

Cómo debe ser la empresa que trabaja con datos

Cómo reinventar su análisis con el almacenamiento de datos creado específicamente para la nube



Contenido

Introducción	03	Herramientas para llevar sus datos a la nube	11
Lo mejor de cada casa: sin límites, escalabilidad rentable y rendimiento	04	Transferencias de datos físicos a gran escala	12
Ampliar, reducir y simultanear: cómo debe adaptar su tamaño un datawarehouse en la nube	05	Coste total de propiedad (TCO): presupuesto y gestión de costes	13
Perfiles de usuario: soluciones a las demandas habituales	07	Retorno de la inversión (ROI): el TCO y muchísimo más	14
La gran pregunta: ¿qué ocurre con la seguridad?	08	Su recorrido de cinco pasos para convertirse en un experto que trabaja con datos	15
Pero espere, que aún hay más: ¿un factor de seguridad adicional?	09	Acerca de PowerData	16
Seguridad y cumplimiento: respeto de los estrictos requisitos de la industria	10		



Introducción

Nunca es fácil ser un profesional de la gestión de datos. Los usuarios de todos los niveles y departamentos quieren obtener información útil a partir de todos los datos de su organización. Por si esto no fuera suficiente, también quieren entender los datos a un nivel que ni tan siquiera alcanza su solución tradicional de datawarehouse, lo que les provoca frustración. Mientras tanto, esperan tiempos de respuesta de smartphones tanto para consultas sencillas como descomunales.

Bombardeado a peticiones, está atrapado en el medio, haciendo malabarismos con tecnología heredada, local, en la nube o en ambas plataformas. Invierte una gran parte de su tiempo en solucionar un caos operativo que no debería suceder. Pero ¿qué ocurriría si su organización pudiese implementar una solución fácil y asequible que ofreciera a sus usuarios de análisis de datos e inteligencia empresarial (BI) lo que quieren cuando lo quieren? ¿Y si pudiera pasar de gestionar problemas diarios a defender realmente iniciativas de datos y tecnologías estratégicas?

Almacene sus datos en un Datawarehouse moderno creado específicamente para la nube

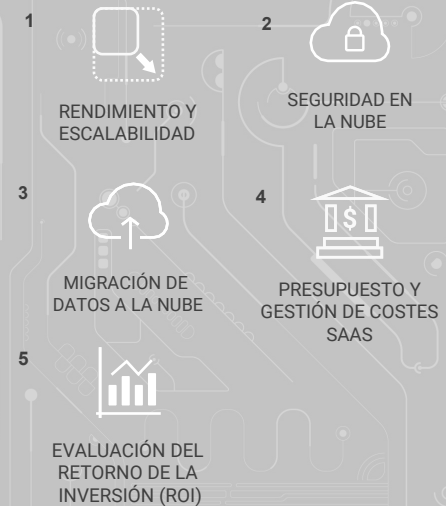
Libere sus datos de las restricciones heredadas con una arquitectura de datos en la nube moderna que ofrece mejoras fundamentales en materia de

rendimiento, simultaneidad y sencillez. Obtenga toda la información de todos sus datos con tiempos de consultas en segundos o minutos, no horas o días. ¿Cuál es el resultado? Dispondrá del tiempo necesario para dedicarse a los aspectos estratégicos.

Haga caso al «boom». Además de numerosas ventajas empresariales y de TI, un datawarehouse en la nube que realmente se beneficie de una arquitectura en la nube moderna puede aumentar el rendimiento hasta 200 veces por una décima parte del coste de los sistemas de datawarehouse heredados ubicados en las instalaciones o simplemente migrados a la nube. A continuación, le indicamos los pasos que necesita conocer para alcanzar este objetivo.

El rápido camino hacia el cambio

Esta guía contiene la información que necesita para entender un almacén de datos en la nube moderno y utilizarlo de manera exitosa. Le llevará por un camino que le permitirá transformar el análisis de datos de su empresa utilizando una «chuleta» basada en cinco temas clave que resultan cruciales para iniciarse en el almacenamiento de datos en la nube:



Lo mejor de cada casa: sin límites, escalabilidad rentable y rendimiento

Analice grandes cantidades de datos diversos de forma rápida

Hoy en día, sus usuarios empresariales lo quieren todo:

- Fácil acceso a conjuntos de datos dispares de gran tamaño.
- Integración de tipos de datos flexibles.
- Capacidad de profundizar a todos los niveles de datos para obtener repuestas rápidamente, independientemente del nivel de implicación de la consulta.

El usuario exige llevar los datawarehouses heredados más allá de sus límites, por lo que se ve empujado a ingeniárselas constantemente para evitar el fallo del sistema. Pero, con demasiada frecuencia, la contienda por recursos ilimitados le obliga a reprogramar trabajos o eliminar por completo la consulta de un usuario. Cuando esto ocurre, no es el experto de nadie.

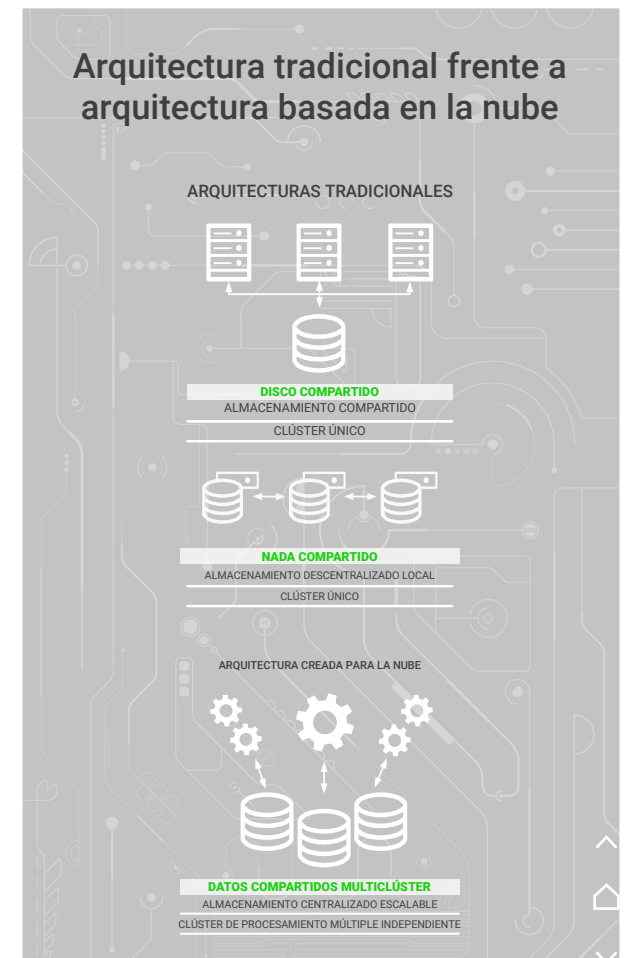
Bienvenido a una flexibilidad de recurso ilimitada

Con la llegada de la arquitectura en la nube, dispone de una nueva forma de pensar en el aumento del almacenamiento y del procesamiento: de forma potente y rentable. El almacenamiento de datos en la nube puede ofrecerle recursos ilimitados y la adaptabilidad que necesita para acceder a una potencia de procesamiento de cualquier

envergadura, pagando únicamente por lo que necesita, ya sea mensualmente, semanalmente, a diario o por horas. Con la nube, puede evitar el problema heredado del sobreaprovisionamiento para picos de demanda, teniendo que soportar el resto del tiempo un sistema infrautilizado. Asimismo, el almacenamiento en la nube puede costar una fracción de los dispositivos de almacenamiento ubicados actualmente en su centro de datos. Pero ¡ojo! Solo un datawarehouse moderno creado específicamente para la nube, uno que realmente separe el procesamiento del almacenamiento, puede beneficiarse de manera eficaz de todo lo que puede ofrecer la arquitectura en la nube.

A continuación, le ofrecemos algunos de los muchos problemas que puede paliar un datawarehouse en la nube moderno:

- **Competencia entre las consultas de los usuarios y las actividades de integración de datos que afecta adversamente al rendimiento.**
- **Usuarios forzados a analizar únicamente agregados y subconjuntos de datos para evitar sobrecargar aún más el sistema.**
- **La complejidad de la migración, la depuración y la transformación de los datos en entornos heredados.**



Ampliar, reducir y simultanear: Cómo debe adaptar su tamaño un datawarehouse en la nube

Varias formas de alcanzar el máximo rendimiento

El almacenamiento de datos en la nube puede ofrecerle recursos de procesamiento ilimitados dinámicamente y sin retrasos. Existen varias formas de ampliar, reducir y simultanear fácilmente el tamaño del datawarehouse para satisfacer la demanda y pagar únicamente por lo que usa:

1. Iniciar un nuevo clúster de procesamiento

Imaginemos que tiene una carga de trabajo que normalmente accede a la misma cantidad de datos todo el tiempo. Ha seleccionado el clúster del tamaño adecuado y ejecuta dichas consultas de manera oportuna. Pero ¿qué ocurre cuando los usuarios desplazan sus consultas de los datos del pasado mes para incluir cinco años de datos? Obtendrá un mejor resultado con un clúster nuevo de mayor tamaño.

No obstante, añadir más nodos de procesamiento al clúster existente puede no ser la mejor solución. La buena noticia es que, con el tipo adecuado de datawarehouse en la nube, puede definir rápidamente e iniciar un nuevo clúster en tan solo unos minutos, tal cual.

2. Mantenga un clúster de procesamiento predefinidos en modo suspendido, pero listo para ponerse en marcha.

Cuenta con un evento regular que requiere una oleada de recursos de procesamiento, las arquitecturas heredadas le obligan a crear o añadir un nuevo clúster desde cero cada vez que lo necesita, lo que consume tiempo y dinero. Por otra parte, el datawarehouse en la nube adecuado puede ofrecer un recurso de procesamiento predefinido en modo suspendido que puede activarse siempre que lo desee. Cuando los usuarios hayan terminado de utilizar el recurso, podrá apagarlo en modo «hibernación» hasta que vuelva a necesitarlo.

Y lo que es aún mejor, puede configurar el recurso para que entre en modo de suspensión automáticamente tras un tiempo predeterminado sin actividad. Una vez que se apaga el clúster, el contador también se detiene, lo que evita que pague por un recurso inutilizado.



3. Inicie un clúster adicional dinámicamente para administrar usuarios simultáneos

El número de consultas y usuarios simultáneos de un clúster puede dispararse repentinamente, dando lugar a una cola. Un verdadero datawarehouse en la nube puede aumentar automáticamente la simultaneidad mediante la creación de manera transparente de un nuevo clúster que equilibre la carga con respecto al primero.

clúster replicado tiene acceso a los mismos datos que el original. Cuando la carga amaine y las consultas recuperen el ritmo normal, el segundo clúster se detendrá automáticamente. De manera ideal, puede determinar esta capacidad de creación de simultaneidades y el coste especificando el número máximo de clústeres que pueden aprovisionarse automáticamente.

4. Cambie el tamaño de un clúster de procesamiento existente de manera manual

Si quiere mantener un estricto control sobre los costes y los recursos de procesamiento, el redimensionamiento manual ofrece una alternativa al ajuste de tamaño automático.

Imaginemos que tiene un clúster existente iniciado con cuatro nodos. Pero sabe que se aproxima un aumento en el número de datos y quiere proporcionar a dicho clúster más potencia de procesamiento, pero durante un período de tiempo específico. Puede cambiar el tamaño del clúster existente, especificando ocho, 16 o 32 nodos mientras aún está ejecutándose. Cuando se reduzca dicho aumento, podrá devolver al clúster su configuración original de cuatro nodos.

Estas son solo algunas de las posibilidades de las que dispone con un datawarehouse que realmente se beneficia de la arquitectura en la nube. Su mejor rendimiento exponencial, en comparación con los datawarehouses locales y «migrados a la nube», está preparado para satisfacer las demandas insaciables de los usuarios de volumen, velocidad y diversidad de datos, inimaginables hace 10 años.

Perfiles de usuario: Soluciones a las demandas habituales

Estrategias de almacenamiento para que salga del modo reactivo

USUARIO

SOLUCIÓN



DIRECTOR FINANCIERO

El director financiero que necesita una respuesta inmediata.

Inicie un nuevo clúster de procesamiento ad hoc rápidamente para que un analista financiero pueda ejecutar consultas en cuestión de minutos, sin que ello afecte a otros usuarios. Sus resultados proceden de la única fuente de datos de la empresa, por lo que está seguro de que son correctos.



VENTAS

El equipo de ventas que realiza un gran esfuerzo al final de cada trimestre.

Mantenga un clúster predefinido para los datos de ventas y de clientes en modo suspendido, pero listo para ponerse en marcha. Active este recurso predefinido en cualquier momento. Cuando el ajetreo del final del trimestre acabe, vuelva a establecerlo fácilmente en modo de suspensión.



MARKETING

El comerciante curioso que busca el recorrido del cliente más rentable.

Cuando los usuarios empiezan a añadir múltiples dimensiones a sus consultas y análisis, la complejidad aumenta rápidamente. Disponer de un clúster independiente para el comerciante garantizará que una consulta compleja ad hoc no afecte a otros usuarios o a otras cargas de trabajo.



CIENTÍFICO DE DATOS (DATASCI)

El científico de datos que de repente quiere someter a una prueba de resistencia su brillante teoría.

Si la cantidad de datos analizada salta de uno a tres años, puede ajustar el tamaño de un clúster existente de manera manual para proporcionar la potencia de procesamiento necesaria. Cuando el científico de datos esté satisfecho con su análisis, podrá devolver el clúster a su tamaño original para reducir costes hasta la realización de la siguiente prueba importante.



ASISTENCIA

El equipo de asistencia que quiere averiguar quién dispone de las puntuaciones de resolución más favorables.

Si un director observa normalmente un determinado conjunto de datos, pero un gran grupo de miembros del equipo aborda el mismo conjunto de datos con consultas adicionales, deberá actuar sin demora, de forma tan rápida que la intervención humana ni siquiera sea una posibilidad. Pero no tiene por qué preocuparse. El ajuste de tamaño de simultaneidad automático puede iniciar un segundo clúster que acceda a los mismos datos y, a continuación, equilibre la carga de las consultas del primer clúster. Cuando la oleada de consultas se reduzca, el segundo clúster se detendrá automáticamente.



ANALISTA

El analista de la cadena de suministro que profundiza en las tendencias de rotación del inventario.

Este analista suele trabajar con un período de seis meses de datos seguidos. De repente, ejecuta consultas en datos que se remontan a dos años atrás. Si recibe quejas de que las consultas son superlentas, puede iniciar un nuevo clúster con un tamaño adecuado en cuestión de minutos, o simplemente redimensionar el clúster existente para satisfacer el aumento de la demanda.



La gran pregunta: ¿y qué ocurre con la seguridad?

La nube puede ser más segura que las soluciones locales

Durante años, las organizaciones han considerado que la nube es incluso más vulnerable que las soluciones locales a pesar de que los hackers informáticos siguen pirateando importantes centros de datos corporativos.

No hay duda de que la seguridad sigue siendo la principal preocupación de las organizaciones que migran datos sensibles a la nube. Si evalúa y selecciona con cuidado las ofertas en la nube que consideran la seguridad una prioridad, puede beneficiarse de algunas de las mejores defensas de seguridad en la industria, mejor que las soluciones que la mayoría de las organizaciones tienen implementadas para proteger sus sistemas legacy (o heredados).

Obtener la seguridad en la nube adecuada resulta crítico. A continuación, le indicamos algunas de las mejores medidas que debe ofrecer un datawarehouse en la nube para disipar los temores de seguridad de las partes interesadas:

Cifrar datos en tránsito y en reposo

Si un usuario no autorizado logra acceder a sus datos, no debe poder leerlos. Punto y pelota. El datawarehouse en la nube moderno debe proteger los datos en tránsito y en reposo,

siempre que se transmitan a través de una red o se guarden en un disco. Esto incluye archivos de datos almacenados persistentemente, resultados de consultas y el contenido de una caché de disco local.

Independientemente de los recursos que dedique a la seguridad para su datawarehouse físico, un datawarehouse en la nube realmente seguro le permitirá redirigir dichos recursos a otras iniciativas de TI estratégicas.

Asimismo, una solución de datawarehouse en la nube moderna debe utilizar los últimos algoritmos de cifrado estándar de la industria. El Estándar de Cifrado Avanzado (AES), con claves de 128 bits, es la práctica recomendada mínima para el cifrado simétrico. Para una seguridad mejorada, los datawarehouses en la nube más robustos utilizan AES-256.

A diferencia de las arquitecturas de seguridad heredadas, el cifrado no debe afectar al rendimiento de la carga o de las consultas de un datawarehouse en la nube. Sin embargo, no es frecuente que los centros de datos empresariales se cifren con este alto grado debido al coste, el tiempo y la falta de experiencia.

SEGURO POR DISEÑO



AUTENTICACIÓN



CIFRADO DE DATOS



CONTROL DEL ACCESO



VALIDACIÓN EXTERNA



Pero, espere, que aún hay más: Factores de seguridad adicionales

Exploración de la gestión de claves de cifrado y de la autenticación de múltiples factores

Además del código de cifrado, existen muchos detalles adicionales que deben tenerse en cuenta. Uno de los más importantes es la gestión de claves.

Gestión de claves

La gestión de claves rige el ciclo de vida de las claves de cifrado, que incluye la generación, el almacenamiento, la distribución, el uso y la eliminación de las claves.

Idealmente, se utiliza una jerarquía de claves de modo que las claves principales cifren las claves secundarias, aproximadamente una por partición de datos. Estas, a su vez, cifran claves aún más específicas como, por ejemplo, una por tabla.

Para cualquier datawarehouse, debe limitar la cantidad de datos cubierta por una clave de cifrado individual y el tiempo que se debe utilizar la clave. Se trata de una práctica recomendada de la industria que se materializa mediante la rotación de claves y la regeneración de claves de datos:

- La rotación de claves es un método que permite generar periódicamente una nueva clave de cifrado para proteger datos recién insertados.

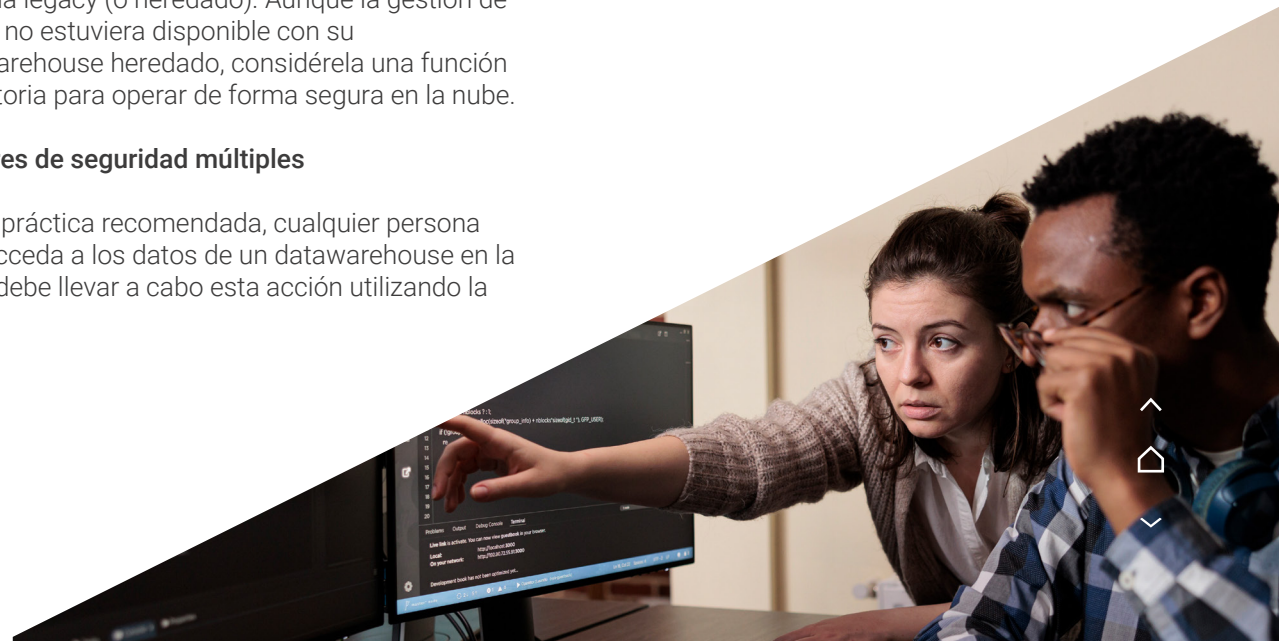
- La regeneración de claves es la capacidad de volver a los datos almacenados previamente, cifrarlos de nuevo con nuevas claves de cifrado recién generadas y deshacerse a continuación de las claves de cifrado anteriores.

Ambos mecanismos son necesarios para gestionar el ciclo de vida completo de las claves de cifrado de conformidad con los más altos estándares de la industria. Un datawarehouse en la nube que ofrezca esta función prácticamente exime a los clientes de la responsabilidad de implementar una gestión y una configuración de cifrado con un sistema legacy (o heredado). Aunque la gestión de claves no estuviera disponible con su datawarehouse heredado, considérela una función obligatoria para operar de forma segura en la nube.

Factores de seguridad múltiples

Como práctica recomendada, cualquier persona que acceda a los datos de un datawarehouse en la nube debe llevar a cabo esta acción utilizando la

autenticación multifactor (MFA). Tras iniciar sesión con un nombre de usuario y una contraseña, el usuario necesitará un segundo mecanismo de autenticación, que puede tratarse de un código aleatorio generado por una aplicación en el smartphone de un usuario. De forma conjunta, estos factores demuestran que los usuarios son quienes dicen ser, por lo que constituye una sólida medida para garantizar que únicamente las partes autorizadas logren acceder a los datos en la nube.



Seguridad y cumplimiento: Respeto de los estrictos requisitos de la industria

La importancia de la verificación de terceros

Los estándares específicos de la industria proporcionan un nivel de garantía adicional para las cuestiones relacionadas con la seguridad de los datos. Si ya está seguro de que su datawarehouse heredado cumple las normas de seguridad o se trata de un requisito al que antes no estaba acostumbrado, un proveedor de datawarehouses en la nube debe cumplir los siguientes estándares:

1 SOC 2:

El Instituto Estadounidense de Contables Públicos Certificados (AICPA) ha desarrollado el informe SOC 2. La finalidad de un informe SOC 2 es evaluar los sistemas de información de una organización en lo que respecta a la seguridad, la disponibilidad, la integridad del procesamiento, la confidencialidad o la privacidad.

2 HIPAA:

La información Sanitaria Protegida (PHI) está sujeta a las normas de seguridad y privacidad en virtud de la Ley de Portabilidad y Responsabilidad de los Seguros Médicos (HIPAA) de Estados Unidos. Los proveedores del servicio en la nube que almacenan PHI deben cumplir la normativa HIPAA para:

- **Seguridad y privacidad**
- **Herramientas de acceso a usuarios**

- **Cifrado**
- **Ubicación de los datos**
- **Devolución de los datos**
- **Planes de contingencia y recuperación ante desastres**
- **Acuerdos de nivel de servicio (SLA)**

En el caso de la mayoría de proveedores de infraestructuras en la nube, tales como Amazon Web Services (AWS), una gestión adecuada de datos PHI requiere controles de seguridad adicionales detallados en el acuerdo de socios comerciales (BAA) de AWS. Estos controles suelen ampliar protecciones críticas, como cifrar datos en reposo, para superar las especificaciones de la HIPAA, ya que los proveedores de servicios en la nube quieren reducir su riesgo. Las organizaciones de asistencia sanitaria que almacenan sus datos en almacenes modernos creados específicamente para la nube disfrutan de los beneficios de estas protecciones adicionales.

3 PCI:

El cumplimiento de la industria de tarjetas de pago (PCI) consiste en el respeto de un conjunto de estándares de seguridad específicos desarrollados para proteger información de tarjetas de crédito durante y después de una transacción financiera. Todas las marcas de tarjetas requieren el cumplimiento PCI. En el caso de información de tarjetas almacenada en un almacén en la nube, el

proveedor debe:

- **Crear y mantener una red segura.**
- **Proteger los datos del titular de la tarjeta.**
- **Mantener un programa de gestión de vulnerabilidades.**
- **Implementar estrictas medidas de control del acceso.**
- **Controlar y probar redes con regularidad;**
- **Mantener una política de seguridad de la información.**

La importancia de la verificación de terceros

Algunas normativas permiten que los proveedores «se autocertifiquen» como prueba de cumplimiento. Para garantizar que sus datos tengan la máxima seguridad, se espera que un proveedor de datawarehouses en la nube utilice los servicios de un proveedor independiente para llevar a cabo las pruebas de penetración. Una «prueba de penetración» constituye un intento de evaluar la seguridad de la infraestructura de TI procurando explorar de manera segura las vulnerabilidades que puedan existir en los sistemas operativos, fallos de aplicaciones y de servicios, configuraciones inadecuadas o comportamiento arriesgado del usuario final. Este tipo de validación de la seguridad ofrece una mayor garantía de que la seguridad de un proveedor de datawarehouses en la nube cumple los estándares de la industria y las expectativas de su organización.



Herramientas para llevar sus datos a la nube...

...sea cual sea su ubicación actual

Introducir los datos en un datawarehouse ha sido históricamente un proceso lento y tedioso, lo que significa que los analistas de negocios no siempre pueden disponer de los datos más actuales. Asimismo, su empresa, al igual que muchas otras organizaciones, puede estar migrando las principales aplicaciones a la nube. Por tanto, ahora también debe enfrentarse al acceso a dichos datos en la nube.

Pero tenemos buenas noticias. Actualmente, existe un gran número de nuevas herramientas que pueden ayudarle a migrar los datos de las aplicaciones existentes en las instalaciones a un datawarehouse en la nube. Entre las opciones populares, se incluyen herramientas de flujo de IU gráfica que le permiten extraer, transformar y cargar (ETL) datos de servicios en la nube y fuentes de bases de datos físicas tradicionales en su almacén en la nube moderno.

El gran paso: Archivos de carga masiva

Un enfoque de carga masiva puede funcionar mejor para su transferencia inicial si tiene muchos terabytes de datos ubicados en dispositivos de almacenamiento dentro de su empresa. Los datos se descargan en miles de archivos manejables y se cargan de forma masiva simultáneamente en el datawarehouse en la nube. Los recursos

complementarios fácilmente implementados para acelerar la carga deben ser parte integrante del datawarehouse en la nube moderno.

Actualizaciones graduales

Una vez que haya transferido sus datos, querrá registrar cambios graduales en el futuro. Se espera que utilice una de las muchas herramientas de ETL o ELT (por sus siglas en inglés, «extraer, cargar y transformar») modernas, disponibles en opciones de software comercial y código abierto. Entre las opciones adicionales, se incluyen las recientes herramientas de replicación de datos avanzada y captura de datos modificados (CDC). En función de la naturaleza y de la fuente de los datos, las nuevas plataformas y las herramientas de transmisión de datos basadas en la nube también pueden resultar útiles. Estas herramientas suelen ser muy adecuadas para realizar migraciones progresivas en lugar de cargas masivas.



Transferencia de datos físicos a gran escala

Cuando sus datos se convierten en algo realmente grande

Si tiene cientos de terabytes o incluso petabytes de datos, resulta lógico utilizar dispositivos físicos de proveedores en la nube para migrar sus datos. Estos dispositivos seguros y resistentes tienen aproximadamente el tamaño de un cajón de leche a la vieja usanza. Puede cargar grandes cantidades de datos cifrados de su centro de datos físico y devolver el dispositivo al proveedor de servicios en la nube, quien, a continuación, cargará los datos cifrados en un área de preparación en la nube.

Este método permite la transferencia de aproximadamente un petabyte de datos a la semana a la nube. A continuación, puede transmitir fácilmente los datos a su proveedor de datawarehouses en la nube.

Los dispositivos de transferencia de datos suelen ser más rápidos y rentables que intentar enviar todos estos datos a la nube a través de Internet. Se trata de una opción lógica si todos sus datos tardasen más de una semana en subirse a un datawarehouse en la nube.

La opción del Exabyte

Actualmente, las empresas generan tal cantidad de datos que suelen medirse en exabytes. (Un único exabyte es equivalente a mil millones de gigabytes). Algunos proveedores de servicios en la nube ofrecen ahora una opción de «incremento» para este nivel de transferencia de datos. En sentido literal, se trata de un dispositivo alojado en un contenedor de gran tamaño, similar a los contenedores de transporte, conectado a un semirremolque. Con diez de estos camiones, puede disponer de un exabyte de datos migrado a la nube en un plazo de seis meses.



Coste total de propiedad (TCO)

Presupuesto y gestión de costes

Planificación del paso al modelo empresarial en la nube

Sin lugar a dudas, la migración de su datawarehouse a la nube le permitirá ahorrar todos los gastos significativos de compra, mantenimiento y seguridad que acarrea un sistema local. Ahora bien, ¿cómo elabora el presupuesto de un modelo de costes totalmente diferente?

SAAS no es solo un cambio tecnológico

Pasar de una solución física a una en la nube también cambia la forma en la que una organización registra y presupuesta la compra. Una solución física suele conllevar un mayor gasto inicial de capital, considerado como un activo y amortizado a lo largo del tiempo. Una solución en la nube constituye una transacción inicial de menor envergadura y se considera un gasto operativo que se deduce cada mes de los ingresos correspondientes de una organización. Y, tras la compra inicial de cualquiera de estas soluciones, ambas requieren gastos adicionales, aunque enormemente diferentes, para que puedan seguir funcionando.

¿Por qué el personal de TI debería tener esto en cuenta? Los directores de TI deben demostrar que una nueva tecnología cumple las necesidades tecnológicas y el presupuesto de una organización. Si cambia el proceso típico para justificar dicha compra, el equipo de TI debe

presentar el nuevo modelo. ¿Y qué significa todo esto? En el centro de la compra de cualquier solución reside el coste total de propiedad (TCO), un cálculo financiero que detalla los costes de un sistema durante su tiempo de vida esperado.

TCO: Comparación entre una solución física y en la nueva

Una cosa es segura: con un cálculo del TCO, los gastos del primer año de una solución en la nube deben ser una fracción del coste de la compra física. Pero, al estimar una propiedad de cinco, 10 o más años, suele surgir un debate que lleva a la siguiente pregunta: «¿El coste de la solución en la nube acabará superando al de la alternativa física?».

Con soluciones de tamaño empresarial o de mayor envergadura, no resulta probable que el coste de la solución en la nube y la alternativa física lleguen a converger en algún momento. Los gastos iniciales y continuos de una solución física seguirán siendo mayores que el conjunto de cargos por el uso de una solución basada en la nube.

La realidad es que, tras los enormes costes iniciales de una compra física, una organización debe tener en cuenta el mantenimiento continuo, las reparaciones y las posibles sustituciones de su infraestructura de TI, incluidos la construcción y el mantenimiento de uno o de varios centros de

datos, lo que puede costar muchos millones de dólares.

Asimismo, existen gastos anuales de mantenimiento y asistencia para los sistemas locales que pueden superar el 20 % del precio de compra inicial del dispositivo o del software del datawarehouse. Incluso pueden derivarse costes relativos a la financiación de la compra. Por último, los sueldos del personal altamente especializado, encargado de prestar asistencia para un centro de datos y desempeñar todas las funciones relacionadas, pueden superar fácilmente el 50 % del TCO. Se trata de un componente del coste sustancial incluso al compararlo con la gran compra inicial de una solución física.

En pocas palabras, un ejercicio del TCO sencillo, pero exhaustivo, para un datawarehouse basado en la nube no solo revelará los avances tecnológicos, sino también el valor monetario que supone cambiar al modelo en la nube.



Retorno en la inversión (ROI): El TCO y muchísimo más

Evaluación del ROI para su datawarehouse en la nube

El TCO y el retorno de la inversión (ROI) son dos caras de la misma moneda que están inextricablemente unidas. Un cálculo del TCO revela cuánto costará un datawarehouse en la nube y cuánto podrá ahorrar una organización en comparación con una alternativa física. Pero estos ahorros representan únicamente un aspecto del ROI. Existen muchos otros factores que deben tenerse en cuenta al pensar en las diferencias de ROI entre la solución física y en la nube. Su cálculo del ROI debe incluir además:

1 Tiempos de implementación

Un datawarehouse en la nube puede ponerse en marcha en semanas o en tan solo unos meses en función del tamaño del proyecto y de la estrategia de migración a la nube. Una organización puede ver los beneficios mucho antes con una solución en la nube y con una menor inversión inicial. Pero tenga presente que los cálculos de una rápida implementación solo hacen referencia a un verdadero datawarehouse SaaS (software como servicio). Las soluciones basadas en los enfoques IaaS (infraestructura como servicio) o PaaS (plataforma como servicio) pueden tardar mucho más tiempo. El tiempo que ahorre se traduce directamente en un ahorro de dinero.

2 Actualización de Software

Las soluciones físicas y locales «migradas a la nube» adoptan un enfoque de desarrollo «en cascada» estándar con respecto a las actualizaciones de las funciones. Para permitir la actualización anual o bianual, el equipo de TI normalmente debe desactivar el sistema o establecerlo en modo de mantenimiento, lo que supone una mayor pérdida de tiempo y dinero. Para evitarlo, el equipo de TI puede quedarse anclado en una versión específica del software, lo que crea otra serie de quebraderos de cabeza completamente diferentes.

Con datawarehouse en la nube moderno, las actualizaciones deben originarse a partir de un enfoque de desarrollo y operaciones (DevOps) flexible (con actualizaciones graduales cada mes que eviten cualquier interrupción a los clientes).

3 El factor Humano

Tal como se ha descrito anteriormente, el número de personas que mantienen un datawarehouse empresarial físico y respaldan la infraestructura puede constituir un gasto enorme. En función de la alternativa en la nube, una organización puede reducir

considerablemente o eliminar prácticamente este gasto dependiendo del nivel de funcionalidad, automatización y gestión de proveedores de la solución.

4 Pague únicamente por lo que se use

Un datawarehouse físico le obliga a adquirir espacio de almacenamiento y potencia de procesamiento suficientes para hacer frente a la demanda el día más ajetreado del año. Pero ¿qué ocurre con los otros 364 días? Con el datawarehouse en la nube adecuado, dispone de una magnífica oportunidad de pagar únicamente por lo que usa, cuando lo usa.

Asimismo, el coste de los recursos de almacenamiento y procesamiento reales debe ser considerablemente inferior con una solución en la nube gracias a las economías de escala de esta tecnología.

Estos factores y muchos otros deben formar parte de sus cálculos del ROI al comparar su opción física con un datawarehouse en la nube. Si tiene pensado utilizar más de un proveedor de datawarehouses en la nube para su próxima compra, no olvide evaluar las diferencias entre estas alternativas, ya que no hay dos datawarehouses iguales, ya sean físicos o en la nube.



Su recorrido de cinco pasos para convertirse en un experto que trabaja con datos

«Chuleta» de aspectos clave de un datawarehouse en la nube





PowerData, es una compañía multinacional de origen español con destacada presencia regional. Desde que se fundó en el año 2000 ha ayudado a más de 100 organizaciones de diversos sectores, categorías y nichos, quienes han depositado su confianza en el expertise y Know-how de PowerData para acelerar su camino hacia la transformación digital.

El alcance de la especialización en gestión de datos en la nube ha sido clave para que una empresa con ambiciones globales como Snowflake haya confiado en PowerData como partner para entregar lo mejor de su tecnología Data Cloud a las organizaciones.

Te invitamos a explorar en todos los proyectos donde aportamos valor con la gestión de datos en powerdata.es

