



MINISTERIO
DE TRABAJO
Y ASUNTOS SOCIALES

SECRETARÍA DE ESTADO
DE LA SEGURIDAD SOCIAL
Gerencia de Informática



INFRAESTRUCTURA DE ALMACENAMIENTO PARA LA ADMINISTRACIÓN ELECTRÓNICA EN LA SEGURIDAD SOCIAL

Centro de Tecnología

Octubre 2007

Control de versiones

Versión	Fecha	Autor	Descripción / Comentarios
1.0	Octubre 2007	Alejandro Ruiz del Portal Antonio de la Paz	Comunicación a Tecnimap 2007

Índice

1	<u>INTRODUCCIÓN</u>	2
2	<u>PLAN DIRECTOR DE ALMACENAMIENTO</u>	3
3	<u>EVOLUCION DE LA RED DE ALMACENAMIENTO EN LA GISS</u>	5
4	<u>REPLICACIÓN</u>	7
4.1	LÍNEAS DE EVOLUCIÓN DE LA REPLICACIÓN	8



1 INTRODUCCIÓN

El reto de conseguir unos servicios públicos modernos y de calidad, dentro del ámbito de la Seguridad Social es un desafío que se plantea dentro del marco de la Administración Electrónica. Según reza en el Plan de Acción para la Administración Electrónica i2010, hay que dar solución a determinadas exigencias, a saber:

- Modernizar y hacer más eficaces los servicios públicos;
- Ofrecer servicios de mayor calidad y más seguros a la población;
- Responder a la demanda de las empresas que desean menos burocracia y más eficacia;

Uno de los principales objetivos de este plan de acción es permitir a los particulares y las empresas obtener con mayor rapidez ventajas concretas en el terreno de la administración electrónica. Así mismo, uno de los ejes prioritarios del plan i2010 es el de alcanzar una mayor eficiencia en la prestación de los servicios por parte de la Administración.

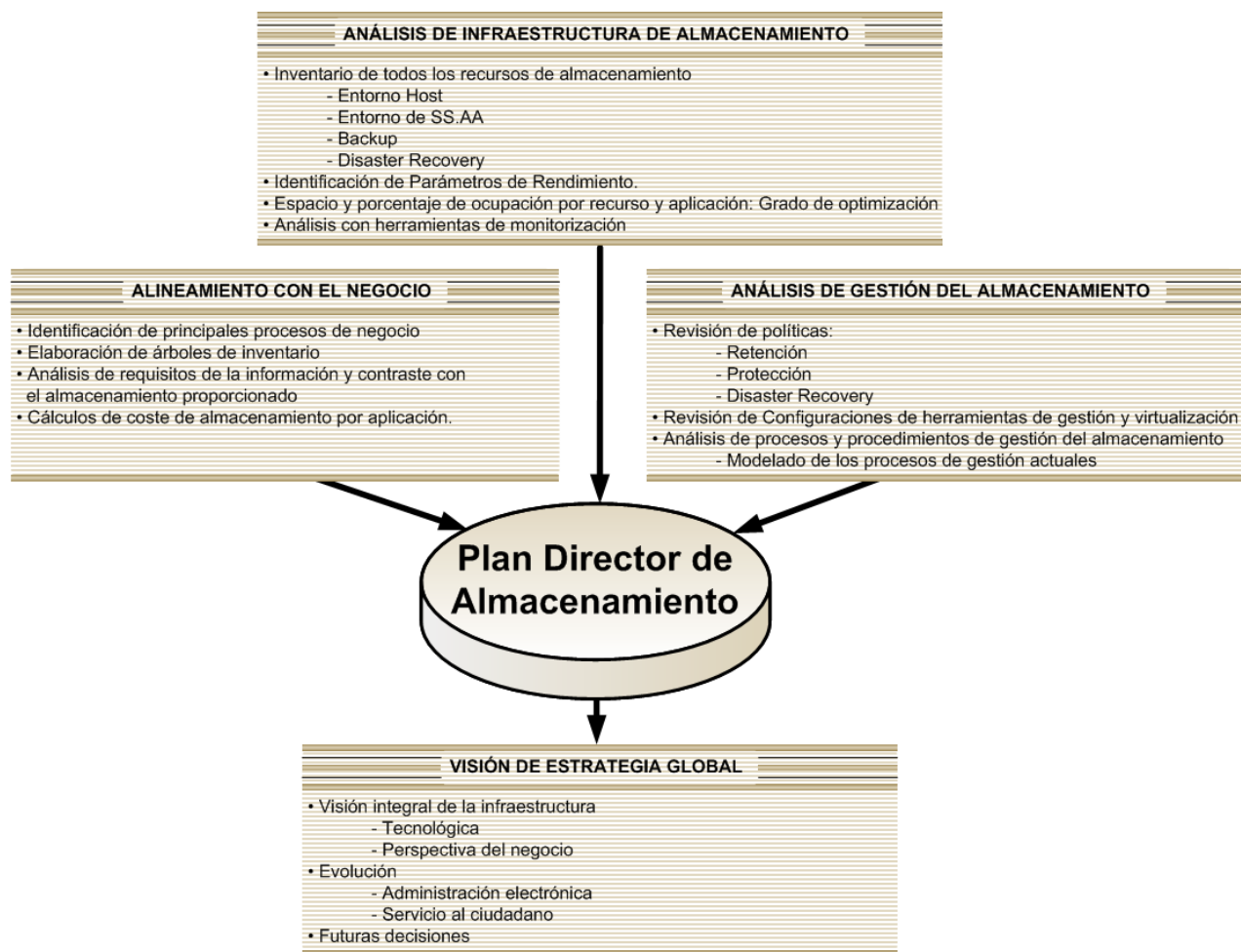
Dentro de la U.E., los estados miembros se han comprometido a lograr una mayor eficiencia, mediante un uso innovador de las tecnologías de la información y las comunicaciones (TIC) , y a reducir sustancialmente la carga administrativa de aquí a 2010.

Las TIC engloban una amplia gama de servicios, aplicaciones, tecnologías y equipos físicos y lógicos, es decir, instrumentos tales como la telefonía e Internet, el aprendizaje a distancia, las televisiones, los ordenadores, las redes y los programas informáticos necesarios para emplear dichas tecnologías.

En este documento y como uno de los componentes necesarios para la obtención de esos servicios eficientes que se pretenden ofrecer, vamos a exponer la estrategia que se ha seguido en cuanto a la evolución de la red corporativa de almacenamiento de la Seguridad Social.

2 PLAN DIRECTOR DE ALMACENAMIENTO

En los últimos años el almacenamiento de información en la Seguridad Social ha experimentado un importante crecimiento, cada vez son más los servicios que se ofrecen y las disposiciones regulatorias que obligan al almacenamiento y custodia de los datos por largos períodos de tiempo. Este sustancial aumento ha supuesto la necesidad de definir unas líneas de actuación adecuadas en el este campo para facilitar la integración de nuevos servicios. Ante esta situación la GISS (Gerencia de Informática de la Seguridad Social) esta trabajando con la idea de establecer un plan director de almacenamiento.



Para conseguir el funcionamiento de una administración electrónica eficaz, ésta tiene que estar sustentada en infraestructuras tecnológicas sólidas, fiables y orientadas al negocio al que están dando servicio. El almacenamiento y su gestión asociada es uno de los puntos clave de estas infraestructuras y una vez asumido esto, en la GISS se ha decidido abordar un estudio para establecer un plan director de almacenamiento a todos los niveles.

Una vez observado el entorno de almacenamiento, existe una problemática común en cualquier gran organización:

- Crecimiento exponencial de datos.
- Costes descontrolados.
- Problemas de escalabilidad.
- Problemas en aplicaciones por el volumen de las bases de datos.
- Ventanas de Backups demasiado grandes y requisitos de infraestructura compleja.

Se hace cada vez más necesario el establecimiento de un plan director de almacenamiento, con el objetivo de alinear el coste de los recursos de almacenamiento asignados a cada dato y el servicio prestado, con el valor que la información almacenada tiene para el negocio, tratando al mismo tiempo de optimizar los procesos de gestión de la infraestructura.

Este plan se debe realizar en base a métricas de servicio, coste y volumen de manera que se establezcan las directrices que marquen una utilización óptima de los recursos de almacenamiento y teniendo en cuenta:

- Marcos de Referencia (ITIL, COBIT, ISO17799)
- Buenas prácticas, comúnmente aceptadas
- Tendencias de la tecnología
- Prácticas habituales en entornos similares

Para su implantación se deben llevar a cabo las siguientes acciones:

1. Análisis de infraestructura de almacenamiento
 - Inventario de todos los recursos de almacenamiento
 - Entorno Host
 - Entorno de SS.AA
 - Backup
 - Disaster Recovery
 - Identificación de Parámetros de Rendimiento.
 - Espacio y porcentaje de ocupación por recurso y aplicación: Grado de optimización
 - Análisis con herramientas de monitorización
2. Análisis de gestión del almacenamiento
 - Revisión de políticas:
 - Retención
 - Protección
 - Disaster Recovery
 - Revisión de Configuraciones de herramientas de gestión y virtualización
 - Análisis de procesos y procedimientos de gestión del almacenamiento
 - Modelado de los procesos de gestión actuales
3. Alineamiento con el negocio
 - Identificación de principales procesos de negocio
 - Elaboración de árboles de inventario
 - Análisis de requisitos de la información y contraste con almacenamiento proporcionado
 - Cálculos de coste de almacenamiento por aplicación.

Como resultado del establecimiento de este plan director de almacenamiento se espera obtener una visión integral de dicha infraestructura, no sólo desde el punto de vista tecnológico sino también desde la perspectiva del negocio, que en nuestro caso se centra en la administración electrónica y en los servicios ofrecidos al ciudadano, consiguiendo alinear ambos puntos de vista.

Esta visión integral permitirá formular una estrategia global en cuanto a la infraestructura de almacenamiento que conducirá las futuras decisiones que se tengan que tomar en este ámbito dentro de la GISS.

3 EVOLUCIÓN DE LA RED DE ALMACENAMIENTO EN LA GISS

A partir del año 2005 la GISS hace una apuesta por modernizar su infraestructura de almacenamiento incidiendo fundamentalmente en dos líneas:

- **División del almacenamiento en niveles**
- **Virtualización del almacenamiento.**

Desde la implementación y puesta en servicio del **almacenamiento estructurado en diferentes niveles**, se ha producido una ampliación de capacidad de los sistemas de discos. Los diferentes niveles de almacenamiento actualmente implantados son:

- Primer Nivel. Formado por cabinas con discos Fibre Channel que ofrecen un mayor rendimiento y fiabilidad.
- Segundo Nivel. Compuesto por cabinas con discos SATA, que ofrecen más capacidad y menor coste, pero tienen menor rendimiento y disponibilidad.
- Tercer Nivel. Constituido por virtualizadores de cinta. Estos son sistemas de discos que emulan cintas de cara al servidor, donde el servidor escribe en el virtualizador y posteriormente el virtualizador se encarga de escribir los datos a cinta sin afectar al rendimiento del servidor, de esta manera se consigue reducir los tiempos de escritura y lectura de este soporte.
- Cuarto Nivel. Establecido por un robot de almacenamiento en cintas de alta capacidad para datos productivos, fundamentalmente backup.

La necesidad de crear niveles de almacenamiento surgió por la aparición de nuevos servicios en el entorno. Se realizó un análisis de los requisitos de las nuevas aplicaciones y se llegó a la conclusión de que el almacenamiento de segundo nivel cumplía con las necesidades de disponibilidad, velocidad de acceso, frecuencia de acceso, etc., de muchos de estos servicios. En consecuencia, gran parte de la ampliación necesaria en disco se realizó sobre la capacidad de almacenamiento de segundo nivel. Esta ampliación supuso con respecto a ampliar la misma cantidad de almacenamiento, en disco de primer nivel, un ahorro del 30% en los costes.

Independientemente del tipo de **virtualización del almacenamiento** que se implemente, ésta aporta al entorno las siguientes características:

- muestra todos los sistemas de discos como un pool de almacenamiento sin tener en cuenta su localización;
- enmascara las diferencias entre dispositivos heterogéneos y de diferentes fabricantes;
- gestiona de un modo común los volúmenes, la seguridad, la replicación y los servicios de copia para cualquier sistema de almacenamiento.

Es decir, mantiene bajo un interfaz único, la gestión y control de entornos heterogéneos, la movilidad de datos entre sistemas heterogéneos (de distintos fabricantes) de forma transparente y sin interrupción del servicio y ajusta los recursos de almacenamiento a las necesidades del negocio.

Además de las dos líneas de actuación principales, se optó por la estrategia de separar los entornos Mainframe y Sistemas Abiertos, manteniendo cada plataforma aislada a nivel de cabinas de discos.

A modo de resumen presentamos en las siguientes figuras la evolución realizada y la prevista en el periodo 2005-2009, tanto para la red de almacenamiento correspondiente al entorno de Sistemas Abiertos, como la correspondiente al entorno Mainframe

EVOLUCIÓN DE LAS REDES S.A.N. EN GISS. < ENTORNO SISTEMAS ABIERTOS >					
	2005	2006	2007	2008	2009
SERVIDORES	AMPLIAR SERVIDORES	INCORPORAR SERVIDORES CON TECNOLOGÍA BLADE.	INCORPORAR TECNOLOGÍA VMWARE. VIRTUALIZACIÓN DE SERVIDORES ESTUDIO DE GESTIÓN DINÁMICA DE CICLO DE VIDA (CORREO ELECTRÓNICO)	AMPLIAR SERVIDORES DE FICHEROS (NAS) AMPLIAR SERVIDORES CON TECNOLOGÍA BLADE	ANÁLISIS DE TECNOLOGÍA WAFS SERVIDORES DE FICHEROS DESCENTRALIZADOS ESTUDIO DE GESTIÓN DINÁMICA DE CICLO DE VIDA DE LA INFORMACIÓN
DWDM		AMPLIAR DWDM. DOBLE RUTA Y AMPLIACIÓN DE PUERTOS DE FIBRA.		AMPLIAR DWDM. TECNOLOGÍA DE 4 GB.	AMPLIAR DWDM. TECNOLOGÍA 10 GB EN ISL.
DIRECTORES DE FIBRA	INCORPORAR DIRECTORES DE FIBRA. TECNOLOGÍA 2 GB.		AMPLIAR DIRECTORES DE FIBRA. TECNOLOGÍA 4 GB.	INCORPORAR NUEVOS DIRECTORES DE FIBRA. TECNOLOGÍA 4 GB	ESTUDIO TECNOLOGÍA ISCSI SAN A TRAVÉS DE REDES IP AMPLIAR DIRECTORES DE FIBRA. TECNOLOGÍA 10 GB.
ALMACENAMIENTO EN CINTA		UNIFICAR LA ROBÓTICA. INCORPORACIÓN DE ROBOT PARA SUSTITUIR ROBÓTICA DISEMINADA			UNIFICAR TECNOLOGÍA DE VIRTUALIZACIÓN EN CINTA
ALMACENAMIENTO EN DISCO	AISLAR LA PLATAFORMA DE ALMACENAMIENTO MAINFRAME – SS. AA. INCORPORAR CABINAS DE DISCOS FIBRE CHANNEL DEDICADAS A SS.AA.	INCORPORAR DISCOS SATA. ALMACENAMIENTO DE 2º NIVEL IMPLANTAR COPIA REMOTA SINCRONA DE DATOS ENTRE CABINAS. C. PRIMARIO -C. SECUNDARIO	AMPLIAR ALMACENAMIENTO DE 2º NIVEL	IMPLANTAR TECNOLOGÍA VIRTUALIZACIÓN EN DISCO. IMPLANTAR 3º COPIA DE DATOS PARA SITUACIONES DE DISASTER RECOVERY ESTUDIO DE TECNOLOGÍA DE ARCHIVADO POR CONTENIDO (CAS)	VIRTUALIZAR A NIVEL DE FICHEROS (NAS) TECNOLOGÍA 4 GB EN DISCO

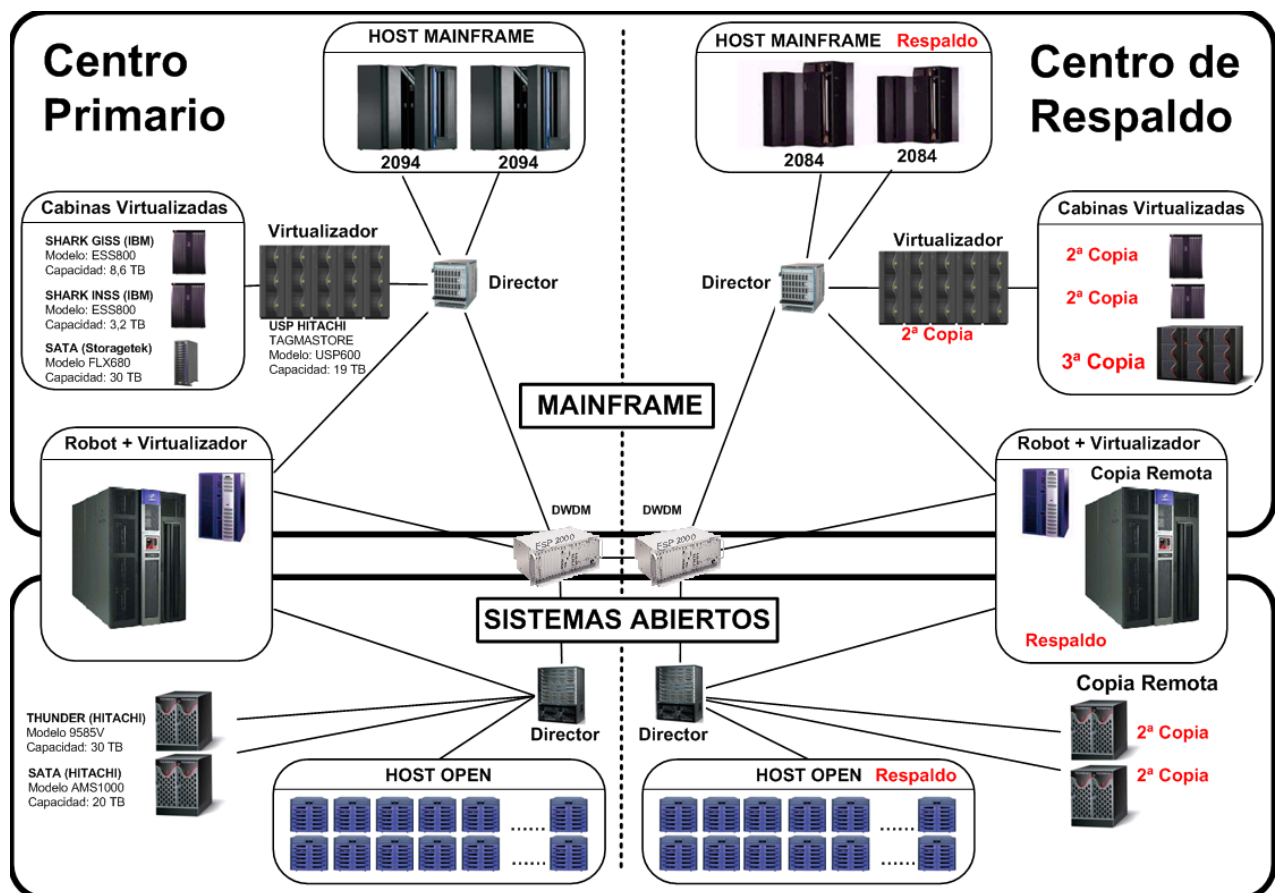
EVOLUCIÓN DE LAS REDES S.A.N. EN GISS. < ENTORNO MAINFRAME >					
	2005	2006	2007	2008	2009
SERVIDORES	INCORPORAR SERVIDORES Z/SERIES 2094.		AMPLIAR POTENCIA DE LOS SERVIDORES Z/SERIES 2094.	AMPLIAR POTENCIA ON DEMAND DE LOS SERVIDORES Z/SERIES 2094.	VALORACIÓN DE NECESIDADES DE AMPLIACIÓN DE SERVIDORES
DWDM		AMPLIAR DWDM. DOBLE RUTA Y AMPLIACIÓN DE PUERTOS DE FIBRA		AMPLIAR DWDM. TECNOLOGÍA DE 4 GB	AMPLIAR DWDM. TECNOLOGÍA 10 GB EN ISL'S
DIRECTORES DE FIBRA	AMPLIAR DIRECTORES DE FIBRA. TECNOLOGÍA 2 GB		INCORPORAR DIRECTORES DE FIBRA. TECNOLOGÍA 4 GB		AMPLIAR DIRECTORES DE FIBRA. TECNOLOGÍA DE 10 GB
ALMACENAMIENTO EN CINTA		UNIFICAR LA ROBÓTICA. INCORPORACIÓN DE ROBOT PARA SUSTITUIR ROBÓTICA DISEMINADA	AMPLIAR LOS SISTEMAS VIRTUALES EN CINTA PARA TODOS LOS ENTORNOS MAINFRAME (GISS, INSS, SICOSS)		UNIFICAR TECNOLOGÍA DE VIRTUALIZACIÓN EN CINTA
ALMACENAMIENTO EN DISCO	AISLAR LA PLATAFORMA DE ALMACENAMIENTO MAINFRAME - SIST. ABIERTOS IMPLANTAR TECNOLOGÍA VIRTUALIZACIÓN EN DISCO INCORPORAR DISCOS DE BAJO COSTE PARA DATOS DE 2º NIVEL	IMPLANTAR 3º COPIA DE DATOS PARA SITUACIONES DE DISASTER RECOVERY AMPLIACIÓN DE SISTEMAS DE VIRTUALIZACIÓN EN CABINA (USP'S) VIRTUALIZAR CABINAS DE DISTINTOS FABRICANTES (ESS800-IBM, 9980V-HDS, SATA-STK)	AMPLIAR EL ESPACIO EN DISCO DE 1º NIVEL	APLICAR MEJORAS EN LA GESTIÓN DE LA 2ª Y 3ª COPIA EN EL CENTRO DE RESPALDO INCORPORAR TECNOLOGÍA DE 4 GB EN CABINAS DE DISCOS	AMPLIAR EL ESPACIO EN DISCO DE 1º NIVEL

4 REPLICACIÓN

En la actualidad la protección de datos se ha convertido en un requisito básico para las organizaciones. Independientemente del sector de la industria, las compañías operan cada vez más en entornos 24x7, lo que ha provocado que conceptos como continuidad de negocio y recuperación ante desastres se hayan convertido en críticos a la hora de implementarlos en sus infraestructuras. Como resultado de estas necesidades, surge la tecnología de copiar los datos en centros de respaldo situados en localizaciones remotas. Esta política, es a día de hoy, la más efectiva para mantener la continuidad de negocio.

La siguiente imagen refleja la infraestructura de los dos centros que tiene actualmente la GISS:

- Centro primario, con los servidores, virtualizadores, cabinas de discos, robótica y equipos de comunicaciones destinados a la producción.
- Centro de Respaldo con todos los equipos dedicados a la replica de datos y preparados para la puesta en marcha en caso de contingencias en el centro principal.



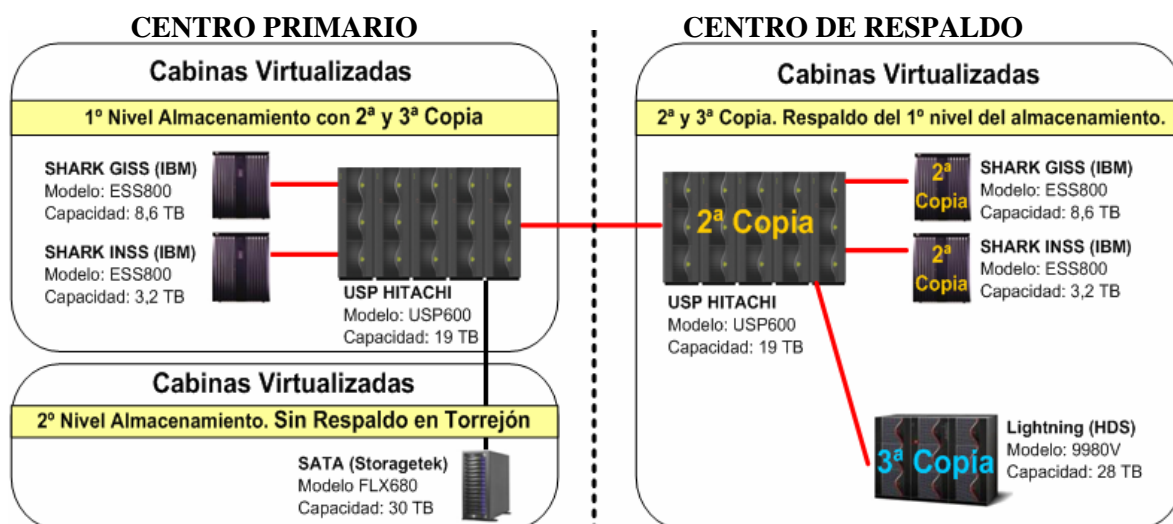
La replicación aumenta la disponibilidad de los datos debido a las siguientes funcionalidades que aporta:

- Automatización general de los procedimientos, reduciendo de esta forma la duración de eventos planificados como, mantenimientos del sistema, pruebas de aplicaciones y backup de datos.
- Permite realizar backup de los datos actuales de producción de forma no disruptiva y sin afectar al rendimiento en la producción.
- La restauración de datos y la recuperación ante una parada no planificada es más rápida que la restauración basada en cinta.

- El sistema secundario puede tomar los procesos del primario, de esta forma evita las paradas planificadas.
- Permite hacer pruebas de recuperación ante fallos de forma no disruptiva por medio de copias online de los datos de producción.

En la GISS debido a la criticidad de los datos, la pérdida de información en caso de fallo debe ser mínima, por tanto, se ha implantado una replicación síncrona contra el centro secundario; de esta forma, se tiene una copia exacta de los datos en todo momento. Además de esta segunda copia de respaldo, la GISS ha implantado una tercera copia o copia local en el centro remoto, con esto se asegura tener al menos un duplicado consistente de los datos remotos. Por tanto, el uso de forma cíclica de copias locales en combinación con cualquier tipo de réplica remota, síncrona o asíncrona, protege los datos de posibles inconsistencias y de fallos en cadena.

PLATAFORMA MAINFRAME. ALMACENAMIENTO EN DISCO



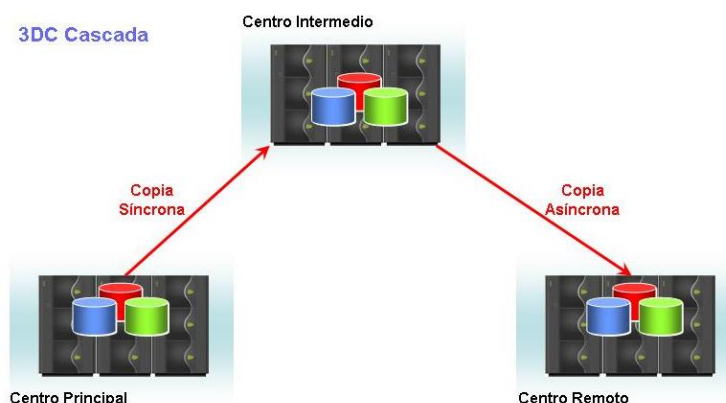
Las copias locales (terceras copias) además de su utilidad para la recuperación de datos, son usadas para realizar backups de datos incluso en momentos de máximo rendimiento de la aplicación sin que afecte a la misma, por aplicaciones de análisis de datos y para pruebas de integración de nuevas versiones de componentes del sistema o nuevos productos antes de ser integrados en el entorno de producción.

4.1 LÍNEAS DE EVOLUCIÓN DE LA REPLICACIÓN

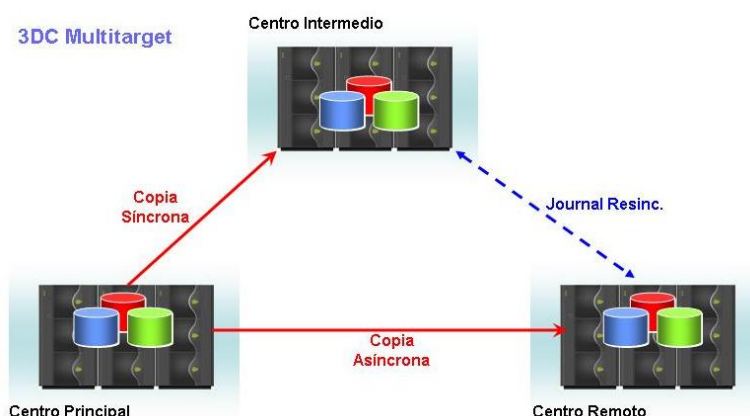
En la actualidad se está realizando un estudio para la implementación de una solución con tres centros de datos (3DC), uno principal y los otros dos de respaldo, donde uno tiene una imagen de los datos mediante la replicación síncrona y el otro con tecnología asíncrona. Esta nueva estrategia aporta lo mejor de la copia síncrona y asíncrona: rápida recuperación y ninguna pérdida de datos ante fallos locales en el centro de producción y proteger los datos frente a desastres regionales. Existen dos topologías la hora de implementar tres centros de datos: en cascada y multi-target.

La topología en cascada inicialmente establece una copia síncrona contra un centro de respaldo situado en la misma región o área (centro intermedio). Este centro de respaldo a su vez realiza una replicación asíncrona de sus datos contra otro centro de respaldo situado fuera de la región o área

(centro remoto). La principal desventaja de esta topología es que si se produce un fallo en el centro intermedio se para todo el proceso de replicación, tanto el síncrono como el asíncrono.

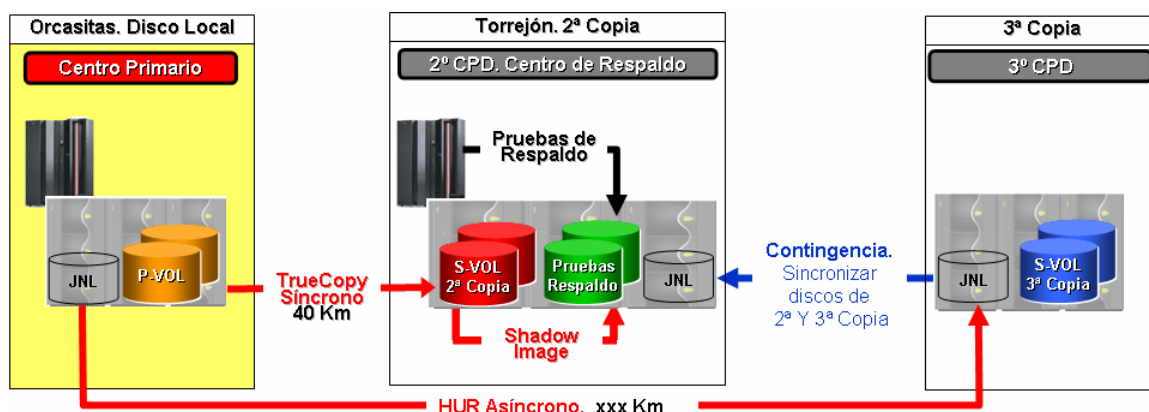


La topología multi-target desde el centro principal o de producción realiza de forma independiente y simultánea una copia síncrona contra un centro de respaldo y una copia asíncrona contra el otro centro de respaldo. Para garantizar la independencia de los dos procesos de replicación, las copias se realizan sobre redes de replicación diferentes.



Esta topología introduce un fichero journal. Este fichero guarda los registros lógicos de las actualizaciones y transacciones en el orden que van ocurriendo en el centro intermedio. En caso de contingencia en el centro de producción, su uso, nos permite resincronizar de forma asíncrona el centro intermedio con el centro remoto, realizando copias diferenciales de los datos, no siendo necesario emprender una copia completa entre ambos centros.

La opción elegida por la GISS es la topología multi-target, debido principalmente a que se evita que un fallo en las comunicaciones entre el centro primario y el secundario deje aislada la 2ª y 3ª copia de respaldo. La configuración desde el punto de vista de los CPDs se muestra en la siguiente imagen:



Con esta evolución, la GISS obtiene los siguientes beneficios:

- Dispondrá de una 2ª Copia consolidada y sincronizada en un CPD próximo a Orcasitas con el objetivo de cubrir contingencias a nivel local en el CPD primario.
- Dispondrá de una 3ª Copia consolidada al 100% de todos los datos en disco del entorno mainframe, en un CPD alejado del CPD primario.
- No existe límite de distancia entre el centro primario y el tercer CPD que contiene la 3ª Copia. En la actualidad, en otras organizaciones se emplea este sistema para realizar copias incluso entre continentes (Europa – Suramérica).
- Con esta configuración se tiene garantizado la integridad de los datos ante cualquier tipo de contingencia de cualquier naturaleza: Desastre natural (fuego, inundaciones, terremotos), desastre humano (terrorismo, fallos intencionados), fallos hardware, fallos eléctricos, etc.