Regulatory Compliance and Database Administration

Una de las regulaciones gubernamentales más visibles es la Ley Sarbanes-Oxley (SOX), conocida como la Ley de Protección de Inversores y Reforma Contable de las Empresas Públicas de USA de 2002. SOX regula las corporaciones para reducir el fraude y los conflictos de interés, mejora la divulgación y la información financiera, y fortalece la confianza en el público. Esta ley exige que los datos financieros de una empresa deben ser completamente recuperables.

Considere la Ley de Responsabilidad y Portabilidad del Seguro de Salud, o simplemente HIPAA, contiene un lenguaje que especifica que los proveedores de atención médica deben proteger la información de atención médica de las personas.

Sin el software de auditoría de bases de datos, es imposible producir una lista de usuarios que observaron una fila específica o un conjunto de filas en cualquier base de datos.

La Ley Gramm-Leach-Bliley (Ley GLB), conocida como la Ley de Modernización Financiera de 1999, es una ley federal promulgada en los Estados Unidos para controlar las formas en que las instituciones financieras manejan la información privada de las personas.

La Ley de Gobierno Electrónico se aprobó en 2002 como respuesta a las amenazas terroristas. El título III de esa ley se denomina Ley Federal de Administración de Seguridad de la Información (FISMA). FISMA básicamente establece que las agencias federales, los contratistas y cualquier entidad que los respalde deben mantener seguridad proporcional al riesgo potencial.

La regulación industrial más visible es ciertamente PCI DSS, que significa Estándar de Seguridad de Datos (DSS) de la Industria de Tarjetas de Pago (PCI).

El incumplimiento de estas leyes puede resultar en enjuiciamiento, lo que puede implicar enormes multas e incluso encarcelamiento. El cumplimiento normativo puede imponer a los ejecutivos de nivel C la necesidad de poder demostrar que los datos corporativos (y, por lo tanto, los sistemas de bases de datos) están protegidos y que los procesos y procedimientos promulgados sobre los datos son precisos y necesarios.

Un enfoque colaborativo para el cumplimiento

TI y las empresas deberían comunicarse regularmente, pero tal vez no se hacen de la manera más efectiva que podrían, pero los tres departamentos son obligatorios (Business, Legal, IT).

Las organizaciones necesitan mapear y clasificar sus datos comerciales de acuerdo con la forma en que cada elemento de datos se ve afectado por las regulaciones.

Una vez mapeados, deben promulgarse controles y políticas que hagan cumplir las normas pertinentes.

¿Por qué los DBA deben preocuparse por el Compliance?

El impacto principal del Compliance en el DBA está en investigar, instalar y administrar la tecnología que respalda el Compliance, en particular con respecto a los datos y el DBMS.

Las tareas relacionadas con el Compliance que afectan la administración de la base de datos incluyen

• Gestión de metadatos y calidad de datos.

• Auditoría de bases de datos y acceso a datos.

• Enmascaramiento de datos y obfuscation

• Retención de datos a largo plazo y archivo de la base de datos.

• Seguimiento más cercano de las tareas tradicionales de DBA (por ejemplo, gestión de cambios, respaldo y recuperación)

Gestión de metadatos, calidad de datos y gobernanza de datos

Garantizar la calidad de los datos es una gran parte del regulatory compliance.

Metadatos

¿Qué son los metadatos?

Los metadatos caracterizan los datos, proporcionando documentación de tal manera que los datos puedan ser entendidos y más fácilmente consumidos por una organización. Los metadatos responden las preguntas de quién, qué, cuándo, dónde, por qué y cómo para los usuarios de los datos.

Se requieren metadatos para colocar los datos en categorías apropiadas para determinar qué regulaciones se aplican.

Sin las definiciones de metadatos adecuadas, es imposible aplicar el regulatory compliance a los datos.

Calidad de datos

¿qué puede hacer un DBA sobre datos de baja calidad? La calidad de los datos es una responsabilidad empresarial, pero el DBA puede ayudar instalando controles tecnológicos. La creación de restricciones en la base de datos puede mejorar la calidad general de los datos, así como definir la integridad referencial en la base de datos.

Deben definirse restricciones adicionales en la base de datos según corresponda para controlar la unicidad, así como los rangos de valores de datos utilizando restricciones y triggers CHECK.

Táctica tecnológica para mejorar la calidad de los datos, creación de perfiles de datos. La creación de perfiles de datos permite examinar los datos existentes en la base de datos y recopilar estadísticas y otra información. Con la creación de perfiles de datos, puede descubrir la calidad, las características y los posibles problemas de la información. Utilizando las estadísticas recopiladas, los analistas de negocios pueden limpiar datos problemáticos en la base de datos.

Gobernanza de datos

Un programa de gobernanza de datos supervisa la gestión de la disponibilidad, usabilidad, integridad y seguridad de los datos empresariales. Un programa sólido de gobernanza de datos incluye un órgano o consejo de gobierno, un conjunto definido de procedimientos y un plan para ejecutar esos procedimientos.

Cuando la gestión de datos se instituye como un mandato oficialmente sancionado de una organización, los datos se tratan como un activo. Eso significa que los elementos de datos se definen en términos comerciales; se asignan administradores de datos; los datos se modelan y analizan; los metadatos se definen, capturan y administran; y los datos se archivan para la retención de datos a largo plazo.

Auditoría de bases de datos y seguimiento de acceso a datos

El aumento de las regulaciones exige la implementación de políticas y procedimientos para proteger información delicada. Cuando los datos confidenciales se almacenan en un DBMS, se debe prestar especial atención a cómo se gobiernan los datos en esas bases de datos.

La auditoría de la base de datos es el proceso de monitorear el acceso y la modificación de los objetos y recursos seleccionados de la base de datos.

Esto se puede lograr utilizando una instalación de auditoría de base de datos. La Tabla 15.1 ofrece una visión general de los requisitos de auditoría de varias reglamentaciones representativas.

Table 15.1. Database Auditing Requirements of Several Regulations

Audit Requirement	SOX	PCI DSS	GLB	HIPAA
Access to sensitive data (SELECT)		X	X	X
Modification of sensitive data (INSERT, UPDATE, DELETE)	X			
Database changes/DDL (CREATE, ALTER, DROP)	X	X	X	X
Security authorizations/DCL (GRANT, REVOKE)	X	X	X	X
Security exceptions (e.g., failed logins)	X	X	X	X

Los datos corporativos confidenciales no pueden protegerse completamente simplemente configurando la autorización de la base de datos utilizando los controles dentro del software.

Todos los principales productos DBMS ofrecen capacidades integradas para auditar bases de datos, pero los ISV ofrecen software más capaz con tecnología de captura más flexible, informes de cumplimiento preempaquetados y soporte multi-DBMS.

Por lo tanto, la mayoría de las instalaciones de auditoría permiten la creación selectiva de registros de auditoría para minimizar los problemas de rendimiento y de almacenamiento.

Los DBMS ofrece diferentes capacidades de auditoría, pero algunos elementos comunes que pueden ser auditados por las instalaciones de auditoría de DBMS son:

• Intentos de inicio de sesión y cierre de sesión (exitosos y no exitosos)

• El servidor de la base de datos se reinicia

• Comandos emitidos por usuarios con privilegios de administrador del sistema

• Intentos de violaciones de integridad (donde los datos modificados o insertados no coinciden con una restricción referencial, única o de verificación)

• Operaciones SELECT, INSERT, UPDATE y DELETE

• Ejecuciones de procedimientos almacenados

• Intentos fallidos de acceder a una base de datos o una tabla (fallas de autorización)

• Cambios en las tablas del catálogo del sistema.

• Operaciones a nivel de fila

La auditoría también se puede utilizar para la recuperación de datos.

Si ha activado la auditoría de bases de datos en su sitio, tenga en cuenta los siguientes consejos:

- La auditoría puede ser un gran consumidor de recursos del sistema.

- Coloque las tablas del catálogo del sistema que almacenan información relacionada con la seguridad en un disco separado e inactivo

- Asegúrese de que el conjunto de datos o la tabla utilizada para almacenar datos de auditoría no se llenen

Técnicas de auditoría de bases de datos

La mejor técnica involucra el monitoreo proactivo de las operaciones de la base de datos directamente en el servidor de la base de datos, esta técnica captura todas las solicitudes de datos a medida que se realizan. Cuando los detalles de la auditoría se capturan a nivel del servidor, el software puede garantizar que se supervise todo el acceso.

Existen cinco técnicas para implementar la auditoría de la base de datos: rastreos (traces), escaneo de registros (log scanning), columnas de auditoría (audit columns), rastreo de paquetes de red (network packet sniffing) y tapping the database server.

Las técnicas como la auditoría basada en rastreo o el análisis de registros de bases de datos pueden perder ciertos tipos de actividades de la base de datos, si una transacción accede a una tabla más de una vez en una sola unidad de recuperación, la traza de auditoría registra solo el primer acceso.

La auditoría basada en registros también es problemática. Debido a que dicha auditoría se basa en el registro de la base de datos, las únicas actividades que se pueden capturar son actualizaciones, inserciones y eliminaciones.

Algunas organizaciones que ahorran dinero intentan implementar la auditoría simplemente agregando columnas de "auditoría" a las tablas, como last_modified_date. Pero este enfoque no es aconsejable, y a los auditores no les gusta porque:

• Las audit trails deben mantenerse fuera de la base de datos

• Si elimina la fila, pierde los datos de auditoría.

• No puede garantizar que los últimos datos modificados sean completamente exactos porque el enfoque se basa en código que puede tener errores, o un programa de actualización simplemente puede pasarse por alto y no modificar la columna last_modified_date

• Alguien puede modificar la columna de auditoría por accidente (o nefastamente)

En rastreo de red, si la actividad no atraviesa la red, un sniffer de paquetes no puede rastrear las acciones.

Hacer tapping en las solicitudes directamente en el servidor es el único método infalible de monitoreo de la actividad de la base de datos que garantiza que todas las acciones auditables quedan atrapadas independientemente de la plataforma, el DBMS y la actividad.

Una solución de auditoría de base de datos debe ofrecer vigilancia de servidor selectiva, integral y no invasiva. Selectivo significa que la solución debe basarse en reglas para permitir la captura de detalles de auditoría solo en los datos específicos que requieren auditoría. Completo significa que la solución debe ser capaz de capturar el escenario completo de información auditable. Y no invasivo significa que el software de auditoría de la base de datos debe poder auditar el acceso a los datos sin incurrir en una costosa degradación del rendimiento.

Auditoría de usuario privilegiado

Los DBA y los administradores del sistema generalmente tienen autorización de alto nivel, como los privilegios DBADM o SYSADM en la base de datos.

La implementación de una auditoría de usuarios privilegiados para rastrear cada acción realizada por dichos usuarios es un curso de acción prudente.

Enmascaramiento de datos y Obfuscation

El enmascaramiento de datos es el proceso de proteger información confidencial en bases de datos que no son de producción ante alguna. Los datos de producción válidos se reemplazan por datos utilizables, referencialmente intactos, pero incorrectos u ofuscados.

Después del enmascaramiento, los datos de prueba se pueden usar al igual que los datos de producción, pero el contenido de la información es seguro. En otras palabras, los datos son precisos en la base de datos, pero enmascarados en la recuperación y visualización.

What Is PII?

Personally Identifiable Information (PII) es la información que se usa para idetificar de manera única, para localizar y contactar a una persona, por ello debe ser protegida, ejemplos de ella pueden ser cédula, placa de carro, fecha de nacimiento, etc.

Técnicas para Data Masking (Enmascaramiento):

Sustitucion: reemplaza data existente con valores random de un data set preparado.
Shuffling: usa la data existente y mueve los valores entre filas de tal forma que no presenta valores en las filas originales.
Varianza de número y fecha: varía los valores existentes en un rango especificado para ofuscarlos. Por ejemplo, los valores de la fecha de nacimiento podrían modificarse dentro de un rango de más o menos 90 días.
Encryption (cifrado): revuelve la data algorítmicamente, no produce data que se ve realística y la puede hacer más grande.
Mulling Out(reflexionando): simplemente quita la data sensible eliminándola.
Sincronizacion tabla a tabla: enmascara datos que aseguran que los resultados sean referencialmente intacto. Si dos tablas tienen columnas con los mismos valores y esas columnas están enmascaradas en una tabla, la segunda es actualizada también.

Si los datos han sido enmascarados no importa quién tenga acceso a ellos porque no es útil.

Database Archiving for Long-Term Data Retention(Archivo de BD para la retencion de datos a largo plazo)
Actualmente se guarda por más tiempo la data, esto porque es necesario realizar procesos analíticos.

El ciclo de vida de la data:

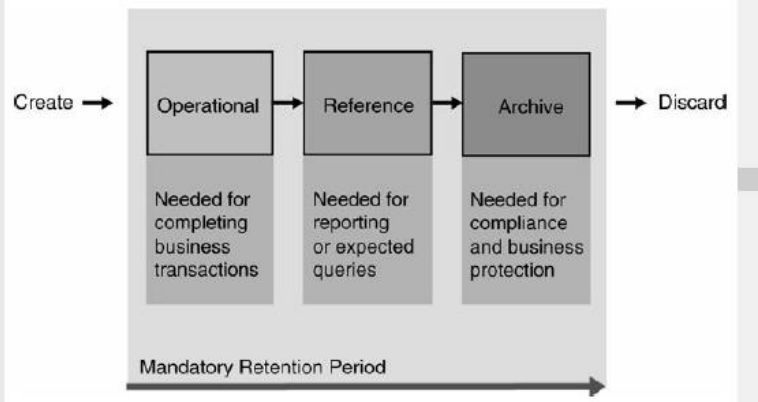

Figure 15.1. The data life cycle

La data primero es creada por una transacción, un periodo de tiempo después entra en el estado operacional, después va al estado de referencia, en este momento todavía se necesita para propósitos internos o externos o que simplemente un cliente la necesite, luego de un tiempo, la data no se necesita para propósitos de negocios y ya no está en cola, pero se debe guardar por propósitos legales, es por ello que se queda en el estado de archivo por último se descarta.

Database Archiving:

Se refiere al proceso de quitar (remove) los registros de data específicamente seleccionados de BDS operacionales que no se necesitan referenciar de nuevo y guardarlas en un archivo de data almacenado estos se almacenan por separado de la BD operativa y no requiere ni al DBMS ni las apps, aun así debe estar disponible cuando se necesite. El archivo debe poder elegir selectivamente piezas particulares de datos relacionados para archivar: ni la base de datos completa, ni una tabla completa, ni siquiera una fila específica. En cambio, todos los datos que representan un objeto comercial se archivan al mismo tiempo.

Archive versus Purge

Los datos **archivados** se eliminan desde el almacén de datos operativos y mantenido en un almacén de datos archivado. Los datos **purgados** son eliminados del almacén de datos operativos y descartados.
Dentro de las empresas se debe contar con políticas para saber qué datos deben ser archivados, ya que si no se cuenta con las reglas necesarias, podría guardarse data innecesaria o bien, por el contrario, se deja por fuera data importante para realizar análisis de la misma.

Database Archiving Requirements
La consideración más importante que se debe tomar en cuenta es que la data debe ser independiente de hardware y software, esto porque si la forma en que se guardó la data fue mucho tiempo atrás, los sistemas de producción que existían antes probablemente o ya no existan o han sido modificados con el paso del tiempo, además las BDS están en constante cambio es por ello que el archivo debe ser capaz de adaptarse a ello.
Además, la solución de archivo debe poder almacenar una gran cantidad de datos además de poder manejarlos por gran cantidad de tiempo(the archived data will outlive the systems and the programmers who generated them and also will outlive the media we store it on), ya que el hardware siempre puede llegar a ser obsoleto, debe ser capaz de ser cambiado de un dispositivo de almacenamiento a otro.
Debe permanecer incambiable, es por esto que el archivo no debe ser capaz de ser cambiado, debe tener only read access.
Finalmente, el archivo requiere que la metadata sea útil.

Components of a Database Archiving Solution

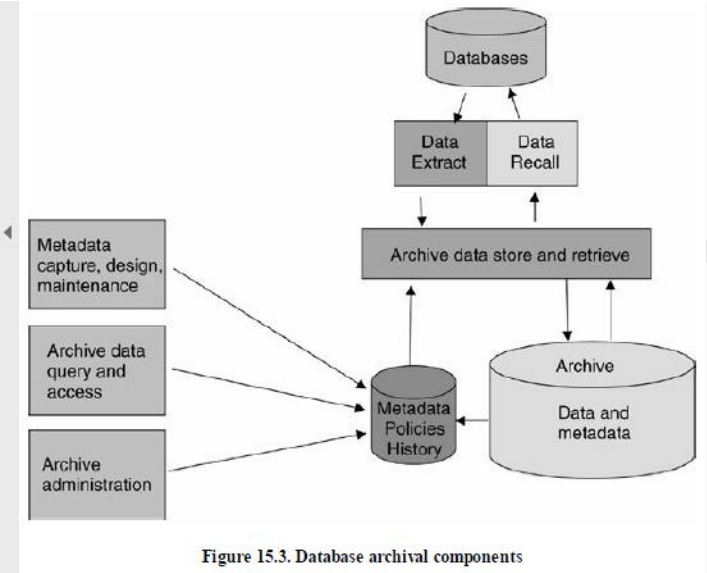

Figure 15.3. Database archival components

Existen dos operaciones que pueden hacer las BD, Data Extract que es la que elimina la data de la BD y Data Recall que restaura la data archivada a la BD.
Todo proceso de archivo necesita metadata para funcionar, por ello debe capturarla, validarla y mejorarla, se necesita conocer la estructura de la BD y la estructura del archivo, además una solución de archivo debe ser basada en políticas de la empresa que determinen cual es la data, cuándo y durante cuánto tiempo debe permanecer antes de descartarse.

También es importante una capacidad de consulta que permita lecturas directas en el archivo, aunque generalmente no son sensibles al performance, finalmente, también es importante tener una capacidad de mantenimiento continuo para los datos archivados. Esto abarca tareas administrativas como seguridad, auditoría de acceso, etc.

Closer Tracking of Traditional DBA Tasks

Database Change Management
Con frecuencia, los DBA simplemente emiten comandos ALTER para cambiar la base de datos en reacción a problemas de performance, sería una práctica mucho mejor requerir que se realicen cambios en la base de datos utilizando una herramienta de administración de cambios los rastrea automáticamente.

Database Backup and Recovery

Las prácticas y procedimientos mejorados de backup y recovery deben ser un componente esencial, asegurar la integridad y disponibilidad de las BDS es el enfoque principal de la

copia de seguridad y planificación de recuperación. Organizations need tools and procedures that help to verify the integrity of database backups.

What Is COBIT?

Es un marco de las mejores prácticas de TI que las empresas pueden usar para mejorar la gestión de sus organizaciones de TI, para mejorar el valor y para garantizar que los objetivos de TI la organización está alineada con los objetivos del negocio.

COBIT and Recovery

La recuperación de la base de datos debe abordarse desde un enfoque de mejores prácticas para permitir que su organización

para hacer el tipo de planificación inicial y monitoreo y evaluación de rutina, aquella organización que tenga COBIT como practica de framework entiende el valor critico de la información y asegura la integridad y disponibilidad de la misma.