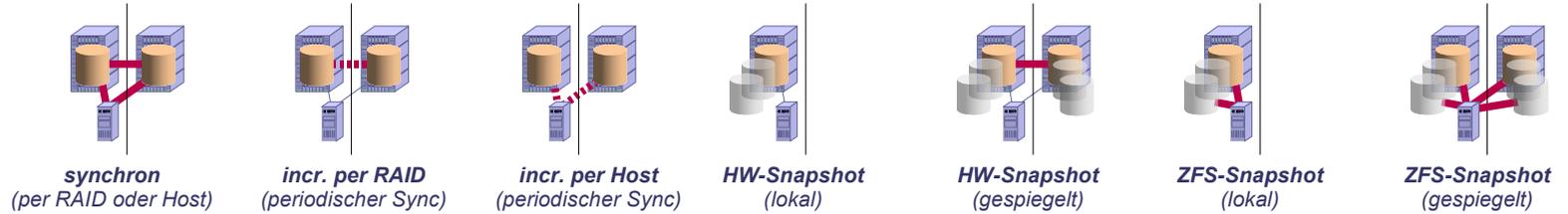


Und warum überhaupt noch Spiegel?

Spiegel vs. Snapshots – Legenden und Fakten



Snapshots und Datenkopien im Vergleich

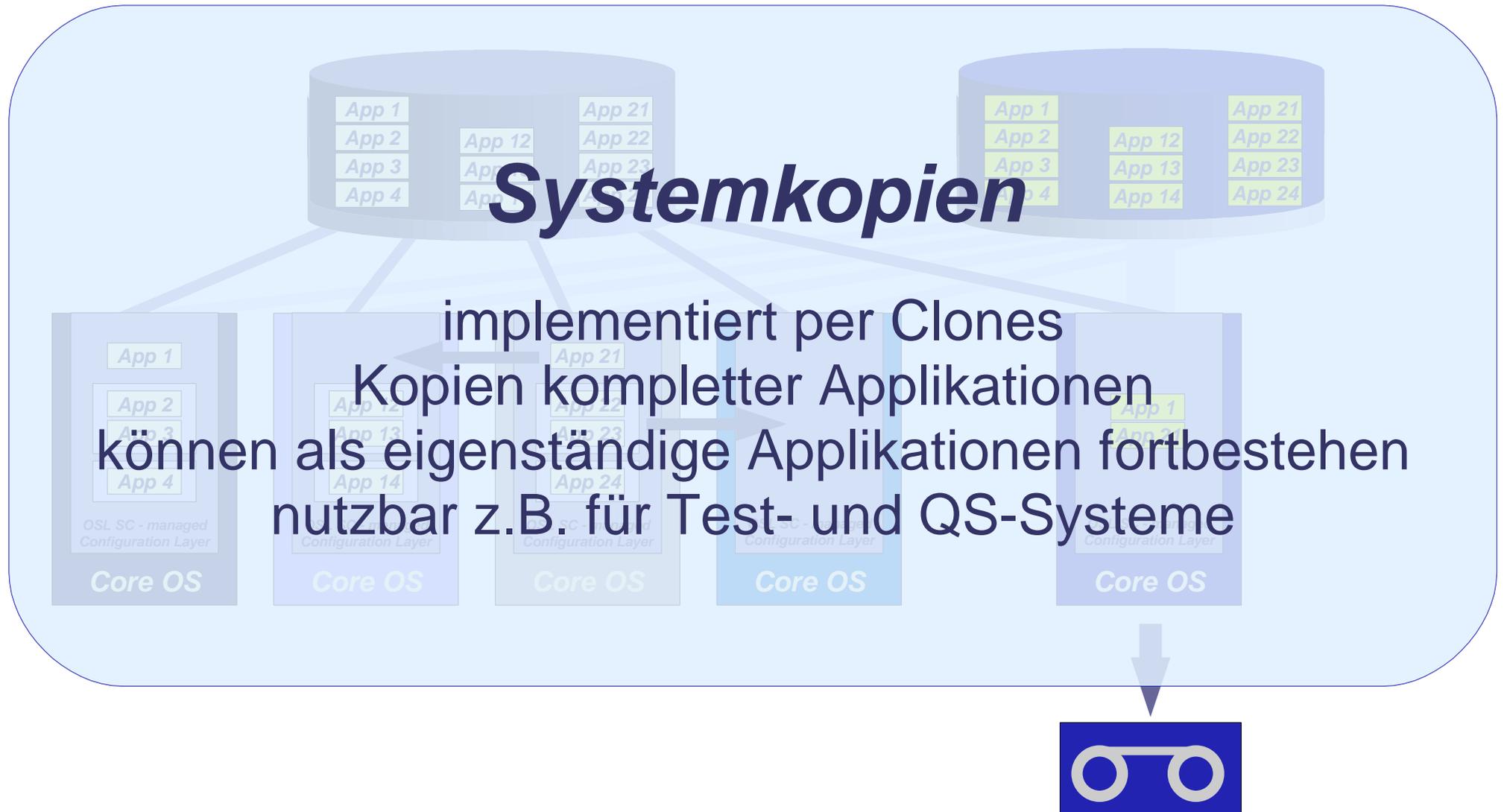


	synchron (per RAID oder Host)	incr. per RAID (periodischer Sync)	incr. per Host (periodischer Sync)	HW-Snapshot (lokal)	HW-Snapshot (gespiegelt)	ZFS-Snapshot (lokal)	ZFS-Snapshot (gespiegelt)
Speicherbedarf	200%	200-300%	200-300%	125-200%	250-400%	125-200%	250-400%
mit o.g. Speicherbedarf mögliche Kopien bzw. Snapshots	1	1-2	1-2	ca. 1-15	ca. 1-15	ca. 1-15	ca. 1-15
OLTP-Performance Original	o	++	++	++	o	++	+
OLTP-Performance bei gleichzeitigem Zugriff auf Original und Kopie/Snapshot	X	++	++	+	o	o	-
simultaner Zugriff auf Original und Kopie/Snap vom gleichen Host	nein	ja	ja	ja	ja	ja	ja
simultaner Zugriff auf Original und Kopie/Snap von verschiedenen Hosts	nein	ja	ja	ja	ja	nein	nein
Backup-Performance Kopie/Snap (bei simultanem OLTP-Betrieb Original)	X	++	++	o	o	o	-
Integration mit Host-OS + Applikationen	o	o	++	-	-	+	+
Handhabung Komplettlösung	o	o	+	-	-	o/+	o/+
Performanceanforderung Original-Speichersystem	hoch	hoch	hoch	sehr hoch	sehr hoch	sehr hoch	sehr hoch
Performanceanforderungen an das Spiegel-Speichersystem (remote)	hoch	mäßig	mäßig	X	sehr hoch	X	sehr hoch
Verfügbarkeit Kopie/Snap nach User- oder SW-Fehler	X	++	+ / ++	++	++	+	+
Schutz gegen Ausfall Ausfall Original-Speichersystem	++	+	+	X	++	X	+ / ++
Brauchbarkeit Snap/Kopie als Sicherung	X	++	++	--	++	-- / -	- / o
Belastung Host	sehr gering	sehr gering	gering	sehr gering	sehr gering	mäßig	mäßig

++ sehr gut + gut o mäßig - schwächer -- schlecht X entfällt/nicht vorhanden

Hostbasierte Spiegelung im Rechenzentrum

Lösungsangebote für vielfältige Problemstellungen



OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Backup-to-Disk

Images mit identischem Namensraum
sofort wieder anlauffähig (ready for restart)
beste Performance mit inkrementeller Synchronisation
keine Belastung der produktiven Hosts

Archivierung

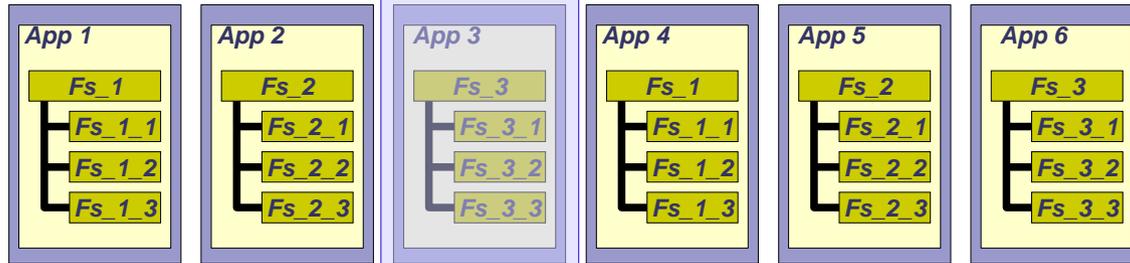
Anschlusslösung für Tape-Archivierung implementierbar

Application Aware Storage Virtualization

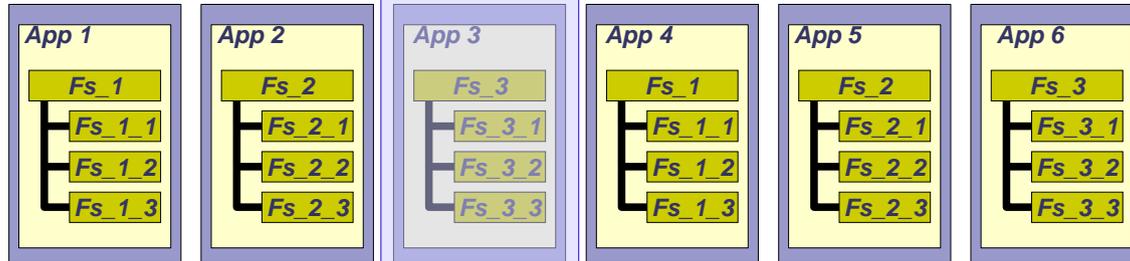
Anwendungsbeschreibungen und Volume Management integriert



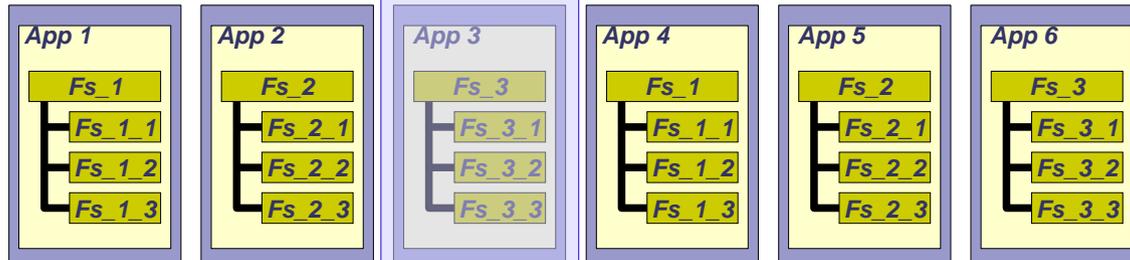
**Universe 0
Produktion**



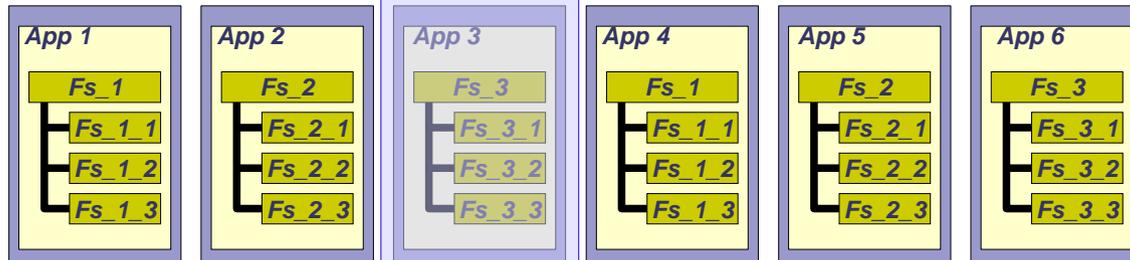
**Universe 1
Backup 1**



**Universe 2
Backup 2**



**Universe 3
DR-Spiegel**

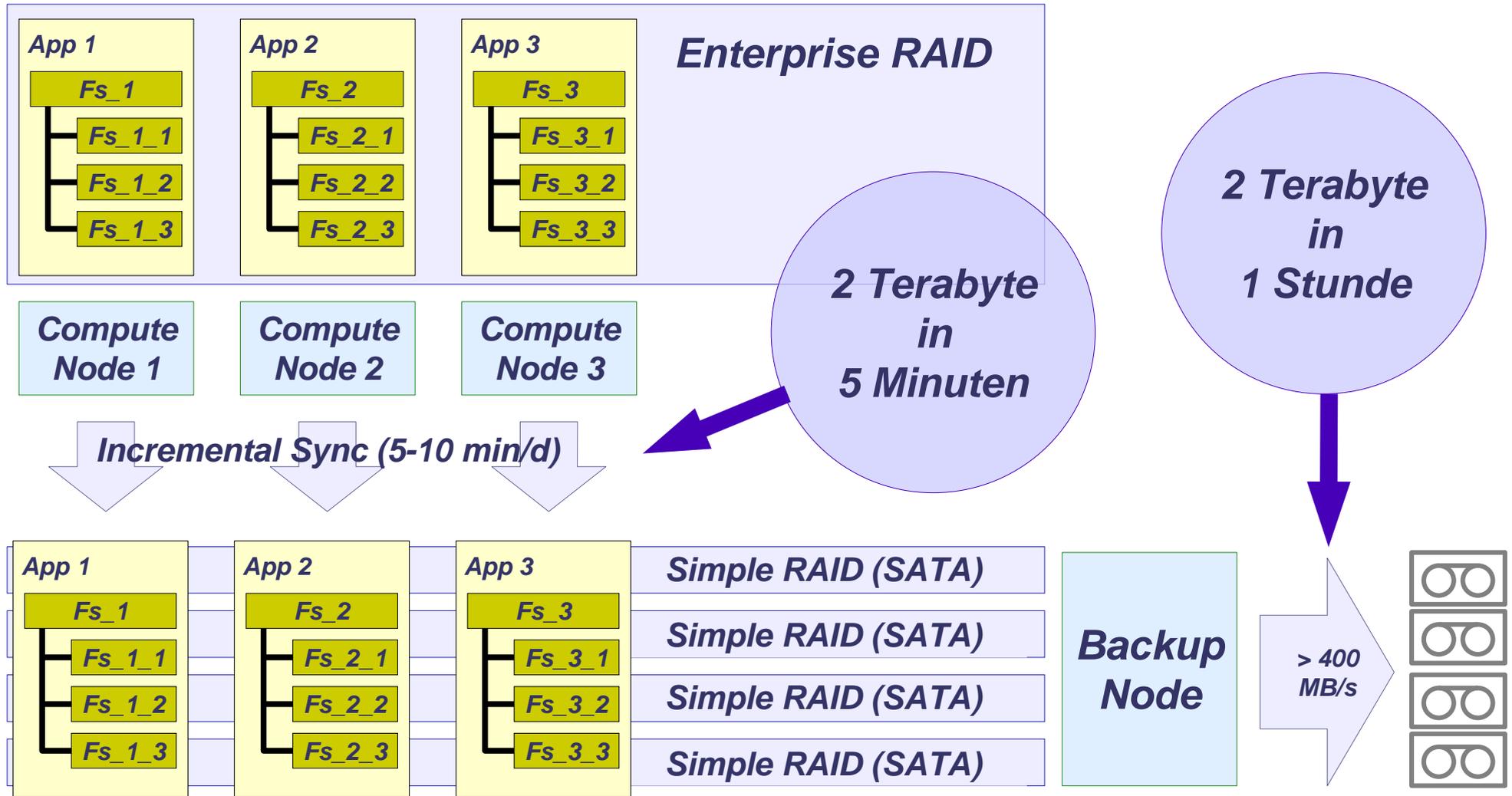


Application Specific View

OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

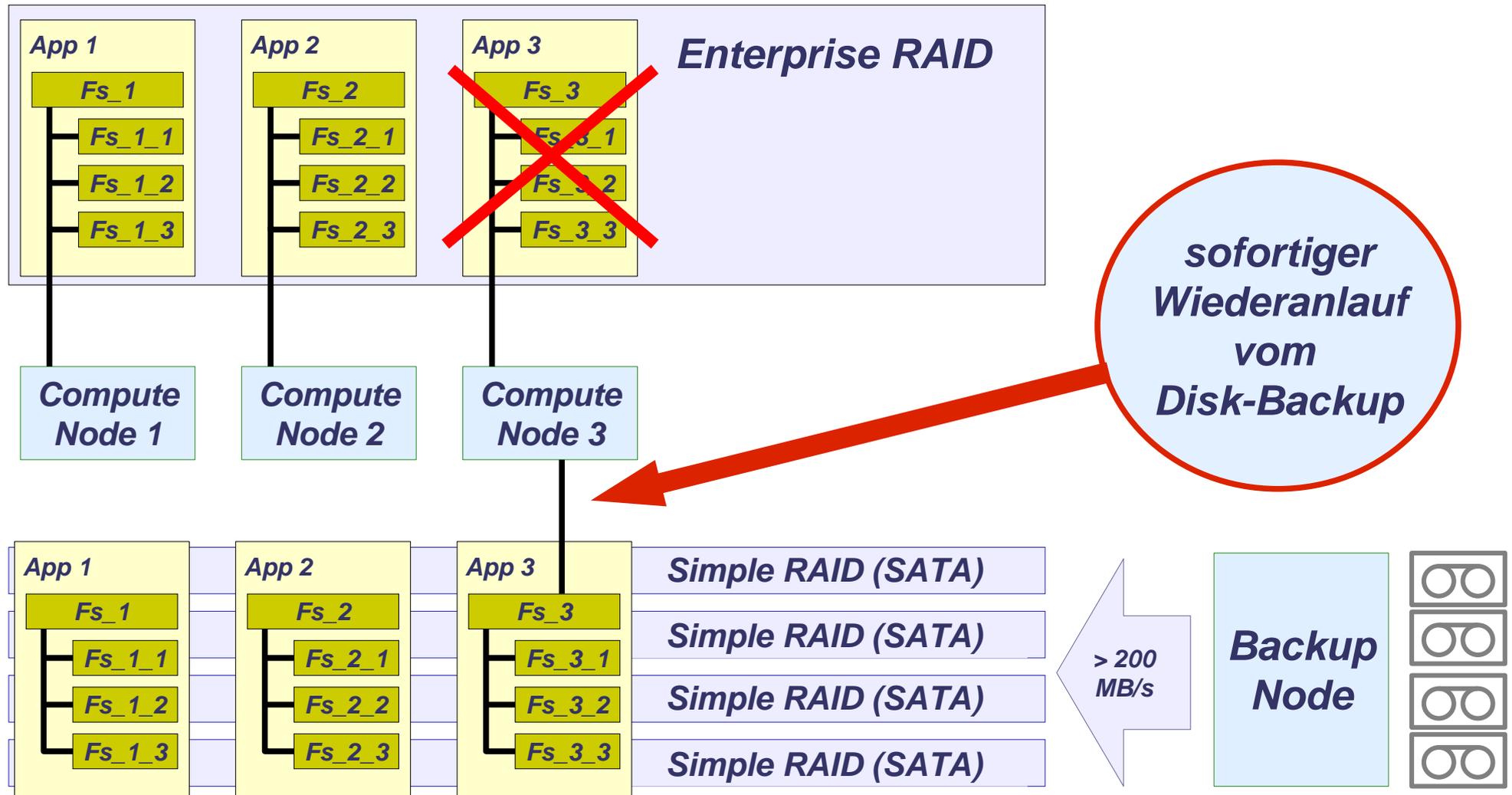
Integriertes Highspeed-Backup ...

Hochverfügbar, kostengünstig, performant



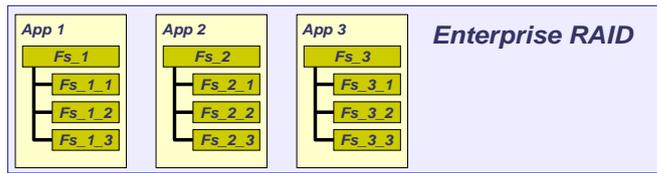
... und Instant Recovery

Keine Ausfallzeiten, Recovery im Hintergrund wieder laufender Anwendungen



Differenzierung zu herkömmlichem Backup

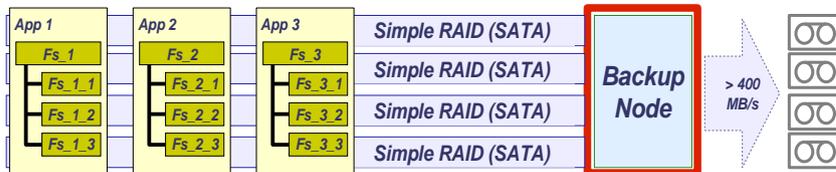
Möglichkeiten von Festplatten konsequent nutzen



Compute Node 1 Compute Node 2 Compute Node 3

Incremental Sync
(5-10 min/d)

Nur ein Backup Node !



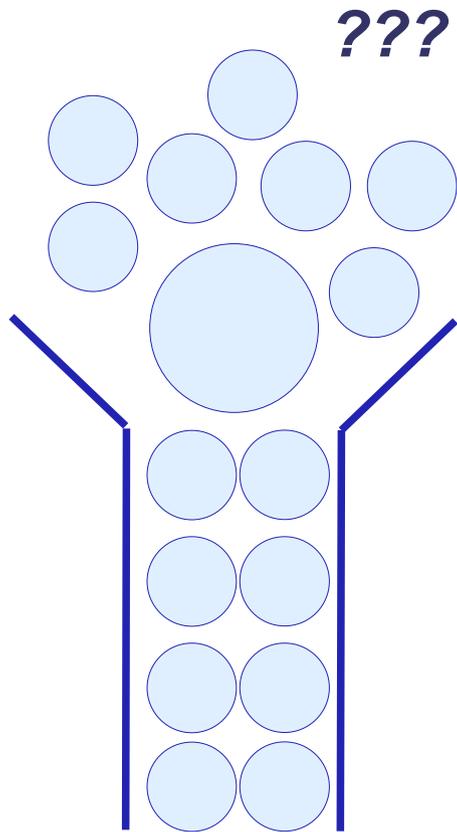
- **Extrem kurzes Backup für produktive Knoten**
- **Minimale CPU-Lastung auf den produktiven Knoten**
(keine Verarbeitung der Daten)
- **Nutzung SAN statt LAN**
- **Kein Backup-Client auf den produktiven Knoten**
(kein dezentrales Pflegen von Konfigurationen)
- **Atomares Backup – Dauer: NULL -> konsistenter Zustand**
- **ready-for-restart-Images der Applikation**
damit extrem schneller Wiederanlauf
- **SW für Tape-Backup nur auf DASI-Server**
- **Zentrale Administration**
- **Extreme Durchsätze bei Tape-Backup/-Restore möglich**
- **Niedrige Anforderungen an Backup-RAID**
 - ermöglicht SATA mit hoher Dichte
 - niedriger Platzbedarf
 - kürzere Backup-Zeiten
 - reduzierter Stromverbrauch / Wärmeabgabe
- **Adaptive Fähigkeiten bzw. "selbstlernend"**
- **Integration mit HV**
- **Leicht zu DR-Umgebung ausbaubar**

Bandbreitensteuerung

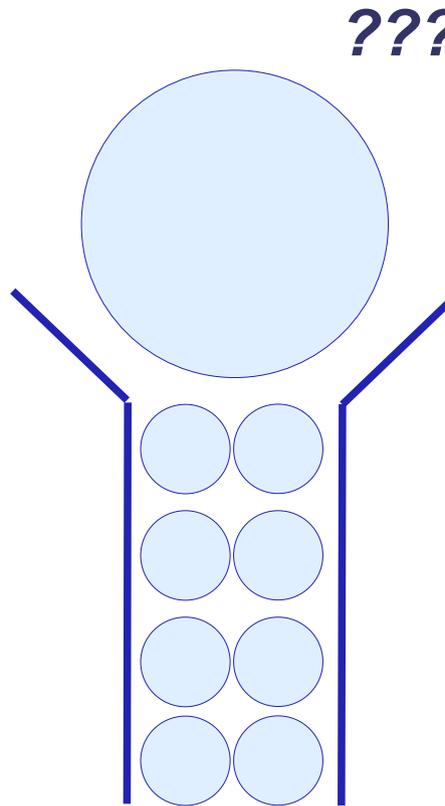
Flaschenhalse vermeiden, Lastspitzen verteilen



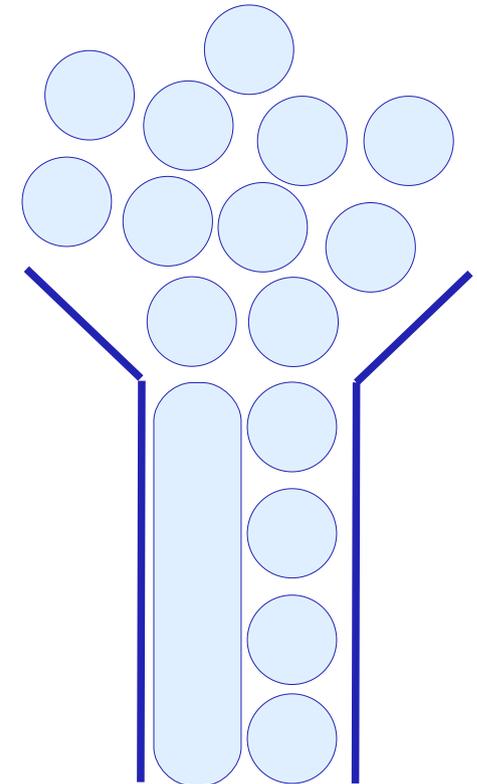
ohne Bandbreitensteuerung



ohne Bandbreitensteuerung



mit Bandbreitensteuerung



OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Bandbreitensteuerung

Vielfältige Möglichkeiten



- **Warum?**

- Sättigung IO-Kanäle
- Sättigung Speichersystem(e)
- Konkurrenz Applikationen

- **Was?**

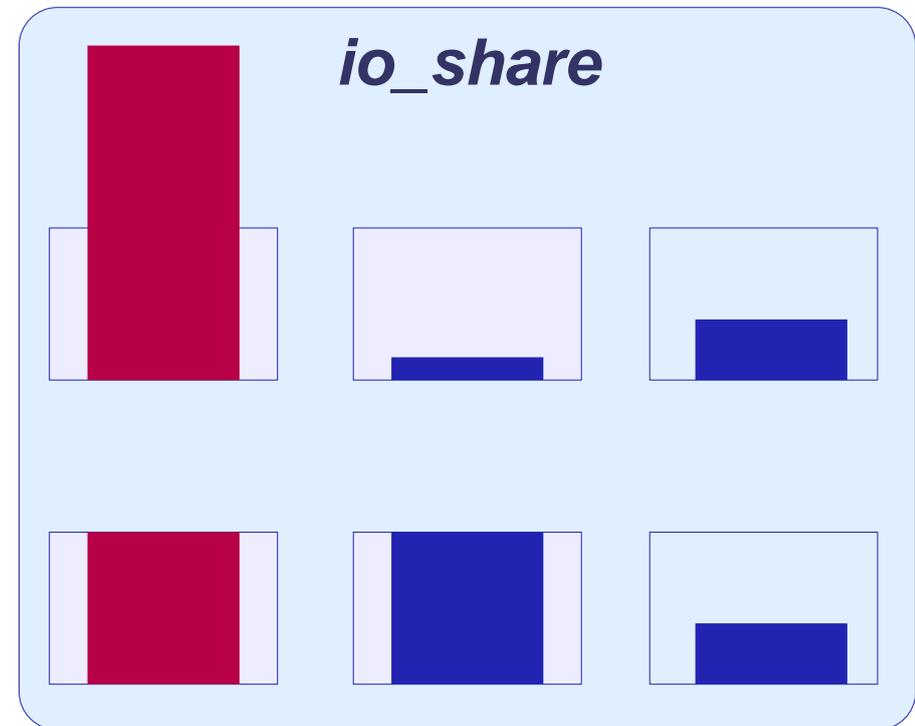
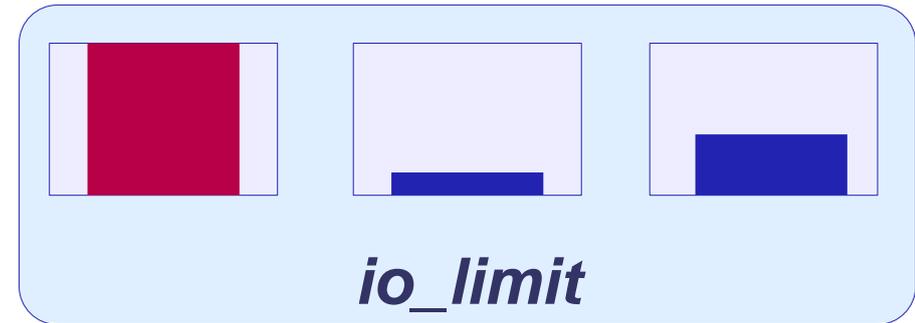
- einzelne Volumes
- Gruppen von Volumes
- Applikationen

- **Wie?**

- absolute Bandbreite (*io_limit*)
- adaptives Konzept (*io_share*)
- Limit für Synchronisationsvorgänge (*sync_limit*)

- **Mit welchem Resultat?**

- verbessertes Antwortzeitverhalten
- faire Verteilung von IO und CPU-Bandbreite
- reduzierte CPU-Belastung
- gesteigerter Gesamtdurchsatz

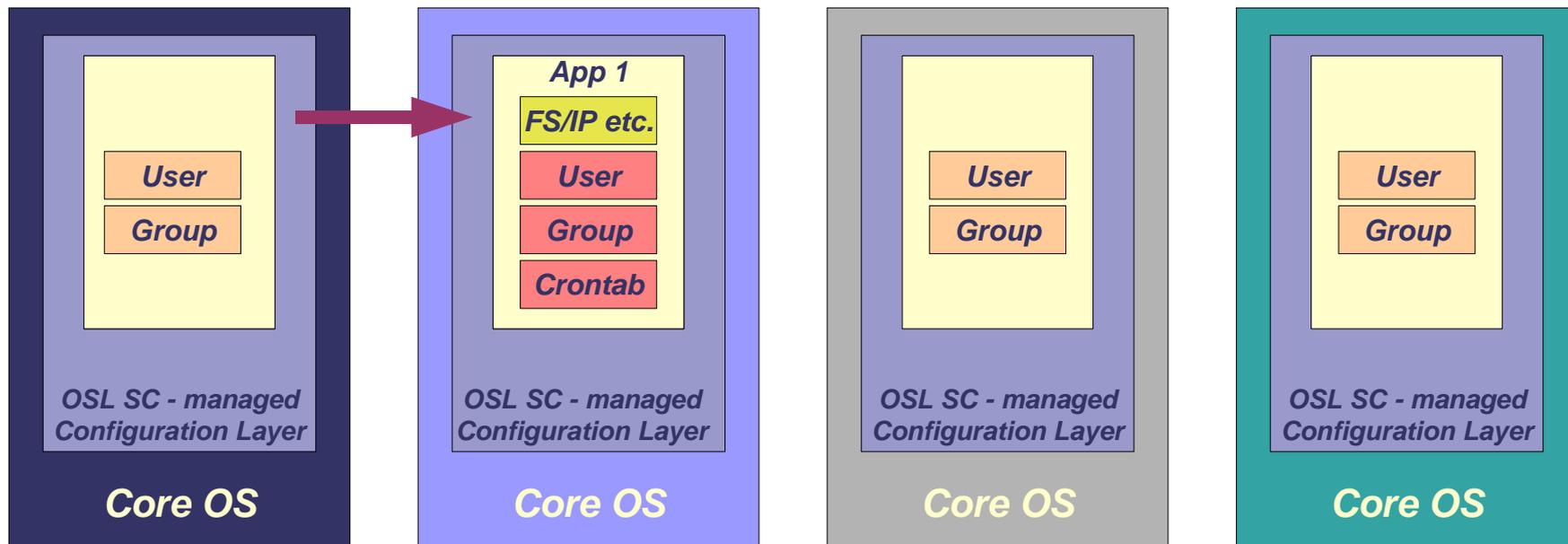


Global User Management

Globale Nutzerverwaltung komplettiert virtuelle Ablaufumgebungen



- Geeignet für Server / Application Service User
- Unabhängig von externen Services wie NIS/LDAP/ADS
- Vermeidung von Konflikten, Synchronisation, automatische Reparatur
- User kann einer Applikation zugeordnet werden
- Crontab und Login-Möglichkeit wandern mit der Applikation
- Auch nach Neuinstallation sofort wieder verfügbar



OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Fileserver und virtueller Anwendungsspeicher

Was OSL Storage Cluster mit verschiedenen Betriebssystemen kann



- *Welche Möglichkeiten gibt es?*
 - *Fileserver für Solaris, Linux und Windows Clients*
 - *Traditioneller Fileserver -> gemeinsames Dateisystem für viele Clients*
 - *Shared raw-Devices -> nutzbar als hochverfügbarer Speicherplatz über das LAN*

- *Welcher Techniken kommen zum Einsatz?*
 - *NFS als Fileserver für Linux oder Unix Clients*
 - *NFSv4 - falls unterstützt (höhere Sicherheit, einfacherer Benutzerrechte)*
 - *Samba als Fileserver für Windows (seit Solaris 10u4 mit ADS Unterstützung)*
 - *iSCSI für raw-Device Shares*

- *Warum im OSL Storage Cluster?*
 - *Möglichkeit eines hochverfügbaren Fileservers*
 - *Über ein Template einfach zu implementieren*
 - *Integriertes Backup und Recovery mit hostbasierten Spiegeln*
 - *Die Größe der Volumes kann einfach, flexibel und nach Bedarf angepasst werden*
 - *Skalierbarkeit*
 - *Als hochverfügbare Anwendung implementierbar – keine dedizierte Hardware notwendig!*

HA Fileserver im OSL Storage Cluster

Warum der OSL SC eine hervorragende Basis für Dateiserver darstellt



Mögliche Konfigurationen

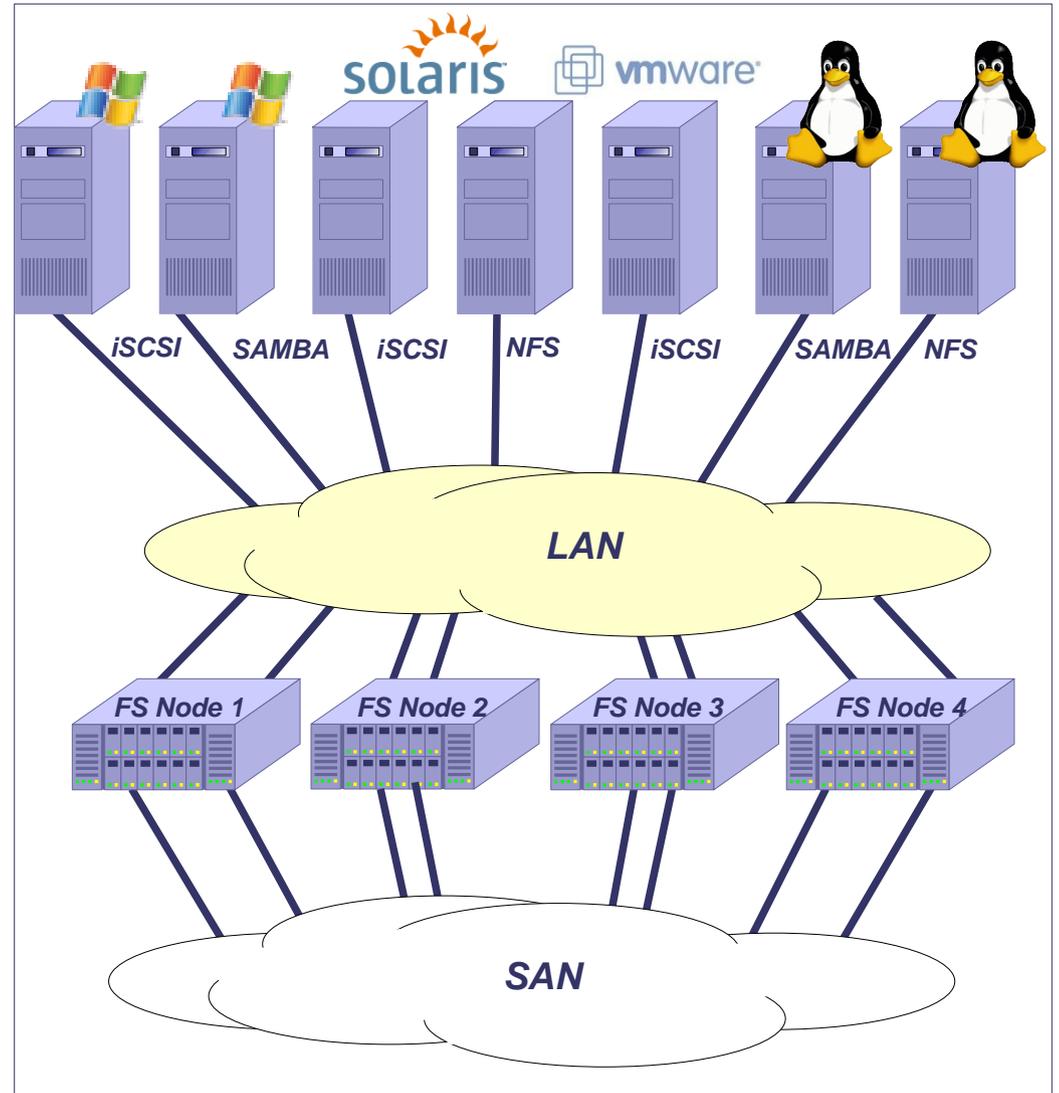
- Skalierbarer Fileserver Cluster
- Per NFS und Samba shared file systems
- Per iSCSI shared raw-Devices

Mögliche Clients:

- Linux: iSCSI, NFS, Samba
- Solaris: iSCSI, NFS
- Windows: iSCSI, SAMBA, (NFS)
- VMware ESX: iSCSI (VMFS), NFS (VMDK)

Vorteile:

- Einfaches Backup mit asynchronen Spiegeln oder ZFS Snapshots
- Ausfallsicherheit durch synchrone Spiegel
- Leicht in heterogene Strukturen integrierbar
- Hochverfügbare FS Applikation im OSL SC
- Einfaches Nutzermanagement dank Clusteruser und -gruppen



OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

OSL Storage Cluster 3.1

Übersicht über die wesentlichen Funktionen



Application Awareness Bandbreitensteuerung User-Management	Application Control Option	Application Mirrors Application Clones B2D / DASI / DR -Tools
	<i>clusterweite Steuerung von Anwendungen</i>	
	<i>virtualisierte (hardwareabstrakte) Ablaufumgebungen</i>	
	<i>Hochverfügbarkeit</i>	
	<i>ressourcenbasiertes Selbstmanagement</i>	
	Application Resource Description	

Clusterfähige Speichervirtualisierung	Extended Data Management
<i>globale (hostübergreifende) Storage Pools</i>	<i>Integration RAID-basierter Datenkopien / Snapshots</i>
<i>Global Disk Inventory</i>	<i>Hostbasierte Datenspiegelung</i>
<i>Global Devices / Global Namespace</i>	<i>Live Data Migration</i>
Cluster-Volumemanager mit automatischer Allokation	<i>Data Clones</i>
<i>Disk Access Management</i>	
<i>IO-Multipathing</i>	

nbH

www.osl.eu

Support & Kundendienst

Qualität und Individualität statt Servicewüste



- *Maintenanceverträge mit bis zu 7x24h Erreichbarkeit und 1h Reaktionszeit über Servicepartner Fujitsu Luxemburg erhältlich*
 - *Darüber hinaus ist jedem Kunden ein Systemingenieur als fachlicher Ansprechpartner bei OSL zugeordnet*
 - *Softwareentwickler stehen so Kunden unmittelbar für Hilfestellungen zur Verfügung*
 - *Vor-Ort-Unterstützung möglich (ggf. kostenpflichtig)*
- > individueller Support in deutsch und englisch*
- > kurze Wege, schnelle Reaktionen*
- > kompetente Ansprechpartner*



OSL Gesellschaft für offene Systemlösungen mbH

***Am Goethepark 18
15566 Schöneiche b. Berlin
Tel.: +49 (0)30 740767-80
Fax: +49 (0)30 740767-89
info@osl.eu***

OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu