

UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA



IMPLEMENTACION DE RED PARA UNA BASE DE DATOS ORACLE 11GR2 EN ALTA DISPONIBILIDAD

INFORME DE SUFICIENCIA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO DE TELECOMUNICACIONES

PRESENTADO POR:

JORGE CHRISTIAN RIVEROS BETETA

**PROMOCIÓN
2009- I**

**LIMA – PERÚ
2013**

**IMPLEMENTACION DE RED PARA UNA
BASE DE DATOS ORACLE 11GR2
EN ALTA DISPONIBILIDAD**

El presente Informe va dedicado principalmente a Dios, por permitirnos seguir adelante en la vida y ayudarnos a superar con éxito todos los obstáculos que se presentan en la búsqueda de nuestra formación personal y espiritual
a mis padres por ser fieles consejeros, por brindarme su apoyo y sabiduría en todo momento.

SUMARIO

El presente informe se inicia con una evaluación ante la necesidad de disponer de una arquitectura de red moderna que permita afrontar los requerimientos de disponibilidad y performance para soportar los datos críticos de una empresa.

Se explica la importancia de una Base de Datos como elemento fundamental para el soporte de la información en las empresas, también se revisa los conceptos básicos de la Base de Datos Oracle 11gR2 y su implementación en alta disponibilidad.

Durante el desarrollo del informe se revisa conceptos teóricos, relacionados a cada una de las tecnologías empleadas para el diseño de la arquitectura de red, además los beneficios de usar los productos involucrados.

Al final del informe se revisa los resultados obtenidos durante la realización de las pruebas de performance y disponibilidad.

INDICE

CAPITULO I

PLANTEAMIENTO DEL PROBLEMA	2
1.1 Descripción del problema	2
1.1.1 Problemáticas en el diseño Lógico.....	2
1.1.2 Problemáticas en el diseño físico	2
1.2 Objetivos del trabajo	2
1.2.1 Objetivo general	2
1.2.2 Objetivos específicos	3
1.3 Evaluación del problema.....	3
1.3.1 Evaluación a la problemática del diseño lógico.....	3
1.3.2 Evaluación a la problemática del diseño físico	3
1.4 Limitaciones del trabajo	4
1.5 Síntesis del trabajo	4

CAPITULO II

MARCO TEORICO CONCEPTUAL	6
2.1 Marco histórico.....	6
2.1.1 Década de 1950	6
2.1.2 Década de 1960	6
2.1.3 Década de 1970	7
2.1.4 Década de 1980	7
2.1.5 Década de 1990	7
2.2 Marco teórico	8
2.2.1 Alta disponibilidad	8
2.2.2 Base de datos relacional.....	9
2.2.3 DAS (Direct attached storage)	9
2.2.4 NAS (Network attached storage)	9
2.2.5 SAN (Storage area network)	10
2.2.6 World wide name (WWN).....	11
2.2.7 Zoning (Zonificación).....	11

2.2.8 Masking	12
2.2.9 Fiber channel protocol (FCP).....	12
2.2.10 Topologías fiber channel (FC)	13
2.2.11 HBA (Host bus adapters).....	14
2.2.12 SCSI (Small computer system interface)	15
2.2.13 iSCSI (Internet SCSI).....	15
2.2.14 LUN (Número de unidad lógica)	15
2.2.15 NPIV (N_Port ID virtualization).....	15
2.2.16 RAID (Redundant array of independent disks).....	16
2.2.17 Virtual switching system (VSS)	18
2.2.18 Linux channel bonding.....	18
2.2.19 Oracle database server.....	19
2.2.20 Automatic storage management (ASM).....	22
2.2.21 Oracle real application clusters.....	23
CAPITULO III	
METODOLOGIA PARA LA SOLUCION DEL PROBLEMA	25
3.1 Alternativas de solución.....	25
3.1.1 Cisco Systems	25
3.1.2 Dell	28
3.1.3 Hitachi.....	30
3.1.4 Brocade communications systems.....	31
3.1.5 Fortinet	32
3.1.6 Oracle Corporation.....	33
3.2 Solución del problema.....	35
3.2.1 Descripción del enclosure PowerEdge M1000e	35
CAPITULO IV	
ANALISIS Y PRESENTACION DE RESULTADOS	72
4.1 Análisis de performance	72
4.2 Creación del usuario de prueba	74
4.3 Creación del procedimiento de prueba	75
4.4 Impacto y rendimiento en el RAC Oracle	76
4.5 Evaluación económica	77
CONCLUSIONES Y RECOMENDACIONES	79
BIBLIOGRAFIA.....	81
LISTA DE ACRONIMOS.....	83
LISTA DE SIGNIFICADO DE PALABRAS	85

PROLOGO

En la actualidad, las empresas buscan constantemente medios que les permitan cubrir la creciente demanda de información, impulsado por el auge que internet ha tenido en la última década. Producto de este crecimiento nuevos servicios se han desarrollado, soportados sobre aplicaciones más complejas y con nuevos requerimientos de operación, obligando a las empresas, a buscar soluciones que sean capaces de mantener niveles óptimos de operación con visión de crecimiento.

Uno de los puntos más críticos para las empresas es mantener la información en el tiempo, para esto diversas empresas han desarrollado sistemas complejos. Oracle es en la actualidad, la empresa líder en brindar soluciones de soporte y gestión de grandes volúmenes de datos. Oracle 11gR2 es el producto bandera de la empresa y el más implementado en el mundo, Oracle RAC 11gR2 (Real Application Cluster) es la versión en alta disponibilidad que satisface las necesidades actuales.

Hoy, Oracle también brinda soluciones de hardware. Pero en el diseño implementado se mostraran la interacción con otros productos, que cubren la parte física, dejando a Oracle solo con la gestión Lógica de los datos.

En la arquitectura física de red se usa productos Cisco, Brocade, Dell, y Hitachi. Se describen diferentes técnicas de implementación de acuerdo a la tecnología existente. A continuación se menciona algunas de las técnicas a emplear:

- Networking: IPV4, routing, VLAN, stack wise, VSS.
- SAN: NPV, FC, LUN, VSAN.

El informe se enfoca en la arquitectura física y los beneficios de implementar una arquitectura redundante. Se toma en cuenta conceptos teóricos básicos relacionados al funcionamiento lógico de la Base de Datos Oracle RAC 11gR2.

CAPITULO I

PLANTEAMIENTO DEL PROBLEMA

1.1 Descripción del problema

1.1.1 Problemáticas en el diseño lógico

El principal problema de desplegar soluciones en base de datos para empresas, es la cantidad de información que puede manejar y soportar una solución cuando se implementa de manera individual.

Esta limitación esta directamente relacionada a la capacidad de procesamiento que ofrecen los actuales servidores, porque son ellos los responsables de soportar las soluciones en administración de base de datos. En la actualidad existe una gran diversidad de procesadores que manejan arquitecturas de múltiples núcleos ofreciendo a los actuales servidores gran capacidad de procesamiento.

Estas nuevas arquitecturas ayudan a solventar parte de la problemática en capacidad de procesamiento, pero ante la creciente demanda de información, se hace evidente la necesidad de multiplicar incluso esta capacidad, además de otorgar la confiabilidad que las empresas demandan.

1.1.2 Problemáticas en el diseño físico

En el diseño de la arquitectura de red que soporte las aplicaciones en base de datos, existen muchos retos tecnológicos, tanto en la implementación desde las conocidas redes LAN hasta las redes de almacenamiento SAN, empleadas exclusivamente para integración de los servidores y storages, otro desafío a considerar es la disponibilidad que estas tecnologías en conjunto deberán brindar a la arquitectura global, también es importante la estrategia de seguridad que la criticidad de estas aplicaciones demandan y su integración al diseño general para proveer una solución escalable, redundante y eficiente.

1.2 Objetivos del trabajo

1.2.1 Objetivo general Diseñar una arquitectura de red para una base de datos relacional en alta disponibilidad, que sirva como base referencial para futuras implementaciones empresariales.

1.2.2 Objetivos específicos

- Estudiar las tecnologías y estándares existentes para la implementación de arquitecturas de redes empresariales.
- Analizar métodos y técnicas estudiadas para la implementación de red para una base de datos en alta disponibilidad.
- Aplicar el diseño de red en una empresa de telecomunicaciones.

1.3 Evaluación del problema

Para el diseño general se tiene en cuenta las limitaciones mencionadas anteriormente y se busca las posibles soluciones de acuerdo a las problemáticas encontradas:

1.3.1 Evaluación a la problemática del diseño lógico

La mejor alternativa para superar la barrera tecnológica que mejora el aumento en la capacidad de procesamiento, consiste en el agrupamiento de múltiples servidores. El termino “cluster” es muy empleado para estos fines como se muestra en la figura 1.1.

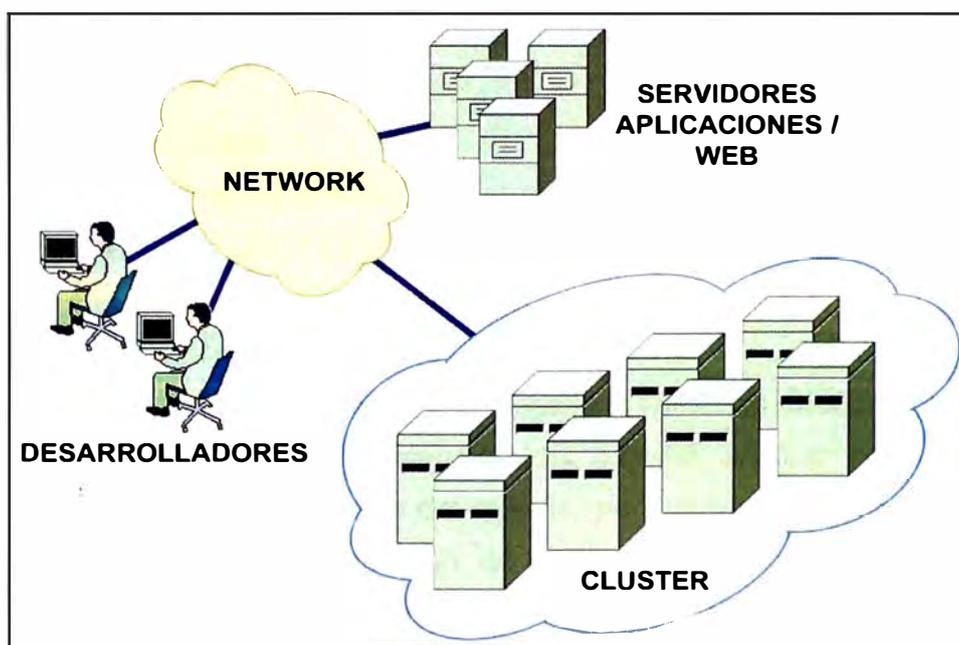


Fig. 1.1 Visión general del diseño.

1.3.2 Evaluación a la problemática del diseño físico

Un requerimiento necesario para construir una verdadera arquitectura redundante es disponer de elementos activos en la red que permitan la conmutación automática en la situación de falla o comportamientos no esperados, esto debe ser considerado en todo el diseño. A continuación se muestra la relación de elementos de red sensibles a este requerimiento en la figura 1.2.

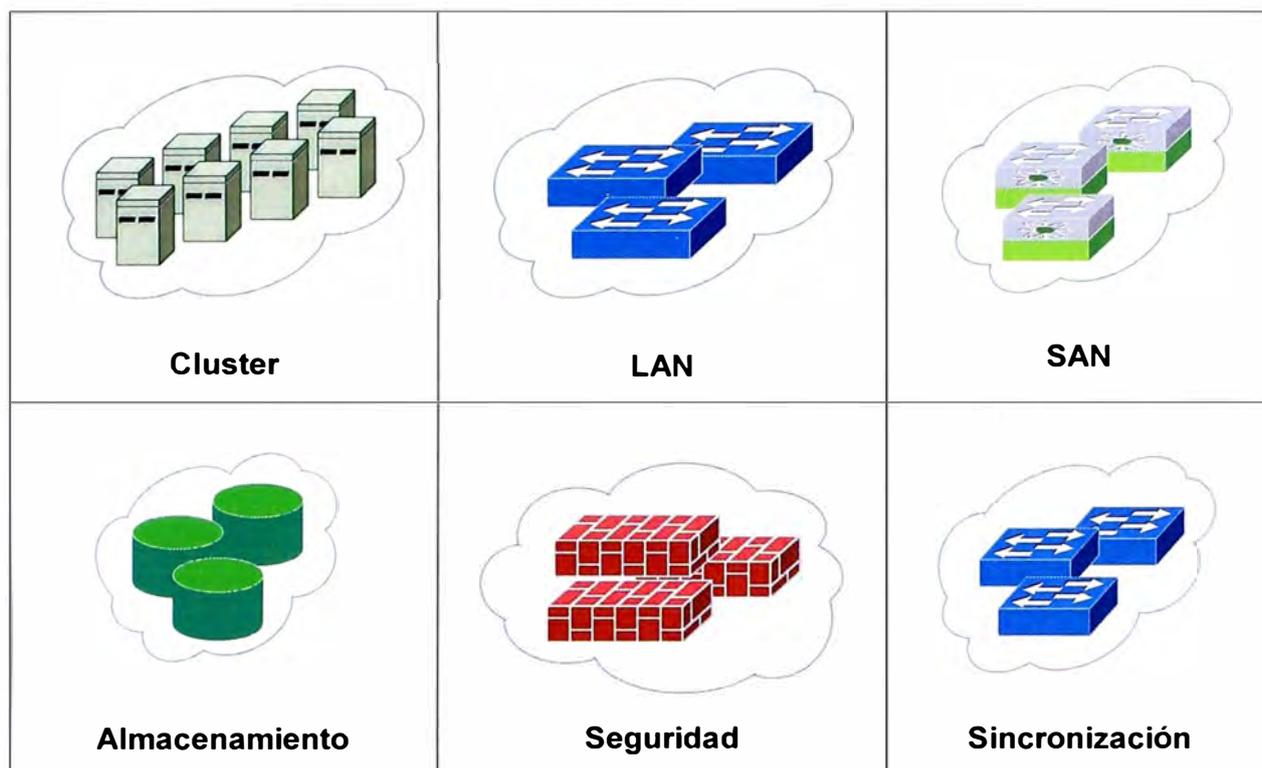


Fig. 1.2 Elementos de red presentes en el diseño.

1.4 Limitaciones del trabajo

En el presente informe se revisa y estudia las tecnologías relacionadas a la implementación de redes LAN y SAN en alta redundancia, así como las consideraciones de seguridad que la aplicación necesita. Los estudios relacionados al funcionamiento lógico de la Base de datos

La implementación de la base de datos Oracle 11gR2 en modo redundante son explicados brevemente, no es la intención del informe abarcar conceptos a nivel aplicación ya que exceden al alcance del estudio, por tratarse de temas netamente de programación y sistemas, la explicación de su funcionamiento básico se considera importante, para el claro entendimiento de la arquitectura general y su interrelación entre las partes del sistema.

1.5 Síntesis del trabajo

Para solventar los requerimientos del diseño de red en alta disponibilidad se resume los puntos críticos que son considerados:

- Revisión de la arquitectura y funcionamiento de la base de datos en alta redundancia.
- Revisión de la plataforma tecnológica donde se instala la aplicación de base de datos, se tiene tres modelos posibles de servidores stand alone, chasis y Virtuales.

- Diseño e implementación de la red LAN en modo redundante.
- Diseño e implementación de la red SAN en modo redundante.
- Diseño e implementación de la seguridad informática en modo redundante.
- Arquitectura global y bloques funcionales del diseño de red.

CAPITULO II

MARCO TEORICO CONCEPTUAL

2.1 Marco histórico

Las bases de datos se han constituido como una de las herramientas más ampliamente difundidas en la actual sociedad de la información, utilizadas como fuentes secundarias en cuanto recuperación y almacenamiento de información en todos los campos de nivel científico, social, económico, político y cultural.

El uso de sistemas de bases de datos automatizadas, se desarrolló a partir de la necesidad de almacenar grandes cantidades de datos, para su posterior consulta, producidas por las nuevas industrias que creaban gran cantidad de información.

Herman Hollerit (1860-1929) fue denominado el primer ingeniero estadístico de la historia, ya que inventó una computadora llamada "Máquina automática perforadora de tarjetas". Para hacer el censo de Estados Unidos en 1880 se tardaron 7 años para obtener resultados, pero Herman Hollerit en 1884 creó la máquina perforadora, con la cual, en el censo de 1890 dio resultados en 2 años y medio, donde se podía obtener datos importantes como número de nacimientos, población infantil y número de familias. La máquina usó sistemas mecánicos para procesar la información de las tarjetas y para tabular los resultados.

2.1.1 Década de 1950

En este lapso de tiempo se dá origen a las cintas magnéticas, las cuales sirvieron para suplir las necesidades de información de las nuevas industrias. Por medio de este mecanismo se empezó a automatizar la información de las nóminas que consistía en leer una cinta o más y pasar los datos a otra, y también se podían pasar desde las tarjetas perforadas. También se pudo simular un sistema de backup, que consiste en hacer una copia de seguridad o copia de respaldo para luego guardarlo en un medio extraíble la información importante. La nueva cinta a la que se transfiere la información pasó a ser una cinta maestra.

2.1.2 Década de 1960

El uso de los discos en ese momento fue un adelanto muy efectivo, ya que por medio de este soporte se podía consultar la información directamente, esto ayudó a ahorrar tiempo. No era necesario saber exactamente donde estaban los datos en los discos, ya que en milisegundos era recuperable la información. A diferencia de las cintas magnéticas, ya no era necesaria la secuencialidad, y este tipo de soporte empezó a ser ambiguo.

Los discos dieron inicio a las Bases de Datos, de red y jerárquicas, pues los programadores con su habilidad de manipulación de estructuras junto con las ventajas de los discos era posible guardar estructuras de datos como listas y árboles.

2.1.3 Década de 1970

Edgar Frank Codd (23 de agosto de 1923 – 18 de abril de 2003), en un artículo “Un modelo relacional de datos para grandes bancos de datos compartidos” en 1970, definió el modelo relacional y publicó una serie de reglas para la evaluación de administradores de sistemas de datos relacionales y así nacieron las bases de datos relacionales.

A partir de los aportes de Codd el multimillonario Larry Ellison desarrolló la base de datos Oracle, el cual es un sistema de administración de base de datos, que se destaca por sus capacidades de transacciones, estabilidad, escalabilidad y multiplataforma.

Inicialmente no se usó el modelo relacional debido a que tenía inconvenientes por el rendimiento, ya que no podían ser competitivas con las bases de datos jerárquicas y de red. Esta tendencia cambió por un proyecto de IBM el cual desarrolló técnicas para la construcción de un sistema de bases de datos relacionales eficientes, llamado System R.

2.1.4 Década de 1980

Las bases de datos relacionales con su sistema de tablas, filas y columnas, pudieron competir con las bases de datos jerárquicas y de red, ya que su nivel de programación era bajo y su uso muy sencillo.

En esta década el modelo relacional ha conseguido posicionarse del mercado de las bases de datos. Y también en este tiempo se iniciaron grandes investigaciones paralelas y distribuidas, como las bases de datos orientadas a objetos.

2.1.5 Década de 1990

Para la toma de decisiones se crea el lenguaje SQL, que es un lenguaje programado para consultas. El programa de alto nivel SQL es un lenguaje de consulta estructurado que analiza grandes cantidades de información el cual permite especificar diversos tipos de operaciones frente a la misma información, a diferencia de las bases de datos de los 80 que eran diseñadas para las aplicaciones de procesamiento de transacciones. Los

grandes distribuidores de bases de datos incursionaron con la venta de bases de datos orientada a objetos. El boom de esta década fue la aparición de la WWW (Word wide web) por éste medio se facilitaba la consulta de las bases de datos. Actualmente esta tecnología tiene una amplia capacidad de almacenamiento de información, también una disponibilidad de siete días a la semana por veinticuatro horas del día sin interrupciones, a menos que se planifique el mantenimiento de las plataformas o sistemas. Si bien es cierto que las bases de datos, son uno de los recursos más importantes para el desarrollo de cualquier área, el contexto histórico ha demostrado que son de una gran influencia en el progreso de los países a nivel mundial, permitiendo un almacenamiento y acceso confiable, eficiente y practico en el uso de la información que se produce.

Las bases de datos no solo han servido como repositorios de información, sino también como una de las herramientas más importantes en el campo de la difusión científica. A nivel mundial existen muchas iniciativas propuestas por profesionales de la información, que a partir de un trabajo colaborativo e interdisciplinario, han dispuesto de herramientas para el uso y desarrollo de aquellos interesados en el área de la ciencia de la Información.

2.2 Marco teórico

2.2.1 Alta disponibilidad

El término “disponibilidad” hace referencia a la probabilidad de que un servicio funcione adecuadamente en cualquier momento.

El término “fiabilidad”, que se utiliza en algunos casos, se refiere a la probabilidad de que un sistema funcione normalmente durante un período de tiempo dado. Esto se denomina “continuidad del servicio”.

La disponibilidad se expresa con mayor frecuencia a través del índice de disponibilidad (un porcentaje) que se mide dividiendo el tiempo durante el cual el servicio está disponible por el tiempo total, la tabla 2.1 muestra valores referenciales.

TABLA N° 2.1 Índice de disponibilidad.

Índice de disponibilidad	Tiempo de inactividad	Niveles de Calidad
98%	7 días	6 (Sin disponibilidad)
99%	3 días y 15 horas	5 (Pésimo)
99,9%	8 horas y 48 minutos	4 (Malo)
99,99%	53 minutos	3 (Tolerable)

99,999%	5 minutos	2 (Bueno)
99,9999%	32 segundos	1 (Excelente)

Cabe señalar que los niveles aceptables en la mayoría de centro de datos son el nivel 1 y nivel 2.

2.2.2 Base de datos relacional

Una base de datos relacional es una base de datos que cumple con el modelo relacional, el cual es el modelo más utilizado en la actualidad, permiten establecer interconexiones (Relaciones) entre los datos guardados en tablas. A través de dichas conexiones se relaciona los datos de ambas tablas, de ahí proviene su nombre "Modelo relacional".

En una base de datos relacional se accede a los datos por medio de relaciones donde las relaciones que almacenan los datos son llamadas "Relaciones base" y su implementación es llamada "Tabla". Otras relaciones no almacenan datos, pero son calculadas al aplicar operaciones relacionales y son llamadas "relaciones derivadas", su implementación es llamada "vista" o "consulta". Las relaciones derivadas son convenientes ya que expresan información de varias relaciones, como se muestra en la figura 2.1.

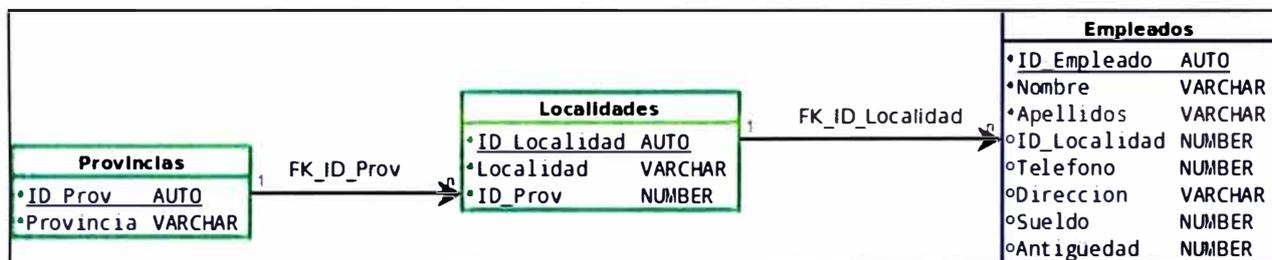


Fig. 2.1 Ejemplo de tablas relacionales.

2.2.3 DAS (Direct attached storage)

Se trata de dispositivos de almacenamiento directamente conectados a las máquinas, como es el caso de discos duros internos, cabinas de disco (en rack en o cualquier otro formato) conectadas directamente a un servidor, o unidades de cinta para backup. Suele basarse en tecnologías SCSI (Small computers system interface), FC (Fiber channel), e IDE. Esta arquitectura de almacenamiento, se relaciona principalmente con la época de los mainframe de IBM, y los miniordenadores UNIX, pues aquellos años se dotaba a estas máquinas de sus propios medios locales de almacenamiento.

2.2.4 NAS (Network attached storage)

Las soluciones NAS se basan en TCP/IP, con protocolos NFS o CIFS por encima. En consecuencia, en la actualidad, un dispositivo NAS será una máquina dedicada con una o más direcciones IP y además estará dotado de una conexión de alta velocidad a la red LAN.

Una arquitectura de almacenamiento NAS puede estar formada por múltiples dispositivos NAS geográficamente distribuidos. En cualquier caso, téngase en cuenta que un servidor NAS utilizará almacenamiento DAS o SAN (almacenamiento interno o almacenamiento externo), evidentemente. Claro está, que existen alternativas que integran soluciones NAS dentro de la propia infraestructura SAN.

2.2.5 SAN (Storage area network)

Es una red de área de almacenamiento, concebida para conectar servidores, arrays de discos y librerías de soporte. Principalmente, está basada en tecnología fiber channel y más recientemente en iSCSI. Su función es la de conectar de manera rápida, segura y fiable los distintos elementos que la conforman.

Una red SAN se utiliza para transportar datos entre servidores y recursos de almacenamiento. La tecnología SAN permite conectividad de alta velocidad, de servidor a storage, storage a storage, en la figura 2.2 se muestra un ejemplo.

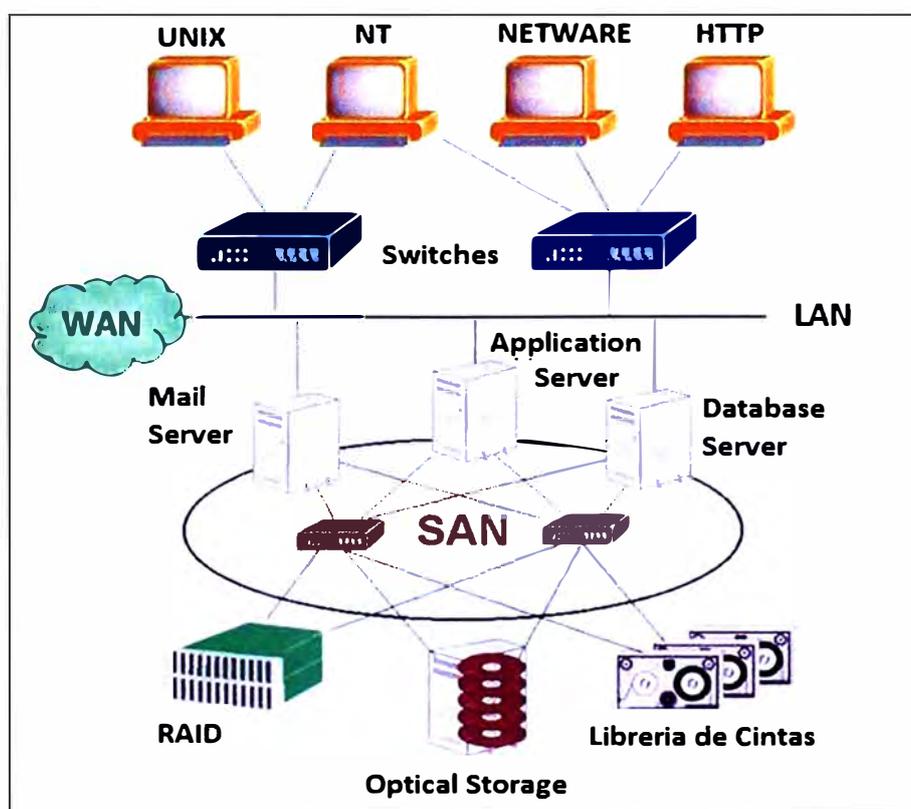


Fig. 2.2 Arquitectura estándar en una red SAN.

- Rendimiento: Permite acceso concurrente por dos o más servidores lo que proporciona un mejor rendimiento.
- Disponibilidad: Se puede hacer una copia exacta de los datos a una distancia de 10Km lo que las hace más seguras.
- Escalabilidad: Como las LAN/WAN puede usar muchas tecnologías. Lo que permite fácil reubicación, seguridad migración y duplicación de datos.
- Seguridad: La seguridad en las SAN ha sido desde el principio un factor fundamental, desde su creación se notó la posibilidad de que un sistema accediera a un dispositivo que no le correspondiera o interfiriera con el flujo de información,

2.2.6 World wide name (WWN)

Es un identificador único dentro de una red de almacenamiento SAN, es decir, es una dirección de 64-bit para identificar elementos en una red fiber channel (FC), similar a una dirección MAC.

Este parámetro posee términos derivados que nos ayudan a identificar cuantitativamente mejor los elementos en una red SAN:

- World wide port name (WWPN). Se refiere a un world wide name (WWN) asignado a un puerto de fibra en una red de almacenamiento SAN, similar a lo que es una dirección MAC en una red ethernet.
- World wide node name (WWNN). Se refiere a un world wide name (WWN) asignado a un nodo o dispositivo fiber channel (FC) de una red de almacenamiento SAN.

2.2.7 Zoning (Zonificación)

Es una tecnología que consiste en que un grupo de elementos se aislen del resto para evitar problemas de visibilidad, la zonificación puede llevarse a cabo por hardware, software o ambas, siendo capaz de agrupar por puerto o por WWN, una técnica adicional se implementa a nivel del dispositivo de almacenamiento que es la presentación, consiste en hacer que una LUN (Logical unit number) sea accesible sólo por una lista predefinida de servidores o nodos. Los conmutadores o switches fiber channel (FC), permiten segmentar sus puertos en diferentes zonas, de forma similar a como ocurre con las VLAN en los conmutadores o switches ethernet, de tal modo, que cada dispositivo sólo podrá comunicarse con el resto de dispositivos de su zona. Es importante tener en cuenta que cada puerto de fibra puede ser miembro de múltiples zonas.

Existen principalmente dos métodos de zoning, por hardware y por software. Así, el zoning por software (Soft zoning) restringe el acceso por nombre, sin embargo, cualquier

servidor podrá acceder a cualquier dispositivo por su dirección de red. Por el contrario, el zoning por hardware (Hard zoning) restringe las comunicaciones en los switches a través de filtrado de tramas, resultando mucho más seguro que el anterior.

2.2.8 Masking

Restringe a servidores o hosts puedan acceder a una determinada LUN (Disco virtual) en una cabina de almacenamiento. En consecuencia, podemos ver el LUN masking, como una configuración o proceso de autorización y seguridad dentro de la red de almacenamiento SAN, y en particular, en las cabinas de almacenamiento, que trabaja en el Nivel 4 del protocolo fiber channel.

2.2.9 Fiber channel protocol (FCP)

El canal de fibra (Fiber channel) es una tecnología de red utilizada principalmente para redes de almacenamiento, disponible primero a la velocidad de 1 Gbps y posteriormente a 2, 4 y 8 Gbps. Actualmente se ha convertido en el tipo de conexión estándar para redes de almacenamiento de ámbito empresarial por las prestaciones inherentes al protocolo y su facilidad de uso. Fiber channel es un protocolo multicapa, en la figura 2.3 se muestra los siguientes niveles que lo integran.

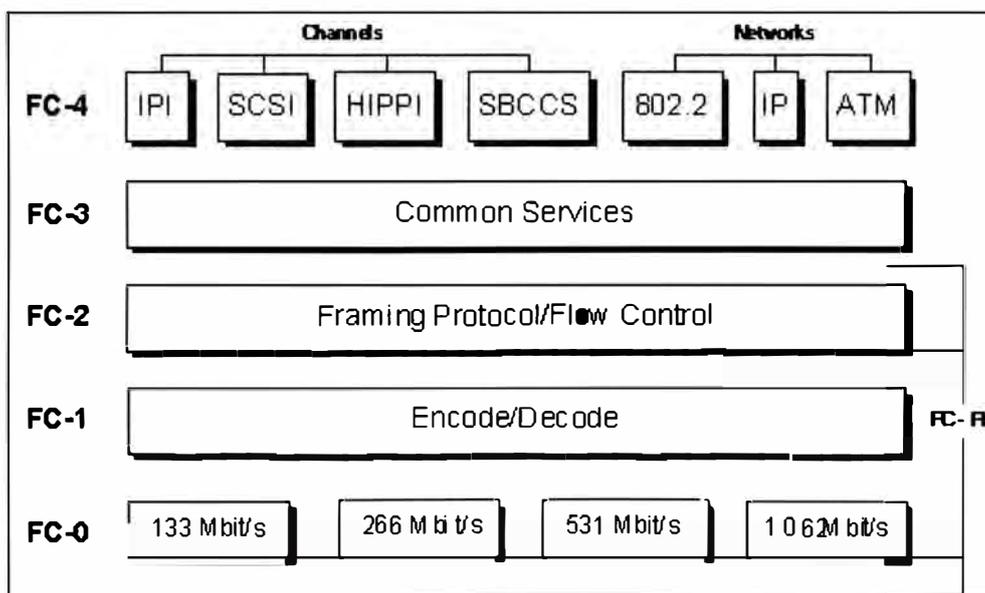


Fig. 2.3 Modelo de OSI – Protocolo fiber channel.

- FC0, capa física (Physical layer). Cables, conectores, etc.
- FC1, capa de enlace (Data link layer). Realiza la codificación y decodificación 8b/10b.

- FC2, capa de red (Network layer). Es el corazón de fiber channel. Define los principales protocolos de fiber channel.
- FC3, capa de servicios comunes (Common services layer). Puede implementar funcionalidades como la encriptación y RAID.
- FC4, capa de mapeo de protocolos (Protocol mapping layer). Es la capa, en la que otros protocolos como SCSI son encapsulados para su entrega a FC2.

En el canal de fibra se definen también los siguientes puertos:

- E_port, es la conexión entre dos switches del canal de fibra. También conocida como puerto de expansión, cuando dos E_port entre dos switches forman un enlace, ese enlace se denomina enlace de "Inter switch" o ISL.
- EX_port, es la conexión entre un router de canal de fibra y un switch de canal de fibra. En el extremo del switch, el puerto es como el de un E_port, pero en el extremo del router es un EX_port.
- F_port, es una conexión de medios en una topología conmutada. Un puerto F_port no se puede utilizar para un bucle de dispositivo.
- FL_port, es la conexión de medios en un bucle público en una topología de anillo arbitrado. También conocido como puerto de bucle. Nótese que un puerto de switch puede convertirse automáticamente en un F_port o un FL_port dependiendo de qué se esté conectando.
- G_port, o puerto genérico en un switch puede operar como E_port o F_port.
- L_port, es el término genérico utilizado para cualquier tipo de puerto de bucle, NL_port o FL_port, también conocido como puerto de bucle.
- N_port, es la conexión de nodo de los servidores o dispositivos de almacenamiento en una topología conmutada. También se conoce como puerto de nodo.
- NL_port, es la conexión de nodo de los servidores o dispositivos de almacenamiento en una topología de anillo arbitrado. También conocido como puerto de bucle de nodo.
- TE_port, es un término utilizado para múltiples puertos E_port unidos juntos para crear un ancho de banda mayor entre switches. También conocidos como puertos de expansión trunking.

2.2.10 Topologías fiber channel (FC)

Esta tecnología ofrece tres posibles topologías como se muestra a continuación:

- Punto a punto (FC-P2P), dos dispositivos se conectan el uno al otro directamente.

Es la topología más simple, con conectividad limitada a dos elementos.

- Anillo arbitrado (FC-AL), todos los dispositivos están en un bucle o anillo, similar a una red token ring. El añadir o quitar un elemento del anillo hace que se interrumpa la actividad en el mismo. El fallo de un dispositivo hace que se interrumpa el anillo. Existen concentradores de canal de fibra que conectan múltiples dispositivos entre sí y que pueden puentear los dispositivos que han fallado. Un anillo también se puede hacer conectando cada puerto al siguiente elemento formando el anillo. A menudo, un anillo arbitrado entre dos dispositivos negociará para funcionar como conexión punto a punto, pero ese comportamiento no es requerido.
- Medio conmutado (FC-SW), todos los dispositivos o bucles de dispositivos se conectan a conmutadores de canal de fibra. Conceptualmente son similares a las modernas implementaciones ethernet. La figura 2.4 muestra los tipos de topologías posibles.

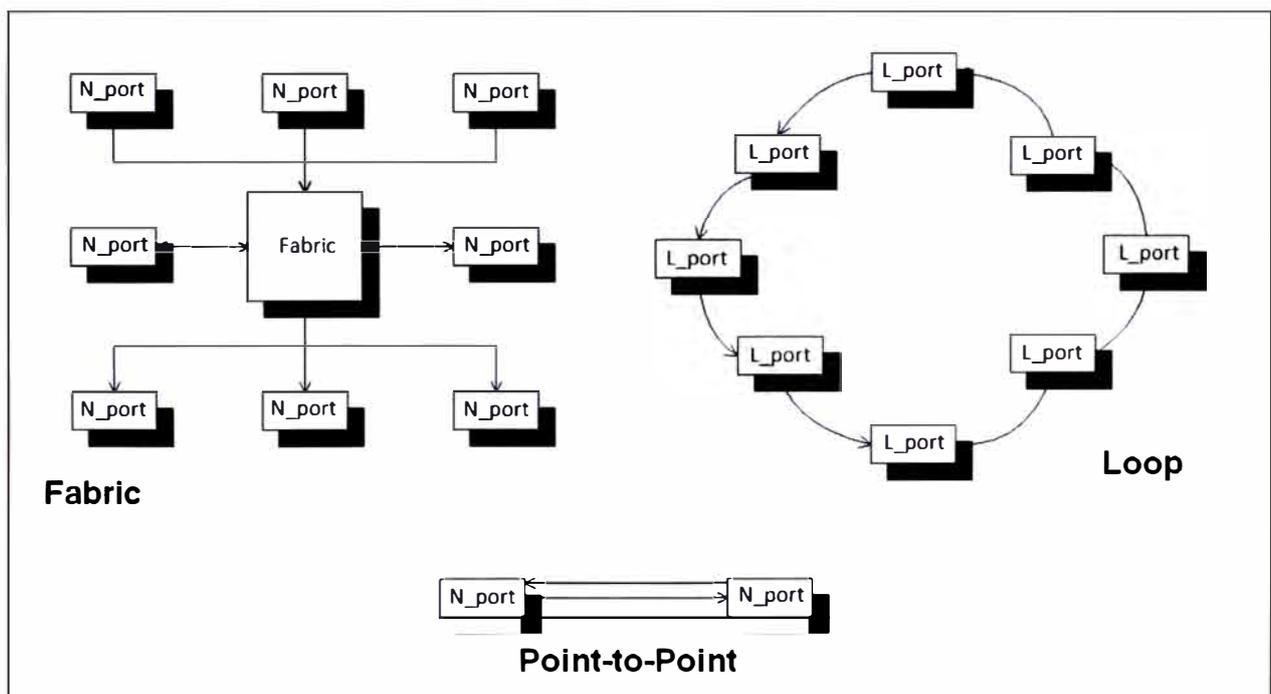


Fig. 2.4 Topologías de redes fiber channel.

2.2.11 HBA (Host bus adapters)

Se refiere a dispositivos donde se conectan otros dispositivos IDE, SCSI, canal de fibra y SATA, pero también se suele utilizar el mismo término para los dispositivos que se conectan a sistemas ethernet, fire wire y USB. Recientemente, la llegada del iSCSI ha dado lugar a HBAs vía ethernet, que se diferencian de las tarjetas de red porque incluyen hardware dedicado para iSCSI.

Cada HBA tiene un identificado único WWN y que le sirve al switch del canal de fibra para identificar las tarjetas HBA que tiene conectadas.

2.2.12 SCSI (Small computer system interface)

Es una interfaz estándar para la transferencia de datos entre distintos dispositivos del bus de la computadora.

2.2.13 iSCSI (Internet SCSI)

Es un estándar que permite el uso del protocolo SCSI sobre redes TCP/IP. iSCSI es un protocolo de la capa de transporte del modelo OSI. iSCSI utiliza TCP/IP para sus transferencias de datos, al contrario que otros protocolos de red diseñados para almacenamiento, como por ejemplo el canal de fibra. Solamente requiere una simple y sencilla interfaz ethernet o cualquier otra red compatible TCP/IP para funcionar.

2.2.14 LUN (Número de unidad lógica)

Es un volumen lógico o una unidad independiente que pueden residir en uno o más dispositivos de almacenamiento simple o conjunto. Son expuestas al mundo exterior con un esquema de direcciones, representa al host como número LUN. Las LUN son los discos o volúmenes que verán los servidores de una empresa, el término LUN es originario del protocolo SCSI, una red de almacenamiento SAN utiliza el protocolo SCSI, trasportándolo a través de fiber channel.

2.2.15 NPIV (N_Port ID virtualization)

NPIV es una utilidad de canal de fibra que permite que un adaptador de canal de fibra tenga muchos ID de puerto N. Cada puerto N tiene una identidad única WWN de puerto y WWN de nodo en la SAN y se puede utilizar para la delimitación de zonas o el enmascaramiento de LUN, la figura 2.5 grafica esta técnica.

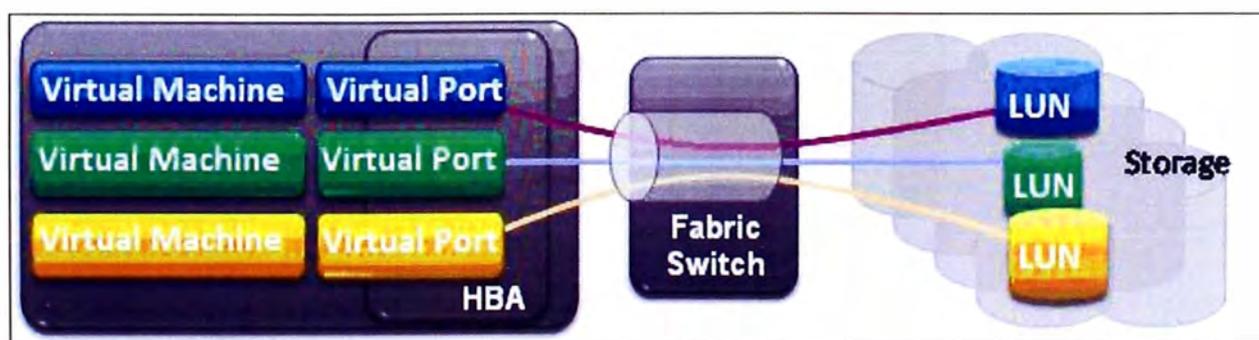


Fig. 2.5 Técnica de virtualización NPIV.

2.2.16 RAID (Redundant array of independent disks)

Hace referencia a un sistema de almacenamiento que usa múltiples discos duros entre los que se distribuyen o replican los datos. Dependiendo de su configuración los beneficios de un RAID respecto a un único disco son los siguientes: Mayor integridad, mayor tolerancia a fallos, mayor rendimiento (Throughput) y mayor capacidad. En sus implementaciones originales, su ventaja clave era la habilidad de combinar varios dispositivos de bajo coste y tecnología más antigua en un conjunto que ofrecía mayor capacidad, fiabilidad, velocidad o una combinación de éstas, que un solo dispositivo de última generación y coste más alto.

- RAID 1, crea una copia exacta (Espejo) de un conjunto de datos en dos o más discos.
- RAID 2, divide los datos a nivel de bits en lugar de a nivel de bloques y usa el código de Hamming para la corrección de errores, figura 2.6.

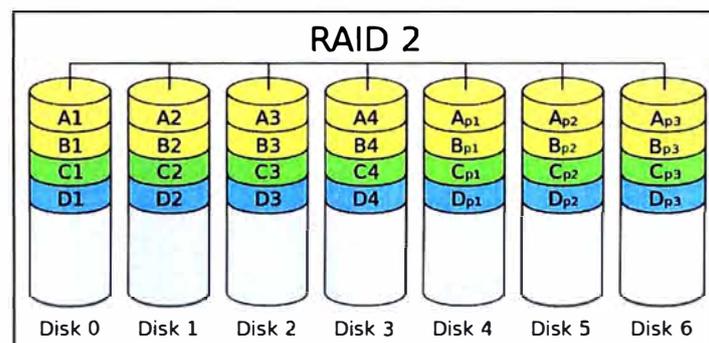


Fig. 2.6 Arreglo de disco – RAID 2.

- RAID 3, usa división a nivel de bytes con un disco de paridad dedicado, figura 2.7.

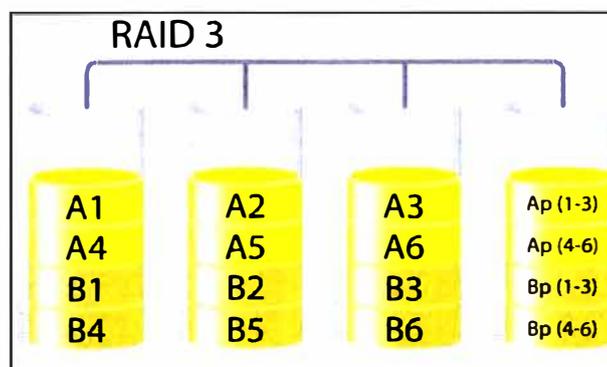


Fig. 2.7 Arreglo de disco – RAID 3.

- RAID 4, usa división a nivel de bloques con un disco de paridad dedicado.

Necesita un mínimo de 3 discos físicos, figura 2.8.

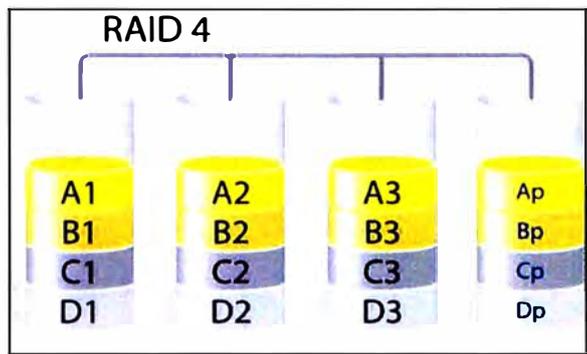


Fig. 2.8 Arreglo de disco – RAID 4.

- RAID 5, es una división de datos a nivel de bloques distribuyendo la información de paridad entre todos los discos miembros del conjunto, figura 2.9.

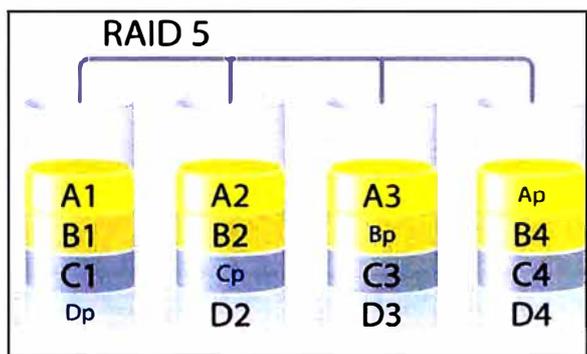


Fig. 2.9 Arreglo de disco – RAID 5.

- RAID 1+0, a veces llamado RAID 10, es parecido a un RAID 0+1 con la excepción de que los niveles RAID que lo forman se invierten, figura 2.10.

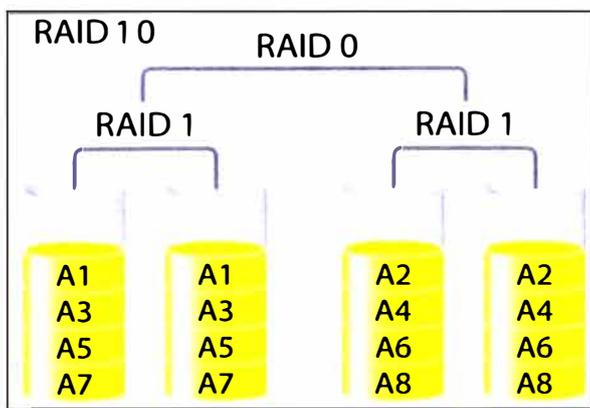


Fig. 2.10 Arreglo de disco – RAID 10.

2.2.17 Virtual switching system (VSS)

VSS es el sistema de red de la tecnología de virtualización que une varios Cisco Catalyst 6500 series switches en un conmutador virtual, aumentando la eficiencia operativa, impulsando las comunicaciones sin escalas, y sistema de mejora de la capacidad de ancho de banda de 1,4 Tbps. El VSS se compone de lo siguiente:

- Miembros de VSS: Cisco Catalyst 6500 series switches desplegados con el supervisor virtual switching 720 10GE.
- Virtual switching link (VSL): Puede unir 10 conexiones gigabit ethernet y soporta hasta 8 miembros o switches.
- Las 10 conexiones gigabit ethernet del enlace VSL tiene que ser de uno o más módulos supervisores de los switches.

VSS permite una funcionalidad sin precedentes y la disponibilidad deseada como se muestra en la gráfica 2.11.

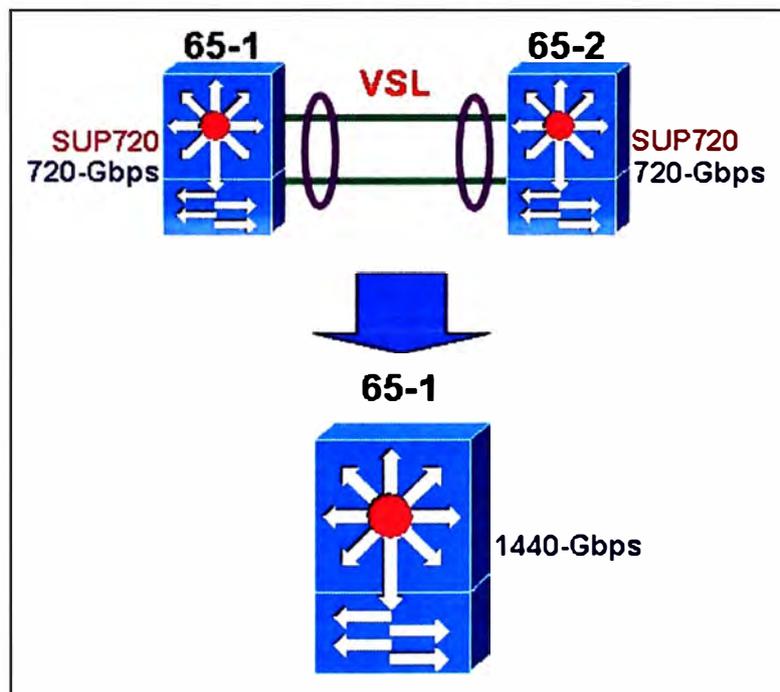


Fig. 2.11 Técnica de virtualización en switches Cisco 6500.

2.2.18 Linux channel bonding

El channel bonding o unión de interfaces de red consiste en simular un dispositivo de red con gran ancho de banda uniendo varias tarjetas de red independientes, de manera que las aplicaciones sólo verán un interfaz de red. Con el channel bonding conseguimos

varias cosas:

- Mayor ancho de banda, el BW de la interfaz virtual será la suma de los anchos de banda de las interfaces reales.
- Balanceo de carga, tendremos balanceo de carga del tráfico de red entre todas las interfaces reales.
- Redundancia, si falla una tarjeta de red los datos irán sólo por las que estén en buen estado.

2.2.19 Oracle database server

Para comprender mejor la base de datos Oracle es necesario conocer su arquitectura. Cuando hablamos de una base de datos no sólo nos referimos a los datos físicos, sino también a los objetos físicos, de memoria y de proceso, detallados a continuación:

a. Componentes de la base de datos

- Archivos de datos (database files), contiene toda la información de la base de datos, datos de usuario y de sistema. Antes de introducir la información se crea un espacio, denominado tablespaces que contiene toda la información. Esto ayuda a organizar la información contenida, por ejemplo, uno para los datos de la aplicación de almacén, otro para la de nóminas. Cada tablespace consta de uno más archivos de disco. Un archivo de datos sólo puede pertenecer a un único tablespace. Cuando Oracle se instala se crea automáticamente los siguiente: System (SYSTEM01.DBF) el cual almacena la información para gestionarse a sí misma, por ejemplo, el diccionario de datos, users (USERS01.DBF) con la información personal de los usuarios, lugar donde el DBA nos deja almacenar las tablas para realizar pruebas, temp (TEMP01.DBF) donde Oracle almacena las tablas temporales y UNDOTBS1 donde Oracle guarda la información de deshacer, para almacenar la imagen anterior de los datos antes de permitir actualizaciones.
- Registros de rehacer o Redo Log, se trata de archivos de datos en los que Oracle registra todas las transacciones o modificaciones (Insert, update y delete) que se producen. Esto permite recuperar la base de datos si existen problemas. Suele haber varios registros de rehacer que están almacenados en los archivos: RED001.LOG, RED002.LOG y RED003.LOG. Un registro de Redo_Log contiene la identificación suficiente para reconstruir el dato modificado.
- Archivos de control (Control files), contienen información sobre los archivos asociados

con una base de datos Oracle. Todas las modificaciones importantes que se hagan en la estructura se registran en el archivo de control, manteniendo así la integridad. Si se daña un disco de control, se podría restaurar utilizando la copia intacta. Los archivos de control creados en la instalación son: CONTROL01.CTL, CONTROL02.CTL y CONTROL03.CTL, además estos archivos de control contienen la siguiente información:

Información de arranque y parada (Nombre de los archivos de la base de datos y Redo_Log), Información sobre los checkpoints (Fecha de creación) y nombre de la base de datos, la figura 2.12 grafica la arquitectura global de las instancias Oracle.

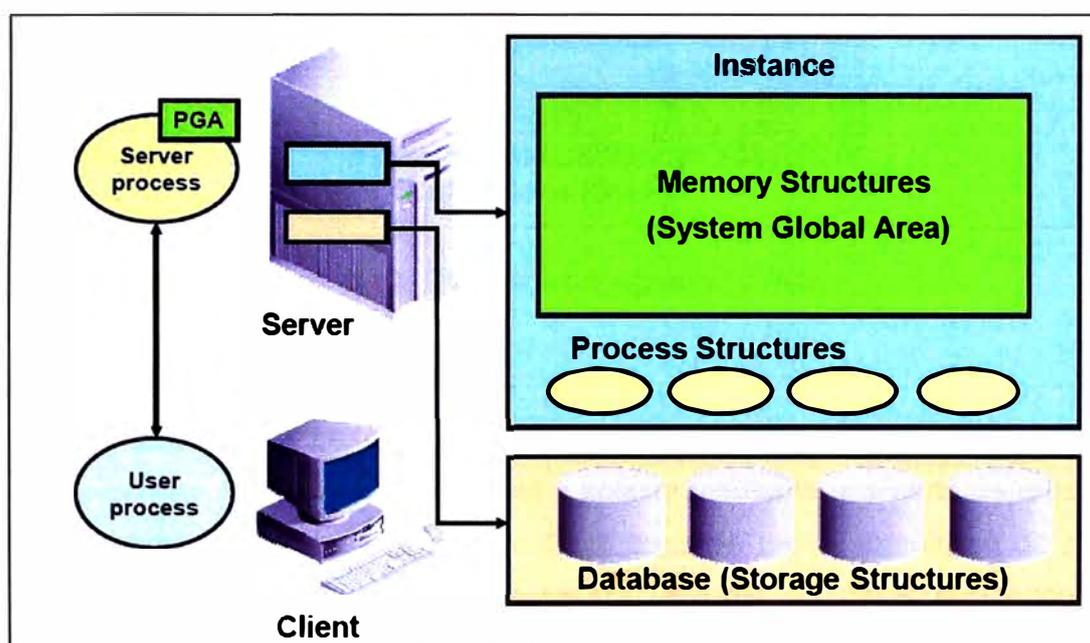


Fig. 2.12 Componentes PGA y SGA Oracle.

b. Memorias SGA y PGA

Son la memoria de los procesos de usuario (Cliente) y servidor, estos se comunican consigo mismos y entre ellos por medio de estructuras de memoria. Oracle utiliza dos tipos de estructuras de memoria: el SGA (System global area) y el PGA (Program global area), la figura 2.13 muestra los procesos existentes.

- Área global del sistema (SGA): Es un grupo de estructuras de memoria que sirven para almacenar los datos de la base de datos que se han consultado recientemente.
- Área global del programa (PGA): Es la zona de memoria utilizada por un único proceso de usuario de Oracle, que contiene datos e información del proceso. La memoria PGA no se comparte.

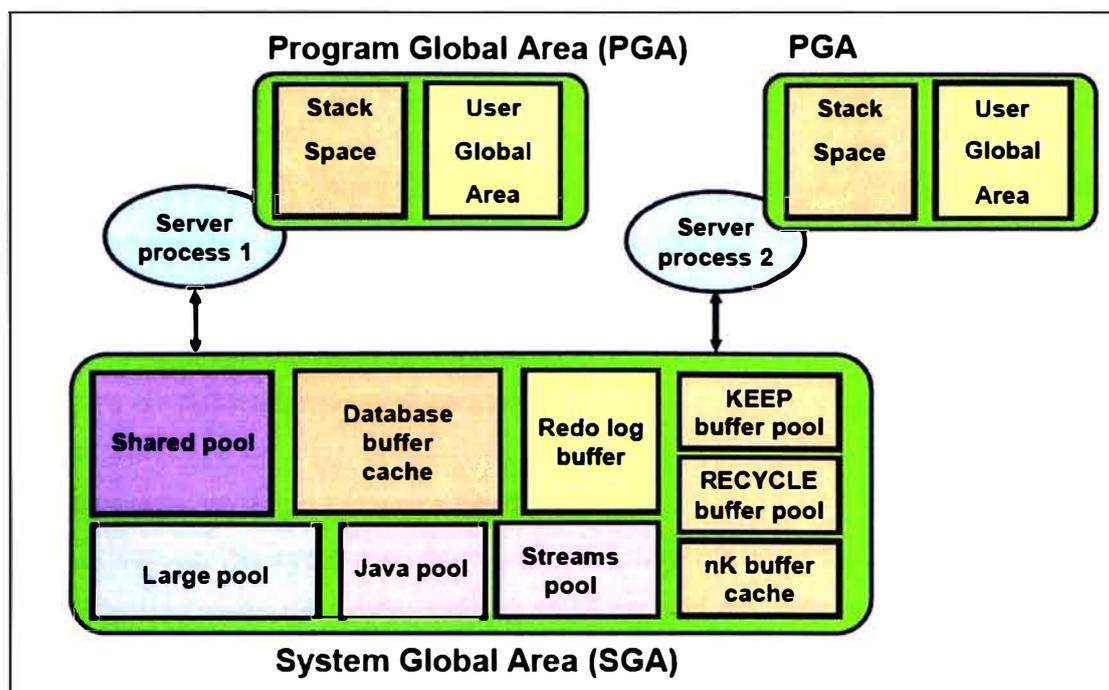


Fig. 2.13 Procesos background Oracle.

c. Procesos background Oracle

Procesos de soporte de la base de datos las relaciones entre las estructuras físicas y de memoria de la base de datos se mantienen y aplican mediante una serie de procesos de soporte. Dependiendo de la configuración de la base de datos, el número de procesos varía. La gestión de estos procesos la realiza la base de datos con poco trabajo administrativo. Hay un conjunto de procesos del servidor que ayudan a la base de datos a funcionar:

- Escritor de bases de datos DBWR (Database writer): Este proceso es el responsable de gestionar el contenido del buffer de datos de la SGA.
- Punto de comprobación o control CKPT (Checkpoint): Este proceso provoca que el DBWR escriba en los archivos de datos todos los bloques que se hayan modificado desde el último punto de control y que actualice las cabeceras de los archivos de datos y los archivos de control para registrar el punto de control.
- Escritor de registros LGWR (Log writer): Gestiona la escritura del contenido del buffer del registro de rehacer de la SGA a los archivos de Redo_Log. Es el único proceso que escribe en los archivos de registro de rehacer y el único que lee los buffers de este registro.
- Supervisor del sistema SMON (System monitor): Es un proceso obligatorio que se ocupa de todas las recuperaciones que sean precisas durante el arranque de la base de

datos.

- Supervisor de proceso PMON (Process monitor): Realiza una limpieza al terminar la ejecución de los procesos
- Archivador ARCH (Archiver): Archiva en disco o cinta una copia de los Redo_Log cuando están llenos para una posible recuperación por fallo de disco o por fallo de sistema.
- Recuperador RECO (Recoverer): Recupera transacciones distribuidas dudosas; se usa en bases de datos Oracle distribuidas.

2.2.20 Automatic storage management (ASM)

ASM es a la vez un file system y un volumen manager integrado, creado específicamente para archivos de bases de datos. Oracle.ASM es una solución muy interesante que nos ofrece la alta performance de los rawdevices y la facilidad de administración de un file system.

ASM permite dividir todo el storage disponible en diskgroups. Luego uno administra los diskgroups y ASM automatiza la ubicación de los archivos dentro de los diskgroups. Cuando se utiliza ASM en las sentencias SQL de creación de estructuras (Tablespaces, control files y archive log files) se especifica la ubicación en términos de diskgroups. ASM crea y administra los archivos dentro de los diskgroups sin necesidad de intervención humana. Además, capacidades adicionales de mirroring y striping, dividiendo los archivos en extensiones de 1Mb. y distribuyéndolos entre los discos que componen un diskgroup. Estas capacidades adicionales optimizan el rendimiento del sistema y la utilización de disco, haciendo que el tuning manual de input output sea prácticamente innecesario.

La forma en que ASM administra el mirroring (espejado) es incluso más flexible que el espejado vía sistema operativo ya que ASM permite especificar el nivel de redundancia a nivel de archivo. De este modo, dos archivos pueden compartir el mismo diskgroup estando uno espejado y el otro no. Oracle administra el espejado a nivel de extensión; si un archivo está espejado, cada extensión de dicho archivo tendrá una o más copias dependiendo del nivel de redundancia configurado para dicho archivo.

Las copias espejadas siempre residen en diferentes discos de distintos diskgroups. ASM soporta tres opciones de espejado a nivel de archivo:

- 2-way mirroring: Cada extensión tiene una copia.

- 3-way mirroring: Cada extensión tiene dos copias.
- Desprotegido: Sin espejado, se utiliza cuando el espejado es provisto por el propio subsistema de discos.

ASM se implementa como una instancia Oracle, con su propia SGA y procesos background. Una configuración "single instance" de ASM puede dar servicio a una o más bases de datos también single instance, todo residente en un solo servidor. Cada diskgroup ASM puede ser compartido por todas las bases de datos del servidor.

En un ambiente "clusterized", cada nodo corre una instancia ASM y se comunican entre ellas mediante una relación peer to peer. Esta arquitectura es válida tanto para un entorno RAC o no RAC, la figura 2.14 grafica la diferencia entre un sistema de archivos clásico y el moderno sistema de administración de discos ASM.

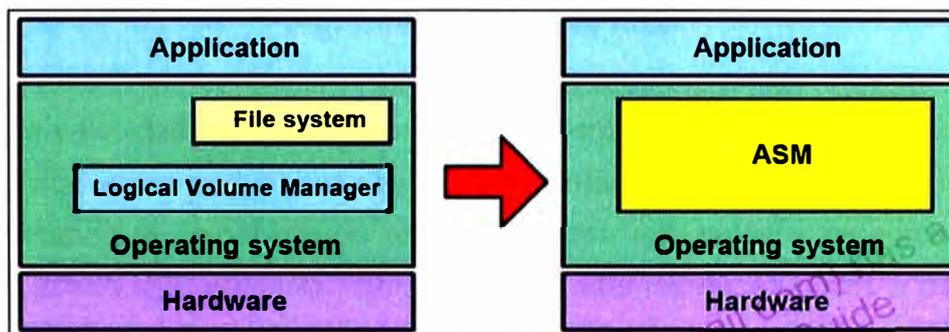


Fig. 2.14 Automatic storage management.

2.2.21 Oracle real application clusters

El software real application clusters (RAC) y una colección de hardware del mismo tipo utiliza el potencial de cada componente para crear un potente entorno de computación. RAC provee escalabilidad y alta disponibilidad debido a que se puede agregar nodos al sistema con ello se proporciona la recuperación de falla si los componentes del nodo fallaran. En un entorno de RAC todas las instancias activas pueden concurrentemente ejecutar transacciones contra una base de datos compartida. RAC coordina cada acceso de las instancias para compartir los datos con lo cual proporciona consistencia e integridad de los datos.

RAC proporciona disponibilidad escondiendo las fallas al usuario final sin interrumpir el acceso a los datos. Transparent application failover (TAF) es una aplicación que re direcciona transparentemente las consultas de los clientes a la base de datos mediante los nodos disponibles, la figura 2.15 muestra la arquitectura RAC.

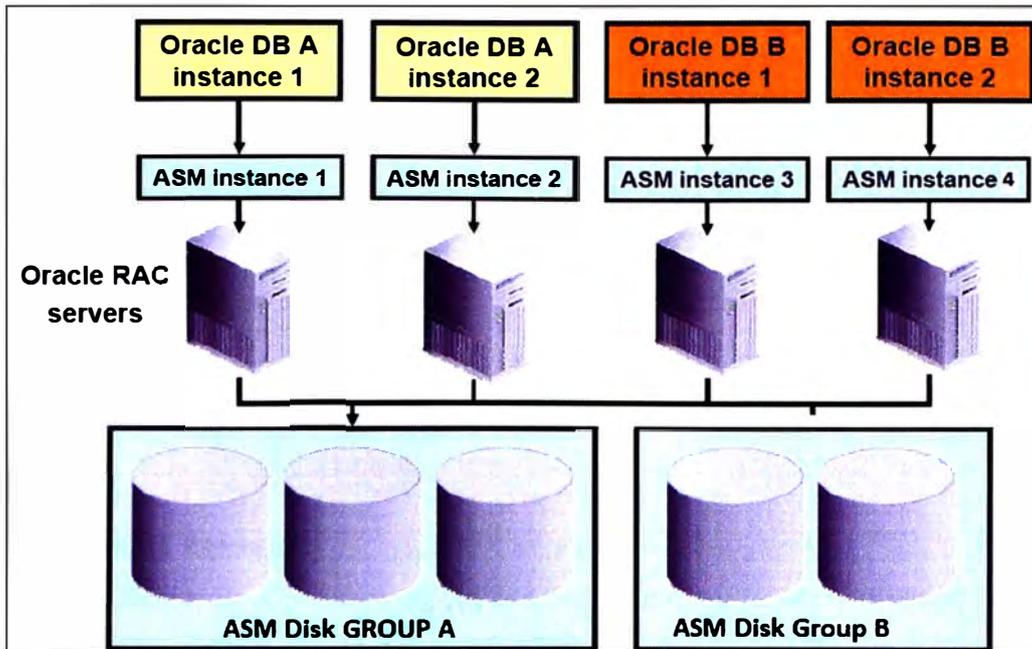


Fig. 2.15 Arquitectura ASM.

RAC usa alta velocidad para el componente de interproceso de comunicación (IPC) para la comunicación entre nodos. Usualmente esta interconexión usa un ancho de banda grande, facilitando la comunicación entre los nodos enlazados y proporcionando baja latencia. Se puede usar ethernet, fiber distributed data interface (FDDI), u otro hardware para su interconexión. El IPC define protocolos e interfaces requeridas en un entorno RAC para transferir mensajes entre las instancias. Los mensajes son la unidad Fundamental de comunicación entre las interfaces. El IPC es funcionalmente construido sobre modelo de mensajes en colas asincrónico.

Generalmente, cada nodo en un clúster de base de datos tiene uno o más CPU. Los nodos con múltiples CPU son configurados típicamente para compartir la memoria principal. RAC requiere que todos los nodos tengan acceso simultáneo a los discos compartidos para dar a las instancias acceso concurrente a la base de datos.

La implementación del subsistema de discos compartidos está basada en su sistema operativo o solución de hardware.

Las configuraciones para RAC son típicamente uniformes. Esto significa que la sobrecarga para cada nodo en el clúster al acceder la memoria es la misma.

Una base de datos RAC tiene los mismos procesos que una sola instancia de base de datos Oracle tales como monitor de proceso (PMON), escritor de base de datos (DBWRn), escritor de log (LGWR), etc.

CAPITULO III

METODOLOGIA PARA LA SOLUCION DEL PROBLEMA

Para encontrar una solución factible y real, es necesario conocer sobre las tecnologías existentes.

3.1 Alternativas de solución

Es importante realizar una revisión de los productos existente en el mercado, y así poder tener una idea clara del alcance de nuestro diseño. Además es importante tener claro que casi todas las nuevas tecnologías y soluciones han sido desarrolladas por las grandes empresas.

3.1.1 Cisco Systems

Cisco Systems es una empresa global con sede en San José,¹ (California, Estados Unidos), principalmente dedicada a la fabricación, venta, mantenimiento y consultoría de equipos de telecomunicaciones tales como:

- Dispositivos de conexión para redes informáticas: Routers, switches;
- Dispositivos de seguridad como corta fuegos y concentradores para VPN;
- Productos de telefonía IP como teléfonos y el CallManager (una PBX IP);
- Software de gestión de red como CiscoWorks, y
- Equipos para redes de área de almacenamiento.

a. Cisco Catalyst 3130 para Dell

El switch Catalyst 3130 de Cisco es la solución de red de nueva generación para entornos de servidores. Está basado en el hardware y el software de la empresa Cisco líderes del sector, el switch Catalyst 3130 de Cisco se ha diseñado con tecnologías exclusivas destinadas de forma específica a satisfacer las exigencias de la infraestructura de aplicaciones basada en servidores modelo chasis. En particular, este switch está diseñado para proporcionar un rendimiento alto y escalable con una conectividad de gran resistencia, además de apoyar las iniciativas de TI en curso para reducir la complejidad y

el coste total de propiedad de la infraestructura de servidor.

El switch Catalyst 3130 de Cisco para Dell M1000E dispone de dos configuraciones. Una configuración permite a los clientes conectar el switch a la estructura de red existente mediante enlaces ascendentes gigabit Ethernet, mientras que la otra configuración permite conectarlo a la estructura de red existente mediante enlaces ascendentes gigabit y 10 gigabit ethernet, la figura 3.1 muestra el modelo a usar.

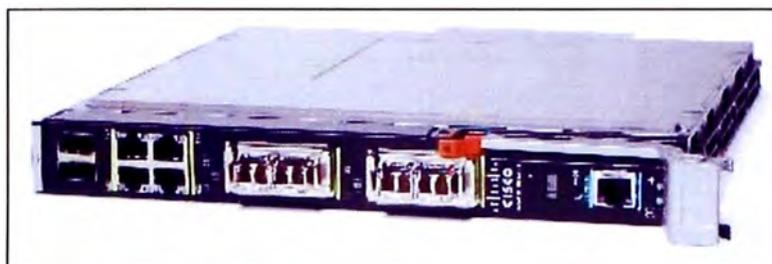


Fig. 3.1 Cisco Catalyst 3130 para Dell.

b. Cisco MDS 9509 multilayer director

Cisco MDS 9509 director son switches SAN de alto performance, estas diseñados para entregar los más altos requerimientos a las redes SAN de las empresas. Ahora, las redes SAN virtuales (VSAN) se utilizan para brindar una mayor seguridad en la red. Tareas específicas, como el aprovisionamiento para almacenamiento de datos dentro de la nueva red SAN se realiza con una mayor flexibilidad. Los servidores y dispositivos de almacenamiento de datos se pueden añadir según sea necesario, sin el gasto adicional y el plazo de tiempo necesario para la instalación de un switch adicional de canal de fibra. Estas implementaciones traen numerosas ventajas, entre las que destacan:

- Mejora del uso del almacenamiento de datos
- Mejora de la gestión y fiabilidad del sistema
- Aprovisionamiento más rápido
- Mayor disponibilidad de los datos
- Reducción de los costes de gestión y de capital
- Gestión simplificada del almacenamiento de datos
- Mayor rendimiento y velocidad de la red
- Mejora de la integridad y seguridad de los datos
- Mejor posicionamiento para el crecimiento del almacenamiento de datos

Cisco MDS 9509 multilayer director proporciona hasta 224 puertos, con rendimiento de 2

Gigabits por segundo y un diseño de alta disponibilidad. Incluye capacidad de SAN virtual (VSAN) para la consolidación de SAN en islas de SAN virtuales en un solo equipo físico. Proporciona seguridad para el despliegue de grandes redes SAN para empresas. La figura 3.2 muestra el modelo a usar.



Fig. 3.2 Cisco MDS 9509 multilayer director

c. Cisco Catalyst 6509-E

La familia Catalyst 6500 ofrece soluciones de conmutación de alto rendimiento para redes empresariales y de proveedores de servicios. Diseñada para afrontar los crecientes requisitos de densidad gigabit, la alta disponibilidad y conmutación inteligente, la figura 3.3 muestra el modelo a usar.



Fig. 3.3 Cisco Catalyst 6509-E.

La familia Catalyst 6500 proporciona una amplia gama de soluciones, que posibilita la existencia tanto de extranets, intranets empresariales como internet para multimedia, datos de tareas críticas y aplicaciones de voz. Además de tener alta tasa de transmisión permite utilizar las nuevas tecnologías de virtualización como el virtual switching system (VSS) permitiendo duplicar la capacidad de transmisión.

d. Cisco adaptive security appliance ASA

Integra una combinación de tecnologías en una única plataforma para la detección proactiva de amenazas y ataques que asedian las redes empresariales.

Detiene los ataques antes de que entren en la red, controla la actividad y provee conectividad VPN flexible. Es una herramienta perfecta para todo tipo de redes corporativas, reduce el coste y la complejidad asociados a la instalación de productos con el mismo nivel de seguridad.

Con la serie 5500 de Cisco ASA se puede proteger de un modo más efectivo y eficiente las redes empresariales gracias a la combinación de la última tecnología proveniente de los PIX 500, la serie IPS 4200 y la serie VPN 3000 de Cisco Systems, añade además, el nuevo equipamiento Anti-X.

Cisco ASA 5510 provee de un alto rendimiento en servicios de Firewall y VPN, tres tarjetas de red integradas 10/100 y, opcionalmente, IPS con AIP-SSM o protección contra malwares con CSC-SSM, la figura 3.4 muestra los modelos existentes.



Fig. 3.4 Cisco adaptive security appliance ASA.

3.1.2 Dell

DELL es una compañía multinacional estadounidense establecida en Round Rock (Texas) que desarrolla, fabrica, vende y soporta a computadoras personales, servidores, switches de red, programas informáticos, periféricos y otros productos relacionados con la tecnología.

a. Power Edge M1000e

Es un servidor modelo chasis diseñado para obtener más eficiencia en materia de tiempo, alimentación, refrigeración, inversión y espacio. Es un diseño innovador, que maximiza la

flexibilidad, la eficiencia térmica y energética, así como la disponibilidad, el rendimiento y la manejabilidad de todo el sistema. El chasis integra las últimas tecnologías de administración, E/S y alimentación y refrigeración en un paquete modular fácil de usar. Esta diseñado desde un principio para admitir las generaciones presentes y futuras de tecnologías de servidores, almacenamiento, redes y administración, el Power Edge M1000e tiene la capacidad necesaria de ampliar su entorno para el futuro.

El Power Edge M1000e brinda una flexibilidad óptima en las arquitecturas de procesadores y chipsets de servidor. Las arquitecturas de servidores Intel y AMD están planeadas para su incorporación dentro de la infraestructura, mientras que los avanzados diseños eléctricos, mecánicos y de interfaces de software habilitan la compatibilidad con varias generaciones de servidores, la figura 3.5 muestra el modelo a usar.



Fig. 3.5 Enclosure Dell M1000e.

b. PowerVault TL2000

La biblioteca de cintas Power Vault TL2000 de Dell ofrece una automatización de cintas sencilla y asequible para las organizaciones cuyas necesidades de protección de datos han superado la tecnología de cargador automático. Es la solución para pequeñas y medianas empresas que requieren respaldo en cinta accesible, fácil de utilizar y automatizado. Contienen varias ranuras para almacenar medios de cinta de almacenamiento y utilizan un mecanismo robótico para mover los medios automáticamente y funcionar de manera independiente y sin supervisión. Para mantener los contenidos debidamente organizados, un lector de código de barras hace un inventario y un seguimiento de todas las cintas almacenadas en la Biblioteca, presentan una tecnología de ubicación óptica de cartuchos para la manipulación y el inventario precisos de los cartuchos.

PowerVault TL2000 son compactas, fáciles de instalar por el cliente y se conectan a un servidor o a una SAN, presenta un intuitivo panel de control de operador, junto con administración web remota, puede realizar la configuración y los diagnósticos, comprobar el estado, supervisar las cintas, la figura 3.6 muestra el modelo de librería de cinta a usar.

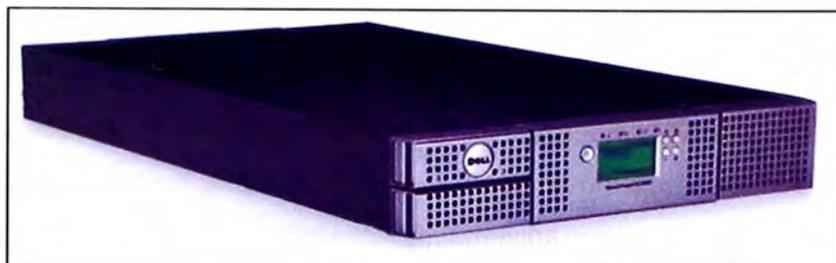


Fig. 3.6 Dell Power Vault TL2000.

c. Dell Power Edge R610

Los servidores Dell ofrecen sistemas de primera clase con una imagen común y una facilidad de usos inspirados en el cliente. El Dell Power Edge R610 saca partido de las características comunes del sistema Dell de forma que, cuando los administradores de TI se familiaricen con el funcionamiento de un sistema, habrán asimilado la lógica necesaria para administrar los servidores Dell de nueva generación. La disposición lógica de los componentes y la ubicación de las fuentes de alimentación facilitan la instalación y las implementaciones futuras. Incluyen una pantalla LCD gráfica e interactiva para supervisar el estado del sistema, emitir alertas y controlar la configuración de administración básica justo delante del servidor. Los clientes disponen de un medidor de alimentación de corriente eléctrica y un termómetro de temperatura ambiente integrados en el servidor, la figura 3.7 muestra el modelo a utilizar en el diseño.



Fig. 3.7 Dell Power Edge R610.

3.1.3 Hitachi

Hitachi es una empresa con bases en Chiyoda, Tokio, Japón. La compañía fue fundada en 1910 como una tienda de reparación de equipos eléctricos. Hitachi produce una gran

variedad de equipos electrónicos y proporciona productos para otras fábricas por ejemplo circuitos integrados y semiconductores.

a. Hitachi adaptable modular storage AMS2500

La nueva serie 2000 de adaptable modular storage supera los límites en eficiencia operativa y redefine el almacenamiento de gama media con la incorporación del primer controlador dinámico de cargas.

Esta nueva plataforma ostenta mejor rendimiento, conectividad, capacidad de ampliación, fiabilidad y facilidad de uso a clientes de medianas empresas, ofrece hasta 4 veces el rendimiento de las generaciones anteriores, además de consolidar el almacenamiento para conexiones iSCSI, NAS, y fiber channel en redes de áreas de almacenamiento (SAN). La línea Hitachi AMS series 2000 está compuesta por tres modelos: Hitachi AMS 2100, Hitachi AMS 2300 e Hitachi AMS 2500. Todos de sistemas de almacenamiento de la serie AMS 2000 satisfacen los estándares internacionales con una disponibilidad del 99,999 por ciento, la figura 3.8 muestra los equipos apilados en una rack.

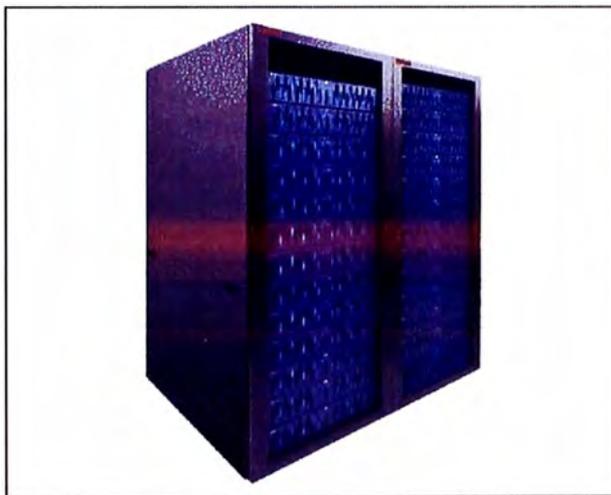


Fig. 3.8 Hitachi adaptable modular storage AMS2500.

3.1.4 Brocade Communications Systems

Brocade Communications Systems es una corporación multinacional americana especializada en redes de datos y productos de almacenamiento informático. La corporación extiende su cartera de productos a través de ethernet (LAN empresariales, WLAN) Interruptores, WAN (Internet) Routers, Switches SAN, controladores de entrega de aplicaciones, aplicaciones de seguridad de red ethernet, almacenamiento adaptadores de red y transceptores. La corporación fue fundada en 1995 y tiene su sede en San José, California, EE.UU. En marzo de 2012, la corporación tuvo la mayor cuota del mercado en

conmutadores SAN.

a. Brocade M5424

El conmutador Brocade M5424 junto al chasis Dell Power Edge M1000e ofrece soluciones sólidas para las implementaciones de redes de área de almacenamiento (SAN) de canal de fibra. Esta propuesta no sólo ayuda a simplificar y reducir la cantidad de componentes de hardware de las SAN necesarios para una implementación, sino que también mantiene la capacidad de ampliación, el rendimiento, la interoperabilidad y la manejabilidad de los entornos SAN tradicionales, la figura 3.9 muestra el equipo a usar.

- Ideal para redes de área de almacenamiento más amplias
- La integración de las funciones de conmutación de las redes de área de almacenamiento (SAN) permite reducir la complejidad y aumentar la manejabilidad de estas redes
- Hasta 16 puertos internos de estructuras y hasta 8 enlaces ascendentes de 2 Gb/4 Gb/8 Gb con detección automática

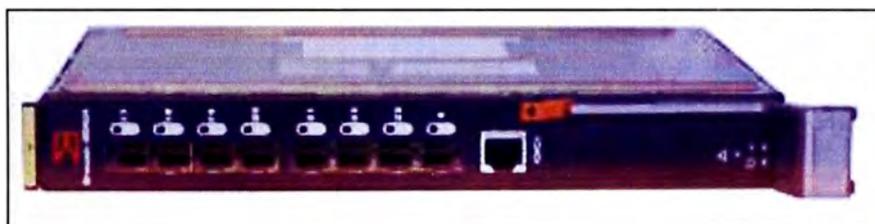


Fig. 3.9 Brocade M5424.

3.1.5 Fortinet

Fortinet es una empresa privada estadounidense, situada en Sunnyvale (California), que se dedica especialmente al diseño y fabricación de componentes y dispositivos de seguridad de redes (firewalls, UTM).

La compañía que fue fundada en el año 2000 por Ken Xie y desde entonces ha tenido una gran proyección en el mundo de la seguridad de las comunicaciones. Actualmente es la marca de referencia en sistemas de seguridad UTM, habiendo superado a Cisco y Checkpoint en su lucha por este mercado.

a. Fortigate 311B

Ofrece hasta 12 Gbps de rendimiento de firewall y 14 interfaces GbE (con módulo opcional AMC). La integración con Forti Manager y Forti Analyzer simplifica la gestión de

la seguridad, la elaboración de informes y análisis, mientras que los gastos de operación se simplifican por su facilidad de uso.

El FortiGate-311B proporciona registro de eventos locales o remotos. En un ambiente remoto se utiliza servidores de almacenamiento, además la alta densidad de los puertos es compatible con múltiples modos de seguridad para las distintas situaciones. De este modo los usuarios cumplen con las políticas de seguridad informática de red establecidos por las empresas, la figura 3.10 muestra al Fortigate 311B.



Fig. 3.10 Fortigate 311B.

3.1.6 Oracle Corporation

Oracle Corporation es una de las mayores compañías de software del mundo. Sus productos van desde bases de datos hasta sistemas de gestión. Cuenta además, con herramientas propias de desarrollo para realizar potentes aplicaciones, como:

- Oracle Designer, orientado a la creación de aplicaciones más escalables,
- Oracle JDeveloper, entorno de desarrollo para programadores.
- Oracle Developer suite, orientado para aplicaciones en lenguaje Java.

La compañía es el proveedor líder mundial en desarrollo de software, con la adquisición de la compañía SUN le otorgó un papel de liderazgo en el campo del software.

a. Oracle database enterprise edition 11gR2

Oracle es un sistema de gestión de base de datos objeto-relacional ORDBMS. La figura 3.11 muestra el logo del producto. Sus principales características son:

- Alto soporte de transacciones.
- Estabilidad.
- Escalabilidad.
- Soporte multiplataforma.



Fig. 3.11 Oracle database enterprise edition 11gR2.

b. Oracle Real Application Clusters (RAC)

Oracle RAC permite que múltiples computadoras ejecuten el software de SGBD de Oracle simultáneamente mientras acceden a una base de datos individual. Esto se llama una base de datos en cluster. La figura 3.12 muestra el logo del producto.

En una base de datos de Oracle no RAC, una base de datos individual es accedida por una instancia individual. La base de datos se considera la colección de ficheros de datos, ficheros de control, y ficheros redo log localizados en disco. La instancia se considera la colección de procesos del sistema operativo y memoria relacionada de Oracle que están ejecutándose en el computador. En Oracle RAC, dos o más computadoras (cada una con una instancia) acceden concurrentemente a una base de datos individual. Esto permite que una aplicación o usuario se conecte a alguno de los computadores y tenga acceso a los mismos datos.

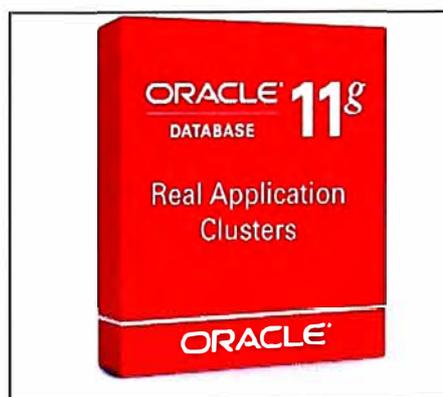


Fig. 3.12 Oracle real application clusters (RAC).

c. Oracle enterprise linux

Oracle linux, conocido anteriormente como Oracle enterprise linux, es una distribución de GNU/Linux basada en Red Hat, re empaquetada y distribuida por Oracle, disponible bajo

la GNU (GPL), desde finales de 2006 Oracle linux puede ser descargado gratuitamente, la figura 3.13 muestra el logo del producto.

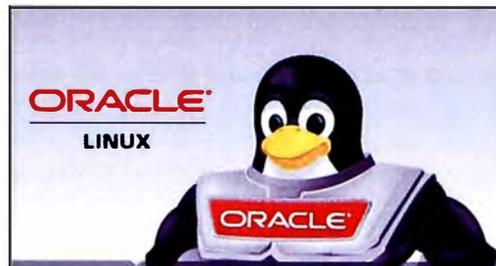


Fig. 3.13 Oracle enterprise linux.

3.2 Solución del Problema

En los apartados anteriores se detalla las tecnologías existentes actualmente en el mercado, el cual servirá como base para diseñar una arquitectura de red, además para entender la arquitectura se explica a continuación el funcionamiento interno de muchos de los dispositivos involucrados.

3.2.1 Descripción del enclosure Power Edge M1000e

La base de datos RAC Oracle 11gR2 está montada sobre una plataforma redundante usando como base el chasis Dell, este equipo es la base para soportar la conectividad interna de los nodos del clúster, la figura 3.14 muestra algunas características físicas.

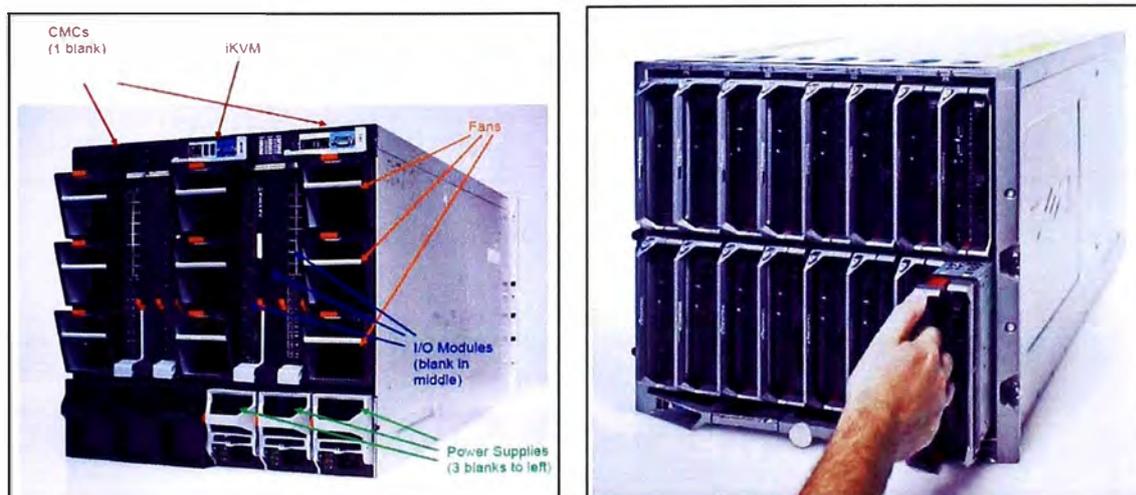


Fig. 3.14 Chasis modular - Enclosure Power Edge M1000e.

a. Módulos de Servidores para el Chasis Dell

El gabinete Power Edge M1000e admite hasta 16 módulos del servidor de media altura,

cada uno de los cuales ocupa una ranura a la que se accede en el frente del gabinete. Los módulos del servidor que no poseen un factor de forma estándar de media altura también se admitirán en el futuro y ocuparán alturas completas de ranuras, anchos dobles de ranuras, o ambos. Las ranuras mecánicas dentro del gabinete admiten servidores que poseen el doble de altura y/o de anchura que el módulo de media altura.

El módulo de servidor es un módulo de lámina de metal que aloja a los discos duros de conexión en caliente del servidor, y está diseñado para no requerir herramientas en ninguna operación de ensamblaje o desensamblado. La extracción de cualquier pieza del servidor es intuitiva y no requiere herramientas. Esto incluye la tapa superior, la instalación del planar, los discos duros, las tarjetas de estructuras y la canastilla extraíble de discos duros, la figura 3.16 describe algunas características internas de esta cuchilla.

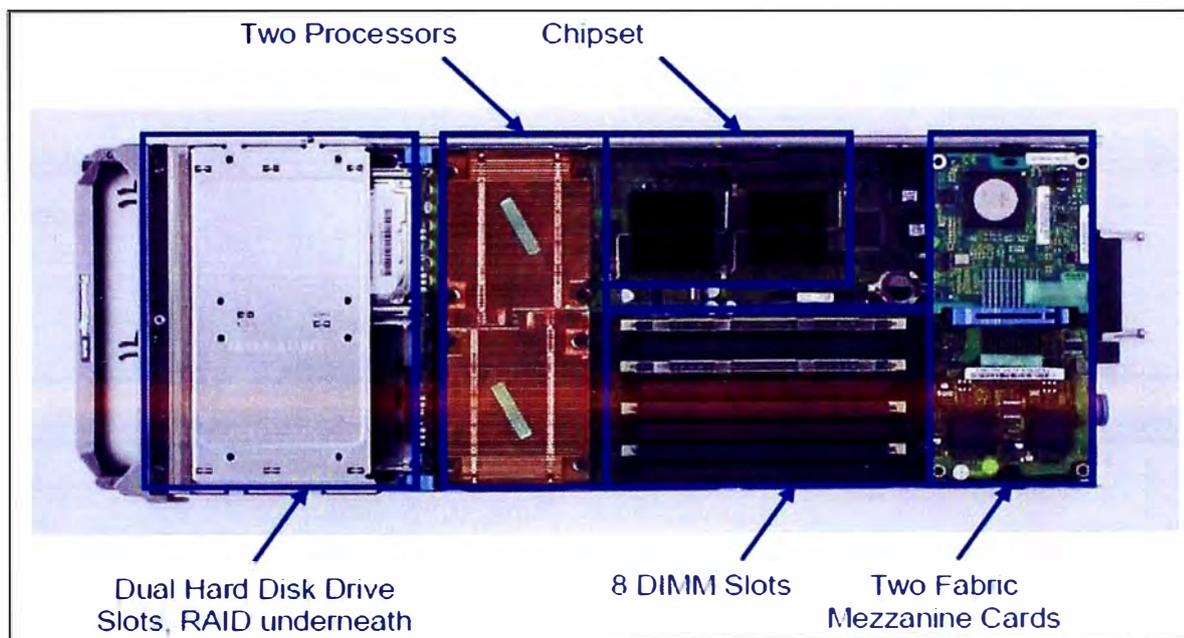


Fig. 3.16 Estructura de los módulos de servidor para el chasis Dell.

Para el diseño se usa los slots 4, 5, 6, 7 como se muestra en la tabla 3.2.

TABLA N° 3.2 Slots usados en el enclosure Power Edge M1000e.

Slot	Descripción
4	Nodo 1 – RAC Oracle 11gR2
5	Nodo 2 – RAC Oracle 11gR2
6	Nodo 3 – RAC Oracle 11gR2
7	Nodo 4 – RAC Oracle 11gR2

Se muestra la distribución final de los nodos en el enclosure M1000e y demás cuchillas en la parte frontal, la figura 3.15 muestra la distribución final en el ambiente real.

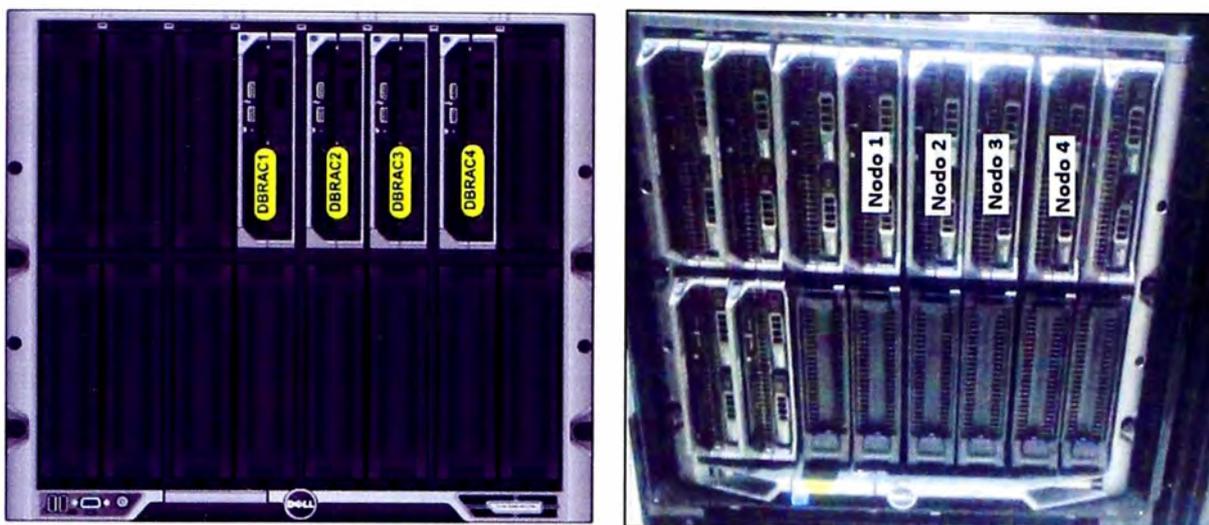


Fig. 3.15 Chasis modular – Nodos de la base de datos Oracle RAC.

La tabla 3.3 resume las características de las cuchillas que son usadas en la elaboración del proyecto.

TABLA N° 3.3 Descripción de los módulos de servidor para el chasis Dell.

Server Module	M600
Processor	Intel Xeon Dual and Quad Core 40W, 65W, 80W, and 120W options
Chipset	Intel 5000P (Blackford)
Memory Slots	8 Fully Buffered DIMMs (667 MHz)
Memory Capacity	32GB (4GB x 8) at launch 64 GB (8GB x 8) planned Q108
LOM	2 x GE with hardware TCP/IP Offload Engine and iSCSI Firmware Boot Upgrade to full iSCSI offload via license key
Fabric Expansión	3 x 8 lane PC Dual port GE w/ TOE Dual Port FC4 (Emulex & Qlogic) Dual Port 4x DDR InfiniBand
Console	Virtual KVM vía iDRAC IPMI Serial over LAN (SoL) vía iDRAC Rear mounted iKVM switch ports (tierable) Front KVM ports on modular enclosure control panel
Operating Systems	Oracle Linux OEL 5u8

b. Arquitectura E/S del chasis Dell

Admite hasta 6 módulos de E/S (entrada y salida) del blade en total para obtener 3 estructuras totalmente redundantes que aprovecha un ambiente completamente potente capaz de ampliarse a 8,4Tb/s. Con las tecnologías actuales de 1/10Gbps ethernet en los servidores se puede obtener hasta 40Gbps en los switches ethernet y 32Gb para los switches de canal de fibra, la figura 3.17 muestra la implementación real.



Fig. 3.17 Ubicación de módulos fabric switch del chasis Dell.

Se muestra la distribución final de switches ethernet y fabric switches del enclosure M1000e en la tabla 3.4.

TABLA N° 3.4 Módulos fabric switch usados en el chasis Dell.

Slot	Descripción
A1	Blade switch 3130 for Dell (LAN public)
A2	Blade switch 3130 for Dell (LAN public)
B1	Blade switch 3130 for Dell (LAN interconnect)
B2	Blade switch 3130 for Dell (LAN interconnect)
C1	Switch fabric Brocade M5424 (SAN)
C2	Switch fabric Brocade M5424 (SAN)

La interconexión de los switches blade con los módulos de servidores, se realiza mediante un BUS interno (M1000e midplane) donde cada server module integra 3 estructuras de interconexión (Tarjetas de Red). La estructura A (Fabric A LOM) viene por defecto y consiste de 2 puertos giga ethernet, las estructuras B y C son idénticas y totalmente personalizables, y están conectadas como dos conjuntos de cuatro canales

desde las tarjetas intermedias en los módulos del servidor a los módulos de E/S en la parte posterior del chasis. En la figura 3.18 describimos cada uno de las Estructuras disponibles para el chasis Dell.

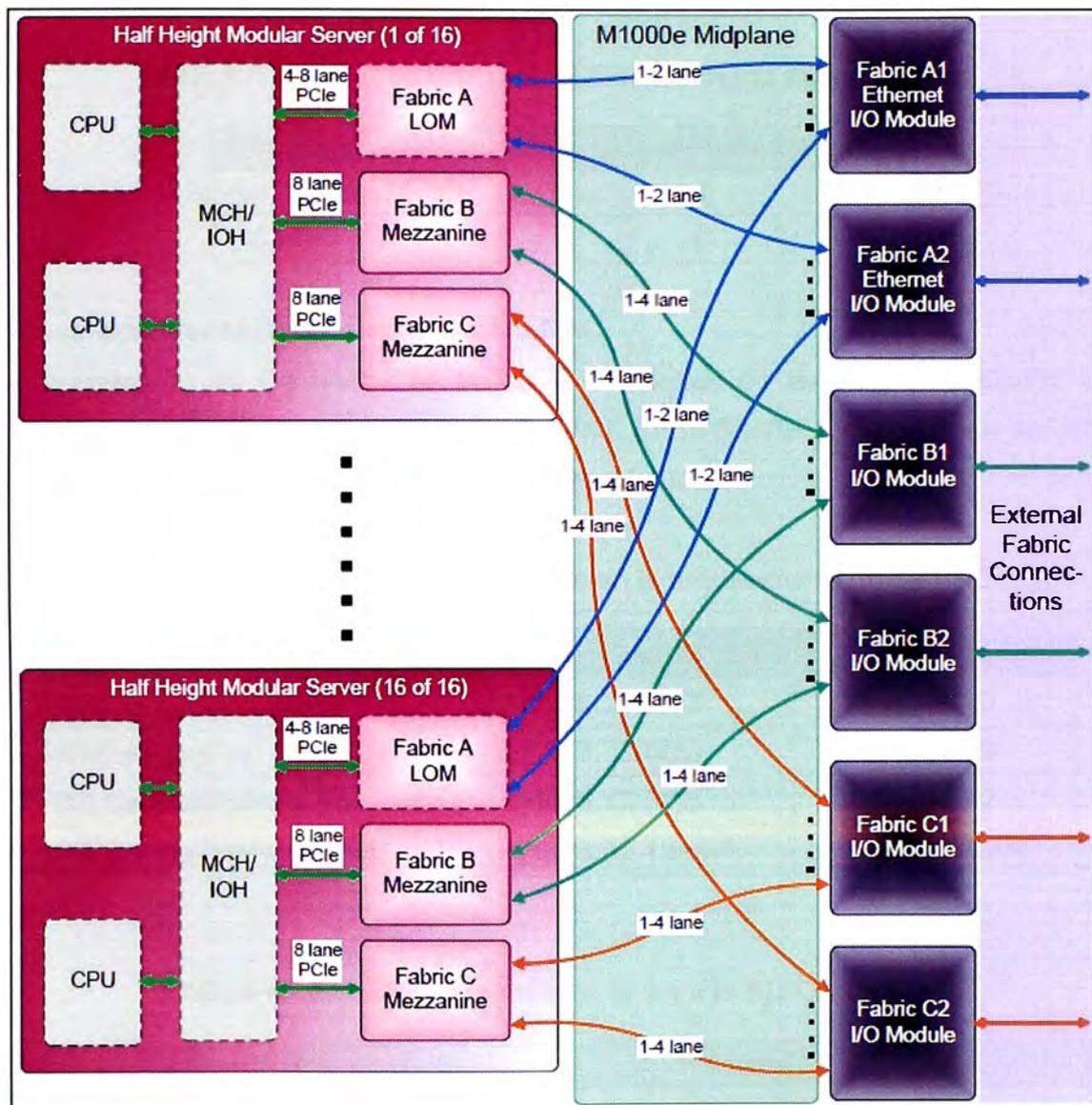


Fig. 3.18 Conexión E/S del chasis Dell.

- La red fabric A es una red fabric ethernet Gb redundante que admite las ranuras A1 y A2 de módulo de E/S. Las controladoras ethernet integradas en cada blade establecen la red fabric A como solo ethernet.
- La red fabric B es una red fabric redundante de 1 a 40 Gb/s que admite las ranuras B1 y B2 de módulo de E/S. Actualmente, la red fabric B admite módulos ethernet de 1 o 10 Gb, infiniband DDR/QDR y fibre channel de 4 u 8 Gbps. Es posible que en el futuro se admitan otros tipos de red fabric.

- La red fabric C es una red fabric redundante de 1 a 40 Gb/s que admite las ranuras C1 y C2 de módulo de E/S. Actualmente, la red fabric C admite módulos ethernet de 1 o 10 Gb, infiniband DDR/QDR y fibre channel de 4 u 8 Gbps. Es posible que en el futuro se admitan otros tipos de red fabric. La tabla 3.5 muestra las velocidades a usar.

TABLA N° 3.5 Velocidades de TX en los fabric switch usados.

Fabric	Data rate per lane (Gbps)
GE1000BASE-KX	1
FC 8G	8

c. Direccionamiento IP de nodos Oracle RAC

Para la conexión de los nodos se aplica lo conceptos de teaming y multipath, de tal manera que se encuentran en alta disponibilidad. En el diseño a implementar se utiliza el direccionamiento IP mostrado en la tabla 3.6 y tabla 3.7.

TABLA N° 3.6 Redes LAN usadas en la arquitectura Oracle RAC.

Nombre de red	Dirección de red	VLAN
Red de gestión SW	172.16.26.0/23	VLAN 49
Red de gestión	172.16.73.0/24	VLAN 73
DB Oracle (Public)	172.16.72.0/25	VLAN 72
DB Oracle (Interconnect)	172.16.72.128/25	No usa VLAN

TABLA N° 3.7 Direccionamiento IP para la BD Oracle RAC.

Equipo	IP LAN	IP virtual	Mascara	Gateway	Team	
Nodo 1	Eth0 Eth1	172.16.72.3	172.16.72.4	255.255.255.128	172.16.72.1	Bond0
	Eth2 Eth3	172.16.72.131		255.255.255.128		Bond1
Nodo 2	Eth0 Eth1	172.16.72.5	172.16.72.6	255.255.255.128	172.16.72.1	Bond0
	Eth2 Eth3	172.16.72.132		255.255.255.128		Bond1
Nodo 3	Eth0 Eth1	172.16.72.7	172.16.72.8	255.255.255.128	172.16.72.1	Bond0
	Eth2 Eth3	172.16.72.133		255.255.255.128		Bond1

Nodo 4	Eth0	172.16.72.9	172.16.72.10	255.255.255.128	172.16.72.1	Bond0
	Eth1					
	Eth2	172.16.72.134		255.255.255.128		Bond1
	Eth3					
Scan			172.16.72.11 172.16.72.12 172.16.72.13	255.255.255.128		Scan

Los servidores se configuran en modo de alta redundancia, para esto se utiliza el término “bonding”, que consiste en el agrupamiento de puertos físicos en un puerto lógico virtual. Esta configuración permite incrementar el ancho de banda disponible proporcional al número de interfaces, también ayuda a mantener la alta disponibilidad de los servidores ante la falla repentina de alguna interface agrupada.

- Direccionamiento en el nodo 1 – RAC Oracle 11gR2, tabla 3.8 y figura 3.19.

TABLA N° 3.8 Configuración – Server module Dell (Nodo 1).

```
[oracle@srviwaybdprod001 ~]$ cat /etc/sysconfig/network-scripts/ifcfg-bond0
DEVICE=bond0
BOOTPROTO=none
NM_CONTROLLED=no
IPADDR=172.16.72.3
NETMASK=255.255.255.128
ONBOOT=yes
USERCTL=no
BONDING_OPTS="mode=active-backup miimon=100"

[oracle@srviwaybdprod001 ~]$ cat /etc/sysconfig/network-scripts/ifcfg-bond1
DEVICE=bond1
BOOTPROTO=none
IPADDR=172.16.72.131
NETMASK=255.255.255.128
ONBOOT=yes
USERCTL=no
BONDING_OPTS="mode=active-backup miimon=100"
NM_CONTROLLED=no

[oracle@srviwaybdprod001 ~]$ cat /sys/class/fc_host/host1/port_name
0x2001848f69299e10
[oracle@srviwaybdprod001 ~]$ cat /sys/class/fc_host/host2/port_name
0x2002848f69299e10
```

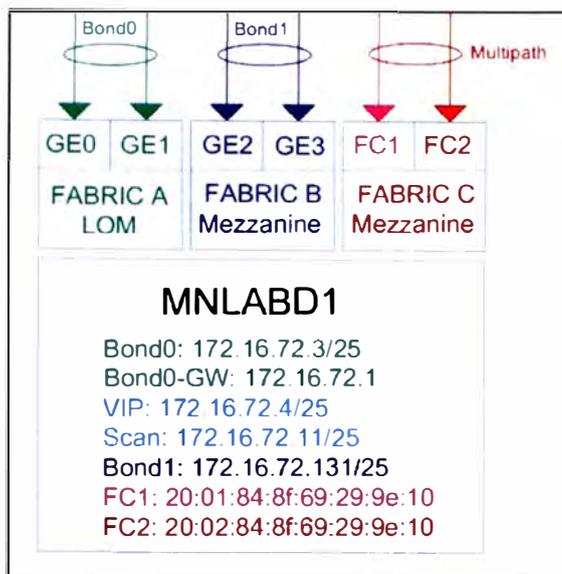


Fig. 3.19 Representación lógica – Server module Dell (Nodo 1).

- Direcccionamiento en el nodo 2 – RAC Oracle 11gR2, tabla 3.9 y figura 3.20.

TABLA N° 3.9 Configuración – Server module Dell (Nodo 2).

```
[oracle@srviwaybdprod002 ~]$ cat /etc/sysconfig/network-scripts/ifcfg-bond0
DEVICE=bond0
BOOTPROTO=none
NM_CONTROLLED=no
IPADDR=172.16.72.5
NETMASK=255.255.255.128
ONBOOT=yes
USERCTL=no
BONDING_OPTS="mode=active-backup miimon=100"

[oracle@srviwaybdprod002 ~]$ cat /etc/sysconfig/network-scripts/ifcfg-bond1
DEVICE=bond1
BOOTPROTO=none
IPADDR=172.16.72.132
NETMASK=255.255.255.128
ONBOOT=yes
USERCTL=no
BONDING_OPTS="mode=active-backup miimon=100"
NM_CONTROLLED=no

[oracle@srviwaybdprod002 ~]$ cat /sys/class/fc_host/host1/port_name
0x2001848f69299e1d
[oracle@srviwaybdprod002 ~]$ cat /sys/class/fc_host/host2/port_name
0x2002848f69299e1d
```

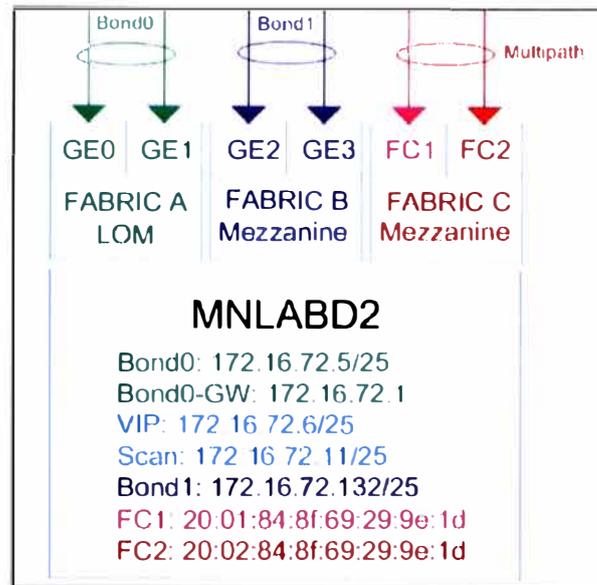


Fig. 3.20 Representación lógica – Server module Dell (Nodo 2).

- Direccionamiento en el nodo 3 – RAC Oracle 11gR2, tabla 3.10 y figura 3.21.

TABLA N° 3.10 Configuración – Server module Dell (Nodo 3).

```
[oracle@srviwaybdprod003 ~]$ cat /etc/sysconfig/network-scripts/ifcfg-bond0
DEVICE=bond0
BOOTPROTO=none
NM_CONTROLLED=no
IPADDR=172.16.72.7
NETMASK=255.255.255.128
ONBOOT=yes
USERCTL=no
BONDING_OPTS="mode=active-backup miimon=100"

[oracle@srviwaybdprod003 ~]$ cat /etc/sysconfig/network-scripts/ifcfg-bond1
DEVICE=bond1
BOOTPROTO=none
IPADDR=172.16.72.133
NETMASK=255.255.255.128
ONBOOT=yes
USERCTL=no
BONDING_OPTS="mode=active-backup miimon=100"
NM_CONTROLLED=no

[oracle@srviwaybdprod003 ~]$ cat /sys/class/fc_host/host1/port_name
0x2001848f69299e2a
[oracle@srviwaybdprod003 ~]$ cat /sys/class/fc_host/host2/port_name
0x2002848f69299e2a
```

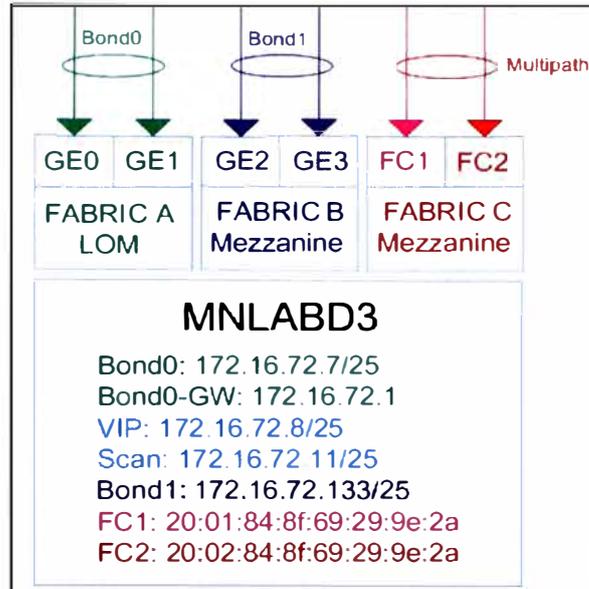


Fig. 3.21 Representación lógica – Server module Dell (Nodo 3).

- Direccionamiento en el nodo 4 – RAC Oracle 11gR2, tabla 3.11 y figura 3.22.

TABLA N° 3.11 Configuración – Server module Dell (Nodo 4).

```
[oracle@srviwaybdprod004 ~]$ cat /etc/sysconfig/network-scripts/ifcfg-bond0
DEVICE=bond0
BOOTPROTO=none
NM_CONTROLLED=no
IPADDR=172.16.72.9
NETMASK=255.255.255.128
ONBOOT=yes
USERCTL=no
BONDING_OPTS="mode=active-backup miimon=100"

[oracle@srviwaybdprod004 ~]$ cat /etc/sysconfig/network-scripts/ifcfg-bond1
DEVICE=bond1
BOOTPROTO=none
IPADDR=172.16.72.134
NETMASK=255.255.255.128
ONBOOT=yes
USERCTL=no
BONDING_OPTS="mode=active-backup miimon=100"
NM_CONTROLLED=no

[oracle@srviwaybdprod004 ~]$ cat /sys/class/fc_host/host1/port_name
0x2001848f69299e37
[oracle@srviwaybdprod004 ~]$ cat /sys/class/fc_host/host2/port_name
0x2002848f69299e37
```

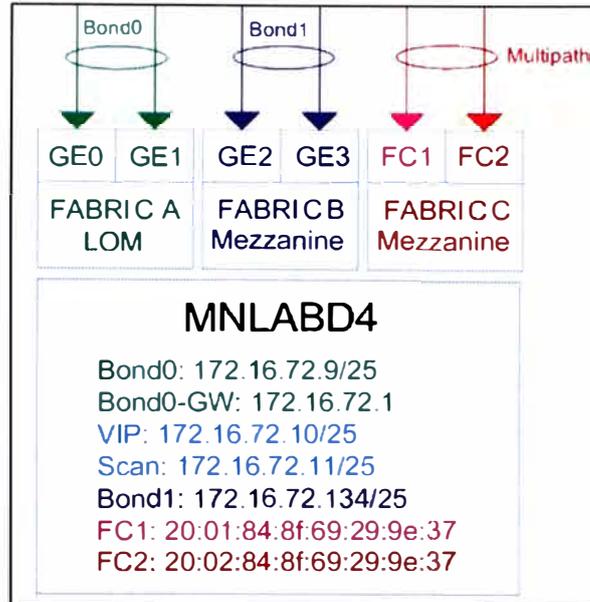


Fig. 3.22 Representación lógica – Server module Dell (Nodo 4).

d. Conectividad E/S LAN del chasis Dell

En la interconexión de los servidores se utiliza los switches Cisco Catalyst 3130G, a continuación se muestra la configuración realizada por switch. Esta configuración también muestra los puertos asociados a cada enlace.

- Switches Cisco 3130G para la red PUBLIC (Switch A1 y switch A2), tabla 3.12.

TABLA N° 3.12 Configuración – Switches Cisco blade A1 y A2.

<pre> A1#sh vl b VLAN Name Status Ports ----- 72 DB active Gi0/4, Gi0/5, Gi0/6, Gi0/7 A1#sh etherchannel summary Group Port-channel Protocol Ports -----+-----+-----+----- 14 Po14(SU) - Gi0/19(P) Gi0/20(P) </pre>			
<pre> A2#sh vl b VLAN Name Status Ports ----- 72 DB active Gi0/4, Gi0/5, Gi0/6, Gi0/7 A2#sh etherchannel summary Group Port-channel Protocol Ports -----+-----+-----+----- 15 Po15(SU) - Gi0/19(P) Gi0/20(P) </pre>			

Estos switches son la cara hacia la red pública, por ello necesitan suficiente ancho de banda para soportar las transacciones, esto se grafica en las figuras 3.23 y figura 3.24.

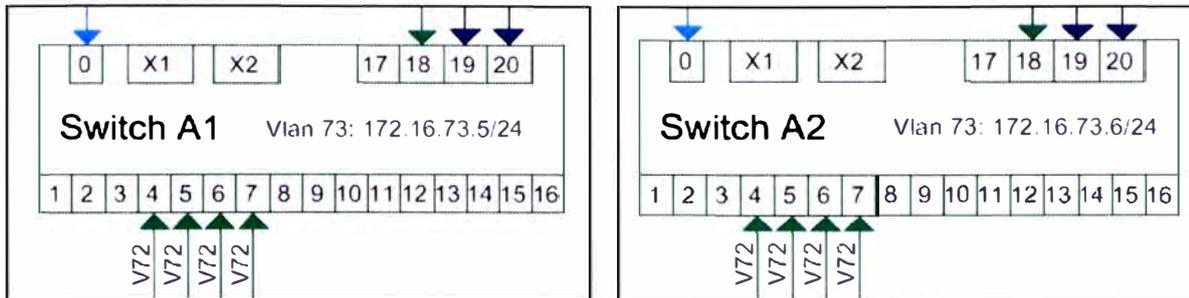


Fig. 3.23 Representación lógica – Switches Cisco blade A1 y A2.

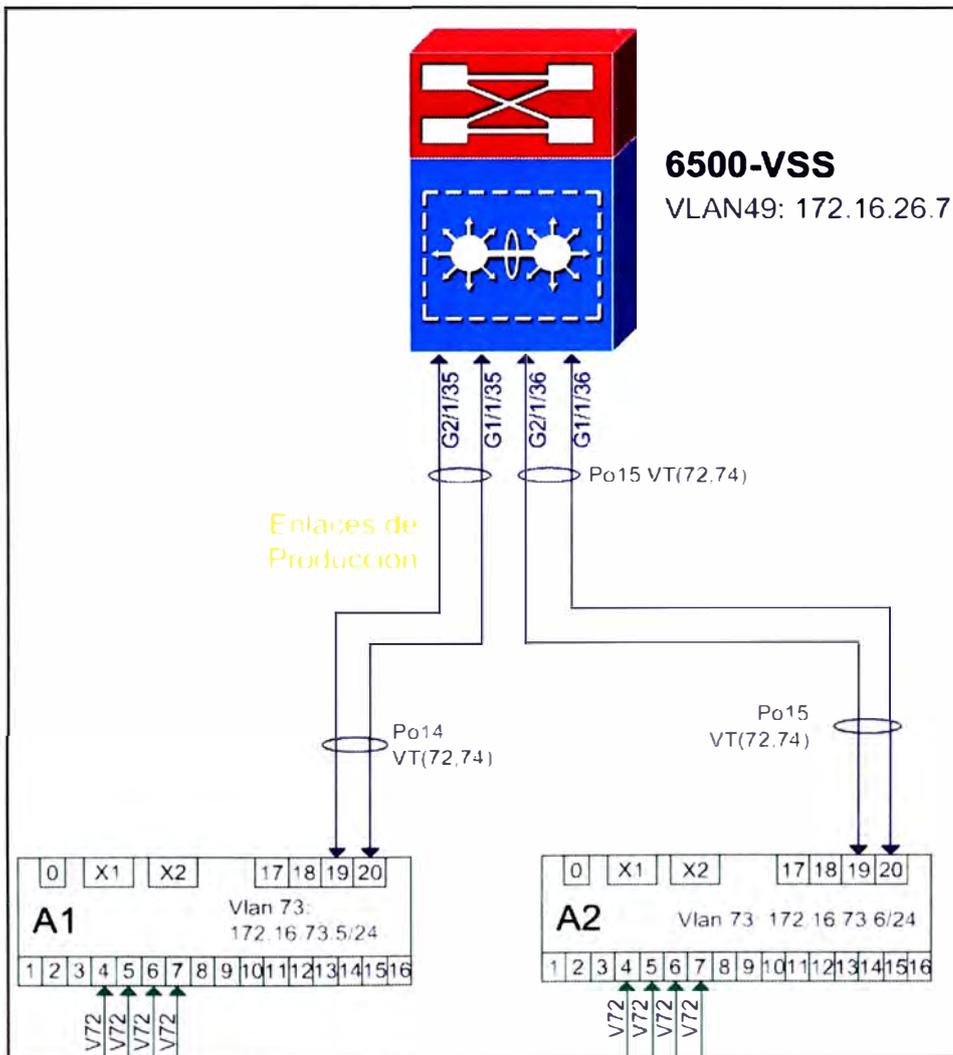


Fig. 3.24 Etherchannel entre el switch core y el enclosure M100e.

- Switches Cisco 3130G para la red INTERCONNECT (Switch B1 y switch B2), la tabla 3.13 y la figura 3.25 muestran los puertos que se usan en el proyecto.

TABLA N° 3.13 Configuración – Switches Cisco blade B1 y B2.

<pre> B1#sh vl b VLAN Name Status Ports ----- 2012 INTERCONNECT active Gi0/4, Gi0/5, Gi0/6, Gi0/7 </pre>	
<pre> B1#sh etherchannel summary ----- 2 Po2(SU) - Gi0/19(P) Gi0/20(P) </pre>	
<pre> B2#sh vl b VLAN Name Status Ports ----- 2012 INTERCONNECT active Gi0/4, Gi0/5, Gi0/6, Gi0/7 </pre>	
<pre> B2#sh etherchannel summary Group Port-channel Protocol Ports ----- 2 Po2(SU) - Gi0/19(P) Gi0/20(P) </pre>	

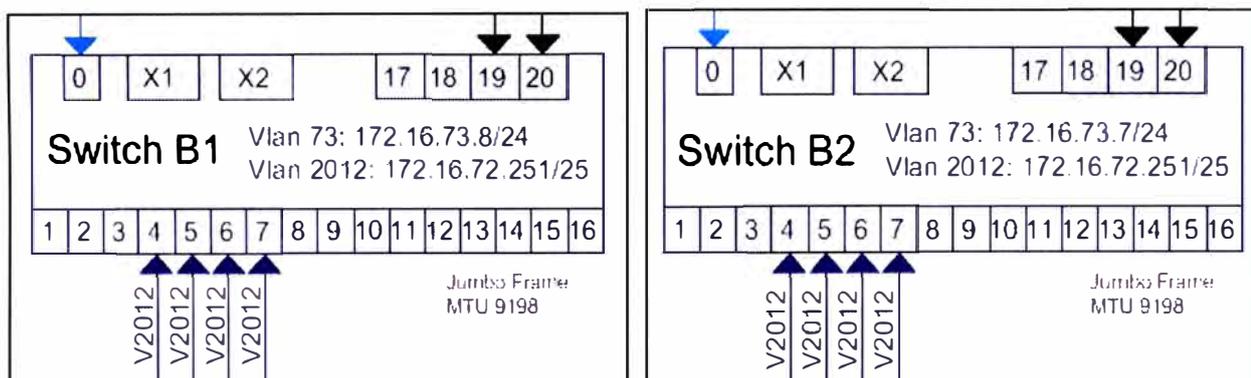


Fig. 3.25 Representación lógica – Switches Cisco B1 y B2.

e. Conectividad E/S SAN del chasis Dell

El switch Brocade M5424 y el chasis blade Dell Power Edge M1000e ofrecen sólidas soluciones para implementaciones SAN de fiber channel. Este switch no sólo ayuda a simplificar y reducir la cantidad de componentes de hardware SAN necesarios para una implementación, sino que mantiene la capacidad de ampliación, la interoperabilidad, el rendimiento y la administración de los entornos SAN tradicionales.

El switch Brocade tiene la capacidad de trabajar en modo access gateway. En esta configuración, el switch SAN simplifica la implementación de SAN al utilizar la tecnología NPIV, la cual se usara con el switch Cisco MDS 9509 director, la figura 3.26 muestra las posibles configuraciones de estos switches.

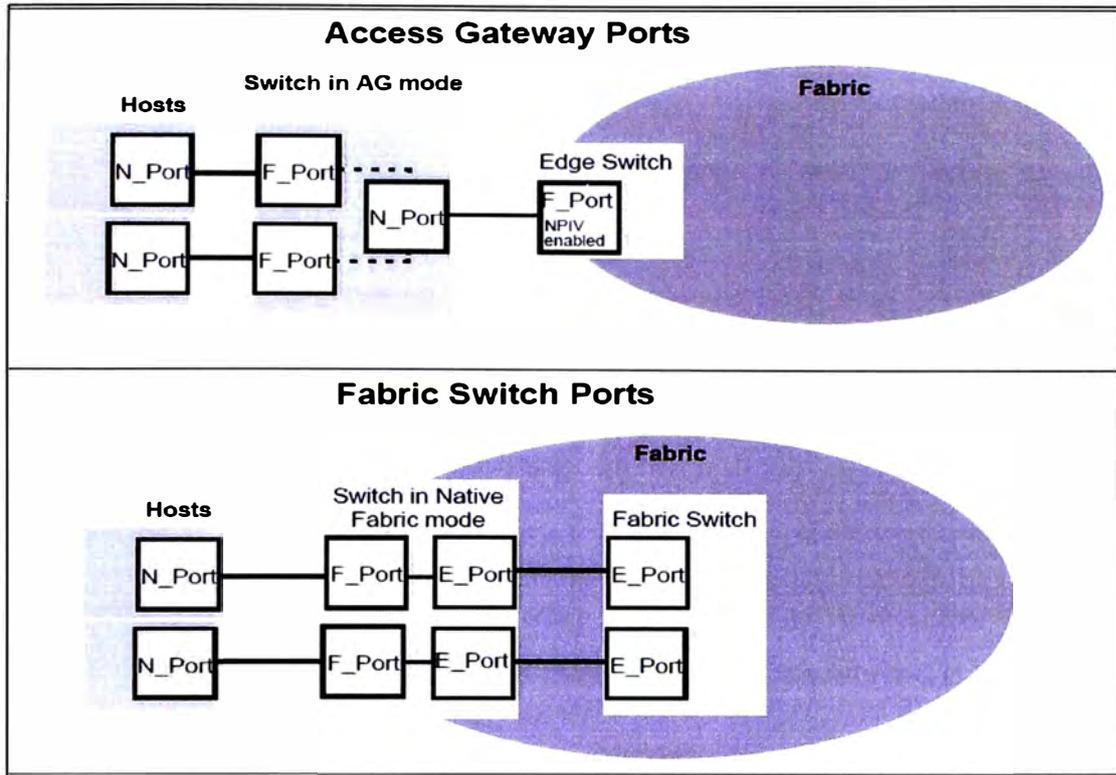


Fig. 3.26 Comparación entre el modo access gateway y fabric switch estándar.

Se utiliza también el concepto de agrupamiento de puertos en este tipo de switch, de esta manera los puertos tipo N pueden ser agrupados en un único puerto virtual, los puertos tipo F también pueden ser mapeados a este puerto virtual. El agrupamiento permite diseñar enlaces redundantes y ampliar el ancho de banda, este comportamiento se muestra en la figura 3.27, tabla 3.14, tabla 3.15, figura 3.30 y figura 3.31

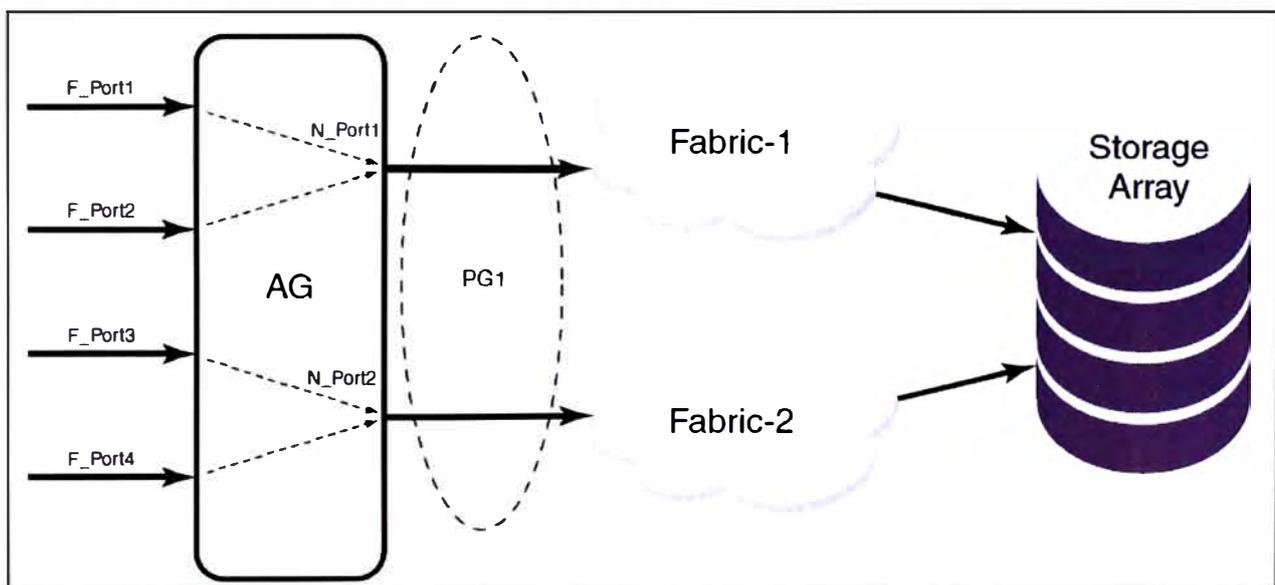


Fig. 3.27 Agrupamiento de puertos (PG) en los puertos tipo N.

- Descripción del switch M5424 (Switch fabric C1 y C2), figura 3.28 y figura 3.29.

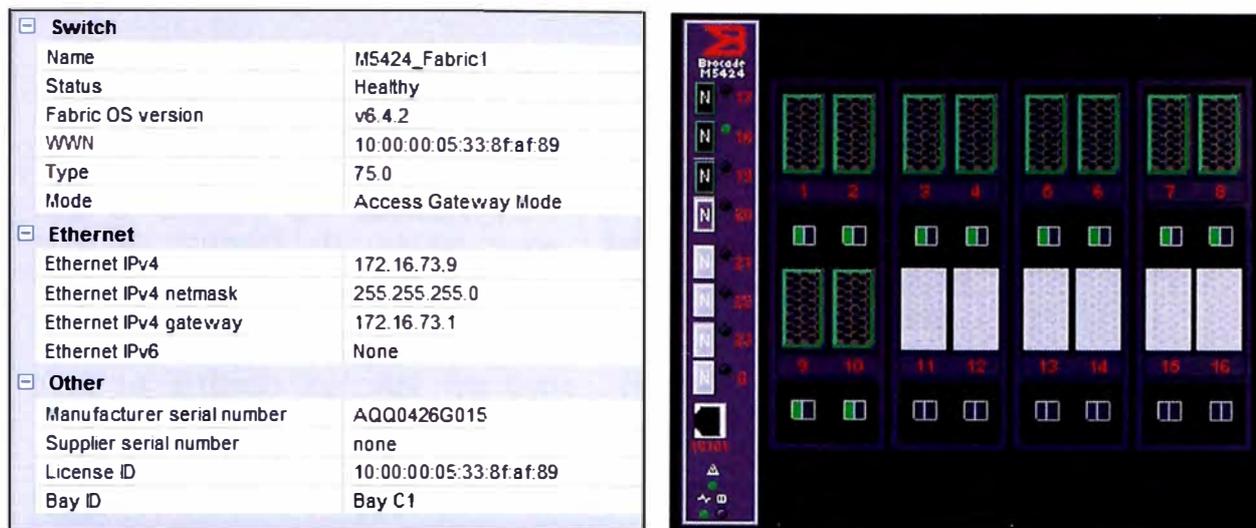


Fig. 3.28 Información del dispositivo – Switch fabric C1.

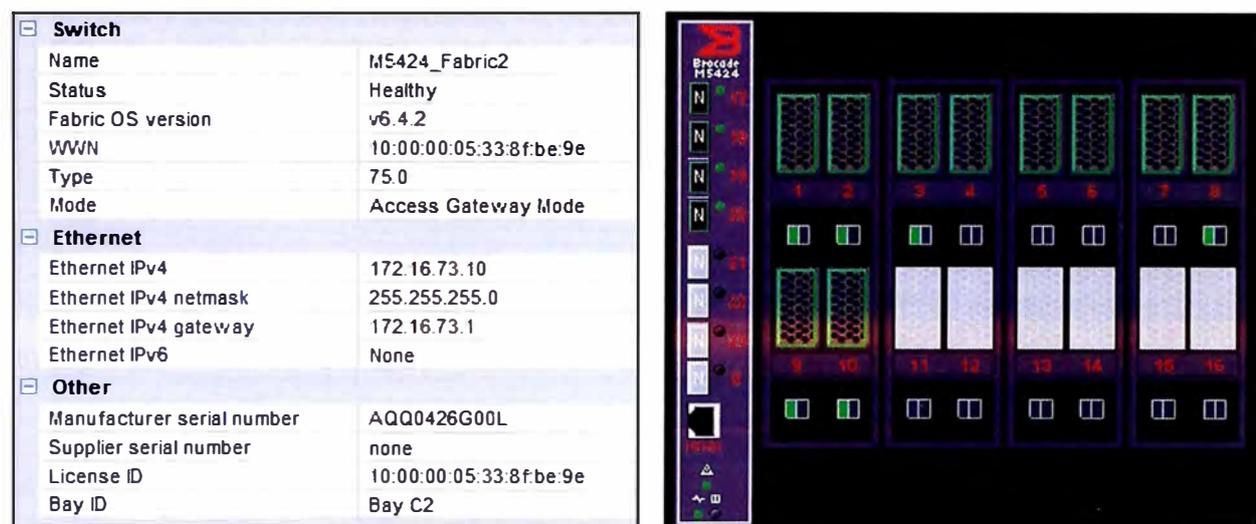


Fig. 3.29 Información del dispositivo – Switch fabric C2.

TABLA N° 3.14 Configuración – Switches fabric Brocade C1.

```

M5424_Fabric1:jorge.riveros>switchshow
switchName:  M5424_Fabric1
switchType:  75.0
switchState: Online
switchMode:  Access Gateway Mode
switchWwn:   10:00:00:05:33:8f:af:89
switchBeacon: OFF

Index Port Address Media Speed State  Proto
=====
0 0 010000 id 4G No Light FC
    
```

```

1 1 010100 cu ANNo_Sync FC Disabled (Persistent)
2 2 010200 cu AN No_Sync FC Disabled (Persistent)
3 3 010300 cu ANNo_Sync FC Disabled (Persistent)
4 4 010400 cu N8 Online FC F-Port 20:01:84:8f:69:29:9e:10 0x0a0002
5 5 010500 cu N8 Online FC F-Port 20:01:84:8f:69:29:9e:1d 0x0a0003
6 6 010600 cu N8 Online FC F-Port 20:01:84:8f:69:29:9e:2a 0x0a0042
7 7 010700 cu N8 Online FC F-Port 20:01:84:8f:69:29:9e:37 0x0a0062
8 8 010800 cu ANNo_Sync FC Disabled (Persistent)
9 9 010900 cu ANNo_Sync FC Disabled (Persistent)
10 10 010a00 cu AN No_Sync FC Disabled (Persistent)
11 11 010b00 cu AN No_Sync FC Disabled (Persistent)
12 12 010c00 cu AN No_Sync FC Disabled (Persistent)
13 13 010d00 cu AN No_Sync FC Disabled (Persistent)
14 14 010e00 cu AN No_Sync FC Disabled (Persistent)
15 15 010f00 cu AN No_Sync FC Disabled (Persistent)
16 16 011000 cu AN No_Sync FC Disabled (Persistent)
17 17 011100 id N4 Online FC N-Port 20:0a:54:7f:ee:03:3a:01 0x0a0000
18 18 011200 id N4 Online FC N-Port 20:0a:54:7f:ee:03:3a:01 0x0a0020
19 19 011300 id N4 Online FC N-Port 20:0a:54:7f:ee:03:3a:01 0x0a0040
20 20 011400 id N8 No_Light FC

M5424_Fabric1:jorge.riveros>ag --pgshow
PG_ID PG_NamePG_ModeN_PortsF_Ports
-----
0 pg0 - None
2 pg1 lb 0;17;18;19;20; 1;2;3;4;5;6;7;8;9;10;11;12;13;14;15;16;
21;22;23
-----

```

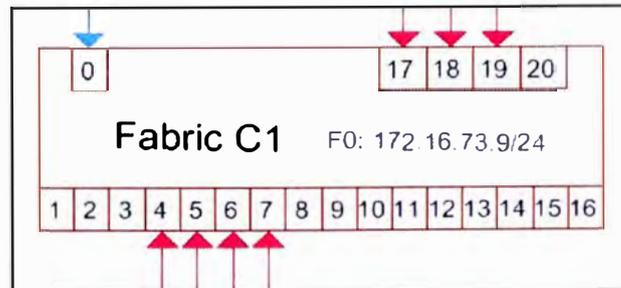


Fig. 3.30 Representación lógica – Switch fabric C1.

TABLA N° 3.15 Configuración – Switches fabric Brocade C2.

```

M5424_Fabric2:Jorge.riveros>switchshow
switchName: M5424_Fabric2
switchType: 75.0
switchState: Online
switchMode: Access Gateway Mode
switchWwn: 10:00:00:05:33:8f:be:9e
switchBeacon: OFF

Index Port Address Media Speed State Proto
=====

```

```

0 0 010000 id N8 No_Light FC
1 1 010100 cu AN No_Sync FC Disabled (Persistent)
2 2 010200 cu AN No_Sync FC Disabled (Persistent)
3 3 010300 cu AN No_Sync FC Disabled (Persistent)
4 4 010400 cu N8 Online FC F-Port 20:02:84:8f:69:29:9e:10 0x140003
5 5 010500 cu N8 Online FC F-Port 20:02:84:8f:69:29:9e:1d 0x140004
6 6 010600 cu N8 Online FC F-Port 20:02:84:8f:69:29:9e:2a 0x140005
7 7 010700 cu N8 Online FC F-Port 20:02:84:8f:69:29:9e:37 0x140023
8 8 010800 cu AN No_Sync FC Disabled (Persistent)
9 9 010900 cu AN No_Sync FC Disabled (Persistent)
10 10 010a00 cu AN No_Sync FC Disabled (Persistent)
11 11 010b00 cu AN No_Sync FC Disabled (Persistent)
12 12 010c00 cu AN No_Sync FC Disabled (Persistent)
13 13 010d00 cu AN No_Sync FC Disabled (Persistent)
14 14 010e00 cu AN No_Sync FC Disabled (Persistent)
15 15 010f00 cu AN No_Sync FC Disabled (Persistent)
16 16 011000 cu AN No_Sync FC Disabled (Persistent)
17 17 011100 id N4 Online FC N-Port 20:14:54:7f:ee:03:3a:01 0x140000
18 18 011200 id N4 Online FC N-Port 20:14:54:7f:ee:03:3a:01 0x140020
19 19 011300 id N4 Online FC N-Port 20:14:54:7f:ee:03:3a:01 0x140060
20 20 011400 id N4 Online FC N-Port 20:14:54:7f:ee:03:3a:01 0x140040
21 21 011500 id N8 No_Light FC
22 22 011600 id N8 No_Light FC

M5424_Fabric2:jorge.riveros>ag --pgshow
PG_ID PG_NamePG_ModeN_PortsF_Ports
-----
0 pg0 - None
1 pg1 lb 0;17;18;19;20; 1;2;3;4;5;6;7;8;9;10;11;12;13;14;15;16;
21;22;23
-----
23 23 011700 id N8 No_Light FC
    
```

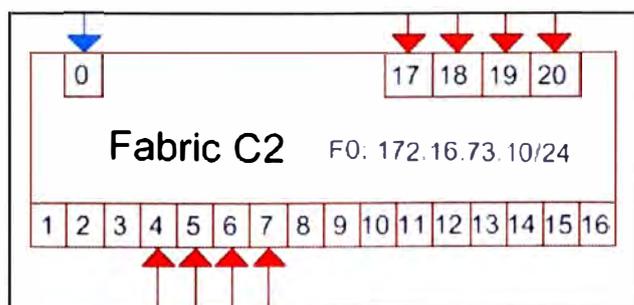


Fig. 3.31 Representación lógica – Switch fabric C2.

f. Conectividad del Cisco MDS 9509 multilayer director switch

El Cisco director MDS 9509 encargado de conectar la red SAN entre el enclosure M1000e, storage Hitachi y la librería de cintas TL2000 se muestra la tabla 3.16 donde se listan los módulos instalados en el chasis, para ello se ejecuta el siguiente comando en la CLI del switch.

TABLA N° 3.16 Módulos usados en el switch Cisco MDS 9509.

```

MDS9509Lurin# show module
Mod Ports Module-Type           Model           Status
-----
1  48  1/2/4/8 Gbps FC Module       DS-X9248-96K9  ok
2  48  1/2/4/8 Gbps FC Module       DS-X9248-96K9  ok
5  0   Supervisor/Fabric-2a         DS-X9530-SF2AK9 active *
6  0   Supervisor/Fabric-2a         DS-X9530-SF2AK9 ha-standby

Mod SwHw      World-Wide-Name(s) (WWN)
-----
1  5.2(1)      1.5  20:01:54:7f:ee:03:3a:00 to 20:30:54:7f:ee:03:3a:00
2  5.2(1)      1.5  20:41:54:7f:ee:03:3a:00 to 20:70:54:7f:ee:03:3a:00
5  5.2(1)      1.3  --
6  5.2(1)      1.3  --

Mod MAC-Address(es)           Serial-Num
-----
1  00-0d-ec-76-e1-20 to 00-0d-ec-76-e1-24 JAF1537BPTQ
2  c8-4c-75-b1-20-24 to c8-4c-75-b1-20-28 JAF1537BQAP
5  c8-9c-1d-41-77-ac to c8-9c-1d-41-77-b0 JAF1538BCTC
6  c8-9c-1d-41-77-b8 to c8-9c-1d-41-77-bc JAF1538BCSL
    
```

El MDS posee doble modulo supervisor en modo activo – pasivo, lo que convierte al MDS en el núcleo de la red SAN, la figura 3.32 muestra la conectividad.

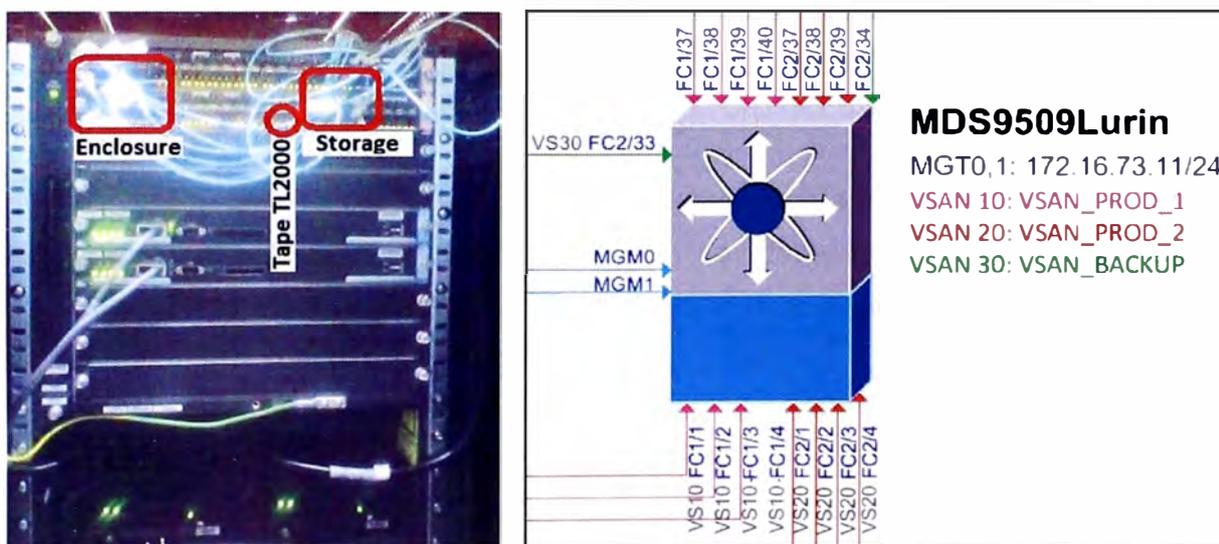


Fig. 3.32 Representación lógica del multilayer director switch.

En la figura 3.33 se muestra el diagrama de flujo necesario para la creación de una VSAN en el Switch Director MDS y la asignación a los puertos tipo F, ya que por defecto el chasis asigna la VSAN 1 a todas las interfaces.

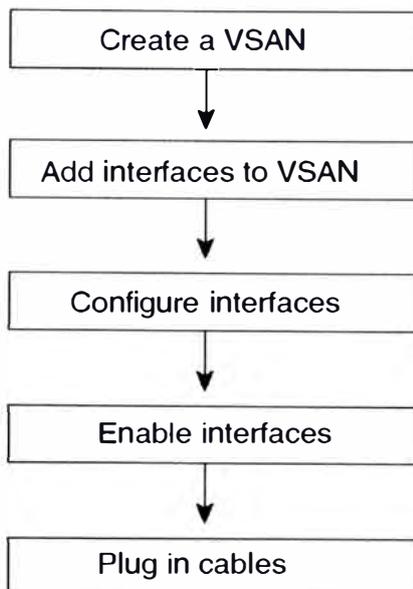


Fig. 3.33 Proceso de creación y asignación de VSANs.

En el diseño se utiliza las VSAN mostradas en la tabla 3.17 donde también está incluido el equipo de librerías de cintas.

TABLA N° 3.17 Lista de VSAN creadas.

VSAN	Nombre	Descripción
10	VSAN_PROD_1	Red de producción SAN A
20	VSAN_PROD_2	Red de producción SAN B
30	VSAN_BACKUP	Red de backup SAN Backup

La figura 3.34 muestra el diseño de red que se usa para la arquitectura SAN, los dispositivos están dibujados en su forma lógica, y el storage muestra el arreglo de LUN empleado para la base de datos, las tablas 3.18, 3.19 y 3.20 muestra la asignación de puertos a sus correspondientes VSAN.

TABLA N° 3.18 Asignación VSAN de puertos FC en el switch Cisco MDS 9509.

```

MDS9509Lurin# show vsan membership
vsan 10 interfaces:
  fc1/1      fc1/2      fc1/3      fc1/4
  fc1/37     fc1/38     fc1/39     fc1/40
vsan 20 interfaces:
  fc2/1      fc2/2      fc2/3      fc2/4
  fc2/37     fc2/38     fc2/39     fc2/40
vsan 30 interfaces:
  fc2/33     fc2/34     fc2/35     fc2/36
  
```

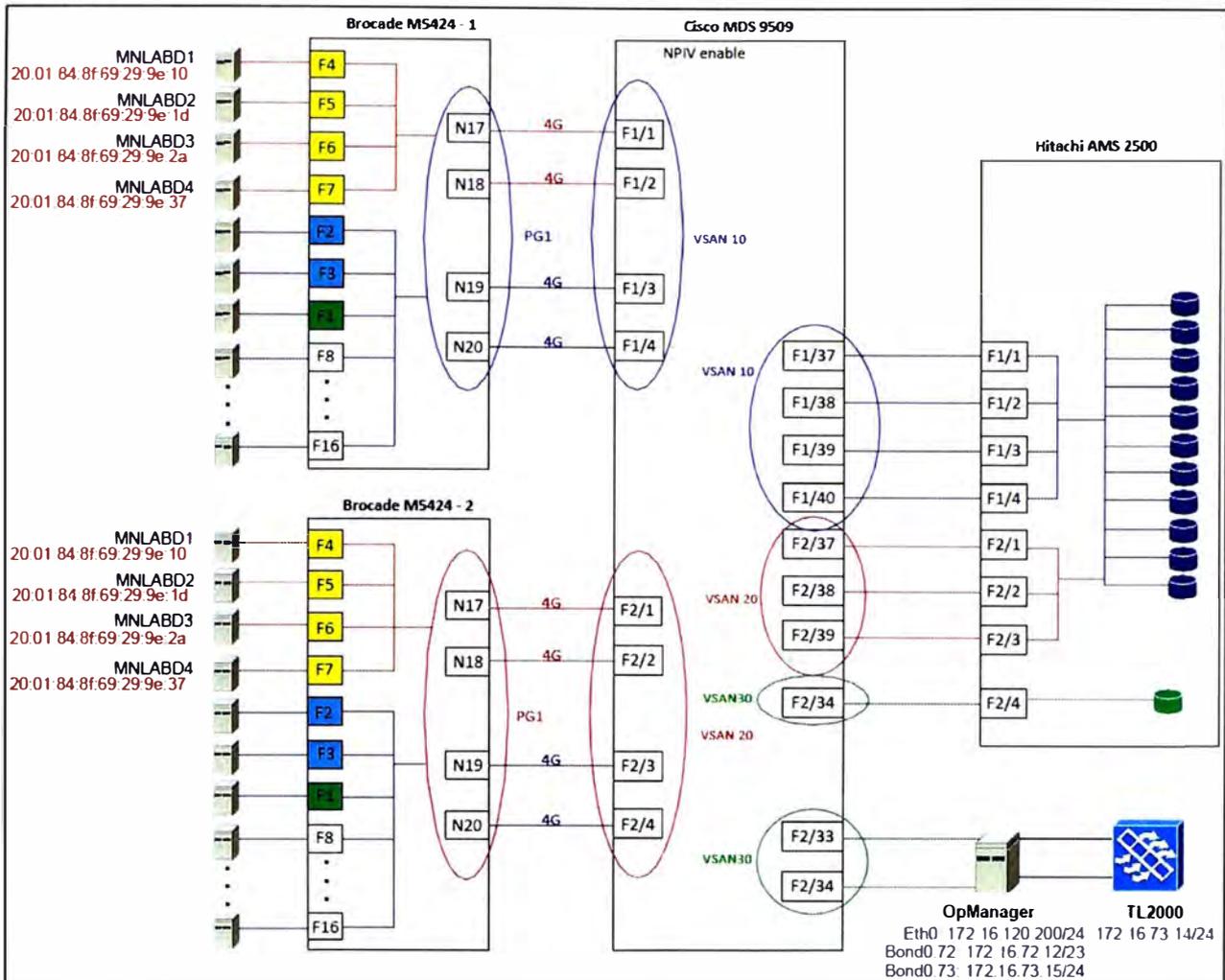


Fig. 3.34 Diseño de topología – Red SAN.

TABLA N° 3.19 Resumen de puertos FC y conexiones SAN – Modulo 1 Cisco MDS.

Slot – Modulo1	Equipo Conectado
FC1/1 (VSAN 10)	Enclosure C1 – G0/17
FC1/2 (VSAN 10)	Enclosure C1 – G0/18
FC1/3 (VSAN 10)	Enclosure C1 – G0/19
FC1/4 (VSAN 10)	Enclosure C1 – G0/20
FC1/37 (VSAN 10)	Storage Con0 – A
FC1/38 (VSAN 10)	Storage Con0 – B
FC1/39 (VSAN 10)	Storage Con0 – C
FC1/40 (VSAN 10)	Storage Con0 – D

TABLA N° 3.20 Resumen de puertos FC y conexiones SAN – Modulo 2 Cisco MDS.

Slot – Modulo2	Equipo conectado
FC2/1 (VSAN 20)	Enclosure C2 – G0/17
FC2/2 (VSAN 20)	Enclosure C2 – G0/18
FC2/3 (VSAN 20)	Enclosure C2 – G0/19
FC2/4 (VSAN 20)	Enclosure C2 – G0/20

FC2/37 (VSAN 20)	Storage Con1 – A
FC2/38 (VSAN 20)	Storage Con1 – B
FC2/39 (VSAN 20)	Storage Con1 – C
FC2/34 (VSAN 30)	Storage Con1 – D
FC2/33 (VSAN 30)	Tape FC0

Para habilitar el control de acceso entre dispositivos SAN, es necesaria la creación de zonas que incrementan la seguridad SAN, la figura 3.35 muestra este procedimiento.

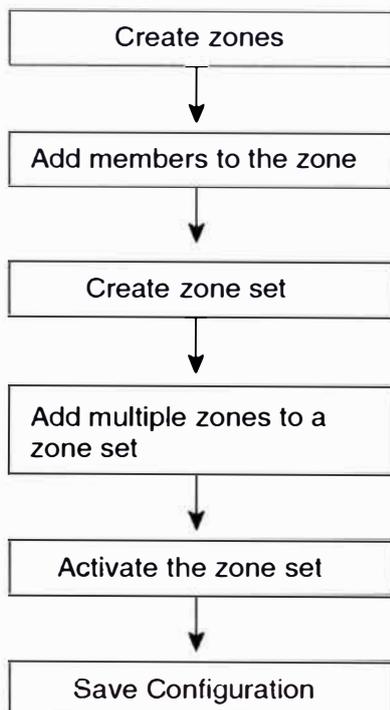


Fig. 3.35 Proceso de creación y asignación de zonas.

Las zonas creadas son listadas en la tabla 3.21 usando un alias, de esta manera se evita escribir las direcciones WWN para el resto de la configuración.

TABLA N° 3.21 Resumen de zonas creadas en el switch Cisco MDS 9509.

```

MDS9509Lurin# show zoneset brief
zoneset name ZONESET_VSAN_PROD_1 vsan 10
zone ZONE_AMS2500_0B_BLADE8_1
zone ZONE_AMS2500_0C_BLADE9_1
zone ZONE_AMS2500_0B_BLADE10_1
zone ZONE_AMS2500_0A_MNLABD1_1
zone ZONE_AMS2500_0A_MNLABD2_1
zone ZONE_AMS2500_0A_MNLABD3_1
zone ZONE_AMS2500_0A_MNLABD4_1

zoneset name ZONESET_VSAN_PROD_2 vsan 20
    
```

```
zone ZONE_AMS2500_1B_BLADE8_2
zone ZONE_AMS2500_1C_BLADE9_2
zone ZONE_AMS2500_1B_BLADE10_2
zone ZONE_AMS2500_1A_MNLABD1_2
zone ZONE_AMS2500_1A_MNLABD2_2
zone ZONE_AMS2500_1A_MNLABD3_2
zone ZONE_AMS2500_1A_MNLABD4_2

zoneset name zoneset_backup vsan 30
zone zone_backup
```

g. Conectividad en la red de backup y gestión

El switch Cisco Catalyst 2960 necesario para el funcionamiento de la red de gestión y el almacenamiento en cinta como se muestra en la figura 3.36.



Fig. 3.36 Switch Cisco Catalyst 2960G de gestión.

A través de este switch se encuentra el tráfico de gestión a la red por medio de la VLAN 73, además a través del enlace etherchannel se transmiten el tráfico de backup.

El almacenamiento en cinta se realiza por medio del TL2000 Dell, el cual es administrado actualmente por el servidor Dell Power Edge, como se muestra en la figura 3.37.

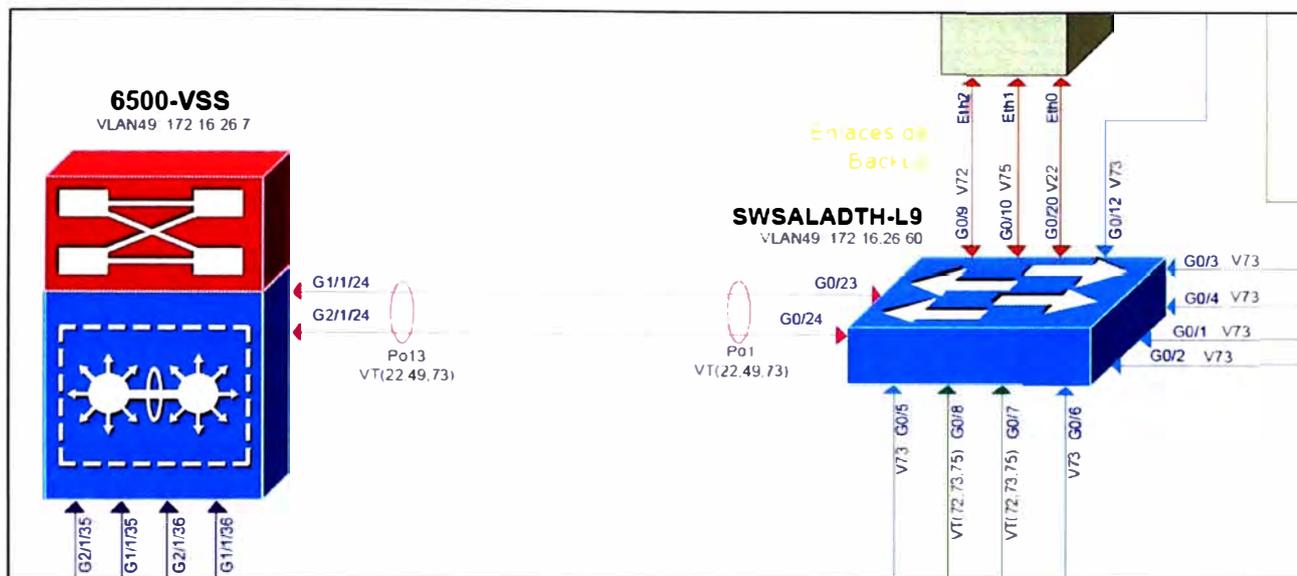


Fig. 3.37 Conectividad en la red de gestión y backup.

Para el direccionamiento IP en la red de gestión se utiliza el rango de IP proporcionado por el área de TI, la tabla 3.22 detalla este direccionamiento para cada uno de los equipos presentes en el diseño.

TABLA N° 3.22 Direccionamiento IP en la red de gestión.

Equipo	Nombre	IP	Mascara	Gateway	Interface
Dell M1000E	CMC	172.16.73.3	255.255.255.0	172.16.73.1	Gb
Cuchilla	Server 2008	172.16.73.4	255.255.255.0	172.16.73.1	Bond0 (Eth0, Eth1)
WS-CBS3032-DEL-F	A1	172.16.73.5	255.255.255.0	172.16.73.1	Vlan 73
WS-CBS3032-DEL-F	A2	172.16.73.6	255.255.255.0	172.16.73.1	Vlan 73
WS-CBS3032-DEL-F	B1	172.16.73.7	255.255.255.0	172.16.73.1	Vlan 73
WS-CBS3032-DEL-F	B2	172.16.73.8	255.255.255.0	172.16.73.1	Vlan 73
M5424 SAN I/O	M5424_Fabric1	172.16.73.9	255.255.255.0	172.16.73.1	F0
M5424 SAN I/O	M5424_Fabric2	172.16.73.10	255.255.255.0	172.16.73.1	F0
Cisco MDS 9509	MDS9509Lurin	172.16.73.11	255.255.255.0	172.16.73.1	Mgt0, Mgt1
Hitachi AMS2500	Controller 1	172.16.73.12	255.255.255.0	172.16.73.1	Con1
Hitachi AMS2500	Controller 2	172.16.73.13	255.255.255.0	172.16.73.1	Con2

-	Tape	172.16.73.14	255.255.255.0	172.16.73.1	Mgt
PowerEdge R610	OpManager	172.16.72.12	255.255.255.12	Rutas	F1.72
		172.16.75.3	255.255.255.0	Rutas	F1.75
WS-C2960G-24TC-L	SWSALADT H-L9	172.16.26.60	255.255.254.0	172.16.26.1	Vlan 73

h. Storage Hitachi AMS2500

El AMS2500 está compuesto por dos controladores que administran el flujo de información entre el servidor y los discos, asignando dos caminos en caso uno de ellos falle. En arquitecturas tradicionales los controladores se gestionan en modo activo – pasivo, con el avance de la virtualización se duplica la velocidad, para ello el Hitachi AMS2500 integra los nuevos controladores simétricos como se muestra en la figura 3.38.

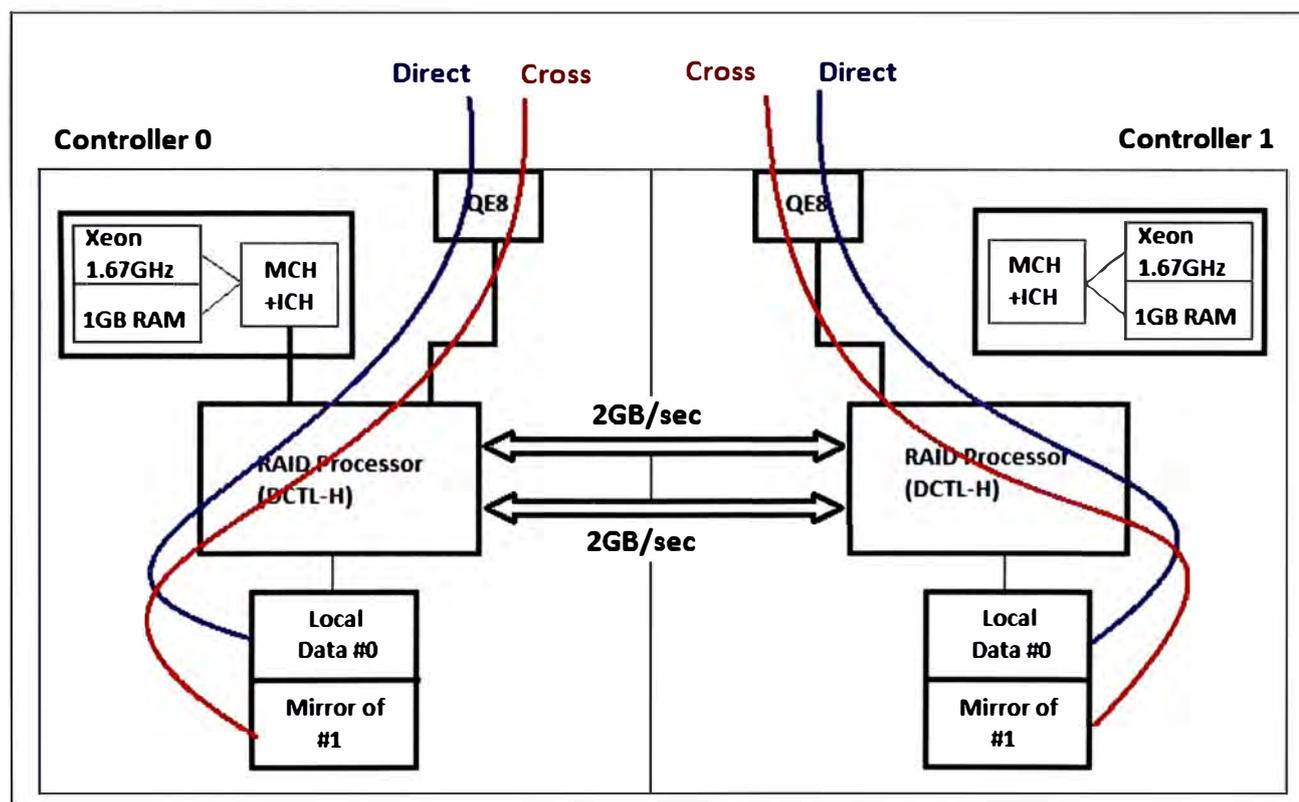


Fig. 3.38 Controlador simétrico en modo activo – activo.

Para la arquitectura desarrollada se usa ambos controladores para proveer una mejor velocidad de acceso entre el MDS 9509 y el AMS2500. La interconexión de los controladores y los servidores se realiza a través del switch director MDS9509, para el medio de transmisión se usa fiber channel con patch cords de fibra óptica a 8Gbps. La conexión se muestra en la figura 3.39.

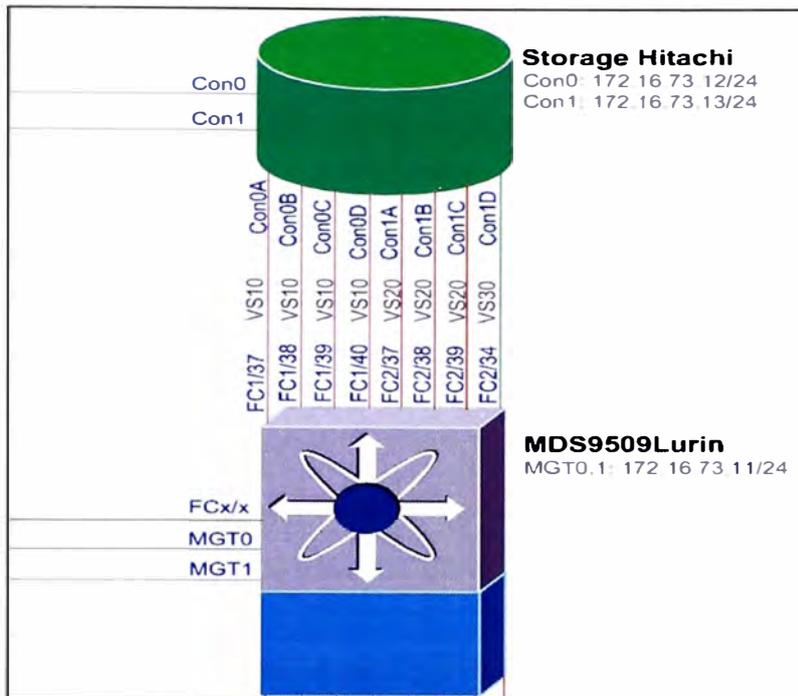


Fig. 3.39 Conexión entre el AMS2500 y el MDS9509.

La administración del storage modular es por medio del software de gestión HSNM2 (Hitachi storage navigator modular 2), y se accede a cualquiera de las IPs de gestión configuradas (<http://172.16.73.4:23015/StorageNavigatorModular/Login>), figura 3.40.

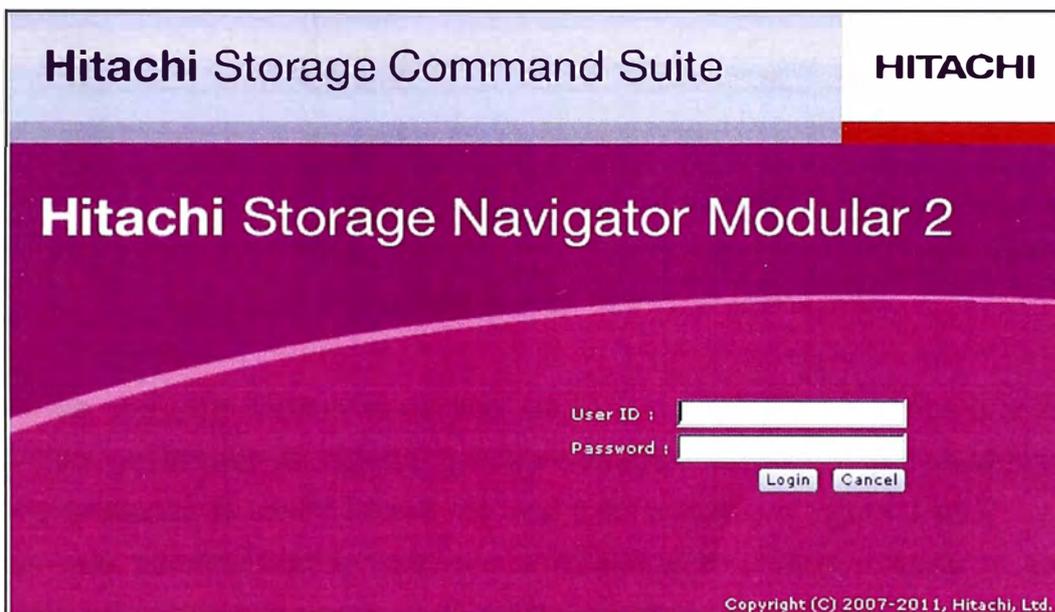


Fig. 3.40 Consola de administración HSNM2.

Dentro del sistema se muestra los array disponibles. En este caso solo se tiene uno, de esta manera se ingresa al AMS2500 disponible en la lista de arrays, figura 3.41.



Fig. 3.41 Arrays disponibles en la consola de gestión HSNM2.

Dentro del array se aprecia las características de los controladores, también una serie de opciones para la creación de nuevos arreglos de discos, figura 3.42.

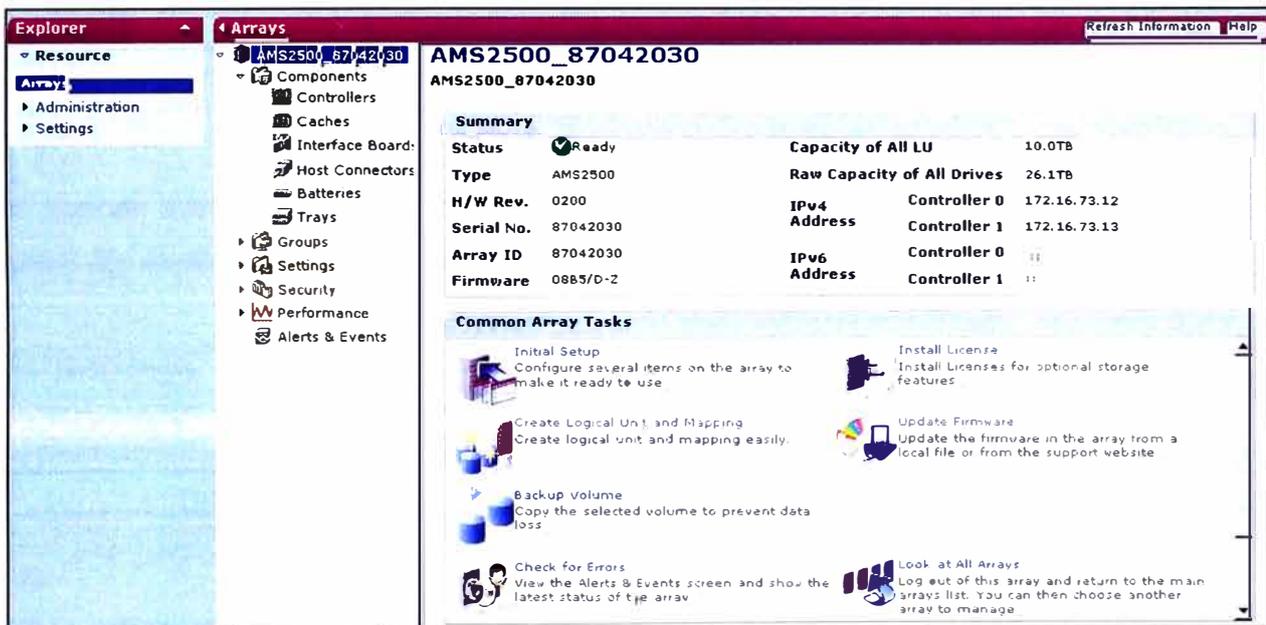


Fig. 3.42 Wizard de configuración en el HSNM2.

En la parte de unidades lógicas se aprecia las LUN o volúmenes creados que luego son agrupados en un arreglo de discos. Estos volúmenes luego son presentados al MDS 9509 donde se realiza la zonificación a los host SAN respectivo, figura 3.43.

El entorno web permite realizar diversas actividades de mantenimiento al arreglo de discos, entre estos están la creación de LUN, formateo de particiones, agrupamiento de disco del tipo RAID en todas sus variedades, también es posible la creación de alarmas de estado, configuración de balanceo de carga entre las controladoras SAN, es posible medir el performance de inserción y extracción de data mediante una aplicación web en tiempo real.

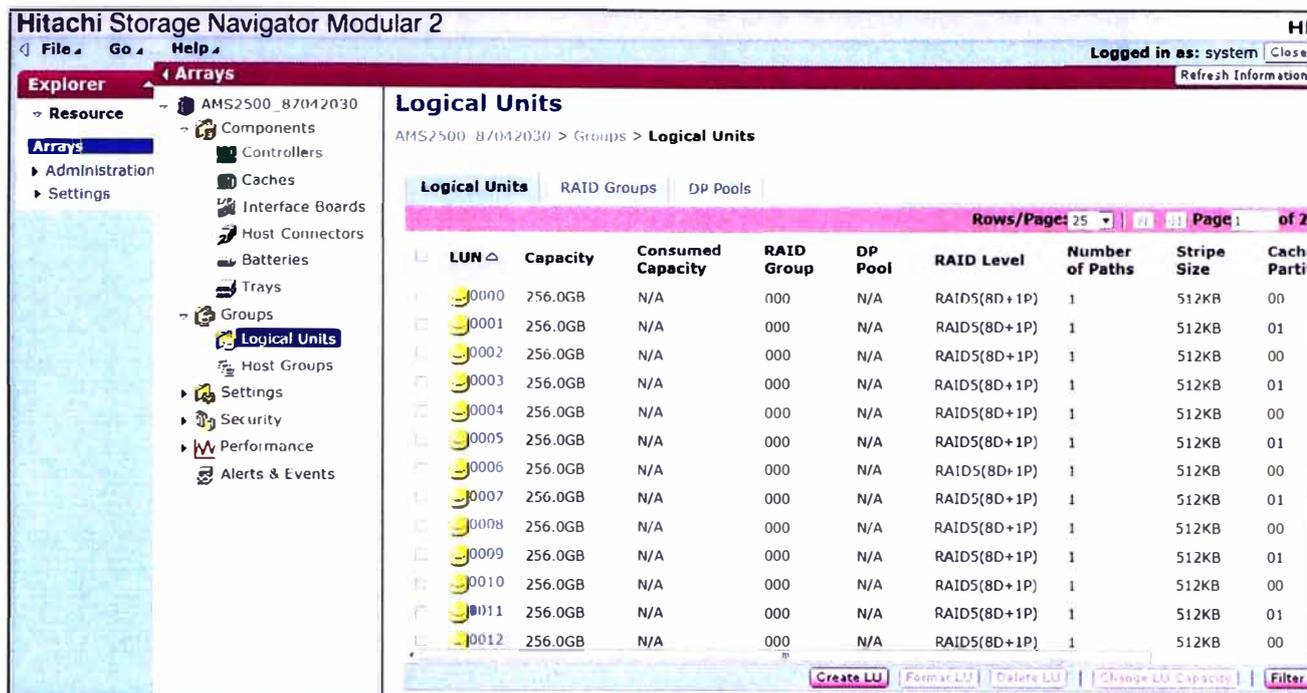


Fig. 3.43 LUNs existentes en el storage AMS2500.

Se conoce a los host ports como los puertos de cada controlador y son nombrados de la forma siguiente: 0A, 0B, 0C, 0D, 1A, 1B, 1C y 1D, se muestra en la figura 3.44.

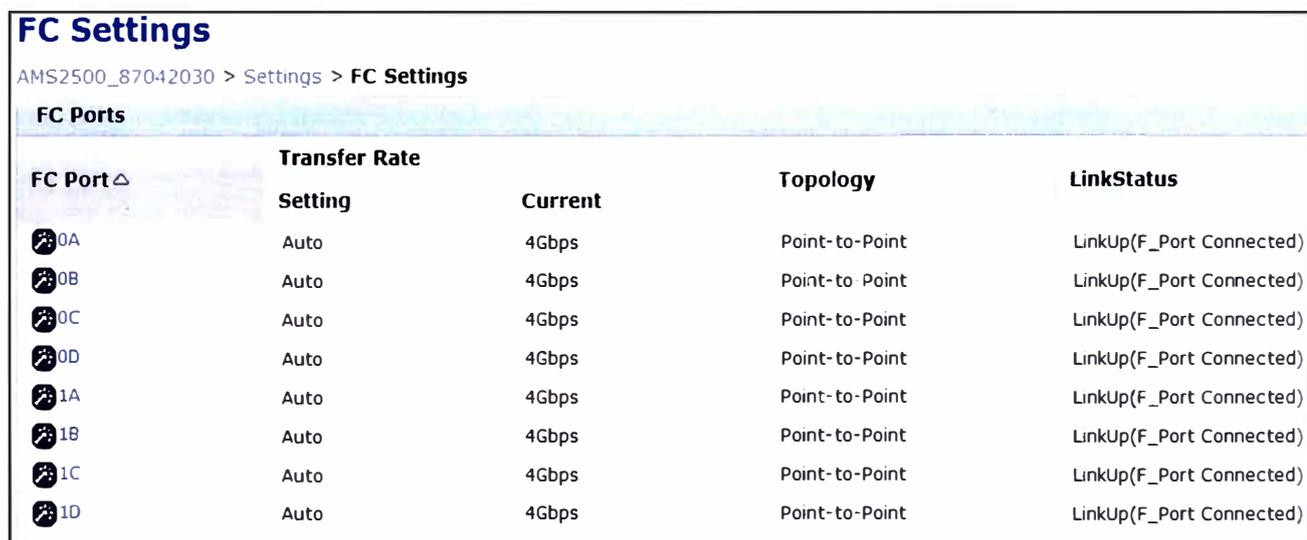


Fig. 3.44 Puertos FC del storage AMS2500.

Estos puertos contienen las LUN alcanzables para el MDS 9509 de acuerdo a la asignación de LUN realizado, en la imagen 3.45 se muestra el host group de nombre Cluster_MNLA_0 y las LUN asignada, figura 3.46. Los Host groups permiten seleccionar las WWNs que podrán acceder al arreglo de discos, figura 3.45, 3.46 y 3.47.

Host Group Property

Enter the information for the host group to be created.

Host Group No.: 001 * Edit to: **Available Ports**

* Name:
 32 characters or less (alphanumeric characters, '!', '#', '\$', '%', '&', "'", '+', '-', '.', ':', '=', '@', '^', '_', '{', '}', '~', '(', ')', '[', ']' or ').

Options:
 Platform: Linux
 Middleware: not specified

Available Ports

- Port
- 0A
- 0B
- 0C
- 0D

Fig. 3.45 Host Group con nombre Cluster_MNLA_0.

Assigned Logical Units

Rows/Page: 25 | Page 1 of 1

<input type="checkbox"/>	H-LUN <input type="checkbox"/>	LUN	Capacity	RAID Group	DP Pool
<input type="checkbox"/>	0000	0016	2.0GB	N/A	000
<input type="checkbox"/>	0001	0017	2.0GB	N/A	000
<input type="checkbox"/>	0002	0018	2.0GB	N/A	000
<input type="checkbox"/>	0003	0019	256.0GB	N/A	000
<input type="checkbox"/>	0004	0020	256.0GB	N/A	000
<input type="checkbox"/>	0005	0021	256.0GB	N/A	000
<input type="checkbox"/>	0006	0022	256.0GB	N/A	000
<input type="checkbox"/>	0007	0023	512.0GB	N/A	000
<input type="checkbox"/>	0008	0024	512.0GB	N/A	000
<input type="checkbox"/>	0009	0025	512.0GB	N/A	000
<input type="checkbox"/>	0010	0026	512.0GB	N/A	000
<input type="checkbox"/>	0011	0039	2.0GB	N/A	000
<input type="checkbox"/>	0012	0040	2.0GB	N/A	000
<input type="checkbox"/>	0013	0041	2.0GB	N/A	000

Fig. 3.46 LUN asignadas al host group para el cluster Oracle RAC.

Selected WWNs

Rows/Page: 25 | Page 1 of 1

<input type="checkbox"/>	Nickname	Port Name	Port <input type="checkbox"/>
<input type="checkbox"/>		2001848F69299E10	0A
<input type="checkbox"/>		2001848F69299E1D	0A
<input type="checkbox"/>		2001848F69299E2A	0A
<input type="checkbox"/>		2001848F69299E37	0A

Fig. 3.47 Host SAN permitidos al arreglo del Cluster_MNLA_0.

i. Implementación del Oracle RAC 11gR2

Para construir la plataforma 11g en un ambiente clúster de Base de datos se instaló el sistema operativo ORACLE LINUX 5.8, figura 3.48.

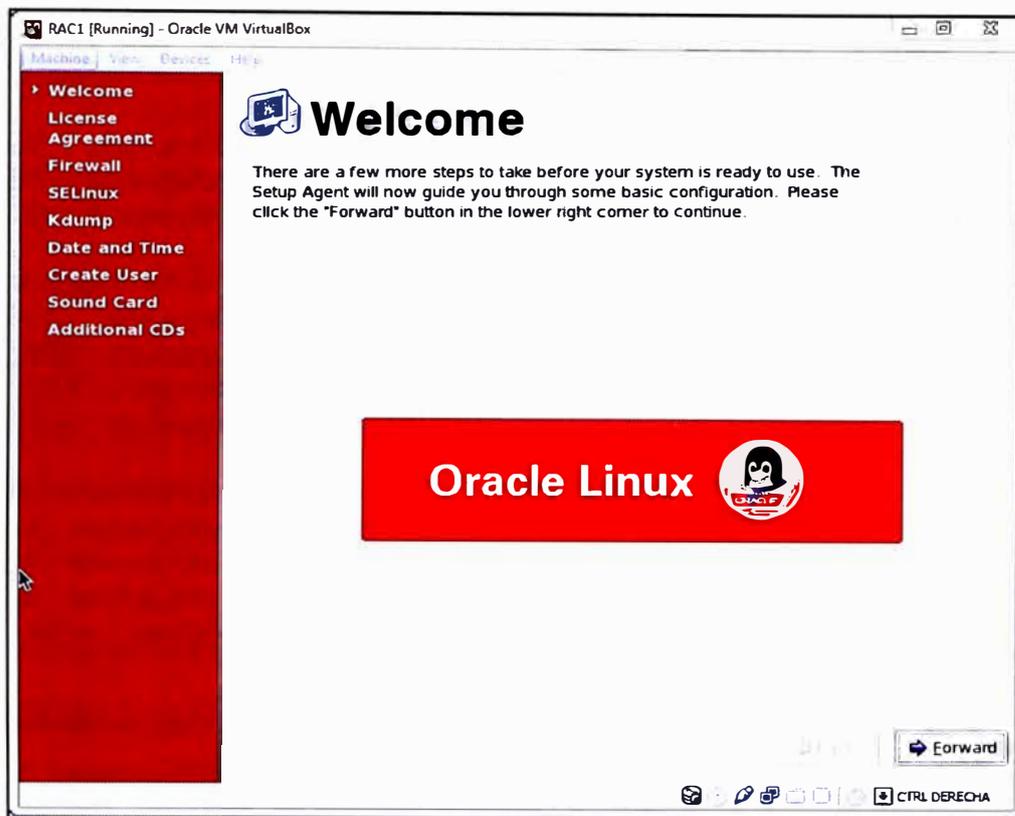


Fig. 3.48 Pantalla de bienvenida – Instalación Oracle linux 5u8.

Se procede a la instalación de los RPM que Oracle solicita en la preinstalación, se muestra los RPM necesarios en la tabla 3.23.

TABLA N° 3.23 Paquetes RPMs instalados en Oracle linux.

```
[oracle@sriwaybdprod001 ~]$ rpm -qa | grep binutils
binutils-2.17.50.0.6-20.el5
[oracle@sriwaybdprod001 ~]$ rpm -qa | grep elfutils
elfutils-libelf-devel-static-0.137-3.el5
elfutils-0.137-3.el5
elfutils-libelf-devel-0.137-3.el5
elfutils-libelf-0.137-3.el5
elfutils-libs-0.137-3.el5
elfutils-libelf-0.137-3.el5
```

Se realiza la creación de un directorio donde se asocia el software del clúster y de la base de datos (**/u01**). Se configura el reconocimiento de IP públicas, IP privadas y enlaces virtuales como se muestra en la tabla 3.24.

TABLA N° 3.24 Resolución de IPs en el nodo 1 – Oracle RAC.

```
[oracle@sriwaybdprod001 ~]$ cat /etc/hosts
# Do not remove the following line, or various programs
# that require network functionality will fail.
127.0.0.1          localhost.localdomain localhost

##### PUBLICA #####
172.16.72.3      sriwaybdprod001.medianetworks.pe  sriwaybdprod001
172.16.72.5      sriwaybdprod002.medianetworks.pe  sriwaybdprod002
172.16.72.7      sriwaybdprod003.medianetworks.pe  sriwaybdprod003
172.16.72.9      sriwaybdprod004.medianetworks.pe  sriwaybdprod004

##### PRIVADA #####
172.16.72.131   sriwaybdprod001-priv.medianetworks.pe sriwaybdprod001-priv
172.16.72.132   sriwaybdprod002-priv.medianetworks.pe sriwaybdprod002-priv
172.16.72.133   sriwaybdprod003-priv.medianetworks.pe sriwaybdprod003-priv
172.16.72.134   sriwaybdprod004-priv.medianetworks.pe sriwaybdprod004-priv

##### VIP #####
172.16.72.4      sriwaybdprod001-vip.medianetworks.pe sriwaybdprod001-vip
172.16.72.6      sriwaybdprod002-vip.medianetworks.pe sriwaybdprod002-vip
172.16.72.8      sriwaybdprod003-vip.medianetworks.pe sriwaybdprod003-vip
172.16.72.10     sriwaybdprod004-vip.medianetworks.pe sriwaybdprod004-vip
```

Se crea 2 usuarios para la gestión de la base de datos

- Grid, usuario que administrara exclusivamente el clúster de base de datos.
- Oracle, usuario que administrara exclusivamente la base de datos.

Se implementa el registro de llaves RSA 1024, de esta manera los nodos no necesitan credenciales para acceder a los recursos de otros nodos, tabla 3.25.

TABLA N° 3.25 Llaves publicas RSA usado entre los nodos del Oracle RAC.

```
[oracle@sriwaybdprod001 ~]$ cat .ssh/authorized_keys
ssh-rsa
AAAAB3NzaC1yc2EAAAADAQABAAQCh7wkVEeN7cveAerqBonMDr/gt7aA7veuUJ
a5CHKr+0uogk+ltNai36Tnlppl/G9+Tz325OrX6gUvwS9BpQWNripBildq3GCQDXjaQFB5
K88YekGhjfdpHMbOtBhxTC3C5/Aiw+L8dlrWb0Z0gu8MygmxFHX1R+DILr0gLdxZ+yw==
Keys generated using jsch
ssh-rsa
AAAAB3NzaC1yc2EAAAADAQABAAQCh7wkVEeN7cveAerqBonMDr/gt7aA7veuUJ
a5CHKr+0uogk+ltNai36Tnlppl/G9+Tz325OrX6gUvwS9BpQWNripBildq3GCQDXjaQFB5
K88YekGhjfdpHMbOtBhxTC3C5/Aiw+L8dlrWb0Z0gu8MygmxFHX1R+DILr0gLdxZ+yw==
Keys generated using jsch
ssh-rsa
AAAAB3NzaC1yc2EAAAABlWAAIAEA4mbM9uAJYDJA4X/5zCKO271S9ojybLJ2M3F+W
LHk2Rvkt8QnSo/HcWqcDzeILzWv1JatJQTjFf1L4ZBhCtLK6j4SDegIWW7QfsIKJCLfLJe0
UlsalwZ75IAfTINHjPjOEOjdl8TSnBHAgHmJjGibmpnJaTfQ6HedETyalgRiq6N0=
```

```
oracle@srviwaybdprod001.medianetworks.pe
ssh-rsa
AAAAB3NzaC1yc2EAAAABIwAAAIEAxGgFwPqJRBn5+R0s0/bPgUP5kGUHJPCvtCPs
dz7dBfilowwRuB+HJt+dDCD8H7RRmCH4jvzhyX+8XJYXCXtQbYcJNZrb5H9j+11B7nK
QIFDdpix88Jo23udJT0AI7KkCMebB6+mCxA+lfABtni88+otNiF0rv7lItJ2vGRO7jk=
oracle@srviwaybdprod003.medianetworks.pe
ssh-rsa
AAAAB3NzaC1yc2EAAAABIwAAAIEA8G3I6C1X/R5+PTnKCdQyy7wMT3yUayT17JwFq
MDnaEpc5OMSybuTTBniIPWJoHtldzQjJtAnSHDuwyjYdJKjaYnwiLZQG2d72Dw4OcSXB
6B+aCPK6YW8Ulfqn0utL8TsPc3shWcdcjdk+er2OjPW/0wV4I6mqlt/4IMsODEkts=
oracle@srviwaybdprod002.medianetworks.pe
ssh-rsa
AAAAB3NzaC1yc2EAAAABIwAAAIEAqNMKri2jky+mY7x/Yc0wUO4ys1I3U7J9fXhm+8M
5UkrVmfqd2Z2TaqfqOltjpR8p8L1+yy+HG4n+SgyhbswP29cZn+ndgFwVEK7oMwk0iTZU
rMPG277YMv2YDg67eFMaGNIJCXDikzrEv6dfgxYY+RWJ1o7mulUgHKx2NiUWOXU=
oracle@srviwaybdprod004.medianetworks.pe
```

Para la instalación en una plataforma RAC en versión 11g se instala una nueva tecnología en Oracle llamada “infraestructura grid” el cual consiste en un software que contiene todas las herramientas y tecnologías para la administración de la base de datos. Antes de la instalación de grid es necesario que el servidor DNS resuelva 3 direcciones IP mediante el nombre **media-scan.medianetworks.pe** a este se conoce como el “Scan Listener”, esto es nuevo en la versión 11gR2 de Oracle y se usa para que las aplicaciones se conecten solo a este nombre, así se realiza la conexión de los nodos de la base de datos, ya no se usa la direcciones IP virtuales, se muestra en la figura 3.49.

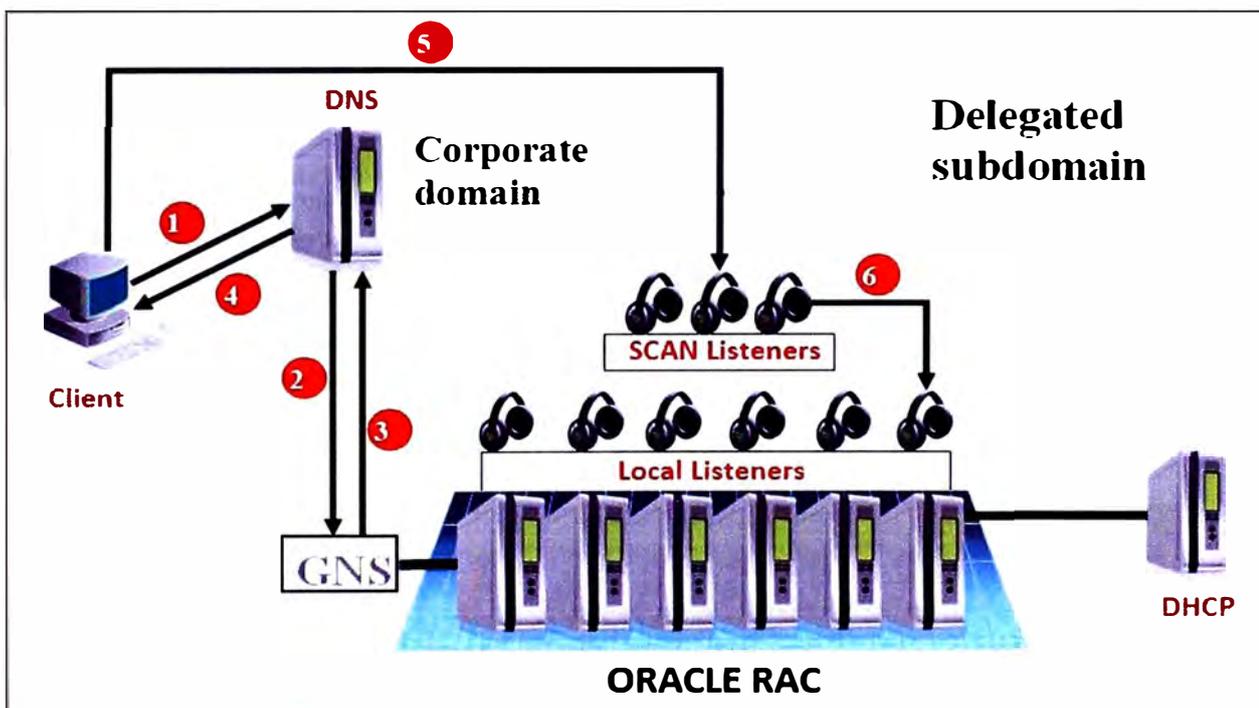


Fig. 3.49 Funcionamiento del SCAN listener.

El SCAN listener correrán en solo 3 nodos diferentes, tabla 3.26.

TABLA N° 3.26 Listener activos en los nodos 1, 2 y 3.

```
SCAN Listener LISTENER_SCAN1 is enabled
SCAN Listener LISTENER_SCAN1 is running on node srviwaybdprod003
SCAN Listener LISTENER_SCAN2 is enabled
SCAN Listener LISTENER_SCAN2 is running on node srviwaybdprod002
SCAN Listener LISTENER_SCAN3 is enabled
SCAN Listener LISTENER_SCAN3 is running on node srviwaybdprod001
```

Una vez que se activa los scan listener se procede a lanzar el instalador de la infraestructura grid y así se crea los recursos conformados por la base de datos, luego de la instalación se revisa al estado de los recursos como se muestra en la tabla 3.27

TABLA N° 3.27 Revisión del CRS (Verificando la disponibilidad de los servicios).

```
[grid@srviwaybdprod001 ~]$ crsctl check crs
CRS-4638: Oracle High Availability Services is online
CRS-4537: Cluster Ready Services is online
CRS-4529: Cluster Synchronization Services is online
CRS-4533: Event Manager is online
```

Después se presenta la LUN desde el storage y se configura el proceso de redundancia llamado "multipath", en la tabla 3.28 se muestra las LUN asignadas al nodo 1.

TABLA N° 3.28 Multipath en el nodo 1 del Oracle RAC.

```
mpath2 (360060e80100afe60053027ee00000011) dm-3 HITACHI,DF600F
size=2.0G features='1 queue_if_no_path' hwhandler='0' wp=rw
|-+- policy='round-robin 0' prio=1 status=active
|`- 1:0:0:1 sdc 8:32 active ready running
`-+- policy='round-robin 0' prio=0 status=enabled
  `- 2:0:0:1 sdq 65:0 active ready running
mpath1 (360060e80100afe60053027ee00000010) dm-2 HITACHI,DF600F
size=2.0G features='1 queue_if_no_path' hwhandler='0' wp=rw
|-+- policy='round-robin 0' prio=1 status=active
|`- 2:0:0:0 sdp 8:240 active ready running
`-+- policy='round-robin 0' prio=0 status=enabled
  `- 1:0:0:0 sdb 8:16 active ready running
mpath9 (360060e80100afe60053027ee00000018) dm-10 HITACHI,DF600F
size=512G features='1 queue_if_no_path' hwhandler='0' wp=rw
|-+- policy='round-robin 0' prio=1 status=active
|`- 2:0:0:8 sdx 65:112 active ready running
`-+- policy='round-robin 0' prio=0 status=enabled
  `- 1:0:0:8 sdj 8:144 active ready running
mpath8 (360060e80100afe60053027ee00000017) dm-9 HITACHI,DF600F
size=512G features='1 queue_if_no_path' hwhandler='0' wp=rw
```

```
|+- policy='round-robin 0' prio=1 status=active
|`- 1:0:0:7 sdi 8:128 active ready running
`+- policy='round-robin 0' prio=0 status=enabled
  `- 2:0:0:7 sdw 65:96 active ready running
mpath14 (360060e80100afe60053027ee00000028) dm-15 HITACHI,DF600F
size=2.0G features='1 queue_if_no_path' hwhandler='0' wp=rw
|+- policy='round-robin 0' prio=1 status=active
|`- 2:0:0:12 sdab 65:176 active ready running
`+- policy='round-robin 0' prio=0 status=enabled
  `- 1:0:0:12 sdn 8:208 active ready running
mpath7 (360060e80100afe60053027ee00000016) dm-8 HITACHI,DF600F
size=256G features='1 queue_if_no_path' hwhandler='0' wp=rw
|+- policy='round-robin 0' prio=1 status=active
|`- 2:0:0:6 sdv 65:80 active ready running
`+- policy='round-robin 0' prio=0 status=enabled
  `- 1:0:0:6 sdh 8:112 active ready running
mpath13 (360060e80100afe60053027ee00000029) dm-14 HITACHI,DF600F
size=2.0G features='1 queue_if_no_path' hwhandler='0' wp=rw
|+- policy='round-robin 0' prio=1 status=active
|`- 1:0:0:13 sdo 8:224 active ready running
`+- policy='round-robin 0' prio=0 status=enabled
  `- 2:0:0:13 sdac 65:192 active ready running
mpath6 (360060e80100afe60053027ee00000015) dm-7 HITACHI,DF600F
size=256G features='1 queue_if_no_path' hwhandler='0' wp=rw
|+- policy='round-robin 0' prio=1 status=active
|`- 1:0:0:5 sdg 8:96 active ready running
`+- policy='round-robin 0' prio=0 status=enabled
  `- 2:0:0:5 sdu 65:64 active ready running
mpath12 (360060e80100afe60053027ee00000027) dm-13 HITACHI,DF600F
size=2.0G features='1 queue_if_no_path' hwhandler='0' wp=rw
|+- policy='round-robin 0' prio=1 status=active
|`- 1:0:0:11 sdm 8:192 active ready running
`+- policy='round-robin 0' prio=0 status=enabled
  `- 2:0:0:11 sdaa 65:160 active ready running
mpath5 (360060e80100afe60053027ee00000014) dm-6 HITACHI,DF600F
size=256G features='1 queue_if_no_path' hwhandler='0' wp=rw
|+- policy='round-robin 0' prio=1 status=active
|`- 2:0:0:4 sdt 65:48 active ready running
`+- policy='round-robin 0' prio=0 status=enabled
  `- 1:0:0:4 sdf 8:80 active ready running
mpath11 (360060e80100afe60053027ee0000001a) dm-12 HITACHI,DF600F
size=512G features='1 queue_if_no_path' hwhandler='0' wp=rw
|+- policy='round-robin 0' prio=1 status=active
|`- 2:0:0:10 sdz 65:144 active ready running
`+- policy='round-robin 0' prio=0 status=enabled
  `- 1:0:0:10 sdl 8:176 active ready running
mpath4 (360060e80100afe60053027ee00000013) dm-5 HITACHI,DF600F
size=256G features='1 queue_if_no_path' hwhandler='0' wp=rw
|+- policy='round-robin 0' prio=1 status=active
|`- 1:0:0:3 sde 8:64 active ready running
`+- policy='round-robin 0' prio=0 status=enabled
  `- 2:0:0:3 sds 65:32 active ready running
mpath10 (360060e80100afe60053027ee00000019) dm-11 HITACHI,DF600F
```

```

size=512G features='1 queue_if_no_path' hwhandler='0' wp=rw
|-+- policy='round-robin 0' prio=1 status=active
|`- 1:0:0:9 sdk 8:160 active ready running
`-+- policy='round-robin 0' prio=0 status=enabled
  `- 2:0:0:9 sdy 65:128 active ready running
mpath3 (360060e80100afe60053027ee00000012) dm-4 HITACHI,DF600F
size=2.0G features='1 queue_if_no_path' hwhandler='0' wp=rw
|-+- policy='round-robin 0' prio=1 status=active
|`- 2:0:0:2 sdr 65:16 active ready running
`-+- policy='round-robin 0' prio=0 status=enabled
  `- 1:0:0:2 sdd 8:48 active ready running
    
```

Luego de presentarse las LUN a los 4 nodos de la base de datos, se procede a configurar los archivos Voting File y OCR propios de la infraestructura grid.

En la tabla 3.29 se observa la distribución de 3 raw devices (archivos para acceder directamente a disco sin usar el sistema operativo).

TABLA N° 3.29 Estado del actual OCR file.

```

Status of Oracle Cluster Registry is as follows:
  Version                          :                3
  Total space (kbytes)              :            262120
  Used space (kbytes)               :             3248
  Available space (kbytes)         :            258872
  ID                                :          1494582677
  Device/File Name                  :          +DATA
                                     Device/File integrity check succeeded
  Device/File Name                  :    /dev/raw/raw2
                                     Device/File integrity check succeeded
  Device/File Name                  :    /dev/raw/raw3
                                     Device/File integrity check succeeded
                                     Device/File not configured
                                     Device/File not configured
  Cluster registry identity check succeeded
    
```

En el Voting File se observa que también se distribuye en 3 raw devices para su contingencia como se muestra en la tabla 3.30.

TABLA N° 3.30 Estado de los raw devices.

#	STATE	File Universal Id	File name	Disk group
1.	ONLINE	077ae696a4f64f85bf448e9b03038f4a	(/dev/raw/raw12)	[DG_VOTING_OCR]
1.	ONLINE	7ea15df69a156fd9863b31c49e97a46	(/dev/raw/raw13)	[DG_VOTING_OCR]
1.	ONLINE	301b313f3e7b4f04bf22fee07e1540b	(/dev/raw/raw14)	[DG_VOTING_OCR]

Luego se montan los grupos de discos para su administración en ASM, tabla 3.31 y 3.32.

TABLA N° 3.31 Grupos de disco montados al ASM.

State	Type	Rebal	Name
MOUNTED	EXTERN	N	FRA/
MOUNTED	EXTERN	N	DG_VOTING_OCR/
MOUNTED	EXTERN	N	DATA1/
MOUNTED	EXTERN	N	DATA/

TABLA N° 3.32 Resumen de DISKGROUPS creados.

Grupo de Discos	Descripción
FRA/	Forma parte del repositorio de backups.
DG_VOTING_OCR/	Forma parte del repositorio de los VotingFile.
DATA/	Almacena el OCR.
DATA1/	Repositorio para los datafiles, spfile y controlfile

Por último se inicia la instalación de la base de datos, luego se revisa nodo por nodo la funcionalidad de los demonios (procesos de servidor), tabla 3.33.

TABLA N° 3.33 Status de procesos PMON en los nodos Oracle RAC.

```
[oracle@srviwaybdprod001 ~]$ ps -ef | grep pmon
grid 6978 1 0 Oct10 ? 00:14:30 asm_pmon_+ASM1
oracle 9691 5088 0 14:05 pts/1 00:00:00 grep pmon
oracle 9977 1 0 Oct26 ? 00:14:12 ora_pmon_MNLA1

[oracle@srviwaybdprod002 ~]$ ps -ef | grep pmon
oracle 10797 1 0 Oct26 ? 00:14:07 ora_pmon_MNLA2
grid 15539 1 0 Oct25 ? 00:10:31 asm_pmon_+ASM2
oracle 19614 19585 0 14:06 pts/1 00:00:00 grep pmon

[oracle@srviwaybdprod003 ~]$ ps -ef | grep pmon
grid 7324 1 0 Nov08 ? 00:06:52 asm_pmon_+ASM3
oracle 7763 1 0 Nov08 ? 00:09:04 ora_pmon_MNLA3
oracle 23399 23364 0 14:05 pts/1 00:00:00 grep pmon

[oracle@srviwaybdprod004 ~]$ ps -ef | grep pmon
oracle 10670 1 0 Oct26 ? 00:14:34 ora_pmon_MNLA4
oracle 13265 13235 0 14:07 pts/1 00:00:00 grep pmon
grid 14944 1 0 Oct25 ? 00:10:30 asm_pmon_+ASM4
```

Las 4 instancias de la base de datos se encuentran operativas, tabla 3.34 y 3.35.

TABLA N° 3.34 Status de los nodos RAC.

```
[oracle@srviwaybdprod001 ~]$ srvctl status database -d MNLA
Instance MNLA1 is running on node srviwaybdprod001
Instance MNLA2 is running on node srviwaybdprod002
Instance MNLA3 is running on node srviwaybdprod003
```

Instance MNLA4 is **running** on node sriwaybdprod004

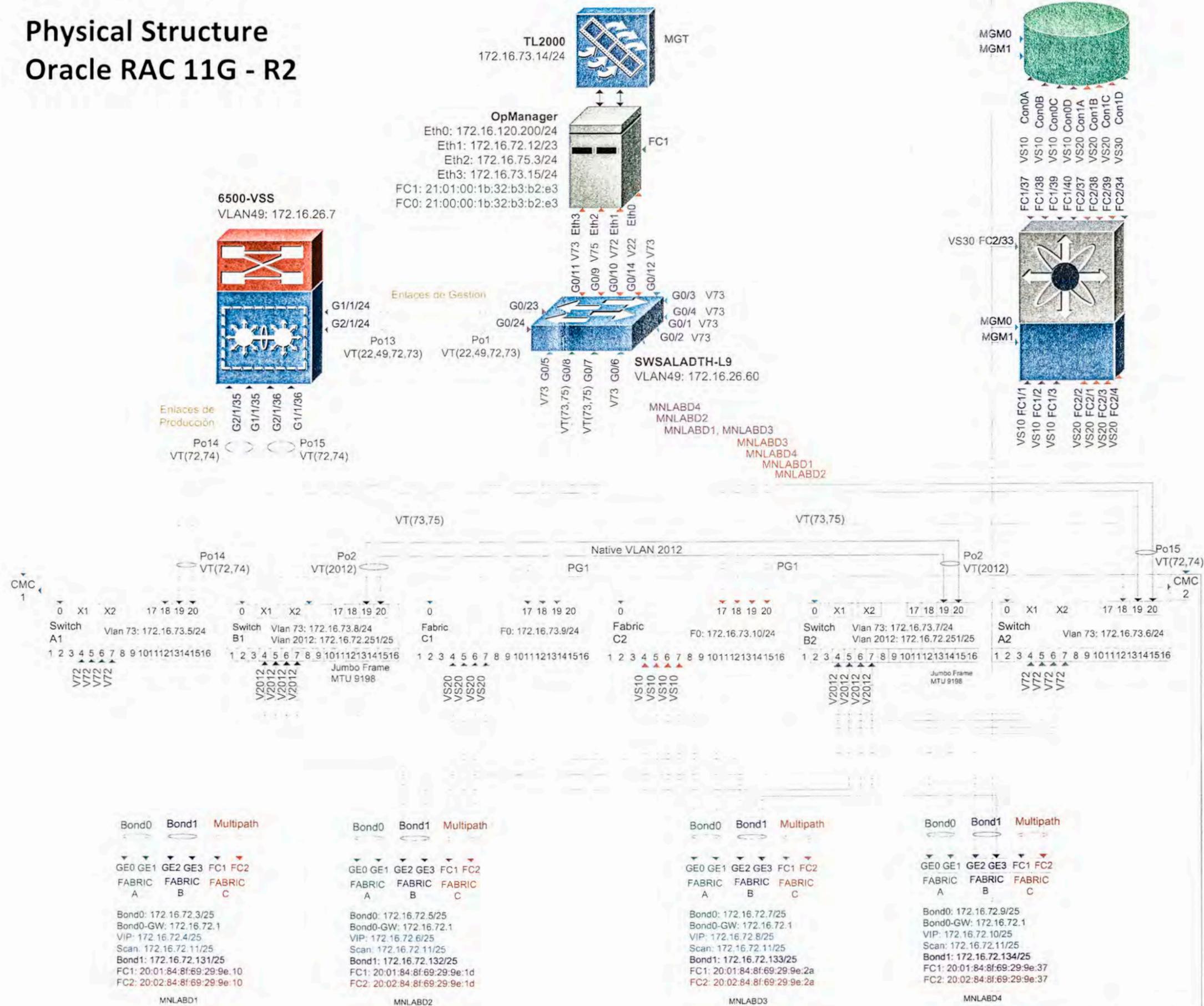
TABLA N° 3.35 Respuesta de actividad a la base de datos.

```
[oracle@sriwaybdprod001 ~]$ tnsping MNLA
Used TNSNAMES adapter to resolve the alias
Attempting to contact (DESCRIPTION = (ADDRESS = (PROTOCOL = TCP)(HOST =
media-scan.medianetworks.pe)(PORT = 1521)) (CONNECT_DATA = (SERVER =
DEDICATED) (SERVICE_NAME = MNLA)))
OK (10 msec)
```

j. Arquitectura de Red General

En la página 71 se muestra el diagrama de la arquitectura general.

Physical Structure Oracle RAC 11G - R2



Storage Hitachi
MGM0: 172.16.73.12/24
MGM1: 172.16.73.13/24
[AMS2500_OA] 50:06:0e:80:10:0a:fe:60
[AMS2500_OB] 50:06:0e:80:10:0a:fe:61
[AMS2500_OC] 50:06:0e:80:10:0a:fe:62
[AMS2500_OD] 50:06:0e:80:10:0a:fe:63
[AMS2500_1A] 50:06:0e:80:10:0a:fe:68
[AMS2500_1B] 50:06:0e:80:10:0a:fe:69
[AMS2500_1C] 50:06:0e:80:10:0a:fe:6a
[AMS2500_1D] 50:06:0e:80:10:0a:fe:6b

MDS9509Lurin
MGT0,1: 172.16.73.11/24
VSAN 10: VSAN_PROD_1 / 20:0a:54:7f:ee:03:3a:01
VSAN 20: VSAN_PROD_2 / 20:14:54:7f:ee:03:3a:01
VSAN 30: VSAN_BACKUP / 20:1e:54:7f:ee:03:3a:01

Zone AMS2500_OA_MNLABD1_1 vsan 10
20:01:84:8f:69:29:9e:10 [MNLABD1_1]
50:06:0e:80:10:0a:fe:60 [AMS2500_OA]

Zone AMS2500_OA_MNLABD2_1 vsan 10
20:01:84:8f:69:29:9e:1d [MNLABD2_1]
50:06:0e:80:10:0a:fe:60 [AMS2500_OA]

Zone AMS2500_OA_MNLABD3_1 vsan 10
20:01:84:8f:69:29:9e:2a [MNLABD3_1]
50:06:0e:80:10:0a:fe:60 [AMS2500_OA]

Zone AMS2500_OA_MNLABD4_1 vsan 10
20:01:84:8f:69:29:9e:37 [MNLABD4_1]
50:06:0e:80:10:0a:fe:60 [AMS2500_OA]

Zone AMS2500_OB_BLADE8_1 vsan 10
20:01:84:8f:69:29:9e:44 [BLADE8_1]
50:06:0e:80:10:0a:fe:61 [AMS2500_OB]

Zone AMS2500_OC_BLADE9_1 vsan 10
20:01:84:8f:69:29:9e:51 [BLADE9_1]
50:06:0e:80:10:0a:fe:62 [AMS2500_OC]

Zone AMS2500_OB_BLADE10_1 vsan 10
20:01:84:8f:69:29:9e:5e [BLADE10_1]
50:06:0e:80:10:0a:fe:61 [AMS2500_OB]

zone_backup vsan 30
21:01:00:1b:32:b3:b2:e3 [BACKUP_SRV]
50:06:0e:80:10:0a:fe:6b [AMS2500_1D]

Enclosure M1000e
Oracle RAC 11gR2

Design by: Christian Riveros	Title: Physical Structure Oracle RAC 11G - R2	
media networks	Location: Lurin, Lima - Perú	Versión: 1.5
	Unit: Gerencia de Seguridad y Monitoreo	31/05/2012

CAPITULO IV

ANALISIS Y PRESENTACION DE RESULTADOS

En este capítulo se muestra los resultados obtenidos en el diseño así como el rendimiento de los nodos de la base de datos Oracle RAC finalmente.

4.1 Análisis de performance.

Para medir la performance y capacidad de la base de datos, se crea un procedimiento que permite ingresar 50 millones de registros en la instancia MNLA, para tal objetivo será necesario la creación de una cuenta de prueba en la base de datos y asignación de privilegios en un determinado tablespace. La herramienta que se usa para medir el performance, es el Enterprise Manager, figura 4.1. Aquí se realiza la administración de todos los recursos de la base de datos. La manera de acceder es mediante el siguiente enlace web: <https://172.16.72.3:1158/em>.

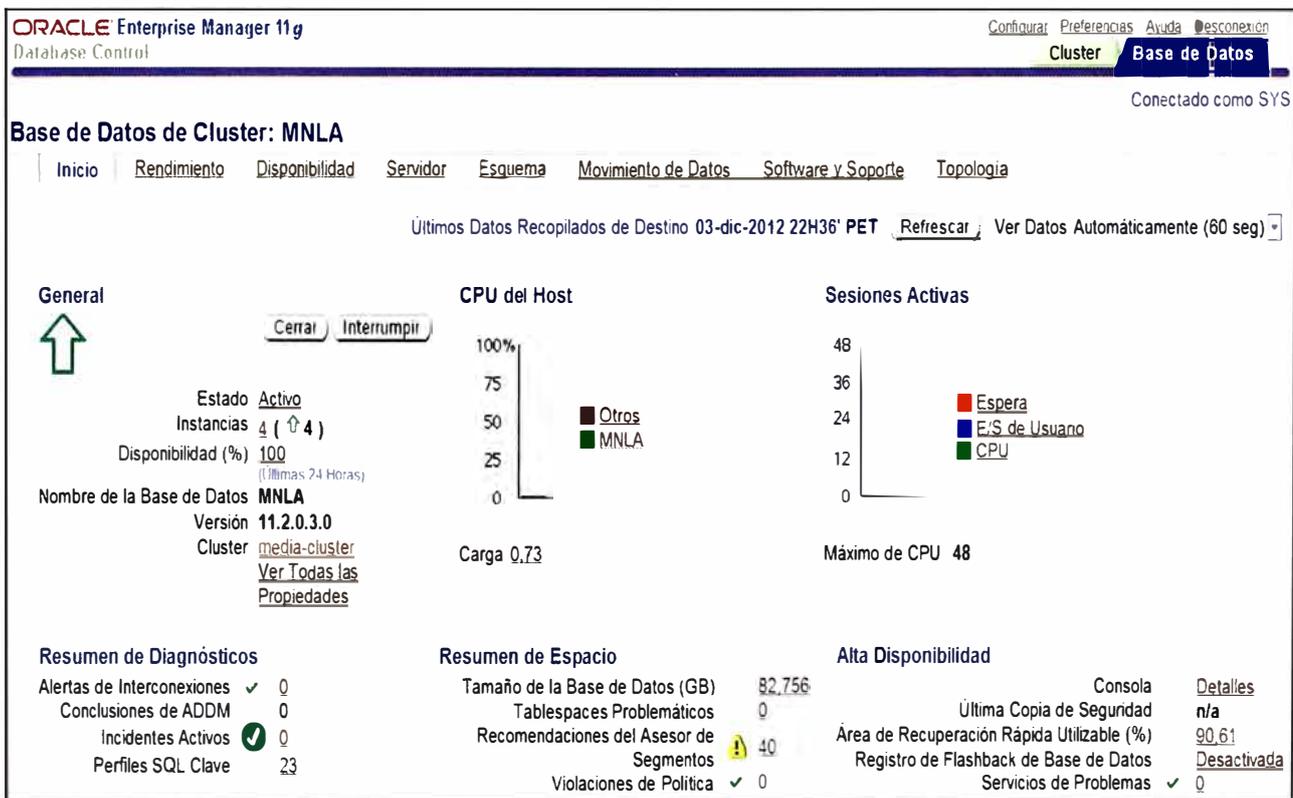


Fig. 4.1 Oracle enterprise manager 11g para base de datos.

Por medio de esta consola podemos observar parámetros ya configurados mediante línea de comandos, como se muestra en la figura 4.2.

Nombre	Tipo	Subred	Tipo de Interfaz
media-cluster	Cluster		
srviwaybdprod001.medianetworks.pe	Host		
bond0	Interfaz	172.16.72.0	Público
bond1	Interfaz	172.16.72.128	Privado
eth0	Interfaz	n/a	Desconocido
eth1	Interfaz	n/a	Desconocido
eth2	Interfaz	n/a	Desconocido
eth3	Interfaz	n/a	Desconocido
srviwaybdprod002.medianetworks.pe	Host		
bond0	Interfaz	172.16.72.0	Público
bond1	Interfaz	172.16.72.128	Privado
eth0	Interfaz	n/a	Desconocido
eth1	Interfaz	n/a	Desconocido
eth2	Interfaz	n/a	Desconocido
eth3	Interfaz	n/a	Desconocido
srviwaybdprod003.medianetworks.pe	Host		
bond0	Interfaz	172.16.72.0	Público
bond1	Interfaz	172.16.72.128	Privado
eth0	Interfaz	n/a	Desconocido
eth1	Interfaz	n/a	Desconocido
eth2	Interfaz	n/a	Desconocido
eth3	Interfaz	n/a	Desconocido
srviwaybdprod004.medianetworks.pe	Host		
bond0	Interfaz	172.16.72.0	Público
bond1	Interfaz	172.16.72.128	Privado
eth0	Interfaz	n/a	Desconocido
eth1	Interfaz	n/a	Desconocido
eth2	Interfaz	n/a	Desconocido
eth3	Interfaz	n/a	Desconocido

Fig. 4.2 Datos de red en la base de datos en el enterprise manager.

En la pestaña inicio del Enterprise Manager se muestra el estado operativo de los nodos de la base de datos, figura 4.3.

Inicio	Rendimiento	Destinos	Administración	Interconexiones	Topología
Hosts					
Nombre	Estado	Estado de Clusterware	Alertas	Violaciones de Política	Porcentaje de Cumplimiento
srviwaybdprod001.medianetworks.pe			0 0	5 2 0	70
srviwaybdprod002.medianetworks.pe			0 0	5 1 0	76
srviwaybdprod003.medianetworks.pe			0 0	5 1 0	76
srviwaybdprod004.medianetworks.pe			0 0	5 1 0	76

Fig. 4.3 Estado de los nodos Oracle RAC.

4.2 Creación del usuario de prueba.

Para acceder a la base de Datos usamos el lenguaje de programación SQL. Luego se ejecuta las sentencias mostradas en la tabla 4.1 para la creación del usuario de prueba.

TABLA N° 4.1 Sentencia SQL usada para la creación del usuario de prueba.

```
SQL> create tablespace prueba datafile '+DATA1/mnla/datafile/prueba.dbf' size 10G;
SQL> create user a identified by a default tablespace prueba quota unlimited on prueba;
SQL> grant dba to a;
SQL> conn a/a
SQL> create table b (data number (30)) tablespace prueba;
```

Con la ejecución de los siguientes comandos se asigna 10 Gigabyte al tablespace **prueba.dbf**, y se otorga privilegios de administrador al usuario **“a”**. Luego realizamos la conexión con la cuenta **“a”** para crear la tabla **“b”** donde se almacenaran los registros. En la figura 4.4 se observa la configuración realizada por los comandos de la tabla 4.1.

Seleccionar	Nombre	Tamaño Asignado (MB)	Espacio Usado (MB)	Espacio Usado Asignado (%)	Ampliación Automática
<input checked="" type="radio"/>	INTRAWAY	200,0	114,7	57,3	YES
<input type="radio"/>	INTRAWAY_INDX	50,0	1,0	2,0	YES
<input type="radio"/>	IWAYDATA	10.300,0	9.802,6	95,2	YES
<input type="radio"/>	IWAYDATA_INDX	100,0	1,1	1,1	YES
<input type="radio"/>	IWAYLOGS	34.368,0	13.235,0	38,5	YES
<input type="radio"/>	IWAYLOGS_INDX	14.850,0	14.138,0	95,2	YES
<input type="radio"/>	IWAYTEMP	100,0	1,4	1,4	YES
<input type="radio"/>	IWAYTEMP_INDX	50,0	1,0	2,0	YES
<input type="radio"/>	PRUEBA	10.240,0	64,0	0,6	NO
<input type="radio"/>	SYSAUX	2.030,0	1.929,2	95,0	YES
<input type="radio"/>	SYSTEM	790,0	785,1	99,4	YES
<input type="radio"/>	TEMP	6.884,0	29,0	0,4	YES
<input type="radio"/>	UNDOTBS1	1.240,0	25,9	2,1	YES
<input type="radio"/>	UNDOTBS2	125,0	20,5	16,4	YES
<input type="radio"/>	UNDOTBS3	125,0	10,9	8,8	YES
<input type="radio"/>	UNDOTBS4	13.575,0	75,8	0,6	YES
<input type="radio"/>	USERS	20,0	2,4	11,9	YES

Fig. 4.4 Tablespace prueba creado para la pruebas de performance.

4.3 Creación del procedimiento de prueba.

La configuración de la tabla 4.2 permite el ingreso de 50 millones de registros usando la cuenta "a" y en la tabla "b".

TABLA N° 4.2 Procedimiento PL/SQL para insertar 50 millones de registros

```
SQL> begin
2 for i in 1..50000000
3 loop
4 insert into b values (i);
5 end loop;
6 end;
7 /
```

En la figura 4.5 se visualiza el contenido del tablespace prueba.dbf, los cuadros en color amarillo representan los registros ingresados.

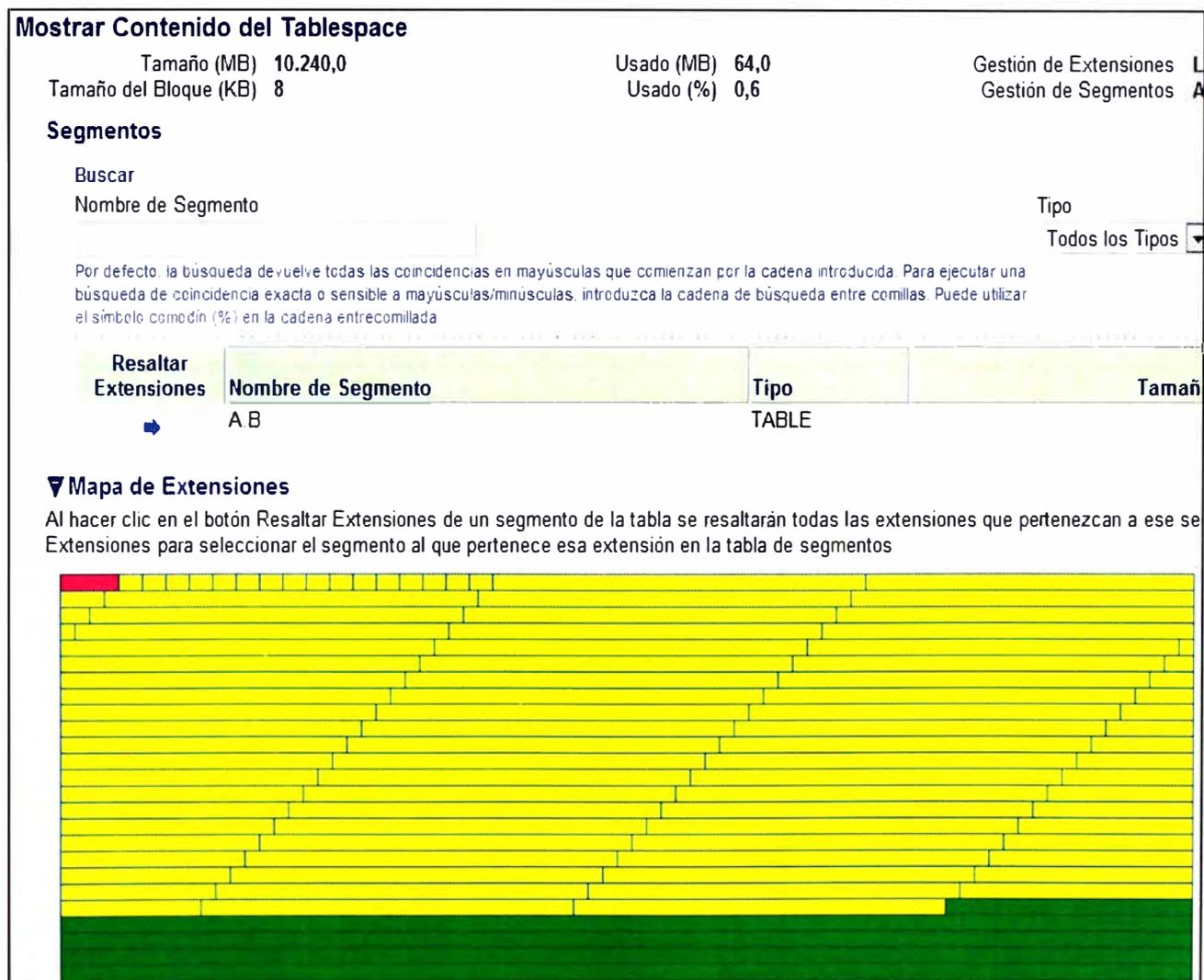


Fig. 4.5 Contenido de tablespace prueba

4.4 Impacto y rendimiento en el RAC Oracle.

Se aprecia un cierto incremento en el consumo de recursos de los nodos. En la figura 4.6 se observa el impacto en todos los niveles tanto en ancho de banda y CPU.

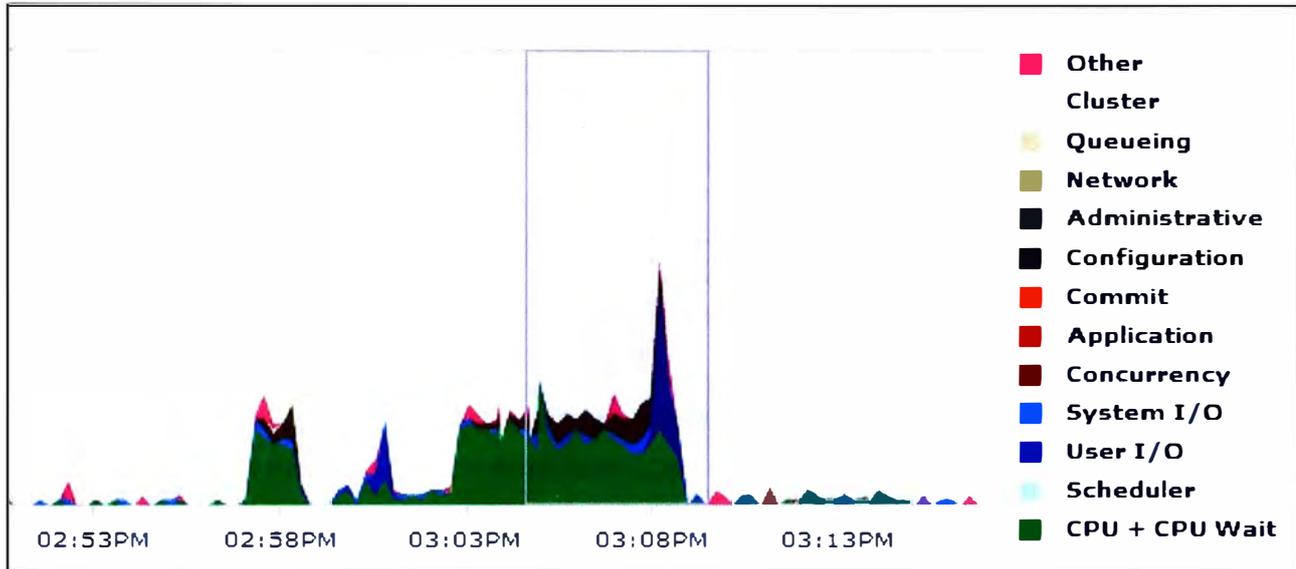


Fig. 4.6 Gráfica de impacto en la base de datos Oracle.

En la figura 4.7 se observa que la sentencia SQL "INSERT" tiene más impacto en la BD.

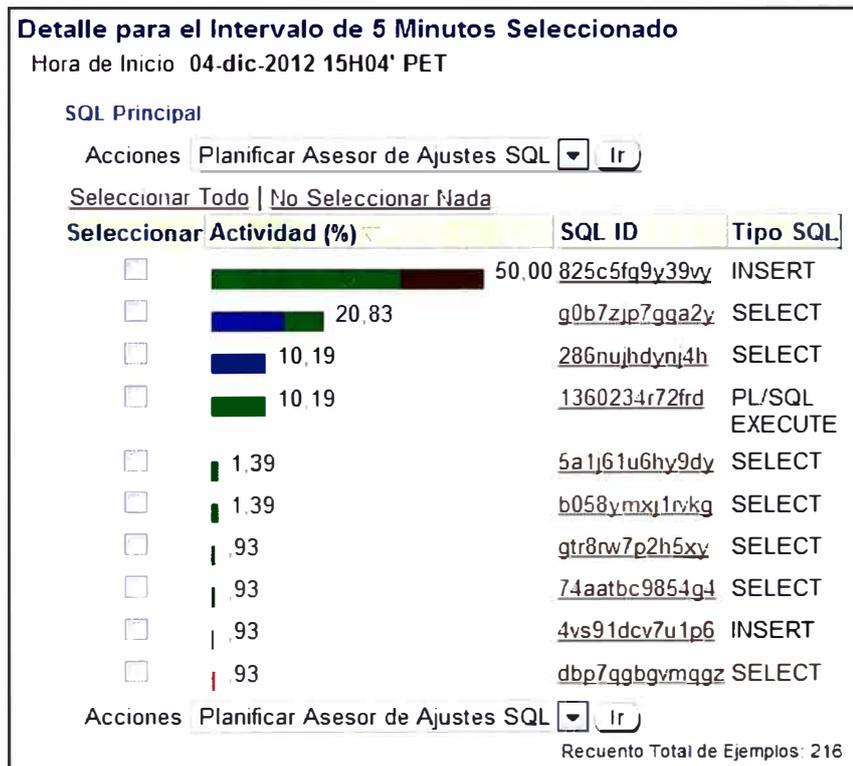


Fig. 4.7 Sentencia SQL con mayor impacto a la base de datos.

En el Enterprise Manager se puede apreciar en detalle la sentencia SQL ejecutada por el procedimiento PL/SQL, figura 4.8.

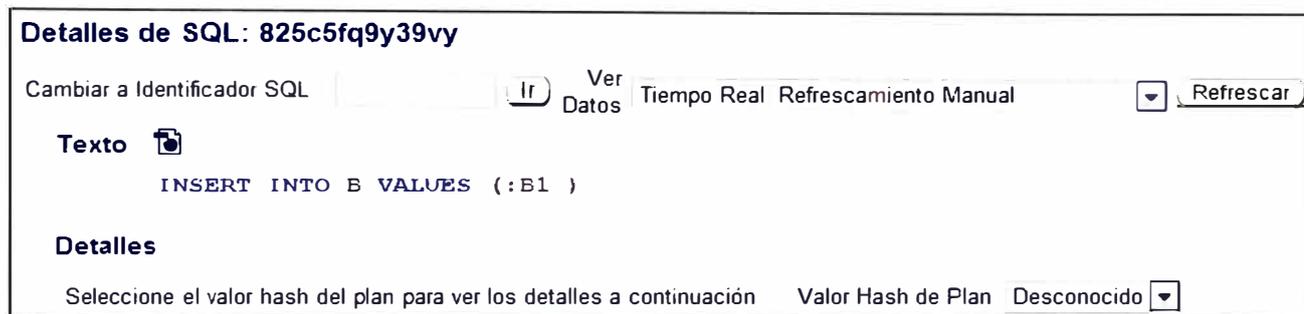


Fig. 4.8 Sentencia SQL ingresada en el procedimiento PL/SQL.

En la figura 4.9 se aprecia el consumo de ancho de banda, donde los picos llegan hasta los 260 Mbps por segundo. Recordar que la base de datos dispone de dos switches de la marca Cisco para la conexión con la red pública y que la arquitectura soporta hasta 4Gbps de transmisión y recepción máxima.

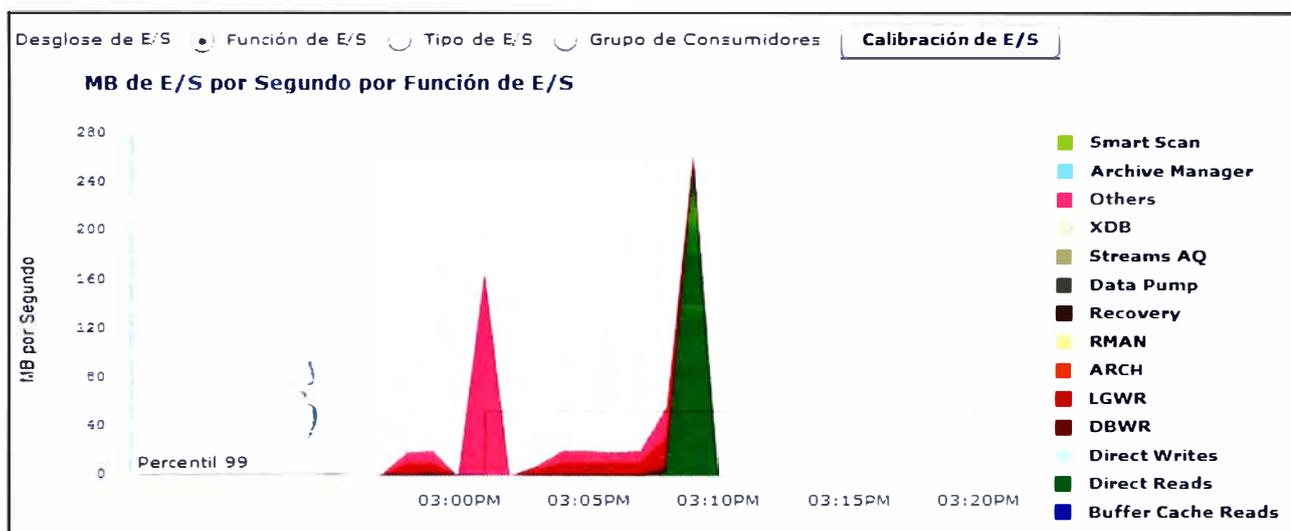


Fig. 4.9 Impacto en el ancho de banda de la base de datos.

4.5 Evaluación económica.

En esta sección se muestra un resumen de los equipos utilizados, únicamente en la infraestructura de red responsable de soportar a la base de datos, no se incluye equipos de la red de transporte responsables de interconectar las distintas sedes de la empresa, así como tampoco se incluye el precio del software de base de datos Oracle.

Se muestra además un valor referencial monetario en dólares según los precios de lista ofrecidos por los fabricantes, según indica la tabla 4.3.

TABLA N° 4.3 Valor referencial unitario según precio de lista de los fabricantes.

Equipo	Descripción	Precio Unitario	Cantidad	Precio Total
Enclosure M1000E	PowerEdge M1000e Blade Enclosure	\$105.000.00	1	\$105.000.00
WS-CBS3032-DEL-F	Cisco Catalyst Blade Switches 3130 for Dell	\$3850.00	4	\$15.400.00
M5424 8Gb SAN I/O	Brocade M5424 Fiber Channel Blade Switch	\$9520.00	2	\$19.040.00
Cisco MDS 9509	Cisco MDS 9509 Multilayer Director	\$95.000.00	1	\$95.000.00
Hitachi AMS2500	Hitachi Adaptable Modular Storage 2000	\$64.900.00	1	\$64.900.00
PowerEdge R610	DELL PowerEdge R610 rackmount server	\$3.500.00	1	\$3.500.00
WS-C2960G-24TC-L	Cisco Catalyst 2960 Series Switches	\$1.800.00	1	\$1.800.00
				\$304.640.00

El valor total en la infraestructura de red es \$304.640.00 como se indica en la tabla 4.3, dentro de este presupuesto no está incluido la mano de obra utilizado para la implementación, ya que es realizado por el personal de la empresa, pero se estima un periodo de 3 meses la duración del proyecto hasta la entrega final al área de sistemas que lleva acabo la instalación de la base de datos y migración desde la versión anterior.

CONCLUSIONES Y RECOMENDACIONES

1. Oracle real application clusters (Oracle RAC), con Oracle Database 11g Enterprise Edition, permiten ejecutar una sola base de datos en un grupo de servidores y proporciona una tolerancia a fallos, un rendimiento y una capacidad de ampliación inigualables, sin necesidad de cambios de aplicaciones.
2. Oracle RAC es un componente que permite agregar alta disponibilidad a nuestro diseño al ofrecer múltiples instancias de una base de datos con una sola unidad de almacenamiento, con la ayuda de un componente de balanceo, alta escalabilidad y disponibilidad de los servicios ofrecidos por la base de datos.
3. Los analistas valoran la creciente importancia de Oracle RAC entre el gran número de clientes de todos los sectores que consolidan sus aplicaciones de proceso de transacciones y almacenamiento de datos.
4. Usando esta arquitectura como referencia se puede diseñar una infraestructura que cumple las necesidades más exigentes en la mayoría de empresas.
5. Como las organizaciones están en proceso de migración a la nube este modelo se adapta completamente a los nuevos requerimientos de virtualización. Se reduce el riesgo introduciendo nuevas tecnologías innovadoras.
6. La solución de alta disponibilidad de base de datos por hardware llamada RAID local, proporciona disponibilidad a través de distintos niveles en los arreglos de discos, con lo cual se evita la pérdida de datos ante la falla de cualquier disco, el AMS modular Hitachi ofrece altos niveles de performance y disponibilidad en conjunto con el Cisco MDS 9509 se obtiene una red SAN robusta diseñada para soportar futuras ampliaciones. Se recomienda su uso en la medida de las necesidades y requerimientos de la empresa.
7. De la arquitectura propuesta, consideramos los siguientes escenarios de indisponibilidad: Falla de switches (10 segundos), falla de nodos RAC (1 minuto), falla de

disco (5 segundos), falla de red (5 minutos), en total obtendremos 6 minutos + 15 segundos de indisponibilidad por año = 0.10416 horas por año. Se considera además que las horas totales por año son 24horas x 365días = 8760 horas, entonces la disponibilidad = $(8760 - 0.10416) / 8760 \times 100 \% = 99.999\%$ (nivel de calidad bueno), esta es la disponibilidad esperada según la tabla 2.1 y la requerida por las aplicaciones más exigentes.

8. Antes de implementar una solución se necesita determinar las necesidades de la empresa que solicita el servicio, tales como disponibilidad, costo, eficiencia, rendimiento, etc.

9. Se recomienda revisar y estudiar todas las opciones disponibles en el mercado, los equipos usados en este informe son productos top en equipos de comunicación y almacenamiento, por ello siempre es deseable encontrar soluciones equivalentes y económicas.

10. Es importante resaltar que la redundancia es un punto clave en la implementación de alta disponibilidad para bases de datos, el software de Oracle en si no proporciona la disponibilidad requerida porque depende de diferentes componentes que interactúan entre sí, por eso es necesario prestar atención al diseño de red en general.

BIBLIOGRAFÍA

- [1] Oracle RAC One Node 11gR2
Ing. Francisco Riccio. Es un IT Specialist en IBM Perú
www.oracle.com/technetwork/es/articles/servers-storage-admin/implementar-rac-one-node-11gr2-1841983-esa.html
- [2] Using VSANs and Zoning in the Cisco MDS 9000 Family
Dan Hersey and Tom Nosella
www.cisco.com/warp/public/cc/pd/ps4159/ps4358/prodlit/zonev_wp.pdf
- [3] Cisco MDS 9000 Family Quick Configuration Guide
Copyright © 2005 Cisco Systems, Inc. All rights reserved
www.cisco.com/en/US/docs/storage/san_switches/mds9000/sw/san-os/quick/guide/qcg.pdf
- [4] Dell PowerEdge M1000e Modular Enclosure Architecture
John Loffink, Engineer/Strategist, January 2008
www.dell.com/downloads/global/products/pedge/en/pedge_m1000e_white_paper.pdf
- [5] Deploying Oracle 11gR2 Enterprise Edition Real Application Cluster
Anantha Adiga September 2011
www.hds.com/assets/pdf/deploying-oracle-real-application-cluster-with-enterprise-linux5.4-on-hitachi-ams.pdf
- [6] Cisco MDS 9000 Family Acceleration Services Enhance Hitachi.
Sarah Hamilton and Vinay Gaonkar.
www.cisco.com/web/partners/downloads/765/other/hitachi_truecopy_synchronous_replication_performance.pdf
- [7] Cisco MDS 9000 Family Cookbook for Cisco MDS SAN-OS Release 3.1
Venkat Kirishnamurthy, October 2007
www.cisco.com/en/US/docs/storage/san_switches/mds9000/sw/rel_3_x/cookbook/MDScookbook31.pdf
- [8] Symmetric Active-Active Controller
Mark Adams and Patrick Allaire, December 2010
www.hds.com/assets/pdf/hitachi-white-paper-dynamic-virtual-controller-technology.pdf
- [9] Storage Reference Architecture from Cisco, Hitachi and Vmware Reference

Architecture

Sean McKeown, Hitachi Data Systems June 7, 2011

www.cisco.com/web/partners/downloads/765/other/storage_reference_architecture.pdf

[10] Access Gateway Administrator's Guide Supporting Fabric OS v6.4.0

Copyright © 2007-2011 Brocade Communications Systems,

www.brocade.com/downloads/documents/product_manuals/B_SAN/AccessGateway_AdminGd_v700.pdf

[11] Technical Brief: How to Configure NPIV on VMware ESX Server 3.5

Brocade Communications Systems, Inc. All Rights Reserved 02/2009

www.brocade.com/downloads/documents/technical_briefs/Brocade_NPIV_ESX3.5_WP.pdf

[12] Cisco Catalyst Blade Switch 3030 for Dell Command Reference

Cisco IOS Release 12.2(55)SE August 2010

www.cisco.com/en/US/docs/switches/blades/3030/software/release/12.2_58_se/command/reference/bootldr.html

[13] Information Storage and Management

G. Somasundaram, 2009

EMC Education Services

[14] Connecting the Dell PowerEdge M1000e to a Cisco SAN Fabric

Brocade Communications Systems, Inc. All Rights Reserved 07/2009

www.dell.com/downloads/global/products/pedge/en/connecting_dell_pedge_m1000e_to_cisco_san_fabric_en.pdf

LISTA DE ACRONIMOS

ACL	Access Control List
ARCH	Archiver
ASM	Automatic Storage Management
BUS	Binary Unit System
CA	Corriente Alterna
CIFS	Common Internet File Services
CKPT	Check Point
CLI	Command Language Interpreter
CPU	Computer Processing Unit
DAS	Direct Attached Storage
DBA	Data Base Administrator
DBWR	Data Base Writer
DNS	Domain Name Service
DMZ	Demilitarized Zone
FC	Fiber Channel
FCP	Fiber Channel Protocol
GNU	GNU's Not Unix
GPL	General Public License
HBA	Host Based Adapter
IBM	International Business Machines
IDE	Identifier Extension
IP	Internet Protocol
IPV4	Internet Protocol Version 4
ISCSI	Internet Small Computer Serial Interface
LAN	Local Area Network
LCD	Liquid Crystal Diode
LGWR	Log Writer
LUN	Logical Unit Number
MAC	Machine Address Code
MDS	Multilayer Director Switch

NAS	Network Accessible Storage
NFS	Network Files Services
NPV	N Port Virtualization
NPIV	N Port ID Virtualization
OCR	Oracle Cluster Registry
ORDBMS	Object Relational Data Base Management System
PGA	Program Global Area
PMON	Process Monitor
RAC	Real Application Clusters
RAID	Redundant Array of Inexpensive Disk
RECO	Recoverer
RPM	Red Hat Package Manager
RSA	Rivest Shamir Adleman
SAN	Storage Area Network
SATA	Serial Advanced Technology Attachment
SCSI	Small Computer Serial Interface
SGA	System Global Area
SGBD	Sistemas Gestores de Base de Datos
SMON	System Monitor
SQL	Search and Query Language
TCP/IP	Transmisión Control Protocol Internet Protocol
TI	Tecnologías de la Información
UNIX	Uniplexed Information and Computing System
USB	Universal Serial Bus
UTM	Unified Threat Management
VLAN	Virtual Local Area Network
VSAN	Virtual Storage Area Network
VSL	Virtual Switch Link
VSS	Virtual System Switching
WAN	Wide Area Network
WLAN	Wireless Local Area Network
WWN	World Wide Name
WWNN	World Wide Node Name
WWPN	World Wide Port Name
WWW	World Wide Web

LISTA DE SIGNIFICADO DE PALABRAS

Array, término usado para guardar o agrupar un conjunto de objetos de la misma clase.

Array de Discos, agrupación de discos para formar uno de mayor capacidad.

Automatic storage management, funcionalidad proporcionado por Oracle para la administración de grupos de discos lógicos.

Background Oracle, procesos de servidor que se ejecutan a nivel del sistema operativo.

Bonding, término usado en Linux para representar el agrupamiento lógico de puertos de red en uno virtual.

Boom, novedad, modelo revolucionario.

Chassis, equipo con múltiples ranuras para soportar diferentes módulos.

Cluster, conglomerado de equipos con las mismas características y que desempeñan una tarea común.

Disk Groups, agrupamiento lógico de discos para formar un volumen más amplio.

Fiber Channel, tecnología de red utilizada en redes de almacenamiento.

Fiber channel protocol, protocolo de transporte para la transmisión de comandos SCSI sobre redes de canal de fibra.

Firewall, equipo de seguridad informática, que permite el bloque de servicios y segmentos de red protegidos.

Infraestructura Grid, arquitectura general para una base de datos Oracle redundante.

Linux Channel Bonding, capacidad de agrupar dos o más puertos físicos en un Puerto virtual de mayor capacidad.

Masking, capacidad de permitir o bloquear el acceso a una LUN desde un servidor.

Mirroring, capacidad de crear redundancia en el arreglo de discos administrador por ASM.

MNLA, nombre usado para la base Oracle descrita en el informe.

Model OSI, es un marco de referencia para la definición de arquitecturas en la interconexión de los sistemas de comunicaciones.

Oracle database server, software desarrollado por la compañía Oracle para la implementación de bases de datos relacionales.

Oracle RAC, versión en alta disponibilidad de la base de datos Oracle.

Oracle real application clusters, software de oracle que se instala en múltiples servidores y que permiten acceder a la misma base de datos.

Patch Cord, cable UTP usado para conectar un dispositivo electrónico a otro.

PL/SQL, lenguaje de programación usado para realizar consultas a una base de datos.

Process background Oracle, servicios o programas que se ejecutan en una implementación de base de datos para monitorear o controlar el funcionamiento del sistema.

Productos Top, equipos informáticos con un alto rendimiento y costo.

Raw Devices, es un método usado para acceder de manera directa a la gestión de discos sin pasar por el sistema operativo instalado.

Red Token Ring, arquitectura de red desarrollada por IBM, con una topología física en forma de anillo.

Router, equipo de comunicación responsable de dirigir los flujos de información de los protocolos de internet.

SCAN Listener, tecnología usada por Oracle para realizar el balanceo de transacciones entre los nodos de un cluster.

Servidor Blade, servidor en forma de chasis con múltiples ranuras de expansión, en la mayoría de modelos soporta hasta 16 servidores por chasis.

Stack wise, tecnología usada por Cisco para unir virtualmente dos o más switches.

Stand Alone, equipo para ser usado de manera independiente.

Striping, capacidad de redistribuir el contenido de uno disco mientras la base de datos esta en producción.

Switch, equipo de comunicación responsable de retransmitir los flujos de información basados en las direcciones MAC.

Tablespace, archivo de sistema usado para almacenar los datos de una tabla lógica en una base de datos relacional.

Throughput, volumen de trabajo o de información que fluye a través de un sistema.

Transparent application failover, permite a la aplicación reconectarse automáticamente a una base de datos, si falla la conexión a la instancia de la base de datos.

Virtual, sistema existente solo en ambiente computacional.

Virtual switching system, tecnología propietaria de Cisco desarrollada para asociar virtualmente dos equipos de comunicación de la serie Catalyst 6500.

Voting File, archivo de sistema usado por Oracle para el funcionamiento del agrupamiento de discos usando ASM.

Worldwide name, hacen referencia a un identificador único en una red de almacenamiento.

Zoning, capacidad de segmentar dos o más puertos en un fabric switch para crear una zona segura de comunicación.