

O'REILLY®

Cortesía de
aws

Una introducción a las bases de datos en la nube

Una guía para administradores

Wendy A. Neu, Vlad Vlasceanu,
Andy Oram y Sam Alapati

INFORME

Deshágase de las bases de datos antiguas

AWS ofrece la selección más amplia de bases de datos creadas para un fin específico, lo que le permite **ahorrar, crecer e innovar más rápido**



Escala empresarial a una décima parte del costo de las bases de datos comerciales



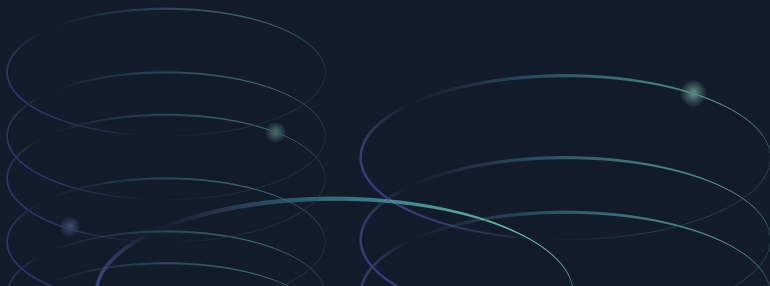
Rendimiento de tres a cinco veces mayor en comparación con otras alternativas populares



Más de 14 motores de bases de datos: más que cualquier otro proveedor

aws databases

Más información: aws.amazon.com/databases



Una introducción a las bases de datos en la nube

Una guía para administradores

*Wendy A. Neu, Vlad Vlasceanu, Andy Oram
y Sam Alapati*

Beijing • Boston • Farnham • Sebastopol • Tokio

O'REILLY®

Una introducción a las bases de datos en la nube

por Wendy A. Neu, Vlad Vlasceanu, Andy Oram y Sam Alapati

Copyright © 2019 O'Reilly Media Inc. Todos los derechos reservados.

Impreso en los Estados Unidos de América

Publicado por O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

Los libros de O'Reilly pueden adquirirse para uso educativo, empresarial o promocional. Las ediciones en línea también están disponibles para la mayoría de los títulos (<http://oreilly.com>). Para obtener más información, comuníquese con nuestro departamento de ventas corporativas/institucionales: 800-998-9938 o a corporate@oreilly.com.

Editor de desarrollo: Jeff Bleiel

Diseño editorial: David Futato

Editor de adquisiciones: Jonathan Hassell

Diseño de la portada: Karen Montgomery

Editora de producción: Katherine Tozer

Ilustradora: Rebecca Demarest

Corrector de textos: Octal Publishing, LLC

Septiembre de 2019: Primera edición

Historial de revisiones de la primera edición

19-08-2019: Primera versión

El logotipo de O'Reilly es una marca registrada de O'Reilly Media, Inc. *Una introducción a las bases de datos en la nube*, la imagen de la portada y las imágenes comerciales relacionadas son marcas registradas de O'Reilly Media, Inc.

Las opiniones expresadas en este libro son propias de los autores y no representan las opiniones de la editorial. Si bien la editorial y los autores han obrado de buena fe para garantizar que la información y las instrucciones incluidas en este libro sean precisas, la editorial y los autores no se harán responsables por la existencia de errores u omisiones, incluida, a título enunciativo, la responsabilidad por los daños y perjuicios que surjan del uso de este libro o la confianza puesta en él. El uso de la información y las instrucciones incluidas en este libro es responsabilidad del lector. Si algunas muestras de códigos u otras tecnologías incluidas o descritas en este libro están sujetas a licencias de código abierto o a los derechos de propiedad intelectual de terceros, es responsabilidad del lector garantizar que su uso cumpla con tales licencias o derechos.

Este libro forma parte de la colaboración entre O'Reilly y AWS. Consulte nuestra [declaración de independencia editorial](#).

978-1-492-04482-6

[LSI]

Tabla de contenidos

1. Opciones de bases de datos en la nube.....	1
Efectos de alto nivel producto de la migración a la nube	2
Bases de datos autoadministradas versus bases de datos administradas	5
Bases de datos nativas de la nube	6
Tipos de bases de datos administradas	7
Cuando una base de datos administrada podría no ser la mejor opción	8
El rol del DBA en una base de datos administrada	9
2. El rol cambiante del DBA en la nube.....	11
Cómo cambian las tareas del DBA en la nube	13
Seguridad para los datos y las aplicaciones en la nube	19
Infraestructura como código: sacarle provecho a la nube	23
3. Migración de sus bases de datos a la nube.....	25
Planificación	26
Migración de datos	32
Optimización	35
Conclusión	38
4. Conclusión.....	38

Opciones de bases de datos en la nube

La adopción rápida de la nube suele medirse en términos de negocios puros, y su popularidad en aumento es ampliamente reconocida. Por ejemplo, un informe reciente de Gartner (que requería la entrega de información empresarial para su revisión) descubrió que las bases de datos crecen a un ritmo del 68 % en la nube, mientras que el crecimiento local es casi inexistente, sin tener en cuenta el aumento de los costos y lo que se conoce como “actualizaciones obligatorias”.

Pero este pronunciado cambio en el acceso a la informática y las bases de datos también aparece en la clase de servicios ofrecidos y la evolución de los empleos profesionales relacionados con la informática. Muchos tipos de bases de datos administradas ya forman parte de la gama de ofertas de los proveedores de nube principales. El uso de estas bases de datos eliminará muchas tareas realizadas por un administrador de bases de datos (DBA) en entornos tradicionales en los que una organización es propietaria de su propio hardware, que denominaremos entornos *locales*. La migración a la nube agregará nuevas tareas, cambiará algunas existentes y proporcionará un contexto con diferencias sutiles para comprender muchas de las tareas.

Los proveedores de nube se empeñan en mantenerse al día con los cambios en el espacio de las bases de datos y varían sus ofertas para cumplir con las diversas necesidades de sus clientes. Las bases de datos relacionales (en forma de almacenes transaccionales y de datos) aparecen junto a aquellas no relacionales, como los almacenes de clave-valor y de documentos.

Este informe ayuda al DBA y al personal relacionado (como los científicos de datos, arquitectos de datos y desarrolladores de aplicaciones) a elegir entre las ofertas de nube. Explica, de manera general, cuáles son las responsabilidades que el DBA tendrá en el entorno de la nube. Por último, ofrece guías para realizar una migración.

En el informe no se abordan los argumentos a favor y en contra de la migración a la nube porque existen otros recursos que apoyan esta decisión, que está vinculada estrechamente a las características particulares de las bases de datos y a cómo se utilizan. Además, no analizamos ni recomendamos ofertas de nube particulares, si bien hacemos referencia a las ofertas de los proveedores de nube más importantes en la actualidad: Amazon Web Services (AWS), Microsoft Azure y Google Cloud Platform.

En este capítulo se incluye lo siguiente:

- Una breve síntesis de las diferencias entre la informática local y la informática en la nube, con un enfoque en los aspectos relevantes para los DBA
- La definición de una base de datos administrada
- Los tipos de bases de datos administradas que ofrecen los proveedores de nube
- Una descripción general del rol de un DBA en la nube

Las ofertas de nube suelen reflejar el software disponible para las implementaciones locales tradicionales. Esto significa que puede migrar los datos desde sus propios equipos hacia una base de datos similar ofrecida por un proveedor de nube. Por ejemplo, puede ejecutar MySQL u Oracle en la nube en lugar de hacerlo en su centro de datos, o puede utilizar una base de datos nativa de la nube, como Amazon Aurora.

También debe ser capaz de controlar ambos territorios (local y la nube) para migrar los datos a la nube a un ritmo cómodo para su organización. Además, puede mantener, para siempre, dos implementaciones, una local y otra en la nube. La nube suele ser un buen lugar para intentar algo totalmente nuevo, como un proyecto de big data que no haya creado anteriormente.

Si bien suponemos que la mayoría de los lectores estarán familiarizados con la nube, con sus beneficios y las varias razones a favor o en contra de su uso, en la siguiente sección se resumen las características de la nube que sustentan las ideas de este informe. Después de eso, podremos concentrarnos en las bases de datos.

Efectos de alto nivel producto de la migración a la nube

A medida que migre sus datos a la nube, encontrará diferencias con las implementaciones locales que afectarán sus decisiones. A continuación, presentamos los cambios clave de interés para un DBA:

- Pague por el uso, con aumentos por hora o mensuales, para ejecutar cada servidor (llamado *instancia* en la jerga de la informática virtual). De esta manera, si posee una aplicación de inteligencia empresarial (BI) que ejecuta todas las noches, puede poner en marcha una instancia de una base de datos por ocho horas durante la noche para evitar pagar por las 16 horas del día restantes. Con algunos servicios (como **AWS Reserved Instances**), también puede ejecutar una base de datos con un descuento durante un período limitado.
- El hardware y la infraestructura relacionada están a cargo del proveedor. Esto puede restarles una enorme cantidad de trabajo a los administradores de sistemas y de bases de datos. En lugar de aprovisionarse de su propio hardware, elige la capacidad de la CPU, la memoria y el almacenamiento que desea incorporar en su instancia de base de datos. Los proveedores suelen tener ofertas especiales que pueden mejorar el rendimiento de manera similar a las estrategias utilizadas localmente: por ejemplo, Microsoft Azure SQL Data Warehouse ofrece almacenamiento en disco de estado sólido (SSD) y Amazon DynamoDB posee una versión de almacenamiento en caché en memoria especial.
- Los proveedores de nube más importantes (AWS, Azure y Google Cloud) no solo ofrecen el aislamiento de las cargas de trabajo en máquinas virtuales (VM), sino también el aislamiento de redes corporativas completas en la nube, un servicio llamado **nube privada virtual (VPC)**. Así como las VM permiten que sus servidores utilicen recursos informáticos con eficiencia y seguridad, una VPC le permite configurar su red corporativa en la nube, lo que podría ayudarlo a ahorrar dinero mientras permite que el proveedor de nube administre una gran parte de la seguridad de la red. Los servidores de bases de datos pueden participar en las VPC. También puede configurar una red privada virtual (VPN) para comunicar la VPC de la nube con sus equipos locales.
- Los proveedores de nube ofrecen varias ubicaciones, lo cual es importante por muchos motivos: para ofrecer servidores que se encuentren geográficamente cerca de las bases de usuario de los clientes, proporcionar redundancia en caso de que ocurran desastres, o permitir balances de carga, entre otras. Algunos países, particularmente aquellos de la Unión Europea, requieren que determinados tipos de datos se almacenen dentro de sus jurisdicciones legales por razones de privacidad u otros temas legales. El concepto de *región*, que puede hacer referencia a gran parte de un continente, y de las *zonas de disponibilidad* (AZ) dentro de dichas regiones figuran en **AWS, Azure y Google Cloud**. Por lo tanto, puede distribuir sus cargas de trabajo en varias AZ dentro

de una región, lo que le permitirá replicar datos y servicios, y configurar el balance de carga a través de copias redundantes de datos para proporcionar resiliencia. Algunos proveedores de nube ofrecen regiones con características especiales para cumplir con los requisitos normativos, especialmente los relacionados con la seguridad. Todos los proveedores de nube importantes también ofrecen servicios de redes de distribución de contenido (CDN).

- Los proveedores de nube ofrecen maneras sofisticadas de escalar servicios, proporcionar tolerancia a las fallas y realizar balances de carga. Si la instancia de base de datos falla y la configuración está establecida para una disponibilidad alta, podría habilitarse una nueva instancia. Quizá, el cliente deba reiniciar las sesiones que se ejecutaban en la instancia anterior cuando ocurrió la falla, pero sus esfuerzos administrativos deben enfocarse en configurar la tolerancia a las fallas en lugar de levantarse a las 2:00 a. m. para reiniciar la base de datos. De manera similar, si la carga en una base de datos aumenta mucho, puede iniciar una nueva instancia para compartir la carga y desactivarla cuando ya no sea necesaria. Esto se llama *escalado elástico*.
- Los proveedores suelen ofrecer opciones de líneas de comandos, interfaces gráficas y API, y cada una posee fortalezas y debilidades que los profesionales informáticos conocen bien. Las API permiten la automatización y, por lo tanto, le quitan un problema de las manos.
- Algunos temas relacionados con la seguridad que debe manejar localmente los controla el proveedor, quien obviamente controla el acceso físico a los servidores. Según cómo utilice sus servicios, también puede protegerlo contra ataques de Internet de bajo nivel, como el escaneo de puertos y los ataques de denegación de servicio (DoS). En capítulos posteriores, abordaremos los aspectos de la seguridad que continúan siendo responsabilidad del DBA.
- Los proveedores de nube incorporan herramientas de monitoreo y rendimiento que puede conectar a sus bases de datos con facilidad y siempre generan nuevas características. Las herramientas de monitoreo de terceros también podrían agregar utilidades y ampliar las capacidades de monitoreo incorporadas de la plataforma.

Muchos de los criterios de selección de un proveedor son generales y no son específicos de las bases de datos: los precios, el ecosistema de los desarrolladores externos relacionado con cada oferta, la estabilidad, el soporte en cuanto a los regímenes de cumplimiento legal, el valor agregado en herramientas de alta gama para inteligencia artificial, entre otros.

Mientras evalúa las ofertas de nube, piense en el futuro de su organización. Por ejemplo: ¿el procesamiento de datos en tiempo real y de streaming serán importantes para usted en el futuro? ¿Qué ofertas de nube le brindarán soporte dentro de dos años?

Bases de datos autoadministradas versus bases de datos administradas

Ahora, nos enfocaremos específicamente en las bases de datos. Las opciones en la nube se dividen en dos categorías principales:

Bases de datos autoadministradas

En esta oferta de nube, el proveedor solo proporciona el hardware, el hipervisor para ejecutar su VM y las API u otras herramientas para administrar las implementaciones. Debe crear una VM que ejecute un sistema operativo y su aplicación. Por supuesto, puede ejecutar lo que desee en la VM, incluido un motor de bases de datos de su preferencia. Debe realizar la mayor parte de las tareas administrativas por su cuenta, como la instalación de actualizaciones y la configuración de todas las opciones de red.

Bases de datos administradas

En esta oferta de nube, el proveedor proporciona el hardware y el software del servidor. La mayoría de los proveedores ofrecen bases de datos tradicionales (como Oracle y MySQL) y bases de datos nativas de la nube, que son específicas de cada proveedor.

Las bases de datos autoadministradas se parecen bastante a las bases de datos locales y las bases de datos administradas cambian mucho más su forma de trabajar. Estas son las diferencias clave:

- Si utiliza una base de datos administrada, no tendrá que descargar, instalar, actualizar, configurar ni realizar copias de seguridad de la base de datos manualmente, ya que el proveedor de nube se encarga de hacerlo. Aún podrá modificar los parámetros que le indican a la base de datos cómo ejecutarse (intercalaciones, cursores, conexiones, etc.) y controlar decisiones, como la instalación de nuevos lanzamientos y durante cuánto tiempo almacenar las copias de seguridad.
- En ocasiones, el proveedor de nube puede proporcionarle una licencia de una base de datos de propiedad exclusiva e incluir su costo dentro del costo de uso del servicio de la nube. En otros casos, la adquisición de una licencia sigue siendo su responsabilidad. Si ya adquirió una licencia de un proveedor de bases de datos de propiedad exclusiva, a veces puede aplicarla a una base de datos administrada.
- Como el proveedor instala y ejecuta la base de datos administrada, estará limitado a las decisiones que tome el proveedor. No puede exigir un motor de base de datos que el proveedor no admita ni una versión que ya no sea compatible. Pero los tres proveedores principales (AWS, Azure y Google) ofrecen varias bases de datos

de propiedad exclusiva y de código abierto, además de sus propias bases de datos nativas de la nube.

- Las bases de datos autoadministradas y administradas ofrecen diferentes parámetros de configuración.
- Con las bases de datos administradas, gran parte de la seguridad la controla el proveedor. Pero incluso con las bases de datos administradas hay decisiones importantes de seguridad que debe tomar usted, como quiénes tendrán cuentas y qué límites establecer en las tablas o columnas. El monitoreo y las auditorías de los intentos de acceso suelen seguir siendo su trabajo. También puede controlar las configuraciones de seguridad, como qué direcciones IP tienen acceso.

Por lo tanto, las bases de datos administradas ofrecen muchas ventajas sobre las implementaciones locales y las bases de datos autoadministradas, pero podría tener razones específicas para ejecutar su propia base de datos. En el resto del informe se abordan las bases de datos administradas porque ofrecen una oportunidad única para que el DBA se enfoque en las actividades que generarán un impacto a largo plazo en la empresa.

Bases de datos nativas de la nube

Las bases de datos administradas también se pueden dividir en dos categorías: tradicionales y nativas de la nube. Los proveedores de nube suelen ofrecer bases de datos tradicionales, como Oracle, SQL Server, MySQL y PostgreSQL como bases de datos administradas. Si ha desarrollado su organización en torno a una de estas bases de datos locales, el traslado a la misma base en la nube simplifica la migración. Es menos probable que deba alterar sus aplicaciones y puede utilizar herramientas con las que está familiarizado para administrar las bases de datos. Además, puede combinar ofertas de diferentes proveedores o mantener la versión local de la base de datos.

Pero los proveedores han invertido mucho esfuerzo en sus propias ofertas novedosas, en ocasiones llamadas bases de datos *nativas de la nube*. Y brindan evidencias de que las bases de datos nativas de la nube tienen un mejor rendimiento, escalan con más facilidad y son más rentables a largo plazo. Por esta razón, los testimonios de **Autodesk** e **InfoScout** sugieren que los ingenieros de AWS han resuelto muchos de los problemas de escalabilidad y eficiencia propios de la administración de bases de datos relacionales en la nube con su propia base de datos: Amazon Aurora. Las bases de datos nativas de la nube también están diseñadas para realizar escalas de gran envergadura, una tarea que ha sido históricamente muy difícil para las bases de datos relacionales.

Tipos de bases de datos administradas

La mayoría de los tipos de bases de datos que se encuentran en el campo también se ofrecen como bases de datos administradas. Además, los proveedores de nube también desarrollaron sus propias bases de datos nativas de la nube, para adaptarse a las tendencias populares del sector y ofrecer beneficios de rendimiento. Los tipos principales de bases de datos admitidas incluyen los siguientes:

Bases de datos relacionales

Como ya mencionamos, algunas bases de datos son versiones administradas de bases de datos populares de uso extendido. Por ejemplo, Azure ofrece el tradicional SQL Server de Microsoft. Amazon admite MariaDB, MySQL, Oracle, PostgreSQL y SQL Server mediante [Amazon Relational Database Service \(RDS\)](#). Estas ofertas lo ayudan a migrar las bases de datos con más facilidad desde las instalaciones locales.

Asimismo, los proveedores han creado bases de datos propias. Por ejemplo, Azure ofrece [Azure Cosmos DB](#), Google ofrece [Cloud Spanner](#) y AWS ofrece [Amazon Aurora](#).

Almacenes de datos

Aunque estas son bases de datos relacionales típicas, difieren de aquellas orientadas a las transacciones internamente y en cuanto a sus ofertas. Por ejemplo, las bases de datos orientadas a las transacciones suelen almacenar todas las columnas de una fila en conjunto, para que pueda recuperar varias columnas sobre un cliente o producto rápidamente. Por otro lado, los almacenes de datos en la nube tienden a ser columnares, lo que significa que almacenan datos por columna en lugar de hacerlo por fila. Esto aumenta significativamente la velocidad de las consultas comunes de los almacenes, como: "Entrégame las edades de todos los clientes que viven en California". Las herramientas que ofrecen los proveedores con los almacenes de datos se concentran en la entrada y extracción rápidas, lo que facilita su uso en aplicaciones de big data. [Amazon Redshift](#), [Google BigQuery](#) y [Azure SQL Data Warehouse](#) son ejemplos de estas ofertas.

Bases de datos no relacionales

Este término se suele utilizar para una diversidad de almacenes de datos que, a diferencia de las bases de datos relacionales, se construyen para casos de uso de aplicaciones especiales. Los proveedores de nube ofrecen varias de estas para diferentes propósitos:

- Las bases de datos de *clave-valor* ofrecen rapidez en el almacenamiento y la recuperación de valores sin soporte para operaciones más sofisticadas.

- Las bases de datos de *documentos* almacenan datos como documentos JSON con un esquema flexible, lo que permite que los datos se almacenen y consulten en el mismo formato utilizado en las aplicaciones.
- Las bases de datos de *gráficos* almacenan relaciones entre objetos, lo que facilita la ejecución de algoritmos, como la búsqueda del objeto con la mayor cantidad de conexiones.
- Las bases de datos de *búsquedas* optimizan la ubicación de documentos que contienen palabras específicas.
- Las bases de datos de *series temporales* registran eventos con marcas de fecha y hora, y están optimizadas para tareas relacionadas con el tiempo, como la creación de gráficos de eventos que ocurren en el tiempo, que son útiles para hacer un seguimiento de eventos, como lecturas o publicaciones web, desde dispositivos.
- Las bases de datos de *libros mayores* son como archivos de registro mejorados y seguros que registran actividades, como transacciones financieras, las firman y las hacen inmutables.

Otras opciones, incluidas las bases de datos de almacenamiento en caché en memoria, también podrían estar disponibles. Los proveedores de nube ofrecen servicios de almacenamiento menos estructurados (como BLOB o archivos). Algunos ejemplos son [Amazon Simple Storage Service \(Amazon S3\)](#), [Google Cloud Storage](#) y [Microsoft Azure Blob Storage](#). Los servicios ofrecen opciones aún más rentables para los almacenes de archivos, junto con herramientas para trasladar datos desde y hacia los almacenes con facilidad.

Inevitablemente, en este capítulo se han establecido las características de las bases de datos en términos generales, con muchas palabras mitigantes como “algunas” o “podría”. Esta reticencia a un mayor compromiso parte de las diferencias sutiles que se pueden notar después de analizar las ofertas de los proveedores en detalle. Algunas ofertas han evolucionado más que otras. Algunas bases de datos ofrecen una administración o migración más sencilla que otras. Debe tener en cuenta todos estos factores para elegir una oferta de nube.

Cuando una base de datos administrada podría no ser la mejor opción

El servicio de bases de datos administradas no resulta adecuado en dos escenarios.

El tamaño de la base de datos y las entradas y salidas por segundo (IOPS) son más grandes que los límites de la base de datos

Las bases de datos nativas de la nube están diseñadas para escalar y cumplir con demandas extremas. Sin embargo, las bases de datos administradas basadas en motores de bases de datos más tradicionales suelen presentar límites más amplios en cuanto a su tamaño y las operaciones de entradas/salidas por segundo que pueden admitir. Muchas organizaciones se ajustan a estos límites. Pero si la base de datos es realmente enorme y posee necesidades de IOPS más grandes que las que puede admitir el proveedor, podría descubrir que el servicio de bases de datos administradas elegido es inviable. Debería ejecutar una base de datos nativa de la nube o sus propias instancias en la nube mediante el alquiler de las VM y el almacenamiento necesarios.

Debe tener el control total de sus bases de datos

Si es importante que tenga el máximo control de las bases de datos, debe administrar sus propias bases de datos en la nube. Si necesita una versión específica de la base de datos que el proveedor de nube no admite, o si necesita utilizar características u opciones que el proveedor no admite, tendrá que arreglárselas por su cuenta.

El rol del DBA en una base de datos administrada

Como mencionamos al principio del informe, la migración a la nube cambiará su trabajo, a veces sutilmente. En el [Capítulo 2](#) se aborda esto en detalle, pero por ahora solo diremos que, cuando haya dejado el mantenimiento del hardware, la configuración del sistema y otras funciones a cargo de las bases de datos administradas, aún tendrá mucho trabajo que hacer. Estas son algunas de las tareas que suelen realizar los DBA en una base de datos administrada:

- Determinar los requisitos, como la capacidad y memoria de la CPU.
- Configurar las ejecuciones de la base de datos, incluidas la automatización del escalado, la recuperación de fallas y la toma de instantáneas de la base de datos.
- Elegir las regiones y AZ donde alojar las instancias de la base de datos.
- Crear y manipular las bases de datos y las tablas mediante la línea de comandos, consola o API que ofrece el proveedor.
- Autorizar las cuentas de usuarios, configurar grupos y controlar el acceso a las bases de datos y sus partes. Algunas bases de datos ofrecen este acceso mediante la administración de identidades

y accesos (IAM) y otras utilizan las interfaces tradicionales que ofrece la base de datos en las implementaciones locales.

- Determinar la división de los datos en particiones cuando deba distribuirlos en varios sistemas.
- Configurar los parámetros de la base de datos que afectan el rendimiento o la resiliencia
- Realizar auditorías de los accesos y monitorear la actividad de la base de datos por razones de seguridad, rendimiento y resiliencia.
- Realizar entradas y transferencias de datos mediante el proceso de extracción, transformación y carga (ETL) tradicional o mediante herramientas de transmisión de datos más recientes. Estas transferencias de datos pueden incluir la aceptación de datos de orígenes externos, la incorporación de datos en herramientas de big data como Spark o Kafka, el almacenamiento de datos en un almacén de datos en la nube o local y cualquier otra relación que deba establecer con otras herramientas para procesar los datos.
- Determinar los esquemas o formatos, junto con los arquitectos de datos y los desarrolladores de aplicaciones, e implementarlos en almacenes de datos relacionales o no relacionales.

Lo más emocionante es que tendrá más tiempo para analizar el contexto empresarial de la base de datos y pensar cómo mejorar el valor de los datos que administra para la organización. Entraremos en detalle acerca de estos temas en el [Capítulo 2](#).

El rol cambiante del DBA en la nube

Tradicionalmente, el administrador de la base de datos (DBA) debía controlar casi toda la gama de operaciones correspondientes al almacenamiento y el uso de los datos de una organización. La lista de tareas del DBA incluía las siguientes funciones:

- Instalación de estanterías y cableado
- Instalación, aplicación de parches y actualización del software de la base de datos
- Creación y configuración de las instancias de la base de datos
- Administración de usuarios, roles y permisos
- Realización de copias de seguridad y recuperaciones
- Aseguramiento de los datos, incluido el cifrado de datos críticos
- Migración, entrada y exportación de datos
- Realización de ajustes de rendimiento
- Monitoreo y solución de problemas de las bases de datos
- Manejo de la alta disponibilidad
- Asistencia a desarrolladores y analistas de bases de datos respecto de las tareas relacionadas con la base de datos

Como puede notar a partir de la lista, el trabajo del DBA suele estar relacionado con las operaciones de las bases de datos, la administración de desastres, la administración de usuarios y los ajustes de rendimiento. El monitoreo y la resolución de problemas suelen ser manuales y reactivos en lugar de planificados.

No queda mucho tiempo para realizar tareas de un nivel superior en la cadena de valor, como el diseño de las bases de datos y la optimización de las aplicaciones que dependen de ellas.

La falta de automatización y un enfoque reactivo a los problemas han hecho que los DBA tengan dificultades para seguir el ritmo de todas las bases de datos bajo su control. La mayoría de sus días laborales están dedicados a lidiar con tareas de extrema urgencia.

A medida que la informática en la nube cambia profundamente los roles laborales y la manera en que las organizaciones dirigen sus negocios, se está reestructurando el trabajo tradicional del DBA. En la nube, muchas de las tareas tradicionales de un DBA simplemente desaparecen o se reducen de manera significativa. El proveedor de nube, por supuesto, se encarga de todo el trabajo de la infraestructura, como el montaje y apilado de los servidores, las redes y el almacenamiento. Las copias de seguridad y la seguridad también se externalizan, en su mayoría, al dominio del proveedor de nube. Entonces, ¿qué hacen los DBA en la nube?

Si ejecuta sus propias bases de datos en la nube, seguirá realizando muchas tareas tradicionales del DBA, como la instalación y la aplicación de parches de software, la creación de copias de seguridad de las bases de datos, etcétera. En este informe, el enfoque está puesto en las bases de datos administradas, como las definimos en **“Bases de datos autoadministradas versus bases de datos administradas” en la página 5**, porque ofrecen bastantes beneficios adicionales en comparación con las bases de datos autoadministradas, además de realizar cambios más importantes en el rol del DBA. En este capítulo, la atención también se centra en las bases de datos relacionales, que suelen requerir una mayor variedad de tareas administrativas.

Cuando utiliza una base de datos administrada en la nube, puede invertir más tiempo en la arquitectura de datos y las aplicaciones que admite la base de datos. La anticipación de necesidades y las mejoras de la base de datos se convierten en las responsabilidades principales del DBA. Estará en una situación mucho más favorable para derivar más valor de los activos de datos de la organización al ayudar a los equipos a entregar nuevas características con más rapidez y realizar ajustes en el rendimiento de las aplicaciones de manera proactiva.

Otras tareas incluyen ayudar a establecer canalizaciones de datos en lote y en tiempo real para entrar y transformar los datos de lote y de streaming. La creación, el mantenimiento y los ajustes de las canalizaciones de datos de alta distribución son funciones importantes de los administradores que trabajan con bases de datos basadas en la nube. También puede dedicar más tiempo a mejorar la seguridad de la base de datos y garantizar su adherencia y cumplimiento con los requisitos impuestos por gobiernos y organismos industriales.

Cuando utiliza un servicio de bases de datos administradas en la nube, el rol tradicional del DBA se invierte. En la **imagen 2-1** se resume la diferencia entre la forma en que los DBA invierten su tiempo localmente y con un servicio de bases de datos administradas.

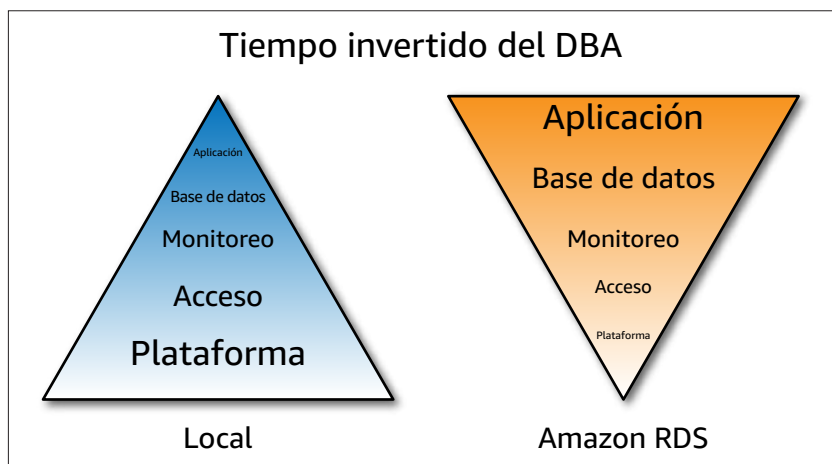


Imagen 2-1. El rol invertido del DBA en la nube

En síntesis, la migración de las bases de datos a la nube libera el tiempo que invertía en la administración y el desarrollo, y que ahora puede dedicar a otras tareas más importantes. Dado que el mantenimiento de rutina, las copias de seguridad y tareas similares están a cargo del proveedor de nube, puede enfocarse más en los objetivos empresariales, como la optimización de procesos comerciales clave.

El factor más importante que determinará su éxito en el nuevo entorno de la nube es su familiarización con todos los servicios relevantes del proveedor de nube. Estos incluyen las diferentes opciones de almacenamiento, las herramientas de monitoreo (que puede utilizar si paga una tarifa adicional), las características de seguridad de la nube y otros servicios incorporados que puede aprovechar para ejecutar las bases de datos en la nube de manera rentable, segura y con un alto rendimiento.

Cómo cambian las tareas del DBA en la nube

Cuando se realiza una migración de las bases de datos a la nube, muchas tareas operativas tradicionales desaparecen, otras permanecen modificadas y surgen tareas nuevas para el DBA. Sin embargo, casi todo lo que haya hecho localmente cambiará de alguna forma, ya que tendrá a su disposición nuevas herramientas y opciones.

Las tareas del DBA que pasan a manos del proveedor

Una diferencia clave en el uso de un servicio de bases de datos administradas en la nube, en comparación con una base de datos local, es que los DBA realizan tareas administrativas sin acceder físicamente a los servidores ni controlar directamente las instalaciones. Al principio, el DBA podría tardar un poco en acostumbrarse a esta nueva realidad. Es importante que los DBA aprovechen la oportunidad para meterse de lleno en las nuevas tareas desarrolladas en este capítulo.

Tareas del DBA que siguen existiendo, pero con cambios

Muchas de las responsabilidades convencionales del DBA aún son necesarias en la nube, pero con diferencias sutiles, sin importar si utiliza un servicio de bases de datos administradas o si ejecuta sus propias bases de datos.

Aprovisionamiento

Los servicios de bases de datos administradas, como Amazon RDS, ofrecen un amplio rango de tipos de instancias para sus servidores de bases de datos. En lugar de ordenar un sistema de hardware con requisitos particulares de memoria y CPU para su centro de datos, elige uno o más tipos de instancias que ofrecen la capacidad de memoria y CPU que cumplen con sus requisitos. La combinación del tipo de instancia con sus cargas de trabajo puede ayudar a reducir la posibilidad de pagar por recursos que no utilizará.

La escalabilidad es otra tarea de aprovisionamiento. El escalado automático, que puede realizar al nivel de las VM y de las bases de datos, permite que el sistema agregue una instancia cuando las cargas se vuelven muy grandes y que desactive una instancia para no desperdiciar el exceso de capacidad cuando las cargas disminuyen. Aunque el escalado automático es una forma de automatización de gran valor, puede ser víctima de un ataque de DoS o de un error de configuración que active una gran cantidad de instancias e infle los costos.

También puede elegir regiones cercanas a los usuarios de las bases de datos para obtener tiempos de respuesta óptimos. El uso inteligente de las ofertas de un proveedor de nube global es un aspecto clave del aprovisionamiento en la nube.

Como ya dijimos, las bases de datos tradicionales, en especial las de procesamiento de transacciones en línea (OLTP), son particularmente difíciles de escalar. A medida que la cantidad de usuarios, el tamaño de los conjuntos de datos y la complejidad de las operaciones de las bases de datos aumentan, los tiempos de respuesta de las consultas

se alargan. El crecimiento empresarial se puede reflejar en todos esos factores, además del crecimiento de las aplicaciones que representan una carga más grande en la base de datos. Un desarrollador de SQL y un DBA expertos pueden solucionar estos problemas hasta cierto punto mediante el establecimiento de parámetros de configuración óptimos de la base de datos y el perfeccionamiento del código SQL.

Sin embargo, los perfeccionamientos de códigos y las configuraciones óptimas de las bases de datos solo son útiles hasta cierto punto. Por lo tanto, la escalabilidad de la base de datos sigue siendo un problema complicado en muchos casos, ya que no puede cambiar la cantidad de servidores, RAM o CPU con facilidad y en el momento para que se correspondan con los requisitos cambiantes de las cargas de trabajo. Todos estos recursos suponen el gasto del presupuesto de capital, que debe ser aprobado con mucha anticipación. Agregar un solo servidor físico demora semanas en muchos de los centros de datos tradicionales.

ETL, entrada y exportación de datos

Una de las tareas más importantes de los DBA es trasladar los datos de alguna forma. Generalmente, los transforman para extraer campos importantes o les dan otro formato. Aunque durante alrededor de 40 años este campo se ha caracterizado por las herramientas ETL, formas recientes de procesamiento de flujos, como **Hadoop** y **Spark** han creado muchas herramientas nuevas de entrada de datos.

Los agentes de mensajes como **Apache Kafka** y **Amazon Kinesis** dirigen los flujos de datos hacia varios consumidores en un modelo de *publicación-suscripción*. La mayoría de los sitios utilizan diferentes generaciones de estas herramientas para diferentes propósitos, y su rol es dominarlas y enseñarles a coexistir pacíficamente.

Los datos de streaming imponen nuevas cargas en las bases de datos tradicionales que obligan a los DBA a ajustar los parámetros de una manera diferente. El streaming suele introducir más procesos que realizan muchas actualizaciones frecuentes y pequeñas en lugar de algunas más grandes. Los controles de concurrencia también se saturan.

Las tareas generales son las mismas, ya sea que sus datos vayan a una base de datos local o a un almacén en la nube. Pero los proveedores de nube ofrecen sus propias versiones de datos de streaming, además de herramientas de entrada y exportación de datos. Por ejemplo, si utiliza el almacén de datos basado en la nube de Microsoft, **SQL Data Warehouse**, puede utilizar tanto el software existente en su conjunto de herramientas o las herramientas ofrecidas en **Microsoft Azure Data Factory (ADF)**. Amazon ofrece **numerosas herramientas para la entrada** de datos, como **Kinesis Data Firehose** y **Snowball**. Google **cuenta con un conjunto de herramientas para entradas** que combina opciones de muchos orígenes.

Copias de seguridad e instantáneas

Las bases de datos en la nube ofrecen dos métodos básicos para realizar copias de seguridad y restaurar las bases de datos. Las copias de seguridad automatizadas se suelen activar de manera predeterminada y se pueden usar para recuperar bases de datos en un punto del tiempo.

Por otro lado, las instantáneas son copias de seguridad que inicia usted. En los entornos locales, los DBA toman instantáneas de forma rutinaria (mediante la copia de los datos de la base de datos mientras están en un estado consistente) para realizar copias de seguridad o replicar el almacén de datos. En la nube, aún puede utilizar una instantánea para realizar copias de seguridad a largo plazo (como para un almacenamiento en cinta), hacer replicaciones en otras regiones de la nube y realizar retenciones por motivos de cumplimiento con las normas.

Los servicios de bases de datos administradas en la nube, como Amazon RDS, facilitan la toma de instantáneas. Puede tomar instantáneas del volumen de almacenamiento de una instancia de la base de datos (que puede incluir varias bases de datos) y crear una instancia mediante la restauración de la instantánea de la base de datos.

Acceso

En la nube, usted es responsable de todas las operaciones exclusivas de sus datos y usuarios: la asignación de cuentas, la conformación de grupos o roles para combinar cuentas de usuarios y la designación del acceso a varias partes de las bases de datos. Hablaremos de esto con más detalle en ["Seguridad para los datos y las aplicaciones en la nube"](#) en la [página 19](#).

Mantenimiento de los entornos de desarrollo y pruebas

Una tarea común del DBA es establecer entornos de desarrollo y pruebas. A menudo, esto demanda mucho trabajo por parte del DBA. Si toma instantáneas de la base de datos regularmente, puede crear un entorno de desarrollo o pruebas que refleje el estado de la base de datos en un punto determinado del tiempo. Puede encontrar ejemplos de tales servicios de [Amazon](#), [Google Cloud](#) y [Azure](#).

Registros

Los DBA necesitan registros de las bases de datos para resolver problemas y por razones de cumplimiento y auditorías. Esto sigue siendo necesario en la nube. En la [imagen 2-2](#), se muestra cómo puede visualizar los registros de una instancia de la base de datos con facilidad mediante la consola de AWS. Cuando cuenta con muchas bases de datos que debe monitorear y administrar, la utilidad de esta característica se vuelve más clara.

DB Instances > sqttest

Details **Recent Events & Logs**

Filter: Search ... X Viewing 22 of 22 Logs

Name	Last Written	Size			
log/SQAGENT2	July 19, 2016 at 3:34:20 PM UTC-4	2.3 kB	view	watch	download
log/SQAGENT.OUT	September 21, 2017 at 2:44:18 PM UTC-4	2.3 kB	view	watch	download
log/SmartAdminEvents_Backup_0_131134304597650000.xel	July 19, 2016 at 9:09:20 PM UTC-4	11.3 kB	view	watch	download
log/SmartAdminEvents_Backup_0_131504930594010000.xel	September 21, 2017 at 2:44:19 PM UTC-4	10.8 kB	view	watch	download
log/SmartAdminEvents_0_131134304590010000.xel	July 19, 2016 at 9:09:20 PM UTC-4	5.6 kB	view	watch	download
log/SmartAdminEvents_0_131504930588000000.xel	September 21, 2017 at 2:44:18 PM UTC-4	5.1 kB	view	watch	download
log/log_5.trc	September 21, 2017 at 2:43:26 PM UTC-4	1 MB	view	watch	download

Imagen 2-2. Visualización de los registros de la base de datos mediante la consola de Amazon RDS

Monitoreo y notificaciones

Los proveedores de nube ofrecen una gran cantidad de información sobre las métricas, como el uso de la CPU, la latencia y el rendimiento de las consultas, y el uso de E/S y de la memoria. Ofrecen una diversidad de capacidades de monitoreo incorporadas, algunas gratuitas y otras por un costo adicional. También puede establecer alertas que le avisen con rapidez si tales medidas superan los límites previstos.

Muchas plataformas en la nube exponen eventos, como cuándo ocurre una falla, cuándo se realiza una copia de seguridad y cuándo se detiene o reconfigura una base de datos. La plataforma proporciona integraciones que le permitirán establecer notificaciones o tomar otras medidas automáticas cuando ocurra un evento.

Optimización de rendimiento

El ajuste del rendimiento de una base de datos es una de las tareas clave de un DBA tradicional: distingue a los buenos DBA de los regulares. Un buen ajuste de rendimiento requiere un conocimiento detallado del almacenamiento, la red, la memoria y el código de la aplicación. La migración de las bases de datos a la nube no libera al DBA de todas sus responsabilidades relacionadas con ellas, y el ajuste del rendimiento, tanto reactivo como proactivo, seguirá siendo una de sus responsabilidades principales. Sin embargo, los proveedores de nube ofrecen algunas maneras sencillas de obtener información, como **AWS RDS Performance Insights**.

En la nube, no pasará mucho tiempo preocupándose por los típicos problemas de rendimiento relacionados con el servidor y el almacenamiento, ya que podrá aumentar o disminuir el escalado de sus bases de datos según las cargas de trabajo. Es más probable que pase su tiempo analizando los tiempos de espera de la base de datos para mejorar sus tiempos de respuesta.

En los entornos locales tradicionales, la misma base de datos proporciona algunas herramientas de rendimiento, como el comando EXPLAIN y el registro de consultas lento de MySQL. Otras herramientas de monitoreo del rendimiento se ejecutan por fuera de la base de datos, y puede establecer un sistema en la nube para ejecutarlas. Pero los proveedores de nube han desarrollado herramientas de rendimiento que se ejecutan de manera nativa en la nube que suelen ofrecerse sin costo adicional. La mejor opción sería utilizar lo que le brinda el proveedor.

¡Las bases de datos que tienen un rendimiento deficiente localmente no comenzarán a rendir mejor por arte de magia solo por haberlas migrado a la nube! Pero, la migración puede ayudar en algunos casos, como cuando el entorno de la base de datos local no contaba con la capacidad para admitir una demanda de la aplicación de una gran cantidad de IOPS.

Manejo de la alta disponibilidad

Esta tarea es fundamental en la actualidad, cuando las personas abandonan los sitios web que no logran **cargarse en dos o tres segundos**. En las instalaciones físicas, esta tarea requiere la adquisición anticipada de una potencia de procesamiento adecuada, además de la configuración de balanceadores de carga. En la nube, puede optar por un escalado automático y configurar su base de datos para que se ejecute en varias zonas. Una base de datos de varias zonas puede **realizar una conmutación por error hacia otra AZ automáticamente**, si la primera zona falla. Asimismo, puede ofrecer sus datos en varias regiones tanto para una disponibilidad alta como para un almacenamiento local.

Nuevas tareas para el DBA en la nube

Cuando ejecuta sus bases de datos en un entorno de nube, debe ocuparse de tareas que, en la mayoría de los casos, no eran su responsabilidad en los centros de datos. Estas tareas incluyen una nueva orientación hacia los costos y el seguimiento de las licencias de las bases de datos en la nube.

Planificación de los costos operativos

Los cálculos fundamentales para el costo total de propiedad (TCO) y el rendimiento de la inversión (ROI) cambian significativamente en la nube. Como mencionamos en **“Aprovisionamiento” en la página 14**, se suele pagar por instancias y tiempos particulares en la nube en lugar de pagar por adelantado por una infraestructura rígida. Sin embargo, debe tener en cuenta muchas otras cosas que no están directamente relacionadas con las bases de datos. Por ejemplo, la ubicación de la

base de datos en regiones cercanas a los usuarios puede reducir los costos de las redes y, quizá, hasta puede generar nuevos negocios.

Los proveedores ofrecen una gran cantidad de opciones que se adaptan al uso que hace de ellas y a sus necesidades, y que reducen significativamente los gastos. Por ejemplo, si compra instancias reservadas en AWS, tal vez pueda reducir los costos drásticamente. En general, una planificación meticulosa permite que prediga sus necesidades en un momento en particular y elimine el sobreaprovisionamiento que deben realizar los centros de datos en las instalaciones.

Efectos en el diseño de la arquitectura y las aplicaciones

La arquitectura de las bases de datos basadas en la nube es intrínsecamente diferente de aquella correspondiente a las bases de datos locales tradicionales presentes en los centros de datos. Debe tener en cuenta varias regiones geográficas y AZ, diferentes tipos de diseños de redes y otras características específicas de la nube cuando planifica las implementaciones de su base de datos. El DBA debe estar en contacto estrecho con los desarrolladores de la aplicación para comprender sus necesidades y brindarles la mejor opción de la nube al precio más bajo posible.

Seguimiento del cumplimiento de la licencia

Las bases de datos de propiedad exclusiva se ofrecen en la nube mediante dos modelos básicos de licencias: con licencia incluida o "traiga su propia licencia" (BYOL). Este último es valioso si ya tiene una licencia y el modelo admite BYOL. Quizá deba consultar con un especialista en licencias o con el proveedor para descifrar si su licencia existente se aplica a la nube.

Seguridad para los datos y las aplicaciones en la nube

Las introducciones a la informática en la nube suelen organizar las funciones de un administrador en una pila, con la instalación de servidores físicos y cableados en la base de la pila y la administración de la aplicación en la parte superior. Las bases de datos administradas le asignan la responsabilidad de todas estas tareas en conjunto al proveedor de nube y solo dejan las tareas del nivel de la aplicación a cargo del cliente. Las bases de datos autoadministradas dejan la mayoría de las tareas en manos del cliente y el proveedor solo proporciona los servidores físicos y las redes básicas.

Las tareas de seguridad también se ajustan a este concepto de pilas y llevan a lo que se conoce comúnmente como un "modelo de seguridad compartida". El proveedor asume más tareas de seguridad en las bases

de datos administradas que en las autoadministradas. Para resumir, el proveedor de nube supervisa la seguridad de la infraestructura, pero usted sigue siendo responsable por la seguridad de sus datos y la información del usuario en la nube. Por lo tanto, las tareas realizadas por el proveedor de nube incluyen las siguientes:

- Asegurar físicamente los centros de datos y los equipos, incluida la supervisión del personal
- Garantizar la redundancia mediante replicaciones y copias de seguridad, conforme a las guías del cliente
- Aplicar parches y actualizar el software, incluida la base de datos en sí, si elige una base de datos administrada
- Ejecutar herramientas que monitoreen y auditen los accesos
- Proporcionar firewalls de redes y aplicaciones
- Garantizar la protección contra los ataques de DoS
- Servicios de administración de identidades, que configura usted mismo

Las tareas de seguridad del cliente incluyen las siguientes:

- Establecer los usuarios y roles mediante el sistema de administración de identidades y accesos (IAM) del proveedor
- Asignar derechos de acceso a los usuarios en la base de datos, lo que incluye la definición de las cuentas y los roles, y el mantenimiento actualizado de las cuentas mediante la eliminación de usuarios cuando ya no se requiere su acceso
- Especificar las reglas del firewall, las copias de seguridad y otros parámetros que ofrezca el proveedor de nube
- Establecer alertas y revisar los registros en busca de accesos no autorizados
- Encargarse de la seguridad del sistema operativo, si utiliza su propia VM en lugar de una base de datos administrada
- Encargarse de la seguridad a nivel de la aplicación, como la prevención de ataques de inyección SQL.
- Cifrar los datos, tanto en reposo como en tránsito
- Autenticar y autorizar a los usuarios que intentan acceder a las bases de datos o aplicaciones de la manera adecuada
- Recopilar e inspeccionar registros
- Establecer alertas de eventos y monitorearlos en busca de anomalías

El conocimiento de la gama completa de características de seguridad en la nube es esencial para ejecutar bases de datos allí. En las siguientes subsecciones, abordaremos las características clave de seguridad en la

nube en las que deberían enfocarse los DBA: control de accesos e IAM, aislamiento de redes y cifrado de datos.

Control de accesos y administración de identidades y accesos

Los proveedores de nube suelen destacarse en cuanto a los mecanismos sólidos de control de acceso. Los proveedores mencionados en este informe dependen de un servicio de IAM centralizado para administrar los usuarios, las credenciales de seguridad (contraseñas, claves de acceso y permisos) y las políticas de autorización que controlan cuáles son los recursos y servicios a los que pueden acceder los usuarios. Los administradores deben dominar IAM solo para obtener acceso a la nube por su cuenta y en nombre de los usuarios. Además, las bases de datos nativas de la nube suelen estar integradas con herramientas generales de IAM en la nube. La conexión de las protecciones de la base de datos con el servicio de IAM de la nube proporciona las políticas de acceso más sencillas y seguras.

Utilizará IAM para definir cuentas de usuarios y, luego, agregar reglas de acceso específicas de la base de datos. Con el uso de IAM, puede otorgar diferentes permisos de usuario para realizar distintas operaciones en la base de datos. También puede establecer controles de acceso pormenorizados (que la mayoría de las bases de datos ofrecen, como restricciones en filas o columnas específicas) mediante IAM.

Las bases de datos tradicionales utilizan las mismas herramientas en la nube utilizadas localmente, como los comandos GRANT en SQL. Pero tal vez pueda **conectarlas a IAM** para utilizar las mismas cuentas de usuario tanto en la nube como en la base de datos y beneficiarse de la seguridad adicional y la conveniencia que ofrece esa integración.

Aislamiento de redes

Algunas características comunes de la nube para proteger sus sistemas en la red son las VPC, los firewalls y las listas de control de acceso de red (ACL).

Como analizamos en **“Efectos de alto nivel producto de la migración a la nube” en la página 2**, una VPC es una red privada dentro de la nube que se usa para la comunicación entre sus servidores. Dentro de una VPC, puede aislar instancias de las bases de datos mediante la especificación de un rango de IP que tenga permitido el acceso a cada una de ellas. Las organizaciones que crean una VPC poseen el control total de sus entornos de redes virtuales y pueden seleccionar sus propios rangos de

direcciones IP, crear subredes y configurar sus propias tablas de ruta y gateways de red.

Además, puede configurar una gateway privada virtual que extienda su red corporativa hasta la VPC y permita el acceso a las instancias de las bases de datos en dicha VPC mediante una VPN que usted elija.

La mayoría de los proveedores de nube ofrecen firewalls incorporados que lo ayudan a controlar el acceso a la red de las instancias informáticas. Quizá, el proveedor de nube también le ofrezca una opción de conectividad privada o dedicada para conectar la oficina del consumidor en la nube y los entornos locales con el entorno de la nube. Puede establecer grupos de seguridad de bases de datos para asegurar las instancias de las bases de datos dentro de una VPC. Los grupos de seguridad son reglas de firewall que controlan el acceso a la red de sus bases de datos en la nube. También puede permitir o denegar que el tráfico de red ingrese y salga de una subred mediante las ACL de la red. Todo el tráfico de red que ingresa o sale de su VPC mediante la VPN se puede inspeccionar con la infraestructura de seguridad local (como un firewall) y con sistemas de detección de intrusos.

NOTA

Conexiones directas

En lugar de conectar una VPN, puede conectar sistemas externos a los sistemas dentro de la nube mediante **Direct Connect** en AWS o **ExpressRoute** en Azure. Estos sistemas aprovechan los enlaces de redes privadas proporcionados por los operadores de telecomunicaciones para crear una conexión directa. El cifrado integral se sigue recomendando y se suele realizar mediante una capa de conexión segura (SSL)/capa de transporte seguro (TSL) estándar.

Cifrado de datos

En la nube, es fácil cifrar los datos para proporcionar una protección adicional a los datos en reposo. Por ejemplo, cuando habilita el cifrado en un clúster de Amazon RDS, la base de datos almacena todos los datos en las tablas que crea en un formato cifrado. El cifrado también se aplica a las copias de seguridad de la base de datos. Asimismo, el cifrado es fácil de configurar cuando transmite datos desde y hacia la nube.

El cifrado es particularmente necesario para las organizaciones que deben cumplir con los requisitos de la industria, como la Ley de Portabilidad y Responsabilidad del Seguro de Salud (HIPAA) para la atención médica, la Ley Sarbanes-Oxley (SOX) para los informes financieros y el Estándar de Seguridad de Datos para la Industria de

Tarjetas de Pago (PCI DSS) para las empresas de comercio electrónico y ventas minoristas. Si protege sus claves de cifrado, la pérdida de los datos cifrados no suele considerarse como un evento de seguridad que se deba notificar.

Los proveedores de nube suelen ofrecer servicios administrados para simplificar la creación, el control y la administración de sus claves de cifrado. Por esta razón, AWS Key Management Service (KMS) ofrece una vista centralizada de todos los usuarios clave de la organización. Se integra con AWS CloudTrail para proporcionar un registro que muestra el uso de las claves en la organización y, de esta manera, satisface varios requisitos clave normativos y de conformidad.

Infraestructura como código: sacarle provecho a la nube

Cuando migra sus bases de datos a la nube, puede surgir la tentación de limitar los cambios administrativos al mínimo porque está acostumbrado a hacer las cosas de una manera determinada. Sin embargo, esta estrategia evitará que capitalice completamente los inmensos beneficios que la nube pone a sus pies.

Un avance tecnológico clave que distingue los sistemas basados en la nube de aquellos en los centros de datos locales es la disponibilidad de los sistemas de configuración que tratan su infraestructura como un código. El tratamiento de la infraestructura como un código permite muchas prácticas de DevOps que, a su vez, facilitan la colaboración estrecha entre los desarrolladores y las operaciones para que puedan automatizar los resultados de la aplicación a escala. A nivel local, definitivamente puede instalar herramientas de automatización como sistemas de administración de configuraciones (Puppet y Chef son ejemplos populares) y herramientas de integración continua (CI), como Jenkins. Pero, la nube ya tiene incorporadas herramientas muy sofisticadas de este estilo.

Por ejemplo, las plantillas proporcionadas con AWS CloudFormation lo ayudan a moldear toda su infraestructura como un código. Puede definir la infraestructura en la nube mediante la creación y configuración de recursos como tablas de bases de datos y buckets de almacenamiento (Amazon S3) y tratar a estos recursos como códigos. Puede controlar las plantillas de AWS CloudFormation en su sistema de control de orígenes y administrarlas de la misma manera en que los desarrolladores administran sus archivos de códigos.

Una herramienta como CloudFormation ofrece los siguientes beneficios:

- Ayuda a establecer una sola fuente de verdad para todos sus recursos de la nube.
- Se puede integrar con herramientas de administración de códigos, como un sistema de control de orígenes.
- Ayuda a automatizar las implementaciones de desarrollo y de pruebas.
- Admite sus planes de recuperación de desastres.

Cuando utiliza los archivos de las plantillas para crear recursos de manera programática, consigue repetibilidad y consistencia en todos sus entornos.

Migración de sus bases de datos a la nube

Como la mayoría de las transformaciones organizacionales, la migración a la nube no se lleva a cabo de un día para el otro. Debe seleccionar un proyecto a utilizar como prueba en su proveedor de nube elegido: puede ser un proyecto existente que proporcione un caso de prueba útil o un proyecto nuevo que esté libre de prácticas heredadas.

Casi todas las organizaciones que migran a la nube lo hacen mediante una migración de prueba de concepto de una base de datos prescindible. Cuando haya elegido la base, o cuando haya migrado con éxito un proyecto pequeño que se beneficie con la nube, podrá extender lo aprendido a otras bases de datos de su organización.

Este capítulo lo ayudará a comprender los criterios que suelen estar asociados con una migración inicial exitosa. Gran parte del material se aplica a las bases de datos relacionales. Aunque cada migración es única, seguramente llevará a cabo la mayoría de los siguientes pasos o todos ellos:

Planificación

- Recopilación de requisitos
- Determinación de las capacidades para abordar los requisitos
- Evaluación de qué bases de datos migrará y qué cambios podrían ser necesarios para la base de datos o las aplicaciones que la utilicen
- Determinación de criterios de éxito y restauración (mecanismos a prueba de errores)

Migración de datos

- Replicación
- Incorporación de los cambios desde la creación de la réplica
- Pruebas de aplicaciones
- Etapas de transición
- Controles posteriores a la migración

Optimización

- Ajustes de rendimiento
- Diseño de la alta disponibilidad
- Determinación de qué eventos registrar y monitorear
- Creación de un plan de recuperación de desastres

Analizaremos cada etapa principal en las siguientes secciones.

Planificación

Esta fase ayuda a la organización a evaluar los siguientes aspectos:

- Los elementos relevantes del entorno actual: aplicaciones, bases de datos y cargas de trabajo críticas
- Si la aplicación o la carga de trabajo se ejecutará adecuadamente en el entorno del proveedor de nube
- Cómo la migración ayudará a alcanzar los objetivos empresariales
- Los requisitos para el estado final del sistema
- El costo de la ejecución del entorno informático actual en la nube, que debería representar un rendimiento de la inversión integral, según lo explicado en ["Planificación de los costos operativos" en la página 18](#)

Factores en una migración

Cuando desea migrar una base de datos que se ejecuta localmente a un servicio administrado en la nube, debe lidiar con problemas físicos, de software y organizacionales. Debe prestar atención a varios niveles de cambios:

Migración física de datos

Debe trasladar los datos a la nube. La transferencia de datos de bases de datos grandes puede consumir mucho tiempo y ser muy costosa, pero existen servicios de migración que pueden ayudar.

Compatibilidad de la infraestructura

Compare cada aspecto del servicio en la nube con sus sistemas heredados locales. Un [caso de estudio en el blog de AWS](#) ofrece un ejemplo interesante de problemas imprevistos que podrían ocurrir durante una migración de prueba debido a incompatibilidades inesperadas del software y errores menores. Esta migración en particular se demoró durante un tiempo debido a las diferencias en las implementaciones SSL de los sistemas locales y el proveedor de nube.

Compatibilidad de las bases de datos

Es probable que la base de datos en la nube difiera de la base de datos local respecto de determinadas características que afectan la migración. Ciertas sutilezas, como los conjuntos de caracteres de la base de datos y los permisos relacionados con los procedimientos almacenados, podrían presentar dificultades. Una migración homogénea (como la de una base de datos de Oracle a otra base de datos de Oracle) seguramente será más sencilla que una migración heterogénea (como la de una base de datos de Oracle a una base de datos de PostgreSQL o a una de las bases de datos de propiedad exclusiva de los proveedores de nube). En la ["Comprobación de incompatibilidades" en la página 31](#), se explica lo que puede hacer para reducir la fricción entre las bases de datos. Para determinadas clases de migraciones, los proveedores de nube podrían proporcionar herramientas que lo ayudarán a alivianar sus responsabilidades.

Si actualmente ejecuta una versión muy antigua de una base de datos local, podría ser el momento de realizar una actualización. En lugar de simplemente elegir una versión más nueva del mismo sistema de administración de bases de datos relacionales (RDBMS), lo que conllevará determinado grado de dificultad, podría aprovechar la oportunidad para intentar llevar a cabo una migración heterogénea a un nuevo RDBMS que resulte prometedor.

Cambios organizacionales

Las personas de diferentes equipos, en especial los DBA y los desarrolladores de aplicaciones, tendrán que invertir tiempo en aprender a manejar las herramientas de la nube y administrar las nuevas instancias mientras mantienen en funcionamiento los

sistemas heredados locales hasta que esté listo para migrar todo a la nube. Podría elegir pasar meses o incluso años realizando una migración completa. También podría mantener algunos de los datos (o una copia de estos) de manera local en forma permanente.

Tareas de migración importantes

Sugerimos que tome los siguientes pasos para migrar con éxito.

Creación de un plan de migración a la nube

El plan de migración a la nube debe enumerar todas las bases de datos y las aplicaciones en el orden en que desea trasladarlas a la nube. El plan de evaluación final debería delinear el plan de migración para todas las bases de datos. Si necesita recursos nuevos (de software, finanzas o personal) en el nuevo entorno de la base de datos, también tendría que incluirlos en el plan.

Determinación de quién realizará la migración

Los DBA y los desarrolladores deberían trabajar juntos en el esfuerzo de migración porque cada equipo tiene algo que ofrecer. Durante la elección del personal, también puede decidir si empleará un servicio de migración proporcionado por el proveedor de nube o por terceros.

Tareas educativas

Incluyen una amplia gama de tareas, como la comprensión del uso de las herramientas del proveedor de nube, las características específicas del motor de la base de datos de destino, el alcance de la migración de la base de datos y la arquitectura de las bases de datos en la nube.

Creación de la arquitectura de la base de datos en la nube

Seleccione los tipos de bases de datos que desee utilizar. Además, debe decidir si utilizará una base de datos administrada o una autoadministrada. Cuando tome esta decisión, tenga en cuenta todos los factores relevantes, como los costos, el rendimiento, la confiabilidad y la escalabilidad.

Elección de un proceso de migración

En **"Migración de la base de datos"** en la [página 32](#), se establecen los criterios de selección entre las replicaciones, las copias de seguridad/restauraciones o un servicio de migración dedicado.

Creación de su infraestructura informática para una base de datos autoadministrada.

Si elige administrar sus propias bases de datos en la nube, cree instancias virtuales e impleméntelas mediante el servicio informático del proveedor, como Amazon Elastic Compute Cloud (Amazon EC2).

Determinación de las oportunidades de realizar cambios en la arquitectura
Podría descubrir que las decisiones que tomó a nivel local ya no resultan adecuadas para el nuevo entorno en la nube. Por ejemplo, quizá pueda consolidar particiones.

Configuración de las cuentas de la base de datos en la nube
Configure cuentas generales para los usuarios en la nube mediante IAM antes de establecer los usuarios, roles y grupos en la base de datos.

Actualización y pruebas de las aplicaciones
Quizá deba reescribir el código para trabajar con la nueva base de datos. Algunas migraciones de aplicaciones podrían automatizarse, al igual que la base de datos en sí.

Evaluación de la preparación

Una evaluación de su preparación lo ayudará a estimar los costos, la arquitectura de la base de datos en la nube, los planes de migración y el impacto de la migración a la nube en las normas de conformidad. El resultado de una evaluación de preparación para migrar a la nube es un informe detallado de la preparación de su empresa para la migración de sus bases de datos a la nube.

A continuación, desarrollaremos descripciones breves de los pasos clave durante una evaluación de la preparación para migrar a la nube.

Entrevistas con las partes interesadas
Hablar con los desarrolladores de aplicaciones, los usuarios comerciales y otras partes vinculadas con el uso de los datos dentro del equipo de su organización lo ayudará a determinar los requisitos relacionados con el rendimiento, la alta disponibilidad y las características y capacidades de las bases de datos basadas en la nube.

Análisis de las bases de datos locales existentes
Analice sus bases de datos existentes para determinar los patrones de crecimiento de los datos, las estrategias para realizar copias de seguridad y de recuperación, las exportaciones e importaciones de datos en curso, etc. La comprensión de los patrones de uso actuales de las bases de datos lo ayudará cuando sea el momento de decidir cuáles utilizar en la nube y si necesitará una oferta de bases de datos específica para obtener las capacidades necesarias.

Priorización de las bases de datos a migrar
En primer lugar, elija las bases de datos que le gustaría migrar a la nube. Un criterio importante tiene que ver con la cantidad de

cambios que necesitará, lo que, a su vez, requiere la coordinación con los desarrolladores.

Análisis del costo de la migración

Realice un análisis exhaustivo de los costos de migrar a la nube en comparación con los costos de permanecer en los centros de datos locales, para comprender totalmente el TCO y el ROI de la migración a la nube.

La reducción de costos mediante el uso de un servicio de bases de datos en la nube podría ser uno de sus objetivos principales. Pero una elección inadecuada de los servicios en la nube podría frustrar ese objetivo.

Seguridad y conformidad

A veces, necesita regiones o AZ especiales para cumplir con estándares como los informes de control de organización de servicios 2 (SOC 2), PCI DSS o la ley HIPAA de Estados Unidos. Afortunadamente, los proveedores de nube suelen combinar regiones y AZ específicas con estos requisitos legales. Además, existen regiones especializadas, como AWS GovCloud (Estados Unidos), que es una región aislada de AWS sujeta a los puntos de referencia altos y moderados de FedRAMP (Programa de Administración de Autorizaciones y Riesgos Federales). Por último, debe verificar si cumple con los acuerdos a nivel de servicio (SLA) de su organización en la configuración elegida de la nube. Estos SLA suelen incluir métricas para los mantenimientos planificados, las copias de seguridad, los objetivos de punto de recuperación (RPO) y los objetivos de tiempo de recuperación (RTO).

Si ha realizado la evaluación de la preparación correctamente, también tendrá un buen panorama del alcance de su migración a la nube. Algunas de sus aplicaciones podrían ser tan antiguas que deberían rediseñarse completamente para migrarlas a la nube. Si no cuenta con los recursos para reescribir las aplicaciones, puede mantenerlas en las bases de datos locales o migrar los servidores completos a la nube.

Pasos para migrar

Debe planificar exactamente cuándo y cómo realizará la migración de los datos, junto con las preparaciones para hacer restauraciones si las pruebas posteriores demuestran que la base de datos no funciona adecuadamente. Establezca pruebas y criterios para alcanzar el éxito en cada paso.

Cambios en las aplicaciones

Debe realizarlos en coordinación con el equipo de desarrollo.

Esto incluye la programación de los cambios en las aplicaciones dentro del proceso de desarrollo y las pruebas de desarrollo.

Automatización de la migración

Esto es muy valioso porque le permite repetir la migración tantas veces como sea necesario y arreglar las fallas cada vez que ocurren o cambiar los parámetros. También puede adaptar y reutilizar el marco de automatización para migraciones posteriores.

Comprobación de incompatibilidades

Si utiliza los mismos RDBMS (como MySQL, por ejemplo) para ambas bases de datos, compruebe las versiones para determinar si está utilizando características que la versión en la nube no admite. Obviamente, es incluso más probable que estas comprobaciones encuentren incompatibilidades si realiza una migración a un nuevo motor de bases de datos.

Los problemas de conversión que encuentre durante una migración heterogénea a una base de datos en la nube son los mismos que encontrará al realizar una migración local. Sin embargo, tendrá acceso a un rango más amplio de soluciones.

Los proveedores de bases de datos suelen ofrecer herramientas que ayudan a migrar los datos a una base de datos en la nube del mismo tipo. Por ejemplo, Oracle incorpora herramientas de migración de datos como Data Pump, RMAN y SQL Developer que lo ayudan a migrar sus bases de datos de Oracle a la nube. Además, los proveedores de nube y otras entidades ofrecen servicios como [AWS Schema Conversion Tool](#) para migrar esquemas. Estos servicios establecen si los esquemas se pueden traducir y le sugieren soluciones alternativas en caso de ser necesario.

Aun así, la mayoría de las migraciones de bases de datos heterogéneas requieren tanto herramientas automatizadas como acciones manuales por parte del DBA. Si realiza una migración a una base de datos de AWS, puede aprovechar las [recomendaciones específicas del sitio web de AWS](#).

Algunas herramientas de migración utilizan un traductor para convertir los objetos de una base de datos (como los procedimientos almacenados de Oracle) en objetos de una base de datos externa a Oracle, como PostgreSQL. Sin embargo, estas herramientas podrían realizar un trabajo incompleto, debido a que cada motor de bases de datos utiliza prácticas de codificación únicas.

Esta es la razón por la que se realizan pruebas estrictas en la base de datos de destino para comprobar su precisión y rendimiento después de migrarla a la nube. Algunas técnicas de programación que utilizan RDBMS específicos

podrían no tener un equivalente exacto en la base de datos de destino. La herramienta de migración podría indicar esta clase de problemas de conversión durante el proceso de migración. En muchos casos, también puede mostrarle el código que quizá deba ejecutar para superar los problemas y completar la migración de la base de datos con éxito.

Migración de datos

Los proveedores de nube están muy alertas a los requisitos de sus clientes potenciales respecto de la migración de sus bases de datos locales a la nube con interrupciones y costos mínimos. Cada proveedor ofrece herramientas para agilizar la migración. AWS, por ejemplo, ofrece numerosos artículos que abordan en detalle cómo realizar una migración desde varias bases de datos locales a AWS, incluidos procedimientos para [Oracle](#), [MySQL](#) y [PostgreSQL](#).

Aun así, las migración a la nube pueden conllevar tiempos de inactividad disruptivos y costosos. Una herramienta de migración (proporcionada por el proveedor o por terceros) debe ser capaz de manejar todos los aspectos de la base de datos, como los esquemas, permisos de usuario, disparadores y procedimientos almacenados.

Le recomendamos que lleve un diario durante las migraciones iniciales porque las lecciones que aprenda en el camino les brindarán información a usted y sus colegas a medida que intenten realizar más migraciones. Si encuentra suficientes aspectos negativos que hagan que cambie de proveedor (o que eviten la migración por completo), el diario proporcionará evidencias importantes que respaldarán la decisión.

Migración de la base de datos

Muchos proveedores de nube y otras entidades ofrecen servicios de migración, como [Microsoft Azure Database Migration Service](#) y [AWS Database Migration Service](#), que vale la pena considerar. Los publicitan como rápidos, de funcionamiento sin contratiempos y fáciles de usar. Sin embargo, quizá prefiera utilizar procesos existentes que ya conozca, como la restauración a partir de una copia de seguridad a un nuevo entorno de bases de datos y el uso de replicaciones.

Si utiliza un servicio de migración de bases de datos, puede mantener las bases de datos de origen completamente operativas durante la migración a la nube y minimizar el tiempo de inactividad de las aplicaciones que dependen de la base de datos. Esto facilita que alcance el objetivo principal de tiempos de inactividad mínimos citado por muchas organizaciones que migran a la nube. En la [imagen 3-1](#) se ilustra cómo AWS Database Migration Service crea las tablas, carga los datos y mantiene las tablas sincronizadas con aquellas de la base de datos de origen.

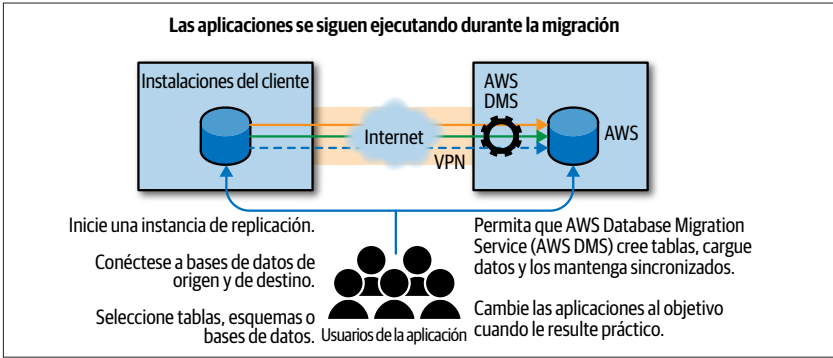


Imagen 3-1. Mantenimiento del funcionamiento de las aplicaciones durante la migración de las bases de datos

El proceso de migración suele centrarse en un proceso típico de copias de seguridad/restauraciones. Las tareas dependen de su capacidad de desactivar la base de datos durante la transferencia. Si debe mantener la base de datos en funcionamiento, esta es la secuencia básica que debe seguir:

1. Realizar una copia de seguridad dinámica de la base de datos original durante su ejecución.
2. Copiar y restaurar la copia de seguridad en la nube.
3. Permitir que las replications incorporen en la versión de la nube todos los cambios realizados en la base de datos original después de la creación de la copia de seguridad.

Claramente, las copias de seguridad/restauraciones se simplifican si puede desactivar la base de datos. Entonces, la migración sigue este patrón:

1. Desactivar la base de datos original.
2. Realizar una copia de seguridad.
3. Copiar y restaurar la copia de seguridad en la nube.
4. Iniciar la nueva versión en la nube.

Sin embargo, es probable que invierta tiempo en realizar pruebas de la nueva versión en la nube antes de estar listo para comenzar su producción con el fin de garantizar que las aplicaciones funcionen, que la seguridad esté habilitada correctamente y que el rendimiento sea el esperado. Así es que quizá quiera reiniciar la versión original antes de estar listo para confiar en la nueva versión. Luego, utilizará la replicación para sincronizar la nueva versión con la original.

Migración de aplicaciones

Existen dos estrategias básicas para migrar las bases de datos y las aplicaciones relacionadas con estas:

Lift and shift

Migre la base de datos completa en su estado actual, con todas las aplicaciones que admite, a la nueva base de datos en la nube. Esta estrategia casi no utiliza las características nativas de la nube.

Rediseño de las bases de datos

Reanalice cómo las bases de datos locales se encargan de las tareas administrativas (como la escalabilidad y la alta disponibilidad) e investigue las ofertas del proveedor de nube en esas áreas, ya que se vuelven cada vez más sofisticadas. Por supuesto, esta estrategia requiere más trabajo y esfuerzo. También podría necesitar recodificar las aplicaciones para ajustarlas a los cambios en la estrategia de la base de datos.

Es común que los clientes comiencen con un proceso “lift and shift” y continúen con el rediseño y otras modernizaciones.

Controles posteriores a la migración

Después de copiar completamente los datos en la base de datos en la nube, valide la base de datos de destino para asegurarse de que todos sus objetos estén presentes allí. Existen varias pruebas para determinar si la migración fue exitosa:

Validación de datos

Puede ser tan simple como controlar la cantidad de filas o ejecutar una suma de comprobación para indicar que no hubo pérdidas ni daños. Los procesos de validación más complejos incluyen los cambios en el esquema inherentes de las migraciones heterogéneas. Por desgracia, esta validación básica suele pasarse por alto.

Funcionalidad básica

Realice pruebas integrales para asegurarse de que la migración se haya realizado correctamente y el sistema funcione como debería. Analice el uso de las características que tienden a diferir entre los motores de las bases de datos y sus versiones, como los disparadores y procedimientos.

almacenados. *Pruebas de rendimiento*

Realice pruebas de resistencia en los nuevos sistemas y compare su rendimiento para tener una idea de dónde puede mejorar los tiempos de respuesta y el rendimiento.

Evaluación de la seguridad

Garantice la seguridad de los sistemas en la nube mediante evaluaciones de vulnerabilidad y pruebas de penetración.

Estos pasos también forman una transición a la siguiente fase: la optimización. Esta fase parte de la migración exitosa de sus bases de datos y aplicaciones locales a la nube. Los elementos cruciales en esta fase son la administración del rendimiento y la optimización de costos. Es fundamental que supervise los nuevos sistemas cuidadosamente y de manera continua para poder regresar a los sistemas antiguos con rapidez en caso de encontrarse con problemas inesperados.

Optimización

En este punto, seguramente todo funcionará sin problemas y podrá felicitar a su equipo por haber hecho un buen trabajo. En esta sección se resumen las tareas principales que deben realizar el DBA y otros miembros del equipo para avanzar.

Disponibilidad

Este, por supuesto, es un requisito fundamental que debe cumplir en cualquier entorno. Afortunadamente, la migración a la nube elimina algunas causas de error (si configura correctamente el reinicio de los servicios que hayan fallado). Además, el proveedor de nube ofrece numerosas herramientas para predecir fallas y alertarlo ante la posibilidad de que surja alguna. Para aprovecharlas, realice la planificación descrita en esta sección (muchas de estas actividades también son útiles para la optimización de rendimiento):

Planificación para la recuperación de desastres (DR)

Con un proveedor de nube grande y sólido, puede realizar conmutaciones por error a una nueva AZ, si la actual falla. Pero necesita tanto un plan como un proceso automatizado para conmutaciones por error. El plan debería abordar el RTO, el RPO y la redundancia geográfica. Aproveche al máximo las capacidades de conmutación por error proporcionadas por el proveedor en lugar de crear las suyas.

Monitoreo de registros y del sistema

Determine los eventos que pueden indicar problemas inminentes, además de los problemas que ya han ocurrido. Estos se pueden incorporar en un seguimiento automatizado. El seguimiento debería transmitir la información necesaria para establecer el origen de las fallas: por ejemplo, si surgen de la acción de un usuario (como el reinicio de un servicio), de un ataque o de otros cambios en el entorno. Algunas fallas pueden ser consideradas normales y puede lidiar con ellas mediante el uso de las herramientas automatizadas. Debería registrarlas, pero no es necesario emitir alertas para el administrador.

Monitoreo de cambios

Los administradores deben saber qué cambios en la configuración de la base de datos, el tamaño de las instancias o la topología de los clústeres pueden afectar la disponibilidad. Los entornos de desarrollo modernos utilizan procesos sólidos para realizar el seguimiento de cambios y el control de versiones a fin de que todos los cambios pasen por un proceso de vetado y puedan ser revertidos.

Pruebas del sistema

Intente determinar los puntos débiles del sistema y anticipar las fallas. Algunos equipos realizan ejercicios “pre-mortem” para identificar y eliminar orígenes potenciales de fallas. Los sitios grandes tienen la capacidad de desactivar sistemas deliberadamente y analizar si la recuperación es adecuada. Esta clase de evaluación, llamada *ingeniería del caos*, se popularizó gracias a **Chaos Monkey** de Netflix. Así como realiza restauraciones regulares para garantizar que las copias de seguridad funcionen, debe realizar pruebas de sus procedimientos de recuperación.

Optimización de rendimiento

El rendimiento puede beneficiarse de las prácticas expuestas en la sección anterior, especialmente del monitoreo. El monitoreo del rendimiento le permite determinar las relaciones entre los eventos y los cambios en las métricas de la base de datos, además de visualizar las discrepancias entre las tendencias de rendimiento predichas y las existentes. Asimismo, el rendimiento se puede mantener y mejorar mediante procesos adicionales.

Pruebas de las cargas de trabajo

El crecimiento del tamaño y la complejidad de los datos, además de los cambios en el comportamiento de las aplicaciones, afectan el rendimiento. Compruebe el rendimiento regularmente para aprender acerca de la degradación antes de que sus clientes le comenten sobre ella. Luego, podrá escalar o realizar otros cambios para adaptarse. La actualización de las estadísticas del caché y las tablas de las bases de datos podría demorarse después de realizar cambios importantes, lo que significa que los planes para realizar consultas y el rendimiento general podrían tardar un poco en regresar a los niveles deseados.

Optimización distribuida

En los entornos de aplicaciones que cambian con rapidez, los desarrolladores comprenden mejor las consultas de las bases de datos que el DBA. Si se encargan de la comprobación de consultas y su optimización, podrán obtener un buen rendimiento más rápido que el DBA, esto lo liberará para que realice optimizaciones y planificaciones a largo plazo.

Como conocen la importancia del rendimiento, los proveedores de nube han insertado rastreos y han expuesto una gran cantidad de estadísticas para los usuarios. Por ejemplo, AWS RDS Performance Insights le permite extraer datos históricos en un panel y visualizar gráficos que muestran la carga impuesta por operaciones particulares, los usuarios o los comandos SQL en la base de datos. Puede consultar el panel para responder preguntas sencillas como: "¿Cuándo se sobrecargan mis alojamientos?", o preguntas más complejas como: "¿Por qué esta consulta produce un rendimiento deficiente?". Azure ofrece tanto monitoreo de rendimiento como **ajustes automáticos**. Otras entidades también producen herramientas de monitoreo y rendimiento en la nube.

Adaptación a las diferencias en la nube

Por último, los proveedores ofrecerán nuevos servicios, nuevas opciones de VM, nuevo hardware (como SSD) y otros cambios que le pueden proporcionar grandes beneficios. Siempre debe estar al tanto de lo que el proveedor puede ofrecerle para poder descargar trabajos en sus procesos estandarizados y comprobados. A medida que lo haga, actualice los procedimientos operativos documentados. Finalmente, puede informar los eventos anormales al servicio de atención al cliente del proveedor, indicando la urgencia del evento.

Conclusión

Una migración a la nube es un proceso a largo plazo. Comience de a poco porque descubrirá que tiene mucho que aprender en el recorrido. Lleve un diario y registre con honestidad todos los errores y problemas encontrados. No se avergüence si los sistemas heredados albergan prácticas deficientes o errores que surgen durante la migración: eso ocurrirá en casi todas las organizaciones. La documentación de los problemas es lo mejor que puede hacer por su empresa.

Con suerte, una o más de sus migraciones iniciales saldrán bien y estará listo para realizar una migración importante a la nube. Hacerlo puede generar beneficios en los costos, la flexibilidad y la seguridad. También es importante destacar que la migración a la nube le proporcionará un entorno informático actualizado que lo ayudará a atraer personal experto que desee participar de los últimos avances a su empresa.

Conclusión

Durante muchos años después de que Amazon.com abrió la primera oferta importante en la nube, la prensa comercial hizo la pregunta que enfrentan los administradores de sistemas y los DBA: "¿Migrar o no migrar a la nube?". Enseguida, las nubes locales y las ofertas híbridas combinaron soluciones puras de la nube como opciones para tener en cuenta. Pero las decisiones siempre fueron más complejas. Como se ha demostrado en este informe, las ofertas se han multiplicado rápidamente. Los DBA deben evaluar las bases de datos en simultáneo con los siguientes ejes:

- Proveedores externos, locales o híbridos
- Variedades relacionales o una de las diferentes variedades no relacionales
- Administradas o autoadministradas
- Nativas de la nube (como Amazon Aurora) o entre plataformas (MySQL)
- Aprovechar o no las mejoras de rendimiento, como los discos de estado sólido y el almacenamiento en caché
- Ubicaciones físicas de las regiones y las zonas de disponibilidad de la nube
- Facilidad de migración
- Habilidades relevantes necesarias y que ya posee el personal
- Otros aspectos del soporte y la reputación del proveedor

No es buena idea comprometerse prematuramente con una opción de un área determinada antes de analizar todas las opciones. Puede ser que ahorre mucho dinero y mejore la experiencia del cliente mediante capacitaciones adicionales o la adopción de tecnologías desconocidas.

Además de establecer los criterios básicos de selección de bases de datos, el objetivo de este informe es ayudarlo a prepararse para migrar a la nube al capacitarlo para enfrentar los cambios que podrían ocurrir en sus responsabilidades y tareas. Algunas responsabilidades y tareas se simplifican o eliminan con la migración a la nube, pero también tendrá que aprender acerca de nuevas tecnologías y comenzar a pensar en sus objetivos, como la alta disponibilidad y la optimización, de otras maneras.

Aprenderá mucho durante su primera migración o al comenzar un nuevo proyecto en la nube. Cada proyecto aclarará el escenario de las bases de datos en la nube y le dará ideas para sus siguientes proyectos. Esperamos que este informe lo ayude respecto de lo que debe tener en cuenta durante el proceso.

Acerca de los autores

Wendy A. Neu es una consultora principal de AWS Professional Services que se encarga de los problemas más difíciles de los clientes mediante la creación de sistemas de alta calidad, escalables y arquitectónicamente sólidos. Contribuye regularmente en AWS Database Blog (blog de bases de datos de AWS) y posee certificados de AWS, Oracle y Microsoft SQL Server. Antes de unirse a Amazon, trabajó como consultora en Cincinnati, Ohio, donde ayudaba a los clientes a traducir las necesidades empresariales en soluciones tecnológicas funcionales.

Vlad Vlasceanu es arquitecto principal de soluciones especialista en bases de datos de AWS, con base en la oficina de Santa Mónica, California. Vlad ayuda a los clientes a adoptar soluciones de bases de datos nativas de la nube, como Amazon Aurora, y a implementar arquitecturas a gran escala y de alto rendimiento en AWS. Se centra en diseñar e implementar cargas de trabajo sustentables, rentables y escalables que aprovechan las prácticas recomendadas y las capacidades más modernas que ofrece la plataforma de AWS. Antes de unirse a AWS, la carrera de Vlad incluyó más de 15 años en diseño y desarrollo de aplicaciones basadas en la web centradas en el consumidor y aplicaciones basadas en datos para el sector energético. Vlad posee una maestría en sistemas de ciencias de la información de la Universidad Baylor.

Andy Oram publicó la serie de Linux de O'Reilly, el innovador libro *Peer-to-Peer* (Entre pares) y el éxito de ventas *Beautiful Code* (Código hermoso). Andy también es el autor de muchos informes acerca de temas técnicos, como los lagos de datos, el rendimiento web y los software de código abierto. Sus artículos se publicaron en *The Economist*, *Communications of the ACM*, *Copyright World*, *el Journal of Information Technology and Politics*, *Vanguardia Dossier e Internet Law and Business*. Ha dado charlas en conferencias como O'Reilly's Open Source Convention (Convención sobre código abierto de O'Reilly), FISL (Brasil), FOSDEM, DebConf y LibrePlanet. Andy participa en la organización de políticas de la Association for Computing Machinery (Asociación de maquinaria informática), USTPC. También escribe para varios sitios web acerca de la TI aplicada a la atención médica, y los problemas en la informática y las políticas.

Sam R. Alapati es administrador de datos en Solera Holdings en Westlake, Texas. Forma parte del equipo de big data y Hadoop. Sam es un Oracle ACE, un reconocimiento que confiere Oracle Technology Network. Es el autor de *Modern Linux Administration* (Administración moderna de Linux) (O'Reilly, 2018), además de más de 20 libros de administración de bases de datos y sistemas. Sam tiene experiencia en el trabajo con los tres proveedores de nube más importantes: AWS, Microsoft Azure y Google Cloud Platform.