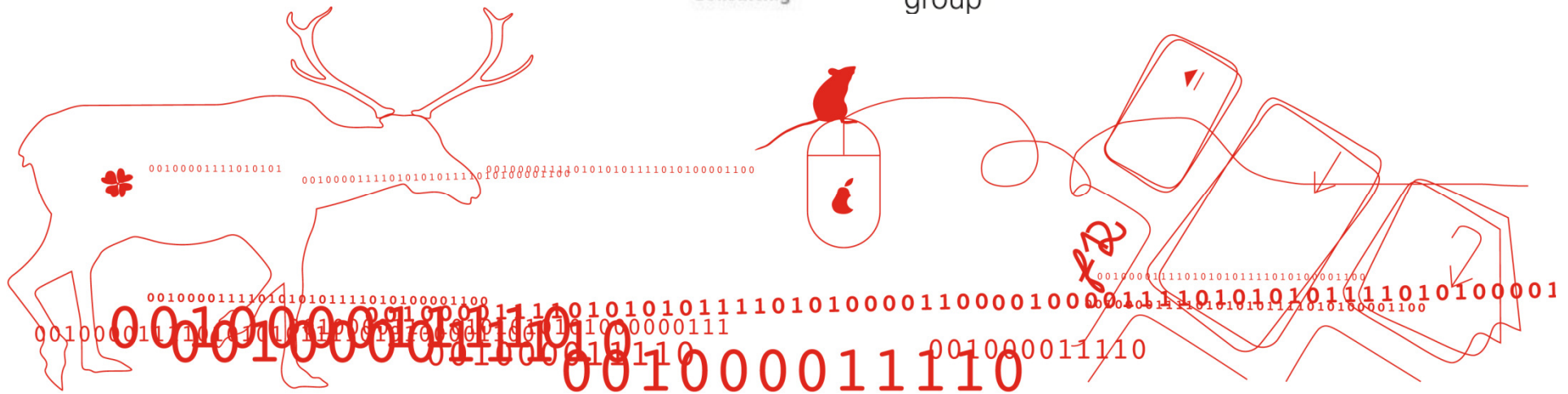
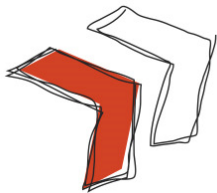


Vortrag:

Storagevirtualisierung und Hochverfügbarkeit

Referent:

Jörg Utesch, Projektvertrieb systemzwo GmbH



• Warum Storagevirtualisierung?	Seite 3- 4
• Active Working Set	Seite 5
• IBM SAN Volume Controller (SVC)	Seite 6 - 7
• Transparent Failover mit SVC	Seite 8 – 11
• Datacore Sansymphony - Funktionen	Seite 12 – 15
• Datacore Sansymphony – Virtuelle Diskmigration	Seite 16
• Datacore Sansymphony – Spiegel und Failover	Seite 17 – 24
• Live Demo	Seite 25

Ausgangssituation:

Ein Kunde setzt ein oder mehrere Storagesysteme ein. Klassischerweise werden hier RAID's aufgebaut und die LUNs physischen Servern bzw. Vmware zur Verfügung gestellt.

Alles prima, aber:

Die Performance und unterschiedlichen Workloads fordern die SAN Systeme.
Bsp: 40 VMs schreiben am Ende des Tages auf bspw. 2 oder 3 physische RAID's.
Die Folge: Performanceengpässe (READ, WRITE, IOs)

Was tun: Mehr „Spindeln“, größerer Cache oder:

Storagevirtualisierung

Ein zentraler „Layer“ –
vergleichbar mit den Hypervisoren der Servervirtualisierung

Was ist Storagevirtualisierung

Zentraler „Layer“ – vergleichbar mit den Hypervisoren der Servervirtualisierung

Beispiel: IBM SAN Volume Controller oder Datacore Sansymphony

-> Dieser Layer verwaltet alle Stagesysteme und stellt sie zentral zur Verfügung

Neue Funktionalitäten werden möglich wie:

- zentrales Caching
- Einsatz von neuen Storage Tiers (Flash, SSD, NL-SAS)
- Möglichkeit eines transparenten Failovers
- Autotiering

Active Working Set, was ist das?

Welche Anforderungen haben meine Applikationen an das Storage im täglichen Betrieb tatsächlich?

Beispiel:

In meinen Datenbanken ändern sich nur x GB pro Tag, das gleiche gilt für Fileserver und Applikationsserver. Muss ich bspw. die gesamte Datenbankgröße auf performantem Storage vorhalten?

Antwort: Nein

Warum?

Dank modernen Tieringkonzepten werden nur die „heißen“ Daten in einem performanten SSD/Flash oder SAS Tier vorgehalten. Da die Systeme die Zugriffe auf Blockgröße monitoren kann eine Datenbank bspw. zu 10% auf Tier1, 30% auf Tier 2 und 60% auf Tier 3-Storage verteilt sein.

Daher: ein vergleichsweise kleiner Tier1 reicht aus um die Workload zu tragen, zukünftige Invest können in günstigeren Tier2/3 (NL-SAS Storage) getätigt werden

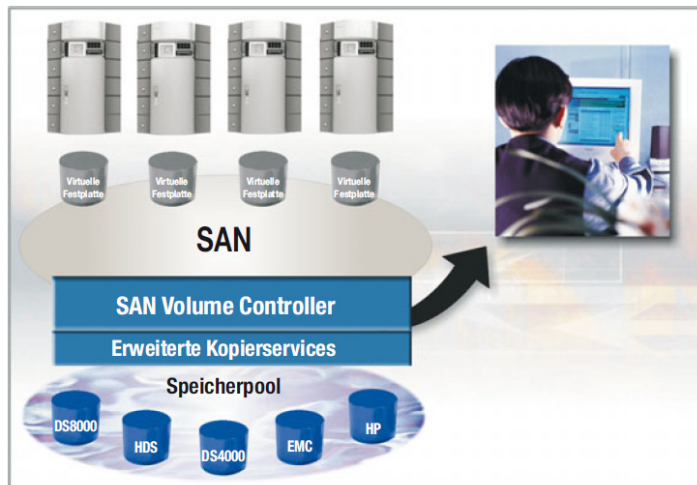


IBM SVC Stretched Cluster

- 24 GB Cache im Controller
- 4x 8Gbit FC Hostports
- Unterstützung für alle namhaften Storagehersteller in einem gemischten Clusterverbund

Funktionen:

- Easytiering (automatisches Storagetiering)
- Intuitive grafische Bedienoberfläche
- Thin Provisioning
- Snapshots (TB lizenziert)
- Synchrone Spiegelung zwischen zwei Storage-Systemen
- Unterstützung von FC, FCoE, 10 Gbit iSCSI (optional)
- Einschränkung:
 - Bei einem Ausfall eines RZ wird ein Quorum Storage benötigt (muß ein IBM Storage System sein)



IBM SAN Volume Controller:

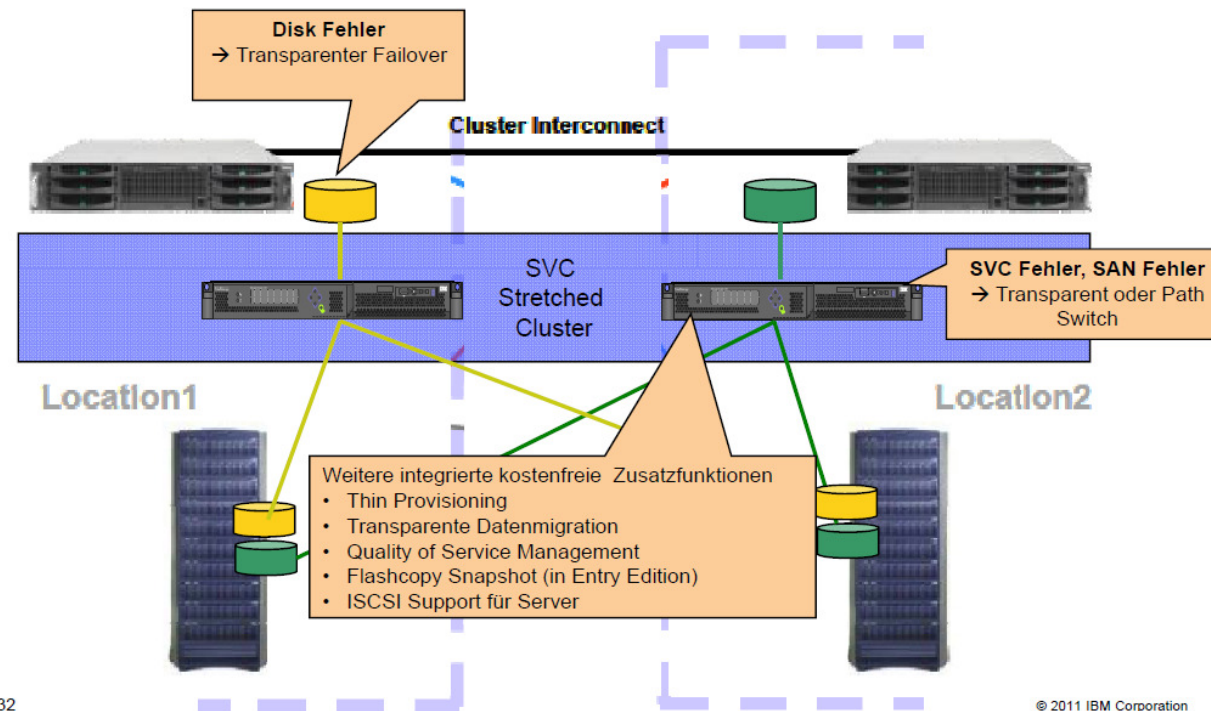
- Ermöglicht die Zusammenfassung von Speichersystemen zu einem zentralen Speicherpool, der sich besser verwalten lässt
- Ausgelegt für bessere Speicherausnutzung – Hostanwendungen erhalten flexibleren Zugang zur benötigten Kapazität
- Steigert die Produktivität der Administratoren durch automatische Bereitstellung und einfache einheitliche Verwaltung zusammengefasster Speichersysteme
- Erhöht die Anwendungsverfügbarkeit, da er Anwendungsausfallzeiten, die durch den Speicher verursacht wurden, fast vollständig beseitigt
- Ermöglicht eine abgestufte Speicherumgebung, in der die Speicherkosten besser an den Wert der Daten angepasst werden können
- Unterstützt erweiterte Kopierdienste von teureren auf billigere Geräte und zwischen Subsystemen verschiedener Hersteller

Einfache Implementierung

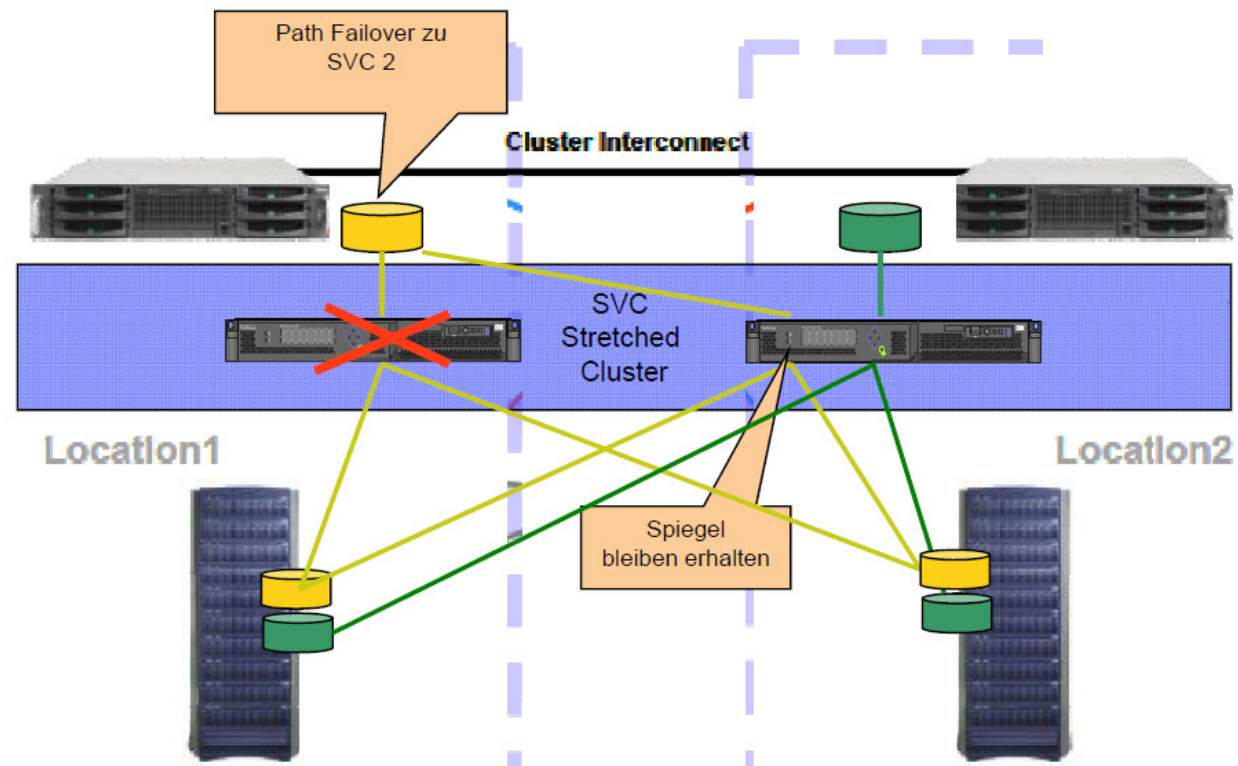
Die SVC Storage Engines werden mit vorinstallierter SVC Software geliefert. So kann diese nach dem Anschluss der Engines an Ihr SAN (Storage Area Network) rasch implementiert werden. SVC Storage Engines basieren auf bewährter IBM System x Servertechnologie und werden stets als redundante Paare implementiert, die eine sehr hohe Verfügbarkeit bereitstellen.

SVC übernimmt die Kontrolle über Ihren bestehenden Speicher und sichert gleichzeitig alle Ihre bestehenden Informationen. Dadurch ermöglicht er eine schnellere und einfachere Implementierung und senkt den Bedarf an zusätzlichem Speicher. Nachdem SVC implementiert ist, können Sie die Konfiguration je nach Bedarf rasch und einfach ändern.

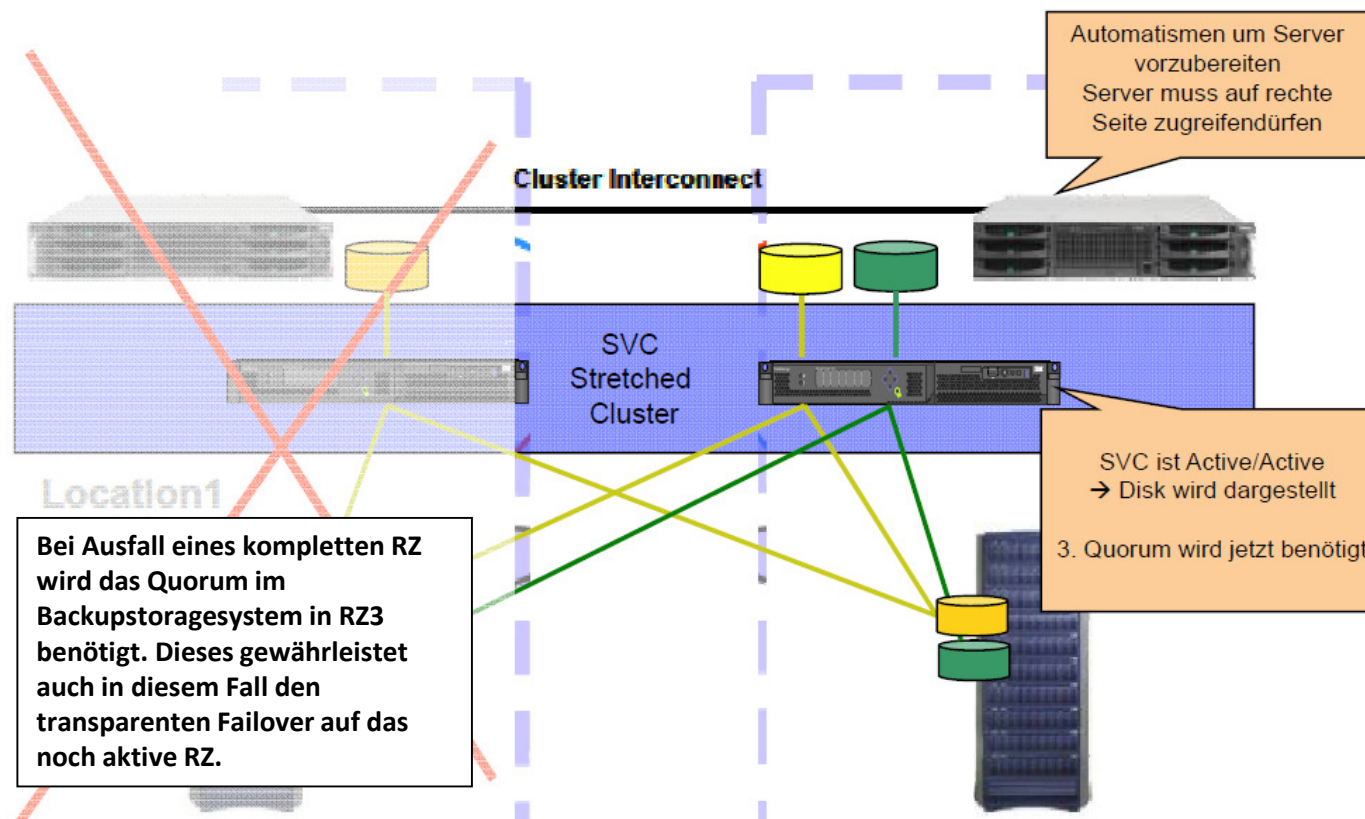
Hochverfügbare SVC Site
Disk Mirror, transparenter Failover

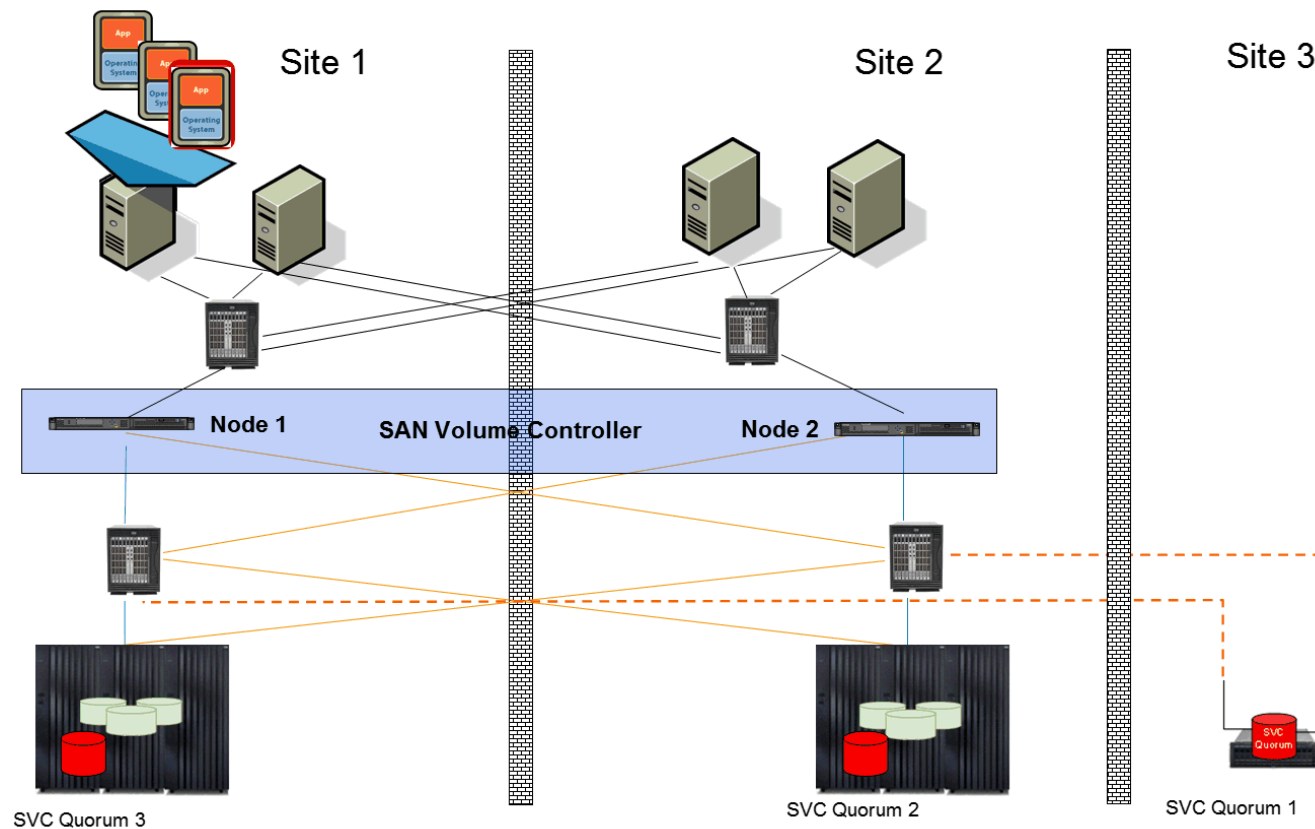


Hochverfügbare SVC Site
Situation Ausfall SVC Knoten oder SAN Komponenten



Hochverfügbare SVC Site





Best Practice:

Ein Quorum Storage ist in einem dritten Brandabschnitt (kleine LUN auf IBM Storage)

Es ist aber auch möglich ein oder mehrere Quorums auf einer LUN auf Node1 bzw. 2 zu erstellen – > höherer Aufwand für Konfiguration und im Fehlerfall besteht die Gefahr das auch auf das Quorum nicht zugegriffen werden kann



Storagevirtualisierung mit Datacore San Symphony

- Herstellerunabhängig
- Serverram = Cache (bis 1TB)
- Virtual Disk Pooling
- Caching
- Thin Provisioning
- Synchronous Mirroring
- Centrally manage nodes in same group
- iSCSI
- **Snapshot**
- Storage Migration & Pass-through Disks
- Cloud Storage Gateway 1TB Virtual Appliance
- **vCenter Plug-in/ System Center Pack**
- Asynchronous Replication CDP 14 Tage
- **Automated Storage Tiering**
- Advanced Site Recovery
- Fibre Channel Protocol





DS3524

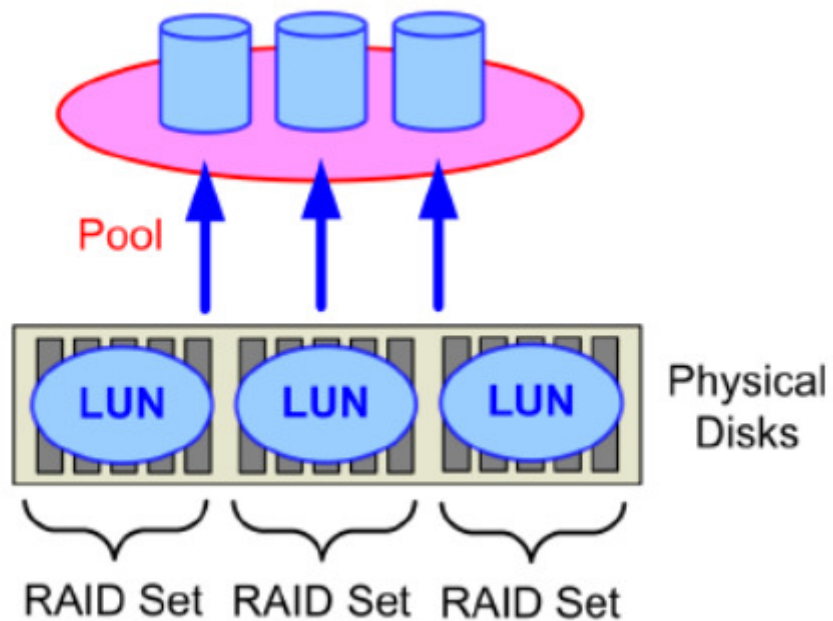


RAID Sets in der DS5324 am Beispiel einer Konfiguration mit 900 GB 10K SAS Harddisks:

- RAID 5-Gruppen 4 + 1 bzw. 2+1 mit 900GB Kapazität je Shelf in Summe ca. 16 TB Netto

Verwaltung der LUNs über Datacore:

- Es werden mehrere kleine Raid-Sets erstellt um mehr I/O für den Pool zur Verfügung stellen zu können
- Datacore erstellt aus den hochgereichten LUNs einen virtuellen Pool mit 16 TB Kapazität die dann den jeweiligen Hosts zur Verfügung gestellt werden können
- Nach best practice von Datacore wird somit durch das striping die Schreiblast auf alle LUNs verteilt und die Notwendigkeit RAID 10 Gruppen im Storagesystem zu erstellen entfällt



Vorteile:

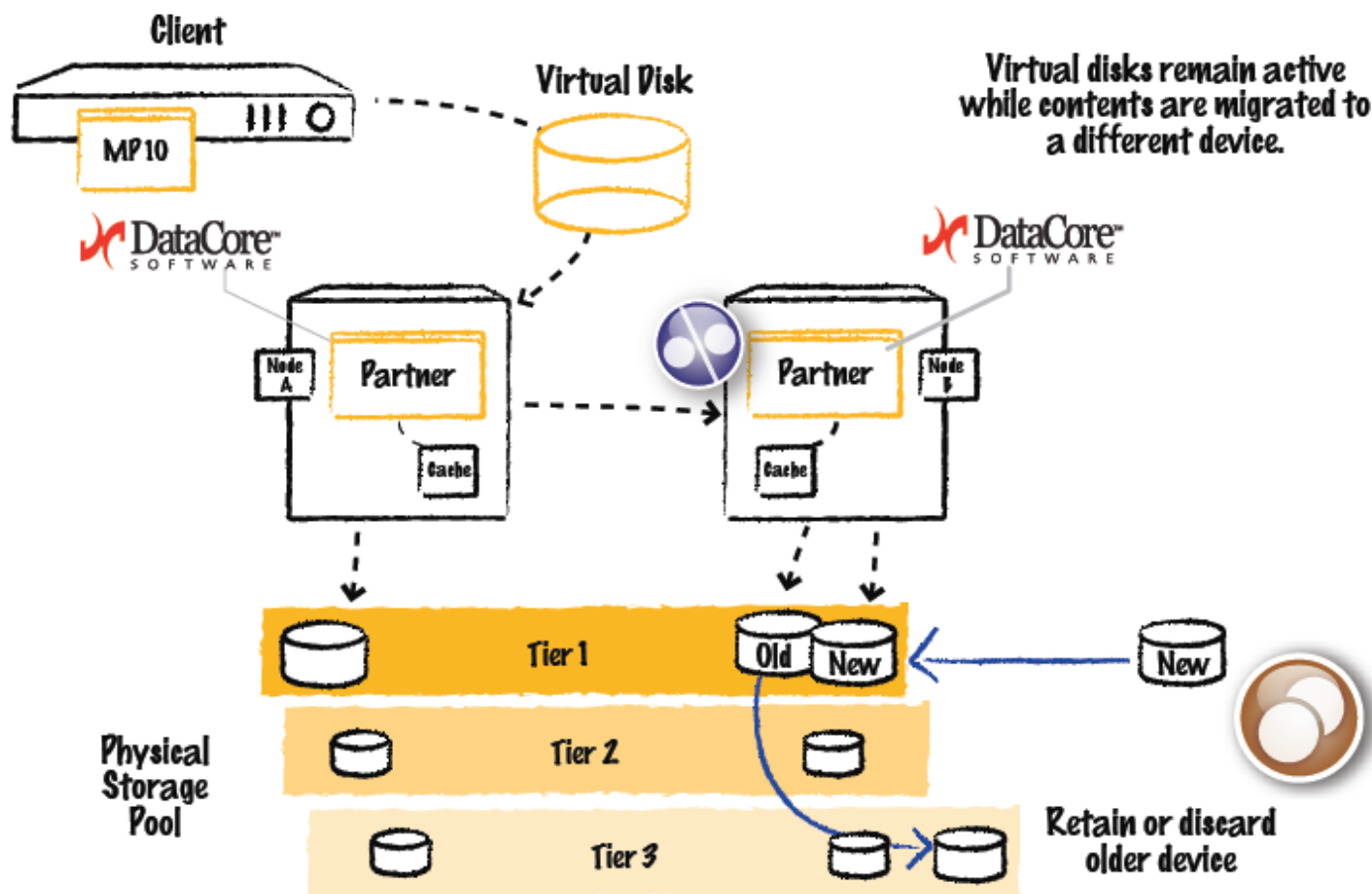
- Datacore hat 3 oder mehr I/O Queues zu den Platten
- Der Verteilungsalgorithmus verteilt die Last über die LUNs und steigert die Performance
- Eine defekte Platte betrifft nur eine LUN
- RAID rebuilds gehen schneller durch die kleineren RAID-Gruppen

Nachteile:

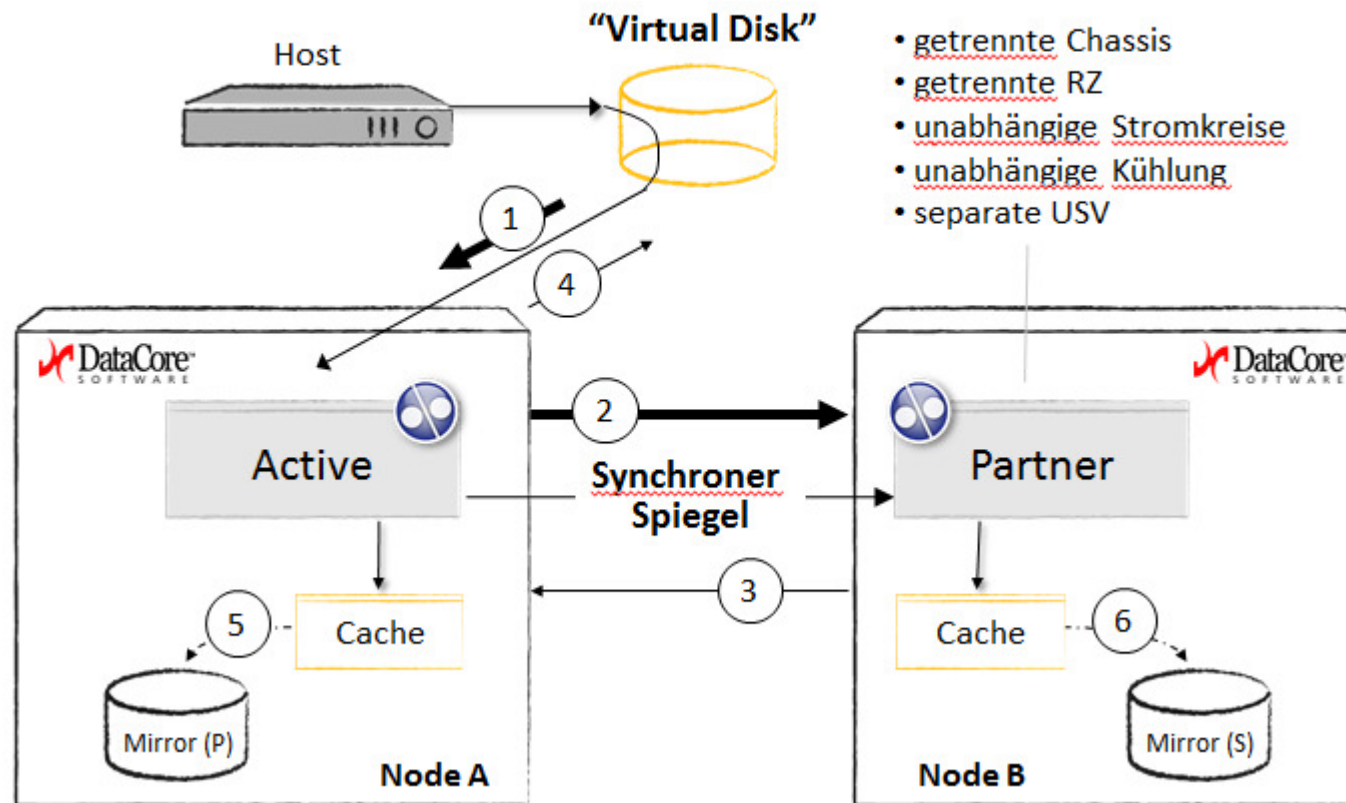
- Größerer RAID-Verschnitt durch den höheren RAID-overhead der kleineren Gruppen

Die Vorteile überwiegen aber deutlich den Nachteil des RAID-Verschnitts. Bezogen auf Performance und Verfügbarkeit die präferierte Konfiguration.

Migrationskonzept mit Virtual Disks Unterbrechungsfreie Migration (schematisch)



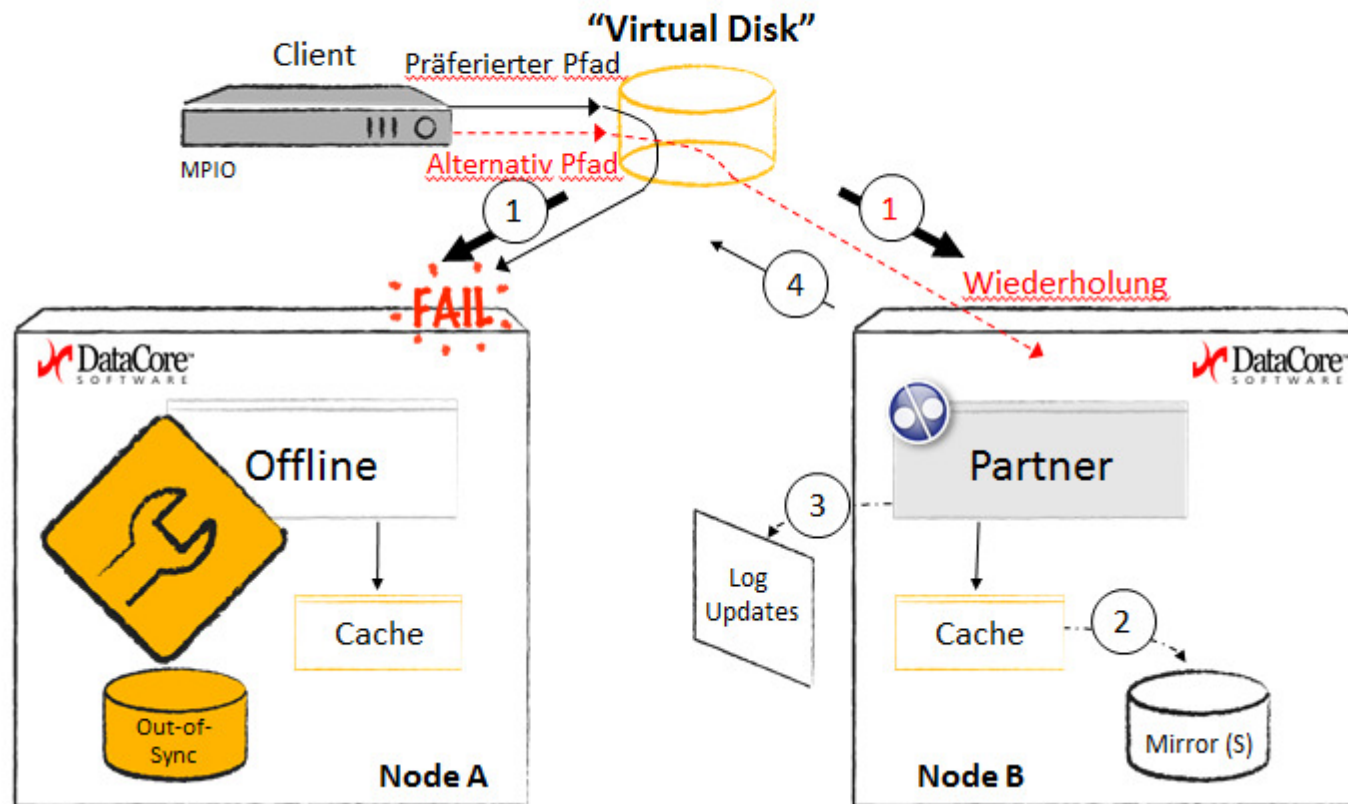
Synchrone Spiegelung mit Datacore



Siehe Zeichnung vorhergehende Seite:

1. Im normalen Betrieb schreiben die Hosts über ihren präferierten Pfad auf die Virtuellen Disks von Datacore.
2. Der Schreibvorgang wird unmittelbar an den zweiten Datacoreserver übermittelt.
3. Das „Write Aknowledge“ wird zurückgegeben wenn der Schreibvorgang auf beiden Datacoreservern im Cache geschrieben ist.
4. Der Host erhält die Information I/O Vorgang abgeschlossen
5. Der Cache wird auf Stagesystem A weggeschrieben
6. Der Cache wird auf Stagesystem B weggeschrieben

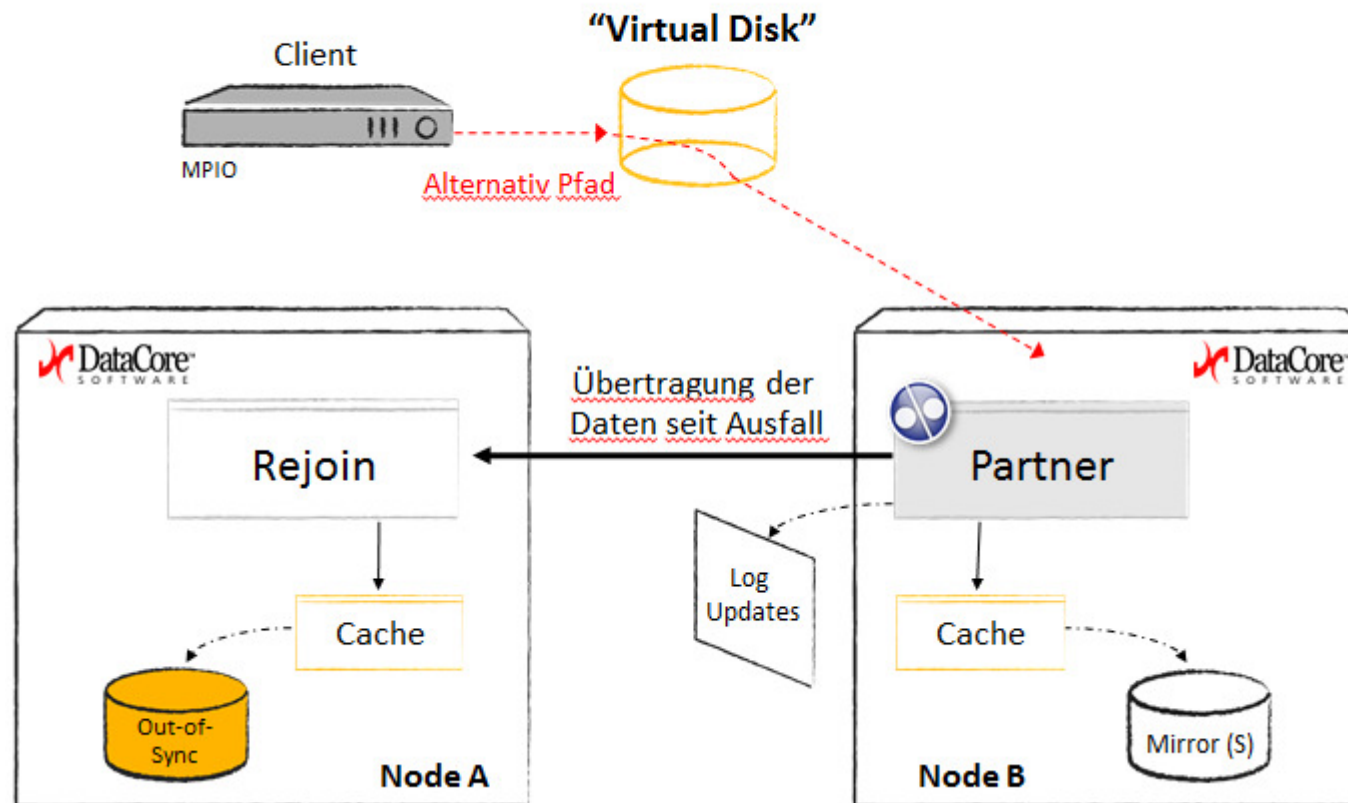
Automatischer Failover via Multipath I/O



Siehe Zeichnung vorhergehende Seite:

1. Sollte der Schreibvorgang über den präferierten Pfad nicht funktionieren und der Host einen Fehler erhalten wird er über den Multipathing Treiber einen Schreibvorgang wiederholen auf dem **alternativen Pfad 1**.
2. Der Write Cache wird auf dem Datacoreserver deaktiviert und der Schreibvorgang direkt auf die Platten weggeschrieben.
3. Parallel dazu werden alle Schreibvorgänge in ein Log geschrieben
4. Der Host erhält sein I/O Erfolg zurückgemeldet

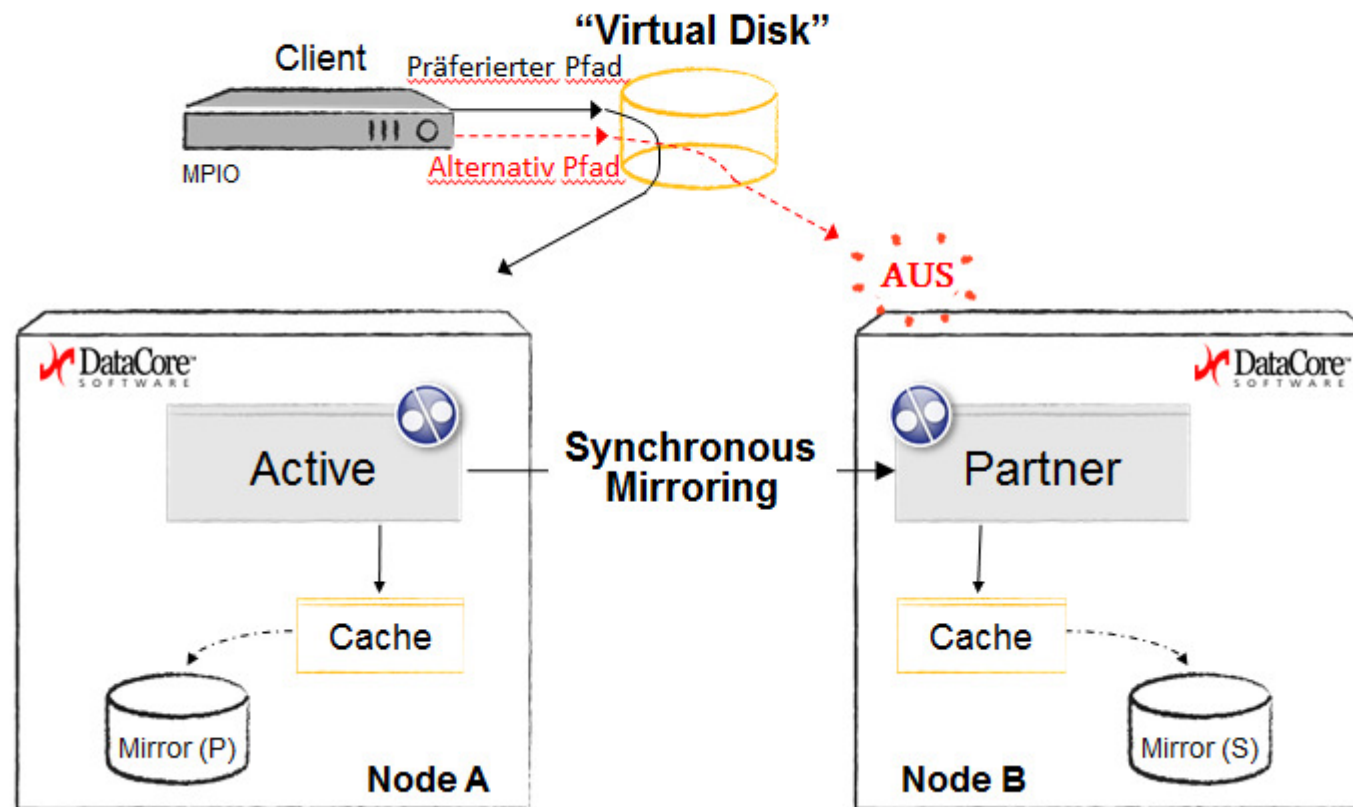
Resynchronisierung der gespiegelten Disks



Siehe Zeichnung vorhergehende Seite:

1. Sobald der Zustand der zum Pfadschwenk geführt hat behoben ist, realisiert Datacore das eine Aufnahme des Spiegels wieder möglich ist. Hierzu transferiert Datacore Node B die Daten seit dem Ausfall an das Stagesystem A.
2. **Während diesem Vorgang loggt Node B weiterhin alle Schreibvorgänge bis die Übertragung abgeschlossen und die beiden Datenbestände wieder synchronisiert sind.**

Herstellung Normalbetrieb



Siehe Zeichnung vorhergehende Seite:

1. Sobald der Vorgang der Resynchronisierung abgeschlossen ist, wird Node A wieder als präferierter Pfad in Betrieb genommen. Dies wird erreicht indem Node B einen Failover initiiert zum alternative (präferierten Pfad). Das führt dazu das die MPIIO Treiber der Host einen Rescan der Pfade durchführen und erkennen, dass der präferierte Pfad wieder zur Verfügung steht und wiederhergestellt wurde. **Von diesem Zeitpunkt an ist der Normalbetrieb hergestellt und der Write-Cache wird auf beiden Nodes wieder eingeschaltet.**

Live-Demo

Fragen?

Sie möchten Storagevirtualisierung unverbindlich testen?

Wir bieten allen Teilnehmern der noon2noon Veranstaltung

die Möglichkeiten einen Voucher für eine kostenlose

Teststellung SVC oder Datacore inkl. Workshop in Anspruch

zu nehmen.

Sprechen Sie uns an!

Vielen Dank

systemzwo GmbH

Jörg Utesch
Projekt Vertrieb

Pfarrer-Weiß-Weg 10
D-89077 Ulm

Tel.: +49 (0) 731 - 14 11 60-20
Fax: +49 (0) 731 - 14 11 60-99
Mobil: +49 (0) 176 - 14 11 60 20

E-Mail: joerg.utesch@sz-group.de
<http://www.sz-group.de>

systemzwo GmbH

Manuel Riedmüller
Systemzwo Professional Services

Pfarrer-Weiß-Weg 10
D-89077 Ulm

Tel.: +49 (0) 731 - 14 11 60-16
Mobil: +49 (0) 176 - 14 11 60 16

E-Mail: manuel.riedmueller@sz-group.de
<http://www.sz-group.de>

systemzwo GmbH

Daniel Dilg
Vertriebsinnendienst

Pfarrer-Weiß-Weg 10
D-89077 Ulm

Tel.: +49 (0) 731 - 14 11 60-22
Fax: +49 (0) 731 - 14 11 60-99

E-Mail: daniel.dilg@sz-group.de
<http://www.sz-group.de>