UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

La Universidad Católica de Loja

MODALIDAD PRESENCIAL

ESCUELA DE CIENCIAS DE LA COMPUTACIÓN

SEGURIDAD EN CLÚSTER DE SERVICIOS DE MAIL Y WEB

Trabajo de fin de carrera previa a la obtención del título de Ingeniero en Sistemas Informáticos y Computación.

AUTOR:

Cuenca Macas Luis Alberto

DIRECTORA:

Ing. Cueva Carrión Samanta Patricia

Loja – Ecuador

2011

CERTIFICACIÓN

Ing. Samanta Patricia Cueva Carrión

DIRECTORA DE TESIS

C E R T I F I C A:

Haber dirigido y supervisado el desarrollo del presente proyecto de tesis previo a la obtención del título de **INGENIERO EN SISTEMAS INFORMÁTICOS Y COMPUTACIÓN**, y una vez que este cumple con todas las exigencias y los requisitos legales establecidos por la Universidad Técnica Particular de Loja, autoriza su presentación para los fines legales pertinentes.

Loja, 3 de octubre del 2011

_____

Samanta Patricia Cueva Carrión

DIRECTORA DE TESIS

CESIÓN DE DERECHOS

Yo, **Luis Alberto Cuenca Macas**, declaro ser autor del presente trabajo y eximo expresamente a la Universidad Técnica Particular de Loja y a sus representantes legales de posibles reclamos o acciones legales.

Adicionalmente declaro conocer y aceptar la disposición del Art. 67 del Estatuto Orgánico de la Universidad Técnica Particular de Loja, que en su parte pertinente textualmente dice: "Forman parte del patrimonio de la Universidad la propiedad intelectual de investigaciones, trabajos científicos o técnicos y tesis de grado que se realicen a través o con el apoyo financiero, académico o institucional (operativo) de la Universidad".

_____

Luis Alberto Cuenca Macas

AUTORÍA

El presente proyecto de tesis con cada una de sus observaciones, análisis, evaluaciones, conclusiones y recomendaciones emitidas, es de absoluta responsabilidad del autor.

Además, es necesario indicar que la información de otros autores empleada en el presente trabajo está debidamente especificada en fuentes de referencia y apartados bibliográficos.

_____

Luis Alberto Cuenca Macas

DEDICATORIA

Con mucho amor y cariño dedico el presente proyecto:

A Dios, por el don de la vida

A mis padres por su amor y apoyo incondicional,

A mi hermano Héctor que me brindo su ayuda incondicional en todo momento.

A Mónica, quien siempre me apoyo en cada momento.

A mis compañeros y amigos, por brindarme su amistad sincera.

Luis

AGRADECIMIENTO

Mi sincero agradecimiento primeramente a Dios por su compañía y bendiciones en todos los momentos de mi vida.

A mis queridos padres quienes me apoyaron incondicionalmente para alcanzar mis metas e ideales planteados.

A la Ing. Samanta Cueva, directora de tesis, quien con sus conocimientos, su orientación, motivación y paciencia brindada, me permitió llegar a la culminación de esta tesis.

A todos quienes conforman la UTPL, en especial al Grupo de Telecomunicaciones, por contribuir al desarrollo de la presente tesis.

A mis compañeros y amigos ya que siempre he recibido motivación y apoyo incondicional.

Luis

ÍNDICE DE CONTENIDOS

Resumen

La seguridad es un término general que interviene en áreas como: la computación y en el procesamiento de la información. En la actualidad las organizaciones dependen a diario de los sistemas computacionales y de las redes para ejecutar sus operaciones y transacciones corporativas, considerando a los datos como un recurso importante dentro de la organización.

Para brindar servicios confiables a los usuarios se hace uso de clústeres para el manejo de la información, pero se debe tener en cuenta que la información que maneja el clúster debe mantener la integridad, asegurando la confiabilidad de datos.

En el presente documento se detalla cada una de las configuraciones a usar en los servicios de web y mail para asegurar la información de los servicios, así como también se detalla la selección de un sistema de replicación de archivos para asegurar la información que fluye a través del clúster de servicios web y mail.

Introducción

En la actualidad se realiza alguna implementación de seguridad después de que ha ocurrido o ha habido alguna intrusión no autorizada en los sistemas.

La seguridad es un término general que cubre una gran área de computación y procesamiento de la información. En la actualidad las organizaciones dependen de sistemas computarizados y redes para ejecutar sus operaciones y transacciones de negocios diarias, consideran sus datos como una parte importante de sus activos generales.

Se hace uso de los clúster para tener servicios confiables, pero para ello es necesario que la información que este en el clúster tenga total integridad y así contar con servicios seguros y confiables.

Además, la protección de la información hoy en día es muy importante, ya que así se puede contar con servicios íntegros y de buena calidad, para lograr esto es necesaria la búsqueda de nuevas tecnologías que permitan la compartición de recursos, eso teniendo en cuenta que la información que fluya sobre estos este protegida.

En la UTPL se cuenta con una implementación de un clúster de balanceo de carga y alta disponibilidad para los servicios de Web y Mail, la cual ofrece:

- Distribuir la carga de trabajo a los servidores reales
- Mantener un servidor de backup que se levante cuando falle el clúster
- Ofrecer alto rendimiento, mediante el servicio ininterrumpido.
- Aprovechar al máximo los recursos existentes.
- Disponer de escalabilidad es decir permitir el incremento del número de nodos sin interrumpir el normal funcionamiento del servicio.

Si bien es cierto la implementación de este clúster es de gran ayuda para brindar los servicios de web y mail, se ha visto la necesidad de asegurar dicho clúster, así como también proteger y encriptar la información, ya que a través de este clúster fluye información confidencial que debe ser manejada de buena manera, para así poder contar con servicios que sean fiables, seguros y de calidad.

CAPÍTULO I: Análisis de la Situación Actual del clúster de servicios de Mail y Web

En este capítulo se realiza una descripción del clúster de servicios de mail y web, la cual nos ayuda a tener un mejor conocimiento acerca de todas las características con las que se cuenta.

Además se analizan las distintas vulnerabilidades que se presentaron en el clúster de servicios de mail y web, tanto en la configuración de los servicios, así como también de la forma de cómo se transmite la información entre los diferentes nodos que lo componen.

1.1. Servidor Mail

Su principal función es el envío y recepción de correo electrónico, el servidor tiene la capacidad de alojar una gran cantidad de cuentas de correo electrónico, utiliza como MTA el Sendmail.

El Servidor mail de la UTPL cuenta con aproximadamente 25000 cuentas de correo las cuales están asignadas a: autoridades, personal administrativo, docentes y estudiantes tanto de la modalidad clásica como la de distancia.

Información Servidor Mail

Nombre del servidor:	gdr3.utpl.edu.ec
Dirección IP:	172.16.X7.1XX
Sistema operativo:	Centos 5.6
Versión de kernel:	2.6.18-238

El servidor Mail posee con las siguientes características de software:

Aplicaciones que corren en el servidor:

Sendmail: Es un MTA[1] que tiene como objetivo encaminar los mensajes de correos de tal manera que estos lleguen a su destino.

POP3: Post Office Protocol 3 es configurado en el servidor para que los usuarios puedan obtener los mensajes de correo electrónico almacenados en un servidor, a través de un cliente para correo electrónico como por Ej. Outlook (Windows), Evolution (Linux), Mail (MAC) entre otros.

IMAP: Es un Protocolo de Acceso a Mensajes en Internet almacenados en un servidor. Mediante IMAP se puede tener acceso al correo electrónico desde cualquier equipo que tenga una conexión a Internet.

MailScanner: Es un sistema que permite realizar un análisis a los correos electrónicos, lo que permitirá identificar si se tratan de spam, también permite configurar opciones para realizar un análisis en busca de virus, lo hace utilizando un determinado antivirus.

[1] **Mail Transport Agent** (Agente de Transporte de Correos) y también Message Transport Agent (Agente de Transporte de Mensajes).

Procmail: Es un MDA[2] que se utiliza para filtrar el correo de entrada. Permite definir "**reglas**" que se asocian con correos entrantes y que realizan funciones concretas, como reencaminar el correo a carpetas o direcciones alternativas. En el servidor se lo emplea para filtro de correos SPAM[3].

NET-SNMP: (Simple Network Management Protocol) es un protocolo ampliamente utilizado en la administración de redes para supervisar la estabilidad del equipo de la red, equipo de cómputo y otros dispositivos.

Webmin: Es una interfaz web que permite la administración remota de sistemas UNIX[4], la cual permite administrar cuentas de usuario, compartir archivos, hacer un shutdown, configuraciones, todo esto de una manera que le resulta fácil de usar al administrador.

Clamav: Es una herramienta antivirus para UNIX, se encarga de hacer un análisis en busca de virus en los adjuntos de los correos electrónicos.

El software presente en los equipos del clúster es:

Software	Aplicación
Sistema Operativo	**Centos 5.6**
Para la Administración del Clúster	**Piranha**
Para Monitoreo del servidor	**Ganglia**
Para gestión de archivos de configuración	**Editores de texto**

Tabla 1.1 Software del clúster

1.2. Hardware del Servidor Mail

En la UTPL se cuenta con un equipo clúster con tecnología Blade[5], la cual ofrece un alto rendimiento ya que está especializada en multiprocesamiento.

Los componentes críticos de un servidor Blade pueden ser transformados en redundantes, ya que permiten el cambio en caliente, de los sistemas de refrigeración, alimentación, controladores y conmutadores Ethernet, unidades de disco rígido y procesadores de servicios.

Clúster Tecnología Blade

Dispositivos	Número	Marcas	Capacidad	Velocidad
Memoria			3GB	266 MHz
Procesador	1	Intel ® Xeon		1.6 GHz
Discos	2	IBM	36 GB	7200 RPM
			860 GB	

Tabla 1.2. Hardware del servidor gdr3

[2] **Mail Delivery Agent** (Agente de Entrega de Correo) es un software que entrega los mensajes de e-mail inmediatamente después de que haya sido aceptado en un servidor
[3] Se llama **SPAM**, **correo basura** a todos los mensajes no solicitados que son enviados en masa.
[4] **UNIX:** Sistema operativo multitarea y multiusuario
[5] **Servidores Blade:** Son servidores finos, con características de cambio en caliente de componentes, caben en un único chasis cada uno de ellos es un servidor independiente, es decir manejan sus propios procesadores, memoria, almacenamiento, controladores de red, sistema operacional y aplicativos.

Nodos Blade

Dispositivos	Número	Marcas	Capacidad	Velocidad
Memoria	1		3GB	266 MHz
Procesador	1	Intel CoreDuo		3.6 GHz
Tarjeta	2	Broadcom		1000 baseTx
Discos	1 a 3	SATA	250MB	

Tabla 1.3 Hardware del servidor gdr3

Los equipos que integrarán el clúster son:

Un servidor maestro

Dos servidores para servicio Mail.

El nodo maestro está incorporado con dos interfaces de red para su adecuada configuración y funcionamiento. Además cuentan con componentes de altas prestaciones para su normal funcionamiento.

Los elementos con los que cuenta el clúster son:

- Un nodo activo: donde corren los servicios
- Dos servidores reales, destinados para servicio SMTP[6].
- Software de Administración

1.3. Servidor Web

La función de un servidor web es la de alojar páginas web. Actualmente en la UTPL se utiliza APACHE como servidor web ya que este es uno de los más confiables y seguros.

Información Servidor Mail

Nombre del servidor: gdr1.utpl.edu.ec

Dirección IP: 200.0.X1.X6

Sistema operativo: Centos 5.6

Versión de kernel: 2.6.18-238

Actualmente el servidor Web cuenta con las siguientes características de software:

Aplicaciones que corren en el servidor:

Apache: Es un software (libre) de código abierto para plataformas Unix (BSD, GNU/Linux, etc.), Windows, Macintosh y otras. Es uno de los servidores Web más populares y utilizados en la actualidad dado el grado de seguridad que este posee.

[6] **Simple Mail Transport Protocol** (Protocolo de Correo Simple) protocolo que se utiliza para la transferencia de correo electrónico en el Internet.

MySql: Es un sistema de gestión de base de datos relacional, multihilo y multiusuario, es uno de los más popular, desarrollado por MySQL AB.

PHP: Es un lenguaje interpretado de propósito general ampliamente usado y que está diseñado especialmente para desarrollo web y puede ser embebido dentro de código HTML. Generalmente se ejecuta en un servidor web, tomando el código en PHP como su entrada y creando páginas web como salida. Puede ser desplegado en la mayoría de los servidores web y en casi todos los sistemas operativos y plataformas.

1.4. Hardware del Servidor Web

Dispositivo	Número	Marca	Capacidad	Velocidad
Memoria	1	---	1 GB	---
Memoria swap[7]	---	---	3 GB	---
Procesador	3	Genuine Intel	---	1.6 GHz
Discos	3	IBM	70 GB	---
			70 GB	---
			230 GB	---

Tabla 1.4 Hardware del servidor gdr1

1.5. Diseño físico del Clúster

El clúster está conformado por un servidor activo como balanceador de carga y dos servidores reales

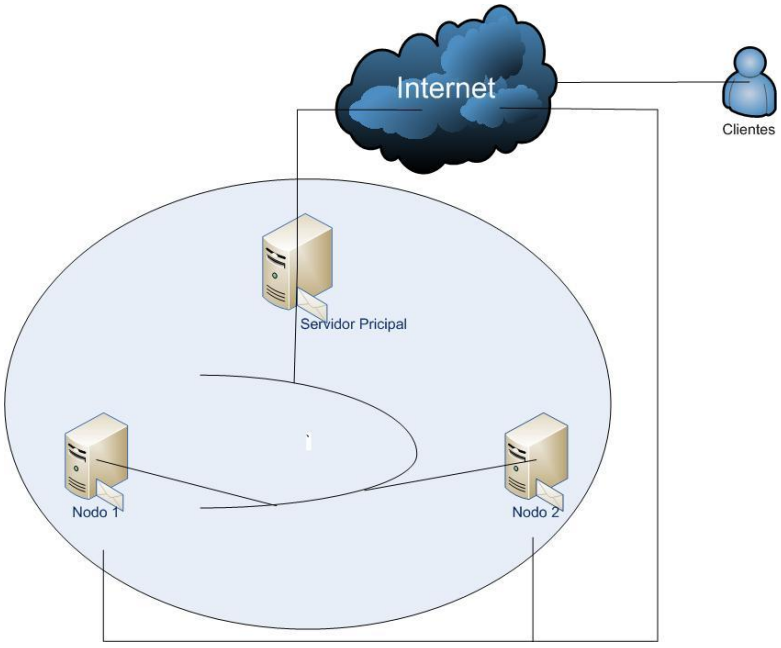


Fig. 1.1 Clúster SMTP con tres Servidores Reales

[7] **swap:** También conocido como memoria virtual, permite utilizar parte del disco duro como memoria RAM.

1.6. Topología del clúster

Para el direccionamiento en el clúster se cuenta con la topología **Direct Routing**, la cual se describe a continuación.

1.6.1. Topología Direct Routing

Este requiere que todos los servidores estén en el mismo segmento físico de red que el balanceador. Este método es el que menos sobrecarga impone al equipo balanceador ya que no tiene la necesidad de reescribir los paquetes (NAT[8]), ni encapsularlos.

El balanceador no se presenta como cuello de botella ya que a través de él pasará únicamente el tráfico en dirección de los clientes al clúster, mientras que el tráfico de salida lo dirigirán directamente los servidores a cada cliente.

Como todos los equipos tendrán configurado un interfaz con la IP pública del clúster, el balanceador hace de punto de entrada al clúster; el resto de equipos estarán conectados al balanceador en la misma red física y en la interfaz conectada a esta red tendrán configurada la IP pública del clúster, pero configurando esta interfaz para que no responda a comandos ARP[9] para de esta manera no interferir con otros protocolos.

Cuando llega una petición al balanceador éste decide a qué servidor enviársela y redirige el paquete a la dirección MAC[10] del servidor elegido. Cuando llega al servidor con destino y como este también tiene configurada la IP pública del clúster, acepta el paquete y genera la respuesta, que enviará directamente al cliente

1.7. Gestión de información del clúster

Como la mensajería electrónica se basa en archivos que cambian constantemente en el servidor de correos, es por ello que se cuenta con un sistema de compartición de recursos NFS[11] ya que así las peticiones de los clientes pueden ser atendidas por cualquiera de los servidores reales y estos podrán ser modificadas en el recurso compartido.

Aquí algunas características del sistema de compartición de recursos NFS que se encuentra implementado actualmente en el sistema.

Características:

- Posee Licencia GPL[12]

[8] **Network Address Translation** (**Traducción de Dirección de Red)** es un mecanismo utilizado por routers IP para intercambiar paquetes entre dos redes que se asignan mutuamente direcciones incompatibles.
[9] **Address Resolution Protocol** (Protocolo de Resolución de dirección) protocolo que provee una dirección a los equipos de que se encuentran en una red local
[10] **Media Access Control Address** es un identificador hexadecimal de 48 bits que se corresponde de forma única con una tarjeta o interfaz de red.
[11] **NFS**: Network File System (Sistema de Archivo de Red)
[12] **GPL**.- (General Public License) Licencia General Publica.

- Varios clientes y servidores tienen la posibilidad de compartir un sistema de archivos común, en este caso el directorio donde se encuentran los archivos de correo de cada uno de los usuarios del servicio de mail.
- Está disponible para la mayoría de las distribuciones de Linux
- Se tiene acceso desde LAN o redes más amplias.
- **NFS** utiliza el mecanismo de protección de *UNIX*, con los bits **rwx.** En **NFS** cada máquina sea cliente y servidor en el mismo instante
- Requiere el demonio portmap

Ventajas

- Su implementación es fácil, así como también la utilización.
- No se necesita de cambio alguno en el sistema de fichero.
- No requiere Hardware especial.
- Flexibilidad de adaptación y escalabilidad

Desventajas

- Seguridad baja, requiere un Firewall.
- Velocidad de acceso baja con respecto a otros esquemas especializados en clustering.
- Su rendimiento se ve afectado al agregar más nodos al clúster.
- Se presenta un punto único de falla si no se cuenta con un servidor de backup para el servidor NFS

1.8. Vulnerabilidades en la configuración del servicio de mail

En la configuración que actualmente la UTPL tiene para el clúster de servicio de mail se identificaron vulnerabilidades, que afectan en el servicio y en la integridad de la información.

A continuación se detalla cada una de las vulnerabilidades:

1.8.1. Cifrado de la mensajería

Los correos que son vistos tanto en la aplicación web como en cualquier cliente de correo, pasan a través de la red en texto plano, lo que implica una grave falla de seguridad ya que estos pueden ser capturados por personas malintencionadas.

En el Anexo A, se puede observar esta vulnerabilidad.

1.8.2. Autenticación para el correo saliente

En la configuración actual no se cuenta con una autenticación para el correo saliente, es decir que para enviar un mail desde una cuenta de correo no es necesario que el usuario se autentique, lo que conlleva a tener una desconfianza del remitente de correo.

Esta vulnerabilidad puede ser usada por un usuario malintencionado ya que puede enviar correos electrónicos tomando el nombre de otros (conocido como suplantación de identidad).

1.8.3. Información que se transmite en el clúster

Actualmente se usa NFS para transmitir la información del clúster hacia los nodos, este sistema no es el adecuado ya que presenta algunas fallas que a continuación se las ira detallando.

NFS, no posee un buen sistema de seguridad ya que esta tarea la deja al cliente.

NFS exporta los directorios hacia determinados dominios, lo que permitiría que un atacante que tenga el control del DNS pueda fácilmente acceder a toda la información que se encuentra compartida.

La sintaxis para escribir la configuración de NFS no es estricta, dejando la posibilidad de que un espacio mal ubicado en el archivo de configuración permita que tal información pueda ser compartida con todo mundo.

1.9. Vulnerabilidades en la configuración del servicio web

En la configuración que actualmente la UTPL tiene para el clúster de servicio de web se identificó una vulnerabilidad, que afecta en la confidencialidad de la información.

A continuación se detalla esta vulnerabilidad:

1.9.1. Cifrado de las páginas de login

Todas las páginas que se utilizan para: ingresar al Entorno Virtual, al sitio ftp no están protegidas, la información que fluyen por estas páginas pasa en texto plano y pueden ser capturadas.

Esto implica un gran problema con la confidencialidad de la información ya que al pasar en texto plano, estos pueden ser capturados por personas malintencionadas para obtener información de nombres de usuario y password.

En el Anexo B, se puede observar esta vulnerabilidad.

CAPÍTULO II: Análisis de los Requerimientos de seguridad y protección de información del clúster de servicios de mail y web

En el presente capítulo se realiza todo el análisis referente a los requerimientos de seguridad y protección de la información que se transmite en el clúster se servicios de mail y web.

En la etapa anterior se revisó varias de las vulnerabilidades que se encuentran presentes en la configuración del clúster, así como también que se cuenta con una configuración de un servicio NFS para compartir los archivos del servidor mail a los nodos del clúster.

Es por ello que se han analizado los siguientes aspectos para lograr una protección e integridad de los datos.

2.1. Asegurar Portmap

El servicio portmap es un demonio de asignación de puertos dinámicos para servicios RPC[13], tales como NIS[14] y NFS. Tiene mecanismos de autenticación débiles y la habilidad de asignar un amplio rango de puertos para los servicios que controla. Por estas razones si está ejecutando servicios RPC, se debería seguir algunas reglas básicas como:

2.1.1. Proteger portmap con TCP Wrappers[15]

TCP Wrappers: permite controlar y proteger los servicios de red, limitando el acceso como sea posible y registrado todas las conexiones para hacer el trabajo de detectar y resolver problemas de forma más fácil.

Los Wrappers son muy utilizados, es por ello que han llegado a formar parte de las herramientas de seguridad por las siguientes razones:

- ✓ La seguridad lógica está concentrada en un solo programa, los wrappers son fáciles y simples de validar.
- ✓ Los wrappers llaman al programa protegido mediante la llamada al sistema estándar exec(), entonces se puede usar un solo wrapper para controlar el acceso a diversos programas que se necesiten proteger.

Por este motivo es importante utilizar TCP wrappers para limitar las redes o máquinas que tienen acceso al servicio portmap puesto que éste no posee autenticación incorporada.

[13] **RPC** (*Remote Procedure Call*, **Llamada a Procedimiento Remoto**) es un protocolo que permite a un programa de ordenador ejecutar código en otra máquina remota sin tener que preocuparse por las comunicaciones entre ambos. El protocolo es un gran avance sobre los sockets usados hasta el momento.
[14] **NIS:** *(Network Information System)* Sistema de información de red.
[15] **TCP Wrappers:** Herramienta simple que sirve para monitorear y controlar el tráfico que llega por la red

Además, solo se debe utilizar direcciones IP cuando se esté limitando el acceso al servicio. Se debe evitar los nombres de hosts ya que estos pueden ser falsificados a través de envenenamiento de DNS[16] y otros métodos.

TCP-Wrappers se compone de 5 programas:

- ✓ **tcpd.** Es el demonio del TCP-Wrappers.
- ✓ **tcpdmatch.** Predice como el tcpd manejaría una petición en específico.
- ✓ **tcpdchk.** Verifica las reglas de control de acceso contenidas en los archivos /etc/hosts.allow y /etc/hosts.deny.
- ✓ **safe-finger.** Versión de finger para implementar el finger reversivo.
- ✓ **try-from.** Programa que permite probar si el sistema es capaz de reconocer qué máquina la está contactando.

2.2. Aseguramiento de la gestión de información del clúster

En el capítulo anterior se realizó un análisis sobre las fallas de seguridad que existen en la gestión de la información, por tal motivo aquí se presentan dos alternativas para dar solución a este problema.

2.2.1. La primera alternativa

Cambiar NFS por un sistema de replicación de archivos que trabaje exclusivamente bajo un ambiente de clúster.

NFS como tal no fue diseñado para trabajar en un ambiente de clúster, además de esto presenta muchas vulnerabilidades que pueden ser aprovechadas por personas malintencionadas, no posee un mecanismo de autenticación y encriptación de los datos que se comparten ya que NFS lo realiza de una manera plana, es por ello que se decidió cambiar NFS por un sistema de replicación de archivos.

En el capítulo tres se realiza un análisis completo de tres sistemas de replicación de archivos que son: GlusterFS, CODA y OpenGFS, en la cual se detallan los beneficios, ventajas y desventajas que se presentan en cada uno, así como también las características de cada uno de ellos.

2.2.2. La segunda alternativa.

Asegurar NFS, a continuación se detalla algunas características de NFS que se deben tomar en cuenta para asegurar el sistema NFS.

[16] **DNS** (Domain Name Service) Servicio de nombres de dominios, servicio del Internet que traduce los nombres de los dominios (gov, edu, net, etc.) en direcciones IP (direcciones numéricas)

NFS: (Network File System) Sistema de Archivos de Red es un servicio RPC que se usa conjuntamente con portmap y otros servicios relacionados para proporcionar sistemas de archivos accesibles a través de la red a las máquinas clientes.

Las características a tomar en cuenta para asegurar NFS son:

2.2.2.1. Planificación adecuada de la red

Debido a que NFS transmite la información sin encriptar sobre la red, es importante que el servicio sea ejecutado detrás de un firewall y en un segmento de red seguro. Cada vez que se transmita la información sobre NFS en una red insegura, hay riesgos de que la información sea interceptada. Un diseño cuidadoso en este aspecto puede ayudar a prevenir aperturas de la seguridad.

2.2.2.2. Errores sintácticos

El servidor NFS determina cuáles sistemas de archivos exportar y a que máquinas exportan los directorios a través del archivo /etc/exports. Se debe tener cuidado con la sintaxis de este archivo, así como también el de no añadir espacios adicionales cuando se esté editando este archivo.

Por ejemplo, la línea siguiente en el archivo /etc/exports, se está especificando que se va a compartir el directorio /tmp/nfs/ a la máquina gdr3.utpl.edu.ec con permisos de lectura y escritura.

/tmp/nfs/ gdr3.utpl.edu.ec(rw)

Por otro lado, esta línea en el archivo /etc/exports, comparte el mismo directorio a la máquina gdr3.utpl.edu.ec con permisos de sólo lectura y lo comparte con todo el mundo con permisos de lectura y escritura debido a un espacio en blanco luego del nombre de la máquina.

/tmp/nfs/ gdr3.utpl.edu.ec (rw)

Se debe siempre de verificar cualquier directorio compartido NFS usando el comando show mount para verificar que está siendo compartido:

show mount –e gdr3.utpl.edu.ec

2.2.2.3. Opción no_root_squash

Cuando se está utilizando los directorios compartidos, NFS por defecto se cambia el usuario root por el usuario nfsnobody, una cuenta de usuario sin privilegios. Así, todos los archivos creados por root son propiedad del

usuario nfsnobody, lo que previene la carga de programas con la configuración del bit seguridad.

Si se utiliza no_root_squash, los usuarios remotos podrán cambiar cualquier archivo en el sistema de archivos compartido y dejar aplicaciones con troyanos para que otros usuarios las ejecuten inadvertidamente.

Por tal motivo no es recomendable utilizar la opción **no_root_squash**

2.2.2.4. NFS y Sendmail

No es recomendado colocar el directorio de correos, /var/spool/mail/, en un volumen compartido NFS.

Debido a que NFS no mantiene un control sobre usuarios e IDs de grupos, dos o más usuarios pueden tener el mismo UID y por tanto recibir y leer los correos electrónicos de otros.

2.3. Asegurando Sendmail

Sendmail es un Agente de transporte de correos (MTA) que utiliza el protocolo de transporte de correos simple (SMTP[17]) para entregar mensajes electrónicos entre otros MTA y a los clientes de correo o agentes de entrega.

Aun cuando muchos MTAs son capaces de encriptar el tráfico entre unos y otros, la mayoría no lo hacen, por tanto el envío de correos electrónicos sobre redes públicas es considerado una forma insegura de comunicación.

Por tal motivo se recomienda tomar en cuenta los siguientes aspectos.

2.3.1. Limitar los ataques de rechazo de servicio

Debido a la naturaleza del correo electrónico, un atacante determinado puede inundar fácilmente el servidor con correos y de esta manera causar un rechazo de servicio. Se puede limitar la efectividad de tales ataques mediante la colocación de límites a las siguientes directivas a /etc/mail/sendmail.mc.

- ✓ **confCONNECTION_RATE_THROTTLE**
  El número de conexiones que el servidor puede recibir por segundo. Por defecto, Sendmail no limita el número de conexiones. Si se establece un límite y este es alcanzado, las conexiones siguientes son retrasadas.

- ✓ **confMAX_DAEMON_CHILDREN**

  El máximo número de procesos hijo que se pueden producir por el servidor. Por defecto, Sendmail no asigna un límite al

[17] **SMTP:** Simple Mail Transport Protocol

número de procesos hijos. Si se coloca un límite y este es alcanzado, las conexiones siguientes son retrasadas.

- ✓ **confMIN_FREE_BLOCKS**

  El número mínimo de bloques libres que debe haber disponible para que el servidor acepte correos. Por defecto es 100 bloques.

- ✓ **confMAX_HEADERS_LENGTH**

  El tamaño máximo aceptable (en bytes) para la cabecera de un mensaje.

- ✓ **confMAX_MESSAGE_SIZE**

  El tamaño máximo aceptable (en bytes) para cualquier mensaje.

En el Anexo C se encuentra la configuración actual que está en el servidor mail y la que se propone en el presente trabajo.

2.3.2. Configurar servicios pop3 e imap seguros

Este punto es muy importante si se quiere que los correos no sean visibles por cualquier intruso y solo puedan ser visibles por el usuario que es dueño de la cuenta y por el usuario que es destinario de tales correos, es por ello que se recomienda configurar los servicios pop3s e imaps, con lo cual se evita que terceras personas puedan interceptar los mensajes de correo electrónico y obtengan la información de los mismos.

2.3.3. Autenticación en el SMTP saliente

Es recomendable implementar una configuración para la autenticación del SMTP saliente de correo electrónico, para evitar que personas malintencionadas hagan uso de cuentas que no son de su propiedad, para enviar correos, así como también se evitará que los spamers hagan uso del dominio utpl.edu.ec para inundar de spam a la red.

2.3.4. Usuarios de correo únicamente

Para prevenir explotaciones del usuario local en el servidor de correo electrónico, los usuarios del mismo solamente deben acceder al servidor usando un programa de correo.

No se debe permitir las cuentas con acceso al shell en el servidor de correo, para ello en el archivo de configuración /etc/passwd[18] se debe cambiar /bin/bash a /sbin/nologin (con la posible excepción del usuario root).

[18] Archivo de configuración de los usuarios de Linux

2.4. Asegurando Apache

Apache es uno de los servidores web más utilizados en todo el mundo.

A continuación se definen algunos conceptos claves para obtener un servidor web seguro.

2.4.1. Definiciones de Seguridad

La seguridad en un servidor web se puede definir de varias formas, a continuación listamos algunas de estas:

Confidencialidad: Se refiere a que la información no puede ser revelada a grupos no autorizados.

Integridad: La información permanece igual en una transmisión o en el repositorio de almacenamiento, hasta que un administrador o grupo autorizado la cambie.

Disponibilidad: Los usuarios deben tener accesos de forma ininterrumpida y oportuna a la información y los recursos.

Evaluación: Analizar el ambiente y los requerimientos de seguridad, se debe crear y documentar políticas de seguridad y crear un plan para la implementación de esa política.

Protección: La implementación de un plan de seguridad que pueden ser:

- Configuraciones seguras.
- Recursos protegidos.
- Mantenimiento.

Detección: Identificación de ataques y violación de políticas, para ello se debe hacer uso de:

- Monitoreo
- Análisis de logs
- Detección de intrusos

Debilidad: No es un aspecto ideal de un sistema, puede ser utilizado por los atacantes de alguna manera para acceder a los mismos. Una debilidad puede ser usada para obtener más información.

Vulnerabilidad: Es un error de programación con futuras consecuencias de seguridad.

Exploit: Es un método para explotar una vulnerabilidad. Esto puede ser utilizado para romper o para aumentar los privilegios de usuario (conocida como la elevación de privilegios).

2.4.2. Modelado de amenazas

El modelado de amenazas permite decidir lo que debe realizar ante una eventual amenaza. Este modelado es realmente útil y gira en torno a las siguientes preguntas:

- ¿Qué tienes que es valioso (activos)?
- ¿Dónde pueden atacar (puntos de entrada)?
- ¿Cómo iban a atacar (las amenazas)?
- ¿Cuál es el costo para proteger las amenazas?
- ¿Qué amenazas hay y como enfrentarse contra ellas (mitigación)?

El mejor momento para empezar a realizar este modelado es al principio y sea de utilidad para el diseño del sistema.

Especialmente útil para la evaluación de la seguridad o como parte de pruebas de penetración (un ejercicio en el que se hace un intento de romper en el sistema como un verdadero atacante). Tras el diseño de varios modelos de amenaza, se verán los patrones recurrentes. Los modelos existentes se pueden utilizar como puntos de partida en un nuevo modelo de amenaza con esto se ahorra tiempo en el diseño de una nueva.

En la tabla 2.1 se propone una lista de las razones por las que se puede atentar contra un determinado servicio, la lista que se propone incluye las razones más importantes.

Motivo	Descripción
Para apoderarse de un bien	Los atacantes a menudo quieren obtener algo valioso, como una base de datos de clientes con tarjetas de crédito, datos confidenciales, información privada, nombres de usuario y password.
Para robar un servicio	Esta es una forma especial de la categoría anterior. Los servidores que tiene su ancho de banda, CPU y espacio en disco duro activos. Algunos atacantes los utilizan para: ○ Almacenar software pirata. ○ Indicadores y puntos de partida para ataques a otros sistemas ○ Utilizarlos como zombis automatizados distribuidos para ataques de denegación de servicio.

Reconocimiento	Los ataques, especialmente a sitios web que son catalogados importantes se realizan con frecuencia para elevar el estatus en su medio.
Emoción	Muchos de los atacantes sienten la emoción de romper las seguridades. Para ellos, el entrar en un sistema más seguro les provoca una mayor motivación.

Tabla. 2.1 Motivos de ataque

En la tabla 2.2 según (Moore, Ellison, & Richard) se muestra una lista de los ataques típicos a un sistema, así como también la forma de cómo tratarlos.

Tipo de ataque	Descripción	Mitigación
Denegación de servicio	Cualquier aplicación, servidor web o red se convierte en una base de ataques que resultan en denegación de servicio, condición en la cual un sistema está sobrecargado y no puede responder con normalidad.	○ Prepararse para los ataques. ○ Inspeccionar la aplicación para eliminar la aplicación basada en puntos de ataque.
Explotación de errores de configuración	Estos son errores del administrador del sistema.	○ Realizar una instalación segura desde el inicio de la implementación ○ Plan de cambios y evaluación del impacto que tendrán los cambios antes de hacerlos. ○ Llevar a la práctica la evaluación de la configuración de forma regular.
La explotación de las vulnerabilidades de Apache	Los parches que son desconocidos, problemas que hay en el servidor web.	○ Parchar con la mayor prontitud. ○ Revisar las actualizaciones.
La explotación de	Los parches que no son conocidos o	○ Evaluar la seguridad de aplicaciones web

vulnerabilidades de aplicación	problemas descubiertos en aplicaciones web, puestas en producción.	antes de que cada aplicación sea puesta en producción.
Los ataques a través de otros servicios	Se trata de una captura de todos los problemas en el servidor web, que no han sido mitigados. Ej. Un servidor de base de datos MySQL corriendo en la misma máquina y abierto al público.	○ No exponer servicios que sean innecesarios.

Tabla. 2.2 Ataques típicos a un sistema

Además de las técnicas de mitigación que se listan en la tabla 2.2, existen otros procedimientos de mitigación los cuales también deben ser practicados como:

- Aplicar un monitoreo y considerar la aplicación de detección de intrusos de manera que se conozca si son atacados.
- Disponer de procedimientos de recuperación ante un eventual ataque.
- Realizar periódicamente copias de seguridad y almacenarlas fuera del sitio para que tenga los datos que se necesitan para los procedimientos de recuperación.

2.4.3. Matriz para protección del servidor web

Un problema que hay con frecuencia es decidir cuales métodos de protección usar en la planificación inicial de la instalación. ¿Cómo decidir qué método es justificable? y ¿Cuál no es?

Para esto se realiza una matriz de decisión para proteger el servidor web, en primer lugar, se lista todos los posibles métodos de protección y clasificados cada uno en términos de complejidad. Todos los sistemas se han clasificado en cuatro categorías que son:

Categoría 1. Crítico (más importantes)

Categoría 2. Producción.

Categoría 3. Desarrollo.

Categoría 4. Pruebas (menos importante)

Entonces se toma una decisión en cuanto a qué método de protección se justifica para utilizar en el sistema.

En la Tabla 2.3 según (Moore, Ellison, & Richard) se presenta una matriz de protección del servidor web.

Técnicas	Categorías			
	4	3	2	1
Instalar parches al kernel				X
Compilar Apache desde el código fuente			X	X
Ajustar la configuración (remover módulos por defecto, restringir ciertos módulos)			X	X
Cambiar la identidad del servidor web.			X	X
Implementar autenticación. Ej. Usar auditoria a la autenticación.			X	X
Implementar SSL[19]			X	X
Desplegar certificados desde un CA[20] reconocido			X	X
Desplegar certificados (cuando sea apropiado)				X
Centralizar logs	X	X	X	X
Uso de mod_security superficialmente			X	X
Uso de mod_security fuertemente				X
Realizar monitoreo al servidor		X	X	X
Realizar monitoreo externo de la disponibilidad			X	X
Realizar una inspección o monitoreo periódico de los logs.	X	X	X	X
Monitoreo en tiempo real de los logs				X
Realizar un análisis periódico de los logs.			X	X
Realizar correlación de eventos.				X
Implementar reglas de firewall		X	X	X
Validar la integridad de los archivos			X	X
Disponer de una evaluación de vulnerabilidades externas.				X
Separar los componentes de aplicación.				X

Tabla 2.3 Matriz de protección del servidor web

[19] **SSL: (Secure Socket Layer) P**rotocolo diseñado por Netscape que posibilita la transmisión segura de información en la Red.

[20] **CA: (Certificate Authority)** certificado de autoridad, compañía autorizada para emitir certificados digitales que autenticarán la identidad de un usuario u organización durante una transacción asegurada en Internet

Este sistema de clasificación ayuda a decidir cuándo aplicar el parche en un sistema después de que un problema se descubre.

Se propone seguir el siguiente plan:

Categoría 1

Parche de inmediato.

Categoría 2

Parche al día siguiente.

Las categorías 3 y 4

Parchar cuando el parche esté disponible o, si el servidor web se ha instalado desde el código fuente.

2.4.4. Cálculo de riesgos

Un plan de parches, como la que se listó en la sección anterior, se asume que tendrá los suficientes recursos para hacer frente a esos problemas y se podrá ocupar de ellos rápidamente.

Esto sólo funciona para los problemas que son fáciles y rápidos de arreglar. Pero, ¿qué ocurre si no hay recursos suficientes para la revisión de todo dentro del cronograma? Algunas a nivel de aplicación y, sobre todo, arquitectónico las vulnerabilidades pueden requerir una gran inversión de recursos. En este punto, se tendrá que tomar una decisión en cuanto a que problemas solucionar ahora y cuales arreglar más tarde. Para ello se tendrá que asignar un riesgo percibido para cada problema y dar la solución a los problemas más grandes primero.

Para calcular el riesgo, en la práctica significa asignar valores numéricos a los siguientes factores para problemas descubiertos:

Explotabilidad

La probabilidad de explotar la vulnerabilidad.

Daños potenciales

La gravedad de la vulnerabilidad

Valor de los activos

El costo de la restauración al estado en que se encontraba, incluyendo los gastos de contratación de terceros para dar solución a los problemas.

Combinados, estos tres factores proporcionan una cuantificación del riesgo. El resultado no significa mucho por sí sola, sino que también servirá para comparar con los riesgos de otros problemas.

Si se necesita una medida para decidir si solucionar un problema o determinar cuánto invertir en medidas de protección, se debe calcular los posibles riesgos anuales. El costo es usado para determinar si se llevará a cabo cualquier acción para mitigar el problema.

2.5.Conexiones cifradas

Hay que potenciar entre los usuarios el uso de programas como ssh/scp o ssl-telnet/ssl-ftp para conectarse a los servidores. La utilización de estos programas es muy simple ya que facilita el trabajo a los administradores, así como también permite realizar conexiones más seguras.

Existen clientes para Windows y MacOS, por lo que todos lo pueden utilizar.

CAPÍTULO III: Sistemas de Replicación de Archivos

En este capítulo se realiza una descripción de los sistemas de replicación de archivos, los cuales ayudan a asegurar la transferencia de información que hay entre los distintos nodos del clúster.

También se encuentra una descripción de tres sistemas de replicación de archivos: CODA, GlusterFS y OpenGFS.

Finalmente se realiza un cuadro comparativo entre estos sistemas, para así seleccionar el que más se acople a los requerimientos que la UTPL necesita para el clúster de servicio de web y mail.

3.1. Introducción

Los Sistemas de archivos distribuidos han crecido en importancia en los últimos años, así como también la dependencia de estos sistemas aumenta, el problema de la disponibilidad se hace más agudo. Hoy en día, una caída del servidor o de la red puede causar inconveniente a muchos usuarios.

La capacidad de compartir los discos, directorios y ficheros a través de la red es uno de los avances más importantes en la computación moderna ya que permiten la reducción de los requisitos de espacio de disco local, haciendo más fácil a los usuarios a colaborar sin tener que al final acabar con cientos de versiones de los mismos archivos. Linux y sistemas Unix tradicionalmente utiliza el sistema de archivos de red NFS.

Es por ello que se hará uso de sistemas para la compartición de ficheros, la cual se la utilizará para tener la dirección /var/spool/mail tal como está en el servidor principal, en todos los nodos clientes del clúster.

NFS es el mecanismo más conocido de intercambio de archivos sobre la red ya que está incluida en la mayoría de los sistemas Unix. NFS se apoya en el núcleo Linux y está relacionado con los servicios públicos que se prestan con cada distribución de Linux. Sin embargo, varios de los modernos mecanismos para el intercambio de archivos y directorios a través de las redes están disponibles para sistemas Linux. Cada uno de estos puede proporcionar importantes ventajas administrativas.

El sistema de intercambio de archivos NFS presenta algunas falencias, inconsistencias en la información, es un sistema de archivos remoto simple ya que exporta la información de permisos al estilo de Unix, así como el identificador numérico de usuario y grupo (*uid/gid*) de cada usuario, pero no realiza ninguna clase de control de acceso, dejando esta labor al sistema cliente. De esta manera se necesita que el sistema cliente obtenga la información adecuada sobre la identidad del usuario y los permisos que el mismo debe tener sobre los archivos exportados.

Es por ello que se ha dado origen a la utilización de otros mecanismos para la replicación de archivos. De entre los cuales tenemos, GlusterFS, OpenGFS, CODA entre otros.

3.2. Descripción de los sistemas de replicación de archivos

Hoy en día se utiliza el término "sistemas de archivos distribuidos" para sistemas de ficheros en red. Esto refleja el hecho de que muchos de estos sistemas de archivos hacen mucho más que la simple exportación de datos a través de una red. Los medios de almacenamiento asociados con estos sistemas de archivos no necesitan estar ubicados en un único sistema, pueden ser distribuidos a través de múltiples ordenadores.

Sistemas de ficheros distribuidos como CODA incluyen su propio mecanismo de gestión de volúmenes que simplifican la gestión del almacenamiento compartido. También apoya a la replicación, que es la capacidad de hacer copias de los volúmenes exportados y almacenar esas copias a otros servidores de archivos. Si un servidor de archivos no está disponible, los datos almacenados en sus volúmenes todavía se pueden acceder desde disposición de las réplicas de ese volumen, esto se conoce como cache.

GlusterFS proporciona servicios de gestión de volúmenes aprovechando esto para montar volúmenes de diferentes servidores de archivos en una jerarquía de directorio centrales que es soportado por el sistema de ficheros. Estas jerarquías de directorio son visibles para todos los clientes de estos sistemas de ficheros distribuidos. Esto permite a los usuarios acceder a sus archivos de datos exactamente de la misma manera que desde cualquier cliente. Si la máquina falla, se puede usar otro ya que los archivos aún están intactos y en condiciones de seguridad en el servidor de archivos.

Los Sistemas de ficheros distribuidos que proporcionan los mismos datos a las diferentes computadoras permiten a los usuarios utilizar el tipo de máquina de escritorio que mejor se adapte a sus necesidades, mientras se está teniendo acceso a un sistema de archivos centralizados. Los usuarios de Macintosh pueden aprovechar las herramientas gráficas disponibles en Mac OS mientras se están guardando sus archivos a los servidores de archivos centralizados. Los usuarios de Windows pueden tener acceso a un robusto sistema de archivos.

Los Sistemas de ficheros distribuidos son especialmente atractivos cuando se trata de trabajo entre grupos ubicados en diferentes ciudades, estados o incluso países. Compartido los datos que están siempre disponibles a través de la red, independientemente de su ubicación.

3.3. CODA

Es un sistema de archivos en red avanzado, está basado en el sistema de archivos **AFS2**[21]

La arquitectura de CODA reconoce tres tipos de máquinas que son: clientes, servidores y una máquina que controla el sistema.

Las máquinas clientes son típicamente usadas por los usuarios para acceder a la información que se encuentra compartida.

[21] **AFS2** (Andrews File System) es un sistema de archivos distribuido.

Las máquinas servidores son seguras, confiables cuyo propósito es el de atender el servicio de compartición de archivos solicitadas por los clientes. Estas máquinas tienen la necesidad de resguardar la información es por ello que utilizan mecanismos de autenticación para los usuarios que requieren acceder a la información compartida.

El tercer tipo de máquina es el de sistema de control de la máquina (SCM) cuyo propósito es el de proporcionar un único punto para el control de administración. Lógicamente el SCM cumple un rol distinto al de los servidores, pero físicamente este puede actuar como uno más de ellos.

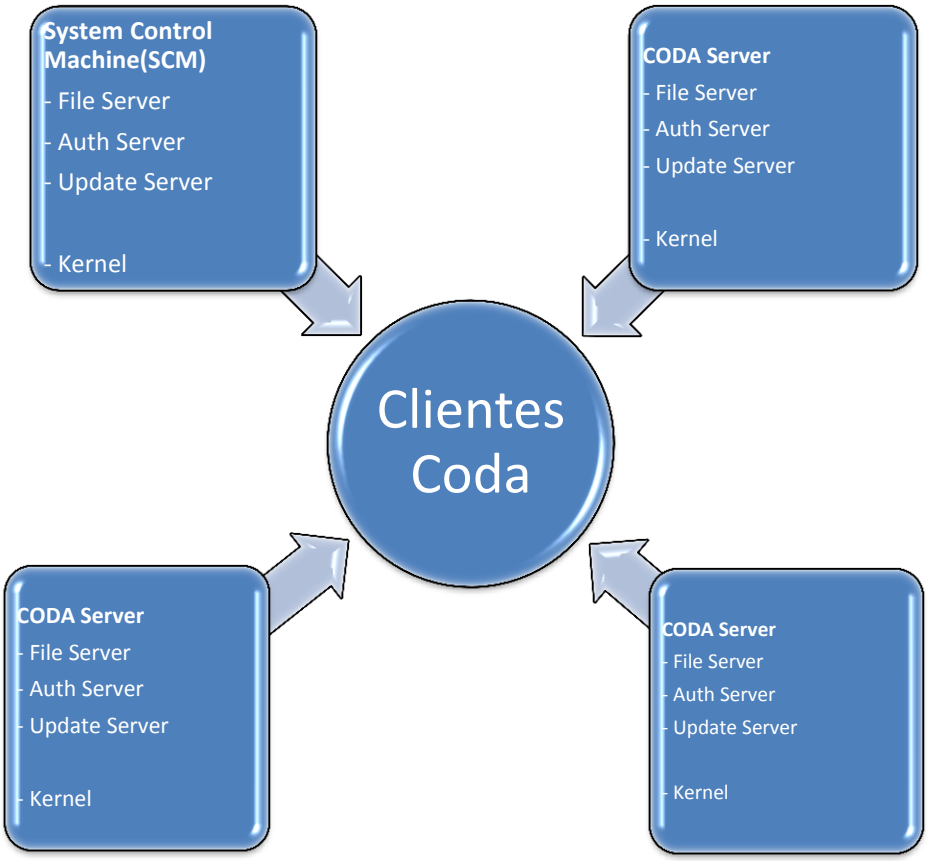


Fig. 3.1 Estructura de Coda

3.3.1. Conceptos claves de CODA

Celdas CODA: Una celda es un grupo de servidores que comparten un conjunto de bases de datos de configuración. Una celda puede constar de un solo servidor o hasta cientos de servidores. A un servidor se designa responsable de la SCM (Sistema de Control de la Máquina).

Volumen CODA: Un volumen es más pequeño que una partición y más grande que un directorio. Los volúmenes tienen una raíz y contienen un árbol de directorio con archivos. Cada volumen está montado en algún lugar bajo /coda y forma un árbol de /coda. Un montaje CODA contiene suficiente información para que el cliente encuentre el servidor que almacena los archivos en el volumen.

Almacenamiento de datos: Los servidores no almacenan ni exportan volúmenes como directorios en el disco local del sistema de archivos, tal como lo hace NFS y Samba. CODA necesita mucho más apoyo en los meta datos para soportar la replicación y las operaciones en modo desconectado. Los servidores CODA almacenan archivos identificados por un número, normalmente todos en un árbol de directorios bajo /vicepa. La meta datos se almacena en un archivo de datos RVM, que puede ser una partición de disco o simplemente un archivo.

RVM: es una operación basada en librerías que hace parte de un espacio de direcciones virtuales de un proceso persistente en el disco. CODA utiliza RVM para gestionar sus metadatos, estos datos están almacenados en un fichero de datos RVM que está asignada en la memoria cuando se inicia.

Validación: al detectar que un servidor es accedido nuevamente, este valida los datos en caché antes de usarlo para asegurarse de que la caché de datos es la última versión del archivo.

3.3.2. Características

- Altas prestaciones mediante caché persistente en los equipos cliente
- Replicación de servidores
- Buena escalabilidad
- Adaptación al ancho de banda
- Funcionamiento continuo ante pérdidas de conexión
- Funcionamiento sin conexión (equipos móviles)
- Modelo de seguridad para autenticación, encriptación y control de acceso
- Es un árbol de directorios con una raíz
- Más pequeño que una partición
- Más grande que un directorio
- Los clientes montan los volúmenes dentro de su directorio /coda (todos los clientes ven lo mismo)

3.3.3. Puntos de montaje/anclaje

Existe un volumen principal que se monta en /coda

Se puede ampliar el árbol insertando otros volúmenes para ello se utiliza el siguiente comando:

cfs mkmount <nv> <pa>

nv: nombre del volumen

pa: Punto de anclaje

Los puntos de anclaje son persistentes (no necesitan ser remontados el reiniciar)

3.3.4. Almacén en el servidor

El contenido de los archivos se almacena en un sistema de archivos local como en NFS o SAMBA

Necesitan más información (metadatos) para los archivos compartidos (propietarios, control de acceso, versiones…)

Los metadatos se almacenan en un archivo especial RVM (Recoverable Virtual Memory), el cual mapea los datos a la memoria al iniciar.

RVM suele ser una partición.

3.3.5. Caché en el cliente

Caché persistente almacenados en **/usr/coda/venus.cache**

El cliente también usa metadatos en un RVM **(/usr/coda/DATA**)

La caché es una réplica de los archivos del servidor

Los accesos de lectura son rápidos ya que son locales (usa caché)

Permite el acceso en desconexión, CODA sincroniza la caché de los clientes con los archivos del servidor.

3.3.6. Autenticación

Uso de un testigo para la autenticación, el testigo se utiliza para comprobar el acceso de los usuarios a los archivos.

Los usuarios hacen un login con el que obtienen una clave de entrada (testigo)

3.3.7. Soporte de CODA en el cliente

CODA necesita un soporte mínimo en el kernel (módulo), trabaja en unión con el administrador de la cache (Venus)

Venus es una aplicación que trabaja en el espacio de usuario y este monta la información que es replicada bajo **/coda**

3.3.8. Utilidades

cfs: Listas control de acceso, caché, puntos de anclaje.

codacon: Realiza la monitorización de operaciones de administración de la caché.

clog: Permite la identificación en el servidor.

cmon: Muestra la lista de servidores.

ctokens: Muestra una lista de tokens CODA de los usuarios que se autenticaron.

3.3.9. Servicios

Los servicios que se ejecutan en el servidor son los siguientes:

codasrv: Servidor principal, trabaja en conjunto con el proceso ***venus*** en los clientes. Estos procesos permiten el intercambio de datos entre cada una de las máquinas que conforman el clúster.

auth2: Acepta peticiones de clog y gestiona los testigos, se ejecuta en todos los servidores, su función es la de validar a los usuarios. Las contraseñas sólo se pueden ser modificadas en el servidor maestro (SCM), por lo cual la copia de la base de datos tiene la propiedad de solo lectura. Las contraseñas se actualizan automáticamente con la ayuda de los demonios ***updateclnt/updatesrv***

updateclnt: Permite mantener actualizadas las copias de las bases de contraseñas en todos los servidores, manteniendo una copia original en el SCM. Para ello, el demonio **updateclnt** realiza una comprobación cada cierto tiempo si los ficheros del SCM han sido actualizados, esto hace que las actualizaciones no sean inmediatas, pues dependen del periodo de comprobación de updateclnt

Servicios que se ejecutan en el cliente

venus: Es un demonio que se ejecuta únicamente en los clientes Coda y tiene como función la de realizar dos aspectos importantes:

- Dialogar con el servidor y realizar el intercambio de datos y de ficheros.
- Dialogar con el controlador del kernel de la máquina local para pasarle éstos datos y que genere el contenido del directorio virtual en ***/coda*****.**

3.3.10. Funcionamiento

El funcionamiento de CODA es el siguiente:

1. Usa **RPC**
2. Si se modifica un archivo se propaga al servidor de forma síncrona
3. Si se detecta que no hay conexión se pasa a modo desconectado
4. En este modo los cambios se almacenan en el CML (registro de modificaciones)

5. Al reconectar se usa el CML para actualizar el servidor

Se tiene dos formas de resolución de conflictos que son:

- **Automático**
- **Manual**

3.3.11. Esquema de funcionamiento según (Braam, Baron, Harkes, & Schnieder)

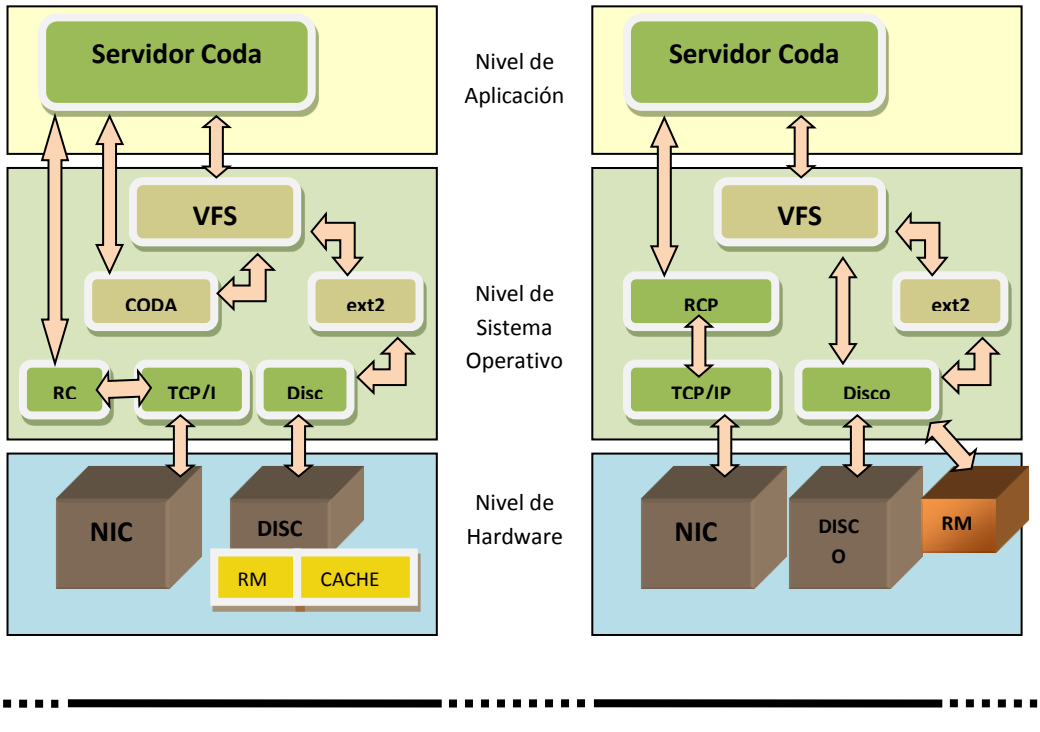


Fig. 3.2 Esquema de funcionamiento

3.4. OpenGFS

Open Global File System es un sistema de ficheros transaccional que permite realizar o soportar el intercambio simultáneo de un dispositivo de almacenamiento de múltiples nodos de computadoras. Esta es una manera de poner en marcha un sistema de archivos en clúster ya que proporciona todos los componentes necesarios para hacerlo.

Las máquinas y dispositivos de almacenamiento están conectados a través de una red de canal de fibra, los nodos se adjuntan a la red como se puede apreciar en la parte superior de la figura No. 3.3 y el repositorio de almacenamiento en la parte inferior.

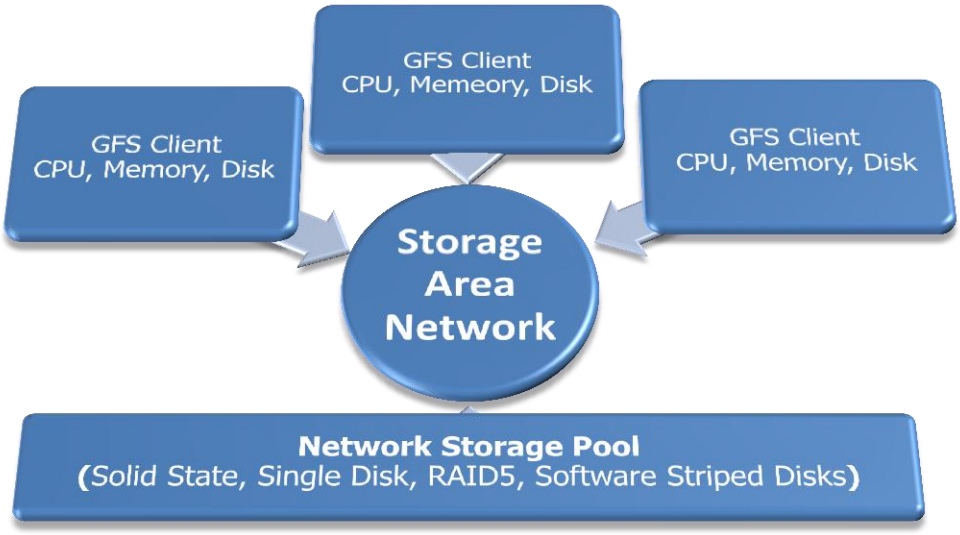


Fig. 3.3 Esquema de OpenGFS

3.4.1. Estructura del sistema de ficheros

Cada sistema de ficheros GFS se divide en varios grupos de recursos (RG). Los Grupos de Recursos están diseñados para distribuir los recursos del sistema de archivos a través de todo el almacenamiento. Existen múltiples Grupos de recursos por dispositivo y pueden ser entrelistados a través de varios dispositivos. Los grupos de recursos son esencialmente mini-sistemas de ficheros. GFS también tiene un superblock que contiene información que no puede ser distribuido a través de los grupos de recursos. Esta información incluye el número de nodos montados en el sistema de archivos, mapas de bits para calcular identificadores únicos para cada nodo, el dispositivo en el que el sistema de archivos está montado y el tamaño de los bloques del sistema de archivos.

3.4.2. Dinodes

Un dinode GFS toma todo un bloque del sistema de archivos. Cada dinode se divide en una sección de encabezado que contiene campos dinode estándar y una sección de enlaces. Cada puntero apunta a un bloque indirecto que a su vez apunta a bloques de datos.

3.4.3. Bloqueos en los dispositivos

Los Bloqueos de dispositivos son mecanismos de los nodos de las máquinas para mantener la exclusión mutua de los datos del sistema de archivos. Estos son implementados en los dispositivos de almacenamiento y se pueden acceder con un comando SCSI. El dispositivo de aplicación de los bloqueos está limitado por:

- ✓ Los comandos de bloqueo de dispositivo son independientes de los demás comandos SCSI.
- ✓ Cada bloqueo requiere cantidades mínimas de controlador de memoria de disco - tan poco como un byte por bloqueo.

Los Bloqueos de dispositivos se utilizan también para ayudar a mantener la coherencia de los metadatos cuando está en caché por varios clientes.

Un cliente de GFS adquiere un bloqueo, lee los datos, modifica los datos, escribe los datos y libera el bloqueo. Esto permite que el sistema de archivos pueda completar las operaciones en las que los metadatos son "atómicas" con respecto a otras operaciones en el mismo metadatos. Cada Dlock también tiene un "número de versión" asociada a ella. Cuando un cliente va a realizar una operación de lectura, modificación o escritura en una sección de metadatos, esta adquiere el bloqueo y al realizar la lectura, modificación o escritura se libera el bloqueo usando una acción de desbloqueo incremental.

Cuando un cliente quiere leer los metadatos, que adquiere el bloqueo, lee los metadatos y libera el bloqueo usando la acción desbloquear.

3.4.4. Estado de los Bloqueos

El estado de cada bloqueo se describe por un bit. Si el bit tiene el valor de 1, el bloqueo se ha adquirido y es propiedad de un nodo de la máquina. Si el bit es 0, el bloqueo está disponible para ser adquirida por cualquier nodo. La acción del comando Dlock establece el conjunto de pruebas para determinar en primer lugar si el valor de bloqueo es 1. Si el valor es 1, el comando devuelve con un estado que indica que el bloqueo ya ha sido adquirido. Si el valor es 0, Dlock establece el bloqueo a 1 y retorna un estado bueno para el iniciador.

3.4.5. Consistencia y Caché

La consistencia se mantiene mediante el uso de operaciones atómicas garantizada por el dispositivo de bloqueos al modificar los datos. Cuando los dispositivos de bloqueo no están implementados en un dispositivo de almacenamiento, los comandos SCSI Reserve y Release se puede utilizar para realizar operaciones atómicas sobre datos. Estos comandos proporcionan acceso exclusivo a todo el dispositivo para un nodo, no por el servicio de las peticiones de otros nodos. Estos comandos garantizan el acceso exclusivo, pero no aportan mucho paralelismo. Con sólo una reserva por dispositivo, muchas solicitudes en conflicto no tienen que esperar hasta que el dispositivo de almacenamiento este libre. En un entorno distribuido tal acceso limitado al sistema disminuye el rendimiento y los tiempos de respuesta. El protocolo SCSI describe

los comandos opcionales Reserve y Release. Estos comandos permiten a los iniciadores de reserva un acceso exclusivo sólo los bloques de datos que pueda necesitar, por lo que estos no son aplicados por la mayoría de fabricantes de dispositivos.

3.4.6. Repositorio de almacenamiento en Red

El repositorio de almacenamiento en red (NSP) el manejador del volumen soporta abstracción de un único espacio de direcciones de almacenamiento unificado para los clientes de GFS. El NSP es implementado en un controlador de dispositivo en la capa superior de la base del dispositivo SCSI y en el manejador del Canal de Fibra. Este controlador traduce los espacios de direcciones lógicas del sistema de archivos a espacio de direcciones de cada dispositivo. Los subrepositorios NSPs dividen en grupos de similares tipos de dispositivo que heredarán los atributos físicos de los dispositivos y conexiones de red.

3.4.7. Grupos de Recursos

GFS distribuye sus metadatos a través de la red de almacenamiento en lugar de concentrar todo en un solo superblock. Múltiples grupos de recursos se utilizan para la partición de metadatos, incluyendo mapas de bits de datos y bloques de datos, en grupos separados a los clientes y así aumentar el paralelismo y la escalabilidad del sistema de archivos, evitar los cuellos de botella y reducir el tamaño medio de las típicas operaciones de búsqueda de metadatos. Uno o varios grupos de recursos pueden darse en un solo dispositivo o un único grupo de recursos pueden incluir múltiples dispositivos. Al igual que los grupos de recursos, los grupos de bloque explotan el paralelismo y escalabilidad al permitir que múltiples hilos de un solo ordenador puedan asignar y liberar bloques de datos, los grupos de recursos de GFS permiten múltiples clientes a hacer lo mismo. GFS también tiene un solo bloque, el superblock, que contiene resumen de metadatos, no se distribuye a través de los grupos de recursos. (El superblock puede repetirse para mejorar el rendimiento y la facilidad de recuperación.)

Esta información incluye el número de clientes montado en el sistema de archivos, mapas de bits para el cálculo de los identificadores únicos de cada cliente, el dispositivo en el que el sistema de archivos está montado y el tamaño de bloque del archivo Sistema. El superblock contiene también un índice estático de los grupos de recursos que describe la ubicación de cada grupo de recursos y demás información de configuración.

3.4.8. VFS Caching

Cuando la capa de VFS necesita información de la capa específica del sistema de archivos, esta hace una llamada de función hacia el sistema de archivos dependiendo de la capa para la información.

Una máquina puede cambiar los datos en el sistema de archivos sin tener que alarmar a otras máquinas. La capa VFS siempre pide a la capa del sistema de archivos específico cuando quiere información. La capa específica del sistema de archivos siempre puede proporcionar un nivel más actualizado de los metadatos. Todos los discos de accesos pasan por la capa VFS. El sistema de ficheros local puede ser muy rápido, porque el VFS evita la sobrecarga de llamar a la función y espera a que la capa específica del sistema de archivos localice y codifique la información solicitada ya que simplemente lee los datos de su propia copia.

Un caching no controlado en un sistema de archivos en red, especialmente en un sistema de archivos de disco compartido, puede dar lugar a incoherencias de datos entre las máquinas.

3.4.9. Mejoras en el sistema de ficheros

Se han introducido mejoras en el sistema de archivos y metadatos. Estos cambios, proporcionan aumento en la escalabilidad de GFS.

3.4.9.1. Consistencia de GFS

Hay que tener mucho cuidado cuando los metadatos son accedidos y actualizados. Ya que si el Dlocks no actúa en el momento oportuno, puede resultar muy fácil la corrupción de los datos y metadatos.

Este nuevo bloqueo también ha aumentado las posibilidades de estancamiento. Hay muchos lugares donde el sistema de archivos debe mantener de dos o más Dlocks para realizar una operación. Por ejemplo, la operación de búsqueda requiere simultáneamente dos bloqueos. La operación de búsqueda toma un directorio y el nombre de un archivo en ese directorio y devuelve el inodo para ese archivo. Dos bloqueos se adquieren para esta operación: un bloqueo se realiza cuando el directorio es leído y cuando el número inodo del archivo es determinado. El otro bloqueo se realiza al tiempo que el inodo se lee. Estos dos bloqueos deben ser considerados, al mismo tiempo. Hay algunos otros lugares donde dos o más bloqueos se realizan en un mismo punto llegando a matar el bloqueo.

3.4.9.2. Buffer Caché

El buffer de memoria caché es un componente importante de los modernos sistemas operativos UNIX, porque evita el exceso de accesos a disco ya que el sistema operativo guarda los bloques de disco utilizado recientemente en una sección de la memoria llamada "buffer caché". Las futuras solicitudes de datos que ya están en el buffer de memoria caché se puedan

completar con rapidez ya que no se requiere el acceso a disco.

Si los datos solicitados no están en el buffer de memoria caché, estos se leen desde el disco y, a continuación, copia en un buffer de memoria caché, así como para el programa de usuario. Esto se aplica tanto a los metadatos de archivos y bloques. En este caso el rendimiento es mucho mayor al utilizar el buffer de caché en vez de acceder al disco. La Caché de bloques de metadatos también mejora el rendimiento para archivos de gran tamaño a causa de las repetidas solicitudes indirectas de referencias de bloque. El uso de cache en el buffer de GFS es complicada por la capacidad de acceder a múltiples clientes y caché de los bloques de disco.

3.4.9.3. Administración del espacio libre

La administración del espacio libre en GFS se basa en el enfoque de mapa de bits. Por cada bloque de sistema de ficheros en un determinado grupo de recursos hay un único bit para representar si el bloque está libre o no. Este método es eficiente, pero el espacio como el sistema de archivos se llena de datos, una búsqueda a través de un creciente número de bits es necesaria, a fin de hallar el necesario espacio libre. La utilización de un esquema de base extendido, puede costar más en términos de espacio, pero proporciona un mejor rendimiento. En lugar de hacer el seguimiento de cada bloque de sistema de archivos en un grupo de recursos, nos limitamos a los bloques libres. Por cada grupo de bloques libres del sistema de archivos en un grupo de recursos habrá una medida que no pierde de vista el punto de partida y el número de bloques en el grupo.

3.4.10. Recuperación de Error

La recuperación de error es importante en un sistema de archivos de discos compartidos, porque ayuda a identificar en donde está fallando y así poder controlar esta falla a tiempo.

Todos los clientes manipulan directamente los metadatos, por lo que el fracaso de cualquier cliente puede dejar los metadatos en un estado inconsistente. Además, como hay tantas máquinas accediendo a los discos, no es práctico para todos ellos a desmontar y esperar a comprobar un sistema de ficheros (fsck) para completar cada vez que un cliente falla. Es importante que las incoherencias causadas por un cliente sean localizadas y fácilmente reparadas mientras que el sistema de ficheros está en funcionamiento.

3.5. GlusterFS

GlusterFS es un sistema de archivos en clúster capaz de escalar a peta bytes y manejar miles de clientes. GlusterFS se puede combinar de forma flexible con productos físicos, virtuales y recursos de la nube para ofrecer almacenamiento empresarial con alta disponibilidad y performance a una fracción del costo de las soluciones tradicionales.

GlusterFS está diseñado para trabajo de alto rendimiento y entornos de nube virtualizados. A diferencia de los centros de datos tradicionales, los entornos cloud requieren multi-alquiler, junto con la capacidad de aumentar o disminuir los recursos bajo demanda.

La ampliación de la capacidad, rendimiento y disponibilidad, se la puede hacer con el uso de nubes públicas y entornos híbridos, ya no hay dependencia de un proveedor específico.

En la Fig. 3.4 se ilustra la forma de funcionamiento de GlusterFS.

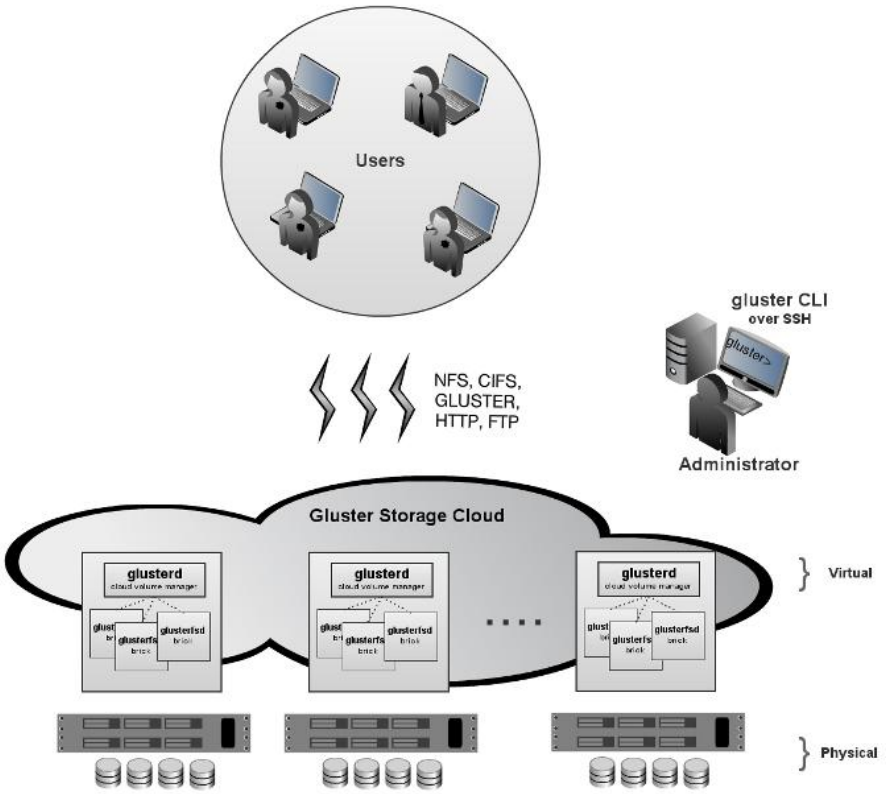


Fig. No. 3.4 Esquema de funcionamiento de GlusterFS tomado de (Gluster)

3.5.1. Principales consideraciones del diseño de GlusterFS

Para el diseño de GlusterFS se toman en cuenta algunos factores como los que se citan a continuación.

- **Capacidad de Expansión:**

  La expansión de GlusterFS es muy grande ya que fácilmente puede sobrepasar un peta-byte de información.

- **Facilidad de administración**

  Utiliza un esquema parecido al NFS (no NFS como tal) debido a que GlusterFS implementa ciertas características de seguridad a la información que se va a replicar.

  La entrada y salida de información no está vinculada a perfiles de hardware y sistema operativo.

- **Fiabilidad**

  Posee un almacenamiento incremental ya que no deja de guardar información.

- **Escalabilidad de flujo I/O**[22]

  Manejador para agrupaciones de I/O

3.5.2. Ventaja sobre el transporte RDMA[23]Características

- ☞ Escalabilidad sobrepasando peta-bytes
- ☞ Soporte para Infiniband, RDMA y TCP/IP
- ☞ Balanceo de carga.
- ☞ Alta disponibilidad.
- ☞ Completamente distribuido.
- ☞ Replicación de archivos
- ☞ Encriptación de los datos
- ☞ Diseñado para trabajar en ambientes de clúster
- ☞ Comunicación segura con los nodos ya que implementa mecanismos de autenticación.
- ☞ La implementación es fácil.
- ☞ Posee Soporte
- ☞ Posee Licencia GPL

3.5.3. Nuevas Características de la versión actual de GlusterFS según (Gluster)

- **Geo-replication**

  Ofrece un servicio de replicación continua, asíncrona, e incremental de un sitio a otro sobre redes de área local (LAN), redes de área amplia (WANs) y a través de Internet.

  Con la utilización de Gluster Geo-replication, se puede establecer redundancia de datos, proporcionando una recuperación ante desastres a través redes LAN, WAN y conexiones a Internet existentes.

[22] **I/O** (Entrada / Salida) se refiere a la entrada y salida de información en un medio.
[23] Remote Direct Memory Access (RDMA) permite mover datos directamente desde la memoria de un computador a otro. Esto permite alto rendimiento, baja latencia de red.

- **Directory Quota**

  Permite establecer cuotas de uso de espacio en disco, directorios o volúmenes. Los administradores pueden controlar el uso de espacio en disco y/o nivel de volumen en GlusterFS mediante la asignación de espacio en disco en cualquier nivel en el volumen y jerarquía de directorios.

- **Top and Profile**

  Permite controlar los diferentes parámetros de la carga de trabajo, lo que ayuda en la planificación de la capacidad y las tareas de optimización del rendimiento del volumen.

- **POSIX ACLs Support**

  Permite asignar diferentes permisos para distintos usuarios o grupos a pesar de que no se corresponden con el propietario original o el grupo propietario.

3.5.4. Componentes GlusterFS

Para la utilización de GlusterFS se hacen uso de dos componentes principales que son:

3.5.4.1. GlusterFS Server

Aquí se exportan los volúmenes para los nodos del clúster, se pueden configurar varios GlusterFS Server para evitar tener un punto único de fallo.

El demonio que se utiliza en el servidor es:

glusterfsd

Archivo de configuración se encuentra bajo:

/etc/glusterfs/glusterfs-server.vol

3.5.4.2. GlusterFS Client

Este componente es instalado y configurado en los nodos del clúster, la función que cumple este es el montar los volúmenes que se encuentren configurados en el servidor.

El demonio que se utiliza en el cliente es:

glusterfs

Archivo de configuración se encuentra bajo:

/usr/local/etc/glusterfs/glusterfs-client.vol

Un aspecto importante a tener en cuenta en GlusterFS es que los clientes deben tener soporte para fuse[24].

[24] **Fuse:** Módulo para el kernel el cual es un soporte para GlusterFS

3.5.4.3. Gluster Console Manager según (Gluster)

Es una utilidad de línea de comandos que simplifica la configuración y gestión del entorno de almacenamiento.

Con el uso de esta consola los administradores pueden crear nuevos volúmenes, iniciar y detener los volúmenes, según como se requiera.

Los administradores también pueden utilizar los comandos para crear scripts para la automatización, así como utilizar los comandos como un API para permitir la integración con aplicaciones de terceros.

Los comandos se pueden invocar directamente desde la consola.

 # gluster COMMAND
 # gluster peer status

También mediante la ejecución de Gluster Console en modo interactivo.

 # gluster
gluster>
gluster> peer status

3.5.5. Propiedades GlusterFS

GlusterFS posee propiedades que permiten mejorar el rendimiento del sistema de archivos.

Estas propiedades están clasificadas de la siguiente manera.

Propiedad	Elementos
– **Rendimiento**	– read ahead – write behind – threaded I/O – IO-cache – stat pre-fetch – booster
– **Replicación**	– afr
– **Clustering**	– stripe – unify
– **Balanceo de carga**	– ALU – NUFA – Random – Round Robin – Switch
– **Depuración**	– trace

Características Extras	– filter – posix-locks – trash – fixed-id
Almacenamiento	– posix
Protocolos Servidor	– tcp/server – ib-sdp/server – ib-verbs/server
Protocolos Cliente	– tcp/server – ib-sdp/server – ib-verbs/server
Autenticación	– auth.ip – auth.login
Encriptación	– rot-13

Tabla 3.2. Propiedades de GlusterFS

3.5.5.1. Rendimiento

Las propiedades de rendimiento trabajan bien cuando son cargadas tanto del lado del servidor como del cliente

read-ahead. Esta propiedad realiza una pre-búsqueda avanzada de una secuencia de bloques de datos basada en predicción.

Es usada en el momento en que la aplicación está ocupada procesando datos que ha leído con anterioridad, para que GlusterFS pueda realizar una pre-lectura del siguiente batch de datos y dejarlo listo para utilizarlo, es decir al momento que se lee un bloque de información, en ese mismo instante también está realizando una pre-lectura del próximo bloque, de este modo las lecturas son más rápidas, adicionalmente cambia el comportamiento como si fuera un canalizador de lectura.

Dentro de read-ahead hay una opción **page-size** que indica el tamaño de un bloque y **page-count** describe la cantidad de bloques a ser pre-buscados.

write-behind. Generalmente la operación de escritura es más lento que la de lectura, es por ello que se utiliza esta propiedad para mejorar significativamente el rendimiento de escritura usando una técnica de escritura en segundo plano o lo que se conoce como background.

De este modo múltiples operaciones de escritura son agregados sin limitación alguna y son escritas en un segundo plano.

La opción **aggregate-size** de write-behind determina el tamaño del bloque a ser escrito.

Se recomienda realizar un benchmark[25] con un incremento en el valor del rango de aggregate-size para así obtener un valor óptimo y lograr un mejor rendimiento.

io-threads. Agrega sincronización para la funcionalidad de lectura/escritura.

Esta propiedad hace que se use de una mejor manera todos los recursos.

io-cache. Ayuda a reducir la carga en los servidores.

3.5.5.2. Replicación

GlusterFS trae consigo una propiedad que permite la replicación de información para ello se utiliza la propiedad **afr**

afr. (Automatic – File – Replication) Replicación automática de archivos, con esta propiedad se evita el tener un punto único de fallo ya que con esto se pueden crear varios espejos.

3.5.5.3. Clúster

Como ya lo dijimos al inicio GlusterFS está diseñado para trabajar en un ambiente de un clúster, es por ello que se cuentan con las siguientes características: **stripe** y **unify.**

stripe. Clasifica los archivos de entrada de acuerdo con un patrón definido por el usuario dentro de un tamaño de bloque dado.

unify. Combina varios nodos de almacenamiento en un solo gran nodo de almacenamiento en el servidor.

3.5.5.4. Balanceo de carga

Permiten decidir cómo se va a distribuir la creación de nuevas operaciones sobre el sistema de archivos del clúster tomando en cuenta la carga, disponibilidad, entre otros factores.

[25] **Benchmark:** Conjunto de procedimientos para avaluar el rendimiento de un servicio

A continuación se describirá cada uno de las propiedades que ayudan al balanceo de carga.

ALU. Es uno de los manejadores más avanzados y disponibles en GlusterFS, este balancea la carga a través de volúmenes, tomando varios factores en cuenta.

NUFA. Permite manejar distintos sistemas de ficheros sin importar que sean de un mismo clúster.

Round-Robin. Crea archivos al estilo round-robin, cada cliente tiene su propio bucle round-robin.

Es una buena opción el utilizarlo cuando los archivos son muy similares tanto en tamaño como en patrones de acceso de I/O.

RR chequea el espacio libre en el servidor antes de ejecutarse, lo que ayudaría a conocer cuando agregar otro bloque en un servidor.

3.5.5.5. Depuración

GlusterFS ofrece la opción de realizar una depuración con el fin de poder encontrar posibles errores. Para ello se cuenta con la propiedad **trace.**

trace. Hace posible la producción de gran cantidad de información para fines de depuración, esta información es escrita en un archivo de logs que es propio de GlusterFS y se lo puede encontrar en */var/log/gluster/glusterfs.log.* (Depende de cómo instaló GlusterFS)

3.5.5.6. Características Extras

Estas características extras permiten aumentar la funcionalidad de GlusterFS ya que permite ampliar las utilidades que ofrece, una de estas características muy importante es **posix-locks** ya que con esta se logra establecer un bloqueo, propiedad que se la debería utilizar para archivos que van a ser accedidos y modificados muchas veces.

Además se tienen otras características que se detallan a continuación.

filter. Permite un filtrado avanzado basado en el nombre del archivo y/o atributo. Hasta la versión actual solo se permite exportar en modo *read-only*.

fixed-id. Muestra el id del usuario y grupo, como especificado para todos los archivos y carpetas.

3.5.5.7. Almacenamiento

posix. GlusterFS confía en sistemas de archivos basados en Disco tales como ext3[26] o XFS[27] para manejar la administración de bloqueo del dispositivo. Entonces la función que cumple **posix** es la de unir el GlusterFS server con el sistema de archivos basado en disco.

3.5.5.8. Autenticación

GlusterFS permite la autenticación entre los nodos, por medio de IP así como también utilizando usuario y password.

3.5.5.9. Encriptación

Esta característica es muy importante ya que con esta se logra tener seguridad de la información que se está transmitiendo entre los diferentes nodos del clúster.

Rot-13. Provee una encriptación y desencriptación de archivos utilizando el algoritmo rot-13[28]. Con lo que se logra proteger la información que se está transmitiendo entre cada uno de los nodos del clúster.

En el Anexo D, se encuentran ejemplos con cada una de estas características.

3.5.6. Reglas para escribir un archivo de configuración

En la versión actual de GlusterFS se cuenta con la consola de administración, con la cual ya no es requerido escribir archivos de configuración. Esta utilidad automáticamente va escribiendo en los archivos de configuración.

GlusterFS utiliza estos archivos de configuración para especificar qué es lo que se va a replicar hacia los nodos clientes.

Para la escritura de un archivo de configuración tanto en el cliente como en el servidor se debe seguir algunas reglas, con esto evitar cualquier tipo problema, así como también tener un orden en cada uno de los archivos de configuración

Las reglas son las que se detallan a continuación:

- Se utiliza el comodín "#" para realizar comentarios dentro del archivo, este se lo debe escribir al inicio del comentario.

[26] **EXT3:** (third extended filesystem o "tercer sistema de archivos extendido") es un sistema de archivos con *registro por diario*. Es el sistema de archivo más usado en distribuciones Linux.
[27] **XFS:** Es un sistema de archivos de 64 bits con journaling de alto rendimiento para su implementación de UNIX llamada IRIX. En mayo del 2000, se liberó XFS bajo una licencia de código abierto.
[28] **ROT-13:** (**rotar 13 posiciones**) Es un sencillo **cifrado por desplazamiento** utilizado para ocultar un texto sustituyendo cada letra por la letra que está trece posiciones por delante en el alfabeto. A se convierte en N, B se convierte en O y así hasta la M, que se convierte en Z. Luego la secuencia se invierte: N se convierte en A, O se convierte en B y así hasta la Z, que se convierte en M.

- Los archivos de configuración de GlusterFS son sensibles a mayúsculas y minúsculas.
- Los campos que no son especificados, GlusterFS la toma como valores por defecto.
- No hay un orden específico para cada una de las opciones que se pueden especificar en un volumen, pero se recomienda tener un estándar para cada uno de los volúmenes.
- Cada opción debería finalizar dentro de una misma línea.
- Todas las líneas que se encuentren en blanco o comentadas, no influirán en la configuración del archivo.
- El valor que se especifica en el nombre de sub-volume debe haber sido configurado antes de usarse en ese volumen, es decir que si se trata de utilizar sub-volume spool, pues el volumen spool tuvo que haber sido creado anteriormente, caso contrario tendremos problemas al levantar el servicio.

3.6. Cuadro comparativo de los sistemas de replicación

A continuación se presenta un cuadro comparativo con cada una de las características de los sistemas de replicación de archivos, esto nos permitirá tener un mejor criterio para seleccionar uno de los sistemas, que se utilizará en el clúster de servicios de mail y web.

Se tomaron en cuenta las características más importantes que debe cumplir el sistema de replicación de archivos para la implementación en el clúster.

No.	Detalle	Valor Ponderado
1	Replicación de archivos a varios equipos.	4
2	Encriptación de la información.	4
3	No posee un punto único de montaje.	4
4	Soporte bloqueo de archivos.	4
5	Punto único de fallo.	3
6	Soporte (Kernel).	3
7	Balanceo de Carga.	3
8	Escalabilidad para transferir grandes cantidades de datos.	3
9	Diseñado para ambientes de clúster.	2
10	Trabajo en modo desconectado.	2
11	Caching local.	2
12	No requiere de un servidor exclusivo para la implementación.	2
13	Licencia GPL.	2
14	Control de cuotas de disco	2

Tabla 3.3 Características cuantificadas

Características	Sistemas de replicación de archivos		
	GlusterFS	**CODA**	**OpenGFS**
1	4	4	4
2	4	4	0
3	4	0	4
4	4	0	4
5	3	3	3
6	3	3	0
7	3	0	0
8	3	0	0
9	2	0	0
10	0	2	0
11	0	2	2
12	2	0	2
13	2	2	2
14	2	0	0
Total	**36**	**20**	**21**

Tabla 3.4 Análisis comparativo de los sistemas de replicación

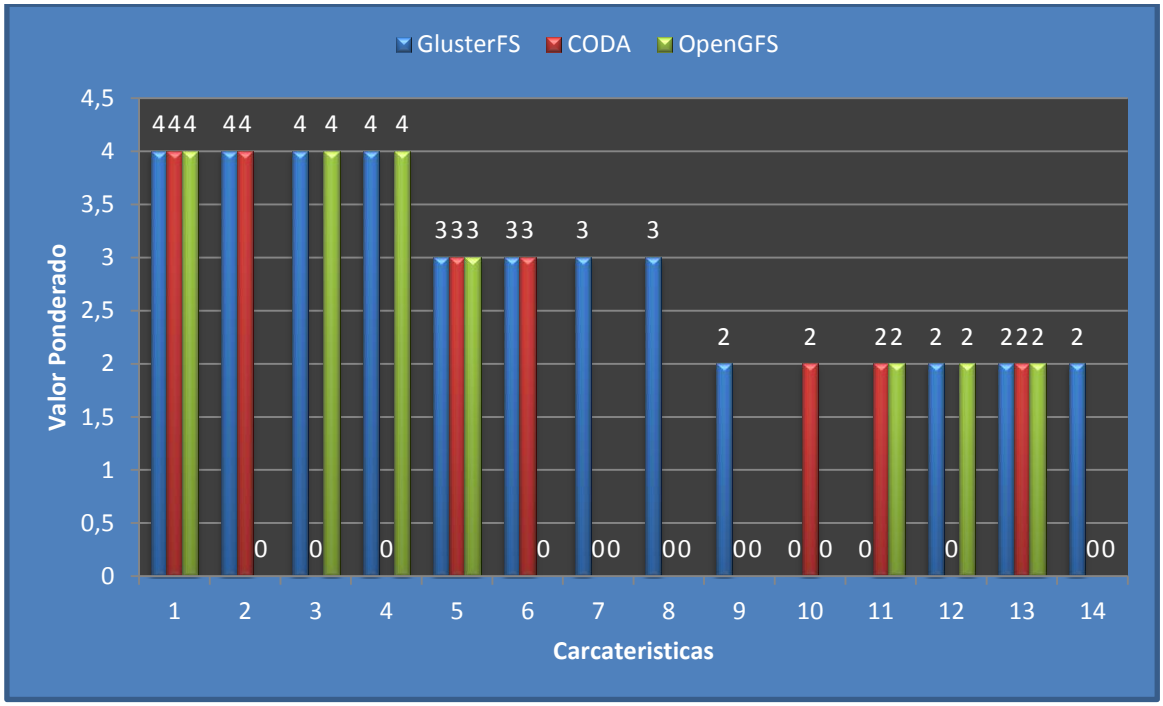


Fig. No. 3.5 Resultados del análisis comparativo

3.7. Conclusiones

El sistema de archivos GlusterFS es el Sistema de Archivos Distribuido que cubre mejor los requerimientos; lo cual se ha concluido por las siguientes razones:

- GlusterFS está diseñado para trabajar en un ambiente de clúster.
- Las características de GlusterFS permiten tener configurado varios sistemas redundantes logrando así evitar tener puntos únicos de fallo.

- Por la escalabilidad que este posee ya que se puede transmitir grandes cantidades de información sin tener ningún problema.
- En cuanto a lo que tiene que ver con los otros dos sistemas de replicación que se describieron no son orientados para los servicios que se pretenden configurar.
- Tanto CODA como OpenGFS presentan algunas desventajas y fueron algunas de ellas las que hicieron que la elección fuera por GlusterFS, tal como se muestra en la tabla 3.3.
- CODA requiere de gran cantidad de recursos como memoria RAM[29] ya que por ejemplo si se pretende replicar 100GB, se debe de separar el 4 % del espacio en disco para almacenar los metadatos, es decir se utilizaran 4GB para tal operación.

3.8. Selección del sistema de replicación de archivos

De lo anteriormente analizado se ha decidido utilizar el sistema de replicación de archivos GlusterFS, debido a que cubre la mayoría de requerimientos planteados en el capítulo dos, con lo cual se puede contar con un clúster de servicios de mail y web con un nivel óptimo de seguridad, tanto en la información que se está transmitiendo en el clúster, como en la información de los usuarios en los servicios web y mail.

[29] **RAM:** (Random Access Memory) Memoria de acceso aleatorio

CAPÍTULO IV: Solución e implementación de seguridad en el clúster de servicios de mail y web

En el presente capítulo se detalla la solución e implementación con la cual se contará en el clúster de servicios de mail y web, cumpliendo con cada uno de los requerimientos especificados en los capítulos anteriores.

4.1. Diseño de la solución

Esquema propuesto para la implementación de la seguridad en el clúster de servicio de mail.

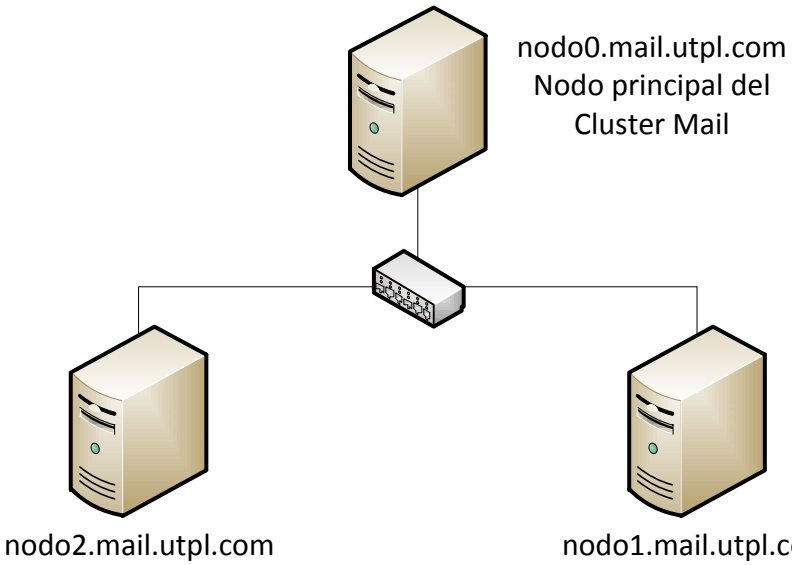


Fig. No. 4.1 Esquema de solución para implementación de seguridad en clúster de servicio de mail.

Características de los equipos que conforman el clúster de servicio de mail.

nodo0.mail.utpl.com	
Característica	**Valores**
IP	172.16.17.161
RAM	2GB
Procesador	Intel Core I7 2.0GHz
Sistema Operativo	Centos 5.6 64Bits

nodo1.mail.utpl.com	
Característica	**Valores**
IP	172.16.17.147
RAM	2GB
Procesador	Intel Core 2 Duo 2.0GHz
Sistema Operativo	Centos 5.6 64Bits

nodo2.mail.utpl.com	
Característica	**Valores**
IP	172.16.17.144
RAM	2GB
Procesador	Intel Core I7 1.8GHz
Sistema Operativo	Centos 5.6 64Bits

Esquema propuesto para la implementación de la seguridad en el clúster de servicio web.

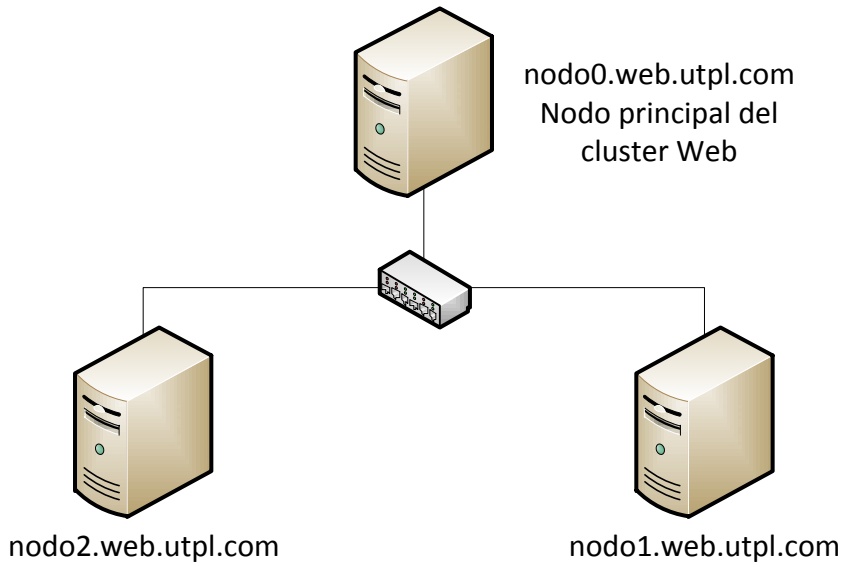


Fig. No. 4.2 Esquema de solución para implementación de seguridad en clúster de servicio web.

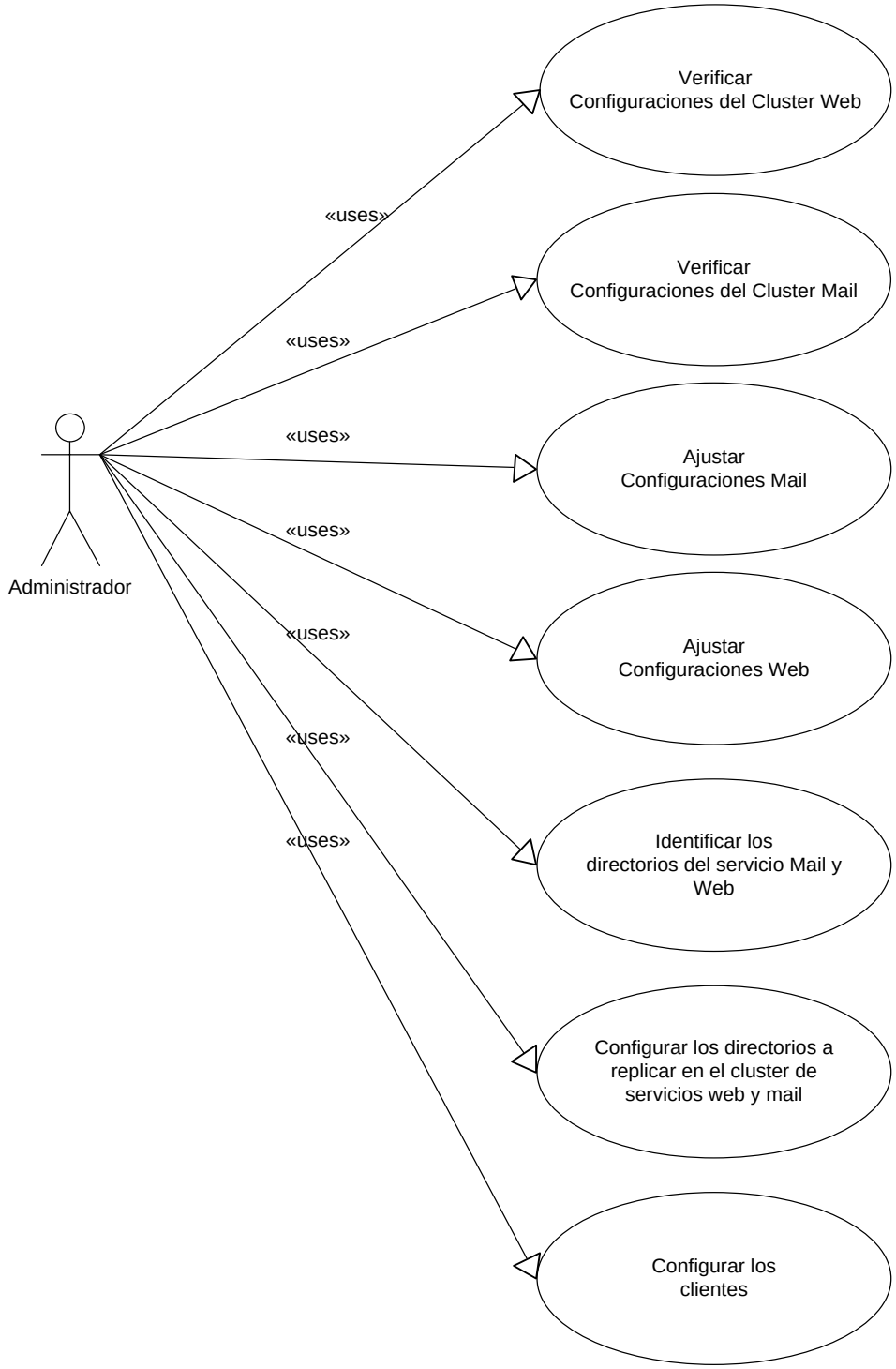
Características de los equipos que conforman el clúster de servicio de mail.

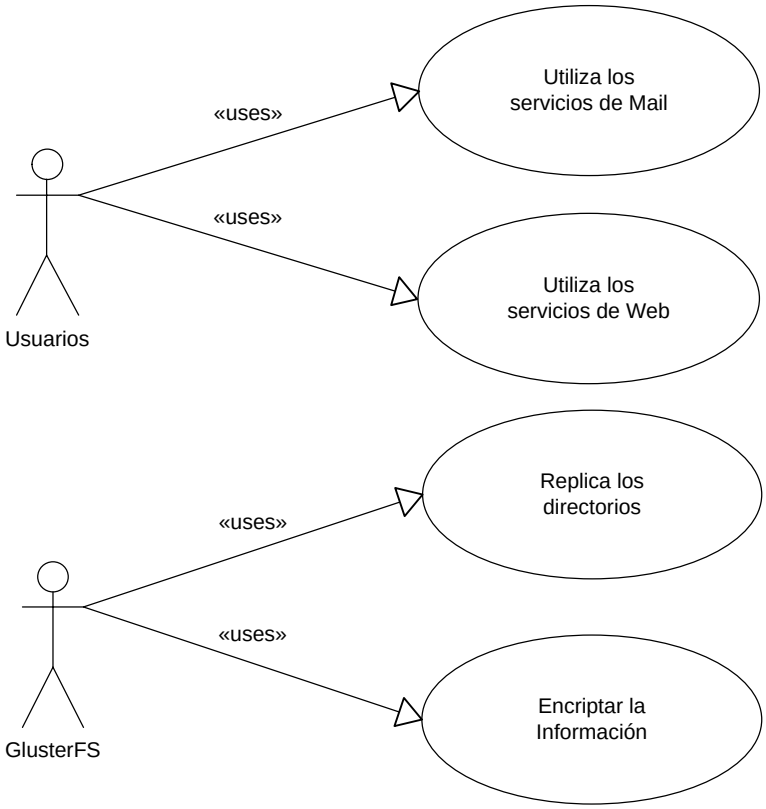
nodo0.web.utpl.com	
Característica	**Valores**
IP	172.16.17.150
RAM	2GB
Procesador	Intel Core I7 2.0GHz
Sistema Operativo	Centos 5.6 64Bits

nodo1. web.utpl.com	
Característica	**Valores**
IP	172.16.17.149
RAM	2GB
Procesador	Intel Core 2 Duo 2.0GHz
Sistema Operativo	Centos 5.6 64Bits

nodo2. web.utpl.com	
Característica	**Valores**
IP	172.16.17.148
RAM	2GB
Procesador	Intel Core 2 Duo 2.0GHz
Sistema Operativo	Centos 5.6 64Bits

Casos de Uso


Verificar
Configuraciones del Cluster Web
Verificar
Configuraciones del Cluster Mail
Ajustar
Configuraciones Mail
Ajustar
Configuraciones Web
Identificar los
directorios del servicio Mail y
Web
Configurar los directorios a
replicar en el cluster de
servicios web y mail
Configurar los
clientes
«uses»
«uses»
«uses»
«uses»
«uses»
«uses»
«uses»
Administrador



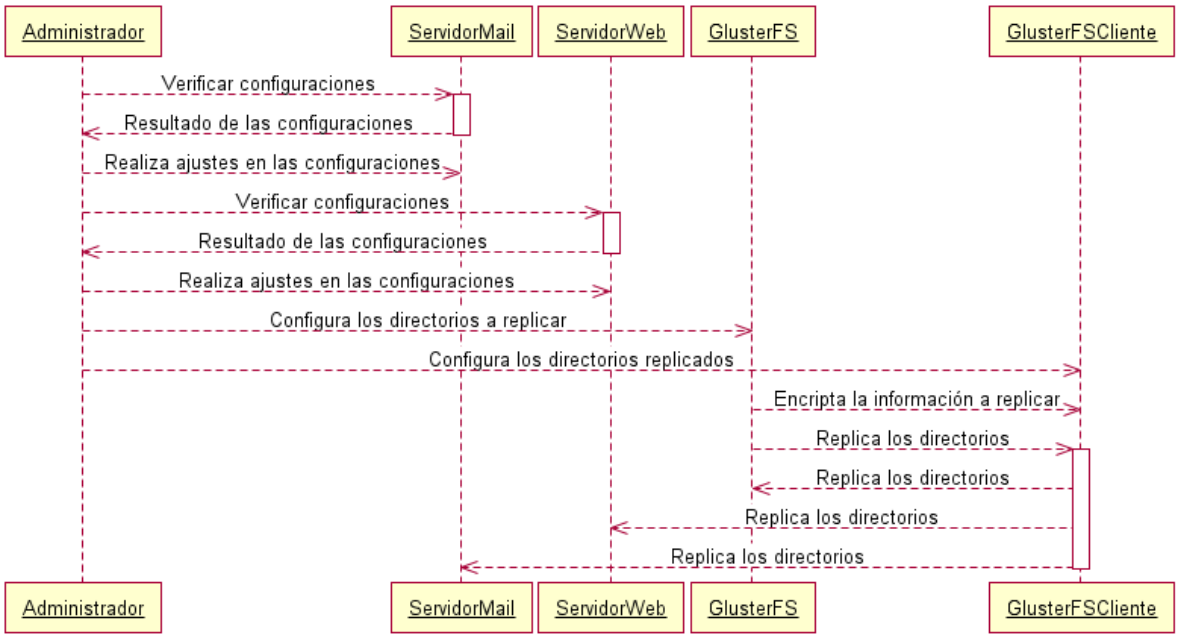
Actor	Casos de Uso
Administrador	Verificar configuraciones del Clúster Mail. Verificar configuraciones del Clúster Web. Realizar ajustes a las configuraciones del servidor Mail. Realizar ajustes a las configuraciones del servidor Web. Identificar los directorios del servicio Mail y Web. Configurar los directorios a replicar en el clúster de servicios de Web y Mail. Configurar los clientes.
Servidor GlusterFS	Replica los directorios. Encriptar la Información.
Usuario	Utiliza los servicios de Mail Utiliza los servicios Web

Descripción de Casos de Uso

Casos de Uso	Descripción
Verificar configuraciones del servidor Mail.	Esta es una tarea que el administrador del servidor mail debe realizar para controlar posibles errores que se puedan generar a partir de ellos.
Verificar configuraciones del servidor Web.	Esta es una tarea que el administrador del servidor web debe realizar para controlar posibles errores que se puedan generar a partir de ellos.
Realizar ajustes a las configuraciones del servidor Mail.	El administrador del servidor Web, de acuerdo con las revisiones realizadas a las configuraciones, deberá hacer los ajustes

	necesarios.
Realizar ajustes a las configuraciones del servidor Web.	El administrador del servidor Mail, de acuerdo con las revisiones realizadas a las configuraciones, deberá hacer los ajustes necesarios.
Identificar los directorios del servicio Mail y Web.	Se debe identificar cada uno de los directorios de los servicios de web y mail, que deberán ser replicados hacia los nodos que conformen en el clúster.
Configurar los directorios a replicar en el servidor.	Una vez que se tienen identificados cada uno de los directorios estos deben ser especificados en el sistema que realizara la replicación de los mismos.
Configurar los clientes.	En cada uno de los clientes se debe configurar cada uno de los directorios que están expuestos en el servidor de replicación.
Replicación de los directorios.	Este es un proceso que se encarga de realizar el servidor de replicación de archivos.
Utilizar los servicios de Mail	Aquí cada uno de los usuarios del servicio de mail accederá a su correo electrónico.
Utilizar los servicios Web	Los usuarios pueden visualizar todo el contenido de las páginas web que están alojadas en el servidor web

Diagrama de secuencia



4.2. Soluciones a problemas detectados en la gestión de información del clúster

Como se observó en el capítulo dos se propuso el cambio del sistema de replicación de archivos, para tal efecto se escogió utilizar el sistema GlusterFS por haber cumplido muchas de las características que se requieren para asegurar el clúster de servicios de mail y web.

4.2.1. Implementación de GlusterFS en el servidor

Primeramente se debe obtener los paquetes necesarios de las siguientes direcciones.

http://download.gluster.com/pub/gluster/glusterfs/3.2/3.2.2/CentOS

Los paquetes para el servidor son:

Paquetes rpm[30]:

glusterfs-core-3.2.2-1.x86_64.rpm
glusterfs-fuse-3.2.2-1.x86_64.rpm

4.2.1.1. Instalación de paquetes

Una vez obtenidos los paquetes se procede a la instalación.

Otra opción que se utiliza para instalar los paquetes es hacerlo desde la compilación del código fuente.

En el anexo E se encuentra detallado el proceso de instalación de estos paquetes.

4.2.1.2. Configuración de GlusterFS

Primeramente se definen los directorios que van a ser replicados, en el caso del servicio del mail son:

Directorios para el servicio de mail

- **/var/spool/mail** Aquí es donde se encuentran el buzón de correo electrónico de cada uno de los usuarios.
- **/public** En este directorio es donde se encuentran cada uno de los usuarios con sus respectivos mensajes indexados para luego ser usados por un cliente de correo electrónico.
- **/etc/mail** Aquí se encuentran los archivos de configuración del mail
- **/etc/passwd** Aquí se encuentran la información de cada uno de los usuarios.

Directorios para el servicio web

- **/mnt/disk2** En este directorio se encuentran alojadas todas las páginas que se visualizaran en el servidor web.

El archivo de configuración que se utiliza para realizar la replicación de archivos se lo encuentra bajo

[30] **rpm:** comando de Linux para la instalación, actualización y desinstalación de paquetes

/etc/glusterfs/gluster-server.vol, en este archivo se configura cada uno de los directorios antes mencionados.

En el anexo F se encuentra detallado el archivo de configuración para el servidor, con cada uno de los directorios que se utilizan en el clúster de servicios de mail y web.

En la version actual de gluster esta configuración se la puede realizar desde la consola de administración gluster de forma sencilla y rápida.

4.2.2. Implementación de GlusterFS en el cliente

Así mismo obtenemos los paquetes necesarios los cuales pueden ser descargadas de las siguientes direcciones.

http://download.gluster.com/pub/gluster/glusterfs/3.2/3.2.2/CentOS

Los paquetes que se deben descargar son:

Paquetes rpm

glusterfs-core-3.2.2-1.x86_64.rpm
glusterfs-fuse-3.2.2-1.x86_64.rpm

Además de los paquetes que se encuentra ahí se debe instalar un paquete fuse para que GlusterFS tenga soporte en el kernel.

4.2.2.1. Instalación del módulo fuse y los paquetes del cliente

Antes de proceder a instalar los paquetes de GlusterFS para el cliente, se debe instalar el módulo fuse.

El proceso para la instalación de estos paquetes se encuentra detallado en el anexo G

Otra opción que se utiliza para instalar los paquetes en el cliente es hacerlo desde la compilación de los códigos fuente, esta se la utilizará si existiere algún problema con los paquetes rpm.

4.2.2.2. Configuración del cliente

Se debe seleccionar cada uno de los volúmenes que se desean montar, esto se lo realiza en un archivo de configuración que se encuentra bajo /etc/glusterfs/glusterfs-client.vol.

En el Anexo H se encuentra detallado el archivo de la configuración de GlusterFS Client.

En la version actual de gluster esta configuración se la puede omitir, ya que la configuración se la realiza desde la consola de administración de gluster y esta se encarga de replicar tal configuración a cada uno de los nodos que conforman el clúster.

4.2.2.3. Configuraciones extras

Para automatizar el trabajo que tendría el administrador ante un eventual reinicio del sistema sobre el cual corre el servicio de glusterfs en el cliente, se ha configurado dos opciones que son:

- Carga automática del módulo fuse
- Montaje automático de los volúmenes

Carga automática del módulo fuse.

Para cargar el módulo de fuse automáticamente se debe editar el archivo que se encuentra bajo /etc/rc.modules y agregar la siguiente línea al final.

/sbin/modprobe fuse

Montaje automático de los volúmenes

Para realizar el montaje automático de los volúmenes editamos el archivo que se encuentra bajo /etc/fstab y al final de este la siguiente línea.

nodo0:/mail-vol /var/spool/mail glusterfs defaults 0 0

nodo0:/mail2-vol /public glusterfs defaults 0 0

nodo0:/ssap-vol /etc/passwd glusterfs defaults 0 0

nodo0:/web-vol /mnt/disk2 glusterfs defaults 0 0

Con esto logramos que tanto el módulo de fuse, como cada uno de los volúmenes se carguen automáticamente, tras un eventual reinicio del sistema.

Para obtener un mejor rendimiento con este sistema de ficheros hay que actualizar dovecot que actualmente es la **0.98** a una nueva versión de **dovecot 1.0** ya que con la que está actualmente no se puede sacar un buen provecho de GlusterFS, mientras que si se actualiza se manejan nuevas características que permites que dovecot trabaje con una sistema de archivos compartidos.

El proceso para realizar esta actualización así como para la configuración de las características nuevas se las encuentra en el anexo I.

4.2.3. Grupos de almacenamiento de confianza – Preparando GlusterFS para la administración.

Antes de configurar un volumen GlusterFS, es necesario crear grupos de almacenamiento de confianza que consiste en los servidores de almacenamiento donde se subirá el volumen.

Un grupo de almacenamiento es una red de los servidores de almacenamiento de confianza.

Para agregar servidores adicionales de almacenamiento, puede utilizar el comando **probe** en uno de los servidor de almacenamiento de confianza que se disponga.

En el anexo J se encuentran los pasos para configurar los grupos de almacenamiento de confianza.

4.2.4. Implementación de reglas de firewall en GlusterFS

Para tener una mayor protección en el clúster de servicios de mail y web se implementó reglas de firewall. De esta manera sólo se da acceso a equipos que están autorizados.

En el anexo K se encuentran las reglas de firewall.

Los puertos filtrados son los siguientes:

Número de Puerto	Descripción
6996	Puerto sobre el que trabaja GlusterFS.
25	Puerto que es utilizado por el SMTP.
111	Puerto utilizado por el portmap.
143	Puerto utilizado por imap.
110	Puerto que es utilizado por pop3.
993	Puerto utilizado por imaps
995	Puerto que es utilizado por pop3s.
443	Puerto utilizado para el servicio web seguro.
80	Puerto utilizado por el servidor web.
22	Puerto utilizado para conexiones remotas seguras.

Tabla 4.1 Lista de puertos filtrados

En anexo L se encuentran capturas de paquetes del servidor mail, en las cuales se verifica cada una de las implementaciones realizadas.

4.3. Soluciones a problemas detectados en la configuración del servicio de mail

A continuación se detalla el proceso para asegurar las diferentes vulnerabilidades de seguridad encontradas en el clúster de servicios de mail.

4.3.1. Aseguramiento de los Servicios POP3 e IMAP

Actualmente los servicios POP3 e IMAP son expuestos tal cual en su configuración por defecto, lo que se pretende ahora es asegurar estos protocolos y no permitir que estos sean vistos por personas no autorizadas.

Para lograr esto procedemos a configurar en el archivo dovecot.conf la opción de pop3s e imaps, con esto se logra que dichos protocolos pasen por un canal seguro, incrementando así la seguridad y también evitamos que estos protocolos puedan ser interceptados por terceras personas.

En el anexo M se encuentra el archivo de configuración de dovecot que está bajo /etc/dovecot.conf con los cambios que se realizaron para asegurar estos protocolos.

4.3.2. Implementación de autenticación SMTP

Esta característica es muy importante ya que si no se emplea la autenticación cualquiera puede tomar el nombre de otro y enviar mails no autorizados (conocido como suplantación de identidad), para poder solucionar este problema se procedió a hacer algunos cambios en los archivos de configuración de Sendmail (/etc/mail/sendmail.mc y /etc/mail/sendmail.cf), así como también se requirió de instalar un paquete que ayudará a implementar esta autenticación, este paquete se llama **cyrus-sasl-2.1.22-5.el5_4.3.x86_64.rpm**.

El proceso para realizar esta implementación la encuentra en el anexo N.

4.4. Soluciones a problemas detectados en la configuración del servicio web

A continuación se detalla el proceso para asegurar la vulnerabilidad de seguridad encontrada la configuración de apache en el clúster de servicios mail.

4.4.1. Implementación de páginas seguras para ingreso a servicios

Como se vio en el capítulo uno, actualmente el servidor de la UTPL no cuenta con páginas cifradas para el ingreso a otros sistemas como el campus virtual, ingreso al sitio ftp, entro otros.

Para ello se implementa el módulo de seguridad SSL en el servidor apache en el archivo de configuración que se encuentra bajo /etc/httpd/conf/httpd.conf. Aquí se debe agregar las siguientes líneas:

#Secure (SSL/TLS) connections

Include conf/extra/httpd-ssl.conf

La primera incluye un comentario para especificar que se está configurando la sección de SSL.

Con la segunda línea se habilita la configuración del módulo de seguridad de SSL.

CAPÍTULO V: Plan de Pruebas y validación

En el presente capítulo se realiza el plan de validación y pruebas del producto en este caso, se validan todas seguridades que se aplicaron al clúster de servicios de mail y web.

5.1. Introducción

Esta es una de las fases más importantes del presente proyecto de investigación de tesis ya que aquí se presentan los resultados del trabajo que se ha realizado y así poder determinar si dichos resultados son los que se esperaba.

Para determinar los resultados de cada una de las implementaciones realizadas sobre el clúster, se realizó un conjunto de pruebas que se listan a continuación:

- Pruebas de Integridad de Datos
- Pruebas de funcionalidad
- Pruebas de Carga
- Pruebas de Stress.

5.1.1. Propósito

El propósito con el que se realiza el plan de pruebas es para verificar y validar cada una de las implementaciones que comprende el producto final de la investigación que se compone algunas implementaciones hechas sobre el clúster de servicios de mail y web.

También verificar que funcione correctamente sobre los escenarios en el que se ejecutará.

Otra parte muy importante es poder medir el desempeño del producto final ante las situaciones que se lleguen a presentar.

Los objetivos del plan de pruebas son:

- Identificar los errores que se encuentren en la implementación.
- Realizar una verificación del comportamiento que tendrá el producto final en los escenarios en los que va a funcionar.
- Identificar cada una las estrategias de pruebas que se emplearan para la consecución de la misma.
- Una de las más importantes es la de corregir inconsistencias que se puedan hallar en el transcurso de las pruebas.

5.1.2. Alcance

Las pruebas se enfocarán a verificar lo siguiente:

- Se cuente con la misma información del servidor principal en todos los nodos del clúster.
- Los usuarios al ingresar a cualquiera de los nodos cuenten con la misma información.
- Los correos electrónicos no se corrompa al replicarse de un servidor a otro.
- La información que se transmite en el clúster esté protegida.
- Verificar que la capacidad del sistema es adecuada para la demanda de trabajo.
- Detectar posibles cuellos de botella
- Determinar el tiempo medio de respuesta que obtendrá el usuario.
- Determinar el máximo número de sesiones concurrentes, inicios de sesión por hora o transacciones por segundo que nuestro sistema es capaz de soportar mientras proporciona un nivel de rendimiento aceptable.
- Determinar cómo reacciona el sistema ante cargas de trabajo que excedan su capacidad.
- Identificar las transacciones más lentas y las más rápidas.

Cabe recalcar que estas pruebas pueden estar limitadas por diversos factores que se pueden suscitar como:

- Disponibilidad del Hardware.
- Personal disponible, se refiere a la capacidad de poder contar con personal suficiente para la concesión de las pruebas ya que estos requieren de tiempo.

5.1.3. Audiencia

La audiencia que se involucrada en este plan es la siguiente:

- Investigadores.
- Usuarios (Administradores, Docentes y Estudiantes)

5.2. Recursos

Para la ejecución de las pruebas se hace uso de los recursos humanos y recursos tecnológicos.

5.2.1. Recursos Humanos

Se ha hecho uso del siguiente recurso humano.

Recurso	Nombre
Diseñador de pruebas	Sr. Luis Cuenca
Probadores / ejecutores de pruebas	Ing. Samanta Cueva (Directora) Sr. Luis Cuenca (Administrador) Srta. Diana Ortega (Estudiante) Sr. Byron Alvarez (Estudiante)

Tabla 5.1 Recursos Humanos

5.2.2. Recursos Tecnológicos

Se ha hecho uso de los siguientes recursos tecnológicos.

Servicio de Mail	
Recurso	Nombre / Tipo
➢ Red/SubNet ➢ Nombre de Servidor ➢ Dirección IP del servidor	172.16.17.0 nodo0.utpl.com 172.16.17.161
➢ Red/SubNet ➢ Nombre de Servidor ➢ Dirección IP del servidor	172.16.17.0 nodo1.utpl.com 172.16.17.147
➢ Red/SubNet ➢ Nombre de Servidor ➢ Dirección IP del servidor	172.16.17.0 nodo2.utpl.com 172.16.17.144
Servicio Web	
Recurso	Nombre / Tipo
➢ Red/SubNet ➢ Nombre de Servidor ➢ Dirección IP del servidor	172.16.17.0 nodo0w.utpl.com 172.16.17.150
➢ Red/SubNet ➢ Nombre de Servidor ➢ Dirección IP del servidor	172.16.17.0 nodo1w.utpl.com 172.16.17.149
➢ Red/SubNet ➢ Nombre de Servidor ➢ Dirección IP del servidor	172.16.17.0 nodo2w.utpl.com 172.16.17.148
➢ 3 portátiles como clientes.	

Tabla 5.1 Recursos Tecnológicos.

5.3. Identificación de los Sistemas a Probar

El producto final de la presente investigación tiene como objetivo cumplir los siguientes requerimientos:

- Replicación correcta y segura de la información que se transmite en el clúster.
- Cifrado de los mensajes de correo electrónico.
- La disponibilidad de la información que los usuarios puedan tener al ingresar a cualquiera de los nodos del clúster

5.4. Estrategia para el Plan de Pruebas

A continuación se dará una explicación de cada una de las pruebas, que se ejecutan.

Las pruebas que se realizaron son las que se listan a continuación:

5.4.1. Pruebas de funcionalidad

Las pruebas de funcionalidad, busca comprobar que la información se replique de un nodo a otro y exista consistencia entre estos nodos.

Objetivo de la prueba:	Verificar que el método escogido para la replicación, permita tener la misma información en todos los nodos.
Técnica:	Envío de Mails hacia una determinada cuenta del servidor de correo. Peticiones al servidor web. Haciendo uso de un sniffer[31] interceptar los paquetes que se transmiten en la red.
Criterio de conclusión:	Verificar que el dueño de la cuenta a la que se envió los mail, pueda verlos en todos los nodos del clúster. Verificar que la información que se muestra es la que se tiene en todos los nodos. Verificar que la información que se transmite en el clúster pase encriptado a través de la red.

[31] **Sniffer:** Programa informático usado para interceptar y detectar información sobre configuraciones en la red.

Para la realización de esta prueba se seleccionó un grupo de PC's clientes que se encuentren en la subred 172.16.17.0/24.

Se realizaron 50 casos de pruebas y tras la ejecución de estas en el clúster web y mail se obtuvieron los siguientes resultados:

Número de peticiones	Resultado
48	Correctas
2	Incorrectas

Tabla. 5.1 Peticiones al servidor web

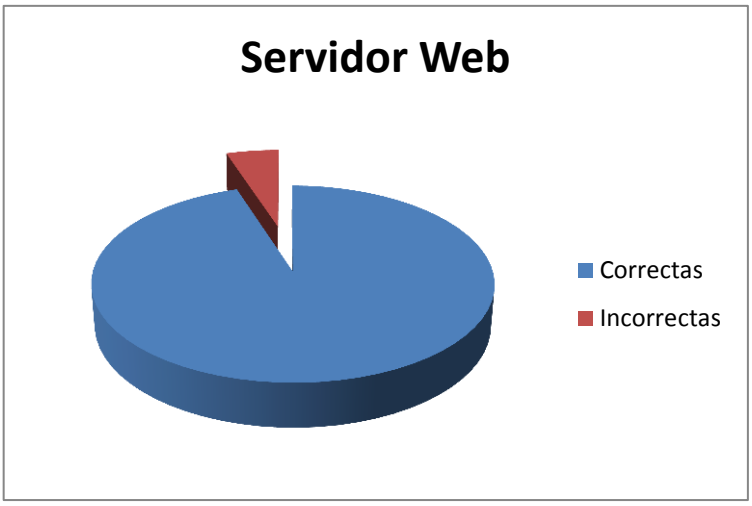


Gráfico. 5.1 Funcionalidad del servidor Web

Número de peticiones	Resultado
49	Correctas
1	Incorrectas

Tabla. 5.2 Peticiones al servidor mail

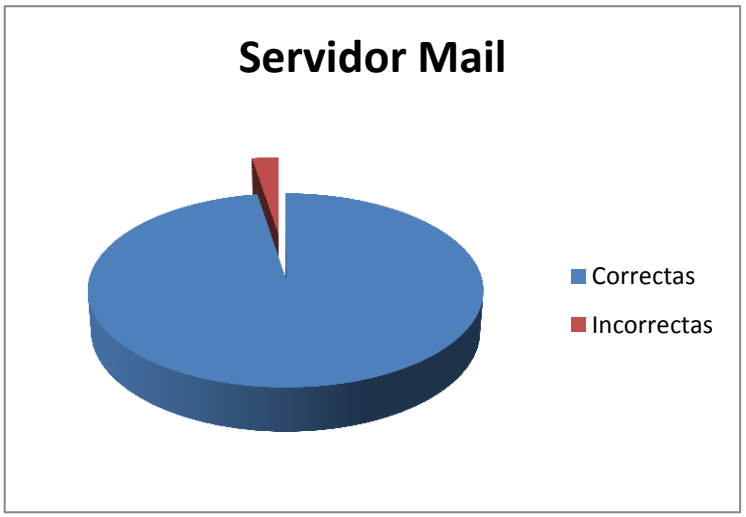


Gráfico. 5.2 Funcionalidad del servidor Mail

Como se puede apreciar en las gráficas el sistema de replicación obtuvo un 96.25% de cumplimiento con esta prueba.

En el anexo L se puede apreciar los resultados que se obtuvieron al utilizar el Sniffer en el momento de realizar las peticiones al servidor web y mail.

5.4.2. Pruebas de integridad de datos

Esta prueba tiene el objetivo de comprobar que durante la replicación de los Mails de un servidor a otro, la información no tenga ninguna alteración.

Los datos tienen que ser los mismos que los que están en los servidores.

Objetivo de la prueba:	Verificar que la información no haya sufrido alteraciones durante la replicación.
Técnica:	Enviar mails con archivos adjuntos, hacia una cuenta del servidor de correos. Utilizar Outlook para hacer el envío de mails y utilizar el Webmail para poderlos revisar. Realizar peticiones hacia el servidor web. Subir contenido al servidor Web.
Criterio de conclusión:	Que la información del mail enviado llegue completa sin sufrir alteración alguna, es decir llegue tal cual se la envió. Que las peticiones al servidor web, muestren la misma información que la de los nodos.

Esta prueba se la realizó sobre el Servidor web y mail.

En el **servidor web** para verificar que los datos se replicaron correctamente y sin ninguna alteración se procedió a subir contenido al servidor web y realizar peticiones al servidor web desde un conjunto de PC's en la subred 172.16.17.0/24 y verificar que en los nodos se muestre la información que fue subida en el servidor web.

En el **servidor mail** para realizar la verificación de esta prueba se procedió a hacer envíos de mail con y sin archivos adjuntos, luego se verificó que los datos enviados fueron recibidos de forma correcta y sin sufrir modificación alguna.

A continuación se ilustran las gráficas que muestran los resultados de la ejecución de las pruebas en el servidor web y mail:

Número de peticiones	Resultado
78	Correctas
2	Incorrectas

Tabla. 5.3 Peticiones al servidor web

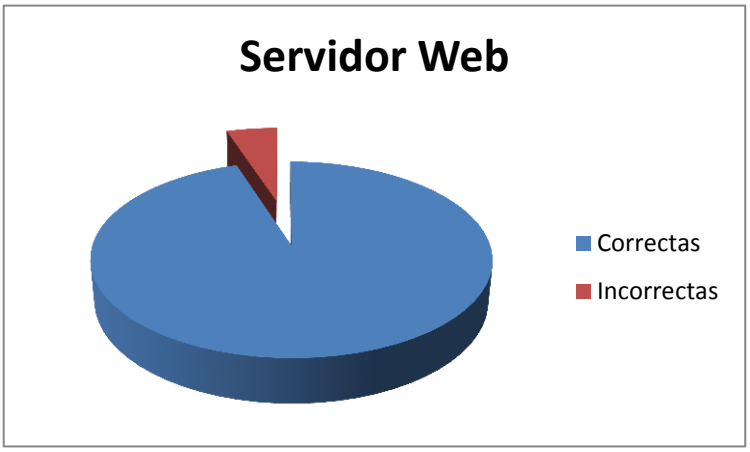


Gráfico. 5.3 Integridad del servidor Web

De las 80 peticiones hechas al servidor mail, las 77 fueron procesadas correctamente y mostraron la información sin alteración, existieron 3 peticiones iniciales en las que se tuvo que transcurrir un lapso de tiempo para poderlos visualizar.

Número de peticiones	Resultado
77	Correctas
3	Incorrectas

Tabla. 5.4 Peticiones al servidor mail

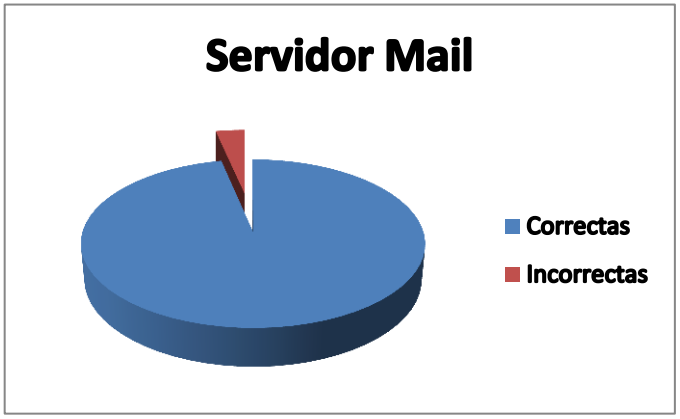


Gráfico. 5.4 Integridad del servidor Mail

Tras la ejecución de estas pruebas, se pudo observar que la replicación de archivos se realiza de una manera correcta sin sufrir ninguna alteración ya que el 97.5% de todas las peticiones hechas al servidor web mostraron la misma información y los correos

electrónicos fueron recibidos satisfactoriamente en un 96.25% sin alteraciones, con esto se está comprobando una vez más la eficiencia y seguridad que posee GlusterFS.

5.4.3. Pruebas de carga

Esta prueba tiene el objetivo de comprobar la capacidad que tiene el sistema de replicar los archivos a todos los nodos.

Objetivo de la prueba:	Obtener un tiempo de respuesta ante determinados tamaños de mails y tamaño de contenido en el servidor web.
Técnica:	Enviar un mail sin adjunto Enviar un mail con un adjunto de 1MB. Enviar un mail con un adjunto de 2MB. Enviar un mail con un adjunto de 6MB. Enviar un mail con un adjunto de 10MB. Subir contenido al servidor web de 10MB. Subir contenido al servidor web de 50MB. Subir contenido al servidor web de 100MB. Subir contenido al servidor web de 500MB.
Criterio de conclusión:	Que el sistema fue capaz de replicar tal cantidad en un determinado tiempo.

En el **servidor mail** para realizar la verificación de esta prueba se procedió a hacer envíos de mail con archivos adjuntos de distintos tamaños desde clientes de correo y desde el Webmail hacia el servidor de correo, luego a través de los mismos se verificó que los mails fueron recibidos de forma correcta con sus adjuntos respectivos.

A continuación se presenta una tabla con los resultados de cada uno de los casos de prueba.

Cantidad	Tamaño MB	Tiempo (seg.)
1	0	Menos de 1
1	1	Menos de 1
1	2	Menos de 1
1	6	Menos de 1
1	10	Menos de 2

Tabla. 5.5 Tiempos de respuesta del servidor mail

En el **servidor web** para realizar la verificación de esta prueba se procedió a subir contenido de distintos tamaños en el servidor web, luego se procedió a realizar peticiones en los nodos clientes para verificar que el contenido fue replicado.

A continuación se presenta una tabla con los resultados de cada uno de los casos de prueba.

Cantidad	Tamaño MB	Tiempo (seg.)
1	10	Menos de 1
1	50	Menos de 1
1	100	Menos de 3
1	500	Menos de 6

Tabla. 5.6 Tiempos de respuesta del servidor web

Con esto se concluye que GlusterFS posee gran capacidad de replicación ya que los tiempos que se pueden ver son menores a 2 segundos para el caso del servidor mail y tiempos menores a 6 segundos para el servidor web.

5.4.4. Pruebas de Stress

Esta prueba tiene el objetivo de ver el comportamiento que tiene el clúster cuando se pretende replicar varios mails en un mismo instante.

Objetivo de la prueba:	Verificar el comportamiento del sistema de replicación de archivos ante envíos simultáneos de mails y peticiones al servidor web.
Técnica:	Desarrollar un script que envíe gradualmente cantidades de mails. Desarrollar un script que envíe peticiones simultáneas al servidor web.
Criterio de conclusión:	Tiempos de respuesta de acuerdo a la cantidad de mail que se envió, así como también comprobar los envíos satisfactorios.

En el **servidor web** para realizar las pruebas de Stress se procedió a desarrollar un script el mismo realiza múltiples peticiones al servidor web.

El script contiene el siguiente código.

```
for id in {1..$np} ; do elinks –dump
http://nodo0w.web.utpl.com ; done
```

En donde:

for: utilizado para crear un ciclo.
$np: especifica el número de peticiones.
elinks: Utilizado para realizar una petición a un servidor web

Este script fue ejecutado desde una consola de Linux en un equipo conectado a la subred 172.16.17.0/24. De la siguiente manera:

[root@nodo0w ~]#./peticionesweb 10

Donde 10 significa el número de peticiones.

Se realizaron varios casos de pruebas con distintas cantidades de peticiones y los resultados son los que se muestran en la siguiente tabla.

Nro. de peticiones	Tiempo (segundos)
1	0
5	1
10	2
20	2.5
40	3
80	5
160	7
320	13
640	25
1280	49

Tabla 5.7 Peticiones al servidor web

Las peticiones se han realizado sobre contenido simple, ya que al realizar peticiones sobre contenido de mayor tamaño va a existir un tiempo de respuesta más largo debido al número de peticiones y al tamaño que ocupara cada petición.

Para las pruebas en el **servidor mail**, también se desarrolló un script el cual enviaba múltiples mails a un destinatario en un mismo instante.

El código del script es el siguiente:

for test int {1..$nm}; do echo "Test de prueba" | mail –s "test #"$test prueba@nodo0.mail.utpl.com; done

Este script fue ejecutado desde una consola de Linux en un equipo conectado a la subred 172.16.17.0/24. De la siguiente manera:

[root@nodo0 ~]#./peticionesmail 40

Donde 40 significa el número de mails enviados.

Luego de ejecutar el script con distintas cantidades se obtuvo los resultados que se muestran en la siguiente tabla.

Nro. de Mails enviados	Tiempo(seg)
10	1
50	4
100	9
200	19
500	52
1000	90
5000	Fallo

Tabla 5.8 Peticiones al servidor mail

Como se puede observar en esta tabla se tuvo un fallo al momento de hacer un envío de 5000 mails simultáneos, esto se debió a que se hizo la prueba en uno de los nodos del clúster y esta tenía que procesar el envío y la recepción de cada uno de los correos. Este inconveniente puede ser resuelto con el aumento de las capacidades de los servidores.

DISCUSIÓN

Conocer y manejar bien las configuraciones de cada uno de los servicios de web y mail es de vital importancia, ya que al poseer este conocimiento se está en la capacidad de poder cambiar ciertas configuraciones que vienen dadas por defecto al momento de instalaros, ya que al no cambiar estas configuraciones genéricas se verían comprometidos los servicios de Mail y Web porque las configuraciones que están por defecto son de conocimiento público por el mismo hecho de que son aplicaciones de código abierto.

La clave para estar preparados ante un eventual ataque a un servicio es el de realizar distintas simulaciones de ataques, ya que con estas simulaciones se tiene una perspectiva de lo que se podría tener en un posible ataque a un servicio.

El crear varios escenarios de ataques permite a los administradores de los servicios tener una base de conocimientos para poder dar solución a problemas de seguridad que se podrían originar en un momento dado. El éxito de estas simulaciones está en recolección de la información que se va generando al momento de ir dando solución a los distintos escenarios que se planteen.

Se contó con herramientas sniffer que permitieron la captura del tráfico que pasa a través del clúster, gracias a esto se logró detectar que toda la información pasa por el clúster de manera insegura, cualquier intruso puede fácilmente interceptar la información de los servicios de web y mail, por esta razón los servicios se ven afectados lo que implica en graves riesgos de seguridad para la organización que en este caso es la UTPL, esto se lo puede apreciar en el anexo A y B.

La información que se maneja en el servidor mail es de carácter confidencial, ya que se trata de información privada de cada uno de los usuarios de este servicio. Teniendo en cuenta esta característica se consideró adecuado el cambio del sistema de archivos para la gestión de la información, ya que el actual NFS no brinda la seguridad adecuada que se debería tener para la manipulación de este tipo de información.

Sería adecuado que el servidor de servicio Web de la UTPL cuente con Certificados Digitales de una entidad certificadora reconocida, esto ayudará a tener un respaldo en el servicio, de esta manera hacer que los servicios sean seguros y confiables para los usuarios que interactúan con los mismos.

La solución que se ha dado para manejar la gestión de información la considero que es muy conveniente ya que el nuevo sistema de replicación de archivos que se utilizó GlusterFS, se adaptó a los requerimientos que se tenían para los servicios, ahora si bien es cierto de que existen muchos sistemas de replicación de archivos se debe realizar un análisis de todas las características que tienen cada uno de los servicios, ya que de ello depende el seleccionar un sistema de replicación de archivos que se adapte a los servicios.

La implementación de GlusterFS es de gran beneficio para el servicio de Mail ya que se comprobó la eficacia que tiene para trabajar en ambientes de clúster, debido a que cuenta con varias características las cuales deben ser aprovechadas al máximo.

Como todo sistema tiene sus ventajas y desventajas, GlusterFS no sería la excepción pues el algoritmo que utiliza para realizar la encriptación es simple, la alternativa que se debe manejar para solventar esta deficiencia es la utilización de canales de conexión seguros utilizando SSL, aprovechando de que el diseño de GlusterFS si lo permite y a más de esta alternativa actualmente se está trabajando para el desarrollo de otro algoritmo de encriptación que sustituirá al que se tiene actualmente.

La nueva version de GlusterFS cuenta con nuevas características, entre las cuales podemos destacar la Geo-Replication que consiste en replicar la información hacia servidores remotos que no necesariamente estén sobre la misma red interna, esto sería de gran ayuda ya que si se tiene la implementación de esta característica sería muy fácil la recuperación de la información ante cualquier eventualidad catastrófica que pueda sufrir los servidores, y con esto se lograría contar con un 99.9% de disponibilidad del servicio.

Conclusiones y recomendaciones

CONCLUSIONES

A continuación se presenta las conclusiones a las que se ha llegado al término de este proyecto.

- En el análisis realizado en los clústers se identificó que existían bajos niveles de seguridad implicando graves riesgos para de seguridad en la información de la organización.

- Con GlusterFS se evita que el administrador tenga que ejecutar scripts de sincronización de datos; ya que el sistema permite la replicación de forma automática y todos los cambios realizados son reflejados de manera automática hacia los demás nodos.

- Con el uso de la consola de administración de GlusterFS se hace más fácil la administración de los directorios a replicar hacia los nodos, lo cual evita cometer al usar archivos de configuración.

- GlusterFS es el que mejor se adapta a los requerimientos del clúster de servicios web y mail de la UTPL, es apropiado para trabajar con ambientes de clúster y posee características que ayudan a mejorar el rendimiento, seguridad y escalabilidad del clúster.

- Con el resultado obtenido en las pruebas de funcionalidad se pudo comprobar la eficiencia de GlusterFS, llegando a obtener un 97.7% en la replicación de archivos, el 2.3% faltante se debe a ajustes que se realizaron sobre la configuración de GlusterFS.

- Se obtuvo un 97% en los resultados de las pruebas de integridad de datos, en la cual se pudo comprobar que no existió alteración de datos, verificando que GlusterFS es un sistema seguro para la replicación de archivos, el 3% faltante se debe a que estas pruebas eran realizadas desde el mismo servidor hacia los nodos, esto ocasionó que existiera mayor procesamiento en el clúster.

- El servidor mail respondió satisfactoriamente en un 96.25% ante las peticiones simultáneas hechas al servidor, el 3.75% faltante se debe a que las peticiones eran enviadas desde el mismo servidor hacia los nodos, esto ocasionó que exista retardo en las respuestas.

- El servidor web respondió satisfactoriamente ante los distintos casos de prueba de peticiones simultáneas hacia el servidor web.

- Las características de GlusterFS permiten tener configurado varios sistemas redundantes logrando así evitar tener puntos únicos de fallo.

- Con la implementación de los protocolos seguros POP3S, IMAPS y HTTPS se logró el aseguramiento en un 98% de los servicios web y mail evitando con esto que puedan ser interceptados por terceras personas, no se logra el 100% ya que cada vez existen nuevas técnicas para burlar la seguridad que se implemente.

- La característica rot-13 que posee GlusterFS para la encriptación de la información, es débil ya que se trata de un algoritmo simple, actualmente el grupo de desarrollo de GlusterFS está desarrollando un mecanismo más fuerte para la encriptación de los datos que se transmiten por el clúster.

RECOMENDACIONES

- Al momento de realizar la configuración de servicios de web y mail, estas se las debe personalizar a los requerimientos de la organización ya que al tener configuraciones por defecto, estas pueden desencadenar en grandes problemas de seguridad que pueden ser aprovechadas por terceras personas, en perjuicio de la organización.

- Los administradores de los servidores deben estar revisando continuamente las nuevas versiones del software que es necesario para el servicio web y mail ya que siempre hay nuevas características que se les puede sacar provecho, para lograr un mejor rendimiento del servicio.

- Adquirir Certificados SSL para el servidor Web a entidades certificadoras como: ***VeriSing, Thawte, RapidSSL, GeoTrust.*** Con esto se garantiza la privacidad de la información, protección de la información durante el proceso de transmisión de datos entre el usuario y el servidor.

- En cada uno de los nodos del clúster se debe realizar una revisión periódica de la integridad de los datos en los nodos del clúster, con esto evitar problemas con los mismos.

- Para seleccionar un Sistemas de Archivos Distribuido, se debe realizar un análisis previo del ambiente sobre el cual será utilizado, ya que hay sistemas de archivos que trabajan para ambientes específicos.

BIBLIOGRAFÍA

Contento, M. & Tinoco, D. Gestión Implementación y capacidad de crecimiento de un clúster para balanceo de carga de los protocolos HTTP y SMTP en los servidores Linux de la UTPL.

Baker, M. *Cluster Computing.* University of Portsmouth.

Braam, P., Baron, R., Harkes, J., & Schnieder, M. *The Coda HOW TO.* School of Computer Science, Carnegie Mellon University.

Cohen, F. (n.d.). *A Preliminary Classification Scheme for Information System Threats, Attacks, and Defenses; A Cause and Effect Model; and Some Analysis Based on That Model".* Retrieved from http://www.all.net/journal/ntb/cause-and-effect.html

Dovecot. (n.d.). Retrieved Julio 28, 2011, from http://wiki.dovecot.org/FrontPage

Gluster. (n.d.). Retrieved Julio 28, 2011, from http://www.gluster.org/docs/index.php/GlusterFS

Gluster. (n.d.). Retrieved Julio 30, 2011, from http://www.gluster.com/community/documentation/index.php/Gluster_3.2_Filesystem_Administration_Guide

Guía de Administración de Redes con Linux. (n.d.). Retrieved from http://www.ibiblio.org/pub/Linux/docs/LuCaS/Manuales-LuCAS/GARL2/garl2/index.html

Informit. (n.d.). Retrieved Julio 28, 2011, from http://www.informit.com/articles/article.asp?p=372008&rl=1

Internautas. (n.d.). Retrieved Julio 28, 2011, from http://www.internautas.org/archivos/port-numbers.txt

Linux Focus. (n.d.). Retrieved Julio 28, 2011, from http://www.linuxfocus.org/Castellano/November2000/article179.shtml

Linux Virtual Server. (n.d.). Retrieved Julio 28, 2011, from http://www.linuxvirtualserver.org/docs/ha/keepalived.html

Manual de seguridad de Red Hat Enterprise Linux 4. (n.d.). Retrieved from http://www.opencontent.org/openpub/

Moore, A., Ellison, R., & Richard, C. (n.d.). *Attack Modeling for Information Security and Survivability.* Retrieved from http://www.cert.org/archive/pdf/01tn001.pdf

Roselló Vicente, J. A. (n.d.). *Clustering de Alta Disponibilidad bajo GNU/Linux.* Retrieved from http://www.bisente.com/documentos/clustering/informe.pdf

Schneier, B. (n.d.). *Attack trees.* Retrieved from http://www.schneier.com/paper-attacktrees-ddj-ft.html

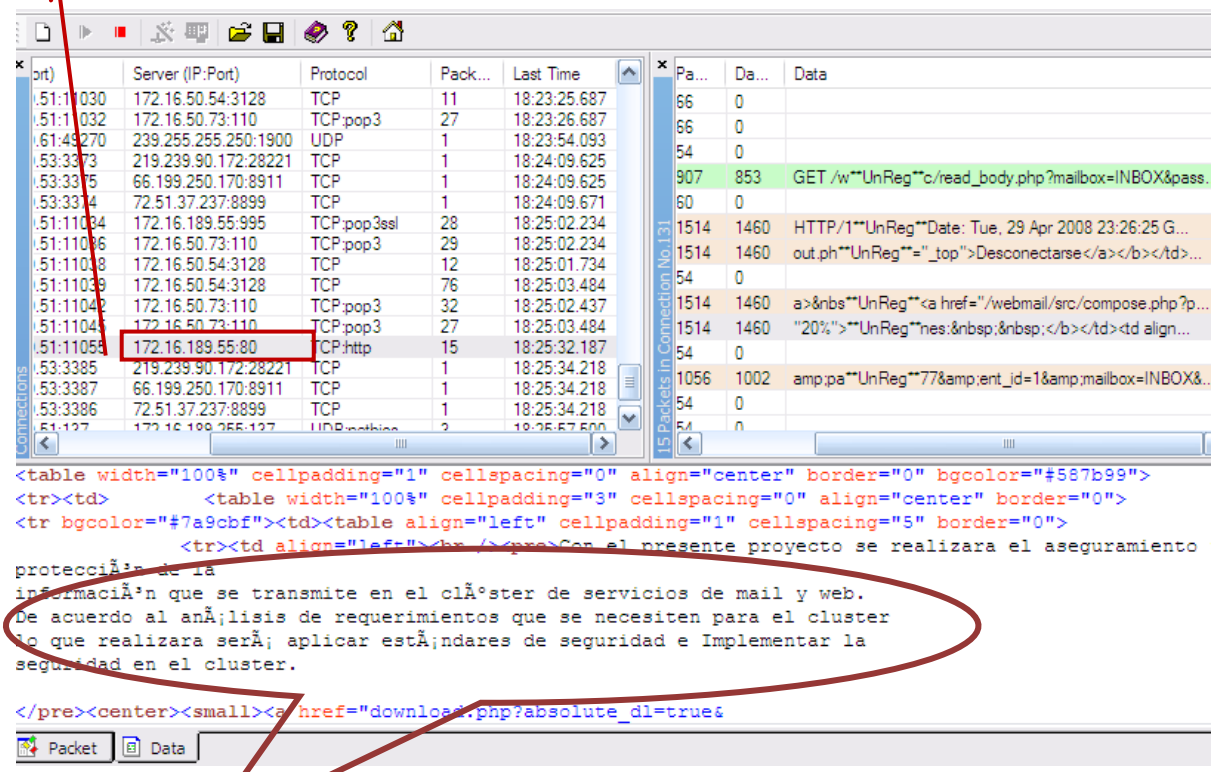
Source Forge. (n.d.). Retrieved Julio 28, 2011, from http://opengfs.sourceforge.net

ANEXOS

ANEXO A

Visualización de correos electrónicos cuando son abiertos por los usuarios en el cliente web (Webmail) Esto se lo realizo utilizando una herramienta que permite la captura de paquetes en la RED

Dirección del servidor.



Cuerpo del mensaje de correo electrónico.

Como podemos observar los mensajes de un usuario de correo electrónico que son abiertos a través de un cliente web son claramente visibles.

Herramientas utilizadas:

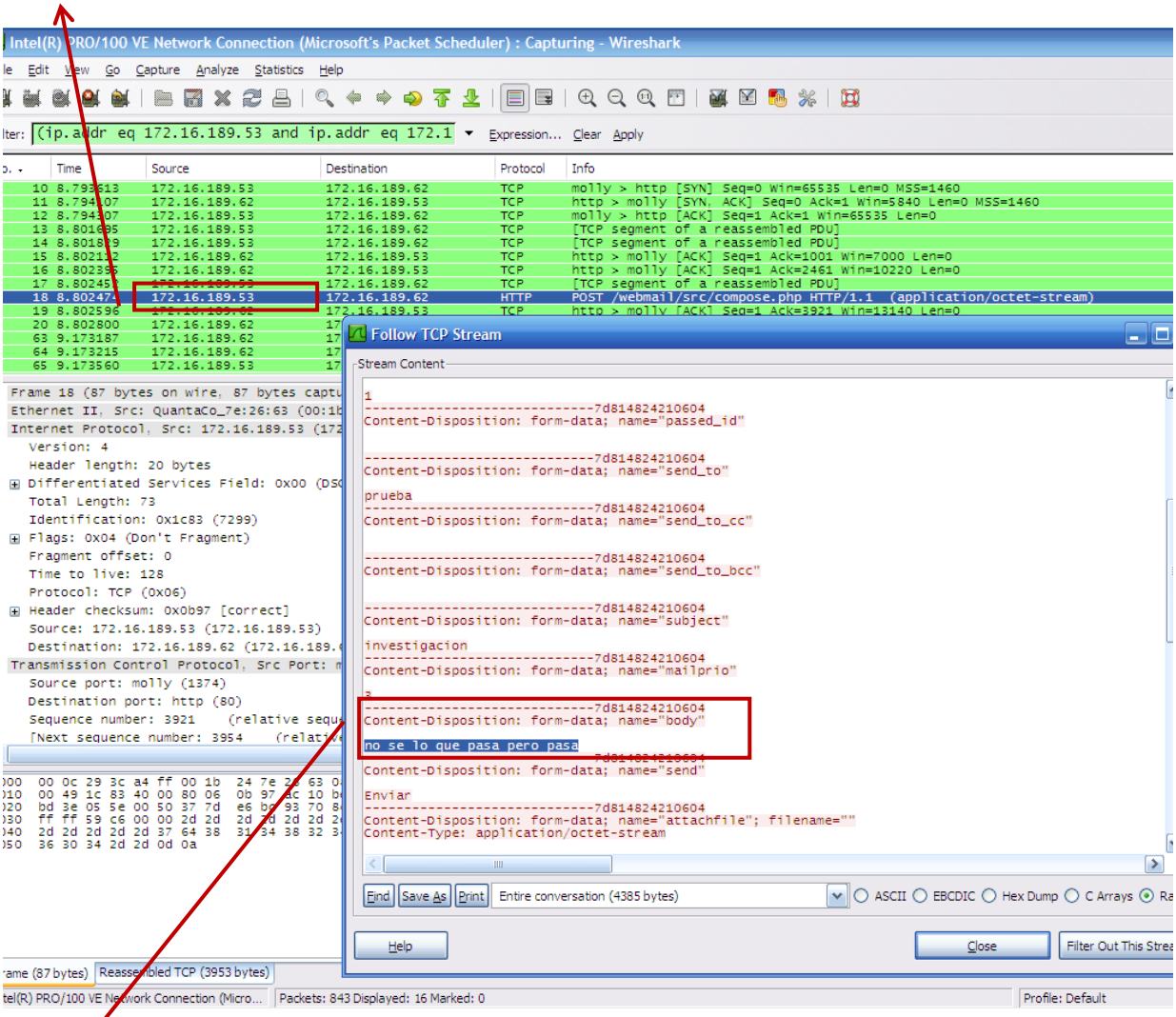
Herramienta utilizada para capturar los paquetes: **EtherDetect Packet Sniffer**

Cliente mail utilizado: **Webmail**

Navegador: **Internet Explorer**

Aquí tenemos otra vista, se está haciendo uso uno de los nodos del clúster y de otro cliente de correo.

Dirección del servidor.



Cuerpo del mensaje de correo electrónico.

Herramientas utilizadas:

Herramienta utilizada para capturar los paquetes: **Wireshark**

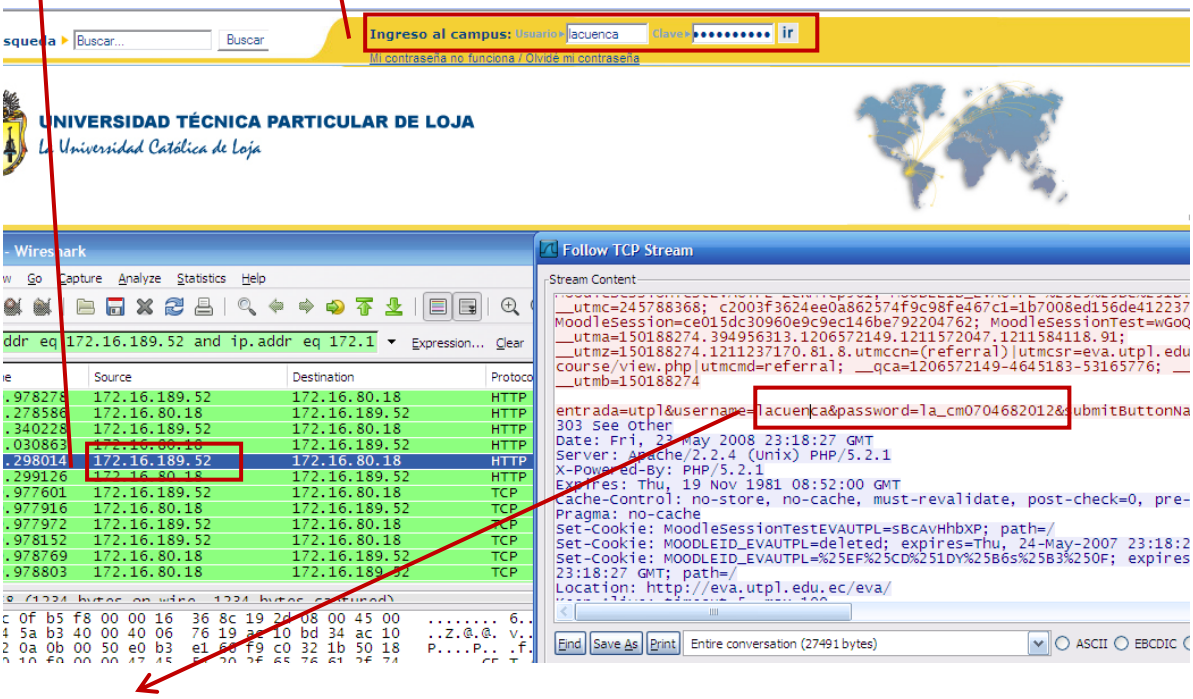
Cliente mail utilizado: **Outlook**

ANEXO B

Para el ingreso a otros servicios como en este caso al campus virtual, se necesita especificar un usuario y un password.

Nombre de usuario y password para ingreso al campus virtual.

Dirección del servidor.



Datos del usuario y password

Como se puede apreciar los datos del usuario y password, pueden ser capturados ya que estos no son encriptados y viajan en texto plano a través de la RED.

Herramientas utilizadas:

Herramienta utilizada para capturar los paquetes: **Wireshark**

Navegador: **Internet Explorer**

ANEXO C

Configuración actual del servidor mail

```
divert(-1)dnl
Dnl #
dnl # This is the sendmail macro config file for m4. If you make changes to
dnl # /etc/mail/sendmail.mc you will need to regenerate the
dnl # /etc/mail/sendmail.cf file by confirming that the sendmail-cf package is
dnl # installed and then performing a
dnl #     make -C /etc/mail
dnl #
include(`/usr/share/sendmail-cf/m4/cf.m4')dnl
VERSIONID(`mail UTPL')dnl
OSTYPE(`linux')dnl
dnl # default logging level is 9 you might want to set it higher to
dnl # debug the configuration
dnl #
dnl define(`confLOG_LEVEL', `9')dnl
dnl #
dnl #
define(`confDEF_USER_ID',``8:12'')dnl
dnl define(`confAUTO_REBUILD')dnl
define(`confTO_CONNECT', `1m')dnl
define(`confTRY_NULL_MX_LIST',true)dnl
define(`confDONT_PROBE_INTERFACES',true)dnl
define(`PROCMAIL_MAILER_PATH',`/usr/bin/procmail')dnl
define(`ALIAS_FILE', `/etc/aliases')dnl
define(`STATUS_FILE', `/var/log/mail/statistics')dnl
dnl define(`UUCP_MAILER_MAX', `2000000')dnl
define(`confUSERDB_SPEC', `/etc/mail/userdb.db')dnl
define(`confPRIVACY_FLAGS', `authwarnings,novrfy,noexpn,restrictqrun')dnl
define(`confAUTH_OPTIONS', `A')dnl
dnl #
dnl #
dnl # Rudimentary information on creating certificates for sendmail TLS:
dnl #     cd /usr/share/ssl/certs; make sendmail.pem
dnl # Complete usage:
dnl #     make -C /usr/share/ssl/certs usage
dnl #
dnl #define(`confCACERT_PATH',`/etc/openldap/cacerts/')
dnl #define(`confCACERT',`/etc/openldap/cacerts/slapd.crt')
dnl #define(`confSERVER_CERT',`/etc/openldap/cacerts/slapd.key')
dnl #define(`confSERVER_KEY',`/etc/openldap/cacerts/slapd.key')
dnl #
define(`confTO_IDENT', `0')dnl
dnl FEATURE(delay_checks)dnl
FEATURE(`no_default_msa',`dnl')dnl
FEATURE(`smrsh',`/usr/sbin/smrsh')dnl
FEATURE(`mailertable',`hash -o /etc/mail/mailertable.db')dnl
FEATURE(`virtusertable',`hash -o /etc/mail/virtusertable.db')dnl
FEATURE(redirect)dnl
dnl FEATURE(always_add_domain)dnl
FEATURE(use_cw_file)dnl
FEATURE(use_ct_file)dnl
dnl #
dnl # The -t option will retry delivery if e.g. the user runs over his quota.
dnl #
FEATURE(local_procmail,`',`procmail -t -Y -a $h -d $u')dnl
```

```
FEATURE(`access_db',`hash -T<TMPF> -o /etc/mail/access.db')dnl
FEATURE(`milter-greylist')dnl
#zepolar add
FEATURE(`greet_pause', `3000')
FEATURE(`blacklist_recipients')dnl
FEATURE(enhdnsbl,`dnsbl.ahbl.org', `"550 Host is on the AHBL - Please see
http://www.ahbl.org/tools/lookup.php?ip="$&{client_addr}"",,`127.0.0.2.',
`127.0.0.3.', `127.0.0.4.',`127.0.0.5.',`127.0.0.6.')dnl
EXPOSED_USER(`root')dnl
FEATURE(dnsbl, `sbl.spamhaus.org',`Spam see http://www.spamhaus.org/SBL/')dnl
define(`confBIND_OPTS', `WorkAroundBrokenAAAA')dnl
FEATURE(`dnsbl', `bl.spamcop.net', `"Spam blocked see:
http://spamcop.net/bl.shtml?"$&{client_addr}')dnl
dnl #
dnl # The following causes sendmail to additionally listen to port 465, but
dnl #starting immediately in TLS mode upon connecting. Port 25 or 587 followed
dnl # by STARTTLS is preferred, but roaming using Outlook Express can't
dnl # do STARTTLS on ports other than 25. Mozilla Mail can ONLY use STARTTLS
dnl # and doesn't support the deprecated smtps; Evolution <1.1.1 uses smtps
dnl # when SSL is enabled-- STARTTLS support is available in version 1.1.1.
dnl #
dnl # For this to work your OpenSSL certificates must be configured.
dnl #
dnl DAEMON_OPTIONS(`Port=smtps, Name=TLSMTA, M=s')dnl
dnl #
dnl #The following causes sendmail to additionally listen on the IPv6 loopback
dnl # device. Remove the loopback address restriction listen to the network.
dnl #
DAEMON_OPTIONS(`Name=MTA-v4, Family=inet, Name=MTA-v6, Family=inet6')
dnl #
dnl # We strongly recommend not accepting unresolvable domains if you want to
dnl # protect yourself from spam. However, the laptop and users on computers
dnl # that do not have 24x7 DNS do need this.
dnl #
dnl FEATURE(`accept_unresolvable_domains')dnl
dnl #
dnl FEATURE(`relay_based_on_MX')dnl
dnl # Also accept email sent to "localhost.localdomain" as local email.
dnl #
MASQUERADE_DOMAIN(utpl.edu.ec)dnl
dnl MASQUERADE_DOMAIN(localhost.localdomain)dnl
dnl MASQUERADE_DOMAIN(mydomainalias.com)dnl
dnl MASQUERADE_DOMAIN(mydomain.lan)dnl
INPUT_MAIL_FILTER(`clamav-milter', `S=local:/var/clamav/clmilter.socket,
F=,T=S:4m;R:4m;E:10m')
MAILER(smtp)dnl
MAILER(procmail)dnl
```

Configuración propuesta en el presente trabajo

```
divert(-1)dnl
dnl #
dnl # This is the sendmail macro config file for m4. If you make changes to
dnl # /etc/mail/sendmail.mc, you will need to regenerate the
dnl # /etc/mail/sendmail.cf file by confirming that the sendmail-cf package is
dnl # installed and then performing a
dnl #
dnl #     make -C /etc/mail
dnl #
include(`/usr/share/sendmail-cf/m4/cf.m4')dnl
VERSIONID(`setup for linux')dnl
OSTYPE(`linux')dnl
dnl #
dnl # Do not advertize sendmail version.
dnl #
dnl define(`confSMTP_LOGIN_MSG', `$j Sendmail; $b')dnl
dnl #
dnl # default logging level is 9, you might want to set it higher to
dnl # debug the configuration
dnl #
dnl define(`confLOG_LEVEL', `9')dnl
dnl #
dnl # Uncomment and edit the following line if your outgoing mail needs to
dnl # be sent out through an external mail server:
dnl #
dnl define(`SMART_HOST', `smtp.your.provider')dnl
dnl #
define(`confDEF_USER_ID', ``8:12'')dnl
dnl define(`confAUTO_REBUILD')dnl
define(`confTO_CONNECT', `1m')dnl
define(`confTRY_NULL_MX_LIST', `True')dnl
define(`confDONT_PROBE_INTERFACES', `True')dnl
define(`PROCMAIL_MAILER_PATH', `/usr/bin/procmail')dnl
define(`ALIAS_FILE', `/etc/aliases')dnl
define(`STATUS_FILE', `/var/log/mail/statistics')dnl
define(`UUCP_MAILER_MAX', `2000000')dnl
define(`confUSERDB_SPEC', `/etc/mail/userdb.db')dnl
define(`confPRIVACY_FLAGS', `authwarnings,novrfy,noexpn,restrictqrun')dnl
define(`confAUTH_OPTIONS', `A')dnl
dnl #
dnl # The following allows relaying if the user authenticates, and disallows
dnl # plaintext authentication (PLAIN/LOGIN) on non-TLS links
dnl #
dnl define(`confAUTH_OPTIONS', `A p')dnl
dnl #
dnl # PLAIN is the preferred plaintext authentication method and used by
dnl # Mozilla Mail and Evolution, though Outlook Express and other MUAs do
dnl # use LOGIN. Other mechanisms should be used if the connection is not
dnl # guaranteed secure.
dnl # Please remember that saslauthd needs to be running for AUTH.
dnl #
dnl TRUST_AUTH_MECH(`EXTERNAL DIGEST-MD5 CRAM-MD5 LOGIN PLAIN')dnl
dnl define(`confAUTH_MECHANISMS', `EXTERNAL GSSAPI DIGEST-MD5 CRAM-MD5 LOGIN
      PLAIN')dnl
dnl #
dnl # Rudimentary information on creating certificates for sendmail TLS:
dnl #     cd /etc/pki/tls/certs; make sendmail.pem
```

```
dnl # Complete usage:
dnl #     make -C /etc/pki/tls/certs usage
dnl #
dnl define(`confCACERT_PATH', `/etc/pki/tls/certs')dnl
dnl define(`confCACERT', `/etc/pki/tls/certs/ca-bundle.crt')dnl
dnl define(`confSERVER_CERT', `/etc/pki/tls/certs/sendmail.pem')dnl
dnl define(`confSERVER_KEY', `/etc/pki/tls/certs/sendmail.pem')dnl
dnl #
dnl # This allows sendmail to use a keyfile that is shared with OpenLDAP's
dnl # slapd, which requires the file to be readble by group ldap
dnl #
dnl define(`confDONT_BLAME_SENDMAIL', `groupreadablekeyfile')dnl
dnl #
dnl define(`confTO_QUEUEWARN', `4h')dnl
dnl define(`confTO_QUEUERETURN', `5d')dnl
dnl define(`confQUEUE_LA', `12')dnl
dnl define(`confREFUSE_LA', `18')dnl
define(`confTO_IDENT', `0')dnl
dnl FEATURE(delay_checks)dnl
FEATURE(`no_default_msa', `dnl')dnl
FEATURE(`smrsh', `/usr/sbin/smrsh')dnl
FEATURE(`mailertable', `hash -o /etc/mail/mailertable.db')dnl
FEATURE(`virtusertable', `hash -o /etc/mail/virtusertable.db')dnl
FEATURE(redirect)dnl
FEATURE(always_add_domain)dnl
FEATURE(use_cw_file)dnl
FEATURE(use_ct_file)dnl
dnl #
dnl # The following limits the number of processes sendmail can fork to accept
dnl # incoming messages or process its message queues to 20.) sendmail refuses
dnl # to accept connections once it has reached its quota of child processes.
dnl #
dnl define(`confMAX_DAEMON_CHILDREN', `20')dnl
dnl #
dnl # Limits the number of new connections per second. This caps the overhead
dnl # incurred due to forking new sendmail processes. May be useful against
dnl # DoS attacks or barrages of spam. (As mentioned below, a per-IP address
dnl # limit would be useful but is not available as an option at this writing.)
dnl #
dnl define(`confCONNECTION_RATE_THROTTLE', `3')dnl
dnl #
dnl # The -t option will retry delivery if e.g. the user runs over his quota.
dnl #
FEATURE(local_procmail, `', `procmail -t -Y -a $h -d $u')dnl
FEATURE(`access_db', `hash -T<TMPF> -o /etc/mail/access.db')dnl
FEATURE(`blacklist_recipients')dnl
EXPOSED_USER(`root')dnl
dnl #
dnl # For using Cyrus-IMAPd as POP3/IMAP server through LMTP delivery uncomment
dnl # the following 2 definitions and activate below in the MAILER section the
dnl # cyrusv2 mailer.
dnl #
dnl define(`confLOCAL_MAILER', `cyrusv2')dnl
dnl define(`CYRUSV2_MAILER_ARGS', `FILE /var/lib/imap/socket/lmtp')dnl
dnl #
dnl # The following causes sendmail to only listen on the IPv4 loopback address
dnl # 127.0.0.1 and not on any other network devices. Remove the loopback
dnl # address restriction to accept email from the internet or intranet.
dnl #
```

```
dnl # DAEMON_OPTIONS(`Port=smtp, Name=MTA')dnl
dnl #
dnl # The following causes sendmail to additionally listen to port 587 for
dnl # mail from MUAs that authenticate. Roaming users who can't reach their
dnl # preferred sendmail daemon due to port 25 being blocked or redirected find
dnl # this useful.
dnl #
dnl DAEMON_OPTIONS(`Port=submission, Name=MSA, M=Ea')dnl
dnl #
dnl # The following causes sendmail to additionally listen to port 465, but
dnl # starting immediately in TLS mode upon connecting. Port 25 or 587 followed
dnl # by STARTTLS is preferred, but roaming clients using Outlook Express can't
dnl # do STARTTLS on ports other than 25. Mozilla Mail can ONLY use STARTTLS
dnl # and doesn't support the deprecated smtps; Evolution <1.1.1 uses smtps
dnl # when SSL is enabled-- STARTTLS support is available in version 1.1.1.
dnl #
dnl # For this to work your OpenSSL certificates must be configured.
dnl #
dnl DAEMON_OPTIONS(`Port=smtps, Name=TLSMTA, M=s')dnl
dnl #
dnl # The following causes sendmail to additionally listen on the IPv6 loopback
dnl # device. Remove the loopback address restriction listen to the network.
dnl #
dnl DAEMON_OPTIONS(`port=smtp,Addr=::1, Name=MTA-v6, Family=inet6')dnl
dnl #
dnl # enable both ipv6 and ipv4 in sendmail:
dnl #
dnl #DAEMON_OPTIONS(`Name=MTA-v4, Family=inet, Name=MTA-v6, Family=inet6')
dnl #
dnl # We strongly recommend not accepting unresolvable domains if you want to
dnl # protect yourself from spam. However, the laptop and users on computers
dnl # that do not have 24x7 DNS do need this.
dnl #
FEATURE(`accept_unresolvable_domains')dnl
dnl #
dnl FEATURE(`relay_based_on_MX')dnl
dnl #
dnl # Also accept email sent to "localhost.localdomain" as local email.
dnl #
LOCAL_DOMAIN(`localhost.localdomain')dnl
dnl #
dnl # The following example makes mail from this host and any additional
dnl # specified domains appear to be sent from mydomain.com
dnl #
MASQUERADE_AS(`mail.utpl.com')dnl
dnl #
dnl # masquerade not just the headers, but the envelope as well
dnl #
dnl FEATURE(masquerade_envelope)dnl
dnl #
dnl # masquerade not just @mydomainalias.com, but @*.mydomainalias.com as well
dnl #
dnl FEATURE(masquerade_entire_domain)dnl
dnl #
MASQUERADE_DOMAIN(mail.utpl.com)dnl
dnl MASQUERADE_DOMAIN(localhost.localdomain)dnl
dnl MASQUERADE_DOMAIN(mydomainalias.com)dnl
dnl MASQUERADE_DOMAIN(mydomain.lan)dnl
MAILER(smtp)dnl
```

```
MAILER(procmail)dnl
dnl MAILER(cyrusv2)dnl
```

ANEXO D

Ejemplos de configuración con cada una de las propiedades que se tiene en GlusterFS

- **read-ahead**

```
volume readahead
  type performance/read-ahead
  option page-size 128kB
  option page-count 4
  option force-atime-update off
  subvolumes <nombre del volumen>
end-volume
```

- **write-behind**

```
volume writebehind
  type performance/write-behind
  option aggregate-size 1MB
  option flush-behind on
  subvolumes < nombre volumen >
end-volume
```

- **io-threads**

```
volume iothreads
  type performance/io-threads
  option thread-count 4     # deault es 1
  option cache-size 32MB
  subvolumes < nombre volumen >
 end-volume
```

- **io-cache.**

```
volume io-cache
  type performance/io-cache
  option cache-size 64MB    # default es 32MB
  option page-size 1MB
  option priority *.h:3,*.html:2,*:1
  option force-revalidate-timeout 2
  subvolumes < nombre volumen >
end-volume
```

- **stat-prefetch.**

```
volume stat-performance
  type performance/stat-prefetch
  option cache-seconds 1
  subvolumes < nombre volumen >
end-volume
```

- **booster**

```
volume booster
  type performance/booster
#option transport-type tcp
#Default es 'unix', which is mostly used when booster is loaded on client side.
  subvolumes < nombre volumen >
end-volume
```

- **afr.**

```
volume afr-example
  type cluster/afr
  subvolumes volumen1 volumen2 volumen3
end-volume
```

- **stripe.**

```
volume stripe
   type cluster/stripe
   option block-size *:1MB
   subvolumes volumen1 volumen2 volumen3
 end-volume
```

- **unify.**

```
volume unify
   type cluster/unify
   subvolumes volumen1 volumen2 volumen3
   option namespace volumen -ns
   # debe ser un nodo que no esté presente en un subvolumen
   option scheduler rr
   # simple round-robin scheduler
end-volume
```

- **trace.**

```
volume trace
  type debug/trace
  subvolumes < nombre volumen >
#option includes open, close, create, readdir, opendir, closedir
#  option excludes lookup, read, write
end-volume
```

- **filter.**

```
volume brick-readonly
  type features/filter
  subvolumes brick
end-volume
```

- **posix-locks.**

```
volume locks
  type features/posix-locks
  subvolumes < nombre volumen >
end-volume
```

- **trash.**

```
volume trash
  type features/trash
  option trash-dir /.trashcan
  subvolumes < nombre volumen >
end-volume
```

- **fixed-id.**

```
volume fixed
  type features/fixed-id
  option fixed-uid 1000
  option fixed-gid 100
  subvolumes < nombre volumen >
end-volume
```

- **posix.**

```
volume posix1
  type storage/posix      # POSIX FS translator
```

```
  option directory /home/export
  # Ruta del directorio a exportar
end-volume
```

- **client.**

```
volume client1
  type protocol/client
  option transport-type tcp/client
   # para transporte TCP/IP
   # option transport-type ib-sdp/client
   # para transporte Infiniband
   # option transport-type ib-verbs/client
   # para transporte Infiniband Verbs
   # option ib-verbs-work-request-recv-size   1048576
   # option ib-verbs-work-request-recv-count  16
   # option ib-verbs-work-request-send-size   1048576
   # option ib-verbs-work-request-send-count  16
   option remote-host 192.168.1.10
   # dirección IP del nodo remoto
   # option remote-port 6996
   # default server port is 6996
   # option transport-timeout 30
   # segundos de respuesta del servidor
   option remote-subvolume < nombre volumen >
end-volume
```

- **server.**

```
volume server
  type protocol/server
  option transport-type tcp/server
  # Para transporte TCP/IP
  # option transport-type ib-sdp/server
  # Para transporte Infiniband
  # option transport-type ib-verbs/server
  # Para transporte  Infiniband Verbs
  #option ib-verbs-work-request-recv-size 1048576
  # option ib-verbs-work-request-recv-count  16
  # option ib-verbs-work-request-send-size   1048576
  # option ib-verbs-work-request-send-count  16
```

```
  # option bind-address 192.168.1.10
  # Por default escucha en todas las interfaces
  # option listen-port 6996      # Por default es 6996
  # option client-volume-filename /etc/glusterfs/glusterfs-client.vol
  subvolumes volumen1 volumen2
  option auth.ip.brick1.allow 192.168.*
  # Permite el acceso al volumen "brick1"
  option auth.ip.brick2.allow 192.168.*
  # Permite el acceso al volumen "brick2"
end-volume
```

- **rot-13.**

```
volume rot-13
  type encryption/rot-13
  encrypt-write [on|off] (on)
  decrypt-read [on|off] (on)
  subvolumes volumen1
end-volume
```

ANEXO E

Instalación de paquetes GlusterFS para el servidor.

1. Ingresamos al sistema como root

   [root@glusterfs ~]#su root

2. Instalamos cada uno de los paquetes en el mismo orden en que fueron listados anteriormente para ello se utiliza el comando **rpm –ivh [nombre del paquete]**

   [root@nodo0 ~]#rpm -ivh glusterfs-core-3.2.2-1.x86_64.rpm

   [root@nodo0 ~]#rpm -ivh glusterfs-fuse-3.2.2-1.x86_64.rpm

ANEXO F

Archivos de configuración de GlusterFS para el servidor.

Nombre de archivo: glusterfs-server1.vol
Ubicación: /etc/glusterfs
Servidor: Mail

Contenido:

```
volume spool
  type storage/posix
  option directory /var/spool/mail
end-volume

volume posix-locks
  type features/posix-locks
  option mandatory on
  subvolumes spool
end-volume

volume io-thr
 type performance/io-threads
 subvolumes posix-locks
end-volume

volume wb
 type performance/write-behind
 subvolumes io-thr
end-volume

volume raae
 type performance/read-ahead
 subvolumes wb
end-volume

volume ra
 type encryption/rot-13
 subvolumes raae
end-volume

volume server
 type protocol/server
 subvolumes ra
 option transport-type tcp/server
 option client-volume-filename /etc/glusterfs/glusterfs-client.vol
 option auth.ip.ra.allow *
end-volume
```

Nombre de archivo: glusterfs-server2.vol
Ubicación: /etc/glusterfs
Servidor: Mail

Contenido:

```
volume casa
  type storage/posix
  option directory /public
end-volume

volume posix-locks
  type features/posix-locks
  option mandatory on
  subvolumes casa
end-volume

volume io-thr
 type performance/io-threads
 subvolumes posix-locks
end-volume

volume wb
 type performance/write-behind
 subvolumes io-thr
end-volume

volume racasaae
 type performance/read-ahead
 subvolumes wb
end-volume

volume racasa
 type encryption/rot-13
 subvolumes racasaae
end-volume

volume server
 type protocol/server
 subvolumes racasa
 option transport-type tcp/server
 option client-volume-filename /etc/glusterfs/glusterfs-client.vol
 option auth.ip.racasa.allow *
end-volume
```

Nombre de archivo: glusterfs-server3.vol
Ubicación: /etc/glusterfs
Servidor: Mail

Contenido:

```
volume confmail
  type storage/posix
  option directory /etc/mail
end-volume

volume posix-locks
  type features/posix-locks
  option mandatory on
  subvolumes confmail
end-volume

volume io-thr
 type performance/io-threads
 subvolumes posix-locks
end-volume

volume wb
 type performance/write-behind
 subvolumes io-thr
end-volume

volume raconfae
 type performance/read-ahead
 subvolumes wb
end-volume

volume raconf
 type encryption/rot-13
 subvolumes raconfae
end-volume

volume server
 type protocol/server
 subvolumes raconf
 option transport-type tcp/server
 option client-volume-filename /etc/glusterfs/glusterfs-client.vol
 option auth.ip.raconf.allow *
end-volume
```

Nombre de archivo: glusterfs-server4.vol
Ubicación: /etc/glusterfs
Servidor: Mail

Contenido:

```
volume user
  type storage/posix
  option directory /etc/passwd
end-volume

volume posix-locks
  type features/posix-locks
  option mandatory on
  subvolumes user
end-volume

volume io-thr
 type performance/io-threads
 subvolumes posix-locks
end-volume

volume wb
 type performance/write-behind
 subvolumes io-thr
end-volume

volume rauserae
 type performance/read-ahead
 subvolumes wb
end-volume

volume rauser
 type encryption/rot-13
 subvolumes rauserae
end-volume

volume server
 type protocol/server
 subvolumes rauser
 option transport-type tcp/server
 option client-volume-filename /etc/glusterfs/glusterfs-client.vol
 option auth.ip.rauser.allow *
end-volume
```

Nombre de archivo: glusterfs-server.vol
Ubicación: /etc/glusterfs
Servidor: Web

Contenido:

```
volume paginas
  type storage/posix
  option directory /mnt/disk2
end-volume

volume posix-locks
  type features/posix-locks
  option mandatory on
  subvolumes paginas
end-volume

volume io-thr
 type performance/io-threads
 subvolumes posix-locks
end-volume

volume wb
 type performance/write-behind
 subvolumes io-thr
end-volume

volume rapaginasae
 type performance/read-ahead
 subvolumes wb
end-volume

volume rapaginas
 type encryption/rot-13
 subvolumes rapaginasae
end-volume

volume server
 type protocol/server
 subvolumes rapaginas
 option transport-type tcp/server
 option client-volume-filename /etc/glusterfs/glusterfs-client.vol
 option auth.ip.rapaginas.allow *
end-volume
```

ANEXO G

- **Instalación del módulo fuse**

Para instalar el módulo de fuse al sistema se realizan los siguientes pasos:

1. Ingresamos al sistema como root

   [root@nodo1~]#su root

2. Ejecutamos el siguiente comando.

3. [root@nodo1~]# sudo yum -y install openssh-server wget fuse fuse-libs openib libibverbs

Con esto se finaliza la instalación del módulo fuse para el equipo cliente.

- **Instalación de paquetes GlusterFS para el cliente.**

Una vez obtenidos los paquetes se procede a la instalación y para ello se realiza los siguientes pasos:

1. Ingresamos al sistema como root

   [root@nodo1~]#su root

2. Instalamos cada uno de los paquetes utilizando el comando rpm –ivh [nombre del paquete]

   [root@nodo1~]#rpm -ivh glusterfs-core-3.2.2-1.x86_64.rpm

   [root@nodo1~]#rpm -ivh glusterfs-fuse-3.2.2-1.x86_64.rpm

ANEXO H

Archivos de configuración de GlusterFS para los clientes.

Nombre de archivo: glusterfs-client1.vol
Ubicación: /usr/local/etc/glusterfs
Servidor: Mail

Contenido:

```
volume client1
  type protocol/client
  option transport-type tcp/client
  option remote-host 172.16.17.55
  option remote-subvolume ra
end-volume

volume stripe1
  type cluster/stripe
  subvolumes client1
  option block-size *:10KB
end-volume

volume iot
  type performance/io-threads
  subvolumes stripe1
  option thread-count 8
end-volume

volume wb
  type performance/write-behind
  subvolumes iot
end-volume

volume ra
  type performance/read-ahead
  subvolumes wb
end-volume

volume ioc
 type performance/io-cache
  subvolumes ra
end-volume

volume iocen
 type encryption/rot-13
 subvolumes ioc
end-volume
```

Nombre de archivo: glusterfs-client2.vol
Ubicación: /usr/local/etc/glusterfs
Servidor: Mail

Contenido:

```
volume client1
  type protocol/client
  option transport-type tcp/client
  option remote-host 172.16.17.55
  option remote-subvolume racasa
end-volume

volume stripe1
  type cluster/stripe
  subvolumes client1
  option block-size *:10KB
end-volume

volume iot
  type performance/io-threads
  subvolumes stripe1
  option thread-count 8
end-volume

volume wb
  type performance/write-behind
  subvolumes iot
end-volume

volume ra
  type performance/read-ahead
  subvolumes wb
end-volume

volume ioccasa
 type performance/io-cache
  subvolumes ra
end-volume

volume ioccasaen
 type encryption/rot-13
 subvolumes ioccasa
end-volume
```

Nombre de archivo: glusterfs-client3.vol
Ubicación: /usr/local/etc/glusterfs
Servidor: Mail

Contenido:

```
volume client1
  type protocol/client
  option transport-type tcp/client
  option remote-host 172.16.17.55
  option remote-subvolume raconf
end-volume

volume stripe1
  type cluster/stripe
  subvolumes client1
  option block-size *:10KB
end-volume

volume iot
  type performance/io-threads
  subvolumes stripe1
  option thread-count 8
end-volume

volume wb
  type performance/write-behind
  subvolumes iot
end-volume

volume ra
  type performance/read-ahead
  subvolumes wb
end-volume

volume iocconf
 type performance/io-cache
  subvolumes ra
end-volume

volume iocconfen
 type encryption/rot-13
 subvolumes iocconf
end-volume
```

Nombre de archivo: glusterfs-client4.vol
Ubicación: /usr/local/etc/glusterfs
Servidor: Mail

Contenido:

```
volume client1
  type protocol/client
  option transport-type tcp/client
  option remote-host 172.16.17.55
  option remote-subvolume rauser
end-volume

volume stripe1
  type cluster/stripe
  subvolumes client1
  option block-size *:10KB
end-volume

volume iot
  type performance/io-threads
  subvolumes stripe1
  option thread-count 8
end-volume

volume wb
  type performance/write-behind
  subvolumes iot
end-volume

volume ra
  type performance/read-ahead
  subvolumes wb
end-volume

volume iocuser
 type performance/io-cache
  subvolumes ra
end-volume

volume iocuseren
 type encryption/rot-13
 subvolumes iocuser
end-volume
```

Nombre de archivo: glusterfs-client.vol
Ubicación: /usr/local/etc/glusterfs
Servidor: Web

Contenido:

```
volume client1
  type protocol/client
  option transport-type tcp/client
  option remote-host 172.16.17.55
  option remote-subvolume rapaginas
end-volume

volume stripe1
  type cluster/stripe
  subvolumes client1
  option block-size *:10KB
end-volume

volume iot
  type performance/io-threads
  subvolumes stripe1
  option thread-count 8
end-volume

volume wb
  type performance/write-behind
  subvolumes iot
end-volume

volume ra
  type performance/read-ahead
  subvolumes wb
end-volume

volume iocpaginas
 type performance/io-cache
  subvolumes ra
end-volume

volume iocpaginasen
 type encryption/rot-13
 subvolumes iocpaginas
end-volume
```

ANEXO I

1. Actualización de dovecot 0.98 a dovecot 1.0.7

Para actualizar dovecot se siguen los siguientes pasos:

1. Obtenemos el paquete dovecot-1.0.7-7.el5.x86_64.rpm de la siguiente dirección:

   http://atrpms.net/name/dovecot/

2. Se utiliza el siguiente comando para actualizar dovecot.

   [root@glusterfs ~]#rpm –Uvh dovecot-1.0.7-7.el5.x86_64.rpm

2. Configuración de las nuevas características de dovecot.

En la nueva versión de dovecot se deben configurar las siguientes características:

mmap_disable= yes

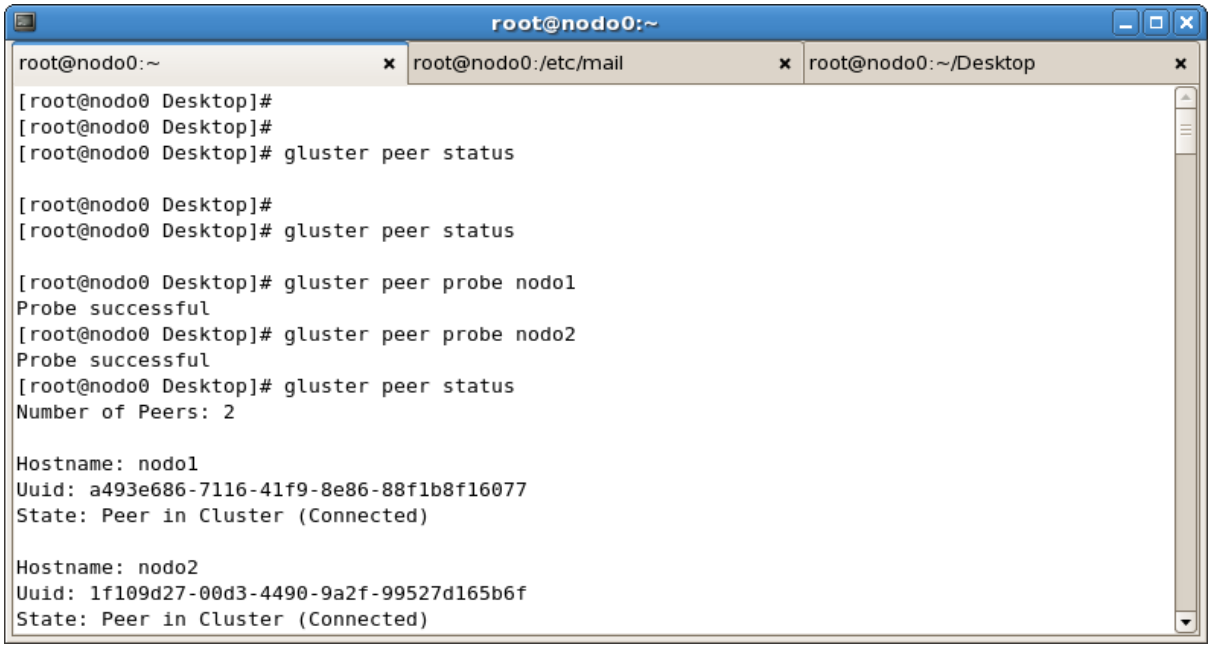
dotlock_use_excl=yes

ANEXO J

Configuración de grupos de almacenamiento de confianza.

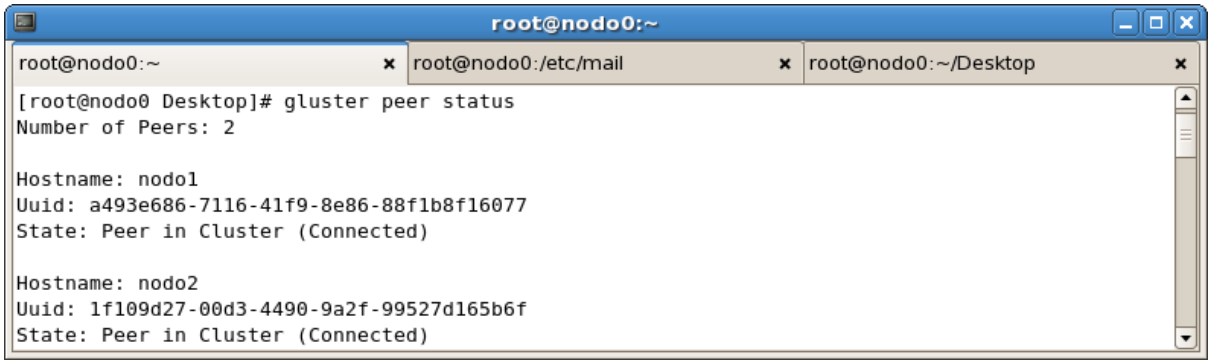
Para la configuración de los grupos de almacenamiento de confianza se hace uso del comando peer, el mismo que puede ser usado desde el shell de linux o desde la consola de administración de GlusterFS.

En la siguiente imagen se puede observar los pasos a seguir para realizar esta configuración.



Para verificar que se configuro correctamente se ejecuta el siguiente comando.

gluster peer status



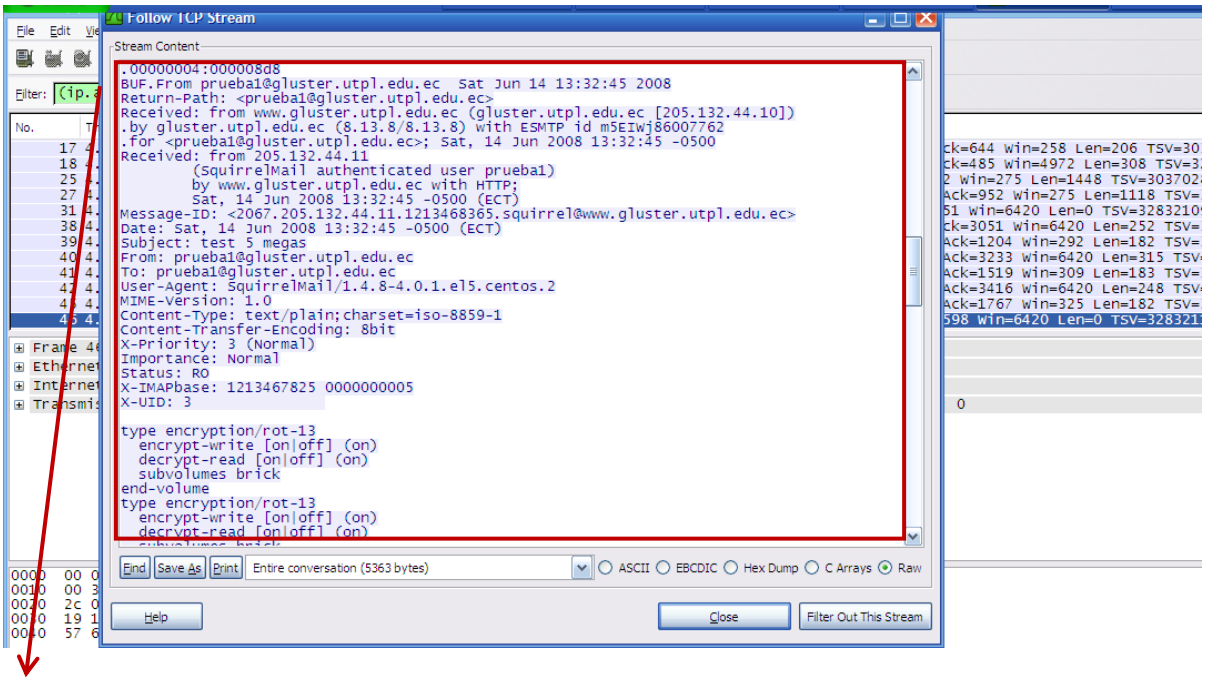
ANEXO K

A continuación se listan las reglas que fueron aplicadas:

```
iptables -X
iptables -F
iptables -P INPUT DROP
iptables -P OUTPUT DROP
iptables -P FORWARD DROP
iptables -A INPUT -i lo -j ACCEPT
iptables -A OUTPUT -o lo -j ACCEPT
iptables -A INPUT -i eth0 -j ACCEPT
iptables -A OUTPUT -o eth0 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s 172.16.17.0/24 -d 172.16.17.55 --dport 6996 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s 172.16.17.0/24 -d 172.16.17.55 --dport 6996 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 -d any/0 --sport 6996 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s any/0 -d 172.16.17.55 --dport 25 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 --sport 25 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s any/0 -d 172.16.17.55 --dport 80 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 --sport 80 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s any/0 -d 172.16.17.55 --dport 22 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 --sport 22 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s any/0 -d 172.16.17.55 --dport 53 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 --sport 53 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p udp -s any/0 -d 172.16.17.55 --dport 53 -j ACCEPT
iptables -A OUTPUT -o eth0 -p udp -s 172.16.17.55 --sport 53 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s any/0 -d 172.16.17.55 --dport 111 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 --sport 111 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s any/0 -d 172.16.17.55 --dport 143 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 --sport 143 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s any/0 -d 172.16.17.55 --dport 110 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 --sport 110 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s any/0 -d 172.16.17.55 --dport 993 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 --sport 993 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s any/0 -d 172.16.17.55 --dport 995 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 --sport 995 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p tcp -s any/0 -d 172.16.17.55 --dport 443 -j ACCEPT
iptables -A OUTPUT -o eth0 -p tcp -s 172.16.17.55 --sport 443 -d any/0 -j ACCEPT
iptables -A INPUT -i eth0 -p udp -s any/0 -d 172.16.17.55 --dport 111 -j ACCEPT
iptables -A OUTPUT -o eth0 -p udp -s 172.16.17.55 --sport 111 -d any/0 -j ACCEPT
```

ANEXO L

Captura de la información que pasa a través del clúster con el esquema actual.



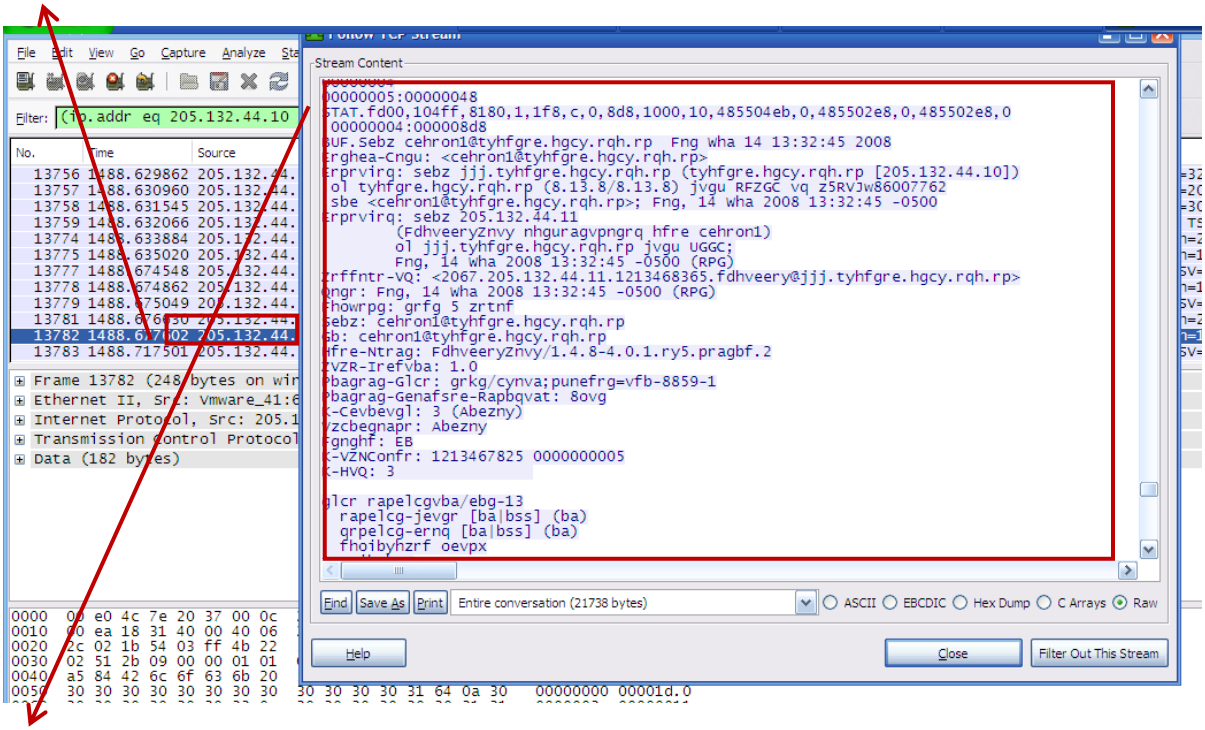
Como se puede apreciar la información que se está transmitiendo con el esquema del clúster sin las nuevas implementaciones puede ser interceptada y legible por intermedio de un sniffer.

Herramientas utilizadas:

Herramienta utilizada para capturar los paquetes: **EtherDetect Packet Sniffer**

Captura de la información que pasa a través del clúster que tiene instalado GlusterFS

Dirección del servidor que tiene implementado GlusterFS



Como se puede ver la información que está pasando a través del clúster esta codificada no es posible verla en texto plano, comprobando de esta manera que la información está viajando de una forma segura y confiable.

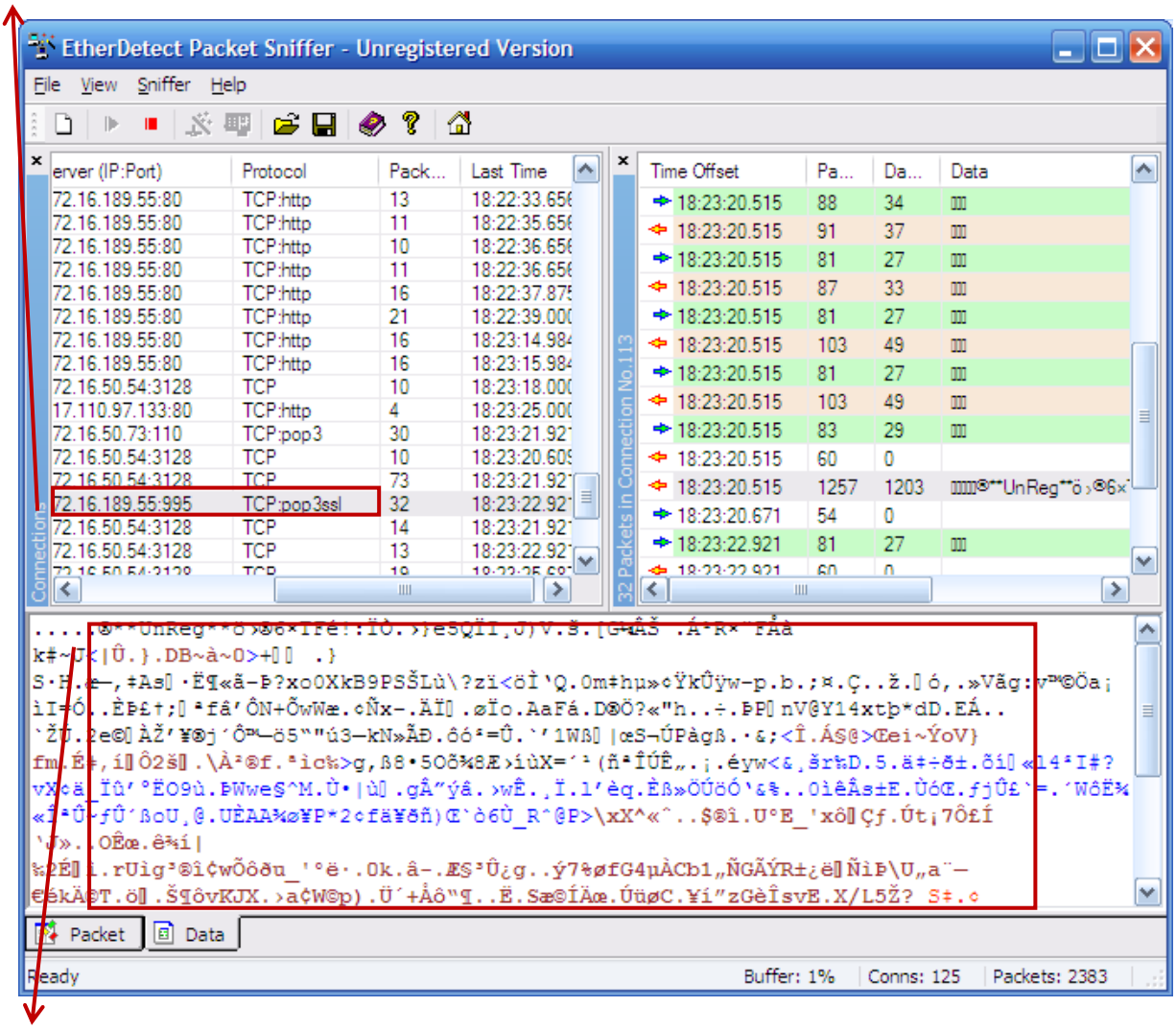
Herramientas utilizadas:

Herramienta utilizada para capturar los paquetes: **EtherDetect Packet Sniffer**

Navegador: **Internet Explorer**

Captura de paquetes mail utilizando protocolos seguros

Dirección del servidor con la implementación de un protocolo seguro.



Cuerpo del mensaje

Como se puede apreciar no es posible visualizar información legible ya que esta está siendo encriptado.

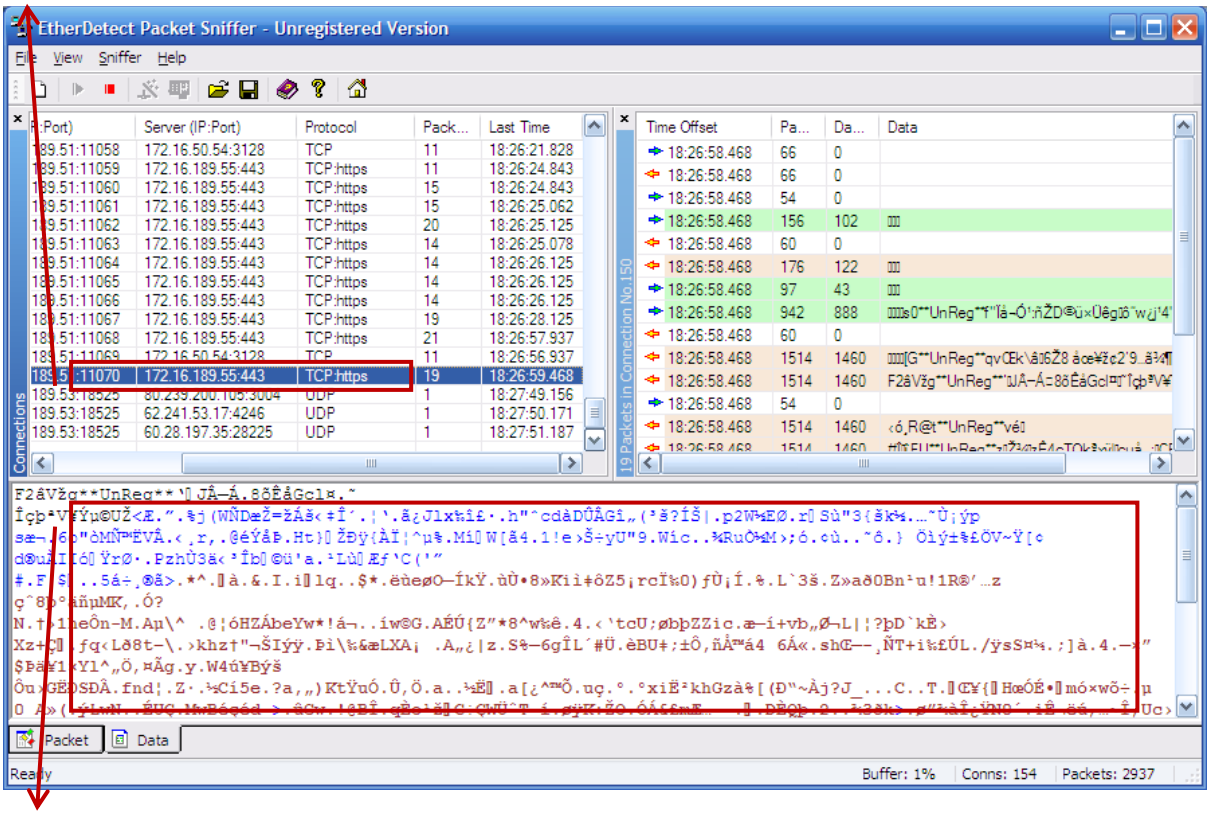
Herramienta utilizada para capturar los paquetes: **EtherDetect Packet Sniffer**

Cliente web: **Outlook**

Navegador: **Internet Explorer**

Captura utilizando protocolos seguros desde Webmail

Dirección del servidor



Cuerpo del mensaje

Herramienta utilizada para capturar los paquetes: **EtherDetect Packet Sniffer**

Cliente web: **Webmail**

Navegador: **Internet Explorer**

ANEXO M

Archivo de configuración del dovecot actual.

En el archivo de configuración se debe modificar la entrada protocols para ahí configurar los protocolos seguros imaps y pop3s

```
## Dovecot 1.0 configuration file

# Default values are shown after each value, it's not required to uncomment # any
of the lines. Exception to this are paths, they're just examples # with real
defaults being based on configure options. The paths listed here # are for
configure --prefix=/usr --sysconfdir=/etc --localstatedir=/var # --with-
ssldir=/usr/share/ssl

# Base directory where to store runtime data.
#base_dir = /var/run/dovecot/

# Protocols we want to be serving:
#  imap imaps pop3 pop3s
protocols = imap pop3 imaps pop3s
```

ANEXO N

Instalación de cyrus-sasl-2.1.22.tar.gz

Con la instalación de este paquete se implementa la autenticación SMTP, los pasos que se deben seguir son los que se listan a continuación:

Pasos para la instalación:

1. Se obtiene el paquete cyrus-sasl-2.1.22.tar.gz
2. Se descomprime el paquete utilizando el siguiente comando:

   [root@glusterfs ~]#tar -xzvf cyrus-sasl-2.1.22.tar.gz
3. Ingresamos al directorio

   [root@glusterfs ~]#cd cyrus-sasl-2.1.22

   [root@cyrus-sasl-2.1.22~]#
4. Ejecutamos ./configure

   [root@cyrus-sasl-2.1.22~]#./configure
5. Ejecutamos make

   [root@cyrus-sasl-2.1.22~]#make
6. Ejecutamos make install

   [root@cyrus-sasl-2.1.22~]#make install

ANEXO 0

Paper SEGURIDAD EN CLÚSTER DE SERVICIOS DE MAIL Y WEB

SEGURIDAD EN CLÚSTER DE SERVICIOS DE MAIL Y WEB

Luis Cuenca[1], Samanta Cueva[2]

RESUMEN

La seguridad es un término general que interviene en áreas como: la computación y en el procesamiento de la información. En la actualidad las organizaciones dependen a diario de los sistemas computacionales y de las redes para ejecutar sus operaciones y transacciones corporativas, considerando a los datos como un recurso importante dentro de la organización.

Para brindar servicios confiables a los usuarios se hace uso de clústeres para el manejo de la información, pero se debe tener en cuenta que la información que maneja el clúster debe mantener la integridad, asegurando la confiabilidad de datos.

En el presente documento se detalla cada una de las configuraciones a usar en los servicios de web y mail para asegurar la información de los servicios, así como también se detalla la selección de un sistema de replicación de archivos para asegurar la información que fluye a través del clúster de servicios web y mail.

Palabras Clave — Clúster, Seguridad, Protocolos Seguros, SSL, Encriptación, NFS, GlusterFS, POP3S, IMAPS, HTTPS.

1. INTRODUCCIÓN

En la UTPL cuenta con un clúster de balanceo de carga y alta disponibilidad para los servicios de Web y Mail, el cual brinda los servicios de:

- Distribución de carga de trabajo a los servidores reales.
- Alta disponibilidad, ya que en el caso de que un servidor falle se levante el servidor de backup.
- Brindar alto rendimiento, mediante el servicio ininterrumpido.
- Aprovechar al máximo los recursos existentes.
- Escalabilidad, permitiendo el incremento del número de nodos sin interrumpir el normal funcionamiento del servicio.

El clúster está conformado por un servidor activo que actúa como balanceador de carga y dos servidores reales.

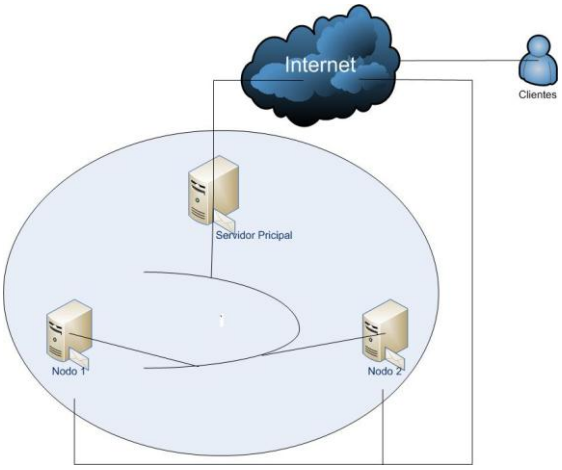


FIGURA I
DISEÑO FÍSICO DEL CLÚSTER

El clúster cuenta con un sistema de compartición de recursos NFS, con esto las peticiones de los clientes pueden ser atendidas por cualquiera de los servidores reales y estos podrán ser modificados en cualquiera de los recursos que se encuentran compartidos.

NFS, no posee un buen sistema de seguridad ya que esta tarea la deja al cliente.

NFS exporta los directorios hacia determinados dominios, lo que permitiría que un atacante que tenga el control del DNS pueda fácilmente acceder a toda la información que se encuentra compartida.

La sintaxis para escribir la configuración de NFS no es estricta, dejando la posibilidad de que un espacio mal ubicado en el archivo de configuración permita que tal información pueda ser compartida con todo mundo.

Para la solución al problema que se plantea se propone las siguientes soluciones:

- Implementar los servicios de correo electrónico y web con protocolos seguros.
- Implementar nuevas directivas de seguridad en las configuraciones del servidor mail.
- Actualizar Dovecot que es un servicio necesario para que funcione el correo electrónico.
- Cambiar el sistema de replicación de archivos que actualmente es NFS por GlusterFS.

[1] UTPL, Loja - Ecuador, lacuenca@utpl.edu.ec, (Tesista)
[2] UTPL, Loja - Ecuador, spcueva@utpl.edu.ec, (Directora de Tesis)

El presente trabajo tiene por objetivo general:

Asegurar y proteger la información que se transmite en el clúster de servicios de mail y web.

Para lo cual se deben:

1) Utilizar mecanismos de Encriptación de información para proteger los datos que se transmiten por el clúster. 2) Definir las tecnologías de compartición. 3) Implementar Seguridad en el clúster.

Para asegurar el clúster se realizó la implementación de GlusterFS como sistema para replicación de archivos, configuración de protocolos seguros para los servicios de web y mail, configuraciones avanzadas para los servicios de web y mail.

2. PROTOCOLOS DEL SERVICIO WEB Y MAIL

Los protocolos que se utilizan en el servicio web y mail son:

- POP3
- IMAP
- HTTP

Estos protocolos están expuestos con las configuraciones por defecto, lo que se pretende ahora es asegurar estos protocolos y no permitir que estos sean interceptados por personas no autorizadas.

2.1. WEB

HTTP: "Es un protocolo de transferencia de hipertexto basado fundamentalmente en el lenguaje HTML" (Salavert Casamor, 2003)

Para habilitar la seguridad en el protocolo HTTP se debe configurar el módulo de seguridad SSL, esto se lo realiza en el archivo de configuración que se encuentra en: /etc/httpd/conf/httpd.conf

2.2. MAIL

POP3: "(Post Office Protocol, Version 3) es un protocolo que funciona a nivel de aplicación según el modelo de referencia OSI y sus especificaciones se describen en la RCF 1939" (Salavert Casamor, 2003)

Este protocolo se configura en el servidor de correo para que los usuarios obtengan los mensajes de correo almacenados en un servidor, a través de un cliente para correo electrónico como:

Outlook (Windows)
Evolution (Linux)
Mail (MAC)

IMAP: "El Protocolo de Acceso a Mensajes de Internet (Internet Message Access Protocol o IMAP) permite a los clientes de correo acceder a mensajes guardados remotamente" (Jorquera, 2008) Mediante IMAP se puede tener acceso al correo electrónico desde cualquier equipo que tenga una conexión a Internet.

Para habilitar la seguridad en los protocolos POP3 e IMAP se procede a configurar la opción de **pop3s e imaps** en el archivo de configuración /etc/dovecot.conf.

Una vez que se realiza estas configuraciones en el clúster, se cuenta con los siguientes protocolos seguros:

- POP3S
- IMAPS
- HTTPS

Con protocolos la información pasa por un canal seguro, incrementando la seguridad y también se evita que la información pueda ser interceptada por terceras personas.

2.3. DIRECTIVAS DE CONFIGURACIÓN DEL MAIL

Debido a la naturaleza del correo electrónico, un atacante puede inundar fácilmente el servidor con correos basura conocido como **SPAM**[3], de esta manera causar un rechazo de servicio.

Se puede limitar la efectividad de estos ataques mediante la especificación de límites a las siguientes directivas de configuración de mail en **/etc/mail/sendmail.mc**

Directiva	Descripción
confCONNECTION_RATE_THROTTLE	El número de conexiones que el servidor puede recibir por segundo. Si se establece un límite y es alcanzado, las conexiones siguientes son retrasadas.
confMAX_DAEMON_CHILDREN	El máximo número de procesos hijo que se pueden producir por el servidor. Si se coloca un límite y es alcanzado, las conexiones siguientes son retrasadas.
confMIN_FREE_BLOCKS	El número mínimo de bloques libres que debe haber disponible para que el servidor acepte correos. Por defecto es 100 bloques.
confMAX_HEADERS_LENGTH	El tamaño máximo aceptable (en bytes) para la cabecera de un mensaje.
confMAX_MESSAGE_SIZE	El tamaño máximo aceptable (en bytes) para cualquier mensaje.

TABLA 1
DIRECTIVAS DEL SERVICIO DE MAIL

[3] S***pam***, mensajes no solicitados, de tipo publicitario, enviados en grandes cantidades que perjudican de alguna manera al receptor.

3. MODELADO DE AMENAZAS

Este modelado es útil para la evaluación de la seguridad o como parte de pruebas de penetración. Tras el diseño de varios modelos de amenaza, se verán patrones recurrentes.

El mantener modelos de amenaza realizados anteriormente es una excelente manera de documentar la evolución del sistema y de esta manera conservar un historial.

Así los modelos existentes se pueden utilizar como puntos de partida para el desarrollo de nuevos modelos de amenazas, evitando realizar un nuevo diseño.

A continuación se presenta una lista de los ataques típicos que puede tener un servidor web.

Ataque	Descripción	Mitigación
Denegación de servicio	Cualquier red, servidor web, o aplicación, se convierte en una base de ataques que resultan en denegación de servicio, condición de un sistema sobrecargado y que no puede responder.	• Prepararse para los ataques. • Realizar una inspección de la aplicación para eliminar los puntos de ataque.
Explotación de los errores de configuración	Estos son errores del administrador del servidor.	• Realizar una instalación segura desde el inicio de la implementación • Realizar un plan de cambios. • Evaluar el impacto de los cambios. • Evaluación de la configuración de forma regular.
Explotación de las vulnerabilidades de Apache	Los parches que son desconocidos, problemas que hay en el servidor web Apache.	• Aplicar parches inmediatamente, Actualización de Apache.
Explotación de vulnerabilidades de aplicación	Los parches que no son conocidos o problemas descubiertos en aplicaciones web, puestas en producción.	• Evaluar la seguridad de las aplicaciones web antes que estén en producción.
Los ataques a través de otros servicios	Captura de los problemas en el servidor web que no han sido mitigados.	• No exponer los servicios que no son necesarios.

TABLA 2
ATAQUES TÍPICOS DE UN SERVICIO

Como se ve, para cada uno de los ataques se tienen técnicas de mitigación, es recomienda hacer uso de otros procedimientos de mitigación como:

- Aplicar monitoreo y considerar la aplicación de detección de intrusos.
- Disponer de procedimientos de recuperación ante un eventual ataque.
- Realizar periódicamente copias de seguridad y almacenarlas fuera, con esto se tienen los datos que se necesitan para los procedimientos de recuperación.

Ahora surge una interrogante para decidir cuál de los posibles métodos de protección es recomendable usar en la instalación inicial del servicio.

Para ello se hace uso de una matriz de decisión para proteger el servidor web, en primer lugar se listan todos los posibles métodos de protección y se los clasifica cada uno en términos de complejidad.

Para esta matriz se utiliza una clasificación que consta de cuatro categorías que son:

Categoría 1. Criticas (muy importantes)
Categoría 2. Producción.
Categoría 3. Desarrollo.
Categoría 4. Pruebas (poco importantes)

Técnicas	Categorías			
	4	3	2	1
Agregar características a Apache desde la compilación del código fuente			X	X
Ajustar la configuración (remover módulos por defecto, restringir ciertos módulos, etc)			X	X
Cambiar la identidad del servidor web.			X	X
Implementar autenticación.			X	X
Implementar SSL			X	X
Desplegar certificados desde un CA reconocido.			X	X
Centralizar logs	X	X	X	X
Usar mod_security			X	X
Realizar monitoreo al servidor		X	X	X
Realizar monitoreo externo de la disponibilidad			X	X
Monitoreo en tiempo real de logs.				X
Realizar un análisis periódico de logs.			X	X
Realizar correlación de eventos.				X
Implementar reglas de firewall		X	X	X
Validar la integridad de los archivos			X	X
Evaluación de vulnerabilidades externas.				X
Separar los componentes de aplicación.				X

TABLA 3
MATRIZ DE DECISIÓN

Se debe hacer uso de esta matriz en el momento que se detecte algún problema en el servidor web, de tal manera que el administrador tome la mejor decisión al momento de aplicar la solución a un problema encontrado.

Una vez que se ha realizado la evaluación con la matriz se debe proseguir con las acciones pertinentes para darles una solución, para ello se propone seguir el siguiente plan dependiendo de la categoría en la que se encuentre:

Categoría 1

Cuando se encuentre en esta categoría, las acciones a tomar deben ser aplicadas de manera inmediata en el servidor.

Categoría 2

En esta categoría se puede tomar un tiempo prudente para realizar las correcciones necesarias en el servidor.

Las categorías 3 y 4

Las acciones a tomar cuando se encuentre en estas dos categorías no son de gran influencia en el servidor, ya que se tratan de actividades de pruebas o desarrollo.

Un aspecto importante en el uso de esta matriz es que se pueden agregar o quitar ciertas técnicas, ya que tendrán su importancia de acuerdo al tipo de servicio que se brinde.

Se deben tomar en cuenta algunas características del servicio al momento de elaborar una matriz de decisión como: Fiabilidad, Confiabilidad, Disponibilidad y Tipo de servicio al que está orientado.

4. MÉTODO

4.1. ASEGURAMIENTO DE LA GESTIÓN DE INFORMACIÓN DEL CLÚSTER

NFS como tal no fue diseñado para trabajar en un ambiente de clúster, además de esto presenta muchas vulnerabilidades que pueden ser aprovechadas por personas malintencionadas, no posee un mecanismo de autenticación y encriptación de los datos que se comparten, ya que NFS lo realiza de una manera plana, es por ello que se decidió cambiar NFS por otro sistema de replicación de archivos.

A continuación se presenta un cuadro comparativo con cada una de las características de los tres sistemas de replicación de archivos analizados, esto permite tener un mejor criterio para seleccionar uno de los archivos, que se utilizará en el clúster de servicios de mail y web.

Se tomaron en cuenta características relevantes de cada uno de los sistemas de replicación:

No	Detalle	Valor Ponderado
1	Replicación de archivos a varios equipos.	4
2	Encriptación de la información.	4
3	No posee punto único de montaje.	4
4	Soporte bloqueo de archivos.	4
5	Punto único de fallo.	3
6	Soporte (Kernel).	3
7	Balanceo de Carga.	3
8	Escalabilidad para transferir grandes cantidades de datos.	3
9	Diseñado para ambientes clúster.	2
10	Trabajo en modo desconectado.	2
11	Caching local.	2
12	No requiere de un servidor exclusivo para la implementación.	2
13	Licencia GPL.	2
14	Control de cuotas de disco	2

TABLA 4
CARACTERÍSTICAS CUANTIFICADAS

Características	Sistemas de replicación de archivos		
	GlusterFS	CODA	OpenGFS
1	4	4	4
2	4	4	0
3	4	0	4
4	4	0	4
5	3	3	3
6	3	3	0
7	3	0	0
8	3	0	0
9	2	0	0
10	0	2	0
11	0	2	2
12	2	0	2
13	2	2	2
14	2	0	0
Total	36	20	21

TABLA 5
CUADRO COMPARATIVO DE LOS SISTEMAS DE REPLICACIÓN DE ARCHIVOS

El sistema de archivos GlusterFS es el Sistema de Archivos Distribuido que cubre mejor los requerimientos; lo cual se ha concluido por las siguientes razones:

- GlusterFS está diseñado para trabajar en un ambiente de clúster.
- Las características de GlusterFS permiten tener configurado varios sistemas redundantes logrando así evitar tener puntos únicos de fallo.
- Por la escalabilidad que este posee ya que se puede transmitir grandes cantidades de información sin tener ningún problema.
- CODA requiere de gran cantidad de recursos como memoria RAM4 ya que por ejemplo si se pretende replicar 100GB, se debe de separar el 4 % del espacio en disco para almacenar los metadatos, es decir se utilizaran 4GB para tal operación.

4.2. GLUSTERFS

GlusterFS es un sistema de archivos que está diseñado para ser utilizado en un ambiente de clúster, es un sistema innovador ya que puede manejar grandes cantidades de información incluso peta-bytes.

Principales consideraciones del diseño de GlusterFS

- **Capacidad de Expansión.**
- **Facilidad de administración.**
- **Geo-replication.**
- **Directory Quota.**
- **POSIX ACLs Support**.
- **Escalabilidad sobrepasando peta-bytes**
- **Balanceo de carga.**
- **Alta disponibilidad.**

4 RAM: (Random Access Memory) Memoria de acceso aleatorio.

- **Completamente distribuido.**
- **Encriptación de los datos**
- **Diseñado para trabajar en ambientes de clúster**
- **Comunicación segura con los nodos ya que implementa mecanismos de autenticación.**

Esquema de Funcionamiento

Para la utilización de GlusterFS se hacen uso de dos componentes principales que son: GlusterFS Server y GlusterFS Client.

GlusterFS Server: Aquí se exportan los volúmenes para los nodos del clúster, se pueden configurar varios GlusterFS Server para evitar tener un punto único de fallo.

GlusterFS Client: Este componente es instalado y configurado en los nodos del clúster, la función que cumple este es el montar los volúmenes que se encuentren configurados en el servidor.

Un aspecto importante a tener en cuenta en el GlusterFS client es que todos los clientes deben tener soporte para fuse.

Gluster Console Manager: Es una utilidad de línea de comandos que simplifica la configuración y gestión del entorno de almacenamiento.

Con el uso de esta consola los administradores pueden crear nuevos volúmenes, iniciar y detener los volúmenes, según como se requiera.

A continuación el esquema de funcionamiento de GlusterFS.

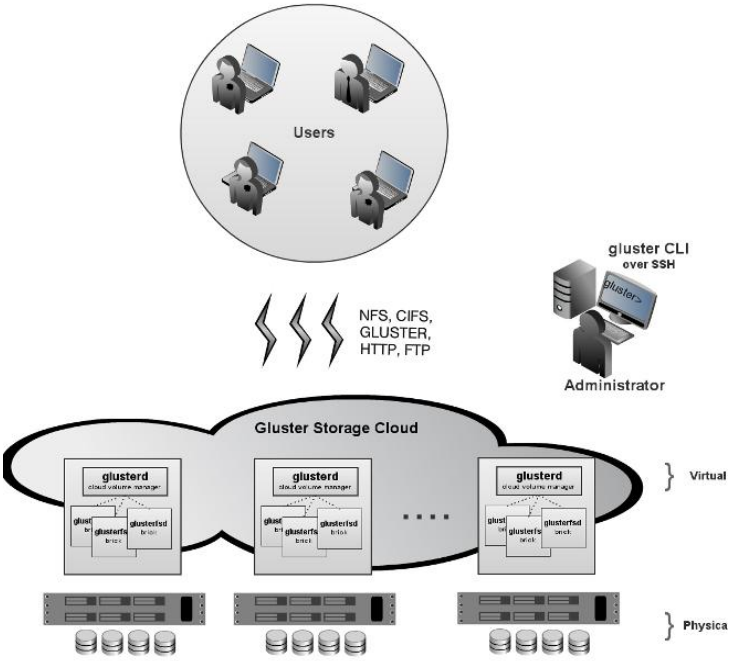


FIGURA 2
ESQUEMA DE FUNCIONAMIENTO DE GLUSTERFS SEGÚN (GLUSTER)

Se planeó los siguientes esquemas para asegurar los clúster de servicio de web y mail.

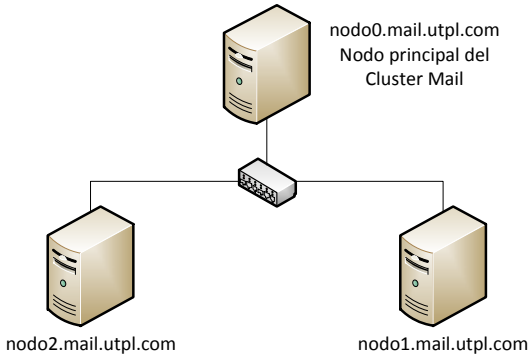


FIGURA 3
ESQUEMA DE SEGURIDAD EN CLÚSTER DE SERVICIO DE MAIL.

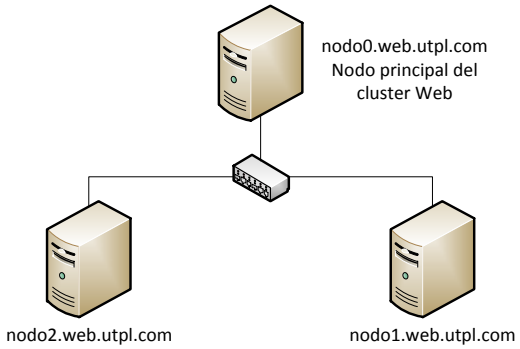


FIGURA 4
ESQUEMA DE SEGURIDAD EN CLÚSTER DE SERVICIO WEB.

5. PRUEBAS

Para determinar los resultados de cada una de las implementaciones realizadas sobre el clúster, se realizó un conjunto de pruebas que se listan a continuación:

- Pruebas de funcionalidad
- Pruebas de Integridad de Datos
- Pruebas de Carga
- Pruebas de Stress.

5.1. PRUEBAS DE FUNCIONALIDAD

Objetivo de la prueba: Verificar que el método escogido para la replicación, permita tener la misma información en todos los nodos.
Técnica: Envío de Mails hacia una determinada cuenta del servidor de correo, peticiones al servidor web.

Haciendo uso de un sniffer interceptar los paquetes que se transmiten en la red.
Criterio de conclusión: Verificar que el dueño de la cuenta a la que se envió los mail, pueda verlos en todos los nodos del clúster.

Para la realización de esta prueba se seleccionó un grupo de PC's clientes que se encuentren en la subred 172.16.17.0/24.

Se realizaron 50 casos de pruebas y tras la ejecución de estas en el clúster web y mail se obtuvieron los siguientes resultados:

Número de peticiones	Resultado
48	Correctas
2	Incorrectas

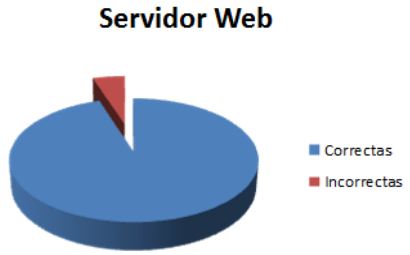


FIGURA 5
FUNCIONALIDAD DEL SERVIDOR WEB

Número de peticiones	Resultado
48	Correctas
2	Incorrectas

FIGURA 6
FUNCIONALIDAD DEL SERVIDOR MAIL

Como se puede apreciar en las gráficas el sistema de replicación obtuvo un 96.25% de cumplimiento con esta prueba.

5.2. PRUEBAS DE INTEGRIDAD DE DATOS

Objetivo de la prueba: Verificar que la información no haya sufrido alteraciones durante la replicación.

Técnica: Enviar mails con archivos adjuntos, hacia una cuenta del servidor de correos, utilizar Outlook para el envío de mails y utilizar el Webmail para revisarlos.
Realizar peticiones hacia el servidor web.

Subir contenido al servidor Web.

Criterio de conclusión: Que la información del mail enviado llegue completa sin sufrir alteración, es decir llegue tal cual se la envió. Que las peticiones al servidor web, muestren la misma información que la de los nodos.

En el servidor web para verificar que los datos se replicaron correctamente y sin ninguna alteración se procedió a subir contenido al servidor web y realizar peticiones al servidor web desde un conjunto de PC's en la subred 172.16.17.0/24 y verificar que en los nodos se muestre la información que fue subida en el servidor web.

En el servidor mail para realizar la verificación de esta prueba se procedió a hacer envíos de mail con y sin archivos adjuntos, luego se verificó que los datos enviados fueron recibidos de forma correcta y sin sufrir modificación alguna.

A continuación se muestran los resultados de la ejecución de las pruebas en el servidor web y mail:

Número de peticiones	Resultado
78	Correctas
2	Incorrectas

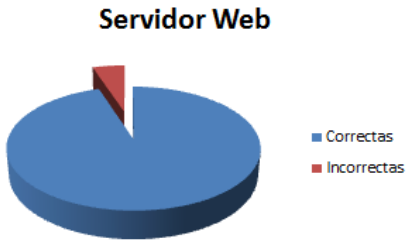


FIGURA 7
FUNCIONALIDAD DEL SERVIDOR WEB

Número de peticiones	Resultado
77	Correctas
3	Incorrectas

FIGURA 8
FUNCIONALIDAD DEL SERVIDOR MAIL

5.3. PRUEBAS DE CARGA

Objetivo de la prueba: Obtener un tiempo de respuesta ante determinados tamaños de mails y tamaño de contenido en el servidor web.

Técnica: Enviar un mail sin adjunto
Enviar un mail con un adjunto de 1MB.
Enviar un mail con un adjunto de 2MB.
Enviar un mail con un adjunto de 6MB.
Enviar un mail con un adjunto de 10MB.
Subir contenido al servidor web de 10MB.
Subir contenido al servidor web de 50MB.
Subir contenido al servidor web de 100MB.
Subir contenido al servidor web de 500MB.

Criterio de conclusión: Que el sistema fue capaz de replicar tal cantidad en un determinado tiempo.

En el **servidor mail** para realizar la verificación de esta prueba se procedió a hacer envíos de mail con archivos adjuntos de distintos tamaños desde clientes de correo y desde el Webmail hacia el servidor de correo, luego a través de los mismos se verificó que los mails fueron recibidos de forma correcta con sus adjuntos respectivos.

A continuación se presenta una tabla con los resultados.

Cantidad	Tamaño MB	Tiempo (seg.)
1	0	Menos de 1
1	1	Menos de 1
1	2	Menos de 1
1	6	Menos de 1
1	10	Menos de 2

TABLA 6
TIEMPOS DE RESPUESTA DEL SERVIDOR MAIL

En el **servidor web** para realizar la verificación de esta prueba se procedió a subir contenido de distintos tamaños en el servidor web, luego se procedió a realizar peticiones en los nodos clientes para verificar que el contenido fue replicado.

A continuación se presenta una tabla con los resultados.

Cantidad	Tamaño MB	Tiempo (seg.)
1	10	Menos de 1
1	50	Menos de 1
1	100	Menos de 3
1	500	Menos de 6

TABLA 7
TIEMPOS DE RESPUESTA DEL SERVIDOR WEB

5.4. PRUEBAS DE STRESS

Objetivo de la prueba: Verificar el comportamiento del sistema de replicación de archivos ante envíos simultáneos de mails y peticiones al servidor web.

Técnica: Desarrollar un script que envíe gradualmente cantidades de mails.

Desarrollar un script que envíe peticiones simultáneas al servidor web.

Criterio de conclusión: Tiempos de respuesta de acuerdo a la cantidad de mail que se envió, así como también comprobar los envíos satisfactorios.

En el **servidor web** para realizar las pruebas de Stress se procedió a desarrollar un script el mismo realiza múltiples peticiones al servidor web.

El script contiene el siguiente código.

for id in {1..$np}; do elinks –dump http://nodo0w.web.utpl.com ; done

En donde:

for: utilizado para crear un ciclo.
$np: especifica el número de peticiones.
elinks: Utilizado para realizar una petición a un servidor web

Este script fue ejecutado desde una consola de Linux en un equipo conectado a la subred 172.16.17.0/24. De la siguiente manera:

[root@nodo0w ~]#./peticionesweb 10

Donde 10 significa el número de peticiones.

Se realizaron varios casos de pruebas con distintas cantidades de peticiones y los resultados son los que se muestran en la siguiente tabla.

Nro. de peticiones	Tiempo (segundos)
1	0
5	1
10	2
20	2.5
40	3
80	5
160	7
320	13
640	25
1280	49

TABLA 8
PETICIONES AL SERVIDOR WEB

Para las pruebas en el **servidor mail**, también se desarrolló un script el cual enviaba múltiples mails a un destinatario en un mismo instante.

El código del script es el siguiente:

for test int {1..$nm}; do echo “Test de prueba” | mail –s “test #”$test prueba@nodo0.mail.utpl.com; done

Este script fue ejecutado desde una consola de Linux en un equipo conectado a la subred 172.16.17.0/24. De la siguiente manera:

[root@nodo0 ~]#./peticionesmail 40

Donde 40 significa el número de mails enviados.

Luego de ejecutar el script con distintas cantidades se obtuvo los resultados que se muestran en la siguiente tabla.

Nro. de Mails enviados	Tiempo(seg)
10	1
50	4
100	9
200	19
500	52
1000	90
5000	Fallo

TABLA 9
PETICIONES AL SERVIDOR MAIL

6. RESULTADOS

De la presente investigación se obtuvieron los siguientes resultados.

- Encriptación de la información que pasa a través del clúster de servicios web y mail de la UTPL.
- Aseguramiento de los protocolos POP3, IMAP, HTTP mediante la implementación de protocolos seguros como: POP3S, IMAPS, HTTPS

- Implementación de nuevas directivas en las configuraciones del servidor de correo electrónico.
- Implementación de un nuevo sistema de replicación de archivos GlusterFS.

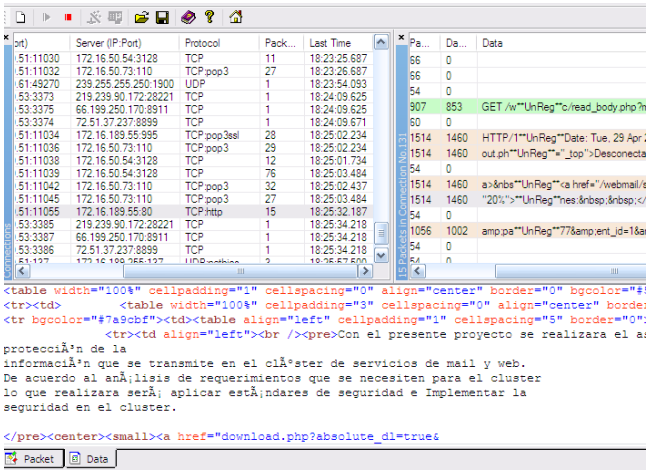


FIGURA 9

CAPTURA DEL TRÁFICO DEL SERVICIO WEB Y MAIL SIN LA IMPLEMENTACIÓN DE PROTOCOLOS SEGUROS

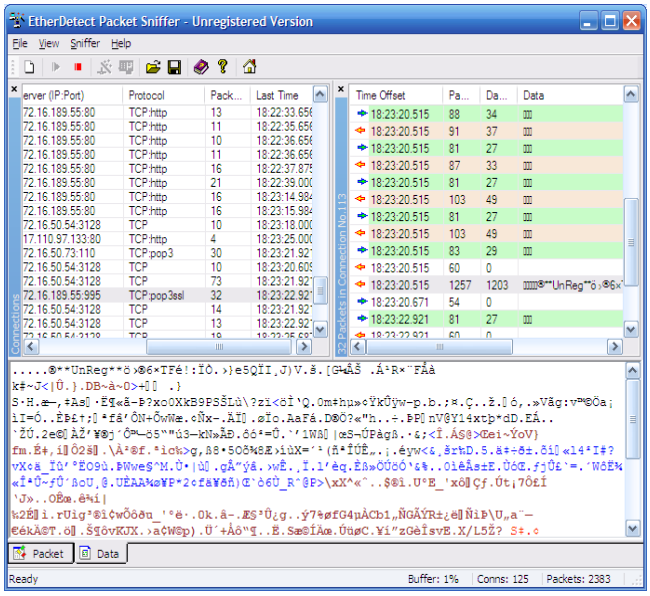


FIGURA 10

CAPTURA DEL TRÁFICO DEL SERVICIO WEB Y MAIL CON LA IMPLEMENTACIÓN DE PROTOCOLOS SEGUROS

Como se puede apreciar el tráfico que pasa por los protocolos seguros está protegida, ya que no se puede interpretar lo que el sniffer detecta.

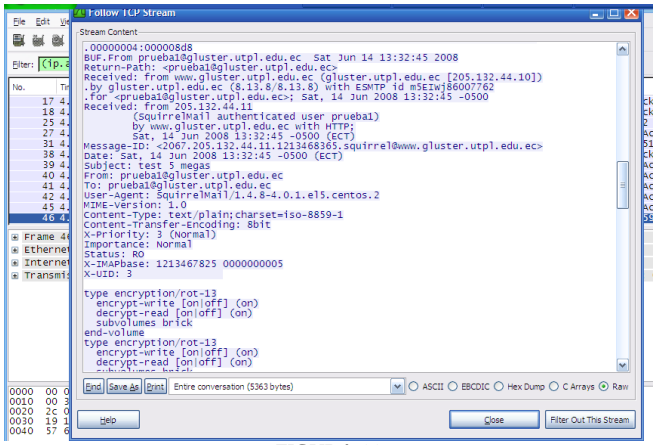


FIGURA 11

CAPTURA DEL TRÁFICO QUE PASA A TRAVÉS DE CADA UNO DE LOS NODOS DEL CLÚSTER

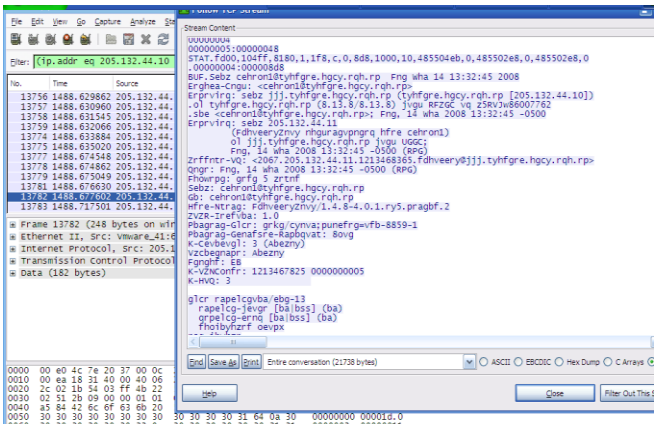


FIGURA 12

CAPTURA DEL TRÁFICO QUE PASA A TRAVÉS DE CADA UNO DE LOS NODOS DEL CLÚSTER

En esta captura podemos apreciar que la información esta encriptado.

7. DISCUSIÓN

Conocer y manejar bien las configuraciones de cada uno de los servicios de web y mail es de vital importancia, ya que al poseer este conocimiento se está en la capacidad de poder cambiar ciertas configuraciones que vienen dadas por defecto al momento de instalaros, ya que al no cambiar estas configuraciones genéricas se verían comprometidos los servicios de Mail y Web porque las configuraciones que están por defecto son de conocimiento público por el mismo hecho de que son aplicaciones de código abierto.

La clave para estar preparados ante un eventual ataque a un servicio es el de realizar distintas simulaciones de ataques, ya que con estas simulaciones se tiene una perspectiva de lo que se podría tener en un posible ataque a un servicio.

El crear varios escenarios de ataques permite a los administradores de los servicios tener una base de

conocimientos para poder dar solución a problemas de seguridad que se podrían originar en un momento dado. El éxito de estas simulaciones está en recolección de la información que se va generando al momento de ir dando solución a los distintos escenarios que se planteen.

La solución que se ha dado para manejar la gestión de información la considero que es muy conveniente ya que el nuevo sistema de replicación de archivos que se utilizó GlusterFS, se adaptó a los requerimientos que se tenían para los servicios, ahora si bien es cierto de que existen muchos sistemas de replicación de archivos se debe realizar un análisis de todas las características que tienen cada uno de los servicios, ya que de ello depende el seleccionar un sistema de replicación de archivos que se adapte a los servicios.

Como todo sistema tiene sus ventajas y desventajas, GlusterFS no sería la excepción pues el algoritmo que utiliza para realizar la encriptación es simple, la alternativa que se debe manejar para solventar esta deficiencia es la utilización de canales de conexión seguros utilizando SSL, aprovechando de que el diseño de GlusterFS si lo permite y a más de esta alternativa actualmente se está trabajando para el desarrollo de otro algoritmo de encriptación que sustituirá al que se tiene actualmente.

La nueva versión de GlusterFS cuenta con nuevas características, entre las cuales podemos destacar la Geo-Replication que consiste en replicar la información hacia servidores remotos que no necesariamente estén sobre la misma red interna, esto sería de gran ayuda ya que si se tiene la implementación de esta característica sería muy fácil la recuperación de la información ante cualquier eventualidad catastrófica que pueda sufrir los servidores, y con esto se lograría contar con un 99.9% de disponibilidad del servicio.

8. CONCLUSIONES

- En el análisis realizado en los clústers se identificó que existían bajos niveles de seguridad implicando graves riesgos para de seguridad en la información de la organización.
- Con GlusterFS se evita que el administrador tenga que ejecutar scripts de sincronización de datos; ya que el sistema permite la replicación de forma automática y todos los cambios realizados son reflejados de manera automática hacia los demás nodos.
- Con el uso de la consola de administración de GlusterFS se hace más fácil la administración de los directorios a replicar hacia los nodos, lo cual evita cometer al usar archivos de configuración.
- GlusterFS es el que mejor se adapta a los requerimientos del clúster de servicios web y mail de la UTPL, es apropiado para trabajar con ambientes de clúster y posee características que ayudan a mejorar el rendimiento, seguridad y escalabilidad del clúster.
- Con el resultado obtenido en las pruebas de funcionalidad se pudo comprobar la eficiencia de GlusterFS, llegando a obtener un 97.7% en la replicación de archivos, el 2.3% faltante se debe a ajustes que se realizaron sobre la configuración de GlusterFS.
- Se obtuvo un 97% en los resultados de las pruebas de integridad de datos, en la cual se pudo comprobar que no existió alteración de datos, verificando que GlusterFS es un sistema seguro para la replicación de archivos, el 3% faltante se debe a que estas pruebas eran realizadas desde el mismo servidor hacia los nodos, esto ocasionó que existiera mayor procesamiento en el clúster.
- El servidor mail respondió satisfactoriamente en un 96.25% ante las peticiones simultáneas hechas al servidor, el 3.75% faltante se debe a que las peticiones eran enviadas desde el mismo servidor hacia los nodos, esto ocasionó que exista retardo en las respuestas.
- El servidor web respondió satisfactoriamente ante los distintos casos de prueba de peticiones simultáneas hacia el servidor web.
- Las características de GlusterFS permiten tener configurado varios sistemas redundantes logrando así evitar tener puntos únicos de fallo.
- Con la implementación de los protocolos seguros POP3S, IMAPS y HTTPS se logró el aseguramiento de los servicios web y mail evitando con esto que puedan ser interceptados por terceras personas.
- La característica rot-13 que posee GlusterFS para la encriptación de la información, es débil ya que se trata de un algoritmo simple, actualmente el grupo de desarrollo de GlusterFS está desarrollando un mecanismo más fuerte para la encriptación de los datos que se transmiten por el clúster.

REFERENCIAS

Baker, M. Clúster Computing. University of Portsmouth.

Braam, P., Baron, R., Harkes, J., & Schnieder, M. The Coda HOW TO. School of Computer Science, Carnegie Mellon University.

Cohen, F. (n.d.). A Preliminary Classification Scheme for Information System Threats, Attacks, and Defenses; A Cause and Effect Model; and Some Analysis Based on That Model". Retrieved from http://www.all.net/journal/ntb/cause-and-effect.html

Gluster. (n.d.). Retrieved Julio 28, 2011, from http://www.gluster.org/docs/index.php/GlusterFS

Guía de Administración de Redes con Linux. (n.d.). Retrieved from http://www.ibiblio.org/pub/Linux/docs/LuCaS/Manuales-LuCAS/GARL2/garl2/index.html

Manual de seguridad de Red Hat Enterprise Linux 4. (n.d.). Retrieved from http://www.opencontent.org/openpub/

Moore, A., Ellison, R., & Richard, C. (n.d.). Attack Modeling for Information Security and Survivability. Retrieved from http://www.cert.org/archive/pdf/01tn001.pdf

Roselló Vicente, J. A. (n.d.). Clustering de Alta Disponibilidad bajo GNU/Linux. Retrieved from http://www.bisente.com/documentos/clustering/informe.pdf

Schneier, B. (n.d.). Attack trees. Retrieved from http://www.schneier.com/paper-attacktrees-ddj-ft.html

Jorquera, D. M. (2008). Administración de servicios de Internet: De La Teoría a La Práctica. Universidad de Alicante.

Salavert Casamor, A. (2003). Los protocolos en las redes de ordenadores. Edicions UPC.