Application Performance Management

User Guide

Issue	01
Date	2021-01-01

HUAWEI TECHNOLOGIES CO., LTD.

Trademarks and Permissions

Notice

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

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

Huawei Technologies Co., Ltd.

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

Website: https://www.huawei.com

Email: support@huawei.com

Contents

1 Overview

What Is APM?

Currently, user experience has become one of core competences of applications. With the increasing application complexity and increasing number of users, application O&M face huge challenges at normal application assurance, fast fault locating, and performance bottleneck identification.

Application Performance Management (APM) monitors and manages the performance of cloud applications in real time. APM provides performance analysis of distributed applications, helping O&M personnel quickly locate and resolve faults and performance bottlenecks.

APM is a cloud application diagnosis service and supports applications based on multiple Java frameworks. It includes powerful analytic tools, displays application status, call process, and operations performed on applications through topology views, tracing, and transactions. This helps you quickly locate faults and performance bottlenecks.

APM Architecture Features

With the emergence of new technologies and methods, enterprises have urgent demands for fast and agile compatibility support, and need to monitor and analyze applications in multi-layer, complex, and hybrid architectures. APM provides automatic and real-time monitoring and analysis capabilities in new IT architectures such as mobile, on-cloud, and distributed systems. It supports proactive O&M and auxiliary optimization to ensure consistent user experience.

To shield the impact of technical changes on the upper-layer computing and storage layer architecture, APM uses the multi-layer decoupling, lightweight, independent extension, and frame-irrelevant data collection access layer to encapsulate bottom-layer technical details and to provide reliable and stable data analysis formats to the upper layer. In this way, the computing and storage layer, and the presentation layer can focus on analysis, calculation, and display of stable data.

2 Basic Concepts

Topology

A topology graphically displays call and dependency relationships between applications. It is composed of circles, arrows, and resources. Each circle represents an application, and each section in the circle represents an instance. The fraction in each circle indicates number of active instance/total number of instances. The data below the fraction indicates the **service latency**, call count, and error count. Each line with an arrow represents a call relationship. Thicker arrows indicate more calls. The values above a line separately indicate the throughput and **overall latency**. Throughput is the number of calls within the selected period. **Application Performance Index (Apdex)** is used in the topology to quantify user satisfaction with application performance. Different colors indicate different **Apdex** ranges, helping you quickly detect and locate faults.

Transaction

A transaction is usually an HTTP request. The process is as follows: user request > web server > database > web server > user request. Transactions are one-off tasks, which are completed by using applications. In the example of an e-commerce application, querying a product is a transaction, and making a payment is also a transaction.

Tracing

APM traces and records service calls, and visually presents the execution tracks and statuses of service requests in distributed systems, so that you can quickly locate performance bottlenecks and faults.

Application

You can put the same type of services into an application for better performance management. For example, you can put accounts, products, and payment applications into the **Mall** application.

Apdex

Apdex is an open standard developed by the Apdex alliance to measure application performance. The application response time is converted into user satisfaction with application performance. The Apdex value ranges from 0 to 1.

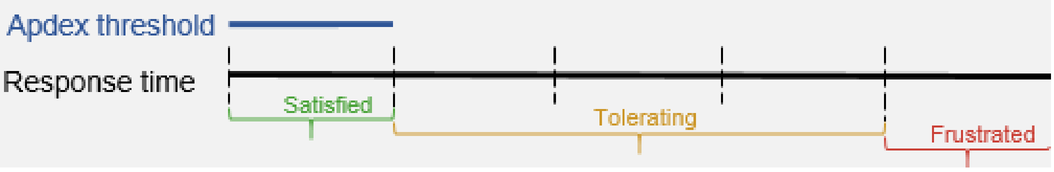
- Apdex principles

Apdex defines the optimal threshold "T" for the application response time. "T" is determined based on performance expectations. Based on the actual response time and "T", user experience can be categorized as follows:

Satisfied: indicates that the actual response time is shorter than or equal to "T". For example, if "T" is 1.5s and the actual response time is 1s, user experience is satisfied.

Tolerating: indicates that the actual response time is greater than "T", but shorter than or equal to "4T". For example, if "T" is 1s, the tolerable upper threshold for the response time is 4s.

Frustrated: indicates that the actual response time is greater than "4T".



- Apdex calculation method

In APM, "T" is the threshold set in **Configuring Apdex Thresholds**, the application response latency equals to the total service latency, and the Apdex value ranges from 0 to 1. The calculation formula is as follows:

Apdex = (Number of normal calls x 1 + Number of slow calls x 0.5)/Total number of calls

In the preceding information:

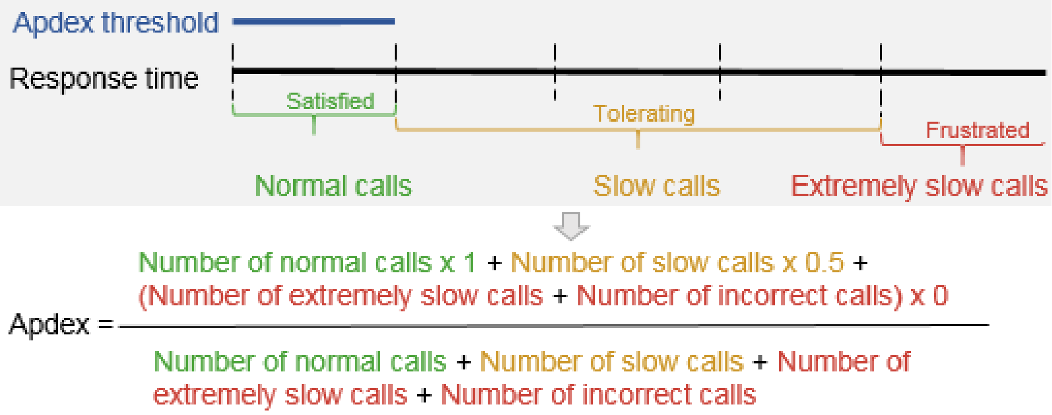
Number of normal calls: indicates the number of successful calls that are completed within a time period of greater than 0 but less than "T".

Number of slow calls: indicates the number of successful calls that are completed within a time period of greater than or equal to "T" but less than "4T".

Number of extremely slow calls: indicates number of successful calls that are completed within a time period of greater than "4T".

Total number of calls: indicates the total number of normal calls, slow calls, extremely slow calls, and incorrect calls.

The Apdex calculation formula is as follows:


Apdex threshold
Response time
Satisfied
Tolerating
Frustrated
Normal calls
Slow calls
Extremely slow calls
Apdex =
Number of normal calls x 1 + Number of slow calls x 0.5 + (Number of extremely slow calls + Number of incorrect calls) x 0
Number of normal calls + Number of slow calls + Number of extremely slow calls + Number of incorrect calls

Apdex value indicates application performance status, that is, user satisfaction with application performance. Different Apdex values are marked by different colors. For details, see **Table 2-1**.

Table 2-1 Apdex values

Apdex Value	Color	Description
0.75 ≤ Apdex ≤ 1	Green	Fast response; good user experience
0.3 ≤ Apdex < 0.75	Yellow	Slow response; fair user experience
0 ≤ Apdex < 0.3	Red	Very slow response; poor user experience

Configuring an Apdex threshold

You can configure the Apdex threshold based on your service requirements. For details, see **Configuring Apdex Thresholds**.

TP99 Latency

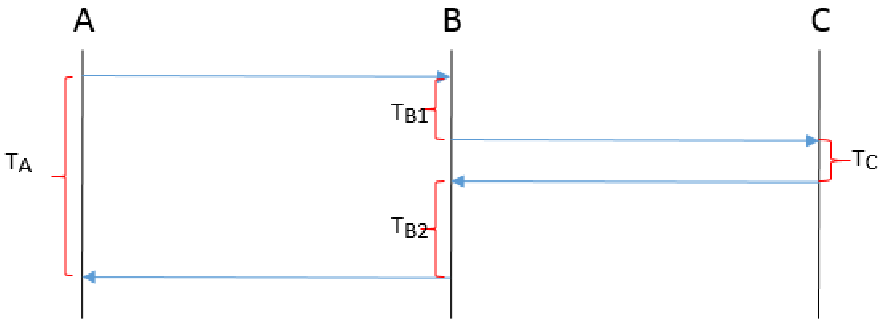
TP99 latency is the minimum time for meeting requirements of 99% requests. In APM, latency refers to TP99 latency.

For example, the time required for processing four requests is 10 ms, 100 ms, 500 ms, and 20 ms respectively.

In the four requests, the number of 99% requests can be calculated by multiplying 4 by 99%, and the rounding value is 4. That is, the number of 99% requests is 4. The minimum time required for the four requests is 500 ms. Therefore, TP99 latency is 500 ms.

Overall Latency/Service Latency

Latency refers to the period from initiating a request to getting a response. In APM, the overall latency refers to the total time consumed by a request, and the service latency refers to the time consumed by a service. For example, assume that service A calls service B, and service B calls service C, as shown in the following figure:



Overall latency = T_A; Latency of service A = T_A; Latency of service B = T_{B1} + T_{B2}; Latency of service C = T_C

Collection Probe

Probes use the bytecode enhancement technology to track calls and generate data. The data will be collected by the ICAgent and then displayed on the UI. If the memory monitoring mechanism is enabled and the instance memory usage is too high, probes enter the hibernation state and stop data collection. For details about the types of data collected by probes, see **Scope and Usage**.

ICAgent

ICAgent is the collection agent of APM. It runs on the server where applications are deployed and collects data obtained by probes in real time. Before using APM, ensure that the ICAgent is installed according to **Installing the ICAgent**.

3 Usage Restrictions

Supported OSs

APM supports multiple operating systems (OSs). When creating an Elastic Cloud Server (ECS), select an OS supported by APM. For details, see **Table 3-1**.

Table 3-1 Supported OSs and versions

OS	Supported Version	Description
SUSE	SUSE Enterprise 12 SP1 64-bit SUSE Enterprise 12 SP2 64-bit SUSE Enterprise 11 SP4 64-bit	-
openSUSE	13.2 64-bit 42.2 64-bit	-
EulerOS	2.2 64-bit	-
CentOS	7.4 64-bit 7.3 64-bit 7.2 64-bit 7.1 64-bit 6.9 64-bit 6.8 64-bit 6.5 64-bit 6.3 64-bit	-

OS	Supported Version	Description
Ubuntu	14.04 server 64-bit 16.04 server 64-bit	-
CoreOS	10.10.5 64-bit	-
Fedora	24 64-bit	The 25 64-bit version has been planned and is being tested.
Debian	To be supported	The 7.5.0 32-bit and 7.5.0 64-bit versions have been planned and are being tested.

Supported Types

Currently, APM can connect to only Java applications. APM supports mainstream Java frameworks, web servers, communication protocols, and databases. For details about the supported types, see **Table 3-2**.

Table 3-2 Supported types

Type	Name	Version
Tool	JDK	JDK 7 and JDK 8
Communication protocol	HTTP client	Apache HttpClient 3, Apache HttpClient 4, and JDK HttpURLConnection
Java framework	CXF Client	2.6.0–3.2.1
	iBatis	2.3.0 and 2.3.4.726
	Jersey	2.0–2.9.1
	MyBatis	1.0.0–1.3.1 (MyBatis-Spring) and 3.0.1–3.4.5 (MyBatis 3)
	Spring	3.1.x–5.0.x
	Spring Boot	1.2.x–1.5.x
	Dubbo	2.5.3–2.5.4 (Dubbo RPC and Dubbo REST)
	CSE	0.4–0.5 (REST over Servlet, REST over Vertx, and Highway RPC)
Database	MySQL	mysql-connector-java 5.1.x
	Oracle	ojdbc5, ojdbc6, and ojdbc14
	Sybase	2.6.0–3.2.1

Type	Name	Version
	MariaDB	1.3.x
	VoltDB	6.x–7.x
	PostgreSQL	9.0.x, 9.1.x, 9.2.x, 9.3.x, 9.4.x, 42.0.x, and 42.1.x
Web server	Tomcat	6.x, 7.x, and 8.x
	Jetty	7.6.x–8.0.0 and 8.1.x–9.x.x
	JBoss	7.0.0–7.1.3 and 7.2.0
	Undertow	1.4.x
Message queue	ActiveMQ	5.6.x–5.15.x
	RocketMQ	4.1.x–4.2.x
	RabbitMQ	1.3.3 and later (spring-rabbit), 2.7.x (amqp-client), 2.6.0, and 3.6.5
	Kafka	0.9.0.1–0.10.0.2
NoSQL	Redis	Jedis 2.0.0–2.9.0
	Memcache	2.9.0–2.12.3 (Arcus)
	MongoDB	3.0.x–3.6.x
	Casandra	2.1.x–3.2.x
	ZooKeeper	1.0.x (com.github.adyliu.zkclient) and 0.1.x (com.github.sgroschupf.zkclient)
	Elasticsearch	2.4.x and 5.1.x
REST Client	Common HTTP	2.x, 3.x, 4.x (httpclient), and ALL (HttpURLConnection)

NOTE

More types are being developed.

4 Permissions Management

If you need to assign different permissions to employees in your enterprise to access your Application Performance Management (APM) resources, Identity and Access Management (IAM) is a good choice for fine-grained permissions management. IAM provides identity authentication, permissions management, and access control, helping you secure access to your cloud resources.

With IAM, you can use your account to create IAM users for your employees, and assign permissions to the users to control their access to specific resources. For example, some software developers in your enterprise need to use APM resources but must not delete them or perform any high-risk operations. To achieve this result, you can create IAM users for the software developers and grant them only the permissions required for using APM resources.

If your account does not need individual IAM users for permissions management, you may skip over this chapter.

IAM can be used free of charge. You pay only for the resources in your account.

APM Permissions

By default, new IAM users do not have any permissions assigned. You need to add a user to one or more groups, and assign permissions policies or roles to these groups. The user then inherits permissions from the groups it is a member of. This process is called authorization. After authorization, the user can perform specified operations on APM.

APM is a project-level service deployed and accessed in specific physical regions. To assign APM permissions to a user group, specify the scope as region-specific projects and select projects for the permissions to take effect. If **All projects** is selected, the permissions will take effect for the user group in all region-specific projects. When accessing APM, the users need to switch to a region where they have been authorized to use this service.

Table 4-1 lists all the system permissions supported by APM.

Table 4-1 System permissions supported by APM

Role	Description	Category
APM FullAccess	Full permissions for APM	System-defined policy
APM ReadOnlyAccess	Read-only permissions for APM	System-defined policy
APM Administrator	Full permissions for APM	System-defined role

Table 4-2 lists the common operations supported by each system-defined policy or role of APM. Choose appropriate policies or roles as required.

Table 4-2 Common operations supported by each system-defined policy or role of APM

Operation	APM FullAccess	APM ReadOnlyAccess	APM Administrator
Obtaining application topology information	√	√	√
Modifying application topology configuration	√	x	√
Deleting application topology configuration	√	x	√
Adding application topology configuration	√	x	√
Obtaining slow SQL analysis results	√	√	√
Obtaining tracing data	√	√	√
Updating tracing configuration	√	x	√
Querying APM configuration	√	√	√

Operation	APM FullAccess	APM ReadOnlyAccess	APM Administrator
Adding APM configuration	√	x	√
Deleting APM configuration	√	x	√
Querying the ICAgent list	√	√	√
Installing the ICAgent	√	x	√
Querying the ICAgent version	√	√	√
Upgrading the ICAgent version	√	x	√
Uninstalling the ICAgent	√	x	√
Delivering an ICAgent event	√	x	√

5 Billing

Package Details

The following table lists the functions supported by different editions of probe products.

Edition	Basic	Enterprise
Version description	50 times; one hour each time If you use APM for more than one hour each time, APM automatically stops data collection and you can only query historical data on the page. You need to manually apply to continue using the free edition or switch to the enterprise edition (paid edition). If you use APM 50 times, you need to switch to the enterprise edition (paid edition).	Open
Data storage duration	7 days	7 days
Application topology	√	√
Tracing	√	√
Transaction analysis	√	√
Slow SQL analysis	√	√
JVM analysis	√	√
AI capability	√	√
Method tracing	√	√

6 Deploying Applications

This section describes how to deploy applications for performance management.

You need to perform operations based on application deployment modes. Currently, APM supports application deployment through:

- ServiceStage. For details, see **ServiceStage Mode**.
- Application Orchestration Service (AOS). For details, see **AOS Mode**.
- Cloud Container Engine (CCE). For details, see **CCE Mode**.
- Elastic Cloud Server (ECS) or Bare Metal Server (BMS) without using AOS, ServiceStage, or CCE. For details, see **VM Mode**.

ServiceStage Mode

ServiceStage is a one-stop DevOps platform service oriented for enterprises and developers. If you select probes when using ServiceStage to create or release applications, APM is automatically connected to the applications. After the applications run for about three minutes, log in to the APM console to view the application information on the **Topology** and **Transactions** pages.

AOS Mode

For AOS, when you add the designer pinpoint to templates during compilation, APM collection probes are added to stacks. After templates are compiled and stacks are created, APM is automatically connected to stack applications. After the stacks run for about three minutes, log in to the APM console to view the application information on the **Topology** and **Transactions** pages.

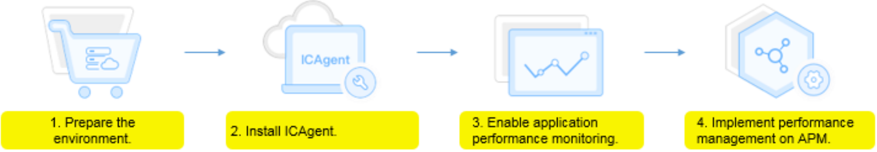
CCE Mode

CCE provides container application management services. If you select probes when creating or upgrading applications, APM collection probes are installed on the applications. After the applications run for about three minutes, log in to the APM console to view the application information on the **Topology** and **Transactions** pages.

VM Mode

Before deploying applications on the ECS or BMS without using AOS, ServiceStage, or CCE, you need to learn the process and prerequisites in advance.

Process


ICAgent
1. Prepare the environment.
2. Install ICAgent.
3. Enable application performance monitoring.
4. Implement performance management on APM.

1. Prepare the environment: Creating a VM and ensure that the applications to be monitored are running properly.
2. Install the ICAgent: Collect application data in real time.
3. Enable application performance monitoring: Modify the startup scripts of the applications to ensure that the ICAgent can collect application data.
4. Implement performance management on APM: After the applications run for about three minutes, APM will automatically discover the application topology and performance data. Then, you can view the topology and tracing data on the APM console.

Prerequisites

An ECS server is available. To use APM, ensure that the following conditions are met:

1. The operating system (OS) of the ECS server and the application type are supported by APM.
2. The ECS server where applications are deployed has been bound to an Elastic IP Address (EIP). For a cluster with multiple ECS servers, ensure that at least one ECS server is bound to an EIP.
3. The Access Key ID/Secret Access Key (AK/SK) have been obtained. The AK/SK are used to install the ICAgent. For details, see **Obtaining the AK/SK**.
4. The time and time zone used by the local browser must be consistent with those of the ECS server.

Procedure

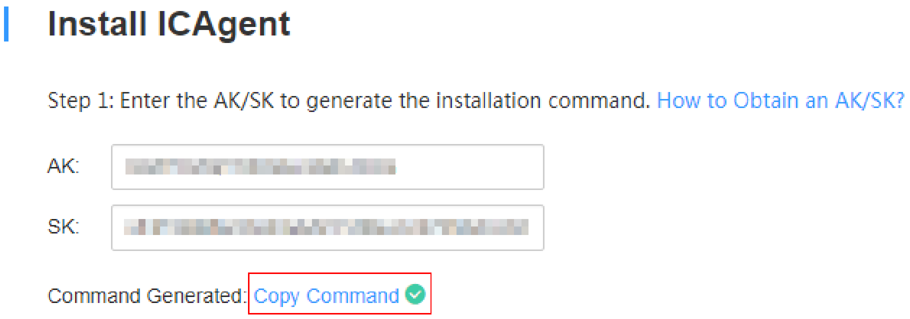
Step 1 Install the ICAgent.

1. Log in to the APM console, choose **Agent** > **Management** in the navigation pane, and click **Install ICAgent**.
2. Generate the ICAgent installation command and copy it.
   a. Enter the obtained AK/SK in the text box to generate the ICAgent installation command.

      NOTE

      Ensure that the AK/SK are correct. Otherwise, the ICAgent cannot be installed.

b. Click **Copy Command**.



Install ICAgent

Step 1: Enter the AK/SK to generate the installation command. How to Obtain an AK/SK?

AK: [redacted]

SK: [redacted]

Command Generated: Copy Command ✓

c. Log in to the ECS server using the EIP as a **root** user through a remote login tool and run the copied command to install the ICAgent.

If the message "ICAgent install success" is displayed, the ICAgent is successfully installed in the **/opt/oss/servicemgr**/ directory.

3. (Optional) When you have multiple ECS servers and have installed the ICAgent on one of the servers, use the inherited installation method to install the remaining servers. For details, see **Inherited Installation**.

Step 2 Enable application performance monitoring.

The **vmall** application and **vmall-product-service** application microservice are used as examples.

After installing the ICAgent, perform the following operations to enable application performance monitoring. After it is enabled, non-intrusive probes are installed on your Java applications.

- For Java applications that are not deployed using JBoss, add the following configuration to the startup scripts of the server where Java applications locate: After the configuration, start the applications for performance monitoring.

  -javaagent:/opt/oss/servicemgr/ICAgent/pinpoint/pinpoint-bootstrap.jar -Dapm_application=vmall -Dapm_tier=vmall-product-service

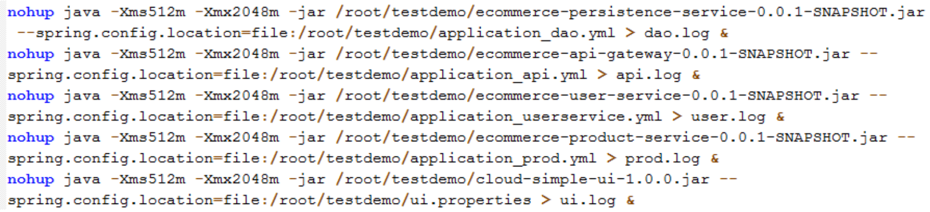
Table 6-1 Configuration description

Parameter	Description	Mandatory	Configuration
-javaagent	JAR package that collection probes depend on.	Yes	The fixed value is **/opt/oss/servicemgr/ICAgent/pinpoint/pinpoint-bootstrap.jar**. **NOTE** To start a Java application as a non-root user, ensure that the user has the read and write permissions on the **/opt/oss/servicemgr/ICAgent/pinpoint/** directory.

Parameter	Description	Mandatory	Configuration
-Dapm_application	Application name.	Yes	Set the name based on actual conditions. For example, if the VMall billing system can form an application, the application name is **vmall**. **NOTE** Naming rules – Only lowercase letters, digits, hyphens (-), and underscores (_) are allowed. The value must start with a lowercase letter or an underscore. – The value must be 1 to 64 characters long.
-Dapm_tier	Application microservice name.	Yes	Set the name based on actual conditions. For example, in the Vmall billing system, the application microservice names are **vmall-product-service** and **vmall-api-service**. **NOTE** Naming rules – Only lowercase letters, digits, hyphens (-), and underscores (_) are allowed. The value must start with a lowercase letter or an underscore. – The value must be 1 to 64 characters long.
-Xbootclasspath	JBoss installation package directory.	Yes	The format is **-Xbootclasspath/p:/JBoss installation package directory**, for example, **-Xbootclasspath/p:/opt/jboss/jboss-as-7.1.0.Final**.

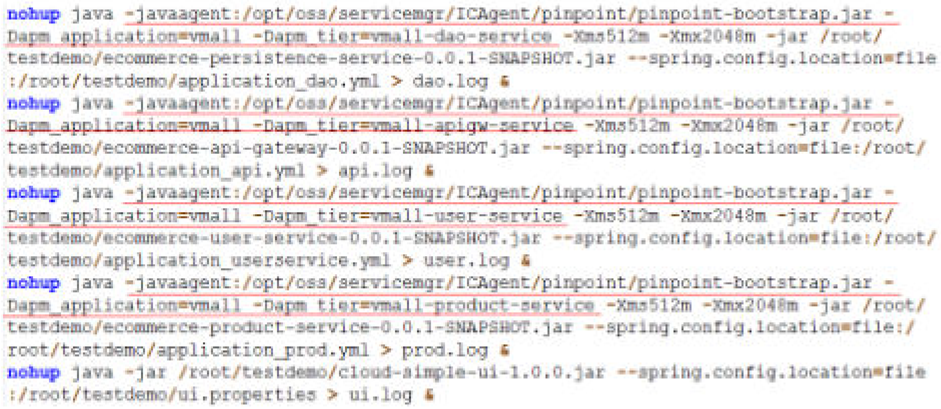
Example

Before configuration:



```
nohup java -Xms512m -Xmx2048m -jar /root/testdemo/ecommerce-persistence-service-0.0.1-SNAPSHOT.jar
 --spring.config.location=file:/root/testdemo/application_dao.yml > dao.log &
nohup java -Xms512m -Xmx2048m -jar /root/testdemo/ecommerce-api-gateway-0.0.1-SNAPSHOT.jar --
spring.config.location=file:/root/testdemo/application_api.yml > api.log &
nohup java -Xms512m -Xmx2048m -jar /root/testdemo/ecommerce-user-service-0.0.1-SNAPSHOT.jar --
spring.config.location=file:/root/testdemo/application_userservice.yml > user.log &
nohup java -Xms512m -Xmx2048m -jar /root/testdemo/ecommerce-product-service-0.0.1-SNAPSHOT.jar --
spring.config.location=file:/root/testdemo/application_prod.yml > prod.log &
nohup java -Xms512m -Xmx2048m -jar /root/testdemo/cloud-simple-ui-1.0.0.jar --
spring.config.location=file:/root/testdemo/ui.properties > ui.log &
```

After configuration:



```
nohup java -javaagent:/opt/oss/servicemgr/ICAgent/pinpoint/pinpoint-bootstrap.jar -
Dapm_application=vmall -Dapm_tier=vmall-dao-service -Xms512m -Xmx2048m -jar /root/
testdemo/ecommerce-persistence-service-0.0.1-SNAPSHOT.jar --spring.config.location=file
:/root/testdemo/application_dao.yml > dao.log &
nohup java -javaagent:/opt/oss/servicemgr/ICAgent/pinpoint/pinpoint-bootstrap.jar -
Dapm_application=vmall -Dapm_tier=vmall-apigw-service -Xms512m -Xmx2048m -jar /root/
testdemo/ecommerce-api-gateway-0.0.1-SNAPSHOT.jar --spring.config.location=file:/root/
testdemo/application_api.yml > api.log &
nohup java -javaagent:/opt/oss/servicemgr/ICAgent/pinpoint/pinpoint-bootstrap.jar -
Dapm_application=vmall -Dapm_tier=vmall-user-service -Xms512m -Xmx2048m -jar /root/
testdemo/ecommerce-user-service-0.0.1-SNAPSHOT.jar --spring.config.location=file:/root/
testdemo/application_userservice.yml > user.log &
nohup java -javaagent:/opt/oss/servicemgr/ICAgent/pinpoint/pinpoint-bootstrap.jar -
Dapm_application=vmall -Dapm_tier=vmall-product-service -Xms512m -Xmx2048m -jar /root/
testdemo/ecommerce-product-service-0.0.1-SNAPSHOT.jar --spring.config.location=file:/
root/testdemo/application_prod.yml > prod.log &
nohup java -jar /root/testdemo/cloud-simple-ui-1.0.0.jar --spring.config.location=file
:/root/testdemo/ui.properties > ui.log &
```

- If you start the program as a non-root user, run the following commands to modify the permissions on the probe file and output directory before enabling application monitoring:

  ```
  chmod -R 777 /opt/oss/servicemgr/ICAgent/pinpoint/
  mkdir -p /paas-apm/collectors/pinpoint
  chmod -R 777 /paas-apm
  ```

- On the ECS server, if you use Tomcat to start the service, you need to add a probe as follows:

  a. Go to the **bin** directory of Tomcat.

  b. Copy the following content to the **catalina.sh** file.

  ```
  JAVA_OPTS="$JAVA_OPTS -javaagent:/opt/oss/servicemgr/ICAgent/pinpoint/pinpoint-bootstrap.jar -Dapm_application=xxx -Dapm_tier=xxx"
  ```

  Set the value of **-Dapm_application** to the application name and that of **-Dapm_tier** to the name of the microservice to be started. Each name can contain up to 64 characters and must start with a lowercase letter or an underscore (_). Only lowercase letters, digits, hyphens (-), and underscores (_) are allowed.

- If JBoss is used to deploy a Java application, add the following information in bold to the corresponding code segment in the **standalone.conf** configuration file before starting JBoss in standalone mode. Note that APM supports the standalone mode only. In the added configuration content, variables are italicized. For details, see **Table 6-1**. After the configuration, start the applications for performance monitoring.

  Example:

  ```
  if [ "x$JBOSS_MODULES_SYSTEM_PKGS" = "x" ]; then
  JBOSS_MODULES_SYSTEM_PKGS="org.jboss.byteman,org.jboss.logmanager,com.navercorp.pinpoint.bootstrap,com.navercorp.pinpoint.common,com.navercorp.pinpoint.exception"
  fi
  # Uncomment the following line to prevent manipulation of JVM options
  # by shell scripts.
  #
  #PRESERVE_JAVA_OPTS=true
  ```

```
#
# Specify options to pass to the Java VM.
#
if [ "x$JAVA_OPTS" = "x" ]; then
JAVA_OPTS="-Xms64m -Xmx512m -XX:MaxPermSize=256m -
Djava.net.preferIPv4Stack=true -Dorg.jboss.resolver.warning=true -
Dsun.rmi.dgc.client.gcInterval=3600000 -Dsun.rmi.dgc.server.gcInterval=3600000"
JAVA_OPTS="$JAVA_OPTS -Djboss.modules.system.pkgs=
$JBOSS_MODULES_SYSTEM_PKGS -Djava.awt.headless=true"
JAVA_OPTS="$JAVA_OPTS -Djboss.server.default.config=standalone.xml"
JAVA_OPTS="$JAVA_OPTS -javaagent:/opt/oss/servicemgr/ICAgent/pinpoint/pinpoint-
bootstrap.jar -Dapm_application=vmall -Dapm_tier=vmall-product-service"
JAVA_OPTS="$JAVA_OPTS -Xbootclasspath/p:/opt/jboss/jboss-as-7.1.0.Final/
modules/org/jboss/logmanager/log4j/main/jboss-logmanager-log4j-1.0.0.GA.jar -
Xbootclasspath/p:/opt/jboss/jboss-as-7.1.0.Final/modules/org/jboss/logmanager/main/
jboss-logmanager-1.2.2.GA.jar -Xbootclasspath/p:/opt/jboss/jboss-as-7.1.0.Final/
modules/org/apache/log4j/main/log4j-1.2.16.jar"
JAVA_OPTS="$JAVA_OPTS -
Djava.util.logging.manager=org.jboss.logmanager.LogManager"
else
echo "JAVA_OPTS already set in environment; overriding default settings with values:
$JAVA_OPTS"
fi
```

Step 3 After the applications run for about three minutes, log in to the APM console to view the application information on the **Topology** and **Transactions** pages.

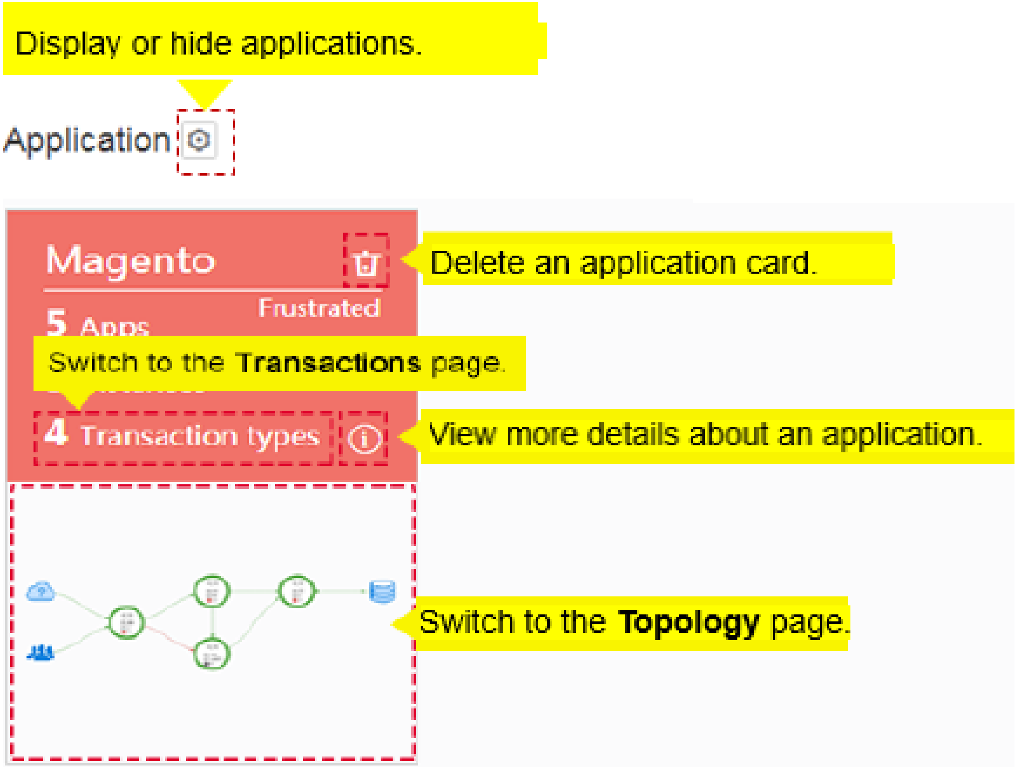
----End

7 Dashboard

You can quickly learn about the health status of applications through the dashboard.

You can perform multiple operations on the **Dashboard** page, as shown in the following figure.

Figure 7-1 Dashboard page


Display or hide applications.
Application
Magento
5 Apps
Frustrated
Delete an application card.
Switch to the **Transactions** page.
4 Transaction types
View more details about an application.
Switch to the **Topology** page.

You can delete a service card in the following scenarios:

- The service connected to Application Performance Management (APM) has been deleted.

- The ICAgent has been uninstalled and service data does not need to be collected.

If the service connected to APM is still running, the service card will be displayed again three minutes after it is deleted.

8 Alarm Center

8.1 Viewing Alarms

Alarms are reported when Application Performance Management (APM) is abnormal or may cause exceptions. Alarms need to be handled. Otherwise, service exceptions may occur.

Viewing Alarms

Step 1 Log in to the APM console.

Step 2 In the navigation pane, choose **Alarm Center** > **Alarm List**.

Step 3 View alarms on the **Alarm List** page.

1. Set a time range to view alarms. There are two methods to set a time range:

   Method 1: Use the predefined time label, for example, **Last 1 hour**, **Last 6 hours**, or **Last 1 day**. You can choose one based on service requirements.

   Method 2: Customize a time range. You can specify 30 days at most.

2. Set filter criteria and click **Search** to view the alarms in the time range.

   You can click **Reset** to reset filter criteria.

Step 4 Perform the operations listed in **Table 8-1** if needed.

Table 8-1 Operations

Operation	Method	Description
Viewing alarm statistics	View alarm statistics that meet filter criteria within a specific time range through a bar graph.	-

Operation	Method	Description
Clearing alarms	Click **Clear** in the **Operation** column to clear a target alarm.	● You can clear an alarm after the problem that causes this alarm is resolved. ● Alarms that are cleared cannot be queried.
Viewing alarm details	Click **View Details** in the **Operation** column to view alarm details.	-

----End

8.2 Viewing Events

Events carry important information, informing you of the changes of Application Performance Management (APM) itself. Such changes do not necessarily cause exceptions. Events do not need to be handled.

Viewing Events

Step 1 Log in to the APM console.

Step 2 In the navigation pane, choose **Alarm Center** > **Event List**.

Step 3 View events on the **Event List** page.

1. Set a time range to view events. There are two methods to set a time range:

   Method 1: Use the predefined time label, for example, **Last 1 hour**, **Last 6 hours**, or **Last 1 day**. You can choose one based on service requirements.

   Method 2: Customize a time range. You can specify 30 days at most.

2. Set filter criteria and click **Search** to view the events in the time range.

   You can click **Reset** to reset filter criteria.

Step 4 Perform the operations listed in **Table 8-2** if needed.

Table 8-2 Operations

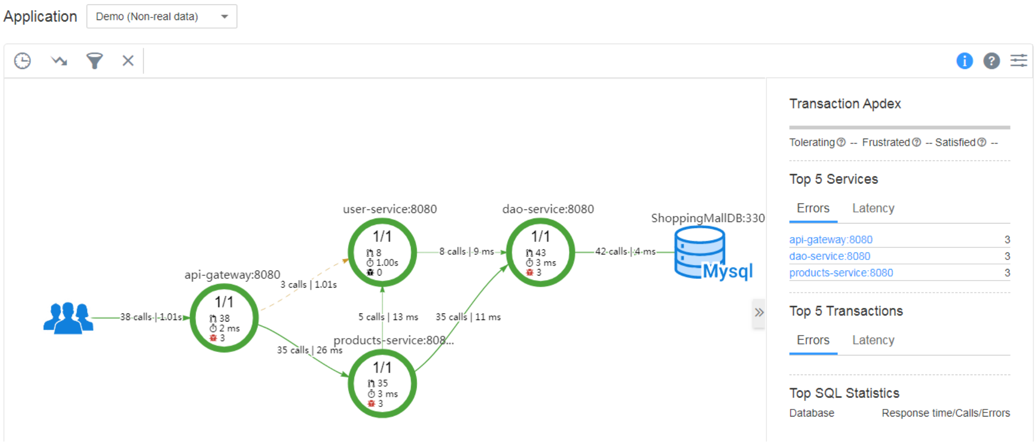
Operation	Method	Description
Viewing event statistics	View event statistics that meet filter criteria within a specific time range through a bar graph.	-

----End

9 Topology

In a topology, each circle represents a service, each section in the circle represents an instance, and each arrow represents a call relationship. In addition, Application Performance Management (APM) can display the call relationships between applications. Each circle can also represent an application. When a circle represents an application, right-click the circle and choose **View Application** to go to the topology page.

Different colors on the circle represent different health statuses. The color is determined by the **Application Performance Index (Apdex)** value. The closer the Apdex value is to **1**, the healthier the application is.



Application Demo (Non-real data)

Transaction Apdex

Tolerating -- Frustrated -- Satisfied --

Top 5 Services

Errors Latency

api-gateway:8080	3
dao-service:8080	3
products-service:8080	3

Top 5 Transactions

Errors Latency

Top SQL Statistics

Database Response time/Calls/Errors

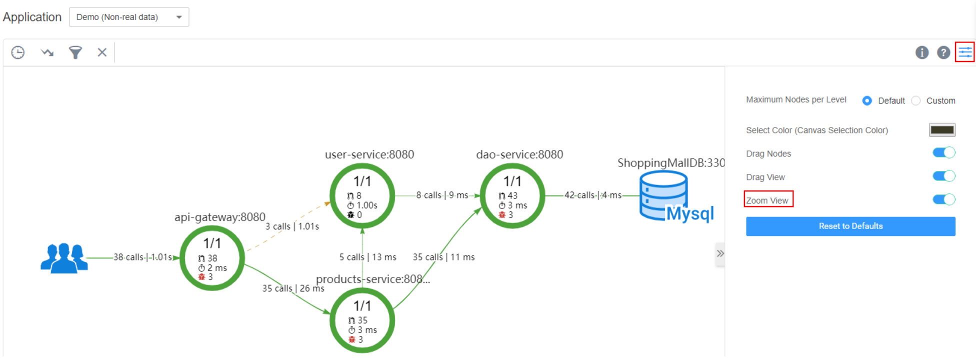
1. **Table 9-1** provides topology description.

Table 9-1 Topology description

Color	Instance	Call
Green	0.75 ≤ Apdex ≤ 1 The instance responds quickly when it is called.	0.75 ≤ Apdex ≤ 1 Quick response.

Color	Instance	Call
Yellow	0.3 ≤ Apdex < 0.75 The instance responds slowly when it is called.	0.3 ≤ Apdex < 0.75 Slow response.
Red	0 ≤ Apdex < 0.3 The instance responds very slowly when it is called.	0 ≤ Apdex < 0.3 Very slow response.
Gray	The instance is not called.	-
Black	The instance has been deleted.	-

2. On the **Topology** page, you can click ☰ to configure the topology. For example, if **Zoom View** is disabled, you cannot zoom in or out the topology.



Application: Demo (Non-real data)

Maximum Nodes per Level ● Default ○ Custom

Select Color (Canvas Selection Color)

Drag Nodes ●

Drag View ●

Zoom View ●

Reset to Defaults

3. On the right of the **Topology** page, set a time range to view the following topology details of an application:
   - Transaction Apdex
   - Top 5 services ranked by errors and latency
   - Top 5 transactions ranked by errors and latency
   - Top 5 SQL statements ranked by response time, calls, and errors
4. In the topology, click a circle (indicating a service) to view metrics, including Service Level Agreement (SLA), basic service metrics, and transaction details.
5. In the topology, click a segment (indicating an instance) in a circle to view metrics, including basic instance metrics, JVM metrics, node metrics, and transaction details.

   NOTE

   Currently, only JVM metrics of the last 15 minutes can be displayed.

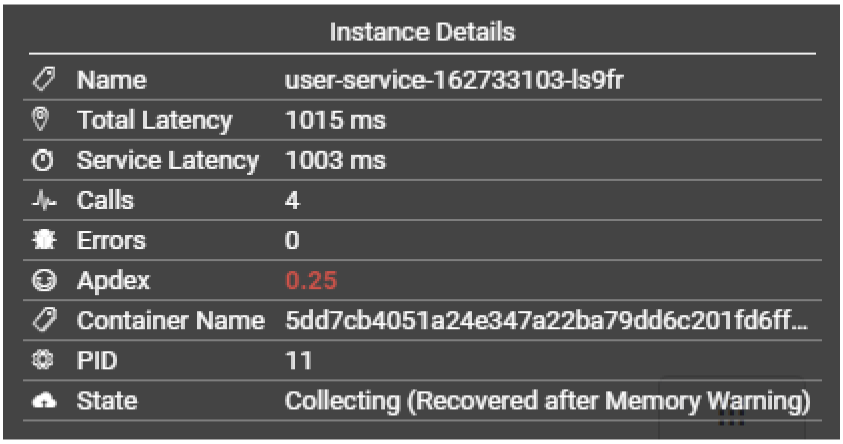
Locating Problems Based on the Topology

The following describes how to locate an instance with a slow response:

Step 1 On the **Topology** page, set the time range during which a problem occurred in the upper right corner.

Step 2 In the topology, view the instance (highlighted in red) with a slow response, as shown in **Figure 9-1**.

Figure 9-1 Abnormal instance



Instance Details

Name	user-service-162733103-ls9fr
Total Latency	1015 ms
Service Latency	1003 ms
Calls	4
Errors	0
Apdex	0.25
Container Name	5dd7cb4051a24e347a22ba79dd6c201fd6ff...
PID	11
State	Collecting (Recovered after Memory Warning)

Step 3 (Optional) For a service containing multiple instances, right-click each instance and choose **Expand** from the shortcut menu to view call relationships to preliminarily identify the abnormal instance.

Step 4 Choose **Find Call-Chain** from the shortcut menu. On the page that is displayed, further locate the problem based on call duration and other parameters.

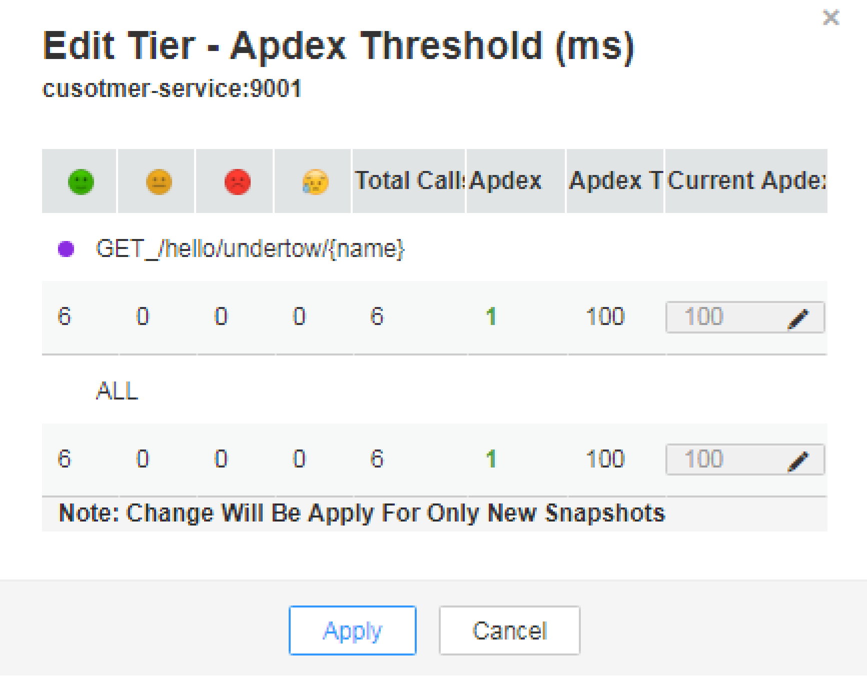
----End

Configuring an Apdex Threshold for a Transaction

The response time of different transactions is different. APM enables you to configure different Apdex thresholds for different transactions. For example, if a login takes more than 50 ms, the response is slow. If a transaction query takes more than 10 ms, the response is slow. In this case, you need to set different Apdex thresholds for the login and query transactions.

Step 1 On the **Topology** page, move the mouse cursor over the circle diagram, right-click it, and click **Edit Threshold**.

Step 2 Modify the Apdex threshold and click **Apply**.



Edit Tier - Apdex Threshold (ms)

cusotmer-service:9001

🙂	😐	☹	😥	Total Call:	Apdex	Apdex T	Current Apde:
● GET_/hello/undertow/{name}							
6	0	0	0	6	1	100	100
ALL							
6	0	0	0	6	1	100	100

Note: Change Will Be Apply For Only New Snapshots

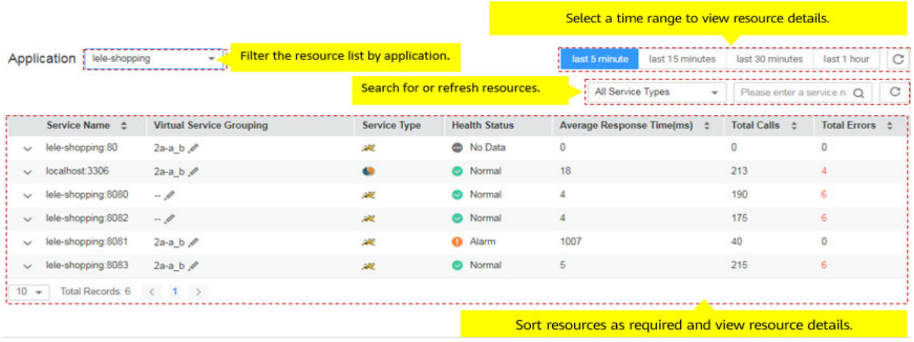
----End

10 Inventory

On the **Inventory** page, application details are displayed, helping you locate faults.

You can perform multiple operations on the **Inventory** page, as shown in the following figure.

Figure 10-1 Inventory page


Select a time range to view resource details.
Application
lele-shopping
Filter the resource list by application.
last 5 minute
last 15 minutes
last 30 minutes
last 1 hour
Search for or refresh resources.
All Service Types
Please enter a service n
Sort resources as required and view resource details.

Service Name	Virtual Service Grouping	Service Type	Health Status	Average Response Time(ms)	Total Calls	Total Errors
lele-shopping:80	2a-a_b		No Data	0	0	0
localhost:3306	2a-a_b		Normal	18	213	4
lele-shopping:8080	--		Normal	4	190	6
lele-shopping:8082	--		Normal	4	175	6
lele-shopping:8081	2a-a_b		Alarm	1007	40	0
lele-shopping:8083	2a-a_b		Normal	5	215	6

10 Total Records: 6 < 1 >

11 Transactions

To complete a transaction, you may call multiple services. Any slow or error call may lead to slow responses. During routine O&M, you can analyze the transactions with slow responses to locate and rectify application problems, thereby improving user experience.

Analyzing Problems Based on Transactions

The following describes how to locate the cause of a transaction with an extremely slow response:

Step 1 On the **Transactions** page, select a transaction with an extremely slow response from the transaction list.

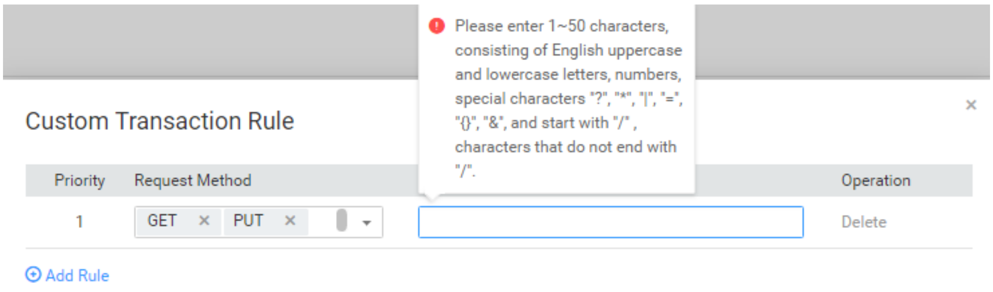
Step 2 Choose **More** > **View Call Relationship** in the **Operation** column. On the page that is displayed, further locate the problem based on call duration and other parameters. Alternatively, switch to the topology page to locate the problem. For details, see **Locating Problems Based on the Topology**.

----End

Customizing Transactions

To precisely define transactions and collect tracing data, use the URI template to customize transactions and classify requests into different transactions. When the collector receives requests, custom transactions will be calculated first.

Step 1 On the **Transactions** page, click **Custom Transaction Rule**. A transaction consists of the request method and regular expression. It is in the format of **{Request Method}_/{pattern}**. Example: When the request methods are **GET** and **POST** and the regular expression is **/{name}**, the transaction is **GET,POST_/{name}**.



Custom Transaction Rule

Please enter 1~50 characters, consisting of English uppercase and lowercase letters, numbers, special characters "?", "*", "|", "=", "{}", "&", and start with "/" , characters that do not end with "/".

Priority	Request Method		Operation
1	GET × PUT ×		Delete

Step 2 Select a request method. Request methods include **GET**, **PUT**, **DELETE**, **POST**, **HEAD**, **CONNECT**, **OPTIONS**, **PATCH**, **TRACE**, and **Select all**. **Select all** indicates all request methods.

Step 3 In the **Regular Expression** text box, enter a transaction rule and click **OK**. In this way, the custom transaction rule is added successfully.

The regular expression uses the **URI template** matching mode of the Spring MVC framework. Example: @RequestMapping(path="/owners/{ownerId}/pets/{petId}", method=RequestMethod.GET), where *ownerId* and *petId* are variables.

To add more custom transaction rules, click **Add Rule**.

NOTE

- A transaction rule must be 1 to 50 characters long. It must start with a slash (/) but cannot end with a slash. Only letters, digits, and special characters (?*|={}&) are allowed.
- Both the question mark (?) and asterisk (*) can be used for fuzzy search. One question mark represents one character, one asterisk represents 0 to N characters between two slashes in a URI, and double asterisks represent infinite characters. Example: When you enter **/first/***, **/first/test** can be returned but **/first/test/test** cannot. When you enter **/first/****, both **/first/test** and **/first/test/test** can be returned.

----End

12 Tracing

12.1 Call Chain

By tracing and recording service calls, Application Performance Management (APM) restores the execution traces and statuses of service requests in distributed systems, so that you can quickly locate performance bottlenecks and faults.

Locating Performance Bottlenecks

Step 1 Log in to the APM console. In the navigation pane, choose **Tracing** > **Call Chain**. Then, select the desired time range, application, and service from three drop-down lists, and click **Search**.

> 🕮 **NOTE**
>
> If you cannot select a service from an application, select another application from the **Application** drop-down list.

Step 2 (Optional) On the **Call Chain** page, click **Advanced** in the upper right corner, specify filter criteria, and click **Search** to search for the desired call chain.

Step 3 Click **View Call Relationship** in the **Operation** column.

Step 4 Identify the call that takes long time based on **Time Line (ms)** and then locate the performance bottleneck.

Step 5 (Optional) View additional information to further locate the fault cause.

Click **View Details** in the **Operation** column to view call details.

----End

Locating Faults

Step 1 On the **Call Chain** page, select the desired time range, application, and service from three drop-down lists, and click **Search**.

Step 2 (Optional) On the **Call Chain** page, click **Advanced** in the upper right corner, specify filter criteria, and click **Search** to search for the desired call chain.

Step 3 Check the application status in the **Status** column and find out the faulty service.

Step 4 Click **View Call Relationship** in the **Operation** column of the faulty service, check whether the return value is normal, and locate the fault.

Step 5 (Optional) View additional information to further locate the fault cause.

Click **View Details** in the **Operation** column to view call details.

----End

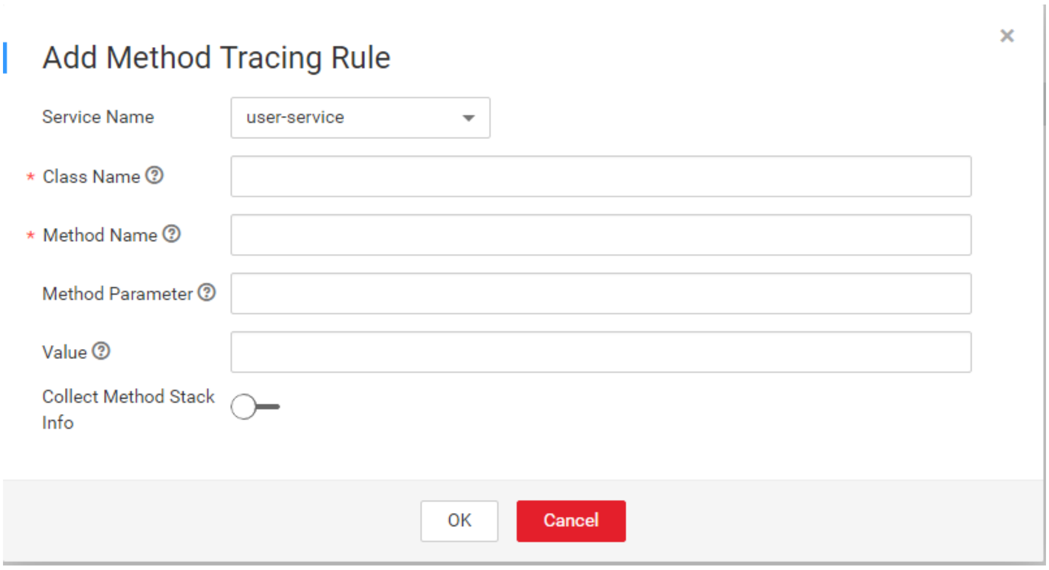
12.2 Method Tracing

Method tracing is used to dynamically trace a method of a class. When the method of this class is called, Application Performance Management (APM) collects the call data of the method based on configured method tracing rules using probes, and displays the call data on the **Call Chain** page. Method tracing is used to help application developers locate method-level performance problems online.

APM traces the APIs of most third-party open-source components, but does not trace specific methods in your applications. To monitor important methods in applications or methods of some third-party open-source components that are not supported by APM, you need to customize method tracing. After the configuration is complete, you can view the call data of the method on the **Call Chain** page.

Step 1 Customize a method tracing rule and start method tracing.

On the **Method Tracing** page, click **Add Method Tracing Rule**, set the parameters as shown in the following figure, and click **OK**.



Add Method Tracing Rule

Service Name: user-service

* Class Name

* Method Name

Method Parameter

Value

Collect Method Stack Info

NOTE

- If **Method Parameter** is not set, the methods of the same method name are used for collection by default.
- If **Value** is not set, the values of the method are not filtered during collection.
- If **Collect Method Stack Info** is enabled, the method stack information is collected.

Step 2 Preliminarily locate service performance problems based on the call duration and status.

Step 3 Click **View Call Relationship** in the **Operation** column to view the method-level call relationships.

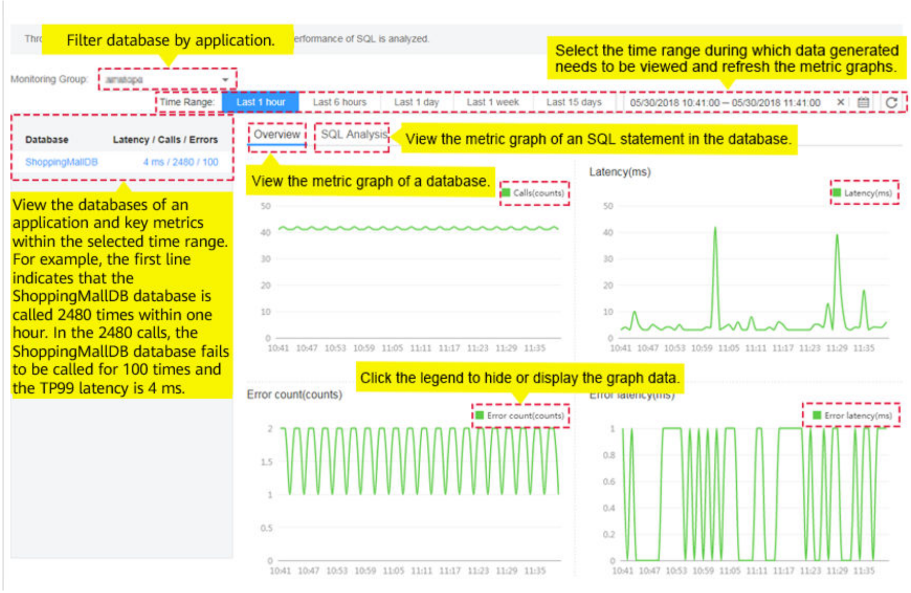
----End

13 SQL Analysis

Application Performance Management (APM) displays key metrics, such as SQL statement calls, response time, and errors for analyzing database performance problems caused by slow or error SQL statements. Currently, SQL analysis supports MySQL, Oracle, and PostgreSQL relational databases only.

SQL Page Description

Figure 13-1 SQL page



Filter database by application.

Select the time range during which data generated needs to be viewed and refresh the metric graphs.

Database	Latency / Calls / Errors
ShoppingMallDB	4 ms / 2480 / 100

View the databases of an application and key metrics within the selected time range. For example, the first line indicates that the ShoppingMallDB database is called 2480 times within one hour. In the 2480 calls, the ShoppingMallDB database fails to be called for 100 times and the TP99 latency is 4 ms.

View the metric graph of an SQL statement in the database.

View the metric graph of a database.

Click the legend to hide or display the graph data.

Analyzing Abnormal SQL Statements

When an SQL statement of a database is abnormal, performance problems such as service timeout may occur. During routine O&M, you can monitor key metrics (such as the error duration and latency) of databases, locate the SQL statements that take a long time to execute, operate at low efficiency, or fail to be called, and then make analysis and optimization accordingly.

The SQL analysis function determines whether to collect SQL data. Before performing the following steps, ensure that this function is enabled. Otherwise, no

SQL data can be queried. This function is enabled by default. If it is disabled, choose **Agent** > **Configuration** in the navigation pane and enable it.

Step 1 On the **SQL Analysis** page, set the time range during which a problem occurred in the upper right corner.

Step 2 On the **Overview** tab page, locate the faulty database in the application based on key metrics. If a database has long latency and many call errors, a performance problem may occur.

Step 3 Analyze the performance problem.

Click the **SQL Analysis** tab, and locate the abnormal SQL statement.

Step 4 Further analyze the cause.

1. Click the abnormal SQL statement to go to the **Call Chain** page and check the impact of this statement on the entire service.
2. Click **View Call Relationship** in the **Operation** column to find out the method of the abnormal SQL statement. Analyze the cause of the abnormal SQL statement in this method. For example, check whether the index is used, data volume is overlarge, syntax is correct, or deadlock occurs. Then, optimize the SQL statement accordingly.

----End

14 JVM Monitoring

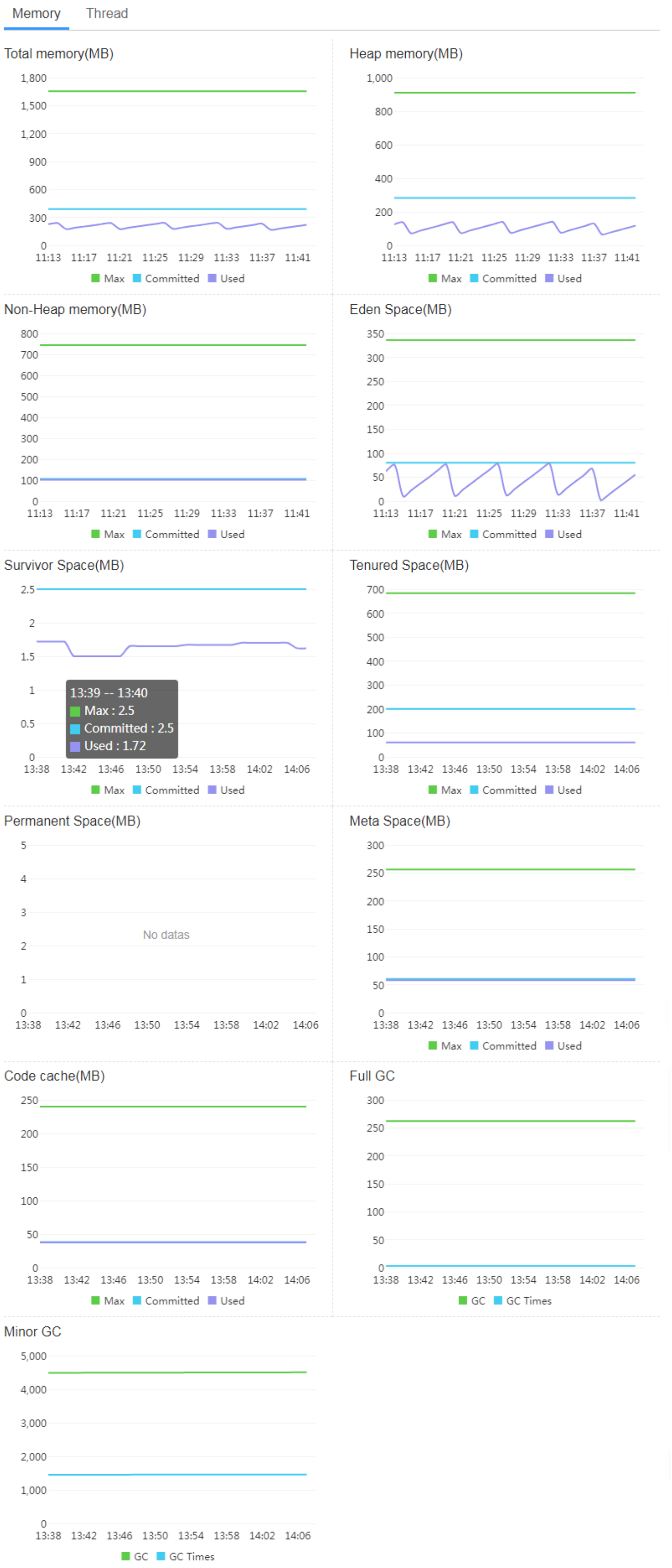
JVM monitoring displays the memory and thread metrics of the JVM operating environment for Java applications. You can monitor metric trends in real time to analyze performance.

On the **Memory** and **Thread** tab pages, you can view the memory and thread graphs to quickly locate problems such as memory leakage and thread exceptions.

Memory Graphs

As shown in **Figure 14-1**, in a selected time range, the trends of the maximum, committed, and used memory in different JVM memory spaces (such as the total memory, heap memory, and non-heap memory spaces) of an instance are displayed. In addition, the garbage collection (GC) duration and times are also displayed.

Figure 14-1 Memory graphs



JVM memory

JVM memory consists of heap and non-heap memory.

- Heap memory: A heap is the data area where the JVM is running. It allocates memory for all class instances and arrays. Heap memory of objects is reclaimed by an automatic memory management system called garbage collector. Heap space consists of eden space, survivor space, and tenured space.
- Non-heap memory: Memory (excluding heap memory) managed by JVM. Non-heap space consists of code cache and permanent space (or meta space).

Java heap is the main area managed by the garbage collector. It is also called garbage collection heap. GC mode includes full GC and minor GC.

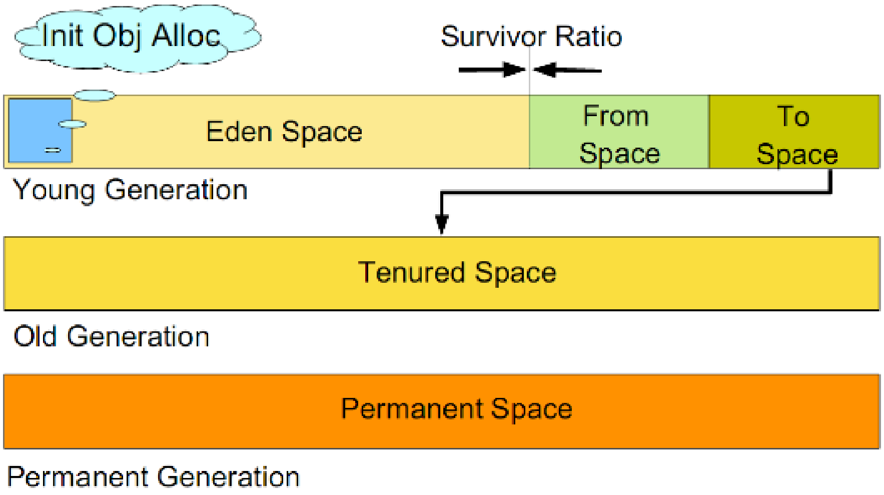
Table 14-1 Memory spaces

Space Name	Description
Eden space	Initially allocates memory from the thread pool to most objects.
Survivor space	Stores the eden space's objects that are not reclaimed during GC.
Tenured space	Maintains the objects which have been stored in the survivor space for a period of time.
Code cache	Compiles and stores local code.
Permanent space	Stores static data of VMs, for example, classes and method objects.
Meta space	Stores local class metadata. In versions later than Java 8, permanent space is replaced by meta space.
Full GC	Indicates the GC performed in the entire heap space (covering young-, old-, and permanent-generation spaces) when the memory space is still insufficient after memory reclamation.
Minor GC	Indicates the GC performed in the young-generation space (including eden and survivor spaces) when the allocated memory is insufficient.

JVM collects garbage based on generations. JVM heap space is divided into old- and young-generation spaces. More than 90% objects that exist only for a short period of time are stored in the young-generation space, while objects that have long life cycles are stored in the old-generation space. Young-generation space is further divided into eden space and two survivor spaces. New objects are initially allocated to the eden space. The survivor spaces are used as the buffer between eden space and tenured space. Objects that are survived after several rounds of

GC in the survivor spaces are then transferred to the old-generation space, as shown in **Figure 14-2**.

Figure 14-2 Memory spaces



NOTE

There are two survivor spaces, which are represented by **from** and **to** pointers. The **to** pointer points to the empty survivor space.

Thread Graphs

As shown in **Figure 14-3**, in a selected time range, the trends of total threads, sticky threads, dedicated threads, and other threads are displayed.

Figure 14-3 Thread graphs

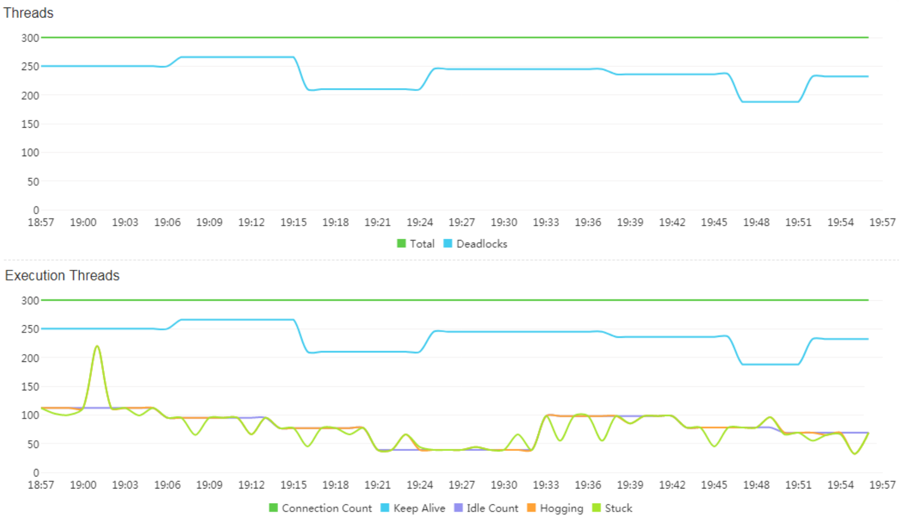


Table 14-2 Threads

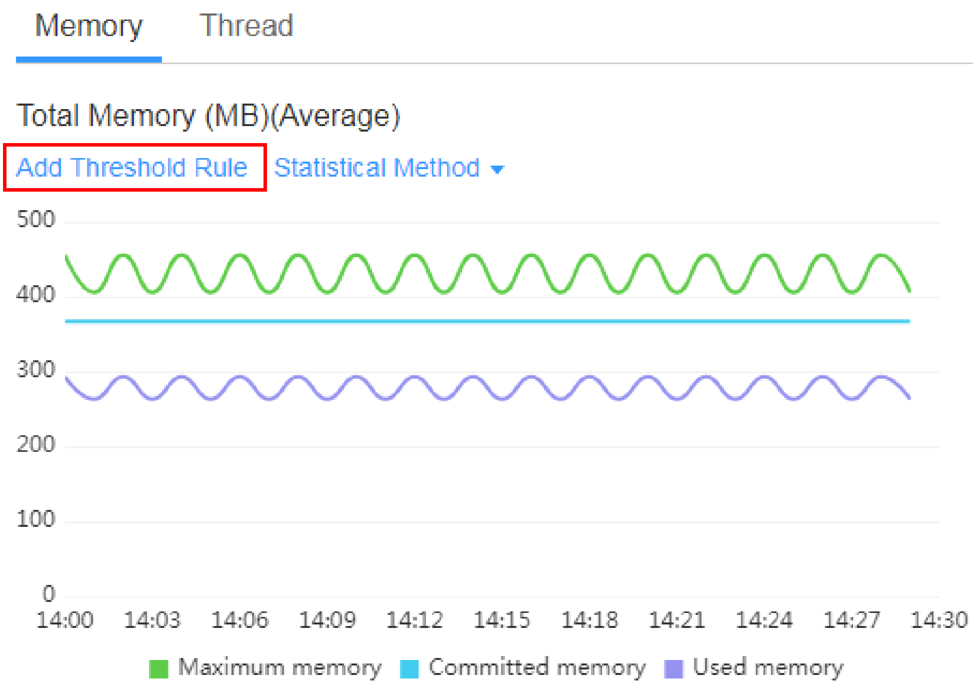
Thread Name	Description
Total threads	Both active and standby threads are included. Sticky threads and dedicated threads become standby threads after being executed.
Deadlock threads	When two or more processes encounter resource conflicts or the communication between them is abnormal, the system enters the deadlock state.
Sticky threads	If the time taken to process a request by a thread exceeds the preset maximum time, the thread is called a sticky thread.
Dedicated threads	If the time taken to process a request by a thread exceeds the normal execution time but does not exceed the maximum time of a sticky thread, the thread is called a dedicated thread.
Total executed threads	Both active and idle threads are included.
Active threads	Sticky threads, dedicated threads, and threads that are being executed are included.
Idle threads	Threads are in idle state. When there is no task, a thread is in the idle state. When receiving a request, the thread pool assigns an idle thread to the request. After the assigned task is completed, the idle thread returns to the thread pool and waits for another task.

Adding a Threshold Rule

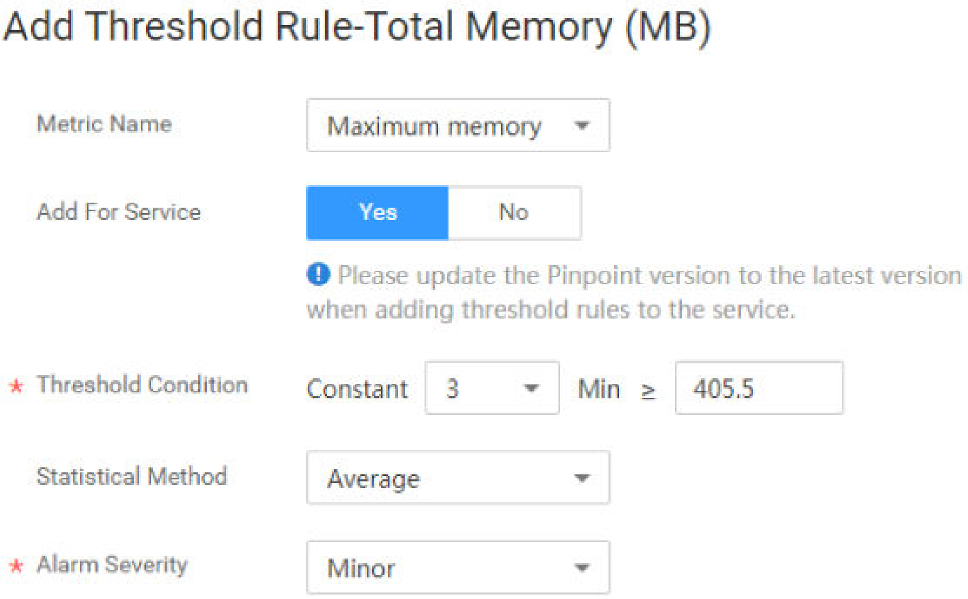
You can add threshold rules for all JVM memory and thread metrics. When the rules are met, alarms are reported, altering you to risks.

Step 1 On the **JVM Monitoring** page, select an application in the upper left corner, and then select an instance.

Step 2 In the trend graph of a memory or thread metric on the right, set a threshold rule. Specifically, click **Add Threshold Rule** on the top of the trend graph.



Step 3 Set rule parameters and click **Submit**, as shown in the following figure.



Add Threshold Rule-Total Memory (MB)

Metric Name: Maximum memory

Add For Service: Yes No

Please update the Pinpoint version to the latest version when adding threshold rules to the service.

* Threshold Condition: Constant 3 Min ≥ 405.5

Statistical Method: Average

* Alarm Severity: Minor

----End

15 Collection Management

15.1 Agent Management

15.1.1 Installing the ICAgent

The following table lists the ICAgent status.

Table 15-1 ICAgent status

Status	Description
Running	The ICAgent is running properly.
Uninstalled	The ICAgent is not installed. For details about how to install the ICAgent, see **Installing the ICAgent**.
Installing	The ICAgent is being installed. This operation takes about 1 minute to complete.
Installation failed	Failed to install the ICAgent on the host. Uninstall the ICAgent according to **Uninstalling the ICAgent Through Logging In to a Server** and then install it again.
Upgrading	The ICAgent is being upgraded. This operation takes about 1 minute to complete.
Upgrade failed	Failed to upgrade the ICAgent. Uninstall the ICAgent according to **Uninstalling the ICAgent Through Logging In to a Server** and then install it again.
Offline	The AK/SK are incorrect. Obtain the correct AK/SK and install the ICAgent again.
Abnormal	The ICAgent is abnormal. Contact technical support.

Prerequisites

Before installing the ICAgent, ensure that the time and time zone of the local browser are consistent with those of the server. If multiple servers are deployed, ensure that the local browser and multiple servers use the same time zone and time. Otherwise, the application topology and tracing data displayed on the UI may be incorrect.

Installation Methods

There are two methods to install the ICAgent. Note that the two methods are not applicable to the container nodes created using ServiceStage, Application Orchestration Service (AOS), or Cloud Container Engine (CCE). To monitor the container nodes using Application Performance Management (APM), perform operations according to **Deploying Applications**. For details, see **Table 15-2**.

Table 15-2 Installation methods

Method	Application Scenario
Initial installation	This method is used when the following conditions are met: 1. An Elastic IP Address (EIP) has been bound to the server. 2. The ICAgent has never been installed on the server.
Inherited installation	This method is used when the following conditions are met: You have multiple servers on which the ICAgent is to be installed. One server is bound to an EIP, but others are not bound to an EIP. You can use this method to install the ICAgent on the servers that are not bound to an EIP.

Initial Installation

After you apply for a server and install the ICAgent for the first time, perform the following operations:

Step 1 Obtain the Access Key ID/Secret Access Key (AK/SK).

- If you have obtained the AK/SK, skip this step.
- For details about how to obtain the AK/SK, see **Obtaining the AK/SK**.

Step 2 Log in to the APM console.

Step 3 In the navigation pane, choose **Agent** > **Management**.

Step 4 Click **Install ICAgent**.

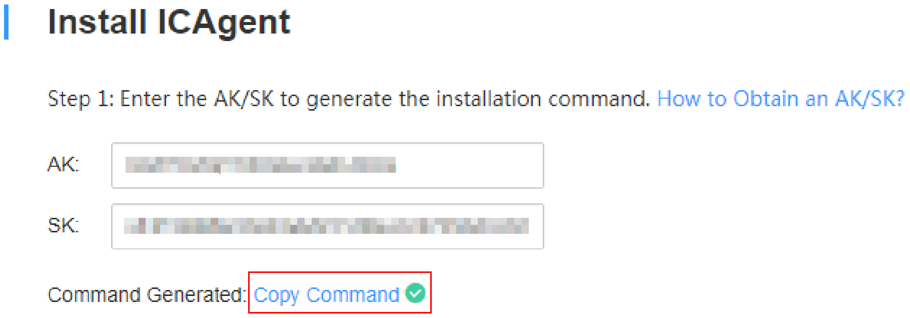
Step 5 Generate the ICAgent installation command and copy it.

1. Enter the obtained AK/SK in the text box to generate the ICAgent installation command.

   NOTE

   Ensure that the AK/SK are correct. Otherwise, the ICAgent cannot be installed.

2. Click **Copy Command**.



Install ICAgent

Step 1: Enter the AK/SK to generate the installation command. How to Obtain an AK/SK?

AK:

SK:

Command Generated: Copy Command ✓

Step 6 Use a remote login tool, such as PuTTY, to log in to the server where the ICAgent is to be installed as the **root** user and run the command copied in **Step 5.2** to install the ICAgent:

NOTE

- If the message "ICAgent install success" is displayed, the ICAgent is successfully installed in the **/opt/oss/servicemgr**/ directory. After the ICAgent is successfully installed, choose **Agent** > **Management** in the navigation pane to view the ICAgent status.
- If the installation fails, uninstall the ICAgent according to **Uninstalling the ICAgent** and then install it again. If the problem persists, contact technical support.

----End

Inherited Installation

If the ICAgent has been installed on a server and the installation package **ICProbeAgent.zip** exists in the **/opt/ICAgent/** directory of the server, use this method to install the ICAgent on a remote server with a few clicks.

Step 1 Run the following command (***x.x.x.x*** indicates the server IP address) on the server where the ICAgent has been installed:

```
bash /opt/oss/servicemgr/ICAgent/bin/remoteInstall/remote_install.sh -ip x.x.x.x
```

Step 2 Enter the password of the **root** user of the server where the ICAgent is to be installed as prompted.

NOTE

- If both the expect tool and the ICAgent have been installed on the server, the ICAgent will be installed on the remote server after the preceding command is executed. If the ICAgent has been installed on the server, but the expect tool has not, enter the information as prompted.
- Ensure that the **root** user can run the **SSH** and **SCP** commands on the server where the ICAgent has been installed to communicate with the remote server where the ICAgent is to be installed.
- If the message "ICAgent install success" is displayed, the ICAgent is successfully installed in the **/opt/oss/servicemgr**/ directory. After the ICAgent is successfully installed, choose **Agent** > **Management** in the navigation pane to view the ICAgent status.
- If the installation fails, uninstall the ICAgent according to **Uninstalling the ICAgent** and then install it again. If the problem persists, contact technical support.

----End

Inherited Batch Installation

If the ICAgent has been installed on a server and the installation package **ICProbeAgent.zip** exists in the **/opt/ICAgent/** directory of the server, use this method to install the ICAgent on multiple remote servers with a few clicks.

NOTICE

1. Ensure that you can run the **SSH** and **SCP** commands on the ECS server where the ICAgent has been installed to communicate with the remote ECS servers where the ICAgent is to be installed.
2. Batch installation scripts depend on Python versions. You are advised to implement batch installation on hosts running Python 2.x. Python 3.x does not support batch installation.

Prerequisites

The IP addresses and passwords of all servers where the ICAgent is to be installed have been collected, sorted in the **iplist.cfg** file, and uploaded to the **/opt/ICAgent/** directory on the server where the ICAgent has been installed. The following is an example of the **iplist.cfg** file, where IP addresses and passwords are separated by spaces.

192.168.0.109 password (Set the password as required.)

192.168.0.39 password (Set the password as required.)

NOTE

- Because the **iplist.cfg** file contains sensitive information, you are advised to clear the information after using it.
- If the passwords of all servers are the same, you only need to list IP addresses in the **iplist.cfg** file and enter the password once during execution. If the password of an IP address is different from those of other IP addresses, you need to list both passwords and IP addresses in the **iplist.cfg** file.
- The batch installation function depends on Python 2.7.*. If the system displays a message indicating that Python cannot be found during the installation, install Python 2.7.* and try again.

Procedure

Step 1 Run the following command on the server where the ICAgent has been installed:

bash /opt/oss/servicemgr/ICAgent/bin/remoteInstall/remote_install.sh -batchModeConfig /opt/ICAgent/iplist.cfg

Enter the default password of the **root** user of the server where the ICAgent is to be installed as prompted. If the passwords of all IP addresses have been configured in the **iplist.cfg** file, press **Enter** to skip this step. Otherwise, enter the default password.

```
batch install begin
Please input default passwd:
send cmd to 192.168.0.109
send cmd to 192.168.0.39
2 tasks running, please wait...
2 tasks running, please wait...
```

```
2 tasks running, please wait...
End of install agent: 192.168.0.39
End of install agent: 192.168.0.109
All hosts install icagent finish.
```

Wait until the message "All hosts install icagent finish." is displayed, which indicates that the ICAgent has been successfully installed on all the hosts listed in the configuration file.

Step 2 After the ICAgent is successfully installed, choose **Agent** > **Management** in the navigation pane to view the ICAgent status.

----End

15.1.2 Upgrading the ICAgent

To ensure better collection experience, Application Performance Management (APM) will continuously upgrade ICAgent versions. When the system displays a message indicating that a new ICAgent version is available, perform the following operations:

NOTE

If the ICAgent has a critical bug, the system will upgrade the ICAgent version.

Step 1 Log in to the APM console.

Step 2 In the navigation pane, choose **Agent** > **Management**.

Step 3 Select **Cluster: XXX** or **Other: user-defined nodes** from the drop-down list on the right of the page.

Step 4 Upgrade the ICAgent.

- If you select **Cluster: xxx** in **Step 3**, directly click **Upgrade ICAgent**. In this way, the ICAgent on all hosts in the cluster can be upgraded at a time.
- If you select **Other: user-defined nodes** in **Step 3**, select a desired host and then click **Upgrade ICAgent**.

Step 5 In the displayed **Upgrade ICAgent** dialog box, click **Yes**. Wait for about 1 minute to complete the ICAgent upgrade. When the ICAgent status changes from **Updating** to **Running**, the ICAgent is successfully upgraded.

----End

15.1.3 Uninstalling the ICAgent

If the ICAgent on a server is uninstalled, server O&M will be affected, making topology and tracing functions unavailable. Exercise caution when performing this operation.

You can uninstall the ICAgent using the following methods:

- **Uninstalling the ICAgent Through the Console**: Applies to the scenario where the ICAgent has been successfully installed and needs to be uninstalled.
- **Uninstalling the ICAgent Through Logging In to a Server**: Applies to the scenario where the ICAgent fails to be installed and needs to be uninstalled for reinstallation.

- **Remotely Uninstalling the ICAgent**: Applies to the scenario where the ICAgent has been successfully installed and needs to be remotely uninstalled.
- **Uninstalling the ICAgent in Batches**: Applies to the scenario where the ICAgent has been successfully installed, and needs to be uninstalled in batches.

Uninstalling the ICAgent Through the Console

Step 1 Log in to the Application Performance Management (APM) console.

Step 2 In the navigation pane, choose **Agent** > **Management**.

Step 3 Select **Other: user-defined nodes** from the drop-down list on the right of the page.

Step 4 Select one or more servers where the ICAgent is to be uninstalled, and click **Uninstall ICAgent**. In the **Uninstall ICAgent** dialog box, click **Yes**.

The ICAgent begins to be uninstalled. This operation takes about 1 minute to complete. When the ICAgent status changes from **Uninstalling** to **Uninstall**, the ICAgent is successfully uninstalled.

----End

Uninstalling the ICAgent Through Logging In to a Server

Step 1 Log in to the server from which the ICAgent is to be uninstalled as the **root** user.

Step 2 Run the following command to uninstall the ICAgent:

```
bash /opt/oss/servicemgr/ICAgent/bin/manual/uninstall.sh;
```

Step 3 Wait until the message "ICAgent uninstall success" is displayed.

----End

Remotely Uninstalling the ICAgent

In addition to the preceding method, you can use a method similar to **Inherited Installation** to remotely uninstall the ICAgent.

Step 1 Run the following command (***x.x.x.x*** indicates the server IP address) on the server where the ICAgent has been installed:

```
bash /opt/oss/servicemgr/ICAgent/bin/remoteUninstall/remote_uninstall.sh -ip x.x.x.x
```

Step 2 Enter the password of the **root** user of the server where the ICAgent is to be uninstalled as prompted.

NOTE

- If both the expect tool and the ICAgent have been installed on the server, the ICAgent will be uninstalled from the remote server after the preceding command is executed. If the ICAgent has been installed on the server, but the expect tool has not, enter the information as prompted.
- Ensure that the **root** user can run the **SSH** and **SCP** commands on the server where the ICAgent has been installed to communicate with the remote server where the ICAgent is to be uninstalled.
- If the message "ICAgent uninstall success" is displayed, the ICAgent is successfully uninstalled. After the ICAgent is successfully uninstalled, choose **Agent** > **Management** in the navigation pane to view the ICAgent status.

----End

Uninstalling the ICAgent in Batches

If the ICAgent has been installed on a server and the installation package **ICProbeAgent.zip** exists in the **/opt/ICAgent/** directory of the server, use this method to uninstall the ICAgent from multiple remote servers in batches with a few clicks.

NOTICE

The servers must belong to the same Virtual Private Cloud (VPC) and network segment.

Prerequisites

The IP addresses and passwords of all servers where the ICAgent is to be uninstalled have been collected, sorted in the **iplist.cfg** file, and uploaded to the **/opt/ICAgent/** directory on the server where the ICAgent has been installed. The following is an example of the **iplist.cfg** file, where IP addresses and passwords are separated by spaces.

192.168.0.109 password (Set the password as required.)

192.168.0.39 password (Set the password as required.)

NOTE

- Because the **iplist.cfg** file contains sensitive information, you are advised to clear the information after using it.
- If the passwords of all servers are the same, you only need to list IP addresses in the **iplist.cfg** file and enter the password once during execution. If the password of an IP address is different from those of other IP addresses, you need to list both passwords and IP addresses in the **iplist.cfg** file.

Procedure

Step 1 Run the following command on the server where the ICAgent has been installed:

bash /opt/oss/servicemgr/ICAgent/bin/remoteUninstall/remote_uninstall.sh -batchModeConfig /opt/ICAgent/iplist.cfg

Enter the default password of the **root** user of the server where the ICAgent is to be uninstalled as prompted. If the passwords of all IP addresses have been configured in the **iplist.cfg** file, press **Enter** to skip this step. Otherwise, enter the default password.

```
batch uninstall begin
Please input default passwd:
send cmd to 192.168.0.109
send cmd to 192.168.0.39
2 tasks running, please wait...
End of uninstall agent: 192.168.0.109
End of uninstall agent: 192.168.0.39
All hosts uninstall icagent finish.
```

Wait until the message "All hosts uninstall icagent finish." is displayed, which indicates that the ICAgent has been successfully uninstalled from all the hosts listed in the configuration file.

Step 2 After the ICAgent is successfully uninstalled, choose **Agent** > **Management** in the navigation pane to view the ICAgent status.

----End

15.2 Collection Configuration

To reduce memory, database, and disk space usage, you can implement collection configurations as required. Collection configurations take effect for the selected application.

Procedure

Step 1 Log in to the Application Performance Management (APM) console.

Step 2 In the navigation pane, choose **Agent** > **Configuration**.

Step 3 Select an application from the **Application** drop-down list.

> **NOTE**
>
> If different applications have different collection configurations, the collection configuration applied to all applications will overwrite that of a specific application.

Step 4 Click ○━ to enable data collection.

> **NOTE**
>
> This function is enabled by default. When you do not need to collect tracing or topology data of a specific application, disable this function to reduce resource usage.

Step 5 Click ○━ to enable the function of collecting normal call chain data.

To reduce the resources consumed by probes, APM collects one more data record every minute when a transaction is abnormal or the latency is greater than **Application Performance Index (Apdex) Threshold**. If this function is enabled, normal call chain data is sampled and collected. If this function is disabled, normal call chain data is not collected.

Step 6 Click ○━ to enable memory monitoring.

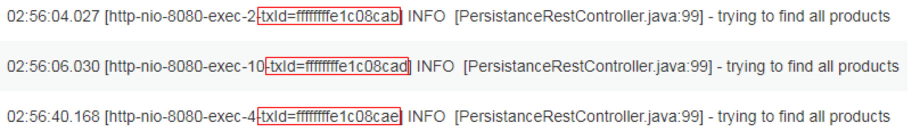
To prevent probes from affecting service performance in peak hours, enable memory monitoring. When the instance memory usage is excessively high, probes enter the hibernation state. You can also click ✎ to set the duration and memory usage.

NOTE

- Memory usage = Memory used by the Java process/Maximum available memory
- Maximum available memory: Use the smaller value between the available memory quota of the container and the maximum heap memory of the JVM. The maximum heap memory of the JVM is the value of **-Xmx**. The default value is 25% of the maximum available memory of the JVM.
- The memory usage during collection suspension must be greater than or equal to that during collection restoration.

Step 7 Click ○ to enable the function of adding trace IDs to logs.

A trace ID uniquely identifies a tracing. When this function is enabled, the system adds trace IDs to logs. You can accurately search for logs based on trace IDs, such as **ffffffffe1c08cab**, **ffffffffe1c08cad**, and **ffffffffe1c08cae**.



```
02:56:04.027 [http-nio-8080-exec-2-txId=ffffffffe1c08cab] INFO  [PersistanceRestController.java:99] - trying to find all products
02:56:06.030 [http-nio-8080-exec-10-txId=ffffffffe1c08cad] INFO  [PersistanceRestController.java:99] - trying to find all products
02:56:40.168 [http-nio-8080-exec-4-txId=ffffffffe1c08cae] INFO  [PersistanceRestController.java:99] - trying to find all products
```

Step 8 Click ○ to enable SQL analysis.

When this function is disabled, no SQL data is affected, but you cannot implement SQL analysis.

Step 9 Set the HTTP response codes to be ignored.

To quickly and accurately locate abnormal tracing, and prevent probes from misreporting normal tracing data, such as custom response codes, set the HTTP response codes to be ignored. Such codes will not be recorded in the error record table. Click ✎, enter the HTTP response codes to be ignored, and click ✓. If multiple HTTP response codes exit, separate them by commas (,).

Step 10 Set the errors and exceptions to be ignored.

To quickly and accurately locate abnormal tracing, and prevent probes from misreporting normal tracing data, set the errors and exceptions to be ignored. Such errors and exceptions will not be recorded in the error record table. Click ✎, enter the errors and exceptions to be ignored, and click ✓. If multiple Java exception classes exist, separate them by commas (,). The default value is null.

----End

16 Configuration Center

Configuring Apdex Thresholds

Step 1 Log in to the Application Performance Management (APM) console.

Step 2 In the navigation pane, choose **Configuration Center**.

Step 3 Select an application from the drop-down list.

Step 4 Set Application Performance Index (Apdex) thresholds. For details about Apdex and Apdex threshold, see **Apdex**.

- Click ✎ next to **Topology Apdex Threshold (ms)**, enter the topology Apdex threshold, and click ✓ to save the threshold.

  NOTE

  The default topology Apdex threshold is 500 ms.

- Click ✎ next to **Transaction Apdex Threshold (ms)**, enter the transaction Apdex threshold, and click ✓ to save the threshold.

  NOTE

  – The default transaction Apdex threshold is 2000 ms.

  – This setting takes effect for all transactions of the application. If an Apdex threshold has been separately set for a transaction, the currently set Apdex threshold takes effect for all transactions except this transaction. To separately set an Apdex threshold for a transaction, do as follows:

    1. In the navigation pane, choose **Transactions**.
    2. In the drop-down list in the upper left corner, select the application to which the transaction belongs.
    3. In the transaction list, click ✎ under the **Apdex Threshold (ms)** column of the transaction, enter an Apdex threshold, and click ✓ to save the threshold.

----End

17 Data Subscription

Application Performance Management (APM) allows you to subscribe to metrics or alarms. After the subscription, data can be forwarded to configured Kafka topics for you to retrieve.

Procedure

Step 1 In the navigation pane, choose **Configuration Center** > **Data Subscription**.

Step 2 Click **Create Subscription Rule** in the upper right corner of the page. Then, set parameters according to **Table 17-1** and click **OK**.

Table 17-1 Subscription rule parameters

Parameter	Description	Example
Rule name	Subscription rule name.	Enter **apm-kafka-test**.
Subscription content	Tracing.	Select **Apm Tracing**.
Subscription Target Type	Custom Kafka, which cannot be changed.	-
Subscription target connection address	Kafka address, which needs to be connected to Internet. Each address must be in the format of "IPv4 address:port". If there are multiple addresses, separate them by commas (,). Example: **192.168.0.1:9092,192.168.0.2:9092**	Set the parameters based on actual requirements.

Step 3 (Optional) Click ○ to enable Kafka SASL_SSL and set the parameters according to **Table 17-2**.

NOTE

Currently, APM supports only Kafka SASL_SSL security authentication. If Kafka SASL_SSL has been enabled for instances, enable this option when configuring data subscription.

Table 17-2 Kafka SASL_SSL parameters

Parameter	Description	Example
User name	SASL username for instance access authentication.	demo
Password	SASL password for instance access authentication. Keep your password secure. The system cannot detect your password.	-
Client certificate	Client certificate in the **.pem** format.	-

Step 4 On the **Rule Details** page, click **Verify and Save Custom Kafka Configuration** to verify the connectivity of the custom Kafka instance.

Step 5 Select the Kafka topic to which the data is to be sent.

Step 6 Click **OK**.

----End

Data Subscription Format

- The tracing data of APM is in the standard Zipkin format. Keywords **appId** and **projectId** need to be parsed from **binaryAnnotations**. The following shows an example.

```
[{
    "traceId": "adb64773d88dfac2",
    "id": "91324c265f7415a3",
    "name": "usg-stun:usg-cce-demo-99592:redis.clients.jedis.binaryjediscluster.subscribe",
    "timestamp": 1599187789769000,
    "duration": 6000,
    "binaryAnnotations": [{
        "key": "SRC-RESOURCE-ID",
        "value": "usg-stun:8080|d74a54d7a25be6552e640b3f658c5ad7"
    },
    {
        "key": "TX-TYPE",
        "value": "subscribe"
    },
    {
        "key": "appId",
        "value": "1011c321b34ff7bf7f2d02ab8a95750b"
    },
    {
        "key": "clusterId",
        "value": "unknown"
    },
    {
        "key": "destinationId",
        "value": "REDIS"
    },
    {
        "key": "monitorGroup",
        "value": "meeting-ulanqab3-mgdc1"
    },
    {
        "key": "namespace",
        "value": "usg"
    },
    {
        "key": "projectId",
        "value": "fd5c4fcd87874b5f85240cd9d93b34e0"
    },
    {
        "key": "result",
        "value": "1"
    },
    {
        "key": "root",
        "value": "true"
    },
    {
        "key": "serviceType",
        "value": "REDIS"
    },
    {
        "key": "transaction.info",
        "value": "false"
    }]
}]
```

- Kafka message example:

```
key:,
value:[{"traceId":"adb64773d88dfac2","id":"91324c265f7415a3","name":"usg-stun:usg-cce-
demo-99592:redis.clients.jedis.binaryjediscluster.subscribe","timestamp":1599187789769000,"duration":
6000,"binaryAnnotations":[{"key":"SRC-RESOURCE-ID","value":"usg-stun:8080|
d74a54d7a25be6552e640b3f658c5ad7"},{"key":"TX-TYPE","value":"subscribe"},
{"key":"appId","value":"1011c321b34ff7bf7f2d02ab8a95750b"},{"key":"clusterId","value":"unknown"},
{"key":"destinationId","value":"REDIS"},{"key":"monitorGroup","value":"meeting-ulanqab3-mgdc1"},
{"key":"namespace","value":"usg"},{"key":"projectId","value":"fd5c4fcd87874b5f85240cd9d93b34e0"},
{"key":"result","value":"1"},{"key":"root","value":"true"},{"key":"serviceType","value":"REDIS"},
{"key":"transaction.info","value":"false"}]}]
```

Follow-up Operations

After the data subscription rule is created, APM will send data to your configured Kafka topic so that you can retrieve the subscribed tracing data.

18 FAQs

18.1 Data Collection

Scope and Usage

When you enable data collection on Application Performance Management (APM), APM only collects service tracing data, resource information, resource attributes, memory detection information, and call request KPIs, but does not collect any privacy data. The collected data is used only for APM performance analysis and fault diagnosis, and is not used for any commercial purposes.

Data Type	Collected Data	Transmission Mode	Storage Mode	Data Purpose
Tracing data	Tracing span data	HTTPS encryption and Access Key ID/ Secret Access Key (AK/SK) authentication for transmission	Project-based isolated storage	Query and display at the tracing frontend

Data Type	Collected Data	Transmission Mode	Storage Mode	Data Purpose
Call request KPIs	Call initiator address, receiver address, API, duration, and status	HTTPS encryption and AK/SK authentication for transmission	Project-based isolated storage	Calculation of transaction call KPI metrics (such as throughput, TP99 latency, average latency, and number of call errors), drawing of application topologies, and display of call metrics and topologies at the frontend.
Resource information	Service type, service name, creation time, deletion time, node address, and service release API	HTTPS encryption and AK/SK authentication for transmission	Project-based isolated storage	Query and display at the resource library frontend
Resource attributes	System type, system startup event, number of CPUs, service executor, service process ID, service pod ID, CPU label, system version, web framework, JVM version, time zone, system name, collector version, and LastMail URL	HTTPS encryption and AK/SK authentication for transmission	Project-based isolated storage	Query and display at the resource library frontend
Memory monitoring information	Memory usage, used memory, maximum memory, remaining memory, memory threshold-crossing time, and memory monitoring configurations	HTTPS encryption and AK/SK authentication for transmission	Project-based isolated storage	Query and display at the resource library frontend

18.2 Obtaining the AK/SK

NOTE

Each user can create a maximum of two Access Key ID/Secret Access Key (AK/SK) pairs. Once they are generated, they are permanently valid.

- AK: unique ID associated with the SK. It is used together with the SK to sign requests.
- SK: secret access key used in conjunction with an AK to sign requests cryptographically. It identifies a request sender and prevents the request from being modified.

Procedure

Step 1 Log in to the management console.

Step 2 Click the username in the upper right corner to access the account center.

Step 3 On the **Basic Information** page, click **Manage**.

Step 4 Obtain the project ID and AK/SK.

1. Obtain the project ID.

   View the project ID in the project list.

2. Obtain the AK/SK.

   a. Choose **Access Keys** in the navigation pane, click **Create Access Key** to create an access key.

   b. Enter the login password and verification code sent to your mailbox or mobile phone.

   c. Click **OK** to download an access key.

   NOTE

   Keep the key secure.

----End