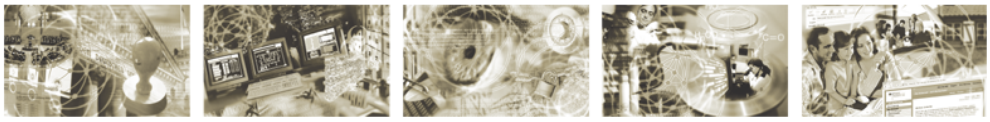
Bundesamt
für Sicherheit in der
Informationstechnik



Band B, Kapitel 6: Speichertechnologien

Bundesamt für Sicherheit in der Informationstechnik
Postfach 20 03 63
53133 Bonn
Tel.: +49 22899 9582-0
E-Mail: hochverfügbarkeit@bsi.bund.de
Internet: https://www.bsi.bund.de

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

1 Speichertechnologien und -verfahren

Schnell wachsende Datenmengen und hohe Anforderungen an die Verfügbarkeit stellen die IT-Verantwortlichen häufig vor die Aufgabe sichere und skalierbare Lösungen für die Speichersysteme zu finden. Dabei stellt sich die Frage, welche Lösungsansätze sind aktuell verfügbar und welche Sicherheits- und Leistungsmerkmale haben diese. Wie auch in den anderen Beiträgen des HV-Kompendiums, werden hier die beschriebenen Konzepte hinsichtlich ihrer Eignung für das HV-Umfeld untersucht.

Die Speicherlösungen werden in diesem Abschnitt entsprechend der nachfolgenden Einteilung (Bottom-Up-Ansatz) vorgestellt:

- Festplatten und Schnittstellen
- Verbindungs-/ Kopplungsprinzipien
- Speicherarchitekturen
- Topologien
- Arbitrationsverfahren
- Speichervirtualisierung
- Datensicherungs- und Wiederherstellungsverfahren
- Dateisysteme
- Speicher-Management

1.1 Festplatten und Schnittstellen

Die Festplatte ist eine Komponente auf der untersten Architekturebene einer Speicherlösung. Bereits auf dieser Ebene können HV-Maßnahmen wirkungsvoll umgesetzt werden. Um Festplatten in Festplattensubsystemen zu verwalten und eine gewisse Fehlertoleranz zu erreichen, werden RAID-Systeme eingesetzt. Die unterschiedlichen Festplattensysteme sind insbesondere durch die eingesetzten Schnittstellentechnologien geprägt. Diese Aspekte werden in den folgenden Abschnitten erläutert.

1.1.1 Festplattensubsysteme

Festplattensubsysteme bestehen aus mehreren einzelnen Festplatten und sind auf Modularität, Skalierbarkeit und Verfügbarkeit optimiert. Sie werden mit unterschiedlichen Gesamtkapazitäten (1 TB bis >300 TB) und Leistungsfähigkeiten auf dem Markt angeboten. Die Festplattensubsysteme besitzen integrierte RAID-Controller, die auf der Backend-Seite (Festplattenseite) entweder ATA-, SCSI- oder FC / AL-Technologie (Fibre Channel Arbitrated Loop) unterstützen. Preiswerte Geräte verfügen lediglich über ein bis zwei Festplattenkanäle. Mittelklasseprodukte und Hochleistungssysteme sind dahingehend umfangreicher ausgestattet und es ist somit möglich, die Festplatten physikalisch mehrpfadig anzuschließen. Der mehrpfadige Anschluss von Festplatten erhöht die Verfügbarkeit des RAID-Verbundes, da so neben der Redundanz der Knoten (Festplatten) auch eine Redundanz der Pfade (Datenverbindungen) realisiert wird. Am Frontend eines Speichersystems befinden sich, über Switches angesteuerte Anschlüsse für Rechnersysteme und / oder Netzkompo-

nenten. In der höheren Preisklasse der Festplattensubsysteme sind i. d. R. Verbindungen (ATA, SCSI, FC) möglich. Auch bei preiswerten Geräten sind Systemkomponenten, wie Stromversorgung, Lüfter und Controller-Baugruppen, redundant ausgeführt.

1.1.2 Festplattenmedien

Festplatten unterscheiden sich nicht alleine durch die Schnittstellentechnologie und den Preis. Vielmehr sind Unterschiede bei der Güte der verwendeten Materialien, der Mechanik und der Elektronik für die Preisdifferenzen verantwortlich. Neben den eher für den Consumer-Bereich und für Desktop-Systeme vorgesehen Geräteklassen sollten im Enterprise- bzw. HV-Umfeld nur Festplatten zum Einsatz kommen, die den Anforderungen an einen dauerhaften und zuverlässigen Server-Betrieb entsprechen. In Tabelle 1 sind die zurzeit gebräuchlichsten Festplatten-Schnittstellentechnologien mit ihren wesentlichen Eigenschaften in einer Übersicht aufgeführt.

	ATA (Advanced Technology Attachment)	***SATA (Serial ATA) 2. Generation***	***SCSI (Small Computer System Interface) Ultra-320***	***SAS (Serial Attached SCSI)***	***Fibre Channel (FC)***
Übertragungs-geschwindigkeit (MB/s)	133	300	320	300	500
Datenbusbreite (Bit)	16	seriell	16	seriell	seriell
Max. Kabellänge (m) des Datenbusses	0,5	1	12	8	5 (Kupfer) 5000 (Glas)
Max. Anzahl an Festplatten am Bussystem	2	4	16	128	Arbitrated Loop: 126 Knoten Switched Fabric: > 1 Million Knoten
Kategorien	Consumer	Consumer/ Enterprise	Enterprise	Enterprise	Enterprise

Tabelle 1: Exemplarische Leistungsdaten der Schnittstellentechnologien bei Festplatten

Die in der Tabelle enthaltenen typischen Leistungsdaten und Eigenschaften dienen der exemplarischen Übersicht. Die tatsächlichen Produkteigenschaften einzelner Festplatten können hiervon abweichen.

1.1.3 Advanced Technology Attachment (ATA)

ATA (Advanced Technology Attachment) wurde als Industriestandard von ANSI für parallele Schnittstellen von Festplatten- und CD-Laufwerken definiert. In dem Standard wird der Anschluss an ein Bus-System und das Signalverhalten spezifiziert.

Eine Weiterentwicklung des Standards stellt SATA (Serial-ATA) dar. Im Gegensatz zum Bus-System bei ATA ist SATA sternförmig, also mit jeweils einem dedizierten Kabel vom Controller zu jeder Festplatte verkabelt. Darüber hinaus ist es möglich, Laufwerke während des Betriebes an- und abzukoppeln (hot plugging). Die Vorteile dieser Technologie sind eine einfachere Verkabelung und höhere Übertragungsleistungen.

1.1.4 Small Computer System Interface (SCSI)

SCSI ist eine (ANSI-) standardisierte, parallele Schnittstelle, welche die Datenübertragung zwischen Geräten auf einem Bus-System ermöglicht. Die SCSI-Technologie wurde kontinuierlich weiterentwickelt. Die verschiedenen SCSI-Versionen unterscheiden sich deutlich hinsichtlich der Leistungsmerkmale und der Anschlusstechnik.

Eine Weiterentwicklung des Standards ist SAS (Serial Attached SCSI). Mit SAS ist eine Höchstleistungs-Plattenanbindung möglich. Die Serialisierung wurde gewählt, da sich beim Standard-SCSI die Synchronisation der einzelnen Datenströme immer komplexer gestaltete. Die Vorteile von SAS gegenüber dem herkömmlichen SCSI liegen in der wesentlich preiswerteren Verkabelung, der höheren Skalierbarkeit, den deutlich gesteigerten Übertragungsleistungen und der nativen Unterstützung von SATA- und SCSI-Geräten.

Eine weitere Variante von SCSI ist iSCSI (Internet Small Computer System Interface). Dies bezeichnet ein standardisiertes Verfahren in dem SCSI-Daten in TCP/IP-Pakete verpackt und über IP-Netze transportiert werden. Die verpackten SCSI-Kommandos gelangen so zu einem SCSI-Router, der auf Basis vorhandener Mapping-Tabellen das entsprechende Zielsystem zur Kommunikation mit der SCSI-Datenquelle auswählt. iSCSI wird eingesetzt, um über eine virtuelle Ende-zu-Ende-Verbindung den Zugriff auf das Speichernetz zu ermöglichen, ohne dass eigene Speichergeräte aufgestellt werden müssen. Vorhandene Netzwerkkomponenten (Switch) können genutzt werden, da keine neue Hardware für die Knotenverbindungen nötig ist.

1.1.5 Redundant Array of Independent Disks (RAID)

Ein RAID (Redundant Array of Independent Disks) ist eine Anzahl gewöhnlicher Festplatten, die im Verbund wie ein einziges großes Laufwerk erscheinen. Durch ein RAID kann sowohl eine Erhöhung der Datensicherheit durch Redundanz, als auch eine Steigerung der Qualitätsmerkmale, wie Transferrate oder Latenz, durch gleichzeitige Ansteuerung von mehreren Festplatten erreicht werden. RAID-Konzepte sind die Grundvoraussetzung, um Ausfälle einzelner Festplatten ohne Datenverlust und Ausfallzeiten zu überstehen, ohne dass kritische Systemzustände eintreten. Hierbei kommen mehrere Redundanzverfahren (siehe Beitrag Prinzipien der Verfügbarkeit) zum Einsatz. Neben der strukturellen Redundanz, aufgrund der mehrfach vorhandenen Festplatten, werden funktionelle Redundanz sowie Informationsredundanz eingesetzt, um eine (nach außen) fehlertolerante Komponente als Gesamtsystem zu realisieren. Es wird zwischen mehreren RAID-Level/-Konzepten unterschieden, die vom RAID-Advisory-Board, einem Industrie-Konsortium, definiert wurden und die oben beschriebenen Redundanzprinzipien in unterschiedlicher Ausprägung nutzen:

RAID 0

Dieser Type von RAID (mindestens 2 Festplatten) steht für Striping (Stripe-Set). Striping bedeutet das Verteilen logisch zusammenhängender Daten auf mehrere Festplatten. Bei diesem Verfahren werden also mehrere Festplatten zu einem großen logischen Laufwerk zusammengefasst. Der Striping-Faktor sagt dabei aus, wie groß die Stücke sind, welche jeweils auf eine der Festplatten geschrieben werden. Der Vorteil von RAID 0 liegt in der Belegung der gesamten Kapazität mit produktiven Daten. Vorteilhaft sind weiterhin der parallele Zugriff auf mehrere Kanäle und der damit realisierte multiple Datendurchsatz. Dieser wird allerdings nur erzielt, wenn seriell gelesen oder geschrieben wird. Fällt eine Festplatte aus dem Verbund aus, so kann auf den gesamten Datenbestand des RAID-Systems nicht mehr zugegriffen werden. RAID 0 realisiert zwar eine

strukturelle Redundanz, aber keine Informationsredundanz, ist daher zur Realisierung hoch verfügbarer Systeme nur sehr eingeschränkt geeignet.

RAID 1

RAID 1 (mindestens 2 Festplatten) realisiert eine Spiegelung (engl. Mirroring, Duplexing) der Informationen. Hier werden alle Daten einer Festplatte auf eine zweite gespiegelt. Der Vorteil der Festplattenspiegelung besteht darin, dass bei einem Ausfall einer Festplatte auf die Daten weiterhin zugegriffen werden kann. Die Leistung beim Schreiben bleibt im Vergleich zu einer einzelnen Festplatte gleich, wenn man davon ausgeht, dass der doppelte Schreibvorgang den Rechner nicht zusätzlich belastet. Die Leseleistung verdoppelt sich im besten Fall, da jetzt die Lesezugriffe auf zwei Festplatten aufgeteilt werden. Der Nachteil von RAID 1 sind die Kosten. Nur die Hälfte der verfügbaren Festplatten steht für Originaldaten zur Verfügung, die andere Hälfte wird für die redundanten Daten benötigt. Diese Strukturelle Redundanz wird zur Umsetzung der Informationsredundanz verwendet.

RAID 2

RAID 2 (mindestens 3 Festplatten) ergänzt RAID 0 (Stripe-Set) um mehrere ECC-Festplatten. Dabei steht ECC für Error Correction Code. Die Originaldaten werden bitweise auf die Festplatten des Stripe-Sets aufgeteilt. Zusätzlich werden ECC-Bits nach dem Hamming-Algorithmus auf die einzelnen Festplatten geschrieben. Die dadurch erreichte Informationsredundanz macht es möglich, bei jedem Lesezugriff die Daten auf Integrität zu überprüfen und nötigenfalls sofort zu korrigieren. Da das Schreiben der Daten und des ECC über separate Kanäle auf getrennte Festplatten geschieht, sind dadurch hohe Transferraten möglich. Allerdings lassen sich diese nur erreichen, wenn die einzelnen Festplatten synchronisiert werden, was mit einem hohen technischen Aufwand und Kosten verbunden ist.

Vorteil dieser Lösung ist, dass auch bei gleichzeitigem Ausfall von zwei Festplatten die Daten nicht verloren gehen. Aufgrund der aufwändigen Implementierung blieb der Einsatz von RAID 2 in der Vergangenheit auf einige wenige Mainframe-Installationen beschränkt.

RAID 3

RAID 3 benötigt mindestens drei Festplatten und stellt eine vereinfachte Implementierung von RAID 2 dar, bei der die ECC-Festplatten durch eine einzelne Festplatte mit Kontrollinformationen (Parity) ersetzt werden. Die Parity-Bits (Informationsredundanz) werden bei jedem Schreibvorgang erzeugt und beim Lesen überprüft (funktionale Redundanz). Zusätzlich zu den Parity-Informationen wird ein Index erstellt, der ebenfalls auf die Parity-Festplatte geschrieben wird. Mit einem Rechenalgorithmus (XOR-Verknüpfung) ist es möglich, beim Ausfall einer Festplatte die fehlenden Daten zusammen mit der Prüfsumme zu rekonstruieren. Um das Erstellen der Kontrollinformationen zu erleichtern, synchronisiert RAID 3 die Kopfpositionen der Festplatten. Dies minimiert den Overhead der Schreibzugriffe, da sowohl Daten als auch Parity-Informationen parallel gespeichert werden. Arbeiten die Anwender mit vielen kleinen und möglicherweise verteilten Datenblöcken, wirkt sich dies negativ auf die Geschwindigkeit aus, da ein häufiges Neusynchronisieren viel Zeit in Anspruch nimmt. Einen weiteren Nachteil kann die Parity-Festplatte darstellen, die bezüglich der Performanz zu einem „Flaschenhals“ werden kann, da sie durch die Erzeugung der Kontrollinformationen an allen Schreiboperationen beteiligt ist. Darüber hinaus ist sie dadurch einer höheren Last als die anderen Festplatten ausgesetzt und kann so die Verfügbarkeit des Gesamtsystems als SPoF beeinträchtigen.

RAID 4

RAID 4 (mindestens 3 Festplatten) entspricht RAID 3, nur mit einem Striping-Faktor von einem Block (oder auch mehr). Auch wird auf das Synchronisieren der Festplattenköpfe verzichtet, um die Nachteile von RAID 3 bei der Verarbeitung kleiner Dateien zu umgehen. Die Kontrollinformationen werden auf eine separate Festplatte geschrieben. Speziell beim Schreiben kleiner Dateigrößen entpuppt sich die Parity-Festplatte für die Performanz als Flaschenhals. Bei jeder Schreiboperation muss zunächst die Checksumme errechnet, die Parity-Information auf der Festplatte gefunden und angesteuert werden. Vorteile bietet RAID 4 in Umgebungen, in denen vor allem Lesezugriffe anfallen. In der Praxis findet dieses Verfahren aber kaum Anwendung.

RAID 5

RAID 5 (mindestens 3 Festplatten) verteilt die Parity-Informationen sowie die Daten in Form von Blöcken auf alle Festplatten. Jede Festplatte ist damit für einen bestimmten Block „Parity-Festplatte“. Leseoperationen werden etwas schneller, weil sie auf noch mehr Festplatten verteilt werden können, Schreiboperationen werden etwas schneller, weil der Flaschenhals Parity-Festplatte wegfällt. Darüber hinaus wird auch die Verfügbarkeit des Gesamtsystems nicht beeinträchtigt, da die zusätzliche Belastung durch die Parity-Zugriffe auf alle Platten verteilt wird. Bei kurzen Transfers ist aber auch RAID 5 auf der Basis von Zugriffen pro Zeiteinheit pro Megabyte einer einzelnen Festplatte unterlegen. RAID 5 bietet ein sehr gutes Preis-Leistungs-Verhältnis, da nur circa ein Drittel der vorhandenen Speicherkapazität für die Redundanz aufgebracht werden muss.

RAID 6

RAID 6 stellt eine Erweiterung des RAID 5 dar. Während in einem RAID 5 lediglich ein Satz an Kontrollinformationen abgelegt wird, sind dies im RAID 6 derer zwei. Dies kompensiert den Ausfall von bis zu zwei Laufwerken. Versagen bei RAID 5 gleichzeitig zwei Disk, sind die Daten hier nicht mehr zu rekonstruieren. Das Schreiben zweier Kontrollinformationssätze beugt diesem Fall vor.

Die nachfolgende Tabelle fasst die wesentlichen Eigenschaften der einzelnen RAID-Prinzipien zusammen.

Level	*Redundanz-verfahren*	*Performanz*	*Ausfallsicherheit*	*Kosten*
RAID 0	nur strukturelle, keine Informations-redundanz	gute Performanz	sehr gering	Kosten für zusätzliche Platten ohne Nutzen
RAID 1	Strukturelle und Informations-redundanz	neutrale Schreib-Performanz, gute Performanz beim Lesen	gering, da nur einfache Redundanz	Verdopplung der Kosten durch Spiegelung
RAID 2	Strukturelle, funktionale und Informations-redundanz	gute Performanz bei synchronisierten Festplatten	hohe Ausfall-sicherheit da von 3 Festplatten bis zu 2 ausfallen können	hoher technischer Aufwand und Kosten
RAID 3	Strukturelle, funktionale und Informations-redundanz	schlechte Performanz	geringe Ausfallsicherheit	günstiges Preis-/ Leistungsverhältnis
RAID 4	Strukturelle, funktionale und Informations-redundanz	geringe Performanz für Schreiboperatio-nen, hohe Perfor-manz beim Lesen	gering, da SPoF Parity-Festplatte	günstiges Preis-/ Leistungsverhältnis
RAID 5	Strukturelle, funktionale und Informations-redundanz	gute Performanz	gute Ausfallsicherheit bei Ausfall von einer Festplatte	günstiges Preis-/ Leistungsverhältnis
RAID 6	Strukturelle, funktionale und Informations-redundanz	gute Performanz	hohe Ausfall-sicherheit, da von 3 Festplatten bis zu 2 ausfallen können	gutes Preis-/ Leistungsverhältnis

Tabelle 2: RAID-Level im Vergleich

Die Funktionalität der einzelnen RAID-Level, insbesondere die zu realisierende funktionale Redundanz der Parity-Algorithmen, kann auf zwei Arten realisiert werden: Entweder kostengünstig über das Betriebssystem als Software-RAID oder mit Hilfe eines RAID-Controllers als Hardware-RAID. RAID-Controller werden für alle gängigen Festplattentechnologien angeboten. Die Vorteile eines Hardware-RAID liegen im Datendurchsatz, in der Unterstützung aller gängigen RAID-Level und der Unterstützung von Hot-Plug- und Hot-Spare-Festplatten. Bei Hot-Spare-Festplatten handelt es sich um Festplatten, die in einem Hot-Stand-By-Betrieb laufen und sofort vom RAID-Controller aktiviert werden, wenn der Ausfall einer Festplatte erkannt wurde. Der Nachteil von RAID-Controllern sind die je nach Hersteller hohen Anschaffungskosten.

Neben den oben genannten RAID-Level, welche sich quasi als Standard etabliert haben, gibt es noch eine Reihe von herstellerspezifischen RAID-Level (RAIDn, RAID 30/50/51), welche aber in der Regel nur Modifikation der oben genanten RAID-Level darstellen. In der Praxis werden hauptsächlich die RAID-Level 0, 1, 5 und 6 eingesetzt.

1.2 Verbindungs-/Kopplungsprinzipien

Die Verbindung der einzelnen Speicherkomponenten lässt sich mittels unterschiedlicher Prinzipien und Technologien realisieren. In diesem Abschnitt werden die Funktionsweise von einigen Koppelelementen und grundlegenden Verfahren erläutert.

1.2.1 Fibre Channel (FC)

Fibre Channel ist eine serielle Verbindungstechnik, die von ANSI (ANSI X3T11) standardisiert wurde. Der Fibre Channel hat weder eine idealtypische Architektur noch eine netzwerktypische; er kombiniert die Vorteile beider Architekturen in einer einzigen I/O-Schnittstelle miteinander. Die Fibre Channel-Verbindungen sind Punkt-zu-Punkt-Verbindungen zwischen zwei Fibre Channel-Controllern des Servers und des Speichergerätes. Fibre Channel hat sich im High-End-Bereich des Hochgeschwindigkeitstransfers etabliert und ist weniger für den universellen LAN-Einsatz mit wechselnden Verkehrsprofilen geeignet. Ideal ist diese Verbindungstechnik für Speichernetze und wegen der vorhersagbaren Übertragungszeiten auch für Video-Übertragungen.

1.2.2 Fibre Channel Switches

Fibre Chanel-Switches (FC-Switches) bilden das Rückgrat eines SANs. Sie steuern den Datenfluss und sind für die reibungslose Kommunikation innerhalb des SANs verantwortlich. FC-Switches sind auf Grund ihrer Leistungsfähigkeit in der Lage zwischen den angeschlossenen Systemen gleichzeitig mehrere, voneinander unabhängige Verbindungen mit voller Bandbreite, zu schalten. Einsteigersysteme besitzen meistens zwischen 4 und 16 Ports. Enterprise-Switches verfügen über bis zu 32 Ports und erlauben aufgrund ihrer hohen Portzahl, den Aufbau von kaskadierten oder vermaschten Infrastrukturen. Insbesondere für den Aufbau von HV-Architekturen werden vermaschte Strukturen realisiert, da diese eine höhere Ausfallsicherheit bieten. Sie erfordern allerdings eine aufwändige, von den einzelnen Switches abzuwickelnde Übertragungssteuerung, was sich negativ auf Latenzzeiten und Performanz auswirken kann.

1.2.3 Host-Bus-Adapter (HBA)

Host-Bus-Adapter sind für die Fibre Channel-Verbindung eines Rechnersystems zum SAN zuständig und ersetzen den ATA-/SCSI-Controller der typischerweise für die Speichersystemanbindung zuständig ist. Die Adapter sind mit PCI-X- oder PCI-Express-Schnittstellen ausgestattet. Je nach PCI-Bus und Busgeschwindigkeit werden Datentransferraten von 2 Gbps (z. B. PCI-X mit 133 MHz) oder 4 Gbps (z. B. PCI-X 2.0 mit 266 MHz) realisiert. Die Kabelanbindung an das SAN geschieht entweder über Kupfer- oder Glasfaserkabeln. Mit Kupferkabel kann eine Kabellänge von bis zu 5 m, abhängig von elektromagnetischen Störungen, erreicht werden. Die maximal erreichbare Kabellänge des Glasfaserkabels ist abhängig von der Wellenlänge (850 nm oder 1310 nm) der Laserdiode des Fibre-Channel-Controllers und des Kabeltyps (Multimode / Monomode) des eingesetzten Lichtwellenleiters. Laserdioden mit einer Wellenlänge von 850 nm werden in Verbindung mit Multimode-Lichtwellenleitern für Kabellängen bis 500 m eingesetzt. Laserdioden mit einer Wellenlänge von 1310 nm und Monomode-Lichtwellenleiter kommen im Bereich bis 5 km zum Einsatz. Mit der Hilfe von so genannten Link Extendern können auch Strecken von über 100 Kilometer überbrückt werden. Da Host-Bus-Adapter selbst aber über kurze Distanzen verkabelt werden, unterstützen sie meist auch nur eine Kabellänge bis 500 m.

1.2.4 Multi-Path

Server- und Speichersysteme können redundant über mehrere Glasfaserkabel an eine SAN-Fabric angeschlossen werden. Die Multi-Path-Funktionalität stellt auf einem Serversystem redundante Datenpfade zum Speicherlaufwerk zur Verfügung und beugt auf diesem Weg einem Ausfall vor. Auf einem Serversystem wird Multi-Path mit Hilfe redundanter Host-Bus-Adapter und dem dazugehörigen Treiber mit Multi-Path-Unterstützung realisiert.

1.3 Speicherarchitekturen

In diesem Abschnitt erhalten Sie einen Überblick über Speicherarchitekturen, die sich hinsichtlich ihrer Eignung für die Integration in eine HV-Lösung unterscheiden. Es wird beschrieben, wie Massenspeicher an Rechnersysteme angeschlossen und betrieben werden. Aufgrund ihrer besonderen Bedeutung und Verbreitung in der IT, werden die drei folgenden Speicherarchitekturen in den weiteren Abschnitten ausführlich dargestellt:

- DAS (Direct Attached Storage),
- NAS (Network Attached Storage)
- und SAN (Storage Area Network)

Die drei Speicherarchitekturen werden anhand der Kriterien Skalierbarkeit, Performanz, Administrierbarkeit sowie IT-Sicherheitsaspekten (Verfügbarkeit, Vertraulichkeit und Authentizität) verglichen.

1.3.1 Direct Attached Storage (DAS)

Als Direct Attached Storage (DAS) werden Massenspeicher bezeichnet, die direkt an ein Rechnersystem angeschlossen sind. Damit sind meistens Festplatten, seltener optische, magnetoptische oder Flash-Speicher gemeint. In den meisten Konfigurationen befinden sich die Festplatten direkt in einem Gehäuse mit Motherboard und Peripheriebaugruppen, etwa Grafikadapter, Netzkarte, Festplattencontroller. Separat an ein Rechnersystem angeschlossene Festplattensubsysteme sind ebenfalls als DAS zu verstehen. Die DAS-Architektur soll primär dazu dienen, einem einzelnen Rechnersystem Festplattenspeicher exklusiv zur Verfügung zu stellen. Die Festplattenzugriffe werden durch das Betriebssystem des Rechnersystems gesteuert und finden blockbasiert statt.

Verfügbarkeit

Redundanz in DAS-Architekturen wird häufig auf Festplatten und Controller-Ebene hergestellt. Konkret bedeutet dies den Einsatz von RAID und von doppelt ausgelegten Festplatten-Controllern. In diesem Fall ist die Verfügbarkeit des Festplattenzugriffs aber abhängig von der Verfügbarkeit der restlichen Systemkomponenten. Fällt das Rechnersystem an dem die Festplatten betrieben werden aus, ist kein Festplattenzugriff mehr möglich. Eine Systemredundanz kann mit Hilfe von zwei Rechnersystemen, die gemeinsam ein Festplattensubsystem nutzen, und einer Cluster-Software hergestellt werden. Eine weitere Möglichkeit Redundanz herzustellen, ist der Einsatz eines verteilten Dateisystems (siehe Abschnitt 1.8.2).

Zugriffsschutz

DAS ist eine verteilte Architektur, bei der die Festplatten bzw. Festplattensubsysteme unter der Betriebssystemkontrolle des Rechners stehen, an den sie angeschlossen sind. Aufgrund dieser Tatsache ist es schwierig, Daten für Benutzer transparent im Netz zur Verfügung zu stellen und den Zugriffsschutz zu verwalten, es sei denn, es wird ein verteiltes Dateisystem, das sich über mehrere DAS-Systeme erstreckt, eingesetzt.

Skalierbarkeit

Die Skalierbarkeit eines DAS-Systems wird begrenzt durch die maximale Anzahl von Festplatten-Controllern und Festplatten, die in einem Rechnersystem oder Festplattensubsystem installiert werden können.

Performanz

Die Festplatten in einem DAS sind über ein I/O-Bussystem (ATA, SATA, SCSI oder SAS) mit dem Festplatten-Controller verbunden. Der Festplatten-Controller wiederum kommuniziert direkt mit dem Prozessor. Diese enge Kombination von Festplatte und Prozessor garantiert einen guten Datendurchsatz.

Administrierbarkeit

Ein DAS wird direkt von dem Rechnersystem, an dem es angeschlossen ist, mit lokalen Werkzeugen des Betriebssystems administriert. Der administrative Aufwand erhöht sich also auf Grund des dezentralen Ansatzes mit Zunahme der Anzahl der DAS-Systeme.

1.3.2 Network Attached Storage (NAS)

Das Konzept von Network Attached Storage (NAS) basiert auf einer zentralisierten Datenhaltung und ist eine an das lokale Netz angeschlossene Massenspeichereinheit. Auch hier werden meistens Festplatten, seltener optische-, magnetoptische- oder Flash-Speicher, damit bezeichnet. Der Zugriff auf ein NAS-System erfolgt über ein IP-Netz auf Dateiebene. Mit Hilfe der Netzprotokolle NFS (Network File Service) und SMB (Server Message Block) / CIFS (Common Internet File System) können Benutzer auf Dateien zugreifen, als ob sie auf ihren lokalen Festplatten abgespeichert wären. Zu diesem Zweck besitzen NAS-Systeme auf Dateitransfer ausgelegte Echtzeitbetriebssysteme. Vom Prinzip ist ein NAS ein auf Schnelligkeit, Skalierbarkeit, Sicherheit und Administrierbarkeit optimierter Dateiserver. Hochleistungs-NAS-Systeme werden oft dazu genutzt, viele kleinere Unix- oder Windows--Dateiserver zu ersetzen. Es wird zwischen der traditionellen NAS-Architektur (siehe Abbildung 1) und kombinierten NAS-/ SAN-Systemen (siehe Abbildung 2) unterschieden.

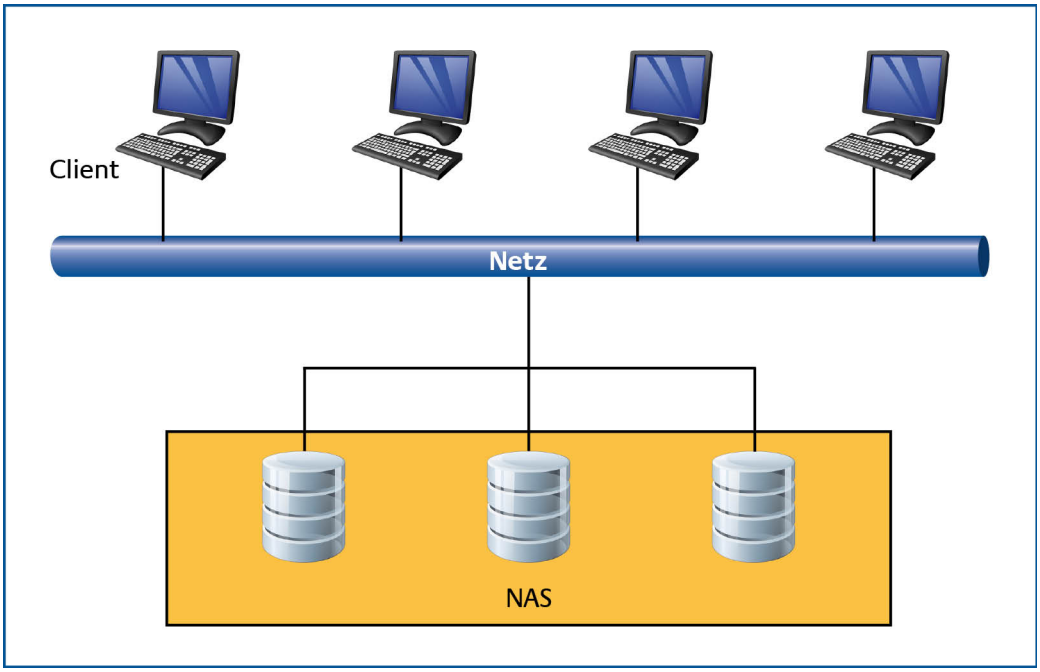


Abbildung 1: Traditionelle Anbindung eines NAS an ein lokales Netzwerk

Die kombinierten NAS- / SAN-Systeme, so genannte NAS-Heads, besitzen neben der normalen Netzschnittstelle zusätzlich einen integrierten Host-Bus-Adapter. Über den Host-Bus-Adapter ist es möglich das NAS-System mit einem SAN zu verbinden. Die Speicherkapazität ist nicht mehr von der NAS-Hardware, sondern von den Grenzen der Speicherkapazität im SAN abhängig. Ein weiterer Vorteil ist die Nutzung des SAN-Backups zur Sicherung des NAS-Systems. Die Nutzung des SAN-Backups führt in der Regel zu einer Entlastung der Netzwerkschnittstelle und zu einem effizienteren Backup.

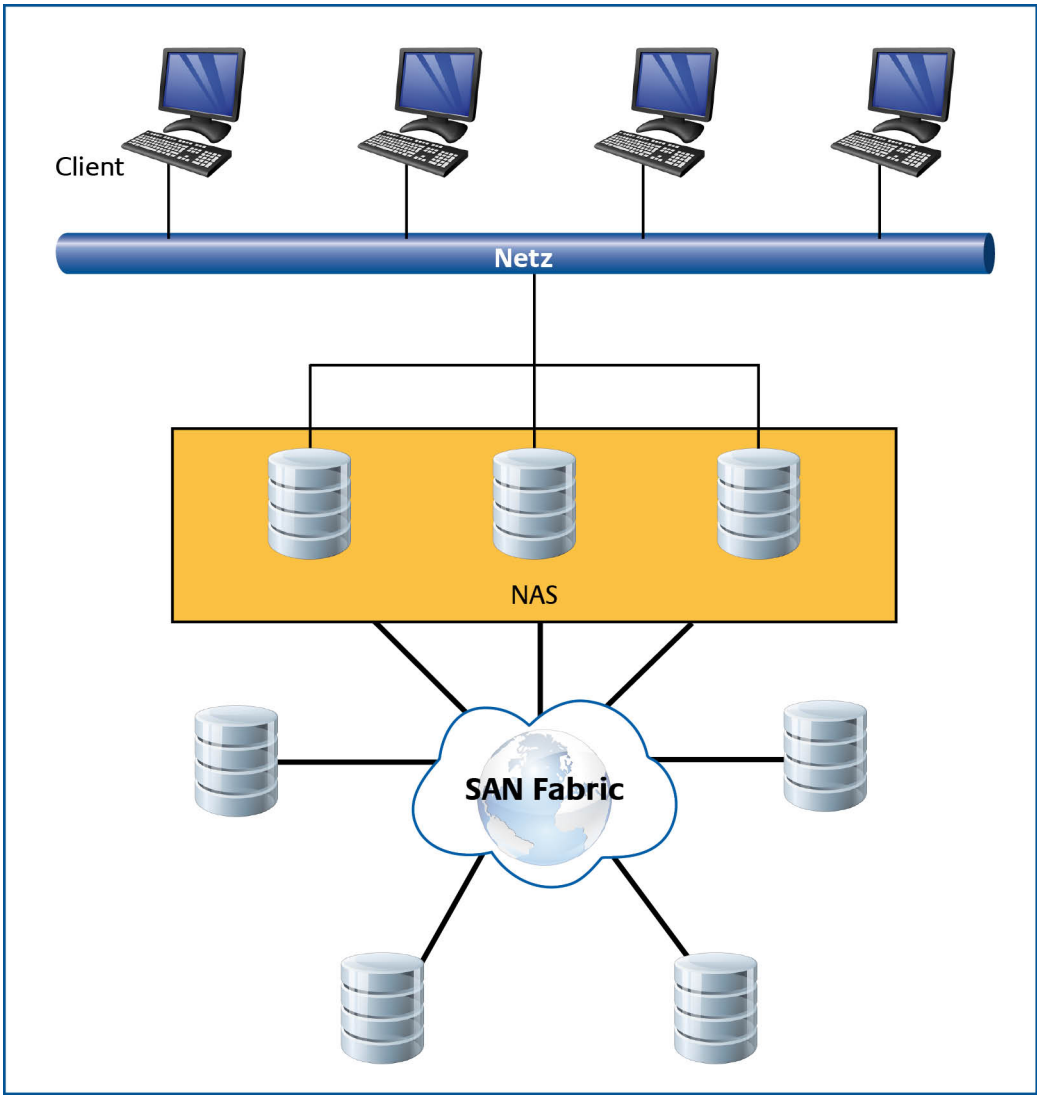


Abbildung 2: Kombinierte NAS-/SAN-Architektur

Verfügbarkeit

In der NAS-Architektur ist der Applikationsserver nicht für die Datenhaltung zuständig, und die Verfügbarkeit der Daten ist von eventuellen Fehlfunktionen der Anwendung oder des Betriebssystems unabhängig. Des Weiteren besitzen die meisten NAS-Systeme redundante Komponenten (Stromversorgung, Lüfter, Netzadapter, Festplatten (RAID) und Festplatten-Controller). Trotzdem kann es innerhalb eines NAS-Systems SPoF (etwa aktive Bussysteme) geben, welche die Verfügbarkeit gefährden können. Aus diesem Grund und zum Zweck der räumlichen Trennung (Backup-Rechenzentrum) kann ein Clustering des NAS-Systems in Kombination mit der Replikation der Daten notwendig sein. Voraussetzung ist dabei ein clusterfähiges NAS-Betriebssystem. Mit dieser Architektur kann Vollredundanz erreicht werden, d. h., es liegt Redundanz auf jeder Ebene der Architektur vor.

Zugriffsschutz

NAS-Betriebssysteme unterstützen sowohl Unix- (Unix File System) als auch Microsoft-Dateirechte (New Technology File System) und lassen sich oft in die Authentifizierungs- und Zugriffsschemata des Netzes über Verzeichnisdienste integrieren.

Skalierbarkeit

Bei der Erweiterung von NAS-Systemen sind die Punkte Speicherkapazität und Leistungsfähigkeit des Netzes zu betrachten. Bei der Speichererweiterung reicht das Spektrum von nicht erweiterbaren Systemen, bis zu Systemen bei denen online Festplatten hinzugefügt werden können. Im Fall, dass ein NAS-System nicht mit Festplatten erweitert werden kann, bestehen zwei Erweiterungsmöglichkeiten. Zum einen kann ein zusätzliches NAS-System im Netzwerk installiert werden, zum anderen besteht bei kombinierten NAS- / SAN-Geräten die Möglichkeit das NAS-System mit einem SAN zu koppeln. Es ist aber zu beachten, dass der Speicherausbau zu einer Erhöhung der Netzlast führen kann und eine Vergrößerung der Leistungsfähigkeit der Netzanbindung oder des Netzwerks erforderlich wird. Im nun folgenden Punkt, „Performanz“ wird das Problem des Netzwerkdurchsatzes näher betrachtet.

Performanz

NAS-Systeme sind dedizierte Lösungen, die ausschließlich für die Bereitstellung von Daten im Netz optimiert sind. Da in einem Netzwerk (LAN oder WAN) auch andere Daten transferiert werden, ist der Datendurchsatz in der NAS-Architektur abhängig von der Leistungsfähigkeit und Güte des Netzes, an dem das NAS-System betrieben wird. Der Overhead des TCP / IP-Protokolls wirkt sich ebenfalls negativ auf den Datendurchsatz aus und das Netz kann zum Flaschenhals werden. Besonders kritisch sind Latenzzeiten im Zusammenhang mit Applikations- und Datenbankservern zu sehen, die eigentlich blockorientiert auf Festplatten zugreifen sollten. Abhilfe kann das Priorisieren des Netzverkehrs oder die Netzanbindung des NAS-Systems über eine Bündelung der Netzanbindung (Trunk) schaffen. Im schlimmsten Fall ist ein Redesign des Netz-Backbones zur besseren Verteilung des Netzverkehrs erforderlich.

Administrierbarkeit

Die NAS-Architektur mit ihrem zentralen Ansatz eines heterogenen Netzspeichers trägt zu einer Verbesserung der Administrierbarkeit der Massenspeicher bei. Mit Hilfe von NAS ist es möglich, einzelne verteilte Speicherlösungen zu einem zentralen System zu konsolidieren. Ein NAS-System wird meistens einfach über einen Web-Browser administriert.

1.3.3 Storage Area Network (SAN)

Ein Storage Area Network (SAN) ist unabhängig vom lokalen Netz und beinhaltet ausschließlich Speichersysteme. Abbildung 3 zeigt den grundlegenden Aufbau der SAN-Architektur. Das SAN-Konzept wurde entwickelt, um das Problem der nur eingeschränkt flexiblen Nutzung der Speicherkapazitäten im DAS und NAS zu lösen. Ein SAN basiert auf Glasfasertechnologie (selten auch auf Kupferkabeln) und erlaubt sehr hohe Transferraten (bis zu 4 Gbps) zwischen Host und Datenspeicher. Mit Hilfe der SAN-Technologie ist es möglich Speichersysteme physikalisch getrennt von den Servern über eine Distanz von bis zu 100 km zu betreiben. In der SAN-Architektur wird zwischen den drei Topologien Point-to-Point, Arbitrated Loop und Switched Fabric unterschieden (siehe hierzu auch Abschnitt 1.4).

Innerhalb der Switched-Fabric-Topologie gibt es die Konzepte Full Mesh (siehe Abbildung 5), Partial Mesh, Simple Core / Edge, Compound Core / Edge und Complex Core / Edge [JuBK01]. In hochverfügbar ausgelegten SANs wird vorwiegend die Switched-Fabric-Topologie eingesetzt, weil sie die besten Leistungsmerkmale hinsichtlich der Verfügbarkeit, Skalierbarkeit und Performanz besitzt. Ein SAN wird ähnlich wie bei einem LAN über Hubs und Switches gebildet. Der Datenzugriff besteht hauptsächlich in der Übertragung von blockbasierten Daten, wie in der Kommunikation zwischen Rechner und Festplatte (ATA und SCSI).

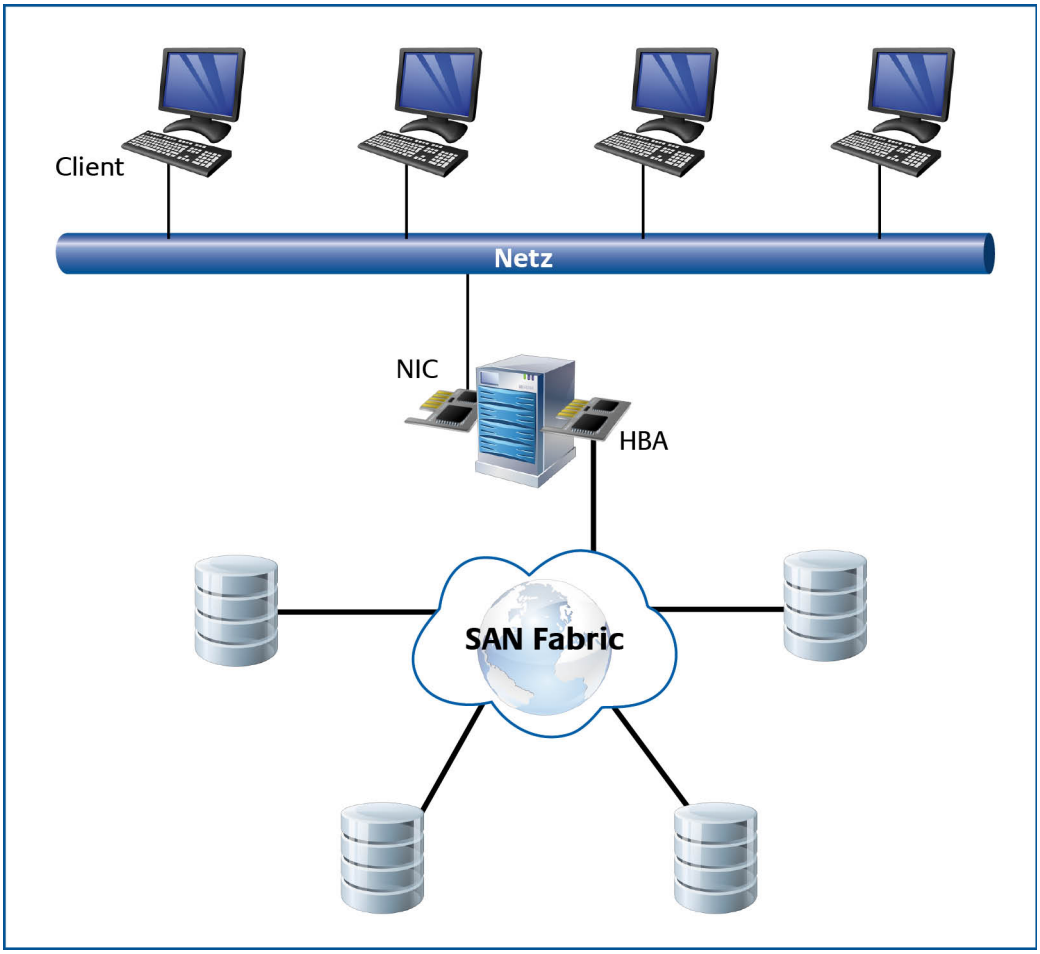


Abbildung 3: SAN-Architektur

Verfügbarkeit

Für die Verfügbarkeit eines SAN ist die SAN-Anbindung der Server, die Komponenten der Festplattensubsysteme sowie die Auswahl und das Design der SAN-Topologie ausschlaggebend. Der Server sollte über redundante Host-Bus-Adapter und eine Multipath-Software für die Unterstützung von redundanten Fibre-Channel-Pfaden verfügen. Innerhalb des SAN müssen redundante Fibre-Channel-Pfade über redundante Hubs oder Switches vorhanden sein. Die Festplattensubsysteme müssen redundante Stromversorgung, Lüfter, Netzadapter, Festplatten (RAID) und Host-Bus-Adapter besitzen.

Zugriffsschutz

Ein SAN ist ein dediziertes Speichernetz und der Zugriff von den Servern auf die Festplatten findet blockbasiert statt. Die Dateisicherheit ist also abhängig von den Betriebs- und Dateisystemen, die auf den Servern für die Verwaltung der Festplatten eingesetzt wird. Innerhalb des SAN kann der Zugriff auf die Speichersysteme mit Hilfe von Zoning und LUN-Masking (Logical Unit Number-Masking) gesteuert werden. Das Zoning ist das Unterteilen eines SAN in logische Subnetze und das LUN-Masking schränkt die Sichtbarkeit von Festplatten ein. Jeder Rechner sieht beim LUN-Masking nur die Festplatten, die ihm zugewiesen sind.

Skalierbarkeit

Die Erweiterbarkeit eines SAN wird bestimmt durch die Auswahl der SAN-Topologie und das nachfolgende Design der ausgewählten Topologie. Die größte Flexibilität hat hierbei die Switched-

Fabric-Topologie mit dem Core / Edge-Konzept. Sie benötigt aber auf jedem Falle entsprechende F-Switches.

Performanz

Durch die modulare Natur des SAN, mit getrennten Festplattensubsystemen und Fabric-Switches, ist ein sehr genaues Tuning möglich. Die größte Flexibilität hat hierbei die Switched-Fabric-Topologie mit dem Core / Edge-Konzept. Ein Core / Edge-Design ermöglicht eine Leistungssteigerung entsprechend dem Wachstum des Unternehmens. Hierfür gibt es verschiedene Möglichkeiten, darunter u. a. das Hinzufügen von Inter Switch Links (ISL) oder modernerer Funktionen zur Leistungssteigerung, wie ISL-Trunking.

Administrierbarkeit

Ein SAN kann, je nach Anforderung, bis zur Größe einer komplexen voll vermaschten Netzinfrastruktur anwachsen (Switched-Fabric-Topologie). Die Hersteller von SAN-Komponenten bieten deshalb umfangreiche Managementsoftware für die zentrale Administration eines SANs an. Es gibt Software für das Monitoring, Fabric-, Speicher- und Datenmanagement. Auf Grund der komplexen SAN-Administration ist es ratsam, einen dedizierten Management-Rechner zu installieren. Dieser Rechner ist durch geeignete Sicherheitsmaßnahmen vor unbefugtem Zugriff zu schützen.

1.3.4 Zusammenfassung Speicherarchitekturen

In Tabelle 3 sind die Eckpunkte der drei Speicherarchitekturen zusammengestellt.

	DAS	*NAS*	*SAN*
Basiert auf Netztechnologie	nein	ja	ja
Basisprotokoll	ATA, SCSI, FC	Ethernet	FC
Performanz	Ultra320-SCSI 320 MByte/s	Ethernet 10 Gbps, PCI-X 133 MHz mit max. 8 Gbps, TCP / IP-Netz max. 40% netto, ergibt max. 400 MByte/s unter optimalen Bedingungen	500 MByte/s
Preisniveau	niedrig	hoch	sehr hoch
Administrierbarkeit	dezentral	zentral	zentral
Hochverfügbarkeit	ungeeignet	eingeschränkt geeignet	geeignet
Skalierbarkeit	schlecht	relativ	sehr gut
Datentypen	alle Daten – blockbasierter Zugriff	nur Dateien – Dateibasierter Zugriff	alle Daten - blockbasierter Zugriff
Sicherheit (Vertraulichkeit)	hoch	mittel	mittel
Einsatzgebiet	Datei-, Applikations- und Datenbankserver	Dateiserver	Datei-, Applikations- und Datenbankserver

Tabelle 3: Zusammenfassung Speicherarchitekturen

1.4 Topologien

Speichernetze stellen topologisch eine Direktverbindung zwischen einem Server- und einem Speichersubsystem dar. Es wird zwischen drei verschiedenen Topologien unterschieden: direkte Verbindung (engl. Point-to-Point), Fibre Channel Arbitrated Loop (FC-AL) und Switched Fabric.

1.4.1 Point-to-Point

Die Point-to-Point-Topologie stellt eine einfache Verbindung zwischen zwei Geräten dar. Für ein Speichernetz bedeutet dies, dass die Point-to-Point-Topologie ein Serversystem mit einem Speichergerät verbindet. Die Point-to-Point-Topologie ist vergleichbar mit der klassischen SCSI-Anschlusstechnik von Speichergeräten. Der Vorteil von Fibre Channel gegenüber SCSI ist aber die höhere Bandbreite und die gesteigerte Distanz. Durch die optische Datenübertragung ist sie ebenfalls robuster gegenüber elektromagnetischen Störungen.

1.4.2 Arbitrated-Loop

Die Arbitrated-Loop-Topologie ist eine Ring-Topologie von zwei bis 126 FC-Geräten. In dieser Ring-Topologie findet die Datenübertragung nur in einer Richtung statt und es ist zu einem Zeitpunkt immer nur zwei Geräten gestattet, Daten miteinander auszutauschen. Das heißt also, dass die Bandbreite unter allen angeschlossenen Geräten aufgeteilt wird. Sind zum Beispiel drei Serversysteme mit Speichergeräten über einen Arbitrated Loop miteinander verbunden, steht jedem Server im Schnitt maximal ein Drittel der gesamten Bandbreite zur Verfügung.

Die Arbitrated-Loop-Topologie ist eine kostengünstige Variante, da keine Switches benötigt werden und sie ist eine gute Wahl für kleine und mittelgroße Konfigurationen. Als problematisch ist die Reihenschaltung der Geräte zu betrachten. Fällt ein Gerät im Arbitrated-Loop aus, steht der ganze Ring nicht mehr zur Verfügung. Die Lösung für solche Probleme bietet ein FC-Hub, der bei Ausfall eines Gerätes den entsprechenden Port schließt und den Ring wieder verbindet. Sollte der Hub ausfallen, so sind natürlich sämtliche Geräte an diesem Hub betroffen und werden nicht mehr verfügbar sein. Um diesen SPoF auszuschließen kann der FC-Hub redundant ausgelegt werden. Zusätzlich sind viele der heute erhältlichen Hubs mit redundanten Komponenten ausgestattet, so dass die Verfügbarkeit der einzelnen Komponenten erhöht wird.

1.4.3 Switched Fabric

Die Fabric-Topologie verwendet einen oder mehrere FC-Switches um Rechnersysteme und Speichergeräte dynamisch miteinander zu verbinden. Die Fabric-Topologie ist hinsichtlich der Switch-Konfiguration einem Ethernet ähnlich. Alle Geräte, die an einem FC-Switch angeschlossen sind, können gleichzeitig Daten übertragen. Ihnen steht zum selben Zeitpunkt auch die volle Bandbreite des Mediums zur Verfügung (siehe Abbildung 4).

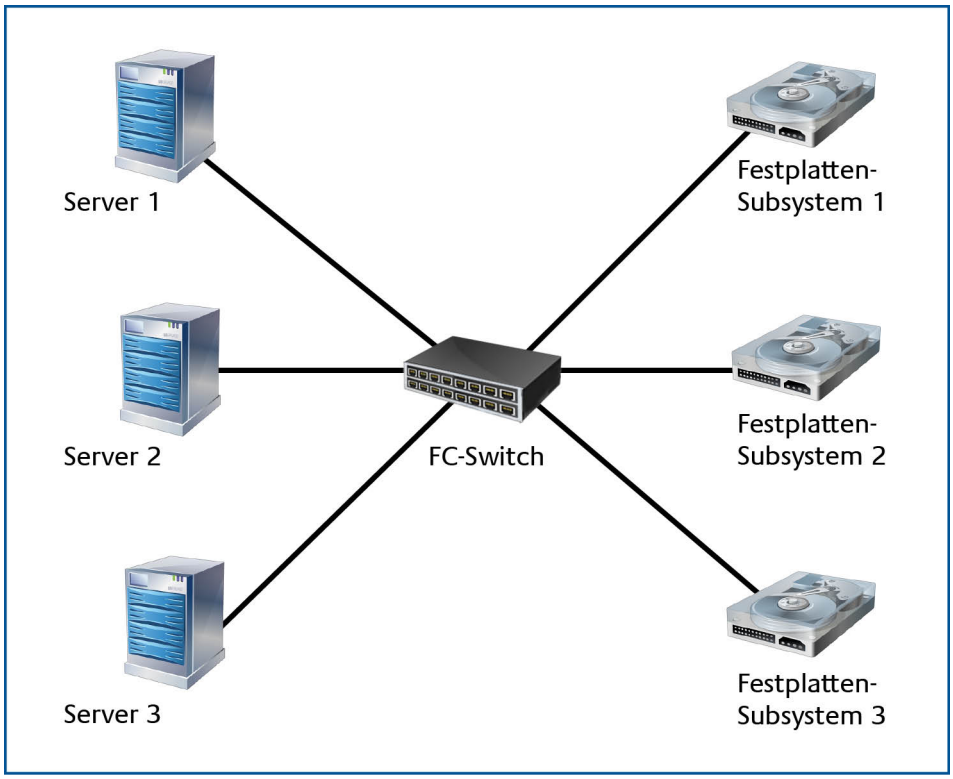


Abbildung 4: Fabric-Topologie

Inter Switch Links (ISL) dienen der Kopplung von FC-Switches. Um die Bandbreite einer Switch-zu-Switch-Verbindung zu erhöhen, ist es möglich, mehrere ISLs zwischen zwei Switches zu installieren. Inter Switch Links dienen nicht nur zur Vergrößerung der Bandbreite zwischen zwei Switches, sondern werden auch für die Konfiguration von doppelten Datenpfaden in Verbindung mit redundanten FC-Switches eingesetzt. Die nachfolgende Abbildung verdeutlicht dies anhand eines Full Mesh-Designs. In diesem Design sind vier FC-Switches über ISLs miteinander voll vermascht. Für die Server- als auch die Festplattensubsysteme stehen mit Hilfe von Multipath redundante Datenpfade zur Verfügung.

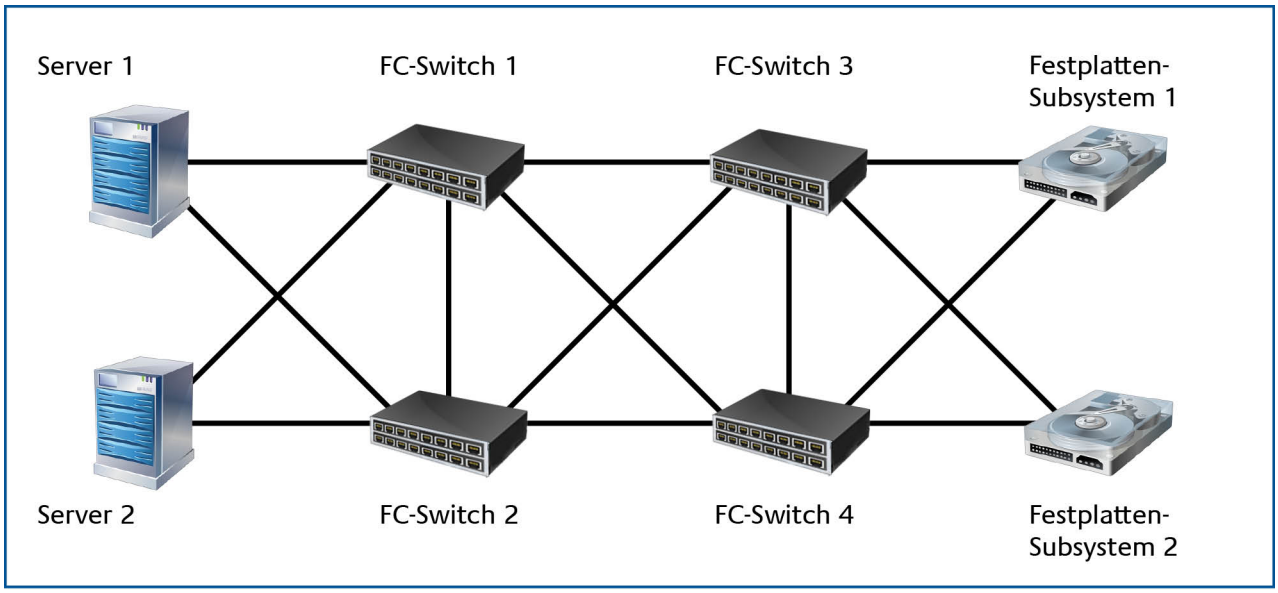


Abbildung 5: Fabric-Topologie in einem "Full-Mesh"-Design

Für eine Fabric-Topologie gibt es mehrere Design-Ansätze, um den jeweiligen Anforderungen an Datendurchsatz, Verfügbarkeit, Skalierbarkeit und Flexibilität gerecht zu werden [JuBK01].

1.5 Arbitrationsverfahren

Die Arbitration beschreibt Verfahren, bei denen sich die Nutzer nach einer gegenseitigen Vereinbarung das Zugangsrecht für eine Ressource bestimmen. Jedes an das Netzwerk angeschlossene Gerät hat generell die gleichen Rechte. Erst die Verhandlung eines Gerätes mit allen anderen sichert diesem den temporären Zugang. Das Verfahren wird u. a. in der FC-Topologie bei der Arbitrated Loop (siehe Abschnitt 1.4) aber auch bei diversen Bussystemen angewendet.

1.5.1 LUN-Masking

LUN-Masking ist eine Eigenschaft von FC-Switches und dient der Zugriffssteuerung innerhalb eines SANs. Die Zugriffssteuerung wird mit Hilfe der Logical Unit Numbers (LUN) realisiert, über die eine feste Beziehung zwischen Server- und logischer Speichereinheit gebildet wird. Diese bilden die kleinste adressierbare Einheit in einem Speichernetz. Ein Server, dem ein bestimmtes RAID-Laufwerke über LUN-Masking zugeordnet ist, kann nur auf dieses Laufwerk zugreifen. Ohne LUN-Masking könnte ein Server über die Fabric verfügbare Ressourcen erfragen und sich Speicher-LUNs aneignen, die zuvor einem anderen Server zugeteilt waren.

1.5.2 Zoning

Eine weitere Methode der Zugriffsteuerung ist das Zoning, das ebenfalls ein Merkmal des FC-Switches ist. Es wird zwischen einem Zoning das auf Ports und einem das auf Fibre Channel World Wide Names (WWN) basiert unterschieden. Da das WWN basierte Zoning relativ unsicher ist, wird darauf im Folgenden nicht mehr eingegangen. Beim portbasierenden Zoning wird die Zugriffssteuerung dadurch geregelt, dass nur bestimmte Ports miteinander kommunizieren dürfen.

1.6 Speichervirtualisierung

Bei der Speichervirtualisierung werden die verschiedenen physischen Speichermedien mit den unterschiedlichen Schnittstellen-Protokollen zu einem einzigen virtuellen Speicherpool zusammengefasst. Aus diesem Speicherpool kann bedarfsabhängig ein virtueller Speicher zusammengestellt werden. Ein solcher virtueller Speicher wird vom Host als lokale Speichereinheit behandelt, die erweiterbar und konfigurierbar ist.

Die Virtualisierung von Speicherressourcen bietet eine höhere Flexibilität und eine einfachere Verwaltung von Speichernetzen. Die Speichervirtualisierung trennt die physischen und logischen Ressourcen voneinander und schafft damit die Voraussetzungen um Speichersysteme von unterschiedlichen Herstellern zentral und flexibel zu verwalten. Darüber hinaus sind die Server-Betriebssysteme nicht mehr an dedizierte Storage-Arrays gebunden, sondern können Ressourcen des zentralen Speicherpools verwenden.

Zur Realisierung der Speichervirtualisierung existieren unterschiedliche Ansätze: In-Band und Out-of-Band sowie Lösungen basierend auf Servern, Speichersystemen und Switches.

1.6.1 In-Band-Methode

Bei der In-Band-Methode ist die steuernde Instanz (Software) im Netzwerk zwischen Host und Speichern installiert. Die Steuer- und Produktionsdaten werden durch diese Instanz geleitet, die sich den Servern als Speichersystem darstellt. Hier findet auch die Zuordnung von Speichersegmenten (logische Volumen) zum Host und die Zugriffsteuerung auf die Daten statt.

Vorteil der In-Band-Methode: der Server benötigt zur Speicherverteilung keine dedizierte Software mehr. Dies erleichtert die Einführung. Allerdings steht und fällt diese Art der Virtualisierung mit der Verfügbarkeit der steuernden Instanz. Aus diesem Grunde sind ist für die steuernde Instanz eine redundante Implementierungen angebracht.

1.6.2 Out-of-Band-Methode

Bei der Out-of-Band-Methode ist die steuernde Instanz außerhalb des Speichernetzes installiert und kommuniziert über das lokale Rechnernetz mit den Fibre-Adaptern des Servers. Dieser Ansatz benötigt lokal geladene Programme, so genannte Agenten. Die steuernde Instanz definiert die logischen Laufwerke, die ein Server anbinden und nutzen darf. Die einschlägigen Informationen über die logischen Blöcke und physischen Zylinder der Laufwerke werden anschließend auf dem Adapter des Hosts gespeichert. Auch hier kann es aber zu empfindlichen Störungen kommen, falls entweder die Hardware oder die Anwendung der steuernden Instanz ausfallen.

1.6.3 Host-basierende Virtualisierung

Die Host-basierenden Lösungen beschränken sich darauf, jeweils eine überschaubare Anzahl unterschiedlicher Server (meist im High-End-Unix-Umfeld) mit einer Anwendung auszustatten, welche die Gleichbehandlung aller Speichersysteme im SAN erlaubt. Dies beschränkt den Anwender allerdings wieder auf eine Server-Umgebung eines einzigen Herstellers mit entsprechenden Nachteilen durch die Abhängigkeit vom Lieferanten. Demgegenüber stehen die Vorteile beim Schutz gegen einen Gesamtausfall des Systems.

1.6.4 Speichersystem-basierende Virtualisierung

Auch auf Storage-Systemen selbst ist eine Virtualisierung möglich. Das Betriebssystem der meisten High-End-Speichergeräte ermöglicht es, nicht nur die vorhandenen Kapazitäten flexibel auf- und zuzuteilen, sondern darüber hinaus auch den Zugriff auf die logischen Laufwerke zu beschränken. Zudem lassen sich diese zwischen Rechnern und auch Anwendungen hin und her bewegen sowie weitere Dienste wie Sicherung, Spiegelung oder Replikation durchführen. Aufgrund der Auslegung von High-End-Speichergeräten mit redundanten Bauteilen, RAID- oder Spiegelschutz und Fernkopie von Daten aus dem Rechenzentrum in andere Standorte ist dies die ausfallsicherste Virtualisierungs-Methode.

1.6.5 Virtualisierung direkt im FC-Switch

Ein neuer Trend in der Virtualisierungsdebatte sind die intelligenten Netzwerkkomponenten wie Router oder Switches. Hier wird die steuernde Instanz nicht auf eine separate Appliance im SAN ausgelagert, sondern befindet sich direkt im FC-Switch. Diese Geräte ermöglichen den Anschluss annähernd aller heute produktiv betriebenen Server- und Speichersysteme und stellen somit den wahrscheinlich zukunftssichersten Ansatz der Virtualisierung dar. SAN-Switches sind sowohl auf Ausfallsicherheit und Verfügbarkeit wie beispielsweise durch Kreuzverkabelung und redundante Komponenten, als auch auf Durchsatz und Leistung ausgelegt. Korrupte oder unvollständige Informationen können bei richtigem Aufbau eines solchen Netzes annähernd ausgeschlossen werden [Herz04].

1.7 Datensicherungs- und Wiederherstellungsverfahren

Wenn präventive Maßnahmen zur Absicherung nicht erfolgreich waren und Datenverlust auftritt, muss auf reaktive Maßnahmen zurückgegriffen werden. Die Datensicherung (Backup) ist die letzte Möglichkeit, verlorene oder korrupte Daten wiederherzustellen und somit auch den funktionalen Prozessablauf wieder zu gewährleisten. Da Informationen zum wichtigsten Kapital von Unternehmen und Organisationen zählen, sind umfassende Datensicherungsmechanismen notwendig, dies vor allem hinsichtlich der Rückversicherung [LiCh03].

Daher ist ein geeignetes Datensicherungskonzept unverzichtbarer Bestandteil jeder IT-Architektur und somit auch der HV-Konzeption. Grundsätzlich sollte das HV-Design auf die Vermeidung von Datenverlusten ausgelegt sein, da anderenfalls hohe Verfügbarkeit nicht gewährleistet ist. Es existieren jedoch Szenarien, in denen das HV-Design an Grenzen stößt.

Diese sind:

- Verlust der Daten oder der Datenintegrität durch Fehlbedienung autorisierter Personen, Sabotage und Malware (da Spiegelungen der Daten ebenfalls betroffen sind).

– Verlust der Daten durch den Totalverlust des HV-Systems (Redundanz ist ebenfalls betroffen).

Ist einer der o. g. Fälle aufgetreten, muss im HV-Umfeld von einem IT-Notfall ausgegangen werden. Muss zur Wiederherstellung (Recovery) der Verfügbarkeit auf eine Datensicherung zurückgegriffen werden, sind die Wiederherstellungszeiten oftmals deutlich größer als die in den HV-Anforderungen tolerierten Ausfallzeiten (Down-Time).

In den nachfolgenden Abschnitten werden verschiedene Verfahren für die Datensicherung und für die Wiederherstellung erläutert. Diese Verfahren sind in unterschiedlicher Art und Weise geeignet, die oben aufgeführten, teilweise gegensätzlichen, Anforderungen zu erfüllen. Im Rahmen des HV-Designs muss deshalb eine sorgfältige Abwägung der Mechanismen erfolgen.

1.7.1 Datensicherungsverfahren

Die entscheidenden Kriterien, nach denen eine Datensicherung durchgeführt werden kann, sind der für eine Datensicherung zur Verfügung stehende Zeitraum und die Art des Zugriffs auf die Daten. Die Zugriffsart ergibt sich meistens aus dem Zeitfenster. Bei hohen Anforderungen müssen Daten permanent verfügbar sein und dürfen deshalb nicht für den Benutzerzugriff gesperrt werden. Ein Verfahren zur Datensicherung, welches das Sperren des Benutzerzugriffs voraussetzt, kann in solchen Fällen nicht eingesetzt werden. Im Folgenden werden die gängigen Arten verschiedener Datensicherungsverfahren beschrieben und deren Einsatz erläutert.

1.7.1.1 Offline-/Online-Sicherung

Bei einer Offline-Sicherung werden alle Benutzer, die eine Verbindung mit dem zu sichernden Server haben, vom Netz getrennt. Für die Dauer der Sicherung können sich Benutzer nicht am System anmelden. Die Offline-Sicherung findet meist über Nacht oder am Wochenende statt und ist damit vom zur Verfügung stehenden Zeitfenster und des zu sichernden Datenvolumens abhängig. Für den Einsatz einer Online-Sicherung ist die richtige Bandrotation (Abschnitt 1.7.3) zur Minimierung der Datensicherungszeit von großer Bedeutung. Bei der Online-Sicherung bleibt die Verbindung zum Server für die Benutzer erhalten. Dadurch kann es zu einer Verminderung der Netz-Performanz kommen. Des Weiteren kann es zu Problemen bei der Sicherung von geöffneten Dateien kommen. Die klassische Offline- / Online-Sicherung wird hauptsächlich mit Bandlaufwerken eingesetzt und ist ein weit verbreitetes Verfahren im DAS-Umfeld.

Um Datensicherungen zeitsparend und sicher durchzuführen, ist eine dem Zeitfenster und Datenvolumen angepasste Sicherungsstrategie erforderlich. Es ist aber nicht nur die Sicherung, sondern auch die Wiederherstellung der Daten im Fall des Datenverlustes zu beachten. Entscheidend ist hier, wie viel Zeit einer IT-Organisation zur Rekonstruktion der Daten zur Verfügung stehen.

In den nächsten Abschnitten werden die klassischen Datensicherungsstrategien vorgestellt.

Vollsicherung

Bei einer Vollsicherung werden sämtliche Daten eines Servers, einer Festplatte, einer Partition oder eines Verzeichnisses gesichert. Eine Vollsicherung kann je nach Datenvolumen und Speichertechnologie sehr zeitaufwändig sein und eignet sich eigentlich nur für kleine Datenmengen.

Differentielle Sicherung

Bei der differentiellen Sicherung werden alle Daten, die nach der letzten Vollsicherung entstanden sind oder verändert wurden, bei jedem Sicherungslauf erneut erfasst. Für eine Datensicherung mit einem Bandlaufwerk bedeutet das, dass für eine Rücksicherung die Vollsicherung und das letzte Band der differentiellen Sicherung benötigt werden.

Inkrementelle Sicherung

Die inkrementelle Sicherung setzt auch auf der Vollsicherung auf und sichert nur die Dateien die seit der letzten Sicherung (voll oder inkrementell) veränderten wurden. Im schlimmsten Fall kann das für eine Wiederherstellung mit Bändern den Einsatz des kompletten Bandsatzes nach sich ziehen. Konkret bedeutet dies, dass eine Vollsicherung und zusätzlich alle inkrementellen Sicherungen eingespielt werden müssen. Das einfachste Verfahren einer inkrementellen Sicherung besteht darin, dass die Sicherungs-Software bei der Sicherung das so genannte Archiv-Bit von Dateien löscht. Das Betriebssystem setzt das Archiv-Bit wieder, sobald eine Datei verändert oder neu erstellt wird.

Virtuelle Vollsicherung

Die virtuelle Vollsicherung ist eine Kombination aus einer Vollsicherung und darauf aufbauenden inkrementellen Sicherungen. Zuerst wird eine Vollsicherung erstellt. Zu einem späteren Zeitpunkt wird eine inkrementelle Sicherung erzeugt. Anschließend bildet die Sicherungssoftware aus der Vollsicherung und den inkrementellen Sicherungen, durch einpflegen der entsprechenden geänderten Blöcke eine neue (virtuelles) Vollsicherung. Die so erstellte virtuelle Vollsicherung reduziert den für die Wiederherstellung notwendigen Zeitraum, weil sie nicht aus einzelnen inkrementellen Sicherungen zusammengesetzt werden muss.

Image-Sicherung

Eine besondere Form der Datensicherung ist die Image-Sicherung. Grundsätzlich entspricht sie der Vollsicherung, da sämtliche Daten gesichert werden. Bei der Image-Sicherung werden aber nicht die einzelnen Dateien und Verzeichnisse einer Festplatte gesichert, sondern die physikalischen Sektoren der Festplatte, d. h. es wird eine ganze Festplatte oder eine ganze Partition gesichert. Bei der Wiederherstellung wird ebenfalls die Festplatte oder die Partition komplett wiederhergestellt. Vielfach können durch diverse Tools auch einzelne Dateien aus Images zurückgeholt werden. Die Image-Sicherung wird meistens für die Betriebssystemsicherung eingesetzt.

1.7.1.2 Datenspiegelung und –Replikation

Die Datenspiegelung ist der Sicherungsprozess, der Daten simultan auf zwei Festplatten schreibt. Diese redundanten Festplatten können sich in einem Rechnersystem oder in zwei räumlich getrennten Geräten befinden. Fällt eine Festplatte aus, können die Daten weiterhin von der zweiten gelesen werden. Eine Datenspiegelung schützt nur gegen Festplattenausfälle, nicht aber gegen Datenverlust, hervorgerufen durch Anwendungsfehler und böswillige Programme (z. B. Virus). Im Fall von fehlerhaften Quelldaten sind auch die gespiegelten Daten fehlerhaft. Die Datenspiegelung ist für den Benutzer völlig transparent. Sie wird häufig für die Aktualisierung von Daten auf Backup-Systemen in Backup-Rechenzentren verwendet. In diesem Fall müssen die Daten möglichst in Echtzeit auf das abgesetzte System repliziert werden. Je nach Art der Replizierung liegt zwischen der Bearbeitung- und Erstellung der Primärdaten und ihrer Replizierung eine gewisse Zeitspanne. Diese Zeitspanne wird als Latenzzeit bezeichnet.

1.7.1.3 Synchrone Replikation

Wenn die Primärdaten und Replikate gleiche Datensätze haben, also keine Latenz vorhanden ist, spricht man von Synchronität. Die primären und replizierten Datensätze sind also identisch. In der Praxis ist aufgrund von Datenlaufzeiten ein hundertprozentig synchrones Replikationssystem nicht möglich. Beispielsweise in dem Fall, wenn primäre und replizierte Daten geographisch sehr weit

auseinander liegen ergeben sich hohe Übertragungszeiten, welche die zulässigen Ausfallzeiten bereits überschreiten.

1.7.1.4 Asynchrone Replikation

Wenn zwischen der Bearbeitung der primären Daten und der Replizierung eine Latenz liegt, spricht man von Asynchronität. Die Daten sind nur zu dem Zeitpunkt der Replikation synchron (identisch). Eine einfache Variante der asynchronen Replikation ist die „File-Transfer-Replication“, der Transfer von Dateien via FTP (File Transfer Protocol) oder scp (Secure Copy).Viele Hersteller bieten sowohl für den NAS- als auch den SAN-Bereich Lösungen zur Datenreplikation an. Die meisten Lösungen sind Kombinationen aus einer Cluster-Software und einer zusätzlichen Softwareoption zur Datenreplikation. Bei allen Lösungen bildet aber das Netz den Flaschenhals. Die Leistungsfähigkeit und Güte der Netzverbindungen entscheidet über einen sinnvollen Einsatz der Datenreplikation. Diese Schwachstelle wird oft durch eine inkrementelle Replikation gemildert. Konkret bedeutet das einen initialen Datentransfer des gesamten Datenbestandes über ein Bandmedium, mit anschließender Replikation nur der geänderten Daten über das Netz.

1.7.1.5 Snapshot

Ein Snapshot ist ein lokal gespeichertes Abbild der Festplattendaten zu einem bestimmten Zeitpunkt. Eine Snapshot-Kopie ist ein „eingefrorenes“ Abbild ohne Schreibzugriff eines logischen Volumes, das einfachen Zugriff auf ältere Versionen von Dateien, Verzeichnishierarchien und / oder LUNs bietet. Der Erstellungsvorgang einer Snapshot-Kopie ist für den Benutzer völlig transparent. Sie kann in kurzer Zeit erstellt werden und eignet sich sehr gut für die Zwischenspeicherung von Daten, um diese anschließend zur Archivierung auf ein Bandmedium zu schreiben. Die Snapshot-Technik wird im NAS- und SAN-Umfeld eingesetzt.

1.7.2 Datensicherungsarchitekturen

Die Datensicherungsarchitekturen beschreiben die Anbindung der Sicherungslaufwerke an die zu sichernden Systeme [Exab99]. In diesem Zusammenhang sind der Datendurchsatz und die Administrierbarkeit wichtig. Gerade die Administration und der Betrieb der Datensicherung stellen einen nicht unerheblichen Aufwand dar.

Lokales Backup

Das lokale Backup sichert mit Hilfe eines direkt am Rechnersystem angeschlossenen Sicherungslaufwerks die lokalen Festplatten. Diese dezentrale Topologie bietet einen guten Datendurchsatz, ist aber relativ unflexibel, da immer nur ein System gesichert werden kann. Auf Grund der festen Zuordnung von Datensicherung und Rechnersystem erhöht sich der Hardware- und Software-Einsatz mit der Zunahme der Anzahl der zu sichernden Rechnersysteme. Der gestiegene Hardware- und Software-Einsatz erhöht wiederum den Administrations- und Betriebsaufwand.

Netz-Backup

Das Netz-Backup ist ein zentraler Sicherungsansatz. Die zu sichernden Rechnersysteme werden über die Netzinfrastruktur auf ein zentrales Sicherungslaufwerk gesichert. Der Datendurchsatz ist abhängig von der Netzbandbreite und der Netzauslastung. Die Datensicherung kann den herkömmlichen Netzverkehr negativ beeinflussen und umgekehrt. Der sinnvolle Einsatz eines Netz-Backups ist abhängig von der Leistungsfähigkeit und Güte der Netzverbindungen über die es betrieben wird. Weitere Faktoren sind das zu sichernde Datenvolumen und das zur Verfügung stehende Zeitfenster. Ein dediziertes Backup-Netz oder ein auf die Erfordernisse des Netz-Backups

ausgelegter Quality of Service (QoS) des Netzes ist Voraussetzung für den Einsatz dieses Verfahrens im HV-Umfeld. Durch die Zentralisierung der Datensicherung reduziert sich der Administrations- und Betriebsaufwand.

SAN-Backup

Storage Area Network (SAN) sind dedizierte hoch performante Speichernetze, die ausschließlich für die Übertragung von Daten zu Massenspeichern gedacht sind. Festplattensubsysteme und Datensicherungssysteme werden über spezielle Protokolle wie Fibre Channel oder iSCSI direkt miteinander gekoppelt, so dass Daten zentral gesichert werden können. Durch den zentralen Ansatz wird der Administrations- und Betriebsaufwand minimiert. Das SAN-Backup ist die effizienteste Datensicherungsarchitektur.

Datensicherungsmedien

In diesem Abschnitt werden der Einsatz von Sicherungsmedien und die Kombinationsmöglichkeiten aufgezeigt. Unter Sicherungsmedien sind Festplatte, Bandlaufwerk, CD und DVD zu verstehen. Für die generelle Betrachtung der Architektur ist es unerheblich, welcher Typ Bandlaufwerk oder welcher Typ Festplatte eingesetzt wird. Entscheidend ist, dass eine Festplatte performanter als ein Bandlaufwerk ist [Exab05].

Backup Disk-to-Disk

In dieser Sicherungsarchitektur werden die Daten vom primären auf ein sekundäres Festplattensystem kopiert. Dieses zweite Festplattensystem sollte aus Gründen der Ausfallsicherheit unabhängig vom primären System sein. Um das Zeitfenster für die Datensicherung möglichst klein zu halten kann die Disk-to-Disk-Sicherung mit Snapshot-Copy kombiniert werden. Die Datensicherung auf Festplatten ist performant, birgt aber auf Grund des mechanischen Datenträgers die Gefahren von Datenverlust. Diese Architektur eignet sich nur für einen kurzen Aufbewahrungszeitraum.

Backup Disk-to-Tape

Die Disk-to-Tape-Sicherung ist die klassische Datensicherung von Festplattendaten. Bandsicherungen sind ein sicheres, aber auch ein langsames Sicherungsverfahren. Je nach Art der Bänder und deren Benutzung ist eine Lebensdauer von 10 – 30 Jahren möglich. Die langen Aufbewahrungszeiträume machen den Vorteil dieser Architektur aus. Bandsicherungen werden meistens mit einer auf das Datenvolumen ausgerichteten Sicherungsstrategie (Bandrotation) eingesetzt, um den Datensicherungszeitraum zu minimieren.

Backup Disk-to-Disk-to-Tape

Die Sicherung von Disk-to-Disk-to-Tape verbindet die Vorteile der beiden erstgenannten Architekturen. Mit Hilfe einer performanten Festplattensicherung werden die Daten für eine sichere Bandsicherung zwischengespeichert.

Backup Disk-to-Disk-to-Any

Bei der Disk-to-Disk-to-Any wird ebenfalls ein Festplattenmedium zur Zwischenspeicherung der Daten benutzt. Das Medium, auf dem die Daten letztendlich gesichert werden, kann je nach Anforderung und Anwendung variieren. Um zum Beispiel geringe Datenmengen im Zusammenhang mit einem Softwareupdate zu sichern, spielt es eine untergeordnete Rolle ob dies auf CD, DVD oder Festplatte geschieht. Ausschlaggebend sind in diesem Fall technische und ökonomische Gründe. Wichtig ist, dass die Sicherung transparent, kontrolliert und nachvollziehbar durchgeführt wird.

1.7.3 Bandrotation

Die Bandrotation bestimmt bei Bandsicherungen im entscheidenden Maße sowohl die Sicherungs- als auch die Wiederherstellungszeit. Mit ihrer Hilfe wird das Datensicherungsintervall in Verbindung mit der Datensicherungsstrategie geplant.

Sechs-Tage-Rotation

Die Sechs-Tage-Rotation ist ein simples und Kosten sparendes Schema, das sich gut für kleinere Organisationen mit wenig Datenvolumen eignet. Mit dieser Rotationsmethode werden mit zwei Freitagsbändern im Wechsel Vollsicherungen und mit vier Bändern an den verbleibenden Tagen inkrementelle oder differentielle Sicherungen durchgeführt.

Grandfather-Father-Son (GFS)

Die so genannte Großvater-Vater-Sohn-Generationssicherung ist ein bekanntes Schema zur Sicherung von Daten. Es wurde entwickelt, um auch im schlimmsten Katastrophenfall eine effiziente und schnelle Wiederherstellung von Daten zu ermöglichen.

Grandfather	1. Generation	Monatssicherung
Father	2. Generation	Wochensicherung
Son	3. Generation	Tägliche Sicherung

Normalerweise benötigt eine GFS-Sicherung 21 Bänder. Jeden Montag wird eine Vollsicherung durchgeführt (Father), eine inkrementelle oder differentielle von Dienstag bis Donnerstag (Son) und eine Vollsicherung am Freitag (Father). Zusätzlich wird an jedem Monatsende eine Komplettsicherung durchgeführt (Grandfather). Es werden vier Bänder für die tägliche Sicherung von Montag bis Donnerstag benutzt, fünf Bänder für die Wochensicherungen und 12 Bänder für die Monatssicherungen. Die Anzahl an Medien kann bei GFS-Sicherungen problemlos erweitert werden, so dass eventuell auch mehr als 21 Bänder im Sicherungskreislauf sind.

Towers of Hanoi (TOW)

Towers of Hanoi ist ein kompliziertes und teures Datensicherungsschema, welches auf dem gleichnamigen Spiel basiert. Das Schema benutzt Bänder unterschiedlich oft, d.h. die ersten Bänder werden bei jeder Rotation genutzt, andere Bänder bei jeder 4. Rotation, wiederum andere Bänder bei jeder 8. Rotation usw. In Tabelle 4 (Towers of Hanoi) ist das Rotationsschema übersichtlich dargestellt.

Tag	*Band*	*Nutzung*	*Sicherung*
1.	A	erstmals genutzt	Vollsicherung
2.	B	erstmals genutzt	Vollsicherung
3.	A	wiederbenutzt	inkrementell
4.	C	erstmals genutzt	Vollsicherung
5	A	wiederbenutzt	inkrementell
6.	B	wiederbenutzt	inkrementell
7.	A	wiederbenutzt	inkrementell
8.	D	erstmals genutzt	Vollsicherung
9.	A	wiederbenutzt	inkrementell
10.	B	wiederbenutzt	inkrementell
11.	A	wiederbenutzt	inkrementell
12.	C	wiederbenutzt	inkrementell
13	A	wiederbenutzt	inkrementell
14.	B	wiederbenutzt	inkrementell
15.	A	wiederbenutzt	inkrementell
16.	E	erstmals genutzt	Vollsicherung
17.	Die Rotation beginnt wieder am 1. Tag		

Tabelle 4: Towers of Hanoi (Quelle: [Exabyte])

Band A wird zuerst für eine Vollsicherung genutzt, anschließend für inkrementelle Sicherungen. Auf Grund der Tatsache das Band A am häufigsten genutzt wird und stärker verschleißt, muss es öfter ausgetauscht werden.

Tabelle 5 stellt noch einmal die drei Typen der Bandrotation mit Vor- und Nachteilen gegenüber.

Rotation	*Vorteil*	*Nachteil*
Grandfather-Father-Son (GFS)	Sichere Datensicherungs-methode. Archiviert einen kompletten Monat.	Dadurch, dass die Rotationsmethode eine große Anzahl an Bänder benötigt, ist sie recht teuer.
Towers of Hanoi (TOW)	Sichere Datensicherungs-methode. Einfache und schnelle Wiederherstellung von kompletten Datenbeständen	Benötigt eine große Anzahl an Bändern. Teure und komplizierte Rotationsmethode.

Tabelle 5: Vor- und Nachteile von Bandrotation (Quelle: [Exab04])

1.7.4 Wiederherstellung

Die Definition des Wiederherstellungsprozesses ist ein wichtiger Bestandteil im Notfallhandbuch einer Behörde oder eines Unternehmens. Dabei geht es neben der geeigneten Strategie auch um die Wiederherstellungszeit. Diese Aspekte können nur bewertet werden, wenn basierend auf der erstellten Datensicherung, entsprechende Wiederherstellungsübungen durchgeführt wurden. Die tatsächliche Wiederherstellungszeit ist mit der maximal tolerierbare Wiederherstellungszeit (Recovery Time Objective - RTO) im Rahmen einer Risikobewertung zu vergleichen. In diesem

Zusammenhang muss berücksichtigt werden, dass die Wiederherstellung der Daten nicht gleichzusetzen ist mit der Wiederherstellung der Anwendung. Bei der Wiederherstellung der Anwendung beansprucht die Wiederherstellung der Daten zwar in der Regel den größten Zeitaufwand, sie ist aber, neben der Fehleranalyse und der Wiederherstellung der Anwendung selbst, nur ein Bestandteil des gesamten Wiederherstellungsprozesses.

Es besteht auch die Möglichkeit bereits vor dem vollständigen Zurückschreiben der Daten, die Anwendung wiederherzustellen und für den Anwender freizugeben. Dies kann zwar aufgrund des Schreibvorgangs die Performanz der Anwendung beeinflussen, reduziert jedoch die Ausfallzeit der Anwendung.

Auf dem Weg zur passenden Sicherungsstrategie müssen IT-Verantwortliche einige grundsätzliche Überlegungen anstellen und sich darüber Klarheit verschaffen, welche Ziele es zu erreichen gilt. Folgende Aspekte sollten berücksichtigt werden:

- In welcher Zeit müssen Daten nach einem Ausfall wiederhergestellt (Recovery Time Objective - RTO) werden?
- Wie aktuell müssen die wiederhergestellten Daten sein? Wie viel Datenverlust kann in Kauf genommen werden (Recovery Point Objective - RPO)?
- Wie groß ist die zu sichernde Datenmenge?
- Welche Backup-Medien / -Orte stehen zur Verfügung?

Diese Fragen lassen sich mithilfe der Ergebnisse aus der SOLL-Analyse beantworten und haben in der Umsetzungskonsequenz direkten Einfluss auf das Speicher-Management (siehe Abschnitt 1.9).

1.7.5 Datensicherung in HV-Umgebungen

Zu Beginn dieses Abschnittes sind Szenarien beschrieben worden (siehe Abschnitt 1.7), welche die Grenzen eines HV-Designs bezüglich Integrität und Verfügbarkeit der Daten zeigen sollten. Auch bei gewissenhafter Planung und Umsetzung umfangreicher Maßnahmen, kann der Verlust der Integrität oder Verfügbarkeit nicht ausgeschlossen werden. Deshalb ist zusätzlich zur funktionellen- und strukturellen Redundanz in einer HV-Umgebung immer die Informationsredundanz in Form von Datensicherungen erforderlich.

Die Verfügbarkeit der Daten bei einem Komponentenausfall in einem HV-Betrieb lässt sich mit einer Spiegelung der Daten auf einem weiteren Speichersystem gewährleisten. Leider ist diese Maßnahme zwecklos bei Totalverlust, Fehlbedienung, Sabotage, Malware oder Softwarefehler wie zum Beispiel die Blockkorruption, da sich diese Beeinträchtigung auch auf die gespiegelten Daten auswirkt.

Nicht integere oder nicht verfügbare Daten in einem HV-Design stellen einen IT-Notfall dar. IT-Notfall heißt u. a., dass SLAs nicht eingehalten werden können, oder schlimmstenfalls die Daten in elektronischer Form verloren gegangen sind und äußerst mühsam rekonstruiert werden müssen. In vielen Fällen wird die Rekonstruktion der Daten nicht möglich sein.

Im IT-Notfall muss also auf eine zuverlässige Datensicherung zurückgegriffen werden können. Das Datensicherungskonzept im HV-Umfeld muss deshalb folgenden Anforderungen genügen:

- persistente (zeitlich dauerhafte) und konsistente (nicht anfällig gegen Datenverluste) Datenhaltung,

- die Datensicherungsmechanismen dürfen den HV-Betrieb nicht wesentlich beeinträchtigen (z. B darf ein Dienst während des Backups nicht deaktiviert werden),
- die Datensicherung muss eine ausreichende Aktualität gewährleisten, um den potentiellen Datenverlust zu minimieren,
- die Datensicherung muss hinreichende Historie (Archivierungshistorie) ermöglichen,
- externer Lagerort der Datensicherung (geographische Trennung),
- insbesondere schnelle Wiederherstellung der Daten (Training und Übung).

Zur Umsetzung der o. g. Anforderungen im HV-Umfeld eignet sich besonders eine Kombination der verschiedenen in diesem Abschnitt dargestellten Verfahren und Architekturen. Die nachfolgende Abbildung veranschaulicht eine im HV-Umfeld häufig vorzufindende Vorgehensweise. Sie zeigt ein mehrstufiges Datensicherungskonzept im HV-Umfeld. Es gewährleistet sowohl die temporale als auch die funktionale Verfügbarkeit.

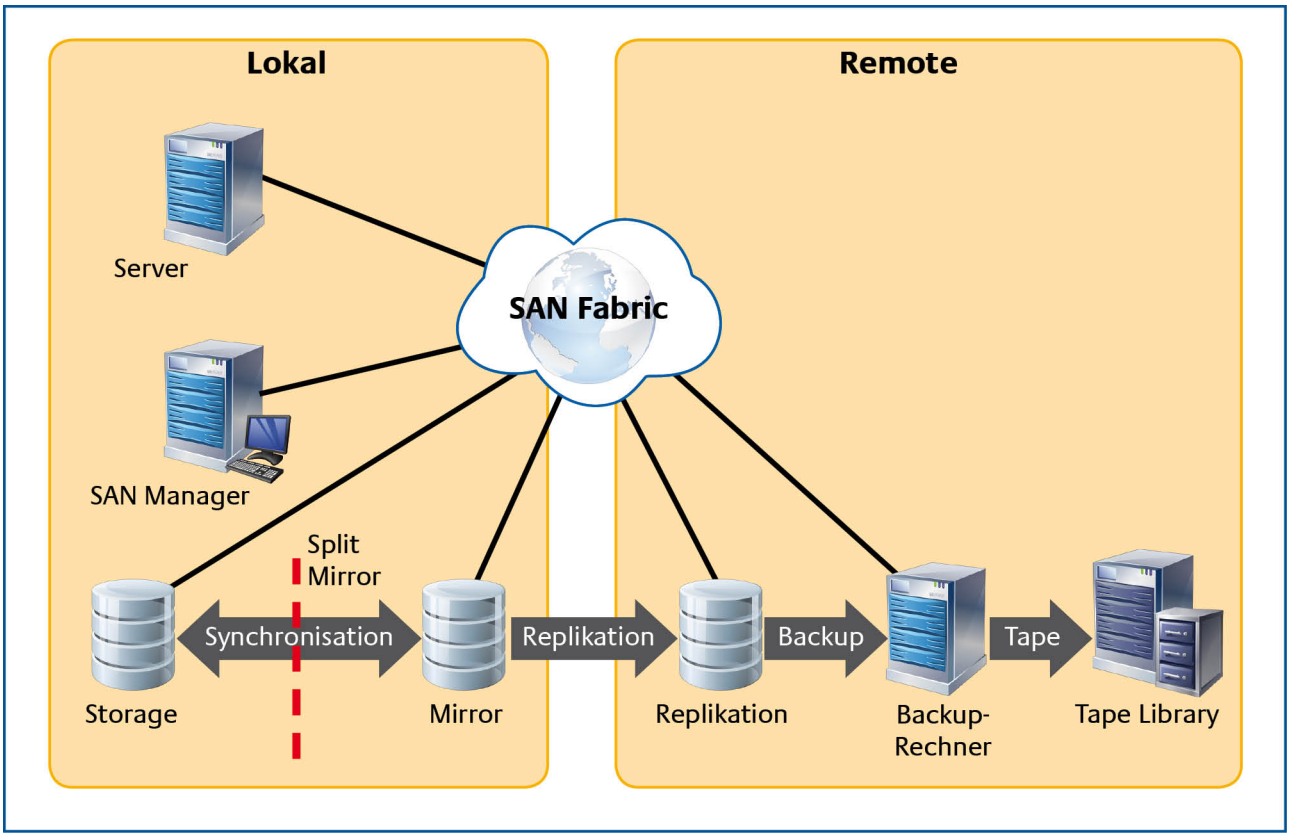


Abbildung 6: Datensicherung im HV-Umfeld

Die temporalen Verfügbarkeitsanforderungen werden in diesem Szenario mittels einem SAN (siehe Abschnitt 1.3.3) und der Spiegelung (siehe Abschnitt 1.7.1) der Daten auf einem weiteren Speichersystem im SAN realisiert. Die Anforderungen hinsichtlich der Datensicherung werden durch eine Kombination von „Replikation – Disk-to-Disk“- (siehe Abschnitt 1.7.1 und 1.7.2), „SAN-Backup“ (siehe Abschnitt 1.7.2) und „Virtueller Vollsicherung“ (siehe Abschnitt 1.7.1) auf Auslagerungsmedien (z. B. Disk-to-Disk-to-Tape, siehe Abschnitt 1.7.2) umgesetzt.

Zunächst wird für eine anstehende Datensicherung die hardwarebasierte Spiegeleinheit kurzzeitig aufgetrennt (*sog. Split Mirror*). Anschließend wird die Datenkopie innerhalb eines SANs mit großer Geschwindigkeit an einen Backup-Rechner übergeben (Replikation auf ein weiteres System). Danach wird die Trennung aufgehoben und der Spiegel resynchronisiert sich mit dem primären System. Die folgende (eigentliche) Datensicherung auf ein Band, kann nun zeitunkritisch erfolgen.

Darf die Auftrennung der hardwarebasierten Spiegeleinheit aufgrund Hochverfügbarkeitsanforderungen in kritischen Anwendungen nur extrem kurzzeitig sein, so wird häufig das Inkrement vom Spiegel zum Backup-Rechner repliziert. Anschließend erfolgt zeitunkritisch eine Vollsicherung und die Übertragung der Daten auf ein Band.

Die Sicherungsmedien sollten nach dem Prinzip der geographischen Redundanz (siehe Beitrag „Prinzipien der Verfügbarkeit“ im HV-Kompendium) in einem geographisch entfernten Ort (mindestens ein weiterer Brandabschnitt) aufbewahrt werden. Sinnvoll erscheint im Zusammenhang mit der oben dargestellten Vorgehensweise, den Backup-Rechner, den Band-Roboter und das Bandarchiv im Auslagerungsort unterzubringen.

1.8 Dateisysteme

Ein Dateisystem ist ein Ordnungs- und Zugriffssystem für die Daten und ermöglicht somit einen gemeinsamen Zugriff auf einen zentralen Datenbestand. Es stellt somit eine Schicht zwischen Festplattensubsystem und Anwendung dar. Zugriffsroutinen für Dateisysteme können Bestandteil eines Betriebssystems sein, wobei das Betriebssystem selbst in einem Dateisystem gespeichert sein kann.

Die Dateinamen sind in speziellen Dateien, den sog. *Dateiverzeichnissen*, abgelegt. Über diese Verzeichnisse kann ein Dateiname und damit eine Datei vom System gefunden werden. Somit identifiziert sich die Datei durch einen entsprechenden Dateinamen und Pfad im Dateisystem.

Für die HV-Umgebung sind verteilte Dateisysteme eine wichtige Voraussetzung, um Ressourcen auf mehrere Rechnerknoten verteilen zu können.

1.8.1 UDF (Universal Disk Format)

UDF steht für „Universal Disk Format“ und beschreibt das Dateisystem (auch High-Level-Formatierung), in welchem die Daten bzw. Dateien auf das Medium geschrieben werden. UDF wurde dabei mit besonderem Augenmerk auf optische Speichermedien entwickelt.

Um die Lebensdauer des Mediums möglichst hoch zu halten ist es daher notwendig, es möglichst gleichmäßig zu belasten.

Zur Nutzung mit Packet-Writing ist das Universal-Disk-Format in Version 1.50 oder höher erforderlich. UDF-Version 1.02 ist ein statisches Dateisystem, vergleichbar mit ISO 9660, das beispielsweise als „ISO/UDF-Bridge“ für DVD-Video zum Einsatz kommt.

Eine der häufigsten Belastungen beim Beschreiben von optischen Medien stellt das Aktualisieren des Inhaltsverzeichnisses dar. Jedesmal, wenn sich eine Datei ändert, und sei es nur der Dateiname, muss diese Änderung ins Inhaltsverzeichnis geschrieben werden. UDF ab Version 1.50 ist deswegen so gut für Packet-Writing geeignet, weil es das Inhaltsverzeichnis an verschiedenen Stellen am optischen Medium speichert und somit die Belastung verteilt.

Ein anderes Problem stellt die Verwaltung des freien Speicherplatzes dar. Auf Festplatten werden Dateien kurzerhand dort geschrieben, wo sonst keine Datei steht – mit anderen Worten dort, wo zum Beispiel gerade eine Datei gelöscht wurde und jetzt wieder freier Speicher ist. Auf optischen Medien würde dies jedoch bedeuten, dass bereits belasteter (da beschriebener) Speicherbereich erneut belastet (da erneut beschrieben) wird. UDF-Versionen 1.50 und höher führen daher eine Liste von Bereichen, welche schon beschrieben wurden und wie oft diese beschrieben wurden. Wenn in einem solchen UDF-Dateisystem eine neue Datei gespeichert wird geschieht dies daher an einer Stelle, die noch nicht oder noch nicht so oft belastet wurde; die Gesamtbelastung des optischen Mediums wird somit in etwa gleich gehalten (sparing table).

Des Weiteren bietet UDF ein Defekt-Management, welches bereits verbrauchte (überbelastete, defekte) Bereiche des verfügbaren Speicherbereichs ausblendet. Das optische Medium kann so weiter verwendet werden, auch wenn bereits Teile davon defekt sind. Die Nutzung dieses Defekt-Managements hängt jedoch von der Software ab, die das UDF-Dateisystem implementiert.

1.8.2 Verteilte Dateisysteme

Der Vorteil sog. verteilter Dateisysteme (oder Netzwerkdateisysteme) ist die Replizierfähigkeit, d. h. die Unterstützung von Kopien von Dateien auf andere Platten und Rechnern, auf die bei

Störungen automatisch zugegriffen werden kann. Ein weiterer Vorteil gegenüber normalen Dateisystemen liegt in der flexiblen Vergabe der Zugriffsrechte, die es unter anderem erlaubt, bestimmte Verzeichnisse ausschließlich ausgewählten Benutzern zugänglich zu machen. Dabei kann zusätzlich zwischen Lese-, Schreib-, Lösch-, Administrations- und Look-Up-Rechten unterschieden werden. Verteilte Dateisysteme stellen dem Benutzer somit Dateiressourcen unabhängig von deren physikalischem Speicherort zur Verfügung (Prinzip der Virtualisierung). Anwender brauchen den tatsächlichen physikalischen Speicherplatz von Dateien nicht zu kennen und anzugeben, um auf die Daten zugreifen zu können. Im nächsten Abschnitt werden bekannte verteilte Dateisysteme kurz dargestellt.

1.8.2.1 Distributed File System (DFS)

Unter einem Distributed File System (DFS) versteht man im Wesentlichen ein Netzwerkprotokoll, welches den Zugriff auf Dateien über ein Rechnernetz ermöglicht. Die Dateispeicherung erfolgt dabei auf mehreren Servern im Netz. Dem Nutzer am Client erscheinen die auf verschiedenen Servern verteilten Dateien einheitlich an einem Ort lokalisiert zu sein. Im Bereich der DFS-Entwicklung hat sich ein kommerzieller und nicht kommerzieller Bereich entwickelt. So existiert im kommerziellen Betriebssystembereich ein entsprechender Serverdienst und im nicht kommerziellen Umfeld eine Implementierung der Open Group[1]. Bezüglich der Verfügbarkeit ist hier bedeutend, dass durch DFS die Nutzer auf redundante Dateien im Netz zugreifen können, so dass ein Ausfall primärer Dateispeicher im Netz keine Auswirkung auf Verfügbarkeit der Datei hat.

1.8.2.2 Andrew File System (AFS)

Das Andrew File System (AFS) wurde zuerst an der Carnegie Mellon Universität (USA) entwickelt. Die dimensionale Ausrichtung geht über die des DFS hinaus. Das AFS skaliert gut (mehrere tausend Workstations, Clients und Fileserver sind keine Seltenheit) und kann über das gesamte Internet entsprechende Ressourcen transparent verwalten. Nicht nur in der Dimension, sondern auch funktional ist das AFS-Konzept eher ganzheitlich orientiert. Neben systemweit angebotenen Funktionen, wie Dateiverwaltung, Benutzerverwaltung und Synchronisation werden im AFS-Konzept funktional die Verwaltung von Clients, Fileserver und Datenbankserver voneinander getrennt, d. h. sie laufen auf physisch verschiedenen Rechnern. Die Authentifizierung der ASF-Clients erfolgt auf dem Fileserver. AFS-Clients sind über alle Betriebssysteme hinweg frei erhältlich. Die Programme, die AFS auf AFS-Server implementieren sind ebenfalls lizenzkostenfrei und unter Linux und anderen Unix-Derivaten lauffähig. Benötigt man spezielle Funktionalitäten auf den Servern, so existieren auch kommerzielle Anbieter.

Wichtige Begriffe im AFS-Umfeld sind Zellen und Volumen. Zellen stellen unabhängige Verwaltungseinheiten aus einem oder mehreren Datenbank- und Fileservern dar, wobei alle AFS-Clients einer bestimmten Zelle zugeordnet sind. Die Verknüpfung von Zellen erfolgt durch sog. Volumeninstanzen. Diese übergeordnete Organisationsstufe mehrerer AFS-Zellen werden relativ autonom verwaltet, können sich aber bezüglich ihres internen Namensraumes gegenseitig vertrauen. Dieses Vertrauen wird durch die gemeinsame Nutzung von Verzeichnissen durch entsprechende Einträge in den Access Control Lists (ACLs) dokumentiert.

1http://www.opengroup.org/

Folgende, die Verfügbarkeit sichernden, Eigenschaften können für das AFS aufgelistet werden:

- Durch die Dateisystemarchitektur ist es möglich Sekundärspeicher ohne Unterbrechung aufzurüsten. Einem Fileserver können mehrere IP-Adressen zugeordnet werden. Dies führt dazu, dass beim Ausfall einer Netzwerkschnittstelle vom AFS-Client aus ein entsprechender Wechsel zu einer anderen Schnittstelle erfolgt.
- Aufgrund der dezentralen Steuerung des AFS (Volumeninstanzen sind autonom), sind Totalausfälle bei entsprechend aufgespanntem AFS sehr selten.

1.8.2.3 Clusterdateisysteme (GFS/GPFS)

Dienste und Anwendungen, welche auf Clustersystemen (siehe Beitrag „Cluster-Verfahren“ im HV-Kompendium) laufen, müssen den parallelen Zugriff auf Daten gewährleisten. Dies kann i. d. R. durch proprietäre Netzwerkdateisysteme wie SAMBA (Implementierung des SMB-Protokolls unter SAMBA siehe M 5.82 (GS-Katalog), oder NFS (siehe M 5.17 (GS-Katalog)) unterstützt werden. Jedoch sind der Datendurchsatz und die Skalierbarkeit, welche diese Protokolle gewährleisten gering. Aus diesem Grund werden im Clusterbereich spezielle Clusterdateisysteme, welche die geforderten Transferraten (> 100 Mbyte/s) und Skalierbarkeit gewährleisten, verwendet.

Clusterdateisysteme müssen folgende Verfügbarkeit sichernde Eigenschaften unterstützen:

- die Verwaltung konkurrierender paralleler Zugriffe durch Sperrmechanismen (Locking),
- die transparente Bereitstellung von redundanten Festplattensystemen innerhalb des SAN (siehe Abschnitt 1.3.3) für die Clients im LAN, denen kein NFS zur Verfügung steht,
- ein ausgefallener Server muss ohne Serviceunterbrechung durch den sog. Fencing-Mechanismus von den Speichersystemen getrennt werden,
- Unterstützung eines clusterweiten Managment-Interfaces zur Erweiterung der Speicherkapazität oder zum Hinzufügen neuer Knoten,
- clusterweite Snapshots,
- synchrone Spiegelung.

Das proprietäre Generell Parallel File System (GPFS) und das frei verfügbare Global File System (GFS) in den jeweils aktuellen Versionen werden diesen Anforderungen zu großen Teilen gerecht. In Abbildung 2 ist ein Implementierungsszenario mit SAN-LAN-Kopplung beispielhaft dargestellt.

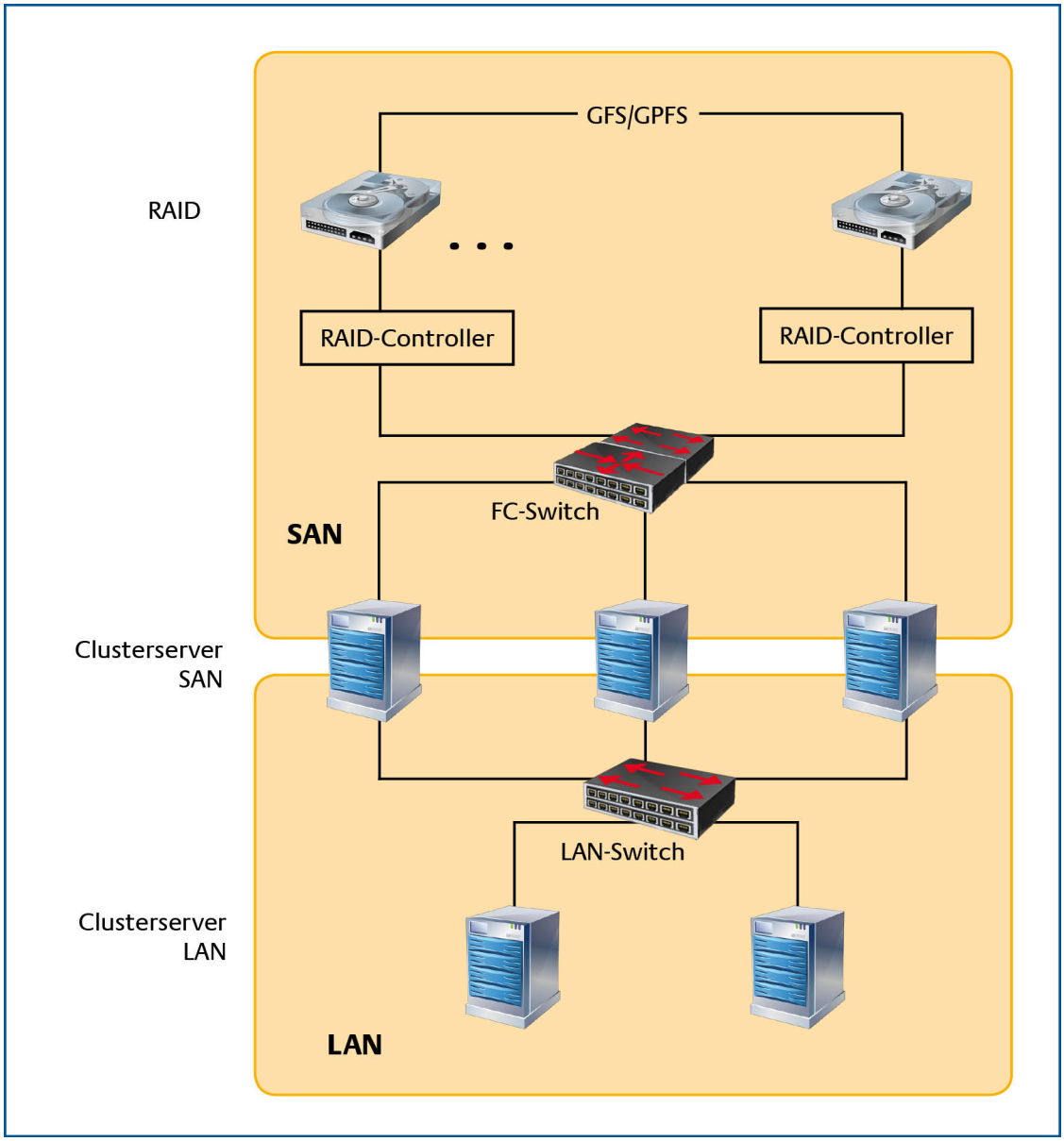


Abbildung 7: Clusterdateiarchitektur (verändert nach [MaTr05])

Wie bei allen Realisierungen im HV-Umfeld ist auch bei der Implementierung von Clusterdateisystemen eine vorausschauende Planung unabdingbar. Dies betrifft die Auswahl und Konfiguration des Speichers (Partitionierung, stripe-size, Blockgröße) ebenso wie die Anpassung der Applikation an die optimale Nutzung des Clusterdateisystems [MaTr05].

1.8.2.4 Network File System (NFS)

Das Network File System (NFS) wird als Erweiterung des Betriebssystems geliefert und wird hauptsächlich auf Unix-Systemen eingesetzt. NFS ermöglicht das Zuordnen von Dateisystemen entfernter Rechner auf lokale Dateisysteme. Dadurch können Benutzer auf Dateien zugreifen als ob sie auf ihren lokalen Festplatten abgespeichert wären. Neben den Dateisystemen können auch Drucker über das Netzwerk anderen Anwendern zugänglich gemacht werden.

1.8.2.5 Common Internet File System (CIFS)

Das Common Internet File System (CIFS) ist eine Weiterentwicklung des SMB-Protokolls, das Microsoft als offenes Protokoll deklariert und an die IETF übergeben hat. CIFS soll dabei eine Ergänzung und Verbesserung für die Standardprotokolle FTP und HTTP als Netzwerk-Dateisystem für Internet-Anwendungen darstellen. CIFS stellt den Benutzern im ganzen Netzwerk genau bezeichnete Festplattenbereiche zur Verfügung.

1.8.2.6 Direct Access Filesystem (DAFS)

Das Direct Access Filesystem wurde von der DAFS-Collaborative entwickelt und bietet gegenüber NFS und CIFS eine verbesserte Verfügbarkeit, Performanz und Zuverlässigkeit. Es beinhaltet einen vom Betriebssystem unabhängigen Memory-to-Memory-Zugriff auf lokale Dateisysteme. Daten können unter „Umgehung“ der CPU direkt vom Speicher eines Systems auf ein anderes System übertragen werden. Sofern die Netzwerkkarten das DAT (Direct-Access-Transport)-Protokoll unterstützen, können Daten über das Netzwerk direkt in den Hauptspeicher geschrieben werden, was eine Steigerung der Performanz und geringere Belastung der CPU bedeutet.

Sowohl NFS als auch CIFS bedienen sich der Standard-Netzwerkprotokolle (TCP/IP) und sind daher verstärkt anfällig gegenüber Latenzzeiten oder unzulänglichen Verbindungen innerhalb des Netzwerkes. Das von DAFS genutzte DAT-Protokoll sowie RDMA (Remote-DMA) bietet hingegen direkten Speicherzugriff und minimalen Overhead im Protokoll. Die Verbreitung von DAFS wurde aber durch neue Varianten von NFSv4 (pNFS) beeinflusst, da entsprechende Vorteile der Implementierungen von DAFS (Steigerung der Skalierbarkeit und Unabhängigkeit von Betriebs- und Speichersystem durch Virtualisierung) durch parallelen Zugriff in pNFS auch möglich wurden. Ein weiteres wesentliches Hemmnis bei der Verbreitung von DAFS ist die Notwendigkeit entsprechender DAFS-fähiger Server [TrEr03].

1.8.2.7 Zusammenfassender Vergleich wichtiger Dateisysteme

Merkmal	*Dateisystem*				
	AFS	***NFS***	***GFS***	***CIFS***	***DAFS***
basiert auf Protokoll	UDP	UDP / TCP /IP	iSCSI/FC	TCP / IP	RDMA/ InfiniBand
Locking	auf Dateibasis	bis NFSv2 kein Locking, ab NFSv4 Locking auf Dateibasis	auf Byte-Ebene	Betriebs-systemebene	auf Dateibasis
Caching	lokal	übers Netz	übers Netz	übers Netz	lokal
Multipathing	nein	nein	ja	nein	nein
Journaling	nein	nein	ja	nein	nein
Access Control	ACL	ACL (NFSv4)	ACL	ACL	lokal
Authentifizierung	Kerberos-Protokoll	DES-Protokoll, RPCSEC (NFSv4)	nein	SMB-Protokoll	ja
Einsatz LAN / WAN	ja / ja	ja / nein	ja / nein	ja / ja (WAFS)	ja / nein
Dateigröße (maximal)	2 GB	2 GB	ca. 72 x 10^6 TByte	~2 GB	~2 GB
Recovery	ja	ja (NFSv4)	ja	nein	nein
Redundanz	ja	ja	ja	nein	nein
Snapshots	ja	ja	ja	ja	nicht bekannt
Lastverteilung	ja	nein	ja	nein	nein
Skalierbarkeit	sehr hoch	gering	mittel	mittel	mittel

Tabelle 6: Vergleich verschiedener verteilter Dateisysteme unter Verfügbarkeit beeinflussenden Parametern

Vergleicht man die vorab dargestellten Dateisysteme (siehe Tabelle 6), so wird deutlich, dass die klassischen Clusterdateisysteme Eigenschaften im Protokoll aufweisen, die die anderen Dateisysteme nur durch zusätzliche, zumeist kommerzielle Software erreichen können. Dazu zählen das Multipathing und Journaling, welche beim GFS ohne zusätzliche Software eine hohe Ausfallsicherheit ermöglichen. Ein Clustering ist hier nicht nötig.

Eine Zugriffskontrolle ist bei nahezu allen Dateisystemen durch entsprechende ACLs realisiert, wobei dies bei GFS nur durch zusätzlichen Protokoll / Software-Aufwand umgesetzt werden kann. Für die Verfügbarkeit von entscheidender Bedeutung ist die Unterstützung von Recovery-Mechanismen. Beim NFSv4 wird dies unterstützt. Hier ist AFS durch implementierte Recovery-Mechanismen auf Dateiebene bis Volumen-Ebene hervorzuheben. Die Dateisysteme AFS, NFSv4 und GFS können zeitnah auf verschiedenen Fileservern replizieren, so dass redundante Daten vorgehalten werden können. Beim GFS sind redundante serverbasierte Cluster und ein entsprechender Log-Manager durch eine zusätzliche Implementierung möglich, aber dann entsprechend komplex zu konfigurieren.

In seinen Eigenschaften bemerkenswert ist das AFS. Sowohl hinsichtlich der Skalierbarkeit, Last-verteilung und weiterer Verfügbarkeit sichernder Parameter ist es heutzutage ein etabliertes Protokoll im WAN-Bereich. Jedoch erscheint die Implementierung und Administration im LAN zu aufwändig. Es wurde eine weitere LAN-Variante des AFS entwickelt (CODA [Brue04]), die weitere Verfügbarkeitsaspekte abdeckt. Ein möglicher SPoF des AFS ist die Authentifizierungskomponente mit Kerberos, welches serverbasiert entsprechend redundant

vorgehalten werden muss. Obwohl verteilte Dateisysteme wie NFSv4 und CIFS ohne entsprechendes Clustern der Server auch zum SPoF werden können, sind sie verbreitet.

1.9 Speicher-Management

In der Einleitung dieses Beitrags wurde bereits auf die hohen Anforderungen an die Speicherlösung im HV-Umfeld hingewiesen. Dies erfordert nicht nur ein ganzheitliches Konzept, sondern auch ein übergeordnetes Management. Die Zielsetzung beim Speicher-Management kann man in drei Kernaussagen zusammenfassen:

1. Speicher-Management soll die Sicherheit der Daten gewährleisten.
2. Speicher-Management soll für die Verwaltbarkeit der Daten sorgen.
3. Speicher-Management soll eine effiziente Datenhaltung ermöglichen.

Erste Anforderungen hinsichtlich der Datensicherheit lassen sich aus den Ergebnissen der SOLL-Analyse ableiten, müssen jedoch für den Speicherbereich konkretisiert werden (siehe auch Überlegungen zur Wiederherstellung im Abschnitt 1.7.4).

Verwaltung und Organisation der Speicherkomponenten, dynamische Zuteilung von Ressourcen, Datensicherung und Wiederherstellung, Archivierung und Optimierung der Speicherinfrastruktur, dies alles sind Aufgaben, die vom Speicher-Management geleistet werden müssen. Darüber hinaus bedient das Speicher-Management an der Schnittstelle zu anderen IT-Prozessen (z. B. Kapazitäts-, Verfügbarkeits- oder Continuity-Management) die allgemeine IT-Organisation mit Management-Dienstleistungen. In der sind die Bausteine dargestellt, welche die wesentlichen Aufgabenbereiche beim Speicher-Management repräsentieren.

1.9.1 Speicher-Ressourcen-Management

Das Speicher-Ressourcen-Management (Storage Ressource Management - SRM) ermöglicht die Verwaltung der Speicherressourcen durch ihre Steuerung und Überwachung. Vor dem Hintergrund der speziellen Verfügbarkeitsanforderungen im HV-Umfeld kann als Ziel in erster Linie die Minimierung der Ausfallzeiten und kontinuierliche Verfügbarkeit der Daten gesehen werden. Mit Blick auf die von den Datensicherungsprozessen betroffenen Anwendungen, wie auch auf die Dauer der Wiederherstellung verlorener Daten, sind hier (wenn überhaupt) nur kleine Zeitfenster für diese Prozesse tolerierbar.

Wesentliche Voraussetzung für eine Vielzahl von Aufgaben der Speicherverwaltung ist ein aktueller Überblick über die vorhandenen Ressourcen. Dies setzt einerseits eine umfassende Katalogisierung aller Speicherressourcen (Plattensysteme, Medien, Netzkomponenten etc.) und andererseits ein auf die eingesetzten Speicherressourcen abgestimmtes Überwachungssystem voraus. Sowohl reaktives als auch proaktives Eingreifen in die Speicherinfrastruktur wird dadurch erst ermöglicht. Wurde eine IT-Strukturanalyse nach IT-Grundschutz durchgeführt, so liegen mit der daraus resultierenden Ergebnisdokumentation bereits die wesentlichen Informationen vor. Mittelfristig sollten alle Speicherressourcen im Rahmen eines Speicher-Ressourcen-Management erfasst, überwacht und analysiert werden.

Eine weitere Ausbaustufe stellt der Einsatz von Automatismen dar. Das Eingreifen aufgrund erkannter Ereignisse erfolgt dann nicht mehr manuell, sondern automatisiert nach zuvor definierten Regeln (siehe auch Abschnitt 1.9.2.3).

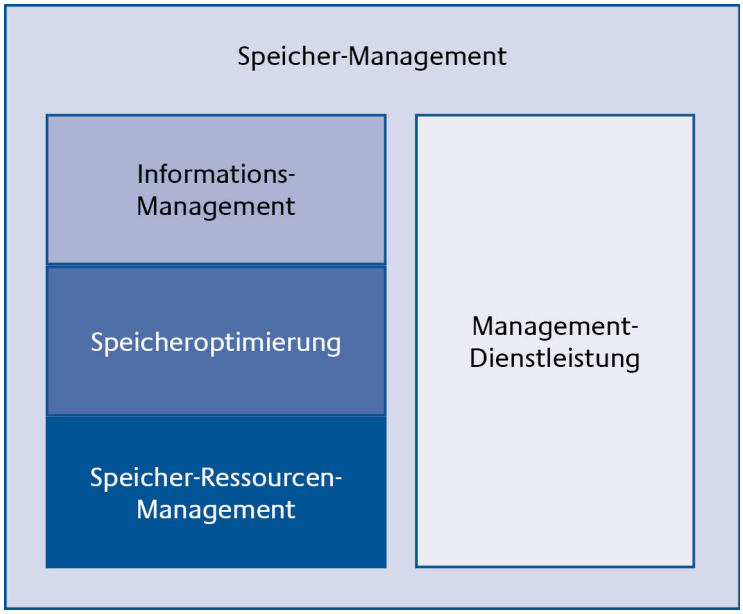


Abbildung 8: Bausteine des Speichermanagements (vgl. [Froe04]

Die nächste Umsetzungsstufe wird durch die Verwendung von Speicher-Management-Tools erreicht. Mit „intelligenten“ Speicher-Management-Tools ist eine einfache (in der Bedienung) und zuverlässige Zuweisung der Ressourcen über das gesamte Netzwerk möglich. Speicherdienste, die in konventionellen Systemen an einzelne Komponenten oder Teilsysteme gebunden waren, können somit netzwerkweit, plattformunabhängig bereitgestellt und bei Bedarf genutzt werden. Bezogen auf die benötigte Information heißt dies, dass zu jeder Zeit, egal mit welchem Gerät, ein Zugriff auf die relevanten Daten möglich ist. Dies setzt jedoch eine entsprechende Speicherarchitektur (z. B. SAN) voraus.

Die marktüblichen Software-Produkte für das Speicher-Ressourcen-Management ermöglichen zusätzlich eine konsolidierte Übersicht aller Speicherressourcen im HV-Verbund. Dabei arbeitet die Software als verteilter, agentenbasierter Datensammler, der alle relevanten Informationen über Speicherressourcen einholt, diese analysiert und in einer zentralen Datenbank ablegt. Diese Daten können für Reports und Trendanalysen verwendet werden und dienen u. a. der Planung des Speicherwachstums. Somit ist das Speicher-Ressourcen-Management auch als Basis des nächsten Bausteins, der Speicheroptimierung, zu sehen.

1.9.2 Speicheroptimierung

Speicheroptimierung konzentriert sich auf die Optimierung bestehender Speicherinfrastrukturen- und -Prozesse. Entsprechend der eingangs formulierten Zielvorstellung, muss das Speicher-Management dafür sorgen, dass die Speicherinfrastruktur den Anforderungen gerecht wird. Mit geeigneten Analysewerkzeugen wird die Datenbasis aus dem Speicher-Ressourcen-Management ausgewertet und die Auslastung der bestehenden Umgebung ermittelt (IST-Feststellung). Zusätzlich wird das Optimierungspotenzial der vorhandenen Lösung untersucht, um Ansätze für eine dauerhafte Verbesserung zu finden (z. B. höherer Automatisierungsgrad oder Virtualisierung). Die gefundenen Lösungsansätze bilden die Grundlage für einen Strategieentwurf.

1.9.2.1 Strategien

Je wichtiger Daten für einen (kritischen) Geschäftsprozess sind, umso höher muss i. d. R. ihre Verfügbarkeit sein. Die Speicherstrategie muss demzufolge auf dem Prinzip der Priorisierung (siehe Beitrag „Prinzipien der Verfügbarkeit“ im HV-Kompendium) aufbauen. Das undifferenzierte Speichern auf ein einziges Speichersystem würde der unterschiedlichen Wertigkeit von Daten nicht gerecht und wäre in hohem Maße unwirtschaftlich. Es hat sich gezeigt, das hierarchische Speicherstrategien das notwendige Kosten- / Nutzen-Verhältnis der zu sichernden Daten garantieren.

1.9.2.2 Speicherhierarchien

Die Einteilung der unterschiedlichen Speichertechnologien erfolgt in drei Stufen. Mit der Darstellung dieser Speicherhierarchien (siehe Abbildung 9) werden Leistungs- und Kostenunterschiede der Technologien deutlich.

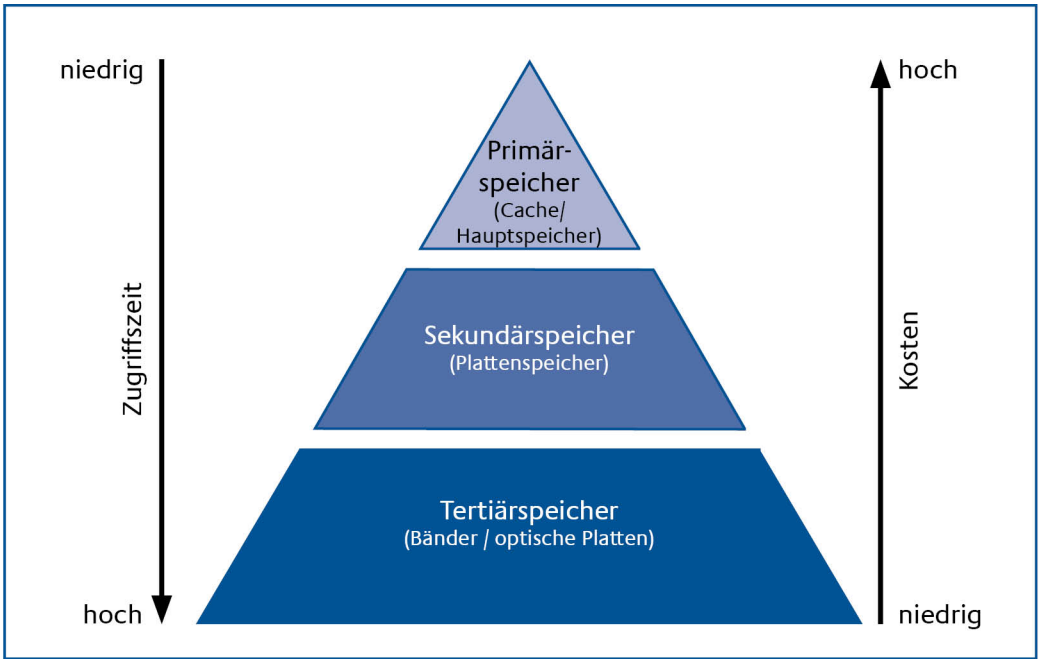


Abbildung 9: Speicherhierarchie

Die Speicherhierarchie beginnt mit dem hoch performanten, primären Speicher (auch Online-Bereich genannt). Die Zugriffszeiten liegen beim Primärspeicher im Nanosekundenbereich und Kosten pro Megabyte sind hier vergleichsweise hoch. Der Sekundärspeicher (auch sekundärer Online-Bereich genannt) bildet in dieser Speicherhierarchie die mittlere Stufe. Der Sekundärspeicher hat ebenfalls Direktzugriffseigenschaften mit Zugriffszeiten im Millisekundenbereich, ist aber mit hoch kapazitiven Festplatten mit geringeren Kosten pro Megabyte aufgebaut. In der Regel basieren diese Systeme auf günstigen Plattentechnologien, wie beispielsweise ATA-Festplatten mit einer geringeren Performanz im Unterschied zum Primärspeicher. Eine gute Eignung des Sekundärspeichers ist bei Backup-Funktionen und in der Archivierung zu sehen. Manche Speicher-Systeme können gleichzeitig Plattenbereiche für die primäre und sekundäre Stufe haben. Das Verhältnis ist wahlweise konfigurierbar und erlaubt die Speicherhierarchie optimal an die Backup- bzw. Archivierungs-Szenarien anzupassen. Am anderen Ende der Hierarchie steht der Tertiärspeicher (auch Nearline-Bereich) mit wesentlich geringerem Preislevel pro Megabyte. In dieser Hierarchiestufe liegen die Zugriffzeiten im Bereich von einigen Sekunden bis in den Minutenbereich. Als Tertiärspeicher können beispielsweise unterschiedliche

Bandtechnologien und magneto-optische Aufzeichnungsverfahren (CD, DVD, Ultra Density Optical etc.) eingesetzt werden.

1.9.2.3 Automatismen und Speichervirtualisierung

Für eine optimale Ausnutzung der Hierarchien aus dem Abschnitt 1.9.2.2 ist ein einheitliches, automatisiertes Speicher-Management erforderlich. Entsprechend der zuvor festgelegten Speicherprozesse, wird mit den Daten gemäß ihrer Wertigkeit verfahren (siehe Abschnitt 1.9.3.1). Die für den Geschäftsprozess erforderlichen Daten können durch gezielte Automatismen auf alle in der HV-Umgebung vorhandenen Speicher verteilt werden. Jedoch nicht nur für die Verteilung der Daten auf die einzelnen Speicherhierarchien sind Automatismen notwendig, sie stellen beispielsweise sicher, dass ein geplanter Datensicherungslauf erfolgreich durchgeführt wird. Wird die Datensicherung aufgrund einer Störung abgebrochen, so ist zumindest eine entsprechende Benachrichtigung an den Administrator erforderlich. Im optimalen Fall wird der Datensicherungslauf jedoch nach Behebung der Störung erneut automatisiert gestartet.

Je automatisierter die Datenbewegung zwischen den Hierarchiestufen in heterogenen Umgebungen erfolgt, desto wirksamer kann Speicher-Management greifen.

Die Komplexität von heterogenen Speicherinfrastrukturen wird durch die Einführung von Speicherhierarchien und -klassen noch gesteigert. Die Speichervirtualisierung stellt ein Verfahren dar, mit dem diese Komplexität „vordergründig“ aufgehoben wird. Dabei werden durch eine zusätzliche (Abstrahierungs-)Schicht sowie durch einen Mechanismus, der steuernd eingreift, die Grenzen einzelner Speichergeräte aufgehoben. So ist es möglich den „Datenpool" homogen zu verwalten, obwohl er heterogen gehalten wird. Dies vereinfacht die Administration und ermöglicht eine flexiblere Verteilung von Speicherressourcen (siehe Abschnitt 1.9.1).

Hersteller- und geräteunabhängige Automatismen zum Aufbau, zur Verwaltung und zum Betrieb von Speicherstrukturen können mittels Speichervirtualisierung zentral verwendet und auf alle Speicherobjekte angewandt werden.

In einer virtualisierten Speicherumgebung ist schließlich auch die automatisierte Verteilung und Bewegung der Daten auf Basis der benötigten Verfügbarkeit und der geforderten Antwortzeit (definiert in Speicherklassen) realisierbar. Eine Klassifizierung der Daten erfolgt im Rahmen des Informations-Managements.

1.9.3 Informations-Management

Informations-Management stellt den Rahmen für die Information-Lifecycle-Management-Lösungen (ILM) dar. Speicher wird hier aus der Sicht der Geschäftsprozesse betrachtet und in Informationskategorien eingeteilt. Die Informationen werden, basierend auf den Anwenderanforderungen, klassifiziert und kategorisiert. Die SRM-Automatismen sorgen anschließend dafür, dass die Daten möglichst effizient auf die jeweils wirtschaftlichste Speicherhierarchie, gemäß definierten Regelwerken verschoben werden.

1.9.3.1 Klassifizierung

Die Einführung von Speicherhierarchien ermöglicht einen differenzierten Umgang mit Daten bzgl. ihres Speicherplatzes. Die Daten selbst müssen nun auch entsprechend der Skalierung bei der Speicherhierarchie in Klassen eingeteilt werden. Die Aufgabe besteht darin, den „Wert“, welchen die Daten für die Organisation darstellen zu ermitteln. Eine Klassifizierung der Daten im HV-Umfeld ist stark durch die Einstufung des Geschäftsprozesses in eine Verfügbarkeitsklasse

(siehe Einführung im HV-Kompendium) geprägt, d. h. der GP gibt bereits den Schutzbedarf hinsichtlich des Sicherheitswertes Verfügbarkeit für die innerhalb des GP zu verarbeitenden Informationen vor.

Daten sind demnach an dem Ort abzulegen, der ihrer Bedeutung für die Organisation / den Geschäftsprozess zukommt (z. B. Daten mit besonderer Bedeutung auf Online-Speichern, Kopien dieser Daten auf Nearline-Arrays etc.). Damit es zu einer realistischen Einschätzung hinsichtlich des Wertes der Daten kommt, müssen die Nutzer (Anwender) und Anbieter (IT-Abteilung) der Speicherdienste gemeinsam an der Klassifizierung arbeiten. Die Prozessbeteiligten müssen vorgeben, welche Daten welche Priorität haben und wie verfügbar diese sein müssen.

Eine Klassifizierung ist zudem unerlässlich, wenn es um die Wiederherstellung nach einem Systemausfall geht. Damit wichtige Prozesse zeitnah nach einem Ausfall zur Verfügung stehen, muss geklärt werden, welche Daten dafür erforderlich sind. Diese sind höherwertig einzustufen und bei der Wiederherstellung bevorzugt zu behandeln.

Eine Klassifizierung nach folgenden Kriterien ist möglich:

- Zugriffshäufigkeit
- Alter
- Dateitypen
- Dateninhalte

Eine automatisierte Klassifizierung unter Verwendung spezieller Klassifizierungsprogrammen ist wünschenswert, aber insbesondere bei einer Bewertung nach dem Inhalt nur schwierig zu realisieren.

Daten können nach einer gewissen Zeit an Wert verlieren und sollten deshalb möglichst automatisch den bevorzugten und teuren Speicherplatz räumen (z. B. „alte“ E-Mails). Durch Änderungen am Geschäftsprozess oder an den IT-Ressourcen entstehen neue Informationen, die wiederum identifiziert und klassifiziert werden müssen. Die Wertigkeit anderer Informationen kann sich dadurch verändern. Die Klassifizierung der Informationen ist somit kein einmaliger Vorgang, sondern ein fortlaufender Prozess. Textanalysetools und Metadatenerhebung können diesen Prozess unterstützen.

1.9.3.2 Information Lifecycle Management (ILM)

Die Zielvorstellung beim Speicher-Management ist, dass das Speicher-Management-System den gesamten Speicher-Lifecycle – von Design und Implementierung der Konfiguration bis hin zur Kapazitätskontrolle und der Optimierung der Performanz – aus einem Guss verwaltet und dabei alle wichtigen Systemplattformen, Datenbanken und Anwendungen abdeckt. Ein solcher Lösungsansatz wird durch die Methodensammlung des Information Lifecycle Managements (ILM) beschrieben. Information Lifecycle Management ist eine Kombination aus Prozessen und Technologien [Bitk04] und setzt sich aus Elementen des Workflow-Management, dem Dokumenten-Management, Content-Management und dem Speicher-Management zusammen. Soll ILM greifen, müssen diese Elemente zu einer ganzheitlichen Lösung zusammengeführt werden, mit dem Ziel, über definierte Zeiträume hinweg jeweils den kostengünstigsten Speicherort oder das ideale Medium zu verwenden. Besonders bei der automatisierten Klassifizierung von Daten kann ein Content-Management-System mit seiner Funktionalität zur Auswertung von Metadaten einen wichtigen Beitrag leisten.

1.9.4 Management-Dienstleistung

Durch die Management-Dienstleistungen werden Aufgaben beschrieben, die sowohl übergeordneter Natur (z. B. Unterstützung beim Incident-Management) oder auch spezielle Einzelaufgaben (z. B. Einrichten und Absichern von Remote-Zugriff) sind. Die Management-Dienstleistungen stellen auch die Verbindung zu anderen Bereichen der HV-Umgebung auf organisatorischer Ebene dar.

Speicher-Management ermöglicht die Strukturierung der Speicherressourcen, setzt aber auch organisatorische Regeln voraus. Nach der Einteilung der Speicherinfrastruktur in Hierarchien und der Klassifizierung der Daten geht es zuletzt darum, geschäftsorientierte Regeln (Policies) auf die gesamte heterogene IT-Infrastruktur anzuwenden und so die richtigen IT-Ressourcen zur richtigen Zeit dem entsprechenden Service-Level automatisiert zuzuordnen. Daten bewegen, Daten sichern und Daten verwalten muss nach standardisierten Prozessen ablaufen, die effektiv und sicher gesteuert werden können. Für jeden Prozess ist eine Beschreibung der Aktionen sowie der Schnittstellen erforderlich. Im Zusammenhang mit der Beschreibung der Speicherprozesse ist auch ein Rollenkonzept zu definieren. Dieses regelt für die einzelnen Prozesse sowohl den Zugriffsschutz als auch die Verantwortlichkeiten.

Ein konsequent ausgestaltetes Speicher-Management führt dazu, dass für diesen Bereich Standard-prozesse und –rollen definiert werden, welche in die Gesamtorganisation (z. B. nach ITIL oder CobiT) der IT-Ressourcen eingebunden sind. Ereignisse aus den Bereichen Service Delivery und Service Support beeinflussen auch immer das Speicher-Management. Erst durch die gemeinsame Nutzung aller Schnittstellen und wichtiger Funktionen in Form von „Common Services“ ist hinsichtlich des Speicher-Management eine Ausbaustufe erreicht, welche die Voraussetzung für eine hoch verfügbare Speicherlösung schafft.

1.9.5 Kombination von Datensicherungsverfahren

Speicher-Management soll die Aufrechterhaltung des Betriebs im Sinne von „Business Continuity“ sicherstellen, zusätzlich aber auch die Systemwiederherstellung im Katastrophenfall unterstützen. Letzteres ist beispielsweise mithilfe geeigneter Datensicherungsverfahren umsetzbar. Gerade in konsolidierten Umgebungen muss die Datensicherung wegen der gewaltigen Kapazitäten und der zahlreichen Anwender höchsten Standards bzgl. der Verfügbarkeit und vor allem auch der Performanz genügen. Die Einführung von Speichervirtualisierung und die Konsolidierung der Speicherinfrastruktur zu einer Lösung mit zentralisiertem Speicher-Managements schafft die Bedingungen für eine nahezu beliebige Kombination von Datensicherungsverfahren (siehe Abschnitt 1.7). Ihre Eignung hängt vom jeweiligen konkreten Anwendungsszenario ab. Grundsätzlich sollte jedoch insbesondere im HV-Umfeld darauf geachtet werden, dass Datensicherung:

1. ohne Störung des Produktivsystems,
2. nach Klassifizierung der Daten (optimale Auswahl der Speichermedien),
3. nach einem abgestimmtem Recovery-Plan,
4. mit geografischer Trennung,
5. und mit kontinuierlicher Speicherung aktualisierter Daten

erfolgt. Ein möglicher Lösungsansatz, der diese Aspekte berücksichtigt, ist bei Continuous Data Protection (CDP) zu finden. CDP ist eine Methode, die kontinuierlich Veränderungen am

Datensatz übernimmt oder mitloggt und diese Veränderungen unter Beibehaltung der Originaldaten speichert. Der Datensatz eines beliebigen Zeitpunkts kann so wieder hergestellt werden.

1.9.6 Auswahl von Technologien und Produkten

Die Elemente für das Speicher-Management setzen sich aus Hardware- und Software-Komponenten zusammen. Die Vielzahl an verfügbaren Technologien und Produkten, die zusätzlich durch die unterschiedliche Herkunft der Anbieter geprägt wird, lässt den Markt für Speicher-Management unübersichtlich erscheinen. Zusätzlich gelten auf jeder Schicht von Datenbank-, Datei- und Volume Manager über Host-Bus-Adapter, Zonen und Logical Units bis hin zur Festplatte verschiedene Regeln und Vorgehensweisen. Diese Tatsache steht dem Bestreben nach Vereinheitlichung und zentraler Verwaltung in einer Software entgegen. Die Storage Management Initiative (SMI) innerhalb der Hersteller-Organisation SNIA (Storage Networking Industry Association) bemüht sich daher eine einheitliche Terminologie und einheitliche Schnittstellen für Speicher-Management-Produkte zu definieren. Zu den Vorteilen SMI-S-fähiger Produkten [Snia06] zählen verbesserte Datensicherheit und -verwaltung durch standardbasiertes Management und herstellerübergreifender Support. Neben Kompatibilität und Administrierbarkeit sind für den Einsatz im HV-Umfeld Funktionalitäten zum Schutz vor dem Systemausfall wichtig. Das Speicher-Management-System muss daher Dienste wie Schranküberwachung (SES) unterstützen. Alle notwendigen Speicherprozesse sollten automatisiert vom Management-System steuerbar sein.

Anhang: Verzeichnisse

Abkürzungsverzeichnis

Ein komplettes Verzeichnis hierzu findet sich in Band AH, Kapitel 5

Glossar

Ein komplettes Verzeichnis hierzu findet sich in Band AH, Kapitel 6

Literaturverzeichnis

Ein komplettes Verzeichnis hierzu findet sich in Band AH, Kapitel 7