

Alibaba Cloud Cloud Monitor

User Guide

Issue: 20190221

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 due to product version upgrades, adjustments, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and the updated versions of this document will be occasionally released through Alibaba Cloud-authorized channels. You shall 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 the document in the context that Alibaba Cloud products and services are provided on an "as is", "with all faults" and "as available" basis. 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 bear any liability for any errors or financial losses incurred by any organizations, companies, or individuals arising from their download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, bear responsibility for any indirect, consequential, exemplary, incidental, 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 content of the Alibaba Cloud website, including but not limited to works, products, images, archives, information, materials, website architecture, website graphic layout, and webpage design, 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 the Alibaba Cloud website, product programs, or content 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 contact Alibaba Cloud directly if you discover any errors in this document.

Generic conventions

Table -1: Style conventions

Style	Description	Example
	This warning information 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.
	This warning information 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 restore business.
	This indicates warning information, supplementary instructions, and other content that the user must understand.	 Notice: Take the necessary precautions to save exported data containing sensitive information.
	This indicates supplemental instructions, best practices, tips, and other content that is good to know for the user.	 Note: You can use Ctrl + A to select all files.
>	Multi-level menu cascade.	Settings > Network > Set network type
Bold	It is used for buttons, menus, page names, and other UI elements.	Click OK.
Courier font	It is used for commands.	Run the <code>cd /d C:/windows</code> command to enter the Windows system folder.
<i>Italics</i>	It is used for parameters and variables.	<code>bae log list --instanceid Instance_ID</code>
[] or [a b]	It indicates that it is an optional value, and only one item can be selected.	<code>ipconfig [-all -t]</code>

Style	Description	Example
<code>{}</code> or <code>{a b}</code>	It indicates that it is a required value, and only one item can be selected.	<code>switch {stand slave}</code>

Contents

Legal disclaimer.....	I
Generic conventions.....	I
1 Dashboard.....	1
1.1 Dashboard overview.....	1
1.2 Manage dashboards.....	2
1.3 Add charts.....	5
2 Host monitoring.....	11
2.1 Host monitoring overview.....	11
2.2 Process monitoring.....	12
2.3 Query GPU monitoring data.....	16
2.4 Host monitoring metrics.....	21
2.5 Alarm service.....	28
2.6 CloudMonitor Java agent introduction.....	29
2.7 Install CloudMonitor Java agent.....	31
2.8 Introduction to the CloudMonitor GoLang agent.....	41
2.9 Install CloudMonitor GoLang agent.....	43
2.10 Agent release notes.....	54
3 Site Monitoring.....	56
4 Alarm service.....	57
4.1 Alarm service overview.....	57
4.2 Manage alarm rules.....	58
4.3 Manage alarm contacts and alarm contact groups.....	60
4.4 Create an alarm callback.....	62
4.5 Use alarm templates.....	64
5 Availability monitoring.....	67
5.1 Manage availability monitoring.....	67
5.2 Local service availability monitoring.....	70
5.3 Status codes.....	72
6 Cloud service monitoring.....	73
6.1 RDS monitoring.....	73
6.2 SLB monitoring.....	77
6.3 OSS monitoring.....	89
6.4 CDN monitoring.....	90
6.5 ApsaraDB for Memcache.....	94
6.6 Global acceleration monitoring.....	97
6.7 High performance time series database hitsdb.....	99
6.8 VPN gateway.....	101
6.9 Elasticsearch monitoring.....	103
6.10 Express Connect monitoring.....	105

6.11 StreamCompute.....	108
6.12 ApsaraDB for HybridDB.....	111
6.13 NAT gateway.....	112
6.14 Open Ad.....	114
6.15 ApsaraDB for PetaData.....	116
7 RAM for CloudMonitor.....	119
8 Application groups.....	122
8.1 Application group overview.....	122
8.2 Create application groups.....	122
8.3 Manage alarm rules.....	125
8.4 Check application group details.....	128
9 Event monitoring.....	133
9.1 Event monitoring overview.....	133
9.2 Report event data.....	139
9.3 Event monitoring best practices.....	147
9.4 Cloud product system event monitoring.....	157
9.5 Use system event alarms.....	168
10 Custom monitoring.....	172
10.1 Custom monitoring overview.....	172
10.2 Report monitoring data.....	175
10.3 Configure a dashboard.....	188

1 Dashboard

1.1 Dashboard overview

The CloudMonitor dashboard provides you with a real time metric visualization solution for a comprehensive overview of your applications and services, enabling you to quickly troubleshoot problems and monitor resource usage.

Display metric trends for multiple instances

The dashboard provides detailed metrics and trends for multiple instances. For example, you can view the metrics of all the ECS instances on which your application is deployed all on one metric chart. This can help you see trends across multiple instances all in one area. Similarly, you can also view the CPU usage of multiple ECS instances over time in one chart.

Display multiple metrics per instance

With dashboards, you can also view several metrics of an ECS instance, such as CPU usage, memory usage, and disk usage all displayed on one metric chart. This visualization solution can help you find exceptions and monitor resource usage efficiently.

Display and sort instance resource usage

Instances can be sorted based on resource usage levels, allowing you to quickly gain insight into resource usage per instance and how usage levels differ between instances. With this information, you can make informed decisions and avoid unnecessary costs.

Display metrics distribution of multiple instances

The CPU usage distribution of an ECS instance group can be visualized with a heat map, allowing you to quickly and accurately discover the real time usage levels of different machines and compare them with each other. These heat maps are not only powerful visualization tools but are also interactive. You can click any one of the color blocks on the heat map to view the metrics and trends of the corresponding machine for a specified period of time.

Display aggregated metrics of multiple instances

With dashboards, you can view the average aggregation value of a particular metric , such as CPU usage of multiple ECS instances, all in one chart. With this capability , you quickly estimate overall CPU usage capacity and check whether the resource usage of different instances is balanced.

Provides full-screen visualization solution

The dashboard supports a full-screen mode that automatically refreshes. In this mode , you can easily add several application and product metrics to the full-screen display , allowing you to have a quick visual overview of all monitored data.

1.2 Manage dashboards

You can easily view, create, and delete dashboards. The procedure for these actions is as follows.

View a dashboard

You can view a dashboard to view and monitor metrics from several different products and instances all within one area.



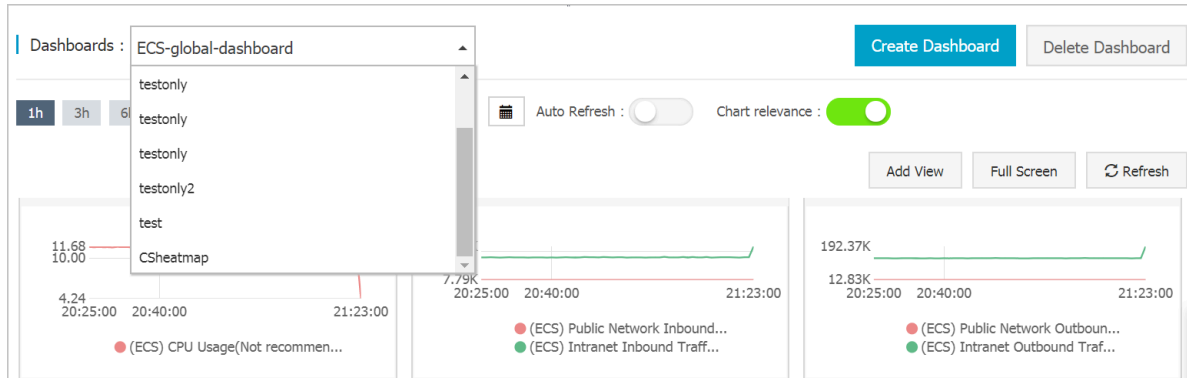
Note:

- CloudMonitor automatically initializes an ECS dashboard and displays ECS metrics.
- CloudMonitor refreshes data measured in one-hour, three-hour, and six-hour periods automatically. However, data measured for more than six hours cannot be refreshed automatically.

Procedure

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, choose Dashboard > Custom Dashboard.

- By default, ECS-global-dashboard is displayed. You can select another dashboard from the drop-down list.



- To view the dashboard in full screen, click Full Screen in the upper-right corner of the page.
- Select a time range. Click the time range button at the top of the page. From there, you can quickly select the time range shown in the charts of the dashboard. The time range you select apply to all the charts on the dashboard.
- Automatic refresh. After you turn on the Auto Refresh switch, whenever you select a query time span of 1 hour, 3 hours, or 6 hours, automatic refresh is performed every minute.
- The units of the metrics measured are displayed in parentheses for the chart name.
- When you rest the pointer over some point on a chart, values at that time point are displayed across all charts.

Create a dashboard

You can create a dashboard and customize the charts for when your business operations grow complex and the default ECS dashboard does not meet your monitoring requirements.



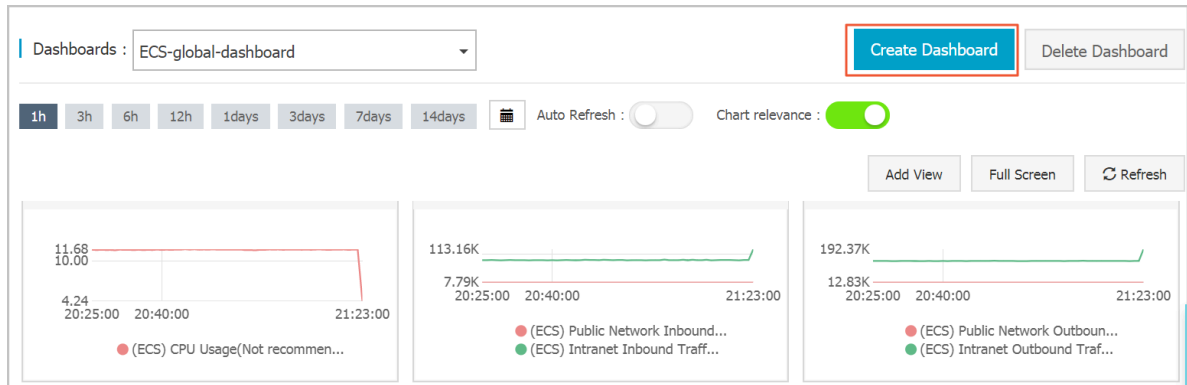
Note:

Up to 20 charts can be created on one dashboard.

Procedure

- Log on to the [CloudMonitor console](#).
- In the left-side navigation pane, choose Dashboard > Custom Dashboard.

3. In the upper-right corner of the page, click Create Dashboard.



4. Enter the name of the dashboard.

5. Click Create. The page is automatically redirected to the new dashboard page where you can add various metric charts as needed.

6. When you rest the pointer over the dashboard name, the Edit option appears on the right hand side. To modify the dashboard name, click Edit.

Delete a dashboard

You can delete a dashboard if you do not need it given changes in your business operations.



Notice:

When you delete a dashboard, all charts that are added to the dashboards are also deleted.

Procedure

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, choose Dashboard > Custom Dashboard.
3. Select the target dashboard from the Dashboards drop-down list.
4. In the upper-right corner of the page, click Delete Dashboard to delete the dashboard.

1.3 Add charts

This topic describes several types of charts common in the CloudMonitor dashboard and how to add a chart.

Scenarios

By default, CloudMonitor creates an initialized ECS dashboard. You can add more charts and tables to the dashboard to view even more data related to your ECS instances.

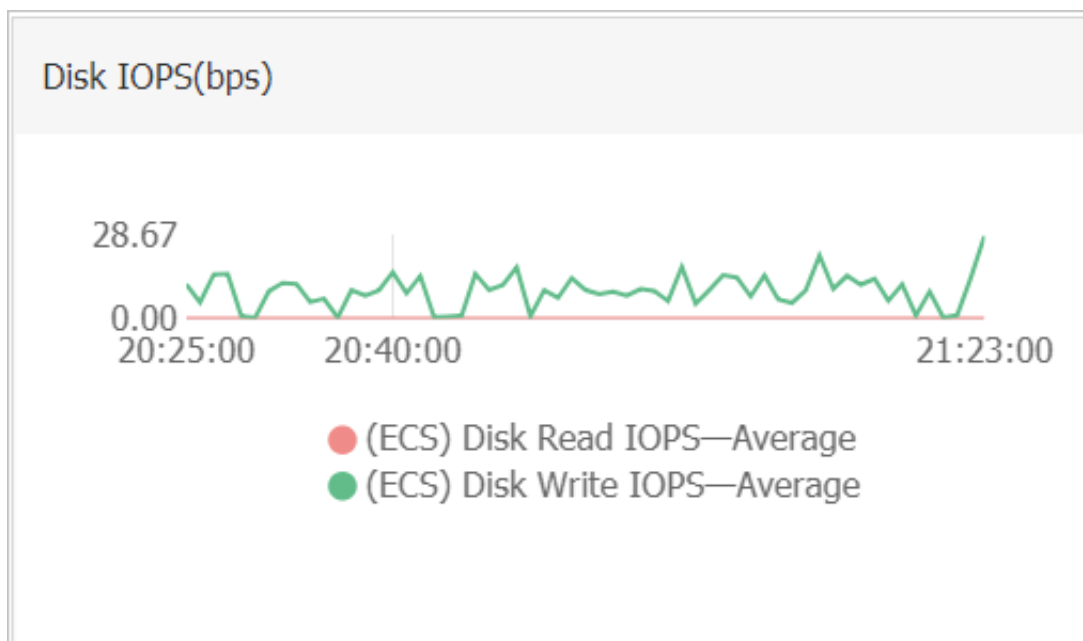
In the case that the ECS dashboard does not meet your monitoring needs, we recommend that you create an additional dashboard to which you can add charts to display custom monitoring data.

Before you begin

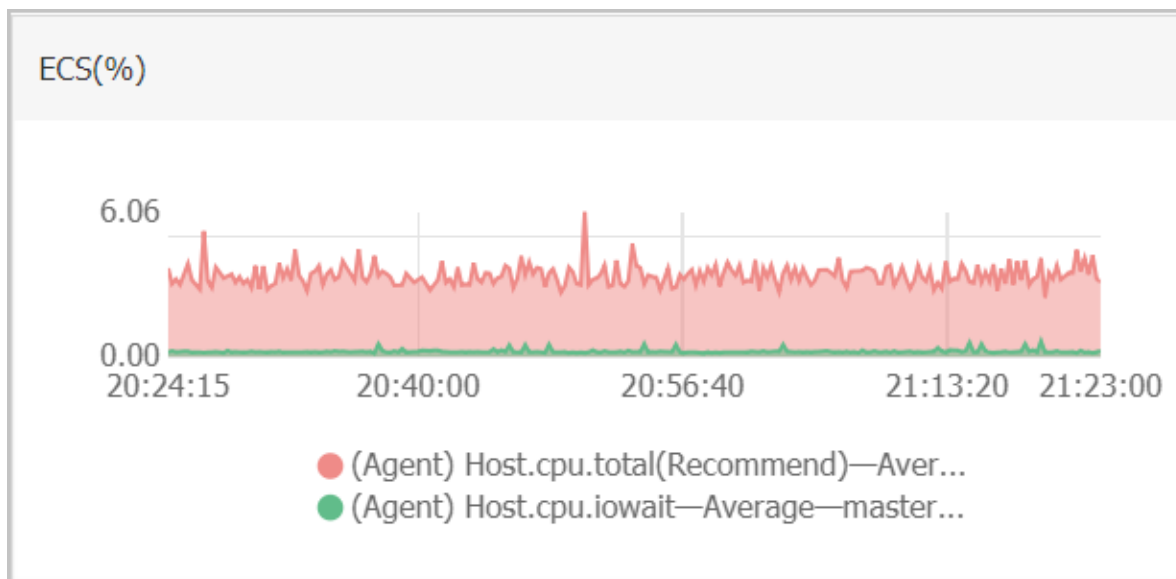
Before you can add a chart, you need to create a dashboard.

Chart types

- **Line chart:** Displays monitoring data on a basis of time series. Multiple metrics can be added.



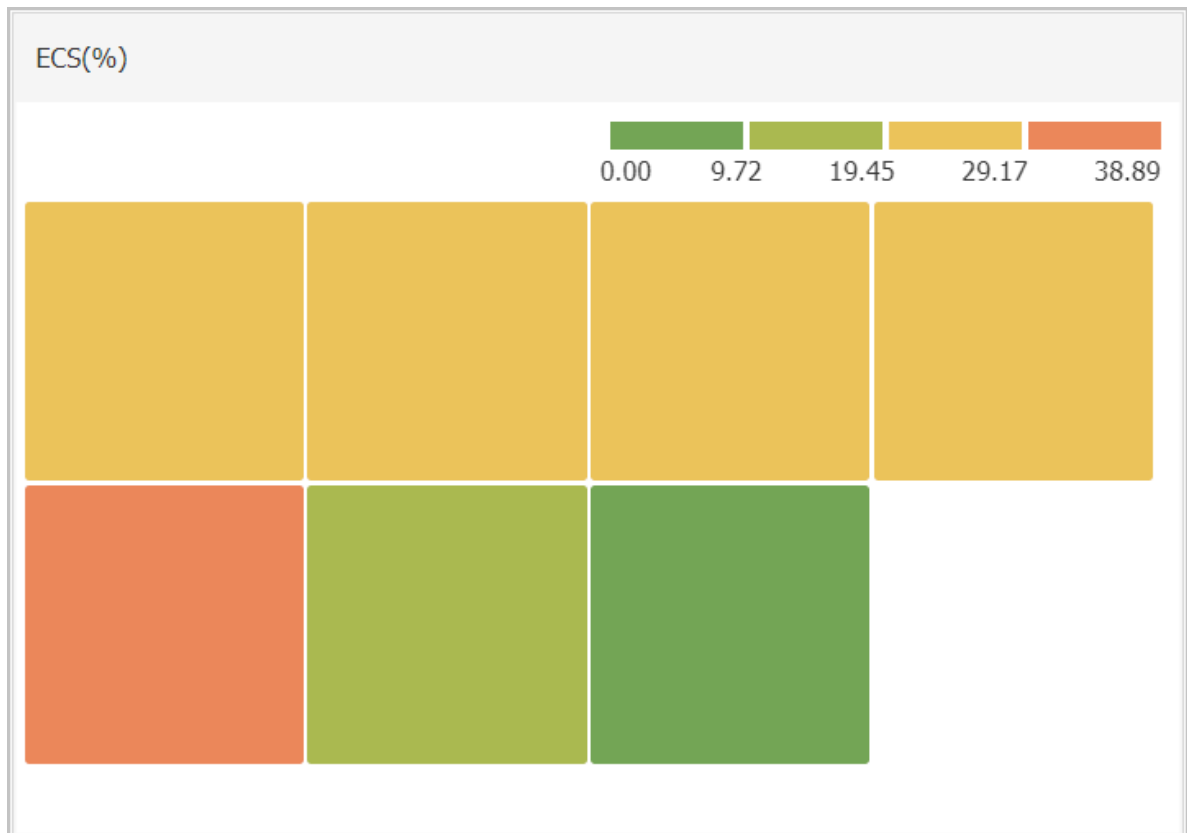
- **Area chart:** Displays monitoring data on a basis of time series. Multiple metrics can be added.



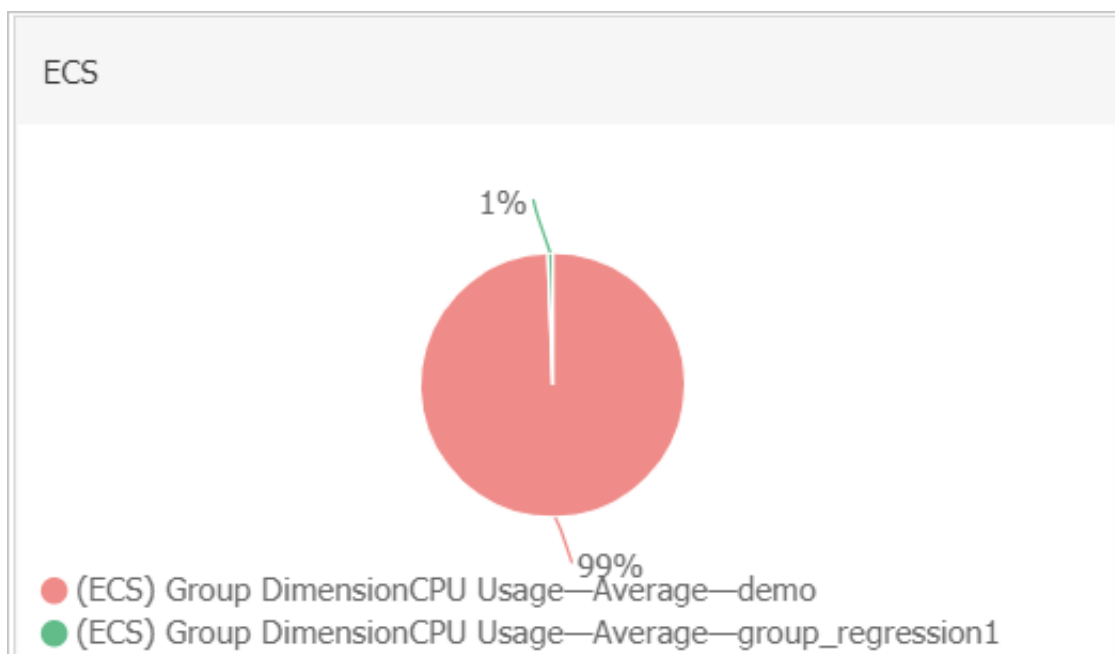
- **Table:** Displays real-time metric data in descending order. Each table displays up to 1,000 data records, which are either the first 1,000 records or the last 1,000 records. Only one metric can be added.

ECS(%)		
Time	Dimensions	Maximum Value
2018-12-06 21:25:00	ESS-asg-yinna_test	100
2018-12-06 21:20:00	node-0003-k8s-for-cs-c9ebd45a41dd645a498a5c06af2b88c53	55.56
2018-12-06 21:25:00	master-02-k8s-for-cs-c9ebd45a41dd645a498a5c06af2b88c53	38.89
2018-12-06 21:25:00	master-03-k8s-for-cs-c9ebd45a41dd645a498a5c06af2b88c53	38.1
2018-12-06 21:00:00	master-01-k8s-for-cs-c9ebd45a41dd645a498a5c06af2b88c53	37.5
2018-12-06 21:00:00	node-0001-k8s-for-cs-c9ebd45a41dd645a498a5c06af2b88c53	35.29
2018-12-06 21:20:00	node-0002-k8s-for-cs-c9ebd45a41dd645a498a5c06af2b88c53	29.41

- **Heat map:** Displays real-time metric data. Heat maps show the distribution and comparison of real-time data of a specific metric for multiple instances. Only one metric can be added.



- **Pie chart:** Displays real-time metric data and can be used for data comparisons. Only one metric can be added.



Add a chart



Note:

- The default ECS dashboard provides the following seven charts: CPU Usage, Network Inbound Bandwidth, Network Outbound Bandwidth, Disk BPS, Disk IOPS, Network Inbound Traffic, and Network Outbound Traffic.
- Up to 20 charts can be added in a dashboard.
- Each line chart can display up to 10 lines.
- Each area chart can display up to 10 areas.
- Each table can display up to 1,000 sorted data records.
- A heat map can display up to 1,000 color blocks.

Procedure

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, choose Dashboard > Custom Dashboard.

3. In the upper-right corner of the displayed page, click Add View.

Add View

1 Chart Type

Line Area Table Heat Map Pie Chart

2 Select Metrics

Dashboards Log Monitoring Custom

ECS ECS Heat Map Gradient Range : 0 auto

No Data

Metrics : (Agent) Host.cpu.total(Recommend) Maximum Value

Resource : q20/i-bp140l3jmql5sfmusa8

+AddMetrics

Save Cancel

4. Select a chart type.

5. Choose from Dashboards, Log Monitoring, and Custom tab pages. In this example, click the Dashboards tab.

6. Select the target Alibaba Cloud product and enter a name for the chart.

7. Select the metric, the statistical method, and the resources.

- Select the metric you want to view.
- Select the statistical method by which the metric data is aggregated. You can choose maximum, minimum, or average.
- Select the resources that you want to monitor.

8. To add a metric, click AddMetrics and repeat the preceding steps.

9. Click Save. The chart is displayed on the dashboard.

10. If you want to resize the chart, drag the right border, lower border, or lower-right corner of the chart.

Metrics

- **Dashboards:** Displays the monitoring data of Alibaba Cloud products.
- **Log monitoring:** metrics added through log monitoring.
- **Custom:** metrics added through custom monitoring.
- **Metrics:** monitoring indicators, such as CPU usage and memory usage.
- **Statistical method:** means by which metric values are aggregated during a statistical period. Some common statistical methods are maximum, minimum, and average.
- **Resource:** You can use an application group or instance to filter resources and view the monitoring data of these resources.

2 Host monitoring

2.1 Host monitoring overview

The host monitoring service of CloudMonitor allows you to monitor your servers in a systematic manner by installing an agent on the servers. Host monitoring currently supports Linux and Windows Operating Systems (OSs).

Scenarios

Host monitoring is available for both Alibaba Cloud ECS servers, and virtual and physical machines provided by other vendors.

Host monitoring collects statistics of a diverse range of OS-related metrics by using the agent, allowing you to retrieve the server resource usage and obtain metrics for troubleshooting.

Hybrid cloud monitoring solution

Host monitoring uses the agent to collect server metrics. You can install the agent on an ECS server or a non-ECS server for monitoring on and off the cloud.

Enterprise-level monitoring solution

Host monitoring also provides an application group function, which allows you to allocate servers from different regions of Alibaba Cloud to the same group for more efficient server management from a business operations perspective. Host monitoring supports group-based alarm management, meaning that you only need to configure one alarm rule for the entire group, which can improve O&M efficiency and the overall management experience.



Note:

- Host monitoring supports Linux and Windows, but does not support Unix.
- Root permissions are required for the agent installation on a Linux OS and administrator permissions are required for that on a Windows OS.

- The TCP status statistics function is similar to the Linux `netstat -anp` command. This function is disabled by default because a large portion of CPU time is consumed when many TCP connections exist.
 - To enable this function in Linux, set `netstat.tcp.disable` in the `cloudmonitor/config/conf.properties` configuration file to `false`. Restart the agent after you modify the configuration.
 - To enable this function in Windows, set `netstat.tcp.disable` in the `C:\Program Files\Alibaba\cloudmonitor\config` configuration file to `false`. Restart the agent after you modify the configuration.

Monitoring capability

Host monitoring provides more than 30 metrics covering CPU, memory, disk, and network to meet your monitoring and O&M requirements. Click [here](#) to view the full list of the metrics.

Alarm capability

Host monitoring provides an alarm service for all metrics, allowing you to set alarm rules for instances, application groups, and all resources. You can use the alarm service according to your business requirements.

You can use the alarm service directly in the host monitoring list or apply the alarm rules to your application groups after you add servers into the groups.

2.2 Process monitoring

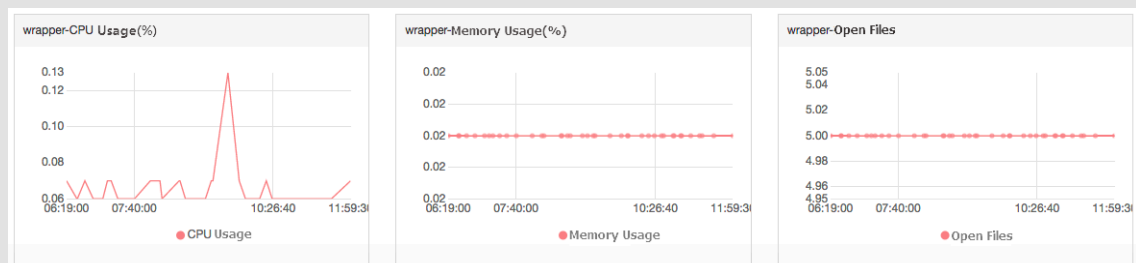
By default, process monitoring allows you to collect information about CPU usage, memory usage, and the number of files recently opened by active processes during some period of time. If you add a process keyword, the number of processes containing the keyword is collected.

View the resource consumption of active processes

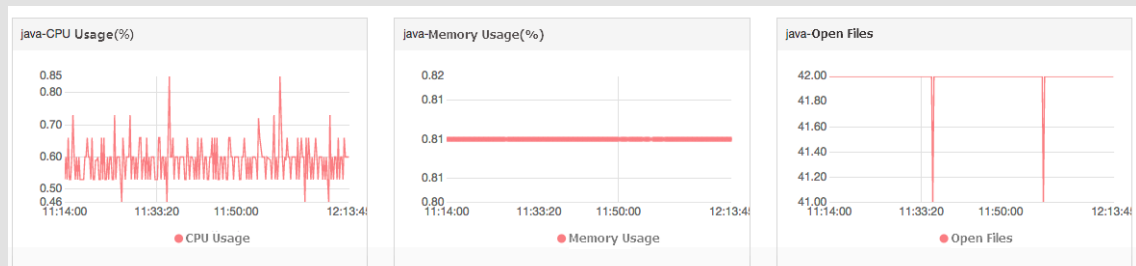
- The CloudMonitor agent filters out the top five processes with the most CPU usage every minute, and records the respective CPU usage, memory usage, and number of files opened by these processes.
- For the CPU and memory usage of a process, see the Linux `top` command.
- For the number of files opened by an active process, see the Linux `lsof` command.

**Note:**

- If your process occupies multiple CPU cores, the percentage shown for CPU usage may exceed 100% because the collected result indicates the total usage of the multiple CPU cores.
- If, during the time period specified for your query, the top five processes have changed, the process list will display all processes that have ever ranked as top five over the specified time period. The times in the list indicate when the processes last ranked in the top five.
- The CPU usage and memory usage, and the number of opened files are collected only for the top five processes. Therefore, if a process has not ranked top five continuously over the time period specified for your query, its data points will appear discontinuous in the charts. The density of the data points for a process indicates its degree of activity on the server.
- As shown in the following figure, the wrapper process has not continuously ranked in the top five processes each time measured. Therefore, the data points in the charts are sparse and discontinuous. The data points in the following charts mean that the process has ranked top five for the particular time measured.



- The following figure shows the charts of the java process. The data points in the charts are dense and continuous. This means that the process continuously ranks in the top five processes with the most CPU usage.



Monitor the number of specified processes

You can learn the number and viability status of key processes by monitoring the number of processes. Specifically, you can add process keywords to the Number of Processes(Count) chart to monitor the number of related processes.

- Add processes for monitoring

For example, assume the following processes run on your server: `/usr/bin/java -Xmx2300m -Xms2300m org.apache.catalina.startup.Bootstrap`, `/usr/bin/ruby`, and `nginx -c /ect/nginx/nginx.conf`. You then add the following six keywords (the keywords can be process names, file paths, parameter names, or other related words), and the corresponding number of processes for each target keyword is output as follows:

- Keyword: `ruby`, number of processes collected: 1
- Keyword: `nginx`, number of processes collected: 1
- Keyword: `/usr/bin`, number of processes collected: 2
- Keyword: `apache.catalina`, number of processes collected: 1
- Keyword: `nginx.conf`, number of processes collected: 1
- Keyword: `-c`, number of processes collected: 1

Procedure

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Host Monitoring.
3. Click the name of the target host, or click Monitoring Charts in the Actions column to access the host monitoring details page.
4. On the displayed page, click the Process Monitoring tab.
5. Rest the pointer over the Number of Processes(Count) chart, and then click Add Process.
6. On the displayed Add Process Monitor page, add the name or keyword of the process you want to monitor and click Add.

- Delete a monitored process

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Host Monitoring.
3. Click the name of the target host, or click Monitoring Charts in the Actions column to access the host monitoring details page.
4. On the displayed page, click the Process Monitoring tab.
5. Rest the pointer over the Number of Processes(Count) chart, and then click Add Process.
6. On the displayed page, find the target process name or keyword and click Delete.

- Set alarm rules

After you configure monitoring for the specified process, you can configure alarm rules for the process. After that, you can receive an alarm notification when the number of the processes changes.

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Host Monitoring.
3. Find the host for which you want to set process monitoring alarm rules, and then click Alarm Rules in the actions column.
4. Click Create Alarm Rule in the upper-right corner of the page.
5. In the Set Alarm Rules area, select (Agent)Host.process.number from the Rule Describe drop-down list, set an appropriate alarm threshold, and then select the process you want to monitor from the processName drop-down list. If multiple processes are configured on the host, the number of processes varies. You can click Add Alarm Rule to configure alarm rules for multiple processes at a time.

2. Set Alarm Rules

Alarm Type: **Threshold Value Alarm** Event Alarm

Alarm Rule:

Rule Describe: (Agent) Host.process.number 1mins Average < 1 Count/Min

processName: Anyprocess Custom

Alarm Rule: Delete

Rule Describe: (Agent) Host.process.number 5mins Average > 6 Count/Min

processName: Anyprocess Custom

[+Add Alarm Rule](#)

Where is the alarm template?

(Agent) Host.process.number—Average—emr_C-7AF9E7BFD87B0EDF_2_RWJW—dfasdf
Alarm Line (Value: 6)

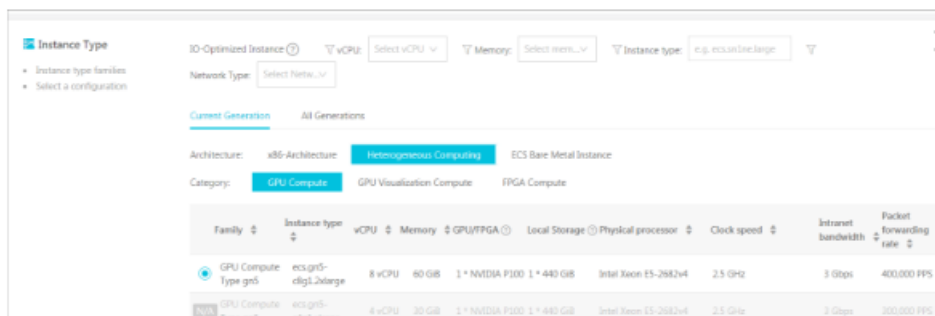
▲ 1/2 ▼

2.3 Query GPU monitoring data

You can query GPU monitoring data in two ways: CloudMonitor console and API

Query GPU monitoring data through the CloudMonitor console

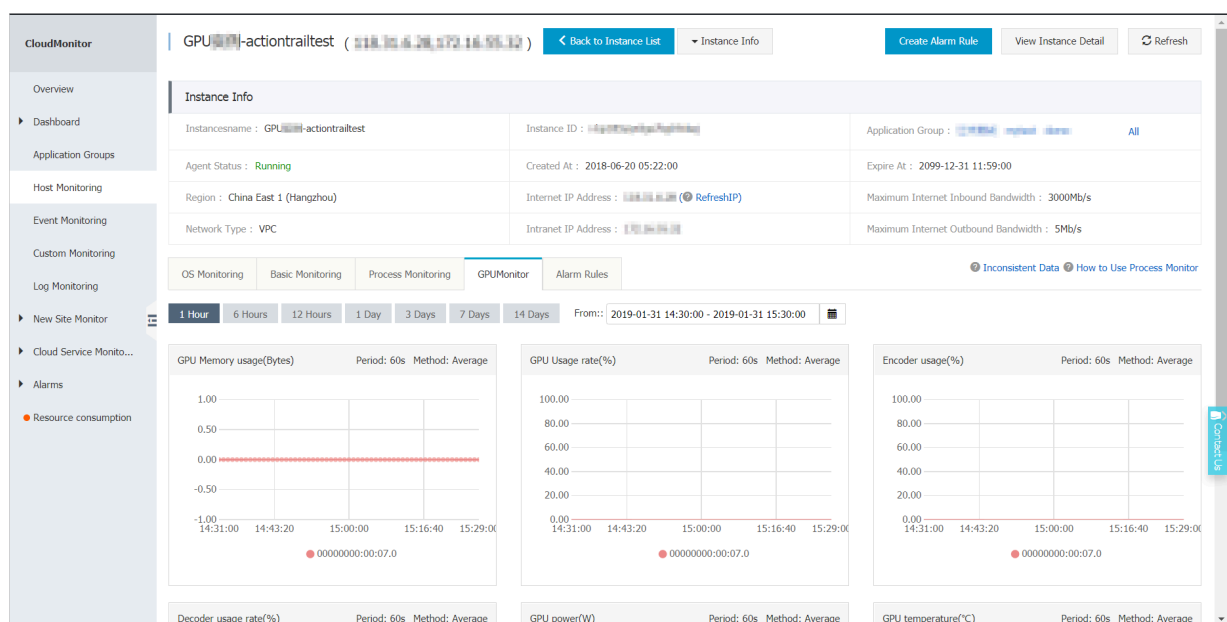
After purchasing an instance of the GPU Compute type for ECS, you only need to install the [GPU driver](#) and the 1.2.28 version of the CloudMonitor plug-in before you can view and configure GPU-related metric charts, or set alarm rules.



View metric charts

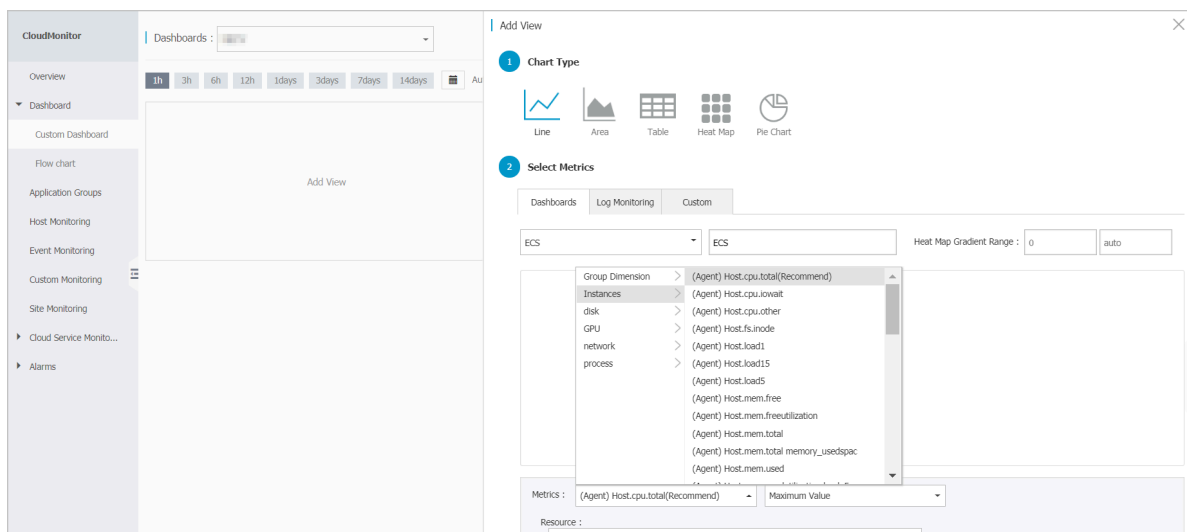
To view GPU-related metric charts, follow these steps:

1. Log on to the [CloudMonitor console](#).
2. Click Host Monitoring in the navigation bar.
3. Query GPU-related metric charts on the GPUMonitor page, as shown in the following figure.

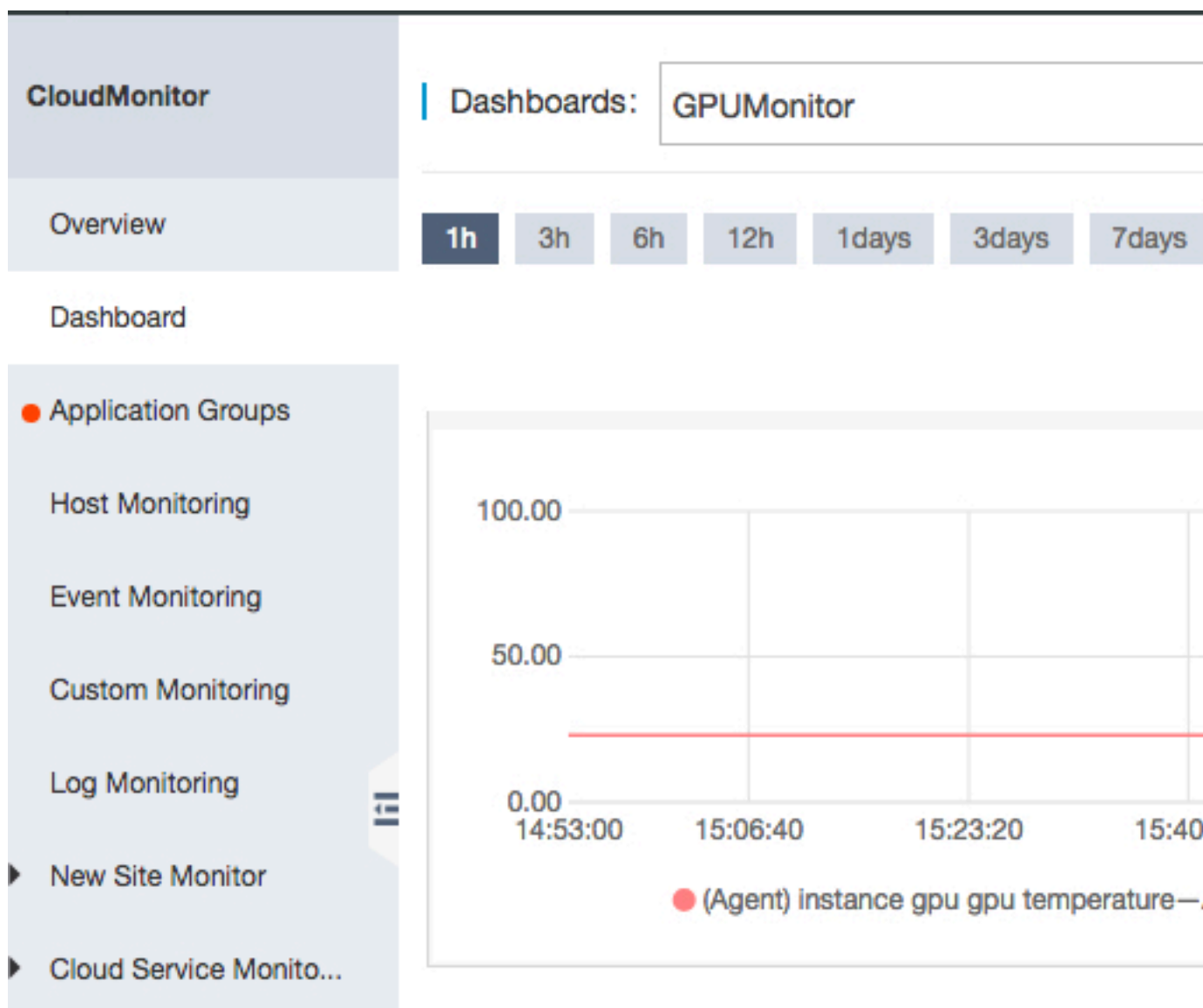


Configure metric charts

1. Log on to the [CloudMonitor console](#).
2. Click Dashboard in the navigation bar.
3. Click Create Dashboard on the Dashboard page.
4. In the pop-up dialog box, enter the name of the new dashboard and click Create.
5. On the refreshed page, click Add View.
6. On the Add View page, select the chart type you want, and then select Metrics, as shown in the following figure.



7. Select the metrics you want from the drop-down menu. The following figure uses the instance-dimension GPU temperature as an example.



Set alarm rules

The method of adding alarm rules for new GPU metrics is the same as that for other ECS metrics. It is recommended that you add GPU alarm rules in batches by applying the template to groups after creating the template. For details, see [Alarm template](#).

Metrics description

GPU-related metrics are classified into three dimensions: GPU, instance, and group.

GPU-dimension metrics

GPU-dimension metrics measure monitoring data on a per GPU basis. The following table lists GPU-dimension metrics.

Metric Name	Unit	Description	Dimensions
gpu_memory_freespace	Byte	GPU-dimension memory free space	instanceId,gpuId
gpu_memory_totalspace	Byte	GPU-dimension memory total space	instanceId,gpuId
gpu_memory_usedspace	Byte	GPU-dimension memory used space	instanceId,gpuId
gpu_gpu_utilization	%	GPU-dimension GPU utilization	instanceId,gpuId
gpu_encoder_utilization	%	GPU-dimension encoder utilization	instanceId,gpuId
gpu_decoder_utilization	%	GPU-dimension decoder utilization	instanceId,gpuId
gpu_gpu_temperature	°C	GPU-dimension GPU temperature	instanceId,gpuId
gpu_power_readings_power_draw	W	GPU-dimension GPU power	instanceId,gpuId
gpu_memory_freeutilization	%	GPU-dimension memory idle rate	instanceId,gpuId
gpu_memory_useutilization	%	GPU-dimension memory utilization	instanceId,gpuId

Instance-dimension metrics

Instance-dimension metrics measure the maximum, minimum, or average value of multiple GPUs on a per instance basis, so that you can query the overall resource usage at the instance level.

Metric Name	Unit	Description	Dimensions
instance_gpu_decoder_utilization	%	Instance-dimension GPU decoder utilization	instanceId
instance_gpu_encoder_utilization	%	Instance-dimension GPU encoder utilization	instanceId

Metric Name	Unit	Description	Dimensions
instance_gpu_gpu_temperature	°C	Instance-dimension GPU temperature	instanceId
instance_gpu_gpu_utilization	%	Instance-dimension GPU utilization	instanceId
instance_gpu_memory_freespace	Byte	Instance-dimension GPU memory free space	instanceId
instance_gpu_memory_freeutilization	%	Instance-dimension GPU memory idle rate	instanceId
instance_gpu_memory_totalspace	Byte	Instance-dimension GPU memory total space	instanceId
instance_gpu_memory_usedspace	Byte	Instance-dimension GPU memory used space	instanceId
instance_gpu_memory_utilization	%	Instance-dimension GPU memory utilization	instanceId
instance_gpu_power_readings_power_draw	W	Instance-dimension GPU power	instanceId

Group-dimension metrics

Group-dimension metrics measure the maximum, minimum, or average value of multiple instances on a per group basis, so that you can query the overall resource usage at the group level.

Metric Name	Unit	Description	Dimensions
group_gpu_decoder_utilization	%	Group-dimension GPU decoder utilization	groupId
group_gpu_encoder_utilization	%	Group-dimension GPU encoder utilization	groupId

Metric Name	Unit	Description	Dimensions
group_gpu_gpu_temperature	°C	Group-dimension GPU temperature	groupId
group_gpu_gpu_utilization	%	Group-dimension GPU utilization	groupId
group_gpu_memory_freespace	Byte	Group-dimension GPU memory free space	groupId
group_gpu_memory_free_utilization	%	Group-dimension GPU memory idle rate	groupId
group_gpu_memory_totalspace	Byte	Group-dimension GPU memory total space	groupId
group_gpu_memory_usespace	Byte	Group-dimension GPU memory used space	groupId
group_gpu_memory_utilization	%	Group-dimension GPU memory utilization	groupId
group_gpu_power_readings_power_draw	W	Group-dimension GPU power	groupId

Query GPU monitoring data through the API

- See [QueryMetricList](#).
- Parameter description: The "Project" parameter should be set to "acs_ecs_dashboard." For the values of the "Metric" and "Dimensions" parameters, see the GPU metrics in the preceding tables.

2.4 Host monitoring metrics

Host monitoring metrics are divided into agent-collected metrics and ECS native metrics. Agent-collected metrics are collected every 15 seconds, and ECS basic metrics are collected every minute.



Note:

The ECS basic metric data may be inconsistent with the operating system (OS) metric data mainly because of:

- **Different statistical frequencies** Metric chart data has the average values collected during measurement periods. The statistical frequency of basic monitoring is one minute, whereas that of OS monitoring is 15 seconds. In case of large metric data fluctuations, basic metric data is smaller than OS metric data because the former data is de-peaked.
- **Different statistical perspectives** The network traffic billing data in basic monitoring does not include the unbilled network traffic between ECS and Server Load Balancer. Whereas, the network traffic statistics in OS monitoring records the actual network traffic of each network adapter. Therefore, the network data in OS monitoring is greater than that in basic monitoring (that is, the agent-collected data is greater than the actual purchased bandwidth or traffic quota).

Agent-collected metrics

- CPU metrics

You can refer to the Linux top command to understand the meaning of the metrics.

Metric	Definition	Unit	remark
Host.cpu.idle	Percentage of currently idle CPUs	%	Percentage of the current CPU is idle
Host.cpu.system	Percentage of the current kernel space used as CPU	%	This metric measures the consumption resulting from system context switchover. A great value indicates that many processes or threads are running on the server.
Host.cpu.user	This metric measures the CPU consumption of user processes.	%	CPU consumption by user processes

Metric	Definition	Unit	remark
Host.CPU.iowait	Percentage of CPUs currently waiting for Io operation	%	This is a relatively high value, which means that there are frequent Io operations.
Host.cpu.other	Other CPU usage percentage	%	Other consumption, calculated in the form of (Nice + softirq + IRQ + stolen) Consumption
Host.cpu.totalUsed	Percentage of total CPU currently consumed	%	The sum of the CPU consumption above, usually used for alarm purposes.

- Memory related monitors

You can refer to the free command to understand the meaning of the indicators.

Metrics	Definition	Unit	Description
Host.mem.total	Total memory	Bytes	Total Server Memory
Host.mem.used	Amount of used memory	Bytes	Memory Used by the user program + buffers + Cache, the amount of memory used for the buffer, and the amount of memory used for the system cache used by the cache
Host.mem.actualused	Memory actually used by the user	Bytes	calculation formula:(used - buffers - cached)
Host.mem.free	Amount of memory remaining	Bytes	Calculated as (total memory-amount of memory used)

Metrics	Definition	Unit	Description
Host.mem.freeutilization	Percentage of memory remaining	%	Calculated as (amount of remaining memory / total amount of memory * 100)
Host.mem.usedutilization	Memory usage	%	Calculated as (actual used / total * 100)

- Metrics of average system load

You can refer to the Linux TOP command to understand what the metrics mean . The higher the value of the monitoring item indicates that the more busy the system is.

Metrics	Definition	Unit
Host.load1	Average system load over the past 1 minute , Windows operating system does not have this metric	None
Host. load5	Average system load over the past 5 minutes , Windows operating system does not have this metric	None
Host. load15	Average system load over the past 15 minutes , Windows operating system does not have this metric	None

- Disk related metrics

- Disk usage and inode usage refer to the Linux DF command.
- Disk read/write metrics can refer to the Linux iostat command.

Metric	Definition	Unit
Host.diskusage.used	Used storage space on disk	Bytes
Host.disk.utilization	Disk usage	%

Metric	Definition	Unit
Host.diskusage.free	Remaining storage space on disk	Bytes
Host.diskusage.total	Total disk storage	Bytes
Host.disk.readbytes	The number of bytes read per second by the disk.	Bytes/s
Host.disk.writebytes	Number of bytes written per second on disk	Bytes/s
Host.disk.readiops	Number of read requests per second on disk	Times/second
Host.disk.writeiops	Number of write requests per second on disk	Times/second

• File System Monitor

Metrics	Definition	Unit	Description:
Host.fs.inode	Inode usage, the Unix/Linux system uses inode numbers to identify files, and the disks are not fully stocked, however, when inode has been assigned, it will not be able to create a new file on disk, windows operating system does not have this metric.	%	Inode number represents the number of file system files, and a large number of small files can cause too high inode usage.

• Network related metrics

- You can refer to the Linux iftop command For a collection of TCP connections, refer to the Linux SS Command.
- The number of TCP connections is collected by default By default, statistics are collected on the number of TCP connections by TCP_TOTAL (total connections), ESTABLISHED (normally established connections), and NON_ESTABLISHED

(connections not in the established state). If you want to obtain the number of connections in each state, follow the subsequent procedure:

■ Linux

Set `netstat.tcp.disable` in the `cloudmonitor/config/conf.properties` configuration file to `false` to enable data collection. Restart the Agent once you modify the configuration. Restart the Agent once you modify the configuration.

■ Windows

Set `netstat.tcp.disable` in the `C:\Program\Alibaba\cloudmonitor\config` configuration file to `false` to enable data collection. Restart the Agent once you modify the configuration.

Metric	Definition	Unit
Host.netin.rate	Number of bits received by the network adapter per second, that is, the uplink bandwidth of the network adapter.	bits/s
Host.netout.rate	Number of bits sent by the network adapter per second, that is, the downlink bandwidth of the network adapter.	bits/s
Host.netin.packages	Number of packets received by the network adapter per second.	packages/s
Host.netout.packages	Number of incoming error packets detected by the drive.	packages/s
Host.netin.errorpackage	Number of outgoing error packets detected by the drive.	packages/s
Host.netout.errorpackages	Number of outgoing error packets detected by the drive.	packages/s

Metric	Definition	Unit
Host.tcpconnection	Number of TCP connections in various states, including LISTEN, SYN_SENT, ESTABLISHED, SYN_RECV, FIN_WAIT1, CLOSE_WAIT, FIN_WAIT2, LAST_ACK, TIME_WAIT, CLOSING, and CLOSED.	

- Process metrics

- For details regarding process-specific CPU usage and memory usage, refer to the Linux top command. CPU usage indicates the CPU consumption of multiple kernels.
- For details about Host.process.openfile, refer to the Linux lsof command.
- For details about Host.process.number, refer to the Linux ps aux |grep 'keyword' command.

Metric	Definition	Unit
Host.process.cpu	CPU usage of a process.	%
Host.process.memory	Memory usage of a process.	%
Host.process.openfile	Number of files opened by a process.	Files
Host.process.number	Number of processes that match the specified keyword.	Processes

ECS metrics

If your host is an ECS server, the following metrics are provided without agent installation once you purchase an ECS instance. The collection granularity is one minute.

Metric	Definition	Unit
ECS.CPUUtilization	CPU usage	%
ECS.InternetInRate	Average rate of Internet inbound traffic.	bits/s

Metric	Definition	Unit
ECS.IntranetInRate	Average rate of intranet inbound traffic.	bits/s
ECS.InternetOutRate	Average rate of Internet outbound traffic.	bits/s
ECS.IntranetOutRate	Average rate of intranet outbound traffic.	bits/s
ECS.SystemDiskReadbps	Number of bytes read from the system disk per second.	Bytes/s
ECS.SystemDiskWritebps	Number of bytes written to the system disk per second.	Bytes/s
ECS.SystemDiskReadOps	Number of times data is read from the system disk per second.	times/s
ECS.SystemDiskWriteOps	Number of times data is written to the system disk per second.	times/s
ECS. internetin	Internet inbound traffic.	bytes
ECS.InternetOut	Internet outbound traffic.	bytes
ECS.IntranetIn	Intranet inbound traffic.	bytes
ECS.IntranetOut	Intranet outbound traffic.	bytes

2.5 Alarm service

Host monitoring provides the alarm service so that you can set alarm rules for a target server, or add servers to an application group and then set alarm rules at the group level. For more information about setting alarm rules for an application group, see [Managing Alarm Rules](#).

Create an alarm rule

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Host Monitoring.
3. Click the Alarm Rules tab.
4. Click Create Alarm Rule in the upper-right corner.

5. In the displayed dialog box, set the parameters. For more information, see [Manage alarm rules](#).
6. Click Confirm to save your alarm rule settings.

Delete an alarm rule

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Host Monitoring.
3. Click the Alarm Rules tab.
4. Find the target alarm rule and click Delete in the Actions column. If you want to delete multiple rules at a time, select the target rules and click Delete under the alarm rule list.

Modify an alarm rule

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Host Monitoring.
3. Click the Alarm Rules tab.
4. Find the target alarm rule and click Modify.

View alarm rules

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Host Monitoring.
3. Click the Instances tab. Then, find the target host and click Alarm Rules in the Actions column to view the alarm rules of the host.
4. To view all the alarm rules, go to the Alarm Rules tab page.

2.6 CloudMonitor Java agent introduction

CloudMonitor provides you with a powerful host monitoring agent that allows you to monitor your servers systematically. The following is a brief introduction to this service, including its installation and resource usage.

Installation path

- **Linux:** `/usr/local/cloudmonitor`
- **Windows:** `C:\Program Files\Alibaba\cloudmonitor`

Process information

After an agent is installed, the following two processes run on your server:

- `/usr/local/cloudmonitor/jre/bin/java`
- `/usr/local/cloudmonitor/wrapper/bin/wrapper`

Port description

- TCP port 32000 of the local host is accessed and listened to for daemons.
- TCP port 3128, 8080, or 443 of remote servers is accessed for heartbeat monitoring and monitoring data reporting. Port 3128 or 8080 is used for Alibaba Cloud hosts, and port 443 is used for other hosts.
- HTTP port 80 of remote servers is accessed for CloudMonitor agent upgrades.

Agent logs

- Logs of monitoring data are located at `/usr/local/cloudmonitor/logs`.
- Logs of startup, shutdown, and daemons are located at `/usr/local/cloudmonitor/wrapper/logs`.
- You can modify `/usr/local/cloudmonitor/config/log4j.properties` to adjust the log level.

Resource usage

- The process `/usr/local/cloudmonitor/wrapper/bin/wrapper` occupies about 1 MB of memory with little to no CPU usage.
- The process `/usr/local/cloudmonitor/jre/bin/java` occupies about 70 MB of memory and 1% to 2% of one core's CPU usage.
- The installation package is 70 MB and occupies about 200 MB of disk space after the installation is complete.
- Logs use a maximum space of 40 MB and are cleared if they use more than 40 MB.
- Monitoring data is sent every 15 seconds, occupying about 10 KB intranet bandwidth.
- Heartbeat data is sent every three minutes, occupying about 2 KB intranet bandwidth.

External dependencies

- The Java agent of CloudMonitor is built in with JRE 1.8.
- Java service wrapper is used for daemons, start up at boot, and Windows service registration.

- The `ss -s` command is used to capture a TCP connection, and if you do not have this command in the current system, you must install `iproute` yourself.

Installation instructions

See [Install CloudMonitor Java agent](#).

Install an agent on a host not provided by Alibaba Cloud

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Host Monitoring.
3. At the top of the displayed page, click Not Aliyun ecs install. In the Monitor Install Guide dialog box that is displayed, select the agent type and host type to view the corresponding installation method.

2.7 Install CloudMonitor Java agent

Install a CloudMonitor Java agent on Linux

Frequently used commands

```
# Running status
/usr/local/cloudmonitor/wrapper/bin/cloudmonitor.sh status

# Start
/usr/local/cloudmonitor/wrapper/bin/cloudmonitor.sh start

# Stop
/usr/local/cloudmonitor/wrapper/bin/cloudmonitor.sh stop

# Restart
/usr/local/cloudmonitor/wrapper/bin/cloudmonitor.sh restart

# Uninstall
/usr/local/cloudmonitor/wrapper/bin/cloudmonitor.sh remove && \
rm -rf /usr/local/cloudmonitor
```

Installation command

This command varies by region. Copy the corresponding command and then run it on your server as a root user.

China North 1 (Qingdao) cn-qingdao

```
REGION_ID=cn-qingdao VERSION=1.3.7 \
```

```
bash -c "$(curl https://cms-agent-cn-qingdao.oss-cn-qingdao-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

China North 2 (Beijing) cn-beijing

```
REGION_ID=cn-beijing VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-cn-beijing.oss-cn-beijing-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

China North 3 (Zhangjiakou) cn-zhangjiakou

```
REGION_ID=cn-zhangjiakou VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-cn-zhangjiakou.oss-cn-zhangjiakou-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

China North 5 (Hohhot) cn-huhehaote

```
REGION_ID=cn-huhehaote VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-cn-huhehaote.oss-cn-huhehaote-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

China East 1 (Hangzhou) cn-hangzhou

```
REGION_ID=cn-hangzhou VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-cn-hangzhou.oss-cn-hangzhou-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

China East 2 (Shanghai) cn-shanghai

```
REGION_ID=cn-shanghai VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-cn-shanghai.oss-cn-shanghai-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

China South 1 (Shenzhen) cn-shenzhen

```
REGION_ID=cn-shenzhen VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-cn-shenzhen.oss-cn-shenzhen-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

Hong Kong (China) cn-hongkong

```
REGION_ID=cn-hongkong VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-cn-hongkong.oss-cn-hongkong-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

US West 1 (Silicon Valley) us-west-1

```
REGION_ID=us-west-1 VERSION=1.3.7 \  

```

```
bash -c "$(curl https://cms-agent-us-west-1.oss-us-west-1-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

US East 1 (Virginia) us-east-1

```
REGION_ID=us-east-1 VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-us-east-1.oss-us-east-1-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

Asia Pacific SE 1 (Singapore) ap-southeast-1

```
REGION_ID=ap-southeast-1 VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-ap-southeast-1.oss-ap-southeast-1-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

Asia Pacific SE 2 (Sydney) ap-southeast-2

```
REGION_ID=ap-southeast-2 VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-ap-southeast-2.oss-ap-southeast-2-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

Asia Pacific SE 3 (Kuala Lumpur) ap-southeast-3

```
REGION_ID=ap-southeast-3 VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-ap-southeast-3.oss-ap-southeast-3-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

Asia Pacific SE 5 (Jakarta) ap-southeast-5

```
REGION_ID=ap-southeast-5 VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-ap-southeast-5.oss-ap-southeast-5-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

Asia Pacific NE 1 (Tokyo) ap-northeast-1

```
REGION_ID=ap-northeast-1 VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-ap-northeast-1.oss-ap-northeast-1-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

Asia Pacific SOU 1 (Mumbai) ap-south-1

```
REGION_ID=ap-south-1 VERSION=1.3.7 \  
bash -c "$(curl https://cms-agent-ap-south-1.oss-ap-south-1-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

EU Central 1 (Frankfurt) eu-central-1

```
REGION_ID=eu-central-1 VERSION=1.3.7 \  

```

```
bash -c "$(curl https://cms-agent-eu-central-1.oss-eu-central-1-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

UK (London) eu-west-1

```
REGION_ID=eu-west-1 VERSION=1.3.7 \ bash -c "$(curl https://cms-agent-eu-west-1.oss-eu-west-1-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

Middle East 1 (Dubai) me-east-1

```
REGION_ID=me-east-1 VERSION=1.3.7 \ bash -c "$(curl https://cms-agent-me-east-1.oss-me-east-1-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

China East 1 Finance Cloud (Hangzhou) cn-hangzhou

```
REGION_ID=cn-hangzhou VERSION=1.3.7 \ bash -c "$(curl https://cms-agent-cn-hangzhou.oss-cn-hangzhou-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

China East 2 Finance Cloud (Shanghai) cn-shanghai-finance-1

```
REGION_ID=cn-shanghai-finance-1 VERSION=1.3.7 \ bash -c "$(curl https://cms-agent-cn-shanghai-finance-1.oss-cn-shanghai-finance-1-pub-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

China South 1 Finance Cloud (Shenzhen) cn-shenzhen-finance-1

```
REGION_ID=cn-shenzhen-finance-1 VERSION=1.3.7 \ bash -c "$(curl http://cms-agent-cn-shenzhen-finance-1.oss-cn-shenzhen-finance-1-internal.aliyuncs.com/release/cms_install_for_linux.sh)"
```

Install a CloudMonitor Java agent on Windows

Installation procedure

1. Download [64-bit agent version](#) or [32-bit agent version](#) based on your operating system version.
2. Create a folder in the path *C:/Program Files/Alibaba* and name it *cloudmonitor*.
3. Decompress the installation package to *C:/Program Files/Alibaba/cloudmonitor* or.
4. Double-click *C:/Program Files/Alibaba/cloudmonitor/wrapper/bin/InstallApp-NT.bat* as an administrator to install CloudMonitor.

5. Double-click `C:/Program Files/Alibaba/cloudmonitor/wrapper/bin/StartApp-NT.bat` as an administrator to start CloudMonitor.
6. After the installation is complete, you can view, start, and stop CloudMonitor through the service panel of Windows.

Uninstall procedure

1. Stop CloudMonitor through the service panel of Windows.
2. Run `C:/Program Files/Alibaba/cloudmonitor/wrapper/bin/UninstallApp-NT.bat` as an administrator to delete CloudMonitor.
3. In the installation directory, delete the entire directory `C:/Program Files/Alibaba/cloudmonitor`.

Download the agent with no Internet connection

If you do not have an Internet connection, you can download the installation package from the intranet. For example, if the region of your host is Qingdao and the host uses a 64-bit system, then the intranet download address is as follows: <http://cms-agent-cn-qingdao.oss-cn-qingdao.aliyuncs.com/release/1.3.7/windows64/agent-windows64-1.3.7-package.zip>.

- For a host in another region, change `cn-qingdao` to the corresponding region ID.
- For a host that uses a 32-bit system, change `windows64` to `windows32`.
- For another version, change `1.3.7` to the corresponding version number.

Security configuration instructions

The following table lists the ports that the CloudMonitor agent uses to interact with your server. If the security software disables these ports, monitoring data may fail to be collected. If your ECS server requires a high level of security, you can add one of the following IP addresses to the whitelist.



Note:

Future version updates and maintenance of CloudMonitor may cause changes to the following IP addresses. To simplify the configuration of your firewall rules, we recommend that you directly allow the 100.100 network segment in the egress direction. This network segment is reserved for the intranet of Alibaba Cloud with no security issues.

Region	IP	Direction	Description
China East 1 (Hangzhou) cn-hangzhou	100.100.19.43:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.45.73:80	Egress	Used to collect monitoring data to CloudMonitor
China North 2 (Beijing) cn-beijing	100.100.18.22:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.18.50:80	Egress	Used to collect monitoring data to CloudMonitor
China North 1 (Qingdao) cn-qingdao	100.100.36.102:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.15.23:80	Egress	Used to collect monitoring data to CloudMonitor
China South 1 (Shenzhen) cn-shenzhen	100.100.0.13:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.0.31:80	Egress	Used to collect monitoring data to CloudMonitor

Region	IP	Direction	Description
Hong Kong (China) cn-hongkong	100.103.0.47:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.103.0.45:80	Egress	Used to collect monitoring data to CloudMonitor
China North 5 (Hohhot) cn-huhehaote	100.100.80.135:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.12:80	Egress	Used to collect monitoring data to CloudMonitor
China North 3 (Zhangjiakou) cn-zhangjiakou	100.100.80.92:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.0.19:80	Egress	Used to collect monitoring data to CloudMonitor
China East 2 (Shanghai) cn-shanghai	100.100.36.11:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.36.6:80	Egress	Used to collect monitoring data to CloudMonitor

Region	IP	Direction	Description
China SW 1 (Chengdu) cn-chengdu	100.100.80.229:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.14:80	Egress	Used to collect monitoring data to CloudMonitor
US East 1 (Virginia) us-east-1	100.103.0.95:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.103.0.94:80	Egress	Used to collect monitoring data to CloudMonitor
US West 1 (Silicon Valley) us-west-1	100.103.0.95:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.29.7:80	Egress	Used to collect monitoring data to CloudMonitor
EU Central 1 (Frankfurt) eu-central-1	100.100.80.241:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.72:80	Egress	Used to collect monitoring data to CloudMonitor

Region	IP	Direction	Description
UK (London) eu-west-1	100.100.0.3:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.0.2:80	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific SE 1 (Singapore) ap-southeast-1	100.100.30.20:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.103.7:80	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific SE 2 (Sydney) ap-southeast-2	100.100.80.92:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.13:80 [47.91.39.6:443]	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific SE 3 (Kuala Lumpur) ap-southeast-3	100.100.80.153:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.140:80	Egress	Used to collect monitoring data to CloudMonitor

Region	IP	Direction	Description
Asia Pacific SE 5 (Jakarta) ap-southeast-5	100.100.80.160:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.180:80	Egress	Used to collect monitoring data to CloudMonitor
Middle East 1 (Dubai) me-east-1	100.100.80.142:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.151:80 [47.91.99.5:443]	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific NE 1 (Tokyo) ap-northeast-1	100.100.80.184:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.137:80 [47.91.8.7:443]	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific SOU 1 (Mumbai) ap-south-1	100.100.80.152:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.66:80	Egress	Used to collect monitoring data to CloudMonitor

Resource consumption

- Installation package size: 75 MB
- Space occupied after installation: 200 MB
- Memory: 64 MB
- CPU: less than 1%
- Network: intranet, with no Internet bandwidth consumption

FAQ

- Where are CloudMonitor logs saved?
 - **Linux:** `/usr/local/cloudmonitor/logs`
 - **Windows:** `C:/Program Files/Alibaba/cloudmonitor/logs`
- What should I do if there is a conflict between the port occupied by the agent and the port used by my service?
 1. Change the port range by modifying the CloudMonitor configuration, with the file location: `/usr/local/cloudmonitor/wrapper/conf/wrapper.conf`.
 2. Restart CloudMonitor.

```
wrapper.port.min=40000
wrapper.port.max=41000
wrapper.jvm.port.min=41001
wrapper.jvm.port.max=42000
```

2.8 Introduction to the CloudMonitor GoLang agent

This topic provides a brief introduction to the CloudMonitor GoLang agent and its installation and resource usage. The GoLang agent can enable you to monitor your servers in a centralized and systematic manner.

Installation path

- **Linux:** `/usr/local/cloudmonitor`
- **Windows:** `C:\Program Files\Alibaba\cloudmonitor`

Process information

After the agent is installed, the following two processes run on your server:

- **Linux 32-bit:** `CmsGoAgent.linux-386`
- **Linux 64-bit:** `CmsGoAgent.linux-amd64`

- Windows 32-bit: CmsGoAgent.windows-386.exe
- Windows 64-bit: CmsGoAgent.windows-amd64.exe

Port description

- TCP port 3128, 8080, or 443 of remote servers is accessed for heartbeat monitoring and the reporting of monitoring data. Port 3128 or 8080 is used for Alibaba Cloud hosts, and port 443 is used for other hosts.
- HTTP port 80 of remote servers is accessed for CloudMonitor agent upgrades.

Agent logs

- Logs of monitoring data are stored in the log directory.
- You can adjust the level of a log by modifying the `cms.log.level` field in the `config/conf.properties` file. If the field does not exist, you can manually create it. The level of a log can be DEBUG, INFO, WARNING, ERROR, or FATAL.

Resource usage

- The agent process occupies a memory of 10 to 20 MB and 1% to 2% of a single core CPU.
- The size of the agent installation package is 10 to 15 MB.
- Logs use up to 40 MB and are cleared if they use more than 40 MB.
- Monitoring data is sent every 15 seconds, occupying about 10 KB of intranet bandwidth.
- Heartbeat data is sent every 3 minutes, occupying about 2 KB of intranet bandwidth.

Installation instructions

For details, see [Install CloudMonitor GoLang agent](#).

Install the agent on a host not provided by Alibaba Cloud

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Host Monitoring.
3. At the top of the displayed page, click Not Aliyun ecs install. In the Monitor Install Guide dialog box that is displayed, select the agent type and host type to view the corresponding installation method.

2.9 Install CloudMonitor GoLang agent

Install a CloudMonitor GoLang agent on Linux



Note:

1. The binary file name of the agent

`CmsGoAgent.linux-${ARCH}`

The value of "ARCH" can be "amd64" or "386" depending on the architecture of your Linux system.

2. Version

In this topic, the version 2.1.55 is used. We recommend that you use the latest version. You can find the number of the latest version on the host monitoring page in the CloudMonitor console.

The screenshot shows the CloudMonitor console interface. On the left, the 'Host Monitoring' tab is selected in the sidebar (marked with a red circle 1). The main panel shows a list of instances with columns for Instance Name, Agent Status, and Agent Version. A red circle 2 highlights the 'Instances' tab. On the right, the 'Monitor Install Guide' dialog is open. It shows 'Go Lang Agent' selected under 'Agent Type' (marked with a red circle 3) and 'Linux' selected under 'OS' (marked with a red circle 4). The 'Install shell' section displays the command to install the agent, with '2.1.55' highlighted as the version.

Frequently used commands

```
# Register the agent as a system service.
/usr/local/cloudmonitor/CmsGoAgent.linux-${ARCH} install
# Remove the agent from system services.
/usr/local/cloudmonitor/CmsGoAgent.linux-${ARCH} uninstall
# Start the agent.
/usr/local/cloudmonitor/CmsGoAgent.linux-${ARCH} start
# Stop the agent.
/usr/local/cloudmonitor/CmsGoAgent.linux-${ARCH} stop
# Restart the agent.
/usr/local/cloudmonitor/CmsGoAgent.linux-${ARCH} restart
# Uninstall the agent.
/usr/local/cloudmonitor/CmsGoAgent.linux-${ARCH} stop && \
/usr/local/cloudmonitor/CmsGoAgent.linux-${ARCH} uninstall && \
```

```
rm -rf /usr/local/cloudmonitor
```

Installation command

Copy the installation command of the region you require and then run the command on your server with root permissions.



Note:

You can also find the command on the Monitor Install Guide page in the CloudMonitor console.

China North 1 (Qingdao) cn-qingdao

```
REGION_ID=cn-qingdao VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-cn-qingdao.oss-cn-qingdao-internal.  
aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

China North 2 (Beijing) cn-beijing

```
REGION_ID=cn-beijing VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-cn-beijing.oss-cn-beijing-internal.  
aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

China North 3 (Zhangjiakou) cn-zhangjiakou

```
REGION_ID=cn-zhangjiakou VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-cn-zhangjiakou.oss-cn-zhangjiakou-  
internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

China North 5 (Hohhot) cn-huhehaote

```
REGION_ID=cn-huhehaote VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-cn-huhehaote.oss-cn-huhehaote-  
internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

China East 1 (Hangzhou) cn-hangzhou

```
REGION_ID=cn-hangzhou VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-cn-hangzhou.oss-cn-hangzhou-internal  
.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

China East 2 (Shanghai) cn-shanghai

```
REGION_ID=cn-shanghai VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-cn-shanghai.oss-cn-shanghai-internal  
.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

China South 1 (Shenzhen) cn-shenzhen

```
REGION_ID=cn-shenzhen VERSION=2.1.55 \  

```

```
bash -c "$(curl https://cms-agent-cn-shenzhen.oss-cn-shenzhen-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

Hong Kong (China) cn-hongkong

```
REGION_ID=cn-hongkong VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-cn-hongkong.oss-cn-hongkong-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

US West 1 (Silicon Valley) us-west-1

```
REGION_ID=us-west-1 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-us-west-1.oss-us-west-1-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

US East 1 (Virginia) us-east-1

```
REGION_ID=us-east-1 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-us-east-1.oss-us-east-1-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

Asia Pacific SE 1 (Singapore) ap-southeast-1

```
REGION_ID=ap-southeast-1 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-ap-southeast-1.oss-ap-southeast-1-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

Asia Pacific SE 2 (Sydney) ap-southeast-2

```
REGION_ID=ap-southeast-2 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-ap-southeast-2.oss-ap-southeast-2-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

Asia Pacific SE 3 (Kuala Lumpur) ap-southeast-3

```
REGION_ID=ap-southeast-3 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-ap-southeast-3.oss-ap-southeast-3-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

Asia Pacific SE 5 (Jakarta) ap-southeast-5

```
REGION_ID=ap-southeast-5 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-ap-southeast-5.oss-ap-southeast-5-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

Asia Pacific NE 1 (Tokyo) ap-northeast-1

```
REGION_ID=ap-northeast-1 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-ap-northeast-1.oss-ap-northeast-1-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

Asia Pacific SOU 1 (Mumbai) ap-south-1

```
REGION_ID=ap-south-1 VERSION=2.1.55 \  

```

```
bash -c "$(curl https://cms-agent-ap-south-1.oss-ap-south-1-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

EU Central 1 (Frankfurt) eu-central-1

```
REGION_ID=eu-central-1 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-eu-central-1.oss-eu-central-1-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

UK (London) eu-west-1

```
REGION_ID=eu-west-1 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-eu-west-1.oss-eu-west-1-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

Middle East 1 (Dubai) me-east-1

```
REGION_ID=me-east-1 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-me-east-1.oss-me-east-1-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

China East 1 Finance Cloud (Hangzhou) cn-hangzhou

```
REGION_ID=cn-hangzhou VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-cn-hangzhou.oss-cn-hangzhou-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

China East 2 Finance Cloud (Shanghai) cn-shanghai-finance-1

```
REGION_ID=cn-shanghai-finance-1 VERSION=2.1.55 \  
bash -c "$(curl https://cms-agent-cn-shanghai-finance-1.oss-cn-shanghai-finance-1-pub-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

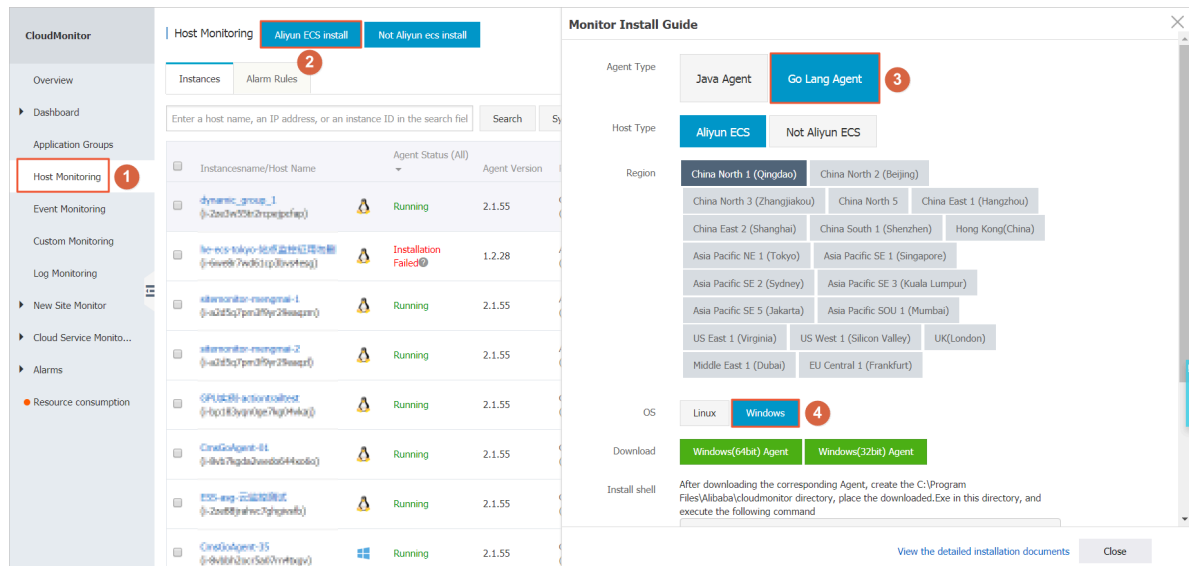
China South 1 Finance Cloud (Shenzhen) cn-shenzhen-finance-1

```
REGION_ID=cn-shenzhen-finance-1 VERSION=2.1.55 \  
bash -c "$(curl http://cms-agent-cn-shenzhen-finance-1.oss-cn-shenzhen-finance-1-internal.aliyuncs.com/cms-go-agent/cms_go_agent_install.sh)"
```

Install a CloudMonitor GoLang agent on Windows

Installation procedure

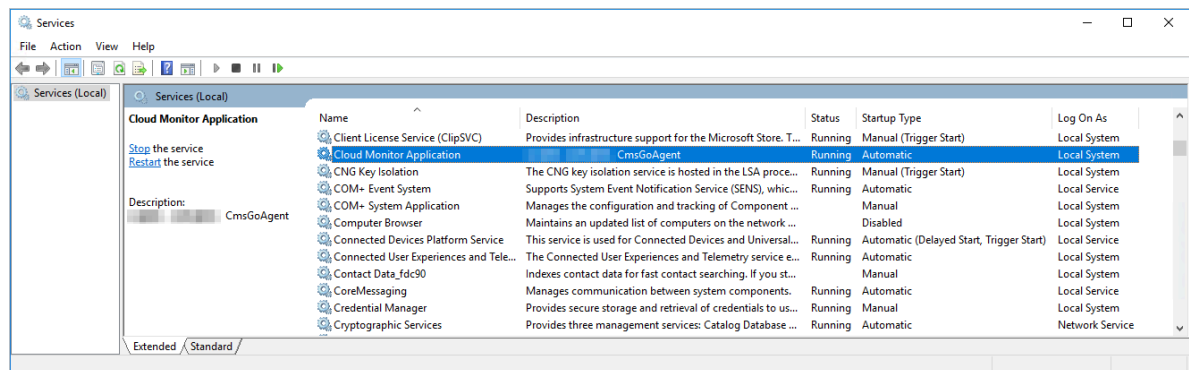
1. Select your region and host type. Then, depending on your operating system version, download a *64-bit agent version* or *32-bit agent version* and save it in C:\Program Files\Alibaba\cloudmonitor.



2. Open the Command Prompt as an administrator.
3. Run the following command:

```
cd"C:\Program Files\Alibaba\cloudmonitor"
CmsGoAgent.windows-amd64.exe install
CmsGoAgent.windows-amd64.exe start
```

4. After the installation is complete, you can use Windows Services to view, start, and stop the agent.



Uninstall procedure

1. Open the Command Prompt as an administrator.
2. Run the following command:

```
cd"C:\Program Files\Alibaba\cloudmonitor"
CmsGoAgent.windows-amd64.exe stop
```

```
CmsGoAgent.windows-amd64.exe uninstall
```

3. Close the Command Prompt, and delete the directory `C:\Program Files\Alibaba\cloudmonitor`.

Download the agent with no Internet connection

If you do not have an Internet connection, you can download the installation package from the intranet. For example, if the region of your host is Qingdao and the host uses a 64-bit system, then the intranet download address is as follows: <http://cms-agent-cn-qingdao.oss-cn-qingdao.aliyuncs.com/cms-go-agent/2.1.55/CmsGoAgent.windows-amd64.exe>.

- For a host in another region, change "cn-qingdao" to the corresponding region ID.
- For a host that uses a 32-bit system, change "amd64" to "386".
- For another version, change "2.1.55" to the corresponding version number.

Security configuration instructions

The following table lists the ports that the CloudMonitor agent uses to interact with your server. Note that monitoring data may not be collected if these ports are disabled by security software. Therefore, in the case that your ECS server requires a higher level of security, we recommend that you add one of the following IP addresses to your whitelist.



Note:

1. Future maintenance and version updates of CloudMonitor may cause changes to the following IP addresses. Therefore, to simplify the configuration of your firewall rules, we recommend that you directly allow the 100.0.0.0/8 CIDR block in the egress direction. This CIDR block is reserved for the intranet of Alibaba Cloud and is free of security issues.
2. The IP addresses in square brackets ([]) are optional. They can be used as backup addresses in the situation that your network connection is poor.

Region	IP	Direction	Description
China East 1 (Hangzhou) cn-hangzhou	100.100.19.43:3128	Egress	Used for monitoring configuration management, and other management and control operations

Region	IP	Direction	Description
	100.100.45.73:80	Egress	Used to collect monitoring data to CloudMonitor
China North 2 (Beijing) cn-beijing	100.100.18.22:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.18.50:80	Egress	Used to collect monitoring data to CloudMonitor
China North 1 (Qingdao) cn-qingdao	100.100.36.102:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.15.23:80	Egress	Used to collect monitoring data to CloudMonitor
China South 1 (Shenzhen) cn-shenzhen	100.100.0.13:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.0.31:80	Egress	Used to collect monitoring data to CloudMonitor
Hong Kong (China) cn-hongkong	100.103.0.47:3128	Egress	Used for monitoring configuration management, and other management and control operations

Region	IP	Direction	Description
	100.103.0.45:80	Egress	Used to collect monitoring data to CloudMonitor
China North 5 (Hohhot) cn-huhehaote	100.100.80.135:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.12:80	Egress	Used to collect monitoring data to CloudMonitor
China North 3 (Zhangjiakou) cn-zhangjiakou	100.100.80.92:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.0.19:80	Egress	Used to collect monitoring data to CloudMonitor
China East 2 (Shanghai) cn-shanghai	100.100.36.11:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.36.6:80	Egress	Used to collect monitoring data to CloudMonitor
China SW 1 (Chengdu) cn-chengdu	100.100.80.229:8080	Egress	Used for monitoring configuration management, and other management and control operations

Region	IP	Direction	Description
	100.100.80.14:80	Egress	Used to collect monitoring data to CloudMonitor
US East 1 (Virginia) us-east-1	100.103.0.95:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.103.0.94:80	Egress	Used to collect monitoring data to CloudMonitor
US West 1 (Silicon Valley) us-west-1	100.103.0.95:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.29.7:80	Egress	Used to collect monitoring data to CloudMonitor
EU Central 1 (Frankfurt) eu-central-1	100.100.80.241:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.72:80	Egress	Used to collect monitoring data to CloudMonitor
UK (London) eu-west-1	100.100.0.3:8080	Egress	Used for monitoring configuration management, and other management and control operations

Region	IP	Direction	Description
	100.100.0.2:80	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific SE 1 (Singapore) ap-southeast-1	100.100.30.20:3128	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.103.7:80	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific SE 2 (Sydney) ap-southeast-2	100.100.80.92:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.13:80 [47.91.39.6:443]	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific SE 3 (Kuala Lumpur) ap-southeast-3	100.100.80.153:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.140:80	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific SE 5 (Jakarta) ap-southeast-5	100.100.80.160:8080	Egress	Used for monitoring configuration management, and other management and control operations

Region	IP	Direction	Description
	100.100.80.180:80	Egress	Used to collect monitoring data to CloudMonitor
Middle East 1 (Dubai) me-east-1	100.100.80.142:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.151:80 [47.91.99.5:443]	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific NE 1 (Tokyo) ap-northeast-1	100.100.80.184:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.137:80 [47.91.8.7:443]	Egress	Used to collect monitoring data to CloudMonitor
Asia Pacific SOU 1 (Mumbai) ap-south-1	100.100.80.152:8080	Egress	Used for monitoring configuration management, and other management and control operations
	100.100.80.66:80	Egress	Used to collect monitoring data to CloudMonitor

Resource usage

- Installation package size: 10–15 MB
- Memory: 10–15 MB, or 20 MB if you include shared space. Actual numbers vary depending on the size of your system memory.
- CPU: 1–2%

- Network: intranet. No Internet bandwidth is used.

Requirements on systems

Operating system	Hardware architecture	Note
Windows 7, Windows Server 2008 R2, or later versions	amd64, 386	None
Linux 2.6.23 or later with glibc	amd64, 386	CentOS/RHEL 5.x are not supported.

FAQ

- Where are CloudMonitor logs saved?
 - Linux: /usr/local/cloudmonitor/logs
 - Windows: C:\Program Files\Alibaba\cloudmonitor\logs

2.10 Agent release notes

1.2.11

Feature optimization and bug fixes. If you are using the local health check function, you will need to upgrade to this agent version.

New features

- Added local and remote protocol detection with support for Telnet and HTTP protocols

Feature optimization and bug fixes

- Fixed the privilege escalation loophole that may have occurred when the tmp directory was used as the temporary download directory of the installation script.
- Fixed the bug of submitting identical device data when the same disk is attached multiple times.
- Fixed the bug that some processes cannot obtain the path and name is fixed.
- Optimized the file download method to prevent monitoring process blocking that may have resulted from the downloading process.

1.1.64

Feature optimization and bug fixes. We recommend those using agent versions after CentOS 7.2 to upgrade to this version.

- Optimized memory usage calculation accuracy by adjusting the memory usage collection logic with MemAvailable filed used for available memory estimation for agent versions after CentOS 7.2.

1.1.63

Feature optimization and bug fixes

- Adjusted the default wrapper log to the info level.
- Optimized fault detection with added error-level log information added.
- Fixed the risk of memory leakage that may have resulted from logs at the debug level.

1.1.62

Feature optimization and bug fixes

- Optimized the HTTP Proxy selection logic to improve the agent installation success rate.
- Added key logs for improved fault detection.

1.1.61

Feature optimization and bug fixes

- Fixed the incorrect collection of topN processes bug that resulted from the abnormal collection of process user names by some systems.

1.1.59

Feature optimization and bug fixes

- Optimized the process count collection method to improve performance.
- Adjusted process monitoring so that two CloudMonitor agent processes are excluded from process count collection.

3 Site Monitoring

4 Alarm service

4.1 Alarm service overview

You can set alarm rules for metrics in host monitoring, instances in cloud service monitoring, and metrics in custom monitoring. Alarm rules can be applied to all resources, to application groups, or to a single instance.

The alarm service supports alarm notifications through various channels such as phone calls, SMS messages, emails, TradeManager, and DingTalk chatbots.

TradeManager only supports alarm notifications through PC clients. You can also install the Alibaba Cloud app to receive alarm notifications in this method.

Host monitoring alarm rules

Alarm rules can be set for all metrics in host monitoring. Alarm detection frequency can be set to a minimum of once per minute.

Cloud service alarm rule

CloudMonitor allows you to set threshold alarms to monitor the consumption of your cloud resources, and set event alarms to monitor the status of instances and services.

Custom monitoring alarm rules

After reporting monitoring data through the custom monitoring API, you can set alarm rules for corresponding metrics. Then, when the value of a metric exceeds the specified threshold, an alarm is triggered and an alarm notification is sent through the specified notification method.

Custom event alarm rules

After reporting event exceptions through custom event API, you can set alarm rules for the events. Then, when an alarm rule is met, an alarm is triggered and an alarm notification is sent with the specified notification method.

4.2 Manage alarm rules

The alarm service provides powerful capabilities to monitor alarms so that you can easily detect metric exceptions and quickly troubleshoot faults.

Parameter description

- **Products:** ECS, RDS, OSS, among others
- **Resource Range:** The range for which an alarm rule takes effect. There are three alarm rule ranges available: All Resources, Application Group, and Instances. When you set Resource Range to All Resources, you can report an alarm for up to 1,000 resources. If the number of your resources exceeds 1,000, alarms cannot be reported for some resources even if they exceed the threshold set in your alarm rules. Therefore, we recommend that you use application groups to divide resources by service before setting up alarm rules to avoid this issue.
 - **All Resources:** Indicates that the specified alarm rule applies to all instances under a user name. For example, if you set the resource range to all resources, and set the alarm threshold for MongoDB CPU usage to 80%, then an alarm is triggered when the CPU usage of any MongoDB instance exceeds 80%.
 - **Application Group:** Indicates that the specified rule applies to all instances under an application group. For example, if you set the resource range to application group and set the alarm threshold for host CPU usage to 80%, then an alarm is triggered when the CPU usage of a host instance exceeds 80%.
 - **Instances:** Indicates that the specified rule only applies to a specific instance. For example, if you set the resource range to instances and set the alarm threshold for host CPU usage to 80%, an alarm is triggered when the CPU usage of the specified instance exceeds 80%.
- **Alarm Rule:** The alarm rule name.
- **Rule Describe:** The main content of the alarm rule where you define the alarm-triggering condition, or value threshold, for related metrics. For example, if you describe the rule as 1-minute average CPU usage $\geq 90\%$, the alarm service will

check every minute whether the average value of metrics within one minute meets or exceeds 90%.

Consider the following example. For the alarm service in host monitoring, a single server metric item reports one data point in 15 seconds, and 20 data points in five minutes. This relates to the following alarm rules.

- 5-minute average CPU usage > 90%: Indicates that the average CPU usage value of the 20 data points for five minutes exceeds 90%.
- 5-minute CPU usage always > 90%: Indicates that the CPU usage values of the 20 data points for five minutes all exceed 90%.
- 5-minute CPU usage once > 90%: Indicates that the CPU usage value of at least one of the 20 data points for five minutes exceeds 90%.
- Total 5-minute Internet outbound traffic > 50 MB: Indicates that the sum of the outbound traffic values of the 20 data points for five minutes exceeds 50 MB.
- Triggered when threshold is exceeded for: An alarm notification is sent if the detected values reach the alarm rule threshold multiple times in a row.
- Effective Period: the period of time for which an alarm rule is valid. The alarm service checks metrics and determines whether to generate an alarm only during this period of time.
- Alarm Contact: a group of contacts who receive alarm notifications.
- Notification Methods: Different notification methods are available based on different alarm levels. Three alarm levels are available: Critical, Warning, and Info.
 - Critical: voice calls, SMS messages, emails, and DingTalk chatbot
 - Warning: SMS messages, emails, and DingTalk chatbot
 - Info: emails and DingTalk chatbot
- Email Remark: supplementary information customized for an alarm email. Remarks are sent as part of the alarm notification email.

Manage alarm rules

CloudMonitor provides three alarm rule management portals: the application group page, metric list page, and alarm rule list page of the alarm service.

- For alarm rule management in application groups, see [Manage alarm rules](#).
- For alarm rule management in host monitoring, see [Manage alarm rules](#).
- For setting alarm rules in custom monitoring, see [Set alarm rules](#).

- You can also set alarm rules in cloud service monitoring.

4.3 Manage alarm contacts and alarm contact groups

Alarm notifications are sent to alarm contacts and alarm contact groups. When creating an alarm rule, you will need to create an alarm contact and an alarm contact group so that you can select the contact and contact group to receive alarm notifications.

Manage an alarm contact

You can create, edit, or delete contact information, such as an email address.

- Create an alarm contact

1. Log on to the [CloudMonitor console](#).
2. In the left navigation pane, click Alarm contacts under Alarms. The Alarm Contact Management page is displayed.
3. Click Create Alarm Contact in the upper-right corner of the page. In the displayed dialog box, enter the contact email address and other information.

The specified email address needs to be verified so that you can avoid entering incorrect information that may cause you to not receive alarm notifications.

- Edit an alarm contact

1. Log on to the [CloudMonitor console](#).
2. In the left navigation pane, click Alarm contacts under Alarms. The Alarm Contact Management page is displayed.
3. Click Edit in the Actions column to edit the contact information.

- Delete an alarm contact

1. Log on to the [CloudMonitor console](#).
2. In the left navigation pane, click Alarm contacts under Alarms. Alarm Contact Management page is displayed.
3. Click Delete in the Actions column.



Note:

Once you delete an alarm contact, CloudMonitor alarm notifications are not longer sent to that contact.

Manage an alarm contact group

An alarm contact group may contain one or more alarm contacts. The same alarm contact can be added to multiple alarm contact groups. , When setting alarm rules, all alarm notifications need to be sent through an alarm contact group.

- Create an alarm contact group

1. Log on to the [CloudMonitor console](#).
2. In the left navigation pane, click Alarm contacts under Alarms. The Alarm Contact Management page is displayed.
3. Click the Alarm Contact Group tab at the top of the page to switch to the alarm contact group list.
4. Click Create Alarm Contact Group in the upper-right corner to open the Create Alarm Contact dialog box.
5. Enter a group name and select the contacts you want to add to the group.

- Edit an alarm contact group

1. Log on to the [CloudMonitor console](#).
2. In the left navigation pane, click Alarm contacts under Alarms. The Alarm Contact Management page is displayed.
3. Click the Alarm Contact Group tab at the top of the page to switch to the alarm contact group list.
4. Click Edit in the Actions column to edit the contact group information.

- Delete an alarm contact group

1. Log on to the [CloudMonitor console](#).
2. In the left navigation pane, click Alarm contacts under Alarms. The Alarm Contact Management page is displayed.
3. Click the Alarm Contact Group tab at the top of the page to switch to the alarm contact group list.
4. Click Delete in the Actions column to delete the contact group.

- Add contacts to a contact group in batches
 1. Log on to the [CloudMonitor console](#).
 2. In the left navigation pane, click Alarm contacts under Alarms. The Alarm Contact Management page is displayed.
 3. Select the contacts that you want to add from the alarm contact list.
 4. Click Add to a contact group at the bottom of the page.
 5. In the displayed dialog box, select the target contact group and click OK.

4.4 Create an alarm callback

The alarm callback feature allows you to integrate the alarm notifications sent by CloudMonitor into existing O&M or notification systems. CloudMonitor pushes alarm notifications to a specified public URL by using the POST request of the HTTP protocol. When you receive an alarm notification, you can take actions based on the notification received.



Note:

The alarm callback retry policy allows for up to three retry attempts with a timeout duration of 5 seconds.

Create an alarm callback

1. Log on to the [CloudMonitor console](#).
2. Select the alarm rule to which you want to add a callback.
3. Enter the URL address to call back in the notification method.

Callback parameters

When an alarm rule calls back a URL address, the pushed POST request is as follows:

Parameter	Data type	Description
userId	string	User ID
alertName	string	Alarm name
timestamp	string	The time stamp for when the alarm is generated

Parameter	Data type	Description
alertState	string	Alarm status. Based on your situation, one of three statuses is returned: OK , ALERT, and INSUFFICIENT_DATA.
dimensions	string	The object that triggered an alarm. For example: <code>[{"userId": "12345", "instanceId": "i-12345"}]</code>
expression	string	Alarm conditions. For example: <code>[{"expression": "\$value>12", "level": 4, "times": 2}]</code> means that an alarm was triggered after the threshold of 12 was exceeded for two times consecutively within the effective period. In these sorts of statements, Level =4 means that the alarms were sent by email, and level=3 means that the alarms were sent by email and SMS. "Times" indicates the number of times the threshold was exceeded consecutively within the effective period.
curValue	string	Current value. The metric value when an alarm is triggered or cleared.
metricName	string	Metric name.
metricProject	string	Product name. For more information about metrics and products, see Preset metric reference .

The following is an example of a POST request.

```
{
  "userId": "12345",
```

```
"alertName":"putNewAlarm_group_a37cd898-ea6b-4b7b-a8a8-de017a8327f6",
"timestamp":"1508136760",
"alertState":"ALARM",
"dimensions":[
  {
    "userId":"12345",
    "instanceId":"i-12345"
  }
],
"expression":["{\"expression\":\"\\\"$Average>90\\\"\",\\\"level\\\":4,\\\"times\\\":2}\""],
"Curvalue": "95 ",
"metricName":"CPUUtilization",
"metricProject":"acs_ecs_dashboard"
}
```

4.5 Use alarm templates

This topic describes how to simplify the creation and management of alarm rules by using alarm templates.

Scenarios

If you have multiple cloud resources (such as ECS instances, RDS services, SLB instances, and OSS buckets), we recommend that you use alarm templates to save alarm rules for these various resources. With having created alarm templates, you can directly apply the templates when creating alarm rules. This process can help you to simplify the creation and management of alarm rules, improving your overall O&M efficiency.

By default, CloudMonitor provides an initialized alarm template that contains common metrics for products such as ECS, RDS, SLB, and OSS, so that you can quickly and easily start to use alarm templates.

Before you begin

Alarm templates are used in combination with application groups. Therefore, we recommend that you create application groups for your resources before you use alarm templates in the creation of related alarm rules. For more information about how to create application groups, see [Create application groups](#).

Create an alarm template



Note:

- Alarm templates can be applied only to application groups.

- Each Alibaba Cloud account can contain up to 100 alarm templates.
- Each alarm template can contain up to 30 metrics.
- The alarm template function is only a shortcut to create multiple alarm rules. Alarm rules are not bound to alarm templates. After an alarm template is modified, alarm rules generated by using this template will remain unchanged. To modify the alarm rules for different application groups in batches, you must apply the modified template to each application group.

Procedure

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, choose Alarms > Alarm Templates.
3. Click Create Alarm Template to go to the Create Alarm Template page.

Create Alarm Template

Basic Information

• Template Name

The name must be within 30 characters and can contain numbe

Description

Up to 64 characters is allowed.

Rule

Rules such as heartbeat alarm in alarm template have been migrated to event monitoring. [Introduction to Cloud Products Events](#)

ECS

Rule Name	Rule Description	Resource Description
+Add Rules		

Products

Add Cancel

4. Enter a Template Name and Description in the Basic Information area.
5. Set an alarm rule. To add more alarm rules, click Add Rules.
6. Click Add.

Use an alarm template

- Use an alarm template when you create an application group

When you create an application group for your resources, you can select an existing alarm template in the MonitorAlarm area. After you have successfully created the application group, CloudMonitor generates alarm rules for this group based on the selected alarm template.

- **Apply an alarm template directly to an existing application group**

If you have created an application group but have not created alarm rules for the group, you can create an alarm template and then quickly apply the template to the group.

5 Availability monitoring

5.1 Manage availability monitoring

Availability monitoring conducts periodical detection tasks to check whether specified local or remote paths or ports respond properly and sends alarm notifications if response timeouts occur or status codes indicate errors based on the conditions specified in your alarm rules. This function can help you to quickly learn if local or remote services are unresponsive or abnormal, improving overall O&M and management efficiency.



Note:

- The CloudMonitor agent must be installed before you can use the availability monitoring function. Check that you have installed the CloudMonitor agent on your specified instances before using this function.
- Once working, monitoring tasks are performed once a minute.

Create availability monitoring tasks

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups.
3. Find the target application group and click the group name.
4. In the left-side navigation pane, select Availability Monitoring.
5. In the upper-right corner of the page, click Create Configuration to open the Create Availability Monitoring page.
6. Enter the task name and select the target server. You can select all servers in the application group to configure the same availability monitoring rules for them, or select some servers in the application group.
7. Select the detection type and the detection target: URL or IP address, ApsaraDB for RDS, and ApsaraDB for Redis are supported.
 - If you selecting ApsaraDB for RDS or ApsaraDB for Redis, relevant instances in the application group and access addresses are displayed.
 - If you select HTTP(S) for the Detection Target, you can configure matching content for HEAD, GET, and POST requests and return values.

8. Set the alarm rules. Status code and response time rules are supported for alarms. Any configuration that meets a condition specified in an alarm rule will trigger an alarm. An triggered alarm is sent as a notification to the alarm contact group that is associated to the application group.

- **Status code alarm:** An alarm that is triggered when the probe status code meets the specified alarm rules.
- **Notification method:** The means by which alarm notifications are sent, such as email or SMS message.
- **Advanced configuration:** Both effective and mute period configurations are supported. Effective period refers to a period in which an alarm rule is effective with alarms possibly triggered in the case that the conditions specified in your alarm rules are met. Mute period refers to a period in which your alarm rules are muted so that alarm notifications will be silenced even if conditions specified in your alarm rules are met.

Viewing availability monitoring tasks

1. Log on to the [CloudMonitor console](#).
2. Click Application Groups in the left-hand navigation bar to go to the application groups page.
3. Select the application groups for which you want to view availability monitoring, then click the application group name to enter the application group details page.
4. Select Availability Monitoring from the left-side navigation pane to go to the Availability Monitoring page. A list displaying the tasks that apply all availability monitoring in the group is displayed.

View monitoring results

1. Log on to the [CloudMonitor console](#).
2. Click Application Groups in the left-hand navigation bar to go to the Application Groups page.
3. Select the Application Groups for which you want to view availability monitoring, then click the application group name to enter the application groups details page.
4. Select Availability Monitoring from the left-side navigation pane to go to the Availability Monitoring page.

5. You can view monitoring results in the list.

- When the task probe does not trigger an alarm, the number of faulty instances in the list is 0.
- When an alarm is triggered for a probe exception, the number of instances that triggered an alarm is displayed in the list, click exception numbers to view the faulty instance details.
- Exception details.

Modify availability monitoring tasks

1. Log on to the [CloudMonitor console](#).
2. Click Application Groups in the left-hand navigation bar to go to the Application Groups page.
3. Select the Application Groups that needs to modify the availability monitoring, click the application group name to go to the app grouping details page.
4. Select availability monitoring on the left-hand menu of the page to enter the management page for availability monitoring.
5. Select the task that needs to be modified, click Modify in the action to go to the modify application groups page.
6. Edit content on the modify application groups page and save the configuration.

View alarm logs

1. Log on to the [CloudMonitor console](#).
2. Click Application Groups in the left-hand navigation bar to go to the Application Groups page.
3. Select the application groups that needs to view the alarm logs, click the application group name to go to the application group details page.
4. Select Alarm Logs on the left-hand menu of the page, and go to the alarm logs page to view the alarm log details.

Enable or disable monitoring tasks

Enabling or disabling monitoring tasks is supported for local health checks. When a task is disabled, health checks are no longer performed and alarms are no longer triggered for the task. However, when a task is enabled, probing is re-started and alarms will be triggered when the conditions specified in alarm rule settings are met.

1. Log on to the [CloudMonitor console](#).

2. Click Application Groups in the left-hand navigation bar to go to the Application Groups page.
3. Select the application groups that needs to be enabled or disabled for availability monitoring, and click the application group name, enter the application group details page.
4. Select availability monitoring on the left-hand menu of the page to enter the task management page for availability monitoring.
5. Select the task that you want to enable or disable, and click enable or disable in the action to modify the task status.

5.2 Local service availability monitoring

This topic describes how to monitor the availability of local service processes and send alarm notifications if response timeouts occur or status codes indicate errors.



Note:

- The CloudMonitor agent must be installed before you can use the availability monitoring function. Check that you have installed the CloudMonitor agent on your specified instances before using this function.
- Monitoring tasks are performed once a minute.
- Before using the availability monitoring function, you must create an application group. For more information, see [Create application groups](#).

Procedure

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups.
3. Find the target application group, and click the group name.
4. In the left-side navigation pane, click Availability Monitoring.

5. Click Create Configuration in the upper-right corner to go to the Create Availability Monitoring page.

Monitoring Configurations:

- **Target Server:** the machine that initiates the detection. The detection source and target for local service availability monitoring are the same machine.
- **Detection Type:** Select URL or IP address.
- **Detection Target:** If you select HTTP(S), the syntax is `localhost:port/path`. If you select TELNET, the syntax is `127.0.0.1:port`. Both are useful for different applications. If you want to detect whether Tomcat is responding properly, select HTTP(S) and enter `localhost:8080/monitor`. If you want to detect the connectivity of MySQL, select TELNET and enter `127.0.0.1:3306`.

Alarm Configuration:

Both Status Code and Response Time are used as the metrics of availability monitoring. An alarm is triggered when either metric value reaches the specified threshold. An alarm notification is sent to the alarm contact group of the corresponding application group. For local availability monitoring, set the status code greater than 400.

- **Status Code:** An alarm is triggered when the returned status code meets the alarm rule.
 - **Notification Method:** the method by which alarm notifications are sent.
 - **Advanced Configuration:**
 - **Muted For:** a period in which your alarm rules are muted so that alarm notifications will be silenced even if conditions specified in your alarm rules are met.
 - **Effective From:** a period in which an alarm rule is effective with alarms possibly triggered in the case that the conditions specified in your alarm rules are met.
6. Click OK to save your settings. When your service does not respond, an alarm notification is sent with the method you specified, such as SMS messages and emails.
 7. To view the details of unhealthy hosts, click the number of Unhealthy Hosts in the availability monitoring task list.

5.3 Status codes

The following is a list of the custom status codes returned whenever an exception is detected after an availability check is completed.

Protocol type	Status code	Definition
HTTP	610	Timeout due to no response within 5 seconds after the HTTP request was issued.
HTTP	611	The detection failed.
Telnet	630	Timeout due to no response within 5 seconds.
Telnet	631	The detection failed.

6 Cloud service monitoring

6.1 RDS monitoring

CloudMonitor displays the ApsaraDB for RDS operation status based on four metrics: Disk usage, IOPS usage, Connection usage, and CPU usage. Once you buy ApsaraDB for RDS products, CloudMonitor will automatically start monitoring the above four metrics without any additional operations.



Note:

- RDS only provides monitoring and alarm services for primary and read-only instances.
- By default, CloudMonitor will create alarm rules for each primary instance and read-only instance. These rules set up the thresholds of CPU usage, connection usage, IOPS usage, and disk usage all to 80%. When metric data exceeds any of the above thresholds, a text message and email will be sent to the contact person for the Alibaba Cloud account.

Monitoring service

- Metric description

Metric	Meaning	Dimension	Unit	Minimum monitoring granularity
Disk usage	The percentage of disk space used by the RDS instance.	Instance	Percentage	5 minutes
IOPS usage	The percentage of I/O requests per second used by the RDS instance.	Instance	Percentage	5 minutes

Metric	Meaning	Dimension	Unit	Minimum monitoring granularity
Connection usage	The connection count is the number of connections that application programs can establish with the RDS instance. Connection usage is the percentage of these connections currently in use.	Instance	Percentage	5 minutes
CPU usage	The percentage of CPU capacity consumed by the RDS instance (CPU performance is determined by the database memory size.).	Instance	Percentage	5 minutes
Memory usage	The percentage of the RDS instance's memory in use. Currently, the memory usage metric is only supported by MySQL databases.	Instance	Percentage	5 minutes

Metric	Meaning	Dimension	Unit	Minimum monitoring granularity
Read-only instance delay	MySQL read-only instance latency	Instance	seconds	5 minutes
Inbound traffic	Inbound traffic to an instance per second	Instance	bit/s	5 minutes
Outbound traffic	Outbound traffic from an instance per second	Instance	bit/s	5 minutes
Instance failure	Event-type metric, for which alarm rules can be set	-	-	-
Instance switchover	Event-type metric, for which alarm rules can be set.	-	-	-

The incoming and outgoing network traffic indicators only support MySQL and SQLServer databases.

- View monitoring data

1. Log on to the [CloudMonitor console](#).
2. Click Cloud Service Monitoring, and then click ApsaraDB for RDS.
3. Click an instance name in the product instance list or click Monitoring Charts in the Actions column to access the instance monitoring details page.
4. Click the Chart Size button to switch to large chart display (optional).

Alarm service

- Parameter description

- **Metric:** It is the monitoring indicator provided by RDS.
- **Statistical Cycle:** the alarm system checks if your monitoring data exceeds the alarm threshold for this cycle. For example, if the statistical cycle of the alarm rule for memory usage is set to one minute, the system checks whether the memory usage has exceeded the threshold value every other minute.
- **Statistic:** This sets the method used to determine if the data exceeds the threshold. You can set Average, Maximum, Minimum, and Sum in Statistics.
 - **Average:** The average value of metric data within a statistical cycle.. The statistical result is the average of all metric data collected within 15 minutes. An average value of over 80% is deemed to exceed the threshold.
 - **Maximum:** The maximum value of metric data within a statistical period. When the maximum value of the metric data collected within the period is over 80%, the value exceeds the threshold.
 - **Minimum:** The minimum value of metric data within a statistical period. When the minimum value of the metric data collected within the period is larger than 80%, the value exceeds the threshold.
 - **Sum:** The sum of metric data within a statistical period. When the sum of the metric data collected within the statistical cycle is over 80%, it exceeds the threshold. The above statistical methods are required for traffic-based indicators.
- **Alarm After Threshold Value Exceeded for Several Times:** Refers to an alarm which is triggered when the value of the metric item continuously exceeds the threshold value in several consecutive statistical cycles.

For example, you may set the alarm to go off when the CPU usage rate exceeds 80% within a 5-minute statistical cycle after the threshold value is exceeded for three times. Alarm notification will not be issued when. No alarm is triggered either when the CPU usage exceeds 80% again in the second detection five minutes later. The third probe still exceeds 80% Therefore, from the first time when the actual data exceeds the threshold to the time when the alarm rule is triggered, the minimum time consumed is: the statistical cycle*(the number of consecutive detections-1), which is $5*(3-1) = 10$ minutes in this case.

- Set alarm rules
 1. Log on to the [CloudMonitor console](#).
 2. Click Cloud Service Monitoring, and then click ApsaraDB for RDS .
 3. Click an instance name in the product instance list or click Alarm Rules in the Actions column to access the instance monitoring details page.
 4. Click Create Alarm Rules at the upper right of the alert policies page to create an alert policy based on the entered parameters.

6.2 SLB monitoring

By monitoring multiple metrics from Server Load Balancer (SLB), such as inbound and outbound traffic and the number of data packets and connections, CloudMonitor helps you to monitor the running status of instances and configure alarm rules accordingly. CloudMonitor automatically collects data from SLB from the time after you create an SLB instance.

Monitoring service

- Metrics
 - Layer-4 metrics

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Port inbound traffic	The traffic consumption for accessing a specified SLB port from the Internet	Port	bit/s	1 minute
Port outbound traffic	The traffic consumption for accessing the Internet from a specified SLB port	Port	bit/s	1 minute

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Number of inbound data packets by port	The number of the request packets (per second) that a specified SLB port receives	Port	count/s	1 minute
Number of outbound data packets by port	The number of the request packets (per second) that a specified SLB port sends	Port	count/s	1 minute
Number of new port connections	The average number of times (per second) the first SYN_SENT status occurs in a TCP three-way handshake during a statistical period	Port	count/s	1 minute
Number of active port connections	The number of connections in ESTABLISHED status during a statistical period	Port	count	1 minute

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Number of inactive port connections	The number of all TCP connections except the connections in ESTABLISHED status during a statistical period	Port	count	1 minute
Number of concurrent port connections	The total number of connections	Port	count	1 minute
Number of backend healthy ECS instances by port	The number of healthy instances reported by a health check	Port	count	1 minute
Number of backend unhealthy ECS instances by port	The number of unhealthy instances reported by a health check	Port	count	1 minute
Number of discarded port connections	The average number of connections discarded per second	Port	count/s	1 minute
Number of discarded inbound data packets by port	The average number of inbound packets discarded per second	Port	count/s	1 minute

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Number of discarded outbound data packets by port	The average number of outbound packets discarded per second	Port	count/s	1 minute
Number of discarded inbound bandwidth by port	The average inbound traffic discarded per second	Port	bit/s	1 minute
Number of discarded outbound bandwidth by port	The average outbound traffic discarded per second	Port	bit/s	1 minute
Number of active instance connections	The number of connections in ESTABLISHED status during a statistical period	Instance	count/s	1 minute
Number of inactive instance connections	The number of connections except those in ESTABLISHED status during a statistical period	Instance	count/s	1 minute
Number of discarded instance connections	The number of connections discarded per second	Instance	count/s	1 minute

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Number of discarded inbound data packets by instance	The number of inbound packets discarded per second	Instance	count/s	1 minute
Number of discarded outbound data packets by instance	The number of outbound packets discarded per second	Instance	count/s	1 minute
Discarded inbound bandwidth by instance	The amount of inbound traffic discarded per second	Instance	bit/s	1 minute
Discarded outbound bandwidth by instance	The amount of outbound traffic discarded per second	Instance	bit/s	1 minute
Number of concurrent instance connections	The total number of connections of the instance (the sum of active and inactive connections)	Instance	count/s	1 minute

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Number of new instance connections	The average number of times (per second) the first SYN_SENT status occurs in a TCP three-way handshake during a statistical period	Instance	count/s	1 minute
Number of inbound data packets by instance	The number of request packets received per second	Instance	count/s	1 minute
Number of outbound data packets by instance	The number of packets sent per second	Instance	count/s	1 minute
Number of inbound bandwidth by instance	The traffic consumption for accessing the SLB instance from the Internet	Instance	bit/s	1 minute

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Number of outbound bandwidth by instance	The traffic consumption for accessing the Internet from the SLB instance	Instance	bit/s	1 minute

- Layer-7 metrics

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Port QPS	The QPS of a specified port	Port	count/s	1 minute
Port RT	The average request latency of a specified port	Port	ms	1 minute
Status codes of the format 2xx	The number of status codes of the format 2xx that SLB returns to the client	Port	count/s	1 minute
Status codes of the format 3xx	The number of status codes of the format 3xx that SLB returns to the client	Port	count/s	1 minute
Status codes of the format 4xx	The number of status codes of the format 4xx that SLB returns to the client	Port	count/s	1 minute

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Status codes of the format 5xx	The number of status codes of the format 5xx that SLB returns to the client	Port	count/s	1 minute
Other status codes	The number of other status codes that SLB returns to the client	Port	count/s	1 minute
Upstream status codes of the format 4xx	The number of status codes of the format 4xx that RS returns to SLB	Port	count/s	1 minute
Upstream status codes of the format 5xx	The number of status codes of the format 5xx that RS returns to the client	Port	count/s	1 minute
Upstream RT	The average request delay from RS to proxy	Port	ms	1 minute
Instance QPS	The QPS of an instance	Instance	count/s	1 minute
Instance RT	The average request latency of the instance	Instance	count/s	1 minute

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Status codes of the format 2xx	The number of status codes of the format 2xx that SLB returns to the client	Instance	count/s	1 minute
Status codes of the format 3xx	The number of status codes of the format 3xx that SLB returns to the client	Instance	count/s	1 minute
Status codes of the format 4xx	The number of status codes of the format 4xx that SLB returns to the client	Instance	count/s	1 minute
Status codes of the format 5xx	The number of status codes of the format 5xx that SLB returns to the client	Instance	count/s	1 minute
Other status codes	The number of status codes of other formats that SLB returns to the client	Instance	count/s	1 minute
Upstream status codes of the format 4xx	The number of status codes of the format 4xx that RS returns to SLB	Instance	count/s	1 minute

Metric	Description	Dimension	Unit	Minimum monitoring frequency
Upstream status codes of the format 5xx	The number of status codes of the format 5xx that RS returns to SLB	Instance	count/s	1 minute
Upstream RT	The average request delay from RS to proxy	Instance	ms	1 minute

**Note:**

The numbers of new connections, active connections, and inactive connections are all based on TCP connection requests from the client to SLB.

· View monitoring data

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, choose Cloud Service Monitoring > Server Load Balancer to go to the SLB instance list page.
3. Select the region to which the target instance belongs.
4. Find the target instance and click the instance name or click Monitoring Charts in the Actions column.
5. On the Monitoring Charts tab page, you can view the monitoring data.
6. To switch to another chart view, click the chart view button in the upper-left corner of the page.

Alarm service

- Set an alarm rule

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, choose Cloud Service Monitoring > Server Load Balancer to go to the SLB instance list page.
3. Select the region to which the target instance belongs.
4. Find the target instance and click Alarm Rules in the Actions column.
5. On the Alarm Rules tab page, click Create Alarm Rules in the upper-right corner.
6. Set the parameters according to the following parameter descriptions and click Confirm.

- Parameters

- Products: ECS, RDS, OSS, among others
- Resource Range: the range for which an alarm rule takes effect. There are two alarm rule ranges available: All Resources and Instances.
 - All Resources: Indicates that the specified alarm rule applies to all SLB instances under your account. For example, if you set the resource range to all resources, and set the alarm threshold for CPU usage to 80%, then an alarm is triggered when the CPU usage of any SLB instance exceeds 80%. When you select All Resources, you can report alarms for up to 1,000 resources. If the number of your resources exceeds 1,000, alarms cannot be reported for some resources even if they exceed the threshold you set in your alarm rule. Therefore, for these scenarios, we recommend that you use application groups to divide resources by service before setting up alarm rules to avoid this issue.
 - Instances: Indicates that the specified rule only applies to a specific instance. For example, if you set the resource range to Instances and set the alarm threshold for CPU usage to 80%, an alarm is triggered when the CPU usage of the specified instance exceeds 80%.
- Alarm Rule: the alarm rule name
- Rule Describe: the main content of the alarm rule where you define the alarm-triggering conditions, or value threshold, for related metrics. For example, if you describe the rule as 5-minute average CPU usage $\geq 90\%$, the alarm service

will check every five minutes whether the average value of CPU usage within five minutes meets or exceeds 90%.

Consider the following example. For the alarm service in host monitoring, one data point is reported in 15 seconds for a single server metric, and 20 data points in five minutes. This relates to the following alarm rules.

- **5-minute average CPU usage > 90%:** The average CPU usage value of the 20 data points in five minutes exceeds 90%.
- **5-minute CPU usage always > 90%:** The CPU usage values of the 20 data points in five minutes all exceed 90%.
- **5-minute CPU usage once > 90%:** The CPU usage value of at least one of the 20 data points in five minutes exceeds 90%.
- **Total 5-minute Internet outbound traffic > 50 MB:** The sum of the outbound traffic values of the 20 data points in five minutes exceeds 50 MB.
- **Muted For:** the period of time that an alarm has been muted so that alarm contacts do not receive any alarm notifications during this period. An alarm rule can be muted up to 24 hours (or 1 day).
- **Triggered when threshold is exceeded for:** An alarm notification is sent if the detected values reach the alarm rule threshold a certain number of times consecutively.
- **Effective Period:** the period of time for which an alarm rule is effective. During this period of time, the alarm service checks metric data and determines whether to generate an alarm.
- **Notification Contact:** a group of contacts who receive alarm notifications.
- **Notification Methods:**
 - **Email and DingTalk chatbot**
 - **Email Subject:** The email subject is set as the product name, metric, and instance ID involved in the alarm by default.
 - **Email Remark:** supplementary information customized for an alarm email. Remarks are sent as part of the alarm notification email.
 - **HTTP CallBack:** Enter a URL accessible through the Internet and CloudMonitor will push the alarm information to the address through a POST request. Currently, only HTTP is supported.

6.3 OSS monitoring

By monitoring the basic service, performance, and metering data of the Object Service Storage (OSS) service, CloudMonitor enables you to gain insights into the overall performance of the OSS service and set alarm rules accordingly. Specifically, this can help you better track requests, analyze usage, collect statistics on business trends, and quickly discover and diagnose system issues.

Monitoring service

- Metrics

The metrics used for monitoring OSS mainly include basic service, performance, and metering indicators. For more information, see [Monitoring indicators reference](#).



Note:

To maintain consistency with the billing policies, the collection and presentation of metering data have the following characteristics:

- Metering data is collected hourly, so that the metering data for your resources is aggregated to a single value each hour. This value represents the overall metering condition of the hour monitored.
- Metering data has an output delay of nearly 30 minutes.
- The metering data time refers to the start time of the relevant statistical period.
- The cutoff time of metering data is the end time of the last statistical period of the current month. If no metering data is produced in the current month, the metering data cutoff time is 00:00 on the first day of the current month.
- For presentation purposes, the maximum quantity of metering data is pushed. For more information about metering data, see [Usage Records](#).

Example

Assuming that you only use PutObject requests to upload data and perform this operation at an average of 10 times per minute. Then, in the hour between 08:00:00 and 09:00:00 on May 10, 2016, the metering result of your PUT requests is 600 times (10 × 60 minutes), the time of metering data is 08:00:00 on May 10, 2016, and the result will be generated at around 09:30:00 on May 10, 2016. If the result is the last data record since 00:00:00 on May 1, 2016, the metering data cutoff time for the current month is 09:00:00 on May 10, 2016. If in May 2016,

you have not produced any metering data, the metering data cutoff time will be 00:00:00 on May 1, 2016.

Alarm service



Note:

The names of OSS buckets are unique. Given this, after you delete a bucket, if you create another one with the same name as the deleted one, the monitoring rules and alarm rules that were previously set for the deleted bucket will also apply to the new bucket.

You can set alarm rules for several metrics in addition to the preceding metering and statistical indicators. You can also add these metrics to your monitoring list. Moreover, multiple alarm rules can be set for a single metric.

Instructions

- For more information about the alarm service, see [Alarm service overview](#).
- For more information about the alarm service for OSS monitoring, see [OSS alarm service user guide](#).

6.4 CDN monitoring

Cloud monitoring by monitoring CDN's QPS, BPS, byte hit ratio, and so on, helps users get domain name usage. After a user adds an accelerated domain name, cloud monitoring automatically begins to monitor it, you are logged in to CDN for cloud monitoring. You can view monitoring details on the page. You can also set alarm rules on monitoring items so that you receive alarm information when the data is abnormal.

Monitoring Services

- Monitoring item description

Monitoring items	Meaning	Dimensions	Unit	Minimum monitor Granularity
Number of visits per second	Total number of visits/time granularity within the time grain	Domain Name	Times	1 minute
Network bandwidth BPS	Maximum network traffic per unit of time	Domain Name	BPS	1 minute
Hit rate	The probability of hit cache for the number of bytes requested in the time grain , note "bytes = number of requests x traffic ", the byte hit ratio more directly feedback back-to-back traffic	Domain Name	Percentage	1 minute
Public network out of traffic	That is, CDN's public network downstream traffic.	Domain Name	Bytes	5 minutes
Return code 4xx	Percentage of HTTP return code 4xx as all return codes within the time grain	Domain Name	Percentage	1 minute

Monitoring items	Meaning	Dimensions	Unit	Minimum monitor Granularity
Return code 5xx share	Percentage of HTTP return code 5xx as all return codes within the time grain	Domain Name	Percentage	1 minute

- Viewing Monitoring Data
 1. Log on to the [CloudMonitor console](#).
 2. Enter the list of CDN instances that the cloud service monitors.
 3. Click the Instance name or the monitor chart in the operation to go to the monitor details page.
 4. Click the size chart toggle button to toggle the larger image display (optional).

Alarm service

- Parameter descriptions
 - Monitor: Monitoring metrics provided by CDN.
 - Statistical Cycle: the alarm system checks if your monitoring data exceeds the alarm threshold for this cycle. For example, to set a memory usage alarm rule,

the statistics cycle is 1 minute, an interval of 1 minute checks if memory usage exceeds the threshold.

- **Statistical Methods:** Statistical Methods refer to settings that exceed the threshold range. You can set the average, maximum, minimum, count, and value in a statistical method.
 - **Average:** the average of the monitored data during the statistical cycle. The statistic result is the average of all metric data collected within 15 minutes. An average value of over 80% is deemed to exceed the threshold.
 - **Maximum:** the maximum value of the monitor data during the statistics cycle. The maximum value of monitoring data collected during the statistical cycle exceeds the threshold of 80.
 - **Minimum:** the minimum value of the monitored data during the statistics cycle. The minimum value for monitoring data collected during the statistical cycle exceeds the threshold of 80.
 - **Value:** Sum of monitoring data during the statistics cycle. When the sum of the metric data collected within the period is over 80%, it exceeds the threshold. Such statistical methods are required for traffic-class metrics.
- **Alarm after several consecutive exceeds threshold:** refers to a number of consecutive statistical cycle monitoring items whose values continue to exceed the threshold to trigger an alarm.

Example: Set CPU usage to exceed 80% Alarm, with a statistical cycle of 5 minutes, alarm three consecutive times after exceeding the threshold, when the first time the detection CPU usage exceeds 80%, no alarm notification is issued. Second Probe in 5 minutes CPU usage exceeds 80%, and no alarm will be issued. The third probe still exceeds 80% Alarm notification will be issued only when. That is, from the first time the actual data exceeds the threshold to the final alarm rule, the minimum time required is the statistical cycle (number of consecutive probes-1) = 5 (3-1) = 10 minutes.

- Set alarm rules

- 1. Log in to the clLog on to the [CloudMonitor console](#).oud monitoring console.
- 2. Enter the list of CDN instances that the cloud service monitors.
- 3. Click the alarm rule in the instance List action to enter the alert Rule Page for the instance.
- 4. Click new alarm rule in the upper-right corner of the alarm Rule Page to create an alarm rule based on the parameter.

6.5 ApsaraDB for Memcache

CloudMonitor provides seven monitoring metrics for ApsaraDB for Memcache, including cache used and read hit rate, to help you monitor the status of the service instances. It also allows you to set alarm rules for these monitoring metrics. After you purchase the Memcache service, CloudMonitor automatically collects data for the previous monitoring metrics.

Monitoring service

- Descriptions of monitoring metrics

Monitoring metrics	Meaning	Dimension	Unit	Minimum monitoring granularity
Cache used	The amount of cache used	Instance	Bytes	1 minute
Read hit rate	The probability of reading key-values (KVs) successfully	Instance	Percentage	1 minute
QPS	Total times of reading KVs per second	Instance	Times	1 minute
Number of records	Total number of KVs in the current measurement period	Instance	KVs	1 minute

Monitoring metrics	Meaning	Dimension	Unit	Minimum monitoring granularity
Cache inbound bandwidth	Traffic generated by accessing the cache	Instance	Bit/s	1 minute
Cache outbound bandwidth	Traffic generated by reading the cache	Instance	Bit/s	1 minute
Eviction	Number of KVs evicted per second	Instance	KVs per second	1 minute

**Note:**

- Monitoring data is saved for up to 31 days.
- You can view metric data for up to 14 consecutive days.

- View monitoring data

1. Log on to the [CloudMonitor console](#).
2. From the Cloud Service Monitoring drop-down list, select ApsaraDB for Memcache.
3. Click an instance name or click Monitoring Charts in the Actions column to access the instance monitoring details page and view various metrics.
4. Click the Time Range quick selection button at the top of the page or use the specific selection function.
5. Click the Zoom In button in the upper-right corner of the monitoring chart to enlarge the chart.

Alarm service

CloudMonitor provides alarm services for all Memcache monitoring metrics. After setting an alarm rule for an important monitoring metric, you can receive an alarm notification once the monitoring data exceeds the set threshold value, so that you can handle the problem rapidly to avoid malfunction.

- **Parameter description**

- **Monitoring metrics:** the monitoring metrics provided by ECS for Redis.
- **Statistical cycle:** the alarm system checks whether your monitoring data has exceeded the alarm threshold based on the cycle. For example, if the statistical cycle of the alarm rule for memory usage is set to one minute, the system checks whether the memory usage has exceeded the threshold value every other minute.
- **Statistical method:** refers to the method used to determine if the data exceeds the threshold. The average value, maximum value, minimum value, and sum value can be set to the statistical method.
 - **Average value:** the average value of monitoring data within the statistical cycle. The statistic result is the average of all the monitoring data collected within 15 minutes. An average value over 80% is deemed to exceed the threshold.
 - **Maximum value:** the maximum value of monitoring data within the statistical cycle. When the maximum value of the monitoring data collected within the statistical cycle is over 80%, it exceeds the threshold.
 - **Minimum value:** the minimum value of monitoring data within the statistical cycle. When the minimum value of the monitoring data collected within the statistical cycle is over 80%, it exceeds the threshold.
 - **Sum value:** the sum of monitoring data within the statistical cycle. When the sum of the monitoring data collected within the statistical cycle is over 80%, it exceeds the threshold. This method is required for traffic metrics.
- **Consecutive times:** an alarm is triggered when the value of the monitoring metrics continuously exceeds the threshold value for the set consecutive cycles.

For example, you have set the alarm to go off when the CPU usage exceeds the threshold value of 80% for three consecutive 5-minute statistical cycles. That is to say, no alarm is triggered when the CPU usage is found to exceed 80% for the first time. No alarm is triggered either when the CPU usage exceeds 80% again in the second detection five minutes later. The alarm is triggered when the CPU usage exceeds 80% again in the third detection. Therefore, from the first time when the actual data exceeds the threshold to the time when the alarm rule is triggered, the minimum time consumed is: the statistical cycle x (the number of consecutive detections - 1), which is $5 \times (3 - 1) = 10$ minutes in this case.

- Set an alarm rule
 1. Log on to the [CloudMonitor console](#).
 2. From the Cloud Service Monitoring drop-down list, select ApsaraDB for Memcache.
 3. Click an instance name or click Monitoring Charts in the Actions column to access the instance monitoring details page.
 4. Click the bell icon in the upper-right corner of the monitoring chart to set an alarm rule for corresponding monitoring metrics of this instance.
- Set batch alarm rules
 1. Log on to the [CloudMonitor console](#).
 2. From the Cloud Service Monitoring drop-down list, select ApsaraDB for Memcache.
 3. Select the appropriate instance on the instance list page. Click Set Alarm Rules at the bottom of the page to add alarm rules in batches.

6.6 Global acceleration monitoring

CloudMonitor monitors multiple monitoring metrics, such as inbound and outbound network bandwidth of Global Acceleration. It helps you monitor the network usage of Global Acceleration and allows you to set alarm rules for the monitoring metrics.

- After you purchase the Global Acceleration service, CloudMonitor automatically collects data on the preceding monitoring metrics.

Monitoring service

- Metrics

metric	Dimension	Unit	Minimum monitor Granularity
Inbound bandwidth	User and instance	Bits/s	1 minute
outbound bandwidth	User and instance	Bits/s	1 minute
Inbound package	User and instance	pps	1 minute

metric	Dimension	Unit	Minimum monitor Granularity
outbound package	User and instance	pps	1 minute

**Note:**

- Monitoring data is saved for up to 31 days.
- You can view the monitoring data for up to 7 consecutive days.

- View monitoring data

1. Log on to the [CloudMonitor console](#).
2. Go to the Global Acceleration instance list under Cloud Service Monitoring .
3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page and view various metrics.
4. Click the Time Range quick selection button from the upper menu of the page or use the specific selection function. You can view the monitoring data for up to 7 consecutive days.
5. Click Zoom In in the upper-right corner of the monitoring chart to enlarge the chart.

Alarm service

- Parameter descriptions

- **Monitoring metrics:** The monitoring metrics provided by Global Acceleration service.
- **Statistical cycle:** The alarm system checks whether your monitoring data has exceeded the alarm threshold based on the cycle. For example, if the statistical cycle of the alarm rule for memory usage is set to one minute, the system checks whether the memory usage has exceeded the threshold value every other minute.
- **Consecutive times:** An alarm is triggered when the value of the monitoring metrics continuously exceeds the threshold value for the set consecutive cycles.

- Set an alarm rule
 1. Log on to the [CloudMonitor console](#).
 2. Go to the Global Acceleration instance list under Cloud Service Monitoring .
 3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page.
 4. Click bell icon in the upper-right corner of the monitoring chart or New Alarm Rule in the upper-right corner of the page to set an alarm rule for corresponding monitoring metrics of this instance.
- Set multiple alarm rules
 1. Log on to the [CloudMonitor console](#).
 2. Go to the Global Acceleration instance list under Cloud Service Monitoring .
 3. Select the appropriate instances on the instance list page. Click Set Alarm Rules to add multiple alarm rules.

6.7 High performance time series database hitsdb

CloudMonitor monitors multiple monitoring metrics, such as HiTSDB disk usage, the number of timelines, and the number of time points. It helps you monitor the network use of NAT Gateway and allows you to set alarm rules for the monitoring metrics. When you purchase hitsdb, cloud monitoring automatically collects data for the hitsdb monitor.

Monitoring service

- Monitoring items

Monitoring items	Dimensions	Unit	Minimum monitor Granularity
Disk usage	User and instance	%	20 seconds
Timeline quantity	User and instance	Count	20 seconds
Point in time the growth rate	User and instance	Count/Second	20 seconds



Note:

- Monitoring data is saved for up to 31 days.

- You can view the monitoring data for up to 14 consecutive days.

- View metric data

1. Log on to the [CloudMonitor console](#).
2. Go to the HiTSDB instance list under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Action column to access the instance monitoring details page and view various metrics.
4. Click a "Time Range" shortcut on the top of the page or use the specific selection function. Up to 14 consecutive days of metric data can be viewed.
5. Click the zoom button in the upper-right corner of the monitor MAP to view the monitor larger image.

Alarm service

- Description

- Monitor: the monitoring indicator provided by hitsdb's service.
- Statistical Cycle: the alarm system checks if your monitoring data exceeds the alarm threshold for this cycle. For example, if the statistical cycle of the alarm rule for memory usage is set to one minute, the system checks whether the memory usage has exceeded the threshold value every other minute.
- Consecutive times: an alarm is triggered when the value of the monitoring metrics continuously exceeds the threshold value for the set consecutive cycles.

- Set an alarm rule

1. Log on to the [CloudMonitor console](#).
2. Go to the HiTSDB instance list under Cloud Service Monitoring.
3. Click the Instance name or the monitor chart in the operation to go To the instance monitoring details page.
4. Click the Bell button in the upper-right corner of the monitoring chart or New Alarm Rule in the upper-right corner of the page to set an alarm rule for corresponding monitoring metrics of this instance.

- Set batch alarm rules
 1. Log on to the [CloudMonitor console](#).
 2. Go to the HiTSDB instance list under Cloud Service Monitoring.
 3. When the instance list page selects the desired instance, click set alarm rule below the page, you can add alarm rules in bulk.

6.8 VPN gateway

CloudMonitor monitors multiple monitoring metrics, such as inbound and outbound network bandwidth of VPN gateway. It helps you monitor the network usage of VPN gateway and allows you to set alarm rules for the monitoring metrics. After you purchase the VPN gateway service, CloudMonitor automatically collects data on the preceding monitoring metrics.

Monitoring service

- Monitoring metrics

CloudMonitor provides the following monitoring metrics:

Monitoring metrics	Dimensions	Unit	Minimum monitoring granularity
Inbound network bandwidth of a bandwidth package	User and instance	Bit/s	1 minute
Outbound network bandwidth of a bandwidth package	User and instance	Bit/s	1 minute
Incoming packet of a bandwidth package	User and instance	PPS	1 minute
Outgoing packet of a bandwidth package	User and instance	PPS	1 minute



Note:

- Monitoring data is saved for up to 31 days.
- You can view the monitoring data for up to 7 consecutive days.

- View monitoring data

1. Log on to the [CloudMonitor console](#).
2. Go to the VPN Gateway instance list under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page and view various metrics.
4. Click the Time Range quick selection button at the top of the page or use the specific selection function. You can view the monitoring data for up to 7 consecutive days.
5. Click the Zoom In button in the upper-right corner of the monitoring chart to enlarge the chart.

Alarm service

- Parameter description

- Monitoring metrics: the monitoring metrics provided by the VPN gateway service.
- Statistical cycle: the alarm system checks whether your monitoring data has exceeded the alarm threshold based on the cycle. For example, if the statistical cycle of the alarm rule for memory usage is set to one minute, the system checks whether the memory usage has exceeded the threshold value every other minute.
- Consecutive times: an alarm is triggered when the value of the monitoring metrics continuously exceeds the threshold value for the set consecutive cycles.

- Set an alarm rule

1. Log on to the [CloudMonitor console](#).
2. Go to the VPN Gateway instance list under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page.
4. Click the bell icon in the upper-right corner of the monitoring chart or New Alarm Rule in the upper-right corner of the page to set an alarm rule for corresponding monitoring metrics of this instance.

- Set alarm rules in batches
 1. Log on to the [CloudMonitor console](#).
 2. Go to the VPN Gateway instance list under Cloud Service Monitoring.
 3. Select the appropriate instance on the instance list page. Click Set Alarm Rules at the bottom of the page to add alarm rules in batches.

6.9 Elasticsearch monitoring

CloudMonitor enables the user to monitor the usage of Elasticsearch services by collecting monitoring metrics such as the cluster status of Elasticsearch, the cluster query QPS, and the cluster writing QPS. Users can also set alarm rules for monitoring metrics. After you purchase the Elasticsearch, CloudMonitor automatically collects data on the preceding monitoring metrics.

Monitoring service

- Monitoring metrics

Monitoring metrics	Dimension	Unit	Minimum monitoring granularity
Cluster status	Cluster		1 minute
Cluster query QPS	Cluster	Count/Second	1 minute
Cluster writing QPS	Cluster	Count/Second	1 minute
Node CPU usage	Node	%	1 minute
Node disk usage	Node	%	1 minute
Node heapmemory usage	Node	%	1 minute
Node: load_1m	Node		1 minute
Node FullGc times	Node	Count	1 minute
Node Exception times	Node	Count	1 minute

Monitoring metrics	Dimension	Unit	Minimum monitoring granularity
Cluster snapshot status	Cluster	-1 indicates that there is no snapshot; 0 indicates success ; 1 indicates in progress; 2 indicates failure	1 minute

**Note:**

- Monitoring data is saved for up to 31 days.
- You can view the monitoring data for up to 14 consecutive days.

- View monitoring data

1. Log on to the [CloudMonitor console](#).
2. Go to the Elasticsearch instance list under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page and view various metrics.
4. Click a Time Range shortcut on the top of the page or use the specific selection function. Up to 14 consecutive days of metric data can be viewed.
5. Click Zoom In in the upper-right corner of the monitoring chart to enlarge the chart.

Alarm service

- Parameter description

- **Monitoring metrics:** the monitoring metrics provided by the Elasticsearch service.
- **Statistical cycle:** the alarm system checks whether your monitoring data has exceeded the alarm threshold based on the cycle. For example, if the statistical cycle of the alarm rule for memory usage is set to one minute, the system checks whether the memory usage has exceeded the threshold value every other minute.
- **Consecutive times:** an alarm is triggered when the value of the monitoring metrics continuously exceeds the threshold value for the set consecutive cycles.

- Set an alarm rule
 1. Log in to the cloud monitoring console.
 2. Go to the Elasticsearch instance list under Cloud Service Monitoring.
 3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page.
 4. Click Bell icon in the upper-right corner of the monitoring chart or New Alarm Rule in the upper-right corner of the page to set an alarm rule for corresponding monitoring metrics of this instance.

6.10 Express Connect monitoring

CloudMonitor monitors multiple metrics, such as the inbound and outbound network traffic of the Express Connect instance. It monitors the network usage of the instance and allows you to set alarm rules for various metrics. Once you buy the Express Connect service, CloudMonitor automatically collects the data for the following metrics.

Monitoring services

- Monitoring items

Monitoring item	Dimension	Unit	Minimum monitoring granularity
Inbound network traffic	User and instance	Bytes	1 minute
Outbound network traffic	User and instance	Bytes	1 minute
Inbound network bandwidth	User and instance	Bits/s	1 minute
Outbound network bandwidth	User and instance	Bits/s	1 minute
Latency	User and instance	ms	1 minute
Packet loss rate	User and instance	%	1 minute



Note:

- Monitoring data is saved for up to 31 days.
- You can view the monitoring data for up to 14 consecutive days.

- View monitoring data

1. Log on to the [CloudMonitor console](#).
2. Enter the instance list of Express Connect under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Action column to access the instance monitoring details page and view metrics.
4. Click a Time Range shortcut on the top of the page or use the specific selection function. Up to 14 consecutive days of metric data can be viewed.
5. Click the zoom-in button in the upper-right corner of the monitoring chart to view a large image.

Alarm service

- Parameter description

- **Metrics:** metric items provided by the Express Connect service.
- **Statistical Cycle:** indicates how often the alarm system checks whether monitoring data exceeds the alarm threshold. For example, if the statistical cycle of the alarm rule for memory usage is set to one minute, the system checks whether the memory usage has exceeded the threshold value every other minute.
- **Statistical Methods:** determines whether the data exceeds the threshold. Average, maximum, minimum, and sum can be set in the Statistical Methods.
 - **Average value:** the average value of monitoring data within the statistical cycle. For example, when the average value of all monitoring data collected within 15 minutes is used as the statistical method, an average value over 80 % is deemed to exceed the threshold.
 - **Maximum value:** the maximum value of monitoring data within the statistical cycle. For example, when the maximum value of all monitoring data collected within 15 minutes is used as the statistical method, a maximum value over 80 % is deemed to exceed the threshold.
 - **Minimum value:** the minimum value of monitoring data within the statistical cycle. For example, when the minimum value of all monitoring data collected

within 15 minutes is used as the statistical method, a minimum value over 80 % is deemed to exceed the threshold.

■ **Sum value:** the sum of monitoring data within the statistical cycle. For example, when the sum value of all monitoring data collected within 15 minutes is used as the statistical method, a sum value over 80% is deemed to exceed the threshold. This method is required for traffic metrics.

- **Consecutive times:** An alarm is triggered when the value of the monitoring metrics continuously exceeds the threshold value for the set consecutive cycles.

For example, you have set the alarm to go off when the CPU usage exceeds the threshold value of 80% for three consecutive 5-minute statistical cycles. That is to say, no alarm is triggered when the CPU usage is found to exceed 80% for the first time. No alarm is triggered either when the CPU usage exceeds 80% again in the second detection five minutes later. The alarm is triggered when the CPU usage exceeds 80% again in the third detection. Therefore, from the first time when the actual data exceeds the threshold to the time when the alarm rule is triggered, the minimum time consumed is: the statistical cycle x (the number of consecutive detections - 1), which is $5 \times (3 - 1) = 10$ minutes in this case.

- Set an alarm rule

1. Log on to [CloudMonitor console](#).
2. Enter the instance list of Express Connect under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Action column to access the instance monitoring details page and view metrics.
4. Click the Bell button in the upper-right corner of the monitoring chart or New Alarm Rule in the upper-right corner of the page to set an alarm rule for corresponding monitoring metrics of this instance.

- Set alarm rules in batches

1. Log on to [CloudMonitor console](#).
2. Enter the instance list of Express Connect under Cloud Service Monitoring.
3. Select instances on the instance list page. Click Set Alarm Rules to add multiple alarm rules.

6.11 StreamCompute

By monitoring service latency, failover rate, and read and write RPS metrics, CloudMonitor helps you gain insights into the overall performance of the StreamCompute services you are using and set alarm rules accordingly. CloudMonitor automatically collects data from StreamCompute from the time when you begin to use this product.

Monitoring service

- Metrics

Metric	Dimensions	Unit	Description	Minimum monitoring frequency
Service latency	Project, job	s	The data processing latency of the current job	1 minute
Read RPS	Project, job	read/s	The average number of data lines read per second for tasks	1 minute
Write RPS	Project, job	write/s	The average number of data lines written per second for tasks	1 minute
Failover rate	Project, job	%	The sum of failover frequency of current job	1 minute



Note:

- Monitoring data is saved for up to 31 days. You can view up to 14 consecutive days of monitoring data.

- View monitoring data

1. Log on to the [CloudMonitor console](#).
2. Go to the StreamCompute instance list under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page and view various metrics.
4. Click the Time Range quick selection button from the upper menu of the page or use the specific selection function. You can view monitoring data from up to 14 consecutive days.
5. Click Zoom In in the upper-right corner of the monitoring chart to enlarge the chart.

Alarm service

- Parameters

- **Metrics:** The monitoring metrics taken from StreamCompute.
- **Statistical cycle:** The recurring period of time in which the alarm system checks whether monitoring data has exceeded the alarm threshold. For example, if an alarm rule for memory usage has a statistical cycle set to one minute, then the system will check whether memory usage has exceeded the threshold value specified in your alarm rule every other minute.
- **Statistical method:** The calculation method and resulting value used to determine whether the data has exceeded the threshold specified in an alarm rule, which can be average, maximum, minimum, or sum. For example, in a statistical period of 15 minutes, an average of 80% of some metric such as memory usage can be specified as the threshold for an alarm to be triggered.
 - **Average value:** The average value of monitoring data within the statistical cycle. For example, when the average value of all monitoring data collected within 15 minutes is adopted as the statistical method, an average value over 80% is deemed to exceed the threshold.
 - **Maximum value:** The maximum value of monitoring data within the statistical cycle. For example, when the maximum value of all monitoring data collected within 15 minutes is adopted as the statistical method, a maximum value over 80% is deemed to exceed the threshold.
 - **Minimum value:** The minimum value of monitoring data within the statistical cycle. For example, when the minimum value of all monitoring data

collected within 15 minutes is adopted as the statistical method, a minimum value over 80% is deemed to exceed the threshold.

■ **Sum value:** The sum of monitoring data within the statistical cycle. For example, when the sum value of all monitoring data collected within 15 minutes is adopted as the statistical method, a sum value over 80% is deemed to exceed the threshold. This method is required for traffic metrics.

- **Consecutive times:** An alarm is triggered after a metric value continuously exceeds the threshold specified in an alarm rule for some set of consecutive cycles. For example, if the consecutive times is set to three, then the conditions specified for an alarm rule must be met for three consecutive statistical cycles before an alarm is triggered.

For example, you have set the alarm to go off when the CPU usage exceeds the threshold value of 80% for three consecutive 5-minute statistical cycles. That is to say, no alarm is triggered when the CPU usage is found to exceed 80% for the first time. No alarm is triggered either when the CPU usage exceeds 80% again in the second detection five minutes later. The alarm is triggered when the CPU usage exceeds 80% again in the third detection. Therefore, from the first time when the actual data exceeds the threshold to the time when the alarm rule is triggered, the minimum time consumed is: the statistical cycle(the number of consecutive detections-1), which is $5(3-1) = 10$ minutes in this case.

- Set an alarm rule

1. Log on to the [CloudMonitor console](#).
2. Go to the StreamCompute instance list under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Actions column.
4. Click the bell icon in the upper-right corner of the monitoring chart or New Alarm Rule in the upper-right corner of the page to set an alarm rule for corresponding metrics of this instance.

- Set multiple alarm rules

1. Log on to the [CloudMonitor console](#).
2. Go to the StreamCompute instance list under Cloud Service Monitoring.
3. Select the appropriate instances on the instance list page. Click Set Alarm Rules to add multiple alarm rules.

6.12 ApsaraDB for HybridDB

Cloud monitoring through monitoring hybriddb's CPU usage, memory usage, and so on, helps the user monitor the usage of the hybriddb instance and enables the user to set alarm rules on the monitor item. After you purchase After hybriddb, cloud monitoring automatically collects data for the above monitoring items.

Monitor

- Monitoring items

Monitoring items	Dimension	Unit	Minimum monitor Granularity
Disk usage	User and instance	%	5 minutes
Connection usage	User and instance	%	5 minutes
CPU usage	User and instance	%	5 minutes
Memory usage	User and instance	%	5 minutes
I/O throughput usage	User and instance	%	5 minutes



Note:

- Monitor Data is saved for up to 31 days.
- Users can view monitoring data for up to 14 days in a row.

- View monitored data.

1. Log in to the [cloud monitoring console](#).
2. Go to the HybridDB instance list under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring information page and view various metrics.
4. Click the time range quick select button or the exact select function at the top of the page, maximum monitoring data support view continuous 14 Monitoring data for days.
5. Click the zoom in button in the upper-right corner of the monitor MAP to view the monitor larger image.

Alarm service

6.13 NAT gateway

By monitoring multiple metrics from NAT Gateway, including SNAT connections and bandwidth package data, CloudMonitor helps you understand the overall network usage and performance of the NAT Gateway services you are using and set alarm rules accordingly. CloudMonitor automatically collects data from NAT Gateway from the time you begin to use this product.

Monitoring Service

· Metrics

Metric	Dimensions	Units	Minimum monitoring frequency
SNAT connections	User and instance	count/minute	1 minute
Bandwidth packets (inbound bandwidth)	User and instance	bit/s	1 minute
Bandwidth packets (outbound bandwidth)	User and instance	bits/s	1 minute
Bandwidth packets (inbound packets)	User and instance	packet/s	1 minute
Bandwidth packets (outbound packets)	User and instance	packet/s	1 minute
Bandwidth packets (outbound bandwidth usage)	User and instance	%	1 minute



Note:

- Monitoring data is saved for up to 31 days.
- You can view up to 14 consecutive days of monitoring data.

- View monitoring data
 1. Log on to the [CloudMonitor console](#).
 2. Go to the NAT Gateway instance list under Cloud Service Monitoring.
 3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page and view various metrics.
 4. Click the Time Range quick selection button from the upper menu or use the specific selection function. You can view monitoring data from up to 14 consecutive days.
 5. Click Zoom In in the upper-right corner of the monitoring chart to enlarge the chart.

Alarm service

- Parameters
 - Metrics: The monitoring metrics taken from NAT Gateway.
 - Statistical cycle: The recurring period of time in which the alarm system checks whether monitoring data has exceeded the alarm threshold. For example, if an alarm rule for memory usage has a statistical cycle set to one minute, then the system will check whether memory usage has exceeded the threshold value specified in your alarm rule every other minute.
 - Consecutive times: An alarm is triggered after a metric value continuously exceeds the threshold specified in an alarm rule for some set of consecutive cycles. For example, if the consecutive times is set to three, then the conditions specified for an alarm rule must be met for three consecutive statistical cycles before an alarm is triggered.
- Set an alarm rule
 1. Log on to the [CloudMonitor console](#).
 2. Go to the NAT Gateway instance list under Cloud Service Monitoring.
 3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page.
 4. Click the bell icon in the upper-right corner of the monitoring chart or New Alarm Rule in the upper-right corner of the page to set an alarm rule for corresponding metrics of this instance.

- Set multiple alarm rules
 1. Log on to the [CloudMonitor console](#).
 2. Go to the NAT Gateway instance list under Cloud Service Monitoring.
 3. Select the appropriate instances on the instance list page. Click Set Alarm Rules to add multiple alarm rules.

6.14 Open Ad

By monitoring more than a dozen metrics taken from Open Ad, including RTB PV and QPS, and ad click PV, CloudMonitor helps you manage and interpret the real-time status of your Open Ad services and set alarm rules accordingly. CloudMonitor automatically collects data from Open Ad from the time when you begin to use this product.

Monitoring service

- Metrics

Metrics	Dimension	Unit	Minimum monitoring frequency
RTB PV	User	count	1 minute
RTB QPS	User	time/s	1 minute
Ad click PV	User	count	1 minute
Ad click QPS	User	time/s	1 minute
Ad click delay	User	ms	1 minute
Ad exposure PV	User	count	1 minute
Ad exposure QPS	User	time/s	1 minute
Ad exposure delay	User	ms	1 minute
DMP active crowd count	User	count/day	1 hour
DMP valid crowd requests	User	time/day	1 hour
Storage space utilized by DMP	User	byte/day	1 hour

Metrics	Dimension	Unit	Minimum monitoring frequency
League and dip effective crowd count	User	count/day	1 hour
Valid audience number in Umeng and DIP	User	time/day	1 hour

**Note:**

- Monitoring data is saved for up to 31 days.
- You can view up to 14 consecutive days of monitoring data.

- View monitoring data

1. Log on to the [CloudMonitor console](#).
2. Go to the Open Ad monitoring page under Cloud Service Monitoring, and view the monitoring data of the Open Ad service.

Alarm service

CloudMonitor provides alarm functions for Open Ad monitoring metrics, so that you can be notified immediately in the case of any metric exceptions.

Set alarm rules

- Method 1

1. Log on to the [CloudMonitor console](#).
2. Go to the Open Ad monitoring page under Cloud Service Monitoring.
3. Click the bell icon in the upper-right corner of the monitoring chart or New Alarm Rule in the upper-right corner of the page to set an alarm rule for corresponding metrics of this instance.

- Method 2

1. Log on to the [CloudMonitor console](#).
2. Go to the Open Ad monitoring page under Cloud Service Monitoring.
3. Click Alarm Rules to go to the Alarm Rules list page. Click Create Alarm Rules in the upper-right corner to create alarm rules.

6.15 ApsaraDB for PetaData

By monitoring several metrics, specifically disk usage, inbound and outbound bandwidth and QPS, CloudMonitor helps you understand the status of your instances in ApsaraDB for PetaData scaling groups and set alarm rules accordingly. CloudMonitor automatically collects data from ApsaraDB for PetaData from the time when you begin to use this product.

Monitoring service

- Metrics

Metric	Dimensions	Unit	Minimum monitoring frequency
Disk usage	User and instance	byte	5 minutes
Inbound bandwidth	User and instance	byte/s	5 minutes
Outbound bandwidth	User and instance	byte/s	5 minutes
QPS	User and instance	count/s	5 minutes



Note:

- Monitoring data is saved for up to 31 days.
- You can view up to 14 consecutive days of monitoring data.

- Viewing monitoring data

1. Log on to the [CloudMonitor console](#).
2. Go to the ApsaraDB for PetaData instance list under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page and view various metrics.
4. Click the Time Range quick selection button from the upper menu of the page or use the specific selection function. You can view monitoring data from up to 14 consecutive days.
5. Click Zoom In in the upper-right corner of the monitoring chart to enlarge the chart.

Alarm service

- Parameters

- **Metrics:** The monitoring metrics taken from ApsaraDB for PetaData.
- **Statistical cycle:** The recurring period of time in which the alarm system checks whether monitoring data has exceeded the alarm threshold. For example, if an alarm rule for memory usage has a statistical cycle set to one minute, then the system will check whether memory usage has exceeded the threshold value specified in your alarm rule every other minute.
- **Statistical method:** The calculation method and resulting value used to determine whether the data has exceeded the threshold specified in an alarm rule, which can be average, maximum, minimum, or sum. For example, in a statistical period of 15 minutes, an average of 80% of some metric such as memory usage can be specified as the threshold for an alarm to be triggered.
 - **Average value:** The average value of monitoring data within the statistical cycle. For example, when the average value of all monitoring data collected within 15 minutes is adopted as the statistical method, an average value over 80% is deemed to exceed the threshold.
 - **Maximum value:** The maximum value of monitoring data within the statistical cycle. For example, when the maximum value of all monitoring data collected within 15 minutes is adopted as the statistical method, a maximum value over 80% is deemed to exceed the threshold.
 - **Minimum value:** The minimum value of monitoring data within the statistical cycle. For example, when the minimum value of all monitoring data collected within 15 minutes is adopted as the statistical method, a minimum value over 80% is deemed to exceed the threshold.
 - **Sum value:** the sum of monitoring data within the statistical cycle. For example, when the sum value of all monitoring data collected within 15 minutes is adopted as the statistical method, a sum value over 80% is deemed to exceed the threshold. This method is required for traffic metrics.
- **Consecutive times:** An alarm is triggered after a metric value continuously exceeds the threshold specified in an alarm rule for some set of consecutive cycles. For example, if the consecutive times is set to three, then the conditions

specified for an alarm rule must be met for three consecutive statistical cycles before an alarm is triggered.

For example, you have set the alarm to go off when the CPU usage exceeds the threshold value of 80% for three consecutive 5-minute statistical cycles. The second time in 5 minutes to detect CPU usage exceeds 80% and no alarm will be issued. The third probe still exceeds 80% Alarm notification will be issued only when. That is, from the first time the actual data exceeds the threshold to the final alarm rule, the minimum time required is statistical cycle X (number of consecutive probes-1) = 5 x (3-1) = 10 minutes.

- Set an alarm rule

1. Log on to the [CloudMonitor console](#).
2. Go to the ApsaraDB for PetaData instance list under Cloud Service Monitoring.
3. Click an instance name or click Monitoring Chart in the Actions column to access the instance monitoring details page.
4. Click the bell icon or New Alarm Rule in the upper-right corner of the monitoring data page to set an alarm rule for corresponding metrics of this instance.

- Set multiple alarm rules

1. Log on to the [CloudMonitor console](#).
2. Go to the ApsaraDB for PetaData instance list under Cloud Service Monitoring.
3. Select the appropriate instances on the instance list page. Click Set Alarm Rules to add multiple alarm rules.

7 RAM for CloudMonitor

[RAM](#) permissions are supported in CloudMonitor. Through the integration of the monitoring console with access control features, you can easily and quickly apply permissions for cloud service monitoring data, alarm rule management, alarm contact and alarm contact groups, and event subscription and related features.



Note:

RAM monitoring data queries are supported for the following cloud products:

- ECS
- RDS
- Server Load Balancer
- OSS
- CDN
- ApsaraDB for Memcache
- EIP
- ApsaraDB for Redis
- Message Service
- Log Service

Permissions

In RAM, if a user is authorized with read-only permissions for CloudMonitor, the user can only view relevant data, such as the monitoring data and alarm services, but cannot write data.

Authentication types

In addition to basic RAM account permission controls, time-based, multi-factor, and IP authentication are supported.

Resources

Fine-grained resource descriptions are not supported by RAM. The “*” wildcard is used for resource authorization.

Operation description

- Monitoring data

Data query actions are divided into two categories: Product instance lists and CloudMonitor metric data queries. When authorizing a RAM account to log on to the CloudMonitor portal and view metric data, you must also grant the account permissions for the corresponding product's instance list and metric data query.

The corresponding actions are listed in the following table.

Product	Action
CMS	QuerMetricList
CMS	QueryMetricLast
ECS	DescribeInstances
RDS	DescribeDBInstances
SLB	DescribeLoadBalancer*
OSS	ListBuckets
OCS	DescribeInstances
EIP	DescribeEipAddresses
Aliyun Cloud for Redis	DescribeInstances
Message Service	ListQueue
CDN	DescribeUserDomains

- Alarm service

The alarm service provides permission controls for alarm rule management, alarm contact and alarm contact group management, and event subscription and related features.

The query-related actions are listed in the following table.

Action	Description
QueryAlarm	Query an alarm rule
QueryAlarmHistory	Query an alarm history
QueryContactGroup	Query a contact group
QueryContact	Query a contact
QuerySms	Query the number of SMSs used

Action	Description
QueryMns	Querying an event subscription configuration

The management-related actions are listed in the following table.

Action	Description
UpdateAlarm	Modify an alarm rule
CreateAlarm	Create an alarm rule
DeleteAlarm	Delete an alarm rule
DisableAlarm	Disable an alarm rule
EnableAlarm	Enable an alarm rule
CreateContact	Create a contact
DeleteContact	Delete a contact
UpdateContact	Modify a contact
SendEmail	Send an email authentication code
SendSms	Send an SMS verification code
CheckEmail	Check an email verification code
CheckSms	Check an SMS verification code
CreateGroup	Create a contact group
DeleteGroup	Delete a contact group
UpdateGroup	Modify a contact group
CreateMns	Create an event subscription
DeleteMns	Delete an event subscription
UpdateMns	Modify an event subscription

8 Application groups

8.1 Application group overview

The application group feature of CloudMonitor allows you to group related resources and monitor these resources in a centralized manner. With application groups, you can easily monitor a group of target resources such as servers, databases, SLB instances, and storage, and apply alarm rules to the application group, thereby improving your overall O&M efficiency.



Note:

- A single account can create up to 100 application groups.
- Up to 1,000 resource instances can be added to one application group.

8.2 Create application groups

This topic describes how to group your cloud resources by creating application groups so that you can manage your resources and alarm rules on a grouped basis.

Scenarios

If you have purchased multiple products on Alibaba Cloud, you can group them together in a centralized manner by creating application groups. With application groups, you can manage resources of different regions and products, such as servers, databases, object storage, and cache, based on your business modules. In addition, you can easily manage alarm rules and view the monitoring data of these grouped resources.

Application group modes

Instances can be added to application groups using dynamic or static mode.

- **Dynamic mode:** When creating an application group, you can set name rules for instances so that instances which meet your name rules will be automatically added into the application group. If you want to add or remove instances to or from the group in the future, you only need to modify the instance names to complete these configurations. Currently, dynamic mode is supported only by ECS, ApsaraDB for RDS, and SLB instances.

- **Static mode:** With static mode, you need to manually add instances to an application group.

Create an application group



Note:

- Up to 1,000 resource instances can be added to each application group.
- Up to 100 application groups can be created under each account.

Procedure

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups.

3. In the upper-right corner of the displayed page, click Create Group.

Create Group

Basic Information

- Product Group Name

- Contact Group

[Quickly create](#)

MonitorAlarm

Select Template

[Go to Create Alarm](#)

Initialize Agent Installation



Event Monitor

- ☒ Subscribe Event notification

After subscription event notification, alarm notification will be sent when service is abnormal within the group. [Introduction to Cloud Products Events](#)

Add Instance dynamically

- ☒ Dynamic rules for ECS instances

- Dynamic rules

☒ All rules ☐ Any rule

instance created in future according with this rule would be added to group

[+Add Rules](#)

4. **Enter Basic Information:** Enter the group name and select one or more contact groups to receive alarm notifications.
5. **Set MonitorAlarm:** Select one or more templates to initialize alarm rules for the instances in the group (optional), and select the notification method. If you turn on the Initialize Agent Installation switch, the CloudMonitor agent will be installed on all servers in the group to collect monitoring data.
6. **Set Event Monitor:** If you select the Subscribe Event notification check box, alarm notifications will be sent when critical-level and warning-level events occur in related resources in the group.
7. **Set Add Instance dynamically.**
 - You can set name rules to automatically add ECS instances that match the name rules to the group. Specifically, instances, including future instances, whose names contain, start with, or end with the words you specify will be automatically added to the group. A maximum of three rules can be added, and the relationship among the rules can be AND or OR.
 - To add rules for ApsaraDB for RDS or SLB instances, click Add Product.
 - To add instances of other Alibaba Cloud products, you need to add them manually after creating the application group.
8. Click Create Application Group.

8.3 Manage alarm rules

You can create, view, modify, enable, disable, and delete threshold alarm rules in application groups.



Note:

When you view alarm rules of an application group, the system displays only the alarm rules applied to this application group. The alarm rules applied to the instances or resources in the group are not displayed.

Create an alarm rule

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups.
3. Find the target group and click the group name.
4. Click Threshold alarm in the upper-right corner.

5. Select the product type, add one or more alarm rules, set the alarm mechanism, select the contact group, and then click Add.

Create alarm rules by using an alarm template

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups.
3. Find the target group and click the group name.
4. In the upper-right corner of the displayed page, click Apply Template to Group.
5. Select the required alarm template and click OK.

Delete an alarm rule

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups.
3. Find the target application group and click the group name.
4. In the left-side navigation pane, click Alarm Rule.
5. Find the target alarm rule, and click Delete in the Actions column to delete this rule. To delete multiple rules at one time, select the rules to be deleted and click Delete under the alarm rule list.

Modify an alarm rule

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups.
3. Find the target application group and click the group name.
4. In the left-side navigation pane, click Alarm Rule.
5. Find the target alarm rule, and click Modify in the Actions column to modify this rule.

Disable or enable alarm rules

If you want to stop a service for application maintenance or upgrades, you can disable all alarm rules of the application group to avoid unnecessary alarm notifications. After the maintenance or upgrades are complete, you can enable the alarm rules.

- Disable all alarm rules of an application group
 1. Log on to the [CloudMonitor console](#).
 2. In the left-side navigation pane, click Application Groups.
 3. Find the target application group and click More in the Actions column.
 4. Select Disable All Alarm Rules.
- Enable all alarm rules of an application group
 1. Log on to the [CloudMonitor console](#).
 2. In the left-side navigation pane, click Application Groups.
 3. Find the target application group and click More in the Actions column.
 4. Select Enable All Alarm Rules.
- Disable some alarm rules of an application group
 1. Log on to the [CloudMonitor console](#).
 2. In the left-side navigation pane, click Application Groups.
 3. Find the target application group and click the group name.
 4. In the left-side navigation pane, click Alarm Rule.
 5. Find the target alarm rule, and click Disable in the Actions column to disable this rule. Repeat this step to disable other alarm rules, or select multiple rules and click Disable under the alarm list.
- Enable some alarm rules of an application group
 1. Log on to the [CloudMonitor console](#).
 2. In the left-side navigation pane, click Application Groups.
 3. Find the target application group and click the group name.
 4. In the left-side navigation pane, click Alarm Rule.
 5. Find the target alarm rule, and click Enable in the Actions column to enable this rule. Repeat this step to enable other alarm rules, or select multiple rules and click Enable under the alarm list.

8.4 Check application group details

The group details page contains the fault list, alarm history, alarm rules, group resources, events, and group resource metric data. You can use this page to monitor the preceding details of your application groups.

Group list

All application groups on CloudMonitor, along with the resources and health status of each group, are displayed on the group details page.

Parameters

- **Group name (or ID):** The name or identification number of an application group.
- **Health status:** The alarm status of any group resource. An application group is healthy when no active alarms are triggered for any of the resources in the group, but unhealthy whenever any metric threshold of a resource in the group is met and an alarm is triggered.
- **Instance count:** The total number of instances in an application group, both ECS and non-ECS instances.
- **Resource types:** The number of resource types in an application group. For example, if an application group contains ECS, ApsaraDB for RDS, and Server Load Balancer instances, then this number is three.
- **Unhealthy instances:** The total number of instances with active alarms in an application group. For example, if two ECS instances and one ApsaraDB for RDS instance have active alarms, the number of unhealthy instances is three.
- **Creation time:** The time when an application group is created.
- **Actions:** The actions that can be applied to an application group. Action types supported are manage, stop notifications, enable and disable all the alarm rules, and delete group.

Fault list

The resources with active alarms in your group are displayed in the fault list to help you to easily view unhealthy instances and quickly troubleshoot the causes.



Note:

- When multiple metrics of a resource have active alarms at the same time, the fault list displays the resource multiple times. Each row of the list shows a metric with an active alarm.
- Once you disable an alarm rule with an active alarm, the resources and metrics associated with the rule no longer appearing on the fault list.

Parameters

- **Faulty resource:** A resource with an active alarm.
- **Start time:** The time when the first alarm is generated for the resource.
- **Status:** Indicates whether a resource has an active alarm.
- **Duration:** The period of time when a faulty resource is in an alarm state.
- **Alarm rule name:** The name of the alarm rule applied to a faulty resource.
- **Actions:** The actions that can be applied to a faulty resource. You can click **Expand** to view the metric trends of a faulty resource with an active alarm over the past six hours, and compare the metric data with the alarm threshold value.

Alarm history

Alarm history provides the account of all the alarm rules applied to a group.



Note:

You can request the alarm history of the last three days. If the interval between the query start time and end time exceeds three days, the system prompts you to re-select the time range.

Parameters

- **Faulty resource:** A resource with an active alarm.
- **Duration:** The time during which a faulty resource is in an alarm state.
- **Occurrence time:** The time when the alarm is generated.
- **Alarm rule name:** The name of the alarm rule applied to a faulty resource.
- **Notification method:** The method by which alarm notifications are sent, which are SMS, email, and TradeManager.
- **Product type:** The product type to which a faulty resource belongs.
- **Status:** The status of the alarm rule, which are alarm status, cleared status, and muted states.
- **Notification target:** The group of contacts who receive alarm notifications.

Alarm rules

A list of all the alarm rules applied to a group is displayed in an alarm rules list. You can select the preferred alarm rule from the list and can enable, disable, or modify the rules based on your requirements.



Note:

The alarm rules list only shows the alarm rules applied to a specific application group. It does not show the alarm rules with Resource Range set to the All Resources or an Instance.

Parameters

- **Alarm name:** Name of an alarm rule specified when the alarm rule was created.
- **Status:** Displays whether the resources associated with the alarm rules have active alarms.
 - **Normal state:** All resources associated with the alarm rules are normal.
 - **Alarm state:** At least one instance associated with the alarm rule has an active alarm.
 - **Insufficient data:** At least one instance associated with the alarm rule has insufficient data and no instance has an active alarm.
- **Enable:** Shows whether the alarm rule is enabled.
- **Product name:** The name of the product to which group resources belong.
- **Alarm description:** A brief description of alarm rules setting.
- **Actions:** The optional operations include Modify, Enable, Disable, Delete, and Alarm History.
 - **Modify:** Click to make changes in the alarm rule.
 - **Disable:** Click to disable the alarm rule. Once the alarm rule is disabled, the alarm service does not check whether metric data exceeds the threshold value.
 - **Enable:** Click to enable the alarm rule. Once you enable a previously disabled alarm rule, the alarm service checks the metric data and determines whether to trigger an alarm based on the alarm rule.
 - **Delete:** Click to delete the alarm rule.
 - **Alarm History:** Click to view the alarm history of the alarm rule.

Group resources

Display all the resources of a group and the health condition of these resource.

Parameters

- **Instance name (or ID):** The instance name or ID of a resource.
- **Health status:** The alarm status of any group resource. An application group is healthy when no alarms are triggered for any of the resources in the group, but unhealthy whenever an alarm is triggered for any resource in the group.

Events

Alarm history and records for alarm rule operation events, such as add, modify, and delete actions, are supported, allowing you to trace any operation performed on a specific alarm rule.



Note:

You can query event information from the last 90 days.

Parameters

- **Occurrence time:** The time when an event occurred.
- **Event name:** The name of an event, which may be an alarm event such as alarm generated or alarm cleared, or an system event such as create alarm rule, modify alarm rule, or delete alarm rule.
- **Event type:** The type of event, which can be divided into system events and alarm events. Types of system events include create alarm rule, delete alarm rule, and modify alarm rule. Types of alarm events include alarm generated and alarm cleared.
- **Event details:** Detailed information associated with an event.

Charts

The lower area of the application group details page displays the monitoring details of group resources. By default, CloudMonitor initializes frequently used metric data. You can choose to customize the area, changing the chart type and metric data displayed.



Note:

To obtain the OS metrics of ECS, you must install the CloudMonitor agent.

Initialized metric data

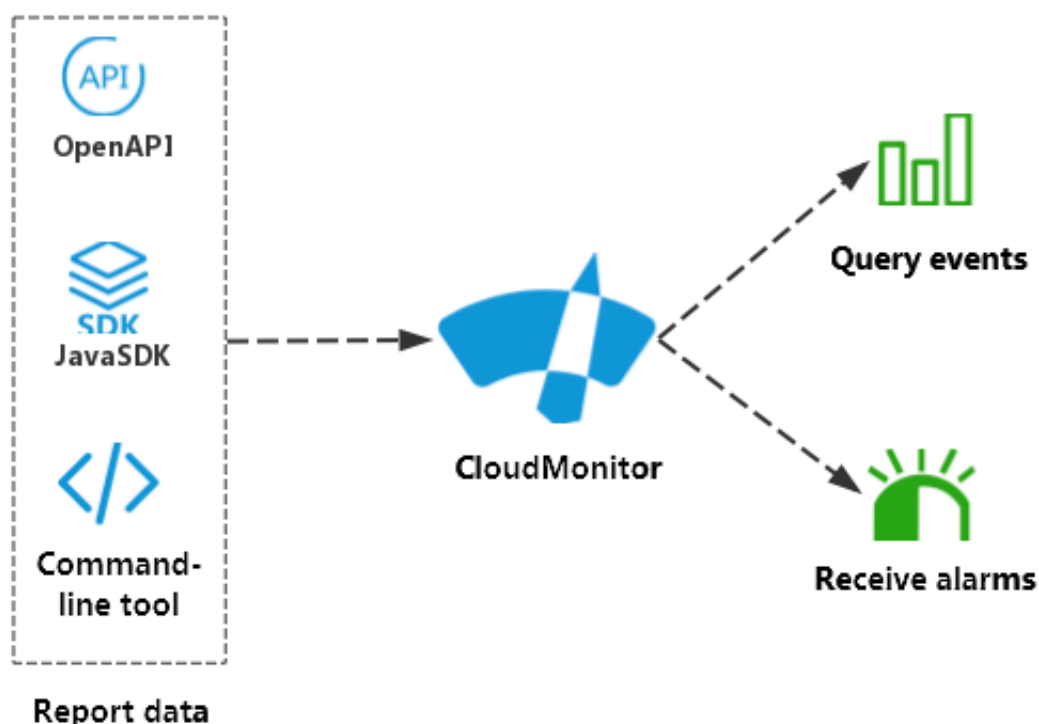
By default, CloudMonitor initiates the following application group data, which are all displayed in line charts. If you want to view more metric data, click **Add Metric Chart** to add more metrics to the data.

Product	Metrics	Chart type	Description
ECS	CPU usage and outbound bandwidth (Internet)	Line chart	Displays the aggregate data of all servers in the group.
ApsaraDB for RDS	CPU usage, disk usage, IOPS usage, connection usage	Line chart	Displays the data of a single database instance.
Server Load Balancer	Outbound bandwidth and inbound bandwidth	Line chart	Displays the data of a single Server Load Balancer instance.
OSS	Storage size and GET/PUT request count	Line chart	Displays the data of a single bucket.
CDN	Downstream bandwidth and hit rate	Line chart	Displays the data of a single domain name.
EIP	outbound bandwidth (Internet)	Line chart	Displays the data of a single instance.
ApsaraDB for Redis	Memory usage, connection usage, and QPS usage	Line chart	Displays the data of a single instance.
ApsaraDB for MongoDB	CPU usage, memory usage, IOPS usage, and connection usage	Line chart	Displays the data of a single instance.

9 Event monitoring

9.1 Event monitoring overview

Event monitoring provides reporting, querying, and alarm monitoring features for event-related data so that you can quickly and easily monitor and report various exceptions and important changes in your business operations. Event monitoring also ensures that you will receive alarm notifications as soon as an event-related exception occurs.



The difference between event monitoring and custom monitoring is as follows:

- Event monitoring reports and queries discontinuous event monitoring data and generates alarms if the conditions specified in your alarm rules are met.
- Custom monitoring reports and queries time-series monitoring data collected periodically and generates alarms if the conditions specified in your alarm rules are met.

Event monitoring processes

- Report event data

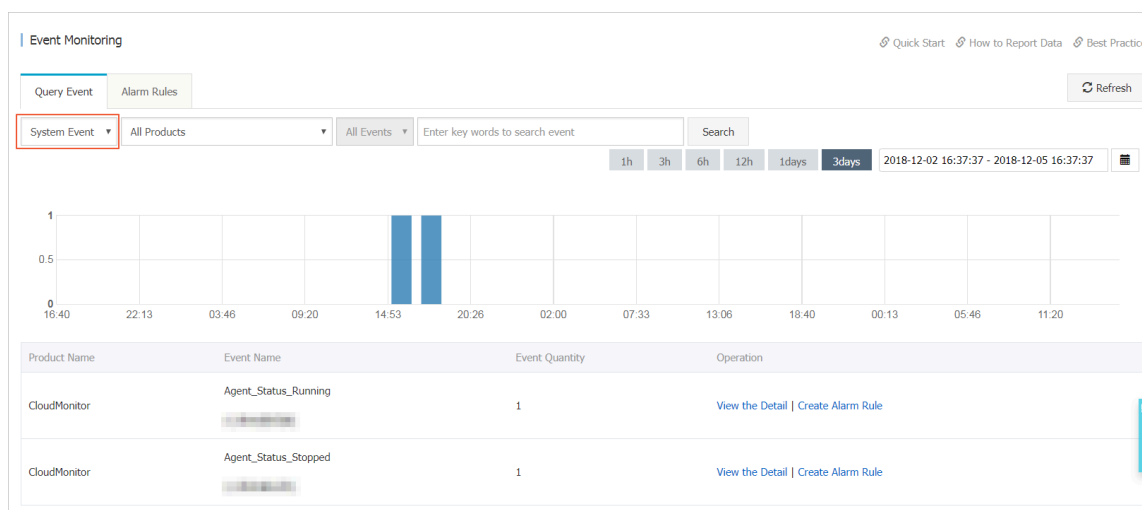
For more information, see [Report event data](#).

- Query event data

In the CloudMonitor console, you can query any reported event data. You can choose to view all the events on the Event Monitoring page, or enter a specific application group to view the events in that group.

To view all the reported events, follow these steps:

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Event Monitoring.
3. On the Event Monitoring page, you can view all the events under System Event or Custom Event, as shown in the following figure.



4. To view the details of a specific event, click View the Detail on the right of the target event.

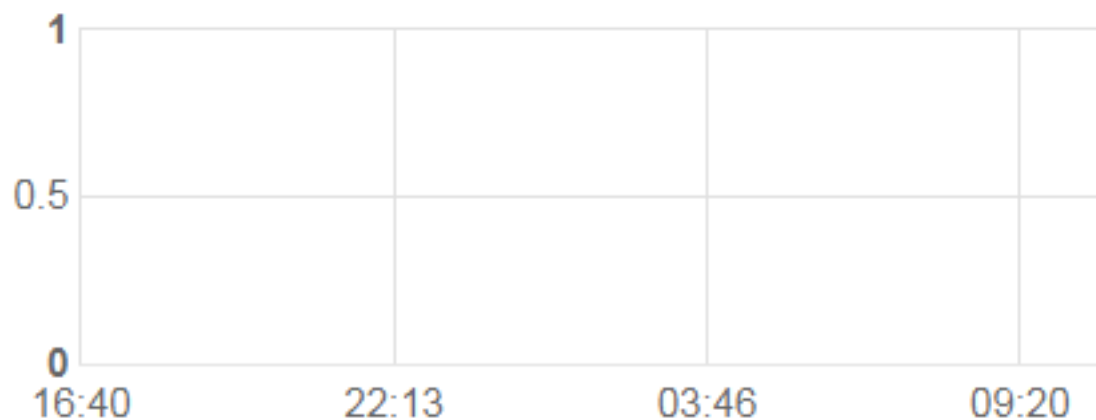
Event Monitoring


Query Event

Alarm Rules

System Event ▼

CloudMonitor ▼



Time	Product Name	Event Name	Event Level
18-12-03 18:40:20	CloudMonitor	Agent_Status_Running 	CR

To query events of a specific group, go to the specific Event Monitoring page of the group.

- Set an alarm rule

Event monitoring provides an alarm reporting feature. When setting an alarm rule, you need to select a corresponding application group. After an alarm is generated

, a notification is sent to the alarm contact group. To set alarm rules for an event, use either of the following two methods:

- Method 1:

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Event Monitoring.
3. Click Create Alarm Rule on the right of the target event.
4. In the displayed Create/modify event alerts dialog box, enter a name for the alarm rule, set the corresponding rules and notification method, and click OK.

- Method 2:

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups.
3. Click the target group name.
4. In the left-side navigation pane, click Event Monitoring.
5. Click Create Alarm Rule on the right of the target event.
6. In the displayed Create/modify event alerts dialog box, enter a name for the alarm rule, set the corresponding rules and notification method, and click OK.

Create / modify event alerts

Basic Information

Alarm Rule Name

Combination of alphabets, numbers and underscore, in 30 character

Event alert

Event Type

☒ System Event ☐ Custom Event

Product Type

Redis

Event Level

CRITICAL ✕

Event Name

Select

Resource Range

☒ All Resources ☐ Application Groups

Alarm type

OK

Cancel

9.2 Report event data

Event monitoring provides APIs for reporting events that can help you quickly and accurately collect and report event-related exceptions to CloudMonitor. Next, by configuring alarm rules for reported events, you can receive alarm notifications as soon as an event-related exception occurs.

CloudMonitor supports reporting event data by using open APIs, Java SDKs, and Alibaba Cloud CLI.

Limits

- Each Alibaba Cloud account can send up to 20 QPS.
- Up to 100 events can be reported at a time.
- Up to 500 kb of data can be reported at a time.

Report data using open APIs

- Service address:

```
https://metrichub-cms-cn-hangzhou.aliyuncs.com
```

- Request syntax

```
POST /event/custom/upload HTTP/1.1
Authorization: <authorizationstring>
Content-Length:<Content Length>
Content-MD5:<Content MD5>
Content-Type:application/json
Date: <GMT date>
HOST: maid
x-cms-signature:hmac-sha1
x-cms-api-version:1.0
x-cms-ip:30.27.84.196
User-Agent:cms-java-sdk-v-1.0
[{"content":"EventContent","groupId":GroupId,"name":"EventName",
"time":"20171023T144439.948+0800"}]
```

- Parameters

Name	Type	Required?	Description
Name	String	Yes	The name of the event
GroupId	Numerical	Yes	The application group ID to which the event belongs
Time	String	Yes	The time when the event occurs
Content	String	Yes	The event details

Request header definition

Signature Algorithm

- Response Element

The system returns the HTTP status code 200.

- Examples

- Request

```
POST /event/custom/upload HTTP/1.1
Host: metrichub-cms-cn-hangzhou.aliyuncs.com
x-cms-api-version:1.0
Authorization:YourAccKey:YourAccSecret
Host:metrichub-cms-cn-hangzhou.aliyuncs.com"
Date:Mon, 23 Oct 2017 06:51:11 GMT
Content-Length:180
x-cms-signature:hmac-sha1
Content-MD5:E9EF574D1AEAAA370860FE37856995CD
x-cms-ip:30.27.84.196
User-Agent:cms-java-sdk-v-1.0
Content-Type:application/json
[{"Content": "123, ABC", "groupid": 100, "name": "event_0", "Time": "loud. 948 + 0800 "}]
```

- Response

```
{
  "Code": 200 ",
  "msg": "// return MSG is empty for reports that are normal
}
```

Report data using the Java SDKs

- Maven dependency

```
<dependency>
  <groupId>com.aliyun.openservices</groupId>
  <artifactId>aliyun-cms</artifactId>
  <version>0.1.2</version>
</dependency>
```

- Sample code

```
public void uploadEvent() throws CMSException, InterruptedException
{
    // Initialize Client
    CMSClient cmsClient = new CMSClient(endpoint, accKey, secret
);
    // Build 2 event reports
    CustomEventUploadRequest request = CustomEventUploadRequest
.builder()
        .append(CustomEvent.builder()
            .setContent("abc,123")
            .setGroupId(101l)
            .setName("Event001").build())
        .append(CustomEvent.builder()
            .setContent("abc,123")
            .setGroupId(101l)
            .setName("Event002").build())
        .build();
    CustomEventUploadResponse response = cmsClient.
putCustomEvent(request);
    List<CustomEvent> eventList = new ArrayList<CustomEvent
>();
```

```

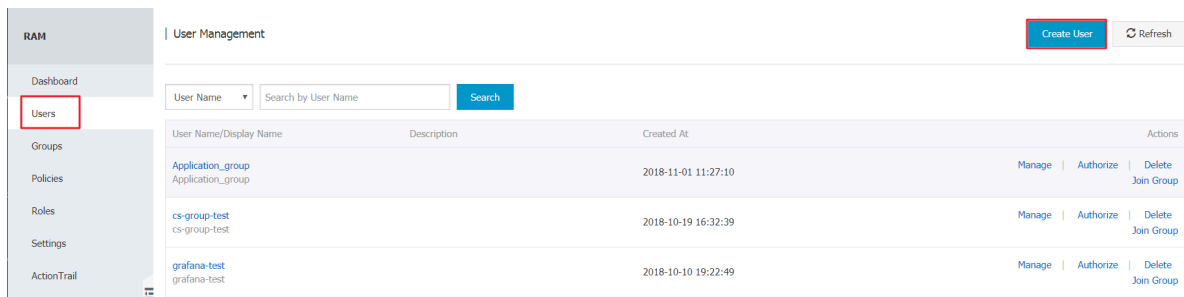
eventList.add(CustomEvent.builder()
    .setContent("abcd,1234")
    .setGroupId(1011)
    .setName("Event001").build());
eventList.add(CustomEvent.builder()
    .setContent("abcd,1234")
    .setGroupId(1011)
    .setName("Event002").build());
request = CustomEventUploadRequest.builder()
    .setEventList(eventList).build();
response = cmsClient.putCustomEvent(request);
}

```

Report data using Alibaba Cloud CLI

Using your primary account, generate a RAM account AccessKey with cloud monitoring privileges (on the premise that a RAM account is more secure than your primary account).

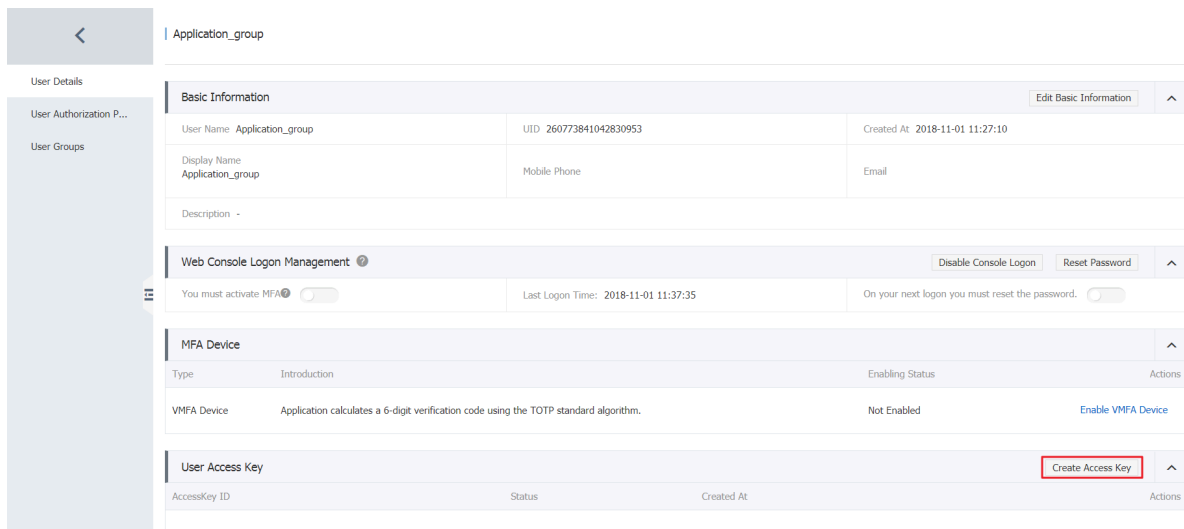
- Create a RAM account.



The screenshot shows the 'User Management' page in the RAM console. On the left sidebar, the 'Users' option is highlighted with a red box. The main area displays a table of users with columns for 'User Name/Display Name', 'Description', 'Created At', and 'Actions'. Three users are listed: 'Application_group', 'cs-group-test', and 'grafana-test'. Each user has 'Manage', 'Authorize', and 'Delete' actions available.

User Name/Display Name	Description	Created At	Actions
Application_group	Application_group	2018-11-01 11:27:10	Manage Authorize Delete Join Group
cs-group-test	cs-group-test	2018-10-19 16:32:39	Manage Authorize Delete Join Group
grafana-test	grafana-test	2018-10-10 19:22:49	Manage Authorize Delete Join Group

- Generate an AccessKeyId and AccessKeySecret for the RAM account.



The screenshot shows the 'Application_group' user details page. The left sidebar has 'User Groups' selected. The main area is divided into sections: 'Basic Information', 'Web Console Logon Management', 'MFA Device', and 'User Access Key'. The 'User Access Key' section at the bottom has a 'Create Access Key' button highlighted with a red box.

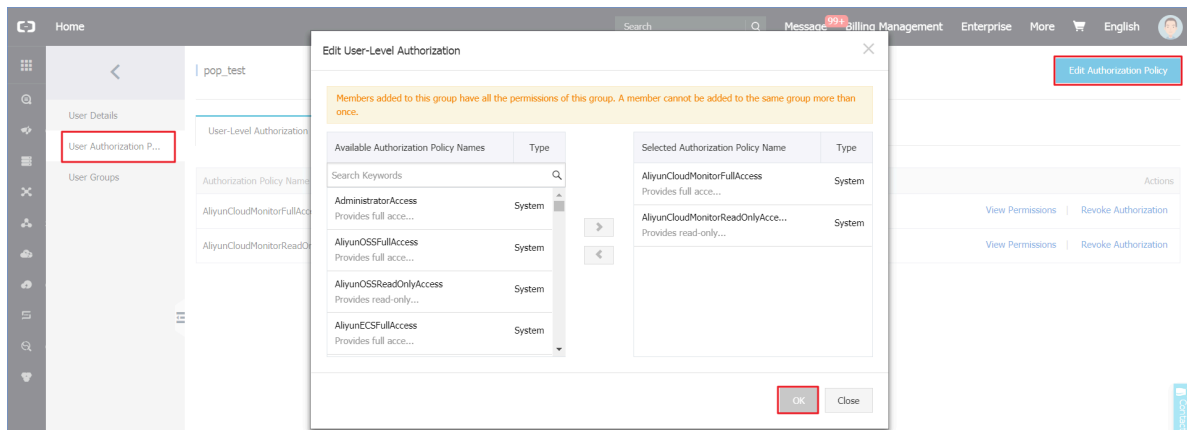
Basic Information			
User Name	Application_group	UID	260773841042830953
Created At	2018-11-01 11:27:10		
Display Name	Application_group	Mobile Phone	Email
Description	-		

Web Console Logon Management		
You must activate MFA	<input type="checkbox"/>	Last Logon Time: 2018-11-01 11:37:35
On your next logon you must reset the password.		<input type="checkbox"/>

MFA Device			
Type	Introduction	Enabling Status	Actions
VMFA Device	Application calculates a 6-digit verification code using the TOTP standard algorithm.	Not Enabled	Enable VMFA Device

User Access Key			
AccessKey ID	Status	Created At	Actions

- Assign CloudMonitor permissions for the RAM account.



Procedure

1. Install the aliyuncli tool.

System requirements: Linux, UNIX, or Mac OS. Environment requirement: Python 2.7.x installed.

a. Install Python

- If Python 2.7.x is already installed on your device, skip this step.
- If Python 2.7.x is not installed on your device, run the following command in the command line window to install Python. Make sure that wget is installed on your device.

```
wget https://www.python.org/ftp/python/2.7.8/Python-2.7.8.tgz (
or download it in other ways and put it in a certain path)
tar -zxvf Python-2.7.8.tgz
cd Python-2.7.8
./configure
make
sudo make install
```

b. Install pip

- If pip is already installed on your device, you can skip this step.
- If pip is not installed on your device, run the following command in the command line window to install pip:

```
curl "https://bootstrap.pypa.io/get-pip.py" -o "pip-install.py"
```

```
sudo python pip-install.py
```

- If the installation is successful, information such as the following is displayed:

```
Successfully installed pip-7.1.2 setuptools-18.7 wheel-0.26.0
```

c. Install the command line tool

If the pip version installed is no longer supported, an error will occur while installing the Alibaba Cloud CLI. You can use the command that follows to upgrade the pip version before performing other operations. We recommend you use pip 7.x or later. If your pip is already a supported version, you can skip this step.

- A. Run the following command in the command line window to upgrade the pip version:

```
sudo pip install -U pip
```

If the installation is successful, information such as the following is displayed :

```
Successfully uninstalled pip-7.1.2  
Successfully installed pip-8.1.2
```

- B. Run the following command to install the Alibaba Cloud command line tool:

```
sudo pip install aliyuncli
```

If the installation is successful, information such as the following is displayed :

```
Successfully installed aliyuncli-2.1.2 colorama-0.3.3 jmespath-0.7.1
```

d. Configure the command-line tool

```
~ Sudo aliyuncli configure  
Aliyun Access Key ID [*****a]: youraccesskeyid  
Aliyun Access Key Secret [*****b]: youraccesskeysecret  
Default Region Id [cn-hangzhou]: cn-hangzhou  
Default output format [json]: json
```

2. Install the CMS SDK

- For Windows, run the following command in the command line window:

```
cd C:\Python27\Scripts
```

```
pip install aliyun-python-sdk-cms
```

- To update the SDK, run the following command:

```
pip install --upgrade aliyun-python-sdk-cms
```

- For Linux, run the following command in the command line window:

```
sudo pip install aliyun-python-sdk-cms
```

- To update the SDK, run the following command:

```
sudo pip install --upgrade aliyun-python-sdk-cms
```

3. Report monitoring data

Use the API PutEvent.

- Reporting example for Windows:

```
aliyuncli.exe cms PutEvent --EventInfo "[{'content':'helloworld','time':'20171013T170923.456+0800','name':'ErrorEvent','groupId':'27147'}]"
```

- Reporting example for Linux

```
aliyuncli cms PutEvent --EventInfo "[{'content':'helloworld','time':'20171023T180923.456+0800','name':'ErrorEvent','groupId':'27147'}]"
```

- If data is reported successfully, the system returns status code 200.

```
{  
  "Code": "200"  
}
```

Status codes

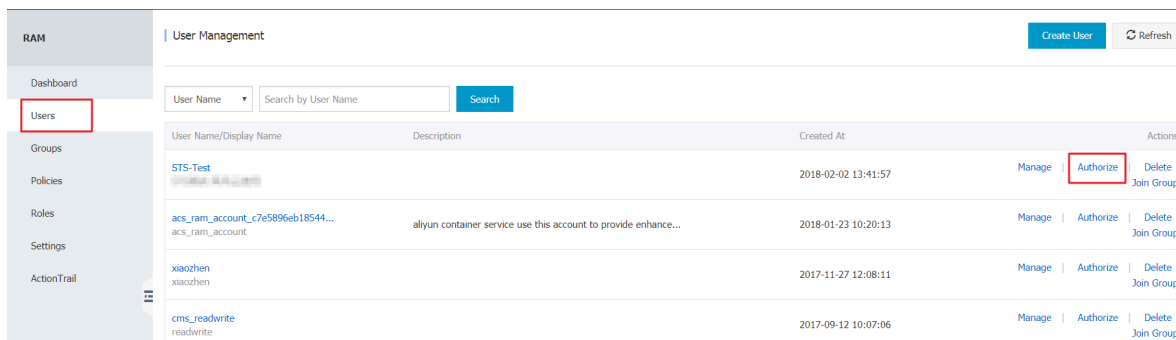
Status code	Description
200	Normal
400	Syntax error in the client request
403	Verification failure, speed limit, or authorization error
500	Internal server error

RAM account authorization

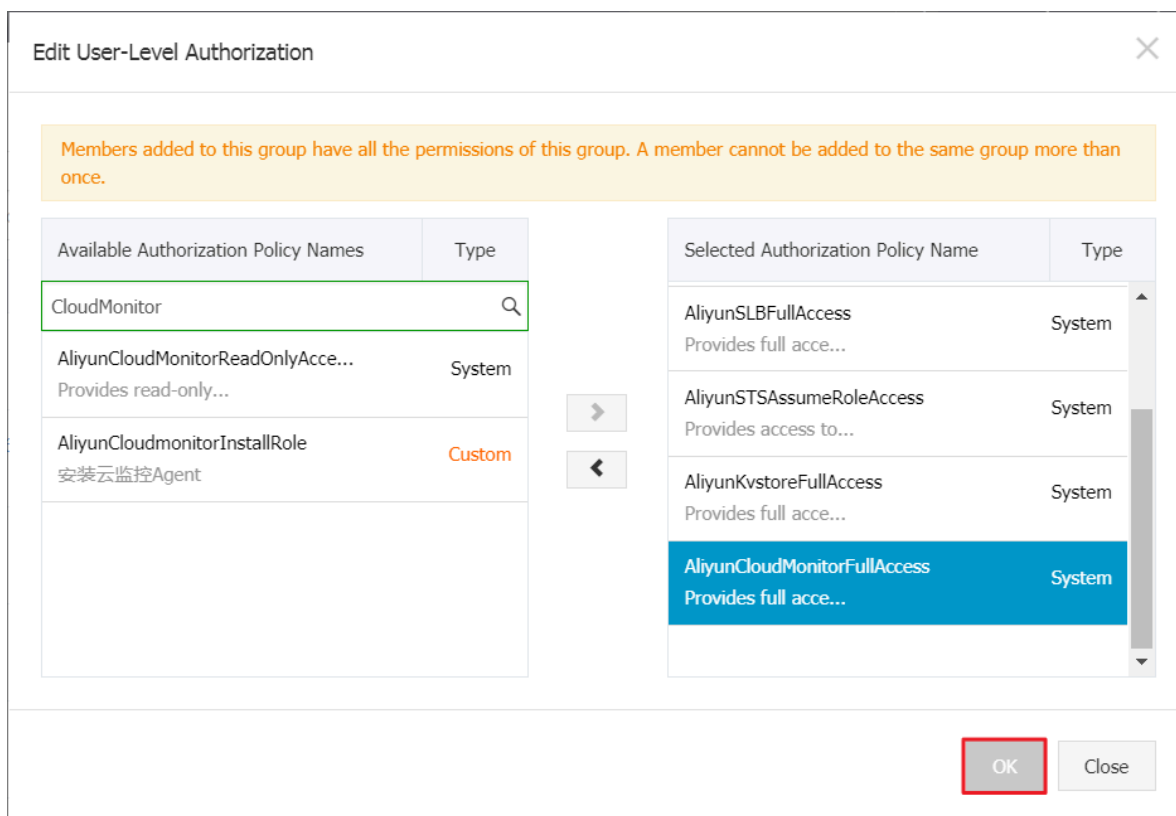
When the AccessKey of a RAM user account is used to report an event, the RAM account must be authorized to manage CloudMonitor. If a RAM account is not authorized to manage CloudMonitor, a prompt will be displayed that indicates that

the event cannot be uploaded because the RAM user account is not authorized for this operation.

1. Log on to the [access control Ram console](#).
2. Enter the user management menu.
3. Click Authorize next to the RAM user account that is used to report data.



4. On the authorization page, select manage permissions for cloud monitoring, and click OK to save the authorization.



9.3 Event monitoring best practices

Use cases

Exceptions may occur when the service is running. Some exceptions can be automatically restored by retry and other methods, while the others cannot. Serious exceptions can even lead to customer business interruption. Therefore, a system is necessary to record these exceptions and trigger alarms when specific conditions are met. The traditional method is to print file logs and collect the logs to specific systems, for example, open-source ELK (ElasticSearch, Logstash, and Kibana). These open-source systems consist of multiple complex distributed systems. The complicated technology and high cost make independent maintenance challenging. CloudMonitor provides the event monitoring feature to effectively solve these problems.

The following examples explain how to use the event monitoring feature.

Case studies

1. Report exceptions

Event monitoring provides two methods for data reporting, namely, Java SDK and Open API. The following describes how to report data by using Java SDK.

a. Add Maven dependency

```
<dependency>
  <groupId>com.aliyun.openservices</groupId>
  <artifactId>aliyun-cms</artifactId>
  <version>0.1.2</version>
</dependency>
```

b. Initialize SDK

```
// Here, 118 is the application grouping ID of CloudMonitor.
Events can be categorized by applications. You can view group IDs
in CloudMonitor application grouping list.
CMSClientInit.groupId = 118L;
// The address is the reporting entry of the event system, which
is currently the public network address. AccessKey and Secret/key
are used for personal identity verification.
```

```
CMSCClient c = new CMSCClient("https://metrichub-cms-cn-hangzhou.
aliyuncs.com", accesskey, secretkey);
```

c. Determine whether to asynchronously report the data.

CloudMonitor event monitoring provides synchronous reporting policy by default. The good thing is that writing code is simple, and the reported events are reliable and free from data loss.

However, such policy also brings some problems as well. Event reporting codes are embedded in business codes, which may block code running and affect the normal business in case of network fluctuations. Many business scenarios do not require events to be 100% reliable, so a simple asynchronous reporting encapsulation is sufficient. Write the event into a `LinkedBlockingQueue` and perform batch reporting on the backend asynchronously using `ScheduledExecutorService`.

```
//Initialize queue and Executors:
private LinkedBlockingQueue<EventEntry> eventQueue = new
LinkedBlockingQueue<EventEntry>(10000);
private ScheduledExecutorService schedule = Executors.newSingleT
hreadScheduledExecutor();
// Report event:
//Every event contains its name and content. The name is for
identification and the content contains details of the event, in
which the full-text search is supported.
public void put(String name, String content) {
    EventEntry event = new EventEntry(name, content);
    // When the event queue is full, additional events are
discarded directly. You can adjust this policy as needed.
    boolean b = eventQueue.offer(event);
    if (! b) {
        logger.warn("The event queue is full, discard: {}", event
);
    }
}

//Submit events asynchronously. Initialize scheduled tasks.
Report events in batch by run every second. You can adjust the
reporting interval as needed.
schedule.scheduleAtFixedRate(this, 1, 1, TimeUnit.SECONDS);
public void run() {
    do {
        batchPut();
    } while (this.eventQueue.size() > 500);
}

private void batchPut() {
    // Extract 99 events from the queue for batch reporting.
    List<CustomEvent> events = new ArrayList<CustomEvent>();
    for (int i = 0; i < 99; i++) {
        EventEntry e = this.eventQueue.poll();
        if (e == null) {
            break;
        }

        events.add(CustomEvent.builder().setContent(e.getContent
()).setName(e.getName()).build());
    }
}
```

```

        if (events.isEmpty()) {
            return;
        }

        // Report events in batch to CloudMonitor. No retry or retry
        // in SDK is added here. If you have high requirement for event
        // reliability, add retry policies.
        try {
            CustomEventUploadRequestBuilder builder = CustomEventUploadRequestBuilder.builder();
            builder.setEventList(events);
            CustomEventUploadResponse response = cmsClient.putCustomEvent(builder.build());
            if (!"200".equals(response.getErrorCode())) {
                logger.warn("event reporting error: msg: {}, rid: {}", response.getErrMsg(), response.getRequestId());
            }
        } catch (Exception e1) {
            logger.error("event reporting exception", e1);
        }
    }
}

```

d. Event reporting demo

- Demo1: http Controller exception monitoring

The main purpose is to monitor if a large number of exceptions exist in HTTP requests. If the number of exceptions per minute exceeds a certain limit, an alarm is triggered. The implementation principle is to intercept HTTP requests by using Spring interceptor, servlet filter and other technologies. Logs are created in case of exceptions and alarms are triggered by setting alarm rules.

The event reporting demo is as follows:

```

// Each event should be informative for searching and locating
// Here, map is used for organizing events and converted to Json
// format as event content.
Map<String, String> eventContent = new HashMap<String, String>();
eventContent.put("method", "GET"); // http request method
eventContent.put("path", "/users"); // http path
eventContent.put("exception", e.getClass().getName()); //
// Exception class name for searching
eventContent.put("error", e.getMessage()); // Error message of
// exception
eventContent.put("stack_trace", ExceptionUtils.getStackTrace(e)); // Exception stack for locating
// Finally submit the events in the preceding asynchronous
// reporting method. Since no retry is performed in asynchronous
// reporting, event loss of small probability may happen. However,
// it is sufficient for alarms of unknown http exceptions.
put("http_error", JsonUtils.toJson(eventContent));

```

```
image.png](http://ata2-img.cn-hangzhou.img-pub.aliyun-inc.com/864cf095977cf61bd340dd1461a0247c.png)
```

- **Demo2: Monitoring of scheduled tasks on the backend and message consumption**

Like the preceding http events, many similar business scenarios require alarms. In the business scenarios such as backend tasks and message queue consumption, the events can be reported by using similar methods to achieve effective monitoring. When any exception occurs, alarms are triggered immediately.

```
//Event organization of the message queue:
Map<String, String> eventContent = new HashMap<String, String>();
eventContent.put("cid", consumerId); // Consumer ID
eventContent.put("mid", msg.getMsgId()); // Message ID
eventContent.put("topic", msg.getTopic()); // Message topic
eventContent.put("body", body); // Message body
eventContent.put("reconsume_times", String.valueOf(msg.getReconsumeTimes())); // The number of retries after message failure
eventContent.put("exception", e.getClass().getName()); // Exception class name in case of exception
eventContent.put("error", e.getMessage()); // Exception message
eventContent.put("stack_trace", ExceptionUtils.getStackTrace(e)); // Exception stack
// Finally, report the event
```



```
put("metaq_error", JsonUtils.toJson(eventContent));
```

Check the event after reporting:

<

Group Resource

●

Dashboards

Fault List

Event Monitor

Availability Monitor...

Log Monitoring

Custom Monitoring

Alarm Logs

Alarm Rule

demo

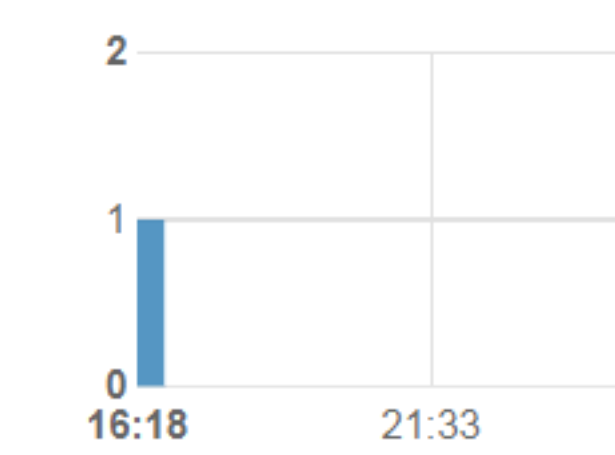
Back to Application

Quick Start

How to Report

System Event

CloudMonitor



Product Name
CloudMonitor
CloudMonitor

- **Set alarms for queue message consumption exceptions:**

<

Group Resource

●

 Dashboards

Fault List

Event Monitor

Availability Monitor...

Log Monitoring

Custom Monitoring

Alarm Logs

Alarm Rule

demo

Back to Application

Quick Start

How to Report

System Event

CloudMonitor

2

1

0

16:18

21:33

Product Name

CloudMonitor

CloudMonitor

- **Demo 3: Record important events**

Another use case of events is to record important actions for later check without sending alarms. For example, operation logs for important business, password change/order change, remote logon, and so on.

<

Group Resource

●

Dashboards

Fault List

Event Monitor

Availability Monitor...

Log Monitoring

Custom Monitoring

Alarm Logs

Alarm Rule

demo

Back to Application

Quick Start

How to Report

System Event

1

0.5

0

16:18

21:33

Time	Product Name
18-11-01 09:28:38	CloudMonitor
18-10-31 09:32:34	CloudMonitor

9.4 Cloud product system event monitoring

System event monitoring allows you to monitor and query system events generated by multiple cloud services, enabling you to gain better insights into your cloud usage.

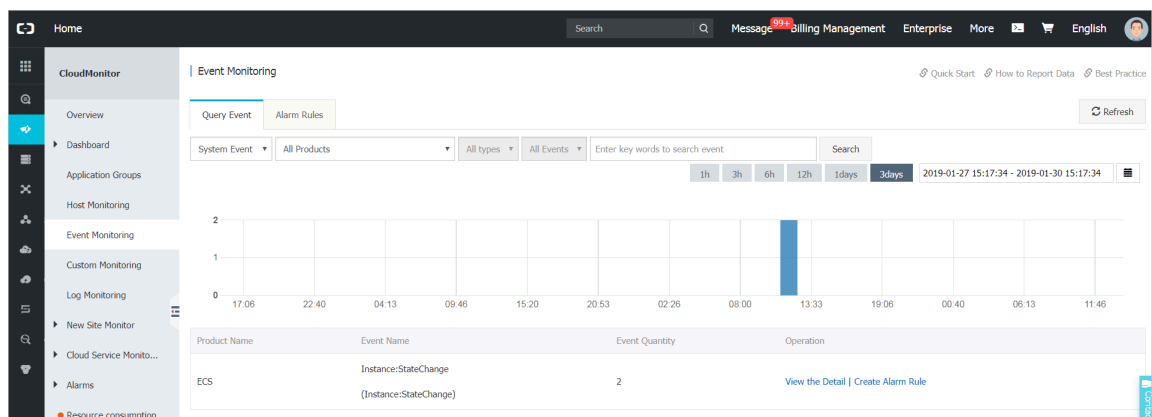
After resources are classified by application group, system events generated by cloud products are automatically associated with the resources in the group, which helps you to integrate various types of monitoring information, so as to quickly analyze and locate problems if any.

At the same time, an alarm function is provided for system events. You can configure alarm rules and receive alarm notifications through different methods, such as email and DingTalk chatbot, according to the event level. You can also use callbacks to learn system events. In this way, you can learn serious events at the earliest possible time and handle them in time, making online operation and maintenance automatic.

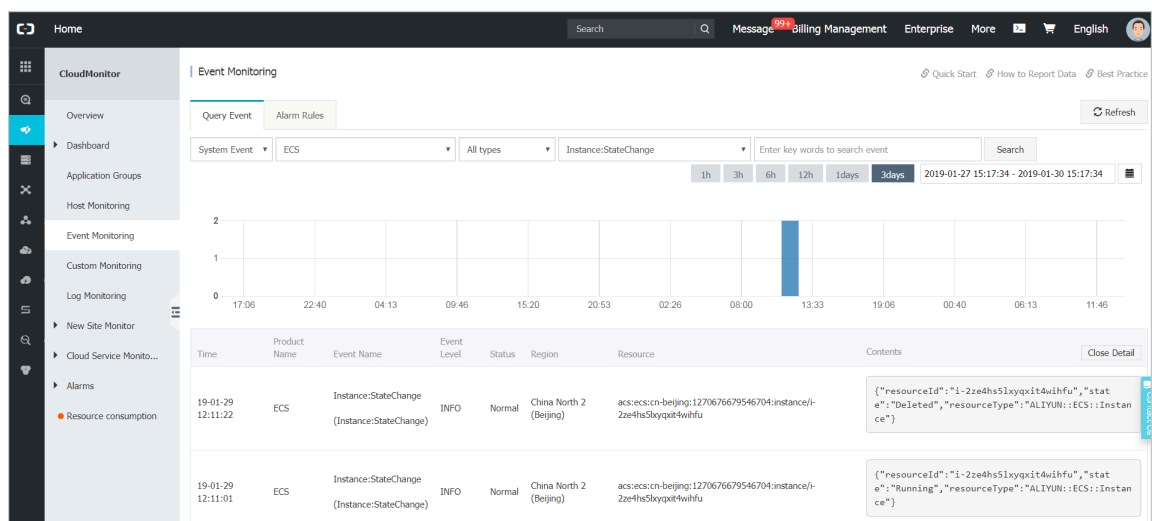
View system events

- Method 1

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Event Monitoring.
3. On the Query Event tab page, select System Event, the target product, the target event type, and the target event from the drop-down lists. Then, select a time period to view the events within this period.



4. Click View the Detail in the Operation column to view the details of an event.



- Method 2

If your resources are allocated into application groups, you can also view system events for the instances in each group on the application group page.

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups.
3. Find the target application group and click the group name.
4. In the left-side navigation pane, click Event Monitor. System events for the instances in the group are displayed.

Use system event alarms

You can set alarm rules for all system events, so that you will be quickly notified in the case of any system events. The following two notification methods are provided:

- Send alarm notifications by email or DingTalk chatbot.
- Notify you of the events through MNS queue, Function Compute service, or URL callback so that you can handle event exceptions according to your service scenarios.

Create an alarm rule

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Event Monitoring.
3. Click the Alarm Rules tab.
4. In the upper-right corner, click Create Event Alerts.
5. In the Basic Information area, enter an Alarm Rule Name.
6. In the Event alert area, set the following parameters:
 - Event Type: Select System Event.
 - Product Type, Event Type, Event Level, and Event Name: Select the target product type, event type, event level, and event name.
 - Resource Range:
 - All Resources: You will be notified of any event that occurs to any of your resources.
 - Application Groups: You will be notified of the events that are associated with the application group you specify.

7. Select the Alarm Type and then click OK.

Create / modify event alerts

Alarm type

☒ Alarm notification

Contact Group [Delete](#)

Notification Method

Info (Email ID+Ali WangWang+DingTalk Robot)

[+Add](#)

☐ MNS queue

☐ Function service ([Best Practises](#))

☐ URL callback

OK **Cancel**

Test an alarm rule

A test function is provided for system event alarms. You can simulate system events to check whether you can be notified of system events as you have specified in alarm rules, for example, whether event alarms can be transferred through MNS queue and whether functions in the Function Compute service can be triggered.

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Event Monitoring.

3. Click the Alarm Rules tab to go to the alarm rule list page.

Event Monitoring

[Quick Start](#)
[How to Report Data](#)
[Best Practice](#)

Query Event

Alarm Rules

Refresh

System Event

Custom Event

Rule Name	Enable	Rule Description	Resource Range	Target	Actions
<input type="checkbox"/> ECSNotWork	Enabled	ECS CRITICAL Instance:InstanceFailure.Reboot:Executed	Application Groups	Alarm notification GPU Info (Email ID+AliWangWang+DingTalk Robot)	Modify test Disable Delete

4. Click Test in the Actions column.

5. Select the event from the Event Name drop-down list, and the content of the event will be displayed in the Content field. You can modify the content, such as the instance ID as needed.

Create event test

Product Type ECS

Event Level :CRITICAL

Event Name

InstanceFailure.Reboot.Executing

Content(JSON)

```
{
  "product": "ECS",
  "resourceId": "acs:ecs:cn-hangzhou:1514026440154742:instance/{instanceId}",
  "level": "CRITICAL",
  "instanceName": "instanceName",
  "regionId": "cn-hangzhou",
  "name": "Instance:InstanceFailure.Reboot:Executing",
  "content": {
    "executeFinishTime": "2018-06-08T01:25:37Z",
    "executeStartTime": "2018-06-08T01:23:37Z",
    "ecsInstanceName": "timewarp",
    "eventId": "e-t4nhcpqcu8fqushpn3mm",
    "eventType": "InstanceFailure.Reboot",
    "ecsInstanceId": "i-bp1a2b3c4d5e6f7g8h9i0j"
  }
},
  "status": "Executing"
}
```

OK

Cancel

6. Click OK. An event will be sent and the alarm notification method you have specified in your alarm rule, such as alarm notification, MNS queue, Function service, and URL callback, will be triggered.

Supported system events of different cloud products

- ECS

Event	Description	Status	Event level
Instance: InstanceFailure.Reboot	Instance reboot due to instance failure has started.	Executing	CRITICAL
Instance: InstanceFailure.Reboot	Instance reboot due to instance failure has ended.	Executed	CRITICAL
Instance: SystemFailure.Reboot	Instance reboot due to system failure has started.	Executing	CRITICAL
Instance: SystemFailure.Reboot	Instance reboot due to system failure has ended.	Executed	CRITICAL
Instance: SystemMaintenance.Reboot	Instance reboot is scheduled due to system maintenance.	Scheduled	CRITICAL
Instance: SystemMaintenance.Reboot	Scheduled instance reboot for system maintenance is avoided.	Avoided	CRITICAL
Instance: SystemMaintenance.Reboot	Scheduled instance reboot for system maintenance has started.	Executing	CRITICAL
Instance: SystemMaintenance.Reboot	Scheduled instance reboot for system maintenance has ended.	Executed	CRITICAL
Instance: SystemMaintenance.Reboot	Scheduled instance reboot for system maintenance is cancelled.	Canceled	CRITICAL

Event	Description	Status	Event level
Instance: SystemMain tenance.Reboot	Scheduled instance reboot for system maintenance has failed.	Failed	CRITICAL
Disk:Stalled	A serious impact on disk performance begins.	Executing	CRITICAL
Disk:Stalled	A serious impact on disk performance stops .	Executed	CRITICAL
Instance: StateChange	Instance status changes.	Normal	INFO
Instance: Preemptibl eInstanceI nterruption	A preemptibl e instance is interrupted.	Normal	WARN

**Note:**

- You can view the instance status in event details. The instance statuses are:
 - Pending: the status after an instance is created and before it begins to run
 - Starting: the status of an instance that is being started
 - Running: the status of an instance that is running
 - Stopping: the status of an instance that is being stopped
 - Stopped: the status of an instance that has stopped or expired or been locked , or that is about to expire, is being recycled upon overdue payment, or is waiting for release
 - Deleted: the status of an instance that has been released

For more information, see [Preemptible instance](#).
- Reasons that a preemptible instance is about to be recycled:
 - Your bid is lower than the current market price.
 - The supply and demand relationship of resources changes.

For more information, see [Preemptible instance](#).

- SLB

Event	Description	Event level
CertKeyExpired_1	The certificate will expire in one day.	WARN
CertKeyExpired_3	The certificate will expire in three days.	WARN
CertKeyExpired_7	The certificate will expire in seven days.	WARN
CertKeyExpired_15	The certificate will expire in 15 days.	WARN
CertKeyExpired_30	The certificate will expire in 30 days.	WARN
CertKeyExpired_60	The certificate will expire in 60 days.	WARN

- OSS

Event	Description	Event level
BucketEgressBandwidth	Bucket downstream bandwidth exceeds the reporting threshold.	INFO
BucketEgressBandwidthThresholdExceeded	Bucket downstream bandwidth exceeds the flow control threshold.	WARN
BucketIngressBandwidth	Bucket upstream bandwidth exceeds the reporting threshold.	INFO
BucketIngressBandwidthThresholdExceeded	Bucket upstream bandwidth exceeds the flow control threshold.	WARN
UserEgressBandwidth	User downstream bandwidth exceeds the reporting threshold.	INFO
UserEgressBandwidthThresholdExceeded	User downstream bandwidth exceeds the flow control threshold.	WARN

Event	Description	Event level
UserIngressBandwidth	User upstream bandwidth exceeds the reporting threshold	INFO
UserIngressBandwidthThresholdExceeded	User upstream bandwidth exceeds the flow control threshold	WARN

- Auto Scaling

Event	Description	Status	Event level
AUTOSCALING:SCALE_IN_ERROR	The scaling-in of a scaling group fails.	Unnormal	CRITICAL
AUTOSCALING:SCALE_IN_SUCCESS	The scaling-in of a scaling group succeeds.	Normal	INFO
AUTOSCALING:SCALE_OUT_ERROR	The scaling-out of a scaling group fails.	Unnormal	CRITICAL
AUTOSCALING:SCALE_OUT_SUCCESS	The scaling-out of a scaling group succeeds.	Normal	INFO
AUTOSCALING:SCALE_REJECT	The scaling of a scaling group is rejected.	Warn	WARN
AUTOSCALING:SCHEDULE_TASK_EXPIRING	Expiration reminder	Warn	WARN
AUTOSCALING:SCALE_OUT_START	The scaling-out of a scaling group starts.	normal	INFO
AUTOSCALING:SCALE_IN_START	The scaling-in of a scaling group starts.	normal	INFO

- IoT

Event	Description	Status	Event level
RuleEngineProcessFa	The rule engine fails.	Failed	WARN

- Smart Access Gateway

Event	Description	Status	Event level
AccessGatewayFailover	Access points switch over.	Agwfailover	INFO
ConnectionDisconnect	The network is disconnected.	Disconnect	CRITICAL
DeviceHacked	Devices are attacked.	Hacked	CRITICAL
DeviceOffline	Devices become offline.	Offline	CRITICAL
DeviceOnline	Devices become online.	Online	INFO

- CloudMonitor

Event	Description	Status	Event level
Group_AddResourcesFailed_QuotaReached	Fails to automatically add a server into an application group because the resource quota is exceeded.	Failed	CRITICAL
Agent_Status_Stopped	The agent fails to respond to the heartbeat check.	Stopped	CRITICAL
Agent_Status_Running	The agent resumes heartbeat.	Running	CRITICAL

- DBS

Event	Description	Status	Event level
CloseContBackup	Incremental backup is disabled.	Failed	INFO
ContBackupFail	An error occurs during an incremental backup.	Failed	WARN
DataRestoreFail	An error occurs during data recovery.	Failed	WARN

Event	Description	Status	Event level
DataRestoreSuccess	Data recovery succeeds.	Running	WARN
FullBackupFail	An error occurs during a full backup.	Failed	WARN
InstancePause	A backup plan pauses.	Failed	INFO
InstanceStart	A backup plan starts.	Running	INFO
OpenContBackup	Incremental backup is enabled.	Running	INFO

- RDS

Event	Description	Status	Event level
Instance_Failover	The master and slave instances switch over.	Executed	WARN
Instance_Failure_Start	Instance fault starts.	Executing	CRITICAL
Instance_Failure_End	Instance fault ends.	Executed	CRITICAL

- Redis

Event	Description	Status	Event level
Instance_Failover	The master and slave instances switch over.	Executed	WARN
Instance_Failure_Start	Instance failure starts.	Executing	CRITICAL
Instance_Failure_End	Instance failure ends.	Executed	CRITICAL

- MongoDB

Event	Description	Status	Event level
Instance_Failure_Start	Instance fault starts.	Executing	CRITICAL

Event	Description	Status	Event level
Instance_Failure_End	Instance fault ends.	Executed	CRITICAL

- Container Service for Swarm

Event	Description	Status	Event level
NodeServiceAbnormal	The node status is abnormal.	Abnormal	CRITICAL
ServiceStatusAbnormal	The service status is abnormal.	Abnormal	CRITICAL

9.5 Use system event alarms

Scenario

To help you to quickly learn of system event exceptions and automate the handling of these event-related exceptions when a system event occurs for one or more of your Alibaba Cloud products, the alarm function of event monitoring provides the following notification methods:

- Alarm notifications for system events are sent as voice, text, or email messages.
- System events are distributed to your MNS queue, function service, and URL callback.

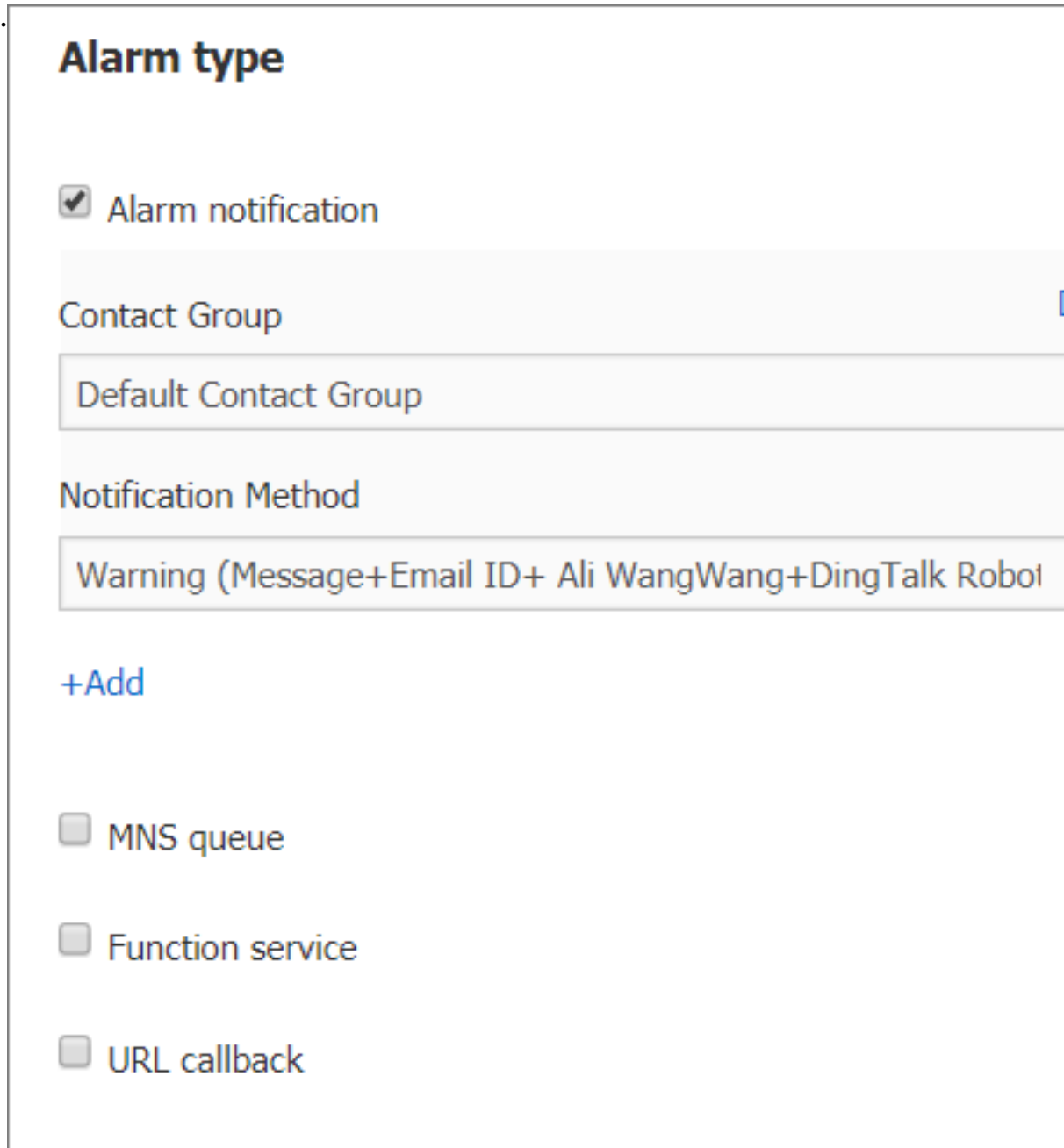
Create an alarm rule

1. Log on to the [CloudMonitor console](#).
2. In the left navigation pane, select Event Monitoring.
3. On the Alarm Rules tab page, click Create event alerts in the upper-right corner.
The Create / modify event alerts dialog box is displayed.
4. In the Basic Information area, fill in the alarm rule name.
5. In the Event alert area, complete the following information:
 - a. Event Type: Select System Event.
 - b. Product Type, Event Level, Event Name: Enter information based on your requirements.
 - c. Resource Range: Select All Resources or Application Groups. If you select All Resources, notifications are sent for any resource-related exceptions. If you

select Application Groups, notifications are sent only when an exception occurs for resources in the application group or groups you specified.

6. Select the Alarm type. CloudMonitor supports four alarm types:

alarm notification, MNS queue, function service, and URL callback.

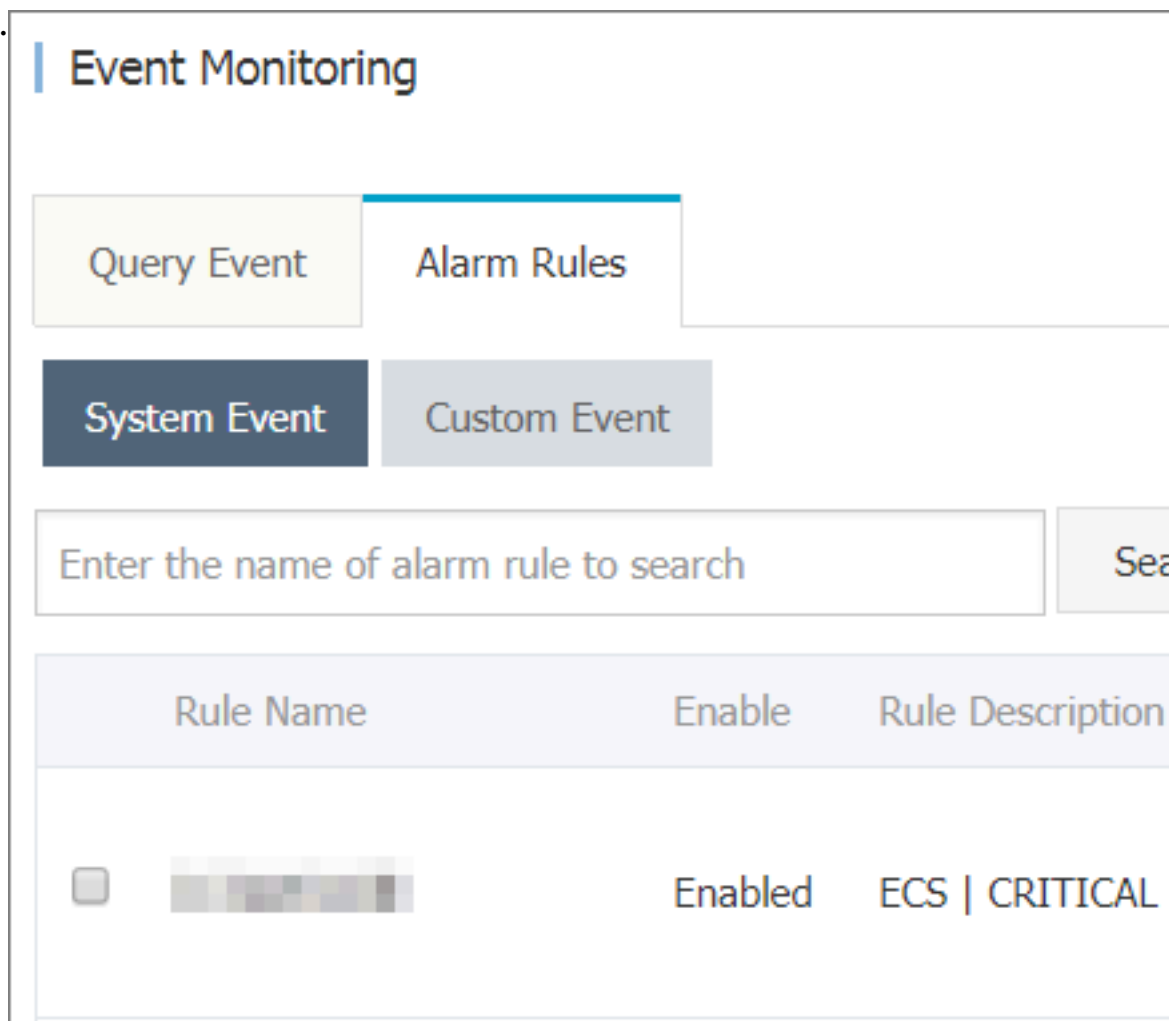


The screenshot shows a configuration window titled "Alarm type". It contains four radio button options: "Alarm notification" (which is selected), "MNS queue", "Function service", and "URL callback". Below the "Alarm notification" option, there are two dropdown menus. The first is labeled "Contact Group" and shows "Default Contact Group". The second is labeled "Notification Method" and shows "Warning (Message+Email ID+ Ali WangWang+DingTalk Robot)". Below these dropdowns is a "+Add" button. The "MNS queue", "Function service", and "URL callback" options are each followed by a small, empty square icon.

Test an alarm rule

To verify that the MNS queues and function service set in your alarm rules function properly, you can use the system event alarm testing function to simulate the occurrence of system events.

1. Go to the Alarm Rules tab page for event monitoring.



2. Click Test in the Actions column.
3. Select an event to be tested. Content corresponding to the event you select will be displayed. You can modify displayed fields, such as Instance ID, so on, based on your requirements.
4. Click OK, and the system will send an event based on the content, triggering the alarm notification, MNS queue, function service,

and URL callback settings that you specified in your alarm rule.

Create event test

Product Type ECS

Event Level :CRITICAL

Event Name



Content(JSON)

```
{
  "product": "ECS",
  "content": {
    "executeFinishTime": "2018-06-08T01:25:37Z",
    "executeStartTime": "2018-06-08T01:23:37Z",
    "ecsInstanceName": "timewarp",
    "eventId": "e-t4nhcpqcu8fqushpn3mm",
    "eventType": "InstanceFailure.Reboot",
    "ecsInstanceId": "i-0t4nhcpqcu8fqushpn3mm"
  },
  "resourceId": "acs:ecs:cn-hangzhou:1270676679546704:instance",
  "level": "CRITICAL",
  "instanceName": "instanceName",
  "status": "Executing",
  "name": "Instance:InstanceFailure.Reboot:Executing",
  "regionId": "cn-hangzhou"
}
```

10 Custom monitoring

10.1 Custom monitoring overview

Application scenarios

Custom monitoring allows you to customize metrics and alarm rules so that you can monitor metrics, report monitoring data, and set alarm rules with your specific requirements in mind.

Custom monitoring is different from event monitoring in that custom monitoring reports and queries time-series data that is collected periodically, whereas event monitoring only reports and queries data that is related to a singular event.

This topic discusses the procedures for operations custom monitoring including reporting, querying, and viewing monitoring data on the console, and how to set alarm rules for custom monitoring.

Procedures

- Report monitoring data.

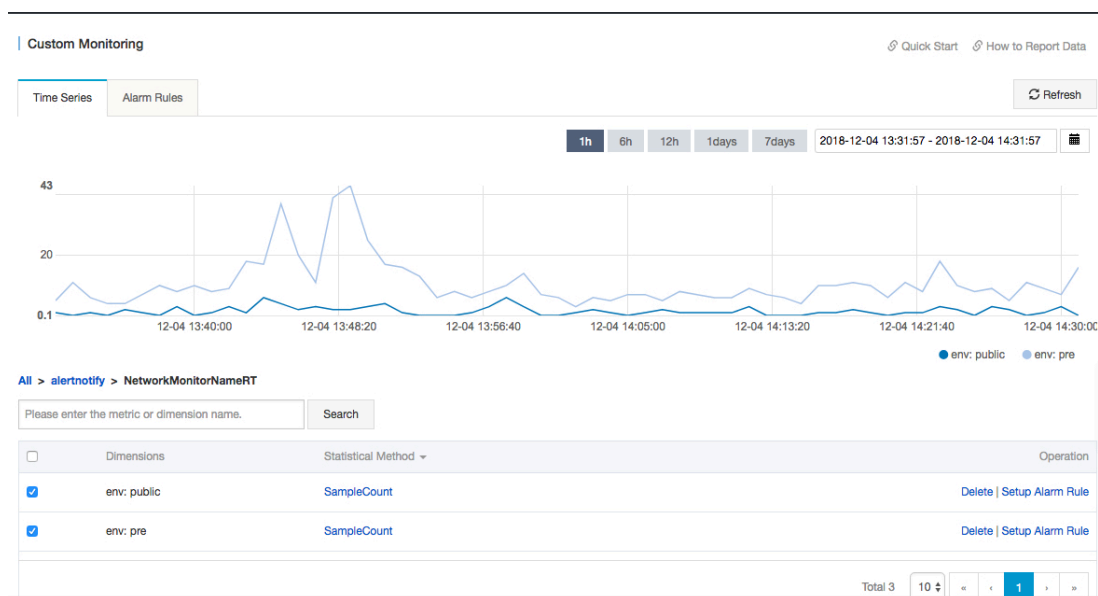
For more information and the specific procedure used, see [Report monitoring data](#).

- Query monitoring data.

After you have reported monitoring data, you can view the reported data in the console. You can choose to view all monitoring data on the custom monitoring page or to view custom monitoring data for one or more application group.

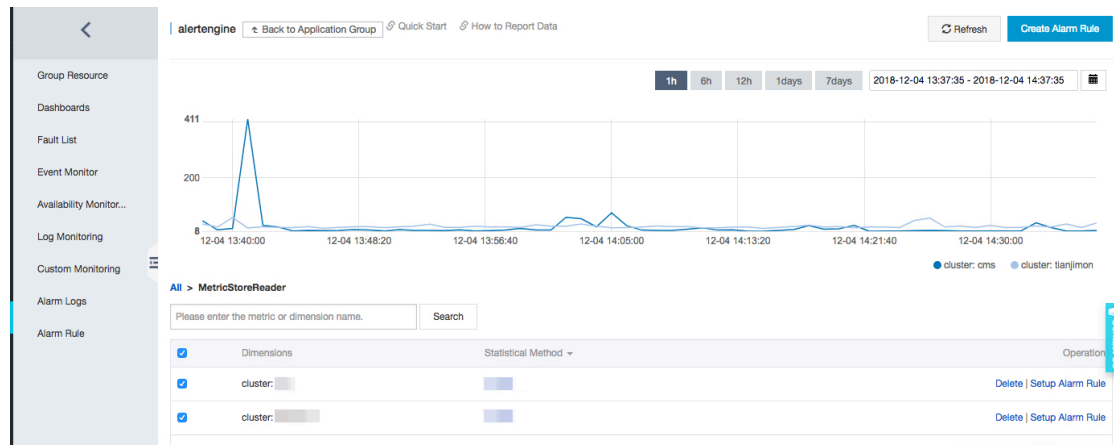
- To view all custom monitoring data, complete the following steps:

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Custom Monitoring. The Custom Monitoring page is displayed.
3. Select the corresponding application group and metric to access the Time Series page.
4. Select the time series you want to view.



- To view the custom monitoring data in an application group, complete the following steps:

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups. The Application Groups page is displayed.
3. Select the target application group.
4. Click Custom Monitoring. The Custom Monitoring page is displayed.
5. Select the target metric. The Time Series page is displayed.
6. Select the time series you want to view.



- Set an alarm rule.

Custom monitoring provides an alarm reporting feature. To set an alarm rule, you need to select an application group. When an alarm is triggered, a notification will be sent to the alarm contacts in the application group. To generate alarms for your monitoring data, set the alarm rule using either of the following two methods:

- Method 1:

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Custom Monitoring. The Custom Monitoring page is displayed.
3. Select the corresponding application group and metric. The Time Series page is displayed.
4. Select the time series for which you want to create an alarm rule, and then click Setup Alarm Rule in the Operation column.
5. On the Setup Alarm Rule page, enter a name for the alarm rule and set the corresponding alarm policy and notification method.

- Method 2:

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, click Application Groups. The Application Groups page is displayed.
3. Select the target application group. The Custom Monitoring page is displayed. Select the time series for which you want to create an alarm rule, and then click Setup Alarm Rule in the Operation column.
4. On the Setup Alarm Rule page, enter a name for the alarm rule and select the corresponding metric, dimension, alarm rule, and notification method.

10.2 Report monitoring data

Custom monitoring provides an API that you can use for reporting monitoring data. Specifically, the API allows you to report the time series data you have collected to CloudMonitor and configure alarm rules to receive alarm notifications.

CloudMonitor provides three methods for reporting data: OpenAPI, Java SDK, and Alibaba Cloud CLI.

Limits

- QPS is limited to 200 in the regions Beijing, Shanghai, and Hangzhou, 100 in the regions Zhangjiakou and Shenzhen, and 50 in all other regions.
- A maximum of 100 data entries can be reported at one time. The body size cannot exceed 256 KB.
- The `metricName` field only supports letters, numbers, and underscores. The field must start with a letter. A non-letter start is replaced with an uppercase "A" and all other invalid characters are replaced with underscores (_).
- The `dimensions` field does not support "=", "&", or ".". Invalid characters are replaced with underscores (_).
- `metricName` and `dimensions` cannot exceed 64 bytes. Otherwise, the key-value is truncated.
- Reporting raw data incurs fees. The free edition supports aggregate data reporting (that is, when reporting data, you need to pass "1" for the `type` field to the request parameter).

Report data by using OpenAPI

After you report raw data by using OpenAPI, CloudMonitor calculates the statistics in one-minute and five-minute intervals by using the following statistical methods:

- Average: the average value
- Maximum: the maximum value
- Minimum: the minimum value
- Sum: sum value
- SampleCount: count
- SumPerSecond: sum/total seconds of the corresponding aggregation period. You can also use the moving average calculation.

- **CountPerSecond**: count/total seconds of the corresponding aggregation period. You can also use the moving average calculation.
- **LastValue**: the last sampling value in the aggregation period, which is similar to gauge.
- **P10**: percentile 0.1, greater than 10% of all sampling data in the aggregation period
- **P20**: percentile 0.2, greater than 20% of all sampling data in the aggregation period
- **P30**: percentile 0.3, greater than 30% of all sampling data in the aggregation period
- **P40**: percentile 0.4, greater than 40% of all sampling data in the aggregation period
- **P50**: percentile 0.5, greater than 50% of all sampling data in the aggregation period, also known as the median
- **P60**: percentile 0.6, greater than 60% of all sampling data in the aggregation period
- **P70**: percentile 0.7, greater than 70% of all sampling data in the aggregation period
- **P75**: percentile 0.75, greater than 75% of all sampling data in the aggregation period
- **P80**: percentile 0.8, greater than 80% of all sampling data in the aggregation period
- **P90**: percentile 0.9, greater than 90% of all sampling data in the aggregation period
- **P95**: percentile 0.95, greater than 95% of all sampling data in the aggregation period
- **P98**: percentile 0.98, greater than 98% of all sampling data in the aggregation period
- **P99**: percentile 0.99, greater than 99% of all sampling data in the aggregation period
- **Service addresses**

Internet address of the service: <https://metrichub-cms-cn-hangzhou.aliyuncs.com>

The intranet addresses of the service is as follows:

Region	RegionId	Endpoints
China East 1 (Hangzhou)	cn-hangzhou	http://metrichub-cn-hangzhou.aliyun.com
China North 3 (Zhangjiakou)	cn-zhangjiakou	http://metrichub-cn-zhangjiakou.aliyun.com
China East 2 (Shanghai)	cn-shanghai	http://metrichub-cn-shanghai.aliyun.com

Region	RegionId	Endpoints
China North 2 (Beijing)	cn-beijing	http://metrichub-cn-beijing.aliyun.com
China North 1 (Qingdao)	cn-qingdao	http://metrichub-cn-qingdao.aliyun.com
China South 1 (Shenzhen)	cn-shenzhen	http://metrichub-cn-shenzhen.aliyun.com
Hong Kong (China)	cn-hongkong	http://metrichub-cn-hongkong.aliyun.com
China North 5 (Hohhot)	cn-huhehaote	http://metrichub-cn-huhehaote.aliyun.com
Middle East 1 (Dubai)	me-east-1	http://metrichub-me-east-1.aliyun.com
US West 1 (Silicon valley)	us-west-1	http://metrichub-us-west-1.aliyun.com
US East 1 (Virginia)	us-east-1	http://metrichub-us-east-1.aliyun.com
Asia Pacific NE 1 (Tokyo)	ap-northeast-1	http://metrichub-ap-northeast-1.aliyun.com
EU Central 1 (Frankfurt)	eu-central-1	http://metrichub-eu-central-1.aliyun.com
Asia Pacific SE 2 (Sydney)	ap-southeast-2	http://metrichub-ap-southeast-2.aliyun.com
Asia Pacific SE 1 (Singapore)	ap-southeast-1	http://metrichub-ap-southeast-1.aliyun.com
Asia Pacific SE 3 (Kuala Lumpur)	ap-southeast-3	http://metrichub-ap-southeast-3.aliyun.com
Asia Pacific SOU 1 (Mumbai)	ap-south-1	http://metrichub-ap-south-1.aliyuncs.com

• Request syntax

```
POST /metric/custom/upload HTTP/1.1
Authorization:<AuthorizationString>
Content-Length:<Content Length>
Content-MD5:<Content MD5>
Content-Type:application/json
Date:<GMT Date>
Host: metrichub-cms-cn-hangzhou.aliyuncs.com
x-cms-signature:hmac-sha1
x-cms-api-version:1.0
x-cms-ip:30.27.84.196
```

```
User-Agent:cms-java-sdk-v-1.0
[{"groupId":101,"metricName":"","dimensions":{"sampleName1":"value1",
"sampleName2":"value2"},"time":"","type":0,"period":60,"values":{"
value":10.5,"Sum":100}}]
```

- Signature algorithm

For more information, see [Signature algorithm](#).

- Request parameters

Name	Type	Required?	Description
groupId	long	Yes	ID of an application group
metricName	string	Yes	Name of a monitoring metric . The name can be up to 64 bytes in length and can contain letters , numbers, and connectors "_-./\". Characters that exceed the limit are truncated.
dimensions	object	Yes	Dimension map . The key-value is a string where the key and the value can be up to 64 bytes in length separately, and can contain letters , numbers, and connectors "_-./\". Characters that exceed the limit are truncated. The maximum number of key-value pairs is 10.

Name	Type	Required?	Description
time	string	Yes	Time of the metric data. It supports "yyyyMMdd'T'HHmmss.SSSZ" and long format timestamps, for example, 20171012T132456.888+0800 or 1508136760000.
type	int	Yes	Type of the reported data . 0 represents raw data, and 1 indicates aggregate data. When you report aggregate data, we recommend that you report the data in both 60s and 300s aggregation periods. Otherwise , you will not be able to query monitoring data that is older than seven days.
period	string	No	Aggregation period in seconds. If type=1, this field is required. The value can be 60 or 300.

Name	Type	Required?	Description
values	object	Yes	Collection of metric values . If type=0, the key must be "value" and raw data is reported . CloudMonitor aggregates raw data into multiple types of data on the basis of aggregation periods, for example, maximum, count, and sum.

Report data by using the Java SDK

- Install the Java SDK

When you install the Java SDK through Maven, the following dependencies must be added:

```
<dependency>
  <groupId>com.aliyun.openservices</groupId>
  <artifactId>aliyun-cms</artifactId>
  <version>0.2.4</version>
</dependency>
```

- Response element

The system returns the HTTP status code 200.

- Examples

- Request example

```
POST /metric/custom/upload HTTP/1.1
Host: metrichub-cms-cn-hangzhou.aliyuncs.com
x-cms-api-version:1.0
Authorization:yourAccessKeyId:yourAccessKeySecret
Host:metrichub-cms-cn-hangzhou.aliyuncs.com"
Date:Mon, 23 Oct 2017 06:51:11 GMT
Content-Length:180
x-cms-signature:hmac-sha1
Content-MD5:E9EF574D1AEAAA370860FE37856995CD
x-cms-ip:30.27.84.196
User-Agent:cms-java-sdk-v-1.0
Content-Type:application/json
```

```
[{"groupId":101,"metricName":"","dimensions":{"sampleName1":"value1","sampleName2":"value2"},"time":"","type":0,"period":60,"values":{"value":10.5,"Sum":100}}]
```

- **Response example**

```
{
  "code":"200",
  "msg":""//The returned msg is null when the reporting is normal
}
```

- **Code example**

- **Report raw data**

```
CMSClinit.groupId = 101L;//Set a common group ID.
CMSClinit cmsClient = new CMSClinit(endpoint, accKey,
secret);//Initialize the client.
CustomMetricUploadRequest request = CustomMetricUploadRe
quest.builder()
    .append(CustomMetric.builder()
        .setMetricName("testMetric")//Metric name
        .setGroupId(102L)//Set a custom group ID.
        .setTime(new Date())
        .setType(CustomMetric.TYPE_VALUE)//The
type is raw data.
        .appendValue(MetricAttribute.VALUE, 1f)//
The key must be this when the type is raw data.
        .appendDimension("key", "value")//Add a
dimension.
        .appendDimension("ip", "127.0.0.1")//Add a
dimension.
        .build())
    .build();
CustomMetricUploadResponse response = cmsClient.putCustomM
etric(request);//Report data.
```

```
System.out.println(JSONObject.toJSONString(response));
```

- Automatically report aggregate data of multiple aggregation periods

SDK supports data reporting after local aggregation. The aggregation periods are one minute and five minutes.

Data type	Description	Aggregated value	Memory usage (excluding names, dimensions, individual time series, and individual aggregation periods)
Value	Typical value	All properties except LastValue	About 4 KB
Gauge	Sample value	LastValue	4 bytes
Meter	Sum and speed	Sum, SumPerSecond	50 bytes
Counter	Count	SampleCount	10 bytes
Timer	Computing time	SampleCount, CountPerSecond, Average, Maximum, Minimum, PXX(P10-P99)	About 4 KB
Histogram	Distribution	SampleCount, Average, Maximum, Minimum, PXX(P10-P99)	About 4 KB

```
//Initialize
    CMSClientInit.groupId = 0L;
    CMSClient cmsClient = new CMSClient(accKey, secret,
endpoint);//Create a client.
    CMSMetricRegistryBuilder builder = new CMSMetricRegistryBuilder();
    builder.setCmsClient(cmsClient);
    final MetricRegistry registry = builder.build();//Create a
registry which contains two aggregation periods.
```



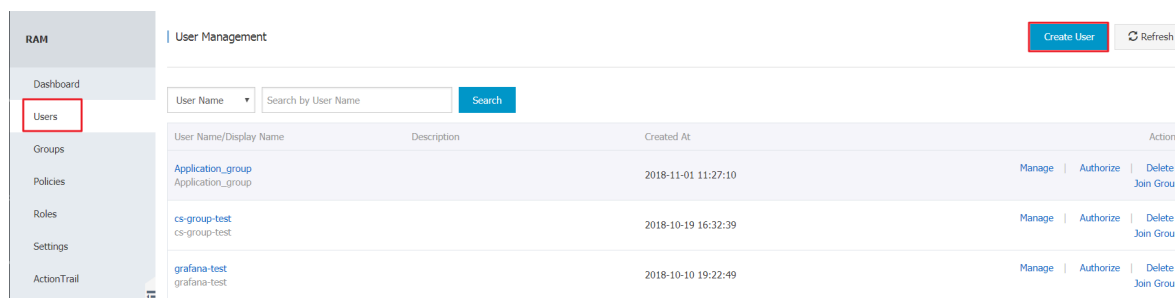
```
//or final MetricRegistry registry = builder.build(
RecordLevel._60S); //Create a registry which contains only the one-
minute aggregation period.
//Use value.
ValueWrapper value = registry.value(MetricName.build("value"));
value.update(6.5);
//Use meter.
MeterWrapper meter = registry.meter(MetricName.build("meter"));
meter.update(7.2);
//Use counter.
CounterWrapper counter = registry.counter(MetricName.build("
counter"));
counter.inc(20);
counter.dec(5);
//Use timer.
TimerWrapper timer = registry.timer(MetricName.build("timer"));
timer.update(30, TimeUnit.MILLISECONDS);
//Use histogram.
HistogramWrapper histogram = registry.histogram(MetricName.build("
histogram"));
histogram.update(20);
//Use gauge.
final List list = new ArrayList();
registry.gauge(MetricName.build("gauge"), new Gauge() {
    @Override
    public Number getValue() {
        return list.size();
    }
});
```

Report data by using Alibaba Cloud CLI

Prepare your Alibaba Cloud account

Make sure that you have created a RAM user under your Alibaba Cloud account and have generated a RAM user Access Key (AK) with CloudMonitor permissions.

- Create a RAM user



- Generate an AccessKeyID and AccessKeySecret for the RAM user.

The screenshot shows the 'User Details' page for a RAM user named 'test123'. The 'User Access Key' section is expanded, showing a table with columns 'AccessKey ID', 'Status', and 'Created At'. A red box highlights the 'Create Access Key' button in the top right corner of this section.

- Grant CloudMonitor permissions to the RAM user.

The screenshot shows the 'Edit User-Level Authorization' dialog box. It contains a table of 'Available Authorization Policy Names' with columns 'Policy Name' and 'Type'. The 'CloudMonitor' policy is highlighted in green, and the 'AliyunCloudMonitorFullAccess' policy is highlighted in red. A red box also highlights the 'Edit Authorization Policy' button in the top right corner of the dialog.

Install Alibaba Cloud CLI

System requirement: Linux, UNIX, or Mac OS

Environment requirement: You have installed Python 2.7.x.

1. Install Python.

- If you have installed Python 2.7.x, skip this step.
- Otherwise, run the following command in your command line interface:



Note:

Make sure that you have installed wget.

```
wget https://www.python.org/ftp/python/2.7.8/Python-2.7.8.tgz (or
download it in some other way and put it in a certain path)
tar -zxvf Python-2.7.8.tgz
cd Python-2.7.8
./configure
make
```

```
sudo make install
```

2. Install pip.

- Run the following command in your command line interface:



Note:

If you have installed pip, you can skip this step.

```
curl "https://bootstrap.pypa.io/get-pip.py" -o "pip-install.py"
sudo python pip-install.py
```

- If the system displays the following or similar information, the installation is successful.

```
Successfully installed pip-7.1.2 setuptools-18.7 wheel-0.26.0
```

3. Install Alibaba Cloud CLI.

- An earlier version of pip can cause installation failure of Alibaba Cloud CLI. Make sure that you use pip 7.x or a later version. You can run the following command in your command line interface to upgrade pip:

```
sudo pip install -U pip
```

If the system displays the following or similar information, the upgrade is successful.

```
Successfully uninstalled pip-7.1.2
Successfully installed pip-8.1.2
```

- Run the following command to install Alibaba Cloud CLI:

```
sudo pip install aliyuncli
```

If the system displays the following or similar information, the installation is successful.

```
Successfully installed aliyuncli-2.1.2 colorama-0.3.3 jmespath-0.7.1
```

4. Configure Alibaba Cloud CLI. Run the following command to configure Alibaba Cloud CLI:

```
~ sudo aliyuncli configure
Aliyun Access Key ID [*****a]: youraccesskeyid
Aliyun Access Key Secret [*****b]: youraccesskeysecret
Default Region Id [cn-hangzhou]: cn-hangzhou
```

```
Default output format [json]: json
```

Install CMS SDK

- The installation method for a Windows system is as follows:

```
cd C:\Python27\Scripts
pip install aliyun-python-sdk-cms
```

- To update the SDK in a Windows system, use the following command:

```
pip install --upgrade aliyun-python-sdk-cms
```

- The installation method for a Linux system is as follows:

```
sudo pip install aliyun-python-sdk-cms
```

- To update the SDK in a Linux system, use the following command:

```
sudo pip install --upgrade aliyun-python-sdk-cms
```

Report monitoring data

Use the `PutCustomMetric` interface to report monitoring data.

- Example for a Windows system

```
aliyuncli.exe cms PutCustomMetric --MetricList "[{'groupId':1,'metricName':'testMetric','dimensions':{'sampleName1':'value1','sampleName2':'value2'},'type':0,'values':{'value':10.5}}]"
```

- Example for a Linux system

```
aliyuncli cms PutCustomMetric --MetricList "[{'groupId':1,'metricName':'testMetric','dimensions':{'sampleName1':'value1','sampleName2':'value2'},'type':0,'values':{'value':10.5}}]"
```

- If the data is reported successfully, status code 200 is returned.

```
{
  "Code": "200"
}
```

Status codes

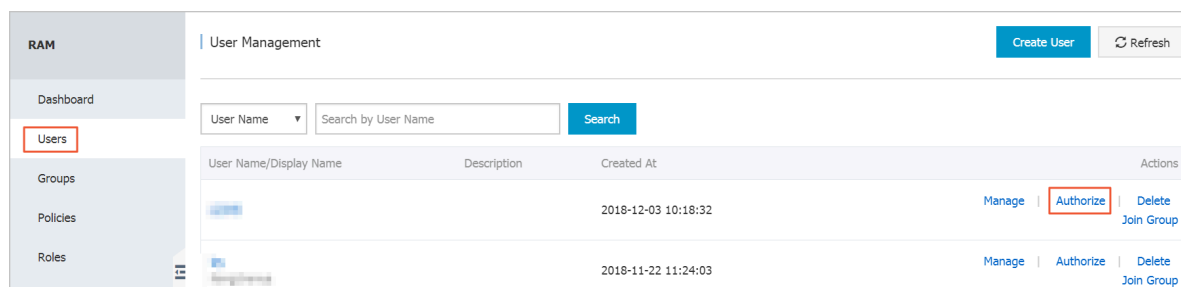
Status code	Description
200	Normal

Status code	Description
206	<p>Partially successful.</p> <p>If "reach max time series num" is returned, it indicates that your time series quota has run out. We recommend that you pay for a higher quota or remove unnecessary time series.</p> <p>If "not allowed original value, please upgrade service" is returned, it indicates that you are using the free edition, which does not support raw data reporting.</p> <p>If "type is invalid" is returned, it indicates that the type value is invalid. Please check if a number other than 0 or 1 is passed in.</p>
400	Syntax errors in the client request
403	Verification failure, speed limit, or not authorized
500	Internal server error

RAM user authorization

You must grant CloudMonitor permissions to the corresponding RAM user before you can report event data by using the RAM user AK. If you do not grant the permissions, when you report data, the prompt "cannot upload, please use ram to auth" is displayed.

1. Log on to the [RAM console](#).
2. In the left-side navigation pane, click Users.
3. Find the target user and click Authorize.



4. On the authorization page, select the AliyunCloudMonitorFullAccess policy and click OK.

Members added to this group have all the permissions of this group. A member cannot be added to the same group more than once.

Available Authorization Policy Names	Type
CloudMonitor	
AliyunCloudMonitorReadOnlyAcce...	System

Selected Authorization Policy Name

Selected Authorization Policy Name	Type
AdministratorAccess	System
Provides full acce...	
AliyunCloudMonitorFullAccess	System
Provides full acce...	

OK Close

10.3 Configure a dashboard

After reporting monitoring data to custom monitoring, you can create a dashboard for easy monitoring and data queries.

- Create a dashboard

1. Log on to the [CloudMonitor console](#).
2. In the left-side navigation pane, choose Dashboard > Custom Dashboard.
3. Click Create Dashboard in the upper-right corner, enter a dashboard name, and click Create to create a dashboard.

Create Dashboard

test_dashboard

Create Close

- Add a chart

1. Click Add Chart in the upper-right corner.
2. In the Select Metrics area, click the Custom tab and enter a chart name.
3. Select the target metric, statistical method, and dimensions.
4. Click Save to save your settings.

2 Select Metrics

Dashboards


Log Monitoring

Custom

Custom Monitoring

online-application

Heat Map Gradient Range: 0 auto



127.24
100.00
50.00
0.00

19:29:00 19:35:00 19:43:20 19:51:40 20:00:00 20:08:20 20:16:40 20:27:00

● MetricStoreReader—Average—cluster:cms

Metrics: MetricStoreReader Average

Dimensions: cluster:cms

+AddMetrics

Save Cancel