2015

De Haagse Hogeschool
Technische Informatica

Afstudeerverslag
Versie 1.0

Marco Stroosnijder
07053126

[PROOF OF CONCEPT: TRADITIONELE STORAGE VERVANGEN DOOR CLOUD STORAGE]

Documenthistorie

Revisiehistorie

Versie	Revisiedatum	Wijzigingen
0.1	11-12-2014	Initiële opzet
0.1.1	29-12-2014	Skelet + Invulling
0.1.3	05-01-2015	Benchmarking / Management GUI
0.1.4	11-01-2015	ArchiMate aanpassingen
0.1.5	17-01-2015	Sprint 1
0.1.6	28-01-2015	Sprint 2
0.1.7	31-01-2015	Sprint 3
0.2.0	1-02-2015	1ste conceptversie
0.3	13-02-2015	Commentaar afstudeerbegeleider werk verwerken
0.4	15-02-2015	Aanvulling schaalbaarheid benchmark
0.5	16-02-2015	Archimate aanpassingen
0.6	18-02-2015	Spelling sessie 1
0.7	19-02-2015	Spelling sessie 2
0.8	20-02-2015	Beoordeelbare versie opleveren
0.9.1	26-02-2015	Commentaar afstudeerbegeleider werk verwerken
0.9.5	09-03-2015	Rewrite adhv Feedback
1.0	26-03-2015	Definitieve versie

Distributie

Naam	Functie	Datum van uitgave	Versie
W.M. Mooijekind	Afstudeerbegeleider HHS	1-02-2015	0.2
W.M. Mooijekind M.W.H. Nieuwland	Afstudeerbegeleider HHS Examinator HHS	28-02-2015	0.9.1
W.M. Mooijekind M.W.H. Nieuwland 3de gecommitteerde	Afstudeerbegeleider HHS Examinator HHS Examinator HHS	30-03-2015	1.0

Inhoud

1 Inleiding

De afstudeeropdracht wordt uitgevoerd binnen de innovatieafdeling van het afstudeerbedrijf. De innovatieafdeling kijkt naar de ontwikkelingen op de markt en probeert actuele problemen binnen het bedrijf op te lossen met behulp van nieuwe technologieën.

Het afstudeerbedrijf maakt veelvuldig gebruik van “proven technology” en COTS-producten (commercial off-the-shelf). Dit is ook van toepassing op de storageoplossingen die gebruikt worden binnen het afstudeerbedrijf. Storage is vakjargon voor opslagcapaciteit. Er wordt gebruik gemaakt van prijzige hardwareoplossingen voorzien van prijzige licentiemodellen. Vaak zijn de storagesystemen overbelast en is horizontaal of verticaal opschalen niet mogelijk vanwege de techniek of het budget.

De afstudeeropdracht bestaat uit een proof of concept (PoC) voor het inzetten van servergebaseerde cloud opslagcapaciteit. Het afstudeerbedrijf heeft afdoende servercapaciteit, de servers zijn vaak voorzien van lokale opslagcapaciteit welke weinig of niet benut wordt. Deze onderbenutting komt met name voort uit het feit dat lokale storage niet geschikt is als “high available” en “shared storage”. De PoC onderzoekt de mogelijkheid om de lokale disks als “high available” “shared storage” aan te bieden in de vorm van “cloud storage”.

Het uit te voeren PoC beperkt zich tot de vervanging van een beperkte hoeveelheid traditionele storage. Wel wordt aandacht geschonken aan de schaalbaarheid van de cloud storage met het oog op kostenefficiënte vervanging van grootschalige traditionele storage oplossing.

Een tweeledige literatuurstudie als onderdeel van dit onderzoek heeft inzicht gegeven in zowel algemene storage- en cloudconcepten als productspecifieke concepten. De algemene storage- en cloudconcepten zijn gebruikt om de eisen en wensen van de opdrachtgever vast te stellen. Deze eisen en wensen zijn vervolgens gebruikt om een productselectie uit te voeren op basis van een MoSCoW analyse. Het productspecifieke deel van de literatuurstudie heeft na deze productselectie plaatsgevonden en heeft inzicht geboden in de voor het onderzoek te realiseren infrastructuur. De technische realisatie vindt plaats op basis van de Scrum ontwikkelmethode nadat de architectuur inzichtelijk is gemaakt in de ArchiMate modeleringstaal.

De technische realisatie van de infrastructuur an sich toont een groot deel van de haalbaarheid aan. Er wordt echter voor de aspecten schaalbaarheid, beschikbaarheid, security, beheerbaarheid, performance en kosten additioneel onderzoek uitgevoerd. De methodieken van deze individuele onderzoeksvragen, de resultaten van de onderzoeksvragen en de significantie hiervan worden in drie opeenvolgende hoofdstukken besproken.

Naast het beschrijven van de onderzoeksmethode wordt de volledige inrichting van de infrastructuur vastgelegd en geautomatiseerd in de configuratiemanagementtool Ansible. Deze beschrijvingen en vastleggingen maken het mogelijk het onderzoek in zijn geheel te reproduceren.

2 Theoretisch kader

In dit hoofdstuk worden kernbegrippen uit de afstudeeropdracht toegelicht. Na de toelichting van deze technische termen licht dit hoofdstuk ook een aantal gebruikte methodieken toe, zoals bijvoorbeeld de ArchiMate taal en de scrum projectmethode. De in het theoretisch kader beschreven concepten komen voor uit het eerste deel van de literatuurstudie. Voor de literatuurstudie is gezocht naar beschikbare boeken met betrekking tot storage en cloud binnen de bibliotheek van het afstudeerbedrijf. De zoekresultaten zijn verfijnd door alleen gebruik te maken van recente publicaties. Het zoeken is dan ook verfijnd door alleen boeken tussen 2011 en het heden weer te geven. Hierna is medewerker van de bibliotheek verzocht de geselecteerde bronnen bij de afstudeerder thuis af te leveren.

2.1 Cloud

Vaak wordt cloud (Schulz, 2011) gezien als een hype of een buzzword, dit komt door de brede definitie van het woord cloud. Cloud is een product, een technologie, een service en een management paradigma. De cloud kan op meerdere manieren gehost worden: public, private of een combinatie tussen deze technologieën, deze combinatie noemen we hybrid.

De locatie van de cloud maakt de definities nog breder: de applicaties en informatie kunnen gescheiden van elkaar staan; deze locatie kan zich verplaatsen; ontsloten zijn via een cloud gateway of een cloud router of zich aanpassen aan de locatie van de eindgebruiker (cloud point of presence).
Daarnaast kan cloud worden geleverd als turnkey oplossing, off-the-shelf, maatwerk, inclusief hardware, software, services of combinaties hiervan.

Na deze opsomming is het duidelijk dat cloud een breed begrip is. Om inzicht te geven in de definitie cloud binnen de context van dit onderzoek is gekozen om de opbouw van cloud services inzichtelijk te maken.

Cloud services zijn opgebouwd uit de 3 componenten: Compute, Network en Storage. In het volgende paragraaf wordt het begrip cloud services verder uitgelicht. In deze paragraaf wordt kort omschreven wat de componenten compute, network en storage voor rol spelen.

Figuur 1 Opbouw cloud services

2.2 Cloud services

Cloud services wordt vaak omschreven met de term *XaaS*. Binnen het acroniem *XaaS* staat de "*X*" voor "de onbekende variable" en *aaS* voor "as a Service". Veel voorkomende vormen van XaaS zijn:

Acroniem	Uitgeschreven afkoring
AaaS	Application as a service
BaaS	Backup as a service
DaaS*	Disk as a Service
DaaS**	Data Storage as a Service
IaaS	Infrastructure as a Service
PaaS	Platform as a Service
SaaS*	Storage as a Service
SaaS**	Software as a Service

Tabel 1: Overzich XaaS.
* Definitie uit (Schulz, 2011)
** Definitie uit (Yoder, 2013)

Opvallend binnen bovenstaande tabel is het dubbele gebruik van de termen DaaS en SaaS. De termen SaaS* en Daas* komen voort uit de bron (Schulz, 2011, p51). De termen SaaS** en Daas** komen voort uit de bronnen (Yoder, 2013, p231, p57). Verder onderzoek op internet toont aan dat beide termen gebruikt worden [1]. De meerderheid gebruikt echter de term SaaS als "Storage as a Service" en DaaS als "Disk as a Service". Daarnaast valt "Data Storage as a Service" buiten het stramien van het acroniem XaaS.

Wat met name van belang is binnen cloud services is dat ze *elastic* zijn of *elasticity* bevatten (Schulz, 2011, p55). Traditionele services werden in statische scenario's gerealiseerd. Zo hebben servers een vaste associatie met een traditionele storage. Cloud brengt elasticiteit, de infrastructuur is in staat zich aan te passen aan wijzigingen in organisaties en de fluctuerende behoefte aan resources.

Deze scriptie heeft met name betrekking op het cloud storage component. Een omschrijving van Cloud storage is te vinden in bijlage GG. Het "compute" component kan worden samengevat als de rekenkracht of de applicatie die de cloud service daadwerkelijk regelt. Het "network" component zorgt voor de communicatie en transport van data. Beide componenten zijn noodzakelijk voor de werking van de cloud storage maar binnen de PoC zijn deze niet elastisch genoeg om als cloud computing of networking te worden bestempeld.

[1] http://www.webopedia.com/TERM/S/SaaS.html

2.3 Storage Concepten

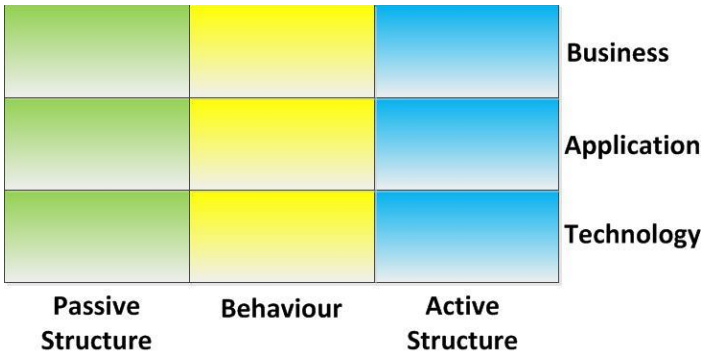
Er zijn drie mogelijke basisopslagstructuren om gegevens in op te slaan op storage componenten. In de Bijlage Opslagstructuren worden de individuele structuren besproken. Het betreft de stuctureren filesystem storage, block storage en object storage.

In de bijlage Bijlage Storage Technieken is informatie terug te vinden over de concepten;

- Snapshotting
- Deduplication
- Data encryption
- RAID
- LVM

2.4 ArchiMate

Voor aanvang van het afstuderen is er door de afstudeerder aangestuurd op het volgen van een enterprise architectuur cursus. In het afstudeerplan was het beschrijven van een architectuur opgenomen. De afstudeerder had hier echter niet de juiste bagage voor. Er is gekozen voor ArchiMate (Van Haren Publishing, 2013), de standaard voor enterprise-architectuur binnen het afstudeerbedrijf. ArchiMate is geen methodiek, ArchiMate is een gestandaardiseerde taal op basis van de IEEE 1471 standaard voor software architectuur. De tweedaagse cursus is afgerond met het succesvol behalen van een examen met certificering van The Open Group. De behaalde certificering is terug te vinden in bijlage A.



Figuur 2 Het archimate framework

Het Archimate model bestaat uit drie lagen (Van Haren Publishing, 2013, p. 8):

1. De businesslaag biedt producten en diensten aan aan externe klanten. De producten en diensten worden gerealiseerd door business processen die uitgevoerd worden door business actors.
2. De applicationlaag biedt ondersteuning aan de businesslaag met applicaties services die worden gerealiseerd door software applicaties.
3. De technologielaag bestaat infrastructuur services zoals processing, storage en communicatie services

Daarnaast zijn er nog drie aspecten: het passive structure aspect; behaviour aspect en het active structure aspect. (Van Haren Publishing, 2013, pp. 8-10)

Binnen ArchiMate zijn 26 viewpoints gedefinieerd, een viewpoint kan bijdragen aan het ontwerp, het maken van keuzes of informeren. Dit is mogelijk in drie abstractieniveaus. Ten eerste het detailniveau, dit niveau wordt over het algemeen gebruikt door en voor software engineers en

proceseigenaren. Ten tweede het samenhangniveau, dat gebruikt wordt op het operationele vlak. En ten slotte het overviewniveau dat gebruikt wordt door enterprise architecten en het management. (Van Haren Publishing, 2013, pp. 97-103)

Voor de architectuur van de IST- en SOLL-situatie is gekozen voor het infrastructure viewpoint, dit viewpoint draagt bij aan het ontwerp en vindt plaats op detailniveau. Het is van toepassing op de technologielaag en maakt gebruik van de aspecten passive, behaviour en active. Deze aspecten geven inzicht in de aard van het concept, zo is een virtualisatie image passief. Het biedt inzicht voor zowel techneut als opdrachtgever in aspecten zoals stabiliteit, security, afhankelijkheden en de kosten van de infrastructuur. (Van Haren Publishing, 2013, p. 123)

Architectuurtekeningen in ArchiMate bestaan uit concepten. Een samenvatting van alle concepten is terug te vinden in bijlage B. Een overzicht conform onderstaand voorbeeld van alle componenten uit de technology layer is terug te vinden in bijlage Q.

"An artifact is defined as a physical piece of data that is used or produced in a software development process, or by deployment and operation of a system." Een artifact maakt onderdeel uit van de passive stucture.

Figuur 3: ArchiMate Artifact

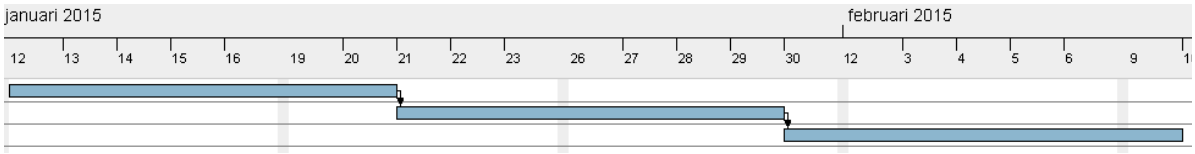
ArchiMate wordt binnen het afstudeerbedrijf gebruikt als communicatiemiddel, het ontwerp is inhoudelijk doorgesproken met een gecertificeerd medecursist. En ArchiMate–technisch (het gebruik van componenten en lijnen) met een masterstudent architectuur. Op basis van de gesprekken kwamen kleine ontwerpfouten aan het licht, deze zijn aangepast in het uiteindelijke ontwerp.

2.5 Scrum

Als ontwikkelmethode voor de technische realisatie is gekozen voor de agile projectmethodiek Scrum (Collaris, 2012) (Verheyen, 2013) Gezien de omvang van de project-groep, slechts 1 projectlid en de korte periode waarin technische realisatie plaats moest vinden is een agile methode de best passende. De werkzaamheden zijn samengevat in 3 sprints die elk 7 werkdagen omvatten.

Naam	Begin datum	Eind datum
Sprint 1 - infrastructuur	12-1-15	20-1-15
Sprint 2 - installatie en configuratie	21-1-15	29-1-15
Sprint 3 - reproduceerbaar	30-1-15	9-2-15

Figuur 4 sprint planning



Figuur 5 sprint planning in grafiek vorm

In bijlage LL Bijlage Scrum concepten is aanvullende informatie te vinden over de begrippen sprint, backlog en standup.

2.6 Ansible

Binnen deze opdracht is gebruik gemaakt van configuratie management. De gebruikte tool hiervoor is Ansible[2]. Ansible is een configuration management engine. Ansible stelt een systeembeheerder in staat configuraties te standaardiseren en periodiek te controleren en indien noodzakelijk aan te passen naar de standaard. Ansbile heeft 5 kern begrippen om die te realiseren: duidelijk, snel, krachtig, efficiënt en veilig. Ansbile maakt geen gebruik van een clientsoftware maar van de (binnen het afstudeerbedrijf) reeds aanwezig SSH toegang. Via deze SSH toegang is Ansbile in staat vanuit een centraal punt aanpassingen door te voeren op individuele servers.

2.7 Ceph

Aan de hand van de productselectie (hoofdstuk 4) is Ceph als meest geschikt aangemerkt. In bijlage MM is een beschrijving te vinden van de Ceph gerelateerde begrippen Journal, OSD, Cephx authenticatie, Pool, Placement groups en CRUSHmaps

[2] http://www.ansible.com/home

3 IST-Situatie

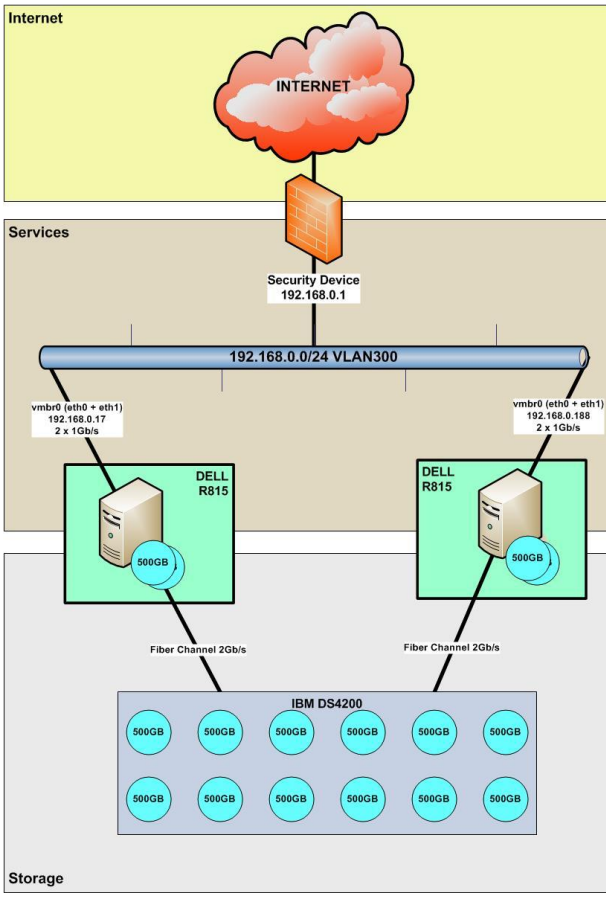
Dit hoofdstuk omschrijft de huidige situatie zoals deze voor aanvang van het afstuderen werd aangetroffen. De afstudeerder werkt op dagelijkse basis met deze omgeving en heeft inzicht in de werking en opzet van deze omgeving. Het geeft inzichten in de architectuur en werking. Daarnaast biedt dit hoofdstuk inzicht in de potentiele uitbreidingsmogelijkheden op basis van de huidige infrastructuur.

3.1 Hoofdlijnen

De huidige infrastructuur levert virtualisatiediensten op basis van het Open Source virtualisatieplatform Proxmox. Proxmox is een gratis alternatief voor het virtualisatieplatform VMWare.

De virtualisatiedienst bestaat uit een internet verbonden cluster van 2 Proxmox nodes. Het cluster maakt het mogelijk virtualisatiehosts te migreren tussen de nodes. Deze migraties maken het mogelijk virtualisatiehosts tijdens onderhoud aan één van beide Proxmox nodes beschikbaar te blijven stellen.

Beide Proxmox nodes zijn op basis van Fibre Channel (FC) redundant verbonden met een IBM DS4200 diskarray. Het IBM DS4200 diskarray is gevuld met 12 500GB disks welke op basis van RAID5 als LUN's worden aangeboden aan beide Proxmox nodes. De Proxmox nodes verdelen deze block storage weer verder onder de individuele virtualisatiehosts al naar gelang de behoefte.



Figuur 6: Proxmox Virtualisatie IST-situatie0

Internet

De virtualisatieomgeving is verbonden met het internet door middel van een 1Gb/s verbinding. Ter verbetering van de beveiliging is er een StateFull firewall geplaatst tussen het internet en de virtualisatieomgeving. Per dienst worden expliciete firewallregels aangemaakt voor de te leveren diensten. Diensten worden van elkaar gescheiden door middel van VLAN technologie. Gelijksoortige diensten worden samengevoegd in een dedicated VLAN. Wanneer diensten niet in een dedicated VLAN worden geplaatst komen deze in het “standaard” VLAN 300.

Services

De virtualisatieomgeving bestaat uit twee Dell R815 servers. De servers in de virtualisatieomgeving maken gebruik van shared storage ontsloten via Fibre Channel. Het gebruik van Fibre Channel beperkt het aantal servers momenteel tot 2. In de paragraaf “Storage” wordt verder uitgeweid over deze beperking. Naast het gebruik van de shared storage ten behoeve van het hosten van virtualisatie images is iedere server voorzien van in totaal 6 500GB harddisks. In de praktijk zijn deze als volgt toegepast: twee harddisken zijn samengevoegd tot een redundante disk(RAID 1) voor het lokale operatingsystem, de overige vier harddisken worden niet gebruikt. Er is geen nut of noodzaak om deze harddisken in te zetten in de huidige situatie, de virtualisatieomgeving heeft alleen baat bij shared storage. Zonder shared storage is het niet mogelijk virtuele servers te verplaatsen naar andere Proxmox node. De aan de betreffende virtuele server gekoppelde lokale storage is slechts beschikbaar op één enkele Proxmox node.

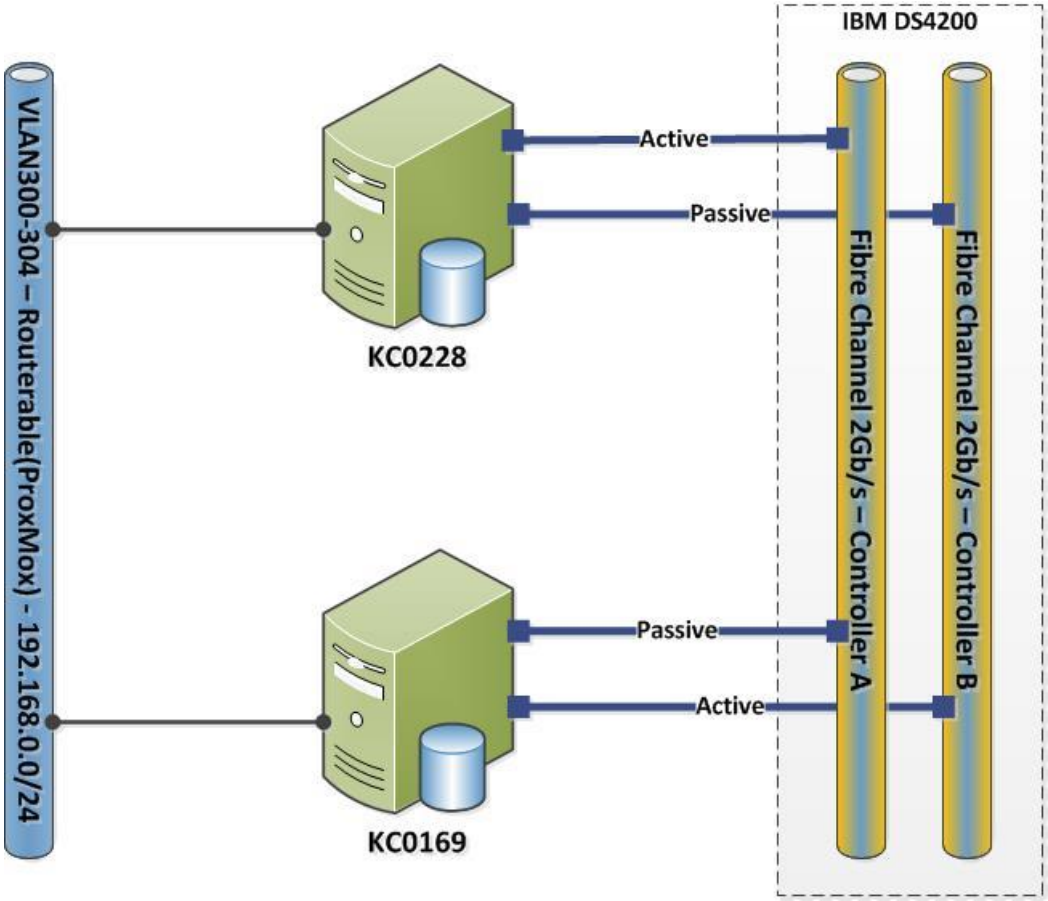
Storage

De shared storage ten behoeve van de virtualisatieomgeving is gerealiseerd op basis van een IBM DS4200 disk enclosure. Het enclosure is inmiddels ±8 jaar oud. Het systeem biedt SAN functionaliteit, met andere woorden de DS4200 biedt block storage aan de virtualisatieomgeving. In het verleden werd het DS4200 disk enclosure gebruikt in combinatie met een tweetal Fibre Channel switches. Beide switches zijn echter defect en afgevoerd. Het ontbreken van de switches beperkt de omgeving qua schaalbaarheid, zo is het slechts mogelijk 2 systemen redundant aan te sluiten op de DS4200 disk enclosure. Metingen binnen dit onderzoek, en dan met name het performance onderzoek, houden rekening met deze beperking van 2 nodes.

Figuur 7: IBM DS4200 Disk Array

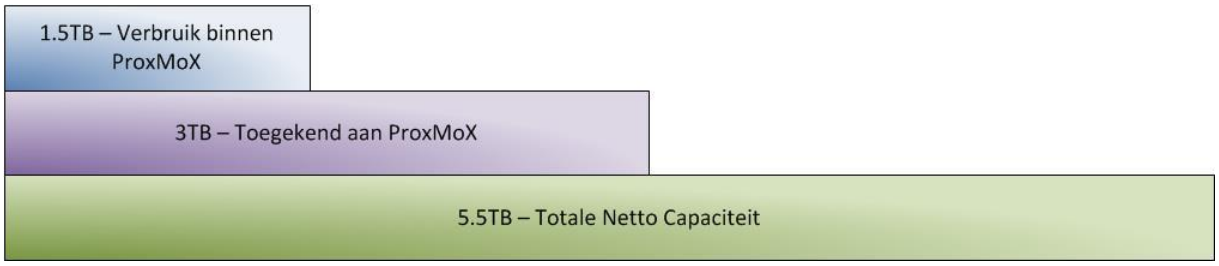
De virtualisatie nodes zijn aangesloten middels Fibre Channel, het boek SNIA Dictionary omschrijft Fibre Channel als volgt (Yoder, 2013, p. 104):
“ A serial I/O interconnect capable of supporting multiple protocols, including access to open system storage(FCP), access to mainframe storage (FICON) and networking (TCP/IP). Fibre Channel supports point to point, arbitrated loop, and switched technologies with a variety of copper and optical links running at speeds from 1 Gb/s to 10Gb/s.”

Aan de hand van deze omschrijving is het mogelijk de huidige inrichting verder te beschrijven:
“De DS4200 oplossing biedt open system storage (FCP) in een point to point opstelling via optische bekabeling met een snelheid van 2Gb/s.“ Hierbij moet opgemerkt worden dat zowel de actieve als de passieve verbinding in staat is op 2Gb/s te communiceren.



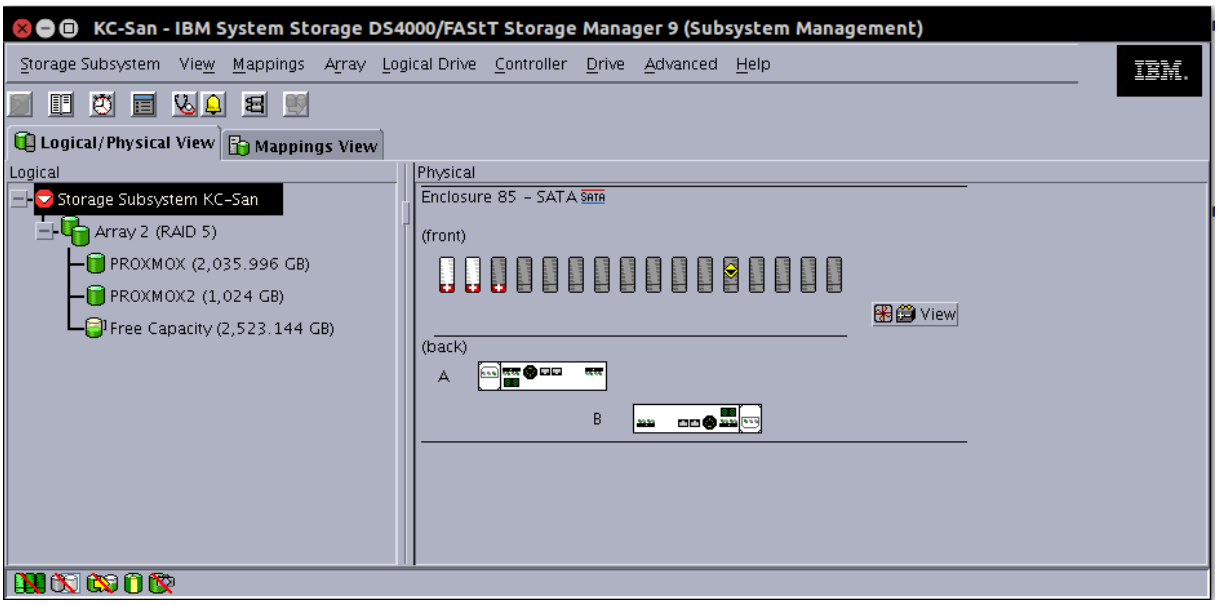
Figuur 8: Schematische weergave van de netwerk- en Fibre Channelaansluitingen

Er zijn 12 harddisks aanwezig in de DS4200 disk enclosure waarvan er 3 als hotspare zijn geconfigureerd. De overige 9 disks leveren in een RAIDs 5 (0) opstelling een totale capaciteit van 5.5TB. Hiervan is 3TB toegekend aan het Proxmox virtualisatiecluster en 2.5 TB vrij bruikbaar. In de praktijk is slechts 1.5TB van de aan Proxmox toegekende capaciteit in gebruik. De aan Proxmox toegekende capaciteit bestaat uit twee delen van respectievelijk 2 en 1TB. Deze volumes zijn op de Proxmox hosts samengevoegd tot één logisch volume van 3TB op basis van LVM (0).



Figuur 9 Verbruik DS4200 storage capaciteit

Deze informatie komt voort uit de management GUI van het IBM DS4200 enclosure en de beheerinterface van Proxmox.



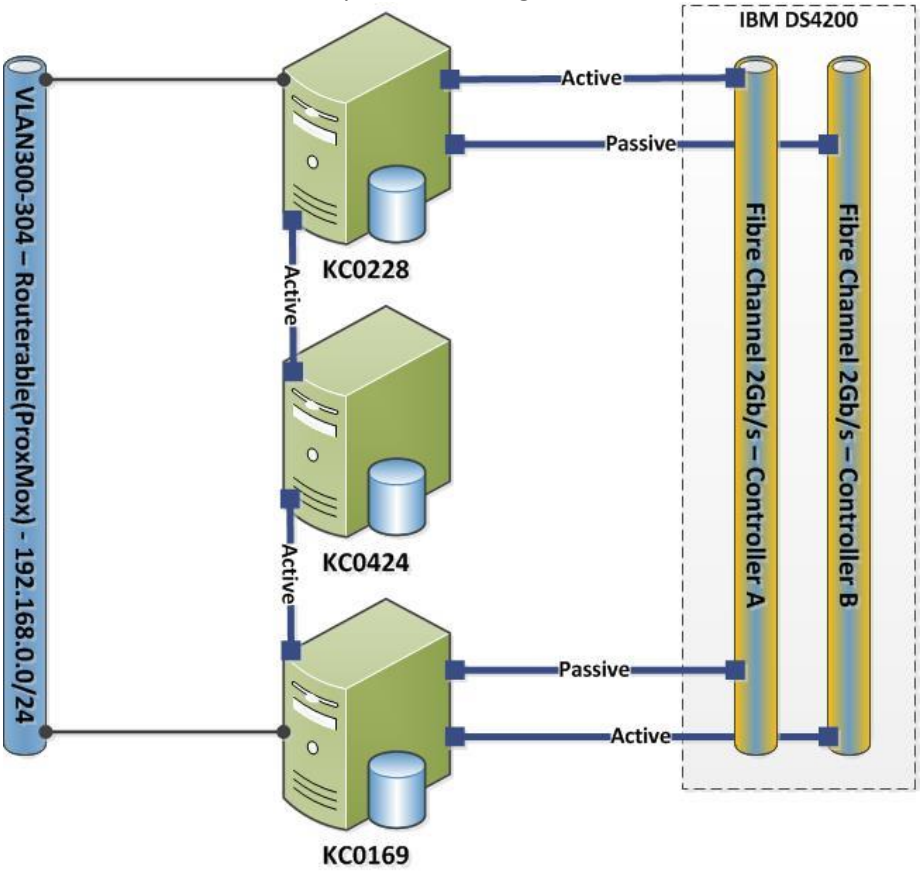
Figuur 10: Beheer interface DS4200 – RAID5, spares en verbruik

3.2 Uitbreidingsmogelijkheden

De opdrachtgever heeft de wens het bestaande End Of Life (EOL) disk enclosure af te stoten. Er zijn echter nog mogelijkheden om het bestaande disk enclosure beter schaalbaar te maken. Dit hoofdstuk omschrijft de uitbreidingsmogelijkheden. Deze uitbreidingsmogelijken zijn in kaart gebracht om kenbaar te maken dat nog wel reële uitbreidingsmogelijkheden zijn binnen de huidige infrastructuur.

3.2.1 Fibre Channel Arbitrated Loop (FC-AL)

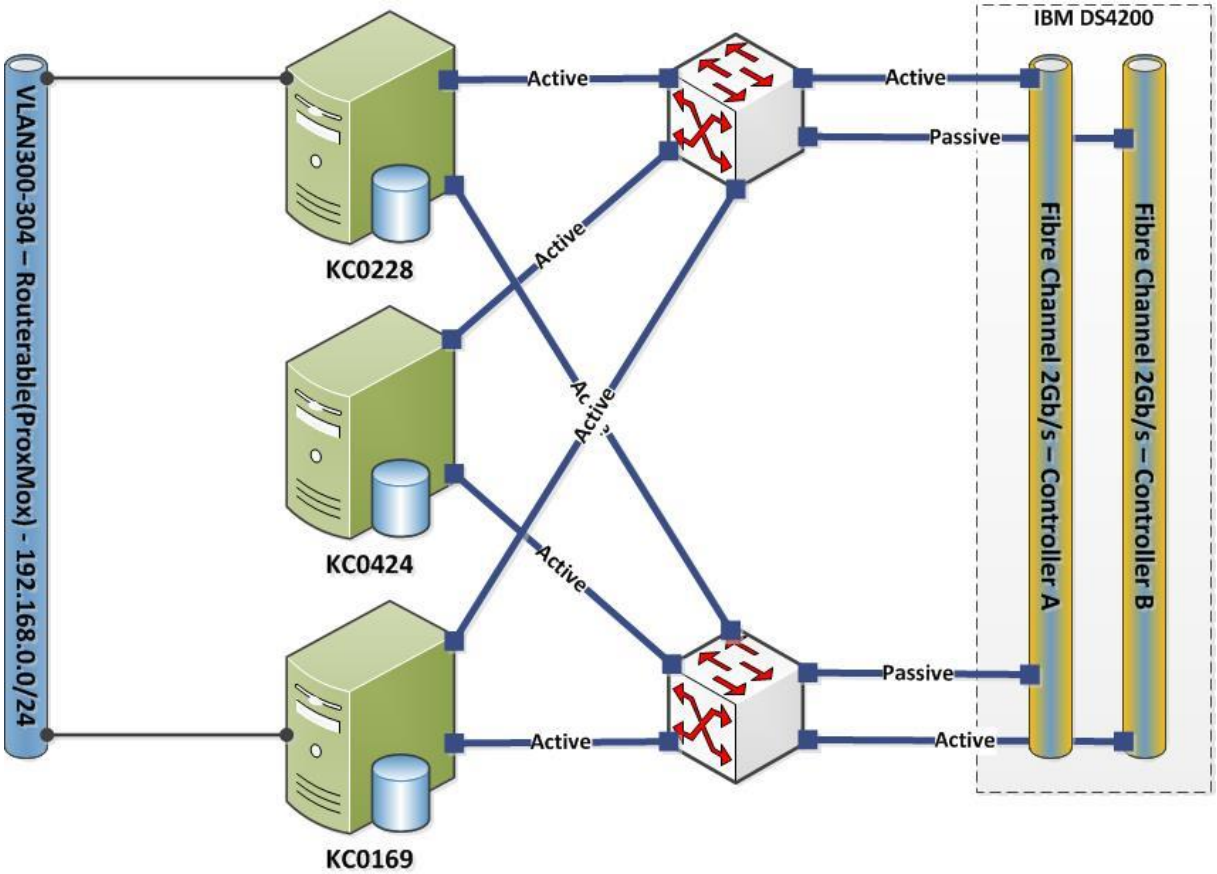
FC-AL (Yoder, 2013, p. 104) maakt het mogelijk tot maximaal 126 system met elkaar te verbinden in een lus. Deze optie maakt de storage oplossing horizontaal schaalbaar. De kosten van deze oplossing zijn relatief laag. Er hoeft alleen geïnvesteerd te worden in additionele FC-adapters en FC-bekabeling. De oplossing levert geen additionele capaciteit of performance. De performance neemt zelfs af al naar gelang er verder horizontaal wordt geschaald. Voorbeeld: de schrijfcapaciteit van het IBM DS4200 enclosure is 120MB/s. Bij 2 nodes is het mogelijk 60MB/s per node weg te schrijven. Bij 3 nodes is het mogelijk 40MB/s per node weg te schrijven. Inzicht in het begrip performance binnen de context van deze afstudeeropdracht is terug te vinden in hoofdstuk 7.



Figuur 11 Oplossing op basis van FC-AR

3.2.2 Fibre Channel switches

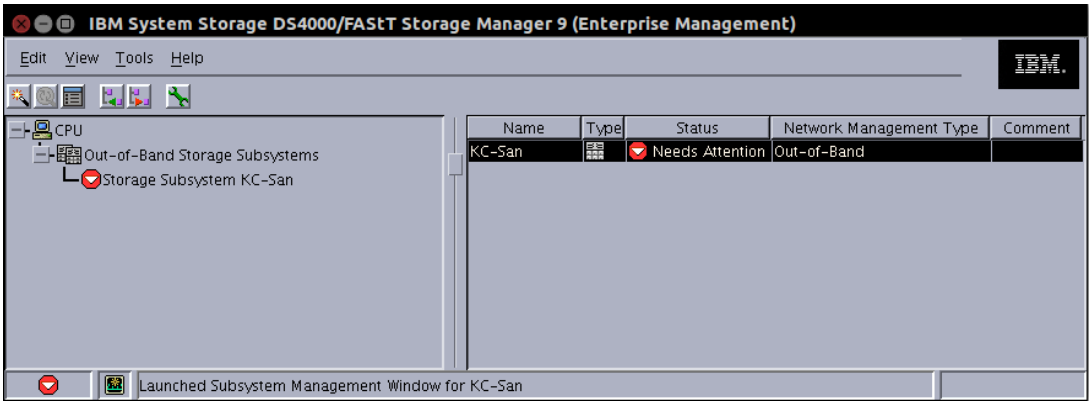
Deze oplossing bestaat uit het plaatsen van twee Fibre Channel switches. In het verleden was dit daadwerkelijk de gebruikte opstelling. Naast gebrek aan kennis binnen de afdeling hebben hardware-defecten bijgedragen aan het uitfaseren van de Fibre Channel switches. Het opnieuw investeren in Fibre Channel switches brengt hoge kosten met zich mee, maar het levert geen additionele capaciteit of performance. De performance neemt zelfs af al naar gelang er verder horizontaal wordt geschaald.



Figuur 12 Oplossing met Fibre Channel switches

3.2.3 1TB disks

Om de capaciteit van het disk enclosure te verhogen is het mogelijk de harddisks in het disk enclosure te **upgraden naar 1TB disks**. Echter de ouderwetse harddisks voor het huidige disk enclosure zijn niet meer nieuw te leveren. De disks die momenteel worden geleverd zijn refurbished. De kosten voor een refurbished 500 GB bedragen momenteel €900. Er zijn wel goedkopere disks beschikbaar met gelijke specificaties echter heeft IBM ervoor gekozen het disk enclosure alleen disks te laten herkennen indien deze is voorzien van IBM firmware.

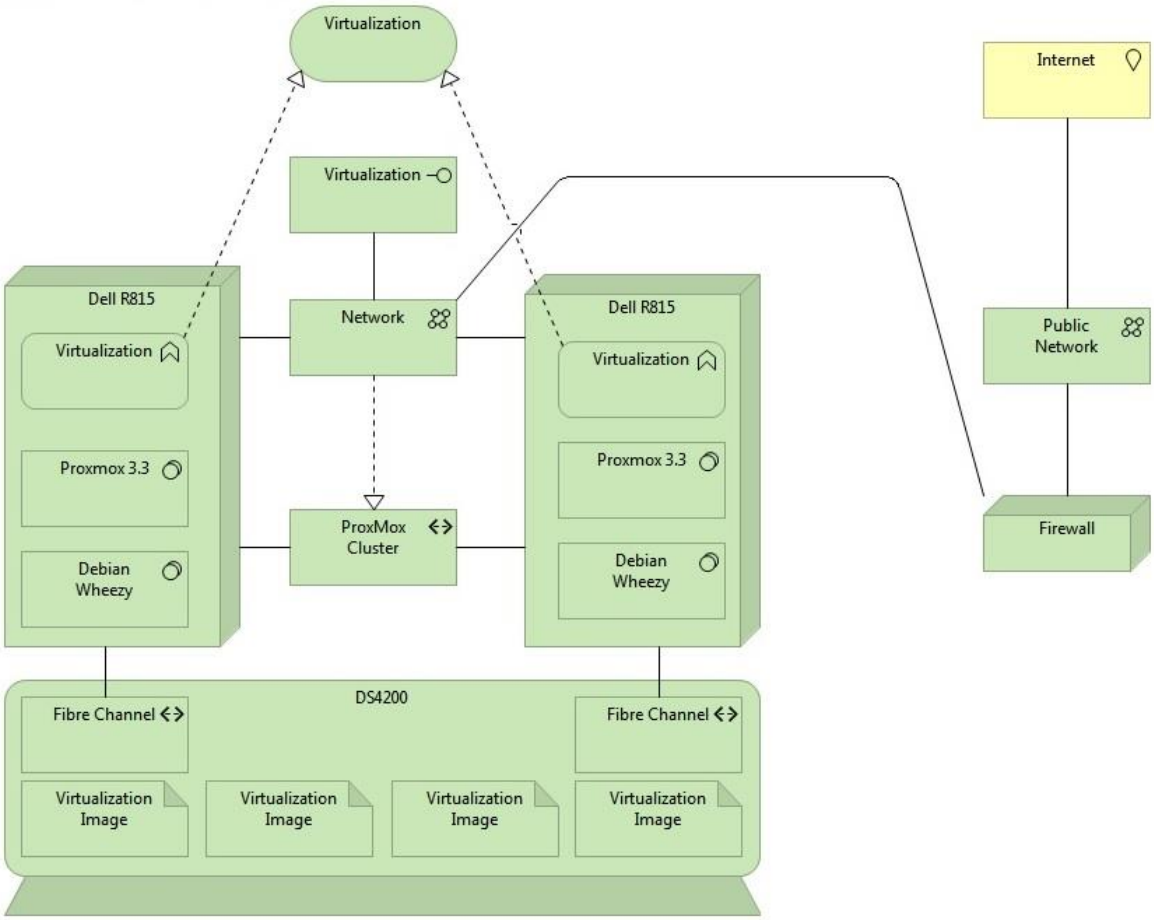


Figuur 13: Weergave defecte disk in beheer GUI

Tijdens het schrijven van mijn scriptie trof ik het systeem aan met een defecte disk. Het was noodzakelijk deze disk te vervangen om een eerlijk vergelijk te doen van de performance. Hoewel er spare refurbished schijven aanwezig waren was het disk enclosure niet in beheer. Na contact met IBM kreeg het afstudeerbedrijf de software om de DS4200 te beheren niet, dit was op basis van het feit dat het afstudeerbedrijf geen lopend service contract heeft voor de DS4200 enclosure. De software werd met een gerichte zoekactie in de homedirectory van (en in overleg met) een oud medewerker die in het verleden verantwoordelijk was voor het enclosure, aangetroffen. Met deze software was het mogelijk het disk enclosure in beheer te nemen voor aanvang van de vervanging van de harddisk. Het vervangen van de disk verliep zonder problemen.

3.3 Architectuur

Aan de hand van de beschrijving van de IST-situatie is op basis van de ArchiMate taal de volgende tekening tot stand gekomen vanuit het infrastructure viewpoint.



Figuur 14 Archimate tekening IST situatie

4 Productselectie

De eisen en wensen die gebruikt worden voor de productselectie zijn tijdens de literatuurstudie tot stand gekomen aan de hand van een overleg met de opdrachtgever. Na de literatuurstudie heeft er nog een iteratie plaatsgevonden van dit overleg. In deze iteratie zijn met name de juiste technische termen inzichtelijk gemaakt voor de opdrachtgever. Ook zijn er enkele potentiele additionele wensen aangedragen bij de opdrachtgever. De meerderheid hiervan is opgenomen als daadwerkelijke wens voor de productselectie. In bijlage FF zijn de eisen en wensen terug te vinden met enkele regels context.
De initiële productselectie is uitgevoerd op basis van drie door de opdrachtgever gestelde eisen. De eerste eis is “Het product dient op basis van vrij verkrijgbare open source software beschikbaar te zijn” en de tweede eis is “het product moet voorzien in het aanbieden van virtualisatieimages” ten derde mochten de oplossingen geen gebruik maken van Fibre Channel. Deze criteria zijn gebruikt om het aantal resultaten te beperken. Indien er geen gebruikt wordt gemaakt van deze criteria waren er honderden potentiele mogelijkheden die op een van deze criteria af zouden vallen. Aan de hand van de drie criteria is onderzoek uitgevoerd op het internet naar beschikbare producten. Aan de hand van Google queries werden de volgende producten aangetroffen.

4.1 Open Source cloud projecten

De volgende open source projecten zijn uit de productselectie voortgekomen. Om een indruk te wekken zijn summiere beschrijvingen van de respectievelijke producten verzameld van de sites van de betreffende projecten(leveranciers).

Gluster[3]

“GlusterFS is an open source, distributed file system capable of scaling to several petabytes (actually, 72 brontobytes!) and handling thousands of clients. GlusterFS clusters together storage building blocks over Infiniband RDMA or TCP/IP interconnect, aggregating disk and memory resources and managing data in a single global namespace. GlusterFS is based on a stackable user space design and can deliver exceptional performance for diverse workloads.”

Ceph[4]

“Ceph is a distributed object store and file system designed to provide excellent performance, reliability and scalability. Ceph was made possible by a global community of passionate storage engineers and researchers. Ceph is open source and freely-available, and it always will be.”

OpenStack Swift[5]/[6]

“Swift is a highly available, distributed, eventually consistent object/blob store. Organizations can use Swift to store lots of data efficiently, safely, and cheaply.”

[3] http://www.gluster.org/
[4] http://ceph.com/
[5] http://www.openstack.org/software/openstack-storage/
[6] http://docs.openstack.org/developer/swift/

Hadoop (2) – HDFS[7]

"The Hadoop Distributed File System (HDFS) is a distributed file system designed to run on commodity hardware. It has many similarities with existing distributed file systems. However, the differences from other distributed file systems are significant. HDFS is highly fault-tolerant and is designed to be deployed on low-cost hardware. HDFS provides high throughput access to application data and is suitable for applications that have large data sets. HDFS relaxes a few POSIX requirements to enable streaming access to file system data. HDFS was originally built as infrastructure for the Apache Nutch web search engine project. HDFS is now an Apache Hadoop subproject."

XtreemFS[8]

"XtreemFS is a general purpose storage system and covers most storage needs in a single deployment. It is open-source, requires no special hardware or kernel modules, and can be mounted **on** Linux, Windows and OS X"

Lustre[9]

"For the world's largest and most complex computing environments, the Lustre file system redefines high performance, scaling to tens of thousands of nodes and petabytes of storage with groundbreaking I/O and metadata throughput."

Sheepdog[10]

"Sheepdog is a distributed object storage system for volume and container services and manages the disks and nodes intelligently. Sheepdog features ease of use, simplicity of code and can scale out to thousands of nodes."

[7] http://hadoop.apache.org/docs/r1.2.1/hdfs_design.html
[8] http://www.xtreemfs.org/
[9] http://wiki.lustre.org/index.php/Main_Page
[10] http://sheepdog.github.io/sheepdog/

4.2 Determineren MoSCoW

De volgende tabel is opgesteld om inzicht te geven in de eisen en wensen, de eisen zijn per definitie een must. De wensen zijn onderverdeeld in should en could. De tabel is teruggekoppeld met de opdrachtgever. De opdrachtgever is akkoord gegaan met de onderverdeling van eisen en wensen in must have, should have en could haves. De must, should en could waarden worden binnen dit document gebruikt voor de productselectie.

	Must	Should	Could	Won't
Het aanbieden van virtualisatieimages	X			
Proxmox dient de storage te ondersteunen	X			
Private cloud storage	X			
Block storage	X			
Filesystem storage		X		
Object storage		X		
Open Source Software (OSS)	X			
Performance		X		
Opslagcapaciteit		X		
Horizontaal schaalbaar	X			
Minimaal 2x maal opgeslagen	X			
Gescheiden (storage) netwerk		X		
Integriteit van de data	X			
Interactie met de cloudservices			X	
Zeer beperkte investeringen		X		
Grafische interface		X		
Rolling upgrades		X		
Snapshotting		X		
Deduplication			X	
Data encryption			X	
Performance degradatie bij replicatie			X	
Geen Fibre Channel	X			

Tabel 2 Onderverdeling Must, Should en Could

4.3 Scoren potentiële oplossingen

In dit hoofdstuk is aan de hand van literatuurstudie inzichtelijk gemaakt welke producten al dan niet voldoen aan de naar must, should en could haves vertaalde eisen en wensen. De must, should en could haves zijn verdeeld over drie tabellen. Aan de hand van de symbolen +, - en de combinatie tussen deze symbolen, "+/-", wordt aangegeven of het product wel(+) of niet(-) voldoet aan de gestelde eisen of wensen. Indien er twijfel is over het wel of niet voldoen, is er gebruik gemaakt van de combinatie "+/-". Indien er twijfel is wordt deze twijfel beargumenteerd. Alle scoring wordt overbouwd met een bronvermelding, over het algemeen met een bron afkomstig van de betreffende leverancier.

Aan de plusjes wordt een waarde van 1 gekoppeld, aan de – een waarde van 0 en aan +/- een waarde van 0.5. Primair vindt productselectie plaats op basis van de scoring van de must haves, bij gelijke scoring worden de should haves meegenomen. Indien er op het vlak van de should haves wederom gelijk gescoord wordt valt de productkeuze terug op de scoring van de could haves. Bij een totaal gelijkspel tussen producten wordt in overleg met de opdrachtgever een product gekozen. In bijlage EE is inzichtelijk gemaakt op basis van welke bron informatie er gebruikt is bij de onderbouwing met een + of -. De bijlage bestaat uit een tabel per criteria per product.

Should	**Filesystem storage**
Product	Ceph
Link	http://ceph.com/ceph-storage/file-system/
Resultaat	+
Motivatie voor resultaat	Ceph provides a traditional file system interface with POSIX semantics.

Tabel 3 Voorbeeld onderbouwing keuze productselectie

4.3.1 Must

Tijdens het realiseren van de must have toetsingstabel kwam aan het licht dat Proxmox gebruik kan maken van alle storage welke op het onderliggende Linux OS ondersteund wordt. Na overleg met de opdrachtgever moest deze must have aangepast worden. De originele eis luidde **Proxmox dient de storage te ondersteunen** , de nieuwe eis luidde **Proxmox dient de storage <u>native</u> te ondersteunen**. In het geval van niet native Proxmox storage wordt er geen gebruik gemaakt van block storage maar van filesystem storage, in de praktijk wordt er een groot block weggeschreven in een directory.
De opdrachtgever maakte kenbaar dat alle must have criteria zogenaamde "knock-out criteria" zijn, dit wil zeggen: indien het product niet aan de gestelde eis voldoet, dan is het product per definitie geen geschikte oplossing.

In onderstaande tabel zijn op de horizontale-as de producten geplaatst. Op de verticale-as de must haves. Het criterium Proxmox dient de storage native te ondersteunen is zoals eerder beschreven een knock-out criterium.Omdat Swift, HDFS, XtreemFS, Lustre en Sheepdog niet native ondersteund worden binnen Proxmox is tijd bespaard door de resterende knock-out criteria niet te onderzoeken. Geen van de oplossingen maakt gebruik van Fibre Channel.

	Gluster	Ceph	Swift	HDFS	XtreemFS	Lustre	Sheepdog
Het aanbieden van virtualisatieimages[11]	+	+	+	+	+	+	+
Proxmox dient de storage native te ondersteunen	+[12]	+[13]	-	-	-	-	+/-[14]*
Private cloud storage	+[15]	+[16]					
Block storage	+[17]	+[18]					
Open Source Software (OSS) [19]	+	+	+	+	+	+	+
Horizontaal schaalbaar	+[20]	+[21]					
Minimaal 2x maal opgeslagen	+[22]	+[23]					
Integriteit van de data	+[24]	+[25]					
Geen Fibre Channel	+	+	+	+	+	+	+

Tabel 4 Must have evaluatie

*https://pve.Proxmox.com/wiki/Storage:_Sheepdog:

Note: Sheepdog is still not stable, so please do not use it in production environments.

Zowel Gluster als Ceph voldoen aan alle eisen.

[11] Initieel selectie criterium
[12] https://pve.Proxmox.com/wiki/Storage:_GlusterFS
[13] https://pve.Proxmox.com/wiki/Storage:_Ceph
[14] https://pve.Proxmox.com/wiki/Storage:_Sheepdog
[15] http://en.wikipedia.org/wiki/Gluster#Private_cloud_deployment
[16] http://ceph.com/community/career/storage-consultant/
[17] https://raobharata.wordpress.com/2013/11/27/glusterfs-block-device-translator/
[18] http://ceph.com/ceph-storage/block-storage/
[19] Initieel selectie criterium
[20] http://www.gluster.org/
[21] http://ceph.com/
[22] http://www.gluster.org/community/documentation/index.php/Gluster_3.1:_Configuring_Distributed_Replicated_Volumes
[23] http://ceph.com/docs/dumpling/rados/operations/pools/
[24] https://access.redhat.com/documentation/en-US/Red_Hat_Storage/2.0/html/Administration_Guide/sect-User_Guide-Managing_Volumes-Self_heal.html
[25] http://wiki.ceph.com/FAQs/How_Does_Ceph_Ensure_Data_Integrity_Across_Replicas

4.3.2 Should

In onderstaande tabel zijn op de horizontale-as de producten geplaatst. Op de verticale-as de should haves. Aangezien de producten Gluster en Ceph gelijk gescoord hebben bij het evalueren van de must haves en de overige producten af zijn gevallen, zijn alleen Gluster en Ceph op de horizontale-as geplaatst.

	Gluster	Ceph
Filesystem storage	+[26]	+[27]
Object storage	+/-[28]*	+[29]
Performance	+[30]**	+[31]**
Opslagcapaciteit	+***	+***
Gescheiden (storage) netwerk	+[32]	+[33]
Zeer beperkte investeringen	+****	+****
Grafische interface	+[34]	+[35]
Rolling upgrades	+[36]	+[37]
Snapshotting	+[38]	+[39]

Tabel 5 Should have evaluatie

*Gluster ondersteunt native geen object storage, wel in combinatie met Swift.
** Er is geen eerlijk vergelijk tussen beide producten te vinden, daarom beide +
*** beiden gebruiken dezelfde hardware
**** beiden gratis verkrijgbaar als open source software

Ceph scoort een fractie beter dan Gluster en daarom valt Gluster af bij het beoordelen van de could haves.

[26] http://www.gluster.org/documentation/About_Gluster/
[27] http://ceph.com/ceph-storage/file-system/
[28] http://www.gluster.org/wp-content/uploads/2012/05/Gluster_File_System-3.3.0-Administration_Guide-en-US.pdf
[29] http://ceph.com/ceph-storage/object-storage/
[30] http://www.networkcomputing.com/storage/gluster-vs-ceph-open-source-storage-goes-head-to-head/a/d-id/1113581
[31] http://www.networkcomputing.com/storage/gluster-vs-ceph-open-source-storage-goes-head-to-head/a/d-id/1113581
[32] http://www.gluster.org/community/documentation/index.php/Network_Configuration_Techniques
[33] http://ceph.com/docs/master/rados/configuration/network-config-ref/
[34] http://www.gluster.org/community/documentation/index.php/Gluster_3.1:_Using_the_Gluster_Management_Console
[35] https://github.com/ceph/calamari
[36] http://www.gluster.org/community/documentation/index.php/Upgrade_to_3.5
[37] http://ceph.com/docs/master/install/upgrading-ceph/
[38] http://www.gluster.org/community/documentation/index.php/Features/Gluster_Volume_Snapshot
[39] http://ceph.com/docs/master/rbd/rbd-snapshot/

4.3.3 Could

Uit de should have analyse is Ceph naar voren gekomen als het beste product voor deze PoC. De opdrachtgever heeft echter verzocht de could haves ook in kaart te brengen voor Ceph. In onderstaande tabel zijn op de horizontale-as Ceph geplaatst. Op de verticale-as de could haves.

	Ceph
Interactie met de cloudservices	+[40/41/42]
Deduplication	-[43]
Data encryption	-[44]
Performance degradatie bij replicatie	+[45]

Tabel 6 Could have evaluatie

Alleen de producten Gluster en Ceph voldoen aan alle gestelde eisen. Op basis van het ontbreken van native object storage binnen Gluster geniet het gebruik van Ceph de voorkeur. Gluster biedt de mogelijkheid tot het aanbieden van object storage indien er naast Gluster gebruik wordt gemaakt van Swift. In overleg met de opdrachtgever is daarom Ceph gekozen als het product wat verder in deze opdracht gebruikt en onderzocht gaat worden.

[40] http://ceph.com/docs/master/radosgw/s3/
[41] http://www.inktank.com/for-service-providers/
[42] https://www.openstack.org/summit/openstack-summit-hong-kong-2013/session-videos/presentation/ceph-the-de-facto-storage-backend-for-openstack
[43] http://lists.ceph.com/pipermail/ceph-users-ceph.com/2013-August/033451.html
[44] http://docs.ceph.com/docs/v0.80/rados/operations/authentication/
[45] http://ceph.com/docs/master/rados/configuration/osd-config-ref/

4.3.4 Relatie met onderzoekscriteria

De eisen en wensen met betrekking tot de PoC zijn getoetst aan de criteria schaalbaarheid, beschikbaarheid, security, beheerbaarheid, performance en kosten. In onderstaande tabel zijn de verbanden tussen de eisen en wensen en de onderzoekscriteria inzichtelijk gemaakt. Het is mogelijk dat een eis of wens een bijdrage levert aan één of meerdere onderzoekscriteria.

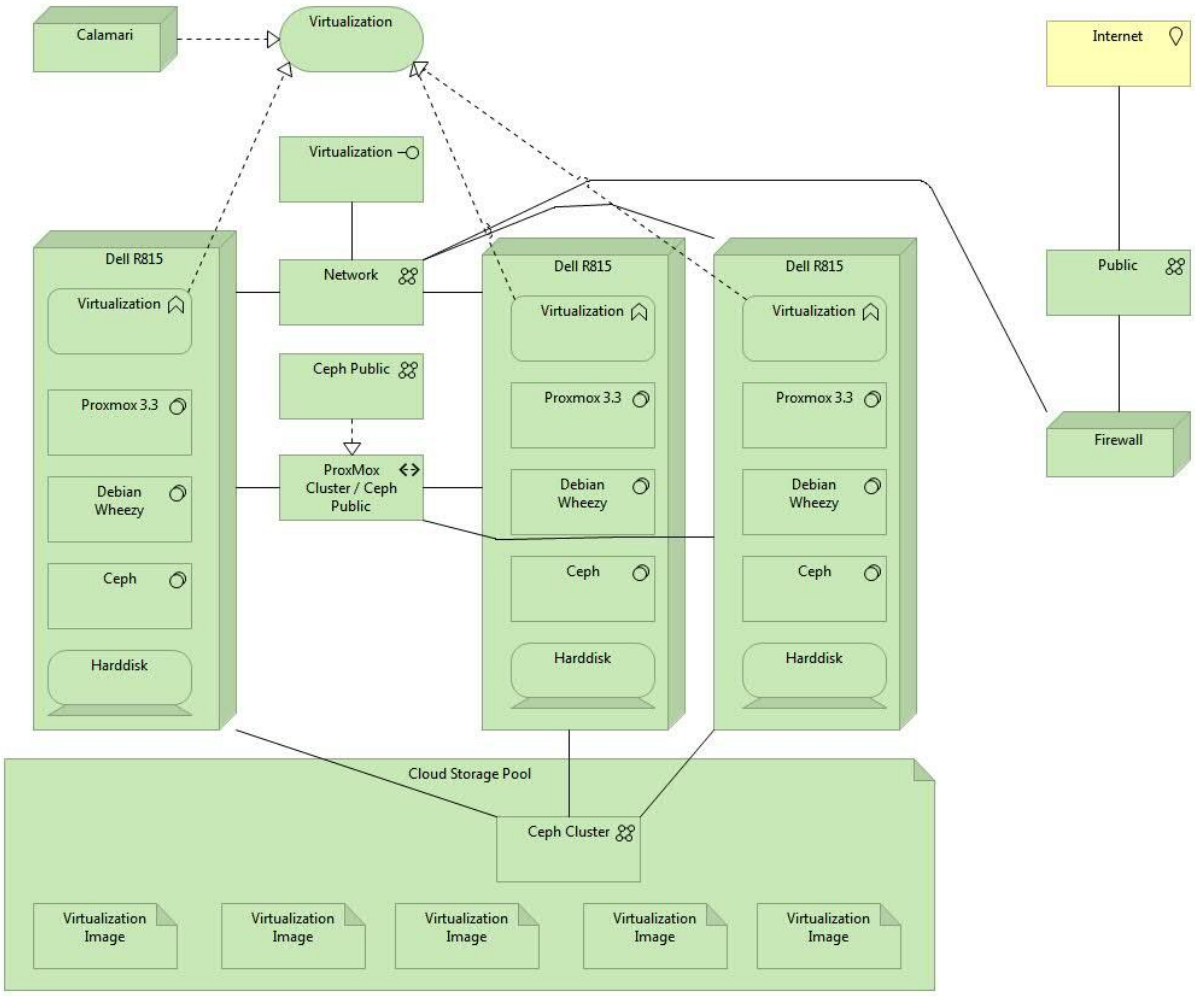
	schaalbaar	beschikbaar	security	beheerbaar	performance	kosten
Het aanbieden van virtualisatieimages		X				
Proxmox dient de storage te ondersteunen				X		
Private cloud storage			X			
Block storage	X			X		
Filesystem storage	X					
Object storage	X					
Open Source Software (OSS)						X
Performance					X	
Opslagcapaciteit	X			X		X
Horizontaal schaalbaar	X	X		X	X	X
Minimaal 2x maal opgeslagen	X	X		X		
Gescheiden (storage) netwerk			X		X	
Integriteit van de data			X	X		
Interactie met de cloudservices	X					
Zeer beperkte investeringen	X					X
Grafische interface				X		
Rolling upgrades		X				
Snapshotting				X		
Deduplication				X		
Data encryption			X			
Performance degradatie bij replicatie		X			X	
Geen Fibre Channel	X					X

Tabel 7 Relatie tot onderzoeksaspecten

5 SOLL-situatie

5.1 Architectuur

Aan de hand van de eisen en wensen van de opdrachtgever kan op basis van de ArchiMate taal de volgende architectuurtekening tot stand worden gebracht vanuit het infrastructure viewpoint.



Figuur 15 ArchiMate tekening SOLL situatie

Schaalbaarheid

Binnen de ArchiMate tekening is een derde node toegevoegd, het is mogelijk additionele nodes toe te voegen aan het Ceph Public Network en het Ceph Cluster Network. Het toevoegen van nodes en daarmee impliciet het toevoegen van het device harddisk zorgt voor schaling van de Cloud Storage Pool. Hier mee wordt de opslagcapaciteit vergroot.

Daarnaast is het mogelijk het software component ProxMox 3.3 van de toegevoegde node deel uit te laten maken van het ProxMox cluster. Hiermee wordt de reken-/virualisatiecapaciteit vergroot.

Beschikbaarheid
Het toevoegen van additionele nodes aan het beide Ceph netwerken vergroot de beschikbaarheid van de systemen. De individuele Ceph netwerken zijn opgebouwd uit meerdere verbindingen op basis van LACP het is mogelijk één of meerdere van deze verbindingen te verwijderen, het Ceph cluster blijft functioneren, naar verwachting treedt er wel performance degradatie op.

Security
Naast de voor de hand liggende toegevoegde beveiligingstechnische waarde van de firewall is er een duidelijk scheiding gemaakt tussen netwerken. Ceph diensten zijn niet beschikbaar op het routeerbare netwerk, dit zorgt voor een kleine attack vector voor Ceph. Het Ceph software component maakt gebruik van Cephx en voorkomt dat ongeauthorizeerde nodes geen toegang hebben tot de Ceph storage, zelfs waneer er toegang is tot de Ceph netwerken.

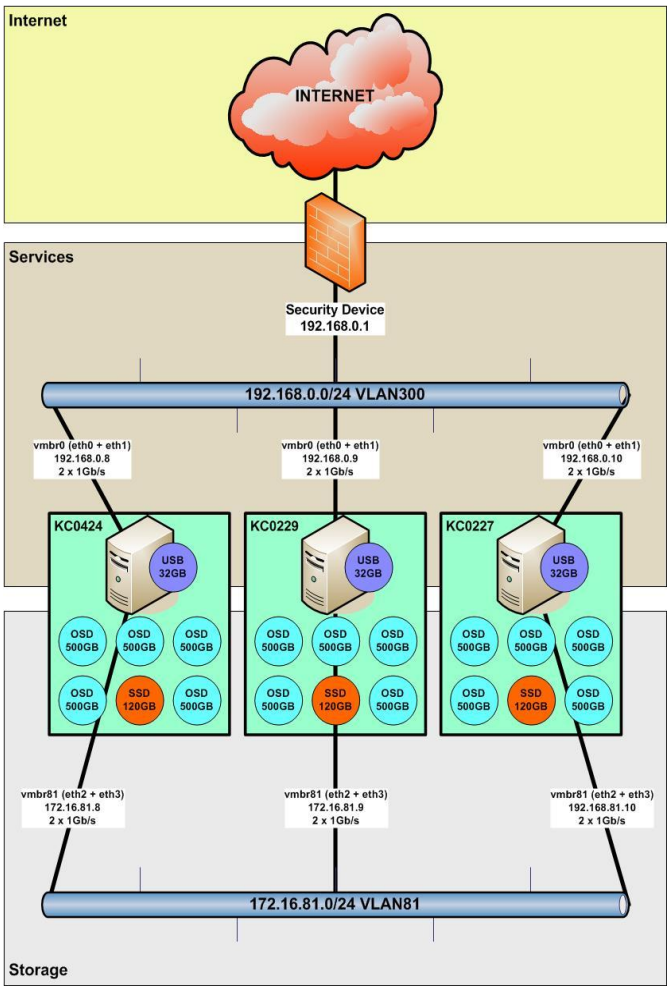
Beheerbaarheid
De gevirtualiseerde calamari host geeft visueel de status van de Ceph storage weer. Dit maakt het mogelijk, ook voor een leek problemen binnen het Ceph cluster te identificeren.

Performance
Het aanmaken van van gescheide netwerken voor de initiele schrijfactie en de replicatie zorgt voor een verbetering in performance. Het upgraden van het netwerk met LACP verdubbeld de performance. Het toevoegen van individuele harddisks en solid state drives zorgt eveneens voor een betere performance.

Kosten
In deze tekening is een additionele node toegevoegd, hier zijn kosten aan verbonden, de kosten van een additionele server lagen bij aanschaf rond de €7000,-. De het grootste verschil in kosten komt voort uit het verwijderen van het device IBM DS4200 disk enclosure (IST-situatie). De kosten van het IBM DS4200 disk enclosure lagen bij aanschaf rond de €35000,-.

5.2 Hoofdlijnen

De SOLL-situatie komt op de volgende punten overeen met de IST-situatie. De omgeving is verbonden met het internet en afgeschermd door een stateful firewall. Er is een Proxmox virtualisatie-cluster die services levert op het services netwerk.



Figuur 16 Proxmox Virtualisatie SOLL-situatie

Het grote verschil zit in het ontbreken van de IBM DS4200 storage. De disk enclosure is vervangen door Ceph cloud storage. Dit heeft de volgende invloeden op de tekening van de Proxmox Virtualisatie SOLL-situatie. De DS4200 is vervangen door een storage netwerk en de Dell R815 servers maken geen gebruik meer van hun interne storage voor het operatingsystem. Het operatingsystem van Dell R815 servers is geplaatst op een USB stick, dit maakt het mogelijk (2x)500GB per server extra ter beschikking te stellen via Ceph. De individuele harddisks worden als Ceph Object Storage Daemon (OSD) aangeboden. Daarnaast zijn de Dell R815 servers voorzien van een solid state disk (zie hoofdstuk 2.3). De lage latency van de solid state maakt deze geschikt voor het gebruik als opslagcapaciteit voor initiële schrijfacties. De solid state disks zijn dan ook ingezet als journals (paragraaf 2.7) Ceph zorgt zelf voor de verdere afhandeling van de initieel geschreven data richting de traditionele harddisks. De traditionele disk uit het slot waar de solid state in is geplaatst is toegevoegd aan de spare voorraad van de innovatieafdeling. Er is een derde server toegevoegd omdat Ceph standaard 3 kopieën op slaat. In overleg met de opdrachtgever is gekozen om deze standaard Ceph instelling te gebruiken binnen het onderzoek.

Alle binnen de SOLL gebruikte hardware was aanwezig in de spare voorraad met als uitzondering de USB sticks. Deze zijn middels het reguliere (zelf)verwervings- en declaratietraject aangeschaft.
Tijdens de technische realisatie hebben enkele wijzigingen plaats gevonden op deze SOLL-situatie. Voortschrijdend inzicht heeft een invloed gehad op met name het netwerktechnisch vlak. De uiteindelijke tekening van de Proxmox Virtualisatie SOLL-situatie is terug te vinden in bijlage BB.

6 Technische realisatie

Dit hoofdstuk geeft inzicht in hoe naast de IST situatie de SOLL situatie is gerealiseerd.

6.1 IST-SOLL en roadmap volgens ArchiMate

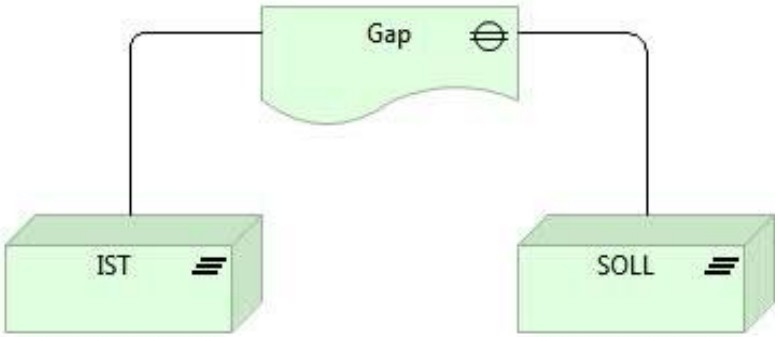
ArchiMate maakt naast de standaard Layer concepts ook gebruik van Implementation and Migration concepts (Van Haren Publishing, 2013, pp. 167-174). Twee van deze concepts passen perfect bij de technische realisatie van PoC. Hoewel er geen daadwerkelijke migratie plaatsvindt in de IST–situatie, is het toch mogelijk het verschil tussen de IST- (Figuur 14) en SOLL- (Figuur 16) situatie in ArchiMate weer te geven. Gebruikmakende van de concepten Plateau en Gap is het mogelijk de migratie te visualiseren.

Figuur 17: ArchiMate Plateau

"A plateau is defined as a relatively stable state of the architecture that exists during a limited period of time."

"A gap is defined as an outcome of a gap analysis between two plateaus."

Figuur 18: ArchiMate Gap



Figuur 19: ArchiMate Migration

De Gap kan als volgt worden omschreven: de verschillen tussen de IST- (Figuur 14) en SOLL-situatie (Figuur 16) bestaan uit het toevoegen van een additioneel netwerk ten behoeve van Ceph en het samenvoegen van het Proxmox Cluster netwerk en het Ceph public netwerk. Daarnaast wordt er een derde host geplaatst.

6.2 Reserveren (bestaande) hardware

Voor deze PoC worden 4 Dell PowerEdge R815 servers gereserveerd. Deze 2U rackmountable servers zijn voorzien van 4 8-core AMD Opteron 6100 Processor met 512K L2 cache per core en 16MB L3 Cache. De 32 geheugensloten zijn gevuld met 8GB modules die samen 256GB werkgeheugen leveren.

Figuur 20: Dell R815 Server

De machines zijn voorzien van een PERC H200 (6Gb/s) diskcontroller, zonder enige vorm van cache. Er kunnen maximaal 6 stuks 2.5” hot-plug SATA SSD, SAS, nearline SAS, of SATA drives worden geplaatst. In theorie kan er maximaal 6TB worden aangeboden met deze opstelling, maar in de praktijk zijn de sloten gevuld met 5x 500GB en 1x 120GB solid state. De volledige specificaties zijn terug te vinden in Bijlage O.

De machines zijn voorzien van 8 1Gb/s netwerkinterfaces die verdeeld zijn over twee adapterkaarten. Naast de Dell R815 servers zijn twee Cisco switches gereserveerd.

6.3 Verwerven hardware

Om de opslagcapaciteit van de server tegen minimale kosten te verhogen is er gekozen voor het plaatsen van een USB-stick in de Dell R815 Servers. Op deze USB-stick wordt het operating system geïnstalleerd. Het installeren van het operatingsystem heeft als voordeel dat er geen interne harddisk wordt gebruikt. De harddisk kan aangeboden worden als Ceph storage, er kunnen op deze manier 4 in plaats van 5 disks worden aangeboden, hiermee wordt de totale opslagcapaciteit verhoogd met 25%. De kosten voor deze toename in capaciteit bedragen € 28,- per server.

Figuur 21: SanDisk 32GB Extreme USB 3.0 Flash Drive

Er is gekozen voor de SanDisk 32GB Extreme USB 3.0 Flash Drive vanwege performance en levenslange garantie. De kosten van 4 USB-sticks, inclusief verzendkosten bedragen: € 112,50. Het intern installeren van de USB verlaagt de diefstalgevoeligheid van de USB-sticks. Gezien de redundantie welke aanwezig is binnen Ceph, is uitval van de USB-stick en daarmee de gehele server, geen probleem. Dit is ook de reden waarom er over 1 stick(disk) wordt gesproken, in de IST-situatie waren 2 500GB disks met elkaar gemirrord.

6.4 Scrum

In totaal zijn er drie **user stories** gedefinieerd binnen dit project. Gezien het kleine aantal user stories is er gekozen om iedere user story toe te kennen aan een individuele sprint.

6.4.1 Product backlog

Story 1: "Als techneut wil ik een infrastructuur zodat ik cloud software kan installeren en configureren."

De infrastructuur uit de story betreft de infrastructuur voor de PoC. Deze infrastructuur bestaat uit de volgende componenten:

- 4 x 19" 2U Dell R815 servers
- 2 x 19" 1,5U Cisco Catalyst 3750G switches

De story bestaat uit de volgende **tasks**:

- De servers dienen te worden geplaatst in een 19" rack.
- Het aansluiten van de iDrac (lights out management) interfaces.
- Het aansluiten van de servers op het bestaande services netwerk.
- Het aanmaken van DNS records.
- Er dienen preseed files aangemaakt te worden voor netwerkinstallatie.
- De servers dienen te worden voorzien van een operatingsystem.
    - Installatie dient plaats te vinden op de interne USB stick.
- De switches dienen te worden geplaatst in een 19" rack.
- De switches dienen als stack te worden geconfigureerd.
- De switches dienen te worden geconfigureerd als storage netwerk.
    - De switches dienen te worden voorzien van management op basis van SSH.
    - De switches dienen te worden voorzien van VLAN's.
    - De servers dienen netwerktechnisch met elkaar te kunnen communiceren.
    - De switches dienen LACP aan te bieden op het public en cluster VLAN.
    - Het testen van de netwerkperformance.
- "Configuration Items" dienen aangepast te worden in de CMDB.
- De fysieke locatie van de servers dient aangepast te worden in de applicatie Rackmonkey.

De **Definition of Done** van deze story luidt als volgt:
"De servers en switches dienen in beheer te zijn. Het moet voor de servers mogelijk zijn met elkaar te communiceren op zowel het services- als het storagenetwerk. Alle aanpassingen dienen te worden geadministreerd."

De definition of done wordt gecontroleerd op basis van de drie laatste genoemde taken:

- Het testen van de netwerkperformance.
- "Configuration Items" dienen aangepast te worden in de CMDB.
- De fysieke locatie van de servers dient aangepast te worden in de applicatie Rackmonkey.

Story 2: “Als techneut wil ik een werkende Proxmox virtualisatieomgeving met cloud storage om testen op uit te voeren”

De story bestaat uit de volgende **tasks**:

- Mirroren van Ceph repositories
- Installatie van de Ceph software
- Het creëren van een Ceph storage cluster
- Partitioneren van de disks
- Aanmaken OSDs
- Het aanmaken van een pool
- Het installeren van Proxmox
- Het creëren van een Proxmox cluster
- Het toevoegen van de Ceph storage aan Proxmox
- Het creëren van een virtuele benchmark host op Proxmox
    - Het installeren van de benchmark suite
    - Het installeren van de benchmark software
    - Het testen van de benchmark software
- Het compileren van de Calamari software
- Het compileren van Diamond*
- Het realiseren van een repository*
- Het installeren van Diamond*
- Het installeren van de Calamari client*
- Het installeren van Calamari server*

* Deze werkzaamheden realiseren een beheerinterface, welke noodzakelijk is voor het onderzoek naar de beheerbaarheid van de storage, zie hoofdstuk 7.

De **Definition of Done** van deze story luidt als volgt:

“Een functionerende virtualisatieomgeving met cloud storage waarop het onderzoek uitgevoerd kan worden.”

De definition of done wordt gecontroleerd op basis van de test:

- Het testen van de benchmark software
  Indien het mogelijk is een benchmark uit te voeren op een op Ceph storage gevirtualiseerde server, kunnen we aannemen dat de Ceph storage functioneert.

Story 3: “Als opdrachtgever wil ik een reproduceerbare inrichting van de cloud storage om mijn bedrijfsvoering toekomstvast te maken.”

De story bestaat uit de volgende **tasks**:

- Het schrijven van scripts voor Ansible (2.6) voor alle technische taken uit sprint 1 en 2
    - De volgende rollen dienen te worden gedefinieerd
        - Common
        - Debian
        - Proxmox
        - Ceph
- Het testen van de scripts voor Ansible.
- Het uitvoeren van herinstallaties.

De **Definition of Done** van deze story luidt als volgt:

“Een geautomatiseerde installatie van een Proxmox virtualisatiecluster met Ceph cloud storage.”

De definition of done wordt gecontroleerd op basis van de test:

- Het uitvoeren van een geautomatiseerde installatie.

6.4.2 Sprint 1 – Story 1

De commando's welke gebruikt zijn tijdens deze sprint zijn terug te vinden in bijlage D en bijlage E. Deze bijlagen behoren toe aan sprint 3, waarin de werkzaamheden van deze sprint zijn geautomatiseerd.

- **De servers dienen te worden geplaatst in een 19" rack.**

De voor de PoC geselecteerde hardware moest verplaatst worden naar een 19" rack met internetverbinding. Bij het afstudeerbedrijf is een duidelijke scheiding tussen het interne en het externe netwerk. Het was dus noodzakelijk de bestaande hardware uit verscheiden 19" racks te verwijderen en indien mogelijk gezamenlijk in een 19" rack te plaatsen. Aan de hand van informatie uit het Rackmonkey[46] 19" rack managementsysteem kon zowel de huidige fysieke locatie van de componenten als de vrije ruimte in de internetkasten worden afgeleid. Aan de hand hiervan werd de hardware verplaatst.

- **De aansluiten van de iDRAC (lights out management) interfaces.**

De geplaatste servers dienen in beheer genomen te worden. Hiervoor is in de Dell R815 een iDRAC[47] (integrated Dell Remote Access Controller) interface beschikbaar. Een iDRAC interface biedt de mogelijkheid via een web interface de server te benaderen, zelfs wanneer deze server uit staat. De interface maakt het dan ook mogelijk om op afstand via de web interface de server op te starten en te installeren via een console.
Het console is te vergelijken met een toetsenbord en beeldscherm combinatie over het netwerk.

Impediments:

Bij het aansluiten van de iDRAC interfaces en het benaderen van de web interfaces werd duidelijk dat niet alle Dell R815 servers dezelfde versie firmware geïnstalleerd hadden staan. De verschillen in firmware kunnen zorgen voor verschillen in performance en gedrag.

Om eenduidige testen uit te kunnen voeren zijn alle firmware versies gelijkgetrokken naar versie: 1.98 (build 8). Aan de hand van deze firmware werden alle BIOS'en versie 2.8.2. Hiermee waren alle binnen deze PoC gebruikte servers identiek qua hardware en firmware. Aangezien het afstudeerbedrijf geen supportcontract heeft was het niet mogelijk alle servers gelijktijdig te voorzien van deze software via een centraal management platform. De software moest gedownload worden en de firmware update moest geëxtraheerd worden uit de gedownloade software.

[46] https://flux.org.uk/projects/rackmonkey/
[47] http://www.dell.com/learn/us/en/555/solutions/integrated-dell-remote-access-controller-idrac

- **Het aansluiten van de servers op het bestaande services netwerk.**

Om deze servers te kunnen installeren en na installatie virtualisatiediensten aan te kunnen bieden is het noodzakelijk de Dell servers te verbinden met het bestaande netwerk voor services. Deze werkzaamheden zijn regulier en leverde dan ook geen problemen op.
Op de switch werden trunk poorten aangemaakt welke werden voorzien van meerdere VLAN's welke gebruikt worden voor virtualisatie. Daarnaast werd een native VLAN toegevoegd om installatie van de servers mogelijk te maken. De aan deze taak gerelateerde Cisco switch configuratie is terug te vinden in bijlage N.

- **Het aanmaken van DNS records.**

Om de diensten te benaderen zijn DNS records aangemaakt. Alle hosts werden voorzien van een PRT en A DNS record met de volgende syntax:
{hardwarenaam}.{domein}.{top level domein}

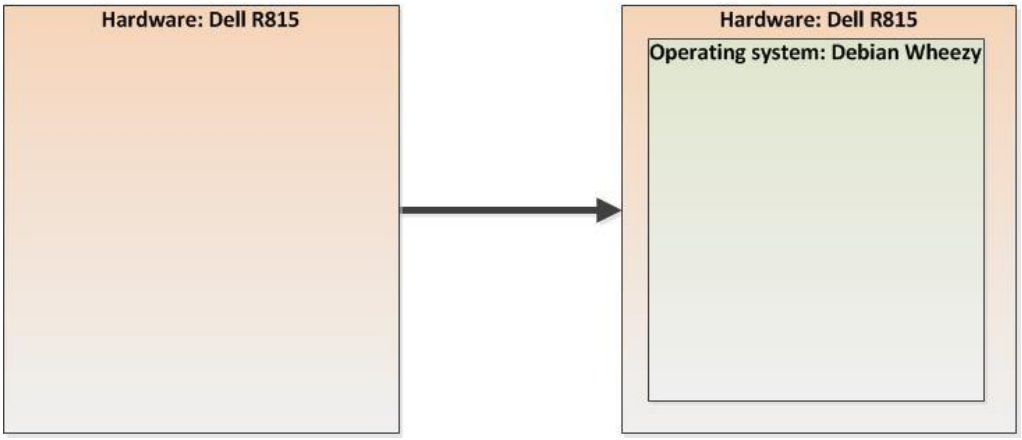
Voor het netwerk waar Ceph gebruik van gaat maken zijn ook DNS records aangemaakt. Hier hebben de A-records echter de volgende syntax:
{hardwarenaam}-CEPH.{domein}.{top level domein}

Impediments:

In sprint 2 bleek Ceph beter ingericht te kunnen worden met 2 dedicated netwerken. Één public netwerk en een cluster netwerk. Het cluster netwerk is bedoeld voor OSD replicatie en het public netwerk voor de monitor hosts en client-server communicatie.
De DNS registratie voor de records met de syntax *{hardwarenaam}-CEPH.{domein}.{top level domain}* zijn dan ook in een later stadium ongedaan gemaakt. Bij de correcte inrichting van de Ceph netwerken is gebruik gemaakt van respectievelijk de volgende syntax:
{hardwarenaam}-MON.{domein}.{top level domein}
{hardwarenaam}-OSD.{domein}.{top level domein}

- **Er dienen preseed files aangemaakt te worden voor netwerkinstallatie.**

Voor de installatie van het OS wordt binnen het afstudeerbedrijf gebruik gemaakt van zogenaamde preseed files. Deze bestanden bevatten een aantal sleutelwoorden die keuzes binnen het installatieproces vertegenwoordigen. De installatiesoftware leest het bestand in bij aanvang van de installatie en voert de installatie zonder menselijke interactie uit.



Figuur 22 Operating system installatie op Proxmox virtualisatiehosts

Er bestond reeds een preseed file voor een virtualisatiehost, deze preseed file is dan ook gebruikt om de installatie te conformeren aan de huidige inrichting van de Proxmox virtualisatiehost. Het gebruik van de identieke preseed file zorgt voor identieke software op de virtualisatiehosts. Identieke software maakt het mogelijk meetresultaten van gelijke metingen één op één met elkaar te vergelijken. Er was één kleine aanpassing noodzakelijk voor de installatie op USB i.p.v. op de harddisk. Deze aanpassing is niet als impediment gemeld omdat deze inherent is aan de installatie op USB. In bijlage C is de gebruikte preseed file terug te vinden.

- **De servers dienen te worden voorzien van een operating system.**

Gebruikmakende van de eerder aangesloten iDRAC interface, het service netwerk en het native VLAN worden de servers voorzien van een Linux installatie. Via de iDRAC interface werd een console gestart, hierna werd de server aangezet via de iDRAC interface. Tijdens het booten van de server werd gekozen voor netwerkinstallatie via PXE. Er verscheen een menustructuur op de console waar operating systems voor installatie kon worden gekozen. Het juiste operatingsystem voor het Proxmox installatieplatform werd gekozen en de betreffende server werd binnen 10 minuten geïnstalleerd. Aan het eind van de installatie werd de server gereboot en was deze via het netwerk bereikbaar middels SSH.

- **De switches dienen te worden geplaatst in een 19” rack.**

Om het netwerk waar normaliter services worden aangeboden niet te verstoren met het netwerkverkeer van de Ceph storage worden dedicated switches geplaatst. Het betreft twee Cisco 3750G switches welke uiteindelijk in een stack opstelling worden geplaatst. Aan de hand van informatie uit het Rackmonkey 19” rack managementsysteem kon zowel de huidige fysieke locatie van de componenten als de vrije ruimte in de internetkasten worden afgeleid. Aan de hand hiervan werd de hardware verplaatst.

- **De switches dienen als stack te worden geconfigureerd.**

Om het Ceph storagenetwerk vanuit het perspectief van de servers redundant uit te voeren zijn twee switches geplaatst, deze switches hebben echter geen kennis van elkaar. Er is gekozen gebruik te maken van een stack cable om de switches aan elkaar te verbinden. Het maken van een stack was niet per definitie noodzakelijk voor Ceph, met het oog op de toekomst zijn deze switches geplaatst

als breder inzetbaar storage backend. Naast de mogelijkheid tot redundantie zorgt de stack voor een verdubbeling van het aantal beschikbare netwerkpoorten.

- **De switches dienen te worden geconfigureerd als storage netwerk.**

    - De switches dienen te worden voorzien van management op basis van SSH.

    De standaard van het afstudeerbedrijf voor remote management van netwerkcomponenten (en unices) vindt plaats op basis van SSH (Secure SHell). Het is mogelijk op afstand instellingen door te voeren op netwerkcomponenten via de versleutelde verbinding. Om remote management mogelijk te maken dienen onderstaande commando's uitgevoerd te worden op het netwerkcomponenten.

```
line vty 0 4
 exec-timeout 15 0
 transport input ssh
```

    Daarnaast moet de switch voorzien zijn van een IP-adres en één of meerdere gebruikers-accounts.

    - De switches dienen te worden voorzien van VLAN's.

    Voor het scheiden van de twee storagesegmenten is het mogelijk slechts gebruik te maken van twee aparte subnets. Communicatie zou plaatsvinden binnen het juiste subnet en het geheel zou zonder problemen functioneren. Om zeker te zijn van scheiding, het effect van menselijke fouten te minimaliseren en ten aanzien van de beveiliging is gekozen voor het aanmaken van VLAN's. Op de storage switches zijn dan ook twee VLAN's aangemaakt VLAN 81[48] ten behoeve van OSD replicatie en VLAN 83 ten behoeve van het monitor netwerk en het wegschrijven van de initiële kopie.

[48] Vlan 81 komt voort uit een impediment in sprint 2

```
VLAN  Name                             Status    Ports
---- -------------------------------- --------- -------------------------------
81    Storage-A                        active    Po1, Po9, Po17
83    ceph-mon                         active    Po2, Po10, Po18
```

Bovenstaande output is van een de Cisco storage switch. De output toont de indeling van de port channels in VLAN's.

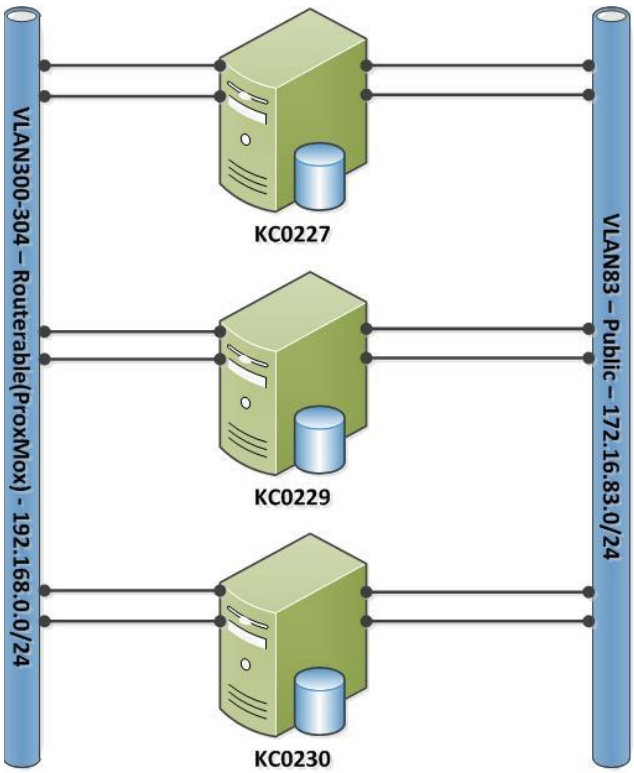
- De servers dienen netwerktechnisch met elkaar te kunnen communiceren

Na het aanmaken van de VLAN's moet het nog steeds mogelijk zijn de hosts met elkaar te laten communiceren.

Dit is getest door met het commando ping ICMP echo pakketten uit te zenden en de response hiervan op te vangen.

Op basis van deze basale test is vastgesteld dat alle servers op de netwerken; 192.168.0.0/24, 172.16.81.0/24,172.16.83.0/24 met elkaar kunnen communi-ceren.

Vlan 81 is later toegevoegd aan de hand van een impediment uit sprint 2. Voor de definitieve VLAN inrichting zie Figuur 31.



Figuur 23 Netwerkverbindingen

- De switches dienen LACP aan te bieden op het storage VLAN.

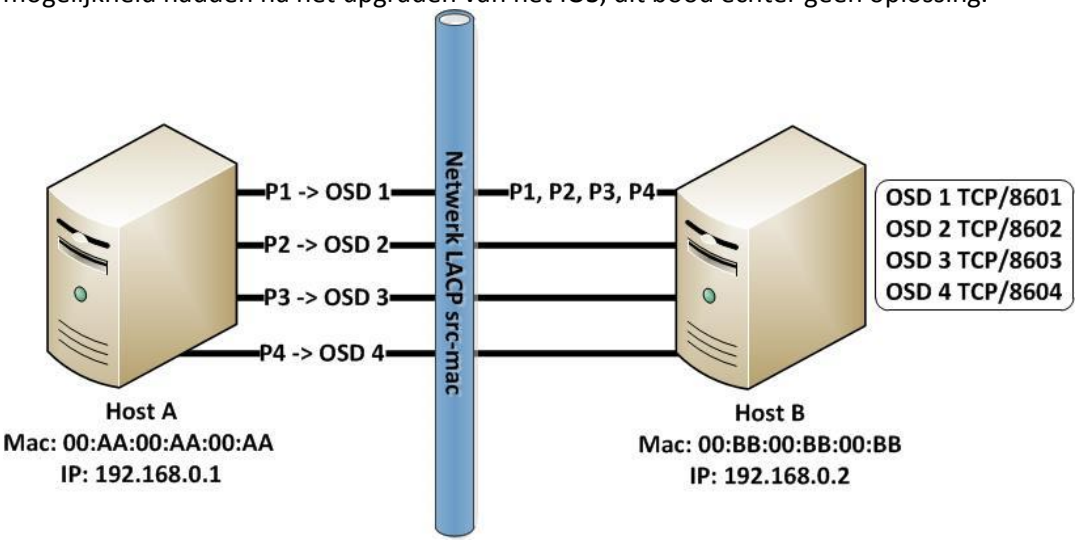
Om afdoende performance te kunnen bieden met de bestaande hardware op de storage netwerken was het noodzakelijk meerdere 1Gb/s verbindingen van de Dell R815 servers te koppelen op basis van het Link Aggregation Control Protocol (LACP). De Cisco configuratie met betrekking tot LACP is terug te vinden in bijlage N.

Om LACP naar behoren te laten functioneren is het noodzakelijk gebruik te maken van het src-dst-port load-balancing algoritme. Dit zorgt dat verkeer van verschillende source poorten en verschillende destination TCP/UDP poorten over verschillende netwerkinterfaces afgehandeld wordt. Dit protocol is het meest efficiënt aangezien iedere OSD een eigen TCP poortnummer heeft. Bij het standaard algoritme src-mac is het met 3 nodes slechts mogelijk met maximaal 2Gb/s te communiceren, ook wanneer het een 4Gb/s LACP port-channel betreft.

```
port-channel load-balance src-dst-port
```

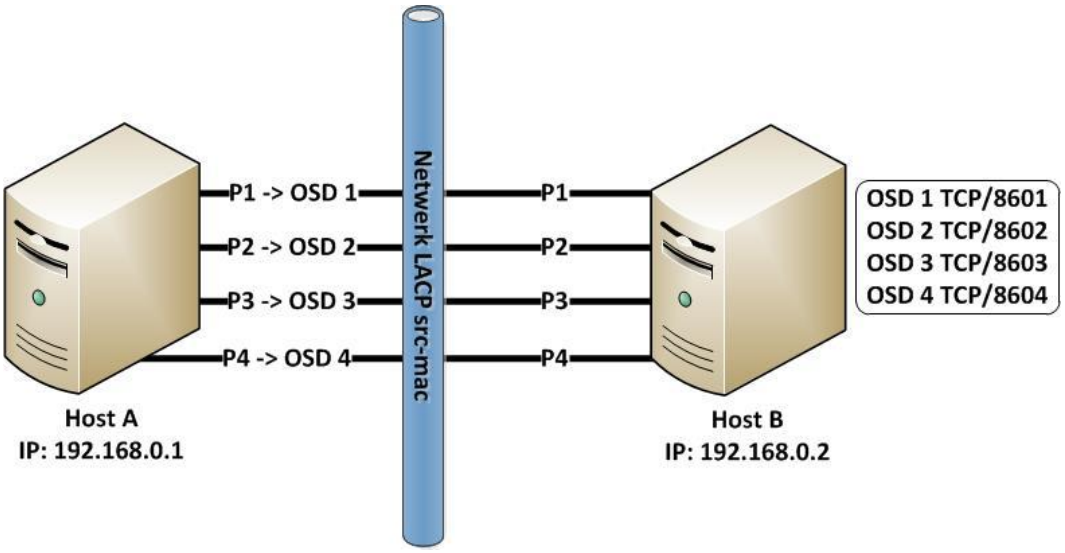
Impediment:

Op het punt van de port-channel load-balancing liep het stuk dat de gebruikte gestackte switches het src-dst-port algoritme niet ondersteunde. Er is onderzocht of de switches de mogelijkheid hadden na het upgraden van het IOS, dit bood echter geen oplossing.



Figuur 24 schematische weergave van LACP met het src-mac algoritme

Uiteindelijk moest er een moderner type switch worden geplaatst om dit algoritme te ondersteunen. Deze switch moest geupdatet worden en geplaatst in het rack. Hierna moest de switch worden geconfigureerd en voorzien van het juiste load-balancing algoritmen. Daarnaast moesten alle kabels omgeprikt worden naar de nieuwe switch. In de bijlage is te zien welke algortimes ondersteund worden door de respectievelijke switches. Zie bijlage R. Deze impediment heeft in totaal 8 uur vertraging opgeleverd.



Figuur 25 Schematische weergave van LACP met het src-dst-port algoritme

Uiteindelijk is gebruik gemaakt van een enkelvoudig uitgevoerde Cisco switch van het model 4948.

- Het testen van de netwerkperformance.

Om het eerder ingestelde src-dst-port algorimte naar behoren te kunnen testen is gebruik gemaakt van de tool iperf[49]. Iperf is een client-server applicatie welke specifiek ontwikkeld is voor het testen van netwerkbelasting.

Op de Dell R815 server die de rol van iperf server op zich nam werden evenveel serverprocessen opgestart als dat er netwerkinterfaces beschikbaar waren in het betreffende netwerk. Ieder serverproces werd met een apart poortnummer gestart om optimaal gebruik te kunnen maken van het src-dst-port algoritme.

Voorbeeld van het starten van de server processen

```
root@KC0229:/tmp# iperf -s -p 2048 -B 172.16.83.9 &
root@KC0229:/tmp# iperf -s -p 2049 -B 172.16.83.9 &
```

Vervolgens werden er gelijktijdig evenredig veel clientprocessen gestart op een van de Dell R815 servers die de rol van iperf client vervulden.

Voorbeeld van het starten van de clientprocessen en de meetresultaten

```
root@KC0230:/tmp# iperf -c 172.16.83.9 -p 2048 & iperf -c 172.16.83.9 -p 2049
[ ID] Interval Transfer Bandwidth
[ 3] 0.0-10.0 sec 1.10 GBytes 942 Mbits/sec
[ ID] Interval Transfer Bandwidth
[ 3] 0.0-10.0 sec 1.10 GBytes 942 Mbits/sec
```

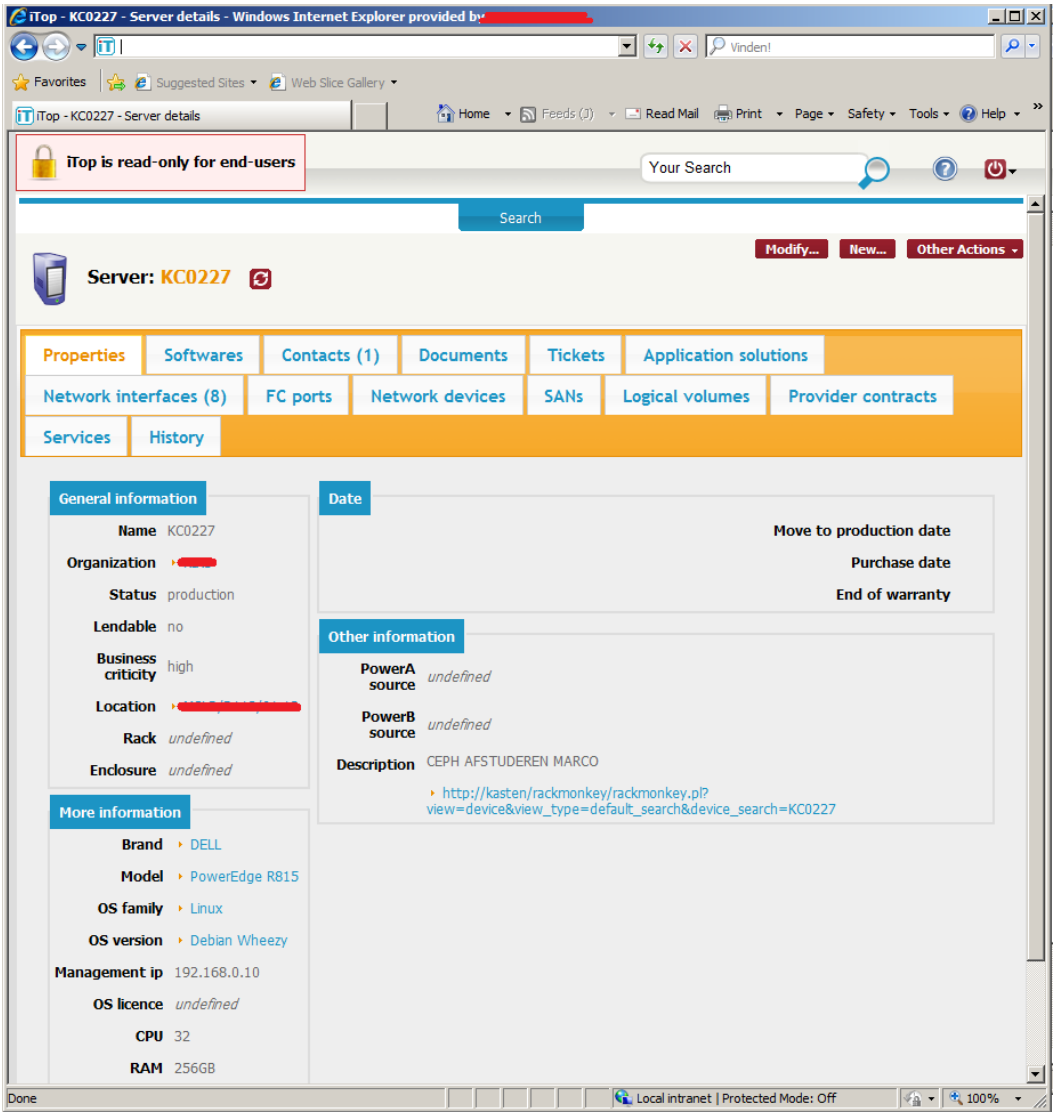
Uit de meetresultaten is af te leiden dat de netwerk bandwidth dicht bij het theoretisch maximum van 1Gb/s komt. Hieruit kunnen we opmaken dat het 2Gb/s netwerk en daarmee het LACP src-dst-port algoritme naar behoren funtioneert.

[49] https://iperf.fr/

Deze test is herhaald op alle netwerken binnen deze PoC. Alle netwerken functioneerden naar behoren.

- **“Configuration Items (CI’s)” dienen aangepast te worden in de CMDB.**

Binnen het afstudeerbedrijf wordt gebruik gemaakt van een Configuration Management DataBase (CMDB) welke volledig Information Technology Infrastructure Library (ITIL) conform is. In het verleden heeft de afstudeerder deze CMDB geïnstalleerd en gevuld met alle CI’s van de innovatieafdeling van het afstudeerbedrijf. De CI’s die binnen het PoC gebruikt zijn, zijn opgenomen in deze CI’s. Niet alleen om te conformeren aan de door het afstudeerbedrijf gestelde standaarden, maar ook om collega’s inzicht te geven in het feit dat de server en switches gebruikt worden voor mijn afstuderen.

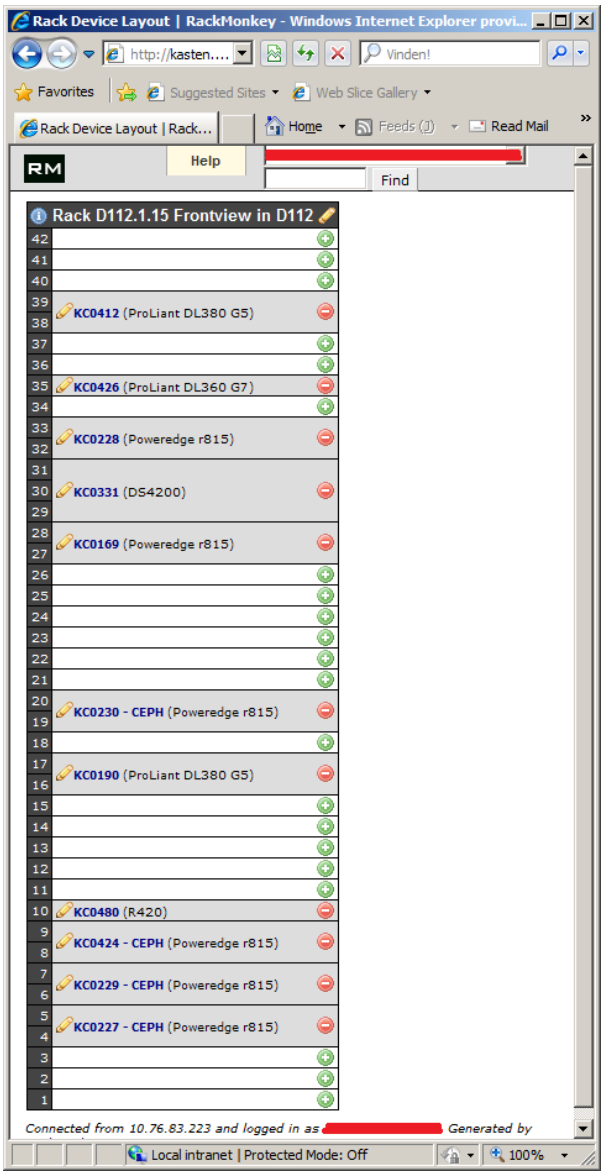


Figuur 26: Itop registratie KC0227

- **De fysieke locatie van de servers dient aangepast te worden in de applicatie Rackmonkey.**

Als aanvulling op de ITop CMDB maakt de innovatieafdeling gebruik van de tool Rackmonkey. Rackmonkey is een basaal 19” rack management systeem. Het biedt de mogelijkheid 19” racks weer te geven inclusief de daarin geplaatste servers. Op basis van de informatie in Rackmonkey kan afgeleid worden hoeveel vrije ruimte er beschikbaar is in een kast en wat de huidige fysieke locatie van een hardware-component is.

De hardwarecomponenten die gebruikt zijn binnen de PoC zijn allemaal in dit systeem opgenomen. Als aanvulling hierop is een gestandaardiseerde URL toegevoegd aan ITop, dit maakt het mogelijk vanuit ITop de fysieke locatie van het CI in 1 klik zichtbaar te maken.



Figuur 27 Frontview weergave 19”rack

6.4.3 Sprint 2 – Story 2

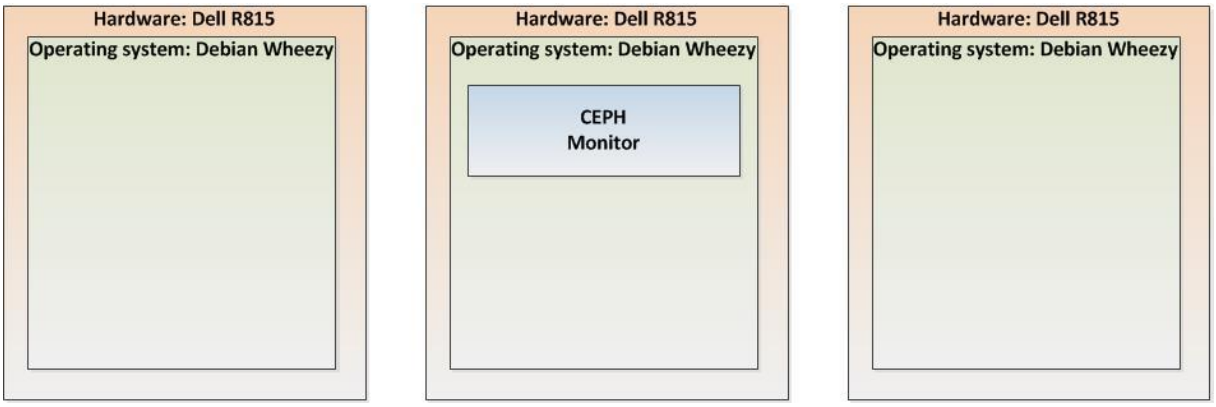
De commando's welke gebruikt zijn tijdens deze sprint zijn terug te vinden in bijlage F en bijlage G. Deze bijlagen behoren toe aan sprint 3, waarin de werkzaamheden van deze sprint zijn geautomatiseerd.

- **Mirroren van Ceph repositories**

Voor het repliceren van de Ceph repositories is gebruik gemaakt van reeds bestaande methodieken binnen het afstudeerbedrijf. De te repliceren Ceph repositories worden toe-gevoegd aan een lijst met softwarebronnen. Iedere nacht worden al deze bronnen automatische gerepliceerd. Na replicatie worden alle bronnen gecontroleerd op mogelijke virus- of malwareinfecties. Hierna is het mogelijk de Ceph repository op te voeren in de toekomstige Ceph servers. Een repositorie is een software bibliotheek.

- **Installatie van de Ceph software**

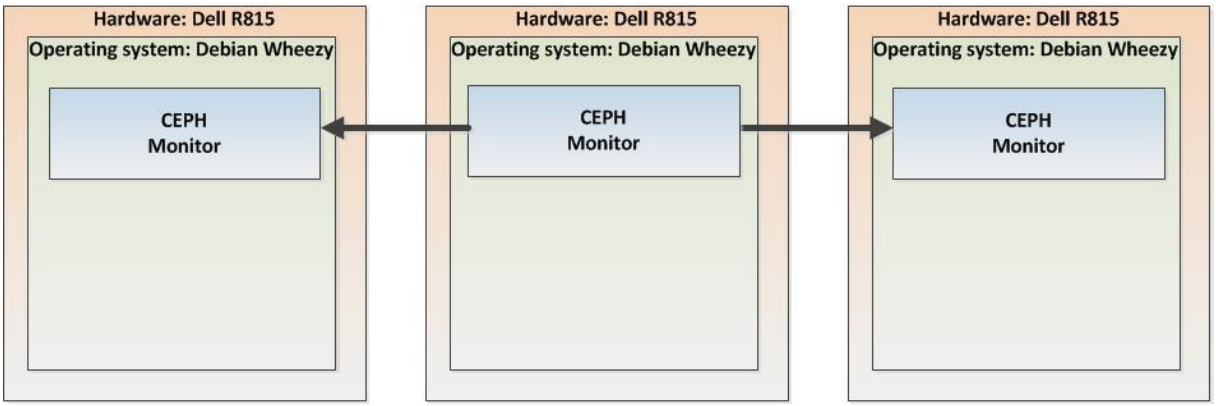
Er is gekozen om Ceph te installeren vanuit de initiële Ceph server(de eerste – maakt niet uit welke van de drie). Voor deze methode was het noodzakelijk een Ceph-gebruiker aan te maken en deze te voorzien van de nodige rechten. Op alle in te richten Ceph servers is de gebruiker Ceph dan ook aangemaakt en voorzien van de juiste sudo rechten. Sudo rechten biedt een niet root gebruiker de rechten om processen uit te voeren als de root gebruiker. Naast het toevoegen van sudo rechten was het noodzakelijk de Ceph-gebruikers onderling met elkaar te laten communiceren via SSH op basis van key authentication. Aanvullende informatie over deze handelingen is terug te vinden op: http://ceph.com/docs/master/rados/deployment/preflight-checklist/#create-a-user



Figuur 28 Ceph software op de eerste nodes

- **Het creëren van een Ceph storage cluster**

Nadat alle Ceph-gebruikers aangemaakt zijn op de Ceph servers en de daarbij behorende sudo en SSH functionaliteit is getest is het mogelijk Ceph te gaan installeren. Vanuit de initiële Ceph server wordt een initiële configuratie aangemaakt, deze wordt daarna vanuit de initiële server gedistribueerd naar de overige Ceph servers. Bij het aanmaken is het mogelijk specifieke settings voor het Ceph cluster mee te geven. Enkele voorbeelden hiervan zijn --public-network en –cluster-network.



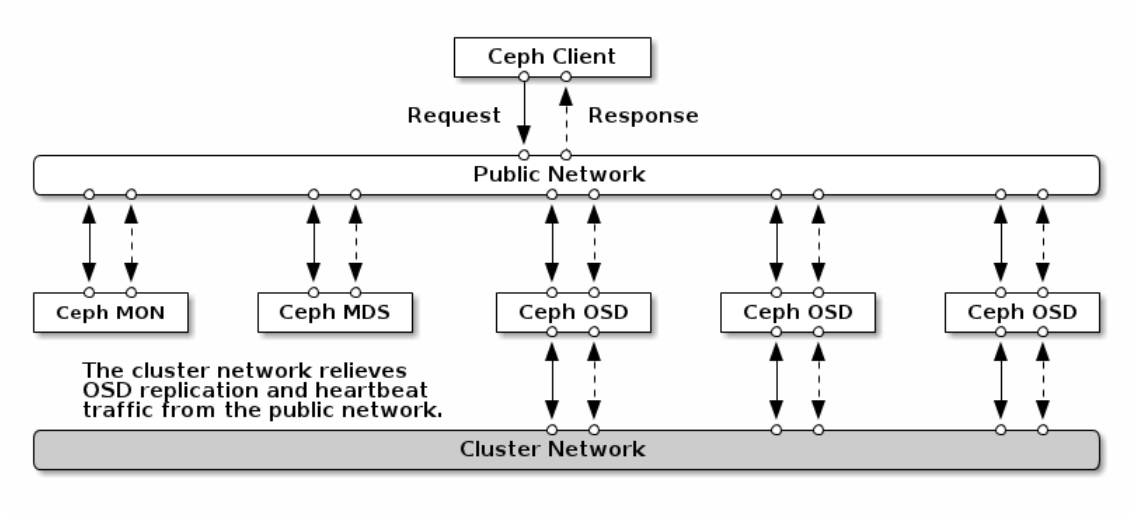
Figuur 29 Ceph cluster na distributie van Ceph software vanaf de initiële node

Impediment:

Een misinterpretatie van de term public network zorgde voor de noodzaak tot een complete redesign van het Ceph cluster. Het public network is niet het publieke routeerbare netwerk waarop communicatie met de Proxmox virtualisatie hosts plaatsvindt. Het public netwerk is het netwerk waarop de initiële write van Ceph plaatsvindt.

De impediment leverde 15 uur extra werk op, daarnaast kwam de impediment op een zaterdag aan het licht en was fysiek herbekabelen noodzakelijk.

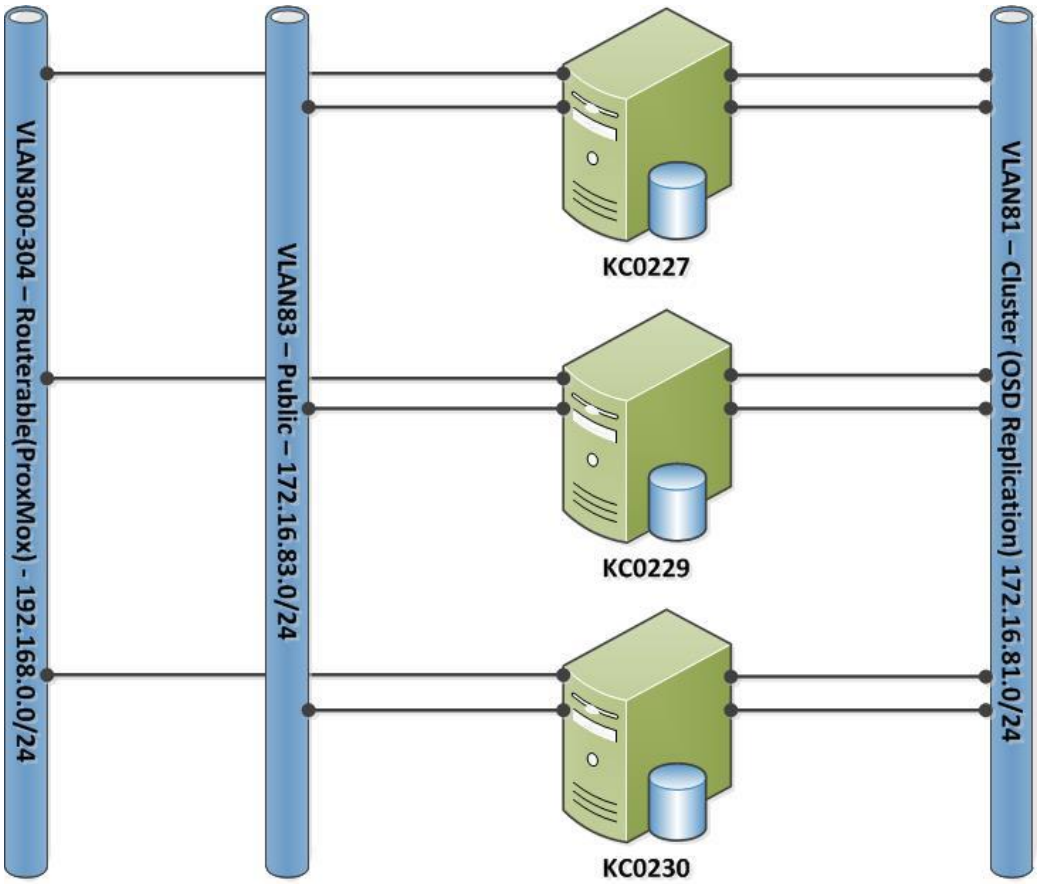
Het netwerk is publiek ten opzichte van de Ceph clients(Proxmox nodes)[50].



Figuur 30 Schematische weergave van het cluster en public network

[50] http://ceph.com/docs/master/rados/configuration/network-config-ref/

Deze impediment leverde uiteindelijk het volgende netwerk design op:



Figuur 31 Schematische weergave na netwerk redesign

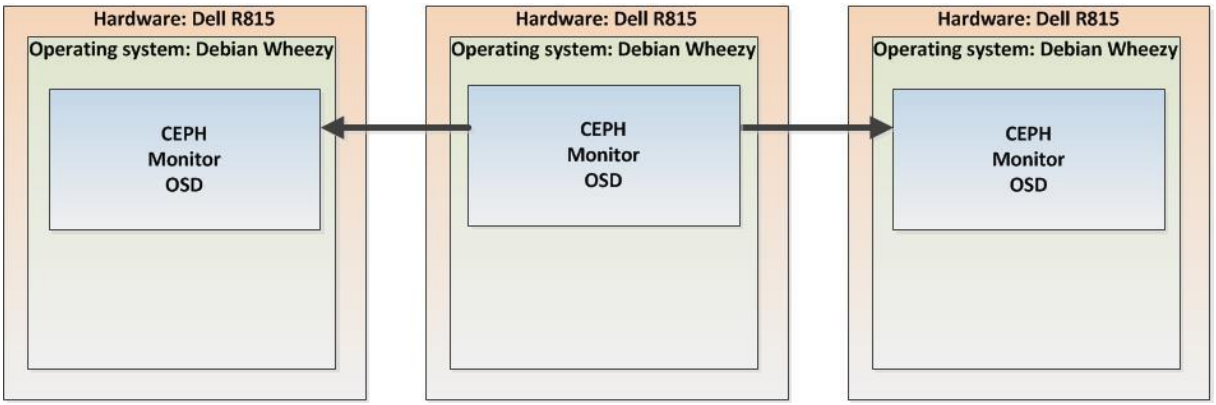
- **Partitioneren van de disks**

In de voorgaande taken uit sprint 2 is een werkend Ceph cluster gerealiseerd. Het cluster heeft echter nog geen storage beschikbaar. Om storage aan te bieden moeten zogenaamde OSD's aangemaakt worden. OSD's bestaan uit 2 delen: een stuk opslagcapaciteit en een journal. De journal neemt als eerste de schrijfacties aan, hierna wordt het weggeschreven naar de traditionele disk. Het is mogelijk de OSD en de journal samen te voegen op een enkele disk, dit komt de performance echter niet ten goede. Er is dan ook gekozen voor het gebruik van een solid state schijf om de write performance te verbeteren. Voor meer informatie over solid state schijven en waarom deze geschik zijn als journal zie: 0. Voordat het mogelijk is om een schijf als OSD aan te bieden dient deze geformatteerd te worden.

- **Aanmaken OSDs**

Na het aanmaken van de partitionering is het mogelijk een OSD aan te maken. De OSD zorgt voor de formattering van de disk. Bij het aanmaken van de OSD's is gekozen gebruik te maken van de standaard Ceph formattering op basis van XFS. Hieronder een voorbeeld van het toevoegen van een OSD. Merk hierbij op dat /dev/sdd een traditionele schijf is en /dev/sde1 een partitie van een solid state schijf is, een journal. Zie Figuur 98 Schematische weergave van fysieke disk naar OSD.

```
ceph-deploy osd create --zap-disk KC0227-OSD:/dev/sdd:/dev/sde1
```



Figuur 32 Het Ceph cluster na het aanmaken van de OSD's

Impediment:

Na een reboot was er nog maar 1 OSD beschikbaar, de rest van de OSD's was "defect" de oorzaak lag in het installeren van Proxmox. Proxmox maakt gebruik van een 2.6 kernel, na installatie van een basis OS wordt er gebruik gemaakt van kernel 3.2. Beide kernels hebben een afwijkende methodiek voor de enumeratie van hardware componenten.

OSD Kernel 3.2		OSD Kernel 2.6	
disk	**journal**	**disk**	**journal**
sdd1	sde1	sdc1	sdi1
sdf1	sde2	sdd1	-
sdg1	sde3	sde1	-
sdh1	sde4	sdf1	-
sdi1	sde5	sdh1	-

Tabel 8 Benaming disk/journal per kernel versie

Om de effecten van deze impediment te voorkomen is gebruik gemaakt van world wide numbers. In iedere harddisk is vanuit de fabrikant een uniek nummer toegekend. Onafhankelijk van de enummeratie van de harddisk blijft dit nummer altijd hetzelfde, hierdoor blijft de verhouding tussen traditionele disk en journal te allen tijde gelijk.

```
ceph-deploy osd create --zap-disk KC0227-OSD:/dev/disk/by-id/wwn-0x5000c5003456e493:/dev/disk/by-id/wwn-0x500253805001c4cb-part1
```

Deze impediment heeft totaal 8 uur vertraging veroorzaakt.

- **Het aanmaken van een pool**

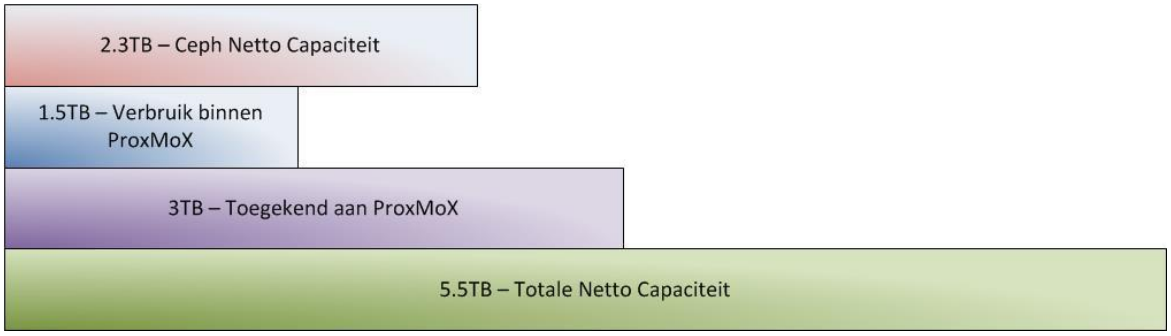
Nu alle OSD aangemaakt zijn op het Ceph cluster is het mogelijk een storage pool aan te maken. Een pool omvat alle binnen het cluster aanwezige OSD's. De totale omvang van de binnen de PoC gerealiseerde POOL is 6990MB (15*466GB) (0). De pool wordt aangemaakt met onderstaand commando. De Pool is aangemaakt van 4096 placement groups. Dit is conform het advies van Ceph[51].

```
ceph osd pool create PROXMOX-CEPH 4096 4096
```

[51] http://ceph.com/docs/master/rados/operations/placement-groups/

Impediment:

Indien er 3 replica's opgeslagen worden is de netto capaciteit van van de pool 2330MB ((15*466GB) / 3) dit is minder dat in de IST-situatie beschikbaar is.

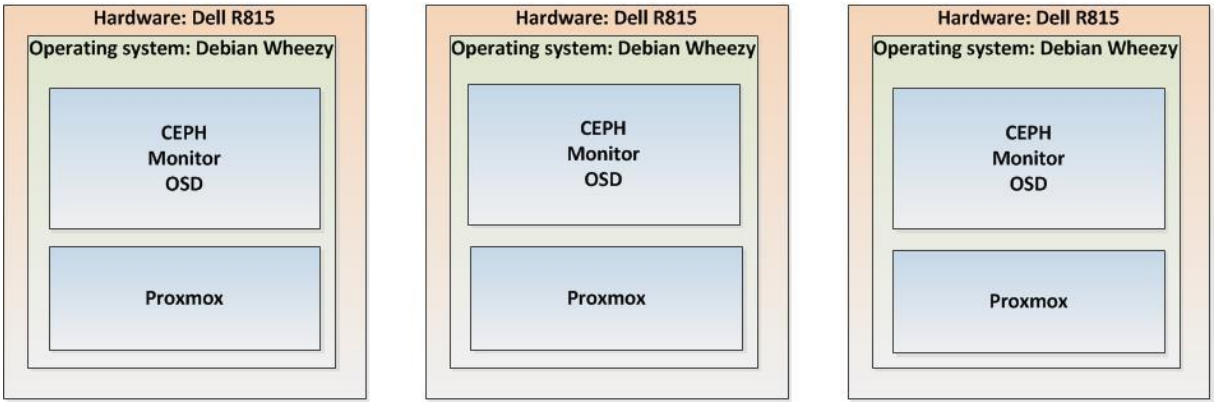


Figuur 33 Capaciteit Ceph in verhouding met IST-situatie

Het impediment is teruggekoppeld aan de opdrachtgever. Deze erkent de tekortkoming maar ziet deze niet als show stopper. Er is aangeboden het aantal replica's terug te brengen naar 2, waardoor de netto capaciteit toe zal nemen, de opdrachtgever heeft dit aanbod afgeslagen.

- **Het installeren van Proxmox**

Het installeren van de Proxmox software is een relatief langdurig proces, het is echter eenvoudig en goed gedocumenteerd binnen de wiki van het afstudeerbedrijf.



Figuur 34 Opbouw van nodes na installatie van Proxmox

- **Het creëren van een Proxmox cluster**

Het creëren van een cluster is eveneens een relatief eenvoudige taak die goed gedocumenteerd is binnen de wiki van het afstudeerbedrijf.

Na het creëren van het Proxmox cluster is het mogelijk de status van het cluster op te vragen met de commando´s pvecm status en pvecm nodes, de output van deze commando´s is terug te vinden in bijlage P.

Impediment:

Na het verhelpen van impediment bij de taak "**Het creëren van een Ceph storage cluster"** is het niet meer mogelijk een Proxmox cluster te realiseren. Tijdens het verhelpen van de impediment m.b.t. het creëren van een Ceph storage cluster was het noodzakelijk de hostnaam van het systeem te verbinden aan het IP van de monitoring interface. Proxmox gebruikt per definitie het IP adres van de vmbr0 interface. Het niet overeenkomen van de hostnaam met het IP van de vmbr0 interface maakt het voor Proxmox onmogelijk een cluster te vormen.

Een herconfiguratie van de netwerkinterfaces maakt het mogelijk Proxmox te installeren. Naast het inzichtelijk maken van het probleem was het noodzakelijk de gehele software stack opnieuw te installeren om het foutief gevormde Proxmox cluster ongedaan te maken. Het impediment heeft in totaal 8 uur vertraging opgeleverd.

- **Het toevoegen van de Ceph storage aan Proxmox**

Om Proxmox toegang te geven tot de gecreëerde Ceph storage dient Proxmox te beschikken over de Ceph cliënt admin keyring.
Dit is mogelijk met een eenvoudige kopie van /etc/ceph/ceph.client.admin.keyring naar het bestand /etc/pve/priv/ceph/PROXMOX-CEPH.keyring
Hierna is het mogelijk via de Proxmox gui een RBD(Rados Block Device, d.w.z. Ceph storage) aan te maken. In bijlage U zijn enkele afbeeldingen te vinden waarin het aanmaken van het RBD wordt weergegeven.

- **Het creëren van een virtuele benchmark-host op Proxmox**

Het installeren van een virtualisatiehost valt onder de reguliere werkzaamheden binnen het afstudeerbedrijf. In bijlage Y zijn enkele screenshots opgenomen om de specifieke kenmerken van de benchmark-host in een later (vervolg)onderzoek te kunnen reproduceren.

  - Het installeren van de benchmark-suite

De installatie van phoronix-test-suite van openbenchmarking.org vindt plaats op basis van het standaard systeemcommando apt-get install. Meer informatie over openbenchmarking.org is terug te vinden in hoofdstuk 0.

  - Het installeren van de benchmark-software

De installatie van de benchmarking tool iozone vindt plaats met het onderstaande commando. Vanuit beveiligingsoogpunt is gekozen geen informatie te delen met de makers van de testsuite m.b.t. de uitgevoerde benchmarks. De keuze voor de benchmarking tool iozone uit de phoronix-test-suite wordt gemotiveerd in hoofdstuk 7.

```
root@KC1038:~# phoronix-test-suite install pts/iozone

Do you agree to these terms and wish to proceed (Y/n): Y
Enable anonymous usage / statistics reporting (Y/n): n
Enable anonymous statistical reporting of installed software / hardware (Y/n): n
```

- Het testen van de benchmark-software

De standaard benchmark welke binnen het onderzoek gebruikt gaat worden voor het meten van de performance van de storage werd getest zonder het opslaan van de resultaten.

Impediment:

Tijdens het uitvoeren van de eerste test met een bestand van 8GB volgde de melding: test failed. Na enig onderzoek bleek de disk vol te staan. Hoewel de hosts geïnstalleerd is op een virtuele 32GB harddisk, is deze harddisk voorzien van Volume Management. De standaard omvang van de root partitie bedraagt 6GB. Met een tweetal commando's was het mogelijk de root partitie en het bijbehorende filesystem te vergroten.

```
root@KC1038:~# lvextend -L +8G /dev/KC1038-vg/
root@KC1038:~# resize2fs /dev/mapper/KC1038—vg-root
```

- **Het compileren van de Calamari-software**

Aan de hand van een handleiding[52] is de Calamari-software gecompileerd. Aanvullende informatie over deze handelingen inclusief commentaar is terug te vinden in bijlage V.
Calamari is een grafische schil voor Ceph. Calamari geeft inzicht in de status van Ceph en maakt het mogelijk beheertaken uit te voeren op het Ceph cluster.

- **Het compileren van Diamond**

Aan de hand van een handleiding[53] is de Diamond-software gecompileerd. Aanvullende informatie over deze handelingen inclusief commentaar is terug te vinden in bijlage W.
Diamond verzameld informatie op de individuele Ceph nodes en verzend deze naar de Calamari server.

- **Het realiseren van een repository[54]**

Om de Calamari en Diamond packages via een gestandardiseerde methode beschikbaar te stellen aan de Ceph en Calamari server(s), is er een repository aangemaakt op de door het afstudeerbedrijf bedoelde locatie. Het aanmaken van de repositories vindt plaats aan de hand van 3 systeemcommando's.

```
apt-get install dpkg-dev
cd /repo/mirror/calamari
dpkg-scanpackages /repo/mirror/calamari /dev/null | gzip -9c > Packages.gz
```

[52] https://marcosmamorim.wordpress.com/2014/07/11/build-package-and-install-calamari-on-debian-weezy/
[53] https://marcosmamorim.wordpress.com/2014/07/11/build-package-and-install-calamari-on-debian-weezy/
[54] https://help.ubuntu.com/community/Repositories/Personal

- **Het installeren van Diamond**

De installatie van Diamond vindt plaats op basis van het standaard systeemcommando apt-get install. Het is hiervoor wel noodzakelijk de eerder aangemaakte repository te initiëren.

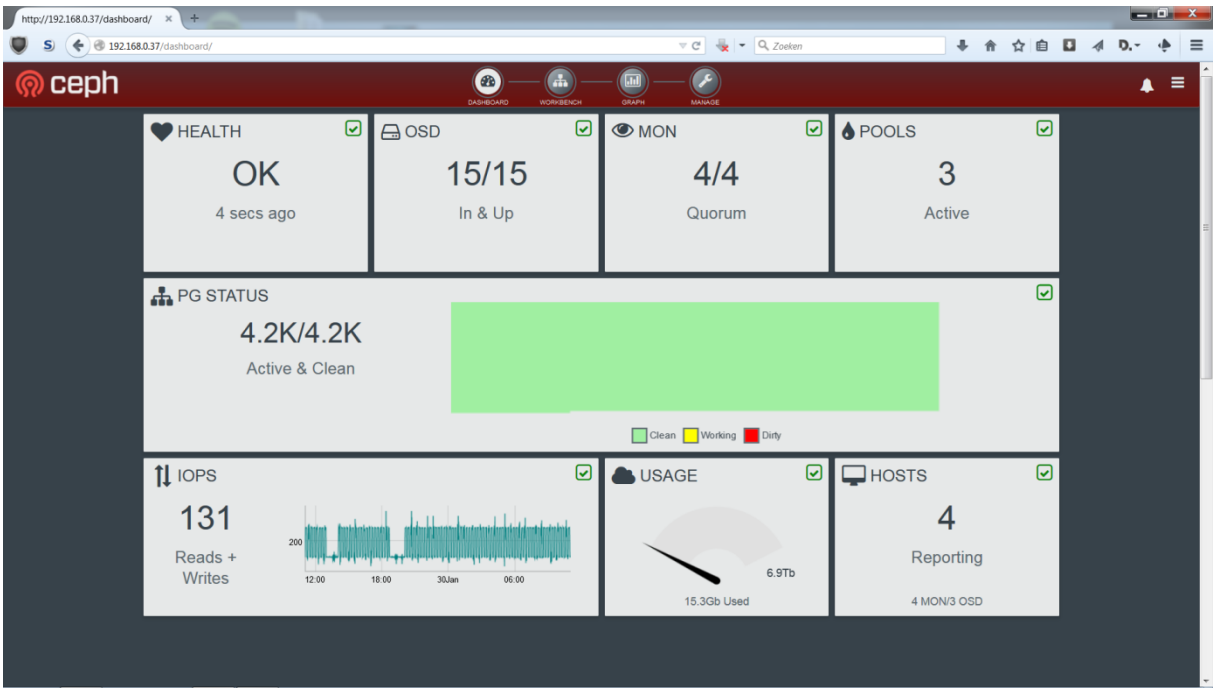
```
apt-get install diamond
```

- **Het installeren van de Calamari-client**

De installatie van Calamari-client vindt plaats op basis van het standaard systeemcommando apt-get install. Het is hiervoor wel noodzakelijk de eerder aangemaakte repository te initiëren.
Calamari-client dient alleen geïnstalleerd te worden op de Calamari-server, de naamgeving kan als misleidend worden ervaren. Het installatieproces is volledig beschreven in bijlage X.

- **Het installeren van Calamari-server**

Het installatieproces is volledig beschreven in bijlage X. Na het uitvoeren moet het mogelijk zijn de Calamari-server webbased te benaderen. Er dient ingelogd te worden met de tijdens de installatie ingevoerde credentials.



Figuur 35 De Calamari-server webinterface

6.4.4 Sprint 3 – Story 3

Deze sprint automatiseert de taken uit sprint 1 en 2. Deze sprint kan worden gezien als monnikenwerk. Alle in sprint 1 en 2 uitgevoerde taken dienen herschreven te worden naar ansbile syntax en in de juiste volgorde te worden geplaatst.

Ansible[55] omschrijft zichzelf als een krachtige automation engine die het mogelijk maakt applicaties eenvoudig te installeren en configuren. Ansbile maakt geen gebruik van custom scripting of custom code, ook wordt er geen agent gebruikt. Ansbile configureert servers en applicaties via het SSH protocol. Het SSH protocol is de standaard voor systeembeheer binnen het afstudeerbedrijf.

- Het schrijven van scripts voor Ansible
  - De volgende rollen zijn gedefinieerd
    - Common
      Deze role bevat platformonafhankelijke scripts. Zo maken alle servers gebruik van kerberosauthenticatie en ntp.

    - Debian
      Deze rol is specifiek voor het Debian OS, het configureert bijvoorbeeld repositories die alleen van toepassing zijn op Debiansystemen.

    - Proxmox
      Deze rol omvat alle Proxmox-gerelateerde zaken. De role is dus geheel onafhankelijk van Ceph.

    - Ceph
      Er is een aparte rol voor Ceph aangemaakt. Hoewel Ceph in enkele eenvoudige stappen vanuit Proxmox geïnstalleerd kan worden is er gekozen voor een duidelijke scheiding. Zo is het mogelijk deze Ceph role ook te gebruiken bij een installatie van overige software, zoals Owncloud. De installatie van Diamond maakt ook deel uit van deze rol.

In de bijlages D,E,F,G zijn per rol alle Ansbile scripts terug te vinden.

Alleen standaardhandelingen zijn opgenomen in Ansible. Zo is het aanmaken van de individuele OSD's de benchmark software en de Calamari server niet meegenomen in Ansible. Ook het aanmaken van de Proxmox-clusters is te specifiek voor Ansible.

[55] http://www.ansible.com/home

Enkele voorbeelden van de flexibiliteit van Ansible creates, when en first_available_file.

Voorbeeld alleen uitvoeren wanneer nog niet eerder uitgevoerd:

```
#
# share ceph key with Proxmox
#
- name: share ceph key with Proxmox
 shell: cp /etc/ceph/ceph.client.admin.keyring /etc/pve/priv/ceph/PROXMOX-CEPH.keyring
 args:
 creates: "/etc/pve/priv/ceph/PROXMOX-CEPH.keyring"
```

Voorbeeld alleen op specifieke host uitvoeren:

```
#
# Install Ceph software
#
- name: Install Ceph software
 shell: su - ceph bash -c "ceph-deploy install --no-adjust-repos KC0227-MON.afstudeer.org KC0229-MON.afstudeer.org KC0230-MON.afstudeer.org KC0424-MON.afstudeer.org"
 when: ( "{{ansible_hostname }}" == "KC0424-MON" )
```

Bestand plaatsen op basis van variable:

```
#
# hostname
#
- name: configure hostname
 template: src={{ item }} dest=/etc/hostname mode=0644 owner=root group=root
 first_available_file:
 - "{{ansible_hostname }}.hostname.j2"
 - hostname.j2
```

- Het testen van de scripts voor Ansible.
  Om de werking van de Ansible-scripts te testen dienen de Ansbile-scripts uitgevoerd te worden. Het is mogelijk alle Ansible-scripts voor alle Ceph-servers uit te voeren met de commando's:

  ```
  cd /etc/ansible
  ansible-playbook site.yml -l ceph_servers
  ```

  Ansible maakt gebruik van kleuren in de output van het playbook commando om de status van de individuele taken inzichtelijk te maken. Ansible gebruikt groen indien een taak succesvol is uitgevoerd, geel indien er een bestand is aangepast en rood indien het niet mogelijk is de role of taak uit te voeren.

- Het uitvoeren van herinstallaties.
  Om het geheel te testen zijn er veelvuldig herinstallaties van het basissysteem uitgevoerd op de Ceph-servers. Hierna is met Ansbile de installatie en configuratie van de applicaties Ceph en Proxmox tot in detail gecontroleerd. Na het uitvoeren van alle Ansible roles werd

automatisch een werkende Proxmox- en Ceph-omgeving opgeleverd. De totale doorlooptijd van herinstallatie tot bruikbare omgeving was 45~60 minuten. Het gebruik van Ansbile toont aan dat de in sprint 1 en 2 uitgevoerde handelingen gewaarborgd zijn en gerepliceerd kunnen worden.

7 Onderzoeksmethode

Om het succes van de PoC inzichtelijk te maken is er kwalitatief en kwantitatief onderzoek (Verhoeven, 2011, p. 30) uitgevoerd op de IST- en SOLL-situatie. In dit hoofdstuk is per onderdeel van het onderzoek (Oost, Een onderzoek voorbereiden, 2012) (Oost, Een onderzoek uitvoeren, 2012) (Oost, Een onderzoek rapporteren, 2011) de gebruikte onderzoeksmethode beschreven.

7.1 Schaalbaarheid

Om de schaalbaarheid van Ceph binnen de PoC omgeving inzichtelijk te maken worden twee aspecten onderzocht: het uitbreiden van het Ceph-cluster en het uitbreiden van de Ceph-cluster opslagcapaciteit. Het uitbreiden van het Ceph-cluster wordt getest door het toevoegen van een 4de node aan het bestaande cluster. Na deze uitbreiding moet het mogelijk zijn additionele OSD's toe te voegen om de opslagcapaciteit van het Ceph-cluster te vergroten. De toe te voegen node is wederom een Dell R815, deze keuze is gemaakt op basis van beschikbaarheid van hardware. In theorie is het mogelijk ieder type en model server toe te voegen.

Het uitbreiden van het Ceph-cluster heeft een grote samenhang met de beschikbaarheid en de performance van de servers. Over de beschikbaarheid kunnen we vanuit de theorie stellen dat deze omhoog gaat aangezien er meer nodes beschikbaar zijn om het totaal van 3 replica's te realiseren. De invloed op de performance wordt aangetoond in het onderzoek naar de performance. Hier wordt een additionele test uitgevoerd met 4 nodes en 17 OSD's in plaats van 3 nodes met 15 OSD's.

7.2 Beschikbaarheid

Om de beschikbaarheid van Ceph inzichtelijk te maken wordt gebruik gemaakt van de Calamari-interface. De Calamari-interface geeft inzicht in de status van het cluster. Om de garantie op beschikbaarheid te testen worden een drietal tests uitgevoerd die direct betrekking hebben op de hardware.

- Het fysiek verwijderen van een traditionele harddisk
- Het fysiek verwijderen van een journal disk, effectief leidt dit tot uitval van 5 OSD's
- Het fysiek verwijderen van beide power cables van één van de Ceph nodes

De tests zijn als volgt opgezet: Het Calamari virtualisatieimage is geplaatst op de Ceph storage. Het feit dat Calamari in staat is om een actuele status van het Ceph-cluster weer te geven is een indicatie dat de Ceph-storage nog actief is. Tijdens iedere individuele test gaat Ceph zijn kopieën herverdelen over de OSD's, om te voldoen aan de eis van 3 kopieën. De duur van de synchronisatieslag na het uitvoeren van de individuele testen gemeten. Hierna wordt het cluster weer in ere hersteld. Ook dan wordt de duur van de synchronisatie gemeten. Tijdens het synchroniseren wordt geen additionele storage verbruikt. Er is echter minimale groei te verwachten door o.a. logfiles. Het begin en einde van de replicatie wordt vastgesteld aan de eerste en laatste logregel met betrekking tot replicatie. Ter ondersteuning worden screenshots gebruikt uit de Calamari-beheerinterface.

Er worden geen testen uitgevoerd met 1 of 2 replica's. Ook worden er geen testen uitgevoerd met het verwijderen van 1 of meer netwerkkabels.

7.3 Security

Er is binnen de technische realisatie van de PoC reeds een duidelijke scheiding aangebracht tussen het netwerk waar Proxmox zijn virtualisatiediensten aanbiedt en de niet-routeerbare twee Ceph-storage netwerken. Dit is beveiligingstechnisch een grote stap in de juiste richting, Ceph zegt hierover[56]: "We prefer that the cluster network is **NOT** reachable from the public network or the Internet for added security." Daarnaast is er gebruik gemaakt van Cephx authenticatie, het is alleen mogelijk met het Ceph-cluster te communiceren indien de juiste sleutels beschikbaar zijn. Ceph zegt over het (gebrek aan)gebruik van authenticatie het volgende[57]: "If you disable authentication, you are at risk of a man-in-the-middle attack altering your client/server messages, which could lead to disastrous security effects."

Voor het vaststellen van de netwerk security worden de volgende drie tests uitgevoerd:

- Op de netwerken 192.168.0.0/24, 172.16.81.0/24 en 172.16.83.0/24 worden nmap scans uitgevoerd naar open service poorten. De scans zoeken naar de 1000 meest voorkomende udp en tcp poorten. Welke service verantwoordelijk is voor de open poort wordt inzichtelijk gemaakt met de tool lsof.
- Op het routeerbare 192.168.0.0/24 netwerk wordt een security scan uitgevoerd met de tool, OpenVAS, een open source vulnerability scanner.
- Sniffen op het storage netwerk: wordt data in plain tekstformaat uitgewisseld?
    - Het schrijven van herkenbare data vanaf een op Ceph-storage gevirtualiseerde host op basis van het commando: dd if=input of=output
    - Het afvangen van het netwerkverkeer op de onderliggende Proxmox host met het commando: tcpdump -nn -X -i vmbr1 -s 0 -tttt -vvvvv | tee /tmp/sniff.txt

Het afstudeerbedrijf is gehouden aan het BIR (Baseline Informatiebeveiliging Rijksdienst) normen-kader voor de rijksoverheid. Bij de technische realisatie is in de sprints rekening gehouden met de binnen dit normenkader gestelde eisen. Aan de hand van het document BIR_Operationele_ Handreiking_v1_0.pdf[58] zijn de eisen voor serverinrichting inzichtelijk gemaakt. Bij de inrichting van een server dient aan de volgende eisen uit het TNK (Technisch Normenkader) te worden voldaan: 11.3.1.1, 10.1.3.4, 11.2.2, 11.1.1, 12.6.1.4, 10.4.1.1, 10.4.1.2 11.7.1.2, 10.4.1.3, 12.6.1.2, 12.4.1.6, 12.6.1.3. Een volledig overzicht van deze eisen is terug te vinden in bijlage 0. In bijlage DD is terug te vinden hoe is voldaan aan deze eisen.

7.4 Beheerbaarheid

Dit onderdeel van het onderzoek bestaat uit het vergelijken van de beheerinterface van de traditionele storage met de Calamari-interface. Er wordt in kaart gebracht welke functionaliteit er in beide interfaces aanwezig is en hoe deze zich met elkaar verhouden.

[56] http://ceph.com/docs/master/rados/configuration/network-config-ref/
[57] http://ceph.com/docs/master/rados/configuration/auth-config-ref/
[58] http://www.earonline.nl/images/earpub/5/5c/BIR_Operationele_Handreiking_v1_0.pdf

7.5 Performance

Er wordt gebruik gemaakt van een phoronix benchmarking suite die geleverd wordt via openbenchmarking.org. Deze gestandaardiseerde benchmarking suite bevat een groot aantal storagegerelateerde testen. Alle testen worden minimaal 4 maal uitgevoerd, indien de standaarddeviatie te hoog is wordt de test vaker uitgevoerd. Deze methodiek heeft veel voordelen vanuit de theorie. In de praktijk wordt het echter lastig deze testen parallel uit te voeren vanaf meerdere nodes, indien de deviatie te groot is bestaat de kans dat één van de nodes zijn testen niet langer parallel uitvoert aan de overige nodes.

Om het testen in met name de complexe gedistribueerde Ceph cloud storage omgeving enigszins overzichtelijk te houden is gekozen om de testen uit te voeren met de applicatie iozone. Iozone is in staat met verschillende blockgroottes sequentieel te lezen en te schrijven. Het effect van random reads en writes wordt niet onderzocht. De sequentiële testen van verschillende blockgrootte hadden een totale duur van 2/2.5 uur.

Alle testen zijn uitgevoerd op exact dezelfde gevirtualiseerde server, met exact dezelfde onderliggende hardware, een Dell R815. Indien er bottlenecks ontdekt worden in de infrastructuur dienen deze verholpen te worden en het effect dient inzichtelijk te worden gemaakt met de gestandaardiseerde test. Om de performance van Ceph op een correcte manier te kunnen vergelijken is een service window aangevraagd waarin alle virtuele images op de IBM DS42000 down mochten worden gebracht. Dit heeft ervoor gezorgd dat alle interactie met de storage afkomstig was van de benchmark host.

Tot slot wordt het effect van 1,2 en 3 replica’s in kaart gebracht. De test met 1 replica is om aan te tonen dat er geen directe indicatie is voor een netwerk bottleneck. Onderzoek naar de invloed van complexe Ceph-instellingen zoals CRUSHmaps maken geen deel uit van dit onderzoek. De opdrachtgever ziet dit wel als een nuttig vervolgonderzoek.

7.6 Kosten

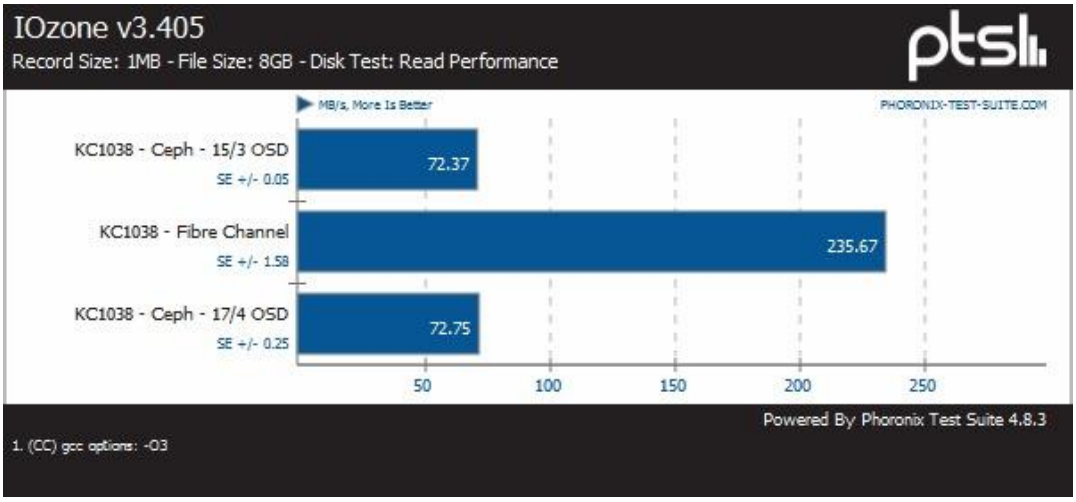
Het is voor de afstudeerder niet mogelijk prijsindicaties op te vragen bij leveranciers. Om toch inzicht te geven de kosten van de Ceph Cloud storage ten opzichte van de traditionele storage in de context van de PoC zullen de variabelen waar kosten aan verbonden zijn in kaart gebracht worden. Van het binnen de PoC gebruikte server component is bekend dat de kosten hiervan bij aanschaf €7000,-. Van de binnen de PoC gebruikte IBM DS4200 disk enclosure is bekend dat de initiële aanschaf kosten €35.000,- bedraagde. Allen op basis van deze van deze twee bedragen is het mogelijk aan te tonen of de vervanging van traditionle storage door cloud storage de in het afstudeerplan begrote besparing van €500.000, - kan realiseren.

8 Onderzoeksresultaten

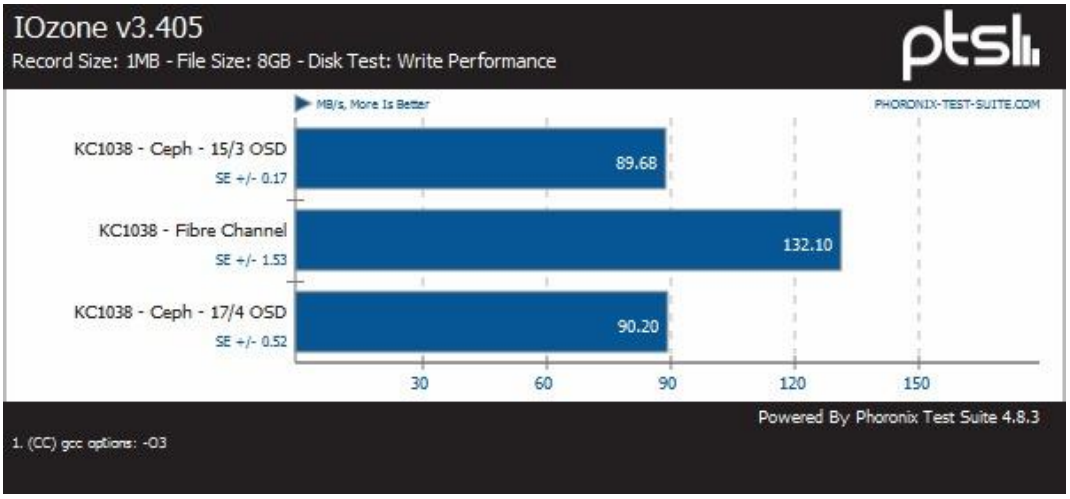
In dit hoofdstuk worden de ruwe meetresultaten weergegeven zoals deze voort zijn gekomen uit de in hoofdstuk 7 vastgestelde onderzoeksmethodieken. In hoofdstuk 9 wordt een waardeoordeel aan deze resultaten gekoppeld.

8.1 Schaalbaarheid

Het toevoegen van een 4de Ceph node verliep probleemloos. Het toevoegen van 2 additionele OSD's op deze 4de node verliep eveneens zonder problemen. Na het toevoegen van de nodes en de OSD's zijn er opnieuw performancemetingen gedaan om de verschillen tussen 3 en 4 nodes en respectievelijk 15 en 17 OSD's inzichtelijk te maken. Onderstaande grafieken geven de lees- en schrijfsnelheid weer van een bestand van 8GB met een blockgrootte van 1MB. Deze test is uitgevoerd voor het upgraden van het netwerk naar 2 en 4Gb/s, zie hoofdstuk 9.5. De 4de machine beschikt niet over afdoende netwerkinterfaces om ook deel uit te maken van een test met een geüpgrade netwerk.



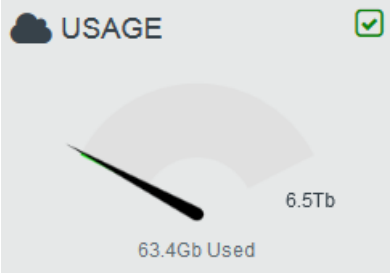
Figuur 36 Leesperformance – bestand 8GB, blockgrootte 1MB



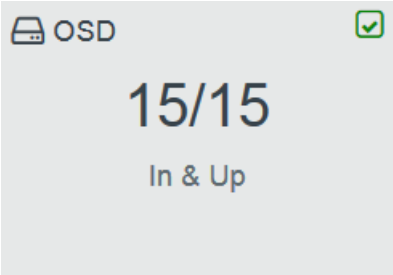
Figuur 37 Schrijfperformance – bestand 8GB, blockgrootte 1MB

8.2 Beschikbaarheid

8.2.1 Het fysiek verwijderen van een OSD disk



Figuur 39 Diskvulling voor aanvang van de test



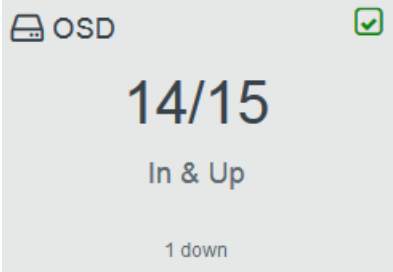
Figuur 38 Beschikbare OSD voor aanvang van de test

(down for 302.300029)

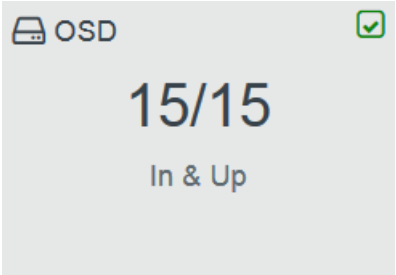
Einde replicatie: 2015-02-07 14:14:33.013518 mon.0 [INF] pgmap v319985: 4224 pgs: 3888 active+clean, 272 active+degraded, 64 active+remapped; 20876 MB data, 65327 MB used, 6546 GB / 6610 GB avail; 2326 kB/s wr, 1074 op/s; 425/16131 objects degraded (2.635%); 210 MB/s, 3 keys/s, 53 objects/s recovering

Duur replicatie na reparatie: **~30 seconden**
Aanvang replicatie: 2015-02-07 14:15:43.473158 mon.0 [INF] osdmap e374: 15 osds: 15 up, 15 in

Einde replicatie: 2015-02-07 14:16:12.440943 mon.0 [INF] pgmap v320063: 4224 pgs: 4224 active+clean; 20876 MB data, 65334 MB used, 7012 GB / 7076 GB avail; 177 kB/s wr, 23 op/s; 153 MB/s, 1 keys/s, 39 objects/s recovering



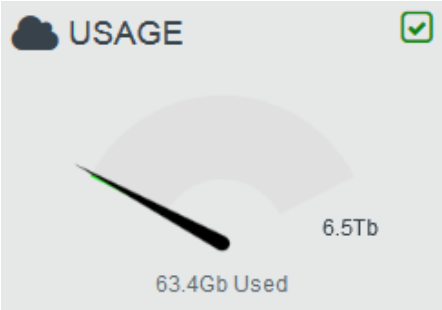
Figuur 40 Beschikbare OSD's tijdens de test



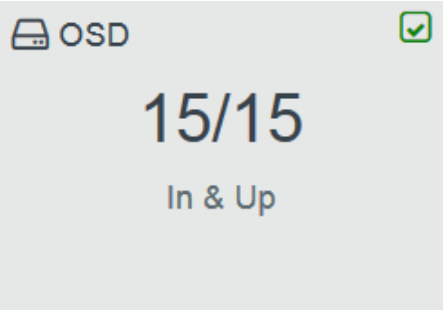
Figuur 41 Beschikbare OSD's na de test

8.2.2 Het fysiek verwijderen van een journal disk,

Het verwijderen van een journal disk leidt effectief leidt dit tot uitval van 5 OSD's.



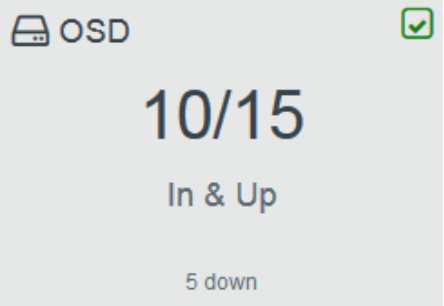
Figuur 43 Diskvulling voor aanvang van de test



Figuur 42 Beschikbare OSD voor aanvang van de test

Duur replicatie na defect: **108 seconden**
Aanvang replicatie: 2015-02-07 14:23:18.712723 mon.0 [INF] osdmap e399: 15 osds: 10 up, 14 in

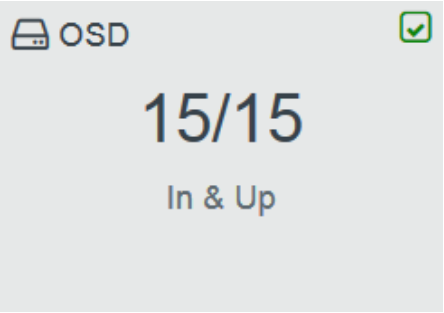
Einde replicatie: 2015-02-07 14:25:06.396211 mon.0 [INF] pgmap v320365: 4224 pgs: 1729 active+degraded, 2495 active+remapped; 20876 MB data, 56850 MB used, 4599 GB / 4655 GB avail; 2580 kB/s wr, 1143 op/s; 4900/16131 objects degraded (30.376%); 191 MB/s, 48 objects/s recovering



Figuur 44 Beschikbare OSD's tijdens de test

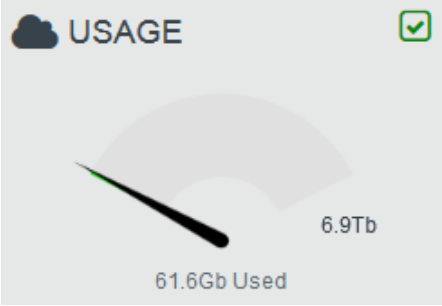
Duur replicatie na reparatie: **57** seconden
Aanvang replicatie: 2015-02-07 14:29:54.918040 mon.0 [INF] osdmap e453: 15 osds: 15 up, 15 in

Einde replicatie: 2015-02-07 14:30:51.319571 mon.0 [INF] pgmap v320670: 4224 pgs: 4224 active+clean; 20877 MB data, 65287 MB used, 7012 GB / 7076 GB avail; 3326 kB/s wr, 1621 op/s; 24781 kB/s, 6 objects/s recovering

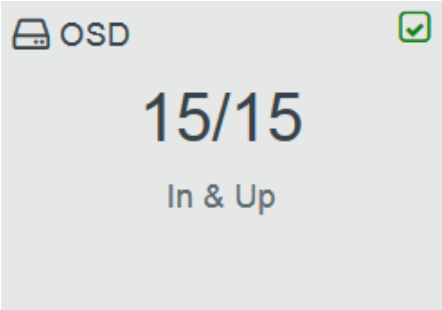


Figuur 45 Beschikbare OSD's na de test

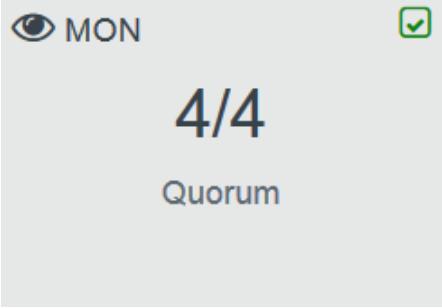
8.2.3 Het fysiek verwijderen van beide power cables van één van de Ceph nodes.



Figuur 46 Diskvulling voor aanvang van de test



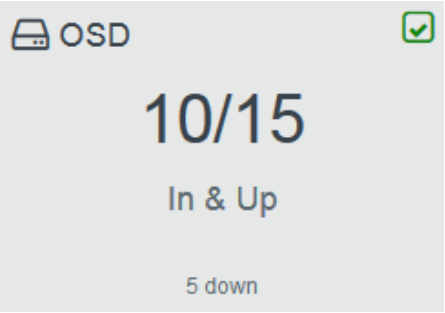
Figuur 47 Beschikbare OSD voor aanvang van de test



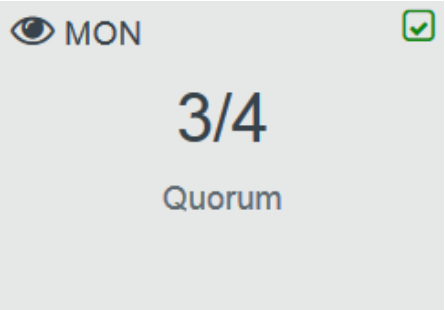
Figuur 48 Aantal beschikbare monitor-hosts voor aanvang van de test

Duur replicatie na defect: **~102 seconden**
Aanvang replicatie:2015-02-07 14:52:39.839903 mon.0 [INF] osdmap e399: 15 osds: 10 up, 14 in

Einde replicatie: 2015-02-07 14:54:26.886507 mon.0 [INF] pgmap v321501: 4224 pgs: 1729 active+degraded, 2495 active+remapped; 20880 MB data, 56749 MB used, 4599 GB / 4655 GB avail; 1887 kB/s wr, 901 op/s; 4898/16131 objects degraded (30.364%); 119 MB/s, 2 keys/s, 31 objects/s recovering



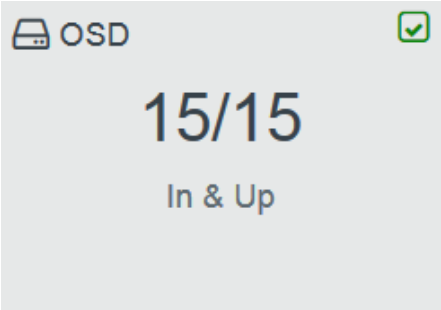
Figuur 49 Beschikbare OSD's tijdens de test



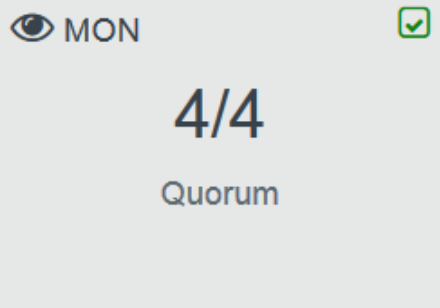
Figuur 50 Beschikbare monitor hosts tijdens de test

Duur replicatie na reparatie: **60 seconden**
Aanvang replicatie:2015-02-07 15:07:51.974130 mon.0 [INF] osdmap e696: 15 osds: 11 up, 11 in

Einde replicatie: 2015-02-07 15:08:52.467908 mon.0 [INF] pgmap v322134: 4224 pgs: 4224 active+clean; 20880 MB data, 65227 MB used, 7012 GB / 7076 GB avail; 1923 kB/s wr, 919 op/s; 145 MB/s, 37 objects/s recovering



Figuur 51 Beschikbare OSD's na de test



Figuur 52 Beschikbare monitor hosts na de test

8.3 Security

In dit hoofdstuk geeft inzicht in de ruwe output van de security scans.

8.3.1 Nmap

De resultaten van de Nmap scan zien er als volgt uit:

```
Starting Nmap 6.40 ( http://nmap.org ) at 2015-02-25 13:45 CET
Nmap scan report for kc0230.afstudeer.org (192.168.0.7)
Host is up (0.00074s latency).
Not shown: 996 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
111/tcp  open  rpcbind
3128/tcp open  squid-http
MAC Address: D0:67:E5:F9:5A:A5 (Dell)
```

Output van alle Nmap scans is terug te vinden in bijlage S.

Open poorten zijn vervolgens met het volgende lsof commando inzichtelijk gemaakt:

```
root@KC0227-MON:~# for i in `netstat -an | grep LISTEN | grep ^tcp | awk '{print $4}' | grep -v :: |
awk -F\: '{print $2}'`; do lsof -n -i4TCP:$i | grep LISTEN; done
```

De resultaten van dit commando zien er als volgt uit:

```
ceph-mon  24791 root   12u  IPv4   60999      0t0  TCP 172.16.83.10:6789 (LISTEN)
pveproxy  25420 www-data   5u  IPv4   69786      0t0  TCP *:8006 (LISTEN)
rpcbind 3217 root   8u  IPv4  16465      0t0  TCP *:sunrpc (LISTEN)
ceph-osd   7589 root   5u  IPv4   27107      0t0  TCP 172.16.83.10:6800 (LISTEN)
ceph-osd  7589 root   7u  IPv4   27109      0t0  TCP 172.16.81.10:6801 (LISTEN)
```

Alle lsof output is terug te vinden in bijlage T.

8.3.2 OpenVAS

De resultaten van de OpenVAS scan, zijn opgenomen in de bijlage Z.

Host	Start	End	High	Medium	Low	Log	False Positive
192.168.0.7	Feb 9, 13:30:13	Feb 9, 15:18:08	0	0	1	16	0
192.168.0.8	Feb 9, 13:30:13	Feb 9, 15:18:06	0	0	1	16	0
192.168.0.9	Feb 9, 13:30:13	Feb 9, 15:18:25	0	0	1	16	0
192.168.0.10	Feb 9, 13:30:13	Feb 9, 15:17:42	0	0	1	16	0
Total: 4			0	0	4	64	0

Figuur 53 Overzicht OpenVAS resultaten

8.3.3 Sniffer

Output van de sniffer bevat de volgende inhoud:

```
root@KC0227-MON:~# grep AAAA /tmp/sniff.txt | more
….
0x0700: 4141 4141 4141 4141 4141 4141 4141 4141 AAAAAAAAAAAAAAAA
0x0710: 4141 4141 4141 4141 4141 4141 4141 4141 AAAAAAAAAAAAAAAA
0x0720: 4141 4141 4141 4141 4141 4141 4141 4141 AAAAAAAAAAAAAAAA
0x0730: 4141 4141 4141 4141 4141 4141 4141 4141 AAAAAAAAAAAAAAAA
0x0740: 4141 4141 4141 4141 4141 4141 4141 4141 AAAAAAAAAAAAAAAA
0x0750: 4141 4141 4141 4141 4141 4141 4141 4141 AAAAAAAAAAAAAAAA
….
```

8.4 Beheerbaarheid

In onderstaande tabel zijn de functionaliteiten van de beheerinterfaces in kaart gebracht. Met behulp van groene vinkjes is voor beide oplossingen inzichtelijk gemaakt of de functioneliteit aanwezig is. Indien de functionaliteit niet aanwezig is binnen de oplossing is dit kenbaar gemaakt met een rood kruis.

Functionaliteit	IBM DS4200	Calamari
Het definiëren van een LUN/Pool	✓	✓
Het weergeven van defecte disks	✓	✓
Het weergeven van de health status	✓	✓
De opbouw van een RAID set/pool	✓	✓
Het aanpassen van systeemwaarden	✓	✓
Inzicht in Raid level / replica's	✓	✓
Totale Capaciteit	✗	✓
Beschikbare capaciteit	✓	✓
Verbruikte capaciteit	✗	✓
Toegekende capaciteit	✓	✓
Authenticatie applicatie	✗	✓
Activiteit (IOPS)	✗	✓
Historisch verbruik	✗	✓

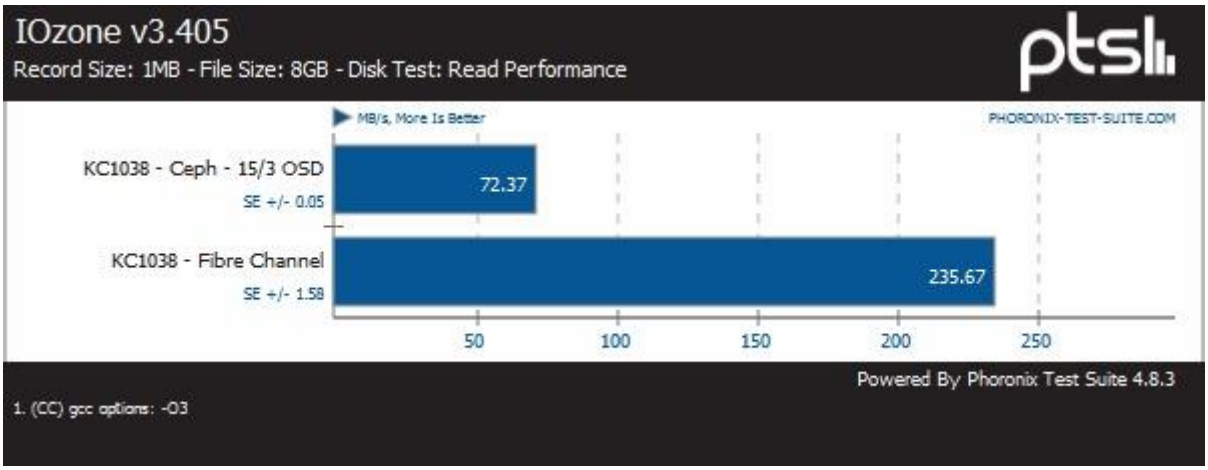
Tabel 9 Vergelijking functionaliteiten

8.5 Performance

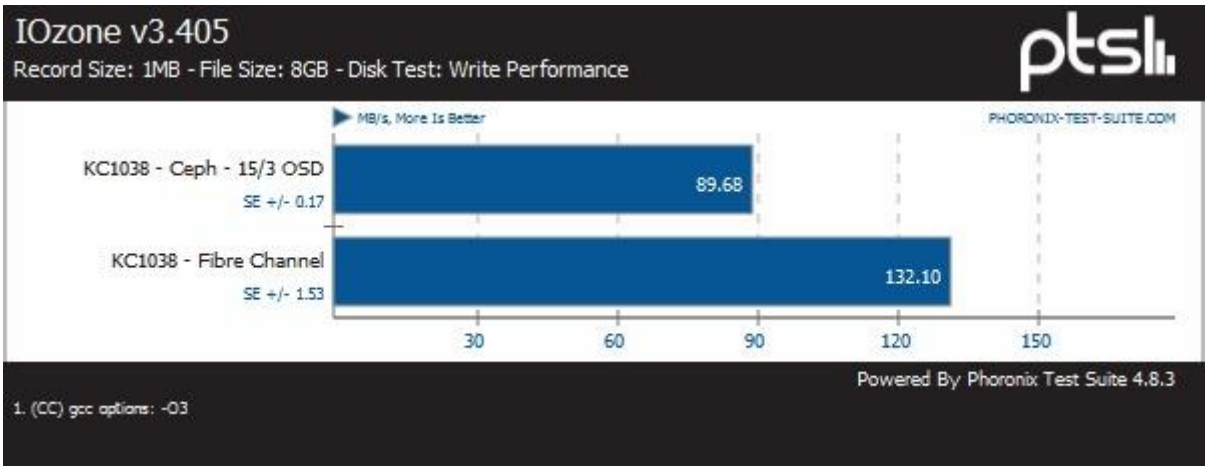
In dit hoofdstuk worden slechts enkele van de meetresultaten weergegeven. Alle resultaten van de uitgevoerde benchmarks zijn terug te vinden in bijlage H,I,J,K,L,M. De interpretatie voor deze verscheidenheid aan benchmarks is terug te vinden in de discussie.

8.5.1 Ceph vs Fibre Channel

Initieel werd de lees- en schrijfsnelheid van de Ceph vergeleken met de lees- en schrijfsnelheid van het traditionele IBM DS4200 SAN.



Figuur 54 Vergelijk in leessnelheid -3 nodes 15 disks Ceph cluster versus traditionele storage-

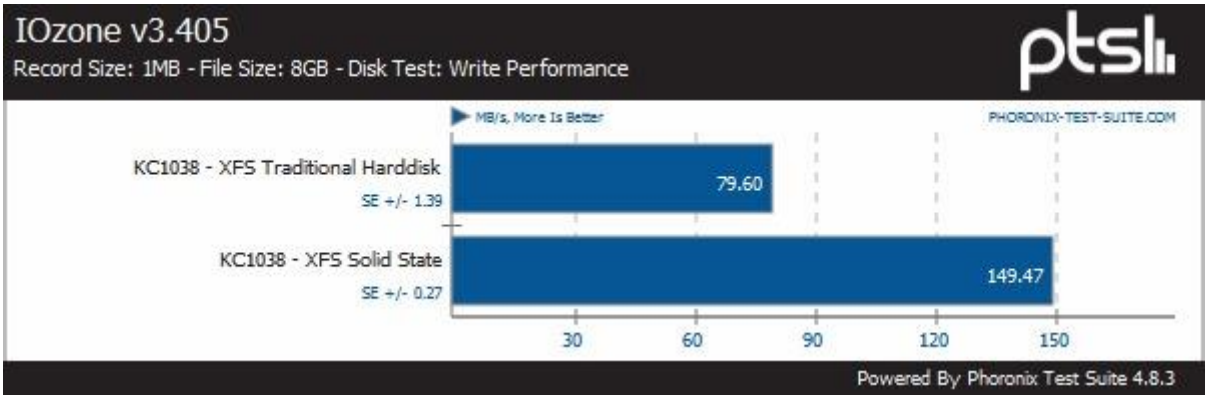


Figuur 55 Vergelijk in schrijfsnelheid -3 nodes 15 disks Ceph cluster versus traditionele storage-

8.5.2 PERC H200 Controller

Na het uitvoeren van van het initiële vergelijk tussen Ceph en de tradionele storage werden de individuele harddisks getest.

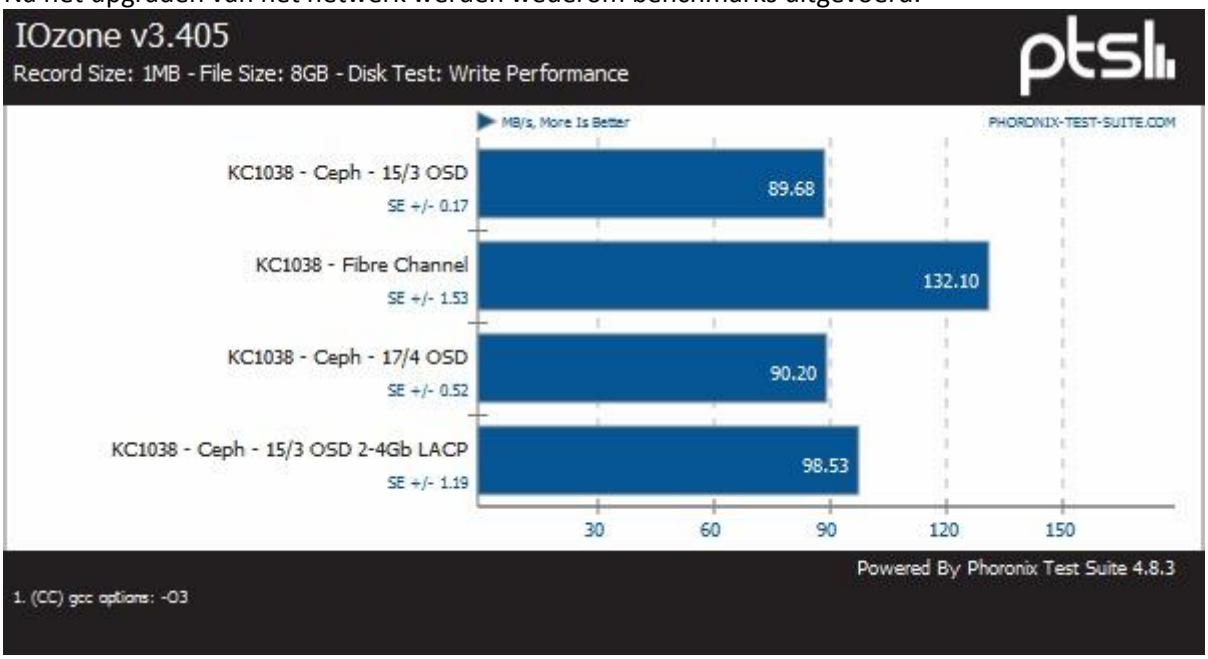
Figuur 56 Vergelijk in leessnelheid –locale tradionele harddisk versus locale Solid State harddisk-



Figuur 57 Vergelijk in schrijfsnelheid –locale tradionele harddisk versus locale Solid State harddisk-

8.5.3 Netwerk upgrade

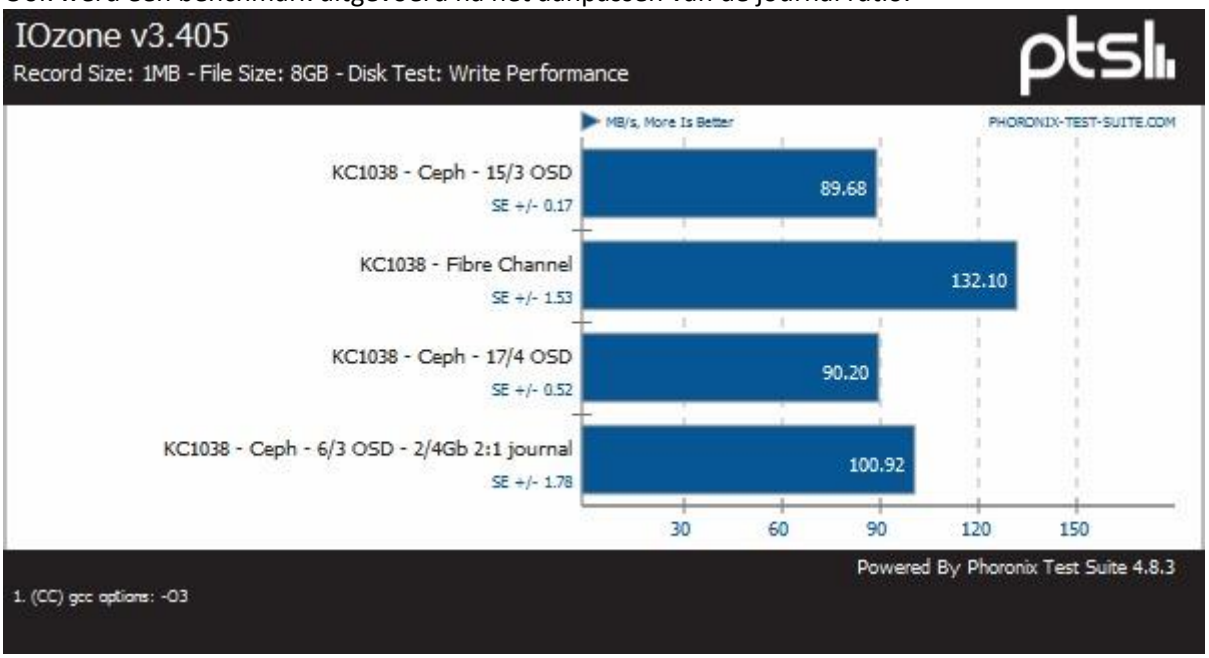
Na het upgraden van het netwerk werden wederom benchmarks uitgevoerd.



Figuur 58 Vergelijk in schrijfsnelheid –na verhelpen netwerk bottleneck-

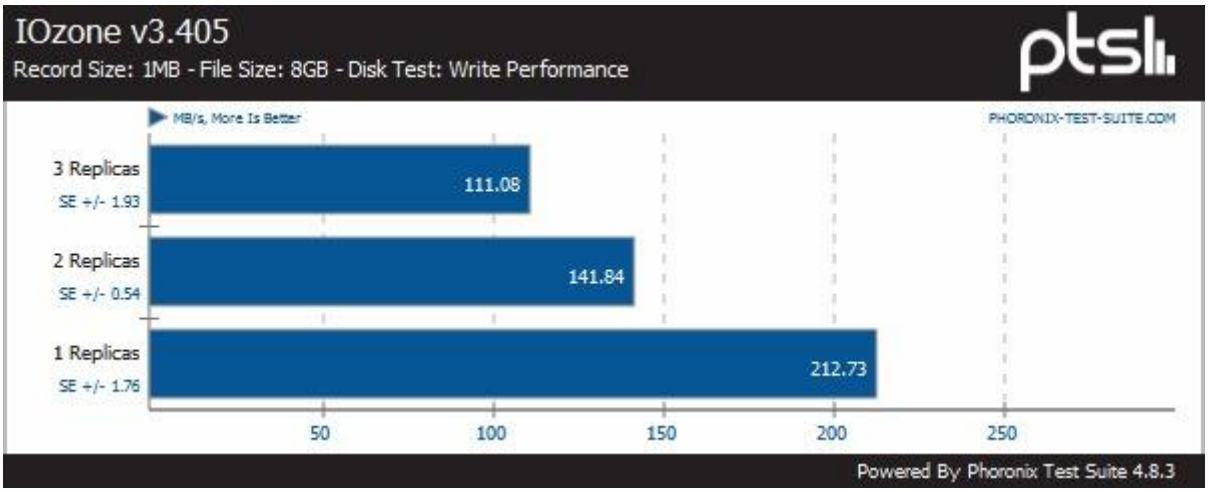
8.5.4 Journal ratio

Ook werd een benchmark uitgevoerd na het aanpassen van de journal ratio.



Figuur 59 Vergelijk in schrijfsnelheid –na optimalisatie journal ratio 3 Ceph nodes 6 disks-

8.5.5 Effect van replica's op performance



Figuur 60 Vergelijk in schrijfsnelheid –bij 3, 2 en 1 replica's-

8.6 Kosten

Binnen het PoC zijn de volgende variabelen aanwezig die gepaard gaan met kosten:

Ceph	Traditionele storage
Calamari Open Source beheerinterface	Support contract noodzakelijk om beheersoftware te mogen gebruiken
Ceph is gratis Open Source software	Support contract noodzakelijk om aanpassingen door te kunnen voeren
Booten van USB stick	4 x ongebruikte disks
Solid State Disks (ratio 1:1)	-
Harddisks om capaciteit te vergroten	Refurbished harddisks om capaciteit te vergroten
8 netwerkpoorten met host	2 Fibre channel adapter om verticaal de schalen op basis van FC-AL
8 netwerkpoorten in server	2 Netwerkpoorten in server
High performance controller	-
Arbeidsintensief	Statisch / Minder arbeidsintensief

Tabel 10 Overzicht kosten variabelen

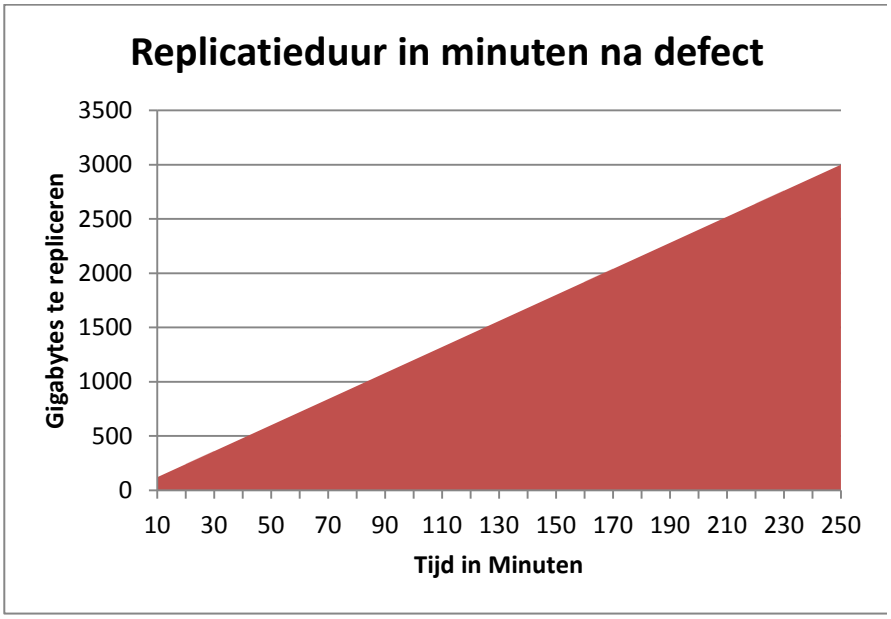
9 Discussie

9.1 Schaalbaarheid

Het is mogelijk additionele nodes toe te voegen aan Ceph. Dit stelt de system administrator in staat met meer rekenkracht gebruik te maken vandezelfde storage. Daarnaast is het mogelijk de capaciteit uit te breiden door OSD's toe te voegen. Hoewel in het schaalbaarheidsonderzoek gebruik is gemaakt van slechts 1 type server, de Dell R815, is het mogelijk ieder type en model server deel uit te laten maken van een Ceph cluster. Nodes kunnen worden toegevoegd aan het cluster als server en bijdragen aan de totale opslagcapaciteit of als client en gebruik maken van de beschikbare shared storage.

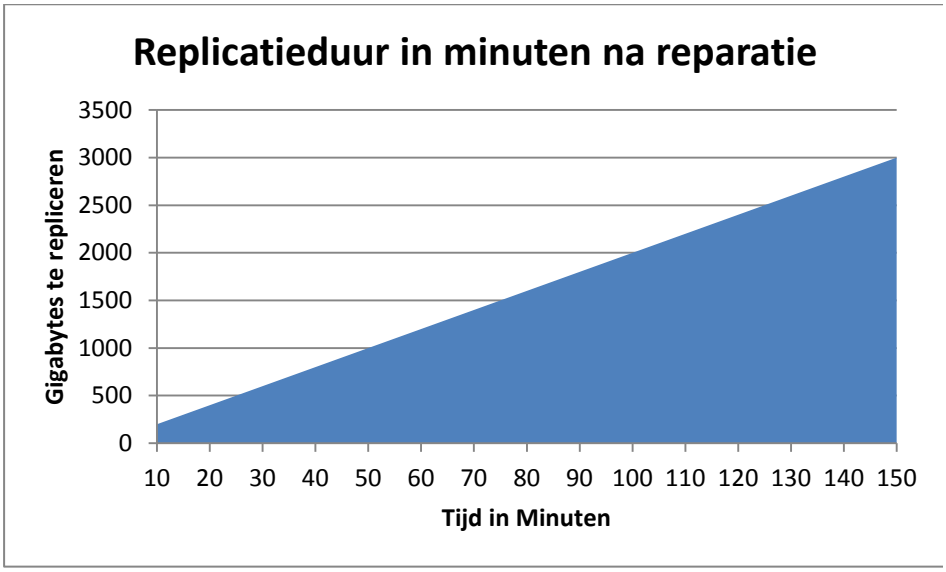
9.2 Beschikbaarheid

Ceph is in staat uitval van één of meerdere disks op te vangen, daarnaast is Ceph in staat de uitval van één of meerdere nodes op te vangen. Bij uitval zorgt Ceph zelf voor het opnieuw repliceren van de opgeslagen data. Indien een uitgevallen disk, disks of server weer beschikbaar is, zorgt Ceph wederom automatisch voor de replicatie. Tijdens de replicatie na het defect van 5 disken en de volledige host (eveneens 5 disken) heeft Ceph ruim 20GB gerepliceerd in ~100 seconden. Uitgaande van deze informatie kunnen we vaststellen dat een maximaal gevulde replica (~2.3TB) in de PoC-opstelling 200 minuten nodig heeft om gerepliceerd te worden na uitval van 1/3 van de clustercapaciteit.



Figuur 61 Duur replicatie na defect adhv omvang

Na reparatie van het defect was Ceph in staat ruim 20GB te repliceren in ~60 seconden. Uitgaande van deze informatie kunnen we vaststellen dat een maximaal gevulde replica (~2.3TB) in de PoC-opstelling 120 minuten nodig heeft om gerepliceerd te worden na herstel van 1/3 van de clustercapaciteit.



Figuur 62 Duur replicatie na reparatie adhv omvang

9.3 Security

De NMAP scans (8.3.1) (Bijlage NMAP scan resultaten) hebben aangetoond dat Ceph niet zichtbaar is op het routeerbare netwerk waar Proxmox zijn diensten op levert. De open poorten zijn op basis van lsof informatie toe te kennen aan respectievelijk het operating system en ProxMox. NMAP laat op de Ceph interfaces poort 6789 als open zien en kent hier die service ibm-db2-admin aan toe. Lsof toont aan dat het hier in werkelijkheid om de service ceph monitor gaat. Tevens toont de lsof aan dat Ceph allen actief is op de daarvoor gecreëerde netwerken.

Aan de hand van de OpenVAS scan (8.3.2) (Bijlage OpenVAS resulaten) kan worden vastgesteld dat het systeem niet vatbaar is voor bekende aanvalsvectoren. Er is per node één melding van het niveau laag. Deze melding is algemeen van aard en houdt geen verband met Ceph.

De resultaten die voortkomen uit de sniffer (0)Security tonen aan dat er geen gebruik wordt gemaakt van encryptie tijdens de uitwisseling van gegevens tussen de clients en de Ceph-storage. De informatie is in plain tekst terug te lezen met een sniffer. De duidelijke scheiding in netwerken zorgt ervoor dat de Ceph-gerelateerde netwerken niet direct te benaderen zijn. Afhankelijk van de gevoeligheid van de data moet overwogen worden data encryptie toe te passen binnen Ceph. Data encryptie zonder hardware offloading heeft naar verwachting echter een impact op de performance.

9.4 Beheerbaarheid

Calamari biedt op hoofdlijnen dezelfde functionaliteit als de beheerinterface van de IBM DS4200 (0). Binnen het onderzoek is gebleken dat Calamari zelfs meer functionaliteit biedt, met name op het gebied van verbruik en historische gegevens.

Hoewel de Calamari interface veel inzicht geeft in de status van het systeem en de mogelijkheid biedt basale handelingen uit te voeren zitten er vanuit het perspectief van de beheerbaarheid nadelen aan het gebruik van Ceph storage ten opzichten van traditionele storage. Zo bestaat het vervangen van een defecte disk in de IBM DS4200 slechts uit het fysiek verwijderen van de defect disk en het plaatsen van een vervangende disk. Binnen Ceph dienen deze handelingen ook uitgevoerd te worden daarnaast dienen nog een groot aantal handelingen(commando's) uitgevoerd te worden om binnen de Ceph software de defecte OSD af te bouwen en de nieuwe OSD op te bouwen en aan te bieden aan het Ceph cluster. Het op- en afbouwen van OSD's maakt geen deel uit van de Calarmari interface. Vergelijkbare functionaliteit is ook niet aanwezig binnen het IBM DS4200 disk enclosure. Binnen het IBM DS4200 disk enclosure wordt dit afgevangen door de hardware.

Daarnaast is duidelijk te zien dat Calamari een interface heeft met een modern design (Figuur 35), in tegenstelling tot de ouderwetse look van de IBM DS4200 grafische interface (Figuur 9 / Figuur 10).

9.5 Performance

De eerste benchmark was een vergelijking tussen Ceph en de Traditionele storage. In deze benchmark is gekeken naar zowel de lees- als schrijfsnelheid. In beide gevallen presteert Ceph slechter dan traditionele storage.

Om de mogelijke oorzaak hiervan te achterhalen is er een benchmark uitgevoerd op de individuele harddisks zonder gebruik te maken van Ceph. De formatting van de disk en van de virtuele disk binnen de gevirtualiseerde benchmark-server zijn geconformeerd aan de formatteringen die gebruikt worden bij het aanmaken van een OSD. Uit de benchmark kunnen de volgende conclusies worden getrokken. De begrenzing van de solid state disk of de PERC H200 controller ligt op 195MB/s lees- en 150MB/s schrijfsnelheid (8.5.2). Dit geeft aan dat de performance van de traditionele disk niet begrensd wordt door de controller, de controller is immers in staat minimaal 195- lees- en 150MB/s schrijfsnelheid te behalen. De traditionele disk leest met 83MB/s en schrijft met 80MB/s (8.5.2).

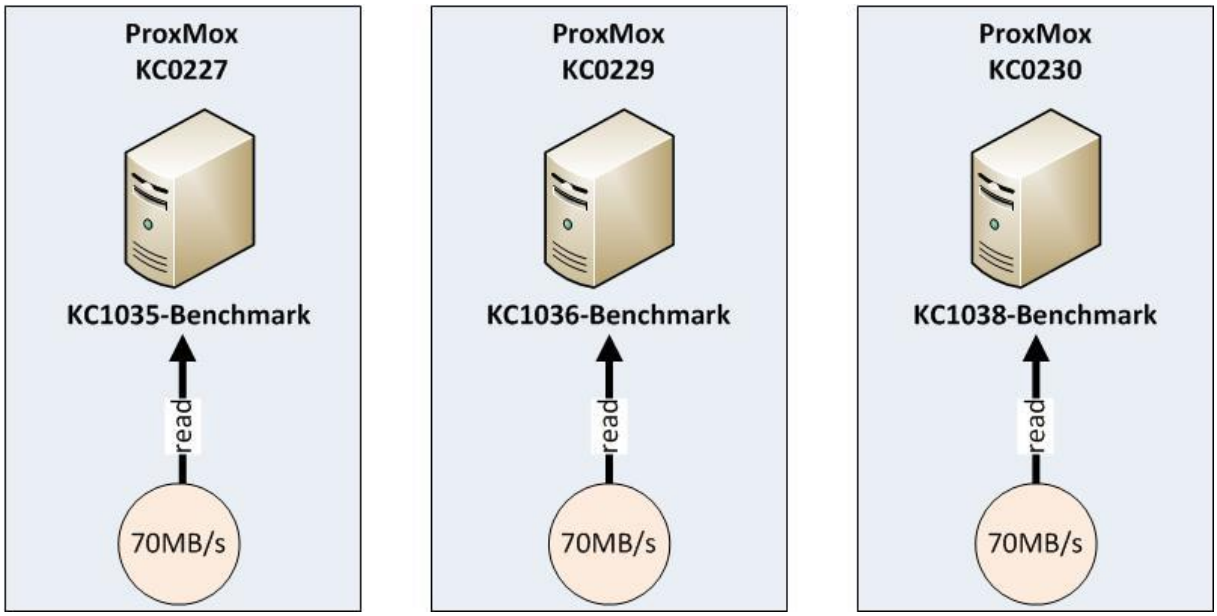
9.5.1 Leessnelheid

De leessnelheid van Ceph is minder dan een derde van de traditionele storage. Ceph leest van slechts één bron. De traditionele storage maakt gebruik van RAID 5 en is in staat van meerdere bronnen te lezen. Van de traditionele storage kunnen we echter stellen dat de leesperformance altijd gedeeld dient te worden door het aantal hosts dat aan het lezen is. Van Ceph kunnen we stellen dat iedere hosts van één initiële bron leest, al dan niet over het netwerk. Één van de ontwikkelaars van Ceph stelt dan ook: "Currently, Ceph does not provide any parallel reads functionality, which means that Ceph will always serve the read request from the primary OSD."[59].

[59] http://www.sebastien-han.fr/blog/2014/02/17/ceph-io-patterns-the-bad/

9.5.2 Schaalbaarheid leessnelheid Ceph

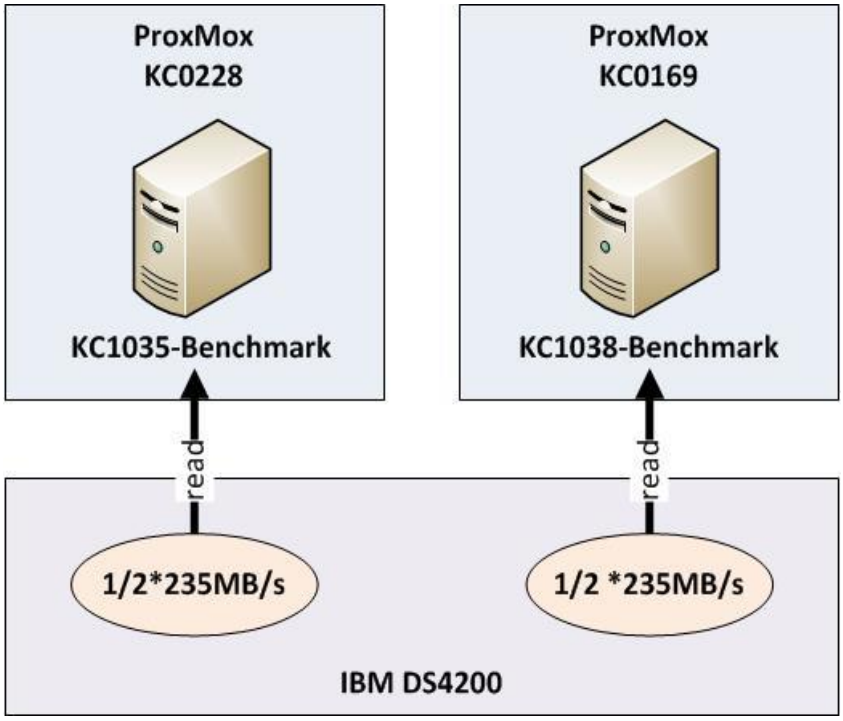
Indien hosts op verschillende Proxmox servers worden gevirtualiseerd mag deze leessnelheid lineair worden vermenigvuldigd. Drie hosts schrijven in theorie dus gezamenlijk met 210MB/s. Om de theorie kracht bij te zetten zijn deze benchmarks toch uitgevoerd. De testresulaten zijn terug te vinden in bijlage AA. Zoals aangegeven in hoofdstuk 0, bestaat de kans bij parallele benchmarks dat door afwijkingen in deviatie de benchmarks niet synchroon verlopen. Dit effect heeft dan ook plaatsgevonden tijdens het uitvoeren van de benchmarks. De benchmarkresulaten zijn dus minder betrouwbaar dan overige benchmarks die op een enkele host uitgevoerd zijn. In de praktijk lezen de hosts met ~165MB/s,espectievelijk 65MB/s, 50MB/s en 50MB/s.



Figuur 63 Theoretische benadering lineaire vermenigvuldiging leessnelheid op Ceph storage

9.5.3 Schaalbaarheid leessnelheid Traditionele storage

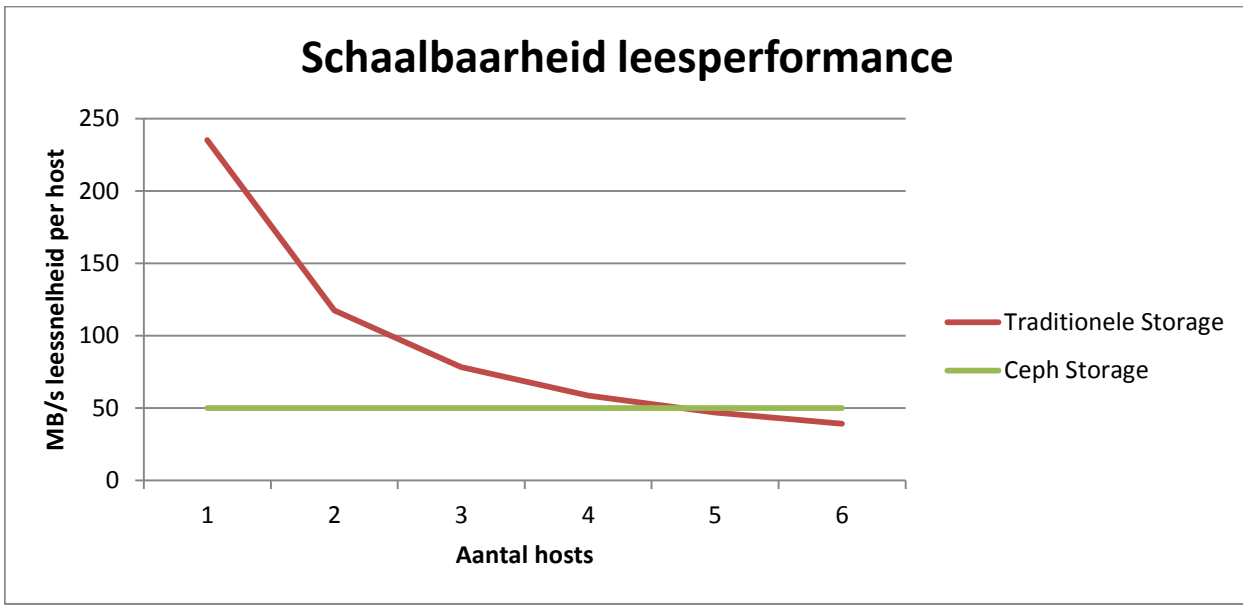
Indien hosts op verschillende Proxmox servers worden gevirtualiseerd mag de leessnelheid in theorie lineair worden gedeeld. Twee hosts schrijven in theorie dus gezamenlijk met 235MB/s. Om deze theorie kracht bij te zetten zijn deze benchmarks toch uitgevoerd. De testresulaten zijn terug te vinden in blijage AA. Zoals aangegeven in hoofdstuk 0, bestaat de kans bij parallele benchmarks dat door afwijkingen in deviatie de benchmarks niet synchroon verlopen. Dit effect heeft bij deze benchmark niet plaatsgevonden. In de praktijk lezen de hosts gezamenlijk met ~245MB/s. Respectievelijk 120MB/s en 125MB/s.



Figuur 64 Theoretische benadering lineaire deling leessnelheid op traditionele storage

9.5.4 (sub)conclusie leessnelheid

Indien de eerder opgedane meetresultaten lineair worden geëxtrapoleerd kunnen we onderstaande grafiek genereren. Uit deze grafiek is af te lezen dat de leessnelheid van Ceph vanaf 5 OSD hosts superieur is aan die van de traditionele storage.

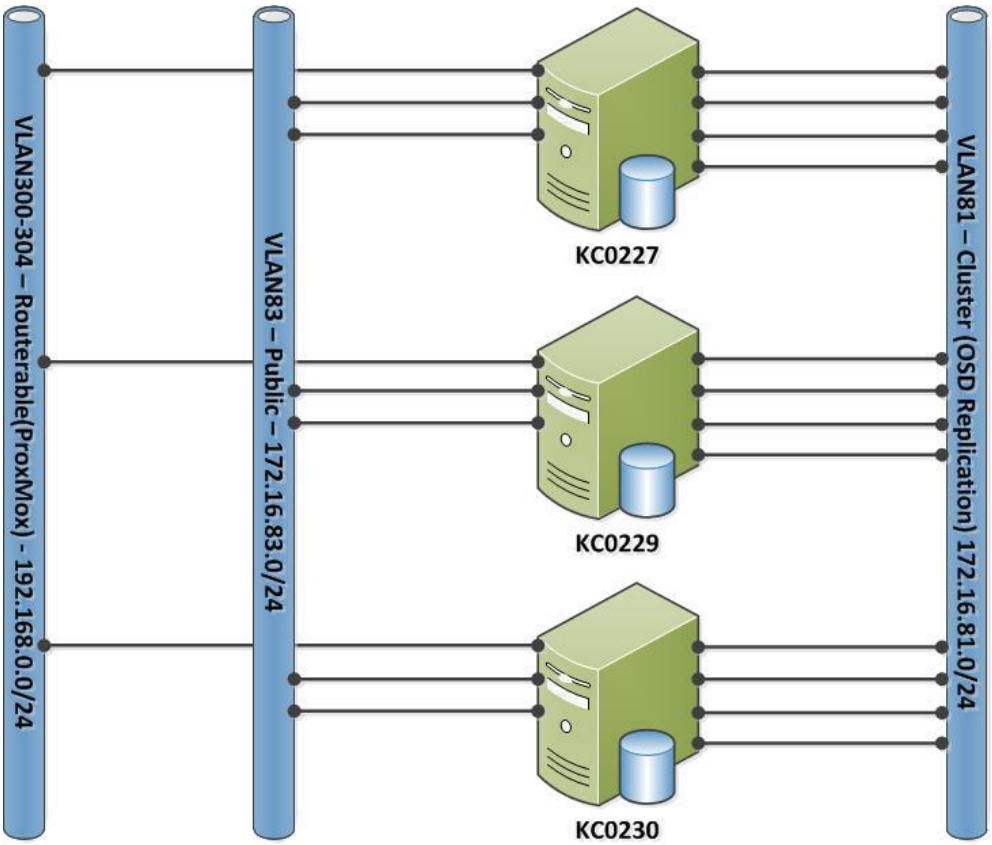


Figuur 65 Schaalbaarheid van de leesperformance

9.5.5 Schrijfsnelheid

De schrijfsnelheid van Ceph ligt op 90MB/s, van de traditionele storage ligt deze op 132MB/s(8.5.1). Een mogelijke oorzaak hiervoor kan worden gezocht in het netwerk, tijdens de literatuurstudie viel de volgende zin mij op: "*Ceph OSD Daemons handle data replication for the Ceph Clients. When Ceph OSD Daemons replicate data more than once, the network load between Ceph OSD Daemons easily* ***dwarfs*** *the network load between Ceph Clients and the Ceph Storage Cluster. This can introduce latency and create a performance problem.*[60]" Om dit uit te sluiten is een upgrade van het netwerk uitgevoerd. Het Ceph public-netwerk is opgewaardeerd naar 2Gb/s en het clusternetwerk is opgewaardeerd naar 4Gb/s.

[60] http://ceph.com/docs/master/rados/configuration/network-config-ref/



Figuur 66 Schematische weergave van het netwerk na de upgrade

Een ander opmerkelijk fenomeen is de lagere schrijfsnelheid van de individueel geteste traditionele disks ten opzichte van Ceph. De traditionele disk heeft een schrijfperformance van 80MB/s(Figuur 57), hierdoor lijkt het onmogelijk dat Ceph met 90MB/s(Figuur 54) gegevens weg kan schrijven. Zoals in het theoretisch kader is vastgelegd schrijft Ceph altijd naar de journal(2.7) en nooit direct naar het traditionele deel van de OSD.

Na de upgrade van het netwerk blijkt de Ceph schrijfsnelheid met 10% te zijn verhoogd, Ceph is nu in staat met 99MB/s (8.5.3) informatie weg te schrijven op disks. Dit komt dichter in de buurt van de 132MB/s (8.5.1) maar mogelijk is er nog meer uit te halen. Hiervoor is gekeken naar de verhouding tussen de traditionele en de solid state disks. Initieel wordt informatie weggeschreven op de solid state disk, later vindt replicatie plaats naar de traditionele disks. Indien het niet mogelijk is informatie vanaf de solid state tijdig weg te schrijven, houdt de OSD nieuwe schrijfverzoeken af. Dit zorgt voor een schrijfpatroon met pieken en dalen.

$$\text{Optimaal aantal journals per SSD} = \frac{\text{Solid State schrijfsnelheid}}{\text{Traditionale disk schrijfsnelheid}}$$

Indien we deze formule toepassen op onze infrastructuur komen we tot een optimaal aantal journals voor onze SSD-disk van 1,80.

$$1{,}80 = \frac{150}{83}$$

Er is gekozen om deze test uit te voeren met twee disks, een verhouding van één solid state per traditionele disk is kostentechnisch niet realistisch. Hierbij mogen we dan ook geen significante verbetering verwachten. In deze formule is geen rekening gehouden met de omvang van de solid

state disks. Op basis van de standaard waarde van “filestore max sync interval”[61] wordt informatie vanuit de journal om de 5 seconden weggeschreven naar traditionele disk. Met een maximale schrijfsnelheid van 150MB/s is het theoretisch mogelijk een journal van 750MB te gebruiken. Binnen de PoC zijn journal van 20GB gebruikt.
De schrijfsnelheid met de aangepaste journal ratio is 101MB/s (0), een verbetering van slechts 2%.

Tot slot is er een poging gedaan om de performance van Ceph op te hogen door het aantal replica’s te reduceren. Bij twee replica’s overtreft Ceph de schrijfperformance van de traditionele storage. Ceph-storage met twee replica’s is in staat 142MB/s (8.5.5) weg te schrijven ten opzichte van 132MB/s (8.5.1) op de traditionele storage. Op basis van het feit dat er met 1 replica snelheden van 212MB/s (8.5.5) worden gehaald is het netwerk uit te sluiten als performance bottleneck. Naar verwachting is de bottleneck terug te vinden in de performance van de PERC H200 controller, dit is echter niet evident.

9.5.6 (Sub)conclusie schrijfsnelheid

Er zijn een aantal verbeteringen toegepast op Ceph om de schrijfsnelheid van de traditionele storage te evenaren. Indien Ceph gebruik maakt van slechts 2 replica’s is het mogelijk de schrijfperformance van de traditionele storage te overtreffen. Dit heeft vanzelfsprekend consequenties voor de beschikbaarheid. Een andere mogelijkheid is het investeren in een betere controller en harddisks tezamen met een lager aantal journals per solid state drive.

9.6 Kosten

- Om Ceph gelijke performance te laten leveren dient geïnvesteerd te worden in additionele solid state drives om de journal ratio te optimaliseren.
- Het verhogen van de journal ratio zorgt voor een verlies in de totale opslagcapaciteit. Om dit te compenseren dient er geïnvesteerd te worden in traditionele schijven met een grotere capaciteit en een betere performance.
- Om deze beter performende disks af te kunnen handelen dient er geïnvesteerd te worden in een high performance controller.
- Indien er groei plaatsvindt is Ceph financieel wel rendabel. Vanaf 5 nodes is de leesperformance gelijk aan die van de traditionele storage. De addionele disks van de overige nodes zorgen voor afdoende capaciteit en een betere schrijfperformance.
- De investeringen in Fibre Channel-adapters en netwerk-adapter en netwerkpoorten zullen nagenoeg gelijk zijn.

Een mogelijk scenario om de kosten inzichtelijk te krijgen is als volgt. Op basis van eerder uitgevoerd performancemeetingen is gebleken dat Ceph financieel rendabel is vanaf 5 nodes. Binnen zowel de IST- als de SOLL-situatie waren al 2 servers aanwezig. Er dienen in de SOLL-situatie dus 3 servers toegevoegd te worden. De kosten van deze 3 additionele servers bedraagt €21.000,- (3*€7000,-). Aangezien het IBM DS4200 disk enclosure niet meer noodzakelijk is mag €35.000,- in mindering worden gebracht. Binnen dit scenario zou het gebruik van Ceph €14.000,- aan hardware kosten besparen. Indien Ceph op grote schaal toegepast wordt binnen het afstudeerbedrijf is de in het afstudeerplan genoemde €500.000,- reëel.

[61] http://ceph.com/docs/master/rados/configuration/filestore-config-ref/

10 Conclusie

Met de technische realisatie van dit Proof of Concept is aangetoond dat het is mogelijk met relatief eenvoudige hardware traditionele storage in de basis te vervangen door Ceph. Hiermee biedt Ceph een potentieel alternatief voor traditionele storage. De geautomatiseerde installatie van de Proof of Concept omgeving stelt het afstudeerbedrijf in staat het binnen deze scriptie beschreven onderzoek te reproduceren. Ook is het mogelijk nieuwe testen en onderzoeken uit te voeren op cloud storage op basis van Ceph.

Uit die onderzoek naar de aspecten schaalbaarheid, beschikbaarheid, security, beheerbaarheid performance en de kosten van Ceph cloud storage zijn de volgende punten naar voor gekomen:

- Ceph is goed schaalbaar en met deze schaalbaarheid komt een noodzakelijke toename in leessnelheid(9.1)(9.5). De is mogelijk additionele Ceph nodes en Ceph storage toe te voegen.
- Ceph is in staat uitval van individuele componenten op te vangen, dit komt de beschikbaarheid ten goede (9.2). Ceph zorgt automatisch voor replicatie van de weggevallen informatie.
- Door het toepassen van de door Ceph aanbevolen netwerkscheiding ontstaat een veilige cloud storage oplossing (9.3).
- Met de gratis Calamari-interface is Ceph beheersbaar, het niet aan kunnen maken van OSD´s is echter een grote tekortkoming (0).
- Ceph is niet in staat de schrijfsnelheid van het traditionele cluster te evenaren indien er gebruikt wordt gemaakt van 3 replica’s (9.5). Bij het gebruik van slechts 2 replica’s is de schrijfsnelheid van Ceph superieur aan die van de binnen de PoC gebruikte traditionele storage (9.5).
- Met geringe investeringen is het mogelijk bestaande componenten efficiënter in te zetten (9.6). Investeringen in solid state drives en netwerkcomponenten leveren een grote bijdrage aan de verbetering van de performance(9.5).

Alleen indien de omgeving opgeschaald wordt naar 5 nodes en het aantal replica’s terug gebracht wordt naar 2 is Ceph een passend alternatief voor de IBM DS4200 disk enclosure(9.5). Het handmatig aanmaken van OSD´s dient dan wel als voldongen feit geaccepteerd te worden.

11 Reflectie

In dit hoofdstuk wordt gereflecteerd op de volgende zaken aspecten van het afstuderen:

- hoe ik het afstuderen had willen doen
- hoe het afstuderen is verlopen
- wat beter had gekund

Planning

De gekozen opdracht sluit in hoofdlijnen nauw aan bij mijn kennisgebied. Open Source, beveiliging, system engineering en deployment zijn dagelijkse kost. Het storage deel van de afstudeeropdracht was nagenoeg in het geheel nieuw voor mij, in het verleden droegen collega's zorg voor dit aspect van de infrastructuur.

In het afstudeerplan zijn volgens de planning 10 dagen gereserveerd voor een literatuurstudie. In deze periode moet ik onderzoeken wat de mogelijkheden zijn van storage in het algemeen en cloud storage. Deze kennis ga ik vervolgens gebruiken bij het uitvoeren van een productselectie op basis van een MoSCoW-analyse.

Het product dat voort komt uit deze productselectie moet vervolgens technische gerealiseerd worden voordat er onderzoek op deze omgeving uitgevoerd kan worden. De omgeving had als doel het onderzoeken van de aspecten schaalbaarheid, beschikbaarheid, security, beheerbaarheid, performance en kosten. Het technisch realiseren van de omgeving moest plaats vinden op basis van de ontwikkelmethode Scrum, dit deel van het afstuderen had een duidelijk proces wat dan ook uitvoerig is vastgelegd.

Verloop

Bij aanvang van de literatuurstudie realiseerde ik me dat er een tweede literatuurstudie noodzakelijk was om de juiste kennis en kunde op te doen van het product dat voort kwam uit de productselectie. Hiervoor had ik geen tijd gereserveerd in het afstudeerplan. De 10 werkdagen zijn dan ook verdeeld over beide literatuurstudies en aangevuld met weekenden en additionele late avonduurtjes. Daarnaast is er vanuit het afstudeerbedrijf geen tot weinig tijd vrij gemaakt voor het uitvoeren van de afstudeeropdracht.

Na het inleveren van het afstudeerplan begon bij mij de verwarring over wat is een "architectuur". Ik zag zelf een netwerktekening als een architectuur, maar wist niet of de HHS hier de zelfde visie op had. Ik heb hoog ingezet door het volgen van een cursus ArchiMate Enterprise architectuur en het behalen van bijbehorende certificering. Uiteindelijk is deze kennis toegepast in het realiseren van een IST- en SOLL-architectuur op basis van ArchiMate. De productselectie was een lang en moeizaam traject dat overigens eveneens uit de planning liep.

Het realiseren van een technische infrastructuur inclusief Ceph cloud storage was een pittige uitdaging, die mij naast veel plezier ook veel ergernissen en frustraties heeft opgeleverd. Alle handelingen die noodzakelijk waren om tot een werkende omgeving te komen zijn tot in detail vastgelegd in de de sprint-verslagen. Ook de tegenslagen zijn hierin vastgelegd als impediment. Na de technische realisatie is uitvoering gegeven aan een onderzoek naar de aspecten schaalbaarheid,

beschikbaarheid, security, beheerbaarheid, performance en kosten van Ceph. Vooral de performance onderzoeken waren langdurig en zorgde herhaaldelijk voor de noodzaak tot aanvullend onderzoek.

Tijdens het afstuderen hebben er een bespreking concept en een tussentijds assessment plaatsgevonden. De resultaten hiervan zijn terug te vinden in respectievelijk bijlage NN en bijlage OO.

Evaluatie

Bij het maken van een projectplanning(afstudeerplan) houd ik in het vervolg rekening met de noodzaak om zich in de verschillende stadia van het project in te lezen in de materie. De tweede literatuurstudie die ik over het hoofd heb gezien heeft mij vanaf het eerste moment van het afstuderen voor veel stress gezorgd. Om deze stress in de toekomst binnen de perken te houden moeten duidelijk afspraken worden gemaakt met het afstudeerbedrijf over de beschikbaarheid voor projecten. En dien ik meer tijd te investeren in het maken van projectplanning. Ook ga ik deze planning in de toekomst laten toetsen door derde. Ik heb bijzonder veel kennis opgedaan tijdens deze intensieve literatuurstudies.

Hoewel de ArchiMate cursus zeer leerzaam wens ik voor aanvang van nieuwe projecten, definities (zoals Architectuur), voor het opleveren van het projectplan inzichtelijk te krijgen in samenspraak met de opdrachtgever. De afstudeerder blijft gebruik maken van deze standaard van het afstudeerbedrijf.

De productselectie was lang en moeizaam omdat twee producten aan nagenoeg alle eisen en wensen van de opdrachtgever voldeden. Hierdoor was het noodzakelijk alle Must, Should en Could haves te beschrijven en onderbouwen. Ik ben dan ook erg te spreken over de MoSCoW methode en wens deze in de toekomst te blijven gebruiken.

Het technisch realiseren van een nieuwe technologie brengt grote risico’s met zich mee. Het feit dat het gelukt is, is te wijten aan mijn brede technische kennis en doorzettingsvermogen. Voor het afstuderen was het nemen van een risico van dusdanige omvang niet verstandig. Vanuit het perspectief van de klant, opdrachtgever en de innovatieve aard van de afdeling van het afstudeerbedrijf wordt het nemen van dit risico omarmd. De door mij gekozen Scrum ontwikkelmethode heeft bijgedragen aan het succesvol uitvoeren van de technische realisatie. Het gebruik van Scrum heeft ervoor gezorgd dat de opdrachtgever om de 7 werkdagen een functioneel product opgeleverd kreeg. Ik ga de Scrum methode implementeren op de afdeling waarop ik leiding geef.

Het onderzoek naar de aspecten schaalbaarheid, beschikbaarheid, security, beheerbaarheid, performance en kosten heeft mijn in aanraking gebracht met een breed aantal methodieken. De variantie in methodieken en het werken met de binnen het afstuderen gerealiseerde technische infrastructuur worden door de afstudeerder als zeer leuk ervaren. Er is op dit vlak in het bijzonder veel kennis opgedaan van het uitvoeren van gestandaardiseerde benchmarks. De gebruikte benchmarkmethodiek zal ik in de toekomst veelvuldig toepassen.

12 Bibliografie

Arias, T. (2012). *Cloud Computing Storage Handbook.* Brisbane: Emereo Publishing.
Collaris, R.-A. (2012). *RUP op Maat.* Den Haag: SDU Uitgevers.
Farley, M. (2013). *Rethinking enterprise storage.* Redmond: Microsoft Corporation.
Gnanasundaram, S. (2012). *Information Storage and Management.* Indianapolis: John Wiley & Sons, Inc.
Jonker, R. (2012). *Wolkbreuk.* Amsterdam: Van Gennep.
Lowe, S. D. (2014). *Software-Defined Storage.* Hoboken: John Wiley & Sons, Inc.
Oost, H. (2011). *Een onderzoek rapporteren.* Amersfoort: ThiemeMeulenhoff.
Oost, H. (2012). *Een onderzoek uitvoeren.* Amersfoort: ThiemeMeulenhoff.
Oost, H. (2012). *Een onderzoek voorbereiden.* Amersfoort: ThiemeMeulenhoff.
Schulz, G. (2011). *Cloud and Virtual Data Storage Networking.* Boca Raton: CRC Press.
Troppens, U. (2009). *Storage Networks Explained.* Heidelberg: John Wiley & Sons, Inc.
Van Haren Publishing. (2013). *ArchiMate 2.1 Specification.* Zaltbommel: Van Haren Publishing.
Verheyen, G. (2013). *Scrum A Pocket Guide.* Zaltbommel: Van Haren Publishing.
Verhoeven, N. (2011). *Wat is onderzoek?* Den Haag: Boom Lemma.
Yoder, A. G. (2013). *SNIA Dictionary 2013 European Edition.* London: Storage Networking Industry Association Europe Ltd.
Yuan, D. (2013). *Computation and storage in the cloud.* London: Elsevier.

13 Afkortingen

COTS[62]	Commercial off the shelf
PoC[63]	Proof of Concept
LVM[64]	Logical Volume Manager
LV	Logical Volume
RAID[65]	Redundant Array of Inexpensive Disks
OSD[66]	Object Storage Device
FC-AR[67]	Fibre Channel Arbitrated Loop
ITIL[68]	Information Technology Infrastructure Library
CI[69]	Configuration Item
CMDB[70]	Configuration Management DataBase

[62] http://en.wikipedia.org/wiki/Commercial_off-the-shelf
[63] http://en.wikipedia.org/wiki/Proof_of_concept
[64] http://en.wikipedia.org/wiki/Logical_Volume_Manager_%28Linux%29
[65] http://en.wikipedia.org/wiki/RAID
[66] http://en.wikipedia.org/wiki/Object_storage_device
[67] http://en.wikipedia.org/wiki/Arbitrated_loop
[68] http://en.wikipedia.org/wiki/ITIL
[69] http://en.wikipedia.org/wiki/Configuration_item
[70] http://en.wikipedia.org/wiki/CMDB

A. Bijlage Certificaat ArchiMate 2

The Open Group Open Certified

ArchiMate® 2 Foundation

This is to certify that

Marco Stroosnijder

has successfully met the requirements of the ArchiMate 2 Certification for People program.

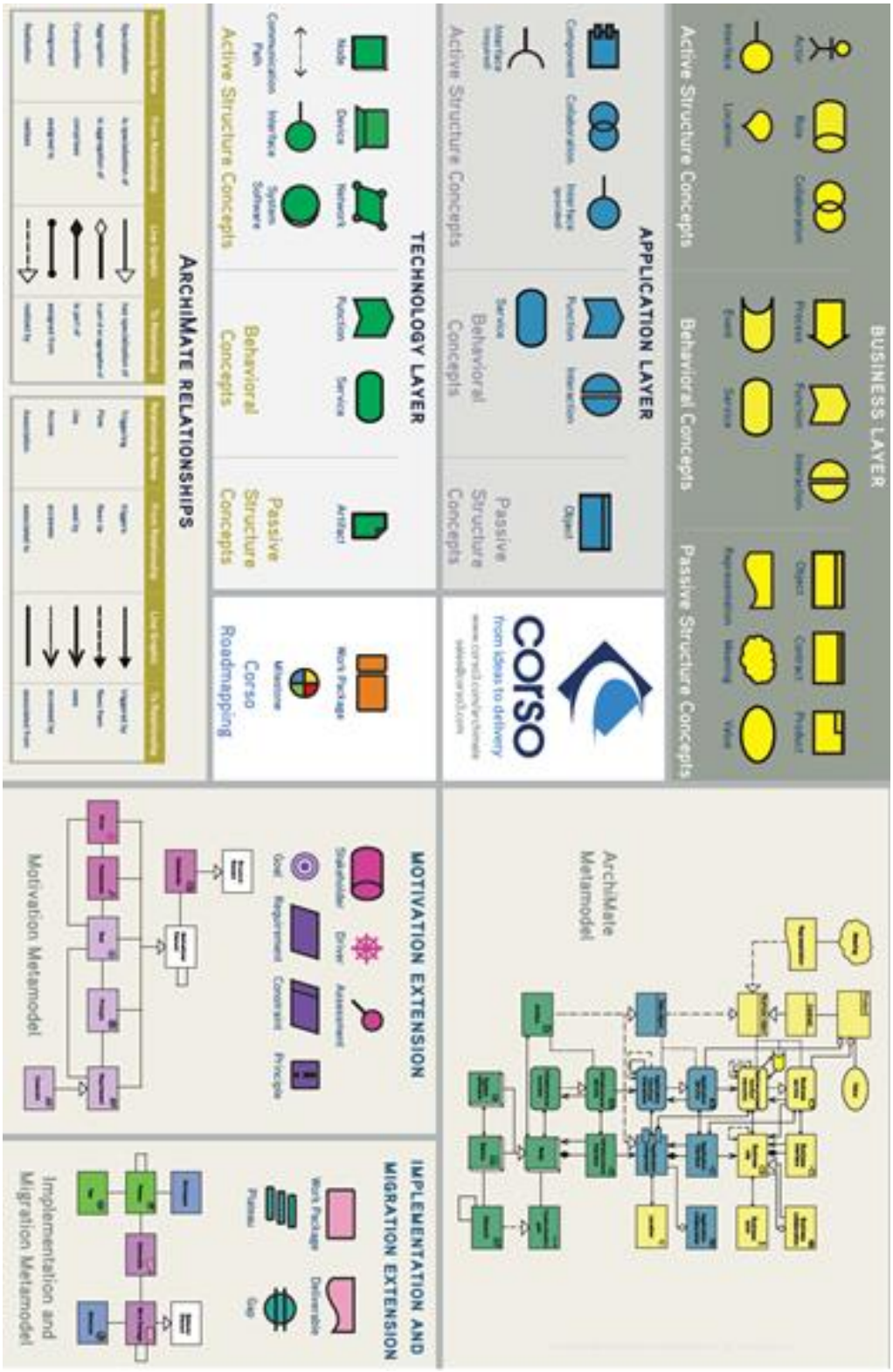
Date registered: 4 December 2014

Allen Brown

Allen Brown, President and CEO, The Open Group

The Open Group Certification mark is a trademark and ArchiMate is a registered trademark of The Open Group. The certification logo may only be used on or in connection with those products, persons, or organizations that have been certified under this program. The directory of certified individuals may be viewed at http://www.opengroup.org/archimate/cert/certified-individuals

B. Bijlage ArchiMate Referentie model



C. Bijlage Preseed file

/var/www/autoinstall/OS/debian/wheezy/wheezy.cfg

```
d-i mirror/country string manual
d-i mirror/http/hostname string kc1000
d-i mirror/http/directory string /linux/ftp.nl.debian.org/debian/
d-i mirror/suite string wheezy
d-i mirror/http/proxy string

d-i console-keymaps-at/keymap select us
d-i keyboard-configuration/xkb-keymap select us

d-i passwd/root-login boolean false

partman-target partman/mount_style select label
d-i partman-auto/disk string /dev/sda
d-i partman-auto/method string lvm
d-i partman-auto/purge_lvm_from_device boolean true
d-i partman-lvm/confirm_nooverwrite boolean true
d-i partman/confirm_nooverwrite boolean true
d-i partman-auto/purge_lvm_from_device boolean true
d-i partman-lvm/device_remove_lvm boolean true
d-i partman-lvm/device_remove_lvm_span boolean true
d-i partman-lvm/confirm boolean true
d-i partman/confirm_write_new_label boolean true
d-i partman/choose_partition select Finish partitioning and write changes to disk
d-i partman/confirm boolean true
d-i partman-auto/init_automatically_partition select Guided - use entire disk and set up LVM
```

```
d-i partman-auto/expert_recipe string \
 boot-root :: \
 40 300 200 ext3 \
 $primary{ } \
 $bootable{ } \
 method{ format } format{ } \
 use_filesystem{ } filesystem{ ext3 } \
 mountpoint{ /boot } \
 . \
 15000 15000 100 ext4 \
 $lvmok{ } \
 method{ format } format{ } \
 use_filesystem{ } filesystem{ ext4 } \
 mountpoint{ / } \
 . \
 2000 2000 2000 linux-swap \
 $lvmok{ } \
 method{ swap } format{ } \
 .

# Amount of volume group to use for guided partitioning:
partman-auto-lvm partman-auto-lvm/guided_size string 28 GB

d-i clock-setup/utc boolean true
d-i time/zone string Europe/Amsterdam
d-i apt-setup/restricted boolean true

d-i passwd/user-fullname string kcbeheer User
d-i passwd/username string kcbeheer
d-i passwd/user-id string 999
d-i passwd/user-password-crypted password CRYPTED-PASSWORD-GOES-HERE
d-i grub-installer/only_debian boolean true
pkgsel pkgsel/update-policy select none

# Encrypt your home directory?
user-setup-udeb user-setup/encrypt-home boolean false
#dont select a desktop
tasksel tasksel/first multiselect standard
d-i pkgsel/language-pack-patterns string
d-i pkgsel/install-language-support boolean false
d-i pkgsel/include string openssh-server

d-i finish-install/reboot_in_progress note
d-i cdrom-detect/eject boolean false
d-i preseed/late_command string \
 in-target mkdir /root/.ssh/ ;\
 echo "SSH KEY GOES HERE" >/target/root/.ssh/authorized_keys
```

D. Bijlage Ansible role common

```
---

#
# user/group management
#
- name: add group kcbeheer
 group: name=kcbeheer gid=1000 state=present

#
#
#
- name: add user kcbeheer
 user:
 name=kcbeheer
 uid=1000
 password=' THERE ONCE WAS A PASSWORD HASH HERE '
 comment='Trisc User'
 group=kcbeheer
 groups=sudo
 home=/home/kcbeheer
 shell=/bin/bash
 state=present

#
#
#
- name: set permissions homedir kcbeheer
 file: path=/home/kcbeheer owner=kcbeheer group=kcbeheer mode=0700 state=directory

#
#
#
- name: configure autorized_keys
 template: src=authorized_keys.j2 dest=/root/.ssh/authorized_keys mode=0644 owner=root
group=root
 when: not "{{ansible_hostname }}" in ["KC0227-MON", "KC0229-MON", "KC0230-MON", "KC0424-
MON"]

#
# motd
#
- name: configure motd
 template: src=motd.j2 dest=/etc/motd owner=root group=root mode=0600
```

```
#
# hostname
#
- name: configure hostname
 template: src={{ item }} dest=/etc/hostname mode=0644 owner=root group=root
 first_available_file:
 - "{{ansible_hostname }}.hostname.j2"
 - hostname.j2

#
#
- name: Make sure the hostname is set
 shell: hostname `cat /etc/hostname`

#
# hosts
#
- name: configure the hosts file
 template: src=hosts.j2 dest=/etc/hosts mode=0644 owner=root group=root

#
# sudoers
#
- name: configure the sudoers file
 template: src=sudoers.j2 dest=/etc/sudoers mode=0440 owner=root group=root

#
# sysctl
#
- name: configure sysctl
 sysctl: name=net.ipv6.conf.all.disable_ipv6 value=1 reload=yes state=present
 when: not "{{ ansible_hostname }}" == "kc1016"

#
# authentication
#
- name: configure krb5.conf
 template: src=krb5.conf.j2 dest=/etc/krb5.conf mode=0644 owner=root group=root

#
#
#
- name: configure ldap.conf
 template: src=ldap.conf.j2 dest=/etc/ldap.conf mode=0644 owner=root group=root
```

```
#
#
#
- name: add homedir nfs mount to the fstab (/export/homes)
 lineinfile: dest=/etc/fstab line="#nfs:/export/homes /home/users nfs nfsvers=3,defaults 0 0"

#
#
#
- name: Create home directory
 file: path=/home/users state=directory

#
#
#
- name: create sssd directory
 file: path=/etc/sssd owner=root group=root mode=0755 state=directory


#
# SSSD
#
- name: configure /etc/sssd/sssd.conf
 template: src=sssd.conf.j2 dest=/etc/sssd/sssd.conf mode=0600 owner=root group=root
```

E. Bijlage Ansible role Debian

```
---
#
# /etc/locale.gen
#
- name: place locales file
 template: src=locale.gen.j2 dest=/etc/locale.gen mode=0644 owner=root group=root

#
# Generate locales
#
- name: generate locales
 shell: locale-gen
 changed_when: False

#
# /etc/apt/sources.list
#
- name: Install Debian repo
 template: src=sources.list.j2 dest=/etc/apt/sources.list mode=0644 owner=root group=root

#
# Install packages
#
- name: install a bunch of needed packages
 apt: pkg={{item}} state=present update_cache=yes force=yes
 with_items:
 - ntp
 - postfix
 - mailutils
 - rsync
 - snmpd
 - sssd
 - resolvconf
 - nfs-common
 - libpam-krb5
 - krb5-user
 - ntpdate

#
# configure postfix to use a relayhost
#
- name: configure mail server
 template: src=main.cf.j2 dest=/etc/postfix/main.cf owner=root group=root mode=0644
```

```
#
# restart postfix
#
- name: start ntp
 service: name=ntp state=restarted enabled=yes
 changed_when: False

#
# nsswitch
#
- name: configure nsswitch.conf
 template: src=nsswitch.conf.j2 dest=/etc/nsswitch.conf mode=0644 owner=root group=root

#
#
#
#- name: make sure /home/users is mounted
# shell: mount -a -t nfs
# changed_when: False

#
# NTP
#
- name: configure ntp.conf
 template: src=ntp.conf.j2 dest=/etc/ntp.conf mode=0644 owner=root group=root

#
#
#
#- name: start ntp
# service: name=ntp state=restarted enabled=yes
# changed_when: False

#
# SNMPD
#
- name: configure /etc/snmp/snmpd.conf
 template: src=snmpd.conf.j2 dest=/etc/snmp/snmpd.conf mode=0600 owner=root group=root

#
#
#
- name: start snmpd
 service: name=snmpd state=restarted enabled=yes
 changed_when: False
```

F. Bijlage Ansible role Proxmox

```
---

#
# user management
#
- name: set root password
 user:
 name=root
 password=' THERE ONCE WAS A PASSWORD HASH HERE '

#
# divert Proxmox enterprise repo
#
- name: divert Proxmox enterprise repo
 shell: dpkg-divert --divert /root/pve-enterprise.list --rename /etc/apt/sources.list.d/pve-enterprise.list
 changed_when: False

#
# Install bridge util and bonding package
#
- name: Install bridge package
 apt: pkg={{item}} state=present update_cache=yes force=yes
 with_items:
 - bridge-utils
 - ifenslave-2.6

#
# Add bonding to /etc/modules
#
- name: load the bondig kernel module on boot
 lineinfile: dest=/etc/modules line="bonding"

#
# modprobe
#
- name: load the bonding kernel module
 shell: modprobe bonding
 changed_when: False

#
# /etc/network/interfaces
#
- name: Configure network interfaces
 template: src={{item}} dest=/etc/network/interfaces mode=0644 owner=root group=root
 first_available_file:
 - "{{ansible_hostname}}.interfaces.j2"
 tags: networking
```

```
#
# restart networking service
#
- name: restart networking
 service: name=networking state=restarted enabled=yes
 changed_when: False
 tags: networking

#
# /etc/resolv.conf
#
- name: Configure resolvance
 template: src=resolv.conf.j2 dest=/etc/resolv.conf mode=0644 owner=root group=root

#
#
#
- name: Add Proxmox repository key
 shell: wget -O- "http://download.Proxmox.com/debian/key.asc" | apt-key add -
 changed_when: False

#
# /etc/apt/sources.list.d/Proxmox.list
#
- name: Add Proxmox repository
 template: src=Proxmox.list.j2 dest=/etc/apt/sources.list.d/Proxmox.list mode=0640 owner=root
group=root

#
#
#
- name: install a bunch of needed packages
 apt: pkg={{item}} state=present update_cache=yes force=yes
 with_items:
 - Proxmox-ve-2.6.32
```

G. Bijlage Ansible role Ceph

```
---

# Ceph PASSWORD = THERE ONCE WAS A PASSWORD HERE

#
# group management
#
- name: add group ceph
 group: name=ceph gid=1234 state=present

#
# user management
#
- name: add user ceph
 user:
 name=ceph
 uid=1234
 password='THERE ONCE WAS A PASSWORD HASH HERE'
 comment='Ceph Cluster User'
 group=ceph
 home=/etc/ceph
 shell=/bin/bash
 state=present

#
# correct incorrect Ceph configuration dir permissions
#
- name: set permissions Ceph configuration
 file: path=/etc/ceph owner=root group=ceph mode=0770 state=directory

#
#
#
- name: generate a ssh key pair for ceph user
 connection: local
 command: ssh-keygen -t rsa -N "" -f
"/etc/ansible/roles/ceph/templates/{{ansible_hostname}}.id_rsa"
 args:
 creates: "/etc/ansible/roles/ceph/templates/{{ansible_hostname}}.id_rsa"

- name: create authorized_keys
 connection: local
 shell: cat /etc/ansible/roles/ceph/templates/*.id_rsa.pub > /etc/
ansible/roles/ceph/templates/authorized_keys
 changed_when: False
```

```
#
# correct incorrect homedirectory permissions
#
- name: set permissions on ceph .ssh
 file: path=/etc/ceph/.ssh owner=ceph group=ceph mode=0700 state=directory

#
#
#
- name: disable StrictHostKeyChecking
 lineinfile: dest=/etc/ceph/.ssh/config line="StrictHostKeyChecking no" state=present create=yes
mode=0600 owner=ceph group=ceph
 tags: deploy

#
# /etc/ceph/.ssh/id_rsa
#
- name: Install id_rsa
 template: src="{{ ansible_hostname}}.id_rsa" dest=/etc/ceph/.ssh/id_rsa mode=0600 owner=ceph
group=ceph

#
# /etc/ceph/.ssh/id_rsa.pub
#
- name: Install id_rsa.pub
 template: src="{{ansible_hostname}}.id_rsa.pub" dest=/etc/ceph/.ssh/id_rsa.pub mode=0644
owner=ceph group=ceph

#
# /etc/ceph/.ssh/authorized_keys
#
- name: Install authorized_keys
 template: src="authorized_keys" dest=/etc/ceph/.ssh/authorized_keys mode=0644 owner=ceph
group=ceph

#
#
#
- name: Add ceph repository key
 shell: wget -q -O- 'https://ceph.com/git/?p=ceph.git;a=blob_plain;f=keys/release.asc' | apt-key add -
 changed_when: False

#
#
#
- name: Add saltstack repository key
 shell: wget -q -O- "http://debian.saltstack.com/debian-salt-team-joehealy.gpg.key" | apt-key add -
 changed_when: False
```

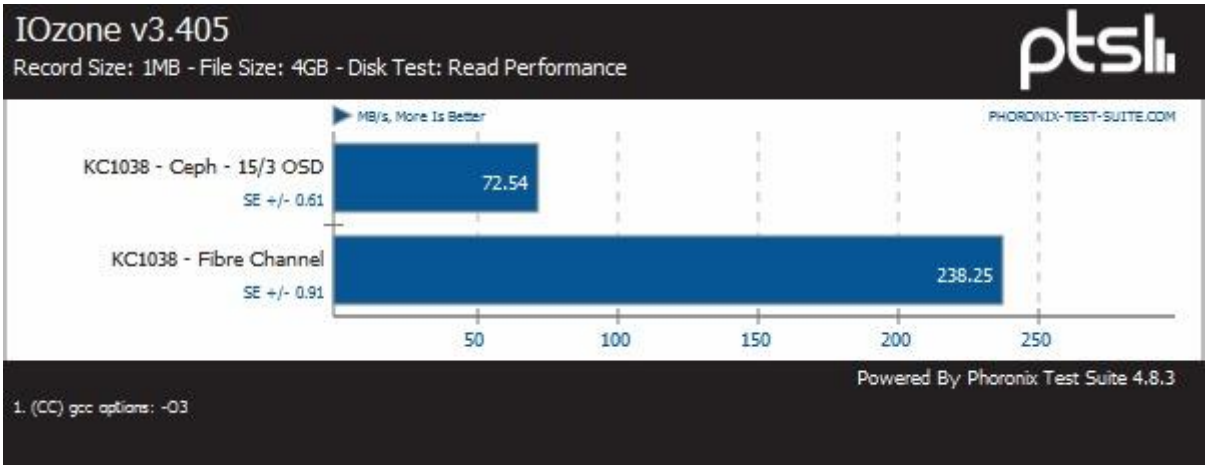
```
#
# /etc/apt/sources.list.d/ceph.list
#
- name: Add ceph repository
 template: src=ceph.list.j2 dest=/etc/apt/sources.list.d/ceph.list mode=0644 owner=root group=root

#
# /etc/sudoers.d/ceph_sudo.conf
#
- name: Install Ceph sudoers file
 template: src=ceph_sudo.j2 dest=/etc/sudoers.d/ceph mode=0440 owner=root group=root

#
# /etc/apt/sources.list.d/saltstack.list
#
- name: Install saltstack repo
 template: src=saltstack.list.j2 dest=/etc/apt/sources.list.d/saltstack.list mode=0644 owner=root group=root

#
# Install Ceph packages
#
- name: Install Ceph packages
 apt: pkg={{item}} state=present update_cache=yes force=yes
 with_items:
 - ceph-deploy
 - ceph
 - gdisk

#
# Create Initial Cluster
#
- name: Install Initial Cluster
 shell: su - ceph bash -c "ceph-deploy new --fsid 8c860ebb-767c-4f82-94bc-dd5400120dc2 --cluster-network 172.16.81.0/24 --public-network 172.16.83.0/24 KC0227-MON.afstudeer.org KC0229-MON.afstudeer.org KC0230-MON.afstudeer.org KC0424-MON.afstudeer.org"
 when: ( "{{ansible_hostname }}" == "KC0424-MON" )
 args:
 creates: "/etc/ceph/ceph.conf"

#
# Install Ceph software
#
- name: Install Ceph software
 shell: su - ceph bash -c "ceph-deploy install --no-adjust-repos KC0227-MON.afstudeer.org KC0229-MON.afstudeer.org KC0230-MON.afstudeer.org KC0424-MON.afstudeer.org"
 when: ( "{{ansible_hostname }}" == "KC0424-MON" )
```

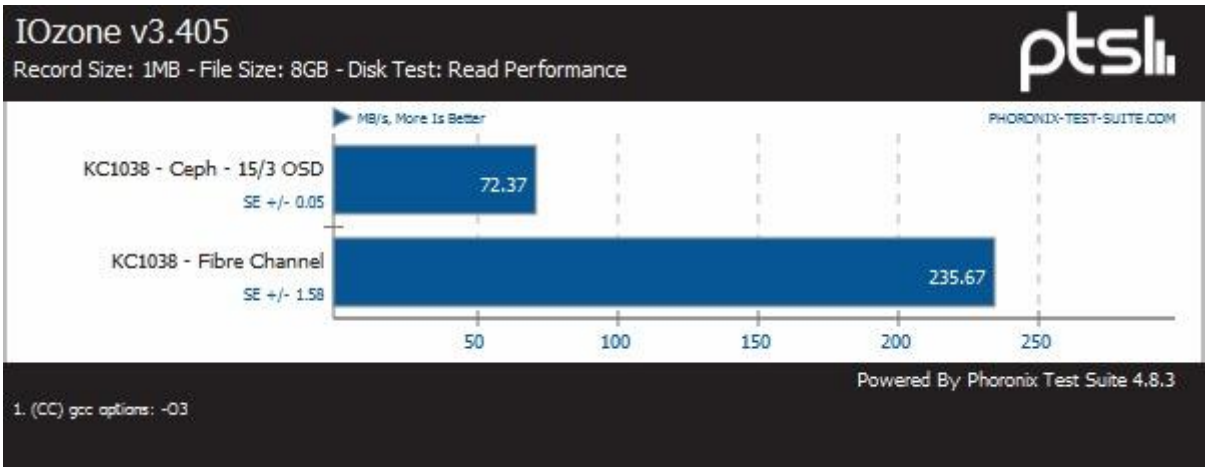
```
#
# Install Ceph monitor nodes
#
- name: Installation of monitor components
 shell: su - ceph bash -c "ceph-deploy mon create-initial"
 when: ( "{{ansible_hostname }}" == "KC0424-MON" )
 args:
 creates: "/etc/ceph/ceph.client.admin.keyring"

#
# Gather Ceph keys
#
- name: Gather Ceph keys
 shell: su - ceph bash -c "ceph-deploy gatherkeys KC0230-MON KC0227-MON KC0229-MON KC0424-
MON"
 args:
 creates: "/etc/ceph/ceph.bootstrap-osd.keyring"

#
#
#
- name: allow ceph user to use the admin keyring
 file: path=/etc/ceph/ceph.client.admin.keyring owner=root group=ceph mode=0640 state=file

#
# Install Ceph admin service
#
- name: Install Ceph admin
 shell: su - ceph bash -c "ceph-deploy admin KC0424-MON.afstudeer.org KC0227-MON.afstudeer.org
KC0229-MON.afstudeer.org KC0230-MON.afstudeer.org"
 when: ( "{{ansible_hostname }}" == "KC0424-MON" )
 changed_when: False
 tags: debugin

#
# create / correct /etc/pve/priv/ceph
#
- name: set permissions Proxmox ceph dir
 file: path=/etc/pve/priv/ceph/ owner=root group=www-data mode=0700 state=directory

#
# share ceph key with Proxmox
#
- name: share ceph key with Proxmox
 shell: cp /etc/ceph/ceph.client.admin.keyring /etc/pve/priv/ceph/PROXMOX-CEPH.keyring
 args:
 creates: "/etc/pve/priv/ceph/PROXMOX-CEPH.keyring"
```

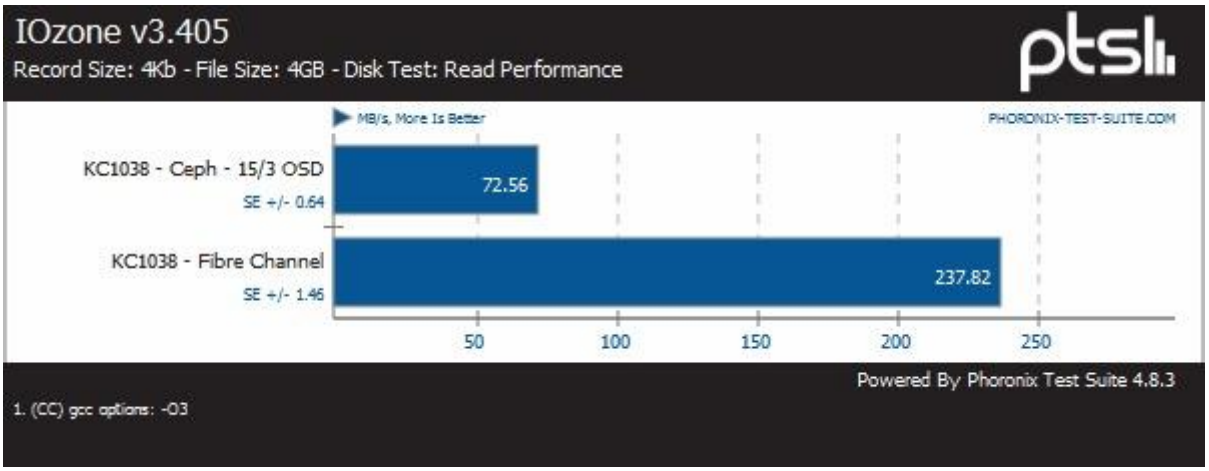
```
#
# /etc/apt/sources.list.d/calamari.list
#
- name: Install calamari repo
 template: src=calamari.list.j2 dest=/etc/apt/sources.list.d/calamari.list mode=0644 owner=root group=root
 tags: calamari


#
# Install Calamari dependencies
#
- name: Install Calamari dependencies
 apt: pkg={{item}} state=present update_cache=yes force=yes
 with_items:
 - salt-minion
 - ipvsadm
 - python-mock
 - python-configobj
 - cdbs
 - diamond
 tags: calamari

#
# Create required symlink
#
- name: Link ipvsadm
 file: src=/sbin/ipvsadm dest=/usr/bin/ipvsadm state=link
 tags: calamari

#
# minion configuration
#
- name: Configure calamari
 template: src=calamari.conf.j2 dest=/etc/salt/minion.d/calamari.conf mode=0644 owner=root group=root
 tags: calamari

#
# create Diamond configuration
#
- name: create Diamond configuration
 shell: cp -a /etc/diamond/diamond.conf.example /etc/diamond/diamond.conf
 changed_when: False
 tags: calamari
 args:
 creates: "/etc/diamond/diamond.conf"
```

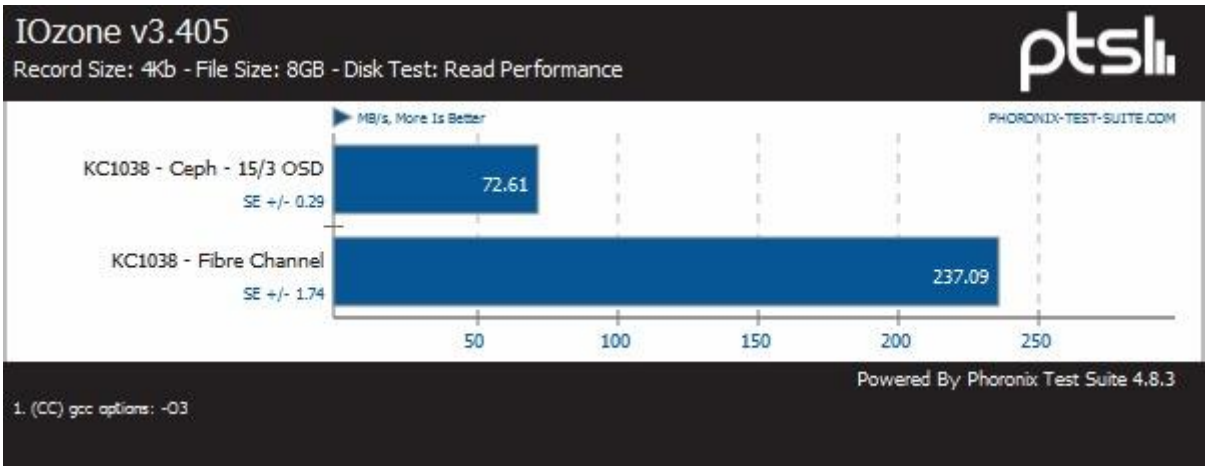
```
#
# restart Diamond
#
- name: start diamond
 service: name=diamond state=restarted enabled=yes
 changed_when: False
 tags: calamari

#
# minion configuration
#
- name: Configure calamari
 template: src=calamari.conf.j2 dest=/etc/salt/minion.d/calamari.conf mode=0644 owner=root group=root
 tags: calamari

#
# restart salt-minion
#
- name: start salt-minion
 service: name=salt-minion state=restarted enabled=yes
 changed_when: False
 tags: calamari
```

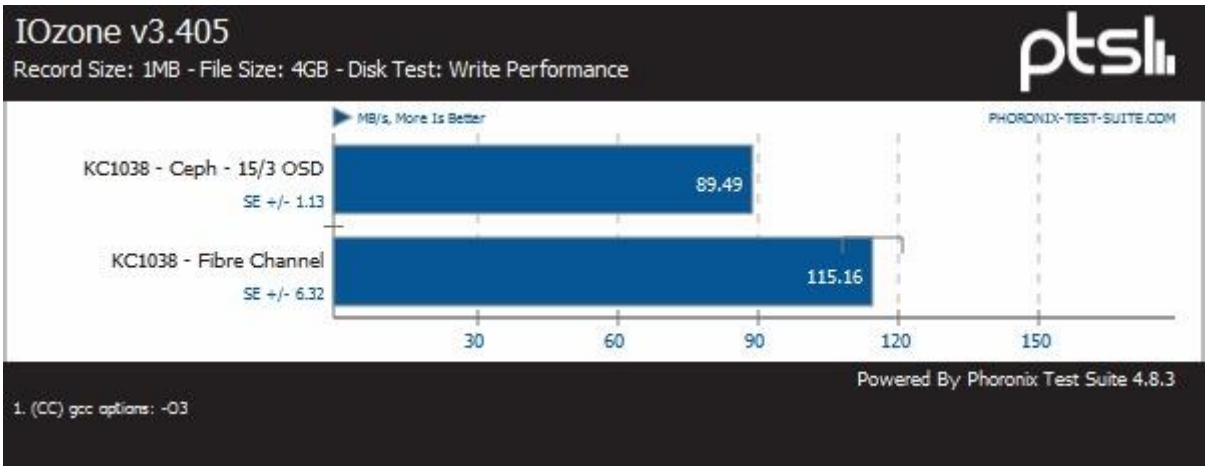
H. Bijlage Benchmark initieel vergelijk

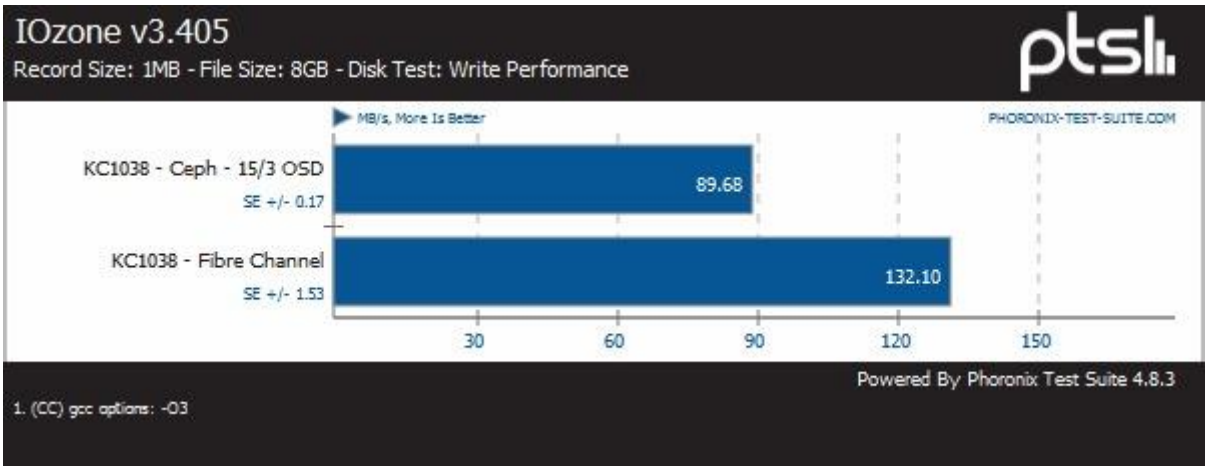


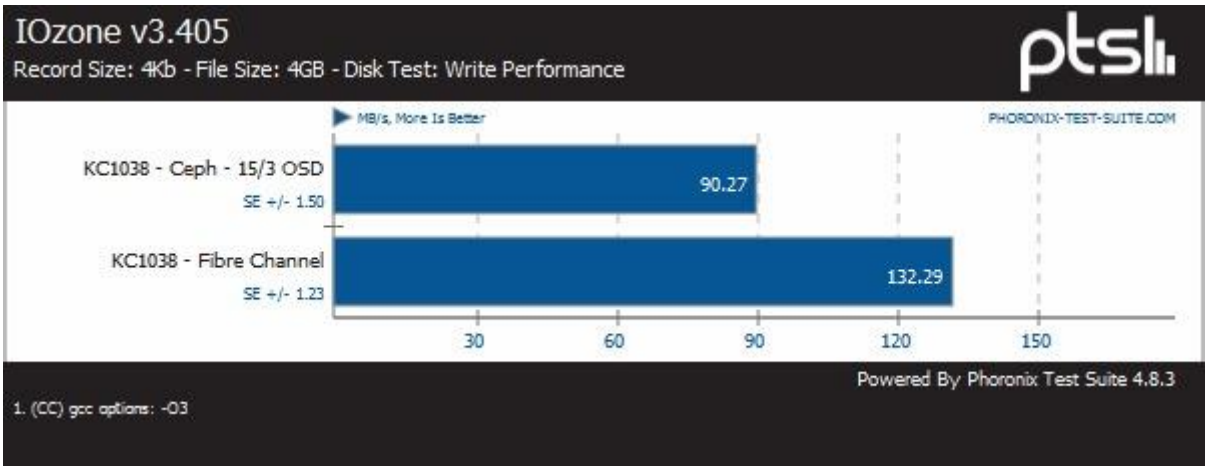


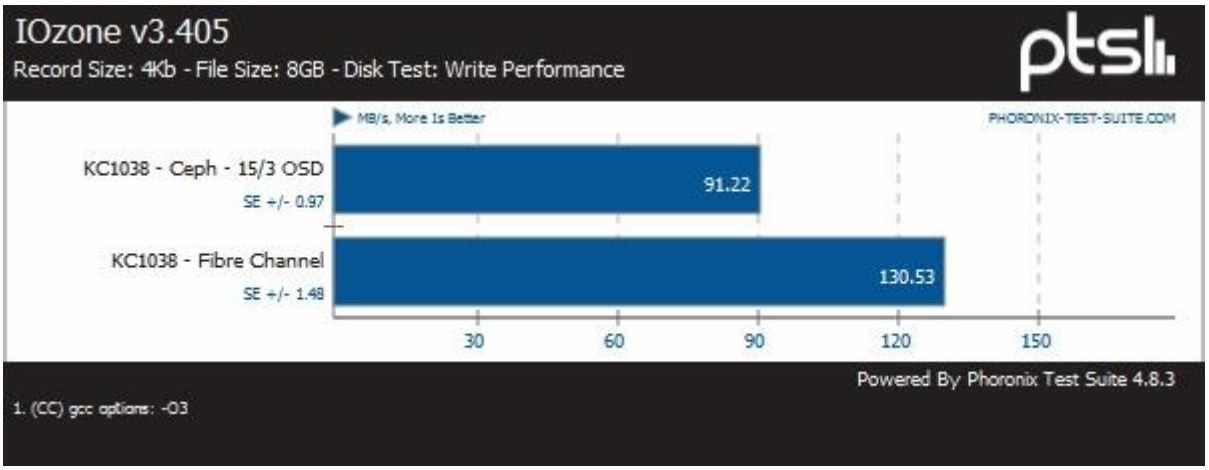


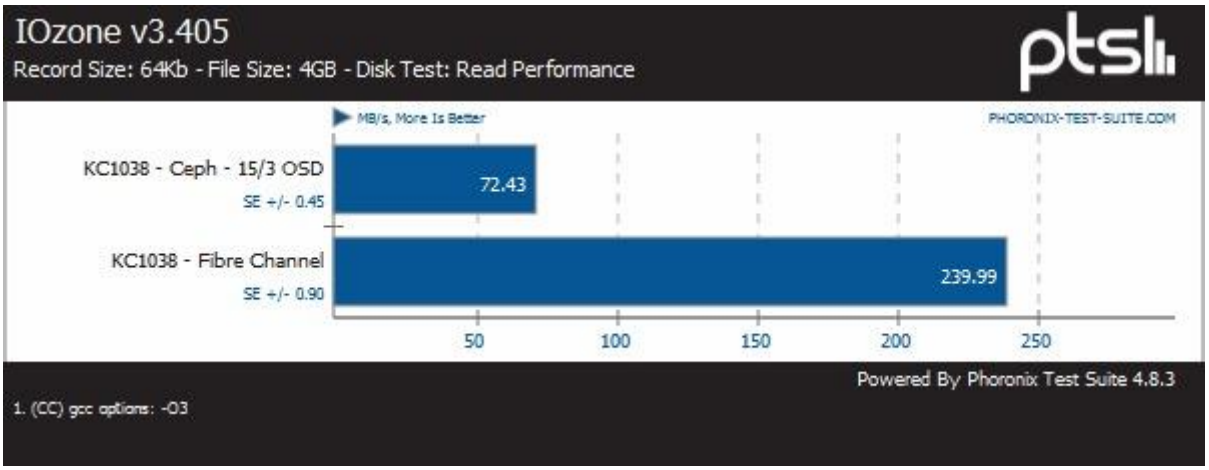

IOzone v3.405
Record Size: 4Kb - File Size: 8GB - Disk Test: Read Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.29
72.61
KC1038 - Fibre Channel
SE +/- 1.74
237.09
50
100
150
200
250
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

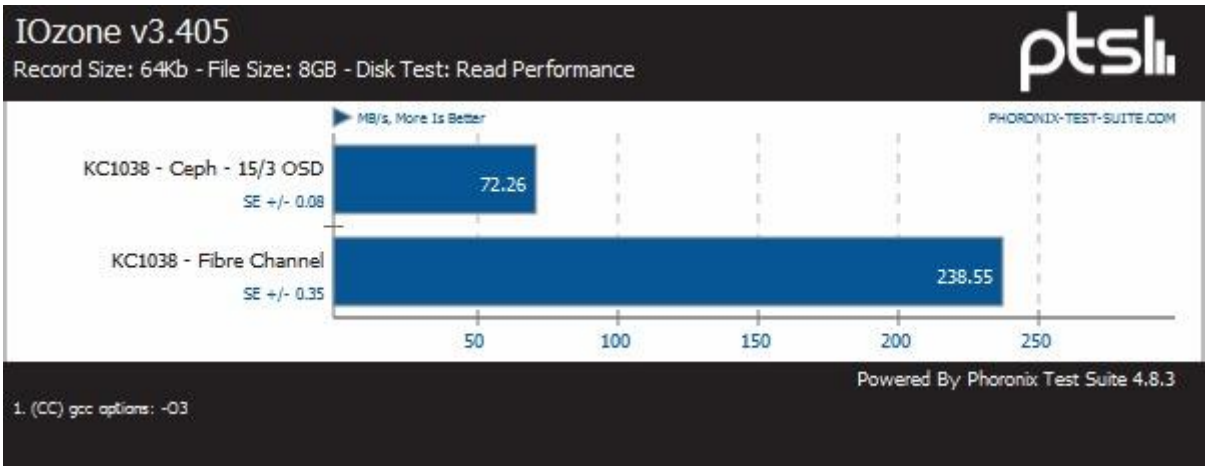

IOzone v3.405
Record Size: 1MB - File Size: 4GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 1.13
89.49
KC1038 - Fibre Channel
SE +/- 6.32
115.16
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

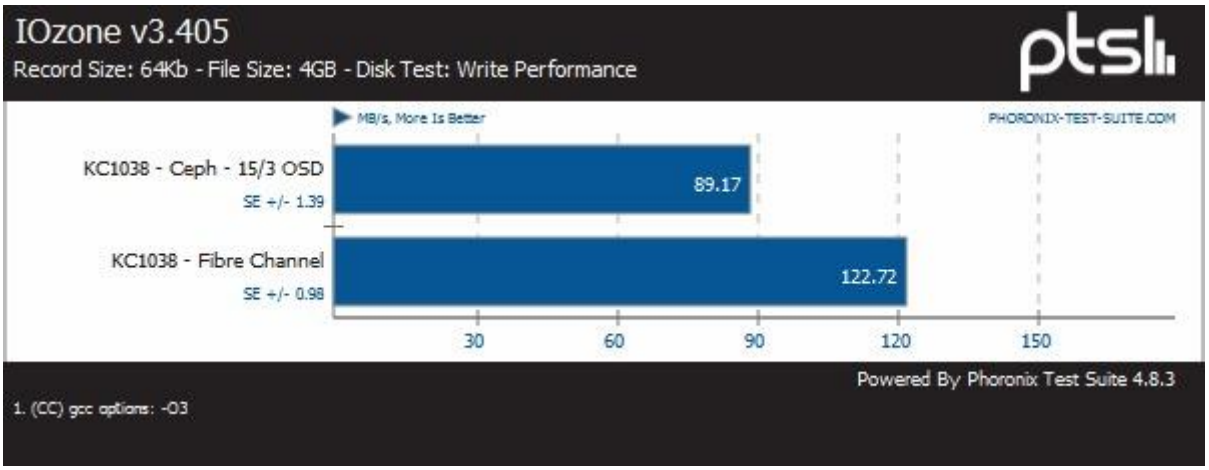

IOzone v3.405
Record Size: 1MB - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.17
89.68
KC1038 - Fibre Channel
SE +/- 1.53
132.10
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

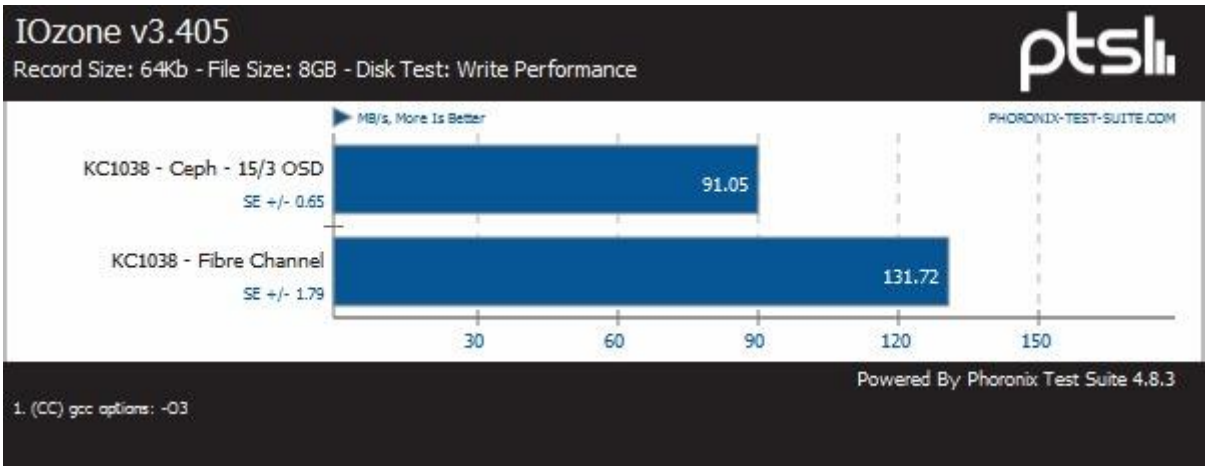

IOzone v3.405
Record Size: 4Kb - File Size: 4GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 1.50
90.27
KC1038 - Fibre Channel
SE +/- 1.23
132.29
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3


IOzone v3.405
Record Size: 4Kb - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.97
91.22
KC1038 - Fibre Channel
SE +/- 1.48
130.53
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

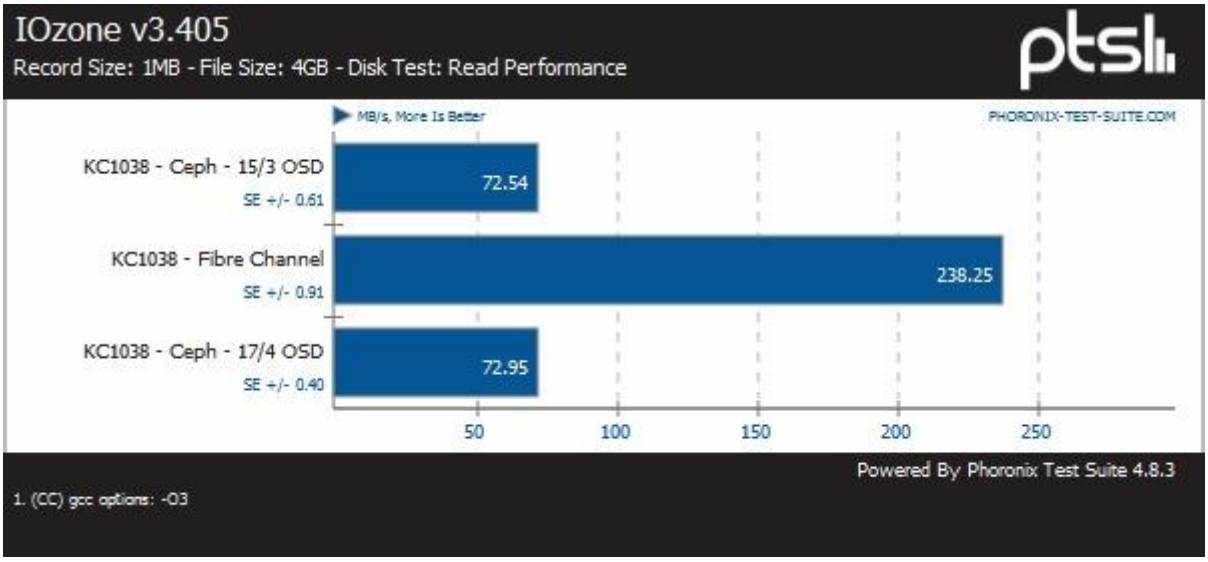

IOzone v3.405
Record Size: 64Kb - File Size: 4GB - Disk Test: Read Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.45
72.43
KC1038 - Fibre Channel
SE +/- 0.90
239.99
50
100
150
200
250
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

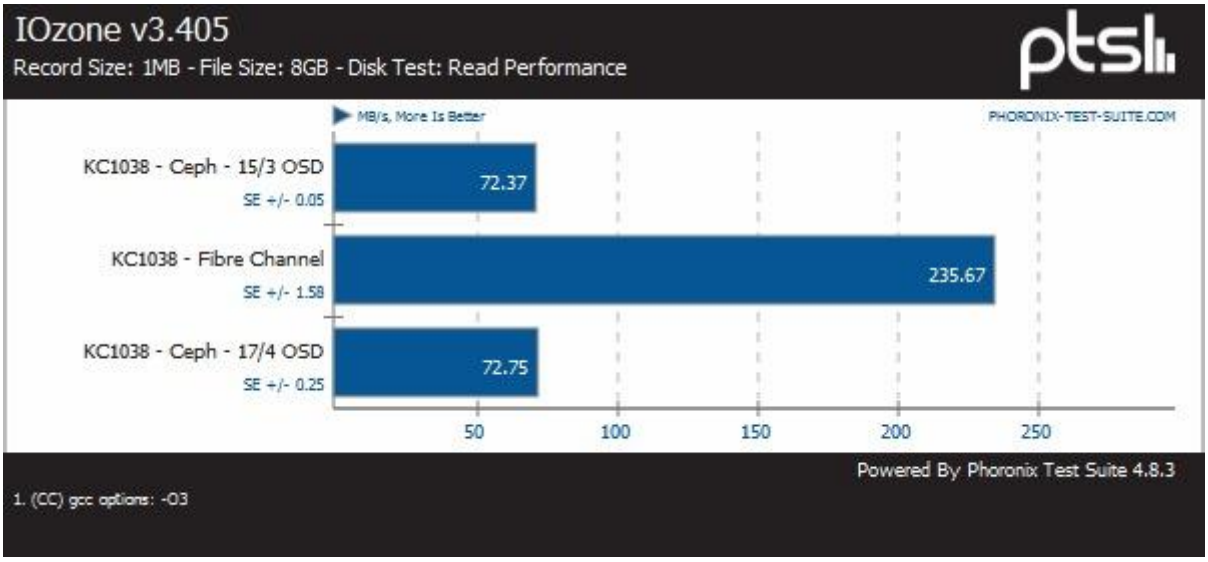

IOzone v3.405
Record Size: 64Kb - File Size: 8GB - Disk Test: Read Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.08
72.26
KC1038 - Fibre Channel
SE +/- 0.35
238.55
50
100
150
200
250
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

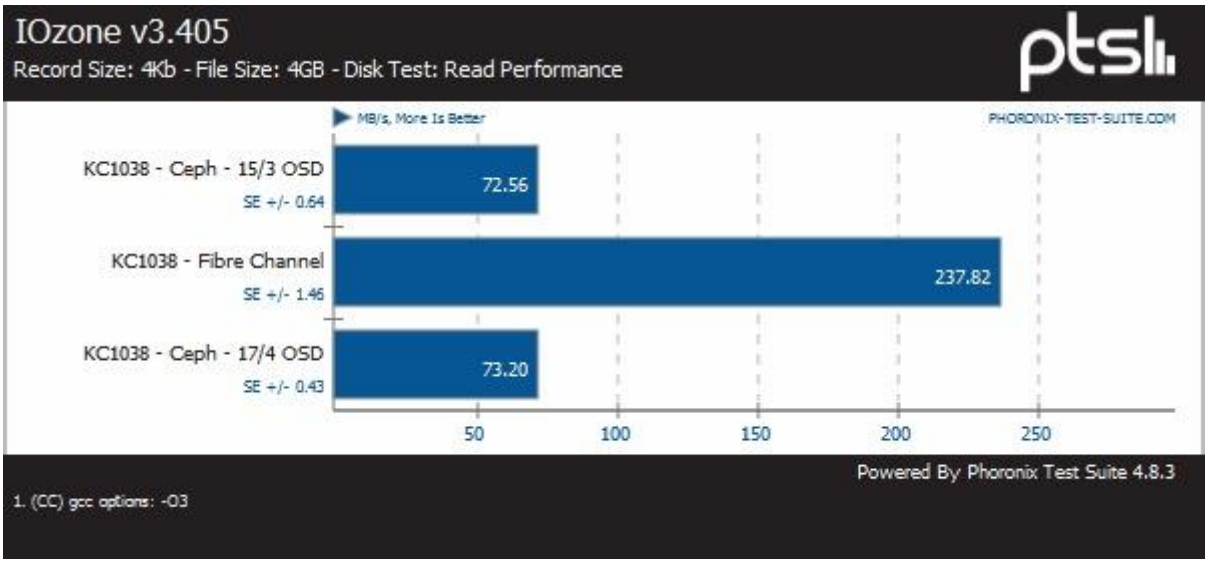

IOzone v3.405
Record Size: 64Kb - File Size: 4GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 1.39
89.17
KC1038 - Fibre Channel
SE +/- 0.98
122.72
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

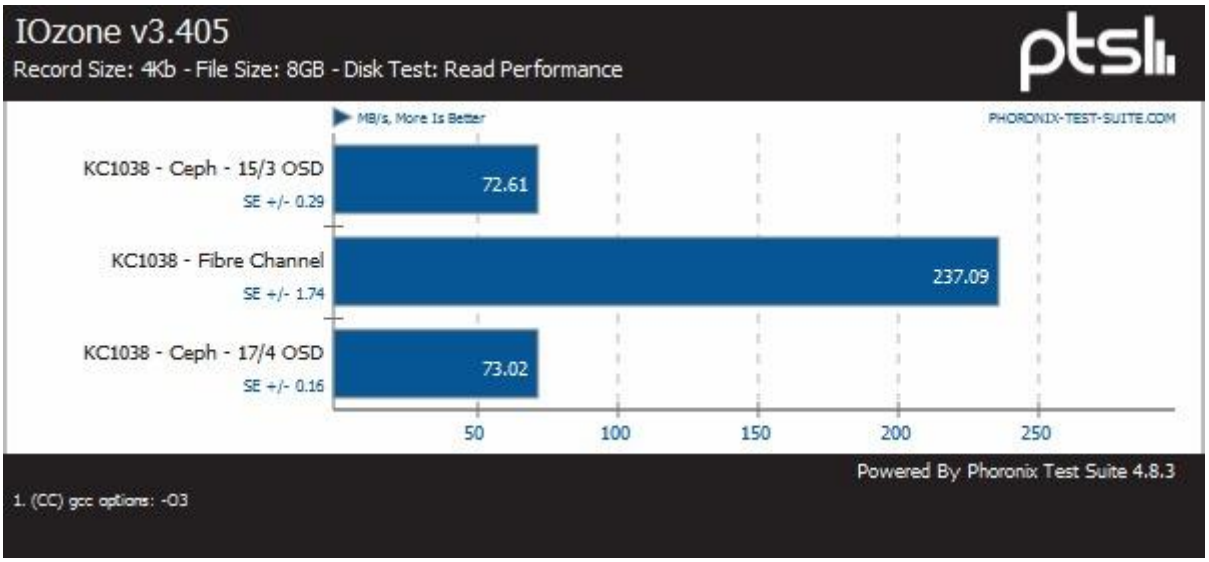

IOzone v3.405
Record Size: 64Kb - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.65
91.05
KC1038 - Fibre Channel
SE +/- 1.79
131.72
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

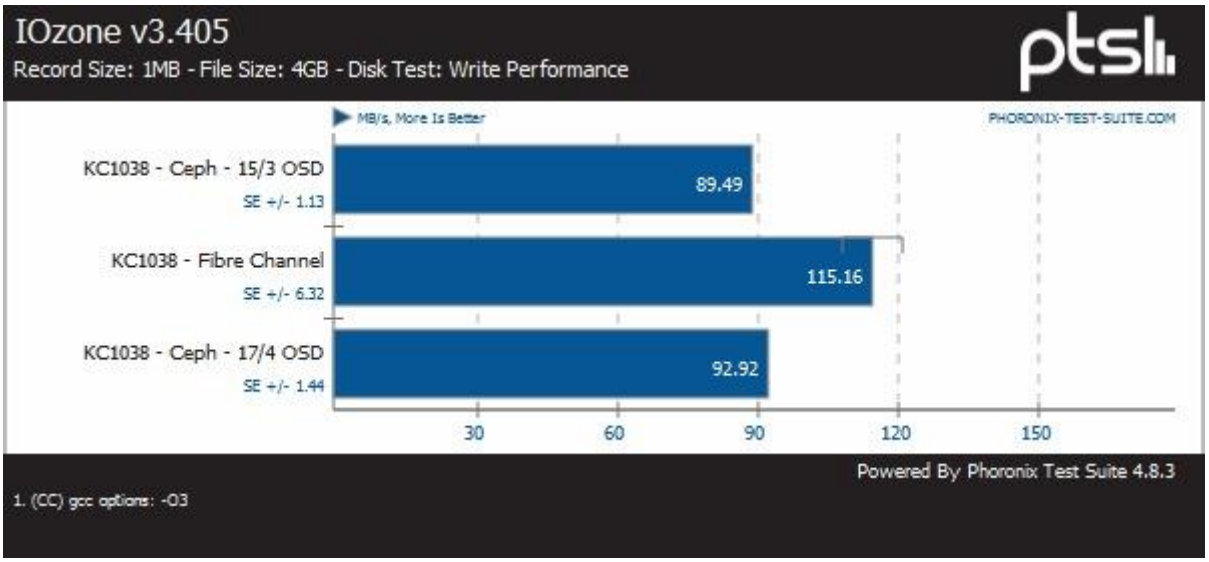
I. Bijlage Benchmark schaalbaarheid

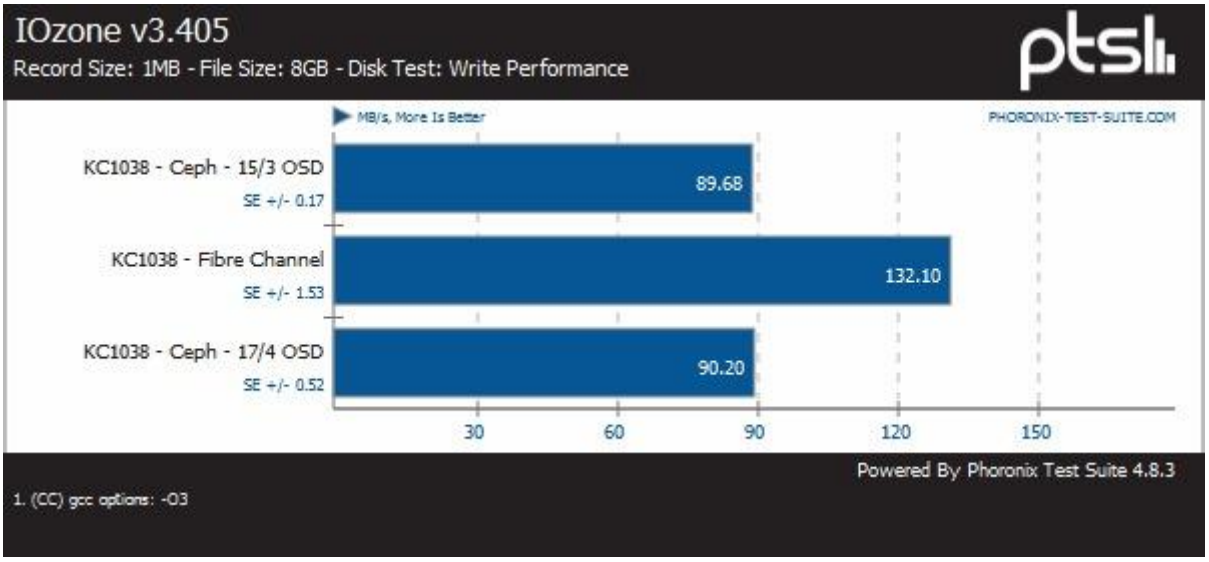


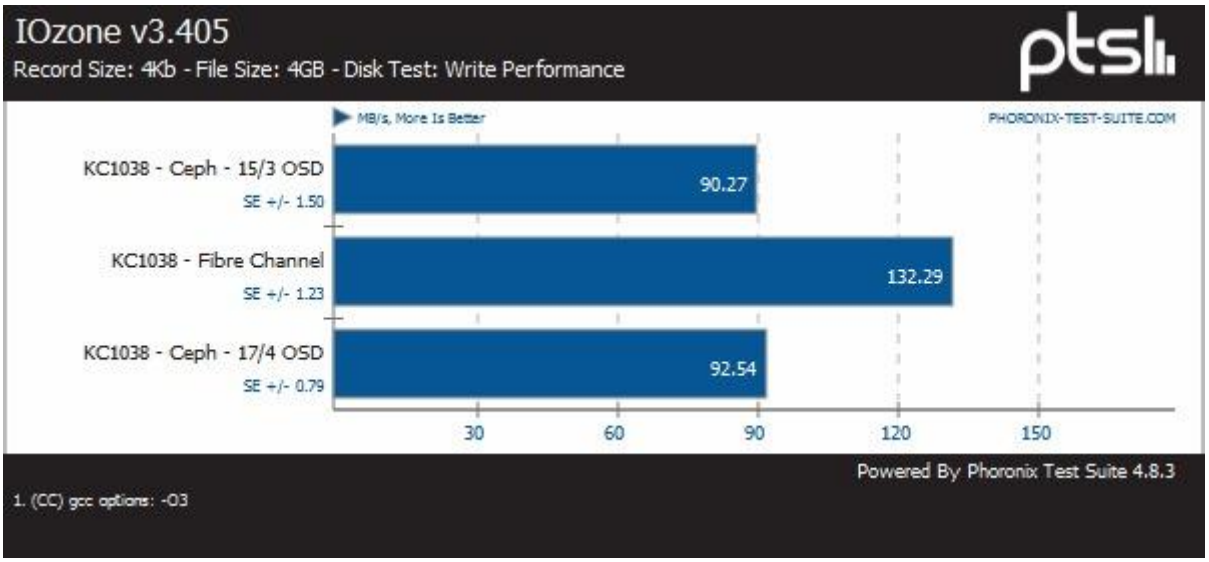


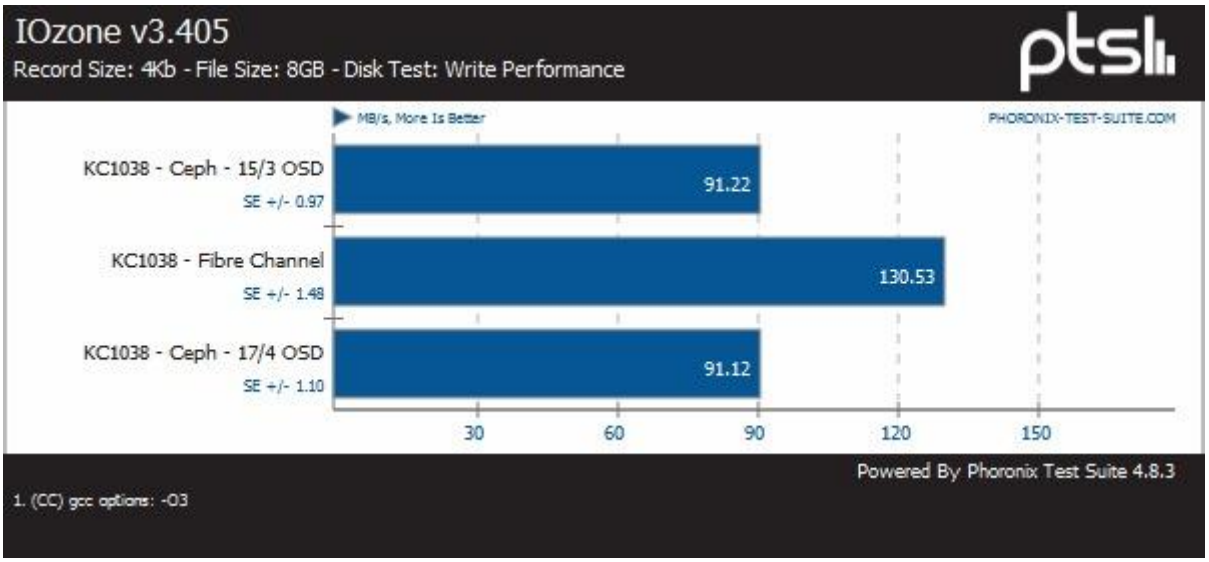


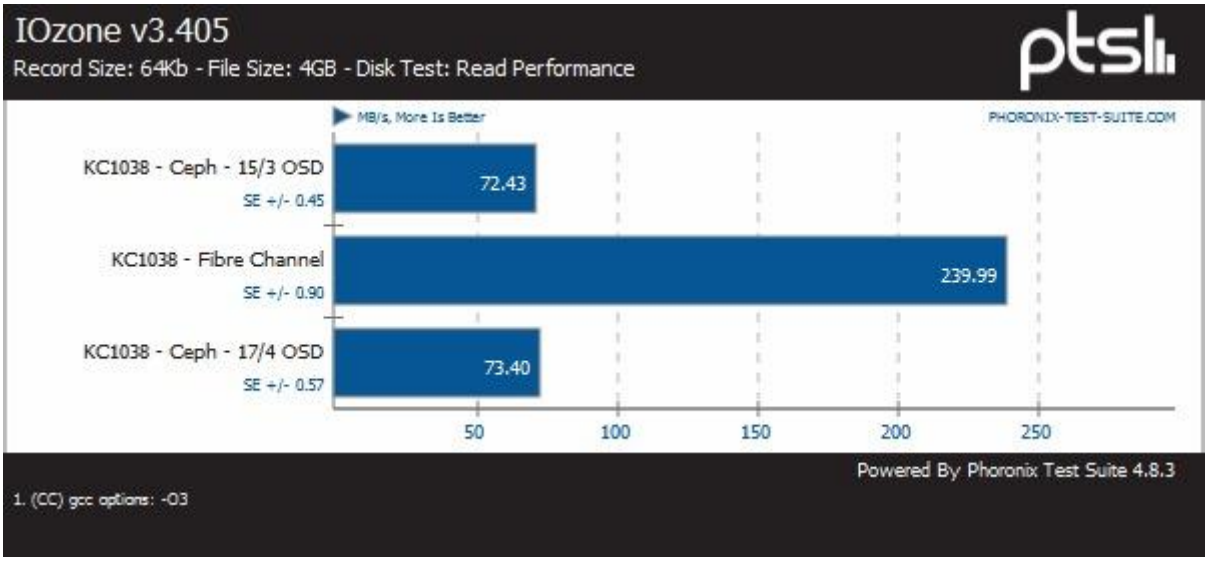

IOzone v3.405
Record Size: 4Kb - File Size: 8GB - Disk Test: Read Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.29
72.61
KC1038 - Fibre Channel
SE +/- 1.74
237.09
KC1038 - Ceph - 17/4 OSD
SE +/- 0.16
73.02
50
100
150
200
250
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

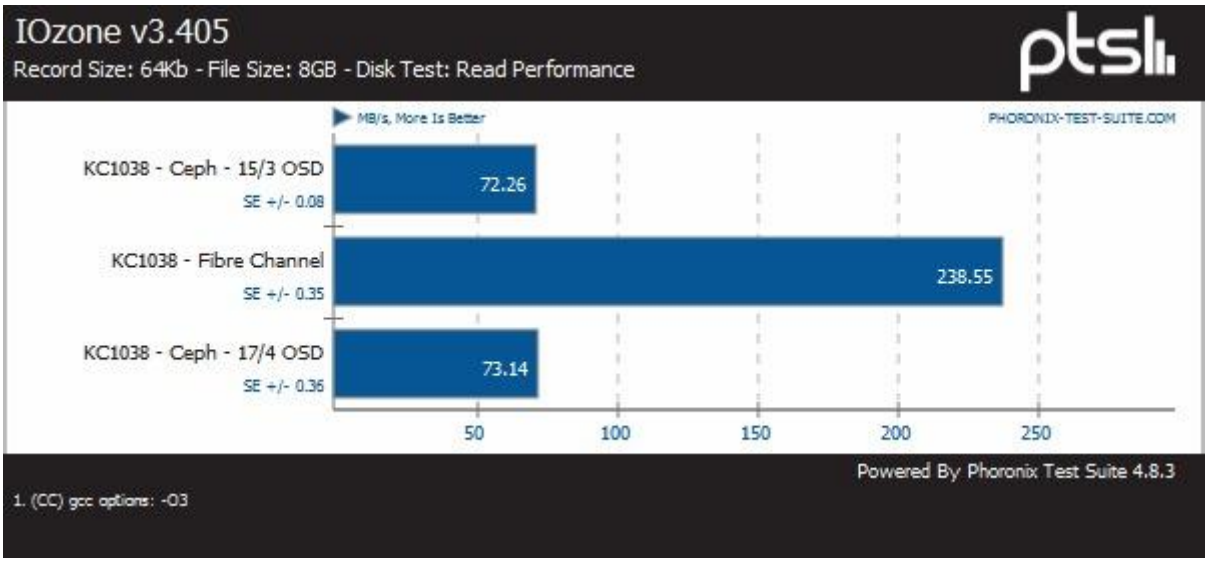

IOzone v3.405
Record Size: 1MB - File Size: 4GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 1.13
89.49
KC1038 - Fibre Channel
SE +/- 6.32
115.16
KC1038 - Ceph - 17/4 OSD
SE +/- 1.44
92.92
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

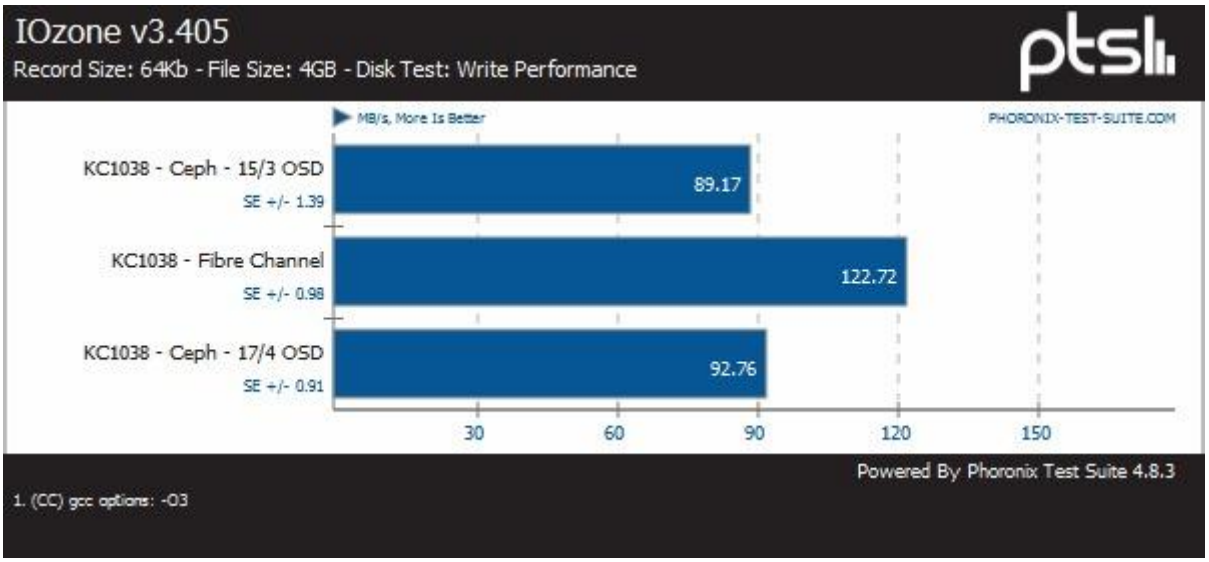

IOzone v3.405
Record Size: 1MB - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.17
89.68
KC1038 - Fibre Channel
SE +/- 1.53
132.10
KC1038 - Ceph - 17/4 OSD
SE +/- 0.52
90.20
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

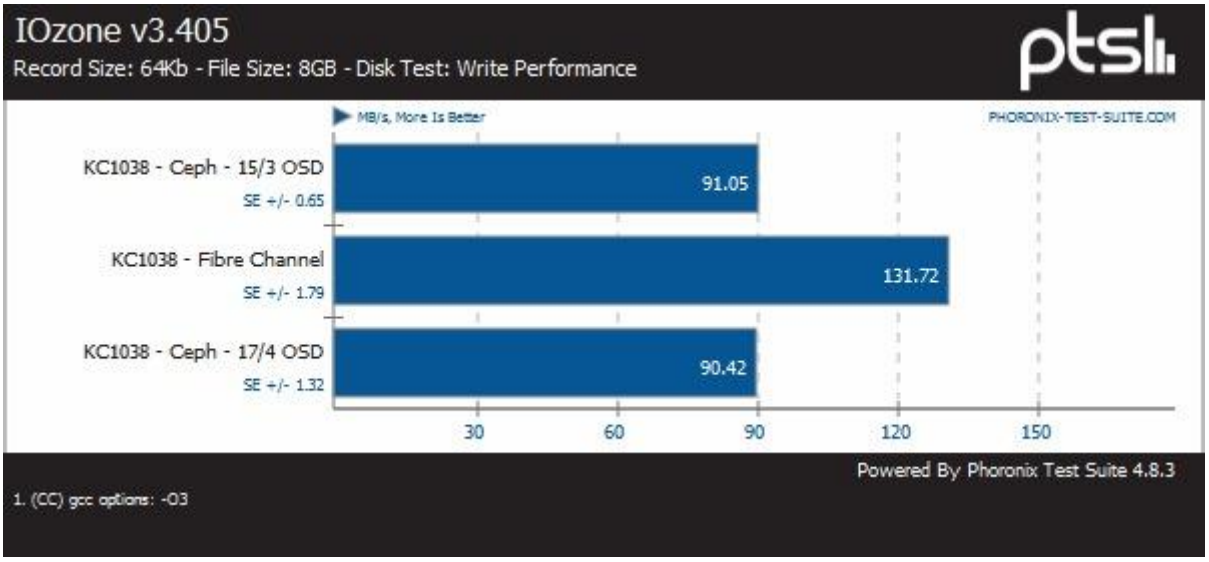

IOzone v3.405
Record Size: 4Kb - File Size: 4GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 1.50
90.27
KC1038 - Fibre Channel
SE +/- 1.23
132.29
KC1038 - Ceph - 17/4 OSD
SE +/- 0.79
92.54
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3


IOzone v3.405
Record Size: 4Kb - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.97
91.22
KC1038 - Fibre Channel
SE +/- 1.48
130.53
KC1038 - Ceph - 17/4 OSD
SE +/- 1.10
91.12
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

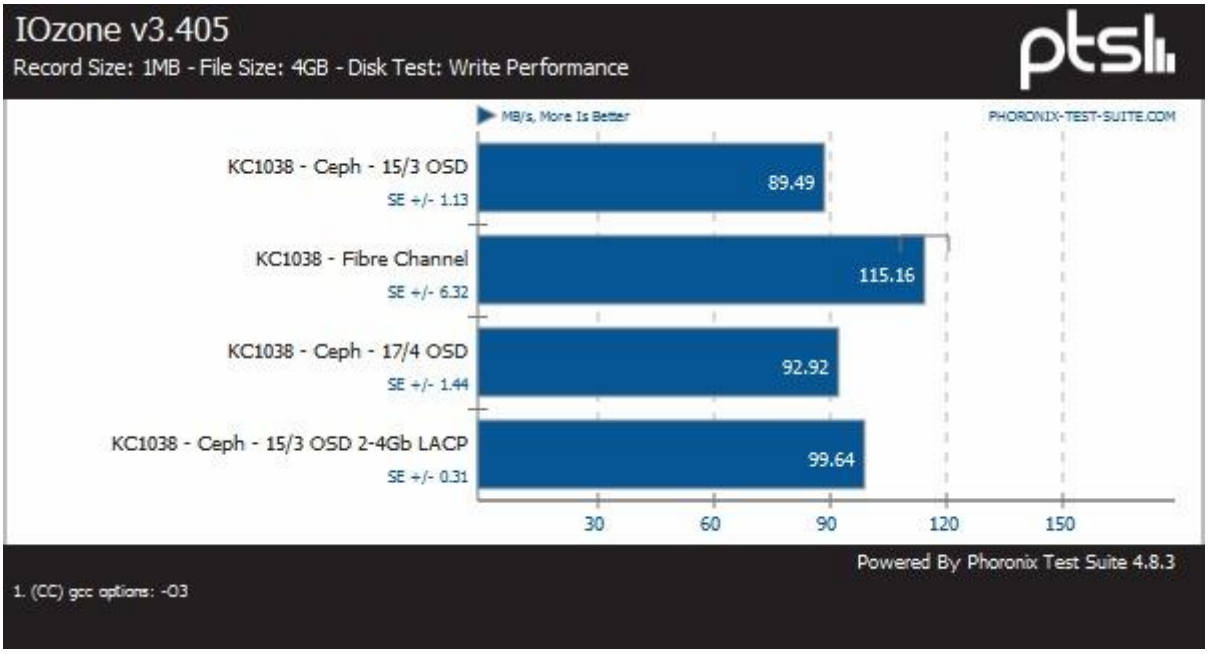

IOzone v3.405
Record Size: 64Kb - File Size: 4GB - Disk Test: Read Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.45
72.43
KC1038 - Fibre Channel
SE +/- 0.90
239.99
KC1038 - Ceph - 17/4 OSD
SE +/- 0.57
73.40
50
100
150
200
250
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

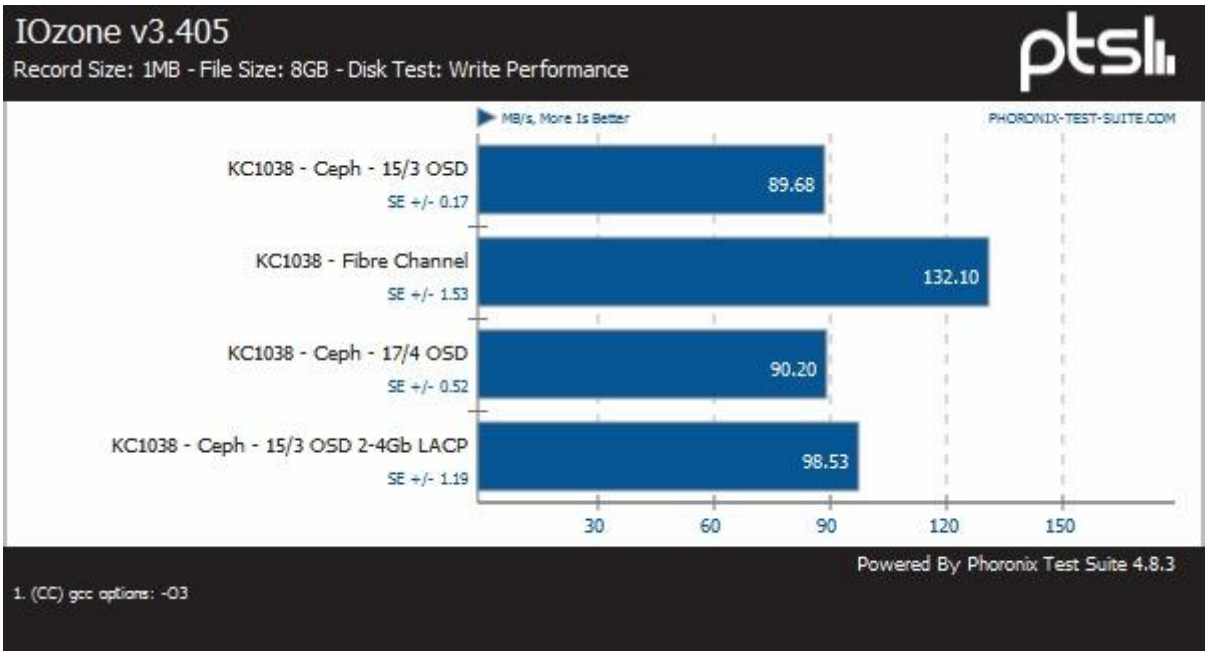

IOzone v3.405
Record Size: 64Kb - File Size: 8GB - Disk Test: Read Performance
ptsl
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.08
72.26
KC1038 - Fibre Channel
SE +/- 0.35
238.55
KC1038 - Ceph - 17/4 OSD
SE +/- 0.36
73.14
50
100
150
200
250
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

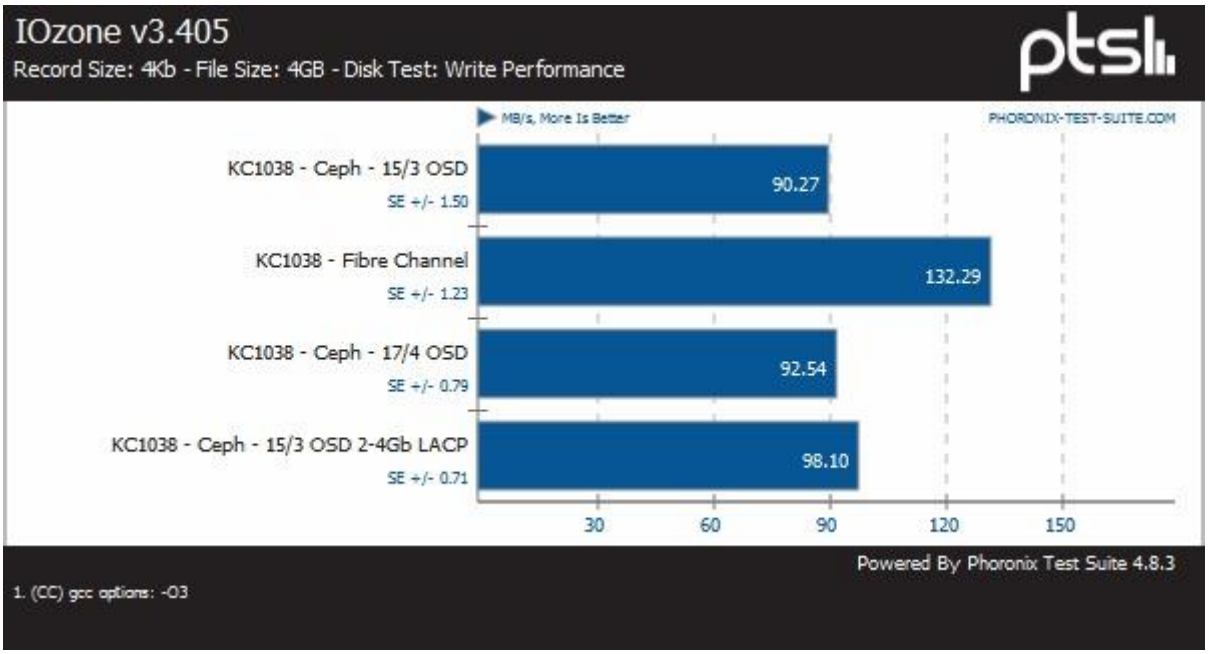

IOzone v3.405
Record Size: 64Kb - File Size: 4GB - Disk Test: Write Performance
ptsl
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 1.39
89.17
KC1038 - Fibre Channel
SE +/- 0.98
122.72
KC1038 - Ceph - 17/4 OSD
SE +/- 0.91
92.76
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

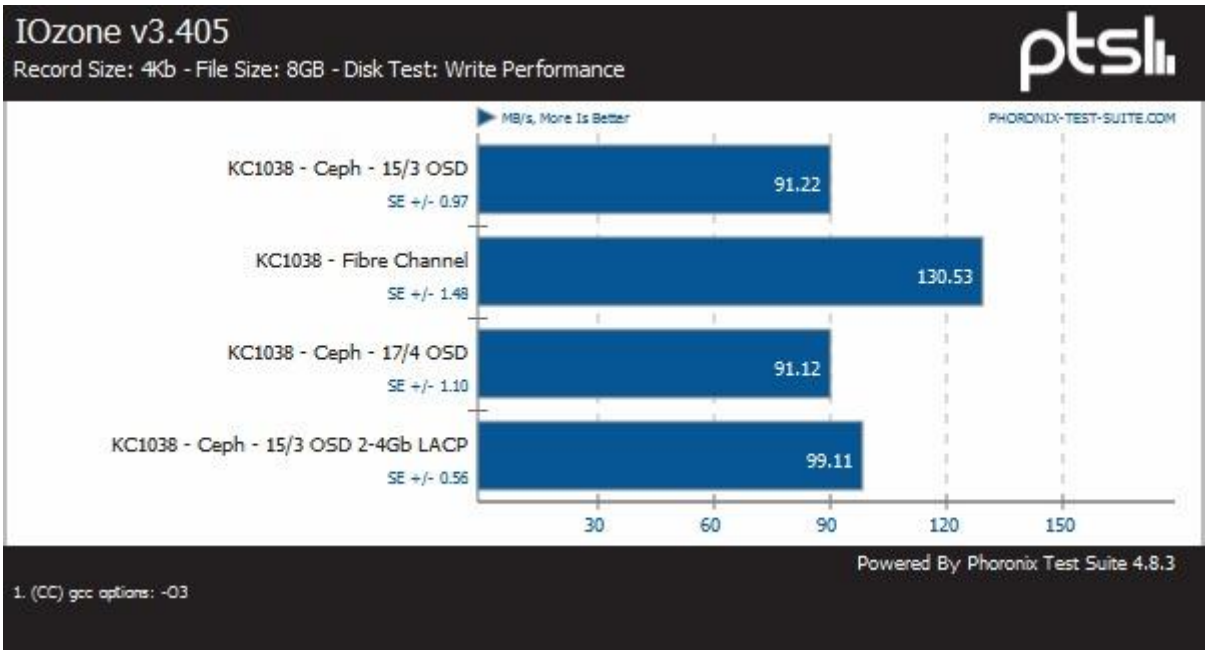

IOzone v3.405
Record Size: 64Kb - File Size: 8GB - Disk Test: Write Performance
ptsl
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.65
91.05
KC1038 - Fibre Channel
SE +/- 1.79
131.72
KC1038 - Ceph - 17/4 OSD
SE +/- 1.32
90.42
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

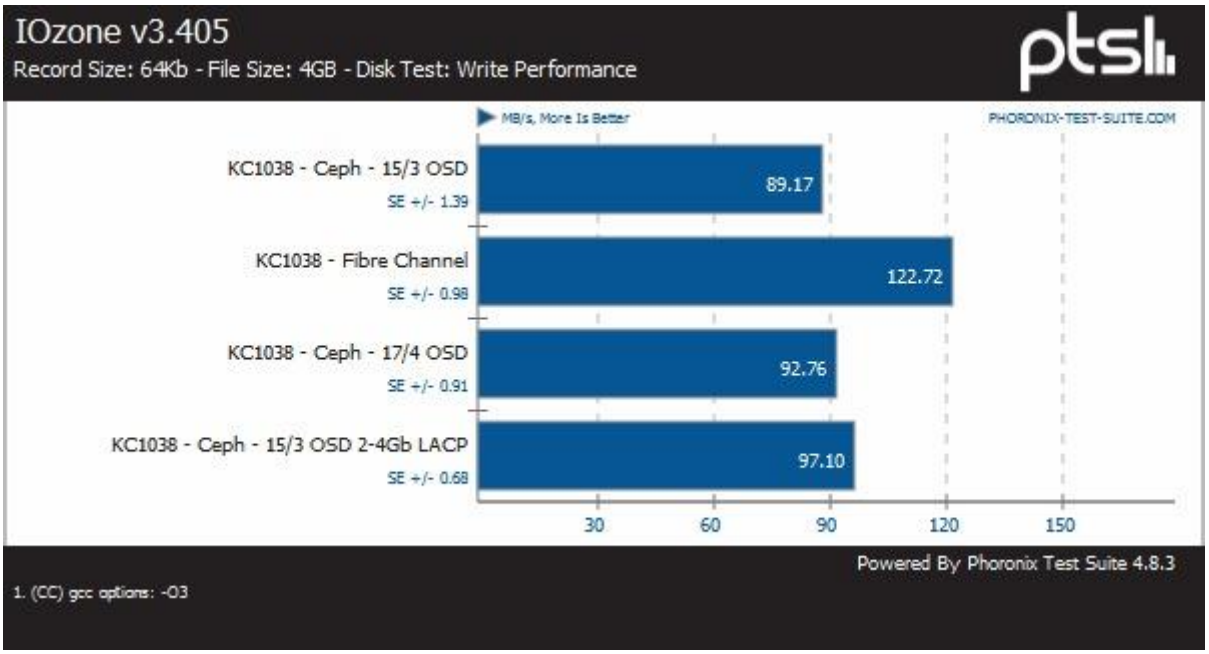
J. Bijlage Benchmark network upgrade

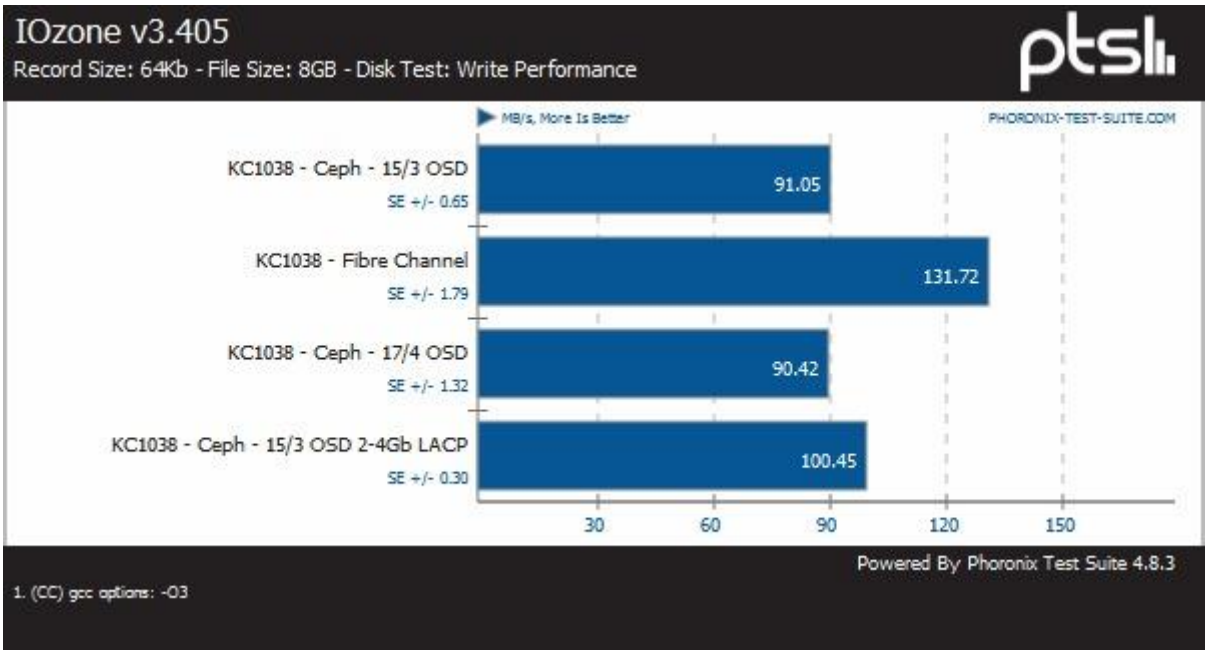




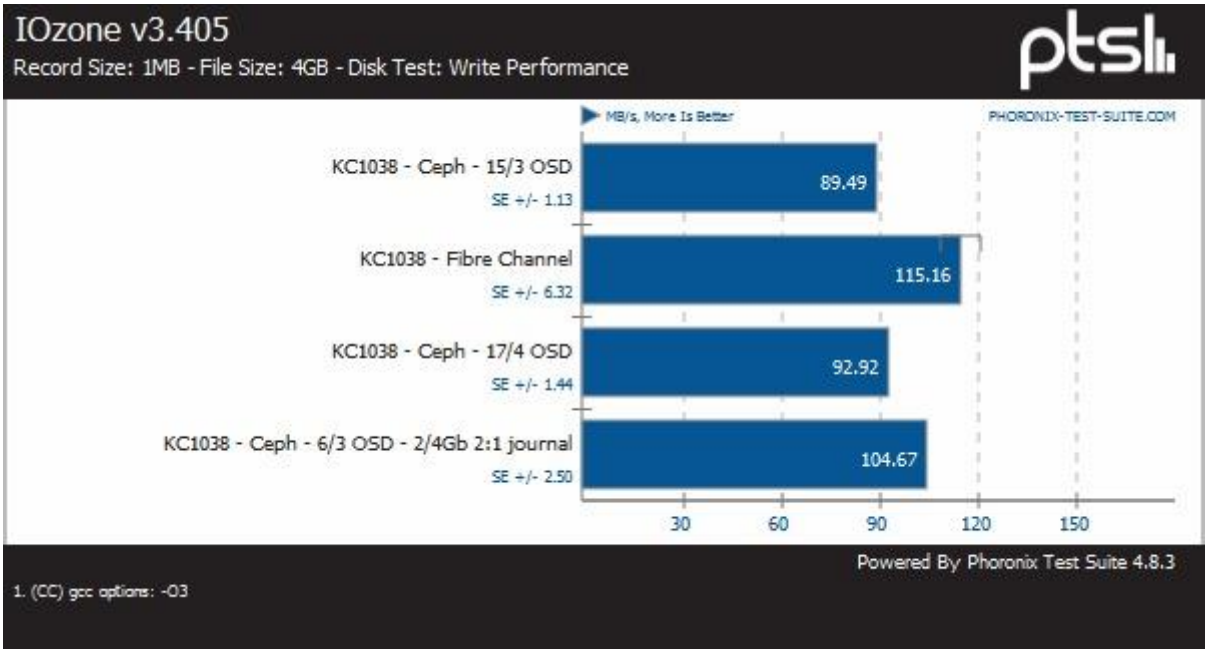

IOzone v3.405
Record Size: 4Kb - File Size: 4GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 1.50
90.27
KC1038 - Fibre Channel
SE +/- 1.23
132.29
KC1038 - Ceph - 17/4 OSD
SE +/- 0.79
92.54
KC1038 - Ceph - 15/3 OSD 2-4Gb LACP
SE +/- 0.71
98.10
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

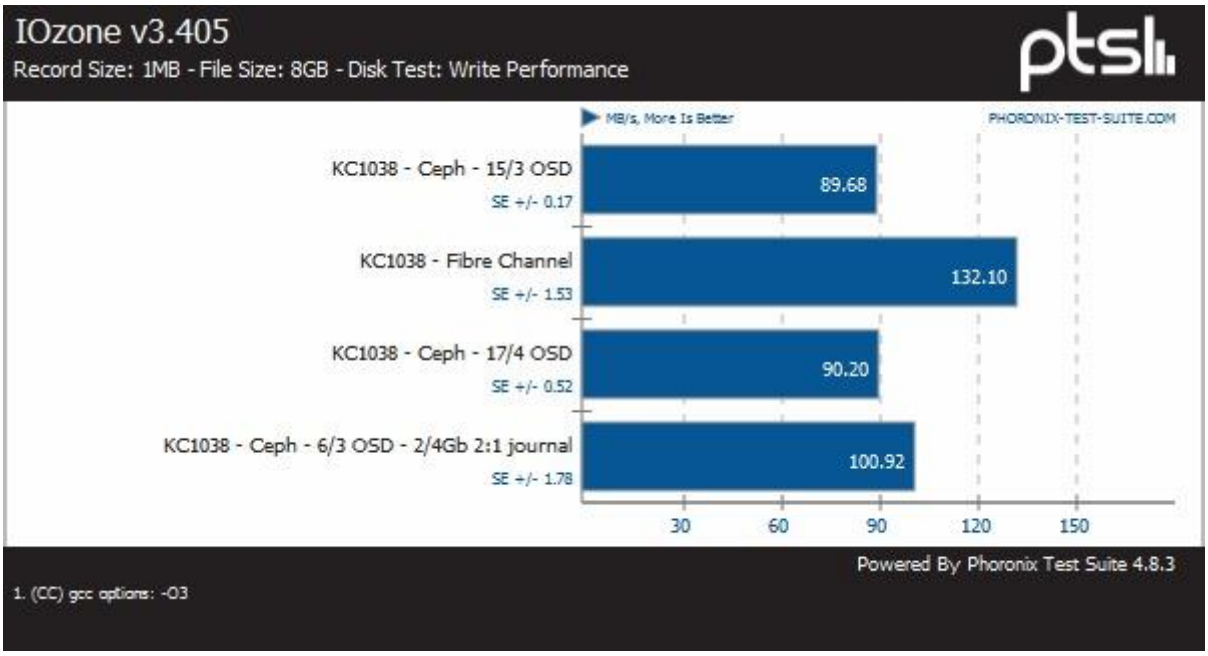

IOzone v3.405
Record Size: 4Kb - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.97
91.22
KC1038 - Fibre Channel
SE +/- 1.48
130.53
KC1038 - Ceph - 17/4 OSD
SE +/- 1.10
91.12
KC1038 - Ceph - 15/3 OSD 2-4Gb LACP
SE +/- 0.56
99.11
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

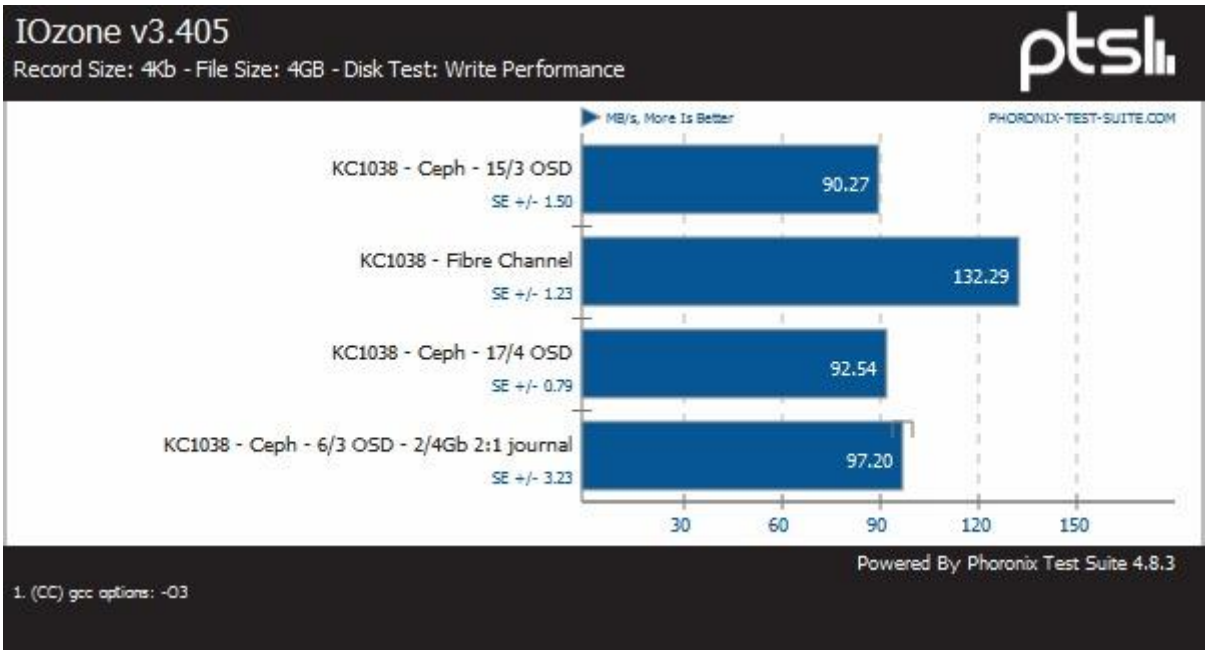

IOzone v3.405
Record Size: 64Kb - File Size: 4GB - Disk Test: Write Performance
ptsl
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 1.39
89.17
KC1038 - Fibre Channel
SE +/- 0.98
122.72
KC1038 - Ceph - 17/4 OSD
SE +/- 0.91
92.76
KC1038 - Ceph - 15/3 OSD 2-4Gb LACP
SE +/- 0.68
97.10
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

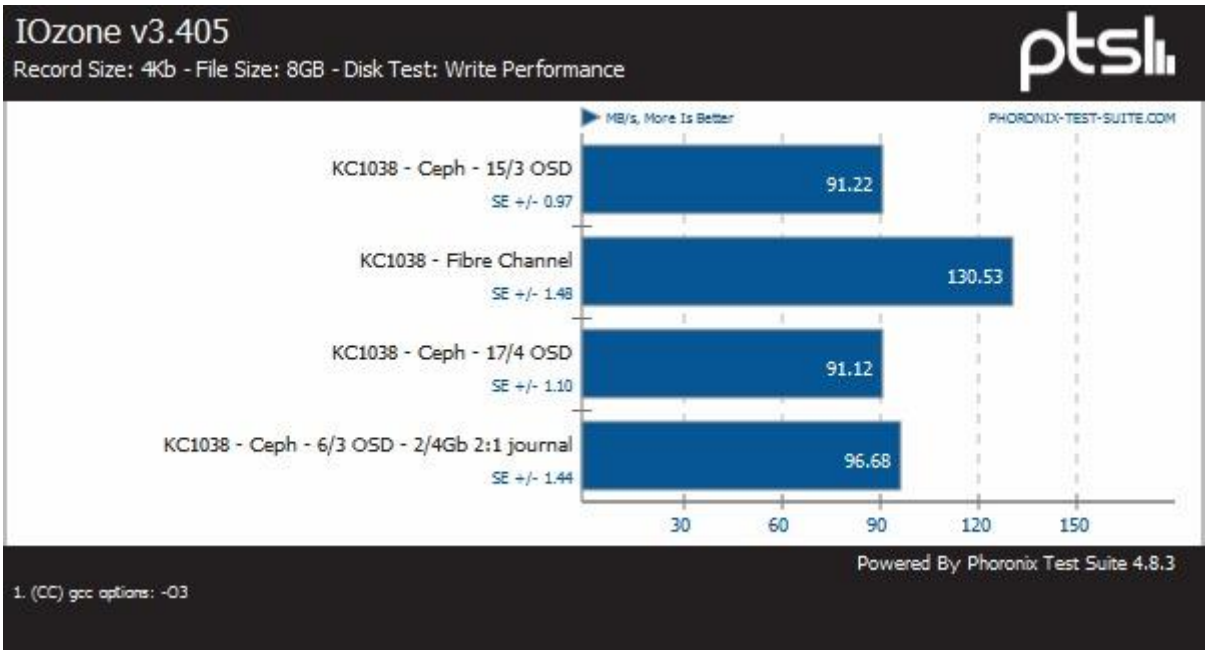

IOzone v3.405
Record Size: 64Kb - File Size: 8GB - Disk Test: Write Performance
ptsl
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.65
91.05
KC1038 - Fibre Channel
SE +/- 1.79
131.72
KC1038 - Ceph - 17/4 OSD
SE +/- 1.32
90.42
KC1038 - Ceph - 15/3 OSD 2-4Gb LACP
SE +/- 0.30
100.45
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

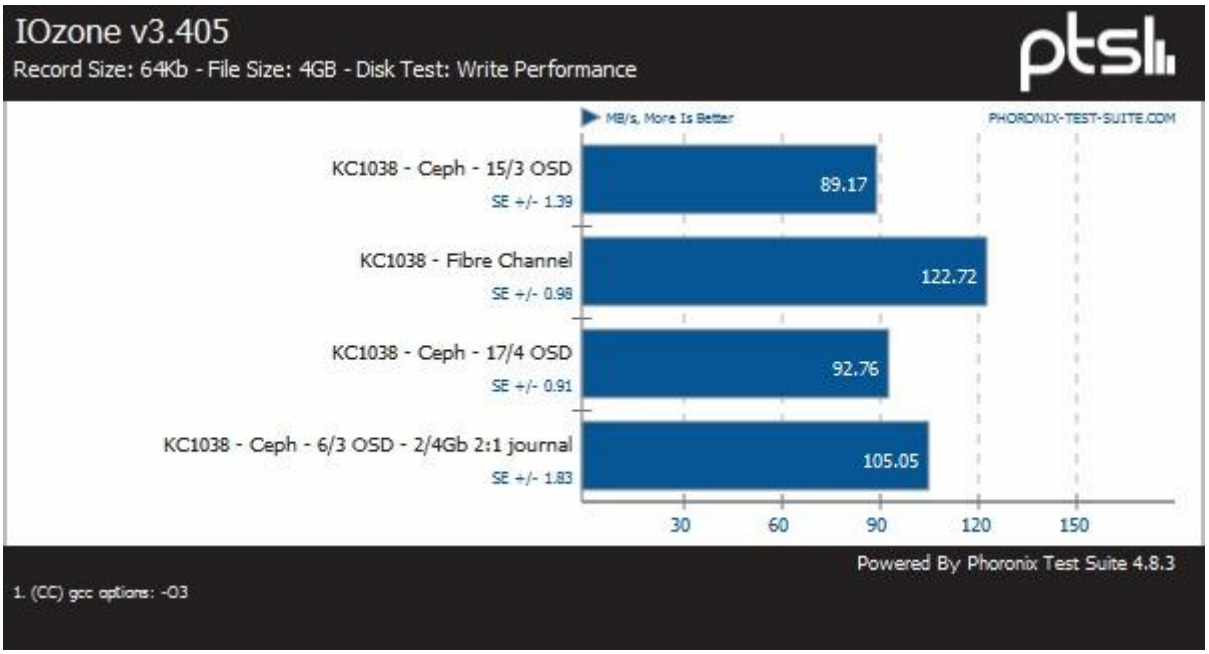
K. Bijlage Benchmark journal ratio

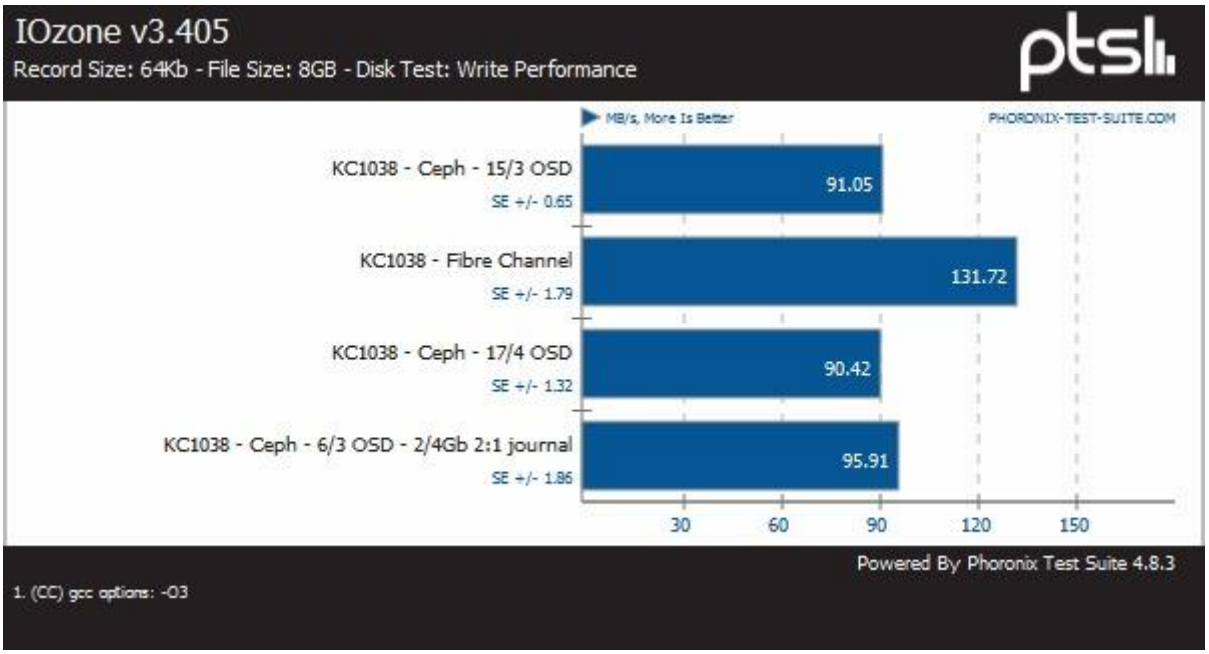




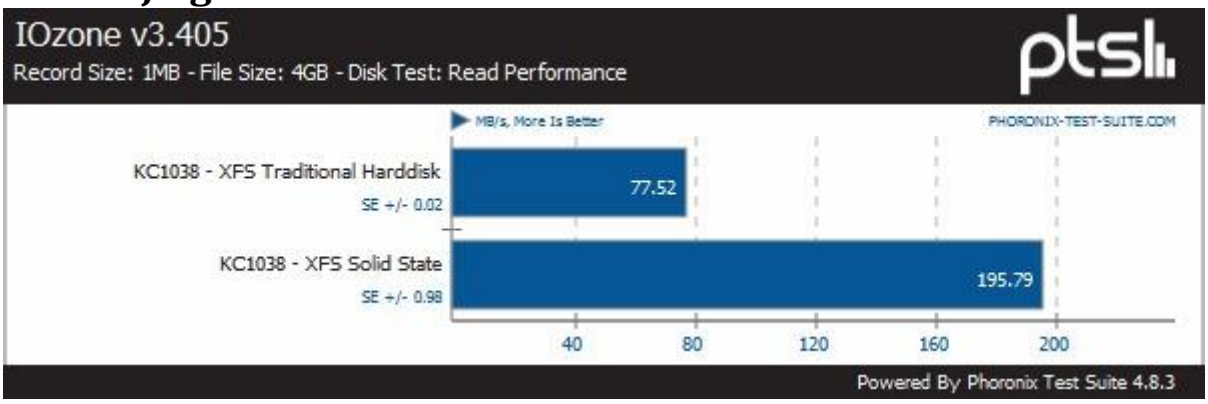

IOzone v3.405
Record Size: 4Kb - File Size: 4GB - Disk Test: Write Performance
ptsl
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 1.50
90.27
KC1038 - Fibre Channel
SE +/- 1.23
132.29
KC1038 - Ceph - 17/4 OSD
SE +/- 0.79
92.54
KC1038 - Ceph - 6/3 OSD - 2/4Gb 2:1 journal
SE +/- 3.23
97.20
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

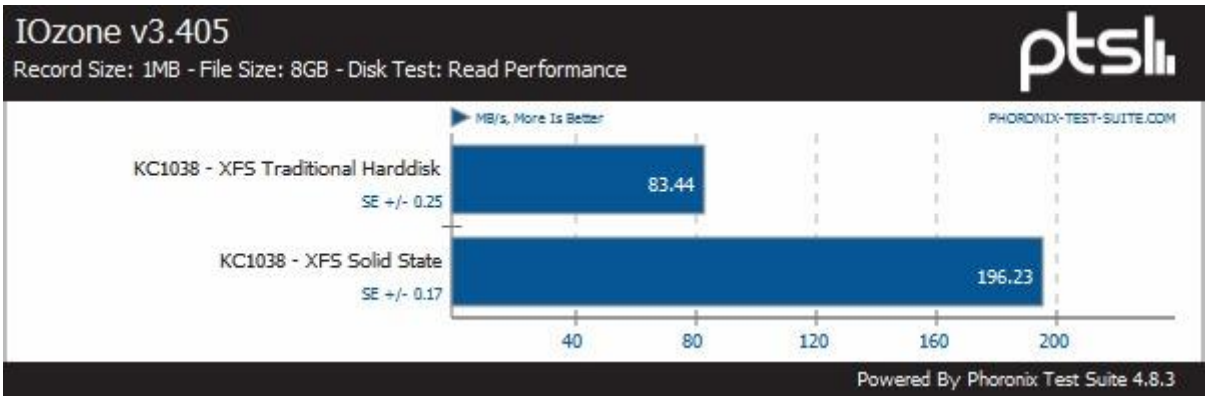

IOzone v3.405
Record Size: 4Kb - File Size: 8GB - Disk Test: Write Performance
ptsl
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.97
91.22
KC1038 - Fibre Channel
SE +/- 1.48
130.53
KC1038 - Ceph - 17/4 OSD
SE +/- 1.10
91.12
KC1038 - Ceph - 6/3 OSD - 2/4Gb 2:1 journal
SE +/- 1.44
96.68
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

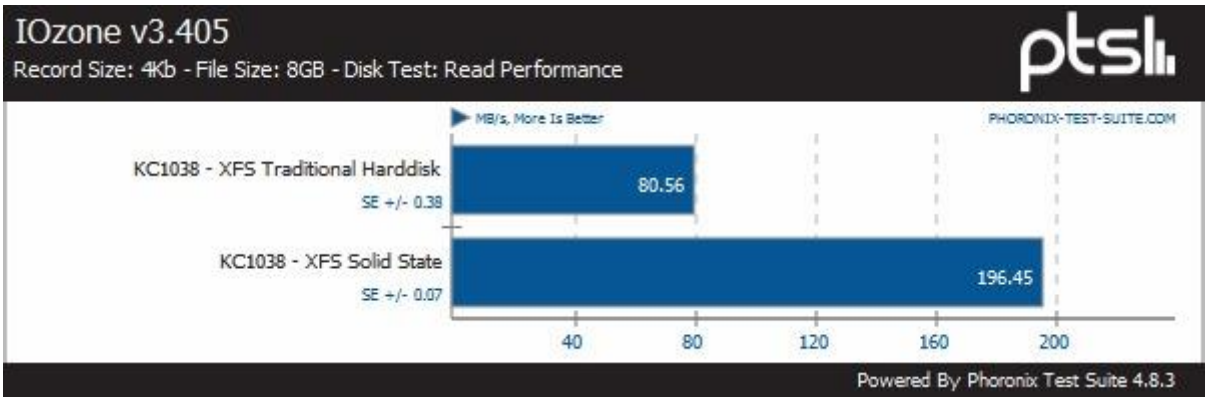

IOzone v3.405
Record Size: 64Kb - File Size: 4GB - Disk Test: Write Performance
ptsl
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 1.39
89.17
KC1038 - Fibre Channel
SE +/- 0.98
122.72
KC1038 - Ceph - 17/4 OSD
SE +/- 0.91
92.76
KC1038 - Ceph - 6/3 OSD - 2/4Gb 2:1 journal
SE +/- 1.83
105.05
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

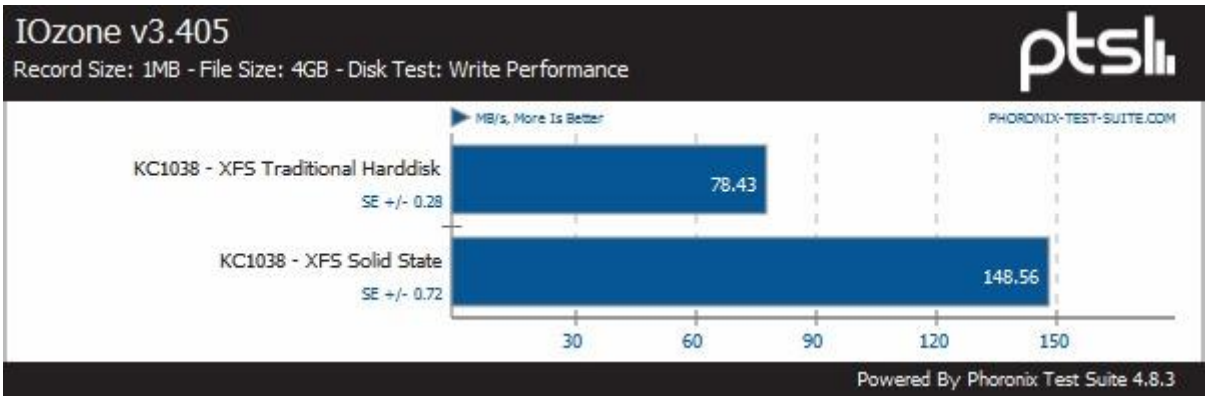

IOzone v3.405
Record Size: 64Kb - File Size: 8GB - Disk Test: Write Performance
ptsl
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - Ceph - 15/3 OSD
SE +/- 0.65
91.05
KC1038 - Fibre Channel
SE +/- 1.79
131.72
KC1038 - Ceph - 17/4 OSD
SE +/- 1.32
90.42
KC1038 - Ceph - 6/3 OSD - 2/4Gb 2:1 journal
SE +/- 1.86
95.91
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3
1. (CC) gcc options: -O3

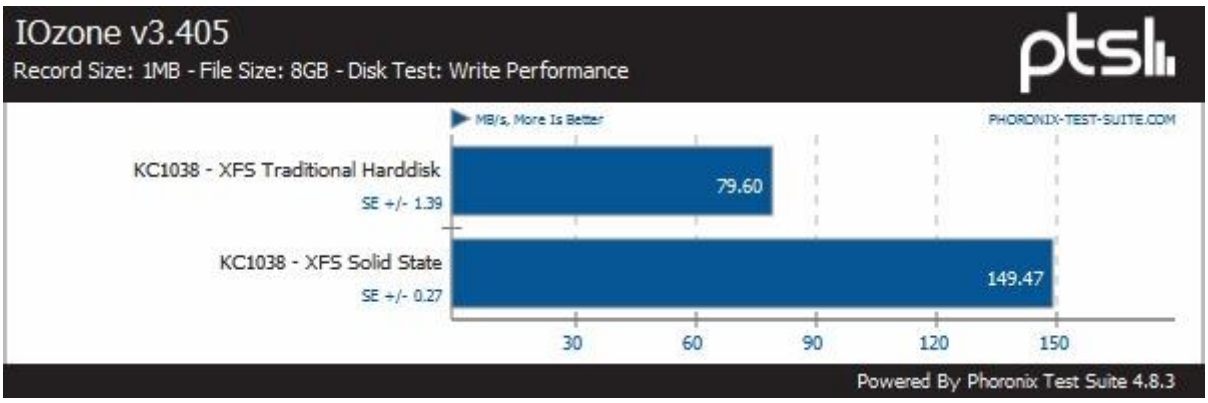
L. Bijlage Benchmark PERC H200 Controller

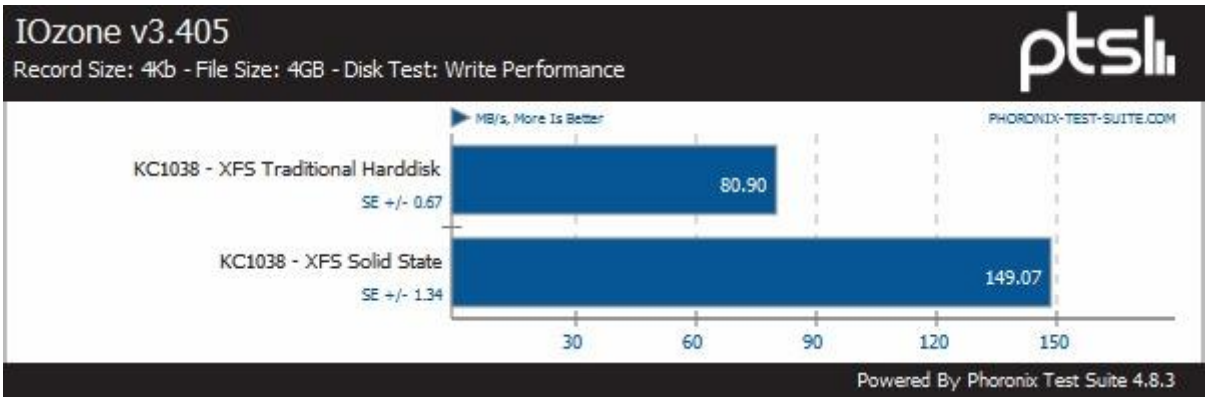


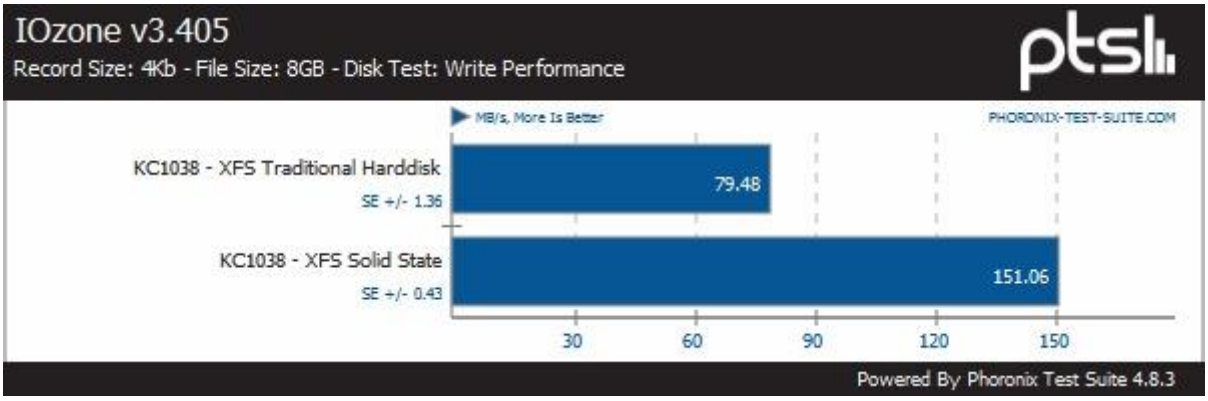


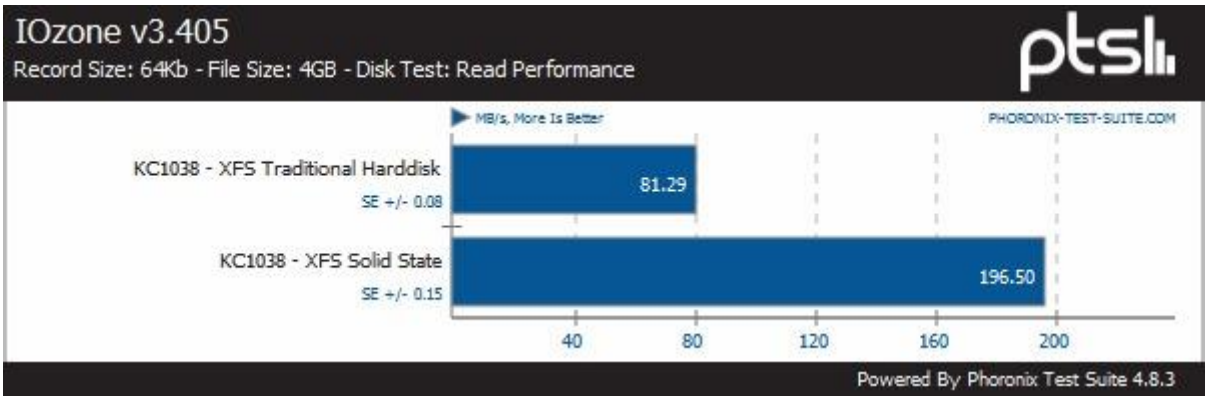


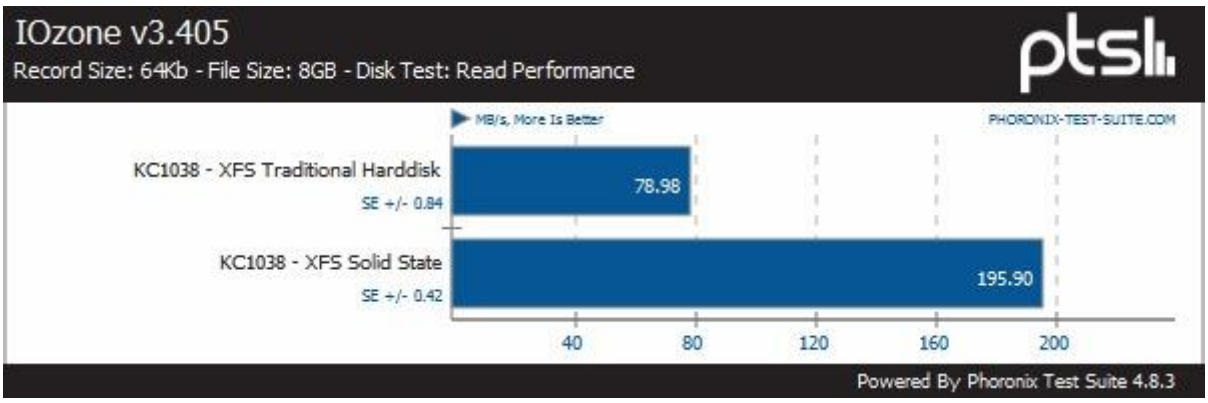

IOzone v3.405
Record Size: 1MB - File Size: 4GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - XFS Traditional Harddisk
SE +/- 0.28
78.43
KC1038 - XFS Solid State
SE +/- 0.72
148.56
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3

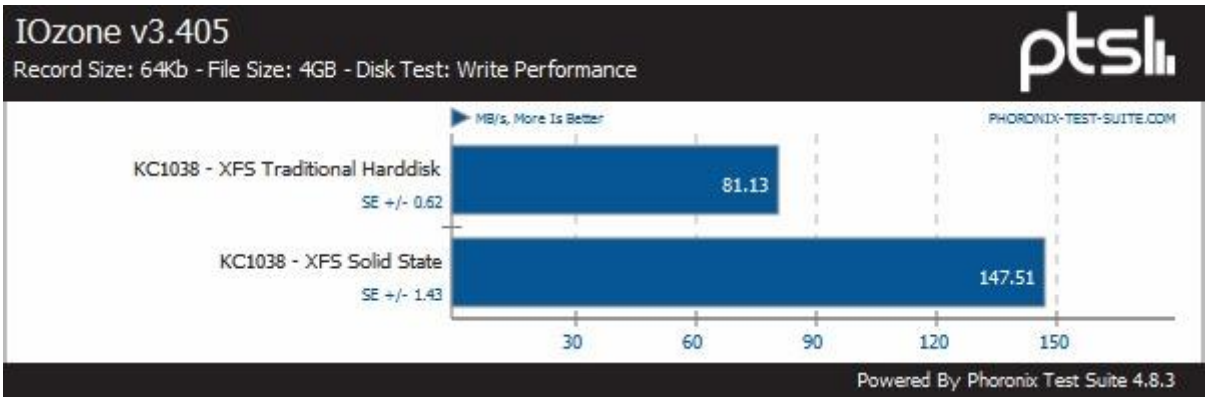

IOzone v3.405
Record Size: 1MB - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - XFS Traditional Harddisk
SE +/- 1.39
79.60
KC1038 - XFS Solid State
SE +/- 0.27
149.47
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3

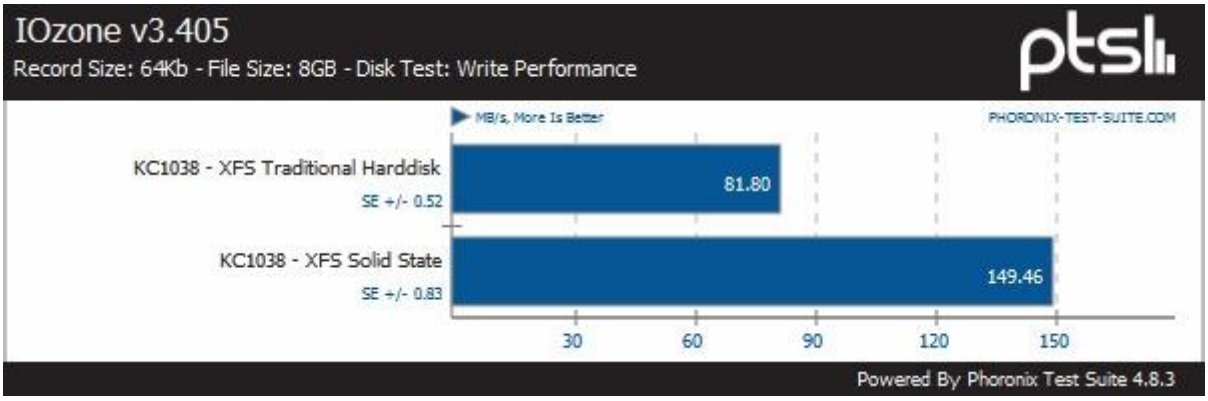

IOzone v3.405
Record Size: 4Kb - File Size: 4GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - XFS Traditional Harddisk
SE +/- 0.67
80.90
KC1038 - XFS Solid State
SE +/- 1.34
149.07
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3


IOzone v3.405
Record Size: 4Kb - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - XFS Traditional Harddisk
SE +/- 1.36
79.48
KC1038 - XFS Solid State
SE +/- 0.43
151.06
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3

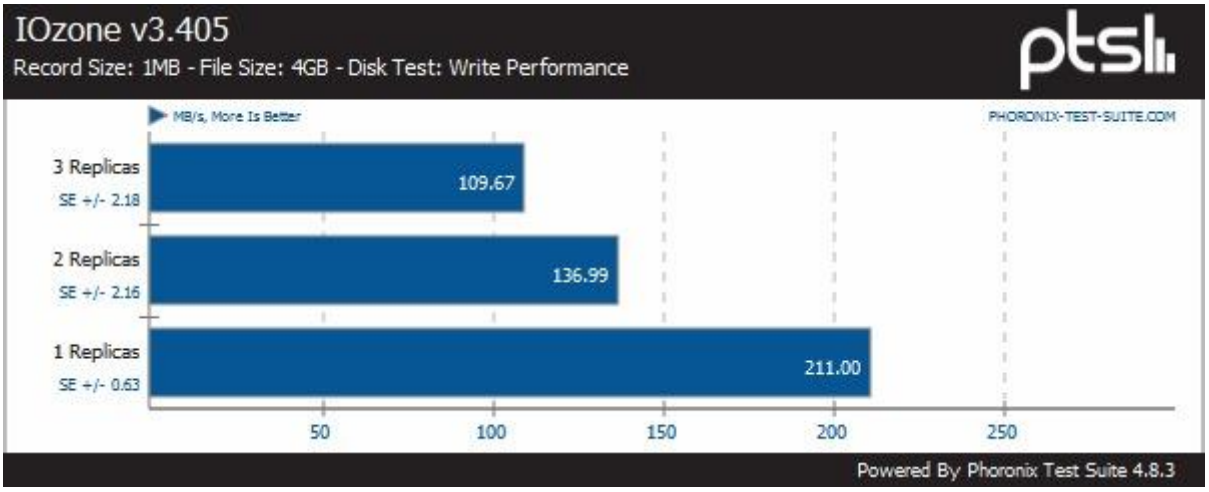

IOzone v3.405
Record Size: 64Kb - File Size: 4GB - Disk Test: Read Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - XFS Traditional Harddisk
SE +/- 0.08
81.29
KC1038 - XFS Solid State
SE +/- 0.15
196.50
40
80
120
160
200
Powered By Phoronix Test Suite 4.8.3

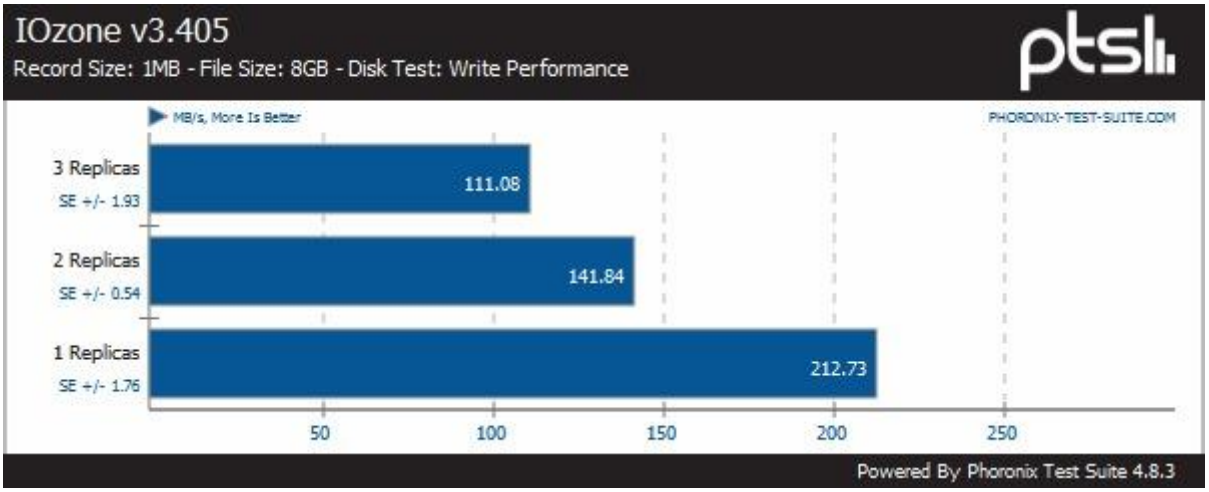

IOzone v3.405
Record Size: 64Kb - File Size: 8GB - Disk Test: Read Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - XFS Traditional Harddisk
SE +/- 0.84
78.98
KC1038 - XFS Solid State
SE +/- 0.42
195.90
40
80
120
160
200
Powered By Phoronix Test Suite 4.8.3

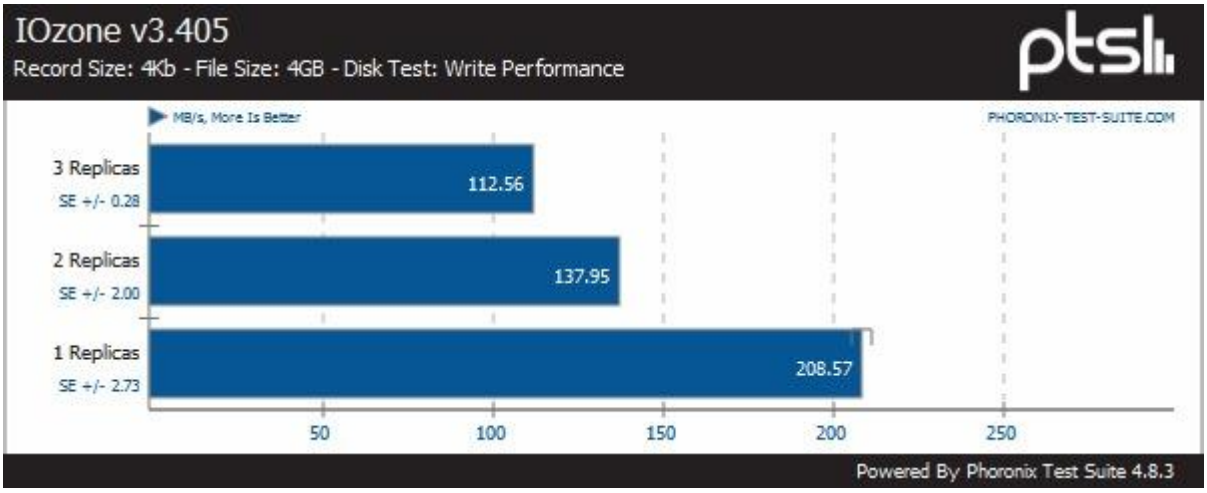

IOzone v3.405
Record Size: 64Kb - File Size: 4GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - XFS Traditional Harddisk
SE +/- 0.62
81.13
KC1038 - XFS Solid State
SE +/- 1.43
147.51
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3

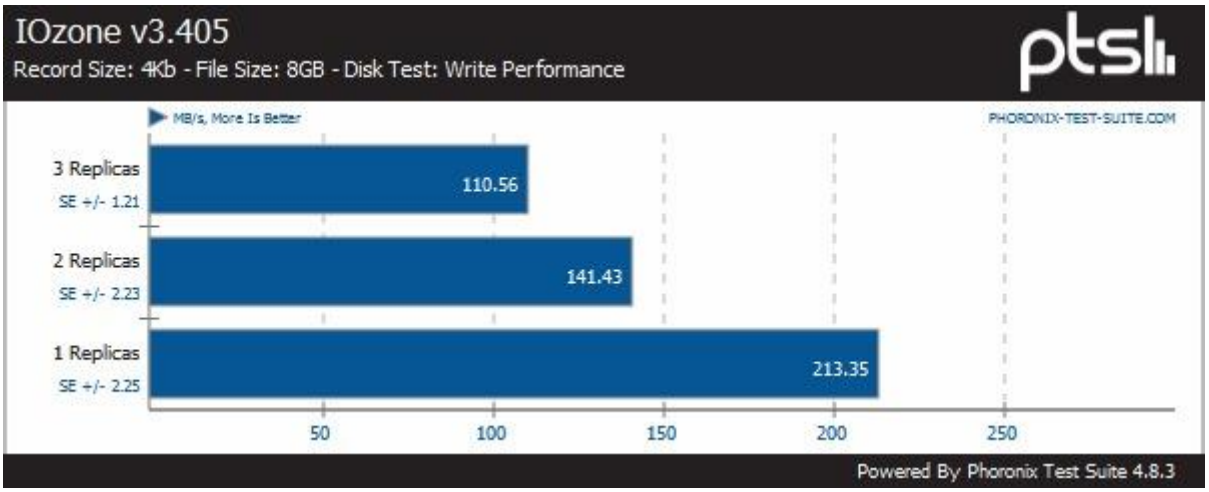

IOzone v3.405
Record Size: 64Kb - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
KC1038 - XFS Traditional Harddisk
SE +/- 0.52
81.80
KC1038 - XFS Solid State
SE +/- 0.83
149.46
30
60
90
120
150
Powered By Phoronix Test Suite 4.8.3

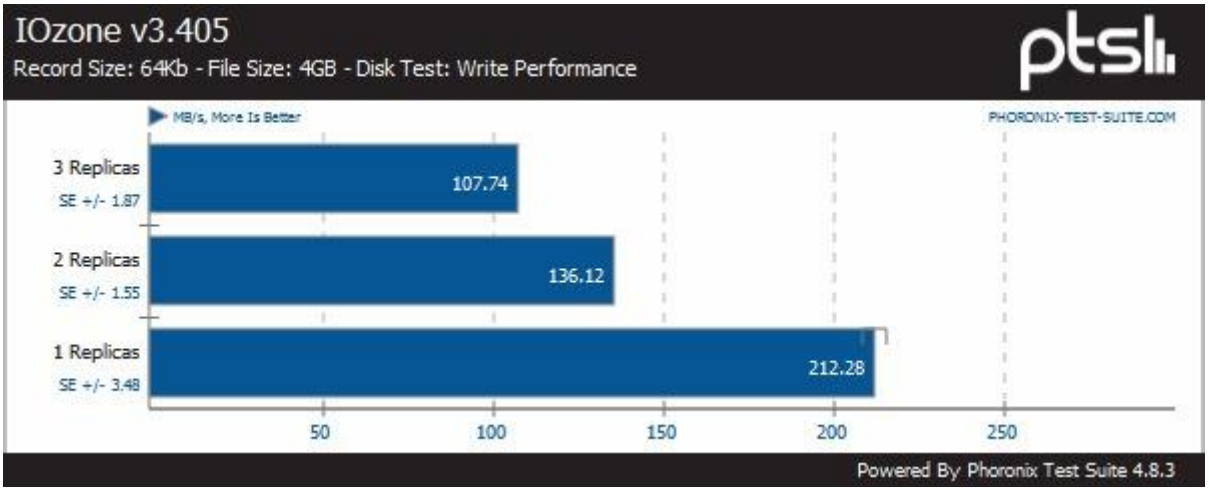
M.Bijlage Benchmark replica's

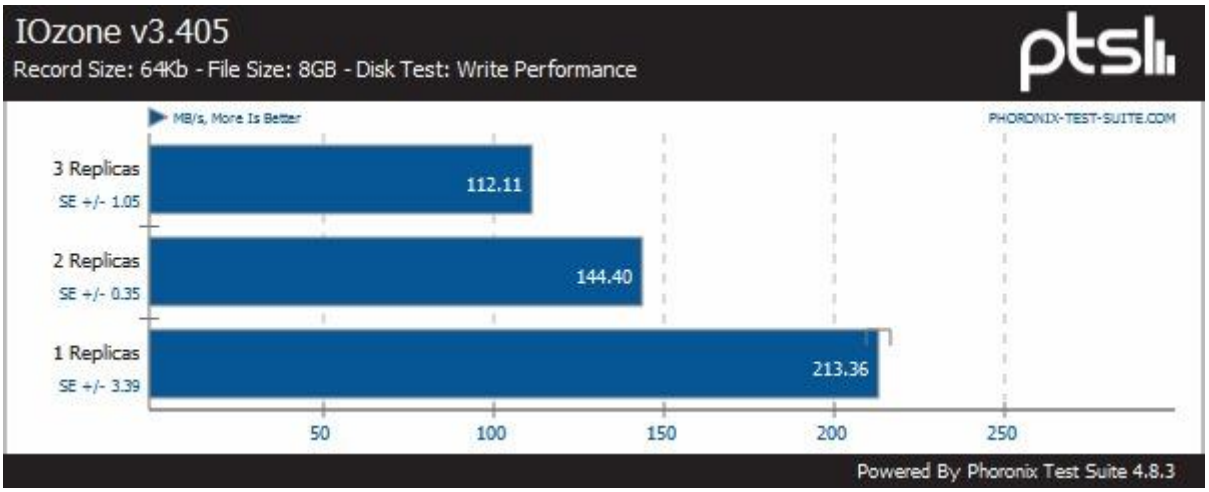







IOzone v3.405
Record Size: 4Kb - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
3 Replicas
SE +/- 1.21
110.56
2 Replicas
SE +/- 2.23
141.43
1 Replicas
SE +/- 2.25
213.35
50
100
150
200
250
Powered By Phoronix Test Suite 4.8.3


IOzone v3.405
Record Size: 64Kb - File Size: 4GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
3 Replicas
SE +/- 1.87
107.74
2 Replicas
SE +/- 1.55
136.12
1 Replicas
SE +/- 3.48
212.28
50
100
150
200
250
Powered By Phoronix Test Suite 4.8.3


IOzone v3.405
Record Size: 64Kb - File Size: 8GB - Disk Test: Write Performance
MB/s, More Is Better
PHORONIX-TEST-SUITE.COM
3 Replicas
SE +/- 1.05
112.11
2 Replicas
SE +/- 0.35
144.40
1 Replicas
SE +/- 3.39
213.36
50
100
150
200
250
Powered By Phoronix Test Suite 4.8.3

N. Bijlage Cisco configuratie LACP

```
interface Port-channel1
 description KC0227-OSD
 switchport
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 spanning-tree portfast
!
interface Port-channel2
 description KC0227-MON
 switchport
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 spanning-tree portfast
!
interface Port-channel9
 description KC0229-OSD
 switchport
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 spanning-tree portfast
!
interface Port-channel10
 description KC0229-MON
 switchport
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 spanning-tree portfast
!
interface Port-channel17
 description KC0230-OSD
 switchport
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 spanning-tree portfast
!
interface Port-channel18
 description KC0230-MON
 switchport
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 spanning-tree portfast
!
interface Port-channel35
 description KC0424
 switchport
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 spanning-tree portfast
!
interface Port-channel43
 description KX-MB-0030
 switchport
 switchport trunk encapsulation dot1q
 switchport mode trunk
 switchport nonegotiate
 no logging event link-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
!
interface GigabitEthernet1/1
```

```
 description KC0227
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 1 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/2
 description KC0227
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 2 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/3
 description KC0227
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 1 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/4
 description KC0227
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 2 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/5
 description KC0227
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 1 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/6
 description KC0227
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 2 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/7
 description KC0227
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
```

```
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 1 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/9
 description KC0229
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 9 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/10
 description KC0229
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 10 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/11
 description KC0229
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 9 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/12
 description KC0229
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 10 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/13
 description KC0229
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 9 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/14
 description KC0229
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
```

```
 channel-group 10 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/15
 description KC0229
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 9 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/17
 description KC0230
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 17 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/18
 description KC0230
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 18 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/19
 description KC0230
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 17 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/20
 description KC0230
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 18 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/21
 description KC0230
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 17 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/22
 description KC0230
 switchport access vlan 83
```

```
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 18 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/23
 description KC0230
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 17 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/31
 description KC0424
 switchport access vlan 83
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 spanning-tree portfast
!
interface GigabitEthernet1/35
 description KC0424
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 35 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/36
 description KC0424
 switchport access vlan 81
 switchport mode access
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 speed 1000
 duplex full
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 35 mode active
 spanning-tree portfast
!
interface GigabitEthernet1/43
 description KX-MB-0030
 switchport trunk encapsulation dot1q
 switchport mode trunk
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 43 mode active
!
interface GigabitEthernet1/44
 description KX-MB-0030
 switchport trunk encapsulation dot1q
 switchport mode trunk
 switchport nonegotiate
 no logging event link-status
 no logging event trunk-status
 no snmp trap link-status
 storm-control broadcast level 0.20
 storm-control action trap
 channel-protocol lacp
 channel-group 43 mode active
```

A. Bijlage OSD commando's

```
ceph-deploy osd create --zap-disk KC0227-OSD:/dev/disk/by-id/wwn-
0x5000c5003456e493:/dev/disk/by-id/wwn-0x500253805001c4cb-part1
ceph-deploy osd create --zap-disk KC0227-OSD:/dev/disk/by-id/wwn-
0x5000c50034513ef7:/dev/disk/by-id/wwn-0x500253805001c4cb-part2
ceph-deploy osd create --zap-disk KC0227-OSD:/dev/disk/by-id/wwn-
0x5000c5003456cac7:/dev/disk/by-id/wwn-0x500253805001c4cb-part3
ceph-deploy osd create --zap-disk KC0227-OSD:/dev/disk/by-id/wwn-
0x5000c50026131b1f:/dev/disk/by-id/wwn-0x500253805001c4cb-part4
ceph-deploy osd create --zap-disk KC0227-OSD:/dev/disk/by-id/wwn-
0x5000c5003456b233:/dev/disk/by-id/wwn-0x500253805001c4cb-part5

ceph-deploy osd create --zap-disk KC0229-OSD:/dev/disk/by-id/wwn-
0x5000c50034568653:/dev/disk/by-id/wwn-0x500253805001c4a7-part1
ceph-deploy osd create --zap-disk KC0229-OSD:/dev/disk/by-id/wwn-
0x5000c5003456d9f3:/dev/disk/by-id/wwn-0x500253805001c4a7-part2
ceph-deploy osd create --zap-disk KC0229-OSD:/dev/disk/by-id/wwn-
0x5000c5003456ecbb:/dev/disk/by-id/wwn-0x500253805001c4a7-part3
ceph-deploy osd create --zap-disk KC0229-OSD:/dev/disk/by-id/wwn-
0x50000395081a6fa8:/dev/disk/by-id/wwn-0x500253805001c4a7-part4
ceph-deploy osd create --zap-disk KC0229-OSD:/dev/disk/by-id/wwn-
0x5000c50041943e23:/dev/disk/by-id/wwn-0x500253805001c4a7-part5

ceph-deploy osd create --zap-disk KC0230-OSD:/dev/disk/by-id/wwn-
0x5000c500345fef4f:/dev/disk/by-id/wwn-0x500253805001c4c4-part1
ceph-deploy osd create --zap-disk KC0230-OSD:/dev/disk/by-id/wwn-
0x5000c5004193f8d7:/dev/disk/by-id/wwn-0x500253805001c4c4-part2
ceph-deploy osd create --zap-disk KC0230-OSD:/dev/disk/by-id/wwn-
0x5000c500345b32db:/dev/disk/by-id/wwn-0x500253805001c4c4-part3
ceph-deploy osd create --zap-disk KC0230-OSD:/dev/disk/by-id/wwn-
0x5000c500345feefb:/dev/disk/by-id/wwn-0x500253805001c4c4-part4
ceph-deploy osd create --zap-disk KC0230-OSD:/dev/disk/by-id/wwn-
0x5000c500345d478f:/dev/disk/by-id/wwn-0x500253805001c4c4-part5

ceph-deploy osd create --zap-disk KC0424-OSD:/dev/disk/by-id/wwn-
0x6848f690dfe905001c01daca09d9ba86:/dev/disk/by-id/wwn-
0x6848f690dfe905001c05cf6811a833d3-part1
ceph-deploy osd create --zap-disk KC0424-OSD:/dev/disk/by-id/wwn-
0x6848f690dfe905001c05cf6811a7b05f:/dev/disk/by-id/wwn-
0x6848f690dfe905001c05cf6811a833d3-part2
```

O. Bijlage Dell R815 Specificaties

Dell PowerEdge R815

The Dell™ PowerEdge™ R815 is a 4-socket rack server offering AMD Opteron™ processors with up to 64 processor cores and an excellent balance of advanced manageability, memory scalability, I/O and redundancy in a space-saving 2U form factor.

Performance-driven technology

Purpose-built to provide a reliable foundation and balanced with AMD Opteron processor technology for stability, consistency and outstanding price for performance, the Dell PowerEdge R815 is designed from the ground up for value and performance. It is an ideal platform for customers looking to maximize data center space and budget.

Purpose-built for reliability

We have listened to our customers and built servers endowed with reliability, availability and quality. Dell's reliability promise is simple — deliver quality products that stand the test of time.

From the robust, metal hard drive carriers and organized cabling to the interactive LCD screen, dual SD hypervisor modules, and embedded diagnostics, all components are purposefully designed and built to improve server reliability and usability. Dell has developed its reliability processes by:

- Employing robust validation and testing processes for durable product design
- Verifying each supplier meets stringent quality standards
- Implementing a "one-touch" build process to ensure that one person is responsible for the entire server build, resulting in greater quality control
- Ensuring that every fully configured Dell server is tested (and re-tested) before it leaves the factory
- Introducing Unified Server Configurator (USC) which offers embedded and persistent diagnostics with no media required to help minimize downtime
- Utilizing robust and durable industrial materials to enable long product lifecycles
- Improving redundancy with each generation by adding features such as dual internal failsafe SD modules to provide failover at the hypervisor level

In addition to these quality measures, all PowerEdge servers are designed with external ports, power supplies, and LED lights or LCD screens in the same location for a more efficient and familiar user experience, as well as easy installation and deployment.

Efficient infrastructure

Limitations on space, power, and cooling capacity combined with rising energy costs present enormous challenges for IT. Dell PowerEdge servers are made to deliver energy efficiency as a design standard while helping to meet the performance and budget goals your infrastructure requires. Energy-efficient system design built with Energy Smart technologies such as power capping, power inventory, and power budgeting help to better manage power within your specific environment.

Efficient fans are included in the PowerEdge R815 chassis, which spin in accordance with server workload demands. In addition, the internal shrouding and logical layout of the internal components aids with airflow direction, helping to keep the server cool. The PowerEdge R815 also includes power-management features such as volt regulators, power-regulating processors, and an interactive LCD screen for access to power consumption information.

Simplified systems management, without compromise

The Dell OpenManage™ systems management portfolio includes Integrated Dell Remote Access Controller 6 (iDRAC6). This embedded feature helps IT administrators manage Dell servers in physical, virtual, local and remote environments, operating in-band or out-of-band, with or without a systems management software agent installed.

OpenManage iDRAC with Lifecycle Controller integrates and connects to leading third-party systems management solutions (such as those from Microsoft, VMware, and BMC Software), so users can maintain a single point of control and capitalize on an existing systems management investment. OpenManage simplifies the lifecycle of deploying, updating, monitoring, and maintaining Dell PowerEdge servers.

Ideal for virtualization, medium-sized databases, high-performance computing, and other highly threaded application implementations.

September 2013

Feature	PowerEdge R815 Technical Specification	
Form factor	2U rack	
Processors	AMD Opteron™ 6100, 6200 and 6300 series processors	
Processor Sockets	4	
Frontside bus or HyperTransport	HyperTransport-3 Links	
Cache	L2: 512K/core L3: 16MB	
Chipset	AMD (SR5650, SR5670 and SP5100)	
Memory[1]	Up to 512GB (32 DIMM Slots): 1GB/2GB/4GB/8GB/16GB/32GB up to 1600MT/s	
I/O slots	6 PCIe 2.0 slots + 1 storage slot: Five x8 slots One x4 slot One x4 storage slot	
RAID controllers	**Internal controllers:** PERC H200 (6Gb/s) PERC H700 (6Gb/s) with 512MB battery-backed cache; 512MB, 1GB Non-Volatile battery-backed cache External HBAs (non-RAID): 6Gb/s SAS HBA SAS 5/E HBA LSI2032 PCIe SCSI HBA	**External controllers:** SAS 5/E with 512MB battery-backed cache PERC H800 (6Gb/s) with 512MB of battery-backed cache; 512MB, 1GB Non-Volatile battery-backed cache PERC 6/E with 256MB or 512MB of battery-backed cache
Drive bays	Up to six 2.5" hot-plug SATA SSD, SAS, nearline SAS, or SATA drives	
Maximum Internal Storage[1]	Up to 6TB	
Hard drives	Hot-plug hard drive options: 2.5" SAS SSD, SATA SSD, SAS (15K, 10K), nearline SAS (7.2K), SATA (7.2K), SAS 512e (15K)	
Communications	Intel® 10GBase-T Copper Single Port NIC, PCI-E x8 Intel Single Port Server Adapter, 10Gigabit, SR Optical, PCI-E x8 Broadcom® BMC57710 10Base-T Copper Single Port NIC, PCI-E x8 Broadcom 10GbE NIC, Broadcom Dual Port 10GbE SFP+ Intel Gigabit ET Dual Port Server Adapter Intel Gigabit ET Quad Port Server Adapter Brocade® CNA Dual port adapter Emulex® CNA iSCSI HBA stand up adapter OCE10102-IX-D Emulex CNA iSCSI HBA stand up adapter OCE10102-FX-D	**Optional add-in HBAs:** QLogic® QLE 2462 FC4 Dual Port 4 Gbps Fiber Channel HBA QLogic QLE 220 FC4 Single Port 4 Gbps Fiber Channel HBA QLogic QLE 2460 FC4 Single Port 4 Gbps Fiber Channel HBA QLogic QLE2562 FC8 Dual-channel HBA, PCIe 2.0 x4 QLogic QLE2560 FC8 Single-channel HBA, PCIe 2.0 x4 QLogic QLE2660 FC16 Single-channel HBA, PCIe 3.0 x4 QLogic QLE2662 FC16 Dual-channel HBA, PCIe 3.0 x4 Emulex LPe-1150 FC4 Single Port 4 Gbps Fiber Channel HBA, PCIe x4 Emulex LPe-11002 FC4 Dual Port 4 Gbps Fiber Channel HBA, PCIe x4 Emulex LPe-12000 FC8 Single Port 4 Gbps Fiber Channel HBA, PCIe 2.0 x4 Emulex LPe-12002 FC8 Dual Port 4 Gbps Fiber Channel HBA, PCIe 2.0 x4 Brocade FC4 and 8GB HBAs
Power supplies	One hot-pluggable non-redundant 1100W power supply	Two hot-pluggable redundant 1100W power supplies
Availability	Hot-plug hard drives, hot-plug redundant power, dual SD modules, ECC memory, interactive LCD screen	
Video	Matrox® G200eW with 8MB memory	
Remote management	iDRAC6 Express (standard), iDRAC6 Enterprise, and vFlash media (upgrade optional)	
Systems management	Dell OpenManage™ BMC, IPMI 2.0 compliant Unified Server Configurator Lifecycle Controller enabled: iDRAC6 Express, optional iDRAC6 Enterprise, and vFlash media Microsoft® System Center Essential (SCE) 2010 v2	
Rack support	ReadyRails™ sliding rails with optional cable management arm for 4-post racks (optional adapter brackets required for threaded hole racks)	
Operating systems	Microsoft Windows Server® 2012 Microsoft Windows Server 2008 SP2, x86/x64 (x64 includes Hyper-V®) Microsoft Windows Server 2008 R2 SP1, x64 (includes Hyper-V) Microsoft Windows® HPC Server 2008 R2 Novell® SUSE® Linux Enterprise Server Red Hat® Enterprise Linux® Oracle® Solaris™	**Virtualization options:** Citrix® XenServer® VMware® vSphere® and ESXi™ Red Hat Enterprise Virtualization® For more information on the specific versions and additions, visit Dell.com/OSsupport.
Featured database applications	Microsoft SQL Server® solutions (see Dell.com/SQL) Oracle database solutions (see Dell.com/Oracle)	

[1] GB means 1 billion bytes and TB equals 1 trillion bytes; actual capacity varies with preloaded material and operating environment and will be less.

Dell Services

Dell Services can help reduce IT complexity, lower costs, and eliminate inefficiencies by making IT and business solutions work harder for you. The Dell Services team takes a holistic view of your needs and designs solutions for your environment and business objectives while leveraging proven delivery methods, local talent, and in-depth domain knowledge for the lowest TCO.

Learn More at Dell.com/PowerEdge

P. Bijlage Proxmox Cluster

```
root@KC0227-MON:~# pvecm status
Version: 6.2.0
Config Version: 4
Cluster Name: CEPH-PROX-CLSTR
Cluster Id: 37366
Cluster Member: Yes
Cluster Generation: 16
Membership state: Cluster-Member
Nodes: 4
Expected votes: 4
Total votes: 4
Node votes: 1
Quorum: 3
Active subsystems: 5
Flags:
Ports Bound: 0
Node name: KC0227-MON
Node ID: 2
Multicast addresses: 239.192.145.136
Node addresses: 172.16.83.10
```

```
root@KC0227-MON:~# pvecm nodes
Node  Sts  Inc  Joined                Name
1     M    8    2015-01-25 22:27:36   KC0424-MON
2     M    4    2015-01-25 22:27:36   KC0227-MON
3     M    12   2015-01-25 22:28:06   KC0230-MON
4     M    16   2015-01-25 22:28:35   KC0229-MON
```

Q. Bijlage Archimate Technology Layer Concepts

Behaviour Layer

"An infrastructure service is defined as an externally visible unit of functionality, provided by one or more nodes, exposed through well-defined interfaces, and meaningful to the environment."

Figuur 67: ArchiMate Infrastructure Service

"An infrastructure function is defined as a behavior element that groups infrastructural behavior that can be performed by a node."

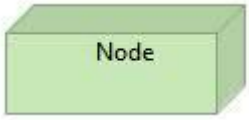
Figuur 68: Archimate Infrastructure Function

Active Layer

"A device is defined as a hardware resource upon which artifacts may be stored or deployed for execution."

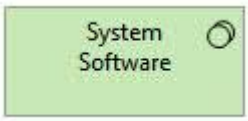
Figuur 69: Archimate Device

"A node is defined as a computational resource upon which artifacts may be stored or deployed for execution."



Figuur 70: Archimate Node

"System software represents a software environment for specific types of components and objects that are deployed on it in the form of artifacts."



Figuur 71: Archimate System Software

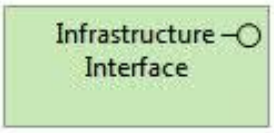
"A communication path is defined as a link between two or more nodes, through which these nodes can exchange data."

Figuur 72: ArchiMate Communication Path

“A network is defined as a communication medium between two or more devices.”

Figuur 73: ArchiMate Network

“An infrastructure interface is defined as a point of access where infrastructure services offered by a node can be accessed by other nodes and application components.”



Figuur 74: ArchiMate Infrastructure Interface

R. Bijlage LACP mogelijkheden

Zonder ondersteuning van src-dst-port

```
KX-MB-0030(config)#port-channel load-balance ?
 dst-ip Dst IP Addr
 dst-mac Dst Mac Addr
 src-dst-ip Src XOR Dst IP Addr
 src-dst-mac Src XOR Dst Mac Addr
 src-ip Src IP Addr
 src-mac Src Mac Addr
```

Met ondersteuning van src-dst-port

```
KC0239(config)#port-channel load-balance ?
 dst-ip Dst IP Addr
 dst-mac Dst Mac Addr
 dst-port Dst TCP/UDP Port
 src-dst-ip Src XOR Dst IP Addr
 src-dst-mac Src XOR Dst Mac Addr
 src-dst-port Src XOR Dst TCP/UDP Port
 src-ip Src IP Addr
 src-mac Src Mac Addr
 src-port Src TCP/UDP Port
```

S. Bijlage NMAP scan resultaten

Netwerk 172.16.81.0/24

TCP

```
root@kc1036:~# nmap -sT 172.16.81.7,9,10

Starting Nmap 6.40 ( http://nmap.org ) at 2015-02-25 12:43 CET
Nmap scan report for KC0230-OSD.afstudeer.org (172.16.81.7)
Host is up (0.00091s latency).
Not shown: 996 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
111/tcp  open  rpcbind
3128/tcp open  squid-http
MAC Address: 00:1B:21:D2:81:00 (Intel Corporate)

Nmap scan report for KC0229-OSD.afstudeer.org (172.16.81.9)
Host is up (0.00096s latency).
Not shown: 996 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
111/tcp  open  rpcbind
3128/tcp open  squid-http
MAC Address: 00:1B:21:D2:80:E0 (Intel Corporate)

Nmap scan report for KC0227-OSD.afstudeer.org (172.16.81.10)
Host is up (0.0011s latency).
Not shown: 996 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
111/tcp  open  rpcbind
3128/tcp open  squid-http
MAC Address: 00:1B:21:D2:60:E0 (Intel Corporate)

Nmap done: 3 IP addresses (3 hosts up) scanned in 0.57 seconds
```

UDP

```
root@kc1036:~# nmap -sU 172.16.81.7,9,10

Starting Nmap 6.40 ( http://nmap.org ) at 2015-02-25 12:43 CET
Nmap scan report for KC0230-OSD.afstudeer.org (172.16.81.7)
Host is up (0.00034s latency).
Not shown: 997 closed ports
PORT    STATE SERVICE
111/udp open  rpcbind
123/udp open  ntp
161/udp open  snmp
MAC Address: 00:1B:21:D2:81:00 (Intel Corporate)

Nmap scan report for KC0229-OSD.afstudeer.org (172.16.81.9)
Host is up (0.00043s latency).
Not shown: 997 closed ports
PORT    STATE SERVICE
111/udp open  rpcbind
123/udp open  ntp
161/udp open  snmp
MAC Address: 00:1B:21:D2:80:E0 (Intel Corporate)

Nmap scan report for KC0227-OSD.afstudeer.org (172.16.81.10)
Host is up (0.00031s latency).
Not shown: 997 closed ports
PORT    STATE SERVICE
111/udp open  rpcbind
123/udp open  ntp
161/udp open  snmp
MAC Address: 00:1B:21:D2:60:E0 (Intel Corporate)

Nmap done: 3 IP addresses (3 hosts up) scanned in 1121.31 seconds
```

Netwerk 172.16.83.0/24

TCP

```
root@kc1036:~# nmap -sT 172.16.83.7,9,10

Starting Nmap 6.40 ( http://nmap.org ) at 2015-02-25 13:08 CET
Nmap scan report for KC0230-MON.afstudeer.org (172.16.83.7)
Host is up (0.0010s latency).
Not shown: 995 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
111/tcp  open  rpcbind
3128/tcp open  squid-http
6789/tcp open  ibm-db2-admin
MAC Address: D0:67:E5:F9:5A:A7 (Dell)

Nmap scan report for KC0229-MON.afstudeer.org (172.16.83.9)
Host is up (0.0011s latency).
Not shown: 995 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
111/tcp  open  rpcbind
3128/tcp open  squid-http
6789/tcp open  ibm-db2-admin
MAC Address: D0:67:E5:F9:68:8A (Dell)

Nmap scan report for KC0227-MON.afstudeer.org (172.16.83.10)
Host is up (0.0011s latency).
Not shown: 995 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
111/tcp  open  rpcbind
3128/tcp open  squid-http
6789/tcp open  ibm-db2-admin
MAC Address: D0:67:E5:F9:63:CB (Dell)

Nmap done: 3 IP addresses (3 hosts up) scanned in 0.57 seconds
```

UDP

```
root@kc1036:~# nmap -sU 172.16.83.7,9,10

Starting Nmap 6.40 ( http://nmap.org ) at 2015-02-25 13:08 CET
Nmap scan report for KC0230-MON.afstudeer.org (172.16.83.7)
Host is up (0.00033s latency).
Not shown: 997 closed ports
PORT    STATE SERVICE
111/udp open  rpcbind
123/udp open  ntp
161/udp open  snmp
MAC Address: D0:67:E5:F9:5A:A7 (Dell)

Nmap scan report for KC0229-MON.afstudeer.org (172.16.83.9)
Host is up (0.00042s latency).
Not shown: 997 closed ports
PORT    STATE SERVICE
111/udp open  rpcbind
123/udp open  ntp
161/udp open  snmp
MAC Address: D0:67:E5:F9:68:8A (Dell)

Nmap scan report for KC0227-MON.afstudeer.org (172.16.83.10)
Host is up (0.00025s latency).
Not shown: 997 closed ports
PORT    STATE SERVICE
111/udp open  rpcbind
123/udp open  ntp
161/udp open  snmp
MAC Address: D0:67:E5:F9:63:CB (Dell)

Nmap done: 3 IP addresses (3 hosts up) scanned in 1075.54 seconds
```

Netwerk 192.168.0.0/24

TCP

```
root@kc1036:~# nmap -sT 192.168.0.7,9,10

Starting Nmap 6.40 ( http://nmap.org ) at 2015-02-25 13:45 CET
Nmap scan report for kc0230.afstudeer.org (192.168.0.7)
Host is up (0.00074s latency).
Not shown: 996 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
111/tcp  open  rpcbind
3128/tcp open  squid-http
MAC Address: D0:67:E5:F9:5A:A5 (Dell)

Nmap scan report for kc0229.afstudeer.org (192.168.0.9)
Host is up (0.00080s latency).
Not shown: 996 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
111/tcp  open  rpcbind
3128/tcp open  squid-http
MAC Address: D0:67:E5:F9:68:88 (Dell)

Nmap scan report for kc0227.afstudeer.org (192.168.0.10)
Host is up (0.00086s latency).
Not shown: 996 closed ports
PORT     STATE SERVICE
22/tcp   open  ssh
25/tcp   open  smtp
111/tcp  open  rpcbind
3128/tcp open  squid-http
MAC Address: D0:67:E5:F9:63:C9 (Dell)

Nmap done: 3 IP addresses (3 hosts up) scanned in 0.52 seconds
```

UDP

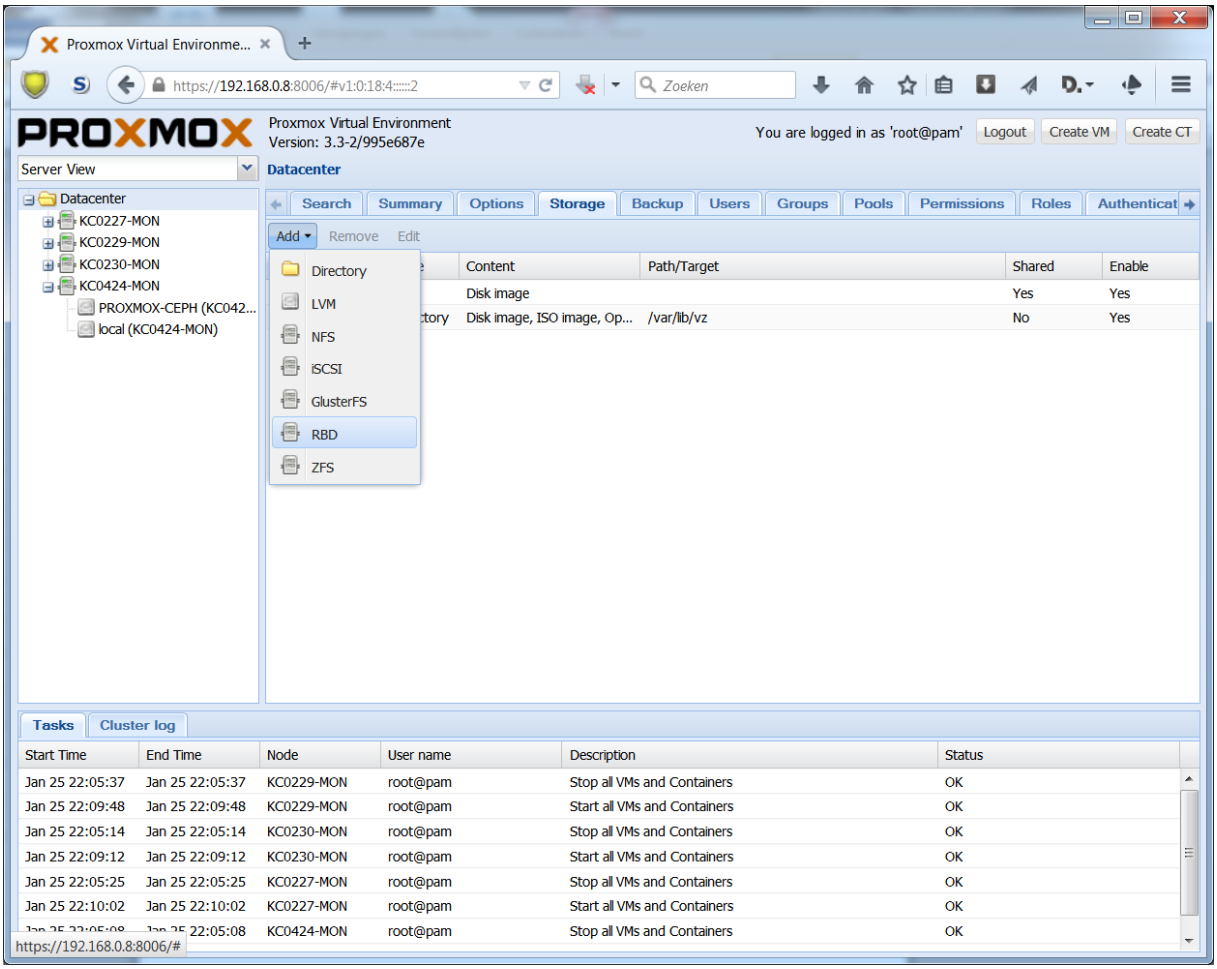
```
root@kc1036:~# nmap -sU 192.168.0.7,9,10

Starting Nmap 6.40 ( http://nmap.org ) at 2015-02-25 13:27 CET
Nmap scan report for kc0230.afstudeer.org (192.168.0.7)
Host is up (0.00040s latency).
Not shown: 997 closed ports
PORT    STATE SERVICE
111/udp open  rpcbind
123/udp open  ntp
161/udp open  snmp
MAC Address: D0:67:E5:F9:5A:A5 (Dell)

Nmap scan report for kc0229.afstudeer.org (192.168.0.9)
Host is up (0.00034s latency).
Not shown: 997 closed ports
PORT    STATE SERVICE
111/udp open  rpcbind
123/udp open  ntp
161/udp open  snmp
MAC Address: D0:67:E5:F9:68:88 (Dell)

Nmap scan report for kc0227.afstudeer.org (192.168.0.10)
Host is up (0.00023s latency).
Not shown: 997 closed ports
PORT    STATE SERVICE
111/udp open  rpcbind
123/udp open  ntp
161/udp open  snmp
MAC Address: D0:67:E5:F9:63:C9 (Dell)

Nmap done: 3 IP addresses (3 hosts up) scanned in 1090.81 seconds
```

T. Bijlage lsof

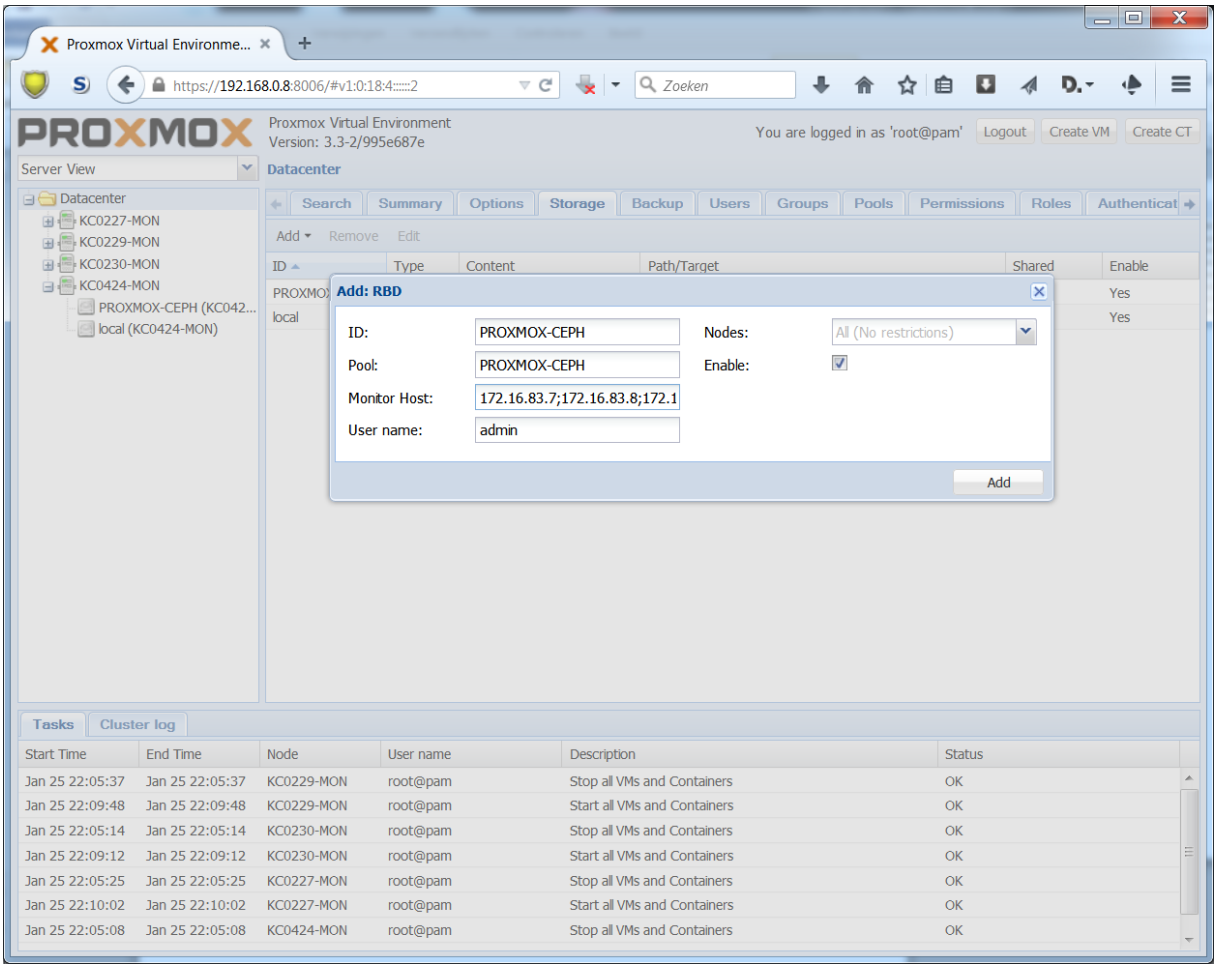
```
root@KC0227-MON:~# for i in `netstat -an | grep LISTEN | grep ^tcp | awk '{print $4}' | grep -v :: |
awk -F\: '{print $2}'`; do lsof -n -i4TCP:$i | grep LISTEN; done
ceph-osd 23114 root   9u  IPv4  52035      0t0  TCP 172.16.83.10:6814 (LISTEN)
rpc.statd 3234 statd   8u  IPv4  16529      0t0  TCP *:51873 (LISTEN)
ceph-mon  24791 root  12u  IPv4   60999      0t0  TCP 172.16.83.10:6789 (LISTEN)
pveproxy  25420 www-data   5u  IPv4   69786      0t0  TCP *:8006 (LISTEN)
pveproxy 171770 www-data   5u  IPv4   69786      0t0  TCP *:8006 (LISTEN)
pveproxy 171772 www-data   5u  IPv4   69786      0t0  TCP *:8006 (LISTEN)
pveproxy 171773 www-data   5u  IPv4   69786      0t0  TCP *:8006 (LISTEN)
rpcbind 3217 root   8u  IPv4  16465      0t0  TCP *:sunrpc (LISTEN)
ceph-osd   7589 root   5u  IPv4   27107      0t0  TCP 172.16.83.10:6800 (LISTEN)
ceph-osd   7589 root   6u  IPv4   27108      0t0  TCP 172.16.81.10:6800 (LISTEN)
ceph-osd   7589 root   5u  IPv4   27107      0t0  TCP 172.16.83.10:6800 (LISTEN)
ceph-osd   7589 root   6u  IPv4   27108      0t0  TCP 172.16.81.10:6800 (LISTEN)
ceph-osd  7589 root   7u  IPv4   27109      0t0  TCP 172.16.81.10:6801 (LISTEN)
ceph-osd  7589 root   8u  IPv4   27110      0t0  TCP 172.16.83.10:6801 (LISTEN)
ceph-osd  7589 root   7u  IPv4   27109      0t0  TCP 172.16.81.10:6801 (LISTEN)
ceph-osd  7589 root   8u  IPv4   27110      0t0  TCP 172.16.83.10:6801 (LISTEN)
ceph-osd  7589 root   9u  IPv4   27111      0t0  TCP 172.16.83.10:6802 (LISTEN)
ceph-osd 10872 root   6u  IPv4   30735      0t0  TCP 172.16.81.10:6802 (LISTEN)
ceph-osd  7589 root   9u  IPv4   27111      0t0  TCP 172.16.83.10:6802 (LISTEN)
ceph-osd 10872 root   6u  IPv4   30735      0t0  TCP 172.16.81.10:6802 (LISTEN)
ceph-osd  10872 root   5u  IPv4   30734      0t0  TCP 172.16.83.10:6803 (LISTEN)
ceph-osd  10872 root   7u  IPv4   30736      0t0  TCP 172.16.81.10:6803 (LISTEN)
ceph-osd  10872 root   5u  IPv4   30734      0t0  TCP 172.16.83.10:6803 (LISTEN)
ceph-osd  10872 root   7u  IPv4   30736      0t0  TCP 172.16.81.10:6803 (LISTEN)
ceph-osd 10872 root   8u  IPv4   30737      0t0  TCP 172.16.83.10:6804 (LISTEN)
ceph-osd 13518 root   6u  IPv4   33767      0t0  TCP 172.16.81.10:6804 (LISTEN)
ceph-osd 10872 root   8u  IPv4   30737      0t0  TCP 172.16.83.10:6804 (LISTEN)
ceph-osd 13518 root   6u  IPv4   33767      0t0  TCP 172.16.81.10:6804 (LISTEN)
pvedaemon  25400    root   5u  IPv4   69764      0t0  TCP 127.0.0.1:85 (LISTEN)
pvedaemon 257832    root   5u  IPv4   69764      0t0  TCP 127.0.0.1:85 (LISTEN)
pvedaemon 362011    root   5u  IPv4   69764      0t0  TCP 127.0.0.1:85 (LISTEN)
pvedaemon 365828    root   5u  IPv4   69764      0t0  TCP 127.0.0.1:85 (LISTEN)
ceph-osd 10872 root   9u  IPv4   30738      0t0  TCP 172.16.83.10:6805 (LISTEN)
ceph-osd 13518 root   7u  IPv4   33768      0t0  TCP 172.16.81.10:6805 (LISTEN)
ceph-osd 10872 root   9u  IPv4   30738      0t0  TCP 172.16.83.10:6805 (LISTEN)
ceph-osd 13518 root   7u  IPv4   33768      0t0  TCP 172.16.81.10:6805 (LISTEN)
ceph-osd  13518 root   5u  IPv4   33766      0t0  TCP 172.16.83.10:6806 (LISTEN)
ceph-osd  15576 root   6u  IPv4   36040      0t0  TCP 172.16.81.10:6806 (LISTEN)
ceph-osd  13518 root   5u  IPv4   33766      0t0  TCP 172.16.83.10:6806 (LISTEN)
ceph-osd  15576 root   6u  IPv4   36040      0t0  TCP 172.16.81.10:6806 (LISTEN)
sshd     3681 root   3u  IPv4   17447      0t0  TCP *:ssh (LISTEN)
ceph-osd 13518 root   8u  IPv4   33769      0t0  TCP 172.16.83.10:6807 (LISTEN)
ceph-osd 15576 root   7u  IPv4   36041      0t0  TCP 172.16.81.10:6807 (LISTEN)
ceph-osd 13518 root   8u  IPv4   33769      0t0  TCP 172.16.83.10:6807 (LISTEN)
ceph-osd 15576 root   7u  IPv4   36041      0t0  TCP 172.16.81.10:6807 (LISTEN)
ceph-osd 13518 root   9u  IPv4   33770      0t0  TCP 172.16.83.10:6808 (LISTEN)
ceph-osd 23114 root   6u  IPv4   52032      0t0  TCP 172.16.81.10:6808 (LISTEN)
```

```
ceph-osd 13518 root    9u  IPv4    33770      0t0  TCP 172.16.83.10:6808 (LISTEN)
ceph-osd 23114 root    6u  IPv4    52032      0t0  TCP 172.16.81.10:6808 (LISTEN)
spiceprox   4691 www-data    5u  IPv4  23191      0t0  TCP *:3128 (LISTEN)
spiceprox 235648 www-data    5u  IPv4  23191      0t0  TCP *:3128 (LISTEN)
master  198935 root   12u  IPv4 14447151      0t0  TCP *:smtp (LISTEN)
ceph-osd  15576 root    5u  IPv4    36039      0t0  TCP 172.16.83.10:6809 (LISTEN)
ceph-osd  23114 root    7u  IPv4    52033      0t0  TCP 172.16.81.10:6809 (LISTEN)
ceph-osd  15576 root    5u  IPv4    36039      0t0  TCP 172.16.83.10:6809 (LISTEN)
ceph-osd  23114 root    7u  IPv4    52033      0t0  TCP 172.16.81.10:6809 (LISTEN)
ceph-osd 15576 root    8u  IPv4    36042      0t0  TCP 172.16.83.10:6810 (LISTEN)
ceph-osd 15576 root    9u  IPv4  36043      0t0  TCP 172.16.83.10:6811 (LISTEN)
ceph-osd  23114 root    5u  IPv4    52031      0t0  TCP 172.16.83.10:6812 (LISTEN)
ceph-osd 23114 root    8u  IPv4    52034      0t0  TCP 172.16.83.10:6813 (LISTEN)
```

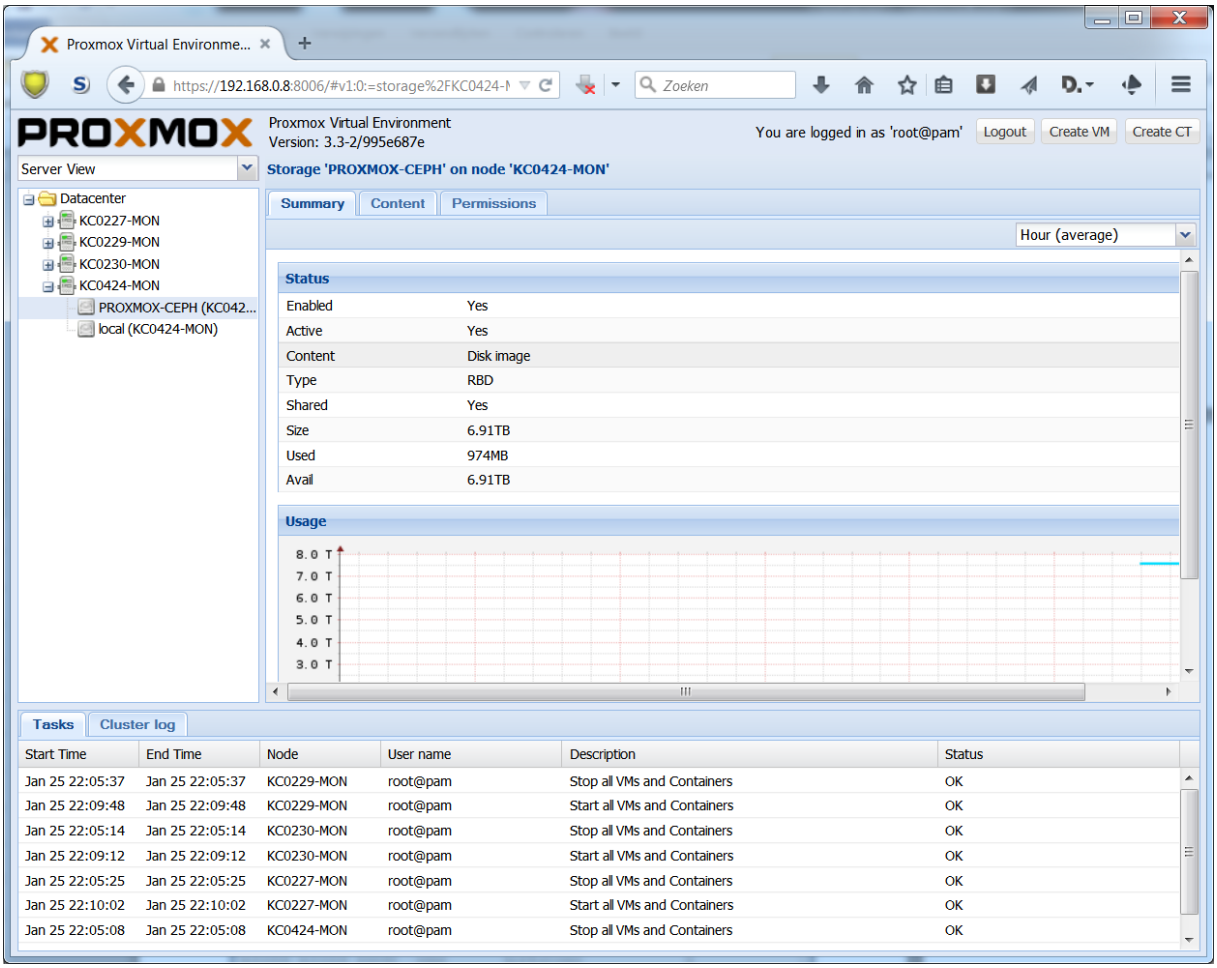
U. Bijlage Aanmaken Rados Block Device in Proxmox



Figuur 75 Selecteer het tab “Storage” en klik op ”Add”, daarna op “RBD”



Figuur 76 Geeft een ID en een Pool op voor de monitoring hosts op gescheiden door het symbool “;”



Figuur 77 Hierna is het RBD beschikbaar

V. Bijlage Compilatie van Calamari software

Voor Ceph Calamari zijn standaard geen packages te downloaden om te installeren. Deze zullen vanuit de source moeten worden gecompileerd. Calamari kent twee installaties; de Calamari-server en de Calamari-client. Voor beide packages zijn een aantal dependencies noodzakelijk voor ze kunnen worden geïnstalleerd;

- git
- curl
- libcairo2-dev
- python-pip
- libpq-dev
- python-dev
- python-virtualenv
- ruby1.9.1
- python-software-properties
- g++
- make
- debhelper
- build-essential
- devscripts

Belangrijk: tijdens het compileren is een verbinding met het internet noodzakelijk omdat er tijdens het proces nog diverse zaken worden gedownload en geïnstalleerd vanuit diverse repositories van github en npm. Voer deze handelingen dan ook uit op een werkplek waar vuil internet beschikbaar is. Zodra de packages klaar zijn zullen ze gescand moeten worden met een virusscanner. Belangrijk: De sources van Calamari spreken over een compilatie met behulp van vagrant – een virtuele omgeving zonder een virtuele machine in te zetten. Uiteindelijk zijn de packages in dit voorbeeld zonder hulp van vagrant gebouwd maar op een eigen virtual machine.

Installeer deze met behulp van apt-get met root-credentials:

```
root@extern:~# apt-get install -y git curl libcairo2-dev python-pip libpq-dev python-dev python-virtualenv
```

Volgende stap is het installeren van nodejs en npm. Op dit moment (augustus 2014) is nodejs is versie 0.10.10 nodig:

```
root@extern:~# git clone https://github.com/joyent/node.git
root@extern:~# cd node
root@extern:~/node# git checkout v0.10.10
root@extern:~/node# ./configure
root@extern:~/node# make -j4
root@extern:~/node# make install
Hierna wordt npm geinstalleerd:
root@extern:~# mkdir /root/npm
root@extern:~# cd /root/npm
root@extern:~/npm# wget --no-check-certificate http://npmjs.org/install.sh
root@extern:~/npm# sh install.sh
Als nodejs en npm aanwezig zijn kunnen er extra dependencies geinstalleerd worden:
root@extern:~# npm install -g bower
root@extern:~# npm install -g coffee-script
root@extern:~# npm install -g grunt-cli
```

```
Gebruik gem om vervolgens compass en sass te installeren:
root@extern:~#gem install compass
root@extern:~# gem install sass
```

Keer nu terug naar de default user.
Haal de Calamari en Calamari-clients op van github:

```
user@extern:~# git clone https://github.com/ceph/calamari.git
user@extern:~# git clone https://github.com/ceph/calamari-clients.git
```

Als eerste compileren we de Calamari-server. Hiervoor moet de source directory in de Debian directory worden verplaatst:

```
user@extern:~# cd calamari/
user@extern:~/calamari# cd debian
user@extern:~/calamari/debian# mv source source.old
user@extern:~/calamari/debian# cd ..
user@extern:~/calamari# dpkg-buildpackage
```

Je kan nu een kop koffie halen. Als het package klaar is moet er in de directory in niveau hierboven de volgende bestanden zijn aangemaakt:
calamari_1.0.0-1.tar.gz
calamari-server_1.0.0-1_all.deb

Vervolgens kan de Calamari-client worden gebouwd. Volgens de documentatie zou er een environment variabele moeten worden aangemaakt die wordt gebruikt tijdens het compileren van het script; tijdens het testdraaien bleek deze environment variabele niet gezien te worden en was de oplossing om deze variabele vooraf een waarde in het script te geven.
Pas daarvoor de Makefile aan:

```
user@extern:~# cd calamari-clients/
user@extern:~calamari-clients# vi Makefile
```

Voeg toe op de derde regel:
REAL_BUILD = y

Sla het document op en start het bouwen van het package

```
user@extern:~/calamari-clients# dpkg-buildpackage
```

Tijd voor de tweede kop koffie. Als het package klaar is zonder fouten dan moet in het bovengelegen niveau nieuwe bestanden zijn geplaatst:
calamari-clients_1.0.0-616-gf17527d.tar.gz
calamari-clients_1.0.0-616-gf17527d_all.deb

W. Bijlage compilatie Diamond

Voor het verzamelen van informatie op de Ceph nodes wordt gebruik gemaakt van een Python Deamon genaamd Diamond.

Deze deamon moet net zoals de Calamari-server en de Calamari-clients nog worden gecompileerd tot een debian package. Hiervoor zijn de volgende handelingen noodzakelijk:

Diamond heeft ook een aantal dependencies nodig, installeer deze eerst.
python-mock
python-configobj
cdbs

```
root@extern:~# apt-get install python-mock python-configobj cdbs
```

Download de Diamond source van github, het is belangrijk dat de Calamari branch wordt gekozen!

```
root@extern:~# git clone https://github.com/ceph/diamond.git --branch=calamari
root@extern:~# cd Diamond
root@extern:~/Diamond# dpkg-buildpackage
```

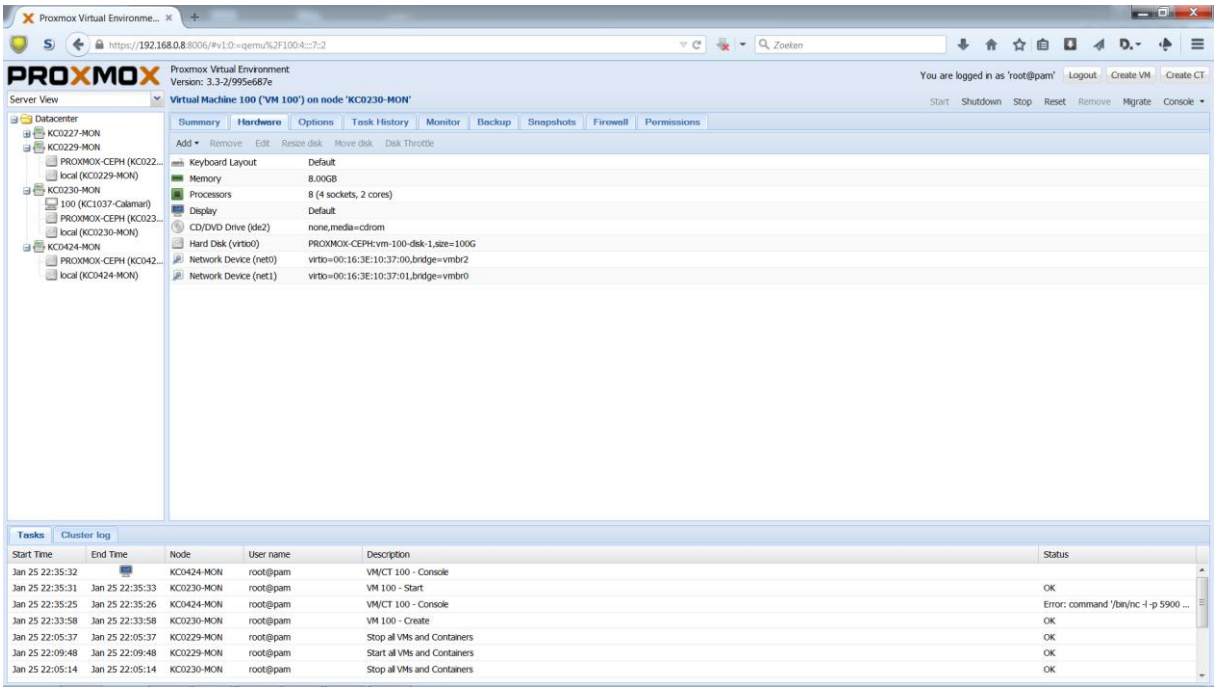
Het package is nu gereed en de volgende bestanden moeten in de bovengelegen directory zijn aangemaakt:
diamond_3.1.0.tar.gz
diamond_3.1.0_all.deb

Deze bestanden hebben we later nodig bij de installatie op de Ceph nodes, wanneer we deze gaan koppelen aan Calamari.

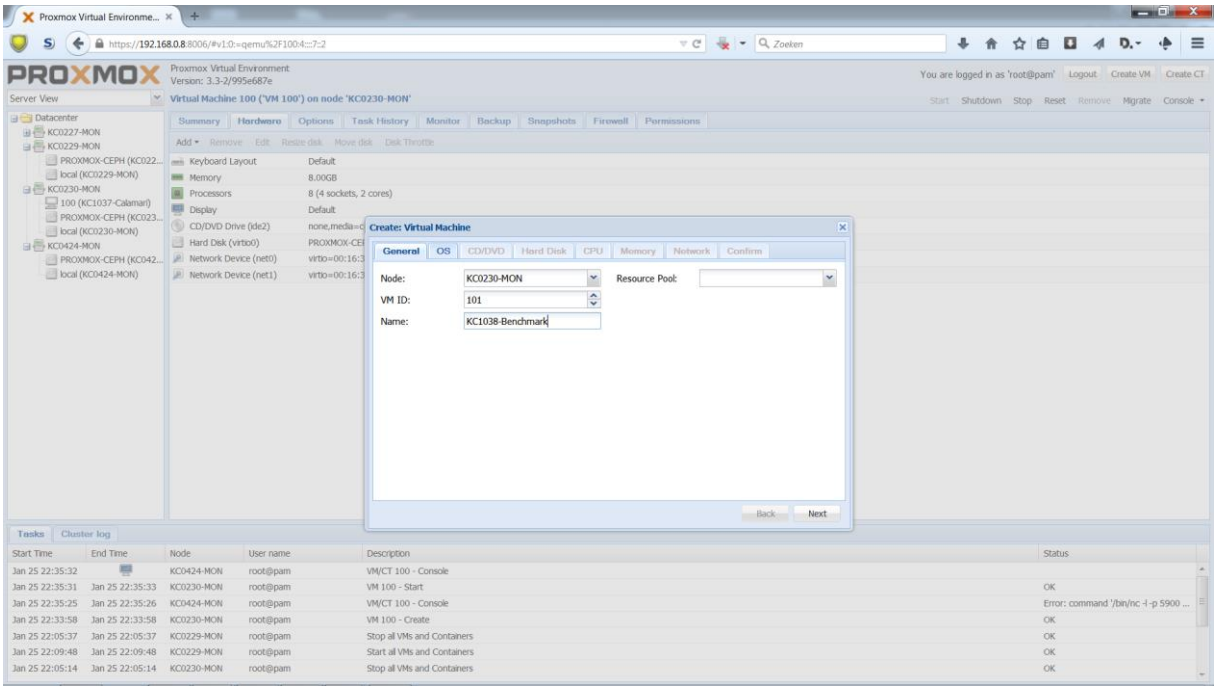
X. Bijlage installatie Calamari Server

```
aptitude install apache2

a2enmod version
service apache2 restart

vi /etc/apt/sources.list
deb http://kc1000.afstudeer.org/linux/calamari/ ./

wget -q -O - "http://debian.saltstack.com/debian-salt-team-joehealy.gpg.key" | apt-key add -
vi /etc/apt/sources.list.d/saltstack.list
deb http://kc1000.afstudeer.org/linux/debian.saltstack.com/debian/ wheezy-saltstack main

aptitude update

aptitude install postgresql libpq5 python-cairo supervisor libcairo2 libapache2-mod-wsgi apache2
python-zope.interface python-software-properties

aptitude install calamari-server calamari-clients


calamari-ctl initialize

vi /etc/network/interfaces
auto eth1
iface eth1 inet static
 address 172.16.83.11
 netmask 255.255.255.0

ifup eth1


#add-apt-repository ppa:saltstack/salt-testing
#aptitude install python-mako
```

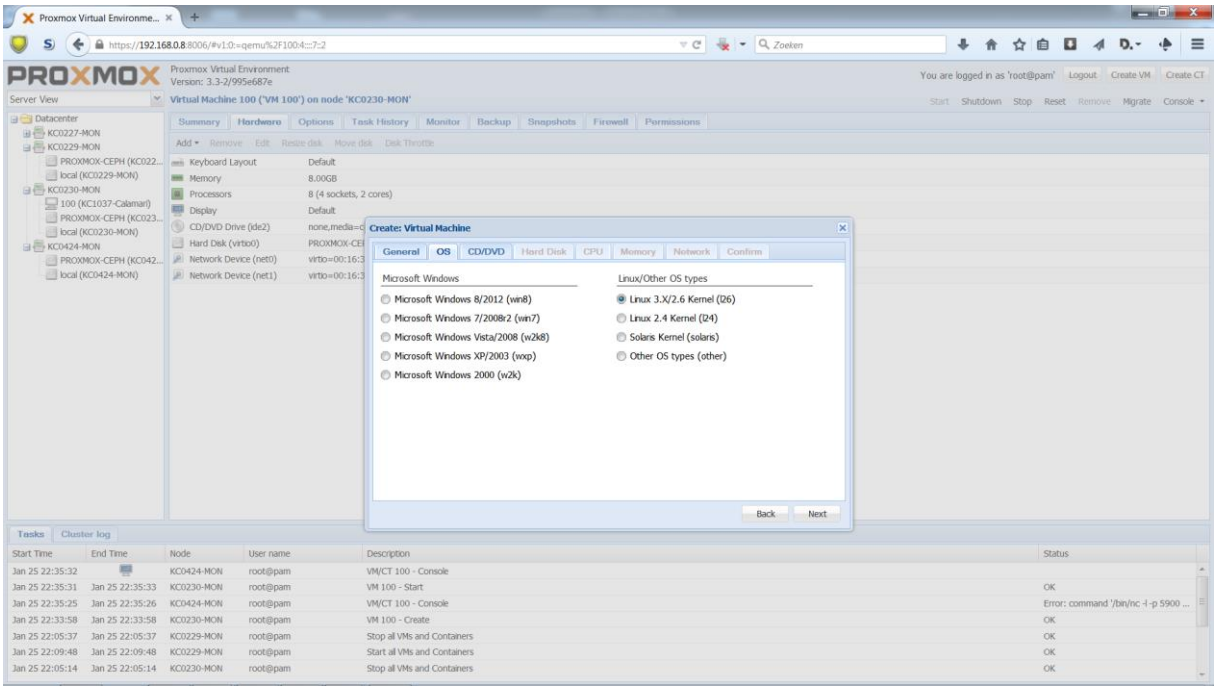
Y. Bijlage aanmaken virtuele host in Proxmox



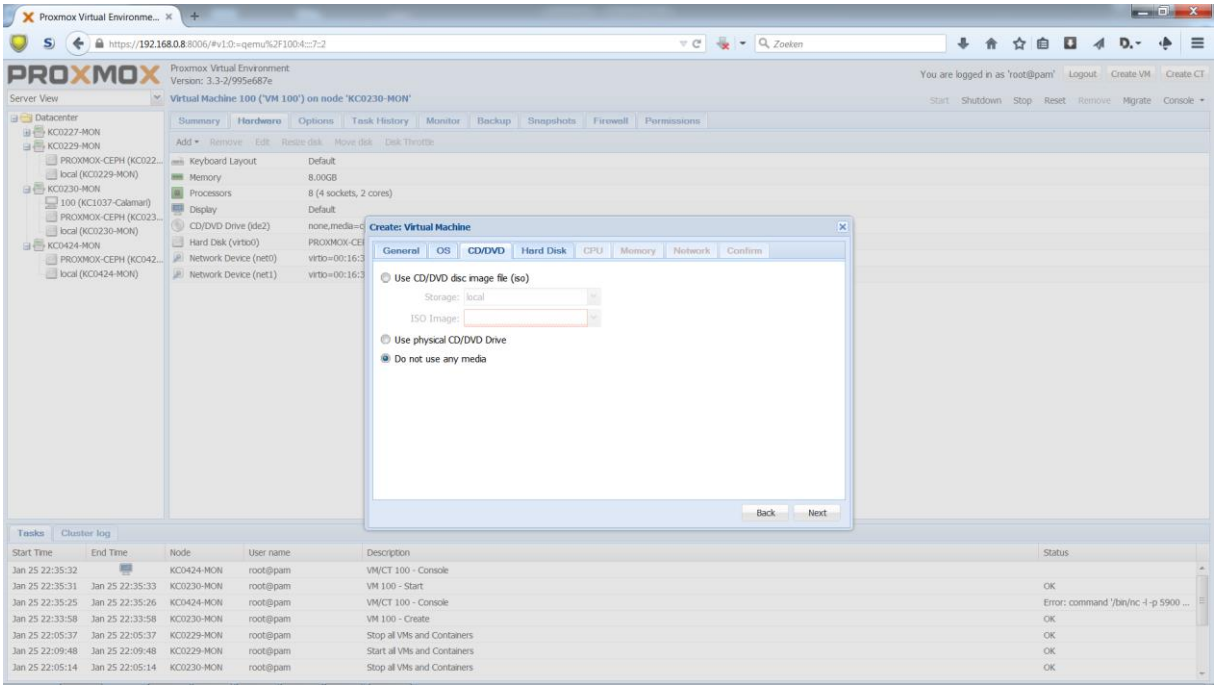
Figuur 78 Klik rechtsboven op "create VM"



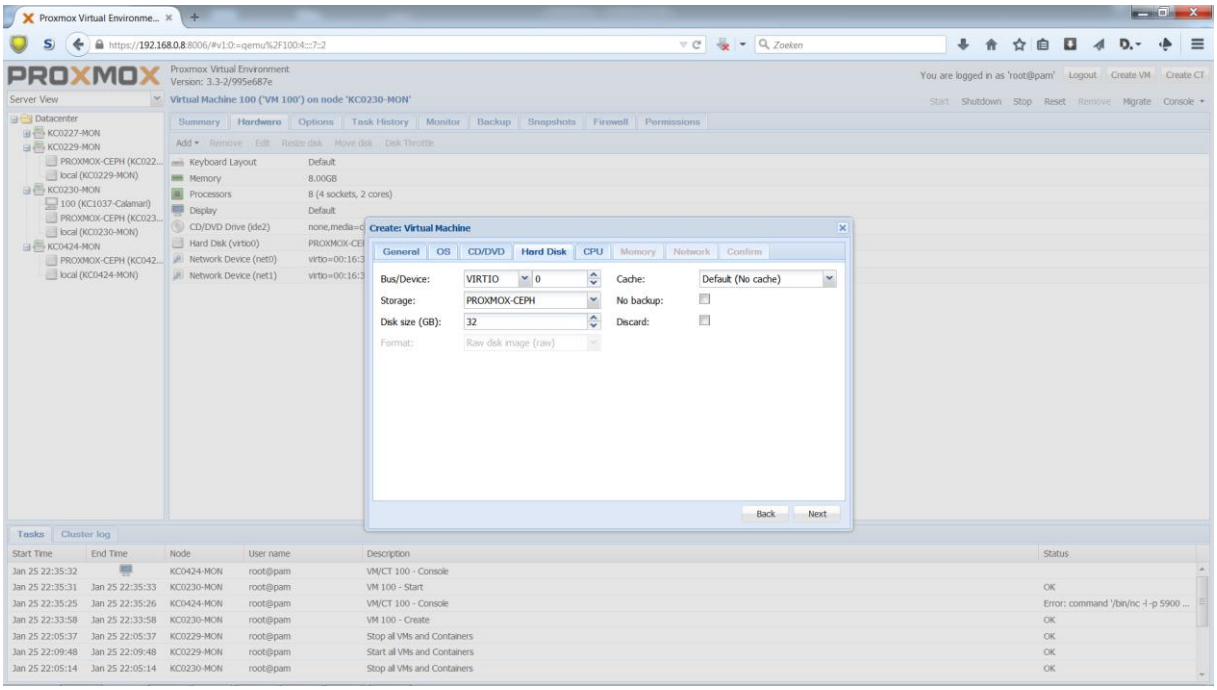
Figuur 79 Geef het systeem een naam



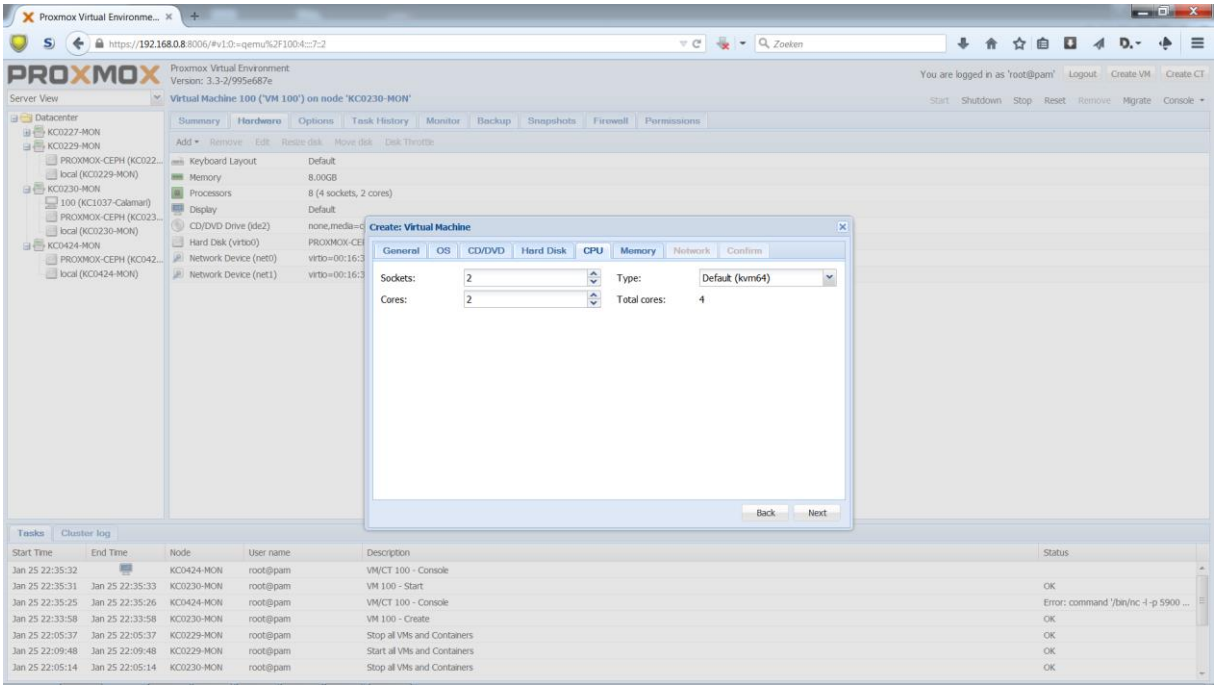
Figuur 80 Selecteer Linux 3.x/2.6 Kernel



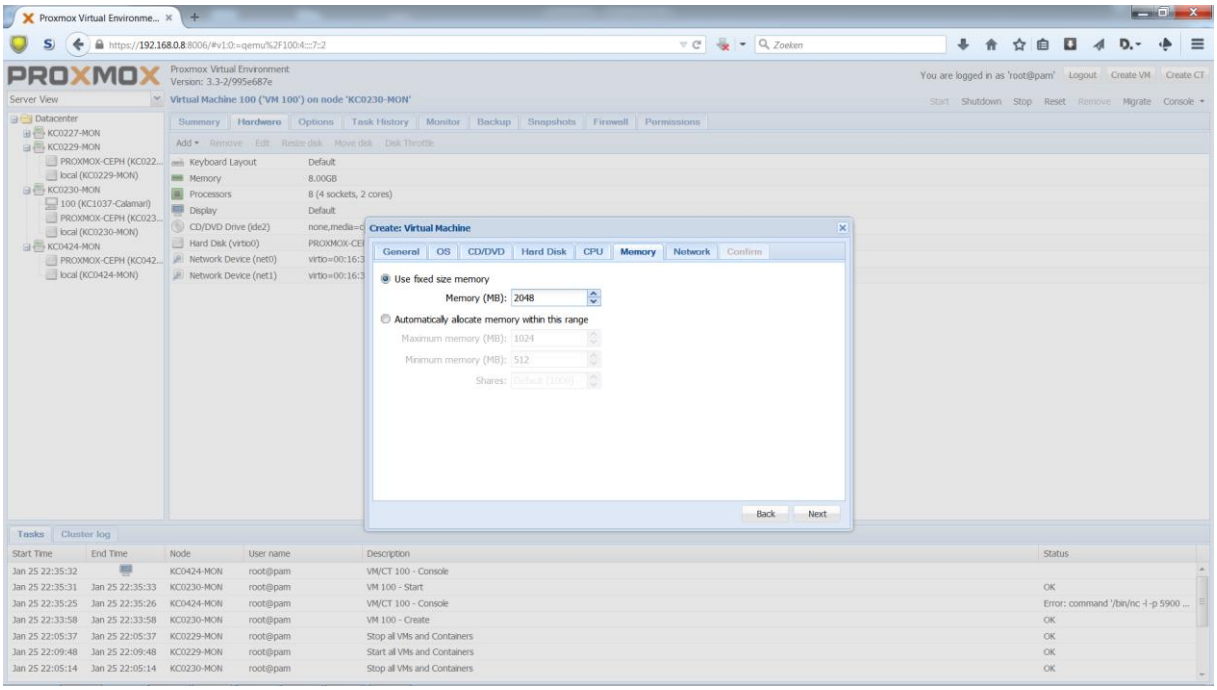
Figuur 81 Maak geen gebruik van CD/DVD tijdens de installatie



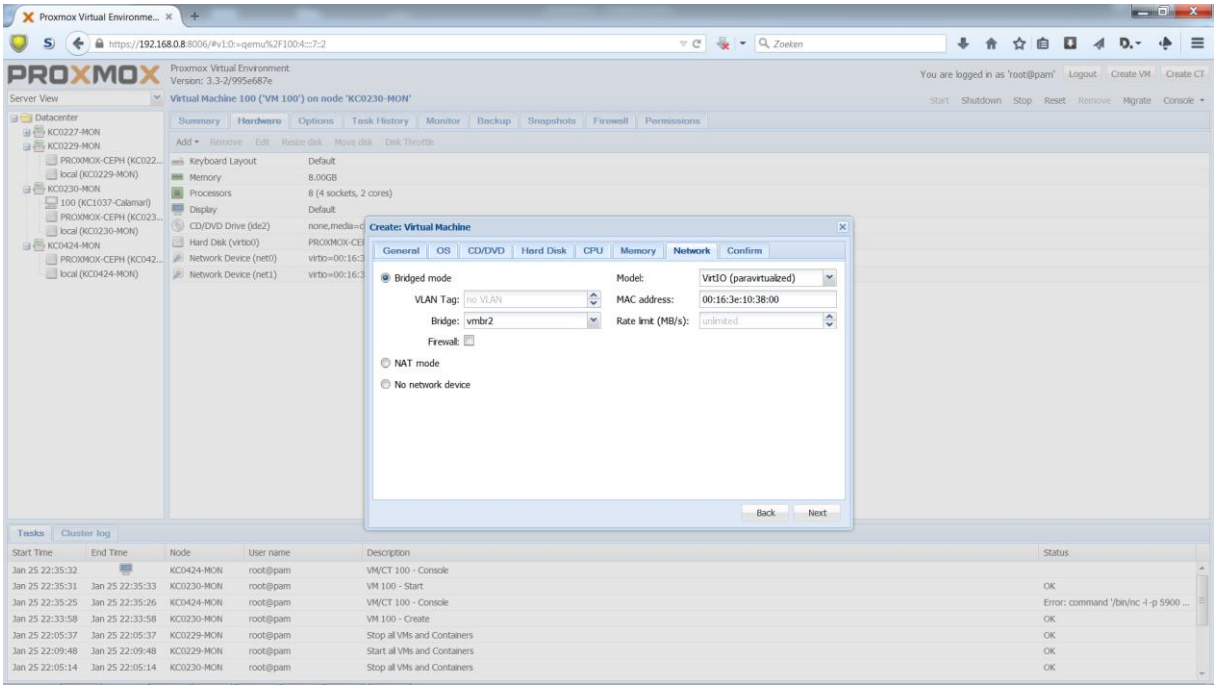
Figuur 82 Creëer een 32GB virtio harddisk aan op de Ceph storage



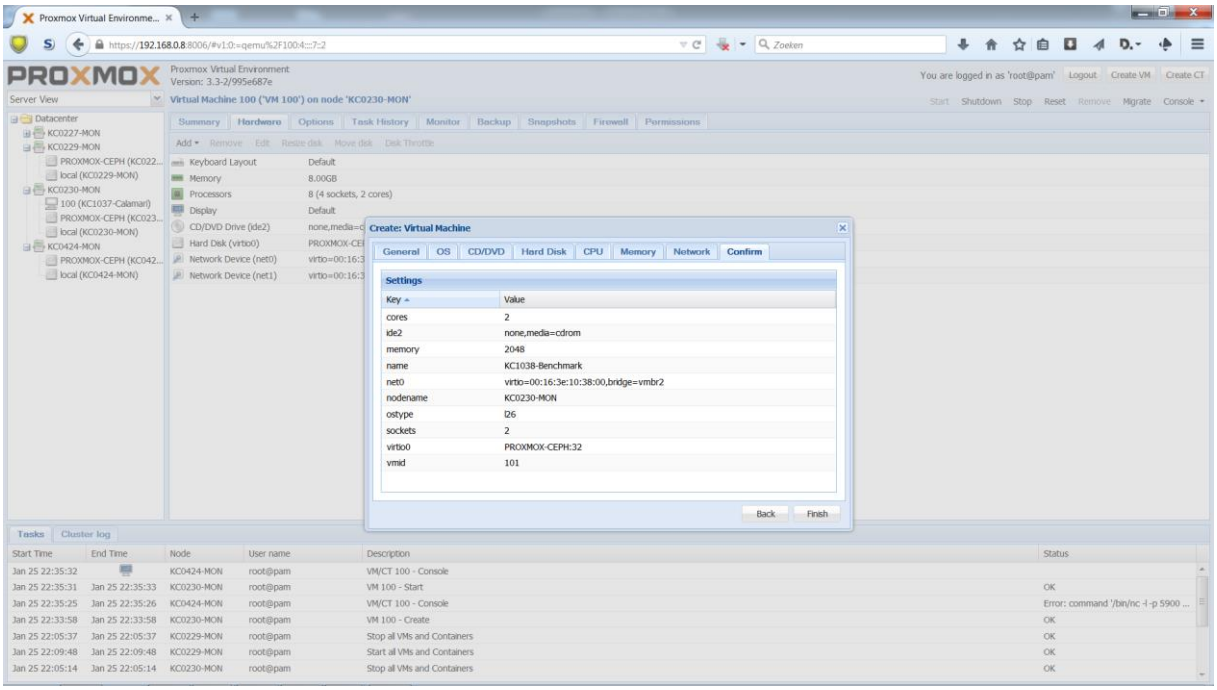
Figuur 83 Ken 2 sockets met 2 cores toe.



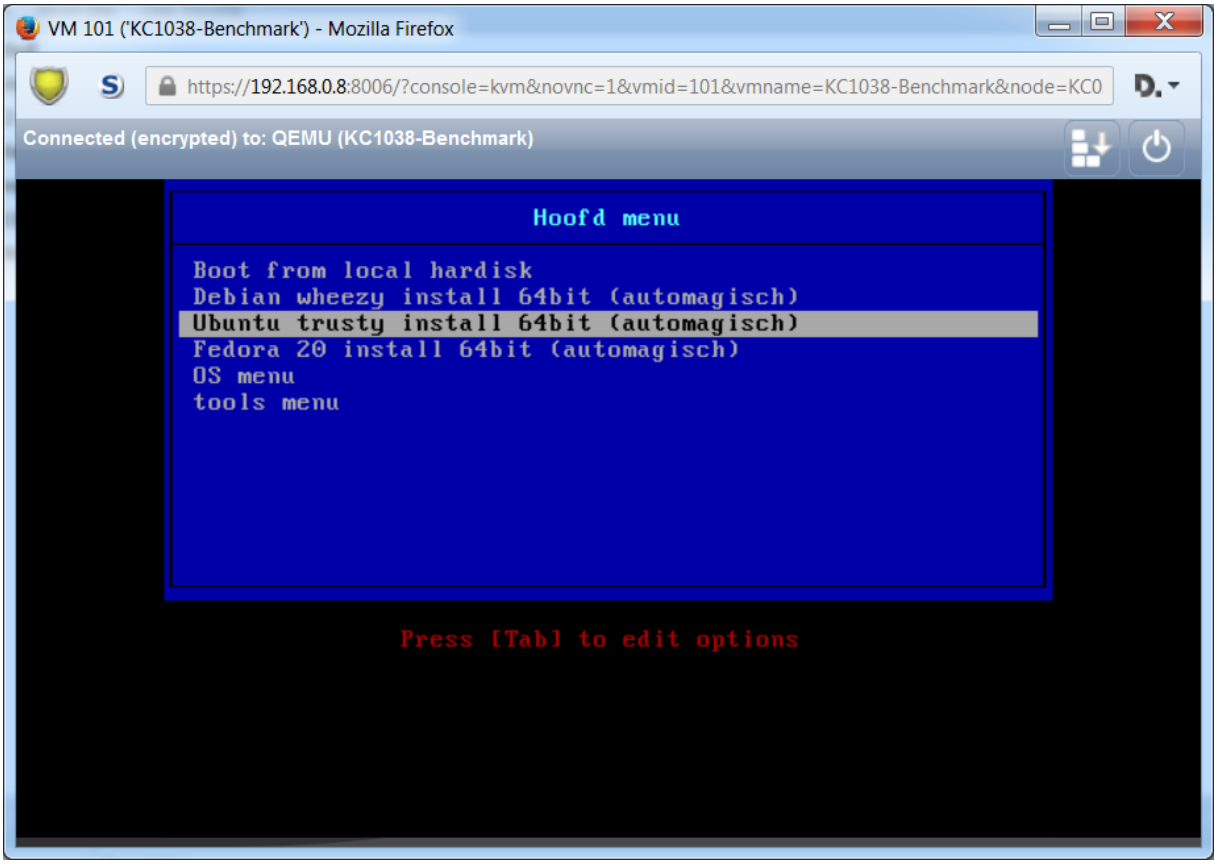
Figuur 84 Ken 2048MB geheugen toe.



Figuur 85 selecteer virtio, vmbr2 en voer het mac adres in



Figuur 86 bevestig de eerder ingevoerde waarden



Figuur 87 Boot PXE en selecteer Ubuntu trusty install 64Bit

Z. Bijlage OpenVAS resulaten

Summary

This document reports on the results of an automatic security scan. The report first summarises the results found. Then, for each host, the report describes every issue found. Please consider the advice given in each description, in order to rectify the issue.

Vendor security updates are not trusted.

Overrides are on. When a result has an override, this report uses the threat of the override.

Notes are included in the report.

This report might not show details of all issues that were found. It only lists hosts that produced issues. Issues with the threat level "Debug" are not shown. Issues with the threat level "False Positive" are not shown.

This report contains all 68 results selected by the filtering described above. Before filtering there were 68 results.

Scan started: **Mon Feb 9 13:30:08 2015**

Scan ended: Mon Feb 9 15:18:26 2015

Host Summary

Host	Start	End	High	Medium	Low	Log	False Positive
192.168.0.7	Feb 9, 13:30:13	Feb 9, 15:18:08	0	0	1	16	0
192.168.0.8	Feb 9, 13:30:13	Feb 9, 15:18:06	0	0	1	16	0
192.168.0.9	Feb 9, 13:30:13	Feb 9, 15:18:25	0	0	1	16	0
192.168.0.10	Feb 9, 13:30:13	Feb 9, 15:17:42	0	0	1	16	0
Total: 4			0	0	4	64	0

Results per Host

Host 192.168.0.7

Scanning of this host started at: 2015-02-09T13:30:13Z

Number of results: 17

Port Summary for Host 192.168.0.7

Service (Port)	Threat Level
general/tcp	Low
general/CPE-T	Log
3128/tcp	Log
25/tcp	Log
22/tcp	Log
123/udp	Log
111/tcp	Log

Security Issues for Host 192.168.0.7

general/tcp

Low (CVSS: 2.6)

NVT: TCP timestamps (OID: 1.3.6.1.4.1.25623.1.0.80091)

Summary

The remote host implements TCP timestamps and therefore allows to compute the uptime.

Vulnerability Detection Result

```
It was detected that the host implements RFC1323.
The following timestamps were retrieved with a delay of 1 seconds in-between:
Paket 1: 1274088670
Paket 2: 1274089863
```

Impact

A side effect of this feature is that the uptime of the remote host can sometimes be computed.

Solution

To disable TCP timestamps on linux add the line 'net.ipv4.tcp_timestamps = 0' to /etc/sysctl.conf. Execute 'sysctl -p' to apply the settings at runtime.

To disable TCP timestamps on Windows execute 'netsh int tcp set global timestamps=disabled'

Starting with Windows Server 2008 and Vista, the timestamp can not be completely disabled.

The default behavior of the TCP/IP stack on this Systems is, to not use the Timestamp options when initiating TCP connections, but use them if the TCP peer that is initiating communication includes them in their synchronize (SYN) segment.

See also: http://www.microsoft.com/en-us/download/details.aspx?id=9152

Vulnerability Insight

The remote host implements TCP timestamps, as defined by RFC1323.

Vulnerability Detection Method

Special IP packets are forged and sent with a little delay in between to the target IP. The responses are searched for a timestamps. If found, the timestamps are reported.

Details: TCP timestamps (OID: 1.3.6.1.4.1.25623.1.0.80091)

Version used: $Revision: 636 $

References

Other: http://www.ietf.org/rfc/rfc1323.txt

general/CPE-T
Log (CVSS: 0.0)
NVT: CPE Inventory (OID: 1.3.6.1.4.1.25623.1.0.810002)

Summary

This routine uses information collected by other routines about CPE identities (http://cpe.mitre.org/) of operating systems, services and applications detected during the scan.

Vulnerability Detection Result

```
192.168.0.7|cpe:/a:openbsd:openssh:6.0p1
192.168.0.7|cpe:/o:debian:debian_linux
```

Log Method

Details: CPE Inventory (OID: 1.3.6.1.4.1.25623.1.0.810002)

Version used: $Revision: 314 $

general/tcp
Log (CVSS: 7.8)
NVT: 3com switch2hub (OID: 1.3.6.1.4.1.25623.1.0.80103)

Summary

The remote host is subject to the switch to hub flood attack.

Description : The remote host on the local network seems to be connected through a switch which can be turned

into a hub when flooded by different mac addresses. The theory is to send a lot of packets (> 1000000) to the port of the switch we are connected to, with random mac addresses. This turns the switch into learning mode, where traffic goes everywhere. An attacker may use this flaw in the remote switch to sniff data going to this host

Reference : http://www.securitybugware.org/Other/2041.html

Vulnerability Detection Result

```
Fake IP address not specified. Skipping this check.
```

Solution

Lock Mac addresses on each port of the remote switch or buy newer switch.

Vulnerability Detection Method

Details: 3com switch2hub (OID: 1.3.6.1.4.1.25623.1.0.80103)

Version used: $Revision: 15 $

general/tcp
Log (CVSS: 0.0)
NVT: DIRB (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.103079)

Summary

This script uses DIRB to find directories and files on web applications via brute forcing.

Vulnerability Detection Result

```
DIRB could not be found in your system path.
OpenVAS was unable to execute DIRB and to perform the scan you
requested.
Please make sure that DIRB is installed and is
available in the PATH variable defined for your environment.
```

Log Method

Details: DIRB (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.103079)

Version used: $Revision: 13 $

general/tcp
Log (CVSS: 0.0)
NVT: arachni (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.110001)

Summary

This plugin uses arachni ruby command line to find web security issues.

See the preferences section for arachni options.

Note that OpenVAS is using limited set of arachni options. Therefore, for more complete web assessment, you should use standalone arachni tool for deeper/customized checks.

Vulnerability Detection Result

```
Arachni could not be found in your system path.
OpenVAS was unable to execute Arachni and to perform the scan you
requested.
Please make sure that Arachni is installed and that arachni is
available in the PATH variable defined for your environment.
```

Log Method

Details: arachni (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.110001)

Version used: $Revision: 683 $

general/tcp
Log (CVSS: 0.0)
NVT: Traceroute (OID: 1.3.6.1.4.1.25623.1.0.51662)

Summary

A traceroute from the scanning server to the target system was conducted. This traceroute is provided primarily

for informational value only. In the vast majority of cases, it does not represent a vulnerability. However, if the displayed traceroute contains any private addresses that should not have been publicly visible, then you have an issue you need to correct.

Vulnerability Detection Result

```
Here is the route from 172.16.41.10 to 192.168.0.7:
172.16.41.10
192.168.0.7
```

Solution

Block unwanted packets from escaping your network.

Log Method

Details: Traceroute (OID: 1.3.6.1.4.1.25623.1.0.51662)

Version used: $Revision: 14 $

22/tcp
Log (CVSS: 0.0)
NVT: SSH Protocol Versions Supported (OID: 1.3.6.1.4.1.25623.1.0.100259)

Summary

Identification of SSH protocol versions supported by the remote SSH Server. Also reads the corresponding fingerprints from the service.

The following versions are tried: 1.33, 1.5, 1.99 and 2.0

Vulnerability Detection Result

```
The remote SSH Server supports the following SSH Protocol Versions:
1.99
2.0
SSHv2 Fingerprint: c5:9c:e5:cf:5e:85:44:6b:2d:5c:9c:68:71:31:d7:b6
```

Log Method

Details: SSH Protocol Versions Supported (OID: 1.3.6.1.4.1.25623.1.0.100259)

Version used: $Revision: 43 $

22/tcp
Log (CVSS: 0.0)
NVT: SSH Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10267)

Summary

This detects the SSH Server's type and version by connecting to the server and processing the buffer received. This information gives potential attackers additional information about the system they are attacking. Versions and Types should be omitted where possible.

Vulnerability Detection Result

```
Detected SSH server version: SSH-2.0-OpenSSH_6.0p1 Debian-4+deb7u2
Remote SSH supported authentication: publickey,password
Remote SSH banner:
(not available)
CPE: cpe:/a:openbsd:openssh:6.0p1
Concluded from remote connection attempt with credentials:
  Login: OpenVAS
  Password: OpenVAS
```

Solution

Apply filtering to disallow access to this port from untrusted hosts

Log Method

Details: SSH Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10267)

Version used: $Revision: 588 $

22/tcp
Log (CVSS: 0.0)

NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
An ssh server is running on this port
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

25/tcp

Log (CVSS: 0.0)

NVT: SMTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10263)

Summary

This detects the SMTP Server's type and version by connecting to the server and processing the buffer received. This information gives potential attackers additional information about the system they are attacking. Versions and Types should be omitted where possible.

Vulnerability Detection Result

```
Remote SMTP server banner :
220 ESMTP
```

Solution

Change the login banner to something generic.

Log Method

Details: SMTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10263)

Version used: $Revision: 339 $

25/tcp

Log (CVSS: 0.0)

NVT: SMTP STARTTLS Detection (OID: 1.3.6.1.4.1.25623.1.0.103118)

Summary

Check if the remote Mailserver supports the STARTTLS command.

Vulnerability Detection Result

```
The remote Mailserver supports the STARTTLS command.
```

Log Method

Details: SMTP STARTTLS Detection (OID: 1.3.6.1.4.1.25623.1.0.103118)

Version used: $Revision: 703 $

25/tcp

Log (CVSS: 0.0)

NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
An SMTP server is running on this port
Here is its banner :
220 ESMTP
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

111/tcp
Log (CVSS: 0.0)
NVT: rpcinfo -p (OID: 1.3.6.1.4.1.25623.1.0.11111)

Summary

This script calls the DUMP RPC on the port mapper, to obtain the list of all registered programs.

Vulnerability Detection Result

```
These are the registered RPC programs:\
\
RPC program #100000 version 4 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100000 version 3 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100000 version 2 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100024 version 1 'status' on port 35302/TCP
RPC program #100000 version 4 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100000 version 3 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100000 version 2 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100024 version 1 'status' on port 58270/UDP
```

Log Method

Details: rpcinfo -p (OID: 1.3.6.1.4.1.25623.1.0.11111)

Version used: $Revision: 41 $

123/udp
Log (CVSS: 0.0)
NVT: NTP read variables (OID: 1.3.6.1.4.1.25623.1.0.10884)

Summary

A NTP (Network Time Protocol) server is listening on this port.

Vulnerability Detection Result

Vulnerability was detected according to the Vulnerability Detection Method.

Log Method

Details: NTP read variables (OID: 1.3.6.1.4.1.25623.1.0.10884)

Version used: $Revision: 487 $

3128/tcp
Log (CVSS: 0.0)
NVT: HTTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10107)

Summary

This detects the HTTP Server's type and version.

Vulnerability Detection Result

```
The remote web server type is :
pve-api-daemon/3.0
```

Solution

Configure your server to use an alternate name like 'Wintendo httpD w/Dotmatrix display' Be sure to remove common logos like apache_pb.gif. With Apache, you can set the directive 'ServerTokens Prod' to limit the information emanating from the server in its response headers.

Log Method

Details: HTTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10107)

Version used: $Revision: 229 $

3128/tcp
Log (CVSS: 0.0)
NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
A web server is running on this port
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

3128/tcp
Log (CVSS: 0.0)
NVT: wapiti (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.80110)

Summary

This plugin uses wapiti to find web security issues.

Make sure to have wapiti 2.x as wapiti 1.x is not supported.

See the preferences section for wapiti options.

Note that OpenVAS is using limited set of wapiti options. Therefore, for more complete web assessment, you should use standalone wapiti tool for deeper/customized checks.

Vulnerability Detection Result

```
wapiti could not be found in your system path.
OpenVAS was unable to execute wapiti and to perform the scan you
requested.
Please make sure that wapiti is installed and that wapiti is
available in the PATH variable defined for your environment.
```

Log Method

Details: wapiti (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.80110)

Version used: $Revision: 14 $

Host 192.168.0.8

Scanning of this host started at:	2015-02-09T13:30:13Z
Number of results:	17

Port Summary for Host 192.168.0.8

Service (Port)	Threat Level
general/tcp	Low
general/CPE-T	Log
3128/tcp	Log
25/tcp	Log
22/tcp	Log
123/udp	Log
111/tcp	Log

Security Issues for Host 192.168.0.8

general/tcp
Low (CVSS: 2.6)
NVT: TCP timestamps (OID: 1.3.6.1.4.1.25623.1.0.80091)

Summary

The remote host implements TCP timestamps and therefore allows to compute the uptime.

Vulnerability Detection Result

```
It was detected that the host implements RFC1323.
The following timestamps were retrieved with a delay of 1 seconds in-between:
Paket 1: 1274091301
Paket 2: 1274092525
```

Impact

A side effect of this feature is that the uptime of the remote host can sometimes be computed.

Solution

To disable TCP timestamps on linux add the line 'net.ipv4.tcp_timestamps = 0' to /etc/sysctl.conf. Execute 'sysctl -p' to apply the settings at runtime.

To disable TCP timestamps on Windows execute 'netsh int tcp set global timestamps=disabled'

Starting with Windows Server 2008 and Vista, the timestamp can not be completely disabled.

The default behavior of the TCP/IP stack on this Systems is, to not use the Timestamp options when initiating TCP connections, but use them if the TCP peer that is initiating communication includes them in their synchronize (SYN) segment.

See also: http://www.microsoft.com/en-us/download/details.aspx?id=9152

Vulnerability Insight

The remote host implements TCP timestamps, as defined by RFC1323.

Vulnerability Detection Method

Special IP packets are forged and sent with a little delay in between to the target IP. The responses are searched for a timestamps. If found, the timestamps are reported.

Details: TCP timestamps (OID: 1.3.6.1.4.1.25623.1.0.80091)

Version used: $Revision: 636 $

References

Other: http://www.ietf.org/rfc/rfc1323.txt

general/CPE-T
Log (CVSS: 0.0)
NVT: CPE Inventory (OID: 1.3.6.1.4.1.25623.1.0.810002)

Summary

This routine uses information collected by other routines about CPE identities (http://cpe.mitre.org/) of operating systems, services and applications detected during the scan.

Vulnerability Detection Result

```
192.168.0.8|cpe:/a:openbsd:openssh:6.0p1
192.168.0.8|cpe:/o:debian:debian_linux
```

Log Method

Details: CPE Inventory (OID: 1.3.6.1.4.1.25623.1.0.810002)

Version used: $Revision: 314 $

general/tcp
Log (CVSS: 7.8)
NVT: 3com switch2hub (OID: 1.3.6.1.4.1.25623.1.0.80103)

Summary

The remote host is subject to the switch to hub flood attack.

Description : The remote host on the local network seems to be connected through a switch which can be turned

into a hub when flooded by different mac addresses. The theory is to send a lot of packets (> 1000000) to the port of the switch we are connected to, with random mac addresses. This turns the switch into learning mode, where traffic goes everywhere. An attacker may use this flaw in the remote switch to sniff data going to this host

Reference : http://www.securitybugware.org/Other/2041.html

Vulnerability Detection Result

```
Fake IP address not specified. Skipping this check.
```

Solution

Lock Mac addresses on each port of the remote switch or buy newer switch.

Vulnerability Detection Method

Details: 3com switch2hub (OID: 1.3.6.1.4.1.25623.1.0.80103)

Version used: $Revision: 15 $

general/tcp
Log (CVSS: 0.0)
NVT: DIRB (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.103079)

Summary

This script uses DIRB to find directories and files on web applications via brute forcing.

Vulnerability Detection Result

```
DIRB could not be found in your system path.
OpenVAS was unable to execute DIRB and to perform the scan you
requested.
Please make sure that DIRB is installed and is
available in the PATH variable defined for your environment.
```

Log Method

Details: DIRB (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.103079)

Version used: $Revision: 13 $

general/tcp
Log (CVSS: 0.0)
NVT: arachni (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.110001)

Summary

This plugin uses arachni ruby command line to find web security issues.

See the preferences section for arachni options.

Note that OpenVAS is using limited set of arachni options. Therefore, for more complete web assessment, you should use standalone arachni tool for deeper/customized checks.

Vulnerability Detection Result

```
Arachni could not be found in your system path.
OpenVAS was unable to execute Arachni and to perform the scan you
requested.
Please make sure that Arachni is installed and that arachni is
available in the PATH variable defined for your environment.
```

Log Method

Details: arachni (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.110001)

Version used: $Revision: 683 $

general/tcp
Log (CVSS: 0.0)
NVT: Traceroute (OID: 1.3.6.1.4.1.25623.1.0.51662)

Summary

A traceroute from the scanning server to the target system was conducted. This traceroute is provided primarily

for informational value only. In the vast majority of cases, it does not represent a vulnerability. However, if the displayed traceroute contains any private addresses that should not have been publicly visible, then you have an issue you need to correct.

Vulnerability Detection Result

```
Here is the route from 172.16.41.10 to 192.168.0.8:
172.16.41.10
192.168.0.8
```

Solution

Block unwanted packets from escaping your network.

Log Method

Details: Traceroute (OID: 1.3.6.1.4.1.25623.1.0.51662)

Version used: $Revision: 14 $

22/tcp
Log (CVSS: 0.0)
NVT: SSH Protocol Versions Supported (OID: 1.3.6.1.4.1.25623.1.0.100259)

Summary

Identification of SSH protocol versions supported by the remote SSH Server. Also reads the corresponding fingerprints from the service.

The following versions are tried: 1.33, 1.5, 1.99 and 2.0

Vulnerability Detection Result

```
The remote SSH Server supports the following SSH Protocol Versions:
1.99
2.0
SSHv2 Fingerprint: 19:48:b0:0f:a2:21:22:88:83:05:57:55:3c:2b:d3:74
```

Log Method

Details: SSH Protocol Versions Supported (OID: 1.3.6.1.4.1.25623.1.0.100259)

Version used: $Revision: 43 $

22/tcp
Log (CVSS: 0.0)
NVT: SSH Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10267)

Summary

This detects the SSH Server's type and version by connecting to the server and processing the buffer received. This information gives potential attackers additional information about the system they are attacking. Versions and Types should be omitted where possible.

Vulnerability Detection Result

```
Detected SSH server version: SSH-2.0-OpenSSH_6.0p1 Debian-4+deb7u2
Remote SSH supported authentication: publickey,password
Remote SSH banner:
(not available)
CPE: cpe:/a:openbsd:openssh:6.0p1
Concluded from remote connection attempt with credentials:
  Login: OpenVAS
  Password: OpenVAS
```

Solution

Apply filtering to disallow access to this port from untrusted hosts

Log Method

Details: SSH Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10267)

Version used: $Revision: 588 $

22/tcp
Log (CVSS: 0.0)

NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
An ssh server is running on this port
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

25/tcp

Log (CVSS: 0.0)

NVT: SMTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10263)

Summary

This detects the SMTP Server's type and version by connecting to the server and processing the buffer received. This information gives potential attackers additional information about the system they are attacking. Versions and Types should be omitted where possible.

Vulnerability Detection Result

```
Remote SMTP server banner :
220 ESMTP
```

Solution

Change the login banner to something generic.

Log Method

Details: SMTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10263)

Version used: $Revision: 339 $

25/tcp

Log (CVSS: 0.0)

NVT: SMTP STARTTLS Detection (OID: 1.3.6.1.4.1.25623.1.0.103118)

Summary

Check if the remote Mailserver supports the STARTTLS command.

Vulnerability Detection Result

```
The remote Mailserver supports the STARTTLS command.
```

Log Method

Details: SMTP STARTTLS Detection (OID: 1.3.6.1.4.1.25623.1.0.103118)

Version used: $Revision: 703 $

25/tcp

Log (CVSS: 0.0)

NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
An SMTP server is running on this port
Here is its banner :
220 ESMTP
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

111/tcp
Log (CVSS: 0.0)
NVT: rpcinfo -p (OID: 1.3.6.1.4.1.25623.1.0.11111)

Summary

This script calls the DUMP RPC on the port mapper, to obtain the list of all registered programs.

Vulnerability Detection Result

```
These are the registered RPC programs:\
\
RPC program #100000 version 4 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100000 version 3 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100000 version 2 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100024 version 1 'status' on port 47911/TCP
RPC program #100000 version 4 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100000 version 3 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100000 version 2 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100024 version 1 'status' on port 58536/UDP
```

Log Method

Details: rpcinfo -p (OID: 1.3.6.1.4.1.25623.1.0.11111)

Version used: $Revision: 41 $

123/udp
Log (CVSS: 0.0)
NVT: NTP read variables (OID: 1.3.6.1.4.1.25623.1.0.10884)

Summary

A NTP (Network Time Protocol) server is listening on this port.

Vulnerability Detection Result

Vulnerability was detected according to the Vulnerability Detection Method.

Log Method

Details: NTP read variables (OID: 1.3.6.1.4.1.25623.1.0.10884)

Version used: $Revision: 487 $

3128/tcp
Log (CVSS: 0.0)
NVT: HTTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10107)

Summary

This detects the HTTP Server's type and version.

Vulnerability Detection Result

```
The remote web server type is :
pve-api-daemon/3.0
```

Solution

Configure your server to use an alternate name like 'Wintendo httpD w/Dotmatrix display' Be sure to remove common logos like apache_pb.gif. With Apache, you can set the directive 'ServerTokens Prod' to limit the information emanating from the server in its response headers.

Log Method

Details: HTTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10107)

Version used: $Revision: 229 $

3128/tcp
Log (CVSS: 0.0)
NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
A web server is running on this port
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

3128/tcp
Log (CVSS: 0.0)
NVT: wapiti (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.80110)

Summary

This plugin uses wapiti to find web security issues.

Make sure to have wapiti 2.x as wapiti 1.x is not supported.

See the preferences section for wapiti options.

Note that OpenVAS is using limited set of wapiti options. Therefore, for more complete web assessment, you should use standalone wapiti tool for deeper/customized checks.

Vulnerability Detection Result

```
wapiti could not be found in your system path.
OpenVAS was unable to execute wapiti and to perform the scan you
requested.
Please make sure that wapiti is installed and that wapiti is
available in the PATH variable defined for your environment.
```

Log Method

Details: wapiti (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.80110)

Version used: $Revision: 14 $

Host 192.168.0.9

Scanning of this host started at:	2015-02-09T13:30:13Z
Number of results:	17

Port Summary for Host 192.168.0.9

Service (Port)	Threat Level
general/tcp	Low
general/CPE-T	Log
3128/tcp	Log
25/tcp	Log
22/tcp	Log
123/udp	Log
111/tcp	Log

Security Issues for Host 192.168.0.9

general/tcp
Low (CVSS: 2.6)
NVT: TCP timestamps (OID: 1.3.6.1.4.1.25623.1.0.80091)

Summary

The remote host implements TCP timestamps and therefore allows to compute the uptime.

Vulnerability Detection Result

```
It was detected that the host implements RFC1323.
The following timestamps were retrieved with a delay of 1 seconds in-between:
Paket 1: 1274054063
Paket 2: 1274055298
```

Impact

A side effect of this feature is that the uptime of the remote host can sometimes be computed.

Solution

To disable TCP timestamps on linux add the line 'net.ipv4.tcp_timestamps = 0' to /etc/sysctl.conf. Execute 'sysctl -p' to apply the settings at runtime.

To disable TCP timestamps on Windows execute 'netsh int tcp set global timestamps=disabled'

Starting with Windows Server 2008 and Vista, the timestamp can not be completely disabled.

The default behavior of the TCP/IP stack on this Systems is, to not use the Timestamp options when initiating TCP connections, but use them if the TCP peer that is initiating communication includes them in their synchronize (SYN) segment.

See also: http://www.microsoft.com/en-us/download/details.aspx?id=9152

Vulnerability Insight

The remote host implements TCP timestamps, as defined by RFC1323.

Vulnerability Detection Method

Special IP packets are forged and sent with a little delay in between to the target IP. The responses are searched for a timestamps. If found, the timestamps are reported.

Details: TCP timestamps (OID: 1.3.6.1.4.1.25623.1.0.80091)

Version used: $Revision: 636 $

References

Other: http://www.ietf.org/rfc/rfc1323.txt

general/CPE-T
Log (CVSS: 0.0)
NVT: CPE Inventory (OID: 1.3.6.1.4.1.25623.1.0.810002)

Summary

This routine uses information collected by other routines about CPE identities (http://cpe.mitre.org/) of operating systems, services and applications detected during the scan.

Vulnerability Detection Result

```
192.168.0.9|cpe:/a:openbsd:openssh:6.0p1
192.168.0.9|cpe:/o:debian:debian_linux
```

Log Method

Details: CPE Inventory (OID: 1.3.6.1.4.1.25623.1.0.810002)

Version used: $Revision: 314 $

general/tcp
Log (CVSS: 7.8)
NVT: 3com switch2hub (OID: 1.3.6.1.4.1.25623.1.0.80103)

Summary

The remote host is subject to the switch to hub flood attack.

Description : The remote host on the local network seems to be connected through a switch which can be turned

into a hub when flooded by different mac addresses. The theory is to send a lot of packets (> 1000000) to the port of the switch we are connected to, with random mac addresses. This turns the switch into learning mode, where traffic goes everywhere. An attacker may use this flaw in the remote switch to sniff data going to this host

Reference : http://www.securitybugware.org/Other/2041.html

Vulnerability Detection Result

```
Fake IP address not specified. Skipping this check.
```

Solution

Lock Mac addresses on each port of the remote switch or buy newer switch.

Vulnerability Detection Method

Details: 3com switch2hub (OID: 1.3.6.1.4.1.25623.1.0.80103)

Version used: $Revision: 15 $

general/tcp
Log (CVSS: 0.0)
NVT: DIRB (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.103079)

Summary

This script uses DIRB to find directories and files on web applications via brute forcing.

Vulnerability Detection Result

```
DIRB could not be found in your system path.
OpenVAS was unable to execute DIRB and to perform the scan you
requested.
Please make sure that DIRB is installed and is
available in the PATH variable defined for your environment.
```

Log Method

Details: DIRB (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.103079)

Version used: $Revision: 13 $

general/tcp
Log (CVSS: 0.0)
NVT: arachni (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.110001)

Summary

This plugin uses arachni ruby command line to find web security issues.

See the preferences section for arachni options.

Note that OpenVAS is using limited set of arachni options. Therefore, for more complete web assessment, you should use standalone arachni tool for deeper/customized checks.

Vulnerability Detection Result

```
Arachni could not be found in your system path.
OpenVAS was unable to execute Arachni and to perform the scan you
requested.
Please make sure that Arachni is installed and that arachni is
available in the PATH variable defined for your environment.
```

Log Method

Details: arachni (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.110001)

Version used: $Revision: 683 $

general/tcp
Log (CVSS: 0.0)
NVT: Traceroute (OID: 1.3.6.1.4.1.25623.1.0.51662)

Summary

A traceroute from the scanning server to the target system was conducted. This traceroute is provided primarily

for informational value only. In the vast majority of cases, it does not represent a vulnerability. However, if the displayed traceroute contains any private addresses that should not have been publicly visible, then you have an issue you need to correct.

Vulnerability Detection Result

```
Here is the route from 172.16.41.10 to 192.168.0.9:
172.16.41.10
192.168.0.9
```

Solution

Block unwanted packets from escaping your network.

Log Method

Details: Traceroute (OID: 1.3.6.1.4.1.25623.1.0.51662)

Version used: $Revision: 14 $

22/tcp
Log (CVSS: 0.0)
NVT: SSH Protocol Versions Supported (OID: 1.3.6.1.4.1.25623.1.0.100259)

Summary

Identification of SSH protocol versions supported by the remote SSH Server. Also reads the corresponding fingerprints from the service.

The following versions are tried: 1.33, 1.5, 1.99 and 2.0

Vulnerability Detection Result

```
The remote SSH Server supports the following SSH Protocol Versions:
1.99
2.0
SSHv2 Fingerprint: f9:ee:79:ce:c6:1a:1e:4f:61:91:20:77:7c:18:19:64
```

Log Method

Details: SSH Protocol Versions Supported (OID: 1.3.6.1.4.1.25623.1.0.100259)

Version used: $Revision: 43 $

22/tcp
Log (CVSS: 0.0)
NVT: SSH Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10267)

Summary

This detects the SSH Server's type and version by connecting to the server and processing the buffer received. This information gives potential attackers additional information about the system they are attacking. Versions and Types should be omitted where possible.

Vulnerability Detection Result

```
Detected SSH server version: SSH-2.0-OpenSSH_6.0p1 Debian-4+deb7u2
Remote SSH supported authentication: publickey,password
Remote SSH banner:
(not available)
CPE: cpe:/a:openbsd:openssh:6.0p1
Concluded from remote connection attempt with credentials:
  Login: OpenVAS
  Password: OpenVAS
```

Solution

Apply filtering to disallow access to this port from untrusted hosts

Log Method

Details: SSH Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10267)

Version used: $Revision: 588 $

22/tcp
Log (CVSS: 0.0)

NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
An ssh server is running on this port
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

25/tcp

Log (CVSS: 0.0)

NVT: SMTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10263)

Summary

This detects the SMTP Server's type and version by connecting to the server and processing the buffer received. This information gives potential attackers additional information about the system they are attacking. Versions and Types should be omitted where possible.

Vulnerability Detection Result

```
Remote SMTP server banner :
220 ESMTP
```

Solution

Change the login banner to something generic.

Log Method

Details: SMTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10263)

Version used: $Revision: 339 $

25/tcp

Log (CVSS: 0.0)

NVT: SMTP STARTTLS Detection (OID: 1.3.6.1.4.1.25623.1.0.103118)

Summary

Check if the remote Mailserver supports the STARTTLS command.

Vulnerability Detection Result

```
The remote Mailserver supports the STARTTLS command.
```

Log Method

Details: SMTP STARTTLS Detection (OID: 1.3.6.1.4.1.25623.1.0.103118)

Version used: $Revision: 703 $

25/tcp

Log (CVSS: 0.0)

NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
An SMTP server is running on this port
Here is its banner :
220 ESMTP
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

111/tcp
Log (CVSS: 0.0)
NVT: rpcinfo -p (OID: 1.3.6.1.4.1.25623.1.0.11111)

Summary

This script calls the DUMP RPC on the port mapper, to obtain the list of all registered programs.

Vulnerability Detection Result

```
These are the registered RPC programs:\
\
RPC program #100000 version 4 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100000 version 3 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100000 version 2 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100024 version 1 'status' on port 35594/TCP
RPC program #100000 version 4 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100000 version 3 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100000 version 2 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100024 version 1 'status' on port 44354/UDP
```

Log Method

Details: rpcinfo -p (OID: 1.3.6.1.4.1.25623.1.0.11111)

Version used: $Revision: 41 $

123/udp
Log (CVSS: 0.0)
NVT: NTP read variables (OID: 1.3.6.1.4.1.25623.1.0.10884)

Summary

A NTP (Network Time Protocol) server is listening on this port.

Vulnerability Detection Result

Vulnerability was detected according to the Vulnerability Detection Method.

Log Method

Details: NTP read variables (OID: 1.3.6.1.4.1.25623.1.0.10884)

Version used: $Revision: 487 $

3128/tcp
Log (CVSS: 0.0)
NVT: HTTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10107)

Summary

This detects the HTTP Server's type and version.

Vulnerability Detection Result

```
The remote web server type is :
pve-api-daemon/3.0
```

Solution

Configure your server to use an alternate name like 'Wintendo httpD w/Dotmatrix display' Be sure to remove common logos like apache_pb.gif. With Apache, you can set the directive 'ServerTokens Prod' to limit the information emanating from the server in its response headers.

Log Method

Details: HTTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10107)

Version used: $Revision: 229 $

3128/tcp
Log (CVSS: 0.0)
NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
A web server is running on this port
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

3128/tcp
Log (CVSS: 0.0)
NVT: wapiti (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.80110)

Summary

This plugin uses wapiti to find web security issues.

Make sure to have wapiti 2.x as wapiti 1.x is not supported.

See the preferences section for wapiti options.

Note that OpenVAS is using limited set of wapiti options. Therefore, for more complete web assessment, you should use standalone wapiti tool for deeper/customized checks.

Vulnerability Detection Result

```
wapiti could not be found in your system path.
OpenVAS was unable to execute wapiti and to perform the scan you
requested.
Please make sure that wapiti is installed and that wapiti is
available in the PATH variable defined for your environment.
```

Log Method

Details: wapiti (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.80110)

Version used: $Revision: 14 $

Host 192.168.0.10

Scanning of this host started at:	2015-02-09T13:30:13Z
Number of results:	17

Port Summary for Host 192.168.0.10

Service (Port)	Threat Level
general/tcp	Low
general/CPE-T	Log
3128/tcp	Log
25/tcp	Log
22/tcp	Log
123/udp	Log
111/tcp	Log

Security Issues for Host 192.168.0.10

general/tcp
Low (CVSS: 2.6)
NVT: TCP timestamps (OID: 1.3.6.1.4.1.25623.1.0.80091)

Summary

The remote host implements TCP timestamps and therefore allows to compute the uptime.

Vulnerability Detection Result

```
It was detected that the host implements RFC1323.
The following timestamps were retrieved with a delay of 1 seconds in-between:
Paket 1: 1274034887
Paket 2: 1274036027
```

Impact

A side effect of this feature is that the uptime of the remote host can sometimes be computed.

Solution

To disable TCP timestamps on linux add the line 'net.ipv4.tcp_timestamps = 0' to /etc/sysctl.conf. Execute 'sysctl -p' to apply the settings at runtime.

To disable TCP timestamps on Windows execute 'netsh int tcp set global timestamps=disabled'

Starting with Windows Server 2008 and Vista, the timestamp can not be completely disabled.

The default behavior of the TCP/IP stack on this Systems is, to not use the Timestamp options when initiating TCP connections, but use them if the TCP peer that is initiating communication includes them in their synchronize (SYN) segment.

See also: http://www.microsoft.com/en-us/download/details.aspx?id=9152

Vulnerability Insight

The remote host implements TCP timestamps, as defined by RFC1323.

Vulnerability Detection Method

Special IP packets are forged and sent with a little delay in between to the target IP. The responses are searched for a timestamps. If found, the timestamps are reported.

Details: TCP timestamps (OID: 1.3.6.1.4.1.25623.1.0.80091)

Version used: $Revision: 636 $

References

Other: http://www.ietf.org/rfc/rfc1323.txt

general/CPE-T
Log (CVSS: 0.0)
NVT: CPE Inventory (OID: 1.3.6.1.4.1.25623.1.0.810002)

Summary

This routine uses information collected by other routines about CPE identities (http://cpe.mitre.org/) of operating systems, services and applications detected during the scan.

Vulnerability Detection Result

```
192.168.0.10|cpe:/a:openbsd:openssh:6.0p1
192.168.0.10|cpe:/o:debian:debian_linux
```

Log Method

Details: CPE Inventory (OID: 1.3.6.1.4.1.25623.1.0.810002)

Version used: $Revision: 314 $

general/tcp
Log (CVSS: 7.8)
NVT: 3com switch2hub (OID: 1.3.6.1.4.1.25623.1.0.80103)

Summary

The remote host is subject to the switch to hub flood attack.

Description : The remote host on the local network seems to be connected through a switch which can be turned

into a hub when flooded by different mac addresses. The theory is to send a lot of packets (> 1000000) to the port of the switch we are connected to, with random mac addresses. This turns the switch into learning mode, where traffic goes everywhere. An attacker may use this flaw in the remote switch to sniff data going to this host

Reference : http://www.securitybugware.org/Other/2041.html

Vulnerability Detection Result

```
Fake IP address not specified. Skipping this check.
```

Solution

Lock Mac addresses on each port of the remote switch or buy newer switch.

Vulnerability Detection Method

Details: 3com switch2hub (OID: 1.3.6.1.4.1.25623.1.0.80103)

Version used: $Revision: 15 $

general/tcp
Log (CVSS: 0.0)
NVT: DIRB (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.103079)

Summary

This script uses DIRB to find directories and files on web applications via brute forcing.

Vulnerability Detection Result

```
DIRB could not be found in your system path.
OpenVAS was unable to execute DIRB and to perform the scan you
requested.
Please make sure that DIRB is installed and is
available in the PATH variable defined for your environment.
```

Log Method

Details: DIRB (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.103079)

Version used: $Revision: 13 $

general/tcp
Log (CVSS: 0.0)
NVT: arachni (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.110001)

Summary

This plugin uses arachni ruby command line to find web security issues.

See the preferences section for arachni options.

Note that OpenVAS is using limited set of arachni options. Therefore, for more complete web assessment, you should use standalone arachni tool for deeper/customized checks.

Vulnerability Detection Result

```
Arachni could not be found in your system path.
OpenVAS was unable to execute Arachni and to perform the scan you
requested.
Please make sure that Arachni is installed and that arachni is
available in the PATH variable defined for your environment.
```

Log Method

Details: arachni (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.110001)

Version used: $Revision: 683 $

general/tcp
Log (CVSS: 0.0)
NVT: Traceroute (OID: 1.3.6.1.4.1.25623.1.0.51662)

Summary

A traceroute from the scanning server to the target system was conducted. This traceroute is provided primarily

for informational value only. In the vast majority of cases, it does not represent a vulnerability. However, if the displayed traceroute contains any private addresses that should not have been publicly visible, then you have an issue you need to correct.

Vulnerability Detection Result

```
Here is the route from 172.16.41.10 to 192.168.0.10:
172.16.41.10
192.168.0.10
```

Solution

Block unwanted packets from escaping your network.

Log Method

Details: Traceroute (OID: 1.3.6.1.4.1.25623.1.0.51662)

Version used: $Revision: 14 $

22/tcp
Log (CVSS: 0.0)
NVT: SSH Protocol Versions Supported (OID: 1.3.6.1.4.1.25623.1.0.100259)

Summary

Identification of SSH protocol versions supported by the remote SSH Server. Also reads the corresponding fingerprints from the service.

The following versions are tried: 1.33, 1.5, 1.99 and 2.0

Vulnerability Detection Result

```
The remote SSH Server supports the following SSH Protocol Versions:
1.99
2.0
SSHv2 Fingerprint: e6:e9:52:f0:b9:ff:f2:56:05:d1:cc:fa:06:71:e6:e0
```

Log Method

Details: SSH Protocol Versions Supported (OID: 1.3.6.1.4.1.25623.1.0.100259)

Version used: $Revision: 43 $

22/tcp
Log (CVSS: 0.0)
NVT: SSH Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10267)

Summary

This detects the SSH Server's type and version by connecting to the server and processing the buffer received. This information gives potential attackers additional information about the system they are attacking. Versions and Types should be omitted where possible.

Vulnerability Detection Result

```
Detected SSH server version: SSH-2.0-OpenSSH_6.0p1 Debian-4+deb7u2
Remote SSH supported authentication: publickey,password
Remote SSH banner:
(not available)
CPE: cpe:/a:openbsd:openssh:6.0p1
Concluded from remote connection attempt with credentials:
  Login: OpenVAS
  Password: OpenVAS
```

Solution

Apply filtering to disallow access to this port from untrusted hosts

Log Method

Details: SSH Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10267)

Version used: $Revision: 588 $

22/tcp
Log (CVSS: 0.0)

NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
An ssh server is running on this port
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

25/tcp

Log (CVSS: 0.0)

NVT: SMTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10263)

Summary

This detects the SMTP Server's type and version by connecting to the server and processing the buffer received. This information gives potential attackers additional information about the system they are attacking. Versions and Types should be omitted where possible.

Vulnerability Detection Result

```
Remote SMTP server banner :
220 ESMTP
```

Solution

Change the login banner to something generic.

Log Method

Details: SMTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10263)

Version used: $Revision: 339 $

25/tcp

Log (CVSS: 0.0)

NVT: SMTP STARTTLS Detection (OID: 1.3.6.1.4.1.25623.1.0.103118)

Summary

Check if the remote Mailserver supports the STARTTLS command.

Vulnerability Detection Result

```
The remote Mailserver supports the STARTTLS command.
```

Log Method

Details: SMTP STARTTLS Detection (OID: 1.3.6.1.4.1.25623.1.0.103118)

Version used: $Revision: 703 $

25/tcp

Log (CVSS: 0.0)

NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
An SMTP server is running on this port
Here is its banner :
220 ESMTP
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

111/tcp
Log (CVSS: 0.0)
NVT: rpcinfo -p (OID: 1.3.6.1.4.1.25623.1.0.11111)

Summary

This script calls the DUMP RPC on the port mapper, to obtain the list of all registered programs.

Vulnerability Detection Result

```
These are the registered RPC programs:\
\
RPC program #100000 version 4 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100000 version 3 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100000 version 2 'portmapper' (portmap sunrpc rpcbind) on port 111/TCP
RPC program #100024 version 1 'status' on port 51873/TCP
RPC program #100000 version 4 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100000 version 3 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100000 version 2 'portmapper' (portmap sunrpc rpcbind) on port 111/UDP
RPC program #100024 version 1 'status' on port 48823/UDP
```

Log Method

Details: rpcinfo -p (OID: 1.3.6.1.4.1.25623.1.0.11111)

Version used: $Revision: 41 $

123/udp
Log (CVSS: 0.0)
NVT: NTP read variables (OID: 1.3.6.1.4.1.25623.1.0.10884)

Summary

A NTP (Network Time Protocol) server is listening on this port.

Vulnerability Detection Result

Vulnerability was detected according to the Vulnerability Detection Method.

Log Method

Details: NTP read variables (OID: 1.3.6.1.4.1.25623.1.0.10884)

Version used: $Revision: 487 $

3128/tcp
Log (CVSS: 0.0)
NVT: HTTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10107)

Summary

This detects the HTTP Server's type and version.

Vulnerability Detection Result

```
The remote web server type is :
pve-api-daemon/3.0
```

Solution

Configure your server to use an alternate name like 'Wintendo httpD w/Dotmatrix display' Be sure to remove common logos like apache_pb.gif. With Apache, you can set the directive 'ServerTokens Prod' to limit the information emanating from the server in its response headers.

Log Method

Details: HTTP Server type and version (OID: 1.3.6.1.4.1.25623.1.0.10107)

Version used: $Revision: 229 $

3128/tcp
Log (CVSS: 0.0)
NVT: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Summary

This plugin attempts to guess which service is running on the remote ports. For instance, it searches for a web server which could listen on another port than 80 and set the results in the plugins knowledge base.

Vulnerability Detection Result

```
A web server is running on this port
```

Log Method

Details: Services (OID: 1.3.6.1.4.1.25623.1.0.10330)

Version used: $Revision: 69 $

3128/tcp
Log (CVSS: 0.0)
NVT: wapiti (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.80110)

Summary

This plugin uses wapiti to find web security issues.

Make sure to have wapiti 2.x as wapiti 1.x is not supported.

See the preferences section for wapiti options.

Note that OpenVAS is using limited set of wapiti options. Therefore, for more complete web assessment, you should use standalone wapiti tool for deeper/customized checks.

Vulnerability Detection Result

```
wapiti could not be found in your system path.
OpenVAS was unable to execute wapiti and to perform the scan you
requested.
Please make sure that wapiti is installed and that wapiti is
available in the PATH variable defined for your environment.
```

Log Method

Details: wapiti (NASL wrapper) (OID: 1.3.6.1.4.1.25623.1.0.80110)

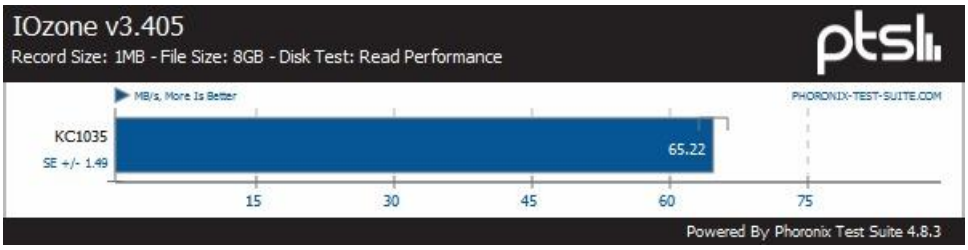
Version used: $Revision: 14 $

This file was automatically generated.

AA. Bijlage Benchmark parallel

Resultaat KC1035 – Ceph

```
IOzone 3.405:
 pts/iozone-1.8.0 [Record Size: 1MB - File Size: 8GB - Disk Test: Read Performance]
 Test 1 of 1
 Estimated Trial Run Count: 3
 Estimated Time To Completion: 13 Minutes
 Started Run 1 @ 14:50:48
 Started Run 2 @ 15:01:45
 Started Run 3 @ 15:10:38 [Std. Dev: 4.77%]
 Started Run 4 @ 15:17:49 [Std. Dev: 4.46%]
 Started Run 5 @ 15:25:57 [Std. Dev: 4.93%]
 Started Run 6 @ 15:34:10 [Std. Dev: 5.59%]

 Test Results:
 65.4951171875
 59.537109375
 63.0009765625
 65.5166015625
 67.8955078125
 69.8916015625

 Average: 65.22 MB/s
```

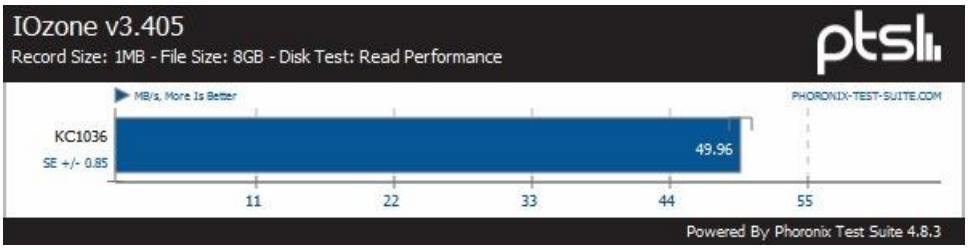


Resultaat KC1036 - Ceph

IOzone 3.405:
pts/iozone-1.8.0 [Record Size: 1MB - File Size: 8GB - Disk Test: Read Perfor mance]
Test 1 of 1
Estimated Trial Run Count: 3
Estimated Time To Completion: 13 Minutes
Started Run 1 @ 14:50:47
Started Run 2 @ 15:03:05
Started Run 3 @ 15:14:26 [Std. Dev: 2.96%]

Test Results:
51.4951171875
49.85546875
48.5439453125

Average: 49.96 MB/s

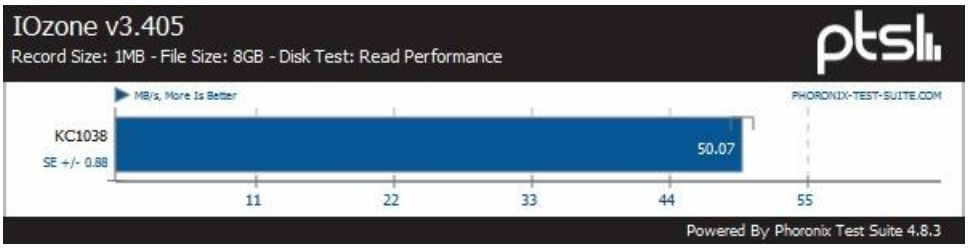


Resultaat KC1038 - Ceph

IOzone 3.405:
pts/iozone-1.8.0 [Record Size: 1MB - File Size: 8GB - Disk Test: Read Performance]
Test 1 of 1
Estimated Trial Run Count: 3
Estimated Time To Completion: 6 Minutes
Started Run 1 @ 14:50:48
Started Run 2 @ 15:03:03
Started Run 3 @ 15:14:26 [Std. Dev: 3.89%]
Started Run 4 @ 15:25:13 [Std. Dev: 3.50%]

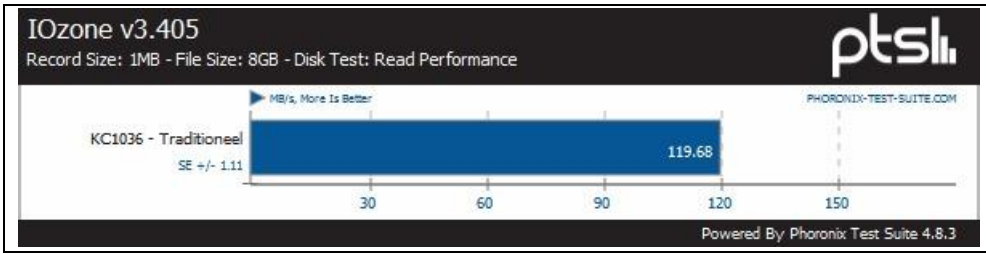
Test Results:
51.583984375
49.7705078125
47.7216796875
51.2119140625

Average: 50.07 MB/s



Resultaat KC1036 – Traditioneel

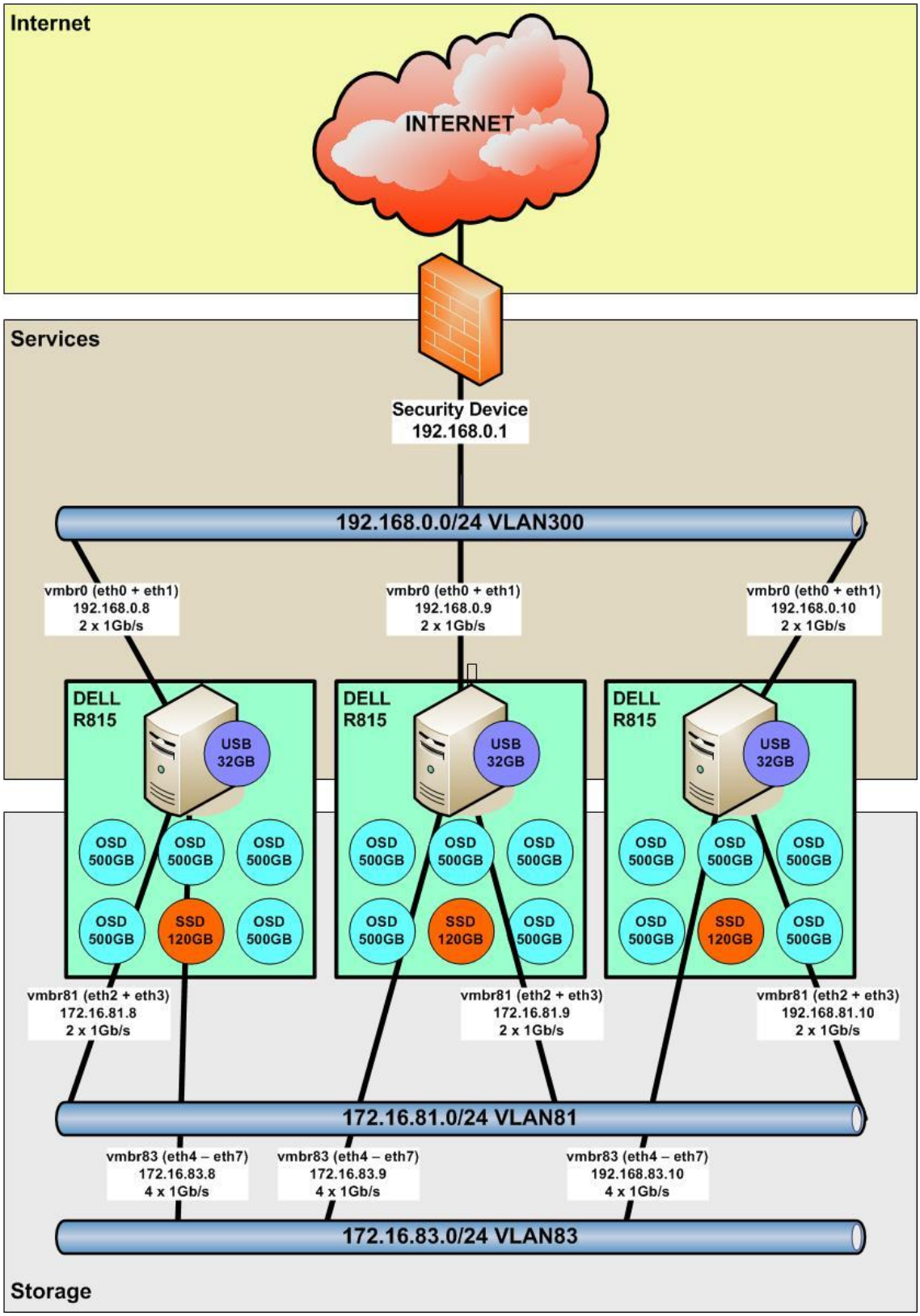
```
IOzone 3.405:
 pts/iozone-1.8.0 [Record Size: 1MB - File Size: 8GB - Disk Test: Read Performance]
 Test 1 of 1
 Estimated Trial Run Count: 3
 Estimated Time To Completion: 13 Minutes
 Started Run 1 @ 16:15:07
 Started Run 2 @ 16:21:17
 Started Run 3 @ 16:27:28 [Std. Dev: 1.61%]

 Test Results:
 119.7578125
 117.72265625
 121.564453125

 Average: 119.68 MB/s
```



Resultaat KC1038 - Traditioneel

```
IOzone 3.405:
 pts/iozone-1.8.0 [Record Size: 1MB - File Size: 8GB - Disk Test: Read Performance]
 Test 1 of 1
 Estimated Trial Run Count: 3
 Estimated Time To Completion: 7 Minutes
 Started Run 1 @ 16:15:07
 Started Run 2 @ 16:21:19
 Started Run 3 @ 16:27:30 [Std. Dev: 0.43%]

 Test Results:
 125.595703125
 125.18359375
 126.265625

 Average: 125.68 MB/s
```

BB. Bijlage Netwerktekening SOLL-definities



Figuur 88 Netwerktekening SOLL-situatie definities

CC. Bijlage TNK(Tactisch Normenkader) BIR

TNK referentie	TNK norm
11.3.1.1	Aan de gebruikers is een set gedragsregels aangereikt met daarin minimaal het volgende: • Wachtwoorden worden niet opgeschreven. • Gebruikers delen hun wachtwoord nooit met anderen. • Een wachtwoord wordt onmiddellijk gewijzigd indien het vermoeden bestaat dat het bekend is geworden aan een derde. • Wachtwoorden worden niet gebruikt in automatische inlogprocedures (bijv. opgeslagen onder een functietoets of in een macro).
10.1.3.4	Verantwoordelijkheden voor beheer en wijziging van gegevens en bijbehorende informatiesysteemfuncties moeten eenduidig toegewezen zijn aan één specifieke (beheerders)rol.
11.2.2	De toewijzing en het gebruik van speciale bevoegdheden behoren te worden beperkt en beheerst
11.1.1	Er behoort toegangsbeleid te worden vastgesteld, gedocumenteerd en beoordeeld op basis van bedrijfseisen en beveiligingseisen voor toegang.
12.6.1.4	Updates/patches voor kwetsbaarheden waarvan de kans op misbruik hoog is en waarvan de schade hoog is worden zo spoedig mogelijk doorgevoerd, echter minimaal binnen één week. Minder kritische beveiligings-updates/patches moeten worden ingepland bij de eerst volgende onderhoudsronde.
12.6.1.4	Updates/patches voor kwetsbaarheden waarvan de kans op misbruik hoog is en waarvan de schade hoog is worden zo spoedig mogelijk doorgevoerd, echter minimaal binnen één week. Minder kritische beveiligings-updates/patches moeten worden ingepland bij de eerst volgende onderhoudsronde
10.4.1.1	Bij het openen van bestanden worden deze geautomatiseerd gecontroleerd op virussen, trojans en andere malware. De update voor de detectiedefinities vindt frequent, minimaal één keer per dag, automatisch plaats.
10.4.1.2	Inkomende en uitgaande e-mails worden gecontroleerd op virussen, trojans en andere malware. De update voor de detectiedefinities vindt frequent, minimaal één keer per dag, (automatisch) plaats.
11.7.1.2	Er zijn, waar mogelijk, voorzieningen om de actualiteit van anti-malware programmatuur op mobiele apparaten te garanderen.
10.4.1.3	In verschillende schakels van een keten binnen de infrastructuur van een organisatie wordt bij voorkeur antivirusprogrammatuur van verschillende leveranciers toegepast.
12.6.1.2	Van softwarematige voorzieningen van de technische infrastructuur kan (bij voorkeur geautomatiseerd) gecontroleerd worden of de laatste updates (patches) in zijn doorgevoerd. Het doorvoeren van een update vindt niet geautomatiseerd plaats, tenzij hier speciale afspraken over zijn met de leverancier.
12.4.1.6	Er is een rollbackstrategie
12.6.1.3	Indien een patch beschikbaar is, dienen de risico's verbonden met de installatie van de patch te worden geëvalueerd (de risico's verbonden met de kwetsbaarheid dienen vergeleken te worden met de risico's van het installeren van de patch

DD. Technische implementatie BIR

Eis	**11.3.1.1**
Implementatie	Er zijn geen gebruikers van het PoC systeem. Potentiele eindgebruikers binnen het afstudeerbedrijf zijn op de hoogte van de regelgeving mbt wachtwoorden.

Eis	**10.1.3.4**
Implementatie	Beheerders kunnen met hun persoonlijke account in loggen. Op het achterliggende authenticatie- en authorisatiesysteem zijn deze gebruikersaccount voorzien van administrator rechten.

Eis	**11.2.2**
Implementatie	De rechten worden per systeem toegekend.

Eis	**11.1.1**
Implementatie	De regelgeving is gedocumenteerd binnen het afstudeerbedrijf.

Eis	**12.6.1.4**
Implementatie	Security patches worden dagelijks automatisch toegepast. Niet security gerelateerde patches worden toegepast indien er geupgrade dient te worden naar een nieuwe OS versie.

Eis	**10.4.1.1**
Implementatie	Alle bestanden worden voor installatie gecontroleerd op malware- en virusinfecties. Bovenliggende systemen worden voorzien van virusscanners.

Eis	**10.4.1.2**
Implementatie	Mail wordt verzonden via een centrale server die voorzien is van virusscanning. De updates van deze scanner worden dagelijks ververst.

Eis	**11.7.1.2**
Implementatie	Er wordt geen gebruik gemaakt van mobiele devices.

Eis	**10.4.1.3**
Implementatie	Voor email / patches en on-access worden verschillende scanners gebruikt.

Eis	**12.6.1.2**
Implementatie	Er wordt gebruik gemaakt van vFense voor het inzichtelijk maken van patch achterstanden.

Eis	**12.4.1.6**
Implementatie	Het betreft een PoC omgeving. Rollback is herinstallatie.

Eis	**12.6.1.3**
Implementatie	Risico’s zijn acceptabel binnen de PoC omgeving. Alle security patches worden geimplementeerd.

EE. Bijlage Motivaties MoSCoW

Must	**Proxmox dient de storage native te ondersteunen**
Product	Gluster
Link	https://pve.Proxmox.com/wiki/Storage:_GlusterFS
Resultaat	+
Motivatie voor resultaat	This new storage plugin supports the use of GlusterFS volumes

Must	**Proxmox dient de storage native te ondersteunen**
Product	Ceph
Link	https://pve.Proxmox.com/wiki/Storage:_Ceph
Resultaat	+
Motivatie voor resultaat	Since Proxmox 3.2, Ceph has been integrated and can be installed and configured through Proxmox specific commands, and run on Proxmox node

Must	**Proxmox dient de storage native te ondersteunen**
Product	Sheepdog
Link	https://pve.Proxmox.com/wiki/Storage:_Sheepdog
Resultaat	+/-
Motivatie voor resultaat	Note: Sheepdog is still not stable, so please do not use it in production environments.

Must	**Private cloud storage**
Product	Gluster
Link	http://en.wikipedia.org/wiki/Gluster#Private_cloud_deployment
Resultaat	+
Motivatie voor resultaat	A typical on-premises, or private cloud deployment will consist of GlusterFS installed as a virtual appliance on top of multiple commodity servers running hypervisors such as KVM, Xen, or VMware; or on bare metal.

Must	**Private cloud storage**
Product	Ceph
Link	http://ceph.com/community/career/storage-consultant/
Resultaat	+
Motivatie voor resultaat	They travel to Inktank customer sites to assist in the design and implementation of storage solutions for public or private cloud clouds.

Must	**Block storage**
Product	Gluster
Link	https://raobharata.wordpress.com/2013/11/27/glusterfs-block-device-translator/
Resultaat	+
Motivatie voor resultaat	Block device translator (BD xlator) is a new translator added to GlusterFS recently which provides block backend for GlusterFS.

Must	**Block storage**
Product	Ceph
Link	http://ceph.com/ceph-storage/block-storage/
Resultaat	+
Motivatie voor resultaat	You can mount Ceph as a thinly provisioned block device! When you write data to Ceph using a block device, Ceph automatically stripes and replicates the data across the cluster.

Must	**Horizontaal schaalbaar**
Product	Gluster
Link	http://www.gluster.org/
Resultaat	+
Motivatie voor resultaat	GlusterFS is a unified, poly-protocol, scale-out filesystem serving many *petabytes* of data.

Must	**Horizontaal schaalbaar**
Product	Ceph
Link	http://ceph.com/
Resultaat	+
Motivatie voor resultaat	Ceph is a distributed object store and file system designed to provide excellent performance, reliability and scalability.

Must	**Minimaal 2x maal opgeslagen**
Product	Gluster
Link	http://www.gluster.org/community/documentation/index.php/Gluster_3.1:_Configuring_Distributed_Replicated_Volumes
Resultaat	+
Motivatie voor resultaat	Distributed replicated volumes replicate (mirror) data across two or more nodes in the cluster. You can use distributed replicated volumes in environments where high-availability and high-reliability are critical. Distributed replicated volumes also offer improved read performance in most environments.

Must	**Minimaal 2x maal opgeslagen**
Product	Ceph
Link	http://ceph.com/docs/dumpling/rados/operations/pools/
Resultaat	+
Motivatie voor resultaat	Replicas: You can set the desired number of copies/replicas of an object. A typical configuration stores an object and one additional copy (i.e., size = 2), but you can determine the number of copies/replicas.

Must	**Integriteit van de data**
Product	Gluster
Link	https://access.redhat.com/documentation/en-US/Red_Hat_Storage/2.0/html/Administration_Guide/sect-User_Guide-Managing_Volumes-Self_heal.html
Resultaat	+
Motivatie voor resultaat	In replicate module, previously you had to manually trigger a self-heal when a brick goes offline and comes back online, to bring all the replicas in sync. Now the pro-active self-heal daemon runs in the background, diagnoses issues and automatically initiates self-healing every 10 minutes on the files which require healing.

Must	**Integriteit van de data**
Product	Ceph
Link	http://wiki.ceph.com/FAQs/How_Does_Ceph_Ensure_Data_Integrity_Across_Replicas
Resultaat	+
Motivatie voor resultaat	Resolution Ceph periodically scrubs placement groups to ensure that they contain the same information. Low-level or deep scrubbing reads the object data in each replica of the placement group to ensure that the data is identical across replicas.

Should	**Filesystem storage**
Product	Gluster
Link	http://www.gluster.org/documentation/About_Gluster/
Resultaat	+
Motivatie voor resultaat	GlusterFS is an open source, distributed file system capable of scaling to several petabytes (actually, 72 brontobytes!) and handling thousands of clients.

Should	**Filesystem storage**
Product	Ceph
Link	http://ceph.com/ceph-storage/file-system/
Resultaat	+
Motivatie voor resultaat	Ceph provides a traditional file system interface with POSIX semantics.

Should	**Object storage**
Product	Gluster
Link	http://www.gluster.org/wp-content/uploads/2012/05/Gluster_File_System-3.3.0-Administration_Guide-en-US.pdf
Resultaat	+/-
Motivatie voor resultaat	Object Storage is built upon Openstack's Object Storage Swift

Should	**Object storage**
Product	Ceph
Link	http://ceph.com/ceph-storage/object-storage/
Resultaat	+
Motivatie voor resultaat	Ceph's software libraries provide client applications with direct access to the RADOS object-based storage system, and also provide a foundation for some of Ceph's advanced features, including RADOS Block Device (RBD), RADOS Gateway, and the Ceph File System.

Should	**Performance**
Product	Gluster
Link	http://www.networkcomputing.com/storage/gluster-vs-ceph-open-source-storage-goes-head-to-head/a/d-id/1113581
Resultaat	+
Motivatie voor resultaat	The art of benchmarking is complex. Enough said. The decision on transfer sizes could itself account for Ceph running faster or slower than Gluster. We can only honestly measure performance is through an independent third party, with tuning input from both teams. This hasn't happened yet, and Red Hat's report is misleading.

Should	**Performance**
Product	Ceph
Link	http://www.networkcomputing.com/storage/gluster-vs-ceph-open-source-storage-goes-head-to-head/a/d-id/1113581
Resultaat	+
Motivatie voor resultaat	The art of benchmarking is complex. Enough said. The decision on transfer sizes could itself account for Ceph running faster or slower than Gluster. We can only honestly measure performance is through an independent third party, with tuning input from both teams. This hasn't happened yet, and Red Hat's report is misleading.

Should	**Gescheiden (storage) netwerk**
Product	Gluster
Link	http://www.gluster.org/community/documentation/index.php/Network_Configuration_Techniques
Resultaat	+
Motivatie voor resultaat	This method lets you add network capacity for multi-protocol sites by segregating traffic for different protocols on different network interfaces.

Should	**Gescheiden (storage) netwerk**
Product	Ceph
Link	http://ceph.com/docs/master/rados/configuration/network-config-ref/
Resultaat	+
Motivatie voor resultaat	We recommend running a Ceph Storage Cluster with two networks: a public (front-side) network and a cluster (back-side) network.

Should	Grafische interface
Product	Gluster
Link	http://www.gluster.org/community/documentation/index.php/Gluster_3.1:_Using_the_Gluster_Management_Console
Resultaat	+
Motivatie voor resultaat	The Gluster Storage Platform Management Console offers an advanced, Web-based interface that you can use to centrally manage your storage cluster.

Should	Grafische interface
Product	Ceph
Link	https://github.com/ceph/calamari
Resultaat	+
Motivatie voor resultaat	Web-based monitoring and management for Ceph

Should	Rolling upgrades
Product	Gluster
Link	http://www.gluster.org/community/documentation/index.php/Upgrade_to_3.5
Resultaat	+
Motivatie voor resultaat	If you have replicated or distributed replicated volumes with bricks placed in the right fashion for redundancy, have no data to be self-healed and feel adventurous, you can perform a rolling upgrade through the following procedure.

Should	Rolling upgrades
Product	Ceph
Link	http://ceph.com/docs/master/install/upgrading-ceph/
Resultaat	+
Motivatie voor resultaat	When upgrading from Dumpling (v0.64) you may perform a rolling upgrade.

Should	Snapshotting
Product	Gluster
Link	http://www.gluster.org/community/documentation/index.php/Features/Gluster_Volume_Snapshot
Resultaat	+
Motivatie voor resultaat	Gluster volume snapshot will provide point-in-time copy of a GlusterFS volume. This snapshot is an online-snapshot therefore file-system and its associated data continue to be available for the clients, while the snapshot is being taken.

Should	Snapshotting
Product	Ceph
Link	http://ceph.com/docs/master/rbd/rbd-snapshot/
Resultaat	+
Motivatie voor resultaat	One of the advanced features of Ceph block devices is that you can create snapshots of the images to retain a history of an image's state. Ceph also supports snapshot layering, which allows you to clone images (e.g., a VM image) quickly and easily.

Could	**Interactie met de cloudservices**
Product	Ceph
Link	http://ceph.com/docs/master/radosgw/s3/
Resultaat	+
Motivatie voor resultaat	Ceph supports a RESTful API that is compatible with the basic data access model of the Amazon S3 API.

Could	**Interactie met de cloudservices**
Product	Ceph
Link	http://www.inktank.com/for-service-providers/
Resultaat	+
Motivatie voor resultaat	Ceph Object Storage combined with interfaces like ownCloud or Citrix Sharefile, allow service providers to offer their own customized Dropbox solutions to consumer and enterprise customers.

Could	**Interactie met de cloudservices**
Product	Ceph
Link	https://www.openstack.org/summit/openstack-summit-hong-kong-2013/session-videos/presentation/ceph-the-de-facto-storage-backend-for-openstack
Resultaat	+
Motivatie voor resultaat	The De Facto Storage Backend for OpenStack

Could	**Deduplication**
Product	Ceph
Link	http://lists.ceph.com/pipermail/ceph-users-ceph.com/2013-August/033451.html
Resultaat	+
Motivatie voor resultaat	There is no active data deduplication, and, again, if I understand the architecture correctly, there probably never will be.

Could	**Data encryption**
Product	Ceph
Link	http://docs.ceph.com/docs/v0.80/rados/operations/authentication/
Resultaat	+
Motivatie voor resultaat	The cephx protocol does not address data encryption in transport (e.g., SSL/TLS) or encryption at rest.

Could	**Performance degradatie bij replicatie**
Product	Ceph
Link	http://ceph.com/docs/master/rados/configuration/osd-config-ref/
Resultaat	+
Motivatie voor resultaat	You can set operations priority weights between client operations and recovery operations to ensure optimal performance during recovery.

FF. Bijlage Requirements Cloud Storage

De vet gedrukte woorden zijn toetsingscriteria op basis waarvan de productselectie voor de POC plaats vindt.

Eisen

De opdrachtgever stelt de volgende eisen aan de PoC:

- De cloud storage oplossing moet minimaal dezelfde functionaliteit bieden als de huidige oplossing. De functionaliteit die de huidige oplossing biedt is het **aanbieden van virtualisatieimages** aan een virtualisatieplatform op basis van Proxmox. De huidige functionaliteit wordt geleverd op basis van **block** storage, de opdracht beperkt zich echter niet tot blocksystem storage, er moet ook gekeken worden naar alternatieven op basis van **filesystem** en **object storage**.

- Het huidige virtualisatieplatform **Proxmox dient de storage te ondersteunen**.

- De oplossing dient te voorzien in **private cloud storage.**

- De gekozen oplossing moet op basis van **open source software** wordt gerealiseerd. Hiervoor worden de volgende drie argumenten aangedragen:
  1 Hoewel de Rijksoverheid het gebruik van open source software niet als norm stelt, stimuleert de Rijksoverheid overheidsorganisaties in het gebruik van open source software: Het “comply or explain” principe .

  2 De innovatieafdeling maakt zo min mogelijk gebruik van closed source en betaalde licentie software;

  3 Gezien de doorlooptijd van verwervingstrajecten is het niet mogelijk binnen de afstudeertermijn betaalde licentie software te verwerven.

- De **performance(**throughput) van de cloud storage moet gelijk of beter zijn ten opzichte van de huidige traditionele storage-oplossing.

- De **opslagcapaciteit** van de cloud storage moet groter of gelijk zijn als de huidige oplossing. De opslagcapaciteit van de cloud storage moet daarnaast **horizontaal schaalbaar** zijn. Alle in de cloud storage opgeslagen gegevens moeten **minimaal 2 maal opgeslagen** worden.

- De cloudoplossing moet uit beveiligings- en performanceoogpunt gebruik maken van een **gescheiden (storage) netwerk**.

- De te realiseren cloudoplossing mag **geen** gebruik maken van **Fibre Channel**.

- Bij uitval van één of meerdere cloud storage nodes moet de **integriteit van de data** en functionaliteit behouden blijven.

Wensen

De opdrachtgever heeft de volgende wensen met betrekking tot de PoC:

- Dat de cloud storage mogelijkheden biedt voor de **interactie met de cloudservices**: Owncloud, OpenStack en Amazon S3.

- De opdrachtgever heeft de wens met **zeer beperkte investeringen** cloud storage te realiseren. De voorkeur gaat uit naar een minimale investering in hard- en software, de initiële grens ligt op €1250,-. Onder deze grens mag personeel namelijk zelfstandig verwerven en declareren. Investeringen boven dit bedrag dienen in overleg met de opdrachtgever plaats te vinden.

- De opdrachtgever wenst een **grafische interface** met vergelijkbare functionaliteit ten opzichte van de huidige grafische interface. We spreken hier over de grafische interface van het storage component zelf, niet over de weergave van de storage binnen Proxmox.

- De opdrachtgever wenst de mogelijkheid tot **rolling upgrades** om downtime te beperken. Rolling upgrade zorgt dat cloud storage software componenten geüpgraded kunnen worden zonder onderbreking van de dienst.

- De opdrachtgever wenst de mogelijkheid tot **snapshotting** van virtualisatie images. Uitleg over de term snapshotting is terug te vinden in 2.3.

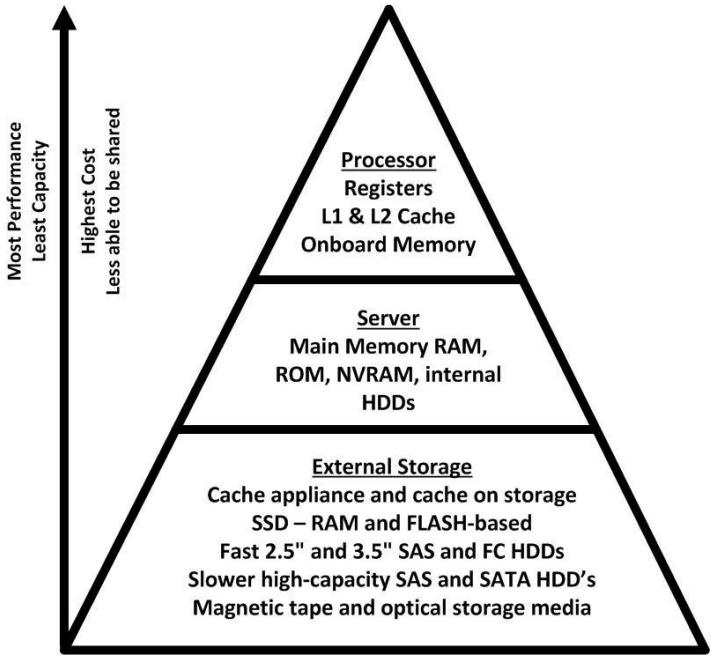
- De opdrachtgever wenst ondersteuning van **deduplication** binnen de cloud storage software. Uitleg over de term deduplication is terug te vinden in 2.3.

- De opdrachtgever wenst de mogelijkheid tot **data encryption** binnen de cloud storage software. Uitleg over de term data encryption is terug te vinden in 2.3.

- Bij uitval van één of meerdere cloud storage nodes moet de eis van minimaal 2x opslaan zo snel mogelijk in ere worden hersteld. Er mag hierbij een **performance degradatie** optreden.

GG. Bijlage Cloud storage

Storage is terug te vinden in alle lagen van een ICT-infrastructuur. Niet alle storage is echter goed bruikbaar als cloud storage, zo kunnen de kosten te hoog zijn of de capaciteit te laag. Greg Schultz beschrijft in zijn boek "Cloud and Virtual Data Storage Networking" een storage hiërarchie. Zijn hiërarchische model bekijkt storage vanuit de tweestrijd tussen kosten en Quality of Service (QoS).



Figuur 89 Storage hierarchie (Schulz, 2011, p26)

Cloud storage doet tot op zekere hoogte afbraak aan dit hiërarchische model. Cloud storage hoort thuis in de laag "External Storage". In de praktijk is deze laag opgebouwd uit het samenvoegen van de interne HDDs (HardDiskDrive) van meerdere componenten uit de laag "Server". Hoe lager de storage oplossing in de hiërarchie zit des te groter de elasticiteit. Cloud storage biedt de elasticiteit door het samenvoegen van meerdere harddisks tot een groot geheel.

Enkele voorbeelden van cloud storage zijn:

I-drive (Arias, 2012, p. 83): De naam is afgeleid van "internet drive". I-drive levert van origine online storage voor het opslaan van mp3 muziekbestanden. I-drive is merkwaardig genoeg al in augustus 1998 opgericht. I-drive levert: "***Storage as a Service***" .

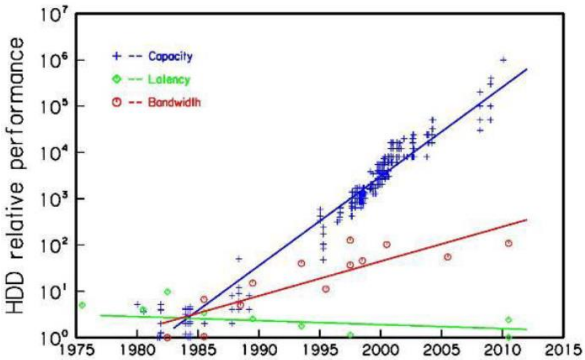
Carbonite (Arias, 2012, p. 15): is een online backup service (***Backup as a Service***) voor consumenten en MKB bedrijven. De online backup service is beschikbaar voor Windows en Apple gebruikers. Bestandstransmissie vindt plaats over een beveiligde verbinding, daarnaast worden bestanden versleuteld opgeslagen.

CTERA Networks (Arias, 2012, p. 41)(Arias, 2012, p41) : Het achterhalen van een ***Disk as a Service*** provider blijkt lastig te zijn. Een dienst die hierbij het meest in de buurt komt is CTERA Networks. CTERA levert een cloud gateway, een fysiek kastje wat het meest weg heeft van een oversized adsl-modem. Het is mogelijk om aan deze gateway USB of eSATA disks aan te sluiten. Deze worden vervolgens gerepliceerd naar CTERA.

Figuur 90: CTERA

HH. Harddisks

Harddisken worden qua capaciteit steeds groter, maar niet evenredig sneller. Dit is vooral te wijten aan het feit dat de mechanische disken gewoonweg niet harder kunnen draaien (Gnanasundaram, 2012, p. 37). Een harddisk is opgebouwd uit één of meer draaiende schijven met cilindrische sporen. De verplaatsingssnelheid van het buitenste cilindrische spoor mag de geluidsbarière niet doorbreken, om schade aan de disk te voorkomen. Daarnaast zijn er nog een aantal principes gekoppeld aan traditionele harddisks. Er is tijd nodig om de juiste locatie op het spoor voorbij te laten draaien en er is tijd nodig om de leeskop naar het juiste cilindrische spoor te verplaatsen. Dit veroorzaakt een vertraging (latency) die wordt uitgedrukt in milliseconden, wat voor computers die op gigahertz snelheid werken, gelijkwaardig is aan een eeuwigheid. Daarnaast gebruiken harddisks altijd stroom (ze draaien altijd) en een harddisk is het component met de hoogste mate van uitval in een datacenter. **Solid State devices** (Arias, 2012, pp. 46-49) (SSD's) op basis van NAND flash technologie zijn daarom een goede ontwikkeling om een directe performance boost te geven en de latency onder de milliseconde grens te brengen. Daarnaast gebruiken SSD's veel minder energie. Bij solid state speelt nog wel een probleem met de slijtage. Ondanks dat ze geen bewegende delen hebben, slijten solid state devices na veelvuldig gebruik wel. De cellen die de bits op moeten slaan, verliezen na verloop van tijd hun capaciteit en worden daardoor onbruikbaar, de elektronica kan een gelimiteerd aantal lees- en schrijfbewerkingen afhandelen.



Figuur 91 Harddisken over de jaren

II. Bijlage Gibibyte versus Gigabyte

Binnen de wereld van opslagcapaciteit (en netwerkcapaciteit) wordt gerekend met andere eenheden dan wij gewend zijn. Zo wordt er niet gesproken over gigabytes maar over gibibytes. De oorsprong van deze verschillende termen is terug te vinden in het formaat van het getal. Een gibibyte is opgebouwd uit machten van 2 en een gigabyte is opgebouwd uit machten van 10.

In onderstaande tabel zijn deze door de IEC (International Electrotechnical Commission) vastgestelde eenheden terug te vinden.

		Base 2			Base 10
kibi	ki	2^10	kilo	k,K	10^3
mebi	Mi	2^20	mega	M	10^6
gibi	Gi	2^30	giga	G	10^9
tebi	Ti	2^40	tera	TB	10^12
pebi	Pi	2^50	peta	P	10^15
exbi	Ei	2^60	exa	E	10^18
zebi	Zi	2^70	zetta	Z	10^21
yobi	Yi	2^80	yotta	Y	10^24

Tabel 11 (SI) meeteenheden storage

Dit wordt inzichtelijk wanneer een nieuwe harddisk wordt geplaatst in een systeem. De binnen deze scriptie gebruikte 500GB harddisken zijn na plaatsing in een systeem maar 466GB. Een zeer klein deel van dit verlies is te wijten aan informatie over partities, maar de meerderheid van de verloren gigabyte is te wijten aan de conversie van gibibytes naar gigabytes.

De fysieke harddisk heeft 500GB op het label staan, 500GB is gelijk aan 10^9 * 500 wat neer komt op 500000000000 bytes. De server maakt echter gebruik van gibibytes, de 500000000000 bytes worden gedeeld door 2^30 (1073741824) dit resulteert in 465,661. Afgerond is dit 466GB.

$$daadwerkelijke\ opslagcapaciteit\ in\ GB = \frac{10^9 * (GB\ op\ etiket\ harddisk)}{2^{30}}$$

Tabel 12 formule daadwerkelijke opslagcapaciteit

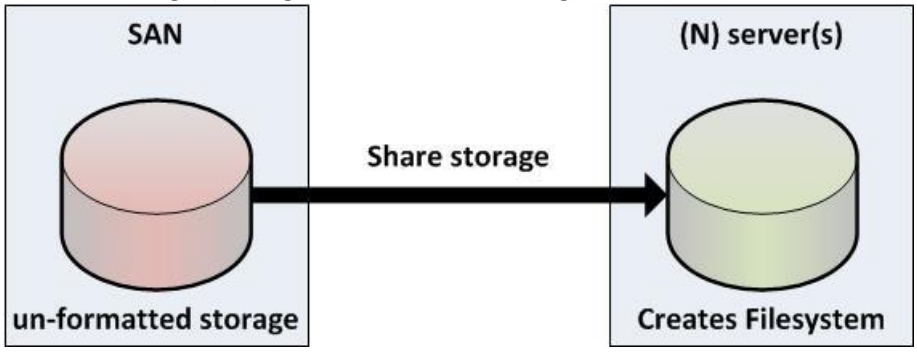
Binnen het onderzoek is bovenstaande formule van toepassing. Dit levert over 15 disks die in dit onderzoek gebruikt zijn een verlies van 510GB op.

JJ. Bijlage Opslagstructuren

De in deze bijlage gebruikte tekeningen bevatten harddisken in de kleuren rood en groen. Rood gekleurde harddisken zijn ongeformatteerd, groen gekleurde harddisken zijn voorzien van een formattering.

Block storage

Block storage (Farley, 2013, p. 93) sluit het best aan bij DaaS –Disk as a Service-. Block storage is het aanbieden van een ruwe disk of RAID set (0) zonder filesystem. Block storage wordt toegepast in SAN-technologie (Storage Attached Networking).

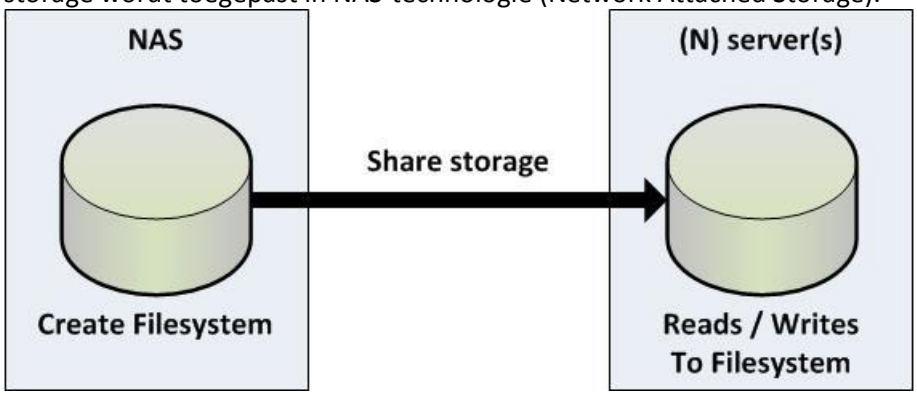


Figuur 92 Schematische weergave block storage

Na het aanbieden van de storage aan één of meerdere servers dient een filesystem aangemaakt te worden voordat het mogelijk is bestanden op te slaan en terug te lezen van de storage. Het aanmaken van het filesystem wordt bewerkstelligd door de storage te voorzien van een formattering.

Filesystem storage

Filesystem storage (Gnanasundaram, 2012, p. 22) sluit het best aan bij SaaS – Storage as a Service-. Filesystem storage is het aanbieden van een reeds geformatteerde disk of RAID set. Filesystem storage wordt toegepast in NAS-technologie (Network Attached Storage).

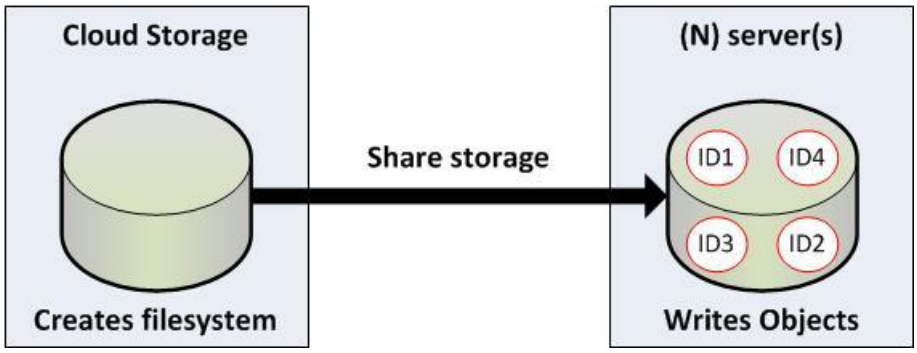


Figuur 93 Schematische weergave filesystem storage

Na het aanbieden van de storage aan één of meerdere servers is het direct mogelijk bestanden te schrijven en terug te lezen van de storage.

Object-based storage

Object based storage (Gnanasundaram, 2012, pp. 179-198) biedt de mogelijkheid informatie op te slaan als objecten op basis van onder andere inhoud, in plaats van op basis van naam en locatie. Een object is een stuk data (meestal een bestand) met alle metadata. Aan objecten wordt een ID meegegeven dat meestal voortkomt uit de inhoud van het object (zowel file als metadata). Het is alleen mogelijk een object op te vragen aan de hand van dit ID. Object storage sluit het best aan bij DaaS -Data Storage as a Service-.



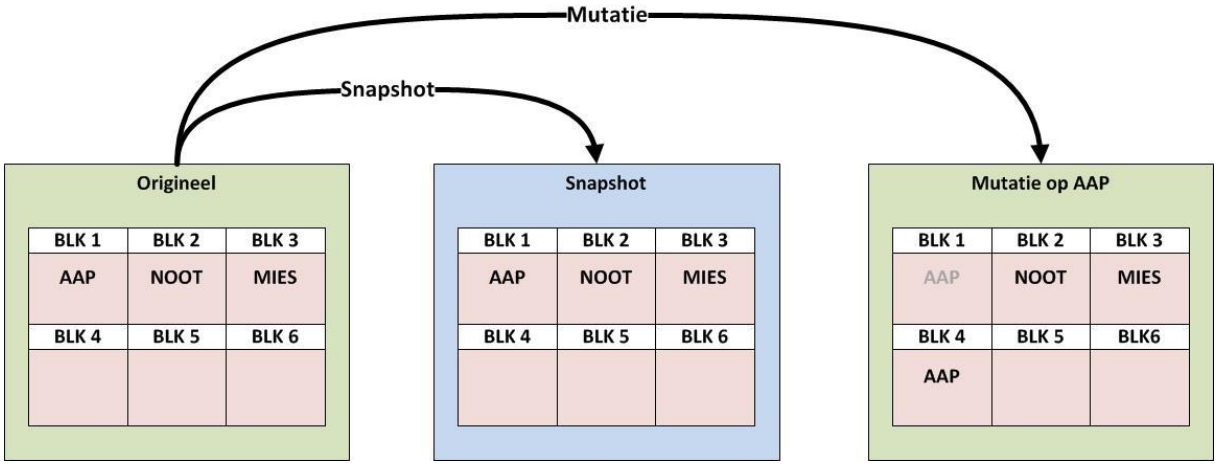
Figuur 94 Schematisch weergave object storage

KK. Bijlage Storage Technieken

Snapshotting

Een snapshot (Gnanasundaram, 2012, p. 271) is een (block)pointer gebaseerde replica*. Het voordeel van een snapshot is, dat de replica zeer weinig capaciteit vereist. Snapshots kunnen worden toegepast op filesystems en Logical Volumes (zie LVM) (0). Een snapshot maakt gebruik van het Copy on Write (CoW) principe. Tijdens het maken van een snapshot wordt een bitmap en een blockmap aangemaakt. De bitmap houdt bij welke wijzigingen er plaatsvinden op het originele filesystem. De Blockmap geeft aan welke adressen(block) uitgelezen dienen te worden indien het snapshot wordt aangesproken.

*functioneert als een replica



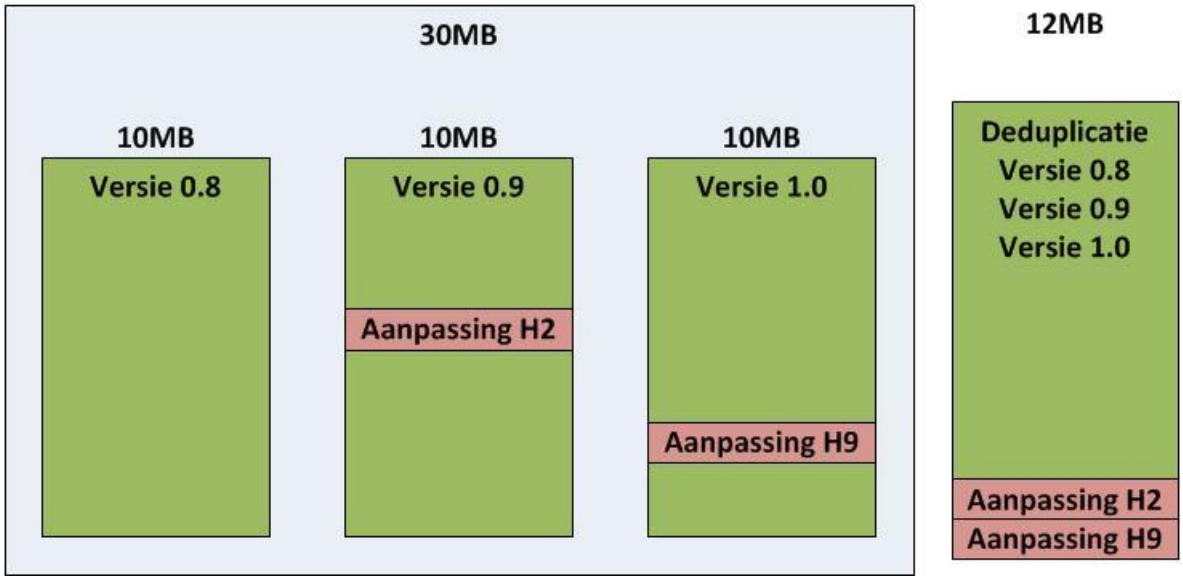
Figuur 95 Snapshotting en mutaties

Indien er een mutatie plaatsvindt (nadat een snapshot is gemaakt) wordt deze weggeschreven op een andere pointer. De originele data blijft behouden, echter alleen de snapshot kent de pointer naar het originele block. Een logisch gevolg van snapshotting is dat er tijdens het maken van een snapshot geen schrijfacties plaats kunnen vinden.

Deduplication

Bestanden worden onderverdeeld in kleine blokjes welke worden voorzien van een hash. De hashes worden onderling met elkaar vergeleken. Indien er meerdere bestanden of blokje zijn die exact gelijk zijn, wordt slechts één kopie van dat blokje of bestand opgeslagen.

Een passend voorbeeld is deze scriptie. Deze scriptie is na veel iteraties tot stand gekomen. Bij met name de laatste iteraties zullen er minieme aanpassingen worden gedaan op het document. Grote delen van de scriptie zijn in iedere versie gelijk. Deduplicatie maakt het mogelijk de overeenkomstige gegevens slechts één maal op te slaan. Hierdoor ontstaat een reductie in het dataverbruik.



Figuur 96 Voorbeeld deduplicatie opslag scriptie

Het voorbeeld van deduplicatie (Yoder, 2013, p. 59) (Gnanasundaram, 2012, pp. 249-252) op de versies van deze scriptie is reëel. Tijdens het realiseren van de scriptie is gebruik gemaakt van dropbox ten behoeve van het delen van documenten met de afstudeerbegeleiders. Dropbox (Arias, 2012, pp. 58-61) maakt daadwerkelijk gebruik van deduplicatie[71].

Data encryption

Data encryption is het versleutelen van data, zowel gegevens “at-rest”op de disk als tijdens de transmissie van data. (Arias, 2012, p. 85) (Gnanasundaram, 2012, pp. 341,342,361). Encryptie van data kan bijvoorbeeld gebruikt worden wanneer er meerdere gebruikers gebruik maken van een enkele gedeelde schijf. Of (voor) het kwijtraken van een USB stick.

[71] https://blogs.dropbox.com/dropbox/2011/07/changes-to-our-policies/

RAID

RAID (Redundant Array of Inexpensive Disks) bestaat uit het samenvoegen van meerdere harddisks met als doel: meer capaciteit, meer performance of meer betrouwbaarheid. De samengevoegde harddisks wordt een RAID set genoemd. Een aantal veel voorkomende RAID opstellingen zijn:

- RAID0 (Troppens, 2009, pp. 25-26)
- RAID1 (Troppens, 2009, p. 26)
- RAID5 (Troppens, 2009, pp. 31-35)
- RAID6 (Troppens, 2009, pp. 35-37)
- RAID10 (Troppens, 2009, pp. 26-31)

Deze opstellingen worden RAID levels genoemd. Deze RAID levels kunnen op basis van software of hardware worden gerealiseerd.

RAID 0
Het samenvoegen van 2 of meer harddisks tot één groot geheel. Naast *capaciteitstoename* is hier een toename in *performance* te verwachten. Het lezen en schrijven van gegevens is mogelijk op beide harddisks, hierdoor is het mogelijk sneller te lezen en te schrijven.

RAID1
Het samenvoegen van 2 harddisks. Deze harddisks vormen een exacte kopie van elkaar. Het gebruik van RAID1 heeft als voordeel dat bij uitval van 1 disk de gegevens behouden blijven. Daarnaast verdubbelt de leessnelheid (in theorie). De schrijfperformance blijft gelijk.

RAID5
Voor het gebruik van RAID5 zijn minimaal 3 disks nodig. De pariteit van de disks wordt onderling verdeeld. Pariteit is additionele informatie die naast de gegevens wordt weggeschreven. Op basis van de additionele informatie zijn de originele gegevens te achterhalen. Dit maakt het mogelijk bij uitval van een enkele disk aan de hand van de pariteit alle gegevens terug te lezen. RAID5 levert een goede leessnelheid aangezien meerdere disks de gegevens bevatten. Het additioneel wegschrijven van pariteit data zorgt voor een lagere schrijfsnelheid.

RAID6
Vergelijkbaar met RAID5. Pariteitsinformatie wordt echter 2 maal opgeslagen. Het tweemaal wegschrijven van informatie heeft een negatief effect op de schrijfsnelheid.

RAID10
Een combinatie van RAID0 en RAID1. Dit RAID level biedt performance en betrouwbaarheid door de kopie van de 2 disks.

LVM (Logical Volume Management)

Wanneer een traditionele harddisk vol is, is er geen eenvoudige manier om de capaciteit van de harddisk te vergroten. Om de harddisk toch te kunnen vergroten is LVM (Gnanasundaram, 2012, pp. 20-22) uitgevonden. LVM stelt een gebruiker in staat twee of meerdere disks zonder RAID controller samen te voegen tot een groter geheel. Het concateneren van de harddisks levert een (LV) logisch volume op. Naast LVM zijn er nog een aantal leverancier afhankelijke implementaties zoals ZVOL en 3PAR.

LL. Bijlage Scrum concepten

Sprints

Binnen die project zijn drie sprints (Verheyen, 2013, p. 102) gedefineerd. Een sprint heeft als doel het realiseren van een functioneel component. Binnen iedere sprint is een enkele user story behandeld. Een user story bestaat uit 3 componenten, “als wie”, “wil ik wat” en het “waarom”. Er zijn twee actoren binnen het project gedefineerd. De opdrachtgever en een techneut. Aan het eind van iedere sprint is een demonstratie gegeven van de gerealiseerde functioneliteit. De opdrachtgever was zeer te spreken over de geboekte resultaten.

Backlog

Alle binnen het project gedefineerde user stories zijn samengevoegd in een product backlog (Verheyen, 2013, p. 102). Aan de hand van deze backlog is serial scrum (Verheyen, 2013, p. 86) toegepast. Serial scrum is het lineair uitvoeren van taken. De reden hiervoor was dat alle kennis binnen het team aanwezig was.

Standups

Hoewel de projectorganisatie zeer klein was vonden er toch daily standups plaats met de Scrum master, gezien de werkdruk was het alleen mogelijk aan de afstudeeropdracht te werken op donderdag, vrijdag, zaterdag en zondag. Daily standups vonden telefonisch plaats op donderdag- en vrijdagmiddag. De Scrummaster was een gecertificeerde bij de Scrum Alliance[72].

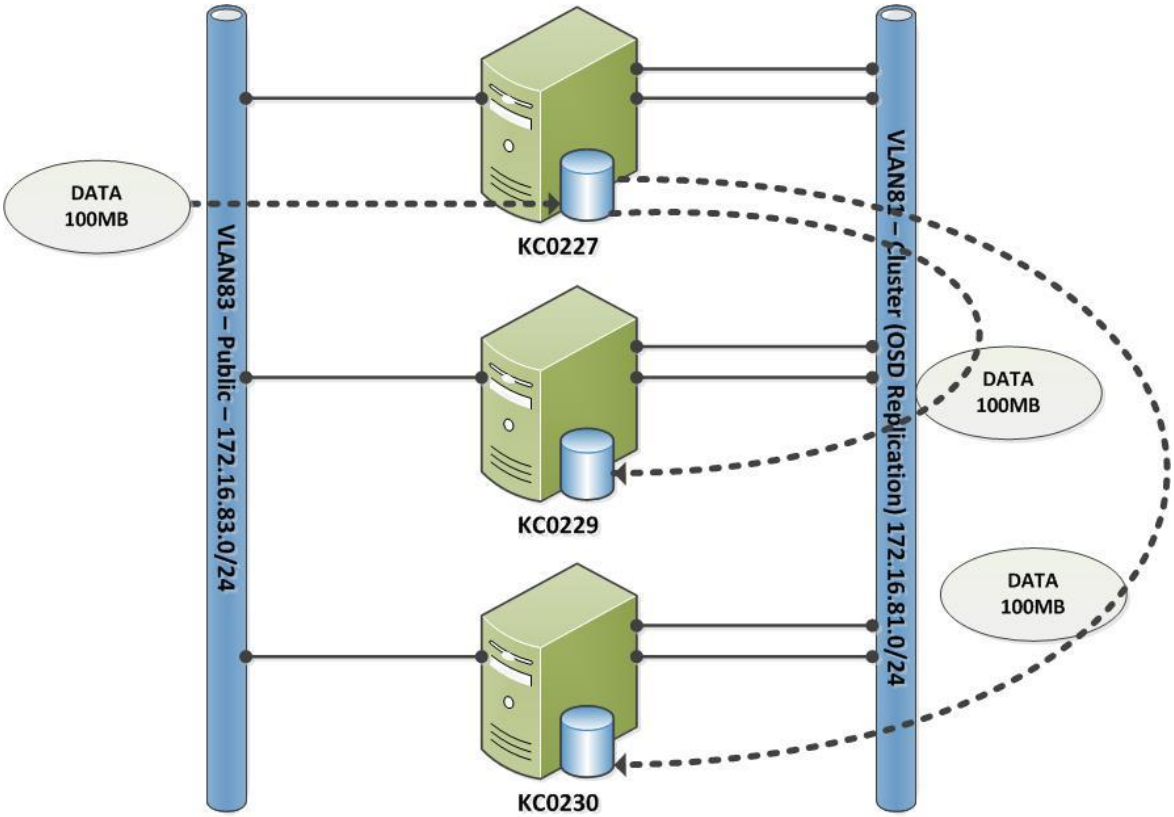
Tijdens de daily standup werd besproken wat er uitgevoerd was, of hier impediments op waren en welke taken er opgepakt gingen worden. De uit te voeren taken zijn voor aanvang in kaart gebracht. Impediments moesten verholpen worden om door te kunnen, alle sprints zijn op tijd afgesloten ondanks de grote hoeveelheid impediments en de complexiteit hiervan.

[72] https://www.scrumalliance.org/certifications/practitioners/certified-scrummaster-csm

MM. Bijlage Ceph concepten

Journal

Ceph slaat standaard drie kopieën op en pas nadat de drie kopieën opgeslagen zijn krijgt de client de melding dat de gegevens opgeslagen zijn. De eerste replica wordt weggeschreven op het public netwerk en de overige twee kopieën worden op het cluster netwerk weggeschreven.

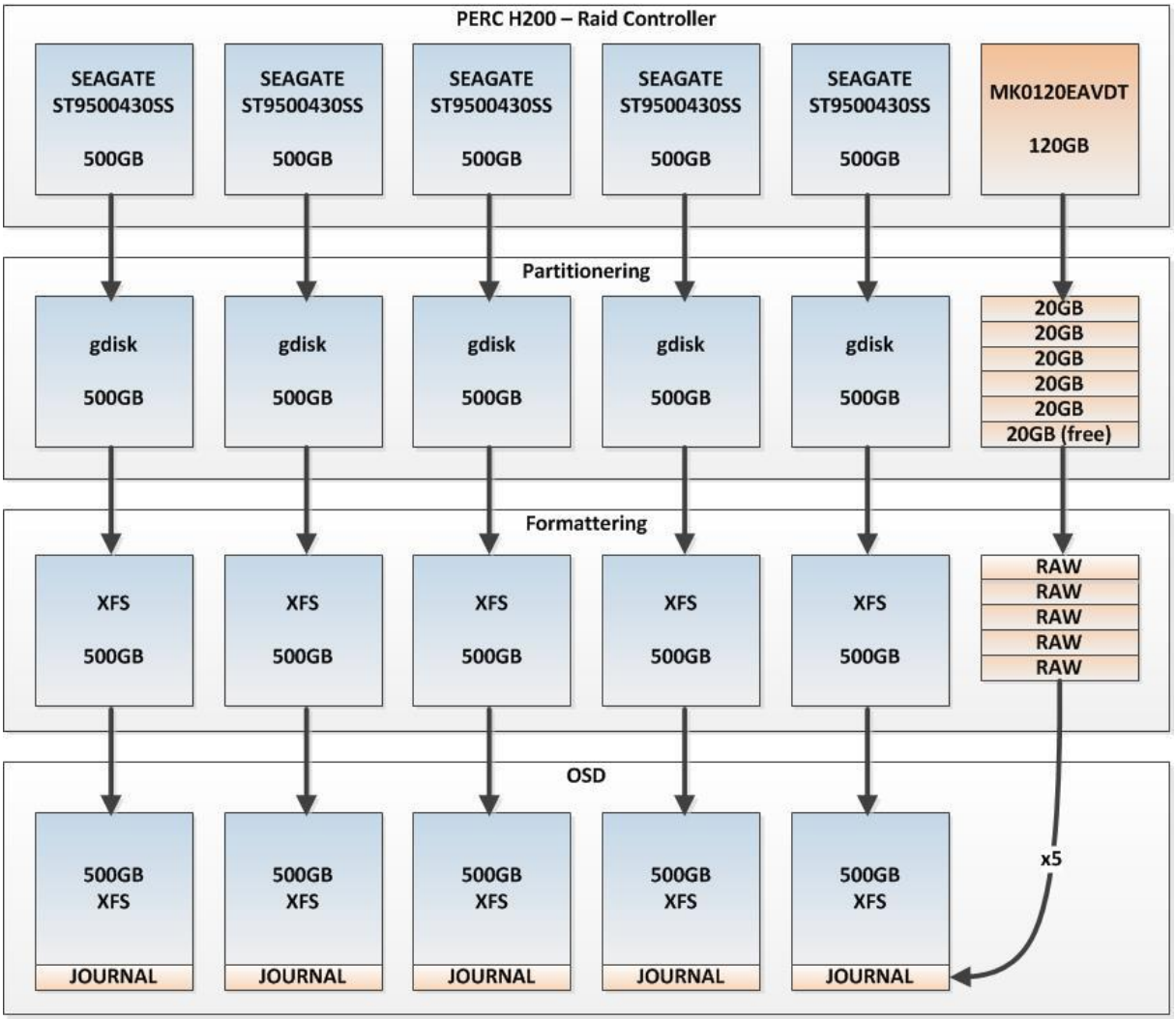


Figuur 97 Schematische weergave writes

Ceph schrijft de te schrijven informatie eerst weg op een zogenaamde journal. Om zo snel mogelijk terug te koppelen dat de informatie is weggeschreven wordt vanuit Ceph aangeraden gebruik te maken van een Solid State Disk.

OSD

Binnen Ceph vindt de opslag van gegevens plaats op één of meerdere Object Storage Daemons (OSD´s). Een OSD bestaat uit een traditionele disk en een journal. Voor het in gebruik kunnen nemen van een OSD dient een harddisk gepartitioneerd en geformatteerd te worden. Hierna is het mogelijk de traditionele disk en de journal samen te voegen tot een OSD. Het schrijven naar de OSD vindt initieel plaats naar de journal, hierna worden deze gegevens door Ceph gerepliceerd van de journal naar de achterliggende traditionele disk. Leesacties vinden te allen tijde plaats vanaf het traditionele deel van de OSD.



Figuur 98 Schematische weergave van fysieke disk naar OSD

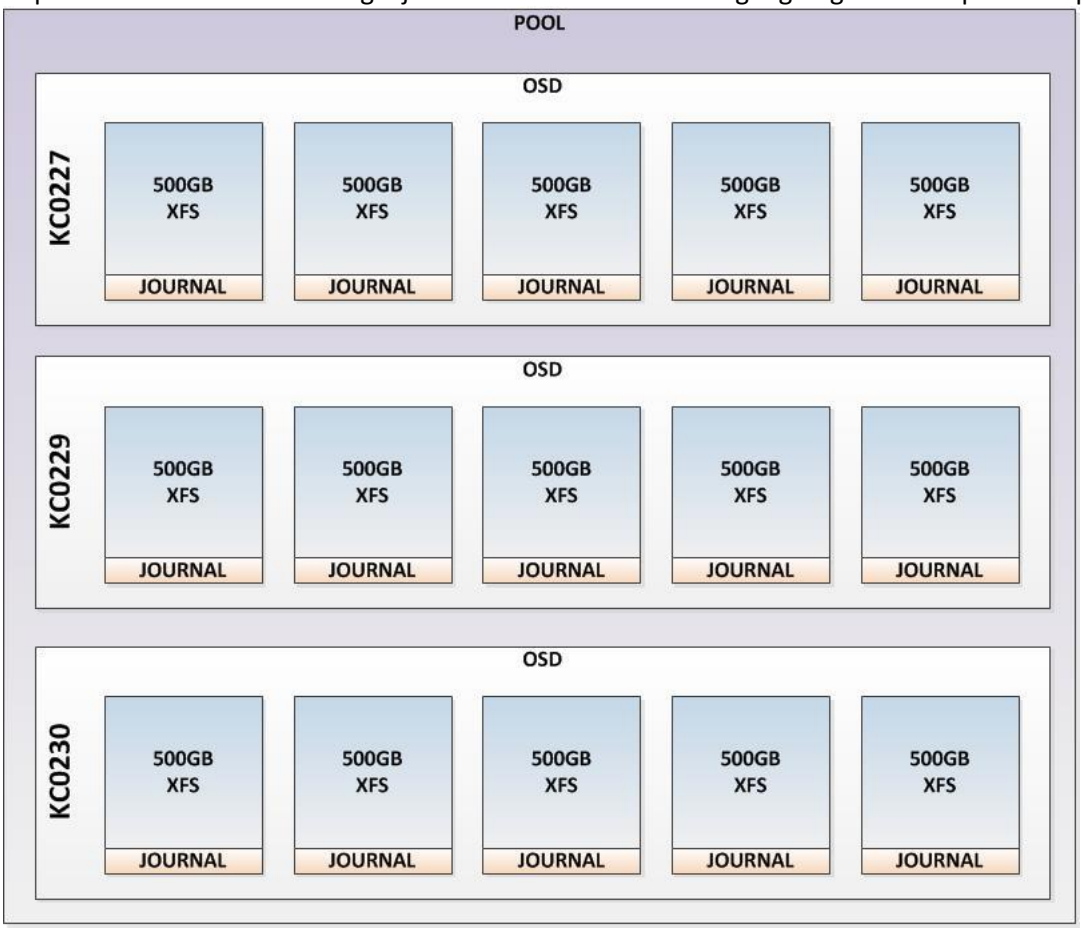
De in de systemen aanwezige RAID controllers bieden de schrijven aan als JBOD (Just a Bunch Of Disks). Er wordt geen gebruik gemaakt van RAID.

Cephx authenticatie

Om gebruik te kunnen maken van de Ceph storage dient gebruik te worden gemaakt van Cephx authenticatie[73]. Cephx is een authenticatiemechanisme op basis van sleutels. De sleutels kunnen zodanig specifiek ingesteld worden dat communicatie alleen mogelijk is met één OSD of één monitor hosts. Indien er geen sleutel aanwezig is kan er niet met Ceph componenten gecommuniceerd worden. Cephx authenticatie staat default aan.

Pool

Een pool[74] is te vergelijken met een RAID set. Een pool voegt alle OSD's tot één groot geheel. Binnen dit geheel worden de drie kopieën opgeslagen. Het is mogelijk meerdere pools aan te maken, om bijvoorbeeld vanuit één Ceph cluster meerdere klanten te voorzien van storage. Aan de hand van Cephx authenticatie is het mogelijk verschillende cliënten toegang te geven tot specifieke pools.

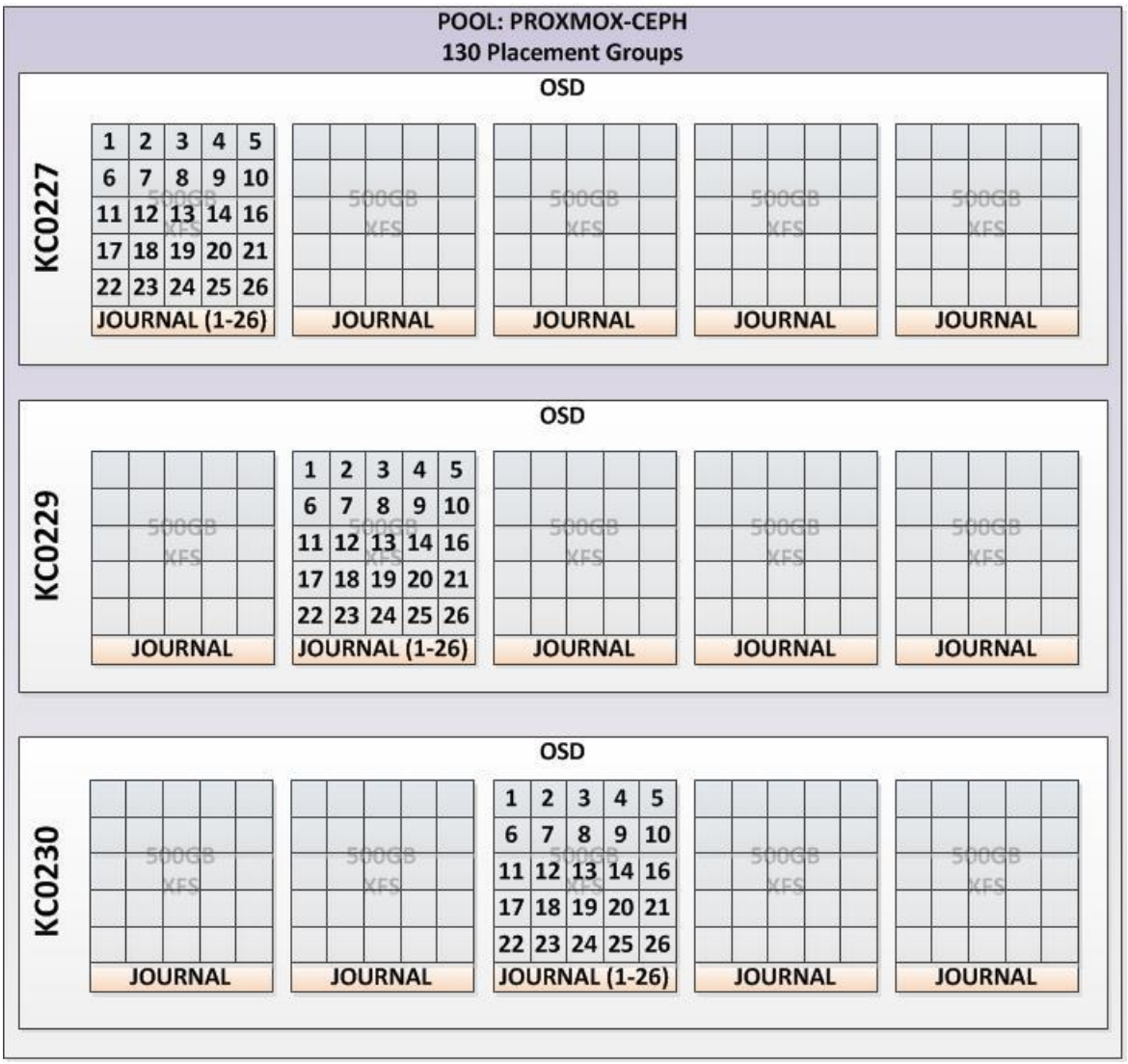


Figuur 99 Schematische weergave van een pool

[73] https://ceph.com/docs/v0.79/rados/operations/auth-intro/
[74] https://ceph.com/docs/v0.79/rados/operations/pools/

Placement groups

Placement groups[75] zorgen voor de verdeling van informatie over de verschillende fysieke nodes van het Ceph cluster. De drie replica's worden verdeeld over de verschillende OSD's. In onderstaand voorbeeld zijn 130 placement groups aangemaakt. De exacte verdeling over de disks is aan de hand van een reeks commando's inzichtelijk te maken. De theorie maakt in ieder geval kenbaar dat de replica's over verschillende machines verdeeld worden.



Figuur 100: Schematische weergave van placement groups binnen een pool.

[75] http://docs.ceph.com/docs/v0.79/rados/operations/placement-groups/

CRUSHmaps

Een CRUSHmap[76] verdeelt de gegevens over de verschillende OSD´s. Crushmaps maken het mogelijk informatie te verdelen over verschillende OSD´s. Zo is het mogelijk zwaardere servers, snellere disks en controllers zwaarder te belasten, of oudere machines te ontzien. Binnen dit onderzoek is gebruik gemaakt van de default CRUSHmap, deze wordt gegenereerd op basis van de opslagcapaciteit.

Hieronder is de CRUSHmap terug te vinden welke binnen het onderzoek gebruikt is. Merk op dat de som van de “weight” van de OSD gelijk is aan de “weight” van de host. De som van de “weight” van de hosts is weer gelijk aan de “weight” van de root.

osd.8 is na vervanging een 600GB harddisk geworden ipv een 500GB, vandaar dat de “weight” 0.55 is ipv 0.45.

```
root@KC0227-MON:~# ceph osd tree
# id    weight  type name       up/down         reweight
-1      6.85            root default
-2      2.25            host KC0227-MON
0       0.45            osd.0           up              1
1       0.45            osd.1           up              1
2       0.45            osd.2           up              1
3       0.45            osd.3           up              1
4       0.45            osd.4           up              1
-3      2.35            host KC0229-MON
5       0.45            osd.5           up              1
7       0.45            osd.7           up              1
6       0.45            osd.6           up              1
8       0.55            osd.8           up              1
10      0.45            osd.10          up              1
-4      2.25            host KC0230-MON
9       0.45            osd.9           up              1
11      0.45            osd.11          up              1
12      0.45            osd.12          up              1
13      0.45            osd.13          up              1
14      0.45            osd.14          up              1
```

[76] https://ceph.com/docs/v0.79/rados/operations/crush-map/

NN. Bijlage Bespreking Concept

Bespreking concept	Tussentijds assessment	Eerste beoordeling

Formulier bespreking concept afstudeerdossier

Student: Marco Stroosnijder **Studentnummer:** 7053126

Datum: 04-02-15

Tijdens de bespreking is het volgende geconstateerd:		**ja**	**nee**
a	*Het voortgangsverslag is ontvangen*	X	
b	*Het afstudeerdossier is digitaal beschikbaar*	X	
c	*Het afstudeerdossier is opgebouwd conform de richtlijnen*	X	
d	*Het goedgekeurde afstudeerplan is aanwezig*	X	
e	*Het plan van aanpak is aanwezig*	X	
f	*Reeds geleverd commentaar is aanwezig*	X	
g	*Het afstudeerdossier geeft voldoende inzicht in de stand van zaken*	X *)	
h	*De afstudeeropdracht is tot nu toe naar behoren uitgevoerd*	X *)	

Verbeterpunten:

Let op niet teveel detail opnemen in het afstudeerverslag, dat kan altijd nog naar bijlagen!

Opmerkingen:

Punt G en H: Met de mondelinge toelichting van Marco tijdens het gesprek is er voldoende vertrouwen in het succesvol afronden.

Aandachtspunt voor Marco is de beschikbare tijd voor het afstuderen naast zijn drukke baan als leidinggevende van een team. Marco geeft nu al aan tot in de late uurtjes aan zijn afstudeerproject te werken, dat lijkt me een risico voor hem.

Naam begeleidend examinator: Wim Mooijekind

Datum: 7-2-2015

Dit formulier wordt door de begeleidend examinator digitaal ingevuld en per email naar de student verstuurd met een cc naar de coördinator van ICT & Media @ Work (A.M.Schipper@hhs.nl). Het formulier dient door de student te worden opgenomen in het afstudeerdossier.

OO. Bijlage Tussentijds Assessment

Bespreking concept	**Tussentijds assessment**	Eerste beoordeling

Formulier tussentijds assessment

Student: Marco Stroosnijder **Studentnummer:** '07053126'

Datum: 05-03-15 **eerste / tweede TTA:** Eerste TTA

Tijdens het tussentijds assessment is het volgende geconstateerd:		**ja**	**nee**
a	*Het voortgangsverslag is ontvangen*	X	
b	*Het afstudeerdossier is digitaal beschikbaar*	X	
c	*Het afstudeerdossier is opgebouwd conform de richtlijnen*	X	
d	*Het goedgekeurde afstudeerplan is aanwezig*	X	
e	*Het plan van aanpak is aanwezig*	X	
f	*Reeds geleverd commentaar is aanwezig*	X	
g	*Het afstudeerdossier geeft voldoende inzicht in de stand van zaken*	X	
h	*De afstudeeropdracht is tot nu toe naar behoren uitgevoerd*	X	

Aanpak	**O**	**T**	**V**	**G**
Passend				X
Theoretisch verantwoord				X
Samenhang uitvoering beroepstaken			X	

Beroepstaken op afgesproken niveau uitgevoerd?		**O**	**T**	**V**	**G**
1	A1 Analyseren van het probleemdomein				X
2	A3 Achterhalen behoefte van belanghebbenden			X	
3	B6 Selecteren van bestaande hardware/softwarecomponent			X	
4	C12 Ontwerpen van een gedistribueerd systeem			X	
5	D19 Realiseren van een gedistribueerd systeem			X	
6	E24 Kwantitatieve analyse maken van de prestaties van systemen				X
7					
8					

Producten	**O**	**T**	**V**	**G**
Tussenproducten			X	
Eindproducten			X	

Effectief communiceren	**O**	**T**	**V**	**G**
Binnen afstudeerbedrijf (nog geen informatie ontvangen van afstudeerbedrijf)				
Afstudeerdossier			X	

Reflectie	**O**	**T**	**V**	**G**
Inzicht in eigen functioneren			X	
Inzicht in eigen leerproces			X	

Toelichting per beoordelingscriterium

Aanpak

De aanpak van Marco mbt zijn afstudeeropdracht/project is goed. Hij heeft zicht uitstekend verdiept in het theoretisch kader mbt de techniek. De opbouw, uitwerking en uitvoering van de afstudeeropdracht is door Marco goed uitgevoerd.

Beroepstaken op afgesproken niveau uitgevoerd?

Ja.
Wat we missen in de requirement analyse is de relatie naar de aspecten benoemd in de doelstelling van de afstudeeropdracht; beveiliging, beheersbaarheid, performance, betrouwbaarheid, kosten en schaalbaarheid.
De architectuurplaat van de Soll situatie is wel opgenomen maar niet verklaard
Het bedrag van 500,00,-- genoemd in het resultaat van de opdrachtbeschrijving komt niet meer terug in het verslag.

Producten

Het Afstudeerverslag is behoorlijk gedetailleerd uitgewerkt. Dat maakt het niet prettig leesbaar. Onze feedback is om kritisch te kijken naar wat bijlages kan om dit te verbeteren. Ook zou het hoofdstuk mbt het onderzoek en de aanpak wat eerder in het verslag moeten komen te staan zodat de opbouw logischer wordt. Je moet er vanuit gaan dat het verslag gelezen en begrepen moet kunnen worden door andere personen dan die in je project hebben geparticipeerd.

Het afstudeerverslag voldoet nog niet op de volgende punten:

1. Persoonlijke reflectie
2. Meer beschrijven van het proces hoe je tot keuzes bent gekomen.

Effectief communiceren
Vanuit het afstudeerbedrijf is nog geen input ontvangen mbt dit punt. Vanuit perspectief school is de communicatie adequaat geweest.

Reflectie
Marco had tijdens het TTA zichtbare moeite met het ontvangen en verwerken van de feedback op zijn afstudeerverslag, waar hij bijzonder trots op is. Je moet het ontvangen van feedback niet zien als een persoonlijk falen maar beschouw het als positief om je verslag nog beter te maken

Advies

X	Inleveren (bindend advies)
	Verlengen (vrijblijvend advies)
	Stoppen (vrijblijvend advies)

Besluit student

Aankruisen welke beslissing de student heeft genomen (alleen na vrijblijvend advies)

X	**Afstudeerdossier wordt op afgesproken datum ingeleverd**	Inleverdatum:30 maart 2015
	Afstudeerperiode wordt verlengd	Inleverdatum:
	Student stopt met afstudeeropdracht	

Naam begeleidend examinator: Wim Mooijekind

Naam tweede examinator: Madelon Nieuwland

Datum: 5 maart 2015

Dit formulier wordt door de tweede examinator digitaal ingevuld, waarna de begeleidend examinator het per email verstuurt naar de student met een cc naar de coördinator van ICT & Media @ Work (A.M.Schipper@hhs.nl). Het formulier dient door de student te worden opgenomen in het afstudeerdossier.