Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas

HERRAMIENTAS DE TI PARA LA CONTINUIDAD DEL NEGOCIO

Carlos Roberto Trujillo Ramírez
Asesorado por el Ing. Andrés González

Guatemala, marzo de 2012

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

FACULTAD DE INGENIERÍA

HERRAMIENTAS DE TI PARA LA CONTINUIDAD DEL NEGOCIO

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

CARLOS ROBERTO TRUJILLO RAMÍREZ
ASESORADO POR EL ING. ANDRÉS ANTONIO GONZÁLEZ LEÓN

AL CONFERÍRSELE EL TÍTULO DE

INGENIERO EN CIENCIAS Y SISTEMAS

GUATEMALA, MARZO DE 2012

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

FACULTAD DE INGENIERÍA

NÓMINA DE JUNTA DIRECTIVA

DECANO	Ing. Murphy Olympo Paiz Recinos
VOCAL I	Ing. Alfredo Enrique Beber Aceituno
VOCAL II	Ing. Pedro Antonio Aguilar Polanco
VOCAL III	Ing. Miguel Ángel Dávila Calderón
VOCAL IV	Br. Juan Carlos Molina Jiménez
VOCAL V	Br. Mario Maldonado Muralles
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

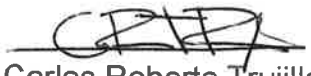
DECANO	Ing. Murphy Olympo Paiz Recinos
EXAMINADOR	Ing. César Augusto Fernández Cáceres
EXAMINADORA	Inga. Virginia Victoria Tala Ayerdi
EXAMINADOR	Ing. Edgar Estuardo Santos Sutuj
SECRETARIA	Inga. Marcia Ivónne Véliz Vargas

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

HERRAMIENTAS DE TI PARA LA CONTINUIDAD DEL NEGOCIO

Tema que me fuera asignado por la Dirección de la Escuela de Ingeniería en Ciencias y Sistemas, con fecha diciembre de 2009.



Carlos Roberto Trujillo Ramirez

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERIA
ESCUELA DE CIENCIAS Y SISTEMAS

Ref: ASESOR 02-02

Guatemala 10 de septiembre de 2010

Señores
Comisión de Revisión de Tesis
Carrera de Ciencias y Sistemas
Facultad de Ingeniería
Universidad de San Carlos de Guatemala
Guatemala, Ciudad

Respetables Señores:

El motivo de la presente es informarles que como asesor del estudiante Carlos Roberto Trujillo Ramírez he procedido a revisar el trabajo de tesis titulado Herramientas de TI para la continuidad del negocio y que de acuerdo a mi criterio el mismo se encuentra concluido y cumple con los objetivos definidos al inicio.

He tenido reuniones periódicas con el estudiante y luego de haber revisado cuidadosamente el trabajo, considero que cumple con los requisitos de calidad y profesionalismo que deben caracterizar a un futuro profesional de la Informática.

Aprovecho para informarle que he leído detenidamente el documento Ref: ASESOR 01-02 y aplicando las recomendaciones que se dan en el mismo procedo a firmar de revisado el trabajo de tesis.

Sin otro particular me suscribo de ustedes,

Atentamente,

Ing. Andres Antonio Gonzalez Leon

Universidad San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Ingeniería en Ciencias y Sistemas

Guatemala, 29 de Septiembre de 2010

Ingeniero
Marlon Antonio Pérez Turk
Director de la Escuela de Ingeniería
En Ciencias y Sistemas

Respetable Ingeniero Pérez:

Por este medio hago de su conocimiento que he revisado el trabajo de graduación del estudiante **CARLOS ROBERTO TRUJILLO RAMIREZ** camé **2004-13252,** titulado: **"HERRAMIENTAS DE TI PARA LA CONTINUIDAD DEL NEGOCIO",** y a mi criterio el mismo cumple con los objetivos propuestos para su desarrollo, según el protocolo.

Al agradecer su atención a la presente, aprovecho la oportunidad para suscribirme,

Atentamente,

Ing. Carlos Alfredo Azurdia
Coordinador de Privados
y Revisión de Trabajos de Graduación

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA - INGENIERIA EN CIENCIAS Y SISTEMAS - COORDINACION TESIS

ESCUELA DE CIENCIAS Y SISTEMAS

UNIVERSIDAD DE SAN CARLOS
DE GUATEMALA

FACULTAD DE INGENIERÌA
ESCUELA DE CIENCIAS Y SISTEMAS
TEL: 24767644

El Director de la Escuela de Ingeniería en Ciencias y Sistemas de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor con el visto bueno del revisor y del Licenciado en Letras, de trabajo de graduación titulado ***"HERRAMIENTAS DE TI PARA LA CONTINUIDAD DEL NEGOCIO"*** *presentado por el estudiante CARLOS ROBERTO TRUJILLO RAMÍREZ, aprueba el presente trabajo y solicita la autorización del mismo.*

"ID Y ENSEÑAD A TODOS"

Ing. Marlon Antonio Pérez Turk
Director, Escuela de Ingeniería Ciencias y Sistemas

Guatemala, 06 de marzo 2012

Universidad de San Carlos
de Guatemala

Facultad de Ingeniería
Decanato

Ref. DTG.124-2012

El Decano de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Ingeniería en Ciencias y Sistemas, al trabajo de graduación titulado: **HERRAMIENTAS DE TI PARA LA CONTINUIDAD DEL NEGOCIO,** presentado por la estudiante universitaria, **Carlos Roberto Trujillo Ramírez,** autoriza la impresión del mismo.

IMPRÍMASE.

Ing. Murphy Olympo Paiz Recinos
DECANO

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
DECANO
FACULTAD DE INGENIERIA

Guatemala, marzo de 2012

/cc
c.c. archivo.

ACTO QUE DEDICO A:

Dios	Por permitirme finalizar una meta más en mi vida.
Mis padres	Mario y Sonia, por todas sus enseñanzas, consejos y amor.
Mis hermanos	Jorge, Joaquín, Patricia, por su apoyo incondicional, en todo momento.
Flor de María	Por todo su cariño, comprensión, y apoyo.

AGRADECIMIENTOS A:

Andrés González	Por el apoyo brindado en la realización de este trabajo.
Empresa ICG	Por permitir la realización de pruebas requeridas para completar el caso de estudio.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES

FIGURAS

TABLAS

GLOSARIO

Alta disponibilidad — Concepto que denota la característica de un sistema de permanecer disponible la mayor parte del tiempo posible.

Ambiente distribuido — Conjunto de sistemas que se complementan entre sí, y que pueden compartir o no sus datos parcial o totalmente.

Blade — Servidor para centros de datos, específicamente diseñado para aprovechar el espacio, reducir el consumo y simplificar su explotación.

Continuidad del negocio — Propiedad de las empresas que asegura que sus operaciones son realizadas sin interrupciones, o con la menor cantidad posible de estas.

Copia instantánea — Copia tomada de un sistema, archivo o conjunto de datos que representa el estado de este recurso en un momento determinado del tiempo.

Dirección IP — Identificador de una interfaz de un dispositivo dentro de una red.

FDDI — Conjunto de estándares ISO y ANSI para la transmisión de datos en redes LAN o WAN.

Heartbeat	Herramienta que provee infraestructura de clúster para distintos servicios.
Nada compartido	Principio de los ambientes distribuidos que dice que ningún dato será compartido por los sitios de replicación en ambientes distribuidos.
Período de convergencia	Tiempo que tardan los sistemas distribuidos en replicar la totalidad de sus datos.
Redundancia	Duplicidad de cualquier elemento de un sistema, los casos pueden ser de *software* o *hardware* e incluye porciones de sistemas, o la totalidad de estos.
Replicación	Procedimiento de copia y sincronización de datos a cualquier nivel (puede ser LUN, discos duros, bases de datos, archivos, etc.) su objetivo es sincronizar una copia remota de los datos en un sitio primario.
Sitio alterno	Sitio que actúa como respaldo del sitio primario, al ocurrir un incidente, este tomará el rol de sitio primario dentro del sistema.

Sitio primario	En un ambiente de replicación, es el sitio que, en el caso ideal, realiza todas las transacciones del sistema, y almacena todos los datos, normalmente, esto cambia al ocurrir un incidente, pasando el sitio secundario a tomar el rol de sitio primario.
STP	Protocolo de red de nivel de capa 2, Su función es la de gestionar bucles en topologías de red.
Todo compartido	Principio de los sistemas distribuidos que dice que todos los datos son compartidos e idénticos para todos los sitios de replicación.

RESUMEN

Actualmente, la tecnología juega un papel importante en las operaciones de las empresas, es común encontrar que los procesos operativos sean irrealizables sin el apoyo de sistemas informáticos. Los planes de continuidad del negocio, deben entonces incluir y hacer énfasis en la continuidad de los servicios de tecnología.

En el presente documento, se estudia la continuidad del negocio desde el punto de vista de tecnología, es decir, se analiza la continuidad de los servicios de tecnología, como pilar básico de la continuidad de las operaciones del negocio. Los temas a tratar abarcan desde la definición de los procedimientos que aseguren la continuidad de los servicios de tecnología, hasta aquellas implementaciones tecnológicas que lleven a elevar el factor de disponibilidad de tales servicios.

Durante el proceso de implementación de soluciones para la continuidad del negocio, el personal de tecnología se verá envuelto en una serie de incógnitas respecto al alcance de cada tipo de implementación, así como su viabilidad y beneficios. Las más importantes de estas incógnitas son estudiadas a lo largo de este documento, con la finalidad de facilitar la toma de decisiones en la implementación.

Se presenta también las posibles opciones de combinación de diversas tecnologías para aumentar el factor de disponibilidad de los servicios de tecnología, así como las implicaciones financieras y de rendimiento sobre los procesos involucrados. Todos los estudios incluidos, fueron realizados pensando en medianas empresas del medio guatemalteco, pero al ser implementaciones bastante genéricas y escalables, pueden ser fácilmente transportadas a medios mucho mayores o menores.

OBJETIVOS

General

Realizar un análisis de las diferentes tecnologías existentes para apoyar a la continuidad del negocio desde el punto de vista de tecnología.

Específicos

1. Analizar herramientas de apoyo para la continuidad de los servicios de tecnología.

2. Establecer cuáles son las mejores opciones para cubrir cada uno de los aspectos de continuidad de los servicios de tecnología.

3. Realizar una comparación entre las diferentes opciones tecnológicas para continuidad del negocio.

INTRODUCCIÓN

El presente trabajo se ha dividido para tratar de manera individual cada uno de los tópicos más importantes a abordar al implementar tecnologías para la continuidad del negocio, a fin de mantener cierto grado de interrelación para mostrar los beneficios de implementaciones complementarias.

En el primer capítulo, encontrará los estándares, normas y marcos de trabajo existentes, tales como COBIT, ITIL, BS25999 etc. Los cuales brindan guías y mejores prácticas para definir el diseño del proceso de continuidad de los servicios de tecnología como la documentación respectiva. En el segundo capítulo, se habla de la redundancia de sistemas de *hardware* y *software*, las implicaciones de redundar en sistemas desde el punto de vista del rendimiento, como de los costos de implementación y mantenimiento de estos sistemas.

El tercer capítulo hace énfasis en la necesidad de implementar tecnologías de alta disponibilidad en la parte de comunicaciones, las ventajas y necesidades de redundar en enlaces, así como la importancia de implementar protocolos de red que soporten redundancia. Se analiza también los dispositivos de red, y las opciones disponibles para crear redes altamente disponibles.

El cuarto capítulo aborda el tema de replicación, separándolo en replicación síncrona y asíncrona, muestra las ventajas y desventajas de cada tipo, así como también, muestra las posibles vías de implementación para estas soluciones. Se analiza la necesidad de adjuntar soluciones como copias instantáneas a los sistemas replicados para asegurar la integridad de los datos.

En el quinto capítulo, se analiza la alta disponibilidad de las bases de datos, las distintas implementaciones que pueden ayudar a asegurar la integridad de los datos almacenados en una base de datos, por ejemplo los sistemas de espejos, bases de datos distribuidas en arquitecturas todo compartido, la recuperación de bitácoras en sitios alternos, etc.

Por último, el sexto capítulo analiza un caso de implementación real centrándose en el procedimiento de promoción del sitio alterno durante una rutina de pruebas realizada por la empresa estudiada. De aquí puede tomarse una idea del tiempo utilizado para restaurar las operaciones del negocio luego de una caída del sitio primario, a la vez se muestran las ventajas de la implementación realizada sobre ambientes totalmente virtualizados.

1. HERRAMIENTAS DE TI PARA LA CONTINUIDAD DEL NEGOCIO

Actualmente, la infraestructura tecnológica soporta gran cantidad de operaciones en los negocios, al punto de que para muchos de ellos resulta imposible continuar operando una vez que los sistemas informáticos sufren una condición que les lleve a interrumpirse. Un breve análisis sobre los costos que implica una caída del sistema para una empresa que base sus operaciones en su infraestructura tecnológica, demuestra la alta importancia que tiene la continuidad de los sistemas informáticos.

Desde el punto de vista de tecnología, la continuidad del negocio se basa en la implementación de una serie de herramientas entre las que se incluye la redundancia de *hardware*, redundancia de software, replicación de sistemas, tecnologías de almacenaje, etc. La implementación de cada una de estas herramientas conllevará a un aumento en la disponibilidad promedio de los sistemas, puesto que aumentan considerablemente las acciones proactivas en detrimento de las acciones reactivas.

Las herramientas de TI para continuidad del negocio, deben ser implementadas pensando en que amenazas desean contrarrestarse, siendo las más comunes:

- Fallos de aplicaciones
- Fallos de sistema operativo

- Fallos de *hardware*
- Problemas de red
- Problemas eléctricos
- Desastres naturales
- Denegación de servicios

Siendo éstas las amenazas principales para los sistemas informáticos, resulta fácil formarse una idea respecto de la dificultad de administrar las herramientas de tecnología en entornos de empresas que trabajan 24 horas al día en un gran porcentaje de días por año.

Todo esto se complica más, si tomamos en cuenta que el 32% de las empresas tendrían serios problemas operativos, si los sistemas tecnológicos dejasen de funcionar durante 4 horas[1].

[1] BAJGORIC, Nijaz. Continuous Computing Technologies for Enhancing Business Continuity. p 12.

1.1. Estándares y normas

Existe una serie de estándares y normas, que pueden consultarse al implementar alta disponibilidad, se describen a continuación.

1.1.1. ISO 27000

ISO 27000 (conocido también como ISO 27002) es una norma que ofrece recomendaciones orientadas a la seguridad de la información. Cuenta con 10 dominios, teniendo uno especialmente dedicado a la administración de la continuidad del negocio. De estos 10 dominios se derivan 36 objetivos de control y 127 controles.

Entre las premisas básicas definidas por el ISO 27000 se tiene que:

- Proteger los procesos críticos para mantener las operaciones comerciales del negocio.

- Deben realizarse planes para la reanudación oportuna de las operaciones esenciales.

- Mitigar los riesgos que puedan afectar a las operaciones esenciales del negocio mediante la limitación de sus alcances.

ISO 27000 también incluye una serie de lineamientos para su implementación:

- Entender los riesgos en términos de probabilidad y tiempo.

- Identificar todos los activos involucrados en procesos críticos.

- Identificar el impacto que tendrían las interrupciones en las operaciones del negocio.

- Identificar y considerar la implementación de controles preventivos adicionales.

- Identificar los recursos disponibles para tratar los requerimientos de seguridad identificados.

- Garantizar la seguridad del personal y la protección de los medios de procesamiento de la información.

- Formular los planes de continuidad del negocio.

- Probar y actualizar constantemente estos planes.

- Asegurar que la gestión de la continuidad del negocio se incorpore a los procesos y estructura de la organización.

1.1.2. Objetivos de control para la información y tecnologías relacionadas

COBIT (siglas en inglés *Control Objectives for information and related technologies*) es un marco de trabajo para la gobernabilidad de TI cuyo fin principal es alinear los objetivos del negocio, con los objetivos de TI.

COBIT cuenta con cuatro dominios, los cuales son:

- Planificación y organización.
- Adquisición e implementación.
- Entrega y respaldo.
- Monitoreo.

Entre los procesos del dominio “entrega y respaldo", existe el proceso llamado “Garantizar la continuidad del servicio”.

La idea de COBIT en cuanto a continuidad del negocio, es establecer controles que aseguren el menor impacto en el negocio, en caso de una interrupción en los servicios de TI, mediante la revisión y prueba continúa de los planes de contingencia, así como el almacenaje de sistemas redundantes en locaciones alternas a la principal. COBIT mide la efectividad de la continuidad del negocio mediante dos indicadores principales:

- Número de horas perdidas por usuario por mes debidas a interrupciones no planeadas.
- Número de procesos críticos del negocio que dependen de TI que no están cubiertos por un plan de continuidad.

COBIT cubre también una serie de controles específicos para asegurar la continuidad de los servicios de TI:

- Crear un marco de trabajo de continuidad.
- Crear planes de continuidad de TI.
- Establecer los recursos críticos de TI, y priorizarlos en cuestión de tiempos máximos de restablecimiento en caso de fallas.
- Mantenimiento a planes de continuidad de TI, que implica la gestión del cambio sobre los planes de continuidad.
- Pruebas regulares sobre el plan de continuidad de TI.
- Entrenamiento de todas las partes involucradas en el plan de continuidad de TI.
- Distribución del plan de continuidad entre las partes involucradas.
- Planear las acciones a tomar durante el período en que los servicios de TI se encuentren recuperándose.
- Almacenaje de respaldos en locaciones alternas.
- Revisión post-reanudación.

COBIT mide el nivel de madurez del proceso en 5 niveles diferentes:

- Nivel 0: no existente
- Nivel 1: inicial
- Nivel 2: repetible pero intuitivo
- Nivel 3: proceso definido
- Nivel 4: administrado y medible
- Nivel 5: optimizado

Para el caso de la continuidad de las operaciones, estos niveles van desde la no existencia de ningún plan de contingencia, y el fallo total de las operaciones en caso de la interrupción de los procesos de TI hasta un entorno en donde los planes de continuidad de TI se encuentran alineados e integrados con los planes de continuidad del negocio, basándose en las mejores prácticas de la industria, y recibiendo un mantenimiento constante. Estos mantenimientos incluyen una gestión cíclica del plan de continuidad, retroalimentando con los resultados de las pruebas globales al personal que gestiona el plan de continuidad (normalmente el CEO) para su actualización. Se garantiza en este nivel que un incidente mayor no ocurrirá únicamente por la falla en un punto.

1.1.3. ITIL

ITIL (Sus siglas en inglés *Information Technology Infraestructure Library*) es un marco de trabajo que define las mejores prácticas para la gestión de los servicios de TI, cuenta con una sección específicamente dedicada a la continuidad del negocio, llamada ITSCM (Gestión de la continuidad de los servicios de TI).

Según ITIL, la planeación de la continuidad del negocio puede tener dos extremos:

- Extremadamente reactivo: donde no existe ningún plan hasta que un desastre ocurre.

- Extremadamente proactivo: donde existe un gasto innecesario de recursos en el planeamiento de la continuidad del negocio.

ITIL se enfoca en proveer continuidad del negocio mediante la introducción de medidas de reducción de riesgos y opciones de recuperación.

Los objetivos de ITSCM son[2]:

- Mantener una serie de planes de continuidad de servicios de TI y planes de recuperación que soporten todos los planes de continuidad del negocio.

- Completar análisis de impacto al negocio regulares, para asegurarse de que los planes de continuidad de los servicios de TI se mantengan alineados con los impactos y riesgos del negocio.

- Proveer información y soporte a todas las demás áreas en todas las cuestiones relacionadas a la continuidad de TI.

- Asegurar que la empresa tenga los mecanismos adecuados para obtener o superar sus objetivos de continuidad.

- Medir el impacto de todos los cambios en los planes de continuidad y recuperación de los servicios de TI.

[2] OGC. ITIL service design. p. 126

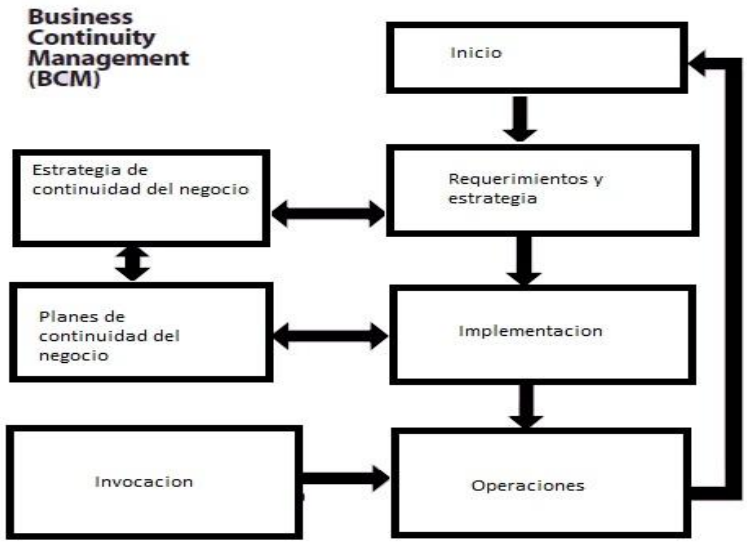
- Asegurarse de que todas las medidas proactivas sean implementadas, una vez que su beneficio costo-efectividad sea evidenciado.

- Negociar con proveedores que puedan prestar servicios necesarios para asegurar la continuidad de los servicios.

ITSCM se orienta a encontrar los procesos que sean críticos para el negocio, e identificar los fallos que puedan llevarle a un desastre real, por ejemplo, en este marco no se incluyen fallos menores como la falla de un disco duro en una computadora cliente o un servidor que no incluya servicios críticos, este tipo de fallos son cubiertos por otras partes de ITIL.

La figura 1 muestra el ciclo de vida de la administración de la continuidad del servicio:

Figura 1. **Ciclo de vida de la gestión de la continuidad del servicio**



Fuente: elaboración propia.

La figura 1 muestra la gestión de la continuidad del negocio según ITIL, se puede ver aquí como el ciclo va desde el inicio del proceso, asegurando durante todo el ciclo que la protección brindada por el plan está actualizada y refleja los cambios necesarios en cualquier momento.

Es importante remarcar que este ciclo incluye las etapas necesarias de un proceso BCM, este incluirá la definición de estrategias básicas para la continuidad del negocio, pero también deberá definirse el rol que juegan los servicios de TI dentro del plan, una vez definido esto, es donde realmente empieza el ciclo de ITSCM.

ITSCM contempla entonces dos fases

- La etapa de requerimientos y estrategia, que realiza análisis de impacto al negocio para entender requerimientos, y en base a esto, diseña una estrategia para reducir el riesgo sobre los procesos, así como los procedimientos de restauración para los servicios críticos. Este análisis de impacto, puede contemplar varios tipos de impactos negativos sobre el negocio, yendo desde impactos puramente financieros hasta efectos negativos sobre la imagen de la empresa, sus relaciones públicas, etc. Y define los impactos de cada problema que pueda surgir en nivel de diversos factores que podrían afectar el impacto que representa para el negocio, por ejemplo, tomando en cuenta la diferencia existente entre que un problema surja en un día de la semana u otro. Así también, se define acá las habilidades necesarias por parte del personal para restablecer los procesos críticos de la empresa, y los tiempos para realizar tales acciones.

- La etapa de iniciación contempla todas aquellas actividades de definición de responsabilidades, creación y comunicación de políticas, definición de alcances, Identificación de los recursos disponibles.

Esta relación es de gran ayuda, puesto que representará escenarios en los que se evalúa el impacto de una falla en cualquier sistema involucrando otras variables, por ejemplo, el tiempo, así, es posible que una empresa sobreviva y continúe operando durante un corto período tiempo si un sistema crítico falla, pero al término de este período, las operaciones de la empresa se verán afectadas seriamente, esto evidencia que las medidas de respuesta pueden tomarse durante ese período de tiempo sin incurrir en serios daños al negocio, lo cual indicará también, la cantidad de recursos que debería invertirse en la ejecución de este plan de restauración en específico, así como su priorización.

Una vez completado el análisis de impacto al negocio, es necesario realizar un análisis de riesgos, clasificándolos según alguna metodología específica para este fin, lo cual arrojará como resultado el listado de aquellas amenazas y el activo sobre el que estas amenazas pueden representar un riesgo. En conjunción el análisis de impacto y el análisis de riesgos permitirán crear una estrategia de continuidad de servicios de TI que corresponda puramente a las necesidades del negocio. Esta estrategia deberá ser un balance entre reducción de riesgos y estrategias para la restauración de los servicios, este balance estará realizado desde la perspectiva de costos, tiempos de respuesta requeridos, e impacto de cualquier falla sobre cualquier sistema.

La etapa de implementación es realizada una vez que los planes de continuidad del negocio se encuentran plenamente alineados con los planes de continuidad de servicios de TI, esto quiere decir que los planes de ITSCM están enfocados a asegurar la continuidad de los servicios de TI que resultan críticos para el negocio, con tiempos de recuperación que son aceptables para el negocio en caso de fallas.

A partir de esta etapa, las pruebas deben realizarse de manera constante, para asegurar la validez del plan de continuidad. Estas pruebas deben ser de varios tipos:

Totales: que incluyen la simulación de fallos en los sistemas, anunciadas o no, deben incluir a proveedores externos si el plan de continuidad lo especifica como necesario, y servirán para medir los tiempos de restauración de los servicios entre otros.

Parciales: pueden abarcar partes del sistema únicamente, por ejemplo un único servidor o un único sistema.

Escenarios: aquí se probaran las reacciones de los planes de continuidad frente a situaciones específicas.

En la etapa de operaciones, se incluyen todas aquellas actividades posteriores a la implementación de los planes, por ejemplo la capacitación del personal asociado a alguna de las actividades del plan, la revisión de los riesgos, impactos, y el plan en general, las pruebas y el manejo del cambio al plan inicial.

La última etapa, la invocación se refiere al momento en que el plan de continuidad de los servicios de TI actúa, esto podrá ser en cualquier momento, lo cual hace que el personal deba estar preparado a toda hora. El plan de continuidad deberá definir claramente las acciones a tomar durante la invocación, por ejemplo, el flujo a seguir para la restauración de respaldos almacenados en media externa, la movilización del personal necesario al sitio de restauración, y la gestión de las comunicaciones con los proveedores involucrados.

1.1.4. CISSP

CISSP (Sus siglas en inglés *Certified Information Systems Security Professional*) considera la continuidad del negocio dividida en dos grandes ramas:

- La planeación de la continuidad del negocio
- La recuperación de desastres

Para CISSP el proceso de BCP incluye 4 pasos principales:

- Delimitación de alcances del proyecto
- Análisis de impactos
- Plan de continuidad
- Implementación

El primer paso a realizar para implementar BCP mediante CISSP debe ser un análisis del negocio, los departamentos involucrados en las operaciones críticas y el rol que juega cada empleado dentro de este contexto. Esto sirve para definir las bases iniciales del proyecto, así como los posibles miembros del equipo para BCP. Este equipo debería estar integrado por al menos:

- Representantes de cada departamento de la organización.
- Representantes de cada departamento que realice acciones críticas para la organización.
- Personal de TI con experiencia a nivel técnico en cada una de las áreas del BCP.
- Responsables de seguridad con conocimientos del BCP.
- Representantes de la alta dirección.
- Representantes legales, familiares con las cuestiones legales de la organización.

Es muy importante también definir los recursos con los que se cuenta para cada una de las fases del BCP. Aunque el mayor porcentaje de recursos utilizados será el trabajo deben considerarse también recursos de *hardware* y *software* que sean utilizados en alguna parte del proceso.

CISSP considera también, como se vio antes, muy importantes aquellas cuestiones legales que afecten la implementación de un plan de continuidad, esto se refiere primordialmente a aquellos servicios que por su nivel de impacto pueden causar problemas legales a la organización, por ejemplo, esto es aplicable en organizaciones que prestan servicios críticos a una comunidad como la policía, bomberos, etc. En donde la falla del servicio inicial, sumado al fracaso de las estrategias definidas en el BCP puede conllevar en ocasiones incluso a la pérdida de vidas humanas.

Estas restricciones legales de continuidad, dependen del medio en que la organización realice sus actividades, puesto que en algunos países, los servicios financieros por ejemplo, son regulados por el gobierno, y tienen la obligación de prestar sus servicios para salvaguardar la seguridad de la economía del país.

La tercera forma de regulaciones legales que pueden afectar el desarrollo del BCP, son los contratos con los clientes de la empresa, debido a que estos normalmente incluirán entre sus cláusulas la disponibilidad de los servicios que la organización presta.

Llegado a este punto, CISSP analiza los resultados de las fases anteriores, y se definen prioridades en los procesos de la empresa, estas prioridades son definidas con base al valor de cada uno de los activos tangibles o intangibles, y este valor es definido en términos cualitativos o cuantitativos según corresponda. Con base en estos activos, el equipo de BCP deberá definir el MTD o tiempo máximo tolerable de inactividad, es decir, el tiempo máximo que la empresa puede detener sus operaciones por fallas no controladas, sin que esta sufra daños irreparables de cualquier tipo.

CISSP cataloga también los riesgos de manera que estos tengan tanto peso como probabilidad de ocurrir, es decir, si la zona en que el negocio opera es muy vulnerable a erupciones volcánicas, este riesgo tendrá una mayor probabilidad de ocurrencia que otros.

CISSP calcula los costos cuantitativos con base en una serie de fórmulas[4]:

Factor de exposición (FE): la cantidad de daño que un riesgo puede ocasionar a un activo.

La expectativa única de perdida (SLE): la pérdida que ocasionaría este riesgo de concretarse sobre el activo.

La pérdida anual para un riesgo (ALE): es la pérdida que se tiene anualmente. Debido al riesgo en cuestión, puede calcularse mediante la fórmula: ALE=SLE*probabilidad de ocurrencia en un año.

Basado en estos cálculos cuantitativos, resulta fácil definir una serie de prioridades en el análisis de riesgos, deberán cubrirse primero aquellos riesgos que eventualmente pudieren representar un riesgo alto para la empresa.

CISSP considera también como parte clave del proceso de BCP la documentación del mismo, por varias razones, incluyendo entre ellas:

La correcta activación del plan en caso de una emergencia, incluso si el jefe del proyecto no se encuentra disponible.

Los miembros que se integren al equipo a futuro pueden entender el porqué de la implementación mediante la documentación.

La gestión del cambio sobre los planes resulta más sencilla si todo se encuentra debidamente documentado.

Por último, se tiene que CISSP define una serie de secciones básicas para el plan de continuidad escrito[4], estas podrán ser modificables de acuerdo al negocio:

- Objetivos del plan
- Declaración de importancia para el negocio
- Declaración de prioridades
- Declaración de responsabilidades organizacionales
- Declaración de tiempos de respuesta
- Evaluación de riesgos
- Mitigación de riesgos
- Guías para la respuesta a emergencias
- Mantenimiento
- Pruebas

1.1.5. BS25999

BS2599 es una norma específicamente diseñada para la continuidad del negocio, esta norma no ve la continuidad del negocio como creación y la gestión de un plan, si no como un sistema de gestión integral que debe integrarse en las actividades diarias de la organización[3].

BS2599 está dividido en dos partes:

BS25999-1: el código de prácticas de la norma

BS25999-2: la especificación de la norma en sí

BS25999 se centra en la creación de un sistema de gestión de la continuidad del negocio (BCMS), así, los pasos principales definidos por la norma son:

- Planeamiento del BCMS: que incluye el análisis de los objetivos de la organización, la toma de requerimientos, la evaluación de riesgos, la definición de responsabilidades de los involucrados, definición de políticas, evaluación y obtención de recursos, entrenamiento del personal, documentación inicial y creación de registros iniciales, acá pueden incluirse entrevistas al personal o estudios.

[3] JHONSON, Perry. Steps to BS25999 Registration. p. 4

- Implementación y operación del BCMS: que incluye el análisis de impacto al negocio, evaluación de riesgos sobre activos, toma de decisiones para tratar los riesgos y estrategias de mitigación, determinación de la estrategia para continuidad, determinación del personal encargado de responder frente a situaciones de fallo y ejercicios con el plan.

- Monitoreo y revisiones del BCMS: que incluye la revisión interna por parte de la organización de su BCMS, para asegurarse de que el plan sigue siendo válido para los objetivos y activos actuales de la empresa.

- Mantenimiento y mejora del BCMS que incluye la toma de acciones tanto preventivas como correctivas para mejorar el BCMS y la mejora continua del mismo.

Uno de los puntos más fuertes de BS25999 es que considera riesgos que otras normas no consideran o no definen a un nivel tan especifico, por ejemplo, considera que pasaría si un empleado de mucha importancia en las operaciones se ve incapacitado, que pasa si los servicios de TI dejasen de funcionar por un periodo de tiempo indefinido o que pasaría en caso de un ataque terrorista que afecte a la organización. A esto se debe agregar, que la primera parte de la norma define prácticas generales para cualquier organización de cualquier tamaño, con lo cual se obtiene una guía para la elaboración del BCMS más conveniente para cada caso específico.

BS25999 hace un bastante énfasis en la gestión de la documentación del BCMS, definiendo incluso una plantilla para la gestión de cambios sobre la documentación, esta deberá incluir como minimo[6]:

- No. de documento

- Autor de la revisión
- Razón para la revisión
- Aprobado por

El contenido básico de un plan de continuidad según BS25999 es:

- Visión general del plan
- Dependencias/Requerimientos
- Información de contactos
- Equipos de recuperación y roles dentro de estos
- Procedimientos de respuesta ante alertas
- Procedimientos de recuperación
- Procedimientos de restauración

La auditoría interna del plan es uno de los puntos más fuertes de BS25999, una auditoría básica debería contar con al menos los siguientes pasos:

- Revisión de los programas

- Análisis de la organización

- Determinar la estrategia de continuidad del negocio

- Diseñar e implementar una respuesta BCM

- Ejercitar y revisar los acuerdos de BCM

Con este tipo de revisiones y rediseños constantes es que la esencia de BS25999 es alcanzada, como se dijo al inicio, se crea una cultura donde los planes de continuidad del negocio son integrados a las actividades diarias de la organización.

1.1.6. BS2577

Desarrollado por el BSI, a partir del PAS 77, es una norma orientada a complementar al BS25999, es una norma destinada puramente a la continuidad de los servicios de tecnología. Esta norma fue publicada siguiendo el método de BS25999 siendo la primera parte BS25777-1 un conjunto de buenas prácticas para la gestión de la continuidad de TI, y BS25777-2 que define los requisitos para establecer, implementar y administrar en si un sistema para la continuidad de los servicios.

Los principales beneficios de BS25777 son:

- Ayuda a identificar las principales amenazas para los servicios de TI

- Identifica el potencial de los fallos en los servicios de TI

- Asegura que la empresa tendrá el nivel de soporte necesario en las herramientas de TI, de acuerdo a su plan de continuidad del negocio.

- Provee herramientas de TI que son adecuadas basándose en sus costos, y la dependencia que el negocio tiene de éstas.

- Provee una conexión entre el BCP y el ITSCM.

- Muestra los posibles riesgos sobre los servicios de TI y la forma de mitigarlas.

1.2. Implicaciones de los periodos de inactividad

Hasta aquí, se ha discutido sobre las pérdidas que un período de inactividad puede ocasionar a una empresa. Estas implicaciones pueden ser pérdidas directas o indirectas. Las pérdidas directas se refieren puramente a pérdidas económicas derivadas de la caída del sistema en un periodo determinado de tiempo. Las pérdidas indirectas comprenden aquellas que pueden significar pérdidas económicas, pero inicialmente a corto o mediano plazo lo representarán, por ejemplo, daños a la imagen, a la credibilidad, etc.

En promedio, un minuto de inactividad costara para una empresa $1400.00, basado en este cálculo, se tiene que 43 horas de inactividad resultarán en un costo de US 3.6 millones, así como que una hora de inactividad representara pérdidas por US 84000[4]

[4] BAJGORIC, Nijaz. Continuous computing technologies for enhancing business continuity. p. 24

Estos análisis de costos para tiempos de disponibilidad, llevan a la necesidad de medir el porcentaje de tiempo que un servicio se encuentra disponible (ya sea desde el punto de vista del negocio o de TI), normalmente el porcentaje de disponibilidad se mide por una cifra como 99.99% o 99.999%, y puede hablarse entonces de “cuantos 9” de disponibilidad se tienen, una estadística se muestra en la tabla I:

Tabla I. **Porcentajes de disponibilidad**

Porcentaje de disponibilidad	Tiempo de inactividad por día	Tiempo de inactividad por mes	Tiempo de inactividad por año
95	72 minutos	36 horas	18.26 días
99	14.4 minutos	7 horas	3.65 días
99.9	86.4 segundos	43 minutos	8.77 horas
99.99	8.64 segundos	4 minutos	52.6 minutos
99.999	0.86 segundos	26 segundos	5.26 minutos

Fuente: BAJGORIC, Nijaz. *Continuous computing technologies for enhancing business continuity*. p. 25.

Un porcentaje del 100% de disponibilidad es prácticamente imposible de conseguir, y en caso de ser posible, resulta demasiado costoso, debido a esto, las empresas toman un porcentaje de disponibilidad lo más cercano al 100% con base a sus necesidades de operación, se debe tomar en cuenta que existen tiempos de indisponibilidad planeados y no planeados, y estos podrían afectar el porcentaje de disponibilidad seleccionado.

Los costos de tiempo de inactividad varían de un sector a otro, esto puede apreciarse en la tabla II.

Tabla II. **Costos por tiempos de inactividad**

Aplicación	**Costo por Minuto**
Telefónicas	US 27000.00
ERP	US 13000.00
SCM	US 11000.00
Comercio electrónico	US 10000.00
Banca en línea	US 7000.00
Servicios personales universales	US 6000.00
Servicio al cliente	US 3700.00
POS	US 3500.00
Mensajería	US 1000.00

Fuente: BAJGORIC, Nijaz. Continuous computing technologies for enhancing business continuity. p. 28.

2. REDUNDANCIA

En sistemas de información, redundar es contar con dos sistemas que sean capaces de efectuar las mismas funciones ya sea con la misma capacidad o con una capacidad reducida, pero garantizando la continuidad de los servicios de TI.

En el ámbito de TI es posible redundar tanto en *hardware* como en *software*, por ejemplo, puede implementarse una arquitectura en la cual exista un servidor puramente dedicado a contingencia, esto es conocido como un clúster activo/pasivo.

El servidor pasivo en este caso, existe puramente para cubrir incidentes, y no entra en funciones si no cuando el servidor principal presenta fallos que le impiden seguir trabajando. En un escenario ideal, el servidor pasivo debería tener las mismas especificaciones que el servidor activo (el que realiza las operaciones normalmente), pero el costo de esto puede ser demasiado alto. Es por ello que, normalmente esta arquitectura se definirá pensando en el plan de continuidad del negocio o en el plan de continuidad de servicios de TI en caso de que este exista, puesto que en estos planes deberán estar definidas claramente cuáles son las operaciones prioritarias de la organización, y en consecuencia cuales son los servicios estrictamente necesarios para continuar estas operaciones.

La redundancia a nivel de *software* debe ser también un aspecto a tomar en cuenta al momento de definir los sistemas informáticos a utilizar para garantizar la continuidad de los servicios de tecnología. Es posible que la redundancia de *software* no represente costos tan altos como representa la redundancia de *hardware*, esto dependerá normalmente de las licencias del software que se utiliza. Las implicaciones que pueda tener la redundancia de software pueden al contrario llegar a ser más graves que en el caso del hardware, puesto que es necesario asegurarse de que los datos existentes en el servidor de redundancia son idénticos a los del servidor principal en un instante de tiempo determinado, esto normalmente implicará cuestiones relativas a la replicación de sistemas de software, tratada en el capítulo tres.

2.1. Redundancia de sistemas de *hardware*

La redundancia a nivel de *hardware*, como se ha dicho antes, deberá ser una opción evaluada minuciosamente antes de ser implementada, principalmente por el alto costo que una implementación de este tipo conlleva. El análisis debería conllevar un estudio de cuáles son los servidores con funciones críticas para el negocio. Este análisis se puede sustentar con el plan de continuidad del negocio, podrá saber el costo aproximado de un período de inactividad en un servidor específico, o, según sea el caso, de una red de servidores que, en conjunto, soporten alguna funcionalidad de la organización. Una vez que se sabe cuál es el costo de la inactividad de un sistema, deberá considerarse también la probabilidad de que esta inactividad suceda, con lo cual se tendra un costo promedio anual (consultar capítulo 1), y, en caso de que este costo sea imposible de asumir (tomando en cuenta no solamente los costos financieros), deberá analizarse el costo de redundar en la arquitectura de *hardware*, y solo una vez que este último costo resulte inferior al primero, se implementará la redundancia.

2.2. Redundancia de sistemas de *software*

La redundancia de sistemas de software, suele ser un caso menos costoso pero más difícil de administrar, puede resultar menos costoso para el *software* desarrollado en casa, puesto que este no tendrá ningún incremento al costo por ser instalado en un servidor adicional, pero esto puede cambiar en el caso de que se use una licencia de terceros, y esta licencia incluya cláusulas que definan algún costo por instalar un servidor extra.

La redundancia de *software* no debe ser confundida con replicación de sistemas de *software*, cuando hablamos de redundancia, nos referimos puramente al *software* que existe más de una vez con los mismos fines en un mismo entorno, aunque es claro que esto muchas veces implicara un sistema de replicación de datos.

Existen al menos dos escenarios en donde puede implementarse redundancia de *software*:

- Sistemas redundantes que necesitan un sistema de replicación: estos sistemas normalmente serán utilizados para brindar alta disponibilidad de servicios de TI, y sus conjuntos de datos deberán ser iguales probablemente incluyendo un retardo en el tiempo.

- Sistemas redundantes en los que no se necesita un sistema de replicación: son sistemas que no necesitan datos replicados para poder trabajar, por ejemplo sistemas cuyo juego de datos se borra después de determinados períodos de tiempo. Estos sistemas normalmente son usados como sistemas de contingencia, pueden ser habilitados para un período de operación, y luego deshabilitados.

- Nodos redundantes que comparten conjuntos de datos: un tercer caso es una implementación en la que existe alta disponibilidad mediante la existencia de varios servidores replicados a nivel de *hardware* y *software*, cuyas instancias pueden encontrarse operando o a la espera (servidor secundario activo o pasivo), y donde estas instancias comparten un juego de datos para operación, siendo que este conjunto de datos no estará alojado en ninguno de los servidores replicados. La disponibilidad de los datos en este tipo de implementaciones será asegurada normalmente con sistemas de almacenamiento de red como SAN, NAS, etc.

2.3. Implicaciones financieras

Para definir las implicaciones financieras de implementaciones redundantes, deberá tomarse en cuenta la redundancia de sistemas de *hardware* como de *software*, basándose en todos los parámetros antes dichos. Es probable que el departamento de tecnología interesado en implementar una solución de este tipo necesite el soporte de otros departamentos dentro de la organización, puesto que normalmente son otros departamentos quienes tienen conocimiento de los costos por período de inactividad para el negocio.

Es importante hacer notar que el trabajo que realiza el área de tecnología será únicamente una parte de la creación de una cultura de continuidad del negocio, por lo cual los departamentos involucrados pueden ser muchos, siendo que los costos operativos en este caso no corresponden al área de tecnología. Esta puede dar únicamente un costo de la implementación de las soluciones y formular un análisis de costo/efectividad basándose en datos que le sean proveídos por otros departamentos.

El cálculo del costo de implantar un sistema redundante deberá realizarse con base en varios rubros:

- Costo del personal que realiza la implementación.
- Servidores.
- Equipo de *hardware* en general.
- Licencias de *software.*
- Personal necesario para la administración del sistema redundante.
- Modificaciones a salas de servidores (por ejemplo cambio en las necesidades de aire acondicionado, etc.).
- Modificaciones necesarias a la estructura de red.
- Necesidad de nuevos proveedores (por ejemplo la compañía que brinda enlaces si se implementan sitios redundantes en locaciones distintas).
- Capacitaciones para el personal involucrado en la administración.

2.4. Rendimiento de los sistemas redundantes

Al analizar las implicaciones de rendimiento para un sistema redundante, ya sea por *hardware* o por *software*, debe tomarse en cuenta la variación de resultados que existirá entre los distintos escenarios de implementación que hasta aquí se han mencionado.

En el escenario en que la redundancia es utilizada sin replicación de datos, por ejemplo, cuando el almacenamiento se da en una SAN, el rendimiento puede verse aumentado al agregar redundancia de *software* y *hardware*, esto porque puede realizarse una configuración de nodos activos del sistema, con lo cual se puede aumentar las capacidades de rendimiento del sistema mediante su crecimiento horizontal.

Se debe considerar también el caso en que uno o más servidores redundantes se incorporan al sistema como servidores pasivos, es decir, servidores que se encuentran disponibles a la espera de un fallo del servidor principal para iniciar su funcionamiento, estos servidores no generan ninguna carga al rendimiento del sistema, hablando por supuesto del caso en que no se implemente replicación de datos.

Probablemente el escenario en que la replicación tenga una implicación más fuerte sobre el rendimiento del sistema, sea aquel en que el sistema maestro y el sistema esclavo tengan juegos de datos separados lo que conlleva la replicación de estos datos y si la replicación es síncrona puede aumentar considerablemente los tiempos de respuesta de los servidores involucrados.

En general, puede decirse que el rendimiento de un sistema redundante dependerá más bien de los componentes del sistema, siendo éstos:

- Capacidad de los enlaces
- Sistema de replicación de datos
- Capacidad de los servidores involucrados
- Rendimiento del *software* utilizado
- Rendimiento del medio de almacenaje (para sistemas sin replicación)

Sabiendo esto, será necesario estudiar en cada caso las capacidades con las que se cuenta, basándose en estos parámetros, y las necesidades del negocio, para decidir si es necesario o no realizar la implementación del sistema redundante.

3. COMUNICACIONES

Cuando se habla de sistemas altamente disponibles, los enlaces que soportan las comunicaciones resultan de vital importancia. Desde los enlaces, que proveen conexión entre los servidores internos de un centro de datos, hasta los que proveen conexión entre distintos centros de datos, resultan sumamente importantes, en tanto que son el medio por el cual los datos serán replicados de un lugar a otro. Es por ello que se analizan diversos métodos para asegurar la disponibilidad de los enlaces primordiales para la continuidad de un negocio, y se debe recordar que este análisis debe hacerse no solo desde el punto de vista de la existencia de comunicación entre dos puntos específicos, si no también incluyendo la capacidad que este enlace tendrá para soportar el trafico al que sería sometido en el peor de los casos.

El desarrollo de un plan de continuidad de servicios de TI deberá contemplar las pruebas de capacidad de los enlaces, así como la redundancia de los enlaces primordiales para el funcionamiento del sistema. Existe una serie de estándares que se pueden estudiar previo a la implementación de una solución para la continuidad de los servicios de TI, entre los cuales se encuentra 803.ad, protocolos como STP, o tecnologías de comunicación como FDDI o canales de fibra óptica. A lo largo de este capítulo se estudiará con detenimiento cada uno de estos temas.

3.1. Redundancia de enlaces

El primer tema a tratar en cuanto a enlaces y comunicaciones, será la redundancia. Básicamente existen tres escenarios en los cuales es aplicable la redundancia:

- Redundancia de enlaces entre centros de datos
- Redundancia de enlaces hacia terceros
- Redundancia de enlaces para la red interna

A continuación se analiza cada uno de estos casos y su importancia.

3.1.1. Redundancia de enlaces entre centros de datos

Probablemente, el tipo de redundancia más importante a contemplar en la implementación de un plan de continuidad de negocio desde el punto de vista de tecnología. Este tipo de enlace proveerá los medios necesarios para replicar los datos existentes en el sitio primario hacia el o los sitios secundarios, en caso de desastre o fallos en el sitio primario, estos enlaces serán los encargados de sincronizar los datos hacia el sitio secundario, en caso de ser necesario.

Este tipo de enlace es normalmente administrado y proveído por terceros, con lo cual la administración del mismo no será un punto demasiado importante para la empresa. Aún así, recae sobre ésta la responsabilidad de elegir las capacidades de los enlaces y la cantidad de los mismos, esto deberá analizarse pensando en diferentes escenarios de implementación y fallas, los cuales se describen a continuación:

3.1.2.1. Replicación síncrona

En la replicación síncrona, la capacidad de los enlaces resulta crítica en tanto que afecta directamente el rendimiento de los sistemas de producción. La replicación síncrona espera tener seguridad sobre los datos que fueron almacenados en todos los sitios involucrados antes de dar por válida cualquier transacción. Con esto, si se utilizan enlaces de capacidades menores a las requeridas, el rendimiento del sistema de producción se verá mermado en gran medida. Las capacidades de estos enlaces se verán afectadas por una serie de factores entre los cuales se puede incluir: cantidad de enlaces redundantes, cantidad de sitios de replicación, cantidad de enlaces redundantes activos/pasivos, cantidad de sitios con replicación síncrona, distancia entre los sitios de replicación, capacidades puramente de los cables utilizados en los enlaces y necesidades de carga por transferencias de datos de las aplicaciones ligadas a la replicación de sistemas.

3.1.2.2. Replicación asíncrona

En este caso, las capacidades de los enlaces no resultan un factor tan crítico como en el caso anterior, el análisis de la capacidad de estos enlaces, es algo muy enfocado al negocio en estudio. Pero en general, se sabe que las capacidades de los enlaces no influirán en el rendimiento de los sistemas de producción. Aun así, se debe considerar el rendimiento desde el punto de vista de que los enlaces no se saturen con envió/recepción de datos, puesto que esto si podría causar caídas en el sistema principal, o la pérdida de datos en el sitio secundario, lo que probablemente anularía los beneficios de los sitios alternos de replicación.

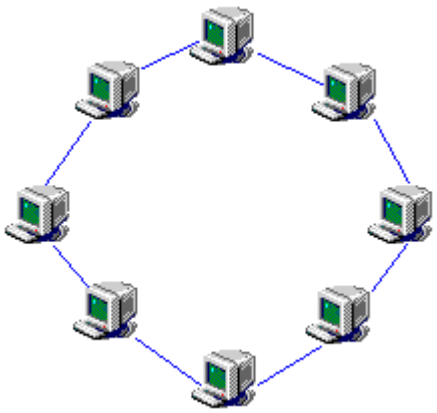
3.1.2.3. Sincronización de datos para recuperación

En esta definición se engloba todas aquellas operaciones que necesiten sincronización de datos, entre centros de datos sin llegar a ser consideradas replicación síncrona o asíncrona. Se tiene por ejemplo, aquellos ambientes secundarios que son sincronizados para ser utilizados como ambientes de pruebas diferidas, almacenamiento de sistemas ODS, o históricos que pudieran en algún momento almacenarse en un centro de datos, dedicado a estos fines. Tomando en cuenta que estas no son operaciones críticas de la empresa, deberá analizarse muy detenidamente la necesidad de redundar en los enlaces hacia estos centros de datos.

Son muchas las formas que existen de interconectar dos o más centros de datos remotos. En Guatemala se puede encontrar empresas como Telecomunique, Pronet o Amnet, las cuales proveen enlaces de fibra oscura a través del área metropolitana. Una de las mayores diferencias que se pueden encontrar entre estos proveedores será la topología que utilizan en su red metropolitana.

Por ejemplo, en el caso de Amnet, se maneja una topología de anillo metropolitano, lo cual conlleva la ventaja de contar con varias rutas, ubicadas en distintas locaciones geográficas para, alcanzar un destino desde un origen cualquiera.

Figura 2. **Topología de anillo**



Fuente: elaboración propia.

La figura 2 muestra una topología típica de anillo, tomando en cuenta que esta red está distribuida en toda el área metropolitana, la probabilidad de la rotura de comunicación en cualquier punto a través de todo el anillo es bastante alta. La experiencia demuestra, al menos en el caso de Guatemala, que los riesgos que afrontan estos enlaces incluyen construcciones o modificaciones a las estructuras de la vía pública, desastres naturales, etc. La ventaja principal del anillo metropolitano, es que de por sí, ya implementa ciertas características de redundancia y contingencia, puesto que es posible llegar desde un punto A, hasta un punto B, de dos formas distintas, atravesando enlaces con locaciones geográficamente muy distintas. Así, si un punto del anillo es fracturado por alguna razón, la conectividad seguirá existiendo para todos los puntos del mismo, así, la red metropolitana normalmente ofrece disponibilidad arriba del 97% de tiempo.

Empresas como Optel o Pronet en Guatemala ofrecen servicios de conectividad, mediante fibra oscura (se debe consultar con los proveedores), en cada caso, es necesario analizar las capacidades de lo que cada una ofrece, la capacidad de los enlaces, y la cantidad de hilos de fibra oscura son factores muy importantes a tomar en cuenta también. Muchas veces, estos requerimientos vienen dados por el *hardware*/*software* que se utilice para alta disponibilidad, buen ejemplo de ello es el caso de las tecnologías para replicación, las cuales definen un mínima de dos hilos de fibra oscura para replicación síncrona/asíncrona entre centros de datos, así como también dictaran el ancho de banda necesario para cada caso según las capacidades del *hardware*/*software* adquirido.

3.2. Alta disponibilidad en redes locales

Es muy importante asegurar las operaciones desde el punto de vista de la red local, entre las opciones para esto, se tienen las siguientes:

3.2.2. STP (Sus siglas en inglés *spanning tree protocol)*

Supóngase una red donde existen enlaces que aseguran la alta disponibilidad mediante redundancia entre los segmentos de red A y B. El problema con esta topología de red, es que un paquete enviado desde un segmento cualquiera, puede verse inmerso en un ciclo infinito, esto sería de la siguiente manera:

Desde el segmento A se envía un paquete que es recibido por el puerto 1 del switch A.

El switch A lo envía hacia el puerto 1 del switch B.

El switch B envía el paquete por su puerto 2.

El switch A recibe el paquete en su puerto dos, y reinicia el proceso.

Esto puede crear un ciclo que deje inutilizable la red, en caso que ocurran muchas incidencias de este caso, lo cual es bastante común, esto se conoce como tormenta de broadcast. Para auxiliarnos existe el protocolo STP encargado de deshabilitar puertos y enlaces que podrían crear ciclos en la red, el protocolo incluye un algoritmo que mantiene deshabilitados los enlaces y puertos redundantes, pero al momento en que los primarios o activos pierdan la conectividad, los primeros serán habilitados.

Redefiniendo la red, en cuestión de cuales enlaces/puertos se encuentran activos y cuáles no lo están. El protocolo se asegura de crear redes sin redundancia activa, pues pretende, como su nombre lo indica, crear una topología de árbol, en la cual sólo existe un camino que interconecta dos segmentos cualesquiera de la red.

El algoritmo mediante el que trabaja STP podría definirse en cuatro pasos:

Definir el switch de mayor prioridad.

Encontrar un camino desde cada switch en la red hasta el switch principal, y calcular el costo de este camino, se elige en base a este cálculo los puertos principales de cada switch.

El cálculo de caminos y puertos debe considerar que no puede bloquearse los puertos que están incluidos en otros caminos.

Si existiera alguna tormenta de broadcast, se bloquea automáticamente el puerto involucrado en cada switch.

3.2.3. Cluster de servicios básicos

Una decisión importante, y no demasiado compleja de implementar, es la creación de clúster para los servicios básicos ofrecidos por la red. Esto dependerá mucho de los servicios que se presten, puede pensarse por ejemplo, si se utiliza un servidor para prestar servicios críticos, podría crearse un clúster activo/pasivo o activo/activo de manera muy sencilla con herramientas incluso gratuitas, por ejemplo, *Heartbeat*, esto reduce en gran medida los costos de recuperación. Se analizara el siguiente caso:

Se tiene dos centros de datos, funcionan con replicación síncrona en configuración activo/pasivo, en el sitio activo ocurre un problema con el servidor, que presta un servicio crítico para la empresa. Este servidor deja de funcionar, en caso que se tenga un servidor de contingencia, este deberá ser promovido por el personal de tecnología, lo cual normalmente tendrá un costo en cuanto a tiempo y otros costos quizás no tan visibles. En caso que no exista el servidor de contingencia en el sitio primario, las cosas se complican aún más, puesto que habría que tomar el sitio secundario, y promover únicamente el servidor afectado en el mismo, lo cual una vez más, conllevará costos de soporte, requerirá la disponibilidad de los enlaces de comunicación, y, en casos, las reconfiguraciones necesarias de red para poder acceder al nuevo servidor primario. Todas estas complicaciones se pueden evitar utilizando *cluster* de servicios en los servidores críticos del sitio principal (en un caso ideal, se replicará esta configuración en el sitio secundario).

Con esta implementación, la recuperación en caso de un fallo en el servidor primario del sitio activo resultará transparente para los usuarios, y más sencilla para el personal de tecnología.

4. REPLICACIÓN

Uno de los métodos más importantes cuando se habla de alta disponibilidad de sistemas informáticos, es la replicación de sistemas. Replicar es, en resumen, crear copias idénticas del sitio primario en un sitio alterno. Esto implica que al momento de perder al sitio primario, en un lapso determinado, las operaciones puedan reanudarse, promoviendo al sitio secundario como sitio primario. Existen dos tipos de replicación: síncrona y asíncrona, cada uno de ellos con sus respectivas implicaciones. En este capítulo se analizan las ventajas de cada uno de estos tipos, así como los diversos métodos, tecnologías utilizadas, e implicaciones de implementar sitios de replicación.

Es importante decir que la replicación de sistemas asegura la existencia de una copia fiel de los datos en el sitio primario, esto implica que los datos serán exactamente los mismos en los sitios secundarios, por ello, en caso de que existiere una corrupción de datos en el sitio primario, esta existirá también en el sitio secundario. Por ende, un sistema de replicación deberá ir acompañado normalmente de un sistema orientado a crear copias instantáneas.

La replicación deberá ser un método para salvaguardar los sistemas de fallos más bien físicos en los servidores o cuartos de servidores. Por ejemplo, puede pensarse en el caso en que ocurra un siniestro sobre el sitio primario, en este caso, podría promoverse un sitio secundario y continuar las operaciones sin mayor retardo.

4.1. Métodos de replicación

Cuando se piensa en replicación de sistemas, la primera pregunta que surge es qué se va a replicar. Puede replicarse todo el sistema operativo, los datos de una aplicación específica, los sectores del disco duro, o incluso todo el servidor, o conjuntos de servidores. Las tecnologías de vitalización aumentan aún más las opciones que se tienen al momento de implementar replicación, así, es posible por ejemplo replicar el almacenamiento de un servidor que contenga varias máquinas virtuales, lo cual simplifica en casos el proceso de promover el sitio secundario.

4.1.2. Replicación de dispositivos de almacenamiento en ambientes virtualizados

Este método está basado en la replicación de los discos duros o medios de almacenamiento, resulta muy útil en implementaciones con virtualización, así, es posible replicar los discos duros que contienen todas las máquinas virtuales utilizadas. La principal ventaja de una implementación con máquinas virtuales y replicación del almacenamiento, será que el proceso de promoción del sitio secundario es más sencillo. Por ejemplo, puede pensarse en un sitio primario con N servidores virtuales, que podrían ejecutarse sobre uno o más *blades* y utilizar como almacenamiento una SAN. Al momento de promover al sitio secundario, se asegura que las máquinas virtuales del mismo estarán actualizadas, tanto en los datos de sus sistemas, como en las configuraciones de los mismos y del sistema operativo. Se debe agregar que, suponiendo que el sitio secundario sea un sitio pasivo, las máquinas virtuales pueden conservar las mismas direcciones IP, sin tener la necesidad de reconfigurar las conexiones a servidores en todos los sistemas.

Todo esto hace más rápido el proceso de restauración, puesto que, al momento de un fallo en los servidores primarios, las tareas a realizar se reducen a:

- Cambiar el papel de los sitios (primario a secundario y viceversa)
- Encender las máquinas virtuales del sitio secundario
- Iniciar los servicios que corran sobre estas máquinas virtuales

Como se dijo antes, esto sólo es posible porque las máquinas virtuales en el sitio secundario se suponen apagadas (configuración activo-pasivo). La experiencia demostrará también el impacto de no necesitar realizar cambios en las configuraciones de conexiones, primordialmente en centros de datos que alberguen muchos sistemas, esto requeriría asistencia de personal técnico para realizar estos cambios, que en el peor de los casos deberían ser aplicados sobre cada una de las máquinas cliente en el sistema que podrían encontrarse en ubicaciones muy distintas geográficamente. Una de las desventajas de esta implementación, es que si existiera una corrupción a nivel de sistema operativo en alguno de los servidores virtuales, esta falla será replicada al sitio secundario, lo cual dejaría inutilizables ambas copias. Otra de las desventajas en esta implementación, puede ser el costo, ejecutar máquinas virtuales requerirá normalmente de servidores *blade*, si a esto se añade que es necesario el medio de almacenamiento, el costo resulta bastante elevado, puesto que la arquitectura de *hardware* debe existir tanto en el sitio primario como en el sitio secundario.

Esta arquitectura cuenta con ventajas que van más allá de las relativas a replicación. Una de ellas y probablemente la mayor, es que, al momento de una falla en el *hardware* de un servidor, los archivos que componen las máquinas virtuales pueden ser copiados a otro servidor incluso local, y podrán ser ejecutados en este nuevo servidor, sin necesidad de promover al sitio secundario. El negocio podría establecer una serie de servicios críticos y agrupar sus blades en función de esta definición.

Las ventajas de virtualizar van desde la restauración rápida de servidores hasta la facilidad de promover un sitio secundario al momento de una emergencia en el primario, sin mayores complicaciones, a esto se debe agregar las facilidades en cuanto a creación de copias instantáneas de los servidores, que pueden ser utilizados para pruebas, certificación de actualizaciones, entre otros sin necesidad de detener los servicios críticos.

4.1.3. Replicación de dispositivos de almacenamiento de los servidores

Otro método de replicación a considerar, es la replicación de los dispositivos de almacenamiento de los servidores utilizados. Este método de replicación, resulta un tanto más complejo de administrar que el anteriormente explicado. Consiste básicamente en replicar los datos contenidos en los servidores utilizados en el sitio primario.

Por ejemplo, si en el sitio primario existen tres servidores, todos podrán existir físicamente o como máquinas virtuales en el sitio primario, estos existirán de la misma manera en el sitio secundario, lo que se replica en este caso son las unidades de almacenamiento de estos servidores, y no los servidores en si como en el caso anterior.

Esto conlleva serias implicaciones en la administración del sistema, por ejemplo, asumiendo que el sistema trabaja en modo activo-pasivo, los servidores en el sitio secundario deberán permanecer encendidos, puesto que es su sistema operativo el que organiza los datos dentro de su almacenamiento.

El hecho de que los dos servidores (primario y secundario) permanezcan encendidos todo el tiempo, implica que estos no pueden tener asignada la misma dirección IP. Con esto, se aumenta de por si el tiempo requerido para promover el sitio secundario.

Por ejemplo, asumiendo un servidor que almacena un DBMS, este en el sitio primario tiene asignada la IP 190.190.X.X, esta ip debe ser única dentro de la red, así que el servidor secundario debería tener asignada una dirección ip distinta. Esto implica que al momento de promover el sitio secundario, se deba actualizar todas las referencias en aplicativos hacia la nueva dirección IP (la del servidor secundario). Este proceso de cambio, en el peor de los casos y dependiendo de la arquitectura de los sistemas utilizados, podría implicar el realizar cambios de configuración en todos los terminales que se conecten al sistema, puede pensarse por ejemplo en sistemas configurados por ODBC, los cuales habrían de ser cambiados manualmente en cada terminal. El costo de hacer esto, puede llegar a ser demasiado alto, puesto que tendría que realizarse cambios en ubicaciones separadas geográficamente, sería necesario asignar personal técnico que realice estos cambios.

Entre las ventajas de esta implementación, puede mencionarse por ejemplo que los sistemas operativos de los servidores no necesariamente van a replicarse, puesto que pueden ser servidores distintos y aislados, que comparten sus datos, pero no así sus sistemas operativos. Entonces, si en algún momento existe una corrupción a nivel de sistema operativo (por actualizaciones fallidas, etc.) esto no se replicará al sitio secundario, con lo cual se tiene mayor tolerancia a fallos.

Aun cuando esto representa una ventaja desde el punto de vista de la continuidad del negocio, conlleva la desventaja que cualquier configuración sobre el sistema operativo del servidor deberá realizarse al menos dos veces: tanto en el sitio primario, como en el sitio alterno.

De las dos opciones planteadas hasta aquí como métodos de replicación, se han analizado las ventajas y las desventajas, y se puede decir que la que conlleva más ventajas es la primera, replicar unidades de almacenamiento que contengan máquinas virtuales, primordialmente por las facilidades que ofrece en administración y promoción de sitios secundarios. La mayor de las desventajas en esta implementación, puede resultar ser el costo a corto plazo de las tecnologías de virtualización, pero al plantearse las ventajas relativas a la continuidad del negocio de este tipo de implementación, así como respecto de otros tópicos en la administración de sistemas (la creación de plantillas para desarrolladores, copias instantáneas para replicar servidores entre ambientes, etc.), estos costos dejan de ser significativos.

4.2. Tipos de replicación

Existen dos tipos de replicación de datos, cada uno de ellos cuenta con sus propias ventajas y desventajas relativas a tiempos de respuesta, necesidades de conexión y tiempos de restauración, estos son descritos a continuación.

4.2.2. Replicación sincrona

En este método de replicación, los datos se replican en el instante en que son recibidos por el sitio primario. El mecanismo consiste en escribir primero los datos al sitio primario, enviarlos al sitio secundario y hasta que ambos sitios hayan escrito sus datos se continúa con la operatoria normal. Conlleva la desventaja de que si existieran retardos sobre el enlace que comunica ambos sitios, puede verse afectado el rendimiento de aquellas aplicaciones que se ejecutan sobre el primario. La mayor de sus ventajas es que al momento de una emergencia sobre el sitio primario, el sitio secundario será una copia casi 100% fiel de éste.

Este tipo de replicación es recomendable solo en sitios que posean enlaces con altas capacidades de transferencia, en otros casos, puede ser suplantado por replicación asíncrona con algunos segundos de tiempo de replicación, este escenario será útil para la mayoría de implementaciones, aunque existirán casos específicos en los que si será necesaria una exactitud de segundos en los sitios primario/alterno.

4.2.3. Replicación asíncrona

Este tipo de replicación, está basado en replicar en tiempo diferido hacia el sitio secundario. Sin importar el método utilizado para replicar, en este tipo se realizan las escrituras a disco sobre el sitio primario, y luego de un período de tiempo definido, estas escrituras son replicadas al sitio secundario. Entre sus ventajas, se puede mencionar que no se afecta de ninguna manera el rendimiento del sitio primario, puesto que el sistema de replicación simplemente toma todos los cambios sobre las unidades de disco cada cierto tiempo y las envía al sitio secundario. Entre las desventajas, se tiene que al momento de un fallo en el sitio primario, dependiendo del tiempo de replicación definido, el sitio secundario no será una copia 100% fiel al sitio primario.

4.3. Promocion de sitios alternos

El proceso de promoción del sitio alterno dependerá de la implementación realizada. Al momento de elegir un método de replicación, el proceso de promoción del sitio alterno será un factor determinante. El proceso de promoción dependerá primordialmente del tipo, tanto como del método de replicación utilizado, así, si se elige replicación síncrona las operaciones serán restablecidas en el sitio secundario en un periodo de tiempo relativamente muy corto, sin causar atrasos a las aéreas operativas, puesto que el sitio secundario contendrá exactamente los mismos datos que el sitio primario. En el caso de la replicación asíncrona, el período de replicación definido afectará el proceso de promoción, puesto que los datos insertados en el sitio primario después de la última sincronización se perderán, esto implicará desarrollar protocolos de contingencia en las aéreas operativas, con el fin de acoplarse a este proceso.

El costo de esto último puede resultar de alguna manera bastante alto, ya que los procesos operativos tendrán que ser rediseñados en alguna medida para tener control sobre las transacciones ya aplicadas en sus sistemas y las que no han sido aplicadas en la línea de tiempo. Por ejemplo, asúmase un sistema bancario que sufre la siguiente serie de sucesos:

- Se aplican un conjunto de transacciones (conjunto A).
- El sitio primario se replica al sitio secundario (replicación asíncrona).
- Se aplica un nuevo conjunto de transacciones (conjunto B).
- Existe un problema con el sitio primario, es necesario promover el sitio secundario.
- Después de la promoción del sitio secundario el conjunto de transacciones B se ha perdido.

Las aéreas operativas deberán tener conocimiento exacto de cuáles transacciones corresponden a qué momento en el tiempo, para conocer cuáles transacciones deberán aplicar nuevamente.

Se analizará ahora el escenario con replicación síncrona:

- Se aplica una transacción A sobre el sitio primario.
- Se aplica la transacción A sobre el sitio secundario.

- Cuando el sitio secundario informa de la transacción aplicada, el sitio primario continúa con el proceso normal.

- Este proceso se repite para un número “n” de transacciones.

- Se aplica una transacción Y en el sitio primario.

- Se aplica la transacción sobre el sitio secundario.

- El sitio primario deja de funcionar, se activa el sitio secundario la transacción nunca se da por cometida en el ambiente puesto que falló durante su creación.

- El sistema transaccional dará la alerta al usuario informando que la transacción resulto fallida.

- Después de promover el sitio secundario, el usuario final debería conocer exactamente cuál transacción estaba realizando al momento del fallo.

- La transacción única será creada nuevamente por el usuario y escrita sobre el sitio secundario que ahora ocupa el papel de sitio primario.

Las operaciones anteriormente descritas deberán ser consideradas, y documentadas plenamente como parte del plan de continuidad del negocio. En caso de una emergencia, las acciones a tomar para promover el sitio secundario deben estar claramente identificadas, descritas y asignadas al personal adecuado.

La toma de una decisión relativa al proceso de promoción de sitios secundarios deberá ser entonces realizada no sólo por el área de informática, puesto que estos procesos afectarán a toda la empresa, deberá ser una decisión consensuada entre todos los departamentos de la misma.

4.4. Pruebas a los sistemas de replicación

Probar un sistema de replicación, y de alta disponibilidad en general, es un punto tan importante como su implementación misma. Aun cuando el sistema de replicación exista, y teóricamente funcione, este no tendrá validez alguna hasta que se realice al menos una prueba, promoviendo al sitio secundario y realizando operaciones sobre el sitio secundario.

Las mejores prácticas indican la realización de pruebas constantes al sistema de replicación. Los planes de pruebas deberían contemplar la desconexión total del sitio primario, y la realización de todas las operaciones de la empresa sobre el sitio secundario durante algún período de tiempo bien definido. Estas pruebas servirán para actualizar la definición de protocolos de contingencia, los procedimientos involucrados en la promoción de sitios, así como la asignación de roles durante la emergencia.

Lo importante de estas pruebas será que, desde que son pruebas y simulaciones, el sitio primario seguirá en estado normal, aunque desconectado, el protocolo de pruebas habría de definir que si existe algún problema con el sitio secundario durante el período de pruebas, se pueda volver al sitio primario sin mayor problema.

Realizar estas pruebas periódicamente resulta extremadamente importante, existirán casos en los que será indispensable migrar las operaciones al sitio secundario aun cuando no se trate de un desastre sobre el primario. Por ejemplo, el *software* utilizado para replicación muchas veces deberá ser actualizado sitio por sitio. Esto implica que el sitio primario deberá ser apagado durante el proceso de actualización de versión, y debido a que los procesos de actualización podrían ser lentos, o por seguridad para evitar corrupciones en la versión en el sitio primario, se debería trabajar durante este tiempo sobre el sitio secundario.

Los procesos de pruebas, como se dijo antes, deberán incluirse en el calendario de actividades de la empresa. El departamento de tecnología deberá realizarlos periódicamente, auxiliándose por los departamentos involucrados. Es posible que la empresa gestione un plan de continuidad del negocio, lo cual implicará pruebas y dictará los requerimientos para las pruebas de continuidad de los servicios de tecnología, en caso contrario, estas deberán ser definidas claramente por personal de todos los departamentos involucrados, en cuanto a períodos de pruebas, alcance de las mismas y el proceso mismo de ejecución de las pruebas, las asignaciones y responsabilidades durante estos períodos no deberían cambiar mucho respecto de las definidas en el proceso de promoción de sitios alternos, el proceso de promoción en pruebas debería ser lo más parecido a la realidad posible, probando escenarios en donde el sitio primario es detenido de manera planificada, y no planificada desde el punto de vista operativo.

Otra ventaja de las pruebas, es que generan métricas sobre el comportamiento de los sistemas replicados, así, en caso de un incidente sobre el sitio primario, se conocerá con bastante exactitud el tiempo de promoción del sitio alterno, conociendo entonces cuál será el impacto sobre las operaciones de la empresa, lo cual junto al conocimiento de costos implicados por detener las operaciones en un espacio de tiempo determinado, mostrará la efectividad del sistema de replicación para amortiguar las pérdidas derivadas de las caídas de sistemas primarios.

4.5. Copias instantáneas

Antes se mencionó que la replicación no es una garantía frente a desastres en el sitio primario, mientras que no lo es para las corrupciones de datos. Por ejemplo, suponga una arquitectura en donde existe un sitio primario y un sitio secundario de replicación. En el caso de que los datos en el sitio primario se corrompan por un fallo en aplicaciones u otra razón. Los datos en el sitio secundario se corromperán también. Esto se puede ver más claramente al analizar el siguiente flujo de ejemplo:

- Tiempo t1, se genera un conjunto de transacciones A en el sitio primario.
- Tiempo t2, el conjunto de transacciones A se replica al sitio secundario.
- Tiempo t3, se genera un nuevo conjunto B de instrucciones.
- Tiempo t4, falla una transacción del conjunto B, existe un error en una aplicación, los datos se corrompen en el sitio primario.

- Tiempo T5 (dependiente del tipo de replicación), se replica las transacciones al sitio secundario.

- Tiempo T6, el sitio secundario resulta corrupto al igual que el primario.

En estos casos, y como se puede deducir del flujo anterior, la replicación no sirve como método de contingencia. Puesto que los errores se propagan al sitio secundario. Es por ello, que un sistema de replicación habría de acompañarse de un sistema de copias instantáneas. El sistema de copias instantáneas provee de puntos de restauración para el estado de los sistemas involucrados.

Por ejemplo, si se tiene un sistema de copias instantáneas, este generará una copia cada determinado período de tiempo y la almacenará en un servidor externo. Para ejemplificar esto, se muestra el flujo de acciones en un sistema con replicación y copias instantáneas:

- Tiempo T1, se genera un conjunto de transacciones A en el sitio primario.

- Tiempo T2, el conjunto de transacciones A se replica al sitio secundario.

- Tiempo T3, se genera un nuevo conjunto B de instrucciones.

- Tiempo T4, se genera una copia instantánea del sitio primario o secundario, se envía a un servidor externo.

- Tiempo T5, falla una transacción del conjunto B, existe un error en una aplicación, los datos se corrompen en el sitio primario.

- Tiempo T6 (dependiente del tipo de replicación), se replica las transacciones al sitio secundario.

- Tiempo T7 el sitio secundario resulta corrupto al igual que el primario.

- Tiempo T8 es posible en este punto restaurar la copia instantánea generada en el tiempo T4.

Como se puede apreciar entonces, las ventajas de implementar un sistema de copias instantáneas sumado a la replicación, son muy notorias. Probablemente el mayor inconveniente o desventaja en esta implementación será el espacio utilizado para almacenar estas copias, Por ejemplo, asumiendo un sistema que trabaje con máquinas virtuales, y replicación de las mismas como se explicó anteriormente, el tamaño mínimo para estas máquinas, incluyendo una plantilla muy básica compuesta de sistema operativo, servidor de BD, servidor de http, etc. Será de aproximadamente 40 GB sin tomar en cuenta el espacio destinado almacenar datos de aplicaciones.

Las copias instantáneas se generarán en lapsos de tiempo determinados primordialmente por las necesidades del negocio en que se realice la implementación. Es común pensar en períodos de una hora, entre copias, pero todo dependerá primordialmente de los costos operativos de la empresa, la relación de costos habría de ser:

$$A<B$$

A: Costo de implementación de sistema para X horas

B: Costos implicados por retrasar X horas las operaciones

A partir de este límite de costos, deberá analizarse la periodicidad con que las copias instantáneas son tomadas. La variable con mayor dominio de valores entre las diferentes implementaciones de estos sistemas, será el espacio de disco utilizado, antes se dijo que un tamaño promedio de un archivo de máquina virtual sería de aproximadamente 40 GB y si se realizan copias instantáneas cada media hora, se requerirá un total de 1920 GB disponibles para almacenar estas copias si se pretende almacenarlas por un periodo de 24 horas desde su creación. Se debe tomar en cuenta que este cálculo corresponde a un único servidor, con almacenamiento muy reducido (40 GB), y en la mayoría de casos, se implementara este tipo de sistemas para más de un servidor. Y probablemente los costos varíen mucho respecto de los ejemplificados acá, debido a que muchos servidores necesitaran almacenar grandes cantidades de datos.

Otra variante de esta implementación, será la elección del origen de las copias instantáneas, estas podrán ser tomadas desde el sitio primario o desde el sitio secundario. El hecho de tomarlas desde el sitio primario, tiene la ventaja de que la última variante que se encontrará en este tipo de implementación, es el sitio donde se almacenan las copias instantáneas. Existen tres alternativas para almacenarlas:

- Sitio primario
- Sitio secundario
- Ambos

Lo ideal sería almacenar las copias en el sitio primario e incluirlas en el sitio de replicación, es la alternativa más segura, pero también la más costosa.

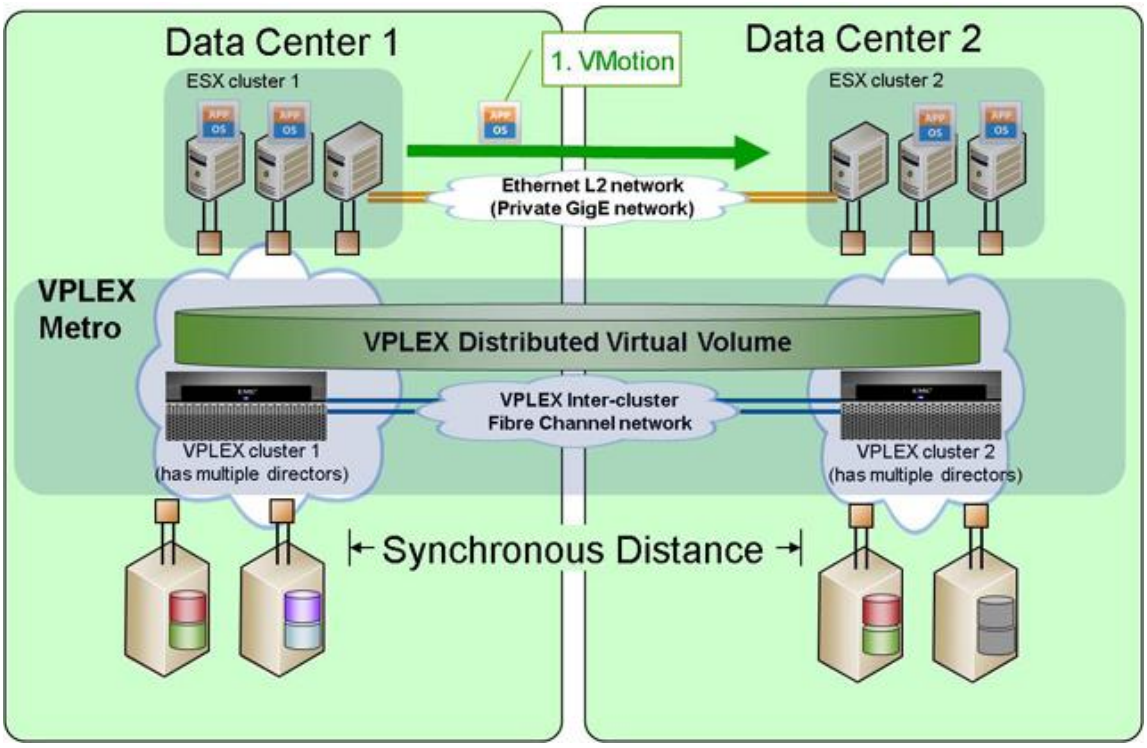
4.6. Replicación en ambientes activo/activo

Hasta aquí se han asumido sitios de replicación que trabajan de manera activo/pasivo. Es decir, el sitio primario se encuentra habilitado y operando, mientras el sitio secundario se encuentra detenido, a la espera de una falla sobre el sitio primario. Esto implica tener recursos sin utilizar en todo el sitio secundario. Una configuración activo/activo, persigue aprovechar al máximo todos los recursos utilizados en la implementación del sistema de replicación, utilizando sistemas de balanceo de transacciones entre ambos sitios. Así, el usuario y los sistemas que utilizan estos recursos, ven a todos los sitios de replicación como uno solo, debido a esto, se conectan a una dirección de red única, y los sistemas de balanceo de transacciones se encargan de distribuir cada una de las peticiones a cualquiera de los sitios, siguiendo algún algoritmo determinado.

Entre las muchas ventajas de este tipo de implementación, se tiene el incremento en el rendimiento de los sistemas en general, puesto que las operaciones de consulta y escritura pueden realizarse en cualquiera de los sitios involucrados, así, no se sobrecarga únicamente los recursos de un sitio.

La figura 3 muestra la implementación de esta tecnología mediante VPLEX de EMC. Puede apreciarse como ambos centros de datos tienen acceso simultaneo a una unidad de almacenamiento virtual distribuida administrada por la tecnología VPLEX. En esta implementación de ejemplo, se utilizan servidores ESX como plataformas para ejecutar maquinas virtuales que se almacenan y replican en este volumen virtual.

Figura 3. **Tecnología EMC-Vplex**



Fuente: www.vmware.com. Consulta: diciembre de 2010

Con la arquitectura mostrada en la figura 3, el desperdicio de recursos de los ambientes activo/pasivo desaparece. Los servidores ESX de cada centro de datos ejecutan sus propios sistemas, y escriben a una unidad virtual, que es percibida como un único disco, el sistema vplex distribuye la carga de lecturas/escrituras entre ambos sitios, y se encarga de replicar la información entre los sitios, con el fin de tener copias idénticas de los datos en ambos sitios.

5. ALTA DISPONIBILIDAD DE BASES DE DATOS

La gran mayoría de los sistemas almacenan sus configuraciones, estado, transacciones, etc. en bases de datos. Es lógico entonces pensar que la replicación de bases de datos resulta de gran ayuda cuando se implementa sistemas para continuidad del negocio. En el capítulo anterior se habló de replicación de sistemas en general, se debe tener claro que la replicación de sistemas, replica absolutamente todo lo que ocurre en el sitio primario. Así, si en el sitio primario se pierde la integridad de los datos, esto sucederá también en el sitio secundario.

Se propuso para subsanar esto la implementación de sistemas de toma de copias instantáneas, las cuales podrían ser restauradas en cualquier momento y contienen toda la información del sistema, básicamente, como se planteó en el capítulo anterior, la copia instantánea será una copia íntegra del servidor que está ejecutando las aplicaciones. Pero es posible que al momento de una corrupción de datos, no sea absolutamente necesario restaurar una copia instantánea del sistema entero. Es posible que el servidor ejecute muchos sistemas, y la corrupción de datos afecte únicamente a uno de estos sistemas. En estos casos, es lógico que se desee restaurar únicamente los datos de los sistemas corruptos, esto se consigue restaurando la base de datos específica de ese sistema, lo cual hace menos costoso el proceso de promoción y reduce riesgos innecesarios.

5.1. Alta disponibilidad de bases de datos

Implementar sistemas de alta disponibilidad para bases de datos, puede resultar tan beneficioso como difícil y costoso. Las principales ventajas que existen en estos sistemas, será el decremento en el tiempo de restauración de las operaciones luego de algún incidente, y el aislamiento del problema, reduciéndolo únicamente a los sistemas involucrados.

Para comprender esto, puede pensarse en la replicación de centros de datos (síncrona o asíncrona), si se asume una falla sobre el servidor que almacena la base de datos, y este es un servidor virtual dentro de un servidor ESX por ejemplo, será necesario promover el sitio secundario en su totalidad para rehabilitar las operaciones. Esto implica detener todos los sistemas que se ejecutan en el sitio primario, aun cuando es posible que no todos ellos presenten problemas. Al implementar soluciones de alta disponibilidad de bases de datos, se abre la posibilidad de restaurar únicamente la base de datos con problemas, sin afectar a los otros sistemas.

Por otra parte, es posible que una empresa pequeña decida no implementar tecnologías de replicación, en este caso, un mínimo de seguridad a sus datos y operaciones sería proveído por la alta disponibilidad en sus bases de datos. Una de las grandes ventajas en estos casos, será que existen empresas en el medio nacional que se dedican a proveer almacenamiento físico para servidores, omitiendo la necesidad de crear un sitio de datos alterno, lo cual reducirá por mucho los costos de implementación.

5.2. Métodos de alta disponibilidad

Existen diversos métodos para asegurar la alta disponibilidad de las bases de datos, los principales son descritos a continuación.

5.2.1. Copias de espejo

Una copia de espejo es una copia completa de la base de datos, que existe en un servidor alterno, esta copia puede ser generada mediante cualquier método, con la restricción de que debe mantener su integridad, es decir solamente se escriben transacciones completadas en la base de datos primaria. El proceso de restauración de esta base de datos secundaria resultará muy sencillo en una configuración activo/pasivo (la que se asume para este método), puesto que consistirá simplemente en re direccionar las aplicaciones hacia esta nueva base de datos.

5.2.2. Replicación mediante archivos de bitácora

La replicación mediante la escritura de la bitácora puede resumirse a las siguientes operaciones:

- La base de datos primaria tiene una lista de transacciones en su lista de sucios.

- La base de datos primaria escribe las transacciones contenidas en su lista de sucios, y almacena registro de estas en su bitácora.

- La base de datos primaria envía su bitácora a la base de datos secundaria.

- La base de datos secundaria realiza un proceso de *roll forward* basándose en la bitácora recibida, con lo cual consigue una copia idéntica de los datos en la base de datos primaria hasta su última escritura a bitácora.

- Si existiese una falla sobre la base de datos primaria, la secundaria contendrá una copia casi 100% fiel, sin contar con las transacciones que aún no se habían terminado al momento del fallo sobre el servidor primario.

5.2.3. Ambientes distribuidos como medios de alta disponibilidad

Otra de las opciones al implementar alta disponibilidad sobre bases de datos, es utilizar tecnologías de bases de datos distribuidas, las cuales podrán ser configuradas para replicar segmentos de bases de datos únicamente, por ejemplo, es posible que bajo ciertas circunstancias únicamente se necesite alta disponibilidad sobre ciertas bases de datos, o incluso sobre ciertas tablas, en este caso, puede implementarse bases de datos distribuidas que existan en modo “todo compartido” en el ambiente.

Las bases de datos distribuidas pueden existir en puntos geográficamente distantes, lo cual aumenta las posibilidades de recuperación, otra de las grandes ventajas, es que la pérdida de un servidor resulta casi transparente para el usuario final. Puesto que la base de datos distribuida actúa como una única base de datos. También se puede mencionar entre sus ventajas que la integridad de los datos se asegura mediante un procedimiento conocido como “commit de dos fases” el cual distribuye cada transacción entre todo el ambiente, y la da por cometida hasta el momento en que todos los servidores la han escrito y dado por finalizada. Esto se comprende mejor con un ejemplo:

- Existe una base de datos distribuida entre los servidores A, B, C.

- Al escribirse una transacción sobre el nodo A, este envía la petición de escritura de esa transacción a los nodos B y C.

- Los nodos B y C escriben su transacción, y envían la respuesta al orquestador del ambiente.

- El orquestador avisa al nodo A para que dé por terminada la transacción.

- Si el proceso de escritura falla por alguna razón sobre el nodo B o C, la transacción es interrumpida y cancelada en todo el ambiente, con lo cual se conserva la integridad de datos sobre todo el ambiente.

5.2.4. Bases de datos a la espera

Es una arquitectura en la cual se tiene una copia íntegra de la base de datos independientemente del método de replicación, como en casi todas las tecnologías para alta disponibilidad de bases de datos, una de las mayores ventajas es que mantiene la integridad de los datos, aunque esto dependerá del método de replicación seleccionado. La ventaja de tener una base de datos a la espera, radica en que el cambio de un servidor por otro será realmente sencillo, no es necesario más que apagar el servidor que actualmente se encuentra configurado como primario, y cambiar la configuración para que el secundario pase a ocupar el lugar de esté.

La exactitud de la copia dependerá del tiempo que se configure para replicación de los datos, pero normalmente este tiempo no habrá de ser demasiado grande. El servidor secundario puede ser utilizado para generación de reportes, en los que no se requiera una extrema exactitud en el tiempo. Con ello, se logra aprovechar los recursos involucrados en el servidor secundario, reduciendo los costos que se desaprovechan y aumentando los beneficios de la implementación.

5.2.5. Copias instantáneas

Como en el caso de la implementación de copias instantáneas en replicación, las copias instantáneas de bases de datos ayudan a asegurar la integridad de los datos en el tiempo. Las copias de espejo, ambientes distribuidos, etc. están sujetos a errores humanos o corrupción por parte de las aplicaciones. Al implementar un sistema de copias instantáneas se tiene la certeza de tener una copia del estado de los datos en distintos puntos en el tiempo.

Los sistemas de copias instantáneas de bases de datos tienen básicamente las mismas características que las copias instantáneas de sistemas replicados que se mencionaron anteriormente. Debe elegirse cuidadosamente el sitio de almacenamiento de las copias generadas, tomando en cuenta varios aspectos:

- El sitio donde se almacenaran las copias.

- La necesidad de replicación de estos datos.

- Las capacidades de los enlaces utilizados hacia el sitio donde se almacenan las copias.

- El tiempo que transcurre entre copias.

- Los procedimientos de restauración.

Como en todas las implementaciones para alta disponibilidad en este caso, es sumamente importante contar con una definición exacta del proceso de restauración, en los casos de bases de datos en específico, es necesario contemplar que un proceso de restauración necesitará ejecutarse sobre un servidor con una serie de herramientas previamente instaladas. Esto se simplificará mucho cuando se trabaja sobre ambientes virtualizados, en los cuales la creación de un servidor, puede basarse en una plantilla previamente construida. Estas plantillas de restauración pueden ser utilizadas en muchos casos en implementaciones de alta disponibilidad, y serán tratadas con más detalle en el capítulo seis.

5.3. Oracle RAC

En el caso de implementaciones con sistemas de bases de datos Oracle, una buena alternativa es Oracle RAC (de sus siglas en inglés *Real Application Cluster*). Esta tecnología provee un ambiente de alta disponibilidad en el que N nodos con su propia instancia de la base de datos. Todos estos nodos, representan una única instancia, y comparten una SAN como medio de almacenamiento. Los nodos se comunican entre sí mediante enlaces de alta velocidad.

Un ambiente RAC es percibido como una única instancia de base de datos para cualquier otro sistema que interactúe con él. Las instancias en los nodos se intercomunican con el fin de distribuirse las transacciones que son recibidas. Estas transacciones son tomadas y ejecutadas en su totalidad por uno de los nodos, el cual escribe los resultados de las mismas en la SAN, los cuales en adelante pueden ser consultados por cualquiera de los otros nodos. Si en medio de una transacción el nodo sufre de algún fallo, sus transacciones pendientes son operadas por otro de los nodos, lo cual asegura la alta disponibilidad de las instancias del DBMS. Los datos son almacenados como se dijo antes en SAN's lo cual asegura su disponibilidad e integridad.

De las muchas ventajas de RAC, se tiene también que aumenta el rendimiento del sistema general, simplifica el crecimiento horizontal, hace transparentes los fallos en cualquier componente de la arquitectura, minimiza las pérdidas del negocio por tiempos inoperativos durante periodos de restauración, puesto que el sistema no se detendrá a menos que exista una cadena de fallos muy improbable (habría de fallar todos los nodos, y la SAN en el sitio primario). Así mismo, con esta tecnología se facilitan las pruebas ya que en cualquier momento en que el sistema no tenga demasiada carga, puede simplemente desconectarse un nodo, y cumplir con el procedimiento que se defina para estos casos. La recuperación luego de esto o de una caída real resultará también sumamente sencilla, puesto que todas las tareas a realizar se reducirán a configurar el nuevo nodo, instalando su instancia de Oracle, y conectarlo a la red de nodos.

5.4. Análisis de caso Navicat

Navicat es una herramienta desarrollada por PremiumSoft, que cuenta, entre otras funcionalidades, con las opciones de sincronizar bases de datos, y tomar copias instantáneas programadas en el tiempo. Básicamente, la implementación que se realizó en las pruebas consiste en sincronizar una base de datos secundaria respecto de la base de datos primaria, inmediatamente al acabar la sincronización de datos, se toma una copia de la base de datos recién sincronizada y se almacena en un servidor alternativo. Todas las pruebas fueron realizadas en ambientes Windows, sobre bases de datos MYSQL.

En el primer escenario de pruebas consiste en un servidor de replicación externo, un servidor local sobre el cual corre Navicat y la base de datos primaria, se insertan filas masivamente sobre el servidor primario y Navicat se encarga de sincronizar al servidor secundario.

En promedio se replica 12310 filas por cada 5 minutos.

Los servidores no presentan problemas de rendimiento perceptibles.

En promedio, después de 5 muestras de tiempos, se tiene que la inserción de una misma cantidad de datos a una base de datos en este escenario con la debida replicación al servidor secundario, tarda:

Sin replicación:

40.3 minutos

Con replicación:

38.5 minutos

Puede verse que los tiempos de inserción varían unos pocos minutos entre sí, esto, pero esta variación es despreciable debido a que no se realizaron estas pruebas sobre servidores dedicados, si no computadoras personales que ejecutan al momento otras tareas.

En el segundo escenario de pruebas se replicó la base de datos de un servidor primario hacia un servidor de pruebas, el software de Navicat fue ejecutado en un tercer nodo, aunque según las pruebas anteriores esto no genera carga representativa si se desea ejecutar Navicat sobre el servidor de replicación. Se replicó la totalidad de las tablas cada 15 minutos, y se tomó una copia instantánea de la base de datos cada hora.

El rendimiento de las aplicaciones ligadas a esta base de datos no se vio afectado desde el punto de vista del usuario final en el tiempo que se realizó la replicación de la base de datos.

Tabla III. **Tiempos de replicación de datos (en minutos)**:

T1:	4.47
T2:	3.16
T3:	3.42
T4:	1.51
T5:	1.5
T6:	2.53
T7:	2.35
T8:	02.02

Fuente: elaboración propia.

En promedio, la sincronización tarda:
2.62 minutos

En conclusión, el análisis de rendimiento no presenta ningún inconveniente en cuanto a los parámetros principales de rendimiento del servidor, siendo que se estudiaron:

- Disco duro
- Memoria principal
- Procesador
- Puertos TCP

Se llega a la conclusión que la implementación de este ambiente no conlleva riesgos significativos al rendimiento del servidor primario.

6. CASO DE ESTUDIO: IMPLEMENTACIÓN DE REPLICACIÓN CON TECNOLOGÍAS EMC

Durante el desarrollo de este trabajo, se realizó un estudio en la empresa Imágenes Computarizadas de Guatemala, misma que implementa un sistema de replicación de datos basándose en tecnologías EMC. El objetivo de este capítulo, es mostrar lo que se ha mostrado en los capítulos anteriores, pero desde un punto de vista más práctico, como referencia para implementaciones que se deseen hacer en el medio nacional, se hace un repaso de los costos aproximados de estas implementaciones, así como los proveedores de servicios que existen en el país.

Se muestra acá también una serie de métricas de tiempos tomadas de los procesos de restauración, implementación y todos los procesos involucrados en las pruebas realizadas por la empresa. Los tiempos de respuesta, los inconvenientes encontrados en las pruebas iniciales, mejores prácticas basadas en lo observado en los procesos de implementación, durante el desarrollo de las pruebas de funcionalidad y la resolución de los inconvenientes encontrados.

Aunque el caso se orienta a una implementación con herramientas específicas, la mayoría de las recomendaciones pueden atenderse en implementaciones con otras herramientas, puesto que el desarrollo de planes de pruebas, planes de implementación, etc. Será similar en muchos casos.

6.1. Descripción general de la implementación

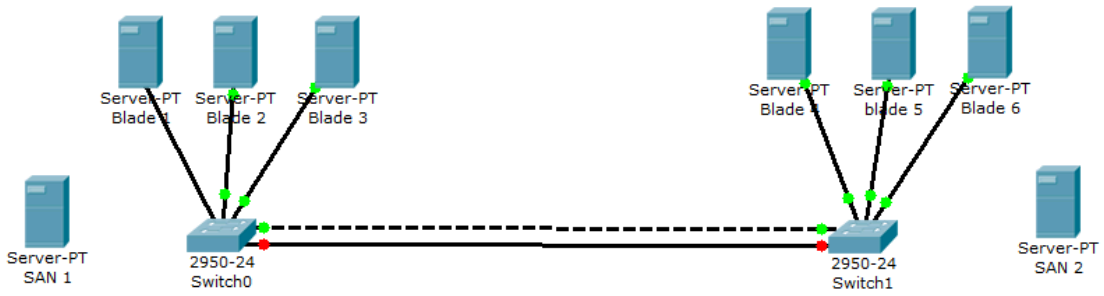
El caso estudiado, como ya se dijo, corresponde a imágenes computarizadas de Guatemala, empresa que implementa un sitio primario en el cual realiza todas sus operaciones y un sitio secundario en una locación distinta, el cual es replicado utilizando tecnologías de EMC. El sitio primario y secundario se encuentran comunicados por dos enlaces que crean redundancia. La replicación que se aplica acá, es asíncrona, el estudio fue realizado durante pruebas programadas por el personal de tecnología, las cuales consistieron en desactivar durante una semana el sitio primario, y activar el sitio secundario para operar en él.

La empresa cuenta con una arquitectura de servidores blade, los cuales ejecutan una serie de máquinas virtuales que a la vez ejecutan todos los servicios requeridos. Estas máquinas virtuales se almacenan en sistemas SAN. Las cuales son replicadas a nivel de LUN hacia el sitio secundario que cuenta con la misma arquitectura. Cabe recordar que las principales ventajas de este tipo de replicación serán al momento de activar el sitio secundario, puesto que los servidores son máquinas virtuales, y éstas serán apagadas al tener un problema en el sitio primario, evitando problemas de direccionamiento o reconfiguraciones necesarias.

Uno de los principales puntos en el diseño de la solución en este y en todos los casos posibles, será el diseño de las pruebas a la implementación, en el caso de estudio las operaciones de la empresa son 24/7, aun así, existe un periodo de tiempo diariamente de aproximadamente 3 horas en las que no se realizan procesos críticos, o al menos pueden ser detenidos los servicios de tecnología sin que el área operativa se vea demasiado afectada.

La figura 4 muestra un esquema básico de cómo están diseñados e interconectados ambos sitios de datos:

Figura 4. **Diseño de sitio de replicación**



Fuente: elaboración propia.

En la figura 5, los *blade* 1, 2,3 pertenecen al sitio primario y los blade 4, 5,6 pertenecen al sitio secundario. Todas las máquinas virtuales ejecutadas por los blade existentes en el sitio primario se almacenan en la SAN 1, la cual es replicada a nivel de LUN hacia la SAN 2 ubicada en el sitio secundario.

Toda la gestión de las máquinas virtuales se realiza con Virtual Center de VmWare. Esto facilita por mucho la gestión de servidores, puesto que ya se cuenta con plantillas de todos los servidores más importantes, mediante las cuales es posible crear un servidor en tiempos muy reducidos.

Para el proceso de restauración del sitio secundario, las ventajas antes explicadas se entienden mejor pensando en la figura 5. Si se tiene la dirección IP A.B.C.D en la máquina virtual 1 que se ejecuta sobre el blade 1 en el sitio primario, la misma máquina virtual está siendo replicada a la SAN 2. Debido a que la máquina virtual es básicamente un archivo binario que incluye todas las configuraciones de la misma, la máquina virtual en el blade del sitio secundario tiene también la IP A.B.C.D, lo cual no crea duplicidad puesto que las máquinas virtuales del sitio secundario se encuentran apagadas.

Al momento de activar el sitio secundario, se asume que el sitio primario no existe por alguna razón, con lo cual todas las máquinas virtuales en el sitio secundario serán encendidas, tomando para si las direcciones IP de las máquinas virtuales en el sitio primario, y evitando cualquier necesidad de reconfiguración de las aplicaciones cliente.

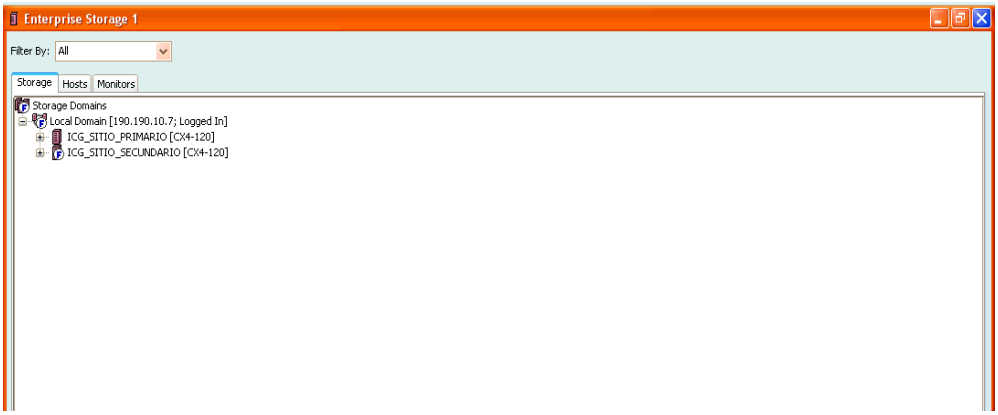
El proceso de activación del sitio secundario se comprende de una serie de pasos:

- Detención del sitio primario
- Promover el sitio secundario
- Iniciar todas las máquinas virtuales en el sitio secundario

Todos estos procesos en el caso de las tecnologías EMC son administrados desde la herramienta Navisphere, cuya interfaz presenta todas las opciones necesarias para configurar los sitios de replicación, así como realizar las operaciones necesarias relativas a la replicación, promoción, etc.

La figura 5 muestra la interfaz de Navisphere donde puede apreciarse los sitios de replicación existentes, puede verse a grandes rasgos que Navisphere reconoce todo el ambiente como “dominios de almacenamiento” puesto que en realidad lo que se replica en este caso es el almacenamiento de cada sitio, en este caso los sistemas SAN.

Figura 5. **Interfaz de NAVISPHERE**



Fuente: elaboración propia.

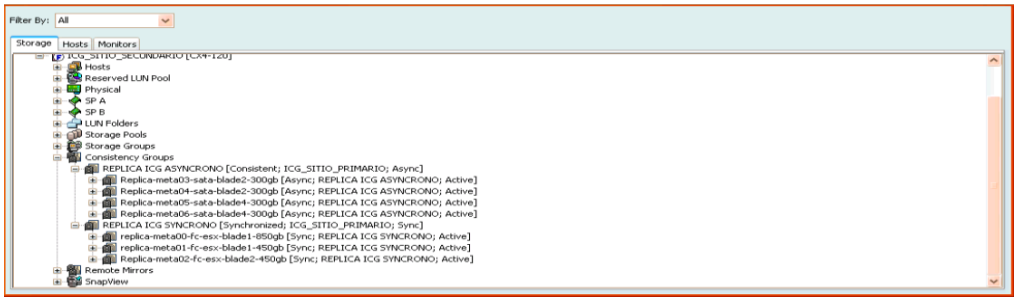
Es importante mencionar también que las herramientas de EMC crean grupos consistentes de LUN’s. Un grupo de consistencia, no es más que un grupo de LUN que se encuentran exactamente iguales en ambos sitios. Básicamente, podríamos definir esto como puntos de consistencia en el tiempo, asúmase que durante la replicación de la SAN 1 a la SAN 2, se han replicado N LUN completos, siendo M el total de LUN en ambos sitios. Si durante el proceso de replicación, el sitio primario falla, probablemente algunos de los LUN que se estaban replicando en ese momento preciso quedaran corruptos en el sitio secundario. Es por esto que el sistema deberá contar con un registro del estado de aquellos LUN que han sido replicados con éxito para devolver todos los LUN hacia un punto de consistencia, estos LUN que han sido replicados con éxito pertenecen entonces a un grupo de consistencia.

6.2. Proceso de promoción del sitio alterno

El proceso de promoción del sitio alterno resulta uno de los procesos más críticos en los sitios replicados tanto en las situaciones reales como en las situaciones propiciadas para pruebas. Básicamente todas las pruebas habrían de centrarse en esto específicamente, ya que si se tiene una implementación de replicación sin realizar pruebas de promoción del sitio alterno de manera periódica, toda la implementación perderá validez en gran medida. Es recomendable que la empresa que implemente replicación síncrona o asíncrona planifique y documente claramente su procedimiento de pruebas, y este deberá incluir necesariamente la promoción del sitio alterno y de ser posible ejecutar ciclos operativos sobre este. En el caso de estudio en específico, las pruebas se han planificado de manera periódica, y consisten en detener por completo el sitio primario y luego promover el sitio secundario.

Es necesario documentar el proceso de promoción del sitio secundario en tanto a tiempos de respuesta, procedimiento explícito, asignaciones y responsabilidades dentro del departamento de tecnología. A continuación se detallan los pasos realizados para promover el sitio secundario.

Figura 6**. Vista general de los grupos de consistencia**

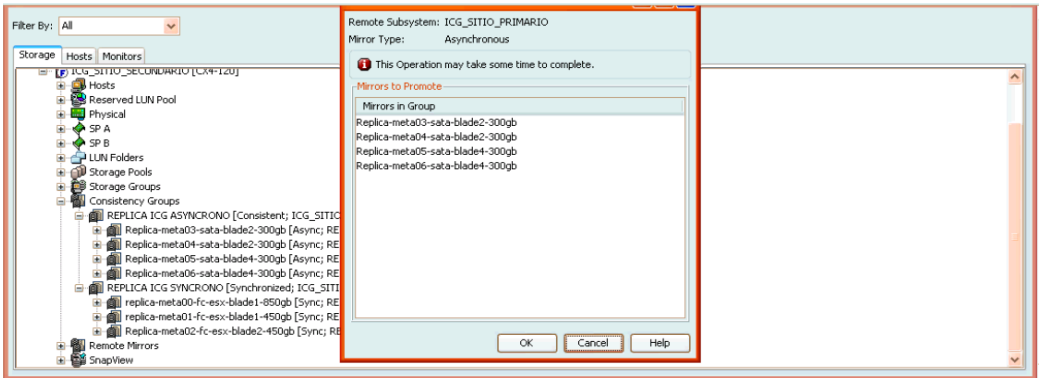


Fuente: elaboración propia.

La figura 6 muestra una vista general de los grupos de consistencia en el sitio secundario, puede verse que existen dos clasificaciones para estos grupos de consistencia: síncronos y asíncronos.

La figura 7 muestra el resultado de ejecutar la operación de promoción al sitio secundario, se alerta sobre el tiempo que la operación puede requerir, así como los *blade* que están involucrados.

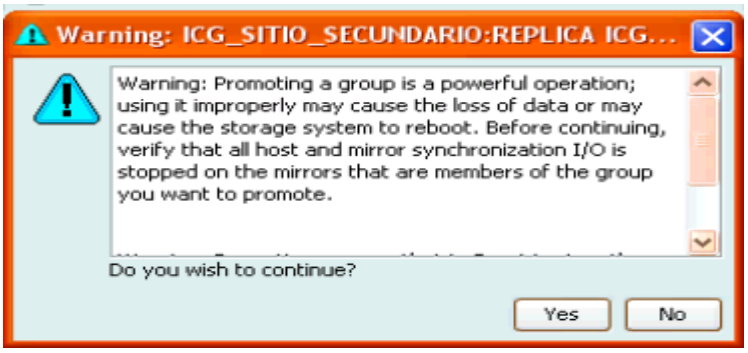
Figura 7**.** **Alertas en la operación de promoción**



Fuente: elaboración propia.

Seguido a esto, se presenta una advertencia de seguridad, indicando la importancia del proceso de promoción del grupo de consistencia alterno, indicando que es una operación riesgosa, y como tal deberá ser ejecutada con pleno conocimiento de las acciones tomadas. También solicita que se verifique que todas las operaciones de entrada y salida a disco en ambos sitios han sido detenidas previamente, debido a ser este un escenario de pruebas, todas las máquinas virtuales fueron apagadas antes de iniciar el proceso. En un caso real, será necesario asegurarse de que no existe ninguna operación de entrada/salida sobre el sitio alterno.

Figura 8. **Alerta de seguridad en promoción**



Fuente: elaboración propia.

Inmediatamente, se alerta al usuario de que al promover el sitio secundario, el sitio primario perderá la sincronización. Es necesario promover el sitio secundario de manera local únicamente para evitar esto, o en caso contrario forzar la promoción.

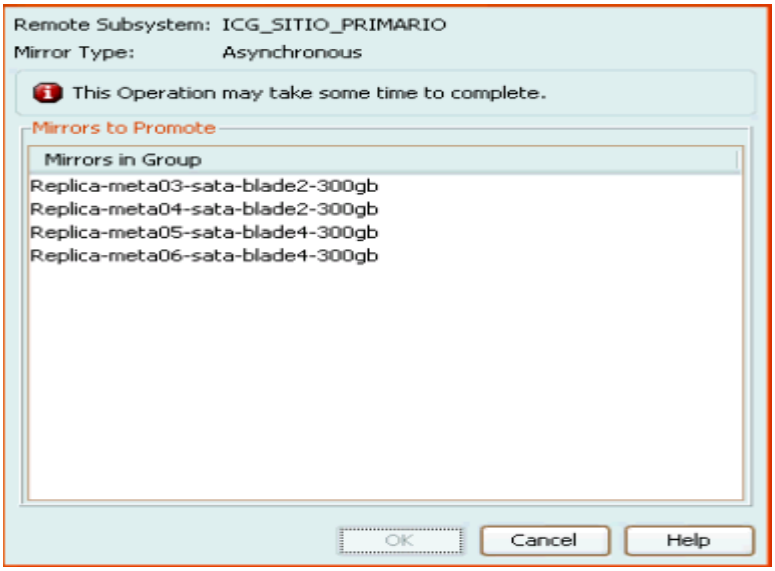
En el caso de las pruebas, se realiza una promoción local, puesto que al término del período de pruebas, es necesario volver al sitio primario actual, y este debe permanecer sincronizado.

Figura 9. **Selección del tipo de promoción**

Fuente: elaboración propia.

La figura 10 muestra el proceso de promoción del sitio, aun cuando la herramienta muestra la alerta que es posible que esta operación tarde mucho tiempo, la operación se realiza en un tiempo de aproximadamente 10 segundos. Se debe tomar en cuenta que se promovió en este paso al grupo de consistencia asíncrono, como puede verse en la figura 10, está compuesto de 4 espejos de replicación, de 300 GB cada uno.

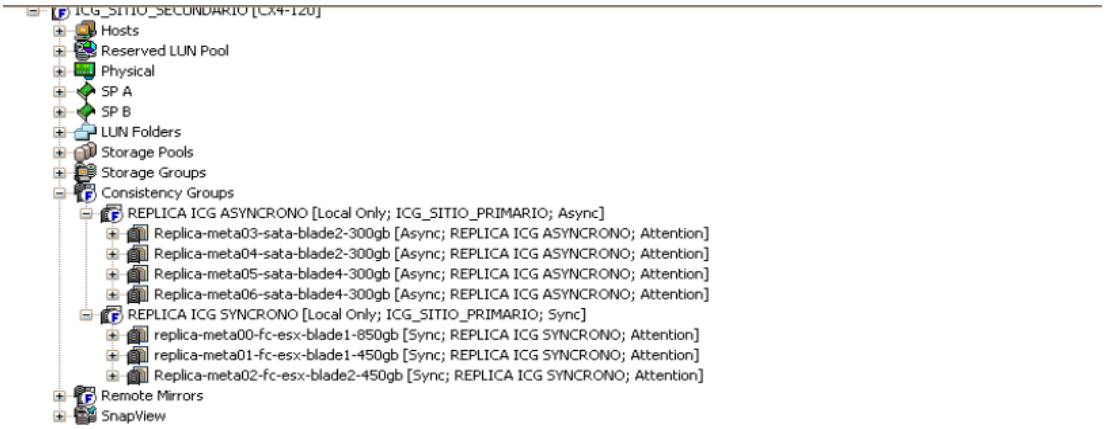
Figura 10. **Progreso de la promoción**



Fuente: elaboración propia.

Debido a la existencia de replicación síncrona y asíncrona en la implementación, es necesario realizar la misma operación de promoción para el grupo de consistencia síncrono. Luego de estas operaciones, ambos grupos de consistencia son mostrados en Navisphere como grupos únicamente locales, puesto que fue la opción elegida.

Figura 11**. Vista general de Navisphere después de la promoción**



Fuente: elaboración propia.

En esta ejecución de pruebas, se enfrentó la necesidad de anexar las LUN en el sitio alterno a los grupos de almacenamiento previamente creados, esto es necesario únicamente la primera vez que se promueve un sitio. Como podrá verse más adelante, en el proceso de vuelta hacia el sitio primario original, este proceso no fue necesario.

Durante el proceso de migración al sitio de replicación alterno, pudo observarse un tiempo de 3 horas aproximadas en la primera migración, este tiempo se vio aumentado primordialmente debido a que no se tomó en cuenta la necesidad de anexar los grupos de almacenamiento a los servidores ESX existentes en el sitio alterno, durante el segundo proceso de migración, consistente en devolver la prioridad al sitio primario original, el tiempo de promoción se redujo a aproximadamente 25 minutos.

De aquí se puede deducir que un sitio de replicación previamente probado y bien configurado puede ser promovido en un promedio de 30 minutos a una hora. En este punto cada empresa debería estudiar el costo de detener sus operaciones durante este periodo de tiempo, contra el costo de implementar un sistema de replicación. Se debe recordar que el sistema de replicación asegura los datos frente a situaciones muy específicas de las que se debe excluir la integridad por errores humanos o de sistemas, mismos que son cubiertos por sistemas de copias instantáneas, los cuales también implementa la empresa de este caso de estudio, misma donde existe un sistema de creación de copias instantáneas de bases de datos, similar al registrado en el capítulo cinco, constituido por copias instantáneas de las bases de datos primarias, seguidas por el proceso de copia de estas copias a un servidor externo, lo cual se realiza periódicamente.

6.3. Valores agregados

Existe una serie de valores agregados en este tipo de implementaciones, entre los cuales habríamos de mencionar:

La escalabilidad horizontal de la solución

El aumento del rendimiento de los servidores que se ejecutan sobre máquinas virtuales resulta muy sencillo.

Es posible con las tecnologías de EMC llegar a trabajar con los sitios de manera activo/activo, para evitar desperdicios de recursos.

El tiempo de respuesta presentado por los servidores del sitio alterno, resultan ser idénticos a los prestados por el sitio primario, desde que la arquitectura es exactamente la misma.

Si se cuenta con replicación activo/pasivo, puede estudiarse detenidamente la asignación de recursos hacia el sitio secundario, por ejemplo, asumiendo que en el sitio primario se tiene dos cortafuegos para maximizar los posibles puntos de falla, en el sitio secundario no será necesario tener también dos cortafuegos, puesto que el sitio secundario operará solamente bajo ciertas circunstancias muy específicos y normalmente de baja probabilidad. Entonces, es posible que el caso de replicación activo/pasivo cree ahorro de recursos en alguna medida, aunque la configuración activo/activo ofrece menos desperdicio de recursos, pero también será importante evaluar la necesidad de recursos para la operación de los servidores.

6.4. Ventajas de la virtualización

Durante el desarrollo de las pruebas realizadas, pudo comprobarse los muchos beneficios que trae la virtualización cuando se habla de continuidad del negocio. Imagine por ejemplo, una empresa que considera crítico su servicio de correo electrónico, si este servidor por alguna razón dejase de funcionar, el personal debería activar su protocolo de promoción del sitio secundario etc.

Si se tiene un entorno virtualizado, la recuperación de un servidor que simplemente presta un servicio (no almacena datos), puede realizarse simplemente creando una máquina virtual nueva a partir de una plantilla preexistente.

Otra de las ventajas notadas en las pruebas, es el hecho de tener más control sobre los servidores, por ejemplo, es posible crear, destruir y manipular en general los servidores, pudiendo agregarles recursos sin necesidad de desarmarlos físicamente y dejarlos fuera de línea mucho tiempo.

CONCLUSIONES

1. La virtualización de sistemas es una tecnología que puede ayudar a alcanzar la alta disponibilidad de los sistemas informáticos, teniendo entre ellas:

   - Replicación de sistemas síncrona y asíncrona, con capacidad de definir diversos de replicación.
   - Creación de planes de contingencia frente a incidentes que afecten directamente a los servidores, haciendo posible que estos sean creados mediante plantillas de máquinas virtuales.
   - Creación de copias de espejo periódicas a servidores críticos.

2. Las soluciones tecnológicas para continuidad de los servicios de tecnología, resultaran muchas veces complementarias entre ellas, siendo que cada una de ellas cumple una característica de disponibilidad, de la siguiente forma:

   - Replicación de centros de datos que cubre desastres y fallos de *hardware* en el sitio primario.
   - Redundancia de *hardware* que cubre fallos de *hardware.*
   - Redundancia de software que complementa a la redundancia de *hardware*.
   - Copias de espejo de sistemas que cubre fallos de *software* o usuario que afecten la integridad de los datos.

- Replicación de bases de datos que puede cubrir fallos de hardware (nodos de *RAC* o cluster), y fallos de *software* que puedan afectar la integridad de los datos.

RECOMENDACIONES

1. La creación de planes de continuidad del negocio desde un punto de vista integral, debe incluir al departamento de tecnología, para que las estructuras de soporte definidas en esta área se acoplen a los protocolos de respuesta definidos en el plan de continuidad.

2. Se insta a las empresas que evalúen la definición de planes de continuidad de los servicios de tecnología, a que estudien los diversos estándares ofrecidos por organizaciones en el mercado, a fin de tomar las mejores prácticas de cada uno de esos estándares, para definir un plan propio y adecuado al giro del negocio.

BIBLIOGRAFÍA

1. British Standards Institute [en linea]. *BS 25999-2*. EEUU: Perry Johnson Registrars, 2007. Disponible en Web: <http://www.bsigroup.co.uk/>. [Consulta: 20 de diciembre de 2010].

2. CONRAD, Erick. *CISSP study guide*. EEUU: Syngress, 2010. 572. p. ISBN: 1597495638.

3. ISACA. *COBIT* [en linea]. 4a ed. EEUU. IT Governance Institute, 2007. Disponible en Web <http://www.isaca.org/Knowledge-Center/cobit/Documents/CobiT_4.1.pdf>.[Consulta: 05 de enero de 2011].

4. BAJGORIC, Nijaz. *Continuous computing technologies for enhancing business continuity*. EEUU: Information Science Reference, 2008. 364. p. ISBN: 1605661600.

5. Office of Government Commerce [en linea]. *ITIL Service Design*. 4a ed. EEUU: Office of Government Commerce, 2007.Disponible en Web <http://www.itil.org.uk/sd.htm>.[Consulta: 05 de enero de 2011].

6. Perry Johnson Registrars. *Steps to BS25999 Registration* [en linea]. EEUU: Perry Johnson Registrars, 2008. Disponible en Web <http://www.pjr.com/downloads/BS25999_steps.pdf>. [Consulta: 10 de diciembre de 2010].