Universidad Técnica Federico Santa María
Departamento de Informática
Valparaiso

Análisis de Arquitecturas Modernas de Data Center

Por

SIMÓN ERNESTO CÁRDENAS ZÁRATE

Memoria para optar al título de
Ingeniero Civil en Informática

Profesor Guía: Javier Cañas Robles
Profesor Correferente: Raúl Monge Anwandter

DICIEMBRE 2017

Resumen

Los data center son estructuras complejas que no solamente se reducen a una habitación llena de servidores, si no que dependen de una completa infraestructura física que incluye sistemas de monitoreos, sistemas de energía y climatización redundantes que constituyen el sistema nervioso de la infraestructura IT sobrepuesta. A su vez, dependiendo del tamaño y de los servicios que ofrece el data center, la presentación de los componentes IT puede variar en múltiples patrones de diseño dependiendo si se apunta a la computación de alto rendimiento, computación en la nube, housing de servidores a externos, almacenamiento para una pequeña organización, etc. Se presenta por lo tanto, un modelo de capas similar al modelo OSI que describe de forma general, todos los componentes que definen un data center, en todos sus niveles, desde su planeación y análisis de costos, hasta la ejecución de sus operaciones. Como caso de estudio, se analiza el data center del Departamento de Informática de la Universidad Técnica Federico Santa María.

Reconocimientos

En primer lugar agradezco profundamente al profesor Javier Cañas quien se preocupó de guiar este trabajo y orientarme en su desarrollo. Agradezco también al personal del Data Center del Departamento de Informática, Miguel Varas, Yonatan Dossow, Eduardo Toledo, y Pamela Parra, por entregarme la información necesaria para poder realizar esta memoria. Agradezco al profesor Claudio Torres, por guiarme en proyectos que me llevaron a conocer el interesante mundo de los data centers. Finalmente agradezco a Mario Larenas, Consultor de Infraestructura TI en Entel S.A. por su fuerte apoyo en la parte de eficiencia energética.

En segundo lugar, agradezco a mis compañeros de universidad, amigos y futuros colegas, quienes me apoyaron en todo el proceso universitario tanto en lo académico como lo personal. Quisiera mencionar en especial a mi grupo de estudios y grandes amigos, Joaquín Meza, Maximiliano Lyon, Matías Yañez, Pablo Pizarro, Sebastián Torres, Fernando Da Silva, Fabián Da Silva, Christian López, Ignacio Oneto, y en especial a mi grupo de la Feria de Software, Hernán Martínez, César Contreras, Carlos Díaz y Mauricio Avendaño. Sin su ánimo, apoyo, y alegrías, probablemente no hubiese llegado a esta etapa.

Finalmente agradezco a mi familia, a mi padre y a mi madre por su apoyo máximo, y su incalculable paciencia, y a mis hermanos, por estar siempre conmigo, aunque estemos a lo lejos.

Índice general

Resumen I

Reconocimientos II

Glosario VI

1. Introducción 1
- 1.1. Definición del problema . . . 2
- 1.2. Objetivos . . . 3
  - 1.2.1. Objetivo general . . . 3
  - 1.2.2. Objetivos específicos . . . 3
- 1.3. Metodología de la solución . . . 3
- 1.4. Estructura del documento . . . 4

2. Estado del Arte 5
- 2.1. Data Center Pulse Stack . . . 6
- 2.2. Open DCME - Data Center Measure of Efficiency . . . 8
- 2.3. Open Compute Project . . . 10
- 2.4. Infraestructuras del futuro . . . 11

3. Modelo de capas de Data Center 13
- 3.1. Estándares y certificaciones . . . 15
  - 3.1.1. Uptime Institute Tier Certification . . . 17
  - 3.1.2. SAS 70 - Statement of Auditing Standards no. 70 . . . 17
- 3.2. Planeación . . . 18
  - 3.2.1. Finanzas: Inversión e impuestos . . . 19
  - 3.2.2. Costo y disponibilidad de energía eléctrica . . . 19
  - 3.2.3. Clima y posibles desastres naturales . . . 20
  - 3.2.4. Enlaces y telecomunicaciones . . . 20
- 3.3. Infraestructura Física . . . 21
  - 3.3.1. Energía . . . 21
  - 3.3.2. Cableado estructurado . . . 24
  - 3.3.3. Enfriamiento . . . 26
  - 3.3.4. Racks . . . 29
- 3.4. Seguridad física de un Data Center . . . 30
  - 3.4.1. Factores Operacionales . . . 31
  - 3.4.2. Componentes Físicos . . . 31
- 3.5. Red . . . 32

3.5.1. Cisco 3-Tier Architecture - Arquitectura de redes de 3 niveles . . . 32
3.5.2. Fat Tree . . . 36
3.5.3. Leaf and Spine . . . 37
3.5.4. Encapsulación de frames . . . 39
3.5.5. SDN - Software Defined Networking . . . 40
3.6. Almacenamiento . . . 41
3.6.1. DAS - Direct Attached Storage . . . 42
3.6.2. NAS - Network-Attached Storage . . . 42
3.6.3. Sistemas RAID - Redundant Array of Independent Disks . . . 42
3.6.4. SAN - Storage Area Network . . . 44
3.6.5. SDS - Software Defined Storage . . . 46
3.7. Virtualización . . . 46
3.7.1. Virtualización basada en hipervisores . . . 47
3.7.2. Virtualización basada en containers . . . 48
3.8. Administración . . . 50
3.8.1. Configuration Management . . . 51
3.8.2. Nubes privadas . . . 52
3.8.3. SDDC - Software Defined Data Center . . . 53
3.9. Aplicaciones . . . 54
3.9.1. Continuous Integration - Integración Continua . . . 54
3.9.2. Continuous Deployment - Entrega Continua . . . 56
3.10. Resumen del modelo . . . 57

4. Caso de Estudio - Departamento de Informática 60
4.1. Estándares . . . 60
4.2. Planeación . . . 61
4.3. Seguridad . . . 62
4.4. Infraestructura . . . 62
4.5. Red . . . 64
4.6. Almacenamiento . . . 65
4.7. Virtualización . . . 66
4.8. Administración . . . 66
4.9. Aplicación . . . 66

5. Propuestas de Mejoramiento 67
5.1. Etapa 1 - Mejoras de Infraestructura (Eficiencia). . . . 67
5.1.1. De los estándares . . . 67
5.1.2. De consumo energético . . . 68
5.1.3. De la climatización . . . 68
5.1.4. Propuestas Etapa 1 . . . 68
5.2. Etapa 2 - Mejoras de IT y administración (Alta disponibilidad). . . . 69
5.2.1. De las redes . . . 69
5.2.2. Del Almacenamiento . . . 69
5.2.3. De la plataforma de virtualización . . . 69
5.2.4. De las aplicaciones y servicios . . . 70
5.2.5. Propuestas Etapa 2 . . . 70
5.3. Etapa 3 - Asistencia de motores de cómputo en la nube (Agilidad). . . . 71

5.3.1. Análisis de costos 72
5.3.2. EC2 versus On-Premise 74
5.3.3. Propuestas Etapa 3 79

6. Conclusiones **80**
6.1. Conclusiones generales 80
6.2. Conclusiones específicas 82
6.3. Trabajo futuro 84

Bibliografía **85**

Glosario

API	**A**pplication **P**rogramming **I**nterface
ASHRAE	**A**merican **S**ociety of **H**eating, **R**efrigerating and **A**ir Conditioning **E**ngineers
ATS	**A**utomatic **T**ransfer **S**witch
AWS	**A**mazon **W**eb **S**ervices
BMS	**B**uilding **M**anagement **S**ystem
CAPEX	**CAP**ital **EXP**enditures
CCTV	**C**losed **C**ircuit **T**ele**V**ision
CDN	**C**ontent **D**elivery **N**etwork
CD	**C**ontinuous **D**eployment
CI	**C**ontinuous **I**ntegration
CM	**C**onfiguration **M**anagement
CRAC	**C**omputer **R**oom **A**ir **C**onditioning
CRAH	**C**omputer **R**oom **A**ir **H**andler
DCP	**D**ata **C**enter **P**ulse
DC	**D**ata **C**enter
DNS	**D**omain **N**ame **S**ystem
DTI	**D**irección de **T**écnologias de la **I**nformación
EC2	**E**lastic **C**ompute **C**loud
EoR	**E**nd **O**f **R**ow
HIPAA	**H**ealth **I**nsurance **P**ortability and **A**ccountability **A**ct
HPC	**H**igh **P**erformance **C**omputing
ISP	**I**nternet **S**ervice **P**rovider
IT	**I**nformation **T**echnology
IoT	**I**nternet **O**f **T**hings
KPI	**K**ey **P**erformance **I**ndicator

LAN	**L**ocal **A**rea **N**etwork
NAS	**N**etwork **A**ttached **S**torage
NOC	**N**etwork **O**perating **C**enter
NPFA	**N**ational **F**ire **P**rotection **A**ssociation
OCP	**O**pen **C**ompute **P**roject
OPEX	**OP**erational **EXP**enditures
PUE	**P**ower **U**sage **E**ffectiveness
RAID	**R**edundant **A**rray of **I**ndependent **D**isks
SAN	**S**torage **A**rea **N**etwork
SAS	**S**tatement **A**uditing **S**tandards
SCADA	**S**upervisory **C**ontrol **A**nd **D**ata **A**cquisition
SDN	**S**oftware **D**efined **N**etworking
SOC	**S**ecurity **O**perating **C**enter
STP	**S**panning **T**ree **P**rotocol
TIA	**T**elecommunications **I**ndustry **A**ssociation
ToR	**T**op **O**f **R**ack
UPS	**U**ninterruptible **P**ower **S**upply
VLAN	**V**irtual **L**ocal **Area** **N**etwork
VxLAN	**V**irtual E**X**tensible **L**ocal **A**rea **N**etwork
VM	**V**irtual **M**achine
VPC	**V**irtual **P**rivate **C**loud

Capítulo 1

Introducción

Herramientas como Facebook, Twitter y Youtube son usadas diariamente para comunicarnos, compartir y mantenernos actualizados. Sin embargo, para poder hacer una búsqueda en Google y obtener como respuesta millones de resultados en una fracción de segundo, o ver el último documental de National Geographic en Netflix, una serie de procesos enorme ocurre por debajo. Una gran cantidad de datos deben ser procesados y almacenados en potentes servidores a lo largo de todo el mundo, los cuales están hospedados en enormes instalaciones conocidas como Data Centers. Estos DC son en realidad grandes habitaciones con múltiples servidores, dispositivos de red, sistemas de enfriamiento, sistemas eléctricos, habitaciones de monitoreo, entre otros. Los DC no solo permiten que podamos visitar perfiles de amigos en las redes sociales, o leer un reportaje en un portal de noticias. Más que eso, los data centers son un elemento esencial de la economía a nivel mundial, ya que prácticamente son los encargados de almacenar y procesar la información digital; por ejemplo datos personales, información bancaria, sistemas de universidades, páginas web, y un sin fin de elementos críticos que nos han permitido entregar y recibir información automáticamente.

Sin embargo, la explosión de la llamada Internet de las Cosas y el Big Data han provocado un gran stress en los data centers, los que se han visto en la urgencia de hacer revisiones a sus tecnologías con tal de proveer un servicio confiable, altamente disponible, rápido y de bajo costo. Un ejemplo claro del poder de estos dispositivos inteligentes fue el último ataque a Dyn en 2016, usando principalmente cámaras de seguridad conectadas

a Internet para generar uno de los ataques de denegación de servicio más grandes de la historia, registrando un tráfico de 1.2 Tbps en contra de su infraestructura DNS [34].

Esta memoria busca generar un modelo conceptual que permita tener una visión general de los DC, y todo lo que significa su implementación y/o modernización, con tal de servir como ayuda a organizaciones y/o empresas en la toma de decisiones. Se describen por lo tanto, arquitecturas que definen un DC en todos sus niveles, desde la elección de la ubicación física, hasta los servicios ofrecidos. Sin embargo, es necesario aclarar que los DC representan sistemas tan complejos, que es casi imposible abarcar todos sus elementos en profundidad. Es por esto que se trata de abarcar todos los niveles de una forma más bien general, de forma que el lector tenga una idea clara de todos los elementos que conforman un DC y como estos elementos se relacionan entre sí.

1.1. Definición del problema

Los DC son grandes y complejas estructuras, en ocasiones muy protegidos del público, y con políticas de seguridad bastante reservadas, es por esto que un análisis a su infraestructura podría resultar difícil de sobrellevar. Si bien existen estándares, y organizaciones como la OCP [30] encargadas de distribuir material para una implementación eficiente, otras empresas deciden resolver sus problemas de IT de forma interna e independiente. Mientras que algunas pequeñas empresas deciden dar el salto de implementar data centers propios con tal de manejar sus datos (de acuerdo a políticas internas o leyes exigidas por el gobierno), otras deciden rentar centros de colocación, sin embargo ambos proyectos pueden presentar un costo elevado en su implementación. Esta desinformación puede llevar a que pequeñas empresas que no manejen muy bien el área de IT, puedan llevarse sorpresas en su presupuesto, debido a que no conocen los desafíos que deben enfrentar los DC en términos de infraestructura física e IT.

Se hace necesario por lo tanto revisar los componentes principales que componen un DC, con tal de facilitar a las pequeñas empresas y organizaciones, una visión general de los costos que podrían generarse en términos de infraestructura y administración.

1.2. Objetivos

1.2.1. Objetivo general

Crear un marco conceptual, mediante un modelo capas, que permita establecer las estructuras que componen un data center, con el fin de servir como ayuda para su comprensión y evaluación. Junto con un estado del arte de las arquitecturas modernas, se busca analizar las últimas tecnologías para diseñar data centers eficientes, de bajo costo, y fácil administración.

1.2.2. Objetivos específicos

- Analizar los problemas y desafíos que están enfrentando los DC en la actualidad.
- Buscar en la literatura actual, las últimas arquitecturas de DC, que describan los diseños de cada capa en particular (principalmente estructura física y TI), que se enfoquen en un diseño costo-eficiente y escalable, y analizar las ventajas y desventajas de unas sobre otras.
- Usar el DC del Departamento de Informática cómo caso de estudio, y ver como encaja con el modelo creado. Luego de esto, se propondrán mejoras a su infraestructura.

1.3. Metodología de la solución

Crear un modelo de capas que defina todos los elementos que componen un DC, es un trabajo de investigación que requiere hacer una revisión a la literatura actual y al estado del arte, con tal de ver cómo han ido evolucionando las tecnologías en términos de eficiencia, escalabilidad y confiabilidad. Por otro lado, usar el modelo y aterrizarlo en un DC como caso de estudio requiere manejar información que podría ser sensible y en algunos casos podrían necesitarse acuerdos de no divulgación de información. Se usará por tanto una metodología de análisis teórico, compuesta por las siguientes etapas:

- Revisión de estado del arte, y últimas tecnologías relacionadas con cada capa del modelo de DC.

- Uso del DC del Departamento de Informática de la UTFSM como caso práctico, con tal de comprender a grandes rasgos sus elementos, y proponer mejoras que sean estado del arte.

1.4. Estructura del documento

En el Capítulo 1, se introduce al lector a los problemas que surgen en los DC, y se describen los objetivos de esta memoria.

En el Capítulo 2, Estado del arte, se describe una pequeña historia de la evolución de los DC hasta hoy, y se describen brevemente algunos estudios que han intentado hacer una descripción general de un DC.

En el Capítulo 3, Modelo de capas de Data Center, se describe de forma detallada, todos los elementos que componen un DC y sus arquitecturas modernas, con el fin de tener una comprensión general de todos sus niveles.

En el Capítulo 4, Caso de estudio, se localiza el modelo de capas al Data Center del Departamento de Informática, con el fin de proponer mejoras y actualizaciones.

En el Capítulo 5, Propuestas de Mejoramiento, se presentan posibles mejoras que podrían ser adoptadas por el Departamento, las cuales podrían servir como trabajo futuro al lector.

En el Capítulo 6, Conclusiones, se cierra este trabajo con el aprendizaje logrado.

Capítulo 2

Estado del Arte

Después de la década del 70, ocurren importantes avances en la fabricación de computadores, dentro de las cuales se encuentran el desarrollo de computadores personales en serie, el surgimiento del protocolo Ethernet, y NFS (Network filesystem): un sistema de archivos montado sobre una red local permitiendo el almacenamiento de datos a través de una arquitectura cliente-servidor. Sin embargo uno de los inventos más revolucionarios es logrado por Tim Berners-Lee en CERN al permitir compartir documentos a través de internet en una tecnología que hoy se conoce como la World Wide Web, surgiendo entonces la necesidad de usar el poder de cómputo de múltiples computadores trabajando como uno, por lo que al final de los años 90 surgen los primeros Data Centers. Un hito histórico en la industria lo consigue VMWare en 2001 al presentar ESX, el hipervisor bare-metal más popular a nivel mundial, disminuyendo los costos de HW en los data centers de forma drástica al optimizar el uso de servidores mediante virtualización de sistemas operativos sobre un software liviano. Sin embargo, el elevado costo de mantener equipamiento IT, y la ventaja competitiva que tenía Amazon durante 2006 como empresa de comercio electrónico, fue lo que los llevó a lanzar AWS, la plataforma Cloud líder a nivel mundial [43], la cual actualmente cuenta con más de 42 data centers alrededor del mundo para proveer servicios cloud y 55 Edge Locations[1] usadas para su red de distribución de contenido (CDN).

[1] Un Edge Location es un nodo de una red de distribución de contenidos de AWS usada para almacenar información de forma temporal con el fin de disminuir latencia a los usuarios finales, y aliviar la carga de servidores centrales.

Los Data Centers han evolucionado a una velocidad sorprendente en los últimos años, debido a las altas demandas del poder de cómputo, y a la necesidad de optimizar los procesos energéticos, por lo que han surgido diferentes acercamientos a la hora de diseñar, construir, y optimizar data centers.

Hasta ahora existen diversas publicaciones y modelos que buscan describir data centers en su totalidad, intentando desligarse de proveedores o tecnologías en particular, con tal de proveer un lenguaje común y estandarizado que permita comprender los elementos y procesos que componen un data center. A continuación se describen los modelos más populares.

2.1. Data Center Pulse Stack

DCP es una organización sin fines de lucro que surge en 2008 con el fin de influenciar a la industria de los Data Centers a partir del conocimiento y colaboración de sus usuarios directos. A comienzos del 2009, y a partir de la explosión del Cloud Computing, DCP llegó a la conclusión de que se hacía necesario un modelo general y común para la comunidad que permita entender, comunicar, desarrollar e innovar data centers, el cual sería mantenido a través de la comunidad y finalmente validado por la industria. Este modelo, similar al modelo OSI, es un modelo compuesto por 6 capas en su versión final, donde cada capa define las diferentes operaciones que componen un DC. Principalmente describe al data center como un conjunto compuesto de

- Bienes raíces: Donde se incluye la ubicación geográfica, implementación modularizada (por ejemplo en containers) y eficiencia de recursos.
- Área física: Que describe la distribución del equipamiento IT.
- Área de plataforma: Donde se describen las arquitecturas de red, almacenamiento y sistemas distribuidos.
- Área de servicios de infraestructura: Donde se describen la optimización del uso de servidores a través del uso de la virtualización, y administración de recursos IT virtuales.

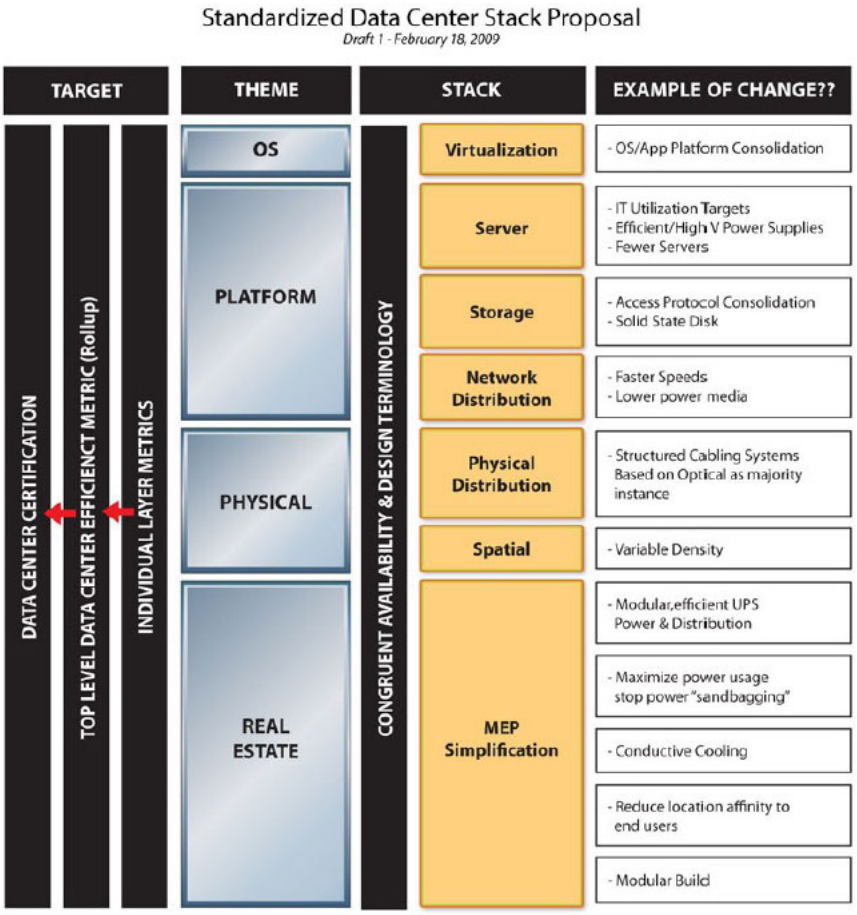

Figura 1, Primer draft DCP Stack.

Fuente: Data Center Pulse.

En este modelo principalmente cada capa tiene diferentes métricas asociadas, las cuales permiten medir los comportamientos de cada bloque de forma estandarizada, por ejemplo eficiencia energética, uso de recursos naturales y eficiencia operacional entre otros. Este modelo no solo cubre las necesidades del cloud computing, si no que también intenta abarcar diferentes mercados que usan recursos computacionales, como lo son los sistemas de control integrado (BMS, SCADA) o data centers modularizados mediante containers. Además este modelo incluye un bloque de validación para cada capa, con el fin de que el propietario del DC pueda certificar que sus instalaciones y operaciones satisfacen las necesidades de sus clientes, por ejemplo a través del sistema de clasificación de Tiers del Uptime Institute, o seguridad de la información a través de SAS 70.

La última versión de este modelo data del 2012, sin embargo este año la organización encargada se disuelve, y el desarrollo del stack pierde su continuación.

2.2. Open DCME - Data Center Measure of Efficiency

Este modelo, diseñado por la empresa de IT holandesa Mansystems, presenta 16 indicadores que permiten medir la eficiencia de un DC, y a la vez separarlo en 4 cuadrantes: Instalación física (Facility), dispositivos IT (IT Assets), administración (Management) y gestión de procesos (Processes). Cada KPI y la forma en que se mide puede ir variando de acuerdo a las necesidades de un DC, y los indicadores deben ser definidos por especialistas en cada una de estas áreas, quienes definirán la forma de medición, y la determinación de *targets*[2].

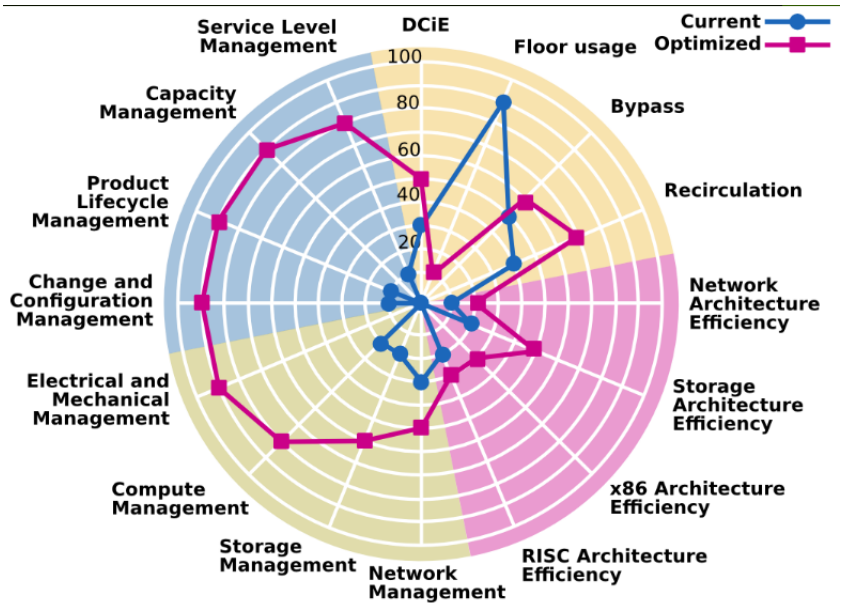

Figura 2, Modelo KPIs de Data Centers.

Fuente: Open DCME.

El cuadrante de **Facility** define 4 KPIs orientados a la optimización del espacio utilizado por los componentes físicos y a la eficiencia energética. El DCIE (Data Center Infrastructure Efficiency), desarrollado por The Green Grid [39] es una métrica que mide eficiencia energética en un DC y se define como:

$$DCiE = \frac{ITPower}{TotalFacilityPower} * 100 \tag{2.1}$$

[2] *Target* se refiere a la meta que se busca lograr con un KPI con tal de diferenciar un resultado positivo de uno negativo

Un indicador similar, también desarrollado por The Green Grid es el PUE (Power Usage Effectiveness) y está relacionado con el DCiE de la forma: $PUE = \frac{1}{DCiE}$, y sirve para medir la cantidad de energía que se está desperdiciando dentro del DC (debido a que cada transformación de voltaje, y distribución provoca pérdidas de energía)

Otros indicadores en el cuadrante de Facility son el uso de suelo, el cual mide la optimización del espacio físico con tal de reducir espacio utilizado (lo que afecta directamente a la optimización del enfriamiento, pues se deben enfriar menos sectores del DC), y Bypass y recirculación, que miden el aire frío desperdiciado, o la mezcla en el interior del recinto con el aire caliente. Estos KPIs deben ser manejados y medidos principalmente por especialistas en eficiencia energética.

El segundo cuadrante, **IT Assets**, busca optimizar el uso de los dispositivos IT, por ejemplo maximizando el uso de discos de almacenamiento, o sacarle el máximo provecho a los servidores usando virtualización. En este cuadrante también se definen 4 KPIs, para medir eficiencia de la red, del almacenamiento y eficiencia de la arquitectura usada (RISC o x86). Medir la eficiencia de estos 4 componentes, depende de varios factores, como por ejemplo el diseño o la arquitectura que se esté siguiendo (3-Tier, SAN, etc), o de la identificación de lo que se busca optimizar (minimizar costos al no tener exceso de discos, maximizar *throughput*, etc). Para medir estos KPIs se comparan los KPIs resultantes del DC versus KPIs de referencia que son definidos por los especialistas IT.

El tercer cuadrante, **Management**, busca optimizar la forma de administrar las tecnologías IT. Los KPIs son principalmente la administración de la red, almacenamiento, computación y energía usando información histórica, y busca medir la forma en la que se moverán los KPIs del segundo cuadrante. Este cuadrante se diferencia del anterior, en que el cuadrante 2 busca eficiencia y maximización del uso de recursos y este cuadrante busca tomar decisiones para no generar bajo/sobre abastecimiento.

El último cuadrante, **Processes**, busca medir los procesos que permiten la existencia de un DC. Los procesos End-to-End son procesos que definen los ciclos de vida de lo que se desea realizar, y en esta etapa se busca medir la eficiencia de esos procesos. Estos KPIs no están dirigidos al funcionamiento del DC en si, sino que están enfocados en proveer un servicio de calidad como lo son acuerdos de nivel de servicio, ciclo de vida de los dispositivos físicos con tal de proveer confiabilidad y seguridad a los clientes antes fallas de HW.

En conclusión, este modelo permite medir la eficiencia de los elementos que componen un DC arbitrario a través de la inclusión de KPIs con el fin de optimizar recursos de acuerdo a las diferentes necesidades. El modelo no define *targets* ni formas de medir, solo recomienda qué medir y con qué fin.

2.3. Open Compute Project

La fundación de el Open Compute Project es sin duda uno de los hitos más importantes de la historia de la informática desde la expansión del software libre. Este proyecto, iniciado principalmente por Facebook en 2009, que actualmente cuenta con el apoyo de Intel, Google, Microsoft, Cisco, y otras grandes empresas, busca diseñar los Data Center más eficientes del mundo, con la premisa de rediseñar todo el HW que compone un DC y liberar estos diseños y especificaciones a la comunidad. Esta iniciativa surge de la necesidad de escalar masivamente la infraestructura de Facebook a un bajo costo, además de minimizar el impacto ambiental, siendo un problema que aún se mantiene vigente y que comparten múltiples empresas de la industria. Se busca también promover la estandarización de las infraestructuras, con el fin de que todos los DC puedan tener un lenguaje común de comunicación independiente de proveedores o arquitecturas. El primer Data Center orientado completamente a OCP es Facebook Prineville [44], el cual es un 38 % más eficiente y 24 % más barato de construir que el resto de sus DC, llegando a alcanzar un extraordinario PUE de 1.07 (es decir, casi el 100 % de la energía suministrada por la red eléctrica utilitaria es usada para energizar los servidores), comparado con un 1.9 PUE en promedio de los data centers comunes. Facebook Prineville destaca por usar 100 % Free Cooling, mediante aire obtenido desde el exterior, eliminando completamente el uso de torres de enfriamiento, *chillers*[3] y unidades de manejo de aire, en los cuales se desperdicia una elevada cantidad de energía mediante el uso de motores mecánicos.

OCP tiene una variedad de proyectos liberada a la comunidad, dentro de los más importantes se encuentran:

- Racks: Open Rack v1 y Open Rack v2, especificaciones de Racks que simplifican el suministro de energía a los servidores.

[3] Máquinas mecánicas (que trabajan en conjunto con las torres de enfriamiento) usadas para enfriar agua a través de condensadores, refrigerantes y anticongelantes.

- Almacenamiento: OpenVault y Bryce Canyon (sistema de almacenamiento de 4U soportando hasta 72 discos), compatibles con Open Rack.
- Servidores: Leopard (Servidor 3-en-1 orientado a optimizar uso de energía) y Open CloudServer, servidor blade orientado a la optimización de consumo energético.
- Networking: Wedge es un switch Top-Of-Rack orientado completamente a SDN, y FBOSS, el controlador SDN para Wedge.

Es de vital importancia que los DC tengan una disponibilidad lo más cercana al 100 %, lo que en estos momentos implica un nivel de redundancia y duplicación excesivo (lo que conlleva a gastos extra, y pobre eficiencia). Sin embargo, los data centers siguen fallando, a pesar de tener múltiples certificaciones de "alta disponibilidad" como lo demostró últimamente un corte de más de 5 horas que tuvo la plataforma AWS, específicamente el servicio de almacenamiento S3, debido a un simple error en la ejecución de un comando en los interiores de US-EAST-1 [40]. Es por esto que OCP hace un cambio radical en sus diseños y cambia su perspectiva de "altamente redundante" a "altamente resiliente", a través de rutinas estándares encargadas de encontrar fallas y resolverlas en el menor tiempo posible.

2.4. Infraestructuras del futuro

La OCP es actualmente una de las iniciativas más importantes en la industria de DC, sin embargo para poder innovar, es necesario correr riesgos. Ejemplos de infraestructuras del futuro, son por ejemplo los DC de Facebook. Facebook Altoona recibe el 100 % de su energía a través de fuentes de energía eólicas. Sin embargo, a pesar de los esfuerzos de OCP de unificar las tecnologías de levantamiento de data centers, otras empresas han optado por innovar de otras maneras.

A mediados de 2016, IBM liberó un artículo en el que demuestra satisfactoriamente que la posibilidad de almacenar datos en 1 átomo es posible, lo que llevaría a crear sistemas de almacenamiento órdenes de magnitud más pequeños que los actuales. Aunque este proyecto aún está lejos de verse en producción en data centers, en el experimento pudo almacenarse satisfactoriamente 2 bits de información en un arreglo de 2 átomos de

Holmio (a través de interruptores magnéticos, simulando un disco duro) a través de un microscopio de efecto túnel [41].

Por otro lado, una propuesta interesante sobre eficiencia general la propone Microsoft con su proyecto Natick en 2015, donde se plantea la posibilidad de desplegar data centers en el fondo submarino, con tal de usar la energía de las olas y aprovechar el agua como método de enfriamiento. Ventajas de esto incluyen espacio semi-ilimitado, debido a que el océano provee un entorno más uniforme que la tierra y más rápidos de construir, ya que se podrían llegar a construir data centers en serie. La propuesta de Microsoft consiste en construir un DC en un tiempo máximo de 90 días, y despacharlo a cualquier lugar del mundo donde sea necesario. Si bien este proyecto presenta varios desafíos, como por ejemplo la difícil reposición de dispositivos defectuosos, manejar la vida marina que podría generarse alrededor del DC, o mantener el interior seco libre de humedad, también presenta múltiples ventajas, como por ejemplo manejar el DC en un 100 % con luces apagadas, y aprovechar la energía proveniente de las olas [42].

Finalmente, una tendencia que se ha hecho notar con la explosión del cloud computing, son los data centers de Hiper-Escala, a los cuales su diseño les permite escalar de forma fácil, eficiente, y a bajo costo a través de infraestructuras definidas por software, donde la hiper-convergencia es un elemento clave. Se prevé que en el futuro, los data centers basados en hiper-convergencia[4] serán el nuevo estado del arte debido a la facilidad de administración y escalabilidad que éstos presentan.

[4]Tipo de infraestructura de muy bajo costo que integra recursos de redes, almacenamiento y virtualización basada en una arquitectura de definición por software

Capítulo 3

Modelo de capas de Data Center

El diseño de un DC es tan complejo que requiere el uso de múltiples disciplinas de la ingeniería para una implementación correcta, eficiente, y altamente disponible. El manejo de alimentación energética, enfriamiento, diseño de redes, almacenamiento, virtualización y aplicaciones distribuidas requieren un sistema firmemente enlazado, donde una falla en un área podría dejar completamente inutilizable a otra. En esta sección se describe desde una perspectiva general, un marco conceptual que define todos los niveles de todos los componentes que conforman los DC actuales, a partir de un modelo organizado por *capas* similar a el modelo OSI. Este modelo está compuesto principalmente por 3 grandes capas, la capa física (Facility), la capa de IT, y la capa de operaciones, además de capas transversales, las cuales consideran estándares, certificaciones, recomendaciones, normas, y buenas prácticas en el diseño de todas las capas superiores[1], y monitoreo de los componentes de cada capa en particular, las cuales se profundizan a lo largo del escrito.

Para permitir una mejor comprensión, se describirán las capas desde una perspectiva $bottom - up$, es decir, desde las capas inferiores a las superiores. Esto es porque es más fácil comprender las capas superiores, si se tiene una buena comprensión (o al menos una comprensión general y/o conceptual) de las capas subyacentes. En la figura 3 se puede observar una descripción muy generalizada de este modelo de capas. Se explicarán

[1] Aunque no existen estándares para desarrollar modelos de capas, el modelo presentado en esta sección se basa en el modelo OSI de redes, el cual describe de forma organizada los elementos y protocolos que componen la comunicación entre sistemas computacionales. Es importante definir la diferencia entre un modelo de capas y una arquitectura como la de Cisco, ya que un modelo de capas resume los componentes de un sistema (en este caso, los componentes de un data center), mientras que una arquitectura es un diseño particular de uno de esos componentes.

a grandes rasgos las 3 grandes capas, para posteriormente detallar cada subcapa en particular.

		Capa	Subcapa	Función
Estándares	Monitoreo y control	Facility	Planeación	Estimar costos totales y gastos operacionales y de capital en la implementación del Data Center. Generar análisis de escalabilidad.
			Seguridad	Ofrecer sistemas de seguridad física de los datos, del recinto, y seguridad en las operaciones.
			Infraestructura	Ofrecer un ambiente adecuado al equipamiento IT.
		IT	Red	Distribuir y balancear la carga entre servidores, mantener una comunicación fluida entre servidor-servidor y servidor-exterior
			Almacenamiento	Ofrecer alta disponibilidad, y tolerancia a fallas. Las arquitecturas de almacenamiento deben ser escalables y manejar un gran volumen de datos.
			Virtualización	Optimizar los recursos dedicados a computo.
		Operaciones	Administración	Manejar una gran cantidad de recursos distribuidos de forma fácil sin impactar las operaciones ni el rendimiento de los servicios.
			Aplicación	Automatizar el proceso de desarrollo, pruebas, modificación y puesta en producción de software.

Figura 3, Modelo de capas de DC según funciones.
Fuente: Elaboración propia.

Dentro de los objetivos de la capa Facility está la reducción de costos de activos fijos, optimización y uso eficiente de recursos, y aumento de ingresos. Una mala planificación, o administración de la infraestructura física podría provocar cortes inesperados en los servicios, afectando directamente al estado financiero de la empresa (por ejemplo en una empresa proveedora de servicios de Cloud, o en un sistema bancario o de intercambio de divisas, un par de minutos abajo se traduce en millones de pesos en pérdidas). La capa física por tanto, presenta la base de esta estructura de capas, ya que son los cimientos de un DC.

La capa de IT es la capa inmediatamente superior. Los componentes principales de la capa de IT son el manejo de redes, servidores, almacenamiento y seguridad, y es donde ocurre el procesamiento. Existen múltiples diseños para estas infraestructuras con tal de

maximizar el *uptime*, enfocadas en escalabilidad y seguridad, y serán cubiertas en esta sección.

Finalmente la capa de Operaciones, es la capa que incluye aplicaciones y servicios ofrecidos. En esta capa se detallan orquestadores, sistemas de configuración en masa y asignación óptima de recursos. Esta capa opera directamente *sobre* la capa de IT, y es la capa final de este modelo de infraestructuras de DC, y da paso al posterior modelo de infraestructuras de la computación en nube.

Si bien las capas tienen cierto sentido de independencia entre ellas (ya que están orientadas a áreas distintas), las capas subyacentes dan soporte, y permiten la existencia a las capas superiores, estableciendo que todas las capas estén relacionadas de alguna manera.

3.1. Estándares y certificaciones

Una de las capa transversales de este modelo de capas de DC es la que define las arquitecturas de las capas superiores basándose en estándares mínimos, recomendaciones y certificaciones. Cuando se está diseñando un DC, se debe cumplir con un conjunto de normas y estándares mínimos para su funcionamiento, conocidos como estándares regulatorios, los cuales se basan principalmente en la prevención de riesgos, protección a trabajadores, y protección de los datos. La NPFA - National Fire Protection Association es una organización encargada de desarrollar estándares y normas para la detección y extinción de incendios (Norma NPFA 75, NPFA 76). También están encargados de mantener el NEC - National Electric Code (tambien conocido como NPFA 70), enfocado en la correcta instalación del cableado y sistemas eléctricos. La ASHRAE por otro lado es una organización encargada de desarrollar estándares mínimos a cumplir para salas de aire acondicionado, temperaturas y humedad. Entes reguladores en Chile son la SEC quienes regulan sistemas eléctricos y la CONAMA encargada de realizar evaluaciones medioambientales.

Las normas en estos casos no son opcionales, y deben respetarse de forma obligatoria, de lo contrario existen riesgos de accidentes laborales, elevado impacto ambiental, multas ante una auditoría, o eventual pérdida de datos impactando directamente a los servicios.

Por otro lado existen otro tipo de estándares, los cuales actúan como buenas prácticas y recomendaciones para el diseño y construcción de data centers, y sirven para fomentar

la universabilidad, y compatibilidad entre componentes. Los estándares más reconocidos por la comunidad son ANSI / TIA 942 y BICSI 002-2014 [46].

El estándar TIA 942 es uno de los estándares más importantes dentro de los diseños de los data center, ya que describe especificaciones para variados elementos dentro del diseño de la infraestructura física, como lo son el diseño de redes, sistemas redundantes de energía y enfriamiento, sistemas de seguridad, protección contra desastres naturales etc. El estándar es mundialmente reconocido debido a que establece una topología de DC modificable, la cual puede ser aplicable a cualquier DC de cualquier tamaño en cualquier lugar. Este estándar también define los 4 niveles de Tiers (profundizados en la siguiente subsección), los cuales fueron después adoptados en las certificaciones del aclamado Uptime Institute para clasificar Data Centers de acuerdo a su disponibilidad. [2]

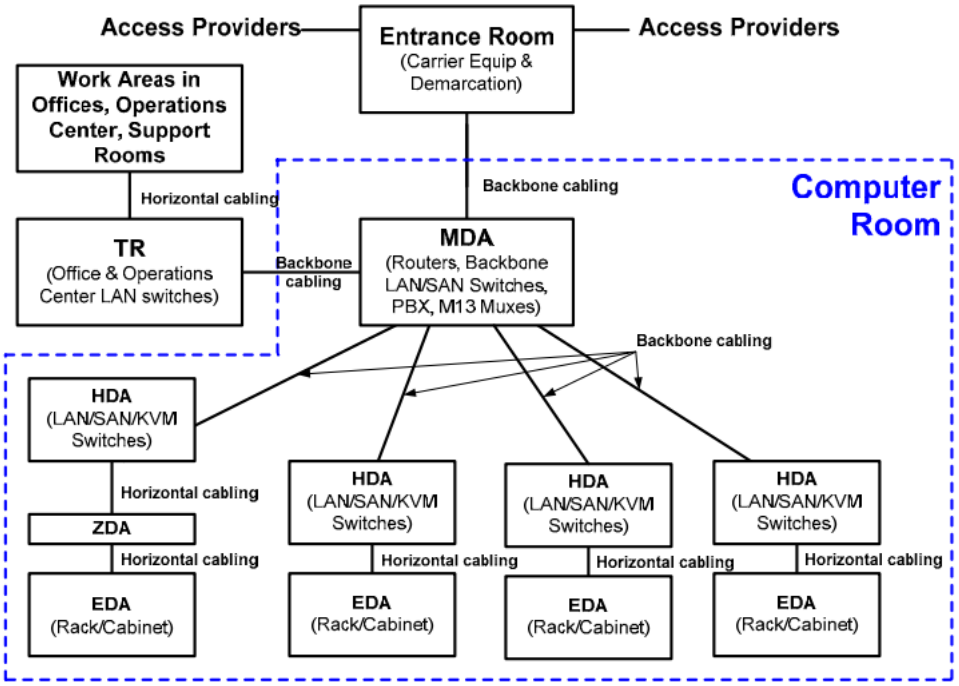

Figura 4, Topología básica de un Data center.

Fuente: TIA-942-A 2015.

Sin embargo, estas especificaciones no dejan de ser recomendaciones o *guidelines* los cuales podrían eventualmente no respetarse y seguir una infraestructura independiente (con los riesgos que estos conlleva). Por otro lado existen las certificaciones, las cuales son

[2]El estándar de Tiers de Uptime Institute fue publicado en 1990 mientras que el estándar TIA 942 fue publicado el año 2005. Uptime está mas enfocada a metas, por lo tanto es más flexible. TIA es más rígida ya que presenta requerimientos técnicos específicos que deben seguirse [47].

confirmaciones que son las que aseguran que los estándares se están cumpliendo, cuyo objetivo es dar credibilidad, confianza, y garantizar que los procesos se están ejecutando de forma acorde.

3.1.1. Uptime Institute Tier Certification

Uptime Institute, al igual que el estándar TIA 942 también posee un sistema de clasificación por Tiers, para evaluar el rendimiento de la infraestructura en cuanto a disponibilidad. Actualmente, Uptime Institute ofrece 3 diferentes tipos de certificaciones: Design (reconoce excelencia de planeación del diseño de infraestructura), Facility (reconoce el cumplimiento de objetivos de la instalación física) y Operations (reconoce excelencia operacional y mitigación de errores operacionales), todas clasificadas en un sistema de 4 Tiers, donde las primeras 2 están enfocadas a la topología e infraestructura física, y la última a la administración y mantenimiento; la cual, además, está separada en las categorías Gold, Silver y Bronze [25].

Si bien tanto el estándar TIA 942 como el sistema de clasificación de Tiers de Uptime Institute catalogan la infraestructura con Tiers de I a IV, ambas clasificaciones presentan diversas diferencias. Es por eso que entre ellas han llegado a un acuerdo con tal de diferenciar claramente sus clasificaciones, en donde TIA 942 eliminará la palabra Tier de su sistema de clasificación [23].

3.1.2. SAS 70 - Statement of Auditing Standards no. 70

Es una certificación orientada al control interno, y reportes financieros, sin embargo en los DC se enfoca en controles de seguridad de redes y seguridad física del recinto para asegurar manejo de los procesos a los clientes. Una certificación más actualizada que cumple con roles similares a SAS 70 es SSAE 16. Estas certificaciones no están orientadas al DC en si, sino que más bien apuntan a los consumidores, o clientes del DC, de esta forma aseguran que sus datos no serán robados, alterados o eliminados de ninguna forma [26].

Existen diversas certificaciones orientadas a diferentes áreas de los DC, dentro de las que también se encuentra, ICREA, ISO 9001, ISO 27001, entre otras [48]. Un artículo que

detalla en más profundidad una mayor cantidad de certificaciones se puede encontrar en [24].

Los estándares y certificaciones presentados en esta sección se enfocan netamente en la arquitectura física, eficiencia energética y optimización de espacio. Para estándares relacionados directamente con el área de TI, Open Compute Project es la organización más completa de desarrollo de este tipo de estándares, sin embargo aún no han trabajado en programas de certificación como tal.

3.2. Planeación

La primera capa de esta arquitectura, constituye lo que se refiere a la localización de un DC, la cual considera la ubicación geográfica, y los costos asociados al sitio donde será implementado el DC.

Esta capa es de principal importancia ya que la ubicación física determina los costos a nivel macro de la instalación en general, constituyendo un elemento principal a la hora de evaluar un proyecto, y hacer análisis de inversión. Distintos continentes, países, regiones o localidades presentan diferentes elementos que podrían afectar la eficiencia/eficacia del levantamiento de un DC ya sea leyes, impuestos, climas, posibilidad de desastres naturales, costo de energía, anchos de banda, etc. Ejemplos de estos países por ejemplo son Islandia, el cual tiene costos de energía relativamente bajos, y el enfriamiento es prácticamente gratis [2]. China por otro lado tiene una demanda extremadamente alta, y enlaces de telecomunicaciones bastante rápidos y confiables. Nueva Zelanda y Holanda, por otro lado son países que obtienen un alto porcentaje de su energía eléctrica a partir de fuentes renovables, provocando una buena imagen (en términos energéticos) a la empresa u organización que planea instalar un DC. Esto provoca que la elección de la ubicación de un DC no sea trivial. En [1] se presenta una lista de los países más convenientes para ubicar un DC, tomando en cuenta los elementos mencionados.

Se destacan los elementos con mayor importancia a considerar en la etapa de planeación de DC:

3.2.1. Finanzas: Inversión e impuestos

Las inversiones en bienes de capitales (CAPEX) de un DC son extremadamente altas superando los 500.000 USD para un DC relativamente pequeño [9], por lo que es necesario realizar una evaluación al proyecto. Implementar un DC requiere construir el sitio completamente, invertir en servidores, racks, sistemas UPS, routers, switches, mantenimiento, personal, etc. Esto sumado a que dependiendo del país/estado/región, pueden haber distintas variaciones de impuestos, provocando un aumento en el costo final. A este tipo de Data Center, del cual se es propietario, y se mantiene un control total, se les conoce como Data Center **On-Premise**. Si bien esta opción puede ser quizás inalcanzable para una pequeña o mediana empresa, una alternativa llamativa es externalizar los servidores a un DC ya establecido y arrendar un espacio dentro de él para colocar dispositivos IT propios. Estos DC se conocen como **centros de colocación**, los cuales se encargan de proveer seguridad, enfriamiento, espacio, energía eléctrica, etc a un costo accesible. Dependiendo del centro de colocación, es posible que se pague solo por el espacio, y/o se comparta el consumo energético, enfriamiento, etc, con otras empresas. La organización/empresa que haga uso de este servicio, sin embargo, debe encargarse de la compra de sus propios servidores y switches, y la administración recae en ellos. Un último acercamiento, y es el que se profundizará en esta memoria, es la externalización a la **computación en la nube**, lo cual se explicará con más detalle en el capítulo 4.

3.2.2. Costo y disponibilidad de energía eléctrica

Los DC son una de las instalaciones que usan más energía eléctrica a nivel mundial, cuyo crecimiento es realmente alarmante llegando a un consumo anual de 416 TWh durante el 2015 [4] (aproximadamente lo que consume todo el reino unido), llegando actualmente a consumir entre 3-4 % de consumo energético a nivel mundial. Es por esto que en el momento de elegir una ubicación para un DC, es preferible elegir un lugar en el que la energía eléctrica sea de fácil acceso, con acceso a múltiples sistemas interconectados (múltiple acceso a redes de energía), a precios bajos. En este momento, los países con

mayor producción energética son China (2640Mtoe[3]), EE.UU. (2012 Mtoe), Rusia, 1341 (Mtoe) Arabia Saudita (650 Mtoe), India, (593 Mtoe), donde además el precio energético por KW/H es bastante bajo debido a que en su mayoría usan energía generada a partir de combustibles fósiles [6], siendo Chile uno de los países más pobres en este ámbito llegando solo a 14 Mtoe [5].

3.2.3. Clima y posibles desastres naturales

Si bien algunos climas fríos y secos proveen una ventaja respecto a climas cálidos, ya que se ahorra bastante en el uso de enfriamiento por aire (Free Cooling), existen lugares que son más propensos a ser golpeados por desastres naturales, como pueden ser inundaciones, terremotos, huracanes, tornados, erupciones volcánicas, incluso desastres provocados por el hombre, como incendios, o explosiones nucleares. Si un DC llega a ser azotado por alguno de estos fenómenos, es muy probable que se pierda parte de los datos almacenados, o se produzcan interrupciones en los servicios. Si bien se podría pensar que estas amenazas no son tan comunes, en los últimos 10 años, al menos el 50 % de los DC han tenido cortes, o fallas debido a desastres naturales más que nada por falta de preparación [7]. Un estándar que se encarga de recomendar buenas prácticas en esta área es el mencionado ANSI/BICSI 002.

3.2.4. Enlaces y telecomunicaciones

Los encargados de proveer Internet a los DC son los conocidos *ISP* (Internet Service Provider). Un ISP no es más que una red con grandes switches/routers y enlaces capaces de soportar enormes velocidades de transmisiones que proveen acceso a Internet. Los ISP están catalogados en 3 niveles, o *Tiers*[4]. Una red o ISP de nivel 1 (Tier-1) se identifica por estar conectado al resto de los ISP de Tier-1 a través de interconexión pública o privada (public/private peering) [8], en el cual ISPs de Tier 1 intercambian datos con otro ISP de características similares en forma de común acuerdo, es decir: sin exigir pagos entre ellos, tener conexiones directas por medio de cables transatlánticos (cables submarinos)

[3]Tonelada equivalente de petróleo. 1Mtoe = 11630 KWatt/hora
[4]No confundir con la clasificación Tier de Uptime o TIA

y estar conectados a múltiples ISPs de Tier-2[5]. Dependiendo del tamaño, un DC puede estar conectado a 1 o múltiples ISPs de distintos Tier para mantener redundancia. En Chile actualmente los únicos ISPs de Tier-1 son Level3 y TIWS (Telefonica International Wholesale Service) [27], que es donde se conectan la mayoría de los ISPs que conocemos en la vida cotidiana (VTR, Movistar, Manquehue, Claro, etc) para salir a Internet. En el momento de elegir una ubicación, es conveniente analizar las ventajas/desventajas de instalar un DC con conexión directa a ISPs Tier-1.

Existen otros elementos a ser tomados en cuenta en la planeación de un DC como por ejemplo la facilidad de acceso a la instalación (aeropuertos, carreteras), calidad de educación de un país y cantidad de profesionales en el área, número de habitantes, planeación de capacidad, sin embargo, estos elementos se alejan del alcance de esta memoria.

3.3. Infraestructura Física

La infraestructura física de un DC define la topología de los componentes críticos de todo el lugar, siendo los principales la energía, el cableado, el enfriamiento y ventilación, y la ubicación de los racks. Se detallan por lo tanto estas 4 áreas en particular, con tal de proveer al lector una visión general de una infraestructura física de un DC.

3.3.1. Energía

Los DC están conectados a una red de distribución de energía eléctrica pública o privada. Si por alguna razón esta red de distribución sufre un corte de servicios inesperados, se podría dejar sin operación al DC por muchas horas. Por lo tanto es necesario poseer un sistema de generación de energía independiente, automático y autónomo.

Los dispositivos UPS son un elemento esencial en los DC, ya que sirven principalmente para corregir distintas fluctuaciones que pueden existir al alimentar un dispositivo electrónico (Si el voltaje varía bruscamente afuera del DC, el sistema UPS se encarga de mantener el voltaje donde corresponde) y para proveer energía durante un breve

[5]Un ISP de Tier 2-3 no está conectado a través de acuerdos a otros ISP Tier 1, sino que debe pagar por usar sus recursos. Aunque generalmente Tier 2 y 3 se usan de forma reemplazable, la diferencia principal es que un Tier 2 está conectado a otros Tier 2 y al mismo tiempo paga por uso de tráfico a ISPs Tier 1, mientras que un Tier 3 solamente paga por uso de tráfico

intervalo de tiempo ante un eventual corte de suministro[6]. Estos sistemas se ubican generalmente entre un ATS (Automatic Transfer Switch) [7] y la carga crítica del DC (generalmente acompañados de una unidad de mantenimiento conocida como Maintenance Bypass Switch la cual sirve para realizar mantenimiento a la unidad UPS sin necesidad de interrumpir el suministro de energía) [16]. Un esquema muy simplificado se puede observar en la figura 5.

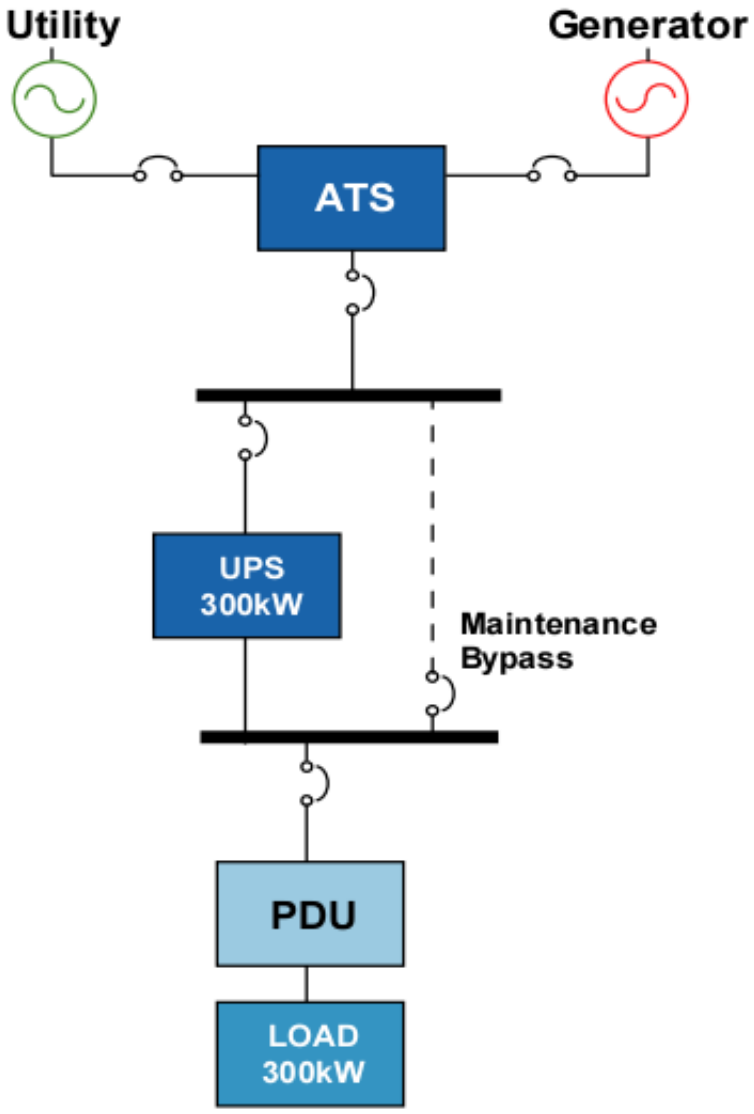

Figura 5, Ubicación de un módulo UPS de un diseño de capacidad según TIA 942. Fuente: www.apc.com.

[6]Este tiempo puede variar entre un par de minutos a un par de horas. Su finalidad no es generar energía, si no que proveer energía a la carga crítica hasta que se enciendan los grupos electrógenos

[7]El interruptor de transferencia es un switch eléctrico encargado de intercambiar entre una fuente de alimentación a otra, generalmente intercambia entre la alimentación principal proveída por un distribuidor de energía público y un generador de emergencia.

Si bien existen distintos sistemas UPS: Standby (offline), Line-Interactive y Online [28], también existen diferentes configuraciones de implementación de sistemas UPS, y aunque existen variantes de estos sistemas, las configuraciones más usadas en la actualidad son las establecidas por el estándar TIA 942:

- **Capacidad:** Los diseños de sistemas UPS se describen usando la nomenclatura "N". N se refiere específicamente a la cantidad de dispositivos para suplir las necesidades requeridas de un diseño de arquitectura. En este caso, se refiere a la cantidad de unidades UPS necesarias para soportar la carga crítica del DC. La carga crítica corresponde a la energía necesaria para que puedan funcionar los componentes críticos para el DC; es decir principalmente los servidores, dispositivos de redes, dispositivos de monitoreo y emergencia, dispositivos de almacenamiento, etc [15]. Un diseño de *capacidad* o también llamado sistema N, es un sistema que no ofrece ningún tipo de redundancia, y usa la cantidad de dispositivos UPS igual al número necesitado por la carga crítica. Esta configuración corresponde a un DC del tipo Tier 1 según el estándar TIA 942.

- **Redundante aislada:** Corresponde a una configuración del tipo N+1. En esta configuración se tiene 1 dispositivo UPS principal el cual se encarga de proveer a toda la carga crítica, y un módulo aislado, el cual solamente entrará en operación cuando falle el módulo principal (en cuyo caso el módulo aislado será el nuevo encargado de proveer energía a toda a la carga crítica). Para esta configuración, no se requiere estrictamente que ambos módulos UPS sean similares, de hecho pueden ser de distinto fabricante/modelo/características. Esta configuración corresponde a un DC del tipo Tier 2 según el estándar TIA 942.

- **Redundante paralela:** Esta configuración corresponde a la paralelización de múltiples unidades UPS de *iguales*[8] características alimentadas por un enlace compartido (lo que en cierta forma es una desventaja ya que convierte a este enlace en un punto único de falla). El sistema se llama N+1 si la energía proveída por el UPS de repuesto es igual a la capacidad de un módulo del sistema, N+2 si la energía de repuesto es 2 veces la capacidad de un módulo, etc. Los módulos UPS en paralelo deben estar sincronizados mediante un panel de paralelización externo, el cual se encarga de controlar un voltaje único de salida. Una de sus ventajas

[8] Deben ser idénticos en cuanto a capacidad, modelo y fabricante

sobre el resto, es que su configuración es escalable a medida que las necesidades energéticas aumenten. Esta configuración corresponde a un DC del tipo Tier 2 según el estándar TIA 942.

- **Redundante distribuida:** Esta configuración presenta 2 o más sistemas UPS independientes, donde cada uno está conectado a diferentes proveedores de distribución. Cada uno de estos sistemas independientes, tiene capacidades de encargarse independiente de la carga crítica a través de STS (Static Transfer Switch) [9]. Esta configuración corresponde a un DC del tipo Tier 3 según el estándar TIA 942.

- **Sistema más sistema - Doble redundancia en paralelo:** Esta configuración, a veces llamada [(N+1) + (N+1)] o 2(N+1), es la configuración más confiable y a la vez más cara de implementar, convirtiendo a todo el sistema tolerante a fallas, mediante la duplicación completa de sistemas redundantes. Si bien es bastante más caro implementar este diseño, en términos de infraestructura y de ingeniería (aunque la configuración sea genérica, las modificaciones pertinentes a esta se hacen de acuerdo a las necesidades), esta configuración afecta directamente al *uptime* del DC, llegando a ofrecer una disponibilidad de 99.9997 % (es decir hasta 1[min], 37[seg] de *downtime* en un año [13]) lo cual influye considerablemente en los ingresos/pérdidas monetarias de la empresa. Esta configuración corresponde a un DC del tipo Tier 4 según el estándar TIA 942.

Aunque es cierto que una configuración altamente redundante (sistemas 2(N+1)) han demostrado ser completamente tolerante a fallas, avances en virtualización y enlaces de gran ancho de banda han permitido que DC con configuración N replicados remotamente, presenten mayor disponibilidad que 1 solo DC completamente redundante.

3.3.2. Cableado estructurado

El cableado dentro del DC puede llegar a ser bastante problemático si no se realiza de la forma correcta, ya que incluye cableado de servidores y dispositivos de almacenamiento *a* dispositivos de redes, y *entre* dispositivos de redes. Problemas en el cableado afectan directamente al rendimiento del DC, produciéndose problemas como asignación

[9] Los interruptores de transferencia estática funcionan similar a los ATS, sin embargo intercambian la energía entre sistemas UPS

incorrecta de vlans, duplex mismatch [10] problemas de DNS, entre otros [17]. Es por esto que se hace necesario seguir un cableado estructurado. Si bien existen diferentes organizaciones internacionales encargadas de crear estándares para el cableado en telecomunicaciones, con el fin de asegurar uniformidad, compatibilidad y disponibilidad, un esquema estructurado está compuesto por los siguientes elementos principales [18][11]

- **Entrance Room - Sala de Entrada**: Consiste en el equipamiento necesario (cables, hardware de conexión, dispositivos de protección) para conectarse al proveedor de servicios ya sea un ISP, telefonía, o cualquier otro sistema externo.
- **Telecomunications Room - Sala de telecomunicaciones**: Es una habitación que mantiene paneles de conexiones, y equipamiento general que interconecta el cableado vertical y horizontal. Estas habitaciones están pensadas para facilitar la administración del cableado horizontal.
- **Backbone Cabling - Cableado vertical**: Sistema de cableado que se encarga de interconectar ISP, salas de comunicaciones y salas de distribución. Generalmente se implementa como topología estrella jerárquica, donde la raíz corresponde a un campus distribuidor, encargado de distribuir el tráfico proveniente del ISP a otras salas de distribución más pequeñas.
- **Horizontal Cabling - Cableado horizontal**: Sistema de cableado que se encarga de conectar servidores ubicados en los racks con un sistema de interconexión cross-connect, el cual funciona como el nodo central en una topología estrella, conocidos como topologías Top-Of-Rack y End-Of-Row

  Tanto para cableado vertical como horizontal se usan cables tipo UTP[12], STP[13] y/o fibra óptica.

Si bien el cableado horizontal y vertical son elementos pertenecientes al estándar TIA-568, existen diferentes arquitecturas de cableado estructurado que también hacen uso de

[10] Duplex mismatch ocurre cuando 2 dispositivos estan conectados al mismo tiempo usando $full-duplex$ y $half-duplex$ llegando a reducir considerablemente el rendimiento hasta 100 kbps o menos

[11] Descritos en el estándar TIA-568

[12] Unshielded Twister Pair, cable de par trenzado sin blindaje, está separado actualmente en 7 categorías: $cat1$, $cat2$,.. $cat7$, usado ampliamente en telefonía y comunicaciones.

[13] Shielded Twisted Pair: cable de par trenzado blindado. Es el mismo que el UTP, salvo que viene recubierto con una capa protectora para protegerlo del ruido eléctrico

estos elementos, y puede encontrarse un resumen general de estas arquitecturas especificado por diferentes estándares en [19]. Estos estándares, además de definir un sistema general de cableado, también se encargan de definir el ciclo de vida, longitud máxima, prácticas de instalación, correcta documentación, etiquetado y cuidado de los cables entre otros.

3.3.3. Enfriamiento

El enfriamiento de los equipos es uno de los procesos más esenciales dentro de los DC, ya que los componentes IT liberan una cantidad elevada de energía, la que se convierte en calor el cual puede tener un impacto negativo en el rendimiento de los equipos IT, e incluso dañarlos. Aunque los dispositivos pueden enfriarse a través de líquido o aire, en esta sección solamente se detallará los sistemas de enfriamiento basados en aire frío, ya que es más fácil mover aire en un entorno cerrado, donde las fugas de aire no presentan mayor peligro ni para el equipamiento ni para el personal, por lo tanto es menos riesgoso (Aunque existen casos en los que el enfriamiento en aceites especiales, al no afectar a los sistemas eléctricos, si se ha hecho notar, como en supercomputadores, o en minería de *criptomonedas* [45]).

Un sistema de enfriamiento en DC (el mismo enfoque se sigue para centrales eléctricas, o plantas en las que se genera calor excesivo), consta de 3 elementos principales, un enfriador de agua, un sistema de distribución y manejo del aire (CRAC - Computer Room Air Conditioning, CRAH - Computer Room Air Handler) [3], y una estrategia de enfriamiento (Por Rack, por Row, o por sala) [14][20].

- **Arquitectura CRAC: Computer Room Air Conditioning** Esta arquitectura usa refrigerantes, o sistemas de condensación, o compresión de aire, por cada unidad manejadora de aire. En este diseño no se usa agua, es orientado a data centers pequeños porque cada unidad manejadora tiene su propio sistema de enfriamiento, lo cual es más caro, pero es más manejable en espacios pequeños.

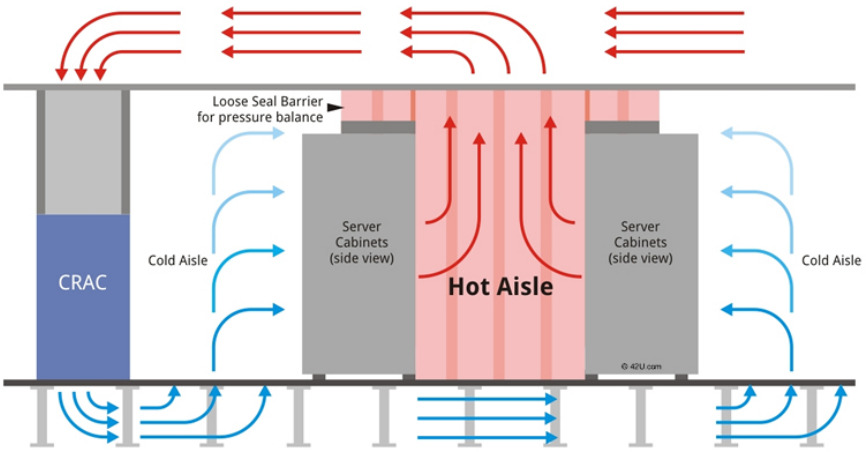

FIGURA 6, CONFIGURACIÓN DE DISTRIBUCIÓN DE AIRE EN LA HABITACIÓN IT BASADA EN PASILLOS FRÍOS Y CALIENTES.

FUENTE: WWW.42U.COM.

- **Arquitectura CRAH - Computer Room Air Handler** Un sistema de manejo de aire usando agua enfriada (Chilled-Water) está compuesto de un suministro de agua[14], una torre de enfriamiento (la cual generalmente se ubica en los techos de los DC, o en las afueras de las instalaciones), y un enfriador (Chiller), los cuales trabajan en conjunto para transportar agua fría (entre 6-8 °C) a los sistemas de manejo de aire dentro del DC, a través de bombas de agua. El agua fría que circula dentro de serpentines de las unidades manejadoras de aire, se encarga de enfriar el aire a su alrededor, el cual es expulsado hacia adentro del data center mediante ventiladores.

[14]Un DC de tamaño medio (15 MW/año) puede llegar a usar hasta 1.3 millones de litros de agua por día. Si bien es un número ínfimo al lado del mercado de la agricultura, se ha fomentado el uso óptimo de agua en DC, por lo que The Green Grid ha creado la métrica WUE - Water Usage Efectiveness, para medir eficiencia del uso de agua, sin embargo este parámetro no ha sido tan popular como el PUE.

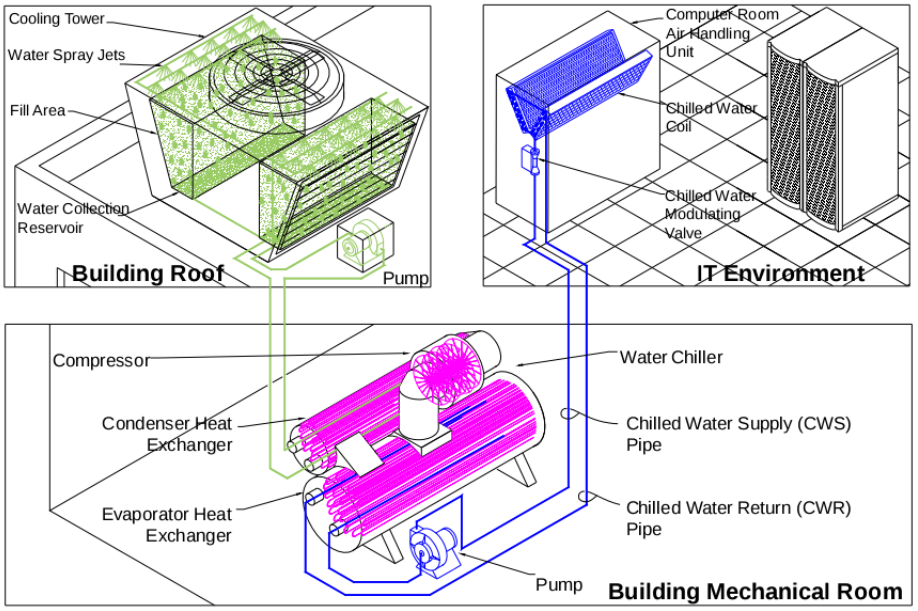

FIGURA 7, SISTEMA DE ENFRIAMIENTO BASADO EN CHILLER.

FUENTE: WWW.APC.COM.

Esa agua fría es transportada al interior del DC a través de ventiladores sistemas de manejo de aire CRAC o CRAH. Estos sistemas se encargan de monitorear y mantener la temperatura de un DC, humedad, y distribuir y regular el aire frio en el interior de la sala IT.

- **Arquitectura Free Cooling** En las arquitecturas anteriores, generalmente el aire es reciclado, es decir las partículas de aire frías se convierten en calientes y viceversa, generando un ciclo infinito. En free cooling, el aire es tomado del exterior, usado para enfriar el equipamiento IT a través de una configuración de ventiladores/extractores, y luego es expulsado hacia el exterior, sin ser reusado. Esta arquitectura hace uso del aire exterior, sin embargo como las características ambientales no son las mismas durante todo un año, esta arquitectura muchas veces es usada como arquitectura auxiliar (Aunque existen data centers de última tecnología como Facebook Altoona, que usan 100 % free cooling)[15]. Es importante destacar que no en todos los lugares del planeta puede implementarse la arquitectura free cooling, ya que depende directamente de la humedad relativa existente en el exterior.

  Esta arquitectura mejora considerablemente el PUE de un DC ya que se elimina por completo la alimentación energética de sistemas de enfriamiento, además de

[15]Entel CDLV3 es el primer data center en latinoamérica en usar 100 % free cooling

reducir considerablemente la emisión de CO2 y los costos asociados a enfriamiento en alrededor de un 67 % [49].

ASHRAE [21] es una organización encargada de liberar estándares respecto al enfriamiento de data centers, además de proponer temperaturas y humedades óptimas referenciales para el equipamiento IT.

3.3.4. Racks

El último elemento de la capa física del modelo de capas de DC, corresponde a la ubicación de los equipos de IT en las salas de datos. Los racks son básicamente cabinas para almacenar equipamiento. Una sala IT o sala de datos está compuesta por filas de racks, que a la vez contienen rejillas de montaje para *enmarcar* servidores, switches, equipos de enfriamiento, paneles de conexiones, etc. La unidad estándar para medir la altura de una rejilla de montaje es un U ($Unit$), la cual equivale a 1.75 pulgadas (44.45mm). Aunque pueden variar, un rack común puede almacenar hasta 42 módulos de 1U [22]. El objetivo de los racks es principalmente optimizar el espacio al permitir apilar los servidores y/o switches de forma horizontal[16], además de proveer un entorno organizado para la distribución de energía.

Figura 8, Servidor 1U, Dell PowerEdge, usado ampliamente en Data Centers. Fuente: www.dell.com.

Existen servidores y switches de distintos tamaños, lo común es que se usen módulos de 1U, 2U y 4U para servidores, dependiendo de su tamaño, mientras que los servidores Blade pueden medir de 4U hacia arriba[17]

[16] *Pizza Box* es un término acuñado por la industria para referirse a este tipo de servidores, por la similitud de estos dispositivos con las cajas de *Pizzas*

[17] Técnicamente un *Gabinete Blade* es un gabinete que puede almacenar múltiples servidores de tamaño muy reducido, los cuales en ocasiones incluyen sus propios switches, y sistemas UPS en un pequeño espacio.

Figura 9, Servidor Blade, Dell m1000e 10U, con 16 servidores independientes. Fuente: www.dell.com.

Aunque no existe ninguna *regla de oro* para la elección de racks, es común que se realice un proceso de selección que analice al menos los atributos del equipamiento a ser instalado y/o accesorios extra, los que incluyen puertas con cerraduras, sujetadores anti-sísmicos, administración de cables, monitoreo de corriente y temperatura. Aunque las medidas de un rack estándar de 19 pulgadas, y 42 rejillas están especificadas en los documentos EIA-310-D, diversas empresas han optado por usar racks alternativos para satisfacer sus necesidades. En [25] se puede encontrar una tabla que presenta diferentes alternativas de racks y sus respectivos beneficios.

Para diseñar layouts de salas IT existen diversas herramientas llamadas DCIM (Data Center Infrastructure Management) las cuales permiten modelar una sala con racks, equipamiento de enfriamiento y suministro de energía, con tal de optimizar el espacio.

3.4. Seguridad física de un Data Center

La seguridad física de un Data Center puede observarse en la literatura en conjunto con la infraestructura física, sin embargo en este trabajo se considera como una capa especial y

separada, ya que la seguridad física es uno de los elementos más importantes al levantar un DC, puesto que cualquier falla en esta capa podría dejar expuesta la información almacenada en el DC, o provocar cortes no planificados, afectando directamente a los servicios, y por ende impactar al negocio u organización.

Existen diversas formas de proveer una capa de seguridad al recinto físico del lugar. Empresas como Amazon y Google han optado por no revelar la ubicación exacta de algunos DC, incluso han llegado a disfrazar sus DC como bodegas de propósito general [10][12], prohibiendo completamente visitas a público exterior. Si bien con un pequeño esfuerzo es posible encontrar estas ubicaciones, esta es una entre muchas otras técnicas de asegurar el recinto físico donde está ubicado el DC.

Para que la seguridad en un DC sea completamente efectiva, o con mayor probabilidad de éxito, deben tenerse 2 elementos en cuenta: los componentes físicos, y los procesos operacionales.

3.4.1. Factores Operacionales

El factor humano representa una de las causas más grandes en violaciones de seguridad, siendo el responsable del 95 % de los incidentes[11], donde se incluyen contraseñas débiles, pérdida de laptops, e-mails mal direccionados con información sensible, etc. Los atacantes conocen esto y por eso buscan atacar el eslabón más débil de esta cadena: las personas. Es por eso que además de los componentes físicos, se deben establecer reglas sobre los procesos operacionales. Dentro de los procesos operacionales podemos encontrar políticas de entrenamiento, contratación, visitas, políticas de correcta desmantelación de servidores/discos que alcanzaron su vida útil (de forma que no se pueda obtener información a través de ningún método).

3.4.2. Componentes Físicos

Dentro de los componentes físicos existen múltiples elementos de seguridad, como lo son muros, cercos eléctricos, sistemas de alarma y cámaras de vigilancia, control de acceso, establecimiento de perímetros de acceso, y elementos de autenticación multifactor [18]. (La

[18] La Autenticación multifactor hace referencia a múltiples elementos necesarios para la autenticación (es decir asegurar que la persona es quien dice ser), generalmente corresponde algo que se *sabe* (contraseñas, o frases de paso), algo que se *tiene* (token de seguridad, dispositivos NFC o RFID) y algo que se *es* (algo inherente a la persona, como huella digital, voz, retina, etc.)

AMF es usada ampliamente en DC tanto para acceder a lugares físicos como sistemas de software, sin embargo, no debe confundirse con la autorización, la cual verifica que los recursos sean manejados solo por personas autorizadas). Para proporcionar seguridad física tanto de los datos como de la instalación, se destacan diversas arquitecturas de diseño, como lo son sistemas SCADA para monitorear estado de los componentes (UPS, generadores, sistemas de enfriamiento) y sistemas BMS para manejo del edificio, dentro de los que se incluyen Circuito Cerrado de Televisión - CCTV y detección y control de incendios.

Certificaciones como SAS 70, SSAE 16 y PCI Security Standard Council [29] aseguran que los procesos de seguridad se estén cumpliendo de forma acorde a los acuerdos de nivel de servicio (SLAs) y a los requerimientos legales.

3.5. Red

Durante muchos años la arquitectura Three-Tier network patentada por Cisco fue considerada estado del arte para redes dentro de DC. Sin embargo, el alto costo de los equipos, y su pobre escalabilidad hicieron necesario un cambio de paradigma. La mencionada explosión del Cloud Computing, IoT, HPC y Big Data han provocado que el tráfico de transferencia de datos entre servidores alcance un 70 % del total del tráfico generado dentro del DC [31] (tráfico conocido como $east - west$), donde el 30 % restante corresponde a tráfico que entra y/o sale del DC hacia Internet u otras WANs (tráfico conocido como $north - south$), por lo tanto las arquitecturas han ido evolucionando a lo largo del tiempo con tal de proveer mayor escalabilidad y eficiencia.

3.5.1. Cisco 3-Tier Architecture - Arquitectura de redes de 3 niveles

Esta arquitectura ha sido una de las arquitecturas más usadas dentro de los DC debido a su robustez, buen rendimiento y diseño modular. Sin embargo, debido al crecimiento exponencial que han tenido los DC, y el surgimiento del paradigma del Cloud Computing, este modelo ya no es viable para grandes DC, debido a su alto costo, baja escalabilidad y baja eficiencia. Además es importante conocer los conceptos claves para poder comprender de mejor forma las arquitecturas modernas. La Arquitectura de 3 niveles está formada por: la capa de acceso, capa de distribución y capa de núcleo. Un esquema simplificado de esta arquitectura puede apreciarse en la figura 10.

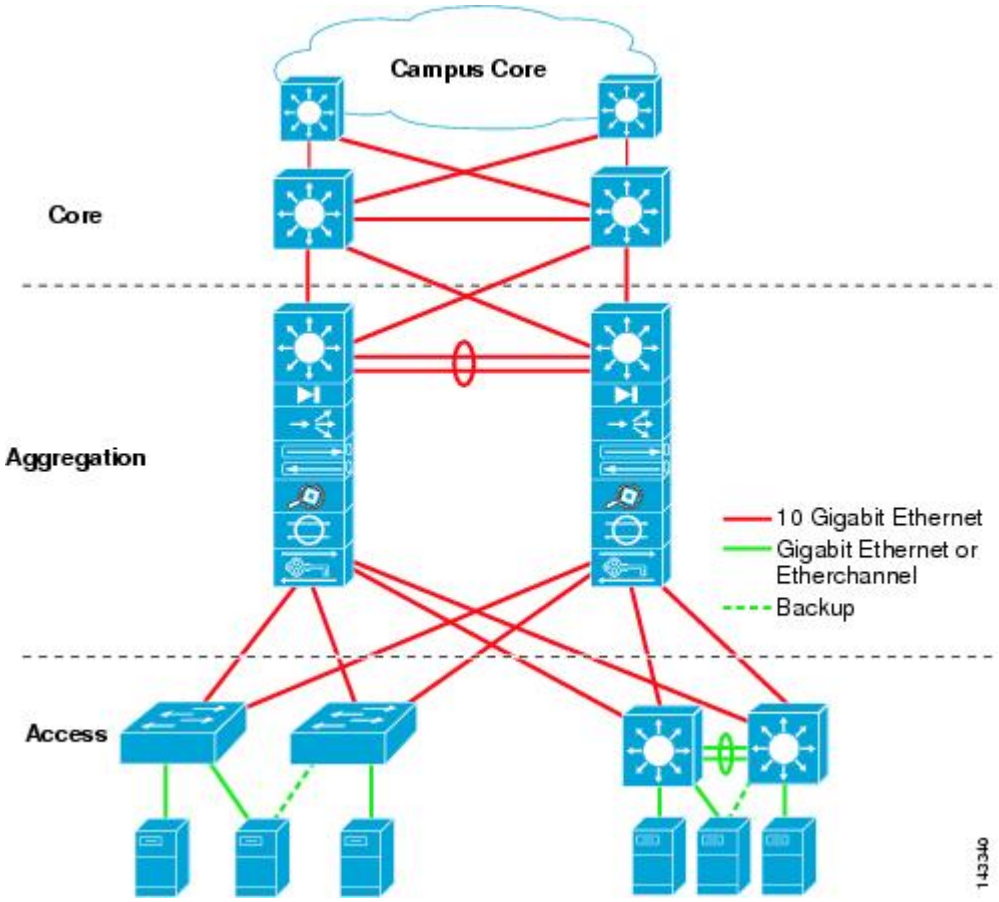

FIGURA 10, ARQUITECTURA DE 3 NIVELES DE CISCO.

FUENTE: WWW.CISCO.COM.

- **Core Layer**

  La capa de Núcleo o Core Layer, está diseñada para transportar grandes cantidades de datos y balancear la carga entre switches de la capa de distribución, con enlaces de alta velocidad, donde se prioriza disminuir latencia y maximizar throughput. Funciona como la puerta de enlace para el resto de la red, conectándose a WANs ya sean, campus universitarios, DCs remotos, o resto de Internet. El equipamiento para esta capa consiste en switches de alta velocidad, costosos tanto energética como monetariamente (un core switch puede costar desde 40.000 USD hacia arriba, además de consumir una tremenda cantidad de energía [32]). Los switches de la capa de núcleo se conectan a través de enlaces etherChannel, la cual es una tecnología que permite la agregación de múltiples enlaces Ethernet en un gran enlace lógico, permitiendo un mayor throughput (por ejemplo permite juntar 2 enlaces de 10Gbps en 1 solo de 20 Gpbs). Se considera una capa de vital importancia, ya que cualquier problema en un switch de esta capa podría causar cortes de servicio en

múltiples servidores dentro del DC, paralizando las operaciones por completo (una gran desventaja, ya que son un punto único de falla de toda la arquitectura), por lo tanto se necesitan enlaces y switches altamente redundantes. Se dice que esta arquitectura presenta una baja escalabilidad ya que escala de forma vertical, es decir deben mejorarse los core y distribution switches para tener un mejor rendimiento. Se verá a continuación que las arquitecturas modernas usan escalamiento horizontal, agregando switches de forma paralela.

- **Distribution - Aggregation Layer**

  La capa de distribución se encarga principalmente de enrutar y filtrar paquetes, usando firewalls y balanceadores de carga, además de dar acceso a WANs a la capa de acceso mediante enlaces redundantes. También es posible agregar políticas de priorización de paquetes para mejorar el QoS (Quality of service)[19]

  La conexión ocurre *entre* switches, que corren el protocolo Spanning Tree (STP) con tal de eliminar ciclos[20]. Ambas capas, tanto Core como Distribution usan switches multicapas, es decir trabajan en la capa 2 y 3 del modelo OSI, por lo tanto manejan protocolos por MAC e IP. En ocasiones se puede ver la capa de Distribución y Core juntas, formando una arquitectura Two-Tier, conocida como Collapsed Core architecture.

- **Access Layer**

  En esta capa los servidores ubicados en los racks se conectan a los switches de la capa de acceso, ya sean 1U servers, blade servers y/o SANs servers. Los switches de la capa de acceso generalmente se presentan en 2 configuraciones posibles:

  Top-Of-Rack: Los servidores dentro del rack se conectan directamente a un switch ubicado dentro del mismo rack (preferentemente en los espacios superiores). El switch se conecta mediante fibra a los switches (redundantes) de la capa de distribución, de esta forma cada rack se ve como si fuera una sola unidad (Este diseño

[19] La calidad de servicio establece la priorización de paquetes por ejemplo multimedia, ya que es más importante que esos paquetes se procesen primero que por ejemplo una transferencia de archivos, en la que no se necesita tanta fluidez

[20] Como el protocolo Ethernet exije enviar mensajes broadcast para descubrir la red, un frame puede quedarse estancado en un bucle infinito entre switches redundantes (al no existir un campo TTL), afectando directamente el rendimiento total de la red.

sin embargo no es escalable para un DC que posea muchos racks, debido a su dificultad de administración, y mantención de las tablas MAC).

End-Of-Row: En este diseño, se dispone de un rack en particular dedicado solamente al manejo de redes, el cual se ubicará físicamente en un sector de una *fila* de racks (preferentemente al final) [33].

Como los enlaces son completamente redundantes entre switches de la capa de acceso con la capa de distribución, pueden producirse ciclos de frames Ethernet, por lo tanto los switches deben correr Spanning Tree Protocol para evitar su formación. Esto provoca que este tipo de arquitecturas tenga un manejo pobre del ancho de banda, ya que STP al crear el árbol de expansión, bloquea una gran cantidad de enlaces, llegando a perder hasta un 50 % de ancho de banda en toda la red [36].

La arquitectura de redes de 3 capas presenta una alta tasa de oversubscription[21], con un diseño típico de 8:1 [46], aunque pueden variar dependiendo de las necesidades. Se sigue este enfoque para no impactar tanto en los costos totales del DC y aumentar la cantidad de puertos a hosts, ya que los enlaces de alta velocidad implican un mayor costo en el diseño (de esta forma se reducen costos tanto en switches como cables.)

[21] Oversubscription: tasa de tráfico que puede obtenerse en el peor caso de transmisión de datos en un switch en particular, y corresponde a la tasa total de tráfico de entrada a un switch divido en la tasa total de tráfico de salida.

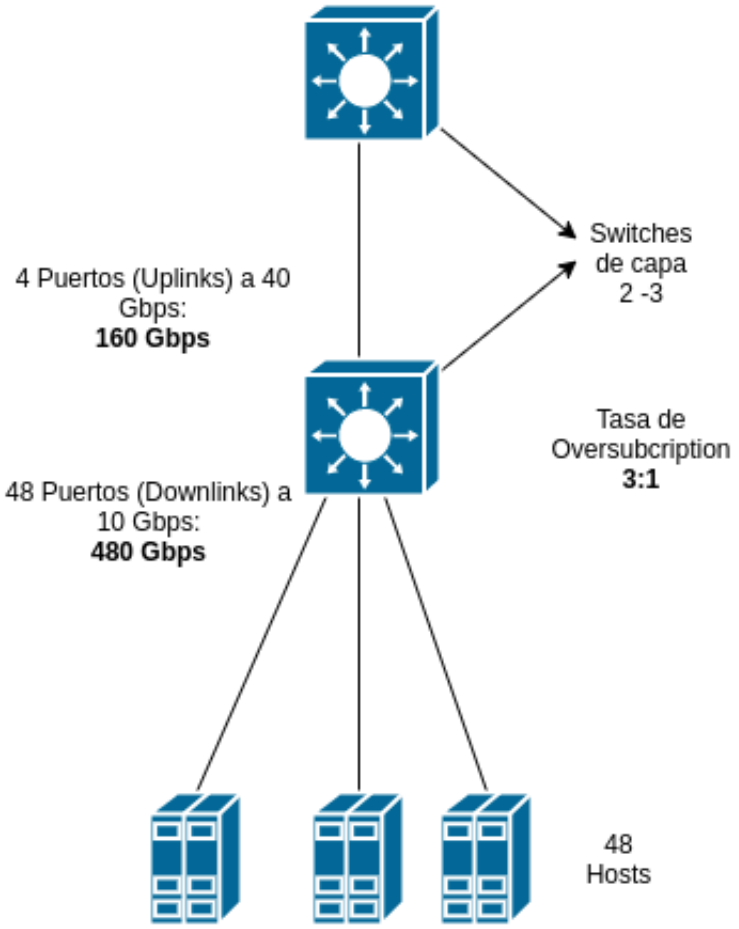

Figura 11, Bosquejo oversubscription 3:1.

Fuente: Elaboración propia.

Si bien una alta oversubscription no es un problema cuando no se genera un gran tráfico (menor a 60 % de la capacidad de los enlaces) [37], ciertamente lo es cuando el tráfico supera el 60 %, produciendo elevadas pérdidas de paquetes.

3.5.2. Fat Tree

La arquitectura Fat Tree es una modificación al modelo 3-Tier, con la diferencia que en este caso se van aumentando los tamaños de los enlaces a medida que se van acercando a la capa de núcleo, con tal de mantener una oversubscription de 1:1.

Sin embargo, la principal desventaja de 3-Tier Architecture y sus derivados, es su baja escalabilidad. Cada switch tiene que almacenar una entrada por cada MAC en la tabla

MAC[22], por lo tanto para una gran cantidad de servidores (miles), este acercamiento no es viable.

Una solución es usar PoDs. Un PoD (Point Of Delivery) es un conjunto de dispositivos de red, cómputo y almacenamiento, encapsulado en una unidad única, cuyo fin es crear una arquitectura de redes basada en módulos [35]. A cada POD se le asigna una pseudo-direccion-MAC, la cual será el identificador del POD. Un POD puede estar compuesto por un simple rack, una fila de racks, o incluso un conjunto de racks. De esta forma, la comunicación ocurre entre PoDs, donde cada pod se encarga de mantener las direcciones de los elementos que lo componen, reduciendo drásticamente la administración de la red. Este acercamiento es ampliamente usado en la arquitectura Leaf and Spine en data centers de alta escalabilidad.

3.5.3. Leaf and Spine

La topología Leaf and Spine es una arquitectura de redes usada en data centers con gran necesidad de escalabilidad, con tal de solucionar los problemas de la arquitectura 3-Tier. Es una arquitectura altamente escalable, con un diseño estándar de oversubscription 3:1 (480 Gbps / 160 Gbps). Esta arquitectura presenta solamente 2 capas, una de *Leaf Switches*, (los cuales son típicamente ToR switches) cuya función es hacer de capa de acceso y filtrado, donde los servidores físicos, firewalls y balanceadores se conectan a ella, y por una capa de *Spine Switches*, los cuales están conectados a cada uno de los leaf switches. De esta forma, los switches de leaf están conectados a lo más a 1 salto de cualquier otro switch en la red, lo cual disminuye considerablemente la latencia. Una topología Leaf and Spine puede apreciarse en la siguiente figura:

[22]Tabla de reenvío de frames también conocida como forwarding table

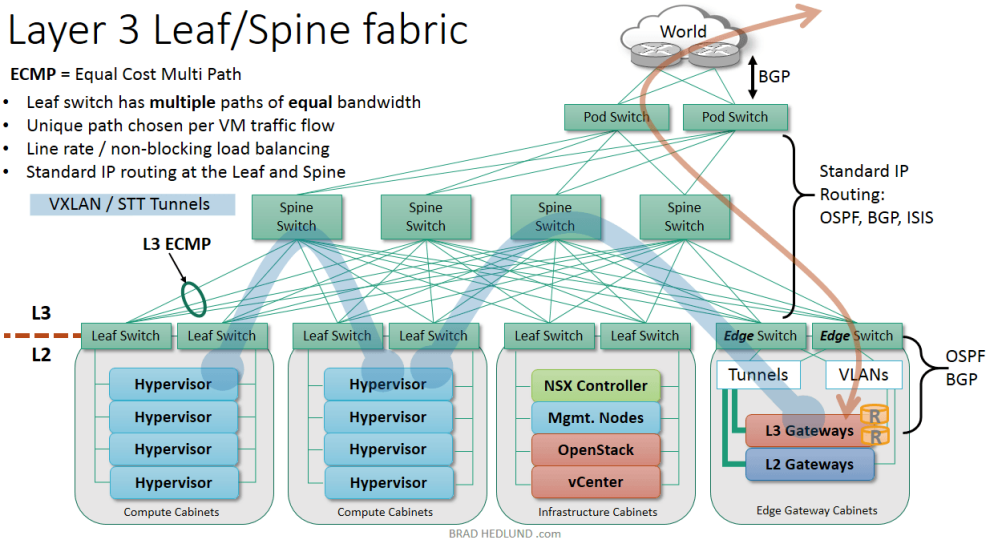

FIGURA 12, ARQUITECTURA LEAF AND SPINE.

FUENTE: HTTP://WWW.NETWORKSBASELINE.IN.

Los switches dejan de procesar frames de la capa 2 OSI, procesando solamente paquetes de la capa 3 OSI[23]. Al no existir frames entre switches, no es necesario bloquear enlaces usando STP como en 3-Tier, por lo que el rendimiento de la red aumenta considerablemente.

Otra ventaja de Leaf and Spine es su tolerancia a fallas. Si un core switch falla en 3-Tier, podrían generarse cortes importantes en toda la red, mientras que si falla un Spine Switch, o experimenta aumentos inesperados de tráfico, solo se apreciaría una leve degradación en los servicios, ya que siempre habrá un Switch Spine disponible para mover datos en la misma cantidad de saltos.

Una implementación a gran escala de esta topología puede observarse en el data center Altoona de Facebook [38], la cual consiste en la organización de múltiples topologías Leaf and Spine ordenadas como PODs, conectadas entre ellas para proveer una arquitectura altamente escalable, rápida, y tolerante a fallas.

Si bien esta arquitectura ha demostrado mejores resultados en tráficos server-to-server en grandes DC, también tiene sus desventajas, dentro de las que podemos encontrar principalmente un alto exceso de cables, ya que todos los Switches entre ambas capas

[23] Aunque Spine and leaf puede ser implementada en la capa 2 usando los protocolos TRILL/SPB los cuales previenen bucles y además permiten usar todos los enlaces, lo más común es ver una implementación de capa 2 OSI entre servidores/leaves y capa 3 OSI entre Leaves y Spines usando el protocolo ECMP (encargado de balancear la carga entre los switches).

deben estar conectados. Además esta nueva topología no disminuye en ningún momento el costo total del DC, ya que los Spine Switches tienen un costo monetario similar a los Core Switches.

Finalmente, como el tráfico entre Leafs y Spines solamente usa paquetes de la capa 3 OSI, se hace más difícil la creación y mantención de VLANs (las VLANs se implementan en frames de la capa 2 OSI) a lo largo de racks, o fuera de PoDs. Han surgido diferentes formas de manejar este problema, y una tecnología que ha sabido manejar esto son las redes definidas por software.

3.5.4. Encapsulación de frames

En un data center común pueden existir cientos de servidores bare-metal, usados para virtualizar otros miles de hosts a través de hipervisores[24], lo cual ha superado completamente la cantidad de hosts soportados en una VLAN (El tag VLAN de 16 bits permite manejar hasta 4096 VLAN ids). Además, mantener y administrar switches con 40k+ hosts es una tarea que podría llevar a muchas fallas. Si bien se pudo apreciar que esto puede resolverse usando PODs en arquitecturas del tipo Leaf-Spine, esta arquitectura se comunica mediante la capa 3, por lo tanto no es posible manejar VLANs a lo largo de racks. Una tecnología que resuelve este problema, es la agregación de protocolos de encapsulación o *tunneling*. VXLAN y NVGRE son protocolos de encapsulación que permiten traspasar información de la capa 2 en un paquete de la capa 3 entre diferentes racks (o subredes).

Este tipo de tecnología se conoce como SDN - Software Defined Networking, específicamente *SDN overlay*, en la cual se separa la parte física de los switches físicos usando switches virtualizados a través de los hipervisores. Si bien esta arquitectura es ampliamente usada en data centers, tiene bastantes restricciones, como por ejemplo falta de programabilidad, o flexibilidad pobre. Un acercamiento más robusto, son las SDN basadas en el protocolo OpenFlow.

[24] Un hipervisor ESXi en promedio virtualiza entre 20 y 40 servidores sin perder rendimiento, aunque es posible superar las 200 VMs por host [50]

3.5.5. SDN - Software Defined Networking

La arquitectura SDN, o redes definidas por software, se define como una arquitectura de redes que separa la parte lógica de la parte física de un dispositivo de red con el fin de facilitar su administración, disminuir errores de configuración y agregar programabilidad a las redes, permitiendo cambiar sus reglas dinámicamente, agregando flexibilidad y elasticidad.

Un dispositivo de red (switch o router) tiene implementado un plano de control encargado de correr protocolos de enrutamiento, reenvio, firewall, entre otros, y un plano de datos, el cual corresponde a la parte física del dispositivo encargada de reenviar paquetes a través de circuitos y señales eléctricas. SDN separa estos planos usando el protocolo OpenFlow. Para lograr esto, se necesitan switches que soporten el protocolo OpenFlow u otro similar. Un esquema conceptual de la arquitectura SDN se presenta en la figura:

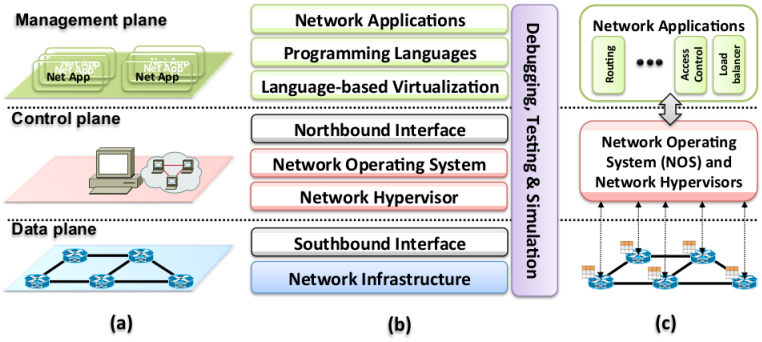

Figura 13, Arquitectura SDN.

Fuente: SDN Fundamentals, Diego Kreutz et al.

El plano de control intercambia información con el plano de datos a través de una API, generalmente llamada interfaz *southbound*. OpenFlow es actualmente la interfaz southbound estándar para redes definidas por software, y actualmente es administrado por la Open Networking Foundation [51]. En el plano de control corre el Controlador SDN, o Network Operating System (por ejemplo Apache FloodLight), el cual implementa toda la lógica de la red, principalmente enrutamiento y/o reenvío de paquetes. Además se encarga de mantener un registro de todos los dispositivos y flowtables[25] de la red.

[25]Una flowtable es una estructura de datos ubicada tanto en el controlador como en los switches OpenFlow, y determina el comportamiento y reenvío de los paquetes en una red SDN

El controlador es el núcleo de la arquitectura SDN, ya que centraliza completamente el manejo de la red. Esto permite manejar la red completamente desde el controlador, lo cual si bien presenta un buen rendimiento, lo convierte en un punto único de falla, ya que un corte o falla en el controlador podría generar una falla en toda la red. Además, un atacante que logre penetrar completamente hasta el controlador podría manejar toda la red a su merced. Existen también controladores distribuidos, como Onix, elasticOn o HyperFlow, los cuales ofrecen mejor tolerancia a fallas ya que eliminan el componente único de falla, sin embargo al estar distribuidos, se pierde consistencia en la información que mantienen los switches en diferentes zonas geográficas.

Este controlador a su vez intercambia información con la capa de aplicación, a través de otra API, conocida como interfaz NorthBound, la cual permite modificar reglas en el controlador, y provee una interfaz para poder programar funcionalidades a la red. Al contrario de la interfaz SouthBound, no existe una interfaz estándar como lo es OpenFlow, ya que cada controlador trae su propia interfaz implementada con bibliotecas diferentes para desarrollar aplicaciones. Dentro de las aplicaciones que finalmente pueden desarrollarse existe la creación de balanceadores de carga, firewall, grupos de seguridad, reglas de entrada/salida, etc.

Una gran ventaja de esta arquitectura, es que SDN se desliga completamente del fabricante de los dispositivos, (los cuales tienen diferentes formas de ser configurados), por lo que mantención y administración recae netamente en el controlador SDN y el protocolo OpenFlow, lo que la hace tremendamente universal al ser completamente independiente de cualquier marca, modelo, o tipo de dispositivo.

3.6. Almacenamiento

Las arquitecturas de almacenamiento son clave en los DC ya que no solo permiten acceder a enormes cantidades de datos en corto tiempo, si no que también son responsables de proveer alta disponibilidad de recursos, principalmente máquinas virtuales.

Existen diferentes soluciones de almacenamiento, y dependiendo del uso que se le quiera dar, ya sea para respaldo a largo plazo, acceso infrecuente, acceso inmediato, o alta disponibilidad. Cuando se implementa una arquitectura de almacenamiento en un DC, deben tenerse diversos elementos en consideración, por ejemplo facilidad de escalabilidad,

análisis de capacidad total sin generar sobreprovisionamiento, disponibilidad, tolerancia a fallas, rendimiento, balanceo de carga, seguridad de la información, presupuesto de la empresa, etc. Se analizan por lo tanto las arquitecturas principales de almacenamiento en data centers que cubren estos elementos.

3.6.1. DAS - Direct Attached Storage

DAS es una arquitectura simple y económica, sin embargo muy limitada, en la cual un dispositivo de almacenamiento está conectado directamente a un servidor físico mediante una interfaz SAS o SATA. Debido a que no existe un área o región compartida, esta arquitectura puede presentar altos porcentajes de desuso. Es por eso que en DC generalmente se presenta el almacenamiento en red.

3.6.2. NAS - Network-Attached Storage

Esta arquitectura es básicamente un servidor de archivos, el cual puede ser accedido a través de una red local, a través de protocolos de transferencia de archivos como FTP, NFS, AFS, SMB, etc. Es un sistema centralizado y escalable, y además permite que los archivos sean compartidos por múltiples hosts, sin embargo al usar interfaces de redes en vez de almacenamiento (como SAS, o SCSI), se pierde bastante rendimiento. Generalmente se implementan como servidores, ya que los dispositivos NAS cuentan con CPU, memoria, tarjetas de red, sin embargo su propósito no es correr grandes procesos, si no que almacenar datos.

3.6.3. Sistemas RAID - Redundant Array of Independent Disks

Los sistemas RAIDs proveen almacenamiento virtual al combinar uno o más discos de forma lógica, ordenados de acuerdo a ciertas configuraciones. Estos sistemas se crearon para proveer una mayor tolerancia a fallas, aumento de capacidad y alta disponibilidad en servidores que debían manejar una alta cantidad de datos (principalmente data centers). Una de las principales ventajas de este tipo de sistemas, es que mantienen hot-swap, lo que permite cambiar discos de forma física de un servidor, sin impactar a los datos ni operaciones. Las configuraciones típicas de sistemas RAIDS son:

- RAID 0 - Disk Stripping: Este tipo de RAID no ofrece ningún tipo de redundancia y/o duplicación, ya que solamente distribuye los discos entre distintos servidores, para maximizar throughput (i.e mayor IOPS). De esta forma se puede estar escribiendo/leyendo 1 solo gran archivo entre varios servidores de forma paralela. El gran problema de RAID 0, es que cualquier tipo de falla en cualquiera de los discos en los que está procesando dicho archivo, genera una corrupción en el archivo total.

- RAID 1 - Mirroring: Este tipo de RAID consiste en crear copias espejo (copias exactas) de un disco, en algún otro servidor, con tal de proveer tolerancia a fallas en caso de que exista algún problema en alguno de los 2 discos, accediendo inmediatamente al disco de respaldo. Una gran desventaja de RAID 1, es que aumenta al doble la cantidad de discos que deben mantenerse, impactando directamente el uso del espacio físico, y aumentando los costos totales del data center.

- RAID 10 - Disk Stripping + Mirroring: Este tipo de RAID es una combinación entre RAID 0 y RAID 1, en la cual se provee el aumento de throughput de RAID 0 además de la tolerancia a fallas de RAID 1, sin embargo se deben tener al menos 4 discos distribuidos para poder ofrecer estas mejoras (2 para Stripping y 2 para mirroring). Si bien este sistema ofrece un alto rendimiento y alta disponibilidad, la capacidad de escalar un sistema RAID 10 es extremadamente baja.

Existen diversas otras configuraciones de RAID, además de las expuestas en esta sección, por ejemplo distribuir datos de acuerdo a paridad, o mezclas híbridas entre paridad, espejo, y separación por bloques, sin embargo escapan del alcance de este trabajo.

Es fácil observar que este tipo de sistemas de almacenamiento tiene un gran problema de escalabilidad, y al ser un sistema de almacenamiento local, no permite distribuir la visibilidad del almacenamiento a través de la red.

Surge entonces la arquitectura estándar de almacenamiento para data centers, la cual se usa para proveer alta disponibilidad de virtualización de servidores, además de ser escalable, y tolerante a fallas, conocida como Storage Area Network.

3.6.4. SAN - Storage Area Network

SAN es una arquitectura de almacenamiento ampliamente usada en data centers, que ofrece escalabilidad, alto rendimiento y alta disponibilidad, basada en almacenamiento a través de redes.

La diferencia principal entre SAN y NAS, es como los servidores *ven* el almacenamiento. En NAS, los servidores, acceden mediante acceso remoto a través de protocolos de la capa de aplicación (como los mencionados FTP, o NFS), sin embargo en la arquitectura SAN, los servidores ven el almacenamiento como si fuera un disco anexado a través de LUNs - Logical Unit Number (también llamado disco lógico), el cual es un identificador de un área o porción de almacenamiento dentro del arreglo SAN. Esta porción puede estar formada por un disco duro, por una partición del disco duro, o por un arreglo de 1 o más discos o tarjetas de almacenamiento. El servidor puede entonces, visualizar esa porción de almacenamiento como un disco montable en el sistema operativo. Una ventaja de este acercamiento, es que dichas porciones de almacenamiento, aunque estén compartidas dentro del mismo arreglo SAN, están aisladas unas de otras en el sentido que una VM no puede acceder a otra LUN. Esta aislación ofrece una mayor seguridad al sistema.

En la figura siguiente se puede observar un esquema típico de las arquitecturas mencionadas, aunque los diseños pueden variar de organización en organización dependiendo de las necesidades.

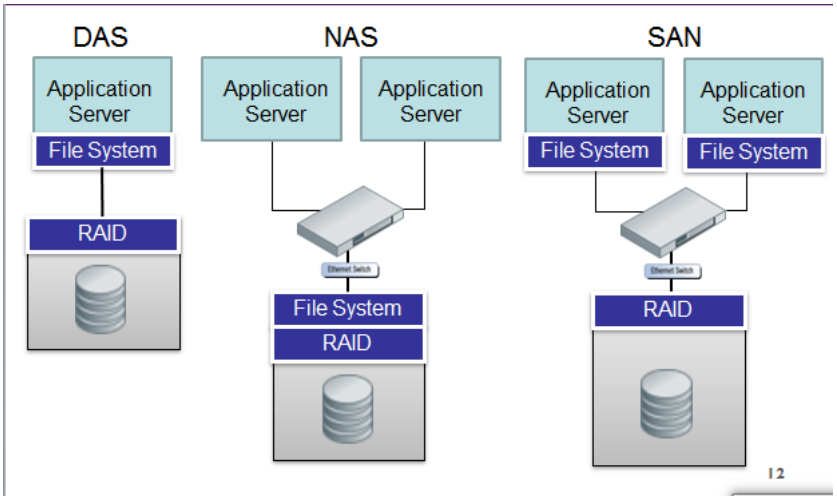

Figura 14, Arquitecturas de almacenamiento en Data Centers.

Fuente: www.Tech-Eye-Tech.com.

Es importante destacar que los arreglos SAN podrían eventualmente generar un punto único de falla. Para evitar esto, muchas veces los sistemas SAN se usan en conjunto con la arquitectura RAID, generalmente RAID 10, para proveer una mayor tolerancia a fallos, y/o mayor throughput, dependiendo de las necesidades.

La conexión entre servidores a dispositivos de almacenamiento se realiza mediante enlaces fiberChannel (fibra de canal), o iSCSI, y aunque la fibra de canal era sinónimo de alta velocidad, alcanzando tasas de transferencia de 4 Gbps a 16 Gbps, los últimos avances en Ethernet han ciertamente desplazado a la fibra de canal, ya que el protocolo iSCSI corre sobre la red local, por lo que alcanza las mismas velocidades a las que puede llegar Ethernet i.e. 1 Gbps, 10 Gbps, 40 Gbps, e incluso 100 Gbps (100GbE), lo que convierte a iSCSI en una solución más económica, flexible y rápida.[26]. Si bien implementar una infraestructura SAN es un poco más complejo y costoso, sus beneficios abarcan optimización del uso de recursos, facilidad de administración, y escalabilidad.

Los sistemas SAN usan diferentes niveles de almacenamiento para poder escalar, por ejemplo usando discos de acceso lento para almacenamiento a largo plazo (cintas, o HDDs), y discos de acceso rápido para alta disponibilidad o caching (SDD, o almacenamiento flash), esto permite equilibrar los costos totales de la instalación, habilidad de escalar y mantener un buen rendimiento.

SAN es una arquitectura de almacenamiento que beneficia directamente en la alta disponibilidad de las máquinas virtuales proveídas por los servidores, al permitir la portabilidad inmediata en el caso de falla de un servidor físico, moviendo el sistema operativo virtual de una máquina a otra con prácticamente 0 downtime, a través de sistemas de archivos distribuidos como VMFS [66].

Aunque los sistemas SAN son arquitecturas centralizadas para mejorar la administración de sus elementos, están netamente ligadas con la tecnología subyacente. Si bien tecnologías como Dell Compellent, Dell EMC Storage, NetApp FAS, etc, ofrecen sistemas SAN completos (es decir tanto servidores físicos dedicados a almacenamiento, como el software de administración), es complejo mantener sistemas de almacenamiento SAN de distintos proveedores. Una solución a esto es el almacenamiento definido por software.

[26]También se puede usar FCoE - FiberChannel over Ethernet, pero se necesita usar switches especiales que soporten este protocolo

3.6.5. SDS - Software Defined Storage

En la actualidad, existen diferentes métodos de virtualizar almacenamiento: A través del mismo hosts OS (LVM - Logical Volume Manager, LDM - Logical Disk Manager) o a través del mismo dispositivo de almacenamiento (por ejemplo un controlador RAID, o un controlador SAN). Sin embargo, cuando existen diferentes proveedores tanto de software como de hardware de almacenamiento, la administración se vuelve compleja, ya que no existe una interoperabilidad estándar entre todos estos componentes. Los data centers por lo tanto se han visto en la necesidad de contar con un sistema centralizado que permita manejar discos y unidades lógicas de almacenamiento y que sea independiente de la tecnología subyacente. Al igual que en SDN, en SDS se abstrae la parte física (data plane, dispositivos de I/O) de la parte lógica (control plane), con tal de administrar los sistemas de almacenamiento de forma centralizada a través de un controlador SDS, el cual debe ser capaz de gestionar almacenamiento heterogéneo, desde cintas, discos, tarjetas, independiente del fabricante, y/o controlador de RAID, con la capacidad de ofrecer almacenamiento por bloques, por objetos, añadir/remover/modificar unidades montables a servidores de forma inmediata, proveer múltiples puntos de control y monitoreo, entre otros.

Aunque no existe un controlador estándar, o una API de comunicación estándar entre data y control plane (como OpenFlow en SDN), existen tecnologías como IBM Spectrum capaces de ofrecer almacenamiento definido por software llegando a reducir los costos por almacenamiento en hasta 50 % [64].

3.7. Virtualización

La virtualización es el elemento más importante de la optimización de recursos de cómputo, cuyo objetivo es usar al máximo un servidor físico y así reducir la inversión en una mayor cantidad de éstos, y por lo tanto el costo total de la infraestructura, permitiendo lanzar múltiples servidores virtuales en 1 solo servidor físico.

Existen diferentes métodos de virtualizar sistemas operativos, como lo son la **paravirtualización**, en la cual el sistema operativo virtualizado puede acceder a los recursos del servidor hosts a través una capa de virtualización (hipervisor), la **virtualización**

completa, en la cual el servidor virtualizado no puede comunicarse con el servidor hosts, solo con la capa intermediaria de virtualización, y la **virtualización basada en contenedores**, la cual permite virtualizar recursos aislados compartiendo un único kernel de un sistema operativo.

3.7.1. Virtualización basada en hipervisores

Tanto la paravirtualización, como la virtualización completa, se logran a través de hipervisores. Un hipervisor es un software que se encarga de abstraer el servidor host (servidor físico) del servidor guest (máquina virtual), proporcionando una interfaz para que los servidores guests puedan interactuar entre ellos y puedan usar los recursos del servidor físico. El hipervisor se encarga de mantener servidores virtuales corriendo en paralelo y aislados, los cuales comparten recursos como CPU, memoria, y dispositivos I/O del servidor hosts. Se distinguen 2 tipos de hipervisores, los llamados tipo 2, que corresponden a software que corren encima de un sistema operativo *hosts*, y pueden virtualizar nuevos sistemas operativos *guests* (por ejemplo VMWare Workstation, o Oracle VM VirtualBox) y los hipervisores llamados tipo 1, o bare-metal, los cuales son los más populares dentro de los data centers, ya que se carga el hipervisor directamente sobre el hardware, por lo tanto al no tener que pasar por un sistema operativo intermedio, se elimina el overhead generado. Como el hipervisor de tipo 1 corre directamente en el hardware del servidor físico, ofrece un mejor rendimiento. El proceso de virtualización es similar para ambos tipos de hipervisores, por lo tanto es posible mover una imagen de una VMs entre hipervisores de tipo 1 y 2 de forma indistiguible.

Un data center al almacenar cientos de servidores, es probable que el manejo de cientos de hipervisores se haga un trabajo tedioso, es por eso que existen diversas *suites* de administración de hipervisores y máquinas virtuales que permiten el lanzamiento o remoción de una VM de un servidor físico de forma fácil y rápida. Las suites vSphere de VMWare y Hyper-V de Microsoft son las suites de virtualización líderes de acuerdo a la tabla Gartner 2016 [65].

Estas suites no solo proveen administración de las máquinas virtuales, si no que también se encargan de mantener alta disponibilidad, balanceo de carga, y optimizan la asignación de recursos entre los servidores.

Este tipo de virtualización genera un importante beneficio en la alta disponibilidad de los servicios, ya que sistemas como vSphere, o Microsoft Hyper-V, se basan en almacenamiento SAN para distribuir sus máquinas virtuales y ofrecer alta disponibilidad. De esta forma si un servidor tiene un corte inesperado, estas suites de virtualización se encargan de replicar todas las máquinas virtuales en otro servidor de forma instantánea. Estas suites también usan software como VMWare DRS (Distributed Resource Scheduler) para distribuir las cargas entre los servidores virtuales.

Antiguamente, el virtualizar un sistema operativo, generaba pérdidas de rendimiento, ya que la abstracción del sistema host a través de un hipervisor generaba un *overhead* suficiente como para afectar al rendimiento. Sin embargo, avances en los hipervisores, han logrado que el rendimiento de sistemas virtualizados sea muy similar al sistema hosts, con la ventaja de poder desplegar cientos de máquinas virtuales en 1 solo servidor físico, con una penalización mínima de rendimiento. Si se tiene en cuenta que el tamaño de una máquina virtual (un sistema operativo completo por cada VM) es de aproximadamente 4 GBs (agregando librerías, módulos, paquetes), y considerando tiempos de lanzar una instancia, actualización de cada VM en particular y otros elementos difíciles de manejar en cientos de VMs simultáneamente, el acercamiento de virtualización por hipervisores ya no es suficiente para manejar servidores de forma distribuida (sin embargo en la sección 8 veremos que existen software encargados del manejo de servidores masivos). La virtualización basada en containers es una tecnología que intenta manejar éstos, y otros problemas.

3.7.2. Virtualización basada en containers

En la virtualización basada en hipervisores, los sistemas guests corren encima de un hipervisor. Este hipervisor se encarga de controlar y manejar el uso de procesamiento, memoria, y almacenamiento. La virtualización basada en containers usa otro acercamiento, en el que la virtualización es llevada a cabo sobre un sistema operativo que corre directamente sobre un servidor físico (sistema operativo host), aislando *mini* máquinas virtuales en un único kernel. Estas *mini* máquinas virtuales en realidad son procesos conocidos como contenedores.

En la figura 15, puede apreciarse de forma general, una diferencia entre ambos tipos de virtualización.

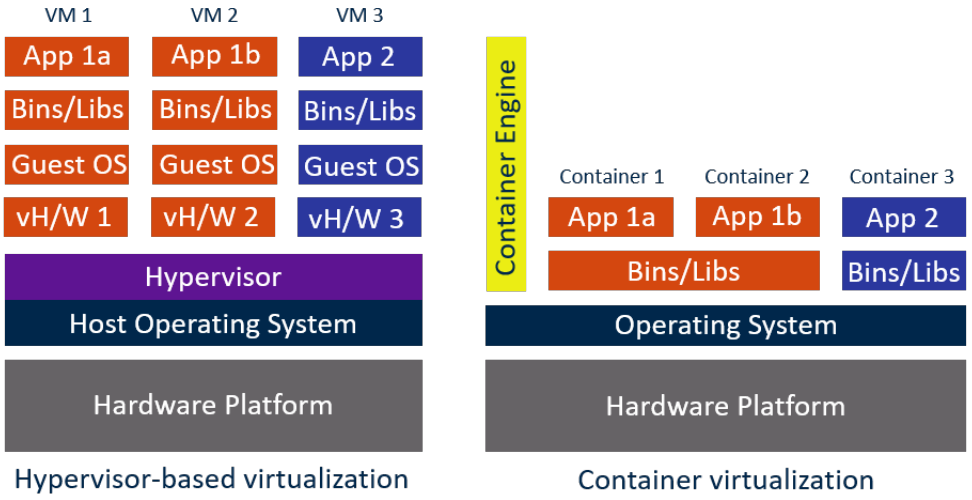

FIGURA 15, DIFERENCIA ENTRE VIRTUALIZACIÓN A TRAVÉS DE HIPERVISORES Y CONTENEDORES, FUENTE: WWW.PEARSONITCERTIFICATION.COM

Actualmente Dockers es el motor de contenedores más popular en la industria, siendo su competidor directo RKT. Está sección profundizará brevemente el uso de Dockers, ya que es una de las tecnologías más potentes en términos de optimización de recursos, ofrecimiento de alta disponibilidad, desarrollo y entrega continua, administración de recursos IT, además de ser ampliamente usado en la industria de data centers.

Para comprender la virtualización basada en containers, primero es necesario definir 2 elementos:

- **Imagen**: Una imagen es un conjunto de recursos: librerías, códigos, archivos binarios, ejecutables y variables de entorno, que permiten el funcionamiento correcto de alguna aplicación. Pueden ser extremadamente livianas (un par de KBs) o tan grandes como un SO completo (por ejemplo una distribución de Linux). Estas imágenes pueden estar almacenadas dentro del mismo servidor donde serán ejecutadas, en un servidor externo, o en un repositorio externo (Dockers Hub, Amazon Containers Registry, Google Container Registry). Esta imagen no debe ser confundida con un OS completo, ya que no existe un kernel, ni módulos de kernel (excepto si la imagen corresponde efectivamente a un SO).

- **Contenedor**: Instancia en ejecución de una imagen. Se habla de virtualización, ya que los contenedores corren completamente aislados unos de otros, y aislados del sistema hosts, a menos que se especifique que se quiere acceder a ciertos recursos

(por ejemplo algún puerto en especifico, fichero dentro del sistema de archivos, o cualquier otro recurso en particular). En [63] puede encontrarse una forma más detallada del proceso de aislación de procesos, archivos, dispositivos, redes y más, en contenedores.

Dockers ofrece una ventaja indiscutible al desplegar aplicaciones de forma rápida y escalables, ya que permite replicar ambientes de desarrollo, prueba y producción en segundos, pudiendo aumentar o disminuir la cantidad de contenedores a medida que sean necesitados. Los contenedores tienen un impacto directo en ambientes de producción, ya que eliminan por completo el riesgo de que los diferentes ambientes del ciclo de producción sean distintos[27]. Aunque dockers no tiene incluida una funcionalidad que permita manejar múltiples contenedores de forma distribuida, tecnologías como Kubernetes [61] o Docker Swarm [62], proveen un sistema de orquestación de containers, permitiendo crear literalmente Data Centers de contenedores con tal de administrar cientos de contenedores, ejecutándolos, deteniéndolos, modificándolos con tal de maximizar el rendimiento, escalabilidad y alta disponibilidad de software en producción. Los contenedores tienen una vital importancia en el ciclo de desarrollo de software conocido como Integración y Entrega continua, los cuales serán revisados en la sección 10: Capa de aplicación.

3.8. Administración

En un data center se deben mantener y administrar cientos de recursos incluyendo servidores, máquinas virtuales, procesos, subredes, etc. Es necesario tener una forma centralizada y/o programable de manejar todos estos recursos, de forma de mantenerlos controlados, y optimizar su uso, evitando su desperdicio. Al igual que SDN que provee una interfaz centralizada y programable para manejar las redes dentro de un data center, se hacen necesario herramientas de administración que permitan manejar la mayor cantidad de recursos, también de forma centralizada y flexible. Para esto existen diferentes acercamientos de administración, los cuales se abordarán en esta sección.

Además de las arquitecturas de administración de data centers, existen herramientas que permiten un nivel de monitoreo más especifico, como por ejemplo herramientas

[27] *It Worked on my machine* - Funcionaba en mi máquina, es una frase popular en el área de desarrollo de software, la cual refleja que los errores en producción se deben ocasionalmente a las diferencias de infraestructura en los ambientes de desarrollo y producción

de monitoreo de redes (PRTG), uso de memoria, CPU, uso de disco (NAGIOS). Sin embargo, no fueron abarcadas en esta sección, ya que se busca profundizar arquitecturas de administración de recursos a nivel macro de un data center.

3.8.1. Configuration Management

Antiguamente la mejor forma de configurar distintos servidores era crear, distribuir y ejecutar scripts en ellos. Sin embargo, cuando se trabajan servidores con distintas características, ya sea de hardware (diferentes arquitecturas del procesador, memoria y almacenamiento disponible, interfaces de red) o software (diferentes sistemas operativos, módulos instalados, servicios ejecutándose) es complejo mantener una infraestructura distribuida de forma sincronizada y que se mantenga operativa. Se hacen necesarios sistemas de orquestación que permitan configurar un gran número de servidores, independiente de su arquitectura física o lógica, de forma fácil, rápida y entendible para cualquier administrador de sistemas.

Las herramientas de gestión de configuraciones son altamente usadas cuando se habla de administración de servidores en data centers, y son usadas principalmente para modificar archivos, instalar paquetes, agregar repositorios, inicializar variables de entorno, manejar servicios y otros, entre múltiples servidores.

También es posible provisionar recursos con este tipo de herramientas (como lanzamiento de máquinas virtuales preconfiguradas), sin embargo no es su función principal. Para este tipo de administración se usan las nubes privadas.

Dentro de sofware de administración de configuraciones, los productos más populares son: Chef [53], Ansible [54], SaltStack [55] y Puppet [56].

Una desventaja importante que tienen estas herramientas, es que al necesitar paquetes, o aplicaciones ubicadas en algún repositorio externo, existe la probabilidad de que esos repositorios no estén disponibles al momento de ejecutar una instrucción de el gestor de configuraciones, pudiendo generar impactos en ambientes de producción. Este tipo de inconvenientes pueden ser tratados a través de contenedores, por ejemplo Dockers. Como se mencionó en la sección anterior, la ventaja que tienen los contenedores, sobre ambientes que deben ser aún configurados, es que un servidor virtualizado puede tardar minutos o hasta horas en estar operativo nuevamente, mientras que los contenedores

al ser extremadamente livianos, pueden ser lanzados instantáneamente, disminuyendo considerablemente los impactos a las operaciones.

Los usos de herramientas de administración de configuraciones pueden mejorar considerablemente los procesos, ya que además de automatizar toda la gestión de la infraestructura, se obtiene mayor visibilidad, permitiendo monitorear fallas y cortes no planeados en ambientes de producción, permitiendo actuar de forma rápida.

3.8.2. Nubes privadas

El término Cloud Computing hace referencia al uso de recursos informáticos a través de acceso remoto, independiente de la arquitectura subyacente. La diferencia entre una nube pública y privada está básicamente definida por el nivel de acceso que se tiene a los recursos. Una nube pública consta de recursos accesibles a través de un data center externo, que se encarga de prestar sus servicios de forma on-Demand. Se profundiza esta arquitectura en el capítulo 4, modelo de capas de Cloud Computing. Una nube privada define recursos en un data center interno de una empresa o negocio con tal de facilitar la administración y acceso a sus recursos de forma fácil, permitiendo flexibilidad y elasticidad. Un tercer tipo de nube a veces se conoce como nube híbrida, la cual es una combinación de nubes privadas y públicas, trabajando en paralelo de forma sincronizada. Este tipo de nubes híbridas ofrece la seguridad de una nube privada, y la escalabilidad de una nube pública.

Las nubes privadas ofrecen servicios similares a las herramientas de administración de configuraciones, sin embargo tienen objetivos distintos. Una nube privada está enfocada en provisionar recursos y mantener control sobre ellos, por ejemplo máquinas virtuales, subredes, grupos de seguridad, aunque también permiten personalizar y configurar recursos para ser lanzados de forma automática (por ejemplo crear una máquina virtual preconfigurada que deba ser lanzada para balancear carga de procesamiento), por lo tanto las nubes privadas también optimizan el flujo de los procesos IT.

Las nubes privadas ofrecen un mayor nivel de seguridad a una nube pública al estar aisladas, permiten personalización directa de los recursos, además de menores tiempos de respuesta al estar ubicados dentro de la red interna.

Ejemplos de software para implementar nubes privadas son: OpenStack [57], Apache CloudStack [58] y VMWare vCLoud [59]. Sin embargo existen nubes públicas que permiten un nivel de seguridad y aislación que simulan nubes privadas a través de Internet, por ejemplo AWS y GCP.

3.8.3. SDDC - Software Defined Data Center

Los Data Centers definidos por software son la cúspide de la tecnología actual respecto al manejo de infraestructura IT. Mezclan todo el potencial de una nube privada, la flexibilidad de las redes definidas por software, además de una interfaz que permita darle programabilidad al Data Center con tal de convertirlo en una infraestructura completamente elástica. Esta arquitectura ofrece la mayor maximización de uso de recursos, ya que permite lanzar/eliminar máquinas virtuales, redes virtuales y sistemas de almacenamiento entre otros, de forma completamente automática y programable, cambiando la topología del data center de forma dinámica y flexible de acuerdo a las necesidades del momento.

Se pueden implementar SDDC a partir de un software que permita diseñar nubes privadas, por ejemplo usando OpenStack [52]. Aunque avances en las tecnologías de las nubes públicas, han permitido desplegar cientos de recursos IT con un par de clicks, por ejemplo a través de Amazon CloudFormation o Terraform[28]. Estas herramientas permiten lanzar un Data Center completamente virtualizado, y replicable en un par de segundos.

Una de las ventajas principales de estas tecnologías es que convierten toda la infraestructura IT, directamente en código, por lo que dejan de depender completamente de cualquier proveedor de nube pública o privada, ya que son replicables en cualquier infraestructura. Permite una mayor tolerancia a fallas al permitir replicar la infraestructura en cualquier data center del mundo. También disminuye directamente los costos asociados a infraestructura, ya que se usan solamente recursos necesitados.

Las herramientas de administración de data center generan un impacto directo en las aplicaciones en ambientes de producción, ya que estas dependen directamente de la infraestructura subyacente para poder funcionar correctamente sin interrupciones.

[28] Estas herramientas constituyen un nuevo grupo de tecnologías conocidas como Infraestructura como Código - IaC.

Al agregar/eliminar funcionalidades, o corregir errores, deben modificarse tanto infraestructura como código fuente de la aplicación, apuntando a 0 downtime en el proceso. La intersección entre estas 2 áreas, desarrollo y operaciones, es conocida coloquialmente como DevOps, la cual busca lanzar software y modificarlo en producción sin afectar a su rendimiento y/o uptime. Las herramientas de administración de infraestructura son el elemento principal a la hora de lanzar aplicaciones en producción, sin embargo también es necesario contar con un método que permita desarrollar código y distribuirlo de forma óptima en esta infraestructura. En la última capa del modelo de data centers, se revisan los 2 métodos principales para desplegar código en servidores distribuidos, sin impactar las operaciones de la organización.

3.9. Aplicaciones

Un Data center puede ofrecer múltiples servicios a través de aplicaciones en las que se encuentran servicios web, análisis de datos, minería de datos, provisión de máquinas virtuales, hosting, etc. Agregar servicios en ambientes de producción de forma continua, sin impactar el rendimiento de las aplicaciones, es uno de los desafíos más importantes que deben manejar los data centers, ya que la confiabilidad en el servicio depende netamente del rendimiento, tiempos de respuesta, y uptimes de las aplicaciones.

Cuando se tienen múltiples servidores distribuidos corriendo una aplicación a la que se debe realizar un cambio, es importante distribuir el código en todas las máquinas que corren la aplicación de tal forma que no impacte ni el rendimiento, ni se produzcan cortes en el proceso.

Para lograrlo de forma eficiente, existen distintas prácticas que permiten llevar aplicaciones de ambientes de prueba a producción. Esta sección se enfoca en 2 importantes áreas de la ingeniería de software, que corresponden a entrega e integración continua.

3.9.1. Continuous Integration - Integración Continua

Cuando se cuenta con un equipo de múltiples desarrolladores en un proyecto, es probable que cada uno trabaje en alguna funcionalidad o arreglo en particular. Para no estropear el código principal, este se bifurca en ramas, en las cuales cada desarrollador trabaja en algún cambio. Estos cambios deben unirse eventualmente con la rama principal, sin

embargo como es probable que una rama del desarrollo tenga varias discrepancias con la rama principal, pueden producirse errores de ejecución o de funcionalidad. La integración continua es una estrategia que asegura que los cambios de todos los desarrolladores se integren correctamente a la rama maestra, y puedan enviarse al entorno de producción sin errores.

La integración continua (CI) es una práctica del desarrollo de software, y un área en particular de la entrega continua (CD), que básicamente consiste en agregar pequeños cambios a una aplicación de forma gradual, y enviarlo desde un ambiente de prueba al repositorio principal, sin errores de desarrollo o infraestructura. Si dicho cambio genera problemas en la aplicación o necesita ser actualizado o revisado, es posible aislarlo y removerlo sin afectar al resto del código, construyendo la aplicación de forma modularizada y continua, eliminando riesgos de fallas en producción.

La implementación estándar del proceso de CI, consta de un equipo de desarrollo, un sistema de control de versiones (Por ejemplo Git o Subversion) asociado a un repositorio, y un servidor dedicado encargado de ejecutar las pruebas de forma automática a través de algún software (Por ejemplo Jenkins).

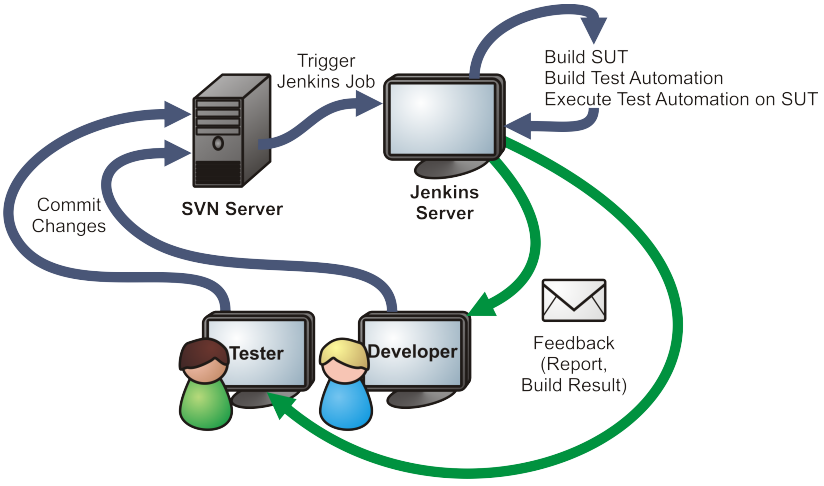

Figura 16, Flujo de trabajo de Integración Continua.
Fuente: www.seleniumframework.com.

Las pruebas son definidas en las primeras fases (antes de comenzar a desarrollar la aplicación) y son cargados en el software de CI. Una vez que los desarrolladores comienzan a enviar nuevas funcionalidades (o corrección de errores) al repositorio maestro, este repositorio se encarga de enviar los cambios directamente al servidor de CI, el cual está

encargado de realizar las pruebas, notificando directamente a los desarrolladores si los tests fueron exitosos o no. Estas pruebas pueden variar desde probar la correcta compilación/ejecución del código, hasta validar que las funcionalidades están funcionando de forma esperada. Cuando las pruebas son finalmente aprobados por los desarrolladores, son enviados nuevamente al repositorio, pero esta vez para ser anexado finalmente a la rama principal.

Este proceso se realiza de forma iterativa, (es decir, se va trabajando sobre la misma funcionalidad hasta que las pruebas entreguen un resultado exitoso), y de forma frecuente (llegando a poder enviarse cambios cada hora), con tal de disminuir las fallas y riesgos cuando los cambios pasen a producción.

La integración continua es una etapa de un proceso mayor, conocido como Entrega Continua (CD), la cual busca agregar los cambios revisados a producción, sin impactar el rendimiento de la aplicación.

3.9.2. Continuous Deployment - Entrega Continua

Los métodos ágiles del desarrollo de software (como SCRUM, Lean programming o eXtreme Programming) consisten en entregar aplicaciones de forma iterativa, libre de errores, y con la capacidad de adaptarse a las necesidades de los usuarios finales, en pequeños intervalos de tiempo. La entrega continua es uno de los principales elementos dentro de la cultura DevOps, que asegura la continuidad de las operaciones en el desarrollo de software sin impactar ambientes de producción.

Esta práctica automatiza completamente el proceso de publicación de software, ya que los procesos de pruebas y revisión de errores, y envío a producción, es realizado por software de manera automática, y no manual[29]. La entrega continua por lo tanto, facilita enormemente el paso a producción, haciendo el proceso más rápido, fácil y automático, además de eliminar el factor del error humano.

[29] Cuando estos cambios se envían a producción y son aprobados manualmente por un desarrollador, y no de forma automática, al proceso se le conoce como Continuous Delivery, sin embargo la idea general del proceso se mantiene.

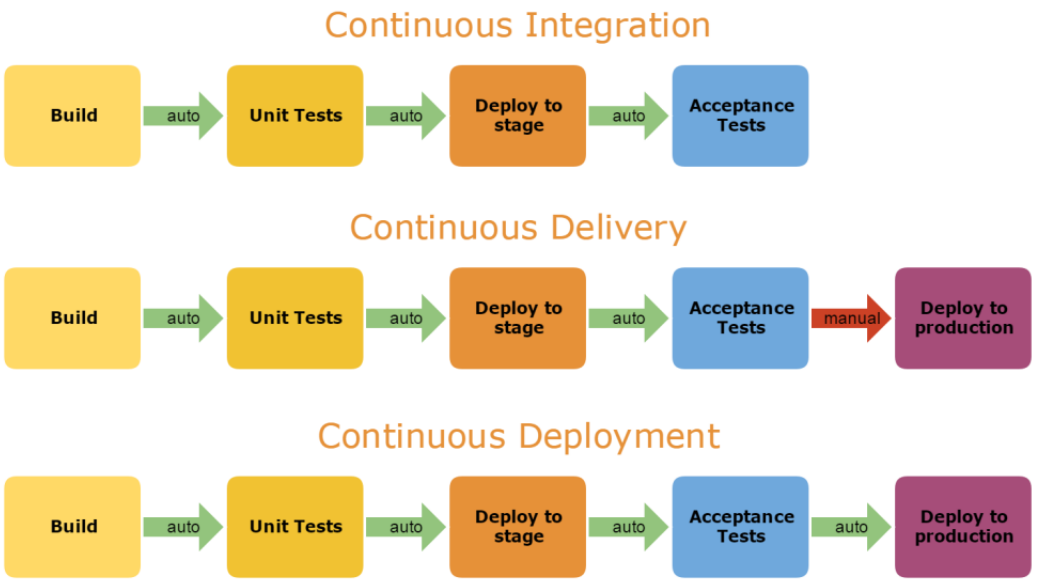

FIGURA 17, ETAPAS DE CI Y CD.

FUENTE: WWW.CODEPROJECT.COM.

Este proceso también puede ser realizado por herramientas de CI, ya que estas herramientas apuntan tanto a la integración continua como a la entrega continua.

Un ejemplo de distribución de código en múltiples data centers para manejar entrega continua puede revisarse en [60], agrupando data centers y servidores en *pools*, usando un servidor único de control como orquestador.

Sin embargo una tecnología que ha sabido abrirse paso en el área de entrega continua, son los contenedores. Como se mencionó en la sección 3.7.1, los contenedores son pequeños y livianos procesos conformados por componentes necesarios para desarrollar, testear y poner aplicaciones en producción. Las herramientas de CI se han adaptado para aceptar dockers como una tecnología estándar, y permiten realizar tests no solo sobre aplicaciones, si no que sobre completos contenedores.

3.10. Resumen del modelo

Se ha presentado en este capítulo, un modelo de capas que describe de forma general los componentes que conforman un data center, con tal de ofrecer al lector una idea de lo complejo que puede resultar su implementación en términos de costos, administración y orquestación.

En la figura 18, puede observarse una tabla que contiene el mismo modelo presentado en la figura 3, sin embargo se detallan las arquitecturas presentes en cada capa.

		Capa	Subcapa	Arquitecturas		Ejemplos de implementación
Estándares	Monitoreo y control	Facility	Planeación	On-Premise		Data Center USM Informática.
				Colocation		Digital Realty, Switch, Entel, Sonda
				Hybrid Cloud		AWS, Azure, GCP
			Infraestructura	• Cooling	CRAC	Liebert
					CRAH	Liebert, Johnson Controls
					Free Cooling	Gesab
				• Power	TIA 942 UPS Systems	Liebert, APC
			Seguridad	• Física	BMS (CCTV, MFA, Fire detection)	Swann Systems, EmpowerID
					NOC, DCIM	StruxureWare
				• Operaciones	Lifecycle MGT	Asset Explorer
					Firewalls	F5 WAF.
		IT	Red	Cisco 3-Tier-Network		Cisco, Juniper, Huawei, OCP Switches
				Leaf & Spine		Cisco, Juniper, Huawei, OCP Switches
				Software Defined Networking		OpenFlow, FloodLight
			Almacenamiento	RAID		Intel RAID Controller.
				NAS		Dell Compellent, Dell EMC2
				SAN		Dell Compellent, Dell EMC2
				Software Defined Storage		IBM Spectrum
			Virtualización	Hypervisor based		vSphere, Hyper-V, RHEV
				Container based		Dockers, RKT
		Operaciones	Administración	CM		Puppet,Chef,Ansible,SaltStack
				Private Cloud		OpenStack, RackSpace
				SDDC		OpenStack, RackSpace
			Aplicación	Continuous Integration		Jenkins, AWS CodePipeline
				Continuous Delivery		Jenkins, AWS CodePipeline

Figura 18, Modelo de capas de data center según arquitecturas y/o diseños de implementación, Fuente: Elaboración propia.

Dentro de la capa de **Facility**, se puede concluir que el punto más importante dentro de la implementación, es el uso de la eficiencia energética. En el caso del enfriamiento, Free Cooling es la arquitectura más eficiente al eliminar completamente los componentes del enfriamiento por manejo de aire, ya sean chillers, torres de enfriamiento, bombas de agua, unidades UPS dedicadas. En el caso de la energía, se puede observar que un indicador esencial para medir eficiencia es el PUE (Power Efficiency Usage), donde PUEs cercanos a 1 indican que de toda la energía suministrada al data center, se está usando el 100 % para hacer funcionar el equipamiento IT. Aunque en la práctica esto no es posible, este indicador muestra la cantidad de energía desperdiciada en transformaciones, pasos de un componente a otro, o pobre manejo del sistema energético, lo cual no solo produce una huella de carbono más alta, si no que genera un gran desperdicio monetario a la empresa.

Dentro de la capa de **IT**, se concluye que nuevamente la optimización de los recursos es el núcleo en las salas de datos, el cual se logra a través de la virtualización. La

virtualización de recursos, tanto de redes, almacenamiento y sistemas operativos, permite hacer un mayor uso de los componentes IT, tratando de desperdiciar lo menos posible los recursos disponibles. Surge también el paradigma de *Arquitecturas Definidas por Software*, el cual busca separar las capas lógicas de las físicas de los componentes IT, con tal de hacerlos independientes de fabricantes o proveedores, de forma que puedan comunicarse entre ellos a través de un protocolo común (por ejemplo OpenFlow en SDN), permitiendo virtualizar componentes, agregarles programabilidad y ofrecer una administración centralizada de ellos.

Finalmente la capa de **Operaciones** se centra netamente en la agilidad y flexibilidad de desplegar software en servidores distribuidos o data centers, sin impactar las operaciones. Esto se logra a partir de una fuerte orquestación de los componentes IT, y técnicas que permitan automatizar los procesos de testeo y deployment.

Capítulo 4

Caso de Estudio - Departamento de Informática

La Unidad de Servicios Computacionales e Internet (uSCI) del Departamento de Informática de la UTFSM, está encargada de proveer servicios de cómputo, almacenamiento, hosting y housing entre otros, tanto al Departamento de Informática (ya sean estudiantes, profesores o personal administrativo), como al público general externo a la Universidad. En la uSCI se encuentra ubicado el Data Center del Departamento de Informática, el cual será utilizado como caso de estudio con tal de localizar el modelo de capas de data center presentado en el capítulo 3, con el objetivo de analizar las arquitecturas presentes y proponer mejoras tanto de infraestructura, optimización de recursos, y disminución de costos con el fin de proveer al lector un caso práctico del modelo presentado.

4.1. Estándares

El data center sigue las normas mínimas establecidas por ASHRAE y NPFA sobre temperaturas, humedad y protección contra incendios (usando extinción mediante químicos ecológicos).

El data center sigue algunos elementos de los estándares populares que definen recomendaciones y buenas prácticas como cableado estructurado y diseños de redundancia, sin embargo para fomentar la escalabilidad, se debe definir una arquitectura modular como

lo establece el estándar TIA-942, además de diseños de redundancia equivalentes a Tier 2 o 3.

El data center de Informática no posee ninguna certificación a través de instituciones como Uptime Institute, o SAS, las cuales están orientadas a organizaciones que requieren un alto nivel de *uptime*, o que trabajan con información altamente sensible (como HIPAA).

4.2. Planeación

El data center sigue una arquitectura del tipo On-Premise, es decir está ubicado directamente en las instalaciones de la universidad, sin usar un centro de colocación, o servicios en nubes públicas. Actualmente provee servicios hacia alumnos y profesores como almacenamiento, procesamiento y hosting de proyectos de alumnos, memoristas, e investigadores, por lo que se está planeando escalar el data center. La siguiente tabla describe los costos aproximados de levantar una infraestructura de data center como la que existe en el Departamento de Informática.

Item	USD
Obras Civiles	5.000 - 10.000
Sistema climatización	50.000 - 60.000
Sistema detección y control de incendios	20.000 - 30.000
Sistema de energía, cableado, racks.	40.000 - 50.000
Sistema de monitoreo (NOC)	5.000 - 10.000
Componentes IT	200.000 - 250.000
Recursos Humanos/mes	5.000 - 10.000

Tabla 1, Costos estimados de una infraestructura similar a la analizada. Fuente: Presupuesto proyecto Data Denter Departamento Informática UTFSM Casa Central

Estos datos serán utilizados nuevamente en la sección 5.3.3 donde se presenta a una nube pública como una extensión de la plataforma de virtualización actual.

4.3. Seguridad

El data center presenta sistemas de CCTV, identificación por contraseña, puertas de seguridad. No se detalla mayor información debido a la confidencialidad de los elementos de seguridad existentes.

4.4. Infraestructura

El data center cuenta con 1 unidad UPS (además de una línea de bypass para realizar mantenciones a la UPS) y con 1 unidad manejadora de aire para climatización del tipo CRAC. No existe redundancia ni duplicación para ninguno de estos componentes, por lo que en la práctica, y de acuerdo al estándar TIA-942 es un Data Center considerado Tier I.

La unidad UPS en promedio se usa al 41 % de su potencia total, y en promedio provee una energía de 5 KWh al data center[1]. Si se considera que el costo de 1 KWh en Chile es de $0.1313 USD (∼$85 CLP) de acuerdo a las tarifas actuales de Chilquinta para empresas que trabajen en alta tensión [68][2], el equipamiento IT genera un gasto mensual alrededor de $472 USD mensual respecto a gasto energético. Sin embargo, existen otros componentes que también deben ser energizados para que el data center pueda funcionar en su totalidad, por ejemplo sistema CCTV, estaciones de trabajo de administración, luz de la sala, sistema de respaldo Bacula, unidad manejadora de aire, etc.

PUE - Power Usage effectiveness

Aunque el concepto de PUE (Power Usage Efectiveness) ya fue introducido en el capítulo 2, es necesario profundizar acá su importancia, ya que indica la eficiencia del uso energético del data center.

En un data center se identifican 2 sectores de consumo energético, la **carga crítica** la cual corresponde a la energía usada netamente por los componentes IT (servidores, switches, routers, firewalls, almacenamiento, respaldo, y estaciones de administración) y

[1] Información entregada por la unidad UPS.

[2] En chile se definen distintas tarifas dependiendo del empalme de redes al que se esté conectado, considerándose alta tensión un cliente conectado a un empalme mayor a 400V. Generalmente estos clientes tienen un cobro por KWh menor al de clientes que trabajan en baja tensión [71]. Se elige por lo tanto en este estudio el menor valor que podría generarse por costo energético, por lo que el costo total estimado es *a lo menos* el obtenido.

la **carga de infraestructura**, la cual está compuesta de la carga crítica más la energía usada en los sistemas de enfriamiento, seguridad perimetral (CCTV), protección contra incendios, luz de la sala, etc. El PUE es el resultado de la carga de infraestructura sobre la carga crítica (Figura 19).[3]

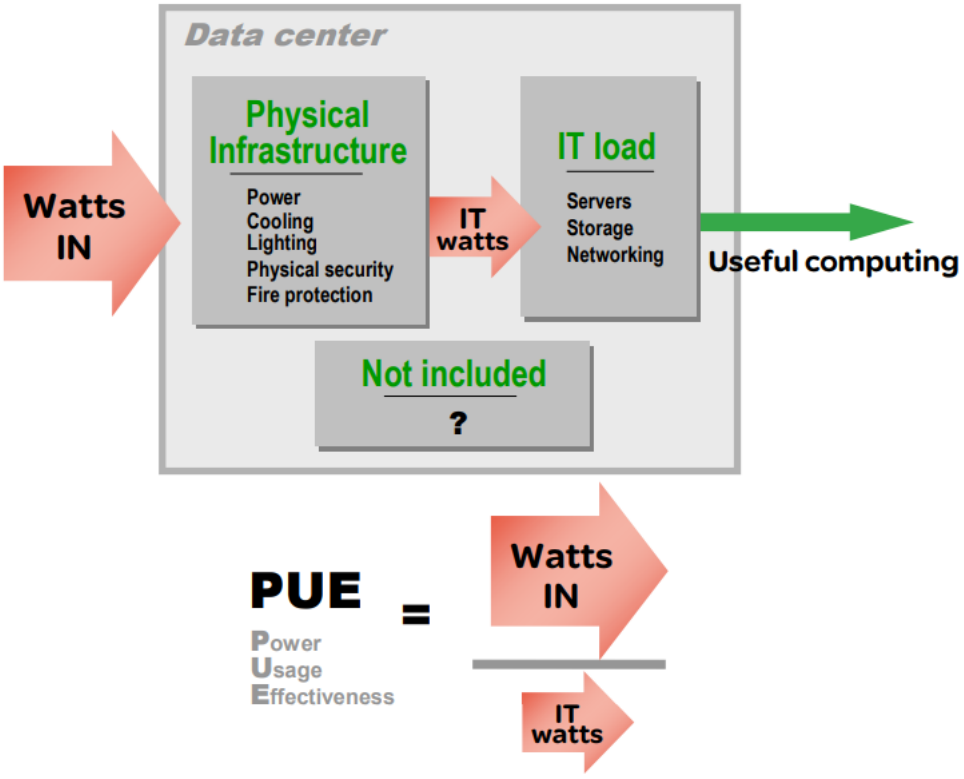

FIGURA 19, DESCRIPCIÓN GRÁFICA DEL PUE.

FUENTE: HTTPS://WWW.APC.COM.

La manejadora de aire se estima que consume un total de 4KWh, y se estima entre la suma de los otros componentes (NOC, luz, sistema CCTV, etc.) un consumo promedio de 1KWh, por lo que en este caso, el data center mantiene un PUE de 2.

$$PUE = \frac{5KWh + 4KWh + 1KWh}{5KWh} = 2 \tag{4.1}$$

En este caso, 2 representa una eficiencia promedio para un data center de pequeña/mediana escala de acuerdo a [70]. Esto significa que cada componente IT que se agregue al data center, debe calcularse que su consumo energético será su consumo promedio

[3]Generalmente la carga crítica se obtiene a nivel de UPS, y la carga de infraestructura se obtiene a nivel del medidor de energía del edificio, sin embargo como el medidor mezcla más componentes que solamente el data center, en este caso debe ser estimado.

multiplicado *al menos* por 2, por lo que el data center genera un costo mensual de $945 USD. Estos valores serán usados en el próximo capítulo, donde se realiza un análisis de costos On-Premise vs Cloud.

Por otro lado, la unidad manejadora de aire, sigue los estándares de la ASHRAE respecto a humedad y temperatura, y separando el data center en un diseño pasillo frío / pasillo caliente, con temperaturas entre 20°C y 25°C. Estas temperaturas no son las temperaturas óptimas para cada pasillo, generando un esfuerzo extra en la unidad manejadora de aire, por lo tanto consume más energía eléctrica que la que debería[4].

4.5. Red

La arquitectura de red que mantiene el DC es una arquitectura 3-Tier, con 3 switchs Top-Of-Rack (switches de acceso), conectándose a un switch de distribución. Este switch de distribución está conectado de forma redundante a los core switches de DTI - Dirección de Tecnologías de la Información de la USM. Además, los switches ToR están conectados a un switch especial para conectividad entre el Data Center y el resto del Departamento de Informática. En la figura 20 puede apreciarse la topología que actualmente presenta la arquitectura de red del data center.

[4]En [72] puede encontrarse una calculadora que entrega las temperaturas óptimas dependiendo de las temperaturas de entrada y salida tanto de los racks como de la sala entera.

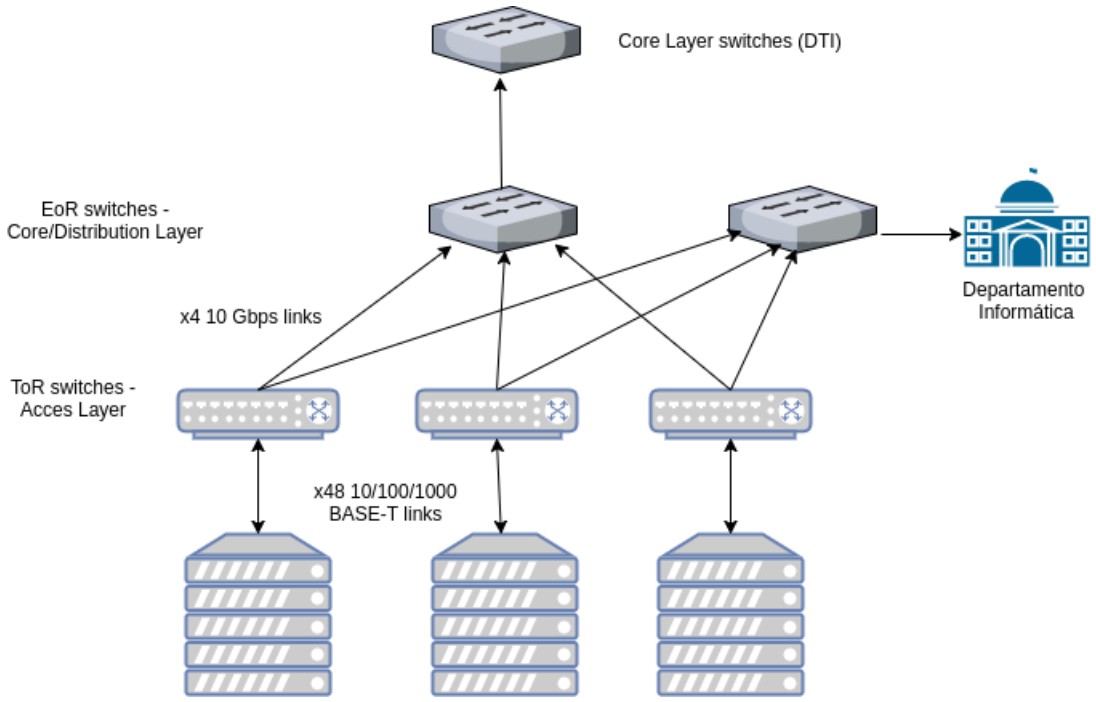

Figura 20, Diagrama de red Data Center.

Fuente: Elaboración propia.

Para aplicaciones que se basan fuertemente en tráfico servidor-a-servidor (bases de datos distribuidas como Hadoop, HPC, o almacenamiento NAS/SAN distribuido para proveer alta disponibilidad), este acercamiento ya no es el óptimo, y se prefieren arquitecturas como Leaf-Spine, sin embargo el costo de estos switchs es bastante elevado, ya que incorporan puertos que soportan mayor ancho de banda (Hasta 10 Gbps).

4.6. Almacenamiento

El data center usa el sistema de almacenamiento de RHEV, para esto utiliza 2 servidores distintos:

- EMC VNXe3150, 24x900GB 10K SAS
- PowerVault MD1000, 20x600GB 10k SAS y 4x400GB SSD SAS

Ambos optimizados para ser usados como servidores SAN y/o NAS, y en distintas configuraciones de RAID. Estos servidores están conectados directamente a los switches Top-Of-Rack.

Como sistema de recuperación y respaldo, se usa la herramienta Bacula [67], el cual es un servidor dedicado específicamente al respaldo (con ciclos de respaldo diarios). Este sistema de respaldo está ubicado en otro sector del mismo edificio del data center, detrás de una unidad UPS independiente.

4.7. Virtualización

El Data center cuenta con un total de 11 servidores usados para virtualización y 11 para propósito general.

El data center usa Red Hat Virtualization Enterprise (RHEV) como su sistema de virtualización principal. Cada hipervisor corre simultáneamente 16 VMs en promedio, usando principalmente servidores Dell Poweredge. Aunque es posible provisionar más de 20 o 30 VMs por servidor host, esto no es recomendando, ya que aunque no se sobrecargue la memoria RAM o la CPU, se pueden apreciar fuertes degradaciones en los acceso a los discos de almacenamiento.

4.8. Administración

Para administrar los recursos principalmente se usa el gestor de la suite RHEV. También se usan herramientas de gestión de configuración, principalmente Puppet y Chef.

4.9. Aplicación

El data center no está orientado al desarrollo, si no que más bien en servicios para los alumnos (como el almacenamiento de datos), y provisión de máquinas virtuales, así como también hosting y housing a externos, por lo que en este DC no aplica el uso extenso de integración/entrega continua.

Capítulo 5

Propuestas de Mejoramiento

En el capítulo 4 se describe como se ajusta el DC del Departamento de Informática con el modelo de capas del capítulo 3. En este capítulo se describen recomendaciones y mejoras con tal de escalar de forma costo-eficiente, y una mejor administración de los recursos del data center de uSCI. Estas propuestas están separadas en 3 etapas, dando prioridad a la eficiencia energética, alta disponibilidad y agilidad/escalabilidad respectivamente.

5.1. Etapa 1 - Mejoras de Infraestructura (Eficiencia).

5.1.1. De los estándares

El DC al no contar con elementos redundantes de infraestructura, está compuesto de múltiples puntos únicos de falla. Esto significa que un problema en la unidad UPS o en la unidad de climatización, podrían generar cortes prolongados de los servicios proveídos. Sin embargo una inversión en renovar la arquitectura energética o de climatización serían bastante elevados, y quizás exagerados para el tamaño del data center, ya que en estos momentos no se espera una completa tolerancia a fallos. Sin embargo, si se desea escalar la infraestructura actual, es necesario guiarse por los lineamientos del estándar TIA-942, la cual define los distintos niveles o *Tiers* de arquitecturas redundantes y componentes duplicados, además de establecer una arquitectura modular de data centers. No son necesarias certificaciones por ejemplo a través de Uptime Institute, SSAE 16, u otras

entidades certificadoras, ya que están orientadas principalmente a clientes que necesiten mantener un alto porcentaje de *uptime* anual o cumplir con restricciones como HIPAA[1].

5.1.2. De consumo energético

La unidad UPS en promedio tiene una tasa de consumo del 41 %, es decir, es posible agregar nuevos servidores y/o switches hasta el doble de la capacidad actual, y el UPS podría soportar dicha carga sin problemas. Sin embargo si se desea agregar más capacidad, es necesario adquirir una nueva unidad UPS.

El agregar una nueva unidad UPS al sistema actual generaría una cierta cantidad de beneficios, dentro de los más importantes se encuentran soportar una mayor carga crítica IT si se fueran a adquirir nuevos servidores y aumento en la disponibilidad.

5.1.3. De la climatización

Después de realizar mediciones de aire y temperatura en el data center, se identifica que existe un cierto porcentaje de recirculación (el aire caliente expulsado por los servidores se mezcla con el aire frío) y bypass (el aire frío no está siendo consumido directamente por los servidores, si no que atraviesa entre los espacios vacíos de los racks). Esto genera que tanto el ventilador de la unidad manejadora de aire, como la máquina condensadora realicen un esfuerzo excesivo, aumentando considerablemente el consumo energético.

5.1.4. Propuestas Etapa 1

Se propone finalmente estudiar y revisar el estándar TIA-942 con tal de implementar el diseño modular establecido. Esto ayudará considerablemente a la escalabilidad del data center, y proveerá una mayor disponibilidad. También se propone la instalación de *Blank Panels* para disminuir la recirculación y *bypass* de aire en los racks. Finalmente se propone consultar con especialistas en eficiencia energética y que identifiquen con mayor precisión los componentes que generan desperdicio energético.

[1] Health Insurance Portability and Accountability Act, reglamento que indica que los registros médicos deben estar altamente protegidos física y digitalmente.

5.2. Etapa 2 - Mejoras de IT y administración (Alta disponibilidad).

5.2.1. De las redes

Como se mencionó en el capítulo 3, la arquitectura de 3 capas de Cisco presenta una topología no sostenible para data centers de hiper-escala (en los que usa la arquitectura Leaf-Spine, o organizada en PoDs), debido al complejo manejo de múltiples switches, y al ancho de banda desperdiciado al existir en ejecución el protocolo *Spanning Tree* (alcanzando hasta un 50 % de pérdida de ancho de banda debido a enlaces bloqueados), sin embargo para el caso de uSCI, la arquitectura de switches de acceso y distribución cumplen con las necesidades del DI, incluso si necesitara escalar varios racks.

Al existir switches esparcidos por todo el Departamento de Informática, un acercamiento a través de SDN si podría ser factible, ya que simplificaría considerablemente la administración de la red a nivel del Departamento, aumentaría el rendimiento considerablemente al no existir enlaces bloqueados, y permitiría la expansión a nuevos laboratorios de forma fácil. Sin embargo, requeriría una actualización del equipamiento actual, ya que este no soporta el protocolo OpenFlow.

5.2.2. Del Almacenamiento

Para escalar el almacenamiento, una opción costo eficiente podría ser apoyarse con las oportunidades que ofrece la computación en la nube, el cual se detalla con más detalle en la etapa 3. Sin embargo, el sistema de almacenamiento usado en uSCI es bastante robusto al usar una combinación de servidores SAN en distintas combinaciones de RAID, para proveer mayor throughput y tolerancia a fallos.

5.2.3. De la plataforma de virtualización

Aunque VMWare vSphere es la suite de virtualización más usada en la actualidad, ya que provee una cantidad superior de componentes, RHEV en este caso es una opción costo-eficiente, ya que los elementos como User Portal, licencias, actualizaciones y componentes extras, no requieren un costo adicional como en el caso de VMWare. Un

cambio en el sistema de virtualización afectaría a otros laboratorios, por ejemplo los que usan la infraestructura de escritorios virtuales. Se presenta un estudio de extensión a la plataforma de virtualización usando la computación en la nube.

5.2.4. De las aplicaciones y servicios

El data center no está enfocado directamente en el desarrollo de grandes aplicaciones, si no que más bien en proveer hosting, housing, y VPS tanto a funcionarios, estudiantes y externos, por lo que no hace uso extensivo de herramientas de CI/CD. Si utiliza por otro lado herramientas de gestión de configuraciones, como Puppet y Chef, sin embargo esas herramientas ya son lo suficientemente potentes como para poder administrar un data center de esta escala.

Por otro lado, los sitios web tanto de uSCI como de estudiantes o profesores, están hospedados en máquinas virtuales. Se recomienda implementar el uso de contenedores, como Dockers, usando un servidor centralizado dedicado a la orquestación de contenedores como Kubernets, o Docker Swarm para proveer alta disponibilidad a estos sitios en caso de falla de algún servidor físico.

5.2.5. Propuestas Etapa 2

Se propone implementar la arquitectura de redes de 3 niveles estándar, es decir, con switches de redundancia, ya que en estos momentos, si falla algún switch de la capa de distribución, se deja sin conectividad a todo el Data Center. En la figura 21 se presenta una arquitectura de red propuesta. Las ‘X‘ representan enlaces bloqueados debido a la existencia del protocolo Spanning Tree, sin embargo el rendimiento comparado con la arquitectura existente se mantiene, ya que la nueva arquitectura está diseñada para proveer mayor disponibilidad en caso de fallas.

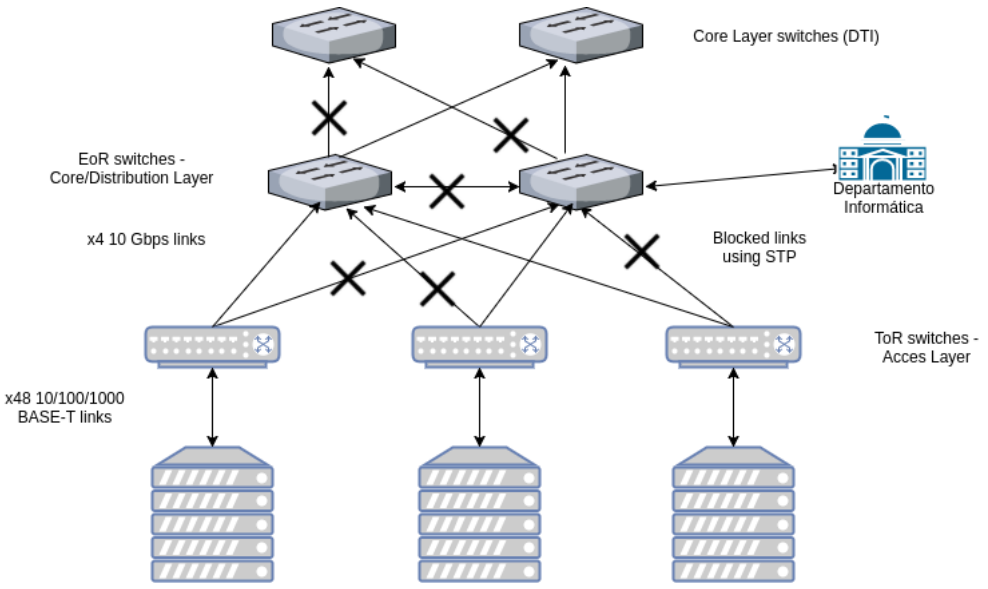

FIGURA 21, PROPUESTA DE ACTUALIZACIÓN A LA ARQUITECTURA ACTUAL DE REDES. FUENTE: ELABORACIÓN PROPIA.

Para el almacenamiento, se propone finalmente apoyarse en la plataforma S3 de AWS usando el modelo de costos de Acceso Infrecuente[2], orientado a almacenar datos que fueran a ser utilizados por máquinas virtuales orientadas a la computación de alto rendimiento.

5.3. Etapa 3 - Asistencia de motores de cómputo en la nube (Agilidad).

Las plataformas Cloud ofrecen una gran oportunidad a emprendimientos y pequeñas y medianas empresas, de poder lanzar y escalar sus productos o servicios, al ofrecer rápidamente infraestructura IT, con tal de permitir a las organizaciones enfocarse netamente en el desarrollo y producción. Sin embargo, las plataformas Cloud también ofrecen una oportunidad de escalar la infraestructura a empresas y organizaciones ya conformadas con infraestructuras On-premises. Dentro de estos beneficios se encuentran:

- Eliminación del estudio de capacidad: Cuando se despliegan sistemas de almacenamiento o servidores On-Premise, es imposible saber con anticipación la cantidad

[2]S3 - Infrequent Access es una clase de almacenamiento en la nube, escalable, altamente disponible, y a bajo costo, orientada a almacenar datos que no sean frecuentemente accedidos, pero que su recuperación deba ser rápida.

exacta de recursos necesarios para proveer un servicio, y lo más probable es que se produzca sobre-abastecimiento (gastos extras innecesario), o abastecimiento insuficiente (recursos que terminan por agotarse generando posibles degradaciones en el servicio). Los sistemas cloud cuentan con data centers escalables, y cuentan con la infraestructura necesaria para proveer recursos de forma continua a sus clientes, simulando entornos de recursos ilimitados.

- Eliminación de infraestructura física: La mantención y administración de equipamiento energético, enfriamiento, sistemas de seguridad, y cableado excesivo quedan completamente eliminados, dejando a cargo al proveedor de la plataforma cloud, de esta forma el consumidor puede enfocarse en el desarrollo y su negocio.
- Estallido dinámico (Dynamic Bursting): Proveer grandes capacidades de cómputo CPU/GPU durante el tiempo necesitado (por ejemplo para tareas de simulaciones, o entrenamiento de redes neuronales), sin necesidad de adquirir los servidores físicos ni configurarlos, haciendo uso eficiente de los recursos monetarios.
- Auto escalamiento: Proveer elasticidad a los recursos de forma que puedan crecer/disminuir de acuerdo a las necesidades. El auto escalamiento permite proveer múltiples máquinas virtuales temporales de forma instantánea para procesamiento en horarios o días en los que se generen *peaks* de solicitudes, permitiendo procesar compras en un *Cyberday* sin colas o degradaciones, o en los primeros días de inscripción de ramos en SIGA sin presentar intermitencias en el servicio.

Por estos motivos, la etapa 3 considera migrar algunos elementos del DC a una nube pública, como caso particular: Amazon Web Services. Se elige esta plataforma por su modelo de negocios flexible, por la oportunidad que ofrece establecer enlaces dedicados entre el data center y la nube pública a través de Level 3 (actual proveedor de Internet de la Universidad), y porque el autor ya posee conocimientos sobre esta plataforma.

5.3.1. Análisis de costos

A continuación se presenta un estudio de costos, en los que se busca desplegar los costos generados por mantener el data center y posibles costos de inversión en el escalamiento de su infraestructura física, versus los costos que genera mantener recursos similares en un entorno cloud.

Es una tarea bastante engorrosa realizar un análisis de costos sobre infraestructuras tan complejas como los data centers o plataformas cloud, ya sea debido a dificultades técnicas, mezcla de recursos que aportan al data center y al mismo tiempo a otros lugares, cálculo de pérdidas de energía en los procesos internos, entre otros, por lo que se realizarán los siguientes supuestos y/o consideraciones:

- No se consideran desastres naturales que puedan destruir parcial o completamente el data center y/o la plataforma cloud de estudio. Ambas plataformas funcionan 24/7 sin cortes.
- En el caso de actualización del equipamiento del data center, no se consideran los costos asociados a instalaciones y/o personal externo necesario para configurar los componentes. No se consideran costos de mantenciones al equipamiento IT, suministro de energía y climatización.
- El espacio físico no presenta mayor problema, y es posible ubicar racks sin inconvenientes. El espacio físico no genera costos extras.
- No se consideran los costos de licencia, ya que tanto en una arquitectura On-premise vs Cloud deben obtenerse licencias.
- No se consideran impuestos, y/o traslado desde el extranjero en el caso de adquirir nuevos componentes IT.
- No se consideran costos de ancho de banda para el data center de informática, pero si para AWS.
- Se consideran solo componentes que están relacionados directamente con la sala de datos, y no con otros laboratorios, y/o salas.

Antes de avanzar en la propuesta, se presentan los 3 modelos de costos que mantiene Amazon para la plataforma de máquinas virtuales EC2:

- **Bajo Demanda**. El modelo Bajo Demanda consiste en proveer máquinas virtuales de distintos tamaños basada en el uso actual. Se cobra por hora de uso, es decir, si se provee una máquina virtual por 2 horas, solo se paga el uso efectivo de esas 2 horas. Este modelo es ideal para levantar máquinas de prueba, simulaciones, o grupos de escalamiento instantáneo temporales.

- **Instancias Reservadas**. El modelo de Instancias Reservadas permite reservar instancias por el plazo de hasta 1 hasta 3 años, disminuyendo los costos de las instancias Bajo Demanda entre 40 y 50 %. Este modelo es ideal para aplicaciones web, o servicios de producción que deban mantenerse constantemente en ejecución.

- **Spot**. El modelo Spot permite proveer máquinas virtuales usando un esquema de oferta y demanda, donde los valores comienzan en hasta un 90 % de descuento respecto de las instancias Bajo Demanda. El cliente define el máximo valor que está dispuesto a pagar. Cuando la instancia que está consumiendo sobrepasa ese valor, Amazon termina inmediatamente esta instancia[3]. Aunque es un poco más complejo administrar este tipo de instancias, este modelo es extremadamente costo-eficiente para realizar tareas orientadas a cálculo de simulaciones, entrenamiento de redes neuronales, o procesamiento en paralelo a gran escala. Usar instancias Spot para cálculos masivos o computación de alto rendimiento es un enfoque costo-eficiente y ágil, al eliminar los elevados costos iniciales, y permitir al personal del Departamento de Informática contar con una infraestructura de elite en un par de segundos.

5.3.2. EC2 versus On-Premise

Se estudiarán 2 tipos de máquinas virtuales, una orientada a propósito general, y una orientada a HPC, usando GPUs.

Caso 1: EC2, t2.micro

Una máquina virtual proveída por el data center al estudiantado tiene en general, las siguientes características:

- CPU: 1vCPU
- Memoria: 1GB
- Almacenamiento: 12 GB.

[3]En la práctica, Amazon envía una notificación con solo 2 minutos de anticipación de que la instancia será terminada, por lo que se recomienda trabajar con funciones Lambda [69] para manejar el término correcto de una instancia terminada automáticamente

En AWS, la instancia EC2: t2.micro tiene características idénticas (aunque en ocasiones el hardware subyacente puede ser distinto). En la tabla 2 puede observarse los distintos costos que genera mantener una máquina virtual de similares características tanto en AWS como en el data center On-premise.

Infraestructura	Valor Hora (USD)	Valor Mensual (USD)
On-Premise	$0.00745	$5.36
EC2: Bajo Demanda	$0.013266	$9.55
EC2: Instancia Reservada (1 año)	$0.0086	$6.19
EC2: Instancia Reservada (3 años)	$0.0066	$4.75

Tabla 2, Costos estimados de mantener una máquina virtual en el data center versus AWS[4]. Fuente: Elaboración propia

Es fácil observar que los costos entre mantener una máquina virtual en el data center, y en instancias reservadas en EC2, son muy similares, por lo que es completamente factible apoyarse en los sistemas de virtualización que ofrece la nube si se requiere una extensión[5].

Caso 2: EC2, p3.x2large

En el caso anterior, se asumió que el data center ya contaba con la plataforma de virtualización necesaria para proveer una MV de las características especificadas. En este momento no existe el hardware necesario para proveer una máquina virtual de tal magnitud para proveer cómputo HPC a nivel de GPUs en el data center, por lo que es necesario adquirir estos nuevos componentes. Dentro de estos componentes se consideran los siguientes gastos de capital:

[4]Los valores de las instancias EC2 incluyen además el valor por almacenamiento en volúmenes EBS

[5]Cabe destacar que en estos momentos el Departamento de Informática de la Universidad no incluye dentro de su presupuesto el uso de la energía eléctrica, el cual para efectos de electricidad, el DI lo ve como un recurso *gratis*, el cual corre por cuenta de la Universidad. Si este determinara que en algún futuro esto no es sostenible, las plataformas cloud presentan una infraestructura factible para extender los recursos IT del Departamento de Informática.

Componente	Costo Unitario estimado(USD)
Rack 42U	$1000
Dell PowerEdge C4130	$4000
NVIDIA Tesla V100 24GB	$8000
Switch L2 48 Ports GbE	$1000
Total Aproximado	14000

Tabla 3, Costos estimados de adquirir una infraestructura orientada a HPC basada en GPU, Fuente: Elaboración propia

Como se puede apreciar, el costo de esta nueva infraestructura requiere un costo total inicial de 14000 $USD, además de un costo pronosticado mensual de 56 USD por consumo energético (se calcula que el nuevo rack consumirá cerca de 0.300 KWh.).

Como ejemplo de comparación, se usa la instancia $p3.x2large$ con las siguientes características:

- GPU: x1 NVIDIA Tesla V100
- CUDA Cores: 5120
- Memoria GPU: 16 GB
- vCPU: x1 Intel Xeon E5-2686 v4
- Memoria VM: 64 GB

En su versión Bajo demanda, el valor es 3.06 USD por hora, en su versión Instancia Reservada el valor es 1.332 USD por hora, y en su versión Spot se asume un valor promedio de 0.65 USD por hora[6]. En la siguiente tabla puede apreciarse los costos de mantener una infraestructura de similares características en los distintos modelos de costos de AWS.

[6]Al momento de escritura, el mínimo valor spot es de 0.5 USD por hora, sin embargo debido a fluctuaciones del mercado es probable que este valor no siempre se mantenga, por lo que asume un valor más alto que el actual.

Infraestructura	Valor Hora (USD)	Valor Mensual
EC2: Bajo Demanda	$3.06	$2203
EC2: Instancia Reservada (3 años)	$1.332	$897
EC2: Instancia de subasta	$0.65	$468

Tabla 4, Costos estimados de mantener una máquina virtual en el data center orientada a GPU versus AWS. Fuente: Elaboración propia

En la figura 22 puede apreciarse que los costos de mantener una instancia EC2 en el modo Bajo Demanda supera considerablemente los costos de adquirir el equipamiento propio al cabo de 1 año, por lo que este tipo de instancia y modelo de costo, no es viable a largo plazo.

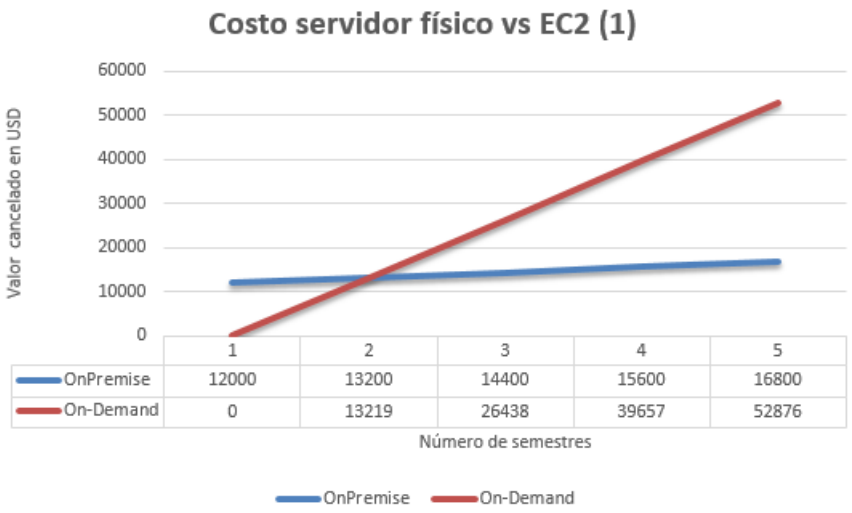

Figura 22, Comparativa de costos de adquirir equipamiento GPU físico versus instancias Bajo Demanda EC2. Fuente: Elaboración propia.

Para el siguiente caso, se usa una Instancia Reservada. Cabe mencionar que el uso de una instancia reservada obliga al consumidor a hacer el pago total de la instancia por hora durante 3 años, a un cierto porcentaje de descuento, sin embargo presenta un modelo de costo más eficiente que el de la instancias bajo demanda. Cabe destacar que estas instancias no son apropiadas para este tipo de computación de alto rendimiento, ya que se enfocan en mantener ambientes de producción que deben estar funcionando 24/7. El costo total de la inversión alcanza al costo generado por la instancia EC2 al término de los primeros 18 meses, por lo que tampoco es viable a largo plazo.

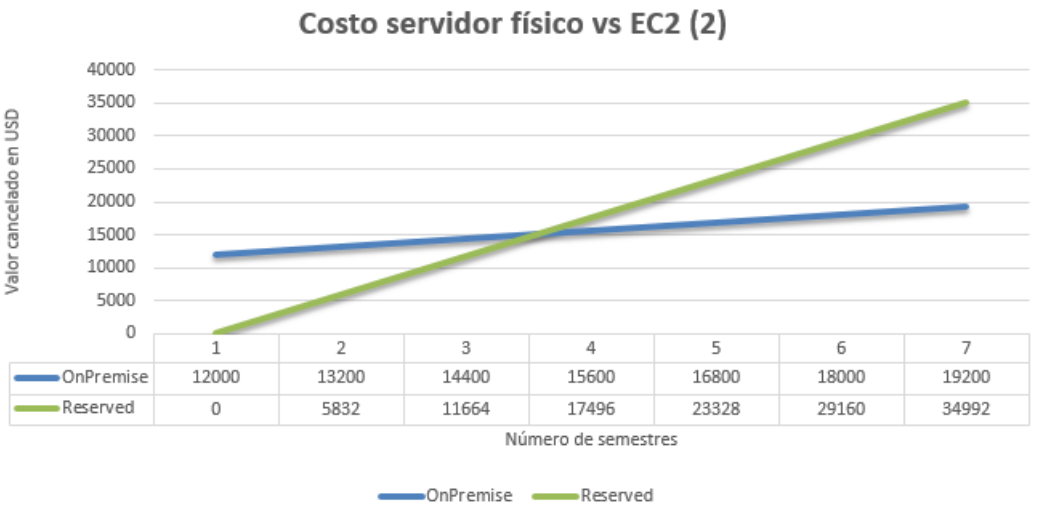

Figura 23, Comparativa de costos de adquirir equipamiento GPU físico versus Instancias Reservadas EC2. Fuente: Elaboración propia.

Para el último caso, se analiza la instancia spot, la cual presenta una considerable mejora en términos de costos, ya que el costo de mantener una instancia EC2 ejecutándose 24/7, recién al tercer año alcanza a los costos de adquirir los componentes físicos.

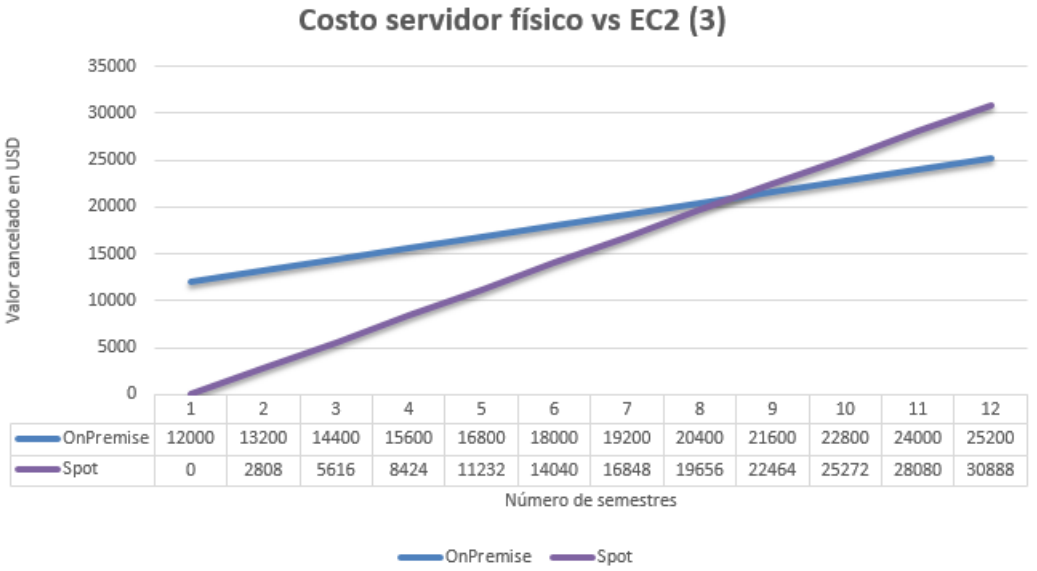

Figura 24, Comparativa de costos de adquirir equipamiento GPU físico versus instancias Spot EC2. Fuente: Elaboración propia.

Puede apreciarse que el mejor modelo de costos en este caso es el modelo Spot, el cual al término del tercer año deja de ser conveniente. Sin embargo es necesario notar que el

caso de estudio se realiza sobre el supuesto de que las instancias están funcionando 24/7. Si el uso de las instancias se disminuye a la mitad (12 horas en promedio por día), el plazo de recuperación de la inversión se extiende a 6 años. Es posible notar que el tiempo de uso que se le dará a las máquinas virtuales es un elemento elemental a considerar en estos tipos de estudios.

Aunque a largo plazo, el uso de la nube podría presentar costos de operaciones más elevados que los que podría presentar una infraestructura on-Premise, es necesario notar las ventajas que provee el uso de las plataformas de computación en la nube:

- Agilidad: Usar instancias en la nube permite provisionar máquinas virtuales potentes al instante, sin necesidad de ningún tipo de logística.
- Flexibilidad. Es posible crear AMIs que se adapten al usuario final, eliminando por completo diferentes configuraciones de la máquina virtual.
- Reembolsables: Si una máquina reservada no se fuera a necesitar más ya sea por motivos de proyectos terminados, o problemas financieros, es posible vender una instancia en AWS MarketPlace, obteniendo una devolución del dinero pagado al inicio.

Queda al criterio del lector definir y evaluar si son convenientes estos beneficios al hacer las comparaciones respectivas.

5.3.3. Propuestas Etapa 3

Se propone la creación de un portal de solicitudes de máquinas virtuales para estudiantes, profesores e investigadores, con sus respectivos requerimientos. Si se fueran a solicitar máquinas virtuales para ejecutar tareas de computación de alto rendimiento, se solicitan tanto instancias On-Demand como instancias de subasta en la plataforma EC2 de AWS, para darle acceso a los usuarios finales a través de SSH o RDP según corresponda.

Capítulo 6

Conclusiones

Se ha presentado un modelo de capas que describe todos los componentes, patrones de diseño y arquitecturas, de la evolución de los data centers desde que se planea su implementación, hasta la entrega de sus servicios. Se concluye que los data centers en la actualidad presentan diversos desafíos, donde los mas importantes son la seguridad, escalabilidad, disponibilidad, y eficiencia energética. Tanto el cloud computing, como los dispositivos IoT, y los motores de bases de datos distribuidas, han empujado a los data centers a acelerar la velocidad en la que se comunican sus servidores de forma interna, tráfico conocido como east-west, el cual puede alcanzar hasta el 70 % del tráfico total en un data center, por lo tanto el escalamiento horizontal es preferido sobre el escalamiento vertical. A continuación se presentan conclusiones generales obtenidas sobre el estado actual de los data centers.

6.1. Conclusiones generales

- Existen diferentes tipos de estándares que permiten una correcta implementación y funcionamiento de un data center, dentro de las que se encuentran las normas mínimas que debe cumplir un data center para que pueda funcionar (temperaturas, humedad, seguridad, protección contra incendios) como las definidas por ASHRAE, o en el caso de Chile: CONAMA para evaluar impacto ambiental. También se presentan diversos estándares que definen estructuras modulares para diseñar data centers de gran escala, como el popular estándar TIA-942, el cual define los

patrones de diseño de redundancia para catalogar a un data center en *Tiers*. Finalmente se describen las organizaciones certificadoras que validan que un data center cumpla con normas mínimas relacionadas con seguridad, eficiencia, disponibilidad, etc, como lo son Uptime Institute, o SAS.

- La explosión del Cloud Computing y de los dispositivos IoT ha empujado a los data centers a elevar de forma considerable el tráfico servidor-a-servidor, por lo que la infraestructura de red de un data center se convierte en su columna vertebral. Un diseño pobre de redes podría llevar a la saturación de los enlaces, y por lo tanto a fuertes degradaciones en los servicios, impactando fuertemente a las aplicaciones que corren sobre el data center.

- Como se establece en el capítulo introductorio, el consumo energético en los data centers llega al 4 % del consumo a nivel mundial, por lo que generan un impacto ambiental considerable. La eficiencia energética en los data centers no solo reduce la huella de carbono, si no que reduce considerablemente los costos generados por energía mal aprovechada, perdiendo millones de dolares al año por concepto de mal manejo energético. El *Free Cooling*, o enfriamiento obtenido del ambiente es uno de los diseños más eficientes en la actualidad, permitiendo disminuir el PUE de las organizaciones a 1.2, 1.1 e incluso menor.

- Aunque los estándares definen que los patrones de diseño redundantes y/o duplicados proveen un uptime de aproximadamente 99.995 %, infraestructuras redundantes no siempre son sinónimo de disponibilidad, ya que las fallas humanas de administración, o ataques de denegación de distribuidos a sistemas DNS siempre podrán dejar a un data center sin servicio.

- La virtualización es una técnica fundamental en la administración de data centers, ya que hace un uso extremadamente eficiente de los recursos IT. No solamente es posible virtualizar servidores a través de máquinas virtuales, también es posible virtualizar redes, volúmenes de almacenamiento, e incluso data centers enteros a través de las nubes públicas.

- Las nubes públicas proveen agilidad, escalabilidad, y alta disponibilidad para desplegar recursos IT de forma rápida. Generalmente son los emprendimientos, o las pequeñas organizaciones quienes optan por las nubes públicas, ya que permiten desplegar una infraestructura IT escalable y tolerante a fallas en cosa de minutos,

permitiéndoles enfocarse netamente en el desarrollo de sus productos, y no en la infraestructura IT subyacente, como lo sería la administración de un data center entero. Dependiendo del caso, podría ser más económico el uso de las nubes publicas incluso a largo plazo, si se manejan de forma correcta los recursos IT y se mantiene control sobre los recursos desplegados.

- Tanto nubes públicas como infraestructuras On-Premise tienen costos ocultos, por ejemplo costo de almacenamiento, ancho de banda en nubes públicas, y costos de mantención, logística, infraestructuras mezcladas (NOC, SOC, IT Room), por lo que es extremadamente complejo manejar todas las variables al hacer una comparación nube/On-Premise.

6.2. Conclusiones específicas

Como caso de estudio, se analizó el data center del Departamento de Informática de la UTFSM, al cual se le encontraron fortalezas, y debilidades, para las cuales se presentan diversas mejoras. Las propuestas fueron presentadas al equipo administrador del Data Center, quienes deben evaluar si implementan las mejoras sugeridas. A continuación se revisan las conclusiones principales obtenidas del modelo de capas enfocadas en el data center del Departamento de Informática.

- Estándares: Si se busca escalar y aumentar el tamaño de la infraestructura del data center, es necesario guiarse por los lineamientos de los estándares reconocidos, como la TIA-942, o BICSI –2014 con tal de proveer arquitecturas modulares, eficientes, escalables, y altamente disponibles.

- Seguridad: Los data centers de gran escala cuentan con muros de protección, cercado eléctrico, protocolos de control de acceso a personal autorizado, destrucción completa mediante químicos de los componentes que termina su ciclo de vida (como discos duros o unidades flash). Ninguno de estos componentes son observados en el data center debido a los costos que conllevan, sin embargo tampoco se hacen altamente necesarios, debido al bajo nivel de confidencialidad de los datos almacenados.

- Infraestructura: Se recomienda realizar un análisis exhaustivo de eficiencia energética con ayuda de profesionales, ya que la energía desperdiciada no solo se traduce en un mayor impacto ambiental, si no que en un gasto monetario importante.

- Redes: Aunque la arquitectura de 3 capas ya no es factible de implementar para data centers que requieran un elevado tráfico servidor-a-servidor (donde se ocupan arquitecturas escalables y más eficientes respecto al uso de enlaces como Leaf Spine), aún es utilizada en miles de data centers en el mundo, siendo el de el DI uno de ellos. Sin embargo la topología del DI carece de redundancia, por lo que se propone agregar 1 switch a la capa de distribución para proveer una mayor disponibilidad.

- Almacenamiento: La arquitectura de almacenamiento cuenta con lo necesario para proveer espacio suficiente a los requerimientos del DI, esto es, información de personal, servidor de correos, e-mail, servidores de archivos, y almacenamiento de VMs. Esta información es replicada diariamente al servidor bacula de respaldo. Un tercer método de respaldo en la nube podría ser factible, siempre y cuando se opte por una opción costo-eficiente, como AWS Glacier.

- Virtualización: La plataforma de virtualización da soporte para las necesidades del DI. Aunque los costos de operación son muy similares a los costos de las plataformas de nubes públicas, se tiene el valor agregado del ancho de banda local, y acceso a repositorios locales. Se analiza una extensión de la plataforma de virtualización usando el servicio EC2, el cual demuestra ser costo eficiente en el uso de las instancias de subasta para computación de alto rendimiento basado en GPUs.

- Administración: El software usado para administrar servidores distribuidos es actualmente usado en data centers de gran escala, por lo que no se proponen mayores recomendaciones en esta área.

- Aplicación: El data center no está orientado al desarrollo y puesta en producción de software que requiera el uso de Entrega Continua, por lo que no se proponen mayores recomendaciones en esta área.

6.3. Trabajo futuro

El modelo de capas presentado está lejos de estar completo, al existir una infinidad de distintas arquitecturas o patrones de diseño de cada capa del modelo de data center. La principal tarea que deja abierta esta memoria, es seguir manteniendo este modelo con arquitecturas que no fueron mencionadas, e ir actualizándolo a medida que aparezcan nuevas. Cabe mencionar que cada una de las capas del modelo permite realizar un trabajo de memoria por si sola, permitiendo la profundización de cada una de ellas.

Por otro lado, el PUE estimado para el data center de Informática en esta memoria es 2, sin embargo se estimó dejando de lado una infinidad de variables, por lo que fácilmente podría llegar a 2.5, o 3, el cual representa un data center altamente poco eficiente. Se propone realizar un trabajo más exhaustivo sobre eficiencia energética en el data center, y proponerlas como mejora al resto de infraestructuras de la Universidad, ya que informática no es el único departamento que mantiene un data center.

Finalmente, se deja como trabajo futuro diseñar una topología de redes basada en SDN por ejemplo usando el software *Mininet*, simulando la infraestructura de redes del Departamento de Informática, y analizar las ventajas y desventajas frente a la estructura existente.

Bibliografía

[1] http://www.zdnet.com/article/top-10-countries-in-which-to-locate-a-data-center/ consultado 24 enero, 2017.

[2] http://www.datacenterdynamics.com/content-tracks/power-cooling/the-top-ten-best-countries-to-locate-a-data-center/32694.fullarticle consultado 20 enero, 2017.

[3] Xiaojing Zhang et al, Power consumption modeling of Data Center IT Room with distributed Air Flow.

[4] http://www.independent.co.uk/environment/global-warming-data-centres-to-consume-three-times-as-much-energy-in-next-decade-experts-warn-a6830086.html consultado 20 enero, 2017.

[5] https://yearbook.enerdata.net/#energy-primary-production.html consultado 20 enero, 2017.

[6] https://www.ovoenergy.com/guides/energy-guides/average-electricity-prices-kwh.html consultado 21 enero, 2017.

[7] https://www.zeniumdatacenters.com/research-reveals-data-center-sector-still-rocked-by-natural-disasters/ consultado 17 enero, 2017.

[8] Mark Winther, Tier 1 ISP, What they are and why they are important, 2006.

[9] https://www.expedient.com/data-center-build-vs-buy-calculator/ consultado 17 enero, 2017.

[10] http://www.theatlantic.com/technology/archive/2016/01/amazon-web-services-data-center/423147/ consultado 24 de enero, 2017.

[11] IBM Research report, IBM Security services 2014, Cyber security Intelligence index, Junio 2014

[12] https://www.wired.com/2012/10/google-data-center-secrets/ consultado 25 enero, 2017.

[13] Kevin mcGarthy et al, Comparing UPS systems design configurations, APC White Paper 75, 2011

[14] Victor Gabriel Galván, Data Center - Una mirada por dentro, 2014.

[15] Victor Avelar, Cálculo del requisito total de potencia para centros de datos, Informe interno 3, 2010.

[16] http://www.riello-ups.com/questions/7-what-are-the-benefits-of-an-external-maintenance-bypass-switch Consultado 29 de enero, 2017.

[17] Fluke networks white paper, Ten top problems network techs encounter, 2009.

[18] TIA Standard-568-C.1, Commercial building Telecommunications Cabling Standard, 2009.

[19] Anixer INC, European Standards Reference Guide, 2012.

[20] Tony Evans, The different technologies for cooling data centers, APC White Paper 59, 2012.

[21] https://www.ashrae.org/

[22] Pearl HU et al, How to choose a IT rack, APC White Paper 201, 2009.

[23] https://es.uptimeinstitute.com/uptime-tia Consultado 21 de febrero, 2017.

[24] http://www.equinix.com/services/data-centers-colocation/standards-compliance/ Consultado 10 abril, 2017.

[25] Uptime Institute, Tier Certification, Certifications Summary, 2014.

[26] https://www.datacenterjournal.com/is-your-data-center-ssae-16-certified/ Consultado 15 Marzo, 2017.

[27] http://www.subtel.gob.cl/normativa-tecnica-internet/ Consultado 17 Marzo, 2017.

[28] Neil Rasmussen, The different types of UPS Systems, APC White Paper 1, 2011.

[29] https://www.pcisecuritystandards.org/

[30] http://www.opencompute.org/

[31] Kashif Bilal et al, Quantitative comparisons of the state of the art data center architectures, Octubre 2012.

[32] http://www.cisco.com, Campus LAN switches. Consultado 28 de Marzo, 2017.

[33] http://bradhedlund.com/2009/04/05/top-of-rack-vs-end-of-row-data-center-designs/ Consultado 28 de Marzo, 2017.

[34] http://dyn.com/blog/dyn-analysis-summary-of-friday-october-21-attack/ Consultado 2 de Abril, 2017.

[35] Big Switch Networks, Core and Pod Data Center Design, Julio 2014.

[36] Xialin Li et al, Green Spine switch management for datacenter networks, Julio 2016.

[37] Mohammand ALizadeh, On the data path performance on leaf spine datacenter fabrics, 2013.

[38] https://code.facebook.com/posts/360346274145943/introducing-data-center-fabric-the-next-generation-facebook-data-center-network/ Consultado 12 de Abril, 2017.

[39] https://www.thegreengrid.org/

[40] https://aws.amazon.com/message/41926/ Consultado 20 de Abril, 2017.

[41] https://www.researchgate.net/publication/305321956_Reading_and_Writing_Single-Atom_Magnets Consultado 21 de Abril, 2017.

[42] Ben Cuttler et al - IEEE Spectrum Magazine, Dunking the Data Center, Marzo 2017.

[43] https://www.gartner.com/doc/reprints?id=1-2G2O5FC&ct=150519 Consultado 25 de Abril, 2017.

[44] https://www.facebook.com/pg/PrinevilleDataCenter/ Consultado 28 de Abril, 2017.

[45] http://www.datacenterknowledge.com/archives/2014/12/09/rise-direct-liquid-cooling-data-centers-likely-inevitable/ Consultado 28 de Abril, 2017.

[46] http://www.datacenterknowledge.com/archives/2016/01/06/data-center-design-which-standards-to-follow/ Consultado 01 de Junio, 2017.

[47] http://searchdatacenter.techtarget.com/tip/Uptime-TIA-and-BICSI-Who-runs-the-data-center-design-standards-show Consultado 01 de Junio, 2017.

[48] http://www.asicomdatacenter.com/datacenter-en-chile-no-todo-lo-que-brilla-es-tier/ Consultado 03 de Junio, 2017.

[49] https://www.masterdc.com/blog/what-is-data-center-free-cooling-how-does-it-work/ Consultado 03 de Junio, 2017.

[50] http://searchdatacenter.techtarget.com/feature/How-many-VMs-per-host-is-too-many Consultado 05 de Junio, 2017.

[51] https://www.opennetworking.org/

[52] Danilo Vergara, Implementación de un Software-Defined-DataCenter usando OpenStack, Junio 2015.

[53] https://www.chef.io/

[54] https://www.ansible.com/

[55] https://saltstack.com/

[56] https://puppet.com/

[57] https://www.openstack.org/

[58] https://cloudstack.apache.org/

[59] https://www.vmware.com/products/vcloud-suite.html

[60] Renato Covarrubias, Diseño de un sistema para distribución de código, sin bajas de servicio, en múltiples data centers, Noviembre 2014.

[61] https://kubernetes.io/

[62] https://docs.docker.com/engine/swarm/

[63] Maximiliano Osorio, Evaluación de Kubernetes, un sistema de orquestación de docker containers para computación en la nube.

[64] https://www.ibm.com/storage/spectrum

[65] Tabla Gartner, https://www.gartner.com/doc/reprints?id=1-3B9FAM0&ct=160707&st=sb Consultado 19 de Agosto, 2017.

[66] Paul Manning, VMWare vStorage Virtual Machine FileSystem, 2010.

[67] https://blog.bacula.org/

[68] Chilquinta, Tarifas suministro eléctrico, Marzo 2017.

[69] http://docs.aws.amazon.com/lambda/latest/dg/welcome.html, consultado 05 de Noviembre, 2017.

[70] http://www.42u.com/measurement/pue-dcie.htm, consultado 05 de Noviembre, 2017.

[71] https://www.chilquinta.cl/seccion/23/opciones-tarifarias.html, consultado 28 de Noviembre, 2017.