Data Security Center

# Best Practices

| | |
|---|---|
| **Issue** | 01 |
| **Date** | 2021-07-06 |



HUAWEI TECHNOLOGIES CO., LTD.

**Copyright © Huawei Technologies Co., Ltd. 2021. All rights reserved.**

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Technologies Co., Ltd.

### Trademarks and Permissions


and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.
All other trademarks and trade names mentioned in this document are the property of their respective holders.

### Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

## Huawei Technologies Co., Ltd.

Address: Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129
People's Republic of China

Website: https://www.huawei.com

Email: support@huawei.com

# Contents

1 Quick Sensitive Data Identification and Masking..... 1

 Copyright © Huawei Technologies Co., Ltd. 

# 1 Quick Sensitive Data Identification and Masking

HUAWEI CLOUD Data Security Center (DSC) provides you with dynamic and static data masking methods to help mask your sensitive data in different scenarios, keeping your data secure and reliable. DSC provides you with over 20 preset data masking rules such as Hash, encryption, character masking, keyword replacement, value change, and roundup. You can use the default masking rules or customize the masking rules to mask sensitive data in the specified database table.

**Sensitive data** refers to the data that may bring serious harm to the society or individuals after being leaked.

NOTE

For individuals, privacy information, such as ID card numbers, home addresses, workplace information, and bank card numbers, is sensitive data. For enterprises or organizations, core information, such as customer information, financial information, technical information, and major decisions, is sensitive data.

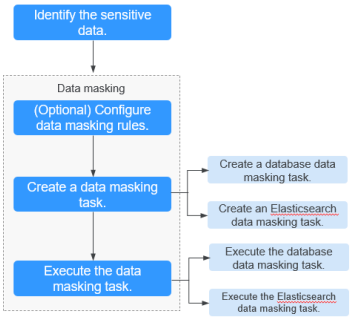
**Static data masking**: DSC can help mask a large amount of data at one time based on the configured data masking rules. Static data masking is used when sensitive data in the production environment is delivered to the development, testing, or external environment for development and testing and data sharing and research. You can create an data masking task on the DSC console to quickly mask sensitive data in databases and big data assets.

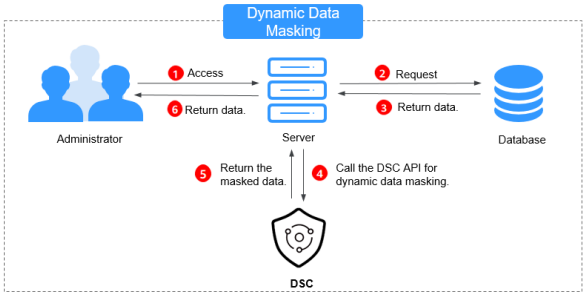
**Dynamic data masking**: DSC provides dynamic data masking APIs to mask the data accessed from the external systems. Dynamic data masking applies to scenarios where data is queried from the external system, such as production applications, data exchange, O&M applications, and precision marketing. For details, see **Dynamic Data Masking**.

## Data Masking Process

With DSC, you can implement dynamic and static data masking to fully protect your sensitive data.

 Copyright © Huawei Technologies Co., Ltd. 




## Data Masking Algorithms

DSC has over 100 sensitive data identification and masking rules and can identify and mask the sensitive information such as personal information (ID card information, bank card information, names, mobile numbers, email addresses, and more), enterprise information (business license numbers, tax registration certificate numbers, and more), and key information (PEM certificates, HEY private keys, and more), device information (IP addresses, MAC addresses, IPv6 addresses, and more), location information (provinces/states, cities, GPS locations, addresses, and more), and common information (dates and others). **Table 1-1** describes the data masking algorithms and application scenarios.

 Copyright © Huawei Technologies Co., Ltd. 

**Table 1-1** Masking algorithms

| Algorithm | Description | Application Scenario |
| --- | --- | --- |
| Hash | Use Hash functions to mask sensitive data. DSC supports SHA-256 and SHA-512.<br>• SHA-256<br>SHA-256, a message-digest algorithm, is used by DSC to compute a digest from a string in the database table.<br>It takes a block of data and returns a fixed-size bit string (hash value). As the value length may exceed the maximum column width allowed in the original table, you can adjust the column width to adapt to the returned SHA-256 hash values.<br>• SHA-512<br>SHA-512, a message-digest algorithm, is used by DSC to compute a digest from a string in the database table.<br>It takes a block of data and returns a fixed-size bit string (hash value). As the value length may exceed the | • Sensitive data: Key information<br>• Application scenario: data storage |

 Copyright © Huawei Technologies Co., Ltd. 

| Algorithm | Description | Application Scenario |
| --- | --- | --- |
| | maximum column width allowed in the original table, you can adjust the column width to adapt to the returned SHA-512 hash values. | |
| Encryption | Use the encryption algorithms and master key to implement data masking. In the encryption and data masking result, the first 16 bytes of an encrypted string is the initialization vector (IV) and the rest is the enciphered text.<br>DSC supports three encryption algorithms: AES-128, AES-192, and AES-256. | • Sensitive data:<br>– Personal data<br>– Enterprise data<br>• Application scenario: data storage |

 Copyright © Huawei Technologies Co., Ltd. 

| Algorithm | Description | Application Scenario |
| --- | --- | --- |
| Character Masking | Use the specified character * or random characters (including numbers, letters, and both number and letters) to cover part of the original content. The following six data masking approaches are supported:<br>● Retain first *N* and last *M*<br>● Retain from *X* to *Y*<br>● Mask first *N* and last *M*<br>● Mask from *X* to *Y*<br>● Mask data ahead of special characters<br>● Mask data followed by special characters<br>**NOTE**<br>DSC has multiple character masking templates. | ● Sensitive data: Personal data<br>● Application scenarios:<br>– Data usage<br>– Data sharing |

 Copyright © Huawei Technologies Co., Ltd. 

| Algorithm | Description | Application Scenario |
| --- | --- | --- |
| Keyword Replacement | Search for keywords in a specified column and replace them.<br>For example, the specified characters are "Zhang San eats at home". After replacement, the characters become "Mr. Zhang eats at home". In the example, "Zhang San" is replaced with "Mr. Zhang".<br>After this algorithm is executed, the value length may exceed the maximum length allowed by the database. In this case, the excess part will be truncated and inserted into the database. | • Sensitive data:<br>– Personal data<br>– Enterprise data<br>– Device data<br>• Application scenarios:<br>– Data storage<br>– Data sharing |

 Copyright © Huawei Technologies Co., Ltd. 

| Algorithm | Description | Application Scenario |
| --- | --- | --- |
| Value Change | Set a specified field to **Null** or left it blank for data masking.<br>● Masking Using the Null Value<br>Set a field of any type to **NULL**.<br>If a field is set to **NOT NULL**, this algorithm changes the attribute of the file to **NULL** when copying the column.<br>● Masking Using a Custom Value<br>Set the target field to a default value.<br>Specifically, a character field is left blank, a numeric field is set to **0**, a date field is set to **1970**, and time field is set to **00:00**. | ● Sensitive data:<br>– Personal data<br>– Enterprise data<br>– Device data<br>● Applicable scenarios<br>– Data storage<br>– Data sharing |

 Copyright © Huawei Technologies Co., Ltd. 

| Algorithm | Description | Application Scenario |
| --- | --- | --- |
| Roundup | Round a date or number.<br>• Date Roundup<br>Roundup of fields after the year field<br>Example: 2019-05-12 -> 2019-01-01 or 2019-05-12 08:08:08 -> 2019-01-01 00:00:00<br>Roundup of fields after the month field<br>Example: 2019-05-12 -> 2019-05-01 or 2019-05-12 08:08:08 -> 2019-05-01 00:00:00<br>Roundup of fields after the day field<br>Example: 2019-05-12 -> 2019-05-12 or 2019-05-12 08:08:08 -> 2019-05-12 00:00:00<br>Roundup of fields after the hour field<br>Example: 08:08:08 -> 08:00:00 or 2019-05-12 08:08:08 -> 2019-05-12 08:00:00<br>Roundup of fields after the minute field | • Sensitive data: General data<br>• Applicable scenarios<br>– Data storage<br>– Data usage |

 Copyright © Huawei Technologies Co., Ltd. 

| Algorithm | Description | Application Scenario |
| --- | --- | --- |
| | Example: 08:08:08 -> 08:08:00 or 2019-05-12 08:08:08 -> 2019-05-12 08:08:00<br>Roundup of fields after the second field<br>Example: 08:08:08.123 -> 08:08:08.000 or 1575612731312 -> 1575612731000<br>• Number roundup Rounds a specified number. | |

**Example:**

Assume that the **dsc_yunxiaoke** table in the **rsd-dsc-test** database stores the information of the following bank employees:

| Name | Birthday | Email | Address |
| --- | --- | --- | --- |
| San Zhang | 1999/6/3 | XXXXXX@163.com | Chengdu, Sichuan |
| Si Li | 1996/3/6 | 55XXXX@qq.com | Beijing |

To identify and mask sensitive data in the table, you can select the preset identification rule template for banking and finance data to identify sensitive data and generate the identification result, and then mask the identified sensitive data using the SHA256 algorithm in **Hash**.

## Step 1 Identifying Sensitive Data

**Step 1** **Buy DSC**.

**Step 2** **Log in to the management console.**

**Step 3** Click ☰ and choose **Security & Compliance** > **Data Security Center**.

**Step 4** In the left navigation pane, choose **Sensitive Data Identification** > **Identification Task**.

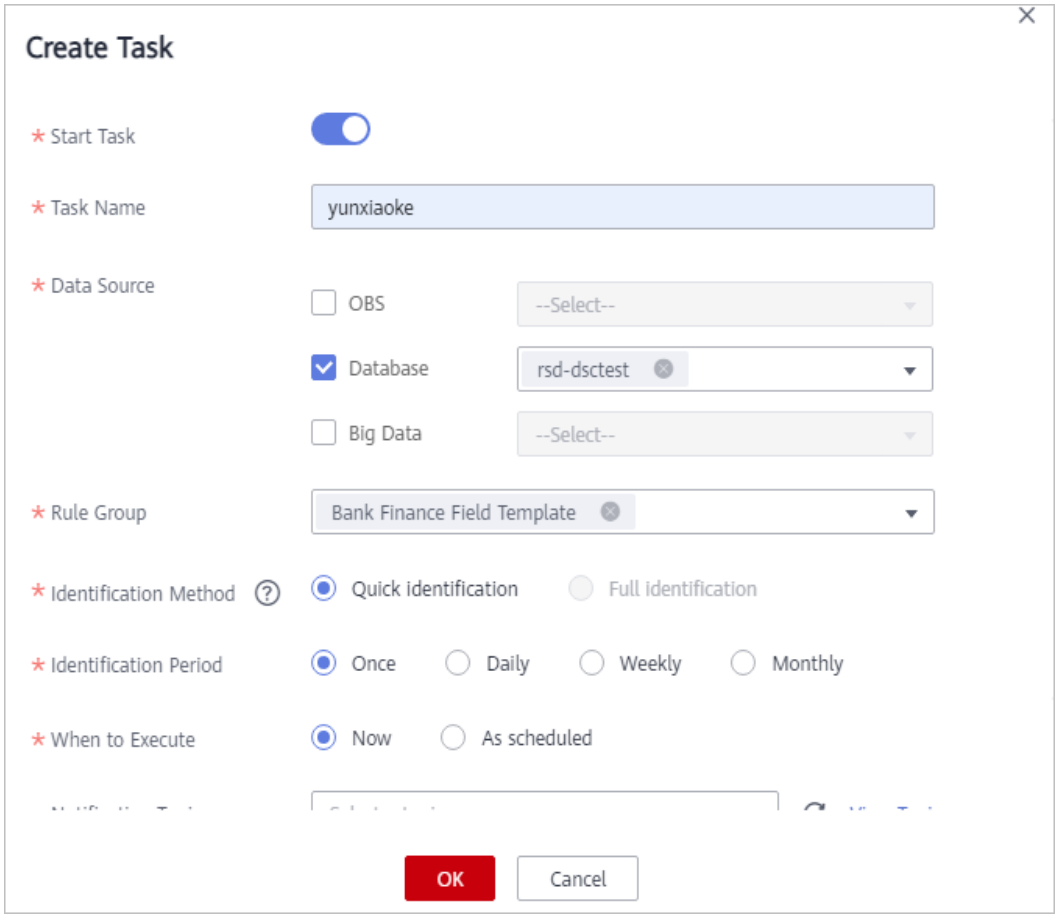
 Copyright © Huawei Technologies Co., Ltd. 

**Figure 1-1** Identification task



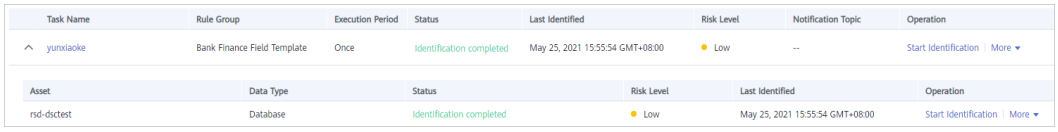

**Step 5** Click **Create Task**. In the displayed dialog box, configure the basic parameters.

**Figure 1-2** Creating a sensitive data identification task




**Step 6** Click **OK**. The sensitive data identification task list is displayed.

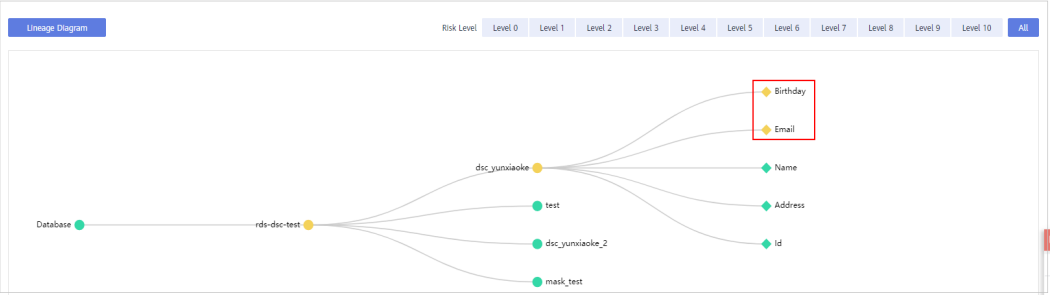
**Figure 1-3** Sensitive data identification task list




**Step 7** When the status of the identification task changes to **Identification completed**. Click **View Result** in the **Operation** column to go to the result details page.

 Copyright © Huawei Technologies Co., Ltd. 

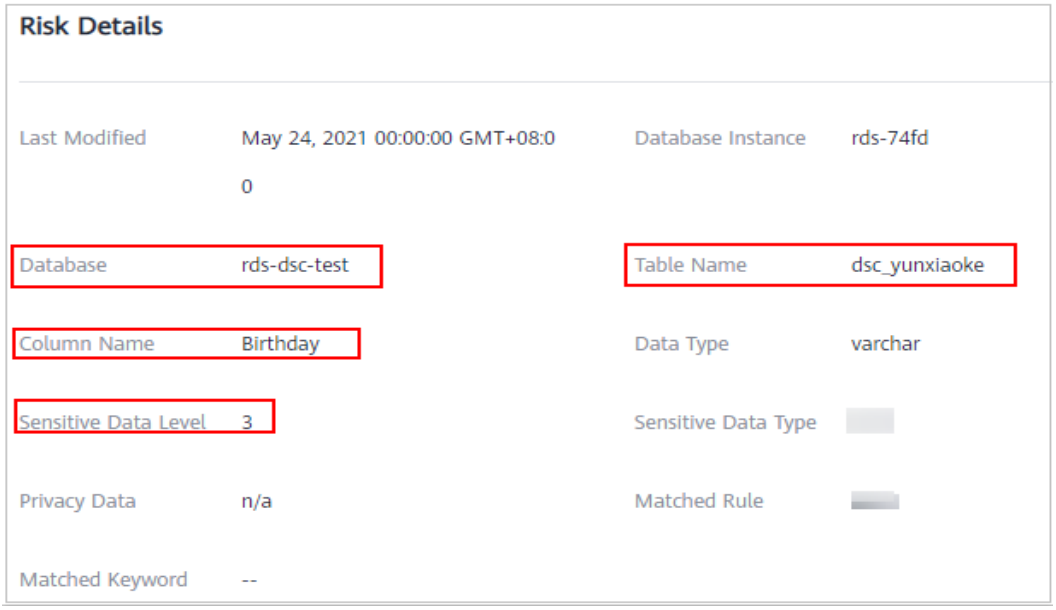
**Figure 1-4** Lineage diagram




In the preceding figure, the birthday dates and email addresses are identified as sensitive data.

**Step 8** Click a sensitive field to view the risk details.

**Figure 1-5** Risk details




Perform operations described in **Step 2 Masking Sensitive Data** to mask the sensitive data in the **Birthday** and **Email** columns of the **dsc_yunxiaoke** table in the **rds-dsc-test** database.

**----End**

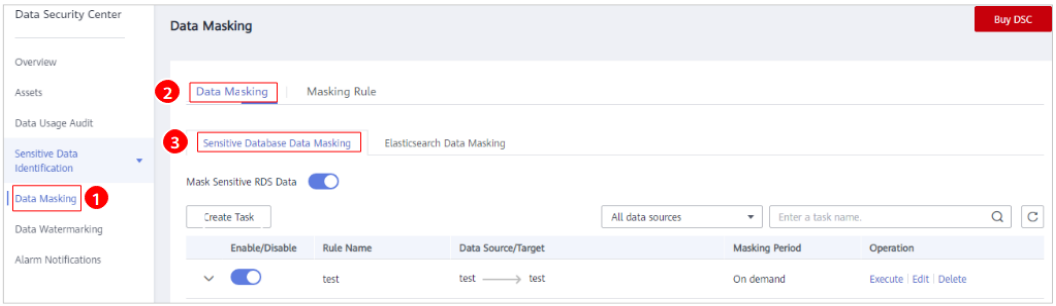
## Step 2 Masking Sensitive Data

You can create data masking tasks for the database and Elasticsearch (the data masking methods are the same). The following describes how to create a database data masking task. For details about how to create a data masking task for Elasticsearch, see **Creating a Data Masking Task for Elasticsearch**.

**Step 1** In the left navigation pane, choose **Data Masking**. The **Data Masking** > **Sensitive Database Data Masking** page is displayed by default.

 Copyright © Huawei Technologies Co., Ltd. 

**Figure 1-6** Accessing the Database Data Masking tab page

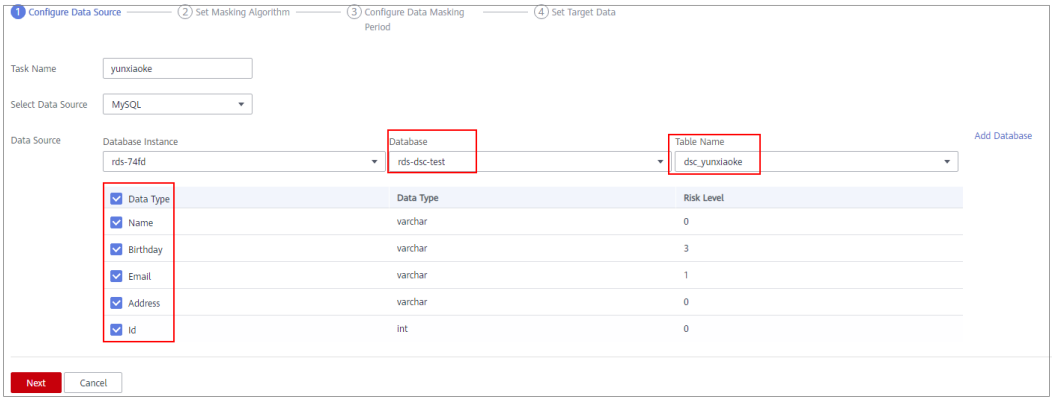



**Step 2** Set **Mask Sensitive RDS Data** to .

**Step 3** Click **Create Task** to configure the data source.

Select all data types if you want a complete table that contains all types of data after the data masking is completed.

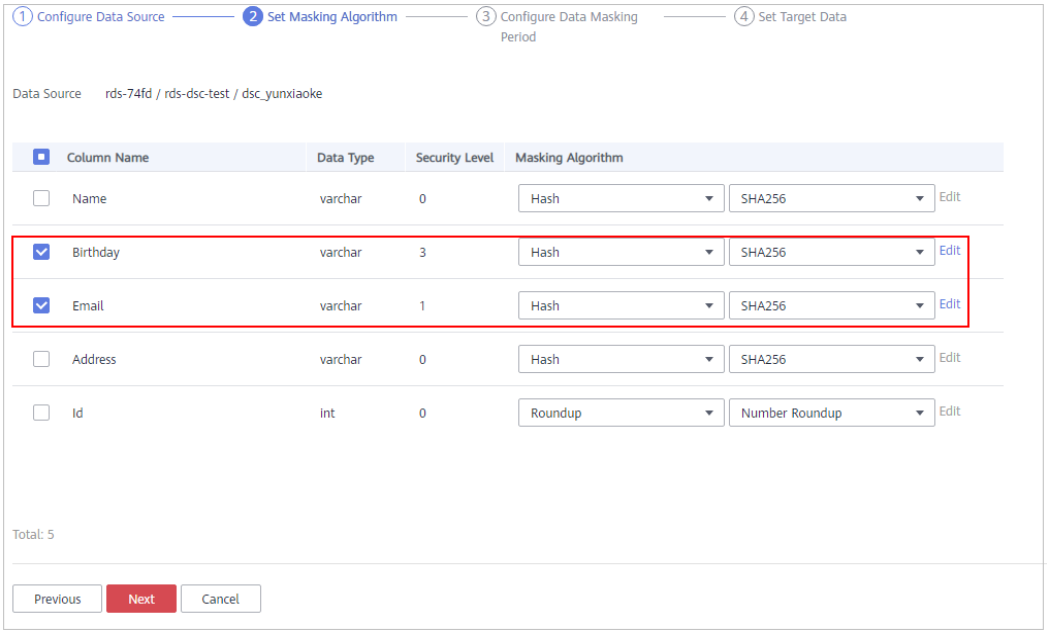
**Figure 1-7** Data source configuration




**Step 4** Click **Next** to switch to **Set Masking Algorithm**.

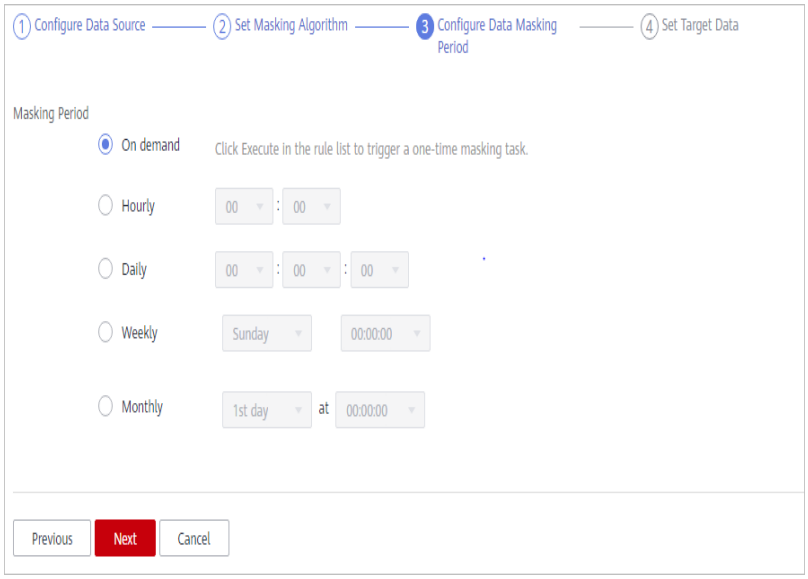
 Copyright © Huawei Technologies Co., Ltd. 

**Figure 1-8** Configuring the data masking algorithm




**Step 5** Click **Next** to switch to the **Configure Data Masking Period** page and configure the data masking period.

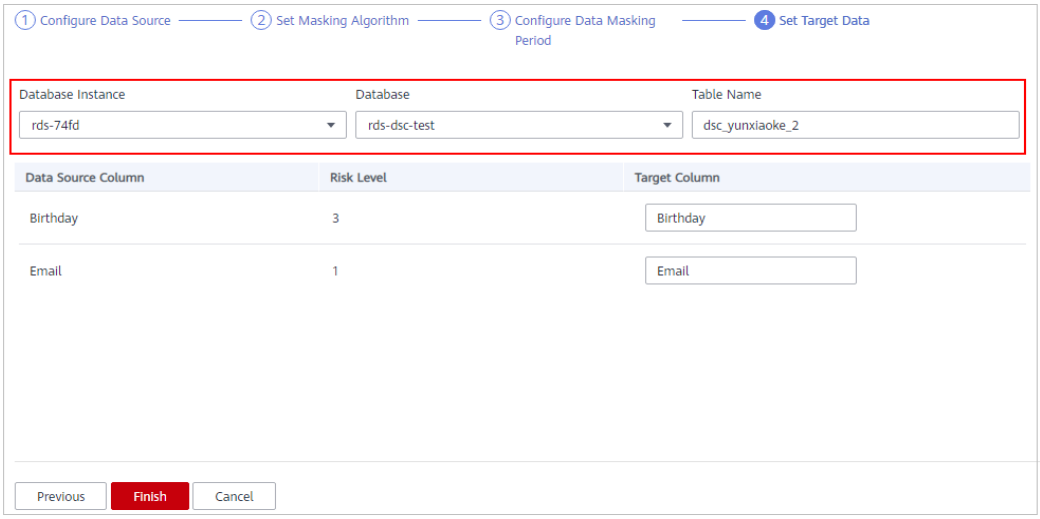
**Figure 1-9** Configuring the data masking period




**Step 6** Click **Next** to the **Set Target Data** page and configure the storage location of the table generated after data masking.

 Copyright © Huawei Technologies Co., Ltd. 

**Figure 1-10** Configuring the storage location of the table generated after data masking

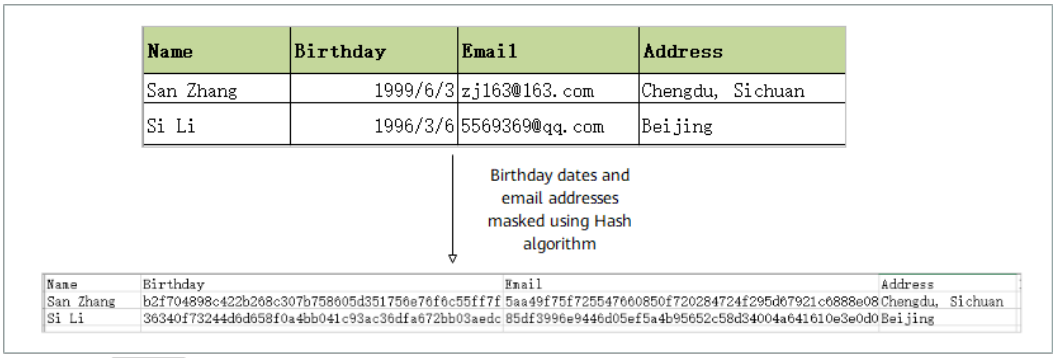



**Step 7** Click **Finish** to return to the database data masking task list. Click  to enable the masking task and then **Execute** in the **Operation** column to execute the task.

If the status changes to **Completed**, the data masking task has been successfully executed.

**----End**

## Verifying the Result




 Copyright © Huawei Technologies Co., Ltd. 