ModelArts

Billing

Issue 01

Date 2024-05-14

Huawei Cloud Computing Technologies Co., Ltd.

Address: Huawei Cloud Data Center Jiaoxinggong Road
              Qianzhong Avenue
              Gui'an New District
              Gui Zhou 550029
              People's Republic of China

Website: https://www.huaweicloud.com/intl/en-us/

Contents

1 Billing Modes

1.1 Overview

ModelArts provides two billing modes: yearly/monthly and pay-per-use to meet your requirements in different scenarios.

- Yearly/Monthly is a prepaid billing mode where resources are billed based on the purchase duration. You get a greater discount if you purchase a longer period. This mode is ideal for scenarios where your compute resource demand is stable during the purchase duration.

- Pay-per-use is a postpaid billing mode where ModelArts compute resources are billed based on the usage duration, in seconds. A bill is generated on the hour. This mode allows flexible resource usage based on service requirements, eliminating the need for pre-provisioning and preventing over or under-allocation of resources. This mode applies if resource requirements fluctuate. You only need to pay for what you have used.

Table 1-1 lists the differences between the two billing modes.

Table 1-1 Differences

Billing Mode	Yearly/Monthly	Pay-per-use
Payment Method	Prepaid Billed by the purchase period specified in your order	Postpaid Billed by usage duration
Billing Cycle	Billed by the purchase period specified in the order	Billed by the second. A bill is generated on the hour.
Applicable Billing Item	Compute resources (vCPUs and GPUs)	Compute resources (vCPUs and GPUs) and EVSs

Applicable Resource Pool	Dedicated resource pools	Public resource pools and dedicated resource pools
Applicable Function	ExeML, notebook, model training, and service deployment	ExeML, notebook, model training, and service deployment
Flavor Change	Supported	Supported
Application Scenario	This cost-effective mode is ideal when the duration of resource usage is predictable. This mode is recommended for long-term users.	This mode is suitable when you want more flexibility and control on compute resource usage.

1.2 Yearly/Monthly

Yearly/Monthly is a billing mode where you pay before using the resources. It is suitable when you have stable resource demands and want to reduce costs. This mode allows you to purchase cloud service resources beforehand with price discounts. This section describes the billing rules of yearly/monthly ModelArts resources.

Application Scenario

The yearly/monthly billing mode requires you to pay for a subscription period before using the resources. It is suitable for long-term users with stable demands. The following are some service scenarios applicable to this billing mode:

- Stable resource demands: For services with long-term, stable resource requirements, such as large model training that requires a long time, the yearly/monthly billing mode offers high cost-effectiveness.

- Long-term projects: For long-term projects, such as scientific model training, the yearly/monthly billing mode ensures stable resource usage throughout the project period.

Constraints

Yearly/Monthly resource pools cannot be used across regions.

Applicable Billing Item

Compute resources (vCPUs and GPUs)

Table 1-2 Applicable billing items

Billing Item		Description
Compute resource	Dedicated resource pools	vCPU and GPU

ModelArts provides yearly/monthly dedicated resource pools. To purchase one, do as follows: On the ModelArts management console, choose **Dedicated Resource Pools** > **Elastic Cluster**. On the **Resource Pools** tab, click **Create**. Then, set **Billing Mode** to **Yearly/Monthly** and view the price details in the lower left corner of the page. The price is calculated based on the resources selected for the dedicated resource pool, including vCPUs and GPUs.

Billing Cycle

The billing cycle of yearly/monthly resources is determined by the purchase duration (UTC+08:00). A billing cycle starts from the time when you enable or renew a resource (accurate to seconds) and ends at 23:59:59 on the expiration date.

For example, if you purchase a dedicated resource pool for one month at 15:50:04 on March 8, 2023, the billing cycle is from 15:50:04 on March 8, 2023 to 23:59:59 on April 8, 2023.

Billing Example

Assume that you purchase a yearly/monthly dedicated resource pool on March 8, 2023 at 15:50:04. The flavor **GPU: 1*nvidia-t4 | CPU: 8 vCPUs 32GB**, and one compute node are used. The billing resource is GPU. Your subscription is one month and you manually renew it for another month before it expires. The billing cycles are as follows:

- The first billing cycle is from 15:50:04 on March 8, 2023 to 23:59:59 on April 8, 2023.

- The second billing cycle is from 23:59:59 on April 8, 2023 to 23:59:59 on May 8, 2023.

You must prepay for each billing cycle. **Table 1-3** shows the billing formula.

Table 1-3 Billing formula

Resource Type	Billing Formula	Unit Price
Compute resource	Specification unit price x Number of compute nodes x Purchase duration	For details, see **ModelArts Pricing Details**.

The fee in the preceding example is calculated as follows:

Fee of the yearly/monthly dedicated resource pool = 625.10 x 1 x 2 = $1250.20 USD

NOTICE

The prices in this document are for reference only. The actual prices are subject to **ModelArts Pricing Details**.

Billing Impacts Following Configuration Changes

If the specifications of a yearly/monthly compute resource do not meet your service requirements, you can change them on the ModelArts management console and pay for the new specifications.

- Upgrades: You change the current specification of a resource to a new, more expensive specification and then pay for the difference.

- Downgrades: You change the current specification of a resource to a new, less expensive specification and Huawei Cloud refunds the difference.

Avoid degrading resource specifications as it will negatively affect ModelArts performance. Example of upgrade fee (without any discount): Assume that you purchase a dedicated resource pool for one month on April 8, 2023. The flavor **GPU: 1*nvidia-t4 | CPU: 8 vCPUs 32GB**, and one compute node are used. On April 18, 2023, you add one compute node. The price of the configuration becomes $1,250.20 USD/month from $625.10 USD/month. The formula is as follows:

Upgrade fee = Price of the new specification x Remaining duration – Price of the old specification x Remaining duration

The remaining duration in the formula is the sum of the remaining days in each calendar month divided by the maximum days in the corresponding calendar month. In this example, the remaining duration = 12 (Remaining days in April)/30 (Maximum days in April) + 8 (Remaining days in May)/31 (Maximum days in May) = 0.6581. The specification upgrade fee is calculated as follows: 1,250.20 x 0.6581 – 625.10 x 0.6581 = $411.37 USD.

For details, see **Pricing of a Changed Specification**.

Impacts of Expiration

Figure 1-1 describes the status of yearly/monthly resources in each period. After purchasing a resource, it remains operational throughout the billing cycle. This can be referred to as the validity period. If the resource is not renewed after expiration, it enters the grace period and then the retention period.

Figure 1-1 Periods of yearly/monthly resources
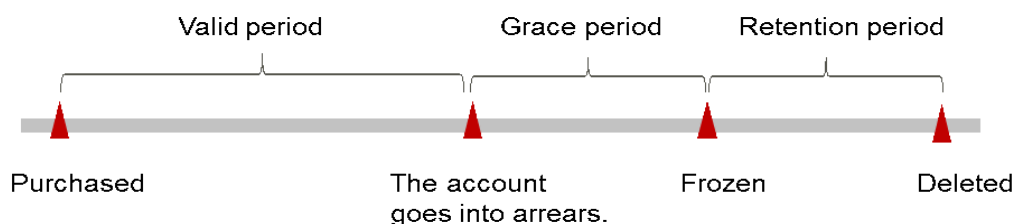
Expiration warning

You will be notified seven days before your yearly/monthly dedicated resource pool expires. Warning messages will be sent by email, SMS, and internal message.

Impacts of expiration

If your yearly/monthly resource pool is not renewed after expiration, it first enters the grace period and its status changes to **Expired**. During the grace period, you can access ModelArts, but the following operations will be restricted:

- Changing resource pool specifications
- Unsubscribing from the resource pool
- Changing the billing mode of the resource pool

If you do not renew the yearly/monthly resource pool within the grace period, the resource enters the retention period and the resource status changes to **Frozen**. You cannot perform any operation on the yearly/monthly resource in the retention period.

If the yearly/monthly resource pool is not renewed after the retention period ends, compute resources (vCPUs and GPUs) will be released and data cannot be restored.

📖 **NOTE**

- Both the grace period and retention period are 15 days.

1.3 Pay-per-Use

Pay-per-use is a billing mode where you pay after using the resources. This billing mode does not require you to make any prepayments or long-term commitments. This section describes the billing rules for pay-per-use resources.

Application Scenario

Pay-per-use is ideal when your resource needs fluctuate. For instance, in the context of AI-generated content (AIGC) inference for To Customer (ToC) services, customer service volumes fluctuate regularly over time. Adopting this billing mode can significantly reduce customers' service costs. You can select resource flavors on the pages for running ExeML jobs, creating notebook instances, creating training jobs, and deploying model services.

Constraints

Pay-per-use resource pools cannot be used across regions.

Applicable Billing Item

Compute resources can be billed in this mode.

Table 1-4 Applicable billing items

Billing Item		Description
Compute resource	Public resource pools	vCPU and GPU
	Dedicated resource pools	

ModelArts allows pay-per-use billing for both public and dedicated resource pools.

- Assume that you need to purchase a pay-per-use dedicated resource pool. On the ModelArts management console, choose **Dedicated Resource Pools** > **Elastic Cluster**. On the **Resource Pools** tab, click **Create**. Then, set **Billing Mode** to **Pay-per-use** and view the price details in the lower left corner of the page. The price is calculated based on the resources selected for the resource pool.
- Pay-per-use public resource pools cannot be purchased on the management console. You can directly select a public resource pool for AI development (including running ExeML jobs, creating notebook instances, creating training jobs, and deploying real-time services). For example, when you create a training job, you can view the price details at the bottom of the page.

Figure 1-2 Price details for creating a training job
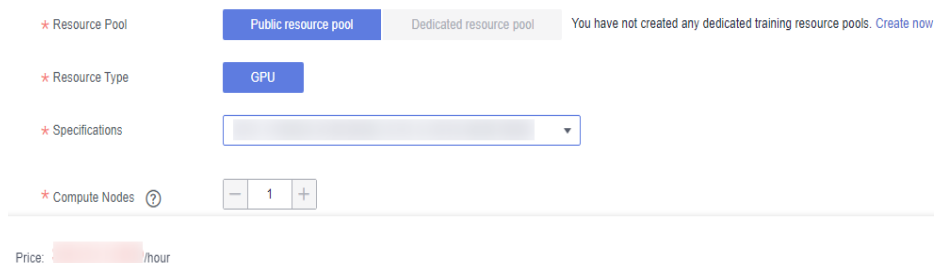
Billing Cycle

Pay-per-use resources are billed by second. A bill is generated on the hour (UTC +08:00). After the bill is generated, a new billing cycle starts.

📖 **NOTE**

- For a dedicated resource pool, the billing starts when it is created, and ends when it is deleted.

  Creating a dedicated resource pool takes time. The billing starts when the dedicated resource pool is enabled, not during its creation process. You can view the time you obtained a dedicated resource pool in the **Basic Information** area on the details page, and view the time when the resource pool is successfully created in the **Occurred At** column of the event whose **Details** is **Pool status changed, from Creating to Running.**

- For a public resource pool, the billing starts when an instance using the public resource pool is created, and ends when the instance is deleted.

  The instance can be:

  - Training jobs and real-time services created for running ExeML jobs

  - Notebook instances

  - Training jobs

  - Real-time services

For example, if you purchase a pay-per-use dedicated resource pool with compute resources (vCPUs) at 08:45:30 and delete it at 08:55:30, the billing cycle is from 08:00:00 to 09:00:00 and fees are generated from 08:45:30 to 08:55:30. The billing duration in this billing cycle is 600 seconds.

Billing Example

Assume that you purchase a pay-per-use dedicated resource pool (specifications: **CPU: 8 vCPUs 32GB**; number of compute nodes: 1)on April 18, 2023 at 09:59:30. The billing resources are compute resources (vCPUs). If you delete the resource pool on April 18, 2023 at 10:45:46, then:

- The first billing cycle is from 09:00:00 to 10:00:00. Fees are generated from 09:59:30 to 10:00:00. The billing duration in this billing cycle is 30 seconds.

- The second billing cycle is from 10:00:00 to 11:00:00. Fees are generated from 10:00:00 to 10:45:46. The billing duration in this billing cycle is 2,746 seconds.

You must pay for each billing cycle. **Table 1-5** shows the billing formula. The **Product Pricing Details** page displays the hourly price of each resource. The price per second of each resource can be obtained by dividing their hourly price by 3600.

Table 1-5 Billing formula

Resource Type	Billing Formula	Unit Price
Compute resource (vCPU)	Specification unit price x Number of compute nodes x Billing duration	For details, see **ModelArts Pricing Details**.

The fee in the preceding example is calculated as follows:

$0.66 USD/hour x 1 x [(30 + 2746)/3600] hours = $0.66 USD

> **NOTICE**
>
> The prices in this document are for reference only. The actual prices are subject to **ModelArts Pricing Details**.

Billing Impacts Following Configuration Changes

If you change the specifications after purchasing a pay-per-use resource pool, a new order is generated and you are billed based on the price of the new configuration. The old order automatically becomes invalid.

If you have purchased a pay-per-use resource pool and then changed the specifications in an hour, multiple billing records will be generated. The duration between the start and end time of each billing record is the validity period of the corresponding configuration.

For example, if you purchase a pay-per-use dedicated resource pool with the flavor **modelarts.vm.cpu.8ud (8vCPUs 16GiB)** and two compute nodes at 09:00:00, and add two nodes at 09:30:00 (four nodes in total after the upgrade), two billing records are generated between 09:00:00 and 10:00:00.

- The first record corresponds to 09:00:00 to 9:30:00 and two compute nodes are billed.
- The second record corresponds to 09:30:00 to 10:00:00 and four compute nodes are billed.

About Arrears

Figure 1-3 describes the status of pay-per-use resources in each period. After purchasing a resource, it remains operational throughout the billing cycle. This can be referred to as the validity period. If your account is in arrears, the resource enters the grace period and then the retention period.

Figure 1-3 Periods of pay-per-use resources
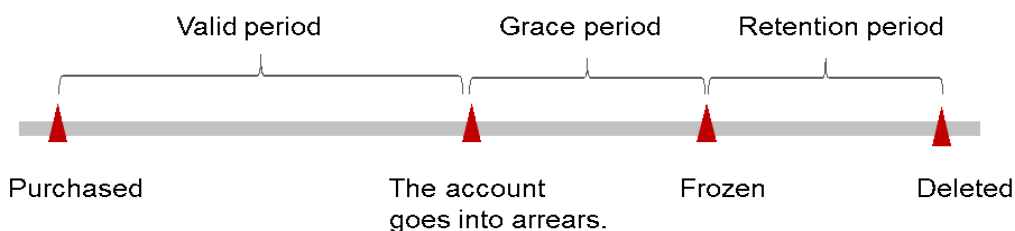
Arrears warning

Fees are deducted based on pay-per-use resources at the end of each billing cycle. When your account is in arrears, you will be notified by email, SMS, and internal message.

Impacts of arrears

After your account is in arrears, the pay-per-use resources will not become unavailable immediately. These resources enter the grace period. You need to pay

fees for pay-per-use resources incurred during the grace period, which are counted as outstanding amount in the Billing Center. Huawei Cloud will automatically deduct the fees when you top up.

If you do not pay the arrears within the grace period, the resources enter the retention period and the resource status changes to **Frozen**. You cannot perform any operation on the pay-per-use resources in the retention period.

If you do not pay the arrears within the retention period, the compute resources will be released and data cannot be restored.

📖 NOTE

- Both the grace period and retention period are 15 days.

Outstanding Balance

If your account is in arrears, some operations will be restricted. Top up your account as soon as possible. **Table 1-6** describes the restricted operations.

Table 1-6 Restricted operations due to arrears

Feature	Restricted Operation
ExeML	Model training and deployment
DevEnviron > Notebook	Creating and starting notebook instances
Training Management > Training Jobs	Creating training jobs
Service Deployment > Real-Time Services	Deploying real-time services
Dedicated Resource Pools	Creating, freezing, unfreezing, and deleting dedicated resource pools

2 Billing Items

2.1 ExeML

Description

Model training and inference in ModelArts ExeML jobs use compute and storage resources, which are billed. For details, see **Table 2-1**.

- Compute resource fee:
  - If dedicated resource pools are used for model training and inference in ExeML jobs, compute resources are not billed.
  - If public resource pools are used for model training and inference in ExeML jobs, compute resources are billed.
- Storage resource fee: fee for storing data in OBS

Table 2-1 Billing items

Billing Item		Description	Billing Mode	Billing Formula
Compute resource	Public resource pools	Usage of compute resources (vCPUs and GPUs) For details, see **ModelArts Pricing Details**.	Pay-per-use	Specification unit price x Number of compute nodes x Usage duration
	Dedicated resource pools	Fees for dedicated resource pools are paid upfront upon purchase. There are no additional charges for running ExeML jobs. For details about fees for dedicated resource pools, see **Dedicated Resource Pool**.	N/A	N/A

Billing Item		Description	Billing Mode	Billing Formula
Stora ge resou rce	Object Storag e Servic e (OBS)	OBS is used to store the input and output data of training and inference. For details, see **OBS Pricing Details**. **CAUTION** OBS resources for storing data are continuously billed. To stop billing, delete the data stored in OBS.	Pay-per-use Yearly/ Monthl y	Creating an OBS bucket is free of charge. You pay only for the storage capacity and duration you actually use.

Billing Example

NOTICE

The following prices are for reference only. For the actual prices, see pricing details for each service.

Example: Running an ExeML job using a public resource pool; billing items: compute resources and standard storage

Assume that you create an ExeML image classification project on April 1, 2023. Data validation is done from 10:00:00 to 10:06:00, image classification is done from 10:06:00 to 11:12, service deployment is done at 11:30:00, and the real-time service is stopped at 12:00:00. A public resource pool is used to run the instance. The compute flavor is **Compute-intensive ECS 1-ExeML (GPU)** (unit price: $6.93 USD/hour) for model training and **Compute-intensive 3 instance (CPU)** for service deployment (unit price: $0.23 USD/hour). The billing items are compute and storage resources. The fee for running this ExeML job is calculated as follows:

- **Compute resource fee** = Specification unit price x Training job running duration + Specification unit price x Service running duration

  Compute resource fee = $6.93 USD/hour x 0.1 hour + $0.23 USD/hour x 0.5 hour = $0.808 USD

- **Storage fee**: ExeML job data is uploaded to or exported from OBS. The storage fee is based on the OBS billing rules.

Fee for running the ExeML job = **Compute resource fee** ($0.808 USD) + **Storage fee**

2.2 Data Management

Description

ModelArts data management allows you to use dataset, data labeling, and data processing functions for free. The details are as follows:

- **Dataset**: Creating datasets in ModelArts is free of charge.
- **Data labeling**: Manual labeling and auto labeling in ModelArts are free of charge.
- **Data processing**: Creating data cleaning, data validation, data selection, and data augmentation tasks in ModelArts are free of charge.

However, you will be billed for storing data in OBS. For details, see **Table 2-2**.

Table 2-2 Billing for storage resources

Billing Item		Description	Billing Mode	Billing Formula
Storage resource	Object Storage Service (OBS)	OBS is used to store the input and output data of training and inference. For details, see **OBS Pricing Details**. **CAUTION** 　OBS resources for storing data are continuously billed. To stop billing, delete the data stored in OBS.	Pay-per-use Yearly/ Monthly	Creating an OBS bucket is free of charge. You pay only for the storage capacity and duration you actually use.

Billing Example

> **NOTICE**
>
> The following prices are for reference only. For the actual prices, see pricing details for each service.

Example: storage fee

Assume that you create a dataset at 10:00:00 on April 1, 2023 and the data is stored in OBS. The billing item is storage resources. The fee for storing data in OBS is calculated as follows:

- **Storage fee**: Data in the data management module is uploaded to or exported from OBS. The storage fee is based on the OBS billing rules. For details, see **OBS Pricing Details**.

In conclusion, **the data management fee** is equal to **the storage fee**.

2.3 Development Environment

Description

Running notebook instances in ModelArts development environments use compute and storage resources, which are billed. For details, see **Table 2-3**.

⚠ **CAUTION**

EVS is still billed even after the notebook instance is stopped. To stop billing, delete the notebook instance.

- Compute resource fee:
  - If dedicated resource pools are used for running notebook instances, compute resources are not billed.
  - If public resource pools are used for running notebook instances, compute resources are billed.
- Storage resource fee: fee for storing data in OBS, EVS, or SFS

Table 2-3 Billing items

Billing Item		Description	Billing Mode	Billing Formula
Compute resource	Public resource pools	Usage of compute resources (vCPUs and GPUs) For details, see **ModelArts Pricing Details**.	Pay-per-use	Specification unit price x Number of compute nodes x Usage duration
	Dedicated resource pools	Fees for dedicated resource pools are paid upfront upon purchase. There are no additional charges for running notebook instances. For details about fees for dedicated resource pools, see **Dedicated Resource Pool**.	N/A	N/A

Billing Item		Description	Billing Mode	Billing Formula
Storage resource	EVS	EVS is used to store data generated for running notebook instances. The default disk size is 5 GB. You are billed by GB from the time when a notebook instance is created to the time when it is deleted. For details, see **EVS Pricing Details**. **CAUTION** 　EVS resources for storing data are continuously billed. To stop billing, delete the data stored in EVS.	Pay-per-use Yearly/ Monthly	Specification unit price x Purchase duration x Purchase quantity
	SFS	SFS is used to store data generated for running notebook instances. For details, see **SFS Pricing Details**. **CAUTION** 　SFS resources for storing data are continuously billed. To stop billing, delete the data stored in SFS.	Pay-per-use Yearly/ Monthly	Pay-per-use: Specification unit price x Purchase duration Yearly/Monthly: Specification unit price x Purchase duration x Purchase quantity
	Object Storage Service (OBS)	OBS is used to store input and output data generated for running notebook instances. For details, see **OBS Pricing Details**. **CAUTION** 　OBS resources for storing data are continuously billed. To stop billing, delete the data stored in OBS.	Pay-per-use Yearly/ Monthly	Creating an OBS bucket is free of charge. You pay only for the storage capacity and duration you actually use.

Billing Example

> **NOTICE**
>
> The following prices are for reference only. For the actual prices, see pricing details for each service.

Example: Using a public resource pool; billing items: compute resources and EVS storage

Assume that you create a notebook instance on April 1, 2023 at 10:00:00. A public resource pool with flavor **CPU: 8 vCPUs 32GB** and one compute node is used,

along with a 5 GB disk. (Unit price: $0.573 USD/hour). Then, you delete the notebook instance at 13:00:00 on the same day. The billing items are compute and storage resources. The fee for running this notebook instance is calculated as follows:

- **Resource fee** = **Compute resource fee** + **Storage fee** = $0.573 USD/hour x 1 x 3 hours = $1.719 USD

In conclusion, the fee for running the notebook instance = $1.719 USD.

Example: Using a dedicated resource pool; billing item: EVS storage

Assume that you create a notebook instance on April 1, 2023 at 10:00:00. A dedicated resource pool and a 5 GB disk are used. You delete the data stored in EVS on April 2, 2023 at 10:00:00. The billing item is EVS storage. The fee for running the notebook instance is as follows:

- **Storage fee** = $0.002 USD/hour x 24 hours = $0.048 USD

In conclusion, the fee for running the notebook instance is $0.048 USD.

2.4 Model Training

Description

Model training in ModelArts uses compute and storage resources, which are billed. Compute resources are billed for running training jobs. Storage resources are billed for storing data in OBS or SFS. For details, see **Table 2-4**.

Table 2-4 Billing items

Billing Item		Description	Billing Mode	Billing Formula
Compute resource	Public resource pools	Usage of compute resources (vCPUs and GPUs) For details, see **ModelArts Pricing Details**.	Pay-per-use	Specification unit price x Number of compute nodes x Usage duration
	Dedicated resource pools	Fees for dedicated resource pools are paid upfront upon purchase. There are no additional charges for model training. For details about fees for dedicated resource pools, see **Dedicated Resource Pool**.	N/A	N/A

Billing Item		Description	Billing Mode	Billing Formula
Storage resource	Object Storage Service (OBS)	OBS is used to store the input and output data of model training. For details, see **OBS Pricing Details**. **CAUTION** OBS resources for storing data are continuously billed. To stop billing, delete the data stored in OBS.	Pay-per-use Yearly/Monthly	Creating an OBS bucket is free of charge. You pay only for the storage capacity and duration you actually use.
	Scalable File Service (SFS) Turbo	When a dedicated resource pool is used for training, multiple SFS Turbo file systems can be mounted for one training job. OBS is used to store the code and input and output data of model training. For details, see **SFS Pricing Details**.	Pay-per-use Yearly/Monthly	Billed based on used storage space of the file system and usage duration Fees = Price per GB-hour x Storage space used x Hours of use
Event notification (billed only when enabled)		This function uses Simple Message Notification (SMN) to send a message to you when the event you selected occurs. To use this function, enable event notification when creating a training job. For pricing details, see **Simple Message Notification Pricing Details**.	Pay by actual usage	• SMS: SMS notifications • Email: Email notifications + Downstream Internet traffic • HTTP or HTTPS: HTTP or HTTPS notifications + Downstream Internet traffic

Billing Example

NOTICE

The following prices are for reference only. For the actual prices, see pricing details for each service.

Example: Using a public resource pool; billing items: compute resources and storage

Assume that you create a training job on April 1, 2023 at 10:00:00. The training job uses a public resource pool with specification **CPU: 8 vCPUs 32GB** and one

compute node (unit price: $0.66 USD/hour) and is stopped at 11:00:00. The billing items are compute and storage resources. The fee for running this training job is calculated as follows:

- **Compute resource fee** = $0.66 USD/hour x 1 hour = $0.66 USD

- **Storage fee**: Training job data is uploaded to or exported from OBS. The storage fee is based on the OBS billing rules.

Fee for running the training job = **Compute resource fee** ($0.66 USD) + **Storage fee**

Example: Using a dedicated resource pool; billing item: storage

Assume that you create a training job using a dedicated resource pool on April 1, 2023 at 10:00:00 and stop the job at 11:00:00. The billing item is storage. The fee for running the training job is as follows:

- **Storage fee**: Training job data is uploaded to or exported from OBS. The storage fee is based on the OBS billing rules.

In conclusion, **the fee for running the training job** is equal to **the storage fee**.

2.5 AI Application

Description

Creating an AI application in ModelArts is free of charge. However, when creating an AI application using a model imported from OBS, you must store the model package file in OBS, which will incur fees. For details, see **Table 2-2**.

Table 2-5 Billing for storage resources

Billing Item		Description	Billing Mode	Billing Formula
Storage resource	Object Storage Service (OBS)	OBS is used to store the input and output data of training and inference. For details, see **OBS Pricing Details**. CAUTION OBS resources for storing data are continuously billed. To stop billing, delete the data stored in OBS.	Pay-per-use Yearly/ Monthly	Creating an OBS bucket is free of charge. You pay only for the storage capacity and duration you actually use.

Billing Example

Example: Storage fee

Assume that you upload the model package required for creating an AI application to an OBS bucket at 10:00 on April 1, 2023. The billing item is storage resources. The fee for creating the AI application is calculated as follows:

- **Storage fee**: The model package required for creating the AI application is uploaded to or exported from OBS. The storage fee is based on the OBS billing rules. For details, see **OBS Pricing Details**.

In conclusion, **the fee for creating the AI application** is equal to **the storage fee**.

2.6 Inference Deployment

Description

Deploying a service in ModelArts uses compute and storage resources, which are billed. Compute resources are billed for running the inference service. Storage resources are billed for storing data in OBS. For details, see **Table 2-6**.

Table 2-6 Billing items

Billing Item		Description	Billing Mode	Billing Formula
Compute resource	Public resource pools	Usage of compute resources (vCPUs and GPUs) For details, see **ModelArts Pricing Details**.	Pay-per-use	Specification unit price x Number of compute nodes x Usage duration
	Dedicated resource pools	Fees for dedicated resource pools are paid upfront upon purchase. There are no additional charges for service deployment. For details about fees for dedicated resource pools, see **Dedicated Resource Pool**.	N/A	N/A
Storage resource	Object Storage Service (OBS)	OBS is used to store the input and output data for service deployment. For details, see **OBS Pricing Details**. **CAUTION** OBS resources for storing data are continuously billed. To stop billing, delete the data stored in OBS.	Pay-per-use Yearly/Monthly	Creating an OBS bucket is free of charge. You pay only for the storage capacity and duration you actually use.

Billing Item	Description	Billing Mode	Billing Formula
Event notification (billed only when enabled)	This function uses Simple Message Notification (SMN) to send a message to you when the event you selected occurs. To use this function, enable event notification when creating a training job. For pricing details, see **Simple Message Notification Pricing Details**.	Pay by actual usage	• SMS: SMS notifications • Email: Email notifications + Downstream Internet traffic • HTTP or HTTPS: HTTP or HTTPS notifications + Downstream Internet traffic
Run logs (billed only when enabled)	**Log Tank Service (LTS)** collects, analyzes, and stores logs. If **Runtime Log Output** is enabled during service deployment, you will be billed if the log data exceeds the LTS free quota. For details, see **LTS Pricing Details**.	Pay by actual log size	After the free quota is exceeded, you are billed based on the actual log volume and retention duration.

Billing Example

> **NOTICE**
>
> The following prices are for reference only. For the actual prices, see pricing details for each service.

Example: Using a public resource pool; billing items: compute resources and storage

Assume that you create a real-time service that uses a public resource pool on April 1, 2023 at 10:00:00 and stop it at 11:00:00. The specifications are as follows: **CPU: 8 vCPUs 32GB** and one compute node. The unit price is $1.3023 USD/hour. The billing items are compute and storage resources. The fee for running this real-time service is calculated as follows:

- **Compute resource fee** = $1.3023 USD/hour x 1 hour = $1.3023 USD
- **Storage fee**: Training job data is uploaded to or exported from OBS. The storage fee is based on the OBS billing rules.

Fee for running the real-time service = **Compute resource fee** ($1.3023 USD) + **Storage fee**

Example: Using a dedicated resource pool; billing item: storage

Assume that you create a real-time service using a dedicated resource pool on April 1 at 10:00:00, 2023 and stop the job at 11:00:00. The billing item is storage. The fee for running the real-time service is as follows:

- **Storage fee**: Training job data is uploaded to or exported from OBS. The storage fee is based on the OBS billing rules.

In conclusion, **the fee for running the real-time service** is equal to **the storage fee**.

2.7 Dedicated Resource Pool

Description

When using ModelArts for AI development, compute resources are billed for running ExeML, development environments, training jobs, and deployed services. For details, see **Table 2-6**.

Table 2-7 Billing items

Billing Item		Description	Billing Mode	Billing Formula
Com pute resou rce	De dic ate d res our ce poo ls	Usage of compute resources (vCPUs and GPUs) For details, see **ModelArts Pricing Details**.	Pay-per-use	Specification unit price x Number of compute nodes x Usage duration
			Yearly/Monthly	Specification unit price x Number of compute nodes x Purchase duration

Billing Example

> **NOTICE**
>
> The following prices are for reference only. For the actual prices, see pricing details for each service.

Example: Using a pay-per-use dedicated resource pool; billing item: compute resource

Assume that you create a pay-per-use dedicated resource pool on April 1, 2023 at 10:00:00 and delete it on May 1, 2023 at 10:00:00. The flavor **CPU: 8 vCPUs 32GB** and one compute node are used. The unit price is $0.66/hour. The billing item is compute resources. The fee for running the dedicated resource pool is calculated as follows:

- **Compute resource fee** = $0.66 USD/hour x 1 x (30 x 24) hours = $475.20 USD

Fee for running the dedicated resource pool = $475.20 USD.

Example: Using a yearly/monthly dedicated resource pool; billing item: compute resource

Assume that you create a yearly/monthly dedicated resource pool on April 1, 2023 at 10:00:00. The flavor **GPU: 1*nvidia-t4 | CPU: 8 vCPUs 32GB (modelarts.vm.gpu.t4u8)** and one compute node are used. The purchase duration is two months. The unit price is $625.10 USD. The billing item is compute resources. The fee for running the dedicated resource pool is calculated as follows:

- **Compute resource fee** = $625.10 USD x 2 = $1,250.20 USD

In conclusion, **the fee for running the dedicated resource pool** is $1,250.20 USD

3 Billing Examples

Scenario 1

At 15:30:00 on March 18, 2023, a user uses a pay-per-use public resource pool for training. The configuration is as follows:

● Flavor: **CPU: 2 vCPUs 8GB 50GB**

● Number of compute nodes: 1

The training job is stopped on March 20, 2023 at 10:30:00. How much is the total cost of the public resource pool in March?

Analysis

From 15:30:00 on March 18, 2023 to 9:00:00 on March 20, 2023, you are billed for flavor **CPU: 2 vCPUs 8GB 50GB** and one node. The unit price is $0.15 USD/hour and the billing duration is 41.5 hours. The fee is calculated as follows:

$0.15 USD/hour x 41.5 hours = $6.225 USD

The total fee generated by the dedicated resource pool in March is $6.225 USD.

Scenario 2

At 15:30:00 on March 18, 2023, a user purchases a pay-per-use dedicated resource pool. The configuration is as follows:

● Flavor: **CPU: 8 vCPUs 32GB (modelarts.vm.cpu.8ud)**

● Number of compute nodes: 1

After a period of time, the user finds that the specifications cannot meet service requirements and adds one node on March 20, 2023 at 9:00:00. At 10:30:00 on March 20, 2023, the user changes the billing mode of the dedicated resource pool to yearly/monthly. The purchase duration is one month. After expiration, the user will unsubscribe from the dedicated resource pool. How much is the total fee generated by the dedicated resource pool from March to April?

Analysis

The usage of the dedicated resource pool can be divided into two phases based on the billing mode. The pay-per-use billing mode is used from 15:30:00 on March 18,

2023 to 10:30:00 on March 20, 2023. The yearly/monthly billing mode is used from 10:30:00 on March 20, 2023 to 23:59:59 on April 20, 2023.

Pay-per-use

From 15:30:00 on March 18, 2023 to 9:00:00 March 20, 2023, you are billed for flavor **CPU: 8 vCPUs 32GB** and one node. The unit price is $0.66 USD/hour and the billing duration is 41.5 hours. The fee is calculated as follows:

$0.66 USD x 41.5 hours = $27.39 USD

From 9:00:00 on March 20, 2023 to 10:30:00 on March 20, 2023, you are billed for flavor **CPU: 8 vCPUs 32GB** and two nodes. The unit price is $1.32 USD/hour and the billing duration is 1.5 hours. The fee is calculated as follows:

$1.32 USD/hour x 1.5 hours = $1.98 USD

Yearly/Monthly

From 10:30:00 on March 20, 2023 to 23:59:59 on April 20, 2023, the yearly/monthly billing mode is used for one month. The unit price is $625.10 USD. The fee is calculated as follows:

$625.10 USD x 1 = $625.10 USD

The total fee of the dedicated resource pool from March to April is 27.39 + 1.98 + 625.10 = $654.47 USD.

Scenario 3

A company needs to use ModelArts to train and develop models and deploy services. A resource pool with flavor CPU: 8 vCPUs 32GB is used for one month. The company wants to know which billing mode is more cost-effective.

Analysis

The following shows the expenditure details of the pay-per-use and yearly/monthly billing modes.

NOTICE

The unit price in this case is for reference only, and the calculated fee is an estimated value. The actual data is subject to that released on the Huawei Cloud official website.

When ModelArts is used for AI development, data is stored in OBS, EVS, or SFS. In this case, storage fees are generated based on the OBS, EVS, and SFS billing standards.

Billing Mode	Calculation Rule	Billing Cycle	Fee Calculation
Pay-per-use	Compute resource fee: Specification unit price x Duration	Hourly fee	Compute resource fee: $1.3023 USD/hour x 720 hours = $937.656 USD
Yearly/ Monthly	Compute resource fee: Specification unit price x Duration	Monthly fee	Compute resource fee: $625.10 USD/month x 1 month = $625.10 USD

This table shows that the yearly/monthly billing is more effective when you have long-term resource demand. The longer the resource package is purchased, the more the discount is.

4 Changing the Billing Mode

4.1 Overview

If the billing mode of your dedicated resource pool cannot meet your service requirements, you can change the billing mode. **Table 4-1** lists the billing items that support billing mode change.

📖 **NOTE**

Public resource pools cannot be purchased separately, and the billing mode cannot be changed.

Table 4-1 Billing items that support billing mode change

Billing Item	Change Description	Reference
Compute resource	Changing the billing mode of a dedicated resource pool will also change the billing mode of its compute resources. • You can change the billing mode of a dedicated resource pool from pay-per-use to yearly/monthly to enjoy price discounts. • You can change the billing mode of a dedicated resource pool from yearly/monthly to pay-per-use to flexibly use ModelArts compute resources. **NOTE** The pay-per-use mode takes effect only after a yearly/monthly period expires.	• **Changing the Billing Mode from Pay-per-Use to Yearly/ Monthly** • **Changing the Billing Mode from Yearly/ Monthly to Pay-per-Use**

4.2 Changing the Billing Mode from Pay-per-Use to Yearly/Monthly

If you want to use a pay-per-use dedicated resource pool for a long time, you can change the billing mode of it to yearly/monthly to reduce costs. By doing so, a new order is generated. After you pay the order, the yearly/monthly resource takes effect immediately.

Assume that you purchase a pay-per-use dedicated resource pool on April 18, 2023 at 15:29:16. Due to service requirements, you change the billing mode from pay-per-use to yearly/monthly on April 18, 2023 at 16:30:30 and a change-to-yearly/monthly order is generated. After you pay for the order, the dedicated resource pool immediately enters the yearly/monthly billing mode. The following bills are generated in the Billing Center:

- Pay-per-use: 2023/04/18 15:00:00 to 2023/04/18 16:00:00. Fees are generated between 15:29:16 and 16:00:00.
- Pay-per-use: 2023/04/18 16:00:00 to 2023/04/18 17:00:00. Fees are generated between 16:00:00 and 16:30:30.
- Yearly/Monthly: 2023/04/18 16:30:30

Constraints

- The billing mode of the target dedicated resource pool must be pay-per-use.
- Only the billing mode of those resources in the **Enabled** status can be changed.
- The billing mode can be changed only for dedicated resource pools, but not for flavors.
- All nodes of the target dedicated resource pool must use the same billing mode (pay-per-use or yearly/monthly).

Procedure

Step 1 Log in to the management console, click ☰ in the navigation pane, and choose **AI** > **ModelArts**.

Step 2 On the ModelArts management console, choose **Dedicated Resource Pools** > **Elastic Cluster** and locate the target dedicated resource pool. Choose **More** > **Change To Prepaid** in the **Operation** column.

Step 3 In the **Change To Prepaid** dialog box, confirm the information and click **OK**.

Figure 4-1 Change To Prepaid
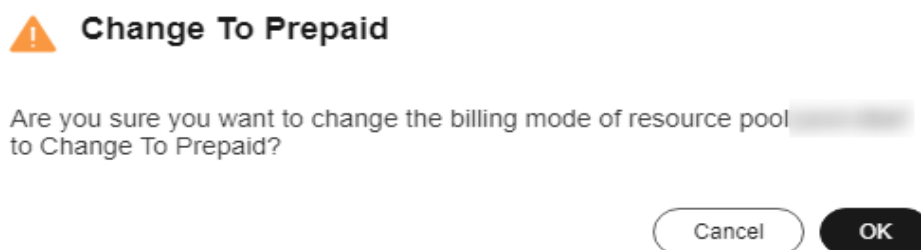
Step 4 Select the usage duration of the elastic cluster, select **Auto-Renew** as required, confirm the estimated expiration time and configuration fee, and click **Pay**.

Step 5 Select a payment method and confirm the payment. After you pay the order, the change is complete.

----End

4.3 Changing the Billing Mode from Yearly/Monthly to Pay-per-Use

You can change the billing mode of a dedicated resource pool from yearly/monthly to pay-per-use for flexible use.

◻ **NOTE**

The pay-per-use billing mode will take effect only after the yearly/monthly subscription has expired.

Constraints

You can change the billing mode from yearly/monthly to pay-per-use only for dedicated resource pools whose status is **Provisioned**.

Procedure

Step 1 Log in to the management console, click ☰ in the navigation pane, and choose **AI** > **ModelArts**.

Step 2 On the ModelArts management console, choose **Dedicated Resource Pools** > **Elastic Cluster** and locate the target dedicated resource pool. Choose **More** > **Change Billing Mode** in the **Operation** column.

Step 3 In the **Change Billing Mode** dialog box, confirm the information and click **OK**.

Figure 4-2 Changing the billing mode
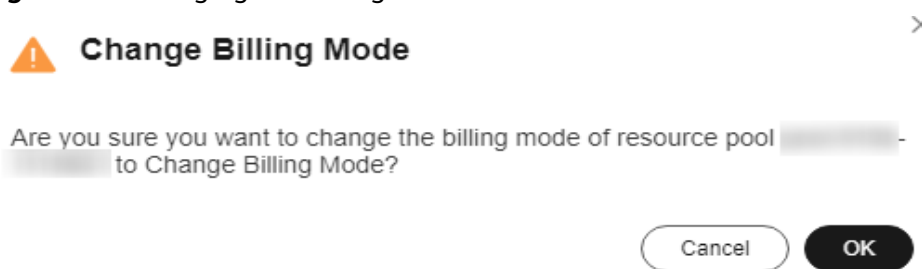
Step 4 Go to the **Billing Center > Renewals** page.

Step 5 Customize search criteria.

- Under the **Pay-per-Use After Expiration** tab, you can search for the resources with billing mode set to be pay-per-use upon expiration.
- Under the **Manual Renewals**, **Auto Renewals**, and **Renewals Canceled** tabs, you can also change the billing mode of the resources to be pay-per-use upon expiration.

Step 6 Change the billing mode of resources from yearly/monthly to pay-per-use after expiration.

- Single resource: Choose **More** > **Change to Pay-per-Use After Expiration** in the **Operation** column of the target instance.

Figure 4-3 Changing the billing mode of a single resource from yearly/monthly to pay-per-use after expiration
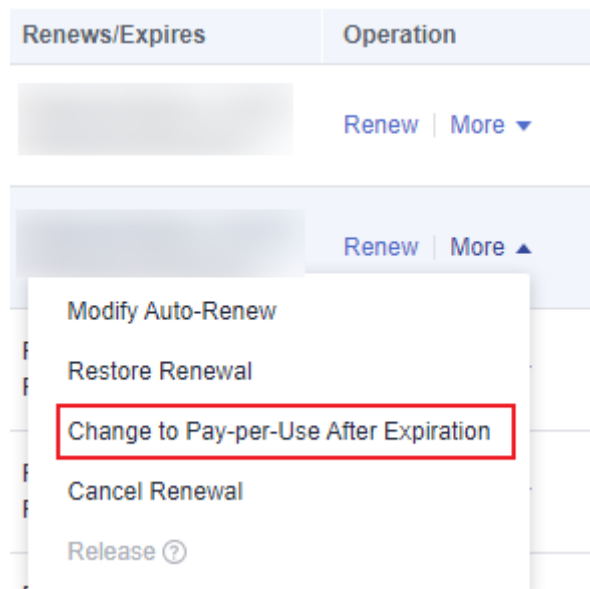
- Multiple resources: Select the instances for which you want to change the billing mode, and click **Change to Pay-per-Use After Expiration** above the instance list.

Step 7 Confirm the change details and click **Change to Pay-per-Use After Expiration**.

----End

5 Renewal

5.1 Overview

Introduction

After a yearly/monthly dedicated resource pool expires, ModelArts services that use this resource pool will not run properly. If you want to continue using the resource pool, renew it within the specified period. Otherwise, vCPU and GPU resources will be automatically released, and data will be lost and cannot be restored.

The renewal operation applies only to yearly/monthly dedicated resource pools. For pay-per-use dedicated resource pools, you only need to ensure that your account balance is sufficient.

If a dedicated resource pool is renewed before it expires, all resources are retained, and ModelArts services using this resource pool will not be affected. For details about the dedicated resource pool status after expiration, see **Impacts of Expiration**.

Functions

Table 5-1 describes the renewal functions for yearly/monthly dedicated resource pools.

Table 5-1 Renewal functions

Function	Description
Manual renewal	You can renew a yearly/monthly dedicated resource pool on the ModelArts management console at any time before it is automatically deleted, in order to extend its validity period.
Auto renewal	After auto renewal is enabled, the dedicated resource pool is automatically renewed before each expiration. This prevents resources from being automatically deleted if you forget to manually renew the resource pool.

You can renew a yearly/monthly dedicated resource pool in any phase of its lifecycle. For details, see **Figure 5-1**.

Figure 5-1 Lifecycle of a dedicated resource pool
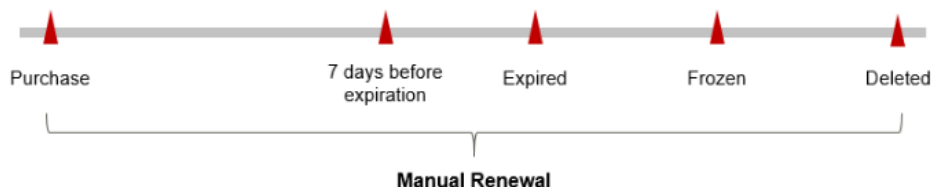
- Before a dedicated resource pool expires, it runs properly and the resource status is **Running**.

- After the resource expires, the resource status changes to **Expired**.

- If the dedicated resource pool is not renewed upon expiration, it enters the grace period. If the resource is not renewed within the grace period, the resource status changes to **Frozen**.

- If the resource pool is not renewed within the grace period, it enters the retention period. If it is not renewed with the retention period, it will be automatically deleted.

☐ NOTE

Both the grace period and retention period are 15 days.

Auto-renewal can be enabled before a dedicated resource pool expires. The system attempts to automatically renew the dedicated resource pool at 03:00 on the seventh day before it expires. If the fee deduction fails, the system attempts to automatically renew the dedicated resource pool at 03:00 every day until the dedicated resource pool expires or the renewal is successful. By default, the dedicated resource pool is renewed seven days before the expiration. You can change the deduction date.

5.2 Manual Renewal

You can renew a yearly/monthly dedicated resource pool on the ModelArts management console at any time before it is automatically deleted, in order to extend its validity period.

## Renewing a Resource in the Billing Center

Step 1 Go to the **Billing Center > Renewals** page.

Step 2 Customize search criteria. On the **Manual Renewals**, **Auto Renewals**, **Pay-per-Use After Expiration**, and **Renewals Canceled** tabs, you can view all resources to be renewed and manually renew the resources.

The **Manual Renewals** tab displays all resources that need to be manually renewed. For details, see **Restoring to Manual Renewal**.

Step 3 Manually renew resources.

- Single resource: Click **Renew** in the **Operation** column of the target resource.

Figure 5-2 Renewing a single resource
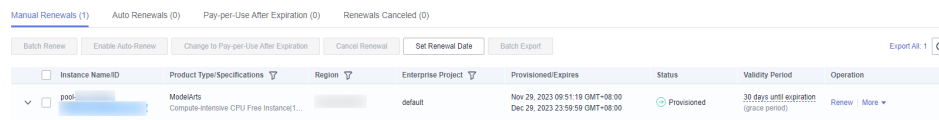
- Multiple resources: Select the check boxes for the target resources, and click **Batch Renew** in the upper left corner.

Step 4 Configure the renewal duration and determine whether to select **Renewal Date** to renew resources on the standard renewal date. For details, see **Standard Renewal Date of Yearly/Monthly Resources**. Confirm the fee and click **Pay**.

Step 5 Select a payment method and confirm the payment. After you pay the order, the renewal is complete.

----End

Standard Renewal Date of Yearly/Monthly Resources

If you have multiple dedicated resource pools with different expiration dates, you can set the expiration dates to the same date for convenient management and renewal.

Figure 5-3 shows a user enabling a renewal date and renewing two resources for one month, each expiring at a different time.

Figure 5-3 Renewal date
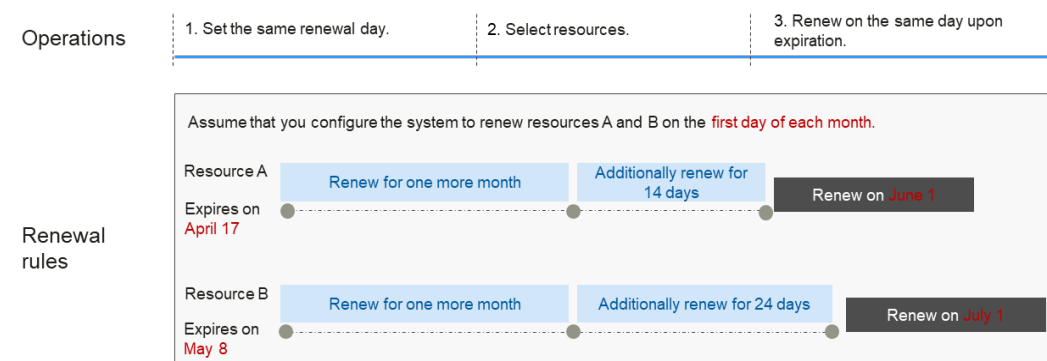
For details about renewal date rules, see **Setting a Renewal Date**.

5.3 Auto-Renewal

Auto-renewal can prevent ModelArts dedicate resource pools from being unavailable if you forget to manually renew them. The auto-renewal rules are as follows:

- The first auto-renewal date is based on when the dedicated resource pool expires and the billing cycle.

- The auto-renewal period of a dedicated resource pool depends on the subscription term. For example, if the subscription term is 3 months, your subscription is automatically renewed for three months every time before expiration.

- Auto-renewal can be enabled before a dedicated resource pool expires. The system attempts to automatically renew the dedicated resource pool at 03:00 on the seventh day before it expires. If the fee deduction fails, the system attempts to automatically renew the dedicated resource pool at 03:00 every day until the dedicated resource pool expires or the renewal is successful.

- After auto renewal is enabled, you can still renew the dedicated resource pool manually. After a manual renewal is completed, auto renewal is still valid, and the renewal charge will be deducted from your account seven days before the new expiry date.

- By default, the renewal charge is deducted from your account seven days before the new expiry date. You can change this auto-renewal payment date as required.

Learn about **Auto-Renewal Rules** before enabling auto-renewal.

Prerequisites

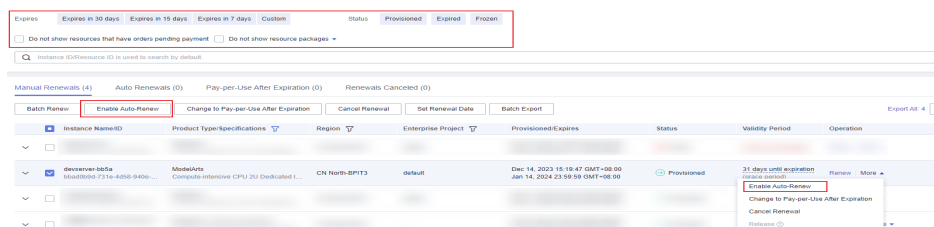The yearly/monthly dedicated resource pool is not expired.

Enabling Auto Renewal on the Renewals Page

Step 1 Go to the **Billing Center > Renewals** page.

Step 2 Customize search criteria.

- You can view the resources for which auto-renewal has been enabled on the **Auto Renewals** page.

- You can enable auto-renewal for resources on the **Manual Renewals**, **Pay-per-Use After Expiration**, and **Renewals Canceled** pages.

**Figure 5-4** Managing renewals

Step 3 Enable auto-renewal for yearly/monthly resources.

- Enabling auto-renewal for a single resource: Select the dedicated resource pool for which you want to enable auto-renewal and click **Enable Auto-Renew** in the **Operation** column.

- Enabling auto-renewal for multiple resources at a time: Select the dedicated resource pools for which you want to enable auto-renewal and click **Enable Auto-Renew** above the list.

Step 4 Select a renewal period, specify the auto-renewal times, and click **Pay**.

Figure 5-5 Enabling auto-renewal
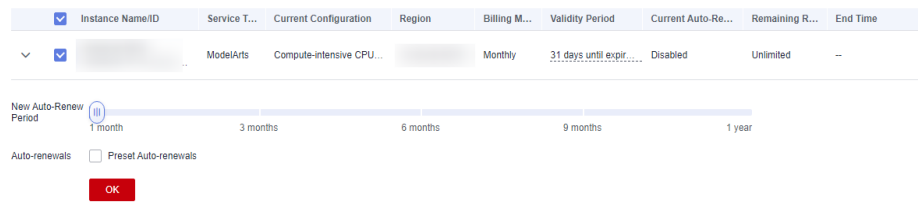
----**End**

Modifying Auto-Renewal on the Renewals Page

Step 1 Go to the **Billing Center > Renewals** page.

Step 2 Select the dedicated resource pool for which you want to modify auto-renewal and click **Modify Auto-Renew** in the **Operation** column.

Step 3 Set the renewal mode, select a renewal period, specify the desired auto-renewal times, and click **OK**.

----**End**

6 Bills

You can choose **Billing** in the Billing Center to view the bill of a resource and learn about the usage and billing information of the resource in a certain period.

Bill Reporting Period

After you pay for a yearly/monthly resource, a bill is reported to the billing system in real time for settlement.

The usage of pay-per-use resources is reported to the billing system at a fixed interval for settlement. Pay-per-use products can be settled by hour, day, or month based on the usage type. Pay-per-use ModelArts dedicated resource pools are billed by hour.

The fee deduction time of pay-per-use resources may be later than the settlement period. For example, if a dedicated resource pool billed by hour is deleted at 8:30, the fee generated from 8:00 to 9:00 is usually deducted at about 10:00. On the **Billing Center** > **Billing** > **Transactions and Detailed Bills** > **Transaction Bills** page, **Expenditure Time** indicates the time when a pay-per-use product is used.

Viewing ExeML Bills

When ExeML projects are running, different bills are generated for training jobs and service deployment.

- Obtain bills generated by training jobs by referring to **Viewing Training Job Bills**.
- Obtain bills generated by service deployment by referring to **Viewing Real-Time Service Bills**.

Viewing Notebook Bills

Step 1 Log in to the ModelArts management console and choose **DevEnviron** > **Notebook**. In the list, click to copy the instance ID.

Step 2 Go to the **Billing > Expenditure Details** page.

Step 3 In the bill list, select **Resource ID** from the filter options, paste the resource ID copied in **Step 1**, and click to obtain the bill.

By default, expenditure details are displayed by usage and the statistical period is a billing cycle. You can also set other statistical dimensions and periods.

----End

Viewing Training Job Bills

To obtain the resource bills for a ModelArts training job, obtain the job ID first. However, the job ID displayed on the ModelArts management console is different from the resource ID reported in the bill. You must learn how to obtain the job ID and the mapping between the job ID and the resource ID reported in the bill. The procedure is as follows:

Step 1 Obtain the resource ID.

Log in to the ModelArts management console. In the navigation pane, choose **Training Management** > **Training Jobs**. Copy the ID of the job whose bills you want to view.

Step 2 Combine the resource ID reported in the bill based on the obtained job ID.

Combination rule: **training_***Job ID*

For example, if the job ID is **32de36ea-26c0-4876-ae48-fdbbb03cd455**, the resource ID reported to the bill is **training_32de36ea-26c0-4876-ae48-fdbbb03cd455**.

Step 3 View the resource bills in the Billing Center based on the resource ID.

1. Go to the **Billing > Expenditure Details** page.
2. In the bill list, select **Resource ID** from the filter options, enter the resource ID combined in **Step 2**, and click Q to obtain the bill.

   By default, expenditure details are displayed by usage and the statistical period is a billing cycle. You can also set other statistical dimensions and periods.

----End

Viewing Real-Time Service Bills

To obtain resource bills of ModelArts real-time services, you must obtain the resource name first. However, the real-time service name displayed on the ModelArts management console is different from the resource name reported in the bill. You must learn how to obtain the resource name and the mapping between the service name and the resource name reported in the bill. The procedure is as follows:

Step 1 Obtain the resource name.

1. Log in to the ModelArts management console and choose **Service Deployment** > **Real-Time Services**. From there, copy the instance name.
2. Click the service name to go to the service details page. On the **Configuration Updates** tab page, click the AI application name to go to the AI application details page.
3. In the **Basic Information** area, copy the ID of the AI application.

Figure 6-1 Obtaining the AI application ID

Step 2 Combine the resource name reported in the bill based on the obtained service name and application ID.

Combination rule: *Real-time service name-AI application ID*

For example, if the real-time service name is **service_predictor_name** and the AI application ID is **b9c97269-c9b6-4c77-b31e-58b0cb4e02e3**, the resource name reported to the bill is **service_predictor_name-b9c97269-c9b6-4c77-b31e-58b0cb4e02e3**.

Step 3 View the resource bill in the Billing Center based on the resource name.

1. Go to the **Billing > Expenditure Details** page.
2. In the bill list, select **Resource Name** from the filter options, enter the resource name combined in **Step 2**, and click 🔍 to obtain the bill.

   By default, expenditure details are displayed by usage and the statistical period is a billing cycle. You can also set other statistical dimensions and periods.

----End

Viewing Dedicated Resource Pool Bills

Step 1 On the ModelArts management console, choose **Dedicated Resource Pools** > **Elastic Cluster**. From there, click on the resource name to go to the resource details page.

Step 2 On the resource details page, click the **Specifications** tab and copy the **Metering ID**.

Step 3 Go to the **Billing > Expenditure Details** page.

Step 4 In the bill list, select **Resource ID** from the filter options, paste the metering ID copied in **Step 2**, and click 🔍 to obtain the bill.

By default, expenditure details are displayed by usage and the statistical period is a billing cycle. You can also set other statistical dimensions and periods.

----End

7 About Arrears

When you use cloud services, if the available credit of your account is less than the bill to be settled, your account is in arrears. Recharge your account promptly to avoid cloud service disruptions due to arrears.

Arrears Reason

- When the used resources exceed the package quota or the resource package is not applicable to the bucket, you will be charged for the used resources on a pay-per-use basis. If your account balance is insufficient, outstanding bills are generated. For more information, see **How Do I View the ModelArts Jobs Being Billed?** Select applicable resource packages or top up your account.

- Your account balance is insufficient to pay for the used resources on the pay-per-use basis, and you do not have any resource packages either.

Impact of Arrears

- Yearly/Monthly

  For yearly/monthly dedicated resource pools, you have prepaid for the resources. Therefore, you can continue to use the existing yearly/monthly dedicated resource pools even if your account is in arrears. However, you cannot perform operations that require fees, such as storing data to OBS or EVS and renewing orders.

- Pay-per-use

  Your account may fall into arrears due to automatic fee deduction for pay-per-use resources. In this case, relevant cloud services will not be stopped immediately. There will be a grace period for pay-per-use resources. You must pay fees for pay-per-use resources incurred during the grace period, which are counted as outstanding amount in the Billing Center. Huawei Cloud will automatically deduct the fees when you top up.

  If you do not pay the arrears within the grace period, the resources enter the retention period and the resource status changes to **Frozen**. You cannot perform any operation on the pay-per-use resources in the retention period.

  If you do not pay the arrears within the retention period, the compute resources (vCPUs and GPUs), EVS disks, and OBS resources will be released and data cannot be restored.

Figure 7-1 Periods of pay-per-use resources
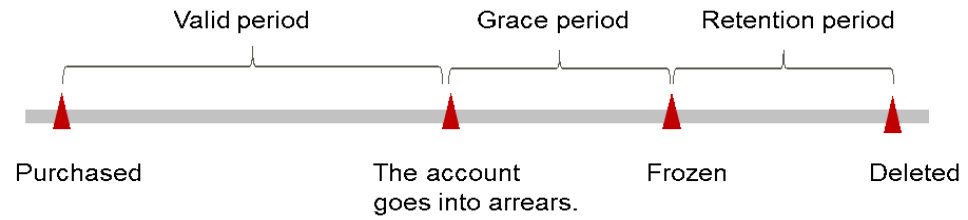
> **NOTE**
>
> Both the grace period and retention period are 15 days.

Avoiding and Handling Arrears

You need to top up your account in arrears.

If you do not use ModelArts anymore, confirm that all billed instances have been stopped or deleted on the **Dashboard** page of the ModelArts management console. Additionally, clear the data stored in EVS disks for running notebook instances and data stored in OBS to stop billing.

You can enable balance alerts on the **Overview** page of the Billing Center. If the sum of your account balance, cash coupons, and flexi-purchase coupons goes below the threshold, the system automatically sends an SMS notification.

8 Stopping Billing

You need to delete or stop the billing items in ModelArts if they are no longer needed. The following are the billing items that need to be stopped or deleted involved in ExeML, development environments, model training, real-time services, and dedicated resource pools:

- ExeML: Stop training jobs and real-time services created for running ExeML jobs. Delete the OBS bucket and data stored in OBS.

- Development environment (Notebook): Delete the notebook instance. Delete the OBS bucket and data stored in OBS.

    📖 **NOTE**

    > EVS is still billed even if the notebook instance is stopped. To stop billing, delete the notebook instance.

- Training job: Stop the training job. Delete the OBS bucket and data stored in OBS.

- Real-time service: Stop the real-time service. Delete the OBS bucket and data stored in OBS.

- Dedicated resource pool: Delete the dedicated resource pool. Pay-per-use dedicated resource pools are billed immediately after they are created and are continuously billed during use. Delete the dedicated resource pools that are no longer used. Yearly/Monthly dedicated resource pools are automatically stopped after they expire.

    📖 **NOTE**

    > OBS resources for storing data are continuously billed. To stop billing, delete the data stored in OBS.

Yearly/Monthly Resources

For yearly/monthly resources, you need to pay for them at a time when purchasing them. The services will be automatically stopped after they expire.

- If a yearly/monthly resource is no longer needed within the billing cycle, you can unsubscribe from it. A certain amount of money will be returned to your account based on whether the resource meets the requirements of five-day unconditional unsubscription and whether cash coupons and discount coupons are used.

- If you have enabled auto-renewal, disable it before the auto-renewal deduction date (seven days before the expiration date by default) to avoid further fees.

Pay-per-Use Resources

If pay-per-use billing mode resources are no longer used, delete them to stop billing.

Searching for Cloud Service Resources Based on Bill Information and Stopping Billing

To ensure that all billed resources can be found and deleted, obtain the IDs of these resources from the bills, find the specific resources with the IDs, and delete them.

You can use the resource ID to search for notebook development environments. The procedure is as follows.

NOTICE

- If a notebook instance is created using a public resource pool, the billing will stop after you delete the instance.
- If a notebook instance is created using a dedicated resource pool, you must delete the dedicated resource pool to stop the billing.
- If you use storage resources, such as OBS and SFS, for running a notebook instance, you must delete the data stored in the storage services accordingly.
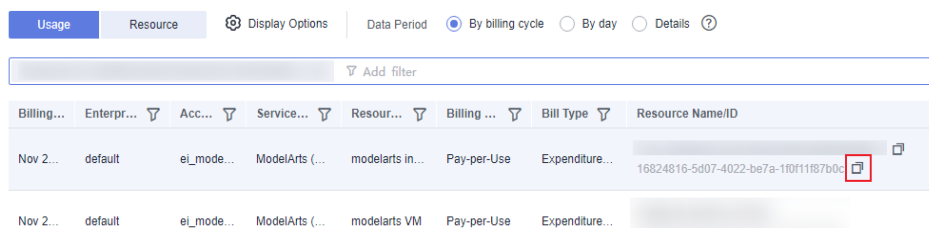

Step 1 Go to the **Billing > Expenditure Details** page.

Step 2 In the bill list, click ⧉ to copy the resource ID.
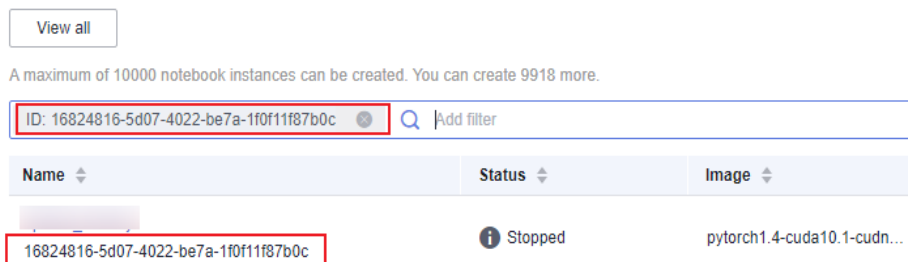
Figure 8-1 Copying a resource ID

Step 3 Log in to the ModelArts management console and choose **DevEnviron** > **Notebook**.

Step 4 To find the resource, select **ID** from the filter options above the notebook list and paste the resource ID copied in **Step 2** in the search box.

Figure 8-2 Searching for a resource

Step 5 Choose **More** > **Delete** in the **Operation** column to delete the notebook instance.

📖 **NOTE**

> You are billed one hour after the resource usage is calculated, so a bill may still be generated after the pay-per-use resource is deleted. For example, suppose you delete some training job resources that are billed by hour at 08:30. Then, you will likely get billed for the usage from 08:00 to 09:00 around 10:00.

----End

**You can use the resource name to search for training jobs. The procedure is as follows.**

NOTICE

- If a training job is created using a public resource pool, the billing will stop after you stop the job.
- If a training job is created using a dedicated resource pool, you must delete the dedicated resource pool to stop the billing.
- If you use storage resources, such as OBS, for running a training job, you must delete the data stored in the storage services accordingly.

Step 1 Go to the **Billing > Expenditure Details** page.

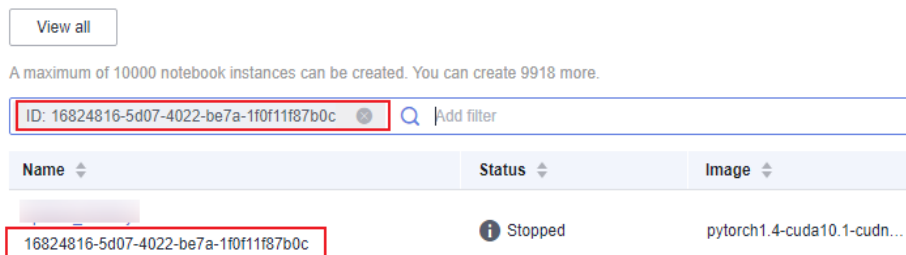**Step 2** In the bill list, click ⧉ to copy the resource name.

Figure 8-3 Copying a resource name
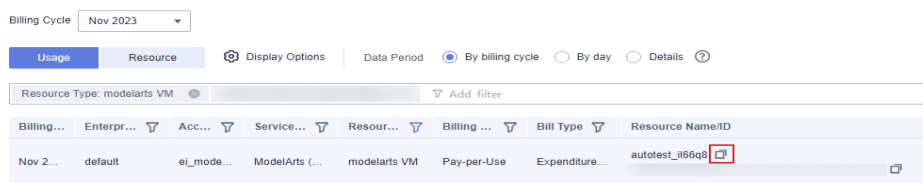
Step 3 Log in to the ModelArts management console and choose **Training Management > Training Jobs**.

Step 4 To find the resource, select **Name** from the filter options above the training job list and paste the resource name copied in **Step 2** in the search box.

Figure 8-4 Searching for a resource

View all

A maximum of 10000 notebook instances can be created. You can create 9918 more.

ID: 16824816-5d07-4022-be7a-1f0f11f87b0c ⊗   🔍 Add filter

Name ⇕	Status ⇕	Image ⇕
16824816-5d07-4022-be7a-1f0f11f87b0c	ⓘ Stopped	pytorch1.4-cuda10.1-cudn...

Step 5 Click **Stop** in the **Operation** column to stop the training job.

> 📖 **NOTE**
>
> You are billed one hour after the resource usage is calculated, so a bill may still be generated after the pay-per-use resource is deleted. For example, suppose you delete some training job resources that are billed by hour at 08:30. Then, you will likely get billed for the usage from 08:00 to 09:00 around 10:00.

----End

For real-time and batch services, the resource ID is different from that reported in the bills. To obtain the bill and stop the billing, do as follows:

> **NOTICE**
>
> - If a real-time or batch service is created using a public resource pool, the billing will stop after you stop the service.
> - If a real-time or batch service is created using a dedicated resource pool, you must delete the dedicated resource pool to stop the billing.
> - If you use storage resources, such as OBS, for running a real-time job, you must delete the data stored in the storage services accordingly.

Step 1 Go to the **Billing > Expenditure Details** page.

Step 2 In the bill list, view the resource ID. The resource ID combination rule is *Real-time service name*-*AI application ID*. This way, you can tell which real-time service the resource ID belongs to.

> 📖 **NOTE**
>
> If the resource ID in a bill contains multiple hyphens (-), you can copy the content before the first hyphen (-).

Step 3 Search for the resource on the cloud service console based on the resource name and delete the resource.

1. On the ModelArts management console and choose **Service Deployment** > **Real-Time Services**.

2. To find the resource, select **Name** from the filter options above the service list and paste the real-time service name copied in **Step 2** in the search box.

3. Choose **More** > **Stop** in the **Operation** column to stop the real-time service.

> **NOTE**
>
> You are billed one hour after the resource usage is calculated, so a bill may still be generated after the pay-per-use resource is deleted. For example, if compute resources billed by hour are deleted at 08:30, the fees generated from 08:00 to 09:00 are usually deducted at about 10:00.
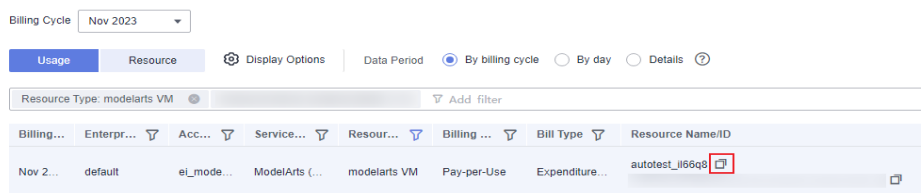
----End

For a resource pool, you can search for a resource by resource name. The procedure is as follows:

Step 1 Go to the **Billing > Expenditure Details** page.

Step 2 In the bill list, click ⧉ to copy the resource name.

Figure 8-5 Copying a resource name

Step 3 Log in to the ModelArts management console. In the navigation pane, choose **Dedicated Resource Pools** > **Elastic Cluster**.

Step 4 Paste the resource name copied in **Step 2** in the search box and click 🔍 to search for the resource.

Step 5 Choose **More** > **Delete** in the **Operation** column to delete the resource pool. Ensure that the resource pool has been deleted from the list.

> **NOTE**
>
> After a pay-per-use resource is deleted, the bill information may still exist because the system usually deducts the fee of the previous settlement period within one hour after the resource is used. For example, if dedicated resource pool resources billed by hour are deleted at 08:30, the fees generated from 08:00 to 09:00 are usually deducted at about 10:00.

----End

9 Cost Management

Cost Composition

ModelArts provides the AI tool chain and AI compute power. The cost consists of the resource cost and O&M cost of AI compute power.

Cost Allocation

ModelArts supports enterprise project management. In this way, costs are managed by project under the same account.

Cost Analysis

Huawei Cloud bills are used to analyze the costs of your account.

Cost Optimization

When you have long-term needs for resources, use the yearly/monthly billing mode to enjoy more discounts. When you temporarily use resources, use the pay-per-use billing mode to avoid waste.

10 Billing FAQs

10.1 How Do I View the ModelArts Jobs Being Billed?

Log in to the ModelArts management console. In the navigation pane on the left, click **Dashboard** and view the jobs that are being billed. Go to the management page and stop them based on site requirements. For example, if a notebook instance is being billed, choose **DevEnviron** > **Notebook**, and stop the running notebook instance.

Figure 10-1 Viewing jobs that are being billed

The following items will be billed when you use ModelArts:

- ExeML: ExeML is billed when it is running. After using ExeML, stop it and the training jobs and services created for running ExeML. Additionally, clear the data stored in OBS.

- Notebook instances:

  - Running notebook instances are billed. To stop billing a notebook instance, stop or delete it. If EVS is used for storage, clear the data stored in EVS.

  - You will be billed for a paid flavor when you try CodeLab. To stop billing the flavor, stop the notebook instance on the JupyterLab page.

- Training jobs: Running training job are billed. To stop billing a training job, stop it. Additionally, clear the data stored in OBS.

- Deployed services: If a model is deployed as a real-time service, the service will be billed. To stop the billing, stop the deployed service. Additionally, clear the data stored in OBS.

- Dedicated resource pools: If you have purchased a dedicated resource pool in ModelArts for AI development and use this resource pool for running ExeML jobs, notebook instances, training jobs, and deployed services, the compute

resources used in these operations will be billed via the dedicated resource pool. A pay-per-use dedicated resource pool is billed continuously since it has been created, so delete it if you do not use it.

> ⚠️ **CAUTION**
>
> In addition to the billing items displayed on the **Dashboard** page of ModelArts, OBS and EVS storage will be separately billed if they are used.
>
> - To stop billing the OBS resources, go to the OBS management console and clear the data in OBS.
> - To stop billing the EVS storage, go to the ModelArts management console and delete the notebook instances with EVS storage. To stop billing the EVS resources, go to the EVS management console and clear the data in EVS.

10.2 How Do I View ModelArts Expenditure Details?

In **Billing Center**, the expenditure details of ModelArts are displayed on an hourly basis. You can switch to the **Expenditure Details** page to view the fees consumed for each job.

Procedure:

1. In **Billing Center**, choose **Billing** > **Expenditure Items**. On the **Expenditure Items** page, click the order or transaction number of a record in the list to view its expenditure details.
2. On the **Expenditure Details** page, the resource usage and fees in the order are displayed. Select display options and data period to view expenditure details.

Figure 10-2 Expenditure details
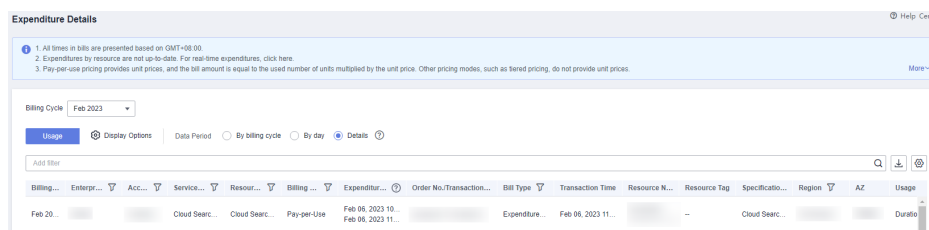
10.3 How Do I Stop Billing If I Do Not Use ModelArts?

The cost of AI development in ModelArts mainly includes the storage fee and resource fee. If ModelArts is no longer used, stop or delete the services running in ModelArts and delete the data stored in OBS and EVS.

Clearing Storage Data

ModelArts data is stored in OBS. To stop storage billing, switch to the OBS console and delete the data and directories.

Clearing Resources

To stop resource billing, check running jobs in ModelArts, and stop or delete the jobs.

Procedure

Log in to the ModelArts console. In the navigation pane, choose **Dashboard**. On the **Dashboard** area, view the jobs that are being billed. Then, stop the jobs as needed.

Figure 10-3 Viewing jobs that are being billed

- In the navigation pane of the ModelArts console, choose **ExeML** and check whether there are any running projects. If so, click **Delete** in the **Operation** column. Then, the billing will be stopped accordingly.

- In the navigation pane of the ModelArts console, choose **DevEnviron** > **Notebook** and check whether there are running notebook instances. If so, click **Stop** in the **Operation** column. Then, the billing will be stopped accordingly. Check whether there are notebook instances using EVS storage. If so, stop and delete the notebook instances. Then, the EVS billing will be stopped accordingly.

- In the navigation pane of the ModelArts console, choose **Training Management** > **Training Jobs**, and check whether there are running jobs. If so, click **Stop** in the **Operation** column. Then, the billing will be stopped accordingly.

- In the navigation pane of the ModelArts console, choose **Training Management** > **Training Jobs**, click the **Visualization Jobs** tab, and check whether there are running jobs. If so, click **Stop** in the **Operation** column. Then, the billing will be stopped accordingly.

- In the navigation pane of the ModelArts console, choose **Service Deployment** > **Real-Time Services** and check whether there are running jobs. If so, click **Stop** in the **Operation** column. Then, the billing will be stopped accordingly.

- In the navigation pane of the ModelArts console, choose **Service Deployment** > **Batch Services** and check whether there are running jobs. If so, click **Stop** in the **Operation** column. Then, the billing will be stopped accordingly.

10.4 Will I Be Charged for Uploading Datasets to ModelArts?

You are not charged for dataset management and labeling in ModelArts. However, datasets are stored in OBS, and you will be billed for the storage on OBS. For details, see **OBS pricing details** and create OBS buckets to store data used by ModelArts.

10.5 What Should I Do to Avoid Unnecessary Billing After I Label Datasets and Exit?

Labeling datasets is free of charge. However, OBS bills you for the storage space used for storing the datasets. You are advised to go to the OBS management console and delete the stored data and OBS buckets to stop billing.

10.6 How Do I Stop Billing for a ModelArts ExeML Project?

- For ExeML jobs created using a **public resource pool**:
  - Log in to the ModelArts management console. In the ExeML job list, delete the target ExeML job. In the training job list, stop the training jobs created for running the ExeML job. In the real-time service list, stop the services created for running the ExeML job. After these operations are complete, the billing is stopped.
  - Log in to the OBS management console, access your OBS bucket, and delete the data stored in the OBS bucket. After these operations are complete, the billing is stopped.
- For ExeML jobs created using a **dedicated resource pool**:
  - Log in to the ModelArts management console. In the ExeML job list, delete the target ExeML job. In the training job list, stop the training jobs created for running the ExeML job. In the real-time service list, stop the services created for running the ExeML job. In the resource pool list, delete the dedicated resource pool that is running the target ExeML job. After these operations are complete, the billing is stopped.
  - Log in to the OBS management console, access your OBS bucket, and delete the data stored in the OBS bucket. After these operations are complete, the billing is stopped.

10.7 How Are Training Jobs Billed?

- If you use a public resource pool, you are charged based on the selected flavor, number of nodes, and running duration. The billing rule is as follows: Unit price of the flavor x Number of nodes x Running duration (accurate to seconds).
- If you use dedicated resource pool, training jobs are not billed separately. You are charged for the dedicated resource pool.

10.8 Why Does Billing Continue After All Projects Are Deleted?

Even if ExeML projects, notebook instances, training jobs, or services of ModelArts are stopped, and there is no charging item is displayed on the **Dashboard** page, the account may still be billed for the OBS storage being used.

The possible causes are as follows:

1. You have uploaded data to OBS for storage when using ModelArts, and the OBS storage is billed. In this case, go to the OBS management console and delete the data, folders, and OBS buckets that are no longer needed.
2. You have selected EVS storage when creating a notebook instance, and the storage is separately billed even after the notebook instance is stopped. In this case, delete the notebook instance.
3. You have switched the flavor to a charging one when experiencing CodeLab.

    In this case, go to the CodeLab page and click ⏻ in the upper right corner to stop the notebook instance.