

Data Warehouse Service

Descripción general del servicio

Edición 01
Fecha 2024-01-05



Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2024. Todos los derechos reservados.

Quedan terminantemente prohibidas la reproducción y/o la divulgación totales y/o parciales del presente documento de cualquier forma y/o por cualquier medio sin la previa autorización por escrito de Huawei Cloud Computing Technologies Co., Ltd.

Marcas registradas y permisos



El logotipo  y otras marcas registradas de Huawei pertenecen a Huawei Technologies Co., Ltd. Todas las demás marcas registradas y los otros nombres comerciales mencionados en este documento son propiedad de sus respectivos titulares.

Aviso

Es posible que la totalidad o parte de los productos, las funcionalidades y/o los servicios que figuran en el presente documento no se encuentren dentro del alcance de un contrato vigente entre Huawei Cloud y el cliente. Las funcionalidades, los productos y los servicios adquiridos se limitan a los estipulados en el respectivo contrato. A menos que un contrato especifique lo contrario, ninguna de las afirmaciones, informaciones ni recomendaciones contenidas en el presente documento constituye garantía alguna, ni expresa ni implícita.

Huawei está permanentemente preocupada por la calidad de los contenidos de este documento; sin embargo, ninguna declaración, información ni recomendación aquí contenida constituye garantía alguna, ni expresa ni implícita. La información contenida en este documento se encuentra sujeta a cambios sin previo aviso.

Índice

1 ¿Qué es GaussDB(DWS)?	1
2 Tipos de almacén de datos	6
3 Ventajas	10
4 Escenarios de aplicación	12
5 Funciones	17
6 Conceptos	23
6.1 Conceptos de gestión de GaussDB(DWS).....	23
6.2 Conceptos de base de datos de GaussDB(DWS).....	23
7 Servicios relacionados	25
8 GaussDB(DWS) Gestión de permisos	27
9 Acceso de GaussDB(DWS)	33
10 Detalles de precios	35
11 Restricciones	39
12 Cuotas de servicio	40
13 Especificaciones técnicas	41
14 Descripción de la versión	43

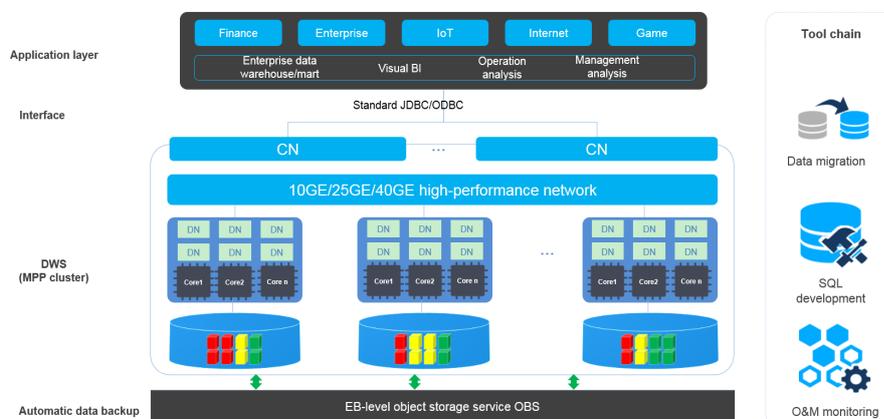
1 ¿Qué es GaussDB(DWS)?

GaussDB(DWS) es una base de datos de procesamiento de datos en línea que se ejecuta en la Huawei Cloud infraestructura para proporcionar un servicio de base de datos analítico escalable, totalmente administrado y listo para usar, liberándolo de la gestión y supervisión complejas de bases de datos. Es un servicio en la nube nativo basado en el almacén de datos convergentes GaussDB de Huawei, y es totalmente compatible con el estándar de ANSI SQL 99 y SQL 2003, así como con los ecosistemas PostgreSQL y Oracle. GaussDB(DWS) proporciona soluciones competitivas para el análisis de big data a nivel de PB en varias industrias.

Arquitectura

GaussDB(DWS) emplea la arquitectura de shared-nothing y el motor de procesamiento masivamente paralelo (MPP), y consiste en numerosos nodos lógicos independientes que no comparten los recursos del sistema, tales como CPUs, memoria y almacenamiento. En una arquitectura de sistema de este variante, los datos de servicio se almacenan por separado en numerosos nodos. Las tareas de análisis de datos se ejecutan en paralelo en los nodos donde se almacenan los datos. El procesamiento de datos masivamente paralelo mejora significativamente la velocidad de respuesta.

Figura 1-1 Arquitectura



- **Capa de aplicación**

Las herramientas de carga de datos, las herramientas de extracción, transformación y carga (ETL), las herramientas de inteligencia empresarial (BI), así como las herramientas de análisis y excavación de datos, se pueden integrar con GaussDB(DWS) a través de estándar de APIs. GaussDB(DWS) es compatible con el ecosistema de PostgreSQL, y la sintaxis SQL es compatible con Oracle, MySQL, y Teradata. Las aplicaciones pueden migrarse sin problemas a GaussDB(DWS) con pocos cambios.

- **API**

Las aplicaciones pueden conectarse a GaussDB(DWS) a través del estándar de Java Database Connectivity (JDBC) 4.0 y Open Database Connectivity (ODBC) 3.5.

- **GaussDB(DWS) (MPP cluster)**

Un clúster de GaussDB(DWS) contiene nodos del mismo variante en la misma subred. Estos nodos proporcionan conjuntamente servicios. Los nodos de datos (DNs) en un clúster almacenan datos en discos. Los coordinadores (CNs) reciben solicitudes de acceso desde las aplicaciones y devuelven los resultados de la ejecución a los clientes. Además, un CN divide y distribuye tareas a los DNs para el procesamiento en paralelo.

- **Copia de seguridad automática de datos**

Las instantáneas de clúster se pueden hacer una copia automáticamente en Object Storage Service (OBS) de nivel EB, lo que facilita la copia de seguridad periódica del clúster durante las horas no pico y garantiza la recuperación de datos después de que se produzca una excepción de clúster.

Una instantánea es una copia de seguridad completa de GaussDB(DWS) en un punto de tiempo específico, incluidos los datos de configuración y los datos de servicio de un clúster.

- **Cadena de herramientas**

Se proporcionan la herramienta de carga de datos en paralelo General Data Service (GDS), la herramienta de migración de sintaxis SQL Database Schema Convertor (DSC) y la herramienta de desarrollo SQL Data Studio. La O&M del clúster se puede supervisar en una consola.

Arquitectura de clústeres lógicos

Figura 1-2 muestra la arquitectura lógica de un clúster de GaussDB(DWS). Para obtener más información sobre las instancias, consulte **Tabla 1-1**.

Figura 1-2 Arquitectura de clústeres lógicos

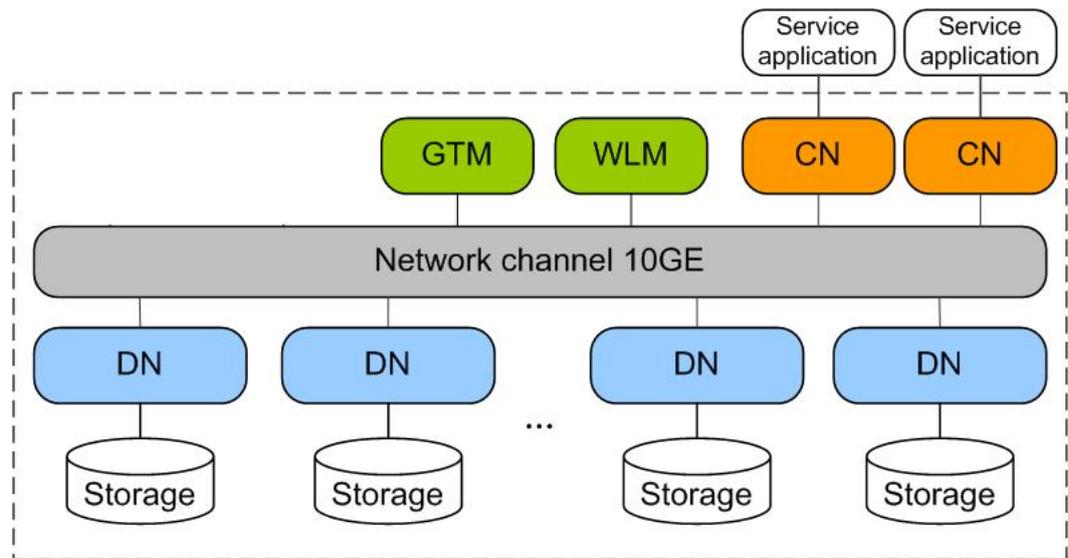


Tabla 1-1 Descripción de la arquitectura de clúster

Nombre	Descripción	Comentarios
Global Transaction Manager (GTM)	Genera y mantiene la información única globalmente, como el ID de la transacción, la instantánea de la transacción y la marca de tiempo.	El clúster incluye solo un par de GTMs: un GTM primario y un GTM de reserva.
Workload Manager (WLM)	WLM controla la asignación de recursos del sistema para evitar la congestión del servicio y el bloqueo del sistema como resultado de una carga de trabajo excesiva.	No necesita especificar los nombres de los hosts en los que se van a implementar WLMs, ya que el programa de instalación instala automáticamente un WLM en cada host.

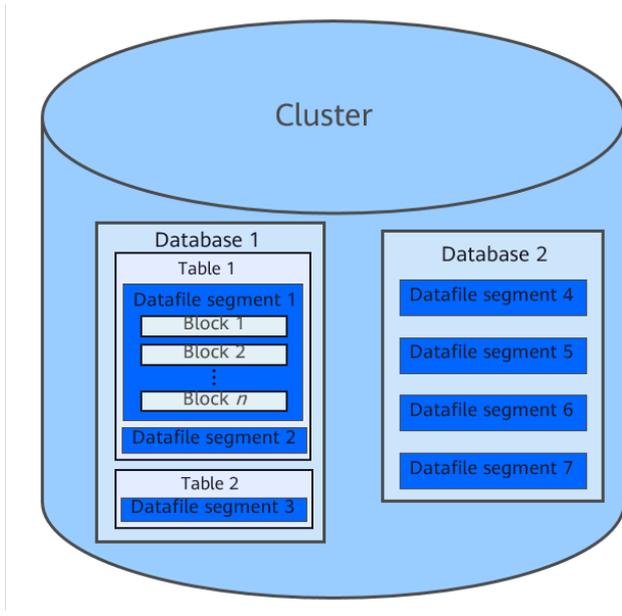
Nombre	Descripción	Comentarios
Coordinador (CN)	Un CN recibe solicitudes de acceso desde aplicaciones y devuelve resultados de ejecución al cliente; divide tareas y asigna fragmentos de tarea a diferentes DN para procesamiento paralelo.	<p>Los CNs en un clúster tienen funciones equivalentes y devuelven el mismo resultado para la misma instrucción DML. Se pueden añadir equilibradores de carga entre CNs y aplicaciones para garantizar que los CNs sean transparentes para las aplicaciones. Si un CN es defectuoso, el equilibrador de carga conecta sus aplicaciones a otro CN.</p> <p>Los CNs necesitan conectarse entre sí en la arquitectura de transacciones distribuidas. Para reducir la carga pesada causada por hilos excesivos en GTMs, no se deben configurar más de 10 CN en un clúster.</p> <p>GaussDB(DWS) maneja la carga de recursos globales en un clúster usando el Central Coordinator (CCN) para la gestión de carga dinámica adaptativa. Cuando se inicia el clúster por primera vez, el CM selecciona el CN con el ID más pequeño como el CCN. Si el CCN es defectuoso, CM lo reemplaza por uno nuevo.</p>
Datano de (DN)	Un DN almacena datos de servicio por columna o fila o en el modo híbrido, ejecuta tareas de consulta de datos y devuelve los resultados de ejecución a los CNs.	Un clúster consta de múltiples DN y cada DN almacena parte de los datos. Si no hay ninguna solución de HA disponible para los DN, no se puede acceder a los datos cuando un DN es defectuoso.
Almacenamiento	Funciona como recursos de almacenamiento local del servidor para almacenar datos de forma permanente.	-

Los DN de un clúster almacenan datos en discos. [Figura 1-3](#) describe los objetos en cada DN y las relaciones entre ellos lógicamente.

- Una base de datos gestiona varios objetos de datos y está aislada de otras bases de datos.
- Un segmento de archivo de datos almacena datos en una sola tabla. Una tabla que contiene más de 1 GB de datos se almacena en varios segmentos de archivo de datos.
- Una tabla pertenece solo a una base de datos.
- Un bloque es la unidad básica de gestión de bases de datos, con un tamaño predeterminado de 8 KB.

Los datos se pueden distribuir en modo de replicación, round-robin o hash. Puede especificar el modo de distribución durante la creación de la tabla.

Figura 1-3 Arquitectura de base de datos lógica



2 Tipos de almacén de datos

Descripción general del tipo de producto

- **Almacén de datos en la nube:** puede analizar datos calientes y fríos y es altamente rentable. Sus recursos de almacenamiento y computación no están limitados, y se pueden escalar y facturar elásticamente por uso. Es adecuado para el análisis convergente que requiere bases de datos integradas, almacenes, mercados y lagos. Es más adecuado para cargas de trabajo OLAP.
- **Almacén de datos de streaming:** proporciona capacidades de análisis de IoT y computación de series de tiempo eficientes basadas en el almacén de datos estándar y soporta la correlación entre datos históricos y en tiempo real. La relación de compresión puede alcanzar 40:1. Se puede utilizar para el análisis en tiempo real de IoT.
- **Almacenamiento de datos híbrido:** proporciona capacidades de procesamiento de transacciones de alta concurrencia, alto rendimiento y baja latencia a bajo costo basado en capacidades de análisis y consulta de datos a gran escala. El almacén de datos se puede utilizar para procesar cargas híbridas HTAP y se puede implementar en modo de nodo único o de clúster.

Restricciones de funciones

Módulo	Función	La Nube	IoT	Híbrido (Cluster)	Híbrido (Standalone)
Menú de navegación	Panel	Sí	Sí	Sí	Sí
	Gestión de clúster	Sí	Sí	Sí	Sí
	Gestión de DR	Sí	Sí	Sí	No
	Gestión de instantánea	Sí	Sí	Sí	No
	Parámetros	Sí	Sí	Sí	Sí
	Gestión de incidentes	Sí	Sí	Sí	Sí
	Gestión de alarmas	Sí	Sí	Sí	Sí

	Gestión de conexiones	Sí	Sí	Sí	Sí
Panel	Recursos	Sí	Sí	Sí	Sí
	Alarmas	Sí	Sí	Sí	Sí
	Eventos recientes	Sí	Sí	Sí	Sí
	Métricas de monitoreo de clústeres (DMS)	Sí	Sí	Sí	No
Gestión de clústeres	Panel de monitoreo (DMS)	Sí	Sí	Sí	No
	Monitoreo de métricas (Cloud Eye)	Sí	Sí	Sí	No
	Reiniciar	Sí	Sí	Sí	Sí
	Ajuste de escala	Sí	Sí	Sí	No
	Redistribución de datos	Sí	No	Sí	No
	Consulta de detalles de redistribución	Sí	No	Sí	No
	Restablecer contraseñas	Sí	Sí	Sí	Sí
	Creación de instantáneas	Sí	Sí	Sí	No, no
	Cancelación de estado de sólo lectura	Sí	Sí	Sí	Sí
	Eliminación	Sí	Sí	Sí	Sí
	Gestión de CNs	Sí	Sí	Sí	No, no
	Ajuste de capacidad del espacio de almacenamiento	Sí	No	Sí	Sí

Información básica	Información básica	Sí	Sí	Sí	Sí
	ELB	Sí	Sí	Sí	No
	Pods	Sí	Sí	Sí	No
	Clúster lógico	Sí	Sí	Sí	No
	Instantánea	Sí	Sí	Sí	No
	Modificaciones de parámetros	Sí	Sí	Sí	Sí
	Ajustes de seguridad	Sí	Sí	Sí	Sí
	Fuentes de datos de MRS	Sí	Sí	Sí	No
	Panel de monitoreo	Sí	Sí	Sí	No
	Etiquetas	Sí	Sí	Sí	Sí
	Administración del nodo	Sí	No	Sí	No
Gestión de DR	Gestión de DR	Sí	Sí	Sí	No
Gestión de instantáneas	Restauración	Sí	Sí	Sí	No
	Eliminación	Sí	Sí	Sí	No
	Copiar	Sí	Sí	Sí	No
Gestión de incidente	Gestión de evento (general)	Sí	Sí	Sí	Sí
Gestión de alarma	Gestión de alarma	Sí	Sí	Sí	Sí
Gestión de conexiones	Gestión de conexiones	Sí	Sí	Sí	Sí
Otros	Inspección	Sí	Sí	Sí	No
	O&M Inteligente	Sí	Sí	Sí	No
	Restauración de nodo	Sí	Sí	Sí	No
	Copia de seguridad cálida al lado del inquilino	Sí	Sí	Sí	No

	Open APIs	Sí	No	No	No
--	-----------	----	----	----	----

 **NOTA**

- Un almacén de datos híbrido implementado en modo independiente no se puede cambiar al modo distribuido. No se admite el escalado, las operaciones de clúster lógico y la gestión de recursos.
- Un almacén de datos híbrido en modo independiente no admite funciones como instantáneas, monitoreo de DMS y fuentes de datos MRS. Estas funciones estarán disponibles en versiones posteriores.
- En un almacén de datos híbrido implementado en modo independiente, se ignorará la sintaxis original para especificar columnas de distribución. No se admiten jobs programados, el tipo de sequence, la importación y exportación de tablas externas de HDFS/OBS, el almacenamiento de varias temperaturas y las tablas particionadas de aumento automático. Estas funciones se van a soportar en versiones posteriores.
- Los almacenes de datos de streaming no se pueden escalar en línea.

3 Ventajas

GaussDB(DWS) utiliza el núcleo de base de datos GaussDB desarrollado por Huawei y es compatible con PostgreSQL 9.2.4. Se transforma de una única base de datos OLTP a una base de datos OLAP distribuida a nivel empresarial orientada al análisis masivo de datos basado en la arquitectura de procesamiento masivamente paralelo (MPP).

A diferencia de los almacenes de datos convencionales, GaussDB(DWS) sobresale en el procesamiento masivo de datos y la gestión general de la plataforma con las siguientes ventajas:

Facilidad de uso

- Gestión visualizada de una sola parada

GaussDB(DWS) le permite completar fácilmente todo el proceso desde el concepto del proyecto hasta la implementación de producción. Con la consola de gestión de GaussDB(DWS), puede obtener un clúster de almacén de datos de nivel empresarial de alto rendimiento y alta disponibilidad en varios minutos. No se requiere software de almacén de datos ni servidores de almacén de datos.

Con solo unos pocos clics, puede conectar fácilmente las aplicaciones al almacén de datos, realizar copias de seguridad de datos, restaurar datos y supervisar los recursos y el rendimiento del almacén de datos.

- Integración perfecta con big data

Sin la necesidad de migrar datos, puede utilizar instrucciones SQL estándar para consultar directamente datos en HDFS y OBS.

- Herramientas de migración de bases de datos heterogéneas

GaussDB(DWS) proporciona varias herramientas de migración para migrar scripts de SQL de Oracle y Teradata a GaussDB(DWS).

Alto rendimiento

- Arquitectura distribuida basada en la nube

GaussDB(DWS) adopta la base de datos basada en MPP de modo que los datos de servicio se almacenan por separado en numerosos nodos. Las tareas de análisis de datos se ejecutan en paralelo en los nodos donde se almacenan los datos. El procesamiento de datos masivamente paralelo mejora significativamente la velocidad de respuesta.

- Respuesta de consulta a billones de registros de datos en segundos

GaussDB(DWS) mejora el rendimiento de la consulta de datos mediante la ejecución de operadores de subprocesos múltiples en paralelo, la ejecución de comandos en registros

en paralelo con el motor de computación vectorizado, y la reducción de las condiciones de juicio redundantes usando LLVM.

GaussDB(DWS) le proporciona una mejor relación de compresión de datos (almacén de columnas), una mejor indexación (almacén de columnas) y un mayor rendimiento de actualización y consulta de puntos (almacén de filas).

- **Carga rápida de datos**

GDS es una herramienta que le ayuda con la carga de datos masivamente paralelos a alta velocidad.

Alta escalabilidad

- Escalamiento horizontal bajo demanda: con la arquitectura abierta de shared-nothing, se pueden agregar nodos en cualquier momento para mejorar las capacidades de almacenamiento, consulta y análisis de datos del sistema.
- Rendimiento lineal mejorado después del escalamiento horizontal: la capacidad y el rendimiento aumentan linealmente con la escala del clúster. La tasa lineal es 0.8.
- Continuidad del servicio: durante el escalado horizontal, se pueden agregar, eliminar, modificar y consultar datos, y se pueden realizar operaciones DDL (**DROP/TRUNCATE/ALTER TABLE**). El escalamiento horizontal a nivel de tabla en línea garantiza la continuidad del servicio.

Fiabilidad robusta

- **ACID**

Soporte para la función de atómica, consistencia, aislamiento y durabilidad (ACID), que garantiza una sólida consistencia de datos para transacciones distribuidas.

- **Diseño integral de HA**

Todos los procesos de software de GaussDB(DWS) están en modo activo/de reserva. Los componentes lógicos tales como los CNs y los DNs de cada clúster también funcionan en modo activo/de reserva. Esto garantiza la fiabilidad y consistencia de los datos cuando se produce cualquier punto único de fallo (SPOF).

- **Alta seguridad**

GaussDB(DWS) es compatible con el cifrado de datos transparente y puede interconectarse con el Database Security Service (DBSS) para proteger mejor la privacidad del usuario y la seguridad de los datos con aislamiento de red y opciones de configuración de reglas de grupo de seguridad. Además, GaussDB(DWS) admite copias de seguridad automáticas completas e incrementales de los datos, lo que mejora la confiabilidad de los datos.

Bajo costo

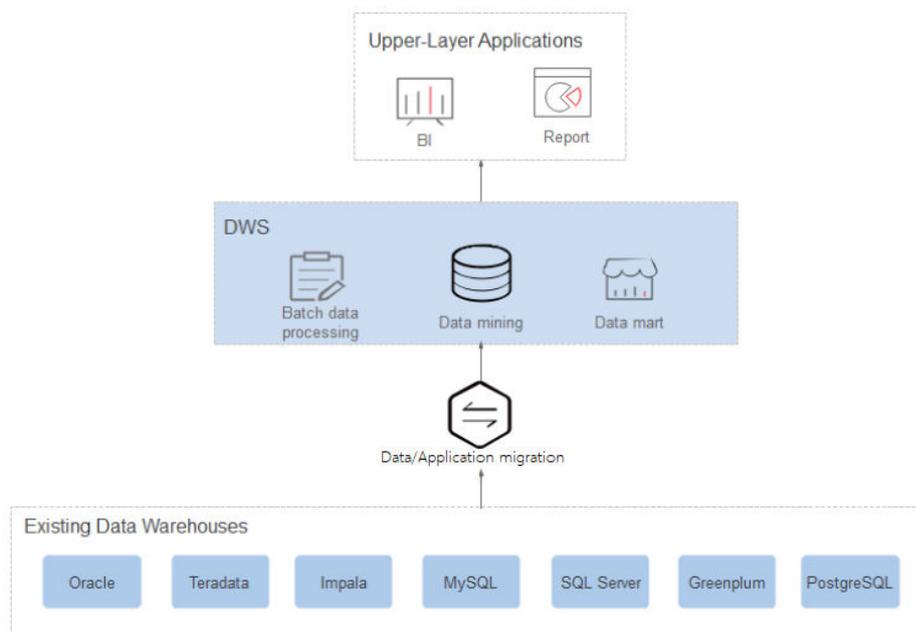
- **Pago por uso:** GaussDB(DWS) se factura en función del uso y la duración del uso. Solo tiene que pagar por los recursos que usa.
- **Inversión flexible en infraestructura:** No es necesario invertir mucho en infraestructura en la etapa inicial. Puede comenzar desde una instancia de almacén de datos con especificaciones bajas y escalarla de forma flexible en cualquier momento.

4 Escenarios de aplicación

Migración de Data Warehouse

El almacén de datos es un importante sistema de análisis de datos para las empresas. A medida que crece el volumen de servicio, el rendimiento de sus propios almacenes de datos no puede satisfacer los requisitos reales de servicio debido a la limitación de escalabilidad y los altos costos. Como almacén de datos de clase empresarial en la nube, GaussDB(DWS) ofrece alto rendimiento, bajo costo y fácil escalabilidad, satisfaciendo los requisitos en la era del big data.

Figura 4-1 Migración del almacén de datos



Ventajas

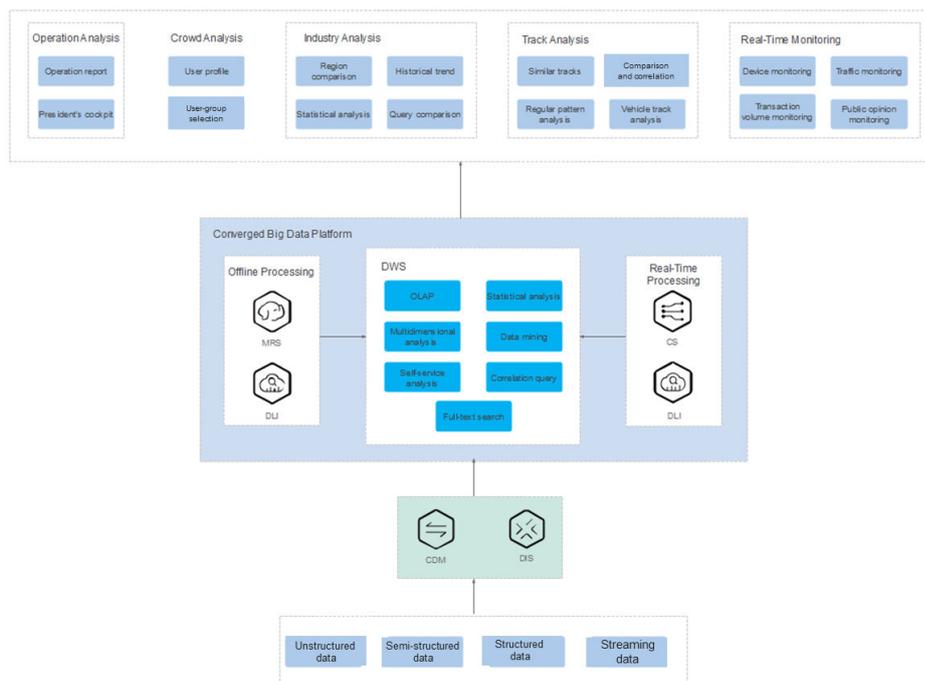
- Migración sin interrupciones
GaussDB(DWS) proporciona varias herramientas de migración para garantizar la migración sin problemas de los sistemas de análisis de datos populares como Teradata, Oracle, MySQL, SQL Server, PostgreSQL, Greenplum e Impala.

- Compatible con almacenes de datos convencionales
GaussDB(DWS) soporta el estándar de SQL 2003 y los procedimientos almacenados. Es compatible con algunas estructuras de datos y sintaxis de Oracle, y se puede interconectar sin problemas con herramientas de BI comunes, lo que ahorra esfuerzos de migración de servicios.
- Segura y confiable
GaussDB(DWS) admite el cifrado de datos y se conecta a DBSS para garantizar la seguridad de los datos en la nube. Además, GaussDB(DWS) admite copias de seguridad automáticas completas e incrementales de los datos, lo que mejora la confiabilidad de los datos.

Análisis de Big Data Convergente

Los datos se han convertido en el activo más importante. Las empresas deben ser capaces de integrar sus recursos de datos y construir plataformas de big data para extraer todo el valor de sus datos. En los casos de uso de análisis predictivo, se deben procesar volúmenes masivos de datos. Huawei GaussDB(DWS) ofrece la potencia de procesamiento necesaria para manejar estos intensos escenarios de cómputo.

Figura 4-2 Análisis de big data convergente



Ventajas

- Entrada de análisis unificado
El GaussDB(DWS) SQL sirve como la entrada unificada de aplicaciones de capa superior, de modo que los desarrolladores de aplicaciones pueden acceder a todos los datos utilizando el SQL.
- Análisis interactivo en tiempo real
El personal de análisis puede obtener información inmediatamente procesable de la plataforma de big data en tiempo real utilizando solicitudes de análisis amplias.

- Escalado automático
La adición de nodos le permite expandirse fácilmente a la capacidad de rango PB mientras mejora el rendimiento de consultas y análisis del sistema.

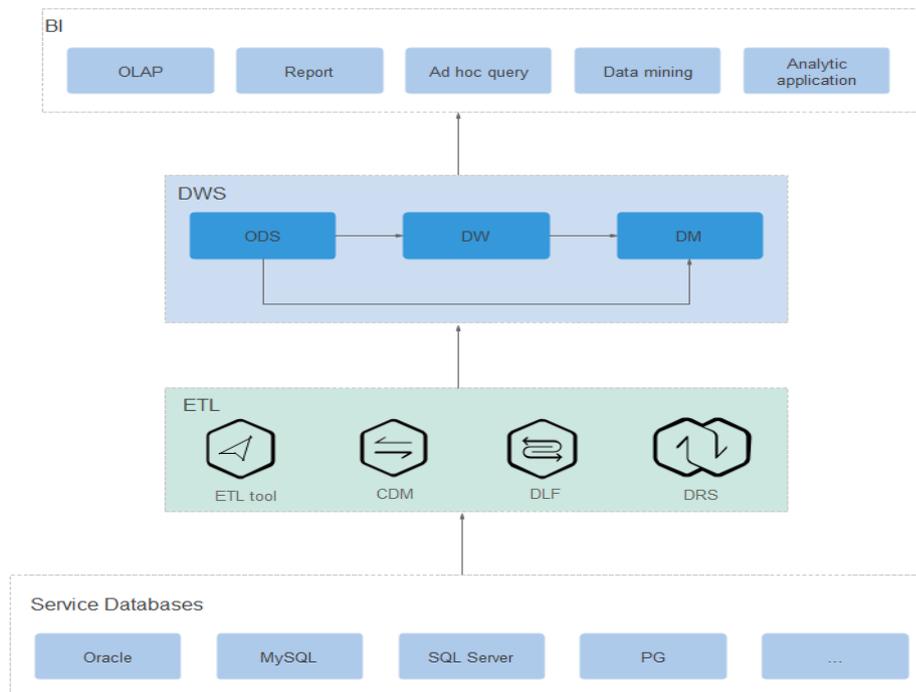
Análisis mejorado de ETL + BI en tiempo real

El almacén de datos es el pilar del sistema de BI para recopilar, almacenar y analizar volúmenes masivos de datos. Impulsa el análisis de decisiones empresariales para las industrias de IoT, finanzas, educación, Internet móvil y Online to Offline (O2O).

Ventajas

- Migración de datos
Capacidad para importar datos en lotes en tiempo real desde múltiples orígenes de datos.
- Alto rendimiento
Almacenamiento de datos rentable a nivel de PB y respuesta al análisis de correlación de billones de registros de datos en cuestión de segundos.
- Tiempo real
Consolidación en tiempo real de datos de servicio para producir información procesable en la toma de decisiones operativas.

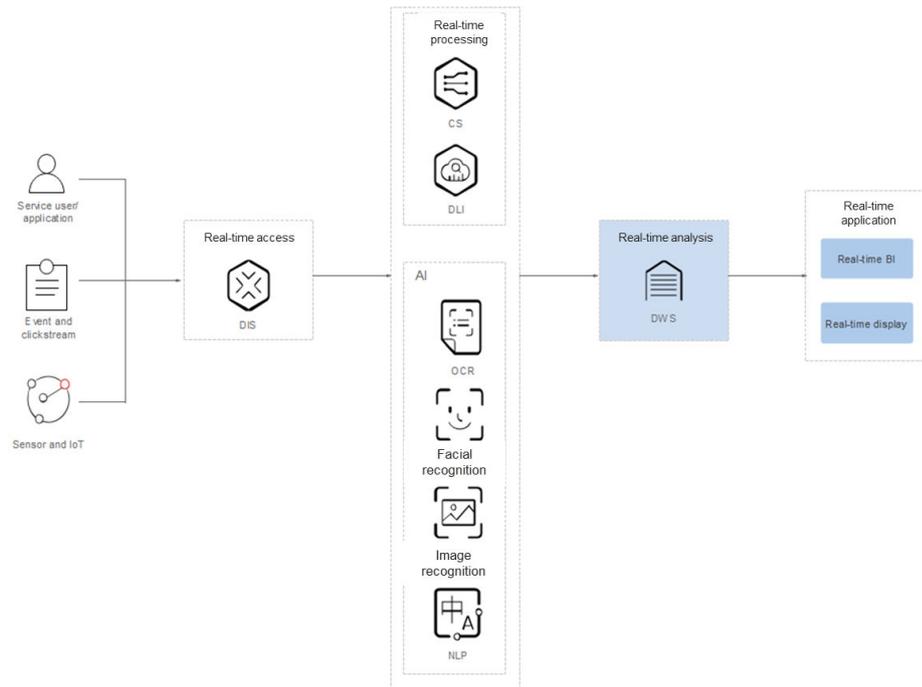
Figura 4-3 ETL mejorado + análisis de BI en tiempo real



Análisis de datos en tiempo real

En los dominios de Internet móvil y IoT, se deben procesar y analizar enormes volúmenes de datos en tiempo real para extraer los datos completos de los datos. Las capacidades rápidas de importación y consulta de datos de GaussDB(DWS) aceleran el análisis de datos para permitir la gestión, el procesamiento y la generación de valor en tiempo real.

Figura 4-4 Análisis de datos en tiempo real



Ventajas

- **Importación en tiempo real de datos de streaming**
Los datos de las aplicaciones de IoT e Internet se pueden escribir en GaussDB(DWS) en tiempo real después de ser procesados por los servicios de computación de flujo e IA.
- **Monitoreo y predicción en tiempo real**
Monitoreo, control, optimización, suministro, autodiagnóstico y autocuración de dispositivos basados en análisis y predicción de datos.
- **Análisis de IA convergente**
El análisis de correlación se puede llevar a cabo en resultados de análisis de datos de imágenes y texto por servicios de IA y otros datos de servicio en GaussDB(DWS).
- **Comercio Electrónico**
Huawei VMALL aprovecha GaussDB(DWS) como su motor de base de datos para el análisis de datos. Los datos de los minoristas en línea se utilizan principalmente para recomendaciones de marketing, análisis de operación y de clientes, y búsqueda de texto completo.
Las ventajas de GaussDB(DWS) son las siguientes:
 - **Análisis multidimensional:** análisis de productos, usuarios, operaciones y regiones
 - **Escalamiento horizontal a medida que crece el negocio:** escalamiento horizontal del clúster bajo demanda a medida que crece el negocio
 - **Alta fiabilidad:** funcionamiento estable a largo plazo del sistema de comercio electrónico
- **IoT**
GaussDB(DWS) le ayuda a analizar cantidades masivas de datos de IoT en tiempo real y realizar una optimización basada en los resultados. Es ampliamente utilizado en IoT industrial, sistema de servicio de O2O y soluciones de IoV.

Las ventajas de GaussDB(DWS) son las siguientes:

- **Archivo en tiempo real de datos de flujo:** importación de datos de flujo desde dispositivos de IoT y la puerta de enlace a GaussDB(DWS) usando Huawei Cloud DIS
- **Monitorización y predicción de dispositivos:** monitorización, control, optimización, suministro, autodiagnóstico y autocuración de dispositivos basada en análisis y predicción de datos
- **Recomendación de información:** recomendación de seguimiento basada en los datos de los dispositivos conectados de los usuarios

5 Funciones

GaussDB(DWS) le permite usar este servicio a través de varios métodos, como la consola de gestión de GaussDB(DWS), el cliente de GaussDB(DWS) y REST APIs. Esta sección describe las funciones principales de GaussDB(DWS).

Gestión de clústeres

Un clúster de almacén de datos contiene nodos del mismo variante en la misma subred. Estos nodos proporcionan conjuntamente servicios. GaussDB(DWS) proporciona una consola de gestión profesional, eficiente y centralizada, que le permite solicitar rápidamente clústeres, administrar fácilmente almacenes de datos y centrarse en datos y servicios.

Las funciones principales de la gestión de clústeres se describen a continuación:

- **Creación de un clúster**

Para utilizar los servicios de almacén de datos en la nube, primero cree un clúster de GaussDB(DWS). Puede seleccionar especificaciones de producto y nodo para crear rápidamente un clúster. También puede comprar un paquete de descuento para crear un clúster.
- **Gestión de instantáneas**

Una instantánea es una copia de seguridad completa que registra datos de configuración y datos de servicio en un punto de tiempo de un clúster de GaussDB(DWS). Una instantánea se puede utilizar para restaurar un clúster en un momento determinado. Puede crear instantáneas manualmente para un clúster o habilitar la creación automática de instantáneas (periódicas). Las instantáneas automatizadas tienen un período de retención limitado. Puede copiar instantáneas automáticas para una retención a largo plazo.

Cuando restaura un clúster a partir de una instantánea, el sistema crea un nuevo clúster con el mismo variante y cantidad de nodos que el original, e importa los datos de la instantánea.

Puede eliminar instantáneas que ya no sean necesarias para liberar el espacio de almacenamiento.
- **Gestión de nodos**

Puede consultar los nodos en un clúster, incluidos el estado, las especificaciones y el uso de cada nodo. Para prepararse para un gran escalamiento horizontal, puede agregar nodos en lotes. Por ejemplo, si se necesitan 180 nodos BMS más para un escalamiento horizontal, agréguelos en tres lotes (60 en cada lote). Si algunos nodos no se pueden

agregar, agréguelos de nuevo. Una vez que todos los 180 nodos se hayan agregado correctamente, utilice los nodos para el escalamiento horizontal del clúster. La adición de nodos no afecta a los servicios de clúster.

- Escalamiento horizontal de un clúster

A medida que aumenta el volumen de servicio, es posible que la escala actual de un clúster no cumpla con los requisitos de servicio. En este caso, puede escalar el clúster agregando nodos de cómputo a él. Los servicios no se interrumpen durante el proceso de escalamiento. Puede habilitar el escalamiento horizontal en línea y la redistribución automática si es necesario.

- Gestión de redistribución

De forma predeterminada, la redistribución se inicia automáticamente después de la escalamiento horizontal del clúster. Para una mayor fiabilidad, deshabilite la función de redistribución automática e inicie manualmente una tarea de redistribución después de que el escalamiento horizontal se realice correctamente. La redistribución de datos puede acelerar la respuesta del servicio. Actualmente, se admite la redistribución sin conexión, la redistribución en línea y la programación sin conexión. El modo predeterminado es la redistribución sin conexión.

- Escalamiento del espacio de almacenamiento

En el escalamiento convencional, los recursos de computación y de almacenamiento están acoplados. Si una empresa escala discos, tiene que agregar CPUs y memoria innecesarias al mismo tiempo. El escalamiento tarda mucho tiempo e interrumpe los servicios. La expansión de la capacidad del disco puede aumentar rápidamente el almacenamiento sin interrupción del servicio. Puede aumentar el espacio en disco sin tener que detener los servicios.

- Gestión de cargas de trabajo

Cuando varios usuarios de bases de datos consultan trabajos al mismo tiempo, algunas consultas complejas pueden ocupar recursos de clúster durante mucho tiempo, lo que afecta al rendimiento de otras consultas. Por ejemplo, un grupo de usuarios de base de datos envía continuamente consultas complejas y que consumen mucho tiempo, mientras que otro grupo de usuarios envía con frecuencia consultas cortas. En este caso, las consultas cortas pueden tener que esperar en la cola para que se completen las consultas que consumen mucho tiempo. Para mejorar la eficiencia, puede utilizar la función de administración de cargas de trabajo de GaussDB(DWS) para manejar estos problemas. La gestión de cargas de trabajo de GaussDB(DWS) utiliza colas de cargas de trabajo como portadoras de recursos. Puede crear diferentes colas de carga de trabajo para diferentes tipos de servicio y configurar diferentes proporciones de recursos para estas colas. A continuación, agregue usuarios de base de datos a las colas correspondientes para restringir sus usos de recursos.

- Clúster lógico

Un clúster físico se puede dividir en clústeres lógicos que utilizan el mecanismo de grupo de nodos. Las tablas de una base de datos se pueden asignar a diferentes nodos físicos por clúster lógico. Un clúster lógico puede contener tablas de varias bases de datos.

- Reiniciar un clúster

El reinicio de un clúster puede provocar la pérdida de datos en los servicios en ejecución. Si tiene que reiniciar un clúster, asegúrese de que no hay servicio en ejecución y que se han guardado todos los datos.

- Eliminación de un clúster

Puede eliminar un clúster cuando no lo necesite. Eliminar un clúster es arriesgado y puede causar pérdida de datos. Por lo tanto, realice esta operación con precaución.

GaussDB(DWS) le permite gestionar clústeres e instantáneas de cualquiera de las siguientes maneras:

- Consola de gestión

Utilice la consola de gestión para acceder a los clústeres de GaussDB(DWS). Cuando haya registrado una cuenta, inicie sesión en la consola de gestión y elija **Data Warehouse Service**.

Para obtener más información acerca de la gestión de clústeres, consulte [Gestión de clústeres](#).

- REST APIs

Utilice REST APIs proporcionadas por GaussDB(DWS) para gestionar clústeres. Además, si necesita integrar GaussDB(DWS) en un sistema de terceros para desarrollo secundario, use APIs para acceder al servicio.

Para obtener más información, consulte la [Referencia de la API de Data Warehouse Service \(DWS\)](#).

Almacenes de datos de nivel empresarial y compatibilidad con estándar de SQL

Después de crear un clúster de almacén de datos, puede utilizar el cliente de SQL para conectarse al clúster y realizar operaciones como crear una base de datos, gestionar la base de datos, importar y exportar datos y consultar datos.

GaussDB(DWS) proporciona bases de datos de alto rendimiento a nivel de petabyte (PB) con las siguientes características:

- Marco de computación MPP, almacenamiento híbrido de columna y fila y ejecución vectorizada, lo que permite responder al análisis de correlación de datos de nivel de mil millones dentro de segundos
- Optimización de computación en memoria basada en Hash Join de Bloom Filter, lo que mejora el rendimiento de 2 a 10 veces
- Comunicación optimizada entre clústeres a gran escala basada en tecnologías de telecomunicaciones, lo que mejora la eficiencia de transmisión de datos entre nodos de computación
- Optimizadores inteligentes basados en costos, que ayudan a generar el plan óptimo basado en la escala del clúster y el volumen de datos para mejorar la eficiencia de la ejecución

GaussDB(DWS) tiene capacidades SQL integrales:

- Soporta estándares de SQL 92 y SQL 2003, procedimientos almacenados, conjuntos de caracteres GBK y UTF-8, y funciones estándares de SQL y funciones de análisis de OLAP.
- Compatible con el ecosistema de PostgreSQL y admite la interconexión con las herramientas de base de datos principales de ETL y BI proporcionadas por proveedores de terceros.

Para obtener más información acerca de la sintaxis de SQL y la guía de operación de la base de datos, consulte la [Guía del desarrollador de Data Warehouse Service \(DWS\)](#).

Diversos modos de importación de datos

GaussDB(DWS) admite la importación eficiente de datos desde múltiples fuentes de datos. The following lists typical data import modes. Para obtener más información, consulte "Importación de datos" en la *Guía del desarrollador del Data Warehouse Service (DWS)*.

- Importación simultánea de datos desde OBS
- Uso de GDS para importar datos desde un servidor remoto
- Ejecución de la instrucción INSERT para insertar datos
- Ejecución de la instrucción COPY FROM STDIN para importar datos
- Uso de un comando meta gsql para importar datos
- Importación de datos de MRS a un clúster de almacén de datos
- Importación de datos de flujo de DIS a GaussDB(DWS)
- Uso de CDM para migrar datos a GaussDB(DWS)
- Uso del Database Schema Convertor (DSC) para migrar scripts de SQL

Además, GaussDB(DWS) admite la importación de datos utilizando herramientas principales de ETL de terceros.

APIs

Puede llamar a las APIs estándares, como JDBC y ODBC, para acceder a las bases de datos en clústeres de GaussDB(DWS).

Para obtener más información, consulte [Uso de los controladores JDBC para conectarse a una base de datos](#) y [Uso de los controladores JDBC para conectarse a una base de datos](#).

Alta fiabilidad

- Soporta redundancia de instancia y datos, asegurando cero puntos únicos de fallo (SPOF) en todo el sistema.
- Soporta múltiples copias de seguridad de datos, y todos los datos se pueden hacer la copia de seguridad manualmente en OBS.
- Aísla automáticamente el nodo defectuoso, utiliza la copia de seguridad para restaurar los datos y reemplaza el nodo defectuoso cuando sea necesario.
- Las instantáneas automáticas funcionan con OBS para implementar la recuperación ante desastres (DR) entre AZ. Si el clúster de producción no proporciona servicios de lectura y escritura debido a desastres naturales en la región especificada o fallas internas del clúster, el clúster de DR se convierte en el clúster de producción para garantizar la continuidad del servicio.
- En el estado **Unbalanced**, el número de instancias primarias en algunos nodos aumenta. Como resultado, la presión de carga es alta. En este caso, puede realizar un switchback primario/de reserva para el clúster durante las horas no pico para mejorar el rendimiento.
- Si la dirección IP interna o EIP de un CN se utiliza para conectarse a un clúster, el fallo de este CN conducirá a un fallo de conexión del clúster. Para evitar fallos de un solo CN, GaussDB(DWS) utiliza el Equilibrio de carga elástica (ELB). Un ELB distribuye el tráfico de acceso a múltiples ECS para el control de tráfico basado en políticas de reenvío. Mejora la capacidad de tolerancia a fallos de los programas de aplicación.
- Después de crear un clúster, el número de CN necesarios varía según los requisitos de servicio. GaussDB(DWS) le permite agregar o eliminar CNs según sea necesario.

Gestión de la seguridad

- Aísla a los inquilinos y controla los permisos de acceso para proteger la privacidad y la seguridad de los datos de los sistemas y usuarios en función del aislamiento de la red y las reglas de grupo de seguridad, así como las medidas de refuerzo de la seguridad.
- Admite conexiones de red de SSL, gestión de permisos de usuario y gestión de contraseñas, lo que garantiza la seguridad de los datos en la capa de red, gestión, aplicación y sistema.

Para obtener más información, consulte [Configuración de la conexión de SSL](#) y [Configuración de la separación de permisos](#).

Monitoreo y Auditoría

- Monitoreo de clústeres
GaussDB(DWS) se integra con Cloud Eye, lo que le permite monitorear nodos de cómputo y bases de datos en el clúster en tiempo real. Para obtener más información, consulte [Monitoreo de clúster](#).
- Monitoreo de base de datos
El DMS es proporcionado por GaussDB(DWS) para asegurar el funcionamiento rápido y estable de las bases de datos. Recopila, monitorea y analiza los datos de métricas de disco, red y OS utilizados por la base de datos de servicio, así como los datos clave de métricas de rendimiento de la ejecución del clúster. También diagnostica hosts de base de datos, instancias y sentencias de SQL de servicio en función de las métricas recopiladas para exponer fallas clave y problemas de rendimiento en una base de datos de manera oportuna, y guía a los clientes para optimizar y resolver los problemas. Para obtener más información, consulte [Monitoreo de base de datos](#).
- Gestión de alarmas
La gestión de alarmas incluye la visualización y configuración de reglas de alarma y la suscripción a la información de alarma. Las reglas de alarma muestran estadísticas de alarmas y detalles de la semana pasada para que los usuarios vean las alarmas del inquilino. Además de proporcionar un conjunto de reglas de alarma de GaussDB(DWS) predeterminadas, esta característica le permite modificar los umbrales de alarma basados en sus propios servicios. Para obtener más información, consulte [Alarmas](#).
- Notificación de evento
GaussDB(DWS) se interconecta con Simple Message Notification (SMN) para que pueda suscribirse a eventos y consultar eventos que se activan. Para obtener más información, consulte [Notificaciones de eventos](#).
- Registros de auditoría
 - GaussDB(DWS) se integra con Cloud Trace Service (CTS), lo que le permite auditar las operaciones realizadas en la consola de gestión y las operaciones de invocación de API. Para obtener más información, consulte [Consulta de registros de auditoría de operaciones clave en la gestión de consola](#).
 - GaussDB(DWS) registra todas las operaciones SQL, incluidos los intentos de conexión, los intentos de consulta y los cambios en la base de datos. Para obtener más información, consulte [Configuración de registros de auditoría de base de datos](#).

Múltiples herramientas de base de datos

GaussDB(DWS) proporciona las siguientes herramientas autodesarrolladas. Puede descargar los paquetes de herramientas en la consola de gestión de GaussDB(DWS). Para obtener más

información sobre las herramientas, consulte la [Guía de herramienta de Data Warehouse Service\(DWS\)](#).

- **gsql**
gsql es una herramienta de cliente de SQL de línea de comandos que se ejecuta en el sistema operativo Linux. Ayuda a conectar, operar y mantener la base de datos en un clúster de almacén de datos.
- **Data Studio**
Data Studio es una herramienta de cliente SQL de Interfaz gráfica de usuario (GUI) que se ejecuta en el sistema operativo Windows. Se utiliza para conectarse a la base de datos en un clúster de almacén de datos, gestionar la base de datos y objetos de base de datos, editar, ejecutar y depurar scripts de SQL, y ver los planes de ejecución.
- **GDS**
GDS es una herramienta de servicio de datos proporcionada por GaussDB(DWS). Funciona con el mecanismo de tabla extranjera para implementar la importación y exportación de datos de alta velocidad.
El paquete de herramientas GDS debe instalarse en el servidor donde se encuentra el archivo fuente de datos. Este servidor se denomina servidor de datos o servidor GDS.
- **Herramienta de migración de sintaxis de SQL DSC**
El DSC es una herramienta de línea de comandos que se ejecuta en el sistema operativo Linux o Windows OS. Se dedica a proporcionar a los clientes servicios de migración de scripts de SQL de aplicaciones de forma simple, rápido y confiable. Analiza los scripts de SQL de las aplicaciones de base de datos de origen utilizando la lógica de migración de sintaxis integrada, y los migra para que sean aplicables a las bases de datos GaussDB(DWS).
El DSC puede migrar scripts de SQL de bases de datos de Teradata, Oracle, Netezza, MySQL y DB2.

6 Conceptos

6.1 Conceptos de gestión de GaussDB(DWS)

Clúster

Un clúster es un grupo de servidores que consta de varios nodos. GaussDB(DWS) se organiza mediante clústeres. Un clúster de almacén de datos contiene nodos del mismo variante en la misma subred. Estos nodos trabajan juntos para proporcionar servicios.

Nodo

Cada clúster de almacén de datos contiene al menos tres nodos, todos los cuales admiten almacenamiento y análisis de datos.

Variante

Debe especificar los variantes de nodo al crear un clúster de almacén de datos. Los recursos de CPU, memoria y almacenamiento varían según los variantes de nodo.

Instantánea

Puede crear instantáneas para hacer una copia de seguridad de los datos del clúster de GaussDB(DWS). Una instantánea se conserva hasta que la elimine en la consola de gestión. Las instantáneas automatizadas no se pueden eliminar manualmente. Las instantáneas ocuparán sus cuotas de OBS.

6.2 Conceptos de base de datos de GaussDB(DWS)

Base de datos

Un clúster de almacén de datos es una plataforma relacional de base de datos orientada al análisis que admite análisis en línea.

OLAP

OLAP es una función importante de los clústeres de almacén de datos. Es compatible con análisis complejos, proporciona soporte para la toma de decisiones adaptado a los resultados del análisis y ofrece resultados de consultas intuitivas.

MPP

En cada nodo del clúster de almacén de datos, la computación de memoria y los sistemas de almacenamiento en disco son independientes entre sí. Con MPP, GaussDB(DWS) distribuye los datos de servicio a diferentes nodos basándose en el modelo de base de datos y las características de la aplicación. Los nodos se conectan a través de la red y procesan en colaboración tareas informáticas como un clúster y proporcionan servicios de base de datos que satisfacen las necesidades de servicio.

Arquitectura de Shared-Nothing

La arquitectura de shared-nothing es una arquitectura de computación distribuida. Cada nodo es independiente para que los nodos no compitan por recursos, lo que mejora la eficiencia del trabajo.

Versión de la base de datos

Cada clúster de almacén de datos tiene una versión de base de datos específica. Puede comprobar la versión al crear un clúster de almacén de datos.

Conexión de bases de datos

Puede utilizar un cliente para conectarse a un clúster de GaussDB(DWS) en Huawei Cloud y a través de Internet.

Usuario de la base de datos

Puede agregar y controlar usuarios quien pueden acceder a la base de datos de un clúster de almacén de datos mediante la asignación de permisos específicos a ellos. El administrador de base de datos generado al crear un clúster es el usuario predeterminado de base de datos.

7 Servicios relacionados

IAM

GaussDB(DWS) utiliza Identity and Access Management (IAM) para la autenticación y autorización.

Los usuarios que tienen los permisos de **DWS Administrator** pueden utilizar completamente GaussDB(DWS). Para obtener los permisos, póngase en contacto con un usuario con los permisos de **Security Administrator** o cree directamente un usuario con los permisos de **DWS Administrator**. Los usuarios a los que se conceden los permisos de acceso a la **DWS Database Access** pueden generar credenciales temporales de usuario de base de datos basadas en usuarios de IAM para conectarse a bases de datos en los clústeres de almacén de datos.

ECS

GaussDB(DWS) utiliza un ECS como nodo de clúster.

BMS

GaussDB(DWS) utiliza un BMS como nodo de clúster.

VPC

GaussDB(DWS) utiliza el servicio Virtual Private Cloud (VPC) para proporcionar una topología de red para que los clústeres aíslen los clústeres y controlen el acceso.

OBS

GaussDB(DWS) utiliza OBS para convertir datos de clústeres y datos externos, satisfaciendo los requisitos de almacenamiento seguro, confiable y rentable.

MRS

Los datos pueden migrarse de MRS a clústeres de GaussDB(DWS) para su análisis después de que Hadoop procese los datos.

CDM

GaussDB(DWS) utiliza CDM para migrar datos desde múltiples fuentes a GaussDB(DWS).

DIS

Puede utilizar Data Ingestion Service (DIS) para ingerir datos de flujo a GaussDB(DWS) en tiempo real.

Cloud Eye

GaussDB(DWS) utiliza Cloud Eye para monitorear las métricas de rendimiento del clúster, entregando información de estado de una manera concisa y eficiente. Cloud Eye admite la personalización de alarmas para que se le notifique de la excepción al instante.

CTS

GaussDB(DWS) utiliza Cloud Trace Service (CTS) para auditar sus operaciones no consultadas en la consola de gestión para asegurarse de que no se realizan operaciones no válidas o no autorizadas, lo que mejora la gestión de la seguridad del servicio.

SMN

GaussDB(DWS) utiliza SMN para enviar activamente mensajes de notificación de acuerdo con sus requisitos de suscripción a eventos, para que pueda recibir inmediatamente una notificación cuando ocurra un evento (por ejemplo, una operación de clúster de claves).

TMS

Con Tag Management Service (TMS), GaussDB(DWS) puede proporcionar funciones centralizadas de gestión de etiquetas y clasificación de recursos entre regiones y servicios. Puede personalizar las etiquetas para clasificar y localizar recursos.

DNS

GaussDB(DWS) utiliza el Domain Name Service (DNS) para proporcionar las direcciones de IP del clúster asignadas desde los nombres de dominio.

8 GaussDB(DWS) Gestión de permisos

Si necesita asignar diferentes permisos a los empleados de su empresa para acceder a sus Huawei Cloud Recursos de GaussDB(DWS), IAM es una buena opción para la gestión detallada de permisos. IAM proporciona autenticación de identidad, gestión de permisos y control de acceso, lo que le ayuda a proteger el acceso a sus Huawei Cloud recursos.

Con IAM, puede usar su Huawei Cloud cuenta para crear usuarios de IAM para sus empleados, y asignar permisos a los usuarios para controlar su acceso a variantes de recursos específicos. Por ejemplo, algunos desarrolladores de software de su empresa necesitan usar recursos de GaussDB(DWS), pero no deben eliminarlos ni realizar operaciones de alto riesgo. Con este fin, puede crear usuarios de IAM para los desarrolladores de software y concederles solo los permisos necesarios para usar los recursos de GaussDB(DWS).

Si su Huawei Cloud cuenta no necesita usuarios individuales de IAM para la gestión de permisos, puede omitir esta sección.

IAM se puede utilizar de forma gratuita. Solo paga por los recursos de tu cuenta. Para obtener más información acerca de IAM, consulte [Descripción general del servicio](#).

Políticas de sistema compatibles

De forma predeterminada, los nuevos usuarios de IAM no tienen permisos asignados. Debe agregar un usuario a uno o más grupos, y adjuntar políticas o roles de permisos a estos grupos. Los usuarios heredan permisos de los grupos a los que se agregan y pueden realizar operaciones específicas en servicios en la nube.

GaussDB(DWS) es un servicio a nivel de proyecto implementado y accedido en regiones físicas específicas. Para asignar permisos de GaussDB(DWS) a un grupo de usuarios, especifique el ámbito como proyectos específicos de región y seleccione proyectos para que los permisos surtan efecto. Si se selecciona **All projects**, los permisos surtirán efecto para el grupo de usuarios en todos los proyectos específicos de la región. Al acceder a GaussDB(DWS), los usuarios necesitan cambiar a una región donde han sido autorizados para usar GaussDB(DWS).

- **Función:** IAM proporciona inicialmente un mecanismo de autorización de grano grueso para definir permisos basados en las responsabilidades del trabajo de los usuarios'. Este mecanismo proporciona solo un número limitado de roles de nivel de servicio para la autorización. Al utilizar roles para conceder permisos, también debe asignar otros roles de los que dependen los permisos para que surtan efecto. Sin embargo, los roles no son una opción ideal para la autorización detallada y el control de acceso seguro.

- **Políticas:** Un variante de mecanismo de autorización detallado que define los permisos necesarios para realizar operaciones en recursos de nube específicos bajo ciertas condiciones. Este mecanismo permite una autorización más flexible basada en políticas, cumpliendo los requisitos para un control de acceso seguro. Por ejemplo, puede conceder a los usuarios de GaussDB(DWS) solo los permisos para administrar un cierto variante de recursos de GaussDB(DWS).

La mayoría de las políticas definen permisos basados en APIs. Para ver las acciones de API admitidas por GaussDB(DWS), consulte [Políticas de permisos y acciones admitidas](#).

Para obtener más información sobre cómo crear una política de permisos detallada, consulte [Creación de una política personalizado de GaussDB\(DWS\)](#)

Tabla 8-1 lista todas las funciones y políticas definidas por el sistema soportadas por GaussDB(DWS).

Tabla 8-1 GaussDB(DWS) permisos del sistema

Nombre de rol/política	Descripción	Categoría	Dependencias
DWS ReadOnlyAccess	Permisos de sólo lectura para GaussDB(DWS). Los usuarios con estos permisos solo pueden ver los datos de GaussDB(DWS).	Política definida por el sistema	N/A
DWS FullAccess	Permisos de administrador para GaussDB(DWS). Los usuarios con estos permisos pueden realizar todas las operaciones en GaussDB(DWS).	Política definida por el sistema	N/A
DWS Administrator	<p>Permisos de administrador para GaussDB(DWS). Los usuarios con estos permisos pueden realizar operaciones en todos los recursos de GaussDB(DWS).</p> <ul style="list-style-type: none"> ● Los usuarios con permisos de la política de VPC Administrator pueden crear VPC y subredes. ● Los usuarios con permisos de la política de Cloud Eye Administrator pueden consultar la información de monitoreo de los clústeres de almacén de datos. 	Rol definido por el sistema	Depende de las políticas del Tenant Guest y Server Administrator , que deben asignarse en el mismo proyecto que la política de DWS Administrator .

Nombre de rol/política	Descripción	Categoría	Dependencias
DWS Database Access	Permiso de acceso a la base de datos de GaussDB(DWS). Los usuarios con este permiso pueden generar las credenciales temporales de usuario de base de datos basadas en usuarios de IAM para conectarse a la base de datos en el clúster de almacén de datos.	Rol definido por el sistema	Depende de la política de DWS Administrator , que debe asignarse en el mismo proyecto que la política de DWS Database Access .

Tabla 8-2 enumera las operaciones comunes admitidas por cada política o rol de GaussDB(DWS) definido por el sistema. Elija las políticas o roles adecuados según sea necesario.

 **NOTA**

- Si utiliza el EIP por primera vez para un proyecto en una región, el sistema le pedirá que cree la agencia de **DWSAccessVPC** para autorizar a GaussDB(DWS) a acceder a VPC. Después de que la autorización es exitosa, GaussDB(DWS) puede cambiar a una VM sana cuando la VM enlazada con el EIP es defectuosa.
- Solo las cuentas o usuarios de nube pública con permisos de **Security Administrator** pueden crear agencias de forma predeterminada. De forma predeterminada, los usuarios de IAM de esas cuentas no pueden crear agencias. Cuando los usuarios utilizan el EIP, el sistema muestra un mensaje indicando permisos insuficientes. Póngase en contacto con un usuario con los permisos de **DWS FullAccess** para autorizar a la agencia en la página actual.

Tabla 8-2 Operaciones comunes soportadas por cada política o rol de GaussDB(DWS) definida por el sistema

Operación	DWS FullAccess	DWS ReadOnlyAccess	DWS Administrator	DWS Database Access
Creación/restauración de clústeres	√	x	√	x
Obtención de la lista de clústeres	√	√	√	√

Operación	DWS FullAccess	DWS ReadOnlyAccess	DWS Administrator	DWS Database Access
Obtención de los detalles de un clúster	√	√	√	√
Configuración de la política de instantáneas automatizada	√	x	√	x
Configuración de parámetros/grupos de parámetros de seguridad	√	x	√	x
Reiniciar clústeres	√	x	√	x
Escalado de clústeres	√	x	√	x
Restablecer contraseñas	√	x	√	x
Aplicación de plantillas de parámetros a clústeres	√	x	√	x
Eliminación de clústeres	√	x	√	x
Configuración de ventanas de mantenimiento	√	x	√	x
Vinculación de EIPs	√	x	√	x
Desvinculación de EIPs	√	x	√	X
Creación de nombres de dominio de DNS	√	x	√	x

Operación	DWS FullAccess	DWS ReadOnlyAccess	DWS Administrator	DWS Database Access
Liberación de nombres de dominio de DNS	√	x	√	x
Modificación de nombres de dominio de DNS	√	x	√	x
Creación de conexiones de MRS	√	x	√	x
Actualización de conexiones de MRS	√	x	√	x
Eliminación de conexiones de MRS	√	x	√	x
Adición/ Eliminación de etiquetas	√	x	√	x
Edición de etiquetas	√	x	√	x
Creación de instantáneas	√	x	√	x
Obtención de la lista de instantáneas	√	√	√	√
Eliminación de instantáneas	√	x	√	x
Copia de instantáneas	√	x	√	x
Creación de plantillas de parámetros	√	x	√	x

Operación	DWS FullAccess	DWS ReadOnlyAccess	DWS Administrator	DWS Database Access
Eliminación de plantillas de parámetros	√	x	√	x
Modificación de plantillas de parámetros	√	x	√	x

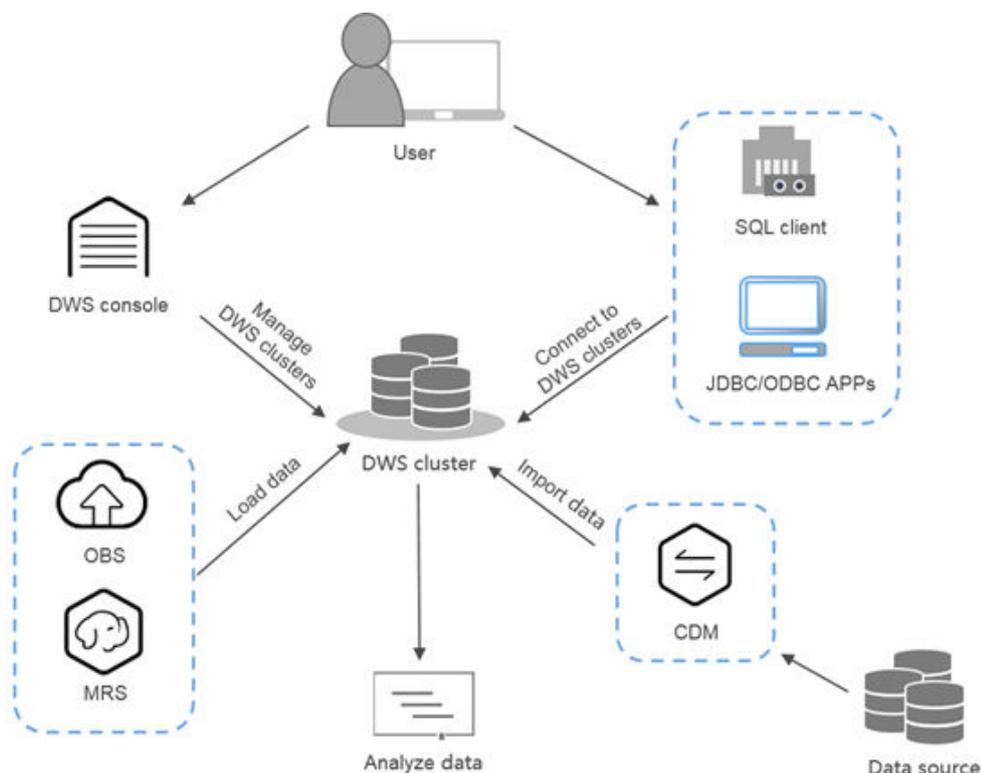
Enlaces útiles

- [Descripción general del servicio de IAM](#)
- [Creación de un usuario y concesión de permisos de GaussDB\(DWS\)](#)
- [Políticas de permisos y acciones admitidas](#)

9 Acceso de GaussDB(DWS)

La siguiente figura muestra cómo usar GaussDB(DWS).

Figura 9-1 Proceso para usar GaussDB(DWS)



Acceso a un clúster

GaussDB(DWS) proporciona una consola de gestión basada en web y APIs compatibles con HTTPS para que pueda gestionar clústeres de almacén de datos.

Acceso a la base de datos en un clúster

GaussDB(DWS) admite el acceso a la base de datos utilizando los siguientes métodos:

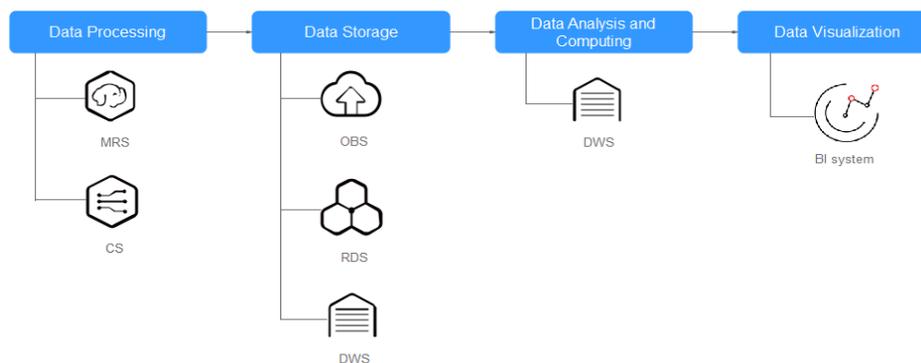
- Clientes de GaussDB(DWS)
Acceda a la base de datos de clúster mediante clientes GaussDB(DWS). Para obtener más información, consulte [Uso del cliente de gsql para conectarse a un clúster](#) y [Uso del cliente de GUI de Data Studio para conectarse a un clúster](#).
- Clientes de PostgreSQL de código abierto
A continuación se enumeran los clientes de código abierto compatibles:
 - PostgreSQL psql 9.2.4 o posterior
Para obtener más información, consulte <https://www.postgresql.org/>.
 - pgadmin
Para obtener más información, consulte <https://www.pgadmin.org/>.
 - dbeaver
Para obtener más información, consulte <https://dbeaver.jkiss.org/download/>.
- Llamadas de API a JDBC y ODBC
Puede llamar a las estándares de APIs, como JDBC y ODBC, para acceder a las bases de datos en clústeres.
Para obtener más información, consulte [Uso de un controlador de JDBC para conectarse a una base de datos](#) y [Uso de un controlador de ODBC para conectarse a una base de datos](#).

Proceso de análisis de datos de extremo a extremo

GaussDB(DWS) se ha integrado perfectamente con otros servicios en el Huawei Cloud, lo que le ayuda a implementar rápidamente soluciones de análisis de datos de extremo a extremo.

La siguiente figura muestra el proceso de análisis de datos de extremo a extremo. También se muestran los servicios en uso durante cada proceso.

Figura 9-2 Proceso de análisis de datos de extremo a extremo



10 Detalles de precios

GaussDB(DWS) admite dos modos de facturación: pago por uso y paquetes de descuento (más económicos). El precio total de un clúster de GaussDB(DWS) se calculará automáticamente para que pueda comprar un clúster con solo unos pocos clics.

Artículos facturables

Para usar Huawei Cloud GaussDB(DWS), solo tiene que pagar por el variante de nodo de almacén de datos que elija y los recursos correspondientes que utilice. Para obtener más detalles de precios de GaussDB(DWS), consulte los [Detalles de precios de GaussDB\(DWS\)](#). Puede utilizar la calculadora de precios de GaussDB(DWS) para obtener rápidamente un precio estimado de un cluster con el variante que elija.

Tabla 10-1 GaussDB(DWS) artículos facturables

Artículo	Descripción
Nodo de almacén de datos	<p>Paga por el variante de nodo que elija. Para obtener más información, consulte los Detalles de precios de GaussDB(DWS). Hay dos modos de facturación disponibles para los variantes de nodo:</p> <ul style="list-style-type: none">● Pago por uso (por hora)● Paquete con descuento <p>Para obtener más información sobre los modos de facturación anteriores, consulte Modos de facturación.</p>
Espacio de almacenamiento de instantáneas	<p>GaussDB(DWS) proporciona espacio de almacenamiento gratuito para que almacene los datos de las instantáneas. Sin embargo, si utiliza más espacio que el espacio de almacenamiento gratuito, la parte excedida se cobra de acuerdo con las reglas de facturación de OBS. Para obtener más información, consulte los Detalles de precios de OBS.</p> <p>El espacio gratuito es el mismo que el tamaño del espacio de almacenamiento total del clúster. (Espacio gratuito = Espacio de almacenamiento de un solo nodo x Número de nodos)</p>

Artículo	Descripción
Almacenamiento de datos en la nube	Puede personalizar el variante de almacenamiento y la capacidad que utiliza el clúster. Los recursos de almacenamiento se cobran por separado. Para obtener más información, consulte los Detalles de precios de GaussDB(DWS) .
(Opcional) EIP y banda ancha	<ul style="list-style-type: none"> ● Vincular un EIP a un clúster de GaussDB(DWS). El EIP se factura según las reglas de precios del EIP de VPC. Para obtener más información, consulte los Detalles de precios de EIP. ● No se le factura el tráfico generado por el clúster GaussDB(DWS) en la red de Huawei Cloud.
(Opcional) Clave de cifrado de la base de datos (DEK)	Si la función de Encrypt DataStore está habilitada al crear un clúster de GaussDB(DWS), se necesitan claves del Servicio de administración de claves (KMS) para cifrar y descifrar la base de datos. Las llaves se facturan según las reglas de facturación DEW. Para obtener más información, consulte los Detalles de precios de DEW .

Modos de facturación

GaussDB(DWS) ofrece dos modos de facturación: pago por uso y paquetes de descuento. Cuanto más tiempo use, menos pagará.

- **Pago por uso (por hora):** En este modo de facturación, puede activar o desactivar GaussDB(DWS) como desee. Se le factura por su duración de uso por hora (el tiempo es exacto a minutos) y el sistema genera una factura cada hora natural y deduce tarifa del saldo de su cuenta. Este modo es aplicable a las pruebas operacionales preliminares y a la prueba de concepto, y al uso a corto plazo.
- **Paquete de descuento:** Puede elegir este modo y hacer un pago único si planea utilizar este servicio durante mucho tiempo. Este modo le ofrece descuentos significativos sobre el modo de pago por uso y le ahorra entre un 39% y un 51% de los costos. Este modo es aplicable al uso a largo plazo después de pruebas y verificaciones.

Usted puede considerar el paquete de descuento como un cupón. El sistema no creará automáticamente un clúster cuando compre un paquete de descuento. Cuando utiliza un clúster, el sistema de facturación compara automáticamente el número de nodos con descuento válidos que compra con el número de nodos en uso de la misma región y variante. Los nodos con descuento no se cobrarán de nuevo y los nodos adicionales se cobrarán sobre una base de pago por uso. Después de que el paquete de descuento caduque, se le facturará en modo de pago por uso si continúa utilizando el servicio.

Los nodos de un paquete de descuento no están vinculados a un clúster:

- Después de comprar un paquete de descuento, el sistema no creará un clúster automáticamente. Si no ha ejecutado ningún clúster, vaya a la consola de gestión de GaussDB(DWS) para crear un clúster con la región, el variante de nodo y la cantidad de nodo correspondientes.
- Si el clúster se factura en modo de pago por uso, el sistema de facturación identifica automáticamente los nodos con descuento de la región, el variante y la cantidad correspondientes y le ofrece descuentos.
- Si desea escalar su clúster y tiene nodos con descuento válidos en la región, el sistema de facturación identifica automáticamente los nodos con descuento y le

- ofrece descuentos. Si no tiene ningún nodo con descuento válido, los nodos recién agregados se facturarán según una base de pago por uso.
- d. Si elimina un clúster existente y crea uno de la misma región y variante, los nodos con descuento aún se pueden utilizar en el nuevo clúster.

AVISO

Preste atención a lo siguiente cuando compre un paquete de descuento:

- Los nodos descontados están enlazados a una región específica.
- Los nodos con descuento están enlazados a un variante de nodo específico. Los nodos de diferentes variantes tienen diferentes precios de paquete.
- Cuando caduca un paquete de descuento o se da de baja del paquete, el sistema le cobrará automáticamente por pago por uso (por hora). El servicio no se interrumpirá mientras el saldo de su cuenta sea suficiente. Por ejemplo, si compra un paquete con tres nodos el 1 de enero al año y la duración del paquete es de un año, entonces el paquete finaliza automáticamente al final del 31 de diciembre del año. Tenga en cuenta que si no ejecuta el clúster dentro del período de validez, el período de validez no se extenderá y la tarifa no se podrá reembolsar. Por lo tanto, para ahorrar costes, se recomienda crear un clúster inmediatamente después de comprar un paquete de descuento o comprar un paquete de descuento después de crear un clúster.
- Un paquete de descuento se paga por adelantado por mes o año. Si planea usar GaussDB(DWS) durante un largo plazo, los nodos de un paquete de descuento son más rentables que los nodos facturados sobre una base de pago por uso.
- Al comprar un clúster por primera vez, se recomienda comprar al menos tres nodos (la escala mínima de un clúster es de tres nodos) y seleccionar el variante de nodo según sea necesario.
- Un paquete de descuento es un concepto de facturación. Después de comprar un paquete de descuento, no se creará automáticamente ningún clúster. Vaya a la consola de gestión de GaussDB(DWS) para [crear un clúster](#) si no existe ningún clúster.
- El proyecto de empresa del paquete de descuento debe ser el mismo que el del clúster. **Todos los proyectos** se aplican a cualquier clúster.

Cambio de configuraciones

GaussDB(DWS) proporciona múltiples variantes de nodo. Después de crear un clúster, puede agregar nodos para escalar el clúster a medida que aumenta el volumen de servicio. Para obtener más información, consulte [Escala horizontal de clústeres](#).

Si el método de cambio de configuración de escalamiento horizontal proporcionado por GaussDB(DWS) no cumple con sus requisitos, puede crear un clúster de nuevo y migrar datos al clúster para realizar el cambio de configuración del clúster.

Renovación

GaussDB(DWS) ofrece dos modos de facturación, pago por uso y paquetes de descuento. En el modo de pago por uso, las tarifas se deducen cada hora y un saldo insuficiente puede conducir a pagos atrasados. En el modo de paquete de descuento, los clústeres deben renovarse antes de que caduque el paquete. De lo contrario, se le asignará un período de retención a sus recursos. Durante este período, los datos se conservarán, pero los clústeres se detendrán.

Para renovar su paquete, visite [Renovaciones](#).

Vencimiento

- El vencimiento no se aplica a los clústeres de pago por uso.
- Después de que el paquete de descuento caduque, se le facturará en modo de pago por uso si continúa utilizando el servicio.

Atrasado

- Un grupo de descuentos no tiene el concepto de atrasos.
- En el modo de pago por uso, las tarifas de clúster se deducen cada hora. Si el saldo de su cuenta es insuficiente para pagar los gastos incurridos en la última hora, su cuenta estará en mora. Si los clústeres se renuevan dentro del período de retención, estarán disponibles y se cobrarán a partir de la fecha de vencimiento original.

11 Restricciones

- Sólo puede administrar clústeres y no puede acceder directamente a los nodos de un clúster. Puede utilizar la dirección IP y el puerto de un clúster para acceder a la base de datos en el clúster.
- No se puede cambiar el variante de un clúster existente. Si necesita nodos con un variante más alto, cree uno nuevo.
- Si utiliza un cliente para conectarse a un clúster, su subred de VPC debe ser la misma que la del clúster.
- Si copia comandos del documento al entorno operativo, el texto se ajusta automáticamente, lo que provoca errores en la ejecución de los comandos. Para resolver el problema, elimine el salto de línea.

12 Cuotas de servicio

Se aplican cuotas para los recursos de servicio en la plataforma para evitar picos imprevistos en el uso de recursos. Las cuotas limitan el número o la cantidad de recursos disponibles para los usuarios.

Tabla 12-1 muestra las cuotas de usuario predeterminadas de GaussDB(DWS). Para obtener más información acerca de cómo consultar y aumentar las cuotas, consulte [Quotas](#).

Tabla 12-1 Cuotas de servicio

Tipo de recurso	Cuota total
Nodos	32

13 Especificaciones técnicas

Esta sección describe las especificaciones técnicas de GaussDB(DWS) en diferentes versiones.

Tabla 13-1 GaussDB(DWS) especificaciones técnicas

Especificaciones técnicas	Valor máximo de 8.0.x	Valor máximo de 8.1.0	Valor máximo de 8.1.1
Capacidad de datos	10 PB	10 PB	20 PB
Número de nodos de clúster	256	256	2048
Tamaño de una sola tabla	1 PB	1 PB	1 PB
Tamaño de los datos en cada fila	1 GB	1 GB	1 GB
Tamaño de una sola columna en cada registro	1 GB	1 GB	1 GB
Número de registros en cada tabla	2 ⁵⁵	2 ⁵⁵	2 ⁵⁵
Número de columnas en cada tabla	1600	1600	1600
Número de índices en cada tabla	Unlimited	Unlimited	Unlimited
Número de columnas en el índice de cada tabla	32	32	32
Número de restricciones en cada tabla	Unlimited	Unlimited	Unlimited

Especificaciones técnicas	Valor máximo de 8.0.x	Valor máximo de 8.1.0	Valor máximo de 8.1.1
Cantidad de conexiones simultáneas	60 para transacciones largas analíticas; 5,000 para transacciones cortas	60 para transacciones largas analíticas; 5,000 para transacciones cortas	80 para transacciones largas analíticas; 5,000 para transacciones cortas
Número de particiones en una tabla particionada	32,768	32768	32768
Tamaño de cada partición en una tabla particionada	1 PB	1 PB	1 PB
Número de registros en cada partición de una tabla particionada	2 ⁵⁵	2 ⁵⁵	2 ⁵⁵

14 Descripción de la versión

Esta sección describe los registros de actualización de la versión del clúster de GaussDB(DWS).

Versión V8.1.1.205

Fecha de lanzamiento: 30 de marzo de 2022

Versión del núcleo: 8.1.1

1. Se resuelven los problemas clave en la red en vivo. No hay actualizaciones de documentos involucradas.

Versión V8.1.1.203

Fecha de lanzamiento: 18 de marzo de 2022

Versión del núcleo: 8.1.1

1. La función de expansión de la capacidad del disco se agrega a la página de la consola. Puede ampliar la capacidad del disco para eliminar rápidamente los cuellos de botella de los recursos de almacenamiento.
2. El almacén de datos híbrido admite la implementación distribuida y ofrece las mismas capacidades que el almacén de datos en la nube. Las capacidades de un almacén de datos independiente no cambian.

Versión V8.1.1.202

Fecha de lanzamiento: 24 de febrero de 2022

Versión del núcleo: 8.1.1

1. Se agregó el tipo de clúster de almacén de datos **Hybrid**, aplicable al procesamiento analítico/transacciones híbridas (HTAP). Una base de datos se puede utilizar tanto para la producción como para el análisis.
2. Se agregó la función switchback a la consola. Puede realizar un switchback primario/de reserva para un clúster en estado **Unbalanced** para restaurar el rendimiento del clúster.
3. El almacén de datos de IoT puede vincularse a Elastic Load Balance (ELB) para evitar fallos de un solo CN en un almacén de datos de IoT.

4. Se ha ampliado el rango de parámetros de GUC admitidos en la consola. Los nuevos parámetros de GUC se pueden modificar en la consola.

Versión V8.1.1.200

Fecha de lanzamiento: 9 de diciembre de 2021

Versión del núcleo: 8.1.1

1. Se agregó la función de monitoreo de bases de datos para admitir monitoreo detallado en nodos, rendimiento, bases de datos, sesiones, consultas, instancias y métricas de carga.
2. Se agregó la función de gestión de alarmas para permitir a los usuarios ver las alarmas activadas por DMS en la consola.
3. Se agregó la función de escalado horizontal basada en fases. Las fases incluyen la adición de nodos, el escalamiento horizontal del clúster y la redistribución de la base de datos.
4. Se pueden configurar más parámetros de GUC en la consola.
5. Los clústeres se pueden enlazar a Elastic Load Balance (ELB).

Versión V8.1.1.100

Fecha de lanzamiento: 30 de julio de 2021

Versión del núcleo: 8.1.1

1. Se agregó la función de gestión de CNs. El número máximo de CNs en un clúster ha cambiado de 5 a 20. Los CNs se pueden agregar o eliminar en la consola de gestión.
2. Se agregó la función de escalado horizontal en línea. Durante el escalamiento horizontal del clúster, los datos se pueden importar continuamente a la base de datos.
3. Operaciones de misión crítica protegidas. Se requiere una credencial para la verificación de identidad si los usuarios desean realizar una operación de misión crítica en la consola de gestión.
4. Se agregó la función DR de doble clúster. Un clúster de DR GaussDB(DWS) homogéneo se puede implementar en otro AZ para garantizar la continuidad del servicio.
5. Lanzamiento del núcleo de base de datos 8.1.1. Para obtener más información, consulte [Novedades](#).

V8.1.0.100

Fecha de lanzamiento: 15 de mayo de 2021

Versión del núcleo: 8.1.0

1. Se agregó la función de clúster lógico. Las tablas de una base de datos se pueden asignar a diferentes nodos lógicos basados en el clúster lógico. Un clúster lógico puede contener tablas de varias bases de datos.
2. Se cambió el nombre predeterminado de la base de datos a **gaussdb** durante la creación del clúster.
3. Lanzamiento del núcleo de base de datos 8.1.0. Para obtener más información, consulte [Novedades](#).

V8.0.1.100

Fecha de lanzamiento: 31 de julio de 2020

Versión del núcleo: 8.0.0.1

1. Se agregó el tipo de clúster de almacén de datos de **Stream** para proporcionar servicios de computación y compresión de series temporales.
2. Se agregó la función de gestión de cargas de trabajo para que se puedan ejecutar dos tipos de trabajos al mismo tiempo sin afectarse entre sí.
3. Se agregaron las herramientas de exportación de datos **gs_dump** y **gs_dumpall**.

V1.7.1

Fecha de lanzamiento: 20 de marzo de 2020

1. Se agregaron las siguientes funciones para el núcleo de la base de datos:
 - UPDATE ahora admite la referencia de alias y nombre de tabla.
 - Puede especificar un schema al crear un índice. El schema especificado debe ser el mismo que el que reside la tabla base.
 - Se agregó la sintaxis de **alter table modify null/not null**.
 - Se agregó el soporte para la operación de **alter sequence maxvalue**.
 - Se limitó el espacio de tabla temporal y el espacio de vaciado del disco del operador.
 - Se eliminaron los registros de subescritura del TOP SQL
 - Se agregó el soporte para la grabación de sentencias CN livianas.
 - Se procesó los datos de antigüedad en las tablas de información de top SQL histórico.
 - Se soportó procedimientos almacenados y sentencias de función de CN retry.

V1.5.200

Fecha de lanzamiento: 8 de agosto de 2019

- Funciones optimizadas
 - a. Aumento de la frecuencia de copia de seguridad de instantáneas automatizadas. De forma predeterminada, las instantáneas automatizadas se crean tres veces al día, a partir de las 00:00 (UTC). Puede establecer el período de ejecución de instantáneas al crear un clúster o modificar la política de instantáneas automatizada existente para un clúster.
 - b. En la lista de clústeres, puede hacer clic en **More** en la columna **Operation** de un clúster para seleccionar la operación de **Create Snapshot** para crear manualmente una instantánea.