

Application Performance Management

Service Overview

Issue 01
Date 2024-04-22



Copyright © Huawei Technologies Co., Ltd. 2024. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Technologies Co., Ltd.

Trademarks and Permissions



HUAWEI and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

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

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

Security Declaration

Vulnerability

Huawei's regulations on product vulnerability management are subject to the *Vul. Response Process*. For details about this process, visit the following web page:

<https://www.huawei.com/en/psirt/vul-response-process>

For vulnerability information, enterprise customers can visit the following web page:

<https://securitybulletin.huawei.com/enterprise/en/security-advisory>

Contents

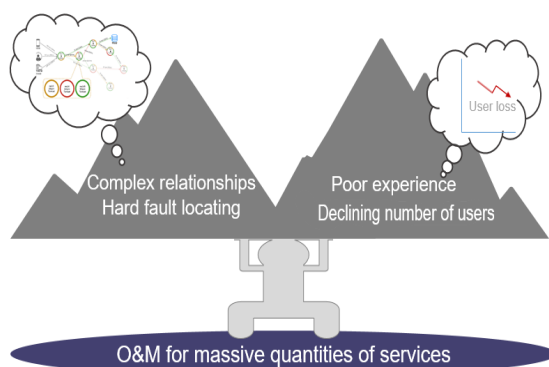
1 Overview.....	1
2 Functions.....	4
3 Application Scenarios.....	7
4 Basic Concepts.....	12
5 Edition Differences.....	16
6 Permissions Management.....	19
7 Privacy and Sensitive Information Protection Statement.....	22
8 Data Collection.....	23
9 Usage Restrictions.....	26
10 Billing.....	29
10.1 Pricing Details.....	30
11 Change History.....	35

1 Overview

O&M Challenges

In the cloud era, there are more and more applications under the distributed microservice architecture. The number of users also increases explosively, facing various application exceptions. In traditional O&M mode, metrics of multiple O&M systems cannot be associated for analysis. O&M personnel need to check application exceptions one by one based on experience, resulting in low efficiency, costly maintenance, and poor stability.

When there are massive quantities of services, O&M personnel face two major challenges.



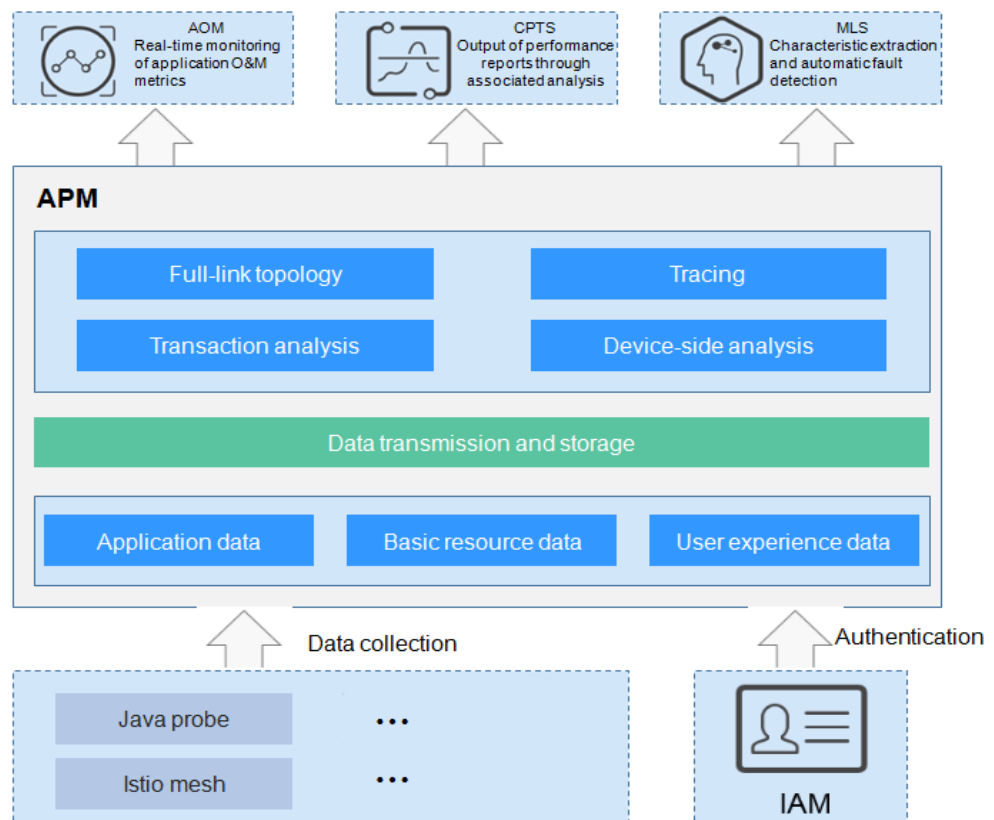
- Large distributed applications have complex relationships, making it difficult to analyze and locate problems. Specifically, O&M personnel face problems such as how to ensure normal application running, and quickly locate faults and performance bottlenecks.
- Users choose to leave due to poor experience. O&M personnel fail to detect and track services with poor experience in real time, and cannot diagnose application exceptions in a timely manner, severely affecting user experience.

Introduction to APM

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.

As a cloud application diagnosis service, APM has powerful analysis tools. It displays the application status, call processes, and user operations based on **topologies**, **tracing**, and **transaction analysis**, so that you can quickly locate and resolve faults and performance bottlenecks.

Figure 1-1 APM architecture



1. Access APM: You can access APM by creating an Identity and Access Management (IAM) agency and implementing Access Key ID/Secret Access Key (AK/SK) authentication.
2. Data collection: APM can collect data about applications, basic resources, and user experience from **Java probes**, and Istio mesh in non-intrusive mode.
3. Service implementation: APM supports **topologies**, **tracing**, and **transaction analysis**.
4. Service expansion:
 - Application Operations Management (AOM) monitors application O&M metrics in real time while APM quickly diagnoses application performance bottlenecks through topologies and tracing.

- Cloud Performance Test Service (CPTS) implements association analysis and generates performance reports after APM identifies performance bottlenecks.
- Based on the historical metric data learned using intelligent algorithms, APM associates metrics for analysis from multiple dimensions, extracts the context data of both normal and abnormal services for comparison, and locates root causes through cluster analysis.

Advantages



Ease of Use

Connects to applications without having to modify code, and collects data in a non-intrusive way. Data comes from:

- Java probe: Collects service call data, service inventory data, and call KPI data in non-intrusive mode based on the pinpoint open-source project.
- Istio mesh: Collects service call data, resource information, and call KPI data in non-intrusive mode through the Kubernetes platform.



High Performance

Delivers high throughput (hundreds of millions of API calls), ensuring premium experience.



Intelligent Analysis

Analyzes root causes using AI-powered threshold detection and machine learning based on historical baseline data.



Open Ecosystem

Opens O&M data query APIs and collection standards, and supports independent development.

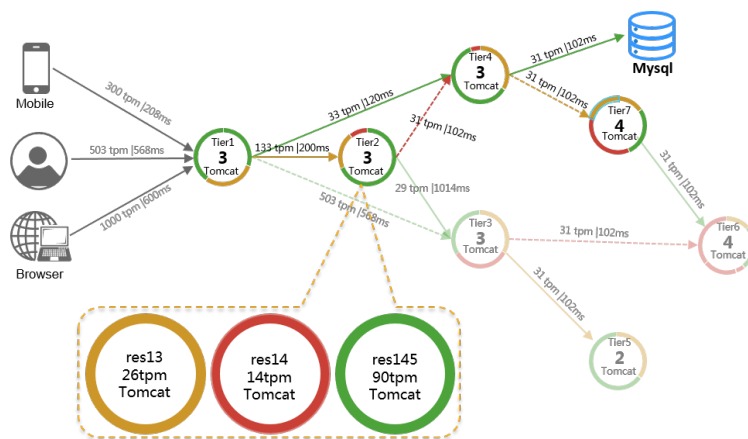
2 Functions

As a cloud application diagnosis service, APM supports full-link topology display, tracing, and transaction analysis.

Topology

- Visible topology: APM displays application call and dependency relationships in topologies. Application Performance Index (**Apdex**) is used to quantify user satisfaction with application performance. Different colors indicate different Apdex value ranges, helping you quickly detect and locate performance problems. **Figure 2-1** shows the application relationships, call data (service and instance metrics), and health status.

Figure 2-1 Topology



Microservice name: Tier2; number of instances: 3; container type: Tomcat

Description: A circle indicates a microservice and is divided into 3 sections based on the number of instances. Different colors indicate different instance statuses. Red indicates abnormal, yellow indicates warning, and green indicates normal.

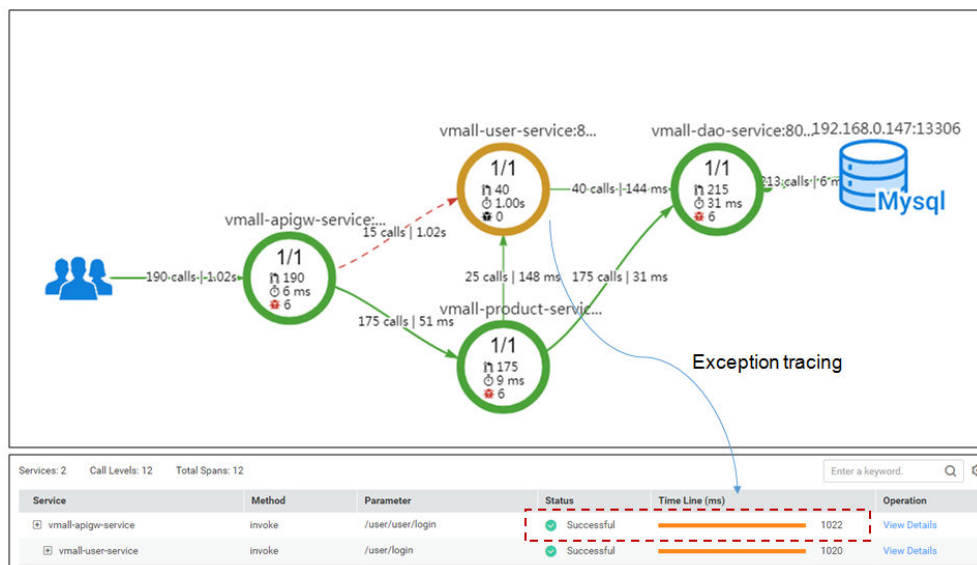
- Inter-application calling: APM can display call relationships between application services on the topology. When services are called across applications, APM can collect inter-application call relationships and display application performance data.

- SQL analysis: APM can count and display key metrics about databases or SQL statements on the topology. APM provides graphs of key metrics such as the number of SQL statement calls, response time, and number of errors. Based on these graphs, you are able to analyze database performance problems caused by slow and incorrect SQL statements.
- JVM metric monitoring: APM can count and display JVM metric data of instances on the topology. APM monitors the memory and thread metrics in the JVM running environment in real time, enabling you to quickly detect memory leakage and thread exceptions.

Tracing

- Tracing: APM comprehensively monitors the number of calls, response time, and number of errors, and displays the service running track and status, helping you quickly demarcate performance bottlenecks and faults. As shown in **Figure 2-2**, if an exception (the red service in the topology) occurs, locate root causes based on the call status, duration, and API call information.

Figure 2-2 Tracing

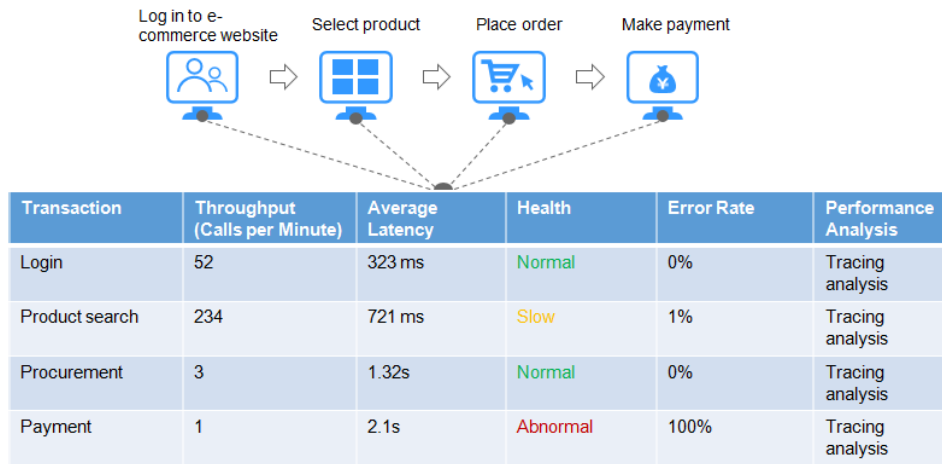


- Method tracing: APM dynamically traces a method of a class. When the method of this class is called, the APM probe collects the call data of the method based on the configured method tracing rule and displays the call data on the tracing page. Method tracing is used to help application developers locate method-level performance problems online.

Transaction Analysis

Transaction analysis: APM analyzes service flows on servers in real time, displays key metrics such as throughput, error rate, and latency of transactions, and uses Apdex to evaluate application performance, intuitively reflecting users' satisfaction with applications. If transactions are abnormal, alarms are reported. For transactions with poor user experience, locate problems through topologies and tracing. **Figure 2-3** shows the transaction status of an e-commerce application. If the health status of a transaction is abnormal, user experience is poor.

Figure 2-3 Transaction analysis



3 Application Scenarios

APM is widely used. You can learn how to use APM based on the following typical application scenarios.

Diagnosis of Application Exceptions

Pain Points

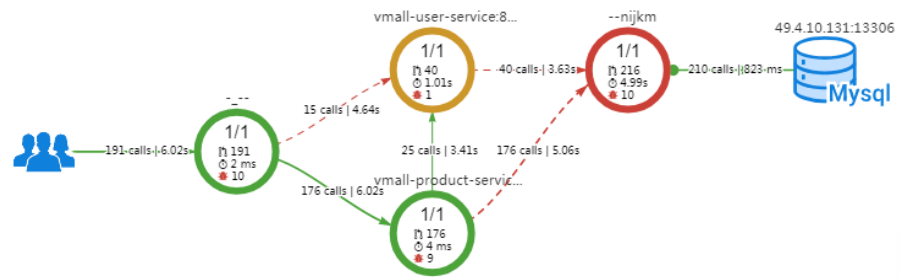
In the distributed microservice architecture, enterprises can develop diverse complex applications efficiently. However, this architecture poses great challenges to traditional O&M and diagnosis technologies. In the example of an e-commerce application, problems are as follows:

- **Difficult fault locating**
After receiving the feedback from customers, customer service personnel submit problems to technical personnel for troubleshooting. In the distributed microservice architecture, a request usually undergoes multiple services/nodes before a result is returned. If a fault occurs, O&M personnel need to repeatedly view logs on multiple hosts to locate the fault. Even for simple problems, troubleshooting requires cooperation from multiple teams.
- **Difficult architecture sort-out**
When service logic becomes complex, it is difficult to find out the downstream services (databases, HTTP APIs, and caches) that an application depends on, and external services that depend on the application from the code perspective. It is also difficult to sort out the service logic, manage the architecture, and plan capacities. For example, enterprises find it hard to determine the number of hosts required for online promotions.

Service Implementation

APM can diagnose exceptions in large distributed applications. When an application breaks down or a request fails, you can locate faults in minutes through topologies and drill-downs.

- **Visible topology:** Abnormal application instances can be automatically discovered on the topology.



- Tracing: You can locate root causes in code through drill-downs after identifying abnormal applications on the topology.

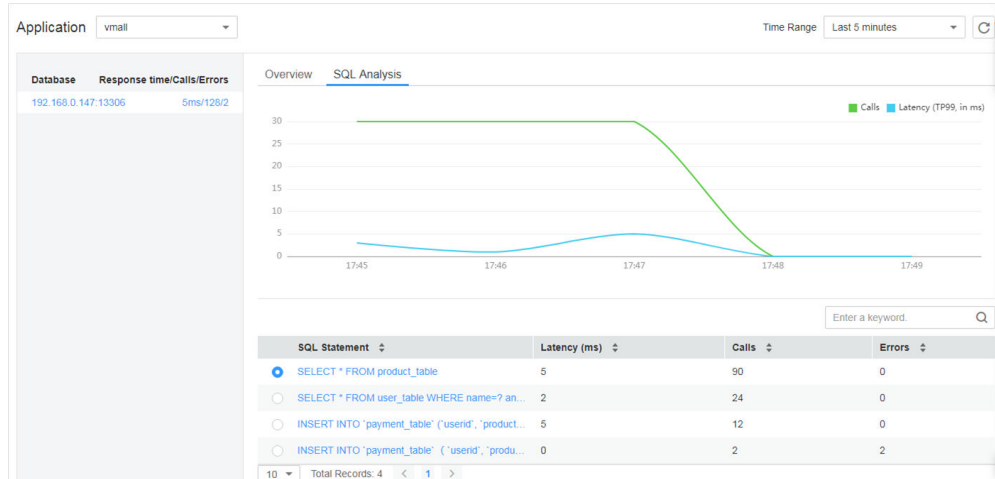
Services: 2 Call Levels: 12 Total Spans: 12 Enter a keyword. 🔍

Service	Method	Parameter	Status	Time Line (ms)	Operation
vmall-apigw-service	invoke	/user/user/login	Successful	1040	View Details
vmall-user-service	invoke	/user/login	Successful	1020	View Details

View Details ✕

Parameter	Value
CLIENT_IP	127.0.0.1
DEST-RESOURCE-ID	vmall-apigw-service:8080 315738...
HOST_IP	127.0.0.1
HOST_NAME	localhost
SRC-RESOURCE-ID	UnmonitoredBrowser
TX-TYPE	ALL_/user/**
USER_AGENT	Mozilla/5.0 (Windows NT 6.1; Wi...
clusterId	UnknownCluster
http.param	workload=1000
http.status.code	200

- SQL analysis: APM displays graphs of key metrics (such as number of SQL statement calls, latency, and number of errors), and supports analysis of database performance problems caused by abnormal SQL statements.



User Experience Management

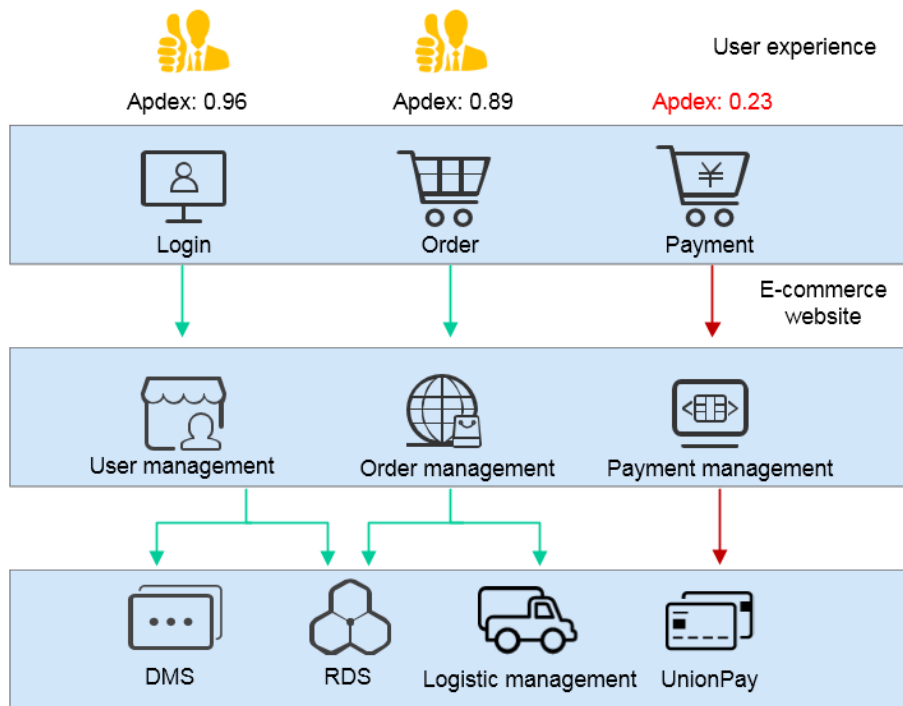
Pain Points

In the Internet era where user experience is of crucial importance, you cannot obtain user access information even if backend services run stably. It is much more difficult to locate frontend problems that occur occasionally. After a system goes online, if users cannot access the system due to errors and you fail to obtain the information in time, you will lose lots of users. If users report page usage problems, how can these problems be reproduced immediately? How can error details be obtained for fast troubleshooting?

Service Implementation

APM provides experience management capabilities. Specifically, it analyzes the complete process (user request > server > database > server > user request) of application transactions in real time, and provides Apdex scores, enabling you to monitor comprehensive user experience in real time. For transactions with poor user experience, locate problems through topologies and tracing.

- Application KPI analysis: KPIs such as throughput, latency, and call success rate are displayed, so that you can monitor user experience easily.
- Full-link performance tracing: Web services, caches, and databases are traced, so that you can detect performance bottlenecks quickly.



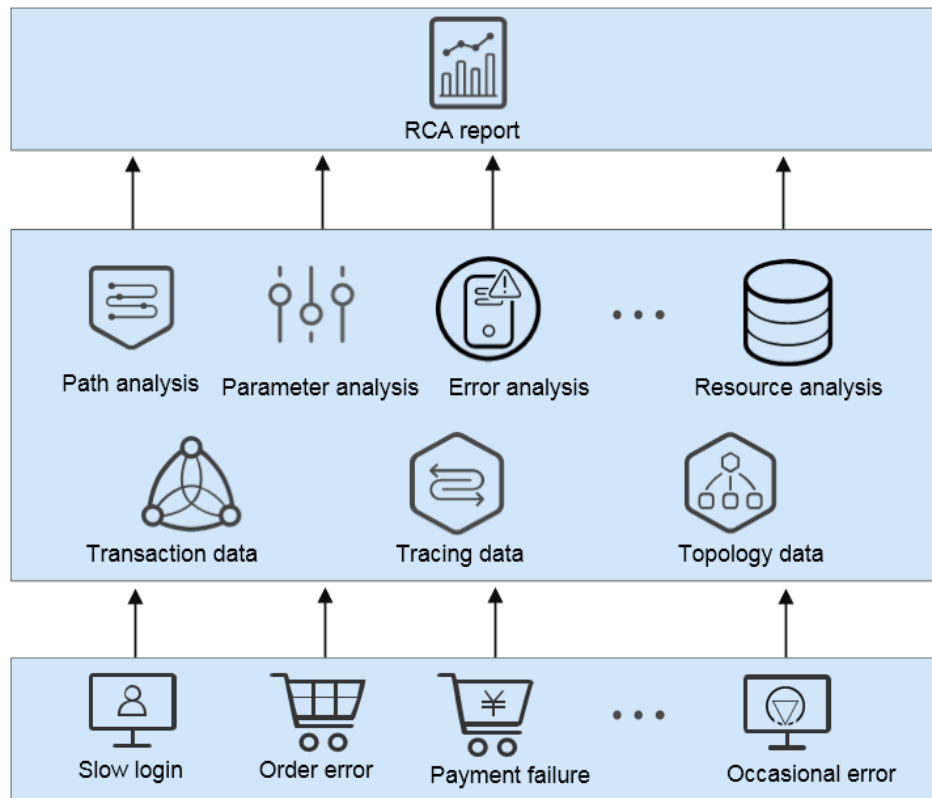
Intelligent Diagnosis

Pain Points

For massive quantities of services, there is rich but unassociated application O&M data, such as hundreds of monitoring metrics, KPI data, and tracing data. How can the system associate metric and alarm data from multiple perspectives (such as applications, services, instances, hosts, and transactions), and automatically complete RCA? How can intelligence analysis be made and possible causes be provided based on the learned historical data and O&M experience library?

Service Implementation

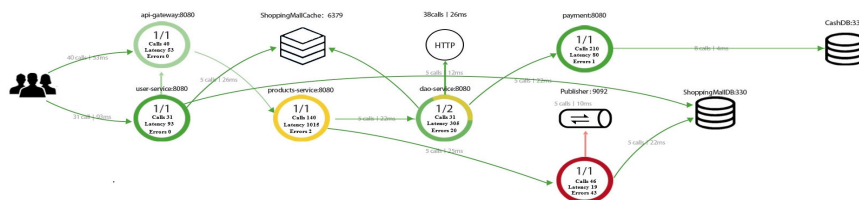
APM supports automatic detection of faults using machine learning algorithms, and intelligent diagnosis. When an exception occurs in a transaction, APM learns historical metric data based on intelligent algorithms, associates exception metrics for multi-dimensional analysis, extracts characteristics of context data (such as resources, parameters, and call structures) when services are normal and abnormal, and locate root causes through cluster analysis. APM can collect and compare the historical data about good and poor experience, and record the environment data that may cause application errors, including input and output parameters, tracing, resource data, and JVM parameters. Based on the Enterprise Intelligent (EI) engine, APM can train historical data online and make predictions.



4 Basic Concepts

Topology

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



Transaction

A transaction is usually an HTTP request (complete process: user request > web server > database > web server > user request). In real life, a transaction is a one-time task. A user completes a task by using an application. 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

An application is a group of the same or similar services categorized based on service requirements. You can put services that fulfill the same function into one application for performance management. For example, you can put accounts, products, and payment services into the **Mall** application.

Apdex

Apdex is an open standard developed by the Apdex alliance. It defines a standard method to measure application performance. The Apdex standard converts the application response time into user satisfaction with application performance in the range of 0 to 1.

- Apdex principle

Apdex defines the optimal threshold (T) for the application response time. T is determined by the performance evaluation personnel 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, the Apdex threshold is the value configured in **Setting Apdex Thresholds**. The application response latency is the service latency. The Apdex value ranges from 0 to 1 and is calculated as follows:

$$\text{Apdex} = (\text{Number of satisfied samples} + \text{Number of tolerating samples} \times 0.5) / \text{Total number of samples}$$

Apdex indicates application performance status, that is, user satisfaction with application performance. Different colors indicate different Apdex ranges, as shown in **Table 4-1**.

Table 4-1 Apdex description

Apdex	Color	Description
$0.75 \leq \text{Apdex} \leq 1$	Green	Fast response; good user experience
$0.3 \leq \text{Apdex} < 0.75$	Yellow	Slow response; fair user experience

Apdex	Color	Description
$0 \leq \text{Apdex} < 0.3$	Red	Very slow response; poor user experience
-	Black	No application, instance, or transaction is invoked.

- Configuring an Apdex threshold

You can configure an Apdex threshold according to [Setting Apdex Thresholds](#).

TP99 Latency

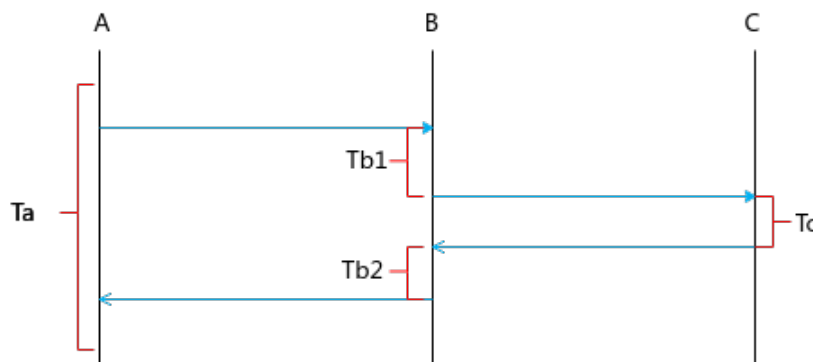
TP99 latency is the minimum time meeting requirements of 99% requests. In APM, latency refers to TP99 latency.

Example: Assume that there are 100 requests, and the time consumed by the requests is 1s, 2s, 3s, 4s...98s, 99s, and 100s. To meet the requirements of 99% requests, at least 99s is required. Therefore, TP99 latency is 99s.

Calculation: Sort all requests by the consumed time in ascending order. TP99 latency = Time consumed by the Nth request. N is the rounded value of $99\% \times$ Total number of requests.

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. The relationship is as follows: Service latency = Overall latency - Latency for calling other services. For example, assume that service A calls service B, and service B calls service C, as shown in the following figure:



- Service A: Overall latency = T_a ; Service latency = $T_a - T_{b1} - T_{b2} - T_c$
- Service B: Overall latency = $T_{b1} + T_{b2} + T_c$; Service latency = $T_{b1} + T_{b2}$
- Service C: Overall latency = T_c ; Service latency = T_c

Probes

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 detection mechanism is enabled and the instance memory is too large, probes enter the hibernation state, that is, stop collecting data. [How Does APM Collect Probe Data?](#)

Mesh

The Istio mesh obtains input and output application program data in non-intrusive mode. Then, the ICAgent and Cloud Container Engine (CCE) Istio mixer obtain and process the mesh data, and report it to APM. You can enable the Istio to collect mesh data. [How Does APM Collect Mesh Data?](#)

ICAgent

ICAgent is a collection agent of APM. It runs on the server where applications are deployed to collect the data obtained by probes in real time. For details about the data collection and purposes, see [APM Service Agreement](#). [Installing the ICAgent](#) is prerequisite for using APM.

5 Edition Differences

APM provides probe and mesh products. Probe products include the basic, professional, enterprise, and platinum editions while mesh products include the basic and professional editions.

 **NOTE**

- The professional, enterprise, and platinum editions of probe products support connection to APM through Java.
- The professional edition of mesh products supports connection to APM through Istio (without language constraints).

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

Edition	Basic	Professional	Enterprise	Platinum
Function 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 professional, enterprise, or platinum edition. If you use APM 50 times, you need to switch to the professional, enterprise, or platinum edition.	Open	Open	Open
Data storage duration	7 days	7 days	30 days	90 days
Application topology	√	√	√	√
Tracing	√	√	√	√
Transaction analysis	√	√	√	√
Slow SQL analysis	√	√	√	√
JVM analysis	√	x	√	√

Edition	Basic	Professional	Enterprise	Platinum
AI capabilities	√	x	√	√
Method tracing	√	x	√	√
Expert support	x	x	24-hour online support	<ul style="list-style-type: none"> • 24-hour online support • One person-day of onsite support in China each quarter
Note: √ indicates supported, and x indicates not supported.				

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

Edition	Basic	Professional
Function 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 professional edition. If you use APM 50 times, you need to switch to the professional edition.	Open
Data storage duration	7 days	7 days
Application topology	√	√
Tracing	√	√
Note: √ indicates supported, and x indicates not supported.		

6 Permissions Management

If you need to assign different permissions to employees in your enterprise to access your 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 cloud 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 cloud 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 cloud account. For more information about IAM, see [Identity and Access Management Service Overview](#).

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 6-1](#) lists all the system permissions supported by APM.

Table 6-1 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 6-2 lists the common operations supported by each system-defined policy or role of APM. Choose appropriate policies or roles as required.

Table 6-2 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	√

7 Privacy and Sensitive Information Protection Statement

All O&M data will be displayed on the APM console. Therefore, you are not advised to upload your privacy or sensitive data to APM. If you need to upload such data, encrypt them.

ICAgent Deployment

When you manually install the ICAgent on an Elastic Cloud Server (ECS), your AK/SK will be used as an input parameter in the installation command. To prevent information leakage, disable historical record collection before installing the ICAgent. After the ICAgent is installed, it will encrypt your AK/SK for storage.

8 Data Collection

When you enable data collection, APM collects service tracing data, resource information, resource attributes, memory detection information, and call request KPI data, but does not collect your personal data. The collected data is used only for performance analysis and fault diagnosis, and is not used for commercial purposes.

Data Type	Collected Data	Transmission Mode	Storage Mode	Function	Storage Period
Service tracing data	Tracing span data	Transmission through HTTPS encryption and Access Key ID/Secret Access Key (AK/SK) authentication	Project-based isolated storage	Query and display at the tracing frontend	Configurable (7 days at most). The data will be deleted upon expiration.

Data Type	Collected Data	Transmission Mode	Storage Mode	Function	Storage Period
Call request KPI data	Call initiator address, receiver address, API, duration, and status	Transmission through HTTPS encryption and AK/SK authentication	Project-based isolated storage	Calculation of transaction call KPI metrics, such as throughput, TP99 latency, average latency, and error calls, drawing of application topologies, and display at the frontend	7 days. The data will be deleted upon expiration.
Resource information	Service type, service name, creation time, deletion time, node address, and service release API	Transmission through HTTPS encryption and AK/SK authentication	Project-based isolated storage	Query and display at the resource library frontend	7 days. The data will be deleted upon expiration.
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	Transmission through HTTPS encryption and AK/SK authentication	Project-based isolated storage	Query and display at the resource library frontend	7 days. The data will be deleted upon expiration.

Data Type	Collected Data	Transmission Mode	Storage Mode	Function	Storage Period
Memory monitoring information	Memory usage, used memory, maximum memory, remaining memory, memory threshold-crossing time, and memory monitoring configurations	Transmission through HTTPS encryption and AK/SK authentication	Project-based isolated storage	Query and display at the resource library frontend	7 days. The data will be deleted upon expiration.

9 Usage Restrictions

Linux OSs Supported by ICAgent

ICAgent supports multiple Linux Operating Systems (OSs). When purchasing an Elastic Cloud Server (ECS), select an OS that is listed in [Table 9-1](#).

Table 9-1 Supported OSs and versions

OS	Version			
SUSE	SUSE Enterprise 11 SP4 64-bit	SUSE Enterprise 12 SP1 64-bit	SUSE Enterprise 12 SP2 64-bit	SUSE Enterprise 12 SP3 64-bit
openSUSE	13.2 64-bit	15.0 64-bit Currently, syslog logs cannot be collected.	42.2 64-bit	
EulerOS	2.2 64-bit	2.3 64-bit	2.5 64-bit	
CentOS	6.3 64-bit	6.5 64-bit	6.8 64-bit	6.9 64-bit
	6.10 64-bit	7.1 64-bit	7.2 64-bit	7.3 64-bit
	7.4 64-bit	7.5 64-bit	7.6 64-bit	7.7 64-bit
	7.8 64-bit	7.9 64-bit	8.0 64-bit	8.1 64-bit
	8.2 64-bit			
Ubuntu	16.04 server 64-bit	18.04 server 64-bit	22.04 server 64-bit	24.04 server 64-bit
Fedora	24 64-bit	25 64-bit	29 64-bit	
Debian	7.5.0 32-bit	7.5.0 64-bit	8.2.0 64-bit	8.8.0 64-bit
	9.0.0 64-bit			

 NOTE

- For Linux x86_64 servers, ICAgent supports all the OSs and versions listed in the preceding table.
- For Linux Arm servers, ICAgent only supports CentOS 7.4 and later versions, and other OSs and versions listed in the preceding table.

Supported Java Types

Currently, APM can connect to Java applications. APM supports multiple mainstream Java frameworks, web servers, communications protocols, and databases. For details, see [Table 9-2](#).

Table 9-2 Supported Java types

Type	Name	Version
Tool	JDK	JDK 7 and JDK 8
Communications protocol	HttpClient	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 and 2.0.4–2.0.9
	Dubbo	2.5.3–2.6.2 (Dubbo RPC and Dubbo REST) and 2.8.4 (Dubbo RPC and Dubbo REST)
	Huawei's Cloud Service Engine (CSE)	1.0.0.B011–1.1.0.B046 (REST over Servlet, REST over Vert.x, and Highway RPC)
	gRPC	1.11.x–1.14.x
Database	MySQL	mysql-connector-java 5.1.X
	Oracle	ojdbc5, ojdbc6, and ojdbc14
	Sybase	2.6.0–3.2.1
	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

Type	Name	Version
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 NOTE When JBoss uses the Java agent probe, special settings are required. For details, see How Do I Connect the JBoss Server in Standalone Mode to APM?	7.0.0–12.0.0
	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
	Cassandra	2.1.x–3.2.x
	ZooKeeper	1.0.x (com.github.adyliu.zkclient) and 0.1.x (com.github.sgroshupf.zkclient)
	Elasticsearch	2.4.x and 5.1.x
Rest Client	Common HTTP	2.x, 3.x, 4.x (HttpClient), and all (HttpURLConnection)

Probe Quantity

APM collects performance metrics, including tracing, SLA, SQL statement call, and JVM metrics. The resources consumed for collecting these metrics are closely related to the number of APM probes, number of inter-service calling times, and sampling ratio. To ensure data collection reliability, you are advised to use a maximum of 20 APM probes on a single node.

10 Billing

Billing

APM supports both pay-per-use and package billing modes. The two modes can be used at the same time. That is, if you use more instances than those included in a package, you will be billed on a pay-per-use basis for the excess instances used. If you use APM without purchasing any package, you will be billed on a pay-per-use basis for all instances. For details, see [Pricing Details](#).

APM provides probe and mesh products. Probe products include the basic, professional, enterprise, and platinum editions while mesh products include the basic and professional editions. After you switch from the basic edition to another edition, the pay-per-use billing mode will apply. If you have purchased a package, instances you already paid for will be preferentially used. After instances included in the package are used up, any additional instances you use will be charged on a pay-per-use basis.

NOTE

- Probes will be billed by hour (rounded up to the nearest one hour).
- A professional-edition probe is one billing unit. An enterprise-edition probe is billed by two billing units. A platinum-edition probe is billed by seven billing units.
- For more information, see [APM Pricing Details](#).

A probe billing unit refers to the basic billing unit of probes during pay-per-use or package billing. Probes of different editions can be converted into certain number of probe billing units for billing. For example, when you purchase a package with 100 professional-edition probes (100 probe billing units), you can choose to use 50 enterprise-edition probes (100 probe billing units/2) or 14 platinum-edition probes (100 probe billing units/7) in one hour.

Package Details

For details about the functions of probe and mesh products, see [Edition Differences](#).

Renewal Details

Packages need to be prepaid. When your package expires or is about to expire, renew it in time. If it is not renewed in time, APM functions may be affected.

Table 10-1 Renewal details

-	Sufficient Balance	Insufficient Balance or Account in Arrears	Retention Period
APM status	You can use APM normally.	Your account is frozen. APM does not collect or display new application data, but still displays the data collected before your account is frozen.	
Renewal details	If you renew your package in time, you can use APM normally. For details, see Manually Renewing a Resource .	If you top up your account, APM automatically unfreezes your account and you will be billed on a pay-per-use basis. If you need a package, renew or purchase one. Otherwise, you will be billed on a pay-per-use basis.	
	If you do not renew your package, APM automatically switches to the pay-per-use billing mode.	If you do not top up your account and the retention period expires, APM releases all your resources and you will not be billed in this period.	

10.1 Pricing Details

APM supports both pay-per-use and package billing. When you subscribe to APM, the package - basic edition is used by default. To use other editions, purchase them as required.

Currently, APM supports Java and PHP applications.

Package Pricing Details of Probe Product - Professional Edition

Table 10-2 Pricing details of the professional edition

Region	Probes	1 Month	1 Year	Currency
CN North-Beijing ⁴ CN East-Shanghai ¹	10	136	1,357	USD
	20	271	2,715	USD
	50	679	6,787	USD
	100	1,357	13,574	USD
	200	2,443	24,433	USD
	500	5,769	57,689	USD
	1,000	10,859	108,592	USD
	2,000	19,004	190,036	USD

Region	Probes	1 Month	1 Year	Currency
CN-Hong Kong	10	220	2,205	USD
	20	441	4,411	USD
AP-Singapore	50	1,102	11,029	USD
AP-Bangkok	100	2,205	22,058	USD
LA-Santiago	200	3,970	39,705	USD
	500	9,375	93,750	USD
	1,000	17,647	176,470	USD
	2,000	30,882	308,823	USD

Package Pricing Details of Probe Product - Enterprise Edition

Table 10-3 Pricing details of the enterprise edition

Region	Probes	1 Month	1 Year	Currency
CN North-Beijing4 CN East-Shanghai1	10	271	2,715	USD
	20	543	5,430	USD
	50	1,357	13,574	USD
	100	2,443	24,433	USD
	200	4,887	48,866	USD
	500	10,859	108,592	USD
	1,000	19,004	190,036	USD
	2,000	38,007	380,071	USD
CN-Hong Kong AP-Singapore AP-Bangkok LA-Santiago AF-Johannesburg	10	441	4,411	USD
	20	882	8,823	USD
	50	2,205	22,058	USD
	100	3,970	39,705	USD
	200	7,941	79,411	USD
	500	17,647	176,470	USD
	1,000	30,882	308,823	USD
	2,000	61,764	617,647	USD

Package Pricing Details of Probe Product - Platinum Edition

Table 10-4 Pricing details of the platinum edition

Region	Probes	1 Month	1 Year	Currency
CN North-Beijing4 CN East-Shanghai1	200	19,004	190,036	USD
	500	38,007	380,071	USD
	1,000	80,765	807,651	USD
	2,000	133,025	1,330,249	USD
CN-Hong Kong AP-Singapore AP-Bangkok LA-Santiago AF-Johannesburg	200	30,882	308,823	USD
	500	61,764	617,647	USD
	1,000	131,250	1,312,500	USD
	2,000	216,176	2,161,764	USD

Package Pricing Details of Probe Product - Basic Edition

Table 10-5 Pricing details of the basic edition

Region	Probes	1 Month	1 Year	Currency
CN North-Beijing4 CN East-Shanghai1 CN-Hong Kong AP-Singapore AP-Bangkok LA-Santiago AF-Johannesburg	Unlimited (Data collection lasts for one hour each time and then automatically stops. Max. data collection times: 50.)	0	0	USD

Pay-per-Use Pricing Details of Different Probe Product Editions

Table 10-6 Pay-per-use pricing details of different probe product editions

Region	Billing Mode	Basic	Professional	Enterprise	Platinum	Pricing Basis (USD)
CN North-Beijing4 CN East-Shanghai 1 CN-Hong Kong AP-Singapore AP-Bangkok LA-Santiago AF-Johannesburg	Pay-per-use	0	0.044	0.088	0.308	Price per instance-hour
ME-Abu Dhabi-OP5	Pay-per-use	0	0.0485	0.09700000 000000002	0.3395	Price per instance-hour

NOTE

- Probes will be billed by hour (rounded up to the nearest one hour).
- A professional-edition probe is one billing unit. An enterprise-edition probe is billed by two billing units. A platinum-edition probe is billed by seven billing units. A probe billing unit refers to the basic billing unit of probes during pay-per-use or package billing. Probes of different editions can be converted into certain number of probe billing units for billing.
- Packages are purchased with one-off payments and take effect immediately upon your payment. You cannot specify a later date for the packages to take effect.
- Multiple packages can be used together.
- During the validity period of a package, if you use more resources than those included in the package, you will be billed on a pay-per-use basis for the excess resources used.
- Packages need to be renewed before they expire. Otherwise, resources will be billed on a pay-per-use basis.
- Package unsubscription is not supported. After a resource package expires, you can continue using APM and your data is still secure.

Billing Rules for Pay-per-Use Resources

- Standard pricing: You pay only for what you use, based on the unit price of the resources you use and how long or often you use them. The cost is accurate to two decimal places.
- Tiered pricing: Pay less by using longer or more. The cost is accurate to two decimal places.
- For pay-per-use billing, decimal numerals on the price calculator are rounded off and are accurate to two decimal places. If the fee is less than \$0.01 USD (after rounding off), \$0.01 USD will be displayed.

11 Change History

Pinpoint Description

Table 11-1 Pinpoint description

Version	Released On	Description
1.0.47	2021-5-22	<ol style="list-style-type: none">1. Supported Kafka clients 2.0.x-2.8.x.2. Supported mariadb-java-client 1.2.x-2.4.x.3. Fixed known Undertow tracing issues.
1.0.43	2021-2-2	<ol style="list-style-type: none">1. Supported Cloud Service Engine (CSE) 2.0.2. Fixed known issues.
1.0.39	2020-11-23	<ol style="list-style-type: none">1. Optimized the transaction matching mechanism.2. Supported monitoring for direct buffer memory.
1.0.38	2020-11-09	Optimized the transaction self-learning capability.
1.0.36	2020-10-10	<ol style="list-style-type: none">1. Optimized ServiceComb tracing.2. Optimized JVM thread monitoring.
1.0.32	2020-8-27	<ol style="list-style-type: none">1. Supported Tomcat monitoring.2. Optimized the gRPC plug-in.
1.0.29	2020-05-09	Supported the Hystrix asynchronous mode of Spring Cloud OpenFeign.
1.0.28	2020-02-12	Fixed the Redis Lettuce bug (that is, memory leakage may occur).

Version	Released On	Description
1.0.27	2019-11-30	<ol style="list-style-type: none">1. Supported the MySQL mysql-connector-java-8.x.x client.2. Supported the Redis Lettuce client.3. Supported the new Jetty version.4. Fixed the Netty bug (that is, memory leakage may occur).5. Optimized probe resource usage.
1.0.14 (latest CCE version)	2019-10-30	Supported Apache Dubbo 2.7.1.
1.0.4 CCE	2019-05-06	<ol style="list-style-type: none">1. Supported gRPC 1.11.x-1.14.x.2. Supported Thrift 0.11.
0.1.36 CCE	2019-03-15	<ol style="list-style-type: none">1. Supported ServiceComb 2.3.25-2.3.52.2. Supported Netty 4.1.22.3. Solved the JDK HTTP tracing interruption.4. Supported Docker application access for non-CCE users.
0.1.25 CCE	2018-12-24	<ol style="list-style-type: none">1. Supported JVM metric monitoring.2. Supported configuration of return codes and exceptions to be ignored.3. Added the http.url attribute to the tracing of the JDK HTTP client.4. Supported parameter collection for method tracing.

Feature Description

Table 11-2 Feature description

Released On	Description
2019-06-11	Supported transaction grouping and rule customization.
2019-06-04	Supported installation, upgrade, and uninstall of the ICAgent for non-Huawei Cloud hosts.
2019-05-22	<ol style="list-style-type: none">1. Connected Docker applications to APM.2. Connected Deployment applications to APM.

Released On	Description
2019-04-18	Obtained an Access Key ID/Secret Access Key (AK/SK) by creating an Identity and Access Management (IAM) agency.
2019-03-15	Supported customization of threshold rules and statistics for JVM monitoring.
2019-03-06	Optimized service and instance metrics, and displayed topology details to enhance full-link topology functions.
2018-12-26	<ol style="list-style-type: none"> Monitoring of JVM memory and thread: Displays metrics about the JVM heap memory, JVM non-heap memory, and thread to quickly detect exceptions such as memory leakage and thread suspension. Exceptions whitelist: Set return codes and exceptions to be ignored.
2018-07-20	Supported both pay-per-use and package billing modes. For more information, see Application Performance Management Pricing Details .
2018-06-11	Launched APM at the Huawei Cloud International website.

ICAgent Description

Table 11-3 ICAgent description

ICAgent Version	Description
5.12.43	Resolved the problem that the ICAgent occasionally restarts during collection of container logs.
5.12.42	Resolved the problem that the number of network connections in close_wait state is sometimes too large during collection of rsyslog logs.
5.12.36	<ul style="list-style-type: none"> Supported display of O&M information by application or component in CCE scenario. Optimized the calculation formula of node CPU usage.
5.12.35	Fixed log truncation in some scenarios.
5.12.32	Solved the problem that the ICAgent may restart during log collection.
5.12.29	Disabled metric collection for D310 chips.
5.12.27	Optimized application discovery rules.

ICAgent Version	Description
5.12.26	Supported collection of text file logs in a specified path. There are no restrictions on file suffixes.
5.12.23	Added the file system usage metric for containers.
5.12.22	Supported log paths containing wildcards for recursive traversal.