ALIBABA CLOUD

Alibaba Cloud

数据安全中心

User Guide

Document Version: 20220217

Alibaba Cloud

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.

2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.

3. The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.

4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).

5. By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.

6. Please directly contact Alibaba Cloud for any errors of this document.

Document conventions

Style	Description	Example
Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	**Danger:** Resetting will result in the loss of user configuration data.
Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	**Warning:** Restarting will cause business interruption. About 10 minutes are required to restart an instance.
Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	**Notice:** If the weight is set to 0, the server no longer receives new requests.
Note	A note indicates supplemental instructions, best practices, tips, and other content.	**Note:** You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click **Settings**> **Network**> **Set network type**.
Bold	Bold formatting is used for buttons , menus, page names, and other UI elements.	Click **OK**.
`Courier font`	Courier font is used for commands	Run the `cd /d C:/window` command to enter the Windows system folder.
Italic	Italic formatting is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	This format is used for an optional value, where only one item can be selected.	`ipconfig [-all|-t]`
{} or {a|b}	This format is used for a required value, where only one item can be selected.	`switch {active|stand}`

Table of Contents

1.Authorize DSC to access Alibaba Cloud resources

Before you use Data Security Center (DSC), you must authorize DSC to access Alibaba Cloud resources. This topic describes how to authorize DSC to access Alibaba Cloud resources.

Prerequisites

DSC is activated.

Context

When you log on to the DSC console for the first time after you activate DSC, the Overview page prompts you to authorize DSC to access Alibaba Cloud resources. DSC can access Alibaba Cloud services, such as Object Storage Service (OSS), ApsaraDB RDS, and MaxCompute, scan your data assets for sensitive data, and analyze the detected sensitive data only after the authorization is complete.

Procedure

1. Log on to the DSC console.
2. On the **Welcome** page, click **Authorize Now**.

   After you click Authorize Now, Alibaba Cloud automatically creates the AliyunServiceRoleForSDDP role for DSC. You can view the created service-linked role on the **RAM Roles** page of the Resource Access Management (RAM) console. You can also query the information about the service-linked role for DSC by using the API or command line interface (CLI). For more information, see Service-linked roles.

   After you authorize DSC to access Alibaba Cloud resources, you must authorize DSC to access specific data assets before DSC can scan the data assets for sensitive data and analyze the detected sensitive data. For more information, see Grant access to data assets.

Service-linked role for DSC

Role name: AliyunServiceRoleForSDDP

Policy name: AliyunServiceRolePolicyForSDDP

Delete the service-linked role

If you no longer need to use DSC, you can delete the service-linked role for DSC. You can delete the AliyunServiceRoleForSDDP role in the RAM console. For more information, see Service-linked roles.

Usage notes

Do not confuse the operation of authorizing DSC to access Alibaba Cloud resources with the operation of authorizing DSC to access data assets. The former operation authorizes DSC to access other data services of Alibaba Cloud, whereas the latter operation authorizes DSC to access specific data assets in these data services. After you authorize DSC to access Alibaba Cloud resources, you must authorize DSC to access specific data assets before DSC can scan the data assets for sensitive data and analyze the detected sensitive data. For more information, see Grant access to data assets.

2.View summary information

After you activate Data Security Center (DSC) and configure risk levels and sensitive data detection rules, you can view the summary information on the Overview page of the DSC console. The summary information includes statistics on authorized data assets, sensitive data detection, static de-identification, and anomalous activities.

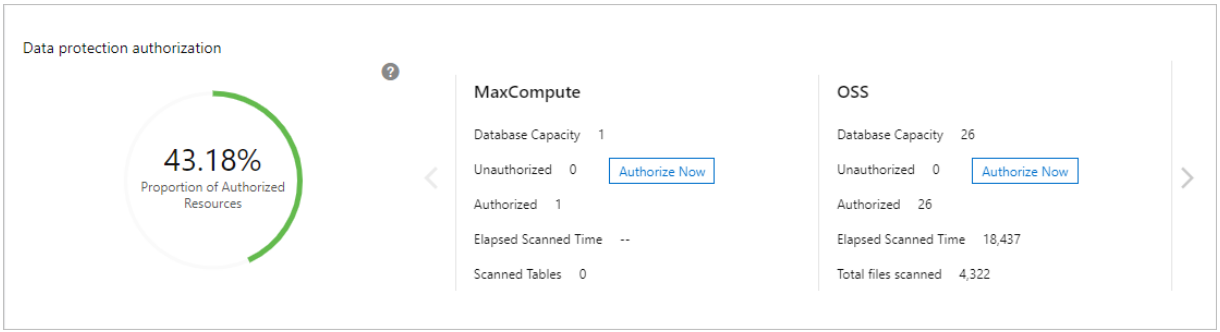
DSC can detect sensitive data in MaxCompute, ApsaraDB RDS, Object Storage Service (OSS), Tablestore, self-managed databases hosted on Elastic Compute Service (ECS) instances, DRDS, and PolarDB.

> **Note** For your data security and privacy, DSC performs only necessary operations such as sensitive data detection and static de-identification. DSC does not store your data.

View statistics on authorized data assets

In the **Data asset authorization** section, you can view statistics on authorized data assets, such as the proportion of data assets that DSC is authorized to access in all your data assets. For each service, this section displays the total number of data assets, the number of data assets that DSC is authorized to access, and the number of unauthorized data assets. This section also displays the total number of times that DSC scans data assets, and the number of objects or tables that DSC scanned.

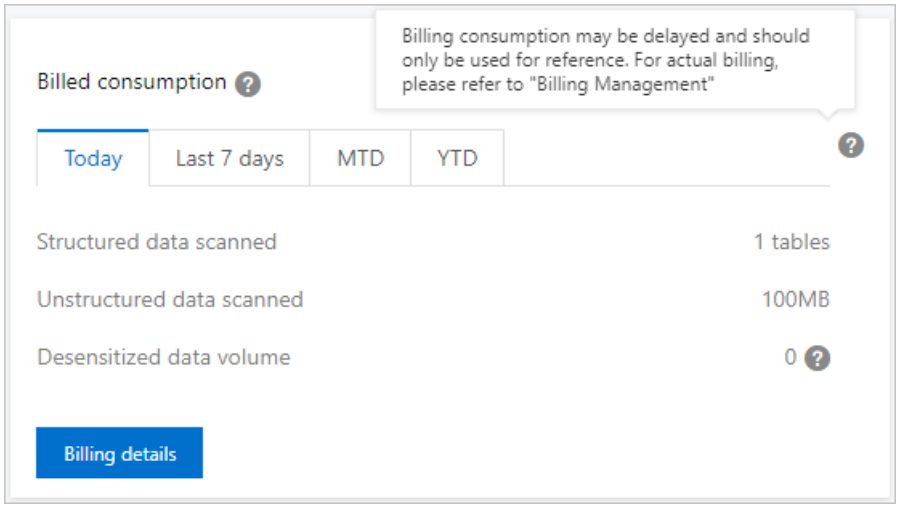
Click **Authorize Now** for a service. The **Cloud hosting** page appears, where you can authorize DSC to access more data assets and view, edit, or delete authorized data assets. For more information, see Grant access to data assets.



View the service usage in pay-as-you-go mode

In the **Billed consumption** section, you can view the size of structured data scanned, unstructured data scanned, and data de-identified by DSC. You can click the tabs to view the data scanned or de-identified on the current day, in the current week, in the current month, and in the current year.

> **Note**
>
> - Structured data refers to data stored in MaxCompute, ApsaraDB RDS, OSS, Tablestore, self-managed databases hosted on ECS instances, DRDS, and PolarDB. Unstructured data refers to data stored in OSS.
> - The billing unit for de-identification is cell.
> - The statistics in this section may not be updated in real time. We recommend that you use the statistics only for reference and check your bills for actual service usage.



Click **Billing details** to go to the **Bills** page in User Center, where you can view the details of each bill.

View statistics on sensitive data detection

In the **Data Identification Result** section, you can view the numbers of sensitive tables, fields, and objects detected in authorized data assets.

DSC detects sensitive data based on the sensitive data detection rules that you have configured. Sensitive data detection rules include built-in rules and custom rules. DSC detects sensitive tables, fields, and objects and classifies them by risk level based on the sensitive data detection rules. For more information, see Manage detection models.

The Data Identification Result section displays the following statistics:

- **Sensitive Tables**: the total number of sensitive tables detected in MaxCompute, ApsaraDB RDS, Tablestore, self-managed databases hosted on ECS instances, DRDS, and PolarDB.
- **Sensitive Fields**: the total number of sensitive fields detected in MaxCompute, ApsaraDB RDS, Tablestore, self-managed databases hosted on ECS instances, DRDS, and PolarDB.
- **Sensitive Files**: the total number of sensitive objects detected in OSS.
- **Risk Level Distribution of Sensitive Tables** and **Risk Level Distribution of Sensitive Fields**: the number of sensitive tables or fields at each risk level in the authorized data assets.

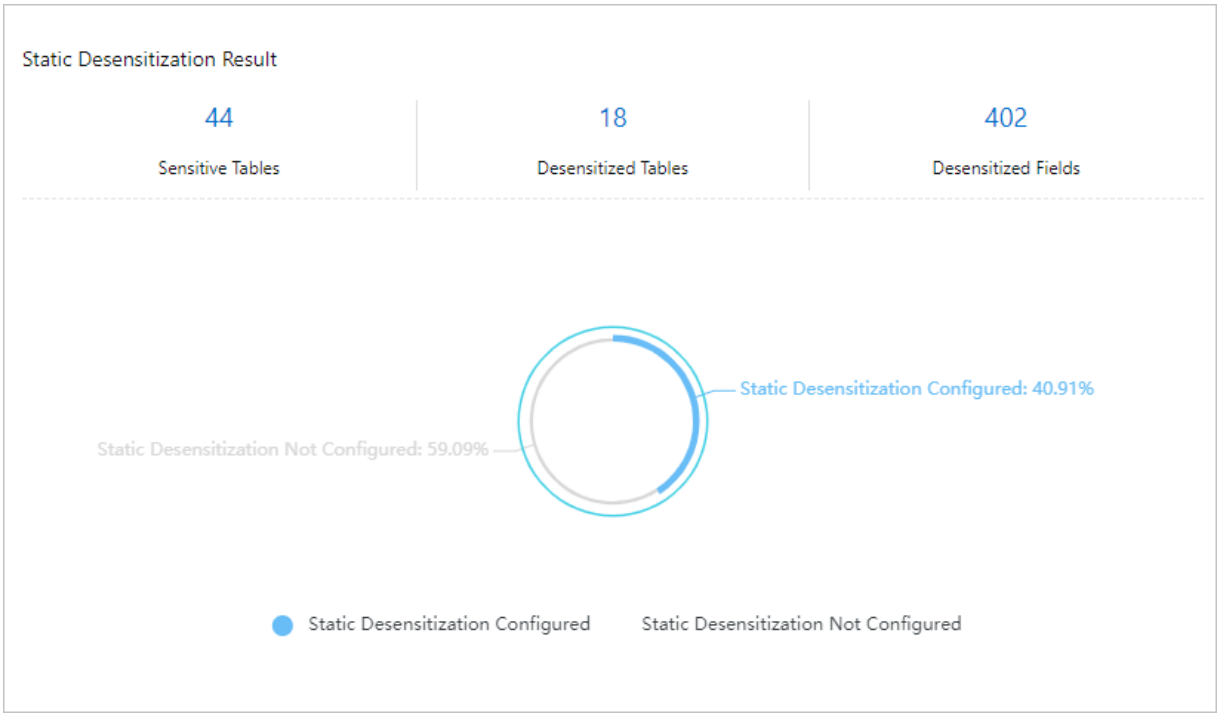
  DSC allows you to define the risk levels of sensitive fields in sensitive data detection rules. DSC supports the following risk levels: S1, S2, and S3. The severity of each risk level sequentially increases in the following way:

  - **S1**: low risk.
  - **S2**: medium risk.
  - **S3**: high risk.

View statistics on static de-identification

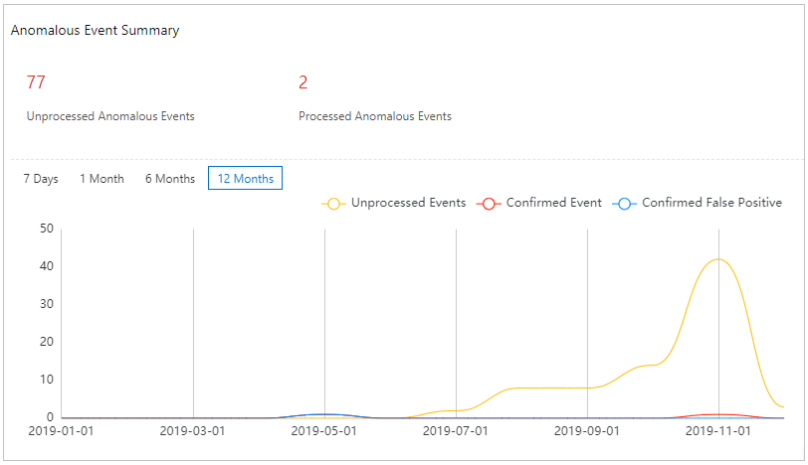
In the **Static Desensitization Result** section, you can view the number of sensitive tables that DSC de-identifies by using de-identification algorithms, number of de-identified tables, and number of de-identified fields. You can also view the proportion of tables for which static de-identification is configured in all sensitive tables.

DSC allows you to create static de-identification tasks to de-identify and protect sensitive data in your data assets. For more information, see Perform static de-identification.



View statistics on and trends in anomalous activities

In the **Anomalous Event Summary** section, you can view the trends in anomalous activities detected in the last seven days, one month, six months, or twelve months.



The Anomalous Event Summary section displays the following statistics:

- **Unprocessed Anomalous Events**: the number of anomalous activities that have not been processed.

DSC can detect anomalous activities that occurred in reading and using sensitive data. Such activities include anomalous permission access, anomalous data flows, and anomalous data operations. DSC detects anomalous activities and generates anomaly alerts based on the anomaly alert configuration.

- **Processed Anomalous Events**: the number of anomalous activities that have been processed.

  DSC allows you to process an anomalous activity by confirming it as a violation or excluding it as a false positive.

Query data assets

Click **Data Asset Search** in the upper-right corner of the **Overview** page. Then, query data assets that contain sensitive data and the risk levels of sensitive data detected by DSC.

You can set the following filters to query data assets:

- **Risk Levels**: Select one or more risk levels of sensitive data to query data assets. If you do not select a risk level, DSC queries data assets that contain sensitive data at all risk levels.

  > **Note** The **N/A** option indicates an unknown risk level.

- **Asset Scope**: Select one or more types of data assets to query, such as MaxCompute project and MaxCompute table.
- **With Sensitive Data**: Select a sensitive data type from the drop-down list to query data assets.

  > **Note** Each option maps a sensitive data detection rule, which can be a built-in rule or a custom rule. You can choose **Sensitive data discovery > Identification Rules** in the left-side navigation pane to view the rules.

- **Asset Name**: Enter the name of a project, instance, database, package, table, or object to query the specific data asset.

3.Methods for authorizing DSC to access data assets

This topic describes the methods for authorizing Data Security Center (DSC) to access different types of data assets.

You can use the following methods to authorize DSC to access different types of data assets: one-click authorization and username and password-based authorization.

- One-click authorization allows you to authorize DSC to access a data asset with one click in the DSC console. You do not need to enter the username and password that are used to access the data asset. During the authorization process, DSC automatically generates a read-only account for the data asset. You cannot use the account to perform de-identification tasks.
- Username and password-based authorization allows you to authorize DSC to access a data asset by entering the username and password that are used to access the data asset. After you authorize DSC to access a data asset, DSC can detect, de-identify, and audit the sensitive data in the data asset. You can de-identify the detected sensitive data based on your business requirements.

Type of data asset	One-click authorization	Username and password-based authorization
ApsaraDB RDS for SQL Server	Supported	Supported
ApsaraDB RDS for PostgreSQL	Supported	Supported
ApsaraDB RDS for PPAS	Supported	Supported
ApsaraDB RDS for MariaDB TX	Supported	Supported
Distributed Relational Database Service (DRDS)	Supported	Supported
AnalyticDB for MySQL	Not supported	Supported
AnalyticDB for PostgreSQL	Not supported	Supported
MaxCompute	Not supported	Supported
Object Storage Service (OSS)	Supported	Supported
PolarDB-X	Not supported	Supported
PolarDB for MySQL	Not supported	Supported
PolarDB for PostgreSQL	Not supported	Supported
Self-managed MySQL databases hosted on Elastic Compute Service (ECS) instances	Not supported	Supported
ApsaraDB OceanBase for MySQL	Not supported	Supported

Type of data asset	One-click authorization	Username and password-based authorization
ApsaraDB OceanBase for Oracle	Not supported	Supported
ApsaraDB for MongoDB	Not supported	Supported

4.Grant access to data assets

Data Security Center (DSC) must be authorized to access specific data assets before it can detect sensitive data in the data assets. Supported data assets include Object Storage Service (OSS) buckets, ApsaraDB RDS databases, ApsaraDB RDS for PPAS databases, Distributed Relational Database Service (DRDS) databases, PolarDB databases, Tablestore instances, self-managed databases hosted on Elastic Compute Service (ECS) instances, MaxCompute projects, AnalyticDB for PostgreSQL databases, AnalyticDB for MySQL databases, ApsaraDB for MongoDB databases, ApsaraDB for OceanBase databases, and ApsaraDB for Redis databases. This topic describes how to authorize DSC to access specific data assets.

Prerequisites

DSC is activated. DSC is authorized to access Alibaba Cloud resources. For more information, see Authorize DSC to access Alibaba Cloud resources.

Background information

You can authorize DSC to access specific data assets in Alibaba Cloud services. If you do not authorize DSC to access the data assets, DSC cannot detect sensitive data in Alibaba Cloud services or de-identify the sensitive data.

Authorize DSC to access OSS buckets

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **OSS** tab, click **Unauthorized**.
4. Select the OSS buckets that you want to authorize DSC to access and click **Batch operation**.

   You can also click **Authorization** in the **Open protection** column for a single OSS bucket to authorize DSC to access the OSS bucket.
5. In the **Batch processing for selected assets** dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:

   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.
   - **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.
   - **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-identification permissions on the selected data assets.
   - **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
     - 0
     - 5
     - 10

- **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
  - 30 days
  - 90 days
  - 180 days

  > **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

6. Click **Ok**.
   After the authorization is complete, DSC scans authorized OSS buckets for sensitive data. When DSC accesses an OSS bucket for the first time, DSC automatically scans all the data in the OSS bucket, and you are charged for the full scan. For more information, see the "How long does it take to scan data in my data asset after I authorize DSC to access the data asset?" section of the Sensitive data scan and detection topic.

   In the list of authorized data assets, you can modify the authorization configuration for a data asset or cancel the authorization for a data asset. After you cancel the authorization for an OSS bucket, DSC no longer scans the OSS bucket.

   > **Note** DSC scans only authorized OSS buckets and analyzes risks of sensitive data detected in these OSS buckets.

Authorize DSC to access ApsaraDB RDS databases

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **RDS** tab.
4. On the **RDS** tab, click **Not authorized**.
5. Find the data assets that you want to authorize DSC to access and enter the username and password that are used to access each data asset in the **Username** and **Password** fields.

   You can also click **Batch password import** to import the logon information for multiple data assets at a time. For more information, see Import logon information for multiple data assets at a time.

   > **Notice** Invalid usernames or passwords cause an authorization failure. Make sure that you enter valid usernames and passwords.

6. Select the data assets that you want to authorize DSC to access and click **Batch operation**.

   You can also click **One-click authorization** or Account Password Authorization in the Actions column for a single data asset to authorize DSC to access the data asset.
7. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:

   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.

- **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.

  SDDP allows you to collect audit logs that cover the generation, update, and use of your data assets. The log information includes the audit rule that is hit for a data asset, the type of the data asset, the type of the operation that hits the audit rule, and the operator account.

  > **Note** After you enable the audit log feature for an ApsaraDB RDS database, SQL Explorer is automatically enabled, and you are charged for using SQL Explorer. You are charged an hourly fee of USD 0.0018 per GB for using SQL Explorer of the non-trial edition. The fee is listed in your bill of ApsaraDB RDS. For more information about how to view the fee, see 查看消费明细. For more information about SQL Explorer, see Use the SQL Explorer feature on an ApsaraDB RDS for MySQL instance.

- **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-identification permissions on the selected data assets.
- **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
  - 0
  - 5
  - 10
- **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
  - 30 days
  - 90 days
  - 180 days

  > **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

8. Click **OK**.

   > **Note** If the authorization fails, check whether the usernames and passwords are correct.

   After the authorization is complete, DSC scans authorized data assets for sensitive data.

   In the list of authorized data assets, you can modify the authorization configuration for a data asset or cancel the authorization for a data asset. When you modify the authorization configuration for an ApsaraDB RDS database, you can modify only the username and password for accessing the database. After you cancel the authorization, DSC no longer scans the database.

Authorize DSC to access ApsaraDB RDS for PPAS databases

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **RDS-PPAS** tab.

4. On the **RDS-PPAS** tab, click **Add data assets**.
5. In the **Add data assets** dialog box, set the parameters as required and click **Ok**.

   The following table describes the parameters for adding an ApsaraDB RDS for PPAS database to DSC.

Parameter	Description
Region	The region where the ApsaraDB RDS for PPAS database that you want to authorize DSC to access resides.
Instance Name	The name of the ECS instance on which the ApsaraDB RDS for PPAS database that you want to authorize DSC to access is hosted.
Database Name	The name of the ApsaraDB RDS for PPAS database that you want to authorize DSC to access.
User name	The username and password of a valid user of the ApsaraDB RDS for PPAS database.
Password	

6. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.
7. Click **OK**.
   After the authorization is complete, DSC scans authorized data assets for sensitive data.

Authorize DSC to access DRDS databases

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **DRDS** tab.
4. On the **DRDS** tab, click **Not authorized**.
5. Find the data assets that you want to authorize DSC to access and enter the username and password that are used to access each data asset in the **Username** and **Password** fields.

   You can also click **Batch password import** to import the logon information for multiple data assets at a time. For more information, see Import logon information for multiple data assets at a time.

   > **Notice** Invalid usernames or passwords cause an authorization failure. Make sure that you enter valid usernames and passwords.

6. Select the data assets that you want to authorize DSC to access and click **Batch operation**.

   You can also click **One-click authorization** or Account Password Authorization in the Actions column for a single data asset to authorize DSC to access the data asset.
7. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:

   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.

- **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.
- **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-identification permissions on the selected data assets.
- **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
  - 0
  - 5
  - 10
- **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
  - 30 days
  - 90 days
  - 180 days

> **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

8. Click **OK**.

> **Note** If the authorization fails, check whether the usernames and passwords are correct.

After the authorization is complete, DSC scans authorized data assets for sensitive data.

In the list of authorized data assets, you can modify the authorization configuration for a data asset or cancel the authorization for a data asset. When you modify the authorization configuration for a DRDS database, you can modify only the username and password for accessing the database. After you cancel the authorization, DSC no longer scans the database.

Authorize DSC to access PolarDB databases

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **PolarDB** tab.
4. On the **PolarDB** tab, click **Unauthorized**.
5. Find the data assets that you want to authorize DSC to access and enter the username and password that are used to access each data asset in the **Username** and **Password** fields.

   You can also click **Batch password import** to import the logon information for multiple data assets at a time. For more information, see Import logon information for multiple data assets at a time.

> **Notice** Invalid usernames or passwords cause an authorization failure. Make sure that you enter valid usernames and passwords.

6. Select the data assets that you want to authorize DSC to access and click **Batch operation**.

You can also click **One-click authorization** or Account Password Authorization in the Actions column for a single data asset to authorize DSC to access the data asset.

7. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:

   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.
   - **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.
   - **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-identification permissions on the selected data assets.
   - **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
     - 0
     - 5
     - 10
   - **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
     - 30 days
     - 90 days
     - 180 days

     > **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

8. Click **OK**.

   > **Note** If the authorization fails, check whether the usernames and passwords are correct.

   After the authorization is complete, DSC scans authorized data assets for sensitive data.

   In the list of authorized data assets, you can modify the authorization configuration for a data asset or cancel the authorization for a data asset. When you modify the authorization configuration for a PolarDB database, you can modify only the username and password for accessing the database. After you cancel the authorization, DSC no longer scans the database.

Authorize DSC to access Tablestore instances

OTS refers to Tablestore.

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **OTS** tab.
4. On the **OTS** tab, click **Unauthorized**.

5. Select the data assets that you want to authorize DSC to access and click **Batch operation**.

   You can also click **One-click authorization** or Account Password Authorization in the Actions column for a single data asset to authorize DSC to access the data asset.

6. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:

   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.
   - **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.
   - **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-identification permissions on the selected data assets.
   - **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
     - 0
     - 5
     - 10
   - **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
     - 30 days
     - 90 days
     - 180 days

     > **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

7. Click **OK**.
   After the authorization is complete, DSC scans authorized data assets for sensitive data.

Authorize DSC to access self-managed databases hosted on ECS instances

A self-managed database hosted on an ECS instance must meet the following requirements before DSC can scan the database:

- The ECS instance on which the self-managed database is hosted resides in a virtual private cloud (VPC) so that DSC can scan the database.
- The self-managed database hosted on the ECS instance is a MySQL, an SQL Server, or an Oracle database.
- The account that you use to connect to the self-managed database hosted on the ECS instance is granted the permissions to remotely access the self-managed database from specified CIDR blocks. You must log on to the self-managed database to complete this authorization before you authorize DSC to access the self-managed database.

1. Log on to the self-managed database hosted on the ECS instance. Grant the account that you use

to connect to the self-managed database the permissions to remotely access the self-managed database from specified CIDR blocks.

For example, run the following command to grant the remote access permissions if the self-managed database hosted on the ECS instance is a MySQL database. If the self-managed database hosted on the ECS instance is of another database type, run the authorization command based on the syntax of the specific database type.

```
GRANT ALL PRIVILEGES ON *.* TO 'Username' @ 'CIDR blocks' IDENTIFIED BY 'Password'
```

Parameter description:

- Username: the username of the account that you use to connect to the self-managed database hosted on the ECS instance.
- CIDR blocks: the CIDR blocks from which the self-managed database hosted on the ECS instance can be accessed.

  You must specify the CIDR blocks in the authorization command based on the region where the data assets reside and the network type of the data assets.

  For more information about the CIDR blocks, see the following CIDR blocks section. In the authorization command, you must specify at least two CIDR blocks of the corresponding region. In addition, the IP address range of the CIDR blocks that you specify can be greater than that of the two CIDR blocks of the corresponding region.
- Password: the password of the account that you use to connect to the self-managed database hosted on the ECS instance.

CIDR blocks

Region	CIDR block
China (Shanghai)	○ 100.104.238.64/26 ○ 100.104.198.192/26
China (Beijing)	○ 100.104.250.0/26 ○ 100.104.51.192/26
China (Hangzhou)	○ 100.104.207.192/26 ○ 100.104.232.64/26
China (Shenzhen)	○ 100.104.247.0/26 ○ 100.104.150.64/26
China (Zhangjiakou)	○ 100.104.37.128/26 ○ 100.104.191.64/26
China (Hohhot)	○ 100.104.234.192/26 ○ 100.104.26.128/26

Region	CIDR block
China (Hong Kong)	○ 100.104.153.64/26 ○ 100.104.65.192/26
Singapore (Singapore)	○ 100.104.158.192/26 ○ 100.104.218.128/26
Malaysia (Kuala Lumpur)	○ 100.104.240.128/26 ○ 100.104.127.0/26
Indonesia (Jakarta)	○ 100.104.127.0/26 ○ 100.104.182.128/26

2. Log on to the DSC console.
3. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
4. On the **Cloud hosting** page, click the **ECS self-built database** tab.
5. On the **ECS self-built database** tab, click **Add data assets**.
6. In the **Asset authorization** dialog box, set the parameters as required and click **Next**.

   The following table describes the parameters for adding a self-managed database hosted on an ECS instance to DSC.

Parameter	Description
Region	The region where the self-managed database that you want to authorize DSC to access resides.
ECS instance ID	The ID of the ECS instance on which the self-managed database that you want to authorize DSC to access is hosted.
Database type	The type of the self-managed database that you want to authorize DSC to access. DSC supports the following types of self-managed databases hosted on ECS instances: MySQL, SQL Server, and Oracle.
Library name	The name of the self-managed database that you want to authorize DSC to access. **Note** To authorize DSC to access other self-managed databases hosted on the ECS instance, click **Add Database** to add the databases.
Port	The port number that is used to access the self-managed database hosted on the ECS instance.

Parameter	Description
User name	The username and password of a valid user of the self-managed database hosted on the ECS instance.
Password	

7. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:

   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.
   - **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.
   - **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-identification permissions on the selected data assets.
   - **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
     - 0
     - 5
     - 10
   - **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
     - 30 days
     - 90 days
     - 180 days

     > **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

8. Click **OK**.
   After the authorization is complete, DSC scans authorized data assets for sensitive data.

Authorize DSC to access a MaxCompute project

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **MaxCompute** tab.
4. On the **MaxCompute** tab, click **Add data assets**.
5. In the **Add data assets** dialog box, set the parameters as required. The following table describes the parameters for adding a MaxCompute project to DSC.

Parameter	Description
Region	The region where the MaxCompute project that you want to authorize DSC to access resides.
Project Name	The name of the MaxCompute project. **Note** Fuzzy search is not supported. You must enter the exact name of the project.

6. Run the following commands on the MaxCompute client to add the DSC account **yundun_sddp** to the MaxCompute project. DSC uses this account to access the MaxCompute project.

```
add user aliyun$yundun_sddp;
grant admin to aliyun$yundun_sddp;
```

Perform one of the following operations based on the returned result:

- If no error message is returned after the preceding commands are run, go to Step 8.
- If an error message is returned after the preceding commands are run, go to Step 7.

7. (Optional)Run the following command to add the service IP addresses of DSC to the IP address whitelist of the MaxCompute project:

```
setproject odps.security.ip.whitelist=11.193.236.0/24,11.193.64.0/24,11.193.58.0/24 odps.security.vpc.whitelist=<VPC ID>;
// 11.193.236.0/24, 11.193.64.0/24, and 11.193.58.0/24 are the CIDR blocks used by DSC on the classic network. They must be added to the IP address whitelist.
// Replace the VPC ID with that of the region where your MaxCompute project resides. The following table describes the VPC IDs of the supported regions.
```

If the IP address whitelist feature is enabled for your MaxCompute project, you must add the service IP addresses of DSC to the IP address whitelist of the MaxCompute project. This prevents authorization failures. You can run the `setproject;` command to check whether the IP address whitelist feature is enabled for your MaxCompute project. If the value of the `odps.security.vpc.whitelist` parameter is empty, the IP address whitelist feature is not enabled. In this case, you can skip this step.

Region	Region ID	VPC ID
China (Zhangjiakou)	cn-zhangjiakou	cn-zhangjiakou_399229
China (Beijing)	cn-beijing	cn-beijing_691047
China (Shenzhen)	cn-shenzhen	cn-shenzhen_515895
China (Shanghai)	cn-shanghai	cn-shanghai_28803
China (Hangzhou)	cn-hangzhou	cn-hangzhou_551733

> **Note** After you configure the IP address whitelist, wait for 5 minutes before you go to the next step.

8. Click **Ok**.

> **Note** If the authorization fails, check whether the authorization parameters are correctly set and whether the DSC account is added to the MaxCompute project.

After the authorization is complete, DSC scans the authorized MaxCompute project for sensitive data.

In the list of authorized MaxCompute projects, you can cancel the authorization for a MaxCompute project. After you cancel the authorization, DSC no longer scans the project.

Authorize DSC to access AnalyticDB for PostgreSQL databases

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **ADB-PG** tab.
4. On the **ADB-PG** tab, click **Add data assets**.
5. In the **Add data assets** dialog box, set the parameters as required and click **Ok**.

   The following table describes the parameters for adding an AnalyticDB for PostgreSQL database to DSC.

Parameter	Description
Region	The region where the AnalyticDB for PostgreSQL database that you want to authorize DSC to access resides.
Instance Name	The name of the ECS instance on which the AnalyticDB for PostgreSQL database that you want to authorize DSC to access is hosted.
Database Name	The name of the AnalyticDB for PostgreSQL database that you want to authorize DSC to access.
User name	The username and password of a valid user of the AnalyticDB for PostgreSQL database.
Password	

6. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:

   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.
   - **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.
   - **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-

identification permissions on the selected data assets.

- **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
  - 0
  - 5
  - 10
- **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
  - 30 days
  - 90 days
  - 180 days

> **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

7. Click **OK**.
   After the authorization is complete, DSC scans authorized data assets for sensitive data.

Authorize DSC to access AnalyticDB for MySQL databases

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **ADB-MYSQL** tab.
4. On the **ADB-MYSQL** tab, click **Unauthorized**.
5. Find the data assets that you want to authorize DSC to access and enter the username and password that are used to access each data asset in the **Username** and **Password** fields.

   You can also click **Batch password import** to import the logon information for multiple data assets at a time. For more information, see Import logon information for multiple data assets at a time.

> **Notice** Invalid usernames or passwords cause an authorization failure. Make sure that you enter valid usernames and passwords.

6. Select the data assets that you want to authorize DSC to access and click **Batch operation**.

   You can also click **One-click authorization** or Account Password Authorization in the Actions column for a single data asset to authorize DSC to access the data asset.
7. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:
   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.
   - **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.

- **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-identification permissions on the selected data assets.
- **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
  - 0
  - 5
  - 10
- **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
  - 30 days
  - 90 days
  - 180 days

  > **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

8. Click **OK**.

   > **Note** If the authorization fails, check whether the usernames and passwords are correct.

   After the authorization is complete, DSC scans authorized data assets for sensitive data.

   In the list of authorized data assets, you can modify the authorization configuration for a data asset or cancel the authorization for a data asset. When you modify the authorization configuration for an AnalyticDB for MySQL database, you can modify only the username and password for accessing the database. After you cancel the authorization, DSC no longer scans the database.

Authorize DSC to access ApsaraDB for MongoDB databases

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **MongoDB** tab.
4. On the **MongoDB** tab, click **Add data assets**.
5. In the **Add data assets** dialog box, set the parameters as required and click **Ok**.

   The following table describes the parameters for adding an ApsaraDB for MongoDB database to DSC.

Parameter	Description
Region	The region where the ApsaraDB for MongoDB database that you want to authorize DSC to access resides.
Instance Name	The name of the ECS instance on which the ApsaraDB for MongoDB database that you want to authorize DSC to access is hosted.

Parameter	Description
Database Name	The name of the ApsaraDB for MongoDB database that you want to authorize DSC to access.
User name	The username and password of a valid user of the ApsaraDB for MongoDB database.
Password	

6. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:

   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.
   - **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.
   - **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-identification permissions on the selected data assets.
   - **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
     - 0
     - 5
     - 10
   - **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
     - 30 days
     - 90 days
     - 180 days

     > **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

7. Click **OK**.
   After the authorization is complete, DSC scans authorized data assets for sensitive data.

Authorize DSC to access ApsaraDB for OceanBase databases

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **OceanBase** tab.
4. On the **OceanBase** tab, click **Unauthorized**.
5. Find the data assets that you want to authorize DSC to access and enter the username and password that are used to access each data asset in the **Username** and **Password** fields.

You can also click **Batch password import** to import the logon information for multiple data assets at a time. For more information, see Import logon information for multiple data assets at a time.

> **Notice** Invalid usernames or passwords cause an authorization failure. Make sure that you enter valid usernames and passwords.

6. Select the data assets that you want to authorize DSC to access and click **Batch operation**.

   You can also click **One-click authorization** or Account Password Authorization in the Actions column for a single data asset to authorize DSC to access the data asset.

7. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:

   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.
   - **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.
   - **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-identification permissions on the selected data assets.
   - **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
     - 0
     - 5
     - 10
   - **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
     - 30 days
     - 90 days
     - 180 days

     > **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

8. Click **OK**.

   > **Note** If the authorization fails, check whether the usernames and passwords are correct.

   After the authorization is complete, DSC scans authorized data assets for sensitive data.

   In the list of authorized data assets, you can modify the authorization configuration for a data asset or cancel the authorization for a data asset. When you modify the authorization configuration for an ApsaraDB for OceanBase database, you can modify only the username and password for accessing the database. After you cancel the authorization, DSC no longer scans the database.

Authorize DSC to access ApsaraDB for Redis databases

ApsaraDB for Redis

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click the **Redis** tab.
4. On the **Redis** tab, click **Unauthorized**.
5. Find the data assets that you want to authorize DSC to access and enter the username and password that are used to access each data asset in the **Username** and **Password** fields.

   You can also click **Batch password import** to import the logon information for multiple data assets at a time. For more information, see Import logon information for multiple data assets at a time.

   > **Notice** Invalid usernames or passwords cause an authorization failure. Make sure that you enter valid usernames and passwords.

6. Select the data assets that you want to authorize DSC to access and click **Batch operation**.

   You can also click **One-click authorization** or Account Password Authorization in the Actions column for a single data asset to authorize DSC to access the data asset.
7. In the Batch operation dialog box, turn on or off the switches to configure the detection, audit, and de-identification permissions for DSC and set the remaining parameters as required.

   Set the following switches and parameters:

   - **Identify permissions**: specifies whether to grant DSC the sensitive data detection permissions on the selected data assets.
   - **Audit permissions**: specifies whether to grant DSC the audit permissions on the selected data assets.
   - **Desensitization permissions**: specifies whether to grant DSC the sensitive data de-identification permissions on the selected data assets.
   - **Display number of sampling**: the number of samples that DSC collects from the selected data assets. DSC collects samples when it detects sensitive data in the data assets. You can use the sensitive data samples to further analyze the sensitive data. Valid values:
     - 0
     - 5
     - 10
   - **Audit log archiving**: the number of days for which audit logs are retained for the selected data assets. Valid values:
     - 30 days
     - 90 days
     - 180 days

     > **Note** You do not need to activate Log Service to archive audit logs that are generated by DSC.

8. Click **Ok**.

   > **Note** If the authorization fails, check whether the usernames and passwords are correct.

   After the authorization is complete, DSC scans authorized data assets for sensitive data.

   In the list of authorized data assets, you can modify the authorization configuration for a data asset or cancel the authorization for a data asset. When you modify the authorization configuration for an ApsaraDB for Redis database, you can modify only the username and password for accessing the database. After you cancel the authorization, DSC no longer scans the database.

Import logon information for multiple data assets at a time

DSC allows you to upload an EXCEL file to import logon information for multiple data assets at a time. This way, you can authorize DSC to access multiple data assets at a time. The data assets include ApsaraDB RDS databases, DRDS databases, and PolarDB databases. To import logon information for multiple data assets at a time, perform the following steps:

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data asset authorization > Data asset authorization**.
3. On the **Cloud hosting** page, click **Batch password import** in the upper-right corner.
4. In the **Batch password import** dialog box, click **DSC Authorization File Template.xlsx**.
5. Open the downloaded template file, enter the username and password used to access each data asset in the **user name** and **password** columns, and then save the template file.

   If you modify the existing usernames and passwords in the template file and upload the file to the DSC console, the usernames and passwords saved in the DSC console are updated.
6. In the **Batch password import** dialog box, click **File Upload** to upload the template file that you have edited.
7. Click **OK**.
   The EXCEL file is uploaded. Then, DSC synchronizes the usernames and passwords that you enter in the file to the **Username** and **Password** columns for the related databases on the **RDS**, **DRDS**, and **PolarDB** tabs. You can authorize DSC to access these databases on the **Cloud hosting** page without the need to manually enter the usernames and passwords for accessing the databases.

Troubleshoot an authorization failure

An authorization failure may occur when you authorize DSC to access your data assets. You can troubleshoot an authorization failure based on the following possible causes:

- **Possible causes of an authorization failure for ApsaraDB RDS**
  - The username or password for accessing the ApsaraDB RDS database is invalid.
  - The service IP addresses of DSC are deleted from the whitelist of the ApsaraDB RDS database.
  - The ApsaraDB RDS database resides on the classic network, but the public endpoint of the ApsaraDB RDS database is inaccessible due to access control.
- **Possible causes of an authorization failure for MaxCompute**
  - The name of the MaxCompute project is invalid.
  - The DSC account fails to be added to the MaxCompute project.

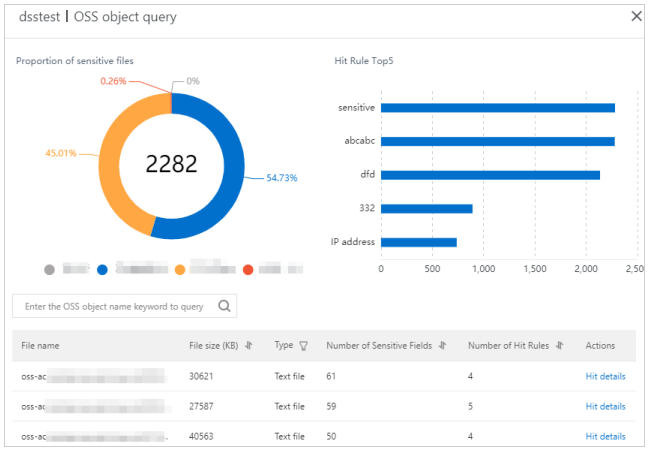
5.Sensitive data discovery

5.1. View sensitive data

This topic shows you how to view sensitive data detected by Data Security Center (DSC) in Alibaba Cloud services such as Object Storage Service (OSS), ApsaraDB RDS, and MaxCompute.

View sensitive data detected in OSS

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Sensitive data assets**.
3. On the **OSS** tab, find the OSS bucket whose details you want to view and click **File details** in the Actions column.
4. In the OSS object query panel, you can view the proportion of sensitive objects at each risk level, top five rules that are hit, and the list of objects with sensitive data detected.



- **Proportions of sensitive objects**

  In the **Proportion of sensitive files** section, you can view a pie chart that displays the numbers and proportions of sensitive objects at the low, medium, high, and unknown risk levels.

- **Top five rules that are hit**

  In the **Hit Rule Top5** section, you can view the top five rules that are hit and the number of times that each rule is hit.

- **List of objects with sensitive data detected**

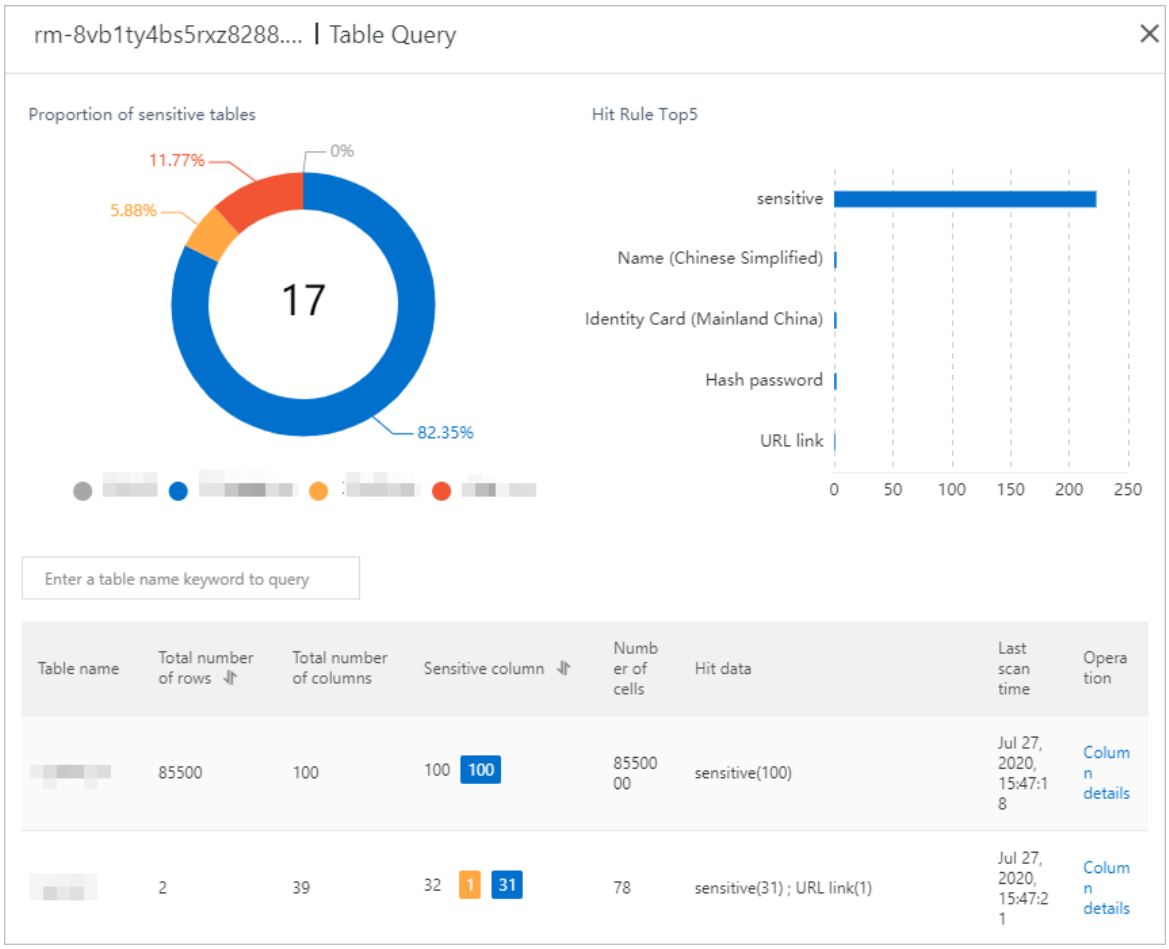
  In the object list, you can view the information about the objects with sensitive data detected, such as the object name, size, type, and number of sensitive fields detected in the object. You can also click **Hit details** in the Actions column for an object to view the details of the sensitive data detection rules that the object hits.

View sensitive data detected in ApsaraDB RDS

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Sensitive data assets**.
3. On the **Sensitive data assets** page, click the **RDS** tab.
4. On the **RDS** tab, find the ApsaraDB RDS database whose details you want to view and click **Table**

details in the Actions column.

5. In the Table Query panel, you can view the proportions of sensitive tables, top five rules that are hit, and the list of tables with sensitive data detected.



- **Proportions of tables**

  In the **Proportion of sensitive tables** section, you can view a pie chart that displays the numbers and proportions of tables at the low, medium, high, and unknown risk levels.

- **Top five rules that are hit**

  In the **Hit Rule Top5** section, you can view the top five rules that are hit and the number of times that each rule is hit.

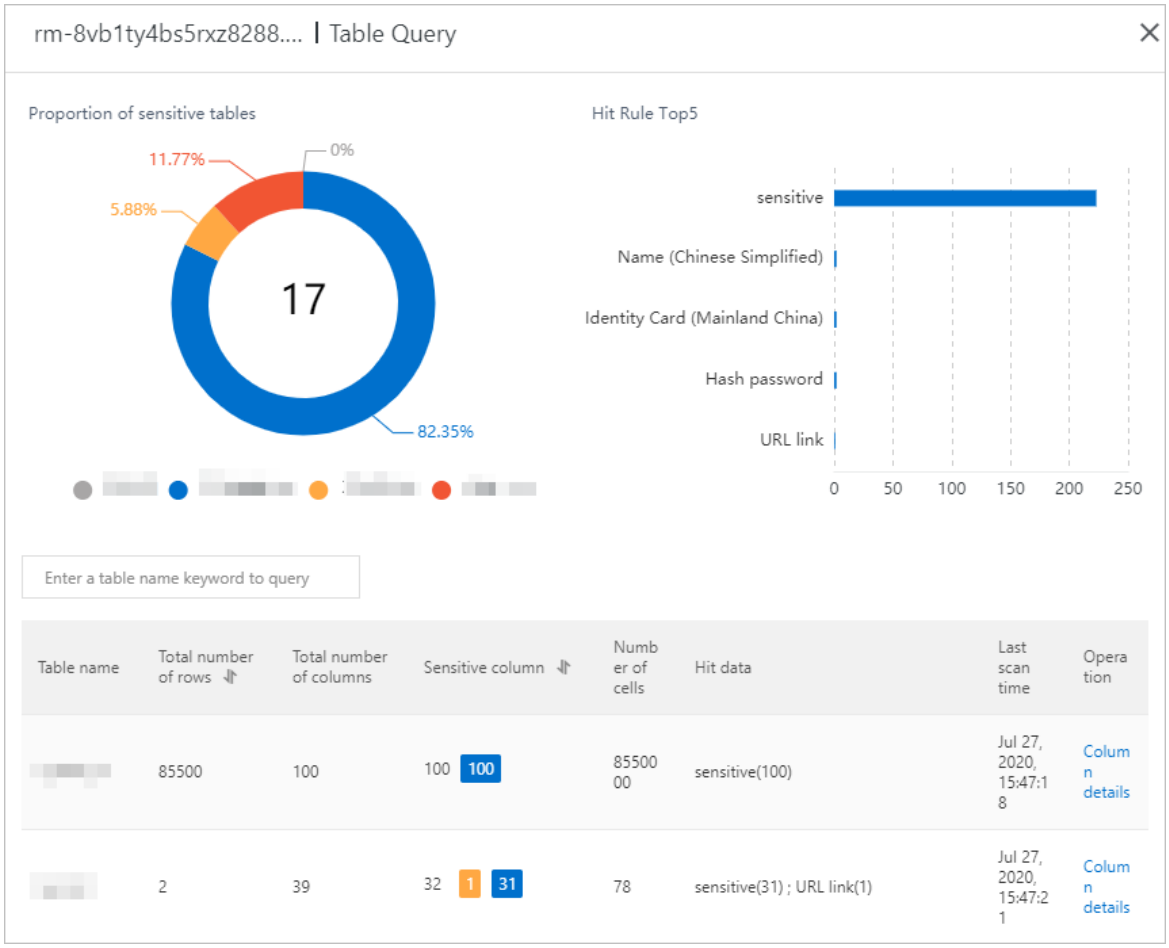
- **List of tables with sensitive data detected**

  In the table list, you can view the information about the tables with sensitive data detected. For example, you can view the table name, total number of rows, total number of fields, number of sensitive fields, and rules that are hit. You can also click **Column details** in the Operation column to view the field details, including the fields that hit sensitive data detection rules and the risk levels of the sensitive fields.

View sensitive data detected in MaxCompute

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Sensitive data assets**.
3. On the **Sensitive data assets** page, click the **MaxCompute** tab.

4. On the **MaxCompute** tab, find the MaxCompute project whose details you want to view and click **Table details** in the Actions column.
5. In the Table Query panel, you can view the proportions of sensitive tables, top five rules that are hit, and the list of tables with sensitive data detected.



- **Proportions of tables**

  In the **Proportion of sensitive tables** section, you can view a pie chart that displays the numbers and proportions of tables at the low, medium, high, and unknown risk levels.

- **Top five rules that are hit**

  In the **Hit Rule Top5** section, you can view the top five rules that are hit and the number of times that each rule is hit.

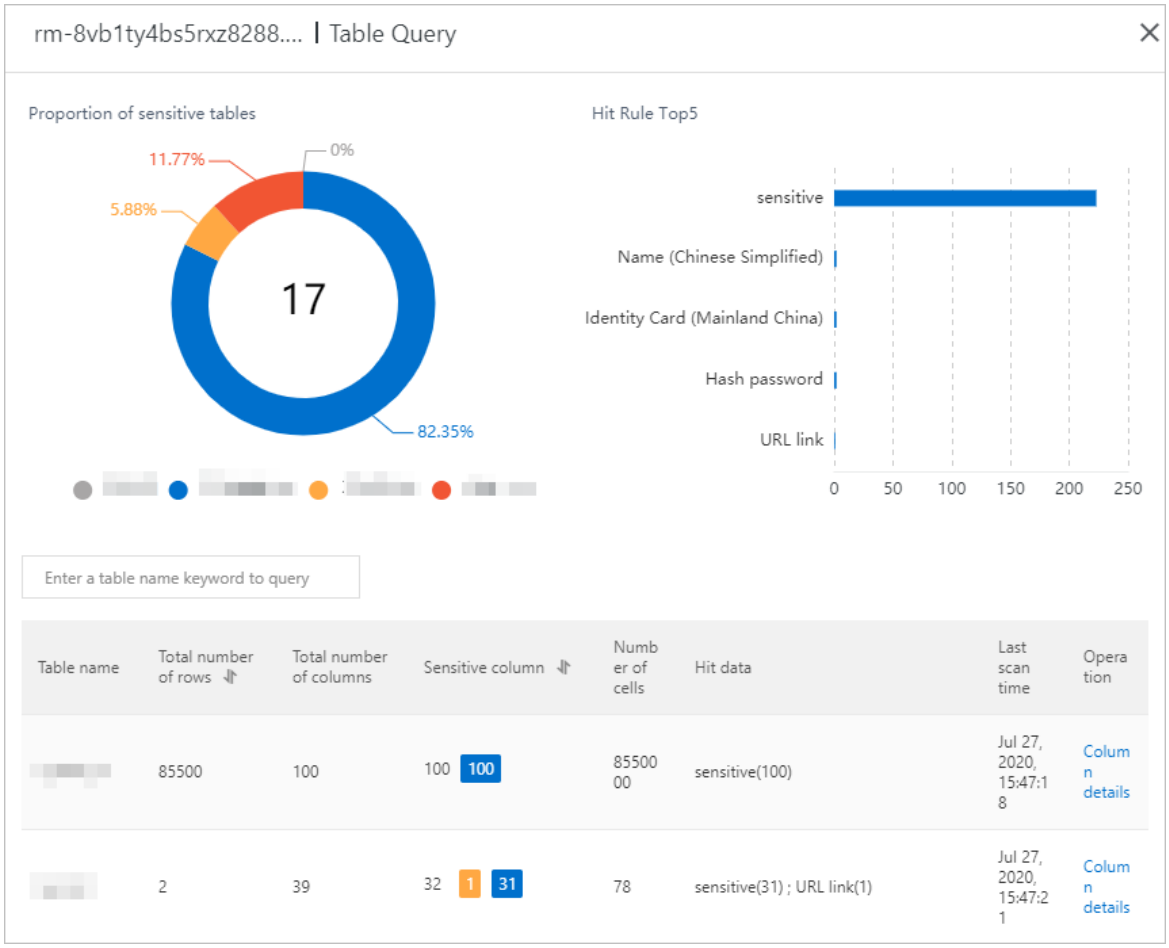
- **List of tables with sensitive data detected**

  In the table list, you can view the information about the tables with sensitive data detected. For example, you can view the table name, total number of rows, total number of fields, number of sensitive fields, and rules that are hit. You can also click **Column details** in the Operation column to view the field details, including the fields that hit sensitive data detection rules and the risk levels of the sensitive fields.

View sensitive data detected in self-managed databases hosted on ECS instances

1. Log on to the DSC console.

2. In the left-side navigation pane, choose **Sensitive data discovery > Sensitive data assets**.
3. On the **Sensitive data assets** page, click the **ECS self-built database** tab.
4. On the **ECS self-built database** tab, find the database whose details you want to view and click **Table details** in the Actions column.
5. In the Table Query panel, you can view the proportions of sensitive tables, top five rules that are hit, and the list of tables with sensitive data detected.



- **Proportions of tables**

  In the **Proportion of sensitive tables** section, you can view a pie chart that displays the numbers and proportions of tables at the low, medium, high, and unknown risk levels.

- **Top five rules that are hit**

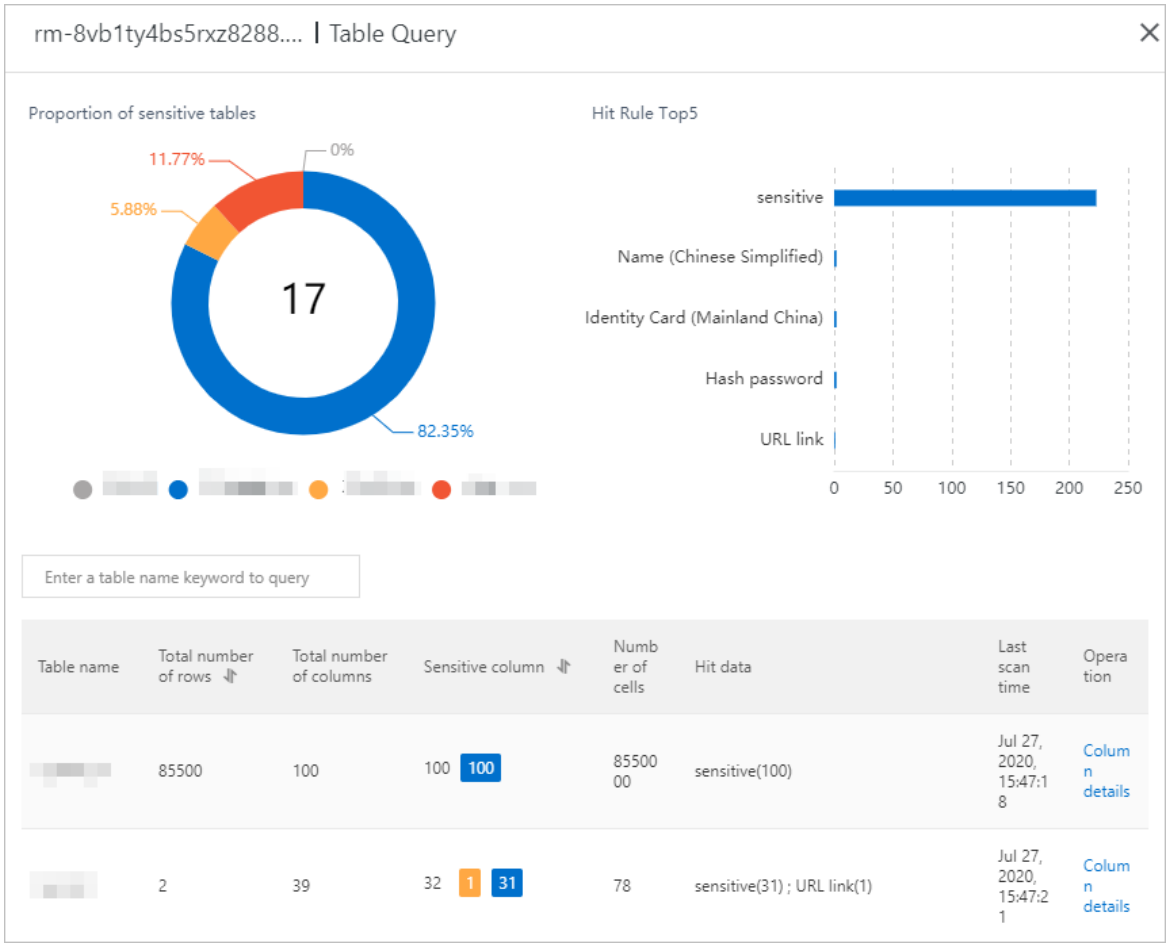
  In the **Hit Rule Top5** section, you can view the top five rules that are hit and the number of times that each rule is hit.

- **List of tables with sensitive data detected**

  In the table list, you can view the information about the tables with sensitive data detected. For example, you can view the table name, total number of rows, total number of fields, number of sensitive fields, and rules that are hit. You can also click **Column details** in the Operation column to view the field details, including the fields that hit sensitive data detection rules and the risk levels of the sensitive fields.

View sensitive data detected in DRDS

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Sensitive data assets**.
3. On the **Sensitive data assets** page, click the **DRDS** tab.
4. On the **DRDS** tab, find the database whose details you want to view and click **Table details** in the Actions column.
5. In the Table Query panel, you can view the proportions of sensitive tables, top five rules that are hit, and the list of tables with sensitive data detected.



- **Proportions of tables**

  In the **Proportion of sensitive tables** section, you can view a pie chart that displays the numbers and proportions of tables at the low, medium, high, and unknown risk levels.

- **Top five rules that are hit**

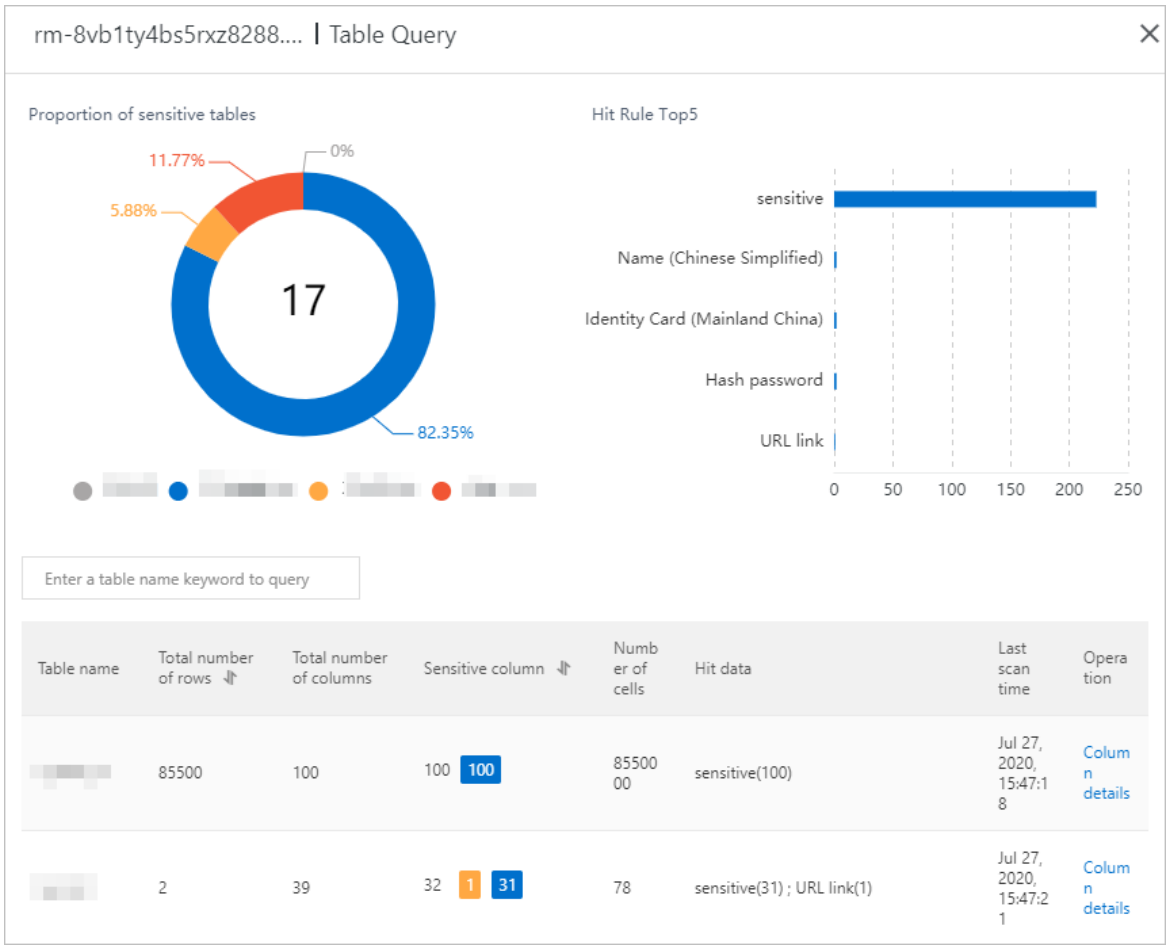
  In the **Hit Rule Top5** section, you can view the top five rules that are hit and the number of times that each rule is hit.

- **List of tables with sensitive data detected**

  In the table list, you can view the information about the tables with sensitive data detected. For example, you can view the table name, total number of rows, total number of fields, number of sensitive fields, and rules that are hit. You can also click **Column details** in the Operation column to view the field details, including the fields that hit sensitive data detection rules and the risk levels of the sensitive fields.

View sensitive data detected in PolarDB

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Sensitive data assets**.
3. On the **Sensitive data assets** page, click the **PolarDB** tab.
4. On the **PolarDB** tab, find the database whose details you want to view and click **Table details** in the Actions column.
5. In the Table Query panel, you can view the proportions of sensitive tables, top five rules that are hit, and the list of tables with sensitive data detected.



- **Proportions of tables**

  In the **Proportion of sensitive tables** section, you can view a pie chart that displays the numbers and proportions of tables at the low, medium, high, and unknown risk levels.

- **Top five rules that are hit**

  In the **Hit Rule Top5** section, you can view the top five rules that are hit and the number of times that each rule is hit.

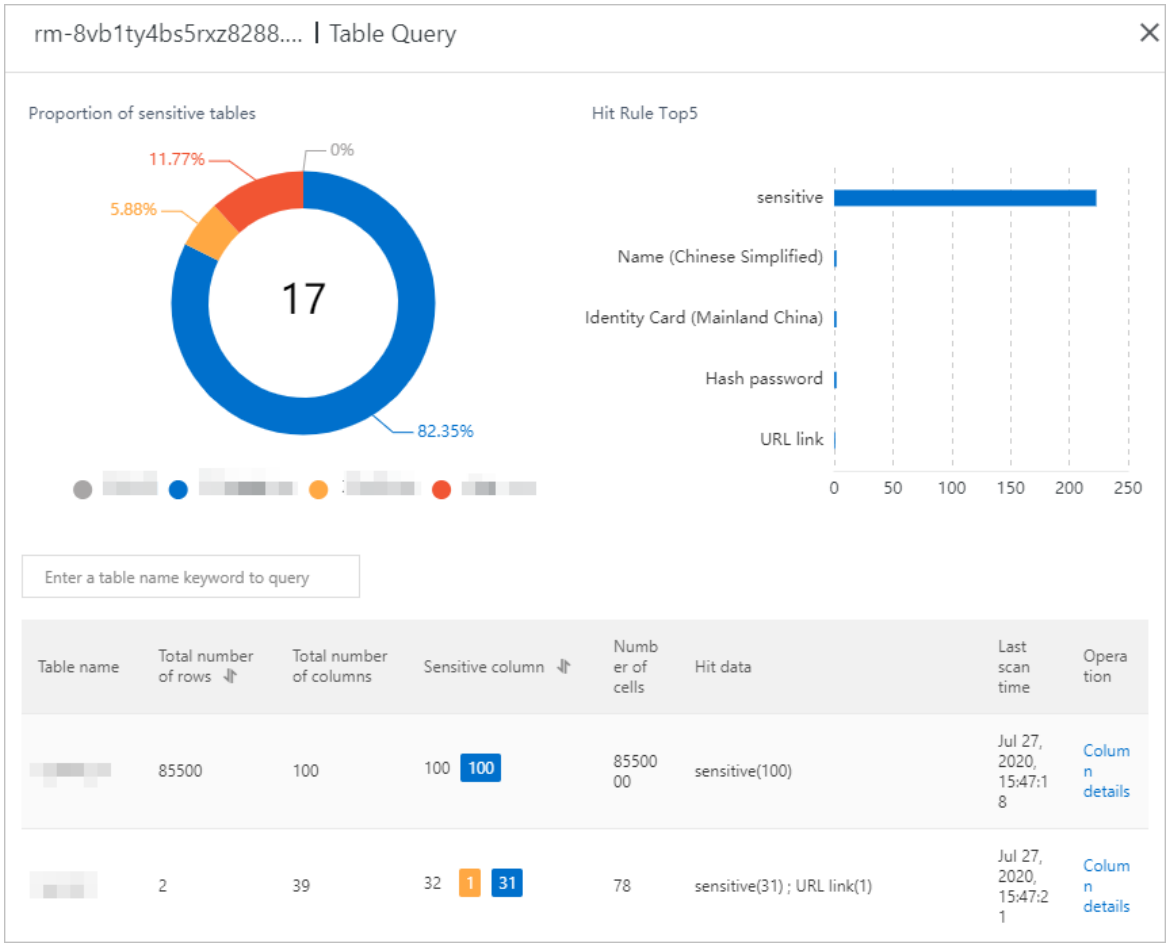
- **List of tables with sensitive data detected**

In the table list, you can view the information about the tables with sensitive data detected. For example, you can view the table name, total number of rows, total number of fields, number of sensitive fields, and rules that are hit. You can also click **Column details** in the Operation column to view the field details, including the fields that hit sensitive data detection rules and the risk levels of the sensitive fields.

View sensitive data detected in Tablestore

OTS refers to Tablestore.

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Sensitive data assets**.
3. On the **Sensitive data assets** page, click the **OTS** tab.
4. On the **OTS** tab, find the Tablestore instance whose details you want to view and click **Table details** in the Actions column.
5. In the Table Query panel, you can view the proportions of sensitive tables, top five rules that are hit, and the list of tables with sensitive data detected.



- **Proportions of tables**

  In the **Proportion of sensitive tables** section, you can view a pie chart that displays the numbers and proportions of tables at the low, medium, high, and unknown risk levels.

- **Top five rules that are hit**

In the **Hit Rule Top5** section, you can view the top five rules that are hit and the number of times that each rule is hit.

- **List of tables with sensitive data detected**

  In the table list, you can view the information about the tables with sensitive data detected. For example, you can view the table name, total number of rows, total number of fields, number of sensitive fields, and rules that are hit. You can also click **Column details** in the Operation column to view the field details, including the fields that hit sensitive data detection rules and the risk levels of the sensitive fields.

References

- Sensitive data scan and detection
- Supported sensitive data
- Supported OSS files

5.2. Export sensitive data

Data Security Center (DSC) helps you detect sensitive data in data assets, and allows you to export the sensitive data in the CSV format. Then, you can view the sensitive data of buckets, tables, or collections in the exported CSV file.

Procedure

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Sensitive data assets**.
3. On the page that appears, click the tab of the data source for which you want to export sensitive data based on your needs.
4. Specify the filter conditions that are used to export sensitive data.

   The filter conditions vary based on the data source. In this example, Distributed Relational Database Service (DRDS) is used as the data source. DSC provides the following filter conditions for DRDS: **Region**, **Instance Name**, **Database type**, **Sensitivity level**, **Start time**, and **End time**. Sensitivity levels include N/A, S1, S2, S3, and S4.
5. Click **Export**.
   After the sensitive data is exported as a file, the file is stored in the default storage directory of your local disk. For example, the default storage directory for Windows is *This PC > Downloads*. By default, the file is named in the format of *instance_Date when the file is exported_Time when the file is exported*. The date format is YYYYMMDD. The time is measured in milliseconds, and the value is a 13-digit timestamp.

5.3. Query sensitive data

The Sensitive data search page displays all the sensitive data that has been detected in your data assets. You can specify one or more types of sensitive data to query and view the distribution of the queried sensitive data across your data assets.

Procedure

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Sensitive data search**.

3. On the **Sensitive data search** page, set filters as required.

   You can set the following filters:

   - **Hit data**: the type of sensitive data. You can select multiple types at a time, such as email address and mobile number.
   - **Enter file name to search/Enter table name to search**: the name of the object or table with sensitive data detected. Fuzzy match is supported.
   - **Region**: the region where the data asset resides.
   - **Bucket/Instance/Project**: the name of the bucket, database, instance, or project with sensitive data detected.
   - Time range: the beginning and end of the time range to query.

   > **Note** If you set multiple filters, SDDP displays the sensitive data that meets all the specified filters.

4. Click **Search**.

Related operations

- **Query sensitive data by risk level**

  On the **OSS-file** tab, set the **Sensitivity level** parameter to S1, S2, or S3 to query sensitive data at the specified risk level.

- **Sort sensitive data based on the total number of rows or sensitive fields in ascending or descending order**

  On a specific tab such as the **RDS-table** tab, click the icon to the right of **Total number of rows** or **Sensitive column** to sort sensitive data based on the total number of rows or sensitive fields in ascending or descending order. The data is sorted in descending order after you click the icon for the first time, and is sorted in ascending order the next time you click the icon.

- **View the details of sensitive data**

  Find the specific data asset and click **Details** or **Column details** in the Operation column. The **Hit query** panel appears if the data asset is an Object Storage Service (OSS) bucket, and the **Column details** panel appears if the data asset is a table. In the Hit query or Column details panel, you can view the following information about the sensitive data detected in the data asset:

  - **Column name**: the name of the sensitive field detected in the table.

    > **Note** This parameter is displayed only in the **Column details** panel for a table in an ApsaraDB RDS database, MaxCompute project, self-managed database hosted on an Elastic Compute Service (ECS) instance, DRDS database, PolarDB database, or Tablestore instance. The **Hit query** panel for an OSS object does not display this parameter.

  - **Hit Rules**: the type and name of the sensitive data detection rule that is hit.
  - **Sensitivity level**: the risk level of the detected sensitive data.

- **Number of hits**: the number of times that the sensitive data detection rule is hit in the object.

  > **Note** This parameter is displayed only in the **Hit query** panel for an OSS object.

- **Data sampling**: the sensitive fields detected by DSC. You can specify the number of samples to be collected by setting the **Sensitive data sampling** parameter on the **Cloud hosting** page. You can set this parameter to 0, 5, or 10. Then, SDDP displays the collected samples based on your setting.

5.4. Manage scan tasks

Data Security Center (DSC) automatically scans the authorized data assets for sensitive data. On the Identify task monitoring page, you can view the details of scan tasks that are created by DSC to scan authorized data assets and rerun the scan tasks as required.

Context

DSC allows you to manage scan tasks that scan your data stored in Object Storage Service (OSS), ApsaraDB RDS, DRDS, PolarDB, Tablestore, self-managed databases hosted on Elastic Compute Service (ECS) instances, and MaxCompute.

After you authorize DSC to access specific data assets, DSC automatically creates and runs scan tasks for these data assets to detect sensitive data. By default, the **automatic scan** feature is enabled for scan tasks that are created by DSC. This feature allows DSC to run a full scan on authorized data assets and then scan the data that is newly written to or modified in these data assets at intervals of 4 hours. In addition, after you create or modify a sensitive data detection rule, DSC automatically reruns scan tasks.

View the details of scan tasks

On the **Identify task monitoring** page, you can view the details of each scan task, such as the related data asset, task status, and time when the task was complete. To view the details of scan tasks, perform the following steps:

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Identify task monitoring**.
3. Click the tab of the service for which you want to view the details of scan tasks.
4. (Optional)Select the region, enter the name of the data asset, specify the beginning and end of the time range to query, and then click **Search**.
5. In the task list, view the details of each scan task, such as the related data asset, task status, and time when the task was complete.

Rescan your data assets

You can rerun scan tasks in the following scenarios:

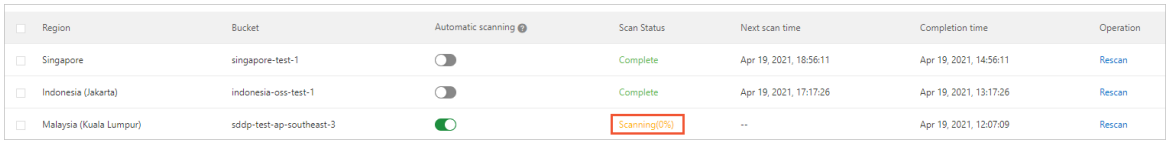
- If you have not enabled the **automatic scan** feature for a scan task, the scan task is not run after it is created. You must rerun the scan task to start it.
- If you enable the **automatic scan** feature for a scan task, DSC automatically reruns the scan task to scan the data that is newly written to or modified in the data asset at intervals of 4 hours. You can also rerun the scan task to scan the data asset for sensitive data immediately after you modify the data in the data asset.

To rescan a data asset for sensitive data, perform the following steps:

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Identify task monitoring**.
3. Click the tab of the service for which you want to rescan data.
4. Find the data asset that you want to rescan and click **Rescan** in the Operation column.
5. In the **Confirm rescan** message, click **OK**.

   Generally, the rescan can be complete within 10 minutes. Wait until the data asset is scanned.

   After the rescan is started, the scan task enters the **Scanning** or **Waiting** state. The percentage that is displayed in the **Scan Status** column indicates the progress of the scan task. After the rescan is complete, the scan task enters the **Complete** state.



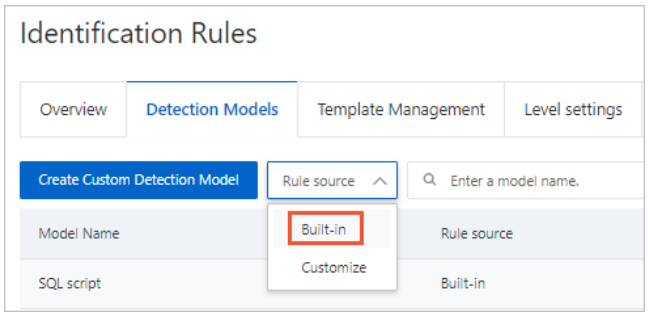
5.5. Manage detection models

Detection models define rules on how to detect sensitive data in your assets. Data Security Center (DSC) provides built-in detection models and allows you to customize models. You can use these models to create your own methods of sensitive data detection. This topic describes how to view built-in detection models and create, edit, or delete custom detection models.

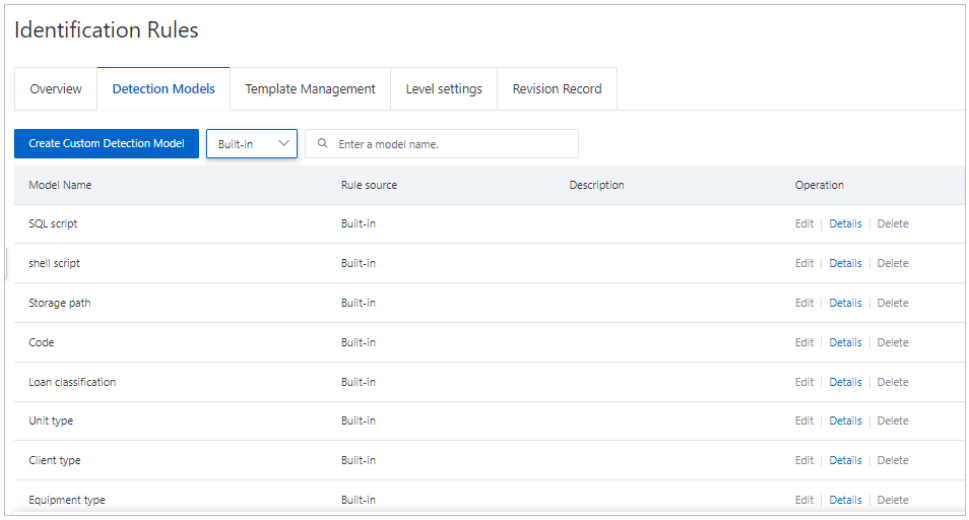
View built-in detection models

The built-in sensitive data detection models provided by DSC apply to regular sensitive data, such as mobile numbers and ID card numbers. You can view the model names, sensitivity levels, and rule information of the built-in detection models provided by DSC. To view the built-in models provided by DSC, perform the following steps:

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Identification Rules**.
3. On the **Identification Rules** page, click the **Detection Models** tab.
4. Select **Built-in** from the **Rule source** drop-down list.



5. View the list of built-in detection models.



You can view information on built-in detection models, such as the model names.

6. To view the details of a specific built-in detection model, find the model and click **Details** in the Operation column.

> **Note** You cannot edit or delete built-in detection models.

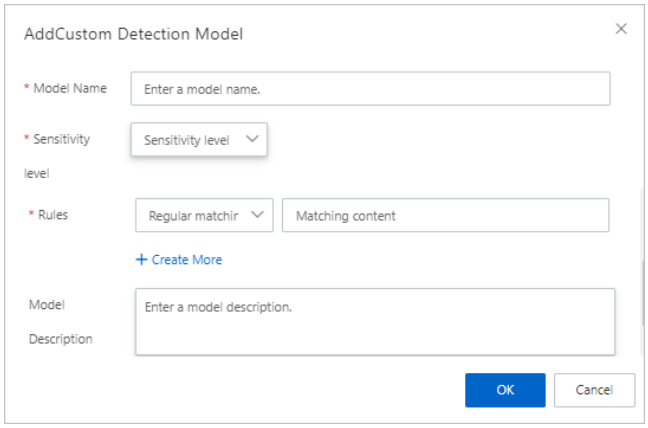
7. In the **ViewCustom Detection Model** dialog box, view the details of the built-in detection model.

   You can view the model name, sensitivity level, and rule information of the model.

Create a custom detection model

DSC detects sensitive data in objects or tables and generates alerts based on sensitive data detection rules defined in detection models. If the built-in detection models cannot meet your business requirements, perform the following steps to create a custom detection model:

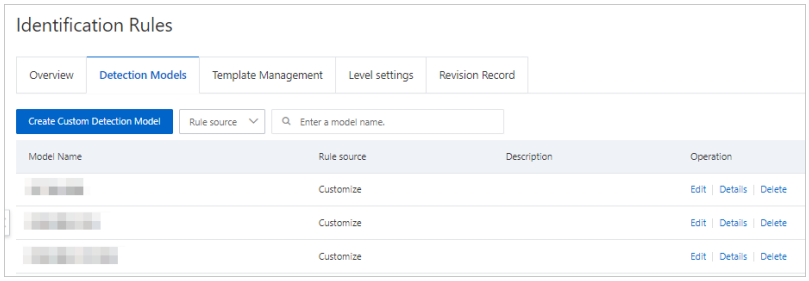
1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Identification Rules**.
3. On the **Identification Rules** page, click the **Detection Models** tab.
4. On the **Detection Models** tab, click **Create Custom Detection Model**.
5. In the **AddCustom Detection Model** dialog box, set the following parameters.



Parameter	Description
Model Name	The name of the custom detection model.
Sensitivity level	The sensitivity level of the sensitive data that is detected based on the rules defined by the custom detection model. Valid values: ○ **S1**: level 1 sensitive data ○ **S2**: level 2 sensitive data ○ **S3**: level 3 sensitive data ○ **S4**: level 4 sensitive data ○ **S5**: level 5 sensitive data **Note** A larger suffix indicates a higher sensitivity level. S5 indicates the highest sensitivity level.

Parameter	Description
Rules	The rules on how to detect sensitive data. Valid values: ○ **Regular matching**: uses a regular expression to detect sensitive data. Examples: ■ *Exampleoo+a*: Data such as Exampleooa, Exampleoooa, and Exampleooooooa is detected as sensitive. The plus sign (*+*) indicates one or more repetitions of the preceding character. ■ *Exampleoo*a*: Data such as Exampleoa, Exampleooa, and Exampleooooooa is detected as sensitive. The asterisk (***) indicates zero or more repetitions of the preceding character. ■ *Exampleo?a*: Data such as Examplea and Exampleoa is detected as sensitive. The question mark (*?*) indicates zero or one repetition of the preceding character. ○ **Does not contain**: detects data that does not contain the specified keyword. ○ **Contains**: detects data that contains the specified keyword. You can create multiple detection rules in a detection model. To create multiple detection rules, click **Create More**. **Notice** ○ If a custom model defines multiple rules, data is detected as sensitive only if the data meets all the rules of the model. ○ The **Does not contain** rules can be used to reduce false positives. We recommend that you use this type of rules together with other rules. ○ The built-in models provided by DSC apply to mobile numbers and ID card numbers. We recommend that you check whether the rules that you want to define have been covered by the built-in models provided by DSC before you create a custom model. For more information, see View built-in detection models.
Model Description	The description of the custom detection model.

6. Click **OK**.
   After you create the detection model, you can view the information of the model in the model list.

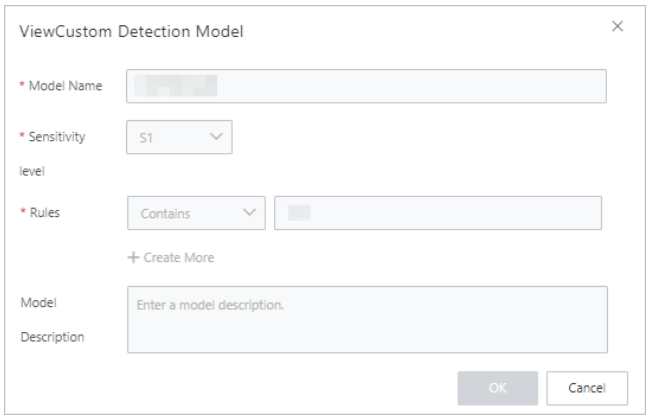


View, edit, and delete a custom detection model

DSC allows you to view, edit, and delete custom detection models. This section describes how to view, edit, and delete a custom detection model.

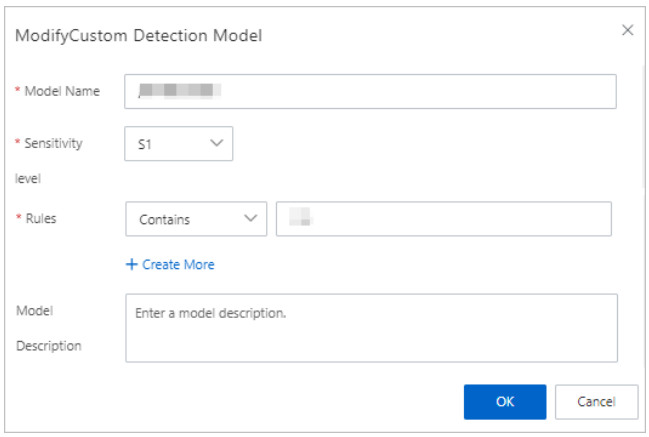
1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Sensitive data discovery > Identification Rules**.
3. On the **Identification Rules** page, click the **Detection Models** tab.
4. Select **Customize** from the **Rule source** dialog box.
5. Find the custom detection model that you want to manage and perform the following operations:
   - **View the details of the custom detection model**

     Click **Details** in the Operation column. In the **ViewCustom Detection Model** dialog box, view the details of the custom detection model.

   - **Edit a custom detection model**

Click **Edit** in the Operation column. In the **ModifyCustom Detection Model** dialog box, modify the parameters and click **OK**. For more information about the parameters, see Parameter description.



Notice If the custom detection model is used by a sensitive data detection template that is enabled, the modification takes effect the next time when DSC scans data. The sensitive data that was detected based on the original model is not affected.

- **Delete a custom detection model**

  Click **Delete** in the Operation column. In the message that appears, click **OK**.

  Notice

  - Delete a custom detection model with caution. After you delete a custom detection model, DSC cannot use this model to detect sensitive data.
  - After you delete a custom detection model, the sensitive data that was detected based on the model is not affected.

6.Data desensitization

6.1. Perform static de-identification

Data Security Center (DSC) allows you to create de-identification tasks to de-identify and protect sensitive data in your data assets. This topic describes how to create and query de-identification tasks.

Prerequisites

DSC is authorized to access your data assets, such as MaxCompute projects, Object Storage Service (OSS) buckets, and ApsaraDB RDS databases. For more information about access authorization, see Grant access to data assets.

Context

DSC supports both static and dynamic de-identification.

- Compared with static de-identification, dynamic de-identification is more flexible and allows you to de-identify specified sensitive data. The size of sensitive data that can be dynamically de-identified at a time must be less than 2 MB. For more information about dynamic de-identification, see Perform dynamic de-identification.
- The static de-identification feature of DSC allows you to use de-identification algorithms to redact, encrypt, or substitute sensitive data detected in authorized data assets and store de-identified data in the location that you specify. For more information, see Supported data de-identification algorithms.

> **Note** DSC allows you to perform static de-identification on OSS objects, ApsaraDB RDS tables, MaxCompute tables, PolarDB tables, and ApsaraDB for OceanBase tables. For more information, see Supported data assets.

Create a de-identification task

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data desensitization > Static Desensitization**.
3. On the **Static Desensitization** page, click **Add Desensitization Task**.
4. Perform the steps in the wizard to create a custom de-identification task.
   i. Enter basic task information and click **Next**.

   > **Note** You can specify a custom task name.

   ii. Specify the type of the data asset that contains the sensitive data to be de-identified, set other parameters as required, and then click **Next**.

- The following table describes the parameters that you need to set if you set the Types of data storage parameter to **RDS table/DRDS table/MaxCompute table/PolarDB table/OceanBase table/ADB-MySQL table**.

Parameter	Description
Types of data storage	The type of the data asset that contains the sensitive data to be de-identified. Set this parameter to **RDS table/DRDS table/MaxCompute table/PolarDB table/OceanBase table/ADB-MySQL table**.
Source Product	The Alibaba Cloud service that contains the sensitive data to be de-identified. Valid values: **RDS**, **DRDS**, **OceanBase**, **MaxCompute**, **ADB-MySQL**, and **PolarDB**.
Source Database/Project	Required. The project or database for storing the table that contains the sensitive data to be de-identified.
Source Table Name	Required. The table that contains the sensitive data to be de-identified.
Source Partition	Optional. The partition that contains the sensitive data to be de-identified. For more information about how to specify a partition, see Specify a partition. You can configure partitions when you create a MaxCompute table. Partitions define different logical divisions of a table to help you efficiently query specific content. For more information, see Partition. **Note** ■ If you set the **Source Product** parameter to RDS or PolarDB, you do not need to set the **Source Partition** parameter. ■ If you leave the **Source Partition** parameter unspecified, DSC de-identifies sensitive data in all partitions of the table.
Sample SQL	Optional. The SQL statement that specifies the sensitive data to be de-identified. If you leave this parameter unspecified, DSC de-identifies all sensitive data in the table. **Note** If you set the **Source Product** parameter to MaxCompute or PolarDB, you do not need to set the **Sample SQL** parameter.

- The following table describes the parameters that you need to set if you set the Types of data storage parameter to **OSS files**.

Parameter	Description
Types of data storage	The type of the data asset that contains the sensitive data to be de-identified. Set this parameter to **OSS files**.
File source	The source of the file that contains the sensitive data to be de-identified. Valid values: **Uploaded Local File** and **OSS Bucket**. **Note** You can upload only a TXT, CSV, XLSX, or XLS file.
OSS Bucket where the source file is located	The OSS bucket in which the source file resides. You can enter keywords to search for the OSS bucket in which the source file resides.
Source file names	The name of the source file. The file name must contain a file name extension. Only TXT, CSV, XLSX, and XLS files are supported. To specify multiple source files of the same format at a time, turn on **Open the pass**. **Note** After you turn on the switch, you can use the wildcard (*) to specify multiple source files at a time. You can use the wildcard only in the file name prefix, for example, *test*.xls*. DSC applies the same de-identification rules to all the source files that are specified at a time. Make sure that the source files share the same schema.
Source file description	The description of the source file. If you set the **File source** parameter to **OSS Bucket**, you do not need to set this parameter.
Separator selection	Optional. The field delimiter. This parameter is required for *CSV* and *TXT* files. Select a field delimiter based on the file format. Valid values: ■ **Semicolon ";" (macOS/Linux default)** ■ **Comma "," (Windows default)**
Table contains header rows	Optional. Specifies whether the sensitive data to be de-identified includes the table header.

iii. Configure a de-identification algorithm and click **Next**.

In the field list, turn on **Desensitization** for each field that requires de-identification and select a de-identification algorithm as required. For more information, see Configure de-identification algorithms.

> **Note**
> - You can click **View and Modify Parameters** next to the selected algorithm to view or modify the rule of the algorithm.
> - A field is de-identified only after Desensitization is turned on for this field.

iv. Enable the data watermarking feature.

After you add a watermark to the data that you distribute, you can trace the source of potential data leaks. For more information about limits on using watermarks, see Limits on using watermarks.

v. Specify the location for storing the de-identified data. Click **Test** to check whether DSC can write data to the specified location. After the test is passed, click **Next**.

vi. Configure the processing logic.

Parameter	Description
How the task is triggered	The mode in which the de-identification task is run. Valid values: ■ **Manual Only**: You must manually run the de-identification task on the Static Desensitization page. ■ **Scheduled Only**: The de-identification task is automatically run at the specified time on an hourly, daily, or monthly basis. ■ **Manual + Scheduled**: The de-identification task is automatically run at the specified time on an hourly, daily, weekly, or monthly basis. You can also manually run the de-identification task as required.
Turn on incremental desensitization	Optional. Specifies whether to enable incremental de-identification. If you turn on this switch, only the data that is added after the last de-identification task is complete is de-identified. You must specify a field whose value increases over time as the incremental identifier. For example, you can specify the creation time field or the auto-increment ID field as the incremental identifier. **Note** DSC supports incremental de-identification only for ApsaraDB RDS databases.

Parameter	Description
Shard field	Optional. The field based on which DSC divides the source data to multiple shards and concurrently de-identifies them. This improves the de-identification efficiency. You can specify one or more shard fields as required. **Note** ■ DSC supports incremental de-identification only for ApsaraDB RDS databases. We recommend that you use the primary key or a unique index as the shard field. ■ If you do not specify a shard field, DSC uses the primary key as the shard field by default to de-identify the source data. You must specify a shard field if the source data does not have a primary key. Otherwise, the de-identification task fails. ■ The query performance and data accuracy may deteriorate if you specify excessive shard fields. Exercise caution when you set this parameter.
Table name conflict resolution	The solution used if the specified destination table exists. Valid values: ■ *Delete the target table and create a new table with the same name.* ■ *Attach data to the target table*: We recommend that you select this option.
Row Conflict Resolution	The solution used if conflicting rows exist in the destination table. Valid values: ■ *Keep conflicting rows in the target table and discard the new data.*: We recommend that you select this option. ■ *Delete conflicting rows in the target table and insert the new data.*

vii. Click **Submit**.
The de-identification task is displayed on the **Desensitization task configuration** tab of the **Static Desensitization** page.

5. Run the de-identification task.

   After the de-identification task is created, find the de-identification task, turn on the switch, and then click **Start** in the Actions column to run the de-identification task.

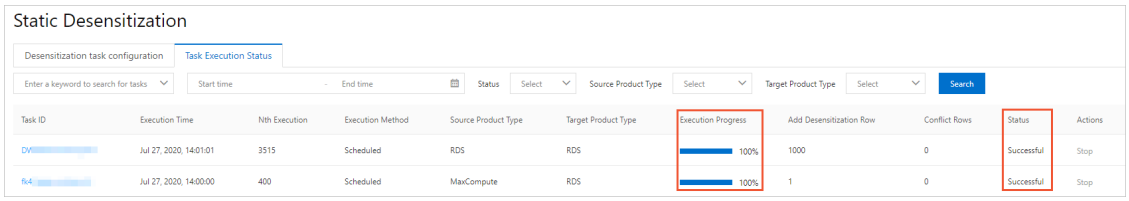
   Note DSC allows you to modify or delete a de-identification task after it is created. However, you cannot modify or delete a de-identification task that is running.

6. View the progress and status of the de-identification task.

   After the de-identification task is started, click the **Task Execution Status** tab. On the **Task Execution Status** tab, click **Search** to update the progress and status of tasks.

> **Note** If you do not click **Search**, the created de-identification task may not be displayed on the **Task Execution Status** tab.

You can check whether the de-identification task is run in the **Status** column of the **Task Execution Status** tab. For more information about the failure causes, see Troubleshoot failures to run de-identification tasks.



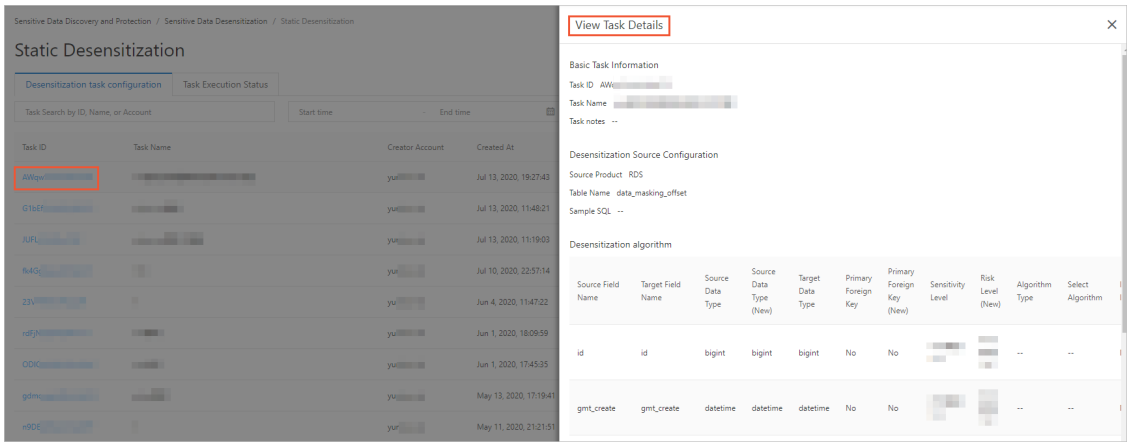
Specify a partition

Partition	Format	Example
N weeks after the specified date	Custom partition field=$[yyyymmdd+7*N]	time=$[20190710+7*1]. It indicates that the data generated in the week after July 10, 2019 is to be de-identified.
N weeks before the specified date	Custom partition field=$[yyyymmdd-7*N]	time=$[20190710-7*3]. It indicates that the data generated in the three weeks before July 10, 2019 is to be de-identified.
N days after the specified date	Custom partition field=$[yyyymmdd+N]	time=$[20190710+2]. It indicates that the data generated in the two days after July 10, 2019 is to be de-identified.
N days before the specified date	Custom partition field=$[yyyymmdd-N]	time=$[20190710-5]. It indicates that the data generated in the five days before July 10, 2019 is to be de-identified.
N hours after the specified time	Custom partition field=$[hh24mi:ss+N/24]	time=$[0924mi:ss+N/24]. It indicates that the data generated in the two hours after 09:00:00 in the 24-hour clock is to be de-identified.
N hours before the specified time	Custom partition field=$[hh24mi:ss-N/24]	time=$[0924mi:ss-1/24]. It indicates that the data generated in the hour before 09:00:00 in the 24-hour clock is to be de-identified.

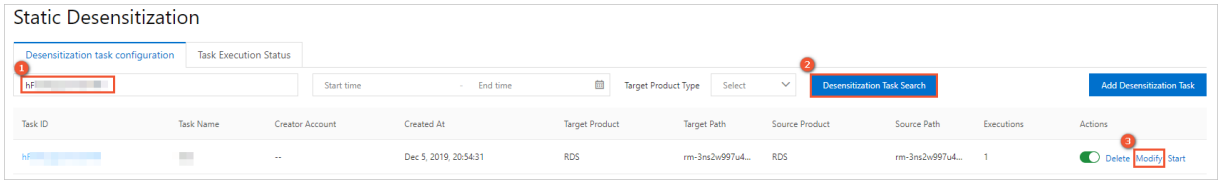
Partition	Format	Example
N minutes after the specified time	Custom partition field=$[hh24mi:ss+N/24/60]	time=$[0924mi:ss+2/24/60]. It indicates that the data generated in the two minutes after 09:00:00 in the 24-hour clock is to be de-identified.
N minutes before the specified time	Custom partition field=$[hh24mi:ss-N/24/60]	time=$[0924mi:ss-2/24/60]. It indicates that the data generated in the two minutes before 09:00:00 in the 24-hour clock is to be de-identified.

Query de-identification tasks

On the **Static Desensitization** page, you can view the details of de-identification tasks that you create. In the task list, click the ID of a de-identification task in the **Task ID** column to view the task details.



To modify an existing de-identification task, enter the ID of the task in the search box and click **Desensitization Task Search**. The de-identification task is displayed. Then, click Modify in the Actions column to modify the de-identification task.



Troubleshoot failures to run de-identification tasks

Error message	Description
The desensitization task does not exist. The task may have been deleted or closed.	The error message returned because the de-identification task is deleted or disabled. If the switch in the **Actions** column is turned off for the de-identification task, the task is disabled.

Error message	Description
Incorrect recurrence configuration of the scheduled task.	The error message returned because the time specified for running the de-identification task daily is invalid.
The desensitization source instance does not exist.	The error message returned because the instance that contains the source table does not exist.
The desensitization target instance does not exist.	The error message returned because the instance is deleted or the permissions to access the instance are revoked.
The desensitization source table does not exist.	The error message returned because the table is deleted or the permissions to access the instance that contains the table are revoked.
Incorrect desensitization algorithm parameter.	The error message returned because the parameters of the de-identification algorithm are invalid.
Empty source table list.	The error message returned because no data exists in the partition column of the source table.
Failed to write data to the target table.	The error message returned because DSC fails to write data to the destination table that you specify.
Failed to query the source table.	The error message returned because the specified data is not found in the source table.
Failed to create the target table.	The error message returned because the destination table does not exist in the specified location.
No primary key has been found.	The error message returned because the primary key is missing in the ApsaraDB RDS source table.
Incorrect ODPS partition field configured for the task.	The error message returned because the specified source or destination partition is invalid when you create the de-identification task.

6.2. Perform dynamic de-identification

You can call the ExecDatamask operation to dynamically de-identify sensitive data.

Context

You can call the ExecDatamask operation to dynamically de-identify sensitive data. When you call this operation, you must specify the ID of a de-identification template. De-identification templates can be used for both static de-identification and dynamic de-identification. You can view the IDs of de-identification templates on the **Desensitization Template** page in the DSC console. You can also create de-identification templates. For more information, see Manage de-identification templates.

Limits

When you call the ExecDatamask operation to dynamically de-identify sensitive data, the size of the sensitive data that is specified by the `Data` parameter must be less than 2 MB.

View the call history of the ExecDatamask operation

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data desensitization > Dynamic desensitization**.
3. On the Dynamic desensitization page, view the call history of the ExecDatamask operation.

   The following table describes the parameters about each call record.

Parameter	Description
Operation for Dynamic De-identification	The name of the API operation.
UID	The unique identifier (UID) of the Alibaba Cloud account or RAM user that was used to call the API operation.
IP address	The IP address of the user who initiated the call.
First call time	The time when the API operation was first called.
Last call time	The time when the API operation was last called.
Cumulative number of calls	The total number of times the API operation was called.

Note Only one record is generated for calls that were initiated by the same account from the same IP address. In this case, the cumulative number of calls is recorded.

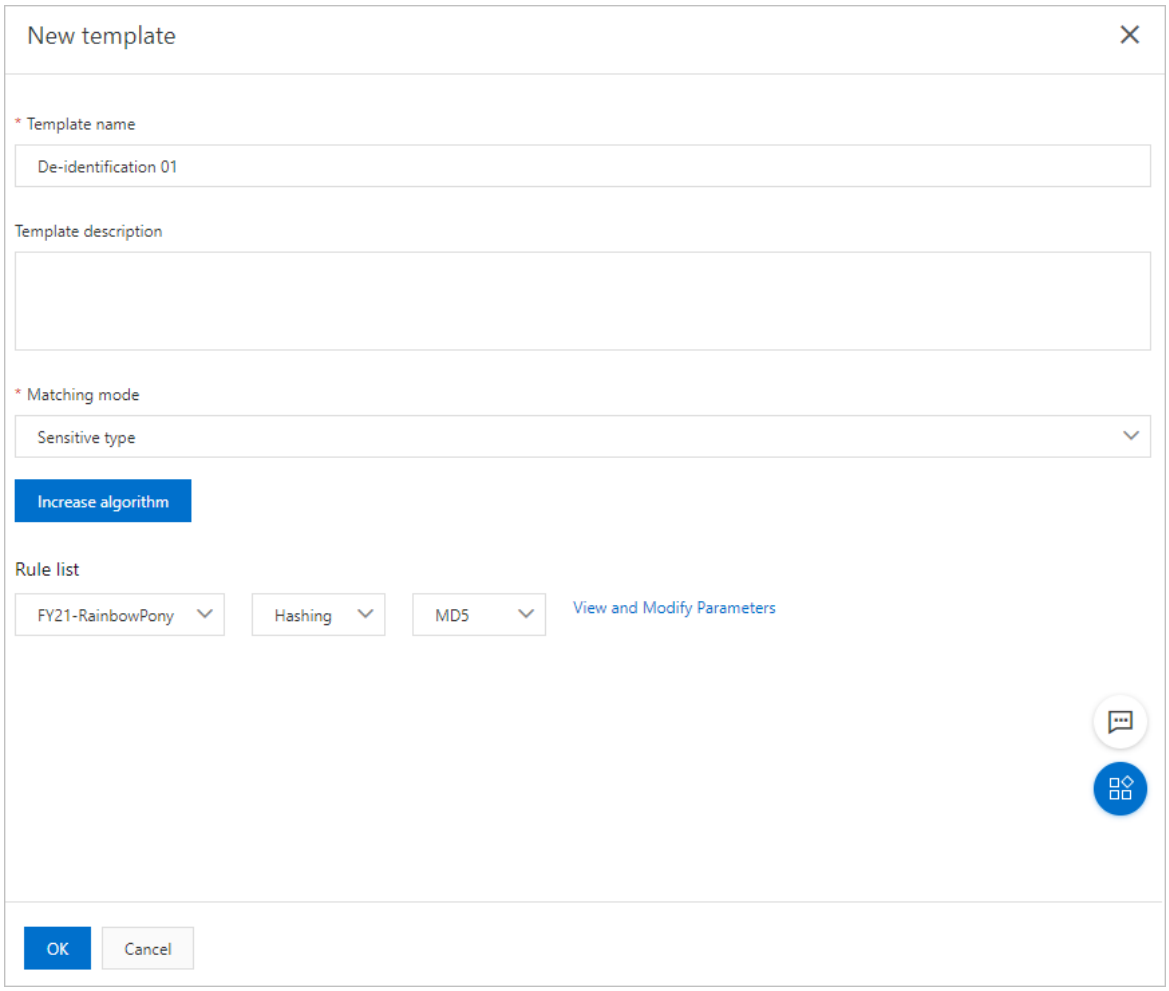
6.3. Manage de-identification templates

Data Security Center (DSC) supports custom de-identification templates. You can create a de-identification template and add de-identification algorithms that are frequently used in the same scenario to the template. This avoids repeated configuration of de-identification algorithms and improves the efficiency in sensitive data processing. This topic describes how to create and manage de-identification templates.

Create a de-identification template

You can create an unlimited number of de-identification templates.

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data desensitization > Desensitization Template**.
3. On the **Desensitization Template** page, click **New template**.
4. In the **New template** panel, set the parameters as required. The following table describes the parameters for creating a de-identification template.



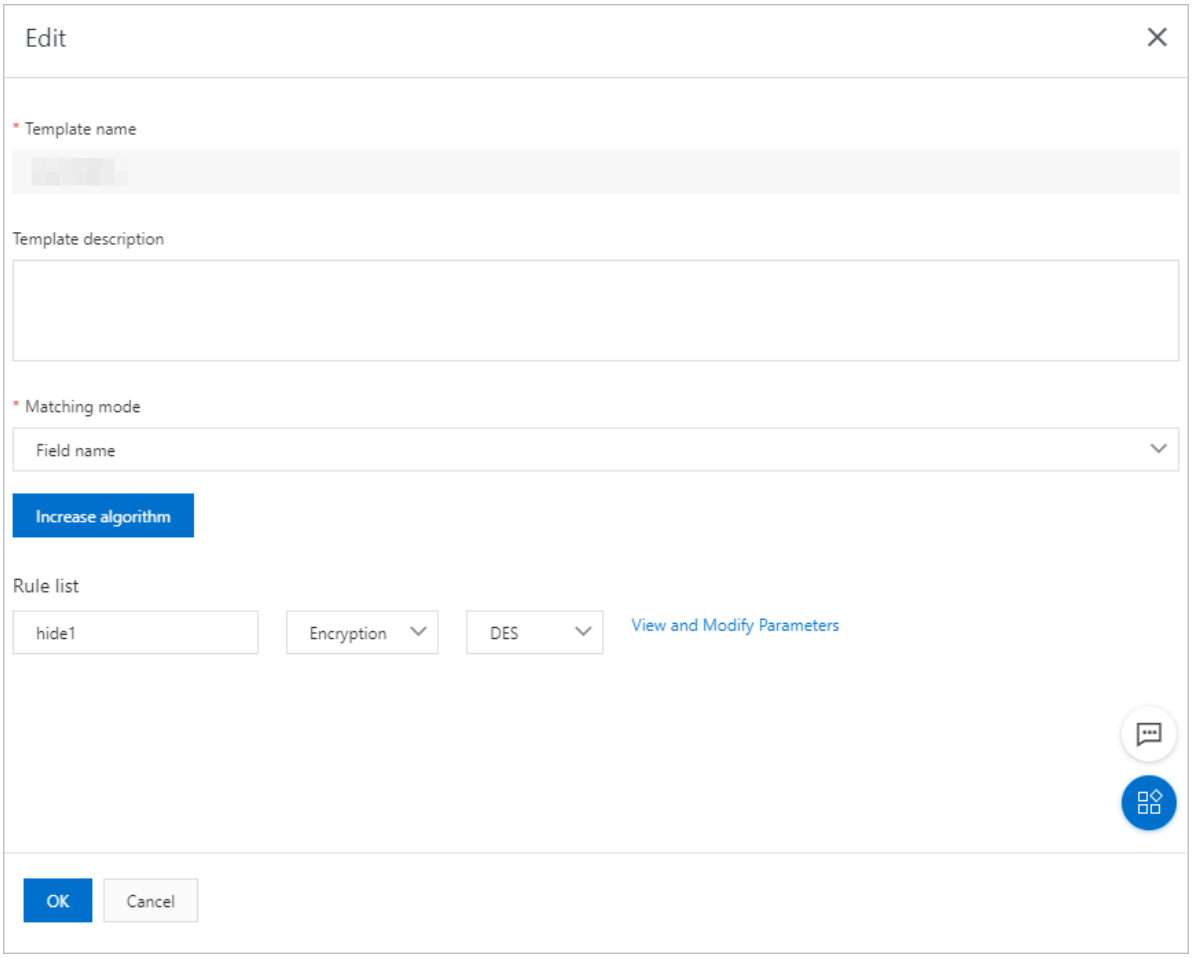
Parameter	Description
Template name	The name of the de-identification template.
Template description	The description of the de-identification template. You can enter information such as the scenario to which the de-identification template is applied.

Parameter	Description
Matching mode	The mode in which the de-identification template finds its matched sensitive data. Valid values: ○ **Sensitive type**: If you select this option, you must select a sensitive data type that is supported by DSC, such as vehicle identification number and unified social credit code, and a de-identification algorithm for the sensitive data type. ○ **Field name**: If you select this option, you must specify the field to be de-identified and select a de-identification algorithm for the field.
Rule list	Select a sensitive data type or enter a field to be de-identified and specify a de-identification algorithm. DSC supports the following de-identification algorithms: ○ Hashing ○ Redaction ○ Substitution ○ Rounding ○ Encryption ○ Shuffling ○ Data decryption For more information, see Supported data de-identification algorithms. You can configure multiple rules in a template. To configure more rules, click **Increase algorithm**.

Manage de-identification templates

- **Modify a de-identification template**

To update the description or rules of a de-identification template, find the template on the **Desensitization Template** page and click **Edit** in the Actions column. In the **Edit** panel, modify the description or rules of the template.



- **Delete a de-identification template**

  To delete a de-identification template that is no longer applicable to the current business scenario, find the template on the **Desensitization Template** page and click **Delete** in the Actions column.

  > **Note** Deleted templates cannot be recovered. Exercise caution when you delete templates.

6.4. Configure de-identification algorithms

This topic shows you how to configure de-identification algorithms and provides related examples.

Context

Data Security Center (DSC) supports hashing, redaction, substitution, rounding, encryption, data decryption, and shuffling. For more information, see Supported data de-identification algorithms.

Procedure

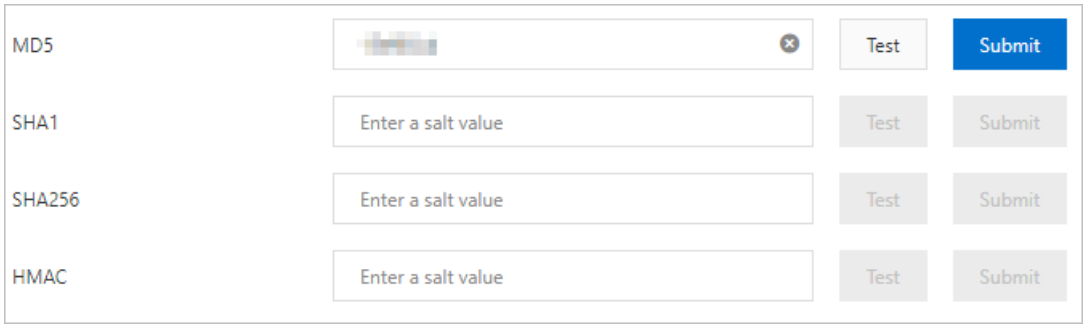
1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data desensitization > Desensitization algorithm**.

3. Click the tab for the de-identification algorithm that you want to use for static de-identification.
4. Configure the de-identification algorithm.

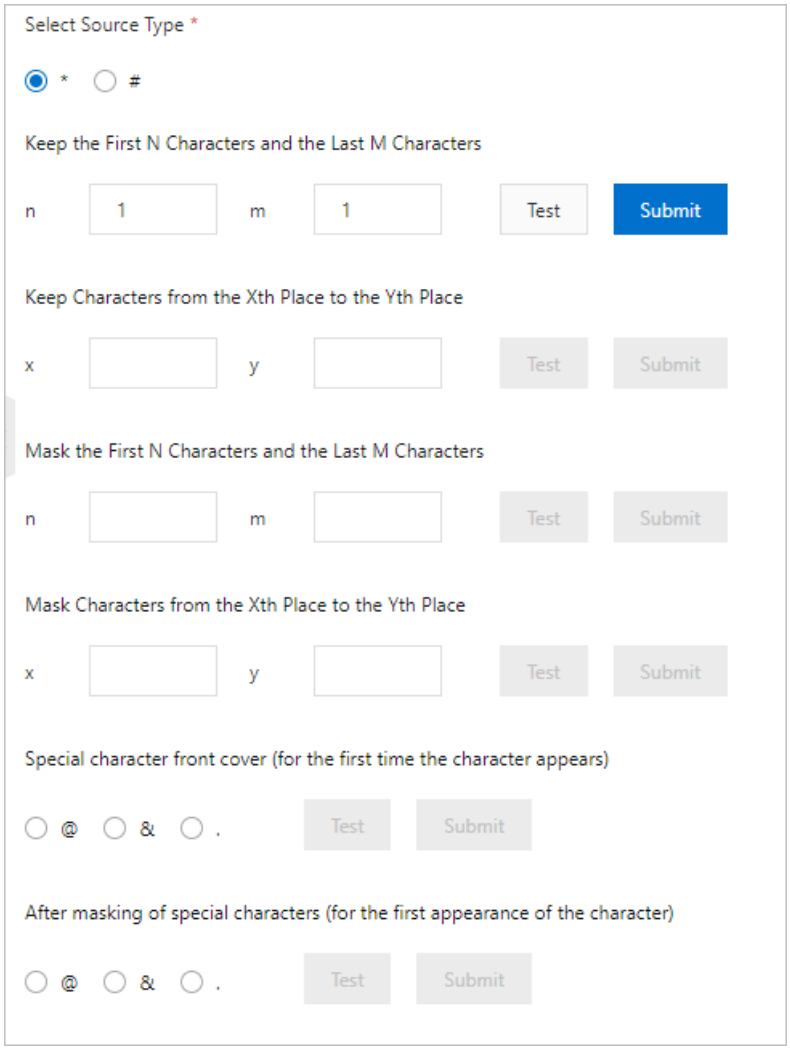
   You can configure each de-identification algorithm in the following way:

   - **Hashing**: Set a salt value for each encryption algorithm.

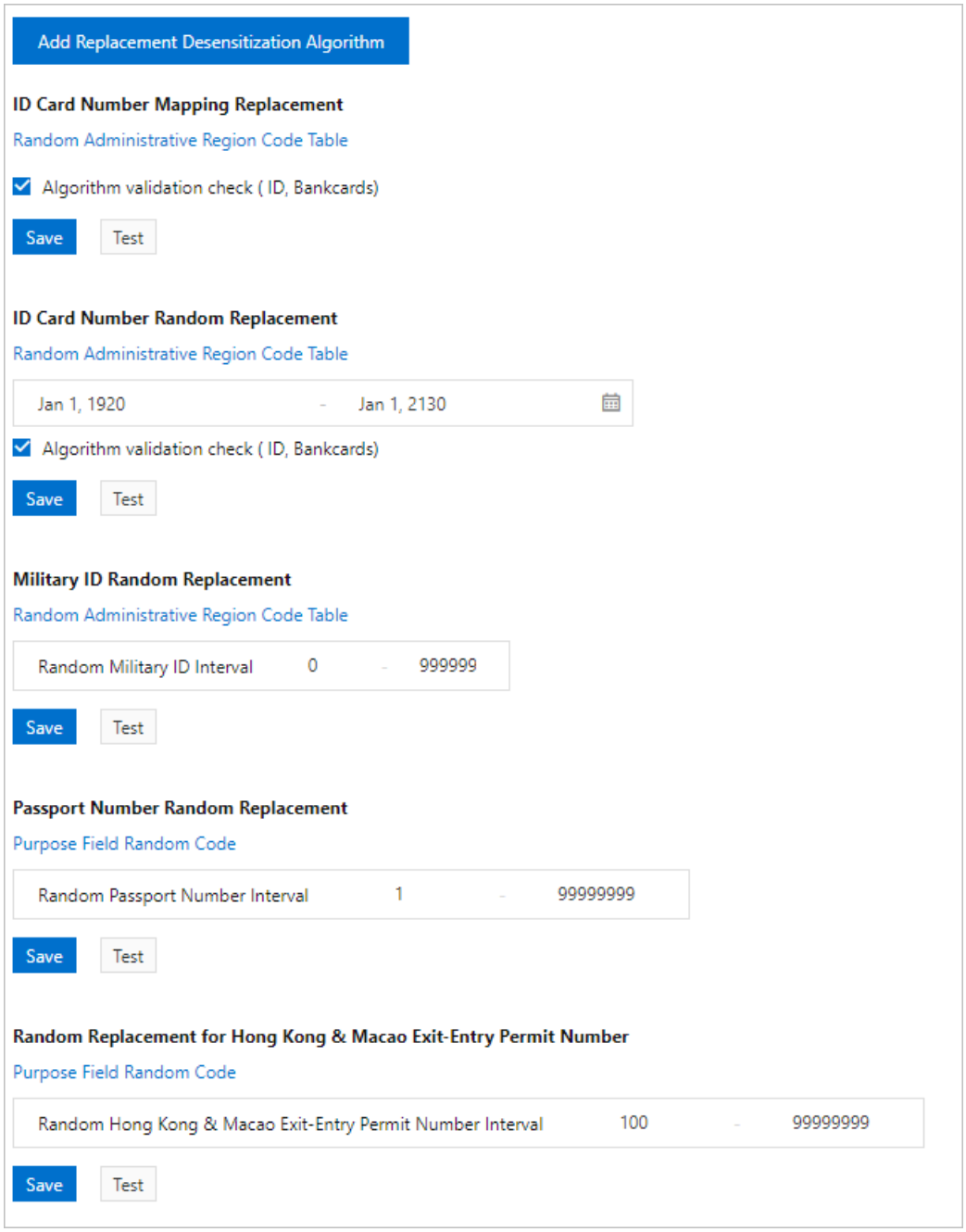
     > **Note**
     >
     > In cryptography, you can insert a specific string to a fixed position in a password so that the hash value of the new password is different from that of the original password. This process is called **salting**. A salt value is the specific string that you insert.

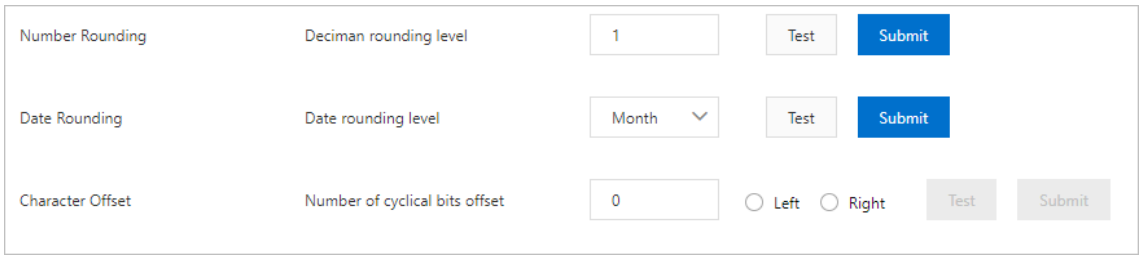
   - **Masking**: Set parameters for the redaction algorithm.



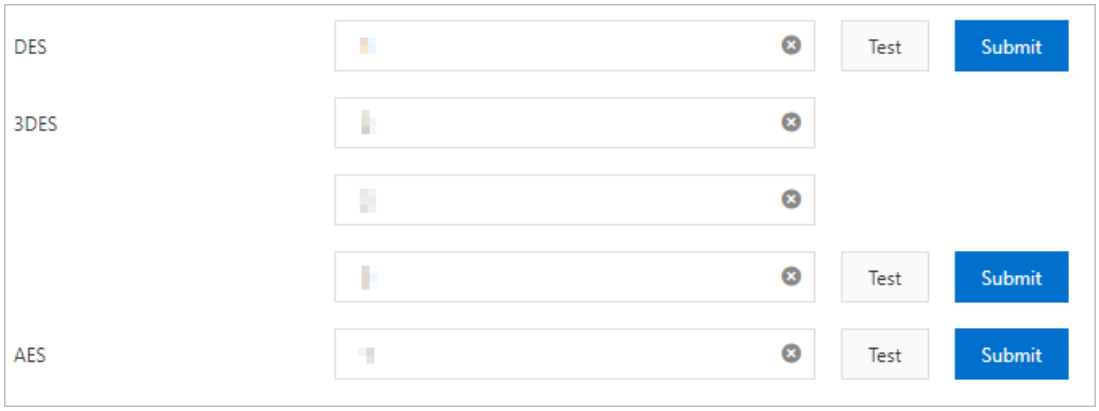
- **Replacement**: Set parameters for the substitution algorithm.



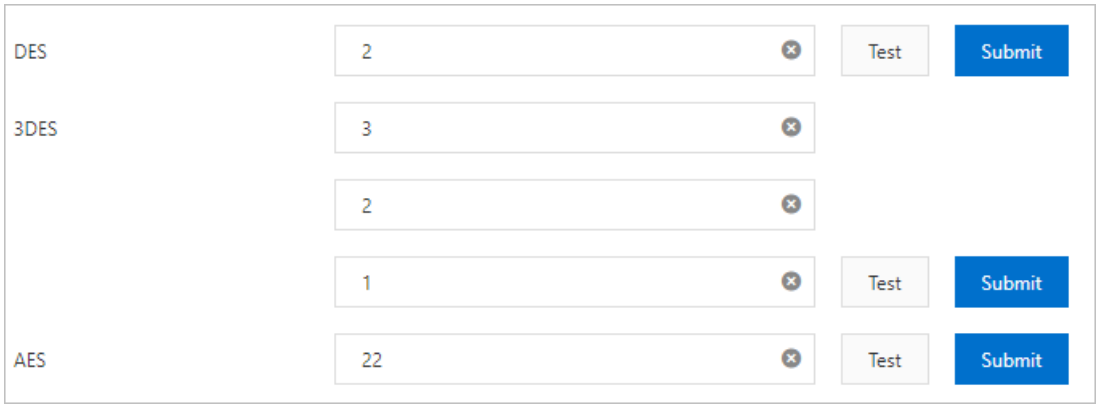
- **Transformation**: Set parameters for the rounding algorithm.



- **Encryption**: Set a key for each encryption algorithm.



- **Data decryption**: Set a key for each decryption algorithm.

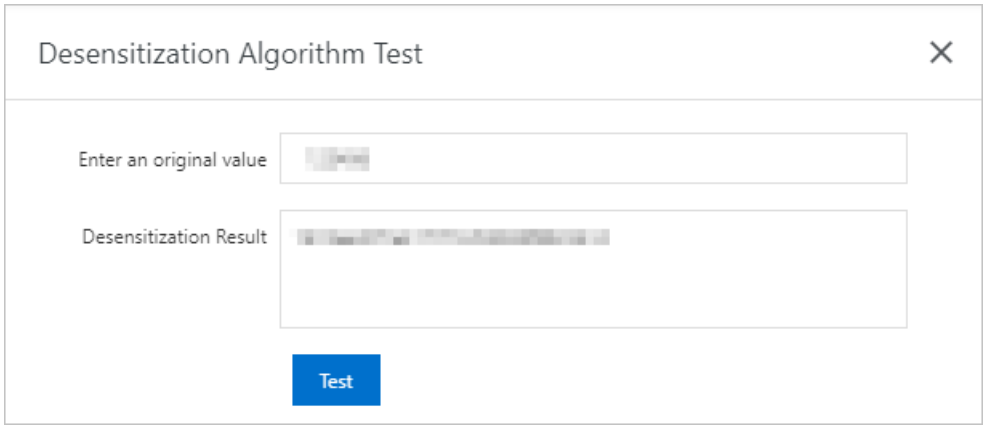


- **Shuffling**: Select a shuffling method.

> **Note** You do not need to test the shuffling method. Click **Submit** directly after you select a shuffling method.

5. Click **Test** for a parameter.

   In the **Desensitization Algorithm Test** panel, check whether the specific parameter for the de-identification algorithm works.



   After the test is completed, close the **Desensitization Algorithm Test** panel.

6. Click **Submit** for the parameter that you have tested.

What's next

After you configure the de-identification algorithm, go to the **Static Desensitization** page to create a de-identification task with the de-identification algorithm or modify the de-identification algorithm of an existing de-identification task. For more information, see Perform static de-identification.

6.5. Extract watermark information from data sources

Data Security Center (DSC) allows you to embed watermark information into the data to be distributed. If a data leak occurs, you can extract watermark information from the leaked data at the earliest opportunity. By reading the extracted watermark information, you can trace the data flow process and identify the organization or user that is responsible for the data leak. The watermark information that is embedded into the data to be distributed does not affect the use of the data. This topic describes how to extract watermark information from a data source.

Context

The watermark information that is embedded into the data in a data source provides the following features:

- **Security**: The watermark information that is embedded into the data is not lost even if the data is modified. This ensures that the watermark information can be accurately identified.
- **Transparency**: The watermark information that is embedded into the data is imperceptible to you and does not affect the use of the data.
- **Detectability**: You can extract the watermark information from data fragments and trace the source of data leaks with a high success rate.
- **Robustness**: You can extract the watermark information from the data even if the data is subject to malicious attacks.
- **Low error rate**: DSC provides a well-designed rule for extracting watermark information. This minimizes the probability of errors in data tracing.

Limits

You can extract watermark information only from data in ApsaraDB RDS databases.

Procedure

1. Log on to the DSC console.
2. In the left-side navigation pane, choose **Data desensitization > Extract watermark**.
3. On the **Extract watermark** page, set the parameters as required.

Parameter	Description
SOURCE product	The Alibaba Cloud service that contains the sensitive data to be de-identified. Only **RDS** is supported.
SOURCE Database/project name	Required. The project for storing the table that contains the watermark information to be extracted.

Parameter	Description
SOURCE table name	Required. The table that contains the watermark information to be extracted.

4. Click **Extract watermark**.

   In the lower part of the **Extract watermark** page, you can view the extracted watermark information. You can configure the watermark information when you add a static de-identification task.

   To copy the extracted watermark information, click **Copy results**.