

**IMPLEMENTACIÓN DE UN CLÚSTER EXPERIMENTAL BAJO  
TECNOLOGÍAS LIBRES PARA PROPORCIONAR ALTA  
DISPONIBILIDAD DE SERVICIOS WEB HTTP**

**RONALD SAMIR ROMERO CARO  
JAVIER RAMÍREZ ANGULO**

**UNIVERSIDAD DE CARTAGENA  
FACULTAD DE INGENIERIA  
PROGRAMA INGENIERIA DE SISTEMAS  
CARTAGENA DE INDIAS D. T. y C.**

**2014**

**IMPLEMENTACIÓN DE UN CLÚSTER EXPERIMENTAL BAJO  
TECNOLOGÍAS LIBRES PARA PROPORCIONAR ALTA DISPONIBILIDAD DE  
SERVICIOS WEB HTTP**

**GRUPO DE INVESTIGACIÓN:  
e-soluciones**

**LINEA DE INVESTIGACIÓN:  
Redes de Computadoras.**

**DIRECTOR TRABAJO DE GRADO  
MSc. Martin Monroy Ríos  
Ingeniero de Sistemas  
Docente Universidad de Cartagena**

**INVESTIGADORES  
RONALD SAMIR ROMERO CARO  
JAVIER RAMIREZ ANGULO**



**UNIVERSIDAD DE CARTAGENA  
FACULTAD DE INGENIERIA  
PROGRAMA INGENIERIA DE SISTEMAS  
CARTAGENA DE INDIAS D. T. y C.  
2014**

**NOTA DE ACEPTACIÓN**

---

---

---

---

---

-----

***Firma del presidente del Jurado***

-----

***Firma del Jurado***

-----

***Firma del Jurado***

## Tabla de Contenido

1.	RESUMEN .....	1
	ABSTRACT.....	2
2.	INTRODUCCIÓN.....	3
2.1	Planteamiento del problema .....	5
2.2	Justificación del estudio .....	5
2.3	Importancia del estudio para el campo profesional .....	7
2.4	Contexto de la investigación .....	7
3.	MARCO TEÓRICO .....	9
3.1	COMPUTACIÓN DISTRIBUIDA .....	9
3.2	CLÚSTER .....	10
3.3	MIDDLEWARES.....	14
3.4	DISPONIBILIDAD INFORMÁTICA.....	19
3.5	SERVIDORES WEB.....	20
3.6	PROTOCOLO HTTP.....	21
3.7	<i>TECNOLOGÍAS LIBRES</i> .....	21
	LINUX.....	22
3.8	APACHE .....	25
3.9	MOODLE.....	26
4.	OBJETIVOS.....	28
4.1	Objetivo General .....	28
4.2	Objetivos Específicos.....	28
4.3	Alcance .....	29
5.	METODOLOGÍA .....	30

5.1	Etapa De Conceptualización.....	31
5.2	Etapa De Diseño .....	31
5.3	Etapa De Implementación.....	32
5.4	Etapa De Análisis.....	33
5.5	Etapa De Comunicación .....	33
6.	RESULTADOS .....	34
	ETAPA DE CONCEPTUALIZACIÓN: .....	34
	ETAPA DE DISEÑO: .....	43
	ETAPA DE IMPLEMENTACIÓN: .....	50
	ETAPA DE ANALISIS: .....	68
	Escenario 1: Fallas en el equipo en donde se encuentra corriendo los Servicios.....	71
	Escenario 2: Fallas en el equipo administrador de la máquina virtual – adición de NODO.72	
	Escenario 3: Fallas en el equipo Master del Pool de Servidores.....	72
	Escenario 4: Falla del soporte NFS o Repositorio Compartido. ....	73
7.	CONCLUSIONES Y RECOMENDACIONES.....	75
8.	BIBLIOGRAFÍA.....	79

## LISTA DE TABLAS

Tabla 1: Índice de Disponibilidad Anual de Sistemas Informáticos.....	20
Tabla 2: Descripción de Equipos computadores componentes del clúster.....	46

## LISTA DE ILUSTRACIONES

Ilustración 1. Etapas para el desarrollo metodológico del proyecto .....	30
Ilustración 2. Modelo de operación Oracle VM Server .....	41
Ilustración 3. Modelo de diseño para PC y red para el clúster HA.....	45
Ilustración 4: Conexiones lógicas y direcciones cluster HA.....	47
Ilustración 5: Conexiones físicas y máquinas hardware componentes del clúster HA.....	48
Ilustración 6: Plataforma de descarga Oracle.....	51
Ilustración 7: Pantalla inicial Instalación de Oracle VM.....	54
Ilustración 8: Selección de Instalación desde una nueva partición.....	55
Ilustración 9: Estructura de particiones de discos Estándar Oracle VM.....	55
Ilustración 10: Selección de Tarjeta de Red eth0.....	56
Ilustración 11 y 12: Configuración de parámetros de Red – Ip, Netmask, Gateway y DNS.....	57
Ilustración 13: Reinicio necesario para confirmar toda la configuración.....	57
Ilustración 14: Inicio y Login de Oracle VM Server.....	58
Ilustración 15: Roles de usuarios Oracle VM .....	62
Ilustración 16: Perfil de Usuario interfaz web de Oracle VM.....	62
Ilustración 17: Pantalla inicial – Creación Server Pool.....	63
Ilustración 18: Parametrización de Server Pool.....	63
Ilustración 19: Totalización y Resultados de creación Server Pool.....	64
Ilustración 20: Confirmación de Medios para creación Máquinas Virtuales.....	65
Ilustración 21: Instalación de Núcleo Linux – Virtual Machine.....	66
Ilustración 22: Servicio web – Servidor Apache corriendo en Máquina Virtual.....	67

## 1. RESUMEN

Coherente con su misión, el programa Ingeniería de Sistemas de la Universidad de Cartagena tiene como política el desarrollo de proyectos de aula, en los cuales los estudiantes tienen la oportunidad de aplicar las competencias adquiridas durante el estudio de su carrera profesional, dando solución a una necesidad de su comunidad.

En la actualidad la Universidad de Cartagena no cuenta con una plataforma informática integral que permita el despliegue de los proyectos ideados y realizados por los estudiantes del programa de Ingeniería de Sistemas, de tal forma que proporcione las herramientas o elementos necesarios para que la comunidad académica los tenga a su disponibilidad, pueda acceder a los servicios web planteados u ofrecidos en tales proyectos y, adicionalmente, haya una clara participación de ésta, en pro del mejoramiento continuo de dichos servicios; por tal razón, se plantea como objetivo la implementación de un clúster informático para proporcionar alta disponibilidad de servicios Web en los proyectos desarrollados en los cursos del programa utilizando tecnologías libres y sin incurrir grandes inversiones tecnológicas se facilita el proceso de retroalimentación inherente en todo sistema abierto, proceso que ayuda a la sostenibilidad de los sistemas en el tiempo.

Este proyecto fue realizado utilizando una metodología de investigación aplicada, basándose en las teorías y conceptos de computación distribuida consultadas en la literatura existente, lo que permitió llegar a la implementación de un clúster mediante la utilización de software libre y hardware con arquitectura x86 / x86-64 de fácil acceso en el mercado, permitiendo aumentar la disponibilidad de los servicios web prestados, sin incurrir en altos costos de servidores o software privados especializados en alta disponibilidad.



## ABSTRACT

It has always been a policy of Systems Engineering major at the University of Cartagena the use of classroom projects developed in academia in order to implement the concepts learned, the vast majority of these projects solve a real need in the community purpose for which they were created and that is consistent with the mission of the System Engineering major.

However, at present Cartagena University does not have a comprehensive software platform to enable the deployment of projects designed and conducted by students of Systems Engineering major, in a way that provide the tools and the information needed to allow the academic community to get them easily, can access web services posed or offered on such projects and, additionally, there is a clear involvement of this, towards the continuous improvement of such services, for this reason, therefore seeks the implementation of a computer cluster to provide high availability of web services projects within major's subject using open source technologies without incurring in large technological investments that facilitates the feedback process inherent at any open system , a process that helps in system sustainability time.

This project is carried out using a research methodology applied, based on the theories and concepts of distributed computing in the literature consulted and reach allowing the implementation of a cluster by using free software and hardware with x86 / x86 -64 easily accessible in the market, allowing to increase the availability of Web services rendered, without incurring in high costs of commercial software or expensive servers specialized in high availability.

## 2. INTRODUCCIÓN

Ante la necesidad de cumplir a cabalidad con las funciones sustantivas de docencia, investigación y proyección social para mejorar la calidad de la formación en el programa de Ingeniería de Sistemas, la Universidad de Cartagena tuvo la necesidad de desarrollar una plataforma que garantice una alta disponibilidad de servicios Web Http para el despliegue de los proyectos de todo tipo que los estudiantes desarrollan y, además, poder llevar el control de éstos por parte del cuerpo docente del Programa. Lo anterior debido a que en la actualidad cada proyecto se realiza en plataformas separadas que pueden no corresponder a escenarios reales de ejecución; por ello se requiere que dichos proyectos se desarrollen en escenarios que modelen la realidad y se facilite su seguimiento, beneficiando directamente el proceso de aprendizaje académico en la medida que el estudiante cuenta con ambientes de practica que se asemejan a los ambientes reales de producción realidad de

Con el desarrollo de este proyecto se pretende beneficiar a la comunidad académica del Programa de Ingeniería de Sistemas de la Universidad de Cartagena, presentando una solución que aunque no es del todo novedosa, se convierte en una herramienta que permite enriquecer el proceso de aprendizaje académico en la medida que el estudiante cuenta con ambientes que son acordes a los escenarios estándares de producción, a los que se enfrentara en su vida laboral; así como también los grupos de investigación del Programa contarán con la herramienta donde tendrán centralizado los proyectos desarrollados, lo cual aumenta la eficiencia en el seguimiento y monitorización de estos.

El objetivo de este proyecto fue implementar un clúster experimental bajo tecnologías libres para proporcionar alta disponibilidad de servicios Web Http que son desarrollados en los proyectos de aula del programa de Ingeniería de Sistemas, mediante la investigación de operaciones, los conceptos tecnológicos requeridos o asociados y la observación directa, para que la Universidad de Cartagena cuente con otra herramienta informática que le permita soportar de forma eficiente los procesos académicos.

Para lograr lo planteado se llevó a cabo una investigación de tipo aplicada, siguiendo cuatro etapas en su metodología: Diseño, Implementación, Análisis de resultados y Comunicación de resultados.

Como resultado de la presente investigación, se desarrolló el Clúster experimental para proyectos académicos con visión de aplicación práctica a futuro, que sean exigentes y de altos requerimientos funcionales y que a su vez. De igual manera, el Clúster ofrece servicios útiles y eficientes sobre las tecnologías de la información y la comunicación, puesto que este tipo de oferta es considerada como el principal objetivo del grupo de investigación E-Soluciones en la línea E-Servicios. Por todo lo planteado, en el proyecto se tuvo en cuenta aspectos conceptuales y teóricos como a) Computación distribuida; b) Tipos de Clúster; c) Disponibilidad Informática; d) Servidores WEB; e) Tecnologías Libres; f) Protocolo HTTP.

Para cumplir cabalmente lo anteriormente descrito, se proyectó unir los recursos de varias computadoras personales soportadas en un sistema operativo flexible y aplicaciones libres, con el fin de conformar una computadora de alto nivel físico y lógico, para que la atención

de peticiones Web de sus clientes se realice en cualquier instante, en el menor tiempo posible y garantizando bajos montos de inversión y costos de operación.

Es de resaltar que el esquema de Clustering usa como principal técnica la redundancia de equipos al conectar o unir varios servidores completos para que sea visto por los usuarios como uno solo, con lo cual se evitan situaciones presentadas en las arquitecturas centralizadas tradicionales en las cuales los errores internos se dan en un solo punto de fallo o (“single point of failure” – SOPF), del cual depende el servicio prestado.

## **2.1 Planteamiento del problema**

¿Cómo proporcionar alta disponibilidad para los servicios desarrollados en los proyectos web de los cursos del programa de Ingeniería de Sistemas de la Universidad de Cartagena, sin incurrir en altas inversiones en tecnología y altos costos asociados?

## **2.2 Justificación del estudio**

En la Universidad de Cartagena existe la necesidad de contar con una plataforma (clúster) que brinde alta disponibilidad y garantice el despliegue de los proyectos que se llevan a cabo en cada curso y son desarrollados por los estudiantes del programa de Ingeniería de Sistemas. Lo que implica beneficios como disponer de una plataforma para el despliegue de los sistemas software desarrollado, que faciliten el uso y seguimiento de dichos proyectos y, se trabaje en escenarios reales de ejecución.

En lo que respecta al ámbito académico, se espera que con la realización del clúster en mención se obtengan beneficios tales como: una formación con escenarios de práctica acordes con las necesidades de la Facultad, desarrollo de habilidades y competencias propias de la carrera, etc., porque se puede brindar la posibilidad de efectuar pruebas a nivel de laboratorio en asignaturas como Ingeniería de Servicios de Internet, Ingeniería de Software, Bases de Datos, entre otras; serviría de soporte a la infraestructura de los servicios prestados por los grupos de investigación, especialmente a los grupos e-Soluciones y Gimática. Esto evidencia los avances realizados por la Universidad para mejorar la gestión académica a realizar, dirigida a reforzar los procesos de registro calificado y acreditación de alta calidad.

En cuanto al proyecto, se considera importante para los investigadores y en general para toda la comunidad académica, dado que es una solución de alta disponibilidad que no existe actualmente en la oferta de servicios del Programa.

Para el programa de Ingeniería de Sistemas representa un proyecto innovador, porque, a pesar de que la tecnología no es nueva, hasta el momento no se ha presentado ninguna iniciativa orientada a proporcionar una solución de clúster que brinde alta disponibilidad para los servicios generados en los distintos escenarios académicos del Programa.

### **2.3 Importancia del estudio para el campo profesional**

La realización de este proyecto brinda una solución viable y factible a la implementación de soluciones de alta disponibilidad durante la prestación de servicios web, en este caso se tomó como objeto de estudio el programa de Ingeniería de Sistemas de la Universidad de Cartagena, pero su documentación hace posible su réplica en cualquier organización que desee implementar una solución a la alta disponibilidad sin incurrir en inversión de altos costos.

Con la puesta en marcha de este proyecto se evidencia el cumplimiento de los objetivos propuestos dentro del programa de Ingeniería de Sistemas, tales como: “Innovar procesos informáticos que presenten soluciones pertinentes a las necesidades de la región y el país, acordes con las tendencias mundiales de la informática, el conocimiento y las comunicaciones”; “Analizar, planear, diseñar e implementar sistemas de información que ayuden a la administración de los recursos informáticos de las organizaciones.”; entre otros.

### **2.4 Contexto de la investigación**

Este proyecto se realizó bajo los lineamientos de los grupos de investigación del programa de Ingeniería de Sistemas de la Universidad de Cartagena, siendo objeto de estudio los artículos y publicaciones más recientes con respecto al tema, teniendo en cuenta las necesidades existentes dentro del Programa.

Una vez analizada toda la documentación pertinente se procedió a la configuración del clúster, a analizar su comportamiento y documentar hallazgos y generar las conclusiones y recomendaciones.

### 3. MARCO TEÓRICO

Para la elaboración de este proyecto se hace necesario tener en cuenta diferentes conceptos que se desprenden de una importante rama de las ciencias computacionales, como lo es la computación distribuida. Dentro de ella veremos la importancia de los clúster y las diferentes tecnologías en que estos se soportan.

#### 3.1 COMPUTACIÓN DISTRIBUIDA

El uso de varios computadores interconectados por redes de comunicaciones dio lugar a la aparición hace varias décadas del concepto de sistema distribuido. Los Sistemas Distribuidos no son más que un conjunto de computadores que trabajan de forma conjunta para realizar una tarea. Los tres componentes principales que forman los sistemas distribuidos son [12]:

- **Computadores o Nodos:** Todo sistema distribuido está formado por más de un computador físico, al que también se le denomina sistema distribuido o simplemente nodo. Los nodos al menos tendrán CPU, memoria y dispositivos de entrada/salida que les permitan comunicarse con el entorno.
- **Red:** Alguno de los dispositivos de entrada/salida de los nodos serán periféricos de comunicaciones que permitirán interconectar a los nodos.



- Estado compartido: Los nodos de un sistema distribuido mantienen un estado compartido. Para mantener este estado de forma correcta, deberán coordinarse entre ellos.

Una de las características más importantes de un sistema distribuido, es su capacidad para tolerar fallos en algunos de sus componentes, a diferencia de lo que ocurre en un sistema centralizado, en un sistema distribuido habrá múltiples unidades de cómputo y de gestión de recursos que serán independientes (en cuanto a su probabilidad de fallo). Por tanto, la probabilidad de que falle el sistema en su totalidad será menor; no obstante se deberán emplear técnicas de alta disponibilidad, como la replicación de componentes, que permitirán que el sistema siga comportándose de acuerdo con sus especificaciones, incluso cuando alguno de sus elementos falle.

### **3.2 CLÚSTER**

Un Clúster es un grupo de equipos independientes que ejecutan una serie de aplicaciones de forma conjunta y aparecen ante clientes y aplicaciones como un solo sistema, propiedad que les permiten aumentar la escalabilidad, disponibilidad y fiabilidad de múltiples niveles de red [5].

- Escalabilidad: Se refiere a la capacidad que tiene un sistema de afrontar carga de trabajo cada vez mayor sin bajar su rendimiento en el medio donde se encuentra.

- Disponibilidad: Es el término que denota el que un sistema este accesible para sus usuarios cuando ellos así lo requieran.
- Fiabilidad: Denota la aptitud que posee un sistema en realizar su tarea sin incurrir en una alta tasa de error.

El término Clúster se aplica a los conjuntos o conglomerados de computadoras construidos mediante la utilización de componentes hardware que no necesariamente son comunes o homogéneos, pero se comportan como si fuese una sola máquina; y tuvo su origen en el verano de 1994 cuando Thomas Sterling y Don Becker conglomeraron 16 procesadores de computadores de sobremesa mediante una red Ethernet de 10 Mbps de manera exitosa. Esta Tecnología surge de diversas tendencias actuales y de la convergencia de ellas, entre las cuales encontramos las redes de alta velocidad y de nueva generación (NGN), el creciente auge y aceptación que han estado teniendo el desarrollo de las tecnologías libres, el desarrollo de herramientas software para la computación distribuida como es el caso de los Middleware, y la disponibilidad de microprocesadores con alta capacidad de procesamiento y a muy bajo costo según la ley de Moore<sup>1</sup>.

---

<sup>1</sup> Ley de Moore, "El número de transistores por unidad de área se duplicaría en 24 meses, al igual que la capacidad de los circuitos integrados". Gordon E. Moore – 1975.

## **TIPOS DE CLUSTERS**

Dependiendo del objetivo por el cual se realice la conglomeración de nodos el clúster podrá ser clasificado en una de las siguientes tres categorías:

- **Clúster de Alto Rendimiento:** Es un conjunto de ordenadores que está diseñado para dar altas prestaciones de servicio en cuanto a capacidad de cálculo. Son Clústeres en los cuales se ejecutan tareas que requieren de gran capacidad de procesamiento, grandes cantidades de memoria, o ambas. El llevar a cabo estas tareas puede comprometer los recursos del clúster por largos periodos de tiempo. Actualmente se libra una batalla por el título de la supercomputadora más poderosa del mundo el cual es ostentado por K Computer, ubicada en el RIKEN Advanced Institute for Computational Science (AICS)- Kobe, Japón y desarrollada por Fujitsu tiene una capacidad de computo de 16 Giga Flops.
- **Clúster de Alta Eficiencia:** Son Clústeres cuyo objetivo de diseño es el de proveer disponibilidad y confiabilidad. Estos tratan de brindar la máxima disponibilidad de los servicios que ofrecen. La confiabilidad se provee mediante software que detecta fallos y permite recuperarse frente a los mismos, mientras que en hardware se evita tener un único punto de fallos.
- **Clúster de Alta Disponibilidad:** Es un conjunto de dos o más máquinas que se caracterizan por mantener una serie de servicios compartidos y por estar constantemente monitorizándose entre sí, de tal forma que cuando una de las

maquinas falla, la otra toma su lugar en los servicios que la primera estaba prestando, brindando de manera transparente al usuario integridad en la información y fiabilidad en el servicio, quien no notara que hubo un fallo en el sistema [3].

En este caso lo que se busca no es exactamente conseguir una gran potencia de cálculo si no conseguir un conjunto de máquinas donde todas realicen la misma función y que, ante el fallo de cualquiera de ellas, las demás puedan asumir sus tareas de una forma transparente y rápida frente al usuario que requiere el servicio. De esta manera lo que se busca con la replicación de máquinas es evitar los puntos únicos de fallo o puntos críticos, que serían aquellas máquinas imprescindibles para el correcto funcionamiento del servicio que se quiere ofrecer; si únicamente tenemos una instancia de cada máquina de este tipo, se convierte en un punto único de fallo ó SPF por sus siglas en inglés (Single Point of Failure)[4] y ante cualquier fallo en este equipo, todo el servicio queda inutilizable.

La teoría sobre este tipo de Clústeres gira en torno a estos SPF y cómo evitarlos, mediante redundancia hardware y el software apropiado, para controlar el correcto funcionamiento de todos los equipos pertenecientes al Clúster y en caso en que una máquina falle hacer que las réplicas entren en funcionamiento de una manera transparente.

### 3.3 MIDDLEWARES

El termino middleware tuvo su primera aparición a finales de los años 1980's cuando se utilizó un software para administrar una conexión entre computadoras, pero fue hasta mediados de 1990 que tuvo su mayor auge con la extensión y consolidación de las redes informáticas. Hoy en día se define como un componente software diseñado para administrar la complejidad y la heterogeneidad inherente en los sistemas computacionales, brindando un control de acceso de manera uniforme a los recursos que se comparten en la computación distribuida. Cuando un conglomerado de computadores se hace ver como un solo recurso computacional, se puede decir que dicho conglomerado posee una sola imagen de sistema (SSI) (Buyya vol.1 1999) la cual es liada a la capa más expuesta del sistema operativo, situándose y sirviendo de interprete entre el sistema operativo y las aplicaciones del usuario.

Es factible hacer una clasificación de Middlewares dependiendo del nivel de abstracción de los mismos.

**3.3.1 Low level:** En un primer nivel podemos encontrar Middleware que se encuentran en la capa más cerca del lenguaje maquina proporcionado por el sistema operativo nativo.

**3.3.2 Mpi:** Message Passing Interface, define un estándar para las transacciones de datos dentro de la interactividad de los procesos en sistema distribuido. No es un lenguaje de programación ni una librería a implementar. Su principal función es

definir cómo deben aparecer las funciones básicas para el intercambio de mensajes. Por medio de éste se definen nombre, secuencias de llamado y resultados. Algunas de estas funciones son definidas punto a punto dentro de los procesos de comunicación.

**3.3.3 Open mpi:** Open MPI es una implementación de código abierto de las versiones 1.2 y 2 estándar de MPI. Su objetivo primordial es la creación de alta eficiencia logrando un equilibrio entre los aspectos de producción y calidad de la librería MPI desarrollada para y por la comunidad HPC.

**3.3.4 Pvm:** Es un conjunto que integra herramientas y librerías que forman un framework para construir una sola maquina a partir de la colección de computadores heterogéneos.

**3.3.5 Parallel file system:** Los ambientes de programación paralela permiten implementar algoritmos que hagan uso de recursos compartidos: CPU, memoria, datos y servicios.

**3.3.6 Mpi-io:** está diseñado para trabajar con conjuntos de datos distribuidos de manera similar a como se realiza el intercambio de mensajes en MPI nativos. La escritura de archivos es similar a enviar mensajes de MPI así como también, la lectura es similar a la recepción de mensajes MPI. Todas las acciones básicas de manipulación de archivos son compatibles es decir, abrir, cerrar, eliminar

y cambiar el tamaño de los archivos. MPI-IO incorpora su propio archivo de representación de estrategia como un conjunto de unidades Etype (tipo de datos elementarios). Etypes de control de acceso a archivos y posicionamiento. Además, las vistas son otro concepto que define qué partes de un archivo son visibles a una tarea y cómo realizar una tarea debe interpretarlos.

MPI-IO implementa diferentes formas de acceder a los datos que introduce flexibilidad en el proceso de E / S. Acceso a los datos se caracteriza por la posición, la sincronización y coordinación. Estas tareas de definir si deben compartir un apuntador de archivo común de manera colectiva o no, si un proceso se bloquea hasta que los datos se escriben, se leen o ninguna de ambas opciones, si el acceso es coordinado o "libre para todos".

**3.3.7 Hadoop:** Apache Hadoop es una plataforma de software que crea un sistema de archivos distribuidos en los discos en un clúster. Está diseñado para unificar los recursos de almacenamiento de grandes grupos de productos de hardware integrado. Basado en Java que permite el desarrollo de aplicaciones que procesan grandes conjuntos de datos. La manipulación de datos es completamente transparente para el usuario y, además, múltiples copias de datos se mantienen en los nodos. Esto proporciona a las aplicaciones la confiabilidad y rutas de acceso a los datos más cortas, dando lugar a un mayor ancho de banda.

**3.3.8 Beowulf:** este tipo de middleware es propicio para crear clúster que procuran lograr alto rendimiento a muy bajo presupuesto. Este tipo de clúster tienen rendimientos comparables con los obtenidos por las supercomputadoras pero a un costo sumamente inferior, siendo esta su principal característica.

Muestra una gran flexibilidad frente a la gran heterogeneidad que se presentan en los componentes de hardware, proporcionando la posibilidad de configurar, reestructurar y optimizar el sistema para ejecutar una aplicación cuando sea necesario. Los nodos en este tipo de clúster son dedicados, lo cual significa que su único propósito es trabajar sobre la aplicación sobre la cual trabaja el clúster y pueden ser configurados físicamente solo con procesador, memoria ram y lógicamente la conexión de red, no son necesario periféricos de entrada y salida incluso unidades de almacenamiento pueden ser omitidas.

**3.3.9 Oscar:** Open Source Cluster Application Resources, es un paquete de software que simplifica el proceso de configuración de un clúster de alto rendimiento, incluyendo todo lo que se puede necesitar para el montaje de uno de estos.

Oscar fue concebido con la idea de llevar la instalación del clúster hasta un estándar unificado, debido a eso dentro de su paquete de instalación se encuentran los software más comúnmente utilizados, liberando a usuario del proceso de descarga, instalación y configuración de paquetes

**3.3.10 Heartbeat:** Como un paquete de software heartbeat trabaja enviando pulsaciones (ping), los cuales verifican si el Server principal está activo o no;



estos pings enviados por heartbeat requieren una respuesta por parte del servidor principal o master, si al cabo de un quantum el servidor no responde dichos ping, heartbeat determina que ese servidor se encuentra inactivo, y automáticamente activa el servidor secundario para que asuma el control de la red y responda las peticiones por parte del cliente.

**3.3.11 Condor:** Es un sistema de software escalable que propicia un ambiente de HTP (High-Throughput Computing) para el cual utiliza una vasta colección de recursos computacionales casi siempre de propiedad distribuida. A diferencia de High Performance Computing donde se produce una alta capacidad de cómputo realizando procesos en corto periodos de tiempo, la HTC se basa en la necesidad de extender la capacidad de cómputo por largos periodos de tiempo.

Condor es un Middleware multifunciones que trabaja por lotes distribuyendo la carga de trabajo en cada uno de sus lotes adscritos. El proyecto condor fue concebido como el software planificador del intento de reconstrucción del genoma humano.

**3.3.12 Globus toolkit:** Globus se basa principalmente en los servicios web para definir las interfaces y la estructura de sus componentes [Foster, 2005]. Por ejemplo, el uso de servicios web basados en XML para describir los mecanismos, descubrir o invocar los servicios de red. Es más, estos protocolos orientados a documentos son muy adecuados para los cálculos de acoplamiento flexible, que son las preferidas en los sistemas distribuidos. Globus Toolkit utiliza los servicios Web

para la mayoría de sus componentes principales. La Asignación de Recursos Grid y Servicios de Gestión (GRAM) implementa las interfaces para el manejo de los elementos de cómputo, servicio fiable de transferencia de archivos (RFT) gestiona las transferencias de datos.

### **3.4 DISPONIBILIDAD INFORMÁTICA**

Se refiere a la continuidad de acceso a los elementos de información almacenados y procesados en un sistema informático, es decir, que la información pueda ser recuperada en el momento que se necesite, esto es, evitar su pérdida o bloqueo, bien sea por ataque doloso, mala operación accidental, situaciones fortuitas o de fuerza mayor. Su objetivo es garantizar el acceso a un servicio o a los recursos.

En cuanto al índice cuantitativo de disponibilidad, se refiere a un porcentaje con el cual se mide el tiempo que se encuentra disponible un sistema y equivale al tiempo disponible sobre el tiempo total; en la siguiente tabla se clasifica teniendo en cuenta el tiempo de inactividad durante un año del sistema [1]:

<b>INDICE DE DISPONIBILIDAD</b>	<b>INACTIVIDAD ANUAL</b>
<b>97%</b>	11 días
<b>98%</b>	7 días
<b>99%</b>	3 días y 15 minutos
<b>99.9%</b>	8 horas y 48 minutos
<b>99.99%</b>	53 minutos
<b>99.999%</b>	5 minutos
<b>99.9999%</b>	32 segundos

Tabla 1: Índice de Disponibilidad Anual de Sistemas Informáticos.

### 3.5 SERVIDORES WEB

Un servidor web es un programa que atiende y responde a las diversas peticiones de los navegadores, proporcionándoles los recursos que solicitan mediante el protocolo HTTP o el protocolo HTTPS (la versión segura, cifrada y autenticada de HTTP) [2]. Un servidor web básico tiene un esquema de funcionamiento muy sencillo, ejecutando de forma infinita el bucle siguiente:

- Esperar peticiones en el puerto TCP (Por defecto es el puerto 80).
- Recibe una petición.
- Busca el recurso de la cadena de petición. (path embebido).
- Enviar la respuesta de la petición por el mismo canal en el cual la recibió.

### **3.6 PROTOCOLO HTTP**

El protocolo HTTP (*Hypertext Transfer Protocol*) es el protocolo base de la Internet. Se trata de un protocolo simple, orientado a conexión y sin estado. La razón de que esté orientado a conexión es que emplea para su funcionamiento un protocolo de comunicaciones TCP<sup>2</sup>, un protocolo que establece un canal de comunicaciones de extremo a extremo, cliente – servidor, por el que pasa el flujo de bytes que constituyen los datos que se están transfiriendo. El protocolo no mantiene estado, en el sentido de que cada transferencia de datos es una conexión independiente de la anterior.

Hoy en día existe y va en aumento una variante de HTTP llamada HTTPS (S por *secure*) que utiliza el protocolo de seguridad SSL (*secure socket layer*) para cifrar y autenticar el tráfico entre cliente y servidor, siendo ésta muy usada por los servidores web de comercio electrónico, así como por aquellos que contienen información personal o confidencial [8][14].

### **3.7 TECNOLOGÍAS LIBRES**

Son tecnologías que permiten su libre reutilización, teniendo en cuenta que los productos o servicios generados con ellas no tienen necesariamente por qué ser gratuitos. Dentro de algunas tecnologías libres que se pretenden utilizar se tienen:

---

<sup>2</sup>*Transport Control Protocol*

## *LINUX*

Parte de la comunidad y numerosos medios prefieren denominar a esta combinación como Linux, aunque GNU/Linux (con las variantes *GNU con Linux* y *GNU+Linux*) es la denominación defendida por el Proyecto GNU y la FSF<sup>3</sup> junto con otros desarrolladores y usuarios para el conjunto que utiliza el sistema operativo Linux en conjunto con las aplicaciones de sistema creadas por el proyecto GNU y por muchos otros proyectos de software.

Desde 1984, Richard Stallman y muchos voluntarios están intentando crear un sistema operativo *libre* con un funcionamiento similar al UNIX, recreando todos los componentes necesarios para tener un sistema operativo funcional. A comienzos de los años 90, unos seis años desde el inicio del proyecto, GNU tenía muchas herramientas importantes listas, tales como editores de texto, compiladores, depuradores, intérpretes de comandos de órdenes etc., excepto por el componente central: el núcleo.

GNU tuvo su propio proyecto de núcleo, llamado Hurd. Sin embargo, su desarrollo no continuó como se esperaba al aparecer el núcleo Linux. De esta forma se completaron los requisitos mínimos y surgió el sistema operativo GNU que utilizaba el núcleo Linux.

El principal argumento de los defensores de la denominación GNU/Linux es resolver la posible confusión que se puede dar entre el núcleo (Linux) y gran parte de las herramientas básicas del resto del sistema operativo (GNU). Además, también se espera que, con el uso del nombre GNU, se dé al proyecto GNU el reconocimiento por haber creado las

---

<sup>3</sup> *Free Software Foundation*

herramientas de sistema imprescindibles para ser un sistema operativo compatible con UNIX, y se destaque la cualidad de estar compuesto sólo por software libre. La primera distribución que incluyó el GNU en su nombre fue en 1992, donde aparecía como *Linux/GNU/X*. La FSF denominó a este sistema "Linux" hasta al menos junio de 1994 y recién a partir de enero de 1995 empezó a llamarlo "GNU/Linux" (también GNU+Linux y *lignux*, términos que han caído en desuso a instancias del propio Stallman) [2].

Algunas de las características técnicas que hacen al sistema operativo Linux apto para trabajos sobre tecnologías de Clustering encontramos: Multitarea, Multiplataforma, Carga de ejecutables por demanda, política de copia en escritura para la compartición de páginas entre ejecutables (varios procesos pueden utilizar la misma zona de memoria para ejecutarse), memoria virtual usando paginación a discos, gestión de memoria como recurso unificado, librerías con carga dinámica y/o estática, consolas virtuales múltiples, soporte para varios sistemas de archivos comunes como Minix-1 y Xenix, y software cliente y servidor NetWare disponible en los núcleos de desarrollo. Todas las características anteriormente destacadas hacen que el sistema OS de Linux sea óptimo para prestación de servicios de gran escala (disponibles, alto rendimiento o eficiencia).

Dentro de las más populares versiones del sistema operativo Linux se encuentran:

- Debian: Siendo una de las más antiguas distribuciones de Linux, *Debian* fue iniciada en 1993 por IanMurdock. Se caracteriza por ser una de las pocas versiones en la que no hay ninguna compañía detrás, la distribución está controlada totalmente por voluntarios vinculados por el contrato social *Debian*. Un líder es elegido cada año desde y por los miembros del proyecto *Debian*.

Se considera la madre de la familia Debian la cual maneja los archivos .deb, los cuales son soportados en su totalidad por la comunidad de código abierto y se basan en dos aspectos claves la estabilidad y la seguridad, criterios con los cuales los nuevos paquetes son probados por un lapso de 2 meses antes de su liberación; la desventaja de este proceso es la obsolescencia que puedan presentar los paquetes en el proceso de optimizar la estabilidad y los fallos de seguridad durante el ciclo de liberación.

- Ubuntu: La distribución que más auge a tenido en los últimos años, con cada vez más usuarios y que más rápido se ha adaptado a las necesidades de los mismos. *Ubuntu* es una distribución basada en *Debian*, y gestionado por *Canonical*. Aunque gran parte del énfasis de la distribución está en el escritorio, también hay una versión de servidor, y es que Canonical ha atraído el apoyo de los agentes comerciales más tradicionales de UNIX, como *Oracle*.
- Red Hat Enterprise: *Red Hat Enterprise Linux (RHEL)* es probablemente la distribución de Linux más conocida y muy popular en cuanto a servidores, además de ser uno de los más veteranos. Ha contribuido a un gran número de aplicaciones para la comunidad Open Source de los años, incluyendo *Red Hat GFS* y su sistema de archivos en clúster. Aunque no es obligatorio, el acceso a soporte y actualizaciones de seguridad requiere que los clientes paguen un honorario por estos “derechos”, al igual que el caso de *SuSe Linux*.

- CentOS: CentOS (abreviatura de Community Enterprise OperatingSystem) es una versión libre disposición de Red Hat Enterprise que no cobra por el acceso a actualizaciones de seguridad. Esto es posible debido a la licencia libre bajo la cual se libera Red Hat. Dado que es una copia casi exacta de RHEL, con únicamente los logotipos y marcas registradas modificadas (y sin el soporte comercial), los binarios son 100% compatibles, es decir, las aplicaciones diseñadas para aplicaciones comerciales de Red Hat se ejecutarán sin modificaciones y con total compatibilidad.
- SuSE Enterprise: *SuSE Linux Enterprise Server (SLES)* es una distribución que toma prestado de *Red Hat* su gestión de paquetes, distribución y su modelo de negocio. Fue creada originalmente en Alemania por un grupo de consultores de *UNIX*, *SuSE* significa “*Software- und System-Entwicklung*” (Desarrollo de software y de sistemas). Y en 2003 fue adquirida por *Novell*, que hicieron de esta una distribución en base a sus ofertas.

### **3.8 APACHE**

En febrero de 1995, el servidor web más popular de Internet era un servidor de dominio público desarrollado en el NCSA (National Center for Supercomputing Applications en la Universidad de Illinois). No obstante, al dejar Rob McCool (el principal desarrollador del servidor) la NCSA en 1994, la evolución de dicho programa había quedado prácticamente



reducida a la nada. El desarrollo pasó a manos de los responsables de sitios web, que progresivamente introdujeron mejoras en sus servidores. Un grupo de éstos, usando el correo electrónico como herramienta básica de coordinación, se puso de acuerdo en poner en común estas mejoras (en forma de “parches”, *patches*). Dos de estos desarrolladores, Brian Behlendorf y Cliff Skolnick, pusieron en marcha una lista de correo, un espacio de información compartida y un servidor en California donde los desarrolladores principales pudiesen trabajar. A principios del año siguiente, ocho programadores formaron lo que sería el Grupo Apache.

De esta manera nace lo que hoy es un servidor web de código libre y robusto, con una arquitectura modular, posee un core<sup>4</sup> al cual se le adicionan módulos funcionales a manera de filtrado de contenido [13], al ser multiplataforma se convierte en el servidor de servicios del protocolo http más utilizado en la internet proporcionado y soportado por un grupo de voluntarios en todo el mundo [6][13].

### **3.9 MOODLE**

Moodle es una aplicación web de tipo Ambiente Educativo Virtual, un sistema de gestión de cursos, de distribución libre, que ayuda a los educadores a crear comunidades de aprendizaje en línea. Este tipo de plataformas tecnológicas también se conoce como LCMS (Learning Content Management System). La versión más reciente es la 2.6

---

<sup>4</sup>Core o Núcleo funcional

Moodle fue creado por Martin Dougiamas, quien fue administrador de WebCT en la Universidad Tecnológica de Curtin. Basó su diseño en las ideas del constructivismo en pedagogía que afirman que el conocimiento se construye en la mente del estudiante en lugar de ser transmitido sin cambios a partir de libros o enseñanzas y en el aprendizaje colaborativo. Un profesor que opera desde este punto de vista crea un ambiente centrado en el estudiante que le ayuda a construir ese conocimiento con base en sus habilidades y conocimientos propios en lugar de simplemente publicar y transmitir la información que se considera que los estudiantes deben conocer.

La primera versión de la herramienta apareció el 20 de agosto de 2002, a partir de allí han aparecido nuevas versiones de forma regular. Hasta julio de 2008, la base de usuarios registrados incluye más de 21 millones, distribuidos en 46.000 sitios en todo el mundo y está traducido a alrededor de 91 idiomas.<sup>1</sup>

Estas herramientas son de gran utilidad en el ámbito educativo, ya que permiten a los profesores la gestión de cursos virtuales para sus alumnos (educación a distancia o e-learning), o la utilización de un espacio en línea que dé apoyo a la presencialidad (aprendizaje semipresencial, blended learning o b-learning).

## **4. OBJETIVOS**

### **4.1 Objetivo General**

Implementar un clúster informático para proporcionar alta disponibilidad de servicios Web en los proyectos desarrollados en los cursos del programa de Ingeniería de Sistemas de la Universidad de Cartagena, utilizando tecnologías libres.

### **4.2 Objetivos Específicos**

1. Determinar los requerimientos funcionales del Clúster mediante el análisis la situación actual de la infraestructura tecnológica que tiene el programa de Ingeniería de Sistemas de la Universidad de Cartagena.
2. Diseñar los componentes del clúster a partir de las tecnologías y protocolos existentes, que garanticen seguridad y alta disponibilidad de servicios web en el contexto académico.
3. Implementar la infraestructura tecnológica del clúster a partir del diseño planteado, para optimizar la disponibilidad de los servicios web.
4. Validar el funcionamiento del clúster implementado, aplicando un caso de estudio.

### **4.3 Alcance**

El proyecto se desarrolla exclusivamente para el contexto del programa de Ingeniería de Sistemas de la Universidad de Cartagena, a partir de los recursos aptos y disponibles en los laboratorios del Programa para el montaje de la plataforma de alta disponibilidad. La implementación del clúster se realiza únicamente con tecnologías de código abierto. La sostenibilidad y mantenimiento del clúster no hacen parte del alcance del proyecto.

## 5. METODOLOGÍA

El proyecto se enmarca en una investigación aplicada de tipo cualitativa, puesto que se implementó un clúster informático o sistema conglomerado de computadoras, basándose en la literatura existente, los casos de éxitos encontrados y la descripción de los fenómenos acontecidos durante su implementación, a la vez que no se realizó recolección de datos cuantitativos. En la siguiente ilustración se describen cada una de las etapas de la metodología utilizada (Ilustración 1).

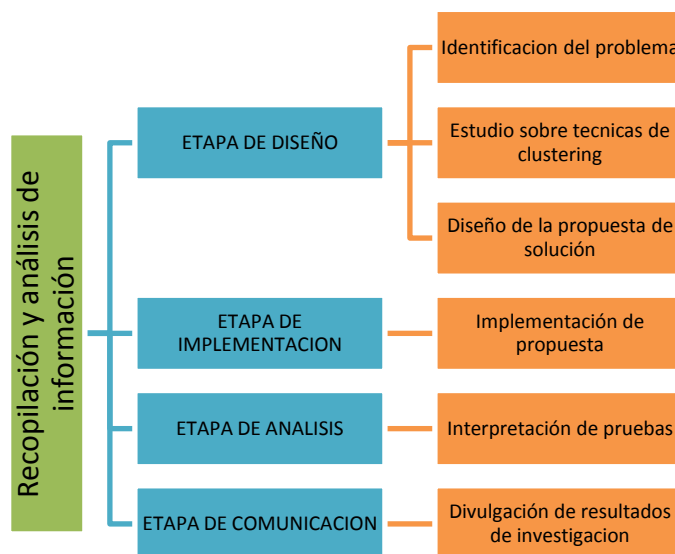


Ilustración 1: Etapas para el desarrollo metodológico del proyecto

La investigación se inició con la etapa de conceptualización comprendiendo una recopilación y análisis de la información que permitió la retroalimentación con las otras etapas consideradas en el proceso de investigación, tales como: la etapa de implementación del diseño, etapa de análisis e interpretación de resultados, para finalizar en la etapa de comunicación de resultados.

A continuación se destacan cada una de las actividades desarrolladas en las etapas del presente proyecto.

### **5.1 Etapa De Conceptualización**

En el proceso de conceptualización se buscó lucrar de información cada una de las etapas derivadas de ésta (diseño, implementación, análisis y comunicación), de modo que las bases teóricas para su desarrollo estén bien estructuradas.

#### Actividades:

- Se estudió variedad de bibliografías que traten sobre la implementación de servicios altamente disponibles con tecnologías libres.
- Se revisaron y confrontaron distintas técnicas para el desarrollo de servicios en red basados en clúster que existen en la actualidad y que permiten su utilización en servicios web.

#### Impacto o resultado obtenido:

- Una base teórica bien conformada y estructurada para la prestación de servicios web de alta disponibilidad.

### **5.2 Etapa De Diseño**

#### Actividades:

- Se identificaron las posibles falencias sobre los sistemas de servicios web en cuanto a su disponibilidad.

- Se revisaron las alternativas de solución para servidores web bajo la técnica de clustering HA<sup>5</sup>.
- Se diseñó de una arquitectura de clúster viable de acuerdo a las necesidades de disponibilidad del servicio prestado.
- Se eligió y analizó las tecnologías físicas y lógicas a utilizar en la implementación del clúster.

Resultados obtenidos:

- Documento sobre el procedimiento de integración de tecnologías hardware y software para brindar una alta disponibilidad en la prestación de servicios web; además de identificar la relación costo-beneficio con la utilización del clúster HA.

### **5.3 Etapa De Implementación**

Actividades:

- Se implementó y configuró el sistema de alta disponibilidad, con un clúster, integrando tecnologías hardware y software con licenciamiento libre, a fin de poner a prueba el diseño y la arquitectura planteada.

Impacto obtenido:

- Un servidor de servicios web funcional, con alto grado de disponibilidad.

---

<sup>5</sup>Clustering High Availability (Clúster de alta disponibilidad)

## **5.4 Etapa De Análisis**

Dentro de esta etapa se puso a prueba tanto los planos de diseño de máquinas y software a utilizar en la etapa de implementación; dichas pruebas se realizaron por medio de pruebas de casos de uso de la Universidad de Cartagena, soportadas en los servicios web actualmente prestados a los estudiantes.

### Resultados obtenidos:

- Documento con pruebas de análisis con las distintas herramientas de testeo utilizadas.
- Limitaciones del sistema, documentación de errores y buenas prácticas.

## **5.5 Etapa De Comunicación**

### Actividades:

- Se seleccionó y recopiló las distintas documentaciones obtenidas de las etapas preliminares, para desarrollar el informe final de la investigación y/o experimentación.

### Impacto Final:

Con el desarrollo de todas las etapas presentadas con anterioridad, basadas en la recopilación y análisis de información, el resultado final o documento a comunicar es el proyecto de grado y la sustentación del proyecto, para informar a la comunidad académica sobre el conocimiento generado mediante el desarrollo de la investigación e implementación del sistema clúster de alta disponibilidad.



## **6. RESULTADOS**

Los resultados del proyecto se presentan teniendo en cuenta las etapas establecidas en la metodología; las actividades planeadas en cada etapa fueron vitales para que el proyecto cobrara vida, basándose en la teoría y la práctica, y despliegue de pruebas de su funcionamiento; todo encaminado en la prestación de servicios con alta disponibilidad. Es importante resaltar que cada etapa corresponde a un objetivo específico del proyecto, por lo tanto en este capítulo se evidencia el cumplimiento de cada uno de ellos. A continuación, se presentan una a una las fases de la implementación del clúster de alta disponibilidad:

### **ETAPA DE CONCEPTUALIZACIÓN:**

Durante esta primera etapa se realizó una investigación sobre proyectos afines a modalidades de clústering, como lo fue lectura y análisis de artículos publicados recientemente en la academia, consulta a revistas nacionales e internacionales relacionadas con la implementación de la tecnología basada en alta disponibilidad obtenidas por medio de consultas a las bases de datos a las cuales se encuentra adscrita la Universidad de Cartagena y externas; y como resultado de dicha investigación, realizada teniendo en cuenta distintos puntos de vista, bibliografías, modelos y teorías (Presentadas en el Marco Teórico) según autores, se clasificaron algunos métodos de implementación (ver etapa de Implementación de Resultados), donde la elección de componentes bases del montaje de la tecnología fue vital, tales como el sistema operativo indicado, versión de servidor web a

utilizar y balanceador con middleware o intermediario entre aplicación de soporte a alta disponibilidad y la plataforma física de máquinas elegida para trabajar.

Como elemento base del clúster a implementar, se estudiaron varios núcleos de sistemas operativos de versión libre, relacionados en el marco teórico, eligiendo la versión de Linux Centos 5 que posee lógicamente las propiedades de Software libre y estabilidad en cuanto a desarrollo de procesos, y servicios web debido a que la comunidad sobre la cual se sustenta aseguran soporte por los próximos 7 años; además de ser concebido por la comunidad Linux como una distribución de la empresa -tal como su nombre lo indica Community ENTerprise Operating System- con un amplio repertorio de aplicaciones de escritorio que lo hacen una distribución altamente recomendable para servidores.

Para la implementación de un buen clúster HA se requiere una plataforma integral, que posea múltiples escenarios de desarrollo con soporte a funcionalidades futuras. Dentro de las propuestas viables para el diseño y posterior configuración de la plataforma altamente disponible que se desea implementar, fueron vistas en detalle 3 soluciones. La primera opción corresponde a la instalación de un servidor HTTP con Apache Tomcat, la segunda opción fue a nivel de VMware<sup>6</sup> con un licenciamiento Express y la opción final, y elegida para operar la infraestructura HA es Oracle VM<sup>7</sup> con soporte en alta disponibilidad de

---

<sup>6</sup> **VMware Inc.**, (VM de Virtual Machine)

<sup>7</sup> **Oracle Virtual Machine**

servicios y máquina, basada repositorios de datos bajo el protocolo NFS<sup>8</sup> con tester de servidores activos interno, mejor conocido como Heartbeat.

Para GNU/Linux existe una gran variedad de proyectos que aportan las características de Alta Disponibilidad, proyectos tanto para la alta disponibilidad en servicios como en datos. Algunos de los proyectos encontrados y destinados a ofrecer alta disponibilidad en servicios son:

Ha-oscar (CEnIT.Latech.edu, 2003). Es un proyecto Open Source cuyo objetivo es proporcionar de manera combinada el poder de la alta disponibilidad con un gran rendimiento de cómputo[16]. El proyecto está basado en un sistema de clustering Beowulf enfocándose por tanto a aplicaciones que requieren un grado elevado de disponibilidad. HAOSCAR tiene en cuenta los puntos débiles por lo que adopta redundancia de componentes. HA-OSCAR incorpora detección de fallos, mecanismos de recuperación, etc.

Kimberlite (Josh Huber, 2013) Es considerado un proyecto único ya que es una completa infraestructura de alta disponibilidad de código abierto que incluye incluso garantía sobre la integridad de los datos. Actualmente es un proyecto abandonado o de muy baja actividad [17].

LifeKeeper (SIOS, 2013). Es una alternativa moderna y flexible comparada con los productos que ofrecen soluciones de alta disponibilidad. Dispone de mecanismos para mantener la integridad de los datos. Una numerosa colección de kits de recuperación para

---

<sup>8</sup> **Network File System** (Sistema de archivos de red)

los distintos servicios y/o aplicaciones permite instalar LifeKeeper en cuestión de horas. Es un producto de la empresa SteelEye [18] .

KeepAlived. El objetivo principal del proyecto es el de proporcionar alta disponibilidad al proyecto Linux Virtual Server. KeepAlived es un servicio que monitoriza el estado del clúster LVS para que en caso de fallo este siga en funcionamiento.

Es recomendable a nivel empresarial e industrial implementar alta disponibilidad en datos para seguridad de la información. Dado que el presente proyecto será desarrollado a nivel académico y universitario, se decidió según la disponibilidad de recursos optar por una unidad de red compartida, que brindara la función de una SAN<sup>9</sup> o un arreglo de discos soportado en SCSI<sup>10</sup>.

Algunos de los proyectos destinados a ofrecer alta disponibilidad en datos, y que podrían ser estudiar para una posible adición de funcionalidades próxima, son:

DRBD (LINBIT, 2013). Software de replicación de dispositivos de bloque formando un RAID 1 a través de la red, es decir, replica los datos en distintas localizaciones. Datos de un sistema de archivos, de una base de datos, etc. Algunas de sus características son una replicación en tiempo real de manera continua, transparencia en las aplicaciones que estén almacenando datos en la unidad, posibilidad de recuperación de los datos ante un desastre, etc [20]. Se complementa perfectamente con Heartbeat.

---

<sup>9</sup> **storage area network (SAN)** - red de almacenamiento integral

<sup>10</sup> **Small Computers System Interface** (Interfaz de Sistema para Pequeñas Computadoras)

FreeNAS (Cochard-Labbé, 2013). Es un servidor NAS “Network-Attached Storage” gratuito, que soporta: protocolos Cifs (Samba), Ftp, Nfs, Rsync, autenticación de usuario local, RAID (0,1,5) por software con un completo interfaz de configuración Web. FreeNAS ocupa menos de 32MB una vez instalado en Compact Flash, disco duro o un dispositivo de memoria usb. FreeNAS está basado en FreeBSD que es un derivado de Unix, permitiendo clientes tanto de GNU/Linux como Windows y por supuesto Unix[21].

Openfiler (Community, 2013). Es un sistema operativo para compartir en red el almacenamiento de tipo SAN-NAS. Provee de buena integración con iSCSI (Internet SCSI) y “fibre channel”, muy útiles en entornos empresariales y en virtualización tal como Xen y VMware. Cuenta con una interfaz Web para su administración y soporta clientes de todos los sistemas operativos [22]. Dispone de soporte empresarial previo pago.

NAS Lite-2. Es un sistema operativo NAS que permite transformar un computador personal en un servidor de archivos Smb/Cifs, Nfs, Afp, Ftp, Http y Rsync. Su principal característica es que NAS Lite está optimizado para ofrecer el máximo rendimiento manteniendo un mínimo consumo de recursos, es decir, NAS Lite es compacto, estable y muy fiable y lo más sorprendente es que se ejecuta directamente sobre la memoria RAM necesitando únicamente 8MB de disco [23].

Para el caso de estudio concreto de la Universidad de Cartagena, dando cumplimiento al primer objetivo específico establecido, se determina que los requerimientos funcionales para la implementación del clúster son:

- Garantizar la gestión de la función de alta disponibilidad del clúster.
- Garantizar el almacenamiento compartido de datos.
- Garantizar la interconexión entre los entes que componen el clúster.

Existen variedad de proyectos dedicados a brindar alta disponibilidad de servicios y datos para la implementación de un clúster de alta disponibilidad, y el criterio principal de selección de una herramienta base para el montaje y configuración de una plataforma clúster debe estar enmarcada por que con ella exista buena documentación, soporte de implementación y además que sea suministrada por un proveedor consolidado en herramientas de Tecnología de la Información; si no se pretende incurrir en altos costos de implementación, lo lógico sería buscar una solución bajo el software libre y gratuito, hoy por hoy una compañía con buen soporte para proyectos experimentales o de laboratorio, con muchas herramientas de licenciamiento sin costo es: Oracle Corporation<sup>11</sup>.

Las razones por las cuales se ha seleccionado esta opción se detallan a continuación:

Oracle provee una gama de software con licenciamiento gratuito y sin restricciones de versiones, es decir, con funcionalidades completas ya sea licenciamiento con ediciones Lite (LE), Personal (PE), Express (XE), Standard (SE) o Enterprise (EE). Actualmente esta compañía brinda soluciones para el desarrollo de aplicaciones propias y potentes como Oracle Designer, JDeveloper - Java, Virtualización VirtualBox, Developer Suit, entre otras; que en su conjunto convierten a este proveedor en una excelente solución para el desarrollo

---

<sup>11</sup> Oracle Corporation – Especialista en tecnologías, bases de datos, etc.

del presente proyecto, ya que brinda un soporte y documentación gratuita hacia sus propias herramientas.

Una solución viable y acorde a los requerimientos de este proyecto, es Oracle VM como plataforma que ofrece un entorno totalmente equipado para un mejor aprovechamiento de los beneficios de la tecnología de virtualización. Dicha plataforma permite implementar sistemas operativos y software de aplicaciones dentro de un entorno de virtualización compatible y ajustada a los requerimientos funcionales necesarios para la implementación del cluster HA; los componentes integrados en la tecnología Oracle VM, son:

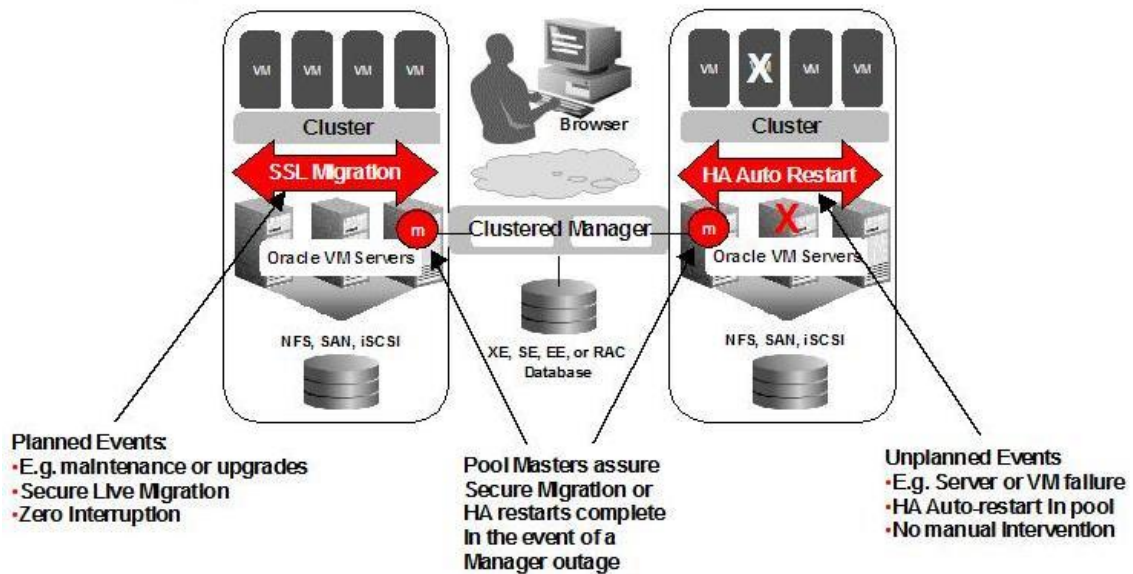
Oracle VM Manager que provee la interfaz de usuario como aplicación estándar ADF (Application Development Framework) web, para gestionar servidores Oracle. Gestiona el ciclo de vida de la máquina virtual; incluyendo la creación de máquinas virtuales desde los medios de instalación o desde una plantilla de máquina virtual, también eliminar, apagar, subir e implementar y migrar en vivo máquinas virtuales. Administra los recursos, incluidos los archivos ISO, plantillas de máquinas virtuales y las unidades de discos que se pueden compartir.

Oracle VM Server como entorno de virtualización autónomo, diseñado para proporcionar una plataforma ligera, segura y basada en un servidor para ejecutar máquinas virtuales. Oracle VM Server se basa en una versión actualizada de la tecnología de hypervisor Xen<sup>12</sup> subyacente, e incluye Agente VM Oracle capaz de conectar medios compartidos. El modelo teórico para el diseño que se estructurara para la Universidad de Cartagena está basado en la siguiente ilustración (Ilustración 2):

---

<sup>12</sup> **Xen**: monitor de máquina virtual de código abierto <http://www.xenproject.org/developers/teams/hypervisor.html>

Ilustración 2: Modelo de operación Oracle VM Server



Se requiere un medio compartido (Shared Storage) hacia donde apuntarán las dos máquinas de operación. Una de ellas siempre estará operativa y la otra trabajará como medio de respaldo ante eventos no planeados (Fallas del primer servidor, reinicio del pool, etc.)

Los requerimientos mínimos de hardware y software para la implementación del clúster de alta disponibilidad soportado por OVM. Dentro de los que se encuentran:

- Procesador: Arquitectura x86/amd64.-1.86 GHz o superior – Con soporte para Virtualización de medios.
- Espacio libre Disco Duro: 10 Gb.
- Espacio de Intercambio (swap space): 2 Gb.
- Memoria: 2 Gb
- Tarjeta de Red



- Red Hat Enterprise Release 4 o superior
- Internet Explorer 8, Mozilla Firefox 2.0, Safari 7.2 o Superior
- Disponibilidad de Puertos 8888 y 8899

Dichos requerimientos son básicos para la implementación, dado que se trabajara con medios virtualizados, implica que se utilizaran recursos para el sostenimiento de la maquina física, y otra parte de recursos hardware para el hospedaje de la máquina virtual en la cual se alojaran todos nuestros servicios web prestados.

Como plataforma inicial se utilizaran 3 computadores, de los cuales dos estaciones son servidores de aplicaciones (un Maestro por defecto del clúster y otro de utilidades y/o servicios), y otro pc es utilizado para la instalación y configuración del manager de aplicaciones, que a su vez tendrá la función de prestar el medio compartido bajo el protocolo NFS.

De esta forma se detallan los requerimientos a nivel de infraestructura para la implementación del clúster altamente disponible:

1. Se necesitan 2 o más equipos que soporten virtualización de Medios con Hypervisor Xen.
  - a. Un equipo será utilizado como Maestro y los demás como servidores nodos.
2. Se necesita un equipo desktop en red, que tendrá funciones de Repositorio de Datos o Medio compartido para el Clúster. También tendrá la función de Manager del Clúster.

3. Un equipo activo de interconexión de computadores en red, sea switch o router.

Como todo sistema de servidores cada máquina debe contar con una dirección IP Estática, desactivando el sistema DHCP o asignación de ip's dinámicas, esto para facilidad en las rutas de operación y conexión.

En este orden de ideas se evidencia el cumplimiento del primer objetivo específico del proyecto; eligiendo la solución de Oracle de Alta disponibilidad con Virtual Machines OVM Server.

### **ETAPA DE DISEÑO:**

Dando cumplimiento al segundo objetivo específico del proyecto, y como se referencio anteriormente en la etapa de conceptualización, se busca implementar y diseñar un modelo de plataforma clúster de alta disponibilidad bajo una solución que no incurra en altos costos de inversión; se elige Oracle Virtual Machine como plataforma para brindar de alta disponibilidad de servicios y máquina, ya que su licenciamiento es gratuito y es netamente propio del proveedor Oracle; no teniendo restricción alguna para descargas desde su Web Site de los software para nodos y núcleos de sistemas operativos necesarios para la instalación, montaje e implementación de la plataforma clúster en las instalaciones de la Universidad de Cartagena.

Basados en el modelo de operación de Oracle VM, se presenta la arquitectura y diseño de plataforma más adecuada para la reactivación de servicios por fallas de uno los nodos componentes instalados en un clúster con alta disponibilidad habilitada; Oracle VM propone una redundancia de nodos apuntando a un medio de datos compartido. En su documentación encontramos que centran su desarrollo en plataformas SAN o arreglos SCSI, pero para los objetivos planteados en la especificación del proyecto se determinó que no se debe incurrir en altos costos de inversión, por consiguiente se implementara bajo una conexión a equipo de red con protocolo NFS activo.

El diseño final que se aplica para la posterior implementación implica la utilización de 3 computadores; de los cuales dos son servidores de aplicaciones y el otro pc es utilizado para la instalación y configuración del manager de aplicaciones. El manager es el ente principal de configuración y gestión del clúster; en su interior se encuentra todo el esquema del grupo de servidores o POOL, de igual modo el medio de instalación de la máquina virtual huésped para los servidores de aplicaciones y por último la máquina virtual como archivo .IMG. El archivo de imagen es trasladado por medio de un agente Oracle que posee cada nodo del POOL de servidores para permitir el funcionamiento del sistema operativo interno (Máquina Virtual OVM), como se observa en la siguiente figura ilustrativa (Ilustración 3).

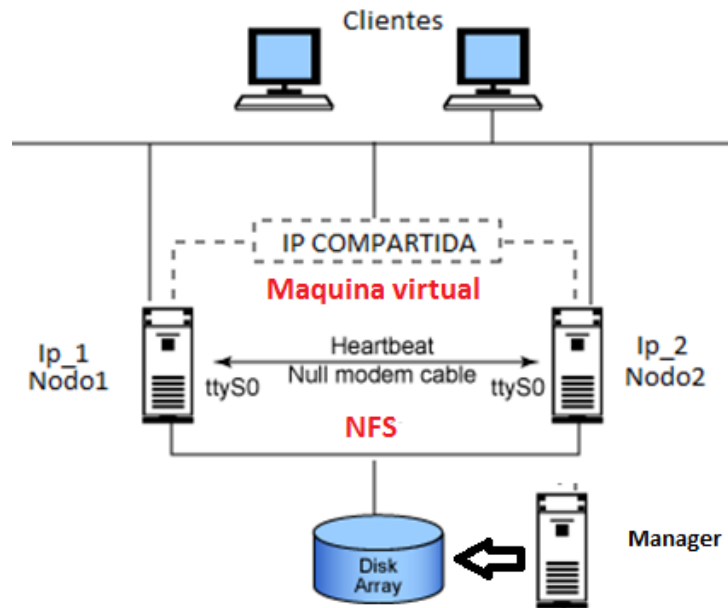


Ilustración 3: Modelo de diseño para PC y red para el clúster HA.

En dicho diseño siempre se realizarán peticiones de servicios a la dirección IP destinada a la máquina virtual, que estará alojándose en uno de los nodos activos; esto quiere decir que para que el clúster preste sus servicios mínimo debe existir una máquina Master activa. Se ha realizado una configuración de direcciones IP estándar, pero que no es obligatoria usarla, dicha configuración es presentada en la siguiente tabla de equipos descriptivos del clúster (Tabla 2).

<b>HOSTNAME</b>	<b>TIPO DE EQUIPO</b>	<b>DIRECCION IP</b>	<b>FUNCIONES</b>
OVMANAGER	Repositorio de Datos	192.168.1.100	Manager para gestión del clúster
OVMSERVER01	NODO1, Master	192.168.1.101	Agente para Hospedar máquina virtual HA.
OVMSERVER02	NODO2, Utilidades y servicios.	192.168.1.102	Agente para Hospedar máquina virtual HA.

**Tabla 2: Descripción de Equipos computadores componentes del clúster.**

A su vez, se implementó un Medio Compartido “SharedVol/” desde una localización remota (en este caso en el mismo equipo donde estará configurado el Oracle VM Manager) para almacenar los datos de la máquina virtual y archivos de configuración de nuestro clúster de servidores; en la siguiente imagen se denotan las conexiones a 3 equipos del montaje “OVMSERVER01, OVMSERVER02 y OVMANAGER” y la conexión hacia máquina virtual “serverweb” alojada en uno de los nodos de operación. En la siguiente fase de la metodología propuesta se muestra paso a paso el desarrollo del diseño propuesto para la implementación del clúster. En las siguientes ilustraciones (Ilustración 4) se muestra físicamente, el diseño planteado anteriormente; en donde 3 veces y un activo de red son los equipos utilizados para que el clúster se encuentre en funcionamiento.



**Ilustración 4: Conexiones Físicas y maquinas hardware componentes del Clúster**

Los equipos deben estar interconectados por medio de un dispositivo de red (switch ó router) de calidad para brindar una buena transmisión de datos y que el motor de middleware pueda obtener respuestas inmediatas ante posibles fallas; cabe anotar que en el mercado existen soluciones Cisco con precios asequibles para el nivel de operación de pymes, con equipos de gama media (routers y switches capa2) pero con funciones óptimas para la operación de prestación de servicios web de calidad, en el presente proyecto se utilizó un router Linksys Cisco<sup>13</sup>, proporcionado por el programa de Ingeniería de Sistemas de la Universidad de Cartagena. En la siguiente Ilustración 5, se muestran las conexiones a los 3 computadores y a la máquina virtual por medio del protocolo SSH<sup>14</sup>, usado por la aplicación PUTTY.exe<sup>15</sup>.

<sup>13</sup> **Cisco:** Cisco System Inc. Fabricante equipos de telecomunicaciones (California – Estados Unidos)

<sup>14</sup> **SSH:** Secure Shell o Interprete de Ordenes Seguras.

<sup>15</sup> **Putty:** Es una implementación libre para conexiones Telnet y SSH para Windows y Linux.

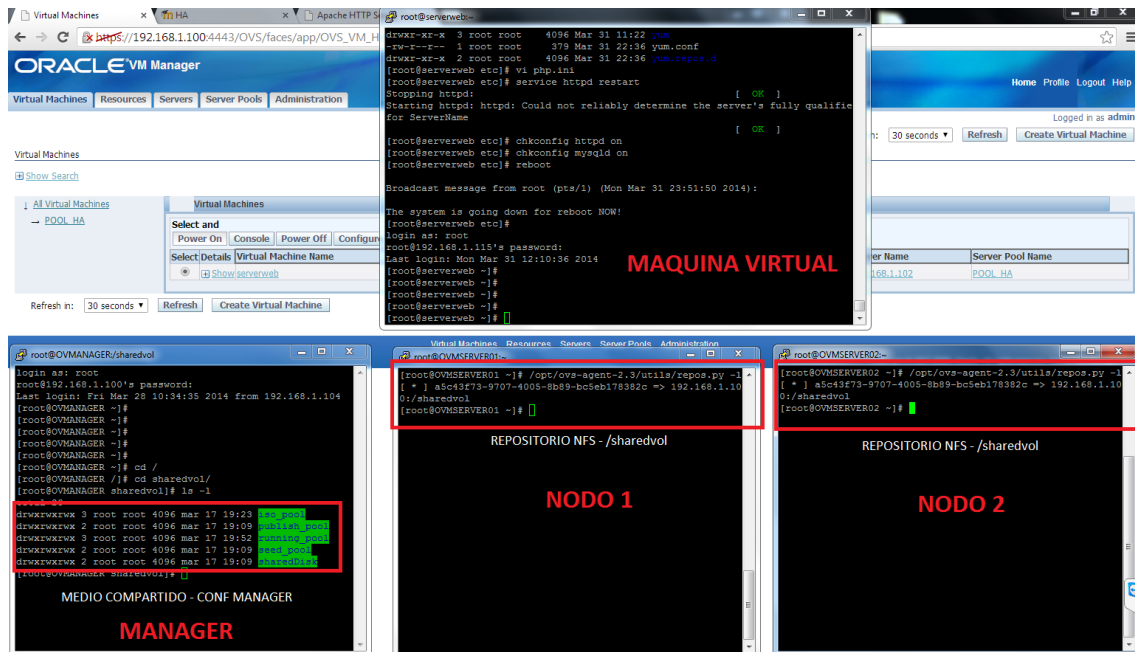


Ilustración 5: Conexiones lógicas y direcciones del clúster HA

Una breve explicación de cómo se compone el protocolo NFS y porque se utilizó, sería basada en que un sistema de archivos distribuidos bajo el protocolo mencionado para entornos heterogéneo permite a los usuarios de ordenadores, acceder a un sistema de archivos distante sin tener que saber los comandos de origen, aun cuando estén en sistemas operativos diferentes. Un sistema de Archivos en Red, permite a los nodos que componente de esta, cargar sistemas de archivos e interactuar con ellos como si estuviesen alojados localmente en la máquina. En el clúster HA, es utilizado como medio disponible para crear un repositorio compartido, en donde se hospeda la máquina virtual que corre en los distintos nodos de clúster.

Actualmente hay 3 versiones de NFS. NFS versión 2 (NFSv2) la más antigua y la que cuenta con más amplio soporte, NFS versión 3 (NFSv3) la cual permite variar el tamaño de los archivos y mejoras en el reporte de errores, aunque no es completamente compatible con NFSv2; Por último se encuentra la versión 4 de NFS (NFSv4), la cual utiliza Kerberos como mejora en la seguridad, trabaja sobre Internet y a través de firewalls. Red Hat Enterprise Linux soporta las 3 versiones del NFS y cuando monta un sistema de archivos lo hace con NFSv4 por defecto.

Algunas de las falencias encontradas sobre los sistemas de servicios web en cuanto a su disponibilidad fueron causadas por manipulaciones sobre los archivos de configuración; además de ser en su mayoría servidores de tipo multiplataforma no se aprovecha la mayoría de funcionalidades que brindan los sistemas operativos. Además, las aplicaciones de servidores web, por defecto, están asociadas a su instalación sobre el puerto 80, así que al momento de instalar otras aplicaciones en donde por error se utilice el mismo puerto, este se verá afectado para dar el output a peticiones de usuarios externos.

La alternativa de utilizar el clustering para dar un servicio web estable y con un alto grado de disponibilidad, brinda soluciones inmediatas con respecto a modificaciones sobre alguno de los nodos; es decir, se tiene una máquina paralela, la cual mantiene el servicio disponible y activo al realizar pruebas o modificaciones en alguno de los nodos.



## **ETAPA DE IMPLEMENTACIÓN:**

Para la fase de implementación del clúster de alta disponibilidad de un servidor Apache bajo Linux, se realizó siguiendo las actividades relacionadas a continuación; es necesario contar con un DVD de instalación CentOS 5.2 y Oracle VM con su respectivo Aplicación Manager y S.O Server en su versión 2.

Durante la instalación de los componentes software se presentaron inconvenientes con la compatibilidad de los equipos con referencia HP Compac 6200 Pro, debido a que el fabricante de dicho equipo no proporciona controladores para la tarjeta de red incluida en el equipo trabajara sobre el núcleo de Linux Red Hat; y no se contaba con tarjetas de red adicionales que trabajaran en puerto adicionales PCI o PCI Express. Por lo anterior, fue necesario recurrir a la utilización de equipos de cómputo con referencia comercial HP 500 MT, teniendo en cuenta que muchos de equipos de esta referencia suministrados por el Programa no contaban con el parámetro de Virtualización en ON y no fue posible acceder a la BIOS para correcta configuración de este importante parámetro en el ámbito de proyecto. Solo fue posible acceder a la configuración de la BIOS en dos equipos con referencia HP 500 MT que, a diferencia del resto de su misma referencia, no requerían contraseña de acceso a la BIOS (Basic Input-Output System)

Los materiales de instalación podrán ser descargados desde el sitio web de desarrollo de Oracle: <https://edelivery.oracle.com/>; a continuación se presenta un pantallazo de la página principal del sitio de desarrollo de Oracle (Ilustración 6), para acceder a esta solo es necesario crear una cuenta gratuita en el sitio y confirmarla por medio del correo

electrónico y luego proceder a la búsqueda de las herramientas necesarias para el montaje del clúster (Oracle Virtual Server y Oracle Virtual Machine)



**Ilustración 6: Plataforma de descarga Oracle. (Oracle.com - 2014)**

Atendiendo al diseño planteado en la etapa anterior, se procede al montaje de las plataformas necesarias para la configuración de las mismas; para ello se presenta el siguiente resumen general del plan de las operaciones que deben ser realizadas en cada estación de trabajo y montajes necesarios para posterior configuración y pruebas:

1. Instalación y configuración OVMSERVER en los dos nodos.
2. Instalación Centos-5 para el Manager y repositorio.
3. Instalación del Manager. No elegir protocolo https para posterior facilidad de acceso /OVS (Repositorio de datos)
4. Actualizar centos-5 con comando: yum -y upgrade
5. Se configura el repositorio con el servicio NFS.

- a. Instalación del servicio nfs con el comando: `yum -y install nfs-utils`
  - b. Configuración de NFS e instalación: `yum -y install system-config-nfs`
  - c. Definir los puertos – modificación de archivo: `vi /etc/sysconfig/nfs`
  - d. Iniciar el servicio con el comando: `service nfslock start`
  - e. Crear una carpeta que será el medio compartido en la raíz: `mkdir /sharedvol`
  - f. Dar permisos total sobre la carpeta: `chmod 777 sharedvol/`
6. Modificar archivo de configuración y acceso por nfs: `vi /etc/exports`
- a. `/sharedvol 192.168.1.102(rw,insecure,sync,no_root_squash)`  
`192.168.1.101(rw,insecure,sync,no_root_squash)`
7. Modificación del archivo hosts en cada nodo. Con el comando “`vi /etc/hosts`”
- a. Agregar la ip del nodo con su nombre
    - i. Ejemplo: `192.168.1.101 OVMSERVER01`
    - ii. Ejemplo: `192.168.1.102 OVMSERVER02`
8. Borrar repositorio local → Server guide -> Managing Storage Repositories.
- a. Comando para lista de repositorios:
    - i. `/opt/ovs-agent-2.3/utils/repos.py -l`
  - b. Comando para borrar repositorio local:
    - i. `/opt/ovs-agent-2.3/utils/repos.py -d “AUID”`
9. Creación del repositorio en los servidores, con el medio compartido /sharedvol. Se utilizó NFS por no contar con SAN, NAS o SCSI.
10. Creación de repositorios en cada nodo, por medio de agente OVS:
- a. Creacion: `/opt/ovs-agent-2.3/utils/repos.py -n 192.168.1.100:/sharedvol`
  - b. Revisar: `/opt/ovs-agent-2.3/utils/repos.py -l`

c. Correr: `/opt/ovs-agent-2.3/utils/repos.py -r "AUID`

d. Activar: `/opt/ovs-agent-2.3/utils/repos.py -i`

11. Con el comando `df-h`, se ve el repositorio apuntando al medio compartido `192.168.1.100:/sharedvol`

12. Crear archivo de clúster en `/etc`:

a. Con el comando `mkdir ocfs2 -> cd /ocfs2`

i. Se crea el archivo `vi cluster.conf` → interconexión de nodos.

13. Creación del pool de servidores con HA. Requiere más de un servidor y un almacenamiento compartido. Solo un nodo puede obtener características Master.

14. Copiar archivo `.iso` en `/sharedvol/iso_pool/ISO`

15. Creación de la máquina virtual o GUEST activando la función de HA.

16. Con el comando: `'Chkconfig --level 35 "nfs or httpd" on` se inicia el servicio de manera automática con el arranque de la máquina.

Obedeciendo al plan de montaje de la plataforma clúster ya planteada en el anterior plan de operación y configuración, fue replicada en máquinas físicas el diseño de plataforma de alta disponibilidad realizado en el presente proyecto.

La instalación de los Oracle VM server, esta soportada según la documentación de Oracle disponible en su sitio web.

- **Fase de instalación de OVM para nodos:**

El montaje de la plataforma da inicio con la instalación de las máquinas de servicio o nodos de operación; estas cuentan con el sistema de virtualización de Oracle OVM Server. Se configura el arranque del sistema desde el medio en que se encuentre disponible sea CD-ROM, USB, etc. En la siguiente imagen (Ilustración 7) se muestra el primer paso de instalación para el sistema de Oracle VM



**Ilustración 7: Pantalla inicial Instalación de Oracle VM.**

Esta instalación consta de un procedimiento básico de instalación donde lo más importante es contar con un borrado completo de los discos duros y asignación de direcciones IP estáticas desde la instalación.

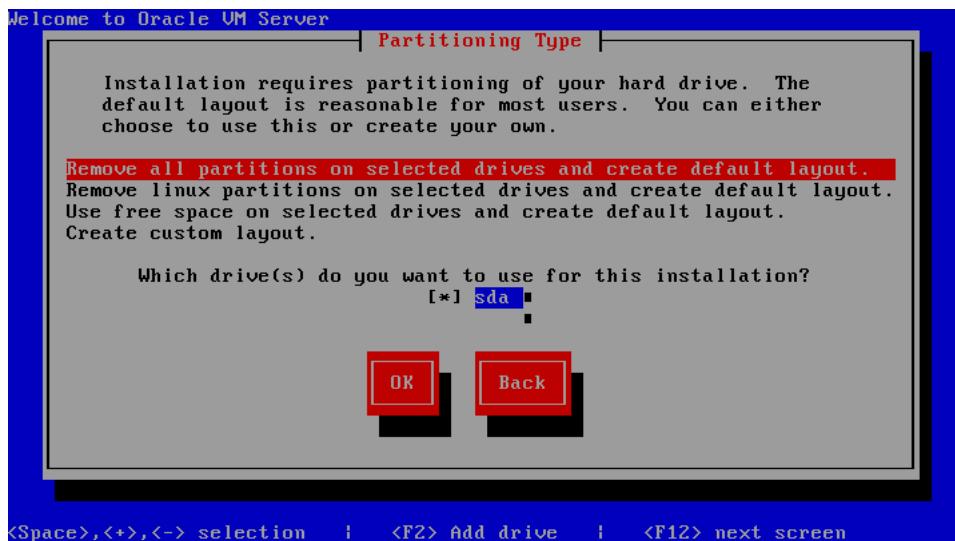


Ilustración 8: Selección de Instalación desde una nueva partición.

La estructura del disco duro quedaría muy parecida a la que se muestra a continuación, en la Ilustración 9; una estructura estándar para el arranque del sistema Linux server:

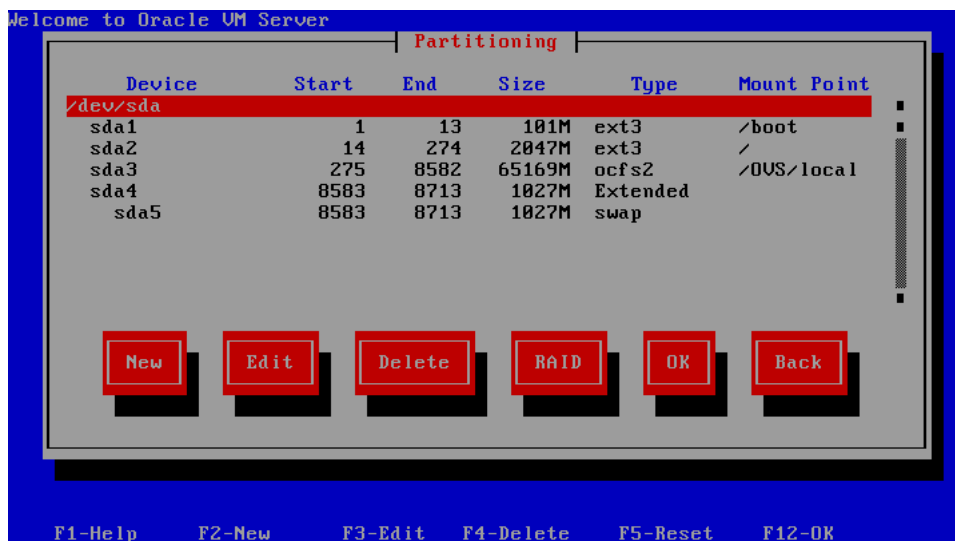


Ilustración 9: Estructura de particiones de discos Estándar Oracle VM.

Luego como paso importante a seguir es la asignación de dirección y configuración de la tarjeta de Red (eth0), como se muestra en la siguiente Ilustración 10, para su posterior configuración:

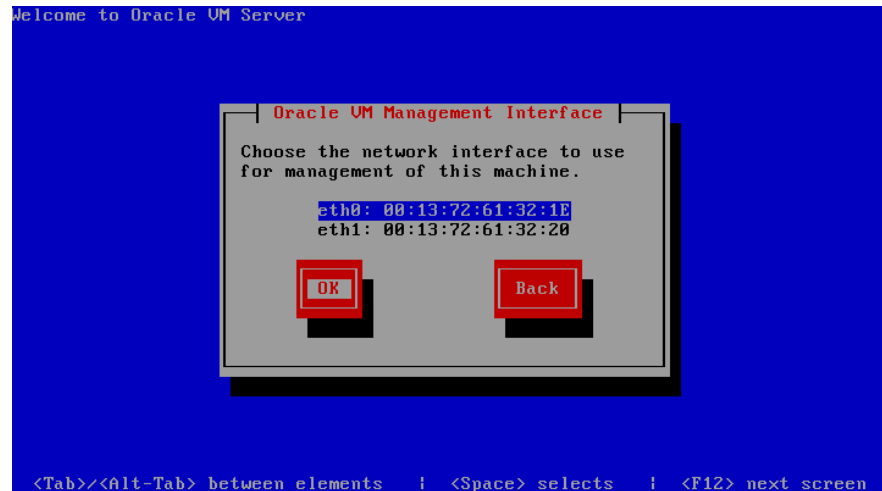
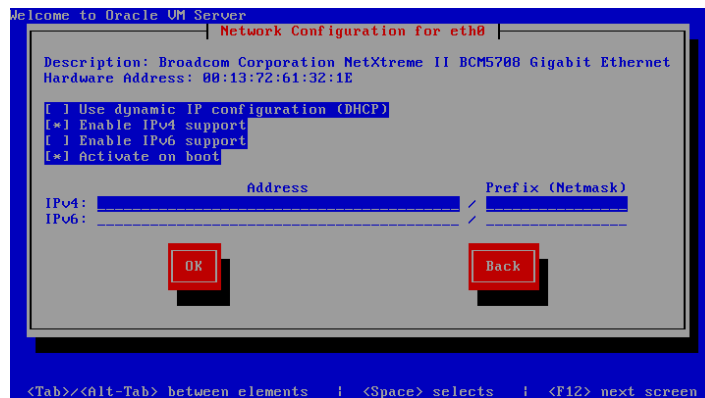
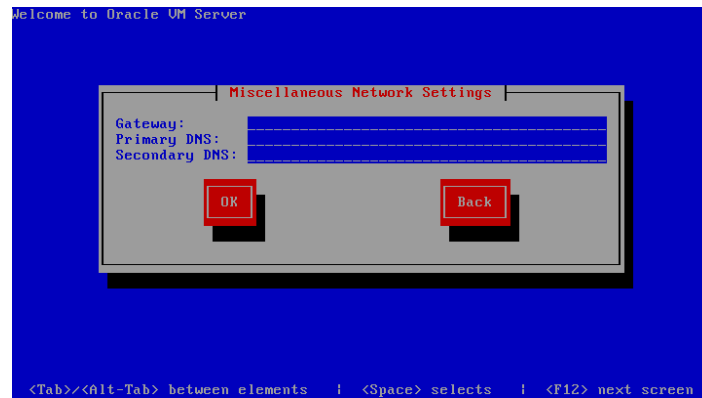


Ilustración 10: Selección de Tarjeta de Red eth0.

Como lo mencionábamos anteriormente asegurarse que las maquinas queden con una asignación de dirección IP manual y configuración de puerta de enlace, mascara de subred y DNS fijo. En las siguientes Ilustraciones 11 y 12, se muestran los campos requeridos para la configuración IP.





**Ilustración 11 y 12: Configuración de parámetros de Red – Ip, Netmask, Gateway y DNS.**

Luego se asignan contraseñas para el usuario root y para el agente de Oracle Virtual Machine. A continuación se debe dar reinicio de la máquina para cargar todos los drivers y aplicaciones, en la siguiente Ilustración 13 se muestra el pantallazo de reinicio de Nodo.



**Ilustración 13: Reinicio necesario para confirmar toda la configuración.**

El Login se realiza con la cuenta de usuario Root y con el Password, como se muestra en la Ilustración 14, asignado durante la instalación del Oracle VM Sever. De esta forma el



agente del sistema VM de Oracle queda activo para la posterior configuración desde su Manager de Procesos.

```

Oracle VM server release 2.1.1
Hypervisor running in 64 bit mode with NO Hardware Virtualization support.

Network :
Management Interface :
If : eth0(Up) Mac : 00:0C:29:EC:9F:29 IP address : 192.168.55.128

Configured Networks and Bridges :
If : eth0      Mac : 00:0C:29:EC:9F:29
If : xenbr0    Mac : FE:FF:FF:FF:FF:FF

CPU :
cpu family    : 6
model         : 15
model name    : Intel(R) Core(TM)2 Duo CPU   E6850  @ 3.00GHz

localhost login:

```

**Ilustración 14: Inicio y Login de Oracle VM Server.**

Posteriormente se configuro todo la interconexión de red para clúster en cada uno los nodos asignándole los siguientes parámetros en el directorio /ocfs2:

- Crear un archivo cluster.conf, con el fin de brindar la interconectividad entre los nodos.

```

node:
ip_port = 7777
ip_address = 192.168.1.101
number = 0
name = OVMSERVER01
cluster = ocfs2

node:
ip_port = 7777
ip_address = 192.168.1.102
number = 1
name = OVMSERVER02
cluster = ocfs2

cluster:
node_count = 2
name = ocfs2

```

La instalación del Oracle VM manager consta de una previa instalación de un núcleo Centos 5.2, cuyo procedimiento de instalación es prácticamente igual a la instalación de un nodo Oracle VM server; no es necesario en este punto reiterar la instalación de un sistema Linux, ya que tanto Centos como Oracle VM server están basados en la versión Red Hat 5 de Linux.

Para el repositorio NFS o medio compartido en donde se encontraran los archivos de imagen para la instalación de la máquina virtual, se crea una carpeta y se asignan los siguientes permisos para el archivo /etc/exports del equipo con el medio:

- /sharedvol 192.168.1.102(rw,insecure,sync,no\_root\_squash)  
192.168.1.101(rw,insecure,sync,no\_root\_squash)

En donde cada dirección IP hace referencia a los nodos componentes del clúster.

- **Instalación del Oracle VM Manager:**

La instalación del Manager para los nodos de procesos cuenta con el núcleo de la aplicación de manejo de equipos y una versión de una base de datos Oracle en su versión express 10g; además un cliente Oracle que soporta J2EE (Java 2 Platform, Enterprise Edition), y librerías de Apache XML-RPC (Remote Procedure Call).

Como primer paso se debe descargar completamente el paquete de instalación desde la plataforma de Oracle: <http://edelivery.oracle.com/oraclevm> y colocar todo el contenido de instalación en una ruta fácil de acceso, en este caso utilizaremos una carpeta en la raíz:

```
# mkdir Manager
# mount -o loop,ro OracleVM-Manager-version.iso /Manager
```

En donde el Manager, es el folder en donde se coloca el ISO de instalación para el Oracle VM Manager. Teniendo nuestro archivo .ISO decomprimido se debe correr el instalador, mediante los comandos:

```
# cd /Manager
# sh runInstaller.sh
```

En adelante el procedimiento es de selección y digitación de contraseñas, para la cuenta administrador del ambiente web y de las funcionalidades adicionales del instalador (Motor BD 10g, Seleccionar la opción 1, para la instalación del Manager:

```
Please enter the choice: [1|2|3]
1. Install Oracle VM Manager
2. Uninstall Oracle VM Manager
3. Upgrade Oracle VM Manager
```

```
Starting Oracle VM Manager 2.2.0 installation ...
```

Luego se instala la base de datos de Oracle 10g:

```
Do you want to install a new database or use an existing one? [1|2]
1. Install a new Oracle XE database on localhost
2. Use an existing Oracle database in my network
```

Se procede a la instalación del paquete de OVS Manager, seleccionando la opción 1 del anterior menu:

```
Installing the ovs-manager package (rpm) ...
```

```
Done
Installing the oc4j package (rpm) ...
Done
```

Obteniendo toda la verificación y confirmación que el sistema de manager se encuentra instalado:

```
Installation of Oracle VM Manager completed successfully.
```

To access the Oracle VM Manager **version** home page go to:

```
http[s]://IP_address:port/OVS
```

To access the Oracle VM Manager web services WSDL page go to:

```
http[s]://IP_address:port/OVSWS/LifecycleService.wsdl
```

```
http[s]://IP_address:port/OVSWS/ResourceService.wsdl
```

```
http[s]://IP_address:port/OVSWS/PluginService.wsdl
```

```
http[s]://IP_address:port/OVSWS/ServerPoolService.wsdl
```

```
http[s]://IP_address:port/OVSWS/VirtualMachineService.wsdl
```

```
http[s]://IP_address:port/OVSWS/AdminService.wsdl
```

To access the Oracle VM Manager help page go to:

```
http[s]://IP_address:port/help/help
```

Con la instalación del Manager completada se procede a iniciar todos los servicios instalados anteriormente mediante los comandos:

```
# /sbin/service oc4j start (inicio del sistema)
# /sbin/service oc4j stop (Para el sistema)
# /sbin/service oc4j status (verificar estatus del Manager)
```

Para la posterior configuración del Manager, se accede a la dirección `http://127.0.0.1:"port"/OVS` la cual muestra un ambiente web donde se realiza un login con el user y password asignados durante la instalación. Los roles que tiene la plataforma del Manager se presenta en las tablas de usuarios disponibles, como son descritos en la Ilustración 15:

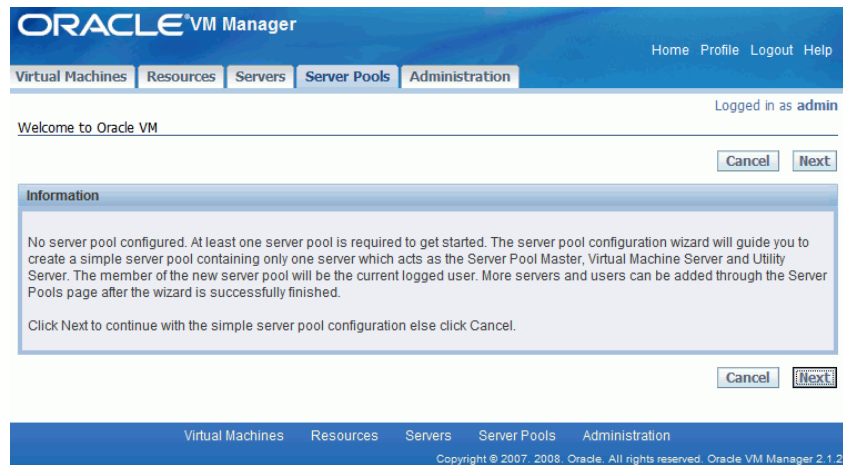
Role	Available Tabs
User	<ul style="list-style-type: none"> <li>■ Virtual Machines</li> <li>■ Resources</li> </ul>
Manager	<ul style="list-style-type: none"> <li>■ Virtual Machines</li> <li>■ Resources</li> <li>■ Servers</li> <li>■ Server Pools</li> </ul>
Administrator	<ul style="list-style-type: none"> <li>■ Virtual Machines</li> <li>■ Resources</li> <li>■ Servers</li> <li>■ Server Pools</li> <li>■ Administration</li> </ul>

Ilustración 15: Roles de usuarios Oracle VM.

The screenshot shows the Oracle VM Manager web interface. At the top right, there is a navigation menu with links for Home, Profile, Logout, and Help. The 'Profile' link is circled in red. Below the navigation menu, there are tabs for Virtual Machines, Resources, Servers, Server Pools, and Administration. Under the 'Virtual Machine Templates' section, there is a search bar and a table with columns for Virtual Machine Template Name, Size (MB), Server Pool Name, Status, and Creation Time. The table currently shows 'No rows yet.' and has refresh and import buttons. The footer of the page contains the text 'Copyright © 2007-2009. Oracle. All rights reserved. Oracle VM Manager 2.2'.

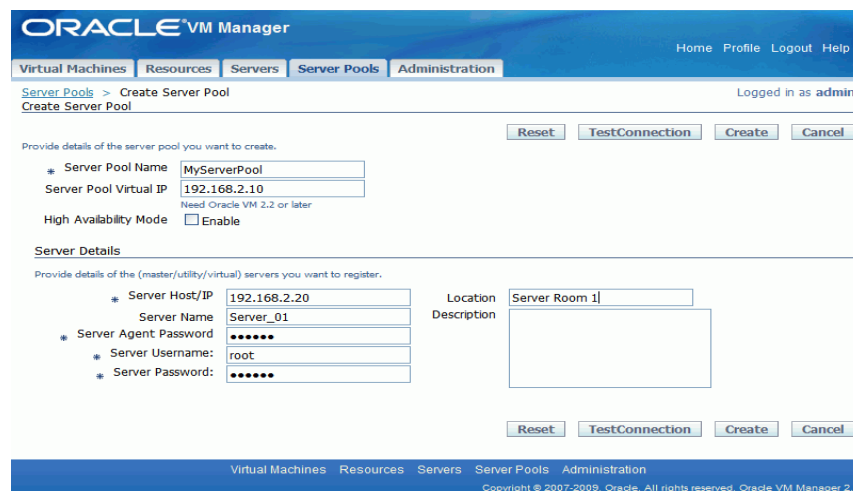
Ilustración 16: Perfil de Usuario interfaz web de Oracle VM.

Los siguientes pasos de configuración serán, la configuración del server pool donde están direccionados los nodos, en la siguiente Ilustración 17 se muestra el primer paso para la creación el pool de servidores o nodos:



**Ilustración 17: Pantalla inicial – Creación Server Pool.**

Configurar parámetros como: Nombre del server pool, IP VIRTUAL para acceso al server pool, se activara el modo de Alta disponibilidad, Asignaran las IP del Server Pool, Nombre y Contraseña del Agente configurados en la instalación de cada uno de los nodos. Obteniendo un pantallazo de resultado parecido al siguiente, y por ultimo crear el server Pool; en la Ilustración 18 se muestra la parametrización de nodos que anteriormente se tabulo (Etapa de Diseño):



**Ilustración 18: Parametrización de Server Pool.**

Al adicionar todos nuestros servidores OVM y tener todos los parámetros de configuración terminados, obtendremos un resultado con la cantidad de máquinas configuradas y usuarios adicionados para cada uno de los servidores o Nodos. En la Ilustración 19 se muestra la totalización de Nodos incluidos al Server Pool, así como los usuarios habilitados para trabajar sobre el Server Pool.

The screenshot shows the Oracle VM Manager interface. At the top, there is a navigation bar with tabs for Virtual Machines, Resources, Servers, Server Pools, and Administration. The 'Server Pools' tab is selected. Below the navigation bar, there are buttons for 'Refresh' and 'Create Pool'. A table displays the details for the 'MyServerPool' server pool. The table has columns for 'Select', 'Server Pool Name', 'Status', 'High Availability Status', 'Servers', 'Users', and 'Logs'. The 'Servers' and 'Users' columns show 'Total: 1' for the 'MyServerPool' row, which are circled in red. There is also a 'View Logs' link circled in red. At the bottom of the interface, there is a footer with the text 'Copyright © 2007-2009. Oracle. All rights reserved. Oracle VM Manager 2.2'.

Select	Server Pool Name	Status	High Availability Status	Servers	Users	Logs
<input checked="" type="radio"/>	MyServerPool	Active	Disabled	Total: 1	Total: 1	View Logs

**Ilustración 19: Totalización y Resultados de creación Server Pool.**

Obtenida y configurada toda la plataforma de servidores OVM, se procede a la instalación de una máquina virtual que será leída por cada uno de los nodos, en donde tendremos toda la configuración del o los servicios web que necesitemos:

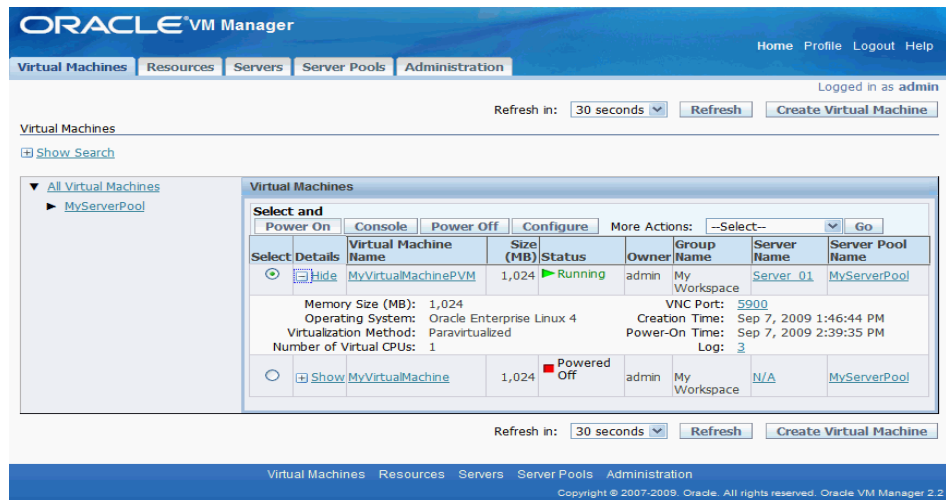
The screenshot shows the Oracle VM Manager interface. At the top, there is a navigation bar with the following tabs: Virtual Machines, Resources, Servers, Server Pools, and Administration. Below this, there is a sub-navigation bar with: Virtual Machine Templates, Virtual Machine Images, ISO Files, and Shared Virtual Disks. The user is logged in as 'admin'. The main content area is titled 'Virtual Machine Templates' and includes a search bar, a 'Refresh in: 30 seconds' dropdown, and 'Refresh' and 'Import' buttons. Below the search bar, there is a table of Virtual Machine Templates with the following columns: Virtual Machine Template Name, Size (MB), Server Pool Name, Status, and Creation Time. The table contains three rows of data. Below the table, there is another 'Refresh in: 30 seconds' dropdown and 'Refresh' and 'Import' buttons. At the bottom of the page, there is a footer with the text: 'Virtual Machines Resources Servers Server Pools Administration Copyright © 2007-2009, Oracle. All rights reserved. Oracle VM Manager 2.2'.

Virtual Machine Template Name	Size (MB)	Server Pool Name	Status	Creation Time
<input checked="" type="radio"/> QVM_EL4U5_X86_HVM_4GB	6145	MyServerPool	Pending	Sep 4, 2009 6:22:22 PM
<input type="radio"/> QVM_EL4U5_X86_PVM_4GB	6145	MyServerPool	Pending	Sep 4, 2009 6:21:19 PM
<input type="radio"/> QVM_EL5U1_X86_64_HVM_4GB	6145	MyServerPool	Active	Sep 4, 2009 6:20:28 PM

**Ilustración 20: Confirmación de Medios para creación Máquinas Virtuales.**

La imagen virtual la podremos crear desde cero, instalando nuestro núcleo Linux desde un .ISO en la opción de ISO Files. El modelo utilizado para la creación de nuestra máquina virtual es mediante la selección de un archivo .iso colocado en nuestro medio compartido en el equipo Manager “/sharedvol:\192.168.1.100”, y posterior configuración por medio de consola gráfica; en la siguiente Ilustración 21 se muestra la interfaz de monitoreo para la máquina virtual, donde se puede gestionar encendido o apagado del sistema operativo virtual.





**Ilustración 21: Instalación de Núcleo Linux – Virtual Machine.**

Luego de instalada y seguir paso a paso la instalación de la máquina virtual desde el Manager, se verá en la plataforma web una configuración estándar de la VM (Virtual Machine). En esta máquina virtual estarán instalados todos los servicios que se requieren con valores de alta disponibilidad; es en este punto se entiende que la máquina virtual siempre debe estar corriendo para que los servicios se encuentren iniciados.

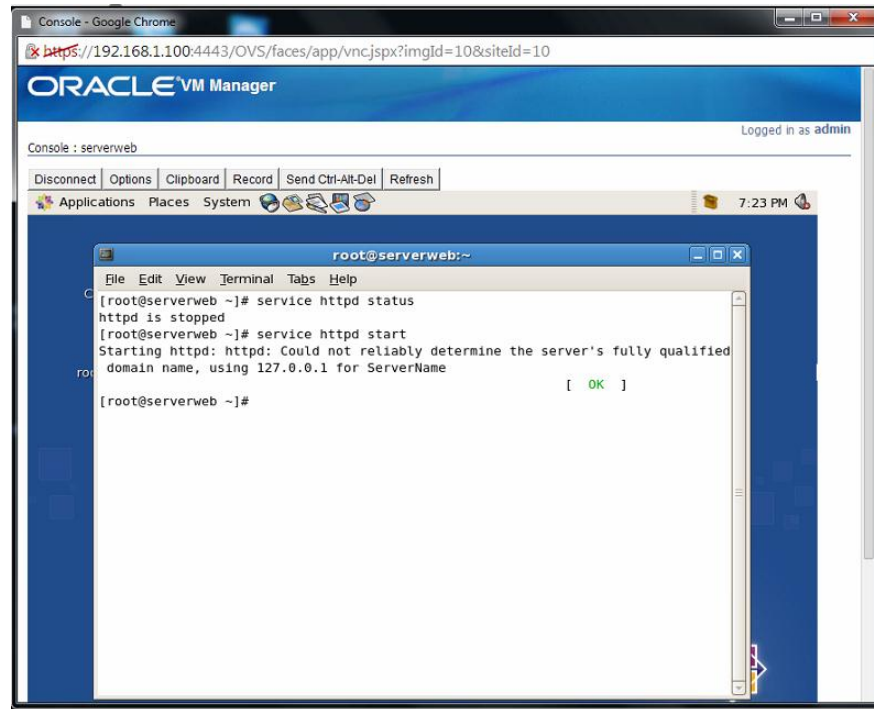
Una ventaja grandísima de esta modalidad en virtualización, es que en tan solo pocos minutos podremos recrear el sistema sin modificaciones, sencillamente instalando un nuevo nodo Oracle VM server de cero y apuntando el archivo .img.

Y por último faltaría encender la máquina virtual para que los nodos comiencen a operar según la cantidad de CPU configuradas.

El servicio de apache server deberá ser configurado en la máquina virtual (cuyo archivo .img está ubicada en la carpeta “/sharedvol/running\_pool”) de proceso, para que pueda ser

compartido para el Pool creado. Dicho servicio podrá ser visualizado por cada uno de los nodos mediante el hospedaje de la maquina por medio del Oracle ovs-agent o agente incluido en cada uno de los servers, y los clientes o peticiones externas podrán llamar por medio de la dirección IP configurada para dicha máquina virtual y el puerto configurado para el servicio web.

A continuación se muestra en la Ilustración 22, como es el acceso a nuestra máquina virtual por medio de la consola de visualización y administración del Manager de configuración:



**Ilustración 22: Servicio web – Servidor Apache corriendo en Máquina Virtual.**

Cada uno de los pasos de instalación está soportados por 4 videos en flash adjuntos al proyecto, los cuales llevan por nombre:

- 1. Acceso al Oracle VM Manager y Creación de una Oracle VM
- 2. Configuración e Instalación de Máquina Virtual
- 3. Escenario Prueba - Falla de desconexión o bloqueo de servidor.
- 4. Escenario Prueba - Migración Máquina Virtual En Vivo

### **ETAPA DE ANALISIS:**

En esta sección se demuestra cómo el clúster garantiza alta disponibilidad sin que se incurra en altos costos, respondiendo a la pregunta de investigación planteada.

En la etapa de conceptualización del presente proyecto se referenciaron varias soluciones para la implementación de una plataforma altamente disponible, pero Oracle VM es la más adecuada para el soporte de la plataforma de alta disponibilidad que se pretende instalar dentro de las instalaciones de la Universidad de Cartagena, ya que además de ser completamente gratuita tiene un gran soporte a inconvenientes antes, durante y luego de la implementación por parte de toda la comunidad desarrolladora bajo los productos de Oracle Corporation. Para la implementación de la herramienta de alta disponibilidad Oracle VM, no es necesario contar con máquinas computadoras de última generación, pero si es necesario que sean de núcleos procesadores con versión de virtualización, los cuales vienen fabricándose en su mayoría desde comienzos del siglo XXI, ya sean de proveedores Intel<sup>16</sup> o AMD<sup>17</sup>. Lo anterior hace de la solución Virtual Machine como plataforma idónea según

---

<sup>16</sup> **Intel Corporation** – Fabricante circuitos integrados desde 1968.

<sup>17</sup> **AMD** - Advanced Micro Devices Inc.

los recursos e infraestructura disponible dentro de las instalaciones y laboratorios de sistemas de la Universidad de Cartagena;

Dado un resumen del montaje, configuración e implementación del sistema clúster HA se realizan pruebas de funcionamiento y confirmación de nuestros objetivos. Se demostraran todas las pruebas de interacción con videoclips que simula el procedimiento y funcionamiento del esquema de operación. Los videoclips pueden ser obtenido en las instalaciones del programa de Ingeniería de Sistemas de la Universidad de Cartagena junto con la documentación del presente proyecto, en caso de inconvenientes presentados comunicar con los autores del proyecto vial e-mail.

Es clave y necesario apuntar que alta disponibilidad no es sinónimo de “nunca tener fallas en nuestra plataforma de servicios”; que tenga fallas un sistema de gestión de servicios, no implica que siempre se debe convivir con esa falla y no reparar los medios afectados, al contrario, es el momento que se debe aprovechar para reparar daños y reactivación de máquinas servers y así continuar la operación normalmente teniendo activo el respaldo una vez más.

El caso de estudio para la realización de pruebas está basado en brindar a la Universidad de Cartagena, un modelo de funcionamiento activo para sistemas web prestados a sus estudiantes; es el caso del funcionamiento de la plataforma Moodle para cursos virtuales académicos y a futuro la implementación para la plataforma académica estudiantil; dando solución a posibles fallas de disponibilidad de dichos servicios. Dentro de la máquina virtual que se alojara en los nodos de operación del clúster, se instalaron los servicios web

necesarios para brindar el acceso a los cursos virtuales de la Universidad como lo fueron: Servidor Apache, motor de base de datos MySQL y el proyecto Moodle.

Los recursos lógicos utilizados para la implementación del proyecto clúster de alta disponibilidad han sido del todo gratuitos, es decir que los componentes de la plataforma Oracle VM no tienen ningún valor y pueden ser descargados sin restricciones desde el sitio principal de Oracle.com. Con lo anterior, cumplimos el ítem mencionado en la pregunta de investigación ya que no incurrimos en ninguna inversión para la descarga de los soportes necesarios de instalación. Agregado a esto es válido mencionar que la plataforma Oracle VM brinda todo el soporte de la comunidad Oracle en caso de que se tengan inconvenientes al implementar su solución de alta disponibilidad.

Para la presente etapa se desarrollaron videoclips demostrando pruebas funcionales de éxito, dando cumplimiento al cuarto objetivo específico, dentro de ellos se encontrara una breve configuración y posible réplica de la plataforma Moodle Cursos virtuales de la Universidad de Cartagena, esta plataforma de cursos virtuales web se ha tomado como base de funcionamiento y prestación de servicios del clúster de alta disponibilidad implementado. En el primer test se tiene la plataforma corriendo y abruptamente desconectamos el server en el que se encuentra trabajando la máquina virtual; en ese momento entra a operar el servicio de alta disponibilidad de Oracle VM, testeando si existe otro server disponible en donde pueda hospedar la máquina virtual con los servicios web de Moodle. Encontrado un server, en cuestión de máximo 1 minuto se reactiva el servicio web de la plataforma de cursos virtuales tal cual como se encontraba corriendo anteriormente.

Los videoclips están denominados de la siguiente forma:

- 3. Escenario Prueba - Falla de desconexión o bloqueo de servidor.
- 4. Escenario Prueba - Migración Maq Virtual En Vivo.

El análisis realizado constó en gran parte, para el estudio de escenarios donde la implementación de la plataforma tuviese escenarios exitosos, limitantes de operación o momentos de rápida reconfiguración para reactivar el servicio en el menor tiempo posible. Es en este punto donde citamos 2 escenarios de éxito, 1 escenarios de modificaciones futuras y 1 escenario final como mejora a nuestro modelo y diseño planteado del clúster HA.

***Escenario 1: Fallas en el equipo en donde se encuentra corriendo los Servicios.***

Para el primer escenario, como se expresó anteriormente se coloca a prueba el funcionamiento del clúster mediante una desconexión inesperada del nodo es que se encuentra corriendo la máquina virtual con los servicios web. En el videoclip “Prueba Falla de desconexión o bloqueo de servidor.” Se observa como pierde conexión el nodo numero 2: OVMSERVER02, y en menos de 1 minuto el sistema logra reactivarse chequeando que el nodo numero 1: OVMSERVER01 está disponible y puede alojar la máquina virtual. Por medio del manager se comprueba que la máquina luego de la desconexión apunta al nodo 1.

A continuación se muestra un pantallazo de último estado de la máquina virtual.

Select	Details	Virtual Machine Name	Memory Size (MB)	Status	Owner	Group Name	Server Name	Server Pool Name
<input checked="" type="radio"/>	Show	Server_HA	1,000	Powered Off	admin	My Workspace	N/A	HAPPOOL
<input type="radio"/>	Show	WebServerHA	1,000	Running	admin	My Workspace	QVMSERVER01	HAPPOOL

***Escenario 2: Fallas en el equipo administrador de la máquina virtual – adición de NODO.***

Para este caso el clúster se recupera automáticamente, perdiendo alrededor de 30 segundos para la migración de la maquina hacia otro VM Server disponible y asignado a nuestro pool. La reconexión es posible gracias a la configuración del POOL de servers teniendo alguna de las máquinas que la componen en modo activo. Este procedimiento de migración de una maquina a otra no debe demorar mucho tiempo, de esta forma cumple con los estándares de disponibilidad de servicios. Los resultados y pruebas en vivo de este escenario pueden ser observados en el videoclip “Migración en vivo de máquina virtual.”

***Escenario 3: Fallas en el equipo Master del Pool de Servidores.***

Para este tipo inconvenientes sencillamente debemos tener más de un respaldo en el clúster, es decir al menos 3 máquinas nodos; reconfigurar de manera rápida el pool de servers asignado las características de Master a la máquina nodo disponible para hospedaje de la máquina virtual (huésped de servicios) es otra posibilidad, pero es posible evitarla con la opción anteriormente mencionada. Cabe anotar en este punto, que el Pool de servers, solo podrá existir un solo Server Master, así como muchas máquinas Utilities y Servidores de aplicaciones. De igual forma, es recomendable correr por defecto nuestra máquina virtual

en un equipo que no sea el Master, así libramos que el Server Master sea dependiente del diseño.

***Escenario 4: Falla del soporte NFS o Repositorio Compartido.***

Al momento de realizar las pruebas de posibles fallas de máquinas disponibles se encontró la posible desconexión del Repositorio Remoto por NFS, o lugar en donde se estén guardando los archivos configurados por el Oracle VM Manager.

Para este tipo de fallas se requiere implementar un sistema de alta disponibilidad en datos o repositorios backup para reconexión de servicios prestados. La propuesta es recrear el sistema como una redundancia entre las máquinas que lo componen; de esta forma se podría automatizar un Job<sup>18</sup> periódico, con componente backup del volumen compartido por NFS hacia los nodos de operación. Con este respaldo se busca que en el menor tiempo posible se pueda reanudar nuestro servicio por medio de una conexión local hacia uno de los nodos (implica destrucción de pool\_server configurado y reconfiguración de un pool local), de esta forma la máquina virtual podremos iniciarla desde cualquier nodo que se encuentre disponible en cuestión de minutos, sin necesidad de reconfigurar los servicios web.

La implementación del diseño planteado por medio de un repositorio lógico remoto, nos permite tener la posibilidad de reactivar nuestro servicio con la instalación de solo un equipo nodo nuevo; ya que toda la configuración de servicios y datos se tiene dentro del archivo .img (máquina virtual) guardada en el directorio /sharedvol.

---

<sup>18</sup> Job de respaldo servidor: tarea de respaldo en otra locación remota.



Se determinó en el marco teórico que la disponibilidad de servicios ideal es un 99% de funcionamiento, dando un tiempo de inactividad de 32 segundos al año. Si se obedece al funcionamiento ideal de OVM cada falla de la plataforma debería ser solucionada mientras que Oracle VM brinda su solución de disponibilidad, y así tener siempre un respaldo para la plataforma de servicios.

Para la universidad de Cartagena es muy útil la implementación del proyecto HA presentado, ya que múltiples de sus servicios son basados en protocolo web; es el caso de la plataforma virtual de notas, plataforma virtual de cursos académicos, catálogo de la biblioteca, entre otros. De esta forma es visión del presente proyecto, que a futuro se logre una integración de dichos servicios en una infraestructura que brinde un índice de disponibilidad de servicios alto, que ante posibles fallas no se produzcan demoras en la reactivación normal de respuestas por parte de los servidores.

Podemos afirmar que la opción de implementación y solución de disponibilidad presentada, es la que brinda mejores marcas de rendimiento en función de respuesta ante posibles bajas de nodos de operación en un clúster; esto se debe a la presencia de un ente verificador dentro de la infraestructura que siempre está verificando que cada uno de los nodos perteneciente al clúster se encuentra operativo; el sistema de clúster altamente disponible es muy parecido a un Latido de Corazón que confirma que cada estación de computo se encuentra disponible para responder a la petición realizada por un tercero a la IP compartida que puede apuntar a todos las estaciones del Clúster.

## 7. CONCLUSIONES Y RECOMENDACIONES

Por medio de la implementación de un clúster HA, se logró proporcionar alta disponibilidad para los servicios desarrollados en los proyectos web de los cursos del programa de Ingeniería de Sistemas de la Universidad de Cartagena, sin incurrir en altas inversiones en tecnología y altos costos asociados, dando respuesta a la pregunta de investigación planteada. Dicha implementación del clúster implicó: definición de tecnologías a utilizar, plantear y estudiar un diseño, realizar una instalación y configuración de servicios y aplicaciones, y por último realizar pruebas de funcionamiento para posterior análisis de resultados exitosos, de recomendación y/o mejoras. Por lo tanto se cumplieron a cabalidad todos los objetivos planteados como se puede verificar detalladamente en el capítulo de resultados y como se resume a continuación.

En miras de contestar la pregunta de investigación planteada en el presente proyecto, en primera instancia se estudió y revisaron propuestas de implementación de alta disponibilidad y se eligió como apoyo y modelo general para el proyecto la virtualización de máquinas bajo Hypervisor Xen, con Oracle VM. Basados en el funcionamiento general del Oracle VM, se desarrolló un diseño de infraestructura hardware y plataforma lógica según los requerimientos funcionales de este, teniendo en cuenta el posible escenario de implementación que se posee en el Programa de Ingeniería de Sistemas de la Universidad de Cartagena.

Luego se planteó un plan de operación e instalaciones obedeciendo al diseño presentado; en donde una a una las máquinas pertenecientes al clúster obtienen un sistema operativo instalado y se instala el ambiente de alta disponibilidad, conectando todos los equipos en un

medio activo de red local. Finalmente se realizaron pruebas de funcionamiento y se obtuvieron varios escenarios de resultados exitosos y otros de mejoras en el futuro que hacen parte de todo el análisis de la plataforma clúster que se propuso, diseñó e implementó bajo la metodología de estudio propuesta. Lo anterior concuerda con el cumplimiento de todos los objetivos planteados para el desarrollo e implementación de una plataforma altamente disponible para prestar servicios web para la Universidad de Cartagena.

También se concluye que una implementación tecnológica implica: una investigación y toma de decisiones técnicas y administrativas para lograr el diseño propuesto, que es evaluado ante un caso de estudio específico al cual se le realizaran pruebas de funcionamiento, para un posterior análisis de resultados; además del proceso de configuración de dispositivos o entidades computacionales. En el presente proyecto se planteó un caso de estudio en la Universidad de Cartagena, basado en los servicios web utilizados por los estudiantes, el cual llevó a realizar una investigación sobre cómo brindar alta disponibilidad de dichos servicios web con los recursos que cuenta el programa de Ingeniería de Sistemas; se planteó un diseño y se realizaron las respectivas instalaciones y configuraciones que en su agrupación generaron el resultado de tener una plataforma clúster altamente disponible para gestión de proyectos y aplicativos de la Universidad de Cartagena, sin incurrir en costos adicionales.

En cuanto a limitaciones para la implementación de la solución de alta disponibilidad con Oracle VM, se puede confirmar que según la bibliografía utilizada y los procedimientos planteados en cada uno de los pasos de la metodología del proyecto, no se tuvieron

limitaciones funcionales ya que se logró el montaje de la plataforma clúster y se realizaron pruebas y análisis de casos de uso que confirman el buen funcionamiento de la misma; ahora bien, los inconvenientes encontrados durante la búsqueda de la implementación del clúster en las instalaciones de la Universidad de Cartagena son catalogados como limitaciones ajenas al proceso de implementación, ya que se cuenta con los recursos necesarios para el montaje del clúster en los laboratorios del programa de Ingeniería de Sistemas. Cabe anotar, que la alta disponibilidad que sugerimos es basada en cuanto a tiempo de funcionamiento de la plataforma sería recomendable, realizar a futuro un indicador de tiempo de disponibilidad del servicio en relación a la tasa de disponibilidad descrita en el marco teórico, dejando la plataforma trabajar al menos 30 días continuos.

Con la idea de mantener el servicio disponible en todo sentido y aumentar la disponibilidad física para el proyecto de clustering de alta disponibilidad, se tiene la opción de adicionar a su infraestructura un sistema de respuesta alterna de alimentación eléctrica, basada en soporte de UPS's y así mantener los equipos o nodos encendidos en caso de tener fallas eléctricas durante la operación. La implementación de estos equipos de respaldo eléctrico aumentaría los requerimientos hardware en su implementación y vale la pena estudiar el caso, pero la visión del proyecto está dirigida a la funcionalidad del sistema lógico y se toma como objetivo secundario, pero no menos importante, la integridad física de equipos.

Se recomienda al Programa de Ingeniería de Sistemas de la Universidad de Cartagena utilizar el clúster para el despliegue de los proyectos académicos realizados por los estudiantes, teniendo en cuenta que se debe contar con un responsable de la administración

y mantenimiento del clúster. Igualmente, se recomienda el desarrollo de futuros proyectos sobre esta temática, que permitan fortalecer la infraestructura tecnológica del programa de Ingeniería de sistemas, para que esté acorde a las exigencias del momento.

Con el desarrollo de este proyecto se beneficia la comunidad académica del Programa de Ingeniería de Sistemas de la Universidad de Cartagena, porque sin ser novedoso, se convierte en una herramienta que permite enriquecer el proceso de aprendizaje académico en la medida que el estudiante cuenta con ambientes que son acordes a los escenarios estándares de producción académica (Proyectos de aula destinados a servicios de la propia Universidad), a los que se enfrentara en su vida laboral; así como también los grupos de investigación del Programa contarán con la herramienta donde tendrán centralizados los proyectos desarrollados, lo cual aumenta la eficiencia en el seguimiento y monitorización de los mismos proyectos, además de contar con mejoras en un futuro por parte de otro cuerpo de estudiantes que demuestren interés por mejorar las aplicaciones y/o desarrollos que se encuentren involucradas dentro la plataforma clúster que se desarrolló.

## 8. BIBLIOGRAFÍA

- [1] Pablo Galdámez Saiz. Hidra: Una Arquitectura para Alta Disponibilidad en Sistemas Distribuidos. Soporte a Objetos. Tesis doctoral departamento de sistemas informáticos y computación, Universidad Politécnica de Valencia. Pag 4,5. 2001.
- [2] Bookman Charles, Linux Clustering (Building and Maintaining Linux Clusters), Ed. NEW RAIDERS, 2004.
- [3] Díaz, G., Chaves J., Mendoza V., y Otros. Adaptación de Clusters de Linux para Servicios de Redes. Universidad de Los Andes, Mérida 5101, Venezuela. 2004.
- [4] Aguilar, V. Clustering de Alta Disponibilidad bajo GNU/Linux. 2001.
- [5] Plaza, E. Cluster Heterogéneo de Computadoras. Enero 24 de 2002.
- [6] Rodríguez, J., Monroy, M. Gestión Empresarial con Tecnologías Open Source. Enero 2009.
- [7] Enciclopedia electrónica de términos computacionales. <http://www.webopedia.com>. Visada 11 junio de 2010.
- [8] Mateau, C., Desarrollo de aplicaciones Web, UOC, Marzo de 2004.
- [9] Cáceres, J., Medina, C. Implementación de un Servidor Web Apache Sobre un Cluster Linux. Universidad Católica del Maule. Talca 2007.
- [10] CLASIFICACION DE PYMES, Ley MiPYME 905 del 2 de Agosto de 2004, <http://www.businesscol.com/empresarial/pymes/>, Visitado el 25 de mayo de 2010.

- [11] Test de stress en servicios de red, Apache. <http://jakarta.apache.org/jmeter/>
- [12] Doreen L. Galli. Distributed Operating Systems: Concepts and Practice. Prentice Hall, Inc, Upper Saddle River, New Jersey, USA, Marzo 1993.
- [13] [http://httpd.apache.org/ABOUT\\_APACHE.html](http://httpd.apache.org/ABOUT_APACHE.html). Visitada 2 agosto de 2010.
- [14] Charles M. Kozierok. No Starch Press, The TCP/IP Guide: A Comprehensive, Illustrated Internet Protocols Reference., Inc. Publicado en 2005, Seccion 3, parte 7.
- [15] Oracle VM, <http://www.oracle.com/technetwork/server-storage/vm/overview/index-160875.html>. Visitada Enero de 2014.
- [16] HA-OSCAR, High Availibity Open Source Cluster Application Resources, <http://xcr.cenit.latech.edu/ha-oscar/>
- [17] Kimberlite, HA Cluster for Linux - <https://lists.debian.org/debian-devel/2001/01/msg00909.html>
- [18] SteelEye™ LifeKeeper®, [http://www.ha-c.org/high\\_availability/components/application\\_availability/cluster/high\\_availability\\_cluster/steel\\_eyes\\_lifekeeper/](http://www.ha-c.org/high_availability/components/application_availability/cluster/high_availability_cluster/steel_eyes_lifekeeper/)
- [19] Keep Alived Loadbalancing & High-Availability, <http://www.keepalived.org/>
- [20] DRBD Software Developement for High Availability Clusters, <http://www.drbd.org/>
- [21] FreeNAS, <http://www.freenas.org/for-business/>

[22] OpenFilter, unified storage. <https://www.openfiler.com/>

[23] NAS Lite-2, File Server. <http://freecode.com/projects/naslite>