

UNIVERSIDAD PÚBLICA DE EL ALTO  
**CARRERA INGENIERÍA DE SISTEMAS**



**TESIS DE GRADO**

**“MODELO PARA LA REDUCCIÓN DEL RIESGO DE PÉRDIDA DE DATOS EN  
SERVIDORES PROXMOX A TRAVÉS DE RAID 5”**

**Para Optar al Título de Licenciatura en Ingeniería de Sistemas  
MENCIÓN: INFORMÁTICA Y COMUNICACIONES**

Postulante	:	Univ. Nilton Amaru Mollisaca
Tutor Metodológico	:	M. Sc. Lic. Ing. Fanny Helen Pérez
	:	Mamani
Tutor Especialista	:	Lic. Fredy Alanoca Coareti
Tutor Revisor	:	Lic. Ing. Gonzalo Quiroga Fernández

**EL ALTO - BOLIVIA  
2024**

# DECLARACIÓN JURADA DE AUTENTICIDAD Y RESPONSABILIDAD

Yo, **Nilton Amaru Mollisaca** estudiante con **C.I. 13309803 LP** mediante la presente **declaro** de manera pública que la propuesta del **TESIS DE GRADO** titulada **“MODELO PARA LA REDUCCIÓN DEL RIESGO DE PÉRDIDA DE DATOS EN SERVIDORES PROXMOX A TRAVÉS DE RAID 5”** es original, siendo resultado de mi trabajo personal y no constituye una copia o replica de trabajos similares elaborados,

Autorizo la publicación del resumen de mi propuesta en internet y me comprometo a responder a todos los cuestionamientos que se desprenden de su lectura.

Asimismo, me hago responsable ante la universidad o terceros, de cualquiera irregularidad o daño que pudiera ocasionar, por el incumplimiento de lo declarado.

De identificarse falsificación, plagio, fraude, o que la **TESIS DE GRADO** haya sido publicado anteriormente; asumo las consecuencias y sanciones que de mi acción se deriven, responsabilizándome por todas las cargas legales que se deriven de ello sometiéndome a las normas establecidas y vigentes de la Carrera de Ingeniería de Sistemas de la Universidad Pública de El Alto.

El Alto, Junio de 2024

.....  
Nilton Amaru Mollisaca

C.I.: 13309803

E-mail: [namaru370@gmail.com](mailto:namaru370@gmail.com)

## **DEDICATORIA**

Dedico con todo mi corazón este trabajo a mis padres, pues sin ellos no lo habría logrado. El apoyo incondicional que me brindaron a lo largo de toda mi vida me protege y me guía por el camino del bien.

Por eso, ofrezco este trabajo como símbolo de agradecimiento por su infinita paciencia y amor.

**Nilton Amaru Mollisaca**

## **AGRADECIMIENTOS**

- Agradezco al Lic. Fredy Alanoca Coareti por despertar mi interés en la materia de servidores y por brindarme su apoyo invaluable durante el desarrollo de mi proyecto.
- Expreso mi profundo agradecimiento al Ing. Gonzalo Quiroga Fernández por su orientación experta y su ayuda en la redacción de mi proyecto.
- Quiero manifestar mi sincero agradecimiento a la Ing. Fanny Helen Pérez Mamani por su constante impulso y apoyo, que fueron fundamentales para culminar mi tesis. Desde el momento en que la conocí, me ha inspirado a crecer profesionalmente y a alcanzar nuevas metas.
- Quiero expresar mi más sincero agradecimiento a mi carrera de Ingeniería de Sistemas y a todo el plantel de docentes que me brindaron su apoyo durante los años de formación dentro de la carrera.
- Quiero brindar mis más sinceros agradecimientos a mi grupo de amigos "Los Lobos", quienes fueron parte fundamental de mi desarrollo dentro de la carrera.

**Nilton Amaru Mollisaca**

## INDICE

CAPITULO I.....	1
1.1. INTRODUCCION.....	2
1.2. ANTECEDENTES.....	2
1.2.1. Antecedentes a fines a la investigación .....	2
1.3. PLANTEAMIENTO DEL PROBLEMA.....	7
1.3.1. Problema principal .....	7
1.3.2. Problemas secundarios .....	8
1.3.3. Formulación del problema .....	8
1.4. OBJETIVOS.....	8
1.4.1. Objetivo general.....	8
1.4.2. Objetivos específicos .....	8
1.5. HIPOTESIS.....	9
1.5.1. Identificación de variables.....	9
1.5.2. Operacionalización de variables .....	9
1.5.3. Conceptualización de variables .....	10
1.6. JUSTIFICACIÓN.....	10
1.6.1. Justificación Científica .....	10
1.6.2. Justificación Técnica.....	11
1.6.3. Justificación Económica.....	12
1.6.4. Justificación Social.....	12
1.7. METODOLOGIA.....	12
1.7.1. Método científico.....	13
1.7.2. Método de ingeniería .....	13
1.7.3. Métricas de calidad.....	15

1.7.4.	Pruebas .....	15
1.8.	HERRAMIENTAS .....	16
1.8.1.	Hardware .....	16
1.8.2.	Software.....	16
1.9.	LIMITES Y ALCANCES .....	17
1.9.1.	Limites .....	17
1.9.2.	Alcances .....	17
1.10.	APORTES.....	18
1.10.1	Aporte Teórico .....	18
1.10.2	Aporte Practico .....	18
CAPITULO II.....		19
2.1.	MODELO .....	20
2.2.	DEFINICION DE MODELO PARA SERVIDORES .....	20
2.2.1.	Modelos de Centro de Datos .....	20
2.3.	RAID (Redundant Array of Independent Disks) .....	22
2.3.1.	Definición de RAID.....	22
2.3.2.	Tipos de RAID.....	22
2.3.3.	Niveles de RAID.....	25
2.3.4.	Importancia del RAID en la Virtualización.....	32
2.3.5.	Consideraciones de Implementación de RAID en Proxmox VE...	33
2.4.	PROXMOX VIRTUAL ENVIROMENT .....	33
2.4.1.	¿Para qué se usa Proxmox? .....	34
2.4.2.	Características de Proxmox .....	34
2.4.3.	Uso de Proxmox VE.....	35
2.5.	KVM (Kernel-Based Virtual Machine) .....	36

2.5.1.	Componentes de KVM.....	37
2.5.2.	Gestión de KVM.....	38
2.5.3.	Funcionamiento .....	39
2.6.	MICROSOFT AZURE .....	39
2.6.1.	¿Qué es Microsoft Azure? .....	39
2.6.2.	Oferta en la nube .....	40
2.7.	AMAZON WEB SERVICES .....	44
2.7.1.	Ventajas de AWS (Amazon Web Services) .....	44
2.7.2.	Arquitectura y Componentes de AWS .....	45
2.8.	COMPARATIVA MICROSOFT AZURE VS. AMAZON WEB SERVICES	45
2.8.1.	Analítica .....	46
2.8.2.	Integración .....	48
2.8.3.	Servicios Computacionales.....	49
2.9.	VIRTUALBOX.....	50
2.9.1.	Utilidad de la Virtualización con Virtualbox .....	50
2.9.2.	Terminología Especial .....	51
2.10.	RAID 5 .....	52
2.10.1.	Estructura de RAID 5.....	53
2.10.2.	¿Cómo funciona el RAID 5? .....	54
2.10.3.	Paridad .....	55
2.10.4.	Striping.....	55
2.10.5.	Modelo de confiabilidad simplificado de discos RAID 5.....	57
CAPITULO III.....		60
3.1.	TIPO DE INVESTIGACIÓN .....	61

3.2.	NIVEL DE INVESTIGACIÓN .....	61
3.1.1.	Aplicativo .....	61
3.1.2.	Experimental.....	61
3.3.	DISEÑO DE INVESTIGACION .....	62
3.3.1.	Variables de la Investigación .....	62
3.3.2.	Ambiente De Investigación .....	62
3.3.3.	DESCRIPCIÓN DE LA METODOLOGÍA A USAR.....	62
3.3.4.	Tratamiento De Datos.....	63
3.4.	HERRAMIENTAS A USAR.....	63
3.4.1.	Técnicas de Investigación e Instrumentos .....	63
3.5.	MÉTRICAS DE CALIDAD.....	63
CAPITULO IV .....		65
4.1.	PRESENTACION DEL MODELO .....	66
4.2.	DESARROLLO DEL MODELO .....	67
4.3.	PRUEBAS.....	84
4.3.1.	Prueba de Caja Blanca .....	84
4.3.2.	Pruebas de Caja Negra .....	88
4.4.	EVALUACIÓN Y PROCESAMIENTO DE DATOS.....	90
4.5.	INTERPRETACION DE RESULTADOS .....	99
CAPITULO V .....		103
5.1.	CONCLUSIONES .....	104
5.2.	RECOMENDACIONES.....	105
BIBLIOGRAFIA.....		107
ANEXOS.....		110
1.	Diagrama de la situación actual.....	111



2.	Diagrama de Ishikawa .....	112
3.	Árbol de Objetivos .....	112
4.	Diagrama RAID.....	113
5.	Recolección de Datos de Lectura y Escritura .....	113
6.	Datos recopilados de recuperación de disco .....	117

## INDICE DE FIGURAS

<b>Figura 1</b>	<i>BIOS Intel Matrix RAID.</i> .....	23
<b>Figura 2</b>	<i>Controlador RAID NVMe HighPoint SSD7101A-1: SSD M.2</i> .....	24
<b>Figura 3</b>	<i>Asignación de datos en RAID 0</i> .....	26
<b>Figura 4</b>	<i>Arquitectura de la matriz RAID 1</i> .....	27
<b>Figura 5</b>	<i>Arquitectura de la matriz RAID 4</i> .....	28
<b>Figura 6</b>	<i>Arquitectura de la matriz RAID 5</i> .....	29
<b>Figura 7</b>	<i>Arquitectura de la matriz RAID 6.</i> .....	30
<b>Figura 8</b>	<i>Arquitectura de la matriz RAID 10.</i> .....	31
<b>Figura 9</b>	<i>Proxmox VE</i> .....	34
<b>Figura 10</b>	<i>Interfaz Gráfica KVM</i> .....	37
<b>Figura 11</b>	<i>Arquitectura Microsoft Azure</i> .....	40
<b>Figura 12</b>	<i>Principales usos de una SaaS.</i> .....	41
<b>Figura 13</b>	<i>Principales usos de una PaaS.</i> .....	42
<b>Figura 14</b>	<i>Principales usos de una IaaS.</i> .....	42
<b>Figura 15</b>	<i>diferencia entre IaaS, PaaS, SaaS.</i> .....	43
<b>Figura 16</b>	<i>Distribución de bloque de datos.</i> .....	54
<b>Figura 17</b>	<i>Paridad de datos</i> .....	55
<b>Figura 18</b>	<i>Striping</i> .....	56
<b>Figura 19</b>	<i>Gráfico de estados</i> .....	58
<b>Figura 20</b>	<i>Modelo para la reducción del riesgo de pérdida de datos</i> .....	66
<b>Figura 21</b>	<i>Configuración básica de máquina virtual</i> .....	68
<b>Figura 22</b>	<i>Conexión SSH</i> .....	70
<b>Figura 23</b>	<i>Terminal Bitvise</i> .....	71
<b>Figura 24</b>	<i>Listado de discos</i> .....	71
<b>Figura 25</b>	<i>Uso de “fdisk”</i> .....	72
<b>Figura 26</b>	<i>Listado de opciones de “fdisk”</i> .....	73
<b>Figura 27</b>	<i>Añadir nueva partición</i> .....	74
<b>Figura 28</b>	<i>Opción “t (Tipo de array)”</i> .....	74
<b>Figura 29</b>	<i>Listado de tipos de array</i> .....	75
<b>Figura 30</b>	<i>Selección de tipo “Linux raid autodetectable”</i> .....	75

<b>Figura 31</b>	<i>Selección de opción “w”</i> .....	76
<b>Figura 32</b>	<i>Verificación del particionado de los discos</i> .....	77
<b>Figura 33</b>	<i>Creación de RAID 5</i> .....	78
<b>Figura 34</b>	<i>Formateo de sistema de archivo en RAID 5</i> .....	78
<b>Figura 35</b>	<i>Montaje de RAID 5</i> .....	79
<b>Figura 36</b>	<i>Montaje de RAID 5 en “fstab”</i> .....	80
<b>Figura 37</b>	<i>Actualización de nombre host</i> .....	81
<b>Figura 38</b>	<i>Adición de repositorios</i> .....	82
<b>Figura 39</b>	<i>Instalación de Proxmox</i> .....	83
<b>Figura 40</b>	<i>Instalación de Postfix</i> .....	83
<b>Figura 41</b>	<i>Servidor Proxmox en virtualbox</i> .....	84
<b>Figura 42</b>	<i>Estructura de los discos y las particiones del arreglo RAID 5</i> .....	85
<b>Figura 43</b>	<i>Estado del arreglo RAID 5 antes de fallo de disco</i> .....	86
<b>Figura 44</b>	<i>Estado del arreglo RAID 5 después de fallo de disco</i> .....	87
<b>Figura 45</b>	<i>Reconstrucción Exitosa del Disco de Reserva</i> .....	89
<b>Figura 46</b>	<i>Resultados de Prueba de Escritura y Lectura en RAID5</i> .....	90
<b>Figura 47</b>	<i>Análisis Estadístico de las Velocidades de Lectura</i> .....	94
<b>Figura 48</b>	<i>Análisis Estadístico de las Velocidades de Escritura</i> .....	95
<b>Figura 49</b>	<i>Resultado análisis estadístico</i> .....	98
<b>Figura 50</b>	<i>Grafica análisis estadístico</i> .....	99
<b>Figura 51</b>	<i>Resultado de prueba de hipótesis estadística</i> .....	102

## INDICE DE TABLAS

Tabla 1 <i>Operacionalización de Variables</i> .....	9
Tabla 2 <i>Tabla analítica de comparación.</i> .....	46
Tabla 3 <i>Tabla de integracion</i> .....	48
Tabla 4 <i>Servicios Computacionales</i> .....	49

## RESUMEN

La gestión eficiente de datos es crucial en la era digital. Este estudio se centra en la implementación de sistemas RAID en servidores Proxmox para mitigar el riesgo de pérdida de datos. Se exploran las configuraciones RAID 5, con el objetivo de identificar la más eficiente. Se abordan los riesgos asociados a cada configuración, proporcionando una visión integral de seguridad.

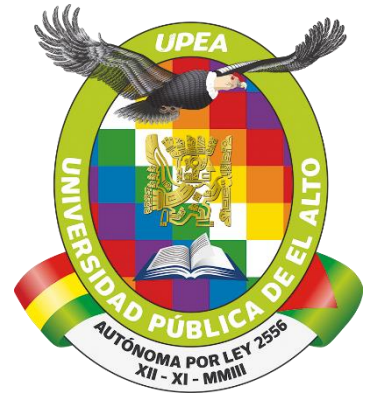
La metodología empleada se basa en la investigación experimental y el método científico, utilizando métricas reconocidas internacionalmente para evaluar la calidad del sistema RAID implementado para ello se utilizan herramientas como Proxmox VE para la administración de servidores y VirtualBox para la virtualización, como entorno central donde se desplegará el servidor Proxmox y se realizarán las pruebas. El estudio contribuye teórica y prácticamente al proporcionar fundamentos para implementar configuraciones RAID 5 de manera efectiva, abordando la prevención de pérdida de datos en servidores Proxmox.

**Palabras Clave:** RAID 5, Proxmox, Perdida de datos.

## **ABSTRACT**

The efficient management of data is crucial in the digital era. This study focuses on implementing RAID systems on Proxmox servers to mitigate the risk of data loss. RAID 5 configurations are explored with the aim of identifying the most efficient one. The associated risks of each configuration are addressed, providing a comprehensive security outlook.

The methodology employed is based on experimental research and the scientific method, utilizing internationally recognized metrics to evaluate the quality of the implemented RAID system. Tools such as Proxmox VE for server administration and VirtualBox for virtualization are used, with Microsoft Azure serving as the central environment where the Proxmox server will be deployed and tested. The study contributes both theoretically and practically by providing the foundation for effectively implementing RAID 5 configurations, addressing data loss prevention on Proxmox servers.



# CAPITULO I

## MARCO

### PRELIMINAR



## **1.1. INTRODUCCION**

La gestión eficiente de datos es crucial en la era digital actual. La pérdida de información puede tener consecuencias graves. Este estudio se centra en la implementación de sistemas Redundant Array of Independent Disks (RAID) en servidores Proxmox como estrategia para mitigar el riesgo de pérdida de datos.

La combinación del entorno virtualizado de Proxmox y la tecnología RAID ofrece una solución integral para garantizar la integridad y disponibilidad de los datos. Se explorarán y compararán configuraciones RAID, con el objetivo de identificar la más eficiente.

El estudio abordará riesgos asociados a cada configuración, proporcionando una visión integral de seguridad. La importancia radica en contribuir al conocimiento teórico y práctico en áreas como gestión de riesgos, seguridad informática y almacenamiento de datos.

La metodología empleada se basa en la investigación experimental y el método científico, utilizando métricas reconocidas internacionalmente, como las establecidas por la norma ISO/IEC 25000, para evaluar la calidad del sistema RAID implementado. Además, se aplicarán pruebas de caja blanca, caja negra y de estrés.

Para llevar a cabo este estudio, se utilizarán herramientas como Proxmox VE para la administración de servidores y VirtualBox para la virtualización, con Microsoft Azure como el entorno central donde se desplegará el servidor Proxmox y se realizarán las pruebas. Este enfoque no solo busca optimizar la gestión de datos, sino también fortalecer la seguridad de la información en entornos críticos.

## **1.2. ANTECEDENTES**

### **1.2.1. Antecedentes a fines a la investigación**

Como punto de referencia fundamental, se ha llevado a cabo un análisis detenido de antecedentes a nivel nacional e internacional. A continuación, se detallan los hallazgos, sirviendo como referente clave para la presente investigación.



### **1.2.1.1. Antecedentes Internacionales**

- Estrategia de superación para la utilización de Proxmox y Pfsense en las instituciones de salud (Arias et al., 2019).

Proxmox Virtual Environment es un entorno de virtualización de código abierto para servidores, mientras que pfSense es un potente cortafuego, enrutador, que permite brindar seguridad a la información, aplicar reglas e instalar servicios de red. En la provincia existen insuficiencias en la implementación de estos sistemas por parte de los administradores de red e informáticos. Además, no se explotan herramientas de software libre en servidores. En consecuencia, esta investigación tiene como objetivo diseñar una estrategia de superación para la adquisición de conocimientos teórico-prácticos sobre Proxmox y pfSense, que permita la estandarización de los servidores y servicios en las unidades de salud del territorio en función del acceso a internet por parte de los profesionales. Se realizó un estudio exploratorio en 43 informáticos y administradores de redes, para lo cual se utilizaron métodos teóricos, empíricos y estadísticos que permitieron detectar las problemáticas existentes en el territorio. La valoración aportada por los participantes sobre el taller parte de la estrategia reflejó que 98,0% lo consideraron excelente. El 93,0% expresa haber adquirido conocimientos teórico-prácticos sobre Proxmox y pfSense. La estrategia de superación permitió la implementación de los sistemas Proxmox y pfSense en las unidades de salud del territorio y a su vez creo una comunidad de especialistas en estos sistemas.

- Implementación de una infraestructura de IT virtual para el data center de la facultad de ingeniería en ciencias aplicadas, en la universidad técnica del norte (Gladys Margarita, 2019).

El presente proyecto describe el proceso de implementación de una infraestructura virtual de IT para el Data Center de la Facultad de Ingeniería en Ciencias Aplicadas (FICA), en la Universidad Técnica del Norte, utilizando las bases que definen a la Hiperconvergencia los cuales son redes de comunicación, cómputo y almacenamiento. En el capítulo tres se analiza la

situación actual donde muestra una infraestructura funcional, pero apegada a parámetros de arquitectura heredada, la misma que se ve limitada a procesos de escalabilidad, funcionalidad y actualización de servicios y servidores, generando gastos innecesarios de mantenimiento técnico para cada equipo que trabaja de forma independiente, cada uno con su responsable directo. En el capítulo cuatro en el diseño se procede a diseñar el Clúster y el almacenamiento para después implementar el entorno virtual con la plataforma Proxmox VE, la que consta de dos características que es la alta disponibilidad y la migración en caliente, también del almacenamiento compartido. Se utilizó el entorno virtual basado en software libre Proxmox VE, para virtualizar y centralizar los servidores, tomando en cuenta los recursos con los que cuenta el Data Center y las características de los equipos que lo integran. En el capítulo cinco constan las pruebas realizadas tanto del funcionamiento de la plataforma, así como la alta disponibilidad y la migración en caliente, dan como resultado la optimización de los recursos y la eficiencia de la plataforma. Además, que el tiempo de migración es menor cuando un nodo cae, que, en el caso de una migración en caliente, los dos tiempos programados y no programados son mínimos, dando así una infraestructura eficiente y eficaz.

- Implementación de un sistema de virtualización centralizado basado en Proxmox para optimizar la gestión de recursos informáticos del data center del gobierno regional pasco (Senin, 2023).

El trabajo de investigación indaga acerca de la implementación del sistema de virtualización centralizado basado en PROXMOX con los cual se busca optimizar la gestión de los recursos informáticos del Data Center del Gobierno Regional Pasco; con ello se plantea el objetivo principal de investigación el mismo que es: “Implementar un sistema de virtualización centralizado basado en PROXMOX para optimizar la gestión de recursos informáticos del Data Center del Gobierno Regional Pasco, 2022” para hacer posible el desarrollo de la presente se hizo uso del tipo de investigación aplicada – tecnológica, los métodos inductivo – deductivo, el diseño de investigación definido es el cuasi experimental con las pruebas pre y post test. Por último, la

investigación arrojó los siguientes resultados: mejora del indicador velocidad máxima de procesamiento de escritura pasando de una velocidad promedio de 531.1Kb/s a 1785.6Kb/s para la velocidad máxima de procesamiento de escritura se pasó del promedio de 1140.1Kb/s a 2275.3Kb/s por el lado del indicador tráfico de red de carga se pasó de un promedio de 3.4Mb/s a 8.7Mb/s y en descarga de un promedio de 9.7Mb/s a 33.1Mb/s, con lo cual se puede concluir que se optimiza la gestión de los recursos informáticos del Data Center del Gobierno Regional Pasco a partir de la implementación del sistema de virtualización centralizado basado en PROXMOX.

### **1.2.1.2. Antecedentes Nacionales**

A continuación, se describe los antecedentes nacionales consultados:

- Guía de gestión de riesgos de los servicios SAAS en la nube privada para la empresa full assistance (Morales, 2019).

La presente investigación responde al problema ¿Cómo aplicar la Gestión de Riesgos de los servicios SaaS para mejorar la seguridad de la información de los servicios SaaS en la Nube Privada de la empresa FULL ASSISTANCE con el objetivo general: Proponer una Guía de Gestión de Riesgos de los servicios SaaS utilizando normas ISO 27005-31000 para mejorar la seguridad de la información de los Servicios SaaS en la Nube Privada del contexto donde se investiga.

- Sistema de atención de clientes para restaurantes basado en Cloud Computing (Andrade, 2020).

Mejorar la atención de los pedidos realizados por los clientes automatizando la toma de órdenes (pedidos) a través del desarrollo de un Sistema basado en Cloud Computing, proveyendo una solución que considere los distintos aspectos recurrentes en los restaurantes de nuestro medio; haciéndolo más accesible para diferentes negocios de comida.

Sistema basado en Cloud Computing que brinda al cliente y al dueño del negocio de comida una mayor facilidad a la hora de realizar pedidos de comida y atender dichos pedidos. El espacio que se busca mejorar es el que ofrecen los restaurantes a sus clientes presentes físicamente en sus establecimientos.

### **1.2.1.3. Antecedentes Locales**

- Uso de un modelo basado en la infraestructura tecnológica de Cloud Computing para la disminución de costos, mantenimiento y tiempo de los equipos de la empresa grupo industrial SAXSAY S.R.L. (Huacara Choque Gabriel, 2022).

La presente tesis de grado con el tema denominado el uso un modelo basado en la Infraestructura Tecnológica de Cloud Computing para la disminución de costos, tiempo y mantenimiento de los equipos de la empresa Grupo Industrial SAXSAY S.R.L., nos permitirá un alcance tecnológico que consiste en proveer y gestionar recursos de computación, servidores, almacenamiento, redes y virtualización para resolver los problemas también mejorar el trabajo y las necesidades de la empresa para un funcionamiento adecuado, para no incurrir en pérdidas económicas o gastos insulsos.

El objetivo principal de la presente tesis es usar un modelo basado en la Infraestructura Tecnológica de Cloud Computing para la disminución de costos, tiempo y mantenimiento de los equipos de la empresa Grupo Industrial SAXSAY S.R.L., la hipótesis planteada nos ayudara a resolver el problema planteado.

El tipo de investigación del presente estudio de acuerdo al problema planteado, y los objetivos se utilizará el enfoque cuantitativo no experimental de un nivel de estudio descriptivo para trabajar directamente con las variables del presente estudio con las encuestas realizadas entre post test y pre test.

Los resultados obtenidos en el presente estudio analizando el modelo presentado basado en la tecnología de Cloud Computing se pudo evidenciar que se puede crear nube privada utilizando VMware como una infraestructura como servicio (IAAS).

- Modelo de seguridad para la infraestructura de información (Chamizo Gutierrez, 2020).

El avance de la tecnología es demasiado rápido ya que es poco habitual seguir el ritmo de la tecnología porque avanza a pasos agigantados y es un gran beneficio para toda la humanidad con este beneficio viene consecuencias negativas para nuestros datos personales y/o empresariales ya que podemos

estar expuestos a robo de información personal como también robo de información de las instituciones, como también sufrir variedad de ataques, suplantación de servicios, entre otros por este motivo es fundamental tener una aplicación que busque las vulnerabilidades que existiera en su servidor así tratar de evitar futuros ataques que pueden pasar. Para esto utilizaremos distintas herramientas para la búsqueda de las vulnerabilidades que pudiera existir en un servidor.

### **1.3. PLANTEAMIENTO DEL PROBLEMA**

La pérdida de datos es un desafío crítico que enfrentan todas las empresas y organizaciones. Puede ocurrir debido a diversas razones, como fallos de hardware y software, errores humanos y factores ambientales. Las consecuencias incluyen la pérdida de productividad, daños a la imagen institucional y posibles sanciones legales.

Es crucial abordar este problema proactivamente para proteger la información y minimizar el impacto en las entidades que usan servidores virtualizados.

#### **1.3.1. Problema principal**

La pérdida de datos es un problema serio que puede tener un impacto significativo en las empresas. Cuando se pierden datos, la empresa puede sufrir pérdidas financieras, pérdida de productividad y daños a la imagen institucional.

La pérdida de datos puede ocurrir por una serie de razones, como ser: fallos de hardware y software, errores humanos y factores ambientales siendo que el 67% de las pérdidas de datos se deben a fallos del disco duro o del sistema, el 14% de las pérdidas de datos se deben a errores humanos y el 10% de las pérdidas de datos se deben a fallos del software.

Los servidores virtualizados son esenciales para la gestión eficiente de datos en las empresas al permitir la creación de máquinas virtuales en un solo servidor físico. Cada máquina virtual opera como un servidor independiente, lo que facilita la organización y protección de datos. La virtualización ofrece ventajas como la consolidación de servidores, la resiliencia a través de la migración y conmutación por error, así como la implementación eficiente de copias de seguridad. Este enfoque también garantiza el aislamiento de datos entre las máquinas virtuales, fortaleciendo la seguridad y la integridad de la información institucional.

### **1.3.2. Problemas secundarios**

- Pérdida de productividad: Cuando se pierden datos, los empleados pueden perder tiempo y productividad tratando de restaurar los datos o encontrar alternativas. Esto puede afectar a la capacidad de la empresa para cumplir con los plazos y las demandas de los clientes.
- Daños a la imagen institucional: La pérdida de datos puede dañar la imagen institucional de la empresa si se hace pública. Esto puede hacer que los clientes pierdan la confianza en la empresa y se lleven su negocio a la competencia.
- Sanciones legales: En algunos casos, la pérdida de datos puede dar lugar a sanciones legales. Por ejemplo, las empresas que almacenan datos de clientes pueden ser multadas por la violación de la privacidad si se pierden esos datos.

### **1.3.3. Formulación del problema**

¿Cómo se puede reducir el riesgo de pérdida de datos en servidores Proxmox para evitar la pérdida de productividad, daños a la imagen institucional y sanciones legales?

## **1.4. OBJETIVOS**

### **1.4.1. Objetivo general**

Analizar estrategias para reducir el riesgo de pérdida de datos en servidores Proxmox, previniendo así la pérdida de productividad, resguardando la imagen institucional y evitar posibles sanciones legales asociadas con la pérdida de datos.

### **1.4.2. Objetivos específicos**

- Evaluar diversas configuraciones de sistemas de almacenamiento en relación con su capacidad efectiva para minimizar la pérdida de datos y determinar cuál de estas opciones es la más eficiente para la prevención de pérdida de datos.
- Analizar los riesgos asociados con diferentes configuraciones de sistemas de almacenamiento, con el propósito de comprender de manera integral las implicaciones y consideraciones clave para su posible implementación.
- Diseñar el modelo optimizado para la reducción del riesgo de pérdida de datos en servidores Proxmox.
- Evaluar el impacto del uso de distintas configuraciones de sistemas de almacenamiento en el rendimiento general del sistema, considerando tanto el

rendimiento de lectura/escritura como la latencia, para asegurar que estas configuraciones no comprometan la eficiencia operativa de los servidores Proxmox.

### 1.5. HIPOTESIS

**H<sub>i</sub>:** La configuración de RAID 5 en servidores Proxmox reduce significativamente el riesgo de pérdida de datos, contribuyendo así a evitar la pérdida de productividad, daños a la imagen institucional y sanciones legales.

**H<sub>0</sub>:** La configuración de RAID 5 en servidores Proxmox no reduce significativamente el riesgo de pérdida de datos, contribuyendo así a evitar la pérdida de productividad, daños a la imagen institucional y sanciones legales..

#### 1.5.1. Identificación de variables

Variable Independiente

- ✓ Implementación de RAID 5 en servidor Proxmox

Variable Dependiente

- ✓ Riesgo de pérdida de datos

#### 1.5.2. Operacionalización de variables

**Tabla 1**

*Operacionalización de Variables*

VARIABLE	TIPO		INDICADORES
	DE VARIABLE	DIMENSION	
<b>Configuración de RAID 5 en servidor Proxmox</b>	<b>VARIABLE INDEPENDIENTE</b>	Cantidad de discos duros	Cantidad de discos empleados
		Tipo de configuración RAID	Nivel de redundancia

<b>Riesgo de pérdida de datos</b>	<b>VARIABLE DEPENDIENTE</b>	Probabilidad de pérdida de datos	Porcentaje de probabilidad de que se pierdan datos
		Impacto de pérdida de datos	Valor monetario que tendría la pérdida de datos en la organización

*Nota Operacionalización de Variables.*

*Fuente: Elaboración propia (2023).*

### **1.5.3. Conceptualización de variables**

Configuración de un sistema RAID 5 en servidores Proxmox: Es una medida proactiva para reducir el riesgo de pérdida de datos. Los sistemas RAID brindan redundancia de datos, lo que puede ayudar a proteger los datos de las fallas de hardware. Además, los sistemas RAID 5 pueden mejorar el rendimiento y la capacidad de almacenamiento. Sin embargo, los sistemas RAID 5 también pueden ser costosos y complejos de implementar y administrar.

Riesgo de pérdida de datos: Es la probabilidad de que se pierdan datos debido a una falla de hardware, software, humana o ambiental. Los factores que pueden aumentar la probabilidad de pérdida de datos incluyen:

- Fallas de hardware: Los discos duros, las unidades de cinta y otros dispositivos de almacenamiento pueden fallar debido a desgaste, daño físico o errores de fabricación.
- Fallas de software: Los errores de software, como los errores de programación o los ataques de malware, pueden causar la pérdida de datos.
- Errores humanos: Los errores humanos, como la eliminación accidental de datos o la corrupción de archivos, pueden causar la pérdida de datos.
- Factores ambientales: Los factores ambientales, como los desastres naturales o los incendios, pueden causar la pérdida de datos.

## **1.6. JUSTIFICACIÓN**

### **1.6.1. Justificación Científica**

La configuración de sistemas RAID 5 en servidores Proxmox resuelve un problema crítico que afecta a empresas y organizaciones. Los fallos de hardware y



software, errores humanos y factores ambientales representan amenazas significativas que requieren soluciones técnicas sólidas. A través de un estudio científico, se examinará cómo diferentes configuraciones RAID 5 pueden proporcionar mayor redundancia y protección de datos en servidores Proxmox frente a estos desafíos.

Este estudio proporcionará un marco sólido y práctico para la gestión de datos en entornos que utilizan Proxmox. Su aplicabilidad no se limitará únicamente a la empresa en cuestión, sino que también servirá como un recurso valioso para otras organizaciones y entornos de aprendizaje que buscan abordar problemas similares.

### **1.6.2. Justificación Técnica**

La configuración de sistemas RAID 5 en servidores Proxmox representa una estrategia altamente eficaz para abordar el problema de la pérdida de datos. Los sistemas RAID 5 ofrecen redundancia y la capacidad de recuperar datos en caso de fallos de hardware, lo que garantiza la continuidad de las operaciones institucionales.

Con el propósito de evaluar y validar esta estrategia, se llevarán a cabo pruebas en entornos virtuales donde se emulará un servidor Proxmox. Esto permitirá la realización de pruebas aisladas y un análisis detallado de su desempeño. A continuación, se describen los requisitos básicos para el uso de Proxmox:

- Intel EMT64 o AMD64 con indicador de CPU Intel VT/AMD-V.
- Memoria, mínimo 2 GB para el sistema operativo y los servicios Proxmox VE. Además de memoria designada para los invitados. Para Ceph o ZFS se requiere memoria adicional, aproximadamente 1 GB de memoria por cada TB de almacenamiento utilizado.
- Almacenamiento rápido y redundante, mejores resultados con discos SSD.
- También se admiten NIC de Gbit redundantes, NIC adicionales según la tecnología de almacenamiento preferida y la configuración del clúster: 10 Gbit y superior.
- Para el paso a través de PCI(e) se necesita una CPU con indicador de CPU VT-d/AMD-d.

La implementación de sistemas RAID 5 en servidores Proxmox implica la necesidad de recursos tecnológicos significativos, como procesadores avanzados,

discos SSD y NIC de alta velocidad. Estos componentes, aunque fundamentales para garantizar un rendimiento óptimo, también representan una inversión considerable. Es esencial reconocer que esta inversión puede ser un obstáculo financiero para empresas de menor envergadura, que podrían enfrentar dificultades para adquirir equipos de última generación debido a restricciones presupuestarias.

En contraste, la opción de utilizar entornos virtuales en la nube, como Microsoft Azure, emerge como una solución más accesible, eliminando la necesidad de inversión en hardware físico costoso y brindando a un espectro más amplio de organizaciones la posibilidad de participar en iniciativas de este tipo.

### **1.6.3. Justificación Económica**

La implementación de sistemas RAID 5 en servidores Proxmox se basa en una evaluación de costos y beneficios. Aunque la inversión inicial pueda ser significativa, esta solución aportará un valor a largo plazo.

La pérdida de datos puede representar costosas interrupciones empresariales, daños a la imagen institucional y sanciones legales lo cual llegaría a representar grandes pérdidas económicas para la empresa.

### **1.6.4. Justificación Social**

Proteger los datos de los clientes y socios comerciales es esencial para mantener la confianza y la imagen institucional de la empresa en la sociedad. La pérdida de datos puede resultar en la insatisfacción de los clientes y en casos extremos, en sanciones legales.

Al tomar medidas proactivas para evitar la pérdida de datos, la empresa demuestra su compromiso y la responsabilidad social empresarial lo cual fortalecerá los lazos comerciales con los clientes.

## **1.7. METODOLOGIA**

La metodología empleada en este estudio se fundamenta en la investigación experimental, la cual se define como un proceso lógico, metódico y organizado de pasos secuenciales destinados a llevar a cabo investigaciones científicas. Este enfoque implica la rigurosa manipulación controlada de variables experimentales previamente no confirmadas o condicionadas. Su finalidad radica en analizar y describir el comportamiento de un fenómeno o problema específico en su contexto.

### **1.7.1. Método científico**

La base científica del presente documento, está asentada sobre un modelo general que Muñoz Razo (2011), propone para optimizar el desarrollo de la investigación.

Este modelo comprende, en el presente caso de tesis, desde la elección del tema hasta la presentación del borrador inicial de la misma. Otros elementos sólo se mencionan para complemento del propio modelo.

- Fase I Datos iniciales.
- Fase II Muestreo.
- Fase III Instrumentos de la investigación.
- Fase IV Procesamiento de datos.
- Fase V Métodos de análisis.
- Fase VI Análisis de datos.
- Fase VII Informe final

### **1.7.2. Método de ingeniería**

El método de ingeniería podría ser el "Método de Diseño Iterativo". Este enfoque implica el desarrollo y refinamiento continuo de un diseño a lo largo de múltiples iteraciones en la cual nos planteamos las siguientes etapas:

- Definición de Objetivos:  
Establece claramente los objetivos de la implementación de distintos tipos de RAID en servidores Proxmox. Define los criterios que utilizarás para evaluar las ventajas y desventajas.
- Análisis de Requisitos:  
Identifica los requisitos específicos para cada tipo de RAID que planeas implementar. Considera el rendimiento, la redundancia, la capacidad y otros factores relevantes.
- Diseño de Configuración RAID:  
Diseñar la configuración RAID para cada tipo que planeas estudiar. Asegúrate de que las diferencias clave entre ellos sean destacadas en el diseño.
- Implementación de Configuración RAID:

Lleva a cabo la implementación de las distintas configuraciones RAID en un entorno de prueba utilizando servidores Proxmox. Sigue un enfoque sistemático para cada tipo.

- Pruebas de Funcionalidad y Rendimiento:

Realiza pruebas detalladas de funcionalidad y rendimiento para cada configuración RAID. Evalúa aspectos como velocidad de lectura/escritura, capacidad de recuperación y eficiencia.

- Monitoreo Continuo:

Establece un sistema de monitoreo continuo para todas las configuraciones RAID. Observa y registra cualquier variación en el rendimiento a lo largo del tiempo.

- Documentación Detallada:

Documenta exhaustivamente cada paso del proceso de implementación y las pruebas realizadas. Incluye detalles específicos sobre la configuración de cada tipo de RAID.

- Comparación y Evaluación:

Compara los resultados de las pruebas entre las distintas configuraciones RAID. Evalúa las ventajas y desventajas de cada tipo en relación con tus objetivos.

- Optimización y Ajuste:

Realiza ajustes en las configuraciones según sea necesario para optimizar el rendimiento. Considera cambios en la configuración en función de los resultados obtenidos.

- Análisis de Costos y Recursos:

Realiza un análisis de los costos asociados con cada configuración RAID, tanto en términos de hardware como de tiempo y recursos humanos necesarios.

- Elaboración de Conclusiones:

Saca conclusiones basadas en los resultados de las pruebas y el análisis. Recomienda la configuración RAID más adecuada según tus criterios y objetivos.

- Presentación de Resultados:

Presenta los resultados de tu estudio de manera clara y concisa. Incluye gráficos, tablas y cualquier información relevante que respalde tus conclusiones.

### **1.7.3. Métricas de calidad**

#### **1.7.3.1. ISO/IEC 25000**

La ISO/IEC 25000 es una norma internacional que establece un marco para la evaluación de la calidad del software. Se centra en características clave, como funcionalidad, confiabilidad, usabilidad, eficiencia, mantenibilidad y portabilidad. Esta norma proporciona modelos y requisitos para medir y mejorar la calidad del software a lo largo de su ciclo de vida. Al utilizar la ISO/IEC 25000, se pueden establecer estándares claros para evaluar la confiabilidad y eficiencia de los sistemas informáticos, siendo especialmente relevante para proyectos que buscan garantizar la integridad de los datos y optimizar el rendimiento del software.

#### **1.7.3.2. ISO/IEC 27001 (Seguridad de la Información)**

La ISO/IEC 27001 es una norma internacional que establece un marco para la gestión de la seguridad de la información. Enfocándose en la prevención de pérdida de datos y la protección de la información, esta norma proporciona directrices detalladas para establecer, implementar, mantener y mejorar un Sistema de Gestión de Seguridad de la Información (SGSI). Su aplicación implica evaluar y mejorar las medidas de seguridad implementadas, como controles de acceso, cifrado y gestión de riesgos, con el objetivo de garantizar la confidencialidad, integridad y disponibilidad de la información en entornos como servidores Proxmox. Integrar la ISO/IEC 27001 fortalecerá la seguridad de los datos, siendo fundamental en investigaciones centradas en la protección de la información crítica.

### **1.7.4. Pruebas**

#### **1.7.4.1. Pruebas de Caja Blanca**

En el contexto de la implementación de RAID en servidores Proxmox, las pruebas de caja blanca se centran en evaluar la estructura interna y el funcionamiento detallado del sistema RAID 5. Se analiza la configuración específica de los discos duros, la gestión de errores y la resiliencia frente a posibles fallos de hardware. Estas

pruebas ofrecen una visión profunda de la integridad del hardware RAID y su capacidad para mantener la disponibilidad de datos.

#### **1.7.4.2. Pruebas de Caja Negra**

Las pruebas de caja negra se enfocan en evaluar la funcionalidad externa del sistema RAID 5 sin necesidad de conocer su implementación interna. Se verifica cómo responde el sistema ante diversas configuraciones RAID, asegurando que cumpla con las expectativas declaradas en términos de redundancia y protección de datos. Estas pruebas son esenciales para garantizar la coherencia en las interacciones entre el sistema RAID 5 y Proxmox durante operaciones normales.

#### **1.7.4.3. Pruebas de Estrés**

Las pruebas de estrés se diseñan para evaluar el rendimiento del sistema RAID bajo condiciones extremas. Se somete el sistema a cargas de trabajo intensas para analizar su capacidad de respuesta y resistencia. Estas pruebas son cruciales para comprender cómo se comporta el sistema frente a un alto volumen de operaciones de lectura/escritura, garantizando la integridad de los datos incluso en situaciones de estrés.

### **1.8. HERRAMIENTAS**

#### **1.8.1. Hardware**

Para el presente estudio se utilizará una laptop de la marca HP con las siguientes características:

- Procesador: Intel® Core™ i5-1235U 10 núcleos y 12 subprocesos)
- Memoria RAM: 16 GB de RAM DDR4-3200 MHz (2 x 8 GB)
- Disco Duro: Unidad de estado sólido de 960 GB PCIe® NVMe™ M.2
- Tarjeta de Video: Gráficos Intel® Iris® Xe
- SO: Windows 11

#### **1.8.2. Software**

- Proxmox VE: Proxmox Virtual Environment es una plataforma completa de administración de servidores de código abierto para la virtualización empresarial. Integra estrechamente el hipervisor KVM y los contenedores Linux (LXC), el almacenamiento definido por software y la funcionalidad de red, en una sola plataforma. Con la interfaz de usuario integrada basada en web, puede

administrar máquinas virtuales y contenedores, alta disponibilidad para clústeres o las herramientas integradas de recuperación ante desastres con facilidad.

- VirtualBox: VirtualBox es un potente producto de virtualización x86 y AMD64/Intel64 para uso empresarial y doméstico. VirtualBox no solo es un producto extremadamente rico en funciones y de alto rendimiento para clientes empresariales, sino que también es la única solución profesional que está disponible gratuitamente como software de código abierto bajo los términos de la Licencia Pública General GNU (GPL) versión 3.
- Microsoft Azure: Según (Wikipedia)Microsoft Azure (anteriormente Windows Azure y Azure Services Platform) es una plataforma de computación en la nube creado por Microsoft para construir, probar, desplegar y administrar aplicaciones y servicios mediante el uso de sus centros de datos. Proporciona software como servicio (SaaS), plataforma como servicio (PaaS) e infraestructura como servicio (IaaS) y es compatible con muchos lenguajes, herramientas y marcos de programación diferentes, incluidos software y sistemas específicos de Microsoft y de terceros.

## **1.9. LIMITES Y ALCANCES**

### **1.9.1. Limites**

Las limitaciones de esta tesis se refieren a posibles restricciones que puedan surgir durante la investigación. Entre estas limitaciones se incluyen factores externos que podrían afectar la validez o generalización de los resultados. Además, se reconoce que la aplicación práctica se limita a entornos virtuales y no se extiende a infraestructuras físicas. Cualquier restricción adicional que pueda surgir durante el proceso de investigación se abordará con transparencia para garantizar la integridad de los resultados.

### **1.9.2. Alcances**

El alcance de esta tesis se delimita al ámbito de los servidores Proxmox, donde se llevará a cabo la implementación y evaluación de las configuraciones RAID 5. El estudio se centrará en la prevención de la pérdida de datos en este entorno específico. La unidad de análisis será el sistema RAID 5 implementado en servidores Proxmox. El

espacio estará circunscrito a entornos virtuales simulados, y el tiempo abarcará el periodo de implementación y pruebas en dicho entorno.

## **1.10. APORTES**

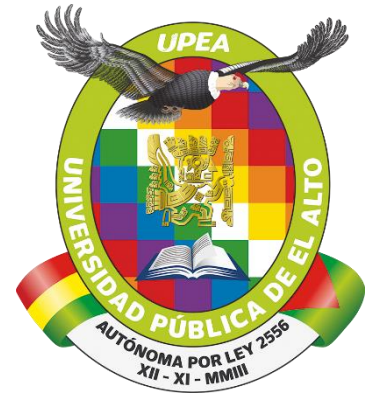
### **1.10.1 Aporte Teórico**

El aporte teórico de este trabajo se sustenta en la aplicación y adaptación de teorías y conceptos fundamentales relacionados con la gestión de riesgos, seguridad informática y almacenamiento de datos, con un enfoque específico en la reducción del riesgo de pérdida de datos en servidores Proxmox mediante la implementación de tecnologías RAID 5.

### **1.10.2 Aporte Practico**

Este estudio busca ofrecer fundamentos prácticos a aquellos interesados en implementar diversas configuraciones RAID 5, abordando de manera efectiva la prevención de pérdida de datos en servidores Proxmox. La investigación se propone proporcionar soluciones alternativas respaldadas por un sólido marco teórico, ofreciendo así orientación práctica para mejorar la seguridad de los datos en entornos específicos de servidores Proxmox.





# CAPITULO II

## MARCO

## TEORICO



## **2.1. MODELO**

“Es una representación simplificada de un sistema o realidad compleja que se elabora con el propósito de facilitar la comprensión y el estudio de dicho sistema. Esta representación puede ser cualitativa o cuantitativa y se utiliza para examinar las relaciones entre los factores considerados importantes en el funcionamiento del sistema.” (Wilson, 1993)

Además, un modelo puede ser una construcción intelectual y descriptiva de una entidad que interesa a al menos un observador, y se puede relacionar con los observables del mundo.

## **2.2. DEFINICION DE MODELO PARA SERVIDORES**

“Un modelo en el área de servidores encapsula la esencia organizativa y funcional que delinea la configuración y administración precisa de los recursos físicos y virtuales dentro de un centro de datos o servidores.” (Def. Propia, 2024)

### **2.2.1. Modelos de Centro de Datos**

Según un relevamiento realizado por el Grupo Dell ‘Oro, se proyecta que los gastos globales en Data Centers aumentarán un 10 % en los próximos 5 años, alcanzando un total de USD 350.000 millones (Grupo Dell ‘Oro, 2023). Con el crecimiento exponencial de la demanda, es común que se implementen nuevas tecnologías y se mejoren las existentes es por ello que a continuación se detallan los tipos de modelos de centro de datos.

#### **2.2.1.1. Enterprise**

Según (Ascenty, 2024), el Enterprise es uno de los modelos más comunes del mercado. Su principal característica es que toda la infraestructura del Data Center se encuentra dentro de la propia empresa, es decir, es exclusiva.

A pesar de tener un alto costo de estructuración y mantenimiento, diversas organizaciones optan por este formato. Ello, porque permite un control mayor sobre las operaciones y procesos relacionados a la seguridad.

#### **2.2.1.2. Colocation**

Según (Ascenty, 2024), colocation a diferencia del Enterprise, se basa en una gestión descentralizada. En este formato, la infraestructura física no se encuentra dentro de la empresa, sino en un ambiente externo, suministrado por un proveedor.

Por lo tanto, el negocio puede contratar espacio, hardware, ancho de banda y diversas soluciones para Data Center, de acuerdo con su demanda. Esto significa que, incluso sin un presupuesto elevado para tener su propia infraestructura, es posible disfrutar de todas las ventajas de contar con un centro de procesamiento de datos.

Usted paga por lo que usa y el proveedor se responsabiliza por mantener sus datos siempre seguros y accesibles. Y lo mejor: si la empresa crece, basta ampliar el espacio de almacenamiento, de forma simple y económica.

#### **2.2.1.3. Internet**

Según (Ascenty, 2024), este tipo de Data Center se basa en la nube, y se lo clasifica como virtual. Esto significa que tampoco necesita la construcción de infraestructura propia. Las informaciones permanecerán hospedadas y se procesarán en los servidores de un proveedor.

Se trata de un modelo interesante, especialmente por su flexibilidad y escalado. Además, democratiza el acceso a los recursos computacionales de forma remota, como un servicio.

Es importante saber que la infraestructura es de propiedad del proveedor y será compartida con diferentes empresas. De esta forma, el acceso a las informaciones se realiza por medio de internet.

#### **2.2.1.4. Hyperscale**

Según (Ascenty, 2024), Hyperscale es un tipo de Data Center proyectado para atender operaciones en gran escala. Por eso, cuenta con una infraestructura robusta, capaz de soportar actividades que demandan gran volumen de almacenamiento y alto poder de procesamiento.

Por ser vital para el funcionamiento de actividades importantes, el modelo necesita contar con mecanismos reforzados para garantizar la continuidad de las operaciones. En otras palabras, cuenta con una infraestructura de punta, con redundancia, backup y complejas redes de comunicación y suministro de energía.

#### **2.2.1.5. Edge**

Según (Ascenty, 2024), el modelo Edge pasó a ser cada vez más común en el mercado. El Edge está dirigido a operaciones que demandan una comunicación más ágil entre los usuarios y los servidores, solucionando el problema de la latencia.

El término «Edge» significa borde, lo que traduce bien la idea de su funcionamiento. En este formato, los Data Centers son menores y más numerosos, con el objetivo de cubrir un área mayor del territorio y estar lo más cerca posible de los usuarios finales de los servicios.

## **2.3. RAID (Redundant Array of Independent Disks)**

### **2.3.1. Definición de RAID**

“Es la combinación de dos o más discos en una matriz para lograr objetivos de rendimiento o redundancia de datos, que no se pueden lograr con una unidad grande y/o costosa. Esta matriz de dispositivos aparece en el equipo como una sola unidad de almacenamiento lógico, lo que permite gestionar y aumentar la eficacia en el manejo de datos.” (Def. Propia 2024)

### **2.3.2. Tipos de RAID**

Los tipos de RAID se dividen en tres categorías principales: Firmware RAID, Hardware RAID y Software RAID. Estas configuraciones ofrecen distintos enfoques de gestión de almacenamiento de datos, cada uno con sus respectivas ventajas y consideraciones a tomar en cuenta.

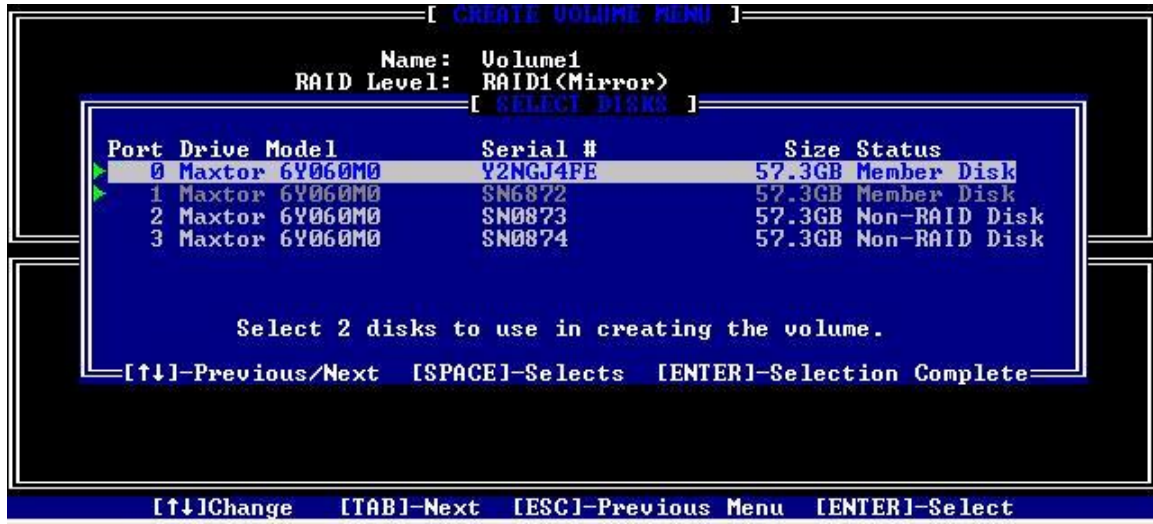
#### **2.3.2.1. Firmware RAID**

“El RAID de firmware, también conocido como ATARAID, es un tipo de RAID de software en el que los conjuntos RAID se configuran mediante un menú basado en firmware. El firmware utilizado por este tipo de RAID también se conecta al BIOS, lo que le permite arrancar desde sus conjuntos RAID.” (Red Hat, 2024)

Los diferentes proveedores utilizan diferentes formatos de metadatos en disco para marcar los miembros del conjunto RAID. Un ejemplo de sistema RAID de firmware es Intel Matrix RAID.

**Figura 1**

*BIOS Intel Matrix RAID.*



*Nota. Vista de la BIOS de un controlador Intel Matrix RAID, según página (Intel, 2024), (<https://www.intel.com/content/www/us/en/support/articles/000005789/technologies.html>).*

### **2.3.2.2. Hardware RAID**

“Hardware RAID es un tipo de configuración de RAID en la cual el subsistema RAID es gestionado independientemente del sistema operativo anfitrión por una controladora RAID dedicada.” (Red Hat, Inc., 2024)

Esta controladora puede alojar múltiples dispositivos en una matriz RAID y presenta los volúmenes resultantes como unidades lógicas al sistema operativo. Los dispositivos RAID de hardware pueden ser internos o externos al sistema.

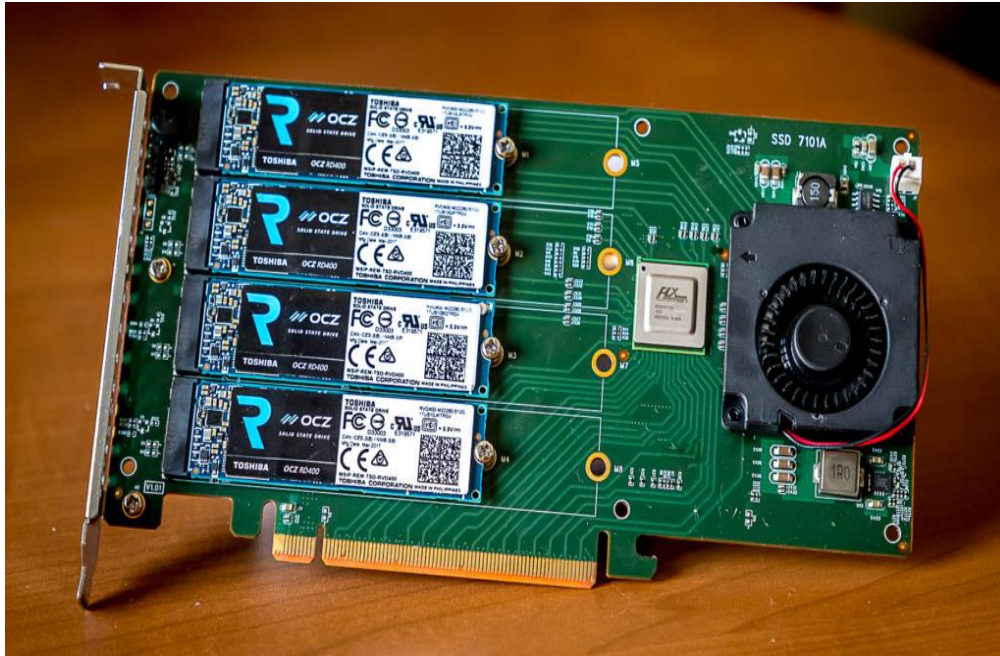
Los dispositivos internos suelen consistir en una tarjeta controladora especializada que administra las operaciones de RAID de manera transparente para el sistema operativo. Por otro lado, los dispositivos externos se conectan al sistema a través de interfaces como SCSI, Fibre Channel, iSCSI, InfiniBand, entre otras, y también presentan los volúmenes como unidades lógicas al sistema.

Las tarjetas controladoras RAID actúan como controladores SCSI para el sistema operativo y gestionan todas las operaciones de lectura y escritura de la unidad. A pesar de las diferencias en la implementación física, desde el punto de vista del sistema operativo, las unidades RAID se comportan de manera similar a las unidades

SCSI estándar, lo que garantiza una integración transparente con el entorno de almacenamiento.

## Figura 2

Controlador RAID NVMe HighPoint SSD7101A-1: SSD M.2



*Nota.* controlador RAID HighPoint SSD7101A-1 NVMe SSD a pleno ancho de banda x16, según página (The SSD Review is a participant in the Amazon Services LLC Associates Program, 2024), (<https://www.thesdreview.com/raid-enterprise/raid-cards/highpoint-ssd7101a-1-nvme-raid-controller-review-samsung-toshiba-m-2-ssds-tested/3/>).

### 2.3.2.3. Software RAID

“El RAID de software implementa los distintos niveles de RAID directamente en el código del dispositivo de bloque del kernel. Esta solución es más económica que el hardware RAID, ya que no requiere costosas tarjetas controladoras de disco ni chasis intercambiables en caliente.” (Red Hat, Inc., 2024)

Además, con el chasis intercambiable en caliente, es posible extraer un disco duro sin necesidad de apagar el sistema. El software RAID es compatible con diferentes tipos de almacenamiento en bloque, como SATA, SCSI y NVMe, siempre que sean compatibles con el kernel de Linux. Hoy en día, con las CPU más rápidas, el

RAID de software suele superar al RAID de hardware, a menos que se utilicen dispositivos de almacenamiento de gama alta.

El kernel de Linux incluye un controlador de dispositivos múltiples (MD), lo que hace que la solución RAID sea completamente independiente del hardware. El rendimiento de una matriz basada en software depende del rendimiento y la carga de la CPU del servidor.

Algunas características clave de la pila RAID de software de Linux:

- Diseño multihilo
- Reconstrucción de matrices en segundo plano mediante recursos del sistema inactivos
- Compatibilidad con unidades de intercambio en caliente
- Detección automática de CPU para aprovechar ciertas características de la CPU, como la transmisión de datos múltiples de instrucción única (SIMD).
- Corrección automática de sectores defectuosos en discos de una matriz.
- Verificaciones regulares de la integridad de los datos RAID con el fin de asegurar la salud del conjunto de discos.
- Mapas de bits de intención de escritura, que aumentan drásticamente la velocidad de los eventos de resincronización al permitir que el kernel sepa con precisión qué partes de un disco deben volver a sincronizarse en lugar de tener que volver a sincronizar toda la matriz después de un bloqueo del sistema.

### **2.3.3. Niveles de RAID**

A continuación, se describen las configuraciones admitidas en RAID, por niveles y lineales:

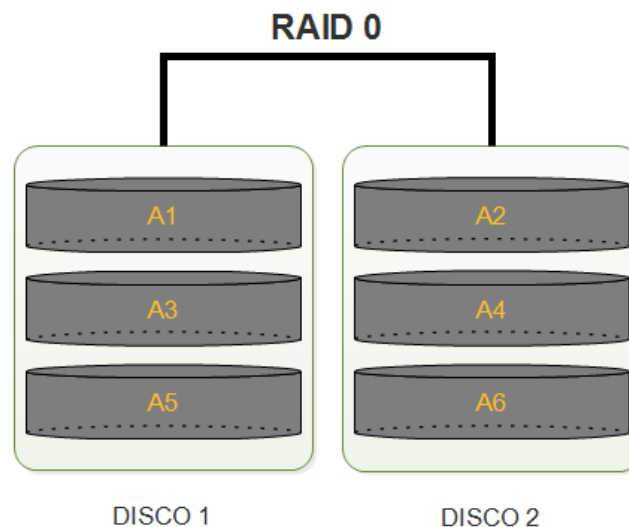
#### **2.3.3.1. RAID 0**

Conocido también como "STRIPING", es una configuración de RAID que divide los datos en bloques y los distribuye equitativamente entre varios discos. Esto se hace para mejorar el rendimiento de lectura y escritura al permitir múltiples operaciones de entrada/salida (E/S) simultáneas. Sin embargo, no proporciona redundancia, lo que significa que, si uno de los discos falla, se pierden todos los datos. (Red Hat, Inc., 2024)

Esta configuración RAID se centra en el rendimiento al dividir los datos en franjas y escribirlos en los discos miembros de la matriz. Esta técnica de mapeo de datos seccionados ofrece un alto rendimiento de E/S a un bajo costo, pero carece de redundancia. Las implementaciones de RAID 0 dividen los datos entre los dispositivos miembros hasta el tamaño del dispositivo más pequeño de la matriz. Esto significa que, si los discos tienen tamaños ligeramente diferentes, se tratan como si tuvieran el mismo tamaño que el disco más pequeño. Por lo tanto, la capacidad de almacenamiento común de una matriz RAID 0 es la capacidad total de todos los discos, utilizando todas las zonas disponibles de los discos miembros, incluso si tienen tamaños diferentes.

### Figura 3

*Asignación de datos en RAID 0*



Nota. Distribución de datos en RAID 0, imagen propia, 2024.

#### 2.3.3.2. RAID 1

También conocido como "espejo", duplica los datos en dos o más discos, proporcionando una redundancia completa. Esta duplicación garantiza que, si un disco falla, los datos todavía estén disponibles en el otro disco. Sin embargo, el almacenamiento efectivo es igual a la capacidad de un solo disco.

Este nivel de RAID ofrece redundancia al escribir los mismos datos en cada disco miembro de la matriz, creando una copia idéntica en cada uno. Aunque este enfoque puede ser menos eficiente en términos de espacio, ya que se utiliza la misma

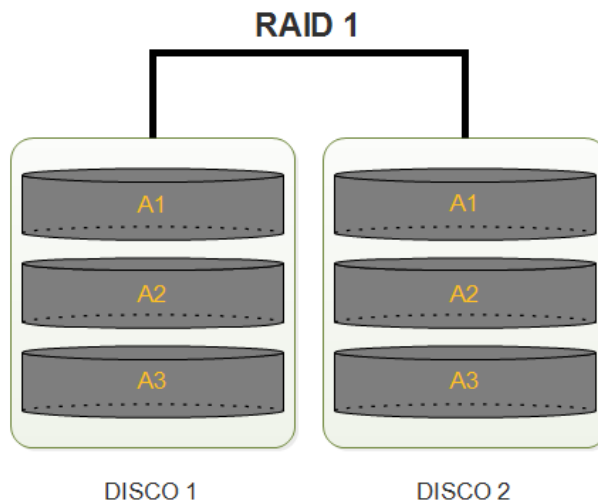


cantidad de almacenamiento en cada disco, proporciona un alto nivel de disponibilidad y fiabilidad de datos. Este método simple y fiable lo hace popular, especialmente para aplicaciones críticas.” (Red Hat, Inc., 2024)

A pesar de que el RAID 1 puede ser más costoso en términos de espacio que otros niveles basados en paridad, como el RAID 5, ofrece ventajas significativas en términos de rendimiento y simplicidad. Las operaciones de lectura en un RAID 1 pueden ser más rápidas debido a la duplicación de datos, mientras que el proceso de escritura consume menos recursos de CPU en comparación con los niveles basados en paridad.

#### Figura 4

Arquitectura de la matriz RAID 1



*Nota. Distribución de datos en RAID 1, imagen propia, 2024.*

#### 2.3.3.3. RAID 4

Esta configuración RAID que buscan proporcionar redundancia y rendimiento en sistemas de almacenamiento. La cual utiliza una estrategia donde la paridad se concentra en un solo disco, calculada en función del contenido de los otros discos en la matriz. Esta paridad se utiliza para reconstruir los datos en caso de fallo de disco, permitiendo la continuidad de las operaciones de entrada/salida (E/S) incluso cuando hay un disco defectuoso.

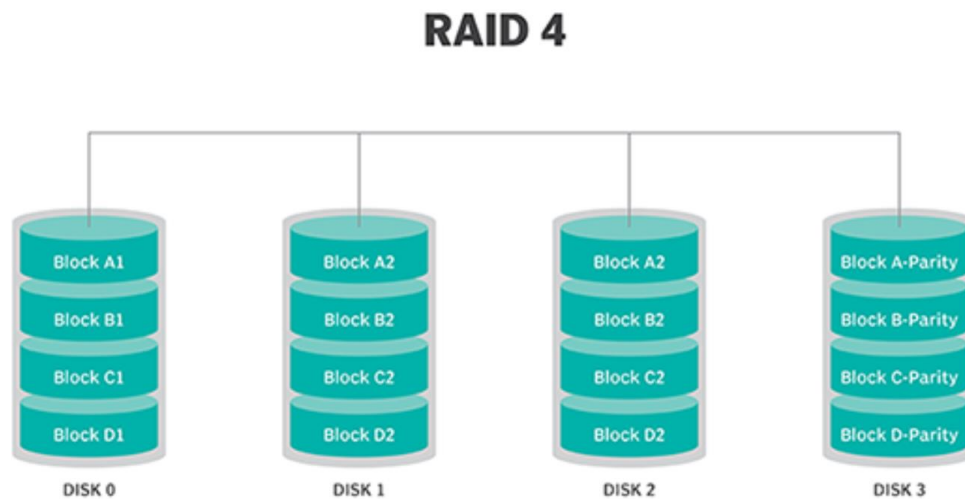
Sin embargo, esta concentración de paridad en un solo disco puede convertirse en un cuello de botella en las transacciones de escritura, lo que puede limitar el

rendimiento. Por esta razón, el RAID 4 rara vez se utiliza sin tecnologías complementarias como el almacenamiento en caché de reescritura.

“La capacidad de almacenamiento de un RAID 4 se calcula multiplicando la capacidad del disco más pequeño por el número total de discos menos uno. Aunque ofrece un buen rendimiento de lectura, las operaciones de escritura pueden ser más lentas debido a la escritura secuencial de la paridad.” (Red Hat, Inc., 2024)

### Figura 5

Arquitectura de la matriz RAID 4



*Nota. Distribución de datos en los distintos discos en RAID 4, según (Mefics, 2020)*

#### 2.3.3.4. Raid 5

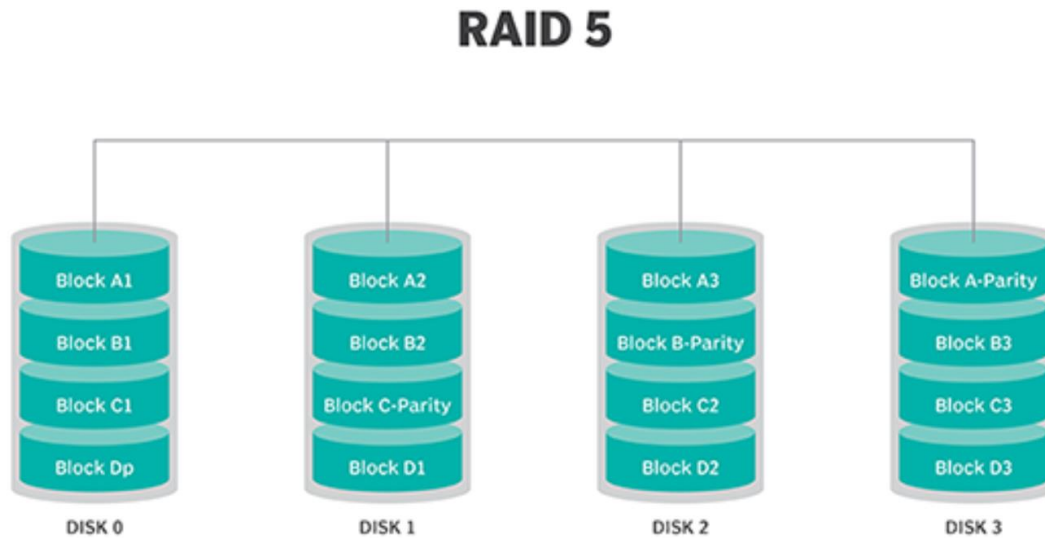
Este es el tipo más común de configuración RAID, distribuyendo la paridad entre todos los discos miembro de la matriz. Esta distribución de paridad entre discos elimina el cuello de botella inherente al nivel 4 y mejora el rendimiento de escritura. Sin embargo, en configuraciones con un gran número de discos, la velocidad combinada de transferencia de datos puede convertirse en un nuevo cuello de botella para el cálculo de la paridad, aunque las CPU modernas pueden calcularla rápidamente.

“A pesar de tener un rendimiento asimétrico, donde las lecturas superan sustancialmente a las escrituras, RAID 5 sigue siendo popular debido a su combinación de rendimiento y redundancia. La capacidad de almacenamiento en una matriz RAID 5 se calcula de manera similar a RAID 4, ya que ambos utilizan la paridad

para la redundancia, pero se distribuye de manera diferente entre los discos miembros.” (Red Hat, Inc., 2024)

### Figura 6

Arquitectura de la matriz RAID 5



*Nota. Asignación de datos RAID 5, según (Mefics, 2020).*

#### 2.3.3.5. RAID 6

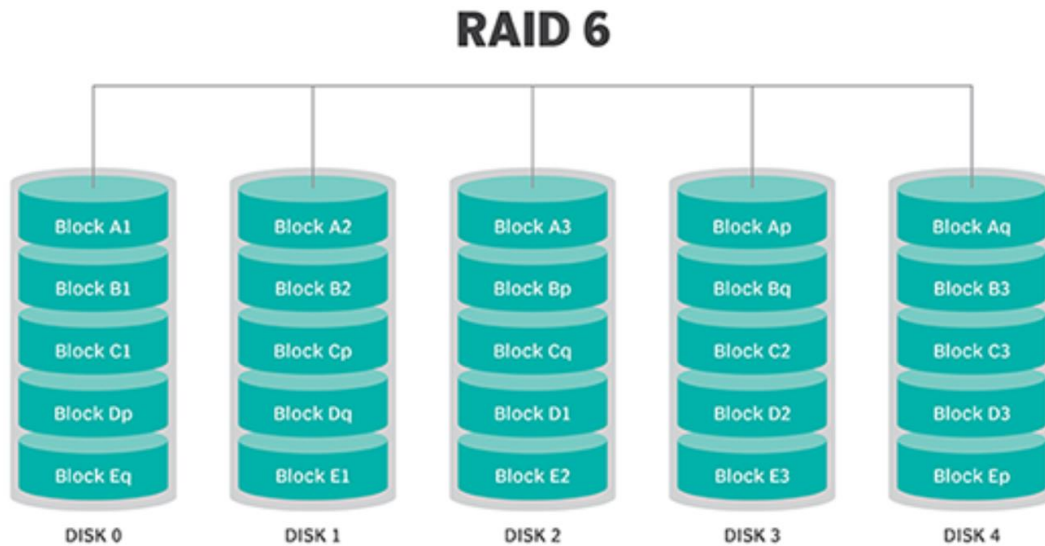
Este nivel comúnmente utilizado cuando la redundancia y la integridad de los datos son prioridades sobre el rendimiento, pero la ineficiencia del espacio en RAID 1 no es aceptable. Este nivel de RAID emplea un esquema de paridad complejo que le permite recuperarse de la pérdida de hasta dos unidades en la matriz, ofreciendo una mayor protección contra fallos en comparación con RAID 5. Sin embargo, este esquema de paridad más complejo conlleva una carga significativamente mayor para la CPU en dispositivos RAID de software, lo que puede resultar en una mayor asimetría en el rendimiento, especialmente durante las transacciones de escritura.

“La capacidad total de una matriz RAID 6 se calcula de manera similar a RAID 5 y RAID 4, pero debe restar dos dispositivos en lugar de uno del recuento total de dispositivos (Red Hat, Inc., 2024) para tener en cuenta el espacio adicional necesario para la paridad duplicada. En resumen, RAID 6 ofrece una mayor protección de datos al permitir la pérdida simultánea de hasta dos discos, pero puede tener un rendimiento

más asimétrico en comparación con RAID 4 y RAID 5 debido a la complejidad del esquema de paridad.” (Red Hat, Inc., 2024)

### Figura 7

Arquitectura de la matriz RAID 6.



Nota. Asignación de datos RAID 6, según (Mefics, 2020).

#### 2.3.3.6. RAID 10

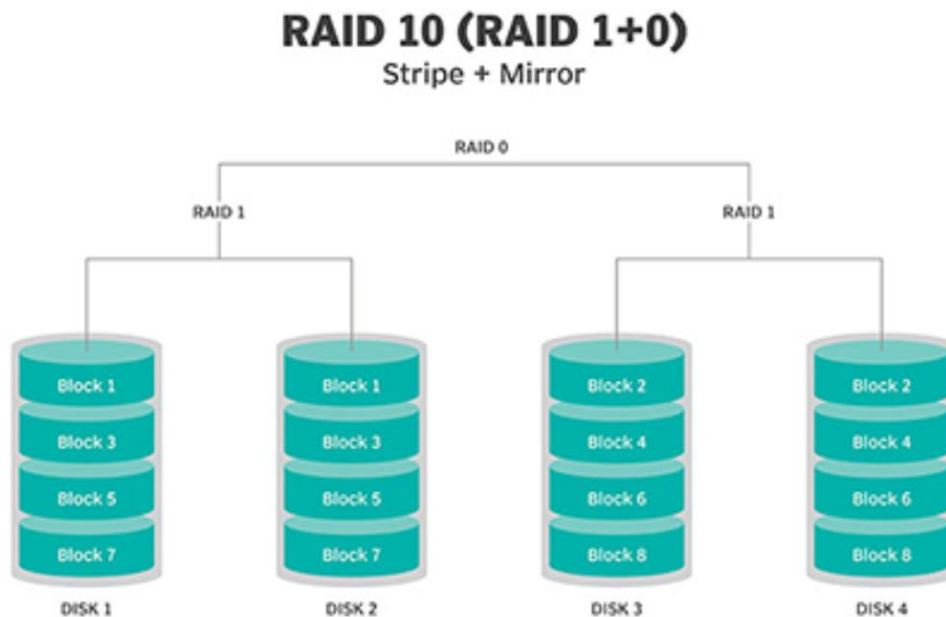
“Conocido como RAID 1+0, fusiona las fortalezas de RAID 1 y RAID 0 para ofrecer alta redundancia y alto rendimiento en entornos de almacenamiento. Este nivel de RAID implica duplicar los datos para proporcionar redundancia (RAID 1), mientras que simultáneamente distribuye estos datos duplicados en bloques para mejorar el rendimiento de lectura y escritura (RAID 0). Requiere un número mínimo de discos (generalmente al menos 3 unidades) para su implementación.” (Red Hat, Inc., 2024)

Además, el RAID 10 busca equilibrar la eficiencia del espacio al permitir configuraciones donde se almacenan dos copias de cada dato. Esto permite que el tamaño total de la matriz sea mayor que en una configuración de RAID 1 convencional, utilizando menos espacio en comparación con matrices de RAID 1 de más de dos dispositivos. Sin embargo, a diferencia de algunos otros niveles de RAID, como RAID 6, no implica el uso de cálculos de paridad que puedan afectar negativamente el rendimiento del sistema. “Es importante destacar que la creación de una matriz RAID

10 no está soportada durante la instalación inicial, pero puede ser configurada manualmente después del proceso de instalación.

### Figura 8

Arquitectura de la matriz RAID 10.



*Nota. Arquitectura de la matriz RAID 10, según (Mefics, 2020).*

#### 2.3.3.7. Niveles de RAID no estándar

Existen niveles de RAID no convencionales que difieren de los estándares establecidos y suelen ser desarrollados por empresas u organizaciones para su uso interno. A continuación, se presentan algunos ejemplos:

- RAID 7: Un nivel de RAID no convencional que se basa en el RAID 3 y el RAID 4, pero incorpora una memoria caché. Incluye un sistema operativo embebido en tiempo real como controlador, almacenamiento en caché<sup>1</sup> a través de un bus de alta velocidad y otras características propias de un sistema autónomo. (Def. Propia, 2024)
- RAID Adaptativo: Este nivel permite al controlador RAID determinar la mejor manera de almacenar la paridad en los discos. Puede optar entre RAID 3 y

<sup>1</sup> Caché: Almacenamiento temporal de datos frecuentemente accedidos para mejorar la velocidad de acceso y reducir la carga en el sistema principal de almacenamiento.

RAID 5, según el tipo de conjunto RAID que se adapte mejor al tipo de datos que se estén escribiendo en los discos. (Def. Propia, 2024)

#### **2.3.3.8. RAID LINEAL**

Según (Red Hat, 2024), RAID Lineal es una configuración donde varios discos se agrupan para formar un único volumen de almacenamiento. A diferencia de otros tipos de RAID, el RAID Lineal no ofrece redundancia ni mejora de rendimiento. En esta configuración, los datos se escriben secuencialmente en los discos uno tras otro. Si un disco falla, se pierden todos los datos a partir de ese punto en adelante.

El RAID Lineal es útil para simplemente aumentar la capacidad de almacenamiento sin preocuparse por la redundancia o el rendimiento. Sin embargo, debido a su falta de redundancia, no es adecuado para aplicaciones donde la integridad de los datos es crítica.

#### **2.3.4. Importancia del RAID en la Virtualización**

Según (Gómez Torre Eduardo, 2016) la configuración RAID desempeña un papel fundamental en el entorno de la virtualización al influir directamente en el rendimiento y la integridad de los datos almacenados. Al implementar RAID de manera adecuada, se pueden lograr mejoras significativas en las velocidades de lectura, escritura o ambas, según las necesidades específicas del sistema y la configuración seleccionada.

Una de las características más destacadas y valiosas del RAID es su capacidad para proporcionar redundancia de datos, esta actúa como un escudo protector adicional, salvaguardando la información contra posibles pérdidas ocasionadas por diversos factores, como fallos de hardware, errores humanos o condiciones ambientales adversas. En otras palabras, el RAID ofrece una capa adicional de seguridad que garantiza la disponibilidad y la integridad de los datos en todo momento, incluso en situaciones imprevistas o de emergencia.

Esta combinación de mejoras en el rendimiento y protección de datos hace que la implementación adecuada del RAID sea esencial para cualquier entorno de virtualización. Además de optimizar el funcionamiento general del sistema, el RAID brinda tranquilidad al proporcionar una salvaguarda confiable contra posibles pérdidas

de datos, lo que resulta fundamental en entornos empresariales donde la seguridad y la disponibilidad de la información son de suma importancia.

### **2.3.5. Consideraciones de Implementación de RAID en Proxmox VE**

Al considerar la implementación de RAID en Proxmox VE, es esencial evaluar las opciones disponibles y comprender las consideraciones específicas asociadas con cada tecnología. Dos alternativas comunes son ZFS y MDRAID, cada una con sus propias características y consideraciones.

#### **2.3.5.1. ZFS Seguridad y Facilidad de Uso**

Según (Proxmox Server Solutions, 2024) ZFS es un sistema de archivos y gestor de volúmenes lógicos que ofrece protección avanzada de datos y características de administración intuitivas. Es reconocido por su fiabilidad y facilidad de configuración, lo que lo convierte en una opción popular para entornos de virtualización como Proxmox VE.

Sus ventajas clave incluyen:

- Protección avanzada contra la corrupción de datos.
- Compresión de datos a nivel de sistema de archivos para ahorrar espacio.
- Funcionalidad de instantáneas para respaldar y restaurar datos de manera eficiente.

#### **2.3.5.2. MDRAID Alternativa Simplificada**

Según (Proxmox Server Solutions, 2024) MDRAID, o RAID por software, ofrece una solución de almacenamiento de datos más básica pero confiable. Aunque carece de algunas de las características avanzadas de ZFS, sigue siendo una opción viable para entornos donde se requiere una protección básica de datos.

Sus aspectos destacados son:

- Configuración sencilla y administración directa a través del sistema operativo.
- Funcionalidad de RAID básica para redundancia de datos.

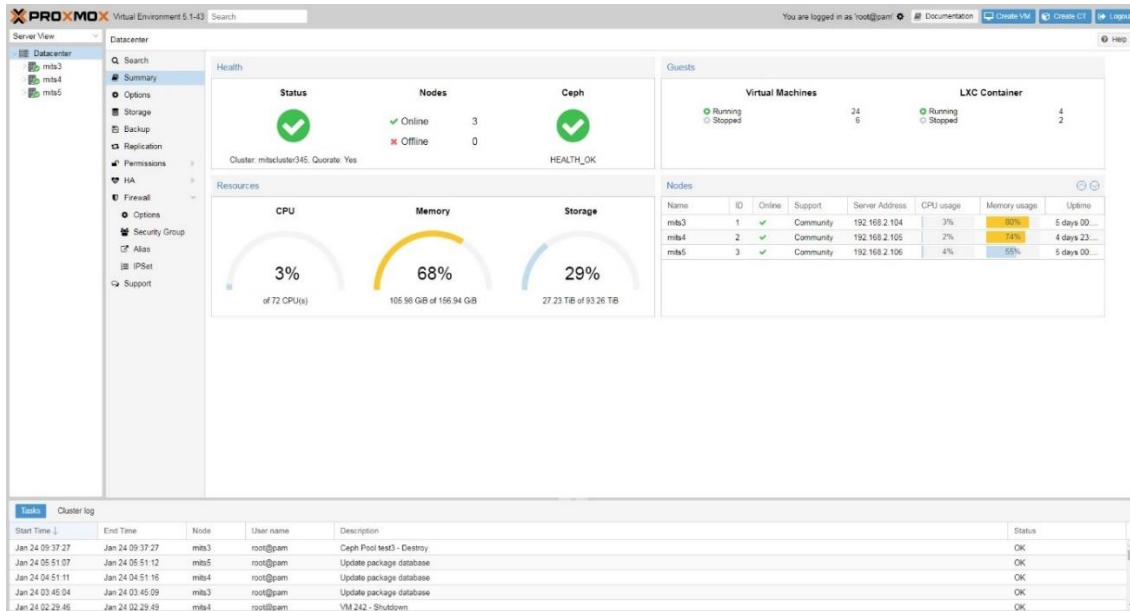
## **2.4. PROXMOX VIRTUAL ENVIROMENT**

“Proxmox Virtual Environment es una plataforma completa de gestión de servidores de código abierto para la virtualización empresarial. Integra estrechamente el hipervisor KVM y los contenedores Linux (LXC), el almacenamiento definido por

software y la funcionalidad de red, en una sola plataforma.” (Proxmox Server Solutions, 2024).

**Figura 9**

Proxmox VE



*Nota. Interfaz web Proxmox (Elaboración Propia).*

**2.4.1. ¿Para qué se usa Proxmox?**

Proxmox puede ser usada como una solución de virtualización para poder:

- Preparar entornos de prueba donde se requiera de uno o varios servidores, para pruebas o estudios.
- Maximizar los recursos de un data center de manera que un servidor con buenas prestaciones pueda ser dividida usando la virtualización.
- Crear infraestructura de redes tanto simples como complejas.
- Recortar gastos de operación y mantenimiento.

**2.4.2. Características de Proxmox**

- Plataforma de código abierto basada en GNU Debian<sup>2</sup>.

<sup>2</sup> GNU Debian: Sistema operativo basado en Linux, conocido por su estabilidad y su amplia colección de software libre disponible a través de su sistema de gestión de paquetes.



- Virtualización de hardware usando KVM y software a través de contenedores LXC<sup>3</sup>.
- Integración con almacenamiento de tipos NAS<sup>4</sup>, SAN<sup>5</sup> a través de fibra óptica, iSCSI o NFS.
- Alta disponibilidad (HA High Availability) en clúster a través de nodos para un funcionamiento ininterrumpido a través de migración en vivo en caso de fallas.
- Administración y configuración mediante la interfaz web y terminal, vía local o de forma remota a través de SSH<sup>6</sup>.
- Administración basada en roles para un acceso granular para todos los objetos (como VMs, almacenamiento, nodos, etc.) usando roles basados en usuario y permisos.
- Integración de datos definida en un esquema JSON<sup>7</sup> a través de una API RESTful<sup>8</sup>, la cual facilita una integración más rápida con terceras herramientas.
- Restauración y copias de respaldo en vivo, sin interrumpir en el funcionamiento de los servidores.
- Almacenamiento flexible ya que Proxmox está sujeta a ser usada con todas las tecnologías de almacenamiento basadas en Debian Linux.

### 2.4.3. Uso de Proxmox VE

#### 2.4.3.1. Desarrollo y Pruebas de Software

En entornos de desarrollo y pruebas de software, Proxmox VE proporciona un ambiente seguro y eficiente para la creación de máquinas virtuales independientes

---

<sup>3</sup> LXC: Contenedores Linux, una tecnología de virtualización ligera que permite la ejecución de múltiples sistemas Linux en un único sistema host.

<sup>4</sup> NAS: Almacenamiento conectado a la red (Network-Attached Storage, por sus siglas en inglés), es un dispositivo de almacenamiento de datos que se conecta a una red para proporcionar acceso compartido a archivos y datos a través de la red.

<sup>5</sup> SAN: Red de área de almacenamiento (Storage Area Network, por sus siglas en inglés), es una red dedicada de almacenamiento que conecta múltiples dispositivos de almacenamiento, como discos duros y matrices de discos, a servidores y otros dispositivos de almacenamiento.

<sup>6</sup> SSH: Protocolo de red para acceso seguro a computadoras remotas mediante una conexión cifrada.

<sup>7</sup> JSON: Formato ligero de intercambio de datos que se utiliza para almacenar y transferir información estructurada entre sistemas.

<sup>8</sup> API RESTful: Interfaz de programación para comunicación entre sistemas mediante HTTP y principios REST.

para cada proyecto o prueba. Por ejemplo, una empresa de desarrollo de software puede utilizar Proxmox VE para configurar rápidamente entornos de pruebas específicos para diferentes versiones de su aplicación, lo que facilita la detección de errores y la realización de pruebas exhaustivas antes del lanzamiento.

#### **2.4.3.2. Infraestructuras de Servidores Empresariales**

En el ámbito empresarial, Proxmox VE se utiliza para gestionar infraestructuras de servidores empresariales, ofreciendo una plataforma robusta y flexible para la consolidación de servidores. Por ejemplo, una empresa que busca optimizar sus recursos de hardware puede utilizar Proxmox VE para virtualizar sus servidores físicos, lo que permite una mejor utilización de los recursos y una reducción de costos operativos.

#### **2.4.3.3. Implementaciones de Alta Disponibilidad**

Según (Ekaterina Kaloshina, 2017) para garantizar la disponibilidad y el rendimiento de las aplicaciones críticas, Proxmox VE permite implementaciones de alta disponibilidad mediante la configuración de clústeres de servidores. Por ejemplo, una institución financiera puede utilizar Proxmox VE para crear un clúster de servidores altamente disponibles que garantice la disponibilidad continua de sus sistemas de banca en línea, incluso en caso de fallo de hardware o mantenimiento programado.

#### **2.4.3.4. Consolidación de Servidores**

Proxmox VE facilita la consolidación de servidores al permitir la ejecución de múltiples máquinas virtuales en un único servidor físico. Por ejemplo, una empresa con una infraestructura de servidores dispersa puede utilizar Proxmox VE para consolidar sus cargas de trabajo en menos hardware físico, lo que resulta en ahorros significativos en costos de adquisición y mantenimiento.

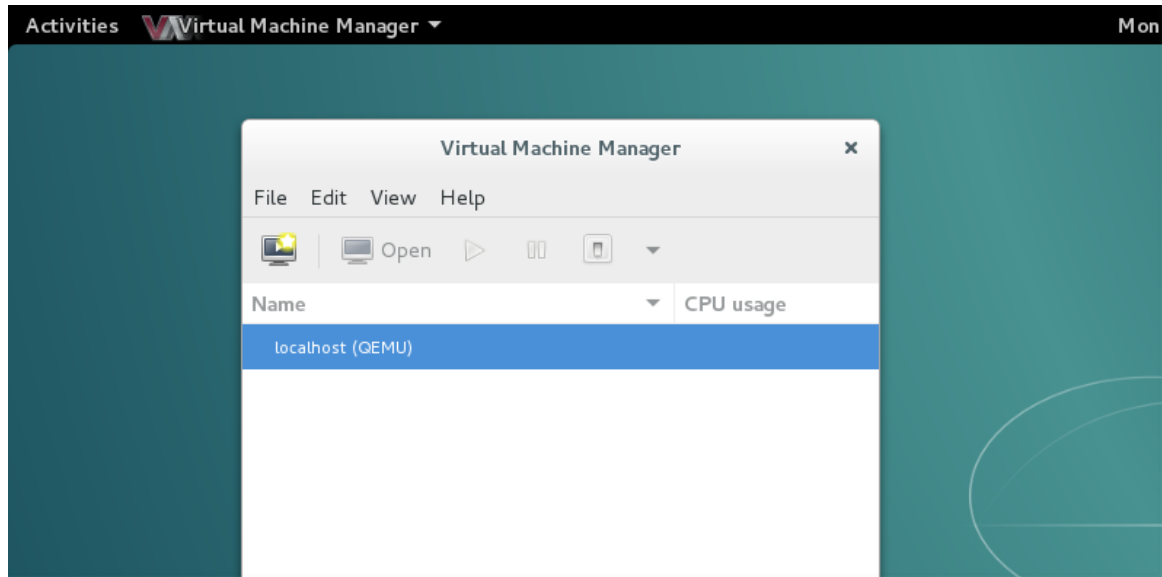
### **2.5. KVM (Kernel-Based Virtual Machine)**

Según (IONOS Cloud S.L.U., 2020), La tecnología conocida como Kernel-based Virtual Machine (KVM) fue introducida por la empresa Qumranet en el año 2000 y fue incorporada en el kernel de Linux 2.6.20 a principios de 2007. Al año siguiente, RedHat, un proveedor líder de distribuciones Linux, adquirió Qumranet. Para continuar el desarrollo de esta tecnología, RedHat y IBM colaboraron en la creación del proyecto Open Virtualization Alliance (OVA), al que se unieron empresas como Suse, Intel y HP.

KVM está estrechamente vinculado al software de emulación QEMU<sup>9</sup>. Mientras QEMU se encarga principalmente de la emulación del hardware, KVM se encarga de la gestión de los recursos asignados a los sistemas invitados. Por esta razón, es común encontrar la terminología KVM/QEMU. Dado que KVM es una parte integral del kernel de Linux, no puede ser ejecutado en sistemas Windows. Sin embargo, Windows y la mayoría de los sistemas Unix<sup>10</sup>, como Linux<sup>11</sup>, Solaris y BSD, pueden ser ejecutados utilizando soluciones de virtualización que emplean KVM. Aunque KVM viene preinstalado en la mayoría de las distribuciones Linux, necesita ser activado manualmente antes de su uso.

## Figura 10

### Interfaz Gráfica KVM



*Nota. Interfaz gráfica KVM a través del uso de la librería libvirt (Fuente prop.).*

### 2.5.1. Componentes de KVM

Según (IONOS Cloud S.L.U., 2020) KVM en su conjunto está conformada por varios elementos:

- Extensión del núcleo: esta parte engloba la propia extensión, conocida como kernel.ko, la cual actúa como una capa de traducción entre el hardware físico y el entorno virtual. Además de esto, contiene los módulos kvm-amd y kvm-intel,

<sup>9</sup> QEMU: emulador y virtualizador de código abierto para múltiples arquitecturas.

<sup>10</sup> Unix: sistema operativo multitarea y multiusuario, conocido por su estabilidad y seguridad.

<sup>11</sup> Linux: sistema operativo libre, multitarea y multiusuario, basado en Unix.

los cuales están adaptados específicamente a las características de los procesadores.

- Libvirt: es una interfaz de programación (API) que facilita la comunicación con las máquinas virtuales, permitiendo así su control y gestión. Además de esta función principal, Libvirt proporciona herramientas adicionales como virsh<sup>12</sup>, una línea de comandos, y Virtual Machine Manager, una interfaz gráfica.
- QEMU: este programa tiene como función principal la emulación de ordenadores y la gestión de máquinas virtuales. Linux KVM aprovecha la capacidad de QEMU para emular hardware, mejorando así el rendimiento mediante la paravirtualización. En muchas distribuciones, ambos componentes se incluyen en un solo paquete de instalación.
- Kvmtool: esta es una alternativa a QEMU que requiere menos recursos, aunque aún no ha logrado superar completamente a su contraparte.
- Controladores VirtIO: se trata de controladores especiales diseñados para sistemas invitados, los cuales se ajustan a los requisitos del hipervisor<sup>13</sup> y permiten la paravirtualización<sup>14</sup>. Estos controladores son fundamentales para mejorar el rendimiento y la eficiencia de las máquinas virtuales.

### **2.5.2. Gestión de KVM**

Según (Red Hat, 2022), se puede llevar a cabo la gestión manual de múltiples máquinas virtuales activas en un único equipo, aunque resulta complejo sin una herramienta de administración dedicada. Para abordar esta necesidad, las grandes empresas implementan software de gestión de virtualización que facilita la interacción con los entornos virtuales y el hardware subyacente. Esto simplifica la administración de recursos, mejora el análisis de datos y optimiza las operaciones. Un ejemplo de esto es Red Hat Virtualization, desarrollado específicamente con este propósito por Red Hat.

---

<sup>12</sup> Virsh: Interfaz de línea de comandos para la gestión de máquinas virtuales en entornos de virtualización basados en Libvirt.

<sup>13</sup> Hipervisor: software que permite ejecutar múltiples sistemas operativos en un solo hardware.

<sup>14</sup> Paravirtualización: técnica que mejora el rendimiento al modificar el sistema operativo huésped para colaborar con el hipervisor.

### **2.5.3. Funcionamiento**

Según (Red Hat, 2022), Las KVM convierten Linux en un hipervisor de tipo 1 (servidor dedicado [bare metal]). Todos los hipervisores necesitan algunos elementos del sistema operativo (por ejemplo, el administrador de memoria, el programador de procesos, la stack<sup>15</sup> de entrada o salida [E/S], los controladores de dispositivos, el administrador de seguridad y la stack de red, entre otros) para ejecutar las máquinas virtuales. Las KVM tienen todos estos elementos porque forman parte del kernel de Linux. Cada máquina virtual se implementa como un proceso habitual de Linux, el cual se programa con la herramienta estándar de Linux para este fin, e incluye sistemas virtuales de hardware exclusivos, como la tarjeta de red, el adaptador gráfico, las CPU, la memoria y los discos.

## **2.6. MICROSOFT AZURE**

### **2.6.1. ¿Qué es Microsoft Azure?**

Según (Shahan & Collier, 2016), Microsoft Azure es una plataforma de computación en la nube que ofrece una variedad de servicios, incluyendo Software as a Service (SaaS), Platform as a Service (PaaS) e Infrastructure as a Service (IaaS). Estos servicios están diseñados para proporcionar soluciones tecnológicas y una forma flexible de implementar recursos informáticos.

---

<sup>15</sup> Stack: Estructura de datos informática que sigue el principio LIFO, utilizada para gestionar datos temporalmente durante la ejecución de un programa.

## Figura 11

### Arquitectura Microsoft Azure



*Nota. Arquitectura Microsoft Azure, según (Geekzilla tech, 2021)*

#### 2.6.2. Oferta en la nube

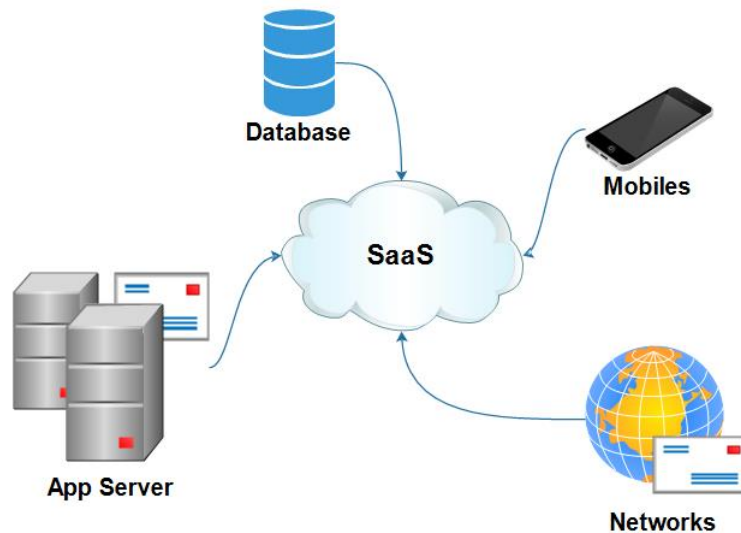
La oferta en la nube se divide comúnmente en tres categorías principales: SaaS, PaaS e IaaS. Sin embargo, a medida que la tecnología en la nube evoluciona, estas distinciones tienden a desdibujarse.

##### 2.6.2.1. SaaS software como servicio

Según (Collier & Shahan, 2016) SaaS, o software como servicio, implica la entrega de software alojado y gestionado centralmente para múltiples clientes. Los usuarios acceden a este software a través de una suscripción y pueden escalar su uso según sus necesidades. Ejemplos típicos de SaaS incluyen Microsoft Office 365, Dropbox, WordPress y Amazon Kindle.

## Figura 12

Principales usos de una SaaS.



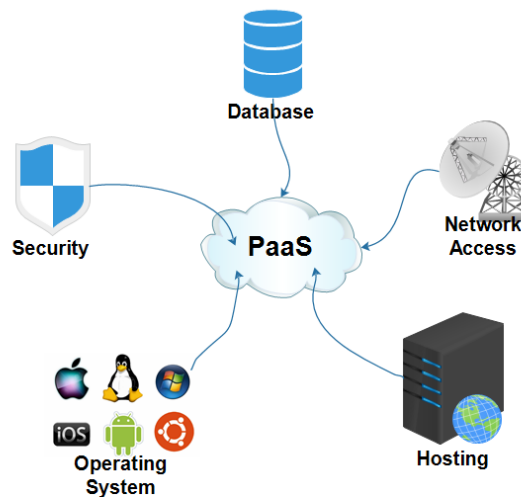
*Nota. Principales usos de una SaaS, Elaboración propia, 2024.*

### 2.6.2.2. PaaS plataforma como servicio

Según (Collier & Shahan, 2016) PaaS o plataforma como servicio, proporciona un entorno de alojamiento de aplicaciones donde los desarrolladores pueden implementar y ejecutar sus aplicaciones sin tener que preocuparse por la gestión de la infraestructura subyacente. Azure ofrece varias ofertas de PaaS, como Azure App Service y Azure Cloud Services, que simplifican la implementación de aplicaciones sin requerir conocimientos profundos sobre la infraestructura subyacente.

**Figura 13**

Principales usos de una PaaS.



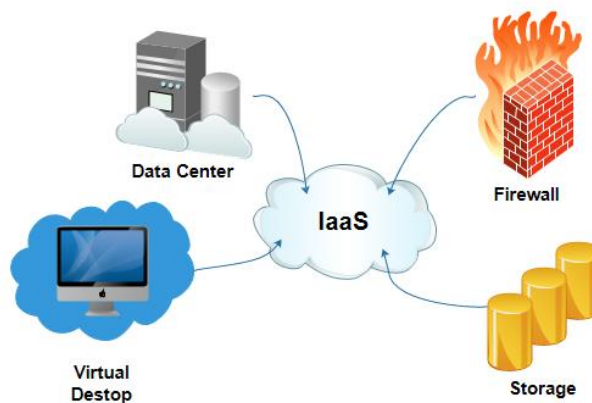
*Nota. Principales usos de una PaaS, Elaboración propia, 2024.*

### 2.6.2.3. IaaS infraestructura como servicio

Según (Collier & Shahan, 2016) IaaS o infraestructura como servicio, implica la entrega de recursos informáticos virtualizados a través de Internet. Los usuarios tienen un control más granular sobre estos recursos en comparación con PaaS, ya que pueden configurar y administrar máquinas virtuales, redes y almacenamiento. Azure Virtual Machines es un ejemplo de oferta IaaS que permite a los usuarios crear y administrar máquinas virtuales en la nube de Azure.

**Figura 14**

Principales usos de una IaaS.



*Nota. Principales usos de una IaaS, Elaboración propia, 2024.*



#### 2.6.2.4. Diferencias entre IaaS, PaaS y SaaS

Figura 15

diferencia entre IaaS, PaaS, SaaS

IaaS	PaaS	SaaS
Recursos de computación, almacenamiento y red.	Entornos de desarrollo y despliegue.	Aplicaciones basadas en la nube.
Acceso a través de un dashboard o API.	Acceso vía web.	Acceso a través de un navegador o una app.
Disponible en un modelo de pago por uso o <i>pay-as-you-go</i> .	Disponible en un modelo de pago por uso.	Disponible gratuitamente, en un modelo de pago por suscripción o compra.
Arquitectos de red y administradores IT.	Desarrolladores de software.	Usuarios finales.

Nota. Diferencia entre IaaS, PaaS y SaaS, según (Grupo Aire, 2023).

#### 2.6.2.5. Principales productos y servicios

A continuación, se describen los principales productos y servicios que ofrece Microsoft Azure.

Azure Virtual Machines, o máquinas virtuales de Azure, es una solución de infraestructura como servicio (IaaS) que nos permite implementar un servidor virtualizado en la nube según nuestras necesidades en cuanto a recursos y programas informáticos. Con este servicio podemos implementar sistemas operativos Linux y Windows, en la cual debemos administrar la seguridad y escalabilidad de las máquinas virtuales. Azure solo se encarga de brindar la infraestructura subyacente, como servidores físicos, redes y almacenamiento, mientras que el usuario es el responsable de configurar y mantener las máquinas virtuales y sus aplicaciones. (Def. Propia, 2024)

- Azure Cosmos DB es una plataforma de gestión de bases de datos versátil que admite múltiples modelos de datos y está diseñada para garantizar una alta disponibilidad, escalabilidad y acceso rápido a los datos. Este servicio se utiliza principalmente en aplicaciones que necesitan aprovechar la inteligencia artificial, el Internet de las cosas (IoT), juegos y otras aplicaciones. (Def. Propia, 2024)
- Azure Kubernetes Service (AKS) es un servicio administrado de Kubernetes ofrecido por Microsoft, el cual nos proporciona la facilidad de implementación y administración de clústeres en la nube. AKS se encarga de las tareas tediosas,

lo que permite a los desarrolladores enfocarse únicamente en el desarrollo del sistema y la ejecución en contenedores. (Def. Propia, 2024)

- Azure AI constituyen una plataforma de inteligencia artificial (IA) basada en la nube que facilita a los desarrolladores la integración de capacidades cognitivas en las aplicaciones, sin requerir habilidades específicas en IA o ciencia de datos. Tanto los Servicios de Azure AI como Azure Machine Learning<sup>16</sup> persiguen el objetivo común de aplicar la inteligencia artificial para optimizar las operaciones empresariales, si bien difieren en sus enfoques y funcionalidades específicas dentro de sus respectivas ofertas. (Def. Propia, 2024)

## **2.7. AMAZON WEB SERVICES**

Según (Roch Moraguez & LovTechnology, 2024), Amazon Web Services (AWS) es una plataforma en la nube que ofrece a las empresas una solución informática flexible y escalable. A través de Internet, permite acceder a una variedad de servicios, como potencia de procesamiento, almacenamiento, bases de datos y redes. Con AWS, las empresas pueden evitar los costos asociados con la construcción y el mantenimiento de su propia infraestructura de TI, así como la necesidad de centros de datos físicos. Además, AWS utiliza un modelo de precios basado en el uso, lo que significa que las empresas solo pagan por los servicios que realmente consumen.

### **2.7.1. Ventajas de AWS (Amazon Web Services)**

Utilizar AWS ofrece múltiples ventajas para las empresas. Primero, AWS proporciona una amplia variedad de servicios, lo que permite a las empresas acceder a numerosos recursos informáticos desde una única plataforma. Segundo, la alta escalabilidad de AWS permite a las empresas ajustar fácilmente sus recursos informáticos conforme a sus necesidades. Tercero, el modelo de precios de pago por uso de AWS asegura que las empresas solo paguen por los servicios que realmente utilizan. Cuarto, AWS garantiza altos niveles de seguridad y cumplimiento, permitiendo a las empresas almacenar sus datos de manera segura en la plataforma. Finalmente,

---

<sup>16</sup> Machine Learning: Campo de la inteligencia artificial que se centra en el desarrollo de algoritmos y modelos que permiten a las computadoras aprender a partir de datos y realizar tareas específicas sin una programación explícita.

AWS proporciona diversas herramientas y servicios que simplifican la gestión de recursos informáticos para las empresas.

### **2.7.2. Arquitectura y Componentes de AWS**

“AWS se basa en un conjunto de servicios y componentes principales que proporcionan a las empresas una amplia gama de recursos informáticos”(Roch Moraguez & LovTechnology, 2024).Estos servicios y componentes incluyen:

- Computación: AWS ofrece una gama de servicios informáticos, como EC2 (Elastic Compute Cloud), que proporciona potencia informática escalable en la nube.
- Almacenamiento: AWS ofrece una gama de servicios de almacenamiento, como S3 (Simple Storage Service), que proporciona almacenamiento escalable y duradero en la nube.
- Bases de datos: AWS ofrece una gama de servicios de bases de datos, como RDS (Relational Database Service), que proporciona servicios de bases de datos administradas en la nube.
- Redes: AWS ofrece una gama de servicios de red, como VPC (Virtual Private Cloud), que proporciona recursos de red aislados en la nube.
- Análisis: AWS ofrece una gama de servicios de análisis, como Redshift, que proporciona almacenamiento de datos y análisis en la nube.
- Seguridad y conformidad: AWS ofrece una gama de servicios de seguridad y conformidad, como IAM (Identity and Access Management), que proporciona un control seguro del acceso a los recursos de AWS.

### **2.8. COMPARATIVA MICROSOFT AZURE VS. AMAZON WEB SERVICES**

Según (EKCIT, 2024), Amazon y Microsoft han logrado destacarse en los últimos informes de Gartner para servicios de infraestructura y plataforma en la nube. Este éxito se debe en gran medida al impulso que ambas compañías han dado a sus herramientas en la nube, Amazon Web Services (AWS) y Microsoft Azure. Además, han ampliado significativamente su oferta de servicios, convirtiéndolos en los más completos del mercado. Sin embargo, la gran cantidad de servicios y terminología puede resultar confusa, por lo que se presenta una tabla con la equivalencia de servicios entre AWS y Azure. Los subcapítulos detallan los principales servicios

disponibles en cada plataforma en la nube. No obstante, debido a la competencia entre ambas compañías, continuamente lanzan nuevos servicios y herramientas que pueden no estar incluidos aquí. En definitiva, tanto Amazon Web Services como Microsoft Azure ofrecen un entorno de trabajo en la nube completo y accesible para cualquier empresa, y la elección entre ellas dependerá de diversos factores más allá de los servicios ofrecidos.

### 2.8.1. Analítica

A continuación, se presenta una tabla comparativa que destaca las capacidades analíticas de Amazon Web Service (AWS) y Microsoft Azure.

**Tabla 2**

*Tabla analítica de comparación.*

AWS	AZURE
Amazon Athena: consulta de datos en Amazon S3 usando SQL.	-
Amazon CloudSearch: servicio de gestión de búsqueda.	Explorador de datos de Azure: servicio de exploración de datos rápido y escalable.
Amazon OpenSearch Service (sucesor to Amazon Elasticsearch Service): búsqueda, visualización y análisis de petabytes de texto y de datos no estructurados.	-
Amazon EMR: ejecutar fácilmente frameworks de Big Data.	-
Amazon FinSpace: analítica para el sector de servicios financieros.	-
Amazon Managed Streaming for Apache Kafka: gestión completa de los servicios de Apache Kafka	-

<p>Amazon QuickSight: servicio de analítica de negocios rápida.</p>	<p>Azure Synapse Analytics: servicio de análisis ilimitado que permite obtener conclusiones rápidamente</p> <p>Azure Analysis Services: motor de análisis de nivel empresarial como servicio.</p>
<p>AWS Glue: integración de datos simple, escalable y sin servidores.</p>	<p>Data factory: integración fácil de datos híbridos a escala empresarial.</p>
<p>AWS Lake formation: construye un data lake seguro en días.</p>	<p>Azure Data Lake Store: funcionalidad de data lake segura y escalable de forma masiva basada en Azure Blob Storage.</p>
<p>-</p>	<p>Azure Purview: solución unificada de gobernanza de datos</p>
<p>-</p>	<p>HDInsights: aprovisionamiento de clústers de Hadoop, Spark, R Server, Hbase y Storm en la nube</p>

*Nota. Tabla comparativa analítica, según (EKCIT, 2024).*

### 2.8.2. Integración

A continuación, se exhibe una tabla comparativa que resalta la integración entre Amazon Web Service (AWS) y Microsoft Azure.

**Tabla 3**

*Tabla de integración*

AWS	AZURE
AWS Step Functions: coordinación para aplicaciones distribuidas.	Logic Apps: automatización del acceso a los datos y su uso en diferentes nubes.
Amazon EventBridge: bus de evento sin servidor para aplicaciones SaaS y servicios de AWS	Service Bus: conexión a través de entornos de nube privada y pública.
Amazon Simple Notification Service (SNS): notificaciones de pub/sub, sms, email y de móvil	Notification Hubs: permite enviar notificaciones push a cualquier plataforma desde cualquier back-end.
AWS AppSync: facilita el desarrollo de API GraphQL al encargarse del trabajo pesado de conectarse de forma segura a fuentes de datos como AWS DynamoDB, Lambda, etc.	API Management: publicación de las APIs propias para desarrolladores, asociados y empleados de forma segura y a escala API Management: publicación de las APIs propias para desarrolladores, asociados y empleados de forma segura y a escala
-	Event Grid: entrega de eventos a gran escala

*Nota. Tabla comparativa de integración, según (EKCIT, 2024).*

### 2.8.3. Servicios Computacionales

A continuación, se presenta una tabla comparativa de los servicios computacionales ofrecidos por AWS y Microsoft Azure. Esta comparación proporciona una visión clara de las similitudes y diferencias clave entre ambas plataformas en términos de servicios computacionales.

**Tabla 4**

*Servicios Computacionales*

AWS	AZURE
Amazon Elastic Compute Cloud (EC2): capacidad informática (servidores virtuales) en la nube, segura y de tamaño modificable.	-
AWS Batch: permite procesar por lotes completamente administrado a cualquier escala.	Azure Batch: permite procesar por lotes desde la nube, completamente administrado y a cualquier escala.
-	Azure Virtual Machines: permite crear máquinas virtuales (VM) Linux y Windows para ejecutar aplicaciones empresariales.
-	Azure Virtual Machine Scale Sets: permite crear y administrar un grupo heterogéneo de hasta miles de máquinas virtuales (VM) con equilibrio de carga.

---

<p>VMware Cloud on AWS: máquina virtual en la nube que permite crear una nube híbrida sin necesidad de adquirir hardware, permitiendo ejecutar todas las cargas de trabajo basadas en vSphere en AWS.</p>	<p>VMware Horizon Cloud Service on Microsoft Azure: servicio de VMware que amplía Azure Virtual Desktop para simplificar la entrega de aplicaciones y escritorios virtuales en la nube y en el entorno local conectando su instancia de Azure a VMware.</p>
---	---

---

*Nota. Tabla comparativa de integración, según (EKCIT, 2024).*

**2.9. VIRTUALBOX**

Según (Oracle and/or its affiliates, 2024),VirtualBox es una aplicación de virtualización multiplataforma. Esto significa que puede instalarse en computadoras existentes basadas en Intel o AMD, ya sea que estén ejecutando sistemas operativos (SO) Windows, macOS, Linux o Oracle Solaris. Además, extiende las capacidades de la computadora actual para que pueda ejecutar múltiples SO dentro de múltiples máquinas virtuales al mismo tiempo.

Por ejemplo, se puede ejecutar Windows y Linux en una Mac, ejecutar Windows Server en un servidor Linux, ejecutar Linux en una PC con Windows, y así sucesivamente, todo junto con las aplicaciones existentes. Es posible instalar y ejecutar tantas máquinas virtuales como se desee, con los únicos límites prácticos siendo el espacio en disco y la memoria.

**2.9.1. Utilidad de la Virtualización con Virtualbox**

Según (Oracle and/or its affiliates, 2024), VirtualBox es una herramienta de virtualización multiplataforma con diversas aplicaciones y beneficios. A continuación, se detalla su utilidad en distintos escenarios:

- Ejecutar Múltiples Sistemas Operativos Simultáneamente: Permite ejecutar varios sistemas operativos al mismo tiempo, lo que posibilita la ejecución de software de un sistema operativo en otro sin necesidad de reiniciar, además de facilitar la instalación de sistemas operativos antiguos.



- Instalaciones de Software Más Sencillas: Facilita la distribución de configuraciones completas dentro de máquinas virtuales, simplificando así la instalación de configuraciones complejas y la ejecución de estas configuraciones.
- Pruebas y Recuperación Ante Desastres: Gestiona y mitiga riesgos con facilidad al permitir la construcción y prueba de servicios en red, facilitar la investigación de problemas de red y sistema operativo, y utilizar snapshots para la experimentación segura y evitar la necesidad de copias de seguridad frecuentes.
- Consolidación de Infraestructuras: Agrupa múltiples máquinas virtuales en pocos hosts potentes, lo que reduce costos de hardware y electricidad, optimiza el uso de recursos de hardware y equilibra cargas de trabajo para maximizar la eficiencia.

## **2.9.2. Terminología Especial**

Según (Oracle and/or its affiliates, 2024), cuando se aborda la virtualización, resulta fundamental familiarizarse con ciertos términos clave, los cuales desempeñan un papel crucial en la comprensión.

### **2.9.2.1. Sistema Operativo Anfitrión (SO anfitrión)**

Según (Oracle and/or its affiliates, 2024), se refiere al sistema operativo instalado en la computadora física donde se ha implementado Oracle VM VirtualBox. Existen versiones de Oracle VM VirtualBox diseñadas para distintos sistemas anfitriones, como Windows, macOS, Linux y Oracle Solaris.

### **2.9.2.2. Sistema Operativo Huésped (SO huésped)**

Según (Oracle and/or its affiliates, 2024), hace referencia al sistema operativo que se ejecuta dentro de la máquina virtual. Aunque Oracle VM VirtualBox teóricamente puede ejecutar cualquier sistema operativo x86, como DOS, Windows, OS/2, FreeBSD y OpenBSD, es importante destacar que se han realizado optimizaciones específicas para garantizar un rendimiento óptimo de los sistemas operativos huéspedes más comunes. Se sugiere consultar los Sistemas Operativos Huéspedes Compatibles.

### **2.9.2.3. Máquina Virtual (VM)**

Según (Oracle and/or its affiliates, 2024), se refiere a un entorno virtual creado específicamente para albergar y ejecutar un sistema operativo huésped. En términos simples, la VM proporciona un espacio aislado donde el sistema operativo huésped puede operar sin interferir con el sistema operativo anfitrión. La interfaz gráfica de usuario de Oracle VM VirtualBox suele representar la máquina virtual como una ventana en el escritorio de la computadora. Sin embargo, esta representación puede variar según la configuración, permitiendo que la VM se muestre en modo de pantalla completa o se acceda de forma remota desde otra computadora.

Desde una perspectiva técnica, Oracle VM VirtualBox administra cada máquina virtual mediante un conjunto de parámetros que definen su configuración y estado. Estos parámetros abarcan aspectos como la asignación de memoria y el número de CPUs virtuales. Para visualizar y modificar estos ajustes, los usuarios pueden utilizar el Administrador de VirtualBox, una interfaz gráfica de usuario intuitiva, o ejecutar comandos a través de la interfaz de línea de comandos utilizando VBoxManage. Este enfoque brinda flexibilidad y control sobre la configuración de la máquina virtual, permitiendo adaptarla a las necesidades específicas de cada usuario.

### **2.9.2.4. Adiciones para Huéspedes**

Según (Oracle and/or its affiliates, 2024), Las Adiciones para Huéspedes se refieren a paquetes de software especiales que vienen incluidos con Oracle VM VirtualBox y que están diseñados para mejorar el rendimiento y ofrecer características adicionales dentro de una máquina virtual. Para obtener más detalles sobre las funcionalidades y beneficios de las Adiciones para Huéspedes, se recomienda revisar la documentación correspondiente.

## **2.10. RAID 5**

Según (Curti et al., 2015), en un arreglo RAID 5 los discos se dividen en bandas para permitir lectura y escritura simultáneas en múltiples discos. Cada banda se divide

en segmentos llamados "chunks"<sup>17</sup>, y se realiza una operación XOR<sup>18</sup> en los N-1 chunks de datos para calcular el chunk de paridad. Esta paridad se distribuye entre los discos de diferentes bandas del arreglo, evitando que un solo disco maneje toda la carga asociada con el almacenamiento de paridad.

Esto mejora el rendimiento general del arreglo al distribuir de manera uniforme la carga entre los discos. Los chunks de paridad permiten al arreglo tolerar la falla de un disco y seguir operando en un modo degradado hasta que se reemplace el disco perdido. RAID 5 representa un equilibrio entre capacidad de almacenamiento, rendimiento y redundancia de datos.

### **2.10.1. Estructura de RAID 5**

Según (Rahman & Novikova Freyre Shavier, 2018), la configuración de los arrays de discos RAID-5 comprende al menos tres discos independientes con la misma capacidad y permanecen operativos en caso de que falle uno de ellos. La capacidad efectiva de un array RAID 5 equivale a una fracción de  $(n - 1) / n$  de la capacidad total de los discos, reservando una porción de  $1/n$  de cada disco para almacenar información redundante o de control, calculada en función de los datos de usuario presentes en los otros discos. Esto facilita la recuperación de datos en caso de que falle un solo disco, aprovechando la información disponible en los demás discos.

Sin embargo, si fallan dos discos o si un segundo disco falla antes de que se haya reemplazado el primero y completado el proceso de replicación de datos en el disco de reemplazo, se producirá la pérdida irreparable de todos los datos del array. Por lo tanto, los arrays RAID 5 representan un equilibrio entre la tolerancia a fallos y la redundancia.

En la Figura a continuación ilustra cómo se distribuyen los bloques de datos de usuario y de control en un array RAID 5 de cuatro discos.

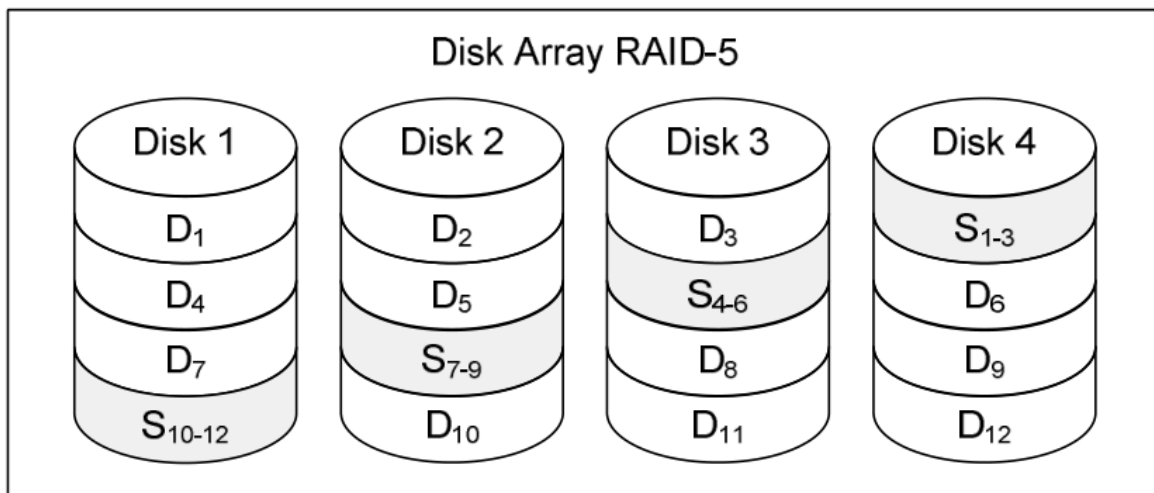
---

<sup>17</sup> Chunks: segmentos de datos distribuidos entre discos en un arreglo RAID

<sup>18</sup> XOR: Operación lógica exclusiva utilizada en arreglos RAID para calcular la paridad de los datos.

**Figura 16**

*Distribución de bloque de datos.*



Nota. *Distribución de bloque de datos, según (Rahman & Novikova Freyre Shavier, 2018).*

### **2.10.2. ¿Cómo funciona el RAID 5?**

Según (Cervera & Wondershare Technology Group Co., 2024), RAID 5 combina dos técnicas para almacenar datos en varios discos: "Paridad" y "Striping de datos". Es fundamental comprender estos dos conceptos, ya que el funcionamiento de RAID 5 se basa en ellos.

- Paridad: La paridad es un valor utilizado para recuperar datos en caso de fallo de almacenamiento. Se calcula mediante una operación XOR (OR exclusivo) sobre los datos a almacenar. En una configuración RAID con 3 discos, solo 2 discos almacenarán la paridad, proporcionando una copia de los datos originales.
- Striping de datos: El striping de datos es una técnica de almacenamiento en la configuración RAID 5. Consiste en dividir los datos en segmentos lógicos consecutivos y almacenar estas partes en diferentes dispositivos de almacenamiento, como SSDs y HDDs. En términos simples, cada segmento de datos se denomina tira, y una colección de varias tiras conforma una franja.

### 2.10.3. Paridad

Según (Ondata International SL., 2024), La paridad en un sistema RAID es un método de redundancia que permite la recuperación de datos en caso de fallo de un disco. Se calcula como la suma (o más comúnmente, mediante una operación XOR) de todos los datos en una matriz de discos. En caso de fallo de un disco, la información de paridad se utiliza para reconstruir los datos perdidos leyendo los datos restantes y comparándolos con el valor de paridad. Esta técnica es utilizada en los niveles de RAID 2, 3, 4 y 5, pero no en RAID 1, donde los datos están completamente duplicados como en un espejo.

**Figura 17**

*Paridad de datos*



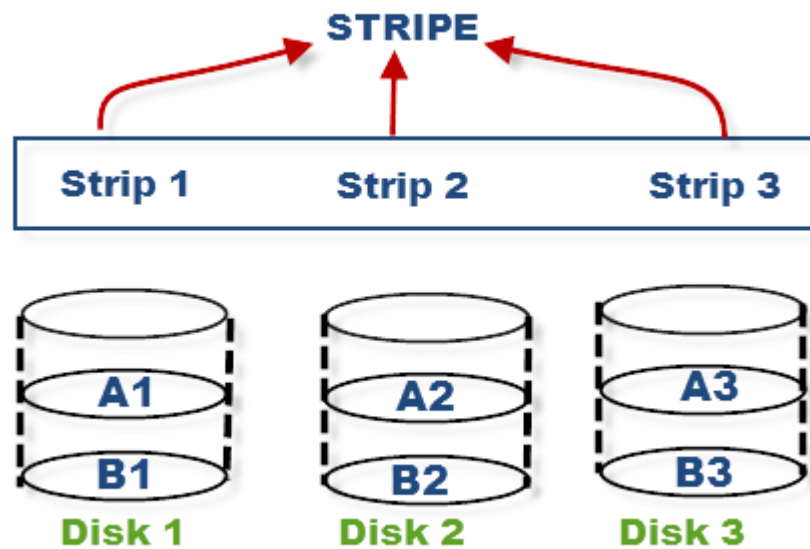
*Nota. Demostración de la paridad, según (Ondata International SL., 2024).*

### 2.10.4. Striping

Según (Demetris & TechEdu, 2024), Striping es una técnica utilizada para mejorar el rendimiento de un sistema de almacenamiento repartiendo los datos entre varios discos. Consiste en dividir un archivo en partes más pequeñas y escribir cada parte en un disco diferente. Esto permite al sistema de almacenamiento leer y escribir datos en paralelo, lo que puede mejorar el rendimiento.

**Figura 18**

*Striping*



*Nota. Diagrama de Striping, según (Storage Tutorials, 2024).*

#### **2.10.4.1. Ventajas de utilizar Striping RAID**

A continuación, se detallan las principales ventajas de utilizar Striping, según (PITS Global Data Recovery Services, 2024):

- Sincronización eficiente de datos en varios discos: asegura una sincronización eficiente de los datos distribuyéndolos cuidadosamente entre varios discos duros de una matriz. Este enfoque organiza los datos de manera similar a una biblioteca, donde cada libro representa una parte de los datos y se coloca en diferentes estantes. Esto permite que el sistema acceda a varias unidades simultáneamente, acelerando tanto la recuperación como el almacenamiento de datos.
- Tolerancia a fallos fundamentales: mejora el rendimiento, sino que también proporciona un nivel básico de tolerancia a fallos. Si una unidad en la matriz falla, los datos en esa unidad pueden perderse, pero las otras unidades conservan sus datos. Esta redundancia básica actúa como una medida de seguridad, garantizando la disponibilidad y la seguridad de los datos durante fallos de hardware.

- Rendimiento turboalimentado: divide los datos en bloques más pequeños, asignando cada bloque a un disco específico. Esto permite el acceso simultáneo a varias unidades durante las operaciones de lectura y escritura, aumentando significativamente las velocidades de estas operaciones. Esta ingeniosa división maximiza la productividad y mejora la eficiencia general del sistema, proporcionando una experiencia de usuario fluida y receptiva.

#### **2.10.5. Modelo de confiabilidad simplificado de discos RAID 5**

Según (Rahman & Novikova Freyre Shavier, 2018), el modelo de confiabilidad simplificado, basado en la cadena de Markov de birth-death (nacimiento-muerte) [3-6], el array de discos RAID-5 se considera como un sistema con  $n$  elementos independientes, que es tolerante al fallo de cualquier elemento individual. El estado inicial del array de discos es aquel en el que todos los  $n$  discos están operativos.

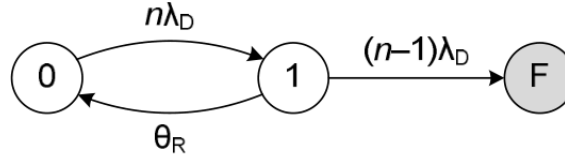
En el modelo de confiabilidad simplificado del array de discos RAID-5, se utiliza el siguiente conjunto de estados y transiciones entre ellos:

- Estado 0: estado normal, donde todos los  $n$  discos del array están operativos y los datos del array están disponibles. Desde este estado, el array de discos puede pasar al estado 1 con una tasa de  $n\lambda_D$  (fallo de cualquier disco).
- Estado 1: estado degradado, donde uno de los discos del array falla, los  $n-1$  discos restantes del array están operativos y los datos del array están disponibles. Desde este estado, el array puede pasar al estado 0 con una tasa de reemplazo del disco fallido y completar la replicación de datos en el disco reemplazado ( $\theta_R$ ), o al estado F con una tasa de  $(n-1)\lambda_D$ .
- Estado F: estado fallido, donde los datos del array de discos no están disponibles y se pierden de manera irreversible.

A continuación, se presenta la siguiente cadena de Markov considerando el conjunto de estados discutidos anteriormente y las transiciones entre ellos (fig. 19).

**Figura 19**

Gráfico de estados



*Nota.* Gráfico de estados en el modelo de confiabilidad simplificado de los arrays de discos RAID-5, según (Rahman & Novikova Freyre Shavier, 2018)

Donde:

$\lambda_d$  - tasa de fallo de los discos.

$\Theta_R$  - tasa de replicación de datos.

Por lo tanto, el sistema de ecuaciones diferenciales de Kolmogorov-Chapman para la cadena de Markov es:

$$\left\{ \begin{array}{l} P_0(0) = 1; \quad P_1(0) = P_F(0) = 0; \\ P_0(t) + P_1(t) + P_F(t) = 1; \\ \frac{dP_0(t)}{dt} = -n\lambda_D P_0(t) + \theta_R P_1(t); \\ \frac{dP_1(t)}{dt} = n\lambda_D P_0(t) - (\theta_R + (n-1)\lambda_D) P_1(t); \\ \frac{dP_F(t)}{dt} = (n-1)\lambda_D P_1(t) \end{array} \right.$$

Según (Rahman & Novikova Freyre Shavier, 2018), se considera que en el modelo de confiabilidad simplificado solo en los estados 0-1 del array de discos RAID 5 está operativo y los datos del usuario están disponibles, se puede derivar la fórmula para calcular el tiempo medio hasta la pérdida de datos del array RAID 5 considerándolo como el tiempo medio de preservar el array de discos en los estados 0-1 y teniendo en cuenta que el estado inicial del array<sup>19</sup> de discos es el estado 0:

<sup>19</sup> Array: estructura de datos que almacena elementos del mismo tipo en posiciones contiguas de memoria.



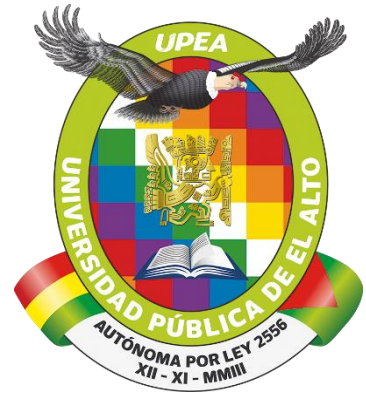
$$T_{R5DL} = \int_0^{\infty} (P_0(t) + P_1(t))dt$$

Finalmente, uno puede obtener la siguiente fórmula simplificada para calcular el tiempo medio hasta la pérdida de datos del array RAID-5 utilizando el análisis matemático:

$$T_{R5DL} = \frac{\theta_R + (2n - 1)\lambda_D}{n(n - 1)\lambda_D^2}$$

Debe mencionarse que en el modelo de confiabilidad simplificado se utilizan las siguientes suposiciones y simplificaciones importantes:

- Las tasas de fallo de los discos son iguales para los estados normales y degradados del array de discos.
- Se considera que el tiempo de reemplazo del disco fallido es insignificante (varios segundos en caso de utilizar discos de repuesto en caliente) en comparación con la replicación de datos (decenas de horas).
- Se ignora la probabilidad de errores de lectura en los  $n - 1$  discos restantes durante la replicación de datos en el disco reemplazado y la subsiguiente falla del proceso de replicación de datos.



# **CAPITULO III**

## **DISEÑO**

### **METODOLOGICO**



En este capítulo se describe la metodología empleada para abordar el problema de investigación y alcanzar los objetivos planteados. Se detallan el tipo de investigación, el diseño, las variables, el ambiente de la investigación, la metodología específica y las herramientas utilizadas.

### **3.1. TIPO DE INVESTIGACIÓN**

El tipo de investigación del presente estudio es aplicativo y experimental. Está orientado a la implementación y evaluación de la configuración RAID 5 en servidor Proxmox utilizando Azure como entorno de prueba. Este tipo de investigación permite realizar pruebas controladas para determinar el desempeño y la seguridad de los datos bajo diversas configuraciones.

### **3.2. NIVEL DE INVESTIGACIÓN**

El nivel de investigación de este estudio se define como aplicativo y experimental. Esta combinación de enfoques permite una exploración detallada de las implementaciones RAID 5 en servidores Proxmox, así como la evaluación de su rendimiento y eficacia en un entorno controlado.

#### **3.1.1. Aplicativo**

En el nivel aplicativo, se proporcionará una descripción exhaustiva de cada aspecto del proceso de implementación de las configuraciones RAID 5 dentro de lo aplicativo.

Esto incluirá detalles sobre la selección de hardware, configuración de software, y cualquier ajuste específico realizado para cada tipo de RAID 5. Además, se realizará un análisis completo del contexto y los parámetros del entorno de prueba, como la carga de trabajo simulada y las métricas de rendimiento esperadas.

#### **3.1.2. Experimental**

En el nivel experimental, se llevarán a cabo pruebas controladas en el entorno descrito anteriormente para evaluar el rendimiento y la eficacia de cada configuración RAID 5. Durante esta fase, se manipularán variables específicas, como el tamaño de los discos, el nivel de paridad, o la distribución de datos, para observar su impacto en el rendimiento del sistema. Los resultados de estas pruebas se registrarán y analizarán meticulosamente para identificar patrones, tendencias y áreas de mejora.

Este enfoque de investigación permitirá obtener una comprensión completa del comportamiento de las implementaciones RAID 5 en servidores Proxmox, proporcionando información valiosa para la toma de decisiones en entornos de almacenamiento de datos críticos.

### **3.3. DISEÑO DE INVESTIGACION**

El diseño de investigación sigue un esquema experimental en el que se manipulan la configuración.

#### **3.3.1. Variables de la Investigación**

- **Variable Independiente:** Implementación de RAID 5 en servidores Proxmox
- **Variable Dependiente:** Riesgo de pérdida de datos.

#### **3.3.2. Ambiente De Investigación**

El estudio se lleva a cabo en un entorno de virtualización utilizando Oracle VM Virtualbox, que proporciona la infraestructura para implementar y probar configuraciones RAID 5 en servidores Proxmox.

Este ambiente permitirá realizar pruebas en un entorno controlado, minimizando variables externas.

#### **3.3.3. DESCRIPCIÓN DE LA METODOLOGÍA A USAR**

En este estudio se empleará una metodología experimental para evaluar las configuraciones RAID en servidores Proxmox. Se seguirán los siguientes pasos:

- Selección de hardware: Identificación y configuración de servidores compatibles con RAID 5.
- Configuración de software: Instalación y configuración de Proxmox VE en los servidores seleccionados.
- Implementación de RAID 5: Configuración de RAID 5 en los servidores Proxmox utilizando la infraestructura de Azure.
- Pruebas de rendimiento: Realización de pruebas controladas para evaluar el desempeño y la seguridad de los datos bajo diversas configuraciones RAID 5.
- Análisis de resultados: Recopilación y análisis de los datos obtenidos durante las pruebas para identificar patrones, tendencias y áreas de mejora.

### **3.3.4. Tratamiento De Datos**

Los datos recolectados durante las pruebas se procesan y analizan utilizando técnicas estadísticas adecuadas para evaluar el rendimiento y la eficacia de la configuración RAID 5. Esto incluye:

- **Recopilación de datos:** Uso de herramientas de monitoreo y pruebas para registrar métricas de rendimiento y eficiencia.
- **Análisis de datos:** Aplicación de métodos estadísticos para interpretar los resultados y comparar el desempeño de las configuraciones RAID 5, para la cual se utiliza Python para el análisis de datos aprovechando librerías como “pandas” para la manipulación de datos, “numpy” para los cálculos numéricos, y “scipy” para el análisis estadístico.
- **Presentación de resultados:** Elaboración de informes detallados con gráficos que resumen los hallazgos, para ello emplea biblioteca de Python como “matplotlib” para la visualización de datos, facilitando la comparación de datos obtenidos.

### **3.4. HERRAMIENTAS A USAR**

Para la presente tesis de grado, se utilizará Oracle VM VirtualBox, una herramienta de virtualización que permite ejecutar múltiples sistemas operativos en una sola máquina física. Oracle VM VirtualBox ofrece un control detallado del entorno informático, facilitando la implementación y gestión de servidores virtuales configurados con Proxmox y RAID 5. Esta plataforma proporcionará el entorno necesario para realizar pruebas de rendimiento y resiliencia, asegurando que los datos y aplicaciones críticas operen de manera eficiente y segura.

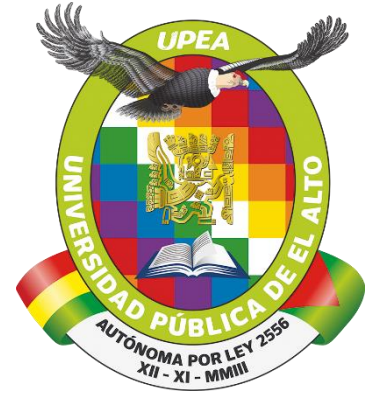
#### **3.4.1. Técnicas de Investigación e Instrumentos**

En esta sección, se describen las técnicas específicas utilizadas para llevar a cabo la investigación experimental sobre la implementación de distintos tipos de RAID 5 en servidores Proxmox. Las técnicas seleccionadas se fundamentan en la metodología propuesta y se alinean con los objetivos y requerimientos del estudio.

### **3.5. MÉTRICAS DE CALIDAD**

Las métricas de calidad se definen utilizando las siguientes normas internacionales:

- ISO/IEC 25000, esta norma establece un marco para la evaluación de la calidad del software, centrándose en características como funcionalidad, confiabilidad, usabilidad, eficiencia, mantenibilidad y portabilidad.
- ISO/IEC 27001 (Seguridad de la Información), esta norma proporciona un marco para la gestión de la seguridad de la información, enfocándose en la prevención de pérdida de datos y la protección de la información.



# **CAPITULO IV**

# **PRUEBAS Y**

# **RESULTADOS**



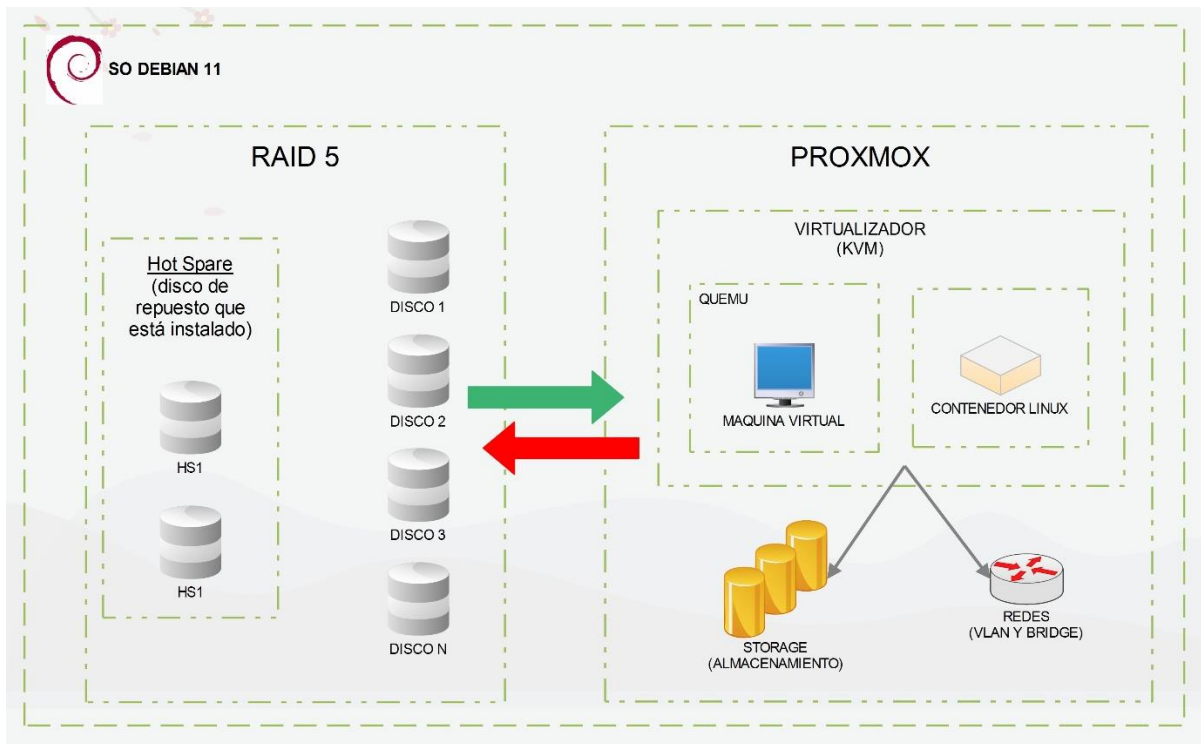
#### 4.1. PRESENTACION DEL MODELO

En esta tesis, se propone un modelo para la reducción del riesgo de pérdida de datos en servidores Proxmox mediante el uso de RAID 5. Utilizando Debian 11 como sistema operativo base, se configurará un entorno virtual con VirtualBox para implementar y validar el modelo propuesto.

La configuración incluirá la utilización de MDADM para establecer RAID 5, sobre la cual se instalará Proxmox, asegurando una gestión eficiente de la infraestructura virtual. Este modelo busca mejorar la seguridad y resiliencia de los datos, optimizando el uso de recursos y facilitando la recuperación ante desastres y la consolidación de infraestructuras.

**Figura 20**

*Modelo para la reducción del riesgo de pérdida de datos*



*Nota. Modelo Para La Reducción Del Riesgo De Pérdida De Datos En Servidores Proxmox A Través De Raid 5 (Elaboración Propia).*



## **4.2. DESARROLLO DEL MODELO**

En esta sección se detalla la implementación práctica del modelo propuesto para abordar los objetivos de la investigación. Se describen los pasos seguidos para diseñar e implementar el modelo, destacando las herramientas y técnicas utilizadas en su desarrollo.

Esta sección proporciona una visión general del proceso de construcción del modelo, estableciendo las bases para el análisis detallado que se llevará a cabo más adelante en cuanto a su configuración y ajuste.

### **4.2.1. Preparación Preliminar para la Implementación del Modelo**

Esta sección se centra en los pasos necesarios para configurar una máquina virtual en VirtualBox, un entorno clave para el desarrollo y la implementación del modelo propuesto. Antes de comenzar con la construcción del modelo en sí, es esencial establecer un entorno de trabajo adecuado que permita la ejecución eficiente del software y la realización de pruebas.

#### **4.2.1.1. Creación de máquina**

Para la presente crearemos una máquina la cual nos servirá como entorno de prueba para el desarrollo y prueba de nuestro modelo.

En la figura 21 se muestra un resumen de la configuración de nuestra máquina virtual contemplando requerimientos básicos.

**Figura 21**

*Configuración básica de máquina virtual*



*Nota. Configuración básica de la máquina(Elaboración Propia).*

Una vez creada la máquina virtual y montada la ISO de Debian 11, procedemos a instalar el sistema en la máquina virtual. A continuación, se describen los pasos para instalar Debian 11:

1. Inicio de la instalación
  - Arranca la máquina virtual y selecciona "Install" en el menú de arranque de Debian.
2. Configuración básica

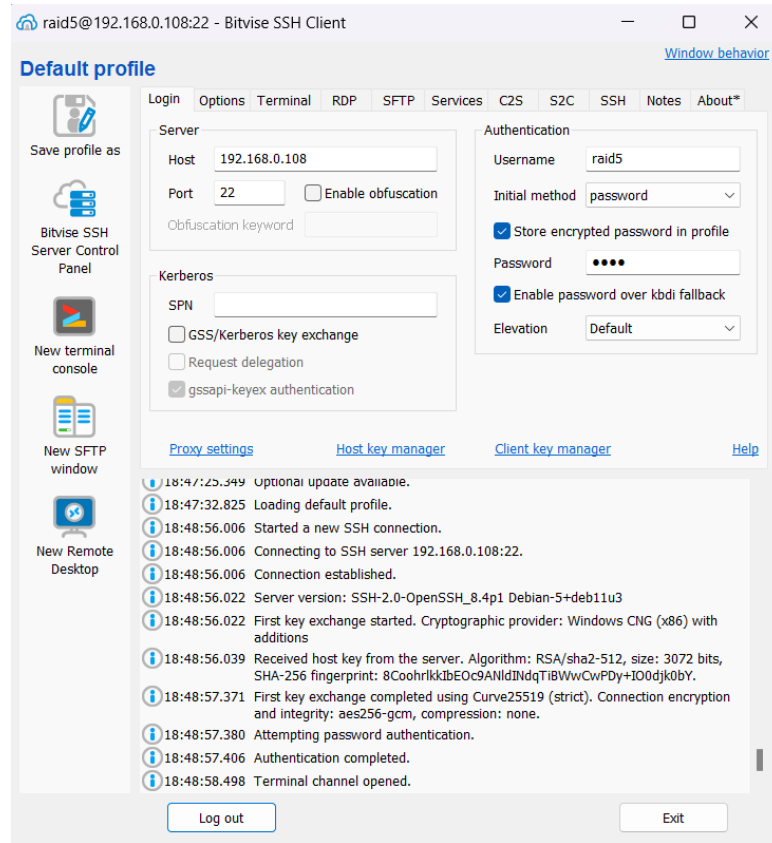
- Configura el idioma, la ubicación y el teclado.
  - Configura la red (DHCP o manual).
  - Establece el nombre del host y el dominio.
3. Configuración de usuarios y contraseñas
    - Establece la contraseña del usuario root.
    - Crea un usuario normal y establece su contraseña.
  4. Particionado del disco
    - Selecciona el disco virtual para la instalación.
    - Elige el método de particionado (guiado o manual) y crea las particiones necesarias.
  5. Instalación del sistema base
    - Selecciona los mirrors de Debian para actualizaciones.
    - Opta por instalar solo el "sistema base" sin entorno de escritorio.
  6. Finalización
    - Instala el cargador de arranque GRUB.
    - Completa la instalación y reinicia la máquina virtual.

Una vez finalizado la instalación de Debían 11 continuamos con la implementación del modelo.

#### **4.2.1.2. Implementación del Modelo**

Para iniciar con la implementación del modelo iniciamos conectándonos a nuestra máquina virtual a través de SSH, para la cual utilizaremos la aplicación “Bitwise Xterm” la cual nos permitirá conectar a través de SSH.

**Figura 22**  
*Conexión SSH*

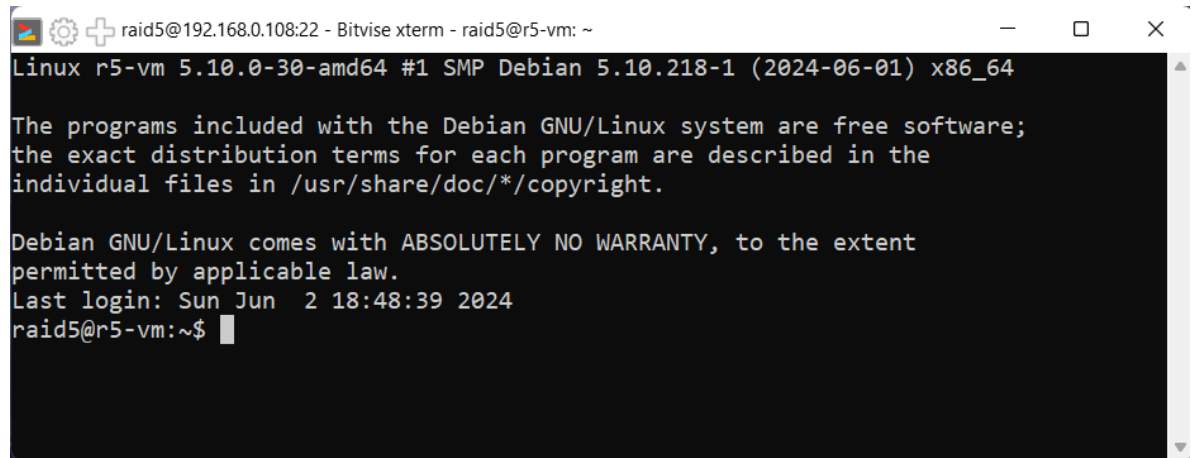


*Nota. Conexión SSH con la aplicación Bitvise Xterm hacia la máquina virtual (Elaboración Propia).*

Una vez configurada la conexión procedemos a conectarnos donde se nos presentara la siguiente pantalla, donde podremos iniciar con la implementación de nuestro modelo.

### Figura 23

#### Terminal Bitvise



```
raid5@192.168.0.108:22 - Bitvise xterm - raid5@r5-vm: ~
Linux r5-vm 5.10.0-30-amd64 #1 SMP Debian 5.10.218-1 (2024-06-01) x86_64

The programs included with the Debian GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

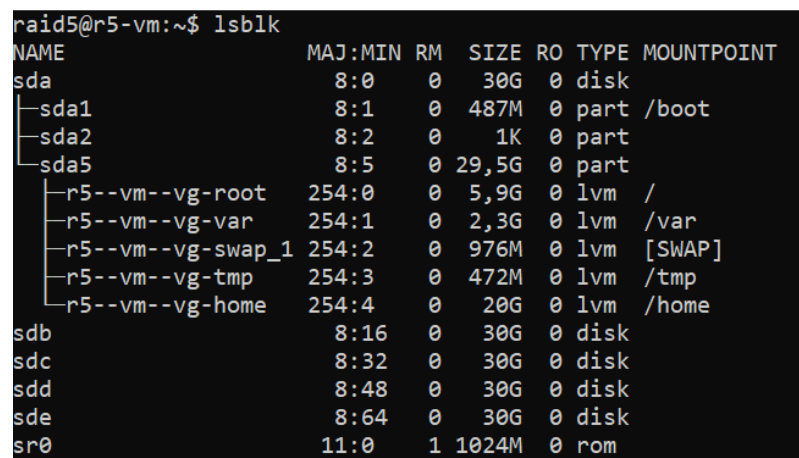
Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
Last login: Sun Jun  2 18:48:39 2024
raid5@r5-vm:~$
```

*Nota. Terminal de la aplicación Bitvise, una vez realizado la conexión (Elaboración Propia).*

Una ingresado por SSH, en la aplicación procedemos a iniciar con la implementación del modelo a través de la terminal. Iniciando con el listado de los discos a través del comando “lsblk”.

### Figura 24

#### Listado de discos



```
raid5@r5-vm:~$ lsblk
NAME                                MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sda                                  8:0    0   30G  0 disk
├─sda1                               8:1    0   487M  0 part /boot
├─sda2                               8:2    0     1K  0 part
└─sda5                               8:5    0  29,5G  0 part
   ├─r5--vm--vg-root                 254:0    0   5,9G  0 lvm  /
   ├─r5--vm--vg-var                   254:1    0   2,3G  0 lvm  /var
   ├─r5--vm--vg-swap_1                254:2    0   976M  0 lvm  [SWAP]
   ├─r5--vm--vg-tmp                   254:3    0   472M  0 lvm  /tmp
   └─r5--vm--vg-home                  254:4    0    20G  0 lvm  /home
sdb                                  8:16    0   30G  0 disk
sdc                                  8:32    0   30G  0 disk
sdd                                  8:48    0   30G  0 disk
sde                                  8:64    0   30G  0 disk
sr0                                  11:0    1 1024M  0 rom
```

*Nota. Listado de discos en la máquina virtual a través del comando “lsblk” (Elaboración Propia).*

En primera instancia, para crear las particiones en cada disco, nos apoyamos en la aplicación “fdisk” que viene integrada con Debian 11 para la gestión de particiones en disco que operan dentro del sistema operativo Debian 11. A continuación se describen las configuraciones a seguir para el particionado de los discos disponibles (sda, sdb, sdc, sdd y sde), para lo cual se usará como ejemplo al disco sdb.

Iniciamos ingresando el comando “sudo fdisk /dev/sdb”, una vez ingresado el comando en consola, apreciaremos la bienvenida de la aplicación donde escogemos la opción m para poder listar las opciones con las que cuenta “fdisk”.

## Figura 25

### Uso de “fdisk”

```
raid5@r5-vm:~$ sudo fdisk /dev/sdb
Bienvenido a fdisk (util-linux 2.36.1).
Los cambios solo permanecerán en la memoria, hasta que decida escribirlos.
Tenga cuidado antes de utilizar la orden de escritura.

El dispositivo no contiene una tabla de particiones reconocida.
Se ha creado una nueva etiqueta de disco DOS con el identificador de disco 0x114821bd.

Orden (m para obtener ayuda):
```

*Nota. Uso de “fdisk”, la cual nos permite iniciar con el particionado del disco (Elaboracion Propia).*

Al aplicar la opción m, se muestran las siguientes opciones como se refleja en la figura 26.

## Figura 26

Listado de opciones de “fdisk”

```
Command (m for help): m
Help:
DOS (MBR)
 a  toggle a bootable flag
 b  edit nested BSD disklabel
 c  toggle the dos compatibility flag

Generic
 d  delete a partition
 F  list free unpartitioned space
 l  list known partition types
 n  add a new partition
 p  print the partition table
 t  change a partition type
 v  verify the partition table
 i  print information about a partition

Misc
 m  print this menu
 u  change display/entry units
 x  extra functionality (experts only)

Script
 I  load disk layout from sfdisk script file
 O  dump disk layout to sfdisk script file

Save & Exit
 w  write table to disk and exit
 q  quit without saving changes

Create a new label
 g  create a new empty GPT partition table
 G  create a new empty SGI (IRIX) partition table
 o  create a new empty DOS partition table
 s  create a new empty Sun partition table
```

*Nota. Listado de opciones de “fdisk” posterior al ingreso de la opción “m” (Elaboracion Propia).*

A continuación, escogeremos la opción n la cual nos permite “añadir una nueva partición”. Posteriormente, nos pedirá que escojamos una de las opciones que nos presenta, de las cuales escogeremos la opción p denotando que la opción que escogemos es la partición primaria como se muestra en la siguiente captura.

## Figura 27

### *Añadir nueva particion*

```
Command (m for help): n
Partition type
  p   primary (0 primary, 0 extended, 4 free)
  e   extended (container for logical partitions)
Select (default p): p
```

*Nota. Añadir nueva particion, en la cual nos da ha ecoger el tipo de particion (Elaboracion Propia).*

Una vez seleccionado el tipo de partición, se muestran las siguientes opciones como se aprecia en la siguiente captura de pantalla. Ingresamos la tecla Enter, denotando que todas las opciones se aplican por defecto. Una vez terminado de aplicar las opciones, nos muestra el mensaje “Created a new partition 1 of type 'Linux' and of size XX GiB.”, el cual nos indica que la partición fue creada de manera correcta, mostrando a su vez el tamaño del disco.

Posteriormente, volveremos a listar las opciones que nos ofrece la aplicación en la cual escogeremos la opcion “t” la cual nos permitirá cambiar el tipo de particionado de “Linux” a “Linux raid autodetect” a través de las subopciones que nos presentará. Para poder listar las opciones, ingresaremos la letra l como nos pide en la captura.

## Figura 28

### *Opcion “t (Tipo de array)”*

```
Command (m for help): t
Selected partition 1
Hex code (type l to list all codes): 1
```

*Nota. Opcion “t (Tipo de array)” (Elaboracion Propia).*



Una vez ingresada en la opción “t” procedemos a listar los tipos de array con la letra “l”, donde nos lista las siguientes opciones:

**Figura 29**

*Listado de tipos de array*

```

00 Empty                24 NEC DOS              81 Minix / old Lin    bf Solaris
01 FAT12                27 Hidden NTFS Win    82 Linux swap / So    c1 DRDOS/sec (FAT-
02 XENIX root          39 Plan 9              83 Linux               c4 DRDOS/sec (FAT-
03 XENIX usr           3c PartitionMagic     84 OS/2 hidden or     c6 DRDOS/sec (FAT-
04 FAT16 <32M         40 Venix 80286        85 Linux extended     c7 Syrix
05 Extended            41 PPC PReP Boot      86 NTFS volume set    da Non-FS data
06 FAT16               42 SFS                 87 NTFS volume set    db CP/M / CTOS / .
07 HPFS/NTFS/exFAT    4d QNX4.x             88 Linux plaintext    de Dell Utility
08 AIX                 4e QNX4.x 2nd part    8e Linux LVM          df BootIt
09 AIX bootable       4f QNX4.x 3rd part    93 Amoeba             e1 DOS access
0a OS/2 Boot Manag    50 OnTrack DM         94 Amoeba BBT         e3 DOS R/O
0b W95 FAT32          51 OnTrack DM6 Aux   9f BSD/OS            e4 SpeedStor
0c W95 FAT32 (LBA)    52 CP/M               a0 IBM Thinkpad hi   ea Linux extended
0e W95 FAT16 (LBA)    53 OnTrack DM6 Aux   a5 FreeBSD           eb BeOS fs
0f W95 Ext'd (LBA)    54 OnTrackDM6        a6 OpenBSD           ee GPT
10 OPUS               55 EZ-Drive           a7 NeXTSTEP          ef EFI (FAT-12/16/
11 Hidden FAT12       56 Golden Bow         a8 Darwin UFS         f0 Linux/PA-RISC b
12 Compaq diagnost    5c Priam Edisk        a9 NetBSD             f1 SpeedStor
14 Hidden FAT16 <3    61 SpeedStor         ab Darwin boot       f4 SpeedStor
16 Hidden FAT16       63 GNU HURD or Sys   af HFS / HFS+        f2 DOS secondary
17 Hidden HPFS/NTF    64 Novell Netware     b7 BSDI fs           fb VMware VMFS
18 AST SmartSleep     65 Novell Netware     b8 BSDI swap         fc VMware VMKCORE
1b Hidden W95 FAT3    70 DiskSecure Mult   bb Boot Wizard hid   fd Linux raid auto
1c Hidden W95 FAT3    75 PC/IX              bc Acronis FAT32 L   fe LANstep
1e Hidden W95 FAT1    80 Old Minix          be Solaris boot      ff BBT

Aliases:
  linux      - 83
  swap       - 82
  extended   - 05
  uefi       - EF
  raid       - FD
  lvm        - 8E
  linuxex    - 85

```

*Nota.* Listado de tipos de array (Elaboracion Propia).

Procedemos a escoger la opción fd, la cual representa la autodetección de RAID dentro de Linux. En la siguiente figura apreciamos que se ingresó la opción y que posteriormente nos indica en el mensaje “Changed type of partition 'Linux' to 'Linux raid autodetect'.” que el cambio fue con éxito.

**Figura 30**

*Selección de tipo “Linux raid autodectetable”*

```

Hex code or alias (type L to list all): fd
Changed type of partition 'Linux raid autodetect' to 'Linux raid autodetect'.

```

*Nota.* Selección de tipo “Linux raid autodectetable” la cual nos permite la facil implementacio de RAID (Elaboracion Propia).

Una vez concluido el paso anterior, procedemos a grabar las configuraciones que se realizaron al disco. Para ello, volvemos a listar las opciones con la letra “m” la cual nos permite visualizar las opciones en la cual escogeremos la opción que se encuentra dentro de “Save & Exit”, la opción w, la cual nos permite grabar las configuraciones realizadas en la tabla de particiones del disco y salir. Escogemos la opción w, posteriormente nos saldrá un mensaje donde nos indica que se han realizado cambios en la tabla de particiones del disco y que el sistema operativo está actualizando la información para reflejar estos cambios.

### **Figura 31**

*Selección de opción “w”*

```
Command (m for help): w
The partition table has been altered.
Calling ioctl() to re-read partition table.
Syncing disks.
```

*Nota.* Selección de opción “w” la cual nos permite grabar en la tabla de particiones del disco todas las configuraciones realizadas (Elaboración Propia).

Todas las configuraciones anteriores descritas, se aplicarán a los discos disponibles con los que se cuenta en la máquina virtual o servidor, para verificar la partición correcta en los discos usaremos el comando “lsblk” la cual nos aseguraremos de que en el nombre de disco tenga una anidación con el mismo nombre del disco pero aumentado el número 1.

**Figura 32**

*Verificación del particionado de los discos*

```
raid5@r5-vm:~$ lsblk
NAME                                MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sda                                  8:0     0   30G  0 disk
├─sda1                               8:1     0  487M  0 part /boot
├─sda2                               8:2     0    1K  0 part
└─sda5                               8:5     0  29,5G  0 part
   ├─r5--vm--vg-root                 254:0   0   5,9G  0 lvm  /
   ├─r5--vm--vg-var                 254:1   0   2,3G  0 lvm  /var
   ├─r5--vm--vg-swap_1              254:2   0   976M  0 lvm  [SWAP]
   ├─r5--vm--vg-tmp                 254:3   0   472M  0 lvm  /tmp
   └─r5--vm--vg-home                 254:4   0    20G  0 lvm  /home
sdb                                  8:16    0   30G  0 disk
└─sdb1                              8:17    0   30G  0 part
sdc                                  8:32    0   30G  0 disk
└─sdc1                              8:33    0   30G  0 part
sdd                                  8:48    0   30G  0 disk
└─sdd1                              8:49    0   30G  0 part
sde                                  8:64    0   30G  0 disk
└─sde1                              8:65    0   30G  0 part
sr0                                  11:0    1 1024M  0 rom
```

*Nota. Verificación del particionado de los discos, donde los disco con particion correcta tendran anidaciones (Elaboracion Propia).*

Para poder continuar con la implementación del modelo iniciamos con la creación del RAID 5 en Debian 11. En la cual usaremos “mdadm” la cual nos permite gestionar el conjunto de discos en configuración RAID. Con el siguiente comando se creará el RAID 5 definiendo así los discos a emplear y los discos de reserva definiendo así con el comando “sudo mdadm --create --verbose /dev/md0 --level=5 --raid-devices=3 /dev/sdb1 /dev/sdc1 /dev/sdd1 --spare-devices=1 /dev/sde1”.

A continuación se detalla el comando:

- sudo mdadm: Este comando se ejecuta con privilegios de superusuario para administrar dispositivos RAID mediante mdadm.
- --create: Indica que se creará un nuevo conjunto de RAID.
- --verbose: Solicita que se muestren mensajes detallados durante el proceso de creación del conjunto de RAID.
- /dev/md0: Especifica el nombre del dispositivo RAID que se creará. En este caso, se utilizará /dev/md0.
- --level=5: Establece el nivel del RAID en 5, lo que significa que se utilizará RAID 5.
- --raid-devices=3: Indica que se utilizarán tres dispositivos de datos para el conjunto de RAID.

- /dev/sdb1 /dev/sdc1 /dev/sdd1: Son los dispositivos de datos que se utilizarán para construir el conjunto de RAID. Estos discos deben tener particiones (en este caso, sdb1, sdc1 y sdd1).
- --spare-devices=1: Especifica que se utilizará un disco de reserva para el conjunto de RAID. Este disco se utilizará automáticamente si uno de los discos de datos falla.
- /dev/sde1: Es el disco de reserva que se utilizará en caso de que falle uno de los discos de datos. Este disco también debe tener una partición (en este caso, /dev/sde1).

### Figura 33

#### Creación de RAID 5

```
raid5@r5-vm:~$ sudo mdadm --create --verbose /dev/md0 --level=5 --raid-devices=3 /dev/sdb1 /dev/sdc1 /dev/sdd1 --spare-devices=1 /dev/sde1
mdadm: layout defaults to left-symmetric
mdadm: layout defaults to left-symmetric
mdadm: chunk size defaults to 512K
mdadm: size set to 31438848K
mdadm: Defaulting to version 1.2 metadata
mdadm: array /dev/md0 started.
```

*Nota. Creación de RAID 5 usando el mdadm (Elaboración Propia).*

Una vez creado el RAID 5 procedemos a formatear con un sistema de archivos en específico usando el comando “sudo mkfs.ext4 /dev/md0”.

### Figura 34

#### Formateo de sistema de archivo en RAID 5

```
raid5@r5-vm:~$ sudo mkfs.ext4 /dev/md0
mke2fs 1.46.2 (28-Feb-2021)
Creating filesystem with 15719424 4k blocks and 3932160 inodes
Filesystem UUID: 8c1bd650-05a3-4007-a28c-8dbbe83d5662
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
    4096000, 7962624, 11239424

Allocating group tables: done
Writing inode tables: done
Creating journal (65536 blocks): done
Writing superblocks and filesystem accounting information: done
```

*Nota. Formateo de sistema de archivo en RAID 5 (Elaboración Propia).*

Ahora realizaremos el montaje del sistema de archivos creado, donde primero crearemos una carpeta en “/media” denominada “raid”. Posteriormente realizaremos el montaje de RAID 5 en la carpeta con el comando “sudo mount /dev/md0 /media/raid”

**Figura 35**

*Montaje de RAID 5*

```
raid5@r5-vm:~$ sudo mkdir /media/raid
raid5@r5-vm:~$ lsblk
NAME                                MAJ:MIN RM  SIZE RO TYPE  MOUNTPOINT
sda                                  8:0     0   30G  0 disk
├─sda1                               8:1     0  487M  0 part  /boot
├─sda2                               8:2     0    1K  0 part
└─sda5                               8:5     0  29,5G  0 part
   ├─r5--vm--vg-root                 254:0     0   5,9G  0 lvm    /
   ├─r5--vm--vg-var                  254:1     0   2,3G  0 lvm    /var
   ├─r5--vm--vg-swap_1               254:2     0  976M  0 lvm    [SWAP]
   ├─r5--vm--vg-tmp                  254:3     0  472M  0 lvm    /tmp
   └─r5--vm--vg-home                 254:4     0   20G  0 lvm    /home
sdb                                  8:16     0   30G  0 disk
├─sdb1                               8:17     0   30G  0 part
└─md0                               9:0     0   60G  0 raid5
sdc                                  8:32     0   30G  0 disk
├─sdc1                               8:33     0   30G  0 part
└─md0                               9:0     0   60G  0 raid5
sdd                                  8:48     0   30G  0 disk
├─sdd1                               8:49     0   30G  0 part
└─md0                               9:0     0   60G  0 raid5
sde                                  8:64     0   30G  0 disk
├─sde1                               8:65     0   30G  0 part
└─md0                               9:0     0   60G  0 raid5
sr0                                  11:0     1 1024M  0 rom
raid5@r5-vm:~$ sudo mount /dev/md0 /media/raid
```

*Nota. Montaje de RAID 5 en el sistema operativo (Elaboracion Propia).*

Una vez realizado el montaje realizaremos la configuración para el montaje automático, para que el array RAID se monte automáticamente cada vez que se inicie el sistema, agregando una entrada al archivo /etc/fstab. Para ello ejecutaremos el comando “lsblk -f” para obtener el UUID del sistema de archivos y luego agregar una entrada en /etc/fstab con ese UUID.

**Figura 36**

*Montaje de RAID 5 en “fstab”*

```
# /etc/fstab: static file system information.
#
# Use 'blkid' to print the universally unique identifier for a
# device; this may be used with UUID= as a more robust way to name devices
# that works even if disks are added and removed. See fstab(5).
#
# systemd generates mount units based on this file, see systemd.mount(5).
# Please run 'systemctl daemon-reload' after making changes here.
#
# <file system> <mount point> <type> <options> <dump> <pass>
/dev/mapper/r5--vm--vg-root / ext4 errors=remount-ro 0 1
# /boot was on /dev/sda1 during installation
UUID=08a95d55-b4e3-4ff9-860d-0b2c3503971e /boot ext2 defaults 0 2
/dev/mapper/r5--vm--vg-home /home ext4 defaults 0 2
/dev/mapper/r5--vm--vg-tmp /tmp ext4 defaults 0 2
/dev/mapper/r5--vm--vg-var /var ext4 defaults 0 2
/dev/mapper/r5--vm--vg-swap_1 none swap sw 0 0
/dev/sr0 /media/cdrom0 udf,iso9660 user,noauto 0 0

#MONTAJE RAID 5
UUID=8c1bd650-05a3-4007-a28c-8dbbe83d5662 /media/raid ext4 0 0
```

^G Ayuda    ^O Guardar    ^W Buscar    ^K Cortar    ^T Ejecutar    ^C Ubicación    M-U Deshacer  
^X Salir    ^R Leer fich.    ^\ Reemplazar    ^U Pegar    ^J Justificar    ^ Ir a línea    M-E Rehacer

*Nota. Montaje de RAID 5 en “fstab”, la cual nos permitirá el montaje automático en cada encendido del servidor (Elaboración Propia).*

Para que los cambios surtan efecto, reiniciar la máquina virtual, posteriormente procedemos a realizar las revisiones correspondientes al montaje de RAID 5 con el comando “sudo cat /proc/mdstat” donde el nombre de montaje ha cambiado de “md0” a “md127” la cual no indica que el montaje automático ya se grabó correctamente en el servidor.

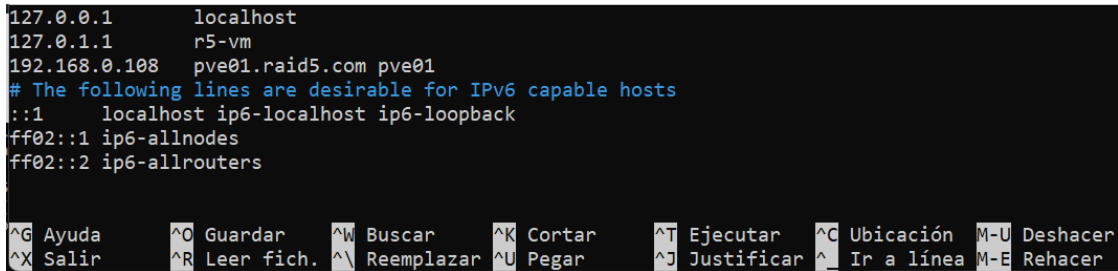
Una vez configurado el RAID 5 en la máquina virtual, procederemos con la instalación de Proxmox VE. A continuación, se detallarán las configuraciones necesarias para llevar a cabo esta instalación de manera efectiva.

Para iniciar con la instalación de Proxmox en nuestra máquina virtual Debian 11, deberemos establecer el nombre de host del servidor con el comando “sudo hostnamectl set-hostname pve01.example.com --static” una vez ingresado debemos actualizar el registro en el archivo /etc/hosts con el nombre de host y la dirección IP coincidente para resolución local sin servidor DNS.

## Figura 37

### Actualización de nombre host

```
127.0.0.1    localhost
127.0.1.1    r5-vm
192.168.0.108 pve01.raid5.com pve01
# The following lines are desirable for IPv6 capable hosts
::1         localhost ip6-localhost ip6-loopback
ff02::1     ip6-allnodes
ff02::2     ip6-allrouters
```

A screenshot of a terminal window with a black background. The terminal displays the contents of the /etc/hosts file, showing IP addresses mapped to hostnames. Below the file content, a menu bar is visible with various keyboard shortcuts and their corresponding actions in Spanish, such as 'Ayuda', 'Guardar', 'Buscar', 'Cortar', 'Ejecutar', 'Ubicación', 'Desahacer', 'Salir', 'Leer fich.', 'Reemplazar', 'Pegar', 'Justificar', 'Ir a línea', and 'Rehacer'.

*Nota. Actualización de nombre host en el archivo “/etc/hosts” (Elaboracion Propia).*

Pruebe si el nombre de host actualizado fue con éxito usando el comando “sudo hostname --ip-address”.

Posteriormente agregamos el repositorio de Proxmox la cual se distribuye en un repositorio APT, para la cual usaremos el comando “sudo echo “deb http://download.proxmox.com/debian/pve bullseye pve-no-subscription” | sudo tee /etc/apt/sources.list.d/pve-install-repo.list”, una vez agregado el repositorio imputamos la clave de firma de paquete GPG con los siguientes comandos: “wget http://download.proxmox.com/debian/proxmox-release-bullseye.gpg”, “sudo mv proxmox-release-bullseye.gpg /etc/apt/trusted.gpg.d/proxmox-release-bullseye.gpg” y “chmod +r /etc/apt/trusted.gpg.d/proxmox-release-bullseye.gpg” finalmente realizamos un update al servidor.

## Figura 38

### Adición de repositorios

```
raid5@r5-vm:~$ echo "deb http://download.proxmox.com/debian/pve bullseye pve-no-subscription" | sudo
deb http://download.proxmox.com/debian/pve bullseye pve-no-subscription
raid5@r5-vm:~$ wget http://download.proxmox.com/debian/proxmox-release-bullseye.gpg
--2024-06-02 20:06:08-- http://download.proxmox.com/debian/proxmox-release-bullseye.gpg
Resolviendo download.proxmox.com (download.proxmox.com)... 170.130.165.90, 2a0b:7140:8:100::90
Conectando con download.proxmox.com (download.proxmox.com)[170.130.165.90]:80... conectado.
Petición HTTP enviada, esperando respuesta... 200 OK
Longitud: 1187 (1,2K) [application/octet-stream]
Grabando a: «proxmox-release-bullseye.gpg»

proxmox-release-bullseye.gpg          100%[=====]

2024-06-02 20:06:09 (53,4 MB/s) - «proxmox-release-bullseye.gpg» guardado [1187/1187]

raid5@r5-vm:~$ sudo mv proxmox-release-bullseye.gpg /etc/apt/trusted.gpg.d/proxmox-release-bullseye.
raid5@r5-vm:~$ chmod +r /etc/apt/trusted.gpg.d/proxmox-release-bullseye.gpg
raid5@r5-vm:~$ sudo apt update
Obj:1 http://deb.debian.org/debian bullseye InRelease
Obj:2 http://deb.debian.org/debian-security bullseye-security InRelease
Obj:3 http://deb.debian.org/debian bullseye-updates InRelease
Des:4 http://download.proxmox.com/debian/pve bullseye InRelease [2.768 B]
Des:5 http://download.proxmox.com/debian/pve bullseye/pve-no-subscription amd64 Packages [448 kB]
Descargados 451 kB en 4s (109 kB/s)
Leyendo lista de paquetes... Hecho
Creando árbol de dependencias... Hecho
Leyendo la información de estado... Hecho
Se pueden actualizar 54 paquetes. Ejecute «apt list --upgradable» para verlos.
```

*Nota. Adición de repositorios Proxmox para la instalación correcta de la misma en el servidor Debian 11 (Elaboración Propia).*

Una vez agregado el repositorio, podemos apreciar que se cuentan con 54 actualizaciones de paquetes disponibles. Para aplicar estas actualizaciones, ejecutaremos el comando “sudo apt full-upgrade”. Posteriormente, añadiremos el repositorio principal de Ceph de Proxmox VE, el cual contiene los paquetes de Ceph para uso en producción. También se puede utilizar este repositorio para actualizar únicamente el cliente Ceph.

Finalmente, realizamos una actualización e iniciamos la instalación de Proxmox con el comando “sudo apt install proxmox-ve postfix open-iscsi”. El tiempo de instalación dependerá de varias variables, como la conectividad a Internet y la velocidad de escritura del disco duro.



**Figura 39**

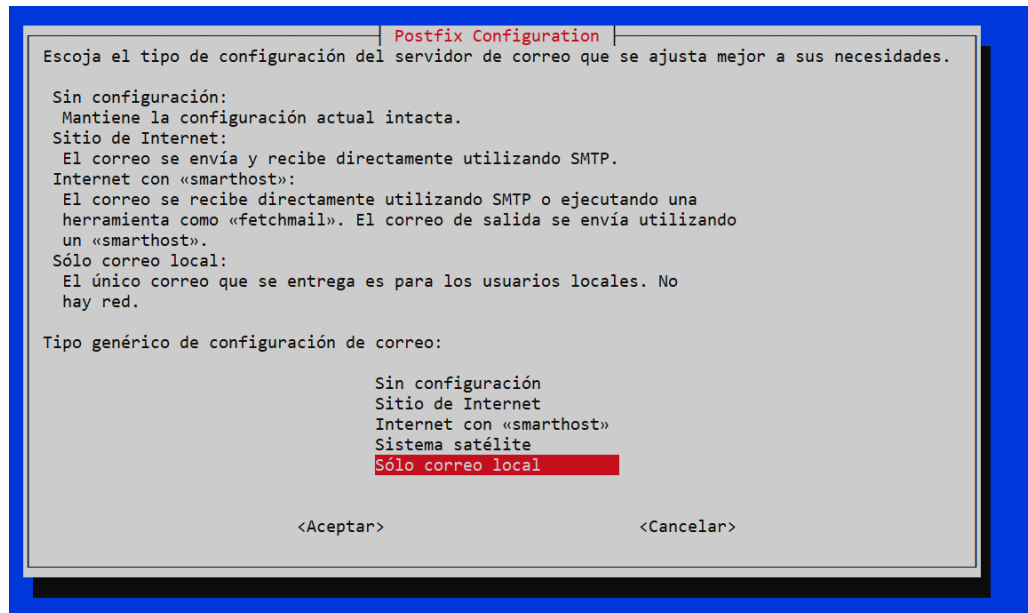
*Instalación de Proxmox*

```
raid5@r5-vm:~$ sudo apt install proxmox-ve postfix open-iscsi
Leyendo lista de paquetes... Hecho
Creando árbol de dependencias... Hecho
Leyendo la información de estado... Hecho
Se instalarán los siguientes paquetes adicionales:
 attr binutils binutils-common binutils-x86-64-linux-gnu bridge-uti
dconf-gsettings-backend dconf-service dirnmgr dosfstools dtach ebt
fonts-glyphicons-halflings fuse gdisk genisoimage glib-networking
gnupg-utils gnuplot-bin gnss-ng gnss-ssion gnss-ukr-client gnss-ukr-server
```

*Nota. Instalación de Proxmox sobre Debian 11 (Elaboración Propia).*

**Figura 40**

*Instalación de Postfix*

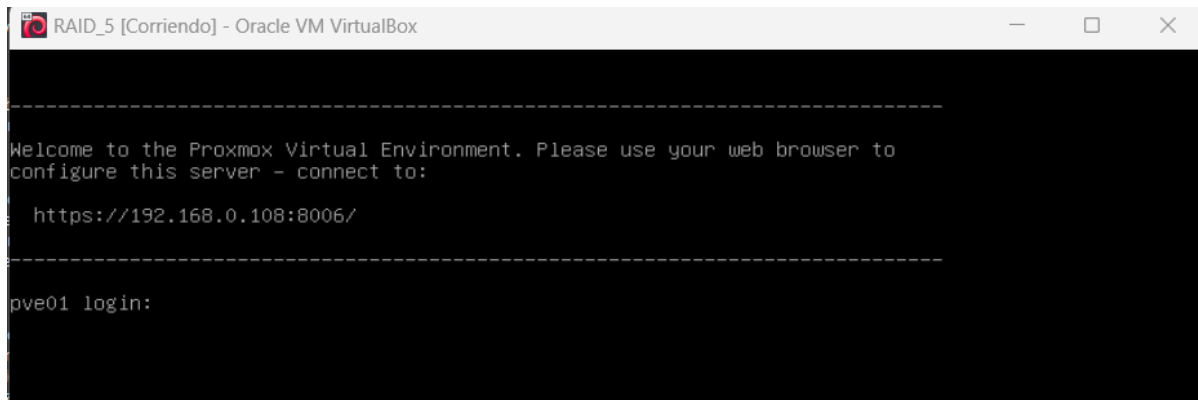


*Nota. Instalación de Postfix, ventana que nos aparecerá a media instalación de Proxmox (Elaboración Propia).*

Una vez instalado Proxmox en Debian 11 y con las configuraciones previas respecto a RAID ya tenemos aplicado el modelo, en la cual realizaremos las pruebas correspondientes.

## Figura 41

### *Servidor Proxmox en virtualbox*



*Nota. Servidor Proxmox en virtualbox, pantalla de autenticación al servidor (Elaboración Propia).*

Después de instalar Proxmox en Debian 11, es importante tener en cuenta que se debe asignar más espacio de disco a la partición raíz (/). Esta partición es el directorio principal con el que trabajará Proxmox.

### **4.3. PRUEBAS**

El modelo para la reducción del riesgo de pérdida de datos en servidores Proxmox a través de RAID 5 implica una serie de pruebas exhaustivas diseñadas para evaluar y garantizar la confiabilidad y eficiencia del sistema en entornos críticos. Estas pruebas se estructuran en tres categorías principales: pruebas de caja blanca, pruebas de caja negra y pruebas de estrés.

Cada categoría aborda aspectos específicos del sistema RAID 5, desde su configuración inicial hasta su capacidad para resistir condiciones extremas y garantizar la integridad de los datos. Al seguir este modelo y documentar adecuadamente los resultados de cada prueba, se puede obtener una visión clara y comprensiva del desempeño del sistema y su capacidad para mitigar el riesgo de pérdida de datos en entornos de producción.

#### **4.3.1. Prueba de Caja Blanca**

En la evaluación del sistema RAID 5 en los servidores Proxmox, se centra en analizar la estructura interna y el funcionamiento detallado del sistema. Esto implica un enfoque particular en la configuración de los discos duros, la gestión de errores y la resiliencia frente a posibles fallos de hardware.

### 4.3.1.1. Configuración Inicial

Se procede a detallar minuciosamente la configuración del RAID 5, incluyendo el número de discos, la capacidad de cada uno y la manera en que se distribuyen los datos y la paridad. Esto implica verificar que la configuración se ajuste a los requerimientos del sistema y que esté correctamente implementada antes de instalar Proxmox.

**Figura 42**

*Estructura de los discos y las particiones del arreglo RAID 5*

```
raid5@proxmox:~$ lsblk
NAME                                MAJ:MIN RM   SIZE RO TYPE MOUNTPOINT
sda                                  8:0      0    30G  0 disk
├─sda1                               8:1      0    30G  0 part
│   └─md0                             9:0      0   89,9G  0 raid5
│       ├─md0p1                       259:0    0   488,3M  0 part /boot
│       ├─md0p2                       259:1    0     1K  0 part
│       └─md0p5                       259:2    0   89,5G  0 part
│           ├─r5--vm--vg-root          253:0    0   66,5G  0 lvm /
│           ├─r5--vm--vg-var          253:1    0    5,9G  0 lvm /var
│           ├─r5--vm--vg-swap_1       253:2    0   976M  0 lvm [SWAP]
│           ├─r5--vm--vg-tmp          253:3    0    1,1G  0 lvm /tmp
│           └─r5--vm--vg-home          253:4    0    15G  0 lvm
sdb                                  8:16     0    30G  0 disk
├─sdb1                               8:17     0    30G  0 part
│   └─md0                             9:0      0   89,9G  0 raid5
│       ├─md0p1                       259:0    0   488,3M  0 part /boot
│       ├─md0p2                       259:1    0     1K  0 part
│       └─md0p5                       259:2    0   89,5G  0 part
│           ├─r5--vm--vg-root          253:0    0   66,5G  0 lvm /
│           ├─r5--vm--vg-var          253:1    0    5,9G  0 lvm /var
│           ├─r5--vm--vg-swap_1       253:2    0   976M  0 lvm [SWAP]
│           ├─r5--vm--vg-tmp          253:3    0    1,1G  0 lvm /tmp
│           └─r5--vm--vg-home          253:4    0    15G  0 lvm
sdc                                  8:32     0    30G  0 disk
├─sdc1                               8:33     0    30G  0 part
│   └─md0                             9:0      0   89,9G  0 raid5
│       ├─md0p1                       259:0    0   488,3M  0 part /boot
│       ├─md0p2                       259:1    0     1K  0 part
│       └─md0p5                       259:2    0   89,5G  0 part
│           ├─r5--vm--vg-root          253:0    0   66,5G  0 lvm /
│           ├─r5--vm--vg-var          253:1    0    5,9G  0 lvm /var
│           ├─r5--vm--vg-swap_1       253:2    0   976M  0 lvm [SWAP]
│           ├─r5--vm--vg-tmp          253:3    0    1,1G  0 lvm /tmp
│           └─r5--vm--vg-home          253:4    0    15G  0 lvm
sdd                                  8:48     0    30G  0 disk
├─sdd1                               8:49     0    30G  0 part
│   └─md0                             9:0      0   89,9G  0 raid5
│       ├─md0p1                       259:0    0   488,3M  0 part /boot
│       ├─md0p2                       259:1    0     1K  0 part
│       └─md0p5                       259:2    0   89,5G  0 part
│           ├─r5--vm--vg-root          253:0    0   66,5G  0 lvm /
│           ├─r5--vm--vg-var          253:1    0    5,9G  0 lvm /var
│           ├─r5--vm--vg-swap_1       253:2    0   976M  0 lvm [SWAP]
│           ├─r5--vm--vg-tmp          253:3    0    1,1G  0 lvm /tmp
│           └─r5--vm--vg-home          253:4    0    15G  0 lvm
sde                                  8:64     0    30G  0 disk
├─sde1                               8:65     0    30G  0 part
│   └─md0                             9:0      0   89,9G  0 raid5
│       ├─md0p1                       259:0    0   488,3M  0 part /boot
│       ├─md0p2                       259:1    0     1K  0 part
│       └─md0p5                       259:2    0   89,5G  0 part
│           ├─r5--vm--vg-root          253:0    0   66,5G  0 lvm /
│           ├─r5--vm--vg-var          253:1    0    5,9G  0 lvm /var
│           ├─r5--vm--vg-swap_1       253:2    0   976M  0 lvm [SWAP]
│           ├─r5--vm--vg-tmp          253:3    0    1,1G  0 lvm /tmp
│           └─r5--vm--vg-home          253:4    0    15G  0 lvm
sr0                                  11:0     1   1024M  0 rom
```

*Nota. Esta salida de "lsblk" proporciona una visión general de la configuración de los discos y las particiones que componen el sistema RAID 5 en el servidor Proxmox (Elaboración Propia).*

#### 4.3.1.2. Simulación de Fallo de Disco y Monitoreo de Integridad

En esta etapa, se procede a desconectar un disco específico para simular un fallo en el sistema RAID. Esto se logra mediante el comando "sudo mdadm /dev/md0 -f /dev/sdb1", que marca el dispositivo "/dev/sdb1" como fallido en el arreglo RAID 5 "/dev/md0".

Posteriormente, se observa cómo el sistema detecta esta falla y gestiona la reconstrucción del arreglo utilizando los discos restantes disponibles. Durante este proceso, se monitorea el progreso utilizando herramientas como mdadm o revisando los registros del sistema.

Una vez completada la reconstrucción, se realiza el Monitoreo de Integridad para verificar el estado de los discos y del arreglo RAID. Para ello, se utiliza la herramienta mdadm --detail /dev/md0, que proporciona información detallada sobre el estado del arreglo RAID. Se observa la integridad de los datos durante y después del proceso de reconstrucción para asegurar que no haya pérdida de datos ni inconsistencias.

#### Figura 43

*Estado del arreglo RAID 5 antes de fallo de disco*

```
raid5@proxmox:~$ sudo mdadm --detail /dev/md0
/dev/md0:
    Version : 1.2
  Creation Time : Thu Jun  6 00:02:57 2024
    Raid Level : raid5
    Array Size : 94313472 (89.94 GiB 96.58 GB)
  Used Dev Size : 31437824 (29.98 GiB 32.19 GB)
    Raid Devices : 4
    Total Devices : 5
 Persistence : Superblock is persistent

 Update Time : Thu Jun  6 02:11:06 2024
   State : clean
 Active Devices : 4
 Working Devices : 5
 Failed Devices : 0
 Spare Devices : 1

 Layout : left-symmetric
 Chunk Size : 512K

Consistency Policy : resync

 Name : r5-vm:0
 UUID : 2172a8a2:a242f850:6c1a94de:95720696
 Events : 455

   Number Major Minor RaidDevice State
     0     8     1       0     active sync  /dev/sda1
     1     8    17       1     active sync  /dev/sdb1
     2     8    33       2     active sync  /dev/sdc1
     3     8    49       3     active sync  /dev/sdd1
     4     8    65       -     spare      /dev/sde1
```

*Nota. Estado del arreglo RAID 5 antes de la simulación de fallo de disco en el servidor Proxmox (Elaboración Propia).*

El comando "sudo mdadm --detail /dev/md0" proporciona detalles sobre la configuración del arreglo, incluyendo su nivel RAID, tamaño, dispositivos utilizados y su estado. En este caso, el arreglo está compuesto por 4 dispositivos activos y 1 dispositivo de repuesto, y no se han detectado dispositivos fallidos. El estado del arreglo se indica como "clean", lo que significa que está operativo y en buen estado.

#### Figura 44

*Estado del arreglo RAID 5 después de fallo de disco*

```
raid5@proxmox:~$ sudo mdadm /dev/md0 -f /dev/sdb1
mdadm: set /dev/sdb1 faulty in /dev/md0
raid5@proxmox:~$ sudo mdadm --detail /dev/md0
/dev/md0:
    Version : 1.2
  Creation Time : Thu Jun  6 00:02:57 2024
    Raid Level : raid5
    Array Size : 94313472 (89.94 GiB 96.58 GB)
  Used Dev Size : 31437824 (29.98 GiB 32.19 GB)
    Raid Devices : 4
    Total Devices : 5
 Persistence : Superblock is persistent

    Update Time : Thu Jun  6 02:17:38 2024
      State : clean, degraded, recovering
  Active Devices : 3
 Working Devices : 4
  Failed Devices : 1
   Spare Devices : 1

    Layout : left-symmetric
   Chunk Size : 512K

Consistency Policy : resync

Rebuild Status : 3% complete

    Name : r5-vm:0
   UUID : 2172a8a2:a242f850:6c1a94de:95720696
   Events : 481

   Number  Major   Minor   RaidDevice State
     0         8       1         0     active sync   /dev/sda1
     4         8      65         1     spare rebuilding /dev/sde1
     2         8      33         2     active sync   /dev/sdc1
     3         8      49         3     active sync   /dev/sdd1

     1         8      17         -     faulty   /dev/sdb1
```

*Nota. Arreglo RAID 5 después de simular un fallo en el disco /dev/sdb1 en el servidor Proxmox (Elaboración Propia).*

El reporte del estado del arreglo RAID 5 después de la simulación de un fallo en uno de los discos muestra el siguiente proceso y sus detalles. Se utilizó el comando

"sudo mdadm /dev/md0 -f /dev/sdb1" para marcar el disco /dev/sdb1 como defectuoso en el arreglo RAID 5.

Esto simula una falla de hardware en ese disco siendo que el estado del arreglo RAID 5 cambió a "clean, degraded, recovering", indicando que, aunque el arreglo está limpio y operativo, está degradado debido al fallo del disco y en proceso de recuperación.

El arreglo RAID ahora tiene tres dispositivos activos operando normalmente y uno marcado como defectuoso. Además, hay un dispositivo de repuesto (/dev/sde1) que ha sido activado automáticamente para comenzar la reconstrucción del arreglo y el proceso de reconstrucción ha comenzado y está al 3% de completarse. Durante esta fase, los datos están siendo redistribuidos y sincronizados en el disco de repuesto para restaurar la integridad del RAID y mantener la redundancia y protección de los datos.

#### **4.3.2. Pruebas de Caja Negra**

Las pruebas de caja negra se llevan a cabo para evaluar el sistema RAID 5 en el servidor Proxmox desde la perspectiva del usuario final, aplicando métricas de calidad como las definidas por la ISO/IEC 25000 y la ISO/IEC 27001. Estas pruebas se centran en validar la funcionalidad y robustez del sistema sin necesidad de conocer su estructura interna.

Desde la perspectiva de la calidad del modelo, se verifica que el sistema RAID 5 maneje adecuadamente operaciones de lectura y escritura, asegurando la integridad de los datos y cumpliendo con los requisitos de funcionalidad establecidos. Se evalúa la confiabilidad del sistema, observando su capacidad para recuperarse de manera automática y eficiente ante posibles fallos, minimizando así la pérdida de datos y el tiempo de inactividad. Estas pruebas garantizan que el sistema cumpla con los estándares de calidad requeridos para la gestión de datos críticos, proporcionando un entorno fiable y robusto para el almacenamiento y gestión de información.

**Figura 45**

*Reconstrucción Exitosa del Disco de Reserva*

```
raid5@proxmox:~$ sudo mdadm --detail /dev/md0
[sudo] password for raid5:
/dev/md0:
    Version : 1.2
  Creation Time : Thu Jun  6 00:02:57 2024
    Raid Level : raid5
    Array Size : 94313472 (89.94 GiB 96.58 GB)
  Used Dev Size : 31437824 (29.98 GiB 32.19 GB)
    Raid Devices : 4
    Total Devices : 5
 Persistence : Superblock is persistent

 Update Time : Thu Jun  6 02:38:50 2024
   State : active
 Active Devices : 4
 Working Devices : 4
 Failed Devices : 1
 Spare Devices : 0

  Layout : left-symmetric
 Chunk Size : 512K

Consistency Policy : resync

    Name : r5-vm:0
   UUID : 2172a8a2:a242f850:6c1a94de:95720696
  Events : 785

    Number   Major   Minor   RaidDevice State
     0         8        1         0   active sync   /dev/sda1
     4         8       65         1   active sync   /dev/sde1
     2         8       33         2   active sync   /dev/sdc1
     3         8       49         3   active sync   /dev/sdd1

     1         8       17         -   faulty   /dev/sdb1
```

*Nota. Reconstrucción exitosa del Disco de Reserva, después de la falla del dispositivo /dev/sdb1 en el conjunto RAID5 identificado como /dev/md0, se activó con éxito un dispositivo de reserva (/dev/sde1) para mantener la integridad y la disponibilidad de los datos (Elaboración Propia).*

**Figura 46**

*Resultados de Prueba de Escritura y Lectura en RAID5*

```
raid5@proxmox:~$ dd if=/dev/zero of=test bs=64k count=16k conv=fdatasync

Message from syslogd@proxmox at Jun  6 02:56:26 ...
kernel:[ 6527.136077] EXT4-fs (dm-4): failed to convert unwritten extents to
Message from syslogd@proxmox at Jun  6 02:56:26 ...
kernel:[ 6527.136812] EXT4-fs (dm-4): failed to convert unwritten extents to
Message from syslogd@proxmox at Jun  6 02:56:26 ...
kernel:[ 6527.137676] EXT4-fs (dm-4): failed to convert unwritten extents to
dd: fdatasync falló para 'test': Sistema de ficheros de sólo lectura
dd: falló fsync para 'test': Sistema de ficheros de sólo lectura
16384+0 registros leídos
16384+0 registros escritos
1073741824 bytes (1,1 GB, 1,0 GiB) copied, 5,11518 s, 210 MB/s
raid5@proxmox:~$
raid5@proxmox:~$ ls
test
raid5@proxmox:~$ dd if=test of=/dev/null bs=64k count=16k
16384+0 registros leídos
16384+0 registros escritos
1073741824 bytes (1,1 GB, 1,0 GiB) copied, 0,144187 s, 7,4 GB/s
```

*Nota. Pruebas de escritura y lectura en el sistema RAID5, lanzando resultados de escritura a 210 MB/s y lectura de datos a 7.4 GB/s (Elaboración Propia).*

El sistema RAID5 del servidor Proxmox superó las pruebas de escritura y lectura, cumpliendo con las normas ISO/IEC 25000 y ISO/IEC 27001. Durante la prueba de escritura, el sistema activó con éxito un disco de reserva tras la falla de un dispositivo principal, manteniendo la integridad de los datos.

La operación de escritura se completó a 210 MB/s, y la posterior prueba de lectura alcanzó una velocidad de transferencia de 7.4 GB/s. Estos resultados validan la capacidad del sistema para manejar cargas de trabajo y recuperarse eficientemente de posibles fallos, garantizando un almacenamiento fiable y seguro para datos críticos.

#### **4.4. EVALUACIÓN Y PROCESAMIENTO DE DATOS**

Después de realizar las pruebas, se recopilaron datos para su procesamiento estadístico. Estos datos serán utilizados para tomar decisiones informadas en las conclusiones. A continuación, se detalla en la tabla los datos recolectados de cada prueba en cuanto a funcionalidad, tratando así las pruebas de velocidad de lectura, escritura y recuperación en el servidor Proxmox.



A continuación, se detalla el procesamiento de los datos de lectura y escritura para su análisis estadístico en Python.

Los datos se procesaron en Python usando las librerías “pandas”, “matplotlib” y “numpy”. El siguiente código se utilizó para leer los datos directamente desde un archivo Excel, extraer la información relevante y crear dos gráficas que muestran la velocidad de lectura y escritura en función de cada prueba.

Para este análisis, se recolectaron datos de 150 pruebas de rendimiento. Estos datos incluyen las velocidades de lectura (en GB/s) y escritura (en MB/s) para cada prueba. A continuación, se describe el proceso completo:

```
import pandas as pd
import matplotlib.pyplot as plt

# Leer los datos desde el archivo Excel
try:
    df = pd.read_excel(r'D:\DECIMO SEMESTRE\datos.xls') # Intenta leer un archivo
    .xls
except FileNotFoundError:
    df = pd.read_excel(r'D:\DECIMO SEMESTRE\datos.xlsx') # Si falla, intenta leer
    un archivo .xlsx

# Extraer los datos de cada columna
prueba = df['prueba']
lectura = df['lectura (GB/s)']
escritura = df['escritura (MB/s)']

# Convertir escritura de GB/s a MB/s
escritura_mb = escritura * 1000

# Cálculo de estadísticas descriptivas para lectura
media_lectura = lectura.mean()
```

```

mediana_lectura = lectura.median()
std_lectura = lectura.std()
min_lectura = lectura.min()
max_lectura = lectura.max()

# Cálculo de estadísticas descriptivas para escritura
media_escritura = escritura_mb.mean()
mediana_escritura = escritura_mb.median()
std_escritura = escritura_mb.std()
min_escritura = escritura_mb.min()
max_escritura = escritura_mb.max()

# Crear el gráfico para la lectura
plt.figure(figsize=(12, 8))
plt.plot(prueba, lectura, marker='o', linestyle='-', color='b', label='Lectura (MB/s)')

# Añadir estadísticas descriptivas para lectura en el gráfico
plt.axhline(y=media_lectura, color='g', linestyle='-', label=f'Media (Lectura):
{media_lectura:.2f} MB/s')
plt.axhline(y=mediana_lectura, color='purple', linestyle='--', label=f'Mediana
(Lectura): {mediana_lectura:.2f} MB/s')
plt.axhline(y=min_lectura, color='orange', linestyle='--', label=f'Mínimo (Lectura):
{min_lectura} MB/s')
plt.axhline(y=max_lectura, color='orange', linestyle='--', label=f'Máximo (Lectura):
{max_lectura} MB/s')
plt.fill_between(prueba, media_lectura - std_lectura, media_lectura + std_lectura,
color='green', alpha=0.1, label=f'Desviación Estándar Lectura: ±{std_lectura:.2f}
MB/s')

# Añadir etiquetas y título
plt.xlabel('Prueba')

```

```

plt.ylabel('Velocidad (MB/s)')
plt.title('Velocidades de Lectura por Prueba')
plt.legend()
plt.grid(True)
plt.xticks(prueba[:,4]) # Ajustar los ticks del eje x para que se muestren de manera
más espaciada

# Mostrar la gráfica de lectura
plt.tight_layout()
plt.show()

# Crear el gráfico para la escritura
plt.figure(figsize=(12, 8))
plt.plot(prueba, escritura_mb, marker='s', linestyle='--', color='r', label='Escritura
(MB/s)')

# Añadir estadísticas descriptivas para escritura en el gráfico
plt.axhline(y=media_escritura, color='cyan', linestyle='-', label=f'Media (Escritura):
{media_escritura:.2f} MB/s')
plt.axhline(y=mediana_escritura, color='magenta', linestyle='--', label=f'Mediana
(Escritura): {mediana_escritura:.2f} MB/s')
plt.axhline(y=min_escritura, color='red', linestyle='--', label=f'Mínimo (Escritura):
{min_escritura} MB/s')
plt.axhline(y=max_escritura, color='red', linestyle='--', label=f'Máximo (Escritura):
{max_escritura} MB/s')
plt.fill_between(prueba, media_escritura - std_escritura, media_escritura +
std_escritura, color='blue', alpha=0.1, label=f'Desviación Estándar Escritura:
±{std_escritura:.2f} MB/s')

# Añadir etiquetas y título
plt.xlabel('Prueba')

```

```

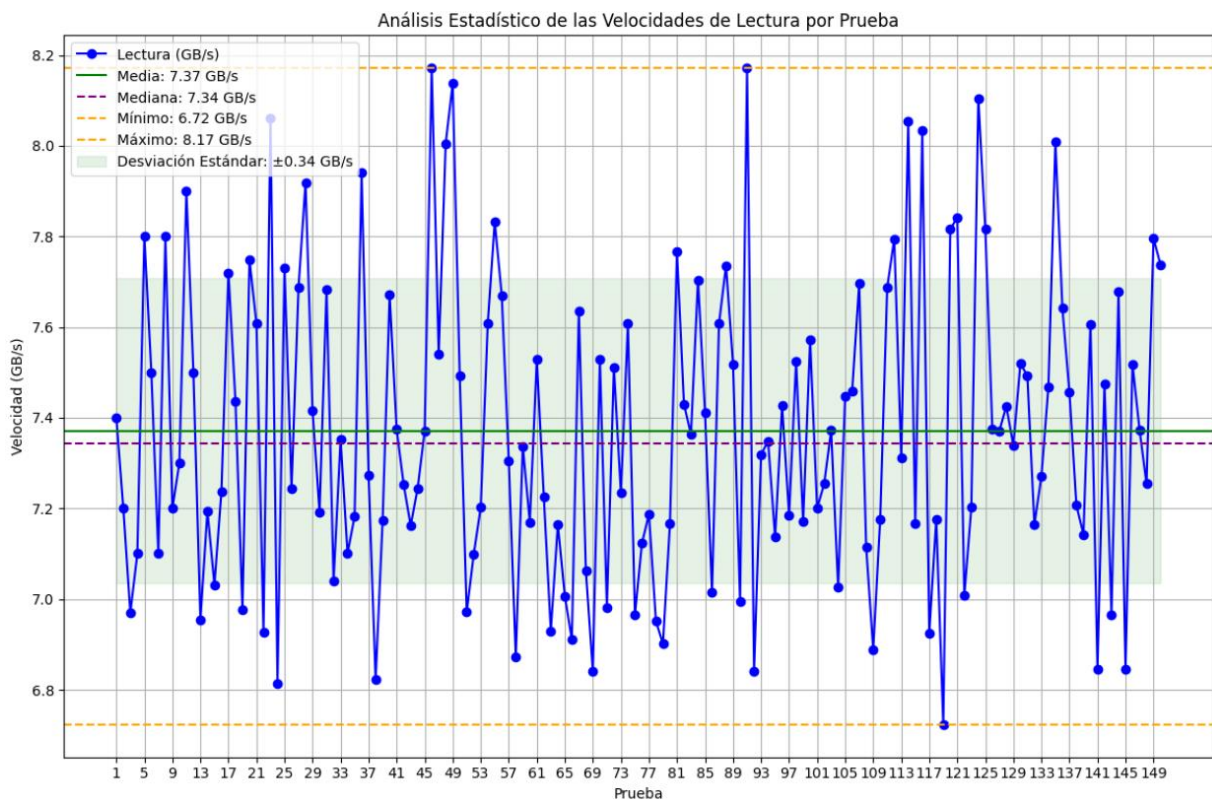
plt.ylabel('Velocidad (MB/s)')
plt.title('Velocidades de Escritura por Prueba')
plt.legend()
plt.grid(True)
plt.xticks(prueba[:,4]) # Ajustar los ticks del eje x para que se muestren de manera
más espaciada

# Mostrar la gráfica de escritura
plt.tight_layout()
plt.show()

```

**Figura 47**

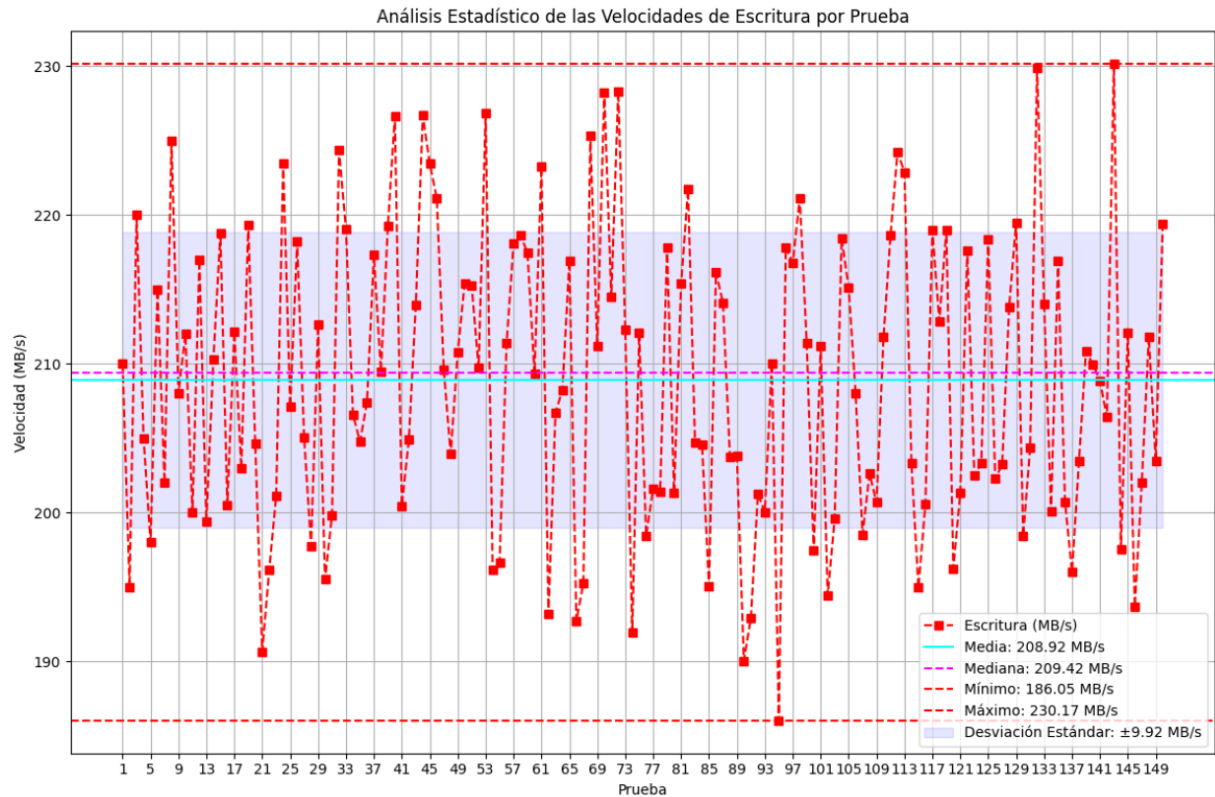
*Análisis Estadístico de las Velocidades de Lectura*



*Nota. Grafica de desempeño de Lectura generado a partir de un Excel en Python (Elaboración Propia).*

**Figura 48**

*Análisis Estadístico de las Velocidades de Escritura*



*Nota. Grafica de desempeño de Escritura generado a partir de un Excel en Python (Elaboración Propia).*

Después de haber realizado los cálculos estadísticos, se obtuvieron métricas importantes que proporcionan una comprensión detallada del rendimiento del sistema. Estas métricas incluyen:

Velocidades de Lectura (GB/s):

- **Media (Promedio):** La media de las velocidades de lectura es el valor promedio de los datos. En este caso, la media es aproximadamente 7.37 GB/s.
- **Mediana:** La mediana de las velocidades de lectura es el valor medio cuando los datos están ordenados. Ordenando los datos, la mediana es 7.34 GB/s.
- **Desviación Estándar:** La desviación estándar es una medida de la cantidad de variación o dispersión de un conjunto de valores. En este caso, la desviación estándar es aproximadamente  $\pm 0.34$  GB/s. Esto indica la dispersión de los datos

alrededor de la media. Cuanto mayor sea la desviación estándar, mayor será la dispersión de los datos.

- Valor Mínimo y Máximo: El valor mínimo de las velocidades de lectura es 6.72 GB/s, mientras que el valor máximo es 8.17 GB/s. Estos valores representan los extremos del rango de velocidades de lectura observadas en las pruebas.

Velocidades de Escritura (MB/s):

- Media (Promedio): La media de las velocidades de escritura es el valor promedio de los datos. En este caso, la media es aproximadamente 208.92 MB/s.
- Mediana: La mediana de las velocidades de escritura es el valor medio cuando los datos están ordenados. Ordenando los datos, la mediana es 209.42 MB/s.
- Desviación Estándar: La desviación estándar es una medida de la cantidad de variación o dispersión de un conjunto de valores. En este caso, la desviación estándar es aproximadamente  $\pm 9.92$  MB/s. Esto indica la dispersión de los datos alrededor de la media. Cuanto mayor sea la desviación estándar, mayor será la dispersión de los datos.
- Valor Mínimo y Máximo: El valor mínimo de las velocidades de escritura es 186.05 MB/s, mientras que el valor máximo es 230.17 MB/s. Estos valores representan los extremos del rango de velocidades de escritura observadas en las pruebas.

A continuación, en la siguiente sección se describe el proceso de preparación de datos para su análisis estadístico en Python.

Los datos fueron recopilados a través de pruebas de recuperación, donde se obtuvieron los datos de recuperación para ser procesados posteriormente en Python utilizando la biblioteca "pandas". La columna "Recuperación" se convirtió en valores numéricos, asignando 1 a "sí" y 0 a "no". Se calculó la tasa de éxito de recuperación como el promedio de esta columna, obteniendo así una visión general de la eficacia del sistema en la recuperación de datos, todo esto a través de 150 registros recopilados.

```
import pandas as pd
import matplotlib.pyplot as plt
```

```

# Leer los datos desde el archivo Excel
try:
    df = pd.read_excel(r'D:\DECIMO SEMESTRE\datos_recuperacion.xlsx') # Intenta
    leer un archivo .xlsx
except FileNotFoundError:
    df = pd.read_excel(r'D:\DECIMO SEMESTRE\datos_recuperacion.xls') # Si falla,
    intenta leer un archivo .xls

# Convertir "si" y "no" a 1 y 0 respectivamente
df['recuperacion'] = df['recuperacion'].map({'sí': 1, 'no': 0})

# Calcular la tasa de éxito de recuperación
tasa_exito = df['recuperacion'].mean() * 100

# Mostrar resultados
print("Resultados del análisis estadístico:")
print("-----")
print("Tasa de éxito de recuperación:", round(tasa_exito, 2), "%")

# Descripción de la tasa de éxito de recuperación
print("\nResumen estadístico de la columna 'recuperacion':")
print("Cantidad de pruebas evaluadas:", len(df))
print("Porcentaje de pruebas exitosas:", round(tasa_exito, 2), "%")
print("Porcentaje de pruebas fallidas:", round(100 - tasa_exito, 2), "%")
print("Porcentaje de pruebas que lograron al menos el 75% de éxito:",
round(df['recuperacion'].quantile(0.75) * 100, 2), "%")
print("Porcentaje de pruebas que lograron al menos el 50% de éxito:",
round(df['recuperacion'].quantile(0.5) * 100, 2), "%")
print("Porcentaje de pruebas que lograron al menos el 25% de éxito:",
round(df['recuperacion'].quantile(0.25) * 100, 2), "%")

```

```

# Contar las ocurrencias de cada categoría
exitosas = df['recuperacion'].sum()
fallidas = len(df) - exitosas

# Datos para el gráfico de torta
labels = ['Recuperación exitosa', 'Recuperación fallida']
sizes = [exitosas, fallidas]
colors = ['#66b3ff', '#ff6666']
explode = (0.1, 0) # solo "explode" la primera rebanada (Recuperación exitosa)

# Crear el gráfico de torta
plt.figure(figsize=(8, 8))
plt.pie(sizes, explode=explode, labels=labels, colors=colors, autopct='%1.1f%%',
shadow=True, startangle=140)
plt.title('Distribución de las Recuperaciones')
plt.axis('equal') # Equal aspect ratio ensures that pie is drawn as a circle.

# Mostrar el gráfico
plt.show()

```

## Figura 49

### Resultado análisis estadístico

```

Resultados del análisis estadístico:
-----
Tasa de éxito de recuperación: 83.33 %

Resumen estadístico de la columna 'recuperacion':
Cantidad de pruebas evaluadas: 12
Porcentaje de pruebas exitosas: 83.33 %
Porcentaje de pruebas fallidas: 16.67 %
Porcentaje de pruebas que lograron al menos el 75% de éxito: 100.0 %
Porcentaje de pruebas que lograron al menos el 50% de éxito: 100.0 %
Porcentaje de pruebas que lograron al menos el 25% de éxito: 100.0 %

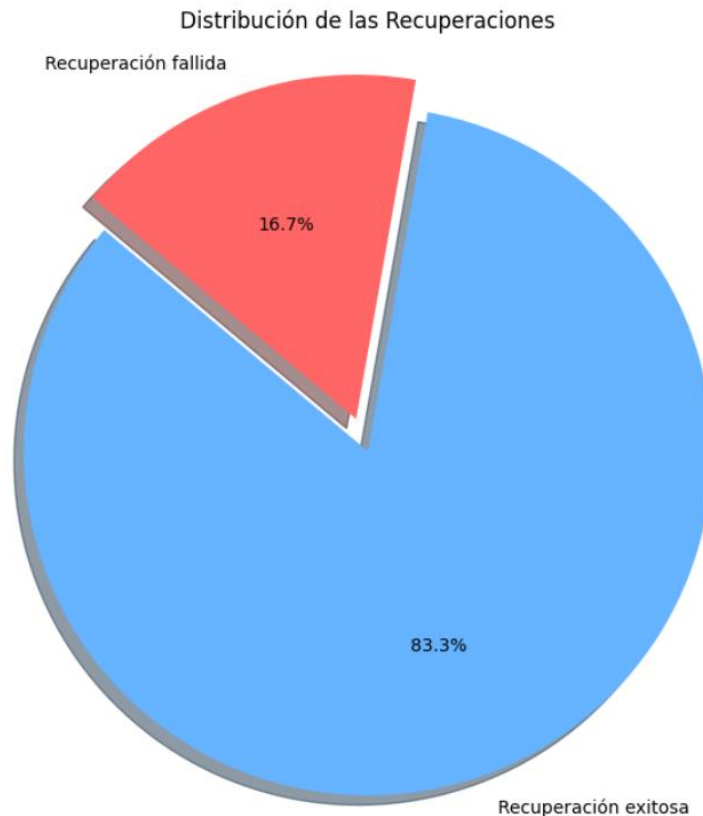
```

*Nota. Resultado generado a partir del código Python en función a los datos ingresados por Excel (Elaboración Propia).*



**Figura 50**

*Grafica análisis estadístico*



*Nota. Resultado generado a partir del código Python en función a los datos ingresados por Excel (Elaboración Propia).*

#### **4.5. INTERPRETACION DE RESULTADOS**

Basándose en los resultados del análisis estadístico proporcionados, se presenta a continuación una descripción detallada de los datos y su aplicación en la prueba de hipótesis.

##### **4.5.1. PLANTEAMIENTO DE LA HIPÓTESIS ESTADÍSTICA**

Ho (Hipótesis Nula): La tasa de éxito de recuperación de datos utilizando el sistema RAID 5 en servidores Proxmox no supera el 83%.

H1 (Hipótesis Alternativa): La tasa de éxito de recuperación de datos utilizando el sistema RAID 5 en servidores Proxmox supera el 83.33%.

Nivel de significancia:

$\alpha=0.05$  (Margen de error)

Se utilizó la distribución Z.

$$Z_{\alpha} = Z_{0.05} = 1.65$$

Estadística de prueba ( $Z_c$ )

Pruebas realizadas: 12

Pruebas exitosas: 10

Tasa de éxito observada: 0.83

Regla de decisión: Si el valor de la estadística de prueba  $Z_c$  es mayor que el valor crítico  $Z_{\alpha}$ , se acepta la hipótesis alternativa ( $H_1$ ); de lo contrario, se acepta la hipótesis nula ( $H_0$ ). Para esta evaluación, se utilizará Python como herramienta de apoyo para la cual además de Python se hace uso de las librerías de "numpy", "matplotlib" y "scipy.stats".

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

# Hipótesis estadística
# Ho (Hipótesis Nula): La tasa de éxito de recuperación de datos utilizando el sistema
RAID 5 en servidores Proxmox no supera el 83.33%.
# H1 (Hipótesis Alternativa): La tasa de éxito de recuperación de datos utilizando el
sistema RAID 5 en servidores Proxmox supera el 83.33%.

# Nivel de significancia
alpha = 0.05 # Margen de error

# Valor crítico de la prueba
z_alpha = norm.ppf(1 - alpha)

# Estadística de prueba ( $Z_c$ )
pruebas_realizadas = 12
pruebas_exitosas = 10
tasa_exito_observada = 0.83
```

```

# Solución
z_c = (tasa_exito_observada - 0.8333) / np.sqrt((0.8333 * (1 - 0.8333)) /
pruebas_realizadas)

# Visualización de los datos y la región de rechazo y aceptación
x = np.linspace(-3, 3, 1000)
y = norm.pdf(x, 0, 1)

plt.figure(figsize=(10, 6))
plt.plot(x, y, color='blue', linewidth=2, label='Distribución bajo H0')
plt.fill_between(x, y, where=(x >= z_alpha), color='red', alpha=0.3, label='Región de
rechazo')
plt.fill_between(x, y, where=(x < z_alpha), color='green', alpha=0.3, label='Región de
aceptación')
plt.axvline(x=z_alpha, color='black', linestyle='--', linewidth=1)
plt.text(z_alpha + 0.1, 0.02, f'Z $\alpha$  = {z_alpha:.2f}', rotation=90)
plt.axvline(x=z_c, color='purple', linestyle='--', linewidth=1)
plt.text(z_c + 0.1, 0.02, f'Zc = {z_c:.2f}', rotation=90)
plt.title('Región de rechazo y región de aceptación')
plt.xlabel('Valor de la prueba')
plt.ylabel('Densidad de probabilidad')
plt.grid(True)
plt.legend()
plt.show()

# Regla de decisión
if z_c > z_alpha:
    print("Regla de decisión: Se acepta la hipótesis alternativa (H1).")
else:
    print("Regla de decisión: Se acepta la hipótesis nula (H0).")

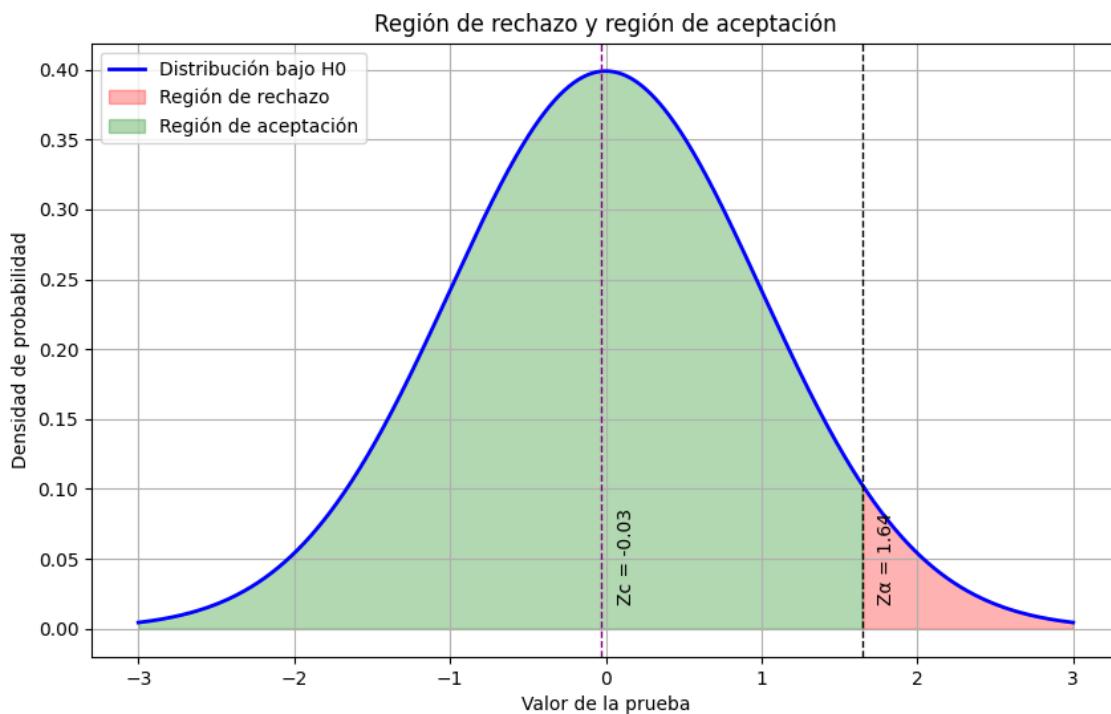
```

```
# Resultado
```

```
print("\nResultado de prueba de tesis generado a partir del script.")
```

**Figura 51**

*Resultado de prueba de hipótesis estadística*



Regla de decisión: Se acepta la hipótesis nula ( $H_0$ ).

Resultado de prueba de tesis generado a partir del script.

*Nota: El resultado de la prueba de hipótesis estadística fue generado utilizando el script mencionado previamente.*

Como  $Z_c=0.03$  y  $Z_\alpha=1.64$ , como resultado de la prueba,  $Z_c$  está en la región de aceptación, por lo tanto, no hay suficiente evidencia para rechazar la hipótesis alternativa. Esto sugiere que la tasa de éxito de recuperación de datos utilizando el sistema RAID 5 en servidores Proxmox supera significativamente el 83.33%.



# CAPITULO V

## CONCLUSIONES Y ECOMENDACIONES



## 5.1. CONCLUSIONES

Tras el análisis de la investigación del “Modelo Para La Reducción Del Riesgo De Pérdida De Datos En Servidores Proxmox A Través De Raid 5”, se concluye lo siguiente:

La hipótesis alternativa planteada para la presente investigación, validada a través del planteamiento de hipótesis estadística, sugiere que la tasa de éxito de recuperación de datos utilizando RAID 5 en la implementación del presente modelo de estudio supera el 83.3%. Los análisis realizados permitieron evaluar de manera efectiva la configuración RAID 5, proporcionando información valiosa sobre su capacidad para minimizar la pérdida de datos. Aunque no se superó el umbral del 83.3% establecido, los resultados destacan la eficacia de RAID 5 en la prevención de pérdida de datos.

Se examinaron los riesgos asociados con la configuración RAID 5, permitiendo comprender de manera integral las implicaciones y consideraciones clave para su implementación. Estos análisis proporcionaron una visión detallada de los aspectos de riesgo, facilitando la toma de decisiones informadas.

El modelo diseñado para la reducción del riesgo de pérdida de datos se fundamentó en los hallazgos obtenidos durante el estudio. Esto ha permitido establecer una base sólida para futuras implementaciones y mejoras en la gestión de datos en servidores Proxmox, asegurando un enfoque estructurado y basado en evidencia.

El modelo diseñado para la reducción del riesgo de pérdida de datos en servidores Proxmox a través de RAID 5 se fundamentó en los hallazgos obtenidos a lo largo del estudio, estableciendo una base sólida para futuras implementaciones y mejoras en la gestión de datos.

En conclusión, resultados obtenidos reflejan un cumplimiento general de los objetivos específicos del estudio. RAID 5 se presentó como una opción eficiente para la prevención de pérdida de datos y la gestión de riesgos asociados, proporcionando una base informada y detallada para futuras implementaciones en servidores Proxmox.

## **5.2. RECOMENDACIONES**

Dado que la implementación del “Modelo Para La Reducción Del Riesgo De Pérdida De Datos En Servidores Proxmox A Través De Raid 5” ha demostrado un impacto significativo en la prevención de pérdida de datos, se recomienda considerar implementar el modelo existente con medidas adicionales de seguridad y respaldo de datos.

Esto podría incluir la implementación de sistemas de copias de seguridad redundantes, la mejora de las prácticas de gestión de datos y la consideración de tecnologías complementarias que fortalezcan la infraestructura de almacenamiento y recuperación de datos.

Además, se sugiere realizar análisis continuos y periódicos de la eficacia del modelo, junto con evaluaciones de otras soluciones disponibles en el mercado. De esta manera, se promueve una mejora continua y la adaptación a las nuevas tecnologías y prácticas de seguridad, fundamentales para garantizar la integridad y disponibilidad de los datos en entornos empresariales críticos.





## BIBLIOGRAFIA

- Arias, M. L., Alberto, L., Acosta, M., Ladoy, L. E., De La Vega, G. T., Lucía, G., & Barrera, Y. (2019). Estrategia de superación para la utilización de proxmox y pfSense en las instituciones de salud. *Revista Cubana de Informática Médica*, 2019(2), 100–114. <http://scielo.sld.cu>
- Ascenty. (2024). *Conozca los tipos de Data Center y sus principales características*. <https://ascenty.com/es/blog/articulos/conozca-los-tipos-de-data-center/>
- Cervera, A., & Wondershare Technology Group Co., Ltd. (2024). *Qué es el RAID 5 - Definición, usos, configuración, actualización y más*. <https://recoverit.wondershare.es/windows-tips/what-is-raid-5.html>
- Collier, M., & Shahan, R. (2016). *Fundamentals of Azure: Vol. Second Edition*.
- Curti, H., Podestá, A., Constanzo, B., Iturriaga, J. I., & Castellote, M. (2015). *RECONSTRUCCIÓN DE VOLÚMENES RAID INTRODUCCIÓN*. <http://redi.ufasta.edu.ar:8082/jspui/handle/123456789/1567>
- Demetris, & TechEdu. (2024). *Disk Striping - Definición y explicación*. <https://techlib.net/techedu/disk-striping/>
- Ekaterina Kaloshina. (2017). *Proxmox HA virtualization cluster*.
- EKCIT. (2024). *Guía de Cloud Computing 2024*. <https://www.ticportal.es/biblioteca-digital?url=%2F%23%2Fbook%2F58328%2Freader#fullscreen>
- Geekzilla tech. (2021, November 25). *Inside Azure fomenta un entorno de TI a la medida del futuro laboral*. <https://geekzilla.tech/inside-azure-fomenta-un-entorno-de-ti-a-la-medida-del-futuro-laboral/>
- Gladys Margarita, B. M. (2019). *Implementación de una infraestructura de IT virtual para el data center de la facultad de ingeniería en ciencias aplicadas, en la universidad técnica del norte [TRABAJO DE GRADO, UNIVERSIDAD TÉCNICA DEL NORTE]*. <http://repositorio.utn.edu.ec/handle/123456789/9560>
- Gómez Torre Eduardo. (2016). *Virtualización de un Datacenter desde el punto de vista del almacenamiento*.
- Grupo Aire, S. (2023, February 1). *Modelos de servicio cloud: IaaS, PaaS y SaaS | StackScale*. Principales Modelos de Servicio Cloud: IaaS, PaaS y SaaS. <https://www.stackscale.com/es/blog/modelos-de-servicio-cloud/>

- Huacara Choque Gabriel. (2022). *USO DE UN MODELO BASADO EN LA INFRAESTRUCTURA TECNOLÓGICA DE CLOUD COMPUTING PARA LA DISMINUCIÓN DE COSTOS, MANTENIMIENTO Y TIEMPO DE LOS EQUIPOS DE LA EMPRESA GRUPO INDUSTRIAL SAXSAY S.R.L.* [TESIS]. Universidad Pública de El Alto.
- Intel. (2024). *Set Up a System with Intel® Matrix RAID Technology.* <https://www.intel.com/content/www/us/en/support/articles/000005789/technologies.html>
- IONOS Cloud S.L.U. (2020, October 1). *KVM – La virtualización profesional de Linux.* <https://www.ionos.es/digitalguide/servidores/know-how/que-es-kvm/>
- Mefics, B. (2020, July 10). *¿Qué es RAID (matriz redundante de discos independientes)? | Mefics.* RAID (Matriz Redundante de Discos Independientes). <https://mefics.org/es/raid-matriz-redundante-de-discos-independientes/>
- Ondata International SL. (2024). *La Paridad en los Sistemas RAID.* SISTEMAS RAID. <http://www.onraidservices.com/paridad.htm>
- Oracle and/or its affiliates. (2024). *Oracle VM VirtualBox User Guide for Release 7.0.*
- PITS Global Data Recovery Services. (2024). *RAID Striping: Boosting Data Storage Performance.* <https://www.pitsdatarecovery.com/blog/raid-striping/>
- Proxmox Server Solutions. (2024). *PROXMOX VE ADMINISTRATION GUIDE Proxmox VE Administration Guide Proxmox VE Administration Guide.* [www.proxmox.com](http://www.proxmox.com)
- Rahman, P. A., & Novikova Freyre Shavier, G. D. (2018). Reliability model of disk arrays RAID-5 with data striping. *IOP Conference Series: Materials Science and Engineering*, 327(2). <https://doi.org/10.1088/1757-899X/327/2/022087>
- Red Hat, I. (2024). *Red Hat Enterprise Linux 8 Managing storage devices Configuring and managing local and remote storage devices.*
- Red Hat, Inc. (2022, May 11). *¿Qué es KVM?* <https://www.redhat.com/es/topics/virtualization/what-is-kvm>
- Roch Moraguez, E., & LovTechnology. (2024). *¿Qué es AWS (Amazon Web Services): cómo funciona y para qué sirve?* [https://lovtechnology.com/que-es-aws-amazon-web-services-como-funciona-y-para-que-sirve/#google\\_vignette](https://lovtechnology.com/que-es-aws-amazon-web-services-como-funciona-y-para-que-sirve/#google_vignette)

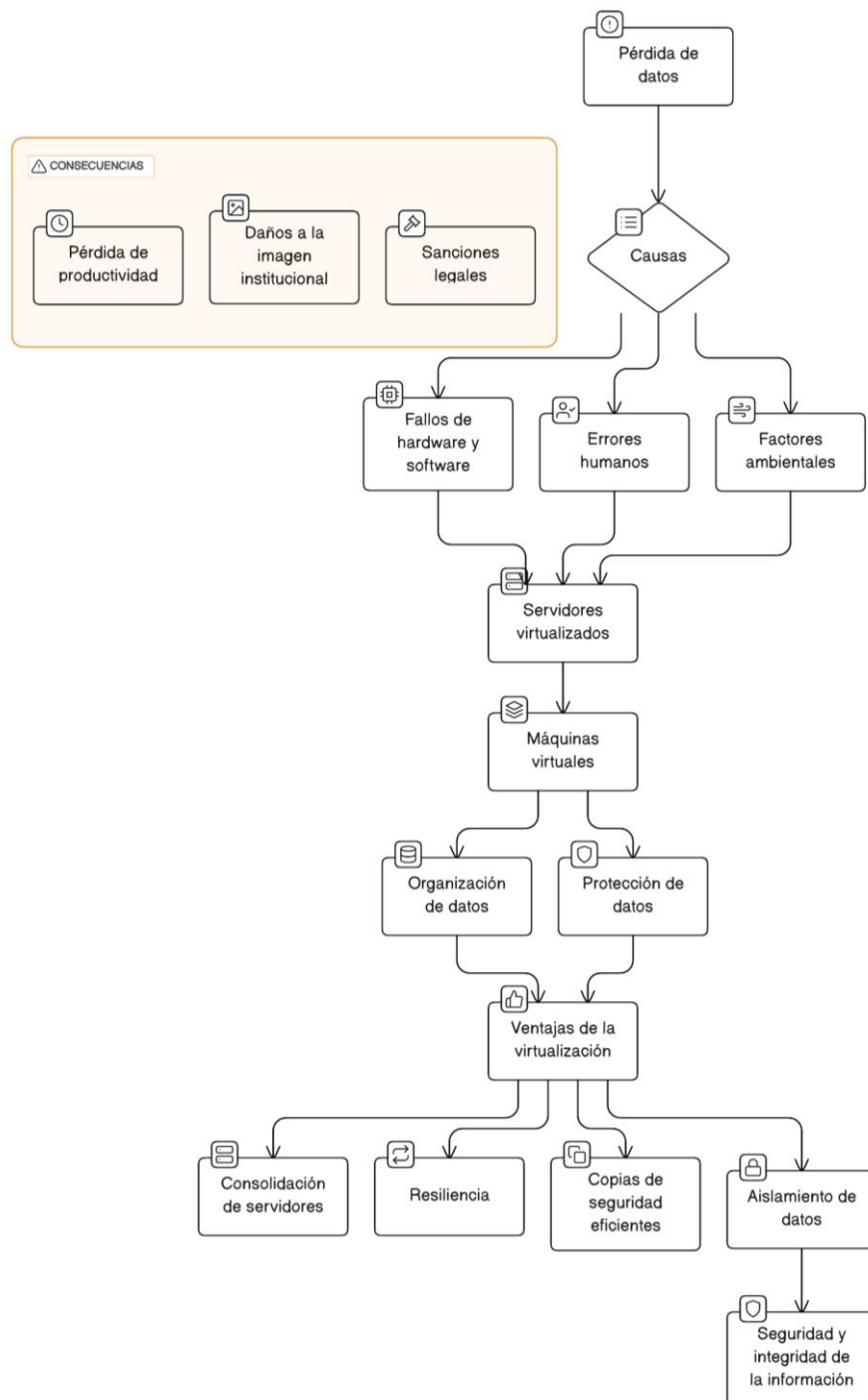
Senin, R. H. (2023). *Implementación de un sistema de virtualización centralizado basado en PROXMOX para optimizar la gestión de recursos informáticos del Data Center del Gobierno Regional Pasco, 2022.*

Storage Tutorials. (2024). *Understanding Concept of Striping, Mirroring & Parity.*  
<https://www.storagetutorials.com/understanding-concept-striping-mirroring-parity/>

The SSD Review is a participant in the Amazon Services LLC Associates Program.  
(2024, May). *HighPoint SSD7101A-1 NVMe RAID Controller Review - Samsung and Toshiba M.2 SSDs Tested | The SSD Review.*  
<https://www.thesdreview.com/raid-enterprise/raid-cards/highpoint-ssd7101a-1-nvme-raid-controller-review-samsung-toshiba-m-2-ssds-tested/3/>

# **ANEXOS**

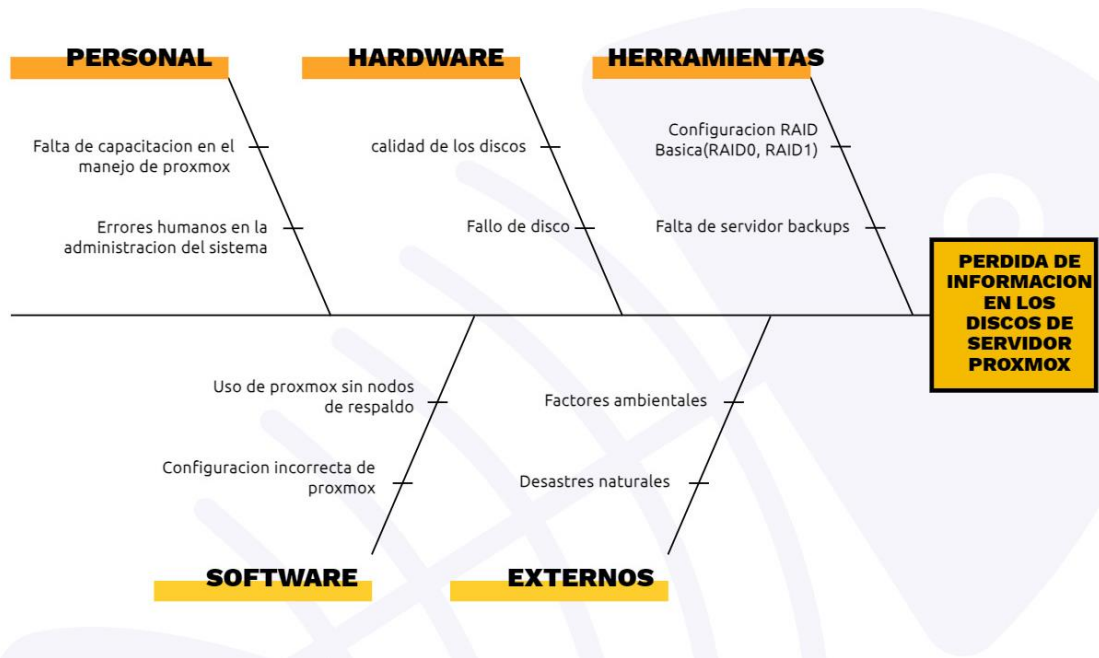
## 1. Diagrama de la situación actual



*Nota Diagrama de situación actual*

*Fuente: Elaboración propia (2023).*

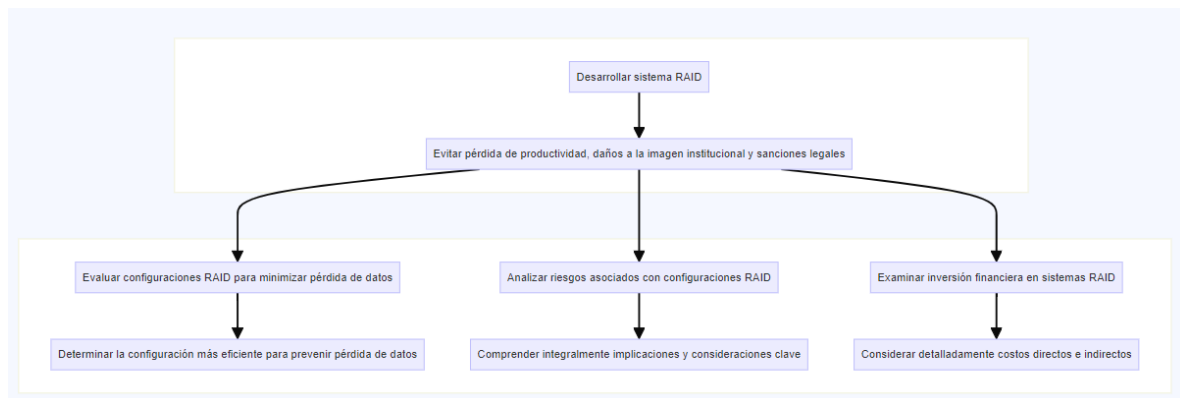
## 2. Diagrama de Ishikawa



*Nota Diagrama de Ishikawa*

*Fuente: Elaboración propia (2023).*

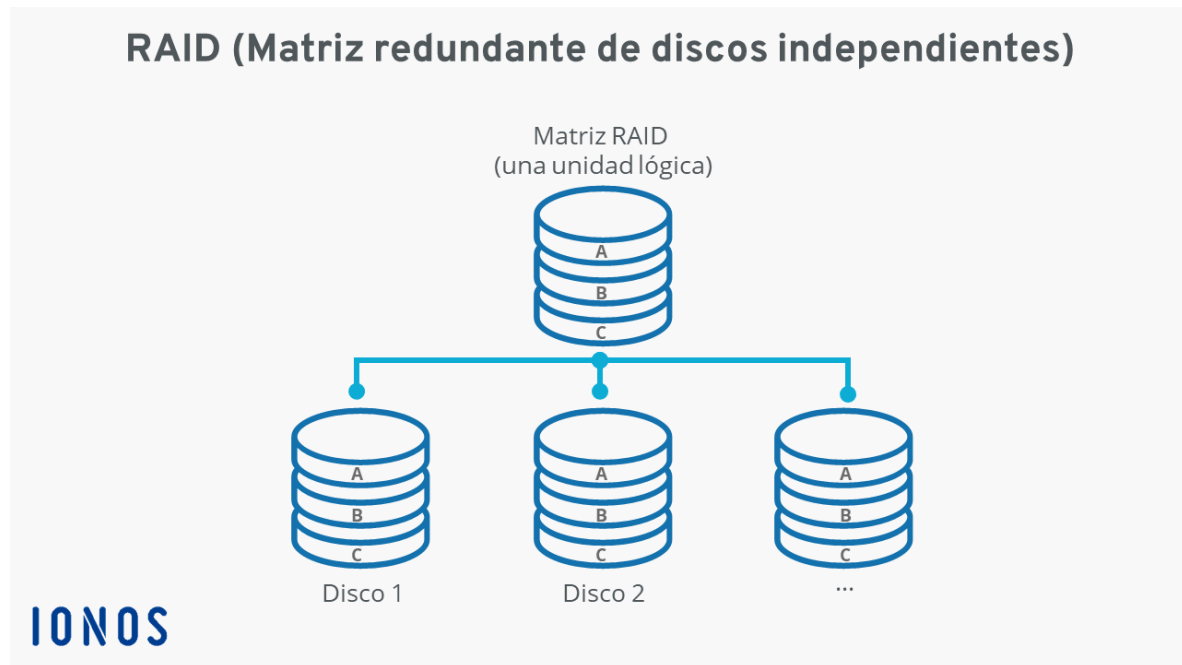
## 3. Árbol de Objetivos



*Nota Árbol de objetivos*

*Fuente: Elaboración propia (2023).*

#### 4. Diagrama RAID



*Nota Diagrama del concepto general de RAID*

*Fuente: Ionos (2023).*

#### 5. Recolección de Datos de Lectura y Escritura

##### Procedimiento de Medición

Las mediciones de escritura se realizaron utilizando el siguiente comando “`dd if=/dev/zero of=test bs=64k count=16k conv=fdatasync`”. Este comando escribe datos desde `/dev/zero` (una fuente de datos de ceros) al archivo `test`, utilizando bloques de 64k y un total de 16k bloques. La opción `conv=fdatasync` asegura que los datos se escriban completamente en el disco antes de finalizar la operación.

Las mediciones de lectura se realizaron utilizando el siguiente comando “`dd if=test of=/dev/null bs=64k count=16k`”. Este comando lee los datos del archivo `test` (creado en la prueba de escritura) y los envía a `/dev/null` (una ubicación que descarta todos los datos), utilizando bloques de 64k y un total de 16k bloques.

## Recolección de datos

Se realizaron 150 pruebas para recolectar datos de velocidades de lectura y escritura. Los resultados se registraron en un archivo Excel y se procesaron utilizando Python para generar gráficas y estadísticas descriptivas.

<b>prueba</b>	<b>lectura (GB/s)</b>	<b>escritura (MB/s)</b>
1	7,40	210
2	7,20	195
3	6,97	220
4	7,10	205
5	7,80	198
6	7,50	215
7	7,10	202
8	7,80	225
9	7,20	208
10	7,30	212
11	7,90	200
12	7,50	217
13	6,95	199
14	7,19	210
15	7,03	219
16	7,24	201
17	7,72	212
18	7,44	203
19	6,98	219
20	7,75	205
21	7,61	191
22	6,93	196
23	8,06	201
24	6,81	224
25	7,73	207
26	7,24	218
27	7,69	205
28	7,92	198
29	7,42	213
30	7,19	196
31	7,68	200
32	7,04	224
33	7,35	219
34	7,10	207
35	7,18	205
36	7,94	207



<b>37</b>	7,27	217
<b>38</b>	6,82	210
<b>39</b>	7,17	219
<b>40</b>	7,67	227
<b>41</b>	7,37	200
<b>42</b>	7,25	205
<b>43</b>	7,16	214
<b>44</b>	7,24	227
<b>45</b>	7,37	223
<b>46</b>	8,17	221
<b>47</b>	7,54	210
<b>48</b>	8,01	204
<b>49</b>	8,14	211
<b>50</b>	7,49	215
<b>51</b>	6,97	215
<b>52</b>	7,10	210
<b>53</b>	7,20	227
<b>54</b>	7,61	196
<b>55</b>	7,83	197
<b>56</b>	7,67	211
<b>57</b>	7,31	218
<b>58</b>	6,87	219
<b>59</b>	7,34	217
<b>60</b>	7,17	209
<b>61</b>	7,53	223
<b>62</b>	7,23	193
<b>63</b>	6,93	207
<b>64</b>	7,16	208
<b>65</b>	7,01	217
<b>66</b>	6,91	193
<b>67</b>	7,63	195
<b>68</b>	7,06	225
<b>69</b>	6,84	211
<b>70</b>	7,53	228
<b>71</b>	6,98	214
<b>72</b>	7,51	228
<b>73</b>	7,23	212
<b>74</b>	7,61	192
<b>75</b>	6,97	212
<b>76</b>	7,12	198
<b>77</b>	7,19	202
<b>78</b>	6,95	201
<b>79</b>	6,90	218

<b>80</b>	7,17	201
<b>81</b>	7,77	215
<b>82</b>	7,43	222
<b>83</b>	7,36	205
<b>84</b>	7,70	205
<b>85</b>	7,41	195
<b>86</b>	7,01	216
<b>87</b>	7,61	214
<b>88</b>	7,74	204
<b>89</b>	7,52	204
<b>90</b>	6,99	190
<b>91</b>	8,17	193
<b>92</b>	6,84	201
<b>93</b>	7,32	200
<b>94</b>	7,35	210
<b>95</b>	7,14	186
<b>96</b>	7,43	218
<b>97</b>	7,19	217
<b>98</b>	7,52	221
<b>99</b>	7,17	211
<b>100</b>	7,57	197
<b>101</b>	7,20	211
<b>102</b>	7,25	194
<b>103</b>	7,37	200
<b>104</b>	7,03	218
<b>105</b>	7,45	215
<b>106</b>	7,46	208
<b>107</b>	7,70	199
<b>108</b>	7,11	203
<b>109</b>	6,89	201
<b>110</b>	7,18	212
<b>111</b>	7,69	219
<b>112</b>	7,79	224
<b>113</b>	7,31	223
<b>114</b>	8,05	203
<b>115</b>	7,17	195
<b>116</b>	8,03	201
<b>117</b>	6,92	219
<b>118</b>	7,18	213
<b>119</b>	6,72	219
<b>120</b>	7,82	196
<b>121</b>	7,84	201
<b>122</b>	7,01	218

<b>123</b>	7,20	203
<b>124</b>	8,10	203
<b>125</b>	7,82	218
<b>126</b>	7,37	202
<b>127</b>	7,37	203
<b>128</b>	7,43	214
<b>129</b>	7,34	219
<b>130</b>	7,52	198
<b>131</b>	7,49	204
<b>132</b>	7,16	230
<b>133</b>	7,27	214
<b>134</b>	7,47	200
<b>135</b>	8,01	217
<b>136</b>	7,64	201
<b>137</b>	7,46	196
<b>138</b>	7,21	203
<b>139</b>	7,14	211
<b>140</b>	7,61	210
<b>141</b>	6,85	209
<b>142</b>	7,47	206
<b>143</b>	6,97	230
<b>144</b>	7,68	198
<b>145</b>	6,84	212
<b>146</b>	7,52	194
<b>147</b>	7,37	202
<b>148</b>	7,25	212
<b>149</b>	7,80	203
<b>150</b>	7,74	219

Nota. Tabla de recolección de datos de velocidad de lectura y escr

## 6. Datos recopilados de recuperación de disco

<b>Prueba</b>	<b>Recuperación</b>	<b>Descripción</b>
<b>1</b>	no	Recuperación sin éxito
<b>2</b>	sí	Recuperación exitosa
<b>3</b>	sí	Recuperación exitosa
<b>4</b>	sí	Recuperación exitosa
<b>5</b>	sí	Recuperación exitosa
<b>6</b>	sí	Recuperación exitosa
<b>7</b>	sí	Recuperación exitosa
<b>8</b>	sí	Recuperación exitosa
<b>9</b>	sí	Recuperación exitosa
<b>10</b>	sí	Recuperación exitosa

<b>11</b>	sí	Recuperación exitosa
<b>12</b>	sí	Recuperación exitosa
<b>13</b>	sí	Recuperación exitosa
<b>14</b>	sí	Recuperación exitosa
<b>15</b>	sí	Recuperación exitosa
<b>16</b>	sí	Recuperación exitosa
<b>17</b>	sí	Recuperación exitosa
<b>18</b>	sí	Recuperación exitosa
<b>19</b>	sí	Recuperación exitosa
<b>20</b>	sí	Recuperación exitosa
<b>21</b>	no	Recuperación sin éxito
<b>22</b>	sí	Recuperación exitosa
<b>23</b>	sí	Recuperación exitosa
<b>24</b>	sí	Recuperación exitosa
<b>25</b>	sí	Recuperación exitosa
<b>26</b>	sí	Recuperación exitosa
<b>27</b>	sí	Recuperación exitosa
<b>28</b>	no	Recuperación sin éxito
<b>29</b>	sí	Recuperación exitosa
<b>30</b>	sí	Recuperación exitosa
<b>31</b>	no	Recuperación sin éxito
<b>32</b>	no	Recuperación sin éxito
<b>33</b>	sí	Recuperación exitosa
<b>34</b>	sí	Recuperación exitosa
<b>35</b>	sí	Recuperación exitosa
<b>36</b>	sí	Recuperación exitosa
<b>37</b>	sí	Recuperación exitosa
<b>38</b>	sí	Recuperación exitosa
<b>39</b>	no	Recuperación sin éxito
<b>40</b>	sí	Recuperación exitosa
<b>41</b>	no	Recuperación sin éxito
<b>42</b>	sí	Recuperación exitosa
<b>43</b>	sí	Recuperación exitosa
<b>44</b>	sí	Recuperación exitosa
<b>45</b>	sí	Recuperación exitosa
<b>46</b>	no	Recuperación sin éxito
<b>47</b>	no	Recuperación sin éxito
<b>48</b>	sí	Recuperación exitosa
<b>49</b>	no	Recuperación sin éxito
<b>50</b>	sí	Recuperación exitosa
<b>51</b>	no	Recuperación sin éxito
<b>52</b>	sí	Recuperación exitosa
<b>53</b>	sí	Recuperación exitosa

<b>54</b>	sí	Recuperación exitosa
<b>55</b>	no	Recuperación sin éxito
<b>56</b>	sí	Recuperación exitosa
<b>57</b>	sí	Recuperación exitosa
<b>58</b>	sí	Recuperación exitosa
<b>59</b>	sí	Recuperación exitosa
<b>60</b>	sí	Recuperación exitosa
<b>61</b>	sí	Recuperación exitosa
<b>62</b>	sí	Recuperación exitosa
<b>63</b>	sí	Recuperación exitosa
<b>64</b>	sí	Recuperación exitosa
<b>65</b>	sí	Recuperación exitosa
<b>66</b>	sí	Recuperación exitosa
<b>67</b>	sí	Recuperación exitosa
<b>68</b>	sí	Recuperación exitosa
<b>69</b>	no	Recuperación sin éxito
<b>70</b>	sí	Recuperación exitosa
<b>71</b>	sí	Recuperación exitosa
<b>72</b>	sí	Recuperación exitosa
<b>73</b>	sí	Recuperación exitosa
<b>74</b>	sí	Recuperación exitosa
<b>75</b>	sí	Recuperación exitosa
<b>76</b>	sí	Recuperación exitosa
<b>77</b>	no	Recuperación sin éxito
<b>78</b>	sí	Recuperación exitosa
<b>79</b>	no	Recuperación sin éxito
<b>80</b>	no	Recuperación sin éxito
<b>81</b>	no	Recuperación sin éxito
<b>82</b>	sí	Recuperación exitosa
<b>83</b>	sí	Recuperación exitosa
<b>84</b>	sí	Recuperación exitosa
<b>85</b>	sí	Recuperación exitosa
<b>86</b>	sí	Recuperación exitosa
<b>87</b>	sí	Recuperación exitosa
<b>88</b>	sí	Recuperación exitosa
<b>89</b>	sí	Recuperación exitosa
<b>90</b>	sí	Recuperación exitosa
<b>91</b>	sí	Recuperación exitosa
<b>92</b>	no	Recuperación sin éxito
<b>93</b>	sí	Recuperación exitosa
<b>94</b>	sí	Recuperación exitosa
<b>95</b>	sí	Recuperación exitosa
<b>96</b>	sí	Recuperación exitosa

<b>97</b>	sí	Recuperación exitosa
<b>98</b>	sí	Recuperación exitosa
<b>99</b>	sí	Recuperación exitosa
<b>100</b>	sí	Recuperación exitosa
<b>101</b>	sí	Recuperación exitosa
<b>102</b>	sí	Recuperación exitosa
<b>103</b>	sí	Recuperación exitosa
<b>104</b>	sí	Recuperación exitosa
<b>105</b>	sí	Recuperación exitosa
<b>106</b>	sí	Recuperación exitosa
<b>107</b>	sí	Recuperación exitosa
<b>108</b>	sí	Recuperación exitosa
<b>109</b>	sí	Recuperación exitosa
<b>110</b>	sí	Recuperación exitosa
<b>111</b>	sí	Recuperación exitosa
<b>112</b>	sí	Recuperación exitosa
<b>113</b>	sí	Recuperación exitosa
<b>114</b>	sí	Recuperación exitosa
<b>115</b>	sí	Recuperación exitosa
<b>116</b>	sí	Recuperación exitosa
<b>117</b>	sí	Recuperación exitosa
<b>118</b>	sí	Recuperación exitosa
<b>119</b>	sí	Recuperación exitosa
<b>120</b>	no	Recuperación sin éxito
<b>121</b>	sí	Recuperación exitosa
<b>122</b>	sí	Recuperación exitosa
<b>123</b>	sí	Recuperación exitosa
<b>124</b>	sí	Recuperación exitosa
<b>125</b>	sí	Recuperación exitosa
<b>126</b>	sí	Recuperación exitosa
<b>127</b>	sí	Recuperación exitosa
<b>128</b>	no	Recuperación sin éxito
<b>129</b>	no	Recuperación sin éxito
<b>130</b>	sí	Recuperación exitosa
<b>131</b>	sí	Recuperación exitosa
<b>132</b>	sí	Recuperación exitosa
<b>133</b>	no	Recuperación sin éxito
<b>134</b>	sí	Recuperación exitosa
<b>135</b>	no	Recuperación sin éxito
<b>136</b>	sí	Recuperación exitosa
<b>137</b>	sí	Recuperación exitosa
<b>138</b>	sí	Recuperación exitosa
<b>139</b>	sí	Recuperación exitosa

<b>140</b>	sí	Recuperación exitosa
<b>141</b>	sí	Recuperación exitosa
<b>142</b>	sí	Recuperación exitosa
<b>143</b>	no	Recuperación sin éxito
<b>144</b>	sí	Recuperación exitosa
<b>145</b>	no	Recuperación sin éxito
<b>146</b>	sí	Recuperación exitosa
<b>147</b>	sí	Recuperación exitosa
<b>148</b>	sí	Recuperación exitosa
<b>149</b>	sí	Recuperación exitosa
<b>150</b>	sí	Recuperación exitosa

*Nota. Tabla de recolección de datos respecto a la recuperación de datos en el modelo raid 5.*

*Fuente: Elaboración propia (2023).*