Copias de seguridad y restauración

CONCEPTOS

Concepto general

- Una copia de seguridad, respaldo, copia de respaldo o copia de reserva (en inglés backup y data backup) en ciencias de la información e informática es una copia de los datos originales que se realiza con el fin de disponer de un medio para recuperarlos en caso de su pérdida.

- La Perdida de Datos en perspectiva

La Perdida de Datos en perspectiva

La información es valiosa

- Las Empresas modernas dependen de la información, y una gran cantidad de ella se almacena en discos duros de servidores y ordenadores personales.
- Los días en los cuales los documentos eran almacenados en filas de archivadores han terminado. Hoy día, simplemente, hay demasiada información para ser almacenada y que su búsqueda dependa de personal humano.
- Y, así como el almacenamiento de papel tiene riesgos (tales como incendios, inundaciones, pérdida, robo, etc…), el almacenamiento electrónico de datos también tiene riesgos.
- Los gestores de empresa competentes entienden el valor de la información que almacenan electrónicamente, y toman las medidas apropiadas para proteger esa información contra la pérdida o destrucción, ya sea deliberada o accidental.
- La analogía es obvia: las organizaciones asegurar sus instalaciones, bienes de equipo, inventario, e incluso el personal clave de las amenazas internas y externas. Los datos, al igual que cualquier activo crítico, debe ser asegurado también.

El valor de los diferentes tipos de datos

- Diferentes activos son protegidos de diferentes formas, dependiendo de lo importante que son y lo fácil que es reemplazarlos.
- Los diferentes tipos de datos , también necesitan diferentes niveles de protección.

- Una simple división separa los datos informáticos en cinco categorías:

1. Código de sistema operativo y utilidades

- Esta categoría incluye los archivos de arranque de sistema, el kernel, controladores de dispositivos, parches y aplicaciones de ayuda necesarias para el correcto funcionamiento del sistema operativo (OS).
- Estos archivos pueden ser muy grandes, pero son fáciles de reemplazar.
- Responsablemente, los administradores de TI hacen copias de los dispositivos de arranque y almacenan las claves de licencia en un sitio seguro para estar preparado si es necesario reinstalar un servidor, un ordenador de escritorio o un portátil.
- Con acceso a la web/internet, las versiones del sistema operativo necesario y los parches son fáciles de recuperar, si ha guardado las claves de licencia.

2. Datos del Sistema Operativo

- Esta categoría se compone de la información, los perfiles de usuario y la personalización de apariencia del sistema operativo.
- Ejemplos de este grupo son los datos de Directorio Activo de Windows, las entradas del registro de Windows, NDS de NetWare, y /etc/inittab en UNIX.
  - El archivo /etc/inittab. Después de que arranque y el núcleo monte el sistema de archivos de root, el primer programa que ejecuta el sistema es init.
- Si se pierden estos datos, un administrador de sistemas expertos en un entorno sencillo podría ser capaz de volver a recrearlos manualmente, pero esto es algo que nunca querrá hacer.
- Tener copia de seguridad de estos datos es muy importante.

3. Aplicaciones

- Al igual que el sistema operativo base, estos datos son bastante fáciles de crear con las descargas de Internet y disponiendo de las claves de licencia.
- Este grupo incluye a los archivos ejecutables, bibliotecas y controladores necesarios para que una aplicación desarrollada por nosotros funcione.

4. Datos de Aplicación

- Estos datos incluyen configuraciones de la aplicación, particiones de base de datos, formatos de tablas, y los datos almacenados en las tablas.
- Esta información suele cambiar con mucha frecuencia, y por lo general sería imposible recrear manualmente.
- Una vez más, las copias de seguridad son fundamentales para estos datos.

5. Datos de Usuario

- Estos datos incluyen los archivos creados y modificados por los usuarios de forma individual.
- Al igual que los datos de aplicación, cambian muy a menudo.
- Estos son los datos que más a menudo son restaurados desde copias de seguridad, ya que son los que se dañan o eliminan de forma accidental más a menudo.

La pérdida de datos

- Con diferencia al resto, la forma más común de pérdida de datos es por errores de los usuarios.
  - Los usuarios pueden eliminar, modificar o mover y perder archivos importantes o partes de los mismos sin darse cuenta, realizar cambios incorrecto en bases de datos o eliminar correo electrónico que desean conservar.
- Los datos pueden ser destruidos o dañados de muchas otras maneras.
  - Los virus pueden infectar archivos.
  - Un error en una aplicación o sistema operativo puede convertir un archivo en buen estado en uno ilegible.
  - Sabotajes deliberado por usuarios malintencionados o intrusos, pueden causar estragos.

… La pérdida de datos

- Fallos de Hardware pueden provocar las mayores pérdidas de datos.
  - Los Discos duros tienen tasas de error medibles.
  - Refrigeración inadecuada puede causar fallos en los discos, y picos de tensión o cortes de energía pueden dañar los datos.
  - La lista es interminable.
- Peor aún, centros de datos enteros, llenos de servidores pueden ser destruidos o resultar inaccesibles, debido a incendios, huracanes, tornados, inundaciones, terremotos, guerras, terrorismo, o un sin número de desastres naturales o provocados por el hombre.

Hacer copias de datos

- Varias tecnologías se han inventado y comercializado para hacer frente a los riesgos de pérdida y corrupción de datos.
- Estas soluciones ofrecen a los gestores una combinación de sistemas en cuanto a fiabilidad, funcionalidad, y costes.
- Algunas de las tecnologías utilizadas incluyen las siguientes **Tecnologías de almacenamiento**:

Copia Espejo (Mirroring)

- Se ha desarrollado para proteger los datos contra fallos de hardware en un sistema de almacenamiento.
- En las copias espejo, se mantienen dos o más versiones idénticas del mismo conjunto de archivos o datos.
- En el caso de que la copia principal de los datos se convierte en inaccesible, el otro –la imagen espejo– está disponible automáticamente.
- Con el costo de duplicar el hardware de discos duros, esta copia es relativamente cara.
- Tampoco proporciona ninguna protección contra una corrupción de datos, infección por virus, o eliminación de archivos.
- Estos errores simplemente se escriben en ambas copias y por lo tanto, ambas estarán mal.

Matriz de discos (RAID – matriz redundante de discos independientes)

- Es una colección de más de dos discos conectados a una controladora especializada y gestionada por software.
- Las diferentes configuraciones de «RAID» incluyen discos en espejo, discos de conjunto a bandas (donde los datos se escriben a través de un número de discos de forma simultánea para reducir las lecturas y escrituras), y los discos con paridad (que proporcionan redundancia si un disco falla, pero sin la necesidad de duplicar totalmente el hardware).
- Todas las diversas configuraciones de RAID proporcionan mejoras en la confiabilidad y el rendimiento, o ambos.
- Una vez más, sin embargo, RAID no protege contra datos corruptos, infección por virus, o eliminación de archivos.

La Replicación de Datos

- Mantiene varias copias de los datos.
- A diferencia de la copia espejo, la replicación tiende a implicar el movimiento de datos a distancia, en un sentido físico o lógico.
- La replicación también implica generalmente que los datos originales y su copia replicada no están necesariamente sincronizados instantáneamente.
- Soluciones de replicación están generalmente basadas en software.

Proveedores de Almacenamiento en la Nube

- Alquilan espacio en disco, generalmente por un costo gigabyte/mes, para las empresas que no quieren comprar y gestionar todo el espacio que necesitan.
- Estos proveedores utilizan tecnologías de replicación y RAID para asegurar que los datos de sus clientes sigue estando disponible.
- Suelen ofrecer otros servicios de gestión de almacenamiento como las copias de seguridad, restauración, archivado o deduplicación.
  - La deduplicacion es una técnica especializada de compresión de datos consistente en eliminar bloques de datos duplicados cuando se realizan y se transfieren copias de seguridad (se detalla mas adelante).
- Los clientes que deseen estos servicios se los deben realizar ellos mismos o buscar proveedores adicionales.

Sistemas de gestión de almacenamiento jerárquico (HSM – Hierarchical Storage Management)

- También hace copias de los datos, pero su objetivo es ahorrar dinero, no para proporcionar redundancia.
- Cuando un archivo no se ha utilizado en un plazo determinado de tiempo, por lo general meses o incluso años, HSM mueve los datos a un medio de almacenamiento más lento y menos costoso, por lo general, una biblioteca de cintas, dejando «tickets de recuperación» en lugar de los archivos de datos originales.
- Estos tickets dejan el archivo original visible para el sistema operativo de forma que cuando un usuario necesita acceder al mismo, el sistema HSM recupera los datos en segundo plano de los medios de copia más lentos, casi en línea y los proporciona a el usuario.
- El propósito principal de este sistema es reducir el costo de almacenamiento de datos, todavía hay sólo una copia de los datos.
- Una solución de copia de seguridad sigue siendo necesaria para proteger contra cualquier tipo de pérdida o corrupción de datos.

Archivado

- Es el siguiente paso después de HSM.
- Después que los datos llegan a cierta edad o cumplen otros criterios definidos por la empresa, se mueve permanentemente del almacenamiento en línea al almacenamiento fuera de línea.
- Los datos archivados se suelen conservar durante largos periodos de tiempo, por lo general más de un año, y se recuperan sólo en circunstancias excepcionales.
- Los administradores de sistemas tienen que mover y colocar los medios físicos donde se guardan dichos datos con el fin de poder restaurarlos el día de mañana.
- Un archivo se identifica no sólo por los datos que contiene, sino por el punto en el tiempo en que se traspaso el archivo al almacenamiento fuera de línea.

Compresión

- Se refiere a una familia de tecnologías diseñadas para ahorrar espacio de almacenamiento mediante el reconocimiento de patrones en los datos almacenados.
- Los datos comprimidos se crean mediante la reescritura de los datos originales en un formato más eficiente.
- Cuando se usan los datos, la descompresión se utiliza para volver a recrear los datos originales.

Deduplicación

- Es una forma de compresión que reconoce cuando un archivo, bloque de datos, o una cadena de bytes es idéntica a otra que ya está almacenado en el sistema. A continuación, se elimina una de las copias, dejando sólo una referencia de la segunda copia a la primera.
- El objetivo de la deduplicación es ahorrar espacio de almacenamiento, o ahorrar ancho de banda cuando es necesario mover o copiar en una red de área amplia (WAN), redes de área local (LAN), o incluso una red de área de almacenamiento (SAN) grandes cantidades de datos.
- Los archivos pueden duplicarse fácilmente cuando se realizan copias para enviar a otro usuario, un sistema de correo electrónico puede contener cientos o miles de copias de un mismo archivo cuando es adjuntado en mensajes.
- Del mismo modo, cuando un archivo se copia y luego se edita, la mayor parte de los bloques de la copia son idénticos a los de la original. La deduplicación puede ahorrar espacio al permitir que las dos copias de archivos compartan los datos que permanecen comunes a ambos.

… Deduplicación

- La deduplicación puede ralentizar la escritura de archivos, o la transmisión de archivos o bloques, porque se necesita tiempo para que el procesador de almacenamiento, optimizador de WAN o dispositivo de deduplicación, analice los datos y reconozca los duplicados.
- Las lecturas de estos archivos también pueden ser más lentas, ya que las piezas modificadas del archivo más unos «punteros» a los bloques comunes, se encuentran dispersos en todo el sistema de almacenamiento.
- Reconstruir el archivo no se llama «de-duplicación», sino «rehidratación». La palabra nos invita a imaginar la adición de agua a los alimentos deshidratados para conseguir comida tan buena como la original.

Copias de seguridad (Backups)

- Tienen dos aspectos fundamentales: los de archivado y los de redundancia de copia, duplicación RAID o replicación.
- Una copia de seguridad es una copia de los datos de producción realizada en un momento determinado en el tiempo.
- A diferencia de la deduplicación, el objetivo de una copia de seguridad es crear una copia separada, conservada en caso de pérdida o deterioro del original.
- La copia de seguridad se pueden almacenar en disco, cinta u otros medios de comunicación, y se puede mantener en línea o fuera de línea, de forma local o fuera de la oficina, lejos de la fuente original de los datos.

Hacer copias de seguridad o backups es más que realizar solo copias

- Por lo general, las copias de seguridad se realizan sobre una base de tiempo regular, por lo general: diaria, semanal o mensualmente.
- Cuando los usuarios o administradores solicitan que se restablezcan los datos de copia de seguridad, se puede elegir entre varias copias realizadas en distintas fechas.
- Estas imágenes de los datos en un tiempo dado son la mejor protección contra errores, borrado o destrucción.
- Los usuarios pueden pedir la restauración de una copia realizada en una fecha en particular, cuando el archivo se sabe que está en buenas condiciones.
- El tiempo de recuperación de datos depende de la cantidad de datos que se van a restaurar, la ubicación de copia de seguridad de datos, el tipo de copia de seguridad, los medios de copia de seguridad, y el paquete de software utilizado.

REQUISITOS DE RESPALDO Y RECUPERACIÓN

Perspectivas

- Los equipos de protección de datos son responsables de algunas de las actividades más críticas de una organización.
- Las políticas, las pruebas y las tecnologías adecuadas son clave para un plan de respaldo y recuperación.
- La protección de datos es una de las actividades más esenciales en TI, y la copia de seguridad y la recuperación de datos son sus componentes clave.
- Seguir algunos requisitos clave de copia de seguridad y recuperación ayudaría a garantizar que los recursos estén fácilmente disponibles y que las actividades de protección de datos sean seguras y protegidas.

… Perspectivas

- Las actividades de protección de datos deben ser repetibles y accesibles para los miembros de TI fuera del equipo de respaldo de datos las realicen si es necesario.

- Prepararnos para posibles auditorías de TI:
  - Las organizaciones con políticas y procedimientos bien documentados,
  - la configuración adecuada de las tecnologías de respaldo y recuperación
  - y las pruebas periódicas de las actividades de respaldo y recuperación refuerzan la confiabilidad de un programa de protección de datos

Lo anterior permite estar preparados para posibles **auditorías de TI**

POLITICAS Y PROCEDIMIENTOS

- El primer requisito, y probablemente el más importante, es tener políticas y procedimientos documentados para la copia de seguridad y la recuperación.

- Si bien muchos empleados de TI responsables del respaldo y la recuperación podrían realizar esas tareas con los ojos vendados, los procedimientos documentados brindan un nivel adicional de comodidad y confianza a los departamentos de TI, especialmente si los miembros del equipo de respaldo y recuperación designados no están disponibles.

- La experiencia con la pandemia de COVID-19 ha subrayado que cualquiera puede contraer y quedar fuera de servicio por una enfermedad.

… POLITICAS Y PROCEDIMIENTOS

- Cualquiera en un departamento de TI podría no poder trabajar de repente, incluso si los miembros de un departamento de TI general tienen experiencia previa en copia de seguridad y recuperación de datos, el acceso a documentos con los procedimientos claramente detallados significa que prácticamente cualquier persona puede realizar copias de seguridad y recuperación de datos.

- Las políticas de respaldo y recuperación de datos no tienen que ser tan específicas y granulares como los procedimientos reales, pero su presencia es importante desde una perspectiva de auditoría.

… POLITICAS Y PROCEDIMIENTOS

- Las políticas muestran que la organización se toma en serio la protección de datos, y especialmente el respaldo y la recuperación, los cuales son elementos críticos en la continuidad del negocio corporativo y los programas de recuperación ante desastres.

- Los auditores de TI generalmente buscan evidencia documentada de políticas y procedimientos.

- Es importante, tanto desde la perspectiva operativa como de auditoría, revisar y actualizar periódicamente las políticas y los procedimientos para verificar que se describen con precisión cómo realizar copias de seguridad y recuperaciones de datos.

TECNOLOGÍA

- Si bien las políticas y los procedimientos son quizás el número uno entre los requisitos de respaldo y recuperación, el acceso a las tecnologías y recursos asociados más relevantes y rentables también es fundamental.

- Estas incluyen aplicaciones de software y repositorios de almacenamiento de datos.

- Las aplicaciones de software de respaldo son componentes integrales de los esfuerzos de una organización para proteger a la empresa de pérdidas de datos, corrupción y robos.

TECNOLOGÍA

- Las aplicaciones de software de copia de seguridad y recuperación de datos ayudan a identificar lo siguiente:
  - archivos de datos, bases de datos y sistemas y aplicaciones críticas;
  - criterios de respaldo y recuperación;
  - ubicaciones de los datos y las copias de seguridad del sistema;
  - horarios de respaldo de datos;
  - verificación de respaldo; y
  - procesos de recuperación de datos, aplicaciones y sistemas.

- Las aplicaciones pueden residir en servidores locales o en ubicaciones alternativas, como en el almacenamiento en la nube.

- Las ubicaciones de almacenamiento de respaldo pueden estar dentro o fuera del sitio.

PRUEBAS

- La prueba de las actividades de protección de datos es la última, pero no menos importante, en esta lista de requisitos clave de copia de seguridad y recuperación.

- Además de las copias de seguridad programadas con regularidad y las actividades de copia de seguridad de emergencia, las pruebas programadas de las actividades de copia de seguridad y recuperación son esenciales.

- Las pruebas de respaldo deben garantizar que los datos que respalda una organización estén en el lugar de almacenamiento designado.

- Más importante aún, las pruebas aseguran que los datos respaldados sean los mismos que los activos de datos primarios.

… PRUEBAS

- Las pruebas comprueban que una organización lleve a cabo medidas de seguridad, como el cifrado cuando los datos están en tránsito, según sea necesario.

- La organización también debe realizar pruebas para confirmar de que puede descifrar y validar los datos. Esto es igualmente cierto con las copias de seguridad de sistemas y aplicaciones: deben ser seguras, sin concesiones y frecuentes.

- Las pruebas periódicas de las capacidades de recuperación de datos informan que los recursos de información de la empresa están disponibles y sean accesibles rápidamente en caso de emergencia.

… PRUEBAS

- Si bien las organizaciones pueden realizar pruebas de respaldo de datos casi a diario, según la frecuencia y los tipos de actividades de respaldo, las pruebas de recuperación de datos se pueden programar a intervalos periódicos, como mensuales, quincenales o incluso semanales.

GRACIAS