COMUNICACIÓN TECNIMAP 2007

Nombre VICTOR
1er Apellido DEUTSCH
2º Apellido FRANCO
NIF (con letra de control) 02736136X
Teléfono 91483 2111
Correo electrónico
VictorEduardo.DeutschFranco@telefonica.es
Organismo / Empresa Telefónica Soluciones
Puesto de trabajo Director Asociado
Dirección de trabajo Ronda de la Comunicación S/N Edificio Norte 2 Planta 1
28050 Madrid

Título de la comunicación

LAS NUEVAS ESTRATEGIAS DE RECUPERACION ANTE CONTINGENCIAS EN EL AMBITO DE LAS AA.PP.

Resumen de la comunicación

Tradicionalmente, la recuperación ante una situación de contingencia se realizaba (y se sigue realizando en muchos casos), mediante la descarga de las cintas de back up en un CPD alternativo.

La consecuencia de esta estrategia es que el tiempo de recuperación del procesamiento de datos es proporción directa del tiempo de descarga de cintas, por definición un proceso lento y no exento de riesgos (deterioro del medio, errores de grabación).

Por otro lado, la pérdida de datos ante una contingencia grave (que implique la pérdida de los discos de explotación), se extiende hasta el último back up realizado. En la mayoría de los casos esto nos lleva normalmente al cierre del día anterior.

Hoy en día muchas AA.PP. funcionan en tiempo real. La nueva ley de Acceso Electrónico de los Ciudadanos profundiza esta tendencia. Unos tiempos de recuperación tan largos y con tanta pérdida de datos pueden tener graves consecuencias, tanto por la interrupción prolongada del servicio público, como por la pérdida de confianza en las instituciones, sin contar con el efecto multiplicador en la actividad económica y social en general.

Tema de la Comunicación Infraestructuras y servicios comunes para el desarrollo de la Administración Electrónica.

"LAS NUEVAS ESTRATEGIAS DE RECUPERACION ANTE CONTINGENCIAS EN EL AMBITO DE LAS AA.PP."

Introducción

El concepto de Disaster Recovery en el ámbito de las AA.PP., tiene bastantes años, y va asociado con la creciente dependencia de las organizaciones públicas de los sistemas informáticos. Desde la década de los años ´60 la rápida informatización de los sistemas de seguridad y defensa de los EE.UU., creó la necesidad de disponer "planes de contingencia" para mantener el funcionamiento de los mismos en situación de "guerra" o "agresión".

A medida que estas organizaciones empezaron a informatizarse aprendieron rápidamente:

- Que la informatización creaba una creciente dependencia de los procesos de negocios, de la fiabilidad y proceso continuo de los sistemas
- Que estos no eran absolutamente fiables, y que era esperable un nivel de fallo
- Que los factores externos de infraestructura o medioambientales afectaban drásticamente, aunque en forma indirecta, a sus procesos. Por ejemplo, un apagón prolongado podía producir sustanciales alteraciones en sus procesos de negocios.
- Que la falta de continuidad o de confiabilidad en el proceso tenía consecuencias económicas inmediatas, sobre todo cuánto más aumentaba la productividad mediante la automatización de los procesos

Debido a esto, bastante pronto en la prehistoria del sector informático, muchas organizaciones civiles empezaron a adoptar modelos de planificación ante contingencias, similares a los que habían comenzado a ejecutar sus colegas militares.

Estos modelos evolucionaron desde el principio general de redundancia, que implica duplicar totalmente las infraestructuras de equipamiento informático y comunicaciones, incluyendo los soportes de datos. De esta forma, en caso de un fallo en la infraestructura original se dispone de una réplica exacta de la misma en otra locación física.

La estructura subyacente de cualquier plan de contingencia incluía, entonces, el arranque ordenado del procesamiento de datos desde la infraestructura duplicada, lo que implicaba generalmente el desplazamiento físico de los soportes de datos y hasta de los propios usuarios finales.

Sin embargo, el progresivo aumento de la complejidad en la arquitectura informática de las empresas (proliferación de servidores departamentales, redes de PC, conexiones entre empresas), trajo consigo un aumento considerable de los riesgos, la complejidad y el coste de implementar los "planes de contingencia" tradicionales. Ya no es aplicable la solución tradicional de disponer de un centro alternativo con un ordenador central similar al propio y descargar allí las cintas de back up.

Hoy en día cualquier organización de tamaño medio utiliza un gama amplia de soluciones tecnológicas que van desde los grandes mainframes, pasando por servidores intermedios, a los terminales y dispositivos móviles, y todos ellos, tienen un grado de criticidad para la ejecución normal de los procesos de negocios.

Más aún, la demanda de los ciudadanos por mayores y mejores servicios ha llevado a las Administraciones a invertir fuertemente en la automatización de sus procesos, lo que
ha aumentado exponencialmente la dependencia de estas organizaciones de sus sistemas informáticos. Actualmente, prácticamente todos los procesos internos de una organización, desde los más elementales, conllevan algún tipo de automatización. Una disrupción en la infraestructura informática, por cualquier causa, tiene un impacto mucho más grave que antaño.

Una consecuencia de esto, es que los tiempos que anteriormente eran asumibles para recuperar el procesamiento normal de una organización afectada por un desastre, actualmente tengan tendencia a

reducirse. Un mayor número de entidades públicas desarrolla hoy operaciones on-line (en tiempo real) y en 24x7, con lo cual desaparecen las "ventanas" de tiempo sin operaciones, que se podían utilizar para recuperar las instalaciones, con escaso impacto en el servicio.

Desde los años ´90 se han desarrollado una serie de tecnologías que nos permiten enfocar el problema de una manera diferente. Entre estos desarrollos se encuentran:

- La revolución en las tecnologías de comunicaciones de los años ´90 y sobre todo la explosión del ancho de banda
- La mayor "inteligencia" de los dispositivos de almacenamiento de datos, que actualmente, pueden interconectarse con diferentes sistemas y con diferentes esquemas de redundancia
- Las nuevas tecnologías de consolidación y virtualización de equipamiento informático, que facilita una mayor economía de escala

La suma de presiones por una mayor exposición de las organizaciones a una catástrofe informática, los menores tiempos asumibles de recuperación y la aparición de estas nuevas tecnologías hacen necesario dar un nuevo enfoque al tradicional planeamiento ante contingencias.

Se puede argumentar que los cambios están empujados por las nuevas formas que están adoptando las arquitecturas tecnológicas en las organizaciones, a partir de la irrupción de estas mismas tecnologías. Vale decir, si se cambia la forma de procesar, por ejemplo, de una arquitectura distribuida a una centralizada, la estrategia de continuidad también debe variar.

Esto es cierto, pero también veremos que, en determinadas circunstancias, aunque no se produzcan cambios significativos en el modelo de la arquitectura tecnológica de una organización, las tecnologías citadas pueden proveernos de un nuevo punto de vista, de nuevas soluciones a viejos problemas, que nos permitan optimizar nuestras estrategias y asegurar un mejor desempeño, ante una situación de crisis.

Sin embargo, son pocas todavía, en proporción, las organizaciones que han comenzado a re-elaborar sus planes de contingencia, a partir de los nuevos principios. También es cierto, que un gran número de organizaciones solamente dispone de soluciones solamente intuitivas, sin una planificación formal, normalmente no por desidia, sino por una menor percepción del riesgo, o por prioridades en la asignación del gasto y las inversiones.

De acuerdo con estudios exhaustivos como el realizado por Hitachi Data Systems, en España, todavía el 19% de las empresas públicas y privadas carece de una estrategia de recuperación de desastre y de continuidad de negocio. Además, el 32% no dispone de un centro remoto de recuperación de desastre.

Muchas veces, el cambio de enfoque no da por el hecho de que las organizaciones se sienten razonablemente seguras con los esquemas de disaster recovery de antaño, que adicionalmente la mayoría de ellas nunca ha utilizado en la práctica. Convengamos que las infraestructuras de respaldo no son la primera prioridad en organizaciones que tienen que lidiar con un entorno cambiante y con múltiples demandas de los agentes sociales.

Todos estos factores contribuirán, entonces, a revisar las estrategias tradicionales de Disaster Recovery, y a la adopción de nuevas y mejores soluciones basadas en las tecnologías expuestas.

Nuevas estrategias

De los aspectos mencionados anteriormente, quizás el más relevante sea el relacionado con las comunicaciones de datos. Actualmente, la disponibilidad de una red de comunicaciones muy robusta, segura y de alta velocidad hace posible:

- la centralización de servidores departamentales en los Centros de Procesamiento de Datos
- la replicación de datos asíncrona (modo templado)

- la restauración de "imágenes" (datos y configuración) de servidores desde diferentes locaciones geográficas
- la división de la plataforma de Producción en dos centros a distancias apreciables (cluster geográfico) en modo activo-activo

Por otro lado, el desarrollo de la tecnología de virtualización en el ámbito de los sistemas x86 ha permitido:

- la posibilidad de respaldar un sistema con equipamiento de otro fabricante diferente sin riesgos de confiabilidad
- mayor flexibilidad para asignar y liberar recursos de procesamiento

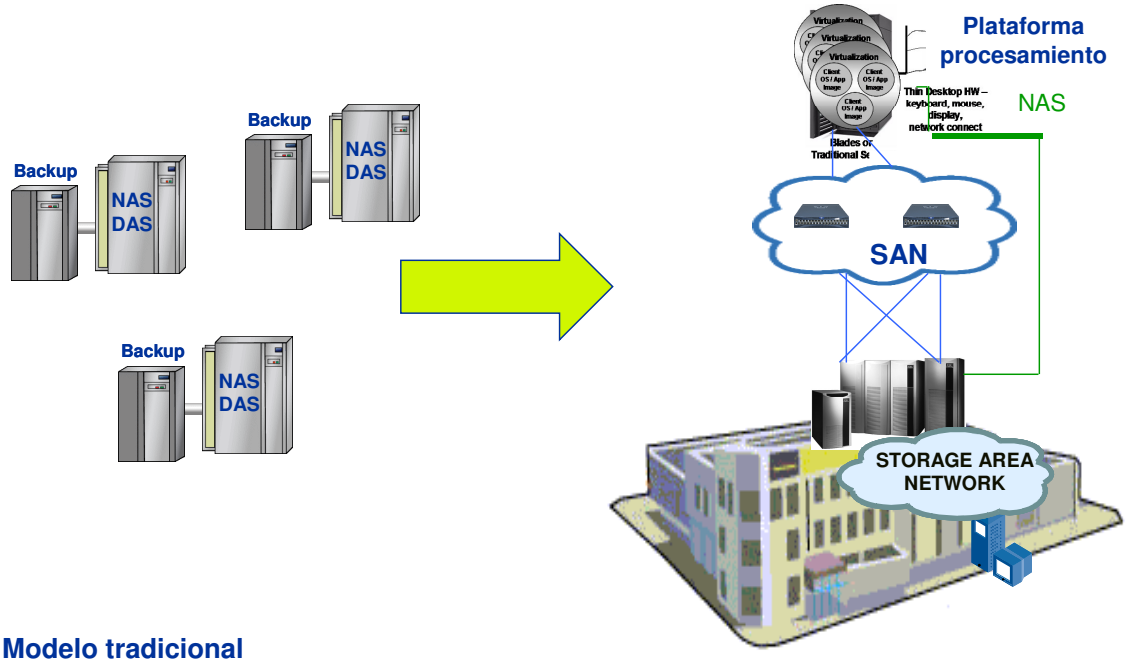
A continuación veremos detalles de cada una de estas soluciones.

Centralización de servidores y aceleración IP

La centralización de los servidores departamentales o situados en oficinas remotas en los CPDs hoy en día es posible sin pérdida de productividad para los usuarios, dado que existen soluciones que permiten lograr velocidades similares a las de LAN a través de redes WAN a grandes distancias.

Esta medida evita duplicar costosas infraestructuras de respaldo para estos equipos (metros acondicionados, medidas de seguridad física, equipos de back up). Al llevar estos servidores al CPD, todos sus datos y aplicaciones quedan protegidos por misma infraestructura física, que respalda a los sistemas centrales.

Estas soluciones se basan en la tecnología de aceleración IP. Consiste en incorporar dispositivos en la red que permiten hacer un mejor aprovechamiento del ancho de banda disponible y reducir sustancialmente la latencia en las comunicaciones. En cada extremo del enlace estos dispositivos utilizan técnicas de compresión y buffering, logrando que la experiencia de un usuario remoto en una WAN sea similar a la de un usuario que trabaja en una LAN.



Modelo tradicional
1 servidor = 1 unidad de almacenamiento

Plataforma de Almacenamiento Gestionado

Esta tecnología está cambiando el diseño de las redes de PC, pasando del modelo distribuido tradicional, a un esquema mucho más centralizado. En el nuevo esquema los datos de estos servidores no se almacenarán en los discos locales de servidores dispersos, sino en grandes infraestructuras de almacenamiento centralizado.

Un ejemplo en el marco de la Administración es la centralización de servidores de los centros periféricos de un sistema sanitario. Dadas las actuales normas de Protección de Datos, los ficheros con datos personales de pacientes, historias clínicas y prescripciones, requieren del máximo nivel de seguridad. Para lograrlo cada centro sanitario, por pequeño que sea, debe realizar una serie de inversiones en elementos de resguardo: barreras de acceso físico y lógico, grabadores de cintas, cajas fuertes o armarios ignífugos y, por supuesto, recursos humanos especializados en sistemas.

Al centralizar todos los datos en un único CPD accesible a través de la red se simplifica la gestión de cada centro sanitario, reduciéndose los riesgos. Este CPD único permite una mayor economía de escala para aplicar soluciones de resguardo de mayor envergadura y un mejor aprovechamiento de los recursos técnicos.

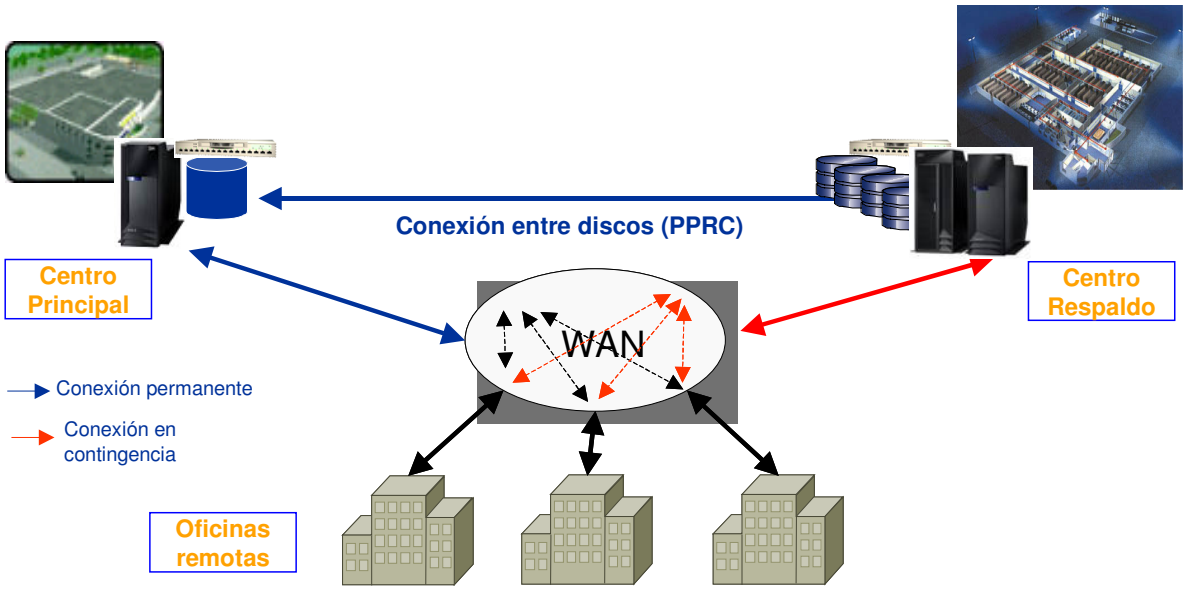
Replicación de discos: posibilidades

Las actuales tecnologías de replicación nos permiten generar un almacenamiento de datos secundario totalmente idéntico al principal en un CPD alternativo, con gran seguridad. De tal forma, el "switching" de un equipo de Producción al de Respaldo se reduce dramáticamente, al eliminarse el tiempo de "download" de las cintas en el segundo. La caída de los costes y la evolución de la tecnología de almacenamiento hace posible este tipo de soluciones, mediante una conexión de banda ancha rápida y robusta.

Existen diferentes tipos de estrategias para la replicación de datos. Vamos a mencionar a continuación sólo algunas de ellas. A priori las podemos clasificar en dos grupos:

- Replicación física
- Replicación lógica

La replicación física o peer-to-peer consiste en replicar los datos a través de la red disco a disco, o mejor dicho, sector a sector de los discos de explotación, en una segunda



infraestructura de almacenamiento.

Esta es una solución relativamente simple y de bajo coste, que garantiza la disponibilidad de la copia remota de los datos, por lo menos con la frecuencia de sincronización, que admita el ancho de banda disponible (algunas horas).

No obstante, tiene algunos inconvenientes:

- Es exigente en requerimientos de comunicaciones, consumiendo todo el ancho de banda disponible
- Errores de grabación en el almacenamiento principal se replican en el de respaldo
- Tiene limitaciones de compatibilidad entre diferentes modelos de cabinas de almacenamiento y, por supuesto, para los equipos que tienen discos internos

El segundo grupo, replicación lógica, lo constituyen las soluciones basadas en software. En este caso, existe un software de replicación que se ejecuta con independencia de la infraestructura de almacenamiento que está por debajo y agrega una serie de facilidades importantes:

- Optimiza el uso de la red, reduciendo la ventana de sincronización, lo que permite una sincronización más frecuente
- Proporciona funcionalidades para la detección y corrección de errores de grabación
- Minimiza los problemas de compatibilidad entre cabinas o entre discos y cabinas

Como contrapartida, la principal desventaja de las soluciones de replicación lógica es que suelen ser más costosas, dada la necesidad de adquirir una licencia de software adicional para implantarlas, junto con los costes derivados de la instalación, configuración, operación y mantenimiento del mismo.

Sin embargo este grupo tiene una composición muy heterogénea, dado que se incluyen en él soluciones específicas como las propias de los gestores de BB.DD., utilidades diversas de diferentes aplicativos (correo electrónico, por ejemplo), y soluciones de "imaging". Nos detendremos en este último grupo, que es el que proporciona una cobertura más general al problema.

Replicación de Imágenes

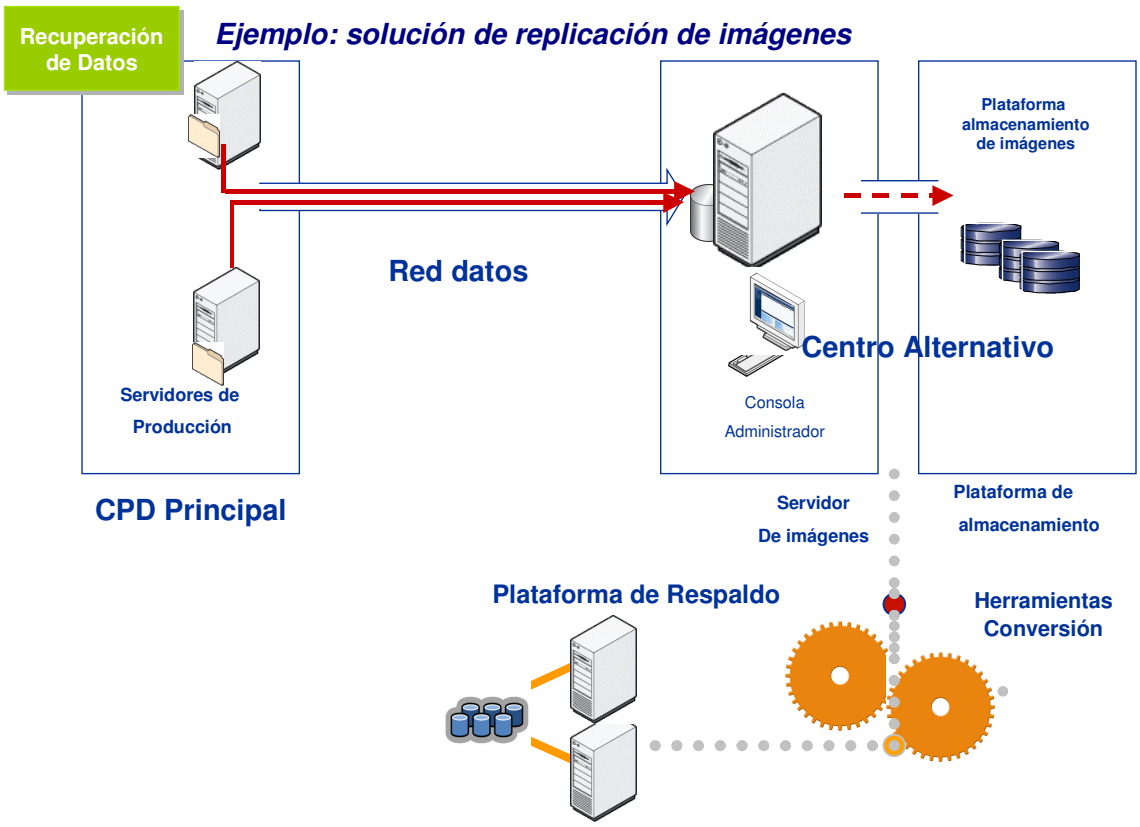
Esta solución es similar a la anterior, pero con el agregado de que, además de replicar los datos, vamos a copiar exactamente la configuración del servidor de Producción, de tal manera que podamos generar una réplica exacta del mismo en el CPD de Respaldo.

De esta forma, se evitan posibles problemas de compatibilidad, al utilizar el conjunto de datos replicados en el servidor de respaldo. Utilizando esta solución la configuración del servidor de Producción se reconstruye completamente al momento de la última sincronización en la plataforma de respaldo.

Esta solución implica:

- Instalar un agente del software de replicación de datos en los servidores de Explotación
- Montar una plataforma de servidores físicos de captura de imágenes en el CPD alternativo, que ejecute el software de replicación de datos
- Establecer un enlace entre cada servidor de Explotación y el servidor de imágenes en el CPD alternativo con el ancho de banda adecuado para realizar la actualización de la copia de resguardo, con la periodicidad deseada

La imagen de los discos de explotación estará siempre almacenada y actualizada en los discos lógicos de la plataforma de respaldo. Mediante las herramientas de gestión de la plataforma, esta "imagen" de los servidores físicos se convertirá, en un almacenamiento de datos en condiciones de ser asociado a los servidores de respaldo, sin pérdida de datos.



Este proceso de conversión normalmente tiene una duración variable, según la misión y características de cada servidor. El tiempo de recuperación debe ser establecido para determinar el tiempo total de recuperación de los sistemas. También es importante destacar que esta solución no es aplicable en todos casos. En ciertas configuraciones singulares la conversión de esta "imagen" en el almacenamiento del servidor de respaldo, puede no ser posible o demorar demasiado tiempo como para que sea practicable.

División de la plataforma de Producción en dos centros

Con una red potente es posible incluso plantearse la subdivisión de la plataforma de producción en dos locaciones distantes entre si un apreciable número de kilómetros, para disminuir el riesgo de catástrofes localizadas en un punto geográfico.

Las implicancias de esta decisión son las siguientes:

- Particionar la capacidad de proceso en dos, instalando cada mitad en un centro diferente.
- Vincular ambos centros con un enlace de alta velocidad como CWDM o DWDM, idealmente con redundancia por dos rutas y también a nivel de equipos de routing
- Establecer un cluster entre las dos plataformas de proceso, de tal forma que el sistema continue en operaciones incluso en caso de una caída completa de uno de los dos
- Sincronizar también la actualización de los datos en disco de tal manera que ambos centros dispongan de la última versión de datos

Con una solución así ante un siniestro completo en uno de los dos centros, la plataforma del centro sobreviviente podrá continuar operando sin que el sistema haya dejado de funcionar en ningún momento

desde la perspectiva de usuario, más allá de alguna pérdida de performance, según el dimensionamiento que tenga la plataforma. Tampoco deberemos soportar pérdida alguna de datos.

Esto se logra al precio de una fuerte inversión en equipamiento y en servicios de comunicaciones, pero que puede justificarse en función del impacto económico, social o institucional que pueda implicar una pérdida de datos o de tiempo de proceso por breve que fuese.

Respaldo con plataformas virtuales

La plataforma x86 ha tenido un desarrollo excepcional desde su aparición en el mercado. Actualmente, un gran número de sistemas críticos para el sector público se soportan bajo su arquitectura, especialmente en lo que se refiere a servicios que se brindan a través de Internet a ciudadanos y empresas.

Uno de los principales problemas al diseñar soluciones de respaldo para esta línea de equipamiento eran las posibles incompatibilidades entre los diferentes fabricantes, a pesar de que todos compartieran una arquitectura básica común.

El desarrollo de plataformas de virtualización o hipervisores nos permite abstraernos de este problema. Hoy en día un hipervisor es capaz de crear un servidor virtual que se ejecute sobre cualquier hardware, independizándose casi completamente de las características físicas de la máquina en la que se aloja.

De esta forma es posible respaldar un servidor de Producción de un determinado fabricante con un servidor virtual, que se ejecuta el hardware de otro, solamente garantizando la compatibilidad a nivel de versiones y parches de sistema operativo y aplicaciones a instalar.

Vale decir que si un servidor de Producción queda indisponible por cualquier motivo, no es necesario que el Centro de Respaldo disponga de otro similar donde descargar los datos. Solamente se necesita activar un servidor virtual (sobre un servidor físico de cualquier otro fabricante) con una versión de SO y parches similar a la del servidor de Producción y descargar allí los datos antes de recuperar el proceso.

Por supuesto, existen consideraciones de performance y algunas restricciones aún que no permiten que esto se de en todos los casos (a veces legales o de licenciamiento), pero hoy día es una alternativa válida para respaldar la amplia mayoría de los servidores x86 en Producción, y una forma diferente de encarar las estrategias de contingencia.

Conclusión

Las organizaciones son cada vez más dependientes de sus sistemas informáticos. Una contingencia que inhabilite los mismos por un tiempo prolongado, actualmente tiene consecuencias más graves desde el punto de vista del negocio que hace algunos años.

Los perjuicios son crecientes, sin hablar de los aspectos de imagen. Los tiempos de recuperación y la pérdida de datos que se podían asumir en el pasado, actualmente son mucho más exigentes.

En este contexto, es necesario revisar las estrategias de contingencia de las organizaciones, que pueden ya no ser válidas, y analizar la factibilidad de implantar nuevas soluciones tecnológicas, particularmente en el dinámico mundo de la recuperación de datos.