



Welten verbinden: RSIO, Solaris und Linux

RSIO und OSL Storage Cluster 4.0

OSL Data Center Technology
Frühjahrstour 2011 • Berlin - Paderborn - Frankfurt

Bert Miemietz

OSL Gesellschaft für
offene Systemlösungen mbH

Neue Plattformen und neue Infrastrukturen

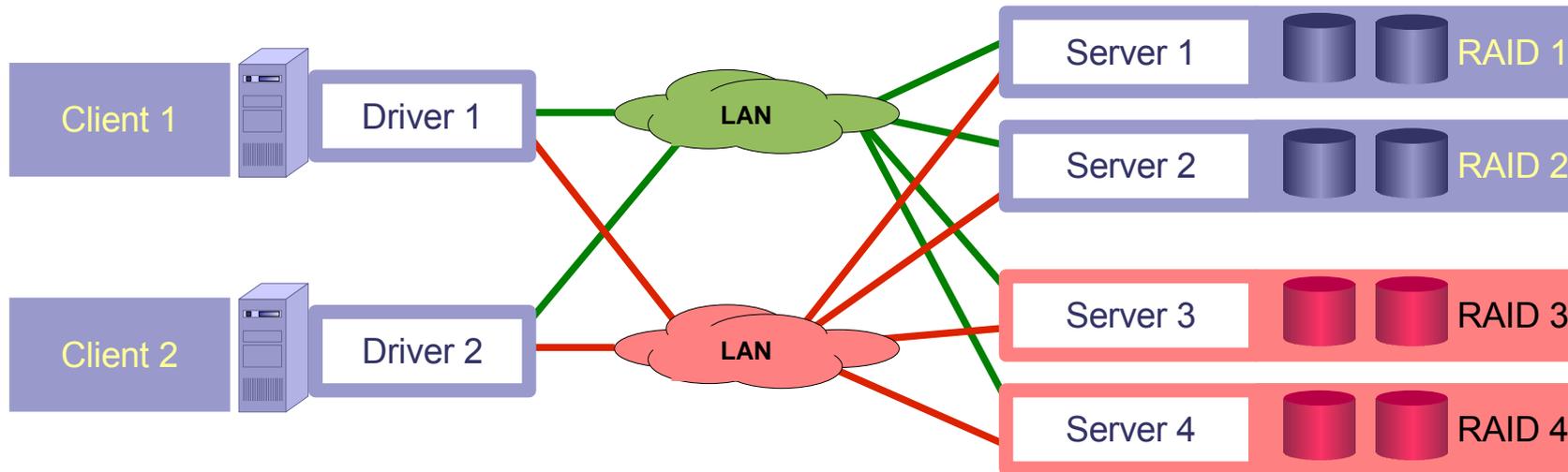
Beide Themen hängen zusammen



- *Bisher nur 3 unterstützte Betriebssysteme:*
SunOS, Solaris, OpenSolaris
- *Nachfragen zu anderen Software-Plattformen:*
*ubuntu, fedora, openSUSE, **debian**, Mandriva, LinuxMint, PCLinuxOS, **slackware**,
gentoo linux, CentOS, **Red Hat**, SLES, SLED*
Gemeinsamkeiten: - *Sammelname Linux*
- *speisen sich aus demselben Kernel-Entwicklungsstamm*
- *Gründe für den Einsatz von Linux:*
- *Hat Linux auch Nachteile?*
- *Daraus folgt: Frage nach der geeigneten I/O-Infrastruktur:*
 - *Fibre Channel -> Preis und techn. Herausforderungen?*
 - *Ethernet -> Warum nicht? Wenn, dann aber Block-I/O*

Block-I/O über Ethernet – einmal anders gedacht

Für vernetzte Strukturen auch Netzwerkparadigmen anwenden



- *I/O-Requests senden
read(), write(), ioctl()*
- *geeignete Kapselung*
- *Verbindungsauf- und Abbau,
Überwachung*
- *Kanal-Multiplexing*

- *I/O-Requests verarbeiten
read(), write(), ioctl()*
- *geeignete Kapselung*
- *Verbindungsauf- und Abbau,
Überwachung*
- *Kanal-Multiplexing*

Ist SCSI die Antwort?

- **Geräteidentifikation und -Beschreibung einfacher möglich (Adresse, Port, Datenstrukturen)**
- **stärkere Berücksichtigung Netzwerk für Admin wünschenswert**
- **viele SCSI-Daten irrelevant, dafür sind viele interessante Funktionen kaum darstellbar**
 - I/O-Requests senden
 - I/O-Requests verarbeiten
- **ohne SCSI keine Wandlung auf Low-Level-Protokoll erforderlich**
- **bestimmte SCSI-Mechanismen im Netz kontraproduktiv (z. B. Bus-Reset)**
 - Verbindungsauf- und Abbau, Überwachung
 - Verbindungsauf- und Abbau, Überwachung
- **reduzierter Kommunikationsaufwand möglich**
 - Kanal-Multiplexing
 - Kanal-Multiplexing

RSIO - Remote Storage I/O

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

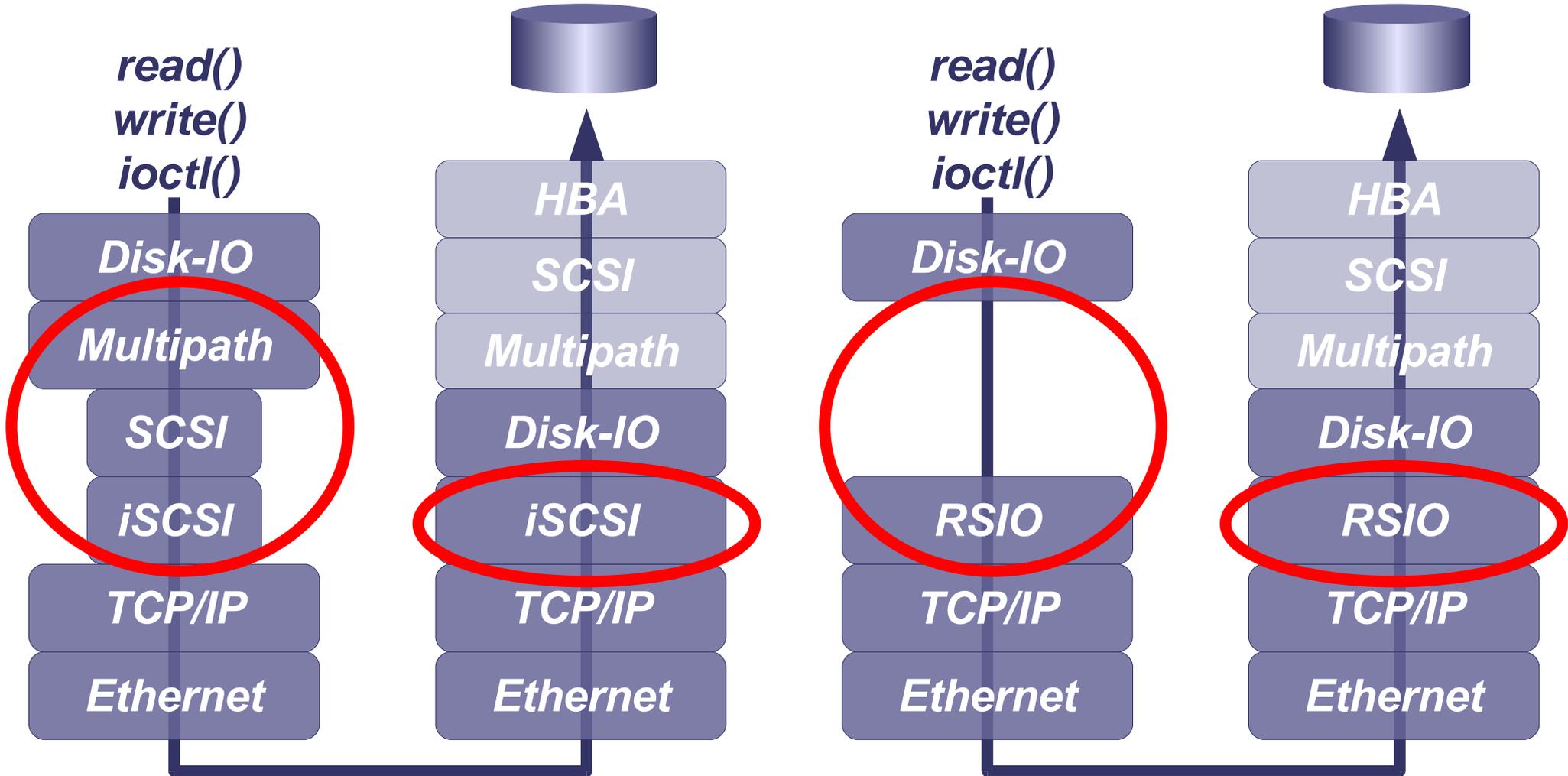


- *neues, von OSL entwickeltes Protokoll*
- *direkter Transport aller relevanten IO-Aufrufe (read, write, ioctl)*
- *integriert Verbindungsaufbau, Überwachung, Path-Multiplexing, Trunking*
- *fähig zu Selbstkonfiguration und Error Recovery*
- *kann alle modernen Storage-Szenarien abbilden:*
 - *einfache 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

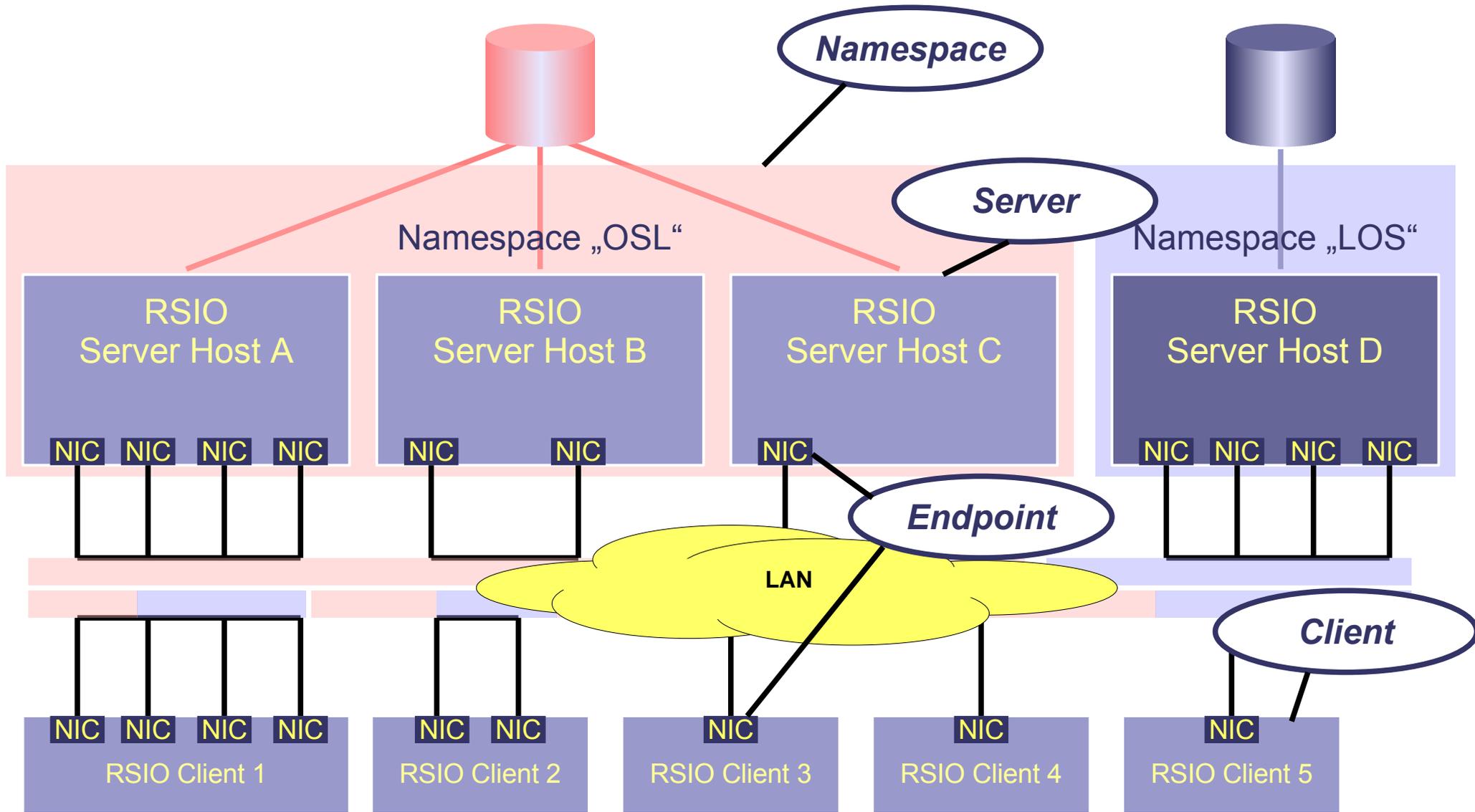
RSIO - Remote Storage I/O

Vergleich der Protokollstacks



RSIO – Architektur im Überblick

Klar gegliedertes und flexibles administratives Konzept



OSL Gesellschaft für offene Systemlösungen mbH

www.osl.eu

Wie OSL RSIO umgesetzt hat

Kombination mit dem OSL Storage Cluster



So meldet sich eine iSCSI-Lun ("format" - Solaris)

```
29. c3t227d0 <DEFAULT cyl 1021 alt 2 hd 64 sec 32>
    /iscsi/disk@0000iqn.1986-03.com.sun%3A02%3A06df3360-bb85-ee33-bf59f2d03474f708.target-00001,0
30. c3t229d0 <DEFAULT cyl 1021 alt 2 hd 64 sec 32>
    /iscsi/disk@0001iqn.1986-03.com.sun%3A02%3A06df3360-bb85-ee33-bf59f2d03474f708.target-00001,0
```

Und so sieht der RSIO-Client Plattenressourcen

```
# rsconfig -q
000 osl
    clt: big-6
    srv: 000 big-5
         0   tvoll          disk          2097152 blocks of 512 bytes
         0   shadow        disk          2097152 blocks of 512 bytes
         0   ora_db         disk          10485760 blocks of 512 bytes
         0   postgres_db    disk          10485760 blocks of 512 bytes
         0   whole_zone     disk          41943040 blocks of 512 bytes
```

Und was ist mit der Performance?

Protokoll erlaubt hohe Performance und beeindruckende Skalierbarkeit



Server-Performance bei Cache Read / 8k

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

Client-Performance Throughput

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

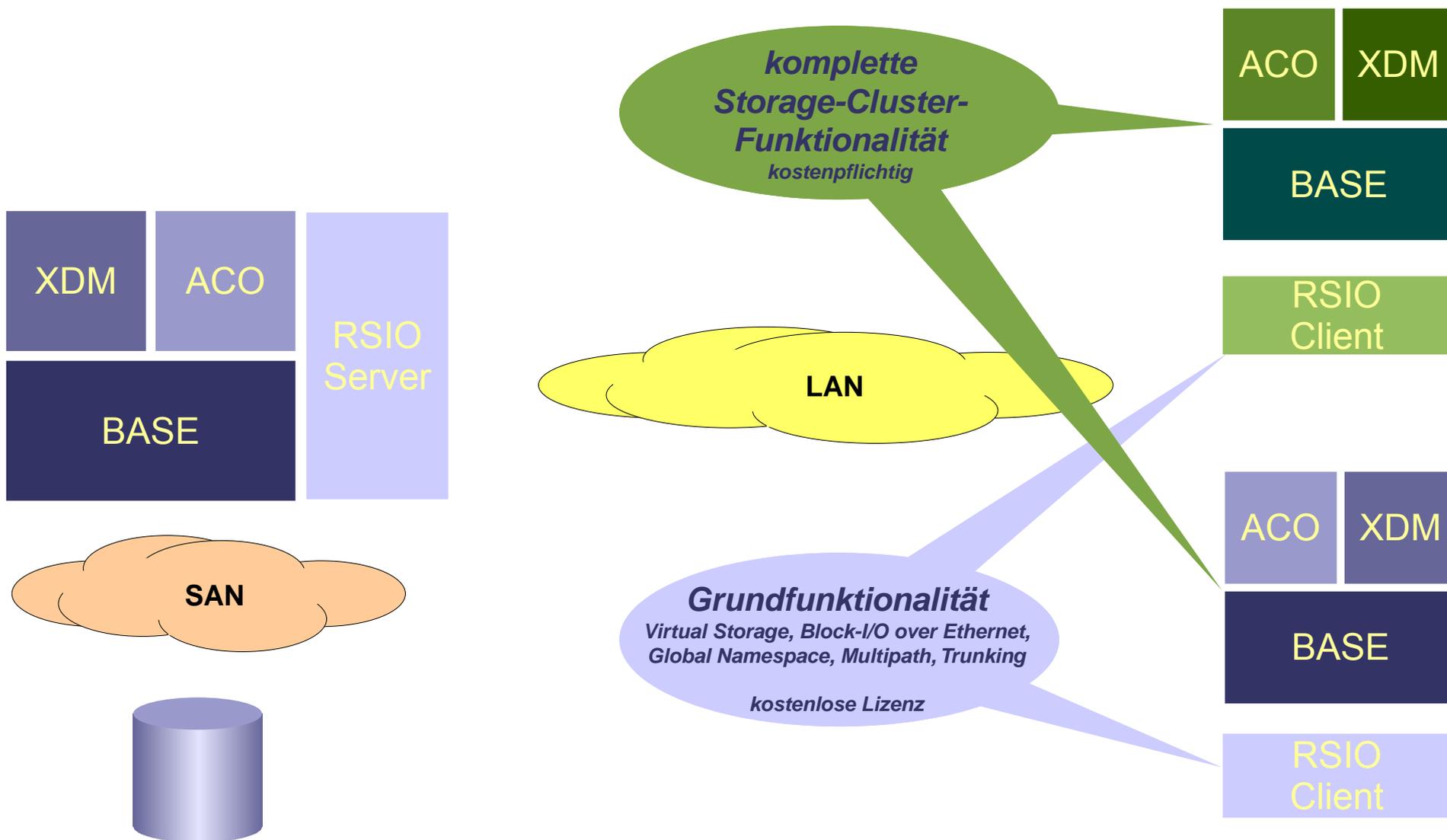
OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Einsatzszenarien

Von der Technologie zur Lösung

Was kann ich mit RSIO aufbauen?

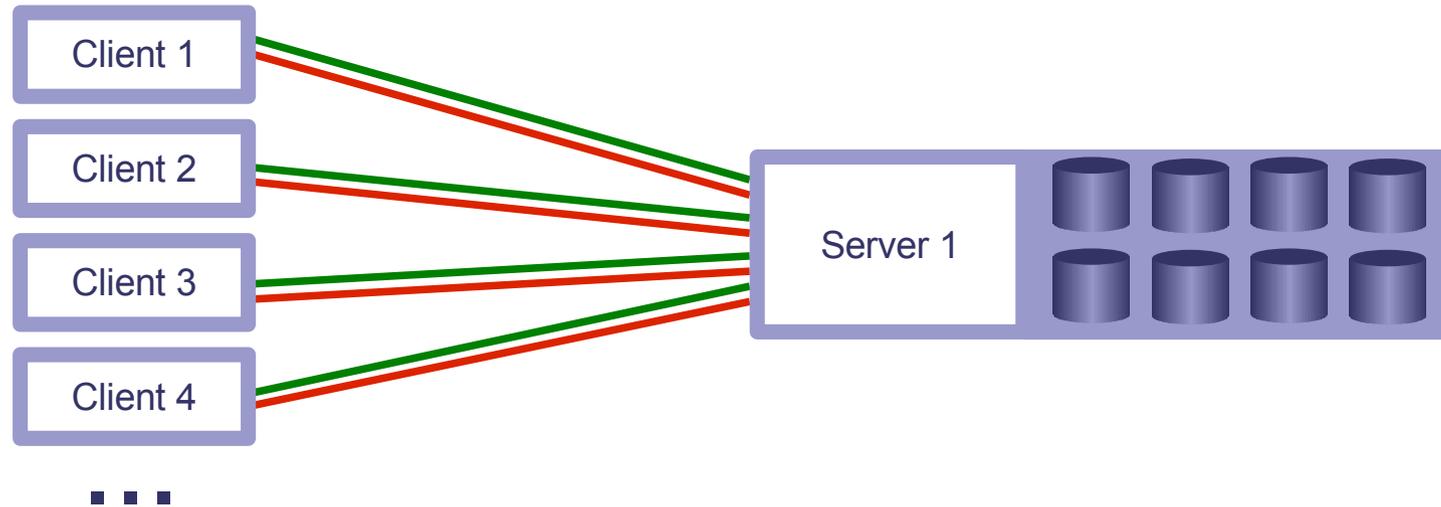
Das Prinzip



OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Was kann ich mit RSIO aufbauen?

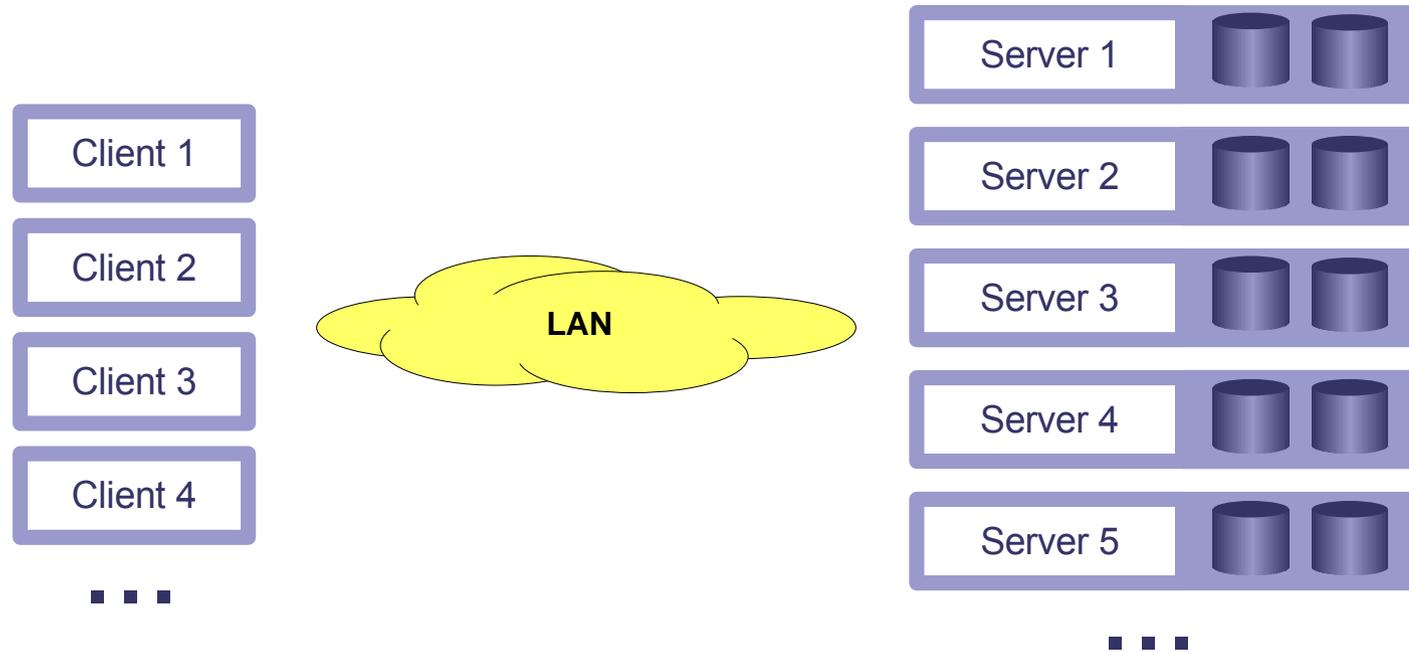
Beispiel 1: Einfacher Zugriff auf Plattenressourcen über LAN



- *Zugriff auf zentrales Speichersystem -> Global Pool, Global Namespace*
- *Virtualisierung und Cluster (HV) auf Clients einfach realisierbar*
- *Möglichkeit der Zentralisierung von Backup, Snapshots ...*
- *sehr preiswerte Speichieranbindung bei guter Performance*
- *redundante Datenpfade, Durchsatz je nach Bedarf skalierbar*

Was kann ich mit RSIO aufbauen?

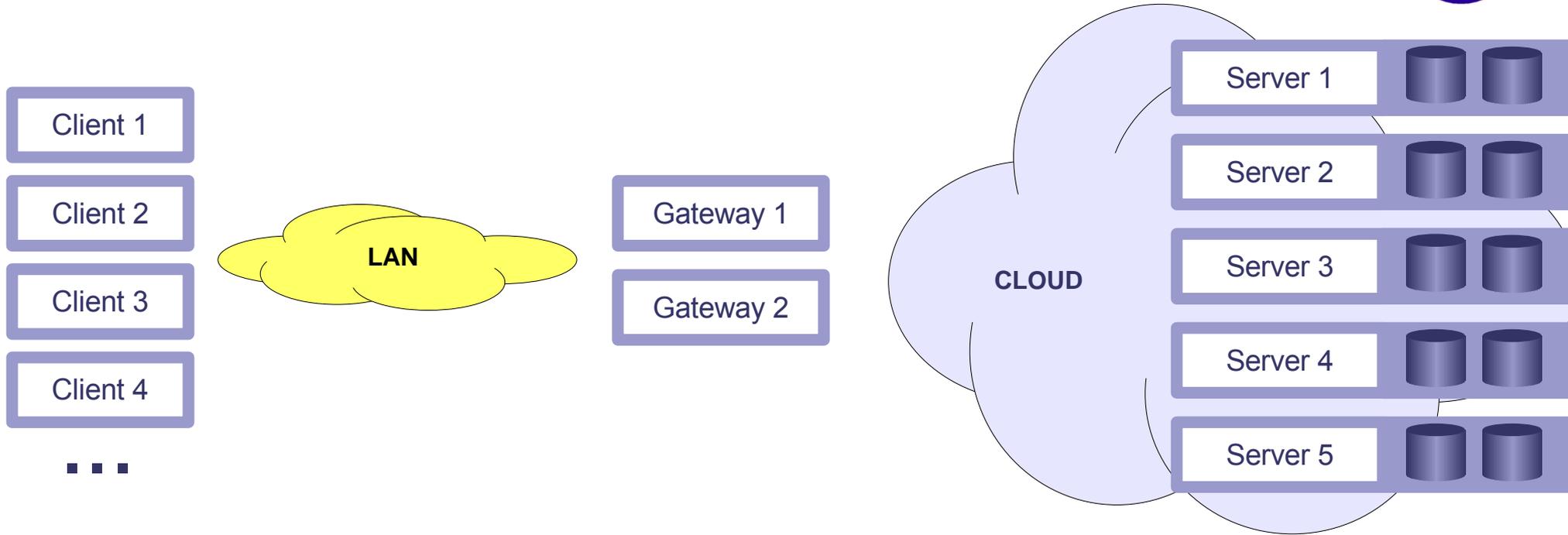
Beispiel 2: Storage Server Farm



- *Skalierung in Speichervolumen und Bandbreite*
- *jeder Server mit eigenem Namespace*
- *Storage-Kapazitäten “einsammeln” und so mit einfachen Mitteln große Kapazitäten und Bandbreiten darstellen*
- *nicht vergessen: Verfügbarkeit in der Server-Farm*

Was kann ich mit RSIO aufbauen?

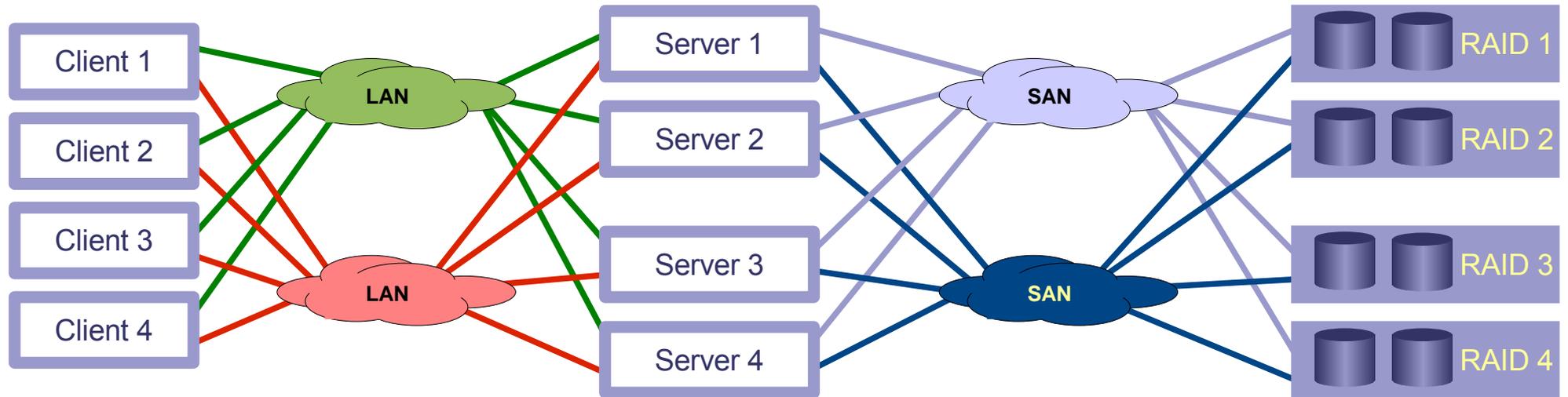
Beispiel 3: Szenario für Cloud Storage



- *Zugriff auf Speicherressourcen jenseits des LAN*
- *Mehrpfadigkeit, Bandbreite, Performance treten in den Hintergrund*
- *Gleichartige Administration wie bei RSIO im LAN*
- *Nutzt prinzipielle Routingfähigkeit von RSIO über IP*

Was kann ich mit RSIO aufbauen?

Beispiel 4: SAN-LAN-Konvergenz und geclusterte Storage-Server

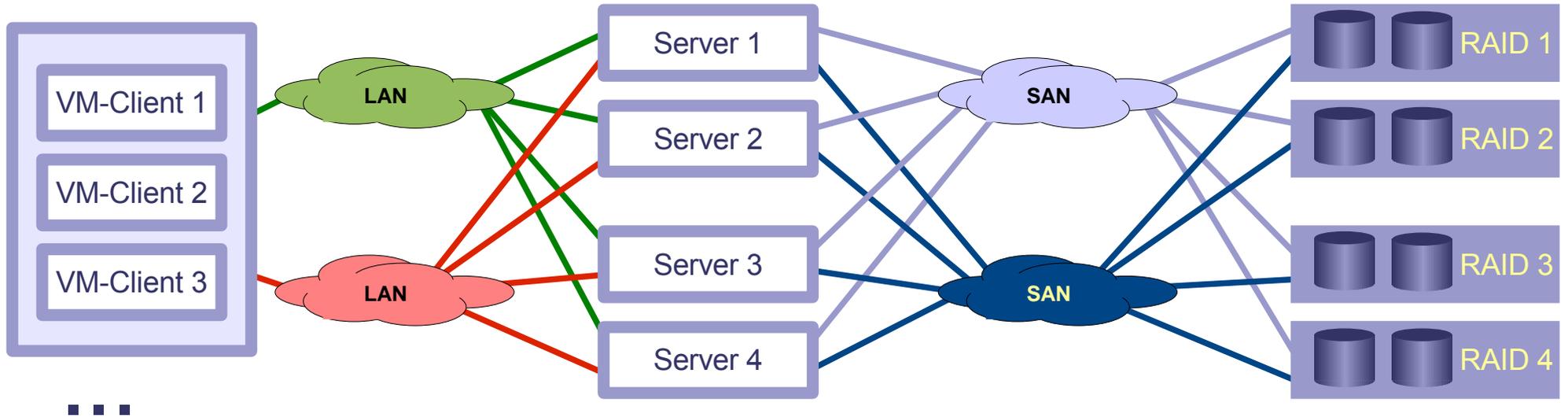


- *SAN ins LAN hinein verlängern*
- *SAN-attached Server reichen “im Hintergrund” Storage-Ressourcen durch*
- *verbesserte Ausnutzung des SANs, Performance-Rightsizing*
- *hohe Performance, hohe Verfügbarkeit bei extrem niedrigen Kosten für RSIO-Clients*
- *weitere Verbesserung von Performance und Systemauslastung möglich z. B. durch Nutzung freien Speichers als Cache*

OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Was kann ich mit RSIO aufbauen?

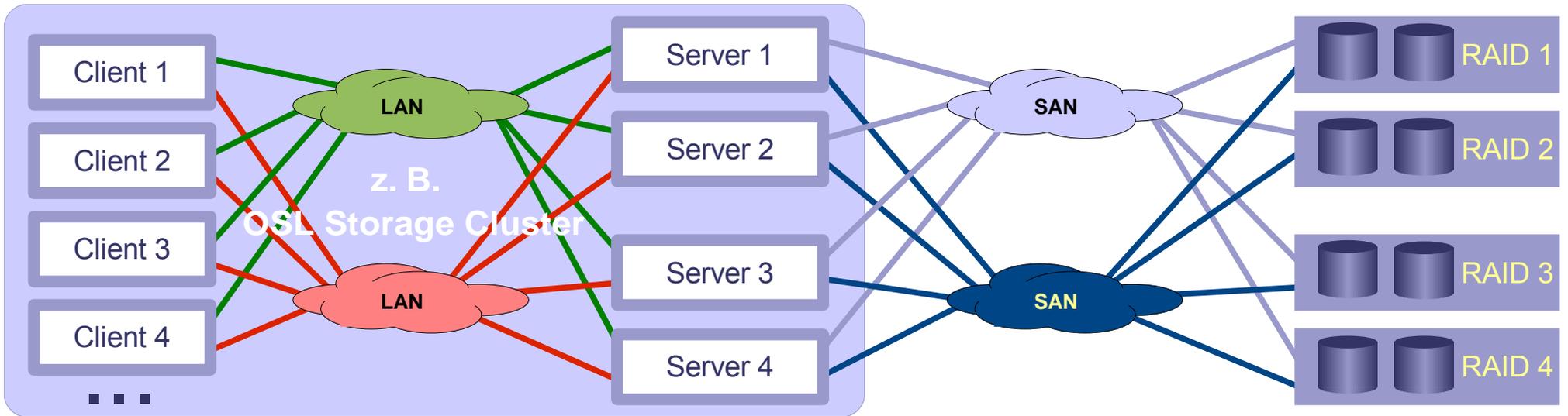
Beispiel 5: Einfacher Zugriff auf die gesamte Storage-Welt aus der VM



- *SAN reicht über die IP-Interfaces bis in die Virtuellen Maschinen hinein*
- *beliebige Devices erreichbar, daneben Selbstkonfiguration ...*
- *Aggregation für 10GbE, Nutzung von VMDq möglich*
- *enorme Vereinfachung*

Was kann ich mit RSIO aufbauen?

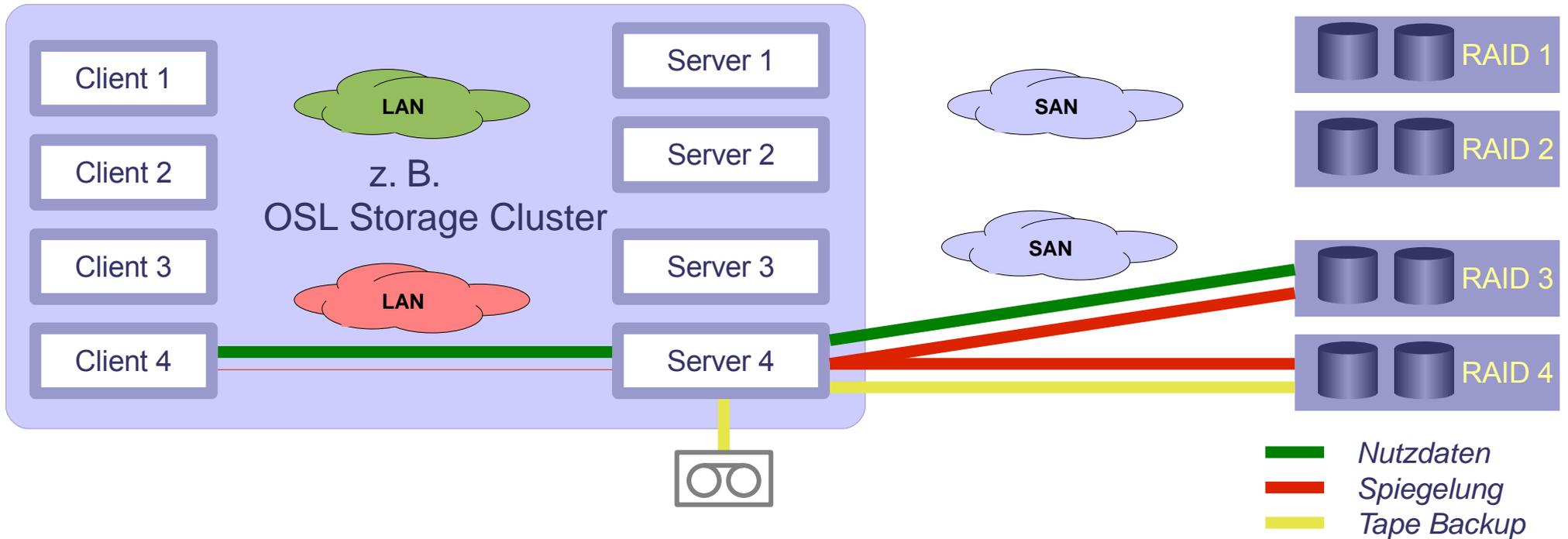
Beispiel 6: Server und Clients in einem Cluster integrieren



- alle Funktionen wie in Beispiel 3 (SAN-LAN-Integration)
- zusätzlich weitere Storage-Management-Funktionen:
 - Storage-Allokation, -Management vom Client aus
 - applikationsbezogene Speichervirtualisierung vollumfänglich auf Client nutzbar
 - Möglichkeit der transparenten Nutzung von Datenspiegelung, Backup to Disk etc.
- Verschmelzung von Client und Server zu einer Einheit
- run applications everywhere

Was kann ich mit RSIO aufbauen?

Beispiel 7: Hochgeschwindigkeits-Backup für LAN-attached Blockdevices



- über das LAN laufen nur Nutzdaten und die Steueranweisungen
- LAN-less Backup:
 - hohe Geschwindigkeit
 - vollständige Steuerung vom Client aus
 - applikationsbezogene Aktionen

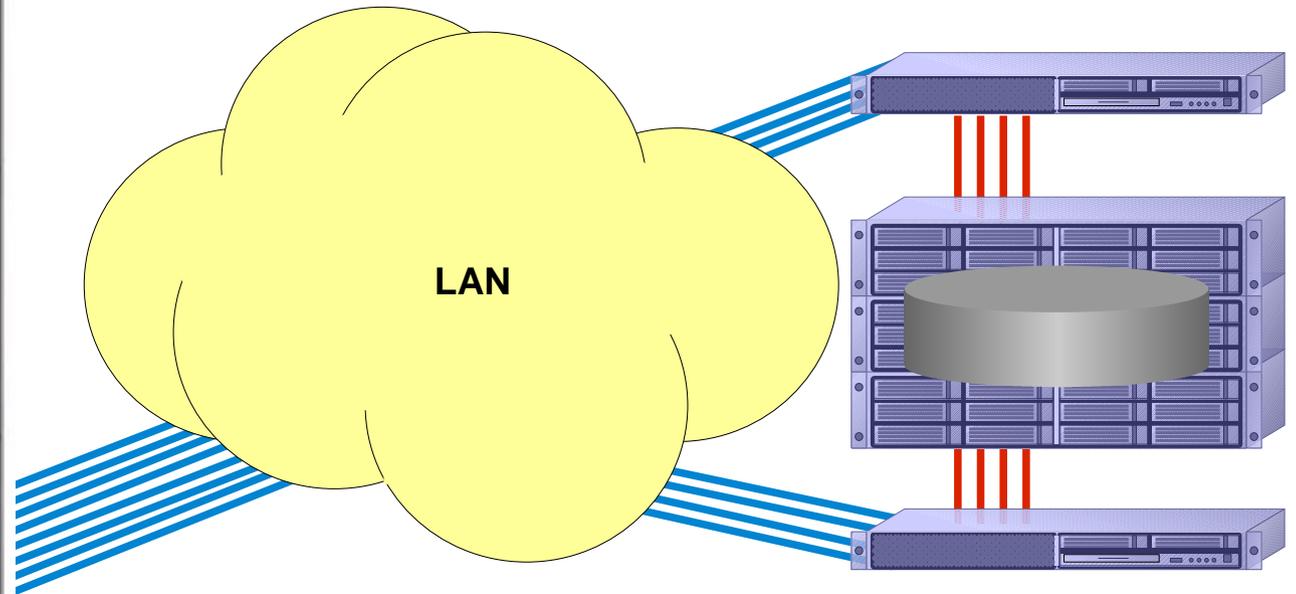
OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Was kann ich mit RSIO aufbauen?

Beispiel 8: 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

Einfach nutzen!

Vorkonfigurierte Lösungspakete

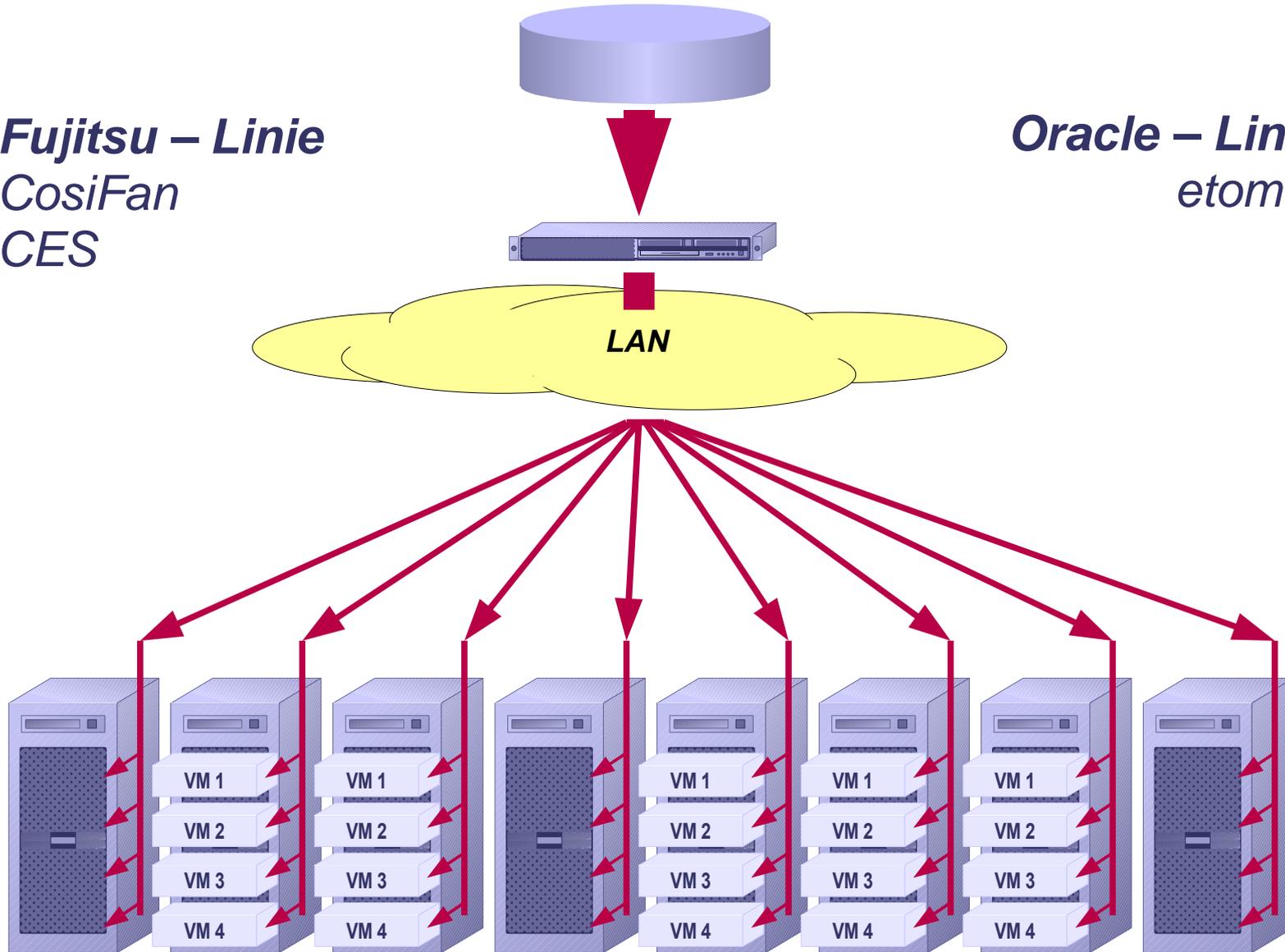
Kann ich eine fertige Lösung mit RSIO kaufen?

Vorzugskonfigurationen in Zusammenarbeit mit Partnern – das Prinzip



Fujitsu – Linie
CosiFan
CES

Oracle – Linie
etomer

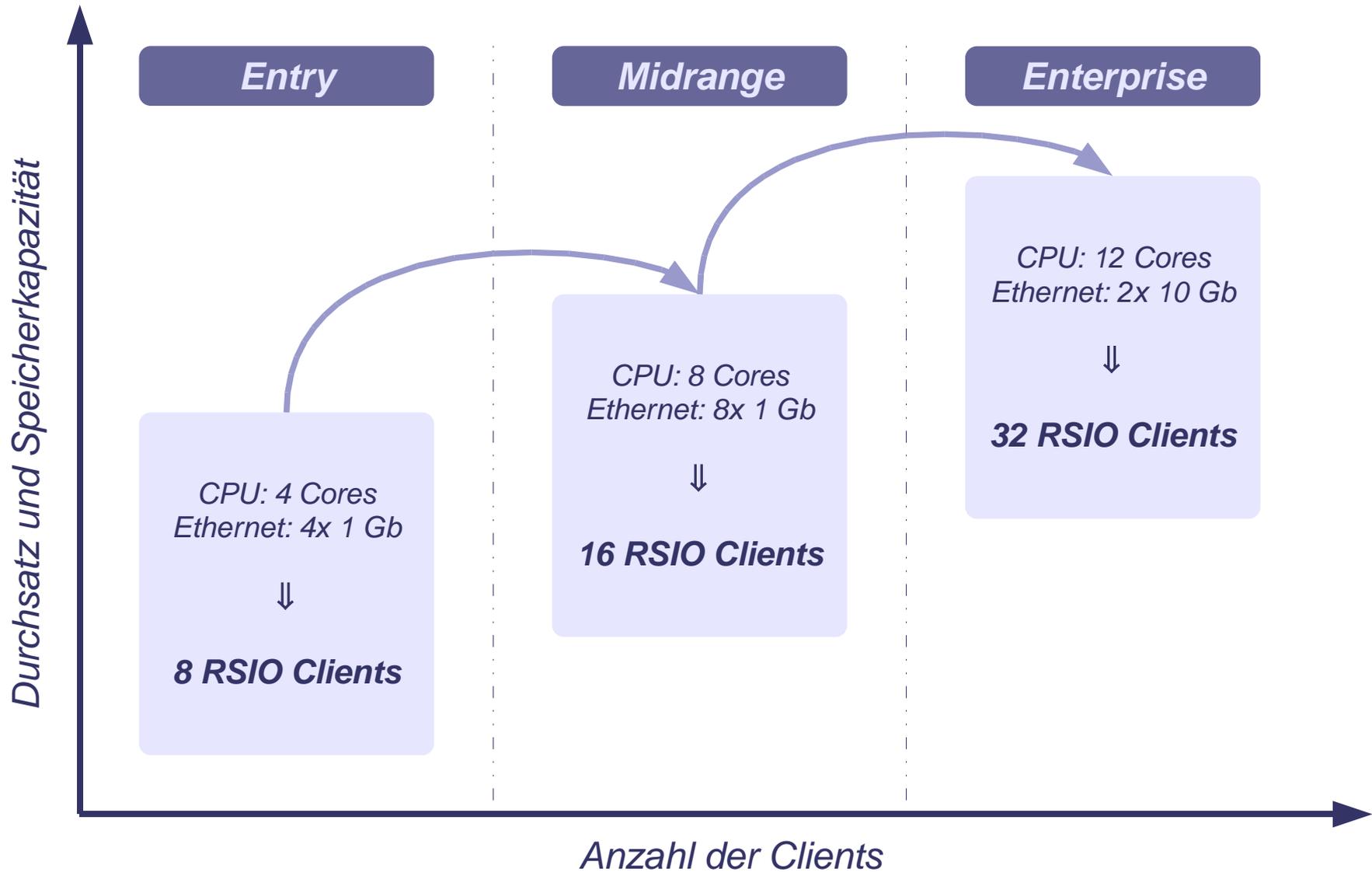


OSL Gesellschaft für offene Systemlösungen mbH

www.osl.eu

RSIO - Vorzugskonfigurationen

Drei jeweils erweiterbare Ausbaustufen



OSL Gesellschaft für offene Systemlösungen mbH

www.osl.eu

RSIO – Vorzugskonfiguration “Entry Level”

Preiswerter Einstieg – geeignet bis 8 Clients



Fujitsu

Eternus DX80 (Single-Controller)



8 Gb Fibre Channel



Primergy RX100 S6

CPU: 1x Xeon X3430 (4 Cores)
LAN: 1 Karte 4x 1 Gb Ethernet

1 Gb Ethernet für RSIO
(8 Clients)

Oracle

ES-8200D (Dual-Controller)



8 Gb Fibre Channel



Sun Fire X4170 M2

CPU: 1x Xeon E5620 (4 Cores)
LAN: 4x 1 Gb Ethernet (intern)

1 Gb Ethernet für RSIO
(8 Clients)

OSL Gesellschaft für offene Systemlösungen mbH

www.osl.eu

RSIO – Vorzugskonfiguration “Midrange”

Stärker ausgebaut – geeignet bis 16 Clients



Fujitsu

Eternus DX80 (Dual-Controller)



8 Gb Fibre Channel



Primergy RX300 S6

CPU: 2x Xeon E5620 (2x 4 Cores)

LAN: 2 Karten 4x 1 Gb Ethernet

**1 Gb Ethernet für RSIO
(16 Clients)**

Oracle

ES-8200D (Dual-Controller)



8 Gb Fibre Channel



Sun Fire X4170 M2

CPU: 2x Xeon E5620 (2x 4 Cores)

LAN: 1 Karte 4x 1 Gb Ethernet

4x 1 Gb Ethernet (intern)

**1 Gb Ethernet für RSIO
(16 Clients)**

OSL Gesellschaft für offene Systemlösungen mbH

www.osl.eu

RSIO – Vorzugskonfiguration “Enterprise”

Hohe Performance und Kapazität – geeignet bis 32 Clients



Fujitsu

Eternus DX80 (Dual-Controller)



8 Gb Fibre Channel



Primergy RX300 S6

CPU: 2x Xeon E5645 (2x 6 Cores)

LAN: 1 Karte 2x 10 Gb Ethernet

**10 Gb Ethernet für RSIO
(32 Clients)**

Oracle

ES-6600 (Dual-Controller)



8 Gb Fibre Channel



Sun Fire X4270 M2

CPU: 2x Xeon X5675 (2x 6 Cores)

LAN: 1 Karte 2x 10 Gb Ethernet

**10 Gb Ethernet für RSIO
(32 Clients)**

OSL Gesellschaft für offene Systemlösungen mbH

www.osl.eu

RSIO – Vorzugskonfiguration: Die Kosten

Vergleich Infrastrukturkosten* SAN (4GB FC) mit RSIO / Ethernet (ohne RAID-System)



	SAN	RSIO
<i>Entry (8 Clients)</i>	<i>ca. 12.700</i>	<i>ca. 5.000</i>
<i>Midrange (16 Clients)</i>	<i>ca. 29.300</i>	<i>ca. 8.500</i>
<i>Enterprise (32 Clients)</i>	<i>ca. 70.000</i>	<i>ca. 18.800</i>

* Stand April 2011, auf Basis von Endkundenpreisen eines Webshops, RSIO ohne Softwarelizenzen

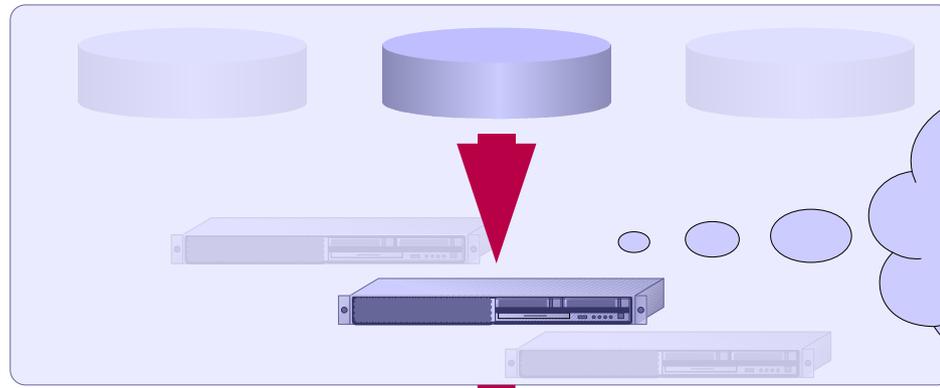
OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Vorzugskonfigurationen

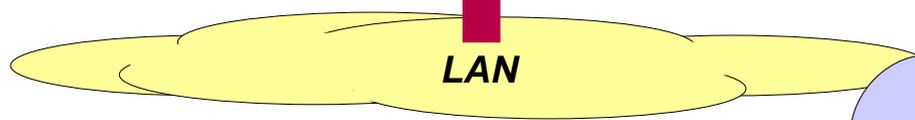
Nochmal zum Prinzip



OSL Storage Cluster
client side automated
storage management,
cluster framework
high availability
etc.



RSIO Server
Virtual Storage
Clone, Mirror, DR,
Bandwidth Control,
Backup ...

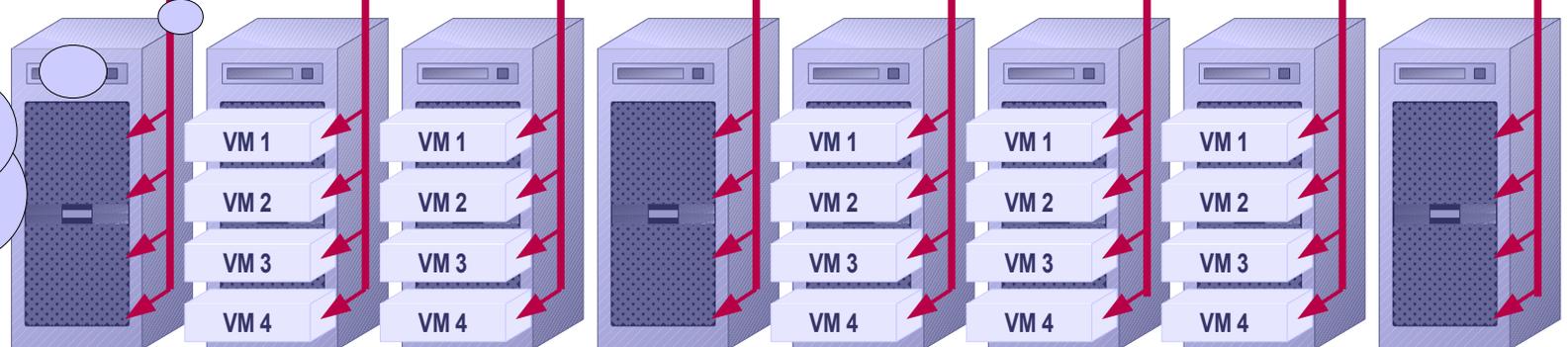


ACO XDM

RSIO - Client
Use Virtual Storage,
scalable,
cluster enabled,
global namespace,
multipathing ...



Virtual Systems
complete
infrastructures
Windows
integration ...



OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

Wo stehen wir?

Parallel: Entwicklung, Pilotierung, Aufbau Linux-Support, Vorbereitung Freigabe



- *RSIO-Pilotierung ab sofort (Solaris / Linux)*
- *Demo-Tour: Mai/Juni*
- *Allg. Freigabe zusammen mit Storage Cluster 4.0 vorauss. Okt. 2011*
- *Technologietage 2011: 14./15. 9. 2011*
- *Anwenderaktiv / Freigabeseminar: vorauss. 29./30. 9. 2011*

sprechen Sie uns an:

+49 (0) 30 / 74 07 67 80

info@osl-it.de

OSL Gesellschaft für offene Systemlösungen mbH
www.osl.eu

OSL SC 4.0 – weitere Themen

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

- *heute Spitzenposition bei Speichervirtualisierung & Clustering*
- *klare Roadmap*
- *Öffnung für weitere Systeme, Integration VMs*
- *enorme Vorteile mit RSIO für Linux und Solaris:*
 - *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*