



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

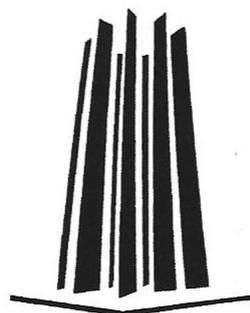
**FACULTAD DE ESTUDIOS SUPERIORES**

**CAMPUS ARAGÓN**

**“ACTUALIZACIÓN TECNOLÓGICA Y OPTIMIZACIÓN  
DE LOS PROCESOS OPERATIVOS PARA LA  
PRESTACIÓN DE SERVICIOS DE VALOR  
AGREGADO”**

**T E S I S**  
PARA OBTENER EL TÍTULO DE:  
INGENIERO EN COMPUTACIÓN  
P R E S E N T A:  
RAQUEL ROMÁN GODÍNEZ

**ASESOR:  
ING. JUAN GASTALDI PÉREZ**



**JUNIO 2008**

**ARAGÓN, MÉXICO D.F.**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*Muchas Gracias:*

*A mí querido Dios, por todas las bendiciones que me ha dado, muy en especial por su protección y guía en mi vida. Sólo a tu lado Señor, se encuentra la sabiduría, el amor y la paz eterna.*

*A mi Papá y Mamá por todo su apoyo y sacrificios para que no sólo mi vida, sino la de mis hermanos, sean vidas sanas y productivas. Ustedes son mi ejemplo de amor, fortaleza, honestidad y respeto a seguir.*

*A mis hermanos Saúl, Katy y René, por estar siempre presentes en mi vida y ser mi aliciente en la conclusión de este trabajo.*

*A mi querida sobrina Dany, por ser una motivación y deseo de superación constante. Gracias pequeña por tu hermosa alegría.*

*A mis queridos amigos, que sin mencionarlos saben que están presentes en mi mente, su amistad es una bendición para mí.*

*A mis compañeros de trabajo por hacer de cada día una maravillosa experiencia en mi vida. Junto a ustedes he aprendido mucho en todos los aspectos de mi vida.*

*A mi asesor Ing. Juan Gastaldi y sinodales de tesis por todo el apoyo que me brindaron en el desarrollo y conclusión de este trabajo. Muy en especial a Oscar Estrada por ser una gran persona y un amigo incondicional.*

*Con sincero agradecimiento, respeto y cariño,*

*Raquel.*

Actualización tecnológica y optimización de  
los procesos operativos para la prestación  
de servicios de valor agregado.

---

# Índice

---

Índice .....	2
Introducción .....	9
<b>Capítulo 1. INTRODUCCIÓN A LOS SERVICIOS DE VALOR AGREGADO .....</b>	<b>13</b>
<b>1.1 Generaciones de telefonía móvil. ....</b>	<b>14</b>
<b>1.1.1 Primera Generación: Sistemas Analógicos.....</b>	<b>14</b>
1.1.1.1 AMPS (Advanced Mobile Phone System) Sistema de Telefonía Móvil Avanzada. ....	14
1.1.1.2 NMT Nordic Mobile Telephones.....	14
1.1.1.3 TACS Total Access Communications System.....	15
<b>1.1.2 Segunda Generación: Sistemas Digitales. ....</b>	<b>16</b>
1.1.2.1 GSM Global System for Mobile Communications.....	16
1.1.2.2 TDMA Time Division Multiple Access.....	17
1.1.2.3 CDMA.....	18
1.1.2.4 PDC Personal Digital Cellular.....	19
1.1.2.5 Servicios de 2G:.....	20
<b>1.1.3 La evolución de la 2G a la 3G: Generación 2.5 y 2.75.....</b>	<b>21</b>
1.1.3.1 Wireless Application Protocol.....	22
1.1.3.2 General Packet Radio Service.....	23
1.1.3.3 EDGE .....	24
<b>1.1.4 Tercera Generación: Sistemas Digitales. ....</b>	<b>25</b>
1.1.4.1 Universal Mobile Telecommunications System.....	25
1.1.4.2 Servicios de 3G.....	26
<b>1.2 Servicios de valor agregado. ....</b>	<b>28</b>
<b>1.2.1 Definición.....</b>	<b>28</b>
<b>1.2.2 Tipos de mercado en telefonía móvil.....</b>	<b>29</b>

1.2.2.1	Mercado masivo.....	29
1.2.2.2	Mercado corporativo.....	29
1.2.3	Tipos de servicios en telefonía celular.....	29
1.2.3.1	Servicio masivo.....	29
1.2.3.2	Servicio corporativo.....	29
1.2.4	Servicios de valor agregado para mercado masivo.....	30
1.2.4.1	Roaming Internacional.....	30
1.2.4.2	Servicios basados en SMS.....	30
1.2.4.3	Servicios basados en transporte de datos.....	31
1.2.4.4	Servicio de Mensajería Multimedia "MMS".....	34
1.2.5	Servicios de valor agregado para mercado corporativo.....	35
1.2.5.1	Oficina móvil.....	35
1.2.5.2	Acceso dedicado.....	35
1.2.5.3	Servicios basados en localización.....	38
1.2.5.4	Servicios corporativos basados en mensajería.....	38
<b>Capítulo 2. DEFINICIÓN DE CONCEPTOS EMPLEADOS EN ARQUITECTURAS DE SERVIDORES</b> .....		<b>40</b>
2.1	Definición de arquitectura.....	41
2.2	Arquitectura cliente / servidor.....	41
2.2.1	Ventajas de la arquitectura cliente-servidor.....	42
2.3	Servidor de aplicación.....	42
2.3.1	Definición.....	42
2.3.2	Características comunes.....	43
2.3.3	Uso.....	43
2.4	J2EE: Java 2 Enterprise Edition.....	44
2.4.1	Servidores de aplicación J2EE.....	44
2.5	Definición de alta disponibilidad.....	45
2.5.1	Cluster de alta disponibilidad.....	45
2.5.1.1	Componentes del Cluster.....	48
2.6	Arquitectura de redes.....	48
2.6.1	Modelo de referencia OSI.....	48
2.6.2	Capas del modelo OSI.....	50
2.6.2.1	Capa física.....	51
2.6.2.2	Capa de enlace de datos.....	51
2.6.2.3	Capa de red.....	52
2.6.2.4	Capa de transporte.....	52
2.6.2.5	Capa de sesión.....	53
2.6.2.6	Capa de presentación.....	53
2.6.2.7	Capa de aplicación.....	53

2.6.3	Ejemplos de protocolos.....	54
2.6.4	Topologías de red.....	55
2.6.4.1	Topología jerárquica.....	55
2.6.4.2	Topología horizontal (bus).....	58
2.6.4.3	Topología en estrella.....	57
2.6.4.4	Topología en anillo.....	58
2.6.4.5	Topología en malla.....	60
2.7	Red de área de almacenamiento.....	61
2.7.1	Características de una SAN.....	62
2.8	Balancedores de carga.....	64
2.8.1	Definición.....	64
2.8.2	Mejoras que aporta el balanceador de carga.....	64
2.8.2.1	Escalabilidad.....	65
2.8.2.2	Disponibilidad.....	65
2.8.2.3	Gestión y flexibilidad en el mantenimiento.....	65
2.8.2.4	Seguridad.....	68
2.8.2.5	Calidad de servicio (QoS).....	68
2.8.3	Tipos de balancedores de carga.....	66
2.8.3.1	Server Load Balancing.....	68
2.8.3.2	Firewall Load Balancing.....	67
2.8.3.3	Cache Load Balancing.....	67
2.8.3.4	Global Server Load Balancing.....	67
2.8.4	Métodos de distribución de carga.....	67
2.8.4.1	Stateless.....	68
2.8.4.2	Stateful.....	69
2.8.4.3	Algoritmos de Scheduling (Hashing).....	70
2.9	Network Address Translation (NAT).....	72
2.9.1	Destination NAT.....	72
2.9.2	Source NAT.....	72
<b>Capítulo 3. ANÁLISIS DE LA SITUACIÓN TECNOLÓGICA Y EJECUCIÓN DE PROCESOS OPERATIVOS INICIAL EN LA PRESTACIÓN DE SERVICIOS DE VALOR AGREGADO.....</b>		<b>74</b>
3.1	Análisis de la situación tecnológica inicial.....	75
3.1.1	Diseño inicial de alta disponibilidad a nivel HW y SW.....	75
3.1.2	Configuración Activo & Pasivo.....	77
3.1.2.1	Ventajas:.....	77
3.1.2.2	Desventajas:.....	77
3.1.3	Application Server de dominio público.....	78
3.1.4	Configuración de Base de Datos.....	80
3.1.5	Procesos aplicativos no distribuidos.....	80

3.1.6	Monitoreo del sistema inicial.....	81
3.1.7	Sistema de respaldos inicial.....	81
3.1.8	Disponibilidad inicial del sistema SVA.....	82
3.2	Ejecución inicial de procesos operativos en los sistemas SVA.....	83
<b>Capítulo 4. DISEÑO E IMPLEMENTACIÓN DE LA ACTUALIZACIÓN TECNOLÓGICA.....</b>		<b>87</b>
4.1	Primera Fase. Cambio de web application server en la arquitectura inicial.....	89
4.1.1	Evaluación de un nuevo web application server.....	89
4.1.1.1	Escalabilidad.....	90
4.1.1.2	Performance.....	90
4.1.1.3	Compatibilidad con diferentes sistemas operativos.....	91
4.1.1.4	Balanceo de carga.....	91
4.1.1.5	Integración con dispositivos móviles.....	91
4.1.1.6	Enterprise application integration (EAI) y J2EE.....	91
4.1.1.7	Seguridad.....	92
4.1.1.8	Supervisión de Cambios.....	93
4.1.1.9	Consola de Administración.....	93
4.1.1.10	Herramientas de Diagnóstico.....	93
4.1.1.11	Log de Eventos.....	94
4.1.1.12	Soporte.....	94
4.1.1.13	Capacitación.....	95
4.1.1.14	Costo.....	95
4.1.2	Resultados obtenidos de la evaluación.....	95
4.1.3	Cambio de Apache Tomcat a BEA Weblogic Server.....	98
4.2	Segunda Fase: Diseño e implementación de la nueva arquitectura.....	99
4.2.1	Diseño propuesto de la nueva arquitectura en dos capas: back-end y front-end.....	101
4.2.2	Diseño propuesto de alta disponibilidad “activo-activo” en los servidores back-end.....	102
4.2.3	Diseño propuesto de alta disponibilidad “activo-activo” en servidores front-end.....	103
4.2.4	Diseño propuesto para la configuración del Dominio de BEA Weblogic.....	105
4.2.4.1	Configuración de maquinas.....	105
4.2.4.2	Función del nodo_admin.....	106
4.2.4.3	Configuración de servidores administrados o instancias.....	107
4.2.4.4	Configuración de node managers.....	107
4.2.4.5	Configuración de clusters.....	108
4.2.4.6	Acceso a las bases de datos desde el dominio de Weblogic Server.....	109

4.2.5	Diseño de direccionamiento IP utilizando balanceadores de carga.	111
4.2.5.1	Direccionamiento IP asignado a la arquitectura inicial.....	112
4.2.5.2	Propuesta de direccionamiento IP en la nueva arquitectura.	113
4.2.6	Diseño de políticas en balanceadores de carga.	114
4.2.6.1	Pool de Servidores:.....	114
4.2.6.2	Servidores Virtuales:.....	115
4.2.6.3	NAT (Network Address Translation):.....	115
4.2.7	Diseño de la solución de monitoreo para la nueva arquitectura.	116
4.2.7.1	Características de HP Open View Operations:.....	116
4.2.7.2	Requerimientos de monitoreo para la nueva arquitectura. ....	117
4.2.7.3	Agentes propuestos para el monitoreo a nivel Hardware y Sistema Operativo. ....	117
4.2.7.4	Agentes propuestos para el monitoreo de las aplicaciones comerciales. ....	118
4.2.7.5	Solución de monitoreo para aplicaciones y desarrollos internos "in-house". ....	123
4.2.7.6	Diagrama de solución final de monitoreo HP OVO.....	123
4.2.8	Diseño de la solución de respaldo de información para la nueva arquitectura.	125
4.2.8.1	Requerimientos para la nueva arquitectura.	125
4.2.8.2	Propuesta de políticas de respaldo	126
4.3	Plan de Trabajo. Actividades realizadas en la fase de implementación.	127
<b>Capítulo 5. MEJORAS EN EL PROCESO DE OPERACIÓN</b>		<b>130</b>
5.1	Introducción a los sistemas de calidad.	132
5.1.1	Características de los sistemas.	132
5.1.2	La percepción de calidad por parte del cliente.	132
5.1.3	Definición de sistema de gestión de la calidad.	133
5.1.4	Razones de aplicación de los sistemas de calidad.	134
5.1.5	La organización y los sistemas.	134
5.1.5.1	¿Qué es una organización?	134
5.1.5.2	Las organizaciones como un sistema.	134
5.1.5.3	La organización como un conjunto de procesos.	135
5.1.5.4	La organización y la necesidad de cambio.	135
5.1.6	La organización como una red de procesos.	136
5.1.6.1	¿Qué es un proceso?	136
5.1.6.2	Características de los procesos.	136
5.1.6.3	¿Qué es una red de procesos?	137
5.1.6.4	Documentación de los procesos.	137
5.1.7	La gestión de la calidad.	138

---

5.1.7.1	¿Qué es gestión?	138
5.1.7.2	El enfoque de gestión dentro de la organización.	138
<b>5.1.8</b>	<b>Normas ISO 9000:2000</b>	<b>139</b>
5.1.8.1	Definición de ISO	139
5.1.8.2	ISO, ¿Para qué?	140
5.1.8.3	Antecedentes de las normas ISO 9000.	140
5.1.8.4	Familia de normas ISO 9000:2000.	141
5.1.8.5	¿En dónde se aplican las normas ISO?	142
5.1.8.6	¿Que se busca con la familia de normas ISO 9000:2000 en la organización?	142
<b>5.1.9</b>	<b>Los 8 principios de la gestión de la calidad</b>	<b>143</b>
5.1.9.1	Enfoque al cliente.	143
5.1.9.2	Liderazgo.	143
5.1.9.3	Involucramiento del personal.	143
5.1.9.4	Enfoque basado en procesos.	144
5.1.9.5	Enfoque de sistemas para la gestión.	144
5.1.9.6	Mejora continua.	144
5.1.9.7	Enfoque basado en hechos para la toma de decisión.	144
5.1.9.8	Relaciones mutuamente beneficiosas con los proveedores.	144
<b>5.1.10</b>	<b>¿Qué es la Famosa Certificación ISO 9000?</b>	<b>145</b>
5.1.10.1	¿Cuáles son los beneficios de la certificación?	145
<b>5.2</b>	<b>Propuesta de mejora en los procesos operativos para prestación de servicios de valor agregado.</b>	<b>146</b>
5.2.1	Organización propuesta en el área operativa.	148
5.2.2	Proceso de alta, baja o modificación en los sistemas de SVA en producción.	150
5.2.3	Proceso de integración de un nuevo sistema SVA.	151
<b>Capítulo 6.</b>	<b>ANÁLISIS DE RESULTADOS</b>	<b>152</b>
<b>6.1</b>	<b>Resultados obtenidos con la actualización tecnológica.</b>	<b>153</b>
6.1.1	Situación inicial.	153
6.1.2	Resultados obtenidos después de la implementación de la primera fase.	154
6.1.3	Resultados obtenidos después de la implementación de la segunda fase.	156
6.1.3.1	Resultados de la implementación de alta disponibilidad en configuración "activo & activo", distribución de tráfico y de procesos aplicativos.	156
6.1.3.2	Resultados de la implementación del monitoreo centralizado.	160
6.1.3.3	Resultados de la implementación del sistema de respaldos automático.	161

6.1.3.4	Fácil de escalar y transparente para el usuario final.....	161
6.2	Resultados obtenidos con la optimización de procesos operativos. ....	163
6.3	Logros alcanzados en la disponibilidad del sistema SVA. ....	165
Conclusiones .....		168
Glosario .....		170
Bibliografía .....		187

# Introducción

---

En la última década, el uso de teléfonos móviles ha aumentando de manera considerable debido a la gran aceptación que han tenido en la sociedad de la información. Estamos por tanto, llegando a un punto en donde la dependencia de la tecnología de comunicación móvil se hace más evidente. Actualmente el área de las comunicaciones móviles, junto con Internet, es la de crecimiento más rápido dentro del sector de las telecomunicaciones a nivel mundial.

En México, más del 60% de la población cuenta con un teléfono móvil, y con ello la oportunidad de acceso a servicios de alta tecnología, como la red de 3G/UMTS, esto es: enlaces móviles de banda ancha de hasta 1.5Mbps, servicio de video llamada, televisión en vivo, video streaming, servicios de información y entretenimiento, entre otros.

Dado lo anterior, y considerando que hasta el primer trimestre del 2007, según el INEGI<sup>1</sup>, el 22.1% de los hogares mexicanos disponía de una computadora, y el 12.3% de estos tenía una conexión a Internet, la tendencia de adquisición tecnológica en nuestro país “teléfono móvil versus computadora personal” indica que el teléfono móvil seguirá siendo el de mayor uso y se convertirá en el medio preferido de acceso a la red de Internet.

Esta tendencia amplía las oportunidades de negocio para las empresas de comunicaciones móviles, pero a su vez, les exige elevar su nivel de calidad de servicio y actualizar constantemente su infraestructura tecnológica.

En este contexto, la empresa de la que formo parte, esta comprometida en trabajar día a día en pro de elevar la calidad de todos los servicios que ofrece a través del trabajo en equipo y dentro de un proceso de mejora continua. Con esta filosofía de trabajo y considerando el incremento en la demanda de los servicios de valor agregado masivos y corporativos que se ha tenido, surgió el proyecto **Actualización Tecnológica** en una de las arquitecturas a mi cargo que brinda servicios de información y entretenimiento a nivel nacional, y que por razones de seguridad, en lo sucesivo la llamaremos *Sistema SVA*.

Adicionalmente, este proyecto contempló en paralelo, la **Optimización de los Procesos Operativos** involucrados, con el fin de facilitar las tareas operativas a realizar en los sistemas de valor agregado.

Por lo anterior, el trabajo presentado en esta tesis en modalidad de un caso práctico, abordó dos objetivos principales:

---

<sup>1</sup> INEGI. Instituto Nacional de Estadística, Geografía e Informática.

1. Contar con una solución a nivel hardware/software que permita **alcanzar y mantener la disponibilidad semanal del sistema SVA al 99.9%**,
2. Optimizar los procesos operativos a aplicar en los sistemas de valor agregado, dado al cambio y ofrecimiento constante de nuevos servicios y aplicaciones.

Tomando en cuenta los objetivos antes mencionados, el presente trabajo se compone de seis capítulos que a continuación describo:

Para iniciar, el Capítulo 1, “Introducción a los servicios de valor agregado”, tiene la finalidad de hacer una breve descripción histórica de las generaciones de telefonía móvil, exponer el concepto de servicio de valor agregado que se utiliza en las operadoras, y describir de manera general algunos de los servicios disponibles actualmente.

El Capítulo 2, “Definición de conceptos empleados en arquitecturas de servidores”, contiene la definición de conceptos empleados en arquitecturas de servidores para entender los términos utilizados en los capítulos siguientes.

En el Capítulo 3, “Análisis de la situación tecnológica y ejecución de procesos operativos inicial en la prestación de servicios de valor agregado”, se presenta el análisis inicial de operación del *sistema SVA* en cuestión, es decir, la configuración inicial de alta disponibilidad, base de datos, procesos aplicativos, sistema de monitoreo y respaldo; así mismo, en este capítulo se describe la ejecución inicial de los procesos operativos que se tenía para este tipo de sistemas.

El Capítulo 4, “Diseño e implementación de la actualización tecnológica”, describe a detalle las decisiones y actividades realizadas para el logro del primer objetivo antes citado.

El Capítulo 5, “Mejoras en el proceso de operación”, expone la propuesta de mejora en los procesos operativos para la prestación de servicios de valor agregado y así impactar en el logro del segundo objetivo definido. Sin embargo, antes de llegar a esta propuesta, se brinda una introducción a los sistemas de calidad, con el fin de entender la filosofía de trabajo que aplicamos en la empresa.

El Capítulo 6, “Análisis de resultados”, muestra los resultados obtenidos una vez implementado el diseño propuesto para la nueva arquitectura y puesto en práctica las mejoras realizadas en los procesos operativos para la prestación de servicios de valor agregado. En este capítulo podrá apreciar los impactos obtenidos en la disponibilidad del *sistema SVA* en cuestión.

Por último, el presente trabajo cuenta con un amplio glosario de términos que espero sean de utilidad para buscar rápidamente cualquier referencia o acrónimo que se pueda encontrar en el mismo. También incluyo una lista de referencias bibliográficas con literatura recomendada a través de la cual usted puede adquirir mayor información y profundizar en alguno de los temas.

# 1 Capítulo 1

---

Introducción a los servicios de valor  
agregado.

## **1.1 Generaciones de telefonía móvil.**

Antes de llegar a definir el concepto de servicios de valor agregado es muy conveniente hacer un breve repaso por las distintas tecnologías que se han venido empleando para implantar el servicio de telefonía móvil y por consecuencia los servicios de valor agregado que existen actualmente.

### **1.1.1 Primera Generación: Sistemas Analógicos.**

En Estados Unidos el primer sistema analógico utilizado fue el AMPS (Advanced Mobile Phone System), mientras que en los países europeos fueron dos los sistemas ampliamente utilizados para la prestación del servicio de telefonía móvil: NMT (Nordic Movil Telecommunications) y TACS (Total Access Communications System).

A continuación se describen de manera muy general estos sistemas:

#### **1.1.1.1 AMPS (Advanced Mobile Phone System) Sistema de Telefonía Móvil Avanzada.**

Es un estándar móvil analógico desarrollado por los Laboratorios Bell en la década de los 70's, ampliamente utilizado en toda América, así como en la costa asiática del Pacífico y Este de Europa. Opera en la banda de frecuencias de 800Mhz y, en su versión digital TDMA, en 800 y 1900MHz.

#### **1.1.1.2 NMT Nordic Mobile Telephones.**

NMT es un servicio de comunicaciones móviles analógico, desarrollado por Ericsson y Nokia, opera en las bandas de 450 y 900 Mhz.

El sistema NMT 450 surge como un servicio normalizado en los países escandinavos (Suecia-Noruega-Dinamarca-Islandia) en el año 1981; éste es un sistema ideal para cubrir una gran extensión de terreno con poca inversión. Otra versión es la denominada NMT 900 que utiliza la banda de frecuencias de 900MHz, permitiendo de esta forma un mayor número de canales, utilizada, inicialmente en aquellos puntos en que el NMT 450 se encontraba saturado.

### 1.1.1.3 TACS Total Access Communications System.

TACS es un sistema de comunicaciones móviles analógico, similar al AMPS, desarrollado por Motorola. Se lanzó comercialmente en el Reino Unido en el año de 1985. Opera en la banda de 900MHz y ha estado ampliamente extendido en Europa. Mediante este sistema se obtiene una mejor calidad del servicio, al mismo tiempo que mejora la relación señal/ruido, al tener una mayor anchura del canal.

La primera generación o sistemas analógicos como AMPS, NMT y TACS, fueron concebidos para transmisión de voz únicamente.

Estos sistemas utilizan el método de Acceso Múltiple por División en Frecuencia (FDMA), una técnica que asigna a cada canal una frecuencia específica, por lo que el cambio de canal implica el cambio de frecuencia. Cada llamada utiliza una banda de frecuencias para transmitir y otra para recibir.

Las frecuencias de utilización de estos sistemas se muestran a continuación:

Sistema	Enlace ascendente estación Móvil a BTS	Enlace descendente BTS a estación Móvil
NMT 450	453 – 457,5 MHz	463 – 467,5 MHz
NMT 900	890 – 915 MHz	935 – 960 MHz
AMPS	869 – 894 MHz	824 – 849 MHz
TACS	935 – 950/960 MHz	890 – 905/915 MHz

Tabla 1.1. Bandas de frecuencia utilizadas por los principales sistemas analógicos.

### **1.1.2 Segunda Generación: Sistemas Digitales.**

Las primeras generaciones de sistemas móviles, tales como NMT, TACS, AMPS, etc., han sido sustituidas por los sistemas digitales como son el GSM en Europa y varios países del mundo, entre ellos México; TDMA en Estados Unidos, CDMA en América del Norte y algunos países de Latinoamérica y PDC en Japón.

Los sistemas digitales, de 2ª Generación, utilizan el método de Acceso Múltiple por División en Tiempo (TDMA), una técnica que divide cada canal en 8 intervalos de tiempo, que juntos forman una trama TDMA, por lo que permiten la transmisión simultánea, pero discontinua, en la misma frecuencia portadora, de ráfagas o paquetes de información. Cada llamada utiliza uno de estos intervalos de tiempo (time slot) con una banda de frecuencias para transmitir y otra para recibir (conexión bidireccional), y cada receptor selecciona la ráfaga con su número e ignora las otras. De esta manera se puede aumentar la capacidad de tráfico, una de las ventajas de utilizar la transmisión digital.

A continuación se describen de manera muy general los sistemas GSM, TDMA, CDMA y PDC considerados como tecnologías de segunda generación:

#### **1.1.2.1 GSM Global System for Mobile Communications.**

GSM o Sistema Global para Comunicaciones Móviles, es un estándar mundial para teléfonos móviles digitales. Es el estándar predominante en Europa, así como el mayoritario en el resto del mundo.

El Sistema Global para Comunicaciones Móviles, tuvo sus inicios en 1982 en el seno de la CEPT Conference Européenne des Postes et Telecommunications en donde se definió el grupo de trabajo denominado GSM (Groupe Special Mobile) encargado de crear las especificaciones para este sistema, el cual, debía soportar la itinerancia entre países, con posibilidad de evolución para ir incorporando nuevas tecnologías, servicios y aplicaciones.

La especificación de la Fase I del GSM concluyó en 1991 con los servicios de voz y las primeras redes se desplegaron inmediatamente. A finales de 1997, el servicio GSM estaba disponible en más de 100 países.

El sistema GSM se planteó como un sistema multioperador. El estándar fue diseñado con la posibilidad de que varios operadores pudieran compartir el espectro. Emplea una interfaz aire TDMA. La tasa de velocidad del estándar es de 9.6Kbps.

Las bandas de frecuencias más comunes son: 900MHz, 1800 MHz y 1900MHz.

- ◆ GSM 900 (transmisión en la banda de frecuencia de 900 MHz) es la red digital predominante de Europa. También se utiliza en la región del Pacífico asiático.
- ◆ GSM 1800 (transmisión en la banda de frecuencia de 1800 MHz) también se utiliza en Europa y Asia, pero su uso no se ha extendido tanto como el de GSM 900.
- ◆ GSM 1900 (transmisión en la banda de frecuencia de 1900 MHz) es el sistema GSM utilizado principalmente en América y Canadá.

Consideradas la tecnología celular digital más avanzada, las redes GSM son líderes en muchos servicios típicamente "digitales", incluido el servicio de mensajes cortos (SMS), la configuración Over the air (OTA) y la localización GSM. Gracias a su tecnología y presencia tanto en el continente Americano como en el resto del mundo, GSM ocupa una buena posición en el roaming global. Muchos teléfonos GSM se denominan "teléfonos globales" porque se pueden utilizar en cualquier país. La tarjeta SIM ("Subscriber Identification Module") es un componente esencial y exclusivo de los teléfonos GSM.

### **1.1.2.2 TDMA Time Division Multiple Access.**

TDMA ofrece servicio inalámbrico digital utilizando la multiplexación por división en el tiempo (TDM). Una frecuencia de radio está dividida en intervalos de tiempo, y luego los intervalos se asignan a múltiples llamadas. De esta forma, una única frecuencia puede admitir varios canales de datos simultáneos.

Además de ser una de las tecnologías celulares digitales más antiguas, TDMA también se considera la tecnología digital menos avanzada, en parte debido a su falta de flexibilidad en comparación con las demás.

TDMA se uso principalmente en Estados Unidos, aunque también se emplea en América Latina, Nueva Zelanda, partes de Rusia y la región del Pacífico Asiático.

### 1.1.2.3 CDMA.

Abreviatura de Code-Division Multiple Access, se trata de una tecnología celular digital que utiliza técnicas de amplio espectro. A diferencia de sistemas competidores que utilizan TDMA, como GSM, CDMA no asigna una frecuencia específica a cada usuario. En lugar de ello, cada canal utiliza el espectro completo que haya disponible.

El CDMA original se conoce como cdmaOne bajo una marca registrada de Qualcomm. A CDMA se le caracteriza por su alta capacidad y celdas de radio pequeño, que emplea espectro extendido y un esquema de codificación especial y, lo mejor de todo es muy eficiente en potencia

Las conversaciones individuales se codifican con una secuencia digital pseudo-aleatorizada. Muchas conversaciones se llevan a cabo simultáneamente enviando todas las comunicaciones en grupos de bits mezclados a la vez, y etiquetando cada grupo que pertenece a una comunicación específica con un código diferente. Por tanto, cada comunicación se puede recomponer en el orden correcto en el otro extremo, utilizando los códigos únicos vinculados a determinados grupos de bits.

Las redes basadas en CDMA están construidas con protocolos basados en IP (Internet Protocol). El estándar cdmaOne ya incorpora en sus terminales los protocolos TCP/IP (Protocolo de control de transmisión/Protocolo de Internet) y PPP (Protocolo punto a punto).

CDMA es la tecnología celular digital más habitual en América del Norte.

Países que cuentan con CDMA:

- ◆ Estados Unidos Sprint Nextel
- ◆ México Iusacell, Unefon
- ◆ Argentina
- ◆ Brasil
- ◆ Venezuela Movilnet, Movistar

La segunda generación o sistemas digitales como GSM, TDMA, CDMA y PDC, son sistemas orientados a soportar, además de la voz, la transmisión de datos a bajas velocidades.

#### **1.1.2.4 PDC Personal Digital Cellular.**

PDC es uno de los tres principales estándares inalámbricos digitales del mundo, junto con GSM y TDMA. Actualmente solo se utiliza en Japón, con los primeros sistemas introducidos por NTT DoCoMo en 1991 para reemplazar a las primitivas redes analógicas. Opera en las bandas de los 800MHz y los 1500MHz, haciendo un uso eficiente del ancho de banda y se basa en la tecnología TDMA.

Con la demanda tan alta de ancho de banda en Japón, el sistema puede operar en dos modos, half-rate y full-rate. Los canales de half-rate reducen la calidad de la voz y las tasas de transmisión de datos, pero proporcionan más canales para ocupar el mismo ancho de banda.

PDC opera dividiendo cada canal en varias ranuras de tiempo y por tanto permitiendo a varios usuarios utilizar a la vez un canal de la misma frecuencia. Cada canal puede soportar 3 usuarios en condiciones normales. Es decir, puede trabajar con 6 canales half-rate o 3 canales full-rate. La tasa de velocidad del estándar es de 9.6Kbps del estilo de otras tecnologías similares como el GSM, pero cuando la codificación es de forma half-rate cae hasta los 5.6Kbps. Aunque esto da una reducción considerable en la calidad de la voz, todavía es adecuado para mantener la inteligibilidad, es decir, que se entienda la voz.

La red PDC soporta muchas más funciones avanzadas "online" que las otras tecnologías móviles de Segunda Generación 2G con su mensajería de texto SMS y la identificación de llamada. Utilizando sus capacidades de Red Inteligente, el PDC también soporta llamadas en prepago, números personales, Números de Acceso Universal, planes de carga avanzados y Redes Privadas Virtuales inalámbricas, VPN's. Las VPN's son grupos cerrados de usuarios que proporciona a los "colegas" trabajar en lugares diferentes para comunicarse a través de la red de telefonía móvil como si estuvieran usando un sistema convencional de telefonía en una oficina.

En Japón la cobertura en interiores es altamente importante, proporcionando una importante diferencia en el servicio para las otras redes. El PDC ha sido diseñado para permitir soluciones de mejora de la congestión en lugares como centros comerciales, oficinas y estaciones de metro. Una red de micro-estaciones y pico-estaciones base puede ser desplegada en interiores, con sistemas de antenas distribuidos y repetidores, sobre todos los edificios que tengan en su planificación el estándar PDC.

Para la transmisión de datos se introdujo el PDC-P (PDC Mobile Packet Data Communication System). Utiliza un sistema basado en la transmisión de paquetes permitiendo a varios usuarios utilizar un canal simple a la vez. La transmisión de datos por conmutación de paquetes es muy conveniente porque el usuario está permanentemente "online" y sólo paga por el volumen de datos transmitido. Al mejorar la eficiencia de la red, el PDC-P da una tasa de transferencia de 28.8Kbps.

#### **1.1.2.5 Servicios de 2G:**

Entre los servicios de valor agregado que destacan en esta segunda generación se encuentran:

##### **Roaming Internacional.**

Los sistemas digitales ofrecen roaming "internacional" para el desplazamiento entre países, portando el mismo terminal. La terminal se registra automáticamente en la siguiente red digital al cambiar de un país a otro, quedando inmediatamente disponible para su uso; sin embargo, es necesario que el operador seleccionado, tenga acuerdos de roaming con operadores de otros países.

##### **SMS ( Short Messaging Service).**

El servicio de mensajes cortos (SMS) es el servicio inalámbrico aceptado globalmente que posibilita el envío y recepción de mensajes de texto hacia y desde teléfonos móviles y a otros sistemas externos tales como buscapersonas, buzones de voz y diversos dispositivos portátiles, donde el texto puede contener palabras, números o una combinación alfanumérica de caracteres. El SMS fue creado para incorporarlo a GSM

Se les llama mensajes cortos porque que tienen entre 100 Y 250 caracteres de largo (en telefonía celular GSM el límite es de 160 caracteres, 70 si se utiliza otro alfabeto distinto del latino, como el chino o el árabe) alfanuméricos, incluidos espacios en blanco, mientras que en CDMA llega hasta 256, generalmente en forma de frases cortas y significativas, disponiendo algunos terminales de la capacidad de "texto predictivo", que simplifica la composición del mensaje. Pero también el uso combinado de caracteres tales como " \* , >, + " Y otros puede ser utilizado para crear mensajes gráficos.

### **1.1.3 La evolución de la 2G a la 3G: Generación 2.5 y 2.75.**

Como se ha mencionado, los primeros sistemas telefónicos móviles analógicos se consideran de 1ª Generación; los sistemas digitales como GSM, son los llamados de 2a Generación; y los próximos, que ya están empezando a aparecer, también digitales, son los denominados de 3a Generación, entre los que el UMTS es su estándar principal, y están destinados a sustituir a todos los anteriores.

Si bien el paso de la 1G a la 2G supuso un cambio en la red y en los terminales, debido a la transición del mundo analógico al digital, se mantuvo el mismo tipo de negocio, que básicamente era el de la voz; por el contrario, el paso de la 2G a la 3G, aunque ambas siguen siendo tecnologías digitales, supone un cambio radical en los elementos de red, en los terminales y en el modelo de negocio. Esto último, debido a que adicional al servicio de voz, cobran una especial relevancia los contenidos de otro tipo, como las imágenes, la música y aplicaciones basadas en el perfil del usuario, sus preferencias y su localización, sin dejar de lado las mayores velocidades que permite y que hacen realidad las aplicaciones multimedia, en donde el tiempo de descarga es un factor crítico.

En la evolución de las redes móviles, como es GSM, se han introducido tecnologías tales como WAP (acceso a contenidos de Internet adaptados para el móvil), GPRS (mayor velocidad de acceso y conmutación de paquetes), EDGE, entre otras, que hacen necesaria la adaptación de las redes actuales con la introducción de nuevos elementos de hardware y software, algunos de ellos aprovechables, no así las terminales, que son nuevas y específicas para la nueva tecnología.

A continuación se describen de manera general las tecnologías WAP, GPRS y EDGE, utilizadas en la prestación de servicios de valor agregado:

### 1.1.3.1 Wireless Application Protocol.

WAP Wireless Application Protocol o Protocolo de Aplicaciones Inalámbricas es un estándar abierto internacional para aplicaciones que utilizan las comunicaciones inalámbricas, por ejemplo: acceso a servicios de Internet desde un teléfono móvil.

Se trata de la especificación de un entorno de aplicación y de un conjunto de protocolos de comunicaciones para normalizar el modo en que los dispositivos inalámbricos, se pueden utilizar para acceder a correo electrónico, grupo de noticias y otros.

El organismo que se encarga de desarrollar el estándar WAP fue originalmente el WAP Forum, fundado por cuatro empresas del sector de las comunicaciones móviles, Sony-Ericsson, Nokia, Motorola y Openwave (originalmente Unwired Planet). Desde 2002 el WAP Forum es parte de la Open Mobile Alliance (OMA), consorcio que se ocupa de la definición de diversas normas relacionadas con las comunicaciones móviles, entre ellas las normas WAP.

WAP define un entorno de aplicación y una pila de protocolos para aplicaciones y servicios accesibles a través de terminales móviles. Consiste en un conjunto de especificaciones, definidas por la Open Mobile Alliance / WAP Forum, que permiten que los desarrolladores diseñen aplicaciones de interconexión para terminales móviles, típicamente teléfonos.

La tecnología WAP permite que los usuarios de estos dispositivos puedan acceder a servicios disponibles en Internet. Sin embargo, existen algunas consideraciones a tener en cuenta al diseñar estos servicios para usuarios móviles, fundamentalmente debidas a las características de los terminales: pantalla significativamente más pequeña que la de un ordenador personal, teclados más limitados que los de un ordenador, limitaciones en la memoria disponible, tanto memoria RAM como memoria para almacenamiento persistente, y limitaciones en la capacidad del procesador, en comparación con la memoria y procesador de un ordenador personal típico. Las redes de telefonía móvil ofrecen también unas prestaciones por lo general menores que los accesos a Internet; sin embargo, con las redes de tercera generación como UMTS las prestaciones mejoran de manera importante.

### **1.1.3.2 General Packet Radio Service.**

GPRS es una tecnología digital de telefonía móvil. Es considerada la generación 2.5, entre la segunda generación (GSM) y la tercera (UMTS). Proporciona altas velocidades de transferencia de datos (especialmente útil para conectar a Internet) y se utiliza en las redes GSM.

GPRS es sólo una modificación de la forma de transmitir datos en una red GSM, pasando de la conmutación de circuitos en GSM (donde el circuito está permanentemente reservado mientras dure la comunicación aunque no se envíe información en un momento dado) a la conmutación de paquetes.

Desde el punto de vista del Operador de Telefonía Móvil es una forma sencilla de migrar la red desde GSM a una red UMTS puesto que las antenas (la parte más cara de una red de Telecomunicaciones móviles) sufren sólo ligeros cambios y los elementos nuevos de red necesarios para GPRS serán compartidos en el futuro con la red UMTS.

GPRS es básicamente una comunicación basada en paquetes de datos. Los timeslots (intervalos de tiempo) se asignan en GSM generalmente mediante una conexión conmutada, pero en GPRS los intervalos de tiempo se asignan a la conexión de paquetes, mediante un sistema basado en la demanda. Esto significa que si no se envía ningún dato por el usuario, las frecuencias quedan libres para ser utilizadas por otros usuarios.

Que la conmutación sea por paquetes permite fundamentalmente la compartición de los recursos radio. Un usuario GPRS sólo usará la red cuando envíe o reciba un paquete de información, todo el tiempo que esté inactivo podrá ser utilizado por otros usuarios para enviar y recibir información. Esto permite a los operadores dotar de más de un canal de comunicación sin miedo a saturar la red, de forma que mientras que en GSM sólo se ocupa un canal de recepción de datos del terminal a la red y otro canal de transmisión de datos desde la red al terminal, en GPRS es posible tener terminales que gestionen cuatro canales simultáneos de recepción y dos de transmisión, pasando de velocidades de 9,6 kbps en GSM a 40 kbps en recepción en GPRS y 20 kbps de transmisión.

Otra ventaja de la conmutación de paquetes es que, al ocuparse los recursos sólo cuando se transmite o recibe información, la tarificación por parte del operador de telefonía móvil sólo se produce por la información transitada, no por el tiempo de conexión. Esto hace posible el desarrollo de aplicaciones en donde un dispositivo móvil se conecta a la red y permanece conectado durante un periodo prolongado de tiempo sin que esto afecte en gran medida a la cantidad facturada por el operador.

Los teléfonos GPRS pueden llevar un puerto bluetooth, IrDA, o conexión por cable para transferir datos al ordenador, cámaras digitales, móviles u otros dispositivos.

### **1.1.3.3 EDGE .**

Enhanced Data rates for GSM of Evolution o Tasas de Datos Mejoradas para la evolución de GSM. También conocida como EGPRS (Enhanced GPRS), es una tecnología de la telefonía móvil celular, que actúa como puente entre las redes 2G y 3G.

EDGE se considera una evolución del GPRS (General Packet Radio Service). Esta tecnología funciona con redes TDMA y su mejora, GSM. Aunque EDGE funciona con cualquier GSM que tenga implementado GPRS, el operador debe implementar las actualizaciones necesarias, además no todos los teléfonos móviles soportan esta tecnología.

EDGE, o EGPRS, puede ser usado en cualquier transferencia de datos basada en conmutación por paquetes (Packet Switched), como la conexión a Internet. Los beneficios de EDGE sobre GPRS se pueden ver en las aplicaciones que requieren una velocidad de transferencia de datos, o ancho de banda alta, como video y otros servicios multimedia.

EDGE puede alcanzar una velocidad de transmisión de hasta 384 Kbps. Para la implementación de EDGE por parte de un operador, la red principal, o core network, no necesita ser modificada, sin embargo, las estaciones bases, BTS, sí deben serlo. Se deben instalar trancceptores compatibles con EDGE, además de nuevas terminales (teléfonos) y un software que pueda decodificar/codificar los nuevos esquemas de modulación.

#### **1.1.4 Tercera Generación: Sistemas Digitales.**

La tercera generación de comunicaciones móviles, llamada 3G, significa un salto enorme respecto de los sistemas actuales. Está pensada para itinerancia global, transmisión de datos a alta velocidad a través de técnicas avanzadas de conmutación de circuitos y de paquetes, soporta tecnología IP, lo que posibilita el acceso a Internet y, en general, aplicaciones multimedia móviles, con servicios personalizados y basados en la localización de los usuarios.

Básicamente, en Europa, cuando se habla de la 3G nos referimos a UMTS, siglas que en inglés hace referencia a los Servicios Universales de Telecomunicaciones Móviles. UMTS, es miembro de la familia global IMT-2000<sup>1</sup> (International Mobile Telecommunications-2000) del sistema de comunicaciones móviles de 3G de la UIT-Unión Internacional de Telecomunicaciones.

UMTS tendrá un papel protagonista en la creación del futuro mercado masivo para las comunicaciones multimedia inalámbricas de alta calidad y velocidad, que alcanzarán a 2.000 millones de usuarios en todo el mundo en el año 2010, donde la denominada 4a Generación empezará a surgir con fuerza.

En los últimos diez años, UMTS ha sido objeto de intensos esfuerzos de investigación y desarrollo en todo el mundo, y cuenta con el apoyo de numerosos e importantes fabricantes y operadores de telecomunicaciones, esto representa una oportunidad única de crear un mercado masivo para el acceso a la Sociedad de la información de servicios móviles altamente personalizados y de uso fácil.

A continuación se describe de manera general el sistema UMTS considerado como tecnología de tercera generación:

##### **1.1.4.1 Universal Mobile Telecommunications System.**

Universal Mobile Telecommunications System (UMTS) es una de las tecnologías usadas por los móviles de tercera generación (3G).

---

<sup>1</sup> IMT-2000 (International Mobile Telecommunications-2000) es una guía que marca los puntos en común que deben cumplirse para conseguir el objetivo de la itinerancia global, es decir, que un terminal de usuario de 3G pueda comunicarse con cualquier red 3G del mundo.

La principal ventaja de UMTS sobre la segunda generación móvil (2G), es la capacidad de soportar altas velocidades de transmisión de datos de hasta 144 kbit/s sobre vehículos a gran velocidad, 384 kbit/s en espacios abiertos de extrarradios y 2 Mbit/s con baja movilidad (interior de edificios). Esta capacidad sumada al soporte inherente del Protocolo de Internet (IP), se combinan poderosamente para prestar servicios multimedia interactivos y nuevas aplicaciones de banda ancha, tales como servicios de video telefonía y video conferencia.

#### **1.1.4.2 Servicios de 3G.**

Los servicios de tercera generación, 3G, combinan el acceso móvil de alta velocidad con los servicios basados en el protocolo IP; ofreciendo una conexión rápida con la WWW (World Wide Web), nuevas formas de comunicarse, de acceder a la información, de hacer negocios, de aprender y de disfrutar del tiempo libre, dejando a un lado las conexiones lentas, los grandes terminales y los puntos de acceso fijos. Con la 3G se pueden realizar múltiples conexiones simultáneamente desde un mismo terminal móvil; así, por ejemplo, un usuario podría conectarse a una base de datos remota para obtener información sin necesidad de interrumpir una sesión de videoconferencia.

Para que los usuarios utilicen los servicios 3G harán falta nuevos teléfonos y otros dispositivos capaces de proporcionarles los servicios que deseen, desde los de telefonía móvil hasta los de multimedia (voz, datos y vídeo). Además, en las redes móviles hay que introducir nuevos sistemas de transmisión por radio, cambiar parte de las plataformas de conmutación y de transmisión e incorporar los nodos de servicio que hagan posibles las prestaciones 3G. El paso de la 2G a la 3G significará, además del cambio de terminal, el cambio de negocio, en donde las aplicaciones personalizadas, de localización y multimedia serán las protagonistas, tanto para los ciudadanos residenciales como para los empresariales, viendo los operadores cómo se incrementa el tráfico en sus redes y se crean nuevas oportunidades de negocio.

Para asegurar el éxito de los servicios de 3G, se ha de proporcionar a los usuarios unas comunicaciones muy eficientes, con una alta velocidad y calidad, fáciles de utilizar. Los sistemas de 3G deben aportar:

- ◆ Transmisión simétrica/asimétrica de alta fiabilidad.
- ◆ Hasta 384 kbit/s en espacios abiertos y 2Mbit/s con baja movilidad.
- ◆ Uso de ancho de banda dinámico, en función de la aplicación.
- ◆ Soporte tanto en conmutación de paquetes como de circuitos.
- ◆ Acceso a Internet (navegación WWW), videojuegos, comercio electrónico y vídeo y audio en tiempo real.
- ◆ Diferentes servicios simultáneos en una sola conexión.
- ◆ Calidad de voz como en la red fija.
- ◆ Mayor capacidad y uso eficiente del espectro.
- ◆ Personalización de los servicios, según perfil de usuario.
- ◆ Servicios dependientes de la posición.
- ◆ Incorporación gradual en coexistencia con los sistemas actuales de 2G.
- ◆ Roaming, incluido el internacional, entre diferentes operadores.
- ◆ Un estándar global y abierto que cubra las necesidades de un mercado de masas.
- ◆ Cobertura mundial, con servicios terrestres y por satélite.

La tercera generación, donde GSM, TDMA y CDMA, evolucionan con GPRS y EDGE, hacia UMTS y CDMA2000. La preocupación por su desarrollo tecnológico estriba en universalizar los servicios junto a las redes que los soportan, estandarizar el terminal de usuario integrado en una sola unidad multifuncional, portátil –de bajo peso y batería de larga duración), optimizar la cobertura de grandes áreas geográficas y atender demandas específicas.

## **1.2 Servicios de valor agregado.**

### **1.2.1 Definición.**

Los servicios de valor agregado son aquellos que utilizan como soporte servicios básicos, telemáticos, de difusión o cualquier combinación de éstos, y con ellos proporcionan la capacidad completa para el envío o intercambio de información, agregando otras facilidades al servicio base o satisfaciendo nuevas necesidades específicas de telecomunicaciones.

Forman parte de estos servicios, entre otros, el acceso, envío, tratamiento, depósito y recuperación de información almacenada, la transferencia electrónica de fondos, el videotexto, el teletexto y el correo electrónico. Sólo se considerarán servicios de valor agregado aquellos que se puedan diferenciar de los servicios básicos.

Para que el servicio de valor agregado se diferencie del servicio básico, es necesario que el usuario perciba de manera directa alguna facilidad agregada a la simple telecomunicación, que le proporcione beneficios de telecomunicaciones adicionales, independientemente de la tecnología o la terminal utilizada; o que el operador de servicios de valor agregado efectúe procesos lógicos sobre la información que posibiliten una mejora, adición o cambio al contenido de la información de manera tal que genere un cambio neto de la misma independientemente del terminal utilizado. Este cambio a su vez, debe generar un beneficio inmediato y directo, que debe ser percibido por el usuario del servicio.

Entre los factores de éxito de los servicios de valor agregado se encuentran la sencillez o facilidad de uso, su disponibilidad y una política de precios aceptable para el consumidor, que, en definitiva, es quien tiene que decidir si los usa, en base al valor que le aporten.

Por lo anterior, un servicio de valor agregado en el ámbito de telefonía móvil, es aquel que se proporciona adicional al servicio básico de voz. Por ejemplo: servicio de buzón de voz, servicio de mensajería "sms", acceso al correo electrónico desde el móvil, etc.

En las siguientes secciones se describirá la funcionalidad y características de algunos servicios de valor agregado que existen actualmente y los cuales son ofrecidos por algunas operadoras nacionales e internacionales.

## **1.2.2 Tipos de mercado en telefonía móvil.**

Para fines del presente trabajo definiremos dos tipos de mercado potencial a través de los cuales una operadora de telefonía celular puede comercializar los servicios de valor agregado:

### **1.2.2.1 Mercado masivo.**

Son todos los clientes que al momento de adquirir a una operadora de telefonía celular el servicio básico de voz, se convierten en clientes activos potenciales para la compra de un servicio de valor agregado, es decir, de un servicio adicional. Ejemplo: servicio de envío y recepción de mensajes cortos.

### **1.2.2.2 Mercado corporativo.**

El mercado corporativo lo conforman todas aquellas empresas que desarrollan sus propias soluciones y/o requieren una solución a sus necesidades específicas de negocio para lo cual hacen uso del transporte de datos y/o servicios que una operadora de telefonía celular ofrece a través de su infraestructura.

## **1.2.3 Tipos de servicios en telefonía celular.**

### **1.2.3.1 Servicio masivo.**

Un servicio masivo es todo aquel servicio de valor agregado que se puede ofrecer al mercado masivo.

### **1.2.3.2 Servicio corporativo.**

Un servicio corporativo es todo aquel servicio de valor agregado que está orientado a cubrir las necesidades específicas de negocio de una empresa.

## **1.2.4 Servicios de valor agregado para mercado masivo.**

### **1.2.4.1 Roaming Internacional.**

El servicio de roaming ha hecho posible que los usuarios de telefonía móvil adquieran una completa libertad de movimiento entre las áreas de cobertura de las diferentes operadoras.

Este servicio consiste en permitir que un usuario que se encuentre en zona de cobertura de una red móvil diferente a la que le presta el servicio pueda recibir las llamadas hechas hacia su número de móvil, sin necesidad de realizar ningún tipo de procedimiento extra, y en muchos casos también permitirle efectuar llamadas hacia la zona donde se contrató originalmente el servicio sin necesidad de hacer una marcación especial. Para alcanzar este fin, ambas compañías, la prestadora original del servicio y la propietaria de la red en la que el cliente esté itinerando, deben tener suscritos un acuerdo de itinerancia, en el que definen qué clientes tienen acuerdo al servicio y cómo se efectuará la conexión entre sus sistemas para enrutar las llamadas.

### **1.2.4.2 Servicios basados en SMS.**

#### **Mensajes escritos.**

El servicio de mensajes cortos o SMS (Short Message Service) permite al usuario enviar y recibir mensajes de texto de su teléfono hacia otro u otros teléfonos celulares.



### Servicios de información.

Este tipo de servicio permite a los usuarios recibir oportunamente toda la información de su interés en cualquier parte del país, a través de mensajes cortos. Entre los servicios de información que actualmente ofrecen las operadoras se encuentran:

- ◆ Noticias nacionales e internacionales.
- ◆ Deportes
- ◆ Información sobre el estado del clima
- ◆ Información de finanzas
- ◆ Notas de espectáculos
- ◆ Horóscopos
- ◆ Sorteos
- ◆ Poemas y versos.
- ◆ Notas de Salud y Dietas
- ◆ Cartelera de cine
- ◆ Lugares de esparcimiento y diversión.
- ◆ Trivial y concursos.
- ◆ Entre otros...



La solicitud es a través de un SMS-MO<sup>2</sup> en donde se indica la palabra clave del contenido de información y la marcación correspondiente del servicio. Una vez enviado el mensaje SMS-MO, la recepción del contenido es a través de un SMS-MT<sup>3</sup>.

#### 1.2.4.3 Servicios basados en transporte de datos.

##### Descarga de contenido vía un portal WAP.

Este servicio permite al usuario comprar contenido como imágenes, sonidos, juegos, videos, temas musicales desde su terminal. Así mismo se puede acceder a una gran variedad de sitios con información de interés general como noticias, sitios turísticos, horóscopos, cine cartelera, deportes, bancos, etc.

---

<sup>2</sup> SMS-MO. Mensaje corto originado en el móvil.

<sup>3</sup> SMS-MT. Mensaje corto terminado en el móvil.

Para hacer uso de este tipo de servicios se requiere que el usuario cuente con:

- ◆ Una terminal que soporte servicios de 2G y/o 3G.
- ◆ Adquirir con su operadora correspondiente algún servicio transporte tal como: GPRS, HSCSD, EDGE, UMTS, etc.

El usuario debe iniciar una sesión de WAP para conectarse al portal wap de su elección y poder descargar el contenido y consultar la información que requiera en el momento que lo desee.



A continuación se muestran los tipos de servicio que se pueden visualizar desde un portal WAP:

### **Noticias**

Estar siempre informado es de gran importancia en todos los ámbitos. Por lo que a través de este medio se pueden consultar las últimas noticias nacionales, internacionales, económicas, culturales, deportivas, etc.



### Música

A través de este medio se puede descargar una gran variedad de contenido para personalizar el celular con tonos, sonidos, canciones, imágenes, fondo de pantalla, etc., además de descargar el karaoke o video del artista favorito.



### Juegos

Otro tipo de contenido que esta disponible vía WAP es la descarga de juegos para divertirse con los amigos y/o poner en práctica las habilidades mentales.



### Información de entretenimiento

También se puede consultar información diaria y actualizada de cada uno de los signos del zodiaco, la compatibilidad entre signos, información de Feng Sui, loterías, etc.



### Turismo

También se puede acceder a un sin fin de información, como información turística de lugares para visitar en vacaciones o paseos que se pueden realizar en un día o en un fin de semana.



### Cartelera de cine

A través de este medio se puede revisar desde el móvil la cartelera de los cines, reseñas de películas, recomendaciones para la semana, dirección y sala en donde se exhibe la película.



#### 1.2.4.4 Servicio de Mensajería Multimedia “MMS”.

**MMS** Multimedia Messaging System es la evolución del servicio de SMS que permite transmitir mensajes que contengan texto, gráficos, imágenes, sonido e incluso vídeo entre cualquier dispositivo móvil utilizando las altas velocidades de transmisión que permiten GPRS, EDGE y UMTS. Para esto, es necesario tener un móvil multimedia con la capacidad y correctamente configurado para poder enviar y recibir mensajes multimedia.

Los teléfonos con MMS cuentan con un editor que permite a los usuarios crear y editar los contenidos de sus mensajes multimedia. Al utilizar este editor junto con, por ejemplo, una cámara, los usuarios pueden realizar mensajes al estilo “power point” o componer postales digitales para enviárselas a sus amigos o clientes.

Así mismo, actualmente las operadoras de telefonía celular también ofrecen a sus clientes servicios de información y entretenimiento vía MMS tales como: postales, información de clima, deportes, poemas, etc.

## **1.2.5 Servicios de valor agregado para mercado corporativo.**

### **1.2.5.1 Oficina móvil.**

Servicio diseñado para asegurar la productividad de los usuarios, aún cuando estén fuera de su oficina. Con la sincronización en tiempo real de toda su información; el usuario podrá comunicarse con sus contactos, confirmar sus juntas, reenviar confirmaciones mediante correos y revisar documentos adjuntos.

Beneficios y Funcionalidades:

- ◆ Brinda operación continua manteniendo una conexión permanente a la red inalámbrica de datos GRPS/EDGE/UMTS permitiendo envío y recepción del correo electrónico en tiempo real.
- ◆ Permite la sincronización inalámbrica del calendario y su organizador personal (libreta de direcciones, lista de tareas y block de notas) que utiliza a diario en la cuenta de correo de su empresa.
- ◆ Tecnología avanzada de cifrado Triple DES a fin de cumplir con las más exigentes pautas de seguridad de la empresa.

### **1.2.5.2 Acceso dedicado.**

Este servicio permite crear al cliente corporativo conexiones privadas de datos entre la red de su empresa y sus dispositivos móviles.

A través de las redes de datos GPRS/EDGE/UMTS el cliente podrá acceder a aplicaciones de su empresa de manera privada, segura y confiable.

Brinda acceso a aplicaciones productivas como son:

- ◆ Cuenta de correo empresarial.
- ◆ Intranet Corporativa.
- ◆ Sistema de levantamiento de pedidos.
- ◆ Telemetría.
- ◆ Sistema de control de inventarios, etc.

Algunos de los beneficios del servicio son:

- ◆ **Confiablez.** Acceso a aplicaciones de su negocio o empresa de forma segura las 24 horas.
- ◆ **Administración.** Mayor seguridad en la asignación de los usuarios que utilizan el acceso.
- ◆ **Control.** Acceso únicamente a aplicaciones específicas de la empresa a través de dispositivos inalámbricos.
- ◆ **Escalabilidad.** Asegura el crecimiento controlado y planeado de su conexión de datos respecto a la expansión de su negocio o empresa.
- ◆ **Cobertura.** Roaming nacional e internacional dependiendo de la operadora.

### 1.2.5.3 Servicios basados en localización.

Los LBS (Location Based Services) o LDIS (Location Dependent Information Services) hacen referencia a servicios basados en localización o para algunos autores simplemente Servicios de Localización.

Los servicios basados en localización buscan ofrecer un servicio personalizado a los usuarios basado en información de ubicación geográfica de estos. Para su operación utiliza tecnología de Sistemas de Información Geográfica, alguna tecnología de posicionamiento bien sea de lado cliente (ej. GPS) o de lado servidor (ej. servicio de posicionamiento suministrado por el operador de la red ) y tecnología de comunicación de redes para transmitir información hacia una aplicación LBS que pueda procesar y responder la solicitud.

Las aplicaciones típicas LBS buscan proveer servicios geográficos en tiempo real. Algunos ejemplos típicos de esto son servicios de mapas, enrutamiento y páginas amarillas geográficas. A continuación se muestra un ejemplo:

### **Servicio de Localización y administración Vehicular.**

Servicio que permite a empresas, desde sus propias instalaciones, localizar, rastrear y monitorear vehículos particulares o utilitarios bajo un concepto avanzado de interacción con el equipo instalado en el vehículo. Además permite la obtención de reportes estadísticos y un nivel sofisticado de funcionalidades desde el sistema y/o a través del envío de mensajes escritos desde un celular previamente definido.

Las funcionalidades que ofrece este servicio son:

- ◆ Localización de un vehículo o grupo de vehículos sobre mapas digitales georeferenciados.
- ◆ Visualización de trayectorias (Reproducción de Históricos).
- ◆ Reportes detallados de localizaciones por distancia y tiempo.
- ◆ Posibilidad de exportar reportes históricos a Excel y/o CSV a través de la página Web.
- ◆ Funcionalidad para medir distancias.
- ◆ Envío de SMS al conductor del vehículo en una pantalla del módem o al equipo celular del conductor.
- ◆ Programación de Rutas.
- ◆ Definición de Puntos de Interés de forma particular para cada empresa adicionales a los Puntos de Interés incluidos en los mapas.
- ◆ Ejecución de comandos enviados desde un teléfono celular vía SMS.
- ◆ Notificación de alarmas vía SMS a un teléfono móvil predefinido.
- ◆ Reportes asociados a rutas que incluyen tiempo y distancia que le tomo al vehículo llegar de un punto A a B, B a C, C a D.

#### **1.2.5.4 Servicios corporativos basados en mensajería.**

##### **Alertas de Sistemas Generadas por Registros.**

Servicio que permite el envío de alertas y/o alarmas desde los sistemas de la empresa reportando algún evento de alta/media trascendencia para el personal de soporte, de campo o personal que requiere de estar informado a través de mensajes de texto (SMS). Las alertas son tomadas desde un registro (archivo) generado por el sistema de la empresa el cual contiene el mensaje que se requiere enviar y una lista de números celulares (hasta 20 números) a los cuales se les notificará el evento.

##### **Mensajería a Grupos.**

Servicio que permite la comunicación vía mensajes de texto desde una PC, por medio de una interface gráfica, hasta un usuario o grupo de usuarios móviles. La interface de comunicación gráfica permite crear un directorio de contactos, crear grupos y listas de distribución, insertar mensajes predefinidos además de ser el medio por el cual se genera el mensaje.

Beneficios y Funcionalidades:

- ◆ Medio efectivo de comunicación para la empresa utilizando el servicio de mensajes de texto (SMS).
- ◆ Comunicación directa desde una PC de la empresa hasta un dispositivo móvil.
- ◆ Mensajes de hasta 150 caracteres.
- ◆ Servicio de comunicación de bajo costo.

##### **Correo vía SMS.**

Este servicio permite que los usuarios móviles reciban avisos vía mensajes de texto (SMS) originados desde cualquier sistema con capacidad de crear un correo electrónico.

Al usuario móvil se le asigna una dirección de la forma número celular@operadora.com la cual permite que el mensaje en formato de correo electrónico sea convertido al formato de SMS para que sea entregado al destinatario móvil.

Beneficios y Funcionalidades:

- ◆ Medio efectivo de comunicación para la empresa utilizando el servicio de mensajes de texto (SMS).
- ◆ Comunicación directa desde cualquier cuenta de correo hasta un dispositivo móvil.
- ◆ Mensajes de hasta 300 caracteres (el mensaje se divide en 2 SMS de 150 caracteres).
- ◆ Integración del servicio con sistemas que manejen el protocolo SMTP (protocolo de comunicación del correo electrónico).
- ◆ Servicio de comunicación de bajo costo.

## **2 Capítulo 2**

---

Definición de conceptos empleados en  
arquitecturas de servidores.

## 2.1 Definición de arquitectura.

Una arquitectura es un conjunto de componentes funcionales interconectados que aprovechando diferentes estándares, convenciones, reglas y procesos, permite integrar una amplia gama de productos y servicios informáticos, de manera que pueden ser utilizados eficazmente dentro de una empresa u organización.

Es importante señalar que para seleccionar el modelo de una arquitectura, hay que partir del contexto tecnológico y organizativo del momento y, que la arquitectura Cliente/Servidor requiere una determinada especialización de cada uno de los diferentes componentes que la integran.

## 2.2 Arquitectura cliente / servidor.

La **arquitectura cliente/servidor** es un modelo para el desarrollo de sistemas de información en donde las transacciones se dividen en procesos independientes que cooperan entre sí para intercambiar información, servicios o recursos. Se denomina **cliente** al proceso que inicia el diálogo o solicita los recursos y **servidor** al proceso que responde a las solicitudes.

Los clientes y los servidores pueden estar conectados a una red local o una red amplia, como la que se puede implementar en una empresa o a una red mundial como Internet.

Bajo este modelo cada usuario tiene la libertad de obtener la información que requiera en un momento dado proveniente de una o varias fuentes locales o distantes y de procesarla como según le convenga. Los distintos servidores también pueden intercambiar información dentro de esta arquitectura.

Entre las principales características de la arquitectura cliente/servidor se pueden destacar las siguientes:

- ◆ El servidor presenta a todos sus clientes una interfaz única y bien definida.
- ◆ El cliente no necesita conocer la lógica del servidor, sólo su interfaz externa.
- ◆ El cliente no depende de la ubicación física del servidor, ni del tipo de equipo físico en el que se encuentra, ni de su sistema operativo.
- ◆ Los cambios en el servidor implican pocos o ningún cambio en el cliente.

### 2.2.1 Ventajas de la arquitectura cliente-servidor.

- ◆ Centralización del control: los accesos, recursos y la integridad de los datos son controlados por el servidor de forma que un programa cliente defectuoso o no autorizado no pueda dañar el sistema.
- ◆ Escalabilidad: se puede aumentar la capacidad de clientes y servidores por separado.

#### Ejemplo:

Visitar un sitio Web es un buen ejemplo de la arquitectura cliente/servidor. El servidor web sirve las páginas web al navegador. La mayoría de los servicios del Internet son tipo de servidores. Algunos ejemplos son servidores de archivo, DNSs, servidores de la impresora, etc.

## 2.3 Servidor de aplicación.

### 2.3.1 Definición.

En informática se denomina **servidor de aplicaciones** a un servidor en una red de computadores que ejecuta ciertas aplicaciones

Usualmente se trata de un dispositivo de software que proporciona servicios de aplicación a las computadoras cliente. Un servidor de aplicaciones generalmente gestiona la mayor parte (o la totalidad) de las funciones de lógica de negocio y de acceso a los datos de la aplicación. Los principales beneficios de la aplicación de la tecnología de servidores de aplicación son la centralización y la disminución de la complejidad en el desarrollo de aplicaciones.

Si bien el término es aplicable a todas las plataformas de software, hoy en día el término *servidor de aplicaciones* se ha convertido en sinónimo de la plataforma J2EE de Sun Microsystems<sup>5</sup>.

---

<sup>5</sup> J2EE- Java Platform, Enterprise Edition. Ver explicación en el punto 2.4 de este capítulo.

### 2.3.2 Características comunes.

Los servidores de aplicación típicamente incluyen también *middleware* (o software de conectividad) que les permite intercomunicarse con variados servicios, para efectos de confiabilidad, seguridad, no-repudio, etc. Los servidores de aplicación también brindan a los desarrolladores una Interfaz para Programación de Aplicaciones (API), de tal manera que no tengan que preocuparse por el sistema operativo o por la gran cantidad de interfaces requeridas en una aplicación web moderna.

Los servidores de aplicación también brindan soporte a una gran variedad de estándares, tales como HTML<sup>6</sup>, XML<sup>7</sup>, JDBC<sup>8</sup>, SSL<sup>9</sup>, etc., que les permiten su funcionamiento en ambientes web (como Internet) y la conexión a una gran variedad de fuentes de datos, sistemas y dispositivos.

### 2.3.3 Uso.

Un ejemplo común del uso de servidores de aplicación (y de sus componentes) son los portales de Internet, que permiten a las empresas la gestión y divulgación de su información, y un punto único de entrada a los usuarios internos y externos. Teniendo como base un servidor de aplicación, dichos portales permiten tener acceso a información y servicios (como servicios Web) de manera segura y transparente, desde cualquier dispositivo.

---

<sup>6</sup> HTML HyperText Markup Language o Lenguaje de Marcas Hipertextuales. Es un lenguaje de marcación diseñado para estructurar textos y presentarlos en forma de hipertexto, que es el formato estándar de las páginas web.

<sup>7</sup> XML. eXtensible Markup Language («lenguaje de marcas extensible»), es un metalenguaje extensible de etiquetas desarrollado por el World Wide Web Consortium (W3C).

<sup>8</sup> JDBC. JDBC es el acrónimo de Java Database Connectivity, un API que permite la ejecución de operaciones sobre bases de datos desde el lenguaje de programación Java independientemente del sistema operativo donde se ejecute o de la base de datos.

<sup>9</sup> SSL. Secure Sockets Layer proporciona autenticación y privacidad de la información entre extremos sobre Internet mediante el uso de criptografía.

## 2.4 J2EE: Java 2 Enterprise Edition.

Java Platform, Enterprise Edition o Java EE (anteriormente conocido como Java 2 Platform, Enterprise Edition o J2EE hasta la versión 1.4), es una plataforma de programación para desarrollar y ejecutar software de aplicaciones en Lenguaje de programación Java con arquitectura de n niveles distribuida, basándose ampliamente en componentes de software modulares ejecutándose sobre un servidor de aplicaciones. La plataforma Java EE está definida por una especificación. Similar a otras especificaciones del Java Community Process, Java EE es también considerada informalmente como un estándar debido a que los suministradores deben cumplir ciertos requisitos de conformidad para declarar que sus productos son conformes a Java EE; no obstante sin un estándar de ISO o ECMA.

### 2.4.1 Servidores de aplicación J2EE.

Como consecuencia del éxito del lenguaje de programación Java, el término *servidor de aplicaciones* usualmente hace referencia a un servidor de aplicaciones J2EE. WebSphere (IBM), Oracle Application Server (Oracle Corporation) y WebLogic (BEA) están entre los servidores de aplicación J2EE privados más conocidos.

Entre los servidores de aplicación J2EE de dominio público se encuentran: JOnAS<sup>10</sup> y JBoss<sup>11</sup>. El servidor de aplicaciones JOnAS, desarrollado por el consorcio ObjectWeb fue el primer servidor de aplicaciones libre en lograr certificación oficial de compatibilidad con J2EE. JBoss es un servidor de aplicaciones de código abierto implementado en Java puro. Al estar basado en Java, JBoss puede ser utilizado en cualquier sistema operativo que lo soporte.

J2EE provee estándares que le permiten a un servidor de aplicaciones servir como "contenedor" de los componentes que conforman dichas aplicaciones. Estos componentes, escritos en lenguaje Java, usualmente se conocen como Servlets, Java Server Pages (JSPs)<sup>12</sup> y Enterprise JavaBeans (EJBs)<sup>13</sup> y permiten implementar diferentes capas de la aplicación, como la interfaz de usuario, la lógica de negocio, la gestión de sesiones de usuario o el acceso a bases de datos remotas.

---

<sup>10</sup> JOnAS. Es un servidor de aplicaciones J2EE de código abierto implementado en Java.

<sup>11</sup> JBoss. Es un servidor de aplicaciones J2EE de código abierto implementado en Java puro. Al estar basado en Java, JBoss puede ser utilizado en cualquier sistema operativo que lo soporte.

<sup>12</sup> JSP. JavaServer Pages es una tecnología Java que permite generar contenido dinámico para web, en forma de documentos HTML, XML o de otro tipo. Esta tecnología es un desarrollo de la compañía Sun Microsystems.

<sup>13</sup> EJB's. Los Enterprise JavaBeans son una de las API que forman parte del estándar de construcción de aplicaciones empresariales J2EE de Sun Microsystems.

La portabilidad de Java también ha permitido que los servidores de aplicación J2EE se encuentren disponibles sobre una gran variedad de plataformas, como Unix, Microsoft Windows y GNU/Linux.

## 2.5 Definición de alta disponibilidad.

Entiéndase como alta disponibilidad de un servicio o aplicación de red, el hecho de prestarlo ininterrumpidamente y con relativa independencia del hardware y/o software que lo sustenta. Se consigue alta disponibilidad redundando los sistemas, como por ejemplo implementando algún tipo de arquitectura en cluster.

### 2.5.1 Cluster de alta disponibilidad.

El término **cluster** se aplica al conjunto de hardware (equipos de cómputo, equipos de red: switches, routers, etc.) o software aplicativo (base de datos, servidores de aplicación, procesos aplicativos, etc.) que se comportan como si fuesen una única entidad.

Para que un cluster funcione como tal, no basta solo con conectar entre sí el hardware y/o software aplicativo, sino que es necesario proveer un *sistema de manejo del cluster*, el cual se encargue de monitorear y gestionar la operación de dicho hardware o software.

Un **cluster de alta disponibilidad** es un conjunto hardware y/o software aplicativo que se caracteriza por brindar el funcionamiento ininterrumpido de una aplicación o servicio de manera transparente al usuario final.

Si se produce un fallo del hardware o de las aplicaciones de alguno de los nodos que conforman el cluster, el *software de alta disponibilidad* es capaz de re-iniciar automáticamente los servicios que han fallado en cualquiera de los otros nodos del cluster. Y cuando el nodo que ha fallado se recupera, los servicios son nuevamente migrados al nodo original. Esta capacidad de recuperación automática de servicios nos garantiza la integridad de la información, porque no hay pérdida de datos, y además evita molestias a los usuarios, que no tienen por qué notar que se ha producido un problema.

### 2.5.1.1 Componentes del Cluster.

En general, un cluster necesita de varios componentes de software y hardware para poder funcionar:

#### **Nodos.**

Pueden ser simples ordenadores, sistemas multi-procesador, estaciones de trabajo (*workstations*), routers, switches de red, balanceadores de carga, etc.

#### **Sistemas Operativos.**

Deben ser multiproceso, multiusuario. Otras características deseables son la facilidad de uso y acceso. Ejemplos: GNU/Linux, Unix: Solaris, HP-UX; Windows NT, etc.

#### **Conexiones de Red.**

Los nodos de un cluster pueden conectarse mediante una simple red Ethernet con adaptadores de red o NIC's, o utilizarse tecnologías especiales de alta velocidad como Fast Ethernet, Gigabit Ethernet, etc.

#### **Middleware (capa de abstracción entre el usuario y los sistemas operativos).**

El **middleware** es un software que generalmente actúa entre el sistema operativo y las aplicaciones con la finalidad de proveer a un cluster lo siguiente:

- ◆ Una interfaz única de acceso al sistema, denominada SSI (*Single System Image*), la cual genera la sensación al usuario de que utiliza un único ordenador muy potente;
- ◆ Herramientas para la optimización y mantenimiento del sistema: migración de procesos, *checkpoint-restart* (congelar uno o varios procesos, mudarlos de servidor y continuar su funcionamiento en el nuevo host), balanceo de carga, tolerancia a fallos, etc.;

- ◆ Escalabilidad: debe poder detectar automáticamente nuevos servidores conectados al cluster para proceder a su utilización.

Existen diversos tipos de middleware, como por ejemplo: MOSIX<sup>14</sup>, OpenMOSIX, Cándor, OpenSSI, etc.

El middleware recibe los trabajos entrantes al cluster y los redistribuye de manera que el proceso se ejecute más rápido y el sistema no sufra sobrecargas en un servidor. Esto se realiza mediante políticas definidas en el sistema (automáticamente o por un administrador) que le indican dónde y cómo debe distribuir los procesos, por un sistema de monitorización, el cual controla la carga de cada CPU y la cantidad de procesos en él.

El middleware también debe poder migrar procesos entre servidores con distintas finalidades:

- ◆ Balancear la carga: si un servidor está muy cargado de procesos y otro está ocioso, pueden transferirse procesos a este último para liberar de carga al primero y optimizar el funcionamiento;
- ◆ Mantenimiento de servidores: si hay procesos corriendo en un servidor que necesita mantenimiento o una actualización, es posible migrar los procesos a otro servidor y proceder a desconectar del cluster al primero;
- ◆ Priorización de trabajos: en caso de tener varios procesos corriendo en el cluster, pero uno de ellos de mayor importancia que los demás, puede migrarse este proceso a los servidores que posean más o mejores recursos para acelerar su procesamiento.

---

<sup>14</sup> MOSIX es un administrador para clusters Linux y multi-clusters.

## 2.6 Arquitectura de redes.

Actualmente las redes de comunicación, están diseñadas de una forma muy estructurada. Éstas, se organizan en capas o niveles con el objetivo de reducir la complejidad de su diseño. Cada una de ellas se construye sobre su predecesora. Tanto el número de capas como el nombre, contenido y función pueden variar de una red a otra, sin embargo, en cualquier red, el propósito de cada capa es ofrecer ciertos servicios a las capas superiores, liberándolas del conocimiento detallado sobre cómo se realizan dichos servicios.

Entre cada par de capas adyacentes hay una **interface**, la cual define los servicios y operaciones primitivas que la capa inferior ofrece a la superior.

Para que dos o más entidades<sup>15</sup> se comuniquen con éxito, se requiere que "hablen el mismo idioma", es decir, debe existir un entendimiento mutuo entre dichas entidades. Este entendimiento es posible gracias al empleo de los protocolos de comunicación. Un **protocolo**, es un conjunto formal de reglas y convenciones que gobiernan el modo en que las entidades involucradas en la comunicación intercambian información.

Al conjunto de capas y protocolos se le denomina **arquitectura de red**. Las especificaciones de ésta deberán contener la información suficiente que le permita al diseñador escribir un programa o construir el hardware correspondiente a cada capa, y que siga en forma correcta el protocolo apropiado.

### 2.6.1 Modelo de referencia OSI.

El modelo de referencia OSI es la arquitectura de red actual más prominente. El objetivo de éste es el de desarrollar estándares para la interconexión de sistemas abiertos (Open System Interconnection, OSI).

---

<sup>15</sup> En general, una entidad, es un dispositivo de red administrable individual. Por ejemplo, una computadora.

Los principios que maneja el modelo OSI son los siguientes:

- ◆ Una capa se creará cuando se necesita un nivel diferente de abstracción.
- ◆ Cada capa deberá efectuar una función bien definida.
- ◆ La función que realizará cada capa deberá seleccionarse con la intención de definir protocolos normalizados internacionalmente.
- ◆ Los límites de las capas deberán seleccionarse tomando en cuenta la minimización del flujo de información a través de las interfaces.
- ◆ El número de capas deberá ser lo suficientemente grande para que funciones diferentes no tengan que ponerse juntas en la misma capa y también lo suficientemente pequeño para que su arquitectura no deba ser difícil de manejar.

El modelo OSI proporciona un modelo conceptual para la comunicación entre entidades, pero el modelo en sí mismo no es un método de comunicación.

OSI es un modelo de 7 capas, donde cada capa define los procedimientos y las reglas (protocolos normalizados) que los sistemas de comunicaciones deben seguir, para poder comunicarse con sus procesos correspondientes de los otros sistemas. Esto permite que un proceso que se ejecuta en una entidad, pueda comunicarse con un proceso similar en otra entidad, si tienen implementados los mismos protocolos de comunicaciones de capas OSI.

## 2.6.2 Capas del modelo OSI.

Según la clasificación del **modelo de referencia de Interconexión de Sistemas Abiertos** (OSI, Open System Interconnection), la comunicación de varios dispositivos ETD<sup>16</sup> se puede estudiar dividiéndola en 7 niveles, que son expuestos desde su nivel más alto hasta el más bajo:

Nivel	Nombre	Categoría
Capa 7	Nivel de aplicación	Aplicación
Capa 6	Nivel de presentación	
Capa 5	Nivel de sesión	
Capa 4	Nivel de transporte	
Capa 3	Nivel de red	Transporte de datos
Capa 2	Nivel de enlace de datos	
Capa 1	Nivel físico	

Tabla 2.1. Capas del modelo OSI.

A su vez, esos 7 niveles se pueden subdividir en dos categorías, las capas superiores y las capas inferiores. Las 4 capas superiores trabajan con problemas particulares a las aplicaciones, y las 3 capas inferiores se encargan de los problemas pertinentes al transporte de los datos.

Los protocolos de cada capa tienen una interfaz bien definida. Una capa generalmente se comunica con la capa inmediata inferior, la inmediata superior, y la capa del mismo nivel en otros computadores de la red. Esta división de los protocolos ofrece abstracción en la comunicación.

Una aplicación (capa nivel 7) por ejemplo, solo necesita conocer como comunicarse con la capa 6 que le sigue, y con otra aplicación en otro computador (capa 7). No necesita conocer nada entre las capas de la 1 a la 5. Así, un navegador web (HTTP<sup>17</sup>, capa 7) puede utilizar una conexión Ethernet o PPP<sup>18</sup> (capa 2) para acceder a la Internet, sin que sea necesario cualquier tratamiento para los protocolos de este nivel más bajo.

<sup>16</sup> ETD. Equipo Terminal de Datos o DTE *Data Terminal Equipment*, es aquel componente del circuito de datos que hace de fuente o destino de la información. Pueden ser un terminal, una impresora o también un potente ordenador.

<sup>17</sup> HTTP. HyperText Transfer Protocol, define la sintaxis y la semántica que utilizan los elementos software de la arquitectura web (clientes, servidores, proxies) para comunicarse.

<sup>18</sup> PPP Point-to-point Protocol, es un protocolo asociado a la pila TCP/IP de uso en Internet. El protocolo PPP permite establecer una comunicación a nivel de enlace entre dos computadoras.

De la misma forma, un router sólo necesita de las informaciones del nivel de red para enrutar paquetes, sin que importe si los datos en tránsito pertenecen a una imagen para un navegador web, un archivo transferido vía FTP<sup>19</sup> o un mensaje de correo electrónico.

### **2.6.2.1 Capa física.**

Esta capa se encarga de activar, mantener y desactivar el enlace físico entre sistemas de redes de comunicaciones. Asimismo, define las especificaciones eléctricas, mecánicas y de procedimiento.

Las especificaciones eléctricas son las que determinan los niveles de tensión (o corriente) y la temporización de los cambios eléctricos que representan los unos y los ceros. Las especificaciones mecánicas, describen los conectores y los hilos de interfaz. Cabe mencionar que las interfaces empleadas en la capa física, se utilizan para conectar dispositivos de usuario al circuito de comunicaciones.

Por último, las especificaciones de procedimiento describen lo que deben hacer los conectores, y la secuencia de eventos necesaria para llevar a cabo la transferencia efectiva de datos a través del interfaz.

### **2.6.2.2 Capa de enlace de datos.**

Es la encargada de proporcionar una transmisión confiable de datos a través del enlace de red. Una de las funciones más importantes de esta capa consiste en detectar errores en la transmisión y en recuperar, por distintos mecanismos, los datos perdidos, duplicados o erróneos; además, alerta a los protocolos de las capas superiores cuando se presenta un error en la transmisión. También se ocupa del direccionamiento físico, el cual, a diferencia del direccionamiento de red, define cómo se nombran los dispositivos en la capa de enlace de datos. Asimismo, se encarga de controlar el flujo de datos para que el dispositivo receptor no se sature con más tráfico del que pueda manejar simultáneamente.

---

<sup>19</sup> FTP. File Transfer Protocol. Es un protocolo de transferencia de archivos entre sistemas conectados a una red TCP basado en la arquitectura cliente-servidor, de manera que desde un equipo cliente nos podemos conectar a un servidor para descargar archivos desde él o para enviarle nuestros propios archivos independientemente del sistema operativo utilizado en cada equipo.

### **2.6.2.3 Capa de red.**

Esta capa se ocupa de las operaciones de encaminamiento por la red y proporciona la comunicación entre distintas redes. Es decir, determina cómo encaminar los paquetes de una entidad origen a una entidad destino, pudiendo tomar distintas soluciones. Lo anterior, se logra a través del direccionamiento lógico de dichas entidades.

El control de la congestión es también problema de este nivel, así como la responsabilidad para resolver problemas de interconexión de redes heterogéneas (con protocolos diferentes, etc.).

Los protocolos de la capa de red son de hecho protocolos de ruteo, sin embargo también otro tipo de protocolos están implantados en la capa de red. La capa de red soporta servicios orientados y no orientados a la conexión de los protocolos de las capas superiores.

### **2.6.2.4 Capa de transporte.**

La función principal de la capa de transporte es aceptar los datos de la capa de sesión, dividirlos en unidades más pequeñas, siempre y cuando sea necesario, pasarlos a la capa de red y asegurar que todos ellos lleguen correctamente a la entidad destino de la manera más eficiente.

Por otra parte, esta capa se necesita para hacer el trabajo de multiplexión transparente a la capa de sesión, es decir, permite que los datos de diferentes aplicaciones sean transmitidos en un enlace físico único. Además se ocupa del establecimiento y liberación de conexiones a través de la red.

Otra de sus funciones habituales de la capa de transporte es el control de flujo entre entidades, esto es, administra la transmisión de datos entre dispositivos para que el dispositivo transmisor no envíe más datos de los que pueda procesar el dispositivo receptor.

#### **2.6.2.5 Capa de sesión.**

Esta capa permite que los usuarios de diferentes máquinas puedan establecer sesiones entre ellos. Una sesión podría permitir al usuario acceder a un sistema de tiempo compartido a distancia, o transferir un archivo entre dos máquinas.

Las sesiones de comunicación constan de solicitudes y respuestas de servicio que se presentan entre aplicaciones ubicadas en diferentes dispositivos de red. Estas solicitudes y respuestas están coordinadas por protocolos implantados en la capa de sesión.

Por otra parte, esta capa se encarga de la administración del testigo y la sincronización entre el origen y destino de los datos.

#### **2.6.2.6 Capa de presentación.**

Se ocupa de los aspectos de sintaxis y semántica de la información que se transmite y no del movimiento fiable de bits de un lugar a otro.

La capa de presentación, brinda una gama de funciones de codificación y conversión que se aplican a los datos de la capa de aplicación. Estas funciones aseguran que la información enviada desde la capa de aplicación de un sistema sea legible por la capa de aplicación de otro sistema

La capa de presentación está relacionada también con otros aspectos de representación de la información. Por ejemplo, la compresión de datos se puede utilizar aquí para reducir el número de bits que tienen que transmitirse, y el concepto de criptografía, el cual se utiliza frecuentemente por cuestiones de privacidad y autenticación.

#### **2.6.2.7 Capa de aplicación.**

Esta capa es la interfaz entre el usuario para interactuar de manera directa con las aplicaciones de software. En este nivel están los programas de administración de bases de datos, los programas de servidores de archivos y de servidores de impresión, los comandos y lenguajes de respuestas de los sistemas operativos. El software de aplicaciones como el procesamiento de textos o hojas de cálculo no están en la capa de aplicaciones, sólo los protocolos que les permiten funcionar.

### 2.6.3 Ejemplos de protocolos.

Capa 1: **Nivel físico**

Cable coaxial, Cable de fibra óptica, Cable de par trenzado, Microondas, Radio, RS-232.<sup>20</sup>

Capa 2: **Nivel de enlace de datos**

Ethernet, Fast Ethernet, Gigabit Ethernet, Token Ring, FDDI, ATM, HDLC.

Capa 3: **Nivel de red**

ARP, RARP, IP (IPv4, IPv6), X.25, ICMP, IGMP, NetBEUI, IPX, Appletalk.

Capa 4: **Nivel de transporte**

TCP, UDP, SPX.

Capa 5: **Nivel de sesión**

NetBIOS, RPC, SSL.

Capa 6: **Nivel de presentación**

ASN.1.

Capa 7: **Nivel de aplicación**

SNMP, SMTP, NNTP, FTP, SSH, HTTP, NFS, Telnet, IRC, ICQ, POP3, IMAP.

---

<sup>20</sup> Ver definiciones de protocolos mencionados de la capa 1 a la 7 en glosario.

## 2.6.4 Topologías de red.

Los nodos de una red necesitan estar conectados para comunicarse, a la forma en que están conectados dichos nodos se le llama **topología**<sup>21</sup>. Una red tiene dos diferentes topologías: una física y una lógica. La **topología física** es la disposición física actual de la red, la manera en que los nodos están conectados unos con otros. La **topología lógica** es el método que se usa para comunicarse con los demás nodos, es decir, la ruta que toman los datos de la red entre los diferentes nodos de la misma.

La forma de una red, no limita los medios de transmisión, es decir, tanto los cables de par trenzado como los cables coaxiales y los de fibra óptica, se pueden adaptar a las diferentes topologías que existen. A continuación se describen las topologías de red más comunes:

### 2.6.4.1 Topología jerárquica.

La estructura jerárquica es una de las más extendidas actualmente. Esta topología proporciona un punto de concentración de las tareas de control y de resolución de errores. En la mayoría de los casos, el DTE<sup>22</sup> situado en el nivel más elevado de la jerarquía es el que controla la red.

Aunque la topología jerárquica resulta interesante por ser fácil de controlar, puede presentar ciertos problemas en cuanto a la posibilidad de aparición de cuellos de botella. En determinadas situaciones, el DTE más elevado, normalmente un gran ordenador central, ha de controlar todo el tráfico entre los distintos DTE. Este hecho no sólo puede crear saturaciones de datos, sino que además plantea serios problemas de fiabilidad.

Las redes con topología jerárquica se conocen también como redes verticales o en árbol.

---

<sup>21</sup> El término topología es un concepto geométrico que alude al aspecto de una cosa.

<sup>22</sup> DTE. Data Terminal Equipment (equipo terminal de datos). Estas siglas se utilizan para referirse de forma genérica al equipo que emplea el usuario final; por ejemplo, las estaciones de trabajo.

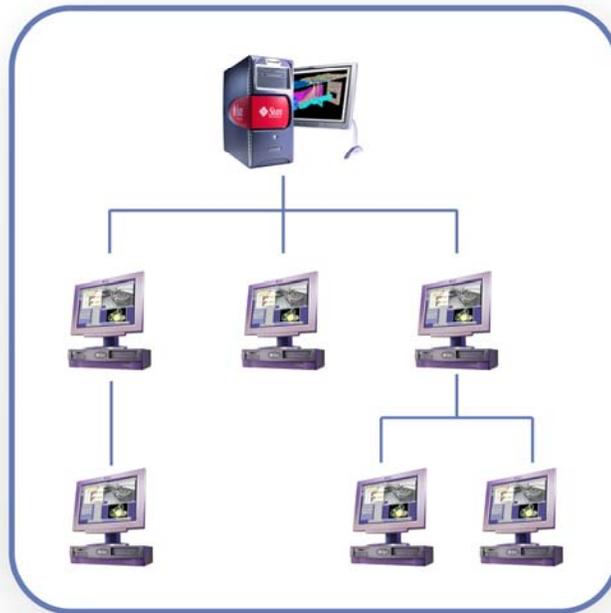


Figura 2.1 Topología jerárquica.

#### 2.6.4.2 Topología horizontal (bus).

Este tipo de topología es frecuente en las redes de área local. Es relativamente fácil controlar el flujo de tráfico entre los distintos DTE, el bus permite que todas las estaciones reciban todas las transmisiones, es decir, una estación puede difundir la información a todas la demás. Ethernet es un ejemplo de una red que usa el método de bus.

La principal limitación de una topología horizontal está en el hecho de que suele existir un sólo canal de comunicaciones para todos los dispositivos de la red. En consecuencia, si el canal de comunicaciones falla, toda la red deja de funcionar. Además, la seguridad podría verse amenazada por un usuario no autorizado.

Una desventaja de la topología de bus es que, por lo general, debe haber un mínimo de distancia entre las derivaciones, de forma que las estaciones de trabajo puedan evadir las interferencias de señales.

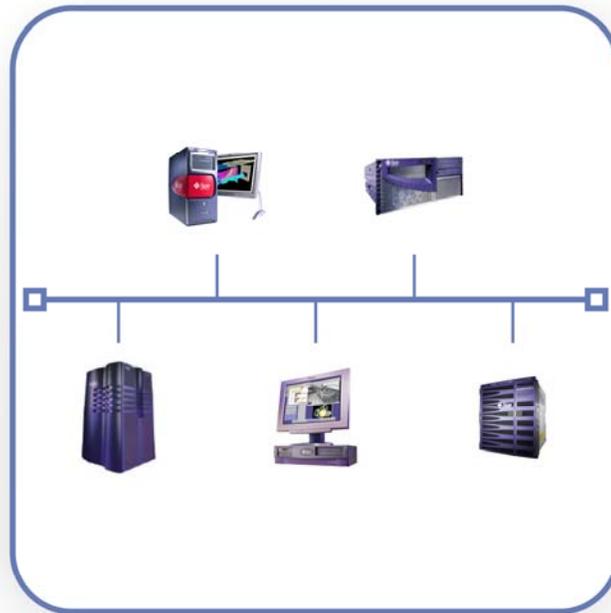


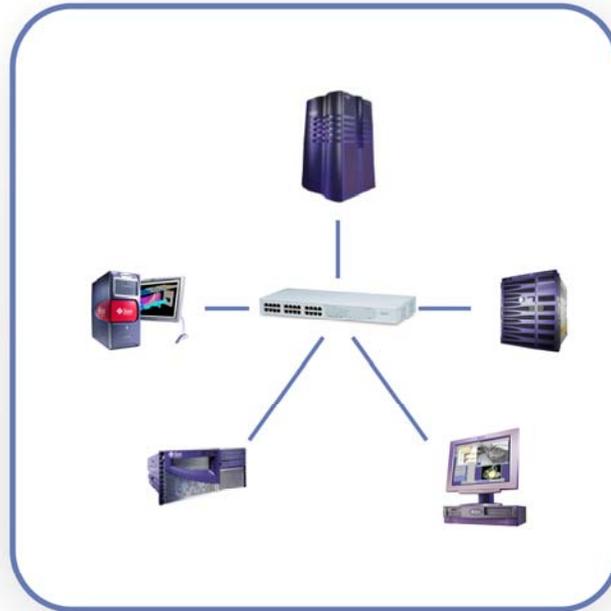
Figura 2.2 Topología horizontal (bus).

#### 2.6.4.3 Topología en estrella.

La topología en estrella es una de las más empleadas en los sistemas de comunicaciones de datos. Todo el tráfico emana del núcleo de la estrella, como se observa en la figura 2.3.

El nodo central, es por lo general un concentrador (hub) o switch, el cual posee el control total de los DTE conectados a éste. Además, es responsable de encaminar el tráfico hacia el resto de los componentes. La localización de fallas, es relativamente una actividad sencilla porque es posible aislar las líneas para identificar el problema.

Esta topología, facilita la adición de nuevos DTE a la red. Todo lo que se requiere es un cable que vaya del punto central de conexión (concentrador) a la tarjeta de interfaz de red de cada DTE.



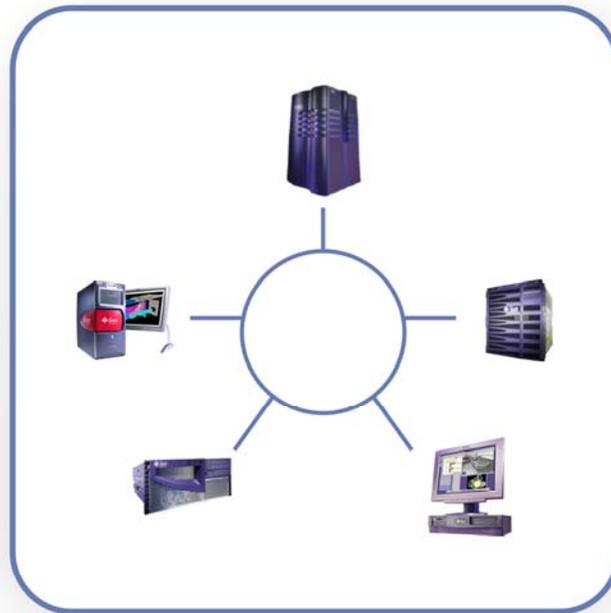
**Figura 2.3 Topología en estrella**

Otra ventaja de la topología estrella es que el administrador de la red puede asignar a ciertos DTE un estatus mayor que a otros. Por tanto, el nodo central tenderá a buscar las señales de estos DTE prioritarias antes de reconocer a otros DTE. Para redes que tengan algunos usuarios clave que requieran respuestas inmediatas a sus solicitudes en línea, esta característica de la topología estrella puede ser de extrema utilidad.

Sin embargo, y al igual que en la estructura jerárquica, la principal deficiencia de esta topología es que puede sufrir saturaciones o hasta la obstrucción total del funcionamiento de la red en caso de avería del nodo central.

#### **2.6.4.4 Topología en anillo.**

La estructura en anillo como se observa en la siguiente figura, se llama así por el aspecto circular del flujo de datos.



**Figura 2.4 Topología en anillo**

En la mayoría de los casos, los datos fluyen en una sola dirección, y cada estación recibe la señal y la transmite a la siguiente del anillo. La organización en anillo resulta atractiva porque con ella son bastante raros los embotellamientos, tan frecuentes en los sistemas en estrella o en árbol. Además, la lógica necesaria para poner en marcha una red de este tipo es relativamente simple. Cada componente ha de llevar a cabo una serie de tareas muy sencillas: aceptar los datos, enviarlos al DTE conectado al anillo o retransmitirlos al próximo componente del mismo.

Sin embargo, como todas las redes, la red en anillo tiene algunos defectos. El problema más importante es que todos los componentes del anillo están unidos por un mismo canal. Si falla el canal entre dos nodos, toda la red se interrumpe. Por eso algunos fabricantes han ideado diseños especiales que incluyen canales de seguridad, por si se produce la pérdida de algún canal. Otros fabricantes construyen conmutadores que redirigen los datos automáticamente, saltándose el nodo averiado, hasta el siguiente nodo del anillo, con el fin de evitar que el fallo afecte a toda la red.

#### 2.6.4.5 Topología en malla.

La topología en malla se ha venido empleando en los últimos años. Lo que la hace atractiva es su relativa inmunidad a los problemas de embotellamiento y averías. Es decir, es posible orientar el tráfico por trayectorias alternas en caso de que algún nodo esté averiado u ocupado, esto gracias a la multiplicidad de caminos que ofrece a través de los distintos DTE y DCE<sup>23</sup>.

Sin embargo, para proporcionar estas funciones especiales, la lógica de control de los protocolos de una red en malla puede llegar a ser sumamente complicada.

A pesar de que la realización de este método es compleja y cara, muchos usuarios prefieren la fiabilidad de una red en malla a otras alternativas.

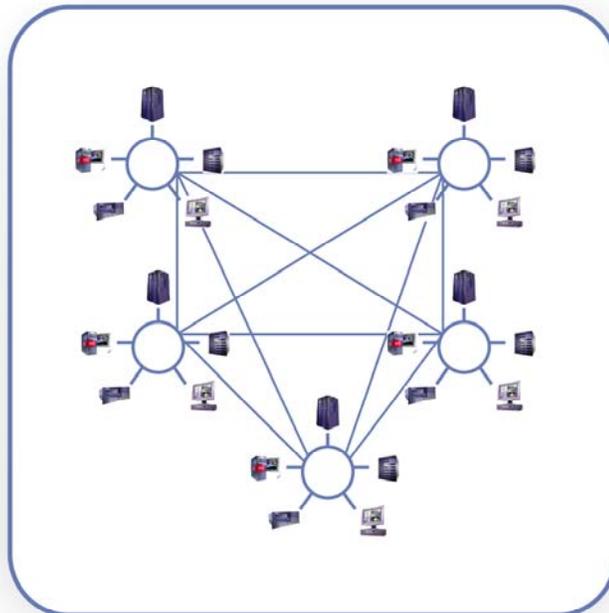


Figura 2.5 Topología en malla

---

<sup>23</sup> DCE. También llamado ETCD *Data Circuit Terminating Equipment* (equipo terminal de circuitos de datos ó equipo de comunicaciones de datos), se emplea para conectar equipos DTE a la línea o canal de comunicaciones, es decir, sirven de interfaz entre el DTE y la red de comunicaciones. Ejemplo: un módem.

## 2.7 Red de área de almacenamiento

Una **red de área de almacenamiento**, en inglés **SAN** (Storage Area Network), es una red concebida para conectar servidores, matrices (arrays) de discos y librerías de respaldo principalmente. Está basada en tecnología fibre channel y más recientemente en iSCSI.

Una red SAN se distingue de otros modos de almacenamiento en red por el modo de acceso a bajo nivel. El tipo de tráfico en una SAN es muy similar al de los discos duros como ATA. En otros métodos de almacenamiento, como NFS, el servidor solicita un determinado fichero, por ejemplo: "/home/usuario/rocks". En una SAN el servidor solicita "el bloque 6000 del disco 4". La mayoría de las SAN actuales usan el protocolo SCSI para acceder a los datos de la SAN, aunque no usen interfaces<sup>24</sup> físicas SCSI.

Una SAN se puede considerar una extensión de Direct Attached Storage (DAS). Donde en DAS hay un enlace punto a punto entre el servidor y su almacenamiento, una SAN permite a varios servidores acceder a varios dispositivos de almacenamiento en una red compartida. Tanto en SAN como en DAS, las aplicaciones y programas de usuarios hacen sus peticiones de datos al sistema de ficheros directamente. La diferencia reside en la manera en la que dicho sistema de ficheros obtiene los datos requeridos del almacenamiento. En DAS, el almacenamiento es local al sistema de ficheros, mientras que en SAN, el almacenamiento es remoto.

SAN utiliza diferentes protocolos de acceso como Fibre Channel y Gigabit Ethernet. En el lado opuesto se encuentra la tecnología Network-Attached Storage (NAS), donde las aplicaciones hacen las peticiones de datos a los sistemas de ficheros de manera remota mediante el protocolo Network File System (NFS). La figura 2.6 muestra un diagrama comparativo entre los diferentes sistemas de almacenamiento: DAS, NAS y SAN.

---

<sup>24</sup> Conexión física y funcional entre dos aparatos o sistemas independientes

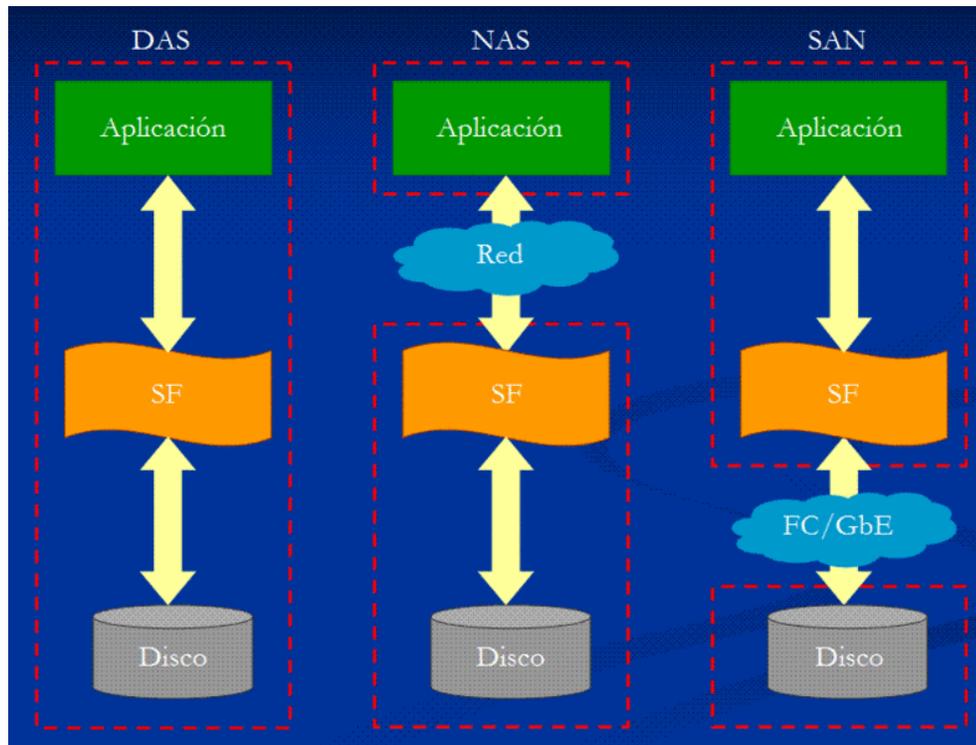


Figura 2.6 Comparación DAS, NAS y SAN.

### 2.7.1 Características de una SAN.

**Latencia**<sup>25</sup> - Una de las diferencias y principales características de las SAN es que son construidas para minimizar el tiempo de respuesta del medio de transmisión.

**Conectividad** - Permite que múltiples servidores sean conectados al mismo grupo de discos o librerías de cintas, permitiendo que la utilización de los sistemas de almacenamiento y los respaldos sean óptimos.

**Distancia** - Las SAN al ser construidas con fibra óptica heredan los beneficios de ésta, por ejemplo, las SAN pueden tener dispositivos con una separación de hasta 10 Km. sin ruteadores.

<sup>25</sup> En redes informáticas de datos se denomina **latencia** a la suma de retardos temporales dentro de una red. Un retardo es producido por la demora en la propagación y transmisión de paquetes dentro de la red.

**Velocidad** - El rendimiento de cualquier sistema de computo dependerá de la velocidad de sus subsistemas, es por ello que las SAN han incrementado su velocidad de transferencia de información, desde 1 Gigabit, hasta actualmente 2 y 4 Gigabits por segundo.

**Disponibilidad** - Una de las ventajas de las SAN es que al tener mayor conectividad, permiten que los servidores y dispositivos de almacenamiento se conecten más de una vez a la SAN, de esta forma, se pueden tener *rut*as redundantes que a su vez incrementaran la tolerancia a fallos.

**Componentes** - Los componentes primarios de una SAN son: switches, HBAs<sup>26</sup>, servidores, ruteadores, gateways, matrices de discos y librerías de cintas.

**ISL** (Inter Switch Link, enlace entre conmutadores) - Actualmente las conexiones entre los switches de SAN se hacen mediante puertos tipo "E" y pueden agruparse para formar una troncal (trunk) que permita mayor flujo de información y tolerancia a fallos.

**Arquitectura** - channel actuales funcionan bajo dos arquitecturas básicas, FC-AL (Fibre Channel Arbitrated Loop) y Switched Fabric, ambos esquemas pueden convivir y ampliar las posibilidades de las SAN. La arquitectura FC-AL puede conectar hasta 127 dispositivos, mientras que switched fabric hasta 16 millones teóricamente.

---

<sup>26</sup> Un HBA (Host Bus Adapter) o adaptador de bus de host, es la tarjeta de interfaz que conecta un host a una SAN (Storage Area Network).

## 2.8 Balanceadores de carga.

### 2.8.1 Definición

El balanceador de carga es el dispositivo de red que se ubica entre los clientes y los servidores (figura 2.7), centralizando la recepción de peticiones. El balanceador de carga se limitará a reenviar las peticiones a los servidores reales, que son los encargados de procesarlas, incrementando de esta manera la capacidad de procesamiento de peticiones que comúnmente se tiene en arquitecturas de un solo servidor. El balanceador de carga toma las decisiones de reenvío de peticiones en función de un algoritmo de scheduling determinado. El método de reenvío puede implementarse de varias formas, en función de la arquitectura de red disponible para comunicar a los servidores reales con el balanceador de carga.

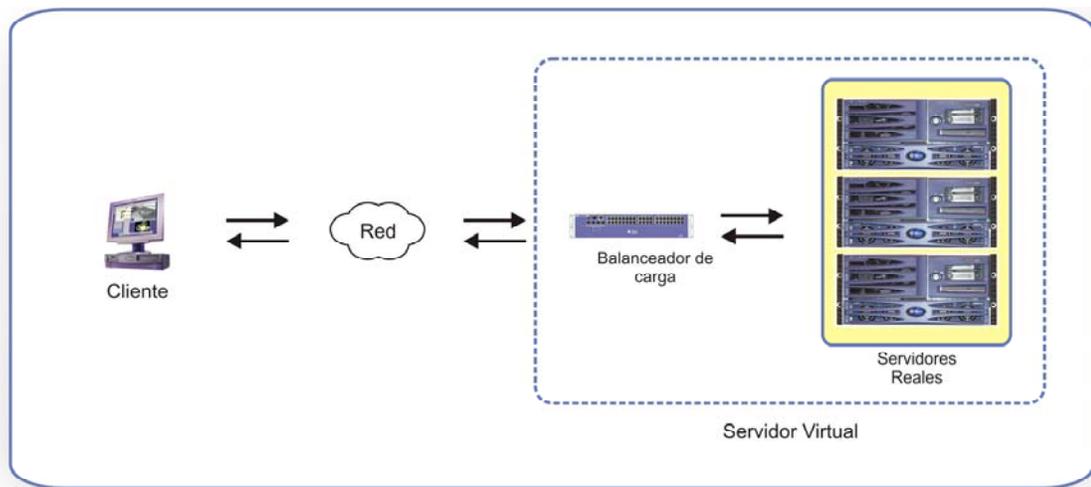


Figura 2.7 Uso de balanceador de carga,

### 2.8.2 Mejoras que aporta el balanceador de carga

El hecho de distribuir el trabajo entre diferentes máquinas permite que la capacidad de carga del servidor virtual sea muy superior a la que tendría un solo servidor real. Además, el balanceador de carga aporta escalabilidad, disponibilidad, facilidad de mantenimiento, seguridad y calidad de servicio.

### **2.8.2.1 Escalabilidad.**

Si la demanda de un servicio es muy elevada, se puede incrementar el número de servidores reales que prestan dicho servicio, de esta manera, con sólo aumentar el número de servidores reales, se aumenta el número de conexiones que puede atender el servidor virtual.

El balanceador de carga distribuye las peticiones de los clientes entre todos los servidores reales disponibles, usando para ello algoritmos de distribución de la carga, que permiten aumentar la capacidad de proceso que tendría un solo servidor. Suponiendo un balanceo perfecto, la capacidad del servidor virtual equivaldría a la suma de las capacidades de cada uno de los servidores reales por separado, pero en términos reales la capacidad del servidor virtual ronda entre el 80% y el 90%.

### **2.8.2.2 Disponibilidad.**

La máquina encargada de balancear la carga hace un chequeo continuo de los servidores reales así como de las aplicaciones que corren en ellos. En caso de que un servidor real o una aplicación dejen de responder, el balanceador ya no le reenviará más peticiones hasta que el servicio esté disponible de nuevo. De esta forma, de cara al usuario final (cliente), el servicio ofrecido por el servidor virtual estará siempre disponible, incluso en el caso de que alguno de los servidores reales deje de prestar servicio. Si el sistema no dispone de los mecanismos necesarios, las conexiones que en ese momento tenía el servidor real fallido se perderán. En este sentido, el balanceador de carga hace automáticamente y con transparencia lo mismo que se haría manualmente, mediante configuraciones estáticas del balanceador de carga, pero disminuyendo notablemente los tiempos de respuesta porque no se requiere de intervención humana para reconfigurar el balanceador.

### **2.8.2.3 Gestión y flexibilidad en el mantenimiento.**

Si se necesita apagar algún servidor real o detener un servicio, debido a un cambio de hardware o a una actualización del software, el uso de un balanceador de carga permite hacerlo de forma controlada, en el sentido de que se dejan de reenviar las peticiones de los clientes a dicho servidor. De esta forma, es posible proceder a realizar las tareas de mantenimiento sin afectar a la disponibilidad del servicio. El proceso es totalmente transparente para los clientes porque siempre encontrarán disponibilidad del servicio en la IP habitual (VIP).

El balanceador de carga aporta flexibilidad en la gestión del servidor virtual porque posibilita cambiar los servicios que está ofreciendo cada uno de los servidores reales en función de las necesidades de cada momento.

#### **2.8.2.4 Seguridad.**

Los balanceadores de carga se encuentran entre los clientes y los servidores reales, pudiendo actuar como filtro ante posibles ataques. Adicionalmente, el uso de NAT incrementa la seguridad debido al hecho de que las direcciones IP privadas de los servidores reales no son alcanzables directamente desde Internet. Los clientes únicamente ven la VIP que proporciona el balanceador, en cuyo interior se aplicarán las reglas de filtrado.

#### **2.8.2.5 Calidad de servicio (QoS).**

Los balanceadores de carga pueden ser usados para distinguir diferentes tipos de clientes y servicios, permitiendo de este modo la implementación de políticas de calidad de servicio (QoS). Los clientes o servicios preferentes pueden disponer de: servicios ofrecidos por un servidor en particular, mayor prioridad de sus paquetes IP y una determinada calidad de servicio o ancho de banda.

### **2.8.3 Tipos de balanceadores de carga.**

Los balanceadores de carga se pueden clasificar dentro de 4 grandes grupos: los balanceadores de carga para servidores (server load balancing), los balanceadores de carga para firewalls (firewall load balancing), los balanceadores de caches (cache load balancing) y los balanceadores de carga entre sistemas servidores remotos (global server load balancing).

#### **2.8.3.1 Server Load Balancing.**

Reparten la carga entre un grupo de servidores aumentando la capacidad del sistema. Solucionan el problema de la caída de un servidor y aumentan la escalabilidad del sistema.

### **2.8.3.2 Firewall Load Balancing.**

Reparten la carga entre un grupo de firewalls. Aumentan la capacidad y escalabilidad del sistema.

### **2.8.3.3 Cache Load Balancing.**

Descargan de trabajo a los servidores web al redireccionar las peticiones a sistemas de cache.

### **2.8.3.4 Global Server Load Balancing.**

Permiten redireccionar las peticiones de los clientes hacia el servidor o centro de datos más idóneo para estos, mejorando los tiempos de respuesta y la calidad del servicio. La selección del servidor idóneo puede realizarse en función de la distancia (un servidor por país o región), o en base a otros factores como el tráfico, calidad del enlace, etc.

## **2.8.4 Métodos de distribución de carga**

Por método de distribución de carga, se entiende la manera en la que el balanceador toma la decisión de a qué servidor, dentro de la granja de servidores, asignar una conexión.

El balanceo de carga puede realizarse de dos formas: **stateless** o **stateful**. Un balanceo de carga que utilice un algoritmo que simplemente distribuya el tráfico a un determinado servidor, sin tener en cuenta el estado de la información de cada sesión, es denominado stateless load balancing. Por el contrario, el balanceador que tiene en cuenta el estado de cada una de las sesiones, para tomar la decisión de a qué servidor dirigir las, se denomina stateful load balancing.

Independientemente de si el balanceador es stateful o stateless, éste utilizará un método de distribución que le ha de permitir determinar qué cantidad máxima de tráfico puede asignar a cada servidor. El objetivo, es realizar un reparto equitativo del tráfico en función de las capacidades de cada servidor porque es posible que dentro de la granja de servidores no todos los equipos tengan las mismas prestaciones.

Se puede clasificar el tráfico de red como TCP, UDP o simplemente IP. Un balanceador de carga capaz de identificar el tráfico como TCP o UDP, podrá también identificar las sesiones TCP o UDP en base a los puertos y direcciones IP origen/destino. El protocolo TCP es orientado a conexión. Las sesiones se establecen y finalizan mediante el intercambio de una serie de órdenes predeterminadas por el protocolo TCP. El balanceador puede entender o no este diálogo y utilizarlo para determinar el inicio y finalización de sesiones.

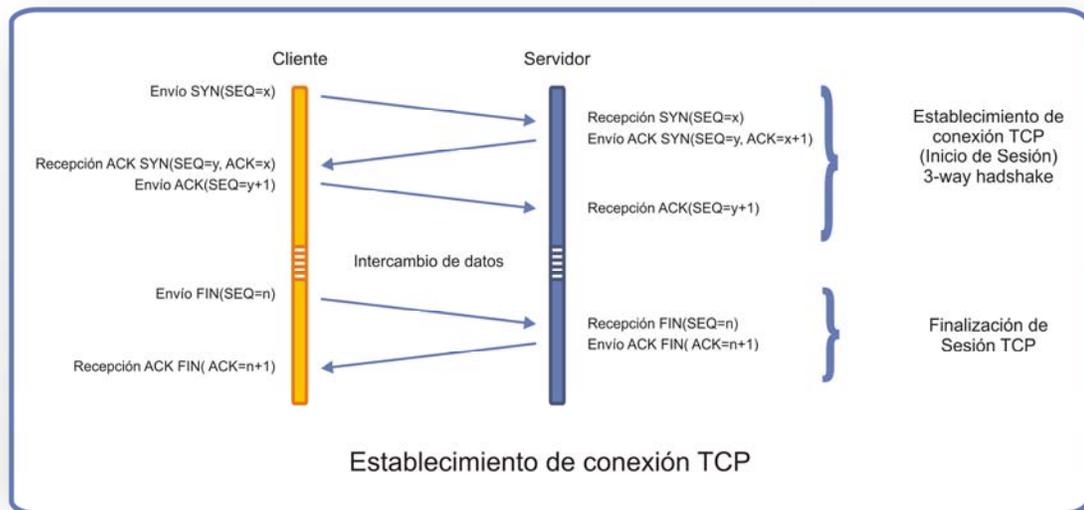


Figura 2.8. Establecimiento de conexión TCP

Tanto el tráfico TCP como UDP funciona sobre IP. Por el contrario, algunos protocolos propietarios sólo pueden correr directamente sobre IP, en estos casos, el balanceador sólo podrá determinar las sesiones en base a las direcciones IP origen y destino.

#### 2.8.4.1 Stateless

Un balanceador de tipo stateless no puede determinar el inicio y fin de sesiones. UDP es un protocolo stateless, dado que no es orientado a conexión. Por el contrario, existen determinados servicios que requieren que todos los paquetes pertenecientes a una sesión UDP, sean enviados al mismo servidor real. Por ejemplo, en aquellos servicios en los que se requiere que se establezcan dos conexiones, una TCP y otra UDP, es requisito indispensable que los paquetes UDP se envíen siempre al mismo servidor real en el que se ha establecido la conexión TCP.

En estos casos, dado que el balanceador identifica la sesión UDP únicamente en base a las direcciones IP y puertos origen/destino, el balanceador crea una regla, que lo configura dinámicamente para reenviar siempre al mismo servidor real los paquetes pertenecientes a esa sesión. Esta regla desaparece automáticamente, si durante un periodo predeterminado de tiempo no se recibe otro paquete que case con ella.

La principal ventaja del balanceador de tipo stateless es la simplicidad de los algoritmos de scheduling utilizados para tomar la decisión de a qué servidor reenviar los paquetes. El hecho de que el balanceador stateless tenga menos elementos en los que basarse para la toma de decisiones de reenvío, repercute en un algoritmo de scheduling más sencillo y por tanto en menores requerimientos para su procesado.

#### **2.8.4.2 Stateful**

El balanceo de carga stateful requiere que el balanceador sepa reconocer y entender el inicio, finalización y diálogo de las sesiones. De esta manera, cuando le llega una petición de conexión, el balanceador crea una regla que lo configura dinámicamente para reenviar todos los paquetes pertenecientes a esa sesión hacia un mismo servidor real. Del mismo modo, cuando el balanceador detecte el fin de esa sesión, borrará la regla que previamente le creó. En el caso de que la sesión no finalice y por un período de tiempo máximo predeterminado, no se reciban más paquetes pertenecientes a la misma, el balanceador también borrará la regla.

El balanceo de carga de tipo stateful es mucho más eficiente que el balanceo de carga stateless, debido a que las decisiones de balanceo son realizadas también en base al estado de la sesión y no sólo en base a su identificación. Esta importante mejora en el balanceo es posible gracias a que los algoritmos de scheduling disponen de más elementos de juicio para la toma de decisiones y que permiten un mejor reparto de las sesiones.

El balanceador stateless considera por igual todas las sesiones pertenecientes a una misma IP, en cambio, el balanceador stateful es capaz de diferenciar la carga de las diferentes sesiones.

Para lograr una mayor comprensión de lo anteriormente expuesto, considérese el siguiente escenario:

Dos clientes efectúan peticiones a un mismo servidor virtual. El primero de ellos, genera 100 paquetes por segundo (60 de http, 30 de FTP y 10 de IRC) hacia el servidor virtual. El segundo cliente sólo genera 5 paquetes (3 de http y 2 de FTP).

Un balanceador stateless considera que los 100 paquetes generados por el primer cliente forman parte de una misma sesión, y por tanto, reenvía los 100 a un mismo servidor real.

Los 5 paquetes del segundo cliente, los considera pertenecientes a otra sesión, por lo que los reenvía a otro servidor real.

Un balanceador stateful es capaz de diferenciar de entre los 100 paquetes enviados por el primer cliente, a qué servicios/sesión pertenecen cada uno de ellos, y por tanto puede repartir, por ejemplo, los 60 paquetes de http al servidor 1, los 30 de FTP al servidor 2 y los 10 de IRC al servidor 3. Dependiendo de las capacidades stateful del balanceador en particular, éste incluso podría llegar a diferenciar entre distintas peticiones http, en función, por ejemplo, de la página solicitada.

Un balanceador stateful identificará cada una de las sesiones de cada cliente y realizará un balanceo de carga mucho más eficiente. Como contrapartida, los balanceadores de carga de tipo stateful requieren de una mayor capacidad de proceso.

### **2.8.4.3 Algoritmos de Scheduling (Hashing)**

Los balanceadores de carga pueden utilizar diferentes métodos para decidir a qué servidor asignar las diferentes conexiones. Algunos de estos métodos son más efectivos que otros, en función del servicio que se ha de balancear. Es por esto, que un balanceador de carga puede utilizar diferentes algoritmos de scheduling para balancear diferentes tipos de servicios.

#### **Secuencial (Round-Robin):**

Es el más sencillo por lo que consume muy pocos recursos.  
Realiza una asignación secuencial y cíclica de los servidores.  
Es poco eficiente porque no toma en cuenta el número de conexiones establecidas en cada servidor.

**De menor número de conexiones (Least Connections):**

El balanceador conoce el número de sesiones establecidas en cada servidor real y reenvía las conexiones nuevas al servidor que menos conexiones tiene en ese momento.

**Distribución Ponderada por pesos (Weighted Distribution):**

Este método permite al administrador asignar un peso a cada servidor, en función de su capacidad y el servicio que ha de prestar.

Se utilizan en combinación con Round-Robin y Least Connections.

**Por tiempo de respuesta (Response Time):**

Complementa a otros algoritmos posibilitando la elección del servidor más rápido.

## 2.9 Network Address Translation (NAT)

La 'Traducción de Direcciones de Red', Network Address Translation (NAT), es un método mediante el cual las direcciones IP son mapeadas desde un dominio de direcciones a otro, proporcionando encaminamiento transparente a las máquinas finales. Existen muchas variantes de traducción de direcciones que se prestan a distintas aplicaciones. Sin embargo, todas las variantes de dispositivos NAT deberían compartir las siguientes características:

- a) Asignación transparente de direcciones.
- b) Encaminamiento transparente mediante la traducción de direcciones (aquí el encaminamiento se refiere al reenvío de paquetes, no al intercambio de información de encaminamiento).
- c) Traducción de la carga útil de los paquetes de error ICMP.

Este apartado se centra en algunos tipos de NAT utilizados en la implementación de balanceadores de carga.

### 2.9.1 Destination NAT

Consiste en cambiar la dirección destino, de la cabecera IP, de los paquetes enviados por el cliente. Esta dirección se sobrescribe en este caso con la dirección IP del servidor real seleccionado por el balanceador de carga. Cuando el paquete retorna al cliente, se restituye esta dirección para que figure como origen. De esta manera, el paso a través del balanceador de carga, se realiza de forma transparente para el cliente. A este tipo de NAT se le conoce también como half-NAT dado que sólo se cambia la dirección destino del paquete.

### 2.9.2 Source NAT

Este tipo de NAT es poco frecuente y sólo se utiliza en situaciones donde el diseño de la red así lo requiere. Esto es, en situaciones donde la topología de la red permite a los paquetes volver al cliente sin pasar por el balanceador de carga.

En determinados diseños de red, pueden existir otras alternativas al source NAT, éstas pasan por utilizar Direct Server Return (DSR) o bien por poner al balanceador como gateway por defecto de los servidores reales. Estas dos alternativas requieren que el balanceador de carga y los servidores reales estén en el mismo dominio de broadcast o capa de enlace.

Source NAT se realiza cambiando las direcciones origen y destino de los paquetes. El balanceador de carga cambia la dirección origen de todos los paquetes por una dirección definida en el balanceador, llamada source IP, antes de reenviar los paquetes a los servidores reales. Cuando los paquetes llegan a los servidores reales, estos últimos los ven como peticiones realizadas por el balanceador de carga, puesto que los servidores no tienen conocimiento de la dirección de los auténticos clientes finales que originaron dichos paquetes. Las respuestas de los servidores reales tienen como dirección destino la IP del balanceador de carga y no la del cliente final. El balanceador recoge estas respuestas y rescribe nuevamente la cabecera IP de manera que la dirección de destino sea la IP del cliente y la de origen la source IP.

Desde el punto de vista del balanceador de carga se establecen dos sesiones. Una con el cliente y otra con el servidor real. Los servidores reales ven al balanceador de carga como un cliente, lo cual implica que el balanceador de carga ha de disponer de un número de puerto para cada una de las sesiones que establece con cada uno de los servidores reales.

El número de puertos máximos que puede abrir un sistema es de 65.536. Existe por lo tanto una importante limitación en cuanto al número de máximo de conexiones simultáneas que el balanceador de carga puede manejar para una sola IP. Para superar esta limitación el balanceador ha de permitir configurar múltiples direcciones source IP.

La principal ventaja de source NAT es permitir la implementación de balanceadores de carga independientemente de la topología de la red, el inconveniente es que los servidores reales no ven la dirección IP original del cliente. Por tanto, aquellas aplicaciones que utilicen algún método de autenticación basado en IP no funcionarán apropiadamente con source NAT.

## **3 Capítulo 3**

---

Análisis de la situación tecnológica y ejecución de procesos operativos inicial en la prestación de servicios de valor agregado.

### 3.1 Análisis de la situación tecnológica inicial.

A continuación se describe el diseño de la arquitectura inicial para proporcionar algunos servicios de valor agregado tanto masivos como corporativos.

#### 3.1.1 Diseño inicial de alta disponibilidad a nivel HW y SW.

Inicialmente se tenía una configuración en alta disponibilidad a nivel hardware y software diseñada de la siguiente manera:

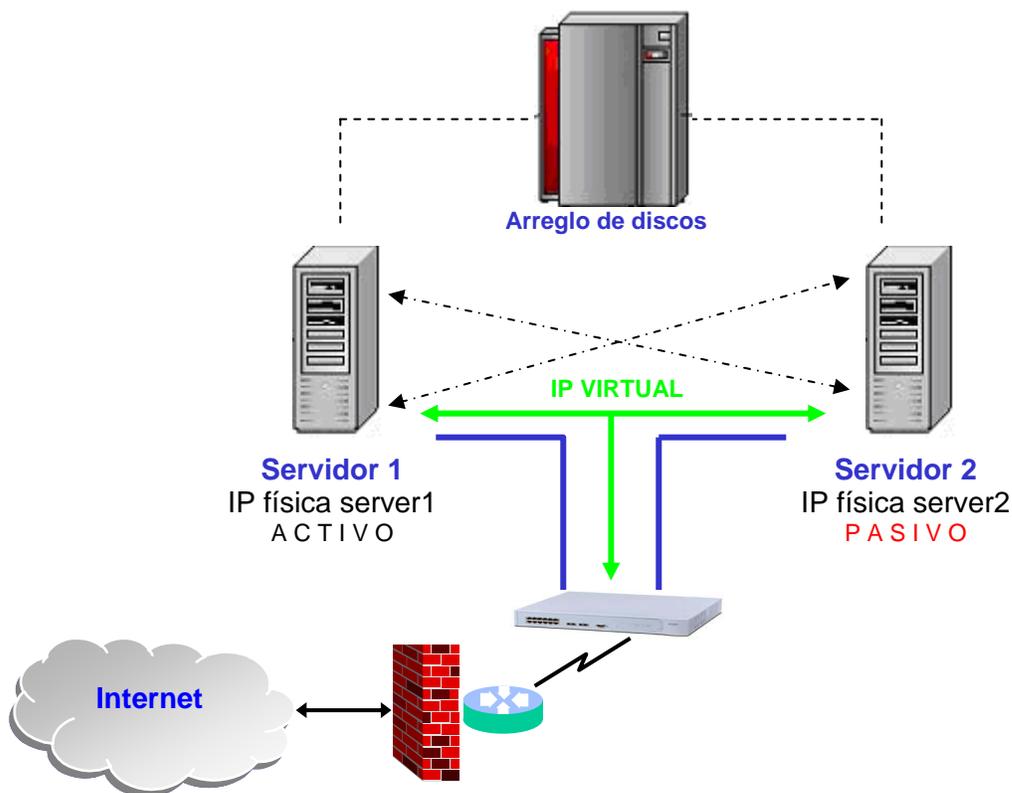


Figura 3.1 Arquitectura inicial del sistema SVA

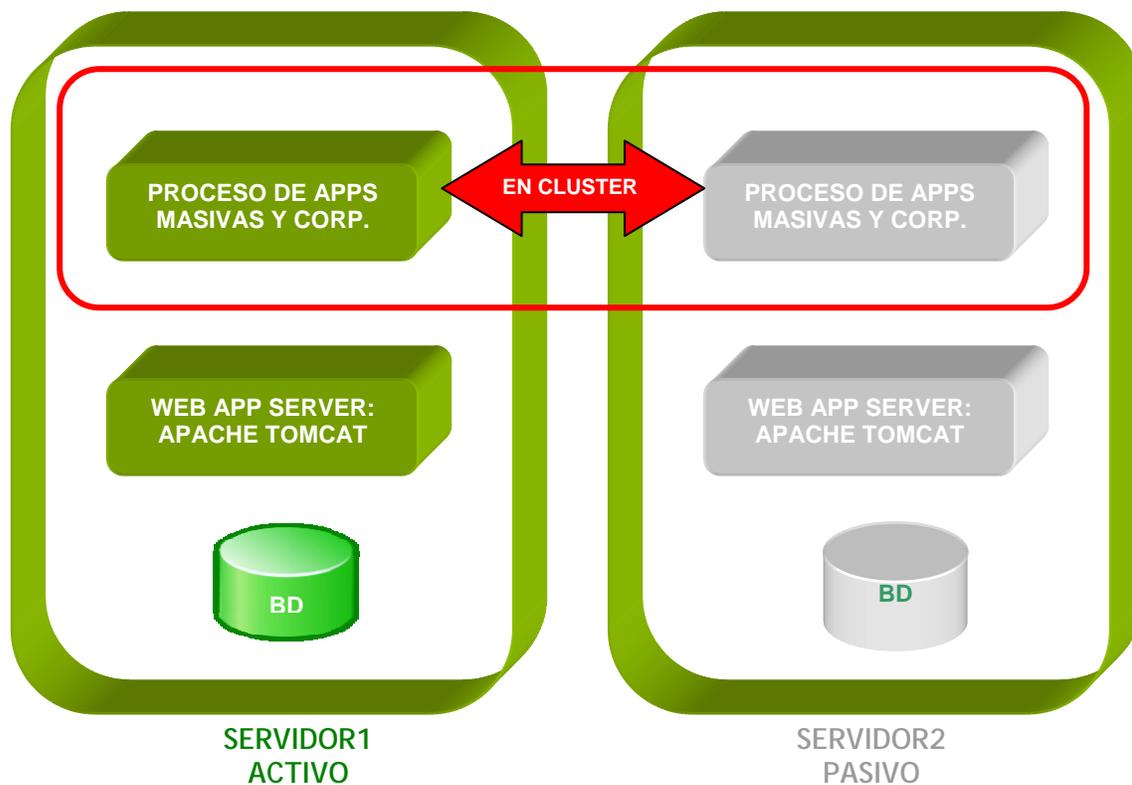


Figura 3.2 Distribución de procesos aplicativos y de base de datos de la arquitectura inicial.

### 3.1.2 Configuración Activo & Pasivo.

Este tipo de configuración permitía mantener la alta disponibilidad a nivel hardware y software del servicio con las siguientes ventajas y desventajas:

#### 3.1.2.1 Ventajas:

- ◆ Existía redundancia a nivel hardware con las mismas capacidades. Esto es, si por alguna razón, algunos de los componentes del hardware del servidor activo se dañaba, automáticamente los procesos aplicativos y de base de datos que se estaban ejecutando en el nodo activo, se levantaban en el nodo pasivo.

#### 3.1.2.2 Desventajas:

- ◆ Se desaprovechaban los recursos del equipo que no se estaba utilizando como nodo activo.
- ◆ No había distribución de tráfico, por lo que todas las peticiones que entraban de los clientes eran procesadas por el servidor activo provocando así, la saturación de sus recursos. (Ej.: cpu & memoria)
- ◆ No se tenía escalabilidad de la arquitectura. Esto es, dada la configuración de red y del cluster, no se podía incrementar el número de servidores activos que prestaban dicho servicio de manera transparente al usuario final.
- ◆ La disponibilidad del *sistema SVA* a nivel procesos aplicativos no era del 100% por dos principales razones:
  - Sólo existía un grupo definido en el cluster para el control de todos los aplicativos. Esto fue una gran desventaja dado que si un proceso aplicativo se corrompía en el nodo activo, este provocaba el failover de todo el grupo y por consiguiente la caída “momentánea” de todos los procesos aplicativos que en ese momento se encontraban trabajando correctamente.

- La base de datos que se tenía para el funcionamiento de los servicios, no se encontraba bajo la operación del cluster. Por lo anterior, no se tenía alta disponibilidad de la misma.
- ◆ La seguridad del sistema era mínima. Esto debido a que el direccionamiento IP virtual del cluster estaba homologado a direcciones publicas con salida a internet, fácil de visualizar desde el exterior.

### 3.1.3 Application Server de dominio público.

Dado el tipo de servicio que reside en esta arquitectura, es necesario contar con un application server que permita la publicación y el correcto funcionamiento de las aplicaciones desarrolladas para dicho servicio. Las características requeridas de este application server son:

- ◆ Contar con respaldo de soporte técnico en línea, las 24 horas del día y los 365 días del año (7X24).
- ◆ Configuración en alta disponibilidad “cluster”
- ◆ Cumplir con el estándar j2ee 1.3 y j2se 1.4+.<sup>27</sup>
- ◆ Configuración y administración sencilla y centralizada.
- ◆ Administración local y remota.
- ◆ Contar con herramientas de tuning y monitoreo en caso de fallas.
- ◆ Compatible con diferentes sistemas operativos, tales como, Solaris, windows NT, etc.
- ◆ Garantía del funcionamiento del producto y capacitación.

El tipo de application server utilizado inicialmente era de dominio público por lo que se tenían los siguientes puntos de falla:

- ◆ No se tenía una garantía de compatibilidad entre el desarrollo de las aplicaciones y las versiones del application server existentes.

---

<sup>27</sup> J2SE 1.4 es prácticamente la versión 1.4 de la maquina virtual de java sobre la que corre el servidor de aplicaciones que implementa el estándar J2EE. En resumen, es la maquina virtual y sus APIs para programar.

- ◆ Sólo se tenía un application server para todos los servicios aplicativos, por lo que si éste fallaba, provocaba la pérdida de todos los servicios.
- ◆ No contaba con herramientas de tuning y monitoreo en línea en caso de falla, lo que provocaba que los tiempos de resolución a las fallas presentadas fueran muy elevados.
- ◆ Al ser de dominio público, no se contaba con un soporte especializado en la resolución de problemas presentados dada la convivencia con el desarrollo de las aplicaciones montadas en el mismo. Por lo anterior, tampoco se tenía un representante en línea del producto a quién escalar y obtener la solución a las fallas presentadas.
- ◆ El tipo de application server instalado inicialmente afectó de manera considerable el servicio, debido a la caída constante de su proceso, que resultaba difícil el logro de nuestro primer objetivo: **“Alcanzar y mantener la disponibilidad semanal del sistema SVA al 99.9%”**.

A continuación se muestra una estadística con los reinicios constantes que se presentaban en este elemento, los cuales afectaban severamente el servicio:

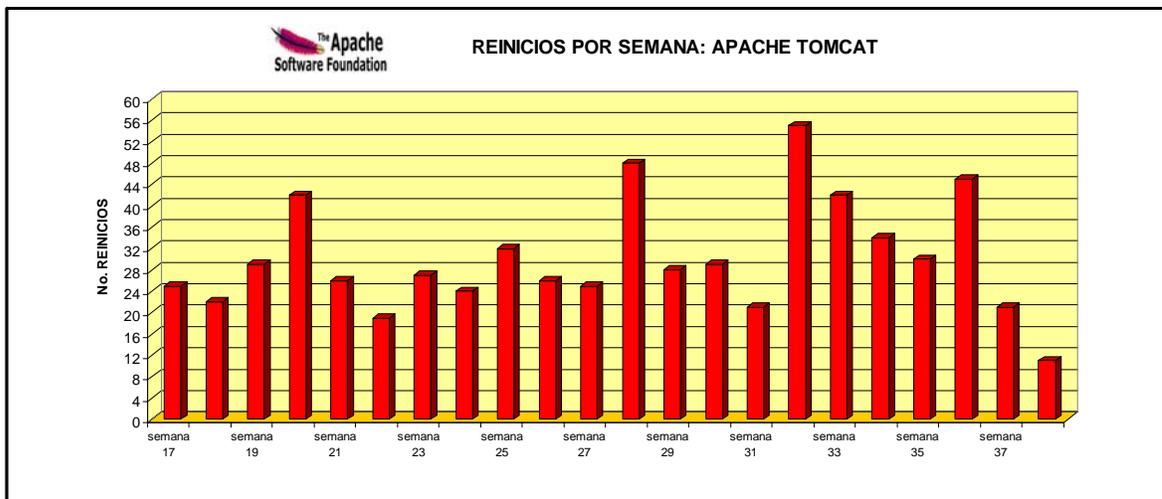


Figura 3.3 Reinicios por semana: Apache Tomcat.

### **3.1.4 Configuración de Base de Datos.**

Dado el dimensionamiento inicial del servicio, se decidió tener una sola base de datos para todas las aplicaciones. Los puntos de falla fueron:

- ◆ No se tenía una configuración de alta disponibilidad en la base de datos. Por lo anterior, si ocurría alguna falla de operación en la base de datos, se impactaban los servicios masivos y corporativos.
  
- ◆ Al momento de realizar mantenimiento a la base de datos, también se afectaba en su totalidad todos los servicios masivos y corporativos.

### **3.1.5 Procesos aplicativos no distribuidos.**

La solución a nivel desarrollo o software aplicativo para el funcionamiento de los servicios estaba compartido, es decir, se utilizaba un solo proceso aplicativo para soportar todo el tráfico de servicios masivos y corporativos.

Lo anterior, no permitía la independencia de los servicios de tal manera que se presentaban los siguientes puntos falla:

- ◆ Solo se tenía un proceso aplicativo para todos los servicios masivos y corporativos; por tanto, si este proceso fallaba, provocaba la pérdida de todos los servicios.
  
- ◆ Al momento de realizar mantenimiento o actualización a un proceso aplicativo, se afectaba en su totalidad a todos los servicios que dependían de dicho proceso. Aún cuando dichos mantenimientos o actualizaciones no se solicitarán para todos los servicios.

### **3.1.6 Monitoreo del sistema inicial.**

En un área de operación es muy importante contar con herramientas de monitoreo y notificación automática que permitan la identificación oportuna de las fallas que se presentan en los sistemas.

Inicialmente, el monitoreo del *sistema de SVA*, se realizaba manualmente debido a que no se tenía un sistema de monitoreo automatizado. Lo anterior, provocaba que no se identificaran a tiempo las fallas ocurridas y que el tiempo de atención y recuperación de una falla se prolongara por varias horas. Esto impactó directamente el servicio que fue imposible alcanzar la disponibilidad del 99.9% que la empresa requería ofrecer al mercado.

Adicional, la operación y mantenimiento de este sistema se llevaba a cabo por un solo administrador impactando también en la disponibilidad del servicio las 24 horas del día los 365 días del año (7X24).

### **3.1.7 Sistema de respaldos inicial.**

Los respaldos de la información del *sistema SVA*, se realizaban a través de las cartucheras internas de los servidores. Esta solución de respaldos no era eficiente por las siguientes razones:

- ◆ El cambio y gestión de cintas de respaldo se realizaba de forma manual.
- ◆ Se utilizaban cintas DDS-120 para respaldar la información. Esto implicaba asignar a una persona responsable de realizar los cambios de cinta y ejecutar el script correspondiente de cada respaldo por un periodo aproximado de 6 horas.
- ◆ Otra desventaja de este sistema de respaldo, es que para realizar un respaldo confiable de la aplicación, se requería dar de baja los procesos aplicativos y de base de datos. Esto implicaba suspender el servicio por un periodo de 6 horas a la semana.

- ◆ Dado que era un requerimiento brindar el servicio de manera continúa el mayor tiempo posible, sólo se realizaba un backup a la semana. Con lo anterior, no se tenía la posibilidad de recuperar hasta la última transacción en caso de desastre o pérdida de información.
- ◆ No se contaba con una solución de respaldo y recuperación de desastres para base de datos.
- ◆ En una situación de desastres, la recuperación de los datos era lenta y complicada en su gestión.

### 3.1.8 Disponibilidad inicial del sistema SVA.

Dado el diseño de la arquitectura inicial, no se tenía una disponibilidad del 99.9%. La disponibilidad promedio del *sistema* SVA estaba alrededor del 93.6%.

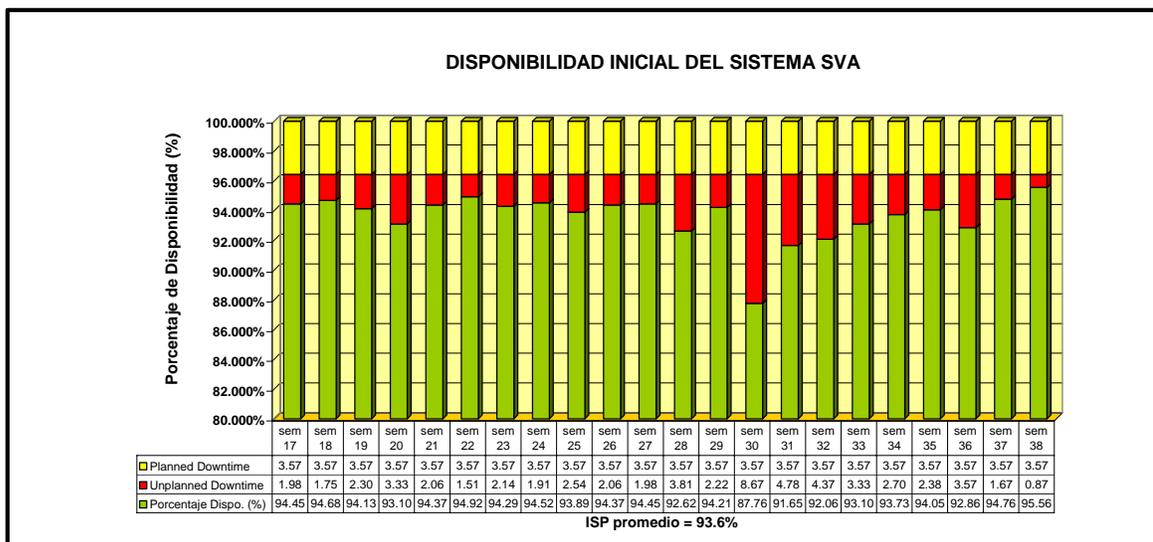


Figura 3.4 Disponibilidad inicial del sistema SVA.

Por todo lo anterior, y dado a que el éxito de los servicios de valor agregado en el mercado incrementó de manera considerable, fue importante aplicar de inmediato las mejoras necesarias a la arquitectura en cuestión para alojar más servicios.

### 3.2 Ejecución inicial de procesos operativos en los sistemas SVA.

Inicialmente la operación y mantenimiento de los sistemas dedicados a los servicios de valor agregado, eran responsabilidad de un solo departamento conformado por un grupo de tres personas. Por lo anterior, toda la operación y mantenimiento de dichos sistemas, se llevaba a cabo por los integrantes de este departamento quienes debían estar disponibles los 365 días del año las 24 horas del día para dar soporte y mantenimiento a los equipos y con esto proporcionar el servicio de manera ininterrumpida a los usuarios finales.

Lo anterior, no era eficiente porque aunado a las tareas de mantenimiento y operación de estos sistemas, el mismo grupo de trabajo, tenía que estar validando, coordinando y ejecutando en horarios de madrugada, las solicitudes de cambios y mejoras en los sistemas de SVA provenientes de un área técnico comercial.

A continuación se muestra el proceso y flujo de trabajo inicial para alta, baja o modificación en los sistemas de servicios de valor agregado:

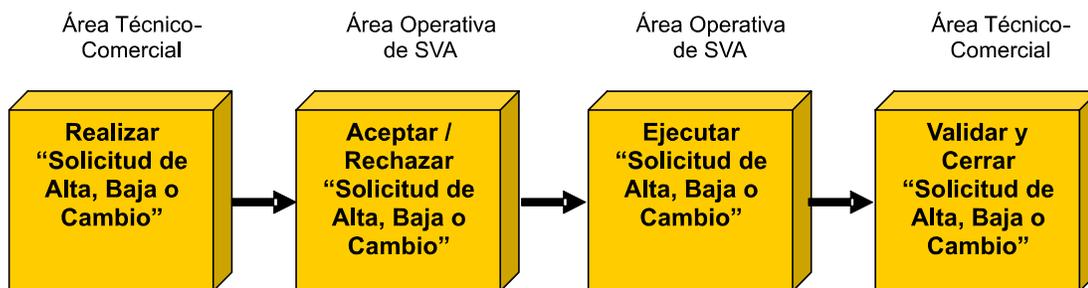


Figura 3.5 Proceso de operación inicial para alta, baja o cambio en los sistemas SVA en producción.

Como puede apreciarse en la figura anterior, el área técnico-comercial era el área responsable de solicitar las altas, bajas o cambios sobre los servicios de valor agregado ya existentes, mientras que el área de operación era responsable de la validación y ejecución de dichas solicitudes, además de proporcionar la operación y el mantenimiento preventivo y correctivo a los servicios en producción.

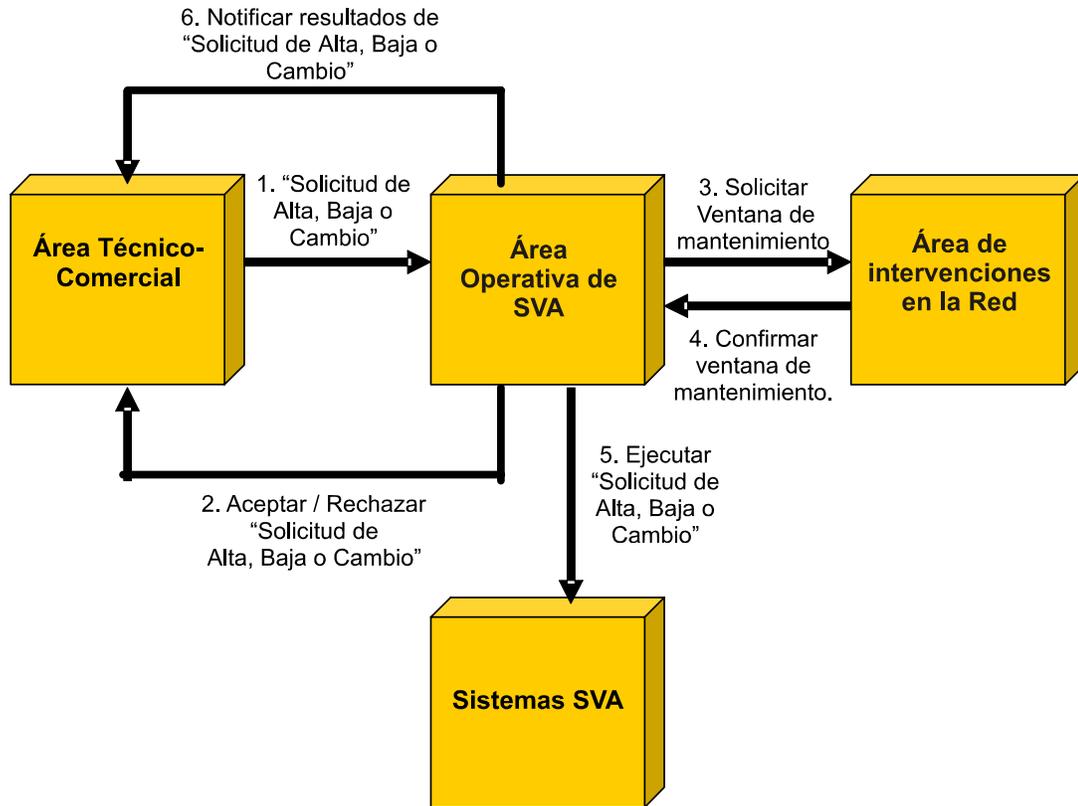


Figura 3.6 Flujo de trabajo inicial para ejecución de solicitudes de alta/baja/cambio en los sistemas SVA.

Dado lo anterior, los problemas e impactos comunes en este proceso de operación inicial fueron:

- ◆ Retardo en la ejecución de solicitudes de alta, baja o modificación de configuración en los sistemas de SVA dada la dependencia de una sola área operativa y la demanda constante de cambios y mejoras en los servicios de valor agregado.
- ◆ No se daba una atención y solución en tiempo real a las fallas presentadas en los sistemas impactando directamente en la disponibilidad del servicio.
- ◆ Falta de seguimiento en tiempo a los programas de mantenimiento preventivos en las sistemas.

- ◆ Bajo análisis de causa raíz a los problemas de operación presentados en las sistemas SVA.
- ◆ Bajo análisis de desempeño y disponibilidad de los sistemas SVA para garantizar la disponibilidad y capacidad de las mismas.

Aunado a lo anterior, no existía un área de ingeniería dedicada a la planeación y evaluación tecnológica de las nuevas soluciones de servicios de valor agregado. Por lo anterior, dicha labor era llevada a cabo también por el área técnico-comercial encargada del diseño y lanzamiento de nuevos servicios de valor agregado con base a las necesidades del mercado masivo y corporativo.

Lo anterior, tampoco era eficiente dada la creciente demanda de nuevos servicios de valor agregado y el crecimiento o expansión requerida en los ya existentes.

A continuación se muestra el proceso de integración inicial de nuevos sistemas de servicios de valor agregado:

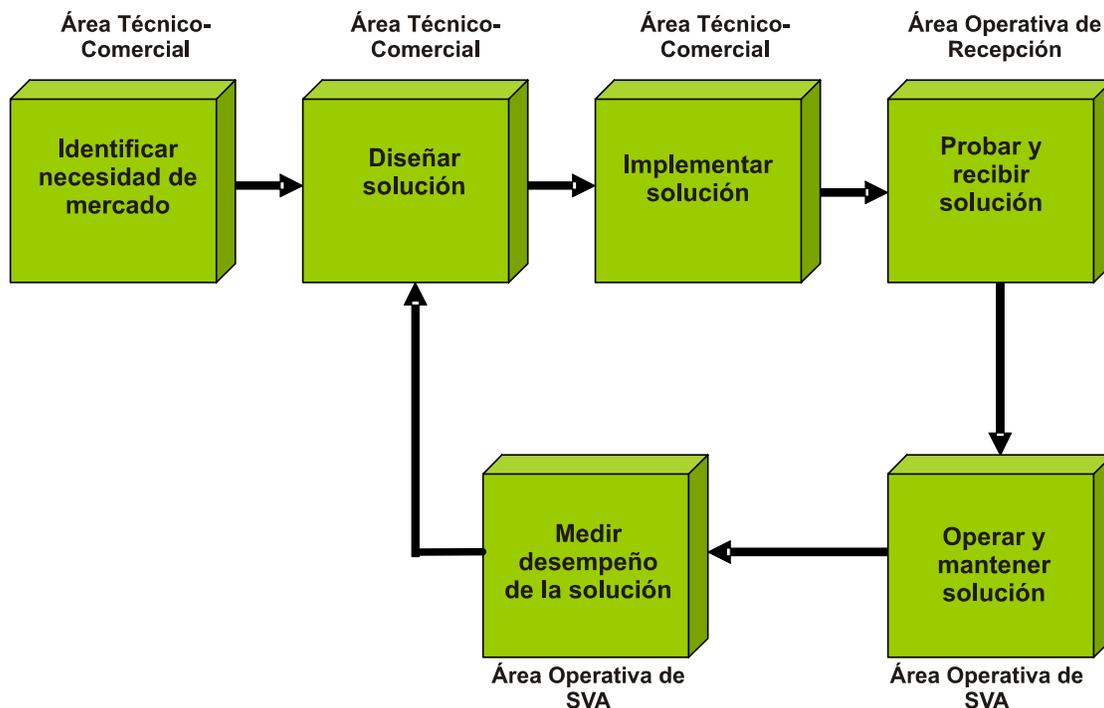


Figura 3.7 Proceso de integración inicial de nuevos sistemas de servicios de valor agregado.

Las principales deficiencias en este proceso de operación inicial fueron:

- ◆ No se consideraban en su totalidad todos los requerimientos del área operativa para la puesta en producción de nuevos sistemas de servicios de valor agregado. Por mencionar algunos de estos requisitos: los sistemas SVA no se integraban con soluciones de monitoreo y respaldo eficientes.
  
- ◆ Tiempos elevados en el proceso de análisis, diseño, implementación y lanzamiento de un nuevo servicio. Lo anterior, dado que existía una misma área realizando estas actividades y a la vez visualizando y diseñando nuevos servicios.

A medida que los servicios y sistemas de valor agregado aumentaron, se tuvieron que implementar mejoras administrativas para mejorar y mantener la operación de los sistemas y lograr así los objetivos de disponibilidad establecidos por la dirección del área, no solo para el sistema objeto de este trabajo sino para el resto de los sistemas dedicados a este tipo de servicios. Ver propuesta de mejora expuesta en el capítulo 5.

## **4 Capítulo 4**

---

Diseño e implementación de la actualización  
tecnológica.

Con el fin de solucionar los problemas de diseño que presentaba la arquitectura inicial descritos en el capítulo anterior y teniendo siempre presente el cumplimiento del primer objetivo: **“Alcanzar y mantener la disponibilidad semanal del sistema SVA al 99.9%”**, se decidió llevar a cabo la actualización tecnológica de dicha arquitectura en dos fases:

**Primera Fase:**

- ◆ Cambio de web application server en la arquitectura inicial.

Aumentar a corto plazo la disponibilidad semanal del *sistema SVA* del 93% al 96% implementando como acción correctiva emergente el cambio del web application server Apache Tomcat a BEA Weblogic Server.

**Segunda Fase:**

- ◆ Diseño e implementación de la nueva arquitectura.

Diseñar e implementar una nueva arquitectura que permitiera alcanzar y mantener la disponibilidad semanal del *sistema SVA* por arriba del 99.9% aun considerando el crecimiento de los servicios de valor agregado a corto (3 meses), mediano (6 meses) y largo plazo (1 año).

A continuación se explica de manera más detallada las actividades realizadas en cada una de las fases antes citadas.

## 4.1 Primera Fase. Cambio de web application server en la arquitectura inicial.

Como se describió en el capítulo 3, la baja disponibilidad del *sistema SVA* se debía principalmente al tipo de web application server utilizado: Apache Tomcat.

Por lo anterior, las acciones correctivas propuestas fueron:

- ◆ Evaluación y adquisición de un nuevo web application server.
- ◆ Cambio de web application server.

### 4.1.1 Evaluación de un nuevo web application server.

Los requerimientos principales de evaluación para la adquisición del nuevo application server fueron:

- ◆ Contar con respaldo de soporte técnico en línea, las 24 horas del día y los 365 días del año (7X24).
- ◆ Configuración en alta disponibilidad “cluster”
- ◆ Cumplir con el estándar j2ee 1.3 y j2se 1.4+.<sup>28</sup>
- ◆ Configuración y administración sencilla y centralizada.
- ◆ Administración local y remota.
- ◆ Contar con herramientas de tuning y monitoreo en caso de fallas.
- ◆ Compatible con diferentes sistemas operativos, tales como, Solaris, windows NT, etc.
- ◆ Garantía del funcionamiento del producto y capacitación.

---

<sup>28</sup> J2SE 1.4 es prácticamente la versión 1.4 de la maquina virtual de java sobre la que corre el servidor de aplicaciones que implementa el estándar J2EE. En resumen, es la maquina virtual y sus APIs para programar.

Tomando como referencia el web application server existente: Apache Tomcat, la evaluación se realizó entre BEA Weblogic Server y SUN Java App Server por ser los dos application servers más recomendados por nuestros desarrolladores de aplicación.

Así mismo, en esta evaluación se consideraron las opiniones críticas de diversos proveedores; así como, los resultados de encuestas que realizan corporaciones tales como: Evans Data Corporation (EDC).

Los criterios que se consideraron en la evaluación del nuevo web application server incluyeron puntos generales como costo, soporte y capacitación; así como características técnicas muy específicas que a continuación se describen:

#### **4.1.1.1 Escalabilidad.**

La escalabilidad se refiere a la habilidad de incrementar la carga de trabajo de una aplicación o sistema sin afectar significativamente el performance del mismo o sus funciones. Los beneficios de un application server altamente escalable son obvios en un ambiente en el cual existe un alto número de transacciones, usuarios, etc. Algunas veces el incremento de la escalabilidad implica hardware adicional, en este momento llega a ser importante balancear la carga a través de varias maquinas.

BEA Weblogic tiene mejor habilidad de escalabilidad en comparación con SUN Java App Server y Apache Tomcat.

#### **4.1.1.2 Performance.**

El performance<sup>29</sup> se refiere a la rapidez con la que un application server puede procesar los requerimientos de las transacciones. Obviamente el performance es afectado por el número y capacidad de las maquinas utilizadas.

En este rubro Apache Tomcat y BEA Weblogic tienen mejor performance que SUN Java App Server.

---

<sup>29</sup> Entiéndase por performance el rendimiento de una maquina o software.

#### **4.1.1.3 Compatibilidad con diferentes sistemas operativos.**

Esto se refiere al número de plataformas sobre las cuales puede ejecutarse un application server. Esta característica es importante porque no es necesario cambiar de web application server cuando se requiera hacer una migración de hardware y/o sistema operativo.

Algunos application server como BEA Weblogic y SUN Java App Server pueden ejecutarse sobre varias plataformas, tales como: Windows, Solaris, HP-UX, y Red Hat Linux. Sin embargo, esta característica es muy fuerte en software libre como: Apache Tomcat, porque ha sido desarrollado para ser compatible con la mayoría de las plataformas existentes.

#### **4.1.1.4 Balanceo de carga.**

Los balanceadores de carga representan una parte integral de la infraestructura web dado que estos distribuyen el tráfico a través de la red de modo tal que un servidor no pueda saturarse.

El application server BEA Weblogic, tiene contemplada la característica de balanceo de carga a través de su funcionalidad de clustering. El cual consiste en asociar varias maquinas bajo el nombre de un mismo servidor web.

#### **4.1.1.5 Integración con dispositivos móviles.**

BEA Weblogic es uno de los mejores application servers que provee a sus usuarios con herramientas “toolkits”, información y soporte para la creación y desarrollo de aplicaciones que cubre un amplio rango de productos móviles.

#### **4.1.1.6 Enterprise application integration (EAI) y J2EE.**

Enterprise application integration (EAI) consiste en la modernización de las aplicaciones viejas. El término también se refiere a la consolidación de datos y mensajes de diferentes sistemas, obteniendo de estos un trabajo conjunto en un ambiente homogéneo. Los estándares y fabricantes de application servers asisten en esta pesada tarea trabajando en proveer una infraestructura común de e-business.

EAI, se le llama así al plan de integrar sistemas o automatizar la integración de sistemas en una empresa, muy frecuentemente la gente se refiere como EAI al producto que les permite integrar los sistemas de su empresa, por ejemplo WebLogic Integration puede ser un EAI, es un producto diseñado para resolver las diferencias de comunicación, protocolo, naturaleza (síncrono y asíncrono) y formato entre sistemas, haciendo posible que distintos sistemas se “comuniquen” o compartan información entre si, de forma muchas veces automática.

J2EE, es el estándar que determina lo que un servidor de aplicaciones java debe ofrecer como mínimo a través de un conjunto de especificaciones. Algunas de las especificaciones que define este estándar son: JDBC's, servlets, EJB, etc.

Cada nueva versión de J2EE implica cambios en algunas de las especificaciones: JDBC, Servlets, EJB, etc. Dichas especificaciones también están versionadas y con cada lanzamiento de una nueva versión J2EE vienen incluidas nuevas versiones de estas especificaciones; no todas las especificaciones suben de versión cuando hay un nuevo J2EE, pero si algunas.

BEA WebLogic figura fuertemente en el trabajo de soporte e integración de ambientes de desarrollo homogéneos a través del estándar J2EE Connector Architecture framework.

BEA Weblogic Server esta certificado bajo el estándar J2EE dado que cumple con las especificaciones de dicho estándar.

#### **4.1.1.7 Seguridad.**

Aspectos de seguridad e infraestructura son requeridos por cualquier application server, dada su integración con base de datos y a que existen sistemas haciendo intrusiones extremadamente problemáticas. El comercio electrónico “e-commerce” también demanda una alta seguridad. Soluciones típicas incluyen SSL /TLS\*, administración de identidad y soporte para Web Services.

#### **4.1.1.8 Supervisión de Cambios.**

Las capacidades de supervisión de cambios en un application server, permiten a los usuarios rastrear las configuraciones del software que han sido alteradas en un tiempo. Los históricos de versiones necesitan ser conservadas en un lugar central y si las modificaciones ocurren dentro del sistema con un usuario específico, varias versiones necesitan ser cargadas en paralelo.

#### **4.1.1.9 Consola de Administración.**

Las consolas de administración proveen una interfaz de administración y operación hacia el application server, dando a los usuarios la habilidad de automatizar tareas administrativas a través de parámetros de configuración del sistema. Así mismo, permiten el acceso remoto a la configuración y administración de sistema sin depender de la ubicación física del mismo.

#### **4.1.1.10 Herramientas de Diagnóstico.**

En un ambiente de producción, es de suma importancia diagnosticar de manera inmediata los problemas de operación que se presenten con el fin de dar solución pronta a los mismos y disminuir la afectación de los servicios.

Los diagnósticos son muy útiles para obtener una vista clara y en línea de los problemas con el servidor, excepciones, performance, etc.

BEA Weblogic Server provee una serie de herramientas de diagnóstico muy útiles para la detección de problemas las cuales pueden ser utilizadas desde la consola de administración.

#### **4.1.1.11 Log de Eventos.**

Es importante que un application server genere logs de eventos en donde se pueda identificar mensajes de alerta o componentes que generan algún problema a fin de enfocar los esfuerzos para resolverlo.

El monitoreo del application server generalmente incluye JVM logs, logs de administración del sistema, logs de instalación, logs de operación, etc.

En los logs se puede encontrar:

- ◆ Recursos de frecuente acceso
- ◆ Actividad por día e intervalos de tiempo
- ◆ Cantidad de datos enviados y recibidos
- ◆ Direcciones IP de quienes acceden al sitio
- ◆ Numero de “hits”
- ◆ Problemas al resolver peticiones

#### **4.1.1.12 Soporte.**

Aunque el soporte no es un componente de cualquier producto de software, éste es un servicio esencial en el funcionamiento del mismo. Lo anterior, debido a que siempre existen problemas con cualquier tipo software, algunos de estos problemas son de configuración, instalación o entendimiento de cómo funciona el software. Aún el mejor software tiene “bugs” de funcionamiento, por lo que es importante contar con un soporte especializado que nos ayude a solucionar o mejorar los problemas que se puedan presentar. Para una empresa de iniciativa privada y la cual base su negocio en software de aplicación, es importante contar con soporte técnico las 24hrs del día los 365 del año.

BEA Weblogic y SUN Java App Server proveen programas de soporte diseñados para grandes empresas incluyendo soporte de misión crítica 7X24 los 365 días del año. En cambio, Apache Tomcat, no cuenta con este tipo de soporte porque es un software de dominio público, es decir, “Open Source applications server” por lo que es soportado por la comunidad de desarrolladores de internet y no garantiza la solución inmediata a cualquier tipo de problema “bug” detectado.

#### **4.1.1.13 Capacitación.**

BEA Weblogic Server y SUN Java App Server proveen programas de capacitación para el manejo y administración del software. Lo anterior, es muy importante para los administradores a cargo de la operación del sistema.

En el caso Apache Tomcat no cuenta con representación comercial y por lo tanto no existen programas de capacitación disponibles para el usuario final.

#### **4.1.1.14 Costo.**

Este es un elemento principal en la elección de un application server. Los “Open Source application servers” están libres de costo. Sin embargo, no es recomendable instalarlos en compañías grandes que requieran de un soporte técnico 7X24X365 para solucionar los problemas que se presentan en la producción, porque puede incrementar la pérdida de ingreso y reputación para la compañía de manera considerable.

Los costos para algunos application servers se calculan con base al número de licencias que se desean adquirir más el costo anual de soporte técnico y mantenimiento. Una vez adquiridas las licencias, solo se pagará anualmente el costo correspondiente al soporte técnico y mantenimiento. Es importante mencionar que para la mayoría de los application servers el número de licencias a utilizar, es igual al número de procesadores con los que cuenta la maquina o servidor a utilizar. Esto es, si la maquina en donde se desea instalar un application server cuanta con 4 cpu's, se deben adquirir 4 licencias del application server seleccionado.

#### **4.1.2 Resultados obtenidos de la evaluación.**

Cada característica evaluada se calificó en comparación con los tres application servers: Apache Tomcat como referencia, BEA Weblogic y SUN Java App Server de la siguiente manera:



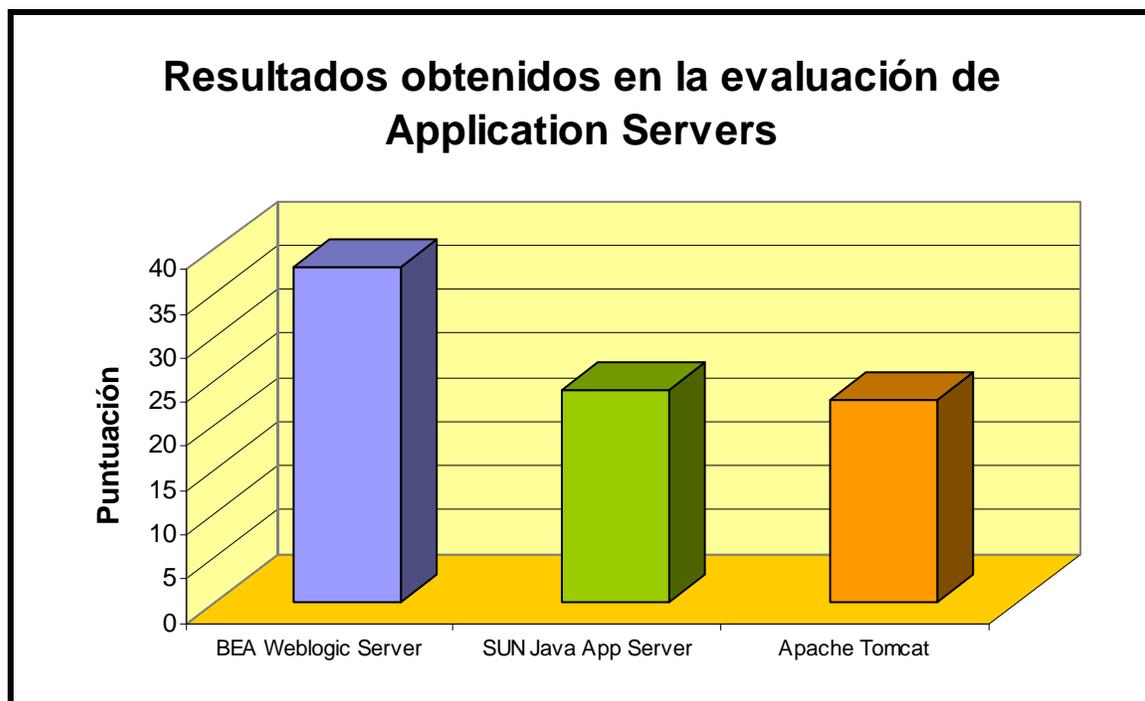


Figura 4.1 Resultados obtenidos en la evaluación de Application Servers.

BEA Weblogic Server	SUN Java App Server	Apache Tomcat Software Actual
37 puntos	23 puntos	20 puntos

Por lo anterior, se decidió adquirir e implementar BEA weblogic server para el funcionamiento de las aplicaciones masivas y corporativas desarrolladas en java.

### 4.1.3 Cambio de Apache Tomcat a BEA Weblogic Server.

Básicamente el cambio realizado en ésta primer etapa consistió en la instalación, configuración y puesta en producción del nuevo web application server en alta disponibilidad Activo-Pasivo a través de Veritas Cluster.

A continuación se muestra el diagrama de configuración final en esta primer etapa:

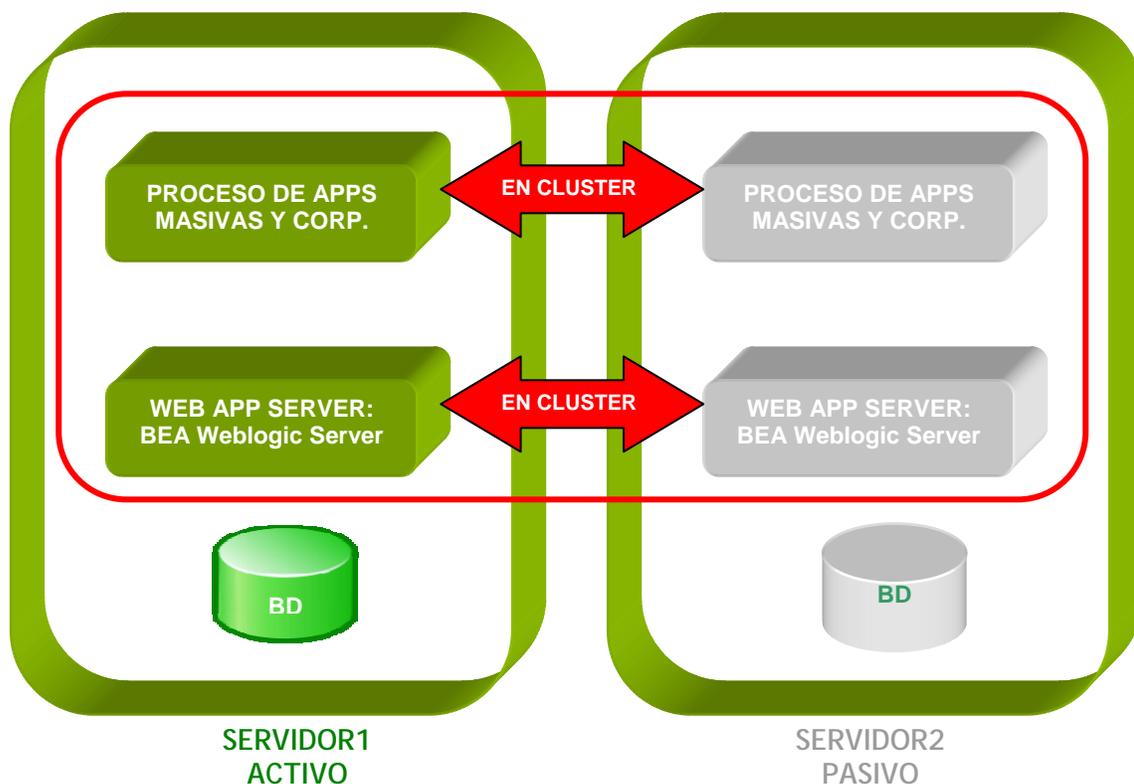


Figura 4.2 Distribución de procesos aplicativos y de base de datos después de la primera fase.

Con la adquisición del producto BEA Weblogic Server se adquirió también el acceso a soporte técnico especializado y derecho a actualizaciones del software para tener siempre las nuevas versiones y por lo tanto hacer uso de las nuevas funcionalidades del producto. Por lo anterior, este web application server se contemplo reutilizarlo en la segunda etapa del proyecto.

## 4.2 Segunda Fase: Diseño e implementación de la nueva arquitectura.

Normalmente, el modelo “cliente-servidor” de un servicio en red funciona de la siguiente manera: los clientes solicitantes del servicio dirigen sus peticiones a un servidor y éste responde ofreciéndoles el servicio solicitado en función de la petición.

Como se mencionó en el capítulo 3, las limitantes de este modelo de servicio en nuestra arquitectura inicial fueron principalmente las siguientes:

- ♦ La **disponibilidad del servicio** dependía directamente del correcto funcionamiento de los procesos aplicativos y de las base de datos. Adicional, no se contaba con una solución de monitoreo automatizado y tampoco existía una solución de respaldo confiable lo que impactaba directamente en la disponibilidad del servicio.
- ♦ La **capacidad de los servidores** que formaban parte del cluster para atender, procesar y servir peticiones, no se aprovechaba al máximo dada la configuración “activo & pasivo” que se tenía. Por lo anterior, tampoco existía una correcta distribución del tráfico y las peticiones que entraban de los clientes eran procesadas por el servidor activo provocando la saturación de sus recursos. Adicional, resultaba urgente incrementar la capacidad de la arquitectura inicial debido al crecimiento esperado de los servicios a corto y mediano plazo.

Por lo anterior, para incrementar la **disponibilidad del sistema SVA**, la propuesta de actualización tecnológica consistió en implementar una nueva arquitectura en dos capas que contemplará las siguientes mejoras:

- ♦ Alta disponibilidad de los servidores y aplicaciones involucradas en configuración “activo & activo” para aprovechar al máximo los recursos de hardware y software adquiridos.
- ♦ Distribución del tráfico entrante al *sistema SVA* para mejorar el procesamiento de las peticiones en cada servidor.
- ♦ Distribución de los procesos aplicativos para obtener una mejor calidad del servicio y una mejor administración en la operación.
- ♦ Monitoreo centralizado de todo el sistema a nivel hardware y software.

- ◆ Integración a un sistema de respaldos confiable y automático.
- ◆ Fácil de escalar y transparente para el usuario final.

Entiéndase por capa “front-end” al conjunto de hardware y software encargados de atender y distribuir las peticiones hacia la capa back-end o hacia otros sistemas externos que forman parte de un servicio.

La definición de capa “back-end” se refiere al conjunto de hardware y software encargados de procesar las peticiones recibidas por parte de la capa front-end. En esta capa residen los procesos aplicativos e instancias de bases de datos encargados del procesamiento y almacenamiento final de las transacciones del sistema.

Dado que la implementación de esta propuesta debía ser transparente a los usuarios finales, fue necesario seguir ofreciendo el servicio sin variar el escenario de red inicial, esto es, se conservó la dirección IP Virtual donde los clientes del servicio dirigían sus peticiones. Por lo anterior, en el diseño propuesto de la nueva arquitectura se consideró la implementación de balanceadores de carga.

#### 4.2.1 Diseño propuesto de la nueva arquitectura en dos capas: back-end y front-end.

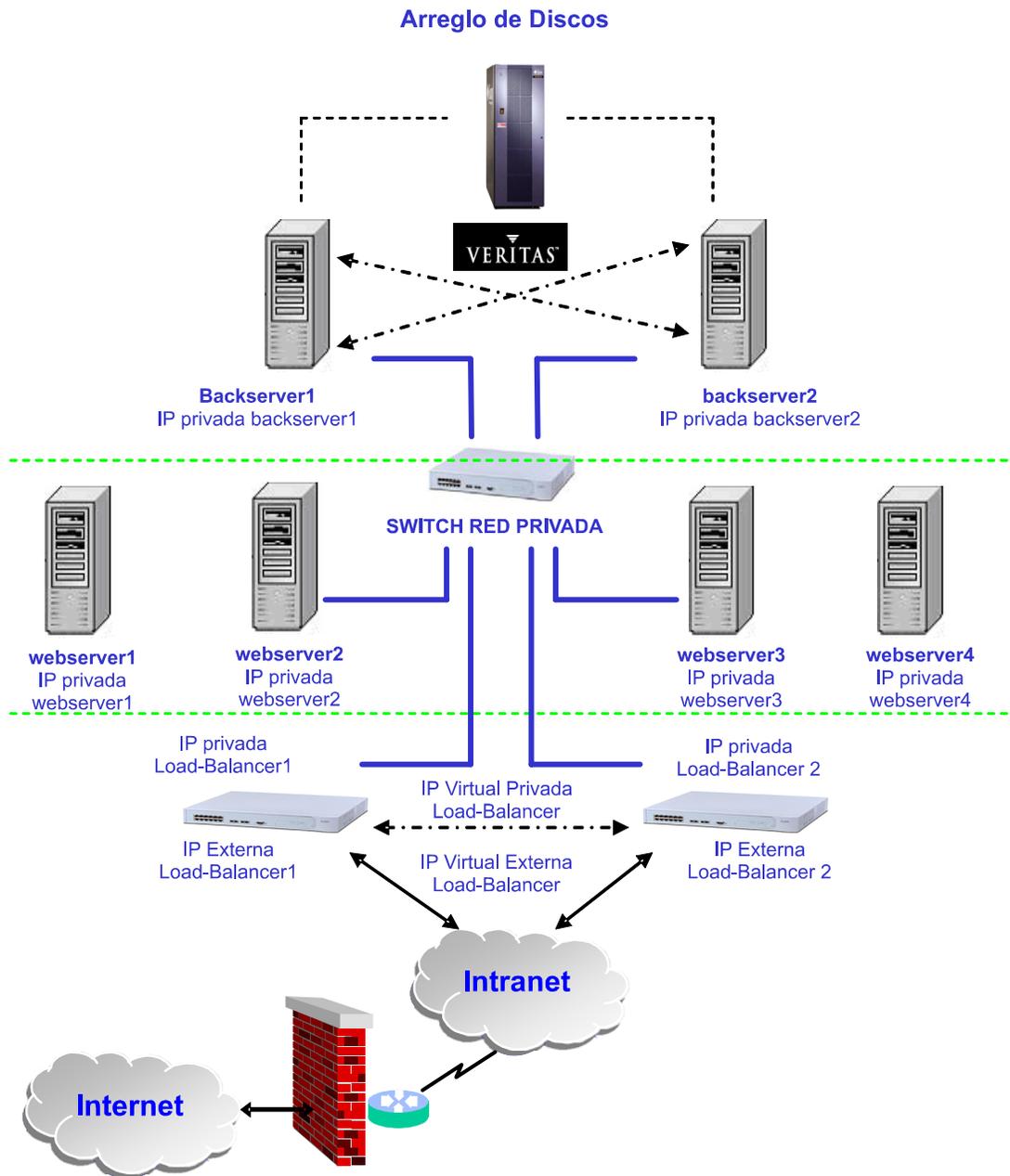


Figura 4.3 Diseño propuesto de la nueva arquitectura.

#### 4.2.2 Diseño propuesto de alta disponibilidad “activo-activo” en los servidores back-end.

En esta propuesta se contempló la distribución y configuración en alta disponibilidad “activo-activo” de los procesos aplicativos e instancias de base de datos encargados del procesamiento y almacenamiento final de las transacciones del sistema. El software propuesto para la configuración en alta disponibilidad fue Veritas Cluster, por lo que a través de éste se crearon grupos para el control independiente de cada base de datos y procesos aplicativos requeridos.

A continuación se muestra la propuesta de distribución y agrupación de las bases de datos y procesos aplicativos en los servidores back-end:

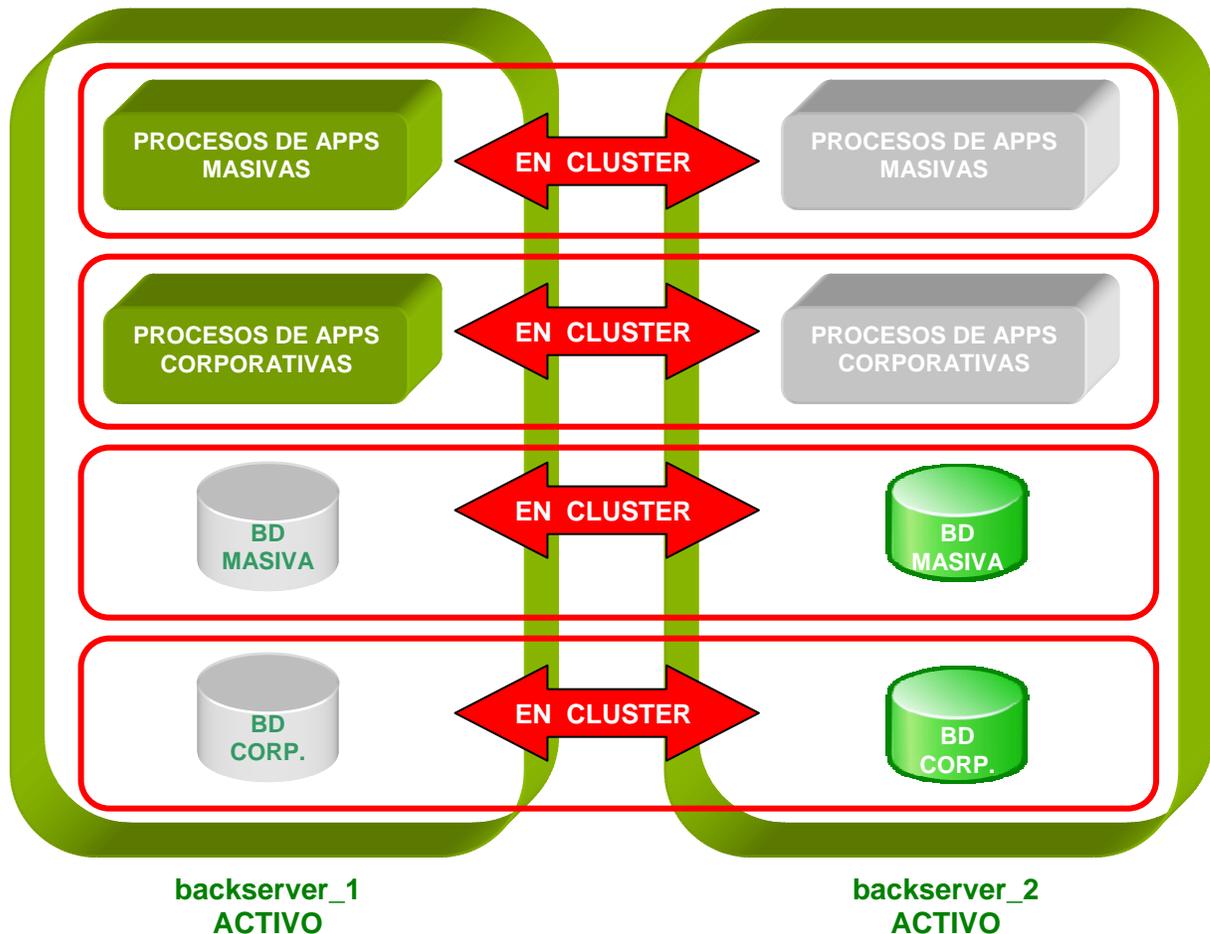


Figura 4.4 Alta disponibilidad “activo-activo” en servidores back-end.

### 4.2.3 Diseño propuesto de alta disponibilidad “activo-activo” en servidores front-end.

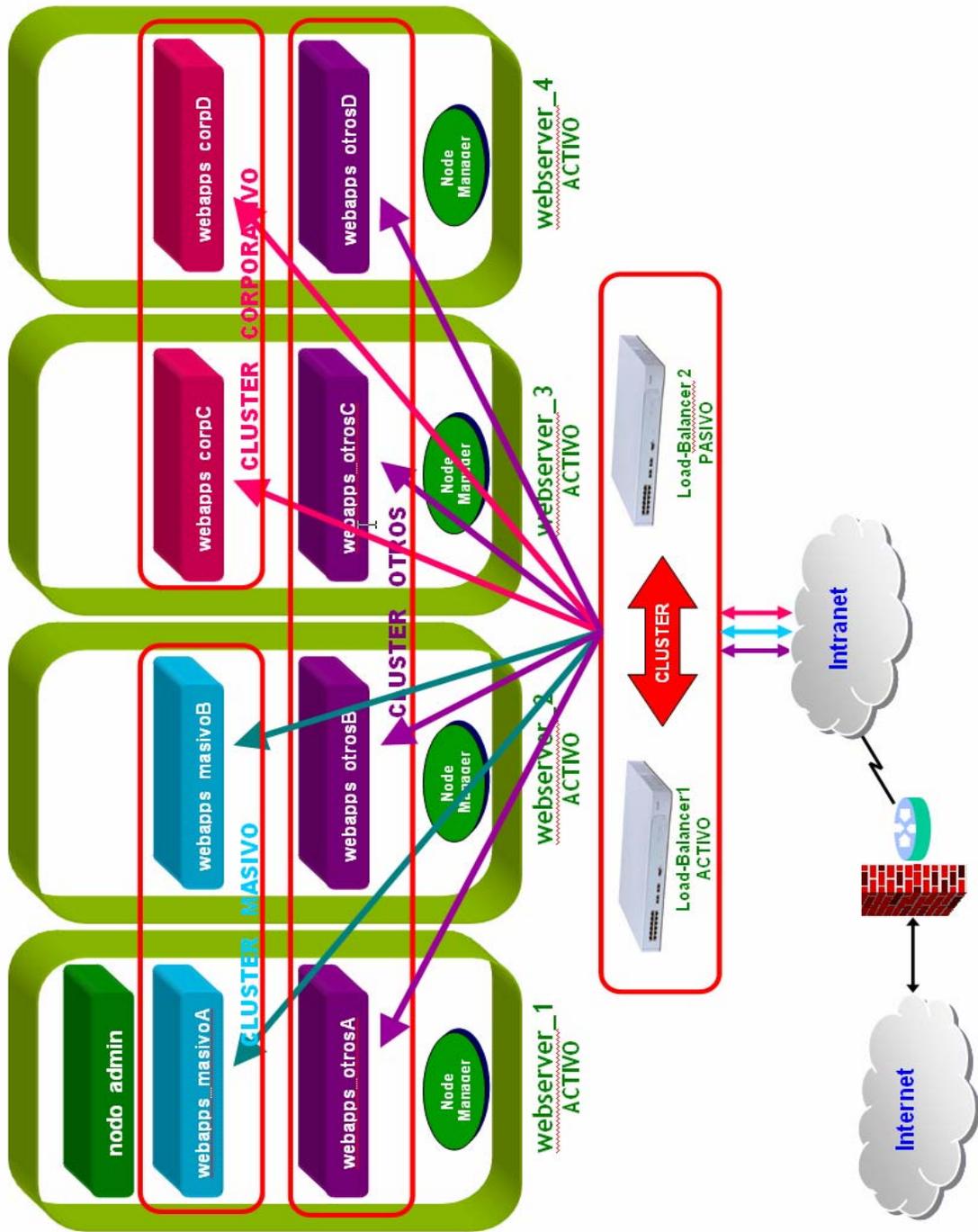


Figura 4.5 Alta disponibilidad “activo-activo” en servidores front-end.

Como se puede apreciar en la figura 4.5, el diseño propuesto de la nueva arquitectura contempló una solución de Balanceo de Cargas en Alta disponibilidad para todas las aplicaciones contenidas en el *sistema SVA*, las cuales se propuso publicar en varios clusters de servidores de aplicaciones BEA WebLogic garantizando así que tanto el contenido como las sesiones de dichas aplicaciones fueran manejadas y replicadas en los clusters de weblogic correspondientes.

Los clusters sugeridos para lograr una mejor distribución de las aplicaciones y así obtener una mejor calidad del servicio y una mejor administración en la operación fueron:

- ◆ **Cluster corporativo.** Este cluster se propuso exclusivamente para la carga de aplicaciones dedicadas a los servicios corporativos.
  
- ◆ **Cluster masivo.** Este cluster se propuso exclusivamente para la carga de aplicaciones dedicadas a los servicios masivos.
  
- ◆ **Cluster otros.** Este cluster se propuso exclusivamente para la carga de otras aplicaciones que dada su naturaleza requieren un ambiente independiente a los servicios masivos y corporativos.

De esta manera las peticiones entrantes al sistema llegan en primer instancia al balanceador de carga activo, el cual direcciona cada petición hacia el servidor front-end más adecuado para atenderla, si la petición requiere por la lógica del servicio, hacer una consulta de base de datos o una petición hacia los procesos aplicativos que viven en los servidores back-end, se establece la conexión desde la instancia de weblogic correspondiente manteniendo la sesión de la petición del cliente hasta su procesamiento final.

#### 4.2.4 Diseño propuesto para la configuración del Dominio de BEA Weblogic.

Un dominio de BEA Weblogic está conformado básicamente de

- ◆ Maquinas.
- ◆ Un servidor administrador: “nodo\_admin”.
- ◆ Servidores administrados “instancias”.
- ◆ Node managers.
- ◆ Clusters.

En un dominio:

- ◆ Solo puede haber un servidor Administrador.
- ◆ Puede haber múltiples servidores administrados.
- ◆ Puede haber múltiples maquinas definidas.
- ◆ Pueden existir múltiples clusters.
- ◆ Pueden existir servidores administrados que forman parte de un cluster y servidores administrados que no están en cluster

La solución propuesta consistió en crear un dominio de servidores de aplicación distribuidos y configurados en cuatro máquinas físicas de la siguiente manera:

##### 4.2.4.1 Configuración de maquinas.

Una maquina es un servidor físico en donde una o más instancias pueden estar trabajando al mismo tiempo, siempre y cuando se configuren en puertos distintos.

<b>Hostname maquina 1:</b>	<b>webserver_1</b>
<b>Servidor Administrador:</b>	nodo_admin
<b>Servidores administrados o instancias:</b>	webapps_masivoA webapps_otrosA
<b>Node_manager de máquina 1</b>	

Tabla 4.1 Instancias propuestas en la maquina “webserver\_1”.

<b>Hostname maquina 2:</b>	<b>webserver_2</b>
<b>Servidores administrados o instancias:</b>	webapps_masivoB webapps_otrosB
<b>Node_manager de máquina 2</b>	

Tabla 4.2 Instancias propuestas en la maquina “webserver\_2”.

<b>Hostname maquina 3:</b>	<b>webserver_3</b>
<b>Servidores administrados o instancias:</b>	webapps_corpC webapps_otrosC
<b>Node_manager de máquina 3</b>	

Tabla 4.3 Instancias propuestas en la maquina “webserver\_3”.

<b>Hostname maquina 4:</b>	<b>webserver_4</b>
<b>Servidores administrados o instancias:</b>	webapps_corpD webapps_otrosD
<b>Node_manager de máquina 4</b>	

Tabla 4.4 Instancias propuestas en la maquina “webserver\_4”.

#### 4.2.4.2 Función del nodo\_admin.

El nodo\_admin es la instancia de administración del dominio (administration server), el cual puede gestionarse de manera centralizada desde una consola web. La configuración de todo el dominio es almacenada en esta instancia en el archivo **config.xml**, la cual, se encarga de actualizarlo y pasar la configuración correspondiente a cada uno de los servidores administrados.

#### 4.2.4.3 Configuración de servidores administrados o instancias.

Un servidor administrado o instancia, es la entidad en donde se cargan y publican las aplicaciones de usuarios finales. A cada instancia se le configura la IP y el puerto a través del cual se publican las aplicaciones, y la memoria asignada para el funcionamiento de dicha instancia.

<i>Instancia</i>	<b>Listening Address</b>	<b>Listening Port</b>	<b>Memoria(MB) Asignada</b>
nodo_admin	IP privada webserver1	9999	384
webapps_masivoA	IP privada webserver1	7010	512
webapps_otrosA	IP privada webserver1	7030	512
webapps_masivoB	IP privada webserver2	7010	512
webapps_otrosB	IP privada webserver2	7030	512
webapps_corpC	IP privada webserver3	7020	512
webapps_otrosC	IP privada webserver3	7030	512
webapps_corpD	IP privada webserver4	7020	512
webapps_otrosD	IP privada webserver4	7030	512

Tabla 4.5 Parámetros de configuración propuestos para cada instancia.

#### 4.2.4.4 Configuración de node managers.

Cada una de las maquinas contempla un proceso de administración de instancias llamado **node manager**, el cual se encarga de reportar a la instancia administradora el estado operativo de las instancias que se encuentran corriendo en esa máquina. El proceso node manager es utilizado en cada máquina para realizar las tareas administrativas de alta, baja y monitoreo de las instancias.

Maquina	Listening Address	Listening Port	Instancias
webserver_1	IP privada webserver1	6666	nodo_admin webapps_masivoA webapps_otrosA
webserver_2	IP privada webserver2	6666	webapps_masivoB webapps_otrosB
webserver_3	IP privada webserver3	6666	webapps_corpC webapps_otrosC
webserver_4	IP privada webserver4	6666	webapps_corpD webapps_otrosD

Tabla 4.6 Parámetros de configuración propuestos para cada node manager.

#### 4.2.4.5 Configuración de clusters.

Un cluster es un grupo de instancias que trabajan coordinadamente. Un cluster provee principalmente:

- ♦ Alta disponibilidad: Asegura la continuidad del servicio cuando una máquina o instancia que forma parte del cluster, no está disponible.
- ♦ Escalabilidad: Permite configurar y agregar más maquina e instancias sin interrumpir el servicio al usuario final.

Maquina	Multicast Address	Multicast Port	Instancias
cluster_masivo	237.0.0.101	7050	webapps_masivoA webapps_masivoB
cluster_corp	237.0.0.102	7050	webapps_corpA webapps_corpB
cluster_otros	237.0.0.103	7050	webapps_otrosA webapps_otrosB webapps_otrosC webapps_otrosD

Tabla 4.7 Parámetros de configuración propuestos para cada cluster.

Las instancias de WebLogic en un cluster se comunican entre ellas utilizando mensajería multicast para enviar heartbeats cada 10 segundos; si 3 heartbeats de un mismo servidor se pierden, entonces es marcado como FAILED y se deja de enviar trabajo a esa instancia.

#### 4.2.4.6 Acceso a las bases de datos desde el dominio de Weblogic Server.

Como se explicó anteriormente las bases de datos se contemplaron en los servidores back-end, por lo que fue necesario considerar su interconexión con el dominio de servidores de aplicación propuesto. Esta interconexión se contempló a través de los data source y pool de conexiones configurables desde la consola de administración de WebLogic Server.

Un **data source** es un objeto por medio del cual se puede acceder al pool de conexiones.

Un **connection pool** es un grupo de conexiones activas y asociadas a una base de datos.

Algunas ventajas de usar pool de conexiones son:

- ◆ La información y configuración de todas las conexiones de un pool está centralizada en la consola de administración.
- ◆ Se puede limitar el número de conexiones que se abren hacia la base de datos.
- ◆ Se aísla la aplicación de la base de datos que se está usando, de tal forma que se puede cambiar la base de datos sin tener que modificar la aplicación.

El JDBC driver es un API que permite la ejecución de operaciones sobre bases de datos independientemente del sistema operativo donde se ejecute o de la base de datos a la cual se quiere acceder.

El siguiente diagrama muestra el flujo de acceso de las aplicaciones que residen en el dominio de Weblogic Server hacia las bases de datos:

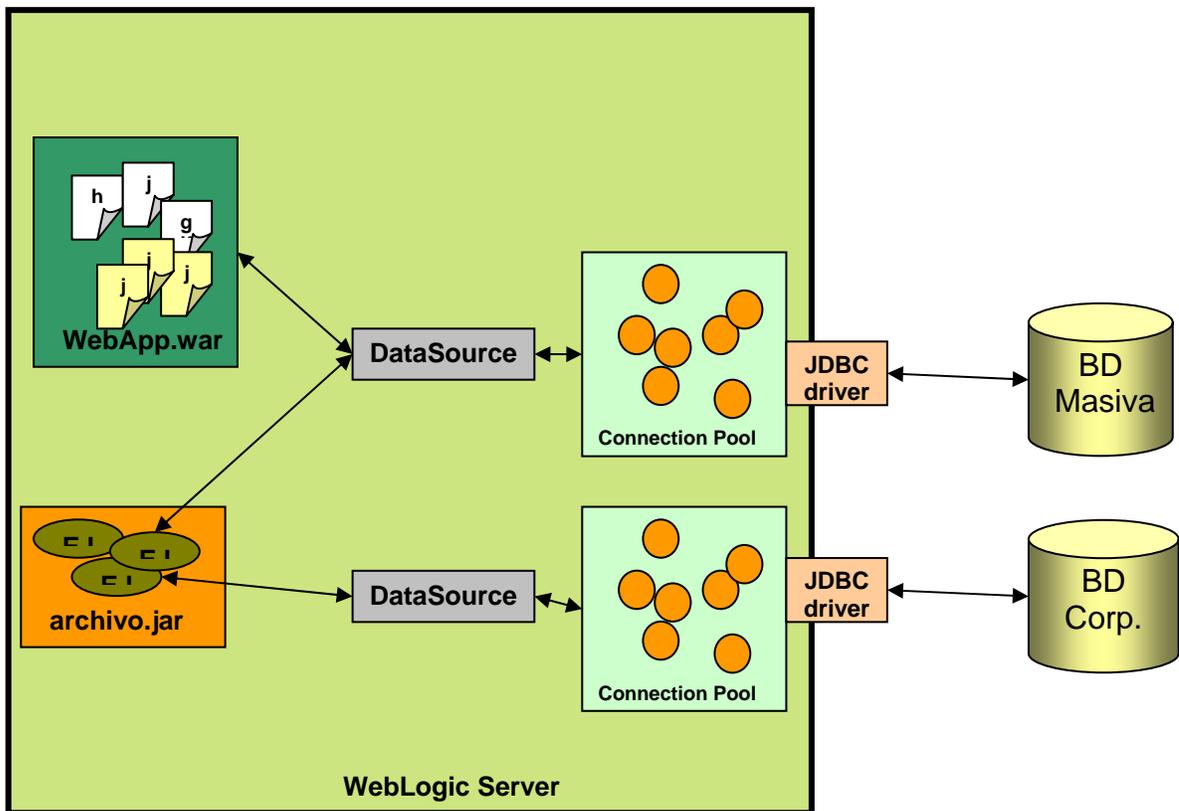


Figura 4.6 Diagrama de flujo para acceso a Bases de Datos desde Weblogic Server.

#### **4.2.5 Diseño de direccionamiento IP utilizando balanceadores de carga.**

En la solución de Balanceo de Cargas en Alta disponibilidad, se contempló la creación de una red LAN “privada” para establecer la comunicación entre todos los elementos propuestos de la nueva arquitectura. Por lo anterior, se asignó a cada equipo una IP válida de este segmento de red LAN “privada”, la cual, se propuso independiente a los segmentos de red válidos en la empresa y por lo tanto no accesible directamente desde una red externa. El acceso para la administración y mantenimiento de cada uno de los servidores propuestos en la nueva arquitectura, así como el acceso a los servicios de las aplicaciones, solo es posible a través del uso del protocolo NAT y de servidores virtuales configurados en los balanceadores de carga.

Con la propuesta de creación de servidores virtuales en los balanceadores de carga, se mantiene el servicio sin cambiar la dirección IP Virtual (VIP) de la arquitectura inicial. De esta manera la actualización tecnológica fue transparente para los usuarios finales.

Las mejoras que proporciona el uso de balanceadores de cargas son:

- ◆ Alta disponibilidad del servicio.
- ◆ Calidad del servicio y una mejor administración en la operación.
- ◆ Facilidad de mantenimiento.
- ◆ Escalabilidad.
- ◆ Seguridad.
- ◆ Mejor distribución de tráfico y por lo tanto un mejor rendimiento en los recursos (cpu y memoria) de los servidores involucrados

El hecho de distribuir el trabajo entre diferentes máquinas permite que la capacidad de carga del servidor virtual sea muy superior a la que tendría un solo servidor físico.

El balanceador de carga toma las decisiones de reenvío de peticiones en función del algoritmo que tiene configurado.

4.2.5.1 Direccionamiento IP asignado a la arquitectura inicial.

<b>Segmento de red LAN en la empresa:</b>	100.100.75.0
<b>Default Router:</b>	100.100.75.1
<b>IP virtual de la arquitectura:</b>	<b>100.100.75.5</b>
<b>IP física del servidor1:</b>	100.100.75.3
<b>IP física del servidor2:</b>	100.100.75.4
<b>Nombre del dominio:</b>	nombre.empresa.com

Tabla 4.8 Direccionamiento IP de la arquitectura inicial.

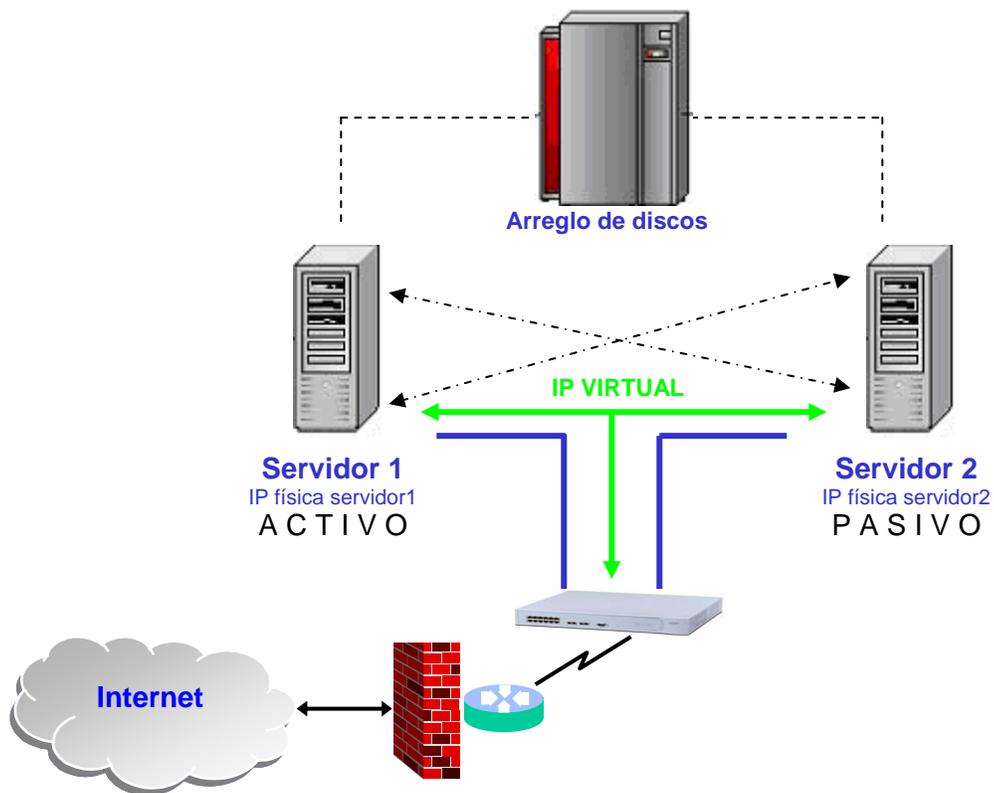


Figura 4.7 Arquitectura inicial del sistema.

## 4.2.5.2 Propuesta de direccionamiento IP en la nueva arquitectura

Descripción	Direccionamiento de Red LAN válida en la empresa	Direccionamiento de Red LAN “privada”
<b>Segmento de red LAN:</b>	100.100.75.0	172.28.1.0
<b>Default Router:</b>	100.100.75.1	172.28.1.1
<b>IP backserver1</b>	100.100.75.10	172.28.1.10
<b>IP backserver2</b>	100.100.75.11	172.28.1.11
<b>IP webserver1</b>	100.100.75.21	172.28.1.21
<b>IP ebserver2</b>	100.100.75.22	172.28.1.22
<b>IP webserver3</b>	100.100.75.23	172.28.1.23
<b>IP webserver4</b>	100.100.75.24	172.28.1.24
<b>IP virtual Apps Masivo</b>	100.100.75.31	172.28.1.31
<b>IP virtual Apps Corporativo</b>	100.100.75.32	172.28.1.32
<b>IP virtual BD Masivo</b>	100.100.75.33	172.28.1.33
<b>IP virtual BD Corporativo</b>	100.100.75.34	172.28.1.34
<b>IP virtual del LoadBalancer</b>	100.100.75.40	172.28.1.40
<b>IP Load Balancer 1</b>	100.100.75.41	172.28.1.41
<b>IP Load Balancer 2</b>	100.100.75.42	172.28.1.42
<b>IP virtual del sistema SVA</b>	<b>100.100.75.5</b>	
<b>Dominio</b>		
<b>Nombre del dominio</b>	nombre.empresa.com	
<b>Direccionamiento IP Homologado</b>		
<b>IP virtual del sistema SVA</b>	100.100.75.5	IP_homologada <sup>30</sup>

Tabla 4.9 Direccionamiento IP propuesto en la nueva arquitectura.

<sup>30</sup> Una IP pública ó homologada, es una dirección IP válida y accesible desde Internet.

#### 4.2.6 Diseño de políticas en balanceadores de carga.

A continuación se muestra el diseño propuesto de las políticas implementadas en los balanceadores de carga referentes a Pools de Servidores, Servidores Virtuales y NAT's:

##### 4.2.6.1 Pool de Servidores:

Un pool de servidores contiene la lista de servidores físicos que atenderán una aplicación así como el puerto de escucha. Además define el modo en cómo debe ser balanceado el tráfico hacia los servidores que integran el pool.

Nombre del pool	Método de balanceo	Miembros del pool	Puerto
pool_masivo	Round Robin	ip_privada_webserver1 ip_privada_webserver2	7010
pool_corporativo	Round Robin	ip_privada_webserver3 ip_privada_webserver4	7020
pool_otros	Round Robin	ip_privada_webserver1 ip_privada_webserver2 ip_privada_webserver3 ip_privada_webserver4	7030

Tabla 4.10 Configuración propuesta para pool de servidores en balanceadores de carga.

#### 4.2.6.2 Servidores Virtuales:

Los servidores virtuales son las direcciones IP y puertos TCP-UDP que deben utilizar los clientes para ingresar a sus aplicaciones y estos están asociados a los pools de servidores.

IP Virtual	Servicio	Pool de Servidores	Estado
100.100.75.5	7010	pool_masivo	Enabled
100.100.75.5	7020	pool_corporativo	Enabled
100.100.75.5	7030	pool_otros	Enabled

Tabla 4.11 Configuración propuesta para servidores virtuales en balanceadores de carga.

#### 4.2.6.3 NAT (Network Address Translation):

Los NAT's definen la relación de direcciones IP externa/IP interna asociadas a los servidores físicos. Se ocupan para que los servidores que contienen las aplicaciones en la red interna puedan ser administrados desde la red externa con los permisos de acceso previamente definidos por los administradores de los servidores.

Descripción	NAT Address	Origin Address	ARP
backserver1	100.100.75.10	172.28.1.10	Enabled
backserver2	100.100.75.11	172.28.1.11	Enabled
webserver1	100.100.75.21	172.28.1.21	Enabled
webserver2	100.100.75.22	172.28.1.22	Enabled
webserver3	100.100.75.23	172.28.1.23	Enabled
webserver4	100.100.75.24	172.28.1.24	Enabled
IP virtual Apps Masivo	100.100.75.31	172.28.1.31	Enabled
IP virtual Apps Corporativo	100.100.75.32	172.28.1.32	Enabled
IP virtual BD Masivo	100.100.75.33	172.28.1.33	Enabled
IP virtual BD Corporativo	100.100.75.34	172.28.1.34	Enabled

Tabla 4.12 Configuración propuesta de NAT's en balanceadores de carga.

#### **4.2.7 Diseño de la solución de monitoreo para la nueva arquitectura.**

El esquema de monitoreo que se propuso para la nueva arquitectura consistió en la configuración y notificación del estado operativo de los equipos que la conforman vía HP Open View Operations.

HP Open View Operations es un producto basado en la tecnología cliente/servidor, donde el servidor actúa como el Nodo Administrador "*Management Server*" y los clientes Nodos Administrados "*Managed Nodes*".

##### **4.2.7.1 Características de HP Open View Operations:**

- ◆ Concentra información del estado operativo de los equipos, así como de los procesos y aplicaciones que residen en un sistema. Esto es posible a través de agentes y herramientas instalados en los nodos administrados "*Managed Nodes*". Ésta información es una fuente excelente para el análisis estadístico de un sistema.
  
- ◆ Distribuye el monitoreo a través de consolas de acuerdo al perfil de cada operador o administrador. La asignación del perfil de usuario es importante porque con base a los privilegios asignados, el operador o administrador, podrá realizar las actividades que le son asignadas y dar seguimiento a un evento de falla. Los eventos pueden ser observados desde la Ventana de Visualización de Eventos (*Message Browser*).
  
- ◆ Una vez que una alarma llega al Nodo Administrador, HP Open View Operations provee mecanismos para reaccionar y dar solución a la misma a través de:
  - ◆ Acciones automáticas o iniciadas por el operador.
  - ◆ Textos de ayuda con instrucciones asociadas al problema específico.
  - ◆ HP Open View Operations también proporciona una interfaz hacia un sistema externo de notificación. Ej. Integración a un sistema de mensajes.

#### **4.2.7.2 Requerimientos de monitoreo para la nueva arquitectura.**

Los requerimientos de monitoreo a cubrir para esta nueva arquitectura fueron:

- ◆ Monitoreo a nivel Hardware.
- ◆ Monitoreo a nivel sistema operativo.
- ◆ Monitoreo a nivel aplicaciones comerciales: Oracle y Weblogic
- ◆ Monitoreo a nivel aplicaciones y desarrollos internos.

Por lo anterior, en el diseño de la solución, se propuso instalar agentes y herramientas en cada equipo para obtener información sobre el estado operativo del sistema.

#### **4.2.7.3 Agentes propuestos para el monitoreo a nivel Hardware y Sistema Operativo.**

Para el monitoreo del hardware y sistema operativo, se contempló el uso dos herramientas de HP: Openview Glance Plus y Openview Performance Manager.

#### **Funcionalidades de Openview Glance Plus.**

- ◆ Monitoreo del performance del sistema en línea de forma gráfica.
- ◆ Permite realizar diagnósticos del sistema mediante una herramienta gráfica.
- ◆ Permite examinar identificar y resolver cuellos de botella en el performance de las máquinas.
- ◆ Generación de alarmas cuando los niveles de utilización de recursos exceden los niveles aceptables de operación. Las alarmas son desplegadas en pantalla.
- ◆ Información de los parámetros de kernel del sistema operativo, tales como: tabla de procesos, tabla de buffers, archivos abiertos, memoria compartida, etc.
- ◆ Generación de gráficas para representar la utilización de los recursos del sistema, incluyendo utilización de CPU, memoria, operaciones de lectura escritura a disco, comunicación de red (interfaces de red), acelerando la detección de “cuellos de botella” en el sistema.

### **Funcionalidades de Openview Performance Manager.**

- ◆ Recopila información sobre el desempeño del sistema con datos colectados en periodos definidos por el usuario.
- ◆ Recolección permanente de métricas de rendimiento del sistema en tres niveles: sistema operativo, aplicación y proceso. Utilización de CPU, memoria, disco, red, monitoreo de entorno, etc.

#### **4.2.7.4 Agentes propuestos para el monitoreo de las aplicaciones comerciales.**

Para el monitoreo de las bases de datos e instancias de weblogic contempladas en la nueva arquitectura, se propuso el uso de los Smart Plug-in para Oracle y BEA WebLogic Server.

#### **SPI – Smart Plug-In.**

El SPI es un grupo de métricas definidas por los desarrolladores y fabricantes del producto que ayudan a aumentar el nivel de monitoreo del mismo.

#### **Funcionalidades de HP OpenView Smart Plug-in for Oracle.**

- ◆ Muestra el estado operativo de las bases de datos.
- ◆ Muestra el estado operativo de los procesos referentes a la BD.
- ◆ Permite conocer estado operativo del Tablespace y el porcentaje de espacio disponible.
- ◆ Estado y desempeño de los segmentos de las BD.
- ◆ Desempeño de las transacciones: Commit, Rollback, Rollback Comit.
- ◆ Procesos en tablas e índices.

- ◆ Encolamientos, accesos de usuarios, timeouts.
- ◆ Uso de memoria asignada en la inicialización.
- ◆ Otros.

#### **Funcionalidades de HP OpenView Smart Plug-in for BEA WebLogic Server.**

- ◆ Monitoreo de los archivos log del sistema, para notificación de errores y mensajes de alarma.
- ◆ Monitoreo de cambios en los archivos de configuración de WebLogic.
- ◆ Estado operativo del servidor WebLogic.
- ◆ Estatus de la conexión de JDBC.
- ◆ Utilización de la máquina virtual de Java.
- ◆ Más métricas predefinidas como los recursos que usa el WebLogic Server, tiempos de ejecución de servlets, time-outs, tasas de peticiones y transacciones, etc.

En las tablas 4.13 y 4.14 se muestra un ejemplo de un pequeño extracto del listado de métricas que se pueden seleccionar para el monitoreo de las instancias de weblogic.

Number	Metric Name	Description	Avail.	WLS Version	Type	Severity	Area	
1	B001_ServerStatus	Status of a server	All	A	Critical	Warning	Availability	
2	B002_ServerStatusRep	Status of a server - reporting	All	R			Availability	
5	B005_JVMMemUtilPct	% of heap space used in the JVM	All	GA	Critical	Major	JVM	
10	B010_ExQueThruRate	# of requests serviced by an execute queue per second	All	GR			Performance	
11	B011_ExQThrdUtilPct	% of threads in use for a server's execute queue.	All	GRA	Critical	Major	Minor	Performance
12	B012_ExQueWaitCnt	# of client requests waiting to be serviced	All	GA	Minor		Performance	
13	B013_SocketTrafficRt	# of socket connections opened per second	All	G			Performance	
14	B014_ActiveSocketCnt	# of socket connections opened	All	GA	Minor		Performance	
15	B015_SrvrRestartsPct	% of permissible restarts	7.0, 8.1, 9	GA	Critical	Warning	Performance	
25	B025_EJBFreePoolWtRt	# of times/min no EJB beans were available from the free pool	All	GRA	Warning		EJB	
225	B225_EJBFreePoolWaitRate	# of times/min no EJB beans were available from the free pool (drill down)	All	A	Warning		EJB	
26	B026_EJBTimeoutRt	# of times/min a client timed out waiting for an EJB bean	All	GRA	Warning		EJB	
226	B226_EJBTimeoutRate	# of times/min a client timed out waiting for an EJB bean (drill down)	All	A	Warning		EJB	
35	B035_EJBTranThruRt	# of EJB transactions per second	All	GRA	Warning		EJB	
36	B036_EJBTranRbRt	# of EJB transactions rolled back per second	All	GRA	Warning		EJB	
238	B238_EJBCacheHitPct	% of EJBs in the cache in use	All	RA	Warning		EJB	

Tabla 4.13 Listado de métricas a considerar en el monitoreo de Weblogic Server.

Number	Metric Name	Description	Aval. WLS Version			Area
			Type	Severity		
70	B070_TransAveTime	Average commit time for transactions	All	GRA	Minor	Transactions
270	B270_ConnectorConnectionPoolUtil	% utilization of available JCA connections in connection pool	7.0, 8.1, 9	RA	Critical Major	Connector
71	B071_TransRollbackPct	% of transactions rolled back	All	GRA	Minor	Transactions
72	B072_TransResErrRbPct	% of transactions rolled back due to resource error	All	GRA	Minor	Transactions
73	B073_TransAppErrRbPct	% of transactions rolled back due to application error	All	GRA	Minor	Transactions
74	B074_TransTimErrRbPct	% of transactions rolled back due to timeout error	All	GRA	Minor	Transactions
75	B075_TransSysErrRbPct	% of transactions rolled back due to system error	All	GRA	Minor	Transactions
76	B076_TransThruRate	# of transactions processed per second	All	GR		Transactions
77	B077_TransHeurCnt	% of transactions returning a heuristic decision	All	GRA	Minor	Transactions
78	B078_ConnectorConnectionPoolLeakedConnRateSum	# of unclosed connector connections and ones that have exceeded their max idle time	7.0, 8.1, 9	G		Connector
278	B278_ConnectorConnectionPoolLeakedConnRate	Rate of leaked connections for the JCA connection pool	7.0, 8.1, 9	RA	Warning	Connector
79	B079_TransCapacityUtil	% utilization of transaction capacity	7.0, 8.1, 9	GRA	Critical Major	Transactions

Tabla 4.14 Listado de métricas a considerar en el monitoreo de Weblogic Server.

Así mismo, a continuación se muestran ejemplos de la descripción a detalle de algunas métricas de monitoreo disponibles en el Smart Plug-in for BEA WebLogic Server

### Metric B001\_ServerStatus

<b>Monitor Template Name</b>	WLSSPI_0001
<b>Metric Name</b>	B001_ServerStatus
<b>Metric Type</b>	Alarming
<b>Description</b>	Status of a server, monitors whether running or not.
<b>Avail. WLS Version</b>	6.1, 7.0, 8.1, 9
<b>Severity: Condition with threshold</b>	Four conditions Critical: WLSSPI-0001.1, threshold 5 Warning: WLSSPI-0001.2, threshold 1.5
<b>Collection Interval</b>	5m
<b>Message Group</b>	WebLogic
<b>Message Text</b>	WLSSPI-0001.1: Server status is unknown (down) WLSSPI-0001.2: Server status: Suspended
<b>Instruction Text</b>	<p>For each server, this metric reports the status (running, shutdown in progress, shutdown pending, suspended, or unknown).</p> <p><b>Probable cause:</b> If the server is not in a running state, the following events may have occurred:</p> <ol style="list-style-type: none"> <li>1 The WebLogic Administrator has selected “Shutdown this server” from the Administration console.</li> <li>2 The WebLogic Administrator has selected “Suspend this server” from the Administration console.</li> <li>3 The server may have gone down for other reasons.</li> </ol> <p><b>Potential Impact:</b> If the server is Shutdown or in the process of shutting down, the server is no longer be available. If the server is Suspended, it only accept requests from the Administration Server. Suspending the WebLogic Server only suspends server responses to HTTP requests. Java applications and RMI invocations are not suspended.</p> <p><b>Suggested action:</b> If the designated server is not running, the WebLogic Administrator should start the server using the appropriate script. It is important to note whether this is the Administration Server or a Managed Server, since the startup script will be different for each type.</p> <p>If the server has been suspended, it may have been placed in this state for a reason. A typical use of this feature would be in a situation where a WebLogic Server is running as a “hot” backup for another server. When it is OK to do so, execute the “Resume this server” command from the Administration console.</p>
<b>Report Type</b>	N/A
<b>Area</b>	Availability

Tabla 4.15 Detalle de la métrica B001\_ServerStatus de Weblogic Server.

#### **4.2.7.5 Solución de monitoreo para aplicaciones y desarrollos internos “in-house”.**

El monitoreo de las aplicaciones y desarrollos internos se propuso a través del monitoreo de sus archivos logs y de sus procesos aplicativos correspondientes.

#### **4.2.7.6 Diagrama de solución final de monitoreo HP OVO.**

El diagrama de la figura 4.8, muestra la interconexión final propuesta de monitoreo centralizado para la nueva arquitectura a nivel hardware y software.

Los agentes de HP OVO y SPI instalados y configurados en cada equipo de la solución, enviarán al nodo administrador de HP OVO el estado operativo correspondiente al Hardware, Sistema Operativo, Bases de Datos y Web Application Servers.

Para conocer el estado operativo de las aplicaciones y desarrollos internos “in-house”, el nodo administrado de HP OVO poleará constantemente el estado operativo de dichos procesos a través de los logs y procesos definidos para este fin. De esta manera, cuando ocurra un evento de falla, el nodo administrador podrá notificar dicho evento a los operadores y administradores de este sistema para resolver de manera inmediata el problema.

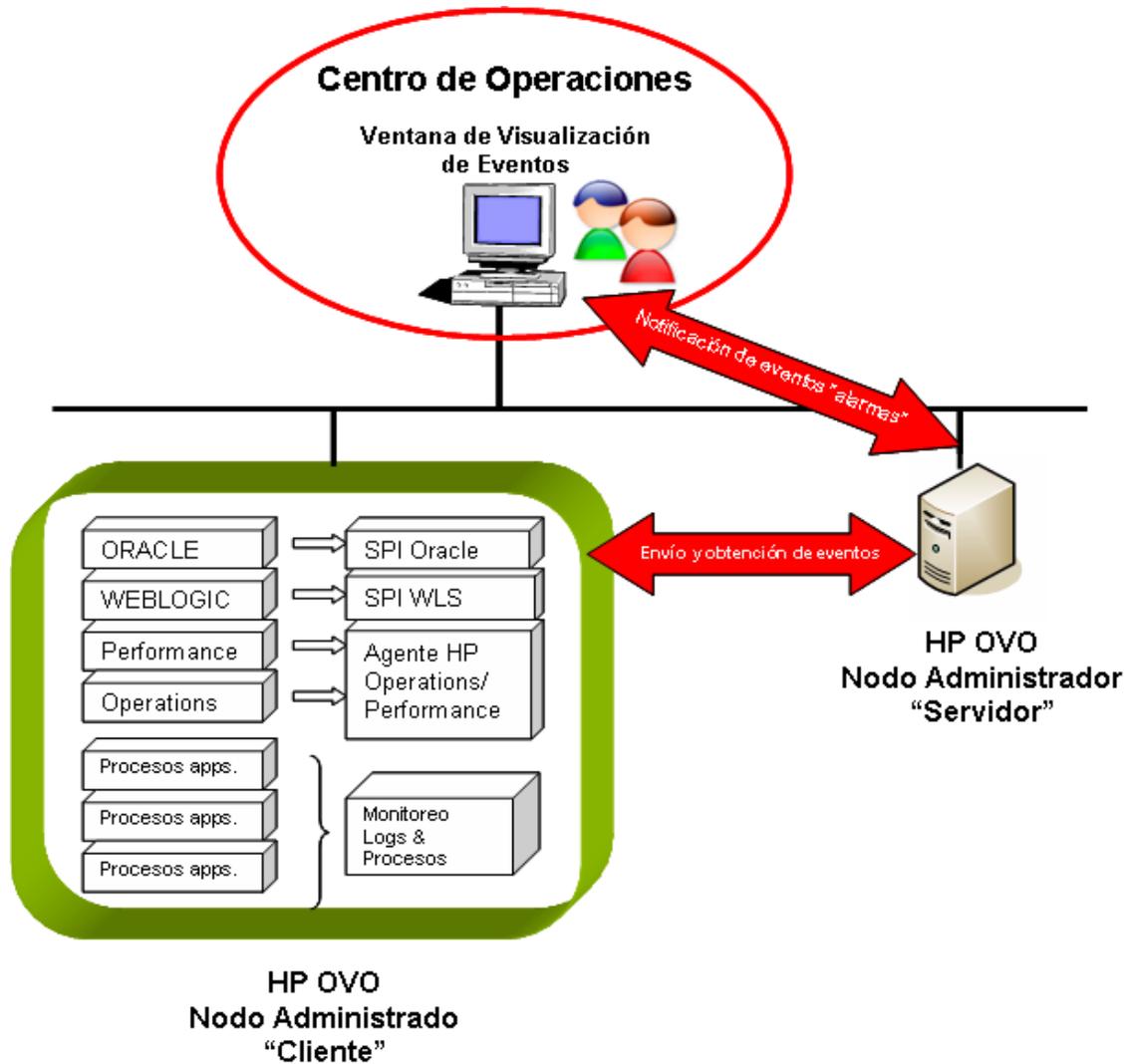


Figura 4.8 Diagrama de solución final para el monitoreo del sistema SVA a través de HP OVO

## **4.2.8 Diseño de la solución de respaldo de información para la nueva arquitectura.**

### **4.2.8.1 Requerimientos para la nueva arquitectura.**

- ◆ Respaldo confiable, automático y en línea de:
  - ◆ Sistema operativo
  - ◆ Software aplicativo
  - ◆ Base de datos.

Por lo anterior, la solución que se propuso consistió en integrar cada uno de los servidores que la componen la nueva arquitectura a una SAN de respaldos.

Con esta solución se proveen principalmente las siguientes ventajas:

- ◆ Al ser una solución de almacenamiento centralizado se pueden calendarizar los respaldos conforme las necesidades y ejecutarse de manera automática.
- ◆ Se elimina la gestión de cintas y con esto la dependencia en sitio de un recurso humano.
- ◆ La información se respalda en línea, esto impacta de manera favorable en la disponibilidad del servicio porque no se requiere dar de baja las aplicaciones y bases de datos para tener un respaldo confiable.
- ◆ Ante una situación de desastre la restauración de información es más sencilla y eficiente.

**4.2.8.2 Propuesta de políticas de respaldo**

Hostname	Calendarización			Periodo de retención		
	SO	BD	APPL	SO	BD	APPL
backserver1	mensual	1 Full diario Archive logs c/8 hrs	1 full semanal	3 meses	3 meses	3 meses
backserver2	mensual	1 Full diario Archive logs c/8 hrs	1 full semanal	3 meses	3 meses	3 meses
webserver1	mensual	N/A	1full semanal	3 meses	3 meses	3 meses
webserver2	mensual	N/A	1full semanal	3 meses	3 meses	3 meses
webserver3	mensual	N/A	1full semanal	3 meses	3 meses	3 meses
webserver4	mensual	N/A	1full semanal	3 meses	3 meses	3 meses

### 4.3 Plan de Trabajo. Actividades realizadas en la fase de implementación.

Conforme la propuesta de actualización tecnológica antes mencionada, el plan de trabajo que se realizó en la segunda fase consistió en:

ID	Task Name	Duration	Start	Finish	% Comp.
1	<b>Actividades de Implantación del Proyecto</b>	<b>148 days</b>	<b>Mon 03/01/05</b>	<b>Wed 27/07/05</b>	<b>100%</b>
2	Kick Off meeting de Implantación	1 day	Mon 03/01/05	Mon 03/01/05	100%
3	<b>Adecuación de Sitio y llegada de equipo a Sitio</b>	<b>56 days</b>	<b>Mon 03/01/05</b>	<b>Mon 21/03/05</b>	<b>100%</b>
4	Revisión de sitio	1 day	Tue 04/01/05	Tue 04/01/05	100%
5	Llegada del equipo a sitio	30 days	Mon 03/01/05	Fri 11/02/05	100%
6	Revisión de inventario	1 day	Mon 14/02/05	Mon 14/02/05	100%
7	Cierre de inventario	1 day	Tue 15/02/05	Tue 15/02/05	100%
8	<b>Instalación Física</b>	<b>11 days</b>	<b>Thu 03/03/05</b>	<b>Thu 17/03/05</b>	<b>100%</b>
9	Rack's	2 days	Thu 03/03/05	Fri 04/03/05	100%
10	Servidores back-end	5 days	Thu 03/03/05	Wed 09/03/05	100%
11	Servidores front-end	5 days	Thu 03/03/05	Wed 09/03/05	100%
12	Arreglo de discos	5 days	Thu 03/03/05	Wed 09/03/05	100%
13	Balanceadores de carga	3 days	Fri 04/03/05	Tue 08/03/05	100%
14	switch de red	2 days	Wed 09/03/05	Thu 10/03/05	100%
15	<b>Interconexión física entre elementos</b>	<b>5 days</b>	<b>Fri 11/03/05</b>	<b>Thu 17/03/05</b>	<b>100%</b>
16	Etiquetado y cableado interno de la solución	5 days	Fri 11/03/05	Thu 17/03/05	100%
17	Retiro de material	2 days	Fri 18/03/05	Mon 21/03/05	100%
18	<b>Fuerza</b>	<b>41 days</b>	<b>Wed 05/01/05</b>	<b>Wed 02/03/05</b>	<b>100%</b>
19	Solicitud de insumos de fuerza	1 day	Wed 05/01/05	Wed 05/01/05	100%
20	Instalación y asignación de posiciones de fuerza	40 days	Thu 06/01/05	Wed 02/03/05	100%
21	<b>Cableado estructurado de datos</b>	<b>27 days</b>	<b>Wed 05/01/05</b>	<b>Thu 10/02/05</b>	<b>100%</b>
22	Solicitud de Segmentos IP / Anchos de Banda	1 day	Wed 05/01/05	Wed 05/01/05	100%
23	Asignación de Segmentos / Anchos de Banda	5 days	Thu 06/01/05	Wed 12/01/05	100%
24	Generación de OT de Redes (enrutamientos, segmentos, nuevos enlaces)	5 days	Thu 13/01/05	Wed 19/01/05	100%
25	Ejecución de OT de Redes	5 days	Thu 20/01/05	Wed 26/01/05	100%
26	Solicitud de cableados de Red de datos	1 day	Thu 27/01/05	Thu 27/01/05	100%
27	Cableado de Red de Datos	10 days	Fri 28/01/05	Thu 10/02/05	100%

ID	Task Name	Duration	Start	Finish	% Comp.
1	<b>Actividades de Implantación del Proyecto</b>	<b>148 days</b>	<b>Mon 03/01/05</b>	<b>Wed 27/07/05</b>	<b>100%</b>
2	Kick Off meeting de Implantación	1 day	Mon 03/01/05	Mon 03/01/05	100%
3	<b>Adecuación de Sitio y llegada de equipo a Sitio</b>	<b>56 days</b>	<b>Mon 03/01/05</b>	<b>Mon 21/03/05</b>	<b>100%</b>
18	<b>Fuerza</b>	<b>41 days</b>	<b>Wed 05/01/05</b>	<b>Wed 02/03/05</b>	<b>100%</b>
21	<b>Cableado estructurado de datos</b>	<b>27 days</b>	<b>Wed 05/01/05</b>	<b>Thu 10/02/05</b>	<b>100%</b>
28	<b>Configuración de HW y SW de la nueva arquitectura con base a los requerimientos y diseño</b>	<b>94 days</b>	<b>Fri 18/03/05</b>	<b>Wed 27/07/05</b>	<b>100%</b>
29	<b>Configuración de servidores Back-end (2 servidores)</b>	<b>23 days</b>	<b>Fri 18/03/05</b>	<b>Tue 19/04/05</b>	<b>100%</b>
30	Configuración SO	5 days	Fri 18/03/05	Thu 24/03/05	100%
31	Particionamiento de arreglo de discos	5 days	Fri 18/03/05	Thu 24/03/05	100%
32	Instalación y configuración de software base: Oracle, Veritas cluster, Veritas Volumen Manager	3 days	Fri 25/03/05	Tue 29/03/05	100%
33	Configuración de BD	5 days	Wed 30/03/05	Tue 05/04/05	100%
34	Instalación y configuración del software aplicativo requerido.	5 days	Wed 06/04/05	Tue 12/04/05	100%
35	Configuración de Alta Disponibilidad	5 days	Wed 13/04/05	Tue 19/04/05	100%
36	<b>Configuración de servidores Front-end (4 servidores)</b>	<b>21 days</b>	<b>Wed 20/04/05</b>	<b>Wed 18/05/05</b>	<b>100%</b>
37	Configuración SO	5 days	Wed 20/04/05	Tue 26/04/05	100%
38	Particionamiento de arreglo de discos	5 days	Wed 20/04/05	Tue 26/04/05	100%
39	Instalación y configuración de software base: Veritas Volumen Manager	3 days	Wed 27/04/05	Fri 29/04/05	100%
40	<b>Instalación y configuración del dominio Weblogic Server en alta disponibilidad</b>	<b>11 days</b>	<b>Mon 02/05/05</b>	<b>Mon 16/05/05</b>	<b>100%</b>
41	Instalación de Weblogic Server	3 days	Mon 02/05/05	Wed 04/05/05	100%
42	Configuración de nodo admin	1 day	Thu 05/05/05	Thu 05/05/05	100%
43	Creación y configuración de nodos administrados "instancias"	2 days	Fri 06/05/05	Mon 09/05/05	100%
44	Creación y configuración de cluster's	2 days	Tue 10/05/05	Wed 11/05/05	100%
45	Configuración de Node Manager's	3 days	Thu 12/05/05	Mon 16/05/05	100%
46	Instalación y configuración de webapps: aplicaciones masivas y corporativas	5 days	Thu 12/05/05	Wed 18/05/05	100%
47	<b>Configuración de Balanceadores de Carga</b>	<b>5 days</b>	<b>Thu 19/05/05</b>	<b>Wed 25/05/05</b>	<b>100%</b>
48	Configuración en alta disponibilidad	3 days	Thu 19/05/05	Mon 23/05/05	100%
49	Creación y configuración de pools de servidores	2 days	Tue 24/05/05	Wed 25/05/05	100%
50	Creación y configuración de servidores virtuales	2 days	Tue 24/05/05	Wed 25/05/05	100%
51	Creación y configuración de NAT's	2 days	Tue 24/05/05	Wed 25/05/05	100%
52	<b>Configuración con otros sistemas</b>	<b>5 days</b>	<b>Thu 26/05/05</b>	<b>Wed 01/06/05</b>	<b>100%</b>
53	Sistemas de corbo	5 days	Thu 26/05/05	Wed 01/06/05	100%
54	Sistemas de facturación	5 days	Thu 26/05/05	Wed 01/06/05	100%
55	Otros sistemas	5 days	Thu 26/05/05	Wed 01/06/05	100%
56	<b>Implementación de la solución de monitoreo</b>	<b>45 days</b>	<b>Thu 26/05/05</b>	<b>Wed 27/07/05</b>	<b>100%</b>
57	Clasificación de alarmas	15 days	Thu 26/05/05	Wed 15/06/05	100%
58	Instalación y configuración de los agentes de HPOV para monitoreo	10 days	Thu 16/06/05	Wed 29/06/05	100%
59	Instalación y configuración de Smart Plug in de Oracle y BEA Weblogic Server	10 days	Thu 30/06/05	Wed 13/07/05	100%
60	Configuración de monitoreo para aplicaciones y desarrollos internos "in-house".	10 days	Thu 14/07/05	Wed 27/07/05	100%
61	<b>Integración a la SAN de respaldos</b>	<b>20 days</b>	<b>Thu 26/05/05</b>	<b>Wed 22/06/05</b>	<b>100%</b>
62	<b>Interconexión de equipos a SAN</b>	<b>10 days</b>	<b>Thu 26/05/05</b>	<b>Wed 08/06/05</b>	<b>100%</b>
63	Instalación de tarjetas y configuración de drivers	10 days	Thu 26/05/05	Wed 08/06/05	100%
64	Instalación de fibra óptica	5 days	Thu 26/05/05	Wed 01/06/05	100%
65	Configuración de puertos en SAN	5 days	Thu 09/06/05	Wed 15/06/05	100%
66	Configuración de pilíticas de backup	5 days	Thu 16/06/05	Wed 22/06/05	100%

ID	Task Name	Duration	Start	Finish	% Comp.
1	<b>Actividades de Implantación del Proyecto</b>	<b>148 days</b>	<b>Mon 03/01/05</b>	<b>Wed 27/07/05</b>	<b>100%</b>
2	Kick Off meeting de Implantación	1 day	Mon 03/01/05	Mon 03/01/05	100%
3	<b>Adecuación de Sitio y llegada de equipo a Sitio</b>	<b>56 days</b>	<b>Mon 03/01/05</b>	<b>Mon 21/03/05</b>	<b>100%</b>
18	<b>Fuerza</b>	<b>41 days</b>	<b>Wed 05/01/05</b>	<b>Wed 02/03/05</b>	<b>100%</b>
21	<b>Cableado estructurado de datos</b>	<b>27 days</b>	<b>Wed 05/01/05</b>	<b>Thu 10/02/05</b>	<b>100%</b>
28	<b>Configuración de HW y SW de la nueva arquitectura con base a los requerimientos y diseño</b>	<b>94 days</b>	<b>Fri 18/03/05</b>	<b>Wed 27/07/05</b>	<b>100%</b>
67	<b>Cursos y/o Workshops técnicos</b>	<b>43 days</b>	<b>Thu 28/07/05</b>	<b>Mon 26/09/05</b>	<b>100%</b>
68	Definición de Capacitación y Participantes	3 days	Thu 28/07/05	Mon 01/08/05	100%
69	Entrega de temario por parte del proveedor para el Workshop	5 days	Tue 02/08/05	Mon 08/08/05	100%
70	Cierre del temario por parte de Telcel para el Workshop	5 days	Tue 09/08/05	Mon 15/08/05	100%
71	Workshop técnico	10 days	Tue 16/08/05	Mon 29/08/05	100%
72	Capacitación Técnica	20 days	Tue 30/08/05	Mon 26/09/05	100%

ID	Task Name	Duration	Start	Finish	% Comp.
1	<b>Actividades de Implantación del Proyecto</b>	<b>148 days</b>	<b>Mon 03/01/05</b>	<b>Wed 27/07/05</b>	<b>100%</b>
2	Kick Off meeting de Implantación	1 day	Mon 03/01/05	Mon 03/01/05	100%
3	<b>Adecuación de Sitio y llegada de equipo a Sitio</b>	<b>56 days</b>	<b>Mon 03/01/05</b>	<b>Mon 21/03/05</b>	<b>100%</b>
18	<b>Fuerza</b>	<b>41 days</b>	<b>Wed 05/01/05</b>	<b>Wed 02/03/05</b>	<b>100%</b>
21	<b>Cableado estructurado de datos</b>	<b>27 days</b>	<b>Wed 05/01/05</b>	<b>Thu 10/02/05</b>	<b>100%</b>
28	<b>Configuración de HW y SW de la nueva arquitectura con base a los requerimientos y diseño</b>	<b>94 days</b>	<b>Fri 18/03/05</b>	<b>Wed 27/07/05</b>	<b>100%</b>
67	<b>Cursos y/o Workshops técnicos</b>	<b>43 days</b>	<b>Thu 28/07/05</b>	<b>Mon 26/09/05</b>	<b>100%</b>
73	<b>Aceptación de la plataforma</b>	<b>28 days</b>	<b>Tue 30/08/05</b>	<b>Thu 06/10/05</b>	<b>100%</b>
74	<b>Actividades de ATPs</b>	<b>28 days</b>	<b>Tue 30/08/05</b>	<b>Thu 06/10/05</b>	<b>100%</b>
75	<b>Entrega de documentos ATP's por parte del proveedor a la empresa</b>	<b>5 days</b>	<b>Tue 30/08/05</b>	<b>Mon 05/09/05</b>	<b>100%</b>
76	ATP Físico: inventario de equipo, intalación de fuerza y cableado estructurado	5 days	Tue 30/08/05	Mon 05/09/05	100%
77	ATP de servidores a nivel Hardware y SO	5 days	Tue 30/08/05	Mon 05/09/05	100%
78	ATP de conectividad a nivel red de datos	5 days	Tue 30/08/05	Mon 05/09/05	100%
79	ATP de BD	5 days	Tue 30/08/05	Mon 05/09/05	100%
80	ATP de Alta Disponibilidad en servidores back-end	5 days	Tue 30/08/05	Mon 05/09/05	100%
81	ATP de procesos aplicativos en back.end	5 days	Tue 30/08/05	Mon 05/09/05	100%
82	ATP de Alta Disponibilidad en servidores front-end con weblogic server	5 days	Tue 30/08/05	Mon 05/09/05	100%
83	ATP de Interconexión con otros sistemas	5 days	Tue 30/08/05	Mon 05/09/05	100%
84	ATP de funcionalidad del sistema a nivel usuario final (casos de tráfico)	5 days	Tue 30/08/05	Mon 05/09/05	100%
85	ATP de balanceo de carga	5 days	Tue 30/08/05	Mon 05/09/05	100%
86	ATP de Cobro y Facturación	5 days	Tue 30/08/05	Mon 05/09/05	100%
87	ATP de Monitoreo	5 days	Tue 30/08/05	Mon 05/09/05	100%
88	ATP de Respaldo	5 days	Tue 30/08/05	Mon 05/09/05	100%
89	Revisión y comentarios al ATP por parte de la empresa	2 wks	Tue 30/08/05	Mon 12/09/05	100%
90	Cierre y entrega de ATP's finales a la empresa	1 wk	Tue 30/08/05	Mon 05/09/05	100%
91	Pruebas internas del proveedor (HW y funcionales)	2 wks	Tue 30/08/05	Mon 12/09/05	100%
92	<b>Aplicación de ATP's para aceptación por parte de la empresa</b>	<b>18 days</b>	<b>Tue 13/09/05</b>	<b>Thu 06/10/05</b>	<b>100%</b>
93	ATP Físico: inventario de equipo, intalación de fuerza y cableado estructurado	1 day	Tue 13/09/05	Tue 13/09/05	100%
94	ATP de servidores a nivel Hardware y SO	2 days	Wed 14/09/05	Thu 15/09/05	100%
95	ATP de conectividad a nivel red de datos	1 day	Fri 16/09/05	Fri 16/09/05	100%
96	ATP de BD	1 day	Mon 19/09/05	Mon 19/09/05	100%
97	ATP de Alta Disponibilidad en servidores back-end	1 day	Tue 20/09/05	Tue 20/09/05	100%
98	ATP de procesos aplicativos en back.end	1 day	Wed 21/09/05	Wed 21/09/05	100%
99	ATP de Alta Disponibilidad en servidores front-end con weblogic server	1 day	Thu 22/09/05	Thu 22/09/05	100%
100	ATP de Interconexión con otros sistemas	2 days	Fri 23/09/05	Mon 26/09/05	100%
101	ATP de funcionalidad del sistema a nivel usuario final (casos de tráfico)	2 days	Tue 27/09/05	Wed 28/09/05	100%
102	ATP de balanceo de carga	1 day	Thu 29/09/05	Thu 29/09/05	100%
103	ATP de Cobro y Facturación	2 days	Fri 30/09/05	Mon 03/10/05	100%
104	ATP de Monitoreo	1 day	Tue 04/10/05	Tue 04/10/05	100%
105	ATP de Respaldo	2 days	Wed 05/10/05	Thu 06/10/05	100%

ID	Task Name	Duration	Start	Finish	% Comp.
1	<b>Actividades de Implantación del Proyecto</b>	<b>148 days</b>	<b>Mon 03/01/05</b>	<b>Wed 27/07/05</b>	<b>100%</b>
2	Kick Off meeting de Implantación	1 day	Mon 03/01/05	Mon 03/01/05	100%
3	<b>Adecuación de Sitio y llegada de equipo a Sitio</b>	<b>56 days</b>	<b>Mon 03/01/05</b>	<b>Mon 21/03/05</b>	<b>100%</b>
18	<b>Fuerza</b>	<b>41 days</b>	<b>Wed 05/01/05</b>	<b>Wed 02/03/05</b>	<b>100%</b>
21	<b>Cableado estructurado de datos</b>	<b>27 days</b>	<b>Wed 05/01/05</b>	<b>Thu 10/02/05</b>	<b>100%</b>
28	<b>Configuración de HW y SW de la nueva arquitectura con base a los requerimientos y diseño</b>	<b>94 days</b>	<b>Fri 18/03/05</b>	<b>Wed 27/07/05</b>	<b>100%</b>
67	<b>Cursos y/o Workshops técnicos</b>	<b>43 days</b>	<b>Thu 28/07/05</b>	<b>Mon 26/09/05</b>	<b>100%</b>
73	<b>Aceptación de la plataforma</b>	<b>28 days</b>	<b>Tue 30/08/05</b>	<b>Thu 06/10/05</b>	<b>100%</b>
106	<b>Entregables de Proyecto</b>	<b>8 days</b>	<b>Fri 07/10/05</b>	<b>Tue 18/10/05</b>	<b>100%</b>
107	<b>Memorias Técnicas</b>	<b>8 days</b>	<b>Fri 07/10/05</b>	<b>Tue 18/10/05</b>	<b>100%</b>
108	Recepción de Memoria Técnica	1 day	Fri 07/10/05	Fri 07/10/05	100%
109	Revisión de Memoria Técnica	1 wk	Fri 07/10/05	Thu 13/10/05	100%
110	Cierre de Memoria Técnica	3 days	Fri 14/10/05	Tue 18/10/05	100%
111	ATPs Firmados	1 day	Fri 07/10/05	Fri 07/10/05	100%
112	<b>Cierre del Proyecto</b>	<b>26 days</b>	<b>Fri 07/10/05</b>	<b>Fri 11/11/05</b>	<b>100%</b>
113	Definición de Plan de Babysitting, Estabilización y Transferencia	3 days	Fri 07/10/05	Tue 11/10/05	100%
114	Puesta en operación (Cutover) o Liberación técnica	1 day	Wed 12/10/05	Wed 12/10/05	100%
115	<b>Babysitting, Estabilización y Transferencia de Conocimientos</b>	<b>20 days</b>	<b>Thu 13/10/05</b>	<b>Wed 09/11/05</b>	<b>100%</b>
116	Ejecución de Plan	20 days	Thu 13/10/05	Wed 09/11/05	100%
117	Liberación y Entrega formal al área de O&M de la empresa	1 day	Thu 10/11/05	Thu 10/11/05	100%
118	Actas de aceptación	1 day	Fri 11/11/05	Fri 11/11/05	100%

## **5 Capítulo 5**

---

### Mejoras en el proceso de operación

Antes de pasar a la propuesta de mejora en los procesos operativos para la prestación de servicios de valor agregado objeto de este trabajo, es importante mencionar que la empresa donde laboro, está comprometida con la satisfacción del cliente y por ello siempre busca la excelencia operacional a través del trabajo en equipo y dentro de un proceso de mejora continua.

Dado lo anterior, nuestras actividades diarias se rigen por el sistema de calidad basado en la norma ISO 9001:2000.

Partiendo de esta premisa y para entender el desarrollo de este trabajo, a continuación se presentan una breve introducción a los sistemas de calidad y a la familia de normas ISO 9000:2000.

## 5.1 Introducción a los sistemas de calidad.

Para comprender mejor el concepto de sistema de calidad debemos empezar por definir qué es un sistema y qué es calidad.

**Sistema.** “Conjunto de partes mutuamente interrelacionadas o que interactúan”.<sup>31</sup>

Es un conjunto de partes coordinadas que operan para lograr los objetivos generales de un todo.

### 5.1.1 Características de los sistemas.

- ◆ Orientación hacia un fin común.
- ◆ Límites establecidos.
- ◆ Dependencia entre elementos.
- ◆ El sistema tiene que ser unicidad. “Comprendido como un todo”
- ◆ Funciones independientes de los elementos.
- ◆ El cambio de los sistemas como una constante.
- ◆ Interrelación entre los elementos

**Calidad.** “Grado en que un conjunto de características inherentes cumple con las necesidades o expectativas establecidas, generalmente implícitas u obligatorias”.<sup>32</sup>

### 5.1.2 La percepción de calidad por parte del cliente.

Cada cliente tiene una percepción particular de la calidad que depende de sus necesidades en un determinado momento.

Los clientes buscan que el producto sirva, que la organización de un buen servicio y a cambio de ello pagar un precio justo.

---

<sup>31</sup> Fundamentos y vocabulario ISO 9000:2000.

<sup>32</sup> Fundamentos y vocabulario ISO 9000:2000.

Las dimensiones claves de la calidad son:

- ◆ El desempeño.
- ◆ Las características (adicionales).
- ◆ La confiabilidad.
- ◆ La conformidad.
- ◆ La durabilidad.
- ◆ La capacidad de servicio.
- ◆ La estética.

Cada una de estas definiciones influye en la percepción de los clientes respecto a la calidad.

### **5.1.3 Definición de sistema de gestión de la calidad.**

“Sistema de gestión para dirigir y controlar una organización con respecto a la calidad”.<sup>33</sup>

Las organizaciones buscan que sus productos sean preferidos por el usuario. Cada día los clientes son más exigentes y la calidad de los productos define el éxito de las organizaciones.

Otra definición:

“Es mantener la mejora de eficacia y eficiencia en la operación e integridad de las organizaciones y no afectar a los clientes en la incorporación de los cambios”.

---

<sup>33</sup> Fundamentos y vocabulario ISO 9000:2000.

#### 5.1.4 Razones de aplicación de los sistemas de calidad.

- ◆ **Mejora de eficacia y eficiencia.** Las organizaciones modernas enfrentan el reto de tener que ser más eficaces y eficientes para satisfacer a los clientes.
- ◆ **Crecimiento.** Para que la organización se desarrolle ordenadamente sin crisis prolongadas. Como por ejemplo para ser una empresa pequeña a mediana, o de mediana a grande.
- ◆ **Complejidad del trabajo.** Cuando en la organización prevalecen contingencias, situaciones complejas y procesos técnicos y administrativos con flujos de comunicación difíciles de comprender.
- ◆ **Diversidad y cambios rápidos.** Cuando la organización proporciona una gran variedad de productos asociados a cambios dinámicos de sus procesos de producción y servicio para afrontar la globalización y la competencia.

#### 5.1.5 La organización y los sistemas.

##### 5.1.5.1 ¿Qué es una organización?

Conjunto de personas e instalaciones con una disposición de responsabilidades, autoridades y relaciones”.<sup>34</sup>

##### 5.1.5.2 Las organizaciones como un sistema.

Las organizaciones se entienden como sistemas, condicionados a las influencias dinámicas del medio donde operan y son sometidos a un proceso de cambio constante.

Las organizaciones no están libres de influencias externas, cuentan con interfaces con otras organizaciones tales como: proveedores de materias primas, la comunidad local, clientes, dependencias gubernamentales, estatales y federales.

---

<sup>34</sup> Fundamentos y vocabulario ISO 9000:2000.

La misión de la organización determina los límites del sistema, estableciendo sus productos y servicios que ofrece.

El personal que pertenece a una organización necesita conocer los objetivos, disponer de los recursos y mantener una dirección consistente para lograr la eficacia y eficiencia.

Los sistemas funcionan por ciclos y tienen retroalimentación.

### **5.1.5.3 La organización como un conjunto de procesos.**

Los productos y servicios determinan el conjunto de procesos que integran una organización, estos procesos se pueden clasificar en operativos y de apoyo.

En una organización existe un conjunto de procesos, cuya interrelación se entiende, en que los productos de unos son los insumos de otros, hasta lograr el producto final.

La combinación adecuada de los procesos nos lleva a la obtención del producto final. El objetivo de todos los procesos es la aceptación de los bienes o servicios en el mercado.

### **5.1.5.4 La organización y la necesidad de cambio.**

Cualquier organización que no se adapte a los cambios tiende a moverse hacia un estado caótico y desorganizado.

La organización que tiene la capacidad de crecer ordenadamente e incorporar nuevas tecnologías en forma continua es competitiva. Una organización se adapta a su ambiente mediante el cambio de sus procesos.

A través del proceso de retroalimentación la organización recibe de su medio externo la información que le permite ajustarse y cambiar.

La organización tiene que recibir recursos suficientes que le permitan mantener su operación e incorporar los cambios.

## 5.1.6 La organización como una red de procesos.

### 5.1.6.1 ¿Qué es un proceso?

“Conjunto de actividades mutuamente relacionadas o que interactúan, las cuales transforman entradas en salidas”.<sup>35</sup>



Figura 5.1 Concepto de proceso.

### 5.1.6.2 Características de los procesos.

En cada proceso se identifican tres aspectos tales como:

- ◆ Insumos
- ◆ Procesos de transformación y
- ◆ Producto final.

Los insumos pueden ser: los materiales, equipos, información, gente, dinero, o condiciones ambientales necesarias para llevar a cabo el proceso.

El proceso de transformación adiciona valor a las entradas o insumos.

---

<sup>35</sup> Fundamentos y vocabulario ISO 9000:2000

Los resultados de un proceso pueden ser: el producto o el servicio que es creado por el proceso, el cual es entregado al cliente.



Figura 5.2 Las seis M's de un proceso.

### 5.1.6.3 ¿Qué es una red de procesos?

Una red de procesos es una representación gráfica de un conjunto de procesos. Por ejemplo: En una organización una red de procesos de trabajo es similar a la red de carreteras.

### 5.1.6.4 Documentación de los procesos.

Los procesos se documentan para facilitar la ejecución del trabajo, unificar criterios y prevenir errores.

Un proceso según su alcance puede ser documentado en uno o varios procesos.

Cuando la complejidad y el riesgo del trabajo lo ameritan a partir de un procedimiento es posible documentar instrucciones específicas para una determinada actividad.

## 5.1.7 La gestión de la calidad.

### 5.1.7.1 ¿Qué es gestión?

“Actividades coordinadas para dirigir y controlar una organización”.<sup>36</sup>

### 5.1.7.2 El enfoque de gestión dentro de la organización.

El proceso de la gestión dentro de una organización es realizado mediante la aplicación sistemática del ciclo: Planear, Ejecutar, Medir y Mejorar.

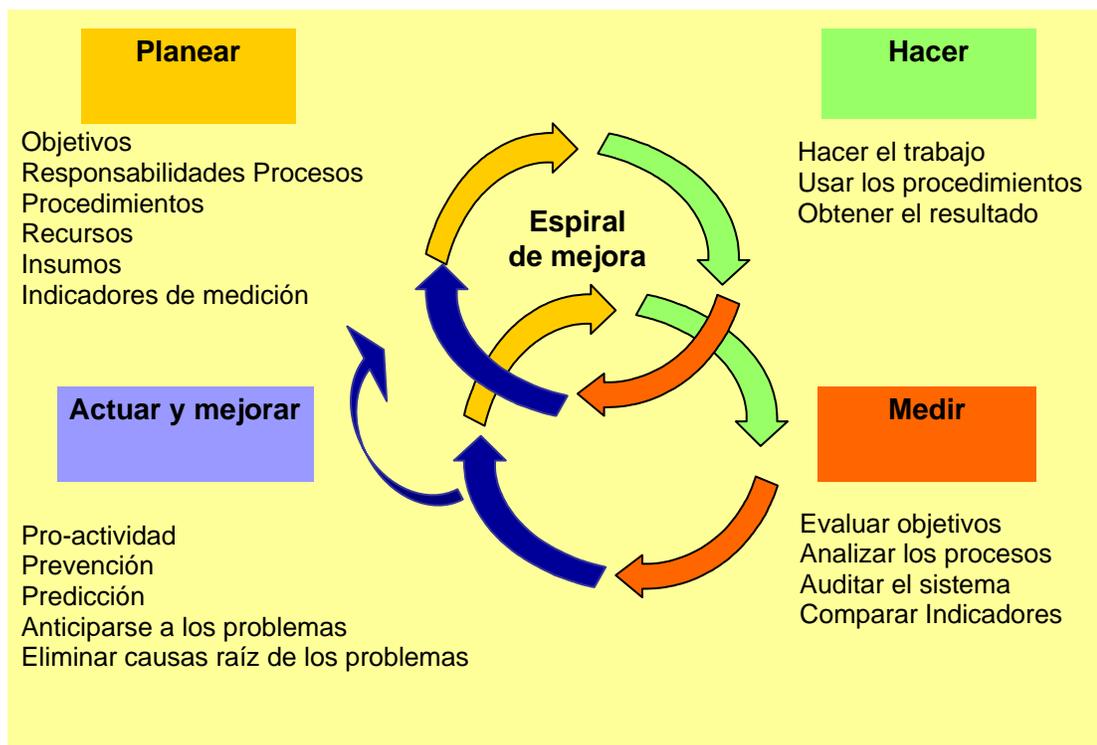


Figura 5.3 Ciclo de mejora continua del sistema de gestión de calidad.

---

<sup>36</sup> Fundamentos y vocabulario ISO 9000:2000.

- ◆ **Planear.** Tiene como finalidad determinar los objetivos a lograr por una organización.
- ◆ **Ejecutar.** Consiste en definir la manera de cómo, quién y con que se lograrán los objetivos definidos para la organización.
- ◆ **Medir.** Es necesario que la organización tenga definidos medios y mecanismos que le permitan conocer si la manera como se han realizado las actividades están siendo congruentes con lo planeado y si estamos en ruta de alcanzar el objetivo de manera eficaz y eficiente. En caso contrario, será necesario realizar las acciones correctivas o preventivas necesarias para asegurar lograr el objetivo.
- ◆ **Mejorar.** Una vez logrados los objetivos planeados es necesario que la organización mejore la manera de hacer las cosas (procesos) y lo que ofrece al usuario final (producto), hasta donde la infraestructura y tecnología disponible en la organización lo permita y con plena satisfacción de las partes interesadas de la misma.

## 5.1.8 Normas ISO 9000:2000

### 5.1.8.1 Definición de ISO



ISO **INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**, es una red de los institutos nacionales de los estándares de 157 países, en base de un miembro por país, con una secretaría central en Ginebra, Suiza, que coordina el sistema.

¿Qué significa ISO 9000?

La denominación ISO-9000 corresponde a un número de identificación que le fue asignado a la familia de normas cuyo tema es la gestión de sistemas de calidad.

### 5.1.8.2 ISO, ¿Para qué?

Una de las razones de aplicación en los sistemas de calidad en las organizaciones es enfrentar el reto de tener que ser más eficaces y eficientes para satisfacer a los clientes.

Las normas de sistema de calidad, establecen buenas prácticas que contribuyen a que una organización trabaje de una manera ordenada y con una orientación a la satisfacción del cliente.

Los beneficios que otorga ISO, es que ha sido implementada en muchas empresas alrededor del mundo, desde grandes corporaciones hasta pequeños negocios todos con un solo objetivo SATISFACER AL CLIENTE.

### 5.1.8.3 Antecedentes de las normas ISO 9000.

La familia de normas ISO 9000, **surge en 1987**, como un esfuerzo orientado a que las organizaciones desarrollen sus sistemas de gestión de calidad en cualquier ámbito de calidad.

El **11 de Diciembre de 1990**, la Dirección de Normas de la Secretaría de Comercio y Fomento Industrial, a través del Diario Oficial de la Federación, aprobó las primeras ocho norma oficiales mexicanas NOM-CC. Con esta acción México, al igual que los países industrializados; adopta el esquema de normalización de la ISO.

En **1994** la familia de normas ISO 9000, es mejorada gracias a la retroalimentación proporcionada por organizaciones de todo el mundo, apoyando la idea de continuar trabajando en la gestión de calidad.

En **Diciembre del año 2000**, se publica una nueva versión cuyo propósito es facilitar el uso de las normas para orientar a las organizaciones en la satisfacción del cliente y la mejora de la eficacia y eficiencia de sus procesos.

#### 5.1.8.4 Familia de normas ISO 9000:2000

- ◆ **ISO 9000:2000.** Sistema de gestión de calidad: Fundamentos y vocabulario.

**Propósito:** Describe los fundamentos de los sistemas de gestión de la calidad y especifica la terminología para los mismos.

- ◆ **ISO 9001:2000.** Sistema de gestión de calidad: Requisitos.

**Propósito:** Especifica los requisitos para los sistemas de gestión de la calidad aplicables a toda la organización que necesite demostrar su capacidad para proporcionar productos que cumplan los requisitos de sus clientes. Su objetivo es orientar a la organización hacia la satisfacción del cliente.

- ◆ **ISO 9004:2000.** Sistema de gestión de calidad: Directrices para la mejora del desempeño.

**Propósito:** El objetivo de esta norma es la mejora del desempeño de la organización y la satisfacción de los clientes y de otras partes interesadas. Proporciona directrices que consideran tanto la eficacia como la eficiencia del sistema de gestión de la calidad.

- ◆ **ISO 19011:2000.** Directrices para las auditorías de sistemas de gestión de calidad y/o medio ambiente.

**Propósito:** Proporciona orientación relativa a la administración de programas de auditorías, ejecución de auditorías y calificación de auditores tanto de sistemas de gestión de la calidad y de gestión ambiental

#### **5.1.8.5 ¿En dónde se aplican las normas ISO?**

Estas normas están diseñadas para diferentes tipos de actividades, y se adaptan a las características de una empresa. Las normas ISO son una guía explícita de lo que buscan las empresas porque se han escrito para que se apliquen a una gran variedad de situaciones.

El propósito de ISO es promover el desarrollo de la estandarización y actividades mundiales relativas a facilitar el comercio internacional de bienes y servicios, así como desarrollar la cooperación intelectual, científica y económica. Los resultados del trabajo técnico de ISO son publicados como estándares internacionales.

#### **5.1.8.6 ¿Que se busca con la familia de normas ISO 9000:2000 en la organización?**

- ◆ Orientar el trabajo de la organización por procesos.
- ◆ Destacar el proceso de mejora continua.
- ◆ Incrementar el compromiso de la alta dirección en el desarrollo y mejora del sistema de gestión de la calidad.
- ◆ Un sistema de gestión de calidad que incluya los requerimientos legales y regulatorios aplicables al producto.
- ◆ Un sistema de gestión de calidad con énfasis en el establecimiento y medición de objetivos en las funciones y niveles relevantes.
- ◆ Medir la satisfacción y/o insatisfacción del cliente.
- ◆ Propiciar la compatibilidad con los sistemas de gestión ambiental.
- ◆ Aplicación de los 8 principios del sistema de gestión de calidad.

### **5.1.9 Los 8 principios de la gestión de la calidad**

Una investigación entre las organizaciones más exitosas a nivel internacional obtuvo como resultado la identificación de 8 principios básicos, todos ellos utilizados como fundamento de la gestión de calidad.

La familia de normas ISO 9000:2000 fue creada tomando como plataforma tanto la versión de ISO-9000: 1994 y mejorando los requisitos para que a través de ellos se utilicen estos 8 principios que son:

- ◆ Enfoque al cliente.
- ◆ Liderazgo.
- ◆ Involucramiento del personal.
- ◆ Enfoque basado en procesos.
- ◆ Enfoque de sistemas para la gestión.
- ◆ Mejora continua.
- ◆ Enfoque basado en hechos para la toma de decisión.
- ◆ Relaciones mutuamente beneficiosas con los proveedores.

#### **5.1.9.1 Enfoque al cliente.**

Las organizaciones dependen de sus clientes y por lo tanto deberían comprender sus necesidades actuales y futuras, satisfacer los requisitos de sus clientes y esforzarse en exceder sus expectativas.

#### **5.1.9.2 Liderazgo.**

Los líderes establecen la unidad de propósito y la orientación de la organización. Ellos deberían crear y mantener un ambiente interno, en el cual el personal pueda llegar a involucrarse totalmente en el logro de los objetivos de la organización.

#### **5.1.9.3 Involucramiento del personal.**

El personal, a todos los niveles, es la esencia de una organización y su total compromiso posibilita que sus habilidades sean usadas para el beneficio de la organización.

#### **5.1.9.4 Enfoque basado en procesos.**

Un resultado deseado se alcanza más eficientemente cuando las actividades y los recursos relacionados se gestionan como un proceso.

#### **5.1.9.5 Enfoque de sistemas para la gestión.**

Identificar, entender y gestionar los procesos interrelacionados como un sistema, contribuye a la eficacia y eficiencia de una organización en el logro de sus objetivos.

#### **5.1.9.6 Mejora continua.**

La mejora continua del desempeño global de la organización debería ser un objetivo permanente de esta. Significa aplicar el ciclo planear, ejecutar, medir y mejorar.

#### **5.1.9.7 Enfoque basado en hechos para la toma de decisión.**

Las decisiones eficaces se basan en el análisis de los datos y en la información. Significa no decidir apoyados en suposiciones.

#### **5.1.9.8 Relaciones mutuamente beneficiosas con los proveedores.**

Una organización y sus proveedores son interdependientes, y una relación mutuamente beneficiosa aumenta la capacidad de ambos para crear valor.

### **5.1.10 ¿Qué es la Famosa Certificación ISO 9000?**

Este sistema de certificación nace para dar respuesta a la necesidad de certificación que tenían aquellas empresas cuyos productos no son certificados, o empresas de servicio. Al mismo tiempo se intenta con ello unificar la forma de evaluación del sistema de calidad de las empresas.

Lograr la certificación no es realmente difícil, es mucho más difícil mantener el sistema y mejorarlo a través del tiempo y la constancia en el uso de las prácticas que están documentadas en el sistema.

#### **5.1.10.1 ¿Cuáles son los beneficios de la certificación?**

Algunos de los beneficios son:

- ◆ Ventajas competitivas.
- ◆ Aumento de eficiencia y eficacia.
- ◆ Reducción de auditorias.
- ◆ Reducción de reclamos de garantías y demandas.
- ◆ Mejoramiento de Control Interno.

## 5.2 Propuesta de mejora en los procesos operativos para prestación de servicios de valor agregado.

Los servicios y sistemas de valor agregado aumentaron rápidamente en la empresa originando así una necesidad de mejora en la organización operativa y administrativa de los mismos. Para cubrir de manera adecuada esta necesidad y cumplir con nuestro objetivo: **“Optimizar los procesos operativos a aplicar en los sistemas de valor agregado, dado al cambio y ofrecimiento constante de nuevos servicios y aplicaciones”**, se propuso tener un equipo de soporte técnico dedicado a la operación y mantenimiento de los sistemas de SVA y un equipo encargado de la gestión de los mismos a nivel corporativo.

El área de Soporte Técnico fue propuesta para cubrir reactivamente cualquier eventualidad de fallas en los sistemas y servicios en horarios 7X24, es decir, durante las 24 horas de los 365 días del año. Dicho equipo de soporte se dispuso con herramientas y procesos definidos para el correcto registro, seguimiento y cierre de fallas presentadas. Entre las responsabilidades de este grupo se encuentran:

- ◆ Atención, seguimiento y solución a reportes de falla críticos, mayores y menores.
- ◆ Escalación de fallas con proveedores de soporte para el seguimiento y solución a las fallas presentadas.
- ◆ Ejecución de órdenes de trabajo correctivo en los sistemas de SVA.
- ◆ Ejecución de órdenes de trabajo planeado en los sistemas de SVA, es decir, alta, baja o cambio en la configuración de los sistemas en producción.

El área Corporativa de Operación fue propuesta para poder cubrir de manera pro-activa, preventiva y predictiva cualquier eventualidad que pudiera presentarse en los sistemas y servicios de valor agregado y ser una interfaz de comunicación entre las áreas de ingeniería y comercial involucradas en la empresa. Entre las responsabilidades de este grupo se encuentran:

- ◆ Definir los requerimientos técnicos de operación para los nuevos servicios y sistemas de valor agregado.

- ◆ Definir los procesos operativos y administrativos a utilizar en los servicios y sistemas de valor agregado.
- ◆ Implementar las acciones de mejora para eliminar la causa raíz de los problemas presentados en la operación de los sistemas de SVA.
- ◆ Planear la ejecución de órdenes de trabajo que impliquen alta, baja o cambio en la configuración de los sistemas de SVA en producción.
- ◆ Planear las rutinas de mantenimiento para seguir una estrategia de prevención en los sistemas de SVA
- ◆ Analizar el desempeño y disponibilidad de los sistemas de SVA para garantizar la disponibilidad y capacidad de las mismas.

### 5.2.1 Organización propuesta en el área operativa.

A continuación se muestra la propuesta de organización para poder tener un mejor soporte operativo y administrativo a nivel correctivo, preventivo y predictivo. Dicha propuesta tuvo como objetivo principal: delimitar y separar los roles y responsabilidades, de tal forma que se puedan atender y solucionar problemas y nuevas necesidades de una manera más eficiente y rápida.

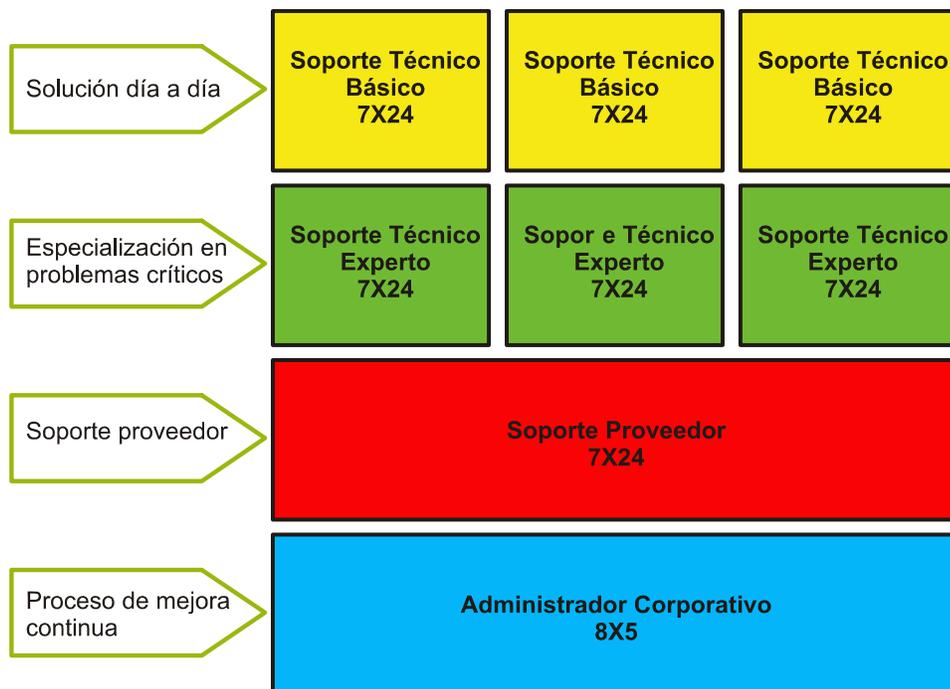


Figura 5.4 Propuesta de organización para soporte operativo y administrativo.

Los Roles de cada posición son los siguientes:

- ◆ **Básico:** Solución de problemas Menores y Mayores.
- ◆ **Experto:** Solución a problemas críticos. Cuenta con experiencia en la solución de problemas críticos y trabaja en equipo con el proveedor de la solución.
- ◆ **Proveedor:** Solución a problemas que el nivel experto no pueda resolver. Cuenta con la mayor experiencia para resolver problemas en el diseño de la solución.
- ◆ **Administrador corporativo:** Aplicación del proceso de mejora continua sobre los sistemas de SVA: Planear, Ejecutar, Medir y Mejorar.

Adicional a esta propuesta de organización en el área de operación, se trabajo en conjunto con otras áreas de la empresa para tener una mejor distribución del trabajo y realizar así las funciones propias de cada área, mejorando dos procesos principales:

- ◆ Proceso de alta, baja o modificación en los sistemas de SVA en producción.
- ◆ Proceso de integración de un nuevo sistema SVA.

### 5.2.2 Proceso de alta, baja o modificación en los sistemas de SVA en producción.

A continuación se describe el proceso para alta, baja o modificación en los sistemas de SVA considerando la mejora en la organización anteriormente descrita y el involucramiento de otras áreas responsables del servicio:

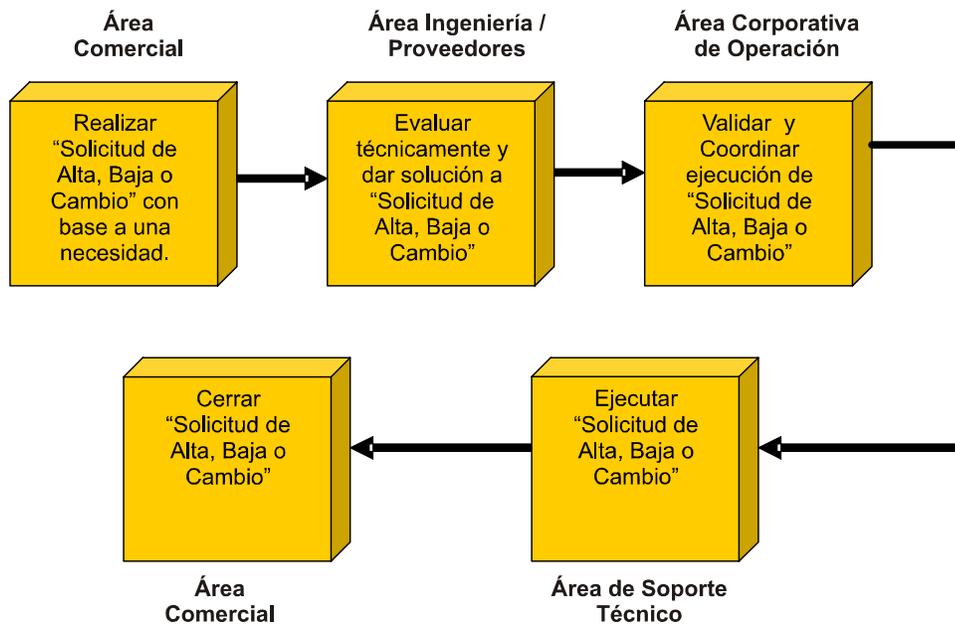


Figura 5.5 Proceso de operación para alta, baja o cambio en los sistemas de SVA en producción.

### 5.2.3 Proceso de integración de un nuevo sistema SVA.

Con el fin de mejorar el proceso de integración de un nuevo sistema SVA y el crecimiento o expansión de las ya existentes, se mejoró el alcance de cada área como a continuación se muestra:

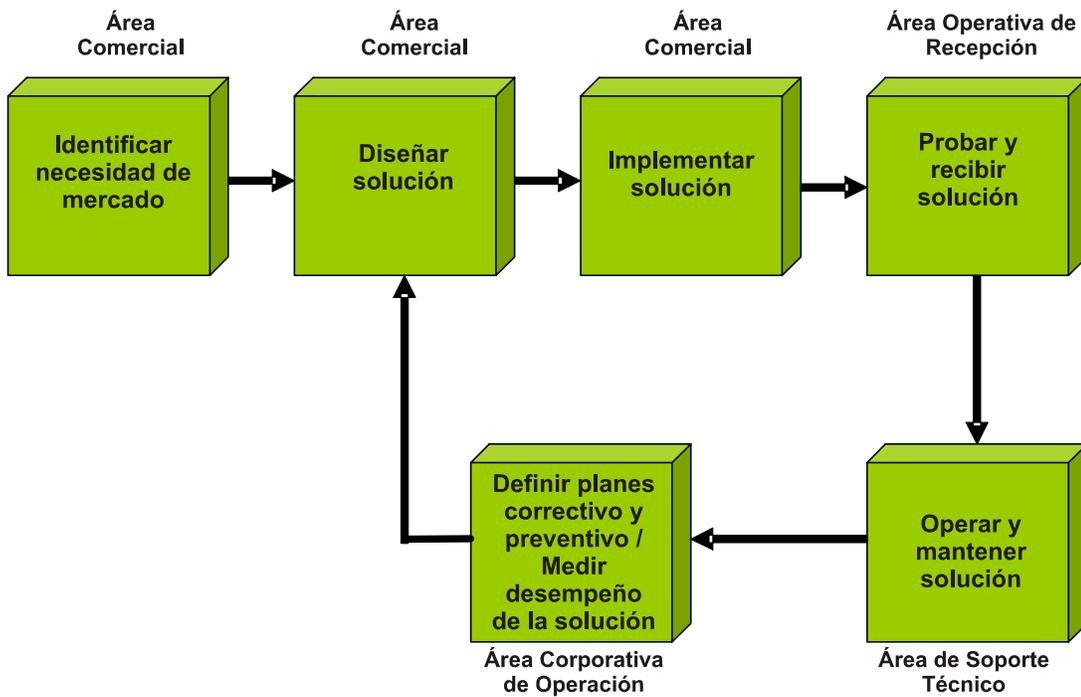


Figura 5.6 Proceso de integración inicial de un nuevo sistema SVA.

## **6 Capítulo 6**

---

Análisis de resultados.

## 6.1 Resultados obtenidos con la actualización tecnológica.

### 6.1.1 Situación inicial.

Como se describió en el capítulo 3, la arquitectura inicial requirió de mejoras en su diseño para el logro de nuestro primer objetivo: “Contar con una solución a nivel hardware/software que permita **alcanzar y mantener la disponibilidad semanal del sistema SVA al 99.9%**”.

La disponibilidad inicial promedio del *sistema SVA* estaba alrededor del 93.6%.

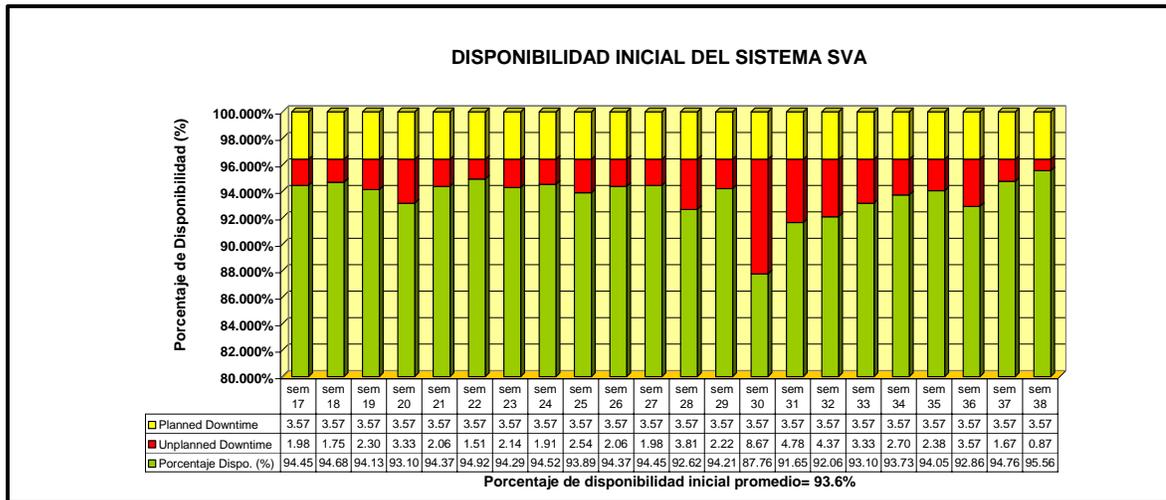


Figura 6.1 Estadística de disponibilidad inicial del *sistema SVA*.

### 6.1.2 Resultados obtenidos después de la implementación de la primera fase.

Con la implementación de la primera fase descrita en el capítulo 4: CAMBIO DE WEB APPLICATION SERVER EN LA ARQUITECTURA INICIAL, los resultados obtenidos fueron inmediatos y reflejaron un aumento en la disponibilidad del *sistema SVA*.

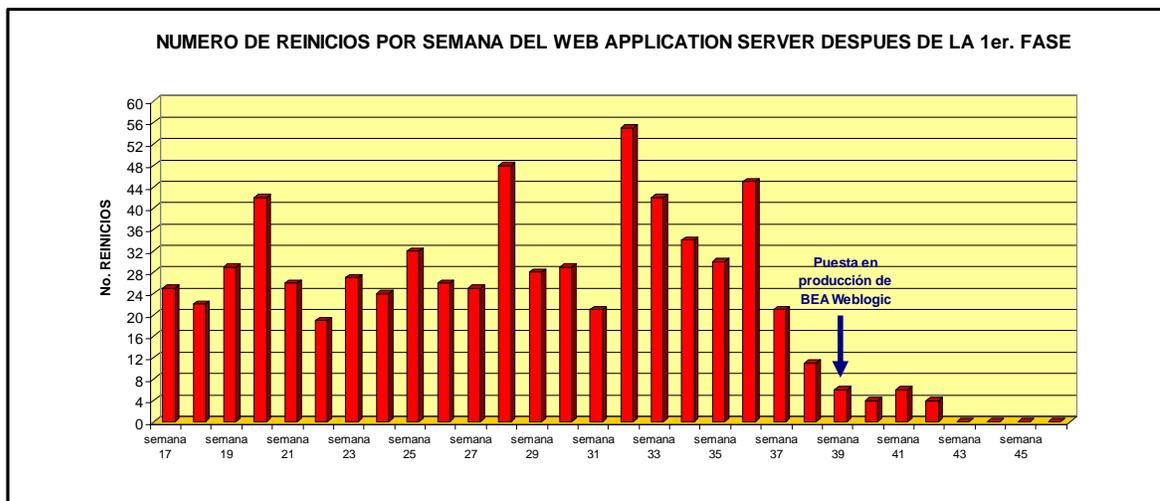


Figura 6.2 Estadística de reinicios por semana después de la 1er. Fase.

Como se puede observar en la estadística de la figura 6.3, la puesta en producción del web application seleccionado: “BEA Weblogic Server (WLS)” reflejó mejoras significativas al *sistema SVA*, es decir, disminuyeron considerablemente los reinicios del web application. Sin embargo, esta implementación requirió de ajustes en los parámetros de configuración del WLS para optimizar su desempeño y reducir a cero el número de reinicios que continuaban afectando la disponibilidad del servicio.

De manera general las modificaciones realizadas al WLS fueron:

- ◆ Aumento de memoria asignada a la instancia de WLS.
- ◆ Implementación de rotación de archivos históricos de accesos a un máximo de tiempo de 24 horas, así como cambiar la ruta de colocación de los archivos.
- ◆ Actualización de la versión del Java Development Kit.
- ◆ Instalación de los parches recomendados para Solaris con la nueva versión de Java.

La implementación de la mejora en WLS se llevó a cabo en las siguientes 3 semanas de su puesta en producción, presentando mejoras significativas y estabilizando el desempeño del WLS lo cual impacto directamente en la mejora de la disponibilidad del sistema.

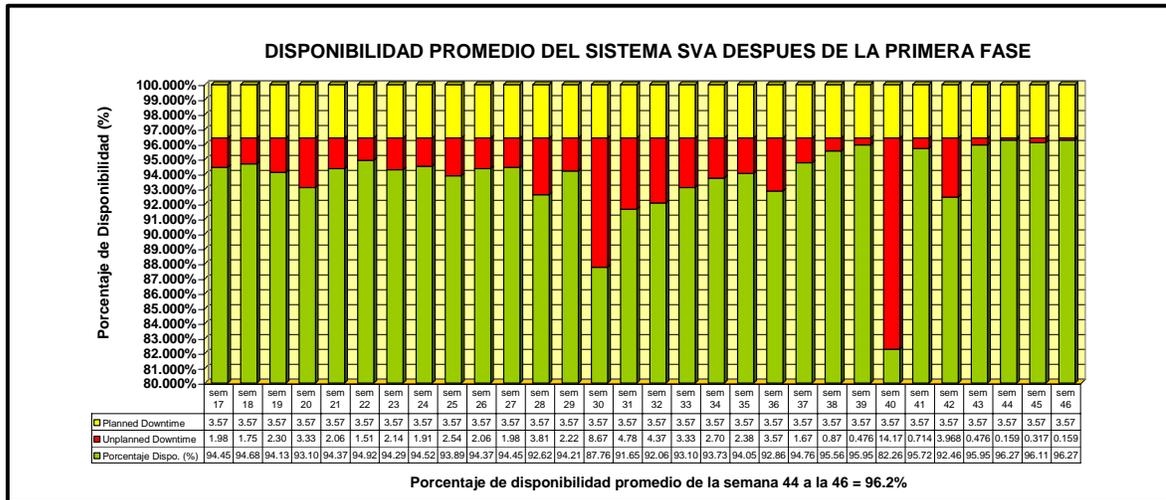


Figura 6.3 Estadística de disponibilidad del sistema SVA después de la 1er. Fase.

Con la conclusión de la primera fase, la disponibilidad promedio que se alcanzó en el sistema SVA fue del 96%. Lo anterior, debido a que si bien se logró la estabilidad del web application server utilizado, la saturación de los recursos de memoria y cpu que se presenciaron en la arquitectura inicial impactó en el correcto funcionamiento del proceso aplicativo dedicado a los servicios masivos y corporativos. Adicionalmente, también impactó en la disponibilidad del sistema SVA, el “downtime” que se generaba semanalmente por el respaldo de la información.

### 6.1.3 Resultados obtenidos después de la implementación de la segunda fase.

Con la implementación de la segunda fase descrita en el capítulo 4: DISEÑO E IMPLEMENTACIÓN DE LA NUEVA ARQUITECTURA, los resultados obtenidos impactaron directamente en el logro del objetivo: “Contar con una solución a nivel hardware/software que permita **alcanzar y mantener la disponibilidad semanal del sistema SVA al 99.9%**” reflejando un aumento en la disponibilidad del *sistema SVA* como a continuación se describe.

#### 6.1.3.1 Resultados de la implementación de alta disponibilidad en configuración “activo & activo”, distribución de tráfico y de procesos aplicativos.

Con la implementación de la alta disponibilidad en configuración “activo & activo” de los servidores y aplicaciones involucradas, la distribución del tráfico entrante al *sistema SVA* usando balanceadores de carga y la distribución de los procesos aplicativos, se logró balancear el procesamiento de las peticiones en cada servidor y por consiguiente aprovechar al máximo los recursos de hardware y de software adquiridos.

A continuación se muestran estadísticas de consumo de cpu y memoria antes y después de la implementación:

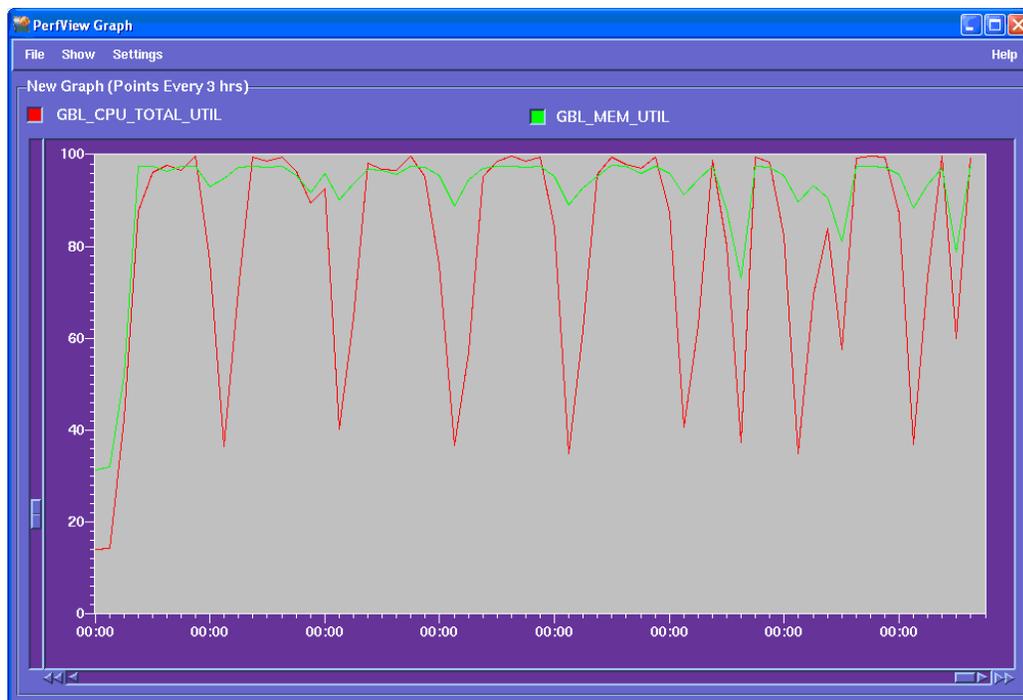


Figura 6.4 Comportamiento de consumo de cpu y memoria en servidor1 - Nodo Activo. Arquitectura Inicial



**Figura 6.5 Consumo de cpu y memoria en servidor2 - Nodo Pasivo.  
Arquitectura Inicial**

Como puede apreciarse en la figura 6.4 y 6.5 los recursos del servidor2 no se aprovechaban al máximo y se presentaba saturación de cpu y memoria en el servidor1 dada la configuración "Activo & Pasivo" que se tenía en la arquitectura inicial.

Las figuras 6.6, 6.7, 6.8 y 6.9, son muestras del consumo de los recursos de memoria y cpu en cada uno de los servidores una vez implementada la segunda fase del proyecto. Es importante mencionar que en el área de operación el umbral establecido de consumo de cpu y memoria en un servidor no debe ser mayor al 75% y dicho umbral no se rebasó en los servidores adquiridos para la capa front-end y back-end aun con el crecimiento de los servicios de valor agregado que se dio a corto (3 meses), mediano (6 meses) y largo plazo (1 año).

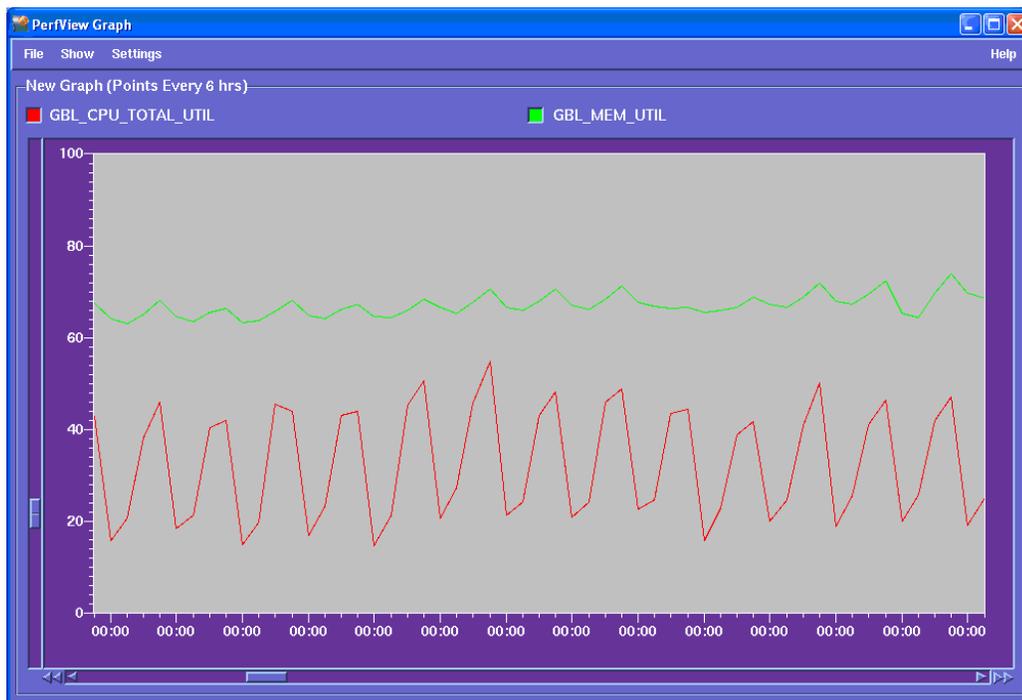


Figura 6.6 Consumo de cpu y memoria en el servidor webserver1. Arquitectura Final.

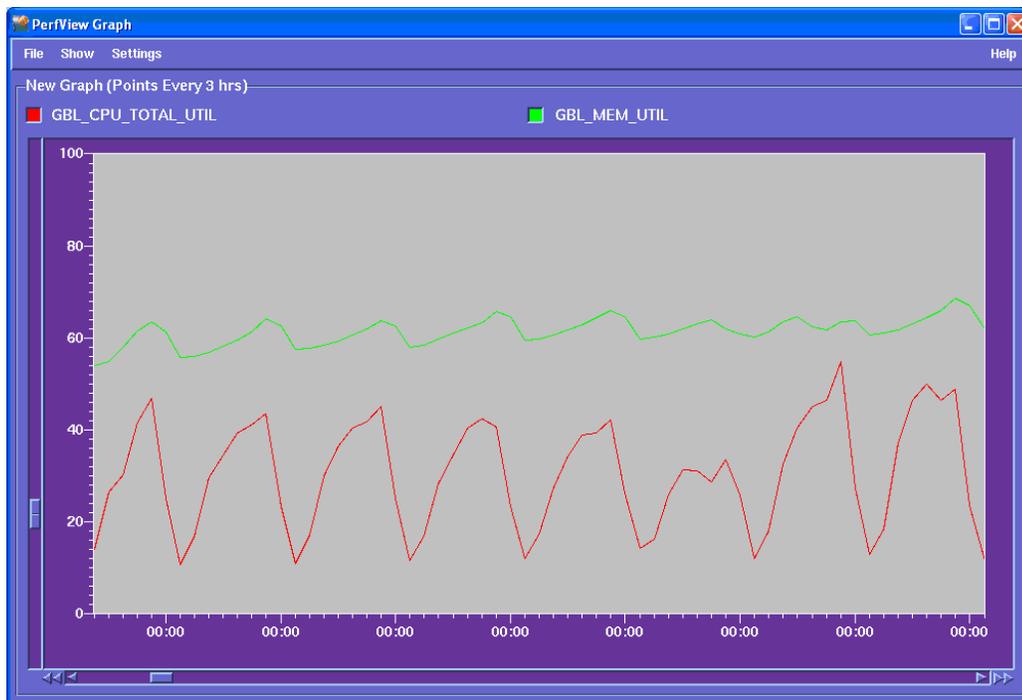


Figura 6.7 Consumo de cpu y memoria en el servidor webserver2. Arquitectura Final.



Figura 6.8 Consumo de cpu y memoria en el servidor backserver 1. Arquitectura Final.



Figura 6.9 Consumo de cpu y memoria en el servidor backserver 2. Arquitectura Final.

### 6.1.3.2 Resultados de la implementación del monitoreo centralizado.

Con la solución de monitoreo centralizado se ha logrado visualizar en tiempo real cualquier falla presentada en el *sistema* SVA y dar una atención inmediata al problema.

A continuación se muestra un ejemplo de la visualización de alarmas a través de la solución de monitoreo HP OVO:

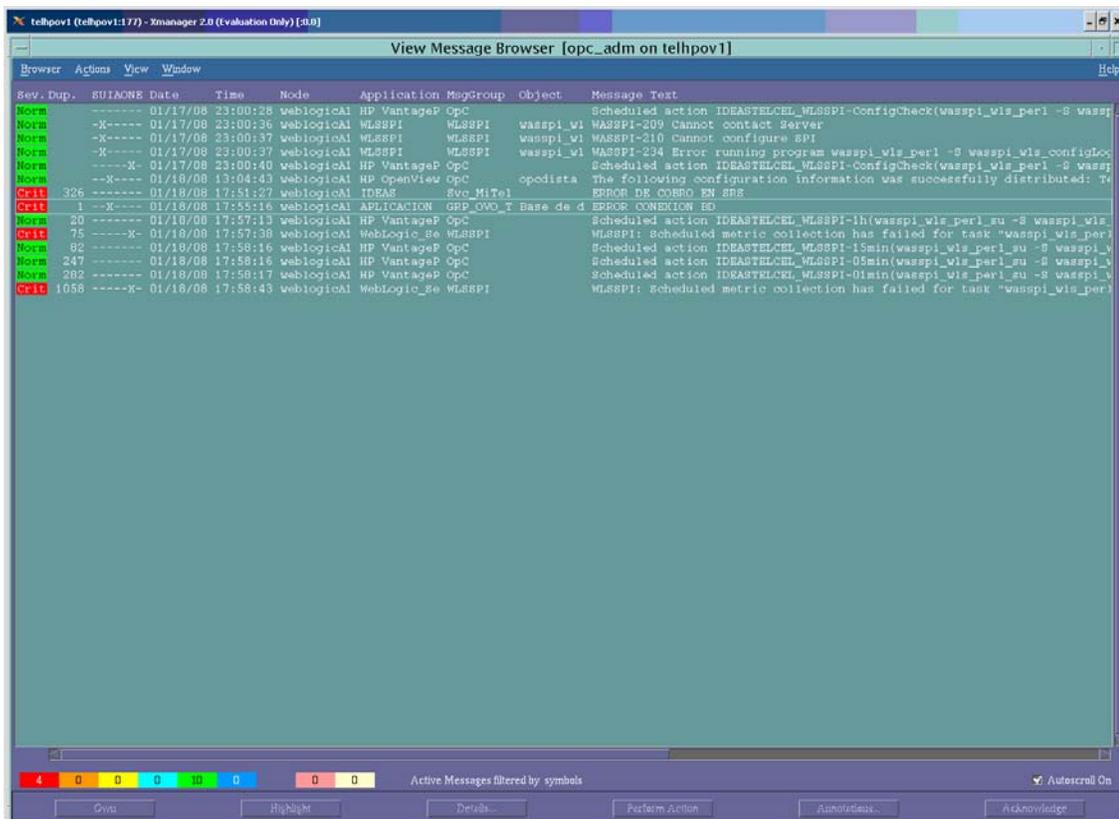


Figura 6.10 Visualización de alarmas a través del sistema de monitoreo HP OVO.

### **6.1.3.3 Resultados de la implementación del sistema de respaldos automático.**

Con la integración del *sistema SVA* a un sistema de respaldos confiable y automático, se logró eliminar el “down time” de 6 horas utilizado para el respaldo de la información. Actualmente, el respaldo se realiza de manera automática y cuenta con mecanismos de respaldo que no requieren dar de baja los aplicativos.

### **6.1.3.4 Fácil de escalar y transparente para el usuario final.**

Con el nuevo diseño del *sistema SVA* es posible escalar sus componentes de manera transparente para el usuario final. Por ejemplo, se ha dado crecimiento a la capa de “front-end” sin la necesidad de afectar la disponibilidad del sistema; esto es, antes de poner en productivo un nuevo nodo, se configura y prueba en un ambiente fuera de producción. Una vez concluido el proceso de implementación, el nuevo nodo se interconecta al sistema, se habilita en el dominio de WebLogic Sever y se procede con la apertura del tráfico en los balanceadores de carga.

Lo mismo pasa en la capa “back-end”, se ha dado la necesidad de instalar y poner en producción nuevos servicios y aplicaciones sin que esto impacte en la disponibilidad del *sistema SVA* siendo transparente para el usuario final.

Ahora bien, los mantenimientos físicos del *sistema SVA* se realizan actualmente haciendo uso de la alta disponibilidad implementada, es decir, si se tiene programado un mantenimiento físico en algún servidor de “back-end” se pasa todo el control de las bases de datos y procesos aplicativos al otro nodo que conforma el cluster para liberar así el que se requiere intervenir.

Lo mismo ocurre para los servidores que conforman la capa “front-end”, pues el servidor al que se desea dar mantenimiento se deshabilita en el dominio de Weblogic Server y se procede con el cierre del tráfico para este equipo desde los balanceadores de carga. Adicionalmente, esto ocurre sin la necesidad de afectar el servicio porque este tipo de trabajos se realizan en horarios de mantenimiento de 01:00 a 06:00 AM y el “switch” o “failover” de los servicios no dura más de 10 minutos.

Por último, las fallas ocurridas en algún proceso o equipo interno ya no impactan de manera directa en la disponibilidad total del *sistema SVA*, dado a que se cuenta con redundancia a nivel físico y lógico del sistema. Esto es, si alguna instancia de “BEA Weblogic Server (WLS)” falla en algún servidor de la capa “front-end”, o el servidor mismo, lo que se procede a ejecutar es desactivar la entrada de tráfico para el nodo en particular que tiene la falla aplicativa o física desde los balanceadores de carga mientras se resuelve el problema en el nodo afectado. Si la falla se presenta en el los servidores del “back-end” se procede a levantar los procesos o bases de datos afectadas en el otro nodo activo.

De esta forma, actualmente se ha logrado escalar y dar mantenimiento preventivo y/o correctivo al *sistema SVA* de manera transparente para el usuario final.

## 6.2 Resultados obtenidos con la optimización de procesos operativos.

Aunado a la mejora de monitoreo en tiempo real, con la mejora en los procesos operativos para la prestación de servicios de valor agregado que se describió en el capítulo 5 se logró una atención a las fallas en los sistemas SVA de 7X24 los 365 días del año con tiempos de respuesta efectivos:

Nivel de Severidad	Criterio	Descripción	Tiempo de atención	Tiempo de recuperación temporal	Tiempo de recuperación final
<b>Crítico</b>	Existe afectación masiva del servicio. Pérdida de funciones prioritarias.	No cursa tráfico. Problemas de aprovisionamiento. Problemas de facturación.	15 min.	1 a 2 horas máximo.	3 días
<b>Mayor</b>	Existe afectación potencial del servicio. Pérdida parcial de funciones	Pérdida de redundancia. Se presenta congestión Disturbios mayores SW & HW. Falla en la obtención de estadísticas.	15 min.	2 a 4 horas.	4 días
<b>Menor</b>	Condición que no afecta tráfico o funcionalidad importante del sistema.	Casos aislados de usuarios debido a fallas poco frecuentes.	15 min.	1 día.	1 semana.

Tabla 6.1 Niveles de de severidad y tiempos de respuesta.

Adicionalmente se obtuvieron los siguientes beneficios:

- ◆ Mejor planeación y programación en la ejecución de solicitudes de alta, baja o modificación de configuración en los sistemas de SVA en un tiempo no mayor a 48hrs.
- ◆ Seguimiento en tiempo a los programas de mantenimiento preventivos en los sistemas de SVA: limpieza física, revisión de desempeño de las aplicaciones, actualización de software, etc.
- ◆ Análisis de causa raíz a los posibles problemas de operación presentados en los sistemas de SVA.
- ◆ Análisis de desempeño y disponibilidad de los sistemas de SVA para garantizar la disponibilidad y capacidad de las mismas.
- ◆ Derivado de las mejoras al proceso de diseño, implementación y lanzamiento de un nuevo sistema de SVA, actualmente se cuenta con una estandarización de los requerimientos operativos para la puesta en producción de nuevos sistemas de SVA.
- ◆ Mejor distribución del trabajo y con ello un mejor análisis, diseño, implementación y lanzamiento de un nuevo servicio.

Lo anterior, contribuyó de manera directa al logro de nuestro segundo objetivo “**Optimizar los procesos operativos a aplicar en los sistemas de valor agregado**, dado al cambio y ofrecimiento constante de nuevos servicios y aplicaciones” y por supuesto contribuyó en la mejora de disponibilidad del *sistema* SVA y del resto de los sistemas de valor agregado.

### 6.3 Logros alcanzados en la disponibilidad del sistema SVA.

Con los resultados obtenidos al término de la segunda fase del presente proyecto y la optimización de los procesos de operación, la disponibilidad del *sistema de SVA* se incrementó considerablemente proporcionando una mejor administración en la operación de la misma y por ende una mejor calidad de servicio para el usuario final.

A continuación se presenta una muestra de disponibilidad semanal obtenida durante 1 año para el sistema objeto del presente trabajo:

Semana	Hrs	Día	Indisponibilidad					Semana	Disponibilidad
			HW	SW	1/2 Tx	Fza	Planeado		
1	24	7	0	0	0	0	0	02 al 08 ene	100.000%
2	24	7	0	0	0	0	0	09 al 15 ene	100.000%
3	24	7	0	0	0	0	0	16 al 22 ene	100.000%
4	24	7	0	0	0	0	0	23 al 29 ene	100.000%
5	24	7	0	0	0	0	0	30 al 05 feb.	100.000%
6	24	7	0	0	0	0	0	06 al 12 feb.	100.000%
7	24	7	0	0.5	0	0	0	13 al 19 feb.	99.702%
8	24	7	0	0.9	0	0	0	20 al 26 feb.	99.464%
9	24	7	0.7	0	0	0	0	27 al 05 mar	99.583%
10	24	7	0	0	0	0	0	06 al 12 mar	100.000%
11	24	7	0	0	0	0	0	13 al 19 mar	100.000%
12	24	7	0	0	0	0	0	20 al 26 mar	100.000%
13	24	7	0	0.4	0	0	0	27 al 2 abr.	99.762%
14	24	7	0	0	0	0	0	3 al 9 abr.	100.000%
15	24	7	0	0	0	0	0	10 al 16 abr.	100.000%
16	24	7	0	0	0	0	0	17 al 23 abr.	100.000%
17	24	7	0	0	0	0	0	23 al 30 abr.	100.000%
18	24	7	0	0	0	0	0	1 al 7 mayo	100.000%
19	24	7	0	0	0	0	0	8 al 14 mayo	100.000%
20	24	7	0	0	0	0	0	15 al 21 mayo	100.000%

21	24	7	0	0	0	0	0	22 al 28 mayo	100.000%
22	24	7	0	0	0	0	0	29 mayo al 4 jun.	100.000%
23	24	7	0	0	0	0	0	4 jun. al 11 jun.	100.000%
24	24	7	0	0	0	0	0	12 jun. al 18 jun.	100.000%
25	24	7	0	0	0	0	0	19 jun. al 25 jun.	100.000%
26	24	7	0	0	0	0	0	26 jun. al 2 jul.	100.000%
27	24	7	0	0	0	0	0	3 al 9 jul.	100.000%
28	24	7	0	0	0	0	0	10 al 16 jul.	100.000%
29	24	7	0	0	0	0	0	17 al 23 jul.	100.000%
30	24	7	0	0	0	0	0	24 al 30 jul.	100.000%
31	24	7	0	0	0	0	0	31 jul. al 6 ago.	100.000%
32	24	7	0	0	0	0	0	7 al 13 ago.	100.000%
33	24	7	0	0	0	0	0	14 al 20 ago.	100.000%
34	24	7	0	0	0	0	0	21 al 27 ago.	100.000%
35	24	7	0	0	0	0	0	28 ago. al 3 sept.	100.000%
36	24	7	0	0	0	0	0	4 al 10 sept.	100.000%
37	24	7	0	0	0	0	0	11 al 17 sept.	100.000%
38	24	7	0	0	0	0	0	18 al 24 sept.	100.000%
39	24	7	0	0	0	0	0	25 sept. al 1 oct.	100.000%
40	24	7	0	0	0	0	0	2 al 8 oct.	100.000%
41	24	7	0	0	0	0	0	9 al 15 oct.	100.000%
42	24	7	0	0	0	0	0	16 al 22 oct.	100.000%
43	24	7	0	0	0	0	0	23 al 29 oct.	100.000%
44	24	7	0	0	0	0	0	30 oct. al 5 nov.	100.000%
45	24	7	0	0	0	0	0	6 al 12 nov.	100.000%
46	24	7	0	0.3	0	0	0	13 al 19 nov.	99.821%
47	24	7	0	0	0	0	0	20 al 26 nov.	100.000%
48	24	7	0	0	0	0	0	27 nov. al 3 dic.	100.000%
49	24	7	0	0	0	0	0	4 al 10 dic.	100.000%
50	24	7	0	0	0	0	0	11 al 17 dic.	100.000%
51	24	7	0	0	0	0	0	18 al 24 dic.	100.000%
52	24	7	0	0	0	0	0	25 al 31 dic.	100.000%
<b>Disponibilidad anual</b>									<b>99.968%</b>

Tabla 6.2 Muestra anual de la disponibilidad alcanzada en el sistema SVA.

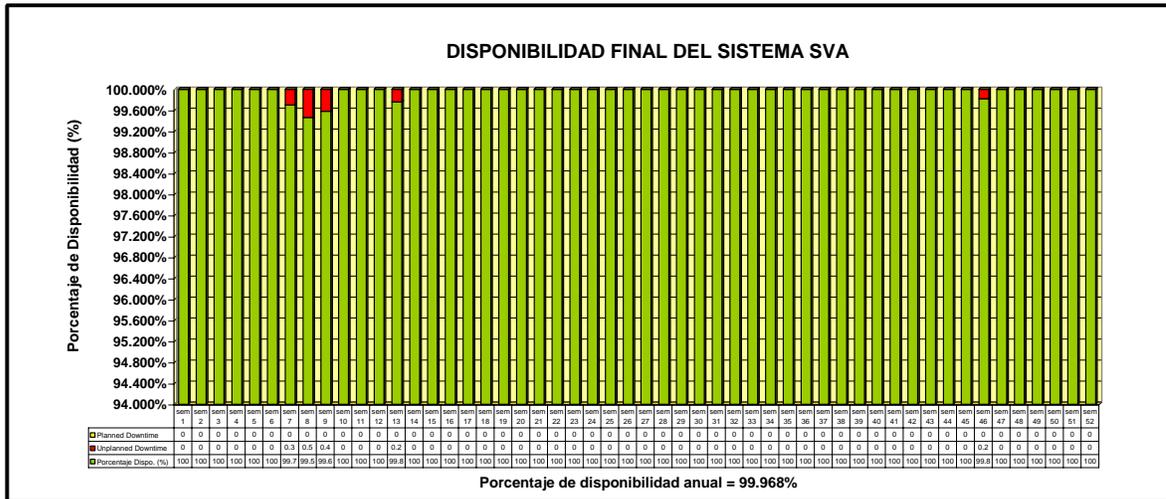


Figura 6.11 Estadística de disponibilidad anual del sistema SVA.

Como se puede observar en la estadística de disponibilidad anterior, el primer objetivo de este proyecto: “Contar con una solución a nivel hardware/software que permita **alcanzar y mantener la disponibilidad semanal del sistema SVA al 99.9%**”, se cumplió para 47 de las 52 semanas que conforman el año. La disponibilidad promedio anual que se alcanzó para el sistema SVA fue del **99.968%**.

Las fallas que impactaron en la disponibilidad del sistema SVA correspondieron principalmente a fallas en nuevos desarrollos aplicativos; sin embargo, se aplicó un mejor control sobre las mismas gracias a la optimización de los procesos operativos y al proceso de mejora continua que se aplica en la empresa.

Por último, con el proyecto de actualización tecnológica, se aseguró la operación del sistema SVA a corto, mediano y largo plazo. Esto es, se han alojado sobre el mismo sistema, nuevos servicios sin la necesidad de invertir grandes cantidades de dinero, tiempo y esfuerzo, ofreciendo a los clientes servicios de vanguardia y calidad.

## Conclusiones

---

Actualmente con la globalización de los mercados, uno de los factores trascendentales para el éxito de una empresa, es la calidad de sus productos ó servicios. Los consumidores de hoy, están exigiendo altos grados de calidad, a precios justos, con rapidez en la entrega y sobre todo con un buen nivel de servicio.

Para una empresa de comunicaciones móviles, es muy importante estar siempre actualizada tecnológicamente y asegurar que sus procesos estén dirigidos a mejorar el nivel de satisfacción del cliente como resultado de sus actividades diarias.

Durante el desarrollo y puesta en producción de este proyecto, logré identificar las áreas de oportunidad que la empresa tenía en el área operativa para los sistemas de servicios de valor agregado. Lo anterior, me llevo a proponer y poner en práctica la actualización tecnológica descrita en el capítulo 4, y las mejoras al proceso de operación que se describieron en el capítulo 5; ambas acciones aportaron grandes beneficios a la empresa, entre los que destacan:

- ◆ Aumento en la disponibilidad del *sistema SVA* del 93 al 99.7% anual.
- ◆ Atención a las fallas en los sistemas SVA de 7X24 los 365 días del año con tiempos de respuesta efectivos.
- ◆ Estandarización de los requerimientos operativos para la puesta en producción de nuevos sistemas de SVA.
- ◆ Mejor distribución del trabajo y con ello un mejor análisis, diseño, implementación y lanzamiento de un nuevo servicio.

Es relevante mencionar, que estos logros se deben en gran parte, al sistema de gestión de calidad que se aplica en la empresa a través de la norma ISO 9001:2000. Esta norma tiene como objetivo, orientar a una organización hacia la satisfacción del cliente utilizando los siguientes principios básicos: enfoque al cliente, liderazgo, involucramiento del personal, enfoque basado en procesos, enfoque de sistemas para la gestión, mejora continua, enfoque basado en hechos para la toma de decisión y, relaciones mutuamente beneficiosas con los proveedores. De este modo, es posible mejorar un producto o servicio de forma constante.

Por lo anterior, puedo concluir, que es y será importante en mi área de trabajo seguir aplicando los principios del sistema de gestión de calidad y buscar siempre la eficiencia de los procesos de operación, con el fin de contribuir en el logro de los objetivos de la empresa y con esto ofrecer un buen servicio al cliente que cumpla y exceda sus expectativas.

## **Glosario**

---

**API. Application Programming Interface** - Interfaz de Programación de Aplicaciones, es el conjunto de funciones y procedimientos (o métodos si se refiere a programación orientada a objetos) que ofrece cierta biblioteca para ser utilizado por otro software como una capa de abstracción. Una API representa un interfaz de comunicación entre componentes software.

**Appletalk** es un conjunto de protocolos desarrollados por Apple Inc. para la conexión de redes. Fue incluido en un Macintosh en 1984 y actualmente está en desuso en los Macintosh en favor de las redes TCP/IP.

**ARP. Address Resolution Protocol.** Protocolo de resolución de direcciones. Es un protocolo de nivel de red responsable de encontrar la dirección hardware (Ethernet MAC) que corresponde a una determinada dirección IP. Para ello se envía un paquete (ARP request) a la dirección de multidifusión de la red (broadcast (MAC = ff ff ff ff ff ff)) que contiene la dirección IP por la que se pregunta, y se espera a que esa máquina (u otra) responda (ARP reply) con la dirección Ethernet que le corresponde. Cada máquina mantiene una caché con las direcciones traducidas para reducir el retardo y la carga. ARP permite a la dirección de Internet ser independiente de la dirección Ethernet, pero esto solo funciona si todas las máquinas lo soportan.

**ASN.1.** Abstract Syntax Notation One (notación sintáctica abstracta 1) es una norma para representar datos independientemente de la máquina que se esté usando y sus formas de representación internas. Es un protocolo de nivel de presentación en el modelo OSI. El protocolo SNMP usa el ASN.1 para representar sus objetos gestionables.

**ATA** Advanced Technology Attachment, controla los dispositivos de almacenamiento masivo de datos, como los discos duros.

**ATM.** El Modo de Transferencia Asíncrona o Asynchronous Transfer Mode es una tecnología de telecomunicación desarrollada para hacer frente a la gran demanda de capacidad de transmisión para servicios y aplicaciones. Con esta tecnología, a fin de aprovechar al máximo la capacidad de los sistemas de transmisión, sean estos de cable o radioeléctricos, la información no es transmitida y conmutada a través de canales asignados en permanencia, sino en forma de cortos paquetes (celdas ATM) de longitud constante y que pueden ser enrutadas individualmente mediante el uso de los denominados *canales virtuales* y *trayectos virtuales*.

**Bluetooth** es el nombre común de la especificación industrial IEEE 802.15.1, que define un estándar global de comunicación inalámbrica que posibilita la transmisión de voz y datos entre diferentes dispositivos mediante un enlace por radiofrecuencia segura, globalmente y sin licencia de corto rango

**Cable coaxial.** El cable coaxial (habitualmente denominado “coax”), está formado por dos conductores que comparten un eje común, de aquí su nombre (“coaxis”). Generalmente, el centro del cable es un hilo de cobre relativamente rígido o un hilo acordonado envuelto a un recubrimiento plástico aislante. El recubrimiento está rodeado por un segundo conductor, un tubo de malla de hilo (algunos incluyen un envoltorio metálico conductor), que sirve como blindaje frente a EMI. Un tubo de plástico duro aislante forma la cubierta del cable.

**Cable de fibra óptica.** El cable de fibra óptica es un medio de transmisión que cada vez se está utilizando más en redes locales que exigen velocidades de transmisiones muy altas y fiables. Las señales luminosas se transmiten a través de un cable (guía de ondas) compuesto por un grupo de cristales o de fibras plásticas. Cada filamento tiene un núcleo central de plástico o cristal con un alto índice de refracción, rodeado de una capa de material similar con un índice de refracción ligeramente menor. El revestimiento aísla las fibras y evita que se produzcan interferencias entre filamentos adyacentes, y al mismo tiempo proporciona protección al núcleo. Todo el conjunto está normalmente protegido por otras capas que no tienen ninguna otra función que la de proteger los filamentos.

**Cable par trenzado.** El cable par trenzado, es el medio de transmisión más antiguo y ampliamente utilizado. Este consiste en dos alambres de cobre aislados, en general de 1 mm de espesor. La forma trenzada del cable se utiliza para reducir la interferencia eléctrica con respecto a los pares cercanos que se encuentran a su alrededor. La aplicación más común del cable par trenzado es en las redes telefónicas. Este tipo de redes propiciaron precisamente el ingreso del par trenzado a los mercados de las redes de computadoras. Actualmente el par trenzado, tiene una amplia difusión no solamente en telefonía, sino también dentro de las redes LAN de computadoras. Existen dos tipos de cable par trenzado: No blindado: UTP (Unshielded Twisted-Pair) y Blindado: STP (Shielded Twisted-Pair)

**Canal de comunicación.** Es el medio de transmisión por el que viajan las señales portadoras de la información que pretenden intercambiar emisor y receptor. Es frecuente referenciarlo también como canal de datos.

**Computación distribuida**, es un nuevo modelo para resolver problemas de computación masiva utilizando un gran número de computadoras organizadas en racimos incrustados en una infraestructura de telecomunicaciones distribuida.

**DAS. Direct Attached Storage**, es el método tradicional de almacenamiento y el más sencillo. Consiste en conectar el dispositivo de almacenamiento directamente al servidor o estación de trabajo, es decir, físicamente conectado al dispositivo que hace uso de él.

**Datagrama** es un fragmento de paquete que es enviado con la suficiente información como para que la red pueda simplemente encaminar el fragmento hacia el equipo terminal de datos receptor, de manera independiente a los fragmentos restantes.

**EJB's.** Los Enterprise JavaBeans son una de las API que forman parte del estándar de construcción de aplicaciones empresariales J2EE de Sun Microsystems (ahora JEE 5.0). Los EJBs proporcionan un modelo de componentes distribuido estándar del lado del servidor. El objetivo de los EJBs es dotar al programador de un modelo que le permita abstraerse de los problemas generales de una aplicación empresarial (conurrencia, transacciones, persistencia, seguridad, etc.) para centrarse en el desarrollo de la lógica de negocio en sí. El hecho de estar basado en componentes permite que éstos sean flexibles y sobre todo reutilizables. No hay que confundir los Enterprise JavaBeans con los JavaBeans. Los JavaBeans también son un modelo de componentes creado por Sun Microsystems para la construcción de aplicaciones, pero no pueden utilizarse en entornos de objetos distribuidos al no soportar nativamente la invocación remota (RMI)

**Estándar.** Un estándar es una especificación que regula la realización de ciertos procesos o la fabricación de componentes para garantizar la interoperabilidad.

**ETD. Equipo Terminal de Datos** o DTE, *Data Terminal Equipment* es aquel componente del circuito de datos que hace de fuente o destino de la información. Puedes ser un terminal, una impresora o también un potente ordenador. La característica definitoria de un ETD no es la eficiencia ni la potencia de cálculo, sino la función que realiza: ser origen o destino en una comunicación. Un ETD fuente por lo general contiene la información almacenada en un dispositivo de memoria principal permanente (que se modifica sin un flujo electrónico continuo), el ETD destino es aquel que recibe una información o datos de manera directa o indirecta, sin alterar el contenido de la información durante el total del proceso. En resumen, se considera ETD a cualquier equipo informático, sea receptor o emisor final de datos.

**Ethernet.** Es el nombre de una tecnología de redes de computadoras de área local (LANs) basada en tramas de datos. El nombre viene del concepto físico de *ether*. Ethernet define las características de cableado y señalización de nivel físico y los formatos de trama del nivel de enlace de datos del modelo OSI. **Ethernet se refiere a las redes de área local y dispositivos bajo el estándar IEEE 802.3** que define el protocolo CSMA/CD.

**Fast Ethernet o Ethernet de alta velocidad.** Es el nombre de una serie de estándares de IEEE de redes Ethernet de 100 Mbps. En su momento el prefijo *fast* se le agregó para diferenciarla de la Ethernet regular de 10 Mbps

**FDDI** Fiber Distributed Data Interface, es un conjunto de estándares ISO y ANSI para la transmisión de datos en redes de computadoras de área extendida o local (LAN) mediante cable de fibra óptica. Se basa en la arquitectura token ring y permite una comunicación tipo Full Duplex. Dado que puede abastecer a miles de usuarios, una LAN FDDI suele ser empleada como backbone para una red de área amplia (WAN).

**FTP.** File Transfer Protocol. Es un protocolo de transferencia de archivos entre sistemas conectados a una red TCP basado en la arquitectura cliente-servidor, de manera que desde un equipo cliente nos podemos conectar a un servidor para descargar archivos desde él o para enviarle nuestros propios archivos independientemente del sistema operativo utilizado en cada equipo.

**Gigabit Ethernet**, también conocida como **GigE**, es una ampliación del estándar Ethernet (concretamente la versión 802.3ab y 802.3z del IEEE) que consigue una capacidad de transmisión de 1 gigabit por segundo, correspondientes a unos 1000 megabits por segundo de rendimiento contra unos 100 de Fast Ethernet.

**HBA** (Host Bus Adapter) o adaptador de bus de host, es la tarjeta de interfaz que conecta un host a una SAN (Storage Area Network).

**HDLC.** High-Level Data Link Control es un protocolo de comunicaciones de datos punto a punto entre dos elementos basado en el ISO 3309. Proporciona recuperación de errores en caso de pérdida de paquetes de datos, fallos de secuencia y otros. Mediante una red de conmutadores de paquetes conectados con líneas punto a punto entre ellos y con los usuarios se constituye la base de las redes de comunicaciones X25. Este es un protocolo de propósito general, que opera a nivel de enlace de datos. Este ofrece una comunicación confiable entre el transmisor y el receptor.

**Hojas de estilo en cascada** (Cascading Style Sheets, CSS), son un lenguaje formal usado para definir la presentación de un documento estructurado escrito en HTML o XML (y por extensión en XHTML). El W3C (World Wide Web Consortium) es el encargado de formular la especificación de las hojas de estilo que servirá de estándar para los agentes de usuario o navegadores.

**HTML** HyperText Markup Language o Lenguaje de Marcas Hipertextuales. Es un lenguaje de marcación diseñado para estructurar textos y presentarlos en forma de hipertexto, que es el formato estándar de las páginas web. Gracias a Internet y a los navegadores como Internet Explorer, Opera, Firefox, Netscape o Safari, el HTML se ha convertido en uno de los formatos más populares y fáciles de aprender que existen para la elaboración de documentos para web.

**HTTP**. HyperText Transfer Protocol. El protocolo de transferencia de hipertexto es el protocolo usado en cada transacción de la Web (WWW). HTTP define la sintaxis y la semántica que utilizan los elementos software de la arquitectura web (clientes, servidores, proxies) para comunicarse. Es un protocolo orientado a transacciones y sigue el esquema petición-respuesta entre un cliente y un servidor.

**ICMP**. Internet Control Message Protocol. El **Protocolo de Mensajes de Control de Internet** o es el subprotocolo de control y notificación de errores del Protocolo de Internet (IP). Como tal, se usa para enviar mensajes de error, indicando por ejemplo, que un servicio determinado no está disponible o que un router o host no puede ser localizado. ICMP difiere del propósito de TCP y UDP porque generalmente no se utiliza directamente por las aplicaciones de usuario en la red. La única excepción es la herramienta ping y traceroute, que envían mensajes de petición Echo ICMP (y recibe mensajes de respuesta Echo) para determinar si un host está disponible, el tiempo que le toma a los paquetes en ir y regresar a ese host y cantidad de hosts por los que pasa.

**ICQ**. *"IGUAL COMO QUIEN "* o "te busco", es un servicio de mensajería instantánea y el primero de su tipo en ser ampliamente utilizado en Internet, mediante el cual es posible chatear y enviar mensajes instantáneos a otros usuarios conectados a la red de ICQ. También permite el envío de archivos, videoconferencias y charlas de voz.

**IGMP**. Internet Group Management Protocol. Es un protocolo de red que se utiliza para intercambiar información acerca del estado de pertenencia entre enrutadores IP que admiten la multidifusión y miembros de grupos de multidifusión. Los hosts miembros individuales informan acerca de la pertenencia de hosts al grupo de multidifusión y los enrutadores de multidifusión sondan periódicamente el estado de la pertenencia.

**Infrared Data Association (IrDA)** define un estándar físico en la forma de transmisión y recepción de datos por rayos infrarrojo. IrDA se crea en 1993 entre HP, IBM, Sharp y otros. Esta tecnología, basada en rayos luminosos que se mueven en el espectro infrarrojo. Los estándares IrDA soportan una amplia gama de dispositivos eléctricos, informáticos y de comunicaciones, permite la comunicación bidireccional entre dos extremos a velocidades que oscilan entre los 9.600 bps y los 4 Mbps. Esta tecnología se encuentra en muchos ordenadores portátiles, y en un creciente número de teléfonos celulares, sobre todo en los de fabricantes líderes como Nokia y Ericsson.

**IEEE** corresponde a las siglas de *The Institute of Electrical and Electronics Engineers*, el **Instituto de Ingenieros Eléctricos y Electrónicos**, una asociación técnico-profesional mundial dedicada a la estandarización, entre otras cosas. Es la mayor asociación internacional sin fines de lucro formada por profesionales de las nuevas tecnologías, como ingenieros eléctricos, ingenieros en electrónica, científicos de la computación, ingenieros en informática e ingenieros en telecomunicación. Su creación se remonta al año 1884, contando entre sus fundadores a personalidades de la talla de Thomas Alva Edison, Alexander Graham Bell y Franklin Leonard Pope. En 1963 adoptó el nombre de IEEE al fusionarse asociaciones como el AIEE (*American Institute of Electrical Engineers*) y el IRE (*Institute of Radio Engineers*).

**IMAP. Internet Message Access Protocol**, es un protocolo de red de acceso a mensajes electrónicos almacenados en un servidor. Mediante IMAP se puede tener acceso al correo electrónico desde cualquier equipo que tenga una conexión a Internet. IMAP tiene varias ventajas sobre POP, que es el otro protocolo empleado para obtener correo desde un servidor. Por ejemplo, es posible especificar en IMAP carpetas del lado servidor. Por otro lado, es más complejo que POP porque permite visualizar los mensajes de manera remota y no descargando los mensajes como lo hace POP.

**IP. Internet Protocol.** El **Protocolo de Internet** es un protocolo no orientado a conexión usado tanto por el origen como por el destino para la **comunicación de datos a través de una red de paquetes conmutados**. Los datos en una red basada en IP son enviados en bloques conocidos como paquetes o datagramas (en el protocolo IP estos términos se suelen usar indistintamente). En particular, en IP no se necesita ninguna configuración antes de que un equipo intente enviar paquetes a otro con el que no se había comunicado antes. El Protocolo de Internet provee un servicio de datagramas no fiable (también llamado del *mejor esfuerzo* (*best effort*), lo hará lo mejor posible pero garantizando poco). IP no provee ningún mecanismo para determinar si un paquete alcanza o no su destino y únicamente proporciona seguridad (mediante *checksums* o sumas de comprobación) de sus cabeceras y no de los datos transmitidos. Por ejemplo, al no garantizar nada sobre la recepción del paquete, éste podría llegar dañado, en otro orden con respecto a otros paquetes, duplicado o simplemente no llegar. Si se necesita fiabilidad, ésta es proporcionada por los protocolos de la capa de transporte, como TCP.

**IPv4** es la versión 4 del Protocolo IP (Internet Protocol). Esta fue la primera versión del protocolo que se implementó extensamente, y forma la base de Internet. IPv4 usa direcciones de 32 bits, limitándola a  $2^{32} = 4.294.967.296$  direcciones únicas, muchas de las cuales están dedicadas a redes locales (LANs).

**IPv6** es la versión 6 del Protocolo de Internet (Internet Protocol), un estándar en desarrollo del nivel de red encargado de dirigir y encaminar los paquetes a través de una red. Diseñado por Steve Deering de Xerox PARC y Craig Mudge, IPv6 está destinado a sustituir al estándar IPv4, cuyo límite en el número de direcciones de red admisibles está empezando a restringir el crecimiento de Internet y su uso, especialmente en China, India, y otros países asiáticos densamente poblados. Pero el nuevo estándar mejorará el servicio globalmente; por ejemplo, proporcionando a futuras celdas telefónicas y dispositivos móviles con sus direcciones propias y permanentes. Al día de hoy se calcula que las dos terceras partes de las direcciones que ofrece IPv4 ya están asignadas. IPv4 soporta 4.294.967.296 ( $2^{32}$ ) direcciones de red diferentes, un número inadecuado para dar una dirección a cada persona del planeta, y mucho menos para cada coche, teléfono, PDA, etcétera; mientras que IPv6 soporta 340.282.366.920.938.463.463.374.607.431.768.211.456 ( $2^{128}$  ó 340 sextillones).

**IPX.** Internetwork Packet Exchange. Intercambio de paquetes interredes. Protocolo de comunicaciones NetWare que se utiliza para encaminar mensajes de un nodo a otro. Los paquetes IPX incluyen direcciones de redes y pueden enviarse de una red a otra. Ocasionalmente, un paquete IPX puede perderse cuando cruza redes, de esta manera el IPX no garantiza la entrega de un mensaje completo. La aplicación tiene que proveer ese control o debe utilizarse el protocolo SPX de NetWare. IPX provee servicios en estratos 3 y 4 del modelo OSI (capas de red y transporte). Actualmente este protocolo está en desuso y solo se utiliza para juegos en red antiguos.

**IRC.** Internet Relay Chat es un protocolo de comunicación en tiempo real basado en texto, que permite debates en grupo o entre dos personas y que está clasificado dentro de la Mensajería instantánea.

**Internet SCSI (iSCSI)** es un protocolo que utiliza TCP/IP para sus transferencias de datos. Al contrario que otros protocolos de red diseñados para almacenamiento, como por ejemplo el canal de fibra (que es la base de la mayor parte de las redes de áreas de almacenamiento), solamente requiere un simple y una sencilla interfaz Ethernet para funcionar. Esto permite una solución de almacenamiento centralizada de bajo coste sin la necesidad de realizar inversiones costosas ni sufrir las habituales incompatibilidades asociadas a las soluciones canal de fibra para redes de área de almacenamiento. Los críticos de iSCSI argumentan que este protocolo tiene un peor rendimiento que el canal de fibra porque se ve afectado por la sobrecarga que generan las transmisiones TCP/IP (cabeceras de paquetes, por ejemplo). Sin embargo las pruebas que se han realizado muestran un excelente rendimiento de las soluciones iSCSI SANs, cuando se utilizan enlaces Gigabit Ethernet.

**ISO.** Organización Internacional para la Estandarización o *International Organization for Standardization*, es una organización internacional no gubernamental, compuesta por representantes de los organismos de normalización nacionales, que produce normas internacionales industriales y comerciales. Dichas normas se conocen como normas ISO y su finalidad es la coordinación de las normas nacionales, en consonancia con el Acta Final de la Organización Mundial del Comercio, con el propósito de facilitar el comercio, facilitar el intercambio de información y contribuir con unos estándares comunes para el desarrollo y transferencia de tecnologías.

**Java** es un lenguaje de programación orientado a objetos desarrollado por Sun Microsystems a principios de los años 1990

**JavaScript** es un lenguaje interpretado, es decir, que no requiere compilación, utilizado principalmente en páginas web, con una sintaxis semejante a la del lenguaje Java y el lenguaje C. Al contrario que Java, JavaScript no es un lenguaje orientado a objetos propiamente dicho, no dispone de Herencia, es más bien un lenguaje basado en prototipos, porque las nuevas clases se generan clonando las clases base (prototipos) y extendiendo su funcionalidad.

**Java Servlet.** La palabra *servlet* deriva de otra anterior, *applet*, que se refería a pequeños programas escritos en Java que se ejecutan en el contexto de un navegador web. Por contraposición, un *servlet* es un programa que se ejecuta en un servidor. Los **servlets** son objetos que corren dentro del contexto de un contenedor de servlets (ej.: Tomcat) y extienden su funcionalidad. También podrían correr dentro de un servidor de aplicaciones (ej.: BEA Weblogic) que además de contenedor para servlet tendrá contenedor para objetos más avanzados como son los EJB (Tomcat sólo es un contenedor de servlets). El uso más común de los *servlets* es generar páginas web de forma dinámica a partir de los parámetros de la petición que envíe el navegador web.

**JBoss.** Es un servidor de aplicaciones J2EE de código abierto implementado en Java puro. Al estar basado en Java, JBoss puede ser utilizado en cualquier sistema operativo que lo soporte.

**JDBC** es el acrónimo de Java Database Connectivity, un API que permite la ejecución de operaciones sobre bases de datos desde el lenguaje de programación Java independientemente del sistema operativo donde se ejecute o de la base de datos a la cual se accede utilizando el dialecto SQL del modelo de base de datos que se utilice.

**JOnAS.** Es un servidor de aplicaciones **J2EE** de código abierto implementado en Java. Forma parte de la iniciativa de código abierto de ObjectWeb, la cual fue lanzada en colaboración con varios socios, entre los que se encuentran Bull, France Telecom e INRIA.

**JSP.** JavaServer Pages es una tecnología Java que permite generar contenido dinámico para web, en forma de documentos HTML, XML o de otro tipo. Esta tecnología es un desarrollo de la compañía Sun Microsystems.

**LAPB. Link Access Procedure, Balanced** es un protocolo de nivel de enlace de datos dentro del conjunto de protocolos de la norma X.25. LAPB está orientado al bit y deriva de HDLC. Modifica a este último para que cualquiera de los dos nodos pueda iniciar la transmisión, por esto se ha denominado **balanceado**.

**MathML** o Mathematical Markup Language, es un lenguaje de marcado basado en XML, cuyo objetivo es expresar notación matemática de forma que distintas máquinas puedan entenderla, para su uso en combinación con XHTML en páginas web, y para intercambio de información entre programas de tipo matemático en general. El soporte de MathML es grande en programas matemáticos (como Maple o Mathcad), pero aún escaso en navegadores (sólo vía plugins en Internet Explorer, parcial en Mozilla).

**MOSIX** es un administrador para clusters Linux y multi-clusters.

**Middleware.** Es un software de conectividad que ofrece un conjunto de servicios que hacen posible el funcionamiento de aplicaciones distribuidas sobre plataformas heterogéneas. Funciona como una capa de abstracción de software distribuida, que se sitúa entre las capas de aplicaciones y las capas inferiores (sistema operativo y red). El Middleware nos abstrae de la complejidad y heterogeneidad de las redes de comunicaciones subyacentes, así como de los sistemas operativos y lenguajes de programación, proporcionando una API para la fácil programación y manejo de aplicaciones distribuidas. Dependiendo del problema a resolver y de las funciones necesarias, serán útiles diferentes tipo de servicios de middleware.

**NetBEUI.** NetBIOS Extended User Interface, en español Interfaz extendida de usuario de NetBIOS, es un protocolo de nivel de red sin encaminamiento y bastante sencillo utilizado como una de las capas en las primeras redes de Microsoft. NetBIOS sobre NetBEUI es utilizado por muchos sistemas operativos desarrollados en los 1990, como LAN Manager, LAN Server, Windows 3.x, Windows 95 y Windows NT.

**NetBIOS,** "Network Basic Input/Output System", es, en sentido estricto una especificación de interfaz para acceso a servicios de red, es decir, una capa de software desarrollado para enlazar un sistema operativo de red con hardware específico. NetBIOS fue originalmente desarrollado por IBM y Sytek como un API/APIS para el software cliente de recursos de una Red de área local (LAN).

**NFS.** Network File System (*Sistema de archivos de red*), es un protocolo de nivel de aplicación, según el Modelo OSI. Es utilizado para sistemas de archivos distribuido en un entorno de red de computadoras de área local. Posibilita que distintos sistemas conectados a una misma red accedan a ficheros remotos como si se tratara de locales.

**NNTP.** Network News Transport Protocol, es una aplicación de Internet que consiste en un protocolo usado para la lectura y publicación de artículos de noticias en Usenet. Su traducción literal al español es "protocolo para la transferencia de noticias en red".

**OSI.** Open System Interconnection. Es un modelo de referencia de Interconexión de Sistemas Abiertos creado por ISO y lanzado en 1984.

**PDA.** Personal Digital Assistant o Ayudante Personal Digital, es un computador de mano originalmente diseñado como agenda electrónica (calendario, lista de contactos, bloc de notas y recordatorios) con un sistema de reconocimiento de escritura. Actualmente se puede usar como una computadora doméstica (ver películas, crear documentos, juegos, correo electrónico, navegar por Internet, escuchar música, etc.).

**Plataforma.** Se refiere al sistema operativo o a sistemas complejos que a su vez sirven para crear programas, como las **plataformas de desarrollo**. Una **plataforma de desarrollo** es el entorno común en el cual se desenvuelve la programación de un grupo definido de aplicaciones.

**POP3.** En informática se utiliza el **Post Office Protocol** en clientes locales de correo para obtener los mensajes de correo electrónico almacenados en un servidor remoto. La mayoría de los suscriptores de los proveedores de internet acceden a sus correos a través de POP3.

**PPP Point-to-point Protocol**, es un protocolo asociado a la pila TCP/IP de uso en Internet. El protocolo PPP permite establecer una comunicación a nivel de enlace entre dos computadoras. Generalmente, se utiliza para establecer la conexión a Internet de un particular con su proveedor de acceso a través de un modem telefónico. PPP facilita dos funciones importantes, la *autenticación*, generalmente mediante una clave de acceso, y la asignación dinámica de IP.

**RARP.** Reverse Address Resolution Protocol. Protocolo de resolución de direcciones inverso. Es un protocolo utilizado para resolver la dirección IP de una dirección hardware dada (como una dirección Ethernet).

**Redes inalámbricas** (en inglés *wireless*), son aquellas que se comunican por un medio de transmisión no guiado (sin cables) mediante ondas electromagnéticas. La transmisión y la recepción se realizan a través de antenas.

**Red por microondas** es un tipo de red inalámbrica que utiliza microondas como medio de transmisión. El protocolo más frecuente es el IEEE 802.11b y transmite a 2.4 GHz, alcanzando velocidades de 11 Mbps (Megabits por segundo).

**Red por radio.** Es aquella que emplea la radiofrecuencia como medio de unión de las diversas estaciones de la red.

**RPC.** Remote Procedure Call, Llamada a Procedimiento Remoto. Es un protocolo que permite a un programa de ordenador ejecutar código en otra máquina remota sin tener que preocuparse por las comunicaciones entre ambos. El protocolo es un gran avance sobre los sockets usados hasta el momento. De esta manera el programador no tenía que estar pendiente de las comunicaciones, estando éstas encapsuladas dentro de las RPC.

**RS-232** (también conocido como Electronic Industries Alliance RS-232C) es una interfaz que designa una norma para el intercambio serie de datos binarios entre un DTE (Equipo terminal de datos) y un DCE (*Data Communication Equipment*, Equipo de Comunicación de datos),

**SCSI**, acrónimo inglés **S**mall **C**omputer **S**ystem **I**nterface, es un interfaz estándar para la transferencia de datos entre distintos dispositivos del bus de la computadora.

**Servlets** son objetos que corren dentro del contexto de un contenedor de servlets (ej.: Tomcat) y extienden su funcionalidad. También podrían correr dentro de un servidor de aplicaciones que además de contenedor para servlet tendrá contenedor para objetos más avanzados como son los EJB (Tomcat sólo es un contenedor de servlets).

**Sistema operativo.** Un **sistema operativo** es un programa o conjunto de programas destinado a permitir la comunicación del usuario con un computador y gestionar sus recursos de una forma eficaz. Comienza a trabajar cuando se enciende el computador, y gestiona el hardware de la máquina desde los niveles más básicos.

**SGML.** Standard Generalized Markup Language o Lenguaje de Marcación Generalizado. Consiste en un sistema para la organización y etiquetado de documentos. La Organización Internacional de Estándares (ISO) ha normalizado este lenguaje en 1986. El lenguaje SGML sirve para especificar las reglas de etiquetado de documentos y no impone en sí ningún conjunto de etiquetas en especial. El lenguaje HTML esta definido en términos del SGML.

**Software no libre.** El software no libre (también llamado software propietario, software privativo, software privado, software con propietario o software de propiedad) se refiere a cualquier programa informático en el que los usuarios tienen limitadas las posibilidades de usarlo, modificarlo o redistribuirlo (con o sin modificaciones), o cuyo código fuente no está disponible o el acceso a éste se encuentra restringido

**SMIL.** Synchronized Multimedia Integration Language o Lenguaje de Integración Multimedia Sincronizada. Es un estándar del World Wide Web Consortium (W3C) para presentaciones multimedia. El lenguaje SMIL permite integrar audio, video, imágenes, texto o cualquier otro contenido multimedia.

**SMTP** Simple Mail Transfer Protocol o protocolo simple de transferencia de correo. Protocolo de red basado en texto utilizado para el intercambio de mensajes de correo electrónico entre computadoras o distintos dispositivos (PDA's, teléfonos móviles, etc.).

**SNMP.** Simple Network Management Protocol. El Protocolo Simple de Administración de Red, es un protocolo de la capa de aplicación que facilita el intercambio de información de administración entre dispositivos de red. Es parte de la familia de protocolos TCP/IP. SNMP permite a los administradores supervisar el desempeño de la red, buscar y resolver sus problemas, y planear su crecimiento.

**SPX. Sequenced Packet Exchange.** Es un antiguo protocolo de red de Novell usado para manejar Novell Internetwork Packet Exchange.

**SSH** Secure SHell, es el nombre de un protocolo y del programa que lo implementa, y sirve para acceder a máquinas remotas a través de una red. SSH trabaja de forma similar a como se hace con telnet. La diferencia principal es que SSH usa técnicas de cifrado que hacen que la información que viaja por el medio de comunicación vaya de manera no legible y ninguna tercera persona pueda descubrir el usuario y contraseña de la conexión ni lo que se escribe durante toda la sesión; aunque es posible atacar este tipo de sistemas por medio de ataques de REPLAY y manipular así la información entre destinos.

**SSL.** Secure Sockets Layer proporciona autenticación y privacidad de la información entre extremos sobre Internet mediante el uso de criptografía. Habitualmente, sólo el servidor es autenticado (es decir, se garantiza su identidad) mientras que el cliente se mantiene sin autenticar; la autenticación mutua requiere un despliegue de infraestructura de claves públicas (o PKI) para los clientes. Los protocolos permiten a las aplicaciones cliente-servidor comunicarse de una forma diseñada para prevenir escuchas (eavesdropping), la falsificación de la identidad del remitente (phishing) y mantener la integridad del mensaje.

**SVG** Scalable Vector Graphics, es un lenguaje para describir gráficos vectoriales bidimensionales, tanto estáticos como animados (estos últimos con ayuda de SMIL), en XML.

**Switch**, en castellano "conmutador", es un dispositivo electrónico de interconexión de redes de ordenadores que opera en la capa 2 del modelo OSI. Un conmutador interconecta dos o más segmentos de red, pasando datos de un segmento a otro, de acuerdo con la dirección MAC de destino de los datagramas en la red.

**Tarjeta SIM.** *Subscriber Identity Module* o Módulo de Identificación del Suscriptor, es una tarjeta inteligente desmontable usada en teléfonos móviles que almacena de forma segura la clave de servicio del suscriptor usada para identificarse ante la red, de forma que sea posible cambiar la línea de un terminal a otro simplemente cambiando la tarjeta.

**Tasa de transferencia de datos.** Es la velocidad con que los datos son transferidos a través de una red.

**TCP** Transmission Control Protocol, en español **Protocolo de Control de Transmisión**. Es uno de los protocolos fundamentales en Internet. Fue creado entre los años 1973 - 1974 por Vint Cerf y Robert Kahn). Muchos programas dentro de una red de datos compuesta por ordenadores pueden usar TCP para crear *conexiones* entre ellos a través de las cuales enviarse un flujo de datos. El protocolo garantiza que los datos serán entregados en su destino sin errores y en el mismo orden en que se transmitieron. También proporciona un mecanismo para distinguir distintas aplicaciones dentro de una misma máquina, a través del concepto de puerto. TCP da soporte a muchas de las aplicaciones más populares de Internet, incluidas HTTP, SMTP y SSH.

**Telnet.** TELEcommunication NETwork, es el nombre de un protocolo de red (y del programa informático que implementa el cliente), que sirve para acceder mediante una red a otra máquina, para manejarla como si estuviéramos sentados delante de ella. Para que la conexión funcione, como en todos los servicios de Internet, la máquina a la que se acceda debe tener un programa especial que reciba y gestione las conexiones. El puerto que se utiliza generalmente es el 23.

**Token Ring.** Arquitectura de red desarrollada por IBM en los años 70's con topología lógica en anillo y técnica de acceso de paso de testigo. Token Ring se recoge en el estándar IEEE 802.5. En desuso por la popularización de Ethernet.

**Tomcat.** También llamado **Jakarta Tomcat** o **Apache Tomcat** funciona como un contenedor de servlets desarrollado bajo el proyecto Jakarta en la Apache Software Foundation. Tomcat implementa las especificaciones de los servlets y de JavaServer Pages (JSP) de Sun Microsystems.

**Trama de red.** Es una unidad de envío de datos. Viene a ser sinónimo de paquete de datos o paquete de red, aunque se aplica principalmente en los niveles OSI más bajos, especialmente en el nivel de enlace de datos.

Normalmente una trama consta de cabecera, datos y cola. En la cola suele estar algún chequeo de errores. En la cabecera van los campos de control de protocolo. La parte de datos es la que quiera transmitir el nivel de comunicación superior, típicamente el nivel de red.

**UDP** User Datagram Protocol, es un protocolo sin conexión que, como TCP, funciona en redes IP. UDP/IP proporciona muy pocos servicios de recuperación de errores, ofreciendo en su lugar una manera directa de enviar y recibir datagramas a través una red IP. Se utiliza sobre todo cuando la velocidad es un factor importante en la transmisión de la información, por ejemplo, RealAudio utiliza el UDP.

**UIT-T.** El **Sector de Normalización de las Telecomunicaciones de la UIT**, es el órgano permanente de la Unión Internacional de Telecomunicaciones (UIT) que estudia los aspectos técnicos, de explotación y tarifarios y publica normativa sobre los mismos, con vista a la normalización de las telecomunicaciones a nivel mundial. Con sede en Ginebra (Suiza) fue conocido hasta 1992 como **Comité Consultivo Telefónico y Telegráfico (CCITT)**.

**Unix** (registrado oficialmente como UNIX®) es un sistema operativo portable, multitarea y multiusuario; desarrollado, en principio, en 1969 por un grupo de empleados de los laboratorios Bell de AT&T, entre los que figuran Ken Thompson, Dennis Ritchie y Douglas Mcllroy.

**Wap Push.** Tecnología que permite acceder mediante uno o varios SMS a la descarga de contenido alojado en un servidor WAP: aplicaciones JAVA, imágenes, melodías polifónicas, videos, etc. El funcionamiento se basa en una estructura XML que contiene la dirección o URL del contenido, la cual es compilada y enviada como un SMS binario. Este mensaje es interpretado como una página WAP que es abierta por el Navegador del teléfono móvil, permitiendo descargar el contenido enlazado.

**W3C.** World Wide Web Consortium, abreviado es un consorcio internacional que produce estándares para la World Wide Web. Está dirigida por Tim Berners-Lee, el creador original de URL (Uniform Resource Locator, Localizador Uniforme de Recursos), HTTP (HyperText Transfer Protocol, Protocolo de Transferencia de HiperTexto) y HTML (Lenguaje de Marcado de HiperTexto) que son las principales tecnologías sobre las que se basa la Web.

**X.25** es un estándar UIT-T para redes de área amplia de conmutación de paquetes. Su protocolo de enlace, LAPB, está basado en el protocolo HDLC proveniente de IBM. Establece mecanismos de direccionamiento entre usuarios, negociación de características de comunicación, técnicas de recuperación de errores. Los servicios públicos de conmutación de paquetes admiten numerosos tipos de estaciones de distintos fabricantes. Por lo tanto, es de la mayor importancia definir la interfaz entre el equipo del usuario final y la red.

**XHTML** Hypertext Markup Language o Lenguaje Extensible de Marcado de Hipertexto. Es el lenguaje de marcado pensado para sustituir a HTML como estándar para las páginas web. XHTML es la versión XML de HTML, por lo que tiene, básicamente, las mismas funcionalidades, pero cumple las especificaciones, más estrictas, de XML. Su objetivo es avanzar en el proyecto del World Wide Web Consortium de lograr una web semántica, donde la información, y la forma de presentarla estén claramente separadas. En este sentido, XHTML serviría únicamente para transmitir la información que contiene un documento, dejando para hojas de estilo (como las hojas de estilo en cascada) y JavaScript su aspecto y diseño en distintos medios (ordenadores, PDAs, teléfonos móviles, impresoras...).

**XML**, sigla en inglés de eXtensible Markup Language («lenguaje de marcas extensible»), es un metalenguaje extensible de etiquetas desarrollado por el World Wide Web Consortium (W3C). XML es un nuevo estándar con una funcionalidad similar a la del SGML aunque más sencillo, y de creación posterior. Por lo tanto XML no es realmente un lenguaje en particular, sino una manera de definir lenguajes para diferentes necesidades. Algunos de estos lenguajes que usan XML para su definición son XHTML, SVG, MathML.

## **Bibliografía**

---

- ◆ **José Manuel Huidobro Moya.** *Comunicaciones Móviles.* Thomson Paraninfo, 2002.
  
- ◆ **Alberto Sendín Escalona.** *Fundamentos de los sistemas de comunicaciones móviles,* McGraw-Hill, 2004
  
- ◆ INEGI. <http://cuentame.inegi.gob.mx/>
  
- ◆ CDMA Development Group. <http://www.cdg.org/>
  
- ◆ 3GPP, 3er. Generation Partnership Project. [www.3gpp.org](http://www.3gpp.org)
  
- ◆ 3GPP2, 3er. Generation Partnership Project 2. [www.3gpp2.org](http://www.3gpp2.org)
  
- ◆ 3G Americas. <http://www.3gamericas.com/>
  
- ◆ Foro UMTS. [www.umts-forum.org](http://www.umts-forum.org)
  
- ◆ BEA Systems. <http://www.bea.com/>
  
- ◆ The Apache Software Foundation. <http://tomcat.apache.org/>
  
- ◆ Evans Data Corporation. <http://www.evansdata.com/>
  
- ◆ Symantec. <http://www.symantec.com/>
  
- ◆ Wikipedia. <http://es.wikipedia.org/>
  
- ◆ Internacional Organization for Estandardization. <http://www.iso.org/>