

Universidad de las Ciencias Informáticas

Facultad 3



**Administración, configuración y optimización de un Sistema
de Bases de Datos Descentralizado en Oracle Database 10g
release 2**

Trabajo de Diploma para optar por el título de
Ingeniero en Ciencias Informáticas

Autor(es): Daniel Fernández Arencibia

Yunieski Fábregas Santos

Tutor: Ing. Rudel Cárdenas Díaz

Asesor: Ing. Alexei Zubizarreta Pérez

Junio - 2007

DECLARACIÓN DE AUTORÍA

Declaramos todos los derechos de autoría sobre el presente trabajo y como tal se autoriza a la Universidad de las Ciencias Informáticas al uso público del mismo.

Para que así conste firmamos la presente a los _____ Días Del Mes de _____ del Año

Daniel Fernández Arencibia

Autor.

Yunieski Fábregas Santos

Autor.

Ing. Rudel Cárdenas Díaz

Tutor.

AGRADECIMIENTOS

Este trabajo pudo ser escrito gracias al apoyo, colaboración y paciencia de muchas personas. Por lo mismo es probable que olvidemos mencionar a algunos de ellos en estas líneas; a ellos les pedimos disculpas.

Agradecemos a la Universidad de las Ciencias Informáticas, por darnos la oportunidad de formarnos en ella durante los últimos cinco años y hacer de nuestros sueños una realidad.

A nuestros padres, amantes de sus hijos, y con esperanza desmedida de la utilidad y significado de la formación académica, nos apoyaron moral y espiritual en todo momento.

A nuestro tutor y asesor Ing. Rudel Cárdenas Díaz e Ing. Alexei Zubizarreta Pérez, que con su seriedad científica debidamente matizada de criticidad y humor, hicieron de esta experiencia académica algo muy satisfactorio.

A los profesores que han compartido sus experiencias durante estos años. Siempre les estaremos agradecidos.

A todos los que de una forma u otra, nos han ayudado durante la realización del trabajo: amigos, compañeros, en fin gracias a todos.

Con cariño y admiración gracias a todos.

DEDICATORIA

A mis padres por hacer de mi lo que soy. No hay una sola línea de este trabajo donde no estén presentes. Nunca me alcanzará el tiempo para hacer por ustedes lo que han hecho por mí. Siempre han confiado ciegamente en todo lo que he hecho y hacer que se sientan orgullosos ha sido el camino que he tomado desde que tuve conciencia de ello. Son los mejores padres del mundo y saben que nunca les fallaré. Vivan muchos años más que los necesito aún.

A Beatriz mi novia quien ha estado conmigo en todo momento, impulsándome a luchar, a no renunciar a mis sueños, ayudándome a confiar y creer en mí, como siempre ella lo ha hecho y enseñándome el verdadero sentido de lo que es amar.

A mis seres queridos, a todos que con su vida y ejemplo forjaron el amor y los valores que ahora forman parte de mí y que pese a no estar presentes nunca más a mi lado físicamente, me continúan guiando día a día y viven y se enorgullecen de mis triunfos y a ellos les dedico este trabajo.

Daniel Fernández Arencibia

DEDICATORIA

A mis padres, gracias por su amor y por ser la luz que me guió durante todos estos años. Espero que se sientan orgullosos de mi como me he sentido yo de ustedes todos estos años por haberme educado con esos principios sin los cuales hoy no podría ser lo que soy.

A mi querida novia Yanisleydi, gracias por soportarme y apoyarme durante estos dos años que han sido maravillosos y por enseñarme lo que es una relación seria basada en el amor.

A mis amigos, en especial mis compañeros de la Vocacional, los mejores amigos que he tenido en mi vida. No importa lo lejos que estemos, siempre seremos el 203.

A mis familiares y a todos los que de una forma u otra han estado presente durante esta larga carrera que es la vida.

Yunieski Fábregas Santos

PROPÓSITO

El documento forma parte de la Solución Informática presentada al Gobierno de la República de Venezuela para la automatización de los Registros y Notarías de Venezuela.

Es una referencia al Diseño e Implementación de una Arquitectura de Bases de Datos Descentralizada utilizando Oracle Database 10g release 2 como software base.

ALCANCE

Está encaminado como futura guía de referencia para Sistemas Informáticos que implementen un entorno Descentralizado de Bases de Datos en Oracle Database 10g release 2.

RESUMEN

El presente documento es un reflejo detallado de la implementación de una Arquitectura de Bases de Datos Descentralizada utilizando Oracle Database 10g release 2, encaminada a automatizar los servicios de los Registros y Notarías de la República Bolivariana de Venezuela. La arquitectura propuesta está sustentada según lo planteado por la “Ley del Registro Público y del Notariado”.

Se hace énfasis a los mecanismos empleados para el diseño, implementación, y posterior integración de cada subsistema a un ambiente descentralizado de Bases de Datos, en el que el intercambio de información es bidireccional.

Se hace alusión a las políticas de seguridad y recuperación a tener en cuenta en caso de producirse desastres que atenten contra la integridad y preservación de la información que se maneja en cada instante de tiempo.

ÍNDICE

INTRODUCCIÓN.....	1
CAPÍTULO I: FUNDAMENTACIÓN TEÓRICA.....	5
Introducción.....	5
1.1 DEFINICIÓN DE BASE DE DATOS.....	5
1.2 MODELOS DE DATOS.....	6
1.2.1 Modelo de Datos Jerárquico.....	7
1.2.2 Modelo de Datos de Red.....	8
1.2.3 Modelo de Datos Relacional.....	9
1.2.4 Modelos de Datos Orientados a Objetos.....	10
1.2.5 Modelo de Bases de Datos Distribuidas.....	11
1.3 SISTEMAS MANEJADORES DE BASES DE DATOS.....	13
1.4 METODOLOGÍA DE DISEÑO DE BASES DE DATOS.....	14
1.4.1 Diseño Conceptual.....	14
1.4.2 Diseño Lógico.....	15
1.4.3 Diseño Físico.....	16
1.5 DISEÑO DE UN CENTRO DE DATOS.....	17
1.5.1 Tendencias de los Centros de Datos.....	17
1.5.2 Planificación y Diseño de un Centro de Datos.....	18
1.5.3 Componentes de un Centro de Datos.....	18
1.6 CARACTERIZACIÓN DE LA SUITE DE PRODUCTOS DE ORACLE.....	23
Conclusiones del Capítulo.....	27
CAPÍTULO II: DISEÑO E IMPLEMENTACIÓN DE UNA ARQUITECTURA DE BASES DE DATOS DESCENTRALIZADA EN ORACLE DATABASE 10G RELEASE 2.....	28
Introducción.....	28
2.1 ARQUITECTURA DE UNA BASE DE DATOS ORACLE.....	28
2.1.1 Estructuras Lógicas y Físicas de una Base de Datos Oracle.....	29
2.2 DISEÑO GENERAL DE LA ARQUITECTURA DE LA BASE DE DATOS.....	30

2.2.1	Diseño de Servidores Locales.	32
2.2.2	Diseño de Bases de Datos Locales para los Registros Públicos y Mercantiles.	36
2.2.3	Diseño Físico de las Bases de Datos Locales.	36
2.3	MECANISMO DE SINCRONIZACIÓN ENTRE BASES DE DATOS LOCALES Y CENTRO DE DATOS.	38
2.3.1	Oracle LogMiner.	39
2.4	SEGURIDAD EN LAS BASES DE DATOS LOCALES.	42
2.4.1	Estándares de Seguridad.	43
2.4.2	Principios de Seguridad Aplicados en las Bases de Datos Locales.	44
2.5	RESPALDO Y RECUPERACIÓN EN LAS BASES DE DATOS LOCALES.	51
2.5.1	Tipos de Backups.	54
2.5.2	Estrategias de Backups. Reglas Básicas.	55
2.5.3	Estrategias de Recuperación. Recovery Manager Oracle.	58
2.6	ARQUITECTURA DE LA BASE DE DATOS CENTRAL.	63
2.6.1	Diseño del Centro de Datos.	63
2.6.2	Selección del Hardware.	66
2.6.3	Software Base Instalado.	69
2.6.4	Diseño e Implementación de un Real Application Cluster (RAC).	73
2.6.5	Qué es Oracle Database 10g Real Application Clusters?	73
2.6.6	Diseño de la Base de Datos Central.	76
2.7	REPLICACIÓN DE DATOS UTILIZANDO ORACLE STREAMS.	79
2.7.1	Arquitectura de Oracle Streams.	80
2.7.2	Mejores Prácticas para Oracle Streams.	85
2.7.3	Estrategias de Respaldo y Recuperación en el Centro de Datos de Registros y Notarias.	88
2.7.4	Modulo de Administración y Control de Backups.	89
2.7.5	Entorno de Integración de HP OpenView Storage Data Protector 6.0.	90
2.7.6	Estrategias de Backups.	91
	Conclusiones del Capítulo.	95
CAPÍTULO III: ESTRATEGIAS DE OPTIMIZACIÓN Y ADMINISTRACIÓN DE BASES DE DATOS EN ORACLE DATABASE 10G RELEASE 2		96
	Introducción.	96

3.1	OPTIMIZACIÓN DE BASES DE DATOS ORACLE DATABASE 10G RELEASE 2.	96
3.1.1	Parámetros de Inicialización de la Base de Datos.	96
3.1.2	Configuración y Diseño del Subsistema de I/O (Entrada/Salida).....	97
3.1.3	Desarrollar Secuencias SQL Eficientes.	98
3.1.4	Uso de Índices.....	100
3.2	PROPUESTA DE HERRAMIENTA PARA LA ADMINISTRACIÓN DE BASES DE DATOS ORACLE 10G.	105
3.2.1	Introducción a la Administración Automática. Oracle Enterprise Manager 10g Grid Control.	107
3.2.2	Características de Oracle Enterprise Manager 10g Grid Control.....	107
3.2.3	Requerimientos para el Despliegue de Oracle Enterprise Manager 10g Grid Control.	108
3.2.4	Oracle Enterprise Manager 10g Grid Control. Arquitectura y Componentes.	109
3.2.5	Opciones de la Consola Oracle Enterprise Manager 10g Grid Control.....	110
3.2.6	Seguridad, Gestión y Administración de Oracle Enterprise 10g Grid Control.....	111
3.2.7	Administración y Monitorización de Sistemas con Oracle Enterprise 10g Grid Control.	112
3.2.8	Monitorización de Recursos.....	112
3.2.9	Gestión de los Métodos de Notificación.....	112
3.2.10	Configuración de Parches.....	112
3.2.11	Gestión de Interrupciones.....	113
3.2.12	Control y Gestión de Cambios.	113
3.2.13	Gestión de Trabajos (Job Schedule).....	113
3.2.14	Baselines y Administración Basada en Políticas.	113
3.2.15	Administración de la Seguridad.....	114
	Conclusiones del Capítulo.....	115
	CONCLUSIONES GENERALES.....	116
	RECOMENDACIONES.....	118
	BIBLIOGRAFÍA.....	119

INTRODUCCIÓN

Cada Estado tiene la responsabilidad de contribuir con el disfrute de los derechos sociales de forma universal y equitativa, atendiendo a la Ley del Registro Público que se promulgó en el año 1993, se autorizó la digitalización y almacenamiento documental de la información contenida en las oficinas de los Registros y Notarías del país. Dada la situación que el Ministerio del Poder Popular para la Relaciones Interiores y Justicia no dispone de una herramienta tecnológica ni procedimental que garanticen efectivamente la ejecución de funciones en los Registros y Notarías, se decidió por parte del organismo constitucional profundizar en un conjunto de acciones encaminadas a garantizar un nivel permisible de la seguridad jurídica de los servicios registrales y notariales de la República.

El trabajo es parte del proceso de Automatización de los Registros y Notarías de la República de Venezuela, el cual tiene como objetivo automatizar todas las operaciones que actualmente se realizan en los Registros.

Según lo establece el “Artículo #28 de la Ley del Registro Público y del Notariado”, se ha diseñado un Centro de Datos que contempla la creación de una Base de Datos en el Distrito Metropolitano de Caracas que consolidará y respaldará la información de todas las materias registrales correspondientes a los Registros Mercantiles y Públicos del País.

A partir del uso de técnicas de diseño de Bases de Datos Relacionales se ha modelado cada sistema como un subconjunto global, en el cual la información trasciende en ambos sentidos utilizando técnicas de replicación soportada por Oracle Database 10g release 2.

La Arquitectura Descentralizada de Bases de Datos implementada es clave en el intercambio de información entre todos los subsistemas que forman parte de la solución informática propuesta.

El Proyecto Registro y Notarías de la República de Venezuela está encaminado a solucionar los problemas que afectan a los Registros Mercantiles y Públicos. Y estos son:

- La automatización de los procesos que actualmente se realizan en cada uno de los registros del país.
- Contribuir con el disfrute de los derechos sociales de forma universal y equitativa.
- Disminuir los tiempos de respuesta al cliente, enfocándonos en ofrecer una atención especializada en cada caso.
- Eliminar vestigios de malversación y corrupción.
- Centralización de la información en un Centro de Datos con el objetivo de establecer controles de auditoría que puedan servir de base para futuras tomas de decisiones.
- Implementación de una solución administrativa centralizada, que permita mantener un monitoreo global de cada subsistema.
- Estandarizar los procesos de gestión en los Registros y Notarías.
- Mejorar los Sistemas de Información y Comunicación.
- Garantizar un sistema altamente disponible, ofreciendo soporte y monitoreo los 7 días de la semana durante las 24 horas (24x7).
- Información certera sobre el tracto sucesivo de los inmuebles y compañías.

Para la implementación de un sistema que responda a las especificaciones descritas anteriormente, este se fragmentó en varios módulos encargados de la modelación de cada situación del negocio en específico y para posteriormente integrar cada subsistema en un solo Sistema Global.

Los módulos que se implementarán son los siguientes:

- Módulo de Público
- Módulo de Mercantil
- Módulo de Servicio Autónomo
- Módulo de Administración Financiera

Para la implementación y operabilidad de cada uno de los módulos mencionados recientemente, se necesita de la modelación del negocio a nivel de Objeto Relacional. La Base de Datos es vital en el correcto funcionamiento de cada módulo. Sus principales funciones son:

- Gestión de la información.
- Fuente de intercambio de información entre las aplicaciones utilizando lenguaje SQL.
- A través de la Arquitectura de Bases de Datos Descentralizada se garantiza el flujo de información entre el Centro de Datos y las Bases de Datos Locales de cada Registro.
- Garantiza disponibilidad de la información los 7 días de la semana durante las 24 horas (24x7).
- Garantiza un total respaldo, integridad y preservación de la información.

Como resultado de lo planteado anteriormente podemos concluir que nuestro Problema Científico está encaminado al diseño, implementación y administración de un Sistema de Bases de Datos Descentralizado sobre Oracle Database 10g release 2.

Nuestro Objeto de estudio está dirigido al Diseño de una Arquitectura de Bases de Datos Descentralizada, definiendo nuestro campo de acción en el diseño, implementación y administración de una Arquitectura de Bases de Datos Descentralizada sobre Oracle Database 10g release 2.

Las interrogantes de investigación que se plantean son las siguientes:

¿Cómo diseñar una Arquitectura de Bases de Datos Descentralizada en Oracle Database 10g release 2 que cumpla con lo establecido en la “Ley del Registro Público y del Notariado”?

¿Cómo definir políticas de respaldo y recuperación que garanticen la seguridad jurídica de los documentos gestionados en cada oficina?

¿Cómo configurar herramientas administrativas para la administración de un Grid de Bases de Datos en Oracle Database 10g release 2?

Dando respuesta a las interrogantes planteadas anteriormente estamos haciendo cumplir los Artículos 27, 28, 29, 30 y 31 de la Ley de Registro Público y del Notariado, además de responder al **Objetivo General** de Diseñar una Arquitectura de Bases de Datos Descentralizada para los Registros y Notarías de la República Bolivariana de Venezuela utilizando como gestor Oracle Database 10g release 2, y a los **Objetivos Específicos** siguientes:

- Diseñar una Arquitectura de Bases de Datos Descentralizada.
- Definir políticas de respaldo y recuperación contra desastres.
- Configurar herramientas administrativas para el monitoreo de un Grid de Bases de Datos en Oracle 10g release 2.

A partir de los elementos expuestos anteriormente, hemos creado un grupo de tareas investigativas encaminadas a dar solución a las interrogantes antes mencionadas. Estas tareas consisten en:

- Investigar sobre métodos de Diseño de Bases de Datos.
- Investigar sobre las herramientas utilizadas para el Diseño de Bases de Datos Relacionales.
- Investigar sobre el Diseño e Implementación de Bases de Datos Descentralizadas.
- Implementar mecanismos de replicación soportados por Oracle Database 10g release 2.
- Optimizar Bases de Datos en Oracle Database 10g release 2 en Red Hat Linux Advanced Server 4 y Microsoft Windows Small Business Server 2003 Standard Edition.
- Administración y Despliegue de un Real Application Cluster (RAC) en conjunto con un Grid de Bases de Datos Locales.
- Configurar herramientas administrativas para la administración de un Grid de Bases de Datos en Oracle Database 10g release 2.

CAPÍTULO I: FUNDAMENTACIÓN TEÓRICA.

Introducción.

En el capítulo se hace mención a un grupo de conceptos que serán referenciados en el resto del documento. Se abordan aspectos relacionados con el uso de las nuevas tecnologías de la informática y las telecomunicaciones. Se reseñan las metodologías empleadas para el análisis, diseño e implementación de bases de datos. Se sintetizan aspectos relacionados con las tendencias, diseño y componentes que forman parte de un Centro de Datos. Se hace un breve enfoque a la suite de productos de Oracle, dirigiendo nuestra atención a Oracle Database 10g como la primera base de datos diseñada para grid computing.

1.1 Definición de Base de Datos.

El término de Base de Datos fue escuchado por primera vez en 1963, en un simposio celebrado en California EUA; con los progresos en la adquisición y almacenamiento de datos en diferentes ambientes como gobierno, industria y aplicaciones científicas, arrojan como resultado el surgimiento de enormes bases de datos, cuyo tamaño se incrementa rápidamente, tanto en número de registros como en la dimensionalidad de los mismos, llegando a alcanzar proporciones de almacenamiento del orden de los terabytes. La necesidad de explorar estas bases de datos y extraer información y conocimiento que sea de interés para los propietarios de las mismas, se ha incrementado en la actualidad.

Una Base de Datos o Banco de Datos se puede definir como un conjunto de datos que pertenecen al mismo contexto almacenados sistemáticamente para su posterior uso[1]. En este sentido, una biblioteca puede considerarse una base de datos compuesta en su mayoría por documentos y textos impresos en papel e indexados para su consulta. La utilización de bases de datos como plataforma para el desarrollo de Sistemas de Aplicación en las Organizaciones se ha incrementado notablemente en los últimos años, se debe a las ventajas que ofrece su utilización, algunas de las cuales se comentarán a continuación:

- Globalización de la información, que permite a los diferentes usuarios considerar la información como un recurso corporativo que carece de dueños específicos.
- Eliminación de información inconsistente.
- Permite mantener la integridad de la información como cualidad altamente deseable y tiene por objetivo almacenar correctamente la información.
- Independencia de datos como factor esencial en la rápida proliferación del desarrollo de Sistemas de Bases de Datos. La independencia de los datos implica un divorcio entre programas y datos.

1.2 Modelos de Datos.

Los Modelos de Datos aportan la base conceptual para diseñar aplicaciones que hacen un uso intensivo de datos, así como la base formal para las herramientas y técnicas empleadas en el desarrollo y uso de sistemas de información[2].

El modelado de datos puede ser descrito como los requerimientos de información y proceso de una aplicación de uso intensivo de datos, construir una representación de la aplicación que capture las propiedades estáticas y dinámicas requeridas para dar soporte a los procesos deseados.

Un Modelo de Datos es por tanto una colección de conceptos bien definidos matemáticamente que ayudan a expresar las propiedades estáticas y dinámicas de una aplicación con un uso de datos intensivo.

Los Modelos de Datos se distinguen en 4 generaciones:

- Modelos de Datos Primitivos (orientados al fichero).
- Modelos de Datos Clásicos.
- Modelos de Datos semánticos.
- Modelos de datos de propósito específico (orientados a la aplicación).

Los modelos de datos clásicos están formados por:

- Modelo Jerárquico.
- Modelo de Red.
- Modelo Relacional.

Los Modelos de Datos no son elementos físicos: son abstracciones que permiten la implementación de un sistema eficiente de bases de datos; por lo general se refieren a algoritmos, y conceptos matemáticos. Algunos modelos con frecuencia utilizados en las bases de datos son: Modelos de Datos Jerárquico, Modelos de Datos de Red, Modelo de Datos Relacional, Modelo de Datos Distribuido.

1.2.1 Modelo de Datos Jerárquico.

Los Modelos de Bases de Datos Jerárquicos como su nombre lo indica, almacenan su información en una estructura jerárquica. En este modelo los datos se organizan en una forma similar a un árbol, donde un nodo padre de información puede tener varios hijos[3]. El nodo que no tiene padres es llamado raíz, y a los nodos que no tienen hijos se los conoce como hojas.

Este modelo es especialmente útil en el caso de aplicaciones que manejan un gran volumen de información y datos muy compartidos, permitiendo crear estructuras estables y de gran rendimiento. Una de las principales limitaciones de este modelo es su incapacidad de representar eficientemente la redundancia de datos. A continuación se mencionan los problemas típicos del Modelo de Bases de Datos Jerárquico. Todos estos problemas se derivan del hecho de que el sistema gestor de Bases de Datos no implementa ningún control sobre los propios datos, sino que queda en manos de las aplicaciones garantizar que se cumplen las condiciones invariantes que se requieran.

Por ejemplo:

Duplicidad de Registros

No se garantiza la inexistencia de registros duplicados. Esto también es cierto para los campos clave. Es decir, no se garantiza que dos registros cualesquiera tengan diferentes valores en un subconjunto concreto de campos.

Integridad Referencial

No existe garantía de que un registro hijo esté relacionado con un registro padre válido. Por ejemplo, es posible borrar un nodo padre sin eliminar antes los nodos hijo, de manera que éstos últimos están relacionados con un registro inválido o inexistente.

Denormalización

Este no es tanto un problema del modelo jerárquico como del uso que se hace de él, sin embargo, a diferencia del modelo relacional, las bases de datos jerárquicas no tienen controles que impidan la desnormalización de una base de datos, por ejemplo, no existe el concepto de campos clave.

1.2.2 Modelo de Datos de Red.

El Modelo de Datos de Red es ligeramente distinto al modelo jerárquico; su diferencia fundamental es la modificación del concepto de nodo: se permite que un mismo nodo tenga varios padres, posibilidad no permitida en el modelo jerárquico[4].

Fue una gran mejora con respecto al modelo anterior, ya que ofrece una solución eficiente al problema de redundancia de datos; pero, aún así, la dificultad que significa administrar la información en una base de datos de red ha significado que sea un modelo utilizado en su mayoría por programadores más que por usuarios finales.

1.2.3 Modelo de Datos Relacional.

El Modelo de Datos Relacional es el modelo más utilizado en la actualidad para modelar problemas reales y administrar datos dinámicamente. Los motivos de este éxito se basan fundamentalmente en que ofrece un sistema simple y eficaz para representar y manipular los datos. Además se basa en un modelo relacional con sólidas bases teóricas. Tras ser postulados sus fundamentos en 1970 por Edgar Frank Codd, de los laboratorios IBM en San José California, no tardó en consolidarse como un nuevo paradigma en los modelos de bases de datos[5]. Su idea fundamental es el uso de relaciones. Estas relaciones podrían considerarse en forma lógica como conjuntos de datos llamados tuplas. La teoría de las bases de datos relacionales la mayoría de las veces se conceptualiza de una manera más fácil de imaginar, pues se piensa en cada relación como si fuese una tabla que está compuesta por registros, que representarían las tuplas y campos.

En este modelo, el lugar y la forma en que se almacenan los datos no tienen relevancia a diferencia de otros modelos como el jerárquico y el de red. Esto tiene la considerable ventaja de que es más fácil de entender y de utilizar para un usuario esporádico de la base de datos. La información puede ser recuperada o almacenada mediante consultas que ofrecen una amplia flexibilidad y poder para administrar la información. Algunas de las grandes ventajas del modelo relacional son:

- Define un álgebra, llamada álgebra relacional a partir de la cual se realizan todas las manipulaciones posibles sobre las relaciones, que se obtienen mediante el uso de operadores.
- Garantía de independencia de los datos.
- Conectividad garantizada con los lenguajes de programación estándar.
- Compatibilidad y estandarización.
- Favorece la normalización por ser más comprensible y aplicable.
- Garantiza la integridad referencial, así al eliminar un registro elimina todos los registros relacionados dependientes.
- Garantiza herramientas para evitar la duplicidad de registros, a través de campos claves o llaves.

Sin embargo, también hay que ser conscientes de los aspectos negativos, o más bien limitaciones, que trae consigo la adopción del modelo de datos relacional.

Existe una serie de desventajas bien conocidas de este modelo, que se ponen de manifiesto especialmente cuando lo comparamos con otros modelos más recientes como por ejemplo el modelo orientado a objetos. Las más obvias son las siguientes:

- Imposibilidad de representar conocimiento en forma de reglas.
- Inexistencia de mecanismos de herencia de propiedades.
- Incompatibilidad entre los tipos de estructuras de datos que se transfieren ó desadaptación de impedancia.
- Dificultad para gestionar datos no atómicos.

1.2.4 Modelos de Datos Orientados a Objetos.

A los modelos que describen el estado y comportamiento de los objetos se les denomina Modelos de Datos Orientados a Objetos.

Este modelo es bastante reciente, y propio de los modelos informáticos orientados a objetos. Una Base de Datos Orientada a Objetos es una base de datos que incorpora todos los conceptos importantes del paradigma de objetos[6]:

- **Encapsulación:** Propiedad que permite ocultar la información al resto de los objetos, impidiendo así accesos incorrectos o conflictos.
- **Herencia:** Propiedad a través de la cual los objetos heredan comportamiento dentro de una jerarquía de clases.
- **Polimorfismo:** Propiedad de una operación mediante la cual puede ser aplicada a distintos tipos de objetos.

En Bases de Datos Orientadas a Objetos, los usuarios pueden definir operaciones sobre los datos como parte de la definición de la base de datos. Las Bases de Datos Orientadas a Objetos proporcionan ventajas tales como:

- La cantidad de información que puede modelarse en una base datos orientada a objetos se incrementa, y es más fácil modelar esta información.

- Los Sistemas de Bases de Datos Orientados a Objetos son capaces de tener mayores capacidades de modelado por medio de la extensibilidad.
- En una Base de Datos Orientada a Objetos, el manejo de versiones está disponible para ayudar a modelar cambios diversos a los sistemas.
- La reutilización de clases juega un rol vital en el desarrollo y mantenimiento más rápido de aplicaciones. Las clases genéricas son potentes, pero más importante es que ellas pueden ser usadas nuevamente.

Las Bases de Datos Orientadas a Objetos brindan beneficios más allá de los modelos tradicionales de bases de datos, pero poseen desventajas que no debieran ser pasadas por alto tales como:

- Los sistemas entidad relación tradicionales han estado en el mercado por un largo tiempo y un cambio se apartaría de las ideas establecidas, requeriría una filosofía diferente de pensar, y en algunos casos los usuarios no tendrían los conocimientos orientados a objetos necesarios para trabajar con sistemas de bases datos orientados a objetos. La educación de la gente en base a la filosofía orientada a objetos es un proceso muy minucioso, requeriría gran cantidad de tiempo, dinero y otros recursos. También, y con motivo del cambio, se requeriría más tiempo para mover los datos a los nuevos sistemas manejadores de Bases de Datos Orientadas a Objetos.
- La necesidad de que los sistemas tradicionales y los sistemas manejadores de bases de datos orientados a objetos se comuniquen y trabajen juntos, debido a que las bases de datos orientadas a objetos frecuentemente son incompatibles entre ellas. Esto hace imposible migrar una aplicación desde una Base de Datos Orientada a Objetos a otra.

1.2.5 Modelo de Bases de Datos Distribuidas.

En un Modelo de Bases de Datos Distribuidas, los datos se almacenan en varios computadores. Los computadores de un sistema distribuido se comunican entre sí a través de diversos medios de comunicación, tales como cables de alta velocidad o líneas telefónicas. No comparten la memoria principal ni el reloj[7].

Un Sistema Distribuido de Bases de Datos consiste en un conjunto de localidades, cada uno de las cuales puede participar en la ejecución de transacciones que accedan a datos de una o varias localidades[7]. La diferencia principal entre los Sistemas de Bases de Datos Centralizados y Distribuidos es que, en los primeros, los datos residen en una sola localidad, mientras que, en los últimos, se encuentran en varias localidades.

1.2.5.1 Consideraciones generales de los modelos de Bases de Datos Distribuidas.

Un Modelo Distribuido de Bases de Datos consiste en un conjunto de localidades, cada una de las cuales mantiene un sistema de bases de datos local. Cada localidad puede procesar transacciones locales, o bien transacciones globales entre varias localidades, requiriendo para ello comunicación entre ellas[7].

Existen varias razones para construir Sistemas Distribuidos de Bases de Datos que incluyen compartir la información, fiabilidad y disponibilidad y agilizar el procesamiento de las consultas. Pero también tiene sus desventajas, como desarrollos de software más costosos, mayor posibilidad de errores y costos extras de procesamiento.

La principal ventaja de los modelos distribuidos es la capacidad de compartir y acceder a la información de forma fiable y eficaz.

Otras ventajas son:

- **Utilización compartida de los datos y distribución del control:** Cada localidad puede controlar hasta cierto punto los datos almacenados localmente.
- **Fiabilidad y Disponibilidad:** Si se produce un fallo en una localidad de un sistema distribuido, no necesariamente implica la desactivación del sistema. El sistema debe detectar cuando falla una localidad y tomar las medidas necesarias para recuperarse del fallo. La disponibilidad es fundamental para los sistemas de bases de datos que se utilizan en aplicaciones de tiempo real.

La desventaja principal de los modelos distribuidos es la complejidad que se requiere para garantizar una coordinación adecuada entre localidades.

El aumento de la complejidad se refleja en:

- Coste del desarrollo de software: es más difícil estructurar un sistema de bases de datos distribuidos y por tanto su coste es mayor.
- Mayor posibilidad de errores: puesto que las localidades del sistema distribuido operan en paralelo, es más difícil garantizar que los algoritmos sean correctos.
- Mayor tiempo extra de procesamiento: el intercambio de mensajes y los cálculos adicionales son una forma de tiempo extra que no existe en los sistemas centralizados.

1.3 Sistemas Manejadores de Bases de Datos.

Un Sistema Manejador de Bases de Datos (DBMS) es un conjunto de programas que se encargan de manejar la creación y todos los accesos a las bases de datos y está compuesto por:

Data Definition Language (DDL): Lenguaje de definición de datos.

Data Manipulation Language (DML): Lenguaje de manipulación de datos.

Structured Query Language (SQL): Lenguaje de consulta estructurado

Algunos de los objetivos de los manejadores de Bases de Datos son:

- Crear una colección integrada de datos disponibles a una amplia variedad de usuarios.
- Proveer calidad e integridad de los datos.
- Asegurar la privacidad a través de medidas de seguridad.
- Permitir un control centralizado de la base de datos, para una administración más eficiente.

Los DBMS más comunes son Oracle, SqlServer, Informix, Sysbase, MySQL, PostgreSql, Magic, Firebird.

En 1964, se conciben los primeros Sistema Gestores de Bases de Datos (SGBD), con estos se pretendió dar un viraje a los sistemas de archivos, los cuales se limitaban a la estructuración del almacenamiento físico de los datos. Con los SGBD se logró por medio de actividades integradas verlos físicamente en un

solo almacenamiento, pero lógicamente se manipulan a través de esquemas compuestos por estructuras donde se establecen vínculos de integridad, métodos de acceso y organización física sobre los datos, permitiendo así obtener valores agregados de utilización. En 1970 se propuso el Modelo de Datos Relacional, este modelo ha marcado la línea de investigación por muchos años y representa al mundo real mediante tablas relacionadas entre sí por columnas comunes. Viendo la necesidad de mejorar este estándar se desarrollaron los Sistemas Gestores de Bases de Datos Relacionales (SGBDR) cuyas características hacen al sistema mucho más eficiente que los sistemas de manejo de archivos.

Un SGBDR es un conjunto de datos relacionados entre sí y un grupo de programas para tener acceso a esos datos, permitiendo concurrencia y recuperación. La persistencia de un SGBDR hace referencia a la conservación de los datos después de la finalización del proceso que los creó y la concurrencia se refiere a la capacidad del sistema para gestionar múltiples usuarios interactuando al mismo tiempo.

1.4 Metodología de Diseño de Bases de Datos.

El Diseño de una Base de Datos es un proceso complejo que abarca decisiones a muy distintos niveles. La complejidad se controla mejor si se descompone el problema en subproblemas y se resuelve cada uno de estos independientemente, utilizando técnicas específicas. Así, el diseño de una base de datos se descompone en diseño conceptual, diseño lógico y diseño físico.

1.4.1 Diseño Conceptual.

El Diseño Conceptual parte de las especificaciones de requisitos de usuario y su resultado es el esquema conceptual de la base de datos. Un esquema conceptual es una descripción de alto nivel de la estructura de la base de datos, independientemente del Sistema Gestor de Bases de Datos (SGBD) que se vaya a utilizar para manipularla. Un Modelo Conceptual es un lenguaje que se utiliza para describir esquemas conceptuales. El objetivo del diseño conceptual es describir el contenido de información de la base de datos y no las estructuras de almacenamiento que se necesitarán para manejar esta información[8].

En esta etapa se debe construir un esquema a partir de la información que se ha capturado del entorno, independientemente de cualquier consideración física. A este esquema se le denomina esquema

conceptual. Al construir el esquema, los diseñadores descubren el significado de los datos. Se encuentran entidades, atributos y relaciones.

El objetivo es comprender:

- La perspectiva que cada usuario tiene de los datos.
- La naturaleza de los datos, independientemente de su representación física.
- El uso de los datos a través de las áreas de aplicación.

El esquema conceptual se puede utilizar para que el diseñador transmita lo que ha entendido sobre la información que se maneja. Para ello, ambas partes deben estar familiarizadas con la notación utilizada en el esquema.

1.4.2 Diseño Lógico.

El Diseño Lógico parte del esquema conceptual y da como resultado un esquema lógico. Un esquema lógico es una descripción de la estructura de la base de datos en términos de las estructuras de datos que puede procesar un tipo de Sistema Gestor de Bases de Datos (SGBD)[9]. Un modelo lógico es un lenguaje usado para especificar esquemas lógicos. El diseño lógico depende del tipo de SGBD que se vaya a utilizar, no depende del producto concreto.

El Diseño Lógico es el proceso de construir un esquema de la información que utiliza la empresa, basándose en un modelo de bases de datos específico, independiente del SGBD concreto que se vaya a utilizar y de cualquier otra consideración física[9].

En esta etapa, se transforma el esquema conceptual en un esquema lógico que utilizará las estructuras de datos del modelo de base de datos en el que se basa el SGBD que se vaya a utilizar, como puede ser el modelo relacional, el modelo de red, el modelo jerárquico o el modelo orientado a objetos. Conforme se va desarrollando el esquema lógico, éste se va probando y validando con los requisitos de usuario.

La normalización es una técnica que se utiliza para comprobar la validez de los esquemas lógicos basados en el modelo relacional, ya que asegura que las relaciones obtenidas no tienen datos redundantes.

El esquema lógico es una fuente de información para el Diseño Físico. Además, juega un papel importante durante la etapa de mantenimiento del sistema, pues permite que los futuros cambios que se realicen sobre los programas de aplicación o sobre los datos, se representen correctamente en la base de datos.

Tanto el Diseño Conceptual, como el Diseño lógico, son procesos iterativos, tienen un punto de inicio y se van refinando continuamente. Ambos se deben ver como un proceso de aprendizaje en el que el diseñador va comprendiendo el funcionamiento de la empresa y el significado de los datos que maneja. El Diseño Conceptual y el Diseño Lógico son etapas clave para conseguir un sistema que funcione correctamente.

1.4.3 Diseño Físico.

El Diseño Físico parte del esquema lógico y da como resultado un esquema físico. Un esquema físico es una descripción de la implementación de una base de datos en memoria secundaria: las estructuras de almacenamiento y los métodos utilizados para tener un acceso eficiente a los datos[10]. Por ello, el Diseño Físico depende del SGBD concreto y el esquema físico se expresa mediante su lenguaje de definición de datos.

El Diseño Físico es el proceso de producir la descripción de la implementación de la base de datos en memoria secundaria: estructuras de almacenamiento y métodos de acceso que garanticen un acceso eficiente a los datos[10].

Para llevar a cabo esta etapa, se debe haber decidido cuál es el SGBD que se va a utilizar, ya que el esquema físico se adapta a él. Entre el Diseño Físico y el Diseño Lógico hay una realimentación, ya que algunas de las decisiones que se tomen durante el diseño físico para mejorar las prestaciones, pueden afectar a la estructura del esquema lógico.

En general, el propósito del Diseño Físico es describir cómo se va a implementar físicamente el esquema lógico obtenido en la fase anterior. Concretamente, en el modelo relacional, esto consiste en:

- Obtener un conjunto de relaciones y las restricciones que se deben cumplir sobre ellas.
- Determinar las estructuras de almacenamiento y los métodos de acceso que se van a utilizar para conseguir unas prestaciones óptimas.
- Diseñar el modelo de seguridad del sistema.

1.5 Diseño de un Centro de Datos.

El término Centro de Datos engloba al menos cuatro diferentes significados para cuatro diferentes tipos de personas. Hay quienes argumentan que un Centro de Datos es el cuarto donde se almacenan los servidores. Otros visualizarán una perspectiva radicalmente diferente. Es verdad que en cierto momento, el Centro de Datos era más pequeño que el cuarto protegido de servidores. Sin embargo, con los avances tecnológicos y los negocios actuales de centrales de información, el término mejor expresado sería Centro de Datos de misión crítica. Los modelos de negocios han pasado por un ciclo completo de ser sitios de datos centralizados a descentralizados y nuevamente centralizados. Los negocios están tomando conciencia de que los datos son su valor más poderoso y que se deben hacer enormes esfuerzos para asegurar su disponibilidad, seguridad y redundancia.

El concepto de Centro de Datos se ha desarrollado dentro de su propio modelo de negocios. Las compañías que proporcionan almacenamiento redundante y fuera de sitio a otras compañías están construyendo instalaciones “state of the art” a escala global. En el corazón de estas instalaciones está la infraestructura IT. Este documento tratará acerca de la infraestructura y los componentes del Centro de Datos del Proyecto Registro y Notarías.

1.5.1 Tendencias de los Centros de Datos.

Los Centros de Datos almacenan la información de aplicaciones Enterprise Resource Planning (ERP), aplicaciones de comercio, Supply Chain Management (SCM), convergencia de video/voz/datos, aplicaciones Business to Business (B2B), así como las aplicaciones de back-office que son críticas en la

operación de una empresa[11]. Un riesgo muy importante es el downtime, que se traduce en pérdidas monetarias. Las empresas que ofrecen equipos y componentes para los Centros de Datos son sensibles a todo esto y han conseguido grandes avances en proveer soluciones viables para los requerimientos de almacenamiento.

1.5.2 Planificación y Diseño de un Centro de Datos.

Los Centros de Datos deben ser cuidadosamente planeados antes de construirse para asegurar su conformidad con todas las normas y reglamentos aplicables. Las consideraciones de diseño incluyen selección de sitio y ubicación, espacio, electricidad, capacidad de enfriamiento, carga de piso, acceso y seguridad, limpieza ambiental, prevención de peligros y crecimiento[11]. Para poder calcular las necesidades anteriores, el arquitecto debe conocer los componentes que tendrá el Centro de Datos.

Es importante conocer el número de usuarios, tipos de aplicaciones y plataformas, unidades de rack requeridas y más importante, considerar el crecimiento esperado. Un Centro de Datos debe tener la capacidad de responder a cualquier cambio de equipo, estándares y demanda, y al mismo tiempo permanecer administrable y sobre todo, confiable.

1.5.3 Componentes de un Centro de Datos.

Los Centros de Datos están compuestos por sistemas de comunicaciones de red de alta velocidad y alta demanda capaz de manejar el tráfico para Storage Area Networks (SAN), Network Attached Storage (NAS), granja de servidores de archivos/aplicaciones/web, y otros componentes localizados en ambientes controlados. El control de ambiente se relaciona a la humedad, inundación, electricidad, temperatura, control de fuego, y por supuesto, acceso físico. Las comunicaciones dentro y fuera del centro de datos se proveen por enlaces WAN, CAN/MAN y LAN en una variedad de configuraciones dependiendo de las necesidades particulares de cada Centro de Datos.

Un Centro de Datos diseñado apropiadamente proporcionará disponibilidad, accesibilidad, escalabilidad, y confiabilidad 24 horas al día, 7 días a la semana, y 365 días al año descontando el tiempo fuera de servicio por mantenimiento.

Los Centros de Datos críticos se monitorean vía Network Operations Center (NOC) el cual puede ser local (in-house) o subcontratado a un tercero. El NOC es el primer lugar donde se realizan las revisiones y el punto de partida para las acciones correctivas. Los NOCs se implementan generalmente durante las horas de operaciones de los Centros de Datos. Los dispositivos de monitoreo de equipos avisarán al NOC de problemas tales como sobrecalentamiento, caídas de equipos, y fallas de componentes por medio de una serie de interruptores que pueden configurarse directamente en el equipo o por medio de un software de monitoreo de terceros el cual corre a través del equipo.

En esta sección definiremos algunos componentes que forman parte del Centro de Datos montado en el Ministerio del Poder Popular para la Relaciones Interiores y Justicia de la República Bolivariana de Venezuela.

1.5.3.1 Definición de Rack.

Los racks son muy útiles en un centro de datos, donde el espacio es escaso y se necesita alojar un gran número de dispositivos.

Estos dispositivos suelen ser:

- Servidores cuya carcasa ha sido diseñada para adaptarse al bastidor. Existen servidores de 1U, 2U y 4U, y recientemente, se han popularizado los servidores blade que permiten compactar más de veinte servidores en una altura de 4U, compartiendo fuentes de alimentación y cableado.
- Conmutadores y enrutadores de comunicaciones.
- Cortafuegos.
- Sistemas de audio.

El equipamiento simplemente se desliza sobre un raíl horizontal y se fija con tornillos. También existen bandejas que permiten apoyar equipamiento no normalizado.

1.5.3.2 Definición de Storage Area Network (SAN).

Una SAN es la aplicación de la tecnología Fibre Channel para crear una red propia dedicada exclusivamente a dispositivos de almacenamiento[12].

Mediante una implementación SAN creamos una red propia de almacenamiento, lo que permite compartir recursos de coste elevado, como puede ser un sistema RAID o una librería para copia de seguridad, a alta velocidad entre varios usuarios sin ningún impacto en la LAN (red de área local).

La implementación de una red SAN nos proporciona la manera más racional de gestionar y administrar los dispositivos de almacenamiento de forma dedicada y especializada, tanto en plataformas homogéneas como heterogéneas, de forma escalable y segura, permitiendo además mantener la inversión efectuada hasta la fecha en dispositivos SCSI de coste elevado[12].

Algunas ventajas que aporta una SAN son:

- **Protección de la inversión actual y futura:** Una ventaja primordial de la SAN es su compatibilidad con los dispositivos SCSI ya existentes, aprovechando las inversiones ya realizadas y permitiendo el crecimiento a partir del hardware ya existente. Mediante el empleo de dispositivos modulares como hubs, switches, bridges y routers, se pueden crear topologías totalmente flexibles y escalables.
- **Gran ancho de banda:** Ancho de banda actual de hasta 200 Mbytes/segundo con doble adaptador Fibre Channel. La tecnología Fibre Channel permite un incremento del ancho de banda efectivo de entre 2,5 y 10 veces la obtenida sobre una plataforma SCSI. Si en la actualidad el ancho de banda es de 1 Gb/s, el nuevo estándar especifica anchos de entre 2 a 4 Gb/s.
- **Centralización del backup/ Copia de seguridad independiente de la LAN:** el sistema de copia se conecta a la SAN, por lo que es posible realizar el backup online, sin afectar al trabajo de los usuarios y ejecutándose en un tiempo mínimo con un impacto prácticamente nulo en el servidor.
- **Tolerancia a fallos permitiendo la ejecución de aplicaciones críticas 24x7:** La utilización de sistemas RAID tolerantes al fallo se ha generalizado en entornos corporativos. Sin embargo y debido a que la velocidad de proceso no debe penalizarse al emplear información redundante, la

potencia de cálculo y proceso de los sistemas RAID hace que tales dispositivos tengan un coste elevado. Una manera de compensar este coste elevado es emplearlo no tan sólo para las aplicaciones críticas corporativas sino para todos los usuarios de la red. Ello requiere que el sistema sea compartido. Fibre Channel por sus características de red y por su elevado ancho de banda convierte a este requerimiento en estándar.

- **Compartición de ficheros entre servidores en entornos heterogéneos:** (Unix, NT, MacOs).
- **Alta escalabilidad y larga distancia entre nodos de la red:** Dependiendo de la topología SAN utilizada podemos interconectar hasta 126 nodos por bucle en Arbitrated Loop, a distancias entre ellos de 30 metros en el caso de utilizar cable de cobre y de hasta 10 Km si se emplea cable de fibra óptica; o más de 16 millones en la topología Fabric, alcanzando capacidades de cientos de Terabytes. Esta escalabilidad es impensable en entornos SCSI y permite planificar una red SAN simplemente añadiendo dispositivos a medida que las necesidades lo requieren.
- **Alta disponibilidad:** Fibre Channel incluye soporte de conexión dual loop. Con ello se proporciona un camino alternativo a la señal en el caso de que un cable falle o sea accidentalmente desconectado. De nada sirve un sistema RAID tolerante al fallo si el único camino para acceder a él se interrumpe, bien sea por el fallo de un componente o por la desconexión accidental de un cable.
- **Fácil administración/ Gestión centralizada:** Debido a que la tecnología Fibre Channel no es la evolución de una tecnología existente sino que parte de cero, ha podido definir libremente una serie de nuevas posibilidades imposibles de implementar hasta ahora. Uno de estos aspectos es la gestión global de todos y cada uno de los dispositivos de almacenamiento que forman parte de la red. Además debido a que Fibre Channel es un medio de transmisión, es independiente del protocolo que transporta. Ello hace que podamos utilizar protocolos ampliamente extendidos en la industria como SES (SCSI Enclosure Services), SAF-TE (SCSI Accessed Fault Tolerant Enclosure), SMART (Self Monitoring Analysis and Reporting Technology), SNMP (Simple Network Management Protocol) y WBEM (Web-based Enterprise Management). La mayoría de las aplicaciones de administración de red emplean estos protocolos, mediante los cuales podemos controlar y optimizar el tráfico de toda la red, diagnosticar de manera más rápida y eficiente problemas, evitando caídas del sistema con el ahorro de costes que ello supone.

- **Fácil Integración:** Las posibilidades de conectividad junto con la posibilidad de realizar hot plug de dispositivos facilitan la adición de componentes sin detener el servidor o servidores. Además, la flexibilidad de poder utilizar indistintamente cable de cobre o fibra óptica en cualquier punto de la red, facilita la instalación e implementación de futuras expansiones.
- **Bajo costo de mantenimiento:** Es fácil deducir que disponer de una red con dispositivos hot swap, con potentes herramientas de gestión y administración, facilita enormemente las tareas de mantenimiento. Además, estas herramientas proporcionan elementos de análisis que permiten diagnosticar, incluso antes de que se produzcan, problemas en la red. Con ello se obtienen ahorros derivados del tiempo de no utilización del acceso a los datos. Además, la gestión centralizada gracias a una SAN reduce drásticamente los gastos de gestión del almacenamiento y aumenta tanto la consistencia y dimensión del control de los administradores del sistema, como la disponibilidad de los datos.

1.5.3.3 Definición de Zona Desmilitarizada (DMZ).

En informática, una zona desmilitarizada (DMZ) o red perimetral es una red local (una subred) que se ubica entre la red interna de una organización y una red externa, generalmente Internet[13]. El objetivo de una DMZ es que las conexiones desde la red interna y la externa a la DMZ estén permitidas, mientras que las conexiones desde la DMZ sólo se permitan a la red externa. Los equipos (hosts) en la DMZ no se pueden conectar con la red interna. Esto permite que los equipos (hosts) de la DMZ puedan dar servicios a la red externa a la vez que protegen la red interna en el caso de que intrusos comprometan la seguridad de los equipos (hosts) situados en la zona desmilitarizada. Para cualquiera de la red externa que quiera conectarse ilegalmente a la red interna, la zona desmilitarizada se convierte en un callejón sin salida.

1.5.3.4 Definición de Virtual LAN (VLAN).

Una VLAN es una red de computadoras lógicamente independiente. Varias VLANs pueden coexistir en un único switch físico[14].

Una VLAN consiste en una red de computadores que se comportan como si estuviesen conectados al mismo cable, aunque pueden estar en realidad conectados físicamente a diferentes segmentos de una red

de área local. Los administradores de red configuran las VLANs mediante software en lugar de hardware, lo que las hace extremadamente flexibles. Una de las mayores ventajas de las VLANs surge cuando se traslada físicamente una computadora a otra ubicación: puede permanecer en la misma VLAN sin necesidad de ninguna reconfiguración de hardware.

1.5.3.5 Definición de HP StorageWorks Enterprise Virtual Array.

HP StorageWorks Enterprise Virtual Array 8000 (EVA8000) es una potente solución de almacenamiento con un alto rendimiento que elimina los costos del almacenamiento tradicional. HP StorageWorks Enterprise Virtual Array permite el almacenamiento con una solución integrada y virtualizada de manera concurrente de los datos tanto de la aplicación como de los archivos.

1.6 Caracterización de la Suite de Productos de Oracle.

La Arquitectura de Bases de Datos Descentralizada tiene como herramienta fundamental para su implementación Oracle Database 10g release 2. A continuación se expondrán algunas de sus potencialidades.

Oracle es básicamente un herramienta cliente-servidor para la gestión de bases de datos, es un producto vendido a nivel mundial, aunque la gran potencia que tiene y su elevado precio hace que solo se vea en empresas muy grandes y multinacionales, por norma general. Es un manejador de Bases de Datos Relacional que hace uso de los recursos de los sistemas informáticos en todas las arquitecturas de hardware, lo que permite garantizar su aprovechamiento en ambientes cargados de información, por su capacidad de almacenar y acudir a los datos de forma recurrente.

El manejador de Bases de Datos Oracle, surgió a final de los años 70 y principio de los años 80. George Koch y su equipo fueron los primeros en desembarcar en el terreno de Oracle en 1982, durante un proceso de evaluación de sistema de gestión de bases de datos para una importante aplicación comercial. Oracle, conocida entonces como Relational Software, tenía poco más de 25 empleados en aquel tiempo y solo unos pocos clientes importantes. Sin embargo, cuando se completo el estudio, Oracle fue declarada vencedora. George Koch afirmó que el Sistema Gestor de Bases de Datos (SGBD) Oracle era técnicamente el mejor producto del mercado.

Actualmente es el mayor y más usado Sistema Manejador de Bases de Datos Relacionales (RDBMS) en el mundo. La Corporación Oracle ofrece este RDBMS como un producto incorporado a la línea de producción. Además incluye varias generaciones de desarrollo de aplicaciones, herramientas de reportes y utilitarios.

Oracle es soportado en computadoras personales (PC), microcomputadoras, computadoras con procesamiento paralelo masivo. Soporta unos 17 idiomas, funciona automáticamente en más de 80 arquitecturas de hardware y software distintos sin tener la necesidad de cambiar una sola línea de código. Esto es porque más del 80% de los códigos internos de Oracle son iguales a los establecidos en todas las plataformas de sistemas operativos.

Oracle es un sistema comercial que aporta un SGBD que ofrece las particularidades básicas para trabajar en entornos multi-usuario. Como sistema gestor de bases de datos, es actualmente uno de los paquetes de software más ampliamente extendidos en todas las compañías que tienen que gestionar una cantidad importante de información.

Oracle es uno de los sistemas más conocidos, que alcanza hoy en día un buen nivel de madurez y de profesionalidad gracias especialmente a:

- Su transportabilidad, funciona hoy en día sobre decenas de plataformas.
- La potencia de sus instrumentos de desarrollo de aplicaciones.
- La riqueza de su diccionario de datos.
- Los mecanismos encargados de la seguridad y la confidencialidad.
- Una experiencia probada sobre el terreno y una buena presencia de Oracle a nivel de formación, consejo y soporte técnico.

La versión 10g de Oracle salió al mercado en febrero del 2004, primero en su versión para UNIX y posteriormente en sus versiones para Linux y Windows. La novedad más llamativa de esta versión es que pone la "g" en el nombre de la versión. Es la capacidad de estos servidores de funcionar según el paradigma de "Grid" o rejilla. La novedad principal de Oracle 10g descansa precisamente en la preparación de dicho software para poder encajar en este modelo. Además, como parte de esta manera de entender el

negocio, Oracle también ha hecho cambios en el área de marketing. Así, la letra "i" que se asociaba desde hace seis años a la marca y que representaba la entrada en la era de Internet de la compañía ha sido sustituida por la "g" de Grid, de tal forma que la anterior versión Oracle 9i cambia a Oracle 10g. No obstante, ciertos usuarios valoran más las mejoras en la administración y la integración de algunos elementos que previamente no funcionaban correctamente juntos.

Al instalar la base de datos, también se instala el nuevo Oracle Enterprise Manager Database Control, basado en la Web, que será la herramienta primaria para manejo de la base de datos y establece un nuevo estándar en cuanto a facilidad de uso, ya que es un entorno visual e intuitivo, sin necesidad de usar el texto como medio de comunicación con la base de datos.

La suite de productos de Oracle contempla las siguientes herramientas:

- **Oracle Database 10g Standard Edition:** Es una base de datos de características completas para pequeñas y medianas empresas que requieren el desempeño, la disponibilidad y la seguridad de la base de datos número uno del mundo a un bajo costo. Disponible en un solo servidor o en servidores en cluster con hasta cuatro procesadores, Oracle Database 10g Standard Edition es la opción segura para desarrollar e implementar de manera económica las aplicaciones de la base de datos[15]. Este producto está considerado una base de datos multiusuario pero con un número limitado de usuarios. Actualmente existe para Windows, Unix y Linux.
- **Oracle Database 10g Enterprise Edition:** Ofrece confiabilidad, escalabilidad y desempeño de primer nivel para configuraciones en cluster y en un solo servidor[16]. Ofrece las más completas características para soportar el procesamiento de transacciones más exigente, inteligencia de negocios, y aplicaciones para la administración de contenido. Oracle Database 10g Enterprise Edition brinda desempeño y escalabilidad de récord absoluto para el procesamiento de transacciones y depósitos de datos de gran escala en Windows, Linux, y servidores UNIX.
- **Oracle Database Lite:** Es un entorno completo e integrado para crear, implementar y gestionar aplicaciones móviles para una amplia gama de dispositivos[17]. Oracle Database Lite incluye una base de datos compacta y relacional para dispositivos móviles y un Mobile Server para sincronización con Oracle Database, aprovisionamiento de aplicaciones y gestión de dispositivos.

- **Oracle TimesTen In Memory Database:** Base de Datos relacional en memoria, y en tiempo real orientada al nivel medio para dar soporte a las necesidades de gestión de datos de las aplicaciones con desempeño clave[18]. Se planea que las versiones futuras incluyan mayor compatibilidad e interoperabilidad con Oracle Database.
- **Oracle Berkeley DB:** Familia de tres bases de datos embebidas, de origen abierto y alto desempeño. Nuevas versiones de Berkeley DB y Berkeley DB Java Edition fueron lanzadas en septiembre de 2006, y Berkeley DB XML 2.3 fue recientemente lanzada en diciembre de 2006[19].
- **Oracle Enterprise manager 10g Grid Control:** Oracle Enterprise manager 10g Grid Control es la consola central de administración y el entorno que automatiza las tareas administrativas para el conjunto de sistemas implicados en un entorno grid[20]. Esta consola ayuda a reducir los costes de administración; con ella, los DBAs pueden agrupar múltiples nodos hardware como bases de datos, servidores de aplicación, servidores Web etc. como si fueran unidades lógicas. Ejecutando trabajos, diseñando políticas, monitorizando el rendimiento y automatizando muchas otras tareas sobre un conjunto de destinos en vez de sobre muchos sistemas individuales, la OEM Grid Control permite escalar un grid fácilmente. Con OEM Grid Control, Oracle 10g automatiza la instalación, configuración y clonación de servidores de aplicación y de bases de datos sobre múltiples nodos. Este entorno puede utilizarse tanto para la adición de nuevos sistemas como para aplicar parches o añadir utilidades a sistemas ya existentes. También mantiene la sincronía entre los nodos.
- **Oracle Application Server 10g:** Oracle Application Server 10g proporciona una plataforma para desarrollar y ejecutar aplicaciones empresariales, integrando muchas funciones por ejemplo un entorno de ejecución para Web Services y J2EE, complementos de Business Intelligence o una Web caché entre otras, aparte de características especialmente enfocadas al grid[21].
- **Real Application Cluster (RAC):** Oracle Real Application Clusters permite que una única base de datos se expanda por múltiples nodos en un grid o red, uniendo los recursos de varias máquinas[22]. Esto que requería un proceso en versiones anteriores del servidor se puede hacer inmediatamente en Oracle 10g, y se puede empezar a balancear el flujo de trabajo hacia la nueva máquina que se incorpora al grid, a la vez que abandonarla cuando ya no es necesaria. Otros Sistemas de Bases de Datos no pueden hacer esto dinámicamente cuando la Base de Datos se encuentra ejecutándose. El nuevo software de cluster en Oracle 10g simplifica el proceso eliminando la necesidad de adquirir, instalar y configurar estas herramientas de terceros. Se pueden añadir servidores a la vez que eliminarlos en un cluster Oracle sin

tiempo de inactividad, es decir, sin detener la base de datos, sin importar tampoco la plataforma donde se encuentra instalado el servidor.

Conclusiones del Capítulo.

El capítulo ha sido un breve recorrido por conceptos y definiciones que son parte intrínseca del trabajo, a partir de los cuales nos apoyaremos como puntos de reflexión para posibles tomas de decisiones. El documento ha centrado su atención en los modelos de datos como base conceptual para el diseño de aplicaciones que hacen un uso intensivo de los datos; en este sentido nos inclinamos para el diseño del modelo de datos en el uso de un Modelo de Datos Relacional orientado a un ambiente distribuido como un simple y eficaz modelo para manipular los datos, además de estar sustentado en sólidas bases teóricas y fundamentos planteados en tópicos anteriores. Como DBMS encargado de manejar la creación y el acceso a los datos nos inclinamos por Oracle Database 10g release 2 por su capacidad como software de funcionar según el paradigma Grid o Rejilla, además de responder a las necesidades de negocio con alta tecnología y bajo costo con el fin de incrementar la eficiencia y la productividad. A partir del alojamiento de aplicaciones críticas que demandan una disponibilidad de 24x7, se propuso el diseño e implementación de un Centro de Datos según los establecen las normas y reglamentos para su construcción. El Centro de Datos está diseñado apropiadamente para proporcionar disponibilidad, accesibilidad, escalabilidad, y confiabilidad las 24 horas al día, los 7 días a la semana, y los 365 días al año descontando el tiempo fuera de servicio por mantenimiento.

CAPÍTULO II: DISEÑO E IMPLEMENTACIÓN DE UNA ARQUITECTURA DE BASES DE DATOS DESCENTRALIZADA EN ORACLE DATABASE 10G RELEASE 2

Introducción.

En el presente capítulo se presentan una serie de elementos importantes para el diseño eficaz de Bases de Datos, tales como las diferentes estructuras que conforman la misma, así como diferentes arquitecturas y pasos a tener en cuenta por el diseñador.

2.1 Arquitectura de una Base de Datos Oracle.

Un servidor Oracle es un Sistema de Gestión, no de administración de Bases de Datos, que provee un abierto, comprensible e integrado acercamiento al manejo de la información. Está compuesto por una o varias instancias y sus respectivas Bases de Datos.

Cada Base de Datos que corre en un servidor Oracle es asociada a una instancia. Cuando se inicia la base de datos, el software de Oracle asigna un área de memoria compartida llamada Área Global del Sistema (SGA) e inician varios procesos en background. Esta combinación del SGA y los procesos de Oracle es lo que conforma una instancia Oracle.

Luego de iniciarse una instancia, esta es asociada con una Base de Datos específica. A este proceso se le conoce como montaje de la Base de Datos. La misma queda lista para ser abierta, lo que permite ser accedida por los usuarios autorizados.

En una misma PC pueden correr concurrentemente múltiples instancias, cada una accediendo a su propia Base de Datos. Todas las estructuras de memoria de Oracle existen en la memoria principal de la computadora donde reside el servidor Oracle. Los procesos de Oracle son tareas que corren en la memoria de estas computadoras y se definen como un hilo de control o mecanismo que ejecuta una serie de pasos.

2.1.1 Estructuras Lógicas y Físicas de una Base de Datos Oracle.

Una Base de Datos Oracle se divide en estructuras de almacenamiento lógicas llamadas tablespaces, que relacionan estructuras que poseen características comunes.

Cada Base de Datos se divide lógicamente en uno o más tablespaces. Uno o más datafiles son explícitamente creados para almacenar físicamente los datos de todas las estructuras lógicas dentro de un tablespace. Si es un tablespace temporal, se crea un archivo temporal en vez de un datafile.

Los esquemas son una colección de objetos que pertenecen a un usuario de una base de datos, son las estructuras lógicas que se refieren directamente a los datos de la base de datos. Incluyen tablas, vistas, secuencias, procedimientos almacenados, sinónimos, índices, clusters y links de bases de datos, en general todo lo que la aplicación crea en la base de datos.

En el nivel más bajo de granularidad Oracle almacena los datos en bloques de datos que corresponden a un número específico de bytes de espacio físico en disco de la base de datos. Se especifica su tamaño para cada tablespace que se crea en la base de datos.

El siguiente nivel de almacenamiento lógico de una Base de Datos Oracle es el extent. Es un número específico y continuo de bloques de datos que son utilizados para almacenar un tipo determinado de información.

El nivel lógico de almacenamiento en Bases de Datos Oracle superior al extent es el segmento. Es un grupo de extents asignados para una determinada estructura lógica.

La Base de Datos Oracle asigna espacio dinámicamente. Cuando los existentes extents de un segmento están llenos, se adicionan más extents adicionales, y estos pueden o no estar continuos en el disco.

Las estructuras lógicas de la base de datos se almacenan en los archivos físicos de la misma. Algunas de las estructuras físicas de una base de datos Oracle son:

- **Archivos de Control:** Contienen información acerca de la estructura física de la base de datos. Son críticos en una base de datos, sin ellos no se puede abrir la base de datos, ni acceder a sus datos.
- **Archivos de datos:** Contienen los datos de un usuario o aplicación de la base de datos Oracle.
- **Archivos de redo:** Permiten la recuperación de la instancia de una base de datos. Si ocurre una ruptura en la base de datos y no se pierde ningún archivo de datos, entonces la instancia puede recuperar la base de datos con la información en estos archivos.
- **Archivos de Resguardo (Backup):** Se usan para la recuperación de la base de datos. Cuando se daña o se borra un archivo debido a alguna falla se puede recuperar los datos al restaurar un archivo de resguardo.

2.2 Diseño General de la Arquitectura de la Base de Datos.

A partir de la necesidad de automatizar los procesos registrales que se realizan en cada uno de los Registro Públicos y Mercantiles del país, unido a la centralización de la información en un Centro de Datos y todo ello enfocado a un entorno de alta disponibilidad 24x7 se propuso dos posibles Diseños de Arquitectura de Bases de Datos.

Arquitectura Centralizada

Esta variante de diseño pretendía el uso de una Arquitectura Centralizada como propuesta de solución. Con la Arquitectura Centralizada se comprendería el diseño de un Centro de Datos que estaría en interacción directa por medio de una VPN (Virtual Private Network) con los doscientos cincuenta y cinco Registros Públicos y Mercantiles de la República Bolivariana de Venezuela.

Algunas ventajas asociadas a esta variante son:

- Menor probabilidad de errores, pues es más fácil garantizar que los algoritmos se ejecuten correctamente.
- Fácil control y administración de la Base de Datos.

- Eliminación de mecanismos de replicación de datos entre las Bases de Datos Locales y el Centro de Datos propensos a errores con bastante frecuencia.
- Menores costos de Software y Hardware, pues es más fácil estructurar un Sistema Centralizado, en lugar de un Sistema Descentralizado.
- Baja disponibilidad y fiabilidad, debido a que la falla de conectividad en una localidad puede producir la desactivación del sistema local.
- Incremento del User Time Response: los tiempos de respuesta al cliente se ven incrementados por la baja latencia de la red.

Arquitectura Descentralizada

La variante de Arquitectura de Bases de Datos Descentraliza propone el diseño de una Base de Datos Central que interactúa con otras doscientas cincuenta y cinco Bases de Datos Locales mediante mecanismos de replicación de datos soportados por Oracle Database 10g release 2. Estas Bases de Datos Locales son las encargadas de gestionar todo el flujo de información en cada Registro Público y Mercantil del país.

Las ventajas que proporcionan la implementación de esta segunda variantes las exponemos a continuación:

- Fiabilidad y disponibilidad, producto de que la falla de conectividad hacia una localidad no produce la desactivación del sistema local.
- Disminución de los niveles de procesamiento en la Base de Datos Central.
- Funcionamiento offline de los sistemas locales: si se produce el fallo de algún componente del Centro de Datos no demanda la paralización de los sitios de información locales.
- Disminución considerable de los tiempos de respuesta al cliente como resultado de una adecuada latencia de red local, y por consecuencia de una mejora ostensible en los tiempos de procesamiento de consultas.

Como resultado de la caracterización de cada variante de diseño, se decidió implementar el diseño de una Arquitectura de Bases de Datos Descentralizada. A pesar de que esta Arquitectura incurre en gastos superiores a la primera variante de diseño y posibles situaciones de errores asociadas a la complejidad que demanda la implementación de un Sistema Descentralizado, existen sólidas bases para la construcción de un Sistema Descentralizado de Bases de Datos que incluya fiabilidad, disponibilidad y agilice el procesamiento de consultas en cada subsistema en correspondencias a las expectativas esperadas por el cliente.

2.2.1 Diseño de Servidores Locales.

Hoy en día, las pequeñas empresas se ven obligadas a reducir el coste de su infraestructura. Debido a los presupuestos limitados, algunas empresas se han visto obligadas a utilizar PCs de sobremesa inapropiados como servidores básicos de pequeñas oficinas. Esto ha provocado problemas de disponibilidad y atraso en los sistemas que en muchos casos han imposibilitado la culminación de los mismos. A continuación brindamos las experiencias en el tema obtenidas durante el diseño de los servidores locales.

2.2.1.1 Selección del Hardware.

A partir del análisis de los aplicativos presentes en cada subsistema, se realizó un estudio de mercado en busca del hardware apropiado, capaz de cubrir los altos niveles de procesamiento que se pudieran presentar en cada subsistema.

Como consecuencia del análisis se decidió la compra de la Serie HP ProLiant ML110. Este es un equipo que pretende resolver el dilema de funcionalidad frente a precio, además de estar diseñado específicamente y sometido a pruebas exhaustivas para ejecutar sistemas operativos para servidores en configuraciones exigentes de servidores. El chasis ergonómico se ha desarrollado para facilitar la futura incorporación de memoria, unidades y dispositivos[23]. Algunas especificaciones del Servidor HP ProLiant ML110 son:

Características principales del procesador

- Intel Pentium 4 630 3 GHz
- Computación de 64 bits
- Velocidad bus de datos - 800 MHz
- Tipo conjunto de chips - Intel E7230
- Hyper-Threading Technology, Intel Extended Memory 64 Technology, Intel Execute Disable Bit, Enhanced Intel Speed Step Technology

Memoria Cache

- Tipo - L2
- Tamaño instalado - 2 MB
- Caché por procesador - 2 MB

Memoria RAM

- Tamaño instalado - 2 GB / 8 GB (máx.)
- Tecnología - DDR II SDRAM – ECC
- Velocidad de memoria - 533 MHz
- Conforme a la especificación de memoria - PC2-3200 y PC2-4200
- Factor de forma - DIMM de 240 espigas

Conexión de redes

- Conexión de redes - Adaptador de red – integrado.
- Protocolo de interconexión de datos - Ethernet, Fast Ethernet, Gigabit Ethernet.
- Protocolo de gestión remota - IPMI 1.5.
- Características - Wake on LAN (WoL).
- Cumplimiento de normas - IEEE 802.3, IEEE 802.3U, IEEE 802.1Q, IEEE 802.3ab, IEEE 802.1p, IEEE 802.3ad (LACP), IEEE 802.1x.

2.2.1.2 Software Base.

En los entornos corporativos de hoy en día, las pequeñas empresas consiguen más a partir de menos. Los profesionales de las tecnologías de la información pueden contribuir a lograr más mediante la implementación de Windows Small Business Server 2003 con el fin de proporcionar una infraestructura informática fiable con una seguridad mejorada. Windows Small Business Server 2003 proporciona una completa solución de servidor empresarial que ofrece funciones de mensajería y colaboración, almacenamiento seguro de datos, impresión confiable y la capacidad de ejecutar varias aplicaciones específicas simultáneamente. Y sus completas características de administración ayudan a los proveedores de tecnología a configurar, ejecutar y mantener las redes de forma productiva y eficiente[24].

Con la utilización de Windows Small Business Server 2003 se pretende que en cada Registro Público y Mercantil del país se pueda:

1. Proteger la información corporativa de forma automática. La información se protege de muchas maneras gracias a Windows Small Business Server 2003. Por ejemplo, el asistente para copias de seguridad garantiza la realización de una copia de seguridad de los datos según una programación periódica.
2. Lograr más cada día, Windows Small Business Server 2003 aporta una infraestructura estable y capaz que sirve de soporte a la informática móvil y que, en combinación con herramientas de colaboración y administración de la información, permite aumentar drásticamente la productividad de los usuarios.
3. Configurar y poner en funcionamiento rápidamente un servidor diseñado para ejecutarse en un pequeño entorno de negocio. Se pueden apreciar en seguida el sólido y ágil método de implementación de Windows Small Business Server 2003, que permite ahorrar tiempo y dinero.

Algunas especificaciones técnicas de Windows Small Business Server 2003 se presentan a continuación:

- Sistema Operativo seguro y confiable que proporciona herramientas y tecnologías destinadas a ayudar a compartir, administrar, proteger y realizar la copia de seguridad de archivos en una red interna.

- Incluye el servicio de directorio de Microsoft Active Directory y sus herramientas.
- Tecnología que ayuda a proteger las conexiones en Internet.
- Infraestructura de comunicación, mensajería y colaboración que contribuye a aumentar la productividad.

Se decidió utilizar Windows Small Business Server 2003, porque brinda las prestaciones para una red pequeña con características de seguridad mejoradas, además de funciones de almacenamiento, impresión, colaboración, correo electrónico, conexión a Internet, compatibilidad para trabajadores móviles y fax.

La utilización como Sistema de Gestor de Bases de Datos Oracle Standard Edition One, en primer lugar, porque además de estar diseñada para operar en máquinas de uno o dos procesadores, situación común en las Oficinas Registrales, la nueva versión tiene procesos de instalación, configuración, puesta en marcha y administración muy sencillos. Por otra parte, desde el punto de vista tecnológico, se trata de la misma Base de Datos Oracle Standard Edition, pero con un esquema de licenciamiento distinto que permite al segmento de las Pymes acceder a esta herramienta a través de dos modalidades: por procesador o bien por usuario designado con un mínimo de cinco[25].

Además Oracle Standard Edition One permite desarrollar todos los tipos de aplicaciones de negocios que requieren las Pymes como sistemas comerciales, de atención al cliente y gestión, ya sea en J2EE con los estándares presentes en el mercado o bajo cualquiera de las herramientas de Oracle disponibles.

Con la implementación de la Base de Datos sobre Oracle Standard Edition One se persigue:

1. Aprovechar el poder y el valor de una Base de Datos de clase empresarial con Oracle Database 10g Standard Edition One, la base de datos completa de Oracle, empaquetada y con un costo que permite su implementación en sistemas de servidores con dos procesadores.
2. Una rápida implementación y administración automatizada, con el mejor desempeño, confiabilidad y seguridad de la Base de Datos Oracle, líder del sector.

3. Escalabilidad, ya que posibilita a las empresas crecer en infraestructura, sin cambiar su tecnología, e ir migrando de manera simple entre las diferentes versiones de Standard Edition hasta llegar a Enterprise Edition.

Por tanto el empleo de **Oracle Standard Edition One** como Sistema Gestor de Bases de Datos, representa una buena estrategia si se desea alcanzar un elevado desempeño, confiabilidad, seguridad y altos niveles de compatibilidad características tradicionales en las herramientas Oracle. Además al usar Oracle Enterprise Edition en el Centro de Datos y presentar un Sistema de Bases de Datos Descentralizado, la comunicación será mucho mas eficiente entre el Centro de Datos y las Oficinas Locales, debido a la compatibilidad entre las versiones de Oracle, que solo difieren en algunas herramientas que son disponibles en una versión y no en la otra.

2.2.2 Diseño de Bases de Datos Locales para los Registros Públicos y Mercantiles.

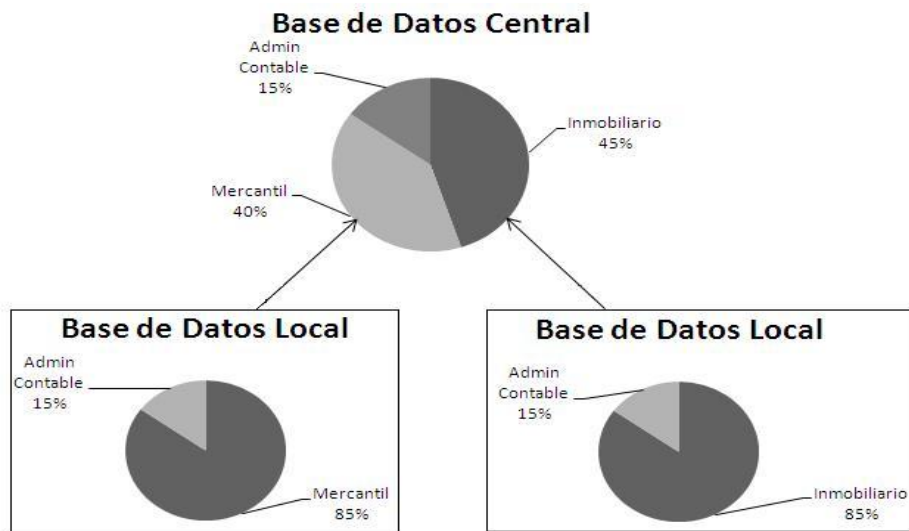
Las Bases de Datos Locales que funcionarán en cada registro del país han sido adaptadas según las necesidades impuestas por los procesos registrales, respetando en todo momento la relación coste-tiempo-beneficio con el objetivo de maximizar el rendimiento en el menor tiempo posible.

Las Bases de Datos Locales son estratégicamente claves en el intercambio de información con el Centro de Datos, permitiendo el envío de información hacia la Base de Datos Central utilizando mecanismos de replicación que no están soportados por la versión Oracle Database 10g Standard Edition One release 2, pero que son compatibles con la misma.

2.2.3 Diseño Físico de las Bases de Datos Locales.

Las Bases de Datos Locales han sido diseñadas estructuralmente idénticas a la Base de Datos Global, pero como un subconjunto de la Base de Datos Global.

Con este diseño se perseguía eliminar cualquier obstáculo estructural que imposibilitara el intercambio de datos entre cada Base de Datos Local y el Centro de Datos. En la figura se esquematiza el diseño implementado para cada Base de Datos Local y el Centro de Datos en general:



En cada base de datos local se asigna cada tabla a un tablespace de acuerdo a sus características, para esto se crearon 2 tablespaces con el objetivo de asignar a estos las tablas que mantienen una cierta relación en cuanto a carga de trabajo:

- Se creó un tablespace en el que se agrupan las tablas que mayor cantidad de operaciones reciben (insert, update, delete) y sobre las que las búsquedas serán mas exhaustivas, estas tablas son las llamadas críticas y pueden alcanzar un alto volumen de datos.
- Se creó un tablespace en el que se agrupan todas las tablas sobre las que se realiza un número menor de operaciones en comparación con las del grupo anterior, en este grupo se encuentran las llamadas nomencladores, sobre las que se realizan, en la mayoría de los casos, solo consultas.

También se creó un tablespace para almacenar todos los usuarios de la base de datos y otro para almacenar todos los índices creados, para el mejor funcionamiento de la misma.

Debido a que la versión de Oracle utilizada en las oficinas (Oracle Standard Edition One) no soporta la funcionalidad de partición de tablas, fue necesario dividir algunas tablas que alcanzan un volumen de datos y de acceso elevados, en varias tablas dependientes de estas para simular el particionamiento que brinda el gestor, esto se hizo con el fin de eliminar posibles bloqueos en la base de datos en un futuro cuando la misma alcance un elevado volumen de datos.

Esta estructura de organización de la base de datos permite que al realizar una operación sobre la misma, de acuerdo a la clasificación de las tablas afectadas, el manejador de la base de datos no tenga que acceder a las estructuras de almacenamiento en disco que tengan mucho volumen de datos y así se optimiza el tiempo de respuesta de la base de datos al usuario. También se consideró crear los tablespaces con la opción de auto incrementarse una vez alcanzaran el tamaño definido inicialmente, para que automáticamente reajustaran su tamaño y evitar pérdida de tiempo en el trabajo sobre las bases de datos.

2.3 Mecanismo de Sincronización entre Bases de Datos Locales y Centro de Datos.

El Sistema de Base de Datos Descentralizado sobre Oracle Database 10g que se propone, impone la necesidad de que toda la información que se gestiona a nivel de oficina se replique hacia la Base de Datos Central. Teniendo en cuenta que adquirir 255 licencias de Oracle Database 10g Enterprise Edition release 2 para cada oficina encarecen enormemente los costos, se decidió la compra de Oracle Database 10g Standard Edition One por usuarios. Esta versión de Oracle 10g diseñada para operar en pequeñas y medianas empresas no dispone de la tecnología Oracle Streams Replication, por lo que se tuvo que implementar un algoritmo basado en minería de Redo Logs.

A continuación se muestra un análisis producto vs precio de Oracle en el mercado mundial:

Oracle Product	Per Processor	Per User	Maximum Processor
Oracle 10g Standard Edition One	\$4,995	\$149 con mínimo de 5 Usuarios.	2
Oracle 10g Standard Edition	\$15,000	\$250	4
Oracle 10g Enterprise Edition	\$40,000	\$400	ilimitado

2.3.1 Oracle LogMiner.

Oracle LogMiner es una herramienta de gran alcance, diseñada para el análisis de los registros transaccionales que han sido registrados en una Base de Datos Oracle, permitiéndoles a los administradores localizar, analizar y restaurar los registros de datos modificados como consecuencia de errores lógicos de los usuarios[26].

LogMiner ofrece dos interfaces:

1. Una interfaz SQL que ejecuta un paquete PL/SQL incorporado donde se especifican los redo logs que serán analizados, el criterio de búsqueda y el diccionario de datos asociado.
2. Un visualizador gráfico (GUI) intuitivo que le permite al administrador configurar el diccionario de datos, especificar los ficheros de transacciones que se analizarán, definir los criterios de búsqueda y por último la visualización de los datos registrados en los ficheros de transacciones.

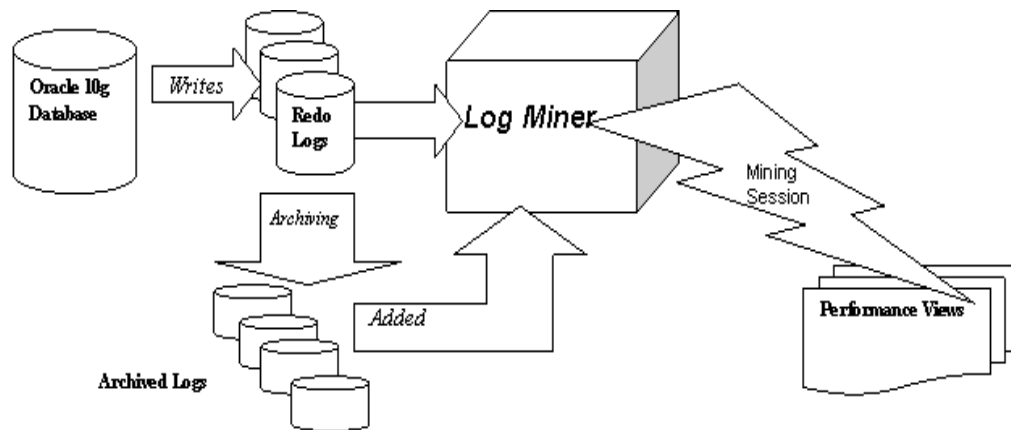
LogMiner puede ser utilizado para:

- Auditorías de sentencias DML en el orden en la cual las transacciones fueron commit y el usuario responsable de su ejecución.
- Análisis histórico del acceso a los datos para la planificación de la capacidad y ajustes en la configuración de la Base de Datos.
- Seguimiento de la evolución del esquema y el impacto en la estructura de los datos.

LogMiner tiene acceso directo a Oracle Redo Logs, que son expedientes completos de todas las actividades realizadas en la Base de Datos, y el diccionario de datos asociado que se utiliza para traducir identificadores y tipos internos de objeto a los nombres y formatos de datos externos. Usando la vista dinámica V\$LOGMNR_CONTENTS LogMiner publica una visión lógica de todas las operaciones realizadas en la Base de Datos, cada fila de la vista contiene un registro detallado de la operación que se efectuó en un determinado momento.

En el siguiente esquema se representa el funcionamiento de Oracle LogMiner:

Componentes del LogMiner.



Oracle Advanced Queueing.

En la versión Oracle 8i se introduce por primera vez el mecanismo de Queueing, con la incorporación de esta poderosa herramienta se logró intercambiar mensajes entre diversas aplicaciones. A esta utilidad se le llamó Advanced Queueing.

Oracle Advanced Queueing (Oracle AQ) es un sistema de encolado de mensajes integrado con la Base de Datos Oracle, que permite almacenar mensajes en cola para posteriormente ser recuperados y procesados por el servidor de Oracle[27].

Las aplicaciones pueden tener acceso a los mensajes en cola mediante el uso de interfaces PL/SQL, Java, C/C++ y Visual Basic, lo que posibilitó una gestión confiable y eficiente de los mensajes en cola, sin hacer uso de software adicionales tales como Transaction Processing (TP) Monitors o Message Oriented Middleware.

Algunas características de Oracle Advanced Queueing son:

- **SQL Access:** Los mensajes son almacenados como tuplas en la base de datos, lo que permite el uso de SQL Standard para acceder a las propiedades de los mensajes.
- **Retención:** Permite la retención de mensajes en cola después de ser procesados, lo que posibilita mantener un historial de mensajes.
- **Propagación de Mensajes:** Permite la propagación de mensajes entre una cola local y una cola remota, que posteriormente serán procesados en la base de datos remota.
- **Historial de Mensajes:** Oracle AQ almacena información de cada mensaje. La información contiene la fecha de ENQUEUE/DEQUEUE y el identificador de transacción que se originó de cada petición.
- **Propagación de Estadísticas:** Oracle AQ permite generar estadísticas en cada evento de propagación. Esta información es utilizada para diagnosticar el rendimiento en cada evento de propagación de mensajes.
- **Múltiples Consumidores de Mensajes:** Un mensaje en cola puede ser procesado por múltiples consumidores.
- **Prioridad de los Mensajes:** Oracle AQ establece 3 posibilidades para que los consumidores puedan acceder a los mensajes:
 1. Definir una propiedad a través de la cual los mensajes se organicen en cola.
 2. Definir una prioridad para cada mensaje.
 3. Asignar una secuencia de desviación para cada mensaje, que permita identificar la posición de un mensaje en relación a otro.

2.4 Seguridad en las Bases de Datos Locales.

En esta sección se mencionan ciertos aspectos que han sido tomados en cuenta cuando se trata de seguridad de la información en Bases de Datos Oracle 10g release 2.

Oracle 10g ofrece capacidades mediante las cuales protege los datos sensibles y críticos de la empresa contra usuarios no autorizados. Características como la encriptación a 128 bits de datos permiten esta protección de los datos de una forma muy segura y confiable[28].

El control de acceso y la autenticación es el mecanismo que asegura que los clientes y los servidores sean reconocidos y aceptados entre ellos según el tipo de transacción requerida.

El primer nivel de protección en un sistema se realiza a través de un código de usuario y una palabra clave. Si la palabra clave del usuario es interceptada, el acceso a toda la información y privilegios del usuario son obtenidos. Si un usuario debe acceder a múltiples servidores y bases de datos se requiere de varias palabras clave. Esta situación genera esquemas de identificación de palabras claves muy simples e inclusive lleva a la escritura en papel de dichas claves.

Otro reto de seguridad es garantizar la autenticación entre servicios, como pueden ser entre dos bases de datos que requieren completar una transacción en común. Idealmente la autenticación en las redes debería permitir y garantizar el envío de las credenciales del usuario.

Técnicas que aseguran la autenticación y control del acceso por parte de los usuarios a la información:

- Es posible utilizar dispositivos que utilizan “testigos” o tokens.
- La utilización de dispositivos biométricos y tecnología RADIUS que garanticen la autenticidad y seguridad.
- Utilización de protocolos de seguridad basados en Infraestructura de Clave Publicas (PK) sobre SSL (Secure Socket Layer).

Oracle provee una solución completa que permite la utilización de los distintos mecanismos de seguridad para la autenticación y control de acceso a la información que va desde el cifrado de los datos en el servidor, pasando por el cifrado de la información que viaja por las redes como también la utilización de códigos de usuario y palabras claves en forma directa o a través de directorios de información que utilizan protocolos LDAP para garantizar su seguridad.

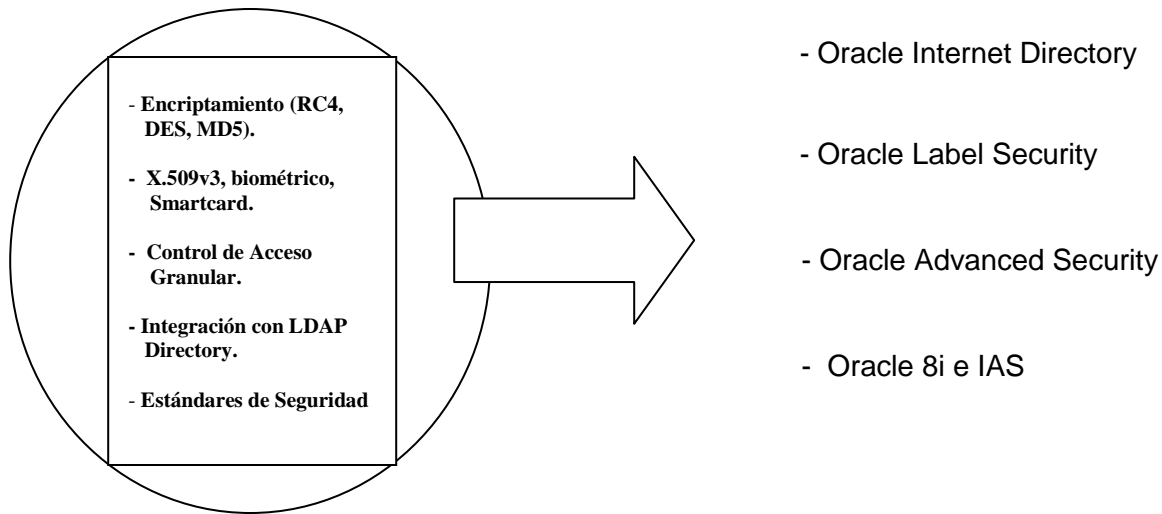
2.4.1 Estándares de Seguridad.

A continuación se especifican los requerimientos estándar de seguridad de la información y las soluciones del mercado disponibles en cada caso.

Requerimientos de Seguridad

Requerimiento de Seguridad	Mecanismo de Seguridad
Privacidad e Integridad en las comunicaciones	Encriptamiento (RC4, DES, MD5).
Autenticación de usuarios	Certificados X.509v3, Smartcards.
Flexibilidad y bajo Costo	Estándares de Seguridad (FIPS 140, Common Criteria).
Control de Acceso	Políticas granulares de Control de Acceso.
Administración de usuario	Integración con LDAP Directory.

Para cada uno de estos requerimientos y según los estándares del mercado Oracle ha adaptado sus productos para dar soporte a estas normas, mecanismo y políticas. Así mismo, Oracle ha recibido la certificación correspondiente como proveedor calificado de productos que cumplen con estas normas de seguridad. En el siguiente cuadro se puede apreciar la propuesta de Oracle para cada uno de los puntos anteriores.



2.4.2 Principios de Seguridad Aplicados en las Bases de Datos Locales.

En esta sección abordaremos detalladamente los Principios de Seguridad expuestos anteriormente.

Algunos de los principios base de seguridad implementados en las Bases de Datos Locales son:

- Disgregación de Responsabilidades.
- Mínimo de Privilegios.
- Estándares de Seguridad de Contraseñas.
- Monitoreo de Actividades Suspicious.

Disgregación de Responsabilidades.

El Principio Disgregación de Responsabilidades sustenta sus bases en los siguientes planteamientos:

- El usuario con privilegio DBA de la Base de Datos debe ser de total confianza, sobre este usuario recaerá la responsabilidad de custodiar la integridad de la Base de Datos.
- El usuario con privilegio DBA y el administrador de sistema no debe ser la misma persona.
- Disgregar las responsabilidades del usuario operador y el DBA de la Base de Datos.

- Las cuentas de usuario en la Base de Datos no debe ser compartidas.

Mínimo de Privilegios.

El Principio de Menos Privilegios plantea lo siguiente:

- Instalación en los servidores el software requerido. El paquete Oracle Database CD contiene un grupo de opciones adicionales que pueden ser instaladas solo si son requeridas. Se recomienda instalar solo aquellos productos y opciones necesarias para el usuario.
- Habilitar solo los servicios requeridos.
- Limitar el acceso de los usuarios a través de la cuenta root o administrador.
- Limitar el acceso de los usuarios con privilegios sysdba y sysoper.
- Autorizar el acceso al Sistema Operativo y a la Base de Datos solo a aquellos usuarios que lo requieran.
- Proteger el diccionario de datos de la Base de Datos, cambiando el valor del parámetro `07_DICTIONARY_ACCESSIBILITY = FALSE`.
- Revocar los privilegios innecesarios del usuario de la Base de Datos Public ejecutando las siguientes sentencias :
- `revoke execute on utl_smtp, utl_tcp, utl_http, utl_file from public.`
- Restringir la autenticación remota a la Base de Datos, configurando el siguiente parámetro `REMOTE_OS_AUTHENT = FALSE`.

Estándares de Seguridad de Contraseñas.

El Principio Estándares de Seguridad de Contraseñas establece:

- Implementar una función que verifique que las contraseñas suministradas por los usuarios corresponde al estándar establecido por la organización.
- Chequear que el mínimo número de caracteres de la contraseña de cada usuario sea mayor a cuatro letras.
- Chequear que la contraseña suministrada por cada usuario no sea idéntico al nombre del usuario.

- Verificar que la contraseña contenga al menos una letra del alfabeto, un número y un carácter especial.
- Comprobar que la contraseña actual del usuario difiere de los anteriores en al menos tres letras.

2.4.2.1 Acceso Restringido desde la Red.

Usar Firewall.

Mantener el servidor de Bases de Datos detrás de un firewall. La infraestructura de red de Oracle, Oracle Net (conocido formalmente como Net8 y SQL*Net), ofrece soporte para una variedad de firewalls. Los firewalls reducen la exposición de asuntos potenciales de seguridad como: apertura de puertos a través del firewall, servidores de sistemas operativos multihilos y la revelación de información crucial en bases de datos que están protegidas por firewalls[29].

Proteger el Listener de Oracle.

Debido a que el listener funciona como puerta de enlace entre el servidor Oracle y la red, es importante limitar las consecuencias de interferencias malignas[29]:

- Restringir los privilegios del listener para que no pueda leer o escribir archivos en la base de datos.

Esta restricción previene que agentes manejadores del listener puedan acceder a dichas lecturas o escrituras. El dueño o usuario del listener no debe ser el usado para instalar Oracle o para manejar la instancia.

Configuración de ejemplo:

```
extproc_listener =
    (description =
        (address =
            (protocol = ipc)(key=extproc)))
    sid_list_extproc_listener =
        (sid_list =
            (sid-desc =
                (sid_name = plsextproc)
            (oracle_home = /u1/app/oracle/product/10.2.0/db_1)
        (program = extproc)))
```

- Garantizar la seguridad en la administración de la Base de Datos haciendo lo siguiente:
 1. Prevenir que el administrador tenga que escribir los privilegios de permiso en el fichero listener.ora y en el de contraseña del listener a la hora de usar administración online:
Para esto adicione o altere esta línea en el fichero listener.ora y luego de hacer esto, volver a cargar la configuración.
admin_restrictions_listener = on
 2. Usar encriptación SSL cuando se administra el listener, haciendo que el protocolo TCPS sea la primera entrada en la dirección:
listener =
 (description =
 (address_list =
 (address =
 (protocol = tcps)
 (host = ed-pdsun1.us.oracle.com) (port = 8281)
))

Para administrar el listener remotamente es necesario definir el listener en el fichero listener.ora en la computadora cliente. Por ejemplo, para acceder al listener USER218 remotamente, usar la siguiente configuración:

```
user281 =  
    (description =  
        (address = (protocol = tcps)  
            (host = ed-pdsun1.us.oracle.com)  
            (port = 8281))  
        )  
    )
```

3. Para versiones anteriores al Oracle Database 10g release 2, es muy importante crear una contraseñas para el TNS listener de Oracle para prevenir la configuración remota del mismo.

A continuación se muestra un ejemplo de cómo hacerlo:

```
lsnrctl> change_password  
Old password: ls8nrc80a  
New password: ls8nrc90a  
Reenter new password: ls8nrc90a  
lsnrctl> set password  
Password:  
The command completed successfully  
lsnrctl> save_config  
The command completed successfully
```

Para Oracle Database 10g release 1 y versiones superiores la autenticación por defecto es a través del sistema operativo, lo que requiere que el administrador sea miembro del grupo local dba group. Establecer una contraseña para el TNS listener en Oracle Database 10g release 1 y versiones superiores simplifica la administración del mismo. Sin embargo esto requiere una buena política de manejo de contraseña para prevenir que usuarios no autorizados puedan adivinar la contraseña y obtengan acceso a operaciones del listener. Se recomienda no establecer contraseñas para el TNS listener en Oracle Database 10g release 1 y versiones superiores.

4. Remover la configuración de procedimientos externos del archivo listener.ora si no se tiene intención de usar dichos procedimientos.
5. Monitorear la actividad del listener.

Monitoreo de Acceso al Sistema.

La autenticación de computadoras clientes a través de internet puede ser problemático. Es mejor usar autenticación de usuario pues evita problemas con direcciones IP falsificadas, aplicaciones o sistemas operativos hackeados y falsificaciones o robos de las identidades de sistemas clientes[29].

Los siguientes pasos mejoran la seguridad de las computadoras clientes:

- Configurar la conexión para usar encriptación SSL (Secure Sockets Layer). De esta forma se habilita los certificados para la autenticación de usuarios y servidores.
- Considerar la creación de certificados de autenticación de usuarios y sistemas tales como:
 - i. Pruebas de aplicación para certificados expirados.
 - ii. Auditorías a listas de revocación de certificados.

Chequeo de direcciones IP.

Usar la herramienta de seguridad de Oracle (Oracle Net valid node checking) para permitir o denegar acceso a los procesos del servidor de Oracle a los clientes de red con direcciones IP específicas[29]. Para usar esta herramienta establezca los siguientes parámetros en el archivo de configuración de red de Oracle protocol.ora:

```
tcp.validnode_checking = yes
```

```
tcp.excluded_nodes = {list of IP addresses}
```

```
tcp.invited_nodes = {list of IP addresses}
```

El primer parámetro es el uso de la herramienta, mientras que los últimos deniegan y permiten respectivamente la conexión a direcciones IP específicas de clientes. Esto ayuda en la prevención potencial de ataques de Denegación de Servicio.

Cifrar el tráfico en la red.

De ser posible, se recomienda usar la Seguridad Avanzada de Oracle para cifrar el tráfico en la red entre clientes, bases de datos y servidores de aplicaciones.

2.4.2.2 Monitoreo de Actividades Suspicious.

El Principio de Monitoreo o Auditoría se basa en la posibilidad de efectuar auditorías automáticas a la Base de Datos Oracle y los distintos tipos de actividades que se realizan en el sistema que son sujetos de auditar, entre ellas[30]:

- Modificaciones al modelos de datos (comandos DDL).
- Modificación de los datos (comandos DML).
- Conexiones a la base de datos.

Las actividades a auditar se configuran de manera flexible y pueden modificarse fácilmente. La generación de pistas de auditoría permite mantener un historial de los cambios que se realizan a los datos e identificar: qué cambió, quién lo realizó y cuándo. A partir del análisis de esta información se pueden llegar a determinar cómo cualquier dato o elemento obtuvo su valor actual.

Las actividades pueden auditarse desde 2 puntos de vistas:

1. Quién las realizó: aquellos comandos ejecutados por cierto usuario. Ej.: conexiones a las bases de datos.
2. Qué objeto afectó: comandos realizados sobre cierto objeto. Ej.: actualizaciones realizadas sobre domicilios.

En una aplicación hay distintas actividades que pueden auditarse:

1. Modificación a la estructura en la que residen los datos (modelo de datos).
2. Modificación a los datos.
3. Conexiones a la base de datos.

Una vez seleccionados los distintos componentes a ser auditados, estos se deben especificar en el módulo de auditoría el cual creará en el repositorio de auditorías los objetos necesarios para registrar la información solicitada.

Los pasos para habilitar la auditoría son:

- 1- Identificar los objetos u estructuras de datos que deben ser auditadas con el fin de satisfacer los requerimientos de información necesarios para su posterior análisis.
- 2- Determinar qué elementos y acciones sobre dichos elementos se desean registrar.
- 3- Utilizar el módulo de auditoría para registrar la información definida en los puntos anteriores y de esta manera habilitar la auditoría sobre dichos objetos.
- 4- Verificar que las pistas de auditoría definidas sirvan para el análisis que se desea realizar.
- 5- Repetir los pasos anteriores hasta lograr que la información generada en el repositorio de auditoría sea la deseada.

El análisis de la información recolectada se realiza mediante el acceso al repositorio de auditorías con herramientas de consulta. Dicho análisis puede realizarse a partir de consultas pre-definidas o realizadas.

2.5 Respaldo y Recuperación en las Bases de Datos Locales.

La seguridad no solo depende de establecer un correcto control de acceso de los usuarios a la Base de Datos. Debemos también protegernos de los posibles errores producidos en la información que da soporte a todo el sistema (meta-información) y la información añadida al repositorio. Oracle dispone de distintos mecanismos de protección para situaciones de caída del sistema, fallos del disco duro. La pérdida de un archivo de la Base de Datos, corrupción de un fichero. Para poder recuperarnos de estos fallos, debemos guardar una copia de seguridad de los datos contenidos en el sistema y un registro de transacciones para ver las ejecuciones que se van realizando por parte de los usuarios.

Planear y comprobar los procedimientos de backup del sistema es la única garantía que existe contra fallos del sistema, del sistema operativo, del software o cualquier otro tipo de circunstancias[31].

Las posibles causas de falla en un Sistema de Bases de Datos pueden agruparse en las siguientes categorías:

- **Físicas:** Son causadas por fallos del hardware, como por ejemplo del disco o de la CPU.
- **Diseño:** Son agujeros en el software, ya sea en el Sistema Operativo o en el Sistema Gestor de Bases de Datos.
- **Funcionamiento:** Son causadas por la intervención humana, debidos a fallos del DBA, configuraciones inapropiadas o mal planteamiento de los procedimientos de backup.
- **Entorno:** Son los causados por desastres naturales, fallos de corriente, temperatura excesiva y otros.

De todas estas posibilidades sólo se puede prevenir los errores de funcionamiento, ya que el resto habitualmente no está dentro de las responsabilidades y capacidades del dba.

Dada la complejidad del sistema y la necesidad cada vez más críticas en la disponibilidad de los sistemas, donde una Base de Datos caída puede causar pérdidas millonarias, es necesario considerar los mecanismos de protección hardware y de redundancia que la tecnología nos proporciona:

- UPS o fuentes de corriente ininterrumpida.
- Discos en Espejo, o tecnología RAID.
- Componentes duplicados.
- Sistemas redundantes.

Una decisión importante es decidir si arrancar la Base de Datos en modo ARCHIVELOG o no. El modo ARCHIVELOG es la forma en que se denomina a una Base de Datos que va a ir almacenando permanentemente en archivos todos los cambios que se van produciendo en su información.

Esta decisión tiene sus ventajas e inconvenientes:

Ventajas:

- Aunque se pierdan los ficheros de datos, siempre se puede recuperar la Base de Datos con una copia antigua de los ficheros de datos y los ficheros de Redo Log archivados.
- Es posible realizar backups en caliente, ya que es un modo de almacenamiento de información de backup online, el archivado de los Redo Log se va produciendo con la base de datos levantada. Esto implica que los usuarios no dejarán de trabajar en ningún momento con la información de la base de datos.

Inconvenientes:

- Se necesitará más espacio en disco.
- El trabajo del Administrador de la Base de Datos se incrementa al tener que determinar el destino del archivado de los Redo Log.

Luego de demostrar que la ejecución de la Base de Datos en modo ARCHIVELOG permite:

- Recuperaciones completas hasta un instante antes de la caída de la base de datos.
- Backups Online.
- Los tablespaces pueden ser puestos offline de forma inmediata.

Se decidió configurar las Bases de Datos Locales en modo ARCHIVELOG para disminuir en caso de desastre los tiempos de recuperación de la Base de Datos.

2.5.1 Tipos de Backups.

Los backups se pueden clasificar en físicos y lógicos.

- **Físicos:** los físicos se realizan cuando se copian los ficheros que soportan la Base de Datos. Entre estos se encuentran los backups del Sistema Operativo, los backups en frío y los backups en caliente[31].
- **Lógicos:** los backups lógicos sólo extraen los datos de las tablas utilizando comandos SQL y se realizan con la utilidad export/import[31].

2.5.1.1 Backups de la Base de Datos en Frío.

Los Backups en Frío implican parar la Base de Datos en modo normal y copiar todos los ficheros sobre los que se asienta[31]. Antes de parar la Base de Datos hay que parar también todas las aplicaciones que estén trabajando con la Base de Datos. Una vez realizada la copia de los ficheros, la Base de Datos se puede volver a arrancar.

2.5.1.2 Backups de la Base de Datos en Caliente.

El Backup en Caliente se realiza mientras la Base de Datos este abierta y funcionando en modo ARCHIVELOG[31]. Este tipo de backup consiste en copiar todos los ficheros correspondientes a un tablespace determinado, los ficheros Redo Log archivados y los ficheros de control.

2.5.1.3 Backups del Sistema Operativo.

Este tipo de backup es el más sencillo de ejecutar, aunque consume mucho tiempo y hace inaccesible al sistema mientras se lleva a cabo[31]. Aprovecha el backup del Sistema Operativo para almacenar también todos los ficheros de la Base de Datos. Los pasos de este tipo de backup son los siguientes:

- 1- Parar la Base de Datos y el Sistema Operativo.
- 2- Comenzar en modo de usuario con privilegios de administración.
- 3- Realizar copia de todos los ficheros del sistema de ficheros.

4- Arrancar el sistema en modo normal y luego la Base de Datos.

2.5.1.4 Backups Lógicos con Export/Import.

Estas utilidades permiten hacer copias de determinados objetos de la Base de Datos, así como restaurarlos o moverlos de una Base de Datos a otra[31]. Estas herramientas utilizan comandos SQL para obtener el contenido de los objetos y escribirlos en ficheros.

2.5.2 Estrategias de Backups. Reglas Básicas.

Es muy importante entender ciertas reglas que determinan la situación de los ficheros y otras consideraciones que afectarán al esquema de backup.

- Es recomendable archivar los ficheros Redo Log en disco, y luego copiarlos a cinta, pero siempre en un disco diferente del que soporta los ficheros de datos y de Redo Log activos.
- Los ficheros copias no deben estar en el mismo dispositivo que los originales. No siempre hay que pasar las copias a cinta, ya que si se dejan en disco se acelera la recuperación. Se deberían mantener diferentes copias de los ficheros de control, colocadas en diferentes discos con diferentes controladores.
- Los ficheros Redo Log en línea deben estar multiplexados, con un mínimo de dos miembros por grupo, residiendo cada miembro en un disco distinto.
- La utilización de un área de recuperación flash (Flash Recovery Area) que permite automatizar la gestión de copias de seguridad de la mayoría de los ficheros. En esta área Oracle se encarga automáticamente de almacenar los ficheros, y de borrarlos cuando ya no son necesarios.
- Siempre que la estructura de la Base de Datos cambie debido a la inclusión, borrado o renombrado de un fichero de datos o de Redo Log, se debe copiar el fichero de control, ya que almacenan la estructura de la Base de Datos. El fichero de control puede ser copiado mientras la Base de Datos está abierta con el siguiente comando:

```
SQL> alter database backup controlfile to <destino>;
```

Teniendo en cuenta las reglas anteriores, se trazó la siguiente estrategia de backup en las Bases de Datos Locales:

1. Tipo de Backup: Backup Total de la Base de Datos.
2. Formato de Backup: Compressed Backup Set
3. Activar el modo ARCHIVELOG.
4. Habilitar Flash Recovery Area.
5. Multiplexar los Archivos de Control, Online Redo Logs, y Archived Redo Logs.
6. Respaldo automático de los Archivos de Control (controlfile) y Parámetros de la Base de Datos (spfile).
7. Configuración de políticas de retención de backup. point-in-time-to-recovery = 7
8. Respaldo todos los ficheros de Redo Log archivados.
9. Configuración de estrategias de backup incrementales diario según el orden siguiente:

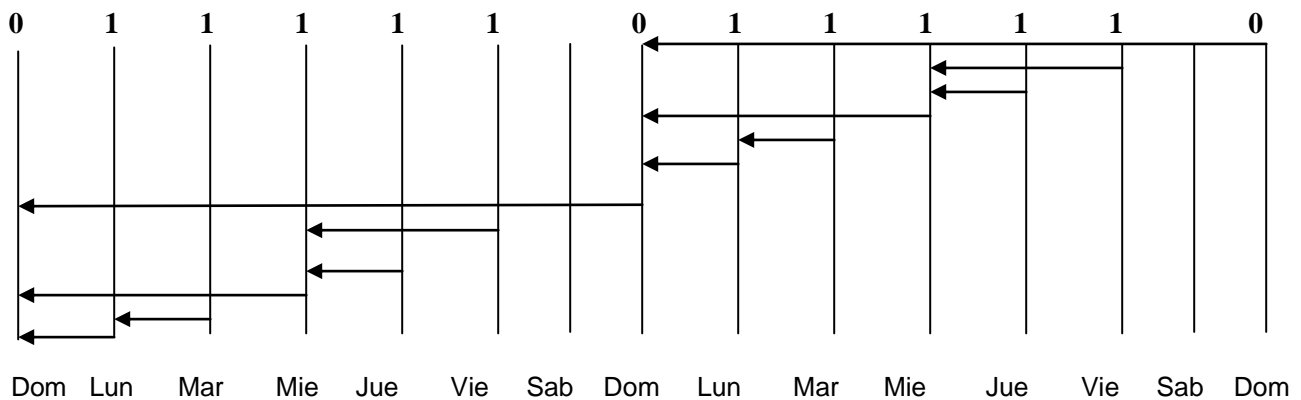
Días de la Semana	Tipo de Backup	Descripción
Sábado.	Backup Incremental Level 0.	Este backup es el punto de partida para los subsiguientes backup incrementales diferencial level. Job Name = backup_dif_incr_level_1_lunes Job Description = incremental diferencial lunes Start -> Later = lunes 08:00 pm Repeat -> Interval -> Frequency = 1 weeks

Lunes.	Backup Incremental Diferencial Level 1.	Backup Incremental Diferencial Level 1. Este backup respaldará todos los bloques de datos que han sido modificados desde el backup incremental level 0 o level 1 más reciente. Job Name = backup_dif_incr_level_1_lune Job Description = incremental diferencial lunes Start -> Later = lunes 08:00 pm Repeat -> Interval -> Frequency = 1 weeks
Martes.	Backup Incremental Diferencial Level 1.	Job Name = backup_dif_incr_level_1_martes Job Description = incremental diferencial martes Start -> Later = martes 08:00 pm Repeat -> Interval -> Frequency = 1 weeks
Miércoles	Backup Incremental Acumulativo Level 1.	Backup Incremental Acumulativo Level 1. Este backup respaldará todos los bloques de datos que han sido modificados desde el Backup Incremental Level 0 más reciente. Job Name = backup_cum_incr_level_1_mier Job Description = incremental acumulativo miércoles Start -> Later = miércoles 08:00 pm Repeat -> Interval -> Frequency = 1 weeks
Jueves.	Backup Incremental Diferencial Level 1.	Job Name = backup_dif_incr_level_1_jueves Job Description = incremental diferencial jueves Start -> Later = jueves 08:00 pm Repeat -> Interval -> Frequency = 1 weeks

Viernes.	Backup Incremental Acumulativo Level 1.	Job Name = backup_cum_incr_level_1_vier Job Description = incremental diferencial viernes Start -> Later = viernes 08:00 Repeat -> Interval -> Frequency = 1 weeks
----------	---	---

En el siguiente gráfico se ilustra la estrategia de respaldo y recuperación que se implementó en las Bases de Datos Locales.

Backup Level



2.5.3 Estrategias de Recuperación. Recovery Manager Oracle.

Oracle Recovery Manager (RMAN) es una herramienta de alta disponibilidad, que permite la planificación y ejecución de complejas estrategias de respaldo y recuperación de Bases de Datos Oracle. Está diseñada para soportar todos los formatos de datos de Oracle. En la versión de Oracle Database 10g fueron incorporadas nuevas funcionalidades convirtiéndola en una potente herramienta de gestión de recuperación simplificando procesos de backup, restauración y recuperación de datos[32].

2.5.3.1 Componentes Funcionales de Recovery Manager (RMAN).

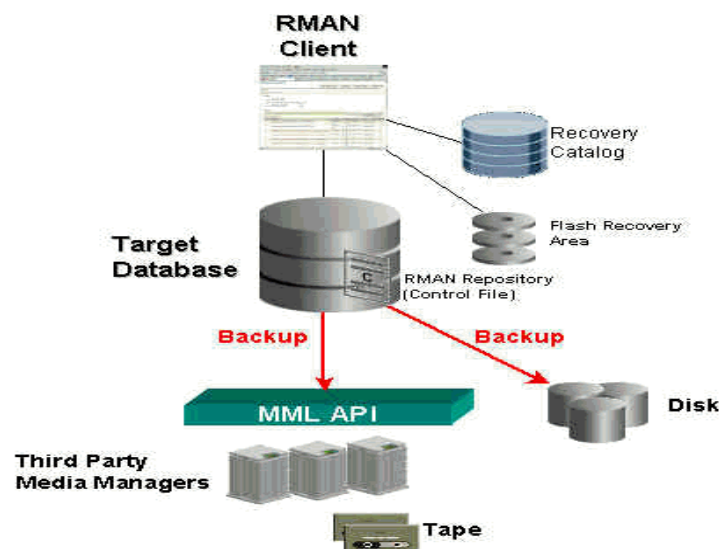
El entorno de RMAN consiste en las utilidades y la Base de Datos que será objeto de backup de los datos. Como mínimo el entorno de RMAN incluye los siguientes componentes:

- La Base de Datos objeto de backup de los datos.
- El cliente de RMAN encargado de interpretar los comandos de respaldo y recuperación, además de la creación en el servidor de procesos usuarios que ejecutarán los comandos que serán registrados en los ficheros de control de la Base de Datos (controlfile).

Algunos entornos tienden a utilizar componentes adicionales soportados por RMAN tales como:

- Flash Recovery Area o Área de Recuperación: Es una localización en disco utilizada por la Base de Datos para el manejo y almacenamiento de los archivos utilizados en los procesos de respaldo y recuperación de datos.
- Recovery Catalog Database: Es un esquema de Bases de Datos separado donde se registra la actividad de RMAN de una o varias Bases de Datos.

En la figura se muestra los componentes soportados por RMAN:



Características Generales de RMAN.

1. Permite respaldar y recuperar los datos exactamente, según su estado inicial.
2. Detención y reparación de los bloques de datos automáticamente.
3. Permite la generación de reportes. Un usuario administrador de la Base de Datos puede mediante el uso de vistas dinámicas, consultar toda la información relativa a los trabajos de respaldo y recuperación que se están ejecutando actualmente o culminaron en un momento dado.
4. Permite la supervisión inteligente de espacio en disco disponible. El administrador de la Base de Datos puede a través del uso de notificaciones enviar alertas conteniendo el poco espacio disponible en el Flash Recovery Area. Estas notificaciones pueden ser enviadas vía SMTP
5. Administración simple y centralizada. El administrador de la Base de Datos pueden interactuar con la herramienta a través de una interfaz de comandos o mediante la consola Web Enterprise Manager.

Existen varios métodos de recuperación, pero todos ellos se basan en la aplicación de los registros de Redo Log.

2.5.3.2 Métodos de Recuperación.

Aplicación de Redo Log.

Cuando una Base de Datos se arranca con el comando startup, la Base de Datos pasa por los estados nomount, mount y open. En este tercer estado, se verifica que se pueden abrir todos los ficheros de log y de datos. Si la Base de Datos se arranca por primera vez después de una caída, se necesitará efectuar una recuperación que consiste en dos pasos: avanzar la Base de Datos hacia adelante aplicando los registros redo log, deshacer las transacciones no confirmadas.

Si después de parar la Base de Datos se reemplaza un fichero de datos por su copia de seguridad, al arrancar la Base de Datos Oracle detecta que el contador de checkpoints del fichero de datos no coincide con el almacenado en el fichero de control. Así, se tendrá que echar mano a los ficheros redo log archivados, empezando por aquel cuyo número de secuencia aparece en la cabecera del fichero de datos.

Recuperación Física.

La utilización de una copia de backup de ficheros de datos siempre necesita de una recuperación física. También es así cuando un fichero de datos se pone offline sin un checkpoint.

Oracle detecta que se necesita una recuperación física cuando el contador de checkpoints de la cabecera del fichero de datos no coincide con el correspondiente contador de checkpoints del fichero de control. Entonces se hace necesario el comando recover. La recuperación comienza en el SCN menor de los ficheros de datos en recuperación, aplicando los registros de redo log a partir de él, y parando en el SCN de final mayor de todos los ficheros de datos.

Existen tres opciones para realizar una recuperación física:

1. Recuperación de Bases de Datos donde se restaura la Base de Datos entera.
2. Recuperación de tablespace donde, mientras una parte de la Base de Datos está abierta, se puede recuperar un tablespace determinado.
3. Recuperación de un fichero de datos específico mientras el resto de la Base de Datos está abierta.

Requisitos para Utilizar Recuperación Física.

La primera condición que se ha de poner para poder recuperar físicamente una Base de Datos es que ésta se esté utilizando en modo ARCHIVELOG. De otro modo, una recuperación completa puede que no sea posible.

Recuperación de la Base de Datos.

La Base de Datos debe estar montada pero no abierta. El comando de recuperación es el siguiente:

```
RECOVER [AUTOMATIC] [FROM 'localizacion'] [BD]
  [UNTIL CANCEL]
  [UNTIL TIME fecha]
  [UNTIL CHANGE entero]
[USING BACKUP CONTROLFILE]
```

Las opciones entre corchetes son opcionales:

- **AUTOMATIC** hace que la recuperación se haga automáticamente sin preguntar al DBA por el nombre de los ficheros redo log. También se puede utilizar para este cometido el comando `set autorecovery on/off`. Los ficheros redo log deben estar en la localización fijada en `LOG_ARCHIVE_DEST` y el formato del nombre de los ficheros debe ser el fijado en `LOG_ARCHIVE_FORMAT`.
- **FROM** se utiliza para determinar el lugar donde están los ficheros redo log, si es distinto del fijado en `LOG_ARCHIVE_DEST`.
- **UNTIL** sirve para indicar que se desea realizar una recuperación incompleta, lo que implica perder datos. Solo se dará cuando se han perdido redo logs archivados o el fichero de control. Cuando se ha realizado una recuperación incompleta la BD debe ser abierta con el comando `alter database open resetlogs`, lo que produce que los redo logs no aplicados no se apliquen nunca y se inicialice la secuencia de redo log en el fichero de control. Existen tres opciones para parar la recuperación:
 - **UNTIL CANCEL** permite recuperar un redo log cada vez, parando cuando se teclea `CANCEL`.
 - **UNTIL TIME** permite recuperar hasta un instante dado dentro de un fichero de redo log
 - **UNTIL CHANGE** permite recuperar hasta un SCN dado.
 - **USING BACKUP CONTROLFILE** utiliza una copia de seguridad del fichero de control para gobernar la recuperación.

Recuperación de un Tablespace.

La Base de Datos debe estar abierta, pero con el tablespace a recuperar offline. El comando de recuperación es el siguiente:

```
RECOVER [AUTOMATIC] [FROM 'localizacion'] TABLESPACE nombre_tablespace [, nombre_tablespace]
```

Recuperación de un Fichero de Datos.

La Base de Datos debe estar abierta o cerrada, dependiendo del fichero a recuperar. Si el fichero a recuperar es de un tablespace de usuario, la Base de Datos puede estar abierta, pero con el fichero a

recuperar offline. Si el fichero es del tablespace SYSTEM, la Base de Datos debe estar cerrada, ya que no puede estar abierta con los ficheros del SYSTEM offline. El comando de recuperación es el siguiente:

```
RECOVER [AUTOMATIC] [FROM 'localizacion'] DATAFILE nombre_fichero [, nombre_fichero]
```

2.6 Arquitectura de la Base de Datos Central.

Con el objetivo de centralizar la información y establecer controles de auditoría que puedan servir de base para futuras tomas de decisiones, se implementó una solución administrativa centralizada que permita mantener un monitoreo global de cada subsistema. Para ello fue necesario el diseño e implementación de un Centro de Datos a través del cual se gestione la información de todos los Registros y Notarías de la República Bolivariana de Venezuela.

2.6.1 Diseño del Centro de Datos.

El diseño exitoso de un Centro de Datos es un reto serio, especialmente cuando el desempeño del Uptime (tiempo de funcionamiento) requiere tolerancia a fallos con capacidad de mantenimiento simultáneo, escalabilidad palpable y la flexibilidad para lograr una larga vida. El Centro de Datos está diseñado apropiadamente para proporcionar disponibilidad, accesibilidad, escalabilidad, y confiabilidad las 24 horas al día, los 7 días a la semana, y los 365 días al año descontando el tiempo fuera de servicio por mantenimiento.

El Centro de Datos contiene los siguientes componentes:

- Infraestructura de cómputo y redes (cableado, fibra, y electrónicos).
- Sistemas eléctricos de distribución, generación y acondicionamiento - UPS, generadores de control ambiental.
- Sistemas de detección y supresión de fuego.
- Seguridad física y prevención de control de acceso, permisos y logging.
- Racks y gabinetes para equipos.
- Equipo de Telecomunicaciones
- Separaciones alrededor del equipo, y terminaciones en paneles y racks.

Es importante predecir el número de usuarios, tipos de aplicaciones y plataformas, unidades de rack requeridas para el montaje de equipos y sobre todo, anticipar el crecimiento y los cambios tecnológicos. Para realizar un estimado del Rendimiento vs. Capacidad del Centro de Datos se hizo un análisis de planeamiento de rendimiento y capacidad.

Planeamiento de Rendimiento

Transacciones de Escritura de Trámites al Centro de Datos (TET).

En cada uno de los doscientos cincuenta y cinco Registros Público y Mercantiles del país se culminan 100 trámites como promedio. Un trámite se considera un transacción completa cuando este haya alcanzado el estado de otorgado o terminado, entonces se envía a la Base de Datos Central.

Por tanto:

100 Trámites en 8 horas.

12.5 TET/h

0.21 TET/min

Transacciones de Lectura al Centro de Datos (TLC).

En cada paso del flujo la aplicación necesita consultar información de la Base de Datos Central, a esta situación súmesele que varios trámites pueden realizarse simultáneamente. En un flujo completo de la aplicación se realizan 20 consultas como promedio por cada trámite.

Por tanto:

100 Trámites x 20 TLC = 2000 TLC en 8 horas

250 TLC/h

4.2 TLC/min

Transacciones de Escritura de Imágenes Overnight (TEO).

Por cada trámite se genera un documento digitalizado que puede contener 10 folios como promedio. Estas transacciones se consideran grandes (long running transactions).

Por tanto:

100 documentos overnight (12 horas)

8.3 TEO/h

0.1 TEO/min

Planeamiento de Capacidad

Cada trámite aproximadamente tiene:

30 Kb en data (según modelo físico de la base de datos)

10 folios * 100 Kb (tamaño de la imagen) = 1000 Kb

Cálculo para un año en un Registro Público o Mercantil

100 trámites * 30 Kb * 22 días * 12 meses = 792000 Kb

= 773.4 Mb

= 0.76 Gb

100 trámites * 1000 Kb * 22 días * 12 meses = 26400000 Kb

= 25781.3 Mb

= 25.2 Gb

Total = 25.9 Gb

Cálculo para un año en los 255 Registros Públicos y Mercantiles

100 trámites * 30 Kb * 22 días * 255 * 12 meses = 201960000 Kb

= 197226.6 Mb

= 192.6 Gb

= 0.2 Tb

$$\begin{aligned} 100 \text{ trámites} * 1000 \text{ Kb} * 22 \text{ días} * 255 * 12 \text{ meses} &= 6732000000 \text{ Kb} \\ &= 6574218.8 \text{ Mb} \\ &= 6420.1 \text{ Gb} \\ &= 6.3 \text{ Tb} \end{aligned}$$

Total = 6.5 Tb

Cálculo para 10 años en los 255 Registros Públicos y Mercantiles

$$\begin{aligned} 100 \text{ trámites} * 30 \text{ Kb} * 22 \text{ días} * 255 * 12 \text{ meses} * 10 &= 2019600000 \text{ Kb} \\ &= 1972265.6 \text{ Mb} \\ &= 1926 \text{ Gb} \\ &= 1.9 \text{ Tb} \end{aligned}$$

$$\begin{aligned} 100 \text{ trámites} * 1000 \text{ Kb} * 22 \text{ días} * 255 * 12 \text{ meses} * 10 &= 67320000000 \text{ Kb} \\ &= 65742187.5 \text{ Mb} \\ &= 64201.4 \text{ Gb} \\ &= 62.7 \text{ Tb} \end{aligned}$$

Total = 64.6 Tb

A partir del análisis realizado de Rendimiento vs. Capacidad se pasó al diseño del Centro de Datos que estará compuesto por un sistema de comunicaciones de red de alta velocidad y demanda, capaz de manejar el tráfico para SAN (Storage Area Networks), NAS (Network Attached Storage), Granja de Servidores de Archivos, Aplicaciones, y otros componentes localizados en un ambiente controlado.

En el [Anexo 6](#) se muestra un pequeño esbozo del Centro de Datos de Registro y Notarías.

2.6.2 Selección del Hardware.

Basados en la definición de alcanzar mayor confiabilidad, confidencialidad y disponibilidad de sus aplicaciones, a menor costo, las aplicaciones se hospedarán en equipos de alta disponibilidad de forma tal que se pueda garantizar que ante una falla eventual el servicio no se verá interrumpido. Para lograr esto, los aspectos de redundancia considerados para mantener el ambiente de operación conllevan desde el abastecimiento eléctrico propio del Data Center, hasta los dispositivos de hardware sobre los cuales se instalan las aplicaciones.

A partir de identificar las aplicaciones esenciales en el proceso como aquellas que cuando no se encuentran disponibles, el cliente no puede continuar con sus tareas debido a no poder disponer en ese momento de las facilidades que proveen dichas aplicaciones se decidió hospedar dichas aplicaciones en equipos que cuentan con niveles de disponibilidad del 99.9%.

A continuación se detallan las especificaciones técnicas de los dispositivos que serán utilizados para alojar los aplicativos que forman parte de la solución propuesta al gobierno de la República Bolivariana de Venezuela.

Características principales de los Dispositivos Empleados.

- **HP ProLiant DL580 G2:** Servidor diseñado para ofrecer la máxima escalabilidad y alta disponibilidad. Su innovador chasis ofrece flexibilidad y capacidad de mantenimiento inigualables con un factor de forma versátil optimizado para bastidor. Este servidor se basa en las sólidas características de los servidores DL580 de generación anterior con nuevas prestaciones, como el procesamiento Dual-Core de 64 bits y memoria RAID de conexión en caliente de acceso frontal, que aportan mayor rentabilidad. Además HP ProLiant DL580 G2 admite los procesadores Intel® Xeon® Dual-Core 7041 y 7030 más recientes para ofrecer la máxima escalabilidad. Con cuatro ranuras PCI Express (x4) estándares y una PCI-X de 64 bits/133 MHz y la posibilidad de añadir ranuras PCI Express o PCI-X de conexión en caliente, el HP ProLiant DL580 G2 proporciona mucho ancho de banda y flexibilidad para expansión. El HP ProLiant DL580 G2 está equipado con prestaciones estándares que mejoran el rendimiento, incluidas las tarjetas de interfaz de red duales Gigabit integradas y el controlador Smart Array P400 con hasta 512 MB de caché de escritura respaldada por batería. Elimina cuellos de botella en el rendimiento del sistema de E/S con buses frontales duales e independientes a 800 MHz, SDRAM DDR2 a 400 MHz y tecnología PCI-X/PCI Express[33].
- **HP ProLiant DL380:** Servidor más vendido del mundo, hace honor a su historial de excelencia en diseño, con una disponibilidad y capacidad de gestión de categoría empresarial, el rendimiento probado del procesador Intel Xeon de 2 vías y gran densidad en un tamaño 2U para una variedad de aplicaciones e implantaciones en bastidor[34]. Rendimiento probado para aplicaciones exigentes de ampliación: procesadores Intel Xeon Dual-Core serie 5000 de vanguardia. Soporta

hasta 32 GB de SDRAM PC2-5300 (DDR2-667) con intercalado de 4 vías para admitir aplicaciones exigentes que hacen un uso intensivo de la memoria; 4 ranuras de expansión PCI-Express y PCI-X opcional ofrecen las más recientes tecnologías de E/S de alto rendimiento; Tarjetas de red Gigabit multifunción dobles con TOE para reducir la latencia de la red. Versatilidad y disponibilidad para una amplia gama de implantaciones: protección total de los datos con HP Smart Array P400 con 256 MB de caché de lectura (modelos básico y de alto rendimiento); o Smart Array P200 con 64 MB de caché de escritura (modelo de gama baja). Ambos controladores son actualizables con caché de escritura respaldada por batería y mayores tamaños de caché de lectura. Diseñado para ofrecer fiabilidad y facilidad de gestión: diseño mecánico limpio, sin uso de herramientas que mejora la fiabilidad y simplifica la configuración y mantenimiento.

- HP StorageWorks Enterprise Virtual Array 8000:** Proporciona una potente solución de almacenamiento para los clientes de gama media con un alto rendimiento que elimina los costes del almacenamiento tradicional[35]. Ideal para grandes departamentos y aplicaciones de misión crítica. Estas cabinas son las soluciones de almacenamiento más extensible, elástico y controlable del mercado. Le ofrecerán una escalabilidad extraordinaria, rendimiento líder de la industria, un conjunto perfectamente integrada de herramientas de gestión y cualidades de protección de datos y soluciones de tolerancia de desastres sin precedentes en entornos heterogéneos.

Características del Sistema	
Sistemas operativos compatibles	HP-UX, HP OpenVMS, HP Tru64 UNIX, Windows 2000 Server & Advanced Server, Windows 2003 Standard/Enterprise (32/64-bit) and Extended/Data Center (64-bit), Sun Solaris, Linux, IBM AIX, Novell NetWare, VMware
Capacidad máxima	120 TB
Interfaz de host	Fibre Channel Switched Fabric
Chasis	42U Cabinet
Configuraciones RAID soportadas	Vraid0, Vraid1, Vraid5
Tamaños/velocidades de discos	Discos FC 300GB/10K, 146GB/10K, 73GB/10K o 15K,

	36GB/15K o 10K Discos FATA 250GB 7.2K
Número máximo de discos	240
Caché máxima	4 GB

2.6.3 Software Base Instalado.

Los Sistemas Operativos Linux han experimentado un decidido movimiento hacia los Centros de Datos corporativos para integrarse en los ámbitos de la informática empresarial, las aplicaciones de misión crítica, las bases de datos y las transacciones. Las nuevas tecnologías y capacidades están generando ventajas competitivas en la relación calidad-precio y en el rendimiento, que hacen que Linux sea una opción cada vez más atractiva y estratégica para implantaciones de mayor envergadura.

A continuación se presenta por qué Red Hat Enterprise Linux 4.0 Advanced Server ha sido la elección ideal para todo el entorno de servidores:

Red Hat Enterprise Linux 4.0 Advanced Server pone al alcance de todos una tecnología de avanzada, al hacerla más simple y accesible. Toda organización desea una mejor disponibilidad, flexibilidad, seguridad y rendimiento. Red Hat Enterprise Linux 4.0 Advanced Server elimina el costo y complejidad de la virtualización de clase de empresas, la alta disponibilidad y el software de gestión de almacenamiento. Estos están incluidos en la plataforma. Sin ningún aumento de precio. Totalmente integrados y con soporte de Red Hat.

Características y Ventajas de Red Hat Enterprise Linux 4.

Red Hat Enterprise Linux v.4, lanzado en febrero de 2005, brinda importantes mejoras tecnológicas. Las áreas de desarrollo específico incluyen las mejoras en las capacidades de seguridad, aumento del desempeño del servidor y la escalabilidad, y capacidades mejoradas de desktop - todo al mismo tiempo que asegura un alto nivel de compatibilidad con las versiones anteriores. Red Hat Enterprise Linux soporta un amplio rango de aplicaciones de hardware y software, y es el entorno Linux líder a nivel mundial centrado en la empresa.

La siguiente sección describe algunas de las principales características de Red Hat Enterprise Linux v.4:

Infraestructura de kernel de Linux 2.6.

Red Hat Enterprise Linux v.4 brinda el producto comercial más estable y robusto basado en el kernel 2.6.9 de la comunidad Linux. Como resultado de ello, el kernel de Red Hat Enterprise Linux v.4 ofrece numerosas mejoras sobre los kernel anteriores, incluyendo muchos algoritmos y características mejoradas:

- Programador de CPU lógico genérico: Maneja CPU de núcleo múltiple y con hyperthread.
- VM de Mapeo invertido basado en objetos: Desempeño mejorado en sistemas de memoria limitada.
- Actualización de la copia de lectura: Optimización del algoritmo SMP para estructuras de datos del sistema operativo.
- Programadores I/O múltiples: Seleccionables en base al entorno de aplicación.
- Soporte mejorado de SMP y NUMA: Escalabilidad y desempeño mejorado para grandes servidores.
- Mitigación de la interrupción de la red (NAPI): Desempeño mejorado para cargas de red pesadas.

Capacidades mejoradas de sistemas de almacenamiento y archivo.

Numerosas mejoras se han incorporado a Red Hat Enterprise Linux v.4 que perfeccionan la escalabilidad y el desempeño de los datos en los sub-sistemas de almacenamiento, incluyendo:

- Desempeño Ext3: Reservas de bloque y directorios Hash Tree que mejoran el desempeño de las operaciones de lectura/escritura I/O y escaneo de directorio.
- Escalabilidad de Ext3: Ahora está soportada la expansión dinámica del sistema de archivos y tamaño de los sistemas de archivo hasta 8TB.
- Administrador de Volumen Lógico: Una actualización integral del LVM provee nuevas características como capturas de pantalla de lectura/escritura y actualizaciones metadatos transaccionales, junto con una nueva GUI flexible de administración. Está programado el mirroring y el multipathing mejorado para ser entregado a mediados de 2005.

- Escalabilidad: Administración mejorada de almacenamiento LUN que posibilita configurar subsistemas de almacenamiento mucho más grandes.
- Automontaje: La inclusión de AutoFSv4 ofrece control de acceso de dispositivo sofisticado, con el soporte de características tales como browsable mounts y servidores replicados.
- Reducción de costos: Soporte para almacenamiento en disco Serial ATA que brinda desempeño mejorado, mayor densidad y costo reducido por megabyte sobre los tradicionales dispositivos IDE.

Nuevas capacidades de seguridad.

El suministro de las capacidades de seguridad sofisticada ha sido un foco de desarrollo específico en la versión de Red Hat Enterprise Linux v.4. Las nuevas características incluyen:

- Control de acceso obligatorio: Security Enhanced Linux (SELinux) brinda una infraestructura MAC que complementa las características de seguridad de Control de Acceso Discrecional existentes suministrada por el entorno Linux estándar. En un entorno basado en MAC, las capacidades de aplicación y los privilegios se establecen por políticas predefinidas y son aplicados por kernel. Esto evita que las aplicaciones errantes comprometan la seguridad del sistema.
- Mejoras en la administración de memoria: Algunas características, incluyendo Exec Shield y Position Independent Executables, se combinan para evitar que las aplicaciones sean explotadas mediante ataques como sobreflujos del buffer.
- Verificación de consistencia de compilación y tiempo de ejecución: Nuevas técnicas de validación en el compilador GCC y la librería Glibc reducen enormemente el riesgo de verse comprometidas por aplicaciones defectuosas.

La utilización como Sistema de Gestor de Bases de Datos Oracle Database 10g Enterprise Edition ofrece confiabilidad, escalabilidad y desempeño de primer nivel para configuraciones en cluster y en un solo servidor. Ofrece las más completas características para soportar el procesamiento de transacciones más exigente, inteligencia de negocios, y aplicaciones para la administración de contenido. Oracle Database 10g Enterprise Edition brinda desempeño y escalabilidad de récord absoluto para el procesamiento de transacciones y depósitos de datos de gran escala en Windows, Linux, y servidores UNIX. Es ideal para ambientes que necesitan soportar un gran volumen de transacciones on-line y aplicativos de data

warehouse que realizan uso intenso de consultas. Ofrece escalabilidad comprobada en todas las configuraciones de hardware y puede ser usado para administrar vastas cantidades de información con el nivel más alto de garantía de seguridad del mercado[36].

Características de Oracle Database 10g Enterprise Edition.

- Soportado en todos los ambientes: Oracle Database 10g Enterprise Edition está disponible en todos los sistemas operativos soportados por Oracle.
- Soporta todos los tipos de datos relacionales estándar, almacenamiento nativo de datos en XML, de texto documentos, imágenes, audio, vídeo y localización.
- Soporta transacciones y consultas distribuidas entre dos o más bases de datos relacionales.
- Ejecuta todos los aplicativos.
- Disponibilidad constante.
- Seguridad comprobada.
- Rápido de instalar, fácil de administrar.
- Desarrollo para el grid.

Beneficios de Oracle Database 10g Enterprise Edition.

- Protección ante las fallas del servidor, fallas del sitio, errores humanos, y reducción del tiempo de baja programado.
- Protección de datos con seguridad única en el nivel de filas, auditoras detalladas, y encriptación transparente de datos.
- Incluye data warehousing de alto desempeño, procesamiento analítico online, y características de extracción de datos.
- Asegura datos a nivel de archivos, auditoria de grano fino y codificación transparente de datos.
- Reduce el tiempo de inactividad planificada.
- Incluye almacenes de datos de alto desempeño, proceso analítico en línea y características de minería de datos.

Por tanto con el empleo de Oracle Database 10g Enterprise Edition pretendemos proveer mejoras en el desempeño, disponibilidad, y seguridad; contribuyendo a lograr una mayor calidad de servicio mientras se reduce los costos y complejidad de la gestión de la información. Además con el incremento en las capacidades de grid computing Oracle Database 10g resulta en un mejor desempeño y mayor disponibilidad y facilidad en la gestión, siendo Oracle Real Application Clusters (RAC) la única tecnología de Bases de Datos en el mercado que permite que el cliente convierta un cluster de servidores estándar de la industria en una plataforma escalable y tolerante a fallas para el funcionamiento de cualquier carga de trabajo de la base de datos.

2.6.4 Diseño e Implementación de un Real Application Cluster (RAC).

Oracle 10g provee la primera infraestructura completa e integrada para potenciar el grid computing. Oracle 10g recoge los atributos fundamentales de esta tecnología de computación:

- **Implementar uno a partir de muchos:** Se trata de conseguir que un número indeterminado de máquinas funcionen como una sola. Para ello, se tiene que aplicar un proceso de virtualización en todas las capas del sistema y una zona común donde se almacenen los recursos.
- **Administrar Muchos como si fueran uno:** Para conseguirlo debemos contar con un software a medida y de un sistema de administración centralizado.
- **Los dos puntos anteriores deben aplicarse a cada elemento del grid:** sistemas de almacenamiento, bases de datos, servidores de aplicación y aplicaciones.

2.6.5 Qué es Oracle Database 10g Real Application Clusters?

Real Application Clusters es un ambiente de computación que aprovecha el poder de procesamiento de múltiples computadoras interconectadas (cluster) para la ejecución de aplicaciones, sin necesidad de cambios en las mismas.

Oracle 10g Real Application Clusters (Oracle 10g RAC) es el producto distintivo de ORACLE que permite configurar arquitecturas de sistemas de Clusters SMP (symmetric multi-processors) para correr sus aplicaciones. En cada uno de los nodos se ejecutan concurrentemente transacciones contra la misma

base de datos, RAC coordina el acceso a los datos compartidos, para proveer consistencia e integridad[37].

2.6.5.1 Arquitectura de un Real Application Clusters (RAC).

Una configuración Oracle 10g RAC mejora los tiempos de respuesta de su aplicación para satisfacer una necesidad de procesamiento creciente. A medida que usted agrega recursos, Real Application Clusters puede utilizarlos y extender su poder más allá de los límites de componentes individuales. Además, la tecnología de Oracle, permite mejorar el tiempo de ejecución de sus consultas gracias al uso de consultas en paralelo en diferentes nodos (multinode parallel query).

Un ambiente adecuadamente configurado tolera fallas con un tiempo de caída mínimo. El servicio de Transparent Application Failover (TAF), de recuperación transparente ante fallas, permite que ante una caída en un servidor del cluster los usuarios de la aplicación no sean afectados. Oracle 10g RAC es un componente fundamental en soluciones de alta disponibilidad.

La figura representa la arquitectura de un Oracle Database 10g Real Application Clusters:

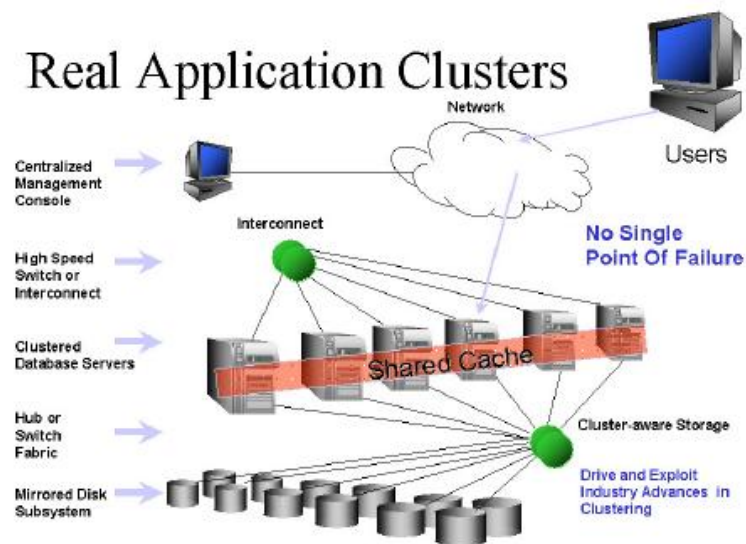


Figure 1 Real Application Clusters Architecture

¿Por qué es Oracle Real Application Clusters la solución correcta?

- Cuando estas arquitecturas paralelas son implementadas correctamente, se obtienen mejoras dramáticas en los tiempos de respuesta y ejecución, en las posibilidades de escalabilidad y crecimiento, y en la disponibilidad de sus Bases de Datos y aplicaciones. Sin embargo, el alto grado de sofisticación de estas tecnologías requieren de un especial cuidado de parte de los administradores del sistema y de las bases de datos.
- El servicio de Fast Track to Oracle 10g RAC de Snoop Consulting esta diseñado para permitir que nuestros clientes, mediante una configuración adecuada, obtengan un uso eficiente y seguro de sus aplicaciones sobre una plataforma de RAC en el menor tiempo posible, con una relación costos-beneficios tangibles.

Con Oracle Real Application Clusters, una arquitectura de caché compartida que le ofrece todas las ventajas del Grid Computing, se puede obtener:

- **Actividad sin interrupciones de las aplicaciones más críticas:** Oracle RAC supera en rendimiento incluso a los sistemas mainframe más rápidos, proporcionándole una alta disponibilidad y una velocidad de máximo nivel. Oracle pone a su disposición las tecnologías necesarias y una arquitectura muy eficaz de alta disponibilidad, que le permitirá mantener protegidos los datos y los sistemas de su empresa frente a factores imprevistos y prolongados.
- **Escalabilidad bajo demanda para adaptarse a los cambios futuros:** Cuando necesite más capacidad y potencia, bastará con añadir otro servidor estándar, lo cual supone una gran flexibilidad para la empresa y un ahorro de hasta el 40% en costes de hardware.
- **Menores costes de administración:** La flexibilidad, la tolerancia a los fallos y el hecho de tener un solo punto de control hacen que con Oracle RAC pueda reducir a la mitad los costes de administración.

En el [Anexo 7](#) se muestra el diseño del Real Application Cluster que se implementó en el Centro de Datos del Proyecto Registro y Notarias.

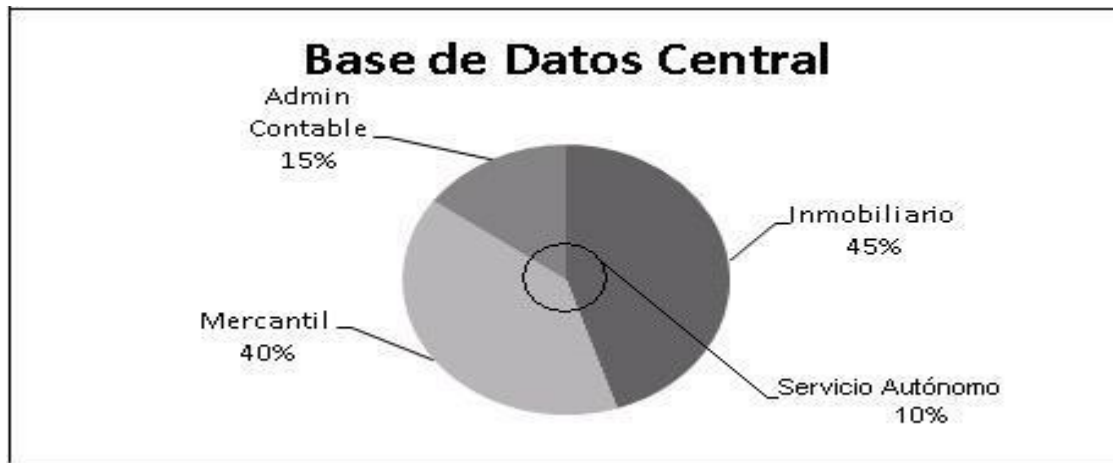
2.6.5.2 Beneficios de Oracle Real Application Clusters.

- **Alta Disponibilidad:** Oracle Real Application Clusters 10g provee una infraestructura de alta para Centros de Datos de alta demanda y disponibilidad. Es un componente esencial en la arquitectura de alta disponibilidad propuesta por Oracle, proporcionando protección para soluciones de alta accesibilidad.
- **Recuperación:** La Base de Datos Oracle incluye un grupo de características que simplifican el proceso de recuperación ante cualquier tipo de anomalía. Si se produce una falla en una de las instancias de la Base de Datos del RAC, entonces la anomalía es notificada a las demás instancias del RAC, que automáticamente ejecutarán un proceso de recuperación en la instancia fallida.
- **Detección de Errores:** Oracle Clusterware automáticamente monitorea la Base de Datos del RAC y provee mecanismos para la detención rápida de fallos tales como: Fast Application Notification que se encarga notificar cualquier anomalía en los componentes del RAC.
- **Funcionamiento Ininterrumpido:** Real Application Clusters provee un servicio de continua accesibilidad para aplicaciones de alta demanda y disponibilidad. Si un nodo del RAC saliera del sistema producto de una anomalía, la Base de Datos permanecería abierta y los aplicativos continuarían sus operaciones de acceso a los datos. Todo este proceso se realiza con total transparencia a los usuarios de la aplicación.
- **Escalable:** Oracle Real Application Clusters permite que una única base de datos se expanda por múltiples nodos en un grid o red, uniendo los recursos de varias máquinas. Se pueden añadir servidores a la vez que eliminarlos en un cluster Oracle sin tiempo de inactividad, es decir, sin detener la Base de Datos, sin importar tampoco la plataforma donde se encuentra instalado el servidor.

2.6.6 Diseño de la Base de Datos Central.

La Base de Datos Central fue concebida como una agrupación de todas las bases de datos locales, en ella se almacenará una copia diaria de todo el flujo de trabajo realizado en las oficinas. De este modo se mantiene un control centralizado de toda la información concerniente a los Registros y Notarías de la República Bolivariana de Venezuela.

En la figura se muestra un esquema de la Base de Datos Central:



2.6.6.1 Diseño Físico de la Base de Datos Central.

En el Centro de Datos se utilizó una estrategia similar a las oficinas. Esta base de datos mantendrá una copia de todas las bases de datos de las oficinas.

Se crearon dos tablespaces, a los que se asignaron todas las tablas atendiendo a sus características comunes en cuanto a volumen de datos.

1. Se creó un tablespace en el que se agrupan las tablas que mayor cantidad de operaciones reciben (insert, update, delete) y sobre las que las búsquedas serán mas exhaustivas, estas tablas son las llamadas críticas y pueden alcanzar un alto volumen de datos.
2. Se creó un tablespace en el que se agrupan todas las tablas sobre las que se realiza un número menor de operaciones en comparación con las del grupo anterior, en este grupo se encuentran las llamadas nomencladores, sobre las que se realizan, en la mayoría de los casos, solo consultas. Además se creó un tablespace para almacenar todos los usuarios de la base de datos y otro para almacenar todos los índices creados para el mejor funcionamiento de la misma.

Con la utilización en el Centro de Datos la versión Oracle 10g Enterprise Edition release 2 fue posible utilizar todas las funcionalidades que brinda la misma, como la de Partición de Oracle, que es un proceso mediante el cual se permite que una tabla, un índice o una tabla organizada por índices puedan ser divididas en partes más pequeñas. Cada partición tiene su propio nombre, y puede opcionalmente tener sus propias características de almacenamiento, tales como ser almacenadas en diferentes tablespaces. Estas particiones se llevan a cabo usando una **llave de partición**, que es un grupo de columnas que determinan en que partición residirá una tupla. Desde la perspectiva del administrador de bases de datos estas particiones brindan una flexibilidad considerable para administrar los objetos de la base de datos. Sin embargo, desde la perspectiva de la aplicación, una tabla particionada es idéntica a una no particionada; no se necesitan hacer modificaciones para acceder a la misma usando comandos SQL / DML.

Estrategia de Almacenamiento de la Base de Datos Central.

Oracle 10g incorpora varias características que simplifican, flexibilizan y automatizan el almacenamiento de la base de datos. Una de las características es Automatic Storage Management (ASM, Administración de Almacenamiento Automático), que es un administrador de volumen integrado verticalmente y un sistema de archivos integrado para los archivos de datos de Oracle. La otra característica que ayuda a simplificar la gestión del almacenamiento de la base de datos es la capacidad de renombrar espacios de tablas.

Gestión de Almacenamiento. Administración Automática del Espacio de Almacenamiento (ASM).

Esta nueva herramienta trata de simplificar la administración del espacio para la Base de Datos Oracle. En vez de manejar numeroso archivos para una base de datos, el Administrador de la Base de Datos sólo gestiona un pequeño número de grupos de disco. Un grupo de disco es un conjunto de dispositivos físicos de disco que Oracle maneja como una única unidad lógica. Un DBA puede definir un diskgroup como el predefinido para una base de datos, y Oracle automáticamente gestiona ese espacio para crear y eliminar los ficheros relacionados con los objetos de la Base de Datos. También permite la integración con sistemas de almacenamiento tipo RAID o Logical Volume Managers (LVMs). Otra ventaja de ASM es que está concebido por Oracle, con lo cual su integración con el resto de productos es mucho mayor.

ASM es una instancia de Oracle en sí misma y administra la capa de almacenamiento con todas las bases de datos que se encuentren en el nodo. A cada nuevo espacio de almacenamiento al que se le da acceso a la instancia ASM se llama un disco ASM. El conjunto de espacios de almacenamiento es un grupo de discos ASM (diskgroup), se da esta agrupación para darles diferente nivel de redundancia principalmente[38].

Ventajas de ASM.

- ASM automatiza y simplifica la disposición óptima de los archivos de datos, archivos de control y archivos de registro.
- Los archivos de bases de datos se distribuyen automáticamente entre todos los discos disponibles y el almacenamiento de la base de datos se vuelve a equilibrar cada vez que se modifica la configuración del almacenamiento.
- ASM proporciona redundancia a través del reflejo de los archivos de la base de datos. ASM se diseñó específicamente para utilizar servidores y discos debajo costo.

Por tanto a partir del análisis anteriormente realizado se decidió implementar la gestión automática de almacenamiento para administrar los archivos de la Base de Datos. Con ASM, no solo se administran los archivos de control y redo log, sino también los datafiles.

En el [Anexo 8](#) se muestra como se implementó la distribución en Grupos de Discos ASM en la Base de Datos Central.

2.7 Replicación de Datos utilizando Oracle Streams.

Oracle Streams, es una tecnología para compartir información. Detecta que información es relevante y quienes la utilizan. Esta tecnología es utilizada por Oracle 10g para propagar los cambios en un ambiente replicado. Se basa en tres acciones aplicadas a la información, Captura, Almacenamiento y Consumo[39].

Para la replicación la Captura se relaciona con un mecanismo que toma los cambios de los Redo Log. El Almacenamiento se vincula cuando los cambios capturados son enviados al área de almacenamiento, y estos cambios son propagados a las áreas de almacenamiento de los equipos remotos con réplica.

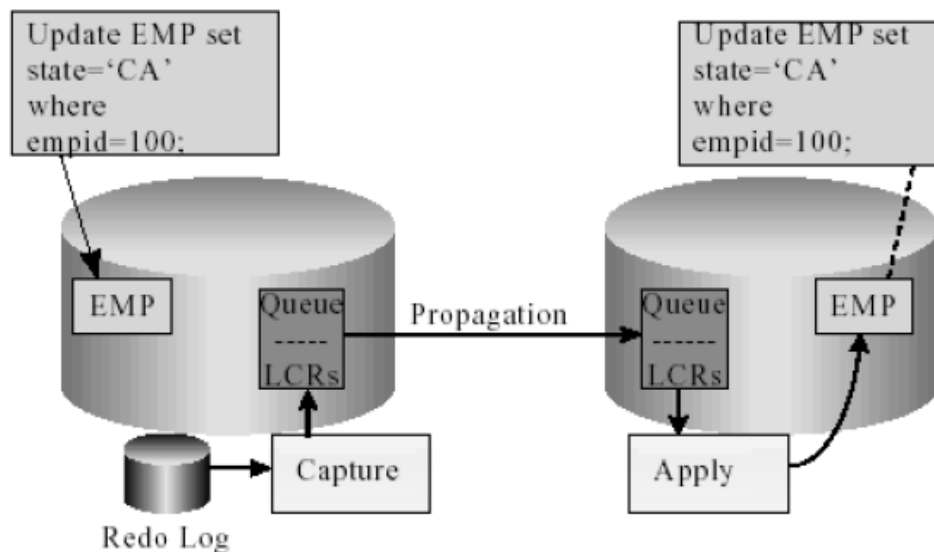
El Consumo es la máquina que se encarga de aplicar los cambios almacenados a la base de datos en cada equipo con réplica.

Las tablas replicadas pueden ser diferentes, Oracle Streams se encarga de transformar la información para ajustarla a la base de datos de cada sitio replicado.

2.7.1 Arquitectura de Oracle Streams.

El Proceso de Captura configurado para recolectar cambios realizados, recupera los mismos desde los Redo Log, formatea la información como Registros Lógicamente Cambiados (LCR logical change records) y coloca los LCR en el área de almacenamiento de la base de datos local. Los LCR son entonces propagados desde el área o cola de almacenamiento de la base de datos origen, hacia las áreas de almacenamiento de las bases de datos destino de la replicación. Las máquinas de consumir de las bases de datos destino son configuradas para recibir los cambios, y cuando estos son propagados a las bases de datos destino las máquinas de consumo locales, se encargan de aplicar automáticamente los cambios en ellas.

La figura ilustra los procesos de captura, propagación y aplicación de una transacción utilizando Oracle Streams:



2.7.1.1 Proceso de Captura.

El Proceso de Captura basado en logs aprovecha el hecho de que los cambios en las tablas son almacenados en los Redo Log para garantizar la recuperabilidad. Capturando los cambios directamente del Redo Log se pretende minimizar la recarga del sistema. El Proceso de Captura recupera los datos cambiados, los formatea como LCR y los coloca en el área de almacenamiento local. El proceso de captura puede filtrar los LCR basado en reglas predefinidas. Los registros lógicamente cambiados (LCR) describen los cambios efectuados en una fila con alguna sentencia DML. Cada LCR incluye, el nombre de la tabla, los nuevos y viejos valores para cada columna cambiada, y los valores de las claves. Para distinguir entre cambios provenientes de diferentes lugares el LCR también almacena información de la base de datos de origen. La búsqueda de información tiene una opción mediante la cual se puede buscar en los archivos del Redo Log, o capturar la información mientras se está escribiendo en el mismo.

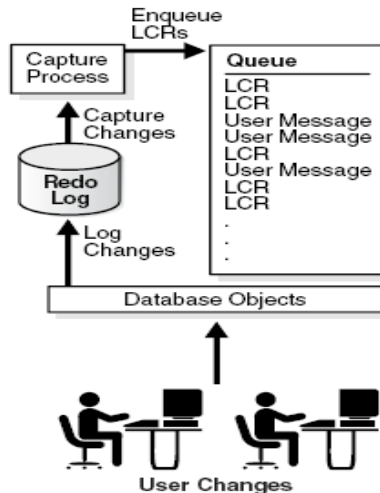
Las aplicaciones de usuario pueden explícitamente colocar mensajes representando estos en el área de almacenamiento. Estos mensajes se pueden formatear como LCR y ser luego consumidos. Cada elemento en el flujo utiliza un conjunto de reglas para distinguir los eventos a ser manejados. Las reglas especificadas por el usuario permiten a la captura ser más selectiva al colocar eventos en el área de almacenamiento. La máquina de consumo aplicará los cambios selectivamente basada en las reglas definidas para ella, del mismo modo se utilizan reglas para la propagación selectiva de eventos entre diferentes áreas de almacenamiento.

Las reglas son expresiones tipo SQL que evalúan a verdadero o falso. Pueden limitar los eventos basadas en el contenido de los propios eventos. Pueden abarcar bases de datos, esquemas o tablas.

Muchas reglas componen un conjunto de reglas y varios conjuntos de reglas pueden tener la misma regla individual. Cada flujo de Oracle se basa en un conjunto determinado de reglas.

En resumen, por cada cambio confirmado a través de triggers se genera un registro LCR en el área de almacenamiento local, el mismo contiene referencias al Redo Log en el que efectivamente se almacena la información sobre el cambio efectuado.

La figura representa un Proceso de Captura Local:



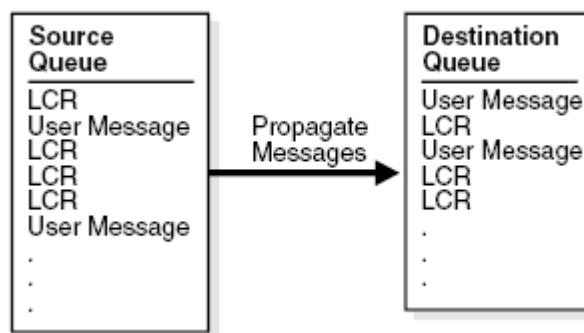
2.7.1.2 Proceso de Almacenamiento o Proceso de Propagación.

Es una cola para almacenar y manejar los eventos capturados. Los cambios en la base de datos convertidos a LCR's son capturados y almacenado en una cola hasta que sean consumidos. La máquina de consumo local, una aplicación del usuario, o el área de almacenamiento de otra base de datos remota, pueden examinar el contenido del área de almacenamiento para ver si les interesa o no un mensaje representando un evento. Opcionalmente pueden evaluar un conjunto de reglas para determinar cuando cierto mensaje se ajusta a un criterio dado y entonces puede ser consumido.

En el caso de una aplicación de usuario la misma quitará de la cola en el área de almacenamiento el mensaje especificado, en el caso de otra área de almacenamiento remota, el mensaje será propagado hacia la misma. Finalmente en el caso de la máquina de consumo local la misma quitará y consumirá dicho mensaje de la cola. Los eventos en el área de almacenamiento en un ambiente replicado pueden ser enviados hacia áreas de almacenamiento de bases de datos remotas. Los eventos se pueden dirigir a través de una o varios sistemas hasta alcanzar el sistema destino. Una base de datos puede ser utilizada como pasarela entre otras dos, aún si en ella no se aplican los cambios que recibe y reenvía. Oracle mantiene información de cada evento en la base de datos de origen para casos en los que se deban diferenciar cambios realizados en diferentes lugares.

Una transformación es un cambio en la forma o en los datos de un objeto participante en una captura o consumo de datos. La transformación se puede representar como una función PL/SQL, que toma como entrada un tipo de datos y devuelve un objeto del tipo de datos destino. Se puede especificar una transformación para la captura de datos, durante el almacenamiento o durante el consumo del mismo.

La figura representa el Proceso de Propagación desde una Base de Datos Local a una Base de Datos Remota:



2.7.1.3 Proceso de Aplicación o Consumo.

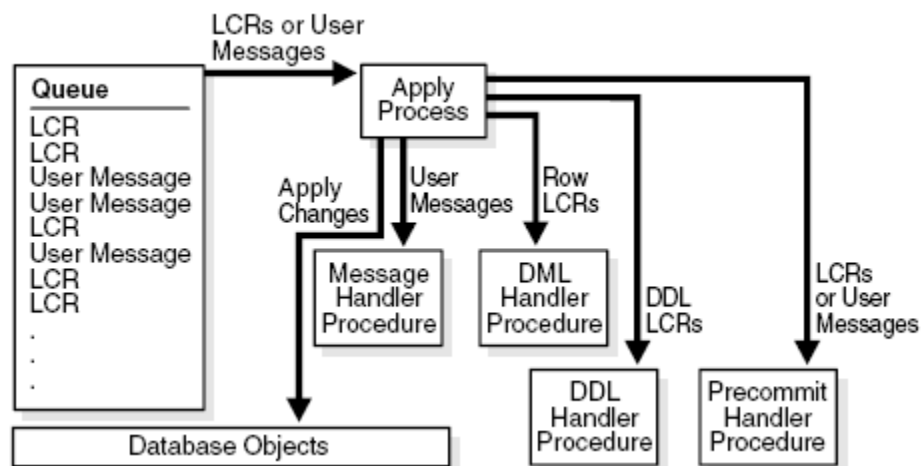
Los mensajes en el área de almacenamiento son consumidos por la máquina de consumo cuando los cambios que ellos representan, son aplicados a la base de datos, o utilizados por una aplicación. El hecho de soportar que los mensajes sean explícitamente quitados de la cola permite que otras aplicaciones sean notificadas de cambios en los datos. La máquina de consumo aplica cambios de DDL, cambios de DML y cambios introducidos por aplicaciones de usuario en forma de LCR's. La máquina de consumo corre localmente en el sistema que alberga la base de datos local. Los conflictos son detectados cuando la fila destino ha sido cambiada y no contiene los valores originales esperados. La máquina de consumo consiste en un coordinador, un buscador y uno o más aplicadores corriendo en segundo plano. El buscador construye las transacciones desde LCR's generados por capturas. El coordinador organiza las dependencias entre transacciones y organiza las dependencias a nivel de las DML's.

Los aplicadores incorporan los cambios a la base de datos local. Los cambios provenientes de múltiples bases de datos pueden ser enviados al área de almacenamiento de una única ubicación. Varios aplicadores harán efectivos los cambios de cada base de datos de forma concurrente.

La máquina de consumo soporta subconjuntos de objetos, por lo tanto una base de datos destino, puede pedir los cambios de apenas un subconjunto, de los datos del origen, seleccionado mediante algún criterio. Las operaciones cuyo LCR no cumple con los criterios establecidos no serán aplicadas en la base de datos de réplica.

La máquina de consumo puede también pasar los LCR a procedimientos definidos por el usuario, por lo tanto se pueden definir procedimientos separados, para cada tipo de operación de DML, por ejemplo. Las aplicaciones de usuario pueden explícitamente, quitar de la cola LCR's u otros mensajes desde el área local de almacenamiento.

La figura representa al Proceso de Aplicación o Consumo de Mensajes:



Oracle 10g en lo que concierne a replicación tiene como principales objetivos, la alta disponibilidad, la escalabilidad, y la mejora en el performance. Siendo el último punto en el que concentra su mayor atención.

2.7.2 Mejores Prácticas para Oracle Streams.

Una Base de Datos fuente (source database) es aquella en la que los cambios capturados por un proceso de captura son generados en los archivos de Redo Log locales. Un proceso de captura local puede capturar los cambios ocurridos en una Base de Datos fuente, mientras que un proceso de captura downstream puede capturar los cambios en una Base de Datos remota.

A continuación se detallan las mejores prácticas de configuración utilizando Oracle Streams en Bases de Datos fuentes que utilizan procesos de captura locales o procesos de captura downstream:

- Activar en modo archivado (archivelog) cada Base de Datos fuentes en un entorno de Oracle Streams, caso de que se presente un ambiente de captura downstreams la Base de Datos donde se generan los archivos de datos fuente debe estar en modo archivado.
- Configuración de Supplemental Logging en cada Base de Datos fuentes en un entorno de Oracle Streams. La configuración de Supplemental Logging es requerida para aquellas tablas cuyos cambios serán capturados por el proceso de captura empleado por Oracle Streams. En Oracle Database 10g release 2 y versiones superiores, el registro suplementario es adicionado automáticamente a las llaves primarias e índices únicos.
- Configuración de una heartbeat tabla en cada Base de Datos fuentes en un entorno de Oracle Streams. El empleo de una tabla heartbeat es para asegurarse de que los cambios están ocurridos en la Base de Datos fuentes están siendo replicados en un ambiente de replicación Streams. Una tabla heartbeat es especialmente útil para bases de datos que tienen un bajo grado de actividad porque se puede asegurar que el ambiente de réplica está funcionando bien aún cuando no haya muchos cambios replicados.

En un entorno de replicación Oracle Streams, la Base de Datos destino es una Base de Datos en la cual son aplicados los cambios que fueron capturados en una Base de Datos fuente a través de un proceso de aplicación o (Apply Process).

A continuación se detallan las mejores prácticas de configuración utilizando Oracle Streams en Bases de Datos destinos:

- Conceder los privilegios requeridos al usuario encargado de configura el o los procesos de aplicación de cambios o (Apply Processes) en la Base de Datos destino.
- Instanciar cada objeto a un valor System Change Number (SCN), este paso se realiza con el objetivo de indicar a partir de que SCN el proceso de aplicación o (Apply Processes) aplicará los cambios sobre un objeto. Este proceso es conocido en un entorno de Oracle Streams como “Set Instantiation SCN Values”.

2.7.2.1 Mejores Prácticas para la configuración de los procesos de captura.

Esta sección describe algunas buenas prácticas para la captura de mensajes en un entorno de replicación Oracle Streams.

- Configuración del parámetro parallelism en el o los procesos de captura de la Base de Datos fuente. Para establecer el paralelismo para cada proceso de captura se especifica el parámetro de parallelism en el procedimiento almacenado DBMS_CAPTURE_ADM.SET_PARAMETER. Este parámetro controla el número de servidores de ejecución paralelos que concurrentemente minan los archivos de redo en busca de cambios. El valor por defecto para el paralelismo de procesos de captura es 1, y es apropiado para la mayoría de las configuraciones de procesos de captura. Asegurarse de que los parámetros de inicialización PROCESSES y PARALLEL_MAX_SERVERS están definidos correctamente cuando se defina el parámetro de paralelismo de los procesos de captura.
- Configuración de tiempo de retención para cada proceso de captura. Periódicamente, un proceso de captura toma un punto de chequeo para facilitar el reinicio de la base de datos más rápido. Estos tiempos de chequeos son mantenidos por defecto en el tablespace SYSAUX. El tiempo de retención de puntos de chequeo para un proceso de captura controla la cantidad de datos de puntos de chequeo que el proceso tiene. El tiempo de retención de puntos de chequeo especifica el número de días previos al SCN (System Change Number) de puntos de chequeo requeridos para retener puntos de chequeo. Cuando un punto de chequeo es más viejo que el periodo de

tiempo especificado, el proceso de captura limpia el punto de chequeo. El valor por defecto del tiempo de retención de puntos de chequeo es de 60 días. Si existen puntos de chequeo disponibles por un tiempo en el pasado, el proceso de captura puede ser usado para recapturar cambios para recuperar una base de datos destino. Se debe definir para el tiempo de retención de puntos de chequeo un valor apropiado para el ambiente en el que se trabaja. La configuración típica es de 7 días.

2.7.2.2 Mejores Prácticas para la configuración de la propagación.

Esta sección describe algunas buenas prácticas para la propagación de mensajes en un entorno de Oracle Streams.

- Propagación de mensajes Queue-to-Queue. En un entorno de replicación Oracle Streams se recomienda la implementación del mecanismo de propagación Queue-to-Queue. La propagación de mensajes Queue-to-Queue permite que cada evento de propagación planificado o en ejecución puedan ser administrados independientemente y por separado.
- Configuración de parámetro latency en el o los procesos de propagación de la Base Datos fuente. Este parámetro representa el tiempo máximo de espera, en segundos para que un mensajes pueda ser propagado desde una Base de Datos fuente a una Base de Datos destino, posterior de ser enviado este a la cola de mensajes. El valor asignado por defecto a este parámetro es 60 segundos, se recomienda un valor inicial de 5s.

2.7.2.3 Mejores Prácticas para la configuración de los Procesos de Aplicación.

La siguiente sección describe algunas buenas prácticas de configuración para los procesos de aplicación o (Apply Processes) en Bases de Datos destinos.

- Configuración del parámetro parallelism en el o los procesos de consumo (Apply Processes) en Bases de Datos destinos. Este parámetro controla el número de ejecuciones en paralelo que

actúan concurrentemente en la aplicación de los cambios. Se plantea que un valor inicial del parámetro entre 1 o 4 es correcto.

- Permitir al Proceso de Consumo o (Apply Processes) continuar su actividad procesamiento de mensajes, aunque detecte errores no fatales durante el proceso de aplicación de los mensajes en cola. Cuando el valor asignado al parámetro de configuración “disable_on_error” es “y” entonces serán descartados todos aquellos errores no fatales que ocurran durante el proceso de aplicación.

Por todo lo anterior descrito podemos afirmar que Oracle Database 10g release 2 presenta un ambiente de replicación flexible y completo, flexible ya que permite la combinación de diferentes técnicas de replicación abriendo un amplio horizonte de posibilidades; y completo ya que las herramientas provistas por Oracle Database 10g release 2 cubren la totalidad del proceso de replicación, permitiendo incluso la administración centralizada del sistema.

2.7.3 Estrategias de Respaldo y Recuperación en el Centro de Datos de Registros y Notarias.

Las consecuencias de la pérdida de datos pueden ser catastróficas. Los datos de un Centro de Datos son decisivos para atender a sus clientes e incrementar su volumen de negocio. Entre las posibles amenazas que afectan a la seguridad de los datos se incluyen cambios en la infraestructura de IT, siniestros en las ubicaciones y virus. Sin embargo, resulta irónico descubrir que los errores humanos y el mal funcionamiento del software y del hardware suelen ser con diferencia las fuentes más comunes de problemas.

Los requisitos clave del mercado global actual son la recuperación inmediata, el mantenimiento de los datos sin pérdidas y la eliminación de ventanas de copia de seguridad para permitir un procesamiento durante veinticuatro horas al día, siete días de la semana. Conscientes de la importancia de disponer de mecanismos preventivos para proteger los datos críticos y poder reanudar sus operaciones en caso de desastre, con el mínimo impacto posible.

2.7.4 Modulo de Administración y Control de Backups.

HP OpenView Storage Data Protector 6.0 es un software altamente escalable de protección de datos que automatiza el respaldo y la recuperación en disco o cinta. La versión 6.0 reduce de manera considerable el tiempo y los recursos necesarios para realizar respaldos al crear un respaldo completo a través de apuntadores, lo que elimina la necesidad de realizar respaldos completos frecuentes al consolidar los respaldos incrementales[40]. El servicio relacionado de respaldo y recuperación de HP ofrece una implementación rápida y sencilla de Data Protector en ambientes SAN y no SAN para poder administrar todo el proceso de integración del cliente, incluyendo planeación, implementación, reportes después de la implementación y actividades de prueba.

Un resumen de las funciones principales de HP OpenView Storage Data Protector 6.0 es:

1- Copias de seguridad avanzadas en disco.

- Rápida restauración a partir de disco y acceso transparente a datos migrados a cinta.
- Complementan las tecnologías de copia de seguridad en disco de Data Protector: copia de seguridad sin tiempo de inactividad, recuperación instantánea y librería virtual de cintas (librería de archivos).
- Fácil y sencilla configuración y obtención de licencias Gestión avanzada de medios.
- Copia de medios automatizada que ahorra espacio.
- Migración de tipos de cinta desde disco a cinta y entre distintos tipos de cinta.
- Consolidación de cintas.
- Sistematización de multiplexado de cintas.
- Gestión mejorada de medios mediante la generación automática de etiquetas a partir de códigos de barras.

2- Duplicación de copias de seguridad.

- Permite duplicar copias de seguridad a distancias ilimitadas en destinos físicos distintos.

- Proporciona el mejor rendimiento de duplicación en una SAN.

3- Configuración y administración más sencillas de SANs de gran tamaño.

- El nuevo dispositivo de rutas múltiples ofrece una vista sencilla de todas las rutas de acceso desde todos los sistemas hasta todos los dispositivos de una SAN de gran tamaño.
- Manejo automático de modificaciones en direcciones SCSI Mayor cobertura del ecosistema.
- Aumenta la amplia compatibilidad actual de sistemas operativos, aplicaciones, unidades, librerías y disk arrays.
- Copias de seguridad sin tiempo de inactividad y la recuperación instantánea de Data Protector a través del estándar SMI-S.
- Ejecución de la consola web de librerías.
- Actualiza las integraciones con HP OpenView, admitiendo las versiones más recientes Data Protector.

4- Media Operations es ahora heterogéneo.

- Además desde HP OpenView Storage Data Protector 6.0, puede hacerse seguimiento y gestionar medios de almacenamiento fuera de línea para Veritas NetBackup.

2.7.5 Entorno de Integración de HP OpenView Storage Data Protector 6.0.

La herramienta HP OpenView Storage Data Protector 6.0 integra en un mismo producto la recuperación basada en disco y en cinta, a través de múltiples aplicaciones, sistemas operativos y arquitecturas de almacenamiento de datos. Es una solución que administra el respaldo y recuperación de plataformas de múltiples servidores. Este permite simplificar y centralizar las operaciones de respaldo integrando una variedad de técnicas, las cuales van desde el respaldo on-line, respaldo open file y recuperación instantánea más avanzada del mercado, basadas en las tecnologías de duplicación y de instantáneas[40].

El entorno de integración de HP OpenView Storage Data Protector 6.0 en el Centro de Datos de Registro y Notarías está formado por:

1. Un Servidor Central o Cell Manager: este servidor cumplen la función de almacenar la configuración de toda la red de backup, como por ejemplo clientes, servidores, unidades robóticas, de cintas. Esta información es utilizada para realizar backup a los servidores registrados a través de la LAN o SAN tanto de archivos planos como Bases de Datos Oracle que residen en la plataforma central.
2. Un Servidor de Instalación o Installation Server (IS): este suele ser utilizado como una estación de trabajo aislada al Cell Manager o formar parte de propio Cell Manager como un componente más del Servidor Central. En la configuración presentada en el siguiente esquema el IS se utilizo como un componente del Cell Manager. Este componente contiene un repositorio de software que se emplea para las instalaciones de estaciones de trabajo clientes. Esta característica de HP OpenView Storage Data Protector 6.0 facilita el proceso de instalación, particularmente en estaciones clientes remotos.
3. Ocho Estaciones de Trabajo Cliente: una celda consiste de un Cell Manager y usualmente varias estaciones clientes. En cada estación cliente de un Cell Manager se instalarán aquellos componentes de HP OpenView Storage Data Protector 6.0 según el rol que desempeña la estación cliente en el entorno de integración. Los componentes instalados en cada estación podrán ser instalados localmente en el sistema o remotamente empleando un servidor de instalación o installation server.

En el [Anexo 9](#) se ilustra el entorno de integración de HP OpenView Storage Data Protector 6.0 en el Centro de Datos de Registros y Notarías.

2.7.6 Estrategias de Backups.

Se han trazado un grupo de estrategias dirigidas a la protección y recuperación de los datos ante desastres, que aseguren la integridad de la información.

Estas estrategias han sido aplicadas sobre la base de las diversas alternativas de recuperación ante desastres que posee el software HP OpenView Storage Data Protector 6.0.

La estrategia de Backup implementada en la Base de Datos del Real Application Cluster es la siguiente:

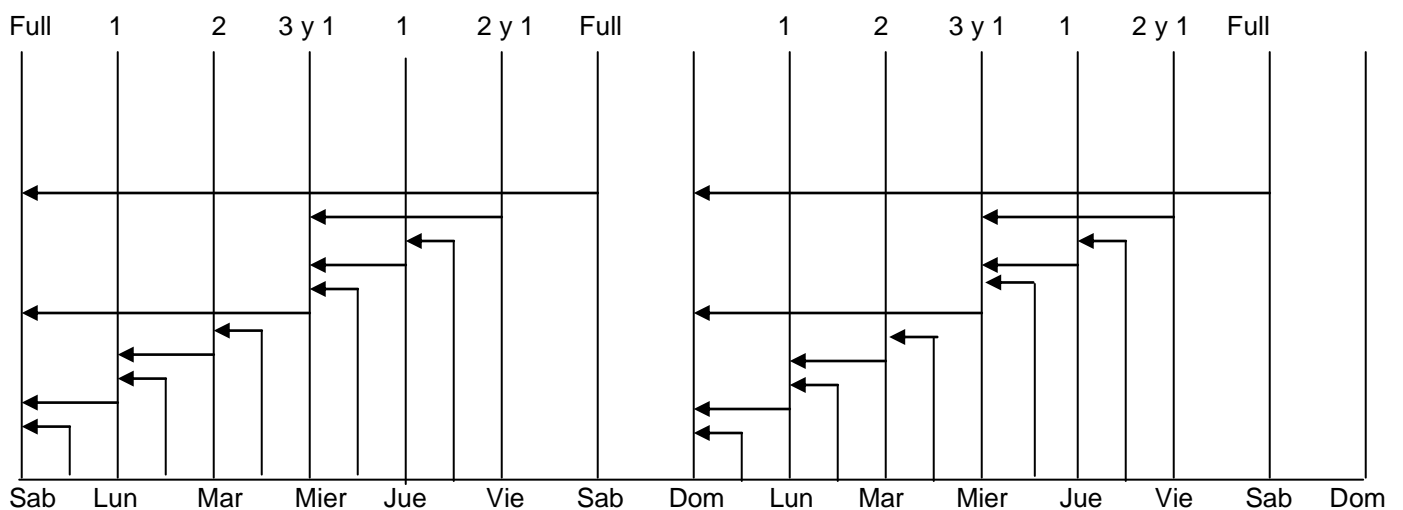
Días de la Semana	Tipo de Backup	Descripción
Sábado.	Backup Incremental Level 0.	Este backup es el punto de partida para los subsiguiente Backup Incrementales Diferencial Level 1. Job Name = backup_full_incr_level_0 Job Description = full incremental sábado Start -> Later = sábado 08:00 pm Repeat -> Interval -> Frequency = 1 weeks
Lunes.	Backup Incremental Diferencial Level 1.	Este backup respaldará todos los bloques de datos que han sido modificados desde el Backup Incremental Level 0 o Level 1 más reciente. Job Name = backup_dif_incr_level_1_lun_1 Job Description = incremental diferencial lunes Start -> Later = lunes 12:00 pm Repeat -> Interval -> Frequency = 1 weeks
	Backup Incremental Diferencial Level 1.	Job Name = backup_dif_incr_level_1_lun_2 Job Description = incremental diferencial lunes Start -> Later = lunes 12:00 am Repeat -> Interval -> Frequency = 1 weeks

Martes.	Backup Incremental Diferencial Level 2.	Job Name = backup_dif_incr_level_2_mar_1 Job Description = incremental diferencial martes Start -> Later = martes 12:00 pm Repeat -> Interval -> Frequency = 1 weeks
	Backup Incremental Diferencial Level 2.	Job Name = backup_dif_incr_level_2_mar_2 Job Description = incremental diferencial martes Start -> Later = martes 12:00 am Repeat -> Interval -> Frequency = 1 weeks
Miércoles.	Backup Incremental Diferencial Level 3.	Este backup respaldará todos los bloques de datos que han sido modificados desde el backup incremental level 1,2 más reciente. Job Name = backup_dif_incr_level_3_mier_1 Job Description = incremental diferencial miércoles Start -> Later = miércoles 12:00 pm Repeat -> Interval -> Frequency = 1 weeks
	Backup Incremental Acumulativo Level 1.	Este backup respaldará todos los bloques de datos que han sido modificados desde el backup incremental level 0 más reciente. Job Name = backup_cum_incr_level_1_mier_2 Job Description = incremental acumulativo miércoles Start -> Later = miércoles 12:00 am Repeat -> Interval -> Frequency = 1 weeks
Jueves.	Backup Incremental Diferencial Level 2.	Job Name = backup_dif_incr_level_2_jueves_1 Job Description = incremental diferencial jueves Start -> Later = jueves 12:00 pm Repeat -> Interval -> Frequency = 1 weeks

	Backup Incremental Diferencial Level 2.	Job Name = backup_dif_incr_level_2_jueves_2 Job Description = incremental diferencial jueves Start -> Later = jueves 12:00 am Repeat -> Interval -> Frequency = 1 weeks
Viernes.	Backup Incremental Diferencial Level 3.	Job Name = backup_dif_incr_level_3_viern_1 Job Description = incremental diferencial viernes Start -> Later = viernes 12:00 pm Repeat -> Interval -> Frequency = 1 weeks
	Backup Incremental Acumulativo Level 1.	Job Name = backup_cum_incr_level_1_viern_2 Job Description = incremental diferencial viernes Start -> Later = viernes 12:00 am Repeat -> Interval -> Frequency = 1 weeks

En el siguiente gráfico se ilustra la estrategia de respaldo y recuperación que se implementó en la Base de Datos del Real Application Cluster.

Nivel de Backup.



Con HP OpenView Data Protector 6.0 se restaura cualquier elemento desde un único archivo hasta un servidor completo mediante una sencilla interfaz gráfica de usuario de fácil administración. Por tanto con la introducción HP OpenView Data Protector 6.0 se automatiza el proceso de restauración y realización de copias de seguridad. Con este software se realizan en un mismo instante copias de seguridad de los datos mientras se ejecutan las aplicaciones.

Conclusiones del Capítulo.

En el Capítulo se han abordado un grupo de aspectos esenciales en el proceso de desarrollo e implementación del proyecto Registro y Notarías.

Dirigimos la atención al diseño de una Arquitectura de Bases de Datos Descentralizada para facilitar la gestión y validación de los datos que se introducen en cada Oficina Registral de la República de Venezuela. Para la implementación de cada subsistema, la Base de Datos que funcionará en cada Registro Público y Mercantil del país han sido diseñadas según las necesidades impuestas por los procesos registrales y respetando en todo momento la relación coste-tiempo-beneficio. Para maximizar su rendimiento y proporcionar una infraestructura informática viable se empleará el hardware y software base adecuado capaz de cubrir los altos niveles de procesamiento resultado de un uso intensivo de los datos.

Como consecuencia de la implementación de un Sistema de Bases de Datos Descentralizado, en el documento se ha enfatizado en un grupo de mecanismos encargados de la gestión y replicación de la información en ambos sentidos (replicación bidireccional desde la Base de Datos Central hasta las Bases de Datos Locales de las Oficinas Públicas y Mercantiles). En relación a los métodos de replicación utilizados, se hace referencia a un conjunto de buenas prácticas de replicación que deben ser aplicadas para obtener un óptimo rendimiento.

Como complemento indispensable del Sistema Descentralizado que se pretende implementar se ha diseñado un Centro de Datos apropiadamente para proporcionar disponibilidad, accesibilidad, escalabilidad las 24 horas al día, los 7 días a la semana, y los 365 días al año descontando el tiempo fuera de servicio por mantenimiento.

Para proteger la información de los posibles errores producidos en la información que da soporte a todo el sistema, en el capítulo se han expuestos los procedimientos y pasos a seguir para realizar los backup del sistema como la única garantía que existe contra fallos del sistema operativo, del software o cualquier otro tipo de circunstancia.

CAPÍTULO III: ESTRATEGIAS DE OPTIMIZACIÓN Y ADMINISTRACIÓN DE BASES DE DATOS EN ORACLE DATABASE 10G RELEASE 2

Introducción.

En el presente capítulo se brinda una solución a los principales problemas para los administradores de bases de datos, específicamente en Oracle, en lo concerniente a la optimización y administración de Bases de Datos.

3.1 Optimización de Bases de Datos Oracle Database 10g release 2.

Optimizar el rendimiento y la disponibilidad de las Bases de Datos Oracle es fundamental para cumplir las expectativas de nivel de servicio y administrar el crecimiento exponencial de la información.

A continuación se abordan un grupo de aspectos orientados a aumentar la disponibilidad, simplificar el control y ajustar el rendimiento de las Bases de Datos Oracle 10g release 2.

3.1.1 Parámetros de Inicialización de la Base de Datos.

Una instancia de Base de Datos en Oracle 10g es configurada usando parámetros de inicialización, los que se especifican en el archivo de parámetros de inicialización. Estos parámetros influyen en el funcionamiento de la instancia de Base de Datos, incluyendo su rendimiento.

A continuación se muestran algunos parámetros de inicialización más relevantes y su descripción:

DB_BLOCK_SIZE: Especifica el tamaño de los bloques de la base de datos almacenados en los datafiles y cacheados en el Área Global del Sistema (SGA)[41]. El rango de valores depende del Sistema Operativo, pero es normalmente 8192 bytes para Sistemas de Procesamiento de Transacciones (OLTP) y mayor para sistemas de Bases de Datos tipo Warehouse o Almacén de Datos.

SGA_TARGET: Especifica el tamaño total de todos los componentes del SGA. Si el parámetro está especificado, entonces a todos los componentes del SGA se les asigna un tamaño automáticamente[41].

PROCESSES: Especifica el máximo número de procesos que pueden ser inicializados por la instancia[41]. Este es el parámetro más importante que inicializar, porque el valor de muchos parámetros es deducido de este.

UNDO_MANAGEMENT: Especifica cual modo de administración de espacio de undo el sistema debe usar. El modo AUTO es el más recomendado[41].

UNDO_TABLESPACE: Especifica el tablespace de undo que el sistema debe usar al iniciarse la instancia de Base de Datos[41].

3.1.2 Configuración y Diseño del Subsistema de I/O (Entrada/Salida).

El rendimiento de muchos sistemas de software está limitado por el I/O a disco. Oracle está diseñado de manera tal que si una aplicación está bien escrita, su rendimiento no debería ser limitado por I/O.

Los siguientes requerimientos de I/O deben ser considerados al Diseñar el Subsistema de I/O de una Base de Datos Oracle 10g.

- Almacenamiento: la capacidad mínima de discos.
- Accesibilidad: la continuidad (24x7) u horarios laborales solamente.
- Rendimiento: tiempo de respuesta de la aplicación.

A continuación se listan algunos parámetros de Oracle y del Sistema Operativo que se pueden usar para especificar el tamaño de I/O:

OS block size: determina el tamaño de I/O para operaciones de los archivos de red y otros.

SORT_AREA_SIZE: Determina el tamaño de I/O y la concurrencia para operaciones de orden.

HASH_AREA_SIZE: Determina el tamaño de I/O para operaciones de hash.

3.1.2.1 Tablas, Índices y Tablespaces Temporales.

Si los archivos con alta actividad de I/O son datafiles que pertenecen a tablespaces que contienen tablas e índices, entonces se recomienda identificar si el I/O de esos archivos puede ser reducido entonando las sentencias SQL o el código de la aplicación.

Si los archivos con alta actividad de I/O son datafiles que pertenecen al tablespace temporal, entonces se recomienda investigar si es mejor entonar las sentencias SQL ejecutando ordenamiento en disco para evitar esta actividad, o si entonar el ordenamiento.

3.1.3 Desarrollar Secuencias SQL Eficientes.

Reestructurando sentencias SQL.

Frecuentemente, reescribir una sentencia SQL ineficiente es más fácil que modificarla. Si se entiende el propósito de una sentencia dada, entonces es posible crear rápida y fácilmente una nueva sentencia que cumpla con dicho propósito.

A continuación se brindan algunos consejos para la creación de sentencias SQL eficientes:

- Se aconseja componer predicados usando (AND) y (=). Cuando en un predicado intervienen varias condiciones unidas por el operador AND, puede obtenerse un ligero aumento de rendimiento situando la condición más selectiva al principio. Esto se debe al orden en que el *parser* de Oracle analiza las sentencias: las condiciones unidas por el operador AND dejan de evaluarse cuando cualquiera de ellas no es cierta.
- Cuando se hace una consulta multi-tabla con joins, el orden en que se ponen las tablas en el FROM influye en el plan de ejecución. Aquellas tablas que retornan más filas deben ir en las primeras posiciones, mientras que las tablas con pocas filas deben situarse al final de la lista de tablas.
- Las consultas más utilizadas deben encapsularse en procedimientos almacenados. Esto es debido a que el procedimiento almacenado se compila y analiza una sola vez, mientras que una consulta

(o bloque PL/SQL) lanzado a la base de datos debe ser analizado, optimizado y compilado cada vez que se lanza.

- Utilizar siempre que sean posibles las mismas consultas. La segunda vez que se ejecuta una consulta, se ahorrará mucho tiempo de parsing y optimización, así que se debe intentar utilizar las mismas consultas repetidas veces.

Evitar usar columnas transformadas en una cláusula WHERE.

Ejemplo:

Usar

```
where a.order_no = b.order_no
```

En vez de:

```
where to_number (substr(a.order_no, instr(b.order_no, '.') - 1))  
  
= to_number (substr(a.order_no, instr(b.order_no, '.') - 1))
```

No usar funciones SQL en cláusulas de predicado o cláusulas WHERE. Cualquier expresión que use una columna, tal como una función que contiene una columna como su argumento, provoca que el optimizador ignore la posibilidad de usar un índice sobre esa columna, aun siendo un índice único, al menos que haya un índice basado en funciones definido que pueda ser usado.

Evitar los siguientes tipos de expresiones complejas:

- `col1 = nvl (:b1,col1)`
- `nvl (col1,-999) =`
- `to_date(), to_number(), etc`

Estas expresiones evitan que el optimizador asigne cardinalidad válida o estimados de selectividad y puede por el contrario afectar el plan general de ejecución de la sentencia.

Cuando se necesita usar funciones SQL en filtros o predicados join, no se debe usar en las columnas sobre las que se quiere usar un índice, en vez de esto, se puede usar en el lado opuesto del predicado.

Ejemplo:

```
to_char (numcol) = varcol
```

en vez de:

```
varcol = to_char (numcol)
```

3.1.4 Uso de Índices.

Los índices son estructuras opcionales asociadas a las tablas que permiten incrementar notablemente la velocidad de ejecución de las sentencias. En principio, las aplicaciones son independientes de los índices: es posible crearlos y borrarlos sin necesidad de reescribir el código y es el gestor el encargado del mantenimiento de los mismos. No obstante, atendiendo a consideraciones de rendimiento, es imprescindible conocer los índices definidos en cada tabla para escribir el código que haga un uso más eficiente de los mismos. La creación de un índice mejora sensiblemente las operaciones de recuperación de datos que puedan beneficiarse de la existencia del mismo. Sin embargo, las operaciones de actualización e inserción de datos se verán penalizadas por el hecho de tener que actualizar también el índice. Es por ello que se recomienda valorar cuidadosamente los posibles beneficios frente a los efectos negativos que implica la existencia del índice. En general, una tabla no debería tener más de dos o tres índices.

3.1.4.1 Elección de índices.

Los índices pueden crearse sobre un sólo campo o sobre varios, en cuyo caso se habla de índices concatenados. Ambos pueden o no ser únicos. Un índice único hace que el gestor no permita duplicidades de campo (o de la combinación de campos en los concatenados). Sobre una misma tabla pueden crearse más de un índice.

La elección de los índices es una de las decisiones más comprometidas del diseño técnico. Una columna será una buena candidata a ser indexada si:

- No existe duplicidad de valores (el mejor caso). Por ejemplo, el D.N.I.
- El campo toma un amplio rango de valores. Por ejemplo el código de cliente en una tabla de pedidos. Una mala elección sería el campo sexo que sólo toma dos valores.
- El campo es usualmente nulo y con frecuencia se seleccionan aquellas filas que tienen un valor.

Si el índice es concatenado, independientemente del orden de los campos en la tabla, la primera columna del índice será el campo más usado o el más selectivo.[42]

3.1.4.2 Uso eficiente de índices.

Oracle puede acceder a las tablas de dos modos:

- Secuencialmente (FULL SCAN).
- Usando índices.

El diseñador crea los índices en función del uso más frecuente de los datos, pero es el programador el que decide, en último término, si en la sentencia en que está trabajando hará que el gestor RDBMS emplee los índices o haga un acceso secuencial.

En general, el acceso por índice es preferible. Sin embargo, si se prevé que un alto porcentaje de los datos será consultado (digamos más de un 25%) un acceso secuencial puede ser más rápido.

3.1.4.3 Condiciones para el uso de índices.

El gestor RDBMS podrá usar un índice sólo si se cumplen estas dos condiciones:

C1: La columna es referenciada en un predicado.

C2: La columna indexada no está modificada por ninguna función u operador aritmético.

Que se cumplan ambas condiciones no implica necesariamente que el índice sea usado. El optimizador de ORACLE decidirá si es apropiado o no emplearlo.

Es importante resaltar que el orden de los campos en un predicado no es significativo para la utilización o no de un índice. Conociendo esta forma de actuar del optimizador de Oracle 10g, se deducen dos reglas para la creación de índices concatenados:

- Ordenar las columnas por su selectividad, eligiendo como primera columna la más selectiva para limitar el impacto de los usos parciales del índice (RANGE SCAN)
- Elegir como primera columna la más frecuente en los predicados.

Un compromiso entre ambas reglas determinará la creación del índice. La primera suele ser prioritaria dado que ya hemos visto que reescribiendo la sentencias adecuadamente es posible, en general, suprimir los efectos de la segunda regla.

3.1.4.4 Fragmentación de Índices.

Los índices en Oracle no son auto-balanceados. Se fragmentan luego de un número grande de operaciones de lenguaje de manipulación de datos (DML), tales como inserts y deletes, lo que puede provocar una significativa degradación de rendimiento.[43] A continuación se da una opción para detectar y reparar estos índices “fuera de forma”.

Extrayendo la información interna de un índice:

La vista USER_INDEXES contiene información estadística que se coloca ahí cuando se usa el comando ANALYZE INDEX.

Ejemplo:

```
analyze index CUSTOMER_LASTNAME_IND compute statistics;
```

Desafortunadamente, la vista USER_INDEXES está diseñada para proveer información al Optimizador Basado en Costo (CBO) de Oracle para la determinación del camino de ejecución SQL. No mantiene

estadísticas del estado interno de los índices de Oracle. Esto se puede lograr usando el comando ANALYZE INDEX con la cláusula VALIDATE STRUCTURE. Esto no cambia las estadísticas en la vista USER_INDEXES, pero en cambio puede ser ejecutado con seguridad sin afectar al CBO.

Esta es la sintaxis:

```
analyze index CUSTOMER_LASTNAME_IND validate structure;
```

Esto llena solamente la vista SYS.INDEX_STATS. Esta vista puede ser accedida con el sinónimo público INDEX_STATS. La vista contiene información de un índice a la vez, nunca tendrá más de una tupla. En consecuencia se debe consultar esta vista antes de analizar el próximo índice.

Para ver el contenido de la vista:

```
select name as IndexName, height, lf_rows, del_lf_rows from index_stats;
```

El resultado es (Ejemplo):

indexname	height	lf_rows	del_lf_rows
customer_lastname_ind	2	5237	2130

Las siguientes columnas de la vista son especialmente útiles:

- **height:** se refiere al máximo número de niveles encontrados en el índice.
- **lf_rows:** se refiere al número total de nodos hojas en el índice.
- **del_lf_rows:** se refiere al número de tuplas hojas que han sido marcadas como borradas, como resultado de una operación DELETE sobre una tabla.

Para determinar si un índice está fuera de balance y necesita ser reconstruido:

La primera regla es que todo índice con una altura (height) mayor a 3 es un buen candidato para ser reconstruido. Existen excepciones, pero generalmente, las tablas con altura mayor de 3 pueden provocar lecturas innecesarias de bloques de índices.[43]

La segunda regla es que el número de tuplas hojas borradas (del_if_rows) debe ser menor que el 20% del número total de nodos hojas (lf_rows). Un número alto de tuplas hojas borradas muestra que la tabla ha sido propensa a un número alto de operaciones DELETE. Como resultado, el índice se vuelve desbalanceado y es un buen candidato para ser reconstruido.

En este ejemplo, del_if_rows es 2130, lf_rows es 5237; el resultado es:

$$(2130*100) / 5237 = 40.7 \%$$

Este índice, simplemente está “rogando” ser reconstruido.

3.1.4.5 Reconstrucción de Índices.

Usando el comando ALTER INDEX REBUILD se logra usar el índice existente como fuente de datos para el nuevo índice en vez de usar la tabla como fuente de datos, mejorando así el rendimiento de la creación de índices.

La sintaxis quedaría así:

```
alter index <nombre_índice> rebuild;
```

Mientras se crea el Nuevo índice, este existirá simultáneamente con el viejo en la base de datos. Como consecuencia de esto, debe haber suficiente espacio libre para almacenar ambos en la base de datos. Cuando se reconstruye el índice, el nuevo queda disponible y el viejo es borrado. Se libera el espacio en la base de datos.

Existen varias opciones disponibles para usar este comando. La más usada es:

```
alter index <nombre_índice> rebuild parallel nologging compute statistics tablespace nombre_tablespace>;
```

La cláusula PARALLEL provoca que Oracle seleccione un grado de paralelismo igual al número de CPUs disponibles en todas las instancias participantes multiplicado por el valor del parámetro de inicialización PARALLEL_THREADS_PER_CPU.

La cláusula NOLOGGING permite reconstruir el índice más rápido porque las operaciones no serán escritas en el archivo de redo.

La cláusula COMPUTE STATISTICS permite recoger estadísticas a un costo relativamente bajo durante la reconstrucción del índice.

La cláusula TABLESPACE especifica el tablespace donde el índice reconstruido, la partición del índice, o la subpartición del índice serán almacenados. Por defecto es el tablespace donde el índice o partición residían antes de ser reconstruido. Especificando un tablespace diferente se puede mover fácilmente el índice a otra posición.[43]

Para más detalles sobre como detectar y reconstruir índices desbalanceados ver [Anexo 10](#).

3.2 Propuesta de herramienta para la administración de bases de datos Oracle 10g.

¿Que es el Grid Computing?

El término Grid Computing apareció a mediados de los años 90. La idea de Grid está enfocada fundamentalmente al acceso remoto a recursos computacionales y pretende ser un paradigma de desarrollo sin centrarse en una tecnología concreta.

El concepto de Grid Computing se enmarca dentro de la tecnología de computación distribuida englobando conceptos como sistemas operativos distribuidos, programación multiprocesador y redes de ordenadores, etc.

Entre los beneficios del Grid Computing en el marco de las tecnologías de la información podemos enumerar:

- Integrar sistemas y dispositivos heterogéneos. Grid Computing proporciona un conjunto de capacidades de integración horizontal que dirige de forma efectiva los recursos de toda la empresa, e incluso extienden la solución entre múltiples organizaciones.
- Mejora del coste efectivo de los entornos operativos. A través de la visualización de la consolidación, reservas, compartición y gestión de recursos a través de las funciones heterogéneas de tecnologías de la información, el Grid Computing ayuda a simplificar los entornos operativos y su gestión reduciendo la administración de su supervisión.
- El Grid Computing es capaz de descubrir dinámicamente y ajustarse a los entornos cambiantes y fluctuantes de las tecnologías de la información. Por esta razón, Grid Computing facilita el establecimiento, restablecimiento y cambios de los parámetros que requieren los negocios respecto a la seguridad y compartición de recursos.
- Incrementa la capacidad de recursos para responder a las fluctuaciones de la demanda, permitiendo a las organizaciones de tecnologías de la información agregar recursos distribuidos y explotar una capacidad no utilizada, los Grids incrementan de forma importante la cantidad de recursos computacionales y de datos disponibles.
- Aumenta la fiabilidad de la infraestructura sacando ventaja de los recursos del Grid como una alternativa ante la recuperación de los desastres tradicionales.

Oracle Grid Computing es una nueva arquitectura de computación designada para direccionar las necesidades de utilidad de computación. Un beneficio que presenta Oracle Grid Computing está relacionado con el ahorro y pronto retorno de inversiones, pues no se necesita de toda una nueva infraestructura para que funcione. Oracle Grid Computing opera bajo Unix, Windows y Linux este último especialmente recomendado para el Grid, además de que pueden funcionar aplicaciones que no estén basadas en tecnología Oracle.

Con la tecnología Oracle Grid Computing Oracle Database 10g, Oracle Application Server 10g y Oracle Enterprise Manager 10g con Grid Control, Oracle provee alta disponibilidad, escalabilidad y una alta capacidad administrativa.

3.2.1 Introducción a la Administración Automática. Oracle Enterprise Manager 10g Grid Control.

Históricamente, el coste de administración de grandes entornos crece linealmente, a veces, exponencialmente con cada nuevo sistema que se añade al entorno. Simplemente, no es posible mantener los costes de administración de grandes entornos con los métodos tradicionales de gestión no escalables.

Con Oracle Enterprise Manager 10g Grid Control release 2 se administran varias aplicaciones, servicios para el usuario final y toda la infraestructura grid de manera tan fácil como administrar una sola aplicación que se ejecuta en una sola computadora.

El principal valor añadido que introduce Oracle Enterprise Manager 10g Grid Control (OGC) es el de mantener constantes los costes de administración, aunque aumente el número de sistemas. OGC 10g está diseñado para simplificar las complejas tareas a las que se enfrentan los administradores de grandes entornos[44]. Dotado de robustas funcionalidades para administrar sistemas, tanto grandes como pequeños, OGC 10g automatiza las operaciones críticas para reducir el tiempo de implementación y el riesgo de errores, que aumentan a medida que crece el número de sistemas. Para mantener el servicio en un gran entorno, los administradores deben asegurarse constantemente de que el rendimiento y la disponibilidad de las aplicaciones se encuentran dentro de los niveles acordados. OGC 10g es la más completa solución de administración para entornos Oracle.

3.2.2 Características de Oracle Enterprise Manager 10g Grid Control.

Oracle Enterprise Manager 10g Grid Control proporciona una interfaz integrada para la administración y monitorización de aplicaciones y sistemas basados en tecnología Oracle.

Oracle Enterprise Manager 10g Grid Control incluye monitorización completa, gestión del rendimiento, bases de datos distribuidas y administración de servidor de aplicaciones, diagnóstico avanzado, tuning automático y una arquitectura que permite a los administradores:

- Administración y supervisión de las aplicaciones y sistemas basados en tecnología Oracle incluyendo la base de datos Oracle y Oracle Fusion Middleware.
- Interfaz 100% basado en web, fácil de desplegar y usar.
- Funciona con la aplicación J2EE Oracle Management Service que tiene un único repositorio y agentes distribuidos a lo largo de todos los sistemas administrados.
- Aprovisionamiento automático.

3.2.3 Requerimientos para el Despliegue de Oracle Enterprise Manager 10g Grid Control.

El proceso de instalación para Oracle Database Enterprise Manager 10g Grid Control se ha extendido para efectuar una evaluación de prerequisites para asegurarse de que el sistema cumpla con los requerimientos de OS, memoria y CPU[45].

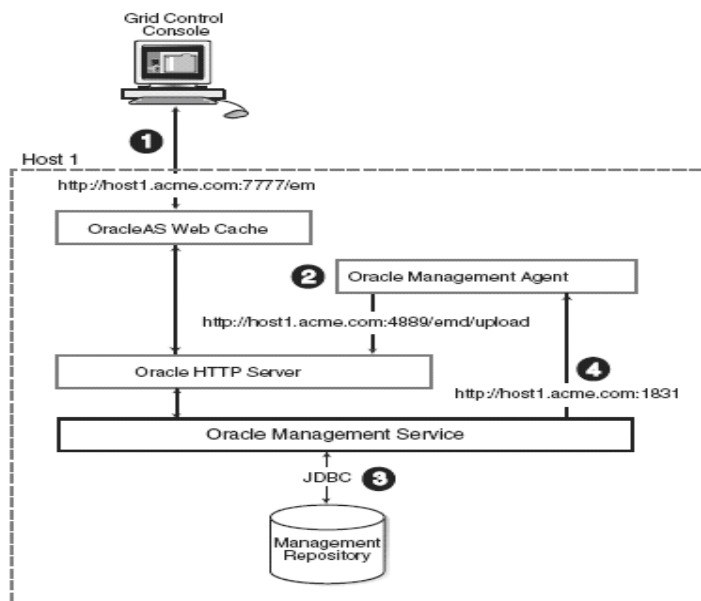
A continuación se muestra los requisitos necesarios:

Tamaño de Despliegue	Hosts	CPUs/Hosts	Memory/Host (GB)
Pequeño (100 objetos monitoreados)	1	1/3 (GHz)	2
Medio (1,000 objetos monitoreados)	1	2/3 (GHz)	2
Grande (10,000 objetos monitoreados)	2	2/3 (GHz)	2

En el [Anexo 11](#) se muestran los requerimientos mínimos de ancho de banda y latencia de red necesarios para lograr una correcta implementación de Oracle Enterprise Manager 10g Grid Control y sus componentes.

3.2.4 Oracle Enterprise Manager 10g Grid Control. Arquitectura y Componentes.

La figura muestra la interacción de cada uno de los componentes asociado a OGC. La configuración que se representa es resultado del proceso de instalación de Oracle Enterprise Manager 10g Grid Control usando una nueva Base de Datos.



Componentes de OGC:

- **Grid Control Console:** La Consola de Grid Control se emplea para monitorear y administrar los targets (host y sus componentes) que son descubiertos por el componente Oracle Management Agent instalado en cada host monitoreado[45].
- **Oracle Management Agent:** Oracle Management Agent se encarga de la captura y envío de información asociada a cada estación de trabajo o targets (hosts) monitoreado, incluido Oracle Management Service y Management Repository Database a través de Oracle HTTP Server[45].
- **Oracle Management Service:** Oracle Management Services (OMS) utiliza JDBC para establecer conexión con Oracle Repository Database, y obtener los datos que serán mostrados al usuario a través de Grid Control Console. OMS emplea Oracle Management Agent URL para monitorear la disponibilidad de Oracle Management Agent[45].

3.2.5 Opciones de la Consola Oracle Enterprise Manager 10g Grid Control.

Configuration Management Pack.

Oracle Configuration Pack permite a la familia Enterprise Manager la automatización de las tareas que más tiempo consumen y descubrir las causas de los fallos que normalmente son la mayor ocupación de los administradores.

Con Configuration Pack se logra:

- Disminuir los costes de manera significativa a la hora de administrar los sistemas.
- Asegurar la consistencia de los despliegues y cumplimiento de los estándares operativos.
- Obtener un software rápido y fiable para instalación, actualización, parcheo y clonación.
- Mejorar las decisiones sobre los despliegues de software. Proporciona información sobre los problemas de configuración.

Provisioning Pack.

Provisioning Pack automatiza los despliegues de software, aplicaciones y parches. Proporciona funcionalidad para el aprovisionamiento de sistemas operativos e imágenes de software, clonación de instalaciones existentes e imágenes de software (tales como CRS/RAC y AS), y parcheo.

Con Provisioning Pack se logra:

- Menor coste relacionado al despliegue de software y aplicaciones a través de estándares y automatización.
- Eliminar errores que normalmente afectan al proceso de despliegue de software.
- Eliminar problemas de seguridad y vulnerabilidades asociados con los parches.

Tuning Pack.

Oracle Tuning Pack proporciona servicios de automatización de procesos y optimización de aplicaciones, permitiendo reducir el coste a la vez que se mejora el rendimiento y la disponibilidad. Oracle Tuning Pack 10g ofrece un conjunto de tecnologías que automatizan el proceso completo de tuning de aplicación, así se reducen los costes de mantenimiento de la base de datos a la vez que se incrementa el rendimiento y fiabilidad.

Con Tuning Pack se logra:

- Obtener recomendaciones automáticas de tuning de sentencias SQL, índices y almacenamiento.
- Obtener consejos basados en el diagnóstico de rendimiento.
- Menor coste en la administración y mayor rendimiento de bases de datos.
- Una utilización de los recursos sensiblemente mejorada.

Diagnostics Pack.

Diagnostics Pack incluye funciones automáticas para la monitorización y diagnóstico integradas en el corazón del motor de base de datos.

Con Diagnostics Pack se logra:

- Menor coste de administración y monitorización identificación de problemas proactiva.
- Una resolución rápida de problemas.

3.2.6 Seguridad, Gestión y Administración de Oracle Enterprise 10g Grid Control.

Con el fin de reducir al mínimo los costos de gestión de los sistemas en un entorno Grid, Oracle ha creado Oracle Enterprise 10g Grid Control auto gestionable, automatizando muchas de las tareas de administración a bajo nivel que tradicionalmente consumen mucho tiempo a los administradores.

3.2.7 Administración y Monitorización de Sistemas con Oracle Enterprise 10g Grid Control.

Se deben utilizar las funciones del administrador para gestionar las cuentas de administración de OGC. Desde la consola de OGC se puede crear y gestionar nuevas cuentas de administración de OGC. Cada cuenta de administrador incluye sus propias credenciales de conexión, así como un juego de roles y privilegios que se asignan a la cuenta. Por defecto, solo los usuarios de la base de datos SYS, SYSTEM y SYSMAN se pueden conectar y utilizar la consola de la base de datos.

3.2.8 Monitorización de Recursos.

Los recursos, como CPU y memoria necesitan ser monitorizados de manera activa y precisa para eliminar la capacidad ociosa y hacer uso óptimo de los recursos existentes.

3.2.9 Gestión de los Métodos de Notificación.

Los métodos de notificación permiten globalmente diferentes mecanismos para enviar notificaciones. Estos métodos incluyen correo electrónico, interrupciones SNMP y archivos de comandos personalizados en ejecución. Una vez definidos, estos métodos se pueden utilizar con las reglas de notificación para enviar notificaciones a los administradores como resultado de las incidencias de alerta. Cada administrador tiene sus propias reglas de notificación.

Mediante la página Métodos de Notificación es posible configurar los servidores de correo saliente si se desea enviar notificaciones de correo electrónico mediante las reglas de notificación y gestionar archivos de comandos e interrupciones SNMP.

3.2.10 Configuración de Parches.

En la página Configuración de Parches puede especificar los parámetros que permiten utilizar las funciones de parches de OGC. Estas funciones simplifican el almacenamiento en zona intermedia y la aplicación de parches y juegos de parches de Oracle a cualquier host en el que esté ejecutando Oracle Manager Agent. También permiten recopilar información sobre asesores de parches críticos de Oracle.

3.2.11 Gestión de Interrupciones.

Las interrupciones permiten suspender el control de uno o más destinos para realizar operaciones de mantenimiento. Si no es súper usuario, solo podrá ver las interrupciones que ha creado en esta página.

3.2.12 Control y Gestión de Cambios.

Los administradores a veces deben averiguar por qué un sistema que antes funcionaba bien de repente no rinde a un nivel aceptable. Grid Control facilita esta tarea haciendo un seguimiento de todos los cambios en las instalaciones hardware y software y sus configuraciones. Esto permite que el administrador vea de manera rápida y sencilla los cambios que se han hecho desde la última vez en que el sistema funcionaba correctamente. Además, la capacidad flashback del servidor de aplicaciones permite a un administrador volver a una configuración previa en un punto en el tiempo para remediar el problema.

3.2.13 Gestión de Trabajos (Job Schedule).

Oracle Grid Control proporciona un configurador de trabajos (Job Schedule) que puede ejecutar tareas automáticamente en periodos específicos de tiempo.

3.2.14 Baselines y Administración Basada en Políticas.

Para reducir la necesidad de labores de administración Oracle Grid Control proporciona la habilidad de definir políticas de administración y establecer límites para esas políticas.

Por ejemplo, el administrador puede definir un límite de rendimiento y definir alertas que se activen si el rendimiento cae por debajo de ese límite. Esta configuración de límites puede usarse como estándar a la hora de crear nuevos sistemas.

3.2.15 Administración de la Seguridad.

La sección seguridad de la pantalla Administración de la Consola Oracle Enterprise Manager 10g Grid Control permite a los administradores autorizados dar de alta y borrar usuarios, roles y perfiles y conceder privilegios y roles a múltiples usuarios de la base de datos.

La sección seguridad tiene tres apartados:

- Usuarios: permite la administración de cuentas de usuario. Los usuarios además de crear, editar, ver, suprimir o realizar cualquier función de comando sobre un determinado usuario, se podrá crear un nuevo usuario tomando como base uno ya creado (Crear como). La página grupo de consumidores nos permite gestionar grupos de usuarios de forma cómoda fijando los privilegios de los mismos.
- Roles: permite la administración y asignación de privilegios a los roles utilizando las potencialidades del lenguaje SQL pero mucho más sencillo debido al uso de una interfaz gráfica.
- Perfiles: permite la administración de perfiles utilizando las posibilidades del lenguaje SQL pero mucho más sencillo debido al uso de una interfaz gráfica.

Lo más interesante de OGC es que evita el uso del lenguaje SQL pudiendo realizar las tareas de administración de seguridad mediante una intuitiva interfaz gráfica.

Por tanto, es incuestionable que Oracle Enterprise Manager 10g Grid Control ofrece una herramienta única de administración y monitorización de no solo el software de Oracle en su Grid, sino también sus aplicaciones Web y hosts. Grid Control está diseñado para administrar grises de aplicaciones donde los clientes pueden administrar muchas aplicaciones comerciales, servicios para el usuario final y toda la infraestructura grid de manera tan fácil como administrar una sola aplicación que se ejecuta en una sola computadora. Es indudable que OGC proporciona una mejor calidad de servicio al menor costo de administración y un mayor valor a través de un mayor alcance de administración.

Conclusiones del Capítulo.

El presente capítulo es una referencia a los principales problemas que presentan los administradores de Bases de Datos en Oracle Database 10g release 2. El documento se enfoca en los elementos esenciales para optimizar el rendimiento y disponibilidad de Bases de Datos Oracle 10g, y en consecuencia cumplir las expectativas de nivel de servicio exigidas por los clientes. Se hace énfasis en el empleo de Oracle Enterprise Manager 10g Grid Control, describiendo las robustas funcionalidades que posee para administrar sistemas, tanto grandes como pequeños automatizando las operaciones críticas para reducir el tiempo de implementación y el riesgo de errores.

CONCLUSIONES GENERALES

Con el desarrollo del trabajo, se considera que han sido cumplido todos los objetivos propuestos inicialmente. Por tanto se puede concluir expresando que:

1. Se describió la implementación de una Arquitectura de Bases de Datos Descentralizada en Oracle Database 10g release 2 como solución propuesta al Ministerio del Poder Popular para la Relaciones Interiores y Justicia con el objetivo de automatizar los Registros y Notarias de la República de Venezuela.
2. La utilización de un Modelo de Datos Relacional para el diseño de la Arquitectura de Bases de Datos Descentralizada que se implementó. Se detallan los elementos por los cuales se decidió la implementación de esta arquitectura, así como el modelo de datos que la sustenta.
3. Se hace referencia al diseño y construcción del Centro de Datos que sustentará la Arquitectura Descentralizada que se propuso; además de consolidar y respaldar la información correspondiente a los Registros Mercantiles y Públicos de la República de Venezuela.
4. Como DBMS encargado del funcionamiento de la Arquitectura Descentralizada se describió el empleo de Oracle Database 10g release 2. Con utilización de este DBMS se logró responder a las necesidades de negocio con alta disponibilidad y bajo costo.
5. Para proteger la información de posibles errores humanos se han descrito un conjunto de estrategias y pasos a seguir para realizar backup al sistema como la única garantía que existe contra fallos del Sistema Operativo, del Software o cualquier otro tipo de circunstancias.
6. Como principal herramienta administrativa se describió la utilización de Oracle Enterprise Manager 10g Grid Control. Se expusieron sus principales características y funcionalidades, que hacen de esta herramienta una excelente opción para la administración de la solución informática que se pretende desplegar en territorio de la República de Venezuela.
7. Para cumplir con lo planteado en el “Artículo #28 de la Ley del Registro Público y del Notariado”, de que toda la información que se gestione en cada Registro Público y Mercantil se replique hacia la Base de Datos Central se expusieron un grupo de novedosas técnicas y mecanismo empleados sobre Oracle Database 10g que permitieron compartir la información en tiempo entre Bases de Datos.

8. Para la protección y seguridad de la información que se encuentra en las Bases de Datos Locales y Central se detallaron un conjunto de estrategias y capacidades soportadas por el software de Oracle Database 10g release 2 para salvaguardar los datos sensitivos y críticos contra usuarios no autorizados.
9. En el documento se abordó como con la utilización de Oracle 10g se evita el sobredimensionamiento de recursos de CPU y almacenamiento, a partir de que no se requiere la necesidad de calcular la capacidad de los sistemas en función de los picos de trabajos. Además se reafirmó como Oracle 10g provee la infraestructura completa e integrada para potenciar el Grid Computing.

Finalmente, se considera que los temas abordados a lo largo del trabajo puedan contribuir al desarrollo futuro de los Sistemas Informáticos que se realicen en la Universidad de las Ciencias Informáticas y que utilicen como software Oracle Database 10g release 2. Además se pretende que el documento constituya una referencia para el Diseño e Implementación de Arquitecturas de Bases de Datos Descentralizadas en Oracle 10g.

RECOMENDACIONES

En el desarrollo del documento se ha profundizado en un conjunto de aspectos que se consideran estratégicos en el desarrollo del trabajo; a su vez otros temas solo han sido analizados brevemente a pesar de su marcada importancia. Por tanto este trabajo no queda exento de ello; por lo que a continuación se relacionan un conjunto de ideas que se consideran necesarias para darle continuidad a este trabajo:

1. Profundizar en los componentes de la arquitectura de Oracle 10g, y su relación con el sistema operativo en el que se implemente; además de conocer el funcionamiento básico de red propuesto por Oracle 10g.
2. Abordar los temas relacionados a la realización de ajustes en la Base de Datos Oracle 10g tan importantes en el rendimiento final de las aplicaciones; además de conocer los puntos críticos de una Base de Datos Oracle 10g y donde se pueden producir los problemas para adelantarse a ellos.
3. Profundizar en la creación y configuración de una Base de Datos en modo cluster. Pasos y técnicas para diseñar de forma adecuada la estructura de un RAC de alta disponibilidad, tanto en procesamiento de transacciones como en sistemas 24x7.
4. Diseño e implementación de una herramienta que se integre a Oracle Enterprise Manager 10g Grid Control que posibilite el monitoreo de los mecanismos de replicación de datos implementados. Con esta integración se garantiza una administración y monitoreo centralizado, y mucho más eficiente.
5. Perfeccionar los mecanismos de replicación implementados, basados principalmente en la minería de logs, así como en el estudio de otros posibles mecanismos de replicación compatibles con la versión Standard Edition One de Oracle que puedan ser más eficientes.

BIBLIOGRAFÍA

1. Base de Datos. Enciclopedia Libre. 2007 [Consultado el: 20 de marzo de 2007];
Disponible en:http://es.wikipedia.org/wiki/Base_de_datos.
2. Ortiz, A.M. Bases de Datos. Modelos de Datos. 2000 [Consultado el: 25 de marzo de 2007];
Disponible en:<http://elies.rediris.es/elies9/index-4.htm>.
3. Ortiz, A.M. El Enfoque Jerárquico. 2000 2007/03/20 [Consultado el: 20 de abril de 2007];
Disponible en:<http://elies.rediris.es/elies9/4-2-1.htm>.
4. Ortiz, A.M. El Enfoque de Red. 2000 [Consultado el: 27 de abril de 2007];
Disponible en:<http://elies.rediris.es/elies9/4-2-2.htm>.
5. Ortiz, A.M. El Enfoque Relacional. 2000 [Consultado el: 20 de marzo de 2007];
Disponible en:<http://elies.rediris.es/elies9/4-2-3.htm>.
6. M. Atkinson, F.B., D. DeWitt, K. Dittrich, D. Maier, and S. Zdonik, The Object-Oriented Database System. 1990.
7. Carballal, F.M. (2007) Bases de Datos Distribuídas.
8. Andrés, M.M.M. Diseño Conceptual. 2001 2007/03/27 [Consultado el: 25 de marzo de 2007];
Disponible en:<http://www3.uji.es/~mmarques/f47/apun/node69.html>.
9. Andrés, M.M.M. Diseño Lógico. 2001 [Consultado el: 28 de marzo de 2007];
Disponible en:<http://www3.uji.es/~mmarques/f47/apun/node70.html>.
10. Andrés, M.M.M. Diseño Físico. 2001 [Consultado el: 29 de marzo de 2007];
Disponible en:<http://www3.uji.es/~mmarques/f47/apun/node71.html>.
11. Solución para Centros de Datos Siemon 10G ip. 2007 [Consultado el: 07 de abril de 2007];
Disponible en:http://www.siemon.com/la/white_papers/SD-03-06-Centros-de-Datos.asp.
12. ¿Qué ventajas me aporta una SAN? 2007 [Consultado el: 07 de abril de 2007];
Disponible en: <http://www.smdata.com/ventajasSAN.htm>.
13. Mitchell, B. DMZ - Demilitarized Zone. 2007 [Consultado el: 02 de abril de 2007];
Disponible en:http://es.wikipedia.org/wiki/Zona_desmilitarizada.
14. Redes LAN y WAN, routers, switches, vlan, OSPF, Bridge Group, T1, E1, and BGP. 2001
[Consultado el: 15 de abril de 2007]; Disponible en:<http://es.wikipedia.org/wiki/VLAN>.

15. Oracle Database 10g Standard Edition. 2007 [Consultado el: 21 de abril de 2007];
Disponible en:http://www.oracle.com/database/Standard_Edition.html.
16. Oracle Database 10g Editions. 2007 [Consultado el: 21 de abril de 2007];
Disponible en:http://www.oracle.com/database/product_editions.html.
17. Oracle Database Lite 10g. 2007 [Consultado el: 20 de abril de 2007];
Disponible en:<http://www.oracle.com/technology/products/lite/index.html>.
18. Oracle TimesTen In-Memory Database. 2007 [Consultado el: 20 de abril de 2007];
Disponible en:<http://www.oracle.com/technology/products/timesten/index.html>.
19. Oracle Berkeley DB Product Family. 2007 [Consultado el: 19 de abril de 2007];
Disponible en:<http://www.oracle.com/technology/products/berkeley-db/index.html>.
20. Oracle Enterprise Manager 10g Grid Control. 2007 [Consultado el: 19 de abril de 2007];
Disponible en:<http://www.oracle.com/technology/products/oem/index.html>.
21. Oracle Application Server. 2007 [Consultado el: 19 de abril de 2007];
Disponible en:<http://www.oracle.com/appserver/index.html>.
22. Oracle Real Application Clusters. Oracle Magazine 2007. 2007 [Consultado el: 19 de abril de 2007]; Disponible en:http://www.oracle.com/database/rac_home.html.
23. Servidores ProLiant ML. HP ProLiant ML110 G3 2006 [Consultado el: 03 de abril de 2007];
Disponible en:<http://h10010.www1.hp.com/wwpc/es/es/sm/WF05a/781-783-380985-380985-12083403-12251632.html>.
24. 10 principales razones para utilizar Windows Small Business Server 2003. 2007 [Consultado el: 28 de abril de 2007]; Disponible en:<http://www.microsoft.com/latam/sbserver/evaluation/top.msp>.
25. Oracle Database 10g Standard Edition One. 2007 [Consultado el: 23 de abril de 2007];
Disponible en:http://www.oracle.com/lang/es/database/Std_One.html.
26. LogMiner. 2007 [Consultado el: 18 de abril de 2007];
Disponible en:<http://www.oracle.com/technology/deploy/availability/htdocs/LogMinerOverview.htm>.
27. Advanced Queuing. 1999 [Consultado el: 15 de mayo de 2007];
Disponible en:<http://lbd.epfl.ch/f/teaching/courses/oracle8i/server.815/a67781/c16queue.htm>.
28. Oracle Database 10g Standard Edition One. 2006 [Consultado el: 05 de mayo de 2007];
Disponible en:http://www.sinux.com.pe/solu_oracle01.htm.
29. Oracle Database Security Checklist. 2006.

30. Montoya, L.C. (2001) Seguridad de la Información.
31. Informática, J.V.D. (1998) Oracle Backup y Recuperación
32. Oracle Recovery Manager. 2007 [Consultado el: 04 de abril de 2007];
Disponible en:http://www.oracle.com/technology/deploy/availability/htdocs/rman_overview.htm.
33. HP ProLiant DL585 G2 Server series - overview. 2007 [Consultado el: 08 de abril de 2007];
Disponible en:<http://h10010.www1.hp.com/wwpc/us/en/en/WF05a/15351-15351-3328412-241644-3328422-3219717.html>.
34. HP ProLiant DL380 Server series - overview. 2007 [Consultado el: 08 de abril de 2007];
Disponible en:<http://h10010.www1.hp.com/wwpc/us/en/en/WF05a/15351-15351-3328412-241644-241475-1121516.html>.
35. HP StorageWorks 8000 Enterprise Virtual Arrays 2007 [Consultado el: 08 de abril de 2007];
Disponible en:<http://h18006.www1.hp.com/products/storageworks/eva/index.html>.
36. Oracle Database 10g Enterprise Edition. 2007 [Consultado el: 01 de abril de 2007];
Disponible en:http://www.oracle.com/database/Enterprise_Edition.html.
37. Oracle Real Application Clusters. 2007 [Consultado el: 15 de abril de 2007];
Disponible en:http://www.oracle.com/lang/es/database/rac_home.html.
38. Automatic Storage Management. 2007 [Consultado el: 12 de abril de 2007];
Disponible en:<http://www.oracle.com/technology/products/database/asm/index.html>.
39. Oracle Streams. 2007 [Consultado el: 10 de abril de 2007];
Disponible en:http://www.oracle.com/technology/products/dataint/htdocs/streams_fo.html.
40. HP Data Protector. 2006 [Consultado el: 02 de mayo de 2007];
Disponible en:<http://h18006.www1.hp.com/products/storage/software/dataprotector/index.html>.
41. Oracle® Database Performance Tuning Guide 10g Release 2. 2006.
42. al., B.L.e., Oracle RDBMS. Performance Tuning Guide. Version 6.0. 1990.
43. Detecting and Fixing Out-of-Balance Indexes. 2001 [Consultado el: 20 de mayo de 2007];
Disponible en:<http://www.devx.com/getHelpOn/10MinuteSolution/16596>.
44. Oracle Enterprise Manager 10g Release 2. 2005 [Consultado el: 10 de mayo de 2007];
Disponible en:http://www.oracle.com/lang/es/enterprise_manager/index.html.
45. Oracle® Enterprise Manager Advanced Configuration 10g Release 2 (10.2). 2005.