**Seguridad de base de datos**

Prevenir la pérdida de datos es una tarea fundamental en la administración de bases de datos y las copias de seguridad son un mecanismo de recuperación con el cual los administradores pueden restaurar una base de datos que ha perdido información por cualquier circunstancia.

La responsabilidad que conlleva garantizar la confidencialidad y privacidad de la información por parte de las organizaciones es relevante en todo aspecto por lo que es necesario mitigar los riesgos que provoquen cualquier tipo de ataque como accesos no autorizados de información o pérdida de datos por fallas humanas o del sistema. La protección de los datos debe llevarse a cabo contra fallos físicos lógicos humanos intencionados o no y cada uno debe ser tratado de forma distinta.

Cuadradas catastróficas del sistema debido a errores de hardware o de software con la consecuente pérdida de contenido en la memoria fallos de soporte físico que ocasiona la pérdida de la información guardada errores de software de las aplicaciones que causan fallo en una o más transacciones de sangre físicos naturales como incendios inundaciones apagones de instrucción negligente o no intencionada de datos o instalaciones por operarios o usuarios y sabotaje o destrucción intencionada de los datos o del hardware.

Es pensable que en una organización se planifiquen estrategias que permitan anticipar la mayoría de los posibles desastres o fallos y así obtener una rápida recuperación mientras que continúe operando para ello existen 2 conceptos que son fundamentales y que influyen en este en este proceso el recovery time objective que determina la mayor cantidad de tiempo que la organización está dispuesta a asumir en la restauración por un error o un fallo mientras una actividad se encuentra sin ejecutar teniendo en cuenta que el impacto dentro del operatividad del negocio será bajo es decir especifica la demora máxima hasta que el servicio de aplicaciones sea restaurado y el recovery point objective establece el volumen de datos en riesgo de pérdida que la organización está dispuesta a asumir que dependerá de la cantidad de transacciones realizadas a la base de datos por unidad de tiempo la recuperación puede requerir revertir a un punto consistente antes del fallo descartando las actualizaciones emitidas después de ese punto de recuperación el objetivo del punto de recuperación es definir la ventana de tiempo máxima permitida para las recientes actualizaciones que pueden perderse y inclusive puede ser cero o sea ninguna pérdida es tolerable.

La política de retención especifica la forma en qué se gestionará la creación de vacas de una base de datos teniendo en cuenta los requerimientos por parte de la organización es decir que la frecuencia y el tipo de copia de seguridad se determinará principalmente con respecto a los cambios la información debe determinarse a tal fin la redundancia que establece un número fijo de backup que debe mantenerse.

La recuperación de fallos que equivale a una restauración de la base de datos a algún estado anterior de modo que sea posible reconstruir un estado correcto cercano al momento de fallo a partir de tal estado anterior sí hay daños físicos el método de recuperación restaurará una copia anterior y reconstruiría el estado rehaciendo las operaciones confirmadas restantes hasta el momento del fallo si la base de datos no presenta daños físicos pero los datos se han vuelto inconsistentes no es necesario restaurar una copia completa de la base sólo las últimas operaciones registradas que pudieran haber causado la inconsistencia se debe entonces deshacer operaciones y rehacerlas a fin de restaurar un estado consistente.

Un sistema gestor de bases de datos debe proporcionar las siguientes funcionalidades como ayuda a la recuperación un mecanismo de copia de seguridad mediante el cual se hagan copias de seguridad periódicas de la base de datos fuera de la base de datos facilidades de registro que mantengan el control del estado actual de las transacciones y de los cambios realizados en la base de datos funcionalidad de puntos de comprobación check points que permita que las actualizaciones de la base de datos que estén llevándose a cabo se hagan permanentes inspector de recuperación que permita al sistema restaurar la base de datos a un estado coherente después del fallo.

Las estrategias típicas de recuperación de una base de datos son básicamente 3 copias de respaldo en otros medios de almacenamiento los backup operaciones específicas de rehacer y deshacer y registrar los movimientos realizados en un archivo diario o bitácora llamado log de las transacciones.

El archivo log es una tabla de la base de datos que contiene una lista secuencial de todas las modificaciones a cada objeto de la base de datos es el administrador de la base de datos quien define qué datos son registrados por ejemplo fecha y hora usuario tabla accedido operación realizada terminal desde dónde se conecta etcétera guardándose el estado anterior y posterior del objeto con quién los datos completos de los cambios desde el último backup y cada registro de cambio de look contiene el identificador de la transacción que ejecute el cambio el identificador del ítem de datos modificados típicamente la ubicación en el disco el valor viejo que fue sobrescrito el valor nuevo luego de la escritura y opcionalmente tiempo de comienzo de la transacción tiempo del comité o sea cuando está la transacción confirmada un tiempo de aborto etcétera.

En este esquema podemos seguir visualmente ese proceso cuando Comienza una transacción de se registra su inicio en el log guardándose el valor existente cuando todas las acciones de la transacción t se ejecutan exitosamente se escribe el cómic en el log con el nuevo valor y por último se realizan las operaciones de escritura en la base de datos.

La operación de recuperación usa 2 primitivas rehacer rehace la actualización registrada en el blog escribe el valor nuevo sobre el ítem de datos modificados y deshacer que deshace la actualización registrada en el blog es decir escribe el valor antiguo sobre el ítem de datos modificados.

Las técnicas de recuperación son 2 vía reprocesamiento en la base de datos se retorna a un estado correcto previo conocido y se reprocesan las transacciones ejecutadas hasta el momento del fallo la otra opción es la recuperación basadas en el contenido del log y puede adoptar 3 estrategias por actualización diferida las las transacciones no cambian la base de datos hasta que alcanzan el punto de confirmación y lo alcanzan cuando el log ha sido actualizado la estrategia en este caso es no deshacer rehacer la actualización inmediata actualiza la base de datos usando el registro de escritura anticipada fuerza las escrituras en el log antes de escribir en la base de datos ejecutando el cómic de la transacción después de escribir en el blog la estrategia es deshacer rehacer Rollback/rollfoward se salva la base de datos periódicamente y se mantiene un archivo lo auxiliar con los cambios producidos en orden cronológico rollback deshace los cambios erróneos y reprocesa las transacciones válidas y rollfoward rehace los cambios en la base de datos usando datos salvados y transacciones válidas desde el último backup.

Cuando se produce un fallo debe determinarse qué transacciones hay que rehacer y cuáles hay que deshacer la infección del archivo log completo es costoso en tiempo y muchas transacciones ya habrán estado confirmadas en la base de datos persistentes para reducir la carga de trabajo que implica la restauración un checkpoint registra antes del cual el sistema gestor de la base de datos estaba en un estado consistente y todas las transacciones fueron confirmadas siguen el siguiente procedimiento se graban todos los registros del log actualmente en memoria principal en un almacenamiento estable se graban en la base de datos todos los bloques de registros almacenados en el buffer y se escribe en el blog entonces un registro de tipo check point.

Los algoritmos de recuperación son técnicas que garantizan las propiedades de atomicidad consistencia y durabilidad de los datos a pesar de las fallas durante la recuperación diferida se consideran los siguientes pasos

1. recorrer hacia atrás desde el final del log hasta alcanzar el checkpoint más reciente
2. continúa recorriendo hasta encontrar el momento de comienzo de control
3. considerar la parte de lo que sigue a este registro
4. para todas las transacciones comenzadas desde el punto de control TC o después que no hayan registrado commit ejecutar deshacer
5. recorrer el log hacia adelante para todas las transacciones comenzadas desde el punto TC que hayan registrado cómic ejecutando rehacer.

Aquí vemos un esquema de este procedimiento el punto el punto de control se marcó en TCY la caída del sistema en TF la transacción t uno ya había confirmado el cómic así como t 2 YT 3 pero no así t cuatro YT 5 por lo tanto t uno puede ignorar si los cambios ya están registrados mediante el jovit en la base de datos antes del momento TC qué 2 YT 3 deben rehacerse porque llegaron a su punto de confirmación después del inicio del punto de control qué cuatro YT 5 deben deshacerse porque ninguna llegó a confirmarse no hay cambios asentados en la base de datos.

El Back up Es una copia de los datos originales que se realiza con el fin de disponer de un medio para recuperarlos en caso de pérdida el objetivo principal de los back ups justamente preservar información por un periodo de tiempo determinado para ser recuperada y restaurada en caso de problemas o fallas que se presenten en la base de datos a nivel físico o lógico él operación ideal consiste en que las copias de seguridad sean generadas sin necesidad de intervención manual la automatización de dicho proceso disminuye la probabilidad de errores que dependan del factor humano. Tenemos los siguientes tipos de back ups en la copia de seguridad parcial- Partial Database Back Up- es una copia de una parte de la base de datos se debe considerar que después de restaurar una copia de este tipo es necesario sincronizarla con el resto de la base de datos el Whole database backup preserva todos los datos generados a la fecha desde el último archivo de control los archivos de control almacenan el estado de la estructura física de una base de datos en ellos se guarda el estado actual de la base de datos y son utilizados para mantener la consistencia interna iniciar las operaciones de recuperación full backup en la copia de seguridad completa también conocido como backup nivel cero es la copia de todos los datos la opción Whole database backup realiza una copia de seguridad de todos los archivos de la base de datos pero esta copia de seguridad no está catalogada es decir no guarda la estructura de directorios lo que significa que no puede utilizarse como referencia para una copia de seguridad incremental la opción full backup sí lo hace y es utilizado para el backup incremental.

El backup incremental también conocido como backup de nivel uno es una copia de los cambios realizados en la base de datos desde el anterior vaca catalogado generado se dividen acumulativo y genera una copia desde el último backup nivel cero full backup a la fecha diferencial que genera una copia desde el último backup nivel uno a la fecha.

El sistema gestor de la base de datos debe proveer técnicas que permitan a ciertos usuarios tener acceso a porciones específicas de una base de datos sin tener acceso al resto mediante un sistema de seguridad autorización de acceso y cifrado de datos se trata de un aspecto que tiene que ver con la protección de datos contra accesos no autorizados y para protegerlos de una posible corrupción durante todo su ciclo de vida el usuario administrador de la base de datos es el responsable de la seguridad global de la base de datos ya que cuenta con funciones de privilegio para crear cuentas con claves de usuario concede el privilegio de acceso revocar o modificar estos privilegios establecer niveles de seguridad para que las cuentas accedan solo al nivel asignado o gestionar el archivo de auditoría donde quedan registradas las operaciones que realizan los usuarios la seguridad es medida según la protección implementada para hacer frente a ataques externos a fallas de software y hardware y a la manipulación de datos por los usuarios ya mencionamos que las medidas de seguridad abarca seguridad física como controlar el acceso a los equipos acceso a lugares donde están instalados terminales clientes servidores y bases de datos horario donde ha permitido realizar las operaciones etcétera seguridad de usuarios acceso sólo al personal autorizado mediante uso de códigos de acceso y palabras clave huella digital reconocimiento de voz etcétera intento de bloqueo fallidos calidad de las claves confidencialidad y tiempo de validez etcétera y la seguridad lógica disco replicado control de calidad y desempeño del hardware encriptación de los datos registro de las operaciones en el archivo de auditoría, etc.

Los siguientes requisitos son esenciales para la seguridad de una base de datos el sistema debe tener capacidad para verificar que sus acciones han sido autorizadas las acciones de los usuarios deben ser supervisadas de modo tal que pueda descubrirse cualquier acción indebida o errónea ningún sistema puede evitar de manera absoluta las intromisiones malintencionadas pero es posible hacer que resulte muy difícil eludir los controles el sistema debe diseñarse entonces a prueba de intromisiones los programadores por ingeniosos que sean no deben poder pasar por alto estos controles los datos deben poder ser sometidos a procesos de auditoría la falta de auditoría en los sistemas de computación ha permitido la comisión de grandes delitos los datos deben ser reconstruirles porque por muchas precauciones que se tomen siempre ocurren accidentes y la base de datos debe ser protegida contra incendio robo agresión y otras formas de destrucción.

Veamos ahora algunos conceptos relevantes a la seguridad de base de datos en qué estación es un procedimiento que utiliza un algoritmo de cifrado con cierta clave llamada clave de cifrado para transformar un mensaje sin atender a su estructura lingüística o significado de tal forma que sea incomprensible o al menos difícil de comprender a toda persona que no tenga la clave de descifrado del alba en esquemas de clave simétrica las claves de cifrado y descifrado son las mismas las partes comunicantes deben tener la misma clave para lograr una comunicación segura en los esquemas de cifrado asimétrico o de clave pública la clave de cifrado se publica para que cualquiera pueda usar y cifrar mensajes sin embargo sólo la parte receptora tiene acceso a la clave de descifrado que permite leer los mensajes presenta tus ventajas principales suprime el problema de transmisión segura de la clave y permite la firma electrónica.

Hackeo vulnerabilidad o intrusión refiere a ataques posibles desde el exterior mediante virus que es un trozo de código que se adosa a un programa legítimo contaminándolo cuando un programa contaminado se ejecuta ejecutará también el código del virus lo que permitirá nuevas reproducciones además de alguna acción que puede ir desde un simple mensaje y molesto hasta la destrucción de los archivos el troyano es un programa aparentemente útil que contiene un trozo de código que al ejecutarlo le brinda a un atacante acceso remoto al equipo infectado la puerta trasera es un punto de entrada secreto dejado por los implementadores del sistema para saltarse los procedimientos normales de seguridad y Pishing deja corriendo en una terminal un programa que pide login y luego password para engañar a los usuarios de modo que estos revelen su clave.

Un Firewal Es un elemento basado en hardware software o una combinación de ambos que controla el flujo de datos que entra y sale de una red. Intenta impedir la diseminación de daño a través suyo con el firewall podemos definir tamaños de paquetes, ips con la que se permite comunicación de esa habilitación de envíos o recogida de paquetes para determinados puertos por ejemplo si una terminal de red intenta enviar un paquete a una dirección ip no autorizada el rechazará este envío impidiendo realizar esta transmisión un proxy es un sistema de software que controla el flujo de datos sólo hacia el exterior.

La seguridad en internet se basa en el uso de protocolos:

SSL(Secure Sockets Layer) que traducido al español significa capa de conexiones seguras es un protocolo que hace uso de certificados digitales para establecer comunicaciones seguras a través de internet recientemente ha sido sustituido por TLS (Transport Layer Security) que está basado en SSL y son totalmente compatible

HTTPS es una combinación del protocolo HTTP usado en cada transacción web con el protocolo SSL/TLS usado para establecer comunicaciones cifradas en sitios web prácticamente todos los principales servicios de internet cifran sus datos desde Gmail hasta Facebook o Twitter incluyendo sistemas operativos como Android en este diagrama podemos ver la secuencia de pasos del algoritmo HTTPS en Facebook en el punto uno cuando el navegador hace una petición al sitio seguro de Facebook este envía un mensaje donde indica que quiere establecer una conexión segura y envía datos sobre la versión del protocolo SSL/TLS otros parámetros necesarios para la conexión en base a esta información enviada por el navegador el servidor web de Facebook responde con un mensaje informando que está de acuerdo en establecer la conexión segura con los datos SSL/TLS proporcionados una vez que ambos conocen los parámetros de conexión el sitio de Facebook presenta su certificado digital para el navegador web para identificarse como un sitio confiable una vez que el navegador tiene el certificado del sitio web de Facebook realiza algunas verificaciones antes de confiar en el sitio por ejemplo integridad del certificado verifica que el certificado se encuentre se encuentre íntegro esto lo hace descifrando la firma digital incluida en él mediante la clave pública y comparándola con una firma del certificado generado en ese momento si ambas son iguales entonces el certificado es válido la vigencia del certificado revisa el periodo de validez del certificado es decir la fecha de emisión y la fecha de expiración que están incluidos en él es verifica el emisor del certificado haciendo uso de una lista de certificados almacenados y que contiene las claves públicas conocidas y de confianza cómo hacía esta lista el navegador revisa que el certificado sea de confianza de no serlo el navegador mostrará una advertencia indicando que el certificado fue emitido por una entidad en la cual no se confía una vez que el certificado cumplió con todas las pruebas del navegador se establece la conexión segura al sitio web