

Speichervirtualisierung

Leitfaden
Version 1

■ Impressum

- Herausgeber: BITKOM
Bundesverband Informationswirtschaft,
Telekommunikation und neue Medien e. V.
Albrechtstraße 10 A
10117 Berlin-Mitte
Tel.: 030.27576-0
Fax: 030.27576-400
bitkom@bitkom.org
www.bitkom.org
- Ansprechpartner: Holger Skurk
Tel.: 030.27576-250
h.skurk@bitkom.org
- Redaktion: Holger Skurk (BITKOM)
- Redaktionsassistentz: Diana Delvalle Silva (BITKOM)
- Gestaltung/Layout: Design bureau kokliko / Anna Menzel (outofgoren) / Astrid Scheibe (BITKOM)
- Titelbild: www.istockphoto.com/justinbalazs
- Copyright: BITKOM 2012

Diese Publikation stellt eine allgemeine unverbindliche Information dar. Die Inhalte spiegeln die Auffassung im BITKOM zum Zeitpunkt der Veröffentlichung wider. Obwohl die Informationen mit größtmöglicher Sorgfalt erstellt wurden, besteht kein Anspruch auf sachliche Richtigkeit, Vollständigkeit und/oder Aktualität, insbesondere kann diese Publikation nicht den besonderen Umständen des Einzelfalles Rechnung tragen. Eine Verwendung liegt daher in der eigenen Verantwortung des Lesers. Jegliche Haftung wird ausgeschlossen. Alle Rechte, auch der auszugsweisen Vervielfältigung, liegen bei BITKOM.

Speichervirtualisierung

Leitfaden

Version 1

Inhaltsverzeichnis

1	Management Summary	3
2	Einführung	6
3	Speichervirtualisierungstechnologien / Methoden	8
3.1	Virtualisierung auf der Speicher-Netzwerk-Ebene	9
3.1.1	In-Band-Architektur	11
3.1.2	Out-of-Band-Architektur	12
3.1.3	Split-Path-Virtualisierung mit speziellen SAN-Switches	13
3.1.4	Vorteile von Virtualisierung auf Speicher-Netzwerkebene	14
3.2	Klassische Virtualisierung innerhalb von Plattensystemen	15
3.3	Virtualisierung von externen Plattensystemen durch Storage-Controller	16
3.4	Grid-basierende Storage-Architekturen	18
3.4.1	Aufbau	18
3.4.2	Funktionsschema	18
3.4.3	Skalierbarkeit	19
3.4.4	Management	19
3.4.5	Vorteile	19
3.5	Speichervirtualisierung auf der Server-Ebene	20
3.6	Tape-Virtualisierung, Virtuelle Tape-Libraries	22
3.6.1	Klassische VTL mit Platten-Cache und Tape als Endlager	23
3.6.2	VTL als reine Disk-Lösung, Disk-VTL	23
4	Datendienste / Service Level Agreements	24
5	Vorteile der Speichervirtualisierung	25
5.1	Optimierte Nutzung der Speicherressourcen	25
5.2	Reduzierung des Verwaltungsaufwands	25
5.3	Investitionsschutz	25
5.4	Neue Funktionen für ältere Systeme	26
5.5	Zukunftsorientierung	26
5.6	Unterbrechungsfreie Datenmigration	26
5.7	Kopierservices und »Business Continuity«	27
5.8	Vereinfachung des Information Life Cycle Management (ILM)	27
5.9	Optimierte Abstimmung zwischen Speicherservice und Anwendungsanforderung	27
5.10	Optimierung der Gesamtspeicherkosten	27
5.11	Fazit	28
	Danksagung	29

1 Management Summary

In der Vergangenheit verdoppelte sich die zu speichernde Datenmenge ca. alle 1,5 Jahre. Eine Abschwächung dieses Trends ist momentan nicht in Sicht. Allein im Jahre 2012 wird weltweit eine Datenmenge erzeugt werden, die einem Festplattenturm von 600km Höhe entspricht. Die Herausforderung für jedes Unternehmen ist es deshalb, die zunehmenden Datenmengen effizient und sicher zu handhaben. Speichervirtualisierung ist dafür ein ideales Konzept.

Virtualisierung von IT-Infrastrukturen ist in aller Munde. Servervirtualisierung verbessert die Auslastung und Flexibilität der Serverlandschaft, Clientvirtualisierung ermöglicht es, durch Automatisierung von Bereitstellungs- und Wartungsprozessen Kosten zu sparen. Durch Kombination dieser beiden Technologien wird der Nutzen maximiert. Speichervirtualisierung ergänzt diese Virtualisierungstechnologien und bietet die dafür benötigte Flexibilität.

Durch Speichervirtualisierung wird die Speicherinfrastruktur besser ausgenutzt und deren Skalierbarkeit erhöht sich. Dadurch sinken die Kosten bei gleichzeitig besserer Anpassbarkeit an die Geschäftsprozesse. Speichervirtualisierung ist also DAS Werkzeug zur Konsolidierung von Speichertechnologien und ein Schlüssel zur Kostenoptimierung und Effizienzsteigerung.

IT Services

Management & Provisioning

Virtualisierung von:	Client	Server	Speicher
Vorteile	<ul style="list-style-type: none"> ■ Einsparung von Hardware und Energie ■ Höhere Verfügbarkeit ■ Größere Flexibilität ■ Zentrale Datenhaltung ■ Einfacheres Desktopmanagement 	<ul style="list-style-type: none"> ■ Einsparung von Hardware und Energie ■ Höhere Verfügbarkeit von Systemen ■ Größere Flexibilität ■ Einfachere Automatisierung 	<ul style="list-style-type: none"> ■ Effektivere Nutzung von Speicherressourcen ■ Unterbrechnungsfreie Datenmigration ■ Einfacheres Management
Technologien	<ul style="list-style-type: none"> ■ Remote Desktop (VDI) ■ Remote Desktop Protokolle ■ Connection Broker ■ Lokale Virtuelle Maschine 	<ul style="list-style-type: none"> ■ Vollvirtualisierung ■ Paravirtualisierung ■ Partitionierung 	<ul style="list-style-type: none"> ■ Host-basierte Virtualisierung ■ Netz-basierte Virtualisierung ■ Controller-basierte Virtualisierung
Auswirkungen / Business Impacts	<ul style="list-style-type: none"> ■ Entkopplung Lebenszyklen von Hardware und Betriebssystem, dadurch maximaler Investitionsschutz ■ Geringere Kosten für Hardware und Energie, vor allem beim Einsatz von Thin Clients ■ Reduktion von Betriebsaufwänden ■ Erhöhte Datensicherheit 	<ul style="list-style-type: none"> ■ Entkopplung Lebenszyklen von Hardware und Betriebssystem, dadurch maximaler Investitionsschutz ■ Minimierung der Downtimes ■ Geringere Kosten für Hardware und Energie ■ Schnellere Umsetzung von Businessanforderungen. 	<ul style="list-style-type: none"> ■ Investitionsschutz ■ Geringere Betriebsaufwände ■ Schnellere Umsetzung von Businessanforderungen
Besonderheiten	<ul style="list-style-type: none"> ■ Synergien mit Server-Virtualisierung 	<ul style="list-style-type: none"> ■ Synergien mit Client-Virtualisierung 	<ul style="list-style-type: none"> ■ Ergänzt Server-Virtualisierung

Virtualisierung von:	Netzwerk	Applikation
Vorteile	<ul style="list-style-type: none"> ■ Einfaches Ressourcensharing ■ Reduktion der Verkabelung ■ Größere Flexibilität ■ Einfacheres Management 	<ul style="list-style-type: none"> ■ Plattformunabhängiger Betrieb von Anwendungen ■ Auflösung von Inkompatibilitäten zwischen Anwendungen untereinander ■ Auflösung von Inkompatibilitäten zwischen Anwendung und Betriebssystemen
Technologien	<ul style="list-style-type: none"> ■ Virtual Local Area Network (VLAN) ■ Virtual Private Network (VPN) ■ Virtuelle Switche 	<ul style="list-style-type: none"> ■ Kapselung der Anwendung in einen Container/einer Sandbox ■ Emulatoren ■ Laufzeitumgebungen
»Auswirkungen / Business Impacts«	<ul style="list-style-type: none"> ■ Reduktion von Betriebsaufwänden ■ Schnellere Umsetzung von Businessanforderungen 	<ul style="list-style-type: none"> ■ »Schnellere und flexiblere Anwendungsbereitstellung und Provisionierung ■ Unterstützt flexibleres Lebenszyklusmanagement«
Besonderheiten	<ul style="list-style-type: none"> ■ Ergänzt Server-Virtualisierung 	<ul style="list-style-type: none"> ■ Kann Client- und Server-Virtualisierung ergänzen

Abbildung 1: Übersicht der IT-Virtualisierungstechnologien

Der vorliegende Leitfaden soll neben den Auswirkungen und Vorteilen für die Unternehmens-IT auch über die heute aktuellen Technologien informieren und somit den IT-Verantwortlichen eine Grundlage bieten, auf der sie eine optimale Entscheidung treffen können.

2 Einführung

Virtualisierung ist in der IT ein wertvolles Instrument, um eine dynamische, agile und zukunftsorientierte Infrastruktur zu errichten. Die Speichervirtualisierung eignet sich hervorragend, um weitere Werkzeuge, wie z. B. Thin Provisioning¹ oder Tiered Storage², gewinnbringend in ein Rechenzentrum zu integrieren. Nutzt ein Unternehmen all diese Methoden, so kann es bereits getätigte Investitionen in Speichersysteme noch rentabler machen.

Zum Beispiel ermöglicht Speichervirtualisierung kombiniert mit Thin Provisioning eine sehr gute Auslastung der Ressourcen und damit eine Optimierung der Speicherkosten. Die daraus resultierende Reduzierung des Verwaltungsaufwands führt zu einer weiteren Kostenreduktion, da die gleiche Anzahl von Mitarbeitern die immer größer werdenden Datenmengen verwalten kann.

Virtualisierung an sich ist zunächst nichts anderes als eine Technologie, die eine Abstraktionsebene zwischen Anwendern bzw. Applikationen und den physischen Ressourcen einführt. Anwendungen greifen dadurch also nicht mehr direkt auf z. B. eine bestimmte Festplatte zu, sondern auf die Virtualisierungsschicht. Sinn dieses Ansatzes ist es, eine Vielzahl an physischen Ressourcen logisch zusammenzuführen. Anwendern gegenüber präsentiert sich das System damit als eine große Einheit. Speicherkomponenten werden dem Anwender in einer logischen Form zur Verfügung gestellt, der vorhandene Speicher kann somit flexibel aufgeteilt werden und ist nicht an die physischen Grenzen gebunden. Dadurch können Speicherressourcen effizient ausgenutzt werden. Weiterhin ist es auch möglich, neue Funktionen über die Virtualisierungsschicht bereitzustellen, die die ursprünglichen Speichersysteme noch nicht hatten. Ältere Speichersysteme können so auch an neuere Anforderungen angepasst und damit länger genutzt werden.

Zu diesen neueren Anforderungen und Funktionen gehören:

- **Thin Provisioning:**
Die Blöcke eines Plattensubsystems bzw. LUNs (Logical Unit Number³) werden nicht beim Erstellen abgelegt, sondern erst, wenn von der Server-Seite geschrieben wird. Unter der Annahme, dass die Auslastung von LUNs selten 100% erreicht, lassen sich auf diese Weise deutlich höhere Kapazitäten zur Verfügung stellen als physikalisch Plattenplatz zur Verfügung steht. Diese Funktion bedingt allerdings ein ausgefeiltes Monitoring der freien Kapazitäten um im Bedarfsfall rechtzeitig reagieren zu können (z. B. Kapazitäten nachschieben)
- **Deduplikation:**
Datenbestände (strukturierte und unstrukturierte Daten) enthalten je nach Anwendung redundante Informationen, was zu einer Vielzahl gleicher Daten-segmente führt. Mit Deduplikation werden Segmente mit gleichem Inhalt nur einmalig vorgehalten. Der Deduplizierungsprozess kann »inline« erfolgen oder über einen nachgelagerten Prozess. Die Effizienz des zur Verfügung gestellten Speichers wird bei diesem Verfahren deutlich erhöht. Nicht zu vergessen ist aber, dass dieser Prozess auch die Anzahl der Schreib- / Leseoperationen deutlich erhöhen kann und man mehr Performance zur Verfügung stellen muss, als eigentlich von der Applikation benötigt wird. Deduplikation darf nur eingesetzt werden, wenn die dahinterliegende Technologie eine 100%ige Datensicherheit gewährleistet.
- **Snapshots und Clones:**
»Snapshots sind zu einem definierten Zeitpunkt eingefrorene Zustände einer LUN, die man schreibbar machen kann«. Snapshots sind in der Regel »space

¹ Siehe nächste Seite oder z. B. hier: http://de.wikipedia.org/wiki/Thin_Provisioning

² http://en.wikipedia.org/wiki/Hierarchical_storage_management#Tiered_storage

efficient« – d.h. es werden zur Darstellung nur die Änderungen auf dem Original abgespeichert. Es lassen sich somit viele logische 100% Kopien erzeugen ohne 100% der Datenblöcke vorhalten oder kopieren zu müssen. In bestimmten Situationen ist es sinnvoll anstatt der Änderungen eine volle Kopie anzulegen, da ein Zugriff auf einen Snapshot Performance von der ursprünglichen LUN abzieht. Diese Point in Time Kopie nennt man Clone, der z. B. bei Applikationsupgrades verwendet wird.

■ **Replikation:**

Unter Replikation versteht man die Fähigkeit, auf Speichersystem-Ebene LUNs oder Gruppen von LUNs zwischen zwei Systemen zu replizieren. Die Replikation erfolgt in der Regel in eine Richtung und kann aus Sicht des Hosts synchron oder asynchron erfolgen. Replikation lässt sich oftmals mit Snapshots und Clones kombinieren.

■ **Compression (Komprimierung):**

Eine Funktion zur Erhöhung der Effizienz zwischen physikalischen und logischen Kapazitäten.

■ **Inline Migration:**

In größeren Speichersystemen lassen sich aufgrund unterschiedlicher Plattenlaufwerke und RAID-Mechanismen sogenannte Speicherklassen abbilden. Über eine »inline« Migration können Daten/LUNs transparent zwischen diesen Speicherklassen verschoben werden, »transparent« bedeutet hier bei laufendem Zugriff von der Server-Seite aus.

Speichervirtualisierung vereinfacht es also auch, Daten für eine bestimmte Zeit auf ein leistungsstärkeres (oder leistungsschwächeres) Speichersystem zu verschieben, um die aktuellen Geschäftsanforderungen zu bedienen. Ein Unternehmen behält die Freiheit, Kapazitäten unterschiedlicher Anbieter im laufenden Betrieb in den

Speicherpool einzubinden, ohne organisatorische Änderungen vornehmen zu müssen. Virtualisierung hilft also nicht nur bei Datenverschiebungen, sondern auch bei »Performance-Transfers«.

Unter Berücksichtigung der Tatsache, dass Speichersysteme in der Regel in Zeiträumen zwischen drei und fünf Jahren erneuert werden und die Daten für die Migrationszeit zu den neuen Systemen, die Monate dauern kann, doppelt gehalten werden, führt diese Funktionalität zu deutlichen Kostenreduzierungen im Unternehmen.

Speichervirtualisierung bildet ein wichtiges Fundament für das sogenannte Information Life Cycle Management (ILM). Unternehmen können den Datentransfer dann über unterschiedliche Speicherklassen von heterogenen Speichersystemen unterbrechungsfrei und schnell umsetzen. Somit können Unternehmensinformationen in jeder Phase ihres Lebenszyklus optimal und kostengünstig verwaltet werden.

Zusammengefasst hat eine Speicherlösung mit Virtualisierungsfunktion folgende Vorteile:

- Reduzierung der Zahl der zu verwaltenden Elemente
- Einsparungen bei Softwarelizenzen und Wartungsleistungen
- Verringert den Administrationsaufwand
- ermöglicht die Zuordnung von freigegebenen Ressourcen für andere Projekte
- vereinfacht die Inbetriebnahme neuer Server
- erleichtert die Speicherung von Daten mehrerer Klassen, eine so genannte »Tiered Storage«-Lösung.

Dies führt zu einer signifikanten Reduktion der Gesamtspeicherkosten.

Bei der Überlegung, eine Virtualisierungslösung einzusetzen, ist die Unternehmensgröße übrigens nicht entscheidend.

³ Bezeichnet eine logische Einheit, i.A. eine virtuelle Festplatte in einem Disk Array, siehe auch Kapitel 3.1 oder <http://de.wikipedia.org/wiki/LUN>

3 Speichervirtualisierungstechnologien / Methoden

Dieses Kapitel beschreibt detailliert die verschiedenen Virtualisierungsmethoden und gibt damit einen fundierten Überblick über die verwendeten Technologien. Das ermöglicht den IT-Verantwortlichen eine Beurteilung der Vor- und Nachteile und damit die optimale Auswahl für das eigene Unternehmen.

Die Virtualisierung des Speichers wird, wie beschrieben, durch Verwendung einer Abstraktionsschicht zwischen

Speicher und Anwendungen erreicht. Diese Abstraktionsschicht, die im Folgenden wegen ihrer Hauptfunktion Virtualisierungsschicht genannt wird, kann an verschiedenen Punkten der Speicherinfrastruktur ansetzen:

Abbildung 2 zeigt die verschiedenen Ansatzpunkte für die Virtualisierungsschicht. Diese begründen verschiedene Virtualisierungstechnologien, welche im folgenden genauer dargestellt werden.

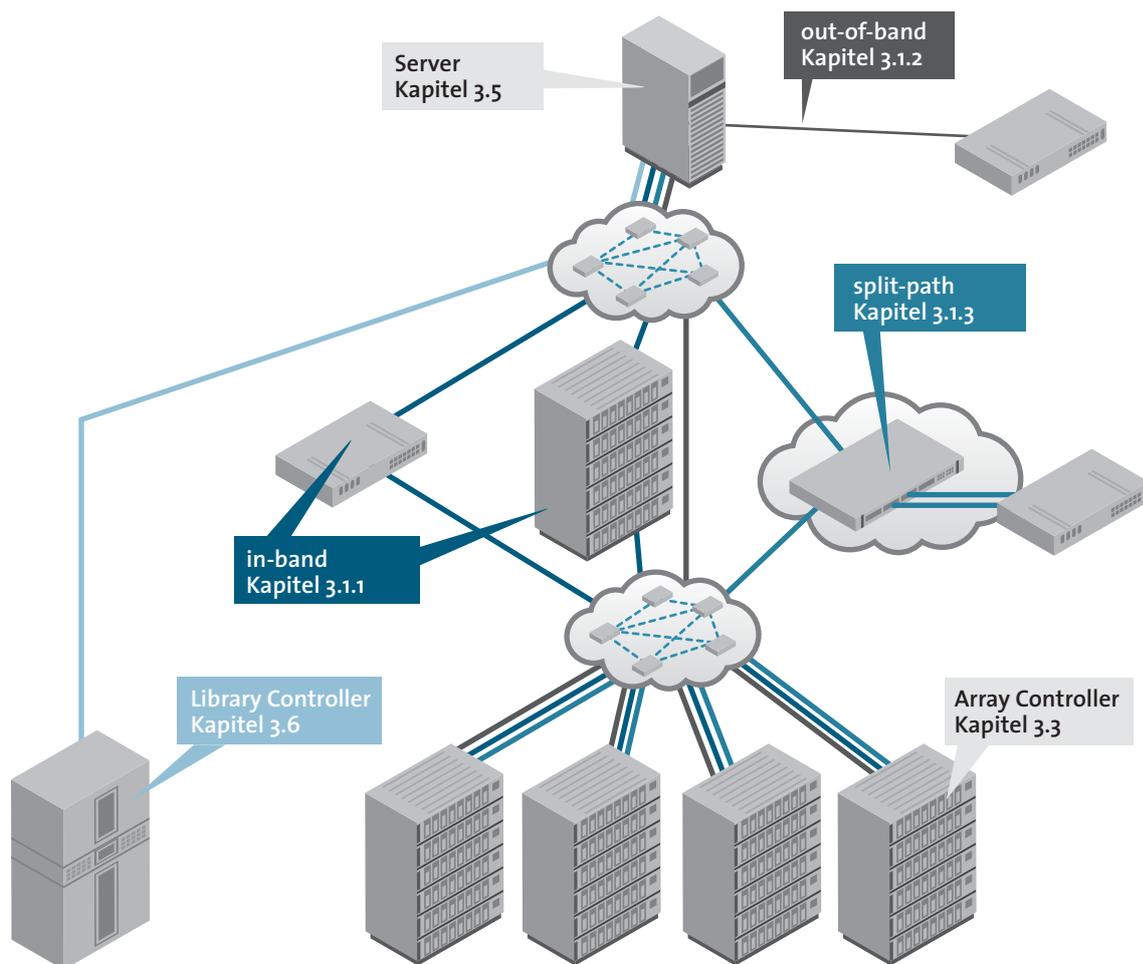


Abbildung 2: Übersicht über die Speichervirtualisierungstechnologien

■ 3.1 Virtualisierung auf der Speicher-Netzwerk-Ebene

Die Prinzipdarstellung dieser Art der Speichervirtualisierung zeigt die folgende Abbildung:

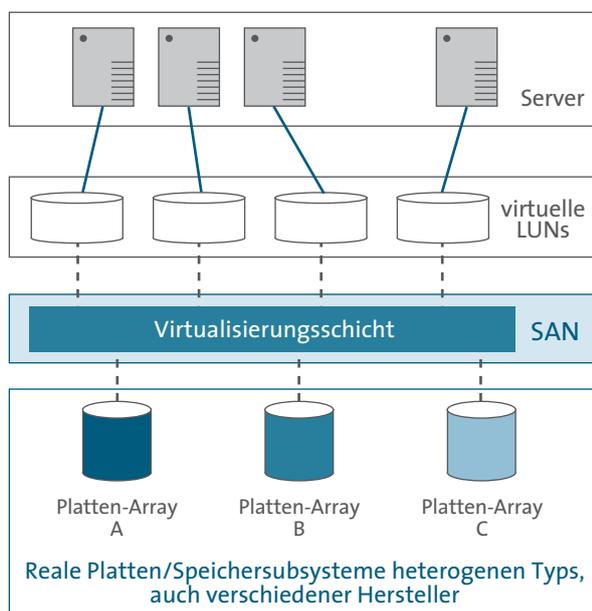


Abbildung 3: Speichervirtualisierung auf Netzwerk-Ebene

Die sich in der Speicher-Netzwerkebene (SAN) befindende Virtualisierungsschicht zeigt den Hosts virtuelle Festplatten bzw. LUNs. Die realen LUNs sind üblicherweise LUNs in realen Speichersubsystemen/Plattenarrays oder – heute eher selten – existieren als pure Plattenlaufwerke (JBOD=Just a bunch of disks). Die Speichersubsysteme können unterschiedlichen Typs sein und von verschiedenen Herstellern stammen.

Speichervirtualisierungen auf Netzwerkebene sind auch über mehrere Standorte realisierbar. So, wie Daten zwischen homogenen intelligenten Plattensubsystemen von einem lokalen auf einen entfernten Standort gespiegelt werden können, können Daten über die lokale Virtualisierungsschicht an eine entfernte Virtualisierungsschicht gespiegelt werden, inklusive Cache-Kohärenz der lokalen

und der entfernten Virtualisierungsschicht. Dadurch ergibt sich eine erhöhte Datenmobilität.

Da sich die von den Hosts gesehenen LUNs in einer so gearteten Architektur in einem Pool von unterschiedlichen physischen Plattensubsystemen befinden, steigt die Flexibilität, mit der Plattenplatz zugewiesen werden kann, erheblich. Können Hosts bei Architekturen ohne diese Technologie ihren Speicher nur aus einem bzw. mehreren strikt getrennten Plattensubsystemen rekrutieren, so erweitert sich dieses Spektrum jetzt erheblich. Alle Funktionalitäten, die sonst auf ein Plattensubsystem beschränkt waren (Snapshots, LUN-Migrationen, Thin Provisioning usw.) funktionieren über Plattensubsystem-Grenzen hinweg und umfassen sogar einfache Platten (JBODs) ohne Zusatzfunktionen. Die Zusatzfunktionen wie Raid oder Thin Provision werden dann von der Virtualisierungsschicht bereitgestellt.

Möglich werden damit für Hosts transparente LUN-Verschiebungen zwischen Plattensubsystemen unterschiedlicher Hersteller, ohne Applikationsunterbrechungen, im laufenden Betrieb. Sonst höchst aufwändige kritische Datenmigrationen sind nun auf einfache Weise realisierbar. Der Anwender ist nicht mehr auf einen Hersteller fixiert und kann ohne erhebliche Aufwände von einer bestehenden auf eine neue Technologie wechseln. Beispielsweise im Rahmen von Unternehmenszusammenschlüssen mit heterogenen Platten-Subsystemen kann sich diese Form von Speichervirtualisierung als hilfreich erweisen.

Bei der Virtualisierung auf Speicher-Netzwerkebene können grundsätzlich drei Architekturen unterschieden werden:

- In-Band-Architektur
- Out-of-Band-Architektur
- Split-Path-Architektur

Diese Unterscheidung richtet sich nach den Übertragungswegen von Kontroll- und Steuerbefehlen einerseits und den eigentlichen Daten andererseits:

Zwischen Server- und Speichersystem gibt es einen SCSI-Datenstrom. Dieser Datenstrom beinhaltet sowohl die

Nutzdaten an sich als auch Metadaten wie SCSI-Kontrollbefehle, Befehle für Speicherdienste etc. Zur Unterscheidung der Architekturen ist der Umgang mit den Kontrollbefehlen und den Nutzdaten geeignet.

Diese Architekturen werden im Folgenden erläutert.

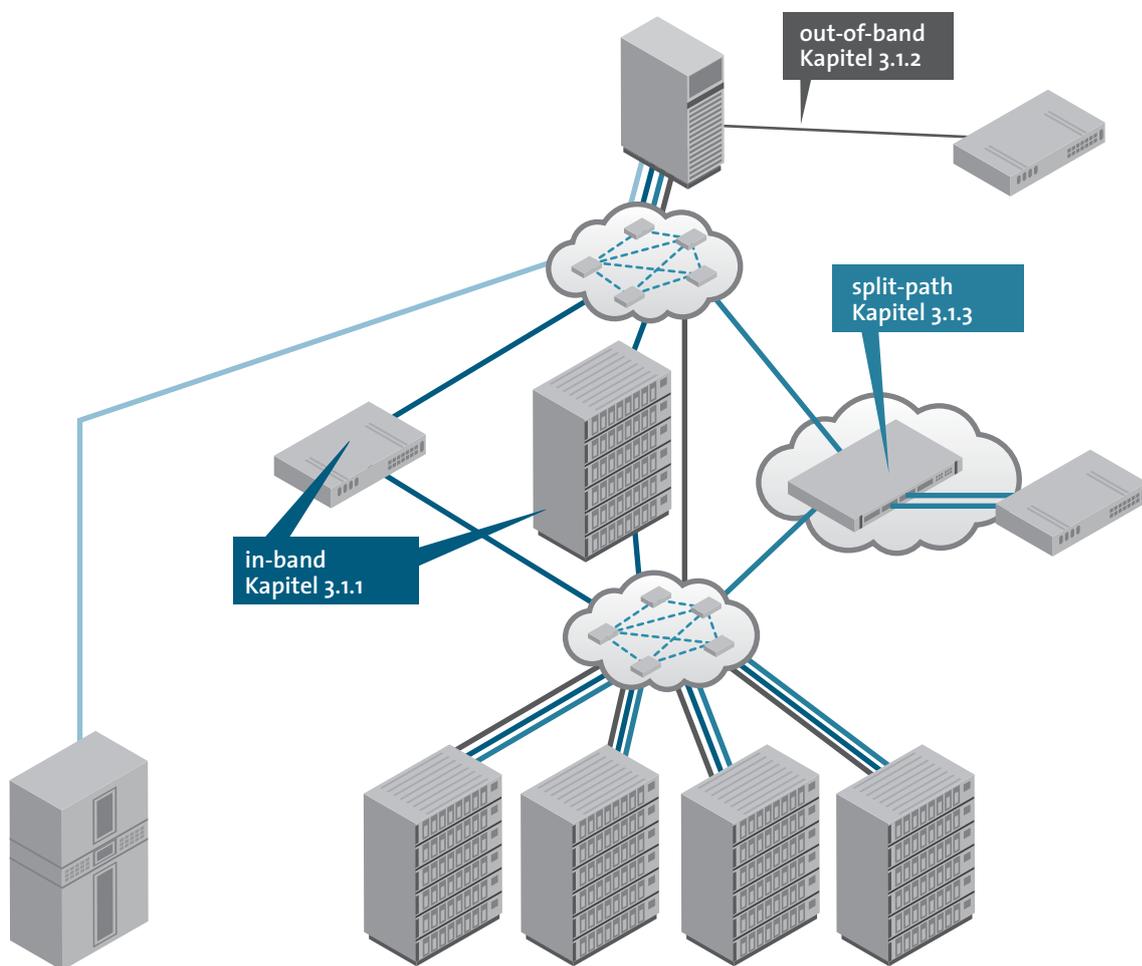


Abbildung 4: Virtualisierung auf Netzwerkebene

3.1.1 In-Band-Architektur

Bei der In-Band-Architektur werden Nutzdaten und Metadaten auf demselben Weg übertragen.

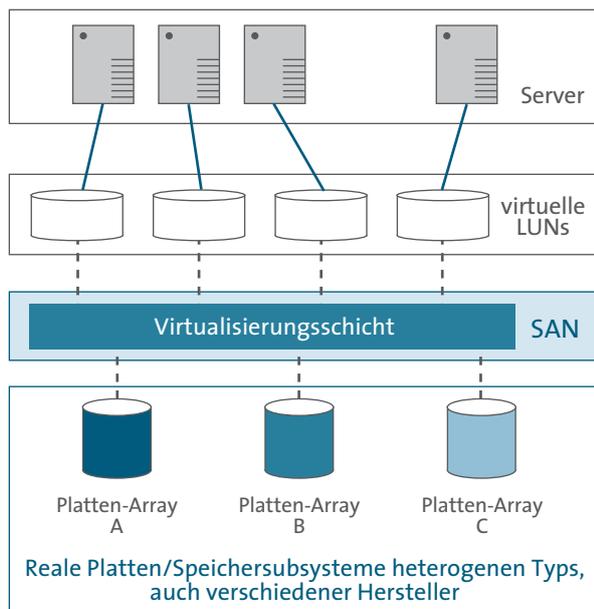


Abbildung 5: Virtualisierung auf Netzwerkebene: In-Band-Architektur

Die Virtualisierung findet in einer zusätzlichen »Appliance« statt. Bei einigen Realisierungen werden zu diesem Zweck Server eingesetzt, die den Hosts virtuelle Platten zeigen. Diese Platten werden durch JBODs oder Platten-Subsysteme realisiert. Die Appliances befinden sich im Datenpfad zwischen den Hosts und den realen Platten (in-Band) und sind im Speichernetzwerk (SAN) -FibreChannel- oder iSCSI-basierend-sowohl mit den Hosts als auch mit den realen LUNs verbunden.

Sowohl alle Kontroll- und Metadaten als auch die eigentlichen Nutzerdaten laufen durch die Virtualisierungs-Appliance (in-Band).

Grundsätzlich müssen auf den Hosts zur Adressierung der virtuellen LUNs keine zusätzlichen Komponenten installiert werden. Allerdings ist im Rahmen von Hochverfügbarkeit das Thema Multipathing zu berücksichtigen. Beim

Multipathing werden zwei redundante Datenpfade zur Kommunikation zwischen Speicher und Server verwendet.

Im Allgemeinen werden Standard-Lösungen wie MPIO (Multipath I/O) unterstützt. Am Markt existieren aber weitere Multipathing-Werkzeuge, deren Möglichkeiten insbesondere in Bezug auf Loadbalancing über die von MPIO hinausgehen. Nachteil dieser proprietären Multipathing-Varianten ist die zusätzliche Abhängigkeit von einem nicht vom Betriebssystem kommenden Software Produkt. Ein eventuell dringend benötigter Applikationswechsel, der einen Releasewechsel des Betriebssystems notwendig macht, kann somit durch eine Software eines dritten Herstellers verzögert werden. Es muss also im Einzelfall unter Berücksichtigung der Vor- und Nachteile geklärt werden, welche Multipath-Variante in Verbindung mit welcher Virtualisierungs-Lösung sinnvoll genutzt werden kann.

Der Markt offeriert neben den Server-basierenden In-Band-Lösungen weitere Ausprägungen.

Man findet zum Beispiel Lösungen, bei denen die Virtualisierung in den Controllern/Storage-Prozessoren von Speichersubsystemen realisiert wird. Neben den »eigenen« Platten können diese Plattensubsysteme auch externe Plattensubsysteme, ggf. anderen Typs und anderer Hersteller, adressieren. Diese Thematik wird im Kapitel 3.3 näher erörtert.

Eine weitere Variante stellen sogenannte Gateway-Lösungen dar. (Abb. 6) Controller/Storage-Prozessoren klassischer intelligenter Plattensubsysteme werden hier als Virtualisierungs-Appliances eingesetzt. Die Controller/Storage-Prozessoren »besitzen« keine »eigenen« Platten im Backend, sondern beziehen ihre LUNs aus externen Plattensubsystemen und nicht, wie üblich, aus integrierten Platten-Einschüben.

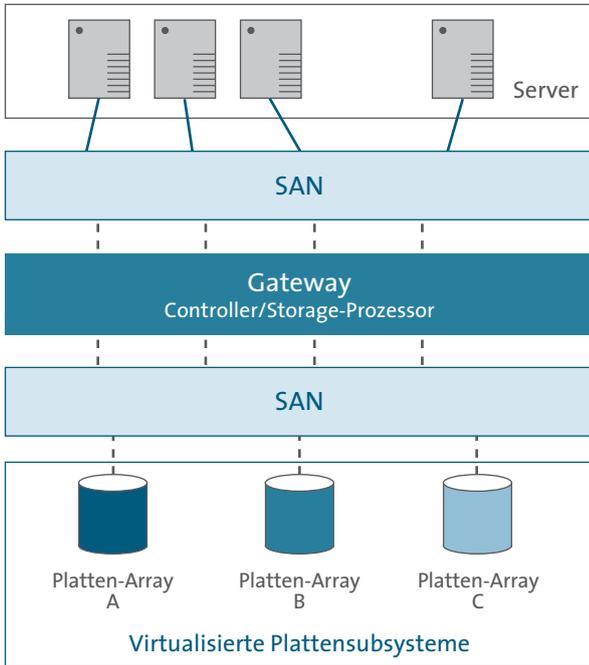


Abbildung 6: Virtualisierung auf Netzwerkebene: Gateway-Lösungen

3.1.2 Out-of-Band-Architektur

Bei Out-of-Band-Lösungen wird die Speichervirtualisierung auf einer Virtualisierungs-Appliance realisiert, die nur Kommandodaten und Kontroll-Metadaten verarbeitet und die Virtualisierung ausschließlich steuert. (Abb. 7)

Die eigentlichen Daten werden direkt über das SAN geschickt. Der Virtualisierungs-Server kann sich auch im lokalen Netz befinden. In einer Out-of-Band - Umgebung muss in allen beteiligten Hosts zusätzliche Interface-Software der Virtualisierungs-Lösung installiert werden. Die Virtualisierungsappliance übermittelt die Adressen der zu übertragenden Datenblöcke an diese Interface-Software, die dann die Datenübertragung anstoßen kann.

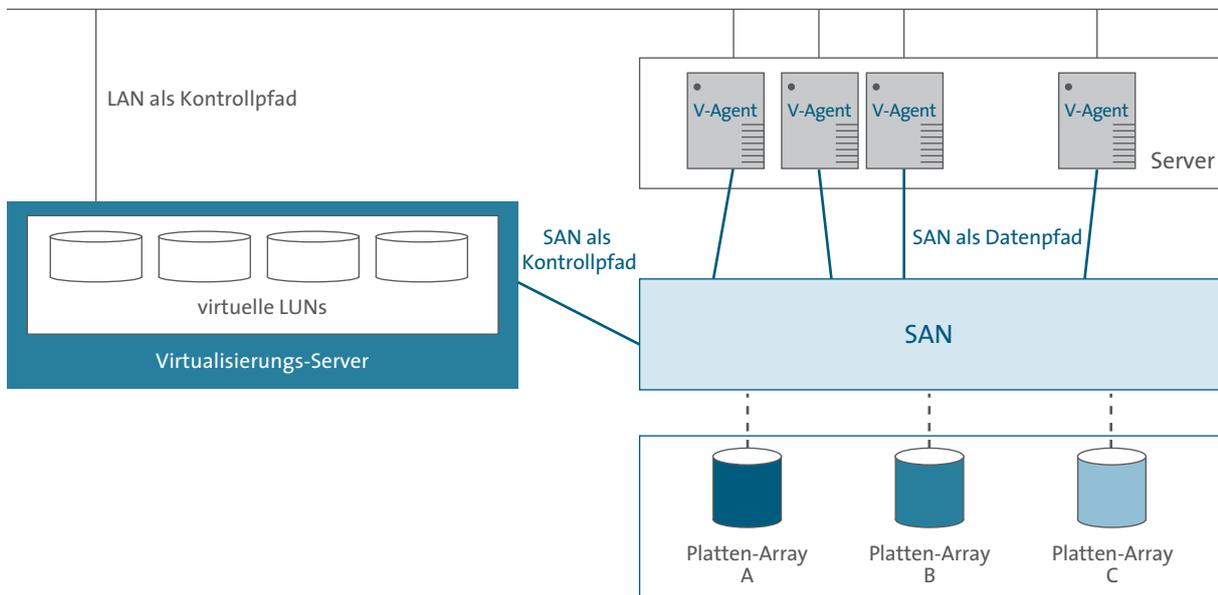


Abbildung 7: Virtualisierung auf Netzwerkebene: Out-of-Band-Architektur

3.1.3 Split-Path-Virtualisierung mit speziellen SAN-Switches

Für diese Architektur werden zur Speichervirtualisierung spezielle SAN-Switches (auch »intelligente« oder »smart« Switches) eingesetzt. Die Virtualisierung findet im SAN-Switch statt. Dies passiert direkt am Port (»Intelligenter Switch-Port« oder »Fast Port«) ggf. in Verbindung mit – intelligenten – Speichervirtualisierungskomponenten innerhalb des Switches.

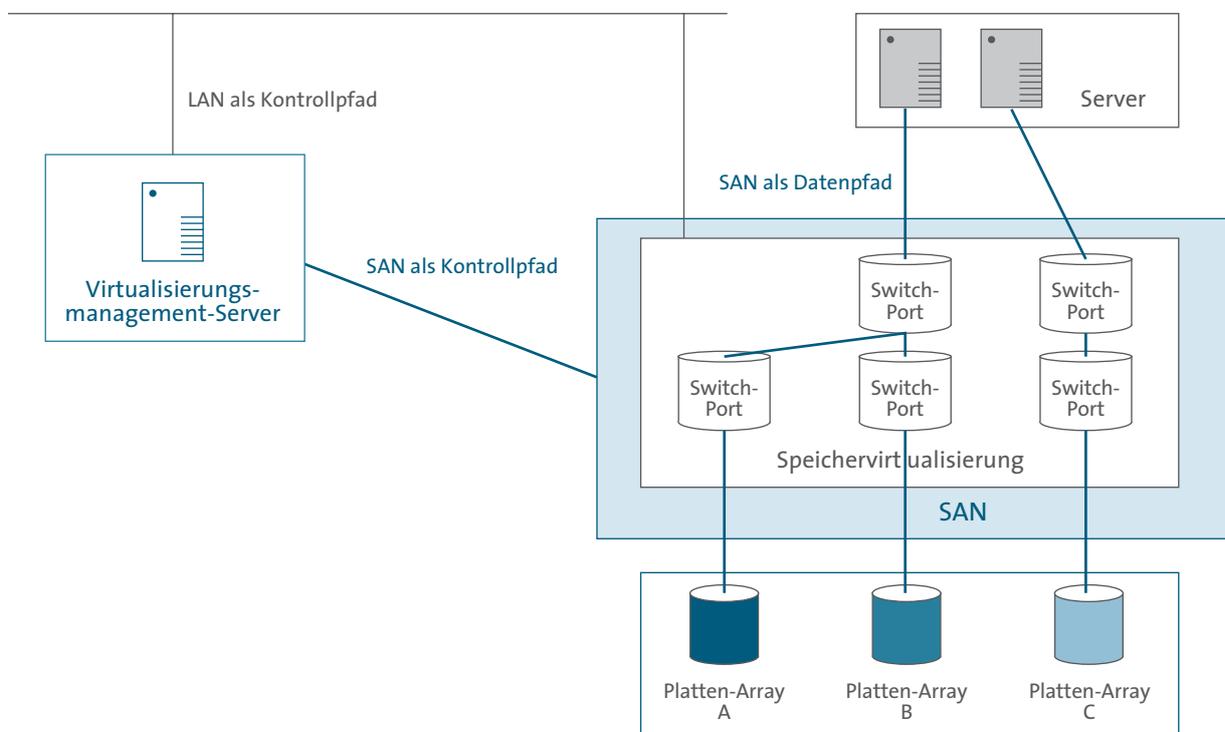


Abbildung 8: Virtualisierung auf Netzwerkebene: Splith-Path

Der Begriff »Split-Path« begründet sich damit, dass der Kontrollpfad vom Datenpfad getrennt wird.

Daten und Kontrollinformationen werden vom Host an den FiberChannel-Port (rote Verbindungen in den Grafiken) geschickt, wo sie voneinander getrennt werden. Die Kontrolldaten werden interpretiert, die Anfrage an virtuelle Volumes wird in eine Anfrage an ein oder mehrere physikalische Volumes übersetzt und an das reale Platten-Array gesendet. Host-seitige Software wie bei

Out-of-Band-Lösungen ist nicht erforderlich. Das Management der Speichervirtualisierung wird von einem anderen Server wahrgenommen, der mit den im Switch angeschlossenen Speichervirtualisierungseinheiten kommuniziert.

Man redet im Zusammenhang mit Split-Path-Lösungen auch von Hybrid-Lösungen, in denen In-Band- und Out-of-Band-Techniken zusammenspielen.

3.1.4 Vorteile von Virtualisierung auf Speicher-Netzwerkebene

Die Hosts sind nicht mehr direkt mit physischen Platten oder Plattensubsystemen verbunden, sondern über eine Virtualisierungsschicht. Damit stellen sich mehrere Platten-Subsysteme wie ein einziges dar. Plattenarrays unterschiedlicher Bauart und verschiedener Hersteller werden zu einem Pool zusammengefasst. Wie mit intelligenten Plattensubsystemen können also Funktionen wie Daten-Verschiebungen oder Daten-Replikationen (Datenspiegelungen usw.) nicht nur innerhalb eines Plattenarrays oder innerhalb homogener Welten, sondern auch über Platten-System - Grenzen hinweg erfolgen.

Die Administration von heterogenen bzw. Mehr-Hersteller-Umgebungen wird durch Speichervirtualisierung einfacher, denn der Speicher-Administrator muss sich dann nur noch mit einem Werkzeug befassen.

Ein Anwender ist nicht mehr abhängig von einer bestimmten Speichersystemplattform, sondern kann Speichersysteme, vorausgesetzt, sie werden von der Virtualisierungslösung unterstützt, frei wählen.

Vergessen werden darf aber nicht, dass sich ein Anwender unter Umständen in gewisse Abhängigkeit vom Lieferanten der Virtualisierungslösung begibt. Ein wichtiges Entscheidungskriterium ist die Möglichkeit, die Virtualisierungslösung zu wechseln oder zu migrieren.

Letzten Endes kann in relativ homogenen Umgebungen auf viele Funktionalitäten moderner Platten-Subsysteme verzichtet werden. Man spart also Kosten für Zusatzoptionen von Plattensubsystemen, sollte diese aber im Rahmen der Aufwände für die Virtualisierungsebene berücksichtigen. In homogenen Umgebungen findet man Virtualisierungslösungen auf Netzwerkebene deshalb eher selten.

Man kann davon ausgehen, dass Plattenspeichervirtualisierungen auf Netzwerk-Ebene auch in Zukunft Lösungen werden bieten können, die mit denen homogener Plattensubsysteme vergleichbar sind. Speichervirtualisierung eignet sich deshalb besonders, dann wenn Plattensubsysteme unterschiedlicher Hersteller oder unterschiedlichen Alters mit neuer Funktionalität ausgestattet werden sollen.

Man beobachtet dort beispielsweise Lösungen zur optimalen Zuweisung von Speicher, wo automatisch und abhängig von aktueller Last LUNs oder sogar Teile davon auf Platten unterschiedlicher Qualität verschoben werden. Es findet Multi-Tiering zwischen sehr schnellen Flash-Disks, klassischen schnellen FibreChannel- und relativ langsamen, aber mit hoher Kapazität ausgestatteten SATA-Platten statt. Thin Provisioning, spielt ebenfalls eine große Rolle.

3.2 Klassische Virtualisierung innerhalb von Plattensystemen

Im Unterschied zur Virtualisierung auf der Ebene des Speichernetzwerkes (SAN) findet im folgenden Ansatz die Virtualisierung im Controller/Storage-Prozessor der Speichersysteme statt, was sie für die angeschlossenen Server und deren Betriebssysteme transparent macht.

Spricht man heute von Speichersystemen, dann sind damit in erster Linie Systeme gemeint, die über RAID-Funktionen verfügen und über ein Speichernetzwerk (SAN) angeschlossen werden. Diese Systeme sind von Natur aus schon virtualisiert, stellen sie doch den angeschlossenen Servern Plattenlaufwerke in Form von LUNs mit verschiedensten Eigenschaften zur Verfügung, vollkommen unabhängig von den physikalischen Laufwerken. (Abb. 9) Neben Skalierung und Durchsatz stellt der Grad der Virtualisierung das wesentlichste Unterscheidungsmerkmal heutiger Speichersysteme dar. Man kann den Funktionsumfang dieser Speichersysteme in Basisfunktionen und erweiterte Funktionen unterteilen:

Basisfunktionen sind:

- Implementierung des RAID-Schutzes (inklusive der Verwaltung von Spare-Platten) Bereitstellen und Verwalten logischer Plattenlaufwerke (LUNs).
- Einfache Systeme verwenden fixe RAID-Gruppen (feste Anzahl Laufwerke je Gruppe), aus deren Nutzdaten-Kapazitäten logische Laufwerke fester Größe erstellt werden.

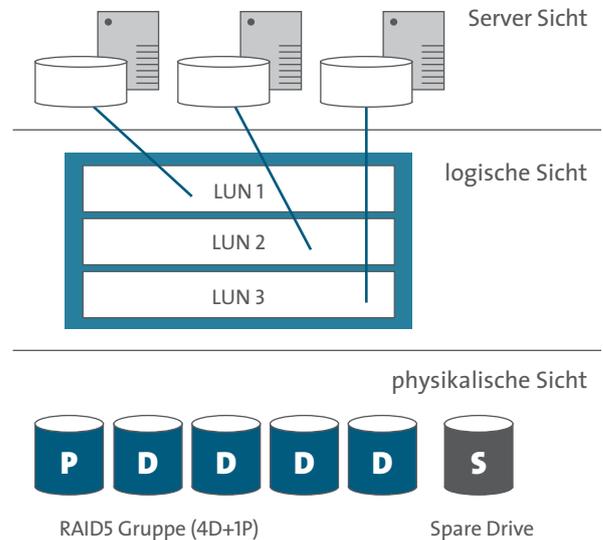


Abbildung 9: Klassische Virtualisierung

Moderne Speichersysteme erzeugen – ähnlich wie die Volume-Manager auf der Server-Ebene – Disk-Pools mit teilweise mehreren Hundert Laufwerken. Aus diesem Kapazitätspool können dynamisch LUNs erzeugt, vergrößert und verkleinert werden. Teilweise kann auf LUN-Ebene der RAID-Schutz definiert werden, oder die Pools sind von sich aus schon gegen Ausfälle abgesichert. (Abb. 10)

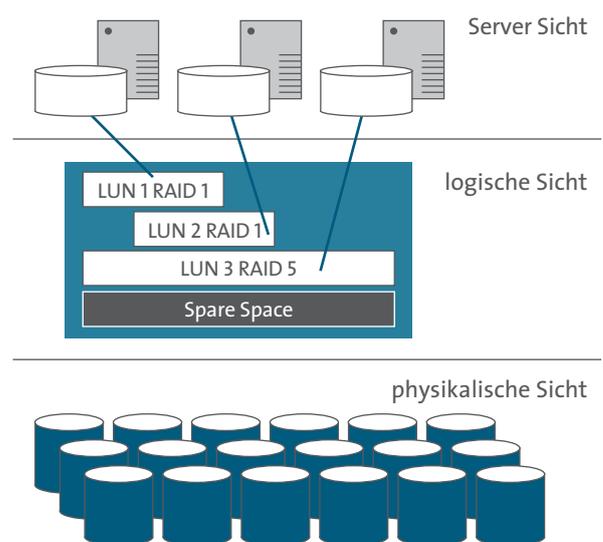


Abbildung 10: RAID-Schutz

Neben dieser Flexibilität bietet dieses Verfahren auch Geschwindigkeitsvorteile:

- selbst die kleinste logische Platte kann die IO-Leistung vieler physischer Plattenlaufwerke nutzen
- Performance-Engpässe können dynamisch durch einfaches Hinzufügen von Laufwerken behoben werden.

Um die Performance weiter zu steigern, kann eine Aufteilung auf mehrere Datenpools sinnvoll sein.

Vorteile der Virtualisierung auf Speichersystem-Ebene

Die hier beschriebenen Virtualisierungsfunktionen sind für die angeschlossenen Server-Systeme in der Regel 100%ig transparent – damit sind diese Funktionen in jeder Betriebssystemumgebung nutzbar. Weitere Vorteile sind:

- Zentrale Administration der logischen und physischen Speicher-Ressourcen Erstellen von Kopien zu Testzwecken und für den Disaster-Schutz
- Möglichkeit, über transparente Inline-Migration Daten im Rahmen eines ILM-Konzepts zwischen Speicherklassen zu verschieben.

3.3 Virtualisierung von externen Plattensystemen durch Storage-Controller

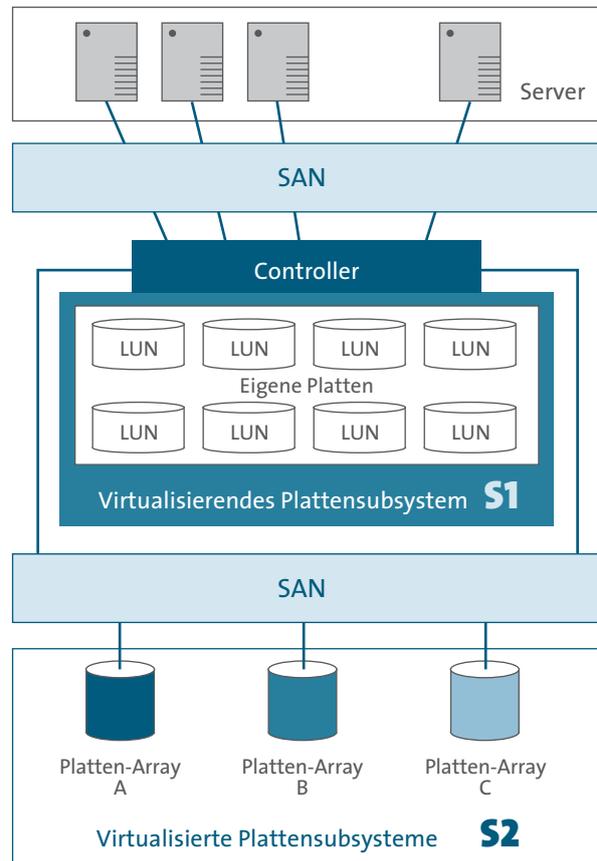


Abbildung 11: Virtualisierung durch Storage-Controller

Bei der Storage-Controller-basierten Virtualisierung (Abb. 11) findet diese direkt an der Stelle statt, an der sich auch die Daten befinden, also im Plattensystem selbst (S1). Neben den im System integrierten Platten können diese Arrays auch extern angeschlossene Speichereinheiten anderen Typs und unterschiedlicher Hersteller adressieren. (S2) Das Prinzip ist dabei sehr einfach: Es erfolgt eine Abbildung, der LUNs der extern angeschlossenen Geräte (S2) auf LUNs im virtualisierenden Speichersystem (S1). Diese logischen Einheiten werden E-LUNs genannt. Hosts sehen immer nur LUNs bzw. E-LUNs egal, wo sich die Daten physisch befinden. Genutzt werden in dieser Architektur alle Funktionen des virtualisierenden Systems (S1), die im Microcode implementiert sind und bei Bedarf

freigeschaltet werden können. Der Storage-Controller hat alle Informationen zu den I/O-Vorgängen, den Applikationen und den Speicherorten. Daher ist der Storage-Controller die Instanz, die die Speichervirtualisierung effizient umsetzen kann.

Die hoch standardisierte Kommunikation zwischen Controllern (S1) und den Plattensystemen (S2) bietet Sicherheitsvorteile im Fall von Fehlern oder bei Schreibaussfällen. Es werden keine neuen Standards und keine Software-Agenten auf Hosts benötigt. SAN und Datenpfad können unverändert und schlank bleiben. Alle zu virtualisierenden Speichersysteme werden an ein universelles Speichersystem direkt angeschlossen. Zusätzliche Schichten sind somit überflüssig und das Konzept bleibt einfach. Heterogene Insellösungen können so in eine homogene Speicherlandschaft überführt werden.

Der größte Vorteil des Speichersystem-Controller-basierten Virtualisierungskonzepts ist, dass jeglicher Speicher – egal ob intern oder extern – jedem Host (Unix, Windows, Linux oder Mainframe) zur Verfügung steht. Dabei spielt es keine Rolle, ob der Host über SAN, NAS, FC, ESCON oder FICON angebunden ist. Zusätzlich zur Speichervirtualisierung stehen alle Datenmobilitätsfunktionen auch für extern angeschlossenen Speicher zur Verfügung: Dazu zählen etwa Snapshots, Remote Mirroring, Server Priority Management oder transparente Volume Migration zwischen verschiedenen Speicherklassen des Speicher-Controllers.

Für die Gewährleistung der Ausfallsicherheit der gesamten Speicherkonfiguration und für unterbrechungsfreie Migrationen der gesamten Infrastruktur – virtualisierende Instanz (S1) und virtualisierter Speicher (S2) – sorgen Clustering-Mechanismen, in denen zwei Speicher-Controller involviert sind.

Der Vorteil dieser Methode liegt darin, einen heterogenen Bestand von Plattensystemen aus Sicht der Hosts homogen darzustellen. Speicher-Administratoren kümmern sich nach erstmaliger Einrichtung nur noch um das virtualisierende System und nutzen nur dessen Werkzeuge. Als Multipathing-Tool werden auf den Hosts der für das

virtualisierende System vorgesehene Mechanismus genutzt. Interessant ist diese Lösung, wenn verschiedene Typen von Plattensystemen vorhanden sind und auf längere Sicht das virtualisierende System zum Standard für den Anwender werden soll. Unternehmen, die bereits das Plattensystem für die Speicherung der Unternehmensdaten erworben haben, können, wenn die Geschäftsprozesse es diktieren, die Virtualisierungsfunktion einschalten, ohne zusätzliche Hardware- und Software installieren zu müssen und ohne das SAN zu verändern.

Diese Methode erlaubt die Bildung von mehreren, autarken, logischen Speicherbereichen, die aus den Ressourcen der involvierten Speichersysteme gebildet werden können. Sowohl für Server, als auch für die Speicheradministratoren erscheint eine solche virtuelle Partition wie ein separates Speichersystem, das z. B. für bestimmte Abteilungen, Unternehmensbereiche oder Kunden unterteilt werden kann. Innerhalb einer virtuellen Partition können individuelle Administrationsrechte vergeben werden. Eine exakt am Bedarf orientierte Aufteilung von Ports, Cache und Speicherkapazität für die virtualisierende Plattform und deren externer Speicherressourcen wird damit ermöglicht. Diese Partitionen lassen sich dynamisch beeinflussen, um die Infrastruktur an sich wandelnde geschäftliche Anforderungen anpassen zu können.

3.4 Grid-basierende Storage-Architekturen

Neben den traditionellen Storagearchitekturen auf Basis redundanter Controller und angeschlossener Festplatten haben sich neuartige Grid-basierende Storage-Ansätze etabliert. Diese Speicher-Grids setzen sich aus mehreren Speicherknoten zusammen.

3.4.1 Aufbau

Jedes Grid kann aus mehr als zwei Speicherknoten bestehen und um einzelne Speicherknoten erweitert werden. Jeder Speicherknoten stellt eine abgeschlossene Einheit dar. Er verfügt über Speicherkapazität, Speicherfunktionalität und Rechenleistung (Prozessor, Cache, Memory, RAID-Schutz und Netzwerkanschlüsse inkl. deren Load-Management). Idealerweise wird bei den Speicherknoten auf leistungsfähige Industriestandard-Server-Hardware zurückgegriffen.

Dies

- sichert permanente Aktualisierung auf den letzten Technologiestand
- spart Kosten,
- vereinfacht Serviceaufgaben und
- erlaubt gemeinschaftliche Managementaufgaben von Server und Storage mittels einheitlicher Managementtools.

3.4.2 Funktionsschema

Mehrere solcher Speicherknoten können zu großen Grids zusammengefügt werden, bilden aber gegenüber der Applikation und dem Management ein einziges Speichersystem, welches über gängige Technologien, wie z. B. FC, FCIP, FCoE, NFS, CIFS und iSCSI, angesprochen wird. Jegliche Ressourcen können dabei als Pool zu einem Logischen Volumen zusammengefasst werden (Speicherkapazität, Performance und Bandbreiten werden dementsprechend aggregiert). (Abb. 12)

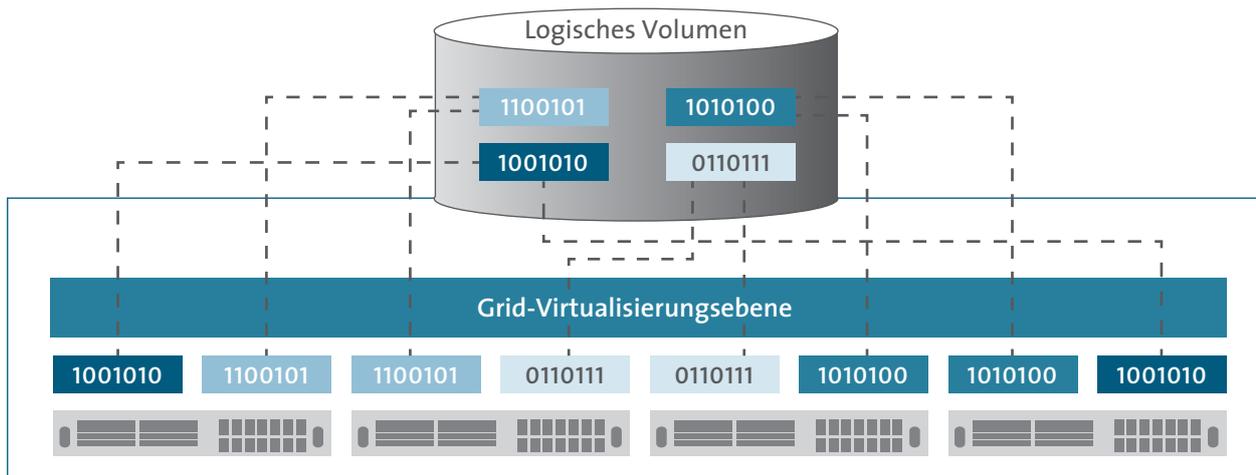


Abbildung 12: Funktionsschema

3.4.3 Skalierbarkeit

Durch Hinzufügen weiterer Speicherknoten lassen sich sowohl zusätzliche Performance, als auch zusätzliche Speicherkapazität dem Grid hinzufügen. Die Migration auf eine neue Hardware kann bei diesem Konzept einfach durchgeführt werden, da das einfache und schnelle Verschieben von virtuellen Ressourcen im Online Betrieb ein reibungsloses Abschalten alter Knoten erlaubt. Migrationsaufwände fallen hier nicht an. Diese Vorgänge können im laufenden Betrieb ohne Störung der Applikationen erfolgen. Die Speicherfunktionalität reorganisiert automatisch das Grid und die Datenblöcke auf alle zur Verfügung stehenden Festplatten, so dass immer die höchst mögliche Performance gewährleistet wird.

3.4.4 Management

Ein zentralisiertes Management erlaubt die einfache Verwaltung des Ressourcenpools. Die Komplexität des Zusammenspiels der Einzelkomponenten wird dabei automatisiert und systemintern gehandhabt. Dies erlaubt, dass bislang komplexe Aufgabenstellungen wie Clustering, Replikationen, applikationsintegrierte Snapshots, Fail-over Handhabung u.a. auch von technisch weniger versierten Administratoren bewältigt werden können.

3.4.5 Vorteile

- **Standardisierung:**
dadurch gleiche HW-Komponenten, Servicestrukturen, Managementtools, Handhabung.
- **Pay-as-you-Grow-Ansatz:**
Durch kapazitätsfreies Thin Provisioning und unterbrechungsfreie Erweiterungsmöglichkeiten kann zusätzliche Kapazität und Performance auch ohne große Bevorratung kurzfristig durch einfaches Hinzufügen weiterer Speicherknoten oder weiterer Storage Kapazität ermöglicht werden.
- **Höchst-Verfügbarkeit:**
Die einzelnen Speicherknoten könnten redundante, hot-plug-fähige Komponenten enthalten. Da die Speicherknoten wiederum selbst aus Sicht des Grids redundant und hot-plug-fähig sind, ergibt sich eine sehr hohe Ausfallsicherheit der Gesamtlösung.
- **Kombination von virtuellen und physischen SAN:**
In bestimmten Anwendungsszenarien (bspw. Außenstandortanbindung) bestehen in den kleinen Standorten die funktionalen Anforderungen an ein SAN, jedoch ist die preisliche Komponente ein Hinderungsgrund für die Umsetzung. Durch die Kombination von Lösungen mit direkt angeschlossenen günstigen Festplatten und einer Replikation auf eine größere Lösung in der Zentrale können gemeinsame Ressourcen Pools gebildet und zentral administriert werden. Dies kann z. B. für eine Backupkonsolidierung verwendet werden.

■ 3.5 Speichervirtualisierung auf der Server-Ebene

Die Notwendigkeit, sich von den engen physikalischen Grenzen von Plattenlaufwerken zu lösen war auf der Betriebssystem-Ebene schon von jeher eine der dringlichsten Anforderungen. Denn einerseits werden Daten-Pools (Dateisysteme) benötigt, deren Kapazität ein Vielfaches einer Festplatte beträgt, andererseits werden kleine dedizierte Datenbereiche gebraucht, für die man keine komplette Platte vorsehen möchte. Die Partitions-Techniken von Festplatten- und Volumen-Manager stellen die technische Implementierung dar und sind seit Jahrzehnten etablierte und erprobte Technologien. Im Bereich der Volume-Manager unterscheidet man:

- Implementierungen, die integraler Bestandteil des Betriebssystems sind
- Softwarelösungen, die auf verschiedenen Betriebssystemen unterstützt werden

Datenbanksysteme und Cluster-Dateisysteme bringen heute vielfach einen eigenen und auf die Belange ihrer Architektur abgestimmten Volume-Manager mit. Das Partitionieren von Festplatten kann man heute fast schon als ein Relikt vergangener Tage bezeichnen. Diese Technologie wird uns dennoch im x86 Server-Bereich noch einige Jahre begleiten, soll aber hier nicht weiter vertieft werden.

Heutige Volume-Manager stellen eine vollständige Virtualisierung von Plattenlaufwerken zur Verfügung. Dabei ist es vollkommen unerheblich, welche Plattentechnologie oder Anschlusstechnik verwendet wird. Das Grundkonzept ist relativ trivial: Die Laufwerke werden zu Pools (Disk- oder Volume-Group genannt) zusammengefasst. Aus solch einem riesigen Pool von Blöcken werden »Logical Volumes« beliebiger Größe gebildet. Ein solches Volume verhält sich wie ein reales Plattenlaufwerk und bildet den Container für ein Dateisystem (Abb. 13):

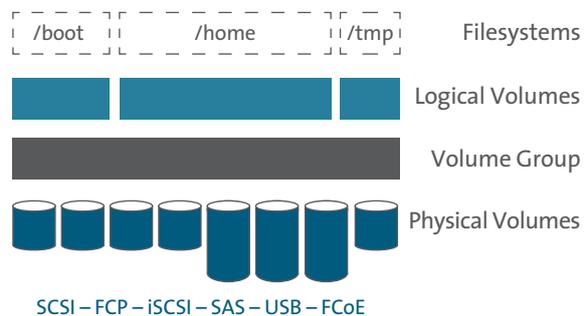


Abbildung 13: Speichervirtualisierung auf Server-Ebene

Neben der flexiblen Bereitstellung von Volumes jeglicher Größe bieten heutige Volume-Manager eine breite Palette zusätzlicher Funktionen:

- Erhöhung der Verfügbarkeit durch RAID-Funktionalitäten: beispielsweise Spiegelung (RAID1) und auch RAID5
- Erhöhung des Durchsatzes durch RAID0-Striping
- Verschlüsselung und Komprimierung
- Erzeugen von Snapshots und Clones
- Synchroner und asynchroner Replikation

Nicht jede Implementierung bietet diesen Gesamtumfang an – auch sind die Funktionalitäten im Detail durchaus unterschiedlich. Dasselbe gilt auch für die Management-Schnittstelle(n). Um den Aufwand für Installation und Betrieb zu minimieren, bietet sich in stark heterogenen Betriebssystem-Umgebungen der Einsatz einer Plattform-unabhängigen Lösung an – auch wenn damit eine komplexere Support-Struktur entsteht und zusätzliche Abhängigkeiten im Bereich der Software-Versionen generiert werden.

Volume-Manager sind in einem heutigen Server-Betrieb – speziell bei großen Umgebungen – nicht mehr wegzudenken. Oftmals werden aus dem Spektrum der Funktionen nur die Basis-Funktionen genutzt, da man z. B. für die Replikation der Disaster-Recovery - Architektur eine »unternehmensweite« Lösung angestrebt.

3.5.1 Vorteile der Virtualisierung auf Server-Ebene

Als integraler Bestandteil vieler Betriebssysteme (vor allem im Bereich der Unix-Systeme) bekommt man Virtualisierungsfunktionen quasi kostenneutral geliefert. Neben der Unabhängigkeit von der Art des physikalischen Speichers bringt diese Technologie eine Reihe weiterer Vorteile mit sich:

- Speicherressourcen können optimiert und flexibel genutzt werden
- Langfristig wird der Aufwand für Speicher-Management-Aufgaben geringer
- Daten-Migrationen können einfach und transparent durchgeführt werden

■ 3.6 Tape-Virtualisierung, Virtuelle Tape-Libraries

Unter Tape-Virtualisierung versteht man die Darstellung von Tape-Libraries, Bandlaufwerken und Bändern durch eine Virtuelle Tape-Library (VTL). Eine VTL ist ein Server, der anderen Servern (meist Backup-Servern) Tape-Libraries/ Bandlaufwerke und Bänder zeigt und diese so emuliert, dass angeschlossene Server sie als real erkennen.

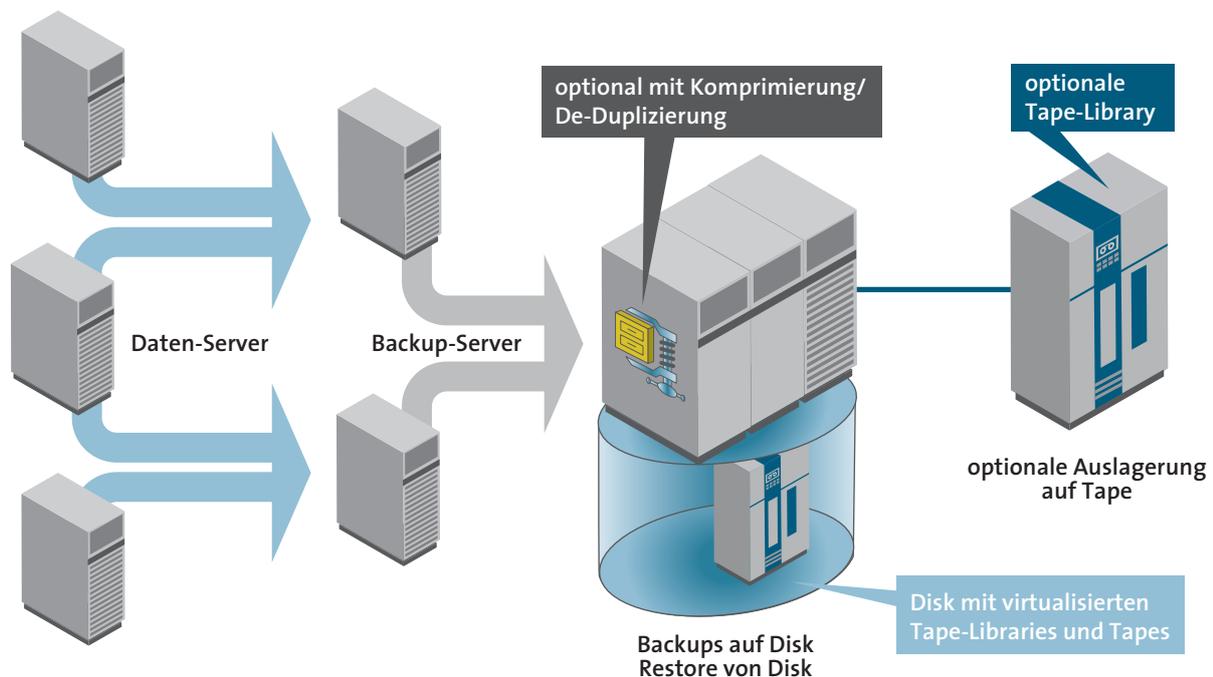


Abbildung 14: Prinzipdarstellung einer VTL

Defacto nutzt eine VTL keine Tapes, sondern Platten zur Abspeicherung der Daten. Realisiert wird auf diese Weise eine Disk-to-Disk-Architektur (D2D). Da VTLs im praktischen Einsatz fast immer für Datensicherungen eingesetzt werden, konzentrieren sich die folgenden Ausführungen auf Backup-Server, obwohl VTLs grundsätzlich für alle Tape-Anwendungen eingesetzt werden können.

3.6.1 Klassische VTL mit Platten-Cache und Tape als Endlager

Diese Architektur existiert seit über zehn Jahren, wurde ursprünglich für Mainframes konzipiert und findet seit einigen Jahren auch außerhalb der Großrechnerwelt eine weite Verbreitung. Sie setzt als Endspeichermedium auf Tape und nutzt Platten als Cache. In Abbildung 14 wird dieser Endspeicherort durch die »Optionale Tape-Library« dargestellt. Die Auslagerung auf Tape erfolgt im Allgemeinen zeitnah, so dass die Datensicherung sich sowohl auf Platte als auch im Cache befindet. Man redet im Zusammenhang mit klassischen VTLs von Disk-to-Disk-to-Tape – Lösungen (D2D2T).

Backup-Server sehen weiterhin nur Tape-Libraries bzw. deren Laufwerke und Tapes. Die End-Speicherung auf Tape ist für die Backup-Server transparent und ist Aufgabe der VTL. Der Shoe-Shining-Effekt kommt bei klassischen VTLs nicht zum Tragen, weil die VTL nicht auf die Liefergeschwindigkeit der Backup-Server angewiesen ist, sondern die Tapes mit großen optimierten Images beschreiben kann.

Dank deutlich gesunkener Platten-Preise kann der Cache immer größer ausgelegt werden, sodass in den allermeisten Fällen von Platte rückgesichert werden kann. Durch Einsatz von Komprimierung und auch De-Duplikation wird dieser Effekt weiter verstärkt.

Die reale Band-Bibliothek kann an einen anderen Ort/ Brandabschnitt verlegt werden und dient als Katastrophen-Absicherung. Klassische VTLs sind darüber hinaus in der Lage, Sicherungs-Duplikate auf einer weiteren Tape-Library zu erzeugen. Auch sind synchrone oder asynchrone Spiegel der Platten-Sicherungen auf räumlich entfernten Platten realisierbar.

Trotz ihrer langen Historie sind klassische VTLs aus genannten Gründen weiter aktuell.

3.6.2 VTL als reine Disk-Lösung, Disk-VTL

Auf die in Abbildung 14 dargestellte »Optionale Tape-Library« wird bei dieser Architektur verzichtet. Wir haben es also mit einer Disk-to-Disk – Lösung zu tun. Reine Disk-VTLs existieren seit einigen Jahren und finden inzwischen dank gefallener Plattenpreise weitere Verbreitung. Man findet sie praktisch immer in Verbindung mit Daten-Deduplizierung. Denn nur so lassen sich die immens gestiegenen Datenumfänge mit D2D bewältigen und riesige Mengen auf im Verhältnis wenig realem Plattenplatz abspeichern. Man redet aktuell von realistisch etwa 90% Platzensparnis. (Klassische VTLs können bei vollem Cache auf Band ausweichen).

Für den Fall, dass ein Anwender auf Tapes nicht vollends verzichten kann, besitzen einige Disk-VTLs die Möglichkeit, Tape-Libraries anzuschließen und diese mit internen »Vault«-Utilities zu bedienen. Wo diese Alternative nicht existiert, wird vorgeschlagen, dass Backup-Server separate Sicherungen auf direkt angeschlossene Tape-Libraries schreiben.

4 Datendienste / Service Level Agreements

Die Speicherinfrastruktur stellt der Unternehmens-IT verschiedene Dienste zur Verfügung, die diese bei der Erfüllung ihrer Aufgaben unterstützt. Es handelt sich im Folgenden also nur um virtualisierungsrelevante, IT-interne Dienste und nicht um die von Kunden wahrgenommene Primärfunktion des zuverlässigen Speicherns von Daten, der ordnungsgemäßen Datensicherung/Archivierung und des Wiederherstellens nach einem Fehlerfall.

Folgende Dienste erbringt die Speicherinfrastruktur:

- **Datenmobilität (transparent)**
Daten können innerhalb einer Speicherlandschaft ohne Auswirkung auf die betroffenen Applikation verschoben werden. Produktfamilien-, hersteller- und speichersystemübergreifend ist diese Funktionalität i.d.R. nur mittels Virtualisierungstechnologien zu realisieren.
- **Desaster tolerante Datenhaltung**
Daten werden an unterschiedlichen Orten mit Hilfe von Spiegelungs- oder Replikationsverfahren konsistent gespeichert. Produktfamilien-, herstellerübergreifend ist diese Funktionalität i.d.R. nur mittels Virtualisierungstechnologien zu realisieren.
- **Transparenter Failover in eine zweite Lokation**
Es kann ohne Auswirkung auf die Applikation auf die Datenkopien bzw. -spiegel in einer zweiten Lokation zugegriffen werden, wenn der Zugriff auf die Daten in der ersten Lokation nicht mehr möglich ist. Produktfamilien-, hersteller- und speichersystemübergreifend ist diese Funktionalität i.d.R. nur mittels Virtualisierungstechnologien zu realisieren.
- **Performancemanagement (Quality of Service)**
Servern wird eine definierte Schreib-/Lese-Performance zugesichert bzw. zugeteilt. Um diese Service darzustellen, werden mehrere Komponenten der Gesamtlösung involviert.
- **Bereitstellung von Datenkopien**
Externen Diensten (z. B. Datensicherung) werden Datenkopien zur Verfügung gestellt. Hier bestehen die Möglichkeiten, das mittels vollständiger Systemkopien oder volumenoptimierter Snapshots zu erreichen. Produktfamilien-, hersteller- und speichersystemübergreifend ist diese Funktionalität i.d.R. nur mittels Virtualisierungstechnologien zu realisieren.
- **Hierarchisches Speicher-Management**
Daten werden in Abhängigkeit vom zeitliche Wert der Daten auf Speicher der entsprechenden Güte gespeichert. Produktfamilien-, hersteller- und speichersystemübergreifend ist diese Funktionalität i.d.R. nur mittels Virtualisierungstechnologien zu realisieren.

5 Vorteile der Speichervirtualisierung

Jede der bisher vorgestellten Virtualisierungstechnologien bietet individuelle Vorteile, die sie für den Einsatz in unterschiedlichen IT-Umgebungen prädestinieren. Allen gemeinsam ist: Speichervirtualisierung ermöglicht eine deutliche Kostenreduktion.

Diese Einsparungen können in verschiedenen Bereichen entstehen, die im Folgenden konkreter erläutert werden.

■ 5.1 Optimierte Nutzung der Speicherressourcen

Bei einer Virtualisierungslösung wird die Kapazität von Speichersystemen gleicher oder verschiedener Anbieter in einem großen Pool zusammengeführt und Teile davon zentral und flexibel den Hostanwendungen zur Verfügung gestellt. Einzelne Bereiche (z. B. Volumes) aus diesem Speicher-Pool können gezielt und bedarfsabhängig bestimmten Anwendungen zugewiesen werden. Größe und Struktur der »virtualisierten« Volumes sind unabhängig von den angeschlossenen Speichersystemen und werden wie herkömmliche Volumes eines Speichersystems behandelt. Die Skalierbarkeit des Gesamtsystems wird wesentlich größer, da sie nicht an die physikalischen Grenzen eines einzelnen Speichersystems gebunden ist.

Durch Speichervirtualisierung wird der vorhandene Speicherplatz sehr viel besser ausgenutzt, wodurch die Kosten stark verringert werden.

■ 5.2 Reduzierung des Verwaltungsaufwands

Die Bündelung der Kapazität mehrerer heterogener Plattenspeichersysteme zu einem einzigen Speicherreservoir erlaubt die Verwaltung von einem zentralen Punkt aus. Die Anzahl der zu verwaltenden Elemente verringert sich, wodurch sich die Komplexität reduziert und der Administrationsaufwand nachhaltig verkleinert wird. Die Abhängigkeit zwischen Servern und Speichereinheiten durch spezielle Treiber wird durch die meisten Virtualisierungslösungen aufgehoben.

Eine Reduzierung des Administrationsaufwands führt zu einer Kostenreduzierung. Die gleiche Anzahl von Mitarbeitern kann die immer größer werdenden Datenmengen verwalten.

■ 5.3 Investitionsschutz

Die Einbindung aller vorhandenen Speichersysteme schützt bereits getätigte Investitionen und nutzt die insgesamt zur Verfügung stehende Kapazität optimal. Selbst ältere Systeme können an die jeweilige Virtualisierungsinstanz angebunden werden und auf diese Weise weiter genutzt werden. Die Virtualisierungslösung, die zum Einsatz kommt, muss allerdings ein möglichst breites Spektrum an Storage-Arrays unterschiedlicher Hersteller unterstützen.

■ 5.4 Neue Funktionen für ältere Systeme

Erweiterte Funktionen, wie z. B. die Erstellung von Point-in-Time Kopien, die Migration von Daten von einem auf einen anderen Datenträger, die synchrone Spiegelung oder das asynchrone Kopieren von Daten an einen entfernten Ort werden in einer virtualisierten Speicherumgebung von der Virtualisierungsinstanz übernommen. Damit profitieren alle, d.h. auch ältere Systeme von den Möglichkeiten, die ein virtualisiertes zentrales Datenreservoir bietet. Bei Systemen, deren Wartung bislang nicht unterbrechungsfrei möglich war, können die Daten für die Dauer des Wartungseingriffs auf ein anderes System und danach wieder zurück kopiert werden, ohne dass eine Downtime entsteht. Diese zusätzlichen Funktionen ermöglichen einen flexiblen Einsatz der vorhandenen Hardware und tragen so verstärkt zum Investitionsschutz bei.

■ 5.5 Zukunftsorientierung

In den meisten Fällen können Administratoren über eine einfache graphische Benutzeroberfläche bequem Konfigurationen und andere Services bei Speicher-Volumes, die auf unterschiedlichsten physikalischen Systemen residieren, durchführen. Um die Flexibilität nachhaltig zu gewährleisten, sollte die Virtualisierungslösung auf offenen Technologien basieren und nicht proprietär sein.

Ein Unternehmen behält die Freiheit, Kapazitäten unterschiedlicher Anbieter im laufenden Betrieb in den Pool einzubinden, ohne organisatorische Änderungen vornehmen zu müssen. Diese Kapazitäten profitieren von der Intelligenz der virtualisierenden Instanz ohne die Notwendigkeit, Software-Lizenzen für die einzelne Systeme erwerben zu müssen.

■ 5.6 Unterbrechungsfreie Datenmigration

Für Unternehmen, die häufige Datenmigrationen durchführen, bietet die Speichervirtualisierung eine effektive Lösung: Sie ermöglicht den dynamischen Transfer der Daten innerhalb des SAN von einem System zum anderen ohne Unterbrechungen der Anwendung.

Virtualisierung hilft nicht nur bei Datenverschiebungen sondern auch bei »Performance-Transfers«. Daten können für eine bestimmte Zeit auf ein leistungsstärkeres Speichersystem verschoben werden, um die aktuellen Geschäftsanforderungen zu bedienen. Das Erhöht die Verfügbarkeit der Unternehmensdaten. Bei Systemen, deren Wartung bislang nicht unterbrechungsfrei war einige Midrange-Systeme, können die Daten für die Dauer eines Wartungseingriffs auf ein anderes System und wieder zurück kopiert werden, ohne dass eine Downtime entsteht.

Unter Berücksichtigung der Tatsache, dass Speichersysteme in der Regel in Zeiträumen zwischen drei und fünf Jahre erneuert werden und die Daten für die Migrationszeit zu den neuen Systemen, die Monate dauern kann, doppelt gehalten werden, führt diese Funktionalität zu deutlichen Kostenreduzierung im Unternehmen.

■ 5.7 Kopierservices und »Business Continuity«

Virtualisierungslösungen ermöglichen zentral die Erstellung heterogener Point-in-Time-Kopien von Daten. Es ist möglich, herstellerunabhängig die Daten von jedem und auf jedes Speichersystem zu kopieren. Auf diese Weise können lokale Kopien über die Grenzen eines einzelnen Systems hinweg erstellt werden und auf einem preisgünstigeren Medium gehalten werden.

Ein großer Vorteil, den die Virtualisierungslösung für die Business Continuity bietet, besteht darin, Daten synchron oder asynchron an geografisch entfernte Standorte zu kopieren. Diese Funktionalität wird für alle Daten des virtualisierten Pools geboten, gleichgültig ob das angeschlossene Speichersystem diese Kopierfunktion »native« unterstützt hätte oder nicht.

■ 5.8 Vereinfachung des Information Life Cycle Management (ILM)

Virtualisierung bildet ein wichtiges Fundament für Information Life Cycle Management. Unternehmen können den Datentransfer über unterschiedliche Speicherklassen von heterogenen Speichersystemen unterbrechungsfrei und schnell umsetzen. Durch Virtualisierung können anwendungsbezogene Bereiche – im Gegensatz zu physisch begrenzten Bereichen – definiert werden, um eine optimale Korrelation zwischen dem Wert der Daten und den entsprechenden Speicherkosten zu bilden.

■ 5.9 Optimierte Abstimmung zwischen Speicherservice und Anwendungsanforderung

Um in einer globalisierten und komplexen Welt wettbewerbsfähig zu bleiben, müssen Unternehmen ihre IT-Ressourcen sorgfältig auf ihre Geschäftsziele abstimmen. Hier spielt das Fundament – die Speicherlandschaft – eine zentrale Rolle. Moderne Virtualisierungslösungen können durch die Bildung virtueller Speichercluster den Anwendungen die Speicherqualität und -quantität liefern, die sie für einen optimalen Betrieb benötigen. Diese virtuellen Partitionen können einer speziellen Anwendung je nach Bedarf die gewünschte Kapazität, den erforderlichen Service (Antwortzeiten) und selbst die im speziellen Fall benötigte Konnektivität variabel zur Verfügung stellen.

■ 5.10 Optimierung der Gesamtspeicherkosten

Eine Speicherlösung mit Virtualisierungsfunktion:

- reduziert die Anzahl der Elemente, die verwaltet werden müssen.
- Einsparungen bei Softwarelizenzen und Wartungsleistungen.
- Verringerung des Administrationsaufwands.
- ermöglicht die Zuordnung von freigegebenen Ressourcen für andere Projekte.
- vereinfacht die Inbetriebnahme (deployment) neuer Server.
- erleichtert die Speicherung von Daten mehrerer Klassen, eine so genannte »Tiered Storage« -Lösung. Somit können Unternehmensinformationen in jeder Phase ihres Lebenszyklus optimal und kostengünstig verwaltet werden.
- steigert die MAK-Effizienz und erhöht die Verfügbarkeit der Unternehmensdaten

Dies führt zu einer signifikanten Reduktion der Gesamtkosten für Speicher.

■ 5.11 Fazit

Die beschriebenen Effekte werden bei jedem Unternehmen zu unterschiedlich hohen Einsparungen führen. Die Unternehmensgröße ist also nicht entscheidend bei der prinzipiellen Überlegung, eine Virtualisierungslösung einzusetzen. Vielmehr bedarf es einer genauen Analyse der vorhandenen Struktur und der anstehenden Herausforderungen. Auch zukünftige Datenspeicher- und Datenmanagement-Aufgaben sollten in die Überlegungen mit einbezogen werden, um auf sich ändernde Geschäftsanforderungen flexibel reagieren.

Danksagung

Diese Veröffentlichung ist im AK Speichertechnologie des BITKOM entstanden. Besonderer Dank gilt den Autoren:

- Guido Klenner, Hewlett-Packard GmbH
- Georg Mey, NetApp Deutschland GmbH
- Dr. Georgios Rimikis, Hitachi Data Systems GmbH
- Dr. Dietrich Schaupp, IBM Deutschland GmbH
- Horst-Wilhelm Stahl, Bull GmbH

Der Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. vertritt mehr als 1.700 Unternehmen, davon über 1.200 Direktmitglieder mit etwa 135 Milliarden Euro Umsatz und 700.000 Beschäftigten. Hierzu gehören fast alle Global Player sowie 800 leistungsstarke Mittelständler und zahlreiche gründergeführte, kreative Unternehmen. Mitglieder sind Anbieter von Software und IT-Services, Telekommunikations- und Internetdiensten, Hersteller von Hardware und Consumer Electronics sowie Unternehmen der digitalen Medien und der Netzwirtschaft. Der BITKOM setzt sich insbesondere für eine Modernisierung des Bildungssystems, eine innovative Wirtschaftspolitik und eine zukunftsorientierte Netzpolitik ein.



Bundesverband Informationswirtschaft,
Telekommunikation und neue Medien e.V.

Albrechtstraße 10 A
10117 Berlin-Mitte
Tel.: 030.27576-0
Fax: 030.27576-400
bitkom@bitkom.org
www.bitkom.org