Treball de fi de grau

Grau d'Enginyeria Informàtica

Anàlisi tècnica per a l'expansió i modernització d'un Hosting ISP

Entrega Final TFC

Estudiant: Raül Aguilera Fornieles

Professors: Javier Panadero Martínez
José Manuel Castillo Pedrosa

Àrea: Administració de xarxes i sistemes operatius

Resum

El treball proposat per aquest TFC consistirà a fer un pla d'empresa per a un Hosting ISP en el qual es contemplarà un estat inicial d'una empresa real en funcionament i es farà una anàlisi tècnica per tal d'assumir un creixement futur notable i un canvi en infraestructures tècniques.

Aquesta anàlisi ens permetrà redissenyar la xarxa interna administrativa de l'empresa tenint present els diferents departaments i les relacions entre ells, la xarxa DMZ pels servidors de gestió interns, les xarxes públiques allotjades a varis CPDs on s'allotjarà amb tecnologia cloud híbrida tan servidors propis de l'empresa com hosting per a clients. En el disseny de les diferents xarxes s'hauran de tenir en compte aspectes com la distribució d'IPs (Networking), tipus d'IPs (IPv4 i/o IPv6), hardware de xarxa, etc.

Un tema important serà el disseny d'aquest entorn cloud híbrid, el seu dimensionament, la tria de les tecnologies d'emmagatzematge i virtualització, tot el sistema de còpies de seguretat amb el disseny dels protocols de còpia i recuperació, el sistema de monitoratge dels sistemes i per tant un sistema d'alertes per controlar l'estat dels servidors i serveis que ens permeti tenir un pla de continuïtat de negoci o de recuperació davant desastres.

Per últim, i relacionat amb els apartats anteriors, també es faria un pla de costos aproximats del que suposaria la implementació d'aquest pla i la planificació de la seva implantació per tal d'inferir el mínim possible en les estructures en producció actuals.

Abstract

The work proposed for this TFC will be to make a business plan for an ISP Hosting in which an initial state of a functioning company will be contemplated and a technical analysis will be carried out in order to assume a remarkable future growth and a change in technical infrastructures.

This analysis will allow us to redesign the internal administrative network of the company taking into account the different departments and the relationships between them, the DMZ network for internal management servers, public networks hosted on several CPDs, where they will be rented with hybrid cloud technology, so much Company's own servers as hosting for clients. In the design of the different networks, aspects such as the distribution of IPs (Networking), types of IPs (IPv4 and / or IPv6), network hardware, etc. should be taken into account.

An important issue will be the design of this cloud hybrid environment, its size, the choice of storage and virtualization technologies, the entire backup system with the design of copying and recovery protocols, the system monitoring system and therefore, a system of alerts to control the status of servers and services that allow us to have a business continuity or disaster recovery plan.

Lastly, and related to the previous sections, an approximate cost plan would also be made which would involve the implementation of this plan and the planning of its implementation in order to infer the least possible in the structures in current production.

Index

Index de figures

Index de Taules

1. Introducció

En aquesta secció intentarem exposar la situació actual de l'empresa per a la qual volem realitzar aquest pla d'empresa i a partir de la qual haurem de desenvolupar el projecte. Haurem de definir els diferents objectius que volem aconseguir i establir els primers passos a prendre per tal de començar a caminar en aquest projecte.

1.1 Context i justificació del treball

El projecte que portarem a terme surt de la necessitat de l'empresa a la qual pertanyo per elaborar un pla d'empresa per a contemplar una expansió de futur i un pas endavant en l'ús de noves tecnologies. Aquesta és una empresa creada a mitjans dels anys 90 per un grup d'empreses dedicades al món del paper i motivades per la creixent derivació dels serveis que oferien al món d'internet. Inicialment orientada a donar accés a internet a empreses i serveis relacionats amb gestions amb la tresoreria, a poc a poc va anar adaptant-se fins a convertir-se en una empresa de serveis d'internet orientats a l'empresa. Serveis com ara allotjament de servidors físics i virtuals privats, allotjament compartit de web i correu, creació d'aplicacions web, contractació i gestió de dominis i marques i qualsevol altra solució informàtica orientada a les empreses. Pel que fa al personal, aquesta empresa comença amb quatre treballadors encarregats, de forma molt cooperativa, de l'àrea tècnica, administrativa i d'atenció al client en una única oficina. Però juntament amb l'empresa, va anar creixent el personal fins a arribar als vint-i-nou treballadors repartits entre la central i dues delegacions, i definits en els diferents departaments que la conformen. Departaments com ara, dominis, administració, gerència, comercial, servei tècnic, sistemes i aplicacions.
Així i tot, en l'actualitat han sorgit nous reptes que obliguen a realitzar un canvi profund que ve motivat per una sèrie de necessitats que la infraestructura actual no ens pot oferir. Aquestes necessitats estan relacionades amb un major dinamisme en la gestió i administració dels recursos, una major escalabilitat de les infraestructures, unificació de tecnologies i millora dels processos de treball del departament IT.
Per altra banda, aquest projecte només contemplarà la part tècnica d'aquesta expansió deixant al marge aspectes com ara personal necessari, campanyes de publicitat o qualsevol altra cosa que no estigui relacionada amb aspectes tècnics.

Al final, amb aquest projecte es vol aconseguir un pla de futur viable i realitzable per tal que l'empresa el porti a terme i no sols pugui fer front a una expansió puntual, sinó que optimitzi les infraestructures de cara a futurs creixements. Per aquest motiu serà molt

important tenir en compte l'escalabilitat dels sistemes a triar tant en l'apartat d'emmagatzemament com en el de virtualització.

1.2 Objectius del treball

Els objectius principals que aquest projecte pretén aconseguir estarien basats en la implementació d'una nova infraestructura per a l'hostatge de màquines virtuals per a l'empresa comentada anteriorment, de tal manera que es pugui substituir i/o actualitzar la ja existent. Per aquest motiu els objectius a aconseguir serien els següents.

- Disseny de les diferents xarxes que formen la infraestructura de l'empresa.
    - Xarxa interna.
    - Xarxa pública.
    - Xarxa de backup.
    - Xarxa d'emmagatzematge.

- Disseny d'un sistema distribuït d'emmagatzematge.
    - Emmagatzematge de backup.
    - Emmagatzematge productiu.

- Disseny d'un sistema distribuït de virtualització.
    - Nivell maquinari.
    - Nivell programari.

- Disseny dels sistemes de seguretat i prevenció
    - Polítiques d'accés.
    - Polítiques de backup.
    - Sistemes de monitoratge.

- Implantació
    - planificació
    - migració i coexistència.

1.3 Enfocament i mètode seguit

L'enfocament del projecte vindrà definit per la necessitat i la posició inicial de l'empresa per la qual es farà el pla d'empresa. Inicialment partim d'unes infraestructures ja existents i en producció.

L'estat inicial de l'empresa passa per una oficina central amb cinc departaments principals i ben diferenciats. Dominis, administració, comercial, tècnic i aplicacions. A més, l'empresa manté dues delegacions connectades amb la central mitjançant VPN[1]. Pel que fa a la part de hosting, aquesta està dividida en dos CPDs separats entre ells per més de 25 km i una petita sala de servidors a les oficines centrals.
L'empresa te adjudicades actualment 8 classes C d'IPv4 distribuïdes entre els dos CPDs i la sala de servidors i una xarxa IPv6 ::/48 a un dels CPDs.

Actualment, en l'estructura associada al hosting, està dividida en diferents subxarxes que intenten aïllar els diferents tipus de hosting. És a dir, els servidors associats a hosting compartit estan aïllats a una subxarxa dels servidors de hosting privat i dels servidors interns de l'empresa. Pel que fa al maquinari, tant per a clients com per ús intern, s'estan fent servir, generalment, servidors aïllats on s'allotgen diferents servidors virtuals. Cada servidor gestiona els seus propis recursos tant de memòria i CPU, com d'espai en disc, el qual està físicament a cada servidor.

Per últim, el programari de virtualització que es fa servir es VMware ESXi en les versions 5.5, 6.0, 6.5 i 6.7. i en algun cas concret gestionats amb un vCenter server.

Sobre aquesta infraestructura s'haurà d'implantar una de nova en la qual ens trobarem amb seccions que esdevindran noves, d'altres que hauran de suplir seccions ja obsoletes i per últim hi haurà seccions de les infraestructures antigues que no es podran tocar, per la qual cosa hauran de coexistir temporalment amb la infraestructura nova.
Aquest requeriment en portarà a triar una estratègia diferent en cada cas.

1.4 Planificació del Treball

La planificació del projecte s'ha fraccionat en 4 fases on s'agrupen els diferents aspectes d'aquest.

- **Fase 1 - Disseny de xarxes**
- **Fase 2 - Disseny sistema de hosting**
- **Fase 3 - Sistemes de seguretat i protecció**
- **Fase 4 – Implantació**

A continuació es passa a detallar cada fase amb el llistat de tasques que inclou i el diagrama de Gantt corresponent.

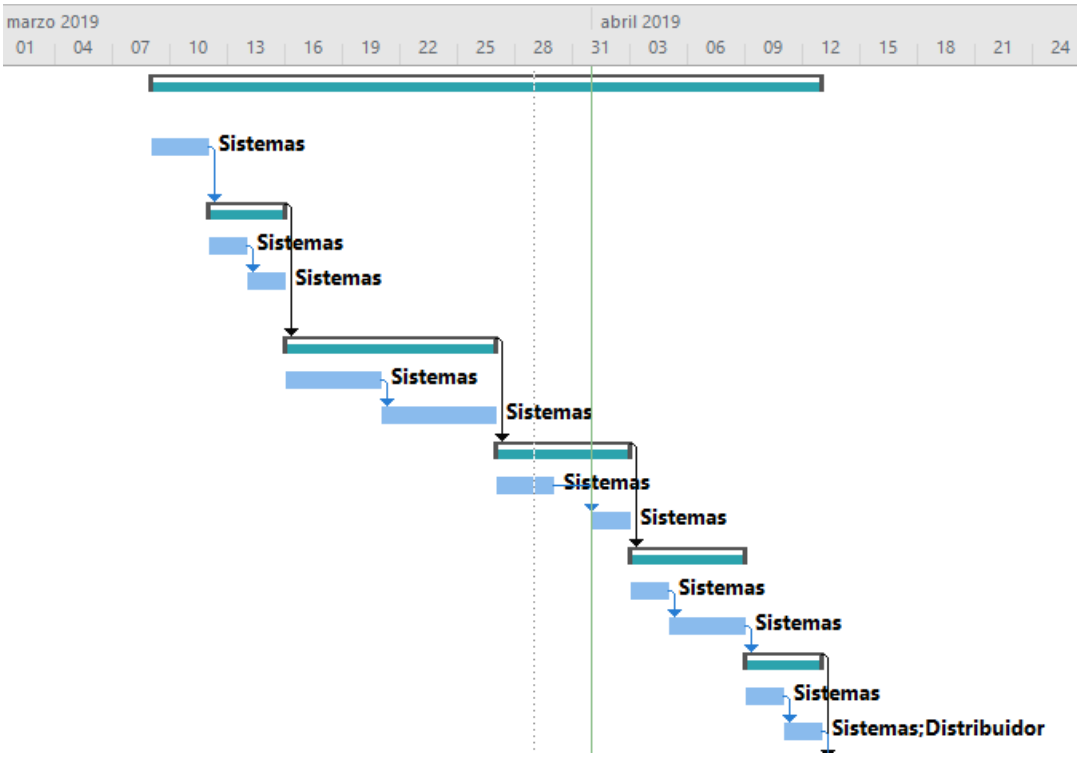
Fase 1 - Disseny de xarxes

En aquesta fase es portarà a terme el disseny de les diferents xarxes diferenciades pel seu propòsit. Inicialment, haurem de portar a terme una primera part on recavarem les dades necessàries pel disseny i una segona que serà el disseny en si. La xarxa interna de la oficina hauria de ser la xarxa més senzilla de dissenyar, ja que només s'haurà d'actualitzar l'estructura actual amb nou maquinari de xarxa i instal·lació de punts d'accés wifi. Per altra banda, pel que fa a les dues oficines connectades amb VPN, també es plantejarà actualitzar el maquinari. Per aquesta raó hauria d'estar completat en quatre dies.
També tenim el disseny de la xarxa pública, aquesta xarxa allotjarà els diferents servidors tant d'ús privat com de clients. En aquesta xarxa s'hauran de segmentar les classes d'IPs assignades per tal de donar servei a diferents propòsits. Aquesta xarxa contemplarà la implantació tant d'IPv4 com IPv6. Per aquest motiu penso que necessitarem uns 8 dies per a dissenyar aquestes xarxes.
Paral·lelament a la xarxa pública haurem de dissenyar una xarxa per a les còpies de backup. Actualment ja existeixen diverses xarxes de backup però amb un nivell molt elevat d'aïllament entre elles que dificulta la seva gestió. Per aquest motiu s'hauran de redissenyar. Penso que amb uns 5 dies seria suficient per a dissenyar-les perquè siguin més eficients.
Per últim, haurem de dissenyar les xarxes on allotjarem l'estructura d'emmagatzemament que connectaran amb els servidors de la xarxa pública per servir storage. Aquestes xarxes no haurien de portar més de 4 dies per al disseny.

Els últims quatre dies d'aquesta fase es dedicaran a l'anàlisi del maquinari de xarxa que es farà servir en les diferents xarxes.

	Modo de	Nombre de tarea	Duración	Comienzo	Fin	Predecesor	Nombres de los recursos
1		**Fase 1 - Disseny de xarxes**	**27 días**	**sáb 09/03/19**	**vie 12/04/19**		**Sistemas**
2		Recopilació de dades (rangs IPs, etc)	2 días	sáb 09/03/19	lun 11/03/19		**Sistemas**
3		**Disseny xarxa interna**	**4 días**	**mar 12/03/19**	**vie 15/03/19**		**Sistemas**
4		- Administrativa	2 días	mar 12/03/19	mié 13/03/19		**Sistemas**
5		- DMZ	2 días	jue 14/03/19	vie 15/03/19	4	**Sistemas**
6		**Disseny xarxa publica**	**8 días**	**sáb 16/03/19**	**mar 26/03/19**	**3**	Sistemas
7		- Serveis propis	4 días	sáb 16/03/19	mié 20/03/19		**Sistemas**
8		- Serveis clients	4 días	jue 21/03/19	mar 26/03/19	7	**Sistemas**
9		**Disseny xarxa backup**	**5 días**	**mié 27/03/19**	**mar 02/04/19**	**6**	Sistemas
10		- Interna	3 días	mié 27/03/19	vie 29/03/19		**Sistemas**
11		- Externa	2 días	lun 01/04/19	mar 02/04/19	10	**Sistemas**
12		**Disseny xarxa emmagatzematge**	**4 días**	**mié 03/04/19**	**lun 08/04/19**	**9**	Sistemas
13		- privada	2 días	mié 03/04/19	jue 04/04/19		**Sistemas**
14		- publica	2 días	vie 05/04/19	lun 08/04/19	13	**Sistemas**
15		**Analisi de maquinari**	**4 días**	**mar 09/04/19**	**vie 12/04/19**	**14**	Sistemas
16		- maquinari xarxa	2 días	mar 09/04/19	mié 10/04/19		**Sistemas**
17		- maquinari hosting	2 días	jue 11/04/19	vie 12/04/19	16	Sistemas;Distribuidor

- Figura 1: Planning Fase 1- Disseny de xarxes.



- Figura 2: Gantt Fase 1- Disseny de xarxes.

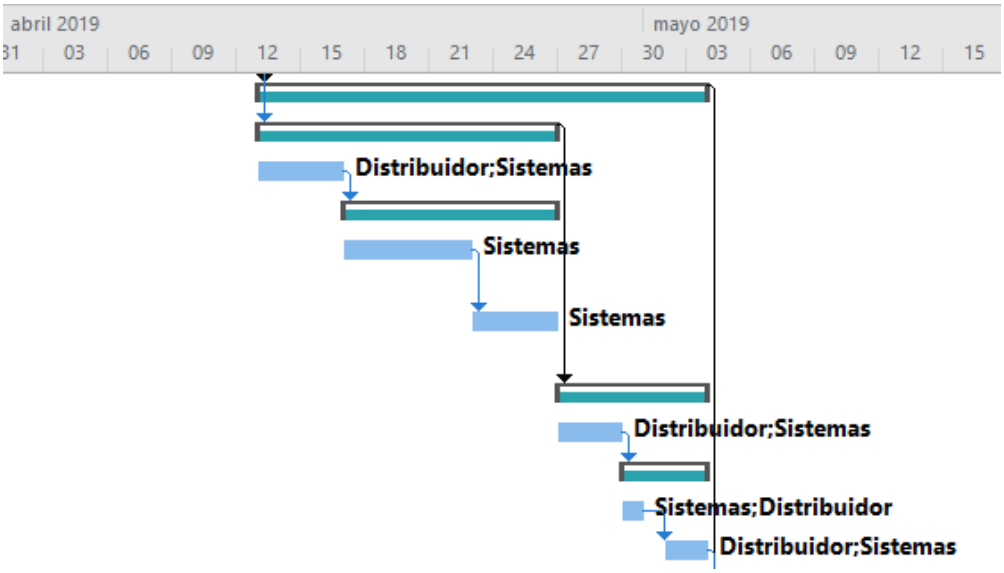
Fase 2 - Disseny sistema de hosting

A la segona fase es portarà a terme el disseny de l'estructura de hosting que serà allotjada als CPD.

En una primera part, s'haurà de dissenyar l'estructura d'emmagatzematge tant pel que fa a maquinari amb tres dies de duració, com de programari on, a més haurem d'analitzar dos sistemes diferents i triar un d'ells per a la implantació amb una duració de nu dies.
A la segona part, haurem de dissenyar el sistema de virtualització que anirà associat a l'estructura d'emmagatzematge anteriorment dissenyada. Igual que en l'apartat anterior, hi haurà una part de maquinari amb dos dies de duració i una part de programari on també haurem d'analitzar dos sistemes diferents per finalment triar aquell més adient. Aquesta segona part té associada una duració de quatre dies.

	Modo de	Nombre de tarea	Duración	Comienzo	Fin	Predecesor	Nombres de los recursos
18		**Fase 2 - Disseny sistema de hosting**	**18 días**	**sáb 13/04/19**	**vie 03/05/19**	**15**	
19		**Disseny sistema emmagatzematge**	**12 días**	**sáb 13/04/19**	**vie 26/04/19**	**17**	
20		- maquinari	3 días	sáb 13/04/19	mar 16/04/19		Distribuidor;Sistemas
21		**- programari**	**9 días**	**mié 17/04/19**	**vie 26/04/19**	**20**	
22		- Analisi sistema emmagatzematge ceph	5 días	mié 17/04/19	lun 22/04/19		Sistemas
23		- Analisi sistema emmagatzematge iSCSI	4 días	mar 23/04/19	vie 26/04/19	22	Sistemas
24		**Disseny sistema virtualització**	**6 días**	**sáb 27/04/19**	**vie 03/05/19**	**19**	
25		- maquinari	2 días	sáb 27/04/19	lun 29/04/19		Distribuidor;Sistemas
26		**- programari**	**4 días**	**mar 30/04/19**	**vie 03/05/19**	**25**	
27		- Vmware ESXi	1 día	mar 30/04/19	mar 30/04/19		Sistemas;Distribuidor
28		- OpenStack	2 días	jue 02/05/19	vie 03/05/19	27	Distribuidor;Sistemas

- Figura 3: Planning Fase 2- Disseny sistema hosting.



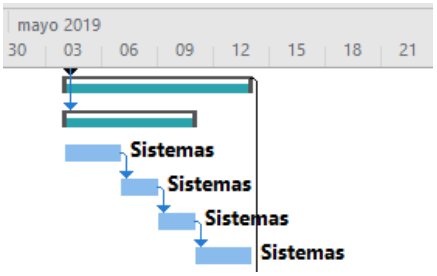
- Figura 4: Gantt Fase 2- Disseny sistema hosting.

Fase 3 - Sistemes de seguretat i protecció

A la fase tres del projecte es farà una anàlisi dels diferents programaris per a les tasques de còpia de seguretat, monitoratge i accés i el disseny de les polítiques associades a cada servei. Aquesta fase no hauria de durar més de 8 dies de cost.

	Modo de	Nombre de tarea	Duración	Comienzo	Fin	Predecesor	Nombres de los recursos
29		**Fase 3 - Sistemes de seguretat i protecció**	**8 días**	**sáb 04/05/19**	**lun 13/05/19**	**18**	
30		**- Analisi de programari**	**6 días**	**sáb 04/05/19**	**vie 10/05/19**	**28**	
31		- backup	2 días	sáb 04/05/19	lun 06/05/19		Sistemas
32		- monitoratge	2 días	mar 07/05/19	mié 08/05/19	31	Sistemas
33		- Accés	2 días	jue 09/05/19	vie 10/05/19	32	Sistemas
34		- Disseny de polítiques	2 días	sáb 11/05/19	lun 13/05/19	33	Sistemas

- Figura 5: Planning Fase 3- Sistemes de seguretat i protecció.



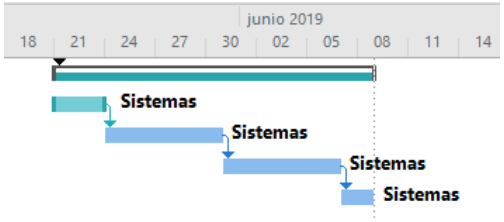
- Figura 6: Gantt Fase 3- Sistemes de seguretat i protecció.

Fase 4 – Implantació

Per últim, tindrem una última fase en la qual es dissenyaran els diferents protocols necessaris per a la implantació del projecte. Aspectes com la planificació de la implantació, decisions de procediment en les diferents estructures a actualitzar o substituir, procediments a portar a terme en les estructures antigues que hauran de conviure amb les estructures noves, la creació de documentació necessària per a la implantació i la gestió de l'estructura una vegada implantada. Aquesta fase, pel fet que es una part important del projecte se l'hi ha assignat una duració de quinze dies.

	Modo de	Nombre de tarea	Duración	Comienzo	Fin	Predecesor	Nombres de los recursos
39		**Fase 5 - Implantació**	**15 días**	**mar 21/05/19**	**sáb 08/06/19**	**35**	
40		- Planificació	3 días	mar 21/05/19	jue 23/05/19		Sistemas
41		- Implantació	5 días	vie 24/05/19	jue 30/05/19	40	Sistemas
42		- Creació de documentació	5 días	vie 31/05/19	jue 06/06/19	41	Sistemas
43		- Posada en marxa	2 días	vie 07/06/19	sáb 08/06/19	42	Sistemas

- Figura 7: Planning Fase 4- Implantació.



- Figura 8: Gantt Fase 4- Implantació.

1.5 Breu sumari de productes obtinguts

Aquest projecte no pretén obtenir un producte físic sinó que intentarà dissenyar un escenari real que es pugui portar a terme en una empresa real per tal d'ampliar i modernitzar les seves infraestructures. Per aquest motiu, una vegada finalitzat el projecte, haurem d'haver obtingut un pla d'empresa per portar a terme en l'empresa a la qual fem referència. Aquest pla d'empresa contindrà una proposta de disseny de la infraestructura que es vol implementar, un possible pressupost d'implantació i la seva planificació.

1.6 Breu descripció dels altres capítols de la memòria

La resta de capítols que es desenvoluparan en aquest projecte tindran a veure amb l'anàlisi de les diferents arquitectures que esdevindran el projecte final, l'anàlisi de requeriments que l'empresa demana, els diferents objectius, la tria de les tecnologies a implementar i la prova pilot en un laboratori virtual d'aquestes.

Els capítols que es desenvoluparan a continuació seran els següents.

Capítol 2. Disseny de xarxes

En el capítol 2, presentarem inicialment els requeriments que ens planteja l'empresa i una vegada establerts dissenyarem les diferents xarxes on s'implantaran les infraestructures per a l'hostatge de servidors virtuals. Com s'ha comentat anteriorment, aquestes xarxes estaran diferenciades per la seva funcionalitat definint així xarxes internes, xarxes públiques per la virtualització, xarxes d'emmagatzematge, xarxes de backup, xarxes privades per a l'empresa i xarxes aïllades per a projectes especials. A més, es definirà la segmentació de xarxa de cadascuna d'elles, el tipus d'IPs que es faran servir i el maquinari proposat per a implementar-les.

Capítol 3. Sistemes d'emmagatzematge

En el capítol 3, inicialment, es presentaran els requeriments que promouen aquest canvi i una vegada establerts es dissenyaran els sistemes d'emmagatzematge tant per a la infraestructura de virtualització com per a la de backup. A més, es farà una breu introducció amb la infraestructura actual i es posaran en relleu els aspectes que ens porten a sospesar un canvi de tecnologia. Una vegada presentat el sistema

en producció es proposarà la infraestructura d'emmagatzematge amb sistema Ceph per a la seva implantació. Per últim, es desenvoluparà un petit escenari en laboratori per tal de mostrar el funcionament del sistema d'emmagatzemament basat en Ceph.

Capítol 4. Sistemes de virtualització

En el capítol 4, igual que en el capítol anterior, inicialment es mostraran els requeriments que ens planteja l'empresa per a la implantació del sistema de virtualització i que ens ha de portar a la tria dels diferents sistemes a implementar. Seguidament es desenvoluparà la instal·lació i configuració del sistema OpenStack pel que fa al seu funcionament, maquinari necessari o la seva implantació sobre un sistema d'emmagatzematge Ceph.
Finalment i sobre el laboratori generat per a Ceph, es desenvoluparà la instal·lació d'un sistema OpenStack sobre el sistema d'emmagatzemament Ceph implementat anteriorment.

Capítol 5. Sistemes de seguretat

En el capítol 5, es definiran els diferents sistemes de seguretat que haurem d'implantar i/o actualitzar per tal de protegir aspectes de la infraestructura com són el maquinari, les dades, la privacitat i per obtenir informació del funcionament de tots ells en temps real tant a escala informativa com alertes en cas d'incidència. En cadascun dels apartats, tallafocs, backups, monitoratge, sistemes d'alertes i SAIs, es farà una breu introducció de l'àrea en la qual estem treballant i es proposaran programari i/o maquinari per a gestionar-les.

Capítol 6. Implantació

En el capítol 6, es proposara una sèrie de passos a seguir per a portar a terme la implantació del sistema proposat. Primerament amb una planificació temporal aproximada, després amb la proposta d'implementació on es detallaran quin o quins sistemes s'hauran de crear, actualitzar i/o mantenir. Per últim, en el cas dels sistemes que s'hagin d'actualitzar i/o mantenir, es proposaran diferents mètodes d'actuació per a la migració als sistemes nous o de manteniment d'aquells sistemes que no es puguin migrar.

2. Disseny de xarxes

Les xarxes que hem de definir en aquest capítol estan basades en les xarxes Ethernet i la informació referent a aquest tipus de xarxes es pot trobar a l'annex I.

2.1 Requeriments de les diferents xarxes

Quan es va començar a parlar a l'empresa sobre un possible creixement en el nombre de clients, es va fer una anàlisi del que comportaria aquest creixement en aspectes com la capacitat d'allotjament, la capacitat d'escalament, quin tipus d'adreçament seria més idoni i el que era més important, quines de les xarxes actuals es podrien reutilitzar o mantenir i quines no. A més, es va plantejar que seria necessari, no només abordar aquest creixement en la xarxa pública que clarament és la més afectada, sinó fer un replantejament de tota l'estructura de xarxes de l'empresa.

2.1.1 Xarxa interna

Actualment, la xarxa privada de l'empresa ve heretada dels inicis quan hi havia poca gent treballant i els departaments no eren encara ben definits. Aquesta infraestructura contempla una única xarxa global amb una classe C d'IPs internes, com ara pot ser 192.168.1.0/24 serveis per comunicar tots els departaments de l'empresa amb tots els servidors interns. A més, existeix una única sortida a internet per tota la xarxa amb permisos especials per a accedir a les xarxes públiques de l'empresa. Aquest sistema és molt senzill de gestionar quan hi ha pocs treballadors i tothom fa una mica de tot, però la situació actual ha canviat i es vol aprofitar aquest projecte per a remodelar la xarxa interna i ferla més operativa i segura.

Els requeriments que es plantegen de cara a l'actualització de les xarxes internes són:

- Aïllament dels departaments a escala de xarxa.
- Aïllament de la secció de servidors interns (DMZ[8]) per departament.
- Diferenciació de les sortides a internet per departament.
- Creació d'una xarxa WIFI[9] per a dispositius de treballadors i clients.

Per últim, en l'àmbit físic, s'actualitzarà el maquinari necessari per passar tota la xarxa a Gigabit. Es planteja l'opció d'aprofitar el canvi per implementar la xarxa directament a 10Gbps, però atès que no hi ha serveis interns de gran consum de tràfic i a més, les sortides a internet de l'oficina no arribaran a 100Mbps, no veu com a necessària la inversió per a aquest sistema. Això comportarà canvi de maquinari mínim de xarxa com ara commutadors i canvi de cablejat a cable UTP[10] de categoria 6 instal·lats ja en la majoria de casos.

Xarxa interna Cablejada

Tipus de maquinari	Característiques rellevants	Espai	Unitats
Commutador principal	1Gbps, Capa 3, 24ports, Administrable	Sala servidors	1
Commutador Planta 1	1Gbps, VLANs, 48ports, Administrable, PoE	Sala servidors	1
Commutador Planta 2	1Gbps, VLANs, 24ports, Administrable, PoE	Sala servidors	1
Encaminador porta d'enllaç	Encaminador i Servidor VPN (IPSEC)	Sala servidors	1
Encaminadors VPN (IPSEC)	Encaminadors delegacions per a VPN (IPSEC)	Delegacions	2
Cablejat UTP	Cablejat UTP cat6	-	-

- Taula 1: Requeriments de la xarxa interna cablejada .

La idea és fer servir un commutador principal de capa 3 per gestionar les VLANs configurades als commutadors de planta i així alliberar d'aquesta funció a l'encaminador. Per a gestionar la connectivitat de les dues delegacions a la xarxa interna es crearan dos túnels Ipsec per a integrar les xarxes d'aquestes. Actualment ja estan creats aquests dos túnels però s'actualitzarà el maquinari. A més, als commutadors de planta, es faran servir dispositius amb capacitat de PoE[15] per a donar alimentació elèctrica als dispositius de punt d'accés de la xarxa WIFI.

Xarxa WIFI

Tipus de maquinari	Característiques rellevants	Espai	Unitats
Punt d'accés wifi	Doble banda, accés invitat i usuari	Centre plantes	2
Cablejat UTP	Cablejat UTP cat6	-	-

- Taula 2: Requeriments de la xarxa interna WiFi .

2.1.2 Xarxa Pública

Un primer tema a tenir en compte era la previsió de creixement. Aquesta es va establir entorn al 100% en els pròxims anys, per la qual cosa s'hauran d'augmentar el nombre d'adreces IP públiques que es gestionen actualment. En aquest aspecte, es parteix de la següent situació a la xarxa pública:

Datacenter	Nº Classes C IPv4	Nº Ips IPv4	Classe IPv6	Nº Ips IPv6	Ample de banda
CPD1	6	1536	-	-	100 Mbps
CPD2	2	512	1 ::/48	65536 xarxes ::/64	100 Mbps

Situació després de la implantació del projecte					
CPD1	6	1536	-	-	200 Mbps
CPD2	4	1024	1 ::/48	65536 xarxes ::/64	200 Mbps
CPD3	4	1024	1 ::/48	65536 xarxes ::/64	1 Gbps

- Taula 3: segmentació xarxa pública .

Com es pot veure a la taula, es planteja l'ampliació d'IPs del segon CPD[11] a 4 classes C IPv4, el que duplicaria el nombre d'IPs. A més, també es contractarà un nou datacenter amb quatre classes C IPv4 que aportarà 1024 IPs més. A més, en aquest CPD3 es contractarà un rang IPv6 ::/48 amb la idea d'implementar-ho de forma permanent tant al CPD2 com al CPD3 per a la majoria d'allotjaments. Aquestes dues xarxes IPv6 aportaran a cada CDP 65536 xarxes del tipus ::/64 de les que cada una aportaran uns 18,5 trilions d'IPs. Per altra banda, s'augmentarà l'amplada de banda al CPD1 fins als 200Mbps i als CPD2 i 3 es contractarà una amplada de banda d'un Gbps d'inici. Amb aquesta ampliació, hauria de ser suficient per poder gestionar el creixement previst.

Un altre aspecte relacionat seria la gestió de la segmentació dels rangs d'IPs en les diferents xarxes. Una de les coses que es panteixa és la de situar el creixement en els CPDs 2 i 3 mantenint el CPD1 amb la mateixa estructura encara que fent una reestructuració de les xarxes per tal d'optimitzar el seu ús, ja que actualment es troben molt segmentades. Al CPD2 s'hauran d'analitzar les xarxes que hi ha configurades, ja que les dues classes C que hi ha actualment ja estan segmentades i s'haurà de veure si es poden unificar xarxes o no.

En conseqüència, als CPDs 2 i 3 es requeriran xarxes àmplies ja que és en aquests on se situaran els sistemes de cloud que hauran d'allotjar els nous clients.

Per aquesta raó, partint de la segmentació actual,

Datacenter	Nº Classes C IPv4	Xarxes /24	Xarxes /25	Xarxes /26	Xarxes /27
CPD2	2	-	2x(128 IPs)	2x(64 IPs)	4x(32 IPs)
CPD3	0	-	-	-	-

- Taula 4: Segmentació xarxa pública – Estat actual .

es planteja la següent segmentació per als dos CPDs.

Datacenter	Nº Classes C IPv4	Xarxes /23	Xarxes /24	Xarxes /25	Xarxes /26	Xarxes /27
CPD2	4	1x(512 IPs)	1x(256 IPs)	-	2x(64 IPs)	4x(32 IPs)
CPD3	4	1x(512 IPs)	2x(256 IPs)	-	-	-

- Taula 5: Segmentació xarxa pública – Proposta.

Amb aquesta segmentació proposada es pretén tenir dues xarxes principals on s'allotjarien els dos sistemes de cloud, els quals haurien de poder allotjar 500 servidors virtuals independents per clients cadascun. La resta de xarxes anirien destinades a servidors privats en els casos de les xarxes /27 i per a projectes especials en els casos de les xarxes /26. En el cas de les xarxes /24 dels dos CPDs es destinaran als servidors d'allotjament compartit on el requeriment d'IPs públiques és menor donat el seu caire compartit.

Per últim, pel que fa a maquinari es requeriran un parell d'encaminadors, commutadors nous per a les xarxes noves i cablejat de categoria 6 per mantenir la velocitat de Gigabit que ja té la resta de xarxes actuals.

Tipus de maquinari	Característiques rellevants	Espai	Unitats
Encaminador	1Gbps, Administrable	CPD2 i CPD3	4
Commutadors xarxes /23	1Gbps, 48ports, Administrable, SFP+	CPD2 i CPD3	2
Commutadors xarxes /24	1Gbps, 24ports, Administrable, SFP+	CPD2 i CPD3	3
Cablejat UTP	Cablejat UTP cat6	CPD2 i CPD3	-

.- Taula 6: Requeriments xarxa pública

2.1.3 Xarxa d'emmagatzematge

En la planificació del projecte es va parlar de crear una secció aïllada per a l'emmagatzemament tant si havia de ser de cabines amb connexions iSCSI[12] o amb sistema CEPH[13]. Aquesta secció/xarxa havia d'estar separada de la part de virtualització per tal d'obtenir una escalabilitat que en aquest moment no existia. Per aquest motiu es planteja una xarxa d'emmagatzemament nova. Inicialment es va dir de fer servir IPs públiques però després d'analitzar-ho, es veu que l'únic benefici era de connectivitat cap a l'exterior i aquest no era un requeriment especialment útil pel nostre cas enfront del cost d'IPs que suposaria. És per això que es decideix crear aquesta xarxa d'emmagatze-mament amb IPs privades, encara que també existeix l'opció d'implementar-la sota IPv6 pel fet que el nombre d'IPs en aquest sistema deixaria de ser un problema. Tot i això, sigui en IPv4 privades o IPv6, la xarxa d'emmagatze-matge ha d'estar aïllada de la part pública i en conseqüència d'internet.

Finalment, pel que fa als requeriments de maquinari, al ser una xarxa nova es proposa la compra de commutadors administrables nous amb capacitat per a sostenir xarxes 10GBASE-T i cablejat UTP de categoria 7.

Tipus de maquinari	Característiques rellevants	Espai	Unitats
Commutadors	10Gbps, 48ports, Administrable, SFP+	CPD2 i CPD3	2
Cablejat UTP	Cablejat UTP cat7	CPD2 i CPD3	-

- Taula 7: Requeriments xarxa emmagatzematge.

2.1.4 Xarxa de backup

Un altre requeriment del projecte és el redisseny del sistema de còpies de seguretat existent actualment i que consisteix en una sèrie de servidors individuals en el que es configuren les còpies d'un nombre limitat de servidors en cadascun d'ells. La proposta que es fa consisteix a unificar la xarxa de backup per CPD de manera que ens permeti tenir un sistema distribuït de còpies amb el que aconseguiríem un major control i dinamisme a l'hora de gestionar les còpies.

Pel que fa a la xarxa, se seguirà el mateix criteri que per les xarxes d'emmagatzemament i es dissenyarà amb adreces privades d'IPv4 o amb IPv6. A més, les dues xarxes, emmagatzemament i backup seran visibles per tal de gestionar les copies directament des del sistema d'emmagatzemament.

Només ens quedaria la part de maquinari que també seguirà els mateixos requeriments que la xarxa d'emmagatzematge amb commutadors administrables amb tecnologia 10GBASE-T i cablejat UTP de categoria 7.

Tipus de maquinari	Característiques rellevants	Espai	Unitats
Commutadors	10Gbps, 48ports, Administrable, SFP+	CPD2 i CPD3	2
Cablejat UTP	Cablejat UTP cat7	CPD2 i CPD3	-

- Taula 8: Requeriments xarxa backup.

2.2 Disseny final de les diferents xarxes

Una vegada introduïts els requeriments que es plantegen en aquest projecte es passa a detallar la proposta de disseny per a les diferents xarxes partint d'una posició inicial de producció en el cas de les xarxes interna i pública, i de creació en el cas de les xarxes d'emmagatzemament i backup.

Un aspecte a tenir en compte és que les IPs indicades en els esquemes són totalment fictícies, ja que per una part, per privacitat de dades no està permès el fet de publicar les reals ja existents i per altra part les adreces de les xarxes noves no estan contractades encara i per tal motiu no podem saber quins rangs ens assignarien.

2.2.1 Xarxa interna

El disseny de la xarxa interna començaria per veure gràficament l'estat inicial de la xarxa. Actualment, com s'ha dit abans, la xarxa interna consisteix en una classe C completa on es connecten tots els PCs de l'empresa i els servidors interns. Tots ells es connecten al mateix nivell i tenen accés a qualsevol indret de la xarxa.

Un dels primers requeriments parla de la necessitat d'aïllar els diferents departaments entre ells. Per aquest motiu, es planteja la creació d'VLANs[14] de manera que cada departament només veurà la part de la xarxa en la qual està situat. Aquesta segmentació aïllarà els Pcs de cada departament de la resta. Fet que aportarà seguretat d'accés, ja que, per exemple, un PC del departament d'aplicacions no podrà veure un PC del departament d'administració i un d'administració no es veurà afectat per proves de xarxa fetes al departament tècnic. A més, aquest aïllament alleujarà el tràfic de la xarxa, ja que cada VLAN tindrà el seu propi domini de difusió i per tant les trames de broadcast disminuiran, ja que es limitaran al seu domini de difusió. Això farà que l'amplada de banda consumit fins a les hores per trames de broadcast disminueixi notablement i per tant augmentarà el rendiment de la xarxa.

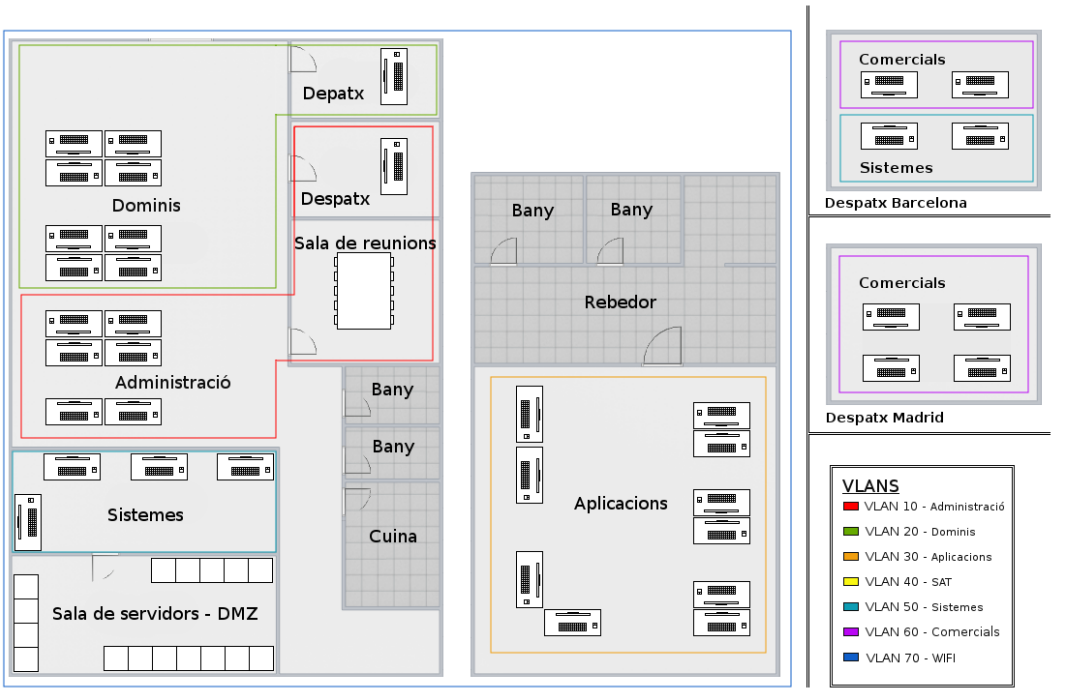
La segmentació proposada es la següent:

Nom VLAN	Descripció	Domini de difusió	host	Max. hosts
10	VLAN Gerència - Administració	192.168.10.0/24	4	254
20	VLAN Dominis	192.168.20.0/24	12	254
30	VLAN Aplicacions	192.168.30.0/24	6	254
40	VLAN SAT	192.168.40.0/24	3	254
50	VLAN Sistemes	192.168.50.0/24	5	254
60	VLAN Comercials	192.168.60.0/24	5	254
70	VLAN WIFI	192.168.70.0/24	Dinàmic	254

- Taula 9: VLANs xarxa interna.

Amb aquesta segmentació es dòna suport a un gran nombre de dispositius a cada xarxa en previsió de futures ampliacions de dispositius, ja que s'assignarà una classe C sencera per a cada departament encara que es podria ajustar a rangs /27 amb 32 IPs per a 30 hosts. Per altra banda es crea una VLAN específica per a dispositius wifi per aïllar-los de les xarxes de treball.

Segons aquestes dades es planteja un esquema de xarxa similar a aquest.



- Figura 9: Esquema oficina xarxa interna.

Pel que fa a la sortida a internet de les VLANS, hauran tres sortides diferents depenent de la VLAN des de la que es connecti. La VLAN de wifi es farà sortir per una IP pública diferent de les altres per tal d'aïllar tant els permisos d'accés com

l'amplada de banda. La resta sortiran per una IP pública comuna amb permisos d'accés limitat a les xarxes públiques de l'empresa excepte les VLANs de SAT i sistemes que sortiran per una IP pública amb permisos per poder administrar els servidors de les xarxes públiques de l'empresa.

La part d'actualització del maquinari es basaria en el canvi de commutadors, manteniment i/o actualització del cablejat necessari i la creació de la infraestructura necessària per implantació de la xarxa WIFI.

Es proposa el següent maquinari:

Xarxa LAN cablejada (+VLANS) (Obert a canvis de marca / model)

Tipus de maquinari	Marca / Model	Unitats
Commutador principal	TP-Link / T3700G-28TQ commutador administrat Gigabit L3 Apilable, 28 ports	1
Commutador Planta 1	Linksys / LGS552P Commutador Gigabit PoE+ administrat de 48 ports amb 2 ports SFP combinats i 2 ports SFP+	1
Commutador Planta 2	Linksys / LGS528 Commutador Gigabit PoE+ administrat de 28 ports amb 2 ports SFP combinats	1
Encaminador porta d'enllaç / Tallafocs	Sonicwall TZ400	1
Encaminadors VPN (IPSEC) / Tallafocs	Sonicwall TZ300	2
Cablejat UTP	Cablejat UTP cat6	-

- Taula 10: Maquinari xarxa interna cablejada.

Xarxa LAN WIFI (Obert a canvis de marca / model)

Tipus de maquinari	Marca / Model	Unitats
Punt d'accés WIFI	tp-link / CAP1200 Punt d'accés Gigabit Inalámbric de Doble Banda AC1200 amb Muntatge de sostre	2
Cablejat UTP	Cablejat UTP cat6	-

- Taula 11: Maquinari xarxa interna WIFI.

2.2.2 Xarxa pública

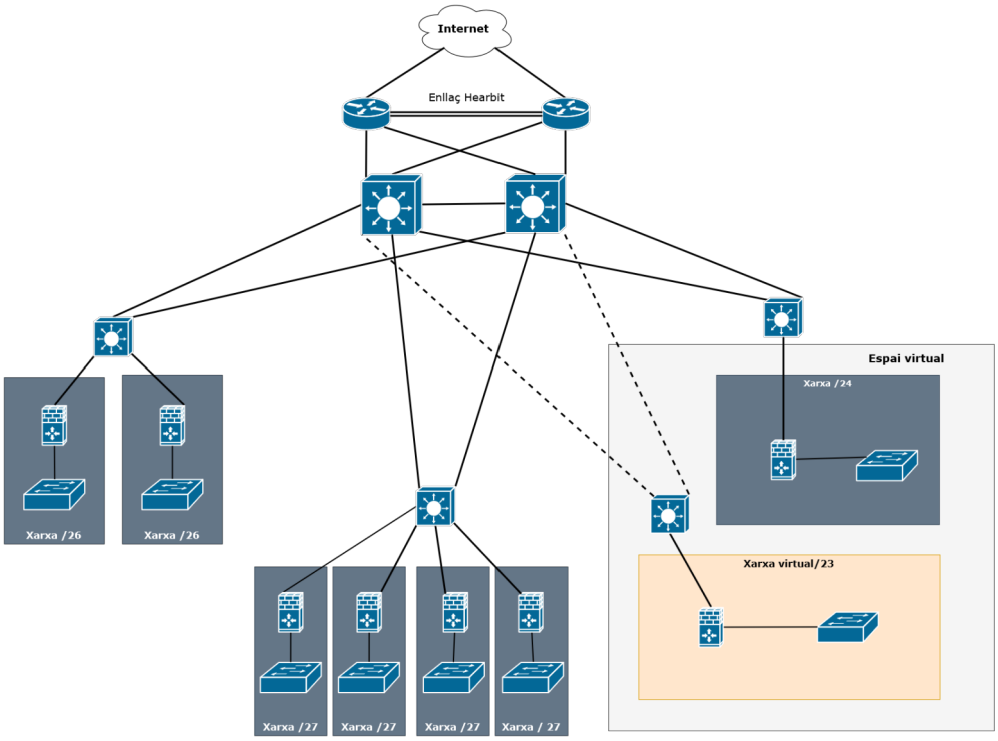
El disseny de la xarxa pública es basarà a definir les diferents xarxes obtingudes de la segmentació ja feta a l'apartat de requeriments i triar un maquinari òptim per a la seva implementació. Segons aquests requeriments, s'ha decidit que tot el treball associat al projecte se centrarà en l'actualització i creació de les xarxes dels CPDs 2 i 3 respectivament, deixant de banda el CPD1 que no es modificarà en aquest projecte.

CPD2 - IPv4

Al CPD2 es proposa inicialment el canvi de l'encaminador actual per dos nous de les marques abans comentades. Els encaminadors es configuraran en mode mestre-esclau de manera que un estarà actiu i l'altre restarà a l'espera amb la mateixa configuració i en cas de caiguda del primer agafarà el comandament perquè el sistema continuí funcionant. A més, es crearà una xarxa /23 per a l'allotjament del sistema de virtualització al qual s'instal·laran una sèrie de servidors físics que donaran servei a múltiples servidors virtuals. La xarxa permetrà l'assignació de 510 adreces IPv4. Una proposta que s'ha de valorar és l'aprofitament de la xarxa /25 per a l'assignació d'adreces dels servidors físics i per tant l'aïllament dels servidors virtuals en la xarxa /23 de tal manera que els programaris de virtualització gestionin la xarxa /23 de manera virtual.

No es contempla contractar una sortida a internet redundant, ja que aquest servei està implementat com servei afegit per part del subministrador d'accés del CPD (subcontractat) que ens aporta una sortida a internet redundant amb cinc proveïdors diferents. D'aquesta manera el CPD ens subministra un punt d'accés a la seva xarxa amb sortida a internet per les portes d'enllaç redundant comentades anteriorment. Per la nostra part s'implementarà un sistema d'encaminadors en mode master-slave per controlar la possible caiguda de l'encaminador principal.

Un possible esquema de la xarxa al CPD2 seria el següent.



- Figura 10: Esquema de la xarxa al CPD2.

Com es pot veure a l'esquema, a part de la replicació dels encaminadors, es replica l'estructura de commutadors principals per evitar la caiguda completa de la xarxa en cas de caiguda del commutador principal.

En un segon nivell s'instal·laran commutadors per agrupar xarxes amb connexió als dos commutadors principals per mantenir la redundància de xarxa. Per altra banda els diferents tallafocs de cada xarxa tindrà una única connexió amb el commutador del segon nivell assignat. En aquest nivell es perd la redundància de xarxa però es minimitza la caiguda de cada commutador i/o tallafocs amb fonts d'alimentació redundants. Per altra banda es mantindrà al CPD maquinari de les mateixes característiques per tal de fer el canvi en cas d'avaria en el menor temps possible.

En la part de l'adreçament es farà l'esquema amb quatre classes que no seran les reals, doncs s'haurà d'esperar a la seva contractació per la seva assignació.

Les quatre classes amb les quals es faran les xarxes són:

(no s'agafa la primera classe C x.2.0.0/24 perquè es vegi bé el subnetting en la xarxa /23)

- `x.2.1.0/24`
- `x.2.2.0/24`
- `x.2.3.0/24`
- `x.2.4.0/24`

Recordem la segmentació triada per al CPD2.

Datacenter	Nº Classes C IPv4	Xarxes /23	Xarxes /24	Xarxes /25	Xarxes /26	Xarxes /27
CPD2	4	1x(512 IPs)	1x(256 IPs)	-	2x(64 IPs)	4x(32 IPs)

- Taula 12: Segmentació al CPD2.

(*Es pot veure la metodologia per la segmentació de les xarxes IPv4 a l'annex X*)

Segons aquesta segmentació les xarxes resultants seran les següents:

Xarxa		Rang d'IPs	Hosts utils
x.2.1.0/24	-->	x.2.1.0 – x.2.1.255 (256 IPs)	254
x.2.2.0/23	-->	x.2.2.0 – x.2.3.255 (512 IPs)	510
x.2.4.0/26	-->	x.2.4.0 – x.2.4.63 (64 IPs)	62
x.2.4.64/26	-->	x.2.4.64 – x.2.4.127 (64 IPs)	62
x.2.4.128/27	-->	x.2.4.128 – x.2.4.159 (32 IPs)	30
x.2.4.160/27	-->	x.2.4.160 – x.2.4.191 (32 IPs)	30
x.2.4.192/27	-->	x.2.4.192 – x.2.4.223 (32 IPs)	30
x.2.4.224/27	-->	x.2.4.224 – x.2.4.255 (32 IPs)	30

- Taula 13: Xarxes resultants al CPD2.

CPD2 -IPv6

Al marge de la xarxa d'IPv4, es proposa implementar la xarxa IPv6 paral·lelament a aquesta. Per a aquesta implementació es parteix d'una classe /48 d'IPv6. Com en el cas d'IPv4, les IPs no seran reals a l'espera de la seva contractació, així que farem servir la següent xarxa:

Xarxa		Xarxa		host
x:y:z::/48	-->	000x:000y:000z	:	0000:0000:0000:0000:0000

- Taula 14: Xarxa IPv6 al CPD2.

(*Es pot veure la metodologia per la segmentació de les xarxes IPv6 a l'annex XI*)

Per tal de segmentar la xarxa, i tal com s'ha explicat a l'annex XI, la segmentació triada serà la següent.

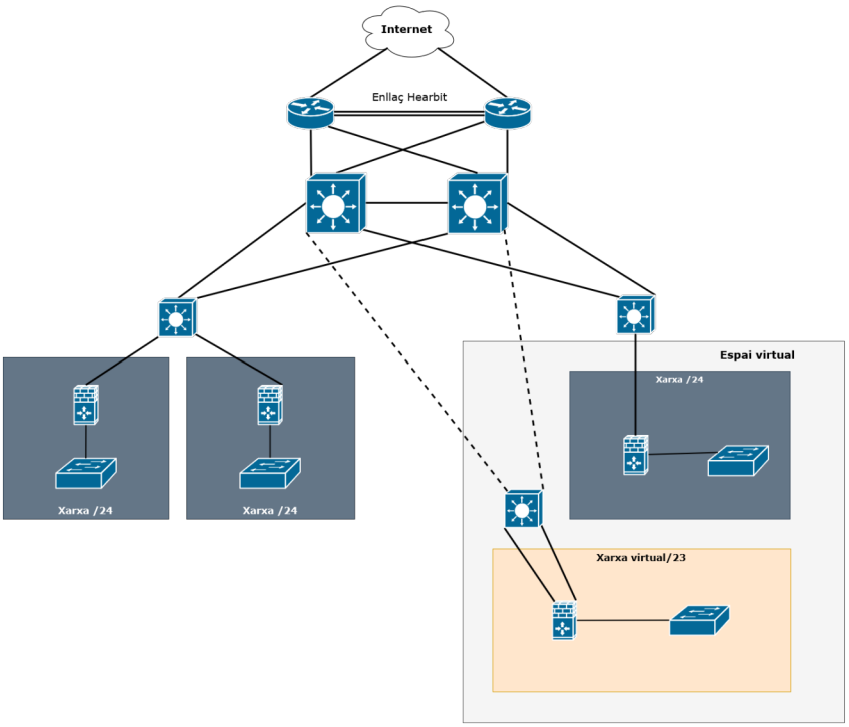
Xarxa		xarxa	Subxarxa	Hosts
x:y:z::0:0/112	-->	000x:000y:000z:0000:0000:0000	:0000	:0000
x:y:z::1:0/112	-->	000x:000y:000z:0000:0000:0000	:0001	:0000
x:y:z::2:0/112	-->	000x:000y:000z:0000:0000:0000	:0002	:0000
x:y:z::3:0/112	-->	000x:000y:000z:0000:0000:0000	:0003	:0000
x:y:z::4:0/112	-->	000x:000y:000z:0000:0000:0000	:0004	:0000
x:y:z::5:0/112	-->	000x:000y:000z:0000:0000:0000	:0005	:0000
x:y:z::6:0/112	-->	000x:000y:000z:0000:0000:0000	:0006	:0000
x:y:z::7:0/112	-->	000x:000y:000z:0000:0000:0000	:0007	:0000

- Taula 15: Xarxes IPv6 resultants al CPD2.

Aquesta segmentació ens aportaria xarxes més petites amb 65536 IPs per xarxa, però totalment suficients per a les xarxes que tenim.

CPD3 – IPv4

Les xarxes del CPD3 no varien molt de les del CPD2 i la segmentació serà encara més fàcil. Es mantindrà la mateixa estructura per als nivells principals de la xarxa, encaminadors i commutadors principals, i s'implementen les tres xarxes definides al disseny.



- Figura 11: Esquema de la xarxa al CPD3.

Datacenter	Nº Classes C IPv4	Xarxes /23	Xarxes /24	Xarxes /25	Xarxes /26	Xarxes /27
CPD3	4	1x(512 IPs)	2x(256 IPs)	-	-	-

- Taula 16: Segmentació al CPD3.

Fent servir el mateix procés que al CPD2 definirem les Classes que farem servir (no seran reals) i en aquest cas l'única acció a fer es la unificació de dues de les classes C per obtenir una xarxa /23.

Les IPs assignades són:

- `x.3.1.0/24`
- `x.3.2.0/24`
- `x.3.3.0/24`
- `x.3.4.0/24`

Aprofitant el subnetting explicat a l'annex X, obtenim les xarxes següents:

Xarxa		Rang d'IPs	Hosts utils
`x.3.1.0/24`	`-->`	`x.3.1.0 – x.3.1.255 (256 IPs)`	254
`x.3.2.0/23`	`-->`	`x.3.2.0 – x.3.3.255 (512 IPs)`	510
`x.3.4.0/24`	`-->`	`x.3.4.0 – x.3.4.255 (256 IPs)`	254

- Taula 17: Xarxes resultants al CPD3.

Aquestes xarxes ens aportaran 1018 IPs útils de les 1024 possibles on tindrem 510 IPs útils per a la xarxa de virtualització, 256 IPs útils per a servidors de l'empresa i 256 IPs útils per a projectes especials.

CPD3 - IPv6

Com també passa amb les xarxes d'IPv4, s'aprofitarà la segmentació realitzada a l'annex XI per implementar en el CPD3. D'aquesta manera, si la xarxa assignada fos aquesta:

Xarxa		Xarxa		host
p:q:r::/48	`-->`	`000p:000q:000r`	:	`0000:0000:0000:0000:0000`

- Taula 18: Xarxa IPv6 al CPD3.

S'implementarà la següent segmentació:

Xarxa		xarxa	Subxarxa	Hosts
`p:q:r::0:0/112`	`-->`	`000p:000q:000r:0000:0000:0000`	`:0000`	`:0000`
`p:q:r::1:0/112`	`-->`	`000p:000q:000r:0000:0000:0000`	`:0001`	`:0000`
`p:q:r::2:0/112`	`-->`	`000p:000q:000r:0000:0000:0000`	`:0002`	`:0000`

- Taula 19: Xarxes IPv6 resultants al CPD3.

Per últim, el maquinari que es proposa per implementar es basa en dispositius TP-Link encara que queda oberta la possibilitat de canviar per altres marques o models.

CPD2

Tipus de maquinari	Marca / Model	Unitats
Encaminador principal	Cisco / ASR 1001-X	2
Commutadors principals	TP-Link / T3700G-28TQ commutador administrat Gigabit L3 Apilable, 28 ports	5
Commutador secundaris	TP-Link / T2700G-28TQ commutador administrat Gigabit L2 Apilable, 28 ports	7
Tallafocs xarxes /23 - /24	WatchGuard Firebox M570 Firewall	2
Tallafocs xarxes /26 - /27	WatchGuard Firebox M370 Firewall	2
Cablejat UTP	Cablejat UTP cat6	-

- Taula 20: Maquinari xarxa pública CPD2.

CPD3

Tipus de maquinari	Marca / Model	Unitats
Encaminador principal	Cisco / ASR 1001-X	2
Commutadors principals	TP-Link / T3700G-28TQ commutador administrat Gigabit L3 Apilable, 28 ports	4
Commutador secundaris	TP-Link / T2700G-28TQ commutador administrat Gigabit L2 Apilable, 28 ports	4
Tallafocs xarxes /23 - /24	WatchGuard Firebox M570 Firewall	2
Cablejat UTP	Cablejat UTP cat6	-

- Taula 21: Maquinari xarxa pública CPD3.

Com veiem al llistat del maquinari, es proposa la instal·lació d'un encaminador marca *Cisco* i model *ASR 1001-X* compatible amb el que ja s'està fent ser en el CPD1. Aquesta gamma d'encaminadors està dissenyada per a ser la porta d'enllaç de proveïdors de serveis i ofereix serveis bàsics com ara tallafocs, seguretat d'encriptació, traducció NAT o una inspecció dels paquets amb un rendiment alt que ens aporta filtratge QoS. A més, a escala de maquinari incorpora doble font d'alimentació per tenir redundància.
(*Més informació sobre l'encaminador Cisco ASR 1001-X a l'annex V*)

Pel que fa als tallafocs WatchGuard ens ofereixen un rendiment més que suficient tant en mode tallafocs, com en mode UTM, per protegir les diferents xarxes. En el cas de les xarxes /23 i /24 s'implementen dos tallafocs individuals model M570, un per cada xarxa, i per les xarxes /26 i /27 del CPD2 s'implementaran dos tallafocs compartits model M370 on les dues xarxes /26 seran gestionades per un tallafocs i

les 4 /27 per l'altre. S'ha triat aquesta marca per sobre d'altres com ara cisco o sonicwall per una banda pels serveis UTM que ofereix i les característiques com ara la velocitat final una vegada aplicades que són millors que les de sonicwall i per l'altre per comparació de preus que són més competitius que qualsevol de les altres.

2.2.3 Xarxa emmagatzematge

La xarxa d'emmagatzematge es planteja com a una xarxa interna amb connexió amb el maquinari de virtualització mitjançant rangs de IPs internes.

(*Els rangs d'IPs internes per IPv4 i IPv6 es poden trobar a l'Annex XII*)

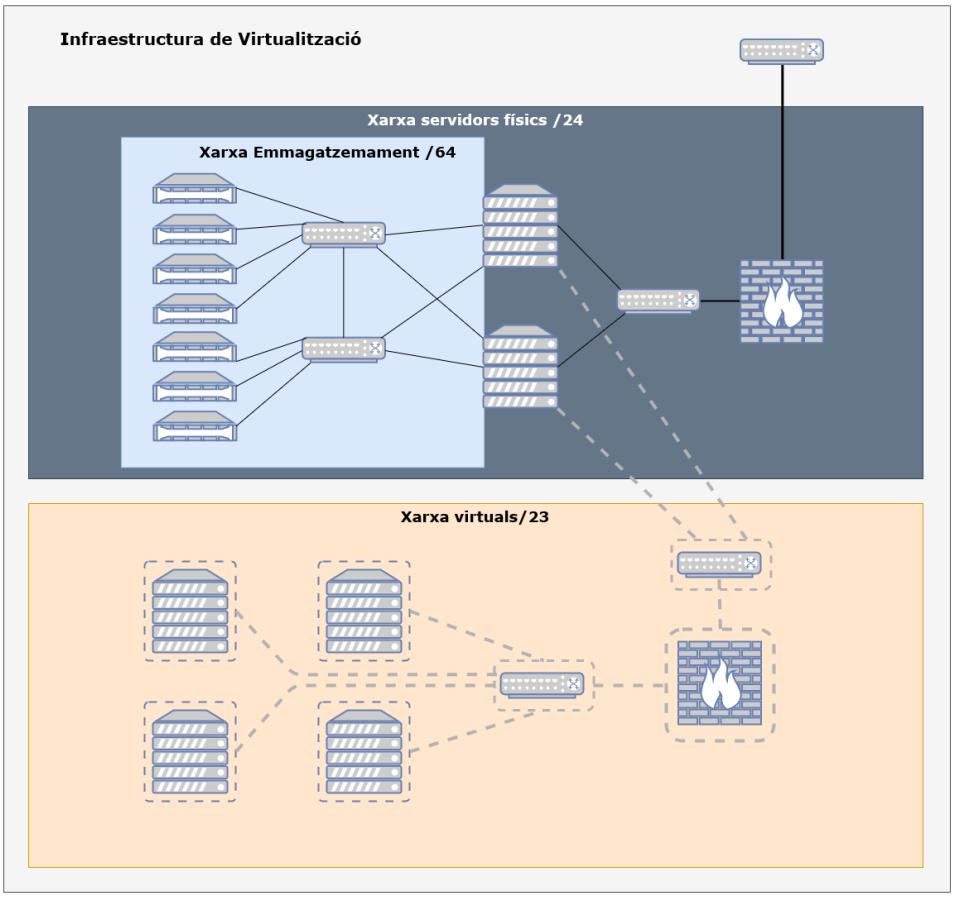
Possible rang triat per la xarxa d'emmagatzemament.

Xarxa /8	prefix	ID global	ID subxarxa	Xarxa /64
fd00::/8	fd	43d7f9fc7b	de4a	fd43:d7f9:fc7b:de4a::/64
Inici de Rang	fd43:d7f9:fc7b:de4a:0:0:0:0			
Fi de Rang	fd43:d7f9:fc7b:de4a:ffff:ffff:ffff:ffff			
Nº de hosts	2^64 IPs (18.446.744.073.709.551.616)			

- Taula 22: Xarxa emmagatzematge IPv6.

Una vegada tenim clar quines IPs es poden configurar a la xarxa d'emmagatzematge s'haurà de decidir si s'implementen IPs d'IPv4 o d'IPv6. Inicialment, com és una xarxa aïllada d'internet es pot perfectament implementar únicament IPs d'IPv6, ja que aquestes ens aporten característiques millorades enfront a l'IPv4 com ara capçaleres més simples que fan més eficient i ràpid el seu processament, paquets IP eficients i extensibles que eviten fragmentació, paquets amb càrrega útil fins a 65536 Bytes, autoconfiguració, etc.
Per tal de connectar amb el programari de virtualització, s'afegiran IPs d'aquesta xarxa als servidors físics.

Per tant s'implantarà la xarxa fd43:d7f9:fc7b:de4a::/64 en la infraestructura d'emmagatzemament següent:



- Figura 12: Esquema infraestructura de virtualització.

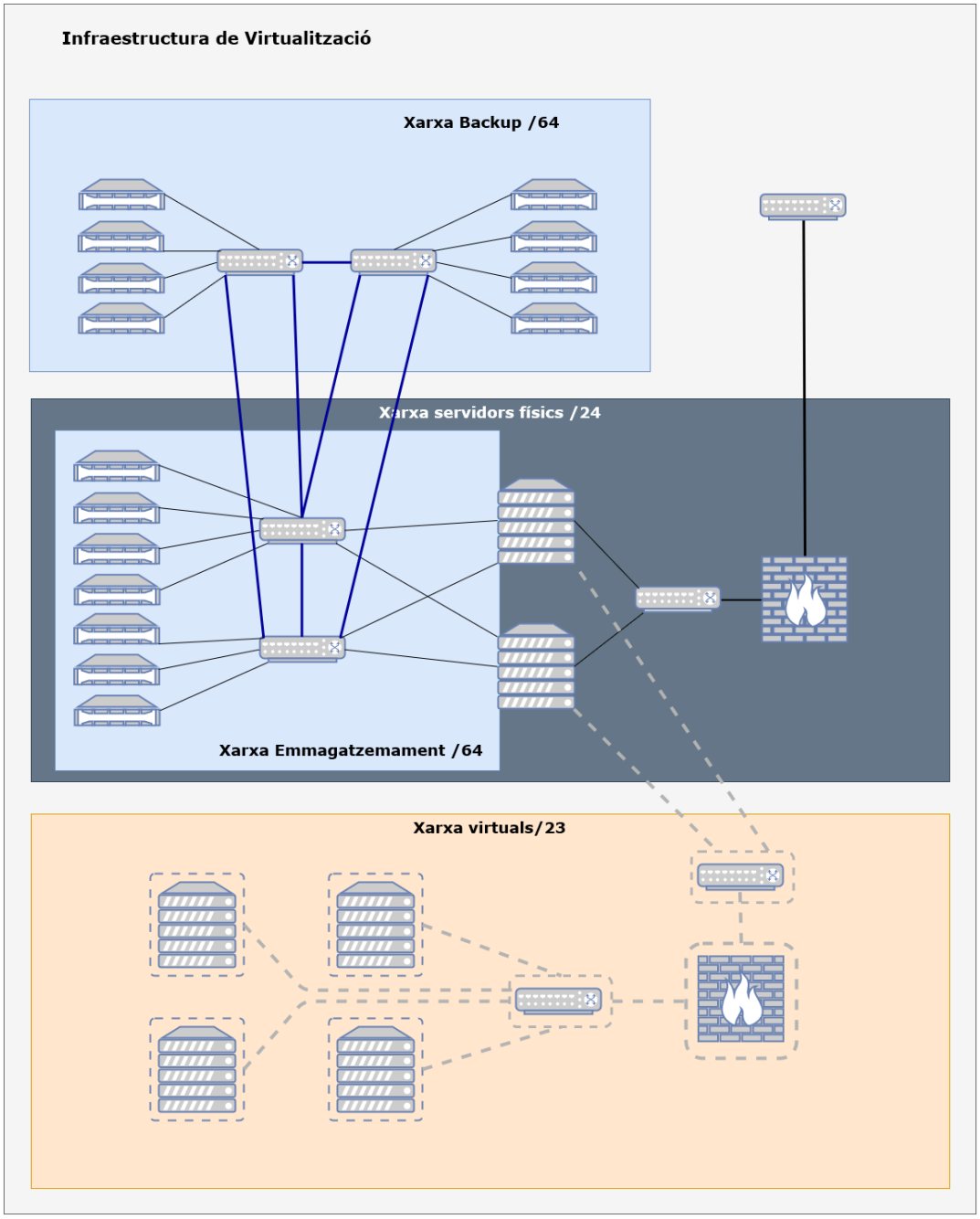
Per últim, el maquinari proposat d'inici consistiria en commutadors de 24 o 48 ports, depenent dels requeriments del sistema d'emmagatzematge. El fet de triar aquests commutadors i no altres com TP-Link, Linksys o Ubiquiti és que els Netgear tenen la característica que accepten xarxes de 10G sota cablejat de coure. Els altres que s'han analitzat no implementaven ports 10G, implementaven pocs i sota ports SFP/SFP+ o si els tenien, eren de fibra.

Tipus de maquinari	Marca / Model	Unitats
Commutadors 52ports	Netgear / NETGEAR XS748T commutador administrat 10Gigabit L2+ / L3, 48ports	1
Commutador 28ports	Netgear / NETGEAR XS728T commutador administrat 10Gigabit L2+ / L3, 28ports	1
Cablejat UTP	Cablejat UTP cat7	-

- Taula 23: Maquinari xarxa emmagatzematge.

2.2.4 Xarxa backup

Igual que passa amb la xarxa d'emmagatzemament, la xarxa de backup s'implementarà amb IPs d'IPv6. A més, per la seva interacció amb aquesta es planteja que s'integri amb la mateixa xarxa d'emmagatzematge per poder realitzar copies tant dels servidors de virtualització com dels d'emmagatzemament. Per tant s'implementarà en la xarxa d'IPv6 `fd43:d7f9:fc7b:de4a::/64`.



- Figura 13: Esquema infraestructura de virtualització amb xarxa de backup.

Respecte al maquinari serà similar al de la xarxa d'emmagatzemament. Commutadors administrables i capacitat per a ports 10G sota cablejat de coure.

Tipus de maquinari	Marca / Model	Unitats
Commutadors 52ports	Netgear / NETGEAR XS748T commutador administrat 10Gigabit L2+ / L3, 48ports	1
Commutador 28ports	Netgear / NETGEAR XS728T commutador administrat 10Gigabit L2+ / L3, 28ports	1
Cablejat UTP	Cablejat UTP cat7	-

- Taula 24: Maquinari xarxa backup.

3. Sistemes d'emmagatzematge

Una de les bases d'aquest projecte és el canvi de tecnologia referent a l'emmagatzematge. En la infraestructura actual, partim d'un sistema de servidors per la virtualització amb VMware, aïllats entre si pel que fa a l'emmagatzemament. És a dir, cada servidor fa servir els seus propis recursos de CPU, RAM i Disc. Això provoca limitacions a l'hora de la gestió de l'emmagatzemament. Per una banda, en cas d'esgotar l'espai assignat al servidor (físic), provocaria migracions de virtuals a altres servidors o haver d'ampliar físicament els discos, que segons el maquinari pot no ser trivial. Per l'altra banda, en cas d'esgotar els recursos de RAM i/o CPU ens comportaria un desaprofitament de l'espai de disc no consumit per la impossibilitat de compartir-ho amb altres maquinaris.

En alguns casos se'n fan servir cabines de discos però són estructures rígides i no fàcilment escalables o únicament destinades a fer còpies de seguretat, que no aporten el caire dinàmic i escalable que es busca en aquest projecte. Per aquesta raó es planteja l'opció alternativa a la continuïtat que implica canviar la infraestructura actual.

3.1 Requeriments del sistema d'emmagatzematge

Com s'ha comentat en altres seccions, l'empresa ens planteja una sèrie de requeriments que es centren principalment en tres aspectes:

- Dinamisme en la gestió de l'emmagatzematge
- Sistema distribuït.
- Sistema fàcilment escalable.

Es vol tenir dinamisme a l'hora de crear, ampliar i/o esborrar els volums per poder oferir la capacitat de gestionar-los mitjançant una API, de manera que es puguin gestionar de forma desatesa i així poder enllaçar-ho amb l'extranet de l'empresa.

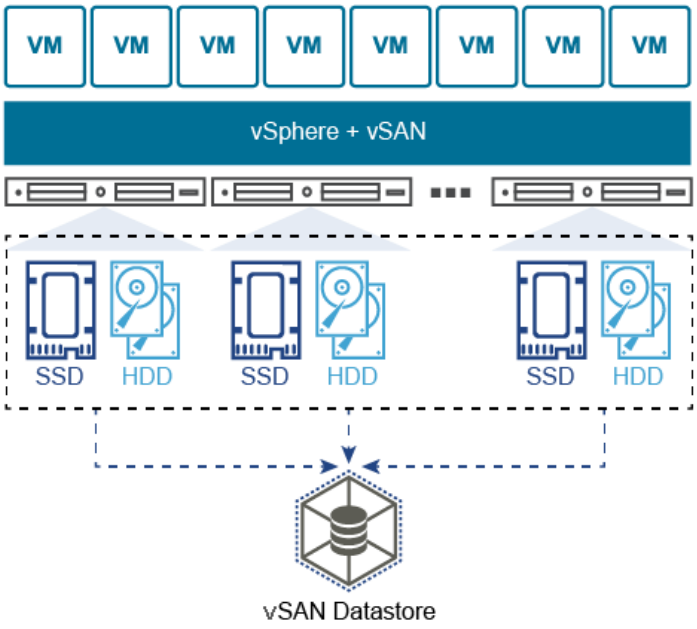
Per altra banda, es vol que la infraestructura que s'implementi sigui de caràcter distribuït. Es parteix d'aquesta premissa per l'experiència amb sistemes monolítics que a més de ser costosos, no són els més òptims en aspectes com pot ser el seu manteniment, actualitzacions i creixement.

Finalment, relacionat amb el punt anterior, es requereix que el sistema sigui fàcilment escalable de manera que ofereixi la possibilitat de fer servir diferents maquinaris tant a l'hora d'ampliar recursos com a l'hora de reemplaçar maquinari espatllat o obsolet i no els limiti a fer servir un maquinari concret.

Inicialment es plantegen dues opcions SDS (Emmagatzemament definit per Programari), ja que aquesta tecnologia és la que pot aportar els requeriments comentats de forma més nativa. Una estaria lligada en cert grau, a la continuïtat dels sistemes de virtualització que seria vSAN[25] i l'altra seria un canvi cap al l'opensource que, tot i ser compatible amb els sistemes actuals, apostaria per un canvi també en el sistema de virtualització. Aquest seria CEPH.

VMware VSAN

Aquest sistema, a diferència de l'utilitzat actualment que seria SAN[26], és un sistema definit per programari que permet una capa abstracta per crear una SAN virtual sense haver de tenir present un maquinari específic. Això facilita aspectes com ampliacions i/o reemplaçament de maquinari a la vegada que abarateix costos respecte a SAN que requeria maquinari específic tant per la seva implementació com per a posteriors actualitzacions. A més, ja és implementat als programaris actuals d'ESXi per la qual cosa no requereix una instal·lació especifica important.



- Figura 14: Esquema VSAN – https://docs.vmware.com/en/VMware-vSAN/index.html

En l'aspecte de gestió de dades, vSAN no treballarà amb LUNs com el tradicional SAN sinó que treballarà per blocks, permetent així una gestió d'aquest més eficient amb aspectes com la deduplicació i la compressió de dades.

L'inconvenient d'aquest sistema és que necessita de llicències del programari VMware per als ESXi, la gestió centralitzada vCenter, el programari de còpies de seguretat Veeam backup, etc. Aquests poden arribar a ser un cost important per l'empresa afegit al hardware necessari. A més, aquest sistema només permet l'ús de recursos locals i per tant només es podran fer servir discos allotjats al cluster. Això ens limita bastant el sistema, ja que els sistemes d'emmagatzematge com a servei com ara AWS S3 no es podran fer servir amb VSAN.

Un exemple d'aquest sistema és l'ofert per VMware junt amb Dell EMC anomenat Dell EMC VxRail.

Més info: https://www.vmware.com/es/products/hyper-converged-infrastructure/dell-emc-vxrail.html

CEPH

Respecte a CEPH, es un sistema d'emmagatzemament per a l'ús de grans volums de dades definit per programari com ara vSAN, però construït a partir de programari lliure que està orientat a l'emmagatzemament per objectes, per blocs i per sistema de fitxers. Es basa en un sistema distribuït i tolerant a fallades d'emmagatzemament d'objectes anomenat RADOS (Reliable Autonomic Distributed Store) que fa servir CRUSH[27], un algoritme hash que determinarà on i com es guarden les dades dintre de RADOS de manera uniforme que evitarà haver de buscar les dades en un index central. A més proporciona emmagatzemament per blocks totalment compatible amb OpenStack.

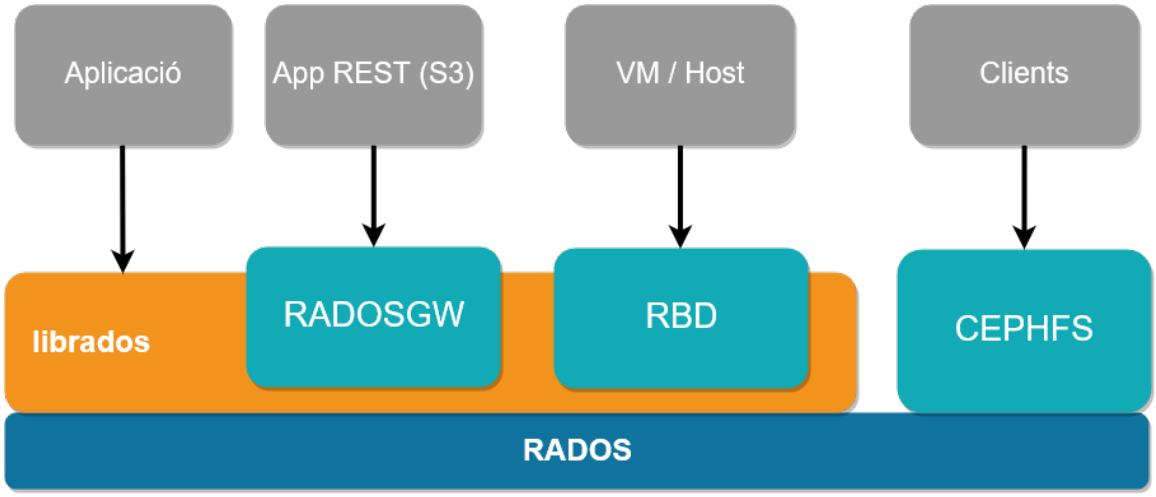
El seu funcionament es basa en quatre tipus de dimonis (daemons) diferents:

- **Monitors (ceph-mon)**: Aquest domini s'encarregarà de gestionar els diferents mapes d'estat del clúster necessaris per a coordinar-se amb els altres dimonis del clúster CEPH (mapa OSD, mapa PG, mapa MDS). El seu nombre por estar a partir d'un únic node monitor en endavant mantenint sempre un nombre imparell (2n+1) de nodes en cas de requerir redundància i alta disponibilitat. En el cas de tenir més d'un únic node els nodes hauran d'arribar a un quòrum per mantenir l'estabilitat del sistema.
- **Managers (ceph-mgr)**: Aquest dimoni s'encarrega de gestionar les mètriques dels diferents aspectes de CEPH, rendiment, ús de l'emmagatzemament, càrrega del sistema, etc. Aquestes dades es poden fer servir per a tenir un monitoratge que pot ser accessible mitjançant web i/o una api REST[28]. El seu nombre pot estar entre un node o dos si volem alta disponibilitat.
- **OSDs (ceph-osd)**: Aquest dimoni s'encarregarà de gestionar l'emmagatzemament de dades, la seva replicació entre els diferents nodes, el seu balanceig, recuperació. Es requereixen almenys tres nodes OSD per a poder oferir redundància i alta disponibilitat.

- **MDSs (ceph-mds)**: Aquest node només serà necessari en cas de voler oferir emmagatzemament a escala de sistema de fitxers CEPH i permetrà executar comandes bàsiques sense provocar un cost elevat al clúster.

A part de tot això, CEPH ens permetrà accedir a l'emmagatzemament , de diferents maneres depenent del tipus d'emmagatzemament que volem.

Es podrà accedir directament a RADOS a través d'aplicacions que ataquin les llibreries *librados*. També es podrà accedir als objectes directament a través del RADOSGW per l'API REST compatible amb Amazon S3 gateway. Per altra banda, es podrà accedir als block/volums a través de RADOS Block Device. Per últim es podrà accedir com a sistema de fitxers (cephfs) a través de clients.



- Figura 15: Esquema RADOS

L'elecció, per diversos aspectes com poden ser el ser de codi obert, la seva tolerància a fallades amb zero punts de fallada, la seva capacitat de treballar sobre maquinari no específic, la capacitat d'accés a l'emmagatzemament per diferents interfícies (objectes, blocs o sistema de fitxers) i una major adaptabilitat a allò que vol l'empresa, junt amb la possibilitat de treballar amb OpenStack de forma gairebé nativa a la vegada que es poden generar connectors iSCSI per a treballar puntualment amb VMware ens fa decantar per a la implementació de Ceph com a sistema d'emmagatzemament principal, encara que conviurà temporalment amb alguns sistemes SAN d'estructures actuals. Un factor molt important ha estat la seva capacitat de ser escalat de forma senzilla amb l'única necessitat d'afegir un nou host OSD amb nous recursos, ja que el mateix sistema s'encarregarà de balancejar i sincronitzar el sistema per fer ús dels nous recursos. A més Ceph està orientat a infraestructures amb grans capacitats d'emmagatzematge (PetaBytes) i això ens dóna la tranquil·litat de no quedar limitats en recursos a curt i mig termini.

En la part que hem comentat referent a les interfícies en què Ceph ens permet oferir emmagatzemament, mentre que VSAN només ofereix emmagatzemament en blocks, Ceph ofereix emmagatzemament com a objectes, en blocks que ens permetrà crear volums i inclòs com a sistema de fitxers que podran ser afegits a clients com a unitats independents mitjançant un client Ceph.

La part econòmica, també és un factor a tenir en compte, ja que Ceph proporciona dues possibilitats interessants en aquest aspecte. La primera, que és un programari de codi obert i per tant, tot i haver-hi solucions de pagament com ara RedHad Ceph Storage, ens permet muntar la infraestructura sense haver de pagar cap llicencia, això si, no sense un grau de dificultat important. Per altra banda, la segona possibilitat interessant en l'aspecte econòmic, és que Ceph permet fer ús de qualsevol classe de maquinari, fet que no ens obligarà a haver de comprar maquinari especialitzat i més costos.

A part de tot això, en un aspecte més orientat a la gestió, però que també és interessant, Ceph proporciona una api pròpia i una REST api compatible amb el format S3 que simplifica moltíssim l'administració i gestió del cluster.

En el nostre cas, instal·larem l'última versió, 14.2.1 anomenada *Nautilus*.

(http://docs.ceph.com/docs/master/releases/nautilus/)

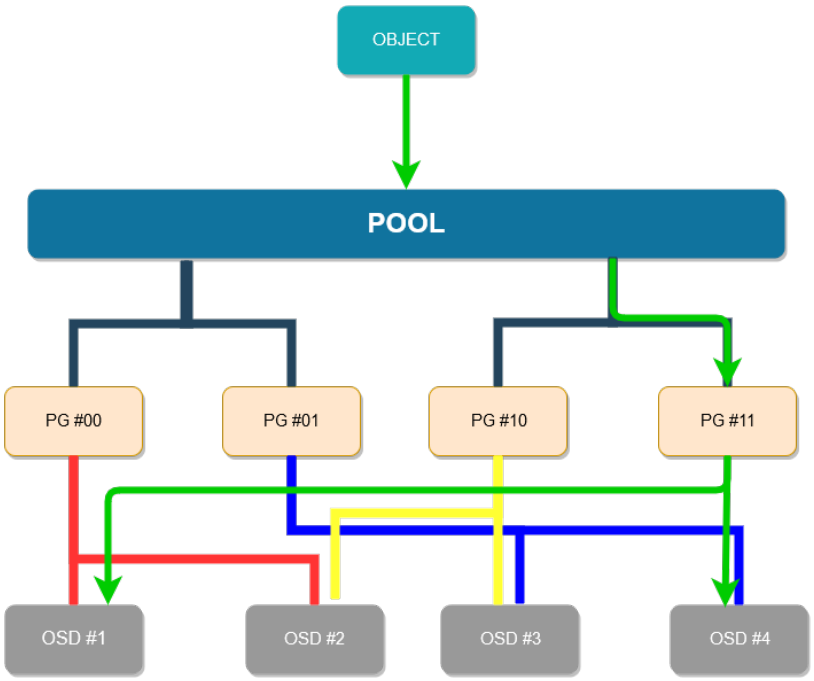
3.2 Desenvolupament del sistema triat

Una vegada hem triat Ceph com a sistema d'emmagatzemament, definirem una sèrie de conceptes que haurem de conèixer per saber com s'estructura Ceph. Com s'ha comentat abans, Ceph funciona a partir d'*objectes* els quals seran la seva unitat bàsica d'emmagatzemament i els que una vegada mapejats seran distribuïdes còpies d'aquests objectes entre altres nodes OSD.

El sistema que seguirà Ceph per replicar les dades vindrà determinat per quin dels dos mètodes de protecció de dades triem. El primer és el mètode clàssic que consisteix en la rèplica completa de l'objecte en el nombre d'OSD indicat. Si algun dels OSD en els que està la còpia cau, el sistema es balancejarà per allotjar la nova rèplica necessària. El segon mètode, anomenat ''Erasure coding'' consisteix en un sistema similar al RAID5, ja que replicarà l'objecte dividit en *n* fragments que es determinaran per k fragments de les dades i m fragments codificats per donar protecció de dades.

Al mètode de rèplica que és el que es farà servir, el nombre de nodes en els quals es guardaran còpies de cada objecte, o dit d'altra manera, el nombre de còpies que existiran de cada objecte, és un paràmetre anomenat «mida de rèplica» (*rep size*) que haurem de definir. Els objectes originals que després s'hauran de replicar en altres OSD estaran associats a uns ''contenidors'' anomenats PG (*Placement Group*) i ens aportaran la representació de com el cluster ha distribuït aquests objectes. El nombre de PG variarà segons els interessos de cada instal·lació encara que la recomanació de Ceph és que estigui

al voltant de 100PG per OSD. Els PG estaran agrupats en contenidors lògics anomenats pools. Aquests pools hauran de ser creats amb l'especificació, tant del nombre de rèpliques com del nombre de PGs encara que existeixen valors per defecte i en el cas dels PG es podria configurar perquè el mateix Ceph s'autogestiones.



- Figura 16: Esquema de funcionament de la replicació d'objectes a CEPH.

El funcionament bàsic seria el següent. Un client guarda un objecte al pool, mitjançant l'algoritme CRUSH comentat abans es decidiria en quin PG s'hauria d'emmagatzemar i mitjançant CRUSH una altra vegada, es decidiria en quins OSD es replicaria.

De totes maneres, la documentació de Ceph recomana un nombre de PG per pool segons el nombre d'OSD que tinguem:

- *pg_num* configurat a 128 per instal·lacions de menys de 5 OSDs
- *pg_num* configurat a 512 per instal·lacions d'entre 5 i 10 OSDs
- *pg_num* configurat a 1024 per instal·lacions d'entre 10 i 50 OSDs
- per a instal·lacions amb més OSDs s'hauria de calcular el nombre de PG optim a partir d'eines com ara la que ofereix la pròpia pàgina de Ceph. (*https://ceph.com/pgcalc/*)

Un altre aspecte important per dissenyar el sistema d'emmagatzemament per l'allotjament de servidors virtuals privats i compartits ha de ser saber la quantitat de virtuals que volem allotjar, la quantitat d'espai que volem destinar a còpies de seguretat i la quantitat d'espai que volem reservar per a ampliacions futures.

Si reprenem el disseny de les xarxes de virtualització, recordem que es van crear dues xarxes /23 de 510 IPs útils per a virtuals de clients i dues més /24 de 254 IPs per a servidors privats als CPD 2 i 3. La previsió en els casos de les xarxes /23 és la d'arribar als 400 virtuals en cada una d'elles i de 100 virtuals en les /24. Ens reservem un tant per cent de les IPs de cada xarxa per casos en què es requereixin servidors amb múltiples IPs i per als servidors físics.

Aquest càlcul ens dóna un total de 500 virtuals que s'allotjaran en cada cluster CEPH, amb una mitja de 100GB de disc per cadascun, necessitaríem un mínim de ±50TB dedicats únicament als virtuals. A més, les recomanacions de CEPH en referència a l'ús de disc, és que ha d'estar sempre per sota del 85% del total del clúster. Això és perquè el sistema necessitarà aquest romanent per gestionar àgilment la relocalització de dades en cas de caigudes de nodes OSD o en una possible ampliació d'aquests. Aquesta política ens marcarà un mínim de 60TB per estar dintre

d'aquest 85% encara que augmentarem la capacitat fins als 100TB per tenir un marge acceptable.

Per altra banda, la idea inicial és la de replicar els clusters entre datacenters amb la qual cosa augmentarem la capacitat necessària fins als 200TB per clúster. 100TB per als pools destinats als virtuals i 100TB destinat a la rèplica de l'altre clúster.

Per obtenir aquesta capacitat per als nodes OSD, farem servir 4 servidors de la marca DELL amb les següents característiques.

Opció	**Selecció**	**Quantitat**
Bàsic	PowerEdge R740 Server	1
Configuració de xassís	Chassis with up to 16 x 2.5" SAS/SATA Hard Drives for 2CPU Configuration	1
Processador	Intel® Xeon® Silver 4110 2.1G, 8C/16T, 9.6GT/s, 11M Cache, Turbo, HT (85W) DDR4-2400	2
Memòria	32GB RDIMM, 2666MT/s, Dual Rank	2
Disc d'inici	BOSS controller card + 2 M.2 Sticks 240G (RAID 1),FH	1
Controlador RAID	PERC H740P RAID Controller, 8Gb NV Cache, Adapter, Low Profile	1
Disc dur	3.84TB SSD SAS Read Intensive 12Gbps 512 2.5in Hot-plug AG Drive, 1 DWPD, 7008 TBW	16
Font d'alimentació	Dual, Hot-plug, Redundant Power Supply (1+1), 750W, Titanium, 200-240VAC	1
Adaptador de xarxa	Broadcom 57416 Dual Port 10GbE BASE-T & 5720 Dual Port 1GbE BASE-T, rNDC	1

- Taula 25: Proposta servidors OSD

Amb aquest servidor tindrem 16 discos SSD SAS de 3,84 TB cada un, dels quals farem servir 14 discos pel cluster, 1 pel journal d'OSD i deixarem 1 com a reserva per activar-ho en cas de fallada d'algun disc. D'aquesta manera aconseguirem 3.84 TB x 14 discos x 4 servidors = 215.04TB al cluster. A més, separarem l'espai en disc per al SO amb dues unitats M.2 de 240G en RAID1 per oferir redundància. Aquestes unitats ofereixen un accés a disc molt més ràpid per al SO.

Una vegada tenim definit la capacitat d'emmagatzemament, també és important dimensionar les CPU i memòria els servidors.

Pel que fa a la CPU, instal·larem 2 CPU Intel Xeon de 8 cores/CPU que seran suficients per la gestió de les dades per part del dimoni OSD, ja que tindrem un core per disc físic, és a dir 1 core/OSD.

Pel que fa a la memòria, la documentació de Ceph recomana un mínim d'1 GB de RAM per cada TB d'emmagatzemament de cada dimoni OSD. Com que tindrem 3.84TB per disc x 14 discos, tindrem 53,76 TB per dimoni (host). Per tant, instal·larem 64GB de RAM.

Pel que fa a la connectivitat, tenim dos ports de 10GE per a l'enviament de dades entre els OSD, ja que aquesta connexió necessitarà una amplada de banda gran. A més tenim dos ports 1GBE per altres connexions com ara amb els monitors.

Una vegada tenim el maquinari definit per als OSD, haurem de calcular el rendiment en tasses de transferències i IOPS[29] que podem suportar al nostre cluster.

Segons les mètriques que ens ofereix DELLEMC veiem que els discos que s'instal·laran als servidors seran un dels següents:

- Toshiba PM4 PX05SR
- Samsung PM1633a de 3840GB

Type	Capacity (GB)	MODEL	Part Number	Endurance DWPD 5 Years	Endurance TBW (TB)	Seq Read 128KB (MB/s)	Seq Write 128KB (MB/s)	Random Read 4KB (IOPS)	Random Write 4KB (IOPS)	Random 4K 70/30
MU	400	Samsung PM1635a	MFC6G	3.0	2,190	950	600	199K	61K	75K
	800	Samsung PM1635a	HF06W	3.0	4,380	950	970	200K	76K	80K
	1600	Samsung PM1635a	W5PP5	3.0	8,760	940	1,010	200K	76K	80K
	480	Toshiba PM4 PX05SV	43PCJ	3.0	2,628	1,050	1,050	205K	60K	110K
	960	Toshiba PM4 PX05SV	503M7	3.0	5,256	1,040	1,040	200K	70K	110K
	1920	Toshiba PM4 PX05SV	V0K7V	3.0	10,512	1,040	1,050	204K	97K	110K
	3840	Toshiba PM4 PX05SV	3DDFT	3.0	21,024					
RI	960	Toshiba PM4 PX05SR	MWGK7	1.0	1,752	1,050	1,050	200K	30K	70K
	1920	Toshiba PM4 PX05SR	0FYFW	1.0	3,504	1,050	1,050	200K	35K	90K
	3840	Toshiba PM4 PX05SR	XCRDV	1.0	7,008	1,050	1,050	190K	35K	90K
	1920	Samsung PM1633a	086DD	1.0	3,504	950	1,000	201K	37K	80K
	3840	Samsung PM1633a	JR1HP	1.0	7,008	950	1,000	201K	37K	80K

- Taula 26: Rendiments discos. https://i.dell.com/sites/csdocuments/Shared-Content_data-Sheets_Documents/en/dell-poweredge-sas-ssd-performance-specifications.pdf

Com veiem, cada disc ens oferirà els següents rendiments I/O:

Mesura d'IOPS	Rendiment Toshiba	Rendiment Samsung
Lectura seqüencial 128KB	1050MB/s	950 MB/s
Escriptura seqüencial 128KB	1050MB/s	1000MB/s
Lectura aleatòria 4KB	190KIOPS	201KIOPS
Escriptura aleatòria 4KB	35KIOPS	37KIOPS
Total 70/30 (lectura/escriptura	90KIOPS	80KIOPS

- Taula 27: Rendiments discos OSD.

Com que tindrem 14 discos a cada servidor, al treballar tots en diferents OSD la limitació de rendiment serà provocada per la controladora Perk H740P que serà capaç d'oferir taxes de transferència de 12GB/s.

Una vegada tenim definit el Hardware per als nodes OSD definirem el nombre d'OSDs de cada node. Segons la documentació de Ceph es recomana que cada OSD estigui configurat en un únic disc, encara que es podria configurar en estructures RAID el que faria que obtinguéssim un rendiment per node més elevat però en detriment del rendiment per OSD que es veuria afectat en major o menor mesura segons el tipus de raid. Per tant, en la part OSD tindrem quatre nodes OSD amb catorze OSD de 3,84TB cadascun. És a dir 56 OSD en tot el clúster i una capacitat aproximada de 215TB.

Amb aquesta informació podem calcular també el nombre de PG que haurem de configurar amb l'eina que hem comentat abans. En aquesta eina, al triar una instal·lació per a OpenStak versió Jewel amb Rados gateway configurat i un total de 56 OSD ens dóna una recomanació de configuració dels diferents pools necessaris segons el tant per cent de dades que allotjarà cada pool d'OpenStack sobre un objectiu de 100PG per OSD. Resultant un nombre total de 2656 PG al clúster.

Pool Name	Size	OSD #	%Data	Target PGs per OSD	Suggested PG Count
.rgw.root	3	56	0.10	100	32
default.rgw.control	3	56	0.10	100	32
default.rgw.data.root	3	56	0.10	100	32
default.rgw.gc	3	56	0.10	100	32
default.rgw.log	3	56	0.10	100	32
default.rgw.intent-log	3	56	0.10	100	32
default.rgw.meta	3	56	0.10	100	32
default.rgw.usage	3	56	0.10	100	32
default.rgw.users.keys	3	56	0.10	100	32
default.rgw.users.email	3	56	0.10	100	32
default.rgw.users.swift	3	56	0.10	100	32
default.rgw.users.uid	3	56	0.10	100	32
default.rgw.buckets.extra	3	56	1.00	100	32
default.rgw.buckets.index	3	56	3.00	100	64
default.rgw.buckets.data	3	56	19.00	100	512
cinder-backup	3	56	18.00	100	256
cinder-volumes	3	56	42.80	100	1024
ephemeral-vms	3	56	10.00	100	256
glance-images	3	56	5.00	100	128
		Total Data Percentage: 100.00%			PG Total Count: 2656

- Figura 17: Exemple pools OpenStack. Taula extreta de la web: https://ceph.com/pgcalc/

A part dels nodes OSD també haurem de configurar els nodes monitors, el RADOS GW i administradors que es muntaran en tres servidors DELL R640 de les següents característiques.

Opció	Selecció	Quantitat
Bàsic	PowerEdge R640 Server	1
Configuració de xassís	2.5" Chassis with up to 8 Hard Drives and 3PCIe slots, 1 or 2 CPU	1
Processador	Intel® Xeon® Gold 5118 2.3G, 12C/24T, 10.4GT/s, 16M Cache, Turbo, HT (105W) DDR4-2400	2
Memòria	16GB RDIMM, 2666MT/s, Dual Rank	8
Controlador RAID	PERC H740P RAID Controller, 8Gb NV Cache, Minicard	1
Disc dur	800GB SSD SAS Mix Use 12Gbps 512e 2.5in Hot-plug AG Drive, 3 DWPD, 4380 TBW (RAID 5)	4
Font d'alimentació	Dual, Hot-plug, Redundant Power Supply (1+1), 750W, Titanium, 200-240VAC	1
Adaptador de xarxa	Broadcom 57416 Dual Port 10GbE BASE-T & 5720 Dual Port 1GbE BASE-T, rNDC	1

- Taula 28: Proposta servidors monitors.

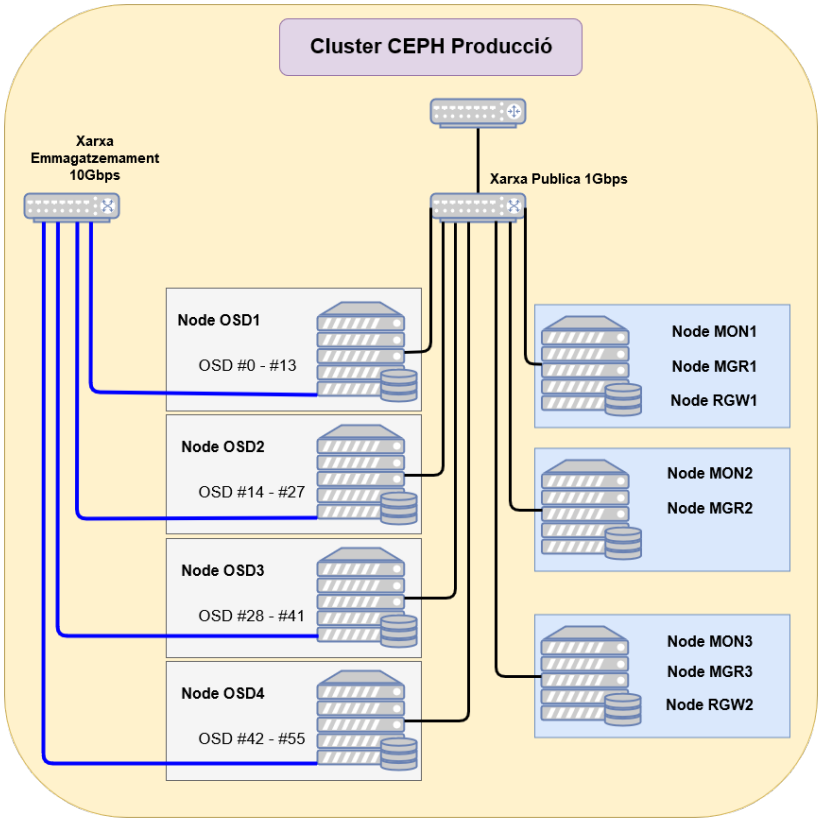
Com veiem, en aquest cas hem seleccionat el model *PowerEdge R640* de DELLEMC per als monitors i altres dimonis (MGR, RGW, MDS) amb un xassís per a 8 discos 2.5''. Quant a CPU, la documentació no dóna molta importància a la capacitat de CPU dels monitors, ja que aquests només s'encarregaran de mantenir una còpia principal del mapa del clúster i no necessiten potència de còmput. Per altra banda, el fet d'aprofitar aquests mateixos servidors per la instal·lació dels dimonis ceph-mgr, ceph-rgw i ceph-mds fa que hàgim de dotar aquests de potència de còmput i per tant muntarem dues CPU *Intel Xeon Gold 5118*.

Pel que fa a la memòria RAM, les recomanacions per als monitors es d'entre 5 i 10 GB per dimoni per entorns grans, pels RGW les recomanacions són d'entre 32 i 64GB per dimoni i pels MDS un mínim de 3GB. Per aquesta raó es muntaran 128GB de RAM per servidor.

Per últim, es muntaran 4 discos SSD SAS de 800GB en RAID 5 per allotjar el SO i els diferents dimonis.

La distribució proposada pels servidors R640 pel que fa als dimonis que gestionarà serà la següent:

- **Monitor ceph-mon**: S'instal·larà ceph-mon en tots tres servidors per obtenir Redundància.
- **Administrador ceph-mgr**: S'instal·larà ceph-mgr en els mateixos servidors que ceph-mon per obtenir redundància. A partir de la versió de Ceph (12.x *Luminous*) aquest component és necessari.
- **Metadata ceph-mds**: D'inici no es muntarà aquest dimoni, ja que de moment no està previst fer servir CephFS.
- **RADOSGW**: S'instal·larà RGW en dos servidors inicialment però segons els rendiments que s'obtinguin, es plantejarà la instal·lació en el tercer servidor.



- Figura 18: Esquema cluster Ceph.

Amb tot això, ja tindríem dissenyat el sistema Ceph d'emmagatzemament i preparat per a la seva implantació. Com a exemple de la seva instal·lació i funcionament es desenvolupa un pilot de característiques més reduïdes.

3.3 Escenari pilot a laboratori

El pilot proposat es portarà a terme amb l'estructura següent:

- 2 servidors *DELLEMC PowerEdge R330* amb SO *ESXi 6.5*
- 6 nodes virtuals amb instal·lació minimal de *CentOS 7*

Dels 6 nodes farem servir 3 per als nodes OSD, 1 pel node monitor, 1 pel node ceph-admin i 1 per instal·lar un client, on a part de provar el muntatge d'espai d'emmagatzemament, s'instal·lara el programari per a monitorar el rendiment i estat del clúster Ceph *Prometheus* i *Grafana*.

Hostname	dirección IP
ceph-admin	192.168.130.120
mon1	192.168.130.121
osd1	192.168.130.131
osd2	192.168.130.132
osd3	192.168.130.133
client	192.168.130.117

- Taula 29: Llistat IPs pilot Ceph.

(*El procés d'instal·lació i les proves fetes, es poden trobar a l'annex VI*)

(*Es poden trobar les proves de rendiment fetes a l'annex XV*)

4. Sistemes de virtualització

L'altre aspecte principal d'aquest projecte és la tria del sistema de virtualització que es farà servir junt amb Ceph per tal de donar servei IaaS[30] als clients i a la mateixa empresa.

Inicialment partim d'una estructura molt dividida on el sistema de virtualització (*Hypervisor*) principal és VMware ESXi i majoritàriament, els diferents servidors aporten físicament els seus recursos de CPU, RAM i disc. Per una part, aquest sistema té una instal·lació molt senzilla, ja que un nou servidor només requereix una instal·lació simple del SO ESXi i poca configuració per funcionar. En contrapartida, no té una estructura centralitzada i per tant per administrar els virtuals de diferents servidors físics es requereixen connexions diferents. A més, com que no tenim una estructura centralitzada d'emmagatzemament compartit, la gestió de l'espai de disc és poc dinàmica i les migracions entre servidors acostumen a ser una tasca molt lenta, atès l'augment de consum de disc experimentat en els últims anys.

Per aquesta raó es planteja un canvi en aquest àmbit per solucionar i/o corregir els problemes de l'estructura actual.

4.1 Requeriments del sistema de virtualització

Els requeriments principals que planteja l'empresa es basen en la simplificació en la gestió de la infraestructura, és a dir, no és productiu haver de tenir una connexió diferent per a cada servidor. Per aquest motiu, es precisa la implementació d'una gestió centralitzada i distribuïda per tal de poder gestionar tota l'estructura des d'una mateixa interfície. A més, ha d'existir una interconnexió real entre servidors que faciliti la interacció entre ells i els servidors virtuals.

Un altre requeriment que planteja l'empresa és que el sistema ha de ser fàcilment escalable. La infraestructura ha de contemplar un creixement futur i per tant ha de ser possible afegir maquinari de forma transparent.

A més, el sistema de virtualització ha d'aportar una API per poder treballar directament amb la finalitat d'enllaçar el sistema amb una extranet on els clients puguin donar d'alta, modificar i donar de baixa virtuals de manera senzilla, dinàmica i en temps real.

Es valorarà positivament l'ús de programari de codi obert tot i que un requisit important serà el suport i la comunitat que tingui aquest programari.

Amb aquests requeriments es plantegen dues opcions per a assolir la part de virtualització de la infraestructura proposada. El primer passa per una continuïtat en l'ús de VMware ESXi tot i que amb el llicenciament del programari necessari per complir els requeriments exposats. L'altra opció proposada és la implementació d'OpenStack com a Hypervisor, ja

que a més de ser un programari de codi obert, en els últims anys està experimentant un creixement exponencial i a més la compatibilitat amb Ceph és absoluta.

Virtualització amb VMware ESXi

La implementació d'aquest programari com sistema de virtualització per al nostre projecte consistiria en la instal·lació de maquinari per a l'allotjament del SO ESXi amb CPU i RAM per a la virtualització i el magatzem de dades (datastore) com a una unitat iSCSI contra Ceph, ja que VMware no és compatible amb el sistema de blocs o objectes de Ceph.

L'estructura de VMware consistiria inicialment en quatre servidors amb ESXi 6.7 instal·lats amb una màquina virtual muntada amb Microsoft Windows server 2016, on s'instal·larà el programari *virtual center (vCenter)* per a la creació i gestió del clúster format pels quatre servidors. VMware ens ofereix dos tipus de llicències per a *vcenter*. Una en la que ens permet gestionar quatre servidors físics (*VMware vCenter Server Foundation*) i una altra en la que el nombre de servidors a gestionar és indiferent, ja que es paga per una llicència de vCenter. (*VMware vCenter Server Standard*). En el nostre cas començaríem per la llicència de tipus *Foundation*. A més, hauríem de llicenciar els ESXi amb el tipus de llicència que més s'adaptés al nostre projecte, ja que primerament es llicencien segons el nombre de CPUs que tingui cada ESXi i segon, el tipus de llicència per obtenir més o menys serveis com ara HA (alta disponibilitat).

Actualment existeixen quatre tipus de llicències d'ESXi, segons els serveis que ofereixen:

- vSphere Standard
- vSphere Enterprise Plus
- vSphere Enterprise plus with Operations Manager
- vSphere Platinum

Derecho de licencia				
	vSphere Standard	vSphere Enterprise Plus	vSphere with Operations Management Enterprise Plus	vSphere Platinum
Descripción	consolidación de servidores y continuidad del negocio	Gestión de recursos, mejora de la disponibilidad y del rendimiento de las aplicaciones	Automatización y gestión de las operaciones inteligentes con análisis predictivo	Visibilidad y seguridad mejoradas, con tecnología de aprendizaje automático e integrada en el hipervisor
Derecho de licencia	Por CPU	Por CPU	Por CPU	Por CPU
vCenter Server (se vende por separado)	vCenter Server Standard	vCenter Server Standard	vCenter Server Standard	vCenter Server Standard
vSphere Integrated Containers	ND	•	•	•
vRealize Operations	ND	N/d	vRealize Operations Standard	N/d
vRealize Log Insight for vCenter Server*	Paquete de 25 OSI por instancia de vCenter Server Standard	Paquete de 25 OSI por instancia de vCenter Server Standard	Paquete de 25 OSI por instancia de vCenter Server Standard	Paquete de 25 OSI por instancia de vCenter Server Standard
VMware AppDefense	ND	N/d	N/d	VMware AppDefense

* Disponible con las versiones de Log Insight for vCenter Server anteriores a la versión 4.6 hasta el 23 de agosto de 2019, día en el que acaba el soporte general de vRealize Log Insight 4.6

- Figura 19: Llicències Vmware vSphere: https://www.vmware.com/es/products/vsphere.html

Un altre aspecte que ens ofereix VMware és la implementació d'una API REST per a la gestió completa del vCenter anomenada *vSphere Automation SDK for Rest* que pot ser cridada amb llenguatges com ara Java o Python. Amb aquesta API compliríem el requeriment de l'API per a interaccionar amb el Hypervisor des d'una extranet per a gestionar els virtuals.

Per últim, VMware és la marca de virtualització per excel·lència i per tant existeix un servei de suport, una comunitat i documentació molt extenses.

Virtualització amb OpenStack

Primer de tot, podem definir OpenStack com un sistema operatiu de núvol format per un compendi de mòduls de codi obert que treballen junts a través d'APIs internes per a crear un sistema de Cloud tipus IaaS. Aquest programari ens permet crear infraestructures Cloud tant privades com públiques de forma molt dinàmica amb maquinari convencional, ja que crea una capa virtual per fer-lo transparent al programari.

Ateses les seves funcionalitats i la seva estructura, OpenStack ens facilita que la gestió dels virtuals que els clients o la mateixa empresa gestionaran sigui molt dinàmica, ja que permet que sigui l'usuari qui pugui agafar aquesta gestió i per tant, decidir en quin moment crear, eliminar, engegar, apagar o modificar les seves màquines (instàncies). Això permet al proveïdor gestionar el pagament dels virtuals per l'ús de recursos i pel temps que el client els fa servir. És a dir, un dinamisme en la contractació de serveis que fins ara no teníem. Relacionat amb això, permet que l'usuari no hagi de dependre del proveïdor per a la instal·lació d'un SO on el client pugui connectar, ja que com s'ha comentat abans, el client rep l'accés a la infraestructura i ell pot crear el seu virtual directament amb els recursos necessaris i instal·lar allò que necessiti.

Per altra banda, OpenStack junt amb Ceph, ens aporta un sistema infinitament escalable en maquinari, però també en virtuals, ja que ens permetrà escalar els recursos d'un virtual de forma immediata.

Un altre aspecte que aporta el caire modular d'OpenStack és la possibilitat d'instal·lar, no només els mòduls que necessitem, sinó que aquests mòduls els podem instal·lar en el mateix servidor o en servidors diferents i de forma única o amb redundància de maquinari. Fet que ens permet l'escalabilitat comentada anteriorment.

Per últim, pel que fa a la comunitat OpenStack és gestionat per l'*OpenStack Foundation* (*https://www.openstack.org/foundation/*) des de setembre de 2012 formant una comunitat de més de 82000 membres de 187 països diferents on podem trobar des de persones fins a companyies com ara IBM, DELL, Red Hat o Mirantis.

Una vegada vistos els dos sistemes IaaS ens centrarem en el desenvolupament d'OpenStack, ja que a part de ser un programari lliure ens aporta un nivell de llibertat i escalabilitat en la gestió de l'emmagatzemament i en la seva implantació que no ens dóna VMware. En ser un Programari modular, ens permet distribuir els diferents mòduls en diferent maquinari segons què convingui més a l'empresa. Per altra banda, tots dos sistemes ens aporten una API per gestionar el sistema però considerem que l'API d'OpenStack està més orientada a la gestió de qualsevol part del sistema tant a escala d'administrador com d'usuari, ja que es pot atacar directament les APIs dels diferents mòduls amb diferents nivells d'accés, mentre que l'API de VMware és un bloc més compacte orientat a l'administració. Per altra banda, tant un com l'altre tenen comunitats molt extenses i moltíssima documentació, però en el cas de VMware, en ser un programari privatiu, en alguns casos aquesta documentació o el suport donat son de pagament, mentre que en OpenStack és lliure. Tot i haver-hi també formació de pagament. A part de tot això, un aspecte important és l'orientació a cloud pública d'openstack que no te VMware més orientat a cloud privada.

Per aquestes raons, i naturalment per la seva compatibilitat amb Ceph, es proposarà per aquest projecte la implementació d'OpenStack per a la creació d'un núvol híbrid i la seva comercialització.

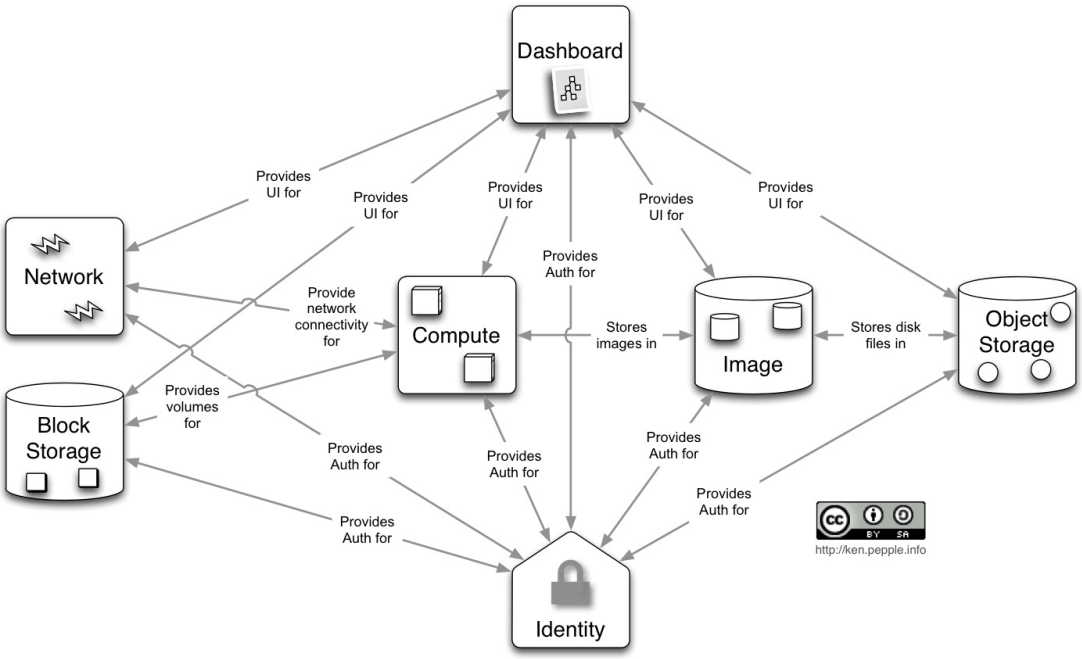
4.2 Desenvolupament del sistema triat

Per a poder desenvolupar la implantació d'OpenStack per a la nostra estructura, primer hem de veure com funciona internament. Primer de tot dir que és un sistema que entra dintre del grup dels **IaaS**, és a dir, Infraestructura com a Servei. Bàsicament consisteix en què un proveïdor de serveis pot oferir una infraestructura virtual al núvol (*projecte o tenant*), que el client pot fer servir per crear un servidor virtual i instal·lar un SO directament com si es tractés d'un maquinari propi.

Els mòduls que el formen estan diferenciats per la seva funcionalitat i els més importants són els següents:

- ***OpenStack Dashboard (Horizon)***: Aquest mòdul ofereix una interfície gràfica (Panell de control) desenvolupat en Django amb diferents rols de funcionament (administrador, usuari) que ens permet la gestió d'OpenStack.
- ***OpenStack Compute (Nova)***: És el mòdul principal d'OpenStack que ens permet desplegar i administrar les màquines virtuals i serveis allotjats a OpenStack mitjançant la gestió de les instàncies d'imatges a través de diferents Hypervisors com ara Xen, KVM, Hyper-V o VMware ESXi.
- ***OpenStack Network (Neutron)***: Aquest mòdul és l'encarregat de la comunicació entre mòduls, gestió d'IPs flotants i a més ens permet crear i gestionar xarxes virtuals que podrem associar als diferents dispositius que tinguem.
- ***OpenStack Block Storage (Cinder)***: Aquest mòdul, equivalent a Amazon EBS, és l'encarregat de la gestió de l'emmagatzemament permanent independent de les instàncies. A més, ens permet vincular i desvincular els volums creats a les instàncies.
- **OpenStack Identity Service (Keystone)**: És el servei encarregat de la gestió dels accessos dels usuaris, projectes, rols i serveis d'OpenStack. Emmagatzema la informació en un catàleg centralitzat com LDAP.
- **OpenStack Image Service (Glance)**: Aquest mòdul gestionarà les plantilles de les imatges dels sistemes a instal·lar a més de les fotos puntuals (Snapshots) de les instàncies. Pot gestionar els formats dels principals hypervisors, com ara *vmdk* de VMware, *vdi* de VirtualBox o *ami* d'Amazon.
- **OpenStack Object Storaje (Swift)**: Aquest component és l'encarregat de la gestió de l'emmagatzemament a escala d'objectes i ens permet aconseguir un sistema d'emmagatzematge massiu escalable i redundant, ja que gestiona la replicació de les dades entre els diferents dispositius d'emmagatzematge instal·lats. A més ens ofereix una API pròpia i una altra compatible amb Amazon S3.
- **OpenStack Orchestation (Heat):** Aquest mòdul s'encarregarà de gestió d'altres serveis d'alt nivell com ara AWS mitjançant HOT (Heat Orchestation Template)
- **OpenStack Telemetry (Celiometer):** Exercirà de monitor permetent l'obtenció de dades útils per a la facturació, informes, escalabilitat, etc.

Esquema d'interconnexió dels mòduls que formen OpenStack.



- Figura 20: Esquema mòduls OpenStack: https://es.wikipedia.org/wiki/OpenStack

Tots aquests mòduls són serveis que s'intercomuniquen entre ells per mitjà d'un sistema de cues de missatgeria anomenat RabbitMQ que es basa en l'estàndard AMQP[31]. I que ens aporta característiques com ara.

- Garantia d'entrega
- Encaminament flexible
- Clusterització
- Federació
- Alta disponibilitat
- Tolerància a fallades

A més, com s'ha comentat abans, donada la modularitat d'OpenStack podrem dissenyar l'estructura de la manera que més ens interessi i per tant distribuir-los en servidors independents, redundant aquells que necessitin més recursos o agrupant d'altres que no siguin imprescindibles. Per exemple, podrem implementar una estructura multinode on separarem nodes controladors i nodes de còmput afegint redundància de maquinari.

Un altre aspecte a tenir en compte ha de ser a distribució que es farà servir per a la implementació. Al ser OpenStack un programari de codi lliure, existeixen infinitat de distribucions possibles, però pel cas que ens ocupa, sent un projecte empresarial les opcions es limiten una mica més.

Les principals distribucions serien les següents:

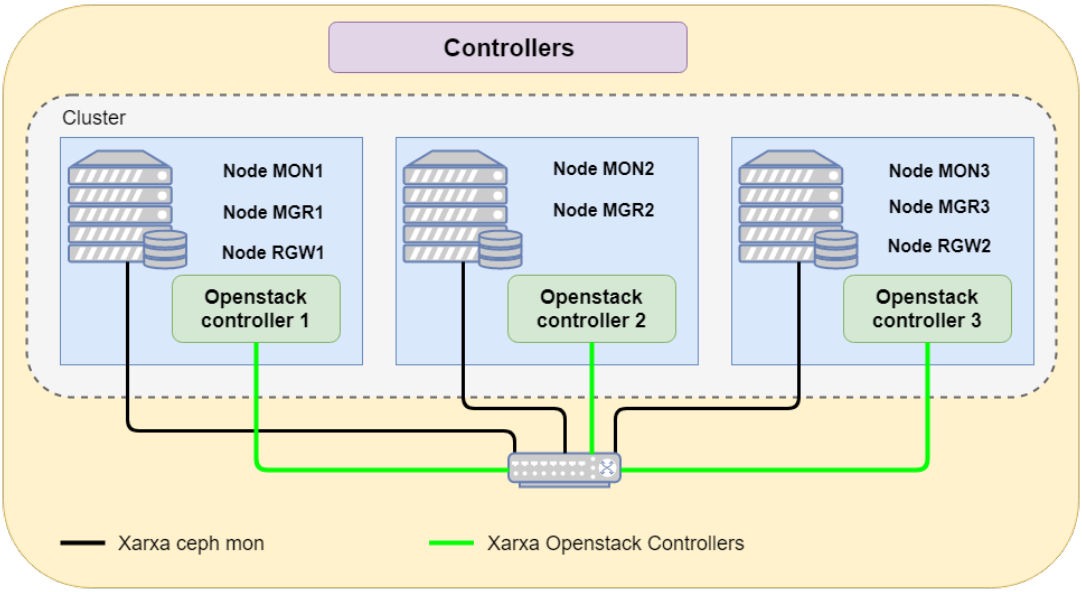
- **VMware Integrated OpenStack**: Distribució creada per VMware per a integrar OpenStacK en vSphere i fer servir un cluster ESXi. El problema és que, naturalment precisa de llicenciament de l'estructura vSphere.
- **Red Hat Openstack Platform**: És una de les distribucions més robusta i amb millor suport encara que Red Hat el subministra sota una llicència de pagament.
- **Mirantis Openstack**: És una de les millors distribucions actualment que manté una gran comunitat. Fa servir un sistema centralitzat de gestió anomenat Fuel que facilita la instal·lació i manteniment del sistema.
- **IBM Cloud Manager with Openstack**: És una distribució pròpia d'IBM per als seus clients però no té una comunitat i documentació suficient per a ser una bona opció.

A part d'aquestes, existeix una altra que és més un projecte que una distribució.

- **RDO project**: Està basada en els treballs d'una comunitat per tal de facilitar la seva instal·lació. Ofereix un instal·lador anomenat *Packstack* que ens facilitarà en la instal·lació. Aquest projecte és recomanable per a entorns de proves o la formació.

En el nostre cas, farem servir RDO project pel pilot i la distribució de Red Hat o la de Mirantis, depenent el tipus d'instal·lació per la qual ens decantem al final. En l'apartat d'implantació es definirà la infraestructura final triada segons criteris de l'empresa.

De totes maneres l'estructura ha de ser la mateixa. La idea és crear un clúster on separem els nodes de control, els de computació i els de xarxa. Pel que fa als **nodes de control**, seran la base del nostre sistema i per tant crearem un sistema d'alta disponibilitat de tipus *Actiu-Actiu* de 3 nodes que inicialment s'allotjaran en els mateixos servidors destinats als monitors de ceph (*Dell PowerEdge R640*) encara que es sotmetran al seguiment del seu rendiment per veure la necessitat de separar-los o no.



- Figura 21: Esquema controllers OpenStack.

Segons la documentació d'Openstack, aquests nodes requereixen una sèrie de recursos que es compleixen en els servidors implementats però que especificarem a continuació igualment.

Recurs	**Requeriments**	**Servidors DellEMC PE R640**
CPU	2 CPUs amb 6 cores com a mínim	2 CPUs amb 12 cores
RAM	Per entorns de producció, • 24GB mínim • 64GB per entorns grans.	16 x 8GB → 128GB
Xarxa	Per entorns de producció, • 2x 10Gbps	• 2 ports 1Gbps • 2 ports 10Gbps
Emmagatzemament	Mínim 1TB	800GB x 4 RAID5 → 2'4T

- Taula 30: Requeriments i proposta controllers OpenStack.

En cada node controlador trobarem els mòduls principals com ara *Keystone*, *glance*, *Nova management*, *Neutron Server, Horizon, Cinder Management i Swift Proxy*. Tot i que l'emmagatzemament es farà amb *Ceph* a través dels connectors natius d'*OpenStack* que fan servir l'emmagatzemament per blocs i objectes. A part d'això, també trobarem serveis de suport com ara MariaDB per bases de dades i RabbitMQ per la gestió de missatges interns.

Pel que fa a la connectivitat, aquests nodes hauran de pertànyer a la xarxa de gestió d'Openstack per comunicar-se amb els altres nodes. Una possible xarxa d'IPv4 seria *10.0.0.0/24* o d'IPv6 com ara *fd43:d7f9:fc7b:de4b::/64*

Als **nodes de computació** seran els nodes encarregats de gestionar els servidors virtuals a través de Hypervisors com KVM o QEMU. En aquests nodes ens trobarem els mòduls bàsics de computació *Nova Hypervisor*, i els agents del mòdul de Xarxes.
A part, gestionaran l'emmagatzemament efímer destinat a les instàncies. És a dir, aquell que desapareix a l'eliminar una instància. A més, com que l'emmagatzemament configurat serà amb Ceph, necessitem tenir connectivitat amb els OSD mitjançant el client Ceph.

En el projecte que estem implementant els requeriments d'aquests nodes i la proposta oferida són les següents.

Recurs	**Requeriments**	**Servidors DellEMC PE R440**
CPU	2 CPUs amb 4 cores com a mínim	2 CPUs amb 22 cores
RAM	64GB	16 x 16GB → 256GB
Xarxa	2x 10Gbps	• 2 ports 1Gbps • 2 ports 10Gbps
Emmagatzemament	2 x 500GB amb RAID1 pel SO	2 x 500GB M.2 RAID1

- Taula 31: Requeriments i proposta computació OpenStack.

Els recursos en aquests nodes són molt importants, ja que seran amb els que s'associaran a les instàncies (*tenants*). Per la qual cosa, inicialment s'implementaran 4 servidors DellEMC PowerEdge R440 en clúster d'alta disponibilitat Actiu-Actiu per tal d'oferir els recursos necessaris pels servidors virtuals.

Per tenir una idea dels recursos amb els quals començarem a treballar tenim.

Recurs	**Recursos per màquina**	**Totals x 4**
CPU	2 CPUs amb 22 cores	8 CPUs amb 22 cores = 176cores
RAM	16 x 16GB → 256GB	1TB
Emmagatzemament	100TB VM + 100TB backup	

- Taula 32: Recursos servidors computació OpenStack.

A més en el cas de la CPU i la Memòria, tenim un petit marge per sobreexplotar-los.

- En el cas de la CPU una mitja de 16:1
- En el cas de la RAM una mitja de 1.5:1
- En el cas del disc no és recomanable.

Això ens permetria obtenir uns recursos per a virtuals de fins a:

Recurs	Recursos reals	Recursos amb sobre explotació
CPU	176 cores	176cores x 16 = 2816 cores
RAM	1TB	1.5 TB
Emmagatzemament	100TB VM + 100TB backup	

- Taula 33: Recursos amb sobreexplotació servidors computació OpenStack.

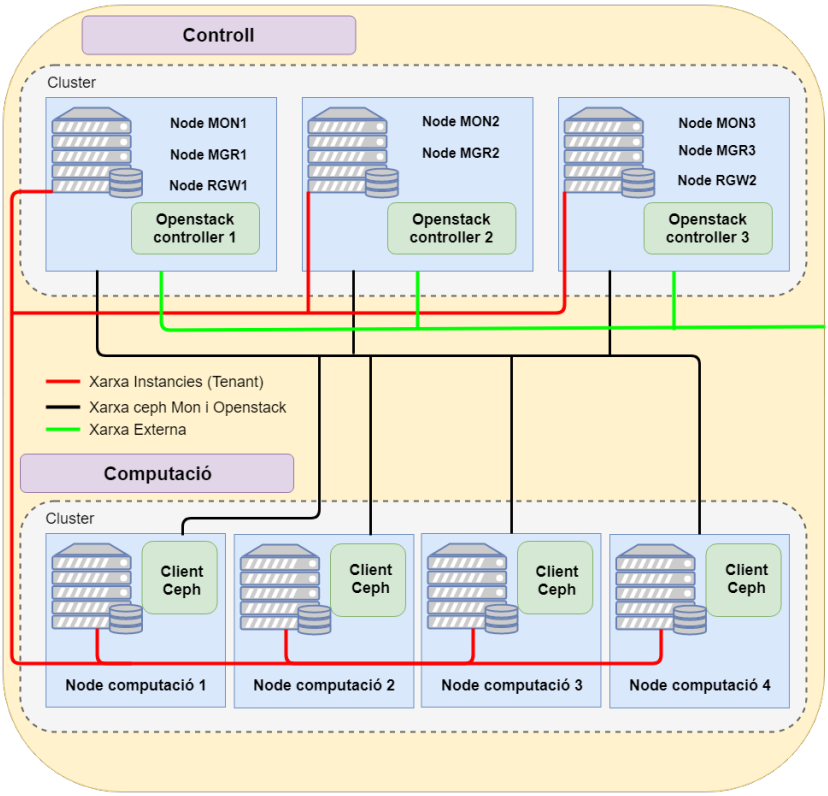
D'aquesta manera, i partint de la premissa que la intenció és arribar al 400 virtuals en aquesta infraestructura tindríem aquest nombre de recursos per virtual.

Recurs	Recursos estimats	Ràtio proposat
CPU	400 virtuals x ±4vcores de mitjana	1600vcores / 176 = 9,09 de ràtio
RAM	400 virtuals x ±3 GB RAM	1.2TB / 1TB = 1.2 de ràtio
Emmagatzemament	100TB VM + 100TB backup	

- Taula 34: Recursos estimats servidors computació OpenStack.

En qualsevol cas, davant la necessitat d'ampliació de recursos, tant Openstack com ceph ens permetrien afegir nodes de manera senzilla i transparent.

Pel que fa a les xarxes necessàries en els nodes de computació, haurem d'implementar la xarxa de gestió per la comunicació amb els nodes de control i xarxes. A més haurem d'implementar la xarxa per a la comunicació entre instàncies que s'implementarà amb túnels VxLAN[32]. Per últim, aquests nodes haurien de pertànyer també a la xarxa pública de Ceph, ja que hauríem d'instal·lar ceph client per tal de fer servir l'emmagatzematge del cluster Ceph.



- Figura 22: Esquema OpenStack.

Una possible xarxa d'IPv4 seria *10.0.1.0/24* o d'IPv6 com ara *fd43:d7f9:fc7b:de4c::/64*

Per últim, els **nodes de xarxa** que seran els encarregats de gestionar les xarxes virtuals, assignació d'IPs als virtuals, etc mitjançant els agents de *Neutron com ara ''neutron openvswitch agent''*, mòdul encarregat de les xarxes. Es planteja separar-los dels nodes de control i còmput però no són nodes que necessitin molts recursos de computació i per tant, inicialment aniran allotjats als nodes de control amb redundància. Això implica afegir dues xarxes al dispositiu. Una per la gestió de xarxes de les instàncies i una per a la connectivitat amb internet.

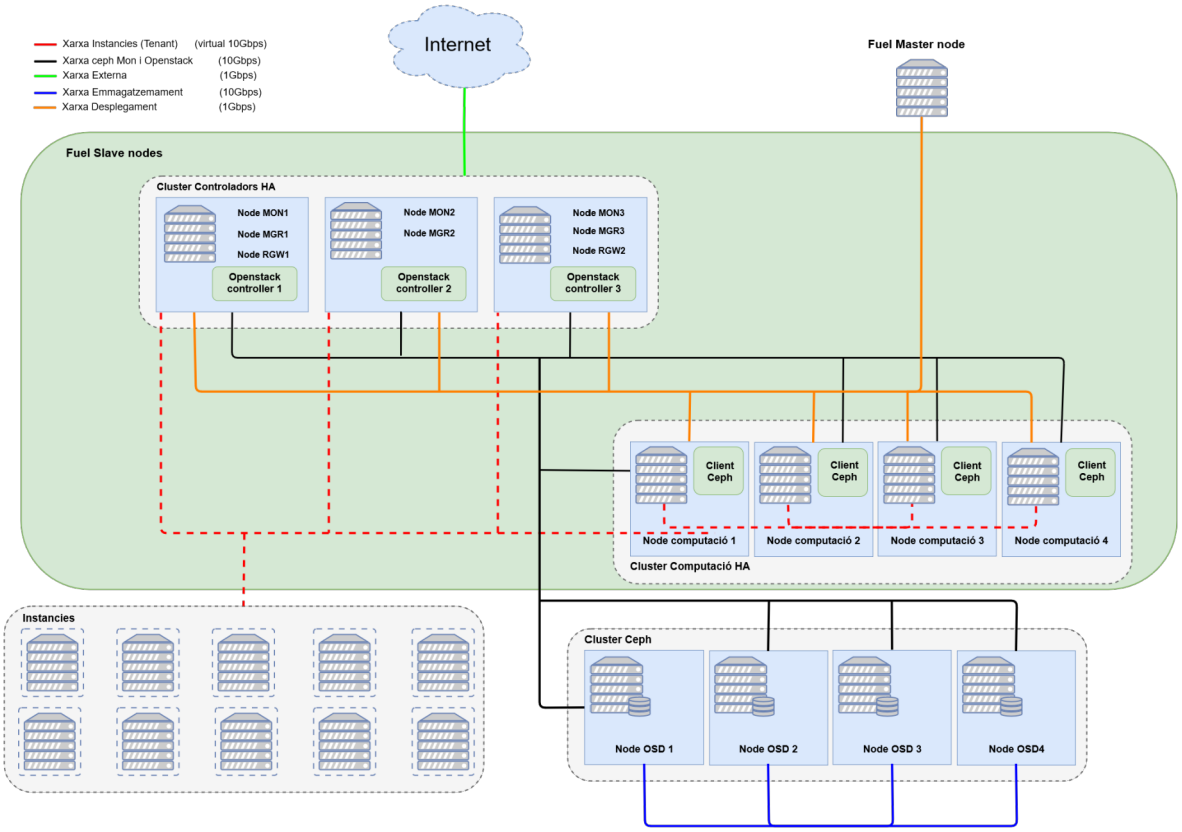
A part de aquests nodes, tant si optem per la instal·lació de Mirantis amb Fuel o la de Red Hat, en ambos casos es necessitarà un servidor addicional per a la instal·lació i gestió. En el cas de Mirantis necessitarem un servidor Fuel Master. Aquest ens ajudarà en la instal·lació i configuració de diferents entorns d'Openstack. Una vegada instal·lat, també pot configurar i comprovar diferents configuracions de xarxa, comprovar la interconnexió entre els diferents mòduls d'Openstack i gestionar l'escalabilitat del sistema permetent-nos afegir o treure nodes del sistema.
En aquest sistema la resta de nodes (Control, computació, xarxa, etc.) serien els nodes Fuel esclaus.

Aquest node Fuel Master requereix uns requeriments mínims que s'indiquen a continuació.

Recurs	Requeriments	Servidors DellEMC PE R240
CPU	1 CPUs amb 4 cores com a mínim	1 CPUs amb 8 cores
RAM	4GB	2 x 8GB → 16GB
Xarxa	10Gbps	• 2 ports 1Gbps • 2 ports 10Gbps
Emmagatzemament	Logs 40%(20GB x 7nodes) 140GB SO 60% 350GB x 0.6 = 210GB **total** **(140x2.5) 350GB**	2 x 500GB M.2 RAID1

- Taula 35: Requeriments node desplegament OpenStack.

Finalment, una vegada tenim l'estructura dissenyada podem mostrar un esquema de com quedaria la nostra estructura de cloud.



- Figura 23: Esquema Ceph + OpenStack.

4.3 Escenari pilot

El pilot proposat es portarà a terme amb l'estructura següent:

- 2 servidors *DELLEMC PowerEdge R330* amb SO *ESXi 6.5*
- 3 nodes virtuals amb instal·lació minimal de *CentOS 7*

Farem servir 3 per als nodes OSD, 1 pel node Controlador, 1 pel node de comput i 1 pel node de xarxa.

Hostname	dirección IP
controller	192.168.130.141
compute	192.168.130.142
network	192.168.130.143

- Taula 43: Llistat IPs pilot OpenStack.

(*El procés d'instal·lació, es poden trobar a l'annex VII*)

5. Sistemes de seguretat

En qualsevol classe de xarxa de l'actualitat, els usuaris mantenen tota mena d'informació important amb la creença que seguirà allà on sigui per sempre i passi el que passi. Però, perquè això sigui cert, els departaments de sistemes que administren aquestes xarxes han de dissenyar i gestionar sistemes de seguretat, tant actius com passius, per a tractar de minimitzar les possibles fallades dels serveis o perdudes d'informació. Aquests sistemes han de gestionar els accessos als serveis que s'ofereixen i aïllar-los de la part pública de la xarxa quan sigui necessari. A més, s'han de protegir les dades que gestionen aquests serveis amb sistemes de còpia per mantenir diferents versions de les dades i obtenir així la capacitat de recuperar-les en qualsevol moment. Per altra banda, tots aquests sistemes han de ser monitorats amb sistemes que ens permetin conèixer en tot moment l'estat de la xarxa, el maquinari, els serveis, etc, i que ens avisin en cas de necessitat d'actuació. Per últim, però igual d'important, s'han d'implementar sistemes de continuïtat de subministrament elèctric que permetin fer front a caigudes de tensió temporals tant per la preservació dels serveis, com per la protecció de dades i maquinari.

En l'apartat de seguretat es dissenyaran els aspectes de la seguretat comentats, com ara l'accés als serveis, la gestió de còpies de seguretat de les dades, el monitoratge dels diferents components de la xarxa tant pel que fa al maquinari com de programari, la gestió d'alertes crítiques i els sistemes de continuïtat de subministrament. Per aquesta raó, cadascuna d'elles serà explicada en una secció diferent.

5.1 Tallafocs

En seccions anteriors s'ha definit la implantació de tallafocs per tal de protegir les diferents xarxes que s'havien dissenyat.

Xarxa LAN

Tipus de maquinari	Marca / Model
Encaminador porta d'enllaç / Tallafocs	Sonicwall TZ400
Encaminadors VPN (IPSEC) / Tallafocs	Sonicwall TZ300

- Taula 36: Encaminadors xarxa LAN.

Xarxa pública

Tipus de maquinari	Marca / Model
Tallafocs xarxes /23 - /24	WatchGuard Firebox M570 Firewall
Tallafocs xarxes /26 - /27	WatchGuard Firebox M370 Firewall

- Taula 37: Encaminadors xarxa pública.

En tots dos casos, els tallafocs són solucions UTM[18] que es caracteritzen per poder treballar en diferents capes del TCPIP i per tant protegiran les xarxes que gestionen en aquestes capes segons la llicència obtinguda. Els dispositius UTM ofereixen per tant, diferents solucions de seguretat en un sol dispositiu. Solucions com ara tallafocs de xarxa, antivirus, antispam, detecció d'intrusions (IDS) o filtratge de continguts. Cadascuna d'aquestes solucions actua en una capa del sistema TCPIP com per exemple el tallafocs de xarxa que actua en capa d'enllaç o el filtratge de continguts, antivirus o antispam que actua en capa d'aplicacions. Però totes aquestes solucions han d'estar equilibrades amb la quantitat i el tipus de tràfic que gestionaran atès que, un excés de protecció i/O filtratge de contingut suposarà un major consum de recursos del dispositiu UTM que pot arribar a convertir-ho en un coll d'ampolla i per tant un problema en el correcte funcionament de la xarxa i dels serveis.

En les dues xarxes exposades es plantejarà una opció diferent, ja que les xarxes que es gestionaran tenen una funcionalitat diferent. En el cas de la xarxa privada, aquesta estarà formada per usuaris que seran consumidors de tràfic, per tant, la seguretat ha d'estar orientada a controlar el contingut que entra a la xarxa. Per aquesta raó farem servir un dispositiu sonicwall TZ400 UTM per la xarxa principal que disposarà de llicència «TotalSecure Advanced Edition» que comportaran solucions com ara:

TotalSecure Advanced Edition
Tallafocs d'última generació amb serveis de seguretat avançats de porta d'enllaç.
Detenció de les amenaces conegudes i desconegudes com el ransomware, virus, spyware, cus, troians i d'altres malwares.
Seguiment de l'activitat de xarxa amb Application Intelligence and Control.
Administració de l'accés a contingut web amb Content Filtering Service.
Assistència en qualsevol moment amb el suport 24×7.

- Taula 38: Llicencia sonicwall: https://www.sonicwall.com/es-mx/products/firewalls/security-services/security-bundles/

A més, de dos tallafocs model sonicwall TZ300 que a més de protegir les oficines externes, gestionaran les dues VPN que integraran les dues oficines en aquesta xarxa.
Més informació dels dispositius Sonicwall triats en l'annex II.

Per altra banda, la xarxa pública estarà integrada per servidors que publicaran continguts i per tant seran productors de tràfic. En aquest cas la seguretat estarà orientada principalment en controlar els accessos externs als serveis publicats (HTTP, FTP, POP3, IMAP, SMTP...) i a denegar qualsevol altre accés. Per aquesta tasca s'implementaran dos dispositius WatchGuard Firebox MXXX i es contractarà inicialment la llicència «Basic Scurity Suite» encara que es farà un seguiment durant el termini de llicència per analitzar si el seu ús és necessari o no. La tria d'un model M570 per a les xarxes /23 i /24, i M370 per a les xarxes /26 i /27 ve determinat pel rendiment de cada dispositiu.

Més informació dels dispositius WatchGuard en l'annex III.

Características de Security Suite

Funcionalidades y servicios	Basic Security Suite
Intrusion Prevention Service (IPS) Escanea todos los puertos y protocolos para proporcionar protección en línea contra ataques	✓
Application Control Bloquea aplicaciones peligrosas, no autorizadas e inapropiadas para garantizar la seguridad y la productividad	✓
WebBlocker (filtrado de contenido y URL) Proporciona filtrado de contenido y URL para bloquear material no apropiado y malware	✓
spamBlocker (filtro de correo no deseado) Bloquea los correos no deseados, sin importar el idioma, el formato o el contenido del mensaje	✓
Gateway AntiVirus (GAV) Utiliza firmas y sofisticados análisis heurísticos para detener amenazas	✓
Reputation Enabled Defense (RED) Utiliza la puntuación de reputación para garantizar una navegación web más rápida y segura con capacidades especiales de detección de botnets	✓
Network Discovery Mapa visual de todos los nodos de la red para identificar fácilmente el riesgo y el comportamiento sospechoso	✓

-Taula 39: Llicencia WatchGuard: https://www.watchguard.com/es/wgrd-products/total-security-suite
(S'han eliminat les dades referents a d'altres llicencies.)

Referent a les polítiques que s'implementaran ene aquests tallafocs, dependran del tipus de serveis que s'allotgin en els servidors, però hi hauran una sèrie de polítiques fixes d'accés per a les IPs públiques de sortida de la xarxa privada. Com havíem comentat anteriorment, la xarxa privada tindrà tres sortides públiques per tal de diferenciar els perfils d'accés als servidors. Una Primera sortida serà per a les connexions WIFI dels dispositius

invitats, aquests no tindran accés als servidors més enllà dels serveis publicats per tothom. Una segona IP serà destinada als llocs de treball i per tant tindrà accés a diferents serveis privats com ara intranets, bases de dades o administracions web.
Per últim, la tercera IP serà destinada als serveis tècnics i de sistemes i per tant tindran accés il·limitat a les diferents xarxes i servidors per tal d'administrar-les.
Al marge d'aquestes, les polítiques de les xarxes noves s'aniran implementant amb la creació de nous servidors/serveis. Encara que es poden plantejar polítiques genèriques per a serveis com HTTP, HTTPS, FTP, SMPT, etc en casos on la creació i eliminació de servidors hagi de ser molt dinàmica. En el cas de xarxes que ja estiguin en producció s'hauran d'analitzar les polítiques ja existents per fer la migració als dispositius nous o eliminar-les en cas que deixin de ser útils.

5.2 Backups

5.2.1 Nivells de còpia

Un altre aspecte important en la seguretat de l'empresa, i mes en una empresa que ofereix serveis d'internet basats en continguts com ara webs, correu, bases de dades, etc, és evitar la perduda de dades sota qualsevol circumstància. Per aquesta raó els diferents departaments, i sobretot els departaments de sistemes hauran de seguir una sèrie de normes i precaucions per evitar-ho el màxim possible. Tot i això, sempre es poden produir casos com ara atacs informàtics, errors humans o fallades del programari o maquinari que provoquin perdudes de dades. Per a aquests casos s'hauran de realitzar còpies de seguretat (Backups) de les dades importants que permetran la recuperació d'aquestes a diferents nivells en cas d'incidència.

A l'hora de realitzar les còpies de seguretat hem de tenir clar que és allò que volem copiar, ja que segons el que volem copiar haurem d'actuar a un nivell o d'altre.

- **Blocs**: Permet copiar volums complets
- **Sistema de fitxers**: Permet copiar fitxers
- **Aplicacions**: Permet copiar dades específiques d'una aplicació

Nivell de blocs

En aquest nivell, el sistema copia els blocs del disc al marge del seu contingut de manera que ens permet copiar volums sencers. A més, ens permet copiar únicament els blocs modificats a partir d'una còpia inicial, fet que disminuirà la mida de les còpies, doncs en el cas de grans fitxers modificats no requerirà la còpia de tot el fitxer sinó que només es copiaran aquells blocs que hagin estat modificats.

Aquest sistema de còpies és recomanable en còpies de volums sencers com ara LUNs iSCSI, CEPH RBD, LVM, para sistemes amb fitxers molt grans i dinàmics com ara bases de dades i para còpies en sistemes remots, ja que minimitza notablement el tràfic a traspassar aportant una reducció de l'amplada de banda consumit i del temps de còpia. Com a contrapartida, en un sistema de còpia a nivell de blocs, la recuperació de fitxers és molt més costosa, atès que els fitxers s'han de reconstruir a partir dels blocs restaurats. A més, en còpies en local no sempre es recomanable, ja que aquest sistema requereix comparacions entre blocs per determinar que s'ha de copiar que poden arribar a ser més costoses que d'altres sistemes de còpia com ara els sistemes a escala de fitxers.

Nivell del sistema de fitxers

En aquest nivell, el sistema copiarà els fitxers/directoris directament del sistema de fitxers mitjançant un agent instal·lat al SO[19], fet que ens permet fer còpies més ajustades, ja que no hem de copiar tot un volum sinó que podem determinar quin/s directoris o fitxers volem copiar evitant així copiar, fitxers del sistema operatiu, temporals o aplicacions que no necessitem. Aquest sistema està orientat a fer còpies de continguts, com ara webs, espais FTP, espais de compartició de documents, configuració, etc on la localització de les còpies serà preferent locals, ja que permet una gestió de l'espai a copiar eficient però pot arribar a consumir una amplada de banda elevada en el cas de fitxers grans.

Pel que fa a la recuperació de dades, permet recuperacions granulars i per tant molt més àgils i ràpides.

Nivell de aplicacions

En aquest nivell, les aplicacions permeten exportar les dades que manté guardades en fitxers de manera que podran ser importats posteriorment en cas necessari. Aquesta exportació/importació es farà sempre a partir de funcions específiques de les mateixes aplicacions o d'aplicacions alternatives sent recomanable que siguin en l'àmbit local i amb programació temporal, ja que poden arribar a consumir molts recursos tant de maquinari com d'amplada de banda. Per altra banda, aquest tipus de còpies acostumen a realitzar-se en local i es combinen amb altres sistemes de còpia com ara sistemes a escala de fitxers.

Pel que fa a la seva recuperació consistirà en el pas invers, ja que s'haurà d'indicar el fitxer a restaurar i l'aplicació importarà les dades contingudes en el.

Un exemple d'aquest tipus de còpies serien els bolcats de les bases de dades com ara MySQL amb la comanda «*mysqldump*» que genera un fitxer de text on es guarden les instruccions SQL necessàries per a replicar la base de dades completa o les taules que es vulguin copiar.

Relacionats amb tots tres nivells de còpia, també es poden definir tres subnivells amb els quals es pot indicar la relació que existirà entre còpies de manera que podem indicar la necessitat de copiar o no dades que ja existeixen en còpies anteriors i no han estat modificades. Aquests nivells de còpia són:

Tipus de copia	Es copiarà...
Completa	Tot allò que s'hagi definit en la configuració de la còpia.
Diferencial	Tot allò que hagi estat modificat des de l'última còpia completa.
Incremental	Tot allò que hagi estat modificat des de l'última còpia de qualsevol tipus.

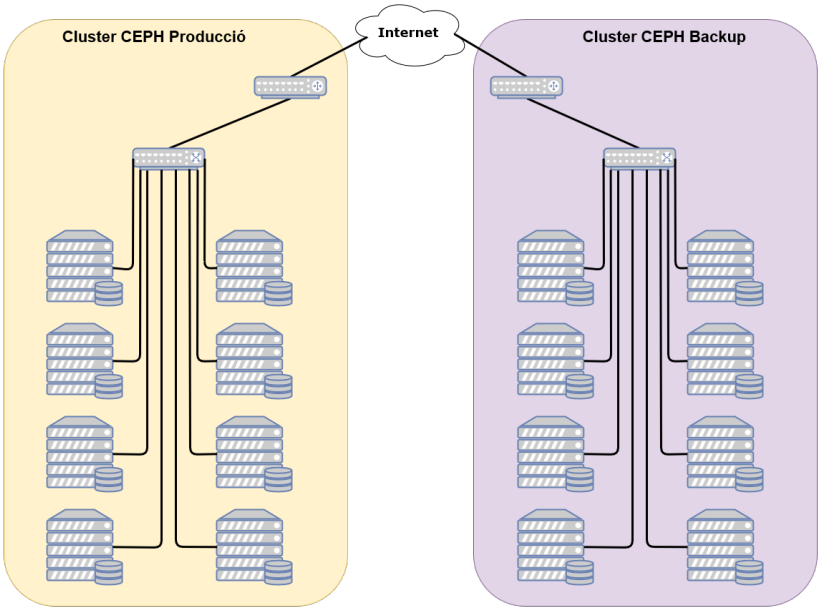
- Taula 40: Tipus de copia

5.2.2 Sistemes de copia

Una vegada explicats els diferents tipus de còpies disponibles podem definir els sistemes de còpies que s'implantaran en les infraestructures dissenyades anteriorment.

Inicialment es diferenciaran entre dues motivacions de còpia diferents. En primer lloc s'hauran de realitzar les còpies, necessàries per a establir un pla de continuïtat de l'empresa on s'haurà de replicar a escala de bloc les principals estructures. Per un costat tot el sistema d'emmagatzemament CEPH haurà d'estar replicat per tal poder restaurar el sistema sencer en cas d'incidència. Aquesta rèplica es pot realitzar de diferents formes tenint en compte capacitats, espai físic i pressupost.

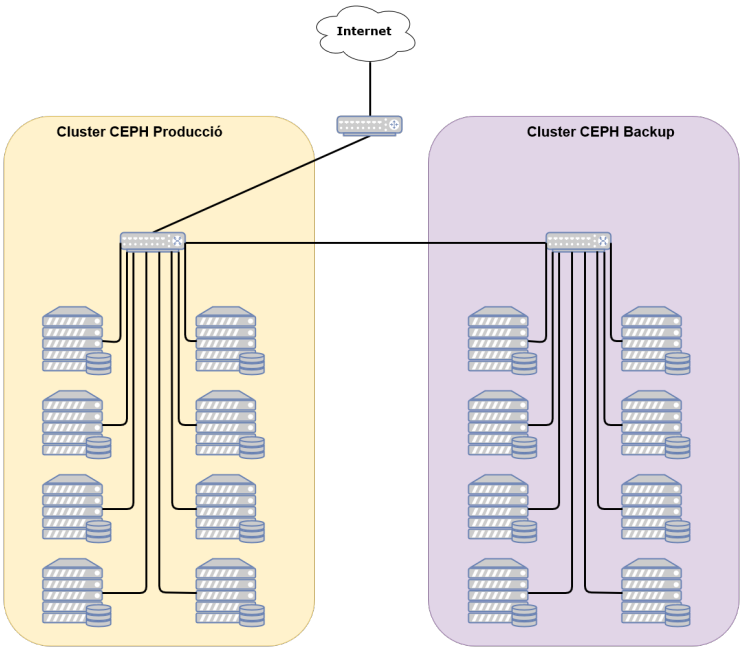
La manera més segura seria la formada per dos sistemes de cabines físiques amb clusters[20] CEPH independents on s'anirien fent còpies incrementals a escala de bloc del pool CEPH en producció al pool CEPH de backup situat en un CPD remot. Aquesta opció és la més costosa, ja que es duplica el maquinari i depèn d'una amplada de banda suficient per a sostenir el tràfic que provoquin les copies, encara que en ser incrementals i a escala de bloc estarien optimitzades.



- Figura 24: Esquema de còpia cluster Ceph 1

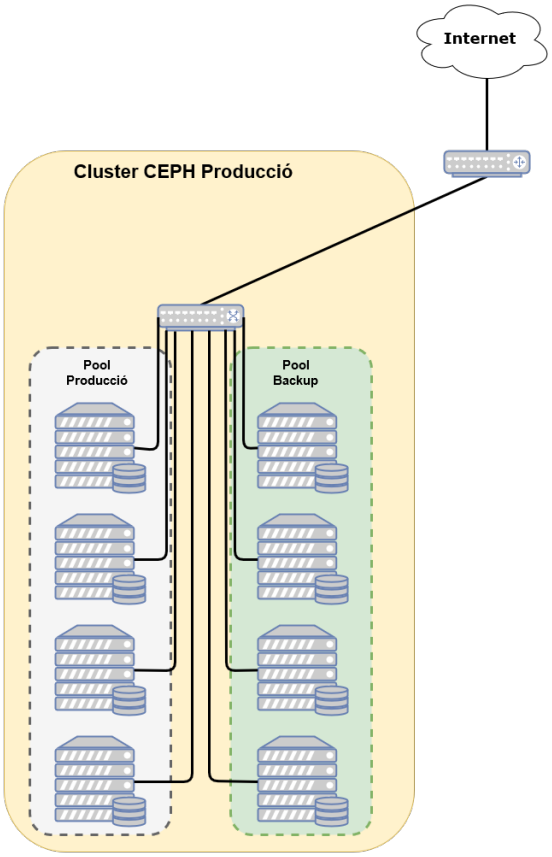
Per altra banda, com es pot veure en el gràfic, aquest sistema ens aporta la capacitat de recuperació davant d'un desastre a l'estructura de producció, ja que el sistema de backup estaria separat del de producció més de 30Km en el cas dels nostres CPDs.

Una altra opció una mica més rendible, estaria formada per dos clústers CEPH independents com en el cas anterior, però situats en un mateix CPD i per tant amb amplada de banda suficient per realitzar còpies de gran mida sota infraestructures 10GBASE-T. En aquest cas, el cost més gran l'aportarà el maquinari duplicat encara que el maquinari destinat a backup es podrà optimitzar fent servir sistemes més econòmics.



- Figura 25: Esquema de còpia cluster Ceph 2

Aquesta solució ja no ens aporta la protecció davant d'un desastre general a les infraestructures de producció, ja que estarien en el mateix CPD, però sí que ens aporta la protecció en cas de desastre general del clúster en producció. Aquesta és la solució que es tria i per tant la que s'implementarà.



- Figura 26: Esquema de còpia cluster Ceph 3

Com a última opció es pot replicar l'opció anterior però fent servir un únic maquinari per a l'estructura CEPH de forma que es configuren dos pools CEPH independents en la mateixa estructura física i es farien les còpies d'un a l'altre també de forma incremental. A més, es realitzaran copies a escala de bloc, de virtuals sencers de serveis principals de l'empresa.

En aquesta última solució, es perden les proteccions davant problemes generalitzats del maquinari però tenim una estructura més assequible econòmicament. Hem de recordar que el sistema d'emmagatzematge sota CEPH és un sistema distribuït que aporta entre altres coses, redundància de dades entre els seus OSD cosa que ens permetrà minimitzar els problemes davant caigudes de nodes puntuals.

Per últim, i en casos especials, es realitzaran còpies a escala de fitxers d'aplicacions web, còpies de bases de dades i documentació de l'empresa per a una restauració rapida. Aquestes còpies es realitzaran en els sistemes allotjats en la xarxa de backup i per tant en maquinari diferent del de producció. A més, es recomana realitzar una còpia remota i encriptada de dades importants com ara bases de dades, aplicacions i/o documentació com a part del pla de continuïtat en cas de desastre.

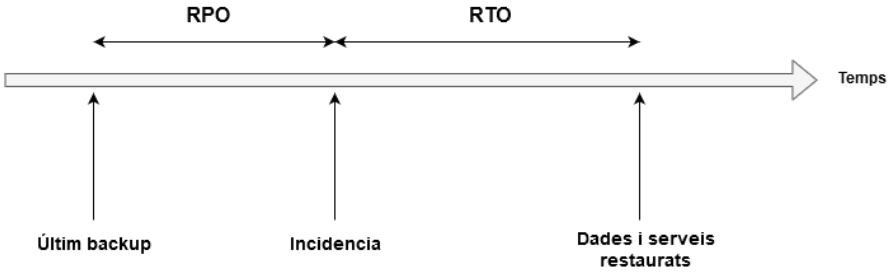
Per altra banda, s'hauran de realitzar totes aquelles còpies contractades pels clients i es classificaran segons el seu abast. Es podran fer copies dels virtuals sencers per tal de poder aixecar-ho completament en cas d'incidència, es podran fer còpies de fitxers o directoris concrets com ara poden ser espais web, espais FTP, espais de compartició de documents o backups de bases de dades.

En tots els casos s'hauran d'establir unes condicions de còpia personalitzades que tindran a veure amb la capacitat de disc disponible, l'amplada de banda en cas de còpies remotes i l'abast en el temps que es vol aconseguir. Tot i això, la recomanació en la majoria dels casos serà la següent:

Tipus de copia	**Programació**	**Històric**
Completa	Mensual	Últims dos mesos
Diferencial	Setmanal	Últim mes
Incremental	Diària	Últimes dues setmanes

- Taula 41: Proposta tipus de copia

Un altra aspecte a tenir en compte a l'hora de dissenyar les polítiques de còpia que s'hauran de fer, és definir quin grau de perduda de dades s'està disposat a assumir al tenir que recorre a la recuperació de dades davant un desastre. Per a definir aquest grau existeixen dos paràmetres que s'hauran d'especificar. Aquests paràmetres són RPO (*Recovery Point Objective*) i RTO (*Recovery Time Objective*). La primera fa referència a la quantitat de dades que es generen entre còpies. Això vol dir que si la còpia es fa cada dia a les 24.00 de la nit i el servidor cau a les 15.00 de l'endemà, totes les transaccions generades fins a aquell moment no s'haurien registrat en cap còpia i per tant s'haurien perdut. Per altra banda, RTO fa referència al temps de recuperació que es pot assumir d'ençà que es produeix la incidència fins que es recupera el servei sense que hi hagi una afectació a la continuïtat del negoci.



-Figura 27: Esquema RPO / RTO

La veritat és que definir uns valors genèrics per a RPO i RTO resulta una mica difícil, ja que per norma general, la varietat de clients i serveis sol ser molt alta i cada un requereix un nivell de seguretat diferent. La majoria de clients poden assumir un RPO de fins a 24 hores atès que acostumen a ser allotjaments amb web i correu no gaire crítics. Per aquesta raó es proposen copies amb freqüència diària. Tot i això, existeixen alguns clients i servidors propis que requereixen un RPO més petit, que en alguns casos arriba a còpies incrementals cada 30 minuts.

En el cas de l'RTO, també varia depenent del crític que sigui el servei. En casos normals definirem un RTO al voltant d'una hora per tenir un servidor operatiu amb l'última còpia aixecada, depenent del tipus de servidor. En els casos més crítics el sistema de backup ens permetrà aixecar l'última còpia com si fos la màquina pròpia en qüestió de minuts.

En qualsevol cas, el software triat serà el següent pels diferents casos.

Tipus de copia	**Openstack + CEPH**	**VMware + LUNs**
Bloc /volum	• Cinder-backup (Ceph backup driver) • Ceph RBD mirroring • Ceph snapshots	• Veeam Backup 9.5 • Vm snapshots
Fitxers	• Bacula backup	
Aplicació	• Cada aplicació generarà les seves còpies • S'inclouran en les copies a nivell de fitxers	

- Taula 42: Comparativa tipus de copia.

(*Els tipus de backup s'expliquen mes profundament a l'Annex XIII*)

5.3 Monitoratge

Pel que fa al monitoratge dels diferents sistemes, aplicarem els programaris i polítiques necessàries per a l'obtenció de les dades que ens permetin portar a terme un manteniment i gestió òptims dels sistemes. El monitoratge dels sistemes dividirem en dos sectors, diferenciant així el monitoratge de les infraestructures de CEPH per un costat i de la resta de servidors per l'altre, siguin propis o de clients, físics o virtuals. Aquesta diferenciació vindrà donada per la necessitat de l'obtenció de dades diferents en un cas i l'altre. Mentre que als servidors de virtualització, hostatge de serveis i virtuals de clients, les dades que volem obtenir estaran relacionades amb el consum de recursos com ara, tràfic, CPU memòria RAM, ús de disc i en casos especials com poden ser alguns servidors de correu, el control de cues, missatges enviats i rebuts, spam rebut, etc, a la infraestructura CEPH voldrem tenir un control molt més ampli i específic, atesa la seva importància. Per aquest motiu haurem de monitorar més a fons tota l'estructura per obtenir dades de l'estat dels diferents nodes, estat dels pools, estat de la xarxa, latències, rendiments, etc.

5.3.1 Estat inicial

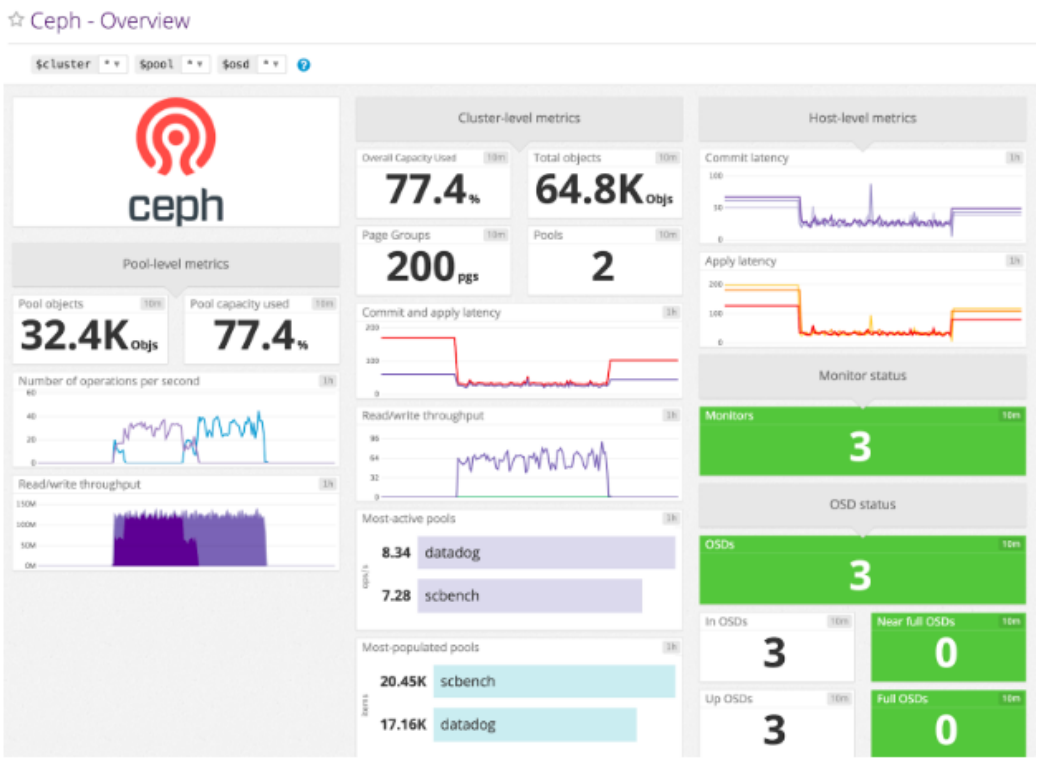
Inicialment, a l'empresa s'està fent servir principalment Cacti pel monitoratge dels servidors dels quals s'obtenen gràfiques de consum de tràfic dels diferents dispositius de xarxa, consum de RAM, CPU i disc. A més, en casos concrets com ara servidors Exchange, o de bases de dades es fan servir els agents per obtenir dades més específiques del servei que ofereixen. Tot i ser útil, Cacti és un software obsolet i cada vegada més limitat. Per aquesta raó es proposa canviar-ho per altres opcions més actuals.

5.3.2 Monitoratge d'infraestructures CEPH.

Pel monitoratge de CEPH es contemplen dues alternatives que integren el monitoratge que necessitem per a CEPH.

- Datadog: https://www.datadoghq.com/

  Aquesta solució SaaS[24] aporta una integració amb més de 250 sistemes, aplicacions i/o serveis com ara CEPH, ja sigui amb RedHat Ceph Storage com Amazon S3, o openstack a través d'agents. A part d'això, ofereix uns panells d'informació interactiva en temps real que es poden personalitzar per adaptar-los a les necessitats de l'usuari.



- Figura 28: Ceph overview: https://www.redhat.com/en/blog/infrastructure-monitoring-service

Com a inconvenient, té que és necessari el pagament d'una quota mensual per al seu ús segons el tipus de servei, nombre de hosts, aplicacions, etc que van des dels 18$/host/mes o els 36$/app/mes.

Es descarta aquesta opció pel seu preu elevat i s'optarà per programari opensource.

- Grafana + prometheus: https://grafana.com/ https://prometheus.io/

Aquest parell de programaris ens aporten el necessari per a l'obtenció de les dades que hem comentat abans sobre Ceph, ja que es complementen. Primerament tenim Prometheus, que es basa en un programari totalment opensource de monitoratge i alertes de sistemes. Aquest programari és l'encarregat d'obtenir les dades dels sistemes a monitorar de forma periòdica i emmagatzemar-los pel seu posterior tractament.

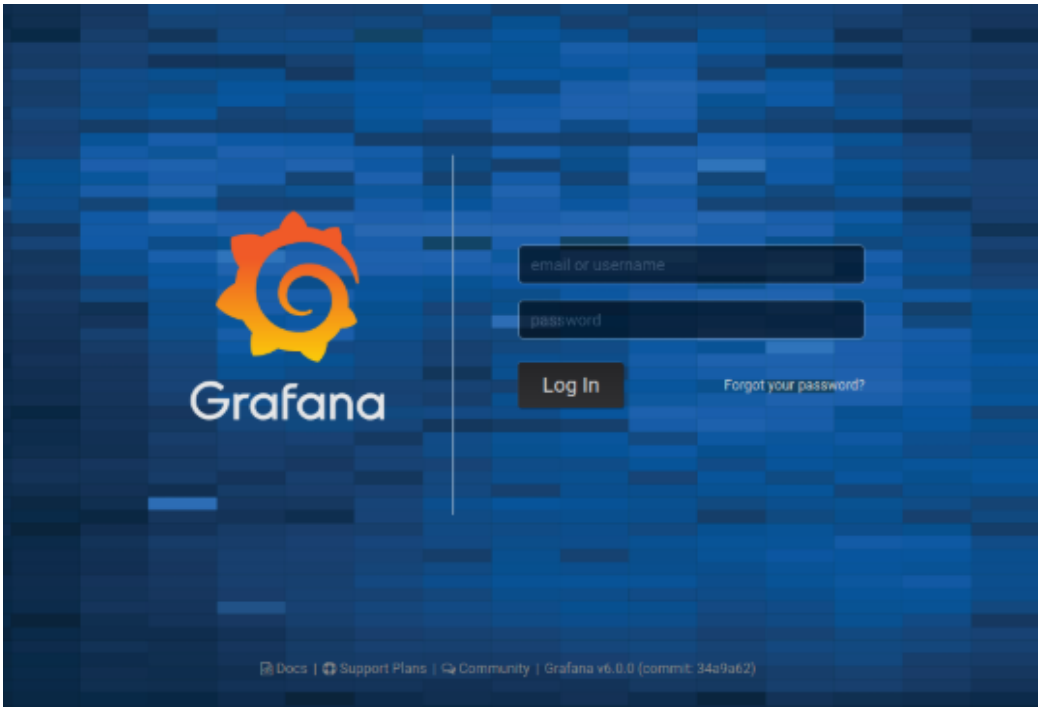
L'estructura de Prometheus consisteix principalment en un servidor central, uns agents anomenats «exporters» que permetran l'obtenció de dades per part del servidor i una gestió d'alertes. Aquest es muntarà en un servidor virtual sota sistema operatiu CentOS 7 amb connectivitat amb els diferents nodes de CEPH i OpenStack.

Un avantatge també de Prometheus és que té una comunitat molt extensa i activa al seu darrere amb diferents canals com ara IRC, llistes de mailing, canal de Twitter i una gestió de versions i suport a través de GitHub.

- https://prometheus.io/community/

A més es pot trobar moltíssima documentació tant a la seva web com als diferents canals d'informació.

Per altra banda, es farà servir Grafana per a la representació de les dades obtingudes per Prometheus. Aquest programari també és opensource i permet gestionar la informació en forma de panells de gràfiques generades en temps real i personalitzables. L'accés a aquests panells es fa a través de navegador web, fet que ens permetrà accedir-hi des de qualsevol lloc sempre que estigui habilitat al firewall.

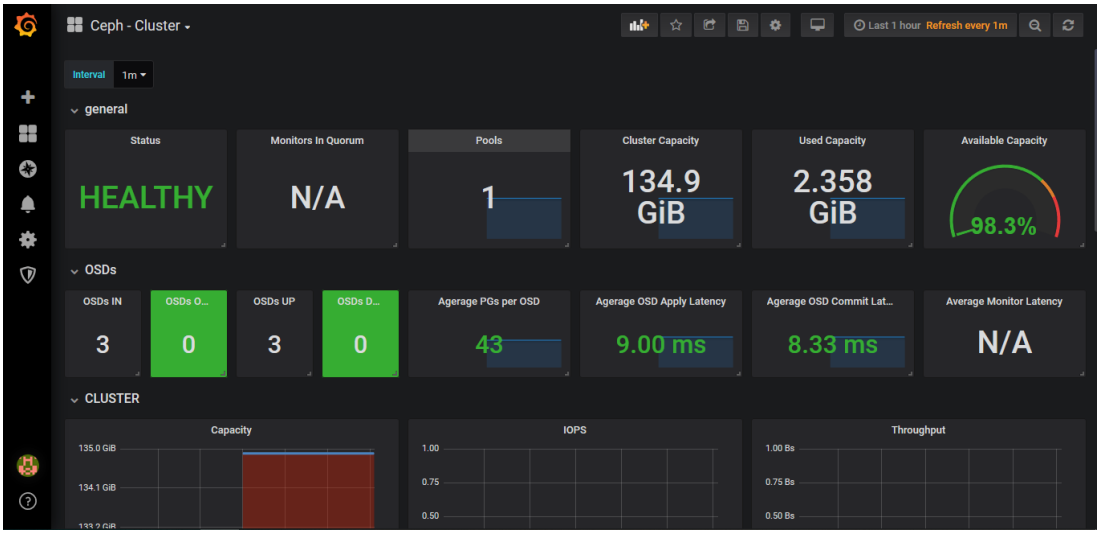


- Figura 29: Grafana login. Imatge extreta d'una captura del sistema instal·lat al laboratori.

Una vegada instal·lat, Grafana és molt configurable a escala de fitxers de configuració però també ens permet configuracions especials i personalitzar l'entrada de dades a través de connectors que es poden instal·lar.

- Figura 30: Grafana Prometheus stats. Imatge extreta d'una captura del sistema instal·lat al laboratori.

A més, com passa en Prometheus, té darrere una gran comunitat i molta documentació que ens ajudarà a ajustar-ho a les necessitats dels nostres sistemes pel monitoratge de CEPH mitjançant plantilles generades per la comunitat.



- Figura 31: Grafana Ceph cluster. Imatge extreta d'una captura del sistema instal·lat al laboratori.

El programari de Grafana es muntarà inicialment en el mateix servidor destinat a Prometheus però s'analitzarà l'opció de separar els dos programaris en cas de notar un rendiment baix. En qualsevol dels casos, només serà accessible per part de la porta d'enllaç destinada a la VLAN de sistemes comentada en l'apartat de disseny de la xarxa interna.

5.3.3 Monitoratge de servidors de virtualització i d'allotjament de serveis

Inicialment, pel monitoratge de la resta de servidors, s'havia optat per mantenir la mateixa estructura que hi ha actualment tot i haver-hi la possibilitat de canviar el programari de Cacti per Nagios, o Cacti i Zabbix per pandora FMS. La primera opció es va desestimar atès que la millora que s'obtenia era mínima comparat amb la feina de migrar el sistema actual. Per altra banda, la segona es va plantejar com valida tot i la migració però l'opció opensource era molt limitada i la versió de pagament tenia un cost molt elevat per a la nostra estructura.

Finalment, s'ha optat per aprofitar l'estructura de monitoratge que es farà servir per a OpenStack i CEPH i es configuraran els diferents servidors i aplicacions per a ser monitorades a través de Prometheus i Grafana.

5.4 Sistemes d'alertes

Inicialment, aquesta tasca és gestionada a través del programari zabbix, el qual s'encarrega de realitzar comprovacions de connectivitat amb els serveis estàndards dels servidors que es volen controlar de manera que en cas de caiguda s'enviarà un missatge d'alerta a la persona encarregada del seu manteniment. Aquest missatge serà de correu o SMS, depenent del grau d'importància del servidor/servei. Actualment es controlen només servidors propis i els d'aquells clients que hagin contractat l'assistència 24 h. Els serveis controlats solen ser els estàndards de qualsevol servidor d'allotjament de serveis d'internet (web, correu i BD).

Aprofitant el sistema implantat pel monitoratge en l'apartat anterior, es planteja retirar zabbix com a servei d'alertes i gestionar-les a partir d'ara amb el mòdul AlertManager que implementa Prometheus. Aquest sistema d'alertes està dividit en dues parts. Primerament el servidor Prometheus on es generen les polítiques d'alertes, i per l'altre, el mòdul AlertManager que les gestiona, agrupa i envia els missatges per diferents canals com ara E-mail, Slack, Pager Duty o SMS.

En la configuració del nostre sistema d'alertes, continuarem amb el mateix sistema de classificació d'alertes que farà servir inicialment els canals d'E-mail i SMS. Aquesta classificació agruparà les possibles incidències en tres grups.

- **Incidències**: Prometheus detectarà la incidència que es mostrarà a Grafana però no enviarà cap missatge. La gestió d'aquestes incidències no és urgent i es portarà a terme en la següent revisió de les gràfiques de Grafana.
- **Notificacions**: Prometheus detectarà la incidència que es mostrarà a Grafana i s'enviarà una notificació al departament de sistemes i servei tècnic mitjançant correu electrònic. La gestió d'aquestes incidències serà urgent però no critica i es portarà a terme pel departament de sistemes com més aviat millor, dintre del seu horari de treball.
- **Alertes**: Prometheus detectarà la incidència que es mostrarà a Grafana, s'enviarà una notificació al departament de sistemes i servei tècnic mitjançant correu electrònic i un SMS al servei de guàrdies. La gestió d'aquestes incidències serà crítica i es portarà a terme pel personal de sistemes en l'horari d'oficina i fora d'aquest horari pel tècnic que estigui de guàrdia tan aviat com es pugui, dintre de les 4 següents hores a partir de l'aparició de la incidència. Aquest servei de guàrdies serà rotatiu i estarà operatiu 24x7 els 365 dies de l'any.

5.5 SAIs (*Sistemes d'Alimentació Ininterrompuda*)

Una altra capa dels sistemes de seguretat relacionada més amb el funcionament del hardware és la que s'encarrega d'assegurar el subministrament elèctric. De res serveix tenir els millors tallafocs, el millor programari de monitoratge, o redundància de hardware si no tenim subministrament elèctric. Per a aquesta tasca instal·larem un maquinari especialitzat per a protegir el maquinari de caigudes de tensió.

Aquest maquinari se'n diu SAI (*Sistema d'alimentació Ininterrompuda*) i s'encarregaran de protegir, com hem dit, de distintes anomalies en el corrent elèctric mitjançant bateries on s'acumularà energia elèctrica que alimentarà els dispositius connectats en cas de falta de corrent elèctric i evitar així la caiguda del sistema. A més, alguns d'aquests dispositius incorporen també un sistema anomenat AVR (*Automatic Voltage Regulator*), per a protegir el maquinari de variacions de la tensió que pogués fer malbé el maquinari connectat. Aquest sistema consisteix en un circuit electrònic que convertirà el flux elèctric variant en un flux continu i estable.

Dintre dels dispositius SAI tenim de tres classes diferents segons el nivell de protecció.

- **SAIs Off-Line**: Únicament emmagatzema l'energia elèctrica a les seves bateries mentre hi ha subministrament elèctric i alimenta els dispositius directament de la línia elèctrica. En cas de caiguda de la tensió elèctrica canvia el seu estat i comença a subministrar corrent elèctric des de les bateries fins que torni el subministrament de la línia o s'esgotin les bateries. Aquest sistema acostuma a necessitar entre 2 i 10 ms per a fer el canvi, de tal manera que dispositius sensibles a aquests temps de

commutació no estarien protegits per aquests dispositius. En el nostre cas, no podríem fer servir aquest tipus de SAIs.

- **SAIs In-Line**: Aquest sistema és similar al Off-Line però incorporarà un sistema de filtratge per a evitar les possibles fluctuacions de la tensió elèctrica. Igual que el sistema anterior, també trigarà entre 2 i 10ms en realitzar la commutació a bateries i per tant tampoc seran efectius pels nostres dispositius.
- **SAIs On-Line**: Finalment tenim els sistemes On-Line que faran servir una tecnologia diferent. Aquests SAIs subministraran corrent als dispositius directament de les bateries en tot moment, fet que evita el temps de commutació en cas de caiguda de la tensió elèctrica. Aquest sistema consisteix a convertir el corrent elèctric que rep de la línia elèctrica en corrent continu en passar per les bateries i després altra vegada a corrent altern per subministrar corrent als dispositius. D'aquesta manera també estabilitzem el corrent elèctric subministrant un flux estable. Aquest sistema sí que és optim per a dispositius sensibles als temps de commutació i per tant serà els dispositius que farem servir.

Una vegada triat la tecnologia de SAIs que necessitem, haurem de calcular la capacitat necessària per poder mantenir el nostre sistema el màxim de temps possible.

(*El càlcul de la capacitat necessària la podem trobar a l'annex IX*)

6. Implantació

Arribats a aquest punt, ja tenim les xarxes i la infraestructura dissenyades i arriba l'hora de la seva implantació. El fet que la implantació succeeixi d'una manera o una altra dependrà de les últimes decisions que falten per prendre.

La primera i més essencial és la de l'estructura a triar.

Inicialment es plantegen tres opcions:

- Dissenyar un cluster CEPH propi a partir de maquinari nou i la instal·lació de CEPH sobre SO CentOS 7.
- Contractar un sistema de claus en mans de RED HAT Ceph Storage amb una estructura similar a la dissenyada al projecte.
- Contractar recursos similars als dels punts anteriors a l'estructura d'Amazon AWS S3.

Després de parlar amb l'empresa, agafa pes l'opció de contractar un sistema Red Hat Ceph Storage de les característiques dissenyades al projecte. Es va demanar un pressupost per tal d'afegir-lo al projecte però ha estat impossible, ja que el temps de resposta per projectes d'aquest tipus és elevat. Al menys fins ara seguim esperant.

Pel que fa als serveis d'Amazon AWS S3 es contempla l'opció de contractar-ho a futur en comptes del clúster Ceph del CPD3. D'aquesta manera tindríem una solució al nuvol del qual no ens hauríem de preocupar pel maquinari. Es demana pressupost d'una estructura semblant per tenir informació de cara a futur.

Per últim, l'empresa planteja la separació del projecte en dues fases. Una primera que contemplarà tota l'estructura dissenyada pel CPD2 i que en cas de portar-se a terme es farà de forma immediata, i una segona fase que contemplarà tota l'estructura dissenyada pel CPD3 que quedarà ajornada, almenys fins a veure resultats de la primera fase. Es planteja la implantació de l'estructura al CPD3, una vegada arribats al 60% d'ús de la primera fase. Per tant la planificació contemplarà només aquesta primera fase.

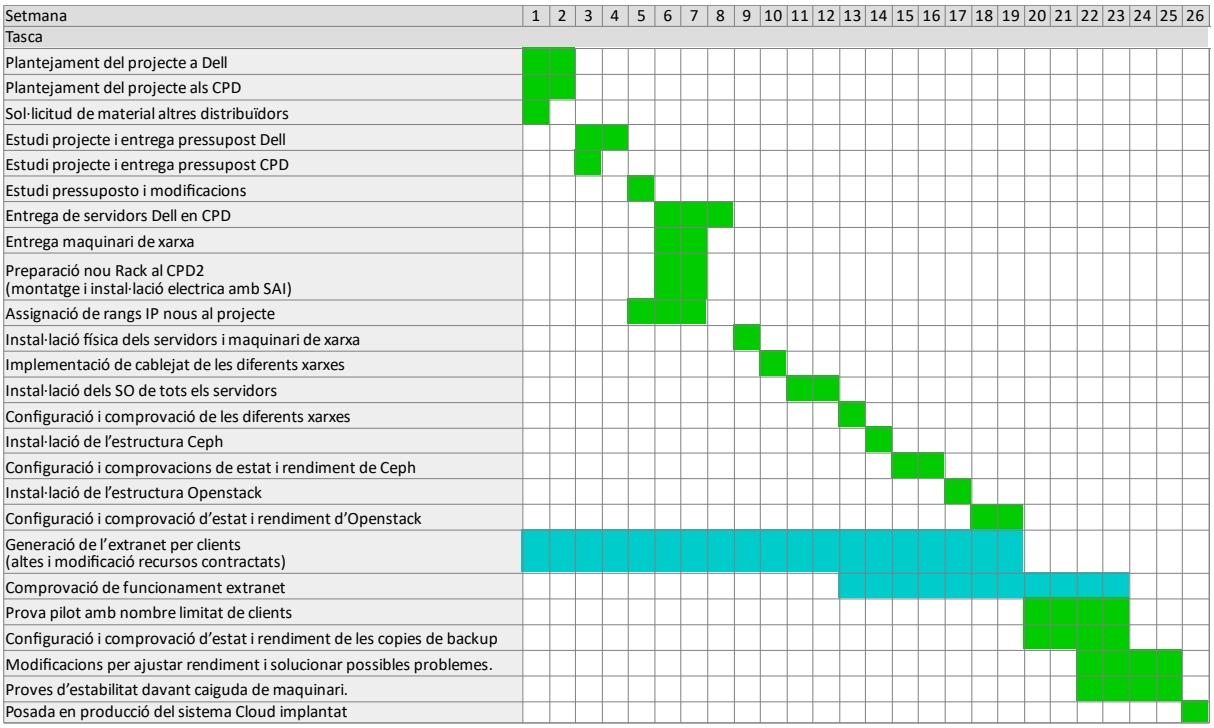
6.1 Planificació

La planificació de les modificacions fetes a la xarxa interna es portaran a terme per els tècnic del departament de sistemes amb l'ajudes de tècnic del departament de suport. Tindrà una duració aproximada de 7 setmanes de les que tres seran per a la compra del material i quatre per a la seva implantació.

Setmana	1	2	3	4	5	6	7
Tasca							
Sol·licitud de material xarxa distribuïdors							
Entrega maquinari de xarxa							
Configuració dels commutadors (VLANS)							
Implementació de cablejat de les diferents xarxes							
Instal·lació física del maquinari de xarxa LAN							
Instal·lació física del maquinari de xarxa WIFI							
Comprobacións de les diferents xarxes							

Pel que fa a la planificació de les tasques relacionades amb els CPDs, es fa en tenint en compte que a les dues el maquinari serà el triat al projecte i que el temps d'instal·lació pot estar més o menys equilibrat en els dos casos. Possiblement en el cas de la instal·lació per part nostra el temps seria més alt, però en el cas de Red Hat Ceph Storage augmentaria el tems d'entrega per ser un projecte relativament gran a causa del fet que haurà de passar per diferents departaments per acceptar-ho, dissenyar-ho implementar-ho, etc.

Es determina el temps en setmanes i no en dates concretes, ja que no es pot determinar encara en quin moment es començarà la implantació.



- Figura 32: Planning implantació.

6.2 Implantació

Pel que fa a la implantació, estarà dividida en quatre fases. La primera fase agrupa totes les tasques de comunicació amb els diferents distribuïdors des de l'inici del projecte fins al moment en què es fa l'entrega de tot el material necessari per a la implementació. Aquesta

fase la portarà a terme, principalment el responsable del projecte i/o del departament de sistemes.

La segona fase estarà formada per totes aquelles tasques de muntatge d'armaris, SAIs, servidors, maquinari de xarxa, cablejat, etc. Implementacions físiques de maquinari que es portaran a terme directament al CPD. Aquesta tasca la portaran a terme principalment, dos tècnics de sistemes amb ajudes puntuals de tècnics de suport. Aquesta fase també contemplarà una primera tanda de comprovacions orientades a l'estat del maquinari.

La tercera fase estarà formada per totes aquelles tasques d'instal·lació i configuració dels SOs i programari que s'ha dissenyat al projecte. Implementació de Ceph, Openstack, sistemes de monitoratge i alertes, sistemes de backup, etc. Aquesta fase la portaran a terme dos tècnics de sistemes amb ajudes puntuals de tècnics de suport. En aquesta Fase també es contemplarà una tanda de comprovacions d'Openstack i Ceph a escala de connectivitat, funcionalitat i rendiments.

La quarta i última fase estarà formada principalment per tasques de comprovació de la infraestructura. Comprovacions a escala de funcionament de l'API, l'extranet, alliberació de recursos, simulacres de recuperació de volums a partir de còpies de seguretat, caigudes i recuperacions de nodes, etc. Aquesta fase la portaran a terme dos tècnics del departament de sistemes junt amb ajudes puntuals dels tècnics de l'àrea de suport.

Una vegada completades aquestes quatre fases, el sistema estarà preparat per a funcionar en producció.

(*A l'annex VIII trobarem una taula amb el cost dels diferents maquinaris i altres aspectes que intervenen al projecte.)*

6.3 Migració i coexistència

Com ja s'ha comentat en els inicis del projecte, la infraestructura dissenyada s'ha d'implementar juntament amb una ja existent. Això provoca que hàgim d'anar amb cura per tal de no afectar el sistema ja existent a la vegada que hem d'anar passant part dels servidors allotjats a l'estructura antiga cap a la nova. Els sistemes actuals es basen principalment en VMware ESXi. La idea és la de migrar aquestes màquines virtuals cap a un node de còmput en el que farem servir VMware ESXi com a Hypervisor, d'aquesta manera anirem buidant els servidors antics de virtuals i retirant maquinari que no sigui útil. En alguns casos, aquest maquinari pot ser relativament nou i per tant aprofitable. Arribats a aquest cas es podria plantejar la d'incorporar aquest maquinari com a nodes OSD a ceph o nodes de computació a Openstack.

Per altra banda, existeixen estructures que no es podran migrar i hauran de coexistir amb el sistema cloud durant el temps que sigui necessari. Aquest no serà un problema atès que aquestes estructures estan en xarxes aïllades que no afectaran el sistema. Per aquesta raó s'han mantingut les xarxes /26 i /27 al CPD2.

7. Conclusions

Una vegada arribats a aquest punt, toca mirar enrere per veure amb perspectiva tot allò que s'ha fet durant els últims quatre mesos, per intentar portar a terme aquest projecte. Penso que ha estat molt emocionant, a la vegada que estressant el fet d'abordar un projecte d'aquestes dimensions, atès que en la fase inicial, portat per l'emoció i les ganes, vaig introduir moltes idees per a portar a terme. Idees que algunes ja em rondaven el cap en l'oficina i d'altres que vaig anant descobrint a partir de la documentació inicial. El fet és que una vegada portades a terme m'he adonat que van ser massa per a un projecte d'aquestes dimensions.

A mesura que he anat desenvolupant aquest projecte, m'he adonat que cada vegada que començava una secció, aquesta se'm feia interminable, dades i més dades que s'acumulaven i a vegades em feien perdre una mica el nord, forçant-me a parar, allunyar-me una mica i tornar a mirar-ho tot amb una mica de distància. Això m´ha permès aprendre a valorar la importància de l'organització, no només a l'hora de desenvolupar el projecte, sinó també en la part de recopilació d'informació. A més, penso que he de treballar més, i suposo que això o portarà l'experiència, en la planificació del temps. La limitació que suposa combinar un projecte com el treball de fi de grau, amb un treball en horari partit i una casa on hi ha dos nens, un d'ells des del març, m'ha portat al límit en alguns moments on, com he comentat abans, he hagut de parar, fer un cafè i reprendre-ho al cap d'una estona.

Pel que fa als objectius, provocat per aquest problema de planificació i limitació del temps, no he pogut assolir tots els objectius que m'havia marcat. Penso que inicialment no vaig tenir en compte les limitacions comentades a l'hora d'anar afegint objectius, i això ha provocat que alguns d'ells, com ara el disseny del departament de suport hagi estat descartat, o els pilots i el desenvolupament de la part d'OpenStack no hagin estat tan treballades com m'hauria agradat.

Al marge d'aquestes incidències, crec que he seguit la planificació bona part del projecte. Als dos primers lliuraments, el seguiment de la planificació va ser escrupolós. Però en començar la tercera entrega, el naixement del meu fill i per tant la baixa de paternitat em van impossibilitar l'accés als servidors que tenia preparats pels pilots. Per aquesta raó, vaig haver de canviar l'ordre en la planificació, dedicant-me primer a la part de sistemes de seguretat i després a la part d'emmagatzemament i virtualització. No va ser un canvi molt crític, donada la relació que hi ha entre els sistemes de seguretat i les xarxes i per tant va ser molt fluït.

Tanmateix, em queden pendents alguns temes que , tot i haver acabat el TFC, estic disposat a desenvolupar amb més profunditat. Primerament, referent als objectius del projecte, m'agradaria poder portar el pilot de *Ceph* i *OpenStack* fins al final. Poder fer més proves de rendiment,

desenvolupar un panell per a atacar l'api d'*OpenStack* i automatitzar algunes tasques com ara la creació d'instàncies, assignació de recursos a les instàncies, creació de *snapshots*, imatges, etc. A més, aprofitant el meu treball, fer una prova amb un parell de projectes per analitzar de forma més precisa el seu rendiment en escenaris reals. A part d'això, també em queda pendent una anàlisi més profunda de programaris com Grafana i/o Prometheus que, tot i que s'han instal·lat al pilot i s'han comprovat el seu funcionament m'han deixat amb la curiositat de veure fins on poden arribar.

Altrament, durant el desenvolupament del projecte m'hauria agradat poder investigar sobre un parell de programaris i que per motius de temps no ha estat possible. El primer és Proxmox, un sistema de virtualització de codi obert de l'estil d'OpenStack i que, després d'una primera ullada, em sembla interessant. A més, també ens dóna l'opció de treballar amb Ceph. El segón és el sistema d'emmagatzemament «Dell EMC VxRail». Una infraestructura VMware amb servidors DellEMC per oferir VSAN que m'agradaria provar, ja que ens ofereix una estructura finalitzada, testejada i estable que s'ha de tenir en compte com a alternativa a Ceph.

Finalment, tot i el patiment experimentat al llarg del projecte, em queda una bona sensació per haver pogut arribar fins aquí, tot i que en alguns moments no ho veia molt clar. Han estat moltes hores d'investigació, recopilació de dades i elaboració de textos, gràfics i taules, que al marge del projecte, s'han convertit en un aprenentatge important del qual poder treure'n profit en l'àmbit laboral.

8. Glossari

Definició dels termes i acrònims més rellevants utilitzats dins la Memòria.

1 – VPN: **V**irtual **P**rivate **N**etwortk – configuració de xarxes que permet a una xarxa externa treballar com si fos dintre de la xarxa a la que es connecta amb un tunel encriptat.

2 – Robert Metcalfe: Enginyer electric dels EE.UU. Va dissenyar Ethernet junt a David Boggs
· https://es.wikipedia.org/wiki/Robert_Metcalfe
· https://en.wikipedia.org/wiki/David_Boggs

3 – Xarxa ALOHA: Sistema per a la comunicació d'ordinadors pioner al 1970.
· https://es.wikipedia.org/wiki/ALOHAnet

4 – model OSI: Model d'interconnexió de sistemes oberts.
· http://www.seaccna.com/modelo-osi-guia-definitiva/
· https://es.wikipedia.org/wiki/Modelo_OSI

5 – IEEE 802: Estandards per a les xarxes d'area local (LAN)
· http://www.ieee802.org/
· http://www.ieee802.org/3/

6 – Redireccions NAT: Network Address Translation
· https://www.cisco.com/c/en/us/support/docs/ip/network-address-translation-nat/26704-nat-faq-00.html

7 – IPv6: Internet Protocol version 6
· https://www.ietf.org/rfc/rfc2460.txt
· https://www.internetsociety.org/resources/2018/state-of-ipv6-deployment-2018
· https://www.google.com/intl/en/ipv6/statistics.html#tab=per-country-ipv6-adoption

8 – DMZ: Desmilitarized zone (Zona desmilitaritzada)
· https://en.wikipedia.org/wiki/DMZ_(computing)

9 – WIFI: Wireless Fidelity (WLAN – Wireless LAN)
· http://www.ieee802.org/11/
· https://es.wikipedia.org/wiki/Wifi

10 – Cable UTP: Cable de par trenat no blindat
· https://es.wikipedia.org/wiki/Par_trenzado_no_blindado

11 – CPD: Centre de processament de dades
· https://es.wikipedia.org/wiki/Centro_de_procesamiento_de_datos

12 – iSCSI: Internet Small Computer System Iinterface
· https://es.wikipedia.org/wiki/ISCSI
· https://es.wikipedia.org/wiki/Small_Computer_System_Interface

13 – CEPH: CEPH File System – Sistema d'arxius distribuït basat en objectes
· https://ceph.com/
· https://es.wikipedia.org/wiki/Ceph_File_System

14 – VLAN: Virtual Local Area Network
· https://es.wikipedia.org/wiki/VLAN
· http://www.seaccna.com/vlan/

15 – PoE: Power Over Ethernet
· http://www.ieee802.org/3/af/ (PoE)
· http://www.ieee802.org/3/at/ (PoE+)
· https://es.wikipedia.org/wiki/Power_over_Ethernet

16 – Subnetting: Segmentació de xarxes IP físiques en d'altres lògiques mes petites a partir de la mascara de xarxa.
· https://www.mikroways.net/2009/06/08/guia-de-subnetting/

17 – MAC address: Adreça física de 48 bits que identifica unívocament a cada dispositiu de xarxa.
· https://es.wikipedia.org/wiki/Direcci%C3%B3n_MAC

18 – UTM: Unified Threat Management (Gestió Unificada d'amenaces)
· https://www.watchguard.com/es/wgrd-products/security-services
· https://searchsecurity.techtarget.com/definition/unified-threat-management-UTM
· https://en.wikipedia.org/wiki/Unified_threat_management

19 – SO: Sistema Operatiu: Programari principal per a la gestió del maquinari.
· https://es.wikipedia.org/wiki/Sistema_operativo
· https://tecnologia-informatica.com/el-sistema-operativo/

20 – Cluster: Conjunt de servidors que es comporten com una única unitat.
· https://es.wikipedia.org/wiki/Cl%C3%BAster_(inform%C3%A1tica)

21 – RBD: CEPH RADOS block Device. Sistema d'emmagatzemament RADOS distribuït basat en blocs de bytes.
· http://docs.ceph.com/docs/jewel/rbd/rbd/

22 – RADOS: Reliable Autonomic Distributed Object Store. Emmagatzemament d'objectes distribuïts autònoms i confiables.
· https://ceph.com/geen-categorie/the-rados-distributed-object-store/

23 – Swift: OpenStack Object Storage. Sistema d'emmagatzemament natiu d'OpenStack.
· https://wiki.openstack.org/wiki/Swift

24 – SaaS: Software as a Service.
· https://es.wikipedia.org/wiki/Software_como_servicio
· https://azure.microsoft.com/es-es/overview/what-is-saas/

25 – vSAN: Virtual Storage Area Network
·https://www.vmware.com/content/dam/digitalmarketing/vmware/es/pdf/products/vsan/vmware-vsan-datasheet.pdf

26 – SAN: Storage Area Network.
· https://www.vmware.com/topics/glossary/content/storage-area-network-san

27 – CRUSH: Controlled Replication Under Scalable Hashing
· http://docs.ceph.com/docs/jewel/rados/operations/crush-map/

28 – REST API: (Representational State Transfer) Transferencia d'Estat Representacional.
· https://es.wikipedia.org/wiki/Transferencia_de_Estado_Representacional

29 – IOPS: Inputs Outputs Per Second
· https://en.wikipedia.org/wiki/IOPS

30 – IaaS: Infrastructure As A Service
· https://en.wikipedia.org/wiki/Infrastructure_as_a_service

31 – AMQP: Advences Message Queuing Protocols
· https://www.amqp.org/

32 – VxLAN: Virtual Extensible Local Area Networking
· https://es.wikipedia.org/wiki/Virtual_Extensible_LAN

9. Bibliografia

- F. Kurose, J.; W. Ross, K.; (2013). Computer Networking - A top-down approach. Ed. Pearson
- Dr. Schulze, W. Director de Global Storage Consulting, Red Hat.(Feb 2016). Ceph CookBook. Quick answers to common problems. (Packt Publishing Ltd)
- Cicileo, G; Gagliano, R; O'Flaherty, C; Rocha, M; Olvera Morales, C; Palet Martinez, J; Vives Martínez, A. (Feb 2011). IPv6 per a tothom. Guia d'ús i aplicació per a diversos entorns. Edició: Fundació puntCAT.
- Navarro, M. (May 2017). Que es OpenStack y cuáles son sus retos. Revista ByteTI.
- D'Atri, A; Bhembre, V: Singh, K. (Oct 2017). Learning Ceph - Second Edition. Editorial: Packt Publishing.

La bibliografia consultada ha estat en la seva majoria documentació a internet, per aquesta raó afegeixo els enllaços consultats agrupats per temàtica a l'annex XIV.

10. Annexos

Annex I – Xarxes Ethernet

Les primeres xarxes Ethernet van ser dissenyades als anys 70 per *Robert Metcalfe*[2] junt amb David Boggs al centre Xerox PARC (Palo Alto Research Center) on va haver de dissenyar una xarxa per connectar diferents ordinadors i impressores a alta velocitat per a l'època. Aquest sistema estava basat en la xarxa ALOHA[3] desenvolupada a la universitat de Hawai el 1970. Inicialment aquesta xarxa desenvolupada per Metcalfe era anomenada *Alto Aloha Network*, però la por a que es pogués limitar el seu ús a computadores *Alto* va provocar el canvi de nom a *Ethernet,* basat en la teoria de la física que afirmava que les ones electromagnètiques es propagaven per un fluid anomenat *ether* i que Metcalfe comparava amb el par de coure de les xarxes que havia dissenyat.

Més tard, sobre el 1985 l' IEEE (Institut d'enginyers elèctrics i electrònics) junt amb el comitè d'estàndards per les xarxes metropolitanes van publicar els estàndards per a les xarxes d'àrea local (LAN). Aquests estàndards van ser els 802[5] i més concretament els 802.3 per a l'Ethernet.

Respecte al seu funcionament, Ethernet operarà a les capes d'enllaç de dades i física del model OSI[4].

Model de capes OSI

7	Capa d'aplicació
6	Capa de presentació
5	Capa de sessió
4	Capa de transport
3	Capa de xarxa
2	Capa d'enllaç de dades
1	Capa física

Aquest tipus de xarxes fan servir un mètode d'accés al medi (capa d'enllaç) amb detecció de portadora anomenat CSMA/CD (Accés Múltiple per detecció de portadora amb detecció de Col·lisions) que consisteix en què cada vegada que un dispositiu vol enviar una trama ha de comprovar que la línia és ocupada o no. Si no està ocupada, començarà a emetre, però si és

ocupada, descartarà l'enviament i esperarà una quantitat aleatòria de temps abans de tornar a intentar-ho. A més el dispositiu mantindrà les comprovacions periòdicament. En el cas que dos dispositius emetin una trama a la vegada es produiria una col·lisió el que provocaria que es descartessin els dos enviaments i els dos dispositius esperarien un període de temps aleatori per tornar a emetre la trama.

Estructura de la trama de 802.3 Ethernet

<table>
<tr><td>Preàmbul</td><td>Delimitador d'inici de trama</td><td>MAC de destí</td><td>MAC d'origen</td><td>802.1Q Etiqueta (opcional)</td><td>Ethertype (Ethernet II) o longitud (IEEE 802.3)</td><td>Payload</td><td>Seqüencia de comprovació (32-bit CRC)</td><td>Gap entre frames</td></tr>
<tr><td>7 Bytes</td><td>1 Byte</td><td>6 Bytes</td><td>6 Bytes</td><td>(4 Bytes)</td><td>2 Bytes</td><td>De 46 (o 42) fins 1500 Bytes</td><td>4 Bytes</td><td>12 Bytes</td></tr>
<tr><td colspan="2"></td><td colspan="6">64–1522 Bytes</td><td></td></tr>
<tr><td colspan="8">72–1530 Bytes</td><td></td></tr>
<tr><td colspan="9">84–1542 Bytes</td></tr>
</table>

Taula extreta de wikipedia. URL: https://es.wikipedia.org/wiki/Ethernet

A la capa física podem trobar diferents medis com poden ser: cable coaxial, cable de par trenat (utp) o fibra òptica. Depenent d'aquests medis i la tecnologia que es faci servir, podem obtenir majors velocitats de i distàncies de transmissió.

En el nostre cas farem servir dues tecnologies principalment, depenent del tipus de xarxes:

Tecnologia	Estàndard IEEE	Medi físic	Xarxa
1000BASE-T	802.3ab - 1999	cable UTP – cat6	Pública, Interna
10GBASE-T	802.3an - 2007	cable UTP – cat7	Emmagatzematge, Backup

Tot i això, les xarxes Ethernet actuals fan servir un model per l'intercanvi de dades anomenat TCP/IP que va ser creat als anys 70 pel departament de defensa dels Estats Units d'America U.S. per tractar de distribuir els seus centres de comandament de manera que un atac a un o varis d'ells no poguessin fer caure el centre de comandament sencer. Per fer això, van inventar aquest model que permetria connexions de computadors d'igual a igual de tal forma que tots els centres habilitats serien capaços de mantenir aquest comandament.

Aquest protocol de transport es va convertir en estàndard molt útil en el moment que es van començar a crear xarxes d'àrea extensa (WAN) en aquest moment els protocols que es feien servir no eren suficients i l'adreçament físic mitjançant les adreces de la capa d'enllaç (MAC) no permetien interconnectar de manera eficient aquestes xarxes, ja que aquestes

adreces anaven associades a un dispositiu concret més enllà de la seva ubicació a la xarxa el que dificultava la comunicació entre dispositius de xarxes diferents. Per aquest motiu es comença a fer servir el protocol TCP/IP el qual implementava adreçament lògic amb adreces de la capa de xarxa (capa 3). Aquest tipus d'adreçament permet aquest tipus de connexions, ja que a més que les diferents xarxes estan agrupades i existeix el concepte de rutes per arribar a qualsevol d'aquests grups, el paquet que s'envia des d'una xarxa mantindrà en tot el camí l'adreça de destí per tal d'anar travessant les diferents xarxes fins a arribar al seu destí.

Aquest protocol de transport és format per dos protocols, TCP i IP amb funcions diferents. El protocol TCP (protocol de control de transport) serà l'encarregat de controlar com les aplicacions envien les seves dades. Per fe això, TCP permetrà dividir aquesta informació en paquets més petits, enviar-los traves la xarxa i tornar-los a compactar en el destí. Mentre que el protocol IP serà l'encarregat de gestionar les adreces de cada paquet per tal que arribin al seu destí correctament.

Paral·lelament al model OSI, TCP/IP implementa el seu propi model en el qual es contemplen només quatre capes relacionades amb les set del model OSI.

Model de capes OSI		**Model TCP/IP**	
7	Capa d'aplicació	4	Capa d'aplicació
6	Capa de presentació		
5	Capa de sessió		
4	Capa de transport	3	Capa de transport
3	Capa de xarxa	2	Capa d'internet
2	Capa d'enllaç de dades	1	Capa física
1	Capa física		

Al model TCP/IP, el protocol TCP treballaria a la capa de transport mentre que el protocol IP treballaria a la capa de xarxa o internet.

Com hem comentat abans, el protocol IP utilitza uns tipus d'adreces lògiques que permeten enviar paquets des d'una xarxa a una altra de manera eficient. Inicialment aquestes IPs estaven definides amb l'estàndard IPv4 que consistia en adreces binàries de 32 bits amb la següent estructura:

```
00000000.00000000.00000000.00000000
```

Aquesta adreça, es fa servir a escala humana en sistema decimal amb el qual veuríem una adreça similar a aquesta:

```
0.0.0.0   o   127.0.0.1
```

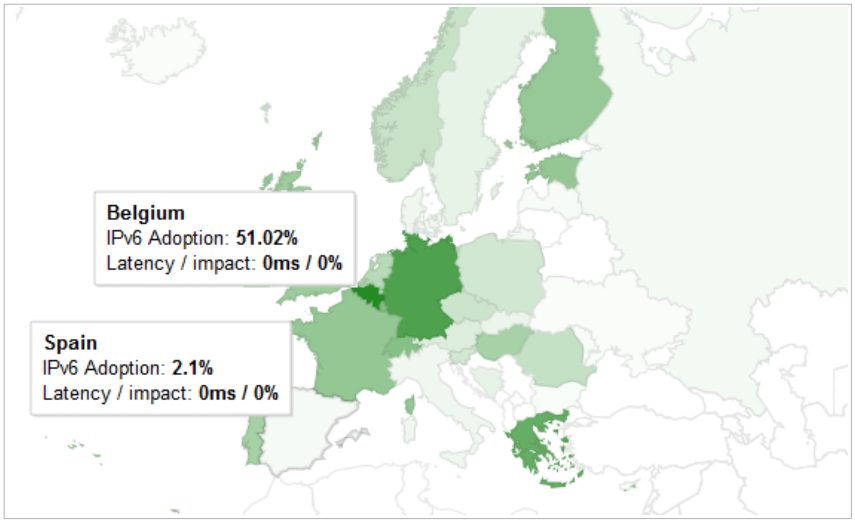
Quatre ternes de 8 xifres amb les quals es podien generar 2^{32} IPs diferents. Encara que la base d'aquest adreçament era que cada IP estaria associada a un dispositiu físic a internet, el creixement exponencial d'Internet i del nombre de xarxes LAN feia perillar la viabilitat d'aquest sistema per un possible esgotament del nombre d'IPs útils. Per tal d'estalviar IPs es van crear uns subgrups d'IPs com a IPs publiques i IPs privades. Aquestes últimes estarien relegades al seu ús local el que permetia que qualsevol pogués fer-les servir en les seves xarxes locals encara que estiguessin configurades en altres xarxes, ja que quedarien ocultes per la/s IPs públiques de la connexió a internet de la xarxa LAN.

Aquest sistema va ser la base de les comunicacions a internet fins a l'actualitat, però tot i la creació de les adreces privades, i d'altres sistemes com ara les redireccions NAT[6], o les adreces d'assignació dinàmica, el nombre d'IPs tard o d'hora acabaria esgotant-se. Actualment, IPv4 està sent substituït a llarg termini per una evolució de l'adreçament IP. El nou sistema anomenat IPv6[7] va ser desenvolupat el 1998 i pretén solucionar el problema que comportava l'aparició dels nous dispositius mòbils, els quals també requereixen una IP, i feia que les 2^{32} IPs d'IPv4 fossin insuficients. Aquest nou sistema es basa en adreces IPs de 128bits amb una representació gràfica en buit nombres de quatre xifres en hexadecimal.

```
2607:f8b0:4002:c09::93  (IPv6 de www.google.com)
```

I ens dóna capacitat per a uns 340 sextilions de possibles adreces.

Al marge de tot això, la implantació d'IPv6 està sent molt lenta, ja que els proveïdors d'accés a internet no estan apostant amb força per aquest sistema. Com a exemple tenim que a Europa, Espanya està al 2.1% d'implantació mentre que a països com Bèlgica estan ja al 51%.



Imatge extreta de l'URL: https://www.google.com/intl/en/ipv6/statistics.html#tab=per-country-ipv6-adoption

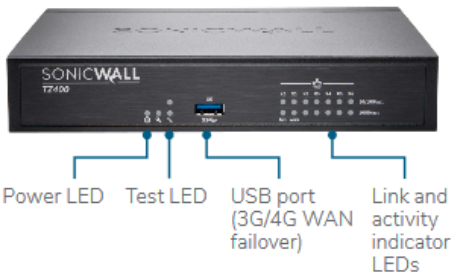
Tot i això, l'expansió de l'IPv6 ja s'està produint i assegura la continuïtat de les xarxes Ethernet per un bon període de temps pel que fa a capacitat d'IPs per habitant.

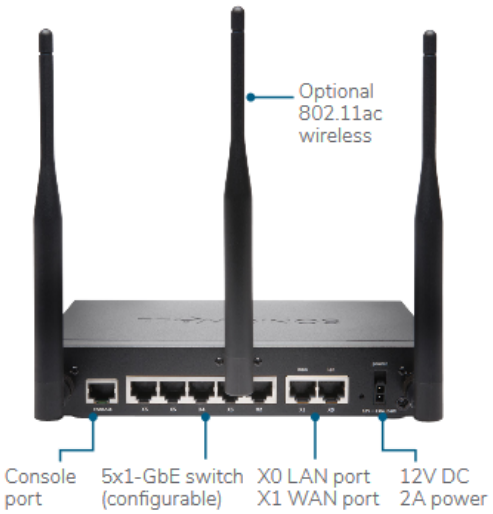
Annex II – Dispositius SonicWall TZ400 i TZ300

SonicWall TZ400 series

Aquesta gamma de tallafocs està dissenyada per a petites empreses que necessiten una protecció de qualitat.

Especificacions	Serie TZ400
Firewall throughput	1.3 Gbps
Threat Prevention throughput	600 Mbps
Anti-malware throughput	600 Mbps
IPS throughput	900 Mbps
Maximum connections	150,000
New connections/sec	6,000



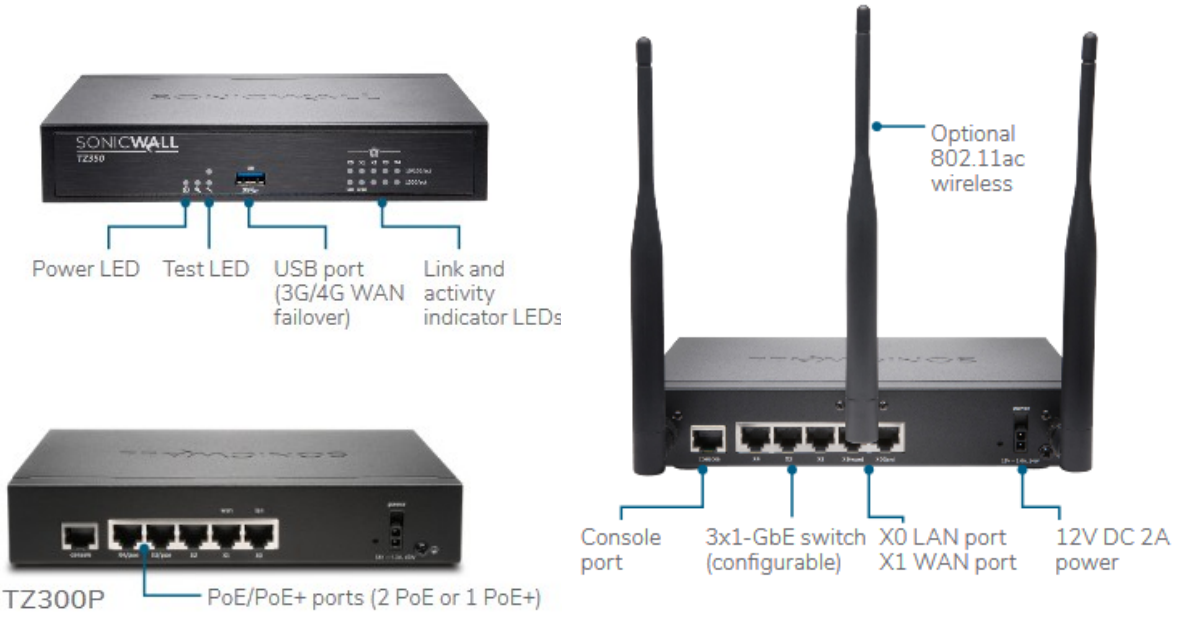


Informació i imatges extretes de l'URL:
https://d3ik27cqx8s5ub.cloudfront.net/media/uploads/2019/03/Datasheet-TZ-Series-US-VG-MKTG5541.pdf

SonicWall TZ300 series

Aquesta sèrie ofereix una solució completa per a xarxes petites per a la protecció mitjançant la prevenció d'intrusions, filtratge de contingut i polítiques de tallafocs. A més ofereix una versió amb wifi i una altra amb PoE.

Especificacions	Serie TZ300
Firewall throughput	750 Mbps
Threat Prevention throughput	235 Mbps
Anti-malware throughput	235 Mbps
IPS throughput	300 Mbps
Maximum connections	100,000
New connections/sec	5,000



Informació i imatges extretes de l'URL:
https://d3ik27cqx8s5ub.cloudfront.net/media/uploads/2019/03/Datasheet-TZ-Series-US-VG-MKTG5541.pdf

Annex III – Dispositius WatchGuard M570 i M370

WatchGuard® Model	WatchGuard Firebox® M370	WatchGuard Firebox® M570
Ideal For	Small to midsize businesses	Midsize business and distributed enterprise
Hardware		
Interfaces	8 Gb	8 Gb (included); Optional 8 Gb, 8 Gb fiber, 4 x 10 Gb fiber
Security		
Application Proxies	HTTP, HTTPS, SMTP, FTP, DNS, TCP-UDP, POP3, POP3S, SIP, H.323	
Intrusion Prevention	Si	Si
User Authentication with transparent Windows authentication	Si	Si
Performance		
Firewall Throughput	8 Gbps	26.6 Gbps
VPN Throughput	4.6 Gbps	5.8 Gbps
AV Throughput	3 Gbps	5.4 Gbps
IPS Throughput	4.8 Gbps	8 Gbps
UTM Throughput	2.6 Gbps	4.4 Gbps
Concurrent Sessions (bi-directional connections)	3,300,000	8,300,000
VPN Tunnels		
Branch Office VPN Tunnels	100	500
Mobile VPN with SSL/L2TP	100	500
Mobile VPN with IPSec	100	500

Informació i imatges extretes de l'URL: https://www.watchguard.com/wgrd-products/appliances-compare/15016/15026

Annex IV – Dispositius SAI Lapara 10000VA

Fitxa Tècnica, Sai rackLapara 10000VA Online LCD

Especificacions	Valor
Capacitat	10000VA / 10000W
Entrada	
Transferència baix voltatge	176 VAC +/- 3 a 100% de carrega o 110 VAC +/- 3 a 50% de carrega
Regrés baix voltatge	186 VAC a 100% de carrega o 120 VAC a 50% de carrega
Transferència alt voltatge	300 VAC +/- 3%
Regrés alt voltatge	290 VAC
Rang de freqüència	46-54Hz o 56-64Hz
Fase	Monofàsic amb toma a terra
Factor de potencia	>= 0.99 a 100% de la carrega
Sortida	
Voltatge de sortida	208/220/230/240 VAC
Regulació voltatge AC (Mode Bateria)	+/- 1%
Rang de freqüència (sincronitzada)	46~54Hz o 56.0~64.0Hz
Rang de freqüència (Mode bateria)	50Hz +/-0.1Hz o 60Hz +/-0.1Hz
Ratí actual de cresta	3:1
Distorsió harmònica	<= 3% THD (carrega lineal) / <= 6% THD (carrega no lineal)
Temps de transferència	Cero / Inverter a Bypass : Cero
Sortida d'ona	Ona sinusoïdal pura
Connexions	
Toma externa	Terminal Block
Ports USB	1

Ports RS232	1
Slot per SNMP	si
Bateria	
Tipus de bateria	12V/9Ah
Nombre de bateries	20
Temps de càrrega	3 hores al 90%
Càrrega (max.)	2.0 A
Voltatge de càrrega	273.0 VDC +/- 1%
Indicadors	
Estat del SAI, nivell de càrrega, nivell de bateria, Voltatge d'entrada/sortida, cronometre de descàrrega i condicions de fallada.	
Alarmes audibles	
Mode bateria	Cada 4 segons
Bateria baixa	Cada segon
Sobrecarrega	2 vegades per segon
Fallada	Contínuament
Dades físiques	
Pes net	SAI : 17 Kg
	Bateries : 63 Kg
Dimensions (Fons x Ample x Alt) mm	SAI : 580 x 438 x 88 mm
	Bateries : 580 x 438 x 133 mm
Altura	SAI : 2U
	Bateries : 3U
Ambient	
Condicions ambientals	20-90% RH @ 0-40°C (No-condensada)
Nivell de soroll	menys de 50 dB a 1 Metre
Gestió Software	
Via USB/RS-232	Suporta Windows 98 SE/ME/NT 4.x/2000/2003/XP/Vista/2008/7, Linux, Unix i Mac
Via SNMP opcional	Gestió de corrent a través del gestor SNMP i navegador web

Informació i imatges extretes de l'URL: https://www.sai-online.es/sai-online/sai-rack-10000-va-10kva-online-lcd-lapara

Annex V – Encaminador Cisco ASR 1001-X

Xassís		
Especificacions físiques	Altura:	43.43 mm
	Amplada:	439.42 mm
	Profunditat:	461.5 mm
	Pes:	11.35 kg
Muntatge en armari	1 U en armari de 19 polsades	
Especificacions		
Memòria	DRAM de 8 GB compartida a través del processador de rutes, ESP i SIP	
Ports Gigabit Ethernet	6 Gigabit Ethernet mitjançant ports SFP	
Ports 10 Gigabit Ethernet	2 10Gigabit Ethernet mitjançant ports SFP	
Ample de banda ESP	2.5 – 20 Gbps	
Encriptació per Maquinari	>8 Gbps	
Font d'alimentació	Doble font d'alimentació redundant	
Més informació		
https://www.cisco.com/c/en/us/support/routers/asr-1001-x-router/model.html#DataSheets		

Annex VI – Instal·lació Cluster Ceph

Com s'ha comentat per a la instal·lació del cluster Ceph haurem d'instal·lar els següents dimonis.

- **Ceph OSDs (ceph-osd)** - Gestiona el magatzem de dades, la replicació i la recuperació de dades. Un clúster Ceph necessita almenys dos servidors Ceph OSD. Farem servir tres servidors OSD sobre CentOS 7 amb un OSD per servidor. A posteriori afegirem un quart OSD en un dels servidors.

- **Ceph Monitor (ceph-mon)** - Supervisa l'estat del clúster, el mapa OSD, el mapa PG i el mapa CRUSH. En aquest pilot farem servir un únic monitor.

- **Ceph Meta Data Server (ceph-mds)** – Aquest node només es necessari per fer servir Ceph com un sistema de fitxers.
- **Ceph Manager (ceph-mgr)** – Aquest node és necessari per a obtenir mètodes de gestió dels nodes monitors.

Prerequisits:

- 2 servidors físics (Dell PowerEdge R330) amb VMware ESXi 6.5 per a la virtualització.
- 6 nodes virtuals amb CentOS 7 minimal instal·lat.
- Accés root en tots els nodes.

Llistat dels nodes virtuals que es muntaran i les IPs assignades.

Hostname	dirección IP
ceph-admin	192.168.130.120
mon1	192.168.130.121
osd1	192.168.130.131
osd2	192.168.130.132
osd3	192.168.130.133
client	192.168.130.117

Tots els nodes OSD necessiten dues particions (en el cas real la partició root anirà allotjada en discos m.2):

- Partició root (/)
- Partició buida que es farà servir com magatzem de dades Ceph.

Pas 1 – Configuració dels Nodes

En aquest pas, configurarem els 6 nodes per preparar-los per a la instal·lació del Cluster Ceph. Hem de seguir i executar totes les comandes indicades a continuació, en tots els nodes.

Creem l'usuari de Ceph

Creem l'usuari '**cephuser**' en tots els nodes.

```
useradd -d /home/cephuser -m cephuser
passwd cephuser
```

Després de crear el nou usuari, necessitem configurar l'usuari '**cephuser**' amb permisos de root. Ha de poder executar comandes com *root* i obtenir privilegis de *root* sense una contrasenya.

Executem la comanda següent per crear un arxiu *sudoers* per l'usuari, i editem l'arxiu */etc/sudoers* amb la utilitat *sed*.

```
echo "cephuser ALL = (root) NOPASSWD:ALL" | sudo tee
/etc/sudoers.d/cephuser
chmod 0440 /etc/sudoers.d/cephuser
sed -i s'/Defaults requiretty/#Defaults requiretty'/g /etc/sudoers
```

Instal·lar i Configurar NTP

Instalem NTP per sincronitzar la data i l'hora en tots els nodes. Executem la comanda *ntpdate* per establir una data i hora a través del protocol NTP, farem servir el servidor NTP de l'organització *ntp.org*. Després iniciarem i habilitarem el servidor NTP perquè s'executi en el moment de l'inici.

```
yum install -y ntp ntpdate ntp-doc
ntpdate 0.es.pool.ntp.org
hwclock --systohc
systemctl enable ntpd.service
systemctl start ntpd.service
```

Instal·lar Open-vm-tools

Aquest no seria un pas necessari en la instal·lació real, ja que serien servidors físics, però en el cas del pilot, en ser nodes que aniran sota ESXi és recomanable la seva instal·lació.

```
yum install -y open-vm-tools
```

Deshabilitar SELinux

Deshabilitem SELinux en tots els nodes editant l'arxiu de configuració de SELinux amb l'editor de seqüències sed.

```
sed -i 's/SELINUX=enforcing/SELINUX=disabled/g' /etc/selinux/config
```

I forcem el valor en la sessió actual sense haver de reiniciar el servidor.

```
setenforce 0
```

Configurar el fitxer de Hosts

Editarem l'arxiu /etc/hosts en tots els nodes amb l'editor *vi* i agreguem línies amb l'adreça IP i els noms de host de tots els nodes del clúster.

```
vi /etc/hosts
```

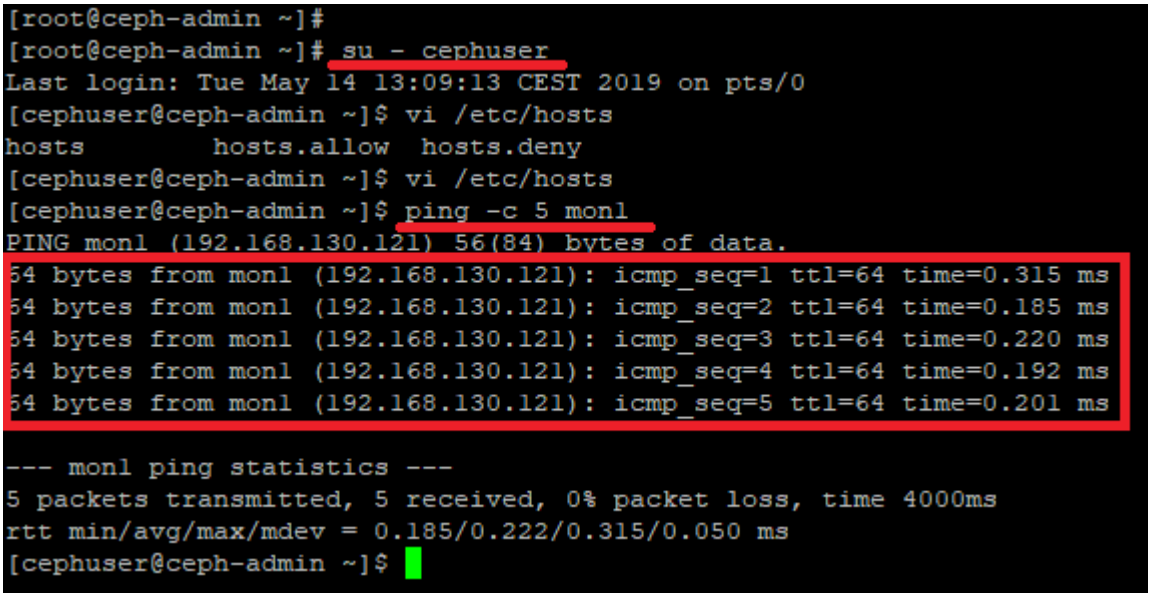
Afegim la configuració al final

```
192.168.130.120   ceph-admin
192.168.130.121   mon1
192.168.130.131   osd1
192.168.130.132   osd2
192.168.130.133   osd3
192.168.130.117   client
```

Guardem el fitxer i sortim de vi.

Ara ja podem provar de fer *ping* entre els servidors amb els seus noms per provar la connectivitat entre ells.

```
ping -c 5 mon1
```



Pas 2 - Configurar el servidor SSH

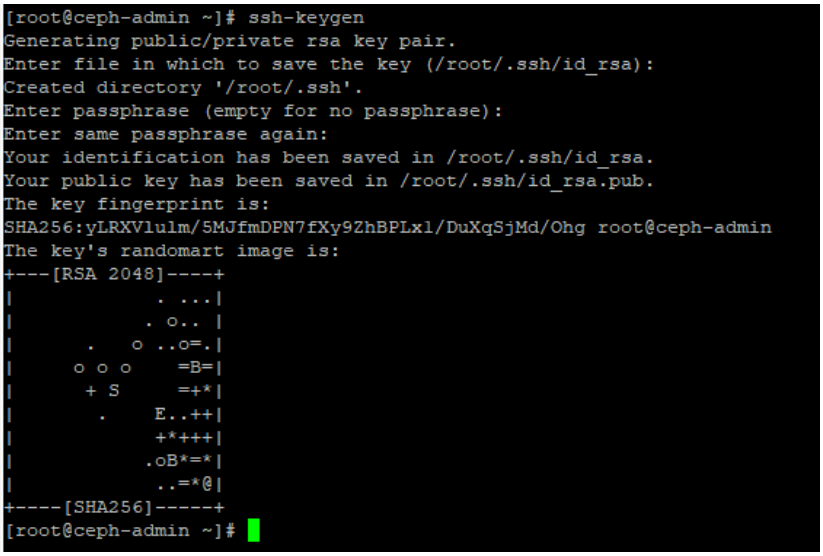
En aquest pas, configurarem el node ***ceph-admin***. El node d'administració es fa servir per configurar el node *monitor* i els nods *osd*. Iniciem sessió al node **ceph-admin** i canviem a l'usuari '**cephuser**'.

```
ssh root@ceph-admin
su - cephuser
```

Com hem dit abans, amb el node d'administració instal·larem i configurarem tots els nodes del clúster, i per tant l'usuari '**cephuser**' del node *ceph-admin* deu tenir privilegis per connectar-se a tots els nodes pe SSH sense contrasenya.

Primer generem les claus ssh per '**cephuser**' i deixarem la contrasenya en blanc per no haver de posar-la cada vegada que connectem.

```
ssh-keygen
```



A continuació, creem l'arxiu de configuració per la configuració ssh.

```
vi ~/.ssh/config
```

I afegim la configuració al final.

```
Host ceph-admin
    Hostname ceph-admin
    User cephuser
Host mon1
    Hostname mon1
    User cephuser
Host osd1
    Hostname osd1
    User cephuser
Host osd2
    Hostname osd2
    User cephuser
Host osd3
    Hostname osd3
    User cephuser
Host client
    Hostname client
    User cephuser
```

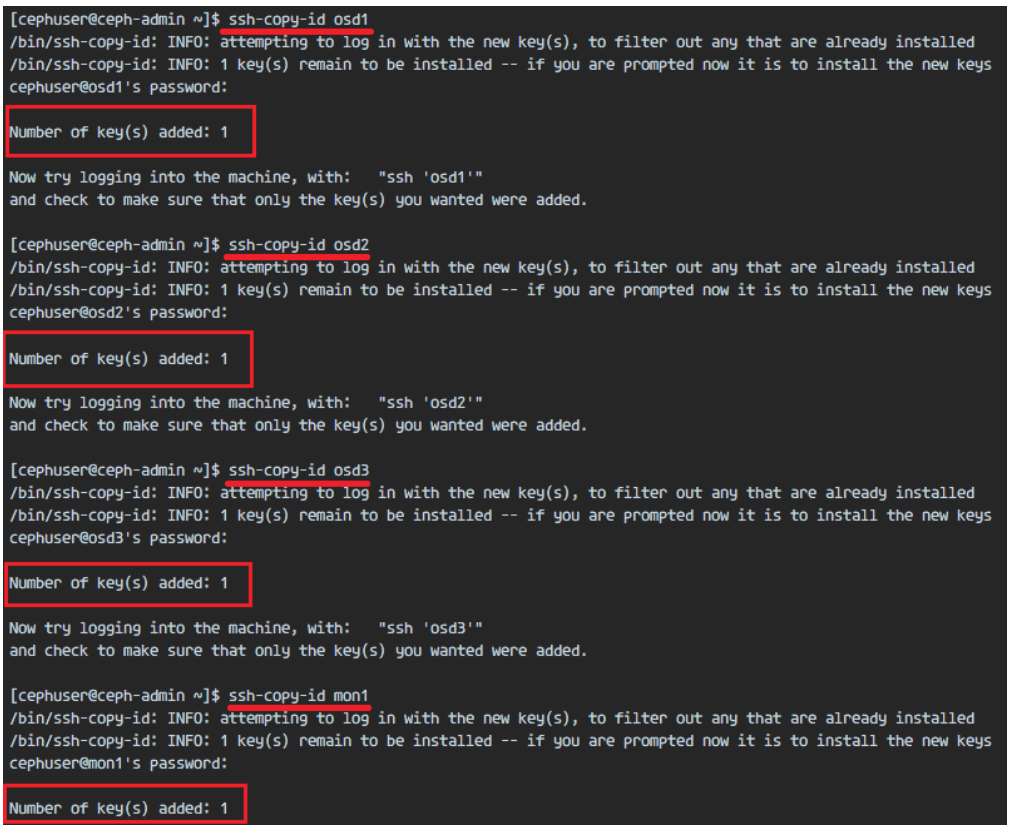
Canviem els permisos del fitxer de configuració.

```
chmod 644 ~/.ssh/config
```

Ara afegim la clau SSH a tots els nodes amb la comanda *ssh-copy-id*.

```
ssh-keyscan osd1 osd2 osd3 mon1 client >> ~/.ssh/known_hosts
ssh-copy-id osd1
ssh-copy-id osd2
ssh-copy-id osd3
ssh-copy-id mon1
ssh-copy-id client
```

Introduïm la contrasenya de l'usuari '*cephuser*' quan ens la demani.



Quan hagi finalitzat, intentem accedir al node *osd1* des de el node *ceph-admin*.

```
ssh osd1
```

Pas 3 – Configurar el Firewalld

Podríem configurar *Firewalld* per protegir el sistema però en el cas que ens ocupa el deshabilitarem en tots els nodes per evitar problemes innecessaris per a una prova pilot.

```
systemctl disable firewalld
systemctl stop firewalld
```

Paso 4 - Configurar els nodes OSD Ceph

En el nostre cas, farem servir 3 nodes OSD i cada node tindrà dues particions(dos discos virtuals).

- **/dev/sda** per a la partició root.
- **/dev/sdb** que serà una partició buida per Ceph - 50GB en el nostre cas.

Des del node *ceph-admin*, iniciem sessió en tots els nodes OSD i formatem la partició /dev/sdb com a XFS.

```
ssh osd1
ssh osd2
ssh osd3
```

Comprovem la partició amb *fdisk*.

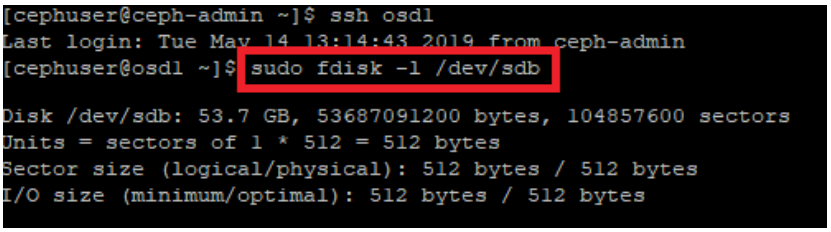
```
sudo fdisk -l /dev/sdb
```

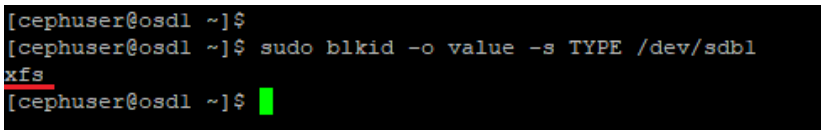
Ara donarem format a la partició */dev/sdb* amb sistema de fitxers XFS i amb una taula de particions GPT fent servir *parted*.

```
sudo parted -s /dev/sdb mklabel gpt mkpart primary xfs 0% 100%
sudo mkfs.xfs /dev/sdb -f
```

Ara comprovem el disc, i obtindrem la informació de la partició xfs */dev/sdb*.

```
sudo blkid -o value -s TYPE /dev/sdb
```





Pas 5 - Muntar el Clúster Ceph

En aquest pas, instal·larem Ceph en tots els nodes des del node *ceph-admin*.

Primer ens connectarem al node ceph-admin i canviarem d'usuari a *cephuser*.

```
ssh root@ceph-admin
su - cephuser
```

Instal·lar ceph-deploy al node ceph-admin

Agreguem el repositori de Ceph i instal·lem l'eina d'implementació de Ceph '*ceph-deploy*' amb la comanda *yum*.

```
sudo rpm -Uhv http://download.ceph.com/rpm-luminous/el7/noarch/ceph-release-1-1.el7.noarch.rpm
sudo yum update -y && sudo yum install ceph-deploy -y
```

Una vegada que s'hagi instal·lat l'eina *ceph-deploy*, crearem un nou directori per la configuració del clúster ceph.

Crear nova configuració de clúster

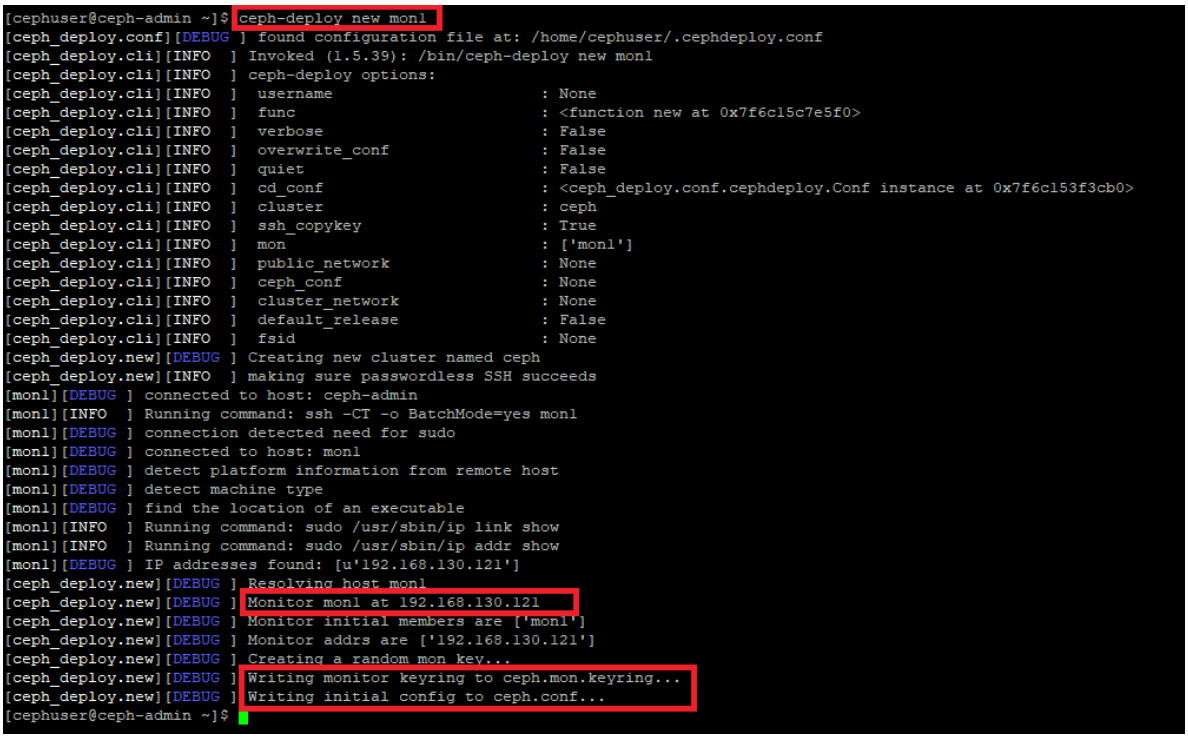
Crearem el nou directori de clúster.

```
mkdir cluster
cd cluster/
```

A continuació, crearem una nova configuració de clúster amb la comanda '*ceph-deploy*', pel node monitor '*mon1*'.

```
ceph-deploy new mon1
```

La comanda generarà l'arxiu de configuració del clúster Ceph '*ceph.conf*' al directori del clúster.



Editem el fitxer de configuració *ceph.conf* amb *vi.*

```
vi ceph.conf
```

Sota el bloc *[global]*, afegirem la següent configuració.

```
# Your network address
public network = 192.168.130.0/24
osd pool default size = 2
```

Guardem el fitxer i sortim de vi.

Instal·lar Ceph en tots els nodes

Ara instal·larem Ceph en tots els altres nodes des del node *ceph-admin*. Això es pot fer amb una sola comanda.

```
ceph-deploy install ceph-admin mon1 osd1 osd2 osd3
```

La comanda instal·larà automàticament Ceph en tots els nodes: *mon1*, *osd1-3* i *ceph-admin*

Una vegada acabi la instal·lació podem desplegar el *ceph-mon* al node *mon1*.

```
ceph-deploy mon create-initial
```

La comanda crearà la clau del monitor que es podrà verificar i obtenir amb la comanda 'ceph'.

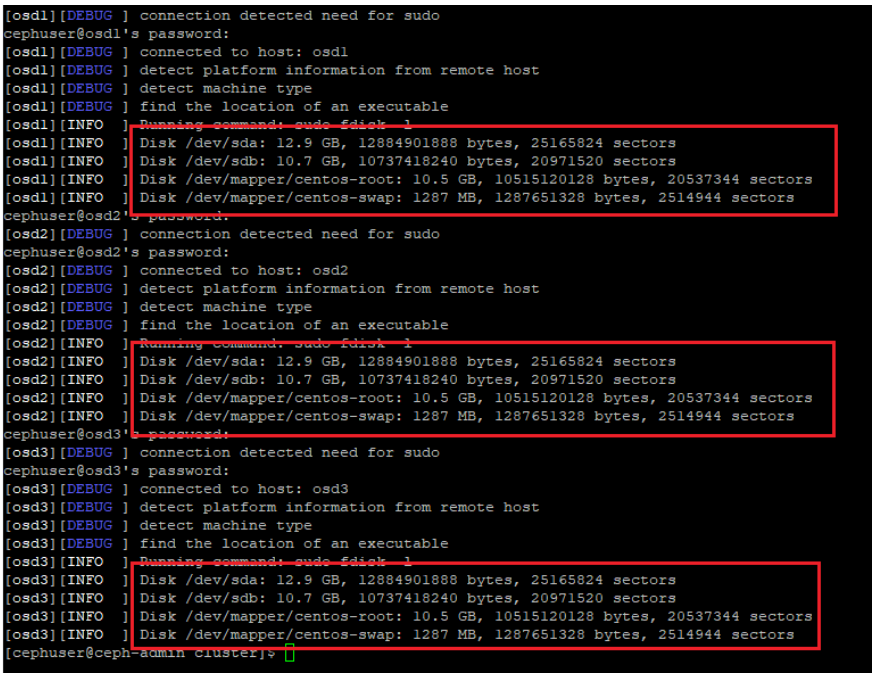
```
ceph-deploy gatherkeys mon1
```

Agregando OSDS al Clúster

Quan Ceph s'hagi instal·lat en tots els nodes, podrem afegir els dimonis OSD al clúster. Els dimonis OSD crearan les seves dades i particions de diari al disc /dev/sdb.

Comprovarem que la partició /dev/sdb estigui disponible en tots els nodes OSD.

```
ceph-deploy disk list osd1 osd2 osd3
```



A continuació, eliminem les taules de partició */dev/sdb* en tots els nodes amb l'opció *zap*.

Para ceph-deploy 1.5.39

```
ceph-deploy disk zap osd1:/dev/sdb osd2:/dev/sdb osd3:/dev/sdb
```

Para ceph-deploy 2.0.0

```
ceph-deploy disk zap osd1 /dev/sdb
ceph-deploy disk zap osd2 /dev/sdb
ceph-deploy disk zap osd3 /dev/sdb
```

La comanda eliminarà totes les dades en */dev/sdb* als nodes OSD de Ceph.

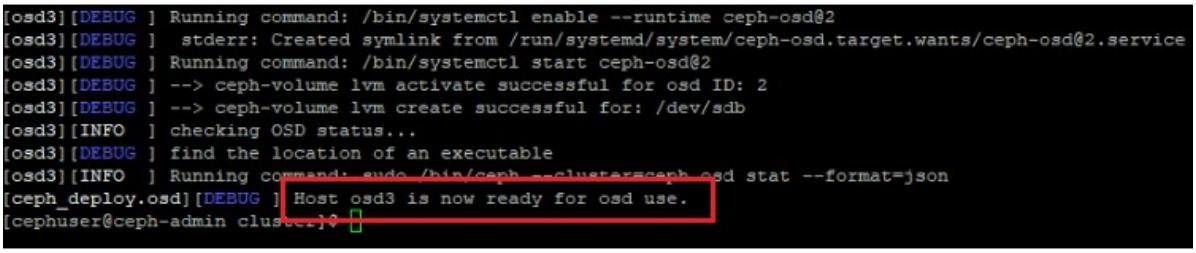
Ara prepararem tots els nodes OSDs. Ens hem d'assegurar que no hagi errors als resultats.

Para ceph-deploy 1.5.39

```
ceph-deploy osd prepare osd1:/dev/sdb osd2:/dev/sdb osd3:/dev/sdb
```

Para ceph-deploy 2.0.0

```
ceph-deploy osd create --data /dev/sdb osd1
ceph-deploy osd create --data /dev/sdb osd2
ceph-deploy osd create --data /dev/sdb osd3
```



Si veiem que *osd1-3* està llest per l'us de l'OSD, llavors la implementació haurà estat correcta.

Ara activarem els OSDs amb la següent comanda:

Solo para Ceph < 2.0

```
ceph-deploy osd activate osd1:/dev/sdb1 osd2:/dev/sdb1 osd3:/dev/sdb1
```

Haurem de comprovar la sortida d'errors abans de continuar. Ara podrem verificar el disc SDB als nodes OSD amb la comanda *list*.

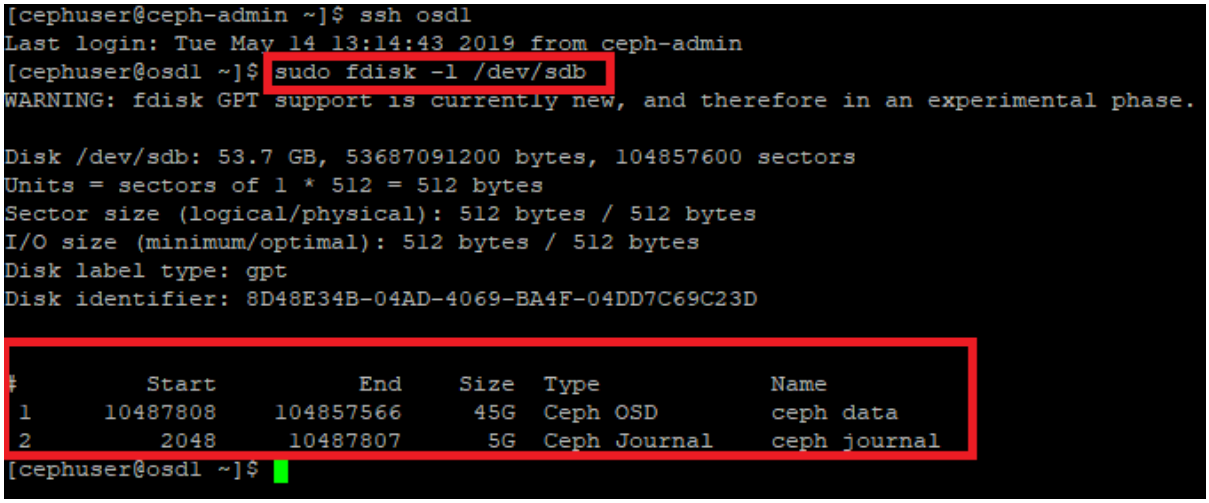
```
ceph-deploy disk list osd1 osd2 osd3
```

El resultat és que */dev/sdb* te ara dues particions:

- **/dev/sdb1** - Ceph Data
- **/dev/sdb2** - Ceph Journal

O es pot verificar directament al node OSD amb *fdisk*.

```
ssh osd1
sudo fdisk -l /dev/sdb
```



A continuació, implementem la clau d'administració en tots els nodes associats

```
ceph-deploy admin ceph-admin mon1 osd1 osd2 osd3
```

Canviem el permís de l'arxiu de clau executant la següent comanda en tots els nodes.

```
sudo chmod 644 /etc/ceph/ceph.client.admin.keyring
```

Arribats a aquest moment, ja tindríem en funcionament el cluster de ceph

Una vegada instal·lat el clúster Ceph, podrem provar-ho i assegurar-nos que no hi hagi errors en la configuració del clúster.

Des del node *ceph-admin*, iniciem sessió en el servidor '*mon1*'.

```
ssh mon1
```

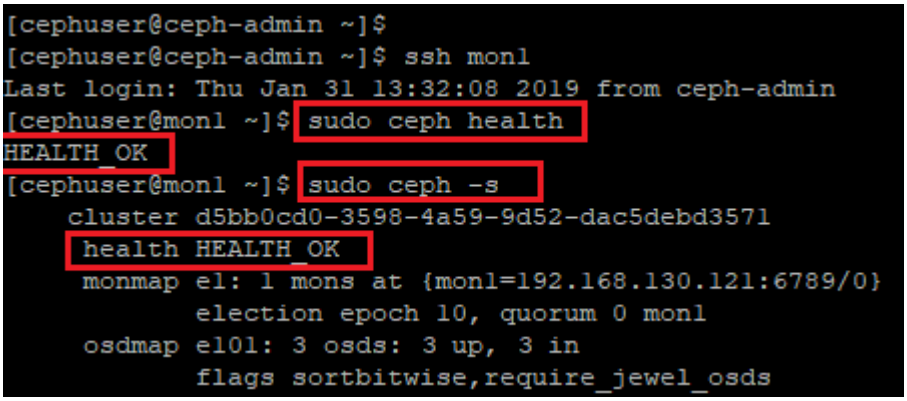
Executem la següent comanda per verificar l'estat del clúster.

```
sudo ceph health
```

Ara comprovem l'estat del clúster.

```
sudo ceph -s
```

I hauríem de veure els resultats com els següents:

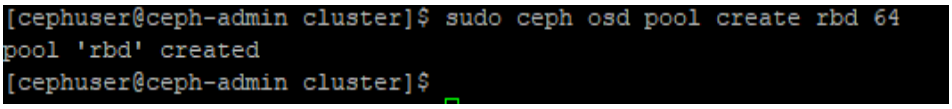


Ens hem d'assegurar que la salut de Ceph estigui bé i que hi hagi un node de monitor 'mon1' amb l'adreça IP '192.168.130.121'. Hauria d'haver-hi també 3 servidors OSD i tots haurien d'estar en funcionament. A més, hauria d'haver-hi un disc disponible d'aproximadament 150 GB - 3x50 GB de partició de dades Ceph.

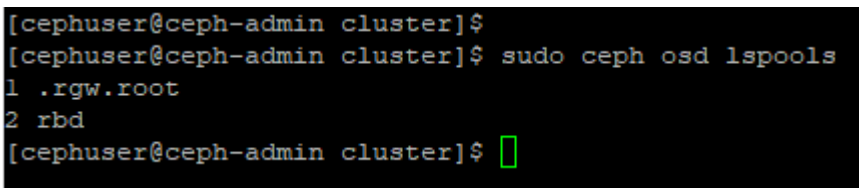
Proves a realitzar:

- **Crear un nou pool.**

  Per crear un nou pool la comanda és molt simple. Només haurem d'indicar el nom del nou pool i el nombre de pg que l'assignarem.



  I llistem els pools per confirmar que s'ha creat correctament.



- **Afegir un nou OSD**

  Una altra comanda important és la d'afegir un nou OSD al cluster. Per fer això haurem d'afegir un nou disc al host OSD en el que volem afegir el nou node OSD.

  Afegirem un nou disc al OSD2, ja que al OSD1 ja tenim dos OSDs.

  Aquesta acció és purament física i en el nostre cas consisteix a afegir un disc virtual amb el VMware web client.

  Una vegada tenim el disk comprovarem que el SO del host OSD on el volem activar ens reconeix el disc amb la comanda fdisc -l

```
[root@osd2 ~]# fdisk -l /dev/sdc

Disk /dev/sdc: 53.7 GB, 53687091200 bytes, 104857600 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
```

  Ara ja podem crear l'osd amb la següent comanda

```
ceph-deploy osd create --data /dev/sdc osd2
```

```
[root@ceph-admin ceph]# ceph-deploy osd create --data /dev/sdc osd2
[ceph_deploy.conf][DEBUG ] found configuration file at: /root/.cephdeploy.conf
[ceph_deploy.cli][INFO  ] Invoked (2.0.1): /usr/bin/ceph-deploy osd create --data /dev/sdc osd2
[ceph_deploy.cli][INFO  ] ceph-deploy options:
[ceph_deploy.cli][INFO  ]  verbose                       : False
[ceph_deploy.cli][INFO  ]  bluestore                     : None
[ceph_deploy.cli][INFO  ]  cd_conf                       : <ceph_deploy.conf.cephdeploy.Conf instance at 0x7f0d9aa113b0>
[ceph_deploy.cli][INFO  ]  cluster                       : ceph
[ceph_deploy.cli][INFO  ]  fs_type                       : xfs
[ceph_deploy.cli][INFO  ]  block_wal                     : None
[ceph_deploy.cli][INFO  ]  default_release               : False
[ceph_deploy.cli][INFO  ]  username                      : None
[ceph_deploy.cli][INFO  ]  journal                       : None
[ceph_deploy.cli][INFO  ]  subcommand                    : create
[ceph_deploy.cli][INFO  ]  host                          : osd2
[ceph_deploy.cli][INFO  ]  filestore                     : None
[ceph_deploy.cli][INFO  ]  func                          : <function osd at 0x7f0d9ac54938>
[ceph_deploy.cli][INFO  ]  ceph_conf                     : None
[ceph_deploy.cli][INFO  ]  zap_disk                      : False
[ceph_deploy.cli][INFO  ]  data                          : /dev/sdc
[ceph_deploy.cli][INFO  ]  block_db                      : None
[ceph_deploy.cli][INFO  ]  dmcrypt                       : False
[ceph_deploy.cli][INFO  ]  overwrite_conf                : False
[ceph_deploy.cli][INFO  ]  dmcrypt_key_dir               : /etc/ceph/dmcrypt-keys
[ceph_deploy.cli][INFO  ]  quiet                         : False
[ceph_deploy.cli][INFO  ]  debug                         : False
[ceph_deploy.osd][DEBUG ] Creating OSD on cluster ceph with data device /dev/sdc
```

```
[osd2][DEBUG ] Running command: systemctl start ceph-osd@4
[osd2][DEBUG ] --> ceph-volume lvm activate successful for osd ID: 4
[osd2][DEBUG ] --> ceph-volume lvm create successful for: /dev/sdc
[osd2][INFO  ] checking OSD status...
[osd2][DEBUG ] find the location of an executable
[osd2][INFO  ] Running command: /bin/ceph --cluster=ceph osd stat --format=json
[ceph_deploy.osd][DEBUG ] Host osd2 is now ready for osd use.
```

Una vegada creat l'osd, si mirem l'estat del clúster, veurem que surt un avís de degradació.

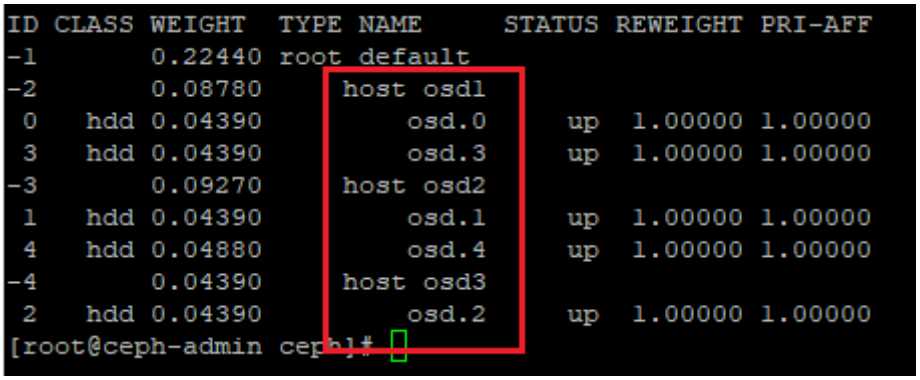
```
[root@ceph-admin ceph]# ceph -s
2019-05-28 12:57:47.095657 7f71f9519700 -1 asok(0x7f71f4000fe0) AdminSocketConfigObs::init: f
nd the UNIX domain socket to '/var/run/ceph/guests/ceph-client.admin.6668.140127401677200.asc
  cluster:
    id:     d5bb0cd0-3598-4a59-9d52-dac5debd3571
    health: HEALTH_WARN
            349/21112 objects misplaced (1.653%)
            Degraded data redundancy: 3828/21112 objects degraded (18.132%), 50 pgs degradec

  services:
    mon: 1 daemons, quorum mon1
    mgr: mon1(active)
    osd: 5 osds: 5 up, 5 in; 1 remapped pgs

  data:
    pools:   4 pools, 320 pgs
    objects: 10.56k objects, 1.09GiB
    usage:   3.72GiB used, 226GiB / 230GiB avail
    pgs:     3828/21112 objects degraded (18.132%)
             349/21112 objects misplaced (1.653%)
             266 active+clean
             50  active+recovery_wait+degraded
             3   active+recovering
             1   active+remapped+backfill_wait

  io:
    recovery: 3.33MiB/s, 15objects/s

[root@ceph-admin ceph]#
```

Això trigarà un temps fins que es corregirà automàticament.

```
  cluster:
    id:     d5bb0cd0-3598-4a59-9d52-dac5debd3571
    health: HEALTH_OK

  services:
    mon: 1 daemons, quorum mon1
    mgr: mon1(active)
    osd: 5 osds: 5 up, 5 in

  data:
    pools:   4 pools, 320 pgs
    objects: 10.56k objects, 1.09GiB
    usage:   3.69GiB used, 226GiB / 230GiB avail
    pgs:     320 active+clean

[root@ceph-admin ceph]#
```

Aquí ja veiem que l'estat és correcte i tenim ja 5 OSDs que veurem amb la comanda tree de ceph.

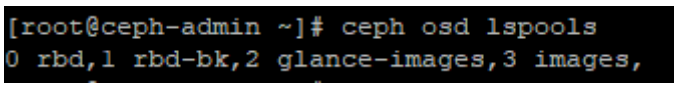
```
ceph osd tree
```



Per últim veurem com crear i esborrar un pool nou i com fer una còpia d'un pool a un altre de nou a partir de snapshots.

Primerament mirarem els pools que tenim amb la comanda següent.

```
ceph osd lspools
```



Com veiem tenim varis pools dels quals esborrarem el rbd-bk i glance-images
Per esborrar-los haurem de fer servir les comandes de ceph següents.

```
ceph osd pool rm rbd-bk
ceph osd pool rm glance-images
```

A l'executar aquestes comandes ceph ens retornarà una advertència de seguretat per indicar-nos el que hem d'afegir per poder esborrar el pool. A més, una vegada executem la comanda correcta ens diu que l'esborrat dels pools està deshabilitat. i haurem de modificar un paràmetre de la configuració de ceph.

```
[root@ceph-admin ~]# ceph osd pool rm glance-images
Error EPERM: WARNING: this will *PERMANENTLY DESTROY* all data stored in pool glance-images.  If you are *ABSOLUTELY CERTAIN* that is what you
 want, pass the pool name *twice*, followed by --yes-i-really-really-mean-it.
[root@ceph-admin ~]# ceph osd pool rm glance-images glance-images  --yes-i-really-really-mean-it
Error EPERM: pool deletion is disabled; you must first set the mon_allow_pool_delete config option to true before you can destroy a pool
```

Per poder esborrar els pools modifiquem el paràmetre per aquesta sessió i ja els podem esborrar.

```
[root@ceph-admin ~]# ceph tell mon.\* injectargs '--mon-allow-pool-delete=true'
injectargs:mon_allow_pool_delete = 'true' (not observed, change may require restart)
[root@ceph-admin ~]#
[root@ceph-admin ~]#
[root@ceph-admin ~]# ceph osd pool rm glance-images glance-images  --yes-i-really-really-mean-it
pool 'glance-images' removed
[root@ceph-admin ~]# ceph osd lspools
0 rbd,1 rbd-bk,3 images,
[root@ceph-admin ~]#
```

Ara farem el mateix pel pool rbd-bk, ja que el tornarem a crear per a fer la còpia de rbd.

```
[root@ceph-admin ~]# ceph osd lspools
0 rbd,3 images,
```

Crearem de nou el pool rbd-bk

```
[root@ceph-admin ~]# ceph osd pool create rbd-bk 64
pool 'rbd-bk' created
[root@ceph-admin ~]#
```

i comprovem que rbd té allotjats objectes, mentre que rbd-bk està buit.

```
[root@ceph-admin ~]# rados -p rbd ls
rbd_data.10236b8b4567.00000000000000c8
rbd_data.10236b8b4567.00000000000000a1
rbd_data.10236b8b4567.000000000000002a
rbd_data.10236b8b4567.0000000000000075
rbd_data.10236b8b4567.00000000000000b1
rbd_data.10236b8b4567.0000000000000018
rbd_data.10236b8b4567.0000000000000054
rbd_data.10236b8b4567.0000000000000036
rbd_data.10236b8b4567.000000000000008d
rbd_data.10236b8b4567.0000000000001405
rbd_data.10236b8b4567.000000000000001f
rbd_data.10236b8b4567.00000000000000c6
rbd_data.10236b8b4567.00000000000000ce
rbd_data.10236b8b4567.0000000000000015
rbd_data.10236b8b4567.00000000000000b6
rbd_data.10236b8b4567.0000000000000002
rbd_data.10236b8b4567.00000000000000da
rbd_data.10236b8b4567.00000000000000de
```

També ho podem veure a escala de volums amb les comandes *rbd ls.*

```
[root@ceph-admin ~]# rbd ls rbd
disk01
[root@ceph-admin ~]# rbd ls rbd-bk
[root@ceph-admin ~]#
```

Arribats a aquest moment ja podem fer un snapshot del volum disk01 en el pool rbd.

```
[root@ceph-admin ~]# rbd snap create rbd/disk01@rbd-snap1
[root@ceph-admin ~]# rbd snap ls rbd/disk01
SNAPID NAME        SIZE TIMESTAMP
     4 snap01     40GiB
     5 rbd-snap1 40GiB Tue May 28 16:04:36 2019
[root@ceph-admin ~]#
```

Ara, per poder copiar aquest snapshot al pool de backup, haurem de protegir el snapshot perquè no hi hagi canvis mentre es mou.

```
[root@ceph-admin ~]# rbd snap protect rbd/disk01@rbd-snap1
```

i ara clonarem el snapshot al pool rbd-bk

```
[root@ceph-admin ~]# rbd clone rbd/disk01@rbd-snap1 rbd-bk/disk01
```

Quan s'hagi clonat, podrem llistar el contingut del pool rbd-bk i veurem que ara si que conté el volum disk01.

```
[root@ceph-admin ~]# rbd ls rbd-bk
disk01
```

Ens hem de recordar de treure la protecció del snapshot de rbd perquè es pugui seguir actualitzant, però si ho fem en aquest moment, ens donarà un error donat que el snapshot de rbd-bk es basa encara en el de rbd i per tant no es pot desprotegir.

```
[root@ceph-admin ~]# rbd snap unprotect rbd/disk01@rbd-snap1
2019-05-28 16:12:01.933940 7f3e47fff700 -1 librbd::SnapshotUnprotectRequest: cannot unprotect: at least 1 child(ren) [25a376b8b4567] in pool '
rbd-bk'
2019-05-28 16:12:01.937267 7f3e47fff700 -1 librbd::SnapshotUnprotectRequest: encountered error: (16) Device or resource busy
2019-05-28 16:12:01.937380 7f3e47fff700 -1 librbd::SnapshotUnprotectRequest: 0x55d70420d260 should_complete_error: ret_val=-16
2019-05-28 16:12:01.942199 7f3e47fff700 -1 librbd::SnapshotUnprotectRequest: 0x55d70420d260 should_complete_error: ret_val=-16
rbd: unprotecting snap failed: (16) Device or resource busy
```

Per poder desvincular els dos snapshots hem d'executar la comanda *flatten* per asseure el snapshot de rbd-bk amb totes les dades de rbd. Aquesta acció trigarà una mica.

```
[root@ceph-admin ~]# rbd flatten rbd-bk/disk01
Image flatten: 100% complete...done.
```

I ja podem desprotegir el snapshot inicial.

```
[root@ceph-admin ~]# rbd snap unprotect rbd/disk01@rbd-snap1
[root@ceph-admin ~]#
```

Annex VII – Instal·lació d'Openstack

Per a la instal·lació d'*Openstack* farem servir El projecte RDO amb el qual instal·larem *Openstack* en tres nodes diferents, *Controller*, *Compute* i *Network*.

Primerament definirem els nodes que necessitarem i la seva configuració.

Node Controller:	Node Compute:	Node Network:
Nom del node: controller.local **Adreça IP:** 192.168.130.141 Mòduls que s'instal·laran: • Keystone • Glance • swift • Cinder • Horizon • Neutron • Nova novncproxy • Novnc • Nova api • Nova Scheduler • Nova-conductor	**Nom del node:** compute.local **Adreça IP:** 192.168.130.142 Mòduls que s'instal·laran: • Nova Compute • Neutron – Openvswitch Agent	**Nom del node:** network.local **Adreça IP:** 192.168.130.143 Mòduls que s'instal·laran: • Neutron Server • Neturon DHCP agent • Neutron- Openswitch agent • Neutron L3 agent

Pas 1 – Instal·lació del Sistema Operatiu.

Es realitzarà una instal·lació minimal de CentOS 7 a tots tres nodes i s'actualitzaran mitjançant yum.

```
yum -y update ; reboot
```

Pas 2 – Actualitzant el fitxer /etc/hosts

Canviarem el nom de cada servidor mitjançant les següents comandes en cas de ser necessari.

```
hostnamectl set-hostname controller

hostnamectl set-hostname compute

hostnamectl set-hostname network
```

Una vegada canviat el nom de la màquina, actualitzarem el fitxer de hosts per associar-ho amb la seva IP, ja que en aquest pilot no tenim un DNS intern on configurar-ho.

```
192.168.130.141    controller.local   controller
192.168.130.142    compute.local  compute
192.168.130.143    local          network
```

Pas 3 – Deshabilitem SELinux i Network Manager als tres nodes

Primer desactivem selinux en la sessió actual i després modifiquem la configuració per fe el canvi permanent en cada reinici.

```
setenforce 0

sed -i 's/SELINUX=enforcing/SELINUX=disabled/g' /etc/selinux/config
```

Després deshabilitarem Network Manager perquè no interfereixi en la configuració de xarxa amb Openstack.

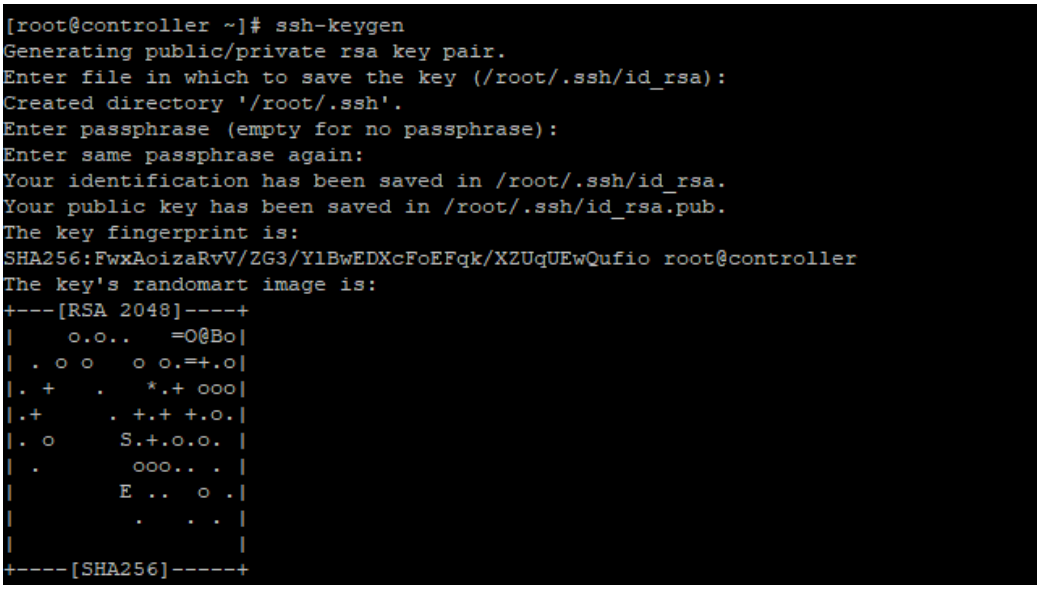
```
systemctl stop NetworkManager

systemctl disable NetworkManager

reboot
```

Reiniciem el servidor per comprovar que els canvis s'han fet correctament.

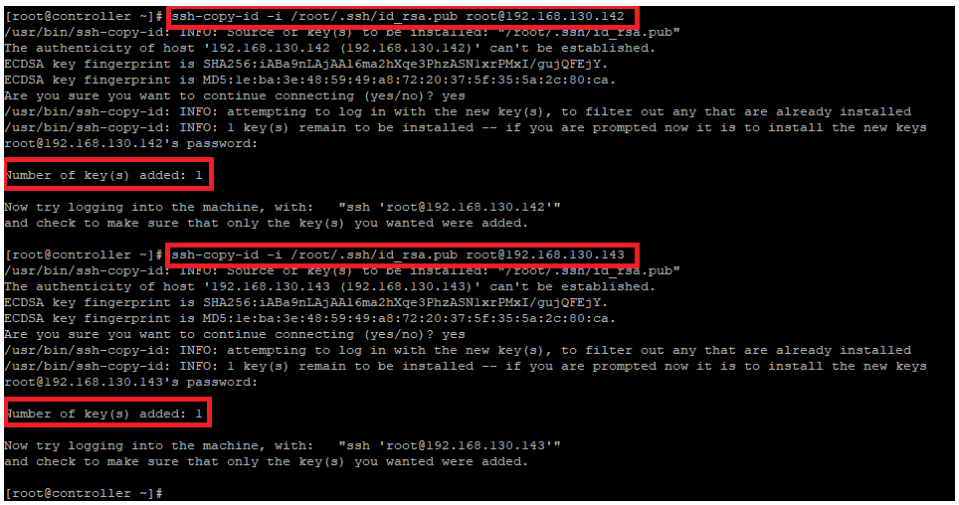
Pas 4 – Configuració per evitar l'autenticació per contrasenya entre el node controlador i els altres dos nodes.

Primer generarem la key del servidor controller i després l'exportarem cap als nodes Compute i Network.

```
[root@controller ~]# ssh-keygen
```



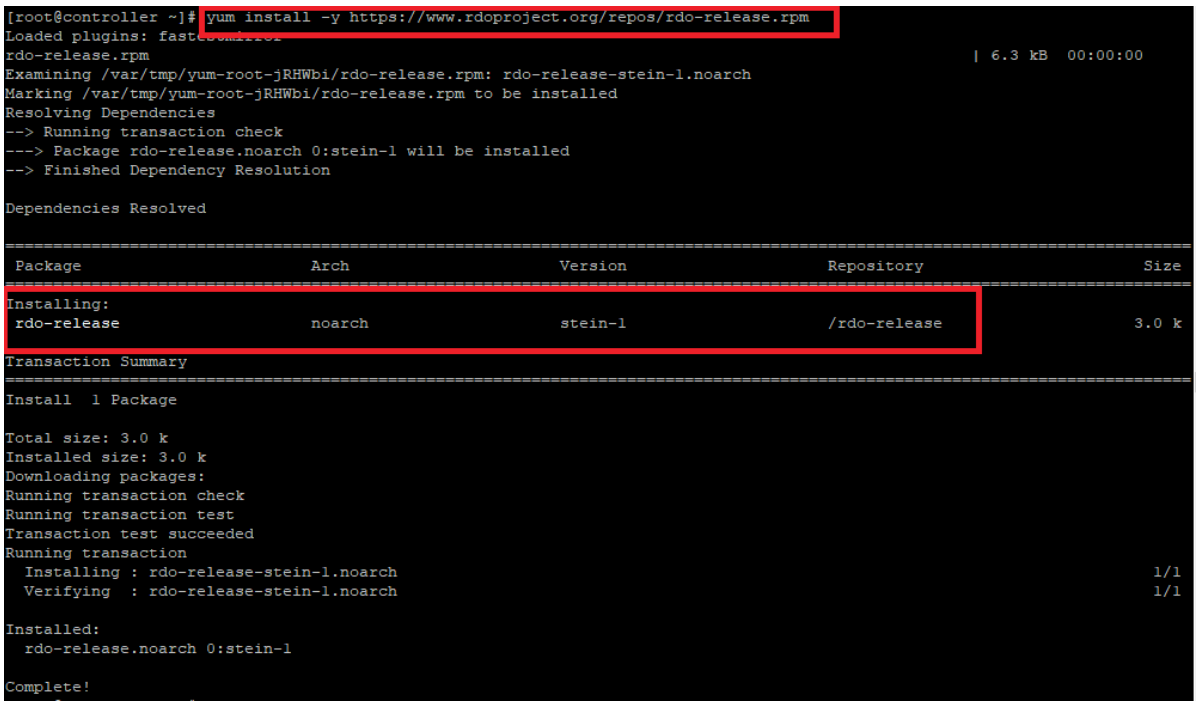

```
[root@controller ~]# ssh-copy-id -i /root/.ssh/id_rsa.pub root@192.168.130.142

[root@controller ~]# ssh-copy-id -i /root/.ssh/id_rsa.pub root@192.168.130.143
```



Pas 5 – Habilitem el repositori RDO per instal·lar packstack

Primer de tot instal·larem el repositori des de rdoproject.org només al node controller i ja podrem instal·lar el paquet d'openstack-packstack.

```
[root@controller ~]#  yum install -y https://www.rdoproject.org/repos/rdo-release.rpm

[root@controller ~]# yum install -y openstack-packstack
```



Pas 6 - Configurem la instal·lació d'Openstack a traves del fitxer answer de packstack.

Primer de tot generarem el fitxer de configuració amb la següent comanda.

```
[root@controller ~]# packstack --gen-answer-file=/root/answer.txt
```

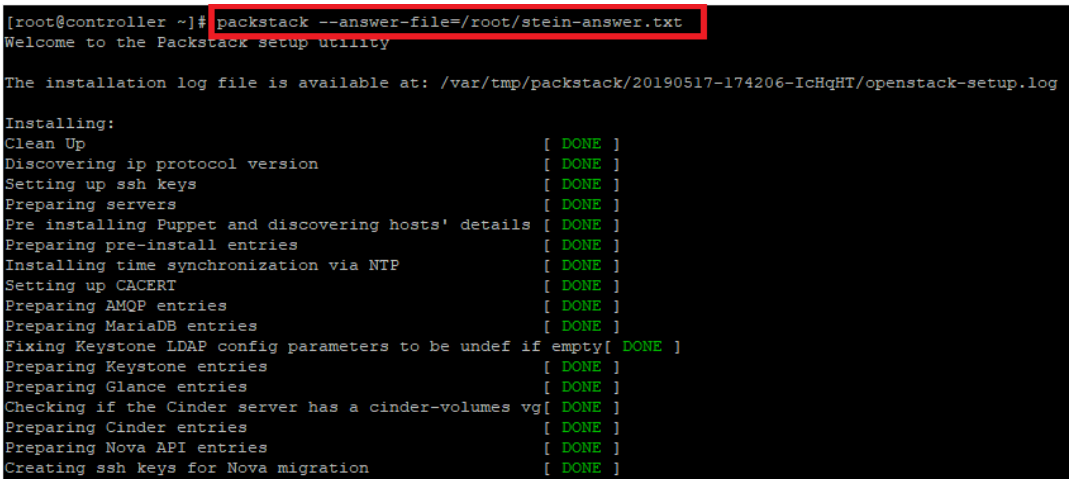
Una vegada generat, l'editem per configurar la instal·lació segons la configuració dels nostres nodes. A més podrem configurar les contrasenyes d'alguns serveis i/o desactivar d'altres com ara Ceilometer, els components de demostració.

```
[root@controller ~]# vi /root/answer.txt

-------------------------------------------------

CONFIG_CONTROLLER_HOST=192.168.130.141

CONFIG_COMPUTE_HOSTS=192.168.130.142

CONFIG_NETWORK_HOSTS=192.168.130.143

CONFIG_PROVISION_DEMO=n

CONFIG_CEILOMETER_INSTALL=n

CONFIG_HORIZON_SSL=y

CONFIG_NTP_SERVERS=es.pool.ntp.org, europe.pool.ntp.org

CONFIG_KEYSTONE_ADMIN_PW='afegim una contrasenya'

-------------------------------------------------
```

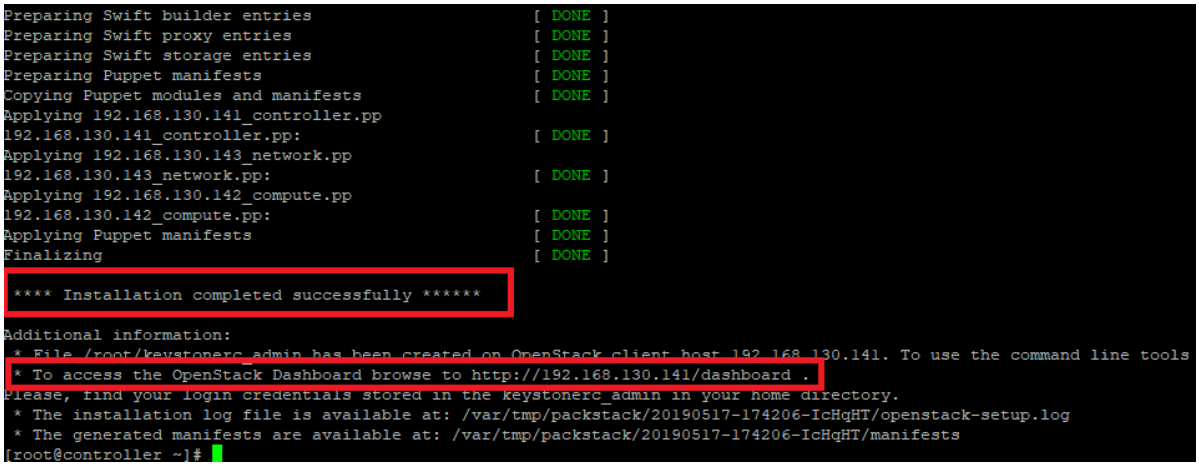
Pas 7 - Instal·lant OpenStack

Ara ja podem instal·lar Openstack des del node Controller.

```
[root@controller ~]# packstack --answer-file=/root/answer.txt
```



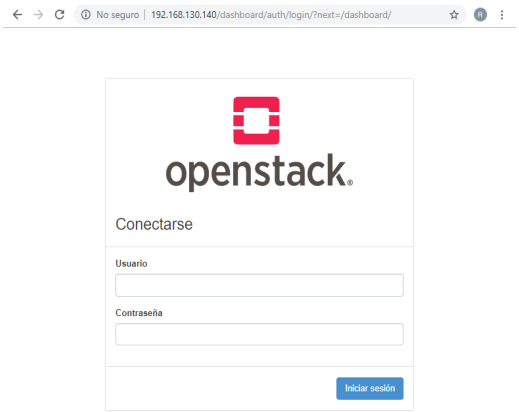
Una vegada acabada la instal·lació ens mostrarà els passos a seguir per accedir al Dashboard.



Pas 8 - Accedint al Dashboard d'Openstack.

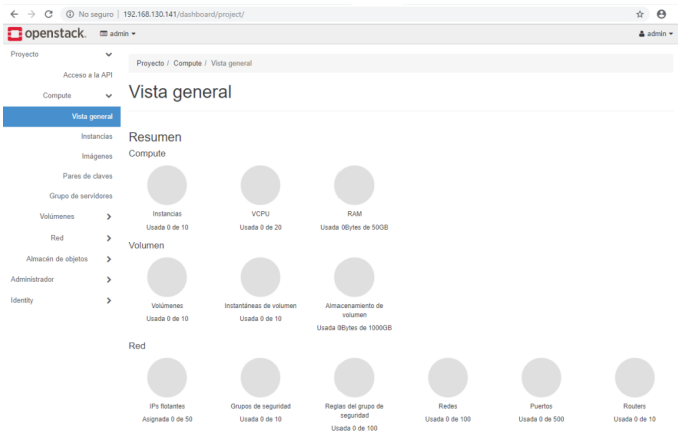
Per accedir al Dashboard d'OpenStack haurem d'introduir la següent URL a qualsevol navegador.

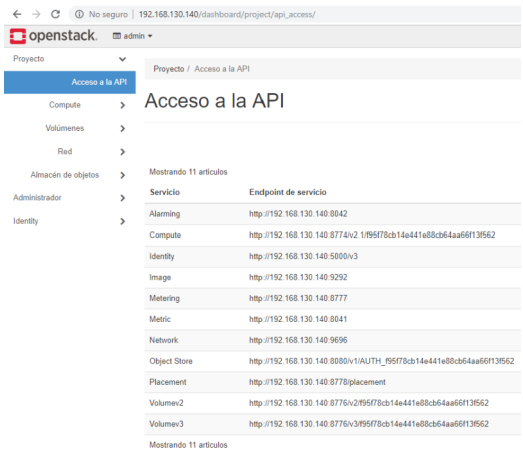
- *https://192.168.130.141/dashboard*



Les dades d'accés les podrem trobar al fitxer «*keystonerc_admin*» que s'haurà creat al directori */root/*. Seran les que hem d'indicar en el parametre *CONFIG_KEYSTONE_ADMIN_PW* del fitxer *answer.txt*.

Una vegada dintre ja podrem accedir a les diferents opcions del panell.





Annex VIII – Estudi de costos

Tenim tres taules diferenciades per localització que ens indiquen els diferents costos que suposarà la implantació del projecte encara que la taula relacionada amb el CPD 3 serà orientativa atès que la implantació d'aquesta part ha estat ajornada.
Els preus indicats no tenen aplicat cap descompte encara, serà a partir de la sol·licitud de pressupost al distribuïdor que ens indicaran els preus ja amb descomptes aplicats. Aquests descomptes poden estar entre el 20% i el 40%.

Xarxa LAN Cablejada			
Tipus de maquinari	**Marca / Model**	**Unitats**	**preu**
Commutador principal	TP-Link / T3700G-28TQ commutador administrat Gigabit L3 Apilable, 28 ports	1	1.742,40€
Commutador Planta 1	Linksys / LGS552P Commutador Gigabit PoE+ administrat de 52 ports amb 2 ports SFP combinats i 2 ports SFP+	1	966,53€
Commutador Planta 2	Linksys / LGS528 Commutador Gigabit PoE+ administrat de 28 ports amb 2 ports SFP combinats	1	334,46€
Encaminador porta d'enllaç / Tallafocs	Sonicwall TZ400 + 1año Total secure	1	1.302,69€
Encaminadors VPN (IPSEC) / Tallafocs	Sonicwall TZ300 (889,83€ unidad)	2	1.779,66€
Cablejat UTP	Cablejat UTP cat6A - Bobina 250m	1	168,19 €
Xarxa LAN WIFI			
Tipus de maquinari	**Marca / Model**	**Unitats**	**Preu**
Punt d'accés WIFI	tp-link / CAP1200 (107,30€ Unidad) Punt d'accés Gigabit Inalámbric de Doble Banda AC1200 amb Muntatge de sostre	2	214,60€
Total			**6.508,53€**

CPD2			
Tipus de maquinari	**Marca / Model**	**Unitats**	**Preu**
Encaminador principal	Cisco / ASR 1001-X (6885,87€ Unidad)	2	13.771,74
Commutadors principals	TP-Link / T3700G-28TQ (1742,40€ Unidad) commutador administrat Gigabit L3 Apilable, 28 ports	5	8.712€
Commutador secundaris	TP-Link / T2700G-28TQ (1136,48€Unidad) commutador administrat Gigabit L2 Apilable, 28 ports	7	7.955,36€
Tallafocs xarxes /23 - /24	WatchGuard Firebox M570 Firewall (6537,26€ Unidad)	2	13.074,52€
Tallafocs xarxes /26 - /27	WatchGuard Firebox M370 Firewall (3211,15€ Unidad)	2	6.422,3€
Cablejat UTP	Cablejat UTP cat6A - Bobina 250m	1	168,19 €
Xarxa Emmagatzemament i backup			
Tipus de maquinari	**Marca / Model**	**Unitats**	**Preu**
Commutadors 48ports	Netgear / NETGEAR XS748T commutador administrat 10Gigabit L2+ / L3, 48ports	1	3.396,31 €
Commutador 28ports	Netgear / NETGEAR XS728T commutador administrat 10Gigabit L2+ / L3, 28ports	1	2.023,56€
Cablejat UTP	Cablejat UTP cat7 (79,15€ 100m)	2	158,3€
Servidors			
Tipus de maquinari	**Marca / Model**	**Unitats**	**Preu**
Servidors OSD	DellEMC PE R740 – 16x2.5- 2CPU 8c- 64GB RDIMM- 56TB (49.568,18€ Unidad)	4	198.272,72€
Servidors Monitors, OS controller	DellEMC PE R640 – 8x2.5- 2CPU 12c- 128GB RDIMM- 2.4TB (10.200,74 € Unidad)	3	30.602,22€
OS Compute	DellEMC PE R440 – 2xM.2- 2CPU 22c - 256GB RDIMM 500GB (28.798,19€ Unidad)	4	115.192,76€
Openstack Fuel Master	DellEMC PE R240 – 2xM.2- 1CPU 8c - 16GB RDIMM 500GB	1	1.757,52 €
Armaris			
Tipus de maquinari	**Marca / Model**	**Unitats**	**Preu**
Armaris (Racks)	Armario rack ImServ 47U 800 x 1200 (1245,43€ Un)	2	2.490,86€
SAIs Online	Lapara 10000VA (2420,19€ Unidad)	2	4.840,38€
Recursos CPD			

Tipus	Marca / Model	Unitats	Preu
Clases IPv4	clase C (256 IP) (98,99/mes)	4	395.96€/mes
Clases IPv6	clase /48 (95,99/mes)	1	95,99€/mes
Ample de banda	1Gbps	1	2100€/mes
Total	**Sense comptar despeses mensuals.**		**401.509,99€**

CPD3			
Tipus de maquinari	**Marca / Model**	**Unitats**	**Preu**
Encaminador principal	Cisco / ASR 1001-X (6885,87€ Unidad)	2	13.771,74
Commutadors principals	TP-Link / T3700G-28TQ (1742,40€ Unidad) commutador administrat Gigabit L3 Apilable, 28 ports	4	6.969,6€
Commutador secundaris	TP-Link / T2700G-28TQ (1136,48€Unidad) commutador administrat Gigabit L2 Apilable, 28 ports	4	1.136,48€
Tallafocs xarxes /23 - /24	WatchGuard Firebox M570 Firewall (6537,26€ Unidad)	2	13.074,52€
Cablejat UTP	Cablejat UTP cat6A - Bobina 250m	1	168,19 €
Xarxa Emmagatzemament i backup			
Tipus de maquinari	**Marca / Model**	**Unitats**	**Preu**
Commutadors 48ports	Netgear / NETGEAR XS748T commutador administrat 10Gigabit L2+ / L3, 48ports	1	3.396,31 €
Commutador 28ports	Netgear / NETGEAR XS728T commutador administrat 10Gigabit L2+ / L3, 28ports	1	2.023,56€
Cablejat UTP	Cablejat UTP cat7	-	158,3€
Servidors			
Tipus de maquinari	**Marca / Model**	**Unitats**	**Preu**
Servidors OSD	DellEMC PE R740 – 16x2.5- 2CPU 8c- 64GB RDIMM- 56TB (49.568,18€ Unidad)	4	198.272,72€
Servidors Monitors, OS controller	DellEMC PE R640 – 8x2.5- 2CPU 12c- 128GB RDIMM- 2.4TB (10.200,74 € Unidad)	3	30.602,22€
OS Compute	DellEMC PE R440 – 2xM.2- 2CPU 22c - 256GB RDIMM 500GB (28.798,19€ Unidad)	4	115.192,76€
Openstack Fuel Master	DellEMC PE R240 – 2xM.2- 1CPU 8c - 16GB RDIMM 500GB	1	1.757,52 €
Armaris			

Tipus de maquinari	**Marca / Model**	**Unitats**	**Preu**
Armaris (Racks)	Armario rack ImServ 47U 800 x 1200 (1245,43€ Un)	2	2.490,86€
SAIs Online	Lapara 10000VA (2420,19€ Unidad)	2	4840,38€
Recursos CPD			
Tipus	**Marca / Model**	**Unitats**	**Preu**
Clases IPv4	clase C (256 IP) (98,99€/mes)	4	395.96/mes
Clases IPv6	clase /48 (95,99€/mes)	1	95,99€/mes
Ample de banda	1Gbps	1	2100€/mes
Total	**Sense comptar despeses mensuals.**		**393.855,16€**

Annex IX – Calcul de capacitat del SAI

Per calcular aquesta capacitat haurem de calcular el cost en wats del nostre maquinari i afegir un mínim d'un 20% més per evitar pics de consum.

CPD2

Dispositiu	Quantitat	Consum per unitat	Consum total
Cisco / ASR-920-4SZ-A	2	75w	150w
TP-Link / T3700G-28TQ	5	63w	315w
TP-Link / T2700G-28TQ	7	63w	441w
WatchGuard Firebox M570 Firewall	2	75w	150w
WatchGuard Firebox M370 Firewall	2	75w	150w
Netgear / NETGEAR XS748T	1	262.8w	261.8
Netgear / NETGEAR XS728T	1	134.9w	134.8w
DELL R740	4	750w	3000w
DELL R640	3	750w	2250w
DELL R440	4	440w	1320w
Consum total			8172,7w
20%			1634,54w
Consum total + 20%			9807,24w

Aquest consum estaria repartit entre els quatre racks contractats al CPD2 on s'hauria d'afegir el consum actual dels servidors que es mantindrien. Tot i això, el consum més elevat que seria provocat pels servidors DELL aniria tot al mateix rack de manera que ens obligaria a instal·lar un SAI de 10000VA en el rack destinat als servidors DELL. La capacitat dels SAIs es mesura en voltsampers i no sempre acostuma a coincidir amb la capacitat en watts. En el SAI que hem triat, el fabricant ens indica la relació (10000W – 10KVA) pel fet que ofereix un factor de potència d'1. Llavors si calculem la relació:

VA = W / FA

VA = 10000w / 1 = 10000VA **= 10KVA**

És un SAI Lapara 10000VA de tipus Online i per instal·lar en un rack de 19''.

Més informació dels dispositius Lapara SAI en l'annex IV.

En els altres racks que ja estan en producció, ja existeixen SAIs adaptats al consum que generen els dispositius que allotgen.

CPD3

Dispositiu	Quantitat	Consum per unitat	Consum total
Cisco / ASR-920-4SZ-A	2	75w	150w
TP-Link / T3700G-28TQ	5	63w	315w
TP-Link / T2700G-28TQ	7	63w	441w
WatchGuard Firebox M570 Firewall	2	75w	150w
Netgear / NETGEAR XS748T	1	262.8w	261.8
Netgear / NETGEAR XS728T	1	134.9w	134.8w
DELL R740	4	750w	3000w
DELL R640	3	750w	2250w
DELL R440	4	440w	1320w
Consum total			8022,7w
20%			1604,54w
Consum total + 20%			9627,24w

En el cas del CPD3 tenim la mateixa situació per la qual cosa instal·larem un SAI de les mateixes característiques en el rack destinat als servidors DELL. Per altra banda, en el cas del CPD3 no hi ha una estructura inicial i per tant s'haurà d'habilitar un segon rack per a la resta de xarxes. Per tal d'assumir un consum similar i evitar haver d'instal·lar més SAIs a posteriori, s'instal·larà un altre SAI de 10000VA al segon rack.
Una vegada tenim els SAIs definits haurem de calcular el temps dels SAIs en mode bateries per poder fer una previsió en les actuacions davant una caiguda de tensió prolongada.

Per poder calcular aquest temps, haurem de fer servir la següent formula:

Temps en minuts de duració d'un SAI / UPS = ((N x V x AH x Eff) / VA) x 60

On:

- **N** = nombre de bateries en el SAI
- **V** = voltatge de les bateries
- **AH** = Amperis-Hora de les bateries
- **Eff** = Eficiència del SAI (pot oscil·lar entre el 90% en mode onlinei el 98% en mode ECO amb un 100% de càrrega)
- **VA** = Volti-Amperis del SAI

En el nostre cas:

- N = 20 bateries
- V = 12v
- AH = 9Ah
- Eff = 0.95% (de mitja)
- VA = 10000VA

Temps en minuts de duració d'un SAI / UPS = ((N x V x AH x Eff) / VA) x 60 =

= ((20 x 12 x 9 x 0.95) /10000) x 60 = 12,312m al 100% de carrega

Finalment explicar que 12,312 minuts pot semblar poca cosa però aquest seria el temps que durarien les bateries amb un consum per part dels dispositius del 100% de la seva capacitat. Inicialment el nostre consum màxim arribarà sobre els 7000w fet que allargaria aquest temps, i a més hem de tenir en compte que el consum calculat és a partir dels màxims que ens donen els fabricants, però aquest no acostuma a produir-se més que en moments puntuals, i la resta del temps poden estar funcionant al 60 o 65% del consum màxim. Per tant, podem estar produint un consum real del 50% de la capacitat del SAI i per tant podríem arribar als 25 minuts.

Per altra banda, aquest tipus de SAI ofereixen la possibilitat d'apagar els servidors connectats de manera segura mitjançant un software que detectarà el temps que queda de càrrega i arribat a un límit establert pot realitzar la tasca que hàgim configurat.

Annex X – Segmentació xarxes IPv4

Per poder saber quines classes es poden unificar per fer la xarxa /23 s'haurà de fer subnetting[16]. Primer de tot traduirem les IPs en format decimal al format binari (el primer terme no influeix i per tant el deixarem com xxxxxxxx)

IP format decimal		IP format binari
x.2.1.0	-->	xxxxxxxx.00000010.00000001.00000000
x.2.2.0	-->	xxxxxxxx.00000010.00000010.00000000
x.2.3.0	-->	xxxxxxxx.00000010.00000011.00000000
x.2.4.0	-->	xxxxxxxx.00000010.00000100.00000000

I la mascara de xarxa.

IP format decimal		IP format binari
23 (255.255.254.0)	-->	11111111.11111111.11111110.00000000
24 (255.255.255.0)	-->	11111111.11111111.11111111.00000000
25 (255.255.255.128)	-->	11111111.11111111.11111111.10000000
26 (255.255.255.192)	-->	11111111.11111111.11111111.11000000
27 (255.255.255.224)	-->	11111111.11111111.11111111.11100000

Ara haurem d'agafar la ip i la seva màscara per veure quines posicions es podran modificar separant la part fixa d'uns de la part modificable de zeros de la mascara.

		Part fixa	Part modificable
IP	-->	xxxxxxxx.00000010.0000000	1.00000000
Mascara	-->	11111111.11111111.1111111	0.00000000

Amb aquesta comparació veiem que podem modificar les últimes 9 posicions i per tant aquesta xarxa estarà formada per les ips compresses per:

```
xxxxxxxx.00000010.0000000 | 0.00000000 x.2.0.0
xxxxxxxx.00000010.0000000 | 0.00000001 x.2.0.1
xxxxxxxx.00000010.0000000 | 0.00000010 x.2.0.2
...
xxxxxxxx.00000010.0000000 | 1.00000000 x.2.1.0
xxxxxxxx.00000010.0000000 | 1.00000001 x.2.1.1
xxxxxxxx.00000010.0000000 | 1.00000010 x.2.1.2
...
xxxxxxxx.00000010.0000000 | 1.11111111 x.2.1.255
```

Com que les classes que formen la xarxa x.2.1.0/23 inclourien la x.2.0.0/24 i la x.2.1.0/24 no es podrien fer servir. Per tant, hauríem de fer servir les dues següents que si que entrarien.

		Part fixa	**Part modificable**
IP	-->	xxxxxxxx.00000010.0000001	0.00000000
Mascara	-->	11111111.11111111.1111111	0.00000000

xarxa formada per les següents IPs:

```
xxxxxxxx.00000010.0000001 | 0.00000000 x.2.2.0
xxxxxxxx.00000010.0000001 | 0.00000001 x.2.2.1
xxxxxxxx.00000010.0000001 | 0.00000010 x.2.2.2
...
xxxxxxxx.00000010.0000001 | 1.00000000 x.2.3.0
xxxxxxxx.00000010.0000001 | 1.00000001 x.2.3.1
xxxxxxxx.00000010.0000001 | 1.00000010 x.2.3.2
...
xxxxxxxx.00000010.0000001 | 1.11111111 x.2.3.255
```

Per tant, seguint el mateix procés de subnetting per a la resta de xarxes obtindrem la segmentació següent:

Xarxa		**Rang d'IPs**	**Hosts utils**
x.2.1.0/24	-->	x.2.1.0 – x.2.1.255 (256 IPs)	254
x.2.2.0/23	-->	x.2.2.0 – x.2.3.255 (512 IPs)	510
x.2.4.0/26	-->	x.2.4.0 – x.2.4.63 (64 IPs)	62
x.2.4.64/26	-->	x.2.4.64 – x.2.4.127 (64 IPs)	62
x.2.4.128/27	-->	x.2.4.128 – x.2.4.159 (32 IPs)	30
x.2.4.160/27	-->	x.2.4.160 – x.2.4.191 (32 IPs)	30
x.2.4.192/27	-->	x.2.4.192 – x.2.4.223 (32 IPs)	30
x.2.4.224/27	-->	x.2.4.224 – x.2.4.255 (32 IPs)	30

Una vegada tenim la segmentació realitzada, podem veure que ens quedaran 1008 IPs utils de les 1024 IPs totals de les 4 classes C. Això es produeix perquè cada xarxa necessita dues IPs per a la seva definició. La primera IP de la xarxa es farà servir com a IP de xarxa (Network), és a dir l'identificador de la xarxa. Per altra banda, l'última IP de la xarxa, anomenada IP de difusió (Broadcast) es fa servir per a la difusió de paquets a tota la xarxa.

Annex XI – Segmentació xarxes IPv6

Es parteix d'una classe /48 d'IPv6 irreal per tal de mostrar un exemple de com s'ha de segmentar la xarxa IPv6

Xarxa		Xarxa		host
x:y:z::/48	-->	000x:000y:000z	:	0000:0000:0000:0000:0000

Aquesta xarxa ens aporta 65536 (ffff en hexadecimal) xarxes del tipus /64

Xarxa		xarxa	Subxarxa	Part modificable
x:y:z:0::/64	-->	000x:000y:000z	0000:	0000:0000:0000:0000
x:y:z:1::/64	-->	000x:000y:000z	0001:	0000:0000:0000:0000
x:y:z:2::/64	-->	000x:000y:000z	0002:	0000:0000:0000:0000
x:y:z:ffff::/64	-->	000x:000y:000z	ffff:	0000:0000:0000:0000

De la part de hosts, a diferents estaments es proposa que la IP assignada estigui formada per l'adreça física (MAC address[17]) però pel caire dinàmic de les xarxes en què ens mourem, ja que acostumen a haver-hi migracions de clients, i també per la comoditat en l'assignació, gestió i manteniment d'aquestes, l'assignació es farà seguint la metodologia que es fa servir a les xarxes d'IPv4.
Es fan dues propostes a l'empresa de cara a la segmentació de les xarxes IPv6. La primera consisteix a fer servir una xarxa /64 per a cada xarxa d'IPv4.

Xarxa		xarxa	Subxarxa	Part modificable
x:y:z:0::/64	-->	000x:000y:000z	0000:	0000:0000:0000:0000
x:y:z:1::/64	-->	000x:000y:000z	0001:	0000:0000:0000:0000
x:y:z:2::/64	-->	000x:000y:000z	0002:	0000:0000:0000:0000
x:y:z:3::/64	-->	000x:000y:000z	0003:	0000:0000:0000:0000
x:y:z:4::/64	-->	000x:000y:000z	0004:	0000:0000:0000:0000
x:y:z:5::/64	-->	000x:000y:000z	0005:	0000:0000:0000:0000
x:y:z:6::/64	-->	000x:000y:000z	0006:	0000:0000:0000:0000
x:y:z:7::/64	-->	000x:000y:000z	0007:	0000:0000:0000:0000

Aquesta proposta ens aportarà una segmentació senzilla i una gran quantitat d'IPs per xarxa que segurament no es faran servir.

L'altra proposta és fer una segmentació similar però més ajustada. Per fer aquesta segmentació es farien servir xarxes /112.

Xarxa		xarxa	Subxarxa	Hosts
x:y:z::0:0/112	-->	000x:000y:000z:0000:0000:0000	:0000	:0000
x:y:z::1:0/112	-->	000x:000y:000z:0000:0000:0000	:0001	:0000
x:y:z::2:0/112	-->	000x:000y:000z:0000:0000:0000	:0002	:0000
x:y:z::3:0/112	-->	000x:000y:000z:0000:0000:0000	:0003	:0000
x:y:z::4:0/112	-->	000x:000y:000z:0000:0000:0000	:0004	:0000
x:y:z::5:0/112	-->	000x:000y:000z:0000:0000:0000	:0005	:0000
x:y:z::6:0/112	-->	000x:000y:000z:0000:0000:0000	:0006	:0000
x:y:z::7:0/112	-->	000x:000y:000z:0000:0000:0000	:0007	:0000

Aquesta segmentació ens aportaria xarxes més petites amb 65536 IPs per xarxa, però totalment suficients per a les xarxes que es necessiten. Per tant es triarà aquesta segmentació.

Annex XII – Rangs d'IPs internes IPv4 i IPv6

IPs Internes IPv4

Rang d'IPs	Nº d'IPs	Nº De xarxes	IPs / xarxa
10.0.0.0 – 10.255.255.255	16.777.214	1	16.777.214
172.16.0.0 – 172.31.255.255	1.048.574	16	65.534
192.168.0.0 – 192.168.255.255	65.534	256	254
169.254.0.0 – 169.254.255.255	65.534	1	65.534

Taula extreta de la URL: https://es.wikipedia.org/wiki/Red_privada

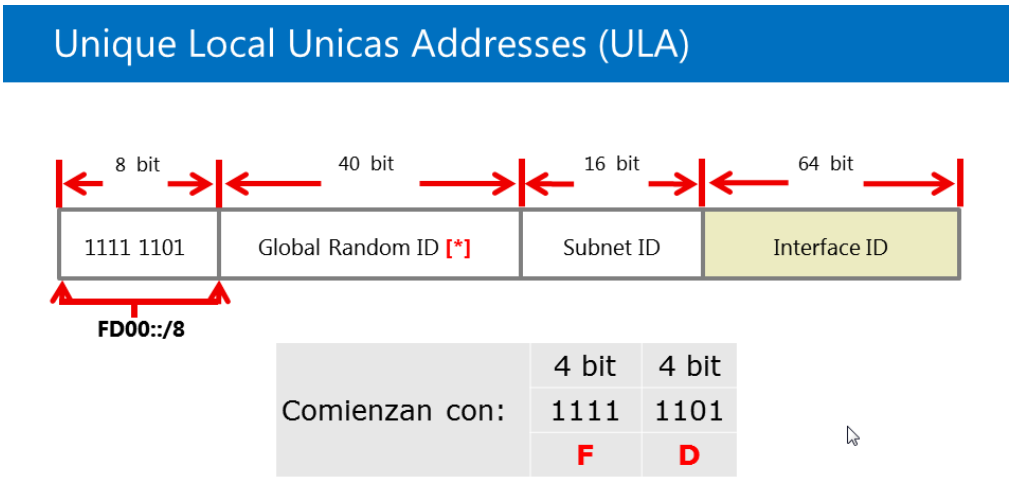
IPs Internes IPv6

Xarxa	Nº d'IPs	Tipo
fd00::/8	2^120 IPs (1,329227995784915872903807060280 3e+36)	Unique Local Address

Host inicio	fd00::1
Host final	fdff:ffff:ffff:ffff:ffff:ffff:ffff:ffff

A més, encara que es podria fer servir tot el rang per a una xarxa interna, el que es recomana és seguir l'estructura de la segmentació explicada abans per crear xarxes més petites de mida /64 o menor.

L'estructura seria la següent:



Imatge extreta de l'URL: https://windowserver.wordpress.com/2013/09/25/aprendiendo-ipv6-clases-de-direcciones-ipv6/

Com veiem, tenim un primer segment de 8bits que identificaria el tipus d'IP com internes. Després tindríem un segon segment de 40bits que seria un identificador en les adreces públiques i que en aquest cas pot ser aleatori, ja que està dintre del rang d'IPs internes. El següent segment de 16bits seria l'identificador de la subxarxa que volem definir i que completaria la part fixa d'una xarxa /64. Per últim, la resta fins als 128bits de l'adreça (64 bits) serien l'identificador únic del dispositiu assignat.
Com a exemple podem tenir una xarxa formada per les següents dades:

Xarxa /8	prefix	ID global	ID subxarxa	Xarxa /64
fd00::/8	fd	43d7f9fc7b	de4a	fd43:d7f9:fc7b:de4a::/64
Inici de Rang	fd43:d7f9:fc7b:de4a:0:0:0:0			
Fi de Rang	fd43:d7f9:fc7b:de4a:ffff:ffff:ffff:ffff			
Nº de hosts	2^64 IPs (18.446.744.073.709.551.616)			

Annex XIII – Nivells de backup

Backup a nivell de Bloc/Volum

A aquest nivell OpenStack + CEPH ofereixen una sèrie d'eines que ens aporten tres tipus de còpia diferent. Per una part, *Cinder Backup* amb *CEPH backup driver,* podrem copiar volums *CEPH RADOS*[22] *block Device (RBD*[21]*)* des d'OpenStack de forma completa, diferencial i incremental. Per defecte, OpenStack està configurat per treballar a Swift[23] com a sistema d'emmagatzemament, però, com s'ha comentat abans, es configurarà per treballar amb CEPH en comptes de Swift. Això comportarà que el sistema de còpies d'OpenStack (*Cinder Backup*) faci servir el *CEPH backup Driver* de forma seminativa.

Configuració del driver de CEPH en el fitxer de configuració de Cinder.
```
# backup_driver = cinder.backup.drivers.ceph
```

Exemples de gestió de backups.

Creació d'un backup
```
# openstack volume backup create [--incremental] [--force] VOLUME
```

Recuperació d'un backup
```
# openstack volume backup restore BACKUP_ID VOLUME_ID
```

A les dues comandes anteriors veiem els exemples de com es crearien còpies d'un volum concret i com es recuperaria aquest volum a partir dels identificadors del backup i el volum. A més, aquesta recuperació ens permetrà restaurar les còpies en el volum original o en un de nou indistintament.

Per altra banda tenim *CEPH mirroring,* que ens aporta la capacitat de replicar el cluster CEPH en una infraestructura secundària. Aquest tipus de rèplica es produeix de forma asíncrona en estructura de primari-secundari o primari-primari. El fet de triar una o d'altra opció ens permetrà sincronitzar les dades en una única direcció cap al secundari o poder sincronitzar en totes dues direccions respectivament. En el cas que ens ocupa optarem per una estructura de clústers (pools) primari-secundari per tal de fer la sincronització sobre el clúster (pool) secundari que no serà en producció. Aquesta sincronització es pot configurar també a dos nivells diferents.

- En mode *Pool*: on es replicarà totes les imatges del pool configurat.
- En mode *Image*: On podrem seleccionar quines imatges del pool volem copiar.

Com a última opció tenim la generació de snapshots (instantànies). Inicialment aquesta opció és utilitzada per fer «fotografies» puntuals que ens permetin tornar enrere en

escenaris d'actualitzacions de sistemes, instal·lacions de programari, o modificacions importants de dades. En cas d'incidència, sempre es podran desfer els canvis tornant a un punt de còpia anterior. Però, CEPH també ens ofereix la possibilitat d'exportar aquestes «fotografies» a clústers externs com a fitxers de diferencies que poden ser guardats o inclús importats en l'estructura CEPH de destí per tal de tenir un clúster sincronitzat. A més, té la capacitat de fer aquestes còpies de forma incremental de tal manera que reduïm la quantitat de dades que es traspassen, optimitzant així el consum d'amplada de banda i per tant la còpia en estructures separades geogràficament.

Un exemple de les comandes que es fan servir per a aquest tipus de còpies seria semblant a aquest:

Creació dels snapshots (es farien amb una separació en el temps suficient)

```
# rbd snap create pool/imatge@snap1

# rbd snap create pool/imatge@snap2
```

Exportació de les diferències dels snapshots.

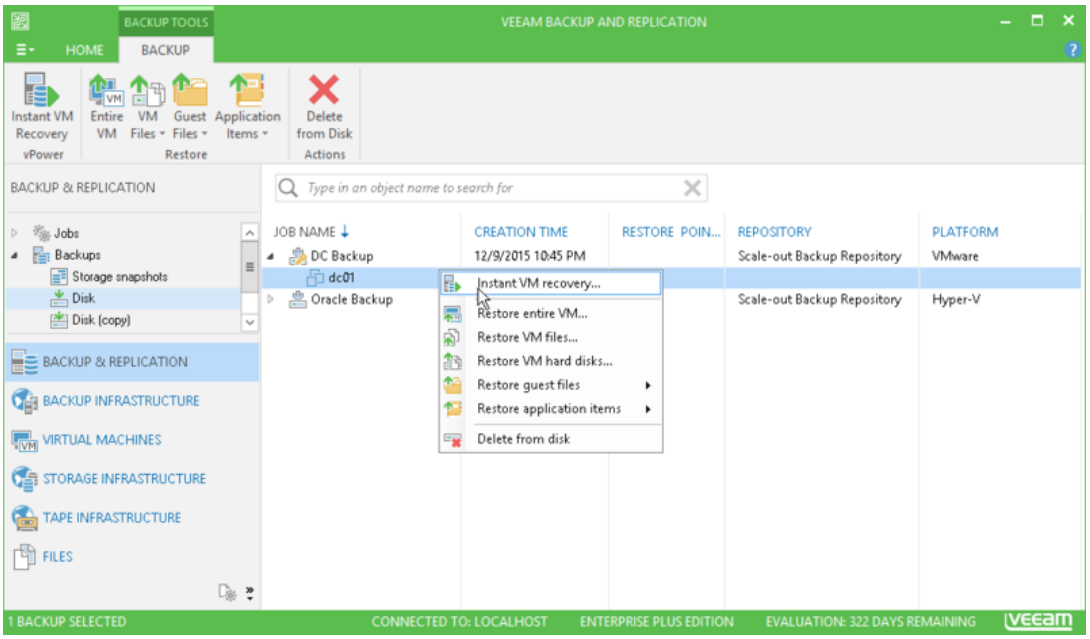
```
# rbd export-diff --from-snap snap1 pool/image@snap2 pool_image_snap1_to_snap2.diff
```

Exportaríem únicament les diferències entre el snapshot «*snap1*» i l'snapshot «`snap2`» de la imatge «`image`» del pool «`pool`» i el guardaríem amb el nom «`pool_image_snap1_to_snap2.diff`». Una vegada tenim el fitxer l'hauríem de traspassar al cluster/pool de backup i importar-ho.

```
# scp ./pool_image_snap1_to_snap2.diff user@backup_cluster:/home/user

En el cluster de backup...

# rbd import-diff /home/user pool_image_snap1_to_snap2.diff
```

Per a gestionar les còpies de qualsevol de les tres opcions s'hauran de programar per tal de realitzar-les periòdicament i mantenir un historial controlat.

Per altra banda tenim Veeam Backup per a entorns de virtualització basats en programari VMware. Aquest programari s'ha d'instal·lar sobre un servidor windows dedicat i connectat al sistema de virtualització. A diferència dels comentats per CEPH que serien opensource, aquest programari necessita llicenciament tant per el mateix programari de backup com pels diferents programaris de virtualització anomenats *VMware ESXi*. Aquest sistema ens proporciona la capacitat de realitzar còpies de les imatges completes de les màquines virtuals de forma complerta o incremental, especificant punts de restauració i amb la capacitat de recuperació de les màquines virtuals senceres o granularment.



Imatge extreta de l'URL: https://en.m.wikipedia.org/wiki/Veeam_Backup_%26_Replication
Especificacions de llicencia: https://en.m.wikipedia.org/wiki/File%3AVeeam-instatnt-vm-recovery.png

En el cas que ens ocupa, aquest sistema ja està en funcionament per a un cloud híbrid que conviurà amb el sistema triat d'OpenStack + CEPH.

Backup a escala de Fitxers

Per a aquesta tasca, el programari triat ha estat bacula primerament per ser programari opensource i segon per la seva versatilitat tant en entorns windows com en entorns linux. El sistema funciona com un sistema Tier-2, és a dir servidor – client on el servidor serà l'encarregat de gestionar les còpies en la seva totalitat i els clients només hauran de permetre les crides del servidor. El programari consta de tres dimonis principals.

- **Director daemon**: Serà l'encarregat de gestionar els clients, la programació de les còpies, i el seu contingut.
- **Storage daemon**: Serà l'encarregat de gestionar l'emmagatzemament on s'han de guardar les còpies.
- **File Daemon**: És el dimoni que formarà part dels clients i permetrà la connexió del servidor amb ells.

Aquest és un sistema molt simple de còpies però molt ràpid i robust per a la tasca que li volem donar. En el nostre cas, només es farà servir per a aquelles còpies contractades pels clients en les que només es vulgui copiar fitxers o directoris específics. A més, aquestes còpies es realitzaran sobre un pool específic del clúster CEPH de producció i per tant també replicat al CEPH de backup.

Backup a escala d'aplicació

En aquest nivell, és difícil definir res, atès que dependrà de l'aplicació des de la qual es faci l'exportació i sempre serà un backup molt personalitzar. Com a exemple més clàssic, tenim una exportació d'una base de dades mysql de la que es vol fer copia. Per a fer aquesta copia/exportació haurem de fer servir la següent comanda que ens permetrà fer còpies de totes les bases de dades o només d'allò que volem copiar.

```
# mysqldump --uadmin -p --all-databases > copiadeseguretat.sql
# mysqldump --uadmin -p base_de_dades > copiadeseguretat_base_de_dades.sql
```

Amb l'execució d'aquestes comandes obtindrem un fitxer en mode text on es guardarà allò que hem especificat en forma de consultes SQL. Fet que ens facilitarà la seva recuperació, ja que només s'hauran d'executar en un client mysql.

A més, aquests fitxers es poden incloure en les còpies a escala de fitxer per a mantenir un històric controlat.

Annex XIV – Enllaços web consultats

Redes Ethernet

- *https://en.wikipedia.org/wiki/Ethernet*
- *https://www.techopedia.com/definition/5280/ethernet*
- *https://searchnetworking.techtarget.com/definition/Ethernet*
- *https://www.linksys.com/es/r/resource-center/que-es-ethernet/*
- *https://smr.iesharia.org/wiki/doku.php/rde:ut2:ethernet*
- *https://es.wikipedia.org/wiki/ALOHAnet*
- *https://www.monografias.com/trabajos105/ethernet-fundamentos-redes/ethernet-fundamentos-redes.shtml*
- *https://es.wikipedia.org/wiki/Ethernet*
- *https://www.prometec.net/tcpip/*
- *https://www.linksys.com/es/r/resource-center/que-es-ethernet/*
- *https://www.lifewire.com/what-is-an-ip-address-2625920*
- *https://searchnetworking.techtarget.com/definition/TCP-IP*
- *https://es.wikipedia.org/wiki/Gigabit_Ethernet*
- *https://es.wikipedia.org/wiki/10_Gigabit_Ethernet*
- *https://www.google.com/intl/en/ipv6/statistics.html#tab=per-country-ipv6-adoption*
- *http://www.seaccna.com/modelo-osi-guia-definitiva/*
- *https://es.wikipedia.org/wiki/Modelo_OSI*

IEEE 802: Estandards per a les xarxes d'area local (LAN)

- *http://www.ieee802.org/*
- *http://www.ieee802.org/3/*

Redireccions NAT: Network Address Translation

- https://www.cisco.com/c/en/us/support/docs/ip/network-address-translation-nat/26704-nat-faq-00.html

IPv6: Internet Protocol version 6

- *https://www.ietf.org/rfc/rfc2460.txt*
- *https://www.internetsociety.org/resources/2018/state-of-ipv6-deployment-2018*

CEPH

- https://www.redhat.com/en/resources/red-hat-ceph-storage-hardware-selection-guide
- https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/3/html-single/red_hat_ceph_storage_hardware_selection_guide/index
- https://ceph.com/geen-categorie/zero-to-hero-guide-for-ceph-cluster-planning/
- http://docs.ceph.com/docs/jewel/start/hardware-recommendations/
- http://docs.ceph.com/docs/master/rbd/iscsi-requirements/
- https://www.supermicro.com/solutions/storage_ceph.cfm
- https://www.supermicro.com/white_paper/white_paper_Ceph-Ultra.pdf
- https://ceph.com/
- http://docs.ceph.com/docs/mimic/rbd/iscsi-target-cli/
- http://docs.ceph.com/docs/mimic/rbd/iscsi-targets/
- http://docs.ceph.com/docs/master/start/hardware-recommendations/#hardware-recommendations
- https://docs.okd.io/3.6/install_config/storage_examples/ceph_rbd_dynamic_example.html
- https://docs.openstack.org/openstack-ansible/latest/user/ceph/full-deploy.html
- https://www.techrepublic.com/article/how-to-deploy-a-ceph-storage-cluster/
- http://docs.ceph.com/docs/master/
- https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/3/
- https://www.virtualtothecore.com/adventures-ceph-storage-part-1-introduction/
- https://searchvmware.techtarget.com/tip/Integrate-Ceph-object-storage-in-a-VMware-vSphere-environment
- https://www.howtoforge.com/tutorial/how-to-build-a-ceph-cluster-on-centos-7/
- https://linoxide.com/storage/setup-red-hat-ceph-storage-centos-7-0/
- https://manuelfrancoblog.wordpress.com/2018/02/19/ceph/
- http://docs.ceph.com/docs/infernalis/install/manual-deployment/#adding-osds

CEPH backup

- https://ceph.com/planet/rbd-ceph-backup/
- http://docs.ceph.com/docs/luminous/rbd/rbd-mirroring/
- https://ceph.com/dev-notes/incremental-snapshots-with-rbd/
- https://github.com/magusnebula/ceph_backup_script/blob/master/ceph_rbd_backup.sh
- https://nicksabine.com/post/ceph-backup/
- https://docs.openstack.org/cinder/latest/admin/blockstorage-volume-backups.html
- https://docs.openstack.org/python-openstackclient/pike/cli/command-objects/volume-backup.html
- https://www.wogri.at/scripts/ceph-vm-backup/

- https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/3/html/block_device_guide/block_device_mirroring#rbd-mirroring-configuring-pool-mirroring
- https://javierin.com/cuantas-iops/

CEPH y OpenStack

- https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/2/html-single/ceph_block_device_to_openstack_guide/index
- https://ceph.com/geen-categorie/ceph-openstack-part-1-2/
- https://www.golinuxcloud.com/steps-to-install-and-configure-controller-node-in-openstack/
- https://docs.openstack.org/openstack-ansible/latest/user/ceph/full-deploy.html
- https://themeanti.me/technology/2018/08/23/ceph_erasure_openstack.html
- http://docs.ceph.com/docs/jewel/rados/operations/erasure-code/
- https://docs.openstack.org/kolla-ansible/rocky/reference/ceph-guide.html#managing-ceph
- https://docs.openstack.org/ocata/config-reference/clustering/api.html

iSCSI

- https://ceph.com/community/new-in-nautilus-ceph-iscsi-improvements/
- http://docs.ceph.com/docs/mimic/rbd/iscsi-overview/
- http://docs.ceph.com/docs/mimic/rbd/iscsi-monitoring
- https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/3/html/block_device_guide/using_an_iscsi_gateway

VMWare

- https://code.vmware.com/web/sdk/6.7/vsphere-automation-rest
- https://www.vmware.com/support/pubs/sdk_pubs.html
- https://www.vmware.com/es/products/vsphere.html

VSAN

- https://docs.vmware.com/en/VMware-vSAN/index.html
- https://virtualizadesdezero.com/vmware-vsan-que-es-y-como-funciona/
- https://www.vmware.com/es/products/vsphere/vsphere-hol.html
- https://www.vmware.com/es/products/hyper-converged-infrastructure/dell-emc-vxrail.html
- https://www.vmware.com/es/products/hyper-converged-infrastructure/dell-emc-vxrail.html

Monitorizacion

- https://ceph.com/planet/the-ceph-monitoring-challenge-prometheus-grafana-and-ansible-rise-to-the-task/
- http://docs.ceph.com/docs/nautilus/mgr/dashboard/#accessing-the-dashboard
- https://www.redhat.com/en/blog/infrastructure-monitoring-service
- https://docs.datadoghq.com/logs/log_collection/?tab=tailexistingfiles#custom-log-collection
- https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/3/html/monitoring_ceph_for_red_hat_enterprise_linux_with_nagios/index
- https://docs.datadoghq.com/getting_started/
- https://blog.pandorafms.org/es/cacti-vs-nagios-vs-pandora-fms/
- https://wiki.pandorafms.com/index.php?title=Pandora:Documentation
- https://access.redhat.com/documentation/en-us/red_hat_ceph_storage/3/html/monitoring_ceph_for_red_hat_enterprise_linux_with_nagios/index

Monitorización CEPH (prometeus + grafana)

- https://maxrohde.com/2018/01/23/setting-up-prometheus-and-grafana-for-centos-rhel-7-monitoring/
- https://computingforgeeks.com/install-grafana-and-influxdb-on-centos-7/
- https://computingforgeeks.com/install-prometheus-server-on-centos-7/
- https://www.enigma14.eu/martin/blog/2017/09/20/prometheus-installation-on-centos-7/
- https://www.fosslinux.com/8328/how-to-install-and-configure-grafana-on-centos-7.htm
- https://www.digitalocean.com/community/tutorials/how-to-install-prometheus-using-docker-on-centos-7
- http://yallalabs.com/linux/how-to-install-grafana-using-mysql-mariadb-database-on-centos-7-rhel-7/
- https://prometheus.io
- https://computingforgeeks.com/how-to-install-grafana-on-rhel-8/
- https://computingforgeeks.com/monitoring-ceph-cluster-with-prometheus-and-grafana/
- https://computingforgeeks.com/how-to-monitor-linux-server-performance-with-prometheus-and-grafana-in-5-minutes/
- https://grafana.com
- https://www.enigma14.eu/martin/blog/2017/09/20/prometheus-installation-on-centos-7/
- https://www.fosslinux.com/8328/how-to-install-and-configure-grafana-on-centos-7.htm
- https://www.fosslinux.com/10398/how-to-install-and-configure-prometheus-on-centos-7.htm

Alertas prometheus

- https://medium.com/@abhishekbhardwaj510/alertmanager-integration-in-prometheus-197e03bfabdf
- https://daenney.github.io/2018/04/21/setting-up-alertmanager
- https://prometheus.io/docs/alerting/configuration/
- https://itnext.io/prometheus-with-alertmanager-f2a1f7efabd6
- https://github.com/messagebird/sachet
- https://medium.com/@zhimin.wen/custom-notifications-with-alert-managers-webhook-receiver-in-kubernetes-8e1152ba2c31
- https://www.robustperception.io/using-slack-with-the-alertmanager
- https://www.robustperception.io/using-pagerduty-with-the-alertmanager

Maquinari DellEMC

- https://www.dell.com/es-es/work/shop/servidores-almacenamiento-y-redes/smart-value-flexi-poweredge-r740-8x25-4110-1x16gb-1x300gb-15k-sas-h330-3y-nbd/spd/poweredge-r740/PER7400
- https://www.dell.com/es-es/work/shop/servidores-almacenamiento-y-redes/smart-value-flexi-poweredge-r640-8x25-4116-1x16gb-1x300gb-15k-sas-h740p-3y-nbd/spd/poweredge-r640/per6402
- https://i.dell.com/sites/csdocuments/Shared-Content_data-Sheets_Documents/en/dell-poweredge-sas-ssd-performance-specifications.pdf

Router Cisco ASR 1001-X

- https://www.cisco.com/c/es_mx/support/routers/asr-1001-x-router/model.html

Switch TPLINK

- https://www.tp-link.com/es/business-networking/managed-switch/t3700g-28tq/#specifications
- https://www.tp-link.com/es/business-networking/managed-switch/t2700g-28tq/#specifications

WiFI UAP-AC PRO AP

- https://store.ui.com/products/unifi-ac-pro
- https://store.ui.com/products/unifi-ac-lr
- UniFi Switch L2 Poe 24p
- https://store.ui.com/collections/routing-switching/products/unifi-switch-l2-poe

- Managed 48-port L2 Gigabit PoE+ Switch
- https://store.ui.com/collections/routing-switching/products/l2-managed-poe-gigabit-switch-with-sfp

Switch L3

- https://www.alliedtelesis.com/products/switches/x610-24ts
- https://www.senetic.es/allied_telesis/allied_telesis_conmutadores/8000gs_series/
- https://dl.ubnt.com/datasheets/unifi/UniFi_AC_APs_DS.pdf

switch 10G

- http://www.downloads.netgear.com/files/GDC/datasheet/en/XS708T_XS716T_XS728T_XS748T.pdf

SAI

- https://www.rackonline.es/content/que-es-un-sai-y-tipos-de-sai
- https://www.dns-system.es/sai-rack-19-online-c-29_198.html
- https://www.ibertronics.com/Alimentacion/UPS/Lapara/On-line/Rack-19
- https://www.sai-online.es/sai-online/sai-rack-10000-va-10kva-online-lcd-lapara

xarxa interna

- https://www.linksys.com/es/p/P-LGS528P/
- https://www.linksys.com/es/p/P-LGS552P/

Firewalls

- https://www.firewalls.com/products/firewalls/watchguard/firebox/firebox-m370
- https://www.firewalls.com/products/firewalls/watchguard/firebox/firebox-m570
- https://www.sonicwall.com/products/firewalls/mid-range/
- https://www.juniper.net/us/en/products-services/security/srx-series/
- https://d3ik27cqx8s5ub.cloudfront.net/media/uploads/2019/03/Datasheet-TZ-Series-US-VG-MKTG5541.pdf
- https://www.watchguard.com/wgrd-products/appliances-compare/15016/15026
- https://www.cisco.com/c/es_es/solutions/enterprise-networks/product-listing.html

Software per el disseny de xarxes.

- https://www.draw.io/
- https://www.ultratools.com/tools/rangeGeneratorResult?globalId=&subnetId=
- https://www.calculadora-redes.com/
- http://www.raid-calculator.com/default.aspx
- https://ceph.com/pgcalc/

Annex XV – Proves de rendiment de Ceph

(Proves extretes de la wiki de Ceph.

- https://tracker.ceph.com/projects/ceph/wiki/Benchmark_Ceph_Cluster_Performance)

Com que tenim un client amb una imatge RBD muntada com a unitat podem fer una prova de rendiment creant un fitxer d'1GB en aquesta unitat per veure com es comporta Ceph.

Sobretot, hem de tenir en compte que aquesta estructura es un pilot que està muntat sobre maquines virtuals i per tant hi haurà limitacions provocades per la compartició de recursos. Per altra banda, la xarxa que connecta els dos servidors del pilot son xarxes a Gigabps.

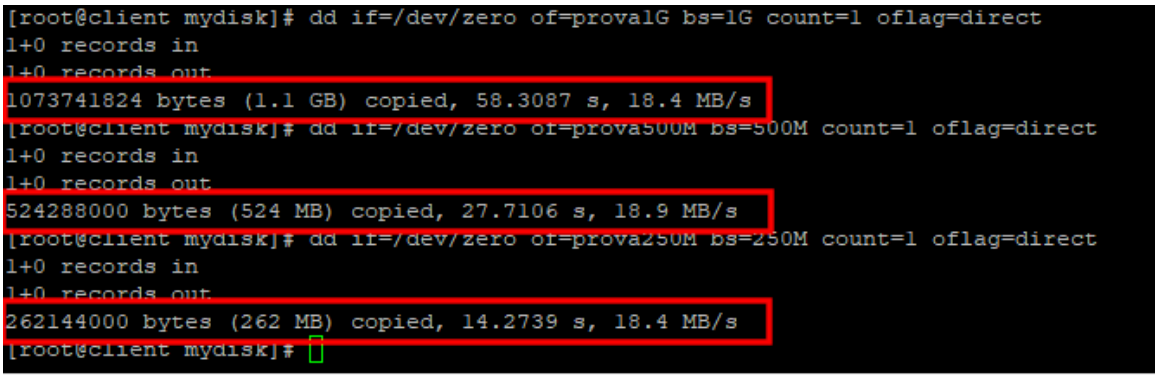
Primerament ens connectarem al servidor client i ens situarem en el directori on tenim muntat la unitat *rbd*.

```
# ssh root@client
# cd /mnt/mydisk
```

Ara executarem la següent ordre *dd* per crear el fitxer d'un Giga directament al directori.

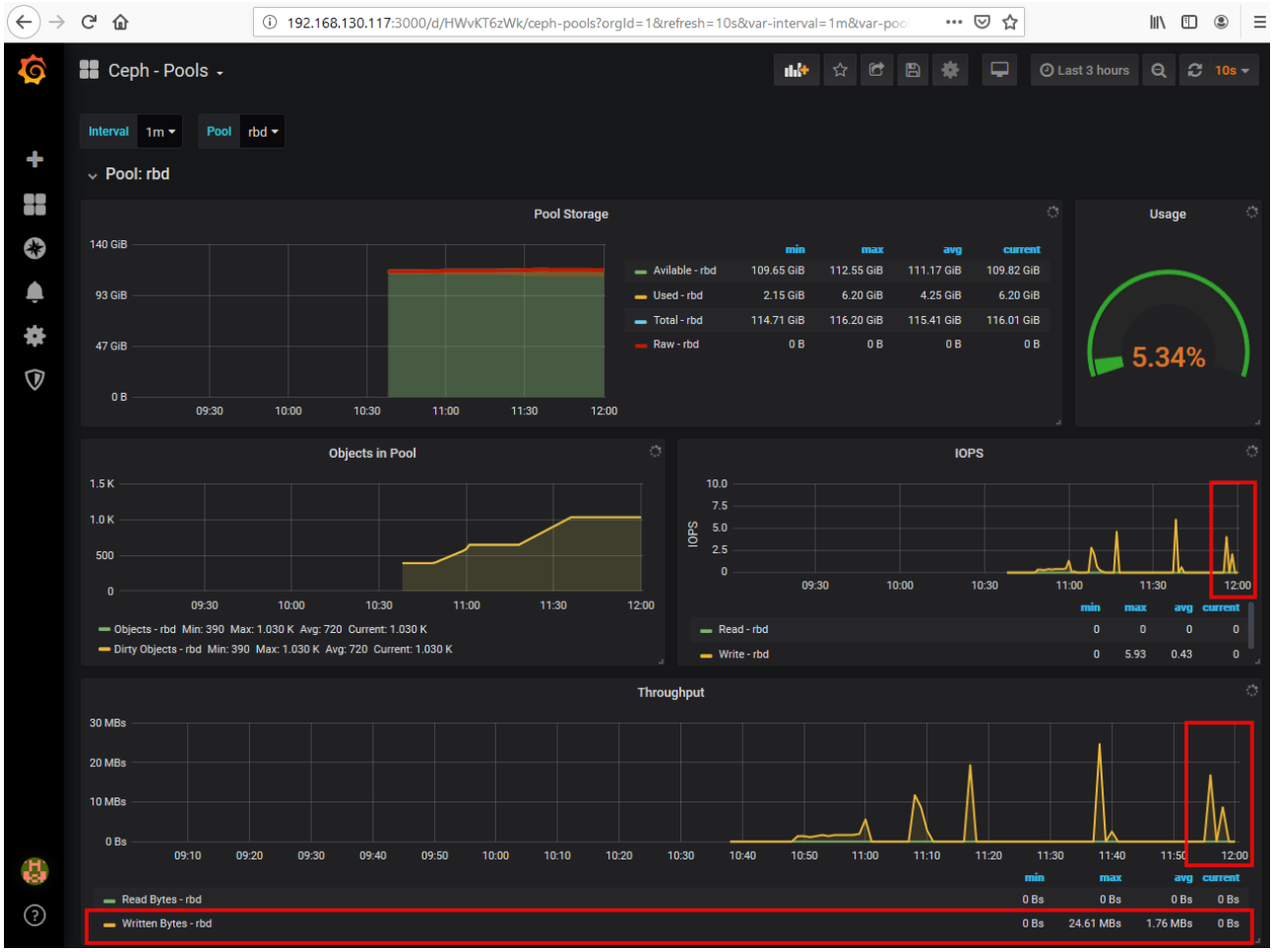
```
# dd if=/dev/zero of=prova1G bs=1G count=1 oflag=direct
# dd if=/dev/zero of=prova500M bs=500M count=1 oflag=direct
# dd if=/dev/zero of=prova250M bs=250M count=1 oflag=direct
```

Després d'una estona, el fitxer s'haurà creat i ens retornarà una estadística de les dades copiades, el temps trigat i la velocitat de transferència aconseguida.
Provarem amb diferents mides de fitxer per tenir una referencia en la velocitat de transferència.



```
[root@client mydisk]# ls -la
total 1816576
drwxr-xr-x  2 root root         55 Jun  6 12:12 .
drwxr-xr-x. 3 root root         20 Jun  4 12:52 ..
-rw-r--r--  1 root root 1073741824 Jun  6 11:54 prova1G
-rw-r--r--  1 root root  262144000 Jun  6 11:59 prova250M
-rw-r--r--  1 root root  524288000 Jun  6 11:57 prova500M
[root@client mydisk]#
```

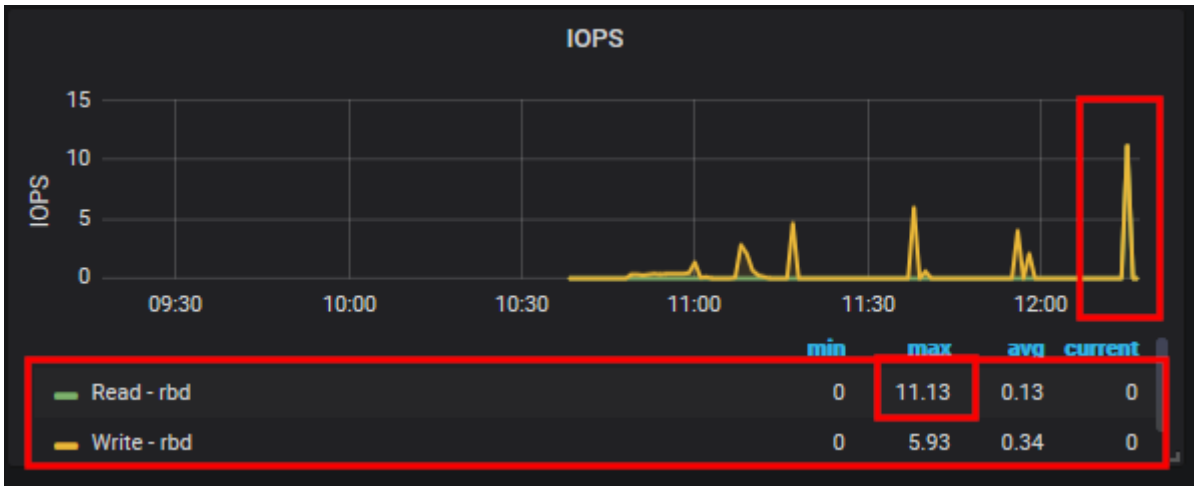
També podem veure les estadístiques capturades per Prometheus i mostrades a Grafana.

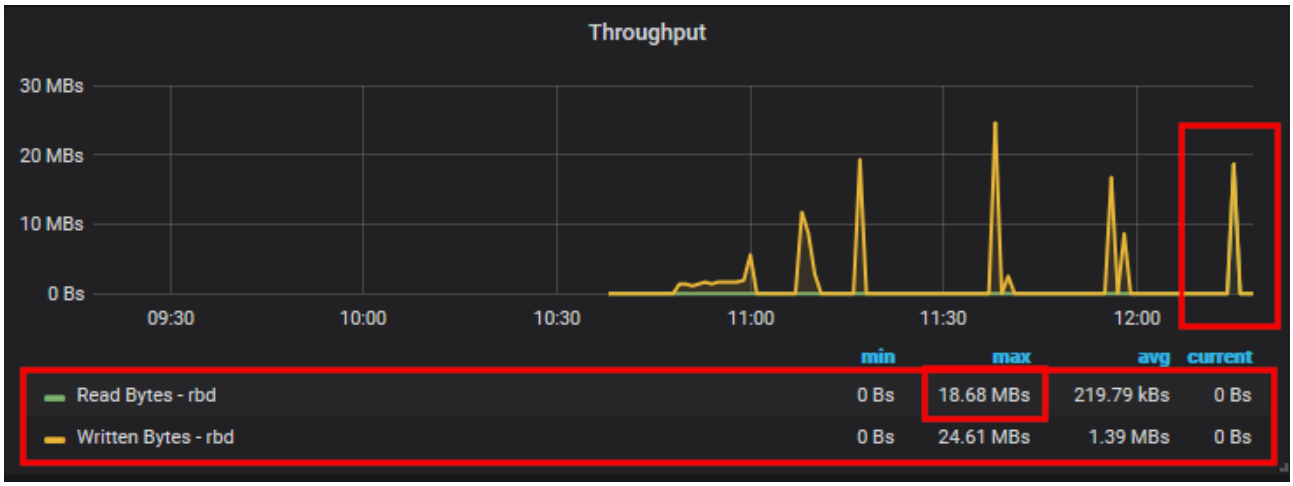


Com es pot veure, la tassa de transferència està al voltant dels 18MBps i es captura coma escriptura.
Ara també podem fer el pas invers copiant els fitxers creats al directori /root que estaria localitzat al disc local. Per tant, estaríem llegint aquestes dades des de la unitat rbd a Ceph.

```
# cp -ax /mnt/mydisk/prova1G /root/
```

Una vegada acabada, veiem que es mantenen les tasses de transferència obtingudes a l'escriptura però es doblen els IOPS ja que la lectura es menys costosa i l'escriptura en local és més ràpida.





També podem provar el rendiment del dispositiu de blocs que tenim muntat al servidor client, des de el node de desplegament mitjançant la comanda *rbd* de Ceph. Amb aquesta ordre podem llançar escriptures contra una imatge al pool indicat i comprovar la tassa de transferència i latències obtingudes.

Aquesta opció necessita d'un *pool* i una imatge per poder treballar que hauríem de crear. En el nostre cas, farem servir la imatge *disk01* del pool *rbd* que ja tenim creada i muntada al node client.

```
# rbd bench-write disk01 --pool=rbd
```

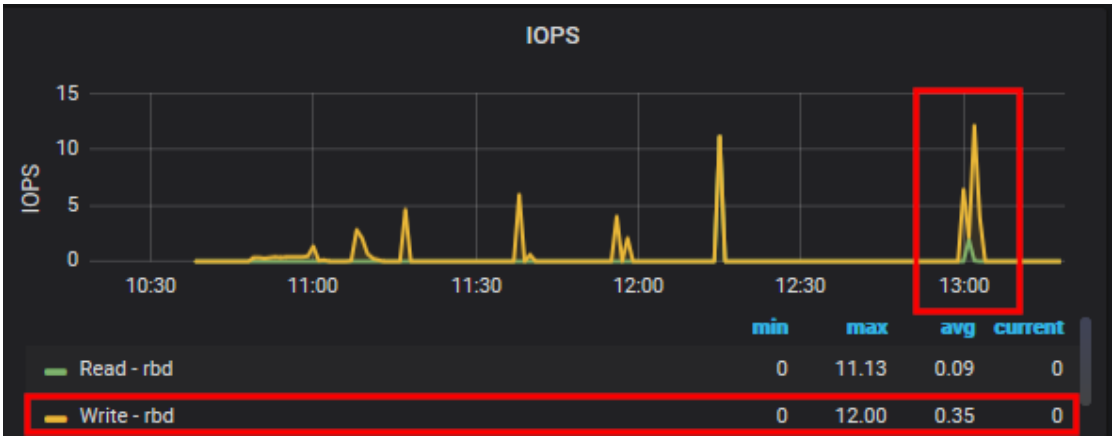
Amb el que obtenim les següents dades.

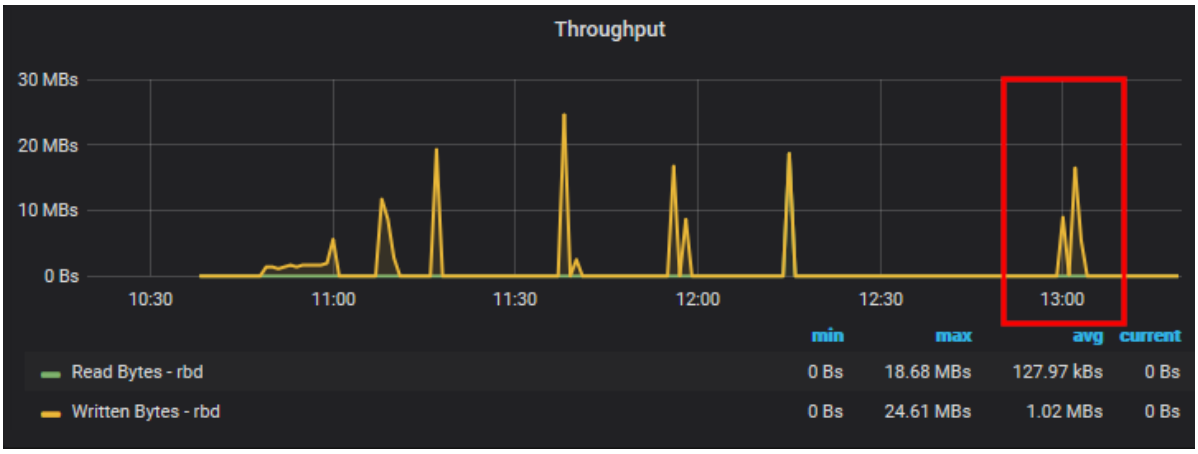
```
[cephuser@ceph-admin ~]$ rbd bench --io-type write disk01 --pool=rbd
bench  type write io_size 4096 io_threads 16 bytes 1073741824 pattern sequential
  SEC       OPS   OPS/SEC   BYTES/SEC
    1      6992   6586.38  26977803.75
    2      9136   4184.67  17140415.72
    3     10560   3348.91  13717135.90
    4     13760   3365.70  13785912.38
    5     15104   3008.91  12324515.34
    6     17040   1994.42  8169125.86
    7     18512   1886.11  7725524.26
    8     20608   2070.02  8478791.94
    9     22672   1769.99  7249863.46
   10     25824   2134.15  8741495.36
```

...

```
   80    214112   2904.20  11895601.90
   81    218848   3465.93  14196453.96
   82    221712   3607.09  14774627.63
   83    224080   3107.55  12728538.22
   84    228080   3432.12  14057963.77
   85    231488   3537.41  14489238.96
   86    234224   3087.51  12646436.09
   87    237696   3098.83  12692824.02
   88    242288   3488.74  14289896.19
   89    245440   3454.68  14150369.50
   90    249264   3404.67  13945512.36
   91    251904   3440.98  14094260.26
   92    254960   3447.24  14119887.41
   93    258304   3389.58  13883736.84
   94    259568   2841.47  11638679.05
elapsed:    95  ops:   262144  ops/sec:  2737.10  bytes/sec: 11211170.46
[cephuser@ceph-admin ~]$
```

Si mirem les gràfiques de Grafana veiem alguna variació, atès la diferencia d'unitat i la freqüència en la que Prometheus fa les consultes de dades.





Altra prova que podem realitzar per tal d'analitzar el rendiment de *Ceph* és fer servir una eina que ofereix *Rados* per fer benchmarking.

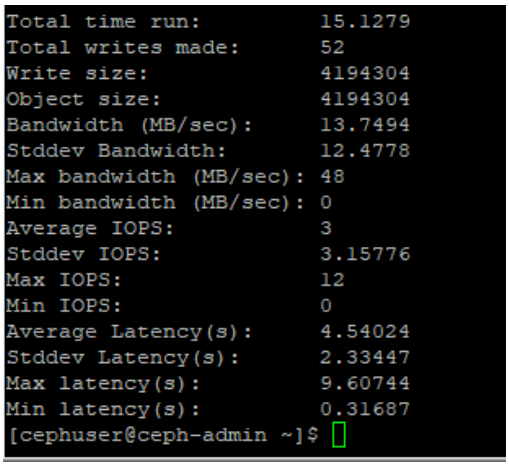
Per portar a terme la prova haurem de crear un pool especific per que *rados bench* treballi contra ell.

```
# ceph osd pool create clusterbench 100 100
# rados bench -p clusterbench 10 write --no-cleanup
```

Els paràmetres que l'indiquem configuren la sentencia per que realitzi les proves contra el pool que hem creat durant deu segons i sense netejar la sortida per tal d'obtenir més dades.

```
[cephuser@ceph-admin ~]$ ceph osd pool create clusterbench 100 100
pool 'clusterbench' created
[cephuser@ceph-admin ~]$ sudo rados bench -p clusterbench 10 write --no-cleanup
hints = 1
Maintaining 16 concurrent writes of 4194304 bytes to objects of size 4194304 for up to 10 seconds or 0 objects
Object prefix: benchmark_data_ceph-admin_6385
  sec Cur ops   started  finished  avg MB/s  cur MB/s last lat(s)  avg lat(s)
    0       0         0         0         0         0           -           0
    1      16        17         1    3.9984         4    0.672933    0.672933
    2      16        17         1   1.99945         0           -    0.672933
    3      16        20         4   5.33212         6     2.93069     2.25743
    4      16        25         9   8.99813        20     3.34589     2.84999
    5      16        29        13   10.3979        16     4.70593     3.34276
    6      16        41        25   16.6633        48     2.61274     3.33497
    7      16        43        27   15.4256         8    0.912958     3.22815
    8      16        45        29   14.4973         8     2.75159     3.19458
    9      16        46        30   13.3308         4     3.91803     3.21869
   10      16        51        35   13.9974        20     4.45227     3.38174
   11      16        52        36   13.0885         4     0.31687     3.29661
   12      16        52        36   11.9979         0           -     3.29661
   13      15        52        37   11.3826         2     6.74305     3.38975
   14      15        52        37   10.5696         0           -     3.38975
   15      11        52        41   10.9314         8     8.70621     3.77717
```

Una vegada acabada la prova s'ens mostrarà una estadística dels aspectes més rellevants, com ara el volum de dades creat, l'amplada de banda consumida, el nombre d'IOPS generats o la latència.



Podrem veure totes les opcions de *rados* al manual oferit per Ceph.

- http://docs.ceph.com/docs/master/man/8/rados/