

Contenido

¿Qué es un DBA?	3
¿Por qué aprender Administración de Bases de Datos?	4
Un punto de vista único	5
Salarios de un DBA	5
Bases de Datos vs DBMS	6
La disciplina administrativa de la administración de bases de datos	7
Políticas de privacidad y datos	8
Un día en la vida de un DBA	9
Evaluación de una oferta de trabajo DBA	11
Administración de Bases de Datos, de Datos y del Sistema	11
Administración de Datos	12
Administración de Bases de Datos	15
Administración del Sistema	15
Tareas del DBA	16
Diseño de Bases de Datos	17
Ajuste y Monitoreo de Rendimiento	17
Aseguramiento de la Disponibilidad	19
Seguridad y Autorización de la Base de Datos	19
Centralización de la Seguridad	20
Gobernanza y Cumplimiento Normativo	20
Respaldo y Recuperación	21
Aseguramiento de la Integridad de los Datos	22
Migración de versión del DBMS	24
Un mil usos	24
Tipos de DBAs	25
DBA del Sistema	25
Arquitecto de Bases de Datos	26
Analista de Bases de Datos	27
Modelador de Datos	27
DBA de Aplicación	27
DBA Orientado a Tareas	28

Analista de Rendimiento	29
Administrador de Data Warehouse	29
Consideraciones del Personal	30
¿Cuántos DBAs?	30
Estructuras de Reporte (Organigrama) de DBA	32
Cuestiones Multiplataforma para el DBA.....	33
Producción vs Pruebas (Desarrollo)	34
El Impacto de las Nuevas Tecnologías en el DBA.....	35
DBA Procedural: Manejo de la lógica de Bases de Datos.....	36
Procedimientos Almacenados.....	36
Triggers (Desencadenadores).....	37
Funciones Definidas por el Usuario (UDF)	37
El Internet: del DBA al eDBA	38
El DBA Personal y la Nube	39
¿Qué es el cómputo en la nube?.....	40
Impacto en los DBAs.....	41
NoSQL, Big Data y el DBA	42
Ejemplos de Ofertas de NoSQL	42
Impacto en los DBAs.....	42
Impacto de las Nuevas Tecnologías en el DBA.....	43
Certificación para DBA	43
Definir la Estrategia de DBMS para la Empresa	44
Selección de un DBMS.....	45
Arquitecturas de DBMS.....	48
Proveedores de DBMS.....	49
Los Tres Grandes	50
El Segundo Nivel.....	50
Otros Actores Importantes.....	51
Ofertas de DBMS de Código Abierto.....	52
Proveedores de DBMS No Relacionales.....	53
Proveedores de DBMS NoSQL.....	53
Proveedores de DBMS Orientados a Objetos	53
Proveedores de DBMS basados en PC	54

¿Qué es un DBA?

Cada organización que maneja datos utilizando un sistema manejador de bases de datos (DBMS) requiere un grupo de administración de bases de datos (DBA) para asegurar el uso y despliegue efectivos de las bases de datos de la compañía. La mayoría de las organizaciones modernas de todos los tamaños usan al menos un DBMS, debido a esto la necesidad por administradores de bases de datos (DBAs) es más grande hoy que nunca antes. Sin embargo la disciplina de la administración de las bases de datos no está muy bien entendida o universalmente practicada de una manera coherente y fácil de replicar.

La necesidad de un administrador de bases de datos es más grande hoy que nunca antes.

Existe una broma frecuentemente repetida acerca de la administración de bases de datos que nos ayuda a entender ambos: la necesidad por un DBA y la ausencia del entendimiento de la función de un DBA. Es algo como esto:

El Jefe de Informática (CIO) de la Corporación Acme contrata una empresa de consultoría de administración para optimizar su tecnología de la información (TI). El consultor, determinado a entender la forma de trabajo de Acme, inicia entrevistando al CIO. Una de sus primeras preguntas fue: “Veo que tiene un DBA en su equipo. ¿Qué es lo que hace?”

El CIO contestó, “Le digo que necesitamos el DBA para asegurar que nuestras bases de datos en Oracle esté en línea. Sé que algo de nuestro proceso crítico de negocio como la entrada de órdenes e inventarios usan Oracle, pero realmente no sé qué hace el DBA. Pero por favor no me diga que necesito otro, porque apenas nos alcanza a pagarle al que ya tenemos”.

Esto es un triste comentario, pero muy frecuentemente verdadero, del estado de la administración de bases de datos en muchas organizaciones. Frecuentemente el DBA es visto como un gurú o mago que utiliza trucos para hacer que las bases de datos y sistemas operen eficientemente. El software de DBMS es tan complejo en estos días que muy pocas personas comprenden más que solo lo básico (como SQL). Pero los DBAs entienden las complejidades del DBMS, haciéndolo un recurso muy valorado. Algunas veces el único recurso que se tiene con conocimientos de administración de bases de datos y desarrollo en la organización es el DBA.

El rol del DBA se puede interpretar como el guardián de los datos o como un activo corporativo. Entonces los DBA en el intento de proteger los datos, a menudo son percibidos con un movimiento lento adverso al riesgo. Los desarrolladores, por otro lado, son cargados con la construcción de nuevas aplicaciones y constantemente son desafiados con hacer cosas rápido y moverse al siguiente proyecto. Obviamente, la diferencia entre estos dos roles y expectativas, con el DBA diciendo “control de cambios, administración de cambios” y los desarrolladores diciendo “implementar ahora, implementar ahora”, puede crear cierta fricción.

Otra crítica frecuente de los DBA, es que pueden ser difíciles de tratar. A veces los DBAs pueden ser cascarrabias con vastos conocimientos técnicos, pero limitadas habilidades personales. Casi cada programador de bases de datos tiene su historia favorita con un DBA, esas famosas

anécdotas inician con “Tengo un problema ...” y terminan con “... y entonces me dijo que dejara de molestarlo y leyera el manual”. Los DBAs simplemente no tienen una imagen cálida.

El hecho de que los DBAs a menudo tienen que sentarse y resolver las cosas por su cuenta, puede ser un factor atenuante en esta mala reputación. Muchos problemas de bases de datos de periodos de análisis y reflexión para resolverlos. Así que a los DBAs por lo general no les agrada que se les moleste. Pero a pesar de que muchos problemas requerirán ser resueltos en la soledad, hay muchos otros que requieren del trabajo en equipo para resolverse. Debido al vasto conocimiento que los DBAs poseen, sus momentos de descanso por lo general no tienen tranquilidad; interrupciones constantes para responder preguntas y resolver problemas son un factor de todos los días.

Los DBAs no deben de ser alentados a ser antisociales. De hecho, los DBAs deben ser entrenados para adquirir habilidades excepcionales de comunicación. Los datos son el elemento vital de las aplicaciones informáticas. Las aplicaciones informáticas son desarrolladas para leer y escribir datos, analizar datos, mover datos, realizar cálculos utilizando datos, modificar datos, y así sucesivamente. Sin datos no habría nada que hacer para las aplicaciones informáticas. El DBA es el centro en el ciclo de vida del desarrollo, asegurando que las aplicaciones tengan un acceso eficiente y preciso a los datos de la corporación. Por lo tanto, los DBAs interactúan con muchos distintos tipos de personas: técnicos, programadores, usuarios finales, clientes y ejecutivos. Sin embargo, muchos DBAs están atrapados en las minucias de los mecanismos internos del DBMS, provocando que nunca lleguen a desarrollar las habilidades requeridas para relacionarse con sus compañeros de trabajo y clientes.

Los DBAs necesitan adquirir habilidades excepcionales de comunicación.

Pero hasta el momento no se ha respondido la pregunta que se hizo al inicio: ¿Qué es un DBA? La respuesta corta a esta pregunta es simple:

Un DBA es el técnico informático responsable de garantizar la funcionalidad y eficiencia operativa de las bases de datos y aplicaciones de una organización.

La respuesta larga a esta pregunta requiere de un libro completo para contestarse.

¿Por qué aprender Administración de Bases de Datos?

Como ya se ha mencionado, los datos son el centro de las aplicaciones de hoy en día, y la organización moderna de hoy simplemente no puede trabajar sin datos. De muchas formas, los negocios de hoy son los datos. Sin datos, los negocios no tendrían la capacidad de manejar sus finanzas, realizar transacciones, o contactar a sus clientes. Las bases de datos fueron creadas para almacenar y organizar esos datos. Cuanto mejor sea el diseño y la utilidad de la base de datos, lo mejor de la organización será posicionado para competir por el negocio. De hecho, uno de los mayores problemas que enfrentan las organizaciones de IT es asegurar la calidad de la administración de bases de datos. De acuerdo a estudios recientes:

Los buenos DBAs son difíciles de encontrar y costosos de contratar. El 70% de los encuestados dicen que se requieren más de tres meses para contratar un DBA, y entrenar a un nuevo DBA en un nuevo entorno toma varios meses más.

La infraestructura de bases de datos que soporta aplicaciones es compleja, crónicamente fragmentada, y difícil de manejar.

Estos dos hallazgos indican claramente que la administración de la base de datos es un trabajo difícil que tiene una alta demanda. Adicionalmente, de acuerdo a la encuesta de salarios de la tecnología Dice 2010-11, se requiere experiencia en Oracle en más de 15,000 puestos de empleo diarios. La demanda por habilidades de Oracle es de hasta 57% año tras año, y el promedio nacional en USA de salarios de profesionales en tecnología con experiencia en Oracle es de \$90,914 USD.

Un punto de vista único

El DBA es el responsable del diseño y mantenimiento de las bases de datos de una empresa, colocándose al DBA de lleno al centro de los negocios. El DBA tiene la oportunidad de aprender acerca de muchas facetas del negocio y como se relacionan entre sí. El DBA puede explorar tecnologías innovadoras como sean adoptadas por la organización. La exposición a las nuevas tecnologías mantiene estimulante el trabajo, pero frustrante si se está intentando averiguar como funciona una nueva tecnología por primera vez. Frecuentemente el DBA trabaja solo en sus esfuerzos; no tiene acceso a experiencia adicional para ayudar cuando surgen los problemas. Sin embargo,

Un buen DBA necesita disfrutar los retos y ser un buen solucionador de problemas.

Salarios de un DBA

Como un técnico no se puede encontrar un empleo más retador en IT que el de DBA. Afortunadamente, el trabajo de DBA también es muy gratificante. Los DBAs son bien pagados. De acuerdo al estudio de salarios realizado por Global Knowledge Tech-Republic, el salario promedio de un DBA es de \$78,468USD anuales en USA, mientras que el promedio de sus gerentes es de \$87,261USD anuales en USA.

Para los empleados de tiempo completo que trabajan como DBAs, los rangos de salarios rondan los \$80 mil USD anuales en USA.

Salarios anuales para DBAs entre 2006 y 2011 en USA	
Año	Salario DBA
2006-7	\$85,441 USD
2007-8	\$85,092 USD
2008-9	\$89,792 USD

2009-10	\$91,283 USD
2010-11	\$88,443 USD

Se debe tener en mente que los salarios que aparecen citados aquí son solo para propósitos ilustrativos y variarán basados en numerosos factores. Como era de esperarse, ya que los años de experiencia se incrementan, lo mismo ocurre con los salarios.

Quizás más importante que el salario es la empleabilidad, y las perspectivas de empleo como DBA son bastante buenas.

El mercado de trabajo para los DBAs crecerá mucho más rápido que el promedio.

Entonces, los DBAs son bien pagados, altamente empleables, poseen empleos retadores, y es probable que se dediquen a los proyectos más visibles e importantes. ¿Qué más se puede pedir? Se espera que los DBAs lo sepan todo, no solo acerca de tecnología de bases de datos, sino de todo lo que esté apenas conectado con ella. Los DBAs casi nunca trabajan solo ocho horas diarias, en vez de eso trabajan largas jornadas con mucho tiempo extra, especialmente cuando el rendimiento está sufriendo o los proyectos de desarrollo se retrasan. De acuerdo a analistas de la industria, los DBAs trabajan en promedio más de 60 horas a la semana. Los DBAs frecuentemente tienen que trabajar en fines de semana y días festivos para mantener las bases de datos durante las horas no pico.

La administración de bases de datos es un trabajo sin parar.

El trabajo de DBA es técnicamente desafiante y gratificante, pero también potencialmente, agotador y frustrante. Pero no deje que eso le asuste. Los aspectos positivos del trabajo son mucho mayores que los negativos.

Bases de Datos vs DBMS

Se asume que se conocen los fundamentos de bases de datos, esto no es algo trivial, pero a veces el problema es que se piensa que se sabe más de lo que actualmente se hace. Por ejemplo "¿Qué es una base de datos?". Se puede apostar que la mayoría de las personas que leen esto creen que saben la respuesta a esa pregunta. Pero algunos (quizás muchos) de ustedes estarían equivocados. SQL Server no es una base de datos; Es un sistema de gestión de bases de datos o DBMS. Se puede usar SQL Server para crear una base de datos, pero SQL Server, en sí mismo, no es una base de datos.

Entonces, ¿qué es una base de datos? Una base de datos es un almacén de datos organizado en el que los datos son accesibles por elementos de datos nombrados (por ejemplo, campos, registros y archivos).

Un DBMS es un software que permite a los usuarios finales o programadores de aplicaciones compartir datos. Proporciona un método sistemático para crear, actualizar, recuperar y almacenar información en una base de datos. Los DBMS también son generalmente responsables de la integridad de los datos, la seguridad de los datos, el control y la optimización del acceso a los datos, la reversión automática, el reinicio y la recuperación.

En términos comunes, se puede pensar en una base de datos como una carpeta de archivos. Puede pensar en el archivador que contiene los archivos junto con las etiquetas de los archivos como el DBMS. Un DBMS gestiona las bases de datos. Usted implementa y accede a las instancias de base de datos utilizando las capacidades del DBMS. Entonces, DB2, Oracle y SQL Server son sistemas de administración de bases de datos. "Su aplicación de nómina utiliza la base de datos de nómina, que puede implementarse utilizando DB2, Oracle o SQL Server.

¿Por qué es eso importante? Si no utilizamos términos precisos en el lugar de trabajo, puede producirse confusión. Y la confusión lleva a proyectos de presupuesto excesivo, sistemas mal desarrollados y pérdida de productividad.

Además de los fundamentos de la administración de bases de datos, los DBA deben ser expertos en los productos DBMS específicos que se están utilizando, y puede haber muchos en la organización. Por ejemplo, una organización grande puede usar DB2 en el mainframe, Oracle e Informix en varias plataformas UNIX diferentes, MySQL en Linux y SQL Server en Windows. Los sistemas heredados más antiguos pueden usar bases de datos IMS, y luego está esa aplicación loca que usa un DBMS marginal como Adabas o Ingres. Y también hay una nueva tecnología de base de datos, como NoSQL, y las ofertas de DBMS del almacén de columnas como Hadoop, como así como sistemas de bases de datos en la nube, como SQL Azure de Microsoft y BigTable de Google.

El grupo de DBA, por lo tanto, debe tener experiencia en cada uno de estos diferentes DBMS y plataformas. Además, el DBA debe ser capaz de determinar qué DBMS y plataforma son los más adecuados para las necesidades de cada aplicación. Este puede ser un trabajo difícil, cargado de política y opiniones conflictivas. El grupo de DBA debe poder actuar como un juez imparcial y basar las decisiones de implementación en las necesidades de la aplicación en comparación con las capacidades y especificaciones de cada DBMS y plataforma.

Los DBA deben implementar decisiones basadas en la mejor adaptación de la aplicación, el DBMS y la plataforma.

La disciplina administrativa de la administración de bases de datos

La administración de bases de datos rara vez se aborda como una disciplina de gestión. El término disciplina implica planificación e implementación de acuerdo a ese plan. Cuando la administración de la base de datos se trata como una disciplina de administración, el tratamiento de los datos dentro de su organización mejorará. Es la diferencia entre ser reactivo y proactivo.

Con demasiada frecuencia, el grupo de DBA está abrumado por solicitudes y problemas. Esto se debe a muchas razones, entre ellas, falta de personal, exceso de compromiso para respaldar proyectos de desarrollo de aplicaciones nuevas (e incluso antiguas), falta de procesos repetibles, falta de presupuesto, y así sucesivamente. Cuando se opera de esta manera, el administrador de la base de datos está siendo reactivo. El DBA reactivo funciona más como un bombero. Su atención se centra en resolver el problema más grande que se le señala. En otras palabras, un DBA reactivo intenta resolver problemas solo después de que surjan los problemas.

Un DBA proactivo puede evitar muchos problemas al desarrollar e implementar un plan estratégico a seguir al implementar bases de datos dentro de su organización. Este plan debe abordar todas las fases del ciclo de vida del desarrollo de la aplicación (ADLC). El DBA proactivo implementa prácticas y procedimientos para evitar problemas antes de que ocurran.

Un DBA proactivo puede evitar muchos problemas.

Un especialista en datos, generalmente el DBA, debe participar en cada fase del ADLC. Durante la fase de inicio y recopilación de requisitos del proyecto, el DBA debe estar disponible para identificar los componentes de datos del proyecto. Él puede ayudar a determinar si los datos requeridos ya existen en otras partes de la organización o si los datos son nuevos. Durante las fases de análisis y diseño, los requisitos de datos rudimentarios deben transformarse en un modelo de datos conceptual y lógico.

Antes de que pueda comenzar el desarrollo, el modelo de datos lógicos debe traducirse a un diseño físico de base de datos que se pueda implementar utilizando un DBMS, como Oracle o DB2. Los datos de muestra se deben completar en la base de datos física para ayudar con las pruebas de la aplicación. Además, el DBA debe desarrollar e implementar un proceso para actualizar los datos de prueba para permitir ejecuciones de prueba repetibles.

Cuando la aplicación pasa del estado de desarrollo al estado operativo, el DBA debe asegurarse de que el DBMS esté preparado para la nueva carga de trabajo. Esta preparación incluye la implementación de medidas de seguridad adecuadas, la medición y modificación de los requisitos de almacenamiento y memoria para la nueva aplicación, y anticipar el impacto de la nueva carga de trabajo en las bases de datos y aplicaciones existentes. El DBA también es responsable de migrar la nueva base de datos del entorno de prueba a la producción.

Mientras la aplicación está operativa, el DBA realiza una gran cantidad de tareas, incluida la supervisión del rendimiento y la disponibilidad, el ajuste, la copia de seguridad y la recuperación, y la administración de la autorización. Pero ninguna aplicación o base de datos permanece estática por mucho tiempo. Debido a que las necesidades del negocio cambiarán, los sistemas de TI que soportan el negocio también cambiarán. Cuando se solicita el mantenimiento, el DBA se involucra en todo el proceso una vez más, desde los requisitos que se acumulan hasta que los cambios se vuelven operativos. Cuando el monitoreo revela una deficiencia en el rendimiento, el DBA puede sugerir alternativas de mejor desempeño para el equipo de desarrollo (si el problema está en el código de la aplicación) o para el equipo de almacenamiento (si el cuello de botella está centrado en la E / S) o recomendar otras soluciones tales como como índices nuevos o mejorados, código SQL alternativo, etc.

Políticas de privacidad y datos

La privacidad de los datos es un problema creciente que se está convirtiendo cada vez más en una carga para las organizaciones, ya que no solo aumenta la cantidad de datos bajo administración, sino que también aumenta la velocidad de llegada de nuevos datos.

En este mundo en red, en el que estamos completamente digitalizados, con nuestras identidades, ubicaciones, acciones, compras, asociaciones, movimientos e historias almacenadas como tantos bits y bytes, tenemos que preguntarnos: ¿Quién está recolectando todo esto? ¿Qué están haciendo con eso? ¿Con quién lo comparten? Más que nada, las personas preguntan: "¿Cómo puedo proteger mi información para que no sea utilizada indebidamente?" Estas son preguntas razonables que se deben hacer; Todos deberíamos querer saber las respuestas.

Existen estatutos federales y estatutos estatales que rigen la privacidad de la información en el país. El enfoque para proteger la privacidad ha sido poco sistemático. La Unión Europea, por otro lado, ha adoptado una Directiva de protección de datos que requiere que sus países miembros adopten leyes que implementen sus términos. La directiva crea derechos para las personas acerca de las cuales se recopila información, conocidos como "sujetos de datos". Las entidades que recopilan información deben notificar a los interesados que explican quién recopila los datos, quién tendrá acceso a ellos y por qué se recopilan los datos. Los sujetos de datos también tienen derecho a acceder y corregir los datos sobre ellos ".

Incluso hay un Día de la Privacidad de los Datos, cuyo objetivo es celebrar la dignidad de la persona expresada a través de información personal.

Finalmente, cuando la aplicación llega al final de su vida útil, el DBA debe ayudar a determinar el estado final de los datos utilizados por la aplicación. ¿Ya no se requieren los datos o otras aplicaciones y procesos también usan los datos? ¿Hay regulaciones que requieren que los datos se almacenen por más tiempo que la aplicación? ¿Tiene la empresa alguna política de privacidad establecida que imponga reglas especiales para el manejo de los datos?

El DBA es responsable de administrar el entorno de base de datos general. A menudo, esto incluye instalar el DBMS y configurar la infraestructura de TI para permitir que las aplicaciones accedan a las bases de datos. Estas tareas deben completarse antes de que se pueda desarrollar cualquier programa de aplicación. Además, el acceso a bases de datos ad hoc es un requisito para muchas organizaciones.

Además, el DBA se encarga de configurar el entorno de consultas ad hoc, que incluye la evaluación e implementación de las herramientas de consulta e informes, el establecimiento de políticas y procedimientos para garantizar consultas ad hoc eficientes y el monitoreo y ajuste de SQL ad hoc.

Como se puede ver, un buen DBA es parte integral del ciclo de vida completo del desarrollo de la aplicación. El DBA está "en demanda" por su conocimiento de los datos y la forma en que los datos son gestionados por las aplicaciones modernas.

Un buen DBA es parte integral del ciclo de vida completo del desarrollo de la aplicación.

Un día en la vida de un DBA

Un día en la vida de un DBA suele ser bastante ocupado. El DBA debe mantener los entornos de producción y prueba, al mismo tiempo que vigila los proyectos de desarrollo de aplicaciones activas, asiste a reuniones de estrategia y diseño, ayuda a seleccionar y evaluar nuevos productos y conecta sistemas heredados a la Web. Y Joe en contabilidad, acaba de presentar esa "consulta

del infierno" de nuevo que está deteniendo el sistema; ¿Puedes hacer algo al respecto? Todas estas cosas pueden ocurrir dentro de una sola jornada laboral de DBA.

Para sumarse al caos, se espera que los DBA sepan todo sobre todo. Desde la jerga técnica y comercial hasta las últimas modas de gestión y tecnología, se espera que el DBA esté "al tanto". Y no espere ningún tiempo privado. Un DBA siempre debe estar preparado para ser interrumpido en cualquier momento para responder a cualquier tipo de pregunta, y no solo sobre bases de datos.

Cuando se producen problemas con la aplicación, el entorno de la base de datos suele ser el primer culpable. La base de datos es "culpable hasta que se pruebe que es inocente", y no al revés.

Un DBA nunca experimentará que un desarrollador de aplicaciones acuda a él en busca de ayuda con una pregunta como: "Tengo un SQL realmente malo aquí; ¿Puede ayudarme a solucionarlo?" "No, en lugar de eso, el desarrollador irrumpe en el cubículo del DBA diciendo en voz muy alta que" hay un problema con DB2 [o inserte su DBMS favorito aquí]; ¿Por qué no lo arreglas para que mi maravilloso programa pueda ejecutarse?

Por lo tanto, el DBA se ve obligado a probar que la base de datos no es la fuente del problema. Debe saber lo suficiente sobre todos los aspectos de la tecnología de la información para localizar errores y exonerar lo que está en su ámbito: el DBMS y las estructuras de base de datos que ha diseñado. Por lo tanto, debe ser un experto en tecnología de bases de datos, pero también conocer los componentes de TI con los que interactúa el DBMS: lenguajes de programación de aplicaciones, sistemas operativos, protocolos y productos de red, procesadores de transacciones, todo tipo de hardware de computadora imaginable, y más. La necesidad de comprender temas tan diversos hace que el DBA sea un recurso muy valioso. También hace que el trabajo sea interesante y desafiante.

Y el trabajo de DBA es un trabajo continuo. Un DBA debe estar constantemente disponible para hacer frente a los problemas, ya que las aplicaciones de base de datos se ejecutan durante todo el día.

La mayoría de los DBA llevan un buscapersonas o un teléfono móvil en todo momento para que puedan contactarse en cualquier momento. Si hay un problema con la base de datos a las 2:00 am, el DBA debe levantarse de la cama, despejarse la cabeza y resolver el problema para que las aplicaciones vuelvan a estar operativas. De lo contrario, se puede producir un tiempo de inactividad de la base de datos, y eso puede cerrar completamente los procesos empresariales.

Los DBA con frecuencia pasan los fines de semana frente a la computadora realizando tareas de mantenimiento de bases de datos y reorganizaciones durante las horas de menor actividad. No puede derribar las bases de datos de misión crítica durante los nueve a cinco días para mantenerlas. Y, francamente, cada vez más empresas tienen días de 24 horas en lugar de los míticos días hábiles de ocho horas de antaño.

Si esto todavía te suena intrigante, sigue leyendo. Y en realidad, no es tan malo como parece. El trabajo es interesante, siempre hay algo nuevo que aprender y, como se mencionó anteriormente, la paga puede ser buena. La única pregunta es: ¿Puede alguien hacer este tipo de trabajo durante

20 años o más sin necesidad de descansar? Y, por cierto, creo que escucho el zumbido de su teléfono móvil, por lo que es posible que desee hacer una pausa aquí para ver qué pasa.

Evaluación de una oferta de trabajo DBA

Como DBA, es casi inevitable que cambie de trabajo varias veces durante su carrera. Al realizar un cambio de trabajo, obviamente tendrá en cuenta requisitos como salario, bonificaciones, beneficios, frecuencia de revisiones y tiempo de vacaciones. Sin embargo, también debe considerar cómo trata la compañía a sus DBA. Diferentes organizaciones dan diferente valor al trabajo de DBA. Es imperativo para el desarrollo de su carrera que busque organizaciones progresistas que comprendan la complejidad y los requisitos de aprendizaje continuo del puesto.

Aquí hay algunas preguntas útiles para hacer:

¿Ofrece la compañía capacitación regular para que sus DBA aprendan nuevas funciones y funciones de DBMS? ¿Qué sucede con la capacitación para tecnologías relacionadas, como programación, redes, comercio electrónico, gestión de transacciones, colas de mensajes y similares?

¿Permite la compañía que los DBA asistan regularmente a grupos de usuarios locales? ¿Qué sucede con las visitas anuales a grupos de usuarios en ubicaciones remotas?

¿Hay DBA de respaldo o será el único que esté disponible las 24 horas del día, los 7 días de la semana?

¿Existen organizaciones de administración de datos y administración de sistemas, o se espera que los administradores de bases de datos realicen todos estos deberes?

¿El grupo de DBA ve su relación con los grupos de desarrollo de aplicaciones como una asociación? ¿O es la relación más antagónica?

¿Se incluyen los DBA en las revisiones de diseño, las discusiones sobre presupuestos y otros comités y funciones de alto nivel de TI?

Cuantas más respuestas de “sí” recibas a estas preguntas, más progresivo será el entorno DBA.

Administración de Bases de Datos, de Datos y del Sistema

Algunas organizaciones definen roles separados para los aspectos comerciales de los datos y los aspectos técnicos de los datos. Los aspectos comerciales de los datos están alineados con una disciplina conocida como administración de datos, mientras que los aspectos más técnicos son manejados por la administración de la base de datos. No todas las organizaciones tienen una función de administración de datos. De hecho, muchas organizaciones combinan la administración de datos en el rol de administración de la base de datos.

Muchas organizaciones combinan la administración de datos en el rol de administración de base de datos.

Algunas veces, las organizaciones también dividen los aspectos técnicos de la administración de datos, ya que el DBA es responsable de usar el DBMS y otra función, conocida como administración del sistema o programación de sistemas, es responsable de instalar y actualizar el DBMS.

Administración de Datos

La administración de datos (DA) separa los aspectos comerciales de la administración de recursos de datos de la tecnología utilizada para administrar los datos. Cuando la función DA existe en una organización, está más estrechamente alineada con los usuarios comerciales reales de los datos. El grupo DA es responsable de comprender el léxico empresarial y traducirlo en un modelo de datos lógico. Con referencia al ciclo de vida del desarrollo de la aplicación, el administrador de datos (DA) estaría más involucrado en las fases de recopilación de requerimientos, análisis y diseño; El DBA, en las fases de diseño, desarrollo, prueba y operación.

Otro factor diferenciador entre DA y DBA es el enfoque de sus esfuerzos. El DA es responsable de cuestiones tales como

- Identificación y catalogación de los datos requeridos por los usuarios de negocios.

- Producción de modelos de datos conceptuales y lógicos para representar con precisión la relación entre los elementos de datos para los procesos de negocios.

- Producción de un modelo de datos empresariales que incorpora todos los datos utilizados por todos los procesos de negocios de la organización

- Configuración de políticas de datos para la organización.

- Identificación de propietarios de datos y administradores.

- Establecimiento de estándares de control y uso de datos.

En resumen, se puede considerar al DA como el Director de Datos de la corporación. Sin embargo, la posición de DA no tiene nada que ver específicamente con la tecnología. Sin embargo, en mi experiencia, el DA nunca tiene un puesto ejecutivo. Eso es realmente muy malo. Muchas organizaciones de TI afirman que tratan los datos como un activo corporativo, pero la falsedad de esa declaración se revela cuando revisa sus acciones. La responsabilidad de la política de datos a menudo queda relegada a los técnicos que no se concentran en los aspectos no técnicos y comerciales de la gestión de datos. Los técnicos hacen un buen trabajo para garantizar la disponibilidad, el rendimiento y la capacidad de recuperación, pero generalmente no son capaces de garantizar la calidad de los datos y establecer políticas corporativas.

De hecho, los datos rara vez se tratan como un verdadero activo corporativo. Piense en los activos que todas las empresas tienen en común: capital, recursos humanos, instalaciones y materiales. Cada uno de estos activos se modela: plan de cuentas, organigramas y jerarquías de informes,

planos de edificios y planos de oficinas, y listas de materiales. Cada uno es rastreado y protegido. Los auditores profesionales se emplean para garantizar que no existan discrepancias en nuestra contabilidad de los activos. ¿Podemos decir lo mismo sobre los datos en la mayoría de las organizaciones?

Una organización de DA madura es responsable de planificar y guiar los requisitos de uso de datos en toda la organización. Esta función abarca cómo se documentan, comparten e implementan los datos. Una gran responsabilidad del personal del DA es garantizar que los elementos de datos se documenten correctamente, generalmente en un diccionario de datos o repositorio. Esta es otra diferencia clave entre DA y DBA. El DA se enfoca en el repositorio, mientras que el DBA se enfoca en las bases de datos físicas y DBMS.

Además, el DA trata con los metadatos, a diferencia del DBA, que trata con los datos. Los metadatos se describen a menudo como "datos sobre datos"; más exactamente, los metadatos son la descripción de los datos y las interfaces de datos requeridas por la empresa. La administración de datos es responsable de la estrategia de metadatos de la empresa.

Los ejemplos de metadatos incluyen la definición de un elemento de datos, los nombres comerciales de un elemento de datos, las abreviaturas utilizadas para ese elemento y el tipo de datos y la longitud del elemento. Los datos sin metadatos son difíciles de usar. Por ejemplo, el número 12 son datos, pero ¿qué tipo de datos? En otras palabras, ¿qué significa eso 12? Sin metadatos no tenemos idea. Considerar:

¿Es una fecha que representa a diciembre, el duodécimo mes del año?

¿O es una fecha que representa el duodécimo día de algún mes?

¿Podría ser una edad?

¿Una talla de zapato?

¿O, Dios no lo quiera, un coeficiente intelectual?

Y así.

Pero también hay otros aspectos más técnicos de los metadatos. Piensa en nuestro número 12 otra vez. Considerar:

¿Es 12 un número grande o pequeño?

¿Cuál es su dominio (es decir, cuál es el universo de valores posibles de los cuales 12 es un solo valor)?

¿Qué pasa con el tipo de datos? ¿Es un número entero o decimal con una escala de 10?

Los metadatos proporcionan el contexto en el que los datos pueden entenderse y, por lo tanto, convertirse en información. En muchas organizaciones, los metadatos no son metódicamente capturados y catalogados; en cambio, existe principalmente en la mente de los usuarios de negocios. Cuando se ha capturado en sistemas, se extiende a través de múltiples programas en

definiciones de archivos, documentación en varios estados de precisión o en especificaciones de programas perdidos por largo tiempo. Algo de eso, por supuesto, está en el catálogo del sistema del DBMS.

Los metadatos proporcionan el contexto en el que los datos pueden entenderse y, por lo tanto, convertirse en información.

Una estrategia de metadatos completa permitirá a una organización comprender los activos de información bajo su control y medir el valor de esos activos.

Una de las mayores contribuciones de DA al activo de datos corporativos es la creación de modelos de datos. Un modelo conceptual de datos describe los requisitos de datos a un nivel muy alto. Un modelo de datos lógico proporciona detalles en profundidad de los tipos de datos, longitudes, relaciones y cardinalidad. El DA utiliza técnicas de normalización para ofrecer modelos de datos sólidos que representen con precisión los requisitos de datos de la organización.

Muchos DBA descartan la administración de datos como mero modelado de datos, requerido solo porque alguien necesita hablar con esos usuarios finales para obtener los requisitos de la base de datos. Pero una verdadera función de DA es mucho más que un simple modelado de datos. Es una disciplina de gestión orientada a los negocios responsable del activo de datos de la organización.

¿Por qué dedicar tanto tiempo a hablar sobre administración de datos en un curso sobre administración de bases de datos? Bueno, pocas organizaciones han implementado y dotado de personal un rol de DA. Cuanto mayor sea la organización, más probable es que exista una función DA. Sin embargo, cuando el rol de DA no está definido en la organización, el DBA debe asumir el manto de planificador y modelador de datos. El DBA generalmente no podrá asumir todas las funciones y la responsabilidad de DA como se resume en esta sección por las siguientes razones.

El DBA tiene muchas otras tareas técnicas que realizar que consumirán la mayor parte de su tiempo.

El gerente del grupo de DBA generalmente no tiene un puesto ejecutivo que le permita dictar políticas.

El DBA generalmente no tiene las habilidades para comunicarse de manera efectiva con los usuarios comerciales y generar consenso.

Francamente, la mayoría de los administradores de bases de datos son más felices al tratar con problemas técnicos y técnicos que con problemas comerciales y no técnicos.

Cuando existen funciones DA y DBA dentro de la organización, los dos grupos deben trabajar muy estrechamente entre sí. No es necesario que ambos tengan el mismo administrador, aunque eso podría facilitar la cooperación entre DA y DBA. En cualquier caso, es imperativo que haya un cierto grado de polinización cruzada de habilidades entre los dos grupos. El DA nunca entenderá la base de datos física como un DBA, y el DBA nunca entenderá los problemas de negocios de los datos como un DA, pero cada función de trabajo sería más efectiva con algún conocimiento sobre el otro.

En resumen, Las organizaciones realmente preocupadas por la calidad, la integridad y la reutilización de los datos implementarán y dotarán de personal invariablemente la función DA.

Administración de Bases de Datos

La administración de la base de datos es el enfoque de todo este curso, por lo que no dedicaré mucho tiempo a definirlo en esta breve sección. El resto del curso lo logrará muy bien. Esta sección resumirá rápidamente las funciones realizadas por el grupo de DBA cuando existe la función DA. Como se ilustra en la imagen siguiente, a un alto nivel, el DBA administra los datos y el DA administra los metadatos. Pero vamos a profundizar un poco más.

El primer deber de DBA es comprender los modelos de datos construidos por DA y poder comunicar el modelo a los desarrolladores de aplicaciones y otros técnicos apropiados. El modelo de datos lógicos es el mapa que el DBA utilizará para crear bases de datos físicas. El DBA transformará el modelo de datos lógico en un diseño de base de datos física eficiente. Es esencial que el DBA incorpore su conocimiento del DBMS que se está utilizando para crear un diseño de base de datos físico eficiente y apropiado a partir del modelo lógico. El DBA no debe confiar en el DA para el modelo físico final, más de lo que el DA debe confiar en el DBA para los modelos de datos conceptuales y lógicos.

El DBA es el conducto para la comunicación entre el equipo de DA y los técnicos y el personal de programación de aplicaciones. Por supuesto, la mayor parte del trabajo del DBA es el soporte continuo de las bases de datos creadas a partir del diseño físico y la administración de las aplicaciones que acceden a esas bases de datos.

Administración del Sistema

Algunas organizaciones, una vez más las más grandes, también tienen un rol de administración de sistemas (SA) o de programación de sistemas que afecta la implementación y las operaciones de DBMS. Cuando el rol de SA existe por separado del rol de DBA, es responsable de la instalación y configuración del DBMS. El administrador del sistema (SA) generalmente no es responsable del diseño y soporte de la base de datos. En cambio, el DBA es responsable de las bases de datos y el SA es responsable de la instalación, modificación y soporte del DBMS.

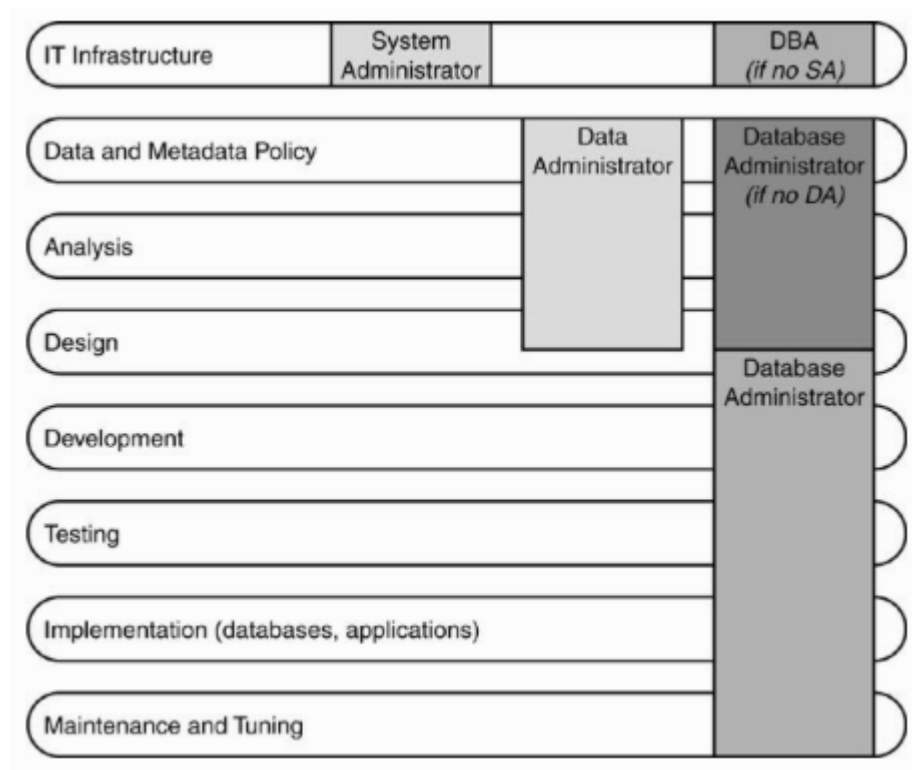
Además, el rol de SA garantiza que la infraestructura de TI se implemente de tal manera que el DBMS esté configurado para funcionar con otro software de sistema habilitante. Es posible que el SA deba trabajar con otros técnicos para configurar los procesadores de transacciones, el software de cola de mensajes, los protocolos de red y los parámetros del sistema operativo para permitir que el DBMS funcione con eficacia. El SA garantiza que la infraestructura de TI esté operativa para el desarrollo de la base de datos configurando el DBMS de manera adecuada, aplicando el mantenimiento continuo del proveedor de DBMS y coordinando la migración a las nuevas versiones y versiones de DBMS.

El administrador del sistema se asegura de que la infraestructura de TI esté operativa para el desarrollo de la base de datos al configurar el DBMS de manera adecuada, aplicar el

mantenimiento continuo del proveedor de DBMS y coordinar la migración a las nuevas versiones y versiones de DBMS.

Al igual que con DA, también debe haber una capacitación cruzada de habilidades entre SA y DBA. El SA nunca entenderá la base de datos física como un DBA, pero es poco probable que el DBA entienda la instalación y las relaciones técnicas profundas de un software de sistema como el SA. Sin embargo, cada función de trabajo será más efectiva con algún conocimiento del otro.

Cuando no existe un grupo de SA independiente, o cuando no existe una SA centrada en DBMS, el DBA asume la responsabilidad de la administración y programación del sistema. El gráfico en la imagen proporciona una delineación rápida de las tareas DA, DBA y SA.



Tareas del DBA

Un DBA debe ser capaz de realizar muchas tareas para garantizar que los datos y las bases de datos de la organización sean útiles, utilizables, estén disponibles y sean correctos. Estas tareas incluyen el diseño, la supervisión y el ajuste del rendimiento, la garantía de disponibilidad, la autorización de seguridad, la copia de seguridad y la recuperación, la integridad de los datos y, en realidad, todo lo que se interconecta con las bases de datos de la empresa. Examinemos cada uno de estos temas.

Diseño de Bases de Datos

La primera tarea que la mayoría de la gente piensa cuando piensa en los DBA es la capacidad de crear bases de datos bien diseñadas. Para diseñar y crear correctamente las bases de datos relacionales, los DBA deben comprender y adherirse a las prácticas de diseño relacional. Deben comprender tanto la teoría relacional como la implementación específica del RDBMS que se utiliza para crear la base de datos. El diseño de la base de datos requiere una buena comprensión de las técnicas conceptuales y lógicas de modelado de datos. La capacidad de crear e interpretar diagramas de entidad-relación es esencial para diseñar una base de datos relacional.

Además, el DBA debe poder transformar un modelo de datos lógico en una implementación de base de datos física. El DBA debe garantizar que el diseño y la implementación de la base de datos permitan una base de datos útil para las aplicaciones y los clientes que la usarán.

De hecho, el diseño de la base de datos es una habilidad significativa para que el DBA la posea. Sin embargo, el trabajo del DBA a menudo se asocia desproporcionadamente con el diseño de la base de datos. Aunque el diseño de bases de datos óptimas es importante, es una porción relativamente pequeña del trabajo del DBA. Un DBA probablemente pasará más tiempo administrando y ajustando bases de datos que en el diseño y la creación de bases de datos.

Sin embargo, de ninguna manera debe interpretar que esto significa que el diseño de la base de datos no es importante. Un diseño relacional deficiente puede dar como resultado un rendimiento deficiente, una base de datos que no satisface las necesidades de la organización y datos potencialmente inexactos.

Ajuste y Monitoreo de Rendimiento

El segundo rol más estrechamente asociado con el DBA es el monitoreo y ajuste del rendimiento. ¿Pero qué significa el término rendimiento de la base de datos? Piense, por un momento, en el rendimiento de la base de datos utilizando los conceptos familiares de oferta y demanda. Los usuarios demandan información de la base de datos. El DBMS proporciona información a los solicitantes. La velocidad a la que el DBMS suministra la demanda de información puede denominarse rendimiento de base de datos. Pero no es realmente tan simple. Cinco factores influyen en el rendimiento de la base de datos: carga de trabajo, rendimiento, recursos, optimización y contención.

La carga de trabajo que se solicita del DBMS define la demanda. Es una combinación de transacciones en línea, trabajos por lotes, consultas ad hoc, almacenamiento de datos y consultas analíticas, y comandos dirigidos a través del sistema en un momento dado. La carga de trabajo puede fluctuar drásticamente de un día a otro, de una hora a otra, de un minuto a otro, e incluso de un segundo a otro. A veces se puede predecir la carga de trabajo (como el procesamiento pesado de fin de mes de la nómina, o el acceso muy ligero después de las 7:30 p.m., cuando la mayoría de los usuarios se han ido para el día), pero en otras ocasiones es impredecible. La carga de trabajo global tiene un gran impacto en el rendimiento de la base de datos.

El rendimiento (del procesamiento) define la capacidad general del hardware y software de la computadora para procesar. Es un compuesto de velocidad de E / S, velocidad de la CPU, capacidades paralelas de la máquina y la eficiencia del sistema operativo y el software del sistema. Las herramientas de hardware y software a disposición del sistema se conocen como los recursos del sistema. Los ejemplos incluyen el núcleo de la base de datos, el espacio en disco, los controladores de caché y el microcódigo.

El cuarto elemento que define el rendimiento de la base de datos es la optimización. Se pueden optimizar todos los tipos de sistemas, pero las consultas relacionales son únicas, ya que la optimización se realiza principalmente de forma interna al DBMS. Sin embargo, hay muchos otros factores que deben optimizarse (formulación de SQL, parámetros de la base de datos, programación eficiente, etc.) para permitir que el optimizador de la base de datos cree las rutas de acceso más eficientes.

Cuando la demanda (carga de trabajo) para un recurso en particular es alta, puede producirse una contención. La contención es la condición en la que dos o más componentes de la carga de trabajo intentan utilizar un solo recurso de manera conflictiva (por ejemplo, actualizaciones duales a la misma parte de datos). A medida que aumenta la contención, disminuye el rendimiento.

Por lo tanto, el rendimiento de la base de datos se puede definir como la optimización del uso de recursos para aumentar el rendimiento y minimizar la contención, lo que permite procesar la mayor carga de trabajo posible.

Cuando una aplicación que usa una base de datos encuentra problemas de rendimiento, el DBA suele ser el primero al que se llama para resolver el problema. Por supuesto, el DBA no puede administrar el rendimiento de la base de datos en un vacío. Las aplicaciones se comunican regularmente con otras aplicaciones, sistemas y componentes de la infraestructura de TI. Una efectiva estrategia de monitoreo y ajuste del rendimiento requiere no solo conocimientos de DBMS sino también conocimientos fuera del alcance de la administración de la base de datos. Muchas tareas de gestión del rendimiento deben compartirse entre el DBA y otros técnicos. En otras palabras, el manejo de los problemas de rendimiento es verdaderamente un esfuerzo de toda la empresa.

El DBA debe estar atento al monitoreo del sistema, la base de datos y el rendimiento de la aplicación. En la medida de lo posible, esto debe lograrse utilizando software y scripts automatizados. Las tablas del sistema de sondeo y las alertas de construcción basadas en umbrales se pueden usar para identificar proactivamente los problemas. Las alertas se pueden configurar para enviar un correo electrónico al DBA cuando las métricas de rendimiento no están dentro de los límites aceptados.

Se requieren muchas tareas y habilidades de los DBA para garantizar un acceso eficiente a las bases de datos. Algunas de estas capacidades incluyen la creación de índices adecuados, la especificación de buffers y cachés suficientemente grandes, la alineación de la implementación de la base de datos con la infraestructura de TI, el monitoreo continuo de las bases de datos y las aplicaciones, la reorganización de la base de datos y la adaptación a los cambios empresariales: más usuarios, más datos, procesamiento adicional, y cambiando requisitos y regulaciones.

Aseguramiento de la Disponibilidad

La disponibilidad de datos y bases de datos a menudo está estrechamente alineada con el rendimiento, pero en realidad es una preocupación separada. Por supuesto, si el DBMS está fuera de línea, el rendimiento será horrible porque no se puede acceder a los datos. Pero asegurar la disponibilidad de la base de datos es un proceso multifacético.

El primer componente de la disponibilidad es mantener el DBMS en funcionamiento. El monitoreo vigilante y las alertas automáticas se pueden usar para advertir sobre las interrupciones de DBMS y solicitar acciones correctivas.

Las bases de datos individuales también deben mantenerse de modo que los datos contenidos en ellas estén disponibles siempre que las aplicaciones y los clientes lo requieran. Hacerlo requiere que el DBA diseñe la base de datos para que pueda mantenerse con interrupciones mínimas, pero también para ayudar a diseñar aplicaciones para minimizar conflictos cuando se requiere acceso concurrente.

Un componente adicional de la disponibilidad es minimizar la cantidad de tiempo de inactividad requerido para realizar tareas administrativas. Cuanto más rápido pueda realizar el DBA las tareas administrativas que requieren que las bases de datos estén fuera de línea, más disponibles estarán los datos. Cada vez más, los proveedores de DBMS y los proveedores de software independientes (ISV - Independent Software Vendors) proporcionan utilidades sin interrupciones que se pueden realizar en bases de datos mientras las aplicaciones leen y escriben desde ellas. Pero estos usualmente requieren más habilidad y planificación por adelantado para implementar.

El DBA debe comprender todos estos aspectos de la disponibilidad y garantizar que cada aplicación reciba el nivel correcto de disponibilidad para sus necesidades.

Seguridad y Autorización de la Base de Datos

Una vez que se haya diseñado e implementado la base de datos, los programadores y usuarios deberán acceder y modificar los datos en la base de datos. Pero solo los programadores y usuarios autorizados deben tener acceso para evitar violaciones de seguridad y modificaciones inapropiadas de datos. Es responsabilidad del DBA garantizar que los datos estén disponibles solo para los usuarios autorizados.

Por lo general, aunque no siempre (consulte la sección "Centralización de la seguridad"), el DBA funciona con las funciones de seguridad internas del DBMS en forma de sentencias de SQL GRANT y REVOKE, así como con las funciones de autorización de grupo del DBMS. La seguridad debe administrarse para muchas acciones requeridas por el entorno de base de datos:

Creación de objetos de base de datos, incluidas bases de datos, tablas, vistas y estructuras de programas

Alterar la estructura de los objetos de la base de datos.

Acceso al catálogo del sistema.

Lectura y modificación de datos en tablas.

Crear y acceder a funciones y tipos de datos definidos por el usuario.

Ejecutar procedimientos almacenados

Iniciar y detener bases de datos y objetos de base de datos asociados

Configuración y modificación de parámetros y especificaciones de DBMS.

Ejecutar utilidades de base de datos como LOAD, RECOVER y REORG

La seguridad de la base de datos también se puede aplicar de otras maneras. Por ejemplo, se pueden crear vistas para bloquear las columnas o filas confidenciales para que no sean vistas por los usuarios finales y los programadores. Y el DBA también interactúa frecuentemente con métodos de seguridad externos cuando afectan la seguridad de la base de datos.

El DBA debe comprender y ser capaz de implementar cualquier aspecto de la seguridad que afecte el acceso a las bases de datos. Un área que debería ser de particular interés debido a las violaciones de datos en las noticias de estos días son los ataques de inyección de SQL y cómo prevenirlos.

Centralización de la Seguridad

Algunas organizaciones han tomado medidas para ensamblar todas las tareas, políticas y procedimientos de seguridad y autorización en un grupo de seguridad de TI centralizado. Tal empresa no es trivial y, como tal, es más común en organizaciones más grandes que en organizaciones más pequeñas.

Las industrias fuertemente reguladas pueden optar por centralizar las operaciones de seguridad. Además, las organizaciones con computadoras centrales tienden a considerar la centralización más a menudo que aquellas sin computadoras centrales.

La transferencia exitosa de la responsabilidad de la seguridad de la base de datos del grupo de administración de la base de datos a un grupo de seguridad centralizado requiere la capacitación del personal de seguridad en técnicas de seguridad de la base de datos. Además, muchas de estas empresas utilizan software que elimina las operaciones de seguridad del DBMS e imita la misma funcionalidad en un paquete de seguridad más tradicional (como RACF o ACF2 en el mainframe).

Incluso teniendo en cuenta estos problemas, la mayoría de las organizaciones todavía confían en el DBA para administrar la seguridad de la base de datos.

Gobernanza y Cumplimiento Normativo

Asegurar el cumplimiento de las regulaciones gubernamentales y de la industria es una tarea adicional que se requiere de la administración de la base de datos, al menos en términos de la implementación de controles adecuados. El DBA debe trabajar con la administración, los auditores

y los expertos de negocios para comprender las regulaciones que se aplican a su industria y la manera en que se tratan los datos.

Ciertos aspectos del cumplimiento normativo abordan los procedimientos operativos estándar de DBA. Por ejemplo, las regulaciones pueden contener lenguaje que aplique procedimientos específicos de seguridad y autorización, requisitos de auditoría, especificaciones de respaldo de datos y procedimientos de administración de cambios. Sin embargo, para garantizar el cumplimiento, es posible que se requiera una documentación más estricta, o quizás un mayor grado de diligencia o automatización (como un seguimiento más profundo de la auditoría).

Otros aspectos del cumplimiento normativo pueden requerir que el DBA adopte diferentes técnicas, tácticas y habilidades. Por ejemplo, las regulaciones de retención de datos pueden requerir que los datos se mantengan mucho tiempo después de que se necesiten para ser almacenados en una base de datos de producción, lo que requiere habilidades de archivo de base de datos. O es posible que ciertos datos deban estar protegidos de la vista, lo que requiere el enmascaramiento de los datos o, en algunos casos, la configuración del cifrado.

Los DBA no deben encargarse de comprender las regulaciones en profundidad, ni deben establecer los estándares mediante los cuales la organización cumple con las regulaciones. Sin embargo, los DBA se involucrarán para ayudar a establecer los controles y procedimientos adecuados para los proyectos de cumplimiento, específicamente y especialmente con respecto al tratamiento de los datos.

Respaldo y Recuperación

El DBA debe estar preparado para recuperar datos en caso de un problema. "Problema" puede significar cualquier cosa, desde una falla del sistema o un error del programa hasta un desastre natural que apaga una organización. La mayoría de las recuperaciones en la actualidad se producen como resultado de un error de software de aplicación y un error humano. Las fallas de hardware no son tan frecuentes como solían ser. De hecho, las estimaciones de los analistas indican que el 80 por ciento de los errores de aplicación se deben a fallas de software y errores humanos. El DBA debe estar preparado para recuperar datos en un punto utilizable, sin importar cuál sea la causa, y hacerlo lo más rápido posible.

El primer tipo de recuperación de datos que generalmente viene a la mente es una recuperación a lo actual, por lo general ante un cierre importante. El resultado final de la recuperación es que la base de datos vuelve a su estado actual en el momento de la falla. Las aplicaciones no están disponibles hasta que se complete la recuperación.

Otro tipo de recuperación tradicional es una recuperación de un punto en el tiempo. La recuperación de un punto en el tiempo por lo general se realiza para tratar un problema de nivel de aplicación. Las técnicas convencionales para realizar una recuperación de un punto en el tiempo eliminarán los efectos de todas las transacciones desde un punto específico en el tiempo. Esto a veces puede causar problemas si hubo algunas transacciones válidas durante ese período de tiempo que aún deben aplicarse.

La recuperación de transacciones es un tercer tipo de recuperación que aborda las deficiencias de los tipos tradicionales de recuperación: el tiempo de inactividad y la pérdida de buenos datos. Por lo tanto, la recuperación de transacciones es una recuperación de aplicaciones mediante la cual los efectos de transacciones específicas durante un período de tiempo específico se eliminan de la base de datos. Por lo tanto, la recuperación de transacciones a veces se denomina recuperación de aplicaciones.

La mayoría de los técnicos piensan en la recuperación para resolver desastres como fallas de hardware. Aunque aún se producen fallas de hardware y los técnicos deben estar preparados para recuperarse de dichas fallas, la mayoría de las recuperaciones actuales son necesarias debido a errores humanos o errores en los programas.

El DBA debe estar preparado para lidiar con todos estos tipos de recuperación. Esto implica desarrollar una estrategia de respaldo para garantizar que los datos no se pierdan en caso de un error en el software, hardware o un proceso manual. La estrategia debe ser aplicable al procesamiento de la base de datos, por lo que debe incluir copias de imagen de los archivos de la base de datos, así como un plan de respaldo / recuperación para los registros de la base de datos. Debe tener en cuenta cualquier actividad de archivo que no sea de base de datos que pueda afectar también a las aplicaciones de base de datos.

Aseguramiento de la Integridad de los Datos

Una base de datos debe estar diseñada para almacenar los datos correctos de la manera correcta sin que dichos datos se dañen o corrompan. Para garantizar este proceso, el DBA implementa reglas de integridad usando las características del DBMS. Tres aspectos de la integridad son relevantes para nuestra discusión de las bases de datos: física, semántica e interna.

Los problemas físicos se pueden manejar mediante las funciones de DBMS como dominios y tipos de datos. El DBA elige el tipo de datos apropiado para cada columna de cada tabla. Esta acción garantiza que solo los datos de ese tipo se almacenen en la base de datos. Es decir, el DBMS impone la integridad de los datos con respecto a su tipo. Una columna definida como "entero" puede contener solo enteros. Los intentos de almacenar valores no numéricos o no integrales en una columna definida como entero fracasarán. Los DBA también pueden usar restricciones para delinear aún más el tipo de datos que se pueden almacenar en las columnas de la base de datos. La mayoría de los productos DBMS relacionales proporcionan los siguientes tipos de restricciones:

Las restricciones referenciales se utilizan para especificar las columnas que definen cualquier relación entre tablas. Las restricciones de referencia se utilizan para implementar la integridad referencial, lo que garantiza que todas las referencias previstas de los datos en una columna (o conjunto de columnas) de una tabla sean válidas con respecto a los datos en otra columna de la misma tabla o una diferente.

Las restricciones únicas aseguran que los valores para una columna o un conjunto de columnas se produzcan solo una vez en una tabla.

Las restricciones de verificación (check) se utilizan para colocar reglas de integridad más complejas en una columna o conjunto de columnas en una tabla. Las restricciones de verificación se definen normalmente mediante SQL y se pueden usar para definir los valores de datos que están permitidos para una columna o conjunto de columnas.

La integridad semántica es más difícil de controlar y se define con menos facilidad. Los DBA deben estar preparados para aplicar políticas y prácticas que garanticen que los datos almacenados en sus bases de datos sean precisos, apropiados y utilizables. Un ejemplo de un problema semántico es la calidad de los datos en la base de datos. El simple almacenamiento de datos que cumplan con las definiciones de integridad física especificadas en la base de datos no es suficiente. Se deben implementar procedimientos y prácticas para garantizar la calidad de los datos. Por ejemplo, una base de datos de clientes que contiene una dirección o número de teléfono incorrecto para el 25 por ciento de los clientes almacenados en la misma es un ejemplo de una base de datos de baja calidad. No existe un método físico y sistemático para garantizar la precisión de los datos. La calidad de los datos se fomenta a través de un código de aplicación adecuado, prácticas comerciales sólidas y políticas de datos específicas. La redundancia es otro tema semántico. Si los elementos de datos se almacenan de manera redundante en toda la base de datos, el DBA debe documentar este hecho y trabajar para garantizar que se implementen procedimientos para mantener los datos redundantes sincronizados y precisos.

El último aspecto de la integridad es un problema interno del DBMS. El DBMS se basa en estructuras y códigos internos para mantener enlaces, punteros e identificadores. En la mayoría de los casos, el DBMS hará un buen trabajo al mantener estas estructuras, pero el DBA debe ser consciente de su existencia y cómo hacer frente cuando falla el DBMS. La integridad del DBMS interno es esencial en las siguientes áreas:

Consistencia del índice. Un índice no es realmente nada más que una lista ordenada de punteros a datos en tablas de base de datos. Si, por algún motivo, el índice no está sincronizado con los datos, el acceso indexado puede no devolver los datos adecuados. El DBA tiene herramientas a su disposición para verificar y remediar este tipo de errores.

Consistencia del puntero. A veces, los objetos multimedia grandes no se almacenan en los mismos archivos físicos que otros datos. Por lo tanto, el DBMS requiere estructuras de puntero para mantener los datos multimedia sincronizados con los datos de la tabla base. Una vez más, estos indicadores pueden perder la sincronización si no se siguen los procedimientos de administración adecuados.

Copia de consistencia. Algunos productos de DBMS ocasionalmente realizan copias de respaldo incorrectas que no pueden usarse efectivamente para la recuperación. Es esencial identificar estos escenarios y tomar acciones correctivas.

En general, asegurar la integridad es una habilidad esencial del DBA.

Migración de versión del DBMS

El DBA también es responsable de administrar la migración de la versión a la versión del DBMS. Los productos de DBMS cambian con bastante frecuencia: las nuevas versiones generalmente se lanzan cada año aproximadamente. La tarea de mantener el DBMS en funcionamiento y actualizado es un esfuerzo continuo que consumirá muchos ciclos de DBA. Cualquier enfoque que se tome debe cumplir con las necesidades de la organización, al tiempo que reduce las interrupciones y minimiza la necesidad de cambiar las aplicaciones.

Un mil usos

Las bases de datos están en el centro de las aplicaciones modernas. Si el DBMS falla, las aplicaciones fallan y si las aplicaciones fallan, todo el negocio puede detenerse. Si las bases de datos y las aplicaciones fallan a menudo, todo el negocio puede fallar. Por lo tanto, la administración de la base de datos es crítica para el éxito continuo de los negocios modernos.

Además, las bases de datos interactúan con casi todos los componentes de la infraestructura de TI. La infraestructura de TI de hoy consiste en cosas tales como

- Lenguajes y entornos de programación como COBOL, Microsoft Visual Studio, C / C ++ / C #, Java y PHP
- Frameworks de software como .NET y J2EE
- Herramientas de diseño de bases de datos y procesos como ERwin, Rational Rose y Visual Paradigm.
- Sistemas de procesamiento de transacciones como CICS y Tuxedo.
- Servidores de aplicaciones como WebSphere, JBoss, Oracle Application Server y EAServer
- Software de cola de mensajes, como MQSeries y MSMQ
- Software y protocolos de red como SNA, VTAM y TCP/IP
- Hardware de red, como puentes, enrutadores, concentradores y cableado
- Múltiples sistemas operativos como Windows, z/OS y MVS, UNIX y Linux, y quizás otros
- Hardware y software de almacenamiento de datos, como servidores de almacenamiento empresarial, Microsoft SMS, IBM DFHSM, SAN y NAS
- Paquetes de seguridad del sistema operativo, como RACF, ACF2 y Kerberos
- Otros tipos de hardware de almacenamiento, por ejemplo, máquinas de cinta, silos y almacenamiento de estado sólido (basado en memoria)
- Técnicas de almacenamiento de archivos y conjuntos de datos que no son DBMS, como VSAM y b-tree
- Productos NoSQL como Hadoop y MongoDB
- Herramientas de administración de bases de datos y cómo interactúan con otras soluciones de administración de sistemas.
- Herramientas y marcos de administración de sistemas como HP OpenView y CA Unicenter
- Software de control operacional, como el software de programación de lotes y subsistemas de entrada de trabajo

- Soluciones de distribución de software para implementar nuevas versiones de software del sistema en toda la red.
- Las bases de datos y aplicaciones habilitadas para Internet y la Web.
- Técnicas de desarrollo del cliente/servidor (multitier, fat server/thin client, thin server/fat client, etc.)
- Tecnologías y técnicas de desarrollo orientadas a objetos y basadas en componentes, como CORBA, COM, OLE / DB, ADO y EJB.
- Dispositivos de tecnología de computación generalizados, como tabletas y teléfonos inteligentes

Aunque es imposible llegar a ser experto en todas estas tecnologías, el DBA debe tener algún conocimiento de cada una de estas áreas y cómo se interrelacionan. Aún más importante, el DBA debe tener los números de teléfono de los expertos para contactar en caso de que algún software y hardware asociado cause acceso a la base de datos o problemas de rendimiento.

Tipos de DBAs

Hay DBA que se centran en el diseño lógico y DBA que se centran en el diseño físico; DBA que se especializan en sistemas de construcción y DBA que se especializan en sistemas de mantenimiento y ajuste; DBA de especialidad y DBA de propósito general. En verdad, el trabajo de DBA abarca muchos roles.

Algunas organizaciones optan por dividir las responsabilidades de DBA en trabajos separados. Por supuesto, esto ocurre con mayor frecuencia en organizaciones más grandes, porque las organizaciones más pequeñas a menudo no pueden permitirse el lujo de tener múltiples DBA especializadas.

Otras empresas simplemente contratan DBA para realizar todas las tareas necesarias para diseñar, crear, documentar, ajustar y mantener los datos de la organización, las bases de datos y los sistemas de administración de bases de datos. Veamos algunos de los tipos más comunes de DBA.

DBA del Sistema

Un DBA del sistema se enfoca en temas técnicos más que empresariales, principalmente en el área de administración del sistema. Las tareas típicas se centran en la instalación física y el rendimiento del software DBMS, incluyendo

Instalar nuevas versiones de DBMS y aplicar arreglos de mantenimiento suministrados por el proveedor de DBMS

Configuración y personalización de los parámetros del sistema.

Ajuste del sistema operativo, la red y los procesadores de transacciones para que funcionen con el DBMS

Asegurar el almacenamiento adecuado para el DBMS

Habilitar el DBMS para trabajar con dispositivos de almacenamiento y software de administración de almacenamiento

Interfaz con cualquier otra tecnología requerida por las aplicaciones de base de datos.

Instalación de herramientas y utilidades DBA

Los DBA del sistema rara vez se involucran con la implementación real de bases de datos y aplicaciones. Pueden involucrarse en los esfuerzos de ajuste de la aplicación cuando se deben modificar los parámetros del sistema operativo o los parámetros complejos de DBMS.

De hecho, el trabajo de DBA del sistema generalmente existe solo si la organización no tiene un departamento oficial de administración de sistemas o programación de sistemas.

Arquitecto de Bases de Datos

Algunas organizaciones designan un trabajo separado, denominado arquitecto de base de datos, para diseñar e implementar nuevas bases de datos. Los arquitectos de la base de datos están involucrados únicamente en el nuevo trabajo de diseño y desarrollo; no se involucran en los esfuerzos de mantenimiento, administración y ajuste de las bases de datos y aplicaciones establecidas. El arquitecto de bases de datos diseña nuevas bases de datos para nuevas aplicaciones o tal vez una nueva base de datos para una aplicación existente.

El fundamento para crear una posición separada es que las habilidades requeridas para diseñar nuevas bases de datos son diferentes de las habilidades requeridas para mantener la implementación de una base de datos existente en funcionamiento. Es más probable que un arquitecto de bases de datos tenga experiencia en administración y modelado de datos que un DBA de propósito general, porque las habilidades de DA son más útiles para desarrollar un diseño de base de datos inicial.

Las tareas típicas realizadas por el arquitecto de base de datos incluyen

- Creación de un modelo de datos lógico (si no existe una posición de DA o de modelador de datos)
- Traducción de modelos de datos lógicos en diseños de bases de datos físicas.
- Implementar bases de datos eficientes, que incluyen características físicas, diseño de índice y asignación de objetos de base de datos a dispositivos de almacenamiento físico
- Análisis de los requisitos de acceso y modificación de datos para garantizar un SQL eficiente y para garantizar que el diseño de la base de datos sea óptimo
- Creación de estrategias de backup y recuperación para nuevas bases de datos.

La mayoría de las organizaciones no tienen una posición de arquitecto de base de datos independiente, en su lugar requieren que los DBA trabajen en proyectos de base de datos nuevos y establecidos.

Analista de Bases de Datos

Otra posición común del personal es el analista de base de datos. Realmente no hay una definición común para el analista de base de datos. A veces, los DBA junior se denominan analistas de bases de datos. A veces, un analista de base de datos desempeña una función similar a la del arquitecto de base de datos. A veces se hace referencia al DA como el analista de base de datos, o tal vez el analista de datos. Y a veces analista de base de datos es solo otro término usado por algunas organizaciones en lugar de administrador de base de datos.

Modelador de Datos

Cuando el rol de DA no está definido o no tiene personal, puede haber un rol de modelador de datos definido. Un modelador de datos suele ser responsable de un subconjunto de las responsabilidades del DA. Las tareas de modelado de datos incluyen

- La recopilación de requerimientos de datos para proyectos de desarrollo.
- Análisis de los requisitos de datos.
- Diseño de modelos de datos conceptuales y lógicos basados en proyectos.
- Creación de un modelo de datos corporativos y actualización del modelo de datos corporativos.
- Trabajar con los DBA para garantizar que tengan una buena comprensión de los modelos de datos

DBA de Aplicación

En contraste directo con el DBA del sistema está el DBA de aplicación. Los DBA de aplicación se centran en el diseño de bases de datos y el soporte y la administración continuos de bases de datos para una aplicación o aplicaciones específicas. Es probable que el DBA de la aplicación sea un experto en la escritura y depuración de complejos SQL y comprende las mejores maneras de incorporar solicitudes de base de datos en los programas de aplicación. El DBA de la aplicación también debe ser capaz de realizar la gestión de cambios de la base de datos, el ajuste del rendimiento y la mayoría de las otras funciones del DBA. La diferencia es el enfoque del DBA de aplicación, no en la implementación general de DBMS y el entorno de base de datos, sino en un subconjunto específico de aplicaciones

No todas las organizaciones tienen DBAs de aplicación dentro de su personal. Sin embargo, cuando existen DBA de aplicación, todavía se requieren DBA de propósito general para soportar el entorno de la base de datos y la infraestructura en general. Cuando los DBA de aplicación no existen dentro de una organización, es probable que los DBA de propósito general se asignen para respaldar aplicaciones específicas mientras simultáneamente se mantiene el entorno de base de datos de la organización.

Hay ventajas y desventajas para la dotación de personal de DBA de aplicación. Los argumentos a favor de los DBA de aplicación incluyen lo siguiente:

Los DBA de aplicación pueden enfocarse mejor en una aplicación individual, lo que puede resultar en un mejor servicio para los desarrolladores de esa aplicación.

El DBA de aplicación se ve más a menudo como un componente integral del equipo de desarrollo y, por lo tanto, está mejor informado sobre los nuevos planes de desarrollo y los cambios en los planes.

Debido a que los DBA de aplicación trabajan constantemente en un conjunto específico de aplicaciones, pueden adquirir una mejor comprensión general de cómo funciona cada aplicación, lo que les permite satisfacer mejor las necesidades de los desarrolladores de aplicaciones.

Con una comprensión más completa de la aplicación, un DBA de la aplicación tendrá una mejor comprensión de cómo la aplicación afecta al negocio en general. Este conocimiento probablemente resultará en la ejecución de tareas de DBA para apoyar mejor a la organización.

Pero no todo es favorable para la aplicación de DBAs. Existen inconvenientes para implementar un rol de DBA de aplicación, incluidos estos:

Los DBA de aplicación pueden perder de vista las necesidades de datos generales de la organización debido a su enfoque limitado en una sola aplicación.

El DBA de aplicación puede aislarse. La falta de comunicación con un grupo de DBA centralizado (si existe) puede resultar en un menor intercambio de habilidades.

Cuando los DBA de aplicación implementan procedimientos útiles, se requiere más esfuerzo para compartir estos procedimientos con los otros DBA.

Debido a la naturaleza centrada en la aplicación de los DBA de la aplicación, pueden perder de vista las nuevas características y funcionalidades que ofrece el grupo DBMS.

En general, cuando se contrata DBAs de aplicación, asegúrese de contar también con un grupo de DBA centralizado. Los DBA de aplicación deben tener la responsabilidad principal de las aplicaciones específicas, pero también deben verse como parte del grupo de DBA centralizado. Al hacerlo, se fomentarán las ventajas de implementar DBA de aplicaciones al tiempo que se desalientan los inconvenientes.

DBA Orientado a Tareas

Las organizaciones más grandes a veces crean DBA muy especializados que se centran en una sola tarea específica de DBA. Pero los DBA orientados a tareas son bastante raros fuera de las grandes departamentos de TI. Un ejemplo de un DBA orientado a tareas es un DBA de copia de seguridad y recuperación que dedica todo su día a garantizar la capacidad de recuperación de las bases de datos de la organización.

La mayoría de las organizaciones no pueden permitirse este nivel de especialización, pero cuando es posible, los DBA orientados a las tareas pueden garantizar que las tareas muy importantes de DBA sean abordadas por especialistas muy bien informados.

Analista de Rendimiento

Los analistas de rendimiento son un tipo específico de DBA orientado a tareas. El analista de rendimiento, es más común que otros DBA orientados a tareas, se centra únicamente en el rendimiento de las aplicaciones de base de datos.

Un analista de rendimiento debe comprender los detalles y los matices de la codificación SQL para el rendimiento, así como tener la capacidad de diseñar bases de datos para el rendimiento. Los analistas de rendimiento saben que el DBMS se está utilizando a un nivel técnico muy detallado para que puedan realizar los cambios apropiados en el DBMS y los parámetros del sistema cuando sea necesario.

Pero el analista de rendimiento no debería ser un DBA del sistema. El analista de rendimiento debe poder hablar con los desarrolladores de aplicaciones en su idioma para ayudarles a facilitar los cambios de programa apropiados para incrementar el rendimiento.

El analista de rendimiento suele ser el miembro más calificado y con mayor experiencia del personal de DBA. Es muy probable que un DBA sénior crezca en este papel debido a su experiencia y al respeto que se ha ganado en los esfuerzos de optimización anteriores.

Administrador de Data Warehouse

Las organizaciones que implementan almacenes de datos para realizar análisis en profundidad de los datos a menudo son empleados de DBA específicamente para monitorear y respaldar el entorno de almacenamiento de datos. Los administradores del almacén de datos deben ser DBA capaces, pero con un conocimiento profundo de las diferencias entre una base de datos que admite el procesamiento de transacciones en línea (OLTP) y un almacén de datos. Las tareas y los requisitos comunes de administración del almacén de datos incluyen

- Experiencia con inteligencia empresarial, análisis de datos y herramientas de consulta e informes.
- Diseño de base de datos para acceso de solo lectura
- Problemas de diseño de almacenamiento de datos, como el esquema en estrella
- Tecnologías de almacenamiento de datos como el procesamiento analítico en línea, u OLAP (incluidos ROLAP, MOLAP y HOLAP)
- Transformación de datos y habilidades de conversión.
- Una comprensión de los problemas de calidad de datos
- Experiencia con formatos de datos para carga y descarga de datos.
- Implementación y administración de middleware.

Consideraciones del Personal

Dotación de personal de la organización DBA no es un asunto simple. Hay varias consideraciones no triviales que deben abordarse, incluido el tamaño del personal de DBA y la estructura de informes para los DBA.

¿Cuántos DBAs?

Una de las cosas más difíciles de determinar es la cantidad óptima de DBA necesarios para mantener las bases de datos de una organización en línea y operar de manera eficiente. Muchas organizaciones intentan operar con la cantidad mínima de DBA en el personal, la idea es que tener menos miembros del personal reduce los costos. Pero esa suposición puede no ser cierta. Un personal de DBA con exceso de trabajo puede cometer errores que causan tiempos de inactividad y problemas operativos que superan el costo del salario de un DBA adicional.

Pero determinar el número óptimo de DBA no es una ciencia precisa. Depende de muchos factores, incluyendo

Número de bases de datos. Cuantas más bases de datos sean compatibles, más complejo será el trabajo de administración de la base de datos. Cada base de datos debe ser diseñada, implementada, monitoreada para disponibilidad y rendimiento, respaldada y administrada. Hay un límite para la cantidad de bases de datos que un DBA individual puede controlar.

- **Número de usuarios.** A medida que los usuarios adicionales se conectan como clientes de las aplicaciones que acceden a las bases de datos, se vuelve más difícil garantizar un rendimiento óptimo de la base de datos. Además, a medida que aumenta el número de usuarios, aumenta la posibilidad de que aumente el volumen de problemas y llamadas, lo que complica aún más el trabajo del DBA.
- **Número de solicitudes.** Una sola base de datos puede ser utilizada por numerosas aplicaciones. De hecho, uno de los principales beneficios del DBMS es permitir que los datos se compartan en una organización. A medida que se conectan más aplicaciones, se ejerce una presión adicional sobre la base de datos en términos de rendimiento, disponibilidad y recursos necesarios, y es posible que se requieran más DBA para admitir la misma cantidad de bases de datos.
- **Acuerdos de nivel de servicio (SLA).** Cuanto más restrictivo sea el SLA, más difícil será para el DBA entregar el servicio. Por ejemplo, un SLA que requiere un tiempo de respuesta de un segundo para las transacciones es más difícil de soportar que un SLA que requiere un tiempo de respuesta de 3 segundos.
- **Requisitos de disponibilidad.** Cuando las bases de datos tienen un período permitido de tiempo de inactividad programado, la administración de la base de datos se vuelve más sencilla porque algunas tareas de DBA requieren una interrupción o son más fáciles cuando se puede tomar una interrupción. Consideraciones como la compatibilidad con transacciones de e-business y la Web hacen que sea necesario contar con una base de datos disponible las 24 horas, los 7 días de la semana.

- **Impacto del tiempo de inactividad.** Cuanto mayor sea el impacto financiero de que una base de datos no esté disponible, más difícil será para el DBA porque la presión se aplicará para asegurar una mayor disponibilidad de la base de datos.
- **Requisitos de desempeño.** A medida que los requisitos de acceso a la base de datos se orientan más hacia el rendimiento y se dicta un acceso más rápido y más frecuente, el DBA se vuelve más complicado.
- **Tipo de aplicaciones.** Las organizaciones implementan todo tipo de aplicaciones. Los tipos de aplicaciones que deben admitirse tienen un impacto en la necesidad de los servicios de DBA. Las necesidades de DBMS y de la base de datos de una aplicación de misión crítica difieren de las de una aplicación que no es de misión crítica. Es más probable que las aplicaciones de misión crítica requieran monitoreo constante y más vigilancia para garantizar la disponibilidad. Del mismo modo, la aplicación OLTP (On-Line Transactional Processing) tendrá diferentes características y requisitos de administración de las aplicaciones OLAP (On-Line Analytical Processing). Es probable que las transacciones OLTP sean de menor duración que las consultas OLAP; Las aplicaciones OLTP realizan operaciones de lectura y escritura, mientras que las aplicaciones OLAP suelen ser de solo lectura. Cada uno tiene desafíos de administración que imponen diferentes procedimientos y necesidades de DBA.
- **Volatilidad.** La frecuencia de las solicitudes de cambio de base de datos es un factor importante en la necesidad de DBA adicionales. Un entorno de base de datos estática que requiera pocos cambios no requerirá el mismo nivel de esfuerzo de DBA que un entorno de base de datos volátil que cambia con frecuencia. Desafortunadamente, el nivel de volatilidad para la mayoría de las bases de datos y aplicaciones tiende a cambiar dramáticamente con el tiempo. Es difícil determinar cuán volátil será el entorno de una base de datos general durante su vida útil.
- **Experiencia del personal de DBA.** La habilidad del personal de DBA existente afectará si se requieren o no DBA adicionales. Un personal altamente calificado de DBA podrá lograr más que un equipo de principiantes. Las habilidades más que la experiencia dictan los requisitos de personal de DBA. Un DBA altamente motivado con dos años de experiencia podría superar fácilmente a un veterano de diez años que está agotado y desmotivado.
- **Experiencia del personal de programación.** Mientras los desarrolladores de aplicaciones menos capacitados estén en la base de datos y en la programación de SQL, los DBA más involucrados deberán estar en el proceso de desarrollo, realizando tareas como la composición compleja de SQL, el análisis, la depuración, la optimización y la conectividad. A medida que aumenta la experiencia del personal de programación, la complejidad de DBA disminuye.
- **Experiencia del usuario final.** Cuando los usuarios finales acceden a las bases de datos directamente con SQL ad hoc, su nivel de habilidad tiene un impacto directo en la complejidad del DBA.
- **Herramientas DBA.** Los proveedores de DBMS y una serie de ISV (Independent Software Vendor) ofrecen herramientas que automatizan las tareas de DBA y facilitan la administración de las bases de datos. Cuantas más herramientas estén disponibles y el grado en que estén integradas, las tareas de DBA serán menos complejas. Los analistas de

la industria han estimado que sin herramientas de DBA se puede requerir hasta el doble de la cantidad de DBA.

A pesar de la lista anterior de problemas de complejidad, es muy difícil combinar todos estos factores en una fórmula que dicte la cantidad óptima de DBA que se emplearán. Aunque la investigación es algo anticuada, los analistas de la industria en el Grupo META crearon una fórmula flexible para calcular el nivel de esfuerzo (LOE) del DBA. La fórmula no es rigurosa, pero llega a el LOE de un DBA aplicando ponderaciones a seis factores: complejidad del sistema, inmadurez de la aplicación, sofisticación del usuario final, funcionalidad del software, disponibilidad del sistema y sofisticación del personal. Al medir cada uno de estos elementos tanto como sea posible para indicar tasas altas o bajas, se insertan los valores en la fórmula y se llega a un número que se traduce en una estimación para la cantidad de DBA requeridos.

Estructuras de Reporte (Organigrama) de DBA

¿A quién debe reportar el grupo de DBAs dentro de la organización? Diferentes compañías han adoptado diferentes enfoques para la estructura de informes de DBA, pero algunas jerarquías de informes son bastante comunes. No hay una respuesta correcta, pero algunas estructuras de informes funcionan mejor que otras. Repasemos algunas de las posibilidades.

Una de las mejores estructuras es crear un grupo de administración de recursos de datos (DRM, por sus siglas en inglés) que esté compuesto por todos los especialistas en información y datos de la organización: DA, DBA, analistas de datos, analistas de rendimiento, etc. Este grupo generalmente informa directamente al CIO (Chief Information Officer), pero puede informar a través de una unidad de programación de sistemas, el centro de datos o el soporte técnico. La siguiente imagen muestra una estructura de reporte típica.

Cuando una organización cuenta con DBA de aplicación, se distribuirán en grupos de aplicaciones, por lo general con una línea de informe directo a los gerentes de programación de negocios. Cada equipo de desarrollo de aplicaciones tiene un recurso DBA de aplicaciones dedicado, como se muestra en la siguiente imagen

Sin embargo, hay problemas con estas dos estructuras de informes. El primer problema es que el DRM se debe colocar más arriba en la jerarquía de informes de TI. Es una buena idea que el grupo DRM informe directamente al CIO. Cuando una organización entiende la importancia de los datos para la salud de la organización, se recomienda colocar un DRM en este nivel.

Además, cuando existen DBA de aplicación, no deben reportar únicamente al administrador de programación de aplicaciones. Una línea secundaria de informes al grupo DRM garantizará que las habilidades de DBA se compartan y se comuniquen en toda la organización. La siguiente imagen delinea la estructura de informes recomendada para el grupo de administración de recursos de datos.

Cuestiones Multiplataforma para el DBA

La administración de un entorno multiplataforma complica el trabajo de la administración de la base de datos. Un conjunto completo de diferentes problemas y cuestiones surgen que deben ser abordados. La primera tarea es definir el alcance del trabajo de cada DBA. ¿Un solo DBA administra todos los distintos DBMS, o cada DBA se enfoca en soportar solo un DBMS?

Este es un tema particularmente espinoso. Por un lado, la funcionalidad de un DBMS es sorprendentemente similar independientemente de la plataforma y el proveedor. Un DBMS está diseñado para almacenar, recuperar y proteger datos. Los programadores, los programas y los usuarios finales interactúan con el DBMS para acceder y modificar los datos. Los problemas de administración son similares (diseño, creación, optimización, etc.), aunque cada DBMS implementa estos elementos de manera diferente. Por lo tanto, se puede afirmar que un DBA debe admitir múltiples DBMS y bases de datos, independientemente de la plataforma o el proveedor.

Por otro lado, cada DBMS ofrece diferentes características, funcionalidad y tecnología. Mantener todas las diferencias y matices rectos es una tarea monumental. ¿No sería mejor desarrollar DBA expertos en plataformas? De esa manera, sus DBA de Oracle pueden enfocarse en aprender todo lo que hay que saber sobre Oracle, sus DBA de DB2 pueden enfocarse en DB2, etc.

Cada organización tendrá que tomar esta determinación basándose en su mezcla particular de DBMS, características y talento DBA. Si su organización usa un DBMS predominantemente, con un uso limitado de otros, puede tener sentido que un solo DBA los respalde a todos, independientemente de la plataforma o el proveedor. El uso disperso de un DBMS generalmente significa menos problemas y, potencialmente, un menor uso de sus funciones más sofisticadas. Al hacer que sus DBA sean de múltiples DBMS y multiplataforma, puede asegurarse de que los DBA más capacitados en su departamento estén disponibles para todos los problemas de administración de bases de datos. Si su organización utiliza muchos DBMS diferentes, probablemente sea prudente crear DBA especializados para las plataformas muy utilizadas y tal vez compartir las tareas de administración para las plataformas menos utilizadas entre otros DBA.

Cuando se compartan las tareas de DBA, asegúrese de documentar cuidadosamente las habilidades y el nivel de conocimiento de cada DBA para cada DBMS que se admite. Tenga cuidado de establecer una rotación de llamada efectiva y justa que no suponga una carga excesiva para ningún DBA o grupo de DBA en particular. Además, utilice la estructura organizativa para promover el intercambio de normas y procedimientos de bases de datos en todos los entornos DBMS compatibles.

También tenga en cuenta que cuando se admiten múltiples DBMS y plataformas, debe considerar la implementación de herramientas DBA, monitores de rendimiento y scripts que puedan abordar múltiples plataformas. Por esta razón, las herramientas de DBA de proveedores externos suelen ser mejores para entornos heterogéneos que las herramientas similares ofrecidas por los proveedores de DBMS.

Cuando su organización admite varios DBMS, el grupo de DBA debe desarrollar pautas para saber qué DBMS se debe usar en qué situaciones. Estas pautas no deben ser reglas duras y rápidas, sino que deben proporcionar una guía para los tipos de aplicaciones y bases de datos mejor admitidas por cada DBMS. Forzar las aplicaciones en un entorno DBMS determinado no es una buena práctica. Las directrices deben usarse simplemente para asegurar el mejor ajuste de la aplicación a DBMS. Estas pautas deben tener en cuenta

- Características de cada DBMS
- Características y características del sistema operativo.
- Capacidades de red de la combinación del DBMS y el sistema operativo.
- Habilidades en el DBMS de los desarrolladores de aplicaciones.
- Soporte de lenguaje de programación
- Cualquier otra cuestión y requisitos organizativos.

Producción vs Pruebas (Desarrollo)

Se deben crear y admitir al menos dos entornos separados para una implementación de calidad de base de datos: producción y prueba (o desarrollo). Los nuevos trabajos de desarrollo y mantenimiento se realizan en el entorno de prueba; Las aplicaciones de funcionamiento operativo se ejecutan en el entorno de producción. Es necesario separar completamente el entorno de prueba del entorno de producción para garantizar la integridad y el rendimiento del trabajo operativo. Si no se separan las pruebas y la producción, las actividades de desarrollo perjudicarán los asuntos cotidianos de su organización. Lo último que se desea es que el código de un programa errante (con errores) en las primeras etapas de desarrollo acceda o modifique los datos de producción. El acceso de prueba a los datos de producción puede causar problemas de rendimiento de producción. Y por supuesto, los programas de prueba que modifican los datos de producción pueden crear datos no válidos.

La separación de los entornos de prueba y producción garantiza la integridad y el rendimiento del trabajo operativo.

El entorno de prueba no necesita ser exactamente el mismo que el entorno de producción. El entorno de producción contendrá todos los datos necesarios para admitir las aplicaciones operativas. Sin embargo, el entorno de prueba solo puede contener un subconjunto de los datos necesarios para facilitar las pruebas de aplicación aceptables. Además, la implementación del DBMS de prueba generalmente no se configurará con la misma cantidad de recursos que el entorno de producción. Por ejemplo, se asignará menos memoria al almacenamiento en búfer y las memorias caché, las asignaciones de conjuntos de datos serán más pequeñas y en menos dispositivos, y el software DBMS puede ser una versión posterior en prueba que en producción (para eliminar cualquier error en el propio código DBMS antes se confía para ejecutar en producción).

Sin embargo, los entornos de prueba y producción se deben juntar de manera similar. El acceso al mismo software del sistema se debe proporcionar en la prueba que se proporciona a la

producción, ya que el personal de programación deberá crear sus aplicaciones en el mismo tipo de entorno en el que se ejecutarán.

Otra diferencia en el entorno de prueba es el número de bases de datos. Es posible que se deban crear múltiples copias de bases de datos para admitir el desarrollo simultáneo de múltiples programadores. El DBA debe planificar y crear este entorno de tal manera que permita al personal de programación controlar el contenido de las bases de datos de prueba. Durante el proceso de desarrollo, los programas que modifican los datos en la base de datos pueden necesitar ejecutarse varias veces. El programador debe poder garantizar, sin embargo, que los datos al comienzo de cada ejecución de prueba son los mismos. De lo contrario, puede invalidar los resultados de la prueba. Por lo tanto, el DBA debe ayudar al personal de programación en la creación de trabajos de carga y descarga de bases de datos para configurar bases de datos de prueba para pruebas de ejecución. Antes de una ejecución de prueba, la base de datos debe estar cargada con los datos de prueba. Después de la ejecución de la prueba, el programador puede examinar la salida del programa y el contenido de la base de datos para determinar si la lógica del programa es correcta. Si no, puede repetir el proceso, cargar para restablecer los datos en la base de datos y volver a realizar la prueba. Se pueden implementar procedimientos automatizados para descargar las bases de datos afectadas por el programa y comparar los resultados con los archivos de carga.

Tratar de predecir el rendimiento de las aplicaciones de prueba una vez que pasan a producción es una tarea difícil. Pero el DBA puede ayudar aquí también. Un DBMS relacional generalmente proporciona un método para recopilar información estadística sobre el contenido de sus bases de datos. El optimizador relacional utiliza estas estadísticas para determinar cómo SQL recuperará los datos. “Pero recuerde, habrá mucho menos datos en las bases de datos de prueba que en la producción. Sin embargo, en algunos casos, el DBA puede configurar secuencias de comandos para leer las estadísticas de producción y copiarlas en el entorno de prueba, lo que permite a los desarrolladores evaluar con mayor precisión el rendimiento de las aplicaciones de prueba en la producción.

Algunas organizaciones implementan más de dos entornos, como se muestra en la siguiente imagen. Si se necesita atención especial para proyectos complejos de desarrollo de aplicaciones, es posible que se deban realizar niveles adicionales de pruebas aisladas. Por ejemplo, puede existir un entorno de prueba unitaria para el desarrollo de programas individuales, luego un entorno de pruebas de integración para garantizar que los nuevos programas trabajen juntos o que los nuevos programas funcionen correctamente con los programas existentes. Es posible que deba establecerse un entorno de control de calidad (QA) para realizar pruebas rigurosas contra programas nuevos y modificados antes de migrar al entorno de producción.

El Impacto de las Nuevas Tecnologías en el DBA

El DBA está en el centro de la acción cada vez que se presentan nuevas formas de hacer negocios y nuevas tecnologías a la organización. Los datos están en el corazón de cualquier aplicación, y la mayoría de las nuevas tecnologías impactan los datos a medida que son adoptados por los desarrolladores de aplicaciones. De hecho, los datos son el alma de los negocios modernos; la base

de datos alberga los datos; y el DBA es el experto que entiende la tecnología de base de datos, y en particular, cómo las bases de datos pueden integrarse con otras nuevas tecnologías.

Examinemos tres nuevas tecnologías específicas que dependen de la administración de la base de datos, al menos en cierta medida, para ser implementadas de manera efectiva: lógica de aplicación acoplada a la base de datos, desarrollo de comercio electrónico habilitado para Internet y computación de mano.

DBA Procedural: Manejo de la lógica de Bases de Datos

Tradicionalmente, el dominio de un sistema de administración de bases de datos era almacenar, administrar y acceder a los datos, de manera adecuada. Si bien estas capacidades básicas aún son necesarias para los productos DBMS modernos, la funcionalidad de procedimiento adicional se está convirtiendo lentamente no solo en una característica agradable, sino en una necesidad. Las características tales como triggers, funciones definidas por el usuario y procedimientos almacenados brindan la capacidad de definir reglas de negocios para el DBMS en lugar de en programas de aplicación separados. Estas características combinan estrechamente la lógica de la aplicación con el servidor de la base de datos.

Dado que todos los productos RDBMS más populares proporcionan características a veces complejas para facilitar la lógica de procedimiento acoplada a la base de datos, se requiere trabajo adicional para administrar y garantizar el uso óptimo de estas características. Esto requiere una expansión de la disciplina de gestión de la administración de bases de datos. Normalmente, a medida que se agregan nuevas funciones, la administración, el diseño y la administración de estas funciones se asignan al DBA de forma predeterminada. Sin una planificación y preparación adecuadas, esto puede llevar al caos. Pero primero examinemos rápidamente cómo se almacena la lógica de la base de datos en un DBMS.

Procedimientos Almacenados

Los procedimientos almacenados se pueden considerar como programas que viven en una base de datos. La lógica de procedimiento de un procedimiento almacenado se mantiene, administra y ejecuta a través de los comandos de la base de datos. La razón principal para usar procedimientos almacenados es mover el código de la aplicación desde una estación de trabajo cliente al servidor de la base de datos. Los procedimientos almacenados suelen consumir menos gastos generales en un entorno cliente / servidor porque un cliente puede invocar un procedimiento almacenado que hace que se ejecuten varias sentencias de SQL. La alternativa, el cliente que ejecuta varias declaraciones SQL directamente, es menos eficiente porque aumenta el tráfico de red, lo que puede degradar el rendimiento general de la aplicación. Un procedimiento almacenado es un objeto de base de datos independiente; no está "físicamente" asociado con ningún otro objeto en la base de datos. Un procedimiento almacenado puede acceder y/o modificar datos en muchas tablas.

Triggers (Desencadenadores)

Los desencadenadores son procedimientos especializados controlados por eventos que se adjuntan a las tablas de base de datos. El RDBMS ejecuta automáticamente el código de activación a medida que los datos cambian en la base de datos. Cada activador se adjunta a una única tabla especificada. Los desencadenantes pueden considerarse como una forma avanzada de regla o restricción escrita mediante lógica de procedimiento. Un disparador no puede ser directamente llamado o ejecutado; es ejecutado automáticamente (o "activado") por el RDBMS como resultado de una instrucción SQL INSERT, UPDATE o DELETE que se emite en su tabla asociada. Una vez que se crea un activador, siempre se ejecuta cuando se produce el evento de "activación".

Funciones Definidas por el Usuario (UDF)

Una UDF, o función definida por el usuario, proporciona un resultado basado en un conjunto de valores de entrada. Las UDF son programas que se pueden ejecutar en lugar de funciones escalares o de columnas SQL estándar incorporadas. Una función escalar transforma los datos para cada fila de un conjunto de resultados; una función de columna evalúa cada valor para una columna particular en cada fila del conjunto de resultados y devuelve un solo valor. Una vez escrito, y definido en el RDBMS, una UDF está disponible como cualquier otra función de base de datos integrada.

La siguiente imagen resume las diferencias entre los procedimientos almacenados, los activadores y las funciones definidas por el usuario.

Una vez que las aplicaciones y los desarrolladores comienzan a confiar en procedimientos almacenados, desencadenadores y UDF, se deben tomar medidas para garantizar que se administran correctamente. Los DBA deben lidiar con los problemas de calidad, facilidad de mantenimiento, eficiencia y disponibilidad. ¿Cómo y cuándo serán probados estos objetos de procedimiento? El impacto de una falla es a nivel de toda la empresa, no relegado a una sola aplicación. Esto aumenta la visibilidad y criticidad de estos objetos. ¿Quién es responsable si fallan? La respuesta debe ser: un DBA.

La función de administrar la lógica de la base de datos de procedimientos debe recaer en alguien experto en esa disciplina. Se requiere un nuevo tipo de DBA para acomodar la administración de la lógica de procedimiento de la base de datos. Este nuevo rol puede definirse como un DBA de procedimiento.

El DBA de procedural debe ser responsable de las actividades de administración de base de datos que requieren soporte de lógica de procedimiento. Esto debe incluir la responsabilidad principal de garantizar que los procedimientos almacenados, los desencadenadores y las funciones definidas por el usuario se planifiquen, implementen, compartan y reutilicen de manera efectiva. El DBA de procedimiento también debe asumir la responsabilidad principal de codificar y probar todos los desencadenadores. Sin embargo, lo más probable es que los procedimientos almacenados y las funciones definidas por el usuario sean codificados por los programadores de la aplicación y revisados por los DBA procedurales para verificar su precisión y rendimiento.

El DBA procedural debe participar y liderar la revisión y administración de todos los objetos de bases de datos de procedimientos: es decir, desencadenadores, procedimientos almacenados y UDF. Si bien es poco probable que los administradores de bases de datos de procedimientos sean tan expertos en programación como los programadores de aplicaciones o analistas de sistemas, deben poder escribir y revisar el código del programa razonablemente bien. El nivel de habilidad requerido depende de qué idiomas soporta el DBMS para crear objetos de procedimiento, la tasa y el nivel de adopción dentro de la organización y si existe o no una organización interna para crear programas comunes y reutilizables. La siguiente tabla proporciona un nivel razonable de participación de DBA procedural para cada tipo de objeto procedural. Además, el DBA procedural debe estar disponible para cualquier problema que se produzca en los objetos de procedurales de base de datos en producción.

El papel del DBA procedural requiere habilidades de comunicación tanto como requiere perspicacia tecnológica (imagen siguiente). Además de administrar y optimizar los objetos de procedimiento de la base de datos, el DBA de procedimiento debe informar a la comunidad de desarrollo sobre nuevos desencadenadores, procedimientos almacenados y UDF. Además, el DBA debe promover la reutilización. Si los programadores no saben que estos objetos existen, nunca serán utilizados. Se pueden asignar otras funciones administrativas de procedimiento al DBA procedural. Dependiendo de la cantidad de DBA y la cantidad de desarrollo de la aplicación que se está realizando, el DBA procedural se puede asignar a funciones adicionales, como

- Participar en revisiones de diseño de códigos de aplicación.
- Revisar y analizar las rutas de acceso de SQL (desde EXPLAIN o SHOWPLAN)
- Depuración de SQL
- Escritura y análisis de sentencias complejas de SQL.
- Reescritura de consultas para una ejecución óptima.

La descarga de tareas relacionadas con la codificación al DBA procedural puede ayudar a los demás equipos de DBAs a concentrarse en el diseño físico real y la implementación de las bases de datos, lo que da como resultado bases de datos mucho mejor diseñadas. El DBA procedural todavía debe reportarse a través de la misma unidad de administración que el DBA tradicional. Al hacerlo, permite un mejor intercambio de habilidades entre los DBA procedurales y los DBA tradicionales centrados en datos. Por supuesto, será necesario que haya una mayor sinergia entre el DBA procedural y los programadores de aplicaciones. La ruta de trabajo típica para el DBA procedural debe ser desde los rangos de programación de la aplicación porque aquí es donde existe la base de habilidades de codificación.

El Internet: del DBA al eDBA

Aunque en estas fechas, Internet difícilmente puede considerarse una tendencia, las empresas y los tecnólogos aún están adaptando sus procesos para alinearse con el comercio electrónico. Las organizaciones de todo tipo y tamaño utilizan tecnologías de Internet para acelerar los procesos

de negocios, y las prácticas y procedimientos de administración de bases de datos se ven afectados por la adopción de bases de datos y aplicaciones habilitadas para Internet.

Los negocios electrónicos deben ser capaces de adaptarse y reaccionar a los cambios constantes. Cuando su negocio está en línea, nunca se cierra. Las personas esperan una funcionalidad completa en los sitios web que visitan, independientemente del momento. Y la web es mundial. Puede que sean las dos de la mañana en la ciudad de Nueva York, pero siempre es la hora pico en algún lugar del mundo. Un negocio electrónico debe estar disponible y preparado para interactuar con los clientes las 24 horas del día, los 365 días del año (366 durante los años bisiestos). No hacerlo puede causar la pérdida de negocio. Cuando un sitio web está inactivo, el cliente irá a otra parte para hacer negocios porque la competencia está a solo un clic del mouse. Por lo tanto, aquellos que manejan un negocio electrónico deben ser expertos, proactivos y siempre vigilantes.

Adoptar el ritmo frenético de un e-business requiere cambios para quienes mantienen el e-business en funcionamiento. DBAs son extremadamente afectados por e-business. La necesidad de integrar la Web con los servicios de TI tradicionales, como el DBMS, coloca altas expectativas en los administradores de bases de datos.

Un DBA que es capaz de administrar aplicaciones basadas en web porque entiende los problemas especiales que surgen debido a Internet es un eDBA. Un eDBA también necesita tener todo el conocimiento y la capacitación de un DBA tradicional, pero estas habilidades deben adaptarse para adaptarse a las aplicaciones y bases de datos que están habilitadas para Internet. Cuando la web se combina con aplicaciones y bases de datos tradicionales, el resultado es una infraestructura compleja (ver la siguiente imagen). El eDBA debe ser capaz de navegar esta infraestructura compleja y heterogénea y proporcionar experiencia donde las bases de datos interactúan dentro de esta infraestructura.

De hecho, hay muchos factores que afectan la administración de la base de datos cuando Internet se combina con la tecnología de la base de datos. Algunas de estas cuestiones incluyen

- Disponibilidad de datos 24/7/52
- Adopción de nuevas tecnologías como Java y XML.
- Conectividad a la web.
- Integración de datos heredados con aplicaciones modernas basadas en web.
- Base de datos y arquitectura de aplicaciones.
- Administración basada en web
- Ingeniería de performance para internet.
- Carga de trabajo impredecible

El DBA Personal y la Nube

Los dispositivos personales, generalmente teléfonos inteligentes pero también PDA (asistentes digitales personales), se están convirtiendo rápidamente en una necesidad para los ejecutivos y empresarios modernos. Un teléfono inteligente es un dispositivo informático de mano. Y a veces tendrá un sistema de gestión de base de datos ejecutándose en él. ¿Por qué es eso interesante? ¿Cambia la forma en que usará su dispositivo? ¿Qué significará eso para su departamento de TI?

Las plataformas populares de computación móvil incluyen Symbian OS, Windows Mobile, iOS y Android.

Los teléfonos inteligentes ofrecen muchos beneficios. Los dispositivos son pequeños y por lo tanto fácilmente transportables. No interfieren con la capacidad de un trabajador móvil para ser móvil. Y debido a que la mayoría de las personas tienen un teléfono móvil en estos días, aumentar sus capacidades para realizar tareas informáticas es una tarea fácil.

Quizás el mayor beneficio de estos dispositivos es su capacidad para ejecutar aplicaciones móviles. No es infrecuente que una aplicación móvil empresarial se base en que la información recuperada de un servidor de computadora principal se envíe al dispositivo móvil de manera generalizada, en cualquier lugar y en cualquier momento en que sea necesario. Las aplicaciones móviles que utilizan la nube como back-end son cada vez más populares.

¿Qué es el cómputo en la nube?

La computación en la nube ofrece un nuevo modelo para la entrega de recursos de TI a los usuarios. La principal característica definitoria de la computación en la nube es dar la ilusión de acceso a petición a una cantidad infinita de recursos informáticos. Un buen ejemplo de un servicio de computación en la nube lo ofrece Salesforce.com, que brinda acceso a una aplicación de CRM a través de la web.

Otro aspecto que prevalece con las ofertas de computación en la nube es que los usuarios pueden alquilar el poder de cómputo sin compromiso. En lugar de comprar un servidor, puede alquilar el uso de uno y pagar solo por lo que usa. Esto solía denominarse informática de servicios públicos porque imita la forma en que las personas pagan los servicios públicos, como el agua o la electricidad. Es un servicio de "pago por uso".

Desde la perspectiva de la base de datos, hay varias ofertas en la nube, incluyendo SimpleDB de Amazon y App Engine Datastore de Google. Además, el SQL Azure de Microsoft admite datos en la nube.

La computación en la nube puede permitir que incluso las organizaciones más pequeñas, o incluso una sola persona, obtengan recursos informáticos de formas que antes no eran posibles.

El diseño e implementación de aplicaciones móviles no es tan sencillo como el desarrollo de aplicaciones de PC de escritorio. Es muy importante que los desarrolladores de aplicaciones móviles consideren el contexto en el que se utilizará la aplicación.

Si bien los beneficios son significativos, existen desafíos que enfrentar a medida que las organizaciones incorporan dispositivos personales en su infraestructura. Los datos en el dispositivo deben administrarse profesionalmente para garantizar la integridad y confiabilidad. Debido a que el dispositivo es remoto, compartir datos puede ser difícil. Los datos de negocios en un teléfono inteligente deben estar sincronizados de manera confiable con los sistemas y bases de datos empresariales existentes. Y desde una perspectiva empresarial y de cumplimiento, puede ser difícil

evaluar el riesgo asociado con una aplicación móvil. Los dispositivos móviles se extravían fácilmente y sin la seguridad adecuada pueden resultar en una violación de datos.

Todos los principales proveedores de DBMS proporcionan versiones de tamaño reducido de sus productos estrella para ejecutar en dispositivos personales. Por ejemplo, los mercados de IBM DB2 Personal Edition, Oracle vende Oracle Database Lite, Microsoft proporciona SQL Server Compact y Sybase ofrece Adaptive Server Anywhere. La idea general es almacenar una pequeña cantidad de datos críticos en el dispositivo móvil en una base de datos. La base de datos local se sincroniza posteriormente con los almacenes de datos a largo plazo en servidores de bases de datos empresariales. Cada DBMS móvil proporciona tecnología para sincronizar los datos desde el dispositivo móvil a las plataformas de servidores empresariales.

Impacto en los DBAs

Los DBA no son necesarios para administrar y trabajar en la base de datos en cada dispositivo móvil, pero el trabajo del DBA se verá afectado por este desarrollo. Una base de datos del tamaño de las almacenadas en un dispositivo móvil no debe requerir el ajuste y la administración en profundidad que se requieren para las implementaciones de bases de datos empresariales. Sin embargo, se solicitará a los DBA que ayuden a diseñar bases de datos implementadas de manera apropiada para dispositivos de factor de forma pequeño como dispositivos móviles. Pero este no es el mayor impacto.

Un gran impacto en el DBA está en la planificación y administración de la sincronización de datos de cientos o miles de PDA. ¿Cuándo debería programarse la sincronización? ¿Cómo afectará a las aplicaciones que usan grandes bases de datos de producción que están involucradas en la sincronización? ¿Cómo puede asegurarse de que los usuarios móviles sincronizarán sus datos de manera confiable y según lo programado?

Además, para las implementaciones en la nube, los DBA probablemente serán responsables de garantizar la disponibilidad confiable de los datos. El diseño y ajuste de una implementación de base de datos para la computación en la nube puede requerir recursos importantes para administrar grandes cantidades de datos y asegurar la disponibilidad las 24 horas del día.

Estos no son problemas menores. Antes de implementar un gran batallón de usuarios de bases de datos móviles que deben sincronizar sus datos, asegúrese de que el equipo de DBA esté preparado para el impacto en sus bases de datos. Como casi todo, la falta de preparación es una receta para un desastre seguro. Pero debemos prepararnos. El equipo de DBA debe estar preparado para brindar soporte a la fuerza laboral móvil al comprender la tecnología de sincronización de datos, la computación en la nube y la posible necesidad de usuarios de bases de datos remotas en su organización.

Un paso adicional en el proceso de preparación para el soporte de bases de datos en dispositivos portátiles es revisar las aplicaciones en su organización e intentar determinar cuáles podrían verse afectadas primero. Las empresas con trabajadores remotos, como una fuerza de ventas distribuida o servicios de seguimiento de entregas, probablemente serán las primeras afectadas. Tómese un

tiempo para revisar los requisitos de datos de esas aplicaciones y cómo una gran afluencia de conexiones remotas podría afectar los sistemas actuales.

La informática de mano y la fuerza laboral móvil están aquí para quedarse. Y el equipo de DBA debe estar preparado para apoyar a estos trabajadores móviles con una infraestructura de datos compartida válida.

NoSQL, Big Data y el DBA

NoSQL es otra tendencia que puede afectar el trabajo realizado por el DBA. NoSQL es un movimiento que, en su forma más básica, se describe por su título. Es decir, un DBMS NoSQL no es compatible con SQL. En un nivel alto, NoSQL implica no relacional, distribuido, flexible y escalable. La mayoría de las ofertas de NoSQL también son de código abierto.

NoSQL surgió de la necesidad percibida de sistemas de bases de datos "modernos" para apoyar iniciativas web. Además, algunos atributos comunes de los DBMS de NoSQL incluyen la falta de un esquema, la simplicidad de uso, el soporte de replicación y una capacidad "eventualmente consistente" (en lugar de la ACID típica: atomicidad, consistencia, aislamiento y durabilidad - capacidad de transacción). Realmente no significa que no hay soporte de SQL. Algunas ofertas de NoSQL han comenzado a ser compatibles con SQL, lo que lleva a algunos expertos a definir NoSQL como "No solo SQL".

Ejemplos de Ofertas de NoSQL

- Cassandra: <http://cassandra.apache.org/>
- CouchDB: <http://couchdb.apache.org/>
- HBase: <http://hbase.apache.org/>
- mongoDB: www.mongodb.org/
- Riak: www.basho.com/

El movimiento NoSQL está ligado al movimiento Big Data. Las bases de datos NoSQL están diseñadas para ofrecer almacenamiento de bajo costo y acceso a grandes cantidades de datos.

Con las implementaciones de NoSQL, los datos generalmente se acceden de una manera y hay poca o ninguna flexibilidad en términos de generar consultas ad hoc y sin complicaciones.

Impacto en los DBAs

Otro sello distintivo de NoSQL es que requiere muy poco trabajo de administración de base de datos. Por supuesto, el DBMS se debe configurar y administrar, y se debe hacer una copia de seguridad de los datos. Así que tenga cuidado cuando alguien trate de decirle que cualquier sistema de base de datos no requiere administración de base de datos. Simplemente no es cierto.

Impacto de las Nuevas Tecnologías en el DBA

A medida que se introduce nueva tecnología en la organización, el grupo de DBA suele ser el primero en examinarla y utilizarla. Las tecnologías anteriores son simplemente ejemplos de tendencias y tecnologías recientes que requieren la administración de bases de datos para una implementación eficiente y efectiva. La mayoría de las nuevas tecnologías tendrán algún impacto en el papel del DBA.

Certificación para DBA

La certificación profesional es una tendencia continua en TI y está disponible para muchos trabajos de TI diferentes. La disponibilidad y los niveles de certificación han crecido a un ritmo alarmante para la administración de bases de datos. Los programas de certificación están disponibles para la mayoría de las plataformas DBMS populares, incluyendo IBM DB2, Microsoft SQL Server y Oracle Database. El concepto detrás de la certificación de DBA es certificar que una persona es capaz de realizar funciones y tareas de administración de base de datos.

Este es un objetivo noble, pero el problema es que pasar una prueba no es un indicador viable de poder realizar un trabajo complejo como DBA. Algunas cosas que solo tienes que aprender haciendo. Ahora no estoy diciendo que la certificación sea inútil. De hecho, tomar la prueba y concentrarse en las preguntas que no responde puede ayudarlo a señalar las áreas de debilidad en las que puede mejorar. Pero, ¿alguien realmente cree que alguien que pasa una prueba formal será tan capaz como alguien con varios años de experiencia como DBA? Las organizaciones deben contratar DBAs según la experiencia pasada que indique un nivel de capacidad. Por supuesto, alguien con experiencia y certificación es mejor que alguien con solo uno de los dos.

Recomiendo que los DBA profesionales se tomen el tiempo para estudiar y aprobar los exámenes de certificación, no porque la certificación lo haga un mejor DBA, sino porque lo hará más empleable. Algunas empresas contratarán solo profesionales certificados. La tendencia hacia el uso de la certificación para guiar las prácticas de contratación aumentará debido a la creciente complejidad de TI. Si crees que podrías cambiar de empleo en algún momento de tu carrera (¿y quién entre nosotros no lo hará?), La certificación es una búsqueda que vale la pena.

Tenga en cuenta que las pruebas de certificación de DBA a veces hacen preguntas de sintaxis arcanas que no son realmente buenos indicadores de las habilidades de un DBA. Obtener la sintaxis 100% precisa es para lo que sirven los manuales y las herramientas de DBA. No hay razón para memorizar la sintaxis porque tiende a cambiar con bastante frecuencia. Es mejor saber dónde encontrar la sintaxis, los parámetros y las respuestas a sus preguntas cuando las necesite, es decir, qué manuales y libros de texto contienen la información necesaria. Los DBA deben poseer un amplio conocimiento general de los conceptos DBMS y los fundamentos de TI, y un buen conocimiento de la forma en que funcionan los sistemas de bases de datos de su organización. Memorizar cada detalle sobre la sintaxis y la estructura de SQL es una pérdida de tiempo porque es complejo y cambia todo el tiempo. En otras palabras, es mejor saber desde el principio que algo se puede (o no) hacer que saber la sintaxis exacta de cómo lograrlo.

Si decide obtener la certificación, tómese el tiempo para prepararse para las pruebas. Hay libros y títulos de software de autoaprendizaje disponibles que pueden ser muy útiles. Estos libros y programas cubren los temas de prueba más probables y proporcionan preguntas de muestra para ayudarlo a prepararse. En muchos sentidos, es como prepararse para un examen de ingreso a la universidad, como los exámenes SAT.

Y una vez que obtenga su certificación, asegúrese de mostrarla con orgullo en su currículum y en su tarjeta de negocios (si su empresa lo permite).

Consulte la siguiente tabla para los sitios web que contienen información sobre la certificación profesional de los productos DBMS más populares.

Definir la Estrategia de DBMS para la Empresa

El proceso de elegir un DBMS adecuado para la gestión de bases de datos empresariales no es tan difícil como solía ser. El número de los principales proveedores de DBMS ha disminuido debido a la consolidación de la industria y la dominación del sector por parte de unos pocos actores muy grandes.

Sin embargo, las organizaciones grandes y medianas generalmente ejecutan varios productos DBMS, desde dos hasta diez. Por ejemplo, no es infrecuente que una empresa grande use IMS o IDMS y DB2 en el mainframe, Oracle y MySQL en varios servidores UNIX diferentes, Microsoft SQL Server en servidores Windows, así como en los dispositivos móviles de otros productos DBMS como Sybase, Ingres, Adabas y PostgreSQL en varias plataformas, por no mencionar los productos DBMS para PC de un solo usuario, como Microsoft Access, Paradox y FileMaker. ¿Quién eligió instalar todos estos DBMS y por qué?

Desafortunadamente, a menudo la respuesta es que no se pensó mucho ni se planificó el proceso de toma de decisiones. A veces, la decisión de comprar e instalar un nuevo DBMS es impulsada por una necesidad comercial o una nueva aplicación. Esto es razonable si su organización no tiene DBMS y debe comprar uno por primera vez. Sin embargo, esto rara vez es el caso.

Independientemente de si existe un DBMS en el sitio, un nuevo DBMS a menudo se considera un requisito para una nueva aplicación. A veces se compra e instala un nuevo producto DBMS sin examinar primero si la aplicación podría implementarse con éxito utilizando un DBMS existente. O, más probablemente, los DBA saben que la aplicación puede implementarse utilizando un DBMS existente pero que carecen del poder o el soporte organizativo para rechazar una nueva propuesta de DBMS.

Hay otras razones para la existencia de múltiples plataformas DBMS en una sola organización. Tal vez la compañía compró un paquete de aplicación comercial que no se ejecuta en ninguna de las plataformas DBMS actuales. A veces, la decisión de comprar un nuevo DBMS es impulsada por el deseo de apoyar la última y mejor tecnología. Por ejemplo, muchas tiendas de mainframe que se mueven desde un modelo de base de datos jerárquico (IMS) o CODASYL (IDMS) al modelo relacional implementado DB2, dan como resultado un DBMS adicional para aprender y respaldar. Luego, cuando la computación cliente / servidor se hizo popular, se implementaron DBMS adicionales en servidores UNIX, Linux y Windows.

Una vez que se instala un DBMS, la eliminación puede ser difícil debido a las incompatibilidades entre los diferentes DBMS y la necesidad de convertir el código de la aplicación. Además, cuando se instala un nuevo DBMS, las aplicaciones y bases de datos antiguas generalmente no se migran a él. El antiguo DBMS permanece y debe seguir siendo compatible. Esto complica el trabajo del DBA.

¿Entonces qué debería ser hecho? Bueno, el grupo de DBA debe estar facultado para tomar las decisiones de DBMS para la organización. Ninguna unidad de negocios debe tener permitido comprar un DBMS sin el permiso del grupo de DBA. Esta es una disposición difícil de implementar y aún más difícil de hacer cumplir. La política empresarial a menudo funciona en contra del grupo de DBA porque con frecuencia posee menos poder organizativo que otros ejecutivos empresariales.

Selección de un DBMS

El grupo de DBA debe establecer una política con respecto a los productos DBMS que se admitirán dentro de la organización. Siempre que sea posible, la política debe minimizar el número de productos DBMS diferentes. Para una departamento con múltiples sistemas operativos y múltiples tipos de hardware, elija un DBMS predeterminado para la plataforma. Desaliente la desviación del valor predeterminado a menos que exista un caso de negocio convincente, un caso de negocio que pase la inspección técnica del grupo DBA.

La mayoría de los principales productos de DBMS tienen características similares, y si la característica o funcionalidad no existe hoy, probablemente lo hará dentro de 18 a 24 meses. Por lo tanto, tenga cuidado antes de decidir elegir un DBMS basado únicamente en su capacidad de soportar una función específica.

Al elegir un DBMS, es aconsejable seleccionar un producto de un proveedor de nivel 1 como se indica en la siguiente tabla. El Nivel 1 representa a los proveedores más grandes que tienen los productos más implementados y soportados en el mercado. No puedes equivocarte con DB2 u Oracle. Ambos son populares y soportan casi cualquier tipo de base de datos. Otro jugador importante es Microsoft SQL Server, pero solo para plataformas Windows. DB2 y Oracle se ejecutan en múltiples plataformas que van desde mainframe a UNIX, así como Windows e incluso dispositivos de mano. La elección de un DBMS diferente a estos tres debe hacerse solo bajo circunstancias específicas.

Después de los tres grandes vienen MySQL, Sybase, Teradata e Informix. La siguiente tabla enumera estos proveedores de DBMS de nivel 2. Todas estas ofertas son productos DBMS de calidad, pero su base instalada es más pequeña, sus productos están diseñados y comercializados para fines específicos, o las empresas son más pequeñas con menos recursos que los Tres Grandes (IBM, Oracle y Microsoft), por lo que hay cierto riesgo al elegir un DBMS de nivel 2 en lugar de nivel 1. Sin embargo, puede haber razones sólidas para implementar una solución de nivel 2, como el alto rendimiento ofrecido por Informix o las capacidades de análisis y almacenamiento de datos de Teradata.

Por supuesto, hay otros productos de DBMS en el mercado, muchos de los cuales son productos finos y dignos de consideración para el procesamiento especializado, ciertas necesidades predefinidas y roles de nicho. Si su empresa está muy involucrada en el movimiento de software de código abierto, PostgreSQL, EnterpriseDB o MySQL podrían ser opciones viables. Si un objeto DBMS es importante para un proyecto específico, puede considerar ObjectDesign o Versant. Y hay una variedad de ofertas de DBMS NoSQL disponibles, como Hadoop, Cassandra y MongoDB.

Sin embargo, para la mayor parte de sus necesidades de gestión de datos, un DBMS de nivel 1, o tal vez de nivel 2, el proveedor de DBMS ofrecerá suficiente funcionalidad con un riesgo mínimo. Una gran cantidad de productos DBMS están disponibles, cada uno con ciertas características que los hacen dignos de ser considerados caso por caso. La elección de cualquiera de los candidatos de nivel inferior, incluso los nombres más importantes como Adabas de Software AG e Ingres de Actian, implica un riesgo adicional.

No quiero que suene como si la selección de un DBMS fuera una obviedad. Necesitará una estrategia y un plan para seleccionar el DBMS adecuado para su situación específica. Al elegir un DBMS, asegúrese de considerar cada uno de estos factores:

- Soporte del sistema operativo. ¿El DBMS es compatible con los sistemas operativos en uso en su organización, incluidas las versiones que está utilizando actualmente y planea usar?
- Tipo de organización. Tenga en cuenta la filosofía corporativa cuando elija un DBMS. Algunas organizaciones son muy conservadoras y les gusta controlar sus entornos; estas organizaciones tienden a gravitar hacia entornos de mainframe tradicionales. Las operaciones del gobierno, las instituciones financieras y las compañías de seguros y salud suelen ser conservadoras. Las organizaciones más liberales a menudo están dispuestas a considerar arquitecturas alternativas. No es raro que las empresas manufactureras, las empresas punto-com y las universidades sean menos conservadoras. Finalmente, algunas empresas simplemente no confían en Windows como un entorno de misión crítica y prefieren usar UNIX; esto descarta algunos proveedores de bases de datos (Microsoft SQL Server, en particular).
- Puntos de referencia. ¿Qué puntos de referencia de rendimiento están disponibles en el proveedor de DBMS y otros usuarios del DBMS? El Transaction Processing Performance Council (TPC) publica puntos de referencia oficiales de rendimiento de la base de datos que se pueden utilizar como una guía para el rendimiento general básico de muchos tipos diferentes de procesamiento de bases de datos. (Para obtener más detalles, consulte la barra lateral "El Consejo de Rendimiento de Procesamiento de Transacción - The Transaction Processing Performance Council"). En general, los puntos de referencia de rendimiento pueden ser útiles como un indicador amplio del rendimiento de la base de datos, pero no deben ser el único factor determinante al seleccionar un DBMS. Muchos de los puntos de referencia de TPC se ejecutan contra implementaciones de bases de datos que no son representativas de la mayoría de los sistemas de bases de datos de producción y, por lo tanto, no son indicativos del rendimiento real de un DBMS en particular. Además, los puntos de referencia se actualizan constantemente para mostrar mediciones de rendimiento nuevas y mejoradas para cada uno de los principales productos DBMS, lo que hace que los "ganadores" de los puntos de referencia se vuelvan obsoletos muy rápidamente.

Los puntos de referencia se actualizan constantemente para mostrar mediciones de rendimiento nuevas y mejoradas.

- Escalabilidad. ¿El DBMS admite la cantidad de usuarios y tamaños de base de datos que pretende implementar? ¿Cómo se construyen, soportan y mantienen las grandes bases de datos, fácilmente o con mucho dolor? ¿Hay usuarios independientes que puedan confirmar las reclamaciones de escalabilidad del proveedor de DBMS?
- Disponibilidad de herramientas de software de apoyo. ¿Están disponibles las herramientas de soporte que necesita para el DBMS? Estos elementos pueden incluir herramientas de consulta y análisis, herramientas de soporte de almacenamiento de datos, herramientas de administración de bases de datos, herramientas de respaldo y recuperación, herramientas de monitoreo de rendimiento, herramientas de planificación de capacidad, utilidades de base de datos y soporte para varios lenguajes de programación.
- Técnicos. ¿Hay un suministro suficiente de profesionales de bases de datos capacitados para el DBMS? Considere sus necesidades en términos de DBA, personal de soporte técnico (programadores y administradores de sistemas, analistas de operaciones, etc.) y programadores de aplicaciones.
- Costo de propiedad. ¿Cuál es el costo total de propiedad del DBMS? Los proveedores de DBMS cobran precios muy variados por su tecnología. El costo total de propiedad se debe calcular como una combinación del costo de la licencia del DBMS; el costo de la licencia de cualquier software de soporte requerido; el costo de los profesionales de bases de datos para programar, dar soporte y administrar el DBMS; y el costo de los recursos informáticos necesarios para operar el DBMS.
- Programa de lanzamientos. ¿Con qué frecuencia el proveedor de DBMS lanza una nueva versión? Algunos proveedores tienen ciclos de lanzamiento rápidos, con nuevos lanzamientos cada 12 a 18 meses. Esto puede ser bueno o malo, dependiendo de su enfoque. Si desea funciones de vanguardia, un ciclo de lanzamiento rápido es bueno. Sin embargo, si su tienda es más conservadora, un DBMS que cambia con frecuencia puede ser difícil de soportar. Un ciclo de lanzamiento rápido hará que las organizaciones conservadoras se actualicen con más frecuencia de la que quisieran o que vivan con un software DBMS obsoleto que probablemente no tenga el mismo nivel de soporte que las versiones más recientes.

Al elegir un DBMS, asegúrese de tener en cuenta la complejidad de los productos. El software DBMS es muy complejo y se está volviendo más complejo con cada nueva versión. La funcionalidad que solía ser compatible solo con software adicional o programas independientes se está agregando cada vez más como funciones del DBMS, como se muestra en la siguiente imagen. Deberá planificar y admitir todas las funciones del DBMS. Incluso si no hay un requisito actual para ciertas características, una vez que implemente el DBMS, los programadores y desarrolladores encontrarán una razón para usar casi cualquier cosa que el proveedor haya incluido. Es mejor planificar y estar preparado que permitir que se utilicen las características sin un plan para darles soporte.

Arquitecturas de DBMS

La arquitectura de soporte para el entorno DBMS es muy crítica para el éxito de las aplicaciones de base de datos. Una elección incorrecta o un componente mal implementado de la arquitectura general puede causar un rendimiento deficiente, tiempo de inactividad o aplicaciones inestables.

Cuando los mainframes dominaban la informática empresarial, la arquitectura DBMS era una preocupación más simple. Todo funcionaba en el mainframe, y eso era todo. Sin embargo, hoy la infraestructura de TI es distribuida y heterogénea. La arquitectura general, incluso para un DBMS de mainframe, probablemente consistirá de múltiples plataformas y software de sistema interoperativo. Un equipo formado por expertos en negocios e informática, en lugar de una sola persona o grupo, debe tomar la decisión final sobre la arquitectura. Los expertos en negocios deben incluir representantes de varios departamentos, así como de contabilidad y asuntos legales para contratos de software. Los representantes de la administración de la base de datos (DA, DBA y SA), así como los miembros del grupo de redes, los expertos en sistemas operativos, el personal de control de operaciones, los expertos en programación y cualquier otra parte interesada, deben incluirse en este equipo.

Además, asegúrese de que el DBMS que seleccione sea apropiado para la naturaleza y el tipo de procesamiento que planea implementar. Hay cuatro niveles de arquitectura DBMS disponibles: empresarial, departamental, personal y móvil.

Un DBMS empresarial está diseñado para la escalabilidad y el alto rendimiento. Un DBMS empresarial debe ser capaz de admitir bases de datos muy grandes, un gran número de usuarios concurrentes y múltiples tipos de aplicaciones. El DBMS de la empresa se ejecuta en una máquina a gran escala, normalmente un mainframe o un servidor de gama alta que ejecuta UNIX, Linux o Windows Server. Además, un DBMS empresarial ofrece todas las "alarmas" disponibles del proveedor de DBMS. El soporte multiprocesador, el soporte para consultas paralelas y otras características avanzadas de DBMS son componentes centrales de un DBMS empresarial.

Un DBMS departamental, a veces conocido como un DBMS de grupo de trabajo, sirve el término medio. El DBMS departamental admite grupos de trabajo pequeños a medianos dentro de una organización; Por lo general, se ejecuta en un servidor UNIX, Linux o Windows. La línea divisoria entre un servidor de base de datos departamental y un servidor de base de datos empresarial es bastante gris. Las actualizaciones de hardware y software pueden permitir que un DBMS departamental aborde tareas que anteriormente solo podían ser realizadas por un DBMS empresarial. La continua disminución del costo de los componentes de hardware y software del departamento contribuye aún más a reducir el costo total de operación y permite que un entorno de grupo de trabajo se amplíe para servir a la empresa.

Un DBMS personal está diseñado para un solo usuario, por lo general en una plataforma de PC de baja a mediana potencia. Microsoft Access, SQLite y FileMaker son ejemplos de software de bases de datos personales. Por supuesto, los principales proveedores de DBMS también comercializan versiones personales de sus soluciones de mayor potencia, como Oracle Database Personal Edition y DB2 Personal Edition. A veces, el bajo costo de un DBMS personal resulta en un intento equivocado de elegir un DBMS personal para una solución departamental o empresarial. Sin embargo, no se deje engañar por el bajo costo. Un producto DBMS personal es adecuado solo

para proyectos de muy pequeña escala y nunca debe implementarse para aplicaciones multiusuario.

Finalmente, el DBMS móvil es una versión especializada de un DBMS departamental o empresarial. Está diseñado para usuarios remotos que generalmente no están conectados a la red. El DBMS móvil permite el acceso y modificación de bases de datos locales en una computadora portátil o dispositivo portátil. Además, el DBMS móvil proporciona un mecanismo para sincronizar los cambios de bases de datos remotos a un servidor de bases de datos departamental o empresarial centralizado.

Un DBMS diseñado para un tipo de procesamiento puede no ser adecuado para otros usos. Por ejemplo, un DBMS personal no está diseñado para múltiples usuarios, y un DBMS empresarial es generalmente demasiado complejo para usuarios individuales. Asegúrese de comprender las diferencias entre el software DBMS empresarial, departamental, personal y móvil, y elija el DBMS adecuado para sus necesidades específicas de procesamiento de datos. Es posible que deba elegir varios tipos de DBMS, es decir, un DBMS para cada nivel, con el uso determinado por las necesidades de cada proyecto de desarrollo.

Si su organización requiere soluciones DBMS a diferentes niveles, favorece la selección de un grupo de soluciones DBMS del mismo proveedor siempre que sea posible. Si lo hace, minimizará las diferencias en el acceso, desarrollo y administración. Por ejemplo, favorecer Oracle Database Personal Edition para sus necesidades de DBMS de un solo usuario si su organización utiliza Oracle como el DBMS de empresa de su elección.

Proveedores de DBMS

Hay muchos proveedores de DBMS para elegir. Sin embargo, existen niveles definidos en términos de popularidad, soporte y liderazgo en el mercado de DBMS. En general, el mercado se puede dividir en los siguientes grupos:

- Los tres grandes: los tres líderes del mercado que constituyen la mayor parte de la base instalada de DBMS, así como la mayor parte de las nuevas ventas.
- El segundo nivel: grandes proveedores de DBMS con productos estables, pero rezagados con respecto a los Tres Grandes en términos de funcionalidad y número de usuarios
- Otros actores importantes: otros proveedores de DBMS con productos viables y aptos para empresas.
- Código abierto: productos DBMS compatibles como software de código abierto (a diferencia de un solo proveedor)
- No relacionales: proveedores que suministran productos DBMS prerrelacionales para respaldar aplicaciones antiguas
- NoSQL: productos DBMS no relacionales para Big Data que son altamente escalables para admitir aplicaciones web modernas
- Orientado a objetos: proveedores de productos ODBMS que se utilizan junto con lenguajes OO y proyectos de desarrollo

- Basado en PC: aunque muchos de los otros proveedores crean versiones para PC de sus productos DBMS, estas compañías o productos se enfocan exclusivamente en la plataforma de PC

Las siguientes secciones presentan un poco más de información sobre cada uno de los grupos del mercado.

Los Tres Grandes

Oracle Corporation, IBM Corporation y Microsoft son los principales proveedores de DBMS, en ese orden. La base instalada para Oracle es mayor que la de cualquier otro DBMS, y es probable que Oracle sea utilizado por más personas que cualquier otro DBMS. El DB2 de IBM está en el segundo lugar y está cobrando fuerza. Con el casi monopolio de DB2 en el mainframe y su creciente base instalada de UNIX y Windows, IBM puede afirmar que es utilizada por casi tantos usuarios como Oracle. En tercer lugar, en general, pero en primer lugar en máquinas Windows, está Microsoft, con Microsoft SQL Server. Microsoft SQL Server solo se ejecuta en plataformas Windows.

Información de contacto:

- Oracle Corporation
500 Oracle Parkway
Redwood Shores, CA 94065
www.oracle.com
- IBM Corporation
New Orchard Road
Armonk, NY 10504
www.ibm.com
- Microsoft
1 Microsoft Way
Redmond, WA 98052-6399
www.microsoft.com

El Segundo Nivel

La categoría de proveedores de DBMS de segundo nivel comprende compañías que fueron líderes del mercado en los años ochenta. Sybase fue un innovador líder en tecnología de base de datos, que introdujo conceptos como el paradigma cliente/servidor, procedimientos almacenados y más.

De hecho, Microsoft SQL Server está basado en la tecnología Sybase. Sybase fue adquirida por el proveedor alemán de aplicaciones empaquetadas SAP en 2010.

Teradata es otro innovador en tecnología de base de datos que enfoca su modelo de negocio principalmente en el almacenamiento y análisis de datos. Teradata se incorporó en 1979 y anteriormente era una división de NCR Corporation, pero se convirtió en una entidad independiente con una escisión de NCR en octubre de 2007.

Informix es otro ex líder que tuvo dificultades con una estrategia de desarrollo equivocada. Todos los activos de DBMS de Informix fueron adquiridos por IBM Corporation en el verano de 2001, e Informix sigue siendo respaldado por IBM y utilizado por muchas organizaciones orientadas al desempeño.

Información de Contacto:

- Sybase Inc. (una empresa de SAP)
6475 Christie Avenue
Emeryville, CA 94608
www.sybase.com
- Teradata Corporation
10000 Innovation Drive
Dayton, OH 45342
www.teradata.com

Otros Actores Importantes

Además de los proveedores mencionados anteriormente, varios proveedores adicionales ofrecen productos relacionales con un alto grado de funcionalidad y con un número razonable de usuarios instalados. El principal de ellos es Ingres, actualmente comercializado por Actian Corporation. Ingres comenzó su vida como la primera implementación de DBMS relacional en la Universidad de California, Berkeley. El código de Ingres fue mejorado en un producto comercial por Relational Technologies. Relational Technologies se convirtió en Ingres Corporation, que fue comprada por ASK Corporation, que luego fue comprada por Computer Associates (ahora CA Technologies). Computer Associates eventualmente escindió a Ingres como una compañía separada, que es una fuente abierta de Ingres. Esa compañía cambió su nombre a Actian en 2011. Ingres aún cuenta con un gran número de usuarios, especialmente en Europa.

Otro jugador importante son los productos DBMS Adabas y Tamino de Software AG. Adabas tiene sus raíces como un DBMS de lista invertida que se aumentó para convertirse en relacional. Tamino es el nuevo DBMS basado en XML de Software AG.

Información de Contacto:

- Actian Corporation
500 Arguello Street, Suite 200
Redwood City, CA 94063
www.actian.com
- Software AG
Uhlandstr. 12
64297 Darmstadt
Germany
www.softwareag.com

Ofertas de DBMS de Código Abierto

La rápida aceptación y el resplandor de los medios que rodean a Linux ha animado a la comunidad de código abierto. El término open source se refiere al software que los usuarios pueden ejecutar, copiar, distribuir, estudiar, cambiar y mejorar de forma gratuita. A menudo, "código abierto" se interpreta como software libre. Esto es comprensible, pero el concepto de código abierto de libre está más cerca de "libertad" que de "gratis". El software de código abierto se adhiere a las siguientes creencias:

- Los usuarios son libres de ejecutar el programa, para cualquier propósito.
- Los usuarios son libres de inspeccionar el código fuente real del programa para determinar cómo funciona.
- Los usuarios son libres de modificar y adaptar el software a sus necesidades específicas.
- Los usuarios son libres de distribuir copias a quien sea.
- Los usuarios son libres de publicar sus mejoras de código al público, para que toda la comunidad se beneficie.

Hay dos productos principales de DBMS de código abierto: PostgreSQL y MySQL. Además de convertirse en un producto comercial, Ingres se mejoró en Berkeley, convirtiéndose finalmente en un DBMS de segunda generación llamado Postgres y, finalmente, en una versión SQL llamada PostgreSQL. Hay varios distribuidores de PostgreSQL, o puede descargarse de forma gratuita en www.postgresql.org/.

MySQL es un verdadero servidor de bases de datos SQL multiusuario y multihilo. Está disponible como descarga gratuita en www.mysql.com, pero técnicamente no es un DBMS de código abierto. Debido a que no obtiene acceso al código fuente, no puede modificarlo. Además, debe pagar una licencia si vende MySQL, si cobra por administrar un servidor o si incluye MySQL como un componente de otro producto.

Oracle Corporation adquirió MySQL en 2009, pero continúa respaldándolo y comercializándolo de la misma manera que MySQL antes de la adquisición.

Proveedores de DBMS No Relacionales

Antes de la aceptación comercial de la tecnología relacional a principios de la década de 1980, el mercado de DBMS se basaba principalmente en mainframe y estaba dominado por dos actores: IBM y Cullinet. El IMS de IBM es un DBMS jerárquico que organiza los datos en estructuras de árbol. Cullinet comercializó IDMS, un DBMS CODASYL que organizó los datos en estructuras de red. Cullinet fue adquirida por Computer Associates a fines de los años 80.

Proveedores de DBMS NoSQL

Otra forma más reciente de DBMS no relacional es el sistema de base de datos NoSQL. En un nivel alto, NoSQL implica no relacional, distribuido, flexible y escalable. Muchos también son de código abierto. NoSQL surgió de la necesidad percibida de sistemas de bases de datos "modernos" para apoyar iniciativas web. Además, algunos atributos comunes de los DBMS de NoSQL incluyen la falta de un esquema, la simplicidad de uso, el soporte de replicación y una capacidad "eventualmente consistente" (en lugar de la capacidad de transacción típica de ACID). Realmente no significa que no hay soporte de SQL. De hecho, hoy en día el "no" se define más comúnmente como "no solo".

Hay una gran cantidad de "sistemas de base de datos" NoSQL disponibles para usar, pero los más populares incluyen

- CouchDB: una base de datos orientada a documentos que se puede consultar e indexar en forma de MapReduce utilizando JavaScript. CouchDB también ofrece replicación incremental con detección y resolución de conflictos bidireccionales.
- MongoDB: un sistema de base de datos orientado a documentos escalable, de alto rendimiento, de código abierto.
- Hadoop y HBase: el proyecto Apache Hadoop desarrolla software de código abierto para computación confiable, escalable y distribuida. HBase es el sistema de base de datos Hadoop. Admite acceso de lectura/escritura aleatorio en tiempo real a "Big Data".

El sitio web en <http://nosql-database.org/> ofrece un buen resumen de los sistemas de bases de datos NoSQL disponibles en el mercado.

Proveedores de DBMS Orientados a Objetos

En un momento dado a fines de la década de 1980 y principios de la década de 1990, parecía que los sistemas de gestión de bases de datos orientados a objetos (ODBMS) iban a ser aceptados rápidamente. Por muchas razones (difícil de consultar, sin un modelo de datos sólido, etc.), esta

tecnología no se adoptó ampliamente. De hecho, muchos de los proveedores de ODBMS han reposicionado sus productos como soluciones de e-business o componentes de base de datos para desarrolladores de aplicaciones orientadas a objetos. Sin embargo, hay pequeños nichos de usuarios ODBMS.

Los siguientes proveedores aún ofrecen productos y componentes ODBMS:

- Almacén de objetos de Progress Software en www.progress.com
- Ontos en www.ontos.com
- Poet en www.poet.com

Proveedores de DBMS basados en PC

Más fáciles de usar que los productos DBMS empresariales completos, los productos DBMS basados en PC proporcionan almacenamiento estructurado para aplicaciones y datos basados en PC. Los siguientes proveedores y productos son los líderes del mercado en el mercado de DBMS basado en PC:

- dBase en www.dbase.com
- FileMaker en www.filemaker.com
- Lotus Approach en www.ibm.com/software/lotus/products/smartsuite/approach.html (incluido en Lotus SmartSuite)
- Microsoft Access en www.microsoft.com/office/access/
- Paradox en www.corel.com (incluida en la edición Profesional de WordPerfect Office)

Tipos de almacenamiento

Almacenamiento de Conexión Directa o DAS

El almacenamiento de conexión directa (direct-attached storage o DAS) es un almacenamiento de cómputo que está conectado a una computadora y no es accesible para otras computadoras. Para un usuario de cómputo individual, un disco duro o una unidad de estado sólido (SSD) es la forma habitual de almacenamiento de conexión directa. En la empresa, las unidades de disco individuales en un servidor se denominan almacenamiento de conexión directa, al igual que los grupos de unidades que son externas al servidor pero que están conectadas directamente a través de la Interfaz de sistemas informáticos pequeños (SCSI), la conexión de tecnología avanzada en serie (SATA), SCSI conectado en serie o Serial-Attached SCSI (SAS), el canal de fibra (FC) o iSCSI.

DAS se puede implementar como discos —unidades de disco duro (HDD) o SSD— dentro de un chasis de servidor, por ejemplo, o como un gabinete de almacenamiento externo o gabinetes conectados directamente a una tarjeta conectada al bus interno de un servidor. También puede ser una unidad individual en una computadora de escritorio o portátil.

Un dispositivo de almacenamiento de conexión directa no está conectado a la red. No hay conexión a través de conmutadores Ethernet o FC que conecten dispositivos de almacenamiento conectados a la red (NAS) y redes de área de almacenamiento (SAN).

Otros tipos de almacenamiento, como los dispositivos ópticos y la cinta, son técnicamente DAS, ya que están conectados directamente a un sistema. Sin embargo, cuando uno se refiere a DAS, generalmente se refiere al almacenamiento primario o secundario interno o externo en forma de HDD y SSD.

Ventajas y desventajas del almacenamiento de conexión directa

DAS puede proporcionar a los usuarios un mejor rendimiento que el almacenamiento en red porque el servidor no tiene que atravesar una red para leer y escribir datos. Es por eso que las organizaciones empresariales a menudo recurren a DAS para ciertos tipos de aplicaciones que requieren un alto rendimiento. Microsoft, por ejemplo, recomienda que las instalaciones de Exchange utilicen DAS.

Problemas de seguridad

Al igual que con todos los sistemas de almacenamiento que alojan las aplicaciones y los datos críticos de una organización, la seguridad es una preocupación primordial. Para proteger los datos alojados en el almacenamiento de conexión directa, debe:

- Realice una evaluación de seguridad para descubrir debilidades. Si no encuentra ninguno, busque más. Puede encontrar que su entorno DAS tiene vulnerabilidades de seguridad, como permisos de usuario laxos, parches faltantes o sistemas mal configurados.
- Ejecute una auditoría de los permisos de usuario y escanee el DAS en busca de información no estructurada accesible para todas las personas en la red. Una vez hecho esto, bloquee los permisos y piense en segmentar su red para proteger mejor los sistemas DAS críticos.
- Haga de los parches una prioridad. No mantenerse al día con los parches de software de aplicación y sistema operativo puede dejar su DAS vulnerable a actos ilícitos. Un intruso que obtenga acceso no deseado al servidor conectado al DAS puede acceder a todos los datos del sistema, y es posible que ni siquiera sepa que el sistema ha sido violado.
- Realice pruebas de continuidad del negocio y tolerancia a fallos comprobando la resistencia del sistema DAS. Concéntrese en escenarios realistas de fallas del sistema, como que el hardware de almacenamiento muera o el centro de datos se dañe o se destruya por completo.

Servidores NAS

Un servidor NAS es un dispositivo de almacenamiento conectado a la red. Su función es la de hacer copias de seguridad de los archivos que tú le indiques en la configuración, tanto los de tu

ordenador personal como los de cualquier otro dispositivo móvil, aunque también tiene muchas otras funcionalidades. Lo único que necesitarás es utilizar las diferentes aplicaciones que tiene cada fabricante.

A efectos prácticos la función principal de estos dispositivos es la de actuar como unidad de almacenamiento, haciendo las veces de disco duro externo o permitiéndote crear tu propio almacenamiento en la nube. La diferencia con las nubes de otras empresas es que en este caso los discos duros donde se almacenan tus datos están en tu propia casa, y no en los servidores pertenecientes a la empresa que te presta el almacenamiento.

Un NAS es un ordenador con su propio sistema operativo y que está adaptado para estar todo el día funcionando. En ellos puedes distinguir dos conjuntos de componentes, estando por una parte lo que es el NAS en sí con su RAM, su procesador y toda su circuitería, y por otra parte los discos duros que puedes añadir a sus ranuras. Dependiendo del modelo o el fabricante, estos discos duros pueden venir incluidos cuando los compres o tendrás que comprarlos aparte.

Aunque su principal funcionalidad sea la misma, la de almacenar tus archivos, no tienes que confundir a los NAS con los discos duros externos. Un disco duro externo es simplemente eso, una unidad de almacenamiento que conectas y desconectas, mientras que el NAS funciona más como un ordenador conectado a la red al que puedes acceder desde varios dispositivos, y cuyas funcionalidades puedes ampliar con aplicaciones.

Este tipo de dispositivo puede tener dos perfiles diferentes. Por una parte están los enfocados a los usuarios domésticos, que priman la facilidad de uso por encima de todo, y luego tienes los que están enfocados a las pequeñas y medianas empresas, y que pueden tener una buena cantidad de ranuras para discos duros y posibilidades para configurarlos.

Los usuarios domésticos también tienen varias opciones al configurar su NAS. Por ejemplo, si tienes uno con dos ranuras de disco duro vas a poder elegir entre aprovechar ambos para sumar el mayor almacenamiento posible o hacer que uno copie todo el contenido del otro para tener una doble copia de seguridad con la que reforzar la seguridad frente a averías. Además, como ya hemos dicho, de ampliar las aplicaciones que puedes instalar para añadirle nuevas funcionalidades.

Red de área de almacenamiento (SAN)

Una red de área de almacenamiento (SAN) es una red de alta velocidad independiente y dedicada que interconecta y suministra depósitos compartidos de dispositivos de almacenamiento a varios servidores. Cada servidor puede acceder al almacenamiento compartido como si fuera una unidad conectada directamente al servidor. Una SAN suele montarse con cableado, adaptadores de bus de host y conmutadores SAN conectados a matrices de almacenamiento y servidores. Cada conmutador y cada sistema de almacenamiento de la SAN debe estar interconectado.

Definición de SAN

Una red de área de almacenamiento (SAN) es una red de dispositivos de almacenamiento a la que pueden acceder varios servidores u ordenadores y que proporciona un depósito de espacio de almacenamiento compartido. Cada ordenador de la red puede acceder al almacenamiento de la SAN como si fueran discos locales conectados directamente al ordenador.

SAN frente a NAS

La SAN y el almacenamiento conectado a red (NAS) son dos tipos diferentes de soluciones de almacenamiento compartido en red. Mientras que una SAN es una red local compuesta por varios dispositivos, un NAS es un único dispositivo de almacenamiento que se conecta a una red de área local (LAN).

¿Qué es un conmutador SAN?

Un conmutador SAN es hardware que conecta servidores a depósitos compartidos de dispositivos de almacenamiento. Se dedica a mover el tráfico de almacenamiento de una SAN.