

# Data Warehouse Service

## Quick Start

Edición 01  
Fecha 2026-09-29



**Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2026. 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.

## **Huawei Cloud Computing Technologies Co., Ltd.**

Dirección: Huawei Cloud Data Center Jiaoxinggong Road  
Avenida Qianzhong  
Nuevo distrito de Gui'an  
Gui Zhou, 550029  
República Popular China

Sitio web: <https://www.huaweicloud.com/intl/es-us/>

---

# Índice

---

1 Uso de DWS para consultar rutas de vehículos en puntos de control de tráfico en segundos.....1

2 Operaciones de SQL básicas..... 5

3 Pasos iniciales con prácticas comunes.....7

# 1 Uso de DWS para consultar rutas de vehículos en puntos de control de tráfico en segundos

Esta práctica muestra cómo analizar los vehículos que pasan por los puntos de control. En esta práctica, 890 millones de registros de datos de los puntos de control se cargan en una única tabla de base de datos en DWS para realizar consultas precisas y difusas, lo que demuestra la capacidad de DWS para realizar consultas de alto rendimiento de datos históricos.

## NOTA

Los datos de ejemplo se han cargado en la carpeta **traffic-data** de un bucket de OBS, y todas las cuentas Huawei Cloud han recibido el permiso de solo lectura para acceder al bucket de OBS.

## Procedimiento general

Esta práctica dura unos 40 minutos. El proceso básico es el siguiente:

1. [Hacer preparaciones](#)
2. [Paso 1: Creación de un clúster](#)
3. [Paso 2: Importación de datos de muestra](#)
4. [Paso 3: Realización del análisis del vehículo](#)

## Regiones admitidas

La [Tabla 1-1](#) describe las regiones donde se han cargado los datos de OBS.

**Tabla 1-1** Regiones y nombres de bucket de OBS

| Región            | Bucket de OBS       |
|-------------------|---------------------|
| CN North-Beijing1 | dws-demo-cn-north-1 |
| CN North-Beijing2 | dws-demo-cn-north-2 |
| CN North-Beijing4 | dws-demo-cn-north-4 |
| CN North-Ulanqab1 | dws-demo-cn-north-9 |

| Región                            | Bucket de OBS           |
|-----------------------------------|-------------------------|
| CN East-Shanghai1                 | dws-demo-cn-east-3      |
| CN East-Shanghai2                 | dws-demo-cn-east-2      |
| CN South-Guangzhou                | dws-demo-cn-south-1     |
| CN South-Guangzhou-InvitationOnly | dws-demo-cn-south-4     |
| CN-Hong Kong                      | dws-demo-ap-southeast-1 |
| AP-Singapore                      | dws-demo-ap-southeast-3 |
| AP-Bangkok                        | dws-demo-ap-southeast-2 |
| LA-Santiago                       | dws-demo-la-south-2     |
| AF-Johannesburg                   | dws-demo-af-south-1     |
| LA-Mexico City1                   | dws-demo-na-mexico-1    |
| LA-Mexico City2                   | dws-demo-la-north-2     |
| RU-Moscow2                        | dws-demo-ru-northwest-2 |
| LA-Sao Paulo1                     | dws-demo-sa-brazil-1    |

## Hacer preparaciones

- Ha registrado una cuenta de DWS y ha verificado el estado de la cuenta antes de usar DWS. La cuenta no puede estar en mora o congelada.
- Obtenga la AK/SK de la cuenta consultando [Creación de una clave de acceso](#).

## Paso 1: Creación de un clúster

Cree un clúster de DWS consultando [Creación de un clúster](#). Seleccione **CN-Hong Kong** para **Region**. Seleccione **Coupled storage and compute** para **Version**. **Esta práctica es para la puesta en marcha del servicio**. Puede optar por no configurar una EIP ni un balanceador de carga.

### NOTA

**CN-Hong Kong** se utiliza como ejemplo. Puede seleccionar otra región según sea necesario. Asegúrese de que todas las operaciones se realicen en la misma región.

## Paso 2: Importación de datos de muestra

Después de conectarse al clúster mediante la herramienta cliente SQL, realice las siguientes operaciones en la herramienta cliente SQL para importar los datos de muestra de los puntos de control de tráfico y realizar consultas de datos.

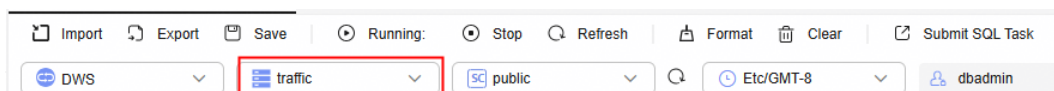
- Paso 1** En la puesta en marcha del servicio, se recomienda utilizar el **SQL Editor** para conectarse al clúster. Para obtener más información, consulte [Uso del SQL Editor para conectarse a un clúster](#).

También puede utilizar otros métodos (como la herramienta de línea de comandos gsql) para conectarse al clúster. Para obtener más información, consulte [Conexión a un clúster de DWS](#).

**Paso 2** En la ventana SQL del editor SQL, ejecute la siguiente sentencia para crear la base de datos **traffic**:

```
CREATE DATABASE traffic encoding 'utf8' template template0;
```

**Paso 3** En la página del editor de SQL, cambie a la nueva base de datos **traffic** en la parte superior.



**Paso 4** En la nueva base de datos, ejecute la siguiente sentencia para crear una tabla:

```
DROP TABLE if exists GCJL;
CREATE TABLE GCJL
(
    kkbh    VARCHAR(20),
    hphm    VARCHAR(20),
    gcsj    DATE,
    cplx    VARCHAR(8),
    cllx    VARCHAR(8),
    csys    VARCHAR(8)
)
with (orientation = column, COMPRESSION=MIDDLE)
distribute by hash(hphm);
```

**Paso 5** Cree una tabla externa, que se utiliza para identificar y asociar los datos de origen en OBS.

#### ATENCIÓN

- `<obs_bucket_name>` indica el nombre del bucket de OBS correspondiente a la región donde se encuentra DWS. Para obtener más información sobre las regiones admitidas, consulte [Regiones admitidas](#). El bucket de OBS y los datos de ejemplo se han preestablecido en el sistema y no es necesario crearlos. En este manual, se utiliza la región **CN-Hong Kong** y se establece `<obs_bucket_name>` en **dws-demo-ap-southeast-1**. No se puede acceder a los datos de bucket de OBS entre regiones. Por ejemplo, si el clúster se encuentra en **CN-Hong Kong**, no puede establecer `<obs_bucket_name>` en el nombre de bucket de otra región.
- Reemplace `<Access_Key_Id>` y `<Secret_Access_Key>` con los valores reales obtenidos en [Hacer preparaciones](#).
- El uso de AK y SK codificadas de forma rígida (hard-coded) o en texto plano representa un riesgo de seguridad. Por seguridad, encripte su AK/SK y guárdelas en el archivo de configuración o en las variables de entorno.

```
DROP FOREIGN table if exists GCJL_OBS;
CREATE FOREIGN TABLE GCJL_OBS
(
    like GCJL
)
SERVER gsmpp_server
OPTIONS (
    encoding 'utf8',
    location 'obs://<obs_bucket_name>/traffic-data/gcxx',
    format 'text',
    delimiter ',',
    access_key '<Access_Key_Id>',
```

```
secret_access_key '<Secret_Access_Key>',  
chunksize '64',  
IGNORE_EXTRA_DATA 'on'  
);
```

**Paso 6** Importe datos de una tabla externa a una tabla de base de datos.

```
INSERT INTO GCJL SELECT * FROM GCJL_OBS;
```

Se necesita algún tiempo para importar datos.

----Fin

### Paso 3: Realización del análisis del vehículo

#### 1. Ejecución de ANALYZE.

Esta sentencia recopila estadísticas relacionadas con tablas ordinarias en bases de datos. Las estadísticas se guardan en el catálogo del sistema **PG\_STATISTIC**. Cuando ejecuta el planificador, las estadísticas le ayudan a desarrollar un plan de ejecución de consultas eficiente.

Ejecute la siguiente sentencia para generar las estadísticas de la tabla:

```
ANALYZE;
```

#### 2. Consulta del volumen de datos de la tabla de datos

Ejecute la siguiente sentencia para consultar el número de registros de datos cargados:

```
SELECT count(*) FROM gcjl;
```

#### 3. Consulta precisa de vehículos

Ejecute las siguientes sentencias para consultar la ruta de conducción de un vehículo por número de placa y segmento de tiempo. DWS responde a la solicitud en segundos.

```
SELECT hphm, kkbh, gcsj  
FROM gcjl  
where hphm = 'YD38641'  
and gcsj between '2016-01-06' and '2016-01-07'  
order by gcsj desc;
```

#### 4. Consulta aproximada de vehículos

Ejecute las siguientes sentencias para consultar la ruta de conducción de un vehículo por número de placa y segmento de tiempo. DWS responde a la solicitud en segundos.

```
SELECT hphm, kkbh, gcsj  
FROM gcjl  
where hphm like 'YA23F%'  
and kkbh in('508', '1125', '2120')  
and gcsj between '2016-01-01' and '2016-01-07'  
order by hphm,gcsj desc;
```

# 2 Operaciones de SQL básicas

Esta sección describe algunas operaciones de SQL básicas de la base de datos de GaussDB(DWS).

## Creación, visualización y eliminación de una base de datos

- Ejecute la instrucción **CREATE DATABASE** para crear una base de datos.  

```
CREATE DATABASE test_db ENCODING 'UTF8' template = template0;
```
- Consulte la lista de la base de datos mediante el metacomando **\l**.  

```
\l
```
- Consulte la lista de la base de datos mediante el catálogo del sistema **PG\_DATABASE**.  

```
SELECT datname FROM pg_database;
```
- Ejecute la instrucción **DROP DATABASE** para eliminar una base de datos.  

```
DROP DATABASE test_db;
```

## Creación, visualización, modificación y eliminación de tablas

- Ejecute la instrucción **CREATE TABLE** para crear una tabla.  

```
CREATE TABLE customer_t1(id INT, name CHAR(40),age TINYINT);
```
- Utilice la función **PG\_GET\_TABLEDEF()** para ver la instrucción de creación de tabla.  

```
SELECT * FROM PG_GET_TABLEDEF('customer_t1');
```
- Ejecute la instrucción **ALTER TABLE** para modificar una tabla.  
Agregue una columna:  

```
ALTER TABLE customer_t1 ADD (address VARCHAR(100));
```

  
Elimine una columna:  

```
ALTER TABLE customer_t1 DROP COLUMN address;
```

  
Modifique un tipo de columna:  

```
ALTER TABLE customer_t1 MODIFY age INTEGER NOT NULL;
```
- Ejecute la instrucción **DROP TABLE** para eliminar una tabla.  

```
DROP TABLE customer_t1;
```

## Creación, consulta y eliminación de índices

- Ejecute la instrucción **CREATE INDEX** o **ALTER TABLE** para crear un índice común.  

```
CREATE INDEX c_id_index on customer_t1(id);  
ALTER TABLE customer_t1 ADD INDEX c_id_index (id);
```
- Utilice el catálogo del sistema **PG\_INDEXES** para ver todos los índices de una tabla.

```
SELECT * FROM pg_indexes WHERE tablename = 'customer_t1';
```

- Ejecute la instrucción **ALTER TABLE** o **DROP INDEX** para eliminar un índice.

```
DROP INDEX c_id_index;
```

```
ALTER TABLE customer_t1 DROP INDEX c_id_index;
```

## Adición, eliminación y modificación de datos de tabla

- Ejecute la instrucción **INSERT INTO** para insertar datos de tabla.

```
INSERT INTO customer_t1 VALUES(1001,'user1',22);
```

- Ejecute la instrucción **SELECT** para consultar los datos de la tabla.

```
SELECT * FROM customer_t1;
```

- Ejecute la instrucción **UPDATE** para actualizar los datos de la tabla.

```
UPDATE customer_t1 SET id = 1009 WHERE id = '1001';
```

- Utilice la instrucción **DELETE** para eliminar datos de tabla.

```
DELETE FROM customer_t1 WHERE id = '1009';
```

# 3 Pasos iniciales con prácticas comunes

Después de crear un clúster, puede utilizar los casos prácticos proporcionados por GaussDB(DWS) para sus servicios.

**Tabla 3-1** Prácticas recomendadas

| Práctica                           |                                                                | Descripción                                                                                                                                                                                                                                                                                                                                                         |
|------------------------------------|----------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Importación y exportación de datos | <b>Importación de datos de OBS a un clúster</b>                | Esta práctica demuestra cómo cargar datos de muestra a OBS e importar datos de OBS a la tabla de destino en GaussDB(DWS), lo que le ayuda a aprender rápidamente cómo importar datos de OBS a un clúster de GaussDB(DWS).<br><br>Puede importar datos en formato TXT, CSV, ORC, PARQUET, CARBONDATA o JSON desde OBS a un clúster de GaussDB(DWS) para su consulta. |
|                                    | <b>Uso de GDS para importar datos desde un servidor remoto</b> | Esta práctica demuestra cómo usar General Data Service (GDS) para importar datos desde un servidor remoto a GaussDB(DWS).<br><br>GaussDB(DWS) le permite importar datos en formato TXT, CSV o FIXED.                                                                                                                                                                |

| Práctica           |                                                                                        | Descripción                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|--------------------|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                    | <b>Exportación de datos de ORC a MRS</b>                                               | GaussDB(DWS) le permite exportar datos de ORC a MRS usando una tabla externa HDFS. Puede especificar el modo de exportación y el formato de datos de exportación en la tabla externa. Los datos se exportan desde GaussDB(DWS) en paralelo usando múltiples DN y se almacenan en HDFS. De este modo, se mejoran los resultados globales de las exportaciones.                                                                                                    |
| Migración de datos | <b>Migración de datos de Oracle a GaussDB(DWS)</b>                                     | Este tutorial muestra cómo migrar datos de Oracle a GaussDB(DWS).                                                                                                                                                                                                                                                                                                                                                                                                |
|                    | <b>Uso de trabajos de Flink DLI para escribir datos de Kafka en DWS en tiempo real</b> | <p>Esta práctica demuestra cómo utilizar trabajos de DLI Flink para sincronizar los datos de consumo de Kafka a GaussDB(DWS) en tiempo real.</p> <p>Esta práctica dura unos 90 minutos. Los servicios de nube utilizados en esta práctica incluyen Virtual Private Cloud (VPC) y las subredes, Elastic Cloud Server (ECS), Object Storage Service (OBS), Distributed Message Service (DMS) for Kafka, Data Lake Insight (DLI) y Data Warehouse Service (DWS)</p> |

| Práctica              |                                                                | Descripción                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-----------------------|----------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Tabla de optimización | <b>Prácticas de optimización de tablas</b>                     | <p>En esta práctica, aprenderá a optimizar el diseño de sus tablas. Comenzará creando tablas sin especificar su modo de almacenamiento, clave de distribución, modo de distribución o modo de compresión. Cargue los datos de prueba en estas tablas y pruebe el rendimiento del sistema. A continuación, siga excelentes prácticas para crear las tablas de nuevo utilizando nuevos modos de almacenamiento, claves de distribución, modos de distribución y modos de compresión. Vuelva a cargar los datos de prueba y el rendimiento de la prueba. Compare los dos resultados de la prueba para averiguar cómo afecta el diseño de la tabla al espacio de almacenamiento y al rendimiento de la carga y consulta de las tablas.</p> <p>Tiempo estimado: 60 minutos</p> |
| Funciones avanzadas   | <b>Mejores prácticas de gestión de datos calientes y fríos</b> | <p>En escenarios masivos de big data, con el crecimiento de los datos, el almacenamiento y el consumo de datos aumentan rápidamente. La necesidad de datos puede variar en diferentes periodos de tiempo, por lo tanto, los datos se gestionan de manera jerárquica, mejorando el rendimiento del análisis de datos y reduciendo los costes de servicio. En algunos escenarios de uso de datos, los datos pueden clasificarse en datos calientes y datos fríos accediendo a la frecuencia.</p>                                                                                                                                                                                                                                                                            |

| Práctica |                                                                           | Descripción                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|----------|---------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|          | <p><b>Mejores prácticas para la gestión automática de particiones</b></p> | <p>Para tablas de particiones cuyas columnas de particiones son tiempo, la función de gestión automática de particiones se puede agregar para crear automáticamente particiones y eliminar particiones caducadas, reduciendo los costos de mantenimiento de la tabla de particiones y mejorando el rendimiento de las consultas. Para facilitar la consulta y el mantenimiento de datos, la columna de tiempo se utiliza a menudo como la columna de partición de una tabla particionada que almacena datos relacionados con el tiempo, como información de pedidos de comercio electrónico y datos de IoT en tiempo real. Cuando los datos relacionados con el tiempo se importan a una tabla particionada, la tabla debe tener particiones de los intervalos de tiempo correspondientes. Las tablas de particiones comunes no crean automáticamente nuevas particiones ni eliminan particiones caducadas. Por lo tanto, el personal de mantenimiento necesita crear periódicamente nuevas particiones y eliminar particiones caducadas, lo que lleva a un aumento de los costos de operación.</p> <p>Abordando esto, GaussDB(DWS) introduce la función de gestión automática de particiones. Puede establecer los parámetros a nivel de tabla <b>period</b> y <b>ttd</b> para habilitar la función de gestión automática de particiones, que crea automáticamente particiones y elimina las particiones caducadas, reduciendo los costos de mantenimiento de la tabla de particiones y mejorando el rendimiento de las consultas.</p> |

| Práctica                  |                                                              | Descripción                                                                                                                                                                                                                                                                                   |
|---------------------------|--------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Gestión de bases de datos | <b>Mejores prácticas de gestión de recursos</b>              | Esta práctica demuestra cómo usar GaussDB(DWS) para la gestión de recursos, ayudando a las empresas a eliminar cuellos de botella en el rendimiento de consultas simultáneas. Los trabajos de SQL pueden ejecutarse sin problemas sin afectarse entre sí y consumir menos recursos que antes. |
|                           | <b>Excelentes prácticas para consultas de SQL</b>            | Basándonos en un gran número de mecanismos y prácticas de ejecución de SQL, podemos optimizar sentencias SQL siguiendo ciertas reglas para ejecutar sentencias SQL más rápidamente y obtener resultados correctos.                                                                            |
|                           | <b>Excelentes prácticas para consultas de sesgo de datos</b> | Esta práctica incluye los siguientes casos de sesgo de almacenamiento: <ul style="list-style-type: none"> <li>● Detección en tiempo real de sesgo de almacenamiento durante la importación de datos</li> <li>● Localización rápida de las tablas que causan sesgo de datos</li> </ul>         |
|                           | <b>Mejores prácticas de gestión de usuarios</b>              | Un clúster GaussDB(DWS) consiste principalmente en administradores de sistemas y usuarios comunes. Esta práctica describe los permisos de los administradores del sistema y los usuarios comunes y describe cómo crear usuarios y consultar información de usuario.                           |

| Práctica                     |                                                                                 | Descripción                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------------------|---------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                              | <b>Consulta de información de tablas y bases de datos</b>                       | <p>Esta práctica demuestra algunos casos básicos de consulta de base de datos:</p> <ul style="list-style-type: none"> <li>● Consulta de información sobre tablas</li> <li>● Consulta del tamaño de la tabla</li> <li>● Consulta de información de base de datos</li> <li>● Consulta del tamaño de la base de datos</li> </ul>                                                                      |
| Análisis de datos de muestra | <b>Análisis de vehículos de punto de control</b>                                | <p>Esta práctica demuestra el análisis de los vehículos pasados en los puntos de control de tráfico. En esta práctica, 890 millones piezas de datos de puntos de control de tráfico se cargan en una única tabla de base de datos en GaussDB(DWS) para realizar una consulta precisa y una consulta difusa. Es un ejemplo de consulta de alto rendimiento de datos históricos en GaussDB(DWS).</p> |
|                              | <b>Análisis de requisitos de cadena de suministro (conjunto de datos TPC-H)</b> | <p>Esta práctica describe cómo cargar el conjunto de datos de ejemplo de OBS a un clúster de almacén de datos y realizar consultas de datos. Este ejemplo comprende análisis de múltiples tablas y análisis de temas en el escenario de análisis de datos.</p>                                                                                                                                     |

| Práctica           |                                                                   | Descripción                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|--------------------|-------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                    | <b>Análisis del estado de operaciones de una tienda minorista</b> | En esta práctica, los datos empresariales diarios de cada tienda minorista se cargan desde OBS a la tabla correspondiente en el clúster de almacén de datos para resumir y consultar KPI. Estos datos incluyen la rotación de la tienda, el flujo de clientes, la clasificación de ventas mensuales, la tasa de conversión mensual del flujo de clientes, la relación precio-alquiler mensual y las ventas por unidad de área. Este ejemplo muestra la consulta y el análisis multidimensional de GaussDB(DWS) en el escenario minorista. |
| Seguridad de datos | <b>Cifrado y descifrado de columnas de datos</b>                  | La encriptación de datos es ampliamente utilizada en sistemas de información para evitar el acceso no autorizado y la fuga de datos. Como el núcleo del sistema de información, el almacén de datos de GaussDB(DWS) también proporciona funciones de encriptación de datos, incluidas las encriptación transparente y encriptación utilizando funciones SQL. En esta sección se describe la encriptación de la función de SQL.                                                                                                            |