

**PROPUESTA DE SOLUCIÓN DE ALTA DISPONIBILIDAD DE LOS
SERVICIOS CRÍTICOS DEL CENTRO DE DATOS DE LA
UNIVERSIDAD DEL CAUCA**



Universidad
del Cauca

**Sandra Milena Pantoja Cárdenas
Juan Javier Imbachí Patiño**

**Universidad del Cauca
Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Telecomunicaciones
Grupo de I+D Nuevas Tecnologías en Telecomunicaciones
Popayán, Julio de 2010**

**PROPUESTA DE SOLUCIÓN DE ALTA DISPONIBILIDAD DE LOS
SERVICIOS CRÍTICOS DEL CENTRO DE DATOS DE LA
UNIVERSIDAD DEL CAUCA**



Universidad
del Cauca

**Sandra Milena Pantoja Cárdenas
Juan Javier Imbachí Patiño**

**Trabajo de grado presentado como requisito para optar al título de
Ingeniero en Electrónica y Telecomunicaciones**

**Director
Ingeniero ALEJANDRO TOLEDO TOVAR**

**Universidad del Cauca
Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Telecomunicaciones
Grupo de I+D Nuevas Tecnologías en Telecomunicaciones
Popayán, Julio de 2010**

TABLA DE CONTENIDO

INTRODUCCIÓN	1
CAPÍTULO 1: ALTA DISPONIBILIDAD PARA LOS SERVICIOS DE UN CENTRO DE DATOS	3
1.1. CONCEPTOS RELACIONADOS CON ALTA DISPONIBILIDAD.....	3
1.1.1. Definición del concepto de Disponibilidad.....	4
1.1.2. Definición del Concepto de Alta Disponibilidad.....	5
1.1.3. Normas de referencia para alta disponibilidad de los servicios.....	6
1.2. IMPORTANCIA DE LA IMPLEMENTACIÓN DE ALTA DISPONIBILIDAD EN CENTROS DE DATOS.....	8
CAPÍTULO 2: EL CENTRO DE DATOS UNIVERSITARIO	12
2.1. DEFINICIÓN.....	13
2.2. MISIÓN.....	14
2.3. OBJETIVOS.....	14
2.4. COMPONENTES DE UN CENTRO DE DATOS.....	15
2.4.1. Equipos e Infraestructura de TI.....	15
2.4.2. Recurso humano.....	15
2.4.3. Servicios y aplicaciones.....	16
2.5. LOS SERVICIOS DEL CENTRO DE DATOS.....	16
CAPÍTULO 3: GUÍA METODOLÓGICA PARA EL DESARROLLO DE UNA SOLUCIÓN DE ALTA DISPONIBILIDAD PARA LOS SERVICIOS CRÍTICOS EN CENTROS DE DATOS UNIVERSITARIOS	18
3.1. FASE I: Visión clara del proyecto de Alta Disponibilidad.....	20
3.1.1. Paso 1. Recopilación de conceptos.....	20
3.1.2. Paso 2. Identificación de la necesidad de Alta Disponibilidad.....	21
3.1.3. Paso 3: Conformación de un grupo de trabajo.....	22
3.1.4. Paso 4. Elaboración de un Plan de trabajo.....	22
3.2. FASE II: Especificación del Entorno.....	23
3.2.1. Paso 1: Reconocimiento de las áreas relacionadas.....	23
3.2.2. Paso 2. Identificación del Tier de Disponibilidad de la infraestructura del CDU.....	24
3.3. FASE III: Clasificación de los servicios críticos del CDU.....	30
3.3.1. Paso 1. Actualización de la Documentación de los servicios del CDU.....	30
3.3.2. Paso 2. Recolección de Estadísticas de funcionamiento de los servicios.....	31
3.3.3. Paso 3. Análisis de funcionamiento de los servicios.....	32
3.3.4. Paso 4. Determinación de los servicios críticos.....	32
3.4. FASE IV: Estado del Arte de las Tecnologías de HA.....	33
3.4.1. Paso 1. Clasificación de las Tecnologías de HA.....	33

3.5. FASE V: Definición del Sistema para Alta Disponibilidad	35
3.5.1. Paso 1. Establecimiento de bases conceptuales sobre mecanismos implícitos en un sistema HA.	35
3.5.1.1. Virtualización.....	35
3.5.2. Paso 2: Identificación de aspectos a tener en cuenta al elegir el sistema HA..	38
3.5.3. Paso 3. Realizar un inventario de componentes involucrados en la implementación del sistema HA.....	38
3.5.4. Paso 4: Definición de la propuesta de solución acorde a las condiciones actuales del CDU.....	43
3.5.5. Paso 5: Profundizar en las tecnologías involucradas en la propuesta.	43
3.6. FASE VI: Implementación del SW y HW de la propuesta.	43
3.6.1. Paso 1: Definir requerimientos hardware y software para la implementación..	43
3.6.2. Paso 2: Capacitación del personal.	43
3.6.3. Paso 3: Instalación y configuración del sistema de almacenamiento	44
3.6.4. Paso 4: Realizar endurecimiento del servidor.	44
3.6.5. Paso 5: Instalación del Software HA.	46
3.6.6. Paso 6: Instalación y configuración de los servicios.....	46
3.7. FASE VII: Evaluación de la disponibilidad	46
3.7.1. Paso 1. Definición de escenarios de pruebas.	46
3.7.2. Paso 2. Medición de disponibilidad en el escenario de pruebas.....	47
3.7.3. Paso 3. Análisis de Resultados.....	47

CAPÍTULO 4: DESARROLLO DE UNA SOLUCIÓN DE ALTA DISPONIBILIDAD ACORDE CON LAS CONDICIONES ACTUALES DEL CENTRO DE DATOS DE LA UNIVERSIDAD DEL CAUCA 50

4.1. FASE I: Visión clara del proyecto de Alta Disponibilidad	50
4.1.1. Paso 1. Recopilación del concepto de Alta Disponibilidad.	50
4.1.2. Paso 2. Identificación de la necesidad de Alta Disponibilidad.	51
4.1.3. Paso 3. Conformación de un grupo de trabajo.	52
4.1.4. Paso 4. Plan de trabajo.....	52
4.2. FASE II: Especificación del Entorno	53
4.2.1. Paso 1. Reconocimiento de áreas relacionadas en el CDU de la Universidad del Cauca.....	53
4.2.2. Paso 2. Identificación del tier de disponibilidad de la infraestructura del CDU de la Universidad del Cauca.	55
4.3. FASE III: Clasificación de los servicios críticos del CDU de la Universidad del Cauca.....	63
4.3.1. Paso 1. Actualización de la documentación de los servicios del CDU.....	63
4.3.2. Paso 2. Recolección de Estadísticas de funcionamiento de los servicios del CDU de la Universidad del Cauca.....	63
4.3.3. Paso 3. Análisis de fallas en el funcionamiento de los servicios del CDU de la Universidad del Cauca.....	65
4.3.4. Paso 4. Determinación de los servicios críticos que necesitan una Solución de Alta Disponibilidad en el CDU de la Universidad del Cauca.	66
4.4. FASE IV: Estado del Arte de las Tecnologías de HA existentes	67

4.5. FASE V: Definición del Sistema para Alta Disponibilidad	71
4.5.1. Paso 1: Establecimiento de bases conceptuales sobre mecanismos implícitos en el sistema HA.....	71
4.5.1.1. Clúster de Alta Disponibilidad [55][56].....	71
4.5.2. Paso 2. Aspectos que se tuvieron en cuenta al elegir el sistema HA.	72
4.5.3. Paso 3: Inventario de componentes disponibles en la implementación del sistema HA.....	73
4.5.4. Paso 4. Definición de la propuesta de solución acorde a las condiciones actuales del CDU.....	74
4.5.5. Paso 5. Profundizar en las tecnologías involucradas en la propuesta.	78
4.6. FASE VI: Implementación del SW y HW de la propuesta.	81
4.6.1. Paso 1: Requerimientos hardware y software para la implementación de la propuesta.....	81
4.6.2. Paso 2: Capacitación del Personal.....	82
4.6.3. Paso 3: Instalación y configuración del sistema de almacenamiento.	82
4.6.4. Paso 4: Endurecimiento del servidor.....	83
4.6.5. Paso 5: Instalación del software Open HA Cluster.	89
4.6.6. Paso 6: Instalación y configuración de los servicios.....	90
4.7. FASE VII: Evaluación de la disponibilidad	90
4.7.1. Paso 1. Definición de escenarios de pruebas.	90
4.7.2. Pasos 2 y Paso3: Medición de disponibilidad en el escenario de pruebas y Análisis de resultados.	93
4.7.2.1. Pruebas de Disponibilidad:.....	93
4.7.2.2. Pruebas con otras herramientas incluidas en el Sistema de Alta Disponibilidad implementado:	97
4.7.2.3. Evaluación de la Disponibilidad.....	102
CAPITULO 5: CONCLUSIONES Y TRABAJOS FUTUROS	105
5.1. CONCLUSIONES.....	105
5.2. TRABAJOS FUTUROS	106
BIBLIOGRAFÍA	107

LISTA DE TABLAS

Tabla 1. Porcentajes de Alta disponibilidad.....	5
Tabla 2. Sectores que sufren más interrupciones	9
Tabla 3.1 Grupo de Trabajo	22
Tabla 3.2. Ejemplo de actividades relacionadas al plan de trabajo.	23
Tabla 3.3. Requisitos para el centro de datos según anexo g del estándar ANSI/TIA 942 30	
Tabla 3.4. Clasificación de tecnologías HA.	35
Tabla 3.5. Sistemas de almacenamiento	42
Tabla 4.1. Grupo de Trabajo.	52
Tabla 4.2. Plan de Trabajo.....	53
Tabla 4.3. Subsistema de Telecomunicaciones.	57
Tabla 4.4. Subsistema Estructural y Arquitectónico.	58
Tabla 4.5. Subsistema Eléctrico.....	60
Tabla 4.6. Subsistema Mecánico CDU Unicauca.	62
Tabla 4.7. Clasificación Tier CDU Universidad del Cauca.....	62
Tabla 4.8. Recopilación Tecnologías HA.	71
Tabla 4.9. Plan de capacitación.	82
Tabla 4.10. Descripción de los equipos involucrados en las pruebas.....	92
Tabla 4.11. Resumen de resultados de interrupción del servicio.....	97

LISTA DE FIGURAS

Figura 3. 1 Fases del desarrollo de una solución de Alta Disponibilidad.	19
Figura. 3.2. Esquema General del desarrollo de una solución de Alta Disponibilidad para los servicios críticos de un CDU.....	49
Figura 4.1. Monitoreo de Disponibilidad de los servicios con Nagios.	64
Figura 4.2. Propuesta de solución acorde a las condiciones actuales del CDU.	75
Figura 4.3. Red SAN-iSCSI.....	76
Figura 4.4. Esquema del sistema HA.....	78
Figura 4.5. Actualización del sistema.....	83
Figura 4.6. Borrar información del sistema.....	83
Figura 4.7. Borrar mensaje de versión del sistema.	84
Figura 4.8. Servicios y puertos abiertos.....	84
Figura 4.9. Procesos de network corriendo.....	85
Figura 4.10. Deshabilitar bind.	85
Figura 4.11. Cambio de puerto.....	86
Figura 4.12. No permitir logueo de root.....	86
Figura 4.13. Configuración usuarios permitidos.	86
Figura 4.14. Prohibición de logueo remoto al usuario root.	87
Figura 4.15. Bloqueo de cuentas administrativas.....	87
Figura 4.16. Archivo deny.....	88
Figura 4.17. Archivo allow.....	88
Figura 4.18. Ejemplo de acceso denegado.....	88
Figura 4.19. Registro en Sun.....	89
Figura 4.20.Cluster HA instalado en servidor Kain.....	90
Figura 4.21. Escenario de pruebas.....	93
Figura 4.22. Generación de tráfico con Siege.....	94
Figura 4.23. Herramientas de prueba actuando sobre los nodos del clúster.....	94
Figura 4.24. Tiempo de interrupción prueba 1.....	95
Figura 4.25. Tiempo de interrupción prueba 2.....	95
Figura 4.26. Tiempo de interrupción prueba 3.....	96
Figura 4.27. Tiempo de interrupción prueba 4.....	96
Figura 4.28. Ping desde os-ohac-1 a os-ohac-2.....	97
Figura 4.29. Captura de los mensajes de ping en os-ohac-2.....	98
Figura 4.30. Utilización de wget para descargar una página web mediante HTTP.....	98
Figura 4.31. Captura de mensajes HTTP de descarga de una página web en el nodo2 ..	98
Figura 4.32. Captura del comando ping cifrado.....	99
Figura 4.33. Captura del comando wget cifrado.....	99
Figura 4.34. Creación de una copia de seguridad del sistema actual.....	100
Figura 4.35. Inicio del sistema con diferentes copias de seguridad.....	100
Figura 4.36. Lista de las copias de seguridad del sistema.....	101
Figura 4.37. Ejecución de prstat en el nodo 1.....	101
Figura 4.38. Ejecución de prstat en el nodo 2.....	102
Figura 4.39. Tiempo de interrupción mínima en Ragnarok.....	102
Figura 4.40. Caída de java en Ragnarok.....	103
Figura.4.41. Log de caída de Java en Ragnarok.....	103

INTRODUCCIÓN

Actualmente, cada día más universidades e instituciones universitarias reconocen la importancia y sobre todo los beneficios que aportan las Tecnologías de la Información (TI) para dar soporte a procesos institucionales y satisfacer requisitos de calidad, para optimizar el uso de los recursos de TI, incluyendo servicios, información, infraestructura y personas. Por tal motivo estas instituciones han implementado servicios de TI, sobre los cuales descargan la responsabilidad de procesos, convirtiendo dichos servicios en un elemento vital para el desarrollo de funciones administrativas, académicas e investigativas.

Bajo este panorama entra a la escena el Centro de Datos Universitario (CDU), como el ente institucional responsable de la administración de los servicios de TI para el beneficio de la comunidad universitaria. Además, es la administración del CDU la responsable de fijar las políticas y estrategias de TI hacia futuro, de tal manera que se aprovechen las ventajas que ofrecen dichos servicios. Sin embargo el acceso de las universidades a la tecnología que hace posible todos estos beneficios, está limitado por factores como el económico y el desconocimiento de estrategias al respecto, principalmente. Esta situación ha conllevado al deterioro en la calidad de los servicios de TI, lo que se refleja en indisponibilidad de los mismos a causa de la interrupción cada vez más frecuente de algunos de ellos, sin que haya una solución viable que permita hacer frente a este problema. Es así como surge la Propuesta de Solución de Alta Disponibilidad para los Servicios Críticos de un Centro de Datos Universitario (CDU), mediante la cual se busca plantear una solución que esté alcance de una institución universitaria.

La propuesta se centra en el concepto de Alta Disponibilidad (High Availability-HA) para los servicios más importantes del CDU, los cuales se pueden catalogar como críticos. Por tanto lo que se busca es minimizar el impacto de las interrupciones tanto planeadas como no planeadas para dichos servicios mediante una estrategia eficiente, efectiva y cuya relación costo-beneficio esté acorde a las condiciones de un CDU. Para lograr dicho objetivo se hace uso de tecnologías complementarias a la HA disponibilidad como virtualización, clustering, raid y espejado de datos, entre otras. Y de mecanismos como la redundancia, eliminación de puntos únicos de fallo (Single Point of Failure-SPOF), estrategias de conmutación por error (failover) y monitoreo de “salud” de servicios y nodos. (Heartbeat), por mencionar los más relevantes, ya que son muchos más.

Adicional a la propuesta se incluye la generación de una guía metodológica, la cual sirve como herramienta para la búsqueda, planteamiento y desarrollo de la propuesta de solución. La guía surge como el resultado de la investigación sobre centros de datos universitarios acerca de las necesidades y problemáticas que enfrentan respecto a la disponibilidad de los servicios. Por tanto, es documento con un enfoque práctico que integra los aspectos más importantes que es necesario tener en cuenta en el momento de buscar una solución de HA para los servicios críticos del CDU.

Acorde con lo mencionado anteriormente, para el logro de los objetivos propuestos en el anteproyecto, el desarrollo del documento del proyecto se realiza mediante cinco capítulos, cuyo contenido general se detalla a continuación:

- **Capítulo 1:** En este capítulo se realiza una introducción respecto al concepto de alta disponibilidad, definiendo su significado, estableciendo sus implicaciones, ventajas y desventajas de su aplicación. Además se enfatiza en la importancia de la alta disponibilidad para los servicio de TI del CDU, desde la perspectiva de normas como ISO/IEC 20000, y de los marcos de trabajo ITIL y COBIT.
- **Capítulo 2:** Como parte del enfoque del proyecto hacia el centro de datos, en este capítulo se muestra de manera general lo que es un Centro de Datos Universitario (CDU). Para ello se abordan aspectos como su definición, misión visión y tipo de servicios ofrecidos entre otros.
- **Capítulo 3:** Este capítulo corresponde específicamente al documento de la guía metodológica propuesta, mediante la cual se busca el desarrollo de la propuesta de solución. Por lo tanto, en él se detalla el contenido de la guía a través de las fases y pasos que la integran.
- **Capítulo 4:** Corresponde al componente práctico del proyecto, ya que es el resultado de la aplicación de la guía metodológica al caso de estudio. En este capítulo se desarrolla la propuesta de solución de alta disponibilidad para el servicio crítico web del Centro de Datos de la Universidad del Cauca (CDUC).
- **Capítulo 5:** Como resultado de la investigación el capítulo final contiene las conclusiones, aportes y el planteamiento de de trabajos futuros.

Para complementar el contenido de los capítulos, se cuenta además con los siguientes anexos:

- **Anexo A: Justificación del Proceso de Desarrollo de la Guía Metodológica.** En este anexo, se justifica la necesidad de la guía y se explica de manera general el proceso de desarrollo de la misma.
- **Anexo B: Encuestas.** Es el anexo que muestra el resultado de las encuestas realizadas como parte del proceso de investigación. Además incluye el análisis de los resultados de las encuestas.
- **Documento C:** El Centro de Datos de la Universidad del Cauca. Este documento contiene información respecto a las áreas que componen el CDUC. Se hace énfasis en el Área de Servidores y Servicios, dada su relación con la temática del proyecto. Esta información se genera como parte de la aplicación de la guía metodológica en el proceso de desarrollo de la propuesta de solución del HA.
- **Anexo D:** Implementación del prototipo de solución. Documento técnico de instalación y configuración de todo lo referente al prototipo de solución.

CAPÍTULO 1

ALTA DISPONIBILIDAD PARA LOS SERVICIOS DE UN CENTRO DE DATOS

La necesidad de alta disponibilidad para los servicios del centro de datos no es reciente, ha estado ahí desde hace ya algún tiempo y ahora su importancia se ha incrementado hasta volverse indispensable, tanto así que si los servicios de Tecnologías de la información (TI) se detienen, la institución que hace uso de ellos, prácticamente también se detiene [1]. Es así como en la actualidad las empresas e instituciones, como es el caso de las universidades, con el fin de dar apoyo y soporte tecnológico a sus procesos internos, se han visto obligadas a hacer uso de los servicios de TI, convirtiéndolos de esta manera en una herramienta clave [2]. Pero ahora, que muchos centros de datos disponen de dicha herramienta, se hallan ante un nuevo desafío, brindar unos servicios de TI de calidad, aumentando su disponibilidad y a un costo aceptable, lo cual es un gran reto actualmente, ya que el mercado no ofrece unas verdaderas opciones de solución que estén al alcance de todo tipo de centros de datos, especialmente de los centros de datos universitarios.

Como respuesta a la anterior situación, se plantea una solución que busca la integración del concepto de alta disponibilidad enfocado en los servicios del centro de datos institucional. Sin embargo, la alta disponibilidad para los servicios de un centro de datos no es un aspecto aislado, puesto que depende de muchos factores y condiciones relacionadas con dichos servicios y su entorno, que es muy importante conocer para saber exactamente de lo que se está hablando y el compromiso que implica asumir un proyecto de este tipo. Cuando la alta disponibilidad es un objetivo a lograr, no se trata sólo de implementar “soluciones” de TI de manera mecánica, lo que se busca en primer lugar es contar con una perspectiva clara de lo que se quiere y cómo se va a alcanzar, lo cual sólo es posible mediante una adecuada planeación que permita comprender los conceptos que maneja la alta disponibilidad y los pasos a seguir para dimensionar el alcance de una posible solución [3].

En este capítulo se articula el marco teórico alrededor del concepto más importante para el proyecto, la alta disponibilidad aplicada a los servicios y su contexto, como lo son: el centro de datos, su infraestructura de TI y la normatividad existente al respecto, buscando generar un marco teórico formal que aporte los conceptos necesarios para el proyecto y posibilite el desarrollo de una propuesta de solución de alta disponibilidad, mediante la generación de una guía metodológica.

1.1. CONCEPTOS RELACIONADOS CON ALTA DISPONIBILIDAD.

Cuando se busca información conceptual respecto al tema de disponibilidad, es posible darse cuenta que no existe una única definición, pues es un concepto muy amplio que

puede aplicarse en muchos campos y áreas, y dependiendo del enfoque, será la definición o concepción. Por esta razón se hace necesario establecer las definiciones y conceptos relacionados con la temática de alta disponibilidad.

1.1.1. Definición del concepto de Disponibilidad.

La disponibilidad como concepto básico, es un indicador que permite expresar cuantitativamente la capacidad de un sistema, aplicación o servicio de estar funcionando durante un porcentaje de tiempo determinado, expresado normalmente en minutos u horas respecto a un intervalo de tiempo [4]. Otra manera de concebirla, es a través de la indisponibilidad, como el porcentaje de tiempo o número de minutos de interrupción del servicio o aplicación, que es posible tener como usuario.

El cálculo formal de la disponibilidad está expresado de manera general mediante la ecuación básica (Ec.1), la cual relaciona el tiempo promedio entre fallas (MTBF-Mean Time Between Failure) y el tiempo medio para la reparación una vez ocurrida la falla (MTTR-Mean Time To Repair), donde A, representa la disponibilidad (Availability) general [5]:

$$A = \text{MTBF} / (\text{MTBF} + \text{MTTR}) \quad \text{Ec.1. Fórmula básica para el cálculo de disponibilidad.}$$

Como la anterior fórmula es muy general, para el cálculo de la disponibilidad aplicada a los servicios se tendrá en cuenta la fórmula (Ec.2), tomada de ITIL (IT Infrastructure Library), dado que ésta expresa de una manera más pertinente la disponibilidad desde el punto de vista de prestación del servicio, donde es muy importante el tiempo de interrupción del servicio respecto al tiempo acordado del mismo. De esta manera dicho cálculo resulta más práctico, ya que es más fácil conocer un tiempo acordado de servicio y el tiempo de interrupción del mismo, en vez del tiempo entre acciones correctivas o de mantenimiento, que resulta más dispendioso y menos familiar. La fórmula según [6], es la siguiente:

$$\% \text{Disponibilidad} = ((\text{AST} - \text{DT}) / \text{AST}) * 100 \quad \text{Ec.2. Fórmula del cálculo de la disponibilidad de los servicios}$$

Donde, AST (Agreed Service Time), corresponde con el tiempo acordado de servicio y DT (Down Time), es el tiempo de interrupción del servicio durante las franjas horarias de disponibilidad acordadas.

Un sencillo ejemplo práctico de cálculo es, si al tener un servicio 24/7 y en el último mes el sistema ha estado caído durante 3 horas por tareas de mantenimiento, la disponibilidad real del servicio fue:

- $\text{AST} = 24[\text{horas/día}] * 30[\text{días}] = 720[\text{horas}]$, que es el tiempo acordado de servicio.
- $\text{DT} = 3[\text{horas}]$, que corresponde al tiempo de caída o interrupción del servicio.

Por lo tanto, $\% \text{Disponibilidad} = ((720 - 3) / 720) * 100 = 99.583\%$

1.1.2. Definición del Concepto de Alta Disponibilidad

Como concepto central, la definición de alta disponibilidad está en función de la naturaleza de su finalidad, así desde la perspectiva de los servicios del centro de datos, se define como: la capacidad de proporcionar acceso a un servicio o aplicación con un mínimo de interrupciones programadas y no programadas, y en caso de que éstas se produzcan, el tiempo de recuperación del servicio debe ser mínimo, mitigando así el impacto del tiempo de inactividad generado [7]. Una manera práctica de expresar el concepto de HA, consiste en que los servicios y aplicaciones estén funcionando 24 horas al día, 7 días a la semana, los 365 días del año; lo que significa que los servicios y aplicaciones deban estar funcionando durante un alto porcentaje del tiempo programado de servicio. A continuación se muestra una tabla con los porcentajes considerados de alta disponibilidad [8]:

PORCENTAJE	TIEMPO	
	DÍA 24 HORAS	DÍA 8 HORAS
99%	87,6 horas (3,65 días)	29,12 horas (1,21 días)
99.9%	8,76 horas	2,91 horas
99.99%	52,56 minutos	17,47 minutos
99,999% (“cinco nueves”)	5,256 minutos	1,747 minutos
99.9999%	31,536 segundos	10,483 segundos

Tabla 1. Porcentajes de Alta disponibilidad.

La alta disponibilidad no es un concepto aislado, existen otros factores con los cuales guarda una estrecha relación, y que es muy importante tener en cuenta, entre los principales están los siguientes:

- La gestión de la continuidad del servicio, que se ocupa de impedir la interrupción de los servicios, debido a desastres naturales u otras fuerzas de causa mayor [6].
- La recuperación de desastres, que significa superar las contingencias que se puedan producir, independientemente de su origen y está sustentada por un plan de recuperación, cuyo objetivo es proporcionar los medios alternos para realizar las funciones normales, cuando los medios habituales no están disponibles debido a una contingencia [9].
- La continuidad del negocio: como enfoque comercial, es un proceso integral que incluye todo lo relacionado con los esfuerzos de las empresas para operar armónicamente en cualquier circunstancia, por tanto la alta disponibilidad, la recuperación de desastres y la gestión de la continuidad del servicio son elementos constitutivos de la continuidad del negocio [9].
- La seguridad: es un componente complementario e imprescindible dentro del concepto de disponibilidad y más cuando se habla de alta disponibilidad para los servicios críticos de un centro de datos. Esta importancia se debe a que muchas de las interrupciones en los servicios del centro de datos son causadas por acciones

externas representadas en intrusión, ataques a servidores, actividades de “hacking”, sabotaje o problemas de virus; e internas como parches de actualización mal aplicados o actividades de configuración de dispositivos de red mal realizadas. En fin, existen un gran número de eventos no autorizados o gestionados indebidamente, que ponen en riesgo la seguridad de las aplicaciones y afectan directamente la disponibilidad de los servicios.

La anterior relación de factores permite mostrar en gran parte la variedad de aspectos que alimentan el concepto de alta disponibilidad, por lo que para lograr un alto nivel de disponibilidad de los servicios se debe comenzar por conocer lo que éste significa e implica. Por ejemplo, estrictamente el término alta disponibilidad tiene un significado definido, pues desde el punto de vista comercial, que es donde surge el concepto, consiste en que los servicios estén funcionando 24 horas al día, los siete días de la semana, los 365 días del año, lo cual se representa mediante la expresión 24x7x365, y que implica un porcentaje de disponibilidad dentro del rango 99%-99.999%, siendo habitual la referencia del nivel máximo, conocido como los cinco nueves de disponibilidad [10], como lo muestra la tabla anterior.

Desde luego, un nivel como el mencionado anteriormente sería el objetivo ideal de alta disponibilidad, pero es muy difícil de lograr incluso para centros de datos de grandes empresas, los cuales cuentan con suficientes recursos económicos, dado el alto costo que implica involucrar integralmente todos los aspectos de un centro de datos, como la infraestructura tecnológica y física, el recurso humano y los recursos económicos que sustentan todo y por tanto fijan el límite del alcance de una solución de alta disponibilidad. Además dicho intervalo de disponibilidad está justificado por un enfoque exclusivamente comercial. Por esta razón, este proyecto buscará entre otros objetivos, determinar un nivel de disponibilidad específico para los servicios críticos de un centro de datos de tipo universitario, que permita catalogarlos como de alta disponibilidad. Dicho nivel será establecido con base en los resultados del análisis de disponibilidad actual de los servicios críticos, que permitan realizar una correcta planificación de la disponibilidad y establecer así unos niveles de disponibilidad adecuados, tanto en lo que respecta a las necesidades reales, como a las posibilidades o disposición de recursos actuales de la institución [6]. Ante lo cual, se asume la alta disponibilidad de servicios como: La capacidad de garantizar la entrega del servicio con un mínimo de interrupciones, durante el mayor porcentaje de tiempo necesario, de tal manera que los servicios estén funcionando adecuadamente en el momento que se soliciten por los usuarios [11].

1.1.3. Normas de referencia para alta disponibilidad de los servicios.

Otro aspecto importante de este proyecto de alta disponibilidad, es el enfoque hacia la estandarización, para lo cual se define un marco de referencia ¹ para el desarrollo del proyecto, basado en aquellas normas que tienen una estrecha relación con los centros de datos, sus servicios y la disponibilidad de los mismos, buscando extraer de ellas los requerimientos y aportes sobre aspectos de TI, que deben y pueden ser aplicados a un centro de datos en un entorno universitario.

¹ Apropriación de los temas, para la construcción de la guía metodológica.

Entre las normas de referencia se identifican cuatro principales: ISO/IEC 20000, ISO/IEC 27000, ITIL (IT Infrastructure Library) y COBIT (Control Objectives for Information and related Technology). Estas normas actualmente marcan la pauta mundial en estándares para entornos de TI, enfocándose especialmente en los servicios y su disponibilidad.

ISO/IEC 20000, es el primer estándar específico para la gestión de servicios de TI, su objetivo es aportar los requisitos necesarios, dentro del marco de un sistema completo e integrado, que permita que una organización provea servicios de TI gestionados y de calidad [12]. Es una norma que va dirigida a organizaciones o instituciones que buscan mejorar sus servicios de TI, mediante la aplicación efectiva de los procesos para monitorizar y mejorar la calidad de los mismos. Para lo cual ISO/IEC 20000 cubre aspectos de TI como: el sistema de gestión de servicios, la planificación e implementación de la gestión del servicio, la planificación e implementación de servicios nuevos o modificados, los procesos de provisión de servicio y los procesos de control, entre otros aspectos. La importancia de esta norma es que, es un estándar con una amplia aceptación e implementación lo que muestra que es muy útil.

Dentro de los estándares se tiene además a COBIT. Actualmente es un estándar ampliamente reconocido y aceptado, ofrece un conjunto de “mejores prácticas” para la gestión de los sistemas de información de las organizaciones, enfocadas especialmente en el control de las TI [13]. El objetivo principal de COBIT es proporcionar una guía de alto nivel sobre puntos en los que establecer controles internos que permitan a las organizaciones determinar si están cumpliendo los requisitos identificados como necesarios, mediante la medición y evaluación de los controles de TI. De este modo, ayuda a evaluar con eficiencia la gestión de la infraestructura y la identificación de potenciales riesgos y mejoras [13] [14].

Adicionalmente está ITIL, el cual provee un marco de trabajo público muy importante, que describe las mejores prácticas en la gestión de los servicios de TI, enfocándolas en la mejora continua de la calidad del servicio [13] [15]. La importancia de este marco de trabajo se debe a que el conjunto de “mejores prácticas” que propone, pueden ser adoptadas y adaptadas según las necesidades, circunstancias y experiencias propias de cada organización, permitiendo una flexibilidad de aplicación.

Por último, aunque no menos importante está la familia de estándares ISO/IEC 27000 ISMS (Information Security Management System), que es un conjunto de normas muy completo, enfocado expresamente en la seguridad de la información. Proporcionan un marco de gestión de la seguridad dirigido a ser utilizado por cualquier tipo de organización [16]. Dentro de este conjunto de estándares se destacan por su importancia dos normas, la ISO/IEC 27001, que trata específicamente la gestión de la seguridad de la información, buscando garantizar su confidencialidad, integridad y disponibilidad. También la norma ISO/IEC 27002, el código de buenas prácticas para la gestión de la seguridad de la información, que complementa al anterior.

Respecto a la seguridad de la información, cuyo propósito es contribuir a que los riesgos sean minimizados, evitando de esta manera eventos que puedan provocar indisponibilidad en los servicios del centro de datos [16] [17], es muy importante dejar claro que aunque seguridad y disponibilidad, son dos conceptos complementarios e interdependientes, son claramente diferenciables entre sí, lo que permite su aplicación por

separado. El presente proyecto se enfocada en los servicios, exclusivamente desde el punto de vista de la disponibilidad.

El objetivo de seleccionar las normas más pertinentes o recomendables en el tema de los servicios de TI, es generar un compendio de las “mejores prácticas”, normalizadas oficialmente y que permitan articular un marco normativo de TI, que sirva de referencia para la construcción de una propuesta de solución de alta disponibilidad para los servicios críticos de un centro de datos universitario. La adopción de las normas mencionadas, está determinada por el objetivo que se persigue, por eso es importante entender que la información que éstas aportan debe primero organizarse, entenderse y apropiarse para luego poder aplicarla en función de las necesidades concretas de cada organización o institución [18].

1.2. IMPORTANCIA DE LA IMPLEMENTACIÓN DE ALTA DISPONIBILIDAD EN CENTROS DE DATOS

La expresión, “en todo momento, en todo lugar”, responde a la mentalidad que el mundo tecnológico ha proyectado a los usuarios, especialmente en lo que se refiere a sistemas de información o comunicaciones, entre ellos están los servicios que ofrece un centro de datos, los cuales ahora deben responder a una exigencia de funcionamiento de 24x7, permitiendo tiempos de inactividad muy pequeños, que se están acortando continuamente, debido al constante aumento de la intensidad de los procesos, de la demanda de aplicaciones en tiempo real y en general de los usuarios. La única manera de enfrentar este reto y poder responder adecuadamente, es recurrir a la alta disponibilidad, cuyos beneficios varían de acuerdo con las organizaciones o tipo de actividades, pues “Cada empresa, al final de cuentas, lo va a traducir en un valor único, como reducción de costos, mejora de servicio, en otra seguridad de la información; no siempre es lo mismo” [10].

Un aspecto muy importante que motiva la implementación de alta disponibilidad, es minimizar la interrupción de los servicios. Pues aún con todas las estrategias posibles, un servicio o aplicación puede quedar total o parcialmente no operativo como consecuencia de una falla [19]. Dentro de las interrupciones según [20], se identifican dos tipos principalmente:

- **Interrupciones previstas** (planned outages): que son el resultado de eventos de mantenimiento, reparación, copia de seguridad, u operaciones de actualización. Por ejemplo reemplazar o eliminar componentes defectuosos o actualizar el hardware o software.
- **Interrupciones imprevistas** (unplanned outages): son el resultado de fallos asociados a componentes hardware o software, problemas de seguridad, desastres naturales, virus, accidentes o caídas involuntarias del sistema, entre otros. En definitiva se caracterizan por su naturaleza aleatoria e impredecible.

De estas interrupciones, las más críticas son las imprevistas, que se presentan inesperadamente y pueden provocar un gran impacto si no se cuenta con una estrategia previa para afrontarlas. Este tipo de interrupciones son una de las razones para buscar

una solución de alta disponibilidad, que permita ante todo minimizar su impacto. A continuación de acuerdo a [21] se muestra una tabla con los sectores más afectados por las interrupciones.

Sector	Porcentaje
Banca y Finanzas	26%
Gobierno, Administraciones Públicas e Instituciones	19,1%
Educación	11,3%
Industria	10,9%
Servicios	9,5%
Comunicaciones	8,2%

Tabla 2. Sectores que sufren más interrupciones

Los datos anteriores permiten evidenciar que los sectores de Instituciones y Educación, se encuentran dentro de los tres sectores con mayor impacto a causa de las interrupciones. No sólo se trata de pérdidas económicas, que son las de más repercusión en el caso empresarial, sino además del trastorno en las actividades vinculadas o soportadas por aplicaciones y servicios que se ven interrumpidos, casi siempre el momento más inoportuno. Por ejemplo, en el caso de un centro de datos universitario, cuando se caen los servicios, el impacto económico no es directo, lo cual no quiere decir que no exista. Pues cuando se presenta una interrupción en el desarrollo de actividades académicas, especialmente de investigación y de actividades administrativas, al final se traduce en pérdida de tiempo, retrasos de proyectos, desaprovechamiento de recursos, y todos estos factores terminan generando unos costos adicionales. Por ejemplo, la interrupción de los servicios del centro de datos, que no permite una conexión a una base de datos, la pérdida de un enlace de videoconferencia, la caída de una plataforma o portal, etc., tienen asociados unos tiempos muertos [22] que no se pueden medir con facilidad, pero se sabe que existen y que generan costos adicionales e insatisfacción en los usuarios.

La importancia de considerar las interrupciones, es tener una perspectiva más amplia que ayude a visualizar el costo que representa la indisponibilidad para un centro de datos, respecto al de la implementación de un sistema de alta disponibilidad.

Ahora, la importancia de dicho sistema tiene dos lados, el del usuario y el del centro de datos mismo. Pues no se busca mejorar los servicios, tan sólo para que el usuario esté a gusto, sino para que además con este proceso de mejora obtener una retribución de beneficios para el centro de datos, como son la optimización en la utilización de recursos de TI, la simplificación de los proceso de gestión de los servicios, ahorro en costos de mantenimiento y operación, y modernización de la infraestructura de TI, entre otros. Todos los anteriores beneficios permiten además al centro de datos responder a los retos planteados ante la exigencia de los nuevos servicios, como el incremento de utilización y capacidad de procesamiento principalmente.

En el centro de datos, cuando se trata de alta disponibilidad, la infraestructura es tan importante como los servicios mismos, pues si no se cuenta con las instalaciones adecuadas, no es posible tener servicios disponibles, ni mucho menos pensar en alta

disponibilidad para los mismos. Por tanto, se reconoce la importancia de la infraestructura física y se tendrá presente como un componente de alta disponibilidad de los servicios. Pues aspectos como el diseño, el espacio, los sistemas de cableado, de alimentación, de refrigeración, el alumbrado, los materiales y una gran cantidad de detalles hacen parte del conjunto de elementos que están presentes en el centro de datos desempeñando una función necesaria e indispensable, por lo que una falla en alguno de ellos, puede ser causa de una interrupción.

Reconociendo la importancia de la infraestructura al momento de buscar una solución de alta disponibilidad para los servicios, surge la necesidad de estandarizar la infraestructura del centro de datos, para que esté alineada con una estrategia integral de alta disponibilidad. Al respecto de estandarización, ya existen normas como TIA-942 y la clasificación Tier del Uptime Institute, que brindan los requerimientos y lineamientos necesarios para el diseño e instalación del centro de datos y así poder entregar una mejor disponibilidad en los servicios que presta, basándose en la infraestructura.

De las normas anteriores, la clasificación Tier es muy importante porque expresa la capacidad física de la infraestructura de TI del centro de datos en función del porcentaje de disponibilidad que puede ofrecer. Este es un enfoque importante ya que permite además realizar un autoexamen de clasificación para conocer en qué categoría se encuentra ubicado un centro de datos, como uno de los puntos de partida para un proceso de alta disponibilidad.

Según [23] [24] [25], esta clasificación tiene cuatro niveles:

- **Tier I: Centro de datos básico:** un centro de datos tipo Tier I puede admitir interrupciones tanto planeadas como no planeadas, la infraestructura del centro de datos estará fuera de servicio al menos una vez al año por razones de mantenimiento y/o reparaciones y errores de operación o fallas en los componentes de su infraestructura causarán la interrupción del centro de datos. La tasa de disponibilidad máxima del centro de datos es 99.671% del tiempo, es decir, el nivel Tier I consigue reducir el tiempo de parada a lo largo de un año a 29 horas como máximo.
- **Tier II: Componentes redundantes:** un centro de datos con componentes redundantes es ligeramente menos susceptible a interrupciones, tanto planeadas como las no planeadas. Su diseño es (N+1), lo que significa que existe al menos un duplicado de cada componente de la infraestructura. La tasa de disponibilidad máxima del centro de datos es 99.741% del tiempo, es decir, consigue reducir el tiempo de parada a lo largo de un año a 22 horas como máximo.
- **Tier III: Mantenimiento concurrente:** las capacidades de un centro de datos de este nivel le permiten realizar cualquier actividad planeada sobre cualquier componente de la infraestructura sin interrupciones en la operación. En este nivel, actividades no planeadas como errores de operación o fallas espontáneas en la infraestructura pueden todavía causar una interrupción del centro de datos. La tasa de disponibilidad máxima del centro de datos es 99.982% del tiempo, es decir, consigue reducir el tiempo de parada a lo largo de un año a 1,5 horas como máximo.

- **Tier IV: Tolerante a fallas:** un centro de datos de este nivel provee capacidad para realizar cualquier actividad planeada sin interrupciones en el servicio, pero además la funcionalidad tolerante a fallas le permite a la infraestructura continuar operando aún ante un evento crítico no planeado. Esto requiere cada sistema con un nivel de redundancia N+1. La tasa de disponibilidad máxima es del 99.995% del tiempo, es decir, el nivel Tier IV consigue reducir el tiempo de parada a lo largo de un año a 26 minutos como máximo.

La clasificación anterior describe de manera muy general los requisitos de la infraestructura de un centro de datos para demostrar un porcentaje de disponibilidad dado, el cual debe acompañar la disponibilidad de los servicios, que tienen igualmente sus propios requerimientos. Por tanto, es importante recalcar que la alta disponibilidad que trata este trabajo de grado es exclusivamente para los servicios críticos de un CDU, es decir el área de servidores y sus procesos y aplicaciones, lo cual será explicado posteriormente.

Reconocer la importancia de la alta disponibilidad para los servicios de un centro de datos, es uno de las principales motivaciones que permite iniciar la búsqueda de una solución de alta disponibilidad. Dicha búsqueda comienza con la apropiación del concepto de alta disponibilidad, a través de una investigación detallada que permita tener claro su significado, su función, ámbito de aplicación, requisitos técnicos, costos y los resultados esperados con su aplicación, entre otros aspectos muy importantes que se debe abordar inicialmente. Sin embargo, dado lo extenso del tema no es posible tratar en su totalidad a profundidad todos los temas, por lo que este primer capítulo aporta los elementos básicos que sustentan teóricamente el desarrollo de la búsqueda de una solución de alta disponibilidad para los servicios críticos de un CDU. Además permite adquirir los criterios necesarios para entrar a analizar lo que es el Centro de Datos Universitario en su conjunto, como los es su estructura organizacional, sus funciones, componentes, áreas, servicios, sus retos y necesidades, y en general todo aquello que permita entenderlo mejor y pueda aportar al desarrollo de la solución de alta disponibilidad.

CAPÍTULO 2

EL CENTRO DE DATOS UNIVERSITARIO

Hasta hace algún tiempo, el centro de datos era considerado tan sólo como el lugar donde se almacenan los servidores [26], pero en el contexto actual esta concepción es muy distinta, ahora el centro de datos es mucho más que sólo servidores, pues ha cobrado gran importancia como el motor de las Tecnologías de la Información (TI) para las empresas e instituciones de hoy en día, principalmente para las universidades. Una de las razones importantes es que los datos, aplicaciones y servicios de una u otra manera forman parte de la vida diaria y se han convertido en una necesidad para los usuarios y en una herramienta vital para el desarrollo de la misión institucional. Ante esta perspectiva, los centros de datos universitarios han sentido la necesidad de infraestructuras de TI cada vez mejores en capacidad, desempeño, fiabilidad y disponibilidad, que permitan proporcionar los servicios institucionales claves con un mejor grado de disponibilidad.

Sin embargo, la realidad es que no todos los centros de datos universitarios actuales fueron diseñados para soportar la creciente exigencia en capacidad de procesamiento de información, almacenamiento de datos y disponibilidad de los servicios y aplicaciones [27]. Una causa importante de esta situación es el panorama de tecnologías de la información, que cambia constantemente, evolucionando hacia nuevos servicios, los cuales demandan un alto consumo de recursos como capacidad de cómputo, ancho de banda y almacenamiento, principalmente. Ante lo cual los centros de datos se ven limitados en su capacidad de integrar nuevos servicios o mejorar la calidad de los ya existentes. Situación que obliga a que los centros de datos actuales deban enfrentar y resolver en el corto plazo desafíos como la escalabilidad, disponibilidad y mantenimiento [28]. Tres conceptos claves que deben estar integrados cuando se busca una estrategia de alta disponibilidad para los servicios críticos de un centro de datos de tipo universitario. Pero lograr en el sentido estricto una alta disponibilidad para los servicios críticos del centro de datos universitario es un gran reto, pues representa según informes de empresas con experiencia en este tema, un alto costo económico, dada la inversión necesaria para cubrir todos los aspectos que implica alcanzar una alta disponibilidad.

Por tanto, según lo expuesto anteriormente es importante definir el centro de datos y sus servicios, comenzando desde los conceptos básicos como parte fundamental del enfoque del proyecto. El objetivo de este capítulo es construir el concepto de “Centro de Datos Universitario (CDU)” y caracterizar los servicios y componentes de infraestructura, básicos e indispensables con que cuenta actualmente un centro de datos de este tipo. Buscando de esta manera obtener una perspectiva general de lo que es un CDU, logrando así un acercamiento real a las necesidades y retos, presentes y futuros, para poder proyectar así una guía metodológica que permita implementar una solución de alta disponibilidad apropiada, con una relación costo-beneficio equilibrada. Para lo cual se ha tenido en cuenta la información obtenida de aspectos como la misión, visión, objetivos y funciones de distintas universidades o instituciones universitarias del país como son: Universidad del

Valle², Universidad Nacional³, Universidad del Cauca⁴, Universidad Autónoma del Valle⁵, Universidad Cooperativa de Popayán⁶ y EAFIT⁷.

2.1. DEFINICIÓN

Para la construcción del concepto de centro de datos universitario se parte de la definición general, ya que permite enfocar el concepto en el contexto específico, resaltando el carácter institucional en el que está inmerso y que define la identidad del mismo. De ahí que el concepto no sea único, ya que éste responde en gran parte al enfoque que tenga el centro de datos, sin embargo es posible lograr una definición que recoja lo esencial. Por eso para el presente proyecto, el centro de datos se define como [29] [30]: el conjunto de recursos físicos, lógicos y humanos necesarios para el procesamiento de datos e información, que permiten la organización y control de todas las actividades relacionadas con dicho procesamiento. Para tal fin, el centro de datos desempeña unas funciones básicas como [29]: almacenar, procesar, e intercambiar información, administrar y proveer aplicaciones y servicios tales como alojamiento web, internet, correo electrónico, ftp, telecomunicaciones e información tecnológica, entre otros.

Junto a la definición para el centro de datos, es necesario contar con una clasificación, que permita enmarcar el tipo de centro de datos mediante características específicas como su finalidad, la actividad que desarrolla y el carácter de propiedad, principalmente. Por lo tanto, desde esta perspectiva es posible según [31], clasificar o catalogar los centros de datos en dos grandes categorías:

- **Centros de Datos Corporativos (Corporate Data Centers-CDCs):** este tipo de centros de datos son aquellos que pertenecen a organizaciones o instituciones para uso privado. Por tanto su propósito principal es dar soporte para el procesamiento de datos y servicios para sus propias organizaciones.
- **Centros de Datos de Internet (Internet Data Centers-IDCs):** es el tipo de centros de datos que pertenecen a empresas de telecomunicaciones o proveedores de servicios para ofrecer un servicio público. Su objetivo es proporcionar servicios de tecnologías de la información como acceso a Internet, VoIP, hosting web o de aplicaciones, gestión de servidores, redes de almacenamiento, distribución de contenidos y compartición de carga, entre otros.

Partiendo de la anterior clasificación y el concepto general, es posible ubicar el centro de datos universitario como un tipo de Centro de datos Corporativo o Institucional (CDCs-Corporative Data Center), por lo tanto se puede definir como la unidad institucional universitaria encargada de gestionar y administrar los servicios de telecomunicaciones (datos, voz y video), así como de proporcionar el soporte técnico, administrativo y

² www.univalle.edu.co

³ www.unal.edu.co

⁴ www.unicauca.edu.co

⁵ www.uao.edu.co

⁶ www.uccpopayan.edu.co

⁷ www.eafit.edu.co

operacional de la infraestructura de TI (Tecnología de la información), velando por su buen uso, la mejora continua de los servicios, la integración de nuevas tecnologías y la actualización de los sistemas, todo ello para el beneficio de la comunidad universitaria [32],[33],[34],[35]. Esta definición permite identificar dos elementos muy importantes para el centro de datos universitario, la misión y objetivos, los cuales guían el desarrollo del centro de datos, impulsando la evolución que le permite adaptarse a nuevos escenarios y exigencias para continuar contribuyendo con la misión institucional, como es su fin.

2.2. MISIÓN

La importancia de retomar la misión del centro de datos universitario es para conocer y comprender lo que motiva su surgimiento y lo que espera lograr hacia el futuro la institución que lo posee. Pues al ser el centro de datos parte de una institución, este debe responder a un fin para el cual ha sido creado, es decir la misión que debe cumplir como unidad institucional. Dicha misión es expresada de muchas maneras, pero en lo que respecta a centros de datos universitarios se puede resumir en [32], [33]: Apoyar la gestión académica, investigativa y administrativa de la universidad, utilizando las herramientas que ofrecen las tecnologías de la información.

La anterior descripción de la misión de un centro de datos universitario, es el resultado de la síntesis de la misión de diferentes centros de datos de universidades. En ella se puede percibir la importancia del centro de datos como encargado de apoyar los aspectos académicos, administrativos e investigativos, que son los principales y alrededor de los cuales se integran los otros aspectos.

2.3. OBJETIVOS

Los objetivos son los elementos que contribuyen al desarrollo de la misión y juegan un papel muy importante, ya que pueden ser traducidos en actividades, procesos, tareas y funciones a realizar, lo que permite aterrizar la misión sobre casos concretos. De esta manera los objetivos se pueden enmarcar de manera generalizada en los siguientes [32], [33], [34]:

- Diseñar, desarrollar y prestar servicios de telecomunicaciones a la comunidad universitaria, velando por su correcta implementación y fomentando su aprovechamiento adecuado.
- Velar por el óptimo funcionamiento de la infraestructura de red y los servicios.
- Brindar soporte a nivel de mantenimiento de equipos de cómputo, servicios, sistemas de información y sistemas de comunicación.
- Analizar, evaluar, planear, diseñar y ejecutar los proyectos que favorezcan el desarrollo de los servicios que presta el centro de datos.
- Establecer la normatividad para el uso de los servicios y sistemas de cómputo, así como para la instalación y mantenimiento de la infraestructura.

2.4. COMPONENTES DE UN CENTRO DE DATOS

En realidad el centro de datos universitario más sencillo y básico posible encierra muchos componentes, algunos de los cuales muchas veces pasan desapercibidos. De manera general a partir de la información obtenida de normas, de estándares y de los mismos centros de datos, es posible identificar tres componentes generales con los que comúnmente cuenta un CDU: infraestructura y equipos, recurso humano y los servicios o aplicaciones.

2.4.1. Equipos e Infraestructura de TI

En este aspecto no es fácil encontrar muchos elementos comunes entre los centros de datos universitarios, especialmente en lo referente a lo que las normas y estándares recomiendan. Aunque, en general los equipos e infraestructura de un centro de datos así como los demás componentes no son una “camisa de fuerza”, mas sin embargo se tienen unos elementos que son básicos y necesarios para que un centro de datos pueda desarrollar sus funciones mínimas, entre ellos se espera contar con los siguientes [30] [35] [36]:

- Sistema de comunicaciones de red de alta velocidad, comúnmente LAN (Local Area Network), WAN (Wireless Area Network), CAN (Campus Area Network).
- Infraestructura de cómputo y redes (cableado, fibra, y electrónicos).
- Equipos de red.
- Servidores.
- Sistemas eléctricos de distribución y generación
- Sistema de climatización o control ambiental.
- Sistemas de detección y supresión de fuego.
- Seguridad física y prevención de control de acceso, permisos y logging.
- Iluminación apropiada.
- Tierra física.
- Racks y gabinetes para equipo.
- Canalizaciones: piso falso y bandejas en techo.
- Circuitos y equipo de proveedores.

2.4.2. Recurso humano

En realidad, el centro de datos es mucho más que lo que una definición puede expresar o describir, pues tan importante como los recursos tecnológicos, lo es el recurso humano, el cual es parte vital del centro de datos, como el motor que impulsa la dinámica de los procesos, actividades y operaciones que se deben desarrollar. Este es un componente muy importante para el centro de datos, ya que sobre él recaen las responsabilidades administrativas, operativas y técnicas que hacen posible su funcionamiento.

2.4.3. Servicios y aplicaciones

Como uno de los componentes más importantes del centro de datos, los servicios son el medio a través del cual es posible desarrollar las funciones para las que fue creado. Pues actualmente las instituciones, como son las universidades hacen uso de los servicios de TI para dar el soporte tecnológico necesario para el desarrollo de sus actividades, las cuales dependen ahora de las TI y sus servicios, de ahí la importancia de la disponibilidad de los mismos.

Además los servicios como herramienta tecnológica le ha permitido a las universidades hacer presencia en nuevos escenarios que rompen con los paradigmas tradicionales de educación, como por ejemplo la educación virtual, que hace posible llegar a lugares donde de otra manera no sería posible, alcanzando así una mayor cobertura educativa que trae consigo un gran beneficio social [37].

En general, los servicios son muy importantes para el centro de datos universitario, por eso su tratamiento se amplía en el siguiente apartado, donde se profundiza su conceptualización y contexto, teniendo en cuenta el objetivo del presente trabajo.

2.5. LOS SERVICIOS DEL CENTRO DE DATOS

Cuando se busca una definición de servicio aplicada a TI, es posible descubrir que un servicio es un concepto que por su intangibilidad, es difícil de definir y como el concepto se aplica en muchos campos, su definición es aún mucho más compleja. Sin embargo, desde el punto de vista de las tecnologías de la información un servicio se concibe como “una aplicación de la red de telecomunicaciones que provee un conjunto definido de funciones que interactúan con la red (recursos) y directa o indirectamente con los usuarios de la misma, para satisfacer necesidades de éstos” [31].

Acorde con la definición anterior, en general un centro de datos universitario desempeña muchas funciones y realiza una gran cantidad de actividades para satisfacer las necesidades de la comunidad universitaria, entre las que se incluyen la prestación de varios servicios, muchos de los cuales son transparentes para los usuarios, pero igualmente importantes. Es así como el desarrollo de la investigación ha permitido identificar los servicios más comunes con los cuenta un CDU básico. En algunos casos se tienen más y en otros menos servicios, esto debido a factores como el número de usuarios, el tipo de institución ya sea pública, privada, profesional, tecnológica, una sede principal o una secundaria y el tiempo de vida como institución, entre otros. Pero en general, un centro de dato universitario cuenta con los siguientes servicios básicos:

- Servicio de Internet
- Servicio de correo electrónico
- Servicio Web
- Servicio Ftp
- Servicio VoIP
- Servicio videoconferencia
- Servicio FTP

- Servicio de Proxy
- Servicio DNS
- Servicio DHCP

Los servicios críticos.

El concepto de servicio crítico va más allá y se emplea para catalogar así a aquellos servicios de tecnologías de la información, que son los más importantes para el desarrollo de funciones vitales para la institución, ya sean administrativas y/o académicas. El carácter de “crítico” se debe principalmente al gran impacto que genera su indisponibilidad sobre actividades o necesidades imprescindibles para la comunidad universitaria, como lo son el acceso a la información y medios de comunicación, principalmente [34] [35].

Aunque el concepto parezca general, en realidad la identificación de un servicio crítico es en parte subjetiva, ya que al tratarse de instituciones, no todas tienen las mismas prioridades y necesidades, lo que hace que un servicio que es considerado crítico para una institución, para otra es posible que no lo sea. Por esta razón, no es conveniente definir cuáles son los servicios críticos de un centro de datos universitario, porque se cae en una generalización. Además este es un resultado que se espera conseguir más adelante como parte del proceso de la aplicación de la guía metodológica para la implementación de una solución de alta disponibilidad para los servicios críticos del centro de datos universitario, actividad que se plantea en el capítulo 3 y se desarrolla en el capítulo 4.

Después de ampliar un poco lo que son la misión y objetivos de los centros de datos universitarios, es más claro ver la importancia de los servicios del centro de datos para el desarrollo de la misión institucional y reafirmar así la necesidad de contar una alta disponibilidad en los servicios más importantes o de misión crítica, que permita a las instituciones prepararse para los retos educativos que debe afrontar de cara a la integración de nuevas tecnologías y modelos educativos. De esta manera se prepara el camino para afrontar el desarrollo de la guía metodológica que permita plantear una propuesta de alta disponibilidad para los servicios críticos del CDU.

CAPÍTULO 3

GUÍA METODOLÓGICA PARA EL DESARROLLO DE UNA SOLUCIÓN DE ALTA DISPONIBILIDAD PARA LOS SERVICIOS CRÍTICOS EN CENTROS DE DATOS UNIVERSITARIOS

La prestación de servicios en Centros de Datos Universitarios (CDU), es de vital importancia para su desempeño tanto de la parte administrativa como de la académica, el quehacer de la Universidad en varios de sus campos de acción como el investigativo, el financiero, la docencia, su avance e impacto con nuevos proyectos, se basa fuertemente en la funcionalidad de su Centro de Datos, debido a esta importancia se hace necesario que los servicios estén disponibles el mayor tiempo posible, y que se preste un servicio continuo, lo cual implica estar preparado ante interrupciones no planeadas. Para solventar esta clase de inconvenientes se debe desarrollar⁸ una solución donde se tengan en cuenta el uso de diversas herramientas y mecanismos que permitan mantener disponibles las aplicaciones de dicho servicio ante eventuales fallas.

Desarrollar una solución de Alta Disponibilidad es una tarea compleja, puesto que depende de las necesidades y condiciones de cada centro de Datos. Se necesita tener claros los conceptos y aspectos que están implícitos en una solución de este tipo, tales como requerimientos, áreas relacionadas, planes de trabajo, recursos, pruebas, etc. Muchos aspectos que son importantes y que no deben obviarse para que el desarrollo de una solución sea lo más óptima posible.

En este capítulo, se ha generado una guía metodológica para el desarrollo de una solución de Alta Disponibilidad para los servicios críticos en Centros de Datos universitarios (CDU), la cual tiene como objetivo servir de apoyo y orientación al administrador del CDU, en cuanto a los lineamientos que debe seguir para cumplir esta tarea satisfactoriamente.

El enfoque de esta guía, es hacia los servidores en los cuales se pueden implementar software, mecanismos y estrategias para desarrollar una solución de la problemática anteriormente expuesta.

En esta guía se plantea una serie de lineamientos a seguir, para que el administrador del CDU, pueda lograr una propuesta de solución de Alta Disponibilidad que sea eficaz, ya que las empresas que brindan soluciones y soporte de alta disponibilidad, tienen sus propios métodos de afrontar este tipo de proyectos. El caso general de quienes requieren esta guía, es más para aquellos administradores de CDU's, que no cuentan con el

⁸ Llevar a cabo, realizar una idea, proyecto, etc.

asequibles por la universidad, y permitiéndole un mejoramiento continuo a la disponibilidad de sus servicios.

La guía se compone de siete fases, en las cuales se van dando los pasos específicos a seguir por el administrador del CDU, estos pasos se han establecido de manera secuencial y se observan en la figura 3.1:



Figura 3. 1 Fases del desarrollo de una solución de Alta Disponibilidad.

3.1. FASE I: Visión clara del proyecto de Alta Disponibilidad

3.1.1. Paso 1. Recopilación de conceptos.

El termino Alta Disponibilidad, resulta ser ambiguo y muy amplio, encierra una diversidad de conceptos y requerimientos de acuerdo al enfoque que se le dé. Por tal motivo antes de introducirse en el desarrollo de una solución de Alta Disponibilidad, es sumamente importante conocer lo que significa e implica este concepto y definir a qué se hace referencia cuando se maneja dicho término, ya que desde un principio, se requiere entender las ventajas que trae implementar este tipo de soluciones en el CDU.

Como se mencionó en el capítulo 1, desde la perspectiva de los servicios del centro de datos, la alta disponibilidad se define como: la capacidad de proporcionar acceso a un servicio o aplicación con un mínimo de interrupciones programadas y no programadas, y en caso de que éstas se produzcan, el tiempo de recuperación del servicio debe ser mínimo, mitigando así el impacto del tiempo de inactividad generado [6]. Esto va de la mano con el objetivo de minimizar los puntos únicos de fallo que representan los factores de mayor impacto e importancia en el momento de implementar Alta Disponibilidad.

La anterior definición da una idea fundamental de lo que es Alta Disponibilidad, pero con el desarrollo de este paso se busca tener clara la filosofía que subyace detrás de este concepto, ya que es muy amplio, por lo que es importante conocer el aspecto teórico en profundidad y obtener una perspectiva del concepto con relación al contexto en el que se va a aplicar, cómo contribuye en el desarrollo de la solución que se quiere lograr, y lo que se necesita para alcanzar dicha meta.

El desarrollo de éste paso implica específicamente investigar los aspectos más relevantes, relacionados con Alta Disponibilidad, comenzando por buscar una respuesta a interrogantes, tales como:

1. ¿Qué es y cómo se define?
2. ¿Cuál es su finalidad?
3. ¿Hacia qué aspectos está enfocada?

Para realizar dicha investigación se debe identificar fuentes de información confiables, que garanticen una información precisa. Se recomienda en lo posible ir directamente a las fuentes de la información, que pueden ser organizaciones como IEEE⁹, ITU¹⁰, ANSI¹¹, ISO, TIA, o de empresas reconocidas del sector que abordan la temática desde la experiencia propia, como por ejemplo Sun Microsystem, HP, VmWare, Novell, entre otras. Para afianzar los conceptos teóricos referentes a HA, se puede tomar como documentos base los capítulos uno y dos del presente trabajo de grado.

El resultado de este paso debe ser específicamente la apropiación del concepto de Alta

⁹ Institute of Electrical and Electronics Engineers- www.ieee.org

¹⁰ International Telecommunications Union-www.itu.int

¹¹ American National Standards Institute-www.ansi.org

Disponibilidad. Conseguir este resultado requiere invertir tiempo en investigación, ya que es un producto fundamental que da la base conceptual para el desarrollo del proyecto.

Este es un paso sumamente importante en el proceso de desarrollo de la solución de Alta Disponibilidad, ya que aportará elementos de juicio en el momento de fijar su alcance y ayuda a realizar un mejor planteamiento del problema, centrándolo específicamente en los servicios críticos del CDU, lo cual corresponde al segundo paso a seguir en esta fase.

3.1.2. Paso 2. Identificación de la necesidad de Alta Disponibilidad.

Este paso corresponde a la identificación de la necesidad de alta disponibilidad en los servicios que presta el CDU. Se debe abordar de manera general el grado de indisponibilidad y falencias en los servicios, y describir el problema o necesidad de Alta Disponibilidad, para tal fin se pueden plantear ciertas preguntas tales como:

1. ¿Se presentan interrupciones frecuentes en los servicios que presta el CDU?
2. ¿Tiene identificado los puntos únicos de fallos en los servicios del CDU?
3. ¿Hay quejas por parte de los usuarios, respecto a los servicios prestados por el CDU? (Es importante ayudarse de los registros que se tienen del área de Helpdesk sobre la quejas recibidas)
4. ¿Hay falencias operacionales en los servicios?
5. ¿Se hacen pruebas para mejorar continuamente la disponibilidad de los servicios o solo se espera a tener las falencias? (esto se refiere a que si se resuelve de manera reactiva los problemas o si existen planes de contingencia para solventar problemas en los servicios).
6. ¿El CDU cuenta con mecanismos en los servidores que ayuden a prever la caída inesperada de un servicio?
7. ¿Hay servicios que continuamente se bloquean o están abajo y se necesita personal que esté pendiente y suba manualmente dichos servicios?
8. ¿Se ha hecho una clasificación acerca del Tier o nivel de disponibilidad que tiene actualmente el CDU?
9. ¿Es conveniente que el CDU esté a la vanguardia de las tecnologías?
10. ¿Cuáles son los servicios críticos que necesitan una solución de HA?
11. ¿Hay implementados mecanismos de redundancia en los servidores, que ayuden a minimizar riesgos en la pérdida de información y en la interrupción del servicio?

Las anteriores preguntas planteadas, ayudan a visualizar la necesidad de implementar una solución de Alta disponibilidad en los servicios que presta el CDU. Generalmente este planteamiento debe hacerlo una persona que tenga un buen conocimiento del funcionamiento de los servicios del CDU, como lo es el administrador del centro de datos.

El resultado de este paso es lograr el planteamiento claro del problema que se va a abordar en el proyecto, de tal manera que permita justificar la necesidad de implementar una solución de Alta Disponibilidad y visualizar el camino hacia la construcción de la misma.

3.1.3. Paso 3: Conformación de un grupo de trabajo

Después de plantear el problema, es adecuado inicialmente conformar un grupo de trabajo en el cual se identifique el recurso humano con el que se cuenta para el desarrollo de la solución, el cual debe incluir al personal que se relacione directamente con el CDU, además se hace necesario, definir los roles que deben desempeñar cada uno de los integrantes del grupo, como lo son: un representante de la parte administrativa, un coordinador, diseñadores, ejecutores y un evaluador de la implementación.

ROL	DESCRIPCIÓN
Coordinador	Persona que tenga un conocimiento profundo del funcionamiento de los servicios como lo es el administrador del CDU.
Representante administrativo	Persona encargada de gestionar la parte de los recursos del CDU, por ejemplo el jefe de la división de sistemas.
Diseñador(es)	Personas que tengan conocimiento de la temática o que trabajen en el área, los cuales inicialmente recopilarán la información, para luego dar aportes en la implementación.
Ejecutor(es)	Personas encargadas de ejecutar los procesos derivados de los respectivos diseños de cada una de las fases implícitas en el desarrollo de la solución. Estos podrían ser los mismos diseñadores, quienes pueden cambiar de rol de acuerdo a los conocimientos que estos tengan y a la cantidad de personal que se tenga disponible para llevar a cabo el desarrollo.
Evaluador	Persona encargada de hacer las pruebas respectivas de la implementación, y de hacer auditoría de que se cumplan con los requerimientos iniciales del proyecto.

Tabla 3.1 Grupo de Trabajo

El resultado de este paso es un **grupo de trabajo constituido y organizado** en el cual se definen roles y responsabilidades para el desarrollo de la solución de Alta Disponibilidad.

3.1.4. Paso 4. Elaboración de un Plan de trabajo

Después de tener claro los roles de cada uno de los integrantes del equipo de trabajo, para llevar a cabo de manera eficaz la implementación de una solución de alta disponibilidad se debe generar un plan de trabajo organizado que permita visualizar las tareas y diferentes objetivos en todas las etapas del desarrollo del proyecto, para establecer la manera cómo se abordará el desarrollo de la solución. Además, se recomienda planear reuniones periódicas, para verificar el cumplimiento y avance de las tareas asignadas.

En este paso, se recomienda seguir los puntos expuestos en la tabla 3.2, en la cual se muestra un ejemplo de actividades relacionadas al plan de trabajo:

OBJETIVOS	ACTIVIDADES	TIEMPO ESTIMADO	RESPONSABLE(S)
1-Recopilación de información actual de los servicios del CDU	<ul style="list-style-type: none"> • Recopilar información del hardware. • Recopilar información del software instalado para cada servicio. • Documentar el funcionamiento de los servicios. • Documentación de la configuración de los servicios. 	Dos semanas	Ejecutor 1

Tabla 3.2. Ejemplo de actividades relacionadas al plan de trabajo.

- **Objetivo:** Se refiere a la finalidad que se debe cumplir en cada uno de los pasos del proyecto, hacia la cual deben dirigirse los recursos y esfuerzos del grupo de trabajo. Estos objetivos se logran gracias a las actividades asignadas, propuestas para realizarse en un determinado período de tiempo.
- **Tiempo estimado:** tiempo previsto para la realización de cada una de las actividades, este tiempo se asignará, según la disponibilidad que tengan los integrantes y de la prioridad que se le dé a cada una de las tareas.
- **Actividades:** tareas asignadas que permita llevar un control del avance y ejecución de cada una de las fases del proyecto.
- **Responsables:** son las personas encargadas de realizar las actividades para cumplir los objetivos, de acuerdo al rol que desempeñen dentro del grupo de trabajo.
- El resultado de este paso es un plan de trabajo que permitirá llevar a cabo los fines del proyecto, mediante una adecuada definición de los objetivos que se pretenden alcanzar, de manera que se utilicen los recursos con eficiencia.

3.2. FASE II: Especificación del Entorno

3.2.1. Paso 1: Reconocimiento de las áreas relacionadas.

Hay muchas áreas relacionadas con los servicios que presta el CDU, en las cuales se podría actuar en el rediseño de operaciones diarias, de funciones y de implantación de nuevos métodos, para mejorar la disponibilidad de los servicios. En este aspecto es necesario definir y tener un conocimiento previo de las áreas más comunes por las que está compuesto el CDU (ver anexo C), entre las cuales se han establecido las siguientes:

- Área de Infraestructura.
- Área de soporte técnico (voz, datos y PC).

- Área de Servidores.
- Área de atención a usuarios (helpdesk).
- Área de desarrollo.
- Área de capacitación.

Es importante especificar detalladamente el área sobre la cual se va a realizar el desarrollo de la solución, puesto que brindar alta disponibilidad a todas las áreas de un centro de datos resultaría ser una tarea muy dispendiosa y sobre todo se necesitarían muchos recursos tanto humanos como económicos para cubrir todas las áreas.

Hay aspectos de cada una de las áreas del CDU que se relacionan de manera indirecta con la puesta en marcha de los servicios, que son complementarias para el buen funcionamiento de los mismos. Por lo que se recomienda tener una información actualizada de las mismas, de cómo están conformadas y qué tareas tienen dentro del CDU. Dichas áreas pueden estar a cargo de un personal distinto y manejar sus propias políticas y recursos; por lo que abarcar todas las áreas sería un proyecto verdaderamente ambicioso en cuestión de disponibilidad de recursos, lo cual no es muy viable, más cuando se trata de un centro de datos universitario.

En esta guía metodológica se recomienda documentar ciertos aspectos fundamentales de las áreas relacionadas, como sus funciones, personal a cargo, recursos disponibles; esto con el fin de que cuando se presenten falencias en el CDU, conocer el área que está directamente relacionada, para hacerlo saber inmediatamente al personal encargado, y en el mejor de los casos contribuir con ideas para su mejoramiento.

El área sobre la cual se va a profundizar para el desarrollo de una solución de Alta Disponibilidad para los servicios críticos de un centro de datos universitario es el área de servidores, puesto que es la que tiene que ver con los métodos que incorporan nuevas tecnologías de HA en los servicios implementados y es el área directamente implicada con la puesta en marcha, mantenimiento y operación del servicio. Por lo tanto, se recomienda documentar exhaustivamente sobre todo lo concerniente a dicha área, lo que implica conocer muy bien sus funciones y estado actual.

El resultado de este paso es la documentación de las áreas del CDU, profundizando en el área de servidores.

3.2.2. Paso 2. Identificación del Tier de Disponibilidad de la infraestructura del CDU.

El estándar TIA-942, establece las características que deben ejecutarse en los componentes de la infraestructura para los distintos grados de disponibilidad. Por lo tanto, inicialmente es recomendable saber en cuál de los cuatro niveles del Estándar TIA-942 (*Telecommunication Infrastructure Standard for Data Centers*) (nombrados en el capítulo 1 del presente trabajo), se encuentra clasificado el centro de datos universitario en cuestión, lo cual da un referente de la disponibilidad que maneja el CDU actualmente y permite tener en cuenta aspectos para mejorar la planificación y gerenciamiento del centro de datos, puesto que para cada uno de los cuatro niveles o Tiers que plantea este estándar,

incluye anexos que describen las recomendaciones para la infraestructura edilicia, de seguridad, eléctrica, mecánica y telecomunicaciones [15]. Los niveles o tiers que maneja este estándar se explican detalladamente en el capítulo dos del presente trabajo y se enumeran a continuación:

- ✓ Tier I: Centro de Datos Básico
- ✓ Tier II: Componentes Redundantes
- ✓ Tier III: Mantenimiento Concurrente
- ✓ Tier IV: Tolerante a Fallas

En esta guía metodológica, se ha generado una tabla de clasificación TIER, con base al estándar TIA-942 (ver tabla 3.3 Requisitos para el centro de datos según anexo g del estándar ANSI/TIA 942), la cual contiene aspectos esenciales y necesarios para clasificar al CDU en un determinado nivel o tier de disponibilidad, la cual le permite al administrador del CDU, tener una visión más amplia en este aspecto, para incluir dentro de su planeación la posibilidad de avanzar hacia un nivel de disponibilidad mayor o dar sugerencias a las otras áreas que no clasifiquen dentro de un nivel apropiado de disponibilidad, ya que hay que tener en cuenta que a mayor número de Tier, mayor grado de disponibilidad y cada área se evalúa independientemente pero la evaluación global se hace respecto a la que tenga menor nivel de clasificación.

Este paso es importante, porque lleva al replanteo de las necesidades de infraestructura de una manera racional y alineada con las necesidades propias de disponibilidad del CDU. Aunque esta guía va enfocada al área de servidores, se hace necesario tener en cuenta esta clasificación en el área de infraestructura, porque afecta la funcionalidad de los servicios que se prestan, haciendo una revisión en este aspecto se pueden generar recomendaciones para manifestarlas al área de infraestructura.

El resultado de este paso es la **clasificación TIER**, en el que se encuentra actualmente el CDU.

TIPO TIER	REQUISITOS EN LOS SUBSISTEMAS		CUMPLE		OBSERVACIÓN
			SI	NO	
TIER 1	TELECOMUNICACIONES				
	1	Tener debidamente etiquetado paneles de conexión, patch cord, puntos de red.			
	2	Emplear norma para realizar etiquetado (ANSI/TIA/EIA-606-A y ANEXO B TIA-942)			
	3	Tener etiquetado todos los bastidores y gabinetes, al frente y atrás.			
	ESTRUCTURALES Y ARQUITECTÓNICOS				
1	Protección contra eventos físicos (intencionales, accidentales, naturales o causados por el hombre)				

TIER 2	MECÁNICOS				
	1	Sistema de refrigeración con una o múltiples unidades de aire acondicionado.			
	2	El sistema de tuberías tiene un único camino o vía.			
	3	Si se tiene un solo generador, todo el equipo de aire acondicionado debe ser alimentado por el sistema generador en espera (standby).			
	ELÉCTRICOS				
	1	Sin redundancia en el sistema de distribución eléctrica.			
	2	Sistema de puesta a tierra debe cumplir con los requisitos mínimos.			
	3	Contar con un método económico de puesta a tierra para satisfacer los requisitos de los fabricantes.			
	TELECOMUNICACIONES				
	1	Tener redundancia de componentes (fuentes de alimentación, procesadores) en equipos críticos como enrutadores, switches para las redes LAN y SAN.			
	2	El cableado del backbone LAN/SAN debe tener redundancia de conexión, ya sea en la fibra o en el par de cobre.			
	3	Debe haber al menos 20m de separación desde el punto de acceso hasta la sala de ingreso.			
	4	Todos los patch cords y jumpers deben estar etiquetados en ambos extremos con el identificador correspondiente a la conexión.			
	ESTRUCTURALES Y ARQUITECTÓNICOS				
	1	La protección contra eventos físicos es mínima.			
2	Todas las puertas de seguridad deben ser de madera sólida con marcos de metal.				
3	Puertas a los equipos de seguridad y salas de control, deben contar con mirillas de 180 grados.				
4	Todos los muros de seguridad deben ser de altura completa (piso a techo).				
MECÁNICOS					
1	El sistema de refrigeración incluye múltiples unidades de aire acondicionado con una unidad en redundancia N+1.				
2	El sistema de tubería tiene una sola vía.				
3	Sistemas de aire acondicionado diseñados para operación 7x24x365.				

	4	Unidades de aire acondicionado de la sala de computadores con redundancia mínima N+1.			
	5	El equipo de aire debe ser alimentado por el sistema generador en espera (standby).			
	6	Sistemas de control de temperatura alimentado con circuitos dedicados redundantes, desde UPS.			
	7	Se debe instalar un sistema generador de reserva para suministrar electricidad al sistema de alimentación ininterrumpida y equipos mecánicos.			
	8	Los tanques de almacenamiento de combustible en el sitio deben estar calculados de un tamaño tal que proporcione un mínimo de 24 horas de funcionamiento del generador en condición de carga de diseño.			
	9	Sistema de almacenamiento de combustible con redundancia y aislamiento.			
	ELÉCTRICOS				
	1	Proporcionar redundancia N+1, para los módulos UPS (Uninterruptible Power Supply).			
	2	Se requiere un generador a la medida para manejar todas las cargas del centro de datos.			
	3	Un circuito no debe servir más que a un bastidor.			
4	La impedancia de tierra debe ser de menos de 5 ohmios.				
TIER 3	TELECOMUNICACIONES				
	1	Debe tener al menos dos proveedores de acceso.			
	2	El cableado de los proveedores de acceso debe estar separado entre sí a lo largo del recorrido al menos 20m.			
	3	Debe contar con dos salas de ingreso (ER-Entrance Room), separadas una distancia mínima de 20m, ubicadas en lados opuestos del centro de datos.			
	4	No se debe compartir equipo del proveedor de acceso, ni entre las dos salas de ingreso.			
	5	Debe haber redundancia en vías de distribución del backbone entre las salas de ingreso, el área de distribución principal y las áreas de distribución horizontal.			
	6	Las conexiones redundantes deben estar en distintas fundas del cable.			
	7	Respaldo en hot standby para todos los equipos críticos de telecomunicaciones, equipos del proveedor de acceso, enrutadores y switches del núcleo de producción y las redes LAN/SAN.			
	8	El sistema de cableado debe estar documentado.			

ESTRUCTURALES Y ARQUITECTÓNICOS				
1	Contar con protección específica contra los fenómenos físicos.			
2	Debe contar con entradas redundantes y puntos de seguridad.			
3	No debe haber ventanas en los muros del perímetro exterior de la sala de computadores.			
4	El edificio debe proporcionar protección contra radiación electromagnética.			
5	Se debe proporcionar separación física o de otro tipo de protección a los equipos redundantes y servicios para eliminar el riesgo de cortes simultáneos.			
6	El perímetro del sitio debe estar protegido por un sistema de detección de intrusos y monitoreado por un Circuito Cerrado de Televisión (CCT).			
7	El acceso al sitio debe ser garantizado por sistemas de identificación y autenticación.			
8	Contar con una sala de seguridad dedicada para proporcionar control central a todos los sistemas de seguridad asociados con el centro de datos.			
MECÁNICOS				
1	El sistema HVAC incluye múltiples unidades de aire acondicionado, con suficientes unidades de redundancia.			
2	El sistema de conductos es de dos vías.			
3	Todas las unidades de aire acondicionad deben estar respaldadas por un generador de alimentación.			
4	Se debe dedicar al centro de datos equipo de refrigeración con redundancia N+1, N+2, 2N ó (2N+1).			
5	Se debe instalar sensores de detección de fugas de agua.			
ELÉCTRICOS				
1	Debe proporcionar al menos redundancia N+1 para el sistema, incluyendo el generador y sistema de UPS.			
2	Al menos dos alimentadores de servicios públicos deben ser prestados para el centro de datos en media o alta tensión de (por encima de 600 voltios).			
3	El combustible en el sitio de almacenamiento debe estar dimensionado para proporcionar un mínimo de 72 horas de funcionamiento del generador en condición de carga.			

	4	Se debe proporcionar una infraestructura de puesta a tierra y un sistema de protección contra rayos.			
	5	Debe estar instalado un supresor de picos de tensión transitorios (TVSS) en todos los niveles del sistema de distribución que alimenta a las cargas electrónicas críticas.			
	6	Se debe proporcionar monitoreo ambiental y de alimentación eléctrica a todos los equipos principales, como los dispositivos de distribución principal, sistemas generadores, sistemas UPS, interruptores automáticos de transferencia estática (ASTS), unidades de distribución de energía, interruptores de transferencia automática, centros de control de motores, sistemas de supresión de aumento transitorio de voltaje, y sistemas mecánicos.			
	7	Deber haber redundancia para servidores de monitoreo y control.			
TIER 4	TELECOMUNICACIONES				
	1	El cableado del backbone debe ser redundante.			
	2	El cableado entre dos espacios debe seguir rutas separadas, con vías comunes sólo en los espacios finales.			
	3	Debe haber respaldo automático para todo equipo crítico de telecomunicaciones.			
	4	El centro de datos debe tener dos áreas de distribución una principal y una secundaria.			
	5	Las dos áreas deben estar ubicadas en lados opuestos del centro de datos y separadas al menos 20m			
	6	Las dos áreas no deben compartir zonas de protección de fuego, unidades de distribución de potencia y equipo de aire acondicionado.			
	7	Los enrutadores y switches de las áreas principal y secundaria deben ser redundantes, de tal manera que aún con una falla en cualquiera de ellas o en unas de las salas de ingreso, la red pueda continuar operando.			
	ESTRUCTURALES Y ARQUITECTÓNICOS				
	1	Contar con protección específica contra todos los potenciales eventos físicos.			
	2	Debe haber un área designada fuera del edificio, lo más cerca posible del generador, para los tanques de almacenamiento de combustible.			
	3	Instalaciones ubicadas en zonas sísmicas 0,1 y 2, deben ser diseñadas de acuerdo con requisitos de una zona sísmica 3. Las ubicadas en zonas			

	sísmicas 3 y 4, deben diseñarse de acuerdo a requisitos para zonas sísmicas 4.			
4	Equipos de datos y bastidores en zonas sísmicas 3 y 4 deben estar fijados en la parte inferior y superior.			
MECÁNICOS				
1	El sistema HVAC incluye múltiples unidades de aire acondicionado, con suficientes unidades de redundancia.			
2	El sistema de tuberías tiene dos vías.			
ELÉCTRICOS				
1	La instalación debe estar diseñada en una configuración 2(N+1) en todos los módulos, sistema y vías de distribución.			
2	Todos los equipos deben soportar traspaso manual en caso de falla o mantenimiento.			
3	En caso de falla, durante el traspaso de alimentación eléctrica, las cargas electrónicas críticas no deben sufrir interrupción en la alimentación.			
4	Se debe proporcionar un sistema de monitoreo de baterías.			
5	El edificio debe tener al menos dos alimentadores de diferente subestaciones de servicios públicos.			

Tabla 3.3. Requisitos para el centro de datos según anexo g del estándar ANSI/TIA 942

3.3. FASE III: Clasificación de los servicios críticos del CDU

Lo que se busca en esta fase es delinear el estado actual de los servicios del centro de datos, para obtener un conocimiento de los servicios que se disponen y de su desempeño, identificando servicios críticos y clasificándolos según su prioridad dentro de la comunidad universitaria, para esto es necesario estudiar la disponibilidad de los servicios que se prestan en el CDU, llevando a cabo los siguientes pasos:

3.3.1. Paso 1. Actualización de la Documentación de los servicios del CDU.

Revisar la documentación actual de los servicios, si no existe información detallada y actualizada, acerca del estado y funcionamiento de los servicios que presta el CDU. Es necesario e indispensable generar esta documentación, con el fin de dar soporte a la investigación y justificar cualquier cambio en los mismos. En este aspecto hay que resaltar que es imprudente que un administrador de un CDU quiera implementar tecnologías más avanzadas y profundizar en el mejoramiento de los servicios cuando no cuenta ni siquiera con la documentación actualizada del funcionamiento y configuración de los servicios que se tienen instalados, lo cual es un error que se comete muy a

menudo según la investigación realizada a diferentes CDU's (ver anexo B). Por lo cual se recomienda tener una buena documentación de la configuración y funcionamiento de los servicios con los que cuenta actualmente el CDU, esto incluye los manuales de instalación, manuales de usuario y una bitácora donde se lleve un registro de fallas comunes y la manera de cómo se resolvieron en cada una de las herramientas instaladas. Es muy importante tener plasmada esta información para que se compartan conocimientos y minimizar esfuerzos entre las personas encargadas del área, ya que mejora la eficiencia y la respuesta ante eventos que puedan causar interrupción de los servicios, contribuyendo de tal manera a la disponibilidad de los mismos.

El resultado de este paso es la documentación actualizada de los servicios del CDU.

3.3.2. Paso 2. Recolección de Estadísticas de funcionamiento de los servicios

Las estadísticas corresponden a los valores arrojados por herramientas de monitoreo, y por resultados de un análisis hecho por las personas que trabajan en el área implicada.

Las herramientas de monitoreo, están implementadas generalmente en el servidor de gestión, las cuales grafican los tiempos de funcionamiento del servicio gracias a los logs que generan cada uno de los servicios en el sistema. Por lo tanto, si por algún motivo no cuenta con los reportes generados por las herramientas de gestión, puede entrar directamente a buscar esta información en los logs del sistema, por ejemplo en el caso de que tenga implementados los servicios en un ambiente Linux, se entraría a la ruta */var/log/* y revisar archivos tales como *syslog*, *error.log*, *access.log*, de cada uno de los servicios instalados en el servidor.

Por otro lado, el análisis de estadísticas puede resultar de un cuestionario realizado por el personal que trabaja en el área de servidores, quienes son los que conocen el funcionamiento de cada uno de los servicios implementados y conocen a fondo el motivo de las fallas de los servicios del CDU; ya que si se hace el test a usuarios normales, a ellos les es transparente el funcionamiento de los mismos y cada uno definirá sus propios servicios críticos según la necesidad o experiencia propia, sin conocer en profundidad las causas de interrupción del servicio, lo cual es lo que se pretende abordar.

En caso de que se lleve un registro de las solicitudes al área, se pueden tener en cuenta la cantidad de solicitudes frente al servicio, cantidad de soporte brindado, tiempo invertido en mantenimiento, tiempo de parada del servicio, etc.

Todos los valores anteriormente mencionados son medibles y ayudan a corroborar de manera cuantitativa la sospecha del administrador sobre el servicio crítico que debe ser mejorado. En esta guía se ha generado un cuestionario para ayuda a la recolección de información relevante pero que está sujeta a cambios según el criterio del administrador del CDU. El cuestionario consta entre otras, de las siguientes preguntas:

1. ¿Por qué cree que es importante implementar mecanismos de Alta Disponibilidad en los servicios?
2. De acuerdo a su experiencia, labor realizada en el área, y según la importancia que éstos representen para el CDU y la comunidad universitaria, califique con un valor de 1 a 5 los servicios implementados en el CDU.
3. ¿Existen servicios que reiteradamente se bloquean o están caídos?
4. ¿Cuáles son los servicios que manualmente debe estar subiendo?
5. De los servicios que anteriormente mencionó, ¿cuáles son los que presentan caídas, con mayor frecuencia?
6. ¿Tiene implementados sistemas de Alta Disponibilidad en los servicios? Especifíquelos.
7. ¿Realiza endurecimiento (hardening) a los servidores antes de la instalación de los servicios?
8. ¿Tiene actualizadas las versiones de los sistemas operativos?
9. ¿Tiene actualizadas las versiones software de los servicios instalados?
10. ¿Tiene documentación actualizada sobre la configuración de los servicios instalados?
11. ¿Ejecuta planes de prueba para el mejoramiento continuo de los servicios? ¿Con qué frecuencia?
12. ¿Cuáles cree que serían los servicios que deberían tener mecanismos de HA (Alta Disponibilidad)?
13. ¿Proyecta un aumento de disponibilidad y mejoría del servicio, definiendo mecanismos y teniendo en cuenta su relación con las otras áreas?
14. Si existen falencias en los servicios, ¿cuáles son los motivos que considera fundamentales en las falencias de dichos servicios?

Por lo tanto, el resultado de este paso es un documento que contenga estadísticas de disponibilidad y de fallas encontradas.

3.3.3. Paso 3. Análisis de funcionamiento de los servicios

El CDU tiene implementados muchos servicios, pero no todos tienen la misma importancia para la universidad, por tanto se debe identificar qué servicios son críticos, para esto, es necesario hacer un seguimiento de las falencias que presentan, conocer si presentan interrupciones y los motivos que las generan e identificar las falencias más comunes, esto permite abordar los problemas de acuerdo a un grado de prioridad.

Gracias al paso dos se recoge información necesaria que ayuda a realizar este paso, que da como resultado la **identificación de las principales fallas** y los servicios implicados.

3.3.4. Paso 4. Determinación de los servicios críticos

Gracias a los pasos mencionados anteriormente, resulta más fácil concluir cuales son los servicios que son críticos para el CDU, ya que se analizaron falencias, necesidades e importancia de cada servicio implementado en el CDU, para así determinar sobre qué servicios se deben implementar un sistema de Alta Disponibilidad.

Por lo tanto el resultado de este paso, es determinar los servicios críticos que necesitan una solución HA, de acuerdo a los resultados obtenidos en los pasos 2 y 3 de esta Fase.

3.4. FASE IV: Estado del Arte de las Tecnologías de HA

Una vez se ha profundizado en la problemática de disponibilidad del centro de datos, se debe buscar información y seleccionar las herramientas software/hardware y mecanismos candidatos que contribuyan al desarrollo e implementación de la solución de alta disponibilidad, como por ejemplo clústeres, consolidación, replicación, virtualización, etc.

3.4.1. Paso 1. Clasificación de las Tecnologías de HA

Es necesario justificar dicha selección, lo cual implica inicialmente conocer el estado del arte y estudiar las soluciones comerciales y libres para implementar Alta Disponibilidad en los servicios de acuerdo a la infraestructura y tecnología con la que cuentan actualmente los servidores en el CDU, esto con el fin de hacer una óptima selección de la solución posible a implementar.

“La selección de software empresarial es una decisión importante para cualquier organización. Las tarifas por las licencias son costosas y las empresas que deben elegir el software que mejor se adapta a su modelo del negocio descubren que se trata de una decisión desalentadora. Las organizaciones deben asumir una enorme pérdida financiera cuando seleccionan un proveedor que, al momento de la verdad, no cumple con lo prometido. Por eso, al seleccionar una solución de software empresarial es necesario considerar muchos factores antes de tomar la decisión” [100]. Teniendo en cuenta, lo anteriormente expuesto, en esta fase se ha generado una tabla con los aspectos más importantes a considerar para la selección (ver tabla 4.1).

Como parte de esta investigación se debe documentar detalladamente los aspectos más relevantes de cada herramienta y mecanismos que sean más a fin según el interés específico de la solución que se esté buscando. Entre algunos factores más importantes se han identificado los siguientes:

- Tipo de licencia del software.
- Precio de compra.
- Trayectoria del producto en el mercado.
- Requisitos hardware requeridos.
- Requisitos software requeridos.
- Tipo de soporte.
- Empresa o desarrollador del producto.

El resultado de este paso es una documentación precisa, actualizada y detallada de tecnologías de HA disponibles, especialmente en los aspectos que se consideran más importantes, ya que ellos determinarán la selección de cada elemento necesario para la solución de Alta Disponibilidad que se está desarrollando.

A manera de sugerencia, la tabla 3.4 propone una manera práctica de presentar la información relevante. En ella se muestran campos que incluyen los factores mencionados, dentro de los cuales se tienen:

- ✓ **Requisitos hardware:** es un campo para mostrar lo que exige o requiere el software, en cuanto a disponibilidad de recursos como velocidad del procesador, capacidad de memoria RAM, tamaño en disco duro, sistema operativo requerido e interfaces de comunicación.

- ✓ **Información del desarrollador:** en este campo se puede consignar información relacionada con la empresa u organización que desarrolla el producto y que es relevante conocer y tener presente, tal como:
 - Nombre de la empresa.
 - Tiempo en el medio o mercado.
 - Página web
 - Correo electrónico de contacto.
 - Localización de la empresa: país de ubicación, sedes con que cuenta.

- ✓ **Tipo de licencia:** la información de este campo, está dada por el carácter de propiedad del producto, Por lo tanto, muestra información relacionada al tipo de licencia, por ejemplo ¿En caso de ser comercial, cuánto tiempo dura la licencia y cuánto cuesta renovarla?

- ✓ **Soporte:** este campo recopila información de acuerdo al tipo de licencia del software, por ejemplo si es de carácter comercial se interesa por aspectos como: ¿está incluido el soporte con el producto, o se debe pagar por él? ¿Cuál es el precio? ¿Qué incluye el soporte: instalación, configuración, actividades de puesta a punto?

- ✓ **Disponibilidad de la información:** Se refiere a la cantidad de información disponible y asequible sobre la tecnología que se está investigando.

- ✓ **Observaciones:** se ha agregado este campo al final para uso libre, está pensado para agregar comentarios sobre información relevante, que no está incluida en los demás campos.

Se recomienda utilizar la tabla 3.4 para la clasificación de las tecnologías HA, la cual permite poner en perspectiva cada uno de los elementos o productos a utilizar, y tener una visión comparativa de las ventajas y desventajas de cada uno, de manera que sea más fácil decidir cuál herramienta es más la apropiada para el desarrollo de la solución, según las condiciones actuales del CDU.

CARACTERÍSTICAS DE TECNOLOGÍAS HA DISPONIBLES										
N°	NOMBRE DEL PRODUCTO	FABRICANTE	Requisitos HW/SW	Precio	Licencia	SOPORTE (Tiempo y Costo)	DISPONIBILIDAD DE LA INFORMACIÓN			OBSERVACIONES (Funcionamiento ventajas y desventajas)
							Baja	Media	Alta	
1										
2										

Tabla 3.4. Clasificación de tecnologías HA.

3.5. FASE V: Definición del Sistema para Alta Disponibilidad

En esta fase ya se tiene conocimiento de los requerimientos de alta Disponibilidad y de los servicios críticos del CDU, por lo tanto, sólo queda escoger cual de las soluciones software y hardware investigadas es la más apropiada para implementar, y seguir una serie de recomendaciones que ayuden a realizar una implementación eficaz.

3.5.1. Paso 1. Establecimiento de bases conceptuales sobre mecanismos implícitos en un sistema HA.

Esta definición implica la recopilación de información de los aspectos relacionados con los mecanismos involucrados en el sistema de HA. Además implica entender los conceptos relacionados que van a servir de base y soporte, tanto en la implementación, como en el mantenimiento y continuidad del servicio. Es necesario resaltar que esta recopilación de información debe quedar documentada y organizada de manera que sea entendible.

El resultado de este paso es la documentación de los mecanismos involucrados en el sistema HA, que permita entender la interacción entre los mismos. En esta guía se mencionan algunos mecanismos más relevantes, con el fin de servir de apoyo inicial en la aclaración de conceptos fundamentales.

3.5.1.1. Virtualización

Es una tecnología de software que permite crear una versión virtual¹² de un dispositivo o recurso. Existen diferentes tipos de virtualización, algunas de ellas son [38], [39]:

- **Virtualización de recursos**

Es la combinación de múltiples recursos con la finalidad de obtener un recurso de mayor capacidad o bien varias entidades virtuales con las mismas características. La segmentación de una red o partición lógica de una única red física es un ejemplo de este tipo de virtualización.

¹² Virtual: existe solo aparentemente y no es real.

- **Virtualización de plataforma**

La virtualización de plataformas consiste en abstraer los recursos de una plataforma mediante software de control conocido como anfitrión o "host", el cual, simula un entorno computacional para el software huésped o "guest". Este software huésped, que generalmente es un sistema operativo, se ejecuta en la plataforma como si fuera el único ejecutándose en ella.

Debido al gran aumento de poder de cómputo de las computadoras actuales resulta más ventajoso tener soporte de máquinas virtuales que tener físicamente máquinas separadas. Los servidores nunca trabajan utilizando a un 100%, con la virtualización se puede aprovechar hasta un 80% la capacidad del CPU y sus recursos. Virtualizar una plataforma permite lograr una utilización de los recursos significativamente alta, así como, en el ahorro de energía, espacio y facilita la administración debido que se reduce el número de servidores físicos.

- **Virtualización del almacenamiento**

La virtualización del almacenamiento consiste en la creación de un contenedor lógico de datos, el cual mediante software aparenta estar físicamente situado en un único servidor, cuando en realidad, los datos pueden estar situados en cientos de discos físicos repartidos por decenas de servidores.

- **Técnicas de virtualización**

La clasificación general de las técnicas de virtualización se divide en virtualización completa y paravirtualización, las cuales se explican a continuación [40], [41]:

- **Virtualización completa**

Un sistema operativo huésped se ejecuta sobre un sistema operativo anfitrión, incluso se pueden ejecutar varios sistemas operativos sin realizar modificaciones sobre ellos. El sistema operativo huésped es una aplicación del sistema operativo anfitrión lo cual hace que su rendimiento sea limitado. Ejemplos: VMWare Workstation/server, Virtual Server, Virtual box, etc.

- **Paravirtualización**

Permite mayor rendimiento comparado con la virtualización completa, eso se debe a que se modifica el sistema operativo invitado, para ejecutarse en el hypervisor (el hypervisor actúa de árbitro para el acceso a los recursos), permitiendo de esta manera, un rendimiento muy cercano al de una máquina real.

- **Hypervisor**

Llamado también Monitor de Máquina Virtual (Virtual Machine Monitor-VMM). Es el software o programa responsable de almacenar y gestionar las máquinas virtuales,

controlando su acceso a los recursos hardware físicos asignados a cada máquina virtual [42].

3.5.1.2. CLÚSTER

Un clúster se define como una agrupación de elementos, en el caso de un clúster de computadoras, hace referencia al conjunto de equipos independientes que están comunicados entre sí y que son usados como un recurso unificado. Existen tres clases de clúster [43] [44] [45] y [46]:

- **Clúster de alto rendimiento**

Clúster diseñado para proporcionar gran capacidad de procesamiento, por lo que está dirigido especialmente hacia ambientes donde se necesita realizar un alto volumen de cálculo, por ejemplo para dar apoyo a actividades de investigación donde se requiere simular un modelo climático que contiene una gran cantidad de variables [44].

- **Clúster de balanceo de carga**

Es el tipo de clúster empleado para distribuir las peticiones de servicio entre los nodos que lo componen [47].

- **Clúster de alta disponibilidad**

El clúster de alta disponibilidad, también conocido como HA Cluster o Clúster de conmutación por error (Failover Cluster), es un grupo de dos o más computadoras interconectadas entre sí. Se implementan con el propósito principal de ofrecer alta disponibilidad para los servicios que ofrece el clúster. Su funcionamiento se basa en tener equipos redundantes o nodos, los cuales se utilizan para proporcionar el servicio cuando falla alguno de los componentes del sistema, de tal manera que el servicio esté disponible el mayor tiempo posible [46].

- **Failover**

Failover o conmutación por error. Se define como la capacidad de conmutación automática, una vez ha ocurrido una falla o terminación anormal de: una aplicación activa, un servidor, un sistema o una red. Dicha conmutación se hace sobre un servidor, sistema, o red, redundante o en espera (standby). La conmutación por error ocurre sin intervención humana y, en general sin previo aviso [47].

- **Heartbeat**

Se denomina así, al método o técnica diseñada para detectar cuando un servicio ha dejado de funcionar. Consiste en la implementación de la comunicación continua, mediante el paso de mensajes entre los nodos de un clúster a través de la utilización de interfaces Ethernet o serie. Dichos mensajes se emplean para verificar el correcto

funcionamiento de los nodos del clúster y de esta manera saber cuándo un nodo está o no disponible [48].

3.5.2. Paso 2: Identificación de aspectos a tener en cuenta al elegir el sistema HA

Este paso describe los aspectos más relevantes a tener en cuenta, en el momento de elegir el sistema HA, algunos de ellos son:

- ✓ **Documentación** existente del producto, debe ser alta la información y documentación sobre la implementación y configuración de las herramientas, puesto que algunos proyectos HA ya se encuentran descontinuados o su información es limitada, lo cual es muy necesario que toda la información esté disponible ya que como va enfocada a una solución que la va a desarrollar el administrador del CDU, necesita tener buenas bases y conocimientos para realizar una implementación eficaz.
- ✓ **Costos y recursos involucrados**, los cuales deben estar disponibles para que el sistema sea viable de implementar y no quede a mitad de camino. La libertad, disponibilidad y costo del software, son aspectos muy importantes tenerlos muy en cuenta, sobre todo cuando el CDU cuenta con grandes limitaciones presupuestales para el proyecto, en este caso, se recomienda el uso de software libre.
- ✓ **Disminución de SPOF** (Single Point Of Failure), el sistema HA debe contribuir a la disminución de los puntos únicos de falla en el servicio.
- ✓ **Proyecto estable**, con un desarrollo y soporte continuo por parte de la comunidad.
- ✓ **Compatibilidad** con el mayor número de servicios y recursos existentes.
- ✓ **Compatibilidad con virtualización**.
- ✓ **Mecanismos de seguridad robustos**, implícitos y disponibles en el sistema.
- ✓ **Tendencia Tecnológica** del producto, puesto que se debe visionar el mejoramiento de la solución, el cual va de la mano con la tendencia de la tecnología implícita en el sistema HA y el poder estar a la vanguardia de la tecnología para aprovechar al máximo las nuevas herramientas tecnológicas.

Por lo tanto el resultado de este paso, es tener claramente identificados los aspectos más importantes que respalden la elección del sistema HA.

3.5.3. Paso 3. Realizar un inventario de componentes involucrados en la implementación del sistema HA.

Muchas veces cuando se habla del centro de datos se tiende a pensar sólo en la parte del cableado, lo cual es un error que se comete a menudo puesto que son varias las áreas que se manejan y muchos los elementos incorporados en ellas, y todos influyen en el

buen funcionamiento del centro de datos. Por lo tanto, es necesario, inicialmente conocer y clasificar los componentes con los que cuenta el CDU, y sobre los cuales se va a implementar un sistema de HA.

Como se mencionó en la fase dos, esta guía hace énfasis en el área de servidores, y sobre ésta área se centra la implementación de un sistema HA. Por lo tanto, es necesario conocer las herramientas y componentes actuales con los que cuenta esta área en el CDU, lo cual ayuda a escoger un sistema viable de implementar. Además permite aprovechar los recursos con los que se cuenta actualmente, como lo son la cantidad de servidores disponibles, el sistema de almacenamiento utilizado para los servicios en cuestión y la disponibilidad de inversión de recursos en esta área.

En este paso se debe analizar el sistema operativo y tipo de almacenamiento adecuado acorde las condiciones actuales y recursos con los que cuenta el CDU.

Para escoger el sistema de almacenamiento, es necesario conocer los tipos de tecnologías existentes relacionadas con esta función, saber cuáles de ellas se pueden implementar y cuáles son los requerimientos y beneficios que conllevan.

En esta guía metodológica, se ha generado una tabla en la cual se resumen los tipos de almacenamiento y sus características más sobresalientes, con el fin de servir de ayuda en la identificación y selección de dicho sistema.

DESCRIPCIÓN	APLICACIÓN	UTILIZACIÓN
DAS (DIRECT ATTACHED STORAGE)		
<ul style="list-style-type: none"> Almacenamiento agregado directo. Suele basarse en tecnologías SCSI (Small Computers System Interface), FC (Fiber Channel), e IDE. 	<ul style="list-style-type: none"> Es ideal para las configuraciones que se basan en el intercambio de archivos localizados y no hay necesidad de transferir archivos a través de largas distancias. Hoy en día, los PCs de escritorio utilizan arquitectura de almacenamiento DAS. 	<ul style="list-style-type: none"> Es económico. Presenta muchos inconvenientes, como es la dispersión del almacenamiento que implica una dificultad en la gestión de los backups y una relativamente baja tolerancia a fallos. Almacenamiento de nivel básico donde los dispositivos (de almacenamiento) están unidos (agregados, conectados) directamente a un equipo servidor o PC, como es el caso de discos duros internos, disqueteras, CD-ROM, arreglos de discos, unidades de cinta para backup, cabinas de disco (en Rack en o cualquier otro formato) y otros.
NAS (NETWORK AREA STORAGE)		
<ul style="list-style-type: none"> Almacenamiento del área de red. Se basa en servidores de archivos con tecnología y 	<ul style="list-style-type: none"> Es un sistema de almacenamiento orientado al servicio de archivos a través de una red de área local. 	<ul style="list-style-type: none"> Sistema específico orientado al servicio de archivos por lo tanto más eficiente, más confiable y disminuye los puntos de falla. Acceso a archivos sobre ambiente de red, no depende de un servidor o

<p>sistemas operativos específicamente desarrollados para éste propósito.</p> <ul style="list-style-type: none"> Se implementa mediante un dispositivo hardware simple, llamado "NAS box" o "NAS head", actúa como interfaz entre el NAS y los clientes. Los clientes se conectan al "NAS head" a través de una conexión Ethernet. 	<ul style="list-style-type: none"> Los sistemas NAS están orientados al manejo de archivos pequeños. 	<p>estación de trabajo.</p> <ul style="list-style-type: none"> Acceso a archivos desde clientes multiplataforma. Es almacenamiento de disco duro (o rígido) que se instala con su propia dirección de red en lugar de asumir la del servidor al cual está conectado.
SAN (STORAGE AREA NETWORK)		
<ul style="list-style-type: none"> Red de área de almacenamiento. Es una red especializada de alta velocidad para transporte y almacenamiento de datos. Su arquitectura permite que diversos dispositivos de almacenamiento se manejen como si fuesen uno solo y estén disponibles para todos los equipos que lo requieran. Suelen basarse en la tecnología FC (Fibre Channel), aunque también pueden basarse en Gigabit Ethernet. 	<ul style="list-style-type: none"> Es una red concebida para conectar servidores, arrays de discos y equipos de respaldo. 	<p>Entre sus ventajas están:</p> <ul style="list-style-type: none"> Mayor velocidad de acceso a datos. Menor tiempo de recuperación ante desastres (los tiempos de Backup y Restore se minimizan). Gestión centralizada, compartida y concurrente del almacenamiento (indiferentemente de la plataforma y sistema operativo de los Host). Reducción de tráfico de red LAN. Una desventaja es su costo (el precio del Gigabyte sale muy caro), y también la existencia de ciertas limitaciones para integrar soluciones y/o dispositivos de diferentes fabricantes.
RAID 0		
<ul style="list-style-type: none"> Consiste en una serie de unidades de disco conectadas en paralelo que permiten una transferencia simultánea de datos a todos ellos, con lo que se obtiene una gran velocidad en las operaciones de 	<ul style="list-style-type: none"> Es una buena alternativa en sistemas donde sea más importante el rendimiento que la seguridad de los datos. Es decir ambientes que puedan soportar una pérdida de tiempo de operación para poder 	<ul style="list-style-type: none"> Ofrece el mejor rendimiento pero no tolerancia a los fallos. Permite acceder a más de un disco a la vez, logrando una tasa de transferencia más elevada y un rápido tiempo de acceso. Por no utilizar espacio en información redundante, el costo por Megabyte es menor. Es el único nivel de RAID que no duplica la información, por lo tanto

<p>lectura y escritura.</p> <ul style="list-style-type: none"> • La velocidad de transferencia de datos aumenta en relación al número de discos que forman el conjunto. • Los datos se desglosan en pequeños segmentos y se distribuyen entre varias unidades. 	<p>reemplazar el disco que falle y reponer toda la información.</p> <ul style="list-style-type: none"> • Por lo tanto es aconsejable en aplicaciones de tratamiento de imágenes, audio, video o CAD/CAM, es decir, es una buena solución para cualquier aplicación que necesite un almacenamiento a gran velocidad pero que no requiera tolerancia a fallos. 	<p>no se desperdicia capacidad de almacenamiento.</p> <ul style="list-style-type: none"> • No existe protección de datos. • No tiene redundancia de datos.
RAID 1 MIRRORING		
<ul style="list-style-type: none"> • Este nivel de RAID usa un tipo de configuración conocido como "mirroring", ya que la información de un disco es completamente duplicada en otro disco. • Este tipo de RAID se conoce también como creación de discos espejo. 	<ul style="list-style-type: none"> • Está diseñado para sistemas donde la disponibilidad de la información es esencial y su reemplazo resultaría difícil y costoso (más costoso que reponer el disco en sí). Típico en escrituras aleatorias pequeñas con tolerancia a fallas. 	<ul style="list-style-type: none"> • El rendimiento de la lectura se mejora pues cualquiera de los dos discos puede leerse al mismo tiempo. • Proporciona el mejor rendimiento y la mejor tolerancia a fallos en un sistema multiusuario. • Se protege la información en caso de falla. • Evita la pérdida de información y las interrupciones del sistema debido a fallas de discos. • Se desperdicia el 50% de la capacidad de almacenamiento. • Sólo maneja dos discos.
RAID 3		
<ul style="list-style-type: none"> • Acceso síncrono con un disco dedicado a paridad. • Se requieren mínimo tres discos y se utiliza la capacidad de un disco para la información de control. • Es mejor para sistemas de un solo usuario con aplicaciones que contengan grandes registros. 	<ul style="list-style-type: none"> • Está especialmente recomendado para aplicaciones que requieran archivos de datos de un gran tamaño (vídeo, imágenes, DataWare House). • Es típico para transferencia larga de datos en forma serial tal como aplicaciones de imágenes o video. 	<ul style="list-style-type: none"> • Proporciona una alta disponibilidad del arreglo, así como una tasa de transferencia elevada, mejorando de ese modo el rendimiento del sistema. • Un disco de paridad dedicado puede convertirse en un cuello de botella porque cada cambio en el grupo RAID requiere un cambio en la información de paridad. • No plantea una solución al fallo simultáneo en dos discos.
RAID 5		

<ul style="list-style-type: none"> • Acceso independiente con paridad distribuida. • Este nivel de RAID es conocido como "striping con paridad distribuida", ya que la información se reparte en bloques como RAID-0, pero un bloque de cada disco se dedica a la paridad. Es decir los datos codificados se añaden como otro sector que rota por los discos igual que los datos ordinarios. • Requiere al menos tres y usualmente cinco discos en el conjunto. 	<ul style="list-style-type: none"> • Es mejor para los sistemas multiusuario en los cuales el rendimiento no es crítico, o que realizan pocas operaciones de escritura. • Es recomendable para aplicaciones intensas de entrada/salida y de lectura/escritura, tal como procesamiento de transacciones. 	<ul style="list-style-type: none"> • Es el esquema de protección de información más usado comúnmente, ya que proporciona un buen rendimiento general con una mínima pérdida de capacidad. • Optimiza la capacidad del sistema permitiendo una utilización de hasta el 80% de la capacidad del conjunto de discos. • Además el sistema tiene suficiente redundancia para ser tolerante a fallos. • Gracias a la combinación del fraccionamiento de datos y la paridad como método para recuperar los datos en caso de fallo, constituye una solución ideal para los entornos de servidores en los que gran parte del E/S es aleatoria, la protección y disponibilidad de los datos es fundamental y el costo es un factor importante.
RAID 10		
<ul style="list-style-type: none"> • Es un nivel de arreglo de discos, donde la información se distribuye en bloques como en RAID-0, adicionalmente cada disco se duplica como RAID-1, creando un segundo nivel de arreglo. Se conoce como "striping de arreglos duplicados". Se requieren, dos canales, dos discos para cada canal y se utiliza el 50% de la capacidad para información de control. • También se le conoce como RAID 0&1. 	<ul style="list-style-type: none"> • Ideal para sistemas de misión crítica donde se requiera mayor confiabilidad de la información, ya que pueden fallar dos discos inclusive (uno por cada canal) y los datos todavía se mantienen en línea. • Es apropiado también para ambientes de escrituras aleatorias pequeñas. 	<ul style="list-style-type: none"> • Este nivel ofrece un 100% de redundancia de la información y un soporte para grandes volúmenes de datos, donde el precio no es un factor importante. • Costo elevado, gran overhead y 100% de redundancia.

Tabla 3.5. Sistemas de almacenamiento

El resultado de éste paso, es un inventario de los componentes y recursos que se tienen a disposición en la implementación del sistema HA.

3.5.4. Paso 4: Definición de la propuesta de solución acorde a las condiciones actuales del CDU.

Se debe tener en cuenta el resultado de todos los pasos anteriormente ejecutados, y realizar una propuesta coherente a las condiciones actuales del CDU y localización donde se va a implementar el sistema HA.

En este paso se debe realizar un esquema o diseño que permita visualizar el funcionamiento del Sistema de alta disponibilidad que se va a implementar. Se recomienda colocar los mecanismos involucrados de manera que se pueda ver la interacción entre los mismos. El resultado de este paso es la propuesta de solución de Alta Disponibilidad para los servicios críticos del CDU acorde a sus condiciones actuales.

3.5.5. Paso 5: Profundizar en las tecnologías involucradas en la propuesta.

Es recomendable, identificar las ventajas más sobresalientes de las tecnologías incluidas en los productos y herramientas que se van a instalar, para aprovechar al máximo las facilidades y funcionamiento de las mismas, por lo tanto el resultado de este paso, corresponde a la definición del funcionamiento y ventajas sobre las tecnologías que están incluidas en el sistema.

3.6. FASE VI: Implementación del SW y HW de la propuesta.

3.6.1. Paso 1: Definir requerimientos hardware y software para la implementación.

El resultado de este paso es el listado de Requerimientos hardware y software, que se van a usar en la implementación de la propuesta, y una pequeña descripción de los mismos.

3.6.2. Paso 2: Capacitación del personal.

Se debe saber si el talento humano con el que se cuenta para la implementación tiene los conocimientos básicos requeridos para la manipulación de la nueva tecnología, o si hace falta capacitación del personal, para de esta forma definir la manera como se va a realizar dicha capacitación, que es ineludible para el buen funcionamiento del software o tecnología que se requiera para complementar y no poner en riesgo la solución de HA que se está implementando.

Se recomienda hacer una evaluación inicial a los integrantes del grupo de trabajo para conocer su nivel de conocimiento con el cual pueden aportar al proyecto. Se recomienda además hacer un listado de los conocimientos previos a la implementación y verificar si es

necesario establecer horarios de socialización o capacitación ya sea entre los mismos integrantes o por medio de personal externo al equipo. El resultado de este paso es el plan de capacitación para el grupo de trabajo.

3.6.3. Paso 3: Instalación y configuración del sistema de almacenamiento

En este paso, se instala y configura el sistema de almacenamiento y el sistema operativo en el cual se va a montar los servicios. Por ejemplo, en el caso de que se vaya a configurar el almacenamiento con un arreglo de discos en RAID5, primero se instala adecuadamente el HW requerido como lo son la tarjeta controladora y los discos, se configura el arreglo desde la BIOS y luego se procede a instalar el sistema operativo sobre el cual se van a instalar los servicios.

El resultado de este paso es el sistema de almacenamiento instalado y configurado con los respectivos manuales de instalación y configuración.

3.6.4. Paso 4: Realizar endurecimiento del servidor.

Primero que todo se instala el sistema operativo en el servidor y luego se procede a realizar el proceso de *endurecimiento o hardening*.

El proceso de *Endurecimiento de Servidores* es la implementación de algunas prácticas de seguridad, que son muy importantes llevar a cabo en un entorno de trabajo de servidores, el cual ayuda a disminuir el riesgo potencial de ataques o intrusiones sobre el servidor que pongan en peligro la confidencialidad e integridad de la información alojada o la buena prestación de un determinado servicio [49].

Existen muchas prácticas de seguridad, así como varios niveles de endurecimiento para servidores en distintos sistemas operativos, por lo tanto, para que el proceso de endurecimiento de los servidores sea satisfactorio se deben seguir básicamente los siguientes pasos [50]:

- Realizar actividades de endurecimiento antes de que el sistema se conecte a la red para prevenir ataques.
- Se debe basar la configuración en el *modelo de menos-privilegios*: en el cual el sistema debe permitir el acceso solo en la medida necesaria para una adecuada funcionalidad. Del mismo modo, se debe dar a los usuarios el mínimo conjunto de derechos de acceso que ellos puedan necesitar.

Las siguientes tareas se basan en los dos puntos anteriores y resumen el proceso general de endurecimiento de servidores, los procesos específicos varían dependiendo del sistema operativo instalado [50][51][52]:

✓ **Actualizar y añadir parches de seguridad al sistema y software instalado**

Realizar Actualizaciones del sistema de forma periódica evita que el servidor se vea afectado por *bugs* conocidos y permite utilizar nuevas características y mejoras del software instalado.

✓ **Evitar suministrar Información del sistema**

Cuando un usuario accede al servidor ya sea que inicia sesión por medio de SSH, SMTP, HTTP etc, algunos archivos brindan información del sistema, por tanto, es necesario borrar, comentar y/o modificar su contenido para evitar suministrar información del mismo.

✓ **Deshabilitar Servicios Innecesarios**

Se debe instalar sólo lo absolutamente necesario y deshabilitar procesos o servicios que no sean fundamentales para el correcto funcionamiento del servidor, por ejemplo en sistemas operativos Linux, el súper demonio *inetd* puede iniciar varios servicios a la vez como POP, IMAP, FTP, etc. Por lo tanto, en la instalación inicial, se debe comentar todas las líneas del archivo de configuración de este demonio.

Así mismo pueden existir otros servicios instalados con el sistema base, que no se necesitan en el servidor, pero que se inician automáticamente por lo tanto es necesario desinstalarlos o deshabilitarlos.

✓ **Asegurar la administración remota**

Una forma muy utilizada para administrar remotamente un servidor, es utilizar SSHv2, pero existen varias vulnerabilidades que pueden hacerlo inseguro, por lo tanto, después de instalar este servicio es importante realizar algunos ajustes para evitar algunos ataques conocidos.

✓ **Implementar firewall**

La implementación de un firewall local en un servidor, que permita únicamente las conexiones desde y hacia puertos TCP/UDP conocidos, es una práctica recomendada en seguridad de servidores puesto que este elemento impide el acceso a posibles aplicaciones inseguras o puertas traseras instaladas en la maquina, para esto se cuenta con diferentes opciones dependiendo del sistema operativo por ejemplo Iptables y TCP Wrappers en Linux o IP Filter en Solaris [51] [53].

✓ **Otras configuraciones seguras**

Dependiendo del o los servicios que se vayan a prestar por determinado servidor, deben tenerse en cuenta las configuraciones seguras recomendadas por los desarrolladores de los mismos, además de procesos estándar como la administración adecuada de permisos en carpetas y archivos del sistema, restricción de shell de usuarios, manejo de contraseñas seguras y un particionado adecuado [51] [52].

El resultado de este paso es el sistema operativo instalado y asegurado, tener en cuenta que la implementación de los pasos en el endurecimiento, va de acuerdo al sistema operativo escogido.

3.6.5. Paso 5: Instalación del Software HA.

Después de haber hecho el respectivo proceso de endurecimiento, se procede a instalar los programas y librerías necesarias para poner en funcionamiento el sistema de HA; se sugiere paralelo a la instalación, documentar los pasos que se siguieron para la instalación, como también la documentación de los inconvenientes encontrados, con el fin de tener a disposición esta información en caso de que se necesite volver a instalar. Además se recomienda quitar el entorno gráfico de los servidores para reducir el consumo de recursos del sistema.

El resultado de este paso, es el sistema HA instalado y configurado y el respectivo manual de la implementación.

3.6.6. Paso 6: Instalación y configuración de los servicios.

Se procede a instalar los servicios críticos definidos en la Fase III, sobre los cuales se va a validar el funcionamiento de Alta Disponibilidad. Se sugiere generar los manuales de instalación y configuración de los servicios, como también la documentación de los inconvenientes encontrados, con el fin de tener a disposición esta información y minimizar esfuerzos cuando se presenten inconvenientes de reinstalación.

Otra recomendación importante aunque no absolutamente necesaria, es que al finalizar la instalación y configuración de los servicios, en el sistema, se proceda a realizar una imagen del sistema, esto con el fin de salvaguardar inconvenientes futuros que dañen la configuración de alguno de los servicios, para así, evitar volver a instalar y a configurarlos desde cero. La imagen permitiría hacer un rollback¹³ al punto en donde estaba funcionando correctamente. Esto se realiza siempre y cuando la tecnología escogida soporte y brinde esta opción.

3.7. FASE VII: Evaluación de la disponibilidad

Esta fase tiene como finalidad evaluar el funcionamiento del sistema con relación a la disponibilidad del servicio y a los objetivos propuestos inicialmente.

3.7.1. Paso 1. Definición de escenarios de pruebas.

Se recomienda por practicidad definir escenarios de pruebas que ayuden a obtener resultados del funcionamiento del sistema implementado y realizar un esquema que permita visualizar con mayor entendimiento dicho escenario. En este escenario de

¹³ Operación que permite devolverse a un estado previo.

pruebas se deben simular condiciones críticas en la prestación de los servicios, como por ejemplo el daño repentino de uno de los servidores, con lo que se pretende comprobar si el sistema automáticamente respalda el servicio caído y si verdaderamente se están disminuyendo, los puntos únicos de falla. También es necesario hacer pruebas de seguridad del manejo de información en el nuevo sistema, por ejemplo si se está implementando un clúster, verificar si la información que se replica entre los nodos, está siendo debidamente encriptada o cumple con un mecanismo de seguridad de respaldo de la información como por ejemplo una imagen de la configuración de los servicios.

El resultado de este paso es la definición del escenario de pruebas y un esquema del mismo.

3.7.2. Paso 2. Medición de disponibilidad en el escenario de pruebas.

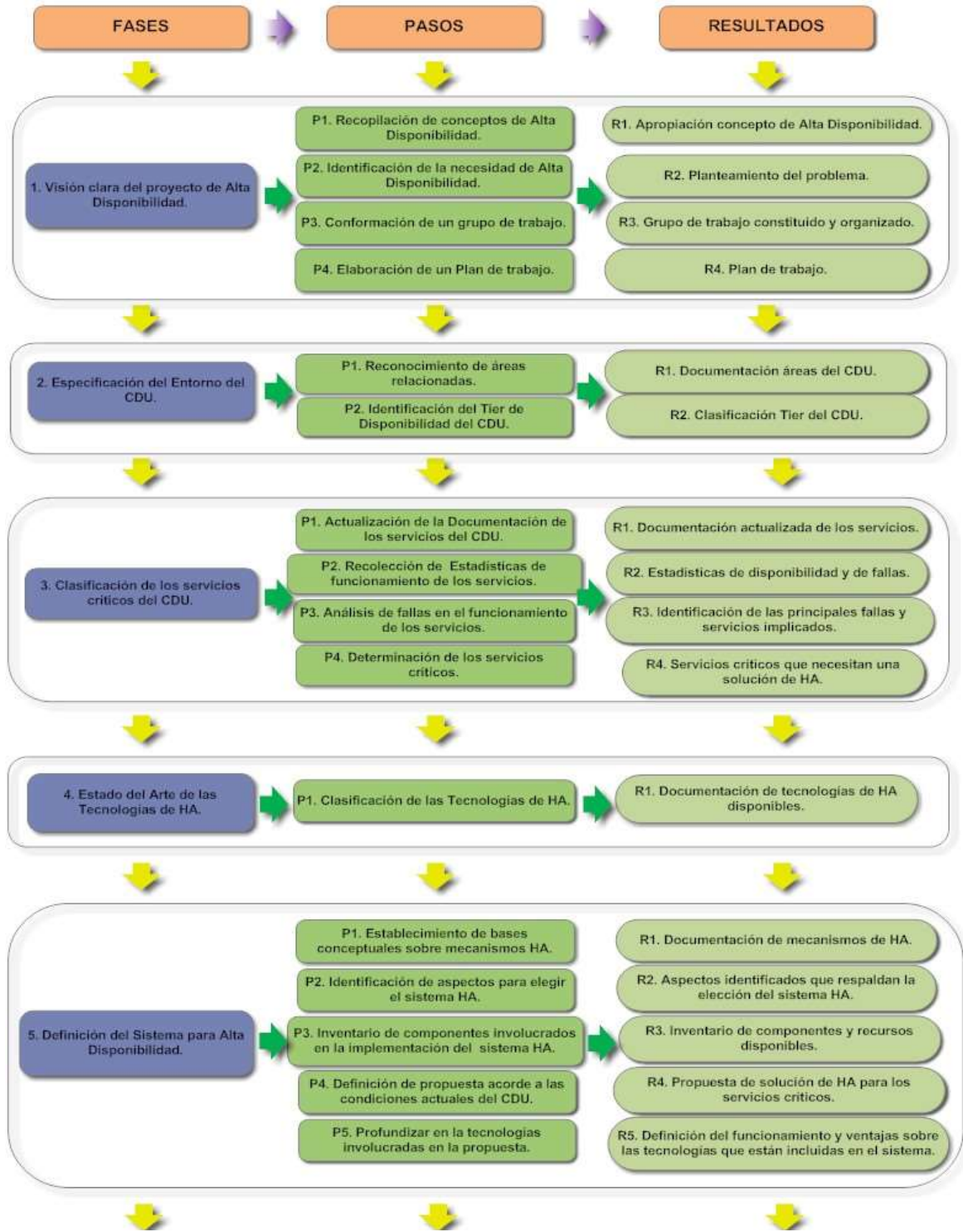
En este paso se recomienda medir tiempos de interrupción del servicio (tiempo que transcurre mientras que automáticamente se estabiliza nuevamente el servicio en caso de una falla), este tiempo multiplicado por el promedio de interrupciones del servicio da como resultado el tiempo total de indisponibilidad del servicio en determinado periodo de tiempo, con el cual se puede aplicar las formula de disponibilidad del servicio (Ec.2) y sacar el porcentaje de alta disponibilidad logrado aplicando el sistema HA que se ha propuesto. Además es necesario realizar pruebas de estrés con los servicios, generar tráfico hacia los servidores implicados para medir el rendimiento de los mismos, en el nuevo sistema implementado. También se recomienda medir el tiempo en que se demora en restaurar la imagen del sistema lo cual es de gran ayuda cuando se ha dañado la configuración del servicio ya que no hay necesidad de comenzar desde cero sino restablecer la imagen guardada.

El resultado de este paso es una lista de mediciones obtenidas en el servidor en el que se implementó el sistema HA, con el fin de comparar los resultados de disponibilidad obtenidos antes de la implementación del sistema.

3.7.3. Paso 3. Análisis de Resultados.

Con los valores obtenidos con el paso anterior, se puede calcular la nueva disponibilidad alcanzada con el prototipo implementado de Alta Disponibilidad, se compara la nueva disponibilidad obtenida y se determina si el aumento de la disponibilidad compensa la implementación de la propuesta. Además se determina si cumple o no cumple con los objetivos iniciales del proyecto y con los aspectos que se tuvieron en cuenta al realizar la elección de los mecanismos HA a implementar, correspondientes al paso dos de la fase cinco. En caso de no cumplirlos el administrador debe analizar si es necesario retroalimentar los pasos en alguna fase de la guía metodológica. El resultado de este paso es un análisis de los resultados obtenidos, para determinar el impacto alcanzado que permitan visualizar el cambio y mejora en la disponibilidad de los servicios. Además, se recomienda hacer un seguimiento de disponibilidad al nuevo sistema implementado, utilizando herramientas de monitoreo que permitan observar y llevar un control de la disponibilidad de los servicios, de manera que permita supervisar continuamente la disponibilidad de los mismos e iniciar los procedimientos oportunos.

A manera de conclusión, se muestra un Esquema general del desarrollo de una solución de Alta Disponibilidad para los servicios críticos en centros de Datos Universitarios (Ver figura 3.2).



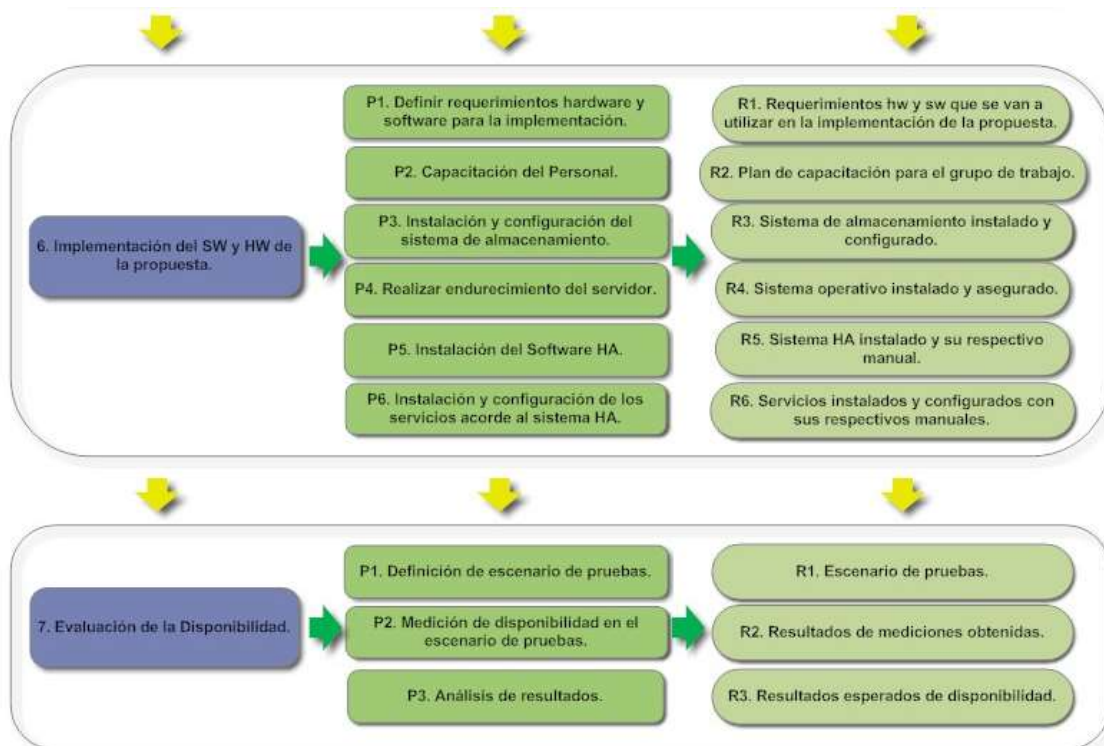


Figura. 3.2. Esquema General del desarrollo de una solución de Alta Disponibilidad para los servicios críticos de un CDU.

CAPÍTULO 4

DESARROLLO DE UNA SOLUCIÓN DE ALTA DISPONIBILIDAD ACORDE CON LAS CONDICIONES ACTUALES DEL CENTRO DE DATOS DE LA UNIVERSIDAD DEL CAUCA

Este capítulo contiene la implementación de cada una de las fases y pasos expuestos en la guía metodológica del capítulo III de este trabajo de grado, con la cual se logra planear desde la propuesta inicial hasta la implementación y validación de una solución de Alta Disponibilidad para el CDU de la Universidad del Cauca (caso de estudio).

En cada una de las fases se van desarrollando una serie de pasos que generan unos productos o resultados que forman parte de la solución propuesta. A continuación se explica detalladamente cada uno de los pasos ejecutados.

4.1. FASE I: Visión clara del proyecto de Alta Disponibilidad

4.1.1. Paso 1. Recopilación del concepto de Alta Disponibilidad.

Para dar cumplimiento a la apropiación del concepto de alta disponibilidad, se comenzó por entender exactamente lo que significa el término Alta Disponibilidad, para posteriormente “construir” una definición que permita tener claro su significado.

La actividad principal de este proceso de apropiación es la documentación respecto al tema de Alta Disponibilidad. Esta tarea permitió construir un marco teórico como parte de una base de conocimiento de referencia, la cual contribuyó a dar el enfoque y dimensionamiento adecuados a la propuesta de solución de alta disponibilidad. Puesto que el concepto de alta disponibilidad es muy amplio y encierra muchos aspectos, los cuales fue necesario identificar para determinar exactamente sobre cuál de ellos se va a trabajar o aplicar el desarrollo de la propuesta de solución. En este caso se estableció que la solución se enfocaría estrictamente en la disponibilidad de los servicios críticos. Dentro de la construcción de la base de conocimiento se llegó a la definición del concepto de alta disponibilidad, el cual quedó así:

“Alta disponibilidad, es la capacidad de proporcionar acceso a un servicio o aplicación con un mínimo de interrupciones programadas y no programadas, y en caso de que éstas se produzcan, el tiempo de recuperación del servicio debe ser mínimo, mitigando así el impacto del tiempo de inactividad generado [6]. Una manera práctica de expresar el concepto de HA, consiste en que los servicios y aplicaciones estén funcionando 24 horas al día, 7 días a la semana, los 365 días del año; lo que significa que los servicios y

aplicaciones deban estar funcionando durante un alto porcentaje del tiempo programado de servicio”.

El concepto anterior fue sólo el primer paso en el proceso de desarrollo de la solución, por delante estaba entender la implicación de aplicarlo en los servicios de un CDU. Como parte del proceso de documentación se generó la respectiva documentación, la cual se encuentra plasmada en el capítulo uno.

4.1.2. Paso 2. Identificación de la necesidad de Alta Disponibilidad.

En el caso de estudio que se aborda en este trabajo de grado, el cual es el centro de datos de la Universidad del Cauca, se han detectado interrupciones continuas en diferentes servicios, por lo cual se necesitan tomar medidas preventivas y correctivas que ayuden a mejorar su disponibilidad.

El Centro de Datos de la Universidad del Cauca, cuenta con una amplia gama de servicios de TI, como son: correo electrónico, web, FTP, Internet, hosting institucional para bases de datos y sitios web, entre otros. Todos estos servicios son soportados mediante la infraestructura de TI del centro de datos. Últimamente dado el crecimiento en el número de usuarios de los servicios, se ha incrementado la utilización de los recursos de la red de datos, cayendo dicha carga sobre los servidores, que deben soportar cada vez mayor procesamiento, exigiéndoles en determinados momentos hasta su capacidad límite, lo que afecta su desempeño, especialmente su disponibilidad.

Ante esta situación actualmente el Centro de Datos de la Universidad del Cauca, no cuenta con un sistema o estrategia, que le permita afrontar estos eventos de manera eficiente, pues algunos procedimientos para restaurar los servicios o levantar un servidor se hacen de forma manual y responden a una estrategia reactiva. Además, cuando el servicio se interrumpe, no existe ningún mecanismo automático de respaldo que minimice su impacto, simplemente se debe esperar a que sea restaurado, lo que no se sabe exactamente cuánto tiempo tomará. Mientras tanto, el servicio no está disponible o funciona bajo condiciones poco óptimas, afectando a la comunidad universitaria. Esto en gran parte se debe a que algunos servicios no cuentan con un sistema redundante que permita minimizar el impacto de la caída en el servicio.

Lo anterior evidencia la necesidad de hacer frente a esta situación, y de proporcionar confianza a los usuarios mediante unos servicios eficientes y altamente disponibles. De aquí surge la necesidad de indagar sobre el tema de Alta Disponibilidad en los servicios, lo cual es la tendencia tecnológica en los sistemas operativos y en el hardware que lo soporta. Por lo tanto, lo que se busca con la guía metodológica es desarrollar una solución de Alta Disponibilidad para los servicios críticos del CDU de la Universidad del Cauca, que le permita mejorar la disponibilidad de los mismos.

Las posibles soluciones pueden ser muchas, pero dado que el enfoque es hacia el Centro de Datos de una Universidad que cuenta con pocos recursos económicos, se requiere que la solución sea viable, económica, y eficiente. En general pueden ser muchos los factores que influyen sobre la indisponibilidad del servicio, tales como la interrupción en los servidores en caso de sobrecalentamiento cuando falla el aire acondicionado,

sobrecarga en la UPS's, o a causa de altos picos generados en el circuito eléctrico al que están conectados, sobrecarga de procesamiento, ataques de denegación de servicio, e incluso la falta de documentación, etc. los cuales se deben detectar y corregir de manera metódica. Para minimizar el efecto de estas interrupciones, es necesario implementar mecanismos de Alta Disponibilidad, que en caso de una eventual falla, proporcionen respaldo a los servicios instalados en el servidor mientras se restablece el servicio, permitiendo así, la continuidad del mismo.

4.1.3. Paso 3. Conformación de un grupo de trabajo.

Rol	Nombre
Coordinador	<ul style="list-style-type: none"> Administrador del CDU
Representante administrativo	<ul style="list-style-type: none"> Jefe de la división de sistemas.
Diseñador(es)	<ul style="list-style-type: none"> Ingenieros a fines en el área de electrónica, telecomunicaciones y/o sistemas.
Ejecutor(es)	<ul style="list-style-type: none"> Monitores del CDU, Tecnólogos o ingenieros a fines en el área de electrónica, telecomunicaciones y/o sistemas.
Evaluador	<ul style="list-style-type: none"> Administrador del CDU. Ingeniero con experiencia en el manejo de procesos y/o proyectos.

Tabla 4.1. Grupo de Trabajo.

4.1.4. Paso 4. Plan de trabajo.

La correcta planificación para encontrar una solución de Alta Disponibilidad permite establecer unos niveles de disponibilidad adecuados en lo que respecta a las necesidades reales del CDU, y a las posibilidades que tiene la universidad para llevar a cabo dichas labores.

OBJETIVOS ESPECÍFICOS	TIEMPO ESTIMADO	ACTIVIDADES	RESPONSABLE(S)
Recopilación de información en el Centro de Datos universitario	seis semanas	<ul style="list-style-type: none"> Conocer la situación actual de disponibilidad de los servicios. Esta información debe ser actualizada. Recolectar información de las herramientas para la monitorización de la disponibilidad. 	Ejecutores. Coordinador.
Recopilación de información actual de los servicios del CDU	4 semanas	<ul style="list-style-type: none"> Recopilar información del Hardware. Recopilar información del software instalado por cada servicio. Documentar funcionamiento de los servicios. Documentación de la configuración de los servicios. 	Ejecutores. Coordinador.

Revisión bibliográfica	Dos semanas	<ul style="list-style-type: none"> • Estado del arte de los métodos y técnicas de análisis a utilizar. • Definiciones relevantes y precisas de las métricas a utilizar. 	Ejecutores
Fase de Diseño	Una semana	<ul style="list-style-type: none"> • Proponer mecanismos para mejorar la disponibilidad de los servicios. 	Diseñadores
Capacitación	4 semanas	<ul style="list-style-type: none"> • Manejo del sistema operativo escogido. • Configuración de maquinas virtuales • Configuración del sistema HA. 	Ejecutores. Coordinador. Evaluador.
Implementación	tres semanas	<ul style="list-style-type: none"> • Instalación de HW y SW. • Realizar manuales de instalación, configuración y de usuario. • Puesta a punto. 	Ejecutores. Coordinador.
Pruebas y Validación	Tres semanas	<ul style="list-style-type: none"> • Escenario de pruebas • Mediciones. • Verificar funcionamiento. 	Ejecutores. Coordinador. Evaluador.

Tabla 4.2. Plan de Trabajo.

4.2. FASE II: Especificación del Entorno

4.2.1. Paso 1. Reconocimiento de áreas relacionadas en el CDU de la Universidad del Cauca.

Como resultado de este paso, se generó una extensa y completa documentación de las áreas del CDU de la Universidad del Cauca, que incluye información actualizada de cada una de ellas, respecto a estructura organizacional, integrantes, funciones, localización, recursos y planes de mejoramiento.

Para tener una idea general de qué tipo de información se debería recolectar, se investigó documentación de otros Centros de Datos Universitarios tanto del país como a nivel Internacional, que tenían publicada la información en sus páginas web, tales como: Universidad del Valle, Universidad de Antioquia, Universidad, EAFIT, Universidad Nacional de Colombia, Universidad Michoacana de San Nicolás Hidalgo (México), El Centro de Cómputo Universitario (CCU Culiacán) de la Universidad Autónoma de Sinaloa (México), Universidad de la Rioja (España).

El documento se generó como aporte al CDU de la Universidad del Cauca, y se entrega a los miembros del área. Debido a que el documento "C" como se le nombró, contiene cierta información precisa de configuración y de equipos, se clasificó de carácter no público y no se deja como anexo en el presente trabajo de grado.

Es necesario destacar que el área sobre la cual se profundizó para el desarrollo de una solución de Alta Disponibilidad para los servicios críticos de un centro de datos universitario, es el área de servidores, puesto que es la que tiene que ver con los métodos que incorporan nuevas tecnologías de HA en los servicios implementados y es el área directamente implicada con la puesta en marcha, mantenimiento y operación del servicio. Por este motivo se han establecido una serie de funciones y tareas rutinarias en el área, que aunque se realizan no se encuentran actualmente documentadas.

Funciones del área de Servidores:

- Instalación y configuración de los servicios que presta el centro de datos, tales como web, proxy, DNS, correo, ftp, DHCP, LDAP.
- Realizar backups de la información de la configuración de los servicios, así como de los usuarios.
- Definir e implementar políticas de manejo y funcionamiento de los servicios que presta el centro de datos.
- Crear cuentas de usuarios para los servicios de correo, ftp, y sitios web.
- Asignación de direcciones IP, ya sean públicas o por NAT.
- Monitoreo y control del buen funcionamiento de los servicios implementados en el centro de Datos.
- Actualizar y mejorar la prestación de los servicios del centro de datos.
- Documentar la implementación, configuración y funcionamientos de los servicios.

Dentro de las tareas fundamentales que se realizan en el área de servidores, está el mantener arriba los servicios que presta el centro de datos, para esto se recomienda un Plan de pruebas rutinario de los servicios que se prestan que permita verificar al inicio de cada jornada (mañana y tarde), el estado actual de servicios del CDU:

1. Revisar el espacio en discos.
2. Revisar que las herramientas de Gestión que posee el CDU, por ejemplo Nagios, awstats, Allot, etc.; estén funcionando correctamente.
3. Comprobar que los DNS estén resolviendo debidamente, esto incluye tanto los DNS externos como los internos.
4. Verificar el buen funcionamiento de los proxys y del servidor DHCP.
5. Verificar si los servidores de correo están funcionando correctamente para el envío y recepción de los mensajes tanto en el correo de la universidad, como a correos externos.
6. Comprobar si las cargas de las CPU estén estables y no al límite.
7. Comprobar que los servicios de web tales como Tomcat, Apache y Mysql; tengan sus respectivos procesos establecidos, y que no estén pegados en un bucle infinito. Esto muchas puede suceder cuando se hacen demasiadas consultas o algún proceso ha quedado bloqueado.
8. Si se posee un servidor de autenticación, como por ejemplo LDAP, se debe verificar que tantas conexiones están establecidas y por supuesto que permita la autenticación de los usuarios a los diferentes servicios.
9. Revisar el ancho de banda de entrada y salida.

10. Revisar la cola de correos de los servidores de correo tanto de estudiantes como de funcionarios, verificar que se pueda enviar correo desde la red local a los servidores de correo tanto internos como externos (y viceversa).
11. Verificar que se pueda Navegar con los tres proxys.
12. Revisar si hay solicitudes pendientes registradas por el área de Helpdesk.

Las tareas anteriores muestran que el mantener los servicios “arriba”, son funciones directamente relacionadas con el área de servidores, por lo tanto es deber de ésta área, estar mejorando la calidad de los servicios que se prestan, esto incluye el deber de actualizarse continuamente y realizar un banco de pruebas que permita estar a la vanguardia de las tecnologías existentes, por lo tanto es función de esta área buscar una solución de Alta Disponibilidad para los servicios que presta el Centro de Datos.

4.2.2. Paso 2. Identificación del tier de disponibilidad de la infraestructura del CDU de la Universidad del Cauca.

Para lograr la identificación del nivel o TIER de disponibilidad del Centro de Datos de la Universidad del Cauca, se procedió a examinar las características en infraestructura de cada uno de los cuatro subsistemas: telecomunicaciones, eléctrico, arquitectónico/estructural y mecánico, según lo indica la plantilla expuesta en el capítulo 3, generada de la norma ANSI/TIA 942. Este proceso permitió determinar en qué nivel (Tier1, Tier2, Tier3 ó Tier4), está ubicada la infraestructura general del centro de datos. Para llevar a cabo esta actividad se examinó cada uno de los puntos expuestos en la plantilla. En la actividad se contó con la colaboración del personal del centro de datos relacionado con cada uno de los subsistemas, permitiendo obtener los resultados mostrados a continuación:

✓ *Subsistema de Telecomunicaciones*

Este es uno de los subsistemas más importantes para el proyecto, ya que permite explorar el estado actual del entorno del Área de Servidores y Servicios, con la que está directamente relacionada. Este subsistema incluye todo aquello del centro de datos que haga uso de telecomunicaciones, entre lo que está:

- El sistema del cableado: que incluye el medio de transporte, su topología física, el etiquetado, la distribución y la documentación.
- Los equipos de red: su redundancia, su alimentación.
- La distribución física del centro datos: las áreas, sus funciones y ubicación.

En el subsistema de telecomunicaciones se verificaron en total 22 características, aplicadas al centro de datos, distribuidas entre los cuatro niveles o TIERS. De este número 17 de ellas arrojaron un cumplimiento negativo, con lo que el resultado final es que el subsistema se ubica en una clasificación TIER 1.

En el centro de datos Unicauca, este subsistema es reconocido por su gran importancia y se realiza continuamente, por parte de la institución, una fuerte inversión para mejorar la infraestructura en equipos de red y demás dispositivos de soporte. Por eso, aunque la

clasificación de este subsistema sea TIER 1, en realidad podría ser mejor, ya que se tiene una muy buena infraestructura en equipos, cableado, proveedores y demás aspectos que darían para un nivel TIER 2. La razón de este resultado, según la investigación desarrollada, se debe principalmente a la manera cómo surgió el centro de datos, ya que no fue concebido ni proyectado para crecer, y aún no se tiene unas políticas claras al respecto. Esta situación ha hecho que vaya creciendo como respuesta a la necesidad inmediata, lo que no da tiempo de planear, para que se haga de manera ordenada bajo unos estándares o recomendaciones básicas. Pues las falencias en cuanto al subsistema de telecomunicaciones son por ejemplo que no se etiqueta el cableado, los paneles de conexión o los bastidores debidamente bajo una norma, en sí no se hace. Otro caso es que no se tiene documentado el cableado, o el hecho de no tener definidos los espacios que sugieren las normas. Lo anterior permite ver que no es que no se tenga lo necesario, sino que está desubicado o desordenado. En la tabla 3, se muestra cada una de las consideraciones examinadas y el respectivo resultado.

REQUISITOS PARA EL SUBSISTEMA DE TELECOMUNICACIONES			
TIPO TIER	CARACTERÍSTICA	CUMPLE	
		SI	NO
TIER 1	Tener debidamente etiquetado paneles de conexión, patch cord, puntos de red.		X
	Emplear norma para realizar etiquetado (ANSI/TIA/EIA-606-A y ANEXO B TIA-942)		X
	Tener etiquetado todos los bastidores y gabinetes, al frente y atrás.		X
TIER 2	Tener redundancia de componentes (fuentes de alimentación, procesadores) en equipos críticos como enrutadores, switches para las redes LAN y SAN.		X
	El cableado del backbone LAN/SAN debe tener redundancia de conexión, ya sea en la fibra o en el par de cobre.		X
	Debe haber al menos 20m de separación desde el punto de acceso hasta la sala de ingreso.		X
	Todos los patch cords y jumpers deben estar etiquetados en ambos extremos con el identificador correspondiente a la conexión.		X
TIER 3	Debe tener al menos dos proveedores de acceso.	X	
	El cableado de los proveedores de acceso debe estar separado entre sí a lo largo del recorrido al menos 20m.		X
	Debe contar con dos salas de ingreso (ER-Entrance Room), separadas una distancia mínima de 20m, ubicadas en lados opuestos del centro de datos.		X
	No se debe compartir equipo del proveedor de acceso, ni entre las dos salas de ingreso.	X	
	Debe haber redundancia en vías de distribución del backbone entre las salas de ingreso, el área de distribución principal y las áreas de distribución horizontal.		X
	Las conexiones redundantes deben estar en distintas vainas del cable.		X

	Debe haber respaldo en hot standby para todos los equipos críticos de telecomunicaciones, equipos del proveedor de acceso, enrutadores y switches del núcleo de producción y las redes LAN/SAN.		X
	El sistema de cableado debe estar documentado.		X
TIER 4	El cableado del backbone debe ser redundante.		X
	El cableado entre dos espacios debe seguir rutas separadas, con vías comunes sólo en los espacios finales.		X
	Debe haber respaldo automático para todo equipo crítico de telecomunicaciones.		X
	El centro de datos debe tener dos áreas de distribución una principal y una secundaria.		X
	Las dos áreas deben estar ubicadas en lados opuestos del centro de datos y separadas al menos 20m		X
	Las dos áreas no deben compartir zonas de protección de fuego, unidades de distribución de potencia y equipo de aire acondicionado.		X
	Los enrutadores y switches de las áreas principal y secundaria deben ser redundantes, de tal manera que aún con una falla en cualquiera de ellas o en unas de las salas de ingreso, la red pueda continuar operando.		X

Tabla 4.3. Subsistema de Telecomunicaciones.

✓ **Subsistema arquitectónico y estructural**

Aunque no lo parezca a primera vista este subsistema al igual que los demás es muy importante en la disponibilidad de un centro de datos. Como su nombre lo indica, incluye todo lo relacionado con la infraestructura física de las instalaciones, tal como:

- La protección contra eventos físicos, como inundaciones, incendios, rayos o cualquier accidente.
- El tipo de materiales que se recomienda emplear en la construcción de los distintos espacios.
- El diseño y ubicación de los espacios de acuerdo a su función.

Para este subsistema se tomaron en total 17 características a ser verificadas, distribuidas entre los cuatro niveles. De ellas 10 arrojaron un cumplimiento negativo, lo que ubica al subsistema arquitectónico y estructural en el nivel TIER 1.

Este resultado se debe principalmente a la razón mencionada para el subsistema anterior, la cual es general, y es que el centro de datos no fue planeado para crecer. En este momento el espacio físico ya no es suficiente, además no está ubicado adecuadamente, pues comparte unas instalaciones que no están diseñadas para el fin, las cuales son muy pequeñas y no permiten realizar todas las modificaciones requeridas para dar cumplimiento a las normas, ya que es muy costoso modificar o remodelar y los recursos son limitados. Con esta situación es muy difícil que el centro de datos pueda tener o cumplir con los requerimientos estructurales y arquitectónicos, y no precisamente porque se desconozcan o por falta de recursos, en realidad es físicamente imposible. Sin embargo se debe reconocer que se cumple con consideraciones básicas y se hacen esfuerzos por mejorar continuamente las instalaciones.

REQUISITO PARA SUBSISTEMA ESTRUCTURAL Y ARQUITECTÓNICO			
Tipo TIER	CARACTERÍSTICA	CUMPLE	
		SI	NO
TIER 1	Existe algún tipo de protección contra eventos físicos (intencionales, accidentales, naturales o causados por el hombre)	X	
TIER 2	La protección contra eventos físicos es mínima.	X	
	Todas las puertas de seguridad deben ser de madera sólida con marcos de metal.		X
	Puertas a los equipos de seguridad y salas de control, deben contar con mirillas de 180 grados.		X
	Todos los muros de seguridad deben ser de altura completa (piso a techo).	X	
TIER 3	Contar con protección específica contra los fenómenos físicos.	X	
	Debe contar con entradas redundantes y puntos de seguridad.		X
	No debe haber ventanas en los muros del perímetro exterior de la sala de computadores.		X
	El edificio debe proporcionar protección contra radiación electromagnética.		X
	Se debe proporcionar separación física o de otro tipo de protección a los equipos redundantes y servicios para eliminar el riesgo de cortes simultáneos.		X
	El perímetro del sitio debe estar protegido por un sistema de detección de intrusos y monitoreado por un Circuito Cerrado de Televisión (CCT).		X
	El acceso al sitio debe ser garantizado por sistemas de identificación y autenticación.	X	
TIER 4	Contar con una sala de seguridad dedicada para proporcionar control central a todos los sistemas de seguridad asociados con el centro de datos.		X
	Contar con protección específica contra todos los potenciales eventos físicos.		X
	Debe haber un área designada fuera del edificio, lo más cerca posible del generador, para los tanques de almacenamiento de combustible.		X
	Instalaciones ubicadas en zonas sísmicas 0,1 y 2, deben ser diseñadas de acuerdo con requisitos de una zona sísmica 3. Las ubicadas en zonas sísmicas 3 y 4, deben diseñarse de acuerdo a requisitos para zonas sísmicas 4.	NA	NA
	Equipos de datos y bastidores en zonas sísmicas 3 y 4 deben estar fijados en la parte inferior y superior.	NA	NA

Tabla 4. 4. Subsistema Estructural y Arquitectónico.

✓ Subsistema eléctrico

Este subsistema es muy importante, dado el impacto que puede generar su funcionamiento en la disponibilidad de los servicios. Pues de nada sirve contar con los mejores equipos y toda la redundancia posible si cuando falla la alimentación principal,

no se cuenta con un sistema auxiliar, e incluso hasta los equipos redundantes se apagan. En este subsistema se examinaron las características relacionadas con:

- El sistema de alimentación: su configuración, tiempo de conmutación, su capacidad de carga, capacidad de operación continua.
- Los elementos del sistema de alimentación: los generadores, su redundancia, el combustible, las UPS, los transformadores.
- El sistema de protección: supresores de picos de tensión y sistema de puesta a tierra.

Se verificaron 19 características, de las cuales, 13 tienen un cumplimiento negativo, lo que clasifica al subsistema eléctrico en el nivel TIER 1.

Este subsistema al igual que los dos anteriores, presenta algunas falencias, pero en general está bien conformado a pesar de lo que indican los resultados. En cuanto a la infraestructura en equipos, es realmente buena, pues se cuenta con un generador diesel con una autonomía de 48 horas de funcionamiento a plena carga y un tiempo de conmutación máximo de 2 minutos. Además se dispone UPS para dar respaldo a los equipos de red mientras se conmuta al generador, adicionalmente se tiene regulador de tensión de 21 KVA, un supresor de picos de voltaje y sistema de puesta a tierra según regulación RETIE (Reglamento Técnico de Instalaciones Eléctricas). Como se puede ver, se cuenta con una infraestructura adecuada, la falencia se ubica en el sistema del cableado eléctrico que no está distribuido siguiendo una norma específica para centros de datos, entonces se instalan circuitos de alimentación por sitios que no debería, por ejemplo están muy cerca al cableado de red. En otros casos no se tiene redundancia en los circuitos eléctricos para cargas críticas como los servidores, enrutadores y switches, entre otros equipos. Pero en general en gran parte es la falta de espacio para realizar todas estas modificaciones, desde luego acompañado del costo, que es el factor más limitante.

REQUISITO PARA SUBSISTEMA ELÉCTRICO			
TIPO TIER	CARACTERÍSTICA	CUMPLE	
		SI	NO
TIER 1	Existe redundancia en el sistema de distribución eléctrico.		X
	Sistema de puesta a tierra debe cumplir con los requisitos mínimos.	X	
	Contar con un método económico de puesta a tierra para satisfacer los requisitos de los fabricantes.	X	
TIER 2	Proporcionar redundancia N+1, para los módulos UPS (Uninterrumpible Power Suply).		X
	Se requiere un generador a la medida para manejar todas las cargas del centro de datos.	X	
	Un circuito no debe servir más que a un bastidor.		X
	La impedancia de tierra debe ser de menos de 5 ohmios.		

TIER 3	Debe proporcionar al menos redundancia N+1 para el sistema, incluyendo el generador y sistema de UPS.		X
	Al menos dos alimentadores de servicios públicos deben ser prestados para el centro de datos en media o alta tensión (por encima de 600 voltios).		X
	El combustible en el sitio de almacenamiento debe estar dimensionado para proporcionar un mínimo de 72 horas de funcionamiento del generador en condición de carga.	X	
	Se debe proporcionar una infraestructura de puesta a tierra y un sistema de protección contra rayos.	X	
	Debe estar instalado un supresor de picos de tensión transitorios (TVSS) en todos los niveles del sistema de distribución que alimenta a las cargas electrónicas críticas.		X
	Se debe proporcionar monitoreo ambiental y de alimentación eléctrica a todos los equipos principales, como los dispositivos de distribución principal, sistemas generadores, sistemas UPS, interruptores automáticos de transferencia estática (ASTS), unidades de distribución de energía, interruptores de transferencia automática, centros de control de motores, sistemas de supresión de aumento transitorio de voltaje, y sistemas mecánicos.		X
	Deber haber redundancia para servidores de monitoreo y control.		X
TIER 4	La instalación debe estar diseñada en una configuración 2(N+1) en todos los módulos, sistema y vías de distribución.		X
	Todos los equipos deben soportar traspaso manual en caso de falla o mantenimiento.		X
	En caso de falla, durante el traspaso de alimentación eléctrica, las cargas electrónicas críticas no deben sufrir interrupción en la alimentación.		X
	Se debe proporcionar un sistema de monitoreo de baterías.		X
	El edificio debe tener al menos dos alimentadores de diferente subestaciones de servicios públicos.		X

Tabla 4.5. Subsistema Eléctrico.

✓ **Subsistema Mecánico**

Es un sistema que incluye un aspecto importante para los servicios, como lo es la refrigeración de los equipos, especialmente los servidores.

En este subsistema se examinaron 19 características en total, 14 de ellas arrojaron un cumplimiento negativo, por lo tanto el subsistema mecánico al igual que los otros tres subsistemas se ubica en un nivel TIER 1.

Al respecto de este subsistema, se tienen consideraciones especiales, ya que se dispone de una infraestructura, que se considera en buen estado, pero que comienza a quedarse pequeña, dado que se han agregado más equipos al centro de datos, en especial servidores de gran desempeño que implican un alto consumo de potencia y por tanto, generan mucho calor. Esta situación no cuenta con el debido monitoreo y control, por lo que no se puede estimar o calcular la necesidad real para un sistema de refrigeración, sin embargo el actual tiene un buen desempeño. Las falencias básicas de este subsistema, están sobre todo en la manera como está diseñado, pues no cuenta con redundancia en

la alimentación, se tienen varias unidades pero no están configuradas en redundancia, estas condiciones hacen que el sistema no tenga un desempeño más eficiente, aún estando en capacidad de realizarlo.

Actualmente la administración del centro de datos, realiza la asesoría necesaria para mejorar el sistema de refrigeración, buscando un sistema debidamente dimensionado, que permita responder a las exigencias actuales y futuras, pero también buscando modernizar los equipos ya que los actuales son un poco antiguos y su eficiencia es baja, lo que genera un consumo innecesario de alimentación especialmente. Además adelantan ampliaciones al sistema de refrigeración de otras áreas, lo que permitirá mejorar el sistema total de refrigeración.

REQUISITO PARA SUBSISTEMA MECÁNICO			
TIPO TIER	CARACTERÍSTICA	CUMPLE	
		SI	NO
TIER 1	Sistema de refrigeración tiene una o múltiples unidades de aire acondicionado.	X	
	El sistema de tuberías tiene un único camino o vía.	X	
	Si se tiene un solo generador, todo el equipo de aire acondicionado debe ser alimentado por el sistema generador en espera (standby).		X
TIER 2	El sistema de refrigeración incluye múltiples unidades de aire acondicionado con una unidad en redundancia N+1.	X	
	El sistema de tubería tiene una sola vía.	X	
	Sistemas de aire acondicionado diseñados para operación 7x24x365.		X
	Unidades de aire acondicionado de la sala de computadores con redundancia mínima N+1.		X
	El equipo de aire debe ser alimentado por el sistema generador en espera (standby).		X
	Sistemas de control de temperatura alimentado con circuitos dedicados redundantes, desde UPS.		X
	Se debe instalar un sistema generador de reserva para suministrar electricidad al sistema de alimentación ininterrumpida y equipos mecánicos.		X
	Los tanques de almacenamiento de combustible en el sitio deben estar calculados de un tamaño tal que proporcione un mínimo de 24 horas de funcionamiento del generador en condición de carga de diseño.	X	
	Sistema de almacenamiento de combustible con redundancia y aislamiento.		X
TIER 3	El sistema HVAC incluye múltiples unidades de aire acondicionado, con suficientes unidades de redundancia.		X
	El sistema de conductos es de dos vías.		X
	Todas las unidades de aire acondicionado deben estar respaldadas por un generador de alimentación.		X
	Se debe dedicar al centro de datos equipo de refrigeración con redundancia N+1, N+2, 2N ó (2N+1).		X

	Se debe instalar sensores de detección de fugas de agua.		X
TIER 4	El sistema HVAC incluye múltiples unidades de aire acondicionado, con suficientes unidades de redundancia.		X
	El sistema de tuberías tiene dos vías.		X

Tabla 4.6. Subsistema Mecánico CDU Unicauca.

En resumen la clasificación de los cuatro subsistemas se muestra en la siguiente tabla:

RESULTADO CLASIFICACIÓN CDU UNICAUCA		
SUBSISTEMA	CLASIFICACIÓN TIER	CLASIFICACIÓN DEL CENTRO DE DATOS
TELECOMUNICACIONES	TIER 1	TIER 1
ELÉCTRICO	TIER 1	
MECÁNICO	TIER 1	
ESTRUCTURAL ARQUITECTÓNICO	TIER 1	

Tabla 4.7. Clasificación Tier CDU Universidad del Cauca.

- **Conclusión y análisis sobre resultados de la plantilla de clasificación Tier para el Centro de Datos de la Universidad del Cauca:**

En general, partiendo de la información obtenida sobre los subsistemas, se estableció que el Centro de Datos de la Universidad del Cauca, según su infraestructura de TI, está ubicado en la categoría TIER 1, la cual corresponde, según la disposición del *Uptime Institute* a una disponibilidad del 99.671% y a un downtime anual de 28.8 horas, que corresponde a un tiempo de caída semanal promedio de 0,6 horas (36 minutos).

Este resultado es además producto de la aplicación del *Anexo G del estándar ANSI/TIA 942*, bajo dos criterios: el primero es que dicho anexo es aplicable de forma independiente a cada subsistema (telecomunicaciones, arquitectónico, eléctrico y mecánico) de la infraestructura. El segundo criterio es que la clasificación global del centro de datos es la correspondiente a la del subsistema con menor clasificación TIER. En el caso, aunque el Centro de Datos Unicauca, cumplía condiciones o características de nivel TIER 2 y 3 para algunos subsistemas, la mayoría de los subsistemas se ubican en el nivel TIER 1, por lo que sin ninguna duda el resultado es claro.

Cabe aclarar, que la clasificación realizada es sólo una herramienta que permite tener una idea del estado actual de la infraestructura de TI del centro de datos y la disponibilidad que puede ofrecer según esta. Además, antes que unos requerimientos exigidos, el estándar ANSI/TIA 942, lo que permite es cotejar unas características estándar pre-establecidas, que identifican cuatro tipos de centros de acuerdo al nivel de disponibilidad que puede proporcionar, dado que cumpla con un perfil de características determinado. A la vez, lo que busca ante todo este estándar, es ayudar a identificar las falencias o

debilidades que posee la infraestructura de TI y poder diseñar un plan para suplirlas, basado en las ventajas que ofrece hacerlo bajo un estándar.

4.3. FASE III: Clasificación de los servicios críticos del CDU de la Universidad del Cauca

4.3.1. Paso 1. Actualización de la documentación de los servicios del CDU.

Se generó información técnica y profundizada que no había sobre los servicios que funcionaban en el área, de cómo están actualmente los servicios del CDU, con el fin de tener a disposición esta información y minimizar esfuerzos.

Esta información se agregó al “Documento C”, de éste proyecto y se entrega como aporte a la División de Sistemas de la Universidad del Cauca, puesto que esta información no es de dominio público, se deja solo para el administrador, encargado del área. Este documento, contiene además, información sobre el funcionamiento de los servicios del CDU, información técnica detallada sobre el área de servidores, configuración, instalación, hardware y software relacionado a los servicios.

4.3.2. Paso 2. Recolección de Estadísticas de funcionamiento de los servicios del CDU de la Universidad del Cauca.

En este paso se tomaron valores de disponibilidad de los servicios, arrojados por la herramienta de monitoreo *Nagios*¹⁴, que actualmente utiliza el CDU de la Universidad del Cauca, ésta herramienta de monitoreo arroja valores de disponibilidad del servicio ya que posee plugins que se comunican con todos los servidores/servicios que se supervisan de manera remota desde el servidor de Gestión, así, de manera, que cada servicio monitoreado, envía información sobre su estado al servidor de gestión, el cual los exhibe en una interfaz gráfica, permitiendo visualizar los tiempos críticos o de indisponibilidad del servicio (franja roja), los cuales son valores de interés en esta recolección de información del funcionamiento de los servicios (ver figura 4.1).

¹⁴ <http://www.nagios.org/>

Service State Breakdowns:

Host	Service	% Time OK	% Time Warning	% Time Unknown	% Time Critical	% Time Undetermined
ftp_emisora	FTP	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
acuariu	FTP	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
	HTTP	99.755% (99.755%)	0.000% (0.000%)	0.000% (0.000%)	0.245% (0.245%)	0.000%
	MYSQL	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
	SSH	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
	TOMCAT	86.054% (86.054%)	0.000% (0.000%)	0.000% (0.000%)	13.946% (13.946%)	0.000%
acuario_viejo	FTP	99.988% (99.988%)	0.000% (0.000%)	0.000% (0.000%)	0.012% (0.012%)	0.000%
	HTTP	94.066% (94.066%)	0.000% (0.000%)	0.000% (0.000%)	5.934% (5.934%)	0.000%
	SI	99.980% (99.980%)	0.000% (0.000%)	0.000% (0.000%)	0.012% (0.012%)	0.000%
	TOMCAT	92.418% (92.418%)	0.000% (0.000%)	0.000% (0.000%)	7.582% (7.582%)	0.000%
afrodita	FTP	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
	HTTP	98.972% (98.972%)	0.000% (0.000%)	0.000% (0.000%)	1.028% (1.028%)	0.000%
	IMAP	99.964% (99.964%)	0.000% (0.000%)	0.000% (0.000%)	0.036% (0.036%)	0.000%
	POP	99.952% (99.952%)	0.000% (0.000%)	0.000% (0.000%)	0.048% (0.048%)	0.000%
	SMTP	94.138% (94.138%)	0.000% (0.000%)	0.000% (0.000%)	5.862% (5.862%)	0.000%
	SSH	99.964% (99.964%)	0.000% (0.000%)	0.000% (0.000%)	0.036% (0.036%)	0.000%
allot	PING	65.264% (65.264%)	0.012% (0.012%)	0.000% (0.000%)	34.725% (34.725%)	0.000%
arges	HTTP	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
atenea	FTP	99.977% (99.977%)	0.000% (0.000%)	0.000% (0.000%)	0.023% (0.023%)	0.000%
	HTTP	99.965% (99.965%)	0.000% (0.000%)	0.000% (0.000%)	0.035% (0.035%)	0.000%
	IMAP	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
	POP	99.964% (99.964%)	0.000% (0.000%)	0.000% (0.000%)	0.036% (0.036%)	0.000%
	SMTP	99.965% (99.965%)	0.000% (0.000%)	0.000% (0.000%)	0.035% (0.035%)	0.000%
	SSH	99.977% (99.977%)	0.000% (0.000%)	0.000% (0.000%)	0.023% (0.023%)	0.000%
atlantis	PING	87.563% (87.563%)	0.000% (0.000%)	0.000% (0.000%)	12.437% (12.437%)	0.000%
cisco4507	PING	99.988% (99.988%)	0.000% (0.000%)	0.000% (0.000%)	0.012% (0.012%)	0.000%
cronos	DHCP	0.000% (0.000%)	0.000% (0.000%)	100.000% (100.000%)	0.000% (0.000%)	0.000%
	DNS	93.172% (93.172%)	0.000% (0.000%)	0.000% (0.000%)	6.828% (6.828%)	0.000%
delos	HTTP	99.953% (99.953%)	0.000% (0.000%)	0.000% (0.000%)	0.047% (0.047%)	0.000%
dns1	DNS	99.917% (99.917%)	0.000% (0.000%)	0.000% (0.000%)	0.083% (0.083%)	0.000%
cisco4507	PING	99.988% (99.988%)	0.000% (0.000%)	0.000% (0.000%)	0.012% (0.012%)	0.000%
cronos	DHCP	0.000% (0.000%)	0.000% (0.000%)	100.000% (100.000%)	0.000% (0.000%)	0.000%
	DNS	93.172% (93.172%)	0.000% (0.000%)	0.000% (0.000%)	6.828% (6.828%)	0.000%
delos	HTTP	99.953% (99.953%)	0.000% (0.000%)	0.000% (0.000%)	0.047% (0.047%)	0.000%
dns1	DNS	99.917% (99.917%)	0.000% (0.000%)	0.000% (0.000%)	0.083% (0.083%)	0.000%
	SSH	99.952% (99.952%)	0.000% (0.000%)	0.000% (0.000%)	0.048% (0.048%)	0.000%
dns2	DNS	96.918% (96.918%)	0.000% (0.000%)	0.000% (0.000%)	3.082% (3.082%)	0.000%
	SSH	96.977% (96.977%)	0.000% (0.000%)	0.000% (0.000%)	3.023% (3.023%)	0.000%
fortianalyzer	PING	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
fortigate	PING	99.096% (99.096%)	0.000% (0.000%)	0.000% (0.000%)	0.904% (0.904%)	0.000%
gaia	HTTP	92.177% (92.177%)	0.000% (0.000%)	0.000% (0.000%)	7.823% (7.823%)	0.000%
hera	LDAP	99.988% (99.988%)	0.000% (0.000%)	0.000% (0.000%)	0.012% (0.012%)	0.000%
hiperion	PING	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
	PROXY	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
hiperion2	PING	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
	PROXY	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
juno	LDAP	99.988% (99.988%)	0.000% (0.000%)	0.000% (0.000%)	0.012% (0.012%)	0.000%
matematicas295c	PING	99.915% (99.915%)	0.000% (0.000%)	0.000% (0.000%)	0.085% (0.085%)	0.000%
netexplorer	PING	97.316% (97.316%)	0.012% (0.012%)	0.000% (0.000%)	2.671% (2.671%)	0.000%
nexus	PING	99.698% (99.698%)	0.000% (0.000%)	0.000% (0.000%)	0.302% (0.302%)	0.000%
	PROXY	99.688% (99.688%)	0.000% (0.000%)	0.000% (0.000%)	0.312% (0.312%)	0.000%
nexus2	PING	99.700% (99.700%)	0.012% (0.012%)	0.000% (0.000%)	0.288% (0.288%)	0.000%
	PROXY	99.724% (99.724%)	0.000% (0.000%)	0.000% (0.000%)	0.276% (0.276%)	0.000%
odin	FTP	99.976% (99.976%)	0.000% (0.000%)	0.000% (0.000%)	0.024% (0.024%)	0.000%
	SSH	99.988% (99.988%)	0.000% (0.000%)	0.000% (0.000%)	0.012% (0.012%)	0.000%
portal	HTTP	99.952% (99.952%)	0.000% (0.000%)	0.000% (0.000%)	0.048% (0.048%)	0.000%
	PING	99.940% (99.940%)	0.000% (0.000%)	0.000% (0.000%)	0.060% (0.060%)	0.000%
	SSH	99.965% (99.965%)	0.000% (0.000%)	0.000% (0.000%)	0.035% (0.035%)	0.000%
ragnarok	HTTP	99.843% (99.843%)	0.000% (0.000%)	0.000% (0.000%)	0.157% (0.157%)	0.000%
	TOMCAT	82.693% (82.693%)	0.000% (0.000%)	0.000% (0.000%)	17.307% (17.307%)	0.000%
recursos3750	PING	99.468% (99.468%)	0.000% (0.000%)	0.000% (0.000%)	0.532% (0.532%)	0.000%
router_emtel	PING	98.738% (98.738%)	0.024% (0.024%)	0.000% (0.000%)	1.238% (1.238%)	0.000%

2010-02-26 00:00:00 to 2010-03-26 17:03:51
Duration: 28c 17h 3m 51s

report page
[Current]

Figura 4.1. Monitoreo de Disponibilidad de los servicios con Nagios.

Además de la recolección de información de la herramienta de monitoreo *Nagios*, se hizo una encuesta al grupo de trabajo del área de servidores del CDU, para obtener información del funcionamiento de los servicios, en la cual se obtuvo como resultado que los servicios de *Tomcat*, *Apache* y *MySQL*, instalados en los servidores web de Ragnarok y acuario, estaban incurriendo en interrupciones frecuentes del servicio y que había la necesidad de estar muy pendiente de dichos servicios, para subirlos de manera inmediata. La encuesta se encuentra en el anexo B del presente trabajo de grado.

4.3.3. Paso 3. Análisis de fallas en el funcionamiento de los servicios del CDU de la Universidad del Cauca.

De acuerdo a la encuesta realizada y a la información suministrada por el personal del área de servidores, se concluyeron aspectos importantes en el funcionamiento de los servicios y se encontraron algunas falencias ocurridas, tales como:

- **Servidores Web:** El alto índice de consultas a los servicios implementados en este servidor, tales como *Apache*, *Tomcat* y *MySQL*, especialmente en la temporada de matrículas hace que los procesos se vuelvan lentos y muchas veces queden bloqueados o algunos de estos servicios quede parado inesperadamente, el administrador o monitor encargado tiene que estar pendiente de que los procesos de este servidor estén “arriba” y funcionando, de lo contrario, tiene que estar “matando” el proceso (en caso de que se quede pegado) o subiendo el proceso (en caso de que se haya caído).

Además en cualquier caso de bloqueo o caída del servicio, no se dispone de un servidor adicional que cumpla su misión mientras éste es restaurado. Aunque es transparente para los usuarios el funcionamiento, detrás de este servidor es muy importante resaltar que siempre debe haber una persona que esté pendiente de subir servicios, sobre todo *Tomcat* y *Apache* que se “caen” constantemente. En el caso específico de la universidad, es una de las tareas más comunes que se presentan. Es importante resaltar que gracias a estos servicios funciona correctamente el portal de la universidad del cauca, el cual es la cara de la institución frente al mundo en internet, ya que en el portal se publica la información masiva a los usuarios, información muy importante como las inscripciones, listas de admitidos, noticias a los usuarios, y demás acontecimientos que suceden en la universidad, convirtiendo este servicio el de mayor importancia para cualquier institución educativa. Además cabe destacar que por medio del servicio web se puede acceder al correo.

- **Servidor FTP:** Fallas de hardware del equipo, específicamente en uno de sus discos duros han hecho que se pierda parte de configuración e información de las cuentas de usuarios del ftp.
- **Servidor Proxy:** La falta de renovación de equipos no ha permitido la actualización a versiones de Linux más recientes, además la falta de redundancia no ha posibilitado hacer pruebas, que mejoren la disponibilidad del servicio.

- **Servidor de Correo:** Recientemente un daño de hardware inesperado produjo la caída del servicio, lo cual afectó masivamente a los usuarios ya que se prolongó el tiempo de reparación y de puesta en marcha del servicio, puesto que no se contaba con un servidor redundante de respaldo y sobre todo que no existía documentación de la configuración de este servicio.
- **Servidor de Gestión:** La falta de documentación de instalación y configuración de algunas herramientas, ha provocado la pérdida de datos importantes, que retrasan nuevamente la instalación de las mismas, como por ejemplo la herramienta de monitoreo de tráfico *MRTG*¹⁵, la cual es una herramienta importante para visualizar el tráfico de entrada y de salida de cada enlace y que necesita volver a instalarse desde un comienzo pero necesita algunos datos como la comunidad a la que pertenece.

Con respecto a las estadísticas obtenidas en la herramienta de monitoreo *Nagios*, se aclara que la indisponibilidad mostrada por *Allot* (equipo gestor de ancho de banda) y el servidor Atlantis, se debe a que estos servicios estaban “caídos” por mantenimiento pero continuaban siendo monitoreados por la herramienta de *Nagios*. Una recomendación por lo tanto, sería que, mientras se vaya a realizar un cambio en un servicio, se desactive el monitoreo de *Nagios* sobre dicho servicio para que no afecte en el reporte de disponibilidad de los servicios.

Como se puede observar en las estadísticas generadas por *Nagios*, Figura 4.1, *Tomcat* representa mayor indisponibilidad del servicio, lo cual está ocurriendo tanto en el servidor de Acuario como en el servidor de Ragnarok, con porcentajes de indisponibilidad del 82.692% y 86.054%, respectivamente, los cuales hacen parte de los servidores web del CDU de la Universidad del Cauca, estos porcentajes son menores a 99%, que es el porcentaje mínimo para ser considerado de Alta Disponibilidad.

4.3.4. Paso 4. Determinación de los servicios críticos que necesitan una Solución de Alta Disponibilidad en el CDU de la Universidad del Cauca.

Los servicios DNS, proxy, correo, y LDAP cumplen un papel importante en el CDU de la universidad del cauca, sin embargo, gracias a los pasos realizados anteriormente de recolección de estadísticas y análisis de fallas, se puede determinar que hay algunos servicios que presentan altos índices de consulta y que necesitan monitoreo y reinicio manual de manera frecuente por parte del administrador o monitores de la red, los servicios con este tipo de inconvenientes son el servidor de aplicaciones java (Tomcat), el servicio web apache y el servicio de base de datos *MySQL*, de los cuales depende el portal y demás aplicaciones importantes que están a disposición de la comunidad universitaria vía web. Debido a la importancia que tienen estos servicios y a los inconvenientes en su funcionamiento se convierten en nuestro objetivo, de darles una solución de Alta disponibilidad, que permita que su disponibilidad se encuentre en el rango 99% a 99.999% (rango que debe tener un servicio, para considerarse de alta disponibilidad).

¹⁵ <http://oss.oetiker.ch/mrtg/>

Por otra parte, y para confirmar la certeza que se tiene sobre los servicios críticos nombrados, según un estudio realizado en el trabajo de grado: “*Criterios Para Establecer Políticas De Seguridad De La Información Y Plan De contingencia, Caso de Estudio El Centro de Datos de La Universidad del Cauca*” [54], desarrollado en la FIET por Carolina Guevara Campo y Fabián Andrés Mera (actual administrador del CDU de la Universidad del Cauca), se hizo la valoración de los activos de los servicios del CDU, y se dio un valor de 10 a los servidores web (la valoración más alta, en una escala de 1 a 10), lo que quiere decir que un fallo en estos servicios, pueden generar grandes pérdidas y grandes consecuencias, ya que cuentan con una alta prioridad por ser la cara visible desde la Universidad hacia Internet. Igualmente el servicio Web, recibió la valoración más alta de RR (Riesgo Residual) entre los servicios, lo cual determina la exposición de los elementos, en cuanto sea más alto el nivel de riesgo, es más necesario el empleo de mecanismos y medios para controlar estos elementos y se encuentra necesaria la implantación de mecanismos para evitar amenazas las cuales pueden afectar la disponibilidad de los mismos.

4.4. FASE IV: Estado del Arte de las Tecnologías de HA existentes

La disponibilidad de los servicios y la gestión de la misma es en este momento un factor crítico en las redes de datos, por eso en el mercado existen muchas empresas que han centrado su interés en el desarrollo de herramientas y productos diseñados para este fin.

La construcción del estado del arte al respecto, recoge las propuestas más consolidadas y conocidas en el mercado, con las cuales se implementó la tabla propuesta en la guía metodológica, con los siguientes resultados:

CARACTERÍSTICAS DE TECNOLOGÍAS HA DISPONIBLES						
SOFTWARE	FABRICANTE	REQUISITOS HW/SW	PRECIO (Dólares)	Licencia	SOPORTE (Tiempo y Costo)	OBSERVACIONES ventajas y desventajas
VMware vSphere 4 VMware vSphere Essentials Plus: Es un producto todo en uno, está proyectado para pequeñas y medianas empresas.	VMWare	Independiente del sistema operativo. \$4999 Sin límites de host \$1495 Limite 3 ESX hosts	\$2995 Fecha consulta Mayo-2010	Propietaria	-Por separado. -Se requiere subscripción. -Tres tipos de soporte: producción, básico y por incidente. -Por 1,2 ó 3 años. Lunes a viernes, desde 6am hasta 6pm. Colombia: 01-007-02057 Prestación vía telefónica, web o mail. El de producción es un soporte 24x7x365. El básico es un soporte 12x5, 12	VENTAJAS: -Control integrado de las aplicaciones. Es Independiente del sistema operativo. -Virtualización de recursos de los servidores (consolidación) -Automatización de los servicios. DESVENTAJAS: -Requiere la renovación de licencia

					horas diarias. El soporte por incidente es 12x5, restringido a un número de incidentes según el producto.	anualmente.
VMware ESXi 4 hypervisor for 32-bit hardware	VMWare	Independiente del sistema operativo.	Gratuito	Propietaria	Telf.: 1-877-486-9273 \$1149 por soporte para 5 incidencias \$749 por tres incidencias al año \$299 por una incidencia al año.	-Cobran por el soporte. Poca documentación para resolver conflictos.
SUSE Linux Enterprise High Availability Extension: Es una suite de tecnología clustering de código abierto que permite implementar clúster físicos y virtuales de alta disponibilidad.	Novell	Soporta las arquitecturas: x86_64 (AMD64 and Intel EM64T) IA64 (Itanium 2) - IBM POWER - IBM System z (64-bit). Se ejecuta sólo bajo Linux.	Depende del tipo de licencia y del volumen de licencias adquiridas.	Propietaria	El soporte y mantenimiento se debe adquirir por separado, mediante suscripción. Depende del tipo de licencia adquirida. Hay dos tipos de mantenimiento: - Prioritario: entrega un servicio 24x7 de clase mundial, con un tiempo general de 4 horas de respuesta ó de una hora cuando hay severidad. - Estándar: es un servicio 12x5 para apoyo solamente, con tiempo de respuesta de 4 horas.	VENTAJAS: -Ofrece agentes de recursos para numerosas aplicaciones de terceros y de código abierto. -Incluye guiones para la monitorización de servicios populares de código abierto, entre ellos: Apache, MySQL, NFS, Postgres, Squid, Tomcat,Xen. -Incluye guiones para aplicaciones de terceros, por ejemplo: Instance y Database de SAP, Oracle, IBM DB2, Informix y WebSphere, Vmware, eDirectorY. - Existe la opción de solicitar una licencia VLA.
IBM iCluster 5.2	IBM	ND	ND	Propietaria	ND	DESVENTAJA: • La licencia se vende para cada procesador.
HA OSCAR (High Availability Open Source Cluster Application Resources)			Gratuito	GPL	ND	DESVENTAJA: • Es un sistema cluster sólo para alto rendimiento.
LVS (Linux Virtual Server)	Wensong Zhang, Comunidad LVS	El nodo director o balanceador debe	Gratuito	GPL	- El soporte es gratuito y se presta a través de mailing list del proyecto.	VENTAJAS: Permite construir clúster de servidores con

		ejecutarse bajo Linux.			- Requiere suscripción al mailing list, la cual es gratuita.	máquinas sencilla, como por ejemplo PCs. DESVENTAJAS : Se necesita modificar (parchar) el sistema operativo.
Ultramonkey	Ultramonkey Org.	Soporte para SO Linux. Virtualmente soporta cualquier SO.	Gratuito	GPL	- Es ofrecido por la comunidad de desarrollo o a través de otros proyectos. - El soporte es gratuito mediante correo electrónico	VENTAJAS: • Es un proyecto de código abierto, hace uso de otros proyectos de código abierto. • Es gratuito. DESVENTAJAS: • El proyecto está cerrado.
HA OpenSolaris Cluster	Sun Microsystem	Actualmente, la versión 2009.06 sólo corre bajo OpenSolaris.	Gratuito	Libre	Gratuito. Brindado por la comunidad OpenSolaris	VENTAJAS: • Es una solución proveniente de un entorno exclusivo para servidores. • Cuenta con una comunidad de desarrollo. • Está integrado con SO Open Solaris 2009.06 el cual es gratuito. DESVENTAJA: • Admite clúster de sólo 2 nodos.
Veritas Cluster Server	Symantec	Soporta Linux, Windows, Solaris, AIX.	ND	Propietaria	Se debe adquirir por separado. El precio depende del tipo de soporte adquirido. Ofrece tres tipos de soporte: Básico: se presta en horario laboral, tiempo de respuesta de 1 hora. Esencial: servicio de 24x7, tiempo de respuesta de 30 minutos. Crítico: servicio 24x7, tiempo de respuesta 15 minutos.	VENTAJAS: • Automatiza y acelera la conmutación ante errores de aplicaciones del centro de datos. • Admite plataformas de SO heterogéneas. • Admite clúster de 32 nodos.

VMWare Server	VMWare	Compatible con la arquitectura x86. Soportado en SO: Windows, Linux y Mac	Gratuito	Libre		VENTAJAS <ul style="list-style-type: none"> • Permite crear múltiples máquinas virtuales sobre un solo host. • Proporciona rápido provisionamiento de servidores. • Es gratuito • Soporta varios SO.
VirtualBox	Sun Microsystems	- Soportado en hardware x86 y AMD64/Intel64. - Se ejecuta en Windows, Linux, Macintosh y OpenSolaris.	Gratuito	GPL		VENTAJAS: <ul style="list-style-type: none"> • Es potente y confiable. • Permite uso comercial como doméstico.
Xen Virtual Machine	Comunidad Xen	- Compatible con arquitecturas x86, x86_64, IA64, ARM. - Soportado en SO Linux.	Gratuito	Libre		VENTAJAS: <ul style="list-style-type: none"> • Se ejecuta directamente sobre el hardware y se convierte en la interfaz para todas las solicitudes hardware. • Es capaz de ejecutar múltiples sistemas operativos de forma segura e independiente.
Quemu	Comunidad Quemu		Gratuito	LGPL		VENTAJAS: <ul style="list-style-type: none"> • Es un emulador que permite ejecutar un sistema operativo bajo cualquier plataforma. • Es un virtualizador que permite portar una máquina virtual sobre cualquier SO.
Virtuozzo	Parallels		ND	Propietaria	<ul style="list-style-type: none"> • El soporte se debe adquirir, es por un año. • Existen dos tipos de soporte: 	VENTAJAS: <ul style="list-style-type: none"> • Permite virtualización y automatización.

					<ul style="list-style-type: none"> • Gold: en horario de oficina, tiempo de respuesta en 4 horas para severidad 1, 8 horas para severidad 2y un día para otro tipo de severidad. • Platinum: servicio 24x7x365, cuatro tipos de severidad con tiempos de respuesta de 2, 4, 8 horas y un día laboral, respectivamente. 	
--	--	--	--	--	--	--

Tabla 4.8. Recopilación Tecnologías HA.

La tabla anterior es la guía de referencia para la exploración y selección, de los componentes software que hacen parte de la solución de alta disponibilidad.

4.5. FASE V: Definición del Sistema para Alta Disponibilidad

Gracias a los pasos anteriores, en este momento ya se tienen previstos los mecanismos y tecnologías que se pueden usar en la propuesta, los cuales corresponden a los mecanismos de Alta Disponibilidad proporcionados por *OpenSolaris*, cuyo sistema operativo tiene a disposición tecnologías muy interesantes e importantes que antes solo estaban disponibles en *Solaris* (las cuales restringían su uso por ser privativas y generar altos costos en su implementación), para confirmar que la solución prevista es una buena selección, se van a evaluar ciertos aspectos en profundidad para tener la certeza y definir claramente el software del sistema de Alta Disponibilidad que se va a implementar.

4.5.1. Paso 1: Establecimiento de bases conceptuales sobre mecanismos implícitos en el sistema HA.

A continuación se nombran algunas bases conceptuales que complementan el paso 1 de la fase cinco en el capítulo tres de la guía metodológica, ya que se hace énfasis en los conceptos relacionados con la solución que se está definiendo en esta fase:

4.5.1.1. Clúster de Alta Disponibilidad [55][56]

En este paso se hace referencia a este concepto, puesto que en la investigación realizada se ha notado que para aumentar altamente la disponibilidad, sobre todo para disminuir los puntos únicos de falla, se utilizan mucho las técnicas de clúster de alta disponibilidad, la cual se define como un conjunto de dos o más servidores, que se caracteriza por compartir el sistema de almacenamiento, y por estar constantemente monitorizándose

entre sí; de esta manera, que si se produce un fallo del hardware o de los servicios de alguna de las máquinas que forman el clúster, el software de alta disponibilidad es capaz de reiniciar automáticamente los servicios que han fallado en cualquiera de los otros equipos del clúster. Y cuando el servidor que ha fallado se recupera, los servicios se migran de nuevo a la máquina original. Las configuraciones más comunes en entornos de clústeres de alta disponibilidad son la configuración activo/activo y la configuración activo/pasivo.

- **Configuración Activo/Activo**

En una configuración activo/activo, todos los servidores del clúster pueden ejecutar los mismos recursos simultáneamente. Es decir, los servidores poseen los mismos recursos y pueden acceder a estos independientemente de los otros servidores del clúster. Si un nodo del sistema falla y deja de estar disponible, sus recursos siguen estando accesibles a través de los otros servidores del clúster.

- **Configuración Activo/Pasivo**

Un clúster de alta disponibilidad, en una configuración activo/pasivo, consiste en un servidor que posee los recursos del clúster y otros servidores que son capaces de acceder a esos recursos, pero no los activan hasta que el propietario de los recursos ya no esté disponible. Este tipo de configuración permite que el rendimiento de los nodos no se vea afectado al cambiar de nodo, por lo cual es la opción que se ha elegido para implementar en el sistema.

- **Componentes de un clúster**

Es necesario aclarar que un sistema clúster está formado habitualmente por componentes hardware y Software tales como los nodos y el middleware. Los nodos, corresponden a cada una de las máquinas que componen el clúster, pueden ser desde simples ordenadores personales a servidores dedicados, conectados por una red. Por regla general los nodos deben tener características similares: arquitectura, componentes y sistema operativo. Otro componente principal es el Middleware del clúster, el cual es el software que actúa entre el sistema operativo y los servicios o aplicaciones finales y representa la parte fundamental del clúster donde se encuentra la lógica del mismo.

4.5.2. Paso 2. Aspectos que se tuvieron en cuenta al elegir el sistema HA.

En este paso se investigó a fondo sobre las prestaciones del sistema HA elegido, se evaluaron características del clúster HA de *OpenSolaris*, y del sistema operativo *OpenSolaris*, teniendo en cuenta los siguientes Aspectos que respaldan su elección:

- ✓ **Alta Documentación:** este aspecto es muy importante para llevar a cabo toda la implementación del sistema HA, cuya información para *OpenSolaris* es muy buena, y hay gran respaldo por parte de la comunidad de *OpenSolaris*.

- ✓ **Costos:** Con el fin de minimizar costos, por pago de licencias y soporte, se optó por utilizar software libre, este aspecto se tuvo en cuenta ya que el CDU de la Universidad del Cauca, cuenta con limitaciones presupuestales para el proyecto. Por lo tanto, el software open Cluster HA cuenta con una Licencia CDDL (Common Development and Distribution License), la cual es libre y el soporte es gratuito. Igualmente el sistema de almacenamiento que se recomienda es *RAID5* ya que el CDU, cuenta con este tipo de hardware.
- ✓ **Disminución del SPOF:** El clúster de alta disponibilidad disminuye drásticamente los puntos únicos de falla en el servicio, puesto que se tienen los servicios funcionando en dos servidores idénticos ubicados en diferentes localizaciones, lo cual mitiga la interrupción del servicio, cuando uno de ellos sufre alguna falla o incidente.
- ✓ **Proyecto Estable:** Proyecto de código abierto patrocinado por Sun Microsystems. Incluye todas las innovaciones y tecnologías claves que ofrece el actual sistema operativo *Solaris 10*, Las futuras versiones de Solaris (11, 12...), serán construidas a partir del código fuente de OpenSolaris.
- ✓ **Compatibilidad** con el mayor número de servicios y recursos existentes. Ofrece gran cantidad de agentes disponibles para la implementación de los servicios.
- ✓ **Tendencia Tecnológica del producto:** Tecnologías novedosas y con grandes prestaciones del sistema HA escogido, tanto en su sistema operativo OpenSolaris que tiene uno de los mejores sistema de archivos *ZFS* (Zetabyte file System), como en las tecnologías incluidas en el sistema *Open HA Cluster de Opensolaris 2009.06*.
- ✓ **Mecanismos de seguridad robustos:** Tiene incorporado protocolos de seguridad avanzados y de alto prestigio como lo son IPsec (Internet Protocol Security), y además permite el manejo de Seguridad con *Tcp_wrappers/IPFilter*.
- ✓ **Compatibilidad con virtualización:** Permite manejar virtualización completa y paravirtualización, además incorpora tecnologías de virtualización como *Crossbow* que permiten Virtualizar las tarjetas de red.
- ✓ **Sistema capaz de generar Copias de Respaldo:** Permite la realización de imágenes de respaldo o snapshots del sistema, permitiendo tener un respaldo del mismo.

4.5.3. Paso 3: Inventario de componentes disponibles en la implementación del sistema HA.

Dentro de los componentes y recursos que existen actualmente en el CDU, y que se podrían aprovechar para el sistema HA, están: dos servidores Dell, en los que actualmente están instalados los servidores web y los switches Ethernet para la interconexión. Sin embargo, no se cuenta con los elementos para implementar un almacenamiento centralizado.

Para el caso específico del prototipo de la solución de Alta Disponibilidad que se está desarrollando, el inventario consta de un solo equipo, por lo tanto, se utilizarán

tecnologías de virtualización para simular todos los elementos que componen el sistema propuesto. Por ejemplo, se instalaran dos máquinas virtuales simulando los nodos manejados en el clúster HA y la interconexión de las redes se harán con switches y tarjetas virtuales, utilizando para ello, la tecnología de *COMSTAR* implícita en *OpenSolaris 2009.06*. El equipo disponible para el prototipo cuenta con las siguientes características: 4 Gigas de RAM, un procesador de 2.86 GHz y un disco duro de 200 Gigas.

4.5.4. Paso 4. Definición de la propuesta de solución acorde a las condiciones actuales del CDU.

- ✓ **Implementación de un clúster de Alta Disponibilidad de dos nodos geográficamente distribuidos**

Permite el manejo de cada nodo de manera independiente, en los cuales se instalen los servicios críticos definidos en el paso cuatro de la fase tres: el servidor de aplicaciones java (*Tomcat*), el servicio web apache y el servicio de base de datos *MySQL*, en dos servidores idénticos ubicados en el IPET y en la División de sistemas respectivamente (ver figura 4.2). De esta manera que cuando se detecta una falla de algún servicio en uno de los nodos, sin intervención manual se migren los servicios al otro nodo físico. Para lograr la redundancia física de los servidores web, se requiere implementar *Open High Availability Cluster (OHAC)*, el cual es la base de código abierto de Solaris Clúster, este tipo de clúster de HA, incluye un conjunto de agentes HA, los cuales trabajan de manera automatizada en caso de que alguno de estos agentes falle en uno de los nodos, el sistema sube inmediatamente el demonio del servicio con problemas, y en caso de que el servicio no responda o el nodo esté caído, el nodo pasivo sube inmediatamente y se activa el grupo de recursos en el nodo redundante.

- ✓ **Implementación de IPMP**

Por otra parte, se propone utilizar doble canal de datos, ya que si por ejemplo, se cae la conectividad se cae todo el clúster; para el manejo de múltiples interfaces físicas, se propone utilizar el sistema IPMP (IP Network multipathing), cuyo sistema permite configurar una o más interfaces físicas en un grupo de múltiples rutas IP o grupo IPMP, el sistema supervisa automáticamente las interfaces del grupo IPMP para detectar posibles errores; en caso de que falle una interfaz del grupo o si se elimina para fines de mantenimiento, IPMP migra automáticamente hacia otra interfaz del grupo o hace que no respondan las direcciones IP de la interfaz fallida. De esta manera, IPMP mantiene la conectividad e impide la interrupción de cualquier conexión.

- ✓ **Periodo operacional continuo y Respuesta automática ante fallas**

Con la implementación de la propuesta de solución, se pretende tener un periodo operacional continuo y minimizar las interrupciones de los servicios por mantenimiento. La migración en vivo de los sistemas completos en ejecución y el monitoreo de las máquinas que pertenecen al clúster (nodos) para detectar fallos de Sistema Operativo o del servicio,

permiten que los nodos se reinicien automáticamente, gracias a las técnicas de heartbeat y Failover implícitas en el sistema HA, obteniendo así redundancia de software y de hardware, lo cual repercute en disminución de los puntos únicos de falla.

✓ **Garantizar confidencialidad mediante IPsec**

En cuanto a la seguridad de la comunicación en el sistema HA, se propone utilizar la pila de protocolos IPsec, para garantizar la confidencialidad de la información que se envía desde un nodo a hacia el otro, de manera que los datos que viajan entre los nodos se envíen cifrados, utilizando fuertes algoritmos de encriptación.

✓ **Creación de imágenes del Sistema Operativo y de la configuración de los servicios:**

El sistema a implementar permite también, hacer snapshots (del sistema operativo), permitiendo ejecutar rollback, lo cual reduce el riesgo ante fallas en la configuración del sistema.

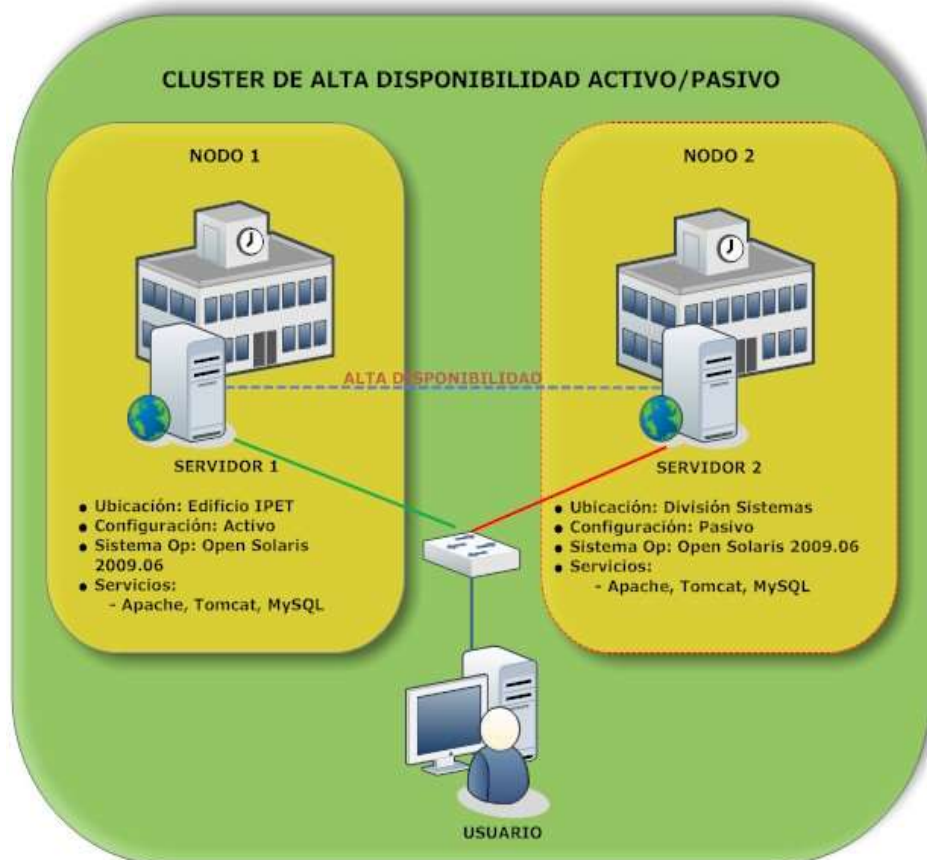


Figura 4.2. Propuesta de solución acorde a las condiciones actuales del CDU.

✓ Manejo Centralizado de los Datos:

En cuanto al almacenamiento de los datos, se recomienda un manejo centralizado de los mismos, se tuvo en primera instancia, como posibles candidatos a los sistemas de almacenamiento tipo NAS (Network Attached Storage) y tipo SAN (Storage Area Network), los cuales aunque parecen idénticos, tienen grandes diferencias: ambos permiten acceder a dispositivos de almacenamiento como arreglos de discos o cintas de backups, conectados a la red, los cuales permiten ser montados por diferentes servidores, siendo posible aplicar políticas de mantenimiento y backups propias a dichos dispositivos.

Los dispositivos NAS, trabajan directamente sobre redes IP, de manera que, comparte archivos, mediante protocolos como SAMBA o NFS; mientras que las redes SAN utilizan el protocolo SCSI o el Fiber Channel, y se comparten dispositivos de bloques, permitiendo la conexión directa entre servidores y dispositivos de almacenamiento (cabinas de discos, dispositivos de cintas magnéticas, etc.). El almacenamiento SAN tiene una relación de uno a uno con el servidor, donde cada dispositivo o LUN (Logical Unit Number) de la SAN es "propiedad" de un solo host o servidor, contrario a NAS, que permite a varios servidores compartir el mismo conjunto de archivos en la red. Una red SAN tiende a maximizar el aprovechamiento del almacenamiento, puesto que varios servidores pueden utilizar el mismo espacio reservado para crecimiento. Otra de las grandes ventajas de las redes SAN, es que al tener mayor conectividad, permiten que los servidores y dispositivos de almacenamiento se conecten por más de una vía, a la SAN, de esta forma, se pueden tener rutas redundantes que a su vez incrementan la tolerancia a fallos.

Una de las mayores limitaciones de utilizar tecnología SAN, es la implementación del canal de fibra, pues requiere de switches de Fibra y conexiones de fibra en cada servidor (lo cual aumenta los costos de su implementación); sin embargo este inconveniente está siendo mitigado gracias a la utilización del protocolo iSCSI (Internet Small Computer System Interface), el cual funciona sobre redes Ethernet, cuenta con el respaldo de los grandes fabricantes de equipos activos y está demostrando ser tan robusto, seguro y rápido como las clásicas redes Fibre Channel [57] [58] [59].

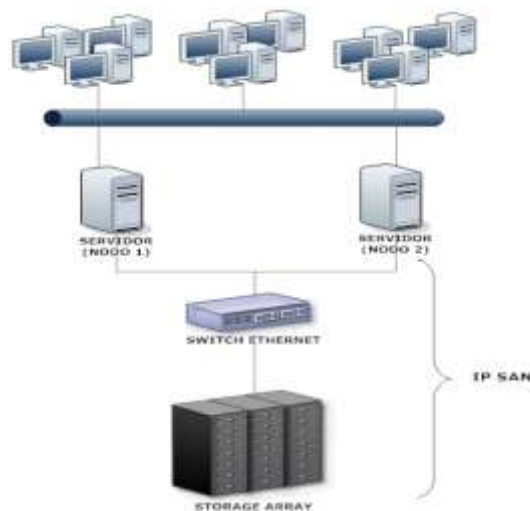


Figura 4.3. Red SAN-iSCSI

Por lo tanto, se propone como parte de la solución de Alta Disponibilidad, en cuanto al almacenamiento se refiere, utilizar idealmente una red SAN iSCSI (internet Small Computer System Interface), como se muestra en la figura 4.3. En la red SAN, se pueden compartir los discos por iSCSI, permitiendo acceder a los datos de una forma bastante transparente (como si el disco estuviera conectado directamente al hardware), ya que importa todo el dispositivo hardware por la red, de manera que en el cliente es detectado como un dispositivo SCSI más. De esta manera, SAN-iSCSI, se convierte en un sistema de almacenamiento ideal para la propuesta del proyecto de alta disponibilidad, ya que los nodos que componen el clúster, tienen acceso a los mismos storage (con un sistema redundante), de manera que si uno de los nodos se cae, entonces el otro, sube con la misma información.

✓ Usar Tecnologías de Virtualización

Se recomienda utilizar tecnologías de virtualización tales como *VirtualBox* o *XEN*, que permitan tener la mayor cantidad de aplicaciones aisladas y protegidas corriendo en máquinas virtuales diferentes. Para el caso de las aplicaciones web, según investigaciones a otros centros de datos, una buena práctica es tener los sitios web aislados en diferentes máquinas virtuales ya que por seguridad si se presenta problemas en uno de ellos, el resto no se verá afectado, buscando virtualizar diferentes servicios y consolidarlos en una misma máquina, de manera que se puedan optimizar los recursos y tener aislados cada uno de los servicios prestados, con el fin de impedir que un problema en uno de los sitios web, afecte al resto. En el anexo D se encuentra el manual de cómo implementar y configurar maquinas virtuales con *VirtualBox*, desde la línea de comandos, el cual puede resultar bastante útil para el administrador del CDU que lo desee implementar, puesto que la configuración y manejo de maquinas virtuales por línea de comandos, en un principio puede resultar algo tedioso, pero es una buena práctica quitar el entorno gráfico a las máquinas virtuales ya que este consume recursos del sistema.

Para la implementación del prototipo y las respectivas pruebas, debido a las limitaciones de equipos se implementará el clúster HA con dos maquinas virtuales, implementadas con *VirtualBox*, las cuales corresponden a los nodos A y B, tal y como se muestra en la Figura 4.4.

La migración en caliente de las máquinas Virtuales entre servidores físicos y la recuperación ante desastres (manejando clúster HA), son funcionalidades implícitas en soluciones comerciales que son muy costosas, ya que representan pagos por licencias, pero en *OpenSolaris* se puede realizar funciones similares haciendo uso de software libre. Por lo tanto esta propuesta representa una solución asequible, que reduce costos de gestión, ahorraría pagos por licencias, asesoría, soporte, y diseño. Además automatizando los procesos se disminuyen tiempos de interrupción y si se lleva a cabo la consolidación de más servicios en máquinas virtuales se podría disminuir los servidores físicos requeridos, lo cual implica también ahorro de energía; adicionalmente, con los servidores liberados se podrían realizar nodos redundantes para disminuir los puntos únicos de fallo, que son la clave para obtener Alta Disponibilidad.

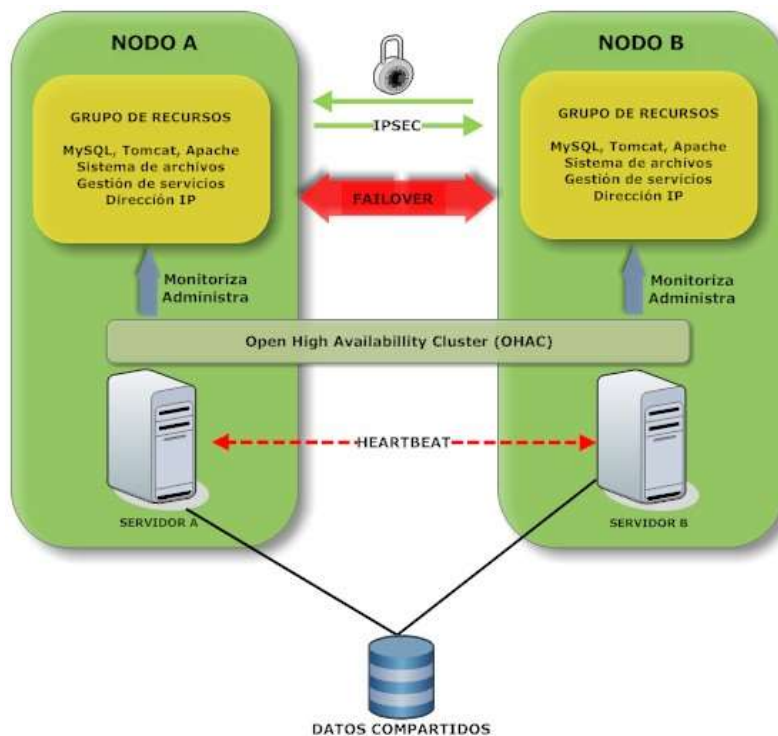


Figura 4.4. Esquema del sistema HA.

4.5.5. Paso 5. Profundizar en las tecnologías involucradas en la propuesta.

En este paso se pretenden resaltar las características más relevantes de las tecnologías que se escogieron y que forman parte de la propuesta, con el fin de tener un conocimiento claro sobre su funcionamiento y si es el caso tratar de aprovechar las facilidades que prestan estas tecnologías, ya que son de software libre, como es el caso de *OpenSolaris*.

✓ Sistema operativo OpenSolaris

OpenSolaris es un Proyecto de código abierto patrocinado por Sun Microsystems, OpenSolaris 2009.06 es la última versión del sistema operativo *OpenSolaris*, es un gran proyecto comunitario que Integra varias tecnologías innovadoras tales como IPS, ZFS, XVM, Zonas, Crossbow entre otras. Algo muy importante de destacar en el uso de este sistema operativo, es que cualquiera puede contribuir en el código y mejorar la tecnología del sistema operativo. *OpenSolaris* es un sistema operativo con aplicaciones de misión crítica en las cuales se puede cumplir con niveles de alta disponibilidad [60].

✓ Gestor de paquetes IPS

El Image Packaging System (IPS), es una herramienta para la gestión de software, enfocada hacia actividades de instalación, actualización y remoción. Mediante su utilización, se pueden realizar tareas como creación y gestión de imágenes, instalación de

nuevos paquetes y actualización de los ya existentes, todo a través de consola o de un gestor gráfico [61].

✓ **Virtualización en OpenSolaris**

Dentro de las tecnologías de virtualización existentes dentro de *OpenSolaris* se tienen *XVM* (Xen), *VirtualBox*, *Zonas* y *Crossbow* [60]

✓ **Zonas OpenSolaris**

Las zonas, son una abstracción del sistema operativo para la partición, lo que permite que varias aplicaciones se ejecuten en forma aislada unos de otros sobre el mismo hardware físico. Este aislamiento evita que los procesos se ejecuten fuera de la zona. Permite ejecutar varias “instancias” del sistema operativo a la vez, todas ellas ejecutándose en un único kernel [62].

Entre sus ventajas se pueden citar:

- Permite realizar snapshots y clones del sistema de archivos en caliente, por lo que el uptime del sistema aumenta considerablemente.
- Soporta ser exportado por NFS, cuotas de disco por grupos o usuarios, ACLs (Access Control Lists), encriptación, etc.

✓ **Redes con Crossbow**

Crossbow es un proyecto incluido en *OpenSolaris*, que proporciona la base para la virtualización de la red y el control de recursos alrededor de cualquier servicio (HTTP, HTTPS, FTP, NFS, etc) o máquina virtual. Mediante su utilización es posible Virtualizar por ejemplo la tarjeta de red, convirtiéndola en un grupo de VNIC'S, que permiten lograr una mayor eficiencia y optimización en el uso de la interfaz de red [63].

✓ **Seguridad en OpenSolaris**

Uno de los pilares de Solaris ha sido siempre la seguridad, característica que ha heredado *OpenSolaris* y sus tecnologías que proveen herramientas para el manejo de usuarios, archivos y el sistema operativo mismo de forma segura. *OpenSolaris* tiene herramientas como ACL, framework Crypto, IPFilter y otras tecnologías que hace realmente seguro el trabajo con *OpenSolaris* [60].

✓ **SMF (Service Management Facility)**

Es el nuevo sistema de arranque que incorpora *OpenSolaris* y que sustituye a los run level de System V. Representa una nueva infraestructura de servicios que permite arrancar los servicios de forma paralela en función de las dependencias que requiera la

aplicación. Una vez arrancado el servicio, el administrador podrá deshabilitar, observar, parar o arrancar servicios de una forma sencilla y cómoda [64].

La gran ventaja de SMF es que ofrece mecanismos por los cuales se establecen unas relaciones de dependencia entre servicios, de forma que un servicio no arrancará hasta que sus dependencias hayan arrancado correctamente. Para cada proceso de arranque de un servicio, se guarda un log, que informará de los pasos que ha seguido el servicio para arrancar y que ayudará a determinar la causa de error, en caso de que un servicio no pueda arrancar.

Se podría decir, que los servicios actúan como objetos, que pueden ser gestionados de forma sencilla por un grupo de comandos de administración, los cuales permitirán parar, arrancar, y ver el estado del servicio y sus dependencias.

Entre otras ventajas de este sistema de arranque se encuentran:

- Permite tomar snapshots de las configuraciones de los servicios, haciendo más sencillo el backup, y restore de cualquier cambio. La configuración es persistente entre reboots, upgrades y patches.
- Se pueden delegar tareas de administración de servicios a usuarios específicos.
- Los sistemas grandes bootan más rápido, porque utilizan el concepto de gestión de servicios en paralelo.

✓ **iSCSI**

Es un protocolo para comunicaciones de dispositivos. En iSCSI, los comandos SCSI que maneja el dispositivo se envían a través de la red, de forma que en lugar de tener un disco SCSI conectado físicamente al equipo, se conecta por medio de la red. iSCSI generalmente se usa para centralizar el almacenamiento en disco e independizar la información de los servidores. Por ejemplo, se podría tener un servidor iSCSI con 1 Terabyte de almacenamiento que centralice todos los dispositivos de almacenamiento de la red y tener los servidores sin particiones para los datos, los discos de datos se conectarán a través de iSCSI por la red y si alguno de los servidores se cayera o hubiera que hacerles mantenimiento, se tendría la información en un disco que se podría enchufar por la red. De manera que si se tuviese un servidor secundario, sería suficiente conectarle el disco iSCSI para tener el servicio funcionando en muy poco tiempo. OpenSolaris contiene el iniciador y el soporte de destino para el protocolo iSCSI que permite utilizar el protocolo SCSI en redes TCP/IP [65].

✓ **COMSTAR**

COMSTAR es un marco de software que le permite convertir cualquier máquina de OpenSolaris en un destino SCSI que se puede acceder a través de la red por los anfitriones de iniciador. De esta manera, COMSTAR rompe la enorme tarea de manipulación de un subsistema de destino SCSI en módulos funcionales independientes [66].

✓ IPsec

Es un protocolo de seguridad de Internet (IPsec), es un conjunto de normas abiertas para proteger las comunicaciones de una red sobre el protocolo internet (IP), a través del uso de los servicios de seguridad criptográfica. IPsec soporta autenticación a nivel de red, la autenticación del origen de datos, integridad de datos, confidencialidad de los datos (encriptación) y protección contra la replicación [67].

4.6. FASE VI: Implementación del SW y HW de la propuesta.

4.6.1. Paso 1: Requerimientos hardware y software para la implementación de la propuesta.

Los Requerimientos para la implementación de la propuesta son:

- 4 Tarjetas de interfaz de red estándar.
- Cableado Ethernet estándar.
- Switch Gigabit o 10 Gigabit Ethernet.
- Arreglo de discos SAN/iSCSI.
- 2 Servidores, con sistema operativo OpenSolaris 2009.06. (Con la restricción de que todos los nodos del clúster deben ejecutarse en la misma plataforma).
- Software Open HA Cluster 2009.06.
- Sistema operativo 2009.06.

Como se mencionó en el paso 3 de la fase 5, debido a las limitaciones de recursos; se implementó un prototipo de la propuesta realizada, con tecnologías de virtualización, que permite ver el funcionamiento del sistema planteado.

El equipo empleado para el prototipo cuenta con las siguientes características: 4 Gigas de RAM, un procesador de 2.86 GHz y un disco duro de 200 Gigas.

• **Open High Availability Cluster (OHAC)**

Es la base de código abierto de Solaris Cluster, una solución de clúster de alta disponibilidad (HA) de Sun Microsystems, Inc. Solaris Cluster incluye la estructura de clúster central, un conjunto de agentes HA para diversas aplicaciones, una estructura de prueba automatizada y una extensión de recuperación frente a desastres.

Requisitos:

- Mínimo de 1 Gbyte de memoria RAM física (2 Gbytes recomendado)
- Un mínimo de 6 Gb de espacio disponible en disco duro

La memoria real física y las características del disco duro están determinadas por las aplicaciones que se instalan.

4.6.2. Paso 2: Capacitación del Personal.

Al tratarse de un proyecto de investigación, este paso consistió en profundizar en la documentación de cada uno de los aspectos, mecanismos y elementos que integran la solución para alta disponibilidad. Uno de los aspectos que requirió mucha investigación fue sobre el sistema operativo (SO) *OpenSolaris*, ya que es un SO nuevo para el grupo, sobre el cual no se había trabajado anteriormente. Entonces se tuvieron casos específicos de cuidado, como por ejemplo el tema de las particiones, repositorios, y sistema de archivos, como también el manejo de maquinas virtuales para el manejo del clúster.

Otra consideración al respecto, es lo referente a la instalación y configuración de *Open HA Solaris*, el software para clustering de *OpenSolaris*, pues requiere de muchos procesos y procedimientos complejos, los cuales se debe realizar por línea de comandos, lo que exige conocer bien la ejecución de comandos en consola. Para ello igualmente fue necesario realizar una fuerte investigación, consultando el sitio web y foros de *OpenSolaris*.

En general la capacitación consta de las actividades enunciadas en la tabla 4.1, la cual se muestra a continuación:

CONOCIMIENTO REQUERIDO	ACTIVIDADES (investigar y socializar)	TIEMPO REQUERIDO
Instalación y configuración de red en <i>OpenSolaris</i> .	Instalación de <i>OpenSolaris</i> . Configuración de red por comandos.	Dos días.
Manejo de comandos	Configuración de repositorios y manejo del gestor de paquetes. Manejo del SMF (Service Management Facility) y de ZFS.	Tres semanas.
Instalación y configuración de maquinas virtuales por consola.	Instalación y configuración de <i>VirtualBox</i> .	Tres semanas.
Instalación y configuración del clúster HA	Configuración de los nodos en el Clúster HA. Creación de grupos y recursos.	Dos semanas.
Instalación y configuración de los servicios.	Instalar y configurar <i>Mysql</i> , <i>Apache</i> y <i>Tomcat</i> . Instalar un prototipo del sitio Web del portal de <i>Unicauca</i> .	Una semana.

Tabla 4.9. Plan de capacitación.

4.6.3. Paso 3: Instalación y configuración del sistema de almacenamiento.

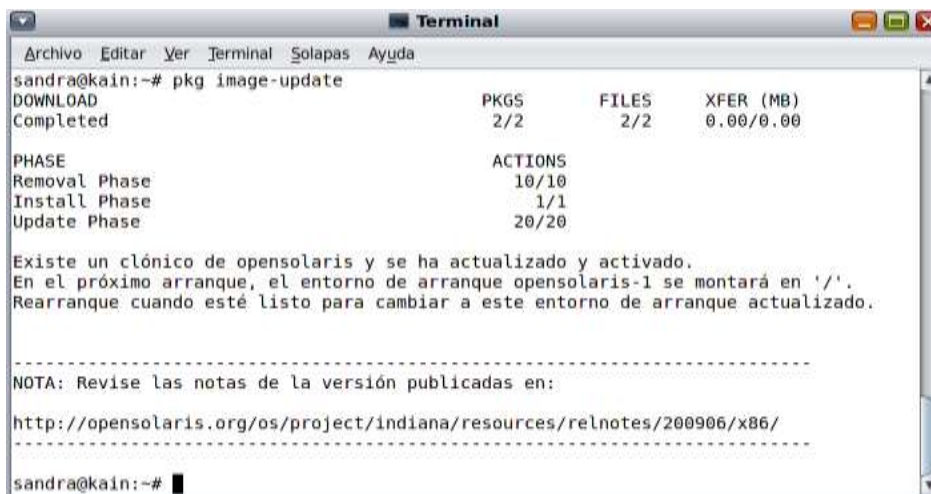
Debido a las limitantes de acceso al hardware, solo se contó con un equipo en el cual se configuró un clúster con máquinas virtuales, creadas con *VirtualBox*. En cada uno de los nodos, se instaló el sistema operativo *OpenSolaris 2009.06*, el cual trae consigo la tecnología *COMSTAR/iSCSI* y se usaron las prestaciones y ventajas que ofrece *COMSTAR*, como son, el permitir convertir cualquier host de *Solaris* en un destino de *SCSI*, de manera que se puede crear arreglos de discos y compartir información entre los

nodos, COMSTAR permite que toda clase de dispositivos SCSI (cinta, disco, etc.) se conecten a un medio de transporte (en este caso por la red ip), gracias a esta herramienta se creó un arreglo de discos con el cual se comparten recursos entre los nodos del clúster (ver anexo D).

4.6.4. Paso 4: Endurecimiento del servidor.

Se instaló el sistema operativo *OpenSolaris 2009.06* y se inicio con el procedimiento de endurecimiento del servidor, de lo cual se muestran a continuación los pasos y evidencias del procedimiento para asegurar el servidor Kain, el cual fue el que se utilizó en la implementación del prototipo [68]:

a. Actualización del sistema



```
Terminal
Archivo Editar Ver Terminal Solapas Ayuda
sandra@kain:~# pkg image-update
DOWNLOAD          PKGS      FILES    XFER (MB)
Completed         2/2       2/2      0.00/0.00

PHASE              ACTIONS
Removal Phase     10/10
Install Phase     1/1
Update Phase      20/20


Existe un clónico de opensolaris y se ha actualizado y activado.
En el próximo arranque, el entorno de arranque opensolaris-1 se montará en '/'.
Rearranque cuando esté listo para cambiar a este entorno de arranque actualizado.

-----
NOTA: Revise las notas de la versión publicadas en:
http://opensolaris.org/os/project/indiana/resources/relnotes/200906/x86/
-----
sandra@kain:~#
```

Figura 4.5. Actualización del sistema

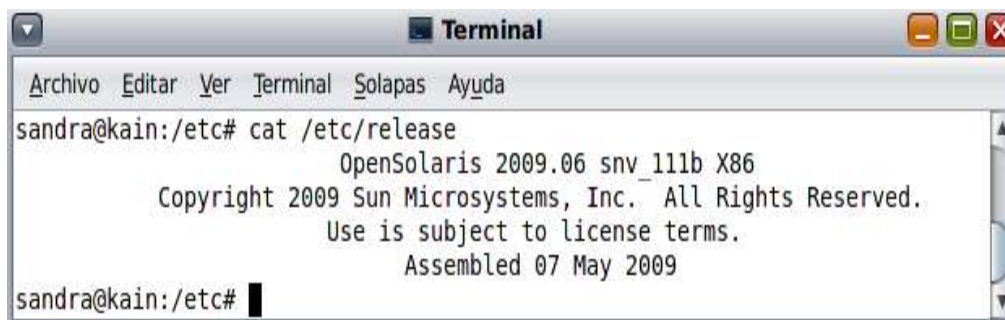
b. Borrar información del sistema

En este paso se debe editar los archivos *modt* y *relese* y borrar información que muestra la versión y binarios del sistema:



```
Terminal
Archivo Editar Ver Terminal Solapas Ayuda
sandra@kain:~# cat /etc/modt
Sun Microsystems Inc.  SunOS 5.11      snv_111b      November 2008
sandra@kain:~#
```

Figura 4.6. Borrar información del sistema.

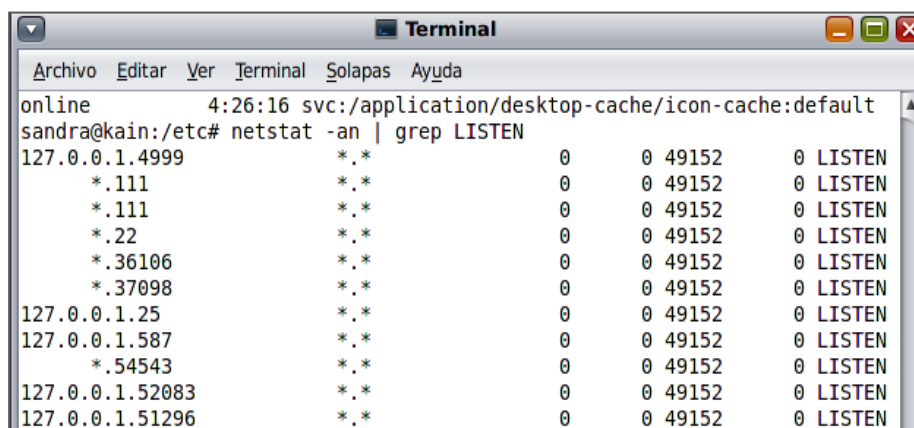


```
Terminal
Archivo Editar Ver Terminal Solapas Ayuda
sandra@kain:/etc# cat /etc/release
      OpenSolaris 2009.06 snv_111b X86
      Copyright 2009 Sun Microsystems, Inc. All Rights Reserved.
      Use is subject to license terms.
      Assembled 07 May 2009
sandra@kain:/etc#
```

Figura 4.7. Borrar mensaje de versión del sistema.

c. Deshabilitar servicios innecesarios.

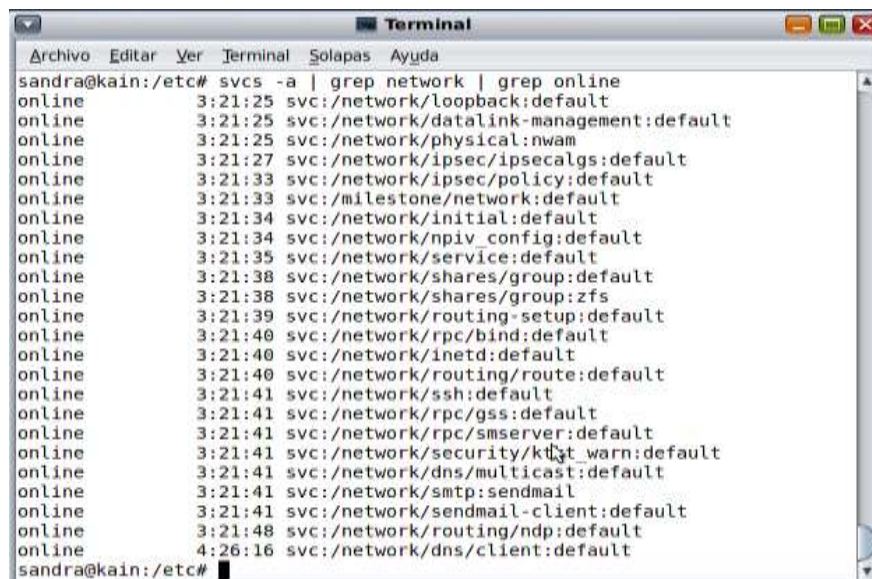
Eliminar procesos que no sean fundamentales: usando el comando: `#svcs -a | grep online`, permite ver todos los servicios que están arriba, pero no es suficiente saber esto, ya que pueden haber procesos que son propios del sistema, por lo tanto se debe usar el comando `#netstat -an | grep LISTEN`, para saber que puertos están abiertos a los cuales se puede acceder desde afuera. Los puertos que están con prefijo 127 son puertos que usan conexiones con el propio equipo, pero si los que tienen asterisco son los puertos que están abiertos hacia afuera por tanto estos son los que se deben tener en cuenta en esta prueba. `netstat -an | grep LISTEN`.



```
Terminal
Archivo Editar Ver Terminal Solapas Ayuda
online 4:26:16 svc:/application/desktop-cache/icon-cache:default
sandra@kain:/etc# netstat -an | grep LISTEN
127.0.0.1.4999      *.*          0           0 49152       0 LISTEN
*.111             *.*          0           0 49152       0 LISTEN
*.111             *.*          0           0 49152       0 LISTEN
*.22              *.*          0           0 49152       0 LISTEN
*.36106           *.*          0           0 49152       0 LISTEN
*.37098           *.*          0           0 49152       0 LISTEN
127.0.0.1.25      *.*          0           0 49152       0 LISTEN
127.0.0.1.587     *.*          0           0 49152       0 LISTEN
*.54543           *.*          0           0 49152       0 LISTEN
127.0.0.1.52083   *.*          0           0 49152       0 LISTEN
127.0.0.1.51296   *.*          0           0 49152       0 LISTEN
```

Figura 4.8. Servicios y puertos abiertos

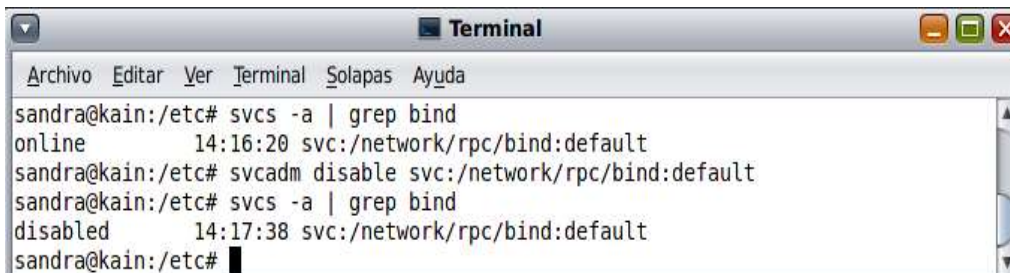
En la captura de pantalla siguiente, se ven los procesos de *network* corriendo, de los cuales se deshabilitan los que no se necesitan por el momento ya que el objetivo es tener una configuración estándar, para que luego a medida que se requiera se habiliten solo los puertos necesarios.



```
sandra@kain:/etc# svcs -a | grep network | grep online
online      3:21:25 svc:/network/loopback:default
online      3:21:25 svc:/network/datalink-management:default
online      3:21:25 svc:/network/physical:nwam
online      3:21:27 svc:/network/ipsec/ipsecalgs:default
online      3:21:33 svc:/network/ipsec/policy:default
online      3:21:33 svc:/milestone/network:default
online      3:21:34 svc:/network/initial:default
online      3:21:34 svc:/network/npiv_config:default
online      3:21:35 svc:/network/service:default
online      3:21:38 svc:/network/shares/group:default
online      3:21:38 svc:/network/shares/group:zfs
online      3:21:39 svc:/network/routing-setup:default
online      3:21:40 svc:/network/rpc/bind:default
online      3:21:40 svc:/network/inetd:default
online      3:21:40 svc:/network/routing/route:default
online      3:21:41 svc:/network/ssh:default
online      3:21:41 svc:/network/rpc/gss:default
online      3:21:41 svc:/network/rpc/smserver:default
online      3:21:41 svc:/network/security/kt_warn:default
online      3:21:41 svc:/network/dns/multicast:default
online      3:21:41 svc:/network/smtp:sendmail
online      3:21:41 svc:/network/sendmail-client:default
online      3:21:48 svc:/network/routing/ndp:default
online      4:26:16 svc:/network/dns/client:default
sandra@kain:/etc#
```

Figura 4.9. Procesos de network corriendo.

Por ejemplo, el proceso *inetd* se deshabilita porque es un demonio que puede iniciar varios servicios a la vez que no son necesarios, ya que el servidor sólo debe tener habilitado el servicio necesario, por ejemplo, si es un servidor de DNS sólo debe tener ese servicio *network/rpc/bind:default*.



```
sandra@kain:/etc# svcs -a | grep bind
online      14:16:20 svc:/network/rpc/bind:default
sandra@kain:/etc# svcadm disable svc:/network/rpc/bind:default
sandra@kain:/etc# svcs -a | grep bind
disabled    14:17:38 svc:/network/rpc/bind:default
sandra@kain:/etc#
```

Figura 4.10. Deshabilitar bind.

Así mismo se deshabilitan otros servicios que no se necesitan. El superdemonio se deshabilita en el archivo */etc/inetd.conf*, para esto, se comentan todas las líneas de ese archivo.

d. Asegurar la administración remota:

- Inicialmente se cambia el puerto por defecto a otro que no sea un puerto conocido, ósea un puerto mayor a 1024:

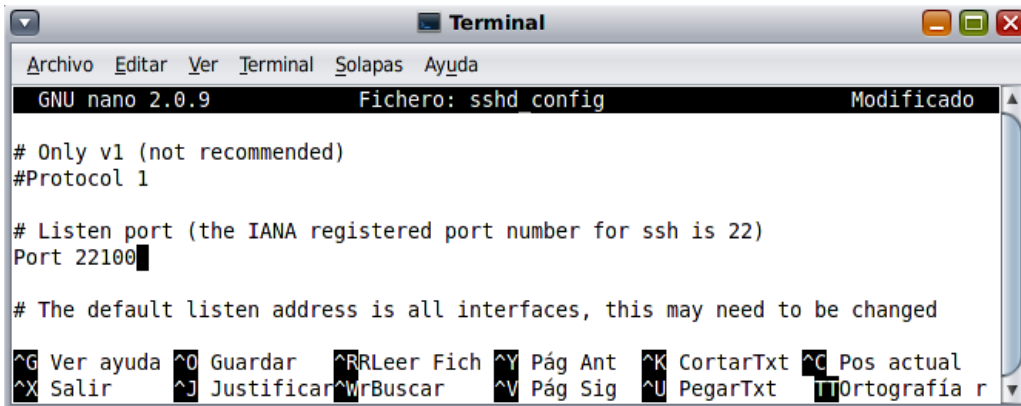


Figura 4.11. Cambio de puerto

- Asegurar que el `root` no se pueda loguear por SSH, ósea que al entrar directamente primero tendría acceder como un usuario normal:



Figura 4.12. No permitir logueo de root.

- Que sólo le permita a los usuarios autorizados acceder por SSH, en este caso al usuario `sandra` , tal como se muestra en la siguiente gráfica:

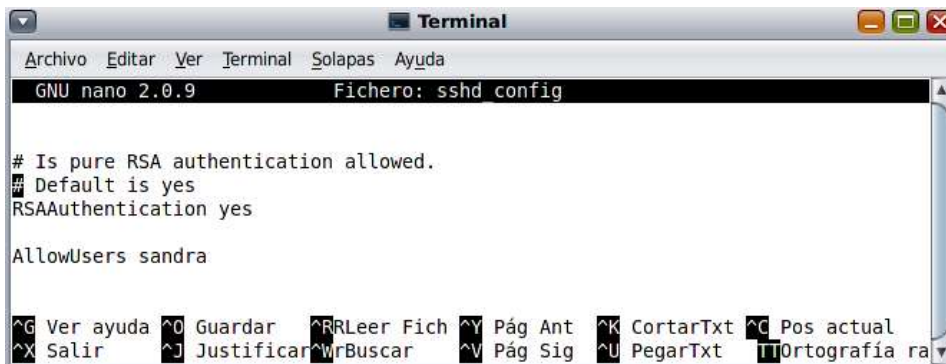


Figura 4.13. Configuración usuarios permitidos.

- Descomentar la línea *CONSOLE* en el archivo */etc/default/login* permite que para cualquier servicio el root pueda loguearse sólo desde el mismo equipo. Esto no sólo aplica a SSH sino a cualquier aplicación de consola remota.

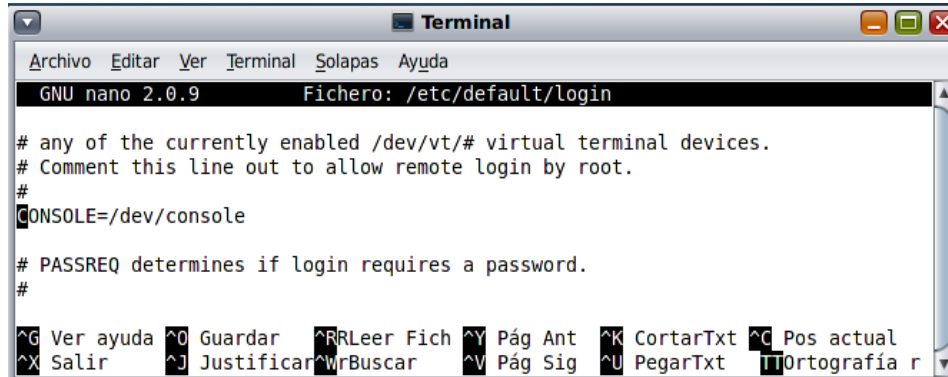


Figura 4.14. Prohibición de logueo remoto al usuario root.

- Bloquear todas las cuentas administrativas para esto en el archivo */etc/shadow* poner **LK** en el campo del *password* de los usuarios que se quiere bloquear *sys*, *uucp* *nuucp* y *listen*, para que en caso de que alguien llegase a conseguir un password de esas cuentas no pueda loguearse, ya que éstos son usuarios del sistema y no deben tener shell, este privilegio lo deben tener sólo los usuarios autorizados.

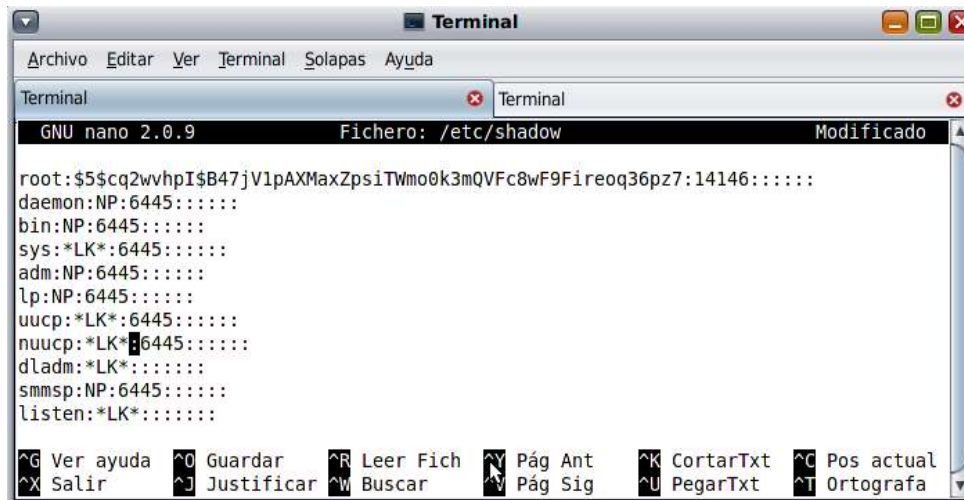
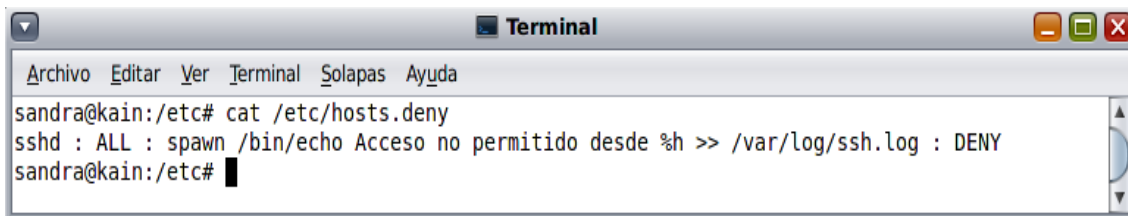


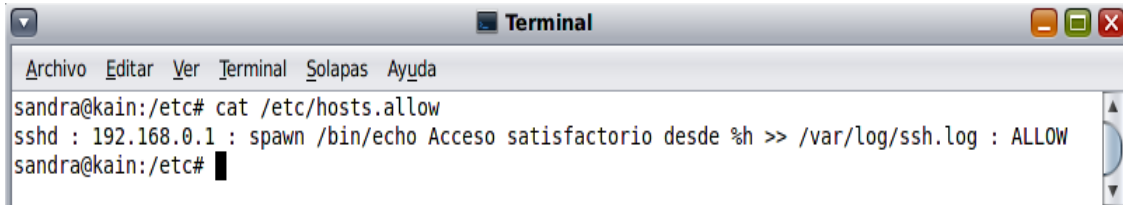
Figura 4.15. Bloqueo de cuentas administrativas.

- Se crean los archivos */etc/hosts.deny* y */etc/hosts.allow* primero se niega todo y luego se va habilitando a los necesarios para que puedan acceder solo desde los equipos autorizados. Estos archivos no están creados por defecto en *OpenSolaris*, por lo tanto hay que crearlos, contrario a como sucede con otros sistemas operativos como Linux.



```
Terminal
Archivo Editar Ver Terminal Solapas Ayuda
sandra@kain:/etc# cat /etc/hosts.deny
sshd : ALL : spawn /bin/echo Acceso no permitido desde %h >> /var/log/ssh.log : DENY
sandra@kain:/etc#
```

Figura 4.16. Archivo deny.



```
Terminal
Archivo Editar Ver Terminal Solapas Ayuda
sandra@kain:/etc# cat /etc/hosts.allow
sshd : 192.168.0.1 : spawn /bin/echo Acceso satisfactorio desde %h >> /var/log/ssh.log : ALLOW
sandra@kain:/etc#
```

Figura 4.17. Archivo allow.

e. Implementación de Firewall [69]

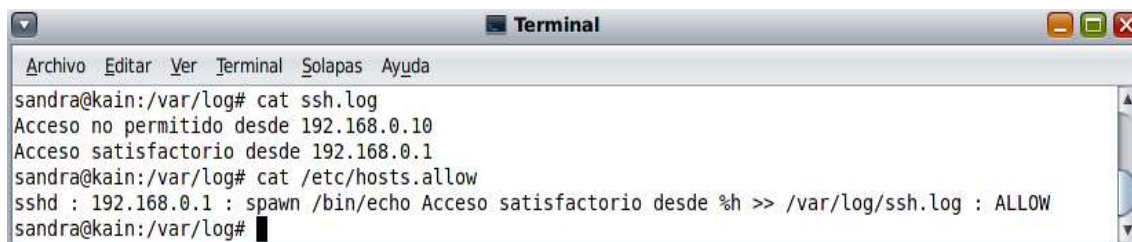
Normalmente en Linux primero se instala y se habilita *tcpwrappers*, pero en *OpenSolaris* ya está *tcpwrappers* habilitado e instalado. La herramienta *tcpwrappers* es un sistema de red ACL(Listas de Control de Acceso), que trabaja en terminales y que se usa para filtrar el acceso de red a servicios de protocolos de Internet que corren en sistemas operativos tipo UNIX.

Para usar *tcpwrappers* los archivos *host.deny* y *host.allow* se pueden configurar de la siguiente manera:

- *sshd: ALL : spawn /bin/echo `bin/date` intento de acceso no autorizado desde %h >> /var/log/ssh.log : DENY*

Explicación de los comandos usados:

- **sshd** corresponde al demonio de ssh (Secure Shell), **all**=indica a quien le aplica la regla ósea a todas las direcciones IP y todos los dominios, y **spawn** es la acción de negarle y que guarde en un log el intento de acceso no autorizado.



```
Terminal
Archivo Editar Ver Terminal Solapas Ayuda
sandra@kain:/var/log# cat ssh.log
Acceso no permitido desde 192.168.0.10
Acceso satisfactorio desde 192.168.0.1
sandra@kain:/var/log# cat /etc/hosts.allow
sshd : 192.168.0.1 : spawn /bin/echo Acceso satisfactorio desde %h >> /var/log/ssh.log : ALLOW
sandra@kain:/var/log#
```

Figura 4.18. Ejemplo de acceso denegado.

4.6.5. Paso 5: Instalación del software Open HA Cluster.

En este paso se instaló todo el sistema de HA en *Open Solaris 2009.06*, del cual se generaron los manuales de instalación, configuración y de usuario, los cuales quedan como anexo D del presente trabajo de grado. Ver Anexo D, el cual contiene Manuales de instalación y configuración del *Open HA Cluster* (Creación de imagen, políticas de seguridad en el tráfico de los nodos con IPsec, virtualización por consola con *VirtualBox* y manejo de Crossbow para la configuración del switch y de las tarjetas de red virtuales).

Sin embargo, en este capítulo se describe un breve resumen de los pasos más relevantes a seguir en la instalación, ya que este procedimiento es bastante extenso:

- 1) Por motivos de que solo se contaba con un servidor, se simularon los nodos mediante máquinas virtuales las cuales se crearon utilizando *VirtualBox*, se instaló el software *OpenSolaris* en cada uno de los nodos del Clúster HA.
- 2) Configuración de los repositorios del *Open HA Cluster*. Para esto fue necesario registrarse en la página de Sun para obtener una cuenta, lo cual permite crear los certificados SSL¹⁶ personales y acceder al contenido de paquetes adicionales del sistema *Open Solaris*, tales como drivers , aplicaciones y otros programas relacionados con los servicios a instalar.



Figura 4.19. Registro en Sun.

Luego se instalaron los respectivos certificados en cada uno de los nodos y se configuraron los repositorios necesarios para la instalación de todos los paquetes.

- 3) Instalación del Software *Open HA Cluster*. El paquete *ha-cluster-completo* contiene los agentes del servicio de datos, *el framework*¹⁷ del Cluster Server y otras herramientas necesarias que se explican detalladamente en el Anexo D.

¹⁶ Secure Socket Layer, protocolo de capa de conexión segura.

¹⁷ Conjunto estandarizado de API's (Application programming Interface), que cumplen funciones específicas.

```

sandra1@os-ohac-1:~$ ls
17-3-capture-crash-dump  Downloads                                P blico
Desktop                  Open_HA_Cluster_2009.06.certificate.pem
Documentos              Open_HA_Cluster_2009.06.key.pem
sandra1@os-ohac-1:~$ date
viernes 5 de marzo de 2010 09:46:02 COT
sandra1@os-ohac-1:~$ pkg info -r ha-cluster-full
Nombre: ha-cluster-full
Resumen: Sun Cluster full installation group package
Categor a: System/HA Cluster
Estado: Instalado

Editor: ha-cluster
Versi n: 2009.6
Versi n: 5.11
Ramificaci n: 0.111
Fecha de empaquetado: Fri May 15 18:29:33 2009
Tama o: 13.24 kB
FMRI: pkg://ha-cluster/ha-cluster-full@2009.6,5.11
-0.111:20090515T182933Z
sandra1@os-ohac-1:~$

```

Figura 4.20.Cluster HA instalado en servidor Kain.

4.6.6. Paso 6: Instalaci n y configuraci n de los servicios.

Se instalaron y configuraron los servicios *Tomcat*, *Apache* y *MySQL*, sobre los nodos del cl ster HA. Luego de corroborar el buen funcionamiento del cl ster con los servicios se instal  el portal de la universidad del cauca. Todas estas evidencias quedan en los manuales de instalaci n y configuraci n como anexo D debido a la complejidad y extensi n de cada uno de los pasos. Ver anexo D: Instalaci n y configuraci n de los servicios en el cl ster HA: balanceo de carga en apache y Failover en Tomcat y MySQL. Adem s se hizo una imagen del sistema para mitigar riesgos por da os en las configuraciones y toda la instalaci n se hizo por comandos, puesto que en un servidor lo m s recomendable es que el entorno gr fico se debe quitar para disminuir el consumo de recursos del sistema.

4.7. FASE VII: Evaluaci n de la disponibilidad

Se instalaron y configuraron los servicios *Tomcat*, *Apache* y *MySQL*, sobre el cl ster HA. Se hicieron las pruebas de funcionamiento de los nodos y luego de corroborar el buen funcionamiento del cl ster con los servicios, se instal  el portal de la Universidad del Cauca. Para evaluar la disponibilidad en el prototipo implementado, se realizaron los siguientes pasos:

4.7.1. Paso 1. Defini n de escenarios de pruebas.

En este paso, se realiza una evaluaci n de disponibilidad con el fin de tener elementos de comparaci n entre el prototipo propuesto y el funcionamiento real del sistema. Para la realizaci n de las pruebas se siguieron los pasos que se describen a continuaci n:

a) Definición del alcance y objetivos

El objetivo de esta evaluación es comprobar que la implementación del prototipo propuesto para la solución de Alta disponibilidad cumpla con los requerimientos de disponibilidad mínimos necesarios en un entorno redundante que le permita minimizar la interrupción del servicio, ante eventuales fallas en el servidor.

Como se está tratando con un sistema que no está en producción, para llevar a cabo las pruebas se simularán situaciones anormales, como por ejemplo la caída de un nodo, lo cual puede ser a causa de problemas eléctricos, o cuando se daña el aire acondicionado o por algún motivo en el que el servidor donde están los servicios de *Apache*, *Tomcat* y *MySQL* dejan de funcionar. Para esto, se ha instalado y configurado un clúster con dos nodos virtuales, en los cuales se matan los procesos en uno de los nodos y se verifica que automáticamente vuelvan a subir en el nodo redundante. Además de esto también se pretende comprobar que la información que circula entre los nodos está siendo debidamente encriptada y si el sistema implementado permite realizar imágenes o snapshots de la configuración de los servicios.

b) Definición de herramientas

Las herramientas utilizadas para llevar a cabo las pruebas mencionadas son:

- **Wireshark:** Es un analizador de protocolos de red, el cual permite configurar la tarjeta de red de un computador en modo “promiscuo” para capturar todos los paquetes que circulen en el mismo segmento de red para identificar y analizar este tráfico de manera detallada.
- **Siege:** Es un probador de estrés. Permite a un servidor web configurarle un número de usuarios concurrentes simulados. Esta herramienta sirvió para generar tráfico http y generar reportes de rendimiento.
- **Nagios:** Herramienta de gestión, que permite ver el estado de cada uno de los servicios que están siendo monitoreados.
- **Snoop:** Es un *sniffer* propio de Solaris para capturar los paquetes que se envían o llegan a cierta interfaz de red.

c) Identificación de los elementos de prueba:

Las pruebas se realizaron sobre la red prototipo implementado, según lo expuesto en la fase 6. La Tabla 4.10 muestra la descripción de las principales características de los cuatro elementos que van a ser evaluados, así como las del equipo auditor.

NOMBRE DEL EQUIPO	DIRECCIÓN IP	UTILIDAD EN LA RED	HERRAMIENTAS INSTALADAS
Os-ohac-1	10.200.2.1	Nodo 1, servidor Web	HA Cluster OpenSolaris, Apache, tomcat, MySQL Server, IPtables, OpenSSH, IPsec, ipfilter
Os-ohac-2	10.200.2.2	Nodo 2, servidor Web	HA Cluster OpenSolaris, Apache, tomcat, MySQL Server, IPtables, OpenSSH, IPsec, ipfilter
Kain	192.168.120.156	Servidor Físico contenedor del clúster HA	Open HA Cluster
Fénix	192.168.120.155	Equipo auditor	Nagios, siege, Wireshark

Tabla 4.10. Descripción de los equipos involucrados en las pruebas.

a) Definición del escenario de pruebas

Para verificar la disponibilidad de los servicios a pesar de que se caiga un nodo, ya sea por motivos de algún daño de hardware, recalentamiento, consumo de procesamiento o bloqueo del sistema, se ha definido el siguiente escenario de prueba (ver figura 4.21. Escenario de pruebas):

Generación de tráfico con la herramienta *Siege* instalada en el equipo fénix en el cual se simulan 1, 10, 50 y 100 usuarios; esta herramienta genera tráfico hacia el puerto 8080 (Tomcat) y 80 (apache) del servidor Kain, el cual contiene los nodos virtuales que forman parte del clúster, y sobre los cuales están los respectivos servicios instalados. En el servidor físico Kain, gracias a la herramienta de *nmap* e *ipfilter* se puede tener acceso a los puertos necesarios tales como *MySQL*, *Apache* y *Tomcat*; para de esta forma poderlos monitorear y manipular. Por otro lado también se instaló la herramienta *Nagios* en el equipo de monitoreo para visualizar la caída y subida de los procesos (de cada prueba se generó el video respectivo), y paralelo a las pruebas se visualizaban los paquetes transmitidos con la herramienta *Wireshark*.

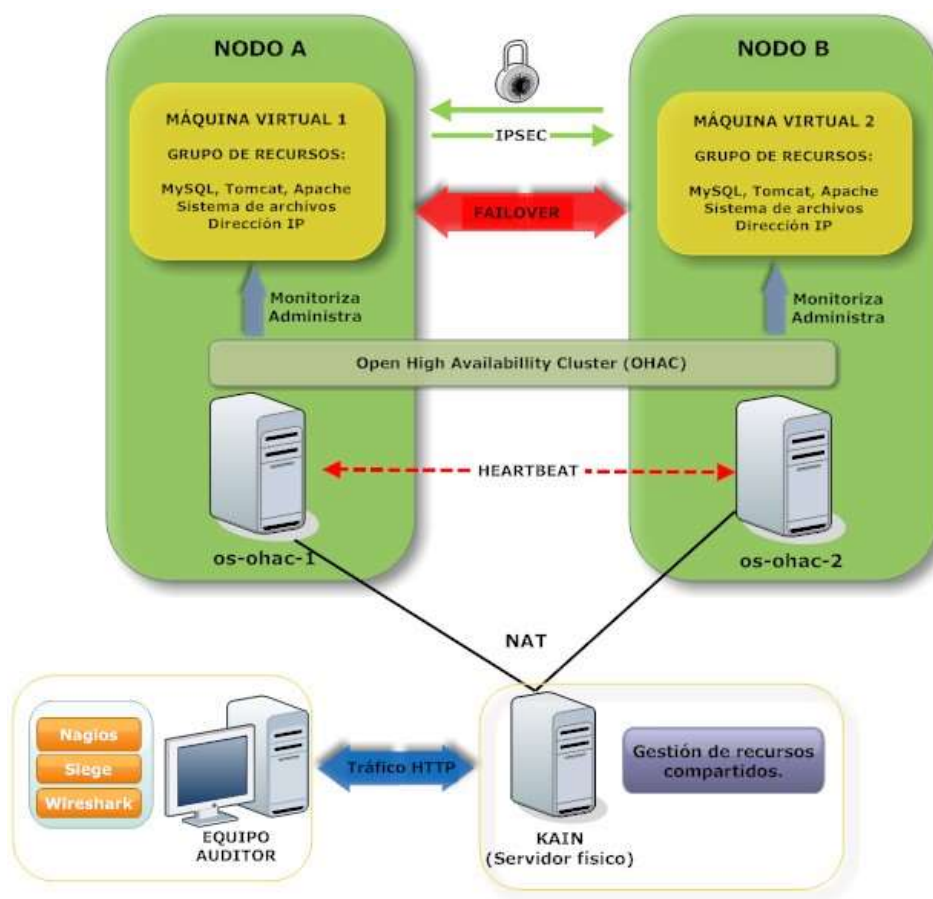


Figura 4.21. Escenario de pruebas.

4.7.2. Pasos 2 y Paso3: Medición de disponibilidad en el escenario de pruebas y Análisis de resultados.

Para la implementación de las pruebas, se usó el equipo que se mencionó en el paso cuatro de la Fase 5; aunque, inicialmente se contó con un servidor de mejores características pero por falta de equipos en el CDU, se tuvo que prescindir de él, puesto que el Centro de Datos de la Universidad del Cauca, no cuenta con equipos redundantes para realizar este tipo de pruebas. A continuación se muestran las pruebas realizadas:

4.7.2.1. Pruebas de Disponibilidad:

a) Primera prueba

En esta primera prueba se simula tráfico generado por 1 usuario, hacia el puerto 8080 durante 10 minutos, como se muestra en la figura 4.23. (ver CD/carpetaAnexo/video 1).

\$siegue -c1 -t10m Kain.unicauca.edu.co:8080. Donde c es la cantidad de usuarios simulados que generan tráfico, t el tiempo en minutos que tarda la generación del tráfico, seguido va el servidor, en este caso Kain, equipo hacia el cual se le está generando tráfico y el puerto que corresponde al tráfico de *Tomcat*. A continuación se muestran las capturas de pantalla referentes a la generación de tráfico con la herramienta *Siege* (ver Figura 4.22) y seguido se muestra la captura de pantalla correspondiente a la ejecución de las herramientas *Nagios* y *Siege*, actuando sobre los dos nodos del clúster que se muestran en consola (ver Figura 4.23).

```
smpantoja@fenix:~$ siege -c1 -t10m kain.unicauca.edu.co:8080
** SIEGE 2.66
** Preparing 1 concurrent users for battle.
The server is now under siege...
HTTP/1.1 200 0.01 secs: 7857 bytes ==> /
HTTP/1.1 200 0.01 secs: 7857 bytes ==> /
HTTP/1.1 200 0.00 secs: 7857 bytes ==> /
HTTP/1.1 200 0.00 secs: 7857 bytes ==> /
HTTP/1.1 200 0.00 secs: 7857 bytes ==> /
```

Figura 4.22. Generación de tráfico con Siege.

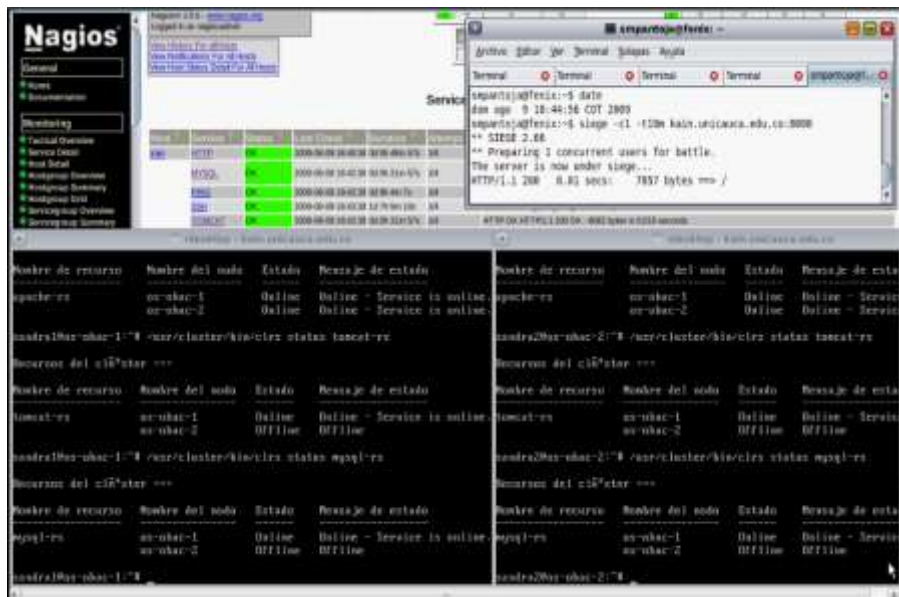


Figura 4.23. Herramientas de prueba actuando sobre los nodos del clúster.

Se observaron los tiempos de caída y recuperación del nodo bajo prueba, de lo cual queda un video generado con los registros respectivos. A las 21:07:33 se cae el nodo 1 y a las 21:07:34 comienza a subir el nodo 2 y a las 21:11:15 sube el nodo 2. Por lo tanto, el tiempo de caída del servicio es igual a: 3 minutos y 42 segundos. Como se mencionó anteriormente, paralelo a esta prueba se ejecutó el sniffer *Wireshark*, el cual permitió visualizar el comportamiento del tráfico durante la caída del nodo (ver Figura 4.24).

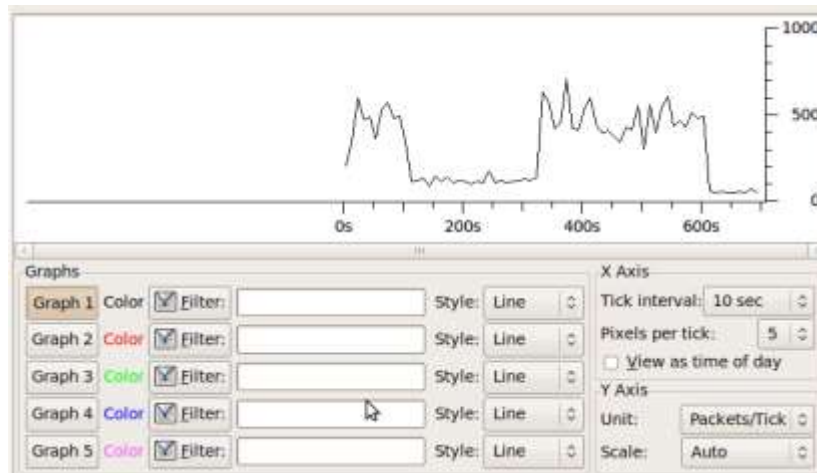


Figura 4.24. Tiempo de interrupción prueba 1.

b) Segunda prueba

Se simula tráfico generado por 10 usuarios, hacia el puerto 8080 durante 10 minutos (ver CD/carpetaAnexo/video 2). Se simula un problema en el nodo 1. A las 22:44:09 se cae el nodo y los servicios y a las 22:47:53, automáticamente sube el nodo 2, generando un tiempo de interrupción= 3 minutos y 44 segundos, como se muestra en la Figura 4.25.

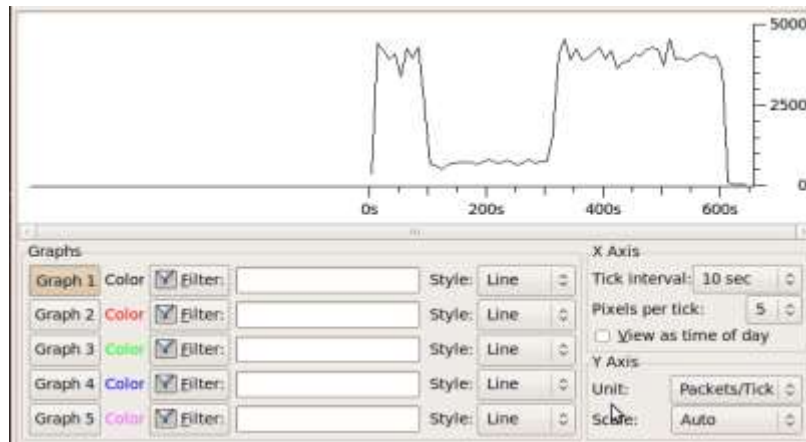


Figura 4.25. Tiempo de interrupción prueba 2.

c) Tercera prueba

Se aumenta el tráfico a 50 usuarios, se simula un problema generado en el nodo 2 y se observa el tiempo de interrupción del servicio mientras sube automáticamente el nodo 1. A las 23:28:41 se cae el nodo 2 y las 23:32:28 subió el nodo 1, Tiempo de interrupción=.3 minutos y 44 segundos (Ver figura 4.26) (ver CD/carpetaAnexo/video 3)

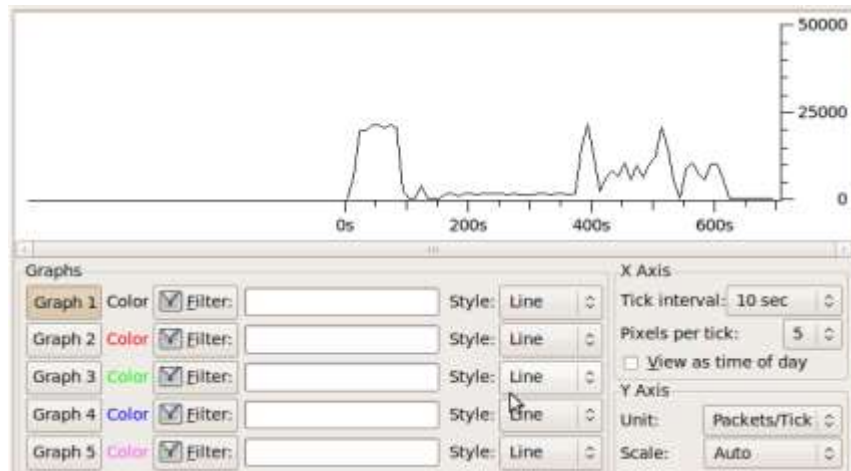


Figura 4.26. Tiempo de interrupcion prueba 3.

d) Cuarta prueba

Se simula un problema en el nodo 1, se genera tráfico *Tomcat* de 100 usuarios, con *Siege*. A las 23:52:06 se cae el nodo1 y a las 23:55:50 subió el nodo2. Tiempo de interrupción= 4 minutos y 9 segundos. (Ver Figura 4.27) (ver CD/carpetaAnexo/video 4).

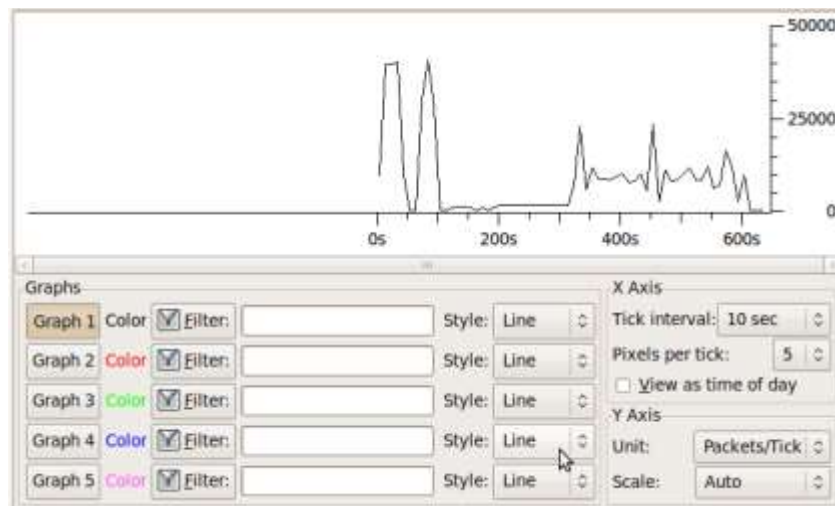


Figura 4.27. Tiempo de interrupcion prueba 4.

Con esta cantidad de usuarios, comenzó a saturarse la conexión y la herramienta *Siege* comienza a mostrar el mensaje “connetion timeout”, por lo que no se pudo simular más conexiones con esta herramienta. A continuación se muestra el resultado de los tiempos de interrupción del servicio (ver Tabla 4.10), cada vez que uno de los nodos presenta un inconveniente, como se puede observar el promedio del tiempo que tardan en migrarse los servicios de un nodo hacia el otro es aproximadamente 4 minutos, tiempo en el cual se

vuelven a levantar automáticamente los servicios, sin necesidad que se haga esta tarea manualmente.

Prueba No.	Número de usuarios	Tiempo de interrupción
Prueba 1	1 usuario	3 minutos y 42 segundos
Prueba 2	10 usuarios	3 minutos y 44 segundos
Prueba 3	50 usuarios	4 minutos y 9 segundos
Prueba 4	100 usuarios	3 minutos y 44 segundos

Tabla 4.11. Resumen de resultados de interrupción del servicio.

4.7.2.2. Pruebas con otras herramientas incluidas en el Sistema de Alta Disponibilidad implementado:

✓ Pruebas de seguridad

En esta prueba se verifica la **confidencialidad** de la información que se envía desde un nodo a otro. Para esto se utiliza la pila de protocolos IPsec, implícita en la instalación de OpenSolaris, lo cual garantiza que los datos entre los nodos se envíen cifrados con el protocolo ESP(Encapsulating Security Payload), este protocolo, utiliza fuertes algoritmos de encriptación como el 3DES(Data Encryption Standard), además, la autenticación se realiza mediante el mecanismo IKEv2(Internet Key Exchange), utilizando avanzadas técnicas para el intercambio seguro de claves.

Inicialmente, antes de la configuración de IPsec, en cada nodo se puede comprobar que la comunicación entre ellos no cuenta con ninguna protección, en la figura 4.28, se utiliza el comando ping desde el nodo1 (os-ohac-1) hasta el nodo2 (os-ohac-2); en la figura 4.29, se utiliza la herramienta *snoop*, un *sniffer* propio de Solaris para capturar los paquetes que se envían o llegan a cierta interfaz de red.

```

os-ohac-1
sandra1@os-ohac-1:~# date
sábado 1 de mayo de 2010 17:34:04 COT
sandra1@os-ohac-1:~# ping -s os-ohac-2
PING os-ohac-2: 56 data bytes
64 bytes from os-ohac-2 (10.0.2.102): icmp_seq=0. time=1.483 ms
64 bytes from os-ohac-2 (10.0.2.102): icmp_seq=1. time=1.706 ms
64 bytes from os-ohac-2 (10.0.2.102): icmp_seq=2. time=0.958 ms
64 bytes from os-ohac-2 (10.0.2.102): icmp_seq=3. time=1.109 ms
^C
----os-ohac-2 PING Statistics----
4 packets transmitted, 4 packets received, 0% packet loss
round-trip (ms)  min/avg/max/stddev = 0.958/1.314/1.706/0.342
sandra1@os-ohac-1:~#

```

Figura 4.28. Ping desde os-ohac-1 a os-ohac-2

```

os-ohac-2
^Csandra2@os-ohac-2:~date
sábado 1 de mayo de 2010 17:34:01 COT
sandra2@os-ohac-2:~# snoop -d e1000g0 host os-ohac-1
Using device e1000g0 (promiscuous mode)
os-ohac-1 -> os-ohac-2 ICMP Echo request (ID: 8537 Sequence number: 0)
os-ohac-2 -> os-ohac-1 ICMP Echo reply (ID: 8537 Sequence number: 0)
os-ohac-1 -> os-ohac-2 ICMP Echo request (ID: 8537 Sequence number: 1)
os-ohac-2 -> os-ohac-1 ICMP Echo reply (ID: 8537 Sequence number: 1)
os-ohac-1 -> os-ohac-2 ICMP Echo request (ID: 8537 Sequence number: 2)
os-ohac-2 -> os-ohac-1 ICMP Echo reply (ID: 8537 Sequence number: 2)
os-ohac-1 -> os-ohac-2 ICMP Echo request (ID: 8537 Sequence number: 3)
os-ohac-2 -> os-ohac-1 ICMP Echo reply (ID: 8537 Sequence number: 3)
sandra2@os-ohac-2:~#

```

Figura 4.29. Captura de los mensajes de ping en os-ohac-2

Como se muestra en la figura anterior (ver Figura 4.29), todos los mensajes ICMP (Internet Control Message Protocol) de petición y de respuesta (ping) son capturados claramente mediante Snoop. Las figuras 4.30 y 4.31, muestran la misma situación, pero esta vez utilizando la herramienta *wget*, con la cual se envía una petición de descarga HTTP desde el nodo1 hasta el nodo2, la cual es capturada con *snoop* en el nodo 2.

```

os-ohac-2
sandra2@os-ohac-2:~# snoop -d e1000g0 host os-ohac-1
Using device e1000g0 (promiscuous mode)
os-ohac-1 -> os-ohac-2 HTTP C port=43303
os-ohac-2 -> os-ohac-1 HTTP R port=43303
os-ohac-1 -> os-ohac-2 HTTP C port=43303
os-ohac-1 -> os-ohac-2 HTTP GET / HTTP/1.0
os-ohac-2 -> os-ohac-1 HTTP R port=43303
os-ohac-2 -> os-ohac-1 HTTP HTTP/1.1 200 OK
os-ohac-2 -> os-ohac-1 HTTP w.unicauca.edu.co/js/validation.js"></script>
os-ohac-2 -> os-ohac-1 HTTP "></
os-ohac-1 -> os-ohac-2 HTTP C port=43303
os-ohac-1 -> os-ohac-2 HTTP C port=43303
os-ohac-2 -> os-ohac-1 HTTP sp; / <a href="http://biblio.unicauca.edu.co" target="_blank">Biblioteca</a
>&nbs
os-ohac-2 -> os-ohac-1 HTTP nviar2" type="button" value="enviar" onClick="javascript:doSubmitLogin('htt
p://w

```

Figura 4.30. Utilización de wget para descargar una página web mediante HTTP

```

os-ohac-1
sandra1@os-ohac-1:~# wget os-ohac-2
--17:43:00-- http://os-ohac-2/
=> 'index.html'
Resolviendo os-ohac-2... 10.0.2.102
Connecting to os-ohac-2|10.0.2.102|:80... conectado.
Petición HTTP enviada, esperando respuesta... 200 OK
Longitud: no especificado [text/html]

[          <<>          ] 29.734      11.20K/s

17:43:16 (1.83 KB/s) - 'index.html' saved [29734]
sandra1@os-ohac-1:~#

```

Figura 4.31. Captura de los mensajes HTTP de descarga de una página web en el nodo2

La figura 4.31, muestra en detalle los mensajes HTTP intercambiados entre los nodos en la descarga de la página web por defecto de os-ohac-2, la cual es una réplica del portal web de Unicauca.

Estas pruebas muestran que no existe confidencialidad en el intercambio de información entre los nodos puesto que un espía podría utilizar una herramienta como *Snoop* para capturar todos los datos enviados entre servidores y de este modo hacerse a datos personales, contraseñas u otro tipo de información confidencial que se envíe sin utilizar seguridad en las aplicaciones.

Después de habilitar IPsec entre las interfaces 10.0.2.101 y 10.0.2.102 de los nodos *os-ohac-1* y *os-ohac-2* respectivamente, se realizaron nuevamente las pruebas mostradas en las figuras 4.28 y 4.30. Las figuras 4.32 y 4.33, muestran las capturas de estas pruebas realizadas con *Snoop* en el nodo2.

```
os-ohac-2
sandra2@os-ohac-2:/etc/inet/secret# date
sábado 1 de mayo de 2010 17:53:05 COT
sandra2@os-ohac-2:/etc/inet/secret# snoop -d e1000g0 host os-ohac-1
Using device e1000g0 (promiscuous mode)
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=19
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=19
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=20
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=21
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=20
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=21
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=22
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=23
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=22
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=23
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=24
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=25
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=26
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=27
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=24
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=25
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=28
```

Figura 4.32. Captura del comando ping cifrado

```
os-ohac-2
^Csandra2@os-ohac-2:/etc/inet/secret# date
sábado 1 de mayo de 2010 17:50:57 COT
sandra2@os-ohac-2:/etc/inet/secret# snoop -d e1000g0 host os-ohac-1
Using device e1000g0 (promiscuous mode)
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=14
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=14
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=15
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=15
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=16
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=16
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=17
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=17
 os-ohac-1 -> os-ohac-2   ESP SPI=0x7b3fbd3 Replay=18
 os-ohac-2 -> os-ohac-1   ESP SPI=0xfc8c5840 Replay=18
sandra2@os-ohac-2:/etc/inet/secret# █
```

Figura 4.33. Captura del comando wget cifrado

Como se aprecia en las figuras anteriores, todos los mensajes que se intercambian entre los nodos son de tipo ESP, es decir se encuentran cifrados con este protocolo de IPsec. Así, todos los mensajes de capa de aplicación como FTP, HTTP, DNS, SQL, etc., serán cifrados en el nodo emisor y enviados al nodo receptor quien los descifra y los envía a la aplicación correspondiente.

✓ Pruebas de la Configuración de respaldo

Los sistemas operativos Solaris y *OpenSolaris* utilizan ZFS como sistema de archivos, una de sus ventajas es la posibilidad de tomar snapshots o clonar un sistema entero de manera sencilla y con una ocupación mínima de espacio en disco, con lo cual se obtiene una copia de seguridad y la capacidad de restauración de todo el sistema en un determinado momento, la herramienta utilizada por *OpenSolaris* para llevar a cabo estas tareas es *beadm*, (Boot Environment Administrator), la cual se encarga de copiar los archivos relevantes del sistema de archivos general y comprimirlos, así como de actualizar el gestor de arranque del sistema (GRUB).

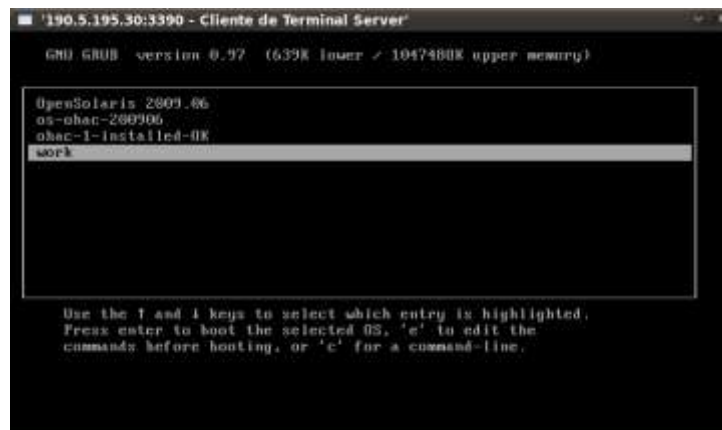
Para crear un nuevo snapshot o copia del ambiente actual es necesario introducir los comandos mostrados en la figura 4.34.



```
os-ohac-1
sandra1@os-ohac-1:~# beadm create work
sandra1@os-ohac-1:~# beadm activate work
sandra1@os-ohac-1:~#
```

Figura 4.34. Creación de una copia de seguridad del sistema actual.

En la gráfica se observa que se crea y se activa el ambiente de inicio denominado “work”, el cual es una copia exacta del sistema actual incluyendo configuración del sistema, aplicaciones y archivos, una vez se reinicia el servidor, el menú de inicio del sistema (Grub) mostrará la nueva opción de inicio como lo muestra la siguiente figura.



```
'190.5.195.30:3390 - Cliente de Terminal Server'
GNU GRUB version 0.97 (639K lower / 1047480K upper memory)

OpenSolaris 2009.06
os-ohac-200906
ohac-1-installed-0K
work

Use the ↑ and ↓ keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

Figura 4.35. Inicio del sistema con diferentes copias de seguridad

Así, la nueva alternativa de inicio “work” es una copia exacta de la actual “os-ohac-1-installed-OK” cuya configuración y archivos se preservan como copia de respaldo en caso de que el ambiente de trabajo sufra un malfuncionamiento por cualquier causa. La imagen muestra también otras copias de respaldo de la configuración del servidor en diferentes estados de instalación, por ejemplo “OpenSolaris 2009.06” es la configuración limpia del sistema operativo, es decir esta copia de seguridad se tomó una vez se instaló el sistema operativo.

Una vez se inicia el sistema en el entorno de trabajo, se puede comprobar los diferentes snapshots del sistema:

```

os-ohac-1
sandra@os-ohac-1:~# beadm list
BE          Active Mountpoint Space  Policy Created
---
ohac-1-installed-OK - -      11.51M static 2010-04-15 14:45
opensolaris - -      5.02M static 2010-03-02 21:50
os-ohac-200906 - -      2.27G static 2010-03-04 21:26
work        NR /      5.96G static 2010-04-19 00:07
sandra@os-ohac-1:~#

```

Figura 4.36. Lista de las copias de seguridad del sistema.

La gráfica anterior muestra además el espacio en disco ocupado por cada copia, se observa que a pesar de ser copias completas de la configuración y archivos del sistema en un momento dado, no ocupan gran espacio en disco duro y es posible volver a cualquiera de ellas simplemente reiniciando la máquina y accediendo a la configuración deseada mediante el menú de arranque.

✓ Pruebas de Rendimiento

En las figuras 4.22 y 4.23, se muestra la ejecución del comando *prstat* en los nodos 1 y 2 respectivamente. Este comando se utiliza para mostrar en tiempo real el consumo de recursos del sistema por las diferentes aplicaciones que se encuentran corriendo en la máquina (Ver Figuras 4.22 y 4.23).

```

os-ohac-1
Archivo Editar Ver Terminal Ayuda
PID USERNAME  SIZE  RSS STATE PRI NICE   TIME    CPU PROCESS/NLWP
15303 root      5388K 3160K cpu0  49  0  0:00:00  0,4% prstat/1
1549  webservd   37M   29M sleep 59  0  0:00:26  0,3% java/17
1030 root      68M   57M sleep 59  0  0:00:30  0,2% java/30
1132 root      62M   16M sleep 59  0  0:00:10  0,1% httpd/1
8 root     14M   10M sleep 59  0  0:00:21  0,1% svc.startd/16
4 root      0K    0K sleep 60  -  0:00:30  0,1% cluster/1148
358 root    8128K 3716K sleep 59  0  0:00:04  0,1% nsd/32
1551 mysql     72M   29M sleep 59  -3  0:00:07  0,1% mysqld/11
788 root     49M  3548K sleep 100 -  0:00:07  0,1% rpc.pmfd/21
10 root     13M   11M sleep 59  0  0:00:22  0,0% svc.configd/25
686 root    2816K 1560K sleep 100 -  0:00:02  0,0% xntpd/1
73 root     23M   13M sleep 59  0  0:00:03  0,0% devfsadm/7
14203 sandra  11M  5764K sleep 59  0  0:00:00  0,0% sshd/1
614 root    6200K 2160K sleep 59  0  0:00:00  0,0% sendmail/1
1 root     2724K 1616K sleep 59  0  0:00:00  0,0% init/1
878 root     15M  5072K sleep 59  0  0:00:01  0,0% cl_eventlogd/5
843 root     18M  4904K sleep 59  0  0:00:01  0,0% cl_eventd/13
1055 root    6844K 3392K sleep 59  0  0:00:00  0,0% gds_probe/1
831 root     21M  5416K sleep 59  0  0:10:30  0,0% scdpmd/13
575 noaccess 2696K 1568K sleep 59  0  0:00:00  0,0% mdnsd/1
626 root     21M   15M sleep 59  0  0:00:02  0,0% fmd/19
Total: 87 processes, 1657 lwps, load averages: 0,11, 0,11, 0,10

```

Figura 4.37. Ejecución de *prstat* en el nodo 1.

PID	USERNAME	SIZE	RSS	STATE	PRI	NICE	TIME	CPU	PROCESS/NLWP
1498	root	5380K	3120K	cpu0	49	0	0:00:00	0,4%	prstat/1
1023	root	68M	57M	sleep	59	0	0:00:25	0,3%	java/24
4	root	0K	0K	sleep	60	-	0:00:00	0,1%	cluster/1142
359	root	8232K	3792K	sleep	59	0	0:00:01	0,0%	nscd/62
181	root	9672K	4544K	sleep	59	0	0:00:00	0,0%	sysventd/19
1194	root	62M	16M	sleep	59	0	0:00:01	0,0%	httpd/1
629	root	20M	13M	sleep	59	0	0:00:01	0,0%	find/19
904	root	15M	3308K	sleep	59	0	0:00:01	0,0%	cl_eventlogd/5
793	root	18M	5004K	sleep	59	0	0:00:01	0,0%	cl_eventd/13
10	root	13M	11M	sleep	59	0	0:00:21	0,0%	svc.configd/25
1426	sandra2	11M	5548K	sleep	59	0	0:00:00	0,0%	sshd/1
844	root	20M	3552K	sleep	100	-	0:00:00	0,0%	rpc.fed/5
812	root	21M	5456K	sleep	59	0	0:08:30	0,0%	scdpmc/13
622	noaccess	2696K	1572K	sleep	59	0	0:00:00	0,0%	mdnsd/1
573	root	3908K	2336K	sleep	59	0	0:00:00	0,0%	hald-addon-acpi/1
608	root	6136K	2160K	sleep	59	0	0:00:00	0,0%	sendmail/1
73	root	24M	13M	sleep	59	0	0:00:01	0,0%	devfsadm/7
249	root	5016K	1664K	sleep	59	0	0:00:00	0,0%	in.mpathd/1
234	root	5556K	2676K	sleep	59	0	0:00:00	0,0%	picld/4
373	root	12M	3540K	sleep	100	-	0:00:00	0,0%	cl_execd/3
713	root	8464K	5352K	sleep	59	0	0:00:00	0,0%	intrad/1

Total: 79 processes, 1629 lwps, load averages: 0,04, 0,04, 0,03

Figura 4.38. Ejecución de prstat en el nodo 2.

Se observa en los recuadros rojos que las aplicaciones más relevantes para estos servidores como java (*Tomcat*), httpd (*Apache*), *MySQL* y clúster, consumen muy pocos recursos en un ambiente normal, incluso el consumo total del sistema es muy bajo, lo cual muestra que este sistema operativo maximiza la eficiencia en el consumo de recursos del sistema.

4.7.2.3. Evaluación de la Disponibilidad

El resultado de este paso es verificar que el porcentaje de disponibilidad alcanzado haya cumplido con los objetivos previstos en las fases de planeación, para lo cual se hace una comparación del prototipo implementado con el funcionamiento real del servidor web, en este caso Ragnarok, en el cual se permitió el acceso para realizar la prueba.

El promedio mínimo en que se caen o bloquean los servicios inesperadamente, es aproximadamente 2 veces en semana, lo cual implica que en un promedio mensual corresponde a 8 veces. Se generó tráfico por medio de *siege* en Ragnarok y se visualizó el tiempo mínimo de interrupción del servicio mientras este se reinicia, el cual es aproximadamente 7 minutos, tal como se muestra en la grafica capturada por *Wireshark* (Ver figura 4.39).



Figura 4.39. Tiempo de interrupción mínima en Ragnarok.

El tiempo mínimo en el que mantiene caído el servicio es de: 6 minutos para que se restablezcan los procesos, sin embargo a este tiempo hay que sumarle los 5 minutos que demora Nagios en refrescar sus datos, lo cual da como resultado 11 minutos de interrupción, esto sin contar el tiempo que se tarda en desplazarse la persona hacia el servidor para reiniciarlo o las veces que no haya una persona disponible en el momento preciso que se caen los servicios u ocurre una falla en el servidor.

Para conocer tiempos aproximados en que puede tardar en subir nuevamente el servicio, se consultaron algunos Logs sobre la caída de los servicios de Ragnarok, como se muestra a continuación, la caída de java en el servidor de Ragnarok, producida el día jueves 25 de marzo de 2010 desde las 07:48:18 am hasta las 08:41:02 am que subió nuevamente el servicio.

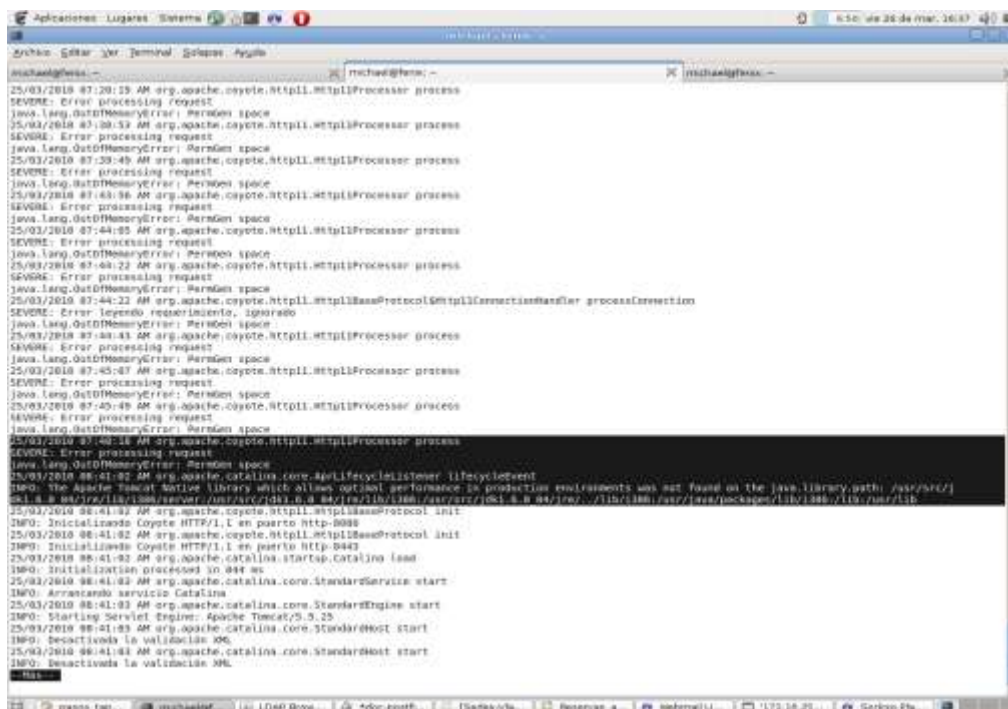


Figura 4.40. Caída de java en Ragnarok.

```

25/03/2010 07:45:49 AM org.apache.coyote.http11.Http11Processor process
SEVERE: Error processing request
java.lang.OutOfMemoryError: PermGen space
25/03/2010 07:48:18 AM org.apache.coyote.http11.Http11Processor process
SEVERE: Error processing request
java.lang.OutOfMemoryError: PermGen space
25/03/2010 08:41:02 AM org.apache.catalina.core.AprLifecycleListener lifecycleEvent
INFO: The Apache Tomcat Native Library which allows optimal performance in production environments was not found on the java.library.path: /usr/src
jdk1.6.0_04/jre/lib/i386/server:/usr/src/jdk1.6.0_04/jre/lib/i386:/usr/src/jdk1.6.0_04/jre/./lib/i386:/usr/java/packages/lib/i386:/lib:/usr/lib
25/03/2010 08:41:02 AM org.apache.coyote.http11.Http11BaseProtocol init
INFO: Inicializando Coyote HTTP/1.1 en puerto http-8080

```

Figura 4.41. Log de caída de Java en Ragnarok.

Como se pudo observar en el anterior log, el tiempo de interrupción del servicio fue de 52 minutos y 8 segundos, lo cual es superior al tiempo mínimo estimado anteriormente, que fue de 11 minutos. Las causas potenciales de caídas del servicio, puede ser la configuración de los servicios o sobrecarga en el procesador, así como también causas inesperadas que han ocurrido o que se pueden presentar, como problemas de sobrecalentamiento de los dispositivos (como cuando hay fallas en el aire acondicionado, o por motivos de sobretensión a causa de los picos de voltaje), fallas en el hardware, pérdida de configuración, y otros factores externos que pueden hacer que el servidor se apague o que los servicios queden bloqueados.

En las pruebas realizadas en el prototipo en el que se instaló el sistema HA, el tiempo de interrupción promedio de los servicios en el prototipo como se observa en la tabla 4.16, es de 4 minutos aproximadamente. Tiempo que corresponde al valor de interrupción, a causa de caída de un nodo del clúster, este sería el tiempo estándar de interrupción del servicio frente a cualquier eventualidad o falla presentada en el servidor, lo cual es una enorme ventaja al no tener la necesidad de que una persona esté pendiente de reiniciar los servicios o el equipo manualmente.

Por lo tanto, si al tener un servicio 24/7 y en el último mes el sistema ha estado expuesto a interrupciones que en promedio son de 8 veces, esto da como resultado un tiempo de interrupción del servicio de $8 \times (4 \text{ minutos}) = 32 \text{ minutos}$.

Cálculo de de la disponibilidad del Servicio en el prototipo:

$$\% \text{Disponibilidad} = ((\text{AST} - \text{DT}) / \text{AST}) * 100$$

AST= $24[\text{horas/día}] * 30[\text{días}] = 720[\text{horas}] = 43200 \text{ minutos} = \text{tiempo acordado de servicio en un mes}$.

DT= 32 minutos, que corresponde al tiempo de caída o interrupción del servicio.

La disponibilidad, calculada con el prototipo arroja los siguientes resultados:

$$\% \text{Disponibilidad} = ((43200 - 32) / 43200) * 100 = 99.925\%$$

Por lo tanto, se demuestra que el porcentaje de disponibilidad de los servicios es muy alto, sobre todo comparado con el registro de disponibilidad capturado por Nagios monitoreando *Tomcat* que es aproximadamente del 83%, como se muestra en la Figura 4.1, de la Fase de Recolección de Estadísticas de Disponibilidad generadas por *Nagios*.

CAPITULO 5

CONCLUSIONES Y TRABAJOS FUTUROS

5.1. CONCLUSIONES

- ✓ Las tecnologías de software libre ofrecen beneficios que normalmente no ofrece el software privativo, por ejemplo no limitar al administrador a depender del soporte, permite explorar a fondo las herramientas para realizar desarrollo o adaptarlas a las necesidades propias del CDU, entre otras estas son algunas de las razones que motivan a escoger la tecnología *HA Open Solaris*.
- ✓ El tiempo de transición obtenido en las pruebas realizadas, donde se simularon distintos tipos de fallas, han demostrado que el sistema de conmutación de los servicios mediante heartbeat y Failover, es una opción estable y eficiente que cumple con las expectativas iniciales de alta disponibilidad, por lo que se recomienda su implementación en la institución.
- ✓ En los procesos de Alta disponibilidad, la información recolectada es fundamental para perfeccionar, definir, dar seguimiento a las actividades, analizar la situación actual del CDU y permitir la toma de decisiones para el mejoramiento continuo, de manera que cada servicio del CDU responda adecuadamente a la revolución tecnológica de las TICs.
- ✓ Clústeres de Alta disponibilidad, Heartbeat, Failover y virtualización, son una combinación necesaria integrar en el desarrollo de soluciones de alta disponibilidad.
- ✓ La aplicación de la guía metodológica al Centro de Datos Unicauca permitió establecer que su aplicación por sí misma no garantiza el logro de una solución de alta disponibilidad en los servicios, pues se requiere principalmente del compromiso firme de la administración CDU y del trabajo conjunto del personal que lo integra.
- ✓ De acuerdo a los resultados obtenidos en caso de estudio, es posible afirmar que seguir los lineamientos que plantea la guía metodológica permiten realizar un proceso ordenado para el plantamiento e implementación de una Solución de Alta Disponibilidad para los servicios de TI de un CDU.
- ✓ Aplicar soluciones de alta disponibilidad a los servicios críticos, es sólo el comienzo de un largo proceso que se debe continuar al resto de aéreas del CDU, puesto que todas son importantes y necesitan de una solución de este tipo.
- ✓ La utilización de la virtualización con *Virtualbox, Xen, Crossbow* y el aprovechamiento de características como realización de copias de respaldo con *beadm*, el manejo *Ipsec* y de tecnologías que permiten manejar protocolos como *iSCSI*, hacen que la solución propuesta reúna con los requisitos de alta disponibilidad necesarios para disminuir los

SPOF de los servicios.

- ✓ Muchas universidades no se han atrevido a implementar soluciones de Alta Disponibilidad debido al desconocimiento del tema y la limitación de los recursos económicos

5.2. TRABAJOS FUTUROS

- ✓ Realizar un estudio comparativo de herramientas de gestión y de monitoreo, que permitan hacer seguimiento a la disponibilidad de los servicios que presta un CDU. Esto con el fin de cuantificar el costo que representa la interrupción o “caída” de un servicio.
- ✓ Según el estándar *TIA/EIA 942*, identificar los aspectos mínimos en infraestructura de TI que deben considerarse en cada subsistema para el diseño de un CDU, que permitan lograr alta disponibilidad para los servicios.
- ✓ Analizar la disponibilidad ofrecida por diferentes herramientas de tecnologías para Clúster de Alta Disponibilidad, en distintos sistemas operativos, por ejemplo *Red Hat*, *Centos*, *Debian*, *OpenSolaris*, *FreeBSD* y *Windows*, y hacer una paralelo de rendimiento respecto al tiempo de respuesta para la restauración de un nodo en cada uno de estos sistemas.

BIBLIOGRAFÍA

- [1] OCG, "Provisión del Servicio," *itsmf.es*, primera edición 2006. [En línea]. Disponible: <http://www.itsmf.es/>. [Consultado: Nov. 2, 2009].
- [2] Y. Medina, D. Rico, "Modelo de Gestión de Servicios para la Universidad de Pamplona: ITIL," *Scientia et Technica*, año XIV, no 39, Septiembre 2008. [En línea]. Disponible: <http://www.utp.edu.co/php/revistas/ScientiaEtTechnica/docsFTP/11732314-319.pdf>. [Consultado: Nov. 2, 2009].
- [3] IT Governance Institute, "Buen Gobierno de las inversiones en TI," *itgi.org*, 2006. [En línea]. Disponible: <http://isaca.org/ContentManagement/ContentDisplay.cfm?ContentID=40792>. [Consultado: Sep. 5, 2009].
- [4] SANCHO, "Disponibilidad," en diccionario de Internet. UIT-T: Base de datos de términos y definiciones de la UIT-T. Disponible: Referencia de UIT-T en línea, <http://www.itu.int/sancho/index-es.asp?lang=es>. [Consultado: Agosto 8, 2009].
- [5] HA Forum, "Providing Open Architecture High Availability Solutions," 2001. [En línea]. Disponible: <http://www.saforum.org>. [Consultado: Sep. 10, 2009].
- [6] Osiatis, "Fundamentos de la Gestión de servicios de TI. Gestión de la disponibilidad, métodos y técnicas," *Osiatis*. [En línea]. Disponible: http://itil.osiatis.es/Curso_ITIL/. [Consultado: Agosto 8, 2009].
- [7] E. Ciurana, "Scalability and high availability," *DZone Refcardz*, version 1 2009. [En línea]. Disponible: http://library.dzone.com/sites/all/files/refcardz/rc043-010d-scalability_3.pdf. [Consultado: Sep. 20, 2009].
- [8] Microsoft Technet, "Descripción de la disponibilidad, la confiabilidad y la escalabilidad," Microsoft 2010. [En línea]. Disponible: <http://technet.microsoft.com/es-es/library/aa996704%28EXCHG.65%29.aspx>. [Consultado: Mayo 15, 2010].
- [9] Software GreenHouse, "Continuidad de Negocios," *swgreenhouse.com*, 2009. [En línea]. Disponible: <http://www.swgreenhouse.com/> [Consultado: Agosto 8, 2009].
- [10] J. Hernández Huerta. "Alta disponibilidad, solución invaluable," *eSemanal*, vol 29, no. 702, Junio 2006. [En línea]. Disponible: <http://www.esemanal.com.mx/>. [Consultado: Agosto. 6, 2009].
- [11] Microsoft Developer Network, "Alta disponibilidad," *Microsoft*, 2009. [En línea]. Disponible: <http://msdn.microsoft.com/es-es/library/bb500217.aspx> [Consultado: Agosto 8, 2009].
- [12] A. M. Pérez, "ISO/IEC 20000 el estándar para la Gestión de Servicios TI,". [En línea]. Disponible: <http://www.uc3m.es/> [Consultado: Agosto 8, 2009].
- [13] IT Governance Institute, "COBIT 4," Information Technology Governance Institute, 2006. [En línea]. Disponible: <http://www.isaca.org>. [Consultado: Agosto 8, 2009].
- [14] NetWorld World España, "Las mejores prácticas para el nuevo centro de datos," *Network World*, 2005. [En línea]. Disponible: <http://www.networkworld.es/>. [Consultado: Agosto 8, 2009].
- [15] ItSMF, "An introductory overview of ITIL V3," *itsmfi.org*, version 1.0, 2007. [En línea]. Disponible: http://www.itsmfi.org/files/itSMF_ITILV3_Intro_Overview_0.pdf. [Consultado: Agosto 8, 2009].

- [16] ISO27000, "Sistema de Gestión de la Seguridad de la Información," *iso27000.com.es*, 2005. [En línea]. Disponible: http://www.iso27000.es/download/doc_sgsi_all.pdf. [Consultado: Agosto 8, 2009].
- [17] Osiatis, "Gestión de la Seguridad, Introducción y Objetivos," *Osiatis*. [En línea]. Disponible: http://itil.osiatis.es/Curso_ITIL/. [Consultado: Agosto 7, 2009]
- [18] CIO, "Mejores prácticas para la gestión de procesos y servicios de Ti," *idg.es/cio*, 2005. [En línea]. Disponible: <http://www.idg.es/cio/estructura/imprimir.asp?id=172583&cat=art>. [Consultado: Sep. 21, 2009].
- [19] Recursos AS/400, "Dossier sobre Alta Disponibilidad, Problemática y Soluciones," [En línea]. Disponible: <http://www.recursos-as400.com/>. [Consultado: Agosto 8, 2009].
- [20] Enrique Vargas, "High Availability Fundamentals," Noviembre de 2000. [En línea]. Disponible: <http://www.sun.com/>. [Consultado: Agosto 8, 2009].
- [21] Departamento de Comunicación de IBM, "Servicios de IBM de continuidad y recuperación de negocio," *ibm.com*, marzo 2009 [En línea]. Disponible: http://www-03.ibm.com/press/es/es/attachment/23443.wss?fileId=ATTACH_FILE1&fileName=Informe_Servicios_Recuperac.pdf. [Consultado: Sep. 22, 2009]
- [22] Software GreenHouse, "Cálculo de Costes de los Tiempos Muertos," *swgreenhouse.com*, 2009. [En línea]. Disponible: <http://www.swgreenhouse.com/>. [Consultado: Agosto 8, 2009].
- [23] W. Pitt Turner, J. H. Seader, K. G. Brill, "TIER Classification define Site infrastructure Performance". [En línea]. Disponible: <http://uptimeinstitute.org/>. [Consultado: agosto 8, 2009].
- [24] AreaData, "Clasificación de la infraestructura de un datacenter- TIA 942," *areadata.com.ar*. 2005. [En línea]. Disponible: <http://www.areadata.com.ar>. [Consultado: Agosto 8, 2009].
- [25] Seguridad CPD, "Niveles de disponibilidad de un centro de procesos de datos (CPD), según el estándar TIA-942," *seguridadcpd.com*, 2005. [En línea]. Disponible: <http://www.seguridadcpd.com/>. [Consultado: Agosto 8, 2009].
- [26] ESTEC, "Conexión para centros de datos Siemon". [En línea]. Disponible: <http://www.estec.cl/descargas/centrodatos.pdf>. [Consultado: Oct. 10, 2009].
- [27] IBM, "Visión de IBM para el Centro de Datos de la Nueva Empresa," *ibm.com*, 2008. [En línea]. Disponible: <http://www-03.ibm.com/>. [Consultado: Agosto 8, 2009].
- [28] A. Obando, "Automatización en el centro de datos," *AdapTive Magazin*, vol 2, p. 24, Julio de 2008. [Online]. Disponible: <http://h20341.www2.hp.com>. [Consultado: Agosto 8, 2009].
- [29] Graphicox, "¿Qué es centro de datos?," *graphicox.com*. [En línea]. Disponible: <http://www.graphicox.com/alojamiento/centrodedatos.html>. [Consultado: Agosto 8, 2009].
- [30] Siemon, "Solución para Centros de Datos Siemon 10G ip," [En línea]. Disponible: <http://www.siemon.com/int/download/brochures/datacenter/datacenter.pdf>. [Consultado: Agosto 8, 2009].
- [31] Anixter Inc, "Data Center Design and Infrastructure Considerations," *anixter.sg/*. Nov 2008. [En línea]. Disponible: <http://www.anixter.sg>. [Consultado: Agosto 8, 2009].
- [32] Oficina de Informática y telecomunicaciones, "A cerca de OITEL", Universidad del Valle, 2008. [En línea]. Disponible: <http://oitel.univalle.edu.co>. [Consultado: Oct. 3, 2009].

- [33] Universidad Nacional de Colombia, "Servicios," 2004. [En línea]. Disponible: <http://www.unal.edu.co> [Consultado: Oct. 3, 2009].
- [34] Universidad del Cauca, "Red de datos,". [En línea]. Disponible: <http://www.unicauca.edu.co/>. [Consultado: Oct. 3, 2009].
- [35] Universidad Michoacana de San Nicolás Hidalgo, "Centro de cómputo universitario", 2009. [En línea]. Disponible: <http://www.umich.mx>. [Consultado: Oct. 5, 2009].
- [36] Anixter Inc, "Data Center Design and Infrastructure Considerations," *anixter.sg/*. Nov 2008. [En línea]. Disponible: <http://www.anixter.sg>. [Consultado: Agosto 8, 2009].
- [37] J. Gumbau, "Hacia la universidad orientada a los servicios: una perspectiva sistémica de cambio permanente por la innovación tecnológica", *Revista de la Universidad y Sociedad del Conocimiento (RUSC)*. vol. 3, n.º 1, Abril 2006. [En línea]. Disponible: <http://www.uoc.edu/rusc/3/1/dt/esp/gumbau.pdf>. [Consultado: Nov. 15, 2009].
- [38] M. Ovilla, "Soporte en hardware para el control de acceso de dispositivos de E/S en un ambiente virtualizado," 16 de Diciembre, 2009.
- [39] E. Marcus, S Hal, "Blueprints for High Availability," 2003. Ed. Willey.
- [40] K. Joel, "Virtual Machine security guidelines, The Center for internet security,". Septiembre 2007.
- [41] Continuidad del negocio basado en virtualización.
- [42] Zorraquino, F.J, "Virtualización: Maquina Virtual". *Astic*, URL, 2006, pp. 68-77
- [43] Red Hat Inc, "Red Hat Cluster Suite Overview," 2008
- [44] Microsoft TechNet, "What Is a Server Cluster?" 2010 [En línea]. Disponible: <http://technet.microsoft.com/en-us/library/cc785197%28WS.10%29.aspx>. [Consultado. Mayo 10, 2010].
- [45] SearchDatacenter, "Definición de cluster," Diccionario en Internet, 2010. [En línea]. Disponible: <http://searchdatacenter.techtarget.com/>. [Consultado: Mayo 2, 2010].
- [46] S. W. Peter, "Clusters for High Availability," Prentice Hall 2001
- [47] Answers, "Definición de failover", Diccionario en Internet, 2010. [En línea]. Disponible: <http://www.answers.com>. [Consultado: Mayo 2, 2010].
- [48] Novell, "The SUSE Linux Enterprise Server Heartbeat Guide,".2008
- [49] Computer Security Institute, "Sistem Hardening, " Computer Security Institute 2004
- [50] A. Reelsen, J. Fernández, S. Peña, "Securing Debian Manual," 2007. [En línea]. Disponible: <http://www.debian.org/doc/manuals/securing-debian-howto/>. [Consultado: Mayo 10, 2010].
- [51] B. Cromwell, "TCP/IP Stack Hardening," 2008.
- [52] K. Fenzi, "Linux Security HOWTO," 2004. [En línea]. Disponible: <http://www.tldp.org/HOWTO/Security-HOWTO/>. [Consultado: Mayo 2, 2010].
- [53] R. Semko, "Hardening Solaris," NEbraskaCERT Conference. 2005
- [54] C. Guevara, F. Mera, "Criterios Para Establecer Políticas De Seguridad De La Información Y Plan De contingencia, Caso de Estudio El Centro de Datos de La Universidad del Cauca," Trabajo de Grado, Popayán, Colombia, feb. 2008.
- [55] Microsoft TechNet, "Server Cluster overview," 2010. [En línea]. Disponible: [http://technet.microsoft.com/en-us/library/cc759183\(WS.10\).aspx](http://technet.microsoft.com/en-us/library/cc759183(WS.10).aspx). [Consultado: Febrero 14, 2010].
- [56] N. Michael, Colin Spence, "Chapter 3: Planning Redundancy and Scaling the SharePoint Environment," *NetWorkWorld*, Agosto 2, 2007, pág. 2. [En línea]. Disponible: <http://www.networkworld.com/subnets/microsoft/080207-ch3-sharepoint-unleashed.html?page=1>. [Consultado: Enero 24, 2010].
- [57] Server and Storage Consolidation with iSCSI Arrays. David Dale, Net App. 2010.

- [58] IP Storage Protocols: Iscsi . Jhon L. Hufferd, consultant. 2010.
- [59] iSCSI SAN's: Ideal Applications, Large and Small. Jasson Brasil, Net App., Gary Gumanow, Dell. 2010.
- [60] Oracle Corporation, "Learn OpenSolaris," 2010. [En línea]. Disponible: www.opensolaris.com. [Consultado: Enero 14, 2010].
- [61] Sun Microsystems Inc, "Introduction to IPS," 2009. [En línea]. Disponible: <http://dlc.sun.com/osol/docs/content/2009.06/IMGPACKAGESYS/>. Consultado. Nov. 13, 2009].
- [62] Oracle Corporation, "OpenSolaris Zones," 2009.[En línea]. Disponible. <http://hub.opensolaris.org/bin/view/Community+Group+zones/WebHome>. [Consultado: Nov. 13, 2009].
- [63] R. Ellard, "Requirements Specification-Colorado Phase I: Infrastructure," OpenSolaris. 2008. [En línea]. Disponible: <http://hub.opensolaris.org/bin/download/Project+colorado/Requirements/colorado-haci.pdf>. [Consultado: Nov. 12, 2009].
- [64] [5].Sun Microsystem, "SystemAdministration Guide:Basic Administration, " págs. 317-395. [En línea].Disponible: <http://dlc.sun.com/pdf/819-2379/819-2379.pdf>. [Consultado: Nov. 11, 2009].
- [65] IETF, "RFC3720: Internet Small Computer Systems Interface (iSCSI)," 2004. [En línea]. Disponible. www.ietf.org. [Consultado: Nov. 13, 2009].
- [66] Oracle, "Introduction to COMSTAR", [En línea] Disponible: <http://hub.opensolaris.org/bin/view/Project+comstar/WebHome>. [Consultado: Nov. 12, 2009].
- [67] Sun Microsystems, "IPSec in the Solaris9 Operating Environment," sun.com. 2002. [En línea]. Disponible: <http://www.sun.com/software/whitepapers/solaris9/ipsec.pdf>. [Consultado: Nov. 10, 2009].
- [68] Sun Microsystem, "OpenSolaris para todos". 2010. [En línea]. Disponible: <http://hub.opensolaris.org>. [Consultado: Nov. 12 2009]
- [69] Sun Microsystem, "TCP-Wrappers on Solaris 10". 2005. [En línea]. Disponible: <http://learningsolaris.com/archives/2005/03/16/tcp-wrappers-on-solaris-10/>. [Consultado: Nov. 11, 2009].