

ProSeminar „Speicher- und Dateisysteme“

Uni HH - WS2010/11

Schriftliche Ausarbeitung

Netzwerkpeichersysteme

Mirko Köster

31.03.2011

Inhalt

1	Einleitung / Motivation	2
1.1	Historie	2
1.2	Einsatzgebiete	3
2	File Server.....	4
2.1	Beschreibung / Charakteristika.....	4
2.2	Topologie.....	5
2.3	Eigenschaften	5
3	NAS (Network Attached Storage).....	6
3.1	Beschreibung / Charakteristika.....	6
3.2	Topologie.....	6
3.3	Eigenschaften	7
4	SAN (Storage Area Network)	8
4.1	Beschreibung / Charakteristika.....	8
4.2	Topologie.....	9
4.3	Eigenschaften	9
5	Cloud.....	11
5.1	Beschreibung / Charakteristika.....	11
5.2	Topologie.....	11
5.3	Eigenschaften	12
6	Vergleich / Fazit.....	13
7	Quellen.....	15

1 Einleitung / Motivation

Vereinfacht gesagt, braucht man Netzwerkspeichersysteme immer dann, wenn man

- Daten mehreren Benutzern / Systemen zur Verfügung stellen will

oder

- den Speicher zentral verwalten will

Die Menge an Daten, die weltweit gespeichert wird, wächst von Jahr zu Jahr exponentiell. Das liegt u.a. an gesetzlichen Vorgaben sowie der elektronischer Kommunikation, aber ganz allgemein auch daran, dass für immer mehr Aufgaben IT genutzt wird.

Dieses Wachstum stellt besondere Anforderungen an Speichersysteme. Bei der Wahl des zur Aufgabe passenden Netzwerkspeichersystems spielen folgende Faktoren eine wichtige Rolle, auf die ich in dieser Ausarbeitung jeweils eingehen werde:

- Preis
- Geschwindigkeit
- Wartbarkeit
- Ausfallsicherheit
- Backup / Restore
- Verwaltung

Der Zugriff auf die Daten eines Netzwerkspeichersystems erfolgt, wie der Name schon sagt, über ein Netzwerk sowie entweder Datei- oder Block-basiert.

Im Detail werde ich auf folgende Netzwerkspeichersysteme eingehen: File Server, NAS (Network Attached Storage), SAN (Storage Area Network) sowie die Cloud.

1.1 Historie

An dieser Stelle will ich keine Historie der Netzwerkspeichersysteme (im eigentlichen Sinne) präsentieren. Vielmehr stelle ich den groben Entwicklungsverlauf dar:

File Server → NAS → SAN (→ Cloud)

Ich habe mich dazu entschieden, da es keine klare lineare Entwicklung gab. Alle genannten Systeme existieren heute parallel. Jedes System hat ein Einsatzgebiet, für das es auch heute noch geeignet ist – und jedes der genannten Netzwerkspeichersysteme wird parallel weiterentwickelt.

1.2 Einsatzgebiete

Der File Server wird häufig in KMU (kleine und mittlere Unternehmen) eingesetzt.

Das NAS kommt überwiegend im SOHO-Bereich (Small Office, Home Office) sowie im privaten Umfeld zum Einsatz.

Ein SAN können sich häufig nur größere Unternehmen oder Rechenzentren leisten; und entsprechend spielt es seine Vorteile auch häufig erst in größeren IT-Umgebungen aus.

Die Cloud hingegen ist noch ein relativ frischer Begriff und daher auch noch nicht klar umrissen. Denkbar für den Einsatz sind aber Szenarien sowohl im privaten als auch im professionellen Bereich. Doch dazu mehr im entsprechenden Kapitel.

2 File Server

2.1 Beschreibung / Charakteristika

Ein Fileserver besteht aus der Serverhardware sowie (in der Regel) einem Serverbetriebssystem und ist an ein Netzwerk angebunden.

Als Serverbetriebssystem kommen häufig vollwertige Betriebssysteme wie Linux / Unix oder Windows Server zum Einsatz. Das wiederum bedeutet, dass der File Server genauso verwaltet werden muss wie alle anderen Server im Unternehmen auch.

Da es sich um einen herkömmlichen Server handelt, ist der Platz- und Energiebedarf entsprechend groß.

Klassische File Server greifen direkt auf ihren Massenspeicher zu (eingebaut / DAS). Es ist aber auch möglich, diesen per SAN zur Verfügung zu stellen.

Die Geschwindigkeit, mit der der File Server seine Daten über das Netzwerk ausliefern kann, hängt von mehreren Faktoren ab:

Zum einen spielt die Geschwindigkeit, mit der der Server an das Netzwerk angeschlossen ist, eine zentrale Rolle. Gigabit-Ethernet sollte es mindestens sein, da dies bei moderner Hardware sonst der limitierende Faktor wäre. Auch sollte auf die Auslastung des Netzwerks geachtet werden.

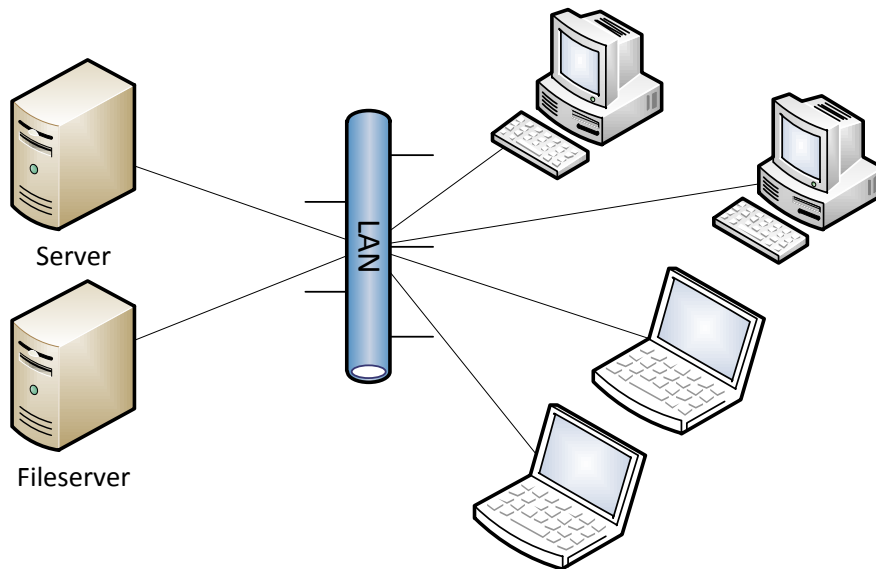
Weiterhin ist die Einzelperformance der Festplatten entscheidend. Meistens werden mehrere davon zu einem RAID-Verbund zusammengefasst, um die Ausfallsicherheit und/oder die Geschwindigkeit zu erhöhen (von RAID0 wird im professionellem Umfeld abgeraten).

Nicht außer Acht gelassen werden sollte die Frage, ob der Server ein dedizierter File Server ist oder noch andere Aufgaben hat. Denn ein File Server, der zu sehr ausgelastet ist, kann entsprechend die Daten auch nur mit vermindertem Tempo ausliefern.

Als Protokoll kommt meist entweder NFS (Unix / Linux) oder SMB/CIFS (Windows) zum Einsatz. Der Zugriff erfolgt somit Datei-basiert. Für die Benutzer hat dies den Vorteil, dass sie auf den ersten Blick keinen Unterschied zu lokalen Dateisystemen sehen.

Werden in heterogenen Umgebungen beide Protokolle benötigt, bedeutet dies einen deutlich erhöhten Verwaltungsaufwand (Zugriffsverwaltung / Rechtemanagement). Es ist aber durchaus möglich, da beide Protokolle sowohl unter Unix / Linux (SMB/CIFS per Samba) als auch Windows (NFS per Dienste für NFS (Network File System)) zur Verfügung stehen.

2.2 Topologie



2.3 Eigenschaften

- Speicherkapazität
 - Mehrere TB möglich
- Speicherdichte
 - Abhängig von eingesetzten Festplatten
- Haltbarkeit
 - Einige Jahre
- Datensicherheit
 - Abhängig von
 - eingesetztem RAID-Level
 - Backup- und Recovery-Strategie
- Performanz
 - Abhängig von
 - eingesetztem RAID-Level
 - Einzelperformanz und Anzahl der Festplatten
 - ca. 100 MB/s (evtl. mehr bei 10Gbit/s-Ethernet)
- Energieverbrauch
 - Mehrere 100W
- Anschaffungskosten
 - Ab ca. 2.000 EUR aufwärts

3 NAS (Network Attached Storage)

3.1 Beschreibung / Charakteristika

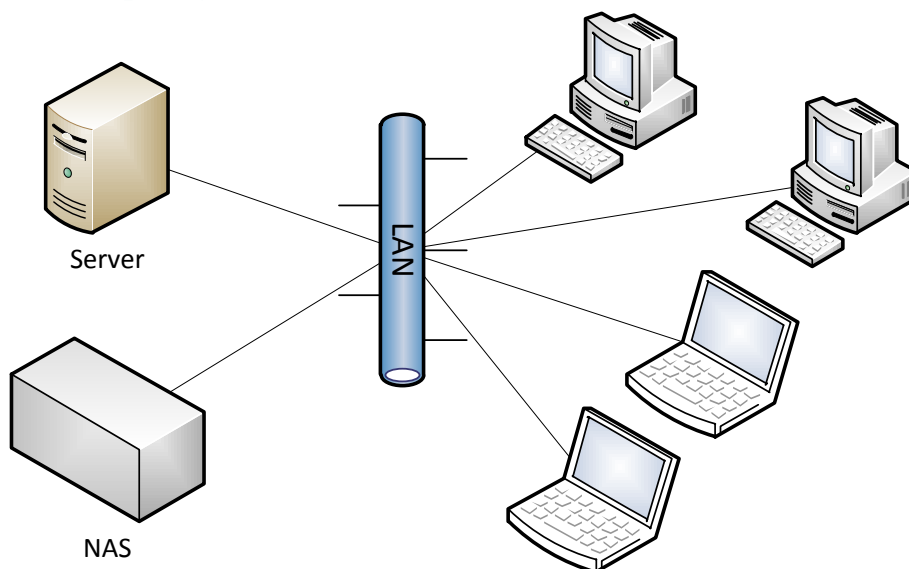
Man kann ein NAS als einen auf das Wesentliche reduzierter File Server betrachten. Im Vordergrund stehen ein geringer Energieverbrauch sowie eine einfache Verwaltung.

Dementsprechend kommt keine Serverhardware zum Einsatz. Oft besteht ein NAS aus einem relativ kleinem Gehäuse, das Platz für ein bis vier Festplatten bietet. Zur Zeit sind NAS auf Intel Atom- oder ARM-Basis verbreitet. Diese Prozessoren verbrauchen im Betrieb wenig Strom bei gleichzeitig hoher Energie-Effizienz (Leistung pro Watt). Die Kehrseite ist, dass diese bei hoher Last schnell überfordert sind.

Als Betriebssystem kommt meist ein angepasstes Linux zum Einsatz, welches von allen nicht benötigten Softwarekomponenten „befreit“ wurde. Zusätzlich bieten diese Systeme eine spezielle Verwaltungsoberfläche in Form eines Web-interfaces. Die Verwaltung ist oft relativ simpel gehalten, so dass sich der Einsatz eines NAS gerade für Privatpersonen oder kleine Büros lohnt.

Trotz der relativ einfachen Verwaltung bieten die meisten NAS als verfügbare Protokolle meist NFS, SMB/CIFS, FTP, HTTP (WebDAV) sowie einen Zugang per SSH. Somit ergibt sich ein klarer Vorteil gegenüber den voll verwalteten File Servern.

3.2 Topologie



3.3 Eigenschaften

- Speicherkapazität
 - Im TB-Bereich
- Speicherdichte
 - Abhängig von eingesetzten Festplatten
- Haltbarkeit
 - Einige Jahre
- Datensicherheit
 - Abhängig von
 - eingesetztem RAID-Level
 - Backup- und Recovery-Strategie
- Performanz
 - Abhängig von
 - eingesetztem RAID-Level
 - Einzelperformanz und Anzahl der Festplatten
 - CPU-Performance
 - ca. 10 bis 100 MB/s
- Energieverbrauch
 - ca. 5 bis 50W
- Anschaffungskosten
 - ca. 100 EUR bis 1.500 EUR

4 SAN (Storage Area Network)

4.1 Beschreibung / Charakteristika

Ein SAN unterscheidet sich grundlegend von den bisher vorgestellten Netzwerkspeichersystemen File Server und NAS. Es bildet ein eigenes Netzwerk, welches getrennt ist vom eigentlichen LAN, mit der Aufgabe, Speicher (und andere Geräte wie Tape-Libraries für Backup/Restore) zu Verwalten und zur Verfügung zu stellen.

SANs bieten einige handfeste Vorteile. Sie sind sehr schnell. Hochwertige SANs sind per Glasfaser (Fibre Channel) verbunden; aber auch herkömmliche IP-Netze können genutzt werden (iSCSI). Durch die Trennung vom eigentlichen LAN wird dieses geschont.

Sie werden zentral verwaltet und schöpfen ihre hohe Ausfallsicherheit / Hochverfügbarkeit aus einem hohen Maß an Redundanz. Es ist möglich, jede einzelne Komponente mehrfach vorzuhalten. Im Fehlerfall übernimmt dann einfach eine andere Komponente.

Zu den Komponenten eines SANs gehören die HBAs (Host Bus Adapter; Funktion vergleichbar mit herkömmlichen Netzwerkadaptern), SAN-Switches und Disk-Arrays (enthalten die Festplatten) sowie die Datenleitungen. Aber auch weitere Komponenten können wie eingangs erwähnt in ein SAN integriert werden.

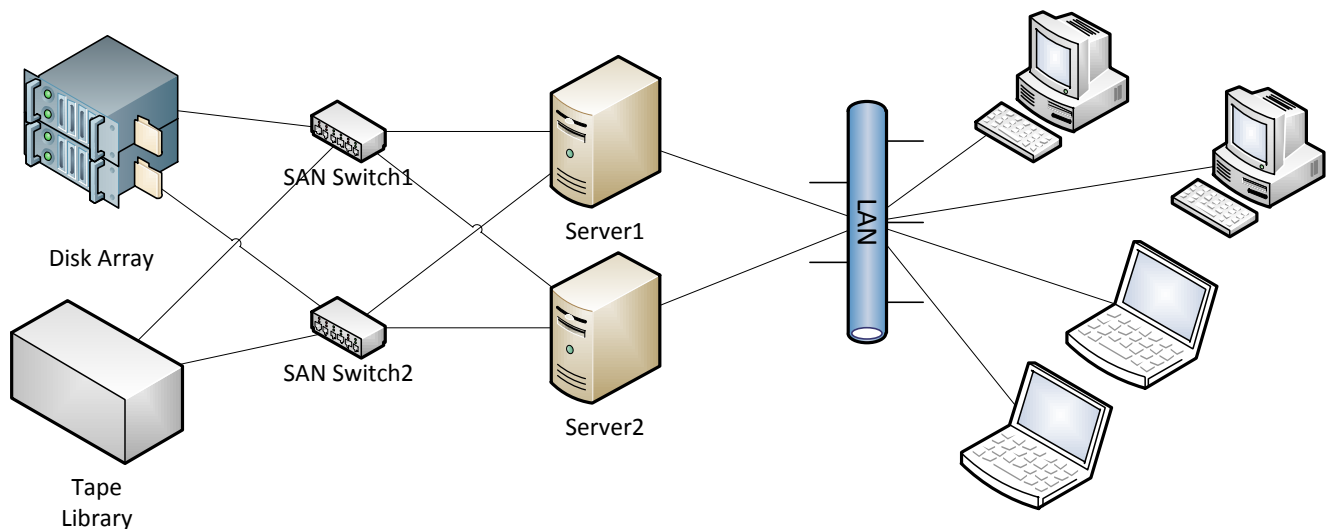
Die Disk-Arrays unterstützen verschiedene RAID-Level. Somit können auch die Festplatten als letzte Komponente redundant ausgelegt sein.

Durch diese hohe Redundanz übersteht ein SAN den Ausfall jeder (einzelnen) Komponente. Dank Multi-Pathing wird der Ausfall eines HBA / einer Leitung / eines Switches verkraftet. Ist auch das Disk-Array redundant ausgelegt, führt auch dieser Ausfall nicht zu Problemen; gleiches gilt für Festplatten (s.o.).

Zu guter Letzt ist es möglich, das ganze SAN redundant auszulegen und z.B. in zwei getrennten Gebäuden zu betreiben, da mit Hilfe von Glasfaserleitungen große Entfernungen überbrückt werden können.

Den Servern, die an das SAN angeschlossen sind, werden virtuelle Festplatten (auch LUNs genannt) zugeordnet. Der Zugriff auf diese erfolgt Block-basiert. Aus Sicht des Betriebssystems verhalten sich diese wie lokale Festplatten.

4.2 Topologie



4.3 Eigenschaften

- Speicherkapazität

ca. 22 TByte / 48 TByte (SAS/SATA)

Beliebig erweiterbar durch weitere Disk Arrays

- Speicherdichte

Abhängig von eingesetzten Festplatten

z. B. 12 Festplatten / 2 HE

- Haltbarkeit

Aufgrund der Redundanz: abhängig vom Supportzeitraum des Herstellers

- Datensicherheit

Abhängig von

- Grad der Redundanz
- eingesetztem RAID-Level
- Backup- und Recovery-Strategie

- Performanz

Abhängig von

- eingesetztem RAID-Level
- Einzelperformanz der Festplatten
- Geschwindigkeit der Anbindung (Glasfaser / Kupferleitung)

ca. 100 MB/s bis max. 1,6 GB/s

- Energieverbrauch

Abhängig von der Konfiguration (mehrere 100 W aufwärts)

- Anschaffungskosten

Ab ca. 10.000 EUR aufwärts

5 Cloud

5.1 Beschreibung / Charakteristika

Die Begriffe Cloud und Cloudcomputing sind noch relativ jung – und zugleich ziemlich abstrakt.

Dabei ist das Thema zur Zeit in aller Munde. So lautete das Top-Thema der CeBIT 2011 „Work and Life with the Cloud“. Dabei bleibt die Frage, ob es sich dabei um einen Hype oder um einen echten Trend für die nächsten Jahre handelt.

Grob gesagt, versteht man unter dem Begriff Cloud eine Reihe von Dienstleistungen im Internet, bei denen die Implementationsdetails vom Kunden verborgen bleiben (Transparenz). Beispiele sind SaaS (Software as a Service), zu mietende Rechenkapazität (Amazon Elastic Compute Cloud (EC2)) oder Speicherkapazität (Amazon Simple Storage Service (S3) oder Dropbox).

Der größte Vorteil der Cloud ist, dass man keine Anschaffungskosten hat und in der Regel kein Personal zum Verwalten benötigt.

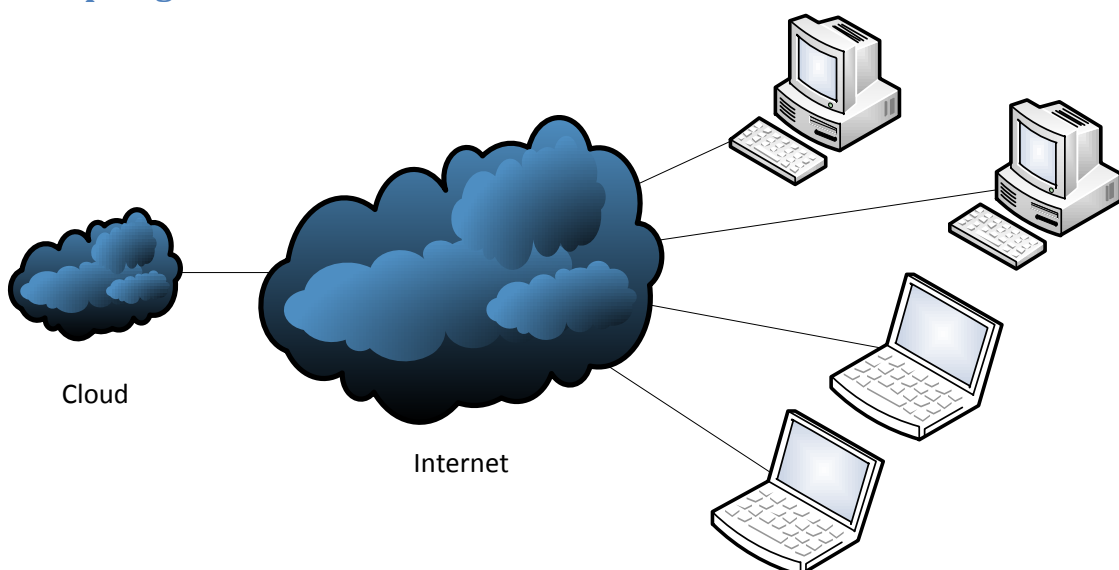
Die Verwaltung obliegt dem Anbieter der Dienstleistung; der Kunde greift über das Internet transparent darauf zu. Im Folgenden werde ich das Anhand des Beispiels Dropbox aufzeigen.

Der Zugriff auf die Daten erfolgt transparent entweder per Webinterface, per Client (welcher die Daten dann im Hintergrund mit der Cloud synchronisiert und unter Windows im Explorer in den „Eigenen Dokumenten“ zur Verfügung stellt) auf dem PC oder per App auf dem Smartphone.

Dabei interessiert den Nutzer nicht, wie und wo die Daten gespeichert sind. So ist es möglich, dass die Daten rund um den Globus in Rechenzentren (evtl. redundant) verteilt sind.

Dem Benutzer wird ein Speicherkontingent zur Verfügung gestellt. In der kostenlosen Variante sind das 2 GB – wer mehr benötigt, muss einen anderen Tarif wählen, der dann entsprechend bezahlt werden muss.

5.2 Topologie

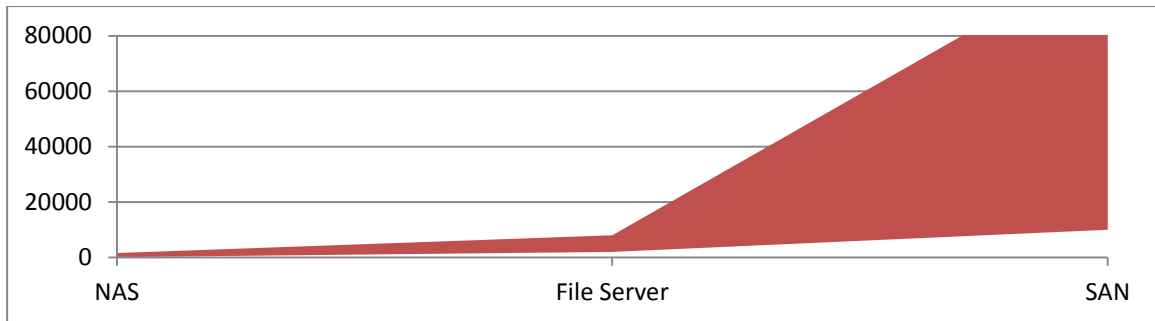


5.3 Eigenschaften

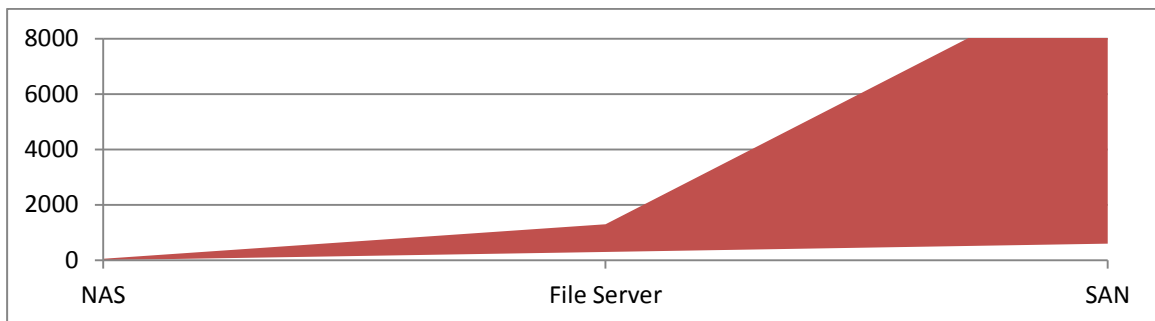
- Speicherkapazität
 - Abhängig vom Tarif
- Speicherdichte, Haltbarkeit und Energieverbrauch
 - Bleibt transparent
- Datensicherheit
 - Muss vertraglich mit dem Anbieter vereinbart werden
- Performanz
 - Meist ist die Internetanbindung der limitierende Faktor
 - Oder der Anbieter drosselt die Geschwindigkeit künstlich (je nach gewähltem Tarif)
- Anschaffungskosten
 - Praktisch keine (evtl. Einrichtungsgebühr)

6 Vergleich / Fazit

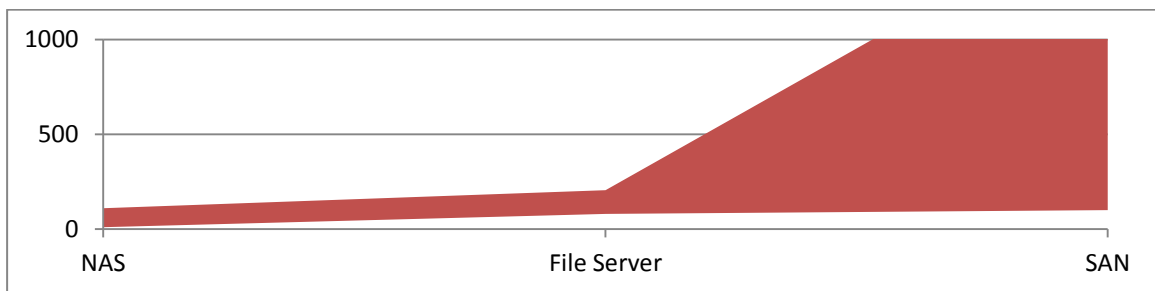
Die Werte in den folgenden Diagrammen sind teilweise geschätzt und hängen meist von der entsprechenden Hardware ab:



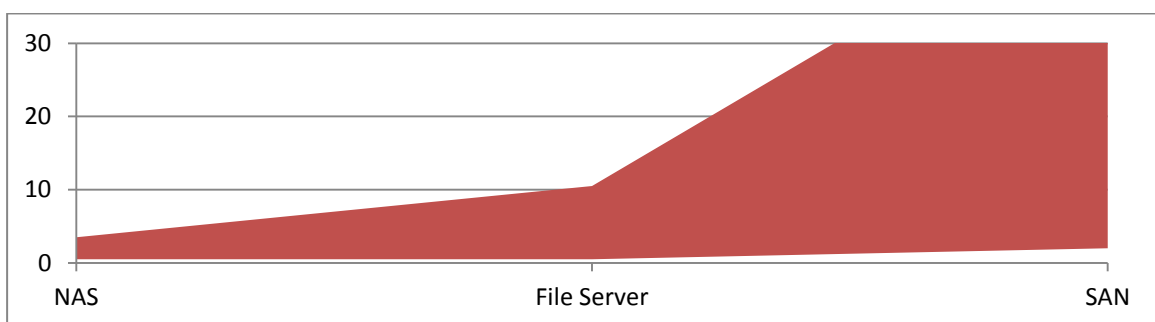
Anschaffungskosten in EUR



Energieverbrauch in W



Performance in MB/s



Speicherkapazität in TB

Abschließend will ich festhalten, dass alle hier vorgestellten Netzwerkspeichersysteme spezifische Vor- und Nachteile haben. Es gibt nicht die eine Lösung für alle Anwendungsfälle.

Alle haben ihre Daseinsberechtigung.

Man muss in jedem Einzelfall abwägen, welches das angebrachteste Produkt ist. So wäre es „mit Kanonen auf Spatzen geschossen“, wenn sich ein Büro mit zwei Arbeitsplatzrechnern ein SAN leisten würde. Andersrum ist für alle größeren Unternehmen ein NAS klar unterdimensioniert.

In den meisten Fällen kann man sich wohl an den im Kapitel Einsatzgebiete beschriebenen Szenarios orientieren, ohne viel verkehrt zu machen.

Erwähnenswert ist noch, dass bei allen Produkten die Preisskala noch oben fast beliebig offen ist (insbesondere beim SAN; wie auf der vorherigen Seite angedeutet).

Wichtige Faktoren, die man bei einer Neuanschaffung berücksichtigen sollte, sind:

- Preis
- Leistung
- Wartungsaufwand
- Verfügbarkeit / Redundanz

Aber auch die Cloud wird wohl immer mehr an Bedeutung gewinnen. Gerade für junge Unternehmen, die noch nicht viel investieren können, scheint diese Lösung in Zeiten, in denen Breitbandanschlüsse fast flächendeckend und günstig verfügbar sind, interessant zu sein.

Die Vorteile sind nicht von der Hand zu weisen. Keine Anfangsinvestitionen, geringe und kalkulierbare laufende Kosten sowie kein bzw. nur ein geringer Wartungsaufwand.

7 Quellen

- Wikipedia
 - http://de.wikipedia.org/wiki/Storage_Area_Network
 - http://de.wikipedia.org/wiki/Network_Attached_Storage
- c't 4/2011 – Netzwerkspeicher
- SAN kurz erklärt
 - <http://www.searchstorage.de/index.cfm?pid=3636&pk=46906>
- Einführung in Storage Area Network (SAN)
 - <http://www.storitback.de/index.html?/service/san.html>