

Bachelorthesis

Konzeptionierung und Realisierung der Erneuerung des Storage Area Network (SAN) eines mittelständischen Unternehmens zur Steigerung der Ausfallsicherheit der IT-Infrastruktur

Vorgelegt am: 18. August 2014

Von: **Dominic Ernst**
August-Bebel-Siedlung 14
04736 Waldheim

Studiengang: Technische Informatik

Studienrichtung: Daten- und Kommunikationstechnik

Seminargruppe: 4TI11-1

Matrikelnummer: 4000033

Praxispartner: ASTA-PICCA IT-Systeme GmbH
Bahnhofstraße 5
04668 Grimma

Gutachter: Dipl.-Ing. (BA) Mathias Schubert
ASTA-PICCA IT-Systeme GmbH

Dipl.-Inf. Falk Puschmann
Staatliche Studienakademie Glauchau

Freigabeerklärung

Hiermit erklären wir uns einverstanden/~~nicht einverstanden~~*), dass die Bachelorthesis / ~~Diplomarbeit~~ *) der/des Studenten/in

Name, Vorname: Ernst, Dominic SG: TI11

zur öffentlichen Einsichtnahme durch den Dokumentenserver der Bibliothek der Staatlichen Studienakademie Glauchau bereitgestellt wird.

Thema der Arbeit:

Konzeptionierung und Realisierung der Erneuerung des
Storage Area Network (SAN) eines mittelständischen Unternehmens
zur Steigerung der Ausfallsicherheit der IT-Infrastruktur
.....

Grimma, 11.08.2014

Ort, Datum

D. Ernst
.....
Unterschrift Student/in

**ASTA
PICCA**
IT-SYSTEME GmbH
Bahnhofstr. 5, 04668 Grimma
Tel. (03437) 71 77-0, Fax 71 77-77

[Signature]
.....
Stempel, Unterschrift Bildungsstätte

*) Nichtzutreffendes bitte streichen

Themenblatt Bachelorthesis

Studiengang Technische Informatik

Student: **Dominic Ernst**
Matrikelnummer: **4000033**
Seminargruppe: **4TI11-1**

Thema der Bachelorthesis

Konzeptionierung und Realisierung der Erneuerung des Storage Area Network (SAN) eines mittelständischen Unternehmens zur Steigerung der Ausfallsicherheit der IT-Infrastruktur

Gutachter/ Betreuer: Dipl.-Ing. (BA) Mathias Schubert
Gutachter (Studienakademie): Dipl.-Inf. Falk Puschmann

Ausgabe des Themas: **20.05.2014**
Abgabe der Arbeit an den SG am: **18.08.2014, 14:00:00**



Prof. Dagmar Menzel
Vorsitzende des Prüfungsausschusses
Technik

www.ba-glauchau.de



Inhaltsverzeichnis

1	Einleitung	1
2	Analyse Ist-Zustand	2
3	Theorie möglicher Techniken	5
3.1	Einführung	5
3.2	SAN-Konzepte im Überblick	7
3.2.1	iSCSI	7
3.2.2	Fibre Channel	9
3.2.3	Fibre Channel over Ethernet	12
3.3	SAN-Topologien	13
3.3.1	Allgemeines zu Topologien	13
3.3.2	Point to Point	13
3.3.3	Loop	14
3.3.4	Fabric	14
3.3.5	Full Fabric	15
3.4	Hochverfügbarkeit	16
4	Projekt-Realisierung	18
4.1	Planung	18
4.1.1	Analyse der Anforderungen	18
4.1.2	Auswahl der Technologie	20
4.1.3	Dimensionierung des SAN	23
4.1.4	SAN-Entwurf	25
4.1.5	Auswahl geeigneter Komponenten	26
4.2	Auftrag	31
4.3	Umsetzung	32
4.3.1	Festlegungen für das SAN-Netz	32
4.3.2	Vorinstallation der Server	33
4.3.3	Vorkonfiguration der Storages	34
4.3.4	Aufbau der Hardware	37
4.3.5	Konfiguration der Komponenten	37
4.3.6	Migration vorhandener Daten	40
4.3.7	Funktionstests	42
4.3.8	Einrichtung des Monitoring	42

4.4	Probleme	43
4.4.1	Migration der Daten	43
4.4.2	Kopiervorgang mit ESXi	43
4.4.3	Übersehene USB-Hardware	43
5	Abschluss	44
5.1	Übergabe an den Kunden	44
5.2	Projekt-Ergebnis	44
5.3	Ausblick	45
	Literaturverzeichnis	46

Abbildungsverzeichnis

Abbildung 1:	Ist-Zustand des Kundennetzwerkes	2
Abbildung 2:	Vergleich Client-Server-Architektur / Datenhaltung im SAN	5
Abbildung 3:	SCSI-Bus	8
Abbildung 4:	Paketkapselung von iSCSI	9
Abbildung 5:	Point to Point-Topologie	13
Abbildung 6:	Loop-Topologie	14
Abbildung 7:	Fabric-Topologie	14
Abbildung 8:	Full Fabric-Topologie	15
Abbildung 9:	Protokolleffizienz der SAN-Konzepte	20
Abbildung 10:	Durchsatz der SAN-Konzepte	21
Abbildung 11:	Maximale Kabellängen von Twisted Pair und LWL	22
Abbildung 12:	SAN-Entwurf	25
Abbildung 13:	Fujitsu PRIMERGY RX300 S8	26
Abbildung 14:	Fujitsu ETERNUS DX100 S3	28
Abbildung 15:	D-Link DXS-3600-16S	30
Abbildung 16:	UFM-Modul	33
Abbildung 17:	Boot von der ESXi-CD-ROM	33
Abbildung 18:	ESXi-Konfiguration mittels vSphere-Client	34
Abbildung 19:	iRMC-Einstellungen im BIOS	37
Abbildung 20:	IP-Konfiguration bei ESXi	38
Abbildung 21:	Datastore-Einrichtung in ESXi	39

Tabellenverzeichnis

Tabelle 1: Verfügbarkeitsklassen	16
Tabelle 2: IP-Adressen im neuen SAN	32

Formelverzeichnis

Formel 1: Verfügbarkeit	16
Formel 2: Protokolleffizienz	20
Formel 3: RAID 5-Kapazität	23
Formel 4: Übertragungsdauer	41

Abkürzungsverzeichnis

ANSI	American National Standards Institute
BIOS	Basic Input Output System
CPU	Central Processing Unit
DDR3	Double Data Rate 3
DHCP	Dynamic Host Configuration Protocol
ECC	Error Correcting Codes
FC	Fibre Channel
FCoE	Fibre Channel over Ethernet
GB	Gigabyte
GBit	Gigabit
HA	High Availability
HBA	Host-Bus-Adapter
HDD	Hard Disk Drive
HTTP	Hypertext Transfer Protocol
IETF	Internet Engineering Task Force
IP	Internet Protocol
iRMC	Integrated Remote Management Controller
iSCSI	internet Small Computer System Interface
LAN	Local Area Network
MAC	Media Access Control
NTP	Network Time Protocol
OSI	Open Systems Interconnection
PCI	Peripheral Component Interconnect
PCIe	PCI Express

RAID	Redundant Array of Independent Disks
RAM	Random Access Memory
RFC	Request For Comments
SAN	Storage Area Network
SAS	Serial Attached Storage
SCSI	Small Computer System Interface
SED	Self Encrypting Disk
SFP	Small Form-factor Pluggable
SNIA	Storage Networking Industry Association
SNMP	Simple Network Management Protocol
SSH	Secure Shell
TB	Terabyte
TCP	Transmission Control Protocol
USB	Universal Serial Bus
USV	Unterbrechungsfreie Stromversorgung
VM	Virtuelle Maschine
VPN	Virtual Private Network
WMI	Windows Management Instrumentation
WWNN	World Wide Node Name
WWPN	World Wide Port Name

1 Einleitung

Die Techniken zur langfristigen Speicherung von Daten haben sich in den letzten Jahrzehnten sehr stark verändert. Viele früher gängige Speichermedien sind fast nicht mehr im Einsatz, andere hingegen gerade dabei sich zu etablieren. Grund dafür ist der immer weiter steigende Zuwachs an Daten und die daraus resultierende Notwendigkeit immer größerer Speicherkapazitäten. Unternehmen, bei denen viele Daten anfallen, speichern heutzutage schon längst nicht mehr auf simplen Netzwerkfestplatten - hier kommen in der Regel ganze Speichernetze zum Einsatz.

Die vorliegende Ausarbeitung beschäftigt sich mit der Erneuerung eines vorhandenen Storage Area Network¹, das den Anforderungen an Speicherkapazität, Datensicherheit und Leistung nicht mehr gewachsen ist. Wobei der Kunde bisher über kein herkömmliches Speichernetz verfügt, sondern seine Daten auf zwei netzwerkfähigen Festplattensystemen ablegt. Das Projekt wird auf den folgenden Seiten von der anfänglichen Analyse des Ist-Zustandes bis zur Übergabe und Auswertung dokumentiert.

Dabei werden vor Projektbeginn die verschiedenen möglichen Technologien betrachtet. Nach einer Analyse der Anforderungen des Kunden wird die für das Projekt beste Technologie ausgewählt. Nach einer groben Dimensionierung werden geeignete Hardware-Komponenten ausgesucht - dabei werden nicht verschiedene Hersteller verglichen, sondern Modelle von unseren Stammlieferanten ausgewählt. Die Einrichtung wird nicht an jedem Punkt bis ins kleinste Detail beschrieben - für unsere Firma häufige Arbeitsschritte wie die Installation eines Virtualisierungsservers werden eher grob umrissen, während die Basiskonfiguration eines Speichersystems detaillierter dargestellt wird.

Ein besonderer Schwerpunkt liegt auf der Sicherheit der gespeicherten Daten und der Verfügbarkeit des Systems. Es werden Maßnahmen zur sicheren Datenhaltung und zur Steigerung der Ausfallsicherheit angeführt und zum Teil - soweit möglich - auch im Projekt umgesetzt.

Zum Abschluss der Ausarbeitung werden aufgetretene Probleme betrachtet und das Erreichen der festgelegten Projektziele ausgewertet.

¹Storage Area Network: Netzwerk zur Speicherung großer Datenmengen

2 Analyse Ist-Zustand

In diesem Kapitel wird der Zustand des Storage Area Network des Kunden vor Projektbeginn betrachtet. Der Auftraggeber des Projektes ist ein Unternehmen, welches Bauteile zur Innenausstattung von Fahrzeugen für die Automobilindustrie entwirft und produziert. Es gibt einen Bürostandort und einen 50 km entfernten Produktionsstandort. Die hier dargestellte Skizze des Unternehmens-Netzwerkes ist stark vereinfacht, sie zeigt nicht alle Client-Computer und beschränkt sich auf den Bürostandort, der die zu fertigenden Bauteile designet. Schwerpunkt der Analyse ist das SAN² dieses Standortes, auf dem alle firmen- und auftragsrelevanten Daten abgelegt sind, der Produktionsstandort greift über VPN³ ebenfalls auf diesen Datenbestand zu.

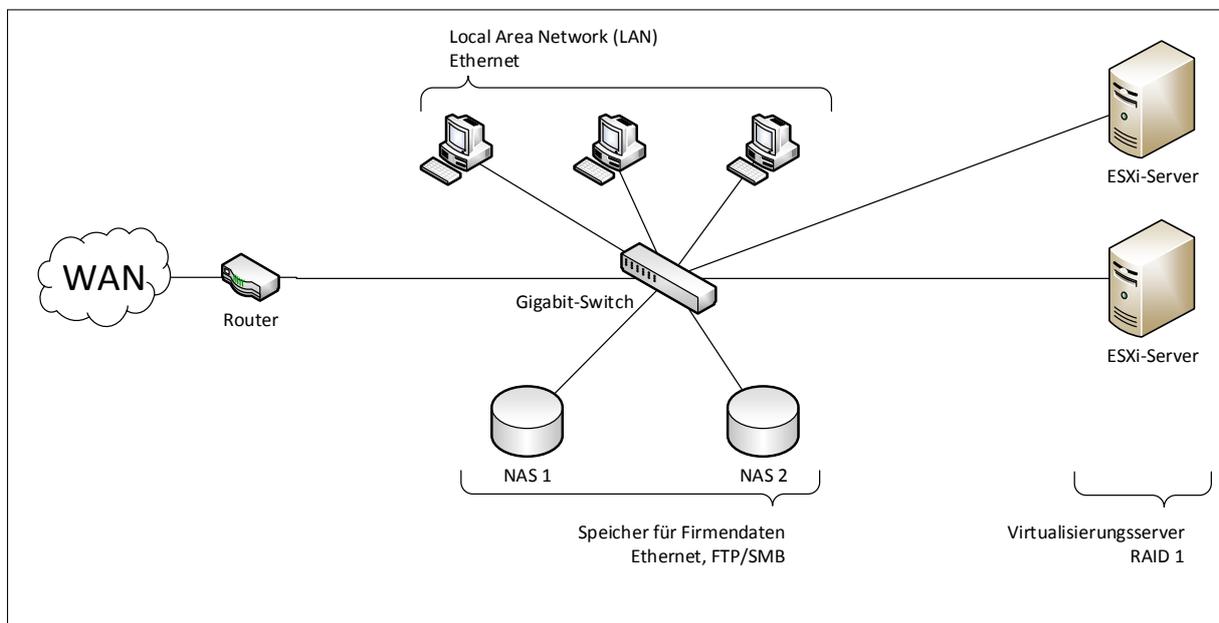


Abbildung 1: Ist-Zustand des Kundennetzwerkes

Den wichtigsten Punkt des Netzwerkes stellen zwei Fujitsu PRIMERGY RX300 S5⁴-Server dar, auf denen die Virtualisierungsplattform VMware ESX⁵ läuft, um jeweils mehrere virtuelle Betriebssysteme über eine physikalische Maschine bereitzustellen. Als Datenspeicher

²SAN: Storage Area Network

³VPN: Virtual Private Network; eine Schnittstelle zur Verbindung von zwei Netzwerken, meist für Fernzugriff auf Arbeitsplatznetzwerke verwendet, i.d.R. Ende-zu-Ende-verschlüsselt

⁴PRIMERGY RX300 S5: Rackserver der PRIMERGY-Reihe von Fujitsu aus dem Jahr 2005

⁵VMware ESX: Plattform zur Virtualisierung eines oder mehrerer Betriebssysteme auf einem physikalischen Computer [Wik14d]

für die VMs⁶ dienen serverinterne Festplatten mit einer Kapazität von 250 GB im RAID⁷ 1⁸. Für firmenbezogene Daten sind zwei NAS⁹-Systeme angeschlossen und als Netzlaufwerk in den VMs eingebunden. Diese beiden Disk-Subsysteme¹⁰ der Marke Synology¹¹ bieten jeweils eine Speicherkapazität von 250 GB¹² im RAID 1, das zweite wurde nachträglich angeschafft, da der Speicher des ersten Systems nicht mehr ausreichend war. Die gesamte Kapazität des der Server und Netzwerkfestplatten beträgt 1.000 GB. Die Vernetzung besteht aus 1 GBit/s¹³-Ethernet¹⁴-Komponenten.

Der zentrale Gigabit-Switch¹⁵ verbindet alle 15 Client-PCs miteinander, mit den Virtualisierungsservern und den NAS-Systemen. Ein Router¹⁶, der als Gateway für Internet und VPN dient, ist ebenfalls angeschlossen.

Die Datensicherung des Servers und einiger ausgewählter Ordner auf den Netzlaufwerken erfolgt jede Nacht auf Band. Die Bänder werden jeweils vor Dienstschluss eingelegt und am nächsten Morgen entnommen. Für jeden Wochentag ist ein eigenes Sicherungsband vorhanden.

Der in der Skizze nicht dargestellte Produktionsstandort ist über VPN mit dem Bürostandort verbunden. Jeweils ab 19 Uhr werden über VPN die Produktionspläne und -daten für den nächsten Werktag von den VMs an den Produktionsstandort übertragen.

⁶VM: Virtuelle Maschine: Betriebssystem, welches auf einer Virtualisierungsplattform wie VMware ESX ausgeführt wird

⁷RAID: Redundant Array of Independent Disks: Verbindung mehrerer Festplatten zu einem logischen Speicher mit höherer Ausfallsicherheit

⁸RAID 1: Level-1-RAID: logischer Datenträger aus zwei Festplatten, die Daten der ersten werden auf die zweite Platte gespiegelt

⁹NAS: Network Accessible Storage: Datenspeicher, der über Netzwerkverbindungen administriert und verwendet wird

¹⁰Disk-Subsystem: Speichereinheit innerhalb eines Storage Area Network

¹¹Synology: ein Hersteller von Speichersystemen

¹²GB: Gigabyte, Speichereinheit (1 GB = 1.024 Megabyte)

¹³GBit/s: Gigabit/Sekunde, Übertragungsgeschwindigkeit: bis zu 1 Milliarde Bit pro Sekunde

¹⁴Ethernet: Technologie zur Datenübertragung über kabelgebundene Netze [Wik14a]

¹⁵Switch: aktives Netzwerkelement zur Verbindung mehrerer Kabel

¹⁶Router: aktive Netzwerkkomponente, die den Zugriff ins Internet ermöglicht

Der dargestellte Ist-Zustand bereitet dem Kunden folgende Probleme:

- Die Datenspeicherung auf den NAS-Systemen ist für die gewachsenen Anforderungen mit 1 GBit/s zu langsam, große Kopiervorgänge verlangsamen die Arbeit enorm.
- Der Speicher der beiden Disk-Subsysteme ist fast erschöpft, eine Erweiterung des Netzwerkes um ein weiteres NAS würde die Performance noch stärker belasten und den Administrationsaufwand weiter erhöhen.
- Alle Daten der Firma lagern in einem Raum und sind nur teilweise gesichert - im Falle einer Katastrophe gibt es keinen kompletten Datenbestand mehr.
- Der Datenbestand ist auf zwei Disk-Subsysteme verteilt, die separat gewartet werden müssen, was hohen Administrationsaufwand bedeutet.
- Die Disk-Subsysteme werden bisher nur reaktiv überwacht, ein Eingriff erfolgt erst bei Ausfall einer Komponente.
- Der zentrale Switch des Firmennetzwerkes und die Disk-Subsysteme stellen Single Points of Failure¹⁷ dar.

Zur Lösung der genannten Probleme wurde das Projekt „Konzeptionierung und Realisierung der Erneuerung des Storage Area Network (SAN) eines mittelständischen Unternehmens zur Steigerung der Ausfallsicherheit der IT-Infrastruktur“ gestartet.

¹⁷Single Point of Failure: Komponente, deren Ausfall das Gesamtsystem in einen arbeitsunfähigen Zustand versetzt

3 Theorie möglicher Techniken

3.1 Einführung

Ein Storage Area Network - kurz SAN - ist ein Hochgeschwindigkeitsnetzwerk zwischen Servern und Speichersubsystemen. Bei herkömmlichen Client-Server-Architekturen erfüllt jeder Server im LAN¹⁸ eine bestimmte Aufgabe und hat zur Datenhaltung ein dediziertes Speichersubsystem. Im Gegensatz dazu kann innerhalb des SAN jede Komponente mit jeder beliebigen anderen kommunizieren - Server-zu-Server, Server-zu-Speichersubsystem und auch Speichersubsysteme untereinander. [Rob01, S. 29] Dieser Unterschied ist zur Verdeutlichung in **Abbildung 2** dargestellt.

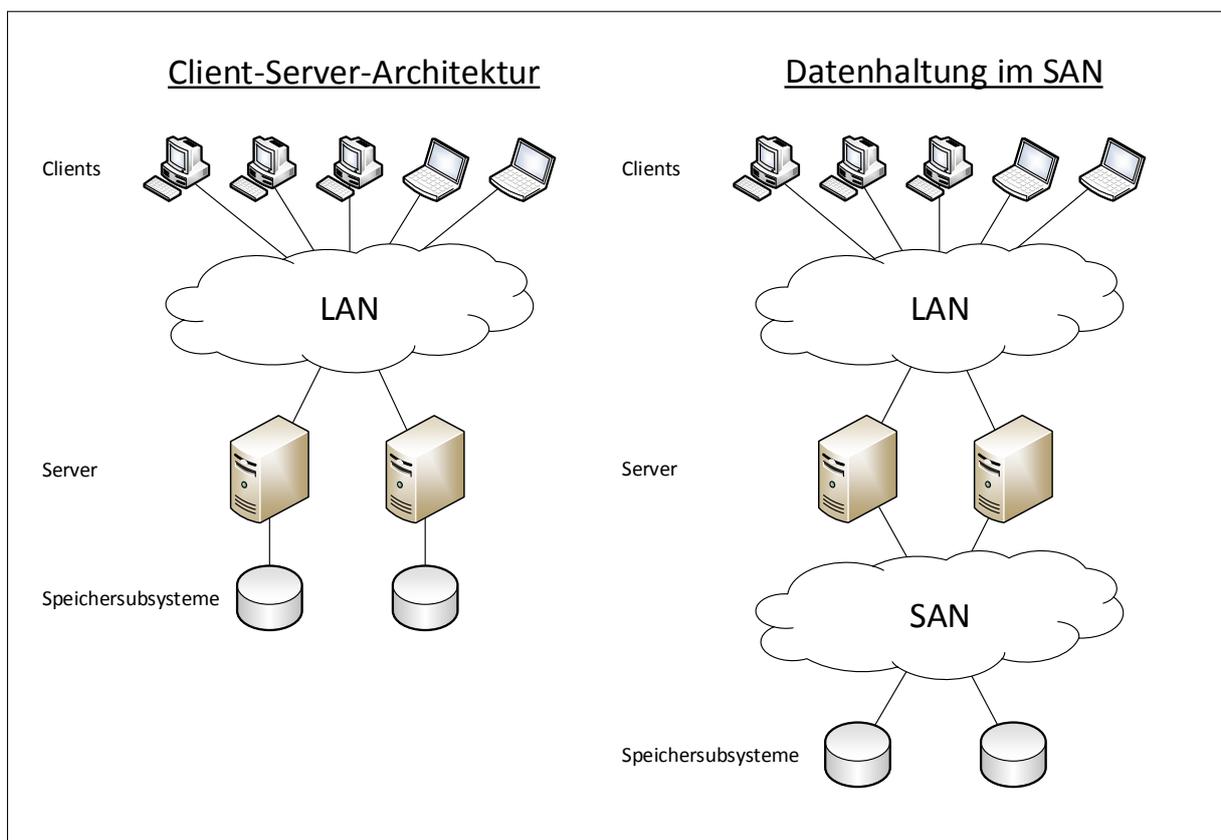


Abbildung 2: Vergleich Client-Server-Architektur / Datenhaltung im SAN

¹⁸LAN: Local Area Network, ein Netzwerk zwischen Computern und anderen aktiven Komponenten

Diese Speichernetzwerke bieten gegenüber der traditionellen Datenhaltung einige wesentliche Vorteile:

- **Leistungsfähigkeit:**

Das SAN gilt als das schnellste verfügbare Kommunikationsmedium zur Datenübertragung - war noch vor einigen Jahren 1 GBit/s die maximale Übertragungsgeschwindigkeit, stellen zur Zeit schon 8, 10 oder 16 GBit/s kein Problem mehr dar.

- **Skalierbarkeit:**

Beim Aufbau eines SAN kann Hardware verschiedenster Hersteller kombiniert werden. Durch internationale Standards ist die grundlegende Kommunikation zwischen allen Komponenten genormt und somit gewährleistet. Erweiterte herstellereigenspezifische Protokolle werden teilweise schon durch Kooperationen von mehreren Herstellern plattformübergreifend unterstützt. Die Infrastruktur eines Speichernetzes ist fast grenzenlos erweiterbar - ein SAN kann etwa aus bis zu 16,78 Millionen Knoten¹⁹ bestehen.

- **Geographische Ausdehnung:**

Zwischen zwei Knoten sind Entfernungen bis zu 10 km möglich - bei Einsatz spezialisierter Hardware sogar mehrere hundert Kilometer. Durch die großen überbrückbaren Distanzen können Storage Area Networks auch gebäudeübergreifend implementiert werden. [Rob01, S. 30]

- **Verfügbarkeit:**

Mit Speichernetzen können Hochverfügbarkeitsstrukturen optimal abgebildet werden. Durch redundante Auslegung von Verbindungswegen und parallel arbeitende Komponenten kann die Ausfallsicherheit und Performance der Infrastruktur weiter verbessert werden.

- **Management:**

Da die Speichersubsysteme von jedem Server im SAN verwendet werden können, gehören Dateninseln der Vergangenheit an. Die komplette Infrastruktur kann von einer einzigen Stelle aus verwaltet werden und Backups müssen nicht zwangsläufig über Nacht durchgeführt werden, um das Netzwerk zu schonen - sie können direkt durchgeführt werden, wenn der Bedarf besteht.

[Rob01, S. 30ff]

¹⁹16,78 Millionen Knoten: ergibt sich aus 2^{24} adressierbaren Geräten (24 Bit Adressraum)

Für den Aufbau eines Storage Area Network stehen verschiedene Technologien zur Verfügung, von denen jede ihre eigenen spezifischen Vor- und Nachteile hat. In den folgenden Abschnitten sollen „iSCSI²⁰“, „Fibre Channel²¹“, und „FCoE²²“ kurz erklärt und miteinander verglichen werden. [uSKG14]

3.2 SAN-Konzepte im Überblick

3.2.1 iSCSI

Die SAN-Technologie iSCSI - kurz für „internet SCSI²³“ basiert auf der Übertragung des SCSI-Protokolls über Ethernet. Das Protokoll wurde durch die SNIA²⁴ entworfen und von der IETF²⁵ in der RFC²⁶ 3720 [Sat04] standardisiert. Die SCSI-Daten werden auf den Ports 860 und 3260 über IP²⁷-Netze transportiert und dazu in TCP²⁸/IP-Pakete verpackt - somit kann bereits vorhandene Netzwerkhardware oder unter Berücksichtigung eines Performance-Verlustes sogar ein vorhandenes Netzwerk für den Aufbau eines iSCSI-SAN verwendet werden. Die Übertragung der Daten erfolgt blockbasiert. [Wik13b]

Im SCSI-Protokoll ist die Kommunikation über einen SCSI-Bus vorgesehen, Initiatoren und Targets können als SAN-Komponenten angeschlossen werden. Der Bus wird an beiden Enden durch einen Terminator begrenzt, dabei handelt es sich um einen aktiven Abschlusswiderstand. **(Abbildung 3)**

Da bei iSCSI der Bus über Ethernet abgebildet wird, entfallen die Terminatoren. Initiatoren sind Server und andere Geräte, die auf Daten zugreifen, Targets hingegen stellen die Daten

²⁰iSCSI: internet Small Computer System Interface, ein SAN-Konzept - wird in Abschnitt 3.2.1 näher erläutert

²¹Fibre Channel: eine SAN-Technologie, näher erklärt im Abschnitt 3.2.2 ab Seite 9

²²FCoE: Fibre Channel over Ethernet: eine weitere SAN-Technologie, erläutert im Abschnitt 3.2.3 ab Seite 12

²³SCSI: Small Computer System Interface - ein Standard zur blockbasierten Datenübertragung

²⁴SNIA: Storage Networking Industry Association - Internationale Handelsvereinigung zur Entwicklung von Standards für Speichernetzwerke [Ass14]

²⁵IETF: Internet Engineering Task Force - Internationale Organisation, welche die Technologien des Internets weiterentwickelt und standardisiert [For14]

²⁶RFC: Request For Comments - Bezeichnung für technische Publikationen über Internet-Technologien

²⁷IP: Internet Protocol - ein etablierter Standard zur Paketvermittlung in Netzwerken

²⁸TCP: Transmission Control Protocol - Protokoll zur paketbasierten Datenübertragung in IP-Netzwerken; arbeitet auf OSI-Schicht 4, siehe Anhang 1

bereit. Jedes (i)SCSI-Gerät wird über seine vorher festgelegte SCSI-ID (eine Zahl) identifiziert. Der zur Vernetzung benötigte Host-Bus-Adapter²⁹ (kurz HBA) kann in Hardware (z.B. als PCI³⁰-Steckkarte) oder in Software (iSCSI-Treiber für normale Netzwerkkarte) realisiert sein.

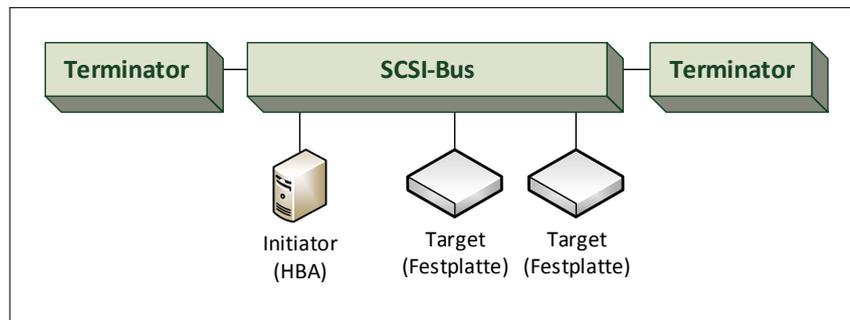


Abbildung 3: SCSI-Bus

Der Datenstrom von iSCSI wird vor dem Versand zunächst in TCP-Datenpakete verpackt - diese nehmen bis zu 1460 Bytes Nutzdaten auf und haben einen Header von 20 Bytes. Da TCP nur ein Protokoll zur verbindungsorientierten Datenübertragung ist und keine Adressierung beherrscht, wird der TCP-Datenstrom wiederum in IP-Pakete gekapselt. Das IP-Protokoll arbeitet auf OSI-Schicht 3 (siehe Anhang 1)) und erhöht den Overhead der Übertragung nochmals mit einem Header von 20 Bytes. Da das für iSCSI genutzte Übertragungsmedium Ethernet ist, werden die IP-Pakete noch in Ethernetrahmen verpackt. Durch den Ethernet-Standard kommt am Anfang der Daten ein weiterer Header von 14 Bytes und am Ende eine Prüfsumme von 2 Bytes dazu, somit wird der Overhead nochmals erhöht. Die mehrfache Kapselung des Datenstromes ist in **Abbildung 4** auf Seite 9 dargestellt. [Nin14]

Vorteile von iSCSI:

- Bereits vorhandene Netzwerkhardware kann verwendet werden.
- iSCSI-Hardware ist im Vergleich zu anderen Technologien in der Regel kostengünstiger.
- Nutzung bewährter Protokolle (Ethernet, TCP/IP)
- Bei kleineren Datenübertragungen etwa so schnell wie Fibre Channel.
- Durch TCP/IP-Basis kann auch das Internet zur Übertragung zwischen Standorten genutzt werden.

²⁹Host-Bus-Adapter: Komponente zum Zugriff auf einen Datenbus

³⁰PCI: Peripheral Component Interconnect - ein computerinterner Anschluss für Erweiterungskarten [Wik14c]

Nachteile von iSCSI:

- Eine Fusion des SAN mit dem vorhandenen LAN ist möglich, wirkt sich jedoch negativ auf die Performance aus. [uSKG13b]
- Im Gigabit-Ethernet niedrigere Datenraten als bei FCoE.
- Höhere Latenzzeit³¹ als Fibre Channel - stark bemerkbar bei Datenbankanwendungen. [uSKG13b] (Da der iSCSI-Datenstrom mehrfach gekapselt wird - siehe **Abbildung 4**, Seite 9 - beansprucht die Verarbeitung der Pakete mehr Rechenleistung, was die Latenzzeit erhöht.)
- Bei Internet-Übertragung ist Verschlüsselung möglich, diese benötigt jedoch zusätzliche Rechenleistung, was sich negativ auf die Datenrate auswirkt. [uSKG13b]



Abbildung 4: Paketkapselung von iSCSI

3.2.2 Fibre Channel

Der Fibre Channel-Protokoll-Standard³² wurde entwickelt, um die von SCSI bekannten Vorteile auszubauen und gleichzeitig die Schwachstellen zu umgehen. Die wichtigsten Entwicklungsziele waren eine höhere Übertragungsgeschwindigkeit, die Überbrückung weitaus größerer Distanzen, die Verringerung von Übertragungsfehlern (Skewing³³ bei SCSI), die Verkürzung von Latenzzeiten sowie die Umsetzung des Protokolls

³¹Latenzzeit: Dauer, die ein anderes Netzwerkgerät zum Senden der Antwort auf eine Anfrage benötigt

³²Fibre Channel-Protokoll-Standard: korrekt bezeichnet „Fibre Channel Physical and Signaling Interface“ [All97]

³³Skewing: bezeichnet Fehler, die entstehen, wenn Daten über mehrere Adern parallel übertragen werden und aufgrund physikalischer Eigenschaften der Leiter zeitversetzt beim Empfänger ankommen

in Hardware als Host-Bus-Adapter zur Verringerung der CPU³⁴-Belastung auf den Knoten.

Ähnlich wie das OSI-Modell³⁵ (Anhang 1) ist das FC³⁶-Protokoll aus mehreren Schichten aufgebaut, diese werden im Folgenden näher erläutert.

Physikalische Schicht (FC 0)

Die unterste Schicht stellt das Übertragungsmedium, Typ, Entfernung und Geschwindigkeit der Übertragung dar. In einem FC-SAN hat jeder Knoten - ähnlich der MAC³⁷-Adresse einer Netzwerkkarte - eine eindeutige Bezeichnung, die WWNN³⁸. Außerdem kann ein Knoten auch mehrere Ports besitzen, diese werden durch die WWPNN³⁹ identifiziert. Die Identifikatoren sind jeweils 64-Bit-Schlüssel. [uSKG13a]

Verschlüsselung (FC 1)

Die Datenübertragung im Fibre Channel ist standardmäßig 8b/10b⁴⁰-kodierte (siehe **Anhang 2**). Dieses Verfahren hat einen sehr geringen Overhead⁴¹, somit kann bei gleicher Geschwindigkeit eine größere Datenmenge übertragen werden als bei Ethernet. Die neueren FC-Switches mit 10 bzw. 20 GBit/s Übertragungsraten verwenden hingegen die 64b/66b⁴²-Kodierung, was den Overhead nochmals verringert. Dadurch wird der Anteil an Nutzdaten bei der Übertragung auf etwa 97% erhöht, somit können über Fibre Channel bei gleicher Geschwindigkeit mehr Daten übertragen werden als beispielsweise über iSCSI.

³⁴CPU: Central Processing Unit - englische Bezeichnung für den Prozessor

³⁵OSI-Modell: Open Systems Interconnection Model - Referenzmodell, welches Netzwerkprotokolle in mehrere Schichten einteilt, siehe Anhang 1

³⁶FC: Abkürzung für Fibre Channel

³⁷MAC: Media Access Control - weltweit eindeutige Identifikation der Netzwerkkarte eines Computers

³⁸WWNN: World Wide Node Name

³⁹World Wide Port Name

⁴⁰8b/10b: je 8 Bit werden in 10 Bit-Werte kodiert und versendet

⁴¹Overhead: Informationen, die zusätzlich zu den Nutzdaten im Netz übertragen werden, z.B. Paket-Kopfdaten

⁴²64b/66b: je 64 Bit werden in 66 Bit-Werte kodiert und versendet

Management und Informationsstruktur (FC 2)

Diese Schicht ist die umfangreichste, sie steuert die Struktur und die Definition der übertragenen Daten und ist in vier Klassen aufgeteilt:

- **Login Session:**

Stellt eine Verbindung zwischen zwei FC-Ports dar. Beim Verbindungsaufbau tauschen die Ports Informationen über ihren Status und unterstützte Operationen aus, anschließend sind die Ausführung von Befehlen höherer Schichten und I/O⁴³-Operationen zwischen den Ports möglich. [Rob01, S. 71]

- **Exchange:**

Ist ein Mechanismus zur Steuerung des Datenaustausches zwischen zwei Ports. Der Exchange wird mit der Datenübertragung gestartet und beendet. Es können mehrere Exchanges parallel laufen, um den Port effizienter zu nutzen. [Rob01, S. 75f]

- **Sequenz:**

Eine Sequenz ist die Übertragung eines großen Datenblocks, sie fasst bis zu 65.536 Frames zusammen. Eine Sequenz von Frames zu übertragen ist effizienter als die Frames einzeln zu senden - bei einer Sequenz müssen die Kopfdaten auf dem Weg zum Empfänger nur einmal ausgewertet werden, statt für jeden einzelnen Frame. [Rob01, S. 76ff]

- **Frames:**

Ein Frame ist das kleinst mögliche Paket zur Datenübertragung im Fibre Channel, kann bis zu 2.112 Bytes Nutzlast enthalten und besitzt einen Overhead von 36 Byte (Header, Prüfsumme, etc.). Man unterscheidet „Link Control Frames“ zur Übertragung von Steuerungsanweisungen und Daten-Frames für die Nutzdaten. [Rob01, S. 78f]

Common Services (FC 3)

Die dritte Schicht stellt herstellerspezifische Zusatzfunktionen zum FC-Standard zur Verfügung. Hier gibt es noch Inkompatibilitäten zwischen Herstellern, da bisher nicht alle Funktionen durch die ANSI⁴⁴ standardisiert wurden. Beispiele dafür sind „Multicast-Sendungen“ (Datenreplikation durch Zustellung der Übertragung an mehrere Empfänger) und „Striping“

⁴³I/O: Input/Output - Eingabe/Ausgabe

⁴⁴ANSI: American National Standards Institute - amerikanische Institution zur Standardisierung von Technologien

(abwechselnder Versand der Daten über verschiedene Ports und Wege zur Erhöhung des Durchsatzes). [Rob01, S. 107ff]

Protokoll Mapping (FC 4)

Die oberste und letzte Protokollschicht von Fibre Channel stellt eine logische Verbindung zwischen verschiedenen Architekturen her, sie ermöglicht beispielsweise die Anbindung von SCSI-Geräten. Dabei werden die SCSI-Befehle und -Datenblöcke in Fibre Channel Frames verpackt und können beim Empfänger zusammengesetzt und wiederum an ein SCSI-Gerät übergeben werden - auf diese Weise wird Fibre Channel als Transportmedium für SCSI und auch andere „Upper Level“-Protokolle verwendet. [Rob01, S. 111ff]

3.2.3 Fibre Channel over Ethernet

Eine weitere SAN-Technologie stellt „Fibre Channel over Ethernet“ - kurz FCoE - dar. Das Ziel der Entwicklung war es, die Vorteile von SCSI und Fibre Channel bestmöglich zu vereinen. Im Wesentlichen basiert das Protokoll auf der Übertragung von FC-Frames über Ethernet - dadurch kann ein bestehendes Ethernet-Netzwerk mit genutzt oder ein neues relativ preisgünstig aufgebaut werden, eine Investition in kostenintensive FC-Vernetzung ist nicht zwingend notwendig. Dennoch sind höhere Übertragungsraten als bei SCSI oder iSCSI möglich. Damit ist Fibre Channel over Ethernet bei nahezu gleicher Performance und Latenzzeit eine kostengünstige Alternative zu Fibre Channel. [Wik13a]

Zur Verwendung des Protokolls benötigen die angebundenen Server eine FCoE-Netzwerkkarte - diese kann zum einen den alltäglichen Datenverkehr als Ethernet-Netzwerkkarte verarbeiten und stellt dem Betriebssystem zum anderen einen Storage-Host-Bus-Adapter zur Verfügung, mit dem Festplatten direkt auf Blockebene angesprochen werden können. Zudem sollten spezielle FCoE-Switches eingesetzt werden, die neben den Ethernet-Diensten auch den FCoE-Transport steuern, beispielsweise den SAN-Datenverkehr priorisieren und die Latenzzeiten gering halten. Moderne FCoE-Switches bieten neben den Ethernet-Ports auch Fibre Channel-Ports an, mit denen Storage Hardware angesprochen werden kann - so ist es möglich, ein SAN ohne den Einsatz teurer FC-Switches aufzubauen. [uSKG08]

3.3 SAN-Topologien

3.3.1 Allgemeines zu Topologien

Eine Topologie beschreibt die Struktur der Vernetzung zwischen Netzwerkgeräten, dabei können sowohl Kabel- als auch Drahtlos-Verbindungen zum Einsatz kommen. Im Falle eines Storage Area Network sind reine Kabel-Vernetzungen typisch. In den folgenden Abschnitten werden einige Topologien beschrieben, die häufig für den Aufbau eines SAN verwendet werden.

3.3.2 Point to Point

Unter einer Point to Point-Topologie (Punkt zu Punkt-Verbindung) versteht man die Verbindung zwischen zwei Knoten eines Storage Area Network. Da diese immer zwischen zwei vorher definierten Ports besteht, ist kein Protokoll zur Verwaltung der Verbindung notwendig. Der „Transmitter“ des einen Ports sendet die Datenframes, der „Receiver“ des anderen Ports empfängt diese. Eine Adressierung wäre theoretisch nicht notwendig, da es nur zwei kommunizierende Ports gibt, wird dennoch für den Fall angewendet, dass ein Knoten weitere Verbindungen zu anderen Topologien unterhält. Eine Point to Point-Topologie ist auf **Abbildung 5** dargestellt. [Nin14]

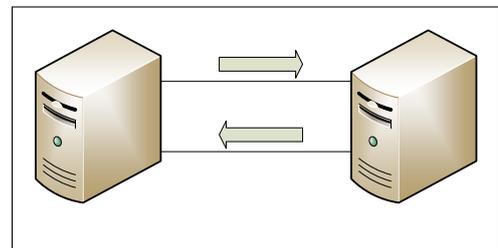


Abbildung 5: Point to Point-Topologie

3.3.3 Loop

Die in der Fibre Channel-Technologie verwendete Ring-Topologie - korrekt bezeichnet als Fibre Channel Arbitrated Loop (FC-AL) - ist mit einem Token-Ring-Netzwerk⁴⁵ vergleichbar. Wie auf **Abbildung 6** ersichtlich, sind alle Komponenten ringförmig vernetzt und die Kommunikation verläuft nur in eine Richtung. Vor jeder Datenübertragung vermitteln (engl. „arbitrate“) die angeschlossenen Geräte, wer als Nächstes senden darf. Die ermittelte Komponente sendet ihre Daten, diese werden jeweils zum nächsten Knoten weitergeleitet, bis sie ihren Empfänger erreicht haben. Die effektive Datenrate fällt dabei geringer aus als die vom Kabel maximal zulässige. Grund dafür ist die Tatsache, dass sich alle Teilnehmer eine Übertragungsleitung teilen, deshalb sollte die Anzahl der aktiven Komponenten in einem Loop möglichst gering gehalten werden. Diese Topologie hat den entscheidenden Nachteil, dass bei Ausfall eines Knotens oder einer Übertragungsleitung keine Kommunikation mehr möglich ist. [Nin14]

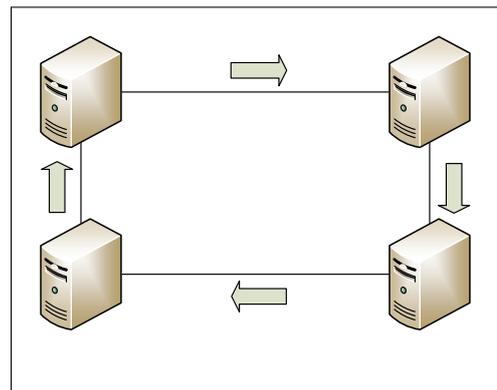


Abbildung 6: Loop-Topologie

3.3.4 Fabric

Eine häufig verwendete SAN-Topologie stellt das „Fabric“ oder auch „Switched Fabric“ dar. Sie zeichnet sich dadurch aus, dass einer oder mehrere Switches zum Einsatz kommen, die jedem Knoten zu jedem beliebigen anderen Knoten eine Punkt-zu-Punkt-Verbindung ermöglichen - so steht jedem Teilnehmer die volle Bandbreite zur Verfügung und durch die Kaskadierung von Switches lassen sich bis zu 16,78 Millionen Geräte anschließen. Ein einfaches Beispiel für ein Fabric mit nur einem Switch ist auf **Abbildung 7** dargestellt. Grundsätzlich gibt es bei

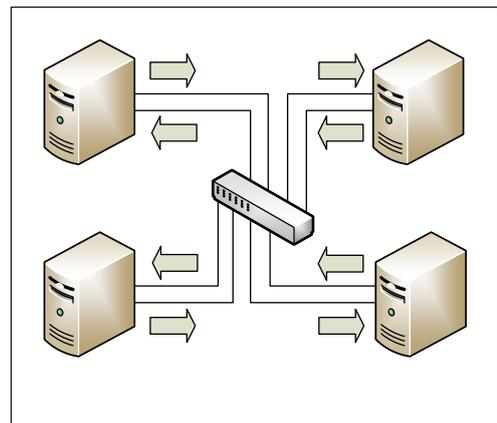


Abbildung 7: Fabric-Topologie

⁴⁵Token-Ring-Netzwerk: Kommunikation verläuft ringförmig immer nur in eine Richtung; dabei darf nur ein Gerät Daten senden, die anderen leiten diese bis zum Empfänger weiter

den Punkt zu Punkt-Verbindungen innerhalb eines Fabric zwei Arten: den verbindungsorientierten Modus, bei dem zwischen zwei Punkten ein exakter Weg festgelegt ist, über den die Daten übertragen werden, und den verbindungslosen Modus, bei dem nur die Adresse des Zielgerätes festgelegt ist und die Switches des Fabric das Routing⁴⁶ übernehmen. [Rob01, S. 35ff]

3.3.5 Full Fabric

Die Fabric-Topologie lässt sich zur so genannten Full Fabric-Topologie erweitern. Dabei wird der zentrale Switch durch mehrere Switches ersetzt, von denen jeder mit jedem vernetzt ist, wie auf **Abbildung 8** dargestellt. Diese Art der Vernetzung bietet den Vorteil, dass bei Ausfall eines Switches nicht das gesamte Netz unterbrochen wird, sondern nur der mit dem betreffenden Switch verbundene Server. Die Ausfallsicherheit kann weiter erhöht werden, in dem jeder Server mit 2 Switches verbunden ist - so können im gezeigten Beispiel bis zu zwei Switches ausfallen und dennoch können alle Knoten miteinander kommunizieren. Der Nachteil dieser Topologie sind die höheren Kosten für die Switches und die Verkabelung. [Nin14]

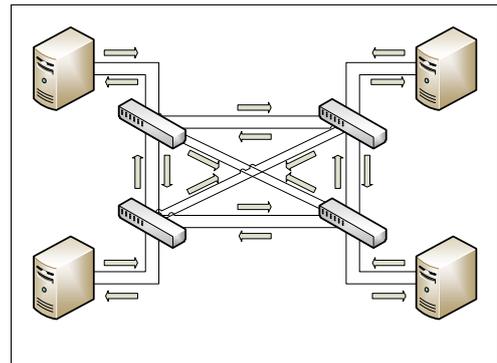


Abbildung 8: Full Fabric-Topologie

⁴⁶Routing: Transport von Datenpaketen zum Empfänger

3.4 Hochverfügbarkeit

Der Begriff der Hochverfügbarkeit (engl. high availability oder kurz HA) bezeichnet die Fähigkeit eines Systems, nach dem Ausfall einer Komponente ohne menschlichen Eingriff weiterhin seinen Betrieb zu gewährleisten. Der Anwender des Systems sollte dabei keine oder nur eine kurze Unterbrechung wahrnehmen. Es gibt mehrere Verfügbarkeitsklassen, diese sind in der folgenden Tabelle aufgeführt. [Wik14b]

Klasse	Verfügbarkeit	Ausfallzeit
1	98%	7,30 Tage/Jahr
2	99%	3,65 Tage/Jahr
3	99,9%	8,76 Stunden/Jahr
4	99,99%	52,6 Minuten/Jahr
5	99,999%	5,26 Minuten/Jahr
6	99,9999%	31,5 Sekunden/Jahr

Tabelle 1: Verfügbarkeitsklassen

Quelle: [TEM08, S. 382]

Die Verfügbarkeit eines Systems lässt sich mit folgender Formel berechnen:

$$V = \left(1 - \frac{t_{down}}{t_{down} + t_{up}}\right) * 100$$

Formel 1: Verfügbarkeit

Dabei steht t_{down} für die Downtime⁴⁷ und t_{up} für die Uptime⁴⁸ des Systems, das Ergebnis V ist die Verfügbarkeit in Prozent. Als Faustregel gilt: je höher die geforderte Verfügbarkeit des Systems, desto kostenintensiver ist dieses in der Anschaffung. [TEM08, S. 381f]

Voraussetzungen für eine hohe Verfügbarkeit sind:

- fehlertolerantes Verhalten der zusammenwirkenden Komponenten
- redundante Auslegung kritischer Komponenten
- automatische Umschaltung bei Ausfall einer redundanten Komponente

⁴⁷Downtime: Zeit, in der ein System nicht arbeitsfähig ist

⁴⁸Uptime: Zeit, in der ein System arbeitsfähig ist

Beispiele für Maßnahmen zur Erhöhung der Verfügbarkeit:

- Absichern der Stromversorgung durch eine USV⁴⁹; weitere Verbesserung der Verfügbarkeit durch redundante Auslegung der USV
- Verwendung von ECC-RAM⁵⁰
- Mehrfache Netzteile pro Server
- Einsatz von RAID-Systemen
- Replikation⁵¹ von Servern und Datenspeichern
- Cold Standby⁵²

Die Einhaltung der geforderten Verfügbarkeit bedingt:

- schnelle Erreichbarkeit von Fachpersonal
- Verfügbarkeit von Ersatzteilen
- proaktive Wartungsmaßnahmen, z.B. Monitoring
- effiziente Kommunikationswege beim Auftreten von Fehlern

⁴⁹USV: Unterbrechungsfreie Stromversorgung - ein Akku, der über kurze Stromausfälle hinweg die Server weiter versorgt

⁵⁰ECC-RAM: Arbeitsspeicher mit ECC; ECC: Error Correcting Codes - Fehlerkorrekturmechanismus

⁵¹Replikation: automatische Kopie von Daten auf einen zusätzlichen Speicher

⁵²Cold Standby: von jeder Komponente ist eine baugleiche zweite vorhanden, die bei Ausfall der ersten durch den Systembetreuer in Betrieb genommen wird; vgl. Hot Standby: Komponente wird bei Ausfall automatisch aktiviert

4 Projekt-Realisierung

4.1 Planung

4.1.1 Analyse der Anforderungen

Im Kapitel „Analyse Ist-Zustand“ wurden auf Seite 4 bereits die Probleme aufgezeigt, die der Kunde mit seinem Storage Area Network hat, die zur Entstehung des Projektes geführt haben. In diesem Abschnitt sollen aus den Problemen konkrete Anforderungen formuliert werden.

SAN-Performance

Das bisher eingesetzte iSCSI-SAN ist mit seiner Übertragungsgeschwindigkeit von maximal 1 GBit/s den Anforderungen des Kunden nicht mehr gewachsen. Größere Kopiervorgänge von mehreren Gigabyte innerhalb des Storage Area Network verlangsamen die Datenverarbeitung der virtuellen Maschinen signifikant, was wiederum die Arbeit der Entwickler ausbremst. Die Geschwindigkeit für Datenübertragungen innerhalb des SAN soll auf 10 GBit/s erhöht werden.

Speicherkapazität

Die im bisherigen SAN zur Verfügung stehende Festplattenkapazität von 1.000 GB ist nahezu erschöpft. Nach der Erneuerung des Speichernetzes sollen - so der Wunsch des Kunden - mindestens 5 TB⁵³ Speichervolumen zur Verfügung stehen, die nachträgliche Erweiterung auf ein Vielfaches sollte problemlos möglich sein.

Hochverfügbarkeit

Die Komponenten des neuen SAN sollen so weit redundant ausgelegt werden, dass bei Ausfall eines Gerätes die Funktionsfähigkeit nicht eingeschränkt wird oder nur eine kurze Unterbrechung bemerkbar ist.

⁵³TB: Terabyte (1 TB = 1.024 GB)

Datensicherheit

Die Sicherheit der Daten wird aktuell nur durch das RAID 1 der Disk-Subsysteme sichergestellt - es gibt keine Absicherung für Katastrophen, wie z.B. ein Feuer im Serverraum. Die neuen Disk-Subsysteme sollten mit einem Hardware RAID-Controller ausgestattet sein, der mindestens RAID Level 5 unterstützt. Zudem soll eine Kopie aller Daten in einem anderen Brandabschnitt des Firmengebäudes abgelegt werden.

Vereinfachtes Management

Statt der zwei Webinterfaces⁵⁴ der iSCSI-Disk-Subsysteme soll es eine einheitliche Oberfläche zur Verwaltung des gesamten Datenspeichers geben, um dem IT-Administrator vor Ort die Arbeit zu erleichtern.

Proaktives Monitoring

Um die Ausfallsicherheit der IT-Infrastruktur weiter zu maximieren, soll das SAN mit einem proaktiven Monitoring-System überwacht werden. Anhand eines Service Level Agreements⁵⁵ werden dem externen IT-Betreuer Berechtigungen eingeräumt, bei bestimmten Fehlern selbst - ohne vorherige Absprache mit dem Kunden - einzugreifen.

Lauffähigkeit von Spezialsoftware

Der Kunde verwendet eine Spezialsoftware auf einem der Server, welche zur Ausführung das Anstecken eines USB-Dongle⁵⁶ erfordert. Das USB-Gerät ist an einen der Virtualisierungsserver angeschlossen und wird an die betreffende virtuelle Maschine durchgereicht. Daraus ergibt sich die Anforderung, dass das Dongle am neuen Server funktionsfähig eingerichtet werden muss.

⁵⁴Webinterface: Verwaltungsoberfläche, die nicht in Form einer lokal installierten Software, sondern als Webseite im Browser zur Verfügung steht

⁵⁵Service Level Agreement: Supportvertrag zwischen IT-Dienstleister und Betreuer, legt fest, welche Leistungen der Kunde in welchem Umfang vertraglich zugesprochen bekommt

⁵⁶USB-Dongle: eine Art Software-Kopierschutz, das betreffende Programm funktioniert nur, wenn ein spezielles USB-Gerät vom Hersteller („Dongle“) am Computer angesteckt ist

4.1.2 Auswahl der Technologie

Nach Festlegung der Ziele gilt es nun, eine geeignete SAN-Technologie für die Realisierung des Projektes auszuwählen, dabei wird das theoretische Wissen aus dem Kapitel „Theorie möglicher Techniken“ in den Entscheidungsprozess mit einbezogen. Verglichen werden die vorgestellten SAN-Konzepte iSCSI, Fibre Channel und Fibre Channel over Ethernet.

Zum Vergleich der Konzepte betrachten wir zunächst die Effizienz der verwendeten Protokolle. Der zugehörige Prozentwert ergibt sich aus dem Verhältnis der Nutzdaten (Payload) zur Größe eines gesamten Datenpaketes inklusive Header. Je größer der Effizienzwert, desto kleiner ist der Overhead des Protokolls - der Wert kann mit folgender Formel berechnet werden:

$$P_{eff} = \left(\frac{Payload}{Payload+Overhead} \right) * 100$$

Formel 2: Protokolleffizienz

Die Effizienz der verglichenen Protokolle ist im Diagramm auf **Abbildung 9** dargestellt.

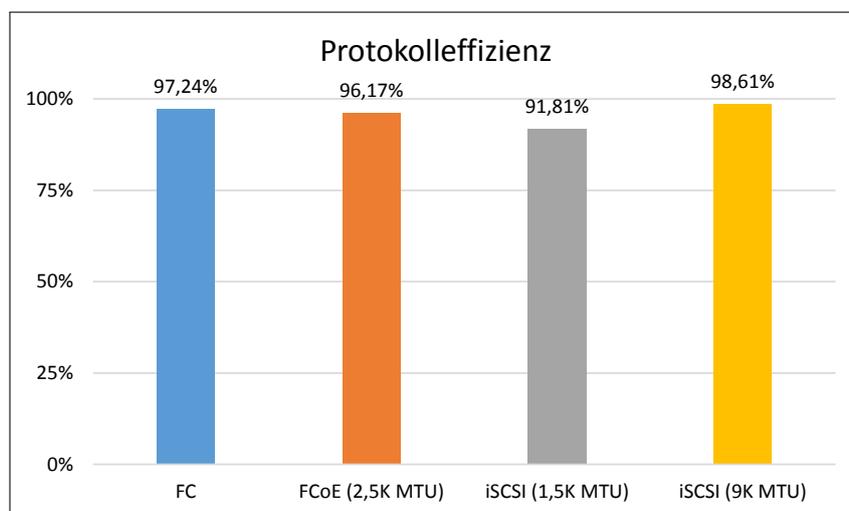


Abbildung 9: Protokolleffizienz der SAN-Konzepte

Quelle: [Raj11, Bild 1], Zahlen übernommen, Grafik neu gezeichnet

Auf den ersten Blick schneiden FC und FCoE in diesem Vergleich besser ab als iSCSI. Jedoch kann das iSCSI-Protokoll mit zwei verschiedenen MTU⁵⁷-Größen betrieben werden. Wenn die zugrunde liegende Netzwerktechnik dies unterstützt, wird iSCSI

⁵⁷MTU: Maximum Transmission Unit - maximale Größe eines Paketes in der Datenübertragung [Nin14]

durch das Aktivieren von Jumboframes (MTU: 9K statt 1,5K) effizienter als FC und FCoE.

Ein weniger theoretisches Vergleichs-Kriterium stellt der Durchsatz bei Lese- und Schreiboperationen dar. Im Rahmen einer Studie der Dell Storage Product Marketing Group aus dem Jahre 2009 wurden die in diesem Projekt betrachteten Protokolle hinsichtlich ihres Durchsatzes verglichen, die Gegenüberstellung ist in **Abbildung 10** dargestellt.

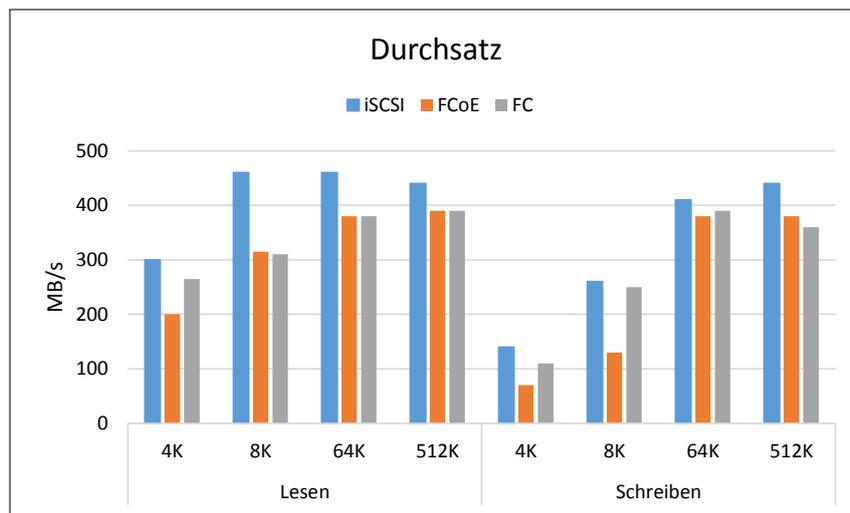


Abbildung 10: Durchsatz der SAN-Konzepte

Quelle: [Raj11, Bild 2], Werte übernommen, Grafik neu gezeichnet

Aus dieser Grafik geht hervor, dass iSCSI bei Einsatz aktueller Hardware (10 GBit/s Ethernet) und aktivierten Jumboframes höhere Übertragungsraten erbringt als FC und FCoE. Da die Protokolle verschiedene Maximalgeschwindigkeiten zur Datenübertragung haben, wurden im Test nur 4 Gigabyte Daten übertragen, um vergleichbare Werte zu erhalten. [Raj11]

Ein weiteres Kriterium ist die maximale Länge einer einzelnen Strecke der Verkabelung. Ein Vergleich ist auf **Abbildung 11** zu sehen.

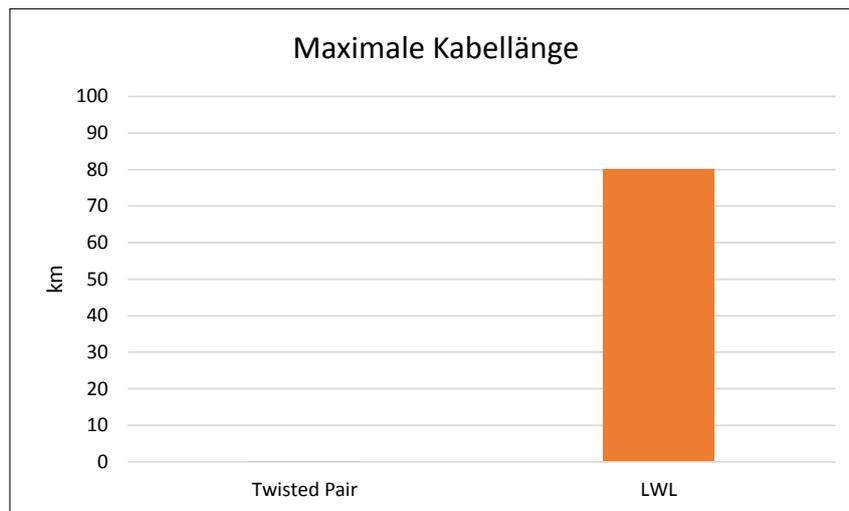


Abbildung 11: Maximale Kabellängen von Twisted Pair und LWL

Quelle: [Wik14a], Werte übernommen, Grafik gezeichnet

Für die Verkabelung mit Twisted Pair-Kabeln darf eine einzelne Strecke nicht länger als 100 Meter sein. Abhängig von Kabeldurchmesser, Medientyp und verwendetem Laser können mit Glasfaserverkabelung problemlos mehrere Kilometer überbrückt werden.

Für die endgültige Entscheidung, welche Technologie eingesetzt werden soll, ist das Kriterium der Kabellänge weniger maßgebend. Der Kunde fordert eine Kopie seines Datenbestandes in einem anderen Brandabschnitt seiner Niederlassung. Der dafür ausgewählte Raum befindet sich ca. 400m vom Serverraum entfernt - diese Entfernung lässt unter Verwendung von 10 Gigabit-Ethernet und Glasfaserverkabelung problemlos überbrücken. Laut der obigen Vergleiche ist iSCSI das Protokoll mit der besten Performance und die Komponenten sind im Vergleich zu Fibre Channel preisgünstiger. Das iSCSI-Protokoll baut zudem auf bewährte Netzwerkprotokolle auf und ist routingfähig, so ist für spätere Erweiterungen beispielsweise die Anbindung der zweiten Niederlassung über VPN möglich. iSCSI ist verglichen mit den anderen vorgestellten SAN-Konzepten am längsten mit auf dem Markt, daher unterstützen sehr viele SAN-Komponenten diese Technologie. Aus diesen Gründen fällt die Entscheidung der zu verwendenden Technologie auf iSCSI.

4.1.3 Dimensionierung des SAN

Um möglichst wenige Technologien zu vermischen und so potenziellen Fehlerquellen und Inkompatibilitäten aus dem Weg zu gehen, wird das gesamte SAN in iSCSI umgesetzt. Es werden zwei identische Storages verwendet, um den Datenbestand in zwei verschiedenen Brandabschnitten vorzuhalten. Die Vernetzung des Storage Area Network erfolgt mit Glasfaserkabel. Um die Verfügbarkeit des Gesamtsystems zu erhöhen, werden die Verbindungsleitungen und alle am Datentransfer beteiligten Switches redundant ausgelegt.

Zum Schutz der Daten gegen Defekte von Festplatten werden die logischen Datenträger auf den Storages mit RAID-Level 5 konfiguriert. Mit der folgenden Formel kann die Kapazität eines Level-5-RAID berechnet werden:

$$C_{R5} = (n - 1) * C_{Disk}$$

Formel 3: RAID 5-Kapazität

Dabei steht n für die Anzahl der zur Verfügung stehenden Festplatten und C_{Disk} für die Kapazität einer einzelnen Platte. In der Formel wird die Anzahl der beteiligten Platten um eine verringert, da im Level-5-RAID die Kapazität einer Platte nur für Paritätsinformationen verwendet wird.

Das RAID 5 übersteht den Ausfall von einer Festplatte ohne Probleme. Die Kapazität einer Festplatte ist für Paritätsinformationen reserviert, welche über alle Festplatten verteilt sind. Fällt eine der Platten aus, können die „verlorenen“ Daten aus den Inhalten der anderen Festplatten und den Paritätsinformationen berechnet werden. Zur weiteren Erhöhung der Ausfallsicherheit wird eine zusätzliche Festplatte als so genannte „Hot Spare“ eingerichtet - diese wird nicht initialisiert und springt bei Defekt einer Platte des Produktivsystems ein. So werden bei Ausfall einer wichtigen Festplatte deren nicht zugreifbare Daten auf der Hot Spare neu aufgebaut. Der defekte Speicher wird durch einen neuen ersetzt und übernimmt von da an die Rolle der Hot Spare. Um Datenverlusten aktiv entgegen zu wirken, springt die Hot Spare-Festplatte auch schon bei einem Pre-Failure⁵⁸ ein. Alternativ kann die Storage auch so konfiguriert werden, dass nach Ersetzen der defekten Festplatte die Daten von der Hot Spare zurückkopiert werden und diese beim nächsten Ausfall einspringen kann.

⁵⁸Pre-Failure: Zustand, der das baldige Auftreten eines Fehlers ankündigt; Beispiel: Festplatte hat eine hohe Lesefehlerrate oder defekte Sektoren

Pro Storage sind drei logische Datenträger geplant:

- 1x **RAID 5 - Variante 3:1**
 - 5x 1,0 TB-Festplatten (3 TB Daten, 1 TB Parität, 1 TB Hot Spare)
 - Kapazität: 3,0 TB
 - Nettokapazität⁵⁹: 2,79 TB
- 2x **RAID 5 - Variante 4:1** mit jeweils
 - 6x 900 GB-Festplatten (3,6 TB Daten, 900 GB Parität, 900 GB Hot Spare)
 - Kapazität: 3,6 TB
 - Nettokapazität: 3,35 TB

Für den Aufbau des SAN werden des Weiteren vier Ethernet-Switches mit SFP+-Modulen eingeplant. Damit ist in jedem der beiden Brandabschnitte mit zwei Switches auch die Redundanz der Anbindung gewährleistet.

Die veralteten Virtualisierungsserver werden durch jeweils zwei neue ersetzt, um eine Replikation in den anderen Brandabschnitt zu realisieren.

Benötigt für das SAN:

- 2x Storage-Einheit
- 4x SFP+-Ethernet-Switches
- 4x ESXi⁶⁰-Server
- 10x 1,0 TB-Festplatten (5x pro Storage)
- 24x 900 GB-Festplatten (12x pro Storage)

⁵⁹Nettokapazität: Hersteller von Festplatten rechnen die Speichereinheiten mit dem Faktor 1.000 um, computerintern wird jedoch mit 1.024 umgerechnet. Die Kapazität ist die Herstellerangabe, die Nettokapazität der reale Wert. Umrechnung: $C_{Netto} = C_{Hersteller} * 0,9313$

⁶⁰ESXi: Virtualisierungssystem von VMware, Nachfolger von ESX

4.1.4 SAN-Entwurf

Im Anschluss an die Dimensionierung des SAN wird ein grafischer Entwurf erstellt. Der fertige Entwurf zeigt nur das Storage Area Network ohne das lokale Netzwerk des Kunden und ist in **Abbildung 12** dargestellt.

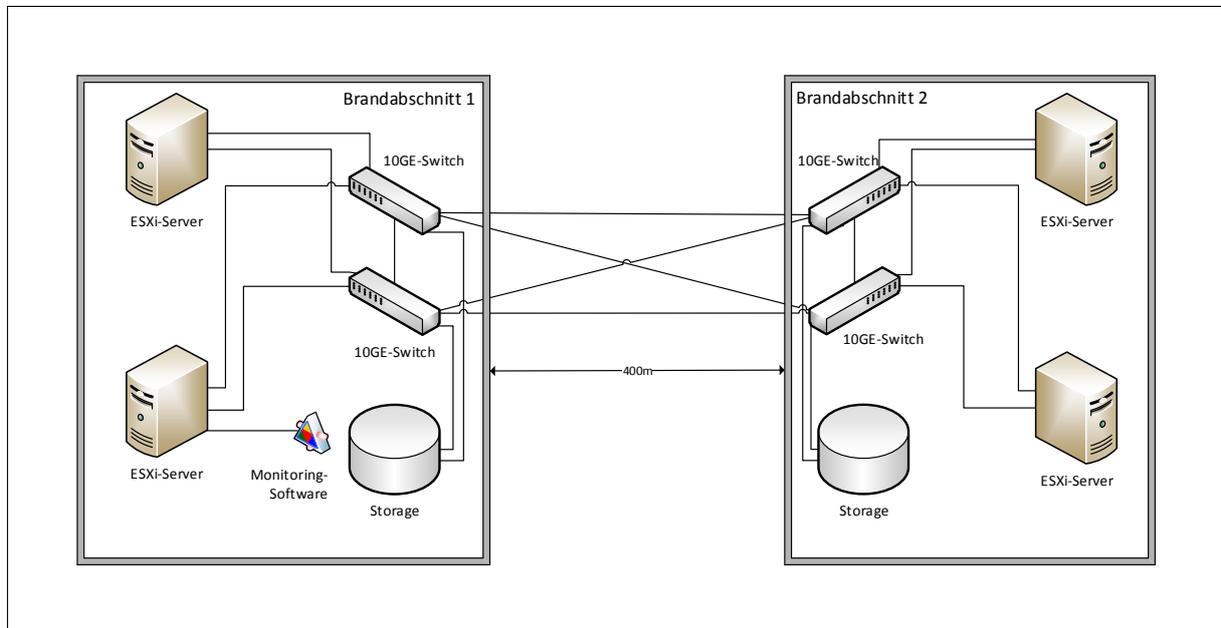


Abbildung 12: SAN-Entwurf

Als Technologie für das SAN wurde iSCSI gewählt, da die Komponenten preisgünstiger als Fibre Channel-Komponenten sind und mit moderner Netzwerkhardware Geschwindigkeiten von bis zu 10 GBit/s erreicht werden - damit wird erhöht sich die bisherige Geschwindigkeit um den Faktor 10 (1 GBit/s zu 10 GBit/s), das SAN ist damit für die Zukunft gerüstet.

Zum Erreichen der vom Kunden gewünschten Datensicherheit werden zwei Storage-Einheiten in verschiedenen Brandabschnitten installiert. Die Storage im Serverraum wird an die ESXi-Server angebunden und repliziert ihre Daten auf die zweite Storage. Die Sicherheit wird weiter erhöht durch die Datenspeicherung im RAID 5 mit zusätzlichen Hot Spare-Festplatten. Im zweiten Brandabschnitt sind zwei weitere Server aufgestellt, welche die Daten der produktiven ESXi-Server replizieren und im Notfall einspringen können.

Die Verwaltungsoberfläche der Storage-Einheiten kann auf jedem beliebigen Computer im Netzwerk aufgerufen werden. Durch die Replikation muss nur noch die Kapazität einer Storage überwacht werden, statt wie bisher von zwei.

Für die Verbesserung der Verfügbarkeit des Storage Area Network werden beide Switches redundant ausgelegt. So funktioniert die Replikation auch noch, wenn in jedem Abschnitt ein Switch ausfällt. Um Fehler wie defekte Festplatten frühzeitig zu erkennen, wird auf einer der virtuellen Maschinen eine Monitoringsoftware installiert.

4.1.5 Auswahl geeigneter Komponenten

Virtualisierungsserver

Für den Betrieb der virtuellen Maschinen werden vier Virtualisierungsserver benötigt, da im Zuge der Erneuerung des Storage Area Network auch die vorhandenen zwei Fujitsu PRIMERGY RX300 S5-Server ausgetauscht werden sollen. Da unsere Firma als Fujitsu Select Partner gelistet ist und unsere Servicetechniker entsprechend geschult sind, fällt die Wahl ebenfalls auf ein Servermodell der Marke Fujitsu. Für das Projekt wurde ein Modell der gleichen Serie wie die vorhandenen Server - der Fujitsu PRIMERGY RX300 S8 - wie in **Abbildung 13** zu sehen, ausgewählt.



Abbildung 13: Fujitsu PRIMERGY RX300 S8

Quelle: Fujitsu Partner Portal [Fuj14c]

Überblick Leistungsdaten

- Chipsatz: Intel® C600
- Prozessor: 2x Intel® Xeon® E5-2650v2
- RAM-Steckplätze: 24x DIMM (12x pro CPU) DDR3
- Netzwerk-Schnittstellen: 2x GBit-Ethernet (RJ45), 1x GBit-Ethernet (RJ45) Management LAN

Die aufgeführten Leistungsdaten sind nur ein Auszug - die komplette Beschreibung ist im Datenblatt des Servers [FUJ14b] zu finden.

Um den Ansprüchen des Projektes besser gerecht zu werden, wurden die Server mit folgenden zusätzlichen Komponenten konfiguriert:

- **12x 8 GB DDR3⁶¹ 1600 ECC**
RAM-Riegel mit einer Kapazität von je 8 GB, auf 1600 MHz getaktet, mit ECC-Unterstützung; Gesamtkapazität: 96 GB
- **2x 500 GB HDD⁶²**
Zwei Festplatten für einen internen Datastore mit Level-1-RAID als Ablage für Arbeitsdaten des ESXi, wie z.B. Betriebssystem-Installations-Datenträger
- **1x 2-Port 10 GBit Ethernet Controller PCIe⁶³ x8 SFP+⁶⁴**
Zusätzliche 10 Gigabit-Netzwerkkarte mit 2 SFP+-Ports, angeschlossen über PCIe x8
- **2x SFP+ Modul 10GbE Multimode**
SFP+ Module (10 Gigabits) zur Anbindung des Servers an das SAN über Glasfaserkabel, verwendet Multimode-Laser
- **1x iRMC⁶⁵ S4 Advanced Pack**
Erweiterungschip zur systemstatus-unabhängigen Fernsteuerung des Servers
- **1x VMware ESXi Embedded UFM Device**
interner USB-Stick zur Installation von VMware ESXi
- **2x VMware vSphere 5.5 Standard**
Lizenzschlüssel für ESXi, 1 Schlüssel pro CPU benötigt

⁶¹DDR3: Double Data Rate 3 - ein Arbeitsspeicher-Typ

⁶²HDD: Hard Disk Drive: Festplatte

⁶³PCIe: PCI Express

⁶⁴SFP+: Small Form-factor Pluggable Plus - Bezeichnung für einen modularen Geräteanschluss

⁶⁵iRMC: Integrated Remote Management Controller

Storage

Ebenfalls aus dem Hause Fujitsu stammt die für das Projekt ausgewählte Storage-Einheit. Die ETERNUS DX100 S3 - zu sehen auf **Abbildung 14** - ist eines der Einstiegermodelle für den SAN-Aufbau. Sie unterstützt Kapazitäten bis zu 1.056 Terabyte und alle gängigen RAID-Level. Im Gegensatz zur nächstkleineren Variante - der DX60 - wird die Replikation auf eine andere Storage-Einheit unterstützt, was für Datenhaltung in zwei Brandabschnitten benötigt wird.



Abbildung 14: Fujitsu ETERNUS DX100 S3

Quelle: Fujitsu Partner Portal [Fuj14c]

Überblick Leistungsdaten

- Anzahl RAID-Controller: 1
- Maximale Cache-Kapazität: 8 GB
- Redundanzen: Lüfter, Netzteil
- Verfügbare RAID-Level: 0, 1, 1+0, 5, 5+0, 6
- Host-Schnittstellen: FC, FCoE, SCSI, Ethernet, SAS
- Anzahl Laufwerke: bis zu 24 pro Einheit

Weitere Spezifikationen zur ETERNUS DX100 S3 sind im zugehörigen Datenblatt [FUJ14a] zu finden.

Wie schon die ESXi-Server werden auch die Storage-Einheiten auf das Projekt abgestimmt. Folgende Komponenten wurden zusätzlich konfiguriert:

- **1x zusätzlicher RAID-Controller**
redundante Auslegung des RAID-Controllers als zusätzliche Ausfallsicherheit
- **2x 4 GB Cache**
pro RAID-Controller 4 Gigabyte Cache-Speicher
- **1x ETERNUS Snapshot Manager**
Software zur Erstellung und Verwaltung von Snapshots des Datenbestandes
- **1x 3 Jahre Vor-Ort-Service**
kostenfreier technischer Support und Ersatzteile vor Ort
- **2x CM w 1xCA iSCSI 10G 2port**
pro RAID-Controller ein iSCSI-Modul mit integriertem SFP+ für Glasfaserkabel

Durch die redundante Konfiguration der Storage mit zwei RAID-Controllern, der Verwendung eines Level-5-RAID bei der Datenspeicherung und der Daten-Replikation in einen anderen Brandabschnitt ist die Hochverfügbarkeit der Speichersysteme gewährleistet.

Switches

Die im SAN verwendeten Switches müssen sorgfältig ausgewählt werden, da sie den gesamten Datenverkehr transportieren und dafür eine ausreichende Geschwindigkeit haben sollten. Die Forderung an den Switch ist eine Übertragungsrate von 10 Gigabit. Da sich diese Geschwindigkeit nur über Glasfaserkabel realisieren lässt, wird ein Gerät mit ausreichend SFP+-Schnittstellen benötigt. Die Wahl fiel auf den DXS-3600-16S des Herstellers D-Link (**Abbildung 15**). Er verfügt über 8 SFP+-Schnittstellen und zwei RJ45-LAN-Ports.

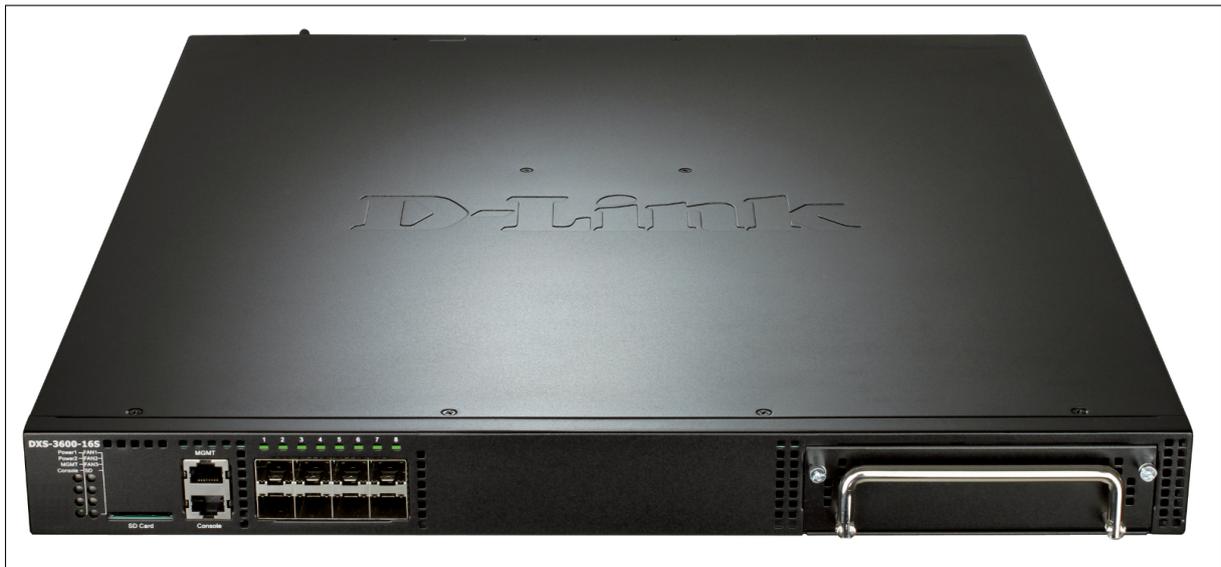


Abbildung 15: D-Link DXS-3600-16S

Quelle: D-Link Deutschland GmbH [D-L13a]

Für die Switches wurden zusätzlich jeweils sieben SFP+-Module „DEM-431XT“ bestellt. Sie werden zum Anschluss der folgenden Komponenten benötigt:

- 2x Anschluss ESXi-Server
- 2x Verkabelung zum 2. Brandabschnitt
- 2x Anschluss der Storage
- 1x Verkabelung zum 2. Switch im gleichen Brandabschnitt

Detaillierte Informationen zum Switch sind im zugehörigen Datenblatt [D-L13b] zu finden.

Festplatten

Die Storage-Systeme wurden mit dem Fujitsu System Architect⁶⁶ konfiguriert, bei diesem Vorgang werden der Typ und die Größe der Festplatte angegeben und die Software schlägt dazu passende Festplatten vor - folgende wurden bestellt:

- 12x **FTS:ETFDB9**

Typ: SAS⁶⁷; Kapazität: 900 GB; Formfaktor: 2,5"; Spindelgeschwindigkeit: 10.000 rpm

- 5x **FTS:ETFNA1**

Typ: NL-SAS⁶⁸; Kapazität: 1,0 TB; Formfaktor: 2,5"; Spindelgeschwindigkeit: 7.200 rpm

4.2 Auftrag

Nach einer umfangreichen Absprache über das geplante Konzept mit dem Kunden wurde durch unsere Firma ein Angebot über die Erneuerung des Storage Area Network und der Virtualisierungsserver erstellt, dieses ist in Anhang 3 zu finden. Im Anschluss an die Bestätigung des Kunden wurde das Projekt zeitnah innerhalb weniger Tage realisiert, dieser Vorgang ist im nächsten Kapitel beschrieben.

⁶⁶Fujitsu System Architect: Software zur individuellen Konfiguration von Fujitsu-Hardware

⁶⁷SAS: Serial Attached Storage: ein Festplatten-Anschluss

⁶⁸NL-SAS: Nearline SAS: Festplatte mit SAS-Anschluss und Leistungsdaten einer SATA-Platte

4.3 Umsetzung

4.3.1 Festlegungen für das SAN-Netz

Schon vor dem Aufbau der Hardware beim Kunden wird für jede Komponente eine IP-Adresse festgelegt. Das LAN mit den Clients hat die Netzadresse 192.168.1.0/24 und das bisherige Storage Area Network die 192.168.2.0/24. Damit es während der Umstellung nicht zu IP-Adresskonflikten⁶⁹ kommt, wird für das neue SAN ein neues Netz mit der Adresse 192.168.0.0/24 gebildet. Innerhalb dieses Netzes werden für die Komponenten folgende IP-Adressen und Hostnamen vergeben.

Zuordnung	Hostname	IP-Adresse
Server 1	ESXi-VM1	192.168.0.11/24
Server 1 iRMC	ESXi-VM1-i	192.168.0.16/24
Server 1 Replica	ESXi-VM1-R	192.168.0.12/24
Server 1 Replica iRMC	ESXi-VM1-R-i	192.168.0.17/24
Server 2	ESXi-VM2	192.168.0.21/24
Server 2 iRMC	ESXi-VM2-i	192.168.0.26/24
Server 2 Replica	ESXi-VM2-R	192.168.0.22/24
Server 2 Replica iRMC	ESXi-VM2-R-i	192.168.0.27/24
Server 3	ESXi-VM3	192.168.0.31/24
Server 3 iRMC	ESXi-VM3-i	192.168.0.36/24
Server 3 Replica	ESXi-VM3-R	192.168.0.32/24
Server 3 Replica iRMC	ESXi-VM3-R-i	192.168.0.37/24
Storage	DX100S3	192.168.0.81/24
Storage Replica	DX100S3-R	192.168.0.82/24
Switch	DXS3600	192.168.0.91/24
Switch Replica	DXS3600-R	192.168.0.92/24

Tabelle 2: IP-Adressen im neuen SAN

⁶⁹IP-Adresskonflikt: entsteht, wenn zwei Geräte die gleiche IP-Adresse haben, i.d.R. können die betroffenen Geräte nicht mehr kommunizieren

4.3.2 Vorinstallation der Server

Direkt nach der Lieferung der bestellten Hardware durch unsere Distributoren wird diese in unserer Firma auf den Einsatz beim Kunden vorbereitet. Für dieses Projekt bedeutet das die Vorinstallation der Virtualisierungsserver und Storage-Einheiten.

Als erster Schritt wird das als Zubehör bestellte UFM-Modul, wie auf **Abbildung 16** gelb hervorgehoben, im Server auf das Mainboard gesteckt - es funktioniert wie ein USB-Massenspeicher, wird jedoch nicht extern mit dem Server verbunden, so ist für die Installation von VMware ESXi keine Festplatte notwendig. Das verwendete Modul hat eine Kapazität von 2 Gigabyte, was für die Systemdaten völlig ausreicht. Anschließend wird der Server eingeschaltet und die BIOS⁷⁰-Einstellungen angepasst. Danach wird die RAID-Konfiguration für die beiden 500 GB-Festplatten im RAID 1 eingerichtet.



Abbildung 16: UFM-Modul

Für die Installation von ESXi wird der Server vom Installationsdatenträger gebootet (**Abbildung 17**), dieser wird bei Bestellung der Lizenzen mitgeliefert. Es handelt sich dabei um eine von Fujitsu gebrandete Version, die bereits alle Treiber für Fujitsu-spezifische Hardwarekomponenten des Servers enthält. Wenn die CD fertig gebootet ist, führt der textbasierte Assistent durch den Rest der Installation. Dabei müssen lediglich der Zielspeicher (UFM-Modul), Regions-, Spracheinstellungen und das root-Passwort festgelegt werden. Die Installation verläuft in der Regel ohne Probleme und ist innerhalb weniger Minuten abgeschlossen.



Abbildung 17: Boot von der ESXi-CD-ROM

⁷⁰BIOS: Basic Input Output System - Firmware eines Computers

Um den fertig installierten ESXi-Server uneingeschränkt verwenden zu können, muss außerdem der Lizenzschlüssel hinterlegt werden. Für jede CPU wird ein Lizenzschlüssel benötigt, daher gibt es pro Server zwei bestellte Schlüssel. Auf der VMware-Webseite im MyVMware-Portal werden die beiden Schlüssel hinterlegt und daraus ein neuer generiert, der beide CPUs umfasst und eingetragen werden kann. Dazu verbindet man sich unter Verwendung des vSphere-Client zum Server und wählt „Konfiguration“ > „Lizenzierte Funktionen“ > „Bearbeiten...“. Daraufhin kann in einem Dialogfenster der generierte Lizenzschlüssel hinterlegt werden (**Abbildung 18**).

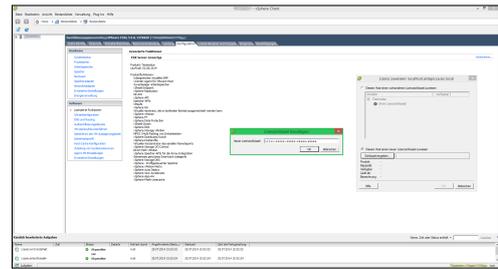


Abbildung 18: ESXi-Konfiguration mittels vSphere-Client

4.3.3 Vorkonfiguration der Storages

Die beiden Storage-Systeme werden ebenfalls vor der Lieferung bei uns in der Firma eingerichtet. Der Prozess der initialen Konfiguration ist in den folgenden Absätzen beschrieben.

Zum Durchführen der Einrichtung wird eine direkte Verbindung zwischen der Storage und einem Notebook per LAN hergestellt - dadurch ist eine Konfiguration über die Web-Oberfläche des Gerätes möglich. Die Standard-Anmeldedaten sind „root“ als Benutzername und als Passwort, die Konfigurationsoberfläche steht nur in englischer Sprache zur Verfügung. Mit einem Klick auf die Schaltfläche „Start Initial Setup“ wird der Konfigurationsassistent gestartet, mit dem die wichtigsten Einstellungen vorgenommen werden.

Im ersten Schritt des Assistenten werden mit Name, Installations-Ort, Administrator und Beschreibung einige Informationen über das Storage-System festgelegt - diese beeinflussen nicht den Betrieb des Gerätes, die Informationen sind in der Verwaltungsoberfläche und in der Management-Software (bei Verwaltung mehrerer Storages) sichtbar und dienen der besseren Unterscheidung.

Anschließend werden Uhrzeit, Datum und Zeitzone eingestellt - die aktuelle Zeit kann händisch eingegeben werden, das ist allerdings nicht notwendig, da ebenfalls die Synchronisation mit einem NTP⁷¹-Server eingerichtet wird. Als Zeitserver wird „ptbti-

⁷¹NTP: Network Time Protocol - Protokoll zur Anpassung von Datum und Uhrzeit an die Vorgabe eines Zeitserverns

me1.ptb.de“ eingestellt, dieser wird von der Physikalisch-Technischen Bundesanstalt betrieben. Die Option „Daylight Saving Time“ (Sommer-/Winterzeit) wird ebenfalls aktiviert, da die vom NTP-Server gelieferte Zeitangabe die Zeitumstellung nicht beinhaltet.

Nach Abschluss der Zeiteinstellungen kann im folgenden Schritt das Passwort des administrativen Benutzers „root“ geändert werden - diese Änderung ist dringend zu empfehlen, da der Benutzer vollen Zugriff auf die Verwaltung der Storage hat und das Standard-Passwort als unsicher gilt, da es allgemein bekannt ist. Das neue Passwort darf eine Länge von mindestens 4 und maximal 64 Zeichen haben.

Im nächsten Schritt kann eine so genannte „Thin Provisioning License“ registriert werden. Wird eine solche Lizenz aktiviert, kann mit der Storage Überprovisionierung betrieben werden, d.h. allen Clients zusammen wird mehr verfügbarer Speicher zugewiesen, als die Storage tatsächlich besitzt - der Speicher wird in diesem Fall dynamisch zugewiesen, wo er gebraucht wird. Da in diesem Projekt keine Überprovisionierung geplant ist, wird die Eingabe der Lizenz übersprungen. [Wik13c]

Nun kann als weitere Option ein Lizenzschlüssel für das Feature „Local Copy“ installiert werden - diese Funktion ermöglicht schnelle Datenübertragung an eine andere Storage als Disaster Recovery⁷². Diese Funktionalität wird nicht benötigt, da eine ständige Replikation zur zweiten Storage eingerichtet wird, die über eine andere Lizenz im nächsten Schritt aktiviert wird - dieser Schritt wird demnach auch übersprungen.

Als Nächstes wird eine Lizenz für „Remote Copy“ eingegeben - diese wird benötigt, um die Replikation auf die zweite Storage zu konfigurieren. Nach Eingabe des Lizenzschlüssels wird mit dem nächsten Schritt fortgefahren.

Eine weitere Option stellt „Self Encrypting Disk“ (kurz: SED) dar, wird diese aktiviert, lassen sich die auf den Platten gespeicherten Daten komplett verschlüsseln - dazu wird in diesem Schritt ein so genannter „SED Authentication Key“ eingetragen. Da die Verschlüsselung der Festplatten im Projekt nicht gefordert ist, wird auch diese Eingabe übersprungen.

Im vorletzten Abschnitt („Set Network Environment“) wird der ETERNUS eine feste IP-Adresse zugewiesen, unter der diese im Storage Area Network erreichbar ist. Außerdem werden bis zu sechs IPv4-Netze anhand von Netzadresse und Subnetzmaske festgelegt, die auf die Storage zugreifen dürfen - Verbindungen anderer Netzwerke werden abgelehnt.

⁷²Disaster Recovery: Methodik zur Datenwiederherstellung im Katastrophenfall

Mit einem weiteren Klick auf „Next“ wird der Einrichtungsassistent abgeschlossen, die Einstellungen gespeichert und angewendet. Nach Abschluss wird der root-Benutzer automatisch abgemeldet, es ist eine Neu-Anmeldung mit dem vorher gesetzten Passwort erforderlich.

Nach der erneuten Anmeldung in der Weboberfläche kann der Assistent nochmals aufgerufen werden, um SNMP⁷³-Einstellungen vorzunehmen - für den Einsatz unserer Monitoring-Software genügen jedoch die Standardeinstellungen. [Aca11a]

Anschließend werden in der Verwaltungsoberfläche die im Abschnitt 4.1.3 (ab Seite 23) definierten RAID-Gruppen angelegt. Die ETERNUS unterstützt das automatische und das manuelle Erstellen dieser Gruppen - eine RAID-Gruppe ist ein Verbund mehrerer Festplatten mit einem festgelegten RAID-Level. Da auf der Storage mehrere RAID-Gruppen eingerichtet werden sollen, wird die manuelle Erstellung verwendet, da sonst die Verwaltungsoberfläche die zu verwendenden Festplatten auswählt. Pro Gruppe wird ein Name vergeben, das RAID-Level (in diesem Fall immer 5) ausgewählt und der Cache Modus auf Automatic gesetzt.

Eine RAID-Gruppe kann mehrere Volumes enthalten, diese stellen die am Ende verfügbaren Datenspeicher dar. Da der Kunde den Speicher der Storage nicht selbst einteilt - lediglich die ESXi-Server legen ihre Daten auf den Volumes ab - werden die Volumes jeweils über die Größe der gesamten RAID-Gruppe erstellt. Für jedes erstellte Volume wird Name, Kapazität, Typ (hier immer „Standard“) und die Zugehörigkeit zur RAID-Gruppe festgelegt. Für jede RAID-Gruppe wird des Weiteren eine Hot Spare-Festplatte zugewiesen, die bei Ausfall einer Platte automatisch deren Funktion übernimmt. Diese werden als „Dedicated Hot Spare“ eingerichtet, sind damit einer RAID-Gruppe fest zugeordnet. Im Gegensatz dazu gibt es auch „Global Hot Spare“, diese Option ist jedoch nur sinnvoll, wenn alle verwendeten Festplatten die gleiche Kapazität haben - so wäre die Festplatte keiner RAID-Gruppe fest zugeordnet. [Aca11b]

⁷³SNMP: Simple Network Management Protocol - Protokoll zur Überwachung und Steuerung von Netzwerkgeräten [Nin14]

4.3.4 Aufbau der Hardware

Eine wichtige Forderung für das Projekt ist es, den Kunden während der Umstellung in seiner Arbeit nicht zu behindern. Deshalb wird die vorhandene Hardware nicht direkt ersetzt. Die neuen Komponenten werden parallel zu den bereits vorhandenen aufgebaut.

Vor dem Anschluss der Glasfaserkabel müssen an den D-Link-Switches und an den PRIMERGY-Servern die zusätzlich bestellten SFP+-Module verbaut werden.

Nach dem Aufbau ist die SAN-Infrastruktur wie im Projektplan eingerichtet, jedoch parallel zur vorhandenen (siehe **Abbildung 1** auf Seite 2) und in einem eigenen Netz. Eine Verbindung zwischen dem neuen SAN und dem Client-LAN des Kunden besteht zu diesem Zeitpunkt noch nicht.

Da jetzt beide Netze parallel existieren, kann das neue SAN direkt in seiner Zielumgebung eingerichtet werden, bevor die Daten übernommen und das alte System abgeschaltet werden.

4.3.5 Konfiguration der Komponenten

Virtualisierungsserver

Auf den Virtualisierungsservern wurde das System VMware ESXi 5.5 bereits in unserer Firma vorinstalliert und der Lizenzschlüssel aktiviert. Jetzt müssen vor Ort einige Einstellungen an das Netz des Kunden angepasst werden.

Die für das Projekt vorgesehenen Server sind mit einem iRMC-Chip ausgestattet. Diese Technologie ermöglicht es uns, den Zustand der Hardware über ein Webinterface einzusehen und den Server fernzusteuern, auch wenn dieser ausgeschaltet ist. Um diese Funktionen zu verwenden, muss iRMC zunächst aktiviert und konfiguriert werden. Dazu wird im BIOS im Reiter „Server Mgmt“ die „iRMC LAN Parameters Configuration“ aufgerufen (**Abbildung 19**), dort wird für den iRMC-Zugriff eine statische IP-Adresse festgelegt. Als iRMC-Schnittstelle

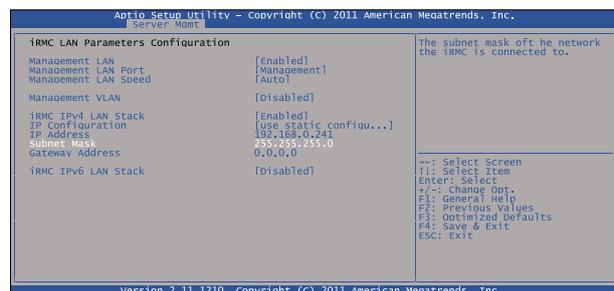


Abbildung 19: iRMC-Einstellungen im BIOS

Quelle: [Aca13] nachgezeichnet

wird der Management-LAN-Port (mit einem Schraubenschlüsselsymbol gekennzeichnet) des Servers verwendet.

Für die ESXi-Installation wurde den Servern in unserer Firma bereits eine feste IP-Adresse zugeteilt. Im produktiven Einsatz sollte für jeden Server eine statische IP-Adresse festgelegt werden, da diese so nicht änderbar ist und nicht von der Verfügbarkeit und Funktionalität eines DHCP⁷⁴-

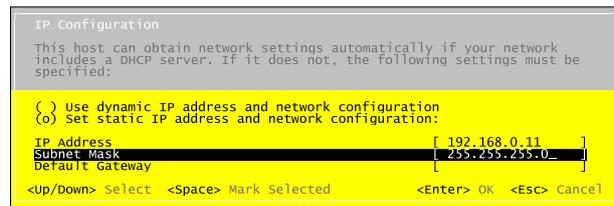


Abbildung 20: IP-Konfiguration bei ESXi

Servers abhängt. Diese Einstellungen werden in der ESXi-Oberfläche unter dem Menüpunkt „Configure Management Network“ (**Abbildung 20**) vorgenommen. Den Servern wurden jeweils die vorgesehenen IP-Adressen aus **Tabelle 2** (Seite 32) zugewiesen. Nach dem Anwenden der Änderungen bietet das System automatisch an, das „Management Network“ (den zuständigen Dienst) neu zu starten - dieser Schritt ist notwendig, damit die neue Konfiguration sofort aktiv ist.

Storage

Die Storages wurden bereits in der Firma vorinstalliert. Nach dem Aufbau und der Vernetzung vor Ort beim Kunden wird mit der Unterstützung eines Storage-Experten von Fujitsu zum Abschluss noch die Replikation auf die zweite ETERNUS eingerichtet.

Switches

Die im Projekt verwendeten Switches sind „managed“ und können sehr anwendungsspezifisch konfiguriert werden. Für die korrekte Funktion des Storage Area Network sind an den Switches keine besonderen Einstellungen nötig. Jedoch wird jedem Switch zur besseren Überwachung eine feste IPv4-Adresse zugewiesen, diese wurde zuvor in **Tabelle 2** auf Seite 32 festgelegt.

⁷⁴DHCP: Dynamic Host Configuration Protocol - Protokoll zur automatischen Vergabe von IP-Adressen

Datastores

Speichermedien auf denen virtuelle Maschinen und andere Arbeitsdaten abgelegt werden, heißen bei VMware „Datastore“. Interne Datastores lassen sich recht einfach konfigurieren, für die externen auf den Storages hingegen muss zunächst eine iSCSI-Verbindung hergestellt werden.

Zur iSCSI-Anbindung wird zunächst in der Netzwerk-Konfiguration des ESXi-Hosts ein virtueller Switch angelegt, der eine Verbindung zwischen der Storage und dem ESXi-System ermöglicht. Durch die Auswahl des VM-Kernel⁷⁵ als Switch-Port wird der virtuelle Switch automatisch als iSCSI-Verbindung gewertet. Anschließend wird in der Speicheradapter-Konfiguration des virtuellen Switches der Speicherort der Storage in Form der IP-Adresse hinterlegt. Weitere Einstellungen wie Authentifizierung sind möglich.

Außerdem wird das so genannte Multipathing konfiguriert. Darunter versteht man, dass zwischen Server und Storage mehrere vordefinierte Verbindungswege bestehen. [DLB10] Dies hat die Vorteile, dass durch parallele Nutzung mehrerer Wege die Latenzzeit verringert wird und redundante Pfade die Ausfallsicherheit des SAN erhöhen. [Rob01] Pro definiertem Pfad werden iSCSI-IP-Adresse des Gerätes, der dazugehörige Port angegeben und falls benötigt die Authentifikationsparameter eingegeben. Es werden alle verfügbaren Pfade eingerichtet, um die maximale Ausfallsicherheit zu erreichen.

Die im SAN erstellten Volumes werden nun als Datastores eingerichtet. Den im nebenstehenden Bild sichtbaren Assistent zur Einrichtung eines Datastore erreicht man im vSphere-Client über den Reiter „Konfiguration“ unter dem Punkt „Speicher“, „Speicher hinzufügen...“. Da es sich um ein iSCSI-SAN handelt, wird als Speichertyp „Festplatte/LUN“ ausgewählt. Im nächsten Schritt wird die aktuelle Dateisystemversion - VMFS5 - gewählt. [Wik14e] Im weiteren Verlauf können die Größe des Datastore und weitere Eigenschaften konfiguriert werden. Zum Abschluss wird der Datastore mit

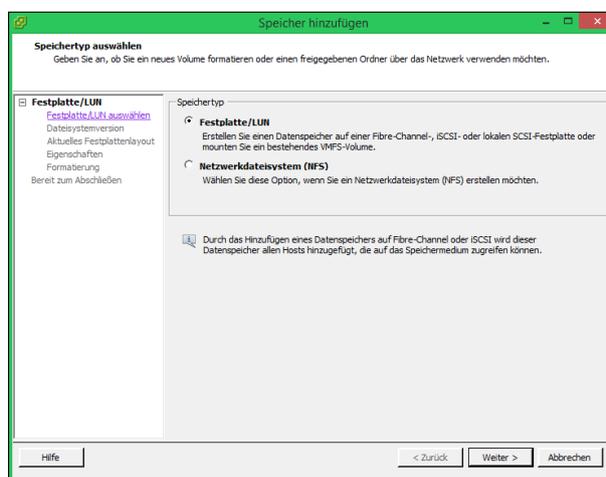


Abbildung 21: Datastore-Einrichtung in ESXi

⁷⁵VM-Kernel: zentrale Softwarekomponente des ESXi

VMFS5 formatiert. Um den gleichen Datastore an den jeweiligen Replica-Server anzubinden, wird auch der Assistent verwendet, allerdings wird die Replica-Storage als Ziel gewählt und das Formatieren am Ende des Assistenten entfällt.

4.3.6 Migration vorhandener Daten

Das neue Speichernetz ist fertig aufgebaut und eingerichtet, nun gilt es, ca. 980 Gigabyte Daten vom alten in das neue SAN zu kopieren. Damit die Arbeit des Kunden durch den Kopiervorgang nicht beeinträchtigt wird, startet dieser an einem Freitagabend, so wird die Umstellung auf das Wochenende verlagert.

Beim Kopieren von Daten sind Schreibvorgänge in der Regel langsamer als Lesevorgänge - würden die Daten direkt vom alten ins neue SAN kopiert, wäre die Ethernet-Verbindung aufgrund der langsamen Komponenten des alten Speichernetzes auf 1 GBit/s limitiert. Deshalb werden die neuen ESXi-Servers an das LAN des Kunden angeschlossen, so stehen für die schnelle Leseoperation nur 1 GBit/s, für die langsamere Schreiboperation jedoch eine Bandbreite von 10 GBit/s zur Verfügung. Die Datenübertragung lässt sich demzufolge mit einem Server als Brücke beschleunigen. Des Weiteren wird ein VPN-Zugang eingerichtet, damit nach Dienstschluss von extern an den Servern gearbeitet werden kann.

Nach Dienstschluss des Kunden startet die Migration der Daten. Dazu wird per VPN mit dem vSphere-Client das alte SAN - welches über GBit-LAN verbunden ist - vorübergehend als Datastore eingerichtet, das Vorgehen dazu ist analog zur Beschreibung im Kapitel „Konfiguration der Komponenten“. Die Verwaltungssoftware vSphere-Client unterstützt zwar das Verschieben von Daten zwischen Datastores, jedoch kein Kopieren.

Um die Dateien dennoch zu kopieren, wird auf dem Host der Dienst SSH⁷⁶ aktiviert, die Einstellung dazu ist unter „Sicherheitsprofil“ in der Konfiguration zu finden. Sobald der SSH-Dienst gestartet ist, wird mit dem Programm PuTTY⁷⁷ eine SSH-Verbindung zum Server hergestellt - die Anmeldedaten sind dabei die gleichen wie im vSphere-Client. Nach dem Login befindet man sich in einer Terminalsitzung auf dem Server und kann ihn mit Befehlen steuern. Mit folgendem Befehl werden die verfügbaren Datastores angezeigt:

```
ls /vmfs/volumes/
```

⁷⁶SSH: Secure Shell - ein Dienst zum Remote-Zugriff auf die Kommandozeile eines Betriebssystems

⁷⁷PuTTY: ein Client für SSH-Verbindungen unter Windows

Das vorhandene Storage Area Network wurde unter dem Namen „SAN-alt“ eingebunden und wird in der Auflistung der Verzeichnisse mit angezeigt. Mit folgendem Befehl wird in das Verzeichnis gewechselt:

```
cd /vmfs/volumes/SAN-alt/
```

Vor der Datenübernahme werden auf den alten Servern alle virtuellen Maschinen ordnungsgemäß heruntergefahren, um Inkonsistenzen zu vermeiden. Anschließend wird der Kopiervorgang mit folgendem Befehl gestartet:

```
nohup cp -rf * /vmfs/volumes/SAN/ &
```

Dabei ist „cp“ der Kopierbefehl von Linux⁷⁸/ESXi. Als erster Parameter erhält der Befehl ein Asterisk, welches als Platzhalter alle Dateien zum Kopieren auswählt. Der zweite Parameter ist der Pfad zum Ziel, hier ist der Datastore im neuen SAN angegeben. Die Option „r“ gibt an, dass rekursiv - mit allen Unterverzeichnissen - kopiert werden soll und „f“ erzwingt das ungefragte Überschreiben, falls im Ziel gleichnamige Dateien vorhanden sind. Der Kopierbefehl wird mit dem Kommando „nohup &“ gekapselt, welches dafür sorgt, dass der Kopiervorgang im Hintergrund stattfindet - so wird dieser auch nicht abgebrochen, wenn die SSH-Verbindung beendet wird.

Da der Lesevorgang bei der Übertragung mit maximal 1 GBit/s langsamer ist als der Schreibvorgang, kann anhand der Lesegeschwindigkeit die ungefähre Dauer der Übertragung mit folgender Formel berechnet werden:

$$t = \frac{1024MB}{125MB/s * 3600s} * 980GB$$
$$t = 2,23h$$

Formel 4: Übertragungsdauer

Der Bruch in der Formel ist zur Berechnung der Übertragungsdauer für ein Gigabyte (1.024 MB) bei 1 GBit/s (entspricht 125 MB/s) in Stunden (3600 s). Der zweite Faktor ist die Gesamtdatenmenge in Gigabyte. So ergibt sich eine theoretische Übertragungsdauer von 2,23 Stunden. In der Praxis dauert die Übertragung länger, da die Formel den Overhead der Übertragungsprotokolle und die Paketgrößen nicht einbezieht.

Nach Abschluss des Kopiervorganges werden ebenfalls über VPN unter Verwendung des vSphere-Client die virtuellen Maschinen in die Bestandslisten der Hosts aufgenommen. Dazu wird im „Datenspeicherbrowser“ der Ordner von jeder virtuellen Maschine geöffnet, die darin enthaltene VMX-Datei markiert und im Kontextmenü der Punkt

⁷⁸Linux: ein freies unixartiges Betriebssystem; bildet die Basis von VMware ESXi

„Zur Bestandsliste hinzufügen“ gewählt. Damit ist die Migration der Daten abgeschlossen.

4.3.7 Funktionstests

Als erster Funktionstest werden die virtuellen Maschinen direkt nach der Migration auf den Servern im Serverraum hochgefahren. Ob wirklich alles reibungslos funktioniert, zeigt sich jedoch erst zu Beginn der neuen Woche, wenn die Anwender aktiv mit dem System arbeiten. Beim Funktionstest durch die Mitarbeiter ist ein Problem aufgetreten, welches jedoch schnell behoben werden konnte, es ist im Kapitel „Probleme“ näher erklärt.

4.3.8 Einrichtung des Monitoring

Unsere Firma setzt zur Überwachung der Kundennetzwerke die Software „PRTG Network Monitor“ aus dem Hause Paessler ein. Für die Einrichtung genügt es, auf einer der virtuellen Maschinen den „PRTG Remote Probe Installer“ auszuführen, dieser installiert eine Anwendung, die durch den PRTG-Server in unserer Firma gesteuert wird, alle relevanten Daten im lokalen Netzwerk erfasst und gebündelt an unseren Server sendet. Nach Installation muss der neuen Probe lediglich ein Name zugewiesen und die öffentliche Adresse unseres PRTG-Servers eingetragen werden.

Die weitere Konfiguration erfolgt firmenintern über die PRTG-Weboberfläche, hier können alle Geräte angelegt und mit verschiedenen Sensortypen (SNMP, WMI⁷⁹, HTTP⁸⁰) versehen werden. Dabei werden durch einen automatischen Suchlauf viele Sensoren von PRTG selbst angelegt.

⁷⁹WMI: Windows Management Instrumentation - Technologie zum Auslesen und Manipulieren von Daten in Windows-Systemen

⁸⁰HTTP: Hypertext Transfer Protocol - Protokoll zur Datenübertragung, vorrangig im Internet für Webseitenaufrufe verwendet

4.4 Probleme

4.4.1 Migration der Daten

Die Migration der Daten gestaltet sich zwischen zwei SAN-Systemen schwieriger als mit traditionellen Speichermedien. Zwar benutzen beide Speichernetze die gleiche Technologie (iSCSI), verwenden jedoch verschiedene IPv4-Netze zur Kommunikation und arbeiten mit unterschiedlichen Übertragungsgeschwindigkeiten. Diese Faktoren haben das Kopieren der Daten erschwert. Als Lösung wurde einer der neuen ESXi-Server als Brücke für die Daten eingesetzt.

4.4.2 Kopiervorgang mit ESXi

Die grafische Oberfläche vSphere-Client erlaubt zwar das Verwalten von Dateien auf Datastores, jedoch nur sehr eingeschränkt. Zwischen zwei Datenspeichern ist nur ein Verschieben von Dateien möglich, kein Kopieren. Ein direktes Verschieben der Daten wurde allerdings ausgeschlossen, diese sollen als Sicherungskopie erhalten bleiben, falls es mit dem neuen SAN Startschwierigkeiten gibt. Der Download und anschließende Upload über den vSphere-Client stellt aufgrund der enormen Datenmenge keine Alternative dar. Das Problem wurde gelöst, indem die Daten wie im Abschnitt 4.3.6 beschrieben über SSH kopiert wurden.

4.4.3 Übersehene USB-Hardware

Während der ersten Erprobung der Arbeit mit dem neuen SAN wurde von einigen Mitarbeitern festgestellt, dass eine auf einem Server eingerichtete Spezialsoftware nicht funktioniert. Die Ursache des Problems war ein zum Start der Software benötigtes USB-Dongle, welches noch am alten Server angeschlossen war. Der Umzug des Dongle an den neuen Server war für das Projekt mit eingeplant, wurde jedoch bei der Umsetzung übersehen. Dieses Problem konnte durch Umstecken des Gerätes an den neuen Server schnell behoben werden.

5 Abschluss

5.1 Übergabe an den Kunden

Nach Abschluss des Projektes und einer einwöchigen Testphase wurde das neue Storage Area Network an den Kunden übergeben. Der IT-Verantwortliche Mitarbeiter des Unternehmens wurde in die Funktionsweise und Wartung des Systems eingeführt und kann so bei kleineren Problemen und Störungen auch selbst erste Maßnahmen ergreifen. Seitens der anderen Angestellten war keine Schulung notwendig, da sich in der Bedienung der Client-Systeme nichts geändert hat - das SAN ist für die Mitarbeiter unsichtbar, sie arbeiten lediglich mit den Daten die über die virtuellen Maschinen zur Verfügung gestellt werden, deren IP-Adressen sich beim Umzug auf die neuen Server nicht verändert haben.

5.2 Projekt-Ergebnis

Mit der Beendigung des Projektes „Konzeptionierung und Realisierung der Erneuerung des Storage Area Network (SAN) eines mittelständischen Unternehmens zur Steigerung der Ausfallsicherheit der IT-Infrastruktur“ wurden folgende Ziele erreicht:

- **Mehr SAN-Performance:**

Das bisherige Speichernetz hatte eine Übertragungsgeschwindigkeit von 1 GBit/s, das neue hingegen 10 GBit/s - die theoretisch maximale Performance wurde um den Faktor 10 gesteigert, damit ist das SAN für zukünftige Belastungsanforderungen gerüstet.

- **Erhöhte Speicherkapazität:**

Die im alten Storage Area Network zur Verfügung stehende Speicherkapazität von 1,0 TB war fast ausgeschöpft. Das neue SAN stellt jetzt eine Kapazität von 9,49 TB zur Verfügung und bietet damit reichlich freien Platz für das Datenwachstum des Kunden. Zudem können die Storageeinheiten einfach erweitert werden.

- **Verbesserte Datensicherheit:**

Der komplette Datenbestand wird voll-redundant in zwei verschiedenen Brandabschnitten vorgehalten. So gehen auch bei einem Brand-, Wasserschaden o.ä. keine Daten verloren.

- **Vereinfachtes Management:**

Es gibt nicht mehr zwei Disk-Subsysteme die separat gewartet und verwaltet werden müssen. Das neue Konzept ermöglicht die Verwaltung der Storage über eine einheitliche Weboberfläche.

- **Höhere Ausfallsicherheit:**

Alle am SAN beteiligten aktiven Netzwerkkomponenten sind redundant ausgelegt, so können eine oder in bestimmten Konstellationen auch mehrere Komponenten ausfallen, ohne dass die Funktionsfähigkeit des Speichernetzes beeinträchtigt wird.

- **Frühzeitige Fehlererkennung:**

Durch den Einsatz einer Monitoring-Software wird das Storage Area Network rund um die Uhr überwacht und protokolliert - so können Fehler frühzeitig erkannt und behoben werden.

- **Lauffähigkeit der Spezialsoftware:**

Durch die - wenn auch leicht verspätete - Migration des USB-Dongle an den neuen Server ist die Spezialsoftware des Kunden uneingeschränkt lauffähig.

5.3 Ausblick

Mit der Durchführung des Projektes wurden alle gestellten Ziele mit nur wenigen Komplikationen erreicht. Dem Kunden steht nun wesentlich mehr Speicherkapazität zur Verfügung und das SAN ist bezüglich Erweiterbarkeit und Belastbarkeit für die Zukunft gerüstet. Da die Daten seit der Umstellung zusätzlich in einen anderen Brandabschnitt repliziert werden und die Virtualisierungsserver mit beiden Storages verbunden sind, ist Transparent Failover⁸¹ möglich. Durch einen Supportvertrag und die Einrichtung des Monitoring hat sich die Kundenbindung weiter intensiviert. Wir freuen uns auf eine weiterhin gute Zusammenarbeit.

⁸¹Transparent Failover: Bei Ausfall einer Komponente springt die dafür vorgesehene Ersatzkomponente ein und übernimmt die Dienste. Dieser Vorgang geschieht für den Anwender unbemerkt („transparent“).

Literaturverzeichnis

- [Aca11a] Fujitsu Corporate Training Academy. Practical Exercise ETERNUS Initial Setup. Unterlagen von Fujitsu-Schulung bei Bytec in Friedrichshafen; nicht öffentlich, November 2011.
- [Aca11b] Fujitsu Corporate Training Academy. Practical Exercise ETERNUS RAID Groups and Volumes. Unterlagen von Fujitsu-Schulung bei Bytec in Friedrichshafen; nicht öffentlich, November 2011.
- [Aca13] Fujitsu Corporate Training Academy. Practical Exercise PRIMERGY Entry Maintenance. Unterlagen von Fujitsu-Schulung bei Bytec in Friedrichshafen; nicht öffentlich, November 2013.
- [All97] Allan, I. Dal; u.a. Fibre Channel Physical and Signalling Interface. Technical report, InterNational Committee for Information Technology Standards, 1997. URL: ftp://www.t11.org/t11/member/fc/ph-3/fcph3_94.pdf Zugriff am 05.08.2014 um 12:54 Uhr.
- [Ass14] Storage Networking Industry Association. SNIA-Website. Internetquelle, 2014. URL: <http://www.snia.org/>, Zugriff am 31.07.2014 um 10:07 Uhr.
- [D-L13a] D-Link Deutschland GmbH. Bild DXS-3600-16S, 2013. URL: <http://www.dlink.com/-/media/Images/Products/DXS/3600/DXS360016SB1Image%20LFront.png> Zugriff am 23.07.2014 um 13:27 Uhr.
- [D-L13b] D-Link (Deutschland) GmbH. Datenblatt DXS-3600-Serie, 2013. URL: http://www.dlink.com/-/media/Business_Products/DXS/DXS%203600/Datasheet/DXS_3600_Datenblatt.pdf Zugriff am 23.07.2014 um 15:01 Uhr.
- [DLB10] Roland Döllinger, Reinhard Legler, and Duc Tanh Bui. *Praxishandbuch Speicherlösungen*. dpunkt.verlag, Heidelberg, 2010.
- [For14] Internet Engineering Task Force. IETF-Website: Mission. Internetquelle, 2014. URL: <http://www.ietf.org/about/mission.html>, Zugriff am 31.07.2014 um 10:12 Uhr.

- [FUJ14a] FUJITSU LIMITED GmbH. Datenblatt FUJITSU ETERNUS DX100 S3 Plattenspeichersystem, 2014. URL: <http://globalsp.ts.fujitsu.com/dmsp/Publications/public/ds-eternus-dx100-s3-ww-de.pdf> Zugriff am 23.07.2014 um 14:26 Uhr.
- [FUJ14b] FUJITSU LIMITED GmbH. Datenblatt FUJITSU Server PRIMERGY RX300 S8 Dual-Socket Rack-Server (2 HE), 2014. URL: <http://globalsp.ts.fujitsu.com/dmsp/Publications/public/ds-py-rx300-s8-de.pdf> Zugriff am 23.07.2014 um 10:21 Uhr.
- [Fuj14c] Fujitsu Technology Solutions GmbH. Fujitsu Partner Portal, 2014. URL: <https://partners.ts.fujitsu.com/> Zugriff am 23.07.2014 um 08:52 Uhr.
- [LAT12] LATTICE Semiconductor. 8b/10b Encoder/Decoder, 2012. URL: http://www.latticesemi.com/~media/Documents/ReferenceDesigns/1D/8b10bEncoderDecoder-Documentation.PDF?document_id=5653 Zugriff am 06.08.2014 um 21:41 Uhr.
- [Nin13] Prof. Dr. Reinhardt Nindel. Vorlesung: Rechnernetze. Staatliche Studienakademie Glauchau, 2013.
- [Nin14] Prof. Dr. Reinhardt Nindel. Vorlesung: Spezielle Netze. Staatliche Studienakademie Glauchau, 2014.
- [Pet08] Peter, J. 8b/10b Coding, 2008. URL: http://www.nikhef.nl/~peterj/KM3net/8B10B_Coding.ppt Zugriff am 06.08.2014 um 20:55 Uhr.
- [Raj11] Ujjwal Rajbhandari. Comparing Performance Between iSCSI, FCoE, and FC. Internetquelle, 2011. URL: <http://en.community.dell.com/techcenter/storage/w/wiki/2722.aspx>, Zugriff am 21.07.2014 um 13:37 Uhr.
- [Rob01] Björn Robbe. *Storage Area Network*. Hanser, München, 2001.
- [Sat04] Satran, J.; u.a. Internet Small Computer Systems Interface (iSCSI). Technical report, IETF, Network Working Group, 2004. URL: <http://www.rfc-editor.org/rfc/pdfrfc/rfc3720.txt.pdf> Zugriff am 05.08.2014 um 12:33 Uhr.
- [TEM08] Ulf Troppens, Rainer Erkens, and Wolfgang Müller. *Speichernetze*. dpunkt.verlag, Heidelberg, 2008.
- [uSKG08] Stor IT Back: Annette Bornemann und Stephan Kranz GbR. FCoE V1.0. Internetquelle, 2008. URL: <http://www.storitback.de/service/fcoe.html>, Zugriff am 18.07.2014 um 14:06 Uhr.

- [uSKG13a] Stor IT Back: Annette Bornemann und Stephan Kranz GbR. Fibre Channel V1.6. Internetquelle, 2013. URL: <http://www.storitback.de/service/fibrechannel.html>, Zugriff am 12.07.2014 um 17:04 Uhr.
- [uSKG13b] Stor IT Back: Annette Bornemann und Stephan Kranz GbR. iSCSI V1.8. Internetquelle, 2013. URL: <http://www.storitback.de/service/iscsi.html>, Zugriff am 12.07.2014 um 16:44 Uhr.
- [uSKG14] Stor IT Back: Annette Bornemann und Stephan Kranz GbR. Einführung in Storage Area Network V1.11. Internetquelle, 2014. URL: <http://www.storitback.de/service/san.html>, Zugriff am 12.07.2014 um 14:54 Uhr.
- [Wik13a] Wikipedia. Fibre Channel. Internetquelle, 2013. URL: http://de.wikipedia.org/wiki/Fibre_Channel_over_Ethernet, Zugriff am 18.07.2014 um 14:42 Uhr.
- [Wik13b] Wikipedia. iSCSI. Internetquelle, 2013. URL: <http://de.wikipedia.org/wiki/ISCSI>, Zugriff am 12.07.2014 um 16:44 Uhr.
- [Wik13c] Wikipedia. Thin Provisioning. Internetquelle, 2013. URL: http://de.wikipedia.org/wiki/Thin_Provisioning, Zugriff am 29.07.2014 um 14:32 Uhr.
- [Wik14a] Wikipedia. Ethernet. Internetquelle, 2014. URL: <http://de.wikipedia.org/wiki/Ethernet>, Zugriff am 29.07.2014 um 10:14 Uhr.
- [Wik14b] Wikipedia. Hochverfügbarkeit. Internetquelle, 2014. URL: <http://de.wikipedia.org/wiki/Hochverf%C3%BCgbarkeit>, Zugriff am 20.07.2014 um 10:50 Uhr.
- [Wik14c] Wikipedia. Peripheral Component Interconnect. Internetquelle, 2014. URL: http://de.wikipedia.org/wiki/Peripheral_Component_Interconnect, Zugriff am 31.07.2014 um 10:21 Uhr.
- [Wik14d] Wikipedia. VMware ESX. Internetquelle, 2014. URL: http://en.wikipedia.org/wiki/VMware_ESX, Zugriff am 25.07.2014 um 15:28 Uhr.
- [Wik14e] Wikipedia. VMware VMFS. Internetquelle, 2014. URL: http://en.wikipedia.org/wiki/VMware_VMFS, Zugriff am 29.07.2014 um 14:51 Uhr.

Anhangverzeichnis

Anhang 1: OSI-Modell	50
Anhang 2: 8b/10b-Kodierung	52
Anhang 3: Projekt-Angebot	55

Anhang 1

OSI-Modell

OSI-Modell

Das OSI-Modell - korrekt bezeichnet als OSI-Basis-Referenzmodell wurde im Jahr 1993 von der ISO standardisiert und teilt die Kommunikation zwischen Computern in mehrere Schichten ein. Es gibt sieben Schichten, angefangen von der Hardware (1) bis zur Anwendung (7), diese sind in der folgenden Tabelle aufgelistet. Für jede Schicht sind die deutsche und englische Bezeichnung, die Aufgabe und ein Protokoll-Beispiel angegeben. [Nin13]

Schicht	Bezeichnung	Aufgabe und Beispiel
7	Anwendungsschicht Application Layer	Funktionsbereitstellung für Anwendungen HTTP
6	Darstellungsschicht Presentation Layer	Systemabhängige Datenformate vereinheitlichen ASCII
5	Sitzungsschicht Session Layer	Steuerung von Verbindungen RPC
4	Transportschicht Transport Layer	Netzunabhängige Flusskontrolle und Segmentierung TCP
3	Vermittlungsschicht Network Layer	Vermittlung von Datenpaketen IP
2	Sicherungsschicht Data Link Layer	Gewährleisten einer fehlerfreien Übertragung Ethernet
1	Bitübertragungsschicht Physical Layer	Festlegung von Übertragungsparametern RS232

Anhang 2

8b/10b-Kodierung

8b/10b-Kodierung

Bei der Übertragung von Daten werden binäre Werte (0, 1) in einem positiven oder negativen elektrischen Pegel ausgedrückt. Folgen in einem Datenstrom zu viele Einsen oder Nullen aufeinander, weicht die Spannungskurve zu stark von einer Wechselstromkurve ab - dadurch ist für den Empfänger schwerer erkennbar, wie viele Takte lang das gleiche binäre Signal übertragen wird. Um diesem Problem entgegen zu wirken, wurde der 8b/10b-Code entwickelt.

Es werden jeweils 8 Bit in 10 Bit kodiert. Ziel ist es, die Wahrscheinlichkeiten des Auftretens von Einsen und Nullen auszugleichen, damit der Empfänger eine eindeutige Wechselstromkurve erhält, die eindeutig interpretierbar ist.

Es gelten folgende Festlegungen:

- Das Eingangssignal von 8 Bit kann 256 verschiedene Zustände haben (2^8).
- Das Ausgangssignal von 10 Bit kann 1024 verschiedene Zustände haben (2^{10}).
- Es gibt Symbole mit folgenden Gleichgewichtsverhältnissen:
 - 252x neutral: 1 und 0 in gleicher Anzahl (jeweils 5).
 - 210x positiv: zwei Einsen mehr als Nullen.
 - 210x negativ: zwei Nullen mehr als Einsen.
- Zusätzlich gibt es 10 frei verwendbare Codewörter, die z.B. bei Fibre Channel als Steuerbefehle genutzt werden.
- In der Übertragung folgen maximal fünf gleiche Bits aufeinander.

Die eigentliche Kodierung ist in eine 3b/4b-Kodierung (die ersten drei Bits) und eine 5b/6b-Kodierung (Bits 4-8) unterteilt. Für die Umwandlung gibt es einfache Ersetzungstabellen, die auf den Datenstrom angewendet werden.

Als Beispiel wird die Zeichenkette „abc“ übertragen. Binäre Darstellung, ein Zeichen pro Zeile mit Unterteilung für beide Kodierungen:

```
011 00001
011 00010
011 00011
```

Der erste Teil für die 3b/4b-Kodierung ist bei allen Zeichen gleich und wird laut Tabelle mit 0011 ersetzt:

0011 00001

0011 00010

0011 00011

Für die 5b/6b-Kodierung sind verschiedene Werte gegeben, diesen wurden im folgenden Block ebenfalls mittels der Referenztabelle ersetzt:

0011 011101

0011 101101

0011 110001

Zum Vergleich sind hier noch einmal die Eingangs- und Ausgangszeichenfolge:

01100001 01100010 01100011

0011011101 0011101101 0011110001

Da nicht alle 1024 zur Verfügung stehenden Codewörter belegt sind, ist auch bedingt eine Bitfehlererkennung möglich.

Die Ersetzungstabellen sind unter [LAT12, S. 3] zu finden. Quelle für Funktionsbeschreibung des Verfahrens ist [Pet08].

Anhang 3

Projekt-Angebot

ASTA-PICCA IT-Systeme GmbH

Computersysteme - Software
Kommunikationssysteme
Consulting



Seit 1990 mit Kompetenz an Ihrer Seite

ASTA-PICCA IT-Systeme GmbH - Bahnhofstraße 5 - 04668 Grimma

[REDACTED]
[REDACTED]
[REDACTED]

Angebot	
Belegnummer	2014-30300
Datum	21.01.2014
Kundennummer	D92553
Bearbeiter	Mathias Schubert

Sehr geehrter Herr [REDACTED],

entsprechend unserer ausführlichen Gespräche über die Erneuerung Ihrer Datenspeicherung unterbreiten wir Ihnen folgendes Angebot:

Pos.	Artikelnr.	Bezeichnung	Menge	ME	Einzelpreis	Gesamtpreis	SC
1	14100060	Server Fujitsu PRIMERGY RX300 S8 Fujitsu-Primergy Basiskonfiguration - 2x 8C Xeon E5-2650v2 (2,60GHz/20MB) - DVD-RW-Laufwerk - 5x PCIe x8 Ip, 2x PCIe x16 Ip - 4+1 redundante Lüfter - 2x Netzteil 450W platin - inkl. Rackmountkit und Kabelarm - 3 Jahre Vor-Ort-Service Zusatzkonfiguration: 12x 8 GB DDR3 1600 R ECC 1Rx4 L 1x Eth. Ctrl 2x10Gbit PCIe x8 D2755 SFP+ 2x SFP+ Module Multi Mode Fiber 10GbE LC 1x iRMC S4 advanced pack 1x VMware ESXi Embedded UFM Device 2x VMware vSphere 5.5 Standard 1x PRIMERGY Einbauservice E1	4	Stck	8.460,00	33.840,00	1
2	14100090	Fujitsu ETERNUS DX100 S3 Storage Fujitsu Eternus Basiskonfiguration - 3 Jahre Vor-Ort-Service Zusatzkonfiguration: 2x DX100 S3 Cache4G, 1x4GB for 1 Ctl. 5x DX1/200 S3 HD 2.5" 1.0TB 7.2krpm x1 12x DX1/200 S3 HD 2.5" 900GB 10krpm x1 1x ETERNUS Snapshot Manager 1x 3 Jahre Vor-Ort-Support 2x DX100 S3 CM w 1xCA iSCSI 10G 2port 1x TP 3J VO Svc,NBD Whz,9x5	2	Stk	39.630,00	79.260,00	1
3	54000557	Switch D-Link DXS-3600-16S Inklusive: 8x DEM-431XT (SFP+-Modul)	4	Stk	9.880,00	39.520,00	1

Übertrag	152.620,00
----------	------------

Geschäftsführer:
Dr. Bernhard Weigel
HRB Leipzig 2013

Bankverbindungen:
Raiffeisenbank Grimma:
IBAN: DE27 8606 5483 0308 0006 72, BIC: GENO DE F1 GMR
Sparkasse Grimma:
IBAN: DE10 8605 0200 1010 0280 29, BIC: SOLA DE S1 GRM

Kontakt: Bahnhofstr. 5, 04668 Grimma
Tel: 03437 7177 0
Fax: 03437 7177 77
e-Mail: info@asta-picca.de
www.asta-picca.de

Angebot 2014-30300 Seite 2 von 2

Pos.	Artikelnr.	Bezeichnung	Menge	ME	Einzelpreis	Gesamtpreis	SC
4	14100068	USV EATON EX 3000 RT2HE Netpack Wechselstrom 230 V - 1.8 kW Stromversorgungsgerät UPS-Technologie: Online Eingangsspannung: Wechselstrom 100/120/160/184-284 V Nötige Frequenz: 40 - 70 Hz Eingabeanschluss/-anschlüsse: 1 x Netzstecker IEC 320 EN 60320 C20 Angaben zu Ausgangsleistungsanschlüssen: 8 x Netzstecker IEC 320 EN 60320 C13 1 x Netzstecker IEC 320 EN 60320 C19 Ausgangsspannung: Wechselstrom 200/208/220/230/240 V (50/60 Hz) Leistungskapazität: 2.7 kW / 3000 VA	4	Stk	1.704,00	6.816,00	1
5	04000028	Verkabelungspauschale	1	Psch	200,00	200,00	1
6	99900037	Arbeitsstunde Ingenieur (HW) IT-Netzwerkingenieur/Systemssoftware - Vorinstallation Virtualisierungsserver - Aufbau und Konfiguration der Komponenten - Einrichtung der Storages - Datenübernahme	32,00	Std.	80,00	2.560,00	1

Zwischensumme					EUR	162.196,00	SC
zzgl. MwSt. mit Steuercode			1	19,00 % von	162.196,00	30.817,24	
Endsumme					EUR	193.013,24	

Alle Einzelpreise verstehen sich zzgl. 19% Mehrwertsteuer.

Unser Angebot gilt 20 Tage.

Für gelieferte Ware gilt erweiterter Eigentumsvorbehalt bis zur kompletten

Bezahlung durch den Kunden. Des Weiteren verweisen wir auf unsere allgemeinen Geschäftsbedingungen, die Sie unserer Internetseite (www.asta-picca.de) entnehmen können.

Wir würden uns freuen, wenn Ihnen unser Angebot zusagt und verbleiben
mit freundlichen Grüßen

ASTA-PICCA IT-Systeme GmbH

Dipl.-Ing. (BA)
Mathias Schubert
Netzwerkingenieur

bestätigt:
(Unterschrift/Stempel)

Zahlungsvereinbarungen:

14 Tage

ohne Abzug

193.013,24 EUR

Ehrenwörtliche Erklärung

„Ich erkläre hiermit ehrenwörtlich“,

1. dass ich meine Bachelorthesis mit dem Thema „Konzeptionierung und Realisierung der Erneuerung des Storage Area Network (SAN) eines mittelständischen Unternehmens zur Steigerung der Ausfallsicherheit der IT-Infrastruktur“ ohne fremde Hilfe angefertigt habe,
2. dass ich die Übernahme wörtlicher Zitate aus der Literatur sowie die Verwendung der Gedanken anderer Autoren an den entsprechenden Stellen innerhalb der Arbeit gekennzeichnet habe und
3. dass ich meine Bachelorthesis bei keiner anderen Prüfung vorgelegt habe.

Ich bin mir bewusst, dass eine falsche Erklärung rechtliche Folgen haben wird.

Ort, Datum

Unterschrift

Thesen

These 1:

Die redundante Auslegung aktiver Netzwerkkomponenten erhöht die Ausfallsicherheit einer Infrastruktur.

These 2:

Die iSCSI-Technologie ist für den Aufbau eines Storage Area Network preisgünstiger als Fibre Channel.

These 3:

Die Datenhaltung in mehreren Brandabschnitten ist ein wichtiges Kriterium für den Schutz des Datenbestandes.

These 4:

Die Virtualisierung von Betriebssystemen bietet mehr Flexibilität als die Installation auf physikalischer Hardware.

These 5:

Speichernetze bieten bessere Ansätze zur nachträglichen Kapazitätserweiterung als traditionelle Technologien zur Datenspeicherung.