

Storage

D. Leeuw

28 februari 2021

v.0.1.1



© 2021 Dennis Leeuw

Dit werk is uitgegeven onder de Creative Commons BY-NC-SA Licentie en laat anderen toe het werk te kopiëren, distribueren, vertonen, op te voeren, en om afgeleid materiaal te maken, zolang de auteurs en uitgever worden vermeld als maker van het werk, het werk niet commercieel gebruikt wordt en afgeleide werken onder identieke voorwaarden worden verspreid.

Over dit Document

Dit document behandelt de opslag van data op de verschillende opslagsystemen voor het middelbaar beroepsonderwijs in Nederland.

Versienummering

Het versienummer van elk document bestaat uit drie nummers gescheiden door een punt. Het eerste nummer is het major-versie nummer, het tweede nummer het minor-versienummer en de laatste is de nummering voor bugfixes.

Om met de laatste te beginnen als er in het document slechts verbeteringen zijn aangebracht die te maken hebben met type-fouten, websites die niet meer beschikbaar zijn, of kleine foutjes in de opdrachten dan zal dit nummer opgehoogd worden. Als docent of student hoeft je je boek niet te vervangen. Het is wel handig om de wijzigingen bij te houden.

Als er flink is geschreven aan het document dan zal het minor-nummer opgehoogd worden, dit betekent dat er bijvoorbeeld plaatjes zijn vervangen of geplaatst/weggehaald, maar ook dat paragrafen zijn herschreven, verwijderd of toegevoegd, zonder dat de daadwerkelijk context is veranderd. Een nieuw cohort wordt aangeraden om met deze nieuwe versie te beginnen, bestaande cohorten kunnen doorwerken met het boek dat ze al hebben.

Als het major-nummer wijzigt dan betekent dat dat de inhoud van het boek substantieel is gewijzigd om bijvoorbeeld te voldoen aan een nieuw kwalificatiedossier voor het onderwijs. Een nieuw major-nummer betekent bijna altijd voor het onderwijs dat men in het nieuwe schooljaar met deze nieuwe versie aan de slag zou moeten gaan. Voorgaande versies van het document zullen nog tot het einde een schooljaar onderhouden worden, maar daarna niet meer.

Document ontwikkeling

Het doel is door middel van open documentatie een document aan te bieden aan zowel studenten als docenten, zonder dat hier hoge kosten aan verbonden zijn en met de gedachte dat we samen meer weten dan alleen. Door samen te werken kunnen we meer bereiken.

Bijdragen aan dit document worden dan ook met alle liefde ontvangen. Let u er wel op dat materiaal dat u bijdraagt onder de CC BY-NC-SA licentie vrijgegeven mag worden, dus alleen origineel materiaal of materiaal dat al vrijgegeven is onder deze licentie.

De eerste versie is geschreven voor het ROC Horizon College.

Versienummer	Auteurs	Verspreiding	Wijzigingen
0.1.0	Dennis Leeuw	Initieel document	

Tabel 1: Document wijzigingen

Inhoudsopgave

Over dit Document	i
1 Inleiding	1
2 Storage devices	3
2.1 Block devices	3
2.1.1 Harddisks	3
2.1.2 SSD - Solid State Drive	3
2.1.3 USB-sticks	3
2.1.4 JOBD - Just a Bunch Of Disks	3
2.2 interfaces	3
2.2.1 IDE - Integrated Drive Electronics	3
2.2.2 SATA - Serial ATA	3
2.2.3 SCSI - Small Computer Systems Interface	3
2.2.4 SAS - Serial Attached SCSI	5
2.2.5 USB - Universal Serial Bus	5
3 Bestandssystemen	7
3.1 FAT - File Allocation Table	7
3.2 Journaling filesystem	7
4 Backup	9
4.1 LTO - Linear Tape Open	9
5 RAID	11
5.1 RAID0	11
5.2 RAID1	11
5.3 RAID5	11
6 DAS - Direct Attached Storage	13

7	SAN - Storage Attached Network	15
7.1	iSCSI	15
7.2	Fibre Channel	15
8	Network Fileserver	17
8.1	NFS - Network File System	17
8.1.1	Security	18
8.1.2	Interoperability	18
8.2	SMB - Server Message Block	18
8.2.1	Security	19
8.2.2	Interoperability	19
8.2.3	Roaming profiles	19
8.3	NAS - Network Attached Storage	19
9	Object Storage	21
10	Web storage	23
10.1	webdav	23
10.2	Amazon S3 bucket	24
10.3	Google Drive	24
10.4	One Drive & Microsoft Azure BLOB	24

Hoofdstuk 1

Inleiding

Onder Storage verstaan we het aanbieden van opslagcapaciteit op opslag-systemen. Dit kan op harddisks in het systeem zijn op een fileserver die gekoppeld is aan het netwerk.

Door een bestandssysteem over het netwerk aan te bieden kan elk systeem dat het netwerk-protocol spreekt gebruik maken van de data, waarmee het delen van bestanden tussen verschillende systemen overbrugt wordt.

Met de opkomst van het Internet en speciaal het web (HTTP) zijn we data (documenten, plaatjes en filmpjes) met elkaar gaan delen, bijna elk apparaat spreekt tegen woorden HTTP of HTTPs en HTML (de opmaak taal van webdocumenten) en daarmee is een belangrijkdeel van de documenten ontsloten voor de hele wereld.

Dit document beschrijft de verschillende opslagsystemen en de technieken om deze systemen te koppelen en toegankelijk te maken.

Hoofdstuk 2

Storage devices

2.1 Block devices

2.1.1 Harddisks

2.1.2 SSD - Solid State Drive

2.1.3 USB-sticks

2.1.4 JOBD - Just a Bunch Of Disks

JBOD is precies wat de naam zegt Just a Bunch Of Disks, ofwel gewoon een hoeveelheid disk. De oude IDE interface in de PC gaf je de mogelijkheid om maximaal 4 harddisks aan te sluiten in je machine. Met de eerst SCSI-interfaces waren dat er 7. Als je meer wilde gebruikte je een cabinet met disken. Omdat daar geen enkele intelligentie in zit moet je elke disk apart kenbaar maken op je systeem. Vandaar de naam Just a Bunch of Disks.

2.2 interfaces

2.2.1 IDE - Integrated Drive Electronics

2.2.2 SATA - Serial ATA

2.2.3 SCSI - Small Computer Systems Interface

SCSI is een parallele bus waaraan meerdere apparaten gekoppeld kunnen worden. Hoewel het meest gebruikt voor het aansluiten van opslagsystemen gebruikte bijvoorbeeld de Apple Macintosh deze interface in de jaren '80

en '90 ook om scanners en printers aan te sturen. De interface kan zowel intern als extern gebruikt worden om devices aan te sluiten op de HBA (Host Bus Adapter). De interne kabel was meestal een ribbon-kabel en de externe kabels waren meestal rond. In tegenstelling tot IDE kent SCSI een command-language waarmee de HBA kan communiceren met de aansloten apparaten.

Op een SCSI-bus worden de apparaten aan elkaar doorgelust, een zogenaamde daisy-chain (madeliefjes-ketting).

De SCSI bus moet getermineerd worden aan beide einden van de bus. Dus het eerste en het laatste device op de bus moeten voorzien zijn van een terminator (weerstandsblok). Dit kan gedaan worden door een terminator aan te sluiten, maar de meeste apparaten hebben ook een schakelaar waarmee de terminatie ge-enabled wordt.

SCSI kent verschillende standaarden die in de loop van de tijd zijn ontwikkeld en behalve voor SCSI-1 is er voor elke variant een "normale" en een "wide" variant. De normale-variant heeft IDs tussen 0 en 7 en de wide-variant heeft IDs tussen 0 en 15. De hoogste prioriteit op de SCSI-bus is ID 7, de HBA heeft meestal dit ID. De prioriteit loopt daarna af van 6 naar 0 en van 15 naar 8.

De traagste variant van SCSI maakt gebruik van de Single Ended (SE) techniek. Deze variant kan gemixed worden met de Low Voltage Differential (LVD) techniek, maar dan neemt de performance van de gehele bus af. LVD is de meest gebruikte techniek. De High Voltage Differential kan niet gemixed worden met de overige twee technieken, maar kan wel langere afstanden overbruggen dan de overige twee.

Standard	Name	Bus width (bits)	Speed (MBps)	SE (m)	LVD (m)	HVD (m)
SCSI-1	SCSI-1	8	5	6	-	25
SCSI-2	Fast SCSI	8	10	3	-	25
	Fast-Wide SCSI	16	20	3	-	25
SCSI-3	Ultra SCSI	8	20	1.5	-	25
	Ultra wide SCSI	16	40	-	-	25
	Ultra2 SCSI	8	40	-	12	25
	Ultra2 wide SCSI	16	80	-	12	25
	Ultra3 Ultra 160 SCSI	16	160	-	12	-
	Ultra 320 SCSI	16	320	-	12	-

SCSI kent een enorme variatie in connectoren.

FIXME: Connectoren CKP

2.2.4 SAS - Serial Attached SCSI

2.2.5 USB - Universal Serial Bus

Hoofdstuk 3

Bestandssystemen

Een hardeschijf heet een block device omdat data gelezen (en geschreven) kan worden in vaste groottes, genaamd blocks, sectors of clusters. Een block is meestal 512 bytes of een veelvoud daarvan. Je kan data rechtstreeks naar een block device sturen, maar meestal wordt een block device geformatteerd zodat er een bestandssysteem wordt aangemaakt. Een bestandssysteem is in het Engels een filesystem en daarvan bestaan er verschillende. Windows gebruikt tegenwoordig NTFS, maar vroeger het FAT-filesystem. Linux kent het extended filesystem (extfs), maar ook bijvoorbeeld BTRFS en Reiserfs. Apple gebruik het Apple File System (APFS).

Je kan geen Apple APFS disk lezen op een Windows systeem en ook geen Windows NTFS disk op de Mac. Onder Linux is NTFS-3G ontwikkeld en daarmee kan Linux NTFS disks lezen en schrijven. NTFS-3G is ook beschikbaar voor Mac OS X. Er is experimentele ontwikkeling gaande om ook APFS te ondersteunen onder Linux. FAT is op bijna elk systeem lees- en schrijfbaar, daar dit systeem al heel oud is en er voor bijna elk besturingssysteem wel drivers voor zijn.

3.1 FAT - File Allocation Table

3.2 Journaling filesystem

Hoofdstuk 4

Backup

4.1 LTO - Linear Tape Open

Hoofdstuk 5

RAID

Ook RAID kan als een vorm van virtualisatie beschouwd worden. Er worden 2 of meer disks samengevoegd tot voor het OS 1 zichtbare disk. De belangrijkste RAID versies zijn RAID 0, 1 en 5.

5.1 RAID0

Met behulp van RAID0 kan je van verschillende schijven één grote disk disk maken.

5.2 RAID1

RAID1 zorgt ervoor dat data die je opslaat 2 keer opgeslagen wordt. Je hebt er dan ook twee disken voor nodig. Op elke disk wordt een kopie van je bestand opgeslagen.

5.3 RAID5

RAID5 is een compromis tussen veiligheid van de data en de maximale opslagcapaciteit die aan de gebruiker beschikbaar gesteld kan worden.

Hoofdstuk 6

DAS - Direct Attached Storage

Direct Attached Storage is een 19-behuizing met disken. Deze kast heeft meestal een SCSI of SAS aansluiting. Via software wordt het systeem als een grote disk of als verschillende partities aangeboden aan een host. De controller die de disken aanstuurt is meestal een RAID controller zodat data redundant wordt opgeslagen. De disken die de host ziet hebben geen enkel verband met de fysieke disken in de 19-behuizing.

DAS en JBOD worden vaak door elkaar gebruikt. Sommigen maken het onderscheid door JBOD de oplossing zonder RAID te noemen en DAS een oplossing met RAID. Letterlijk vertaald: Direct verbonden opslag, maakt dat elke disk die direct aan een systeem verbonden is, intern of extern, onder de titel van DAS kan vallen. Vandaar vermoedelijk de verwarring over deze termen.

Hoofdstuk 7

SAN - Storage Attached Network

Een Storage Attached network is zoals de naam al zegt een netwerk van storage applicances. Dit netwerk maakt het mogelijk de totale opslag te delen met verschillende servers. Zo kan er efficiënter met de ingekochte storage om gegaan worden en kan een beheerder makkelijk de storage van de ene host omzetten naar een andere host.

De meest gebruikte technieken op een storage netwerk zijn ethernet met iSCSI of Fibre Channel.

7.1 iSCSI

iSCSI is een methode om de SCSI commando's over een IP netwerk te versturen. De i staat dan ook voor Internet. Ethernet switches worden gebruikt om het netwerk op te zetten waarover TCP/IP en iSCSI gebruikt wordt om disken of partities te delen.

Het is mogelijk om iSCSI ook over het Internet te gebruiken, maar meestal wordt het gebruikt over losse staande netwerken om de optimale performance uit het netwerk te halen.

7.2 Fibre Channel

Fibre Channel draait over glasvezel, dit maakt het voor veel bedrijven onaan-trekkelijk omdat de meeste bedrijven een koper infrastructuur met ethernet hebben liggen. Het voordeel van glas zijn de hogere snelheden en de veel langere afstanden die overbrugd kunnen worden.

De laatste Fibre Channel standaard stamt uit 2008. We zien dat iSCSI en ethernet de markt van Fibre Channel aan het over nemen zijn.

Hoofdstuk 8

Network Fileserver

Desktop machines hebben een beperkte opslagcapaciteit en hebben weinig tot geen back-up mogelijkheden. Door de data te centraliseren op een fileserver kunnen we de capaciteit van de desktop uitbreiden zonder daar fysiek toegang tot te hebben. Daarnaast kunnen we de opgeslagen data op de fileserver eenvoudig centraal backupperen.

Een fileserver is een computer, server, in het netwerk die via een bepaald protocol bestanden aanbiedt aan gebruikers. In Microsoft Windows netwerken kan dit een SMB, Server Message Block, zijn en in een Linux omgeving bijvoorbeeld een NFS, Network File System, server zijn.

Het kenmerk van een fileserver is dat er een bepaald protocol (SMB, NFS) nodig is dat een bestandssysteem aanbiedt. Een gebruiker mount of mapped het filesysteem naar zijn lokale systeem en kan erop werken alsof het onderdeel is van zijn eigen machine.

8.1 NFS - Network File System

NFS stamt uit 1984 en is bedacht door Sun Microsystems. NFS is een netwerk bestandssysteem dat het mogelijk maakt om bestanden over het netwerk op een andere server te benaderen alsof het lokale bestanden zijn. Dit in tegenstelling met het tot dan toe veel gebruikte FTP waarbij je de een server actief moet benaderen om een bestand te downloaden voordat je het gebruiken kan.

NFS heeft twee protocollen, een om het bestandssysteem te koppelen (mount) en een om het gemounte bestandssysteem te benaderen (nfs). NFS gebruikt een aantal Remote Procedure Calls (RPCs) voor de toegang tot het bestandssysteem zoals het lezen en schrijven van bestanden.

NFS server houdt geen status bij van de clients, alleen welke client welke share gemount heeft. De client moet dus bij elke opdracht alle gegevens

meesturen. Dit heeft als groot voordeel dat je een NFS server kan herstarten zonder dat clients hiervan een verstoring ondervinden. Een nadeel is dat NFS dus niet zelf bijhoudt wie welk bestand benaderd en er dus twee gebruikers hetzelfde bestand kunnen schrijven. Het zogenaamde file-locking moet door een extern proces gebeuren (Network Lock Manager). Vanaf NFS 4 is locking wel een onderdeel van het protocol.

8.1.1 Security

Tot versie 4 kende NFS geen authenticatie methode. Toegang tot het bestandssysteem werd volledig overgelaten aan het onderliggende bestandssysteem zelf, dus de rechten op files en directories. NFS was dus alleen geschikt voor gebruik op een lokaal netwerk waar elke server en client over dezelfde gebruikersdatabase beschikt. Sinds versie 4 kan er gebruik gemaakt worden van kerberos voor authenticatie.

8.1.2 Interoperability

NFS is vanaf versie 2 (RFC1094) beschreven in RFCs en dus een open protocol dat door iedereen geïmplementeerd kan worden. Daarnaast is het een heel simpel protocol waardoor er voor bijna elk operating system een NFS client is en er is vaak ook een NFS server beschikbaar.

8.2 SMB - Server Message Block

SMB maakt het mogelijk om bestanden en printers te delen met het netwerk en heeft tevens een interproces communicatie mogelijkheid in de vorm van een door Microsoft ontwikkelde versie van RPC genaamd MSRPC.

Het SMB-protocol stamt uit 1983 en was oorspronkelijk ontwikkeld bij IBM. Microsoft gebruikte een verder door ontwikkelde versie in zijn Windows Operating System. SMB1 die in 2003 ook bekend was onder de naam CIFS (Common Internet File System) was een erg chatty protocol, waardoor het veel bandbreedte in gebruik nam en dus niet erg geschikt was voor gebruik over Internet links. Voor het delen van lokale resources werd het echter veel gebruikt. SMB1 zou inmiddels niet meer gebruikt moeten worden omdat het vele kwetsbaarheden bevat. SMB2 verminderde de chattiness van het protocol

8.2.1 Security

Al vroeg in de geschiedenis van SMB1 (Windows NT4 service pack 3) kreeg SMB de mogelijkheid van SMB signing. Dit betekende dat van een packet de authenticiteit vastgesteld kon worden. Hiermee werd een man-in-the-middle attack bijna onmogelijk.

SMB 3.0 bracht end-to-end encryption en SMB 3.0.2 geeft je de mogelijkheid om SMB1 niet meer te ondersteunen.

8.2.2 Interoperability

Buiten de Microsoft wereld wordt SMB gebruikt in het SAMBA-project, wat een Open Source project is om Unix-achtige systemen zoals Linux te verbinden met de Windows-wereld. Ook Apple heeft zijn eigen SMB-protocol implementatie. Eerst werd SAMBA gebruikt, maar sinds OS X 10.7 gebruiken ze een eigen implementatie met de naam SMBX.

8.2.3 Roaming profiles

8.3 NAS - Network Attached Storage

Een NAS of een Network Attached Storage is eigenlijk een fileservers in een doosje.

Hoofdstuk 9

Object Storage

Met het groeien van filesystemen wordt het steeds moeilijker om data aan te bieden. Op het moment dat we honderden terabytes of zelfs petabytes aan opslag hebben hebben we ook enorm veel bestanden. Mensen verdelen de data die zij opslaan niet evenredig over een bestandssysteem, sommige mappen bevatten veel bestanden en andere heel weinig. Het indexeren van al deze bestanden, als we bijvoorbeeld een listing opvragen van het bestandssysteem, duurt met de groei van het systeem steeds langer. De oplossing voor dit probleem is het niet meer gebruik maken van een bestandssysteem, maar van andere technieken. Een van die technieken is het gebruik van object storage.

Object storage wordt voornamelijk gebruikt als backend voor applicaties. Als je een bestand in een object storage systeem opslaat krijg je van het systeem een ID terug. Meestal in de vorm van een URL, bijvoorbeeld: <https://my.nextcloud.local/index.php/s/xLiiLba2gximHBt> Dit is een verwijzing naar het object op het systeem. Zoals je kan zien is dit niet lekker makkelijk te onthouden, vandaar dat dit vaak door applicaties gebruikt wordt, die dan een database gebruiken om een relatie te leggen tussen het object en het door de gebruiker opgeslagen bestand.

Hoofdstuk 10

Web storage

Een oplossing voor het object storage probleem is door gebruik te maken van het web of zoals dat tegenwoordig heet de cloud om je data in op te slaan. De webbrowser dient als frontend voor het opslagsysteem en het maakt dan niet meer uit wat we gebruiken als storage.

10.1 webdav

Het oorspronkelijke idee van het World Wide Web (WWW) was om bestanden met elkaar te delen en er gezamenlijk aan te kunnen werken. Het werd echter een alleen lezen systeem tot in 1996 Jim Whitehead het W3C overtuigde om een paar sessies te houden over de mogelijkheid van het schrijven van documenten. Daaruit kwam WebDAV voort wat staat voor Web-based Distributed Authoring and Versioning. Kort gezegd komt het erop neer dat WebDAV een protocol is om documenten te schrijven en de voorzien van een versie. Dat versioning bleek echter voor de eerste versie te ingewikkeld en men concentreerde zich op het gezamenlijk werken aan documenten. Later is er alsnog een versioning standaard gekomen.

WebDAV voegt een aantal request methods toe aan HTTP zodat het gebruikt kan worden als bestandssysteem voor lezen en schrijven. De "methodes" die WebDAV toevoegt zijn:

1. PROPFIND haalt de XML properties van een object op. Dit kan ook een complete filesystem tree zijn.
2. PROPPATCH wijzig 1 of meer properties van een object in 1 actie
3. MKCOL maakt een collectie aan, een collectie kan een map of directory zijn

4. COPY kopieert een object van een URI naar een ander
5. MOVE hernoemt een object van een URI naar een ander
6. LOCK vergrendel een object
7. UNLOCK ontgrendel een object

De functie om een bestand te schrijven naar een webserver bestond al in de PUT-methode.

10.2 Amazon S3 bucket

10.3 Google Drive

10.4 One Drive & Microsoft Azure BLOB

Index

- Amazon S3 bucket, [24](#)
- ATA, [3](#)
- Azure BLOB, [24](#)
- Backup, [9](#)
- Block devices, [3](#)
- DAS, [13](#)
- Direct Attached Storage, [13](#)
- Fibre Channel, [15](#)
- FTP, [17](#)
- Google drive, [24](#)
- Harddisks, [3](#)
- IDE, [3](#)
- Integrated Drive Electronics, [3](#)
- iSCSI, [15](#)
- JBOD, [3](#)
- Just a Bunch Of Disks, [3](#)
- Linear Tape Open, [9](#)
- LTO, [9](#)
- NFS, [17](#)
- Object storage, [21](#)
- One drive, [24](#)
- Parallel Advanced Technology
 Attachement, [3](#)
- PATA, [3](#)
- RAID, [11](#)
- RAID0, [11](#)
- RAID1, [11](#)
- RAID5, [11](#)
- Roaming profiles, [19](#)
- RPC, [17](#)
- SAN, [15](#)
- SAS, [5](#)
- SATA, [3](#)
- SCSI, [3](#)
- Serial ATA, [3](#)
- Serial Attached SCSI, [5](#)
- Small Computer Systems
 Interface, [3](#)
- SMB, [18](#)
- Solid State Drive, [3](#)
- SSD, [3](#)
- Storage Attached Network, [15](#)
- Universal Serial Bus, [5](#)
- USB, [5](#)
- USB-sticks, [3](#)
- Web storage, [23](#)
- Webdav, [23](#)