ALIBABA CLOUD

Alibaba Cloud

AnalyticDB for PostgreSQL
Data Warehouse Development

Document Version: 20220104

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.Use DataWorks to schedule ETL tasks of AnalyticDB for Postgr... 05

1.Use DataWorks to schedule ETL tasks of AnalyticDB for PostgreSQL

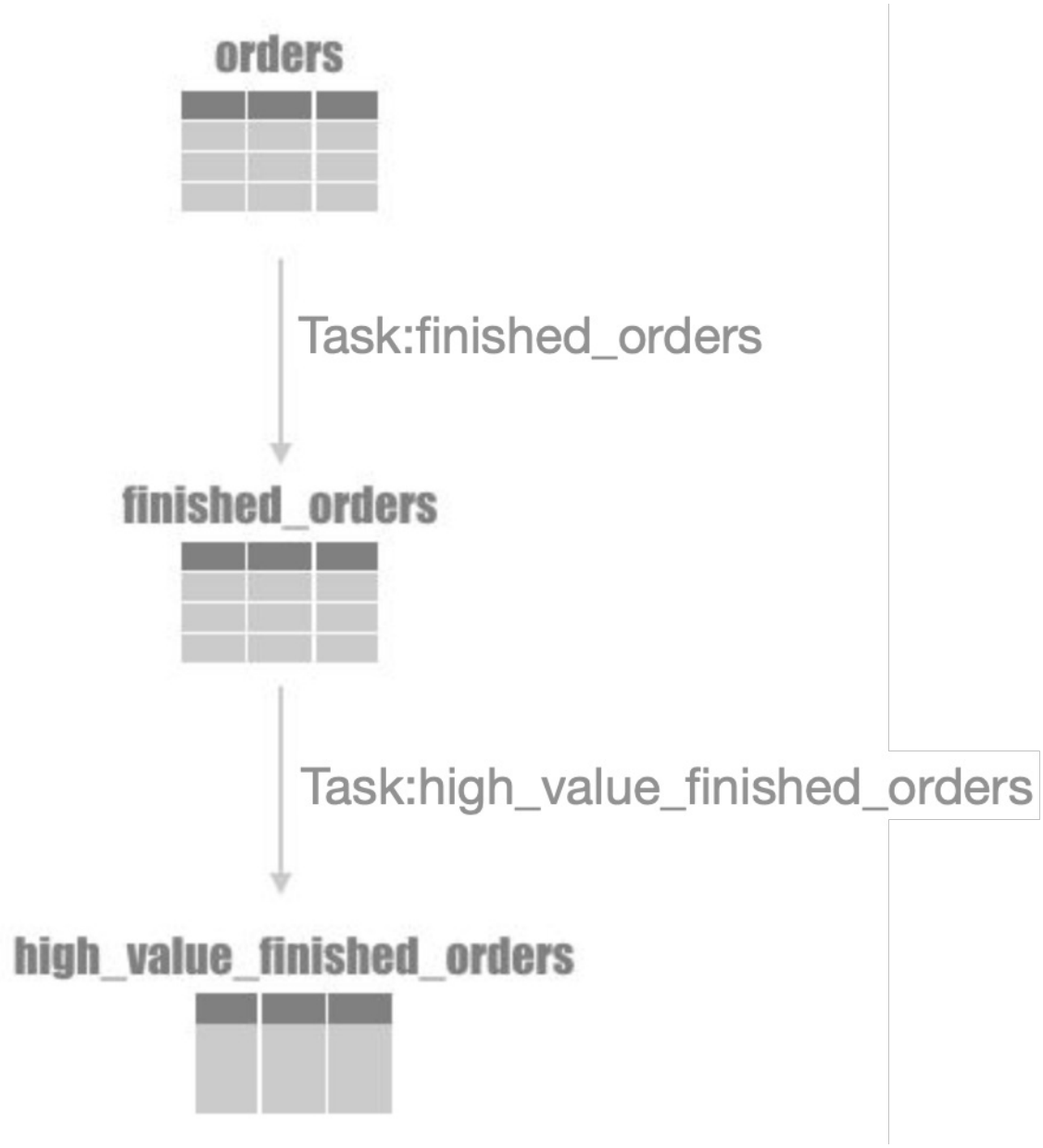
This topic describes how to use DataWorks to schedule the script tasks of AnalyticDB for PostgreSQL. DataWorks allows you to develop, schedule, and maintain the tasks of AnalyticDB for PostgreSQL and manage dependencies among tasks. It enhances the extract, transform, load (ETL) capabilities of AnalyticDB for PostgreSQL.

Prerequisites

- Test data is obtained from the TPCH test data set.
- Data is imported to AnalyticDB for PostgreSQL. For more information, see Introduction to data migration and synchronization solutions.

Overview

Dependencies among tasks are important for task scheduling. For example, the relationships between two AnalyticDB for PostgreSQL tasks created in DataWorks and AnalyticDB for PostgreSQL tables are shown in the following figure.



- Task: finished_orders

  AnalyticDB for PostgreSQL uses `o_orderstatus = 'F'` to filter finished orders in the orders table and then writes the orders into the finished_orders table.

- Task: high_value_finished_orders

  AnalyticDB for PostgreSQL uses `o_totalprice > 10000` to filter orders whose total price is greater than 10,000 from the finished_orders table and writes the orders into the high_value_finished_orders table.

Create tasks

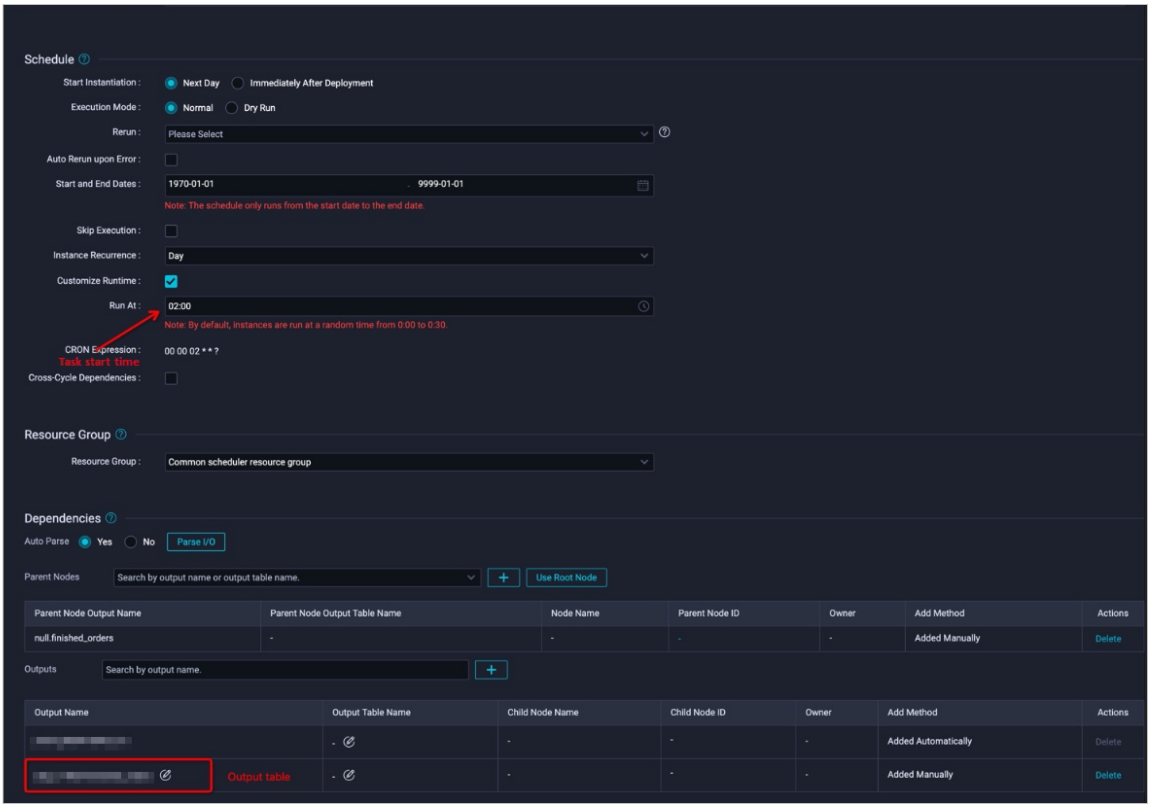
For more information, visit AnalyticDB for PostgreSQL node.

Configure dependencies among tasks

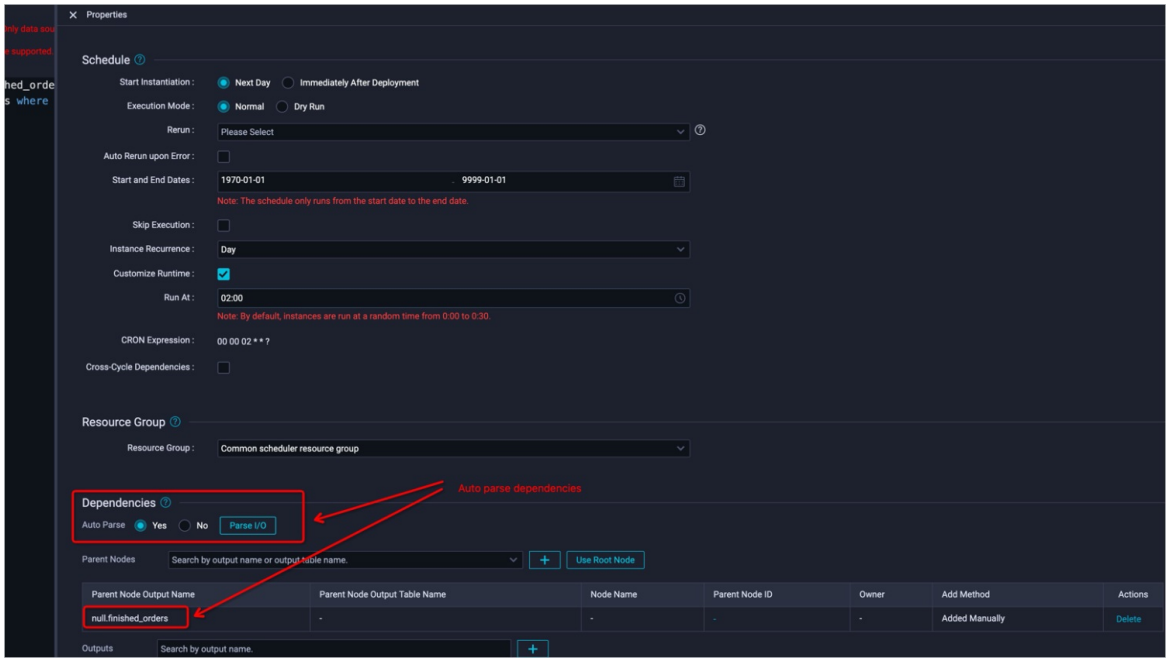
The core of task scheduling is that multiple tasks run at a specific time based on specified dependencies.

For example, the `finished_orders` task runs at 02:00 every day, and the `high_value_finished_orders` task runs after the `finished_orders` task succeeds. To configure a dependency between the two tasks, follow these steps:

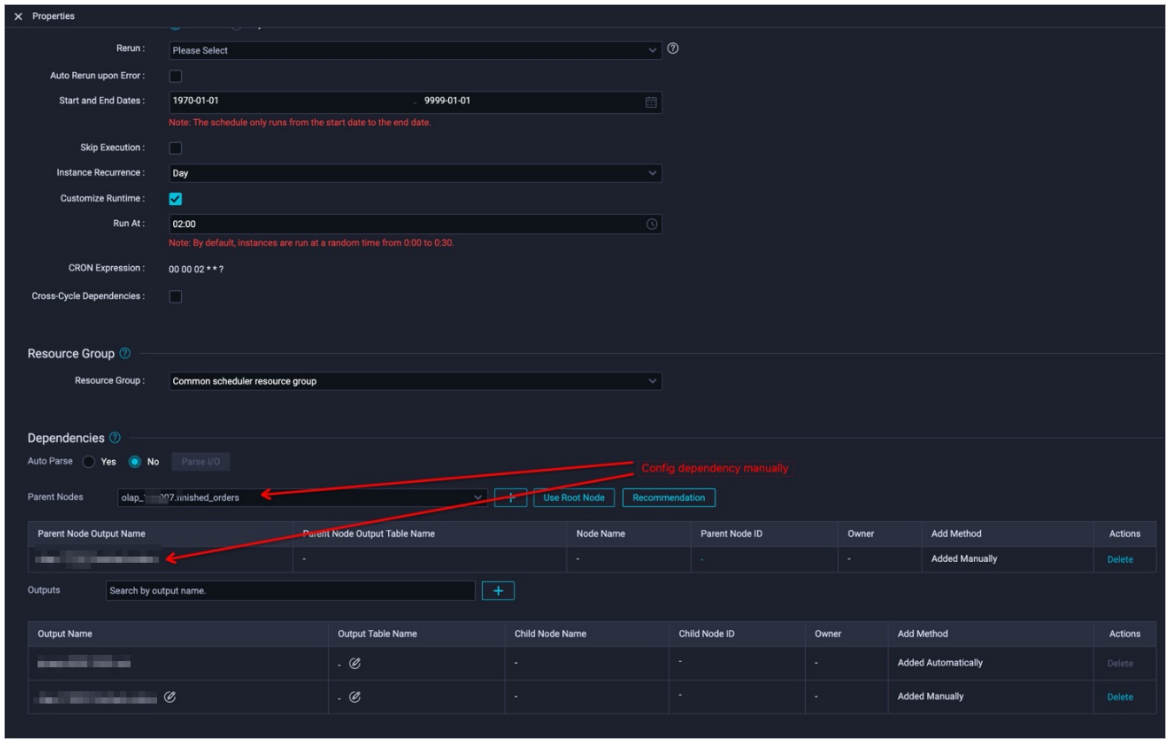
1. Set the runtime of the `finished_orders` task.



2. Configure a dependency for the `high_value_finished_orders` task.

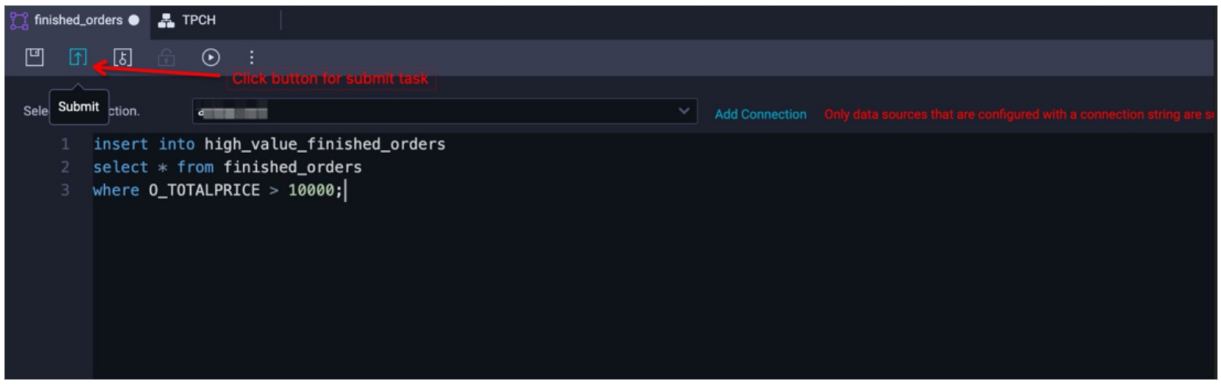


If a task dependency is not parsed automatically, you must manually specify an upper-level dependent node first.

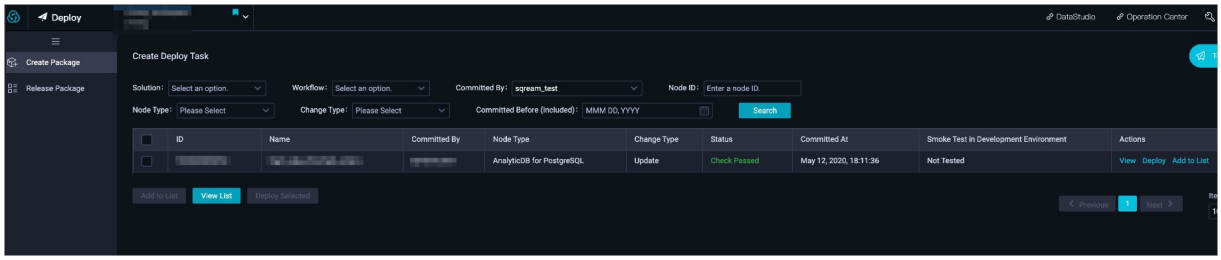


Publish tasks

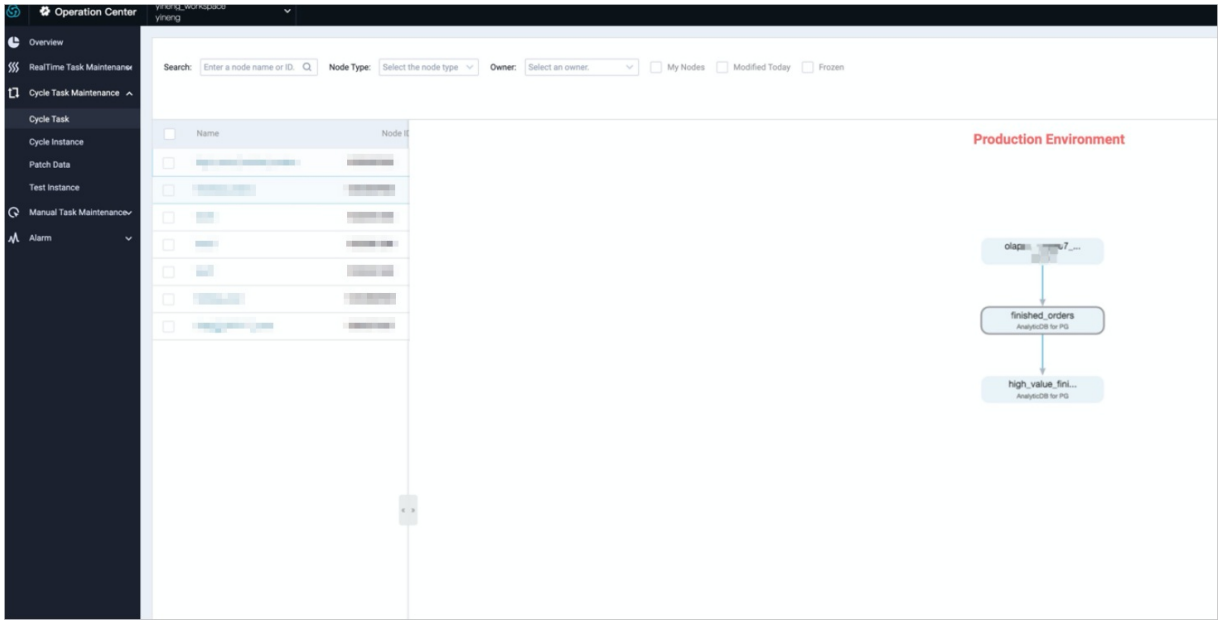
Select the submitted tasks and click the publish icon.



You can view published tasks on the published task list page.



After the tasks are published, you can maintain them on the task O&M page.


Operation Center
Overview
RealTime Task Maintenance
Cycle Task Maintenance
Cycle Task
Cycle Instance
Patch Data
Test Instance
Manual Task Maintenance
Alarm
Search:
Enter a node name or ID.
Node Type:
Select the node type
Owner:
Select an owner.
My Nodes
Modified Today
Frozen
Name
Production Environment
finished_orders
high_value_fini...