

Desarrollo de una estrategia de gestión del ciclo de vida e integración de datos para un entorno híbrido



Tabla de contenido

Introducción: Imponer orden en el caos	3
Los cuatro pilares	4
Convertir los datos en información: Definición de una estrategia del ciclo de vida e integración	7
¿Dónde están los datos que mi organización necesita?	8
¿Cómo se agruparán los datos para crear la información que mi organización necesita?	10
¿Realmente necesito guardar todos estos datos?	18
Soluciones de IBM para la gestión del ciclo de vida e integración de datos	21
Pasos a seguir: Continuación del debate sobre el gobierno de cloud	23

Introducción: Imponer orden en el caos

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud



Los datos en cloud presentan una gran cantidad de información potencial para aquellas organizaciones que buscan crear y mantener una ventaja competitiva en su sector. Sin embargo, para ello, y tal y como se expone en [“The truth about information governance and the cloud,”](#) muchas organizaciones tendrán que reconciliar sus datos en local con los nuevos datos en cloud de terceros. En estos entornos “híbridos” es donde reside la información para tomar decisiones importantes.

Los entornos híbridos tienden a crecer sin mucha planificación, lo que dificulta la gestión de los crecientes almacenes de datos, pero existe un modo de darle sentido a este caos. Como siempre, el primer paso consiste en comprender la naturaleza del problema. El foco de atención se centra en los propios datos, en lugar de los orígenes de los datos o los sistemas utilizados para gestionarlos. Si prioriza los datos y la propiedad de la información extraída de ellos, todo lo demás se ordena rápidamente.

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Los cuatro pilares

¿Cómo puede su organización materializar los beneficios del cloud garantizando a su vez que la información recopilada del cloud sea segura y fiable? La respuesta yace en el gobierno.

Un buen gobierno de la información híbrida se basa en cuatro prioridades clave para la TI y la empresa:

- 1. Amplio consenso sobre qué significa la información,** incluyendo los metadatos sobre políticas comunes y reglas en lenguaje simple para la información que la empresa necesita y cómo se manejará.
- 2. Acuerdo claro sobre cómo se mantendrán y supervisarán los activos de información**—por ejemplo, las reglas de calidad de los datos operativos para la gestión de datos en sistemas locales.

- 3. Prácticas estándar a nivel de departamento y empresa para proteger los activos de información estratégicos,** como articular el acceso a información basado en roles, crear reglas para determinar cómo se compartirá la información y proteger la información sensible ante terceros.

- 4. Una estrategia de integración de datos empresariales,** que incluya la gestión del ciclo de vida y el diseño de una arquitectura para definir cómo circularán los datos y cómo se integrarán en la información estratégica, así como comprender cómo se mantendrá dicha información con el paso del tiempo.

Introducción: Imponer orden en el caos

Los cuatro pilares

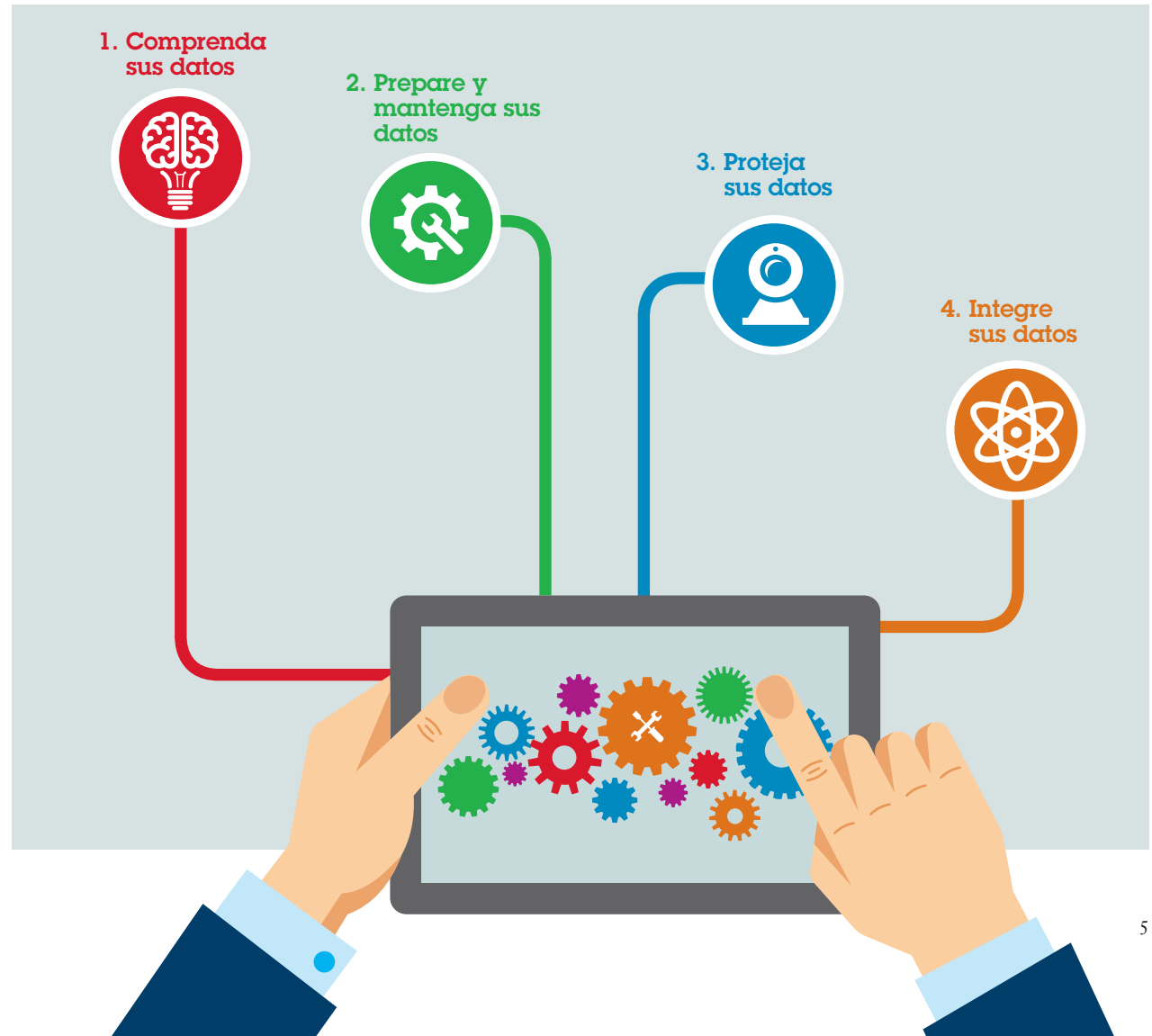
Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Pase el ratón por los siguientes iconos para obtener más información sobre las prioridades para un buen gobierno de la información híbrida.



Introducción: Imponer orden en el caosLos cuatro pilares

Convertir los datos en información:**Definición de una estrategia del ciclo de vida e integración**

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Estos componentes constituyen las bases de un gobierno de información en un entorno híbrido. En cada caso, necesitará combinar distintos habilitadores técnicos, organizativos y de procesos para alcanzar el éxito. Estos pilares dotarán a su organización de la flexibilidad necesaria para moverse con velocidad y seguridad.

Este libro electrónico se centra en el cuarto pilar: crear y ejecutar una estrategia de gestión del ciclo de vida e integración de datos.

Cómo apoderarse de la información estratégica

La adopción de un entorno híbrido no implica necesariamente que su estrategia de TI tenga que ser ya completamente funcional. De hecho, los aspectos del entorno relacionados con el cloud cambiarán rápidamente en función de las prioridades de negocio. No obstante, incluso si solo un pequeño porcentaje de los datos está circulando desde fuentes en cloud, la TI debe definir un plan para la integración y la seguridad de los datos, así como garantizar que la organización sea propietaria de toda la información creada a partir de los datos y el procesamiento, independientemente de su ubicación. La infraestructura híbrida y la informática descentralizada son medios definitivos para la creación de activos de información estratégicos.

Introducción: Imponer orden en el caos
Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

- ¿Dónde están los datos que mi organización necesita?
 - ¿Cómo se agruparán los datos para crear la información que mi organización necesita?
 - ¿Realmente necesito guardar todos estos datos?
-

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Convertir los datos en información: Definición de una estrategia del ciclo de vida e integración

Ya ha reconocido la importancia de ser propietario de los datos estratégicos para su organización, lo cual a su vez genera la necesidad de contar con una estrategia de gestión del ciclo de vida e integración de datos claramente articulada. Esta estrategia definirá qué orígenes de datos son aceptables, de qué elementos de datos ser propietario, cómo se almacenarán o mantendrán los datos en propiedad y cuánto tiempo sobrevivirán los datos en la organización.

Para tomar estas decisiones, puede empezar respondiendo a estas tres preguntas:

1. **¿Dónde están los datos que mi organización necesita?**
2. **¿Cómo se agruparán los datos para crear la información que mi organización necesita?**
3. **¿Realmente necesito guardar todos estos datos?**

Abordemos cada una de estas tres preguntas de manera individual.



Introducción: Imponer orden en el caos
Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

¿Dónde están los datos que mi organización necesita?

Los líderes de negocio persiguen el potencial de los datos, independientemente de dónde residan. Sin embargo, antes de adentrarse en el extenso mundo de los datos, tenga en cuenta que cuanto más crezcan los volúmenes de datos, más costará garantizar que la información sea fiable y precisa. Si no se aborda esta cuestión, los usuarios finales podrían perder la confianza en la información generada a partir de sus datos, provocando la pérdida de oportunidades o generando vulnerabilidad ante amenazas.

Es evidente que la gran cantidad y complejidad de datos y fuentes de datos inhabilita los métodos manuales tradicionales para descubrir, gobernar y corregir información. Deberá



implementar una solución de integración y gobierno de la información (IGI) para dar soporte a las distintas aplicaciones de datos, almacenes e iniciativas para aumentar el almacén de datos—independientemente de si la información se almacena en cloud o en local.

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:**Definición de una estrategia del ciclo de vida e integración**

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Una buena estrategia de integración de datos contará con muchas topologías de datos, integrando la información en el punto de creación a medida que se revelen los orígenes de datos importantes. **Un programa de IIG sólido debe incluir descubrimiento automatizado, creación de perfiles y gestión de metadatos para proporcionar contexto a los diversos conjuntos de datos y mejorar las decisiones de los empleados.**

La implicación es clara: no basta con poner datos en local y prepararlos para su análisis, sino que en un entorno híbrido se necesitan distintas tecnologías y prácticas de integración

en cloud y en local (como Plataforma como Servicio). Tendrá que integrar y gobernar los datos en local con tecnologías en local e integrar y gobernar los datos en cloud con tecnologías en cloud.

Además, tanto si los datos están en local como en cloud, las tecnologías de integración de datos deben ser lo suficientemente ágiles para acomodar una amplia variedad de formatos de datos y para integrarse fácilmente con distintas tecnologías de almacenamiento de datos, desde data marts hasta sistemas Apache Hadoop.

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:**Definición de una estrategia del ciclo de vida e integración**

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

Si las tecnologías de integración de datos se destinan a las fuentes de datos, ¿cómo reunir los distintos elementos para crear los activos de información que necesita la organización? Los datos se transfieren desde una variedad de orígenes, bajo demanda y a gran velocidad, por lo que el rendimiento es clave.

Los datos en constante cambio deben volcarse rápidamente a las distintas aplicaciones del sistema para que los líderes de negocio puedan adaptarse a las cambiantes condiciones del mercado lo más pronto posible. Para manejar estos datos, se necesita una solución de integración de datos empresarial que sea:

- **Dinámica** para cumplir sus requisitos de rendimiento actuales y futuros



- **Ampliable** y particionada para ofrecer una escalabilidad rápida y sencilla
- **Integrada** con Hadoop u otras tecnologías de almacenamiento de big data

Introducción: Imponer orden en el caos
Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Una base muy escalable

El requisito más importante para procesar grandes volúmenes de datos híbridos para la integración de datos es la escalabilidad masiva de datos (MDS). MDS permite procesar grandes cantidades de datos en paralelo, reduciendo drásticamente la cantidad de tiempo dedicado a manejar varias cargas de trabajo. A diferencia de otros modelos de procesamiento, los sistemas MDS optimizan los recursos de hardware disponibles, permitiendo procesar la máxima cantidad de datos por nodo (ver Figura 1).

MDS es importante porque el procesamiento de grandes volúmenes de datos híbridos permite resolver por primera vez muchos problemas de negocio de alto valor, garantizando que la plataforma de hardware genere beneficios previsibles.

Los sistemas MDS comparten 4 aspectos:

1. Presentan una arquitectura *shared-nothing*
2. Se implementan utilizando flujo de datos de software

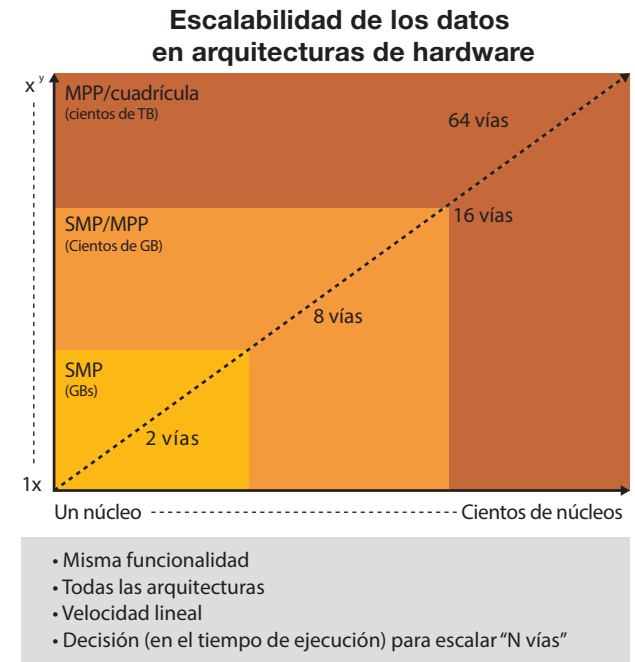


Figura 1. Características básicas para dar soporte a los requisitos de MDS con el objetivo de habilitar el procesamiento de volúmenes de datos ilimitados.

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

3. Utilizan el particionamiento de datos para la escalabilidad lineal de datos
4. Aplican un entorno de aislamiento de diseño

El software con arquitectura **shared-nothing** está diseñado para explotar una arquitectura de procesamiento paralelo masivo (MPP) de nada compartido particionando conjuntos de datos en los distintos nodos informáticos y ejecutando una única aplicación con la misma lógica de aplicaciones en cada partición de datos (ver Figura 2). De esta manera, no existe ningún punto de contención, ni cuello de botella de procesamiento, en el sistema y, por tanto, no hay limitación máxima sobre el volumen de datos, el rendimiento del procesamiento ni el número de procesadores y nodos.

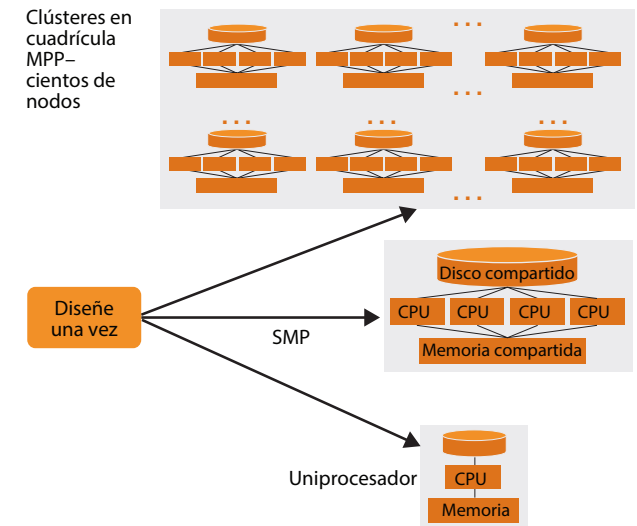


Figura 2. Ejemplo de una arquitectura *shared-nothing*.

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

El flujo de datos de software permite sacar el máximo partido de la arquitectura *shared-nothing*, facilitando la implementación y la ejecución de particionamiento y canalización de datos en un nodo y entre nodos. El flujo de datos de software también oculta la complejidad que implica crear y ajustar aplicaciones paralelas para los usuarios.

El flujo de datos de software es la mejor arquitectura para aplicar procesadores de múltiples núcleos en un servidor de multiprocesamiento simétrico (SMP), que sería la escalada vertical de aplicaciones, y para escalar a múltiples máquinas (escalada

horizontal de aplicaciones)—que será la arquitectura de sistemas más utilizada en entornos de datos híbridos en local y cloud. La arquitectura:

- Admite el paralelismo particionado y canalizado en y entre nodos SMP
- Proporciona un único mecanismo para la paralelización en todas las arquitecturas de hardware, eliminando complejidad
- Reduce la dificultad de crear, ajustar y ejecutar aplicaciones paralelas
- No tiene límite máximo sobre el volumen de datos, el rendimiento de procesamiento ni el número de nodos de procesamiento

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

El particionamiento de datos significa que los conjuntos de datos se particionan en distintos nodos y un único trabajo ejecuta la misma lógica de aplicación en todos los datos particionados (ver Figura 3). Otros métodos, como el particionamiento de tareas, no proporcionan una escalabilidad lineal de los datos a medida que crecen los volúmenes de datos porque la cantidad de datos que se puede ordenar, fusionar y agregar está limitada a lo que puede procesarse en un nodo. Esto también significa que se pueden asignar nodos exclusivos a datos de distintos orígenes, en caso necesario.

Las características de sistemas con particionamiento de datos incluyen:

- Distribuye particiones de datos entre nodos
- Ejecuta un trabajo en paralelo entre nodos
- Habilita la canalización y el reparticionamiento entre etapas y nodos sin acabar en disco
- Utiliza hardware de cuadrícula de bajo coste para big data

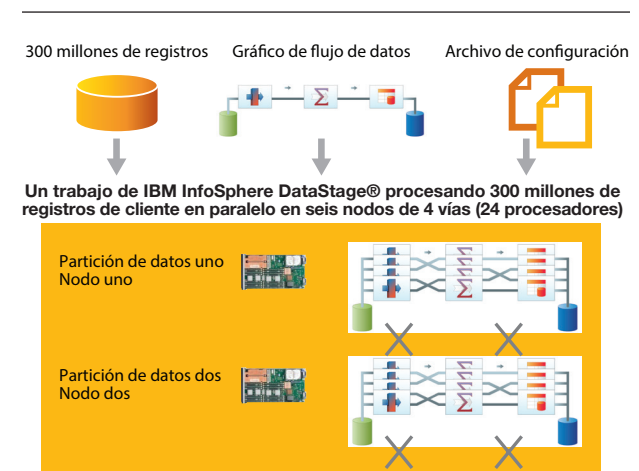


Figura 3. Arquitectura de particionamiento de datos.

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Finalmente, el aislamiento del diseño permite a un desarrollador diseñar un procesamiento de datos una vez y utilizarlo en cualquier configuración de hardware sin tener que rediseñar y reajustar el trabajo. Las características y ventajas incluyen:

- Crear una vez y ejecutar en cualquier lugar sin modificaciones
- Reducir la complejidad con un mecanismo unificado para la paralelización
- Lograr una separación clara entre el desarrollo de un trabajo y la expresión de paralelización en el tiempo de ejecución
- Eliminar la necesidad de ajustar el rendimiento cada vez que cambie la arquitectura de hardware
- Añadir hardware sin límite máximo de escalabilidad de datos

Todas las plataformas más escalables (IBM® Netezza®, IBM PureData™ System, IBM DB2® Database Partitioning Feature, Teradata, Hadoop e IBM InfoSphere® Information Server) se han creado desde cero para integrar estas cuatro características y aplicar arquitecturas en cuadrícula de comodidad y MPP.

Utilice la réplica de datos para acelerar la información de valor

Si quiere maximizar la cantidad de información extraída de orígenes de datos híbridos, deberá emplear distintos estilos de distribución de datos en función del caso de uso.

Determinados casos de uso requieren una vista de los datos actualizada cada minuto (o segundo) para tomar decisiones fiables. Por ejemplo, en el caso de la detección de fraude, el análisis de inventario entre canales o la analítica operativa en tiempo real, basar las decisiones en datos de un mes, una semana o incluso un día de antigüedad es contraproducente.

Recopile y aproveche los datos más rápido que la competencia y ofrezca una mayor continuidad de servicios para generar experiencias del cliente excepcionales. Como consecuencia de ello, se han ampliado los requisitos de transformación y distribución de datos para incluir la transferencia de datos en tiempo real en base a funcionalidades de réplica de datos, especialmente en torno a la captura de datos con cambios.

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Existen tres factores a tener en cuenta a la hora de maximizar el rendimiento y la escalabilidad de la integración de datos en tiempo real:

1. El método aplicado para capturar un cambio en el origen u orígenes: la opción más flexible y eficiente para capturar cambios en el origen es un proceso de réplica que captura los cambios a medida que se escriben en el registro de un origen. Cuando se modifican los datos de un origen, el mecanismo detecta la alteración y reenvía los datos modificados, sin apenas impacto sobre la base de datos de origen ni la aplicación, minimizando la necesidad de contar con grandes ventanas de lotes.

2. El mecanismo utilizado: se pueden utilizar muchos mecanismos para la réplica de datos. Si se implementa correctamente, un método de captura basado en el registro tiene un impacto inferior sobre la base de datos de origen, lo que implica un rendimiento superior.

3. Persistencia de datos temporal: si existe persistencia de datos temporal, también afecta al rendimiento de la réplica de datos. Lo mejor es que una organización sea capaz de distribuir los cambios sin persistencia para incrementar el rendimiento porque, de esta manera, los datos no tienen que grabarse en el disco para que después acceda a ellos un motor de transformación.

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:**Definición de una estrategia del ciclo de vida e integración**

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos**Pasos a seguir: Continuación del debate sobre el gobierno de cloud**

El rendimiento no es el único factor importante en las soluciones de integración de datos.

Algunas consideraciones adicionales incluyen:

- **Flexibilidad de la solución:** los mecanismos flexibles dan soporte a múltiples plataformas, orígenes y destinos, incluyendo archivos Hadoop Distributed File System (HDFS) para grandes volúmenes de datos, además de una gran variedad de topologías.
- **Impacto sobre los procesos y la infraestructura de TI existentes:** la solución de réplica de datos debe integrarse fácilmente en sus procesos de gestión de cambios, permitiendo una automatización sencilla a través de scripts o lenguajes de programación comunes como Java.



- **Facilidad de uso:** las curvas de aprendizaje muy grandes afectan a la integración de datos. Potentes interfaces gráficas facilitan la configuración y supervisión, agilizando la extracción de información de valor real de big data.

Introducción: Imponer orden en el caos
Los cuatro pilares

Convertir los datos en información:**Definición de una estrategia del ciclo de vida e integración**

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

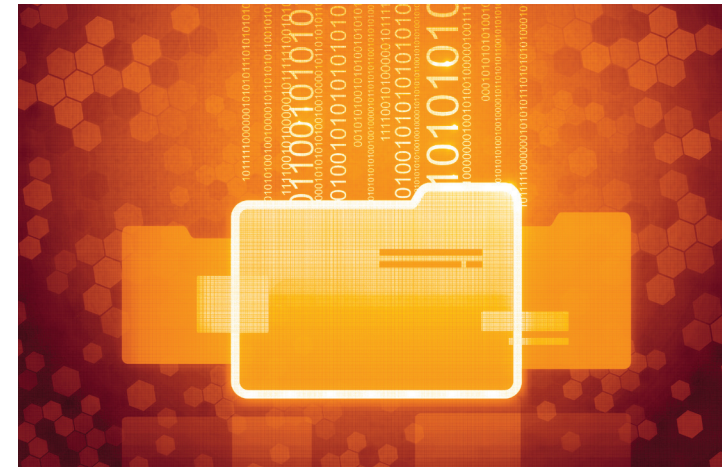
Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

¿Realmente necesito guardar todos estos datos?

Muchos usuarios de analítica son partidarios de guardar todos los datos, porque nunca se sabe qué información se podrá extraer en un futuro de los datos que ahora no son útiles. Sin embargo, este enfoque puede implicar riesgos y costes elevados para la empresa.

Bajo la premisa “vamos a guardarlos por si alguien los necesita después”, muchas organizaciones conservan demasiados datos históricos, muchos de los cuales carecen de valor. Un exceso de almacenamiento y retención solo agrava la situación. Al mismo tiempo, las organizaciones tienen que garantizar la privacidad y la seguridad de los crecientes volúmenes de información confidencial. Las normativas gubernamentales y sectoriales exigen a las organizaciones proteger la información personal, independientemente de donde resida, incluso en entornos de pruebas y desarrollo.



Además, en grandes volúmenes de datos híbridos se ralentizan los tiempos de respuesta y se degrada el rendimiento de las aplicaciones cuando los usuarios realizan consultas. Si no se aborda este problema, el crecimiento continuo de los datos estirará los recursos sobrepasando su capacidad y afectará al tiempo de respuesta de procesos críticos de informes y consultas.

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:**Definición de una estrategia del ciclo de vida e integración**

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos**Pasos a seguir: Continuación del debate sobre el gobierno de cloud**

Estos problemas pueden afectar a los entornos de producción y obstaculizar las tareas de actualización, migración y recuperación tras desastre. Por ello, es vital implementar una gestión de datos inteligente para datos históricos inactivos.

El crecimiento de los datos puede disparar los costes operativos y de infraestructura, que a menudo consumen la mayor parte del presupuesto destinado a la gestión o el almacenamiento de datos. Incrementar los volúmenes de datos implica más capacidad, por lo que probablemente deberá comprar hardware adicional y gastar más en mantener, supervisar y administrar la infraestructura. Los grandes almacenes de datos y los entornos de datos complejos requieren servidores, dispositivos y entornos de prueba de mayor tamaño, lo que incrementa también los costes de licencia de software para la base de datos y sus herramientas, sin mencionar la mano de obra, la energía y los costes legales.

**La función del archivado**

El ciclo de vida de los datos integra múltiples fases: los datos se crean, utilizan, comparten, actualizan, almacenan y finalmente se archivan o se eliminan. La gestión del ciclo de vida de los datos juega un papel importante en la fase de archivado.

Las políticas de retención, como parte de una estrategia de archivado, se diseñan para

Introducción: Imponer orden en el caos

Los cuatro pilares

Convertir los datos en información:

Definición de una estrategia del ciclo de vida e integración

¿Dónde están los datos que mi organización necesita?

¿Cómo se agruparán los datos para crear la información que mi organización necesita?

¿Realmente necesito guardar todos estos datos?

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

guardar elementos de datos importantes como referencia y para uso futuro, además de eliminarse datos que ya no se necesitan para dar soporte a los requisitos legales o empresariales de una organización. La gestión efectiva del ciclo de vida de los datos incluye la inteligencia para archivar datos en contexto, lo que puede incluir información de docenas de orígenes de datos, pero también para archivarla en base a parámetros específicos o reglas de negocio, como la edad. La gestión del ciclo de vida de los datos también ayuda a los administradores de almacenamiento a desarrollar una estrategia de almacenamiento por niveles automatizada que les permita archivar los datos inactivos en un almacén de datos, mejorando así el rendimiento general de las aplicaciones analíticas.

Las ventajas de una gestión efectiva del ciclo de vida de los datos para el negocio y la TI incluyen:

- **Incremento del margen:** Costes de capital e infraestructura inferiores, mayor productividad y menos defectos en las aplicaciones durante el ciclo de desarrollo
- **Reducción de riesgos:** Menos periodos de inactividad de las aplicaciones, menos interrupciones del servicio y rendimiento y conformidad con los requisitos de retención de datos
- **Mayor agilidad del negocio:** Comercialización más rápida y mayor rendimiento y calidad de las aplicaciones

Introducción: Imponer orden en el caos

Convertir los datos en información:
Definición de una estrategia del ciclo de vida e integración

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Soluciones de IBM para la gestión del ciclo de vida e integración de datos

IBM lleva décadas proporcionando soluciones para gestionar los datos empresariales procedentes de múltiples orígenes. La compañía es líder en soluciones de integración de datos, gestión, seguridad y analítica, que destacan por su fiabilidad, flexibilidad y escalabilidad.

[IBM InfoSphere Information Server](#) es una plataforma de integración de información líder del mercado que ayuda a las organizaciones a comprender y gobernar los datos, crear y mantener la calidad de los datos y transformar y distribuir datos. Junto con [InfoSphere Data Replication](#) e [InfoSphere Federation Server](#), InfoSphere Information Server ofrece potentes funcionalidades de IIG:

- **Integración de datos:** transforma los datos en cualquier estilo y los distribuye a cualquier sistema, garantizando una rápida generación de valor y un menor riesgo para la TI. Este paquete incluye la característica InfoSphere Data Clic, que admite la integración de datos de autoservicio.



- **Réplica de datos:** da soporte a los requisitos de réplica de datos en tiempo real y enriquece las aplicaciones móviles, la analítica de negocio y los proyectos de big data integrando datos replicados.
- **Catálogo de gobierno de la información:** ayuda a comprender y gobernar la información a través de un método estandarizado para descubrir activos de TI y definir un lenguaje empresarial común que le permita alinear mejor los objetivos de negocio y TI.

Introducción: Imponer orden en el caos

**Convertir los datos en información:
Definición de una estrategia del ciclo de vida e integración**

**Soluciones de IBM para la gestión del ciclo de vida e
integración de datos**

**Pasos a seguir: Continuación del debate sobre el gobierno
de cloud**

- **Calidad de los datos:** le permite analizar, limpiar, supervisar y gestionar datos, añadiendo un importante valor que ayudará a mejorar la toma de decisiones y la ejecución de los procesos de negocio.
- **Federación de datos:** distribuye datos de forma flexible y da soporte a los requisitos del centro de datos virtual.

Las funcionalidades de InfoSphere IIG también son compatibles con la plataforma de big data de IBM, que incluye herramientas para virtualización y descubrimiento, analítica basada en Hadoop, stream computing, almacenamiento de datos y análisis de texto.

Las soluciones [IBM InfoSphere Optim™](#) complementan las funcionalidades de IIG con funciones de gestión del ciclo de vida de los datos, que se escalan para cubrir

las necesidades empresariales. Tanto si implementa las soluciones InfoSphere Optim para una única aplicación o almacén de datos como para un entorno de datos complejo, obtendrá una estrategia coherente para agilizar la gestión del ciclo de vida de los datos. El exclusivo motor de relaciones de InfoSphere Optim proporciona un único punto de control para guiar las actividades de procesamiento de datos, tales como archivar, crear subconjuntos y recuperar datos.

Las soluciones InfoSphere Optim le ayudan a cumplir con los requisitos gubernamentales y de IIG, así como a abordar los retos asociados a los entornos de datos híbridos de múltiples orígenes. Al archivar datos antiguos de grandes entornos de almacenamiento de datos, recuperará una valiosa capacidad de almacenamiento que le permitirá mejorar los tiempos de respuesta y reducir costes.

Introducción: Imponer orden en el caos

Convertir los datos en información:
Definición de una estrategia del ciclo de vida e integración

Soluciones de IBM para la gestión del ciclo de vida
integración de datos

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Pasos a seguir: Continuación del debate sobre el gobierno de cloud

Los datos en cloud y los servicios de procesamiento presentan una gran oportunidad que los usuarios de negocio no pueden pasar por alto. La TI, por su parte, se encarga de mantener la integridad de los sistemas de creación de informes y transaccionales en local e internos. Por ello, definir una estrategia de gobierno para un entorno híbrido se convierte en una prioridad.

Este libro electrónico expone uno de los cuatro pilares para gobernar con éxito un entorno híbrido: la gestión del ciclo de vida y la integración de los datos en un entorno híbrido. **Para explorar los demás pilares, descargue los distintos libros electrónicos de la serie:**

- [La verdad sobre el gobierno de la información y el cloud](#)
- [Cómo dar sentido a los datos](#)

- [Cómo preparar y mantener los datos](#)
- [Protección de datos en cloud y sobre el terreno](#)

Para obtener más información sobre el liderazgo de pensamiento de gobierno de IBM y las tecnologías de soporte, visite: ibm.com/software/data/information-integration-governance

Además, IBM Global Financing le ayuda a adquirir las funcionalidades de software que su negocio necesita de la forma más estratégica y rentable. Colaboramos con clientes con crédito positivo para personalizar una solución de financiación que se adapte a sus objetivos de desarrollo y negocio, facilitar la gestión eficiente del efectivo, así como mejorar el coste total de propiedad. Financie su inversión en TI e impulse su negocio con IBM Global Financing. Para obtener más información, visite: ibm.com/financing



© Copyright IBM Corporation 2015

Software Group
Route 100
Somers, NY 10589

Producido en los Estados Unidos de América
Abril de 2015

IBM, el logotipo de IBM, ibm.com, DataStage, DB2, InfoSphere, Optim, y PureData son marcas de Internacional Business Machines Corp., registradas en numerosas jurisdicciones de todo el mundo. Otros nombres de productos y servicios pueden ser marcas registradas de IBM u otras empresas. Existe una lista actualizada de marcas registradas de IBM en la Web, en el apartado “Copyright and trademark information” de ibm.com/legal/copytrade.shtml

Netezza es una marca registrada de IBM International Group B.V., una Compañía IBM.

Java y todas las marcas registradas y logotipos de Java son marcas registradas de Oracle y/o sus filiales.

Este documento se considera actualizado en la fecha inicial de su publicación y puede ser modificado por IBM en cualquier momento. No todas las ofertas están disponibles en todos los países en los que opera IBM.

LA INFORMACIÓN PROPORCIONADA EN ESTE DOCUMENTO SE DISTRIBUYE “TAL CUAL”, SIN GARANTÍA ALGUNA, YA SEA EXPRESA O IMPLÍCITA, INCLUYENDO TODA GARANTÍA DE COMERCIALIZACIÓN, IDONEIDAD PARA UN FIN CONCRETO O INFRACCIÓN DE DERECHOS DE TERCEROS. Los productos IBM están garantizados de acuerdo con los términos y condiciones de los contratos con arreglo a los cuales son facilitados.

El cliente es responsable de asegurar que cumple con la legislación y la normativa aplicable. IBM no proporciona asesoramiento legal ni manifiesta o garantiza que sus servicios o productos aseguren el cumplimiento por parte del cliente con ninguna legislación o normativa.



Recicle este documento
