Graph Engine Service

Getting Started

Issue 01

Date 2022-09-30

HUAWEI

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 Service Introduction

Graph Engine Service (GES) is an in-house distributed native graph engine for commercial use. It is the first of its kind in China that owns independent intellectual property rights. It facilitates query and analysis of multi-relational graph data structures. It is particularly well suited for scenarios requiring analysis of rich relationships, including social network analysis, enterprise relationship analysis, risk control, marketing recommendations, social listening, and fraud detection.

2 Creating a Graph

The following content describes how to create a graph on GES console.

Two creation modes are available: **Customize Graph** and **Use Industry-Specific Graph Template**. By default, the system displays the **Customize Graph** tab.

Creating a Custom Graph

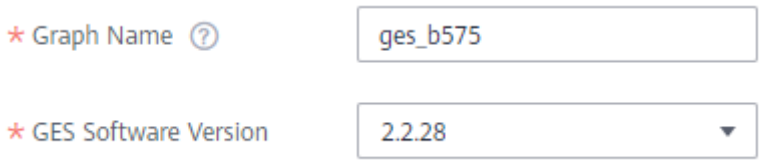
Step 1 Log in to the GES console and click **Create Graph** in the upper right corner of the home page. The **Create Graph** page is displayed.

Step 2 Select the **Region** where GES works from the drop-down list in the upper left corner of the page.

Step 3 On the **Create Graph** page, click the **Customize Graph** tab and set the following parameters:

1. In the **Configure** tab, set the graph name and software version.

Figure 2-1 Graph name and software version

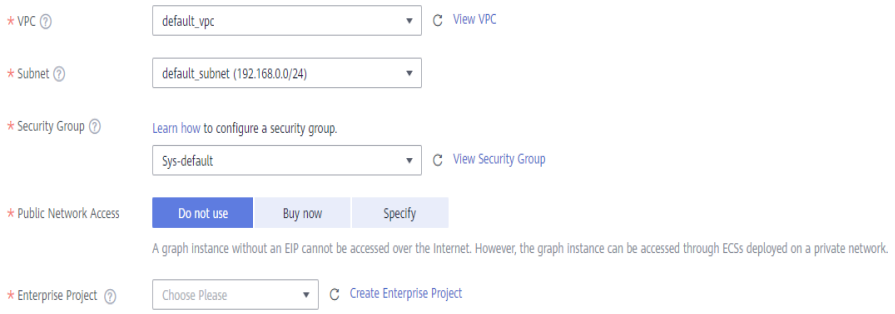


Parameter	Description
Graph Name	You can set a name or use the default name. After a graph is created, its name cannot be changed. The graph name must: – Contain 4 to 50 characters and start with a letter. – Be case-insensitive. – Contain only letters, digits, and underscores (_).

Parameter	Description
GES Software Version	The system uses the latest version by default. Currently, only the default version is available.

2. Specify the network information, including **VPC**, **Subnet**, **Security Group**, **Public Network Access**, and **Enterprise Project**.

Figure 2-2 Network Information

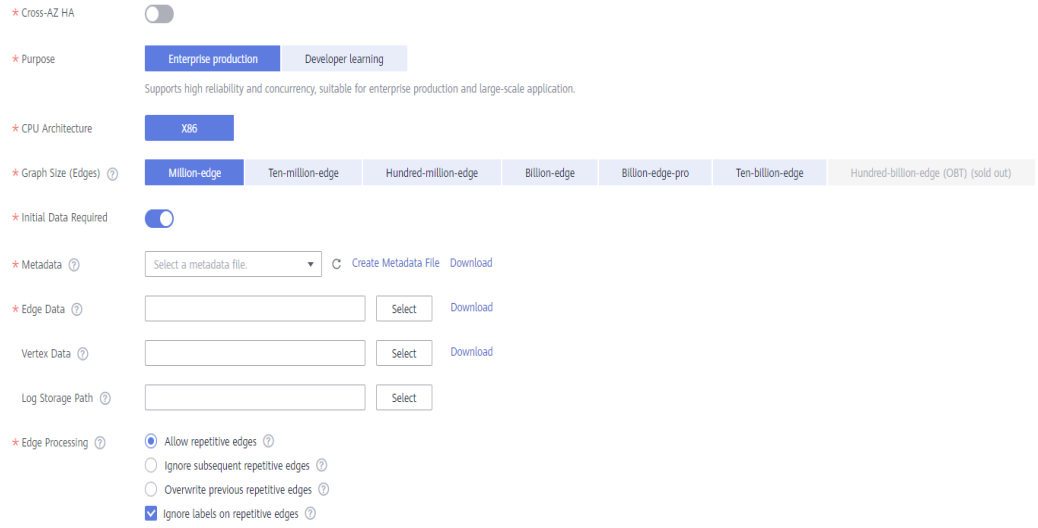


Parameter	Description
VPC	A VPC is a secure, isolated, and logical network environment. Select the VPC for your graph. You can click **View VPC** to view the name and ID. **NOTE** If your account has VPCs, a VPC will be automatically selected. You can change it as needed. If no VPC is available, you need to create a VPC. After the VPC is created, it will be automatically selected.
Subnet	A subnet provides dedicated network resources that are logically isolated from other networks for network security. Select the subnet for your graph. You can access the VPC service to view the name and ID.
Security Group	A security group implements access control for ECSs that have the same security protection requirements in a VPC. – Click **Learn how to configure a security group.** to get instructions. – Click **View Security Group** to learn security group details.

Parameter	Description
Public Network Access	Public network access to the graph. Set this parameter as you need. **Do not use**: A graph instance without an elastic IP (EIP) cannot be accessed over the Internet. However, the graph instance can be accessed through ECSs deployed on a private network. **Buy now**: GES automatically allocates an EIP with exclusive bandwidth to the graph instance so that the graph instance can be accessed over the Internet using the EIP. In addition, GES uses the tenant permission to automatically create an agency with the prefix of **ges_agency_default** in the project to support EIP binding. **Specify**: Select an EIP to allow the graph instance to be accessed over the Internet. Click **Create EIP** to access the VPC management console and create an EIP.
Enterprise Project	Centrally manages cloud resources and members by project. Click **Create Enterprise Project** to go to the **Enterprise Project Management** page.

3. Set graph parameters.

Figure 2-3 Graph parameters



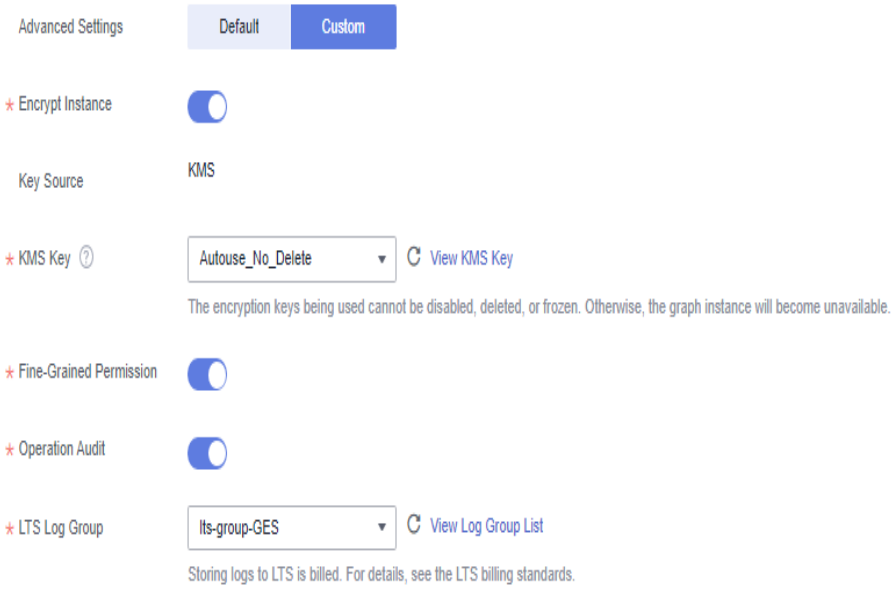
Parameter	Description
Cross-AZ HA	Whether to support cross-AZ cluster. If this function is enabled, graph instances are distributed in different AZs for better reliability.
Purpose	Purpose of the graph to be created. **Enterprise production**: High reliability and concurrency are supported, suitable for production and large-scale applications. **Developer learning**: A complete function experience is offered, suitable for developer learning.
CPU Architecture	**CPU Architecture**: Select **X86** or **Kunpeng**.
Graph Size (Edges)	The system only displays the available sizes based on your resource quota. **Enterprise production** and **Developer learning** support different graph sizes. – **Enterprise production**: For the production purpose, the following graph sizes are available: **Million-edge**, **Ten-million-edge**, **Hundred-million-edge**, **Billion-edge**, **Billion-edge-pro**, **Ten-billion-edge**, **Hundred-billion-edge** (OBT). – **Developer learning**: Only **Ten-thousand-edge** is available.

Parameter	Description
Initial Data Required	This parameter is disabled by default. You can create a graph first and then import data. If you enable this option, set the following parameters: – **Metadata**: Metadata file If no metadata is available, click **Create Metadata File**. For details about how to create metadata, see . If you have metadata files, import them to GES. – **Edge Data**: Edges that form the graph, including information about the edge structures, labels, and properties – **Vertex Data**: Vertices that form the graph, including information about all vertex IDs, labels, and properties. If you leave it blank, the vertices in the **Edge Data** set are used as the source of **Vertex Data**. **NOTE** – The edge and vertex data sets can only be stored in English paths and folders. – Currently, you can import the edge and vertex data sets only from OBS. Therefore, store data files on OBS in advance.. – The sequence of labels in the selected edge or vertex data set and the sequence of properties in the labels must be the same as those in the selected metadata file. Otherwise, **The edge/vertex data file does not match the metadata file** is prompted in the upper right corner and the graph fails to be created. For specific content of the format of the GES graph data, refer to – Import the graph data (including the metadata file, and edge and vertex data sets) in the format specified in the corresponding template. The template contains a copy of movie data. You can click **Download** to download and import it. – **Log Storage Path**: Stores vertex and edge data sets that do not comply with the metadata definition, as well as detailed logs during graph import. Storage on OBS may incur fees, so delete the data in a timely time if you do not need to use it anymore. – **Edge Processing**: Repetitive edges have the same source and target vertices. When labels are considered, repetitive edges must have the same source and target vertices and the same labels. ▪ **Allow repetitive edges**: Multiple edges may exist between a source vertex and a target vertex. ▪ **Ignore subsequent repetitive edges**: If there are multiple edges between a source vertex and a target vertex, only the first edge read is retained.

Parameter	Description
	▪ **Overwrite previous repetitive edges**: If there are multiple edges between a source vertex and a target vertex, only the last edge read is retained. ▪ **Ignore labels on repetitive edges**: If labels are ignored, edges with the same source vertex and target vertex are repetitive edges.

4. **Advanced Settings**: Set this parameter to **Default** or **Custom**.

Figure 2-4 Advanced settings



- **Default**: Use the default settings.
- **Custom**: Set whether to enable **Encrypt Instance**, **Operation Audit**, or **Fine-Grained Permission**.

Parameter	Description
Encrypted Instance	Whether to encrypt a graph instance. **Key Source** is default to **KMS**. **KMS Key**: Select the corresponding key. **NOTE** Disabling or deleting a KMS key affects the instance functions.
Fine-Grained Permission	Whether to enable fine-grained permission management. If this function is enabled, the traverse, read, and write permissions can be set for specific attributes each label.

Parameter	Description
Operation Audit	Whether to enable operation audit **LTS Log Group**: Select the corresponding log group. Click **View Log Group List** to view log information on the log management page. **NOTE** You will be billed for storing logs to LTS. For details, see the LTS billing standards.

Step 4 Click **Next**. The confirmation page is displayed.

Step 5 Confirm the information and click **Submit** to create the graph.

Step 6 After the submission is successful, the **Finish** page is displayed. You can click **Back to Task Center** to view the status and running result of the created graph.

----End

Using an Industry-Specific Graph Template to Create an Instance

Step 1 Log in to the GES console and click **Create Graph** in the upper right corner of the home page. The **Create Graph** page is displayed.

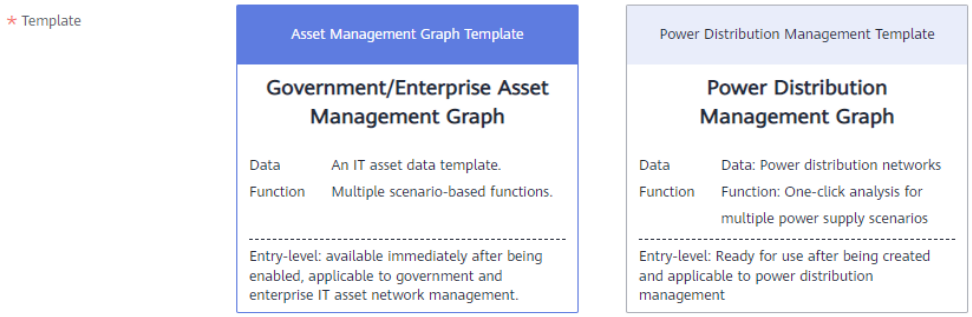
Step 2 Select the **Region** where GES works from the drop-down list in the upper left corner of the page.

Step 3 On the **Create Graph** page, click the **Use Industry-Specific Graph Template** tab and configure the following parameters:

In the **Configure** tab, select a template and set network and graph information:

1. Select the desired template to create your graph. Currently, **Asset Management Graph Template** and **Power Distribution Management Template** are available.

   Figure 2-5 Selecting a template



2. Set network information. Set the parameters by referring to the **Creating a Custom Graph** section.

Step 4 Click **Next**. In the **Confirm** tab, confirm the specifications and click **Submit**. The system automatically creates the graph of the selected specifications and inserts the selected template data (schema and sample data).

Step 5 After the submission is successful, the **Finish** tab page is displayed. You can click **Back to Task Center** to view the status of the created graph.

NOTE

- You do not need to set the name for a graph created based on a template. By default, the name of the template is used as the prefix of the created graph, for example, **assets_management**.
- After the graph is created, the name of the created graph is in **assets_management**_*XXXX* format, where *XXXX* is the unique identifier automatically generated by the system and cannot be modified.

----End

Accessing a Graph

1. You can view all created graphs on the **Graph Management** page and click **Access** in the **Operation** column of a graph.
2. On the displayed graph editor, you can query and analyze the graph. For details, see **Querying and Analyzing a Graph**.
3. For details about the GES management console and graph editor, see the .

3 Querying and Analyzing a Graph

This following content describes how to query a graph.

Accessing a Customized Graph

You can query and analyze graphs in any of the following ways if a graph is created without using any template:

- **Gremlin commands**
  a. Enter the query command in the Gremlin text box in the lower part of the page, for example, **g.V().limit(100)**.

     NOTE

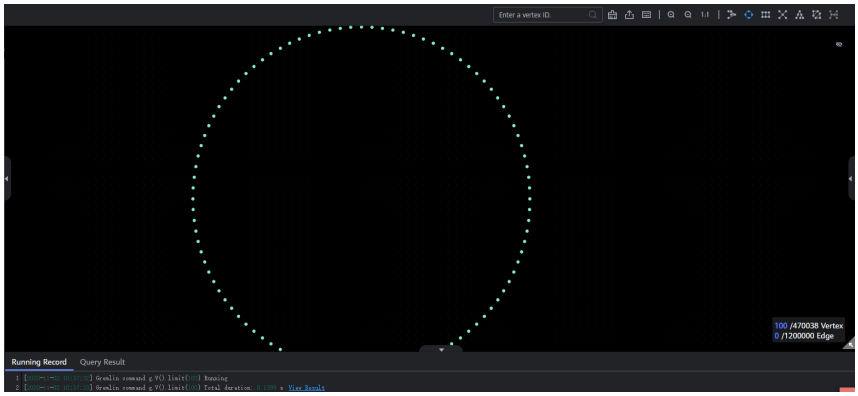
     It may take a long time for the system to return a large amount of data. You are advised to add the **limit** parameter and set it to less than **1,000**.

     Figure 3-1 Gremlin query

  b. Press **Enter** to run the Gremlin command. The query result is displayed in both the canvas and result pane.

     Figure 3-2 Gremlin query result

- **Cypher commands**

a. Enter the query command in the Cypher text box in the lower part of the page, for example, **match (n) return n limit 100**.

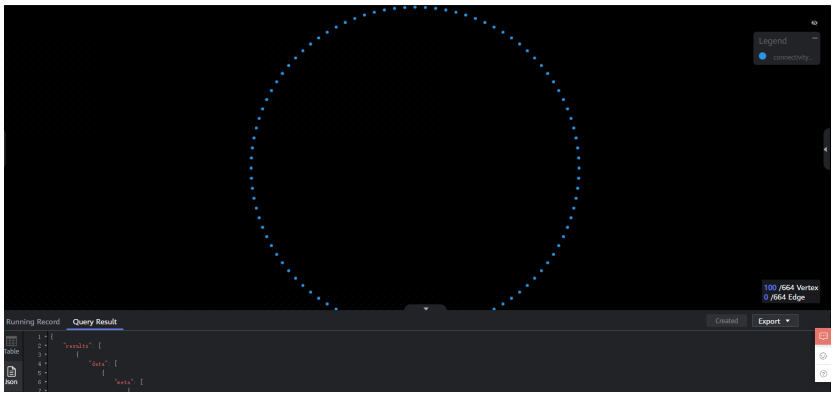
Figure 3-3 Cypher query

NOTE

- If this is your first time using Cypher, click **Create Index** in the result pane. You do not need to perform this operation in subsequent operations.
- **match (n) return n limit 100** queries details about 100 vertices.

b. Press **Enter** to run the Cypher command. The query results are displayed in both the canvas and result pane.

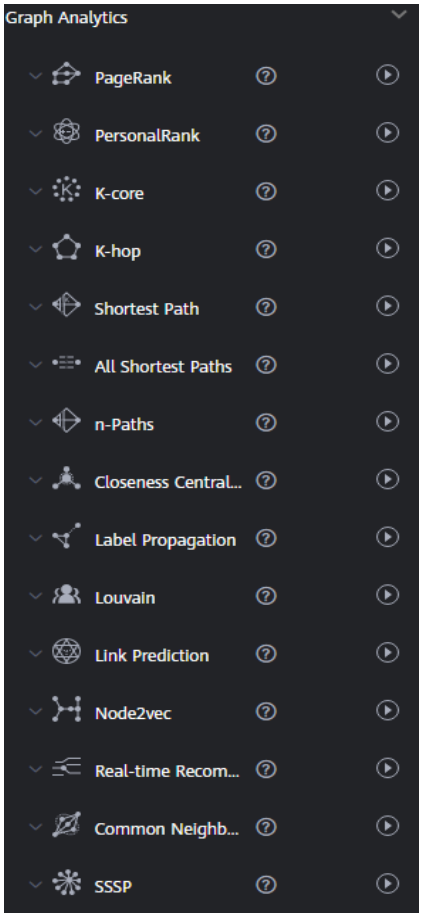
Figure 3-4 Query result



- **Algorithm analysis**

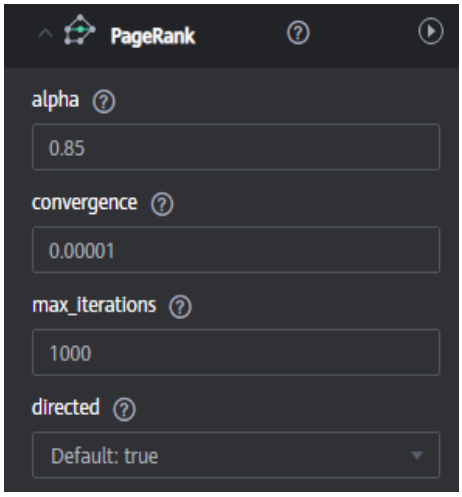
a. Available algorithms are displayed in the left pane of the graph editor. Select the analysis algorithm you want to use.

Figure 3-5 Algorithm list



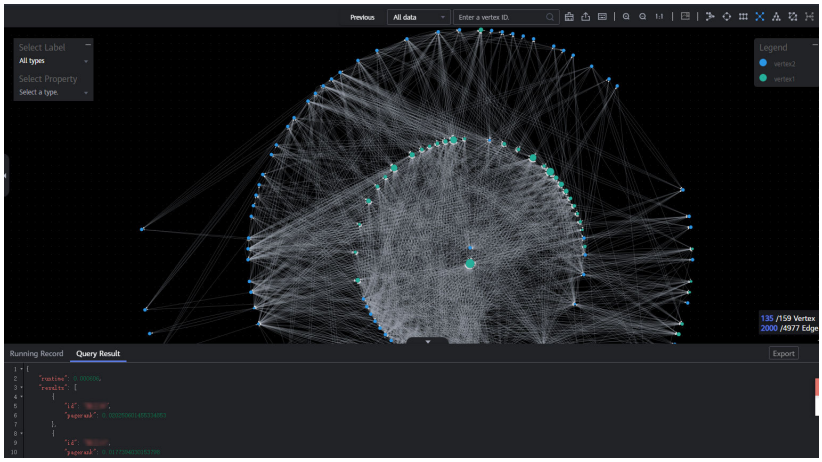
b. Expand the algorithm by clicking and configure the parameters as required. Take PageRank as an example. **alpha** indicates the weight coefficient, and its value is **0.85**. **convergence** is the convergence coefficient, and its value is **0.00001**. **max_iterations** indicates the maximum iterations and its value is **1,000**. **directed** indicates whether to consider the edge direction and the default value is **true**.

Figure 3-6 Setting PageRank algorithm parameters



c. Click  to execute the algorithm. The analysis result is displayed in both the canvas and result pane.

Figure 3-7 Analysis results



Accessing a Graph Created Using a Template

In addition to queries using Gremlin commands, Cypher commands, and algorithms, the following queries are also available for a graph created using an industry-specific template:

- **Asset Management Graph Template**

  Query 1: List administrators and their applications.

  - Command: **g.V().hasLabel('admin').outE().otherV().hasLabel('application').path()**
  - Procedure: Click the run button. The query result is displayed in the graph on the canvas.

  Query 2: Find all paths from administrator *x* to equipment room *x*.

  - Principle: **get${Administrator *x*}** is the source vertex and **${Equipment room *x*}** is the target vertex. Find out all paths using the n-Paths algorithm, where **source** is **Administrator** *x*, **target** is **Equipment room** *x*, **direct** is **false**, **n** is **10**, and **k** is **10**.
  - Procedure: Click the run button. In the dialog box that is displayed, select administrator *x* and equipment room *x* from the drop-down list box. After the operation is complete, the graph is displayed on the canvas.

  Query 3: Find all VMs and physical machines on which application *x* depends.

  - Principle: **get${Application *x*}** is the input to find all points at layer 2. Use the K-Hop algorithm, where **k** is **2**, **source** is *x*, and **mode** is **in**.
  - Procedure: Click the run button. In the dialog box that is displayed, set the application *x* (select the application from the drop-down list box and drag the vertex of the application entity to the drop-down list box). After the operation is complete, the graph is displayed on the canvas.

  Query 4: Find all vertices with a specified label, for example, all administrators, equipment rooms, and physical machines in the graph.

  - Principle: Search vertices by label.
  - Procedure: Click the run button. In the dialog box that is displayed, select **label** from the drop-down list box. The query result is displayed in the graph on the canvas.
- **Power Distribution Management Template**

  Query 1: Query the power supply range.
  - The sub-queries are as follows:
    i. Locate power substation busbars.
    ii. Query the power supply scope of a busbar.
    iii. Query user points within the power supply scope of a busbar.
  - Procedure: Click the run button. In the dialog box that is displayed, select the busbar value. The query result is displayed in the graph on the canvas.

  Query 2: Analyze the outage fault.
  - The sub-queries are as follows:
    i. Trace back the faulty point from the power-off user point.
    ii. Query faulty point details.
  - Procedure: Click the run button. In the dialog box that is displayed, select parameters of the faulty point or the power-off user point. The query result is displayed in the graph on the canvas.

  Query 3: Query the ring circuit power supply.
  - The sub-queries are as follows:
    i. Locate the ring circuit supply line.
    ii. Locate contact switches in the ring circuit supply line.
  - Procedure: Click the run button. The query result is displayed in the graph on the canvas.

Viewing the Analysis Result

You can view the running record and query result in the result pane or click the **Export** button on the right to download the analysis result.

Figure 3-8 Execution result