

MÉTRICAS Y PRUEBAS DE VALIDACIÓN PARA SISTEMAS DE ALMACENAMIENTO

Ing. Freddy A. Poll García¹, Ing. Alejandro Santoyo González², Msc. Lilia Rosa García Perellada³,
Dr.C. Alain Abel Garófalo Hernández⁴

¹Dirección de Tecnologías y Sistemas (DTS), Avenida 5ta B y 66, ²SOLINTEL S.A., Calle 1ra Esq. B, ³Departamento de Telecomunicaciones y Telemática. Facultad de Ingeniería Eléctrica, Instituto Superior Politécnico José Antonio Echeverría (ISPJAE), Calle 114, No. 11901. e/ Ciclovía y Rotonda, Marianao, La Habana,

⁴Departamento de Telecomunicaciones y Telemática. Facultad de Ingeniería Eléctrica, ISPJAE

¹fpoll@ds.mn

²e-mail: santoyo1288@gmail.com

³e-mail: lilianrosa@electronica.cujae.edu.cu

⁴e-mail: aagarofal@gmail.com

RESUMEN

El Almacenamiento en Nube es uno de los servicios de infraestructura que se puede brindar en una Nube Privada, requiriendo la interoperabilidad entre las tecnologías de los Sistemas de Almacenamiento para que los usuarios puedan compartir sus datos. Por esta y otras razones, en una época donde ya se comienza a hablar de términos como almacenamiento compartido entre Nubes Privadas, se hace tan importante definir un conjunto de pruebas que permitan validar las potencialidades de la gran variedad de modelos de Sistemas de Almacenamiento que existen, independientes de las tecnologías subyacentes, así como definir un conjunto de métricas con sus respectivos umbrales que permitan monitorizar el estado de dichos sistemas. Sin embargo, a pesar de los avances realizados, aún no se cuenta con un conjunto de pruebas de validación estandarizadas, ni se han definido métricas para evaluar las potencialidades de los Sistemas de Almacenamiento en general. Dada la situación problemática anterior se trazó el objetivo de proponer el conjunto de métricas y pruebas de validación para evaluar el Sistema de Almacenamiento de una Nube Privada que brinde Infraestructura como Servicio. El resultado obtenido fue la propuesta de un conjunto de pruebas de validación, métricas y sus umbrales para la validación del desempeño de un Sistema de Almacenamiento.

PALABRAS CLAVES:Sistema de Almacenamiento, Nube Privada, Infraestructura como Servicio, Métrica, Pruebas.

ABSTRACT

Standardization in the area of cloud storage systems is a project that is in full swing involving the participation of many leading international organizations. In the current scenario, developing performance evaluation tests, transparent to the implemented underlying technology is a step that is difficult. This paper is a study of current standards in regard to the metrics and validation tests for modern and traditional cloud storage systems developed by major standardization organizations like SNIA, ITU, NIST, DMTF and some other leading Cloud vendors such as Microsoft, Oracle, VMware, OpenNebula and CloudStack.

KEYWORDS:Metric, Standard, Performance, Throughput.

INTRODUCCIÓN

El Almacenamiento en Nube [1][2] es uno de los servicios de infraestructura que se puede brindar en una Nube Privada. Actualmente existen entidades conocidas como Proveedores de Servicios de Almacenamiento (SSP1) que dedican su Nube Privada exclusivamente a la prestación de este servicio, que debe permitir que los usuarios puedan compartir sus datos. Lo anterior es logrado mediante la interoperabilidad entre las tecnologías de los Sistemas de Almacenamiento subyacentes.

Por otra parte, para garantizar que los Sistemas de Almacenamiento empleados ofrezcan elevados niveles de disponibilidad, seguridad, desempeño, así como otros requerimientos técnicos que aseguran la calidad de los servicios de datos, es imprescindible definir un conjunto de métricas y pruebas de validación, que permitan tener el control sobre las características funcionales y operativas de estos sistemas y que certifiquen sus capacidades reales. Actualmente esto no existe de manera formal y única, sino que cada fabricante o proveedor propone métricas y pruebas de validación que se ajustan a su equipamiento y tecnologías. Otras organizaciones en cambio, limitan su alcance a determinadas categorías como disponibilidad o desempeño. Por esta y otras razones, en una época donde ya se comienza a hablar de términos como almacenamiento compartido entre Nubes Privadas (IPCS2), se hace tan importante definir un conjunto estandarizado de pruebas que permitan validar las potencialidades de la gran variedad de modelos de Sistemas de Almacenamiento que existen, independientemente de las tecnologías subyacentes, así como definir un conjunto estandarizado de métricas con sus respectivos umbrales que permitan monitorizar el estado de dichos sistemas.

Uno de los principales aportes de este proceso de estandarización lo constituye la incorporación de elementos que permitan discriminar de forma eficiente entre diferentes propuestas de soluciones de almacenamiento, decisión que debe estar basada y sólidamente fundamentada en resultados obtenidos de procesos de evaluación y certificación de las mismas. La importancia de contar con un conjunto de métricas y sus respectivos umbrales que permitan emitir un criterio verídico, de forma tal que sea independiente a la solución que se está evaluando, es muchas veces subvalorada, pero la realidad es que este aspecto es un elemento vital.

Este trabajo presenta una propuesta de métricas y pruebas de validación para evaluar el Sistema de Almacenamiento de una Nube Privada con soporte para Infraestructura como Servicio (IaaS3), independiente de la tecnología subyacente. En las pruebas fueron definidos los objetivos, resultados esperados, herramientas de Software Libre y Código Abierto (SLCA) y el método para la ejecución de las pruebas. La propuesta permite valorar el diseño de los Sistemas de Almacenamiento de Nubes Privadas y gestionar proactivamente dicho bloque funcional.

¹ Siglas del inglés "Storage Service Provider".

² Siglas del inglés "Inter-Private Cloud Storage".

³ Siglas del inglés "Infrastructure as a Service".

MÉTRICAS DE ALMACENAMIENTO

Para obtener la propuesta de las métricas, y sus respectivos umbrales, se realizó un estudio de las métricas empleadas por algunos de los principales proveedores de servicios en la Nube como Opennebula, Cloudstack, el ElasticUtility Computing Architecture for Linking your Programs to Useful Systems (Eucalyptus), Microsoft, Vmware y el International Business Machines (IBM). También se analizaron las recomendaciones de organizaciones de estandarización como el Instituto Nacional de Estándar y Tecnología (NIST4), la Unión Internacional de Telecomunicaciones (UIT), la Distributed Management TaskForce (DMTF) y la Asociación de Industrias del Almacenamiento en Red (SNIA5) [3][4][5][6].

La DMTF y el NIST no aportaron referencias sólidas que indiquen la existencia de métricas recomendadas para medir el rendimiento de Sistemas de Almacenamiento, sin embargo, a pesar de tampoco encontrar un conjunto de métricas estandarizadas en la SNIA, la documentación asociada al software hIOMon resultó relevante [7]. En el caso de la UIT el aporte fundamental estuvo centrado en las cuatro categorías de métricas que deben ser usadas para evaluar cualquier Sistema de Almacenamiento (Storage Capacity Metrics, Virtual Machine (VM)/Virtual Disk Utilization Metrics, Snapshot Monitoring, In/Out (I/O) Performance metrics) [8], lo cual constituyó un significativo aporte a esta investigación. En el análisis de los proveedores de soluciones de almacenamiento, se encontraron útiles referencias en todos los casos que se indagaron, con excepción de IBM. Fueron detectadas como métricas más empleadas para monitorizar el estado general de estos sistemas: el Número de Operaciones de Lectura/Escritura (L/E) en un Segundo (IOPS6), el ancho de banda efectivo, el throughput, la latencia, la disponibilidad y las métricas de utilización de diferentes recursos tanto físicos como virtuales. Se apreció además, tanto en el análisis de los proveedores como en las recomendaciones, una marcada tendencia a agrupar las métricas según el componente del Sistema de Almacenamiento para el cual están diseñadas [9], lo cual es sumamente útil debido a que con esto se pretende cubrir las principales particularidades de cada uno de los tipos de soluciones de almacenamiento. [10]

Propuesta de métricas para un Sistema de Almacenamiento de una Nube Privada.

La propuesta del conjunto de métricas concebida por los autores del presente trabajo se compone por un conjunto de parámetros para caracterizar la configuración de un Sistema de Almacenamiento y cinco categorías de métricas divididas por: métricas para monitorizar el estado del conjunto de recursos, métricas para monitorizar el estado de los nodos de un clúster, métricas para monitorizar el estado y desempeño de los Sistemas de Ficheros (FS)7, métricas para medir el desempeño de volúmenes lógicos y métricas para medir el desempeño de los nodos de un clúster; en función del estado del arte estudiado. A su vez en cada una de las categorías anteriores las métricas han sido clasificadas en tres categorías: desempeño, disponibilidad y escalabilidad. Las métricas englobadas en la categoría de desempeño abarcan las siguientes subcategorías: capacidad, utilización, throughput, tiempo de

⁴ Siglas del inglés "National Institute of Standards and Technology".

⁵ Siglas del inglés "Storage Networking Industry Association".

⁶ Siglas del inglés "In/Out Operations per Second".

respuesta y exactitud. Esta última categorización permite la toma de decisiones en función de los requerimientos técnicos que se deseen.

En toda solución de almacenamiento es de vital importancia conocer la capacidad para guardar información, por lo que la subcategoría de capacidad agrupa un total de seis métricas que permiten evaluar el estado de las capacidades. La subcategoría de utilización, sumamente útil para los proveedores y los usuarios que arriendan capacidades de almacenamiento, engloba un total de 24 métricas que permiten conocer en todo momento el estado de utilización de los recursos de almacenamiento, teniendo en cuenta que este es otro de los puntos fundamentales durante la explotación de un Sistema de Almacenamiento.

Independientemente de la tecnología que se esté utilizando para desplegar la solución de almacenamiento, los valores que aportan las métricas de transferencias, englobadas en la subcategoría de throughput, constituyen un criterio importante ya que son estas las que brindan el conocimiento de cuán rápido puede un usuario interactuar con el sistema, guardar y obtener su información. Además existen aplicaciones para las cuales se deben cumplir otros requerimientos más estrictos como muy bajas demoras en las transferencias, valores de demoras sostenidos, por lo que se incluye la subcategoría de tiempo de respuesta.

Se incluye en la propuesta la categoría disponibilidad. Este engloba un total de tres métricas, permitiendo caracterizar un Sistema de Almacenamiento desde el punto de vista de la confiabilidad y el aseguramiento de que los valores de tiempo del acceso a los datos se encuentren entre los límites definidos por el cliente. También es muy importante para los proveedores conocer el límite de usuarios que puede soportar su sistema antes de que este comience a degradar su rendimiento debido a la concurrencia de operaciones. Para cubrir este punto se incluyó la categoría de escalabilidad, que abarca un total de seis métricas. Las Tablas 1, 2, 3, 4 y 5 muestran un subconjunto de las métricas propuestas por componentes del Sistema de Almacenamiento.

Tabla 1: Métricas para monitorizar el estado del almacén de recursos. [10]

Categoría	Subcategoría	Clasificación UIT	Métrica	Descripción	Respaldada por
Rendimiento	Capacidad	Storage Capacity Metrics	POOL_TOTAL	Espacio de almacenamiento total con que cuenta un almacén de recursos.	CloudStack
			GRUPOS	Número de grupos de volúmenes que hay creados en un almacén de recursos.	Autores
	Utilización		VOLÚMENES_XFS	Número de volúmenes compartidos mediante protocolos XFS.	Autores
			VOLÚMENES_SCSI	Número de volúmenes asignados mediante almacenamiento en bloques.	Autores
Escalabilidad			INCREMENTO	Tamaño mínimo de las porciones de aprovisionamiento del sistema para satisfacer el incremento de la demanda. Generalmente se corresponde con la	Autores

Disponibilidad			capacidad de los volúmenes más pequeños.	
		FACTOR Fm	Indica qué fracción del tiempo total de trabajo se utilizó en el dispositivo mejorado.	Autores
		MTBF	Intervalo de tiempo promedio entre dos fallas consecutivas sufridas por el sistema.	Autores
		MTTR	Intervalo de tiempo promedio necesario para reparar las fallas ocurridas en el sistema.	Autores

Tabla 2: Métricas para monitorizar el estado de los nodos de un clúster de una nube. [10]

Categoría	Subcategoría	Clasificación UIT	Métrica	Descripción	Respaldada por
Rendimiento	Capacidad		DISCO_NODO_TOTAL	Espacio de almacenamiento asignado por el sistema a un nodo del clúster.	Opennebula, CloudStack, VMware.
		VM/Virtual Disk Utilization Metrics	MV	Número de Máquinas Virtuales (MV) que hospeda un nodo específico.	Opennebula
	Utilización		DISCO_NODO_% ASIGNADO	Porcentaje del espacio de almacenamiento de un nodo que ha sido asignado a las MV que hospeda. Esta métrica debe generar señales de alerta al superar el umbral del 98%.	CloudStack, VMware
			DISCO_NODO_% OCUPADO	Porcentaje del espacio de almacenamiento de un nodo que se encuentra realmente ocupado. Se deben generar alertas de atención al sobrepasar el umbral del 85% y alertas de estado crítico al sobrepasar el 95%.	Oracle, Opennebula, VMware, CloudStack

Tabla 3: Métricas para monitorizar el estado y desempeño de los FS. [10]

Categoría	Subcategoría	Clasificación UIT	Métrica	Descripción	Respaldada por
Rendimiento	Capacidad	VM/Virtual Disk Utilization Metrics	FS_TOTAL	Espacio de almacenamiento total con que cuenta el FS.	Oracle
	Utilización		FS_%OCUPADO	Espacio de almacenamiento no disponible en el FS. Se deben generar alertas de atención al sobrepasar el umbral del 85% y alertas de estado crítico al sobrepasar el 95%.	Oracle

			FS_%LIBRE	Espacio de almacenamiento disponible en el FS.	Oracle
	Throughput	I/O Performance metrics	ANCHO DE BANDA EFECTIVO	Capacidad de transferencia de datos promedio. Define la capacidad promedio de datos transferidos por unidad de tiempo.	SNIA, Oracle, VMware, Microsoft.
			THROUGHPUT	Capacidad promedio de transferencia de datos que llegan sin errores al destino. Debe mantenerse por encima del umbral del 96% del valor de la métrica ANCHO DE BANDA EFECTIVO [11].	Oracle, Microsoft, VMware.
			IOPS	Velocidad promedio en Operaciones por Segundo (OPS) de las operaciones de L/E. Se define en función del número de operaciones de L/E durante el tiempo de observación. En pruebas de lectura pura IOPS = LPS y en pruebas de escritura pura IOPS = EPS.	Oracle, Microsoft, SNIA.
	Tiempo de respuesta		LATENCIA	Tiempo de respuesta promedio para las operaciones de L/E. Debe mantenerse bajo el umbral de los 30 ms [12].	SNIA, Oracle, VMware.
			JITTER	Variabilidad en el tiempo de respuesta para las operaciones de L/E. No debe exceder el umbral del 2% de la métrica LATENCIA [13].	Autores
Disponibilidad			MTBF	Definida en la Tabla 1.	Autores
			MTTR	Definida en la Tabla 1.	Autores
Escalabilidad			USUARIOS	Número de usuarios que usan las capacidades del sistema. Se refiere tanto a nodos de un clúster como a ordenadores independientes acoplados a los recursos exportados del sistema.	Autores
			USUARIOS_MAX	Número de usuarios a partir del cual el desempeño del sistema se degrada drásticamente. Esta degradación se aprecia en las métricas IOPS, THROUGHPUT y LATENCIA.	Autores

Tabla 4: Métricas para evaluar el desempeño de volúmenes. [10]

Categoría	Subcategoría	Clasificación UIT	Métrica	Descripción	Respalda por	
Rendimiento	Capacidad	Storage Capacity Metrics	VOLUMEN_TOTAL	Espacio de almacenamiento total que ha sido asignado a un determinado volumen lógico.	Autores	
	Utilización		VOLUMEN_% OCUPADO	Porcentaje del espacio de almacenamiento de un volumen que se encuentra ocupado. Se deben generar alertas de atención al sobrepasar el umbral del 85% y alertas de estado crítico al sobrepasar el 95%.	Autores	
			VOLUMEN_% LIBRE	Porcentaje del espacio de almacenamiento de un volumen que se encuentra disponible.	Autores	
	Throughput	I/O Performance metrics	ANCHO DE BANDA EFECTIVO	Definida en la Tabla 3.	SNIA, Oracle, VMware, Microsoft.	
			THROUGHPUT	Definida en la Tabla 3. Debe mantenerse por encima del umbral del 96% del valor de la métrica ANCHO DE BANDA EFECTIVO.	Oracle, Microsoft, VMware.	
			Tiempo de respuesta	LATENCIA	Definida en la Tabla 3. Debe mantenerse bajo el umbral de los 30 ms.	SNIA, Oracle, VMware.
				JITTER	Definida en la Tabla 3. No debe exceder el umbral del 2% de la métrica LATENCIA.	Autores
			Exactitud	PRECISIÓN	Tiempo que transcurre desde que se escribe un dato hasta que está listo para ser leído.	Autores
				BER	Probabilidad de que un bit resulte erróneo durante el transporte de datos de las operaciones de L/E.	Autores

Tabla 5: Métricas para evaluar el desempeño de los nodos de un clúster. [10]

Categoría	Subcategoría	Clasificación UIT	Métrica	Descripción	Respalda por
Rendimiento	Utilización		ALM_TIPO	Define si un nodo en específico está utilizando el almacenamiento local, el secundario o ambos.	Autores
			SIST_TIPO	Define si el almacenamiento secundario de un nodo en específico está siendo	Autores

				exportado por un FS o por un sistema de almacenamiento en bloques.	
	<i>Throughput</i>	<i>I/O Performance metrics</i>	ANCHO DE BANDA EFECTIVO	Capacidad de transferencia de datos promedio. Define la capacidad promedio de datos transferidos por unidad de tiempo.	SNIA, Oracle, VMware, Microsoft.
			THROUGHPUT	Definida en la Tabla 3. Debe mantenerse por encima del umbral del 96% del valor de la métrica ANCHO DE BANDA EFECTIVO.	Oracle, Microsoft, VMware.
	Tiempo de respuesta		LATENCIA	Definida en la Tabla 3. Debe mantenerse bajo el umbral de los 30 ms.	SNIA, Oracle, VMware.
			JITTER	Definida en la Tabla 3. No debe exceder el umbral del 2% de la métrica LATENCIA.	Autores
	Exactitud		PRECISIÓN	Definida en la Tabla 4.	Autores
			BER	Definida en la Tabla 4.	Autores
Disponibilidad			MTBF	Definida en la Tabla 1.	Autores
			MTTR	Definida en la Tabla 1.	Autores
Escalabilidad			NODOS	Número de nodos que usan las capacidades del sistema.	Autores
			NODOS_MAX	Número de nodos a partir del cual el desempeño del sistema se degrada drásticamente. Esta degradación de aprecia en las métricas IOPS, THROUGHPUT y LATENCIA.	Autores

PRUEBAS DE VALIDACIÓN PARA SISTEMAS DE ALMACENAMIENTO

Amazon, una de las grandes entidades que posee Nubes para brindar IaaS, ha realizado algunas pruebas de desempeño a su Nube Amazon Elastic Compute Cloud (Amazon EC2) basándose en los métodos y las métricas definidas en las siguientes benchmarks: LMbench, Bonnie, CacheBench y HPC ChallengeBenchmark [14], pero todas las pruebas estuvieron basadas en las cinco instancias del entorno de Amazon EC2 (m1.small, m1.large, m1.xlarge, c1.medium y c1.xlarge) como bien se describe en el documento referenciado, lo cual indica que esto no constituye un estándar para métricas ni para evaluación de desempeño de sistemas, de hecho se conoce que en otras Nubes que brindan IaaS como GoGrid, Mosso y ElasticHost (EH) no emplean las mismas benchmarks para sus pruebas de desempeño. Por otro lado la compañía Microsoft ha publicado un conjunto de pruebas de validación para Sistemas de Almacenamiento que soportan clústeres de conmutación por error, las cuales tienen un perfil bastante general, aplicables a una gran variedad de Sistemas de Almacenamiento, pero tampoco constituyen un estándar [8]. Por su parte, a pesar de los avances realizados, las organizaciones de

estandarización como la SNIA, la UIT y el NIST, no han obtenido un conjunto de pruebas de validación estandarizadas para evaluar las potencialidades de los Sistemas de Almacenamiento en general.

A continuación se exponen el conjunto de pruebas de validación propuestas por los autores del presente trabajo, el que está basado en las pruebas que realizan los principales proveedores de servicios Nube.

Propuesta de pruebas de validación.

El conjunto de pruebas propuestas se encuentra dividido en cuatro categorías fundamentales: configuración, desempeño, disponibilidad y escalabilidad. Estas abarcan todas las métricas propuestas por los autores. [10] La Tabla 6 muestra el conjunto de pruebas propuestas, junto a una breve descripción de las mismas.

Tabla 6: Pruebas de validación y sus categorías.

Categorías	Subcategorías	Nombre de la prueba propuesta	Descripción
PRUEBAS DE DESEMPEÑO	PRUEBAS DE ESTADO DE UTILIZACIÓN	Monitorización	Las pruebas propuestas bajo esta categoría quedan respaldadas por su importancia a la hora de conocer los niveles de utilización de las capacidades del sistema.
	PRUEBAS DE ESTRÉS	Exposición del Sistema de Almacenamiento a tareas exhaustivas	Las pruebas propuestas bajo esta categoría quedan respaldadas por su importancia a la hora de conocer cómo se desempeña un Sistema de Almacenamiento bajo una carga de trabajo extrema, lo cual es cada vez más común en Sistemas de Almacenamiento para servicios telemáticos en la actualidad, con la creciente demanda sobre los sistemas computacionales.
	PRUEBAS DE TRANSFERENCIA	Validación de las transferencias de datos	Las pruebas propuestas bajo esta categoría quedan respaldadas por su importancia a la hora de conocer cuán rápida es transferida una determinada carga de datos hacia o desde un Sistema de Almacenamiento de forma exitosa, lo cual da una clarísima idea de su rendimiento y de la Calidad del Servicio (QoS ⁸).
		Validación de la precisión de las transferencias	
	PRUEBAS DE DEMORAS	Validación de los tiempos de respuestas	Las pruebas propuestas bajo esta categoría quedan respaldadas por su importancia a la hora de conocer cuán rápido son procesadas las peticiones al Sistema de Almacenamiento. Además esta constituye uno de los pilares fundamentales para dar una idea de la QoS que brinda el Sistema de Almacenamiento.
	PRUEBAS DE L/E	Validación de las operaciones de lectura	Las pruebas propuestas bajo esta categoría quedan respaldadas por su importancia a la hora de conocer cuán rápido se suceden las peticiones exitosas de L/E en el Sistema de Almacenamiento, lo cual tiene influencia directa en el rendimiento.
		Validación de las operaciones de escritura	

⁸Siglas del inglés “Quality of Service”.

PRUEBAS DE ESCALABILIDAD	Validación de la persistencia del rendimiento	Las pruebas propuestas bajo esta categoría quedan respaldadas por su importancia a la hora de conocer cuánto se afecta el desempeño del sistema al aumentar el número de usuarios que reciben el servicio.
PRUEBAS DE DISPONIBILIDAD	Validación de la disponibilidad	Las pruebas propuestas bajo esta categoría quedan respaldadas por su importancia para garantizar que los datos estén disponibles para los usuarios aun cuando ocurran fallos en el sistema.
PRUEBAS DE CONFIGURACIÓN	Validación de los procesos de configuración	Las pruebas propuestas bajo esta categoría quedan respaldadas por su importancia para verificar los procesos de creación, configuración y eliminación de recursos virtuales.

Ejemplo de ejecución de Prueba de Configuración.

Del conjunto de pruebas ejecutadas en el trabajo se ejemplificará la Prueba de Configuración. La prueba se desarrolló en un Sistema de Almacenamiento de poca complejidad y recursos, implementado con dispositivos de almacenamiento reales. La interconexión de los dispositivos fue realizada a través de la red interna del ISPJAE. La topología del escenario de pruebas se muestra en la Figura 1. La Tabla 7 muestra las características de los servidores empleados. La red es Ethernet con velocidades de hasta 100 Mbps, aunque el tráfico entre los dispositivos de este sistema, principalmente los que pertenecen a subredes diferentes, no alcanza ese valor. Esto se debe a que esta red no es dedicada al Sistema de Almacenamiento y por lo tanto el ancho de banda es compartido con el resto de los servicios implementados sobre la misma, ya que no ha sido posible implementar las políticas de distribución y aislamiento de tráfico pertinentes para garantizar mejores velocidades de transferencia. Es importante resaltar que el objetivo de esta prueba es verificar la contundencia y precisión de las pruebas de validación propuesta por los autores, no del escenario en sí.

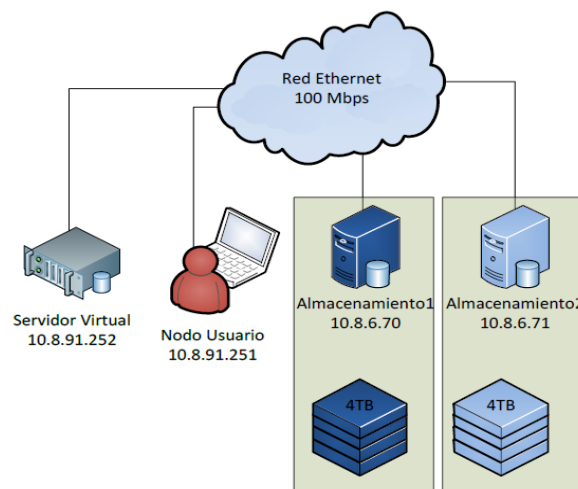


Figura 1: Escenario para la prueba de validación de la configuración.

Tabla 7: Resumen de las propiedades de los dispositivos.

Elemento	Almacén1	Almacén2	Servidor Virtual	Usuario
Memoria RAM.	3.85GB	3.85GB	1GB	2.8GB
CPU.	Pentium(R) Dual-Core CPU E5200	Pentium(R) Dual-Core CPU E5200	AMD Dual-Core A6-4400M	Intel Core 2 Duo
Núcleos.	2	2	1	2
Frecuencia CPU.	2.5 GHz	2.5 GHz	2.7 GHz	2 GHz
Tamaño y número de discos.	4 discos de 931.51 GB	4 discos de 931.51 GB	4 discos de 1.76 TB, 1 disco de 4 GB	1 disco de 320 GB
Almacenamiento total.	≈3.55 TB	≈3.55 TB	≈7 TB	320 GB
Sistema Operativo.	Openfiler 2.99 x64	Openfiler 2.99 x64	Openfiler 2.3 x86	Ubuntu 12.04 x64

En esta prueba se comprobó el correcto funcionamiento del sistema durante todo el procedimiento de la gestión de los recursos virtuales de almacenamiento, principalmente grupos de volúmenes y volúmenes lógicos de almacenamiento. Para esto se crearon, utilizaron, monitorizaron y finalmente se eliminaron cada uno de estos elementos virtuales. Es importante aclarar que en esta prueba se abordó parte de la prueba definida como "Monitorización", dentro de la categoría de "Pruebas de Estado de Utilización", ya que además de chequear la correcta creación de los elementos virtuales, se monitorizaron los niveles de utilización y se emplearon los umbrales de algunas métricas definidas en aquella prueba para definir el criterio de aceptación de esta prueba como tal. Al comprobarse la correcta creación de los elementos de almacenamiento virtuales, se comenzó el proceso de monitorización de los parámetros y la utilización de los mismos. La prueba se consideró exitosa siempre que ninguna de las métricas sobrepasara sus respectivos umbrales. La Tabla 8 muestra las métricas y parámetros que se monitorizan en esta prueba con sus respectivos umbrales. Estos umbrales se establecieron mediante algunas modificaciones hechas a las recomendaciones de umbrales de utilización de la compañía Oracle [9].

Tabla 8: Métricas y umbrales para la prueba.

Métrica	Umbral Alerta	Umbral Alarma
GRUPOS.	X	X
VOLÚMENES.	X	X
VOLÚMENES/GRUPO.	X	X
VOLÚMENES_XFS/GRUPO.	X	X
VOLÚMENES_SCSI/GRUPO.	X	X
GRUPO_TOTAL.	X	X

VOLUMEN_TOTAL.	X	X
GRUPO_%LIBRE.	X	X
GRUPO_%ASIGNADO.	X	↑99%
GRUPO_%OCUPADO.	↑85%	↑95%
VOLÚMEN_%LIBRE.	X	X
VOLÚMEN_%OCUPADO.	↑85%	↑95%

Las herramientas empleadas fueron: el sistema operativo Openfiler versión 2.99 [15] en el cual están sustentadas la mayoría de las métricas monitorizadas; en el nodo usuario se emplearon herramientas como "open-iscsi" y "nfs-client", ya que son extensiones necesarias que debe tener instalado el sistema operativo Ubuntu 12.04 para el proceso de acoplamiento de los recursos exportados mediante los protocolos Sistema para Pequeñas Computadoras de Internet (iSCSI9) y Sistema de Archivos de Red (NFS10).

Durante la prueba se creó un almacén de recursos con una capacidad aproximada a los siete TB, a partir del almacenamiento brindado por los nodos de almacenamiento. Todo el almacén es presentado a los usuarios por el servidor virtual. Luego a través del servidor virtual se creó un grupo de volúmenes llamado "finalgroup", de una capacidad total de aproximadamente siete TB sobre un arreglo de discos utilizando la tecnología de Conjunto Redundante de Discos Independientes (RAID11) - 0. En el mismo se crearon dos volúmenes lógicos llamados "iscsi" y "nfs" de 15 GB cada uno, de los cuales el primero es exportado mediante iSCSI para almacenamiento en bloques y el segundo mediante NFS para acceso basado en ficheros. Luego, el nodo usuario se acopló al almacenamiento exportado mediante ambos protocolos y solo entonces, comienza a explotar los recursos cuya utilización debe ser monitorizada. Finalmente se desacoplaron el almacenamiento del nodo usuario y se eliminaron todos los recursos virtuales creados.

Mientras el almacenamiento se encuentra acoplado al usuario se copió información en ellos y se monitorizó su estado de utilización, el cual nunca superó los umbrales preestablecidos. En la Figura 2 se puede apreciar subrayado con línea roja el nivel de utilización del volumen "nfs", creado en el grupo de volúmenes "finalgroup". Por otro lado, subrayado en azul, se pueden apreciar los niveles de utilización de las particiones usadas por el sistema operativo. En la Figura 3 se puede apreciar subrayado con línea roja el nivel de utilización del volumen iscsi desde el nodo de usuario.

Mount	Type	Partition	Percent Capacity	Free	Used	Size
/exportando	xfs	/dev/mapper/finalgroup-nfs	60%	5.79 GB	8.85 GB	14.65 GB
/mnt/finalgroup/nfs	xfs	/dev/mapper/finalgroup-nfs	60%	5.79 GB	8.85 GB	14.65 GB
/boot	ext3	/dev/sda1	9%	84.70 MB	8.92 MB	98.72 MB
/	ext3	/dev/sda3	17%	2.56 GB	577.11 MB	3.29 GB
/dev/shm	tmpfs	tmpfs	1%	504.72 MB	0.00 KB	504.72 MB
Totals :			51%	8.93 GB	9.43 GB	18.53 GB

Figura 2. Correcta monitorización de los volúmenes XFS.

⁹Siglas del inglés "Internet Small Computers System Interface".

¹⁰Siglas del inglés "Network File System".

¹¹Siglas del inglés "Redundant Array of Independent Disks".

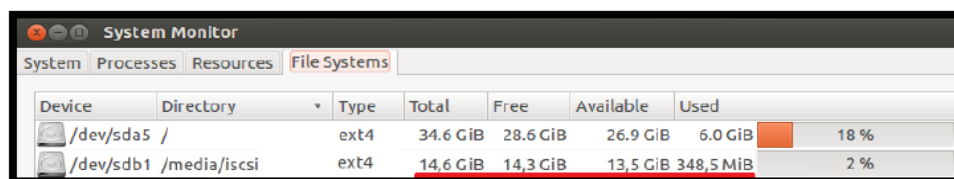


Figura 3: Correcta monitorización del volumen iSCSI.

Con el desarrollo de la prueba de configuración, con resultados satisfactorios, se determinó la buena eficiencia en los procesos de creación, utilización y eliminación de elementos virtuales del Sistema de Almacenamiento bajo prueba, lo cual es útil a la hora de emitir un criterio de bases sólidas para el diseño final del mismo, el cual es uno de los principales propósitos de la prueba diseñada. En esta prueba se pudieron obtener el valor de todas las métricas a medir como muestra la Tabla 9, algunas brindadas directamente por el sistema operativo y otras calculadas utilizando las expresiones 1 y 2.

$$\text{GRUPO}\% \text{OCUPADO} = \frac{\sum \text{CAPACIDAD OCUPADA DE LOS VOLÚMENES QUE CONTIENE}}{\text{GRUPO TOTAL}} \times 100\% (1)$$

$$\frac{\text{VOLÚMENES}}{\text{GRUPO}} = \frac{\text{VOLÚMENESXFS}}{\text{GRUPO}} + \frac{\text{VOLÚMENESSCSI}}{\text{GRUPO}} (2)$$

Tabla 9: Valores reportados para las métricas medidas.

Métrica	General	Volumen "nfs"	Volumen "iscsi"	Extraída del software
GRUPOS.	1	-	-	Si
VOLÚMENES.	2	-	-	Si
VOLÚMENES/GRUPO.	2	-	-	No
VOLÚMENES_XFS/GRUPO.	1	-	-	Si
VOLÚMENES_SCSI/GRUPO.	1	-	-	Si
GRUPO_TOTAL.	7 TB	-	-	Si
VOLUMEN_TOTAL.	-	15 GB	15 GB	Si
GRUPO_%LIBRE.	7064.66 GB	-	-	Si
GRUPO_%ASIGNADO.	29.31 GB (0.414%)	-	-	Si
GRUPO_%OCUPADO.	8.88 GB (0.125%)	-	-	No
VOLÚMEN_%LIBRE.	-	5.79 GB (40%)	14.3 GB (98%)	Si
VOLÚMEN_%OCUPADO.	-	8.85 GB (60%)	348.5 MB (2%)	Si

Los resultados de esta prueba permitieron comprobar la efectividad de la prueba propuesta, validando el método, las métricas y las herramientas, indicando la obtención de valores confiables.

CONCLUSIONES.

En este trabajo se obtuvo como resultado la propuesta de un conjunto de pruebas, métricas y sus umbrales para la validación del desempeño de un Sistema de Almacenamiento independientemente de la tecnología subyacente. En la propuesta se definieron un conjunto de métricas para cada uno de los componentes de un Sistema de Almacenamiento, por lo que se puede obtener una visión completa del estado, uso y potencialidades del mismo. Este resultado se logró a partir del estudio de las métricas y pruebas de validación más utilizadas en la actualidad por las principales organizaciones y compañías desarrolladoras de soluciones de almacenamiento, y del análisis de las ventajas y desventajas de cada una de estas propuestas. La clasificación de las métricas propuestas según el tipo de almacenamiento y las pruebas según las áreas que abarcan, le aportan a la propuesta final flexibilidad y adaptabilidad muy necesario en este tipo de soluciones versátiles. Por último, se mostró un ejemplo de ejecución de dos de las pruebas propuestas en un escenario de prueba totalmente real y los resultados obtenidos, permitió comprobar la efectividad de la prueba propuesta, validando el método, las métricas y las herramientas, indicando la obtención de valores confiables.

RECONOCIMIENTOS

Los autores desean agradecer a todos los que de una forma u otra colaboraron y ayudaron a la realización de este trabajo.

REFERENCIAS.

- [1] L. M. Kaufman, "Data Security in the World of Cloud Computing," *Secur. Priv. IEEE*, vol. 7, no. 4, pp. 61 –64, Aug. 2009.
- [2] K. Bakshi, "Considerations for cloud data centers: Framework, architecture and adoption," 2011, pp. 1 –7.
- [3] P. M. Timothy Grance, "The NIST Definition of Cloud Computing," 2011.
- [4] Focus Group on Cloud Computing, "FG Cloud Technical Report Part 1- Introduction to the cloud ecosystem: definitions, taxonomies, use cases and high- level requirements," 2012.
- [5] Focus Group on Cloud Computing, "FG Cloud Technical Report Part 5: Cloud security," 2012.
- [6] Focus Group on Cloud Computing, "FG Cloud Technical Report Part 4: Cloud Resource Management Gap Analysis," 2012.
- [7] "hIOMon Disk I/O Ranger User Guide," 4450 Arapahoe Avenue, Suite 100 Boulder, Colorado USA, 2013.

- [8] Focus Group on Cloud Computing, "FG Cloud Technical Report Part 3: Requirements and framework architecture of cloud infrastructure," 2012.
- [9] "Automatic Storage Management," 17-Jan-2005. [Online]. Available: http://docs.oracle.com/cd/E24628_01/em.121/e25160/osm_instance.htm. [Accessed: 17-Jan-2005].
- [10] Freddy A. Poll García, "Métricas y Pruebas de validación para Sistemas de Almacenamiento," ISPJAE, La Habana, Cuba, 2013.
- [11] YiqiXu, Dulcardo Arteaga, Ming Zhao, Yonggang Liu, Renato Figueiredo, SeetharamiSeelam. vPFS: Bandwidth Virtualization of Parallel Storage Systems. 2012.
- [12] Seagate. Recomendaciones para las pruebas de rendimiento de unidades SSD. 2012. Available from: <http://www.seagate.com>
- [13] Priscilla Oppenheimer. Top-Down Network Design. Third ed. 2010.
- [14] Simon Ostermann, Alexandrulosup, NezhYigitbasi, RaduProdan, Thomas Fahringer, Dick Epema. A Performance Analysis of EC2 Cloud Computing Services for Scientific Computing.
- [15] Openfiler Administration Guide.