



POSGRADOS

Maestría en **TELEMÁTICA**

RPC-SO-01-NO.025-2021

Opción de Titulación:

Proyecto de titulación con componentes de investigación aplicada y/o de desarrollo

Tema:

Análisis del impacto de la computación de alto rendimiento en ambientes virtualizados basado en técnicas de Benchmark para el data center de la Universidad Politécnica Salesiana campus sur.

Autor(es)

Patricio Ulices Valencia Asqui

Director:

Jorge Enrique López Logacho

QUITO – Ecuador
2023



Autor(es):



Patricio Ulices Valencia Asqui
Ingeniero en Sistemas Mención Gestión.
Candidato a Magíster en Telemática por la Universidad Politécnica Salesiana – Sede Quito.
pvalencia@est.ups.edu.ec

Dirigido por:



Jorge Enrique López Logacho
Profesor Titular de la Universidad Politécnica Salesiana
Ingeniero de Sistemas
Magíster en Redes de Información y Conectividad
jlopez@ups.edu.ec

Todos los derechos reservados.

Queda prohibida, salvo excepción prevista en la Ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra para fines comerciales, sin contar con autorización de los titulares de propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual. Se permite la libre difusión de este texto con fines académicos e investigativos por cualquier medio, con la debida notificación a los autores.

DERECHOS RESERVADOS

2023 © Universidad Politécnica Salesiana.

QUITO– ECUADOR – SUDAMÉRICA

PATRICIO ULICES VALENCIA ASQUI

ANÁLISIS DEL IMPACTO DE LA COMPUTACIÓN DE ALTO RENDIMIENTO EN AMBIENTES VIRTUALIZADOS BASADO EN TÉCNICAS DE BENCHMARK PARA EL DATA CENTER DE LA UNIVERSIDAD POLITÉCNICA SALESIANA CAMPUS SUR.

DEDICATORIA

A Dios, a mis padres, a mi hermano por ser el impulso y el eje central para luchar ante toda situación difícil en la vida, por darme el ánimo de continuar estudiando y sobre todo haber forjado en mi los valores y las ganas de jamás rendirme. A mi hija Lía Valencia dado que en ella veo el reflejo de la perseverancia y la constancia en este camino que ha no resultado nada fácil, a sus besos, a sus abrazos y al cariño incondicional que ella pone en mí, por considerar a su Padre como un ejemplo en su pequeña y corta vida, a las ganas y al amor que se demuestra en todas sus acciones. A todas las personas que de alguna manera han hecho posible este logro.

AGRADECIMIENTO

A Dios por darme el conocimiento y las ganas de seguir adelante, a mis padres, a mi hermano, que me han regalado la fuerza para seguir adelante y no desmayar. A nuestros catedráticos especialmente al ingeniero: Jorge Luis López Logacho que fue pilar fundamental en el desarrollo de la tesis planteada y desarrollada. A la Universidad Politécnica Salesiana por formar líderes con conciencia social. A todas las personas que estuvieron y están en este aprendizaje académico, que Dios les bendiga.

Tabla de Contenido

Resumen	8
Abstract	9
1. Introducción	10
2. Determinación del Problema.....	12
3. Objetivos.....	14
3.1. Objetivo general	14
3.2. Objetivos específicos	14
4. Justificación	14
5. Marco teórico referencial.....	16
5.1. Virtualización	16
5.1.2. Características de la virtualización	16
5.1.3. Tipos de virtualización	17
5.2. Computación de alto rendimiento (HPC)	19
5.2.1. Importancia de la HPC (Computación de Alto Rendimiento).....	20
5.2.2. Funcionamiento de la HPC (Computación de Alto Rendimiento).....	21
5.2.3. Usos de la HPC (Computación de Alto Rendimiento).....	23
5.2.4. El futuro de la HPC.....	24
5.3. Benchmark.....	24
5.3.1. ¿Dónde se puede aplicar el Benchmark?	25
5.3.2. Enfoques de los Benchmark	26
5.3.3. Herramientas y técnicas de Benchmark.....	27
5.4. Arquitectura del data center de la Universidad Politécnica Salesiana.....	28
6. Materiales y metodología.....	31
7. Resultados y discusión.....	39
7.1. Análisis de datos - Resultados	39
7.1.1. Escenario 1.....	39
7.1.2. Escenario 2.....	45
7.1.3. Escenario 3.....	49
7.1.4. Escenario 4.....	50
7.2. Análisis del rendimiento del almacenamiento.....	52
7.2.1. Escenario 1.....	52

7.2.2. Cálculos de rendimiento para almacenamiento.	57
8. Conclusiones.....	60
Referencias	62
Anexos	64

Análisis del impacto de la computación de alto rendimiento en ambientes virtualizados basado en técnicas de Benchmark para el data center de la Universidad Politécnica Salesiana campus sur

Autor(es):

VALENCIA ASQUI PATRICIO ULICES

Resumen

La computación de alto rendimiento permite estudiar y encontrar respuestas a algunos de los mayores problemas del mundo en ciencia, tecnología y negocios. Hoy en día se utiliza para resolver problemas complejos y de alto consumo energético, y las empresas trasladan cada vez más las cargas de trabajo de HPC a la nube. Bajo este argumento el objetivo del proyecto es analizar el impacto de la computación de alto rendimiento en ambientes virtualizados basado en técnicas de Benchmark para la data center de la Universidad Politécnica Salesiana campus sur. Para el cumplimiento del objetivo se recurre a cuatro fases que comprende el desarrollo de las actividades siendo el uso de herramientas y técnicas de Benchmark; elección de la herramienta y definición de las métricas de evaluación; diseño del o de los posibles escenarios; simulación y validación de las métricas y de la herramienta o script. Bajo este proceso se concluye que el uso de técnicas de Benchmark Phoronix Test Suite facilitan la obtención de resultados en los componentes de las computadoras de alto rendimiento, sin embargo, en ambientes virtualizados de estas computadoras no están libres de problemas, uno de ellos, la ejecución de instrucciones al no contar con un sistema operativo nativo. Por lo tanto, el análisis de la infraestructura de la data center de la Universidad Politécnica Salesiana en ambientes virtualizados da como resultado final un valor de 69.08 TFlops sin contar los clústeres y un valor de virtualización total de 76.89 TFlops contando los clústeres. Teniendo en cuenta que el impacto en ambientes virtualizados (clúster) los valores obtenidos no tienen un impacto negativo.

Palabras clave:

Computación de alto rendimiento, ambientes virtualizados, Benchmark, data center.

Abstract

High-performance computing makes it possible to study and find answers to some of the world's biggest problems in science, technology and business. Today it is used to solve complex and energy-intensive problems, and companies are increasingly moving HPC workloads to the cloud. Under this argument the objective of the project is to analyze the impact of high-performance computing in virtualized environments based on Benchmark techniques for the data center of the Universidad Politécnica Salesiana south campus. In order to achieve the objective, four phases are used, including the development of the activities being the use of Benchmark tools and techniques; choice of the tool and definition of the evaluation metrics; design of the possible scenario(s); simulation and validation of the metrics and the tool or script. Under this process it is concluded that the use of Benchmark Phoronix Test Suite techniques facilitates the obtaining of results in the components of high-performance computers, however, the virtualized environment of these computers is not free of problems, one of them is the execution of instructions since they do not have a native operating system. Therefore, the analysis of the data center infrastructure of the Salesian Polytechnic University in a virtualized environment gives as a final result a value of 69.08 TFlops without counting the clusters and a total virtualization value of 76.89 TFlops counting the clusters. Considering that the impact on virtualized environments (cluster) the values obtained do not have a negative impact.

Palabras clave:

High performance computing, virtualized environments, Benchmark, data center.

1. Introducción

La computación de alto también conocida como HPC (por sus siglas en inglés High-Performance Computing) es una disciplina de la informática que se enfoca en utilizar sistemas y recursos computacionales avanzados para resolver problemas complejos y realizar cálculos intensivos en tiempo reducido. Realizan tareas complejas que requieren procesar los datos, crear modelos para su estudio y comprender mejor los problemas complejos [1].

Por otro lado, la computación de alto rendimiento no solo se limita a ambientes privativos (Windows, Mac OS) o con un licenciamiento definido de manera mensual o anual, en los últimos años los ambientes free o conocidos como Software Libre es una vía de investigación, al no tener que pagar por una licencia o software. Una de las ventajas de la HPC se desenvuelve en ambientes virtualizados, que haciendo uso de software libre se puede medir cual es el procesamiento, rendimiento y determinar las ventajas o falencias que se pueden dar al virtualizar una arquitectura de sistemas, en ambiente de desarrollo, o simplemente crear un ecosistema informático haciendo uso de sistemas libres.

Una de las formas posibles de medir este rendimiento en ambiente virtualizados es usando técnicas de Benchmark, de hecho, este término hace referencia a métodos de cálculo computacionales que son aplicables en un ambiente seguro, medible e identificable de tal manera que no existan diferencias con el ambiente del computador[2].

Bajo esta línea, la técnica de medición de un Benchmark es una prueba manejada y controlada sobre dispositivos tecnológicos para comprobar su rendimiento en base a los recursos disponibles. Es decir, se trata de programa de valoración comparativa de computadoras que trabaja estableciendo una serie de pruebas definidas en el PC para medir su rendimiento, una de las herramientas utilizadas es Phoronix Test Suite [3].

En este contexto, al identificar que el Data Center de la Universidad Politécnica Salesiana dispone de una computadora de alto rendimiento con dos tarjetas gráficas, se ha establecido analizar el rendimiento de los recursos de la cual hace uso y las técnicas de Benchmark permitan obtener criterios puntuales sobre el rendimiento de esta máquina.

A su vez, la necesidad de la investigación es conocer de qué forma trabajan las herramientas y técnicas de Benchmark, así como establecer las métricas para analizar el impacto del rendimiento dentro de los discos, tarjetas gráficas en software libre y entornos virtualizados.

2. Determinación del Problema

La HPC incorpora varios clústeres (incorporación de varios computadores) para trabajar de manera simultánea en el procesamiento de datos y ejecutar cálculos que una máquina con prestaciones limitadas no podría realizar [4]. Para tener una perspectiva por ejemplo un computador de escritorio con un procesador de 3 GHz logra ejecutar unos 3.000 millones de transacciones por segundo, las computadoras de alto rendimiento pueden realizar cuatrillones de cálculos por segundo.

No obstante, el rendimiento de las HPC se puede evaluar por sus características de Hardware que posee, como es la tarjeta de video, la memoria RAM, la tarjeta gráfica, los servidores interconectados.

En ambientes virtualizados medir el rendimiento del HPC involucra la creación de máquinas virtuales que estén conectadas con la infraestructura del data center para poder ejecutar técnicas que arrojen una métrica de evaluación, de esta manera se podrá obtener métricas de evaluación haciendo uso de un Benchmark (comparativas de rendimiento) esto se lo realiza con el fin de determinar el comportamiento de los elementos de una infraestructura de sistemas en base a los diferentes test que son ejecutados por el Benchmark, [5] la Universidad Autónoma de Barcelona enfatiza que las técnicas que emplea los Benchmark para medir el rendimiento a través de la ejecución de pruebas como carga de información, estrés, rendimiento dentro una infraestructura son estrictas en su ejecución lo que permite conocer las características técnicas de los dispositivos involucrados y determinar a través de los resultados expuestos si se puede mejorar para obtener un mejor performance.

Por ejemplo, Carpena Latour [6], da a conocer que la mayor parte de software aprovecha uno o máximo dos núcleos de procesamiento y no hace uso de los demás recursos disponibles, se debe a que en gran parte a que el desarrollo de los códigos fuentes con los que se realizan los softwares siguen una serie no definida y repetible

de cadenas de texto que dificulta su interpretación por parte de la máquina y no se logra un resultado totalmente eficiente.

En este contexto de rendimiento las tarjetas gráficas tienen un papel fundamental en el procesamiento de las computadoras de alto rendimiento HPC, no solo para ambientes virtualizados sino también para la infraestructura de un data center, en el estudio de Luszczek et al., [7] dan a conocer que las estructuras de las tarjetas gráficas están dadas por un GPU o unidad gráfica de procesamiento con una gran aumento en el tamaño del pixel de los gráficos, un aumento de procesamiento de datos, aplicación de algoritmos, razón por lo cual al contar con dispositivos limitados en relación al procesador, memoria RAM, almacenamiento, dificulta en gran medida que se pueden llevar a cabo las tareas de análisis, conteo de datos, recopilación, creación de ambientes virtuales.

En el presente trabajo se pretende utilizar las técnicas de Benchmark haciendo uso de ambientes virtualizados (máquinas virtuales) para su ejecución, que estén conectas con la infraestructura del data center de la Universidad Politécnica Salesiana, y que a su vez permita el uso de Software libre como base de su sistema operativo.

El ambiente virtualizado depende de las capacidades del hardware del data center, esto involucra la cantidad de memoria RAM con la que se va a trabajar, el tamaño de almacenamiento, y la instalación de software libre que se realice, en este caso Linux distribución Ubuntu versión 20.0 LTS

Dentro de la elección del Benchmark, se establecerá fases que se las describe en la metodología y herramientas, esto permita definir de mejor manera el Benchmark que se adapte a los objetivos propuestos.

Dentro de esta investigación se usó un Benchmark en un sistema de 8 núcleos con procesadores Core i7 utilizando Open MPI, se trabajó sobre hardware y software virtualizado haciendo uso de los elementos definidos e implementados en el centro de datos de la Universidad Politécnica Salesiana campus Sur.

Como se esperaba, los resultados de HPC tuvieron gastos generales en el entorno virtual instalado y la sobrecarga se vuelve más significativa en base a los test ejecutados por el Benchmark, es decir, a mayor cantidad de test ejecutados mayor información obtenida.

3. Objetivos

3.1. Objetivo general

Analizar el impacto de la computación de alto rendimiento en ambientes virtualizados basado en técnicas de Benchmark para la data center de la Universidad Politécnica Salesiana campus sur.

3.2. Objetivos específicos

- Estudiar cómo trabajan las herramientas y técnicas de Benchmark y como establecer las métricas para analizar el impacto del rendimiento dentro de una tarjeta en software libre y entornos virtualizados.
- Definir las métricas de evaluación para las tarjetas gráficas con los que cuenta la data center de la UPS campus sur apoyándose de las técnicas de Benchmark.
- Implementar una herramienta de Benchmark para obtener las métricas de medición que se definieron en el objetivo anterior.
- Analizar los resultados obtenidos en base a las métricas y a las técnicas Benchmark que se escogieron para analizar las tarjetas gráficas de la computación de alto rendimiento con la que cuenta la data center.

4. Justificación

El presente proyecto técnico pretende realizar una investigación para poder obtener medidas de rendimiento, escalabilidad, demanda de recursos en la Universidad Politécnica Salesiana, predominando el rendimiento de los elementos

que conformar la infraestructura del data center, el incremento de requerimientos computacionales, equipos físicos ha proporcionado la necesidad de virtualización con el fin de aumentar la demanda y le performance de la arquitectura definida del data center. Por este motivo, es necesario realizar un análisis de toda la infraestructura teniendo en cuenta el procesamiento y el almacenamiento en ambientes virtualizados haciendo uso de todos sus elementos.

Un ves definidos los escenarios en base a los resultados obtenidos con técnicas de Benchmark se puede medir el impacto de los ambientes virtualizados bajo software libre, poder analizar el total de la infraestructura obteniendo resultados con unidades de medida en TFlops para el rendimiento e lops para el almacenamiento, con esta medición y análisis del rendimiento del data center haciendo uso de técnicas de Benchmark en entornos virtualizados se puede definir si la arquitectura elegida así como sus componentes son viables para la utilización, creación de entornos virtuales que puedan servir para el procesamiento de datos al crear diferentes máquinas virtuales.

5. Marco teórico referencial

En este apartado se hace referencia al análisis de los elementos que se han detallado en la investigación, definición de criterios de implementación, elaboración de escenarios de medición, de este modo, se pretende cumplir con los objetivos planteados.

5.1. Virtualización

La virtualización se refiere a la tecnología informática que simula las funciones del hardware físico para ofrecer servicios informáticos asentados en software, tales son aplicaciones y redes [8]. Al crear una versión virtual de un recurso o dispositivo (como un ordenador) en un único sistema informático, las organizaciones se benefician de la reducción de precio y del aumento de la eficiencia.

En este sentido, la virtualización utiliza software para simular funciones de hardware y fundar métodos informáticos implícitos, haciendo uso de los elementos de una infraestructura tecnológica o de las características técnicas de los equipos que se disponga.

Gómez y Leyva [9] afirman que, la virtualización es un soporte de hardware que facilita o acelera el funcionamiento de ordenadores virtuales en un ordenador real. Con la virtualización, se pueden ejecutar varios sistemas de software en un solo hardware. Pueden ser diferentes sistemas operativos, por ejemplo.

5.1.2. Características de la virtualización

- **Flexibilidad**

El software de virtualización ofrece más flexibilidad a la hora de realizar pruebas y asignar recursos. Dado que las máquinas virtuales se pueden respaldar y restaurar fácilmente, el departamento de TI puede probar y ensayar fácilmente nuevas

tecnologías, como es el objeto de este estudio. La virtualización también sienta las bases de una habilidad de almacenamiento que agrupa los medios de los artefactos virtuales para su empresa [13]. Esta construcción asentada en la nube permite al dispositivo de TI controlar y acceder a los medios y dispositivos.

- **Fiabilidad**

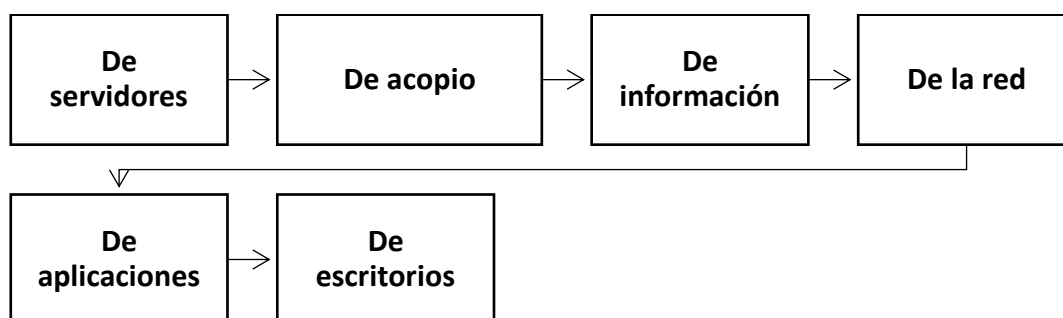
Hacer una copia de datos y reponer es fácil con la virtualización, utilizando instantáneas de la máquina virtual en los servidores existentes. La copia de seguridad también puede automatizarse fácilmente para que todos los datos estén siempre actualizados. En caso de recuperación de un desastre con la copia de virtual, se puede transmitir a un nuevo host en sólo unos minutos [14]. Esta recuperación sin complicaciones de los fallos o pérdidas se traduce en mejora del lugar y de las catástrofes.

- **Eficiencia**

Con la virtualización, un dispositivo físico puede dividirse en varios aparatos implícitos. No se necesitan un número menor de servidores, sino que los existentes se pueden utilizar de forma óptima. Este aumento de la eficiencia se traduce en un ahorro incluso en el mantenimiento. También consiente en hacer varios tipos de sistemas operativos en un único dispositivo, en lugar de tener que utilizar ordenadores separados [13]. Lo que significa que no está atado a proveedores específicos y que puede gestionar sus recursos físicos de manera mucho más eficiente, haciendo que la TI sea más productiva.

5.1.3. Tipos de virtualización

La virtualización surge en los años 60 y se refería inicialmente a la instauración de aparatos virtuales y la concepción del hardware. Hoy en día, se localizan varios tipos de virtualización en los entornos informáticos modernos, que se refieren a la contemplación de recursos informáticos, tal es el caso del software, el almacenamiento, la información o las unidades de red. Por lo tanto, se distingue entre:

Figura 1*Tipos de virtualización*

Nota: La figura muestra los diferentes tipos de virtualización que existen. Tomado de García [16]

I. Virtualización de servidores

Este primer tipo de virtualización son potentes artefactos perfilados para realizar labores específicas y complejas. En los servidores se maneja un hipervisor para dividir los asistentes físicos en variados asistentes virtuales, donde todos tienen su sistema operativo individual [16]. Esto consiente interesar al máximo los servidores para comprimir los costes de hardware y actividad.

II. Virtualización del almacenamiento

El acopio material de varios aparatos de una red se combina en un punto de conexión de acumulación virtual que se gestiona a través de una consola central. En este tipo se requiere un software que reconozca la cabida favorable de los dispositivos materiales y los combine en un medio virtual [15]. El usuario final no nota ninguna diferencia entre el acopio tácito y un disco duro real.

III. Virtualización de datos

Para García [16] este tipo de virtualidad consiente que una aplicación a los datos y se utiliza sin necesidad de conocer información como la ubicación física o el tamaño de la información. Esto expresa que se pueden representar datos de diversos principios sin moverlos o copiarlos. Esta añadidura de información utiliza un

software de virtualidad para constituir y representar virtualmente la información mediante un panel de inspección. Esto admite acceder a grandiosos datos vinculados comenzando en un punto de paso central, libremente dónde estén almacenados.

IV. Virtualización de la red

Divide el espacio de banda utilizable en conductos autónomos, a los cuales se asigna a un punto de conexión según sea necesario. Además, facilita la programación y el aprovisionamiento de la red, incluyendo el equilibrio de carga y el cortafuegos sin tener que cambiar la infraestructura subyacente [13].

V. Virtualización de aplicaciones

Permite a las aplicaciones ejecutarlas por separado e independientemente del sistema operativo que utilicen. Se suele utilizar para ejecutar aplicaciones de Microsoft Windows en Linux o macOS [16].

VI. Virtualización de escritorio

Permite aparentar una estación de compromiso para ingresar a un escritorio de manera antigua a través de un aparato acoplado, como un consumidor liviano [16]. En consecuencia, los escritorios virtuales proporcionan el ingreso más seguro y móvil a los recursos en el foco de la información.

5.2. Computación de alto rendimiento (HPC)

HPC es una tecnología que hace uso de clúster de potentes procesadores trabajando en paralelo para procesar conjunto de datos masivos de diferentes dimensiones (big data por ejemplo) [17].

5.2.1. Importancia de la HPC (Computación de Alto Rendimiento)

La HPC no es nueva, puesto que, las estaciones de trabajo HPC y los superordenadores han desempeñado un papel importante en la investigación académica durante décadas, resolviendo problemas complejos e impulsando el descubrimiento y la innovación [20].

Los volúmenes de datos han aumentado rápidamente en los últimos años, y muchas aplicaciones nuevas se benefician de la potencia de la HPC: la capacidad de realizar operaciones de cálculo intensivo en recursos compartidos para producir resultados en menos tiempo y a menor coste que la informática tradicional. Al mismo tiempo, el hardware y el software de HPC se han vuelto más asequibles y generalizados.

De acuerdo con Naiouf et al., [21] científicos, ingenieros e investigadores están recibiendo HPC para una serie de casos de uso, como la previsión atmosférica, la mecánica cuántica, la indagación de carburante y éter, la electricidad y otras áreas de investigación académica y aplicaciones comerciales.

Por ello, mediante la información, se realizan revelaciones indiscutibles, se generan inventos autorizados e incremento la calidad de vida de millones de individuos a nivel internacional. La HPC es teoría y práctica ayuda del progreso científico, industrial y social.

Procesos conocidos como Internet de las Cosas (IoT), inteligencia artificial (IA) e iconografías 3D siguen evolucionando. En consecuencia, la dimensión y el volumen de la información con los que deben trabajar las empresas está desarrollarse para múltiples intenciones, como la minería de datos, la creación de monedas electrónicas basadas en algoritmos inteligentes, probar nuevos productos o analizar las disposiciones de las existencias y el desplazamiento de la información en tiempo real es fundamental [22].

Para adelantarse a la competitividad, las empresas requieren instalaciones informáticas rápidas y muy fiables para procesar, recolectar y estudiar conjuntos de información difíciles de medir y analizar.

5.2.2. Funcionamiento de la HPC (Computación de Alto Rendimiento)

La HPC actúa en un solo módulo, pero su auténtica fuerza proviene del enlace de distintos nodos HPC en un mismo clúster o superordenador que tiene capacidad de proceso de información en paralelo. Los clústeres son capaces de calcular simulaciones a un nivel abultado, deducciones de IA y observación de información que no serían posibles en un solo procedimiento [23].

Pocos de los iniciales y conocidos superordenadores estuvieron perfeccionados por Cray e IBM. Los superordenadores actuales son clusters de HPC a gran nivel formados por CPUs, aceleradores, estructuras de información de alta ventaja y retentiva avanzada que trabajan juntos en múltiples nodos para impedir que se generen cuellos de botella y obtener un beneficio óptimo.

En efecto, Domínguez [4] menciona que, la HPC o computación de alto rendimiento puede ejecutarse en muchos tipos de cargas de trabajo. Sin embargo, las dos más comunes son las cargas de trabajo intrínsecamente Paralelas y Estrechamente Acopladas:

Figura 2

Tipos de cargas de trabajo - funcionamiento

Cargas de trabajo intrínsecamente paralelas

- Se trata de problemas computacionales que se dividen en tareas pequeñas, sencillas e independientes que pueden ejecutarse simultáneamente, a menudo con poca o ninguna comunicación entre ellas.

Cargas de trabajo estrechamente acopladas

- Esto suele implicar una gran carga de trabajo compartida que se distribuyen en trabajos pequeños que se comunican continuamente entre sí. En otras palabras, los diferentes nodos del clúster se comunican entre sí mientras realizan su procesamiento.

Nota. La figura 2 muestra los diferentes tipos de carga en el funcionamiento dentro del trabajo para la HPC. Tomado de Domínguez [4].

Por otro lado, las soluciones HPC apalean tres mecanismos importantes:

- Informática
- Red
- Almacenamiento

Frente a estas soluciones, los programas y algoritmos de software se establecen paralelamente en los asistentes del clúster. El clúster está acoplado al almacén de datos para detener la salida. Estos elementos actúan al detalle o al requerimiento del procesamiento de la data para realizar diversas tareas [21].

Cada componente debe trabajar en sintonía con los demás para funcionar de manera óptima. Los componentes de stock deben analizar el requerimiento, procesarlo y entregar datos a los servidores, capturarlos lo más rápidamente posible [24]. Del mismo modo, los dispositivos de la red deben poder sobrellevar la transferencia de información de alta celeridad entre los servidores de la computadora y el acopio de datos. Si un dispositivo no logra procesar la data de igual manera que los demás componentes, el rendimiento de toda la infraestructura de HPC se afectado.

5.2.3. Usos de la HPC (Computación de Alto Rendimiento)

De acuerdo con Gómez, Umaña y Arce [25] las soluciones de HPC se despliegan en las instalaciones, en la periferia o en la nube y se utilizan en diferentes sectores con distintos fines. Algunos ejemplos son:

- I. **Laboratorios de investigación:** se maneja para auxiliar a los investigadores a localizar fuentes de energía renovable, percibir el progreso del universo, profetizar y rastrear tempestades y desplegar nuevos materiales.
- II. **Aeroespacial:** creación de disimulos complicadas, tal cual el flujo de aire ante las alas de las aeronaves.
- III. **Medios de comunicación y diversión:** se manipula para reproducir largometrajes, implantar deslumbrantes efectos específicos y pasar programas en vivo en todos los países.
- IV. **Petróleo y gas:** se recurre para determinar con mayor exactitud el lugar que se debe perforar nuevos pozos, mejorando el resultado de la fabricación de los pozos encontrados.
- V. **Inteligencia artificial/aprendizaje automático:** se maneja para descubrir estafas con tarjetas de crédito, suministrar ayuda técnica de autoservicio, enseñar a conducir vehículos autónomos o incluso mejorar los métodos sanitarios.
- VI. **Servicios financieros:** La HPC se utiliza para seguir las tendencias bursátiles en tiempo real y automatizar las operaciones.
- VII. **La sanidad:** Investigación de medicamentos, producción de vacunas y desarrollo de terapias innovadoras para enfermedades raras y comunes. Investigar curas de padecimientos tal es el caso de la diabetes o el Alzheimer y para consentir diagnósticos más rápidos y precisos en los pacientes.
- VIII. **Comercio minorista:** análisis de grandes cantidades de información de consumidores para ofrecer recomendaciones de bienes más determinados y una mejor asistencia al consumidor.

- IX. **Educación:** Se lo utiliza también para las operaciones administrativas y técnicas de instituciones educativas, tales como colegios, universidades, institutos, entre otros. [25]

5.2.4. El futuro de la HPC

Empresas e instituciones de una amplia gama de sectores confían en la HPC para impulsar un crecimiento que se espera que continúe durante muchos años. Se estima que el mercado internacional de HPC crezca de 31.000 millones de USD en 2017 a 50.000 millones en 2023 [26].

Se espera que gran parte de este crecimiento provenga de las implementaciones de HPC basadas en la nube, ya que el rendimiento de la nube continúa mejorando y se vuelve más confiable y estable, lo que liberará a las organizaciones de la necesidad de invertir millones en infraestructuras de centros de datos y gastos asociados.

Además, se espera que siempre esté presente lo siguiente:

- Rendimiento máximo
- Experiencia con HPC
- Flexibilidad para moverse
- No haya costes ocultos

En un futuro próximo, se espera que Big Data y HPC converjan, utilizando el mismo clúster de computación de gran tamaño para realizar análisis de Big Data y ejecutar simulaciones u otras cargas de trabajo de HPC [24]. Cuando estas dos tendencias converjan, se producirá un aumento de la potencia y la capacidad de cálculo para cada una de las dos aplicaciones. Esto admitirá una exploración e invención más sediciosas en el futuro.

5.3. Benchmark

En principio, la evaluación comparativa es una comparación o un estándar de comparación. Esto significa que determinados objetos, procesos, métodos,

productos o servicios se comparan con un punto de referencia previamente definido. Se trata, pues, de un instrumento de comparación y análisis.

En concreto, esto significa que se compara el propio rendimiento con el de una empresa externa o con otros rendimientos propios. Los estándares utilizados para la comparación también se denominan puntos de referencia, estos puntos de referencia son las mejores prácticas que también sirven de guía dentro del análisis del rendimiento. El proceso de evaluación comparativa no es algo puntual, sino un análisis continuo [27].

Esta herramienta se lo usa para medir el rendimiento en sistemas operativos libres o que se necesita licencia como Windows, en esencia son aplicaciones que miden la cantidad de información que puede procesar un sistema operativo, servidor, aplicación, etc., con la finalidad de aportar métricas de evaluación de la tarjeta gráfica, de su procesador o del conjunto de elementos que estén definidos por su arquitectura.

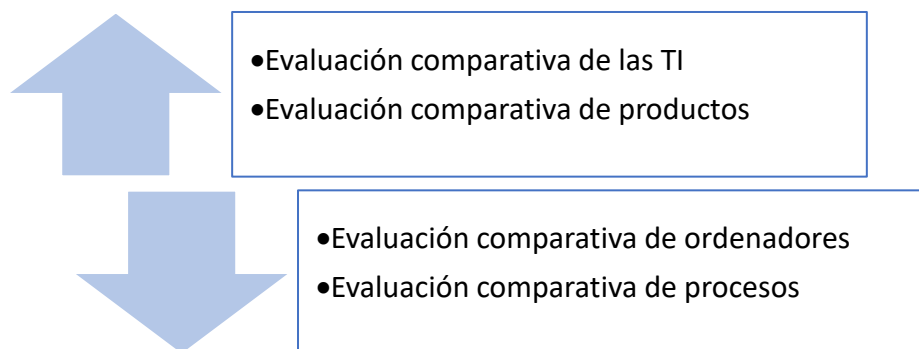
El objetivo principal de un Benchmark es, por supuesto, la optimización continua de los propios procesos, servicios, productos, etc. Además, todo ello sirve para una especie de gestión de la calidad. A través de las comparaciones internas y externas, usted recibe un estado de su objeto como valor real y el valor objetivo del objeto de comparación [28]. De este modo, podrá ver qué potencial puede explotar y dónde debe mejorar.

5.3.1. ¿Dónde se puede aplicar el Benchmark?

Los puntos de referencia y la evaluación comparativa se aplican en varios ámbitos. En este contexto, se ha aceptado hoy en día que la valoración se subdivide en diferentes métodos, como, por ejemplo:

Figura 3

Aplicación del Benchmark



Nota. La figura 3 muestra los diferentes lugares en donde se puede aplicar el Benchmark. Tomado de Poovey et al., [29]

5.3.2. Enfoques de los Benchmark

El enfoque de Benchmark se puede basar a los escenarios en concreto:

Benchmark sintéticos: los que suponen la ejecución de un programa real o producen ActionScript en un programa real también son simplificados porque no hay aleatoriedad en la prueba.

Benchmark real: es un procedimiento de experimento muy disperejo ya que asume la aleatoriedad cuando se emplea una presentación para evaluar su rendimiento. Las pruebas con este tipo de Benchmark se puede ajustar a los parámetros de calidad de una aplicación, software y a las posibilidades del hardware, cabe recalcar que la unidad de medida es el Teraflow [29].

Ahora bien, dentro de los ambientes de software libre se tiene algunos Benchmark con sus respectivas clasificaciones:

Benchmark para CPU: Las pruebas para este tipo pueden variar, pues existen variaciones como medir solamente el CPU o realizar una prueba de estrés para comprobar su estabilidad, dentro de ella tenemos:

- **Hardinfo:** programa de tipo AIDA64, que se usa para exponer indagación del hardware, y tiene otra parte que se usa para probar la CPU.

- **Sysbench:** es otra opción al primero, donde se logra evaluar el rendimiento o la carga de la CPU.
- **Stress y Stress-NG:** forma parte de una prueba de estrés simple.
- **S-tui:** es una pantalla de la diligencia de la CPU, pero asimismo acepta su subrayado

Benchmark para GPU: Existe algunas herramientas para evaluar la tarjeta gráfica en ambientes libres estas son:

- **Unigine Benchmark:** es una compilación de los experimentos (Superposition, Valley, Heaven, etc) más competentes y que está dispuesto para Linux.
- **Glxgears:** principal prueba para OpenGL.
- **GLMark2:** análoga a la antepuesta, de igual forma para OpenGL.
- **Vkmark:** es una opción a las antepuestas, direccionada a la API Vulkan.
- **BaseMark GPU:** instrumento para pruebas de ventaja de la GPU asentada en las excelentes APIs esquemas, como Vulkan.

5.3.3. Herramientas y técnicas de Benchmark

Las herramientas de Benchmarking son útiles para diagnosticar problemas de rendimiento de un PC o para cuantificar una comparación entre el rendimiento de dos máquinas. Dada la utilidad del software de Benchmarking que haya innumerables utilidades de benchmarking disponibles en el mercado. [30]

La técnica de comparación de un Benchmark es una prueba que se utiliza para contrastar productos similares. Se trata de programa de evaluación comparativa de ordenadores que funciona ejecutando una serie de pruebas bien definidas en el PC para medir su rendimiento, en este caso la herramienta de rendimiento libre a utilizar es Phoronix Test Suite.

Phoronix Test Suite de aquí en adelante PTS es la plataforma de pruebas y evaluación más perfeccionada utilizable que suministra un marco extensivo al que se logran agregar cómodamente nuevas pruebas. El software está diseñado para

llevar a cabo de forma efectiva pruebas cualitativas y cuantitativas de forma limpia, reproducible y fácil de usar. PTS se puede utilizar para comparar el rendimiento de su ordenador con el de otros o se puede manejar dentro de la organización para fines de garantía de calidad interna, validación de hardware y gestión de integración continua de rendimiento. [31]

Además, hace que el proceso de llevar a cabo pruebas automatizadas sea increíblemente sencillo. Siempre y cuando tenga PHP instalado, en la totalidad de los procesos es únicamente cuestión de ejecutar un comando como phoronix test suite Benchmark unigine valley, y el PTS se encarga del resto, desde la gestión de dependencias hasta la instalación de pruebas, la ejecución y la agregación de resultados. [31]

5.4. Arquitectura del data center de la Universidad Politécnica Salesiana

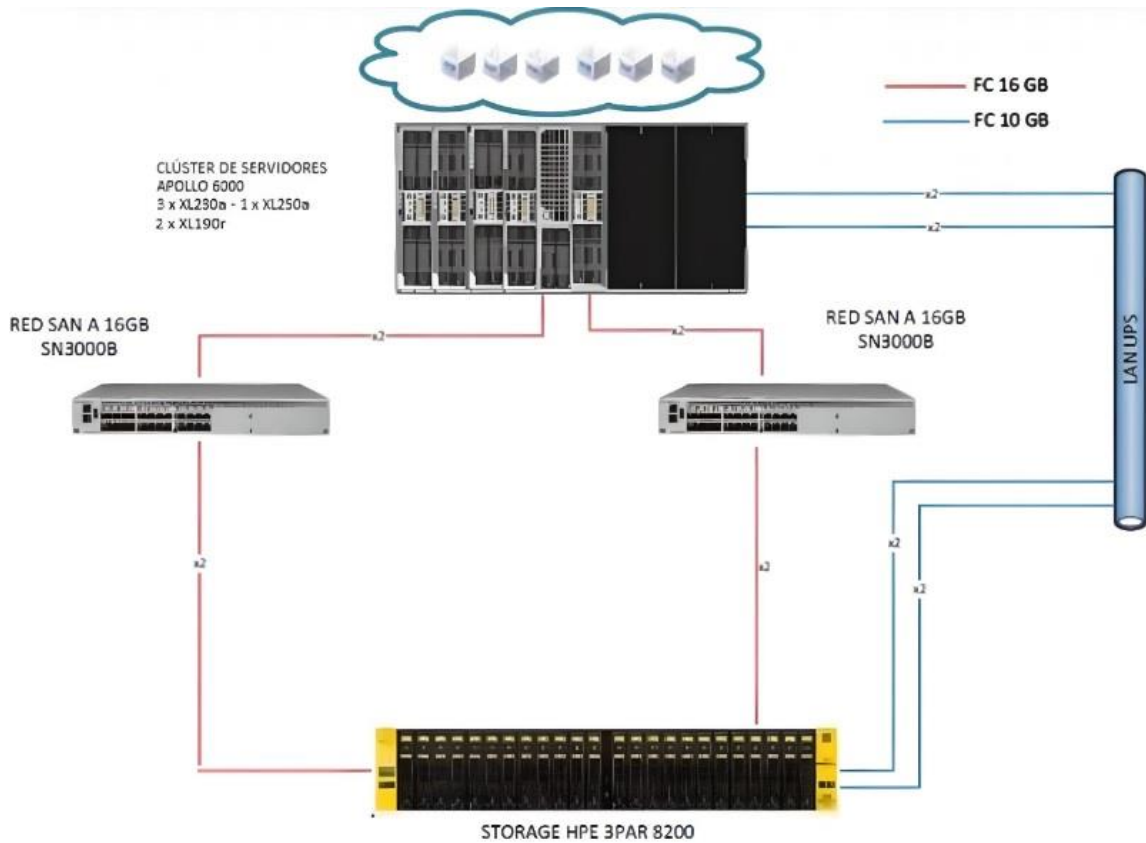
El data center de la universidad Politécnica salesiana está formado por los siguientes elementos, como lo muestra la Figura 4.

- 6 Servidores.
- Storage HPE 3PAR 8200
- Tarjetas Gráficas NVIDIA V100 Y K80

Diagrama de topología física del CPD.

Figura 4

Diagrama de topología física del CPD.

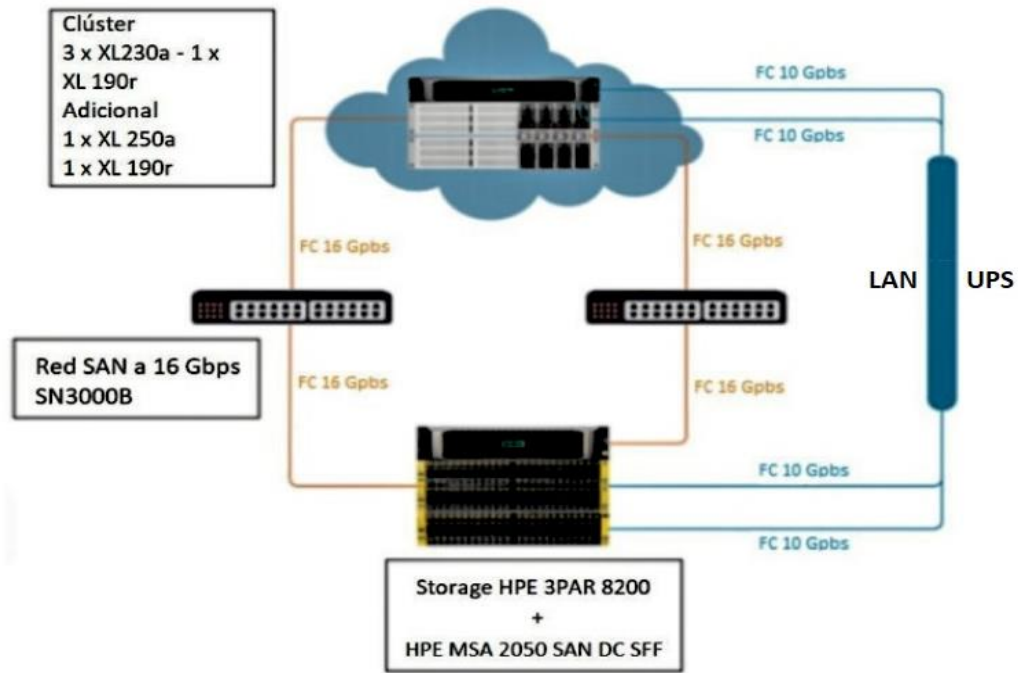


Nota. La figura 4 muestra el diagrama de topología física del CPD. Fuente: Tomado de la memoria técnica de CPD realizado por Akros.

La topología física muestra 6 servidores, con un total de 64 núcleos, en cada uno de los servidores se tiene dos procesadores, con 32 núcleos cada uno.

Figura 5

Diagrama de topología lógica del centro de datos.



Nota. La figura 5 muestra el diagrama Topología lógica del centro del procesador del centro de datos. Fuente: Tomado de la memoria técnica de CPD realizado por Akros.

Actualmente el data center maneja dos tipos de modelos de procesadores, sin embargo, uno de ellos no está dentro de la infraestructura de procesamiento del clúster, pero sigue formando parte de la misma en todo el data center. Los escenarios que se detallan a continuación de la implementación del benchmark hacen referencia a la topología lógica dado que aquí se encuentra detallada cada uno de los procesadores, así como también sus unidades de almacenamiento, en este caso el 3PAR y el MSA. El benchmark actúa sobre el clúster y cada uno de los procesadores.

6. Materiales y metodología

Para cumplir con el objetivo del trabajo se consideró necesario que se ejecuten actividades que permitan obtener los resultados, y para conseguirlo se ha seguido cuatro fases que comprende el desarrollo de esas actividades.

Fase 1: Herramientas y técnicas de Benchmark.

- Estudio sobre el funcionamiento de las herramientas y técnicas de Benchmark sobre software libre, haciendo hincapié especialmente en las tarjetas gráficas en entornos virtualizados.
- Probar las herramientas existentes que se ajusten a las necesidades del objetivo propuesto, al existir diferentes herramientas es necesario definir con cual se va a trabajar y cuál de todas se ajusta al objetivo.
- Investigar y desarrollar, teniendo en cuenta que es probable que las técnicas de Benchmark que existen no se ajusten a los objetivos planteados, sería necesario desarrollo, programar un script personalizado para medir el rendimiento de las tarjetas gráficas.

Fase 2: Elección de la herramienta y definición de las métricas de evaluación.

- En base al estudio que se realizó en la fase 1, fue importante definir que, métricas usar para poder analizar el impacto del rendimiento en tarjetas gráficas de ambientes virtualizados en software libre, para esto se ha apoyado en el análisis y estudio de la base uno pues es en esta fase donde se define que herramienta se ajustan a los objetivos y en base a eso que métricas se puede usar.

Fase 3: Diseño del o de los posibles escenarios.

- Definir los escenarios en base a los ambientes virtualizados donde se interactúe con la computación de alto rendimiento de la data center de la

universidad Politécnica Salesiana dado que pueden varían y tocaría ajustar la solución a esos casos.

- Validación de las métricas de evaluación que se escogieron en la fase 2.
- De ser necesario y de no adaptarse ninguna herramienta de Benchmark diseñar un script que englobe todas las métricas ya definidas.

Fase 4: Simulación y validación de las métricas y de la herramienta o script.

- Validación de las métricas que se escogieron para medir el impacto.
- Ejecución y validación de la herramienta o script que se haya generado o escogido para ejecutar las métricas definidas.
- Análisis de resultados.

Explicación de cada una de las fases con las implementaciones que se realizaron

Detalle Fase 1: Herramientas y técnicas de Benchmark.

En esta fase se realizó la investigación de los diferentes test que ejecuta el Benchmark para software libre, es decir, se recopiló información de las diferentes técnicas y herramientas que existen en el mercado tecnológico para determinar cuál es la que mejor se adapta a los propósitos de la investigación, dentro del análisis se investigó:

- **Geekbench:** Herramienta que puede ejecutarse en software libre o bajo licencia, muestra los resultados de un solo núcleo.
- **Phonorix Test Suite:** Herramienta de evaluación que, mediante la ejecución de pruebas de estrés, carga, rendimiento, da a conocer con porcentajes y medidas cada una de los test que fue ejecutado, cuenta con interfaz gráfica, tiene acceso a más de 100 pruebas que pueden ejecutarse de manera continua sin que esto afecte al performance de la infraestructura, permite hacer ajustes en cada prueba y se ejecución dependerá de la cantidad de test que se ejecuten, la recopilación de información la muestra en un solo archivo en formato PDF, su tiempo de ejecución en un ambiente controlado es de 15 días promedio.

- **Sysbench:** Ejecuta en simultaneado varios test de carga y rendimiento, no soporta más de 3 núcleos a la vez, necesita de un ambiente con fuertes condiciones como por ejemplo una memoria RAM de 20 gigas, un tamaño de disco de 200 GB, los resultados no los muestra en total si no por cada prueba ejecutada, tiempo de ejecución sin importar la cantidad de test de 20 a 30 días.
- **CLI Stress-ng:** Herramienta para estresar la infraestructura, se lo ejecuta en consola, no posee un adecuado informe de las pruebas ejecutadas, no analiza los núcleos en su totalidad.

En conclusión, Phonorix Test Suite es la herramienta que mejor se adapta a los objetivos de la investigación, con este Benchmark se tiene la totalidad de información recopilada de cada prueba a la que fue sometida la infraestructura, se tiene resumen de la información y su instalación no es compleja.

Haciendo uso de Phonorix Test Suite y verificando todas las ventajas que este Benchmark posee, no fue necesario realizar un Script para medir el rendimiento.

Instalación de Phonorix Test Suite en máquina virtual Linux.

- Abrir una consola de administración.
- Ejecutar el comando: `sudo gdebi phoronix-test-suite_9.8.0_all.deb`
- Ejecutar: `phoronix-test-suite`.

Especificaciones de la máquina virtual.

- Sistema Operativo: Linux – Distro Ubuntu 20.04 LTS.
- RAM: 16 Gb.
- Disco Duro: 200 Gb.
- CPU: 4 Núcleos.

Detalle Fase 2: Elección de la herramienta y definición de las métricas de evaluación.

Como se explica en la fase 1, Phonorix Test Suite es la mejor herramienta que se adapta a cumplir con los objetivos de la investigación, no se limita solo a medir el rendimiento del procesador, al contrario, da información de todos los elementos de la infraestructura que se esté analizando, en el capítulo 7 se detalla los test que se definieron para realizar el análisis, esta herramienta en comparación a las destalladas en la fase 1 no se limita al número de núcleos que tiene un procesador, los analiza de forma independiente y a la par analiza los clúster que conforma la HPC.

Las métricas de evaluación que se definieron son:

- Medir el rendimiento de los clústers en unidades Teraflops.
- Medir el rendimiento sin los clústers en unidades Teraflops.
- Medir el rendimiento de las tarjetas de video con los procesadores en unidades Teraflops.
- Medir el rendimiento de los discos en unidad lops

Cada uno de estos valores se los obtendrá con el análisis y la recopilación de los diferentes test que ejecuta el Benchmark.

Detalle Fase 3: Diseño del o de los posibles escenarios.

Los escenarios que se definen se ajustan directamente a las métricas de evaluación, es decir, para cada métrica se plante un escenario dado que las mediciones varían dependiendo de la tarjeta gráfica con la que se esté trabajando, los servidores que estén en clusterización, como se detalla en la siguiente tabla.

Tabla 1

Definición de escenarios en base a los clúster.

Servidor	Modelo	Procesador	Numero de núcleos	Clúster	Escenario
----------	--------	------------	-------------------	---------	-----------

S1	ProLiant XL230a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	SI	1	
	ProLiant XL230a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32			
S2	ProLiant XL230a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	SI		
	ProLiant XL230a Gen	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32			
S3	ProLiant XL230a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	SI		
	ProLiant XL230a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32			
S5	ProLiant XL190r Gen10	Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz	32	SI		
	ProLiant XL190r Gen10	Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz	32			
S6	ProLiant XL190r Gen10	Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz	32	NO		2
	ProLiant XL190r Gen10	Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz	32			

S4	ProLiant XL250a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	NO	
	ProLiant XL250a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32		

Nota. La tabla 1 muestra la definición de los escenarios en base a los diferentes clústers de los servidores.

Tabla 2

Definición de escenarios en base a las tarjetas gráficas.

Tarjeta gráfica	Modelo	Escenario
NVIDIA V100	V100	3
NVIDIA K80	K80	

Nota. La tabla 2 muestra la definición de los escenarios en base a las tarjetas gráficas.

Tabla 3

Definición de escenarios en base a los discos.

Almacenamiento	Discos SAS	Discos SSD	Escenario
MSA	24		4
3PAR-U	14	10	
3PAR-M		40	
Total	38	50	

Nota. La tabla 3 muestra la definición de los escenarios en base a los diferentes discos de almacenamiento disponible en la infraestructura.

Cada uno de estos escenarios tendrá un nombre definido que se los tratará a mayor detalle en el capítulo de resultados y análisis.

- Escenario 1: Análisis de rendimiento del procesamiento
- Escenario 2: Análisis del rendimiento de la infraestructura de procesamiento del centro de datos
- Escenario 3: Análisis del rendimiento del GPU para tarjetas gráficas.

- Escenario 4: Cálculos de rendimiento para almacenamiento.

En esta fase no fue necesario generar ningún Script de ejecución dado que el Benchmark arroja resultados en base a los test que se ejecutan.

Detalle Fase 4: Simulación y validación de las métricas y de la herramienta o script.

La simulación en un primer inicio de pruebas se las realizo en una maquina personal con software libre, Linux Distribución Ubuntu, con las siguientes características:

- Disco Solido de 1 Tb.
- 16 Gigas de RAM.
- Procesador Intel Core I5, séptima generación.
- Tarjeta de video MSI Force GT 710 Low Profile Graphics.

Una vez que el Benchmark fue instalado y ejecutado siguiendo los pasos de la fase 2, los resultados obtenidos fueron positivos, es decir, se tenía mediciones reales, se aportaba información de los test que fueron ejecutados, y se podía medir el rendimiento de la computadora personal.

En base a esta prueba se definió ejecutar el Benchmark en un ambiente virtualizado que esté conectado a la infraestructura del data center de la Universidad politécnica Salesiana, con los recursos que explica en el capítulo 7 bajo un mismo ambiente de software libre y la mismo distribución en nuestro caso Ubuntu.

En al ambiente de prueba (maquina personal), las métricas de evaluación que se definieron en la fase 2, fueron valederas, los resultados que arrojaba el Benchmark se adaptaba a las mediciones plateadas, por lo tanto, se hizo válidas las métricas para el ambiente virtualizado que está conectado al data center de la Universidad Politécnica Salesiana.

Cabe recalcar que no se pretende comparar las mediciones obtenidas, solo se aborda el tema de análisis con las métricas de evaluación que se definieron y los resultados que estos arrojaron y si se tiene un impacto al usar ambientes

virtualizados, o si la infraestructura del data center de la Universidad Politécnica Salesiana puede soportar dichos ambientes.

7. Resultados y discusión

7.1. Análisis de datos - Resultados

7.1.1. Escenario 1

a) Análisis de rendimiento del procesamiento

Esta sección se centra en las pruebas de rendimiento de GPC y GPU, la herramienta del apartado 5.3.3 ofrece diferentes test de ejecución con la aplicación en su total de la herramienta tenemos un reporte (ver Anexo 1) de toda la infraestructura de la data center.

La ejecución de las pruebas del Benchmark Phoronix Test Suite se realizaron durante 15 días, en una máquina virtual que tiene las siguientes especificaciones:

- Sistema Operativo: Linux – Distro Ubuntu 20.04 LTS.
- RAM: 16 Gb.
- Disco Duro: 200GB.
- CPU: 4.
- Benchmark: Phoronix Test Suite.

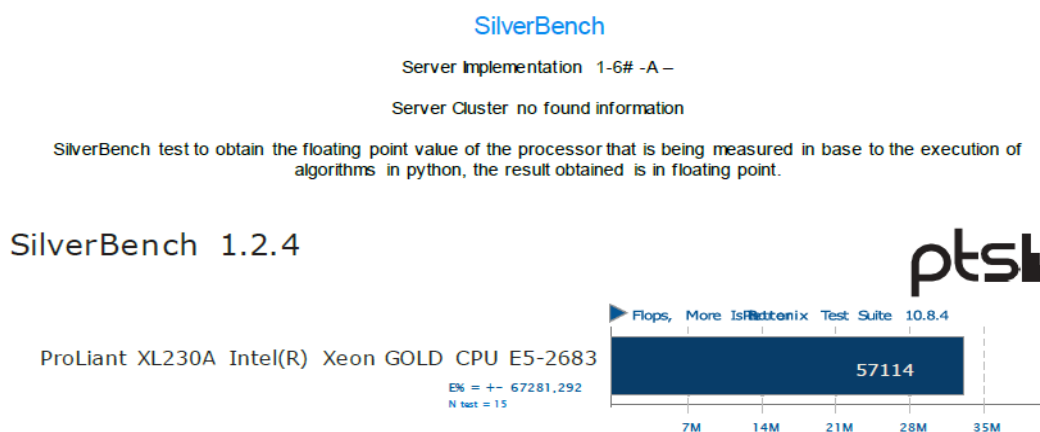
Cabe mencionar que toda la infraestructura fue sometida a prueba con la misma herramienta (Benchmark Phoronix Test Suite) teniendo un total de 295 pruebas presentadas en un PDF de 65 hojas. En estas 265 pruebas se tiene todo el resultado de los diferentes test a los que fue sometida la infraestructura en base a la ejecución del Benchmark que se expuso en los materiales y la metodología, así como el tiempo, el porcentaje de error, tipo de disco duro, tipo de tarjeta gráfica, unidad de medición, especificación de cada prueba realizada con diferentes escenarios, para una base de datos extensa para un mejor análisis.

b) Recopilación de resultados de los diferentes servidores

- **Resultado ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz**

Figura 4

Resultados



Nota. La figura 4 muestra los primeros resultados de la ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU E5-2683.

Tabla 4

Resultados ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU

Nombre	Valor punto de como flotante	% error GFlops	Número de pruebas
ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	57114	17.8%	15
Valor en GFlops	67281.29		

Nota. La tabla 4 muestra los primeros resultados de la ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU E5-2683.

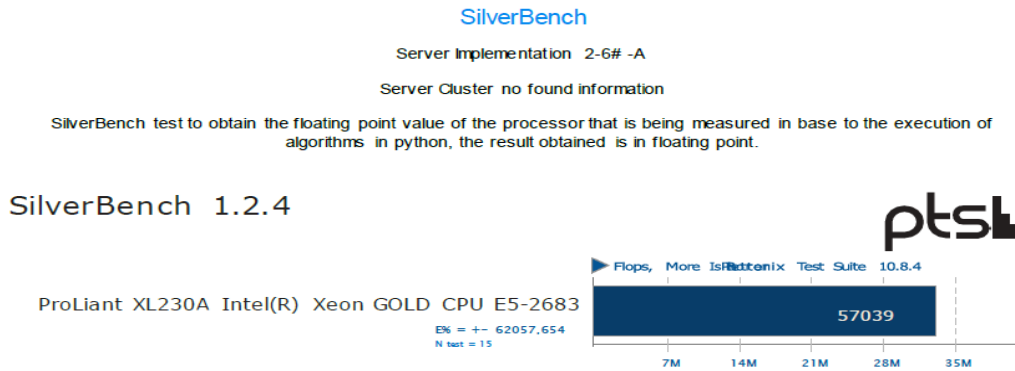
Cabe recalcar que el porcentaje de error entre ambientes virtualizados y la parte tangible es 17.8%, a fin de trabajar con este importante valor se aplicará la siguiente formula en todos los valores obtenidos en la herramienta:

$$\% \text{ de error} = \text{punto de coma flotante} + 17.8\%$$

- **Resultado: ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHZ**

Figura 5

Resultados ProLiant XL230 a Gen 9 Intel (R) Xeon (R)



Nota. La figura 5 muestra los resultados de ProLiant XL230 a Gen 9 Intel (R) Xeon (R).

Tabla 5

Resultados ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU

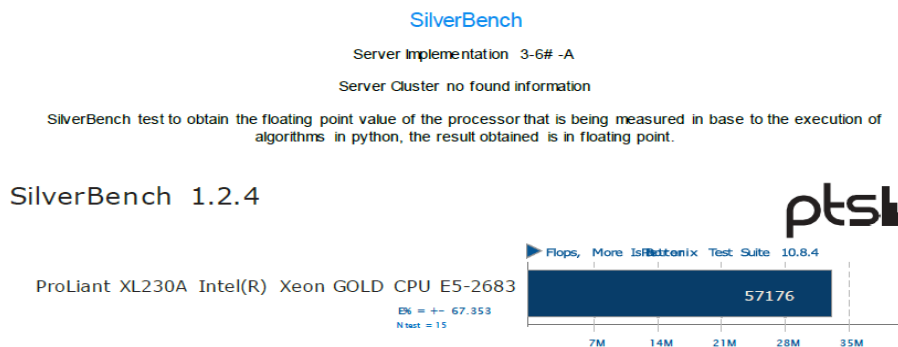
Nombre	Valor de punto de como flotante	% ERROR GFlops	Número de pruebas
ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	57039	17.8%	15
Total GFlops	62057.65		

Nota. La tabla 5 muestra los resultados de ProLiant XL230 a Gen 9 Intel (R) Xeon (R).

- **Resultado: ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHZ**

Figura 6

Resultados ProLiant XL230 a Gen 9 Intel (R) Xeon (R)



Nota. La figura 6 muestra los resultados de ProLiant XL230 a Gen 9 Intel (R) Xeon (R).

Tabla 6

Resultados ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU

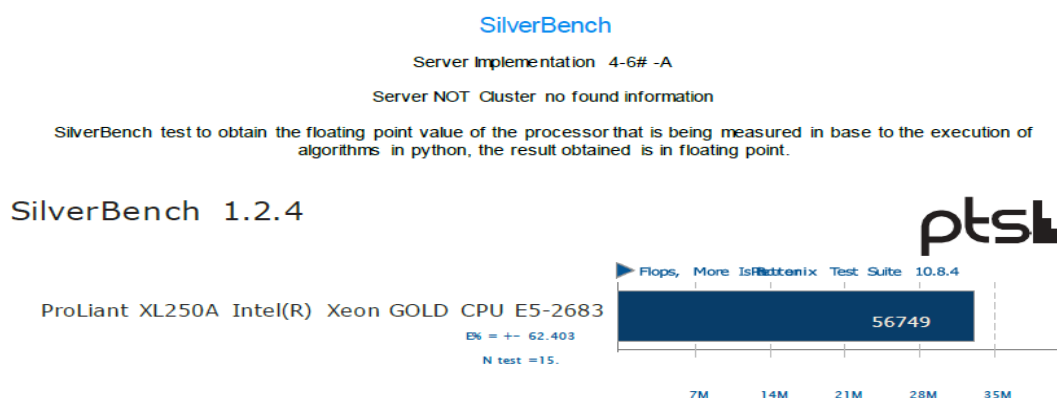
Nombre	Valor de punto de como flotante	% ERROR GFlops	Número de pruebas
ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	57176	17.8%	15
Total GFlops	67353.08		

Nota. La tabla 6 muestra los resultados de ProLiant XL230 a Gen 9 Intel (R) Xeon (R).

- **Resultado: ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHZ**

Figura 7

Resultados ProLiant XL230 a Gen 9 Intel (R) Xeon (R)



Nota. La figura 7 muestra los resultados de ProLiant XL230 a Gen 9 Intel (R) Xeon (R).

Tabla 7

Resultados ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU

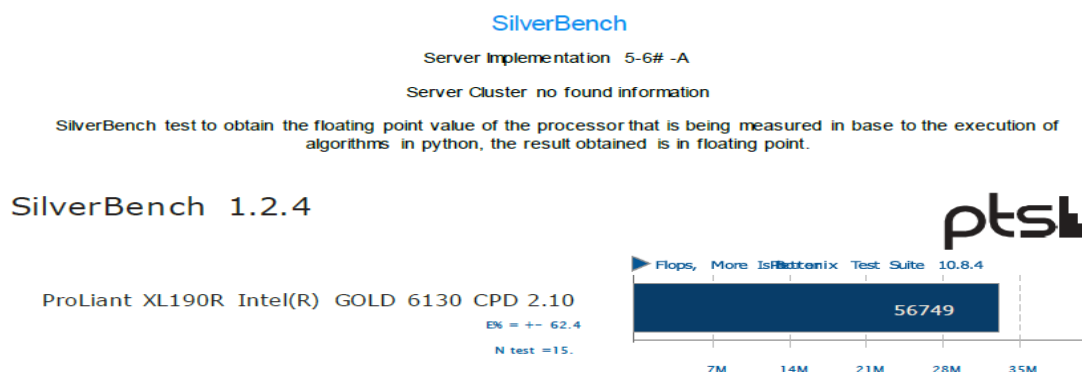
Nombre	Valor de punto de como flotante	% ERROR GFlops	Número de pruebas
ProLiant XL250a Gen9 Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	56749	17.8%	15
Total GFlops	62403.11		

Nota. La tabla 7 muestra los resultados de ProLiant XL230 a Gen 9 Intel (R) Xeon (R).

- **ProLiant XL190r Gen10 Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz**

Figura 8

Resultados ProLiant XL230 a Gen 9 Intel (R) Xeon (R)



Nota. La figura 8 muestra los resultados de ProLiant XL230 a Gen 9 Intel (R) Xeon (R).

Tabla 8

Resultados ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU

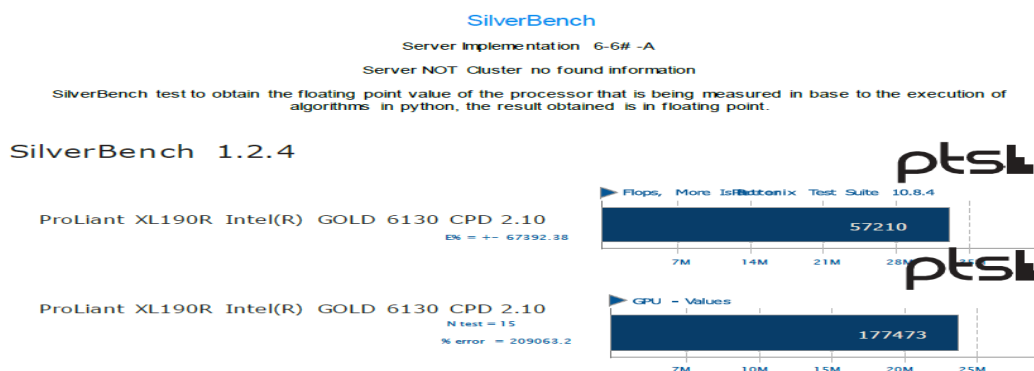
Nombre	Valor de punto de como flotante	% ERROR GFlops	Número de pruebas
ProLiant XL250a Gen9 Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	56749	17.8%	15
Total GFlops	62403.24		

Nota. La tabla 8 muestra los resultados de ProLiant XL230 a Gen 9 Intel (R) Xeon (R).

- **ProLiant XL190r Gen10 Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz**

Figura 9

Resultados ProLiant XL230 a Gen 9 Intel (R) Xeon (R)



Nota. La figura 9 muestra los resultados de ProLiant XL230 a Gen 9 Intel (R) Xeon (R).

Tabla 9

Resultados ProLiant XL230a Gen9 Intel(R) Xeon(R) CPU

Nombre	Valor de punto de como flotante	% ERROR GFlops	Número de pruebas
ProLiant XL250a Gen9 Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	57210	17.8%	15
Total GFlops	67392.38		

Nota. La tabla 9 muestra los resultados de ProLiant XL230 a Gen 9 Intel (R) Xeon (R).

b) Cálculo de rendimiento para procesamiento

Una vez obtenidos los valores de cada uno de los procesadores de los servidores que arroja el Benchmark, se aplica la siguiente formula:

$$rendimiento = \frac{\% \text{ error}}{\text{número de pruebas}}$$

Tabla 10

Rendimiento para posicionamiento

Servidor	Modelo	Total	Rendimiento en GFlops
S1	XL230a Gen9	67281.29	4485.42
S2	XL230a Gen9	62057.65	4137.18
S3	XL230a Gen9	67353.33	4490.22
S4	XL250a Gen9	49048.063	3269.87
S5	XL190r Gen10	62403.372	4160.23
S6	XL190r Gen10	67393.380	4492.89

Nota. La tabla 10 muestra los resultados de rendimiento de posicionamiento.

Para obtener el cálculo total del procesamiento es importante trabajar en unidades TFlops, teniendo esto en consideración se establece la siguiente transformación:

1 GFlop = 0.001 TFlops, es decir:

$$TFlops = GFlops * 0.001$$

Tabla 11

Rendimiento para posicionamiento TFlops

Servidor	Modelo	Total	Rendimiento en GFlops
1	XL230a Gen9	4485.42	4.485
2	XL230a Gen9	4137.18	4.137
3	XL230a Gen9	4490.22	4.490
4	XL250a Gen9	3269.87	3.270
5	XL190r Gen10	4160.23	4.161
6	XL190r Gen10	4492.89	4.494
Total			25.0358

Nota. La tabla 11 muestra los resultados de rendimiento de posicionamiento TFlops.

El total del rendimiento para procesamiento de los servidores es de **25,0358059 TFlops**.

La implementación del Bechmark se aplica a toda la infraestructura que se define en el apartado 5.4, para la topología física del CPD y para cada procesador del centro de datos.

7.1.2. Escenario 2

Análisis del rendimiento de la infraestructura de procesamiento del centro de datos

a) Rendimiento del CLUSTER

Para medir el rendimiento del clúster nos vamos en dos escenarios:

1.- Rendimiento con los datos de cada uno de los servidores:

Para esta métrica se toma en cuenta la información detalla en la tabla 1, donde se detalla cada uno de los servidores, sus especificaciones y si estos conforman el clúster.

Tabla 12

Cantidad de procesadores de la data Center

Servidor	Modelo	Procesador	Numero de núcleos	Clúster
S1	ProLiant XL230a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	SI
	ProLiant XL230a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	
S2	ProLiant XL230a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	SI
	ProLiant XL230a Gen	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	
S3	ProLiant XL230a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	SI
	ProLiant XL230a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	
S4	ProLiant XL250a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	NO
	ProLiant XL250a Gen9	Intel(R) Xeon(R) CPU E5-2683 v4 @ 2.10GHz	32	
S5	ProLiant XL190r Gen10	Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz	32	SI
	ProLiant XL190r Gen10	Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz	32	
S6	ProLiant XL190r Gen10	Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz	32	NO
	ProLiant XL190r Gen10	Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz	32	

Nota. La tabla 12 muestra cantidad de procesadores de la data Center.

Obteniendo los datos de los servidores que están en el clúster podemos sacar el rendimiento aplicando los siguientes conceptos:

Cada procesador que se diferencia en la tabla 1 obtiene los siguientes datos, en la columna por nombre Servidor muestra los otros servidores que dispone el centro de datos, para ello se toma en los servidores que son parte del clúster, cada uno de estos servidores cuentan con 8 procesadores de 32 núcleos, teniendo un total de 256 GHZ, para saber el rendimiento total se aplica la siguiente formula:

Tener en cuenta que la Suma (Servidor Clúster) tiene un valor por procesador que se toma en consideración en la tabla # de 2.89 TFlops lo que da como resultado total 17.27 TFlops.

$$Total\ Clúster = \frac{Suma(Número\ de\ Núcleos) * Suma(Servidor\ Clúster)}{Número\ de\ Núcleos\ de\ Prueba}$$

$$Total\ Clúster = \frac{256Ghz * 17.273\ TFlops}{64}$$

$$Total\ Clúster = 69.08\ TFlops$$

b) Rendimiento con los datos al ejecutar el Benchmark Phoronix Test Suite:

Teniendo en cuenta los valores obtenidos en las pruebas de Benchmark se obtiene lo siguiente:

NAS Parallel: un conjunto de puntos de referencia que tienen como objetivo la evaluación del rendimiento de las supercomputadoras tomando en cuenta para su cálculo 8 clases de ejecución o puntos de referencia:

- Five kernels:
 - IS - Ordenación de enteros, acceso aleatorio a la memoria
 - EP - Embarassingly Paralelo
 - CG - Gradiente conjugado, acceso a memoria irregular y comunicación
 - MG - Multi-Grid en una secuencia de mallas, comunicación de larga y corta distancia, uso intensivo de memoria
 - FT - Transformada de Fourier rápida 3D discreta, comunicación de todos a todos
- Three pseudo application:
 - BT - Bloque Tri-diagonal solver
 - SP - Solucionador escalar pentadiagonal
 - LU - Solucionador de Gauss-Seidel inferior-superior

Tabla 13

Tipos de prueba

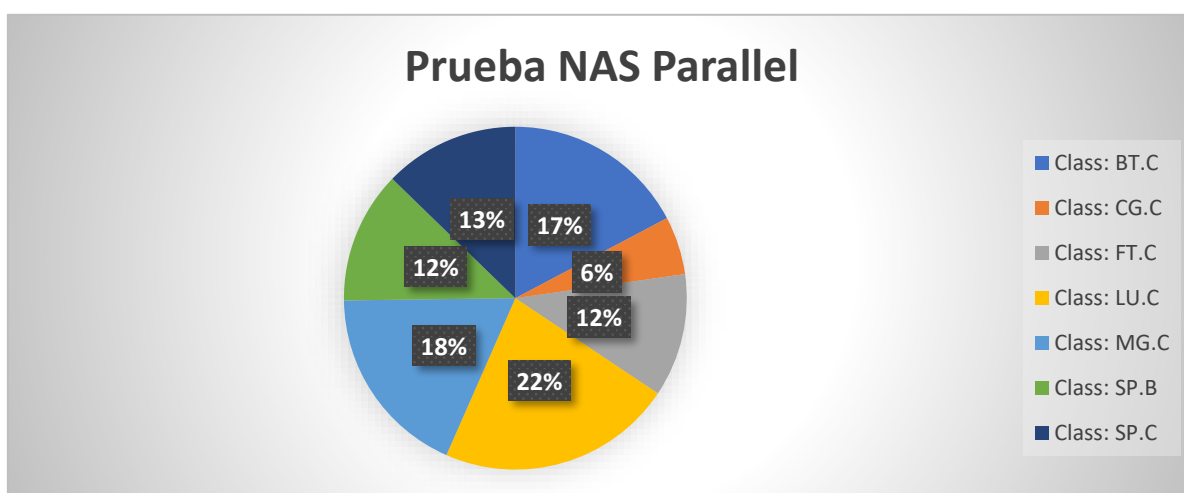
Prueba	Unidad MOPS	Tipo de prueba	Unidad TFlops
NAS Parallel	11942,50	Class: BT.C	11,94
	3787,40	Class: CG.C	3,79
	8038,29	Class: FT.C	8,04

	15355,20	Class: LU.C	15,35
	12580,91	Class: MG.C	12,58
	8619,80	Class: SP.B	8,62
	8808,63	Class: SP.C	8,81

Nota. La tabla 13 muestra los tipos de pruebas.

Figura 10

Resultados expresados en porcentaje de la prueba NAS Parallel



Nota. La Figura 10 muestra la ponderación en porcentaje del resultado de la prueba NAS Parallel, tomando en consideración que las pruebas con mas ponderación en unidad TFlops son las clases LU.C y MG.C.

Total del rendimiento del Clúster: 69, 13273 TFlops

Adicional, comprando estos dos valores, podemos afirmar que la medición mediante las 7 pruebas de NAS Parallel que realizo el bechmark para los servidores en conjunto son exitosas pues no existe diferencia entre los valores tomados de la virtualización versus los valores calculados.

c) Rendimiento total Cluster de la infraestructura Data Center

Para obtener el rendimiento total de la infraestructura se procede a sumar los dos servidores independientes que son parte de la data center, en este caso S4 y S6 que se detalla en la tabla 1.

Para ello aplicaremos la siguiente formula:

$$Rendimiento\ Total = Rendimiento\ Clúster + clúster\ independientes,$$

Donde:

$$Clúster\ independientes = 3,2698709 + 4.492892\ TFlops\ (Mirar\ tabla\ 8)$$

$$Rendimiento\ Total = 69,13273 + 7.7627629$$

$$Rendimiento\ Total = 76.8954\ TFlops$$

7.1.3. Escenario 3

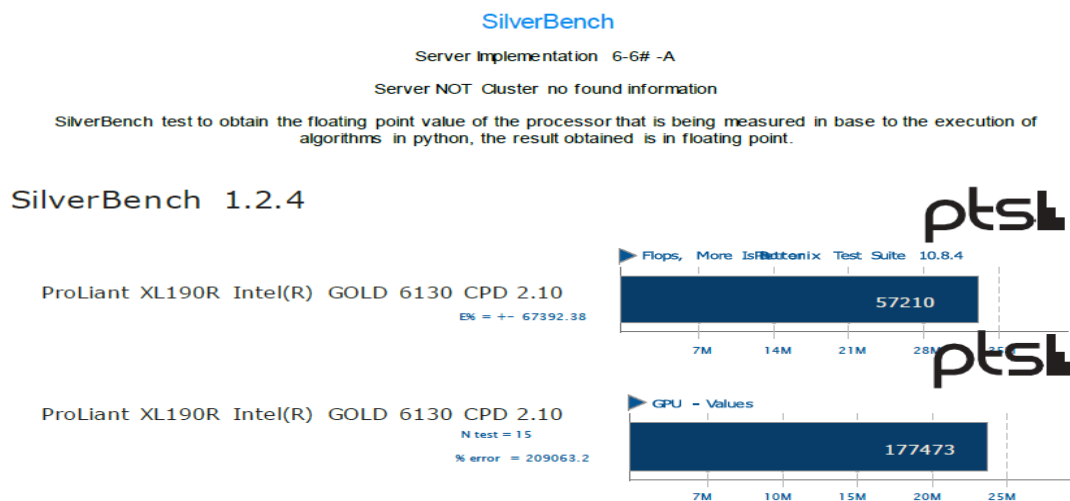
Análisis del rendimiento del GPU para ProLiant XL190r Gen10 con la tarjeta envidia V100.

Actualmente la data center cuenta con dos servidores HPE ProLiant XL190r con tarjeta gráfica NVidia Testa V100, con alrededor de 5000GPU, del cual no está en el cluster y el otro sí, es así que se divide los escenarios y se saca el total de rendimiento de los servidores clúster y de los que no están en clúster, información en tabla 1.

Benchmark Phoronix Test Suite nos aporta (ver anexos) los valores para obtener el cálculo del rendimiento de ProLiant XL190r Gen10 con la tarjeta NVIDIA V100.

Figura 11

Resultados Benchmark Phoronix Test Suite



Nota. La figura 11 muestra los resultados del Benchmark Phoronix Test Suite.

Tabla 14

Resultados Benchmark Phoronix Test Suite

NAME	Punto coma flotante	% de Error	GPU	% de Error GFlops	Número Pruebas
ProLiant XL190r Gen10	57210	67393.38	177473	17.8%	15
Total GFlops			209063.2		

Nota. La tabla 14 muestra los resultados del Benchmark Phoronix Test Suite.

Para calcular el valor de rendimiento de la GPU se aplica la siguiente fórmula con los datos que nos arroja el Benchmark:

$$\text{Valor GPU} = \frac{\% \text{ de error GFlops}}{\text{número de pruebas}}$$

$$\text{Valor GPU} = \frac{209063.2}{15}$$

$$\text{Valor GPU} = 13937,55 \text{ GFlops}$$

Expresando los valores en TFlops, se obtiene:

$$\text{Valor GPU} = 13937,55 \text{ GFlops} * 0.001$$

$$\text{Valor GPU} = 13,94 \text{ TFlops}$$

Obteniendo un resultado de 13,94 Tflops en el beneficio de GPU del Servidor ProLiant XL190r Gen10 que es parte de con una tarjeta gráfica NVIDIA V100.

7.1.4. Escenario 4

Análisis del rendimiento del GPU para ProLiant XL250A con la tarjeta envidia K80.

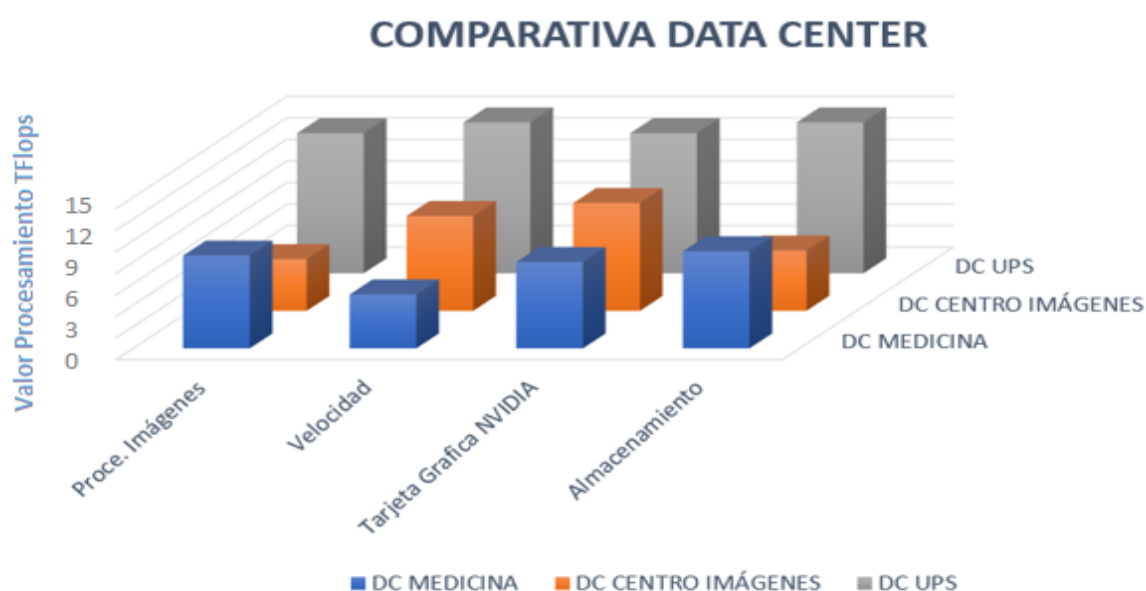
Al realizar la recolección de datos de los 6 servidores presentes dentro de la infraestructura, el xl250 cuya grafica no es calculada (GPU) ya que según NVIDIA, para ellos es sustancial un sistema operativo propio. La tarjeta K80 de acuerdo con NVIDIA manifiesta que hasta 2,91 Teraflops de beneficio en sistematizaciones de

exactitud duplo con NVIDIA GPU Boost y hasta 8,73 Teraflops de beneficio en ordenamientos de exactitud simple con NVIDIA GPU Boost.

Cabe recalcar que el presenta análisis se obtiene el un sistema operativo Linux distro Ubuntu virtualizado, al trabajar en ambiente virtualizados el bechmark no ejecuta todas las pruebas con relación a tarjetas gráficas, en este caso la K80.

Figura 12

Cuadro comparativo de datas center vs data center UPS



Nota. La figura 12 indica la comparación con diferentes datas center vs los resultados obtenidos

Contrarrestando los resultados obtenidos y la imagen presentada, y una vez obtenido todo el rendimiento de la arquitectura del HPC podemos ver que los valores que obtuvieron no están alejados de la realidad, por ejemplo el procesamiento de un data center que ocupa procesamiento de imágenes obtiene un valor de 6.X Tflops, el valor obtenido con el Bechmark de estudio y con las tarjetas gráficas que se está trabajando tiene un total de 13.94 Tflops estando por encima de los valores mostrados en la gráfica de diferentes data center promedios.

7.2. Análisis del rendimiento del almacenamiento

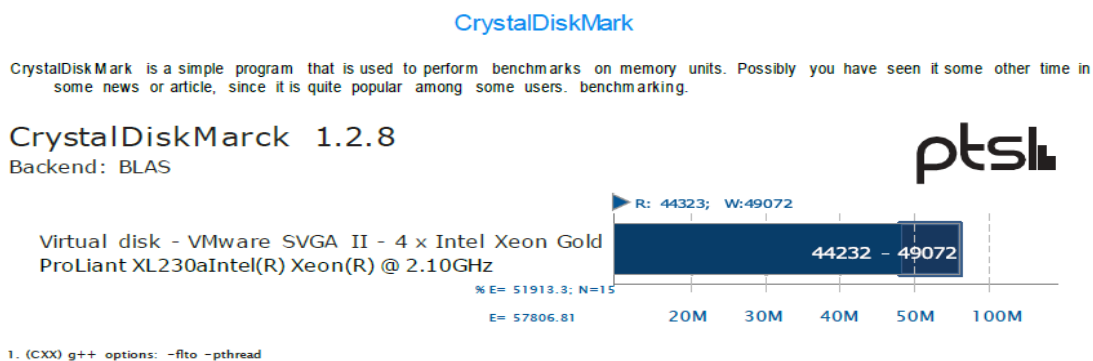
7.2.1. Escenario 1

Para el análisis del rendimiento del almacenamiento se utilizará las métricas obtenidas en el Benchmark CrystalDiskMark (ver anexo 1), donde muestra los valores de medición de cada uno de los discos de los servidores.

a) ProLiant XL230a Gen9 Intel® Xeon® CPU E5-2683 v4 @ 2.10GHz

Figura 13

Resultados ProLiant XL230a Gen9 Intel® Xeon®



Nota. La figura 13 muestra los resultados del ProLiant XL230a Gen9 Intel® Xeon®.

Tabla 15

Resultados ProLiant XL230a Gen9 Intel® Xeon®

Total de Pruebas	Lectura	% de error	Escritura	% de error
15	44423	17.8%	49072	17.8%
Total	51913.2		57806.81	

Nota. La tabla 15 muestra los resultados del ProLiant XL230a Gen9 Intel® Xeon®.

b) ProLiant XL230a Gen9 Intel® Xeon® CPU E5-2683 v4 @ 2.10GHz

Figura 16

Resultados ProLiant XL230a Gen9 Intel® Xeon®

CrystalDiskMark

CrystalDiskMark is a simple program that is used to perform benchmarks on memory units. Possibly you have seen it some other time in some news or article, since it is quite popular among some users. benchmarking.

CrystalDiskMarck 1.2.8

Backend: BLAS



Nota. La figura 16 muestra los resultados del ProLiant XL230a Gen9 Intel® Xeon®.

Tabla 16

Resultados ProLiant XL230a Gen9 Intel® Xeon®

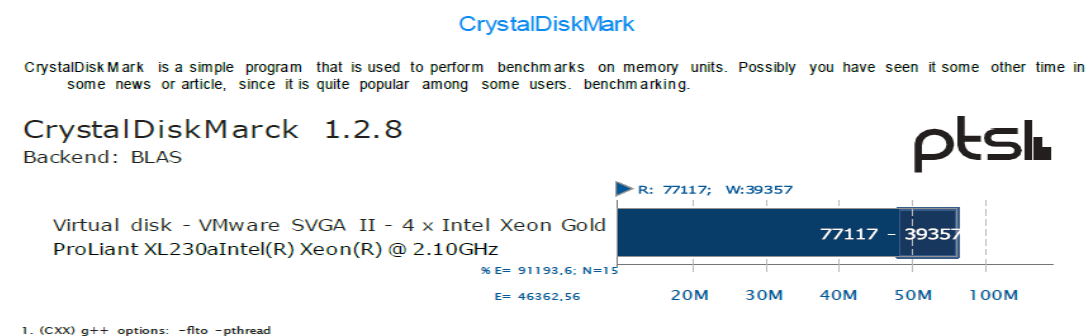
Total de Pruebas	Lectura	% de error	Escritura	% de error
15	37117	17.8%	48606	17.8%
Total	43723.79		57257.82	

Nota. La tabla 16 muestra los resultados del ProLiant XL230a Gen9 Intel® Xeon®.

c) ProLiant XL230a Gen9 Intel® Xeon® CPU E5-2683 v4 @ 2.10GHZ

Figura 17

Resultados ProLiant XL230a Gen9 Intel® Xeon®



Nota. La figura 17 muestra los resultados del ProLiant XL230a Gen9 Intel® Xeon®.

Tabla 17

Resultados ProLiant XL230a Gen9 Intel® Xeon®

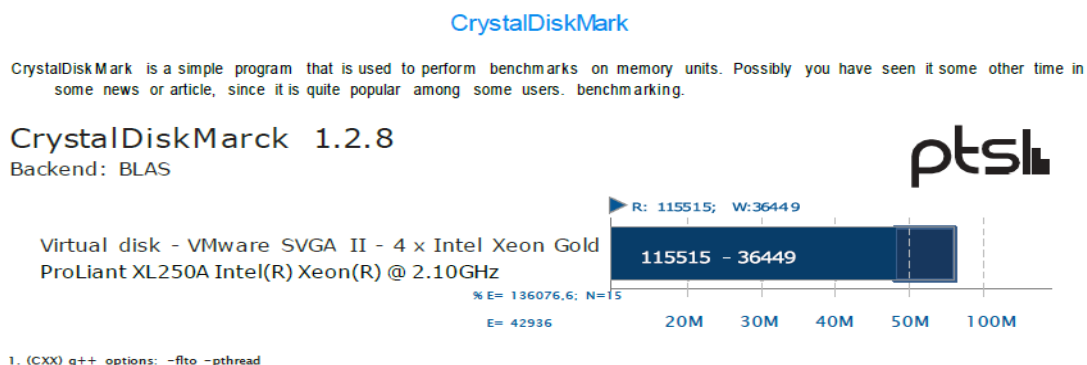
Total de Prueba	Lectura	% de error	Escritura	% de error
15	77117	17.8%	39357	17.8%
Total	91193.69		46362.56	

Nota. La tabla 17 muestra los resultados del ProLiant XL230a Gen9 Intel® Xeon®.

d) ProLiant XL250a Gen9 Intel® Xeon® CPU E5-2683 v4 @ 2.10GHz

Figura 18

Resultados ProLiant XL250a Gen9 Intel® Xeon®



Nota. La figura 18 muestra los resultados del ProLiant XL250a Gen9 Intel® Xeon®.

Tabla 18

Resultados ProLiant XL250a Gen9 Intel® Xeon®

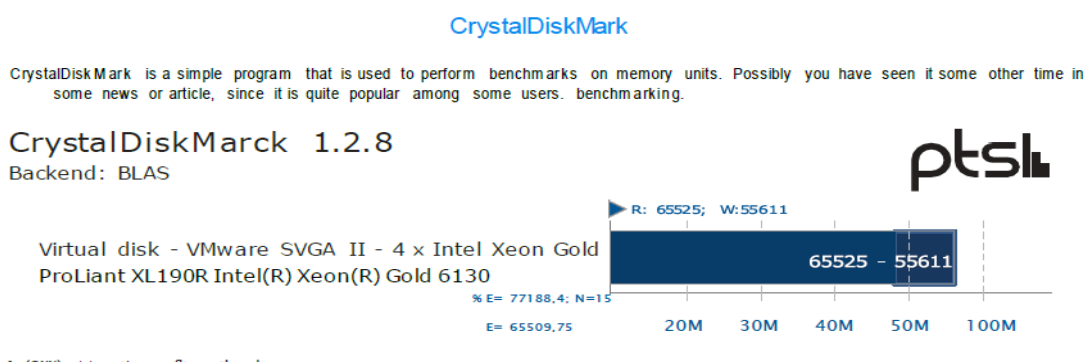
Total de Pruebas	Lectura	% de error	Escritura	% de error
15	115515	17.8%	36449	17.8%
Total	136076.66		42936.93	

Nota. La tabla 18 muestra los resultados del ProLiant XL250a Gen9 Intel(R) Xeon(R).

e) ProLiant XL190r Gen10 Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz

Figura 19

Resultados ProLiant XL190a Gen10 Intel(R) Xeon(R)



Nota. La figura 19 muestra los resultados del ProLiant XL190a Gen10 Intel(R) Xeon(R).

Tabla 19

Resultados ProLiant XL190 Gen10 Intel(R) Xeon(R)

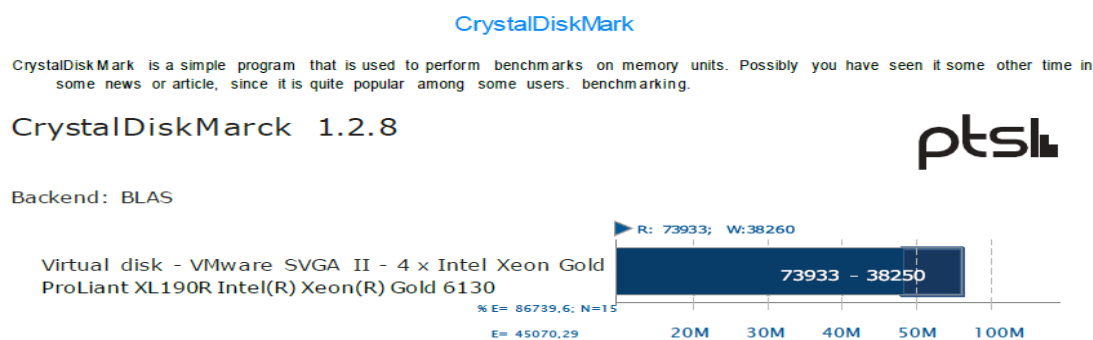
Total de Pruebas	Lectura	% de error	Escritura	% de error
15	65525	17.8%	55611	17.8%
Total	77188.49		65509.75	

Nota. La tabla 19 muestra los resultados del ProLiant XL0190a Gen10 Intel(R) Xeon(R).

f) ProLiant XL190r Gen10 Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz

Figura 20

Resultados ProLiant XL190a Gen10 Intel(R) Xeon(R)



Nota. La figura 20 muestra los resultados del ProLiant XL190a Gen10 Intel(R) Xeon(R).

Tabla 20

Resultados ProLiant XL190 Gen10 Intel(R) Xeon(R)

Total de Pruebas	Lectura	% de error	Escritura	% de error
15	73933	17.8%	38260	17.8%
Total	86739.65		45070.29	

Nota. La tabla 20 muestra los resultados del ProLiant XL0190a Gen10 Intel(R) Xeon(R).

Cabe recalcar, que los porcentajes de error tanto en lectura y escritura la herramienta ya los calcula, teniendo en cuenta que para un error de lectura se suma un 17.8%, de la misma manera en escritura, esto ya se considera en cada una de las tablas de los discos de cada servidor.

Total, de almacenamiento:

Tabla 21

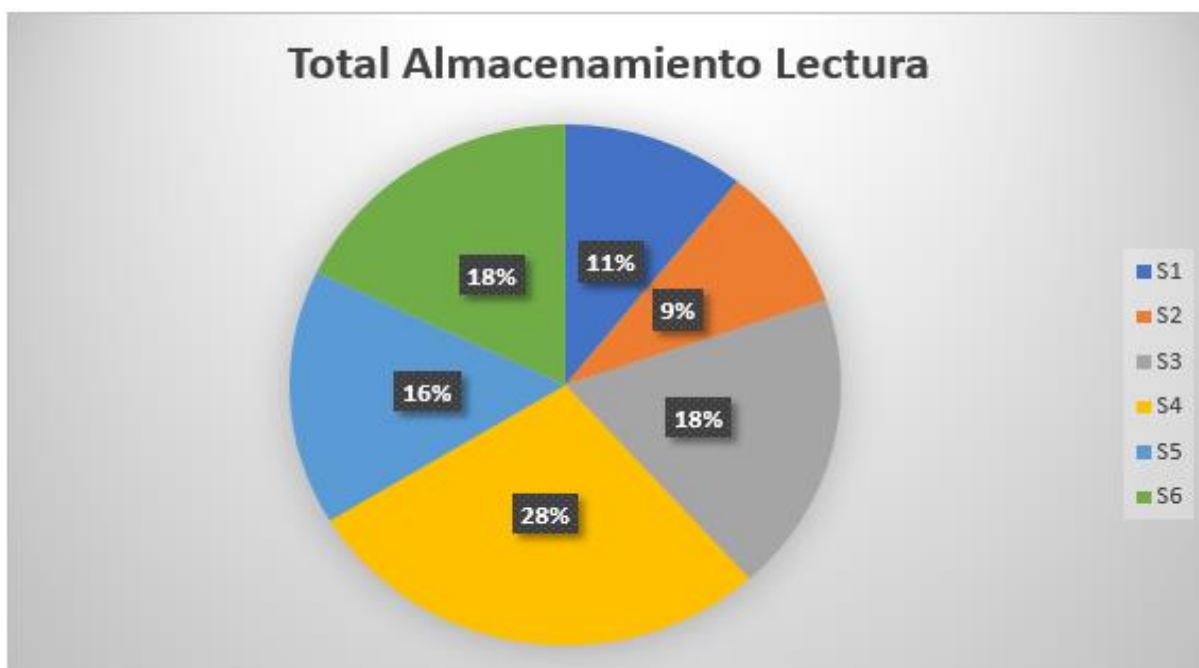
Tabla total de almacenamiento

Servidor	Modelo	Total Read	% E Lectura	Total Write	% E Escritura
S1	XL230a Gen9	44423	51.91%	49072	58.81
S2	XL230a Gen9	37117	44.79%	8606	58.82
S3	XL230a Gen9	77117	91.69%	39357	46.56
S4	XL250a Gen9	115515	14.66%	36449	43.93
S5	XL190r Gen10	65525	77.49%	55611	66.75
S6	XL190r Gen10	73933	87.65%	38260	46.29
Total		413630	48.68 %	267355	31.50 %

Nota. La tabla 21 muestra el total de almacenamiento.

Figura 21

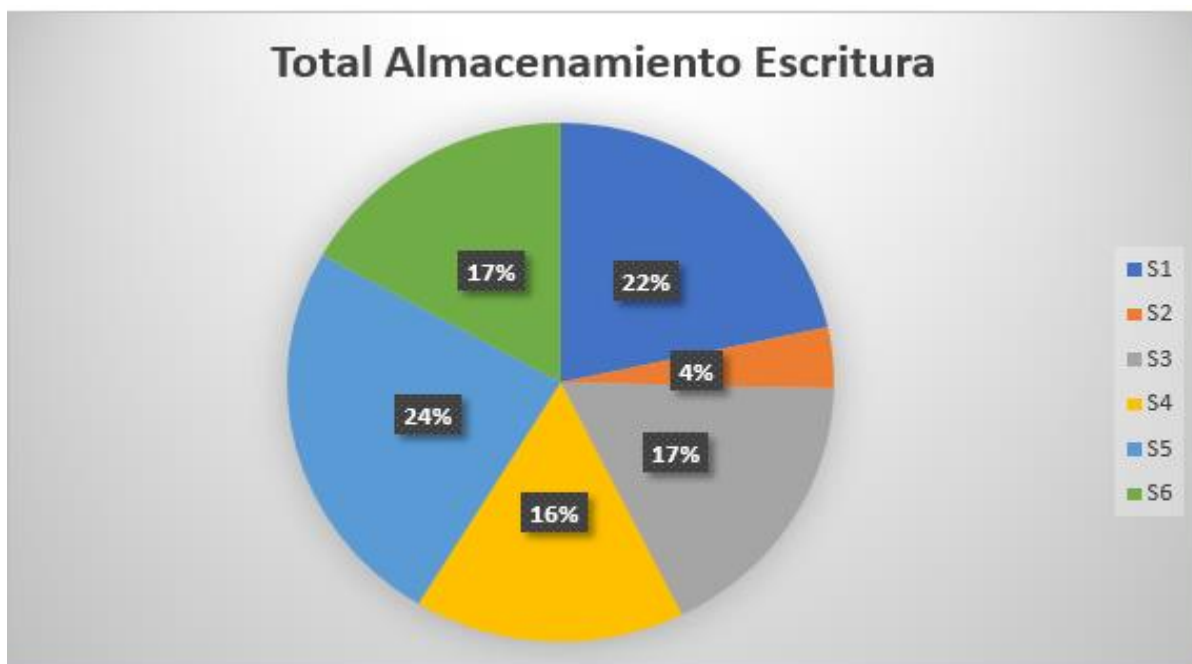
Grafica en porcentaje del almacenamiento en lectura.



Nota. La figura 21 muestra los resultados en porcentaje de las mediciones del almacenamiento en lectura siendo el mayor con un 28% el S4.

Figura 22

Grafica en porcentaje del almacenamiento en escritura.



Nota. La figura 22 muestra los resultados en porcentaje de las mediciones del almacenamiento en escritura siendo el mayor con un 24% el S5.

Para el cálculo de almacenamiento se trabaja en unidades lops de E/S. Huawei.

7.2.2. Cálculos de rendimiento para almacenamiento.

Para el cálculo de almacenamiento se trabaja en unidades lops de E/S. Huawei manifiesta que existen ciertas fórmulas para establecer el dígito real de los lops. También da a conocer que existe una penalidad sobre los RAID. Los recursos de almacenamiento de la data center se encuentran alojados en un RAID 5, por lo tanto, se tendrá una penalidad de 4.

Fórmulas para el cálculo:

$$(1) \text{ Total RAW IOPS} = \text{VEL Disco IOPS} * \text{NUM Discos}$$

$$(2) \text{ Fu IOPS} = \left(\frac{\text{Total RAW IOPS} * \% \text{ Error Escr}}{\text{Penalidad RAID}} \right) + (\text{Total RAW IOPS} + \% \text{ error Lect})$$

Cálculos de rendimiento para almacenamiento.

Para el cálculo de almacenamiento se trabaja en unidades lops de E/S. Huawei manifiesta que existen dos fórmulas para determinar el número real de todos los lops. También da a conocer que existe una penalidad sobre los RAID. Los recursos de almacenamiento de la data center se encuentran alojados en un RAID 5, por lo tanto, se tendrá una penalidad de 4.

Fórmulas para el cálculo:

$$(1) \text{ Total RAW IOPS} = \text{VEL Disco IOPS} * \text{NUM Discos}$$

$$(2) \text{ Fu IOPS} = \left(\frac{\text{Total RAW IOPS} * \% \text{ Error Escr}}{\text{Penalidad RAID}} \right) + (\text{Total RAW IOPS} + \% \text{ error Lect})$$

Para obtener la Velocidad que necesitamos en (1), se lo logra del dígito de revueltas por minuto de cada rueda de las unidades de almacenamiento. La data center posee ciertos tipos de disco, uno con interfaz SAS y otro SSD, el número de discos que se detalla:

Tabla 22

Número de discos

Almacenamiento	Discos SAS	Discos SSD
MSA	24	
3PAR-U	14	10
3PAR-M		40
Total	38	50

Nota. La tabla 22 muestra el número de discos utilizados.

Los discos SAS (Serial Attached SCSI) no son tan rápidos como los SSD (Solid-State Driver) dado que no contienen partes móviles, son hasta 10 veces más rápidos y por esta razón en número de IOPS serán mayores que los SAS, lo que esto conlleva que los discos SSD se deben calcular de forma separa con la misma fórmula ante descrita.

Tabla 23

Tabla de conversión de RMP a IOPS

Unidad de almacenamiento	Modelo	RPM	IOPS
--------------------------	--------	-----	------

SAS	SAS	10000	160
SSD HPE-3PAR	ARFX1920S5xnNTRI	150000	1750
SSD HPE-3PAR	ARFX1920S5xnNTRI	150000	1750

Nota. La tabla 23 muestra la conversión de RMO a IOPS.

Con los datos obtenidos se reemplaza en las fórmulas (1) y (2) teniendo en cuenta los cálculos por separado de SSD y SAS, y los cálculos obtenidos en la Tabla total de almacenamiento

Para SAS:

$$(1) \text{ Total RAW IOPS} = 160 * 38 = 6.080 \text{ Iops}$$

$$(2) \text{ Funcional IOPS} = \left(\frac{6.080 * 31.4944\%}{4} \right) + (6.080 * 48.6835)$$

$$(2) \text{ Funcional IOPS SAS} = 3.4389 \text{ IOPS}$$

Para SSD:

$$(1) \text{ Total R IOPS} = 1750 * 50 = 87500 \text{ Iops}$$

$$(2) \text{ Funcional IOPS} = \left(\frac{87.500 * 31.4944\%}{4} \right) + (87.500 * 48.6835)$$

$$(2) \text{ Funcional IOPS SSD} = 49.5026 \text{ IOPS}$$

Para obtener el cálculo del rendimiento total de almacenamiento, se sumará los totales obtenidos en los cálculos de SSD Y SAS, tanto para el total IOPS como para el Fun IOPS, teniendo los siguientes resultados:

$$\text{Total RAW SSD SAS} = \text{Total RAW IOPS SSD} + \text{Total RAW IOPS SAS}$$

$$\text{Total RAW SSD SAS} = 6.080 + 87.500 = 93.580 \text{ IOPS}$$

Para calcular los IOPS funcionales finales se debe sumar los cálculos obtenidos para SSD Y SAS, obteniendo el resultado final de IOPS funcionales:

$$\text{Total funcional final} = \text{funcional IOPS SSD} + \text{funcional IOPS SAS}$$

$$\text{Total funcional final} = 3.4389 + 49.5026 = 52.9416 \text{ IOPS}$$

8. Conclusiones

La utilización de técnicas de Benchmark son de gran utilidad al momento de seleccionar las métricas de medición mas adecuadas para ejecutar las 15 pruebas de procesamiento y rendimiento, al igual que determinar el porcentaje de error entre máquinas virtuales y las maquinas físicas dado que sin definir el valor de porcentaje (17.8%) los valores obtenidos no hubieran sido precisos.

El uso de elementos de medidas es un punto crítico para la generación de valores existentes y medibles, dentro de las pruebas que se ejecutaron en el ambiente virtualizado y con la ayuda del Benchmark se pudo delimitar la obtención de las unidades de medidas, de no ser este el caso llevara más tiempo el tratar de analizar e interpretar la información que arroja el Benchmark, lo esencial para este tipo de cálculos y análisis es trabajar con unidades “FLOP” para medir el rendimiento y IOPS para el almacenamiento de operaciones de E/S.

Tras el análisis de la infraestructura del data center perteneciente a la Universidad Politécnica Salesiana en el ambiente virtualizado con el uso de técnicas de benchmark da como resultado final un valor de 69.08 TFlops sin contar los clúster, un valor de virtualización total de 76.89 TFlops contando los clúster, un valor de procesamiento de la tarjeta gráfica NVIDIA tesla V100 de 13.94 TFlops e integrada por 2 servidores autónomos añadiendo un coste de 7.67 TFlops, teniendo en cuenta estos resultados el impacto en ambientes virtualizados (clúster) los valores obtenidos no tiene un impacto negativo.

Para el cálculo del rendimiento total, fue indispensable identificar los tipos de discos que hace uso de la data center de la UPS, con esta información y los resultados obtenidos en Benchmark Phoronix Test Suite se procedió a conjeturar y convertir la paridad de revoluciones por minuto a lops, adquiriendo como Total de lops 529416 lops de E/S son los utilizables intrínsecamente en la infraestructura de acopio.

Para ilustrar mejor los resultados se utilizó las gráficas de la herramienta Benchmark Phoronix Test Suite, donde indica a detalle el tipo de prueba que se realizó, los valores obtenidos a detalle y las métricas, así como el nombre de los componentes involucrados.

Referencias

- [1] M. Tomašević, L. Lapuh y Z. Stević, «Evaluation of criteria for the implementation of high-performance computing (HPC) in Danube Region countries using fuzzy PIPRECIA method,» *Sustainability*, vol. 12, nº 07, p. 3017, 2020.
- [2] D. Romero, «Benchmarks: ¿Qué es? ¿Para qué sirve? Historia, tipos y consejos,» *Profesionalreview*, 12 mayo 2018. [En línea]. Available: <https://www.profesionalreview.com/2018/05/12/benchmarks/>. [Último acceso: 24 agosto 2022].
- [3] T. Lynn, «Understanding the determinants of cloud computing adoption for high performance computing,» *51st Hawaii International Conference on System Sciences (HICSS-51)*, vol. 1, nº 1, pp. 3894-3903, 2018.
- [4] D. Domínguez, «Computación de alto rendimiento sobre un entorno de virtualización basado en contenedores,» Universidad Autónoma de Madrid, Madrid, 2016.
- [5] J. Echeverría, «Tajetas gráficas para acelerar el cómputo complejo,» *La tecnología y el cómputo avanzado*, vol. 1, nº 1, pp. 45-50, 2008.
- [6] L. A. Carpena, «Configuración y testeo de un clúster GPU.,» Escuela Politécnica Superior, Alicante, 2017.
- [7] P. Luszczek, J. Dongarra, D. Koester, R. Rabenseifner, B. Lucas y J. Kepner, «Introduction to the HPC Challenge Benchmark Suite,» University of Tennessee Knoxville, 2005.
- [8] J. Filippi, G. Lafuente y C. Ballesteros, «Experiencia de virtualización en la UNLPam,» *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, Vols. %1 de %2-, nº 9, p. 9, 2020.
- [9] L. Gómez y L. Leyva, «La virtualización como alternativa para la educación de posgrado.,» *Revista Cubana de Educación Superior*, vol. 39, nº 12, p. 19, 2020.
- [10] A. Galán, «Virtualización; eficiencia y escalabilidad,» *Revista Ingeniería y Ciencia*, vol. 2, nº 9, pp. 24-28, 2015.
- [11] M. Domínguez, «Virtualización mediante entornos Open Source,» Universidad Politécnica de Valencia, Valencia, 2018.
- [12] M. C. Campos, «OpenStack: Estudio y despliegue de herramientas de virtualización en la nube,» Universidad de Extremadura, 2019.
- [13] M. Mora y G. Galvis, «Ventajas de la virtualización en las organizaciones internacionales,» *Revista Visión Internacional (Cúcuta)*, vol. 2, nº 1, pp. 56-66, 2019.
- [14] M. Cavallo, N. Fattore y M. Geli, «Ventajas y desventajas de la virtualización de la educación en pandemia: miradas de los estudiantes de la FCEYE,» Universidad Nacional de Rosario, 2021.
- [15] F. Corrovilla, L. Salvatierra y A. Lara, «Gestión de Seguridad en Virtualización de Servidores,» *Ecuadorian Science Journal*, vol. 5, nº 4, pp. 150-162., 2021.

- [16] A. García, «Herramientas de virtualización. Tipos, sinergias y oportunidades,» UNED, 2018.
- [17] P. Bordón y F. Crespo, «NLHPC: La experiencia de trabajar usando recursos de computación de alto desempeño,» *Observatorio Económico*, Vols. %1 de %2-, nº 166, pp. 14-15, 2022.
- [18] D. González, C. Estevez, J. León y N. Galvan, «Inteligencia artificial y computación de alto rendimiento en la solución de problemas de gran complejidad,» *Universidad Industrial de Santander*, vol. 33, nº 3, pp. 1-8, 2018.
- [19] D. Medel, M. Murazzo, A. Molina, F. Sánchez y M. Cornejo, «La computación de alta performance como soporte a los sistemas altamente distribuidos,» *Red de Universidades con Carreras en Informática (RedUNCI)*, vol. 1, nº 1, pp. 1-5, 2019.
- [20] C. García, R. González y A. Jaime, «Renderizado de imágenes 3D mediante una computadora de alto rendimiento,» *Pädi Boletín Científico De Ciencias Básicas E Ingenierías Del ICBI*, vol. 10, nº 20, pp. 1-8, 2022.
- [21] M. Naiouf, A. De Giusti, L. De Giusti, F. Chichizola y V. Sanz, «Algoritmos paralelos y evaluación de rendimiento en plataformas de HPC,» *Workshop de Investigadores en Ciencias de la Computación*, vol. XXVII, nº 1, pp. 674-679, 2021.
- [22] R. Bevilacqua, «Breve historia de la evolución de HPC en una institución de investigación en la Argentina,» *SAHTI, Simposio Argentino de Historia, Tecnologías e Informática*, vol. 42, nº 1, pp. 35-42, 2020.
- [23] P. Luszczek, E. Meek, S. Moore, D. Terpstra, V. Weaver y J. Dongarra, «Evaluation of the HPC Challenge Benchmarks in Virtualized Environments,» University of Tennessee Knoxville, 2011.
- [24] M. Barrionuevo, F. Constantini, J. Escalante y M. Lopresti, «Tecnologías HPC híbridas en la solución de problemas complejos,» *Workshop de Investigadores en Ciencias de la Computación*, vol. XXIII, nº 1, pp. 744-748, 2021.
- [25] S. Gómez, J. Umaña y M. Arce, «Infraestructura de HPC del Colaboratorio Nacional de Computación Avanzada, Kabré.,» *Revista Tecnología en Marcha*, vol. 23, nº 1, p. 112, 2020.
- [26] F. Mendoza, A. Uribe, O. Moncada, D. Morales y L. Mantilla, «Implementacion de las GPUs para un futuro inteligente,» *Universidad Industrial de Santander*, vol. 3, nº 3, pp. 1-9, 2018.
- [27] L. Ke, W. Gang, C. Gong, M. Liqiu y H. Junwei, «Object detection in optical remote sensing images: A survey and a new benchmark,» *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, nº -, pp. 296-307, 2020.
- [28] J. Cows, A. Tsamados, M. Taddeo y L. Floridi, «A definition, benchmark and database of AI for social good initiatives,» *Nature Machine Intelligence volume*, vol. 3, nº -, p. 111–115, 2021.
- [29] J. Poovey, T. Conte, M. Levy y S. Gal, «A Benchmark Characterization of the EEMBC Benchmark Suite,» *IEEE Micro*, vol. 29, nº 5, pp. 18 - 29, 2009.
- [30] T. Veliz, «Utilidad del benchmarking como estrategia de mejora empresarial,» *Polo del Conocimiento: Revista científico-profesional*, vol. 6, nº 3, pp. 2026-2044, 2021.

- [31] M. López, «Comprobando el rendimiento del PC con Phoronix Test Suite,» 2018. [En línea]. Available: <https://hipertextual.com/2018/02/rendimiento-pc-benchmark-phoronix-test-suite>.
- [32] Y. Fernandez, «Tarjeta gráfica: qué es, qué hay dentro y cómo funciona,» 2020. [En línea]. Available: <https://www.xataka.com/basics/tarjeta-grafica-que-que-hay-dentro-como-funciona>.
- [33] Amda, «Tarjetas gráficas AMD Radeon™ RX,» 2022. [En línea]. Available: <https://www.amd.com/es/graphics/radeon-rx-graphics>.
- [34] T. Ding, X. Chen y H. Cao, «Principles of loop thermosyphon and its application in data center cooling systems: A review,» *Renewable and Sustainable Energy Reviews*, vol. 150, nº 1, pp. 1-16, 2021.
- [35] A. Shirmarz y A. Ghaffari, «Performance issues and solutions in SDN-based data center: a survey,» *The Journal of Supercomputing*, vol. 76, nº -, p. 7545–7593, 2020.
- [36] E. Masanet y A. Shehabi, «Recalibrating global data center energy-use estimates,» *SCIENCE*, vol. 367, nº 6481, pp. 984-986, 2020.
- [37] J. Zhang, R. Yu y T. Huang, «Load Balancing in Data Center Networks: A Survey,» *IEEE Communications Surveys & Tutorials*, vol. 20, nº 3, pp. 1-13, 2018.
- [38] K. Martijn y W. Fons, «Usage impact on data center electricity needs: A system dynamic forecasting model,» *Applied Energy*, vol. 291, nº -, pp. 1-19, 2021.
- [39] J. Calles-García y P. González-Pérez, *La Biblia del Footprinting*, 2011.
- [40] www.elhacker.net, «www.elhacker.net,» [En línea]. Available: https://www.elhacker.net/trucos_google.html.
- [41] S. Copertari, «Entrevista: VIRTUALIZACIÓN DE LA EDUCACIÓN EN TIEMPOS DE PANDEMIA,» *Revista Científica Educ@ção*, vol. 4, nº 7, pp. 891-895, 2020.
- [42] M. F. Candel, «Caracterización del sistema de memoria de una GPGPU.,» Universidad Politécnica de Valencia, Valencia , 2014.
- [43] S. Razzaq, «Scheduling algorithms for high-performance computing: An application perspective of fog computing,» *Recent Trends and Advances in Wireless and IoT-enabled Networks*, vol. 1, nº 1, pp. 107-117, 2019.

Anexos

PRUEBAFINAL

pruebafinal

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold

Processor: 4 x Intel Xeon Gold 6130 (7 Cores), Motherboard: Intel 440BX (6.00 BIOS), Chipset: Intel 440BX/ZX/DX, Memory: 16GB, Disk: 215GB Virtual disk, Graphics: VMware SVGA II 2GB, Network: VMware VMXNET3

OS: Ubuntu 20.04, Kernel: 5.4.0-26-generic (x86_64), Desktop: GNOME Shell 3.36.9, Display Server: X Server 1.20.8, Display Driver: NVIDIA, OpenGL: 4.5 Mesa 21.2.6 (LLVM 12.0.0 256 bits), Vulkan: 1.1.182, Compiler: GCC 9.4.0, File-System: ext4, Screen Resolution: 1225x656, System Layer: VMware

Kernel Notes: Transparent Huge Pages: madvise

Compiler Notes: --build=x86_64-linux-gnu --disable-vtable-verify --disable-werror --enable-checking=release --enable-clocale=gnu --enable-default-pie --enable-gnu-unique-object --enable-languages=c,ada,c++,go,brig,d,fortran,objc,obj-c++,gm2 --enable-lto --enable-libstdcxx-debug --enable-libstdcxx-time=yes --enable-multiarch --enable-multilib --enable-nls --enable-objc-gc=auto --enable-offload-targets=nvptx-none=/build/gcc-9-Av3uEd/gcc-9-9.4.0/debian/tmp-nvptx/usr,hsa --enable-plugin --enable-shared --enable-threads=posix --host=x86_64-linux-gnu --program-prefix=x86_64-linux-gnu- --target=x86_64-linux-gnu --with-abi=m64 --with-arch=32=i686 --with-default-libstdcxx-abi=new --with-gcc-major-version-only --with-multilib-list=m32,m64,mx32 --with-target-system-zlib=auto --with-tune=generic --without-cuda-driver -v

Disk Notes: MQ-DEADLINE / errors=remount-ro,relatime,rw / Block Size: 4096

Processor Notes: CPU Microcode: 0x2006a09

Graphics Notes: BAR1 / Visible vRAM Size: 256 MIB - vBIOS Version: 00.00.00.00

Python Notes: Python 3.8.10

Security Notes: itlb_multihit: KVM: Vulnerable + 11tf: Mitigation of PTE Inversion + mds: Mitigation of Clear buffers; SMT Host state unknown + meltdown: Mitigation of PTI + spec_store_bypass: Mitigation of SSB disabled via prctl and seccomp + spectre_v1: Mitigation of usercopy/swaps barriers and user pointer sanitization + spectre_v2: Mitigation of Full generic retpoline IBPB: conditional IBRS_FW STIBP: disabled RSB filling + tsx_async_abort: Not affected

IOR

IOR is a parallel I/O storage benchmark making use of MPI with a particular focus on HPC (High Performance Computing) systems. IOR is developed at the Lawrence Livermore National Laboratory (LLNL).

Block Size: 2MB - Disk Target: Default Test Directory

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

Block Size: 4MB - Disk Target: Default Test Directory

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

Block Size: 8MB - Disk Target: Default Test Directory

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

Block Size: 16MB - Disk Target: Default Test Directory

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

Block Size: 32MB - Disk Target: Default Test Directory

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

Block Size: 64MB - Disk Target: Default Test Directory

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

Block Size: 256MB - Disk Target: Default Test Directory

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

Block Size: 512MB - Disk Target: Default Test Directory

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

Block Size: 1024MB - Disk Target: Default Test Directory

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

SHOC Scalable Heterogeneous Computing

The CUDA and OpenCL version of Vetter's Scalable Heterogeneous Computing benchmark suite. SHOC provides a number of different benchmark programs for evaluating the performance and stability of compute devices.

Target: OpenCL - Benchmark: S3D

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test run did not produce a result. The test run did not produce a result. The test run did not produce a result.

Target: OpenCL - Benchmark: Triad

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test run did not produce a result. The test run did not produce a result. The test run did not produce a result.

Target: OpenCL - Benchmark: FFT SP

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test run did not produce a result. The test run did not produce a result. The test run did not produce a result.

Target: OpenCL - Benchmark: MD5 Hash

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test run did not produce a result. The test run did not produce a result. The test run did not produce a result.

Target: OpenCL - Benchmark: Reduction

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test run did not produce a result. The test run did not produce a result. The test run did not produce a result.

Target: OpenCL - Benchmark: GEMM SGEMM_N

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test run did not produce a result. The test run did not produce a result. The test run did not produce a result.

Target: OpenCL - Benchmark: Max SP Flops

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test run did not produce a result. The test run did not produce a result. The test run did not produce a result.

Target: OpenCL - Benchmark: Bus Speed Download

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test run did not produce a result. The test run did not produce a result. The test run did not produce a result.

Target: OpenCL - Benchmark: Bus Speed Readback

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test run did not produce a result. The test run did not produce a result. The test run did not produce a result.

Target: OpenCL - Benchmark: Texture Read Bandwidth

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test run did not produce a result. The test run did not produce a result. The test run did not produce a result.

High Performance Conjugate Gradient

HPCG is the High Performance Conjugate Gradient and is a new scientific benchmark from Sandia National Labs focused for super-computer testing with modern real-world workloads compared to HPCC.

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: cat: 'HPCG-Benchmark*.txt': No such file or directory

HPL Linpack

HPL is a well known portable Linpack implementation for distributed memory systems. This test profile is testing HPL upstream

directly, outside the scope of the HPC Challenge test profile also available through the Phoronix Test Suite (hpcc). The test profile attempts to generate an optimized HPL.dat input file based on the CPU/memory under test. The automated HPL.dat input generation is still being tuned and thus for now this test profile remains "experimental".

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

NAS Parallel Benchmarks

NPB, NAS Parallel Benchmarks, is a benchmark developed by NASA for high-end computer systems. This test profile currently uses the MPI version of NPB. This test profile offers selecting the different NPB tests/problems and varying problem sizes.

NAS Parallel Benchmarks 3.4

Test / Class: BT.C



Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold

SE +/- 6.58, N = 3



- (F9X) gfortran options: -O3 -march=native -pthread -lmpl_usempif08 -lmpl_mpihf -lmpl
- Open MPI 4.0.3

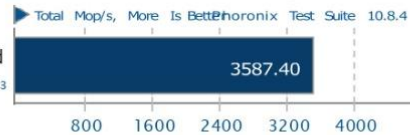
NAS Parallel Benchmarks 3.4

Test / Class: CG.C



Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold

SE +/- 21.96, N = 3



- (F9X) gfortran options: -O3 -march=native -pthread -lmpl_usempif08 -lmpl_mpihf -lmpl
- Open MPI 4.0.3

Test / Class: EP.C

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

Test / Class: EP.D

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. E: There are not enough slots available in the system to satisfy the 7

NAS Parallel Benchmarks 3.4

Test / Class: FT.C



Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold

SE +/- 29.17, N = 3



1. (F9X) gfortran options: -O3 -march=native -pthread -lmpi_usempif08 -lmpi_mpifh -lmpi
2. Open MPI 4.0.3

Test / Class: IS.D

Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold: The test quit with a non-zero exit status. The test quit with a non-zero exit status. The test quit with a non-zero exit status.

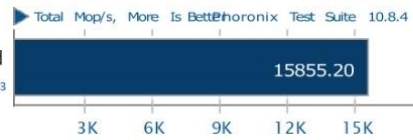
NAS Parallel Benchmarks 3.4

Test / Class: LU.C



Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold

SE +/- 43.21, N = 3



1. (F9X) gfortran options: -O3 -march=native -pthread -lmpi_usempif08 -lmpi_mpifh -lmpi
2. Open MPI 4.0.3

NAS Parallel Benchmarks 3.4

Test / Class: MG.C



Virtual disk - VMware SVGA II - 4 x Intel Xeon Gold

SE +/- 15.61, N = 3



1. (F9X) gfortran options: -O3 -march=native -pthread -lmpi_usempif08 -lmpi_mpifh -lmpi
2. Open MPI 4.0.3