



**REPÚBLICA DE PANAMÁ
UNIVERSIDAD INTERNACIONAL DE CIENCIA Y TECNOLOGÍA
FACULTAD DE CIENCIAS DE LA COMPUTACIÓN Y TECNOLOGÍA**

**COMPROBACIÓN DE LA DEPENDENCIA DE LA EFICIENCIA DEL COSTO
DEL ALMACENAMIENTO DE DATOS EN LA NUBE AWS CON EL DISEÑO
DEL CICLO DE VIDA**

**PROYECTO DE TRABAJO PARA OPTAR AL GRADO DE LICENCIADO EN
INGENIERÍA EN REDES DE COMUNICACIONES CON ÉNFASIS EN
TECNOLOGÍAS INALÁMBRICAS**

Tutor: Erick A. Ramos Sánchez

Autor: Axel González Crespín

Ciudad de Panamá, enero 2020



**REPÚBLICA DE PANAMÁ
UNIVERSIDAD INTERNACIONAL DE CIENCIA Y TECNOLOGÍA
FACULTAD DE CIENCIAS DE LA COMPUTACIÓN Y TECNOLOGÍA**

**COMPROBACIÓN DE LA DEPENDENCIA DE LA EFICIENCIA DEL COSTO
DEL ALMACENAMIENTO DE DATOS EN LA NUBE AWS CON EL DISEÑO
DEL CICLO DE VIDA**

**PROYECTO DE TRABAJO PARA OPTAR AL GRADO DE LICENCIADO EN
INGENIERÍA EN REDES DE COMUNICACIONES CON ÉNFASIS EN
TECNOLOGÍAS INALÁMBRICAS**

Autor: Axel González Crespín

Ciudad de Panamá, enero 2020



Ciudad de Panamá, __ de enero de 2020

Profesor (a)

Nagib Yassir

Coordinador Comité de Titulación de Estudios de Licenciatura.

Presente.

En mi carácter de Tutor del Trabajo de Grado presentado por el Bachiller Axel González Crespín, documento de identidad (cédula o pasaporte) N. [REDACTED] para optar al grado de Licenciado en Ingeniería en Redes de Comunicaciones con Énfasis en Tecnologías Inalámbricas, considero que el trabajo: Comprobación de la dependencia de la eficiencia del costo del almacenamiento de datos en la nube AWS con el diseño del ciclo de vida reúne los requisitos y méritos suficientes para ser sometido a la presentación pública y evaluación por parte del Jurado examinador que se designe.

Atentamente,

[REDACTED]

Erick Ramos Sánchez

Documento de identidad (Pasaporte), No. [REDACTED]

Línea de Investigación: Ingeniería y sistemas de comunicación.

ÍNDICE GENERAL

PORTADA INTERNA	ii
CARTA DE APROBACIÓN DEL TUTOR	iii
ÍNDICE GENERAL	iv
ÍNDICE DE IMÁGENES	vii
ÍNDICE DE TABLAS	vii
RESUMEN	viii
ABSTRACT	ix
Introducción.....	1
1 CAPÍTULO I <i>PLANTEAMIENTO DEL PROBLEMA</i>	2
1.1 Definición del problema.....	2
1.2 Preguntas de investigación	3
1.2.1 Objetivo general.....	3
1.2.2 Objetivos específicos.....	3
1.3 Justificación.....	4
2 CAPÍTULO II <i>MARCO TEÓRICO</i>	5
2.1 Antecedentes del estudio	5
2.2 Definiciones de términos básicos usados en la investigación.	6
2.3 Historia del cómputo en la nube.....	11
2.3.1 Historia de AWS	14
2.3.2 Historia de Azure	15
2.4 Ventajas de la computación en la nube.....	15
2.5 Desventaja del cómputo en la nube	19
2.6 Características de la computación en la nube.....	20

2.7	Modelos de computación en la nube.....	20
2.8	Hosting.....	22
2.9	Colocation o housing.....	23
2.10	Modelo tradicional de infraestructura de tecnología de la información	23
2.11	Modelos económicos de la infraestructura de TI tradicional	24
2.12	Modelos económicos del cómputo en la nube	25
3	CAPÍTULO III <i>MARCO METODOLÓGICO</i>	26
3.1	Tipo de investigación.....	26
3.2	Paradigma.....	26
3.3	Población y muestra.....	27
3.4	Técnicas e instrumentos para la recolección de datos.....	27
4	CAPÍTULO IV <i>ANÁLISIS DEL ESCENARIO</i>	27
4.1	Estrategias de selección de almacenamiento	31
4.2	Definición del ciclo de vida del dato	33
4.3	Niveles de servicio en el ciclo de vida del dato	34
4.4	Requerimientos	36
4.5	De criterios de evaluación en la comparación de los distintos tipos de almacenamiento	37
4.6	Almacenamiento tradicional evaluado.....	38
4.6.1	Características del almacenamiento Unity 450F.....	38
4.6.2	Configuración del almacenamiento Unity 450F	40
4.6.3	Valor agregado del almacenamiento Dell Unity 450F.....	42
4.7	Almacenamiento en nube evaluado	44
4.7.1	Beneficios de usar volúmenes EBS.....	45
4.7.2	Beneficios de usar simple storage service – S3	48

4.8	Análisis comparativo de rendimiento de los distintos tipos almacenamiento evaluados	51
4.9	Análisis comparativo del costo de los distintos tipos de almacenamiento evaluados	53
4.9.1	Para los primeros 11 Terabytes de almacenamiento se tiene:	53
4.9.2	Desde los 11 Terabytes hasta los 89 Terabytes.....	54
4.9.3	Almacenamiento de data por el 1er año.	55
4.9.4	Almacenamiento de data en reposo, con más de 1 año de haber sido creada.....	55
4.9.5	Escenario de pago por consumo	56
4.10	Comparación de costos a lo largo del ciclo de vida del dato.....	57
4.11	Conclusiones	62
4.12	Recomendaciones	64
5	REFERENCIAS.....	65

ÍNDICE DE IMÁGENES

Imagen 1. RQ, (2018). Modelo de servicios de nube. Extraído de https://www.quisted.net/index.php/2018/08/30/dc-cloud-computing/	22
Imagen 2. (2020). Detalles del almacenamiento usado para la comparación en este trabajo.	41
Imagen 3. (2020). Detalles del costo del almacenamiento usado para la comparación en este trabajo.....	41
Imagen 4. González, Axel (2020). Relación de pagos mes a mes durante 26 meses.	59
Imagen 5. González, Axel (2020). Costos acumulados a lo largo de 26 meses.....	59

ÍNDICE DE TABLAS

Tabla 1. Comparativa de costos y rendimiento de los distintos niveles de almacenamiento s3.	51
Tabla 2. Comparativa de rendimiento para los primeros 11 TB de almacenamiento.	52
Tabla 3. Comparativa de rendimiento luego de los primeros 11 TB de almacenamiento	53
Tabla 4. Comparativa de costos para los primeros 11 TB de almacenamiento.	54
Tabla 5. Comparativa de costos para data almacenada durante el primer año.	55
Tabla 6. Comparativa de costos durante el primer año del ciclo de vida del dato.....	57
Tabla 7. Pago mensual e inversión de cada solución durante el ciclo de vida del dato.	58



**REPUBLICA DE PANAMÁ
UNIVERSIDAD INTERNACIONAL DE CIENCIA Y TECNOLOGÍA
FACULTAD DE CIENCIAS DE LA COMPUTACIÓN Y TECNOLOGÍA**

**COMPROBACIÓN DE LA DEPENDENCIA DE LA EFICIENCIA DEL COSTO
DEL ALMACENAMIENTO DE DATOS EN LA NUBE AWS CON EL DISEÑO
DEL CICLO DE VIDA**

**Autor: Axel González Crespín
Tutor: Erick Ramos Sánchez
Año: 2020**

RESUMEN

La Computación en la Nube o Cloud Computing, por su nombre en idioma Inglés, es un modelo de prestación de servicios de tecnología, que permite al usuario acceder a un catálogo de servicios estandarizados y responder con ellos a las necesidades de su negocio, de forma flexible, adaptativa y escalable, pagando por el consumo efectuado o de forma gratuita bajo circunstancias particulares. Entre los servicios ofrecidos por los proveedores de Cloud Computing, el almacenamiento es de los primeros que comenzó a hacerse popular, y gracias a su evolución como servicio es posible habilitar grandes espacios de almacenamiento de manera muy rápida, confiable y segura. En consecuencia, este trabajo busca demostrar, con carácter descriptivo y con base bibliográfica, que es posible lograr ahorros significativos de costos por almacenamiento haciendo uso de la nube pública Amazon Web Services (AWS) en comparación con un almacenamiento de gama media/alta de la marca Dell-EMC apalancado en un esquema de ciclo eficiente de vida de grandes volúmenes de datos, pudiendo traducirse en una manera económica de mantener y asegurar la data, sin necesidad de sobredimensionar infraestructura de tecnología ni de hacer pagos anticipados. Para esta comparativa de costos y de rendimiento técnico de ambos tipos de almacenamiento se hizo uso de información pública de fabricantes y proveedores de almacenamiento y computo en la nube. La demostración técnica y económica puede resultar un poco más compleja cuando se agregan varias variables adicionales al análisis, como pueden ser la seguridad, niveles de servicio, mantenimiento e inversiones asociadas, por ejemplo. La eficiencia de este esquema depende de un diseño apropiado del Ciclo de Vida de los datos.

Descriptores: Computación en la Nube, Eficacia de la Computación en la Nube, Eficacia en el almacenamiento de datos en la nube, ciclo de vida de datos, elasticidad.

Línea de investigación: Ingeniería y sistemas de comunicación.



**REPUBLIC OF PANAMA
INTERNATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY
FACULTY OF COMPUTING AND TECHNOLOGY SCIENCES**

**CHECKING THE DEPENDENCY OF COST EFFICIENCY
OF STORAGE DATA IN THE AWS CLOUD WITH THE DESIGN
THE CYCLE OF LIFE**

**Author: Axel González Crespín
Tutor: Erick Ramos Sánchez
Year: 2020**

ABSTRACT

Cloud Computing is a model for the provision of technology services, which allows the user to access a catalog of standardized services and respond with them to the needs of their business, so flexible, adaptive and scalable, paying only for the consumption made, or free of charge under particular circumstances. Among the services offered by cloud computing providers, storage is one of the first that began to become popular, and thanks to its evolution as a service it is possible to enable large storage spaces very quickly, reliably and safely. In this work we sought to demonstrate, on a descriptive basis and on a bibliographic basis, that it is possible to achieve significant savings in storage costs by using the Amazon Web Services (AWS) public cloud compared to a medium / high-end storage of the brand Dell-EMC leveraged in an efficient life cycle scheme of large volumes of data. The feasibility of storing large amounts of data in a cloud computing service provider can be translated into an economical way to maintain and secure the data, paying for consumption, without the need to oversize technology infrastructure or make advance payments. For this, a cost and technical feasibility comparison was made using public information from manufacturers and storage providers and compute in the cloud. The technical and economic demonstration can be more complex when several additional variables are added to the analysis, such as security, service levels, maintenance and associated investments, for example. The efficiency of this scheme depends on an appropriate design of the Data Life Cycle.

Descriptors: Cloud Computing, Efficiency of Cloud Computing, Efficiency in cloud data storage, Data life cycle, Elasticity.

Research line: Engineering and communication systems.

Introducción

Durante años los departamentos de tecnología han apoyado muchas de sus decisiones basadas en malas experiencias, aun cuando puede ser beneficioso pues de los errores se aprende, muchas de las malas experiencias también tienen origen en malos contextos y en malas implementaciones. Para ello se ha traído este planteamiento para contextualizar estas primeras líneas.

El crecimiento exponencial de la digitalización, el Internet y la telefonía celular, ha hecho posible llegar a casi todos los rincones del planeta. Esto ha potenciado la necesidad de almacenar información cada vez en mayores proporciones. Ante esto, las tecnologías que hacen frente a esta disrupción son las que son democráticamente accesibles desde cualquier sitio, solo basta con tener acceso a Internet, y esto es parte de la oferta de la computación en la nube.

Es por esta razón que se presenta este trabajo a fin de que consideremos la evaluación de una tecnología tradicional, pero muy vigente y desarrollada, en contraste con la computación en la nube y las ventajas que este ofrece.

Es de notar que este trabajo expone distintos puntos de vista y conclusiones respecto al cómo abordar análisis comparativos de las tecnologías de almacenamiento, y así poder determinar la factibilidad del uso de estas tecnologías en un entorno empresarial, institucional o académico, centrándose en Amazon Web Services, por ser pionero en este tipo de servicios.

Esto sirve como preludio a la creciente demanda de análisis de datos y al cambio de paradigma en el desarrollo de servicios centrados en los usuarios y clientes, por encima de otras consideraciones, para potencializar el crecimiento de los negocios.

1 CAPÍTULO I PLANTEAMIENTO DEL PROBLEMA

1.1 Definición del problema

Uno de los grandes retos de cualquier gerencia de tecnología de una empresa es proveer recursos y servicios de forma eficiente y al menor costo posible.

En cuanto a la retención de datos a largo plazo, esto se convierte en un gran problema con el pasar de los años. En el contexto de una empresa que requiere por distintas razones, inclusive por regulación legal, mantener datos por periodos largos de tiempo, la estrategia de almacenamiento es crítica, las empresas bajo este contexto deben analizar variables como: niveles de servicios de acceso al dato, accesibilidad, confiabilidad, disponibilidad, y cuando se trata de mantener estos datos en centro de datos tradicionales, surgen otras consideraciones importantes, como por ejemplo, que el equipo de almacenamiento deba mantenerse por largos periodos de tiempo bajo cobertura de garantía (aun cuando ya esté depreciado y posiblemente en obsolescencia) que se deben pagar cuotas de mantenimiento, que requiere manos expertas para su administración y manipulación. Este está sujeto a pérdida de información, dada la obsolescencia de los discos donde este se almacena y que por tratarse en oportunidades de equipos mecánicos o de cintas magnéticas, es necesario contar con un respaldo del respaldo.

Todo lo anterior reta a los tecnólogos y al modelo tradicional por el que durante años ha sido seguido por las empresas. Es necesario focalizar nuevos paradigmas, hacer uso de nuevas tecnologías, pero siempre manteniendo el norte ofreciendo servicios de manera eficiente, bajo le premisa de la mejor relación costo/beneficio.

1.2 Preguntas de investigación

Algunas de las preguntas que ha motivado el desarrollo de este estudio son:

- ¿Es eficiente el uso de almacenamiento en la nube?
- En el contexto de almacenamiento de grandes volúmenes de datos ¿Es económicamente factible pagar por este servicio a los proveedores de nube?
- ¿Cuáles son las consideraciones para hacer una evaluación justa entre el almacenamiento tradicional y la nube?
- ¿En qué momento deja de ser eficiente mantener grandes volúmenes de datos en la infraestructura actual?

1.2.1 Objetivo general

Comprobar que apalancado en un esquema de ciclo eficiente de vida para grandes volúmenes de datos es posible lograr ahorros significativos de costos por almacenamiento haciendo uso de la nube pública Amazon Web Services (AWS) en comparación con un almacenamiento de gama media/alta de la marca Dell-EMC.

1.2.2 Objetivos específicos

1. Describir el ciclo de vida del dato, de tal manera que se explique cómo va evolucionando el tipo de almacenamiento donde este se aloje con el pasar de los días, meses, semanas y años, haciendo que el dato que por periodos de tiempo más largos se almacena pierda la propiedad de ser recuperado con mayor prioridad, potenciando la factibilidad de ahorro económico cuando estos datos se llevan a la nube AWS.
2. Comparar el costo de almacenamiento de los servicios de almacenamiento en Amazon Web Services y un almacenamiento de gama media de la marca Dell-EMC, para corroborar la existencia de un beneficio económico según describa el ciclo del dato.

3. Analizar los costos asociados a los distintos servicios de almacenamiento para sustentar el ciclo de vida del dato y el tipo de almacenamiento a escoger.

1.3 Justificación

La demostración de la dependencia de la eficiencia del costo del almacenamiento de datos en la nube, con un diseño apropiado del ciclo de vida de los datos, permitirá establecer nuevos estándares y procedimientos para cualquier industria que requiera almacenar grandes volúmenes de datos fríos, pagando solo lo que se consume y sin necesidad de hacer grandes inversiones de infraestructura de tecnología. Con este trabajo se pretende cerrar brechas respecto a tecnologías emergentes, con énfasis en la computación en la nube. Para los tecnólogos de manera general, pudieran significar un argumento concluyente que los ayude a tomar decisiones en pro de sus empresas o de los departamentos que ellos gerencian.

Hacer uso eficiente de los recursos tecnológicos ayuda a potenciar los beneficios de las empresas, a simplificar las operaciones e impulsar el hecho de que para tener acceso a las mejores tecnologías del mercado y del mundo ya no es necesario hacer grandes inversiones en tecnología.

Es conocido que cuando se inicia un proyecto de tecnología se hacen provisiones con expectativa de crecimiento, pero ¿Qué ocurre al cabo de un mes si después de realizada una gran inversión el proyecto es descartado? Muy probablemente el dinero haya sido desperdiciado. Por razones como esta, es de gran importancia explorar el cómputo en la nube, cuyo modelo económico permite el pago por uso, bajo demanda.

2 CAPÍTULO II MARCO TEÓRICO

En este capítulo se abordarán conceptos que permitirán contextualizar el trabajo presentado, lo que permitirá comprender las distintas tecnologías y sus términos. Todas estas definiciones ayudarán a comprender de mejor manera las dos tecnologías que se estarán comparando a lo largo de este documento.

2.1 Antecedentes del estudio

Tolosa, C. y González, J., (2014) investigaron “*Amazon web services: alternativa para el almacenamiento de información*” y concluyeron que es de gran importancia asesorarse y valorar cuidadosamente el impacto que la adopción de modelos como la de *IaaS* tendría en una organización, también resalta la importancia de establecer niveles de servicio así como establecer un acuerdo de nivel de servicio con el proveedor que esté muy bien detallado para gestionar adecuadamente los servicios y conseguir las máximas ventajas sin que implique incurrir en algún tipo de riesgo en el normal operar del negocio.

Castro, J., (2016) trabajó sobre el “*Análisis costo beneficio de I.T. in situ vs. IaaS. caso de estudio empresa Smartpro s.a*” concluyó que los costos mensuales que mantienen los proveedores internacionales de nube son accesibles y menores a los proveedores nacionales, según se detalla en los resultados de los cuadros comparativos que realizó, de todas las configuraciones realizadas siempre la nube, con servicios de tipo de infraestructura como servicio de estos proveedores, fueron menos costosos.

Vásquez, C., (2018) investigó sobre “*Plan metodológico comparativo entre invertir en un data center vs infraestructura como servicio cloud computing empresas pymes – caso de estudio empresa Bupartech*” y en una de sus

conclusiones cita: *“Según la evaluación económica realizada en el presente trabajo a diferencia de los sistemas de hardware y software tradicional cloud no requiere de una inversión inicial para comenzar a utilizar las soluciones”*.

De los trabajos anteriores se desprenden iniciativas para dar la bienvenida a la tecnología nube. En principio, el primer trabajo citado hace referencias muy valiosas sobre el cómo ha de ser la adopción del modelo de nube y lo importante, para la fecha de ese trabajo, de la consideración de que no existía ninguna regulación respecto a la limitante del uso y procesamiento de datos en la nube. En ese trabajo ya se hacía énfasis en que el costo de los servicios en la nube es muy competitivos y atractivos para los usuarios. En el segundo trabajo, referencia a lo atractivo del consumo de Infraestructura como servicio, y que la comparación entre proveedores de hosting y de nube existen brechas importantes en cuanto a funcionalidad y costos .En el tercer trabajo, gran aporte señalando que la inversión inicial de adopción de servicios en la nube es equivalente a la demanda de uso de los servicios, lo cual suele ser muy bajo, pues el modelo de pago por consumo a demanda no es posible sino solo a través de proveedores de cómputo en la nube.

2.2 Definiciones de términos básicos usados en la investigación.

Almacenamiento de datos o datastorage: es la retención de información mediante el uso de tecnología desarrollada especialmente para guardar esos datos y mantenerlos lo más accesibles posible.

Almacenamiento de gama media/alta: es un tipo de hardware diseñado para para almacenar grandes volúmenes de datos, que trabaja a grandes velocidades, cuenta con redundancia de sus elementos más críticos y que de manera general brindan rendimiento y capacidades muy superiores al promedio de lo que oferta el mercado.

Arreglo de discos o RAID (del inglés redundant array of independent disks): es típicamente implementado para la protección de la información o incremento del desempeño al acceso de los discos duros. Existen varios tipos de arreglos y los

más usados en la industria son: 0, 1, 5 y el 0+1 ó 10, siendo este último el de mayor desempeño, protección y costo.

Azure; Microsoft Azure, antes conocido como Windows Azure, es la plataforma de computación en nube pública de Microsoft. Proporciona una gama de servicios en la nube, incluidos los de computación, analítica, almacenamiento y redes. Los usuarios pueden elegir entre estos servicios para desarrollar y escalar nuevas aplicaciones, o ejecutar aplicaciones existentes, en la nube pública.

AWS, Amazon Web Services: es la plataforma en la nube más adoptada y completa en el mundo, que ofrece más de 165 servicios integrales de centros de datos a nivel global. Millones de clientes, incluso las empresas emergentes que crecen más rápido, las compañías más grandes y los organismos gubernamentales líderes, confían en AWS para el funcionamiento de sus infraestructuras, el aumento de su agilidad y la disminución de los costos.

Caché de disco: es un tipo de memoria de acceso muy rápido que almacena los datos que se han leído recientemente, con lo que se agiliza el acceso a esos mismos datos en un futuro, así evitar que se tarde mucho tiempo en cargar esos mismos datos.

Ciclo de vida del dato: de principio a fin es clave para que se extraiga la máxima utilidad de los datos. Desde su creación o captura y almacenamiento, pasando por su uso en distintas actividades de procesamiento, hasta que estos se vuelven obsoletos y se eliminan. Todo lo anterior define el ciclo de vida del dato.

Cómputo en la Nube o Cloud Computing: son servicios entregados a través de la conectividad y gran escala de Internet. La computación en la nube democratiza el acceso a recursos de software de nivel internacional, pues es una aplicación de software que atiende a diversos clientes. La multilocación e hiperescalamiento es lo que diferencia la computación en la nube de la simple tercerización y de modelos de proveedores de servicios de aplicaciones más antiguos. Ahora, las pequeñas empresas tienen la capacidad de dominar el poder de la tecnología avanzada de manera escalable.

Datos de acceso frecuente o Hot Data: Es un tipo de dato que requiere ser accedido de manera frecuente, lo que demanda un tipo de almacenamiento particular que responda a las necesidades de acceso del dato.

Datos de acceso infrecuente o Cold Data: Es un tipo de dato que es accedido con baja frecuencia, que no demanda un tipo de almacenamiento particular para cubrir necesidad de rendimiento.

De archivo o archiving: Es un tipo de dato que demanda una frecuencia de acceso muy baja, muy inferior al cold data. Normalmente este tipo de data se ubica en un almacenamiento muy económico.

Deduplicación: es una técnica especializada de compresión de datos para eliminar copias duplicadas de datos repetidos. Un término relacionado con la deduplicación de datos es la compresión inteligente de datos. Esta técnica se usa para optimizar el almacenamiento de datos en disco y también para reducir la cantidad de información que debe enviarse de un dispositivo a otro a través de redes de comunicación.

Dell-EMC son dos empresas que unieron fuerzas en la fusión de tecnología más grande de la historia para convertirse en Dell Technologies. Dell Technologies está conformada por siete líderes en tecnología (Dell, Dell EMC, Pivotal, RSA, Secureworks, Virtustream y VMware) que juntos forman una empresa capaz de impulsar la transformación digital y lograr resultados reales todos los días para los clientes y las personas que se asocian con ella.

Economía de escala: se refiere al poder que tiene una empresa cuando alcanza un nivel óptimo de producción para ir produciendo más a menor coste, es decir, a medida que la producción en una empresa crece, sus costes por unidad producida se reducen. Cuanto más produce, menos le cuesta producir cada unidad.

GB o Gigabyte: es una unidad de medida aproximadamente igual a 1 billón de bytes. Un gigabyte es igual a 1,000MB (realmente 1.024 megabytes).

Hiperescalamiento: en el contexto de centros de cómputo, se trata de la capacidad de asignación o aprovisionamiento masivo de recursos a un servicio

de cómputo en minutos. Por ejemplo, que un cluster de servidores pueda crecer de 2 nodos a cientos de nodos en minutos.

Host spare: es un disco de reserva es un disco duro que se puede utilizar en los sistemas RAID tanto hardware como software.

Infraestructura como servicio o Infrastructure as a Service, IaaS por sus siglas en inglés: es un método para ofrecer funcionalidades de computación, almacenamiento, redes y de otros tipos a través de Internet. La IaaS permite a las empresas utilizar sistemas de funcionamiento, aplicaciones y almacenamiento basados en la web sin tener que comprar, administrar y brindar soporte a la infraestructura de nube subyacente.

Inputs Outputs Per Second o IOPS, por sus siglas en inglés: se ocupa de determinar a qué velocidad puede un disco duro leer datos y escribirlos. Puede usarse tanto para los discos tradicionales mecánicos como para las unidades de estado sólido.

Logical Unit Number o LUN, por sus siglas en inglés: es una dirección para una unidad de disco duro y por extensión, el disco en sí mismo. En una Storage Area Network o SAN, por sus siglas en inglés, un LUN no es normalmente un disco entero sino una partición virtual (o volumen) dentro de un conjunto RAID.

Multilocación: Es la capacidad de estar en más de dos ubicaciones de manera simultánea. En el contexto de este trabajo, tiene más que ver con que los datos y servicios están replicados y disponibles en más de dos ubicaciones.

Niveles de acuerdo de servicio o Service Level Agreement, SLA por sus siglas en inglés: es, simplemente, un acuerdo contractual entre una empresa de servicios y su cliente, donde se define, fundamentalmente, el servicio y los compromisos de calidad. Este tipo de contrato existe, en su forma más elemental, desde hace mucho tiempo; por ejemplo, los contratos de suministro de agua o electricidad. Pero han sido las empresas de telecomunicaciones y de informática las que han popularizado su uso y las siglas SLA. La razón es que los servicios de estas empresas son enormemente flexibles y versátiles, con lo que la selección de las prestaciones adecuadas es difícil y el control de calidad es

complejo. La consecuencia de todo esto es que las expectativas y concepto de calidad entre prestador y cliente van a diferir si no se fijan y documentan a priori.

Nube híbrida: combina nubes privadas, nubes públicas y servidores dedicados.

Nube pública: es una basada en el modelo estándar de la computación en la nube, en la que un proveedor de servicios pone a disposición del público en general recursos, como aplicaciones y almacenamiento, a través de Internet. Los servicios de nube pública pueden ser gratuitos u ofrecidos en un modelo de pago por uso. Una nube pública tiene muchos clientes/usuarios y ofrece escalabilidad de pago por consumo. Es una buena opción para los sitios web que operan de cara al público y que tienen un tráfico impredecible o elevado.

Nube privada: es un tipo de computación en la nube que ofrece ventajas similares a la nube pública, incluida la escalabilidad y el autoservicio, pero a través de una arquitectura propietaria. A diferencia de las nubes públicas, que ofrecen servicios a varias organizaciones, una nube privada está dedicada a una sola organización. Una nube privada ofrece un ambiente dedicado a través de servidores virtuales o dedicados.

TB o Terabyte: es una unidad de medida informática cuyo símbolo es el TB, y es equivalente a 2 a la 12 bytes. Comúnmente se acepta que un terabyte es equivalente a 1,000 GB lo cual implica 1,000,000 de Mb o 1,000,000,000 de Kb.

Tolerancia a fallas: Es la capacidad que tiene un sistema para continuar operando aun cuando uno o varios de sus componentes hayan fallado.

Plataforma como servicio o Platform as a Service, PaaS por sus siglas en inglés: es un entorno de desarrollo e implementación completo en la nube, con recursos que permiten entregar todo, desde aplicaciones sencillas basadas en la nube hasta aplicaciones empresariales sofisticadas habilitadas para la nube. Usted le compra los recursos que necesita a un proveedor de servicios en la nube, a los que accede a través de una conexión segura a Internet, pero solo paga por el uso que hace de ellos.

Servidores dedicados: están dirigidos a los que necesitan ambientes huéspedes únicos. Por lo general, las empresas con operaciones esenciales para el negocio que exigen un alto rendimiento además de un estricto cumplimiento

de las normas confían en los servidores dedicados. Son requeridos cuando los servicios han de ser desplegados y deben cumplir con normas de certificación, en cuando a la privacidad y manejo de dato.

Snapshot o instantánea: también denominadas instantáneas de volumen o VSS (volumen snapshot service) son un elemento de seguridad informática complementarias a las copias de seguridad o respaldos. Son muchos más rápidas de restaurar que un respaldo tradicional.

Software como servicio o Software as a Service, SaaS por sus siglas en inglés: permite a los usuarios conectarse a aplicaciones basadas en la nube a través de Internet y usarlas. Algunos ejemplos comunes son el correo electrónico, los calendarios y las herramientas ofimáticas (como Microsoft Office 365). (Microsoft) lo detalla como una solución de software integral que se adquiere de un proveedor de servicios en la nube mediante un modelo de pago por uso. Usted alquila el uso de una aplicación para su organización y los usuarios se conectan a ella a través de Internet, normalmente con un explorador web. Toda la infraestructura subyacente, el middleware, el software y los datos de las aplicaciones se encuentran en el centro de datos del proveedor. El proveedor de servicios administra el hardware y el software y, con el contrato de servicio adecuado, garantizará también la disponibilidad y la seguridad de la aplicación y de sus datos. SaaS permite que una organización se ponga en marcha y pueda ejecutar aplicaciones con un costo inicial mínimo.

2.3 Historia del cómputo en la nube

Aunque muchos especulan que Amazon Web Services ha sido el padre y creador del cómputo en la nube, la verdad es que sus inicios se remontan a muchos años atrás. (Cabaca, 2018) señala que:

“El concepto cloud computing se ha convertido es una commodity no solo en el sector tecnológico, sino en cualquier ámbito de nuestra vida personal o profesional. Todos asociamos el concepto cloud con la modernidad, pero, ¿Sabías que sus orígenes se

remontan a los años 60? En este artículo descubrimos el origen de la nube y de aquellos que la hicieron posible.

La historia de la tecnología está llena de grandes genios injustamente olvidados. El protagonista de la nuestra, John McCarthy, es un buen ejemplo de ello; creador de Lisp, el lenguaje elegido para el desarrollo de la Inteligencia Artificial (término que acuñó el mismo) y todavía en uso, también está considerada la primera persona que expresó el concepto de computación colectiva.

Durante su discurso para celebrar el centenario del MIT, en 1961, McCarthy introdujo la computación en tiempo compartido, imaginando un escenario futuro donde la capacidad de procesamiento de los ordenadores se vendería como servicio, tal y como se proporcionaba el agua o el gas. En una época donde un ordenador era tan grande como una habitación y valía tanto como una casa, nadie podría imaginar que en unos años todos tendríamos uno miles de veces más potente y en nuestro bolsillo.

En informática, el uso del tiempo compartido indica la capacidad para compartir de forma concurrente un recurso computacional (el uso del procesador, la memoria, el espacio en disco.) gracias a las tecnologías de multiprogramación y la inclusión de interrupciones de reloj como parte del sistema operativo. De esta forma, gracias al software, era posible limitar los recursos que estaban disponibles para cada usuario y optimizar el uso de los equipos.

La idea de «alquilar tiempo de computación» se hizo muy popular durante los primeros años de la década de los sesenta, pero las limitaciones técnicas de la época hicieron muy complicado llevarlo a la práctica. Sí existieron algunas empresas como Tymshare, National CSS o Dial Data que proporcionaban servicios de computación basados en tiempo compartido.

A finales de los sesenta se ofrecían paquetes que incluían editores de texto, entornos de desarrollo para lenguajes de programación, paquetes de ofimática, almacenamiento de archivos y soluciones de impresión a cambio de un alquiler que iba en función del tiempo de conexión y el uso de los recursos.

Para acercarnos un poco más al cloud computing tal y como lo conocemos hoy debemos avanzar hasta 1962, cuando Joseph Carl Robnett Licklider (J.C.R. Licklider, para algunos, el verdadero inventor del cloud computing) introdujo la idea de miles de ordenadores interconectados para poder acceder a recursos y datos desde cualquier lugar.

«Consideren un caso en el que diferentes centros de datos están conectados, con su propio lenguaje y su forma de hacer las cosas. ¿No sería deseable, incluso necesario, que todos se pongan de acuerdo para usar el mismo lenguaje o, al menos, tener una convención para preguntar en que lenguaje habla el otro?» afirmó Licklider en un discurso que ahora se torna casi profético.

Unos meses después, en abril de 1963, formuló un esbozo de una red de ordenadores y la presentó ante la plana mayor de ARPA (Agencia de Investigación de Proyectos Avanzados), para terminar en el Pentágono, donde convenció a los mandos militares de la importancia de una red invulnerable a un ataque soviético. Esta fue la clave para que, en plena guerra fría, se iniciara el desarrollo de ARPAnet, precursor de la Internet que todos conocemos.

De forma casi paradójica, la espectacular bajada de precios que experimentó la informática durante los años setenta y ochenta relegaron a un discreto segundo plano los avances relacionados con el cloud computing. No sería hasta 1996 cuando un ejecutivo de Compaq llamado George Favaloro y un técnico llamado Sean O'Sullivan presentaron el plan de negocio que podéis ver en este enlace: oficialmente, estamos ante la primera vez en la historia donde se usa el término «cloud computing».

A finales de los 90 Salesforce introdujo el concepto de aplicaciones empresariales a través de Internet y, en 2002, los ingenieros de Amazon decidieron que podían utilizar parte de su hardware para ofrecer almacenamiento cloud. La llegada de Gmail, Google Docs, Windows Azure o iCloud, por citar algunos ejemplos conocidos, han ido formando el ecosistema actual, donde un gran porcentaje de lo que hacemos dentro y fuera de Internet depende, en gran medida, de tecnología basada en cloud computing”.

2.3.1 Historia de AWS

(AWS, 2019) “En 2006, Amazon Web Services (AWS) comenzó a ofrecer servicios de infraestructura de TI a las empresas en forma de servicios web, ahora comúnmente conocidos como computación en la nube. Uno de los beneficios clave de la computación en la nube es la oportunidad de reemplazar las inversiones iniciales de infraestructura con costos variables bajos que se escalan con el negocio. Con la nube, las empresas ya no necesitan planificar y adquirir servidores y otra infraestructura de TI con semanas o meses de anticipación. En cambio, pueden activar instantáneamente cientos o miles de servidores en minutos y entregar resultados más rápido.

En la actualidad, Amazon Web Services ofrece una plataforma de infraestructura en la nube altamente confiable, escalable y de bajo costo que impulsa a cientos de miles de empresas en 190 países de todo el mundo. Con ubicaciones de centros de datos en EE. UU., Europa, Brasil, Singapur, Japón y Australia, los clientes de todas las industrias están aprovechando los siguientes beneficios de costos.

AWS ofrece precios bajos de pago por uso sin gastos iniciales ni compromisos a largo plazo. Se puede construir y administrar una infraestructura global a escala, y pasarle los beneficios de ahorro de costos en forma de precios más bajos. Con la eficiencia de la escala y experiencia, se han podido bajar los precios en 15 ocasiones diferentes en los últimos cuatro años”.

2.3.2 Historia de Azure

(Lello, 2016) señala que:

“Microsoft Azure (anteriormente Windows Azure y Azure Services Platform) es una plataforma ofrecida como servicio y alojada en los Data Centers de Microsoft. Anunciada en el Professional Developers Conference de Microsoft (PDC) del 2008 en su versión beta, pasó a ser un producto comercial el 1 de enero de 2010. Windows Azure es una plataforma general que tiene diferentes servicios para aplicaciones, desde servicios que alojan aplicaciones en alguno de los centros de procesamiento de datos de Microsoft para que se ejecute sobre su infraestructura (Cloud Computing) hasta servicios de comunicación segura y federación entre aplicaciones”.

En definitiva, proporciona un entorno gestionado para la ejecución e implementación de aplicaciones y servicios en la nube; para los desarrolladores brinda un entorno de computación bajo demanda y almacenamiento alojado en los centros de datos de Microsoft para aplicaciones en la web”.

2.4 Ventajas de la computación en la nube

a. Reducción de costos

La empresa prescinde de inversiones en infraestructura de Tecnología propia y licencias de software. Para empezar a trabajar no es necesario instalar ningún tipo especial de hardware. Por el contrario, adquirir una infraestructura tecnológica propia es mucho más costoso y no todos los negocios pueden permitírselo. ¿Por qué? porque no se trata solamente de la adquisición tecnológica, sino de todo lo que esto conlleva: gastos de mantenimiento, energéticos, personal técnico etc.

Es posible que con la gestión desde la nube un negocio pueda obtener ahorros significativos. Por otro lado, la empresa se puede centrar exclusivamente en su actividad sin necesidad de tener que dedicar tiempo y recursos a mantenimiento

tecnológico, que a veces puede salir tremendamente costoso. Muchas empresas dedican muchas horas hombre al año para solventar fallas en los equipos de cómputo, lo que nos da una idea de lo que puede suponer para un negocio el costo de mantenimiento de infraestructura si lleváramos el control de horas hombre bajo un presupuesto.

Si bien es cierto que suele asegurarse que la nube siempre es más económica, lo cierto es que la respuesta más sensata sería “depende del escenario”.

b. Movilidad: Acceso desde cualquier dispositivo y lugar

La movilidad se ha convertido en una gran ventaja competitiva, tanto para trabajar o atender clientes desde cualquier lugar, como para favorecer la flexibilidad laboral de un trabajador que puede organizar toda su actividad sin depender de una oficina. Simplemente basta con tener conexión a Internet para acceder a las aplicaciones o a la información. Además, los documentos están alojados en la nube, no en equipos individuales, por lo tanto, distintos usuarios pueden compartirlos o trabajar sobre ellos sin necesidad de estar físicamente juntos, inclusive es posible que trabajen sobre un mismo archivo de manera colaborativa, estando los dos editando el mismo documento al mismo tiempo. Muchas empresas que aprovechan la movilidad, y con esta forma de trabajo, aumentan su productividad en un porcentaje bastante importante.

c. Pago por uso y gasto bajo control

La computación en la nube se basa en modelos de pago por uso. Las empresas contratan solamente los servicios que necesitan por “instantes” de tiempo, permitiendo la posibilidad de ajustar los gastos a necesidades reales. El cliente añade o elimina servicios en función de lo que requiera, evitando tener que invertir en infraestructura propia que con el tiempo quedaría obsoleta. Es decir, se paga por lo que necesita y cuando se necesita. Si determinados días del mes se requiere más capacidad o almacenamiento de datos, basta con ampliar los servicios en línea e instantáneamente. Y si, por ejemplo, la empresa cambia el número de trabajadores o su estructura, solo tiene que añadir o restar funciones bajo demanda.

d. Tecnología siempre actualizada

Poder disfrutar siempre de las últimas versiones del software y de las más modernas aplicaciones hasta hace poco tiempo significaba un “lujo” solo al alcance de las grandes compañías, por el alto costo esto significaba. Con la computación en la nube, los usuarios se aseguran una tecnología siempre actualizada y optimizada. Las actualizaciones, desde el punto de vista del usuario, se hacen automáticamente y simplemente estarán disponibles la próxima vez que inicie su sesión en la nube. La idea de contar con servicios que ya están desarrollados y que son de avanzada antes suponían un imposible por costos de adquisición, como, por ejemplo: reconocimiento de voz, reconocimiento facial, inteligencia artificial, reconocimiento de texto. Ahora se paga por el uso de esa tecnología que no requiere ser desarrollada, sino que requiere que se pague por usarla por el tiempo que se requiera y luego prescindir de ella.

e. Capacidad de almacenamiento ilimitada

Hoy en día se manejan grandes cantidades de datos y la nube ofrece un almacenamiento prácticamente ilimitado. Por ejemplo, si un equipo tiene 500 gigabytes de capacidad de almacenamiento en disco duro, es de considerar que eso es infinitamente pequeño comparado con los terabytes disponibles en la nube. Es decir, un usuario no está obligado a realizar ampliaciones en sus propios equipos cada poco tiempo, con todo el gasto que esto puede significar.

f. Respeto al medio ambiente

Aunque la computación en la nube casi siempre se habla en términos de rentabilidad o productividad, este punto es también muy importante. Hacer uso de la nube, reduce la huella de carbono según (Networld, 2013), haciendo posible ahorrar recursos que pasan de estar almacenados en componentes físicos a ser virtuales. Esto supone un considerable ahorro en consumo de energía, lo que se traduce en importantes beneficios para el medio ambiente. La virtualización puede llegar a reducir el consumo de energía y sus niveles de contaminación en porcentaje importante, considerando que a través de la virtualización es posible reducir equipos de cómputo y consumo energético.

g. No se necesita “lo último” en computadores

Se pueden utilizar computadores con discos duros más pequeños, menos memoria y procesadores. Por ejemplo, en este contexto, un computador, no necesitaría siquiera una unidad de CD o DVD, ya que no requiere instalar programas y los datos o documentos tampoco necesitan ser salvados necesariamente en el disco duro. Además, con menos programas acaparando la memoria en el computador, este rendirá mucho mejor porque tiene menos procesos cargados en memoria.

h. Seguridad

A diferencia de la informática tradicional, en la que cualquier falla en un computador puede destruir todos los datos, con el cómputo en la nube si el disco duro de un computador dejara de funcionar esto no implicaría afectación en la disponibilidad de los datos. Y siendo menos “drásticos”, si simplemente se da el caso en el que un computador se bloquea, todos los datos seguirán disponibles en la nube y totalmente accesibles desde otro dispositivo. En cuanto a las copias de seguridad de la información, que está claro que son vitales para cualquier negocio, pero la realidad es que no todo el mundo las hace, o al menos no de forma regular. En el entorno de cómputo en la nube el usuario se despreocupa de las copias de seguridad porque la nube las realiza automáticamente y con un cifrado seguro, a prueba de cualquier ataque de hackers, solo es necesario activarlos con algunos pocos clics.

i. Se acaba la dependencia de un computador

Con la nube dejas de estar atado a un computador particular. Es decir, se pueden cambiar los computadores sin miedo, porque las aplicaciones y los datos seguirán estando en la nube. Los documentos, datos y aplicaciones son los mismos sin importar qué computador o dispositivo se esté utilizando para acceder a ellos.

j. Igualdad

Por último, y sin duda una de las ventajas más importantes, es que el servicio en la nube permite que las pequeñas y medianas empresas estén en las mismas condiciones que las grandes potenciando un mercado más justo e igualitario en el

acceso a la tecnología. La computación en la nube ha “democratizado” el acceso a las nuevas tecnologías.

2.5 Desventaja del cómputo en la nube

- a. Es necesaria una conexión permanente a Internet. Es imposible acceder a los servicios de cómputo en la nube si no es a través de una conexión a Internet. Durante el período en que una conexión a Internet esté caída no se podrá trabajar con las aplicaciones ni acceder a sus datos.
- b. Resulta indispensable que la conexión a Internet sea de buena velocidad, de esto depende la calidad del servicio que se contrata. Por ejemplo, una conexión a través de líneas telefónicas hace inviable el cómputo en la nube.
- c. Dependiendo de la arquitectura de la solución de tecnología, puede parecer que el cómputo en la nube sea lento, pues hay soluciones que son óptimas cuando están instaladas en el computador local.
- d. Los datos almacenados pueden no estar seguros. La información debe recorrer diferentes nodos para llegar a destino, cada uno de ellos (y sus canales) pueden ser un foco de inseguridad si no se consideran las posibles vulnerabilidades. Si se utilizan protocolos seguros, Hypertext Transfer Protocol Secure (HTTPS, por sus siglas en inglés) por ejemplo, la velocidad total disminuye debido a la sobrecarga que estos requieren para encriptar y desencriptar los datos. Sin embargo, todo depende de cómo estos servicios son administrados, los proveedores de nube suelen dar directrices muy claras sobre cuáles son las mejores prácticas para asegurar la data. Es responsabilidad de quien administra. La llave de seguridad está en manos del quien contrata el servicio.
- e. Crea dependencia por parte de los usuarios hacia los proveedores de servicios, debido a la centralización de las aplicaciones, a la gran cantidad de información almacenada en la Nube y al costo que representa mover grandes cantidades de datos fuera de la nube.

- f. Escalabilidad a largo plazo. A medida que más usuarios empiecen a compartir la infraestructura de la Nube, la sobrecarga en los servidores de los proveedores aumentará, en consecuencia, el costo.

2.6 Características de la computación en la nube

1. Ubicuidad: Posibilidad ofrecida a los usuarios de acceder desde cualquier dispositivo a los servicios contratados de nube, en cualquier lugar y en cualquier momento, siempre cuando se disponga de conexión Internet.
2. Tecnología de la virtualización: Ofrece a los usuarios el nivel de flexibilidad necesario con los dispositivos para la realización de las actividades diarias.
3. Virtualización: Consiste en la creación, a través de software, de una versión virtual de algún recurso tecnológico. Es la tecnología que sustenta el cómputo en la nube.
4. Facilidad de uso: La complejidad tecnológica de la nube es transparente para el usuario final, no hay necesidad de conocer la infraestructura que hay detrás de la nube para hacer uso de sus servicios. La tecnología es presentada de manera sencilla y muy bien documentada para hacer uso de ella.
5. Escalabilidad: Provisión elástica en función de la demanda. Capacidad que consiste en aumentar o disminuir las funcionalidades ofrecidas al cliente, en función de sus necesidades puntuales sin que sean necesarios nuevos contratos, eliminando así el sobrecoste derivado de mantener los sistemas infrautilizados fuera de las puntas de demanda.

2.7 Modelos de computación en la nube

Existen 3 modelos que definiremos a continuación:

- a. Software como servicio – (SaaS): El proveedor de servicio no solo ofrece poder de cómputo y los entornos de ejecución necesarios, sino también los productos

de software. De manera que las aplicaciones informáticas están alojadas en el servidor y se accede a ellas por lo general a través de un navegador web, sin que sea necesario instalar las mismas. De esta manera se libera al cliente de operaciones de mantenimiento, de soporte, y se ofrece una disponibilidad permanente. El cliente no se preocupa de la administración del hardware ni de la actualización de software. Es un servicio que se consume “como está”.

- b. Infraestructura como servicio – (IaaS): Es uno de los modelos de servicios en la nube que depende del aprovisionamiento de recursos informáticos bajo demanda, como hardware virtual (máquinas virtuales, almacenamiento, red, balanceadores de carga, etc.) lo que permite a los usuarios, como administradores de sistemas, instalar sistemas operativos, servidores, bases de datos, que crean el camino para la construcción y despliegue de aplicaciones. El consumidor alquila los recursos de hardware en lugar de comprarlos e instalarlos en su propio centro de datos, lo que le permite ir variando el consumo de los recursos en función de sus necesidades.
- c. Plataforma como servicio – (PaaS): ofrece todos los beneficios y la funcionalidad de IaaS, así como un marco en el que es posible desarrollar rápidamente diferentes tipos de aplicaciones en la nube. Las ofertas de plataforma como servicio comprenden un paquete de servicios administrados o componentes precodificados diseñados para acelerar el proceso de desarrollo de principio a fin. El modelo de entrega en la nube de más rápido crecimiento es PaaS, y tiene como objetivo agilizar el desarrollo de aplicaciones al proporcionar todas las capacidades necesarias para construir, implementar, probar y actualizar aplicaciones dentro de un entorno único e integrado. Es ideal para llevar las aplicaciones al mercado rápidamente, al tiempo que permite que los equipos de desarrollo se centren menos en la codificación de bajo nivel, y más en la eficiencia y la producción.

Hay autores que definen dos tipos de servicios más en la nube, como lo son Function as a Service (FaaS, por sus siglas en inglés) y Container as a Service (CaaS, por sus siglas en inglés), ambas con origen en PaaS.

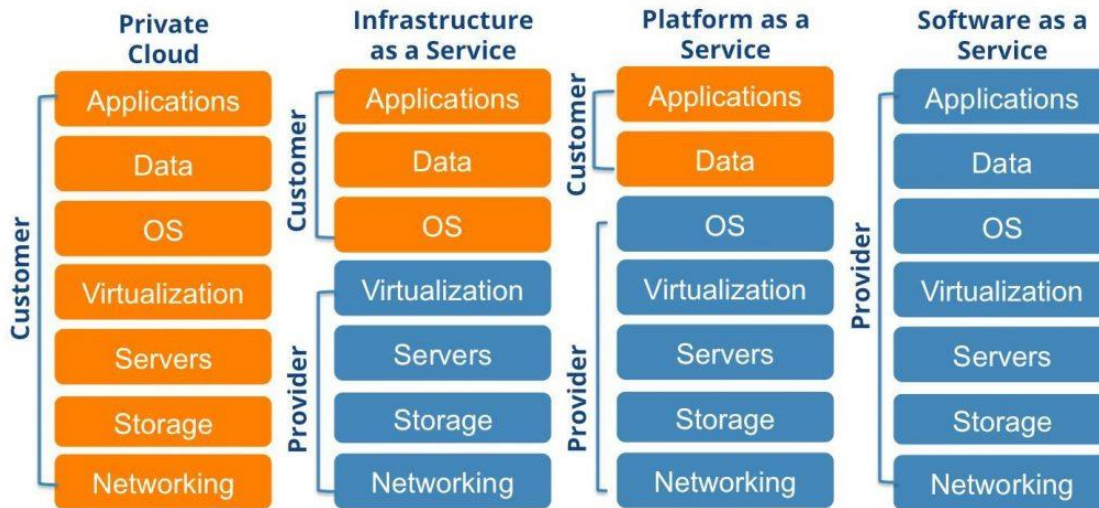


Imagen 1. RQ, (2018). Modelo de servicios de nube. Extraído de <https://www.quisted.net/index.php/2018/08/30/dc-cloud-computing/>

2.8 Hosting

Es un servicio en línea que permite a empresas e individuos publicar o habilitar aplicaciones o servicio de tecnología informáticos que sean accedidas vía Internet o vía enlaces de conexión privada web en Internet. Un hosting provee las tecnologías y servicios necesarios para hacer posible una publicación de este tipo. Se trata, básicamente, de alquilar un espacio en un servidor donde se puede almacenar todos los archivos y datos necesarios para que las aplicaciones funcionen correctamente. Es posible compararlo con IaaS, solo que no ofrece las bondades tecnológicas que ofrecen los proveedores de cómputo en la nube. El hosting nació para hospedar páginas web, principalmente.

La mayoría de los proveedores de hosting ofrecen diferentes tipos de hosting web para poder satisfacer las necesidades de diferentes clientes. Los 2 tipos de hosting más comunes son:

- Hosting Compartido: Servidor compartido con otros clientes de quien ofrece el servicio.
- Hosting VPS: Servidor dedicado.

2.9 Colocation o housing

Es un servicio en el que se ofrece alquiler de un espacio físico en un centro de datos para instalar servidores. La infraestructura tecnológica del centro de datos está especialmente acondicionada para hospedar servidores en condiciones óptimas, tanto de calidad como de conectividad.

El housing es un término utilizado en países hispanoparlantes, también se le conoce como “DcaaS”, es decir, Data Center as a Service por sus siglas en inglés.

La diferencia entre housing y un hosting, es que el hosting es infraestructura tecnológica alquilada en la que los equipos físicos son propiedad del proveedor y no de quien paga por el servicio, mientras que el housing consiste en alquilar un espacio físico en un centro de datos para albergar equipos de cómputo y la propiedad de esos equipos es del cliente que paga por el servicio.

2.10 Modelo tradicional de infraestructura de tecnología de la información

(Grupo Garatu, s.f.) señala que:

“La infraestructura TI hace referencia a un conglomerado de servicios, dispositivos físicos y aplicaciones de software que forman el sostén de los sistemas, funcionamiento de tareas y comunicaciones en cualquier compañía. Esta infraestructura sirve de soporte a la estrategia de negocios y sistemas de información de la empresa, así como en servicios más ágiles para la comunicación con los clientes”.

La plataforma de Infraestructura de TI está conformada por:

- Los componentes de hardware.
- Las funcionalidades de procesamiento de los servidores.

- Los dispositivos de almacenamiento de datos.
- Las funcionalidades de red para permitir que los datos fluyan a través de la Plataforma de Infraestructura de TI.

Para el común de los casos, las empresas inician haciendo inversiones en los componentes que se han listado anteriormente.

2.11 Modelos económicos de la infraestructura de TI tradicional

Los modelos tradicionales de TI son rígidos y complejos, que implican la adquisición de recursos sustanciales y la necesidad de gastar miles de dólares para ofrecer nuevos servicios. Bajo este modelo, para lograr implementar nuevos servicios, se debe establecer una configuración compleja, que después debe ser sometida a pruebas para luego ser implementada, este es un proceso que puede llevar semanas o meses. De la misma manera y para implementar un cambio, se deberá trabajar con diversos miembros de los diferentes equipos, con diferentes configuraciones de hardware y con herramientas de gestión dispares, todas ellas con cierta tendencia a aumentar los tiempos de implementación y configuración, con tal de mantenerse trabajando bajo un esquema controlado y tradicionalista que ha demostrado por años ser funcional.

Los modelos tradicionales no les permiten a las empresas actuar conforme el mercado va exigiendo nuevos productos, para ser competitivos hoy en día es indispensable que las empresas adopten nuevos modelos que les permita mayor dinamismo, que ayude a mejorar el tiempo que demoran en terminar un producto de tecnología consumible por el usuario final.

Lo que ocurre normalmente cuando se trabaja bajo un modelo de infraestructura de TI tradicional, es requerido adquirir equipos con capacidades superiores a lo que demanda el momento y antes de esta adquisición ya se ha pasado por un proceso espeso y complicado de selección de proveedores, evaluación de propuestas económicas, evaluación de costos y beneficios, sustentación de la inversión, sostenibilidad de la inversión en el mediano plazo, y en algunos casos se sustenta como se monetizará la inversión realizada.

Una parte importante del modelo económico tradicional es que también se requiere contemplar año tras año la renovación de contratos de soporte, mantenimiento y garantía de los equipos adquiridos.

2.12 Modelos económicos del cómputo en la nube

Existen distintos tipos de proveedores de cómputo en la nube que ofrecen distintos modelos de servicios y costos asociados. Principalmente se hará mención a los que tienen mayor adopción.

- I. Pago por consumo bajo demanda: Se paga solo por el consumo real de los recursos en la nube que se tenga contratado, por ejemplo, uso de CPU, Mb de transferencia, almacenamiento en disco, etc.

El monto de lo que se tiene que pagar se calcula por los recursos que se consuman, incluyendo el de los usuarios que usen los servicios.

Teniendo en cuenta que todos los servicios en la nube son escalables y elásticos y con costes muy ajustados, siempre se podrá pagar solo los recursos que necesiten y que se consuman.

- II. Pago por reservas de infraestructura: Las reservas de infraestructura ofrecen un descuento importante (de hasta el 75 %) en comparación con los precios de infraestructura bajo demanda.

Para las aplicaciones o servicios con estado constante o uso previsible, la opción de infraestructura reservada puede suponer un ahorro considerable. Los proveedores de nube más populares ofrecen reservas de 1 año y 3 años.

Para ambos casos también permite la opción al usuario de utilizar licenciamiento de software propio en un modelo Bring Your Own License (BYOL por sus siglas en inglés), ya sea para sistemas operativos o base de datos, por nombrar lo más común. Este modelo permite al usuario reutilizar el licenciamiento adquirido.

3 CAPÍTULO III MARCO METODOLÓGICO

3.1 Tipo de investigación

La investigación de este trabajo es bibliográfica como diseño, donde se analizó información relacionada con ellos tipos de almacenamiento que ofrece la nube Amazon Web Services, sus bondades, sus capacidades y sus costos, de la misma manera se analizó un tipo de almacenamiento comercial, costos y características, que permitió el desarrollo de este trabajo. El tipo de investigación de este trabajo es de carácter descriptivo, por el nivel de detalle ampliado que fue necesario para hacer comparaciones justas de funcionalidades de productos. (Shuttleworth, 2008) define la investigación descriptiva así:

“La investigación descriptiva es un método científico que implica observar y describir el comportamiento de un sujeto sin influir sobre él de ninguna manera. Muchas disciplinas científicas, especialmente las ciencias sociales y la psicología, utilizan este método para obtener una visión general del sujeto o tema. La investigación descriptiva es frecuentemente usada como un antecedente a los diseños de investigación cuantitativa, representa el panorama general destinado a dar algunos valiosos consejos acerca de cuáles son las variables que valen la pena probar cuantitativamente. Los experimentos cuantitativos suelen ser costosos y requieren mucho tiempo, así que es resulta razonable primero tener una idea de qué hipótesis son dignas de análisis”.

3.2 Paradigma

El paradigma sobre el cual se desarrolla el trabajo es cuantitativo; utiliza un método de análisis causal, correlaciona. Da más importancia a lo objetivo, entendido como lo medible. Este es un enfoque orientado a la comprobación, confirmatorio, reduccionista, inferencial e hipotético deductivo.

3.3 Población y muestra

El trabajo va dirigido a profesionales de tecnología que requieran algún elemento comparativo sobre la relación costo beneficio del almacenamiento de la nube vs almacenamientos tradicionales, sirve de guía para la toma de decisiones, definición de criterios y para entender como es el modelo de consumo y facturación del servicio de almacenamiento en la nube. La muestra tomada para este desarrollo; AWS Simple Storage Service y almacenamiento comercial DELL de gama media\alta, como lo es el Unity 450F.

3.4 Técnicas e instrumentos para la recolección de datos

Los instrumentos para utilizar durante esta investigación se fundamentan en información pública de fabricantes y proveedores de servicios de nube y de DELL, disponible en Internet. Adicionalmente se recolectó información relacionada al tema en distintos sitios de Internet para dar contexto a las conclusiones del trabajo.

4 CAPÍTULO IV ANÁLISIS DEL ESCENARIO

La política de almacenamiento de datos es una parte imprescindible de la administración de la información, y resulta ser una tarea compleja en la que el primer paso consiste en la identificación de los sistemas de almacenamiento de información en una empresa. Es necesario realizar un análisis detallado del entorno operativo e identificar los datos que requieren ser preservados a través de un programa de gestión del ciclo de vida de la información, para clasificar la información en función del valor que aporta al negocio, requisitos de protección, etc.

La elección del tipo de almacenamiento es clave, ya que no todos los medios de almacenamiento ofrecen el mismo nivel de seguridad, durabilidad y fiabilidad. Hay que asegurar que los datos son copiados íntegramente del origen y en unos soportes fiables y en buen estado. La política de copias de seguridad, o back-up, es otro aspecto importante de la gestión de los sistemas de almacenamiento. Por ello, se tendrán que

determinar las soluciones tecnológicas elegidas, la frecuencia de copias, el plazo de guardado, la monitorización de las aplicaciones, plazos de recuperación, etc.

A continuación, se listan los tipos de almacenamiento de información digital más populares:

- Disco Compacto y Disco Versátil Digital (CD y DVD por sus siglas en inglés): Con una vida estimada de entre 10 y 35 años, siempre que sean conservados en óptimas condiciones. Estos soportes ópticos no cuentan con suficientes capacidades para cubrir las necesidades de las organizaciones. La degradación del material con el que son fabricados y su vulnerabilidad a situaciones de humedad, altas temperaturas, también son cuestiones para tener en cuenta. No es el tipo de almacenamiento que se utilice en una organización en donde el almacenamiento y cuidado de la data es crítica. Un CD tiene una capacidad de almacenamiento de 700MB y un DVD 4.7 GB. Se estima que los discos ópticos comienzan a tener problemas a partir de los 7-10 años. Además, al pasar de los 15 años los daños son muy graves.
- Unidades de cinta magnética: es un tipo de soporte de almacenamiento de información que permite grabar datos en pistas sobre una banda de material magnético (como óxido de hierro o algún cromato). Puede grabarse cualquier tipo de información de forma digital o analógica. Las cintas magnéticas son dispositivos de acceso secuencial, pues si se quiere tener acceso al enésimo (n) bloque de la cinta, se tienen que leer antes los n-1 bloques anteriores, por lo que la recuperación y lectura de los datos en este medio suele ser muy demorada.

En 2014, Sony consiguió que una cinta magnética, que puede almacenar 148 Gbit por pulgada cuadrada o 23 Gbit por cm², permitiendo que en una cinta se almacenen 185 TB (185000 GB). En tanto en mayo de 2014 la empresa Fujifilm anunció que desarrolló un cartucho que almacena 154 TB.

Se estima que las cintas magnéticas tienen una vida útil media de 30 años, aunque podría alargarse o acortarse dependiendo de cómo cuidemos los cartuchos.

- Unidades de discos duros (*Hard Drive Disk* o HDD) son discos duros mecánicos, constituyen una opción muy popular. Tienen como principales problemas la necesidad de mantenimiento, se degradación, tienen una vida media de 5 años. Lo anterior nos hace pensar que no es la solución definitiva para el almacenamiento de la información de una organización a largo plazo. En los años 80s existían enormes discos duros con capacidades de menos de 100 megabytes (MB), ahora existen discos duros de 30 terabytes (TB) o 30.000 Gigabytes (GB).
- Unidades de estado sólido (*Solid State Drive* o SSD, por sus siglas en inglés) son discos que guardan la información en microchips con memorias flash interconectadas. Tienen bastantes ventajas con respecto a los. Son más rápidos en el procesamiento, se calientan menos, son más pequeños, pero son más costosos. Tienen problemas similares a los HDD, son similares en capacidad de almacenamiento. Se estima que tienen 15 años de vida útil.
- El almacenamiento basado en tecnologías cloud computing; es un sistema por el que la información es alojada en servidores externos a la empresa. Habitualmente, conlleva gestión de copias de seguridad y redundancia de la información en lugares geográficamente distantes para asegurar que, en caso de catástrofe, la información sea recuperable. Por esta cuestión, es una de las principales apuestas para el almacenamiento de datos en empresas. La vida útil del almacenamiento de momento es infinita, el límite de capacidad de almacenamiento es infinito y las ofertas de este tipo de servicio es muy amplia.

Luego tenemos dos tecnologías que son comúnmente utilizadas en las empresas, que se apoyan en el tipo de almacenamiento que se ha descrito anteriormente, para resguardar su data y que hacen parte de una infraestructura de TI tradicional. Muchos de los equipos que usan esta tecnología cuentan con servicios de integración a los sistemas de almacenamiento tipo nube; lo que da paso a infraestructuras híbridas, donde conviven tecnologías tradicionales de infraestructura de TI y la nube. Estas dos tecnologías son:

Network Attached Storage (NAS, por sus siglas en inglés) Un servidor NAS es un dispositivo de almacenamiento conectado a la red. A efectos prácticos la función principal de estos equipos es la de actuar como unidad de almacenamiento.

Un NAS es un computador con su propio sistema operativo y que está adaptado para estar todo el día funcionando. En ellos puedes distinguir dos conjuntos de componentes, estando por una parte lo que es el NAS en sí con su RAM, su procesador y toda su circuitería, y por otra parte los discos duros que puedes añadir a sus ranuras. Dependiendo del modelo o el fabricante, estos discos duros pueden venir incluidos cuando los compres o tendrás que comprarlos aparte. De manera que la capacidad de almacenamiento puede variar, pero tienen un máximo definido, que varía según el modelo y el fabricante.

Aunque su principal funcionalidad sea la misma, la de almacenar archivos, no debe confundirse a un NAS con los discos duros externos. Un disco duro externo es simplemente eso, una unidad de almacenamiento que conectas y desconectas, mientras que el NAS funciona más como un computador conectado a la red al que puedes acceder desde varios dispositivos, y cuyas funcionalidades puedes ampliar con aplicaciones.

Storage Area Network (SAN por sus siglas en inglés) es una red dedicada de alta velocidad que brinda acceso al almacenamiento a nivel de bloque. Las SAN se adoptaron para mejorar la disponibilidad y el rendimiento de las aplicaciones al segregar el tráfico de almacenamiento del resto de la una red local.

Las SAN permiten a las empresas asignar y administrar más fácilmente los recursos de almacenamiento, logrando mayor eficiencia. En lugar de tener capacidades de almacenamiento aisladas en diferentes servidores, se puede compartir un conjunto de capacidades entre varias cargas de trabajo diferentes y dividirse según se necesite.

Una SAN consta de hosts, conmutadores y dispositivos de almacenamiento interconectados. Los componentes se pueden conectar usando una variedad de protocolos. Fibre Channel es el protocolo de transporte original de elección. Otra opción es Fibre Channel over Ethernet (FCoE, por sus siglas en inglés), que permite a

las organizaciones trasladar el tráfico Fibre Channel a través de Ethernet existente de alta velocidad, almacenamiento convergente y protocolos IP en una sola infraestructura. Otras opciones incluyen Internet Small Computing System Interface (iSCSI por sus siglas en inglés), comúnmente utilizada en organizaciones pequeñas y medianas, e InfiniBand, comúnmente utilizada en entornos informáticos de alto rendimiento.

Tanto una SAN y como una NAS, son soluciones de almacenamiento basadas en red. Una SAN generalmente usa conectividad Fibre Channel, mientras que NAS generalmente se conecta a la red a través de una conexión Ethernet estándar. Una SAN almacena datos en el nivel de bloque, mientras que NAS accede a los datos como archivos. Para un SO cliente, una SAN generalmente aparece como un disco y existe como su propia red separada de dispositivos de almacenamiento, mientras que NAS aparece como un servidor de archivos.

La SAN suele estar asociada al almacenamiento de datos de tipo estructurada, como las bases de datos, mientras que NAS generalmente se asocia con datos no estructurados, como videos, imágenes, archivos de voz, archivos de texto.

4.1 Estrategias de selección de almacenamiento

(Lopez, 2018) señala que:

“La variedad de oferta en temas de almacenamiento digital puede llevar a algunas empresas a escoger cualquier producto (porque es el que más le suena, o porque es el más económico) sin una adecuada reflexión sobre el uso que va a dársele. Asimismo, una vez instalado, no todas las compañías realizan las copias de seguridad de forma adecuada. Y aún son más las que no tienen un plan previsto en el caso de una pérdida de información relevante. Las fugas de información no son algo que solo ocurra a las grandes industrias del cine cuando van a lanzar un estreno. Los ciberdelincuentes salen a ‘pescar’ y lanzan ataques masivos que pueden afectar a

cualquiera. Sin olvidarnos de los accidentes que pueden deteriorar nuestros sistemas. Así que sí: es necesario contar con un plan y una estrategia”.

Parte de una estrategia es hacer una estimación del volumen de datos con el que se cuenta. El presupuesto del que se dispone también debe ser tomado en cuenta, aunque la disponibilidad de oferta hoy día permite hacer una adecuada gestión de datos a precios muy competitivos.

“También se debe considerar las normativas existentes en cuanto a protección de datos. Ese es uno de los motivos que lleva a muchas empresas a utilizar el backup online para datos que requieren especial protección, y que deben mantener una copia fuera de las instalaciones de la compañía. Por otra parte, todo departamento de tecnología de una empresa es responsable de asegurarse de conocer en todo momento varias cuestiones: qué datos se están copiando; dónde se están almacenando; con qué Tiempo Objetivo de Recuperación (RTO, por si siglas en inglés) y Punto Objetivo de Recuperación (RPO, por sus siglas en inglés) cuentan (esto marcará qué información se recuperará en caso de desastre); qué dice el plan de recuperación de desastres (Disaster Recovery o DR, por sus siglas en inglés) sobre los pasos a seguir (por todos y cada uno de los miembros de la empresa) tras un incidente grave y una pérdida de datos; ¿Cuánto se tardará en volver a la normalidad?; se tiene copias de seguridad, pero ¿Funcionan adecuadamente en el momento de proceder a reinstaurar la información?, etcétera. Todos estos cuestionamientos son fundamentales a la hora de la escogencia del tipo de almacenamiento”.

Hoy nadie pone en duda que el mundo ha cambiado y que vivimos en una era digital. En el mundo el uso de Internet permite la transformación de las industrias, las economías y abre nuevas formas para que la gente se comuniquen. El proceso de transformación digital ha pasado de ser teoría a convertirse en un hecho tangible, y la velocidad en la que está avanzando es vital para el futuro de la economía mundial.

4.2 Definición del ciclo de vida del dato

(Protección de Datos, 2019) señala que:

“Aunque existen diferentes modelos, podríamos considerar que el ciclo de vida de un dato o de la información se distribuye en seis principales, que irían desde su creación hasta su destrucción. Aunque normalmente estas fases se distribuyen secuencialmente, una tras otra, en ocasiones tras pasar la fase de creación, un archivo puede moverse entre fases sin restricciones, o incluso que no pase por todas las etapas. Estas son las etapas:

1. Creación

Implica la generación de nuevo contenido digital, o la alteración, actualización o modificación de un contenido existente.

Aunque esta etapa no hará parte de la evaluación de este trabajo, es importante señalarla.

2. Almacenamiento

Es el acto de ubicar los datos digitales en algún tipo de repositorio para su custodia y normalmente ocurre de forma prácticamente simultánea a su creación. Esta etapa suele ser la que más costo asociado tiene y genera, y dependiendo del tipo de dato que se maneje el que más aporta al negocio.

3. Uso

Los datos son visualizados, procesados, o utilizados de otro modo en algún tipo de actividad, no incluyendo su modificación.

En esta etapa el tipo de almacenamiento que se usa es indispensable, de ello depende el acceso a la data, la velocidad y calidad de entrega.

4. Compartición

La información se hace accesible a otros, puede ser a otros usuarios, clientes, proveedores, etc.

Dependiendo del tipo de almacenamiento y de políticas de uso o seguridad esta actividad puede ser sencilla o muy compleja.

5. Archivado

Los datos dejan de ser usados activamente y pasan a conservarse en un almacenamiento de largo plazo.

En este punto el dato pierde un poco de valor, dado que es poco probable que sea requerido ser accedido, por lo que es importante considerar el tipo de almacenamiento a utilizar, pues ya no se demanda criticidad.

6. Destrucción

Los datos son destruidos de forma permanente usando medios físicos o lógicos. Es en esta fase donde se cierra el ciclo de vida del dato o contenido digital, puesto que tras esta fase ya no estará accesible. El dato ha de ser completamente desechado”.

4.3 Niveles de servicio en el ciclo de vida del dato

En el mercado existen varias tecnologías de discos, se hará descripción breve de 3 de las tecnologías más usadas:

1. Serial Advanced Technology Attachment (SATA, por sus siglas en inglés): Es una tecnología que permite que los discos sean capaces de almacenar y leer grandes volúmenes de información a través de una interfaz de transferencia de datos entre discos duros, con una placa base que permite una mayor velocidad y capacidad para conectar unidades de disco a la vez. La tecnología más reciente de este tipo de disco duro tiene una velocidad de hasta 600 Megabits por segundo. Su frecuencia es de 6000 MHz. También son llamados SATA 6 Gb. Es una tecnología que se usa para conectar discos electromecánicos. En cloud descritos como Discos estándar o, simplemente, Hard Drive Disk (HDD por sus siglas en inglés).

2. Serial Attached SCSI (SAS por sus siglas en inglés): Es una tecnología, la evolución de la SATA, que permite almacenar y leer grandes volúmenes de información a altas velocidades por medio de pequeños electroimanes (también llamadas cabezas de lectura y escritura), sobre un disco recubierto de limadura magnética. Los discos vienen montados sobre un eje que gira a altas velocidades. El interior del dispositivo está totalmente libre de aire y de polvo, para evitar choques entre partículas y por ende, pérdida de datos. Los discos duros SAS tienen capacidad de almacenamiento entre 72 GigaBytes (GB, por sus siglas en inglés) hasta 2 TeraBytes (TB, por sus siglas en inglés) disco permanece girando todo el tiempo que se encuentra encendido. Este tipo de disco duro tiene una velocidad de hasta 6 Gigabits por segundo. En cloud descritos como Discos estándar o, simplemente, Hard Drive Disk (HDD por sus siglas en inglés).
3. Solid State Disk (SSD por sus siglas en inglés): Sin discos duros que son hasta diez veces más rápidos que los discos duros convencionales, tienen menos desgaste debido a que los discos SSD no tienen partes mecánicas, son sólidos. Los discos duros SSD también desprenden menos calor al alcanzar menos temperatura y no hacen absolutamente ningún ruido. Son más livianos y consumen menos energía, pueden ser hasta 10 veces más rápidos que los discos de tecnología SATA y SAS.

Para efectos de este trabajo se estará haciendo la evaluación de la evolución del costo y el nivel de servicio de cada una de las etapas por las que transita el dato. Para ello es necesario definir el tiempo en el que el dato pasa por cada una de las etapas descritas anteriormente. El volumen de datos que se generan diariamente es de 8GB.

Existe una métrica importante para medir el rendimiento de los discos, para ello se usa el acrónimo IOPS (Inputs Outputs Per Second o Entradas Salidas Por Segundo, en español). Otra métrica importante es la capacidad de transferencia de datos, llamada throughput, suele medirse en Megabits por segundo.

A continuación, se describe brevemente que ocurre con los datos y que demandan durante cada una de las etapas:

- Creación: Durante esta etapa es indispensable que el tipo de almacenamiento utilizado ofrezca un rendimiento que satisfaga la demanda de operaciones de escritura al momento de su creación. Durante esta etapa se definirá un espacio de almacenamiento que sea capaz de mantener disponibilidad de la data por volumen, lo que quiere decir que la data permanecerá almacenada en esta etapa mientras no se ocupe más del 80% de espacio dispuesto para la creación y o manipulación.
- Almacenamiento: En esta etapa el requerimiento es cumplir con la expectativa de almacenar toda la data que se genera, que demanda un rendimiento alto. Adicionalmente el espacio donde se almacenará conforme los datos vayan perdiendo vigencia y criticidad.
- Uso y compartición: Estas etapas son las más exigentes porque es cuando se requiere su consumo, cuando más usuarios están accediendo a ella o cuando los procesos de una empresa los consume para sus reportes o análisis.
- Archivado: En esta etapa es cuando los datos migran a una nueva estructura de almacenamiento de menor rendimiento y más económico.
- Destrucción: En esta etapa la data ya no tiene valor. Comúnmente, las organizaciones se apoyan en procesos enmarcados en normas internacionales para eliminar y destruir los datos.

4.4 Requerimientos

Es importante destacar que, independientemente de la tecnología, para los niveles de servicio y medidas de rendimiento se utilizaran los mismos criterios y parámetros de medición. Los parámetros descritos están relacionados con el rendimiento promedio de un tipo de almacenamiento de funcionalidades avanzadas y gama media.

A continuación, se presenta el requerimiento o estructura base usada para la comparación de los distintos niveles de servicio y costos de las tecnologías de almacenamiento.

- Creación, almacenamiento, uso y compartición: En estas 4 etapas usaremos el mismo tipo de almacenamiento. Este debe cumplir con los siguientes parámetros de rendimiento:

- Rendimiento mínimo 60.000 IOPS.
- Disponibilidad inmediata de la data generada de los últimos 3 meses. La data que tenga más de 1 año de creada ha de ser reclasificada a la siguiente etapa.
- Al menos 80 TB de almacenamiento.
- Capacidad de transferencia de al menos 1000 Megabits por segundo.
- Archivado: En esta etapa se debe cumplir con almacenaje por espacio de 10 años de histórico de data, desglosado en 1 respaldo full mensual. La data que tenga más de 10 años de almacenada deberá ser destruida.
- Destrucción: En esta etapa se contará con un proceso de eliminación de datos bajo normativa con estándar internacional.

Para el caso del almacenamiento evaluado, la data siempre permanecerá dentro del equipo evaluado. En tanto que, usando la nube, esta migrará de almacenamiento conforme la data vaya depreciándose.

4.5 De criterios de evaluación en la comparación de los distintos tipos de almacenamiento

Los criterios utilizados sobre los que se enmarca esta evaluación son:

- Costos del almacenamiento. Mensual y anual.
- Nivel de disponibilidad de la data, medida mes a mes.
- Rendimiento. IOPS y Throughput.
- Costos de soporte del fabricante.
- Costo de reemplazo de piezas dañadas.
- Complejidad de la solución.
- Actualización del almacenamiento.

4.6 Almacenamiento tradicional evaluado

Como muestra para este trabajo, se ha tomado en consideración un equipo de almacenamiento de gama media, equipo que puede ser utilizado como almacenamiento y que ofrece gran valor a sus consumidores. El equipo evaluado es el UNITY 450F, del fabricante DELL-EMC que cumple con los requerimientos exigidos.

Este sistema de almacenamiento Dell EMC Unity All-Flash se fundamenta en una arquitectura integrada para bloques, archivos y volúmenes virtuales para ambientes virtualizados con compatibilidad simultánea para los distintos protocolos nativos de conexión a almacenamiento. Cada sistema aprovecha dos procesadores de almacenamiento, conectividad completa de 12 Gb y un ambiente operativo multi-core diseñado y patentado de Dell EMC con el fin de ofrecer eficiencia y rendimiento inigualables. Se agrega capacidad de almacenamiento adicional a través de los gabinetes de arreglos de discos y, para un rendimiento adicional, hay disponibles actualizaciones de controladoras en línea.

4.6.1 Características del almacenamiento Unity 450F

A partir de este equipo de almacenamiento elegido para la comparación es posible hacer múltiples configuraciones. A continuación, se listarán los máximos posibles configurables de este equipo. Admite hasta 8 puertos de conexión de alta velocidad de 1/8/10 o 16 Gigabits por segundo, 3.9 Petabytes, o 3900 Terabytes, de almacenamiento crudo, que no es necesariamente la capacidad obtenida con la configuración ideal. El tamaño máximo de unidad lógica configurable es de 256TB, al igual que el tamaño máximo de sistema de archivo soportado. Hasta 1500 snapshots o instantáneas por volumen lógico de almacenamiento. Y hasta 260.000 IOPS, a partir de una configuración muy particular del volumen de dato. Altura: 190.8 cm (75 in); ancho: 61.1 cm (24.0 in); profundidad: 99.2 cm (39.0 in); peso vacío: 173 kg (380 lb).

Los límites que no deben traspasarse para que los equipos tengan el funcionamiento más confiable y, al mismo tiempo, obtengan un funcionamiento del centro de datos con valores razonables de eficiencia energética. En cuanto a temperatura De 18 °C a 27 °C (de 64.4 °F a 80.6 °F) con un punto de condensación de 5.5 °C (41.9 °F) a una humedad relativa del 60 % y con un punto de condensación de 15 °C (59 °F).

El fabricante define un Rango de funcionamiento permitido ampliado Durante algunos períodos del día o del año, las condiciones de entrada de los equipos pueden quedar fuera del rango permitido continuo, pero aún dentro del rango improbable ampliado. El funcionamiento de los equipos se limita a un máximo del 10 % de las horas de funcionamiento al año en este rango. De 5 °C a 10 °C y de 35 °C a 40 °C (sin incidencia de luz solar directa sobre los equipos) con un punto de condensación de -12 °C y una humedad relativa de entre el 8 % y el 85 % con un punto de condensación de 24 °C (temperatura máxima de bulbo húmedo). Fuera del rango permitido continuo (de 10 °C a 35 °C), el sistema puede funcionar con un valor mínimo de 5 °C o con un valor máximo de 40 °C para proporcionar un máximo del 10 % de las horas de funcionamiento al año. Para las temperaturas de entre 35 °C y 40 °C (de 95 °F a 104 °F), reducción de la capacidad de la temperatura máxima permitida de bulbo húmedo a 1 °C por 175 m sobre 950 m (1 °F por 319 pies sobre 3,117 pies).

Adicionalmente, el fabricante señala que existen excepciones al rango de funcionamiento permitido ampliado, y es que durante algunos períodos del día o del año, las condiciones de entrada de los equipos pueden quedar fuera del rango permitido continuo, pero aún dentro del rango excepcional ampliado. El funcionamiento de los equipos se limita a un máximo del 1 % de las horas de funcionamiento al año en este rango. De 5 °C a 10 °C y de 35 °C a 40 °C (sin incidencia de luz solar directa sobre los equipos) con un punto de condensación de -12 °C y una humedad relativa de entre el 8 % y el 85 % con un punto de condensación de 24 °C (temperatura máxima de bulbo húmedo). Fuera del rango permitido continuo (de 10 °C a 35 °C), el sistema puede funcionar con un valor mínimo de 5 °C o con un valor máximo de 45 °C para proporcionar un máximo del 1 % de las horas de funcionamiento al año. Para las temperaturas de entre 35 °C y 45 °C (de 95 °F a 104 °F), reducción de la capacidad

de la temperatura máxima permitida de bulbo húmedo a 1 °C por 125 m sobre 950 m (1 °F por 228 pies sobre 3,117 pies).

Estas condiciones mencionadas, son referenciadas debido a la gran importancia que tiene mantener estos equipos en condiciones apropiadas para que su vida útil no se vea disminuida, para lo cual es necesario contar con un equipo de enfriamiento sofisticado.

4.6.2 Configuración del almacenamiento Unity 450F

La configuración del equipo de almacenamiento es la siguiente:

- 128 GB de cache de primer nivel (L1).
- Grupo de 14 discos tipo SSD de 3.84 Terabytes en capacidad cada uno, configurados bajo un arreglo de discos tipo RAID 5, distribuidos de la siguiente manera: 12 discos tipo SSD para datos + 1 Redundancia + Host Spare. Para un total de 77.32 Terabytes en almacenamiento usables para datos, con capacidad de picos de hasta 97600 IOPS.
- Grupo de 10 discos tipo SSD de 400 Gigabytes, configurados bajo un arreglo de discos de tipo RAID 5, distribuidos de la siguiente manera: 8 discos tipo SSD de datos + 1 Redundancia + Host Spare. Para un total de 11.24 Terabytes en almacenamiento usables para datos, con capacidad de hasta 75000 IOPS.
- 8 puertos FC a 16 gigabits por segundo.
- Fuentes de Poder redundantes.
- Kit para bastionado en rack.
- Contrato de mantenimiento por 26 meses.

El costo de este equipo, incluyendo traslado a sitio, armado, configuración y puesta en marcha es de \$189.125,00.

Los detalles de este equipo han sido obtenidos a partir de una propuesta económica por parte de un partner Dell, que tiene muchos años de experiencia comercializando la marca y que nos ha pedido no ser citado en este trabajo.

COMPONENTES DE LA SOLUCIÓN – UNITY 400F

DELLEMC
PARTNER
PLATINUM



- Sistema Unity All Flash 450F:
 - Controladora redundante de 2U y 25 slots para discos de 2.5"
 - CPU: 2 x Intel 10-core, 2.2GHz por controladora
 - 128 GB de cache de primer nivel (L1)
 - Pool de 14 SSD de 3.84TB en RAID-5.
 - 12 SSD de datos + 1 Redundancia + Host Spare
 - 77,32 TB Efectivos usables para datos, 97600 IOPS
 - Pool de 10 SSD de 400GB en RAID-5.
 - 8 SSD de datos + 1 Redundancia + Host Spare
 - 11,24 TB Efectivos usables para datos, 75000 IOPS
 - 8 Puertos FC a 16 gbps
 - Fuentes de Poder redundantes
 - Kit para bastionado en rack
 - Contrato de mantenimiento por 26 meses

Imagen 2. (2020). Detalles del almacenamiento usado para la comparación en este trabajo.

RESUMEN DE INVERSIONES – UNITY 450F

DELLEMC
PARTNER
PLATINUM

Componentes	Unity 450F
Hardware / Software / Garantías	176,752
Envío de Planta	
Seguro Internacional	
Servicios Profesionales	
Sub Total	176,752
ITMBS (7%)	12,373
COSTO TOTAL SOLUCIÓN	189,125

Imagen 3. (2020). Detalles del costo del almacenamiento usado para la comparación en este trabajo.

4.6.3 Valor agregado del almacenamiento Dell Unity 450F

4.6.3.1 Software de administración:

- Unisphere: Element Manager. Software de administración del equipo de almacenamiento.
- Unisphere Central: Alertas y tablero consolidados. Software de monitoreo y alertas.
- CloudIQ: Analítica de almacenamiento basada en la nube.
- Aprovisionamiento delgado o thin provisioning. Permite dar apariencia de tener más espacio asignado disponible de lo que realmente tiene.
- Pools dinámicos. Permite técnicas avanzadas y de distribución de almacenamiento que permite una forma óptima de administración y asignación de espacio.
- Reducción de datos Compresión/deduplicación (bloques y archivos).
- Proactive Assist: Permite configurar el soporte remoto, el chat en línea, abrir una solicitud de servicios, etc.
- Calidad de servicio.
- Adaptador de Dell EMC Storage Analytics: Característica de valor para la virtualización basada en el fabricante vmware, ofrece mucho valor para el análisis de la asignación de almacenamiento.
- Archiving y organización en niveles de archivos y bloques en la nube pública o privada (Cloud Tiering Appliance).

4.6.3.2 Protocolos unificados

- Archivo: Archiva los datos en una estructura jerárquica. Los datos se guardan en archivos y carpetas y se presentan tanto al sistema que lo almacena como al sistema que lo recupera en el mismo formato.
- Bloque: es un tipo de almacenamiento de datos que se suele utilizar en entornos SAN donde los datos se almacenan en volúmenes, también denominados bloques.

- VVols: Es una funcionalidad compatible con el fabricante VMWARE que permite asignación de volúmenes de datos compatibles con múltiples protocolos de transporte.

4.6.3.3 Protección local

- Cifrado basado en la controladora (opcional) con administración de claves de autoadministración o externa,
- Copias en un punto en el tiempo locales (instantáneas/snapshots y clones delgados).
- AppSync Basic. Se encarga de orquestar y automatizar el proceso de generar y consumir aplicaciones consistentes a partir de copias de datos de producción.
- Dell EMC Common Event Enabler, Antivirus Agent, Event Publishing Agent.

4.6.3.4 Protección remota

- Replicación de bloques y archivos asíncrona nativa.
- Replicación de bloques y archivos síncrona nativa.
- Envío de instantáneas/snapshots.
- Dell EMC RecoverPoint Basic.Softwrae que permite hacer backup y restauración sobre ambientes virtualizados, permitiendo recuperación granular a través del tiempo.
- Dell EMC RecoverPoint for VMs. Permite hacer backup y restauración de máquinas virtuales.

4.7 Almacenamiento en nube evaluado

Para este trabajo se toma como referencia los servicios de almacenamiento de Amazon Web Services (AWS). Este proveedor de servicios en la nube ofrece distintos productos con distintos niveles de servicio y rendimiento.

Para satisfacer los distintos niveles de rendimiento, servicio y costos se conjugarán varios servicios de AWS para poder estar alineado a los requerimientos demandados.

Uno de los servicios es Elastic Block Storage (EBS, por sus siglas en inglés). Un EBS es un dispositivo de almacenamiento de nivel de bloque duradero que se puede adjuntar a un único equipo de cómputo. Es posible usar volúmenes de EBS como almacenamiento principal de los datos que requieren actualizaciones frecuentes, como la unidad del sistema de un servidor o el almacenamiento de una aplicación de base de datos. También se pueden utilizar con las aplicaciones de uso intensivo que realicen exploraciones de discos continuas. Los volúmenes de EBS son persistentes. Y el otro es Simple Storage Service (S3, producto definido en AWS) es un servicio de almacenamiento para Internet. Está diseñado para facilitar a los desarrolladores recursos de computación escalables basados en Web.

Los volúmenes de EBS son flexibles. En el caso de los volúmenes de la generación más nueva adjuntados a los tipos de servidores virtuales de la generación más nueva, es posible aumentar el tamaño de forma dinámica, modificar la capacidad de IOPS provisionada y cambiar el tipo de volumen de los volúmenes de producción activos.

Amazon EBS proporciona los siguientes tipos de volumen: SSD de uso general (general purpose 2, gp2 por sus siglas en inglés), SSD de IOPS provisionadas, HDD con velocidad optimizada, Cold HDD y Magnético (standard, un tipo de generación anterior). Se diferencian en las características de rendimiento y en el precio, y permiten adaptar el rendimiento y el costo del almacenamiento a las necesidades de las aplicaciones.

4.7.1 Beneficios de usar volúmenes EBS

4.7.1.1 Disponibilidad de datos

(AWS, 2019) señala:

“Cuando se crea un volumen de EBS en una zona de disponibilidad (ubicaciones definidas en AWS para que los clientes desplieguen servicios), se replica automáticamente dentro de dicha zona para evitar la pérdida de datos por errores en cualquiera de los componentes del hardware. Cuando crea un volumen; puede adjuntarse a un servidor de la misma zona de disponibilidad. Después de adjuntar un volumen, este aparece como dispositivo de bloques nativo, similar a un disco duro u otro dispositivo físico. En ese punto, el servidor puede interactuar con el volumen tal y como lo haría con una unidad local. A través del servidor se puede formatear el volumen de EBS con un sistema de archivos, como ext3 y, a continuación, instalar aplicaciones.

Un volumen de EBS se puede asociar solo a un servidor a la vez, pero se pueden adjuntar varios volúmenes al mismo servidor. Si adjunta varios volúmenes a un dispositivo que ha nombrado, puede distribuir los datos entre los volúmenes para un mayor rendimiento y velocidad de E/S.

El volumen de EBS y el servidor al que se va a adjuntar deben estar en la misma zona de disponibilidad.

Es posible obtener datos de monitorización para los volúmenes de EBS, incluidos los volúmenes de dispositivo raíz para servidores respaldados por EBS, sin cargo adicional.

4.7.1.2 Persistencia de datos

Un volumen de EBS es un almacenamiento fuera del servidor que puede persistir independientemente del tiempo de vida de este. Se continúa pagando por el uso del volumen mientras persistan los datos.

Los volúmenes de EBS que están asociados a un servidor en ejecución se pueden desasociar automáticamente del servidor con sus datos intactos cuando se apaga el

servidor si se desactiva la casilla de verificación “DeleteOnTermination” (Eliminar al terminar) al configurar los volúmenes de EBS para un servidor a través de la consola de administración de servidores. A continuación, el volumen se puede volver a adjuntar a un nuevo servidor, permitiendo una recuperación rápida. Si la casilla de verificación “DeleteOnTermination” está marcada, los volúmenes se eliminarán al dar de baja al servidor donde están adjuntados. Si se utiliza un servidor respaldado por EBS, se puede detener y reiniciar dicho servidor sin que ello afecte a los datos almacenados en el volumen adjunto. El volumen permanece adjunto durante todo el ciclo de detención e inicio. Esto le permite procesar y almacenar los datos en el volumen de forma indefinida, utilizando solo los recursos de procesamiento y almacenamiento cuando sea preciso. Los datos persisten en el volumen hasta que este se elimine explícitamente. El almacenamiento de bloques físico utilizado por los volúmenes de EBS eliminados sobrescribe con ceros antes de asignarse a otra cuenta. Si se trabaja con información confidencial, es posible cifrar los datos manualmente o almacenarlos en un volumen protegido por Cifrado de Amazon EBS.

De forma predeterminada, el volumen raíz de EBS que se crea y adjunta a un servidor durante su configuración se elimina cuando se da de baja dicho servidor. Es posible modificar este comportamiento cambiando el valor de la marca “DeleteOnTermination” a falso al desplegar un servidor. Este valor modificado hace que el volumen persista incluso después de dar de baja el servidor y permite adjuntar el volumen a otro servidor.

De forma predeterminada, los volúmenes de EBS adicionales que se crean y adjuntan a un servidor durante su configuración y despliegue no se eliminan cuando se da de baja dicho servidor. Es posible modificar este comportamiento cambiando el valor de la marca “DeleteOnTermination” a verdadero al desplegar un servidor. La modificación de este valor provoca que los volúmenes se eliminen cuando al servidor se le da de baja.

4.7.1.3 Cifrado de datos

Para un cifrado de datos simplificado, es posible crear volúmenes de EBS cifrados con la característica Cifrado de Amazon EBS. Todos los tipos de volúmenes de EBS

admiten el cifrado. Es posible utilizar volúmenes de EBS cifrados para satisfacer numerosas necesidades de cifrado de datos en reposo para datos y aplicaciones regulados/auditados. Cifrado de Amazon EBS utiliza algoritmos Advanced Encryption Standard de 256 bits (AES-256, por sus siglas en inglés) y una infraestructura de claves administrada por Amazon. El cifrado se realiza en el servidor físico que aloja el servidor virtual, lo que proporciona cifrado de los datos en tránsito entre el servidor virtual y el almacenamiento de Amazon EBS.

Cifrado de Amazon EBS utiliza claves maestras de AWS Key Management Service (AWS KMS, por sus siglas en inglés) al crear volúmenes cifrados y cualquier instantánea (copia de dato en formato virtualizado) creada a partir de ellos. La primera vez que cree un volumen de EBS cifrado en una región, se crea automáticamente una clave maestra. Esta clave se utiliza para Cifrado de Amazon EBS a menos que seleccione una clave maestra del cliente (CMK, por sus siglas en inglés) que haya creado por separado con AWS KMS. Crear su propia CMK, AWS ofrece una mayor flexibilidad, incluida la opción de crear, rotar, desactivar y definir los controles de acceso y auditar las claves de cifrado utilizadas para proteger los datos.

4.7.1.4 Instantáneas o Snapshots

Amazon EBS ofrece la posibilidad de crear instantáneas (copias de seguridad) de cualquier volumen de EBS y guardar una copia de los datos del volumen en Amazon Simple Storage Service (S3, según AWS), donde se almacena de forma redundante en varias zonas de disponibilidad. No es necesario adjuntar el volumen a un servidor en ejecución para tomar una instantánea. Mientras se sigan grabando datos en un volumen, es posible crear periódicamente una instantánea del volumen para usarla como referencia para nuevos volúmenes. Estas instantáneas se pueden utilizar para crear varios volúmenes de EBS nuevos o mover volúmenes por diferentes zonas de disponibilidad. Las instantáneas de volúmenes de EBS cifrados se cifran automáticamente.

Cuando se crea un nuevo volumen de EBS a partir de una instantánea, es una copia exacta del volumen original en el momento en que se tomó la instantánea. Los volúmenes de EBS que se restauran a partir de instantáneas cifradas se cifran

automáticamente. De manera opcional, si se especifica una zona de disponibilidad distinta, puede usar esta funcionalidad un volumen duplicado en dicha zona. Las instantáneas se pueden compartir con cuentas específicas de AWS o ser públicas. La creación de instantáneas implica cargos en Amazon S3, en función del tamaño total del volumen. Para las instantáneas sucesivas del volumen, solo se cobrará por cualquier dato adicional por encima del tamaño total del volumen.

Las instantáneas son backups incrementales, lo que significa que solo se guardan los bloques del volumen que han cambiado después de la instantánea más reciente. Si tiene un volumen con 100 GigaBits de datos, pero solo 5 GigaBits han cambiado desde su última instantánea, solo se graban en Amazon S3 los 5 GigaBits de datos que se han modificado. Aunque las instantáneas se guarden de forma incremental, su proceso de eliminación está diseñado para que solo tenga que retener la instantánea más reciente para restaurar el volumen.

Para categorizar y administrar los volúmenes y las instantáneas, pueden etiquetarse con los metadatos que se elijan.

4.7.1.5 Flexibilidad

Los volúmenes de EBS admiten cambios de configuración activos mientras se encuentran en producción. Es posible modificar el tipo de volumen, el tamaño del volumen y la capacidad de IOPS sin interrupciones del servicio.”

4.7.2 Beneficios de usar simple storage service – S3

(AWS, 2019) señala:

“Amazon S3 cuenta con diversas características que puede utilizar para organizar y administrar sus datos a fin de permitir casos de uso específicos, obtener rentabilidad, reforzar la seguridad y satisfacer los requisitos normativos. Los datos se almacenan como objetos dentro de recursos llamados “buckets” y un solo objeto puede tener un tamaño de hasta 5 terabytes. Entre las características de S3 se incluyen capacidades para anexar etiquetas de metadatos a los objetos, mover y almacenar datos entre los

tipos de almacenamiento de S3, configurar y aplicar controles de acceso a datos, proteger los datos frente a usuarios no autorizados, ejecutar análisis de big data y monitorizar los datos en los niveles de objeto y bucket.

4.7.2.1 Administración del almacenamiento

Los nombres de los buckets de S3, prefijos, etiquetas de objetos y S3 Inventory ofrecen diversas formas de clasificar los datos en categorías, también es posible configurar otras características de S3 para realizar acciones. S3 Batch Operations facilita la administración de datos en Amazon S3 a cualquier escala, independientemente de que sean miles de objetos (un objeto puede ser un archivo) o miles de millones los que se almacenan. S3 Batch Operations le permite copiar objetos entre buckets, reemplazar conjuntos de etiquetas de objetos, modificar los controles de acceso y restaurar objetos archivados desde Amazon S3 Glacier, con tan solo unos clics en la consola de administración de Amazon S3.

Amazon S3 también admite características que ayudan a mantener el control de versiones de los datos, impedir el borrado accidental y replicar datos en las mismas regiones de AWS o en otras.

Con la replicación de S3, puede replicar objetos (así como sus metadatos y etiquetas de objeto respectivos) en otras regiones de AWS o en las mismas para lograr una latencia reducida, conformidad, seguridad, recuperación de desastres y muchos otros casos de uso. La replicación entre regiones de S3 se configura en un bucket de S3 de origen y replica los objetos en un bucket de destino de otra región de AWS. La replicación en la misma región de S3 replica objetos entre buckets dentro de la misma región. El control de tiempo de replicación de Amazon S3 (S3 RTC, según AWS) ayuda a cumplir requisitos de conformidad para la replicación de datos mediante el suministro de un acuerdo de nivel de servicio y visibilidad en relación con los tiempos de replicación.

También es posible aplicar políticas de escritura única y lectura múltiple (WORM, por sus siglas en inglés) con S3 Object Lock. Esta característica de administración de S3 bloquea la eliminación de una versión de objeto durante un período de retención

previamente definido, lo que le permite aplicar políticas de retención como una capa adicional de protección de datos o para cumplir obligaciones de conformidad.

4.7.2.2 Monitorización del almacenamiento

Es posible utilizar las características de S3 y otros servicios de AWS para monitorizar y controlar la forma en que se utilizan los recursos de S3. Se pueden aplicar etiquetas a los buckets de S3 para asignar costos en múltiples dimensiones de negocio (por ejemplo, centros de costos, nombres de aplicación o propietarios) y, después, utilizar los informes de asignación de costos de AWS para ver el uso y los costos que agregan las etiquetas de bucket. También se puede utilizar Amazon CloudWatch para realizar un seguimiento del estado operativo de los recursos de AWS y configurar alertas de facturación que se envíen cuando los cargos estimados alcancen el umbral definido por el usuario. Otro servicio de monitorización de AWS es AWS CloudTrail, que hace un seguimiento de las actividades de nivel de bucket y objeto e informa sobre ellas. Se pueden configurar notificaciones de eventos de S3 para activar flujos de trabajo y alertas, así como invocar a otros servicios en AWS cuando se realice un cambio específico en los recursos de S3. Las notificaciones de eventos de S3 se pueden utilizar para transcodificar automáticamente archivos multimedia al cargarlos en Amazon S3, procesar archivos de datos a medida que están disponibles o sincronizar objetos con otros almacenes de datos.

4.7.2.3 Clases de almacenamiento S3

En Amazon S3, se pueden almacenar datos con distintas clases de almacenamiento de S3: S3 Estándar, S3 Intelligent-Tiering, S3 Estándar - Acceso poco frecuente (S3 Standard-IA), S3 Única zona – Acceso poco frecuente (S3 One Zone-IA), Amazon S3 Glacier (S3 Glacier) y Amazon S3 Glacier Deep Archive (S3 Glacier Deep Archive).

Cada una de las clases de almacenamiento de S3 admite un nivel de acceso a datos específico con su costo correspondiente. Esto significa que se pueden almacenar los datos de producción críticos en S3 estándar para un acceso frecuente, almacenar los datos de acceso poco frecuente en S3 Standard-IA o S3 One Zone-IA para guardar los costos y archivar datos al precio más reducido en los tipos de almacenamiento de archivado (S3 Glacier y S3 Glacier Deep Archive). Es posible utilizar el análisis de

clases de almacenamiento en S3 para monitorizar patrones de acceso entre los objetos y detectar datos que deben cambiarse a tipos de almacenamiento de menor costo. Después, se puede utilizar esta información para configurar una política de ciclo de vida de S3 que realice la transferencia de datos. Las políticas de ciclo de vida de S3 también pueden utilizarse para provocar el vencimiento de los objetos al final de su ciclo de vida. Para almacenar datos con patrones de acceso cambiantes o desconocidos, se pueden utilizar las capas inteligentes de S3, que mueven los datos automáticamente basándose en patrones de acceso cambiantes entre una capa de acceso frecuente y otra de acceso poco frecuente de menor costo que le permite ahorrar.”

Tabla 1. Comparativa de costos y rendimiento de los distintos niveles de almacenamiento s3.

	S3 Estándar	S3 Estándar – Acceso poco	S3 Única zona – Acceso poco	S3 Glacier	S3 Glacier Deep Archive
Diseñado para ofrecer durabilidad	999999999999% (11 nueves)	999999999999% (11 nueves)	999999999999% (11 nueves)	999999999999% (11 nueves)	999999999999% (11 nueves)
Diseñado para ofrecer disponibilidad	99,99%	99,9%	99,5 %	99,99%	99,99%
SLA de disponibilidad	99,9%	99%	99%	99,9%	99,9%
Zonas de disponibilidad	≥3	≥3	1	≥3	≥3
Cargo mínimo de capacidad por objeto	N/D	128 KB	128 KB	40 KB	40 KB
Cargo mínimo por duración de almacenamiento	N/D	30 días	30 días	90 días	180 días
Tarifa de recuperación	N/D	por GB recuperado	por GB recuperado	por GB recuperado	por GB recuperado
Latencia de primer byte	milisegundos	milisegundos	milisegundos	minutos u horas seleccionados	horas seleccionadas
Tipo de almacenamiento	Objeto	Objeto	Objeto	Objeto	Objeto
Transiciones del ciclo de vida	Sí	Sí	Sí	Sí	Sí

Fuente <https://aws.amazon.com/es/s3/faqs/>

4.8 Análisis comparativo de rendimiento de los distintos tipos almacenamiento evaluados

El ciclo de vida de la data que yace en el almacenamiento UNITY 450F no migra a otro almacenamiento, la data permanecerá a lo largo de los 10 años en el mismo equipo, esto a efectos de esta comparación. A diferencia de la data que yace en la nube, que

durante el primer año se mantiene en el volumen de almacenamiento EBS y luego migra a AWS S3 Deep Glacier.

Durante el análisis se logró determinar algunas limitantes importantes de señalar. Una de ellas está relacionada con la capacidad de IOPS que ofrece el almacenamiento UNICYT 450F y el que ofrece AWS. Mientras el UNITY 450F ofrece un máximo de 97600 IOPS en el arreglo de discos más rápido y de hasta 60000 para el arreglo de discos con menor desempeño, este IOPS se refiere al límite que ofrece el arreglo de discos en sí, lo que se traduce en que todos los volúmenes de discos que sean generados en ese arreglo de discos la sumatoria de IOPS entre todos esos volúmenes no podrá ser mayor al límite establecido. En el caso de AWS estos límites son por volumen, cada volumen podrá lograr un máximo de 64000 IOPS, en otras palabras, si se provisionaran dos discos en AWS cada uno de ellos tendría la capacidad independiente de proveer 64000 IOPS, pero por otro lado en AWS existe una limitante por máquina virtual de un máximo posible para soportar hasta 80000 IOPS. Para el cálculo de los costos de AWS se ha de considerar mantener al tope los IOPS el 100% del tiempo, que no se ajusta a un escenario real.

Para los primeros 11 Terabytes de almacenamiento, este es el performance:

Tabla 2. Comparativa de rendimiento para los primeros 11 TB de almacenamiento.

Almacenamiento	IOPS RAID 5	Read IOPS Por Disco	Write IOPS Por Disco	Throughput
UNITY 450F	97,600	N/A	N/A	2,100MB/S
EBS SSD	N/A	64,000	64,000	1,000MB/S
AWS S3 Deep Glacier	N/A	Lo determina la conectividad	Lo determina la conectividad	Lo determina la conectividad

Fuente: González, Axel (2020)

Para el almacenamiento que sobrepasa los 11 Terabytes y que llegan hasta los 89 Terabytes, este es el performance:

Tabla 3. Comparativa de rendimiento luego de los primeros 11 TB de almacenamiento

Almacenamiento	IOPS RAID 5	Read IOPS Por Disco	Write IOPS Por Disco	Throughput
UNITY 450F	75,000	N/A	N/A	750 MB/S
EBS SSD	N/A	64,000	64,000	1,000MB/S
AWS S3 Deep Glacier	N/A	Lo determina la conectividad	Lo determina la conectividad	Lo determina la conectividad

Fuente: González, Axel (2020)

Al hacer la comparativa respecto al tiempo, retención de 1 año y luego a 10 años, los resultados en cuanto rendimiento se mantienen según los cuadros anteriores.

A simple vista pareciera que el UNITY 450F ofrece un mayor rendimiento de IOPS para todos los casos, sin embargo, el performance que ofrece está asociado al equipo en sí, lo que quiere decir que independientemente de la cantidad de volúmenes de discos que este tenga en uso el límite sigue estando definido por los IOPS que se muestran el recuadro. En cambio, los IOPS que ofrece EBS SSD de AWS están asociados al volumen como tal, es decir, ese rendimiento puede ofrecerlo por cada disco. Por lo que resulta ventajoso en este sentido AWS. Para el caso de AWS S3 Deep Glacier, esta data puede ser accedida condicionada al tipo de conexión TCP\IP que se tenga, porque el acceso a esta data no se hace a través de protocolos de red de almacenamiento. Existe otra limitante en EDS SSD, y es que el volumen de disco asignado no puede ser menor a 4GigaBytes y no mayor a 16 TeraBytes.

4.9 Análisis comparativo del costo de los distintos tipos de almacenamiento evaluados

Al igual que en el análisis de rendimiento, el ciclo de vida de la data que yace en el almacenamiento UNICYT 450F no migra a otro almacenamiento, la data permanecerá a lo largo de los 10 años en el mismo equipo, esto a efectos de esta comparación. A diferencia de la data que yace en la nube, que durante el primer año se mantiene en el volumen de almacenamiento EBS y luego migra a AWS S3 Deep Glacier.

4.9.1 Para los primeros 11 Terabytes de almacenamiento se tiene:

Tabla 4. Comparativa de costos para los primeros 11 TB de almacenamiento.

Almacenamiento	Hasta 50 TB/Mes	Hasta 450TB /Mes	Costo por IOPS	Costo por GB /Mes	Costo Total
UNITY 450F	N/A	N/A	N/A	N/A	\$189,125.00
EBS SSD	N/A	N/A	\$0.065	\$0.125	Depende del uso
AWS S3 Deep Glacier	\$0.023 x GB/MES	\$0.022 x GB/mes	\$0.09 x GB/mes	\$0.09 x GB/mes	Depende del uso

Fuente: González, Axel (2020)

Haciendo un ejercicio rápido, para obtener el costo del almacenamiento en nube, se fijará inversión de 26 meses, que es el tiempo de garantía y soporte del equipo UNITY 450F. Entonces, El costo de tener provisionado 11 Terabytes con el nivel de rendimiento más alto posible durante los 26 meses sería: Total de provisión de almacenamiento en GB x \$0.125 + Costo por IOPS provisionado x \$0.065. Esto daría el costo mensual. Luego se multiplicaría por 26 meses.

$$11,000 \text{ Gigabytes} \times \$0.125 + 64,000 \times \$0.065 = \$5,535.$$

$$\$5,535 \times 26 = \$692,250.$$

$$\underline{\text{Almacenamiento total UNITY 450F} = \$189,125.}$$

$$\underline{\text{Almacenamiento total EBS SSD} = \$143,910.}$$

En este ejercicio parece que la solución de AWS es inviable por la vía financiera si se considera que aún no ha sido provisionado el total de almacenamiento, sin embargo, es necesario considerar que en un escenario real la demanda de almacenamiento crece conforme pasa el tiempo, por lo que resultaría justo hacer una comparación de costos en esos términos. Más adelante se demostrará.

4.9.2 Desde los 11 Terabytes hasta los 89 Terabytes

Si se hace el mismo ejercicio que en punto anterior, resultaría en gran desventaja AWS. Haciendo el ejercicio rápido, este sería el resultado:

Tota, de provisión de almacenamiento en GB x \$0.125 + Costo por IOPS provisionado x \$0.065. Esto daría el costo mensual. Luego se multiplicaría por 26 meses.

78,000 Gigabytes x \$0.125 + 64,000 x \$0.065 = \$13,910

\$13,910 x 26 = \$361,660.

Almacenamiento total UNITY 450F = \$189,125.

Almacenamiento total EBS SSD = \$361,660.

4.9.3 Almacenamiento de data por el 1er año.

Si consideraremos el almacenamiento altamente demandante por el término de 1 año, Para el cálculo de AWS tendríamos: Total de provisión de almacenamiento en GB x \$0.125 + Costo por IOPS provisionado x \$0.065. Esto daría el costo mensual. Luego se multiplicaría por 12 meses.

11,000 Gigabytes x \$0.125 + 64,000 x \$0.065 = \$5,535.

\$5,535 x 12 = \$66,420.

Almacenamiento total UNITY 450F = \$189,125.

Almacenamiento total EBS SSD = \$66,420.

Este resultado es mucho más favorable para el almacenamiento de AWS, si consideramos que el costo del almacenamiento UNITY no es posible llevarlo a 1 año, dado que se trata de un pago del total de la inversión por la adquisición del equipo.

4.9.4 Almacenamiento de data en reposo, con más de 1 año de haber sido creada.

Tabla 5. Comparativa de costos para data almacenada durante el primer año.

Almacenamiento	Hasta 50 TB/Mes	Hasta 450TB /Mes	Costo por IOPS	Costo por GB /Mes	Costo Total
UNITY 450F	N/A	N/A	N/A	N/A	\$189,125.00
EBS SSD	N/A	N/A	\$0.065	\$0.125	Depende del uso
AWS S3 Deep Glacier	\$0.023 x GB/MES	\$0.022 x GB/mes	\$0.09 x GB/mes	\$0.09 x GB/mes	Depende del uso

Fuente: **González, Axel (2020).**

En este escenario las variables cambian un poco. Considerando que el nivel de servicio de un dato en reposo no demanda ser accedido de manera inmediata, por lo que no hay una exigencia en cuanto a operación de IOPS. El almacenamiento escogido en AWS para este tipo de dato es distinto al usado en los ejercicios anteriores. Se procede con el ejercicio rápido, según hemos realizado en los puntos anteriores.

Provisión 50 Gigabytes x \$0.023 + Provisión 28,000 Gigabytes x \$0.022. Esto daría el costo mensual. Luego se multiplicaría por 26 meses.

$50,000 \text{ Gigabytes} \times \$0,023 + 28,000 \text{ Gigabytes} \times \$0,022 = \$1,766.$

$\$1,766 \times 14 = \$24,724$

Almacenamiento total UNITY 450F = \$189,125.

Almacenamiento total AWS S3 Deep Glacier = \$24,724.

Resultado muy favorable para AWS, si consideramos que el costo del almacenamiento UNITY no es posible llevarlo a 1 año, dado que se trata de un pago del total de la inversión por la adquisición del equipo.

4.9.5 Escenario de pago por consumo

En los ejercicios anteriores se hizo provisión total de recursos, aun cuando posiblemente no se necesitaban usar en su totalidad. Como sabemos en un escenario real la data de una empresa va en constante crecimiento, para AWS es el escenario ideal el pago por consumo, pues se paga solo lo que demanda el negocio. Veamos cómo sería el comportamiento a lo largo del tiempo, durante 12 meses, considerando que se demanden un volumen de datos nuevos de 5,000 Gigabytes por mes, tomando en cuenta el ciclo de vida del dato.

A lo largo del primer año el resultado sería el siguiente:

Tabla 6. Comparativa de costos durante el primer año del ciclo de vida del dato.

Almacenamiento	Mes 1	Mes 2	Mes 3	Mes 4	Mes 5	Mes 6	Mes 7	Mes 8	Mes 9	Mes 10	Mes 11	Mes 12
UNITY 450F	\$189,125	-	-	-	-	-	-	-	-	-	-	-
EBS SSD	\$4,785.00	\$ 5,410.00	\$ 6,035.00	\$ 6,660.00	\$ 7,285.00	\$ 7,910.00	\$ 8,535.00	\$ 9,160.00	\$ 9,785.00	\$10,410.00	\$11,035.00	\$11,660.00
AWS S3 Deep Glacier	-	-	-	-	-	-	-	-	-	-	-	-

Fuente: González, Axel (2020).

Almacenamiento total UNITY 450F = \$189,125.

Almacenamiento total AWS EBS SSD = \$98,670.

En este escenario hay gran diferencia, notablemente favorable para AWS, solamente si comparamos el primer año.

4.10 Comparación de costos a lo largo del ciclo de vida del dato

Este análisis permite tener visibilidad del costo a lo largo del ciclo de vida del dato, en esta última comparación se someterá a consideración el tiempo de vigencia del equipo de almacenamiento UNITY 450F (26 meses según propuesta económica). Durante estos 26 meses analizaremos cómo evoluciona el dato en el almacenamiento de AWS.

En este ejercicio estimaremos que por mes se generan 3 Terabytes de data nueva, de manera que a lo largo de 26 meses se logre ocupar 78 Terabytes de datos. Veamos a continuación en la gráfica cómo va la curva de inversión en AWS en comparación con la inversión por adelantado del almacenamiento UNITY 450F.

La siguiente tabla resume la inversión mes a mes y el monto total de inversión a lo largo de 26 meses. El costo de AWS es 9,78% superior al del almacenamiento UNITY 450F.

Tabla 7. Pago mensual e inversión de cada solución durante el ciclo de vida del dato.

Mes	Espacio en GB ocupado	Costo EBS SSD	Costo S3 DeepGlacier	Total AWS	Total UNITY 450F
1	3000	4535	0	4535	189125
2	6000	4910	0	4910	0
3	9000	5285	0	5285	0
4	12000	5660	0	5660	0
5	15000	6035	0	6035	0
6	18000	6410	0	6410	0
7	21000	6785	0	6785	0
8	24000	7160	0	7160	0
9	27000	7535	0	7535	0
10	30000	7910	0	7910	0
11	33000	8285	0	8285	0
12	36000	8660	0	8660	0
13	39000	8660	69	8729	0
14	42000	8660	138	8798	0
15	45000	8660	207	8867	0
16	48000	8660	276	8936	0
17	51000	8660	345	9005	0
18	54000	8660	414	9074	0
19	57000	8660	483	9143	0
20	60000	8660	552	9212	0
21	63000	8660	621	9281	0
22	66000	8660	690	9350	0
23	69000	8660	759	9419	0
24	72000	8660	828	9488	0
25	75000	8660	897	9557	0
26	78000	8660	966	9626	0
				\$ 207,655.00	\$ 189,125.00

Fuente: González, Axel (2020).

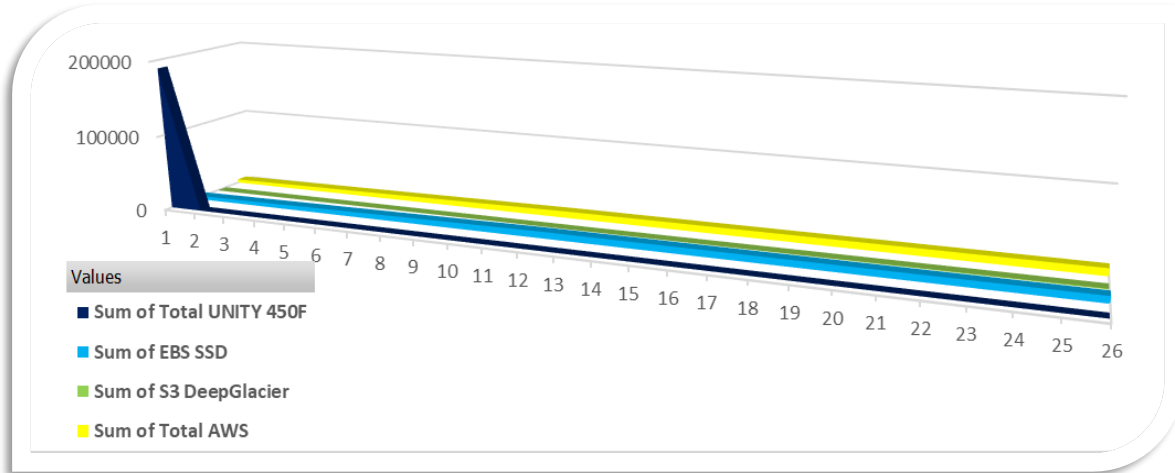


Imagen 4. González, Axel (2020). Relación de pagos mes a mes durante 26 meses.

Esta gráfica muestra la relación de pago mes a mes de los servicios. Este gráfico permite apreciar lo que significa provisionar de entrada el monto total de la inversión en comparación a poder hacerlo mes a mes.

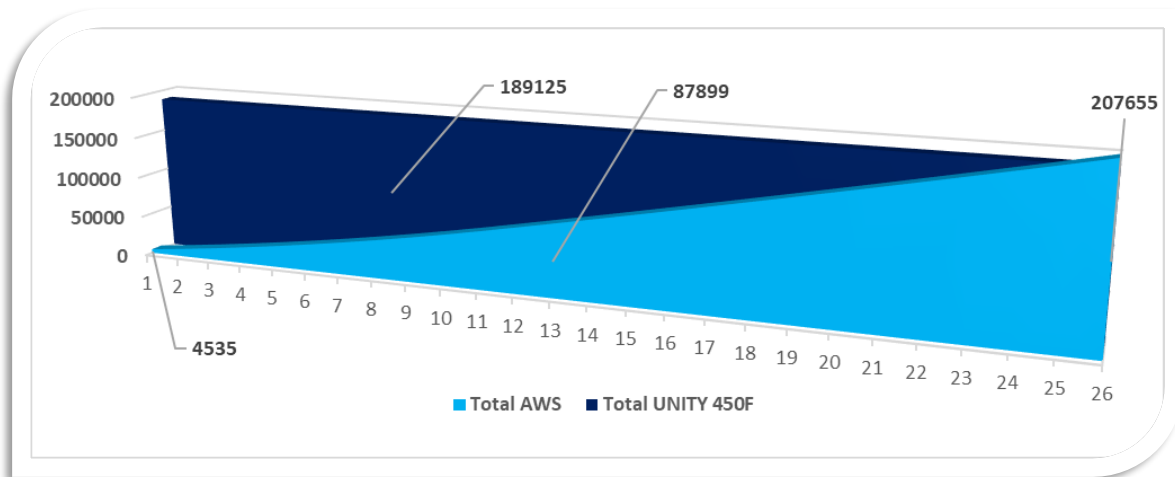


Imagen 5. González, Axel (2020). Costos acumulados a lo largo de 26 meses.

Si se considera mantener el respaldo por 10 años, el costo a razón de Gigabytes de almacenamiento en AWS sería de \$0.023 al mes. Este costo es difícil de mejorar, para cualquier solución de almacenamiento tradicional, con la tecnología comercial actual.

Otro punto importante y fundamental a evaluar es la disponibilidad de la data, o su tolerancia a fallas. En el caso del UNITY 450F cada arreglo de discos es capaz de soportar la falla de hasta dos discos para continuar operando y permitir que la data esté disponible, pero la disponibilidad de la información no solo depende del equipo, también depende de la ubicación del centro de datos, por lo que es necesario evaluar la redundancia de la alimentación eléctrica, de un banco de baterías que apoye al transfer de la carga eléctrica, en caso de que exista una falla, capacidad de estar disponible después de un siniestro, tal como un terremoto, incendio, inundación. Todas esas variables permitirían calcular la disponibilidad del servicio, sin embargo, para este ejercicio no es posible. A diferencia de AWS que, si garantiza un nivel de disponibilidad para cada uno de los almacenamientos, en el caso de EBS SSD es de 99,99% al mes. Para S3 Deep Glacier AWS garantiza la disponibilidad de la data está fijado en 99,999999999% al mes. En ambos casos, si esa disponibilidad no se cumple, AWS contempla créditos de servicios para mitigar el no cumplimiento del acuerdo.

Resulta bastante interesante hacer múltiples ejercicios de consumo y de costos, sin embargo, resulta muy complicado hacer una comparación equitativa solamente considerando que una comparación equitativa tendría que considerar más variables, por ejemplo: costos de consumo eléctrico, costo de horas hombre de instalación y puesta en marcha de los equipos, costo de conectar el equipo o servicio (caso AWS) a la red de la empresa, por ejemplo.

A nivel de costos se pudo constatar que puede resultar como producto con menor costo el almacenamiento UNITY 450F, condicionado a que no se evalúe el resto de costos asociados a la habilitación de un almacenamiento de este tipo, como lo es acondicionamiento del espacio físico para colocarlo, adecuación eléctrica, conectorización del equipo a través de switches de fibra, el cableado y el consumo eléctrico. Todo apunta a que, si se considera con detalle minucioso con estos parámetros, el costo global del UNITY 450F sería más alto.

A nivel técnico se observó que ambas soluciones pueden cubrir las expectativas en cuanto a rendimiento, sin embargo, la solución del almacenamiento de UNITY 450F puede que esté sobredimensionado o que la comparación con AWS haya sido injusta

dado que se considera entre los costos que los discos estarán operando al 100% de sus capacidades el 100% del tiempo, este escenario está muy alejado de uno real. Si bien es cierto que para AWS resulta muy provechoso el hecho de ajustar la asignación de recursos según se vayan requiriendo, también lo es para toda organización que busca optimizar costos.

4.11 Conclusiones

En General, se ha comprobado que es posible lograr ahorros significativos de costos por almacenamiento haciendo uso de la nube pública AWS en comparación con un almacenamiento de gama media/alta de la marca Dell-EMC, apalancado en un esquema de ciclo eficiente de vida de grandes volúmenes de datos. Siendo AWS una oferta muy atractiva para nuevos emprendedores y para empresas con presupuesto reducido.

Con una adecuada definición de un ciclo de vida de datos, es posible lograr ahorros significativos. Básicamente, considerando que los datos pierden la necesidad de ser accedidos con mayor frecuencia conforme pasa el tiempo, lo que quiere decir que un dato creado recientemente es más importante, desde el punto de vista de ser accedido, que un dato que ha sido creado 6 meses atrás, como ejemplo referencial. Teniendo esta definición clara, entonces es posible reclasificar el tipo de almacenamiento donde los datos reposan a almacenamientos de menor rendimiento y más económicos, entonces así se logra ahorro.

Cuando se hace la comparación de los distintos tipos de almacenamiento evaluado, es posible comparar y corroborar que a lo largo del paso de los meses existe un beneficio económico, sobre todo cuando se trata de AWS, en donde la relación de costo del almacenamiento por GB es más económico y de mayor rendimiento que uno DELL-EMC.

Durante el análisis se pudo observar que hay otro factor influyente para lograr hacer una comparación más equitativa y está relacionado con que se ha asumido que la provisión de almacenamiento será usada al 100%, en cuanto a espacio y en cuanto a rendimiento. Una gran verdad es que es muy difícil que una empresa que demande las capacidades de almacenamiento de datos ilustrado en este trabajo. Que una empresa demande 80 Terabytes de datos que requieran 60.000 IOPS de manera permanente en un solo equipo y en una sola ubicación, solo ocurría en grandes empresas de analítica de datos. Esto nos lleva a pensar que la solución del

almacenamiento en AWS estaría sobredimensionada, respecto a un requerimiento objetivo.

Otro resultado del análisis tiene que ver con es que la comparación de performance no ha sido ajustada a la necesidad del negocio sino a una “carrera” de IOPS, cuando lo que importa es lo que demanda el negocio, por lo que, ajustar una propuesta a una necesidad real de negocio correspondería, casi con total seguridad, a un punto en el que el escenario sería completamente favorable a AWS. El almacenamiento de EBS SSD de AWS es capaz de ofrecer 64000 IOPS a un disco siempre que, en tamaño este sea mayor a 1280 Gigabytes.

Toda esta cantidad de interacciones ha hecho que el crecimiento del volumen de datos haya supuesto un cambio en la infraestructura y la tecnología que los recopila. Cada vez es más fácil considerar la nube como solución, no solo por los costos altamente competitivos sino también por todo el valor agregado que hay alrededor de la nube.

4.12 Recomendaciones

La disponibilidad de la data no solo está sujeto a los equipos, también depende de otros factores, entre ellos las fallas del suministro eléctrico que se hacen cada vez más comunes y frecuentes, sobre todo en los países en vías de desarrollo. Las fallas eléctricas son sencillas de simular, basta con manipular un interruptor o breaker, por ello se recomienda que otros estudiantes desarrollen un nuevo estudio sobre la variación de la eficiencia del UNITY 450F en comparación con EBS y S3 de AWS bajo fallos del suministro eléctrico.

Se recomienda que otros estudiantes desarrollen una nueva investigación donde se compare la eficiencia del almacenamiento en nube EBS y S3 de AWS con una plataforma basada en Hewlett Packard (HP), o también, donde se compare la eficiencia de la nube AWS con una plataforma basada en Google Cloud Storage, algunos de estos servicios de nube que, a pesar de ser pagos, poseen periodos de prueba gratuitos que facilitan hacer pruebas gratuitas por periodos de tiempo limitados.

En otro orden de ideas, después de culminar esta investigación y reconocer el crecimiento de la demanda de servicios en la nube por parte de los usuarios y empresas se hace evidente recomendar que la Universidad Internacional de Ciencia y Tecnología (UNICyT) adopte una línea de investigación más específica en torno a la Computación en la nube y sus tecnologías. Esto le facilitaría a la comunidad académica de UNICyT aproximarse a las grandes corporaciones que desarrollan la tecnología, como Amazon, Google, Microsoft, IBM, HP, Oracle, entre otras para solicitar fondos para apoyar la investigación científica en el área.

5 REFERENCIAS

- Alegsa, L. (21 de Junio de 2016). *Definición de Cinta magnética (de almacenamiento de datos)*. Obtenido de http://www.alegsa.com.ar/Dic/cinta_magnetica.php
- Amazon Web Services. (2019). *Características de Amazon S3*. Obtenido de <https://aws.amazon.com/es/s3/features/?nc=sn&loc=2>
- Andrade, D. (2019). *efi Economía y Finanzas Internacionales*. Obtenido de <http://puceae.puce.edu.ec/efi/index.php/economia-internacional/14-competitividad/20-economias-de-escala>
- Asensio, R. M.-B. (s.f.). Obtenido de Universidad de Murcia: <http://dis.um.es/~barzana/enlaces/luz1.htm>
- AWS. (2019). *Amazon EBS features*. Obtenido de <https://aws.amazon.com/ebs/features/>
- AWS. (2019). *AWS*. Obtenido de <https://aws.amazon.com/es/what-is-aws/>
- AWS. (2019). *Volúmenes de Amazon EBS*. Obtenido de https://docs.aws.amazon.com/es_es/AWSEC2/latest/UserGuide/EBSVolumes.html
- Azad, A. (Septiembre de 2018). *mediate.com*. Obtenido de Mediation vs. Arbitration: <https://www.mediate.com/articles/azad-mediation-arbitration.cfm>
- B., G. (5 de Noviembre de 2019). *Hostinger*. Obtenido de <https://www.hostinger.es/tutoriales/que-es-un-hosting>
- Bednarz, A. (03 de Mayo de 2018). *networkworld*. Obtenido de <https://www.networkworld.es/networking/que-es-una-san>
- Bejerano, P. G. (s.f.). *Blog Thinking Big*. Obtenido de <https://blogthinkbig.com/como-se-fabrica-la-fibra-optica/>
- Cabaca, T. (Julio de 2018). Obtenido de Mycomputerpro: <https://www.muycomputerpro.com/2018/07/03/historia-cloud-computing>

- Conectart. (s.f.). *Conectart*. Obtenido de <https://blog.conectart.com/que-es-cloud-computing/>
- Cosmed, X. (02 de Octubre de 2016). *Conserva tus datos para siempre con M-Disc y LTO*. Obtenido de <https://lamanzanamordida.net/datos-tiempo-m-disc-lto/>
- Cristhian Leonardo Tolosa Cuadrado, J. S. (2014). *Revistas udistrital*. Obtenido de <https://revistas.udistrital.edu.co/index.php/revcie/article/download/6500/9185?inline=1>
- DELL EMC. (2019). *DELL EMC*. Obtenido de <https://corporate.delltechnologies.com/es-ar/about-us/who-we-are.htm>
- discosduros.org. (2019). *Discos duros SATA*. Obtenido de <https://discosduros.org/disco-duro-sata/>
- DiscTech. (2019). *Seagate 1200 ST400FM0053 400GB Enterprise SAS SSD*. Obtenido de <https://www.disctech.com/Seagate-1200-ST400FM0053-400GB-Enterprise-SAS-SSD>
- DiscTech. (2019). *Seagate XS3840SE10113 3.84TB SED SAS Solid State Drive*. Obtenido de <https://www.disctech.com/Seagate-XS3840SE10113-3.84TB-SED-SAS-Solid-State-Drive>
- Física, S. C. (s.f.). *fisic*. Obtenido de <https://www.fisic.ch/contenidos/ondas-y-la-luz/teor%C3%ADas-de-la-luz/>
- García, J. (7 de Noviembre de 2012). Obtenido de SEAS, ESTUDIOS SUPERIORES ABIERTOS: <https://www.seas.es/blog/automatizacion/reflexion-y-refraccion-de-la-luz-en-transmisiones-de-fibra-optica/>
- Gomez, S. (28 de Abril de 2015). *La Mediación: Solución de conflictos*. Obtenido de <https://www.youtube.com/watch?v=dlr8ncmfis>
- Grupo de alumnas de 1ªA de Comunicación Audiovisual, d. I. (17 de Abril de 2017). *Cloud Computing*. Obtenido de

<https://cloudcomputing728.wordpress.com/2017/04/27/desventajas-del-cloud-computing/>

Grupo Garatu. (s.f.). Obtenido de GaratuCloud: <https://garatucloud.com/tecnologia-informatica-infraestructura-ti-redes-virtualizacion/>,

Hewlett Packard Enterprise. (2019). *Hewlett Packard Enterprise*. Obtenido de Hewlett Packard Enterprise: <https://www.hpe.com/es/es/what-is/data-storage.html>

Kiziryan, M. (2019). *Economías de escala*. Obtenido de Economipedia haciendo fácil la economía: <https://economipedia.com/definiciones/economias-de-escala.html>

Lello, R. D. (Abril de 2016). *Radians*. Obtenido de <https://www.radians.com.ar/blog/?p=3089>

Lopez, A. (Marzo de 2018). *Infordisa*. Obtenido de <https://www.infordisa.com/es/estrategia-de-almacenamiento-ti/>,

Manuel Calvo, T. P. (2010). Principales características de la mediación.

masadelante.com. (2019). *masadelante.com*. Obtenido de masadelante.com: <https://www.masadelante.com/faqs/gigabyte>

Microsoft. (2019). *Microsoft Azure*. Obtenido de Microsoft Azure: <https://azure.microsoft.com/es-es/overview/what-is-paas/>

Microsoft. (2019). *Microsoft Azure*. Obtenido de Microsoft Azure: <https://docs.microsoft.com/es-es/azure/storage/blobs/storage-blob-storage-tiers>

Microsoft. (2020). *Microsoft Azure*. Obtenido de <https://azure.microsoft.com/es-es/overview/what-is-saas/>

Muñoz, J. C. (Octubre de 2016). *Repositorio Pontificia Universidad Católica del Ecuador*. Obtenido de <http://repositorio.puce.edu.ec/bitstream/handle/22000/13158/Javier%20Alfredo%20Castro%20Munoz.pdf?sequence=1&isAllowed=y>

- Networkworld. (2019). *Networkworld from IDG*. Obtenido de SLA: Qué son, para qué sirven: <https://www.networkworld.es/archive/sla-que-son-para-que-sirven>
- Network. (29 de Julio de 2013). *Los servicios cloud contribuyen a reducir las emisiones de carbono*. Obtenido de <https://www.networkworld.es/cloud-computing/los-servicios-cloud-contribuyen-a-reducir-las-emisiones-de-carbon>
- Oviedo, E. M. (15 de Octubre de 2015). Obtenido de Think Big / Empresas: <https://empresas.blogthinkbig.com/10-ventajas-del-cloud-computing/>
- Perez, G. (s.f.). *Espectrometria*. Obtenido de https://www.espectrometria.com/espectro_electromagnitico
- PowerData. (s.f.). *Blog Powerdata*. Obtenido de Blog Powerdata: <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/claves-de-una-estrategia-del-ciclo-de-vida-del-dato>
- Protección de Datos. (19 de Abril de 2019). *Ciclo de vida del dato*. Obtenido de <https://www.protecciondatos.org/ciclo-de-vida-del-dato/>
- Rackspace. (2019). *Rackspace*. Obtenido de Biblioteca en la nube de Rackspace: <https://www.rackspace.com/es/library/what-is-iaas>
- Rodriguez, A. (10 de Junio de 2012). *Fibra Óptica Hoy*. Obtenido de <https://www.fibraopticahoy.com/fibra-optica-que-es-y-como-funciona/>
- Salesforce Latin América. (2000). *SalesForce*. Obtenido de SalesForce: <https://www.salesforce.com/mx/cloud-computing/>
- Search Datacenter. (2019). *Search Datacenter en Español*. Obtenido de <https://searchdatacenter.techtarget.com/es/definicion/Nube-privada-nube-interna-o-nube-corporativa>
- Search Datacenter. (2019). *Search Datacenter en Español*. Obtenido de <https://searchdatacenter.techtarget.com/es/definicion/Nube-publica>

- Search Datacenter. (2019). *Search Datacenter en Español*. Obtenido de <https://searchdatacenter.techtarget.com/es/definicion/Microsoft-Azure-Windows-Azure>
- Shuttleworth, M. (2008). *Explorable*. Obtenido de <https://explorable.com/es/disenio-de-investigacion-descriptiva>
- Sol, A. S.-K. (2019). Socio LatamLex Guatemala.
- Taylor, S. (2019). *alllaw.com*. Obtenido de What are Mediation and Arbitration?: <https://www.alllaw.com/articles/legal/article9.asp>
- Tecnología hecha palabra. (2017). *Tecnología hecha palabra*. Obtenido de Tecnología hecha palabra: <http://www.tecnologiahechapalabra.com/salud/miscelanea/articulo.asp?i=600>
- Vásquez Benítez, C. A. (2018). *Pontificia Universidad Católica del Ecuador*. Obtenido de <http://repositorio.puce.edu.ec/handle/22000/15077>
- VMWARE. (9 de Septiembre de 2015). *Virtual Volumes y protocolos de almacenamiento*. Obtenido de <https://docs.vmware.com/es/VMware-vSphere/6.0/com.vmware.vsphere.storage.doc/GUID-AAA99054-4D81-49F8-9927-65E9B08577AD.html>
- Westland, S. (2001). Obtenido de Imagen digital: http://www.gusgsm.com/nanometro_nm
- Xataka. (10 de 2018). *Servidores NAS: qué son, cómo funcionan y qué puedes hacer con uno*. Obtenido de <https://www.xataka.com/basics/servidores-nas-que-como-funcionan-que-puedes-hacer-uno>

ANEXO



UNIVERSIDAD INTERNACIONAL DE CIENCIA Y TECNOLOGÍA FACULTAD DE CIENCIAS DE LA COMPUTACIÓN Y TECNOLOGÍA

INFORME DE ACTIVIDADES DE TUTORÍA OPCIÓN DE TITULACIÓN II

Estudiante: Axel Gonzalez Crespín . Cédula de identidad o pasaporte No. 8-1159-1973

Tutor: Prof. Erick A. Ramos Sánchez. Cédula de identidad o pasaporte No. 134163825

Correo electrónico del participante: axel.gonzalez@unicyt.net Celular No. 6277-3416

Título tentativo del trabajo de grado (TG) y de pasantía profesional (PEOP). Comprobación de la dependencia de la eficiencia del costo del almacenamiento de datos en la nube AWS diseño del ciclo de vida

Línea de Investigación: Ingeniería y sistemas de comunicación

SESIÓN	FECHA	HORA REUNIÓN.	ASPECTO TRATADO	OBSERVACIÓN
1.	21/09/2019	09:02 a.m.	Contacto inicial: Exploración de área de investigación; Aceptación de Tutoría	Se estableció la modalidad a distancia como método de asesoría de tesis. (del 21/09/2019 al 28/09/2019)
2.	03/10/2019	05:01 p.m.	Se exploraron posibles temas de tesis propuestas por el asesor y por el estudiante	(del 03/10/2019 al 04/10/2019). Vía email.
3.	05/10/2019	09:00 a.m.	Selección del tema de investigación y se discutió un plan preliminar	Esta reunión fue presencial en la sede de UNICYT. Aula 8
4.	15/10/2019	01:55 p.m.	Revisión de anteproyecto	Correcciones varias (del 15/10/2019 al 16/10/2019). Vía email.
5.	21/10/2019	09:48 a.m.	Revisión y aprobación de Anteproyecto	Observaciones y correcciones varias (del 21/10/2019 al 01/11/2019). Vía email.
6.	21/11/2019	03:21 p.m.	Bases teóricas	Sugerencias de bases teóricas; Observaciones y correcciones varias (del 21/11/2019 al 24/11/2019). Vía email.
7.	10/12/2019	2:56 p.m.	Aspectos metodológicos y revisión general del avance de la tesis	Observaciones y correcciones varias (del 10/12/2019 al 04/01/2020). Vía email.

8.	17/01/2020	10:15 a.m.	Revisión final de la tesis	Observaciones y correcciones varias (del 17/01/2020 al 20/01/2020). Vía email.
9.				

Título definitivo: Comprobación de la dependencia de la eficiencia del costo del almacenamiento de datos en la nube aws diseño del ciclo de vida

Comentarios finales acerca de la investigación: Declaramos que las especificaciones anteriores representan el proceso de dirección del trabajo de grado arriba mencionado.

Firma

Firma