Alibaba Cloud Log Service

User Guide

Issue: 20180925

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.
2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company, or individual in any form or by any means without the prior written consent of Alibaba Cloud.
3. The content of this document may be changed due to product version upgrades, adjustments, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and the updated versions of this document will be occasionally released through Alibaba Cloud-authorized channels. You shall pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.
4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides the document in the context that Alibaba Cloud products and services are provided on an "as is", "with all faults" and "as available" basis. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not bear any liability for any errors or financial losses incurred by any organizations, companies, or individuals arising from their download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, bear responsibility for any indirect, consequential, exemplary, incidental, special, or punitive damages, including lost profits arising from the use or trust in this document, even if Alibaba Cloud has been notified of the possibility of such a loss.
5. By law, all the content of the Alibaba Cloud website, including but not limited to works, products, images, archives, information, materials, website architecture, website graphic layout, and webpage design, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade

secrets. No part of the Alibaba Cloud website, product programs, or content shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates).

6. Please contact Alibaba Cloud directly if you discover any errors in this document.

Generic conventions

Table -1: Style conventions

Style	Description	Example
	This warning information indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	 **Danger:** Resetting will result in the loss of user configuration data.
	This warning information indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	 **Warning:** Restarting will cause business interruption. About 10 minutes are required to restore business.
	This indicates warning information, supplementary instructions, and other content that the user must understand.	 **Note:** Take the necessary precautions to save exported data containing sensitive information.
	This indicates supplemental instructions, best practices, tips, and other content that is good to know for the user.	 **Note:** You can use **Ctrl** + **A** to select all files.
>	Multi-level menu cascade.	**Settings** > **Network** > **Set network type**
Bold	It is used for buttons, menus, page names, and other UI elements.	Click **OK**.
`Courier font`	It is used for commands.	Run the `cd /d C:/windows` command to enter the Windows system folder.
Italics	It is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	It indicates that it is a optional value, and only one item can be selected.	`ipconfig` *[-all|-t]*
{} or {a|b}	It indicates that it is a required value, and only one item can be selected.	`swich` *{stand | slave}*

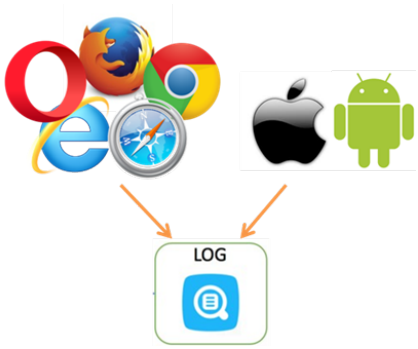
Contents

Legal disclaimer I
Generic conventions I
1 Data Collection 1
1.1 Web Tracking 1
1.2 Logstash 5
1.2.1 Quick installation 5
1.2.2 Custom installation 5
1.2.3 Set Logstash as a Windows service 7
1.2.4 Create Logstash collection configurations 9
1.2.5 Advanced functions 11
1.2.6 Logstash error processing 12
1.3 SDK collection 12
1.3.1 Producer Library 12
1.3.2 Log4j Appender 15
1.3.3 C Producer Library 15
1.4 Collection acceleration 16
1.4.1 Overview 16
1.4.2 Enable Global Acceleration 21
1.4.3 Disable Global Acceleration 26
2 Logtail collection 28
2.1 Limits 28
3 Index and query 32
3.1 Overview 32
3.2 Data type of index 36
3.2.1 Text type 36
3.2.2 Value type 38
3.3 Query syntax 38
3.4 Context query 43
3.5 Other functions 46

1 Data Collection

1.1 Web Tracking

Log Service supports collecting logs from HTML, H5, iOS, and Android platforms by using Web Tracking, and customizing dimensions and metrics.



As shown in the preceding figure, you can collect user information from various browsers, iOS apps, and Android apps (apart from *iOS/Android SDK*) by using Web Tracking. For example:

- Browsers, operating systems, and resolutions used by users.
- Browsing behaviors of users, such as the clicking behaviors and purchasing behaviors on the website.
- The staying time in the app for users and whether the users are active or not.

Note:
Using Web Tracking means that this Logstore enables the anonymous write permission of the Internet, and dirty data may be generated.

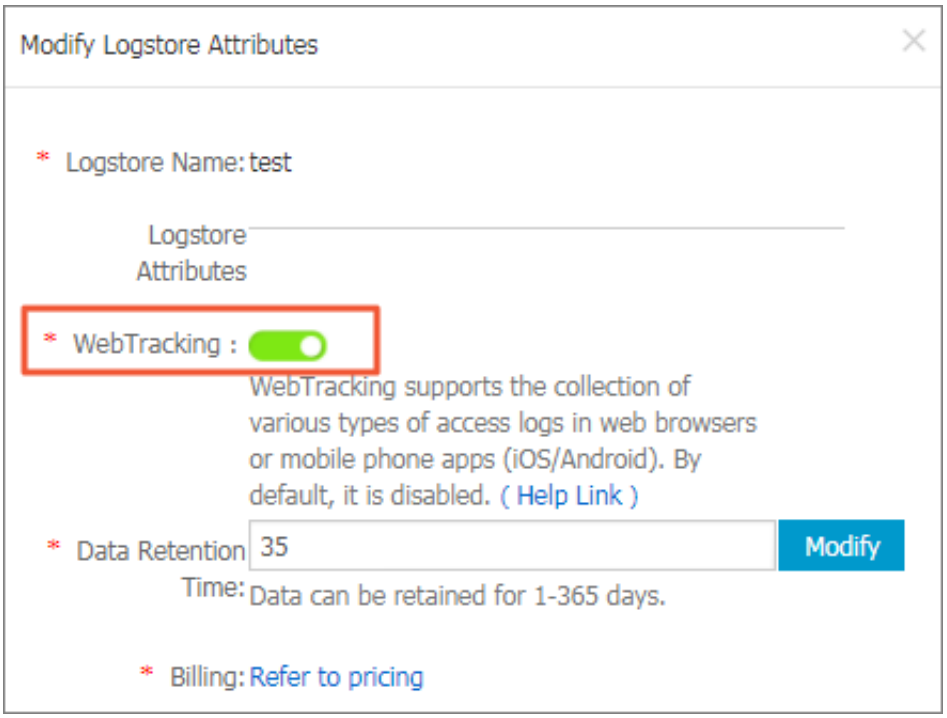
Procedure

Step 1. Enable Web Tracking

You can enable Web Tracking in the console or by using Java SDK.

- **Enable Web Tracking in the console**

  1. On the Logstore List page, click **Modify** at the right of the Logstore that must enable the Web Tracking function.
  2. Turn on the Web Tracking switch.



- **Enable Web Tracking by using**

  Java SDK:

```
import com.aliyun.openservices.log.Client;
import com.aliyun.openservices.log.common.LogStore;
import com.aliyun.openservices.log.exception.LogException;
public class WebTracking {
  static private String accessId = "your accesskey id";
  static private String accessKey = "your accesskey";
  static private String project = "your project";
  static private String host = "log service data address";
  static private String logStore = "your logstore";
  static private Client client = new Client(host, accessId,
accessKey);
  public static void main(String[] args) {
      try {
          //Enable the Web Tracking function on the created Logstore
.
          LogStore logSt = client.GetLogStore(project, logStore).
GetLogStore();
          client.UpdateLogStore(project, new LogStore(logStore,
logSt.GetTtl(), logSt.GetShardCount(), true));
          //Disable the Web Tracking function.
          //client.UpdateLogStore(project, new LogStore(logStore,
logSt.GetTtl(), logSt.GetShardCount(), false));
          //Create a Logstore that supports the Web Tracking
function.
          //client.UpdateLogStore(project, new LogStore(logStore, 1
, 1, true));
      }
      catch (LogException e){
          e.printStackTrace();
      }
  }
```

```
}
```

Step 2. Collect logs

After the Web Tracking function is enabled for Logstore, you can use any of the following three methods to upload data to the Logstore.

- **Use HTTP GET request**

```
curl --request GET 'http://${project}.${host}/logstores/${logstore}/
track? APIVersion=0.6.0&key1=val1&key2=val2'
```

The parameter meanings are as follows.

Field	Meaning
`${project}`	The name of the project created in Log Service.
`${host}`	The domain name of the region where your Log Service is located.
`${logstore}`	The name of the Logstore with the Web Tracking function enabled under `${project}` .
`APIVersion=0.6.0`	The reserved field, which is required.
`__topic__=yourtopic`	Specify the log topic, reserved fields (optional).
`key1=val1`, `key2=val2`	The key-value pairs to be uploaded to Log Service. Multiple key-value pairs are supported, but you must make sure that the URL length is less than 16 KB.

- **Use HTML img tag**

```
<img src='http://${project}.${host}/logstores/${logstore}/track.gif
? APIVersion=0.6.0&key1=val1&key2=val2'/>
<img src='http://${project}.${host}/logstores/${logstore}/track_ua.
gif? APIVersion=0.6.0&key1=val1&key2=val2'/>
```

The parameter meanings are the same as those in Use HTTP GET request.

- **Use JS SDK**

1. Copy *loghub-tracking.js* to the *web* directory, and introduce the following script on the page:

Click to download.

```
<script type="text/javascript" src="loghub-tracking.js" async></
script>
```

Note:

To keep page loading running, the script sends HTTP requests asynchronously. If data must be sent several times in the page loading process, the subsequent request overwrites the preceding HTTP request, and the browser shows the tracking request exits. Sending requests synchronously can help to avoid this problem. To send requests synchronously, replace the statement in the script.

Original script:

```
this.httpRequest_.open("GET", url, true)
```

Replace the last parameter to send requests synchronously:

```
this.httpRequest_.open("GET", url, false)
```

2. Create a Tracker object.

```
var logger = new window.Tracker('${host}','${project}','${logstore
}');
logger.push('customer', 'zhangsan');
logger.push('product', 'iphone 6s');
logger.push('price', 5500);
logger.logger();
logger.push('customer', 'lisi');
logger.push('product', 'ipod');
logger.push('price', 3000);
logger.logger();
```

The parameter meaning are as follows:

Field	Meaning
`${host}`	The domain name of the region where your Log Service is located.
`${project}`	The name of the project created in Log Service.
`${logstore}`	The name of the Logstore with the Web Tracking function enabled under `${project}`.

After running the preceding commands, you can see the following two logs in Log Service:

```
customer:zhangsan
product:iphone 6s
price:5500
```

```
customer:lisi
product:ipod
price:3000
```

After data is uploaded to Log Service, you can use Log Service to *ship* data to Object Storage Service (OSS). You can also use the Consumer Library provided by Log Service to consume data.

1.2 Logstash

1.2.1 Quick installation

You can choose to install logtash quickly on your server by default.

Context

Log Service provides an installation package based on Logstash 2.2.2, which integrates with JRE 1.8, Log Service write plug-in, and NSSM 2.24. The deployment process by using this package is simpler than *Custom installation*. You can select the custom installation for complex requirements.

Procedure

1. Download and extract the *installation package* to the C: drive.
2. Confirm the Logstash startup program path is *C:\logstash-2.2.2-win\bin\logstash.bat*.

1.2.2 Custom installation

You can install Logstash by using quick installation or custom installation methods.

Context

When you have other requirements for logstroudsburg's installation configuration, you can choose how you want to customize the installation, modify the default installation configuration.

Procedure

1. Install Java
   1. Download the installation package.

      Go to the *Java official website* to download JDK for installation.

2. Sets the environment variable.

   Add or modify environment variables in advanced system settings.

   - **PATH:** *C:\Program Files\Java\jdk1.8.0_73\bin*
   - **CLASSPATH:** *C:\Program Files\Java\jdk1.8.0_73\lib;C:\Program Files\Java\jdk1.8.0_73\lib\tools.jar*
   - **JAVA_HOME:** *C:\Program Files\Java\jdk1.8.0_73*

3. Perform verification.

   Run `PowerShell` or `cmd.exe` for verification.

```
PS C:\Users\Administrator> java -version
java version "1.8.0_73"
Java(TM) SE Runtime Environment (build 1.8.0_73-b02)
Java HotSpot(TM) 64-Bit Server VM (build 25.73-b02, mixed mode)
PS C:\Users\Administrator> javac -version
javac 1.8.0_73
```

2. Install Logstash

   1. Download the installation package from the official website.

      Select version 2.2 or later on the *Logstash* home page.

   2. Install Logstash.

      Extract *logstash-2.2.2.zip* to the *C:\logstash-2.2.2* directory.

      Confirm the Logstash startup program path is *C:\logstash-2.2.2\bin\logstash.bat* .

3. Install the plug-in used by Logstash to write logs to Log Service

   Install the plug-in online or offline based on the network environment where the machine resides.

   - Online installation

     The plug-in is hosted by RubyGems. For more information, see *here* .

     Run `PowerShell` or `cmd.exe` to go to the Logstash installation directory.

```
PS C:\logstash-2.2.2> .\bin\plugin install logstash-output-logservice
```

   - Offline installation

     Download from the official website. Go to the *logstash-output-logservice* page and click **Download** in the lower-right corner.

If the machine from which logs are collected cannot access the Internet, copy the downloaded gem package to the *C:\logstash-2.2.2* directory of the machine. Run *PowerShell* or *cmd.exe* to go to the Logstash installation directory. Perform the following command to install lLogstash:

```
PS C:\logstash-2.2.2> .\bin\plugin install C:\logstash-2.2.2\
logstash-output-logservice-0.2.0.gem
```

- Perform verification.

```
PS C:\logstash-2.2.2> .\bin\plugin list
```

Verify that logstash-output-logservice exists in the installed plug-in list of the machine.

4. Install NSSM

Download from the official website. Go to the *NSSM official website* to download the NSSM installation package.

fter you download the installation package to the local machine, extract it to the *C:\logstash-2.2.2\nssm-2.24*.

1.2.3 Set Logstash as a Windows service

When logstash.bat is started in PowerShell, the Logstash process is working in the frontend. Logstash is generally used for testing configurations and debugging collections. Therefore, we recommend that you set Logstash as a Windows service after the debugging is passed so as to enable Logstash to work in the backend and start automatically when power-on.

Besides setting Logstash as a Windows service, you can also start, stop, modify, and delete the service by using command lines. For more information about how to use NSSM, see *NSSM official document*.

Add Logstash as a Windows service

This operation is generally performed when Logstash is deployed for the first time. If Logstash has been added, skip this step.

Run the following command to add Logstash as a Windows service.

- 32 -bit system

```
C:\logstash-2.2.2-win\nssm-2.24\win32\nssm.exe install logstash "C:\
logstash-2.2.2-win\bin\logstash.bat" "agent -f C:\logstash-2.2.2-win
\conf"
```

- 64. -bit system

```
C:\logstash-2.2.2-win\nssm-2.24\win64\nssm.exe install logstash "C:\
logstash-2.2.2-win\bin\logstash.bat" "agent -f C:\logstash-2.2.2-win
\conf"
```

Start the service

If the configuration file in the Logstash *conf* directory is updated, stop the Logstash service and then start it again.

Run the following command to start the service.

- 32 -bit system

```
C:\logstash-2.2.2-win\nssm-2.24\win32\nssm.exe start logstash
```

- 64 -bit system

```
C:\logstash-2.2.2-win\nssm-2.24\win64\nssm.exe start logstash
```

Stop the service

Run the following command to stop the service.

- 32 -bit system

```
C:\logstash-2.2.2-win\nssm-2.24\win32\nssm.exe stop logstash
```

- 64 -bit system

```
C:\logstash-2.2.2-win\nssm-2.24\win64\nssm.exe stop logstash
```

Modify the service

Run the following command to modify the service.

- 32 -bit system

```
C:\logstash-2.2.2-win\nssm-2.24\win32\nssm.exe edit logstash
```

- 64 -bit system

```
C:\logstash-2.2.2-win\nssm-2.24\win64\nssm.exe edit logstash
```

Delete the service

Run the following command to delete the service.

- 32 -bit system

```
C:\logstash-2.2.2-win\nssm-2.24\win32\nssm.exe remove logstash
```

- 64 -bit system

```
C:\logstash-2.2.2-win\nssm-2.24\win64\nssm.exe remove logstash
```

1.2.4 Create Logstash collection configurations

Context

Related plug-ins

- **logstash-input-file**

  This plug-in is used to collect log files in tail mode. For more information, see*logstash-input-file*.

  Note:
  path indicates the file path, which must use UNIX separators, for example, *C:/test/multiline/*.log*. Otherwise, fuzzy match is not supported.

- **logstash-output-logservice**

  This plug-in is used to output the logs collected by the logstash-input-file plug-in to Log Service.

Parameters	Description
endpoint	Log Service endpoint. Example: *http://regionid.example.com*. For more information, see Log Service endpoint.
project	The project name of Log Service.
logstore	The Logstore name.
topic	The log topic name. The default value is null.

Parameters	Description
source	The log source. If this parameter is set to null, the IP address of the current machine is used as the log source. Otherwise, the log source is subject to the specified parameter value.
access_key_id	The AccessKey ID of the Alibaba Cloud account.
access_key_secret	The AccessKey Secret of the Alibaba Cloud account.
max_send_retry	The maximum number of retries performed when data packets cannot be sent to Log Service because of an exception. Data packets with retry failures are discarded. The retry interval is 200 ms.

Procedure

1. Create collection configurations

   Create a configuration file in the `C:\logstash-2.2.2-win\conf\` directory and then restart Logstash to apply the file.

   You can create a configuration file for each log type. The file name format is `*.conf`. For easier management, we recommend that you create all the configuration files in the `C:\logstash-2.2.2-win\conf\` directory.

   Note:
   The configuration file must be encoded in UTF-8 format without BOM. You can use Notepad++ to modify the file encoding format.

   - IIS logs

     For more information, see *Use Logstash to collect IIS logs*.
   - CSV logs

     Use the system time of log collection as the log uploaded time. For more information, see CSV log configuration.
   - Logs with built-in time

     Take CSV log format as an example. Use the time in the log content as the log uploaded time. For more information, see *Use Logstash to collect CSV logs*.
   - General logs

By default, the system time of log collection is used as the log uploaded time. Log fields are not parsed. Single-line logs and multiline logs are supported. For more information, see *Use Logstash to collect other logs*.

2. Verify configuration syntax

   1. Run `PowerShell` or `cmd.exe` to go to the Logstash installation directory:

      ```
      PS C:\logstash-2.2.2-win\bin> .\logstash.bat agent --configtest --config C:\logstash-2.2.2-win\conf\iis_log.conf
      ```

   2. Modify the collection configuration file. Temporarily add a line of rubydebug configuration in the output phase to output the collection results to the console. Set the type field as per your needs.

      ```
      output {
      If [type] = "***"{
        stdout { codec => rubydebug }
        logservice {

        }
      ```

   3. Run `PowerShell` or `cmd.exe` to go to the Logstash installation directory and start the process:

      ```
      PS C:\logstash-2.2.2-win\bin> .\logstash.bat agent -f C:\logstash-2.2.2-win\conf
      ```

   After the verification, end the *`logstash.bat`* process and delete the temporary configuration item rubydebug.

What's next

When *`logstash.bat`* is started in PowerShell, the Logstash process is working in the frontend. Logstash is generally used for testing configurations and debugging collections. Therefore, we recommend that you set Logstash as a Windows service after the debugging is passed so as to enable Logstash to work in the backend and start automatically when power-on. For how to set Logstash as a Windows service, see *Set Logstash as a Windows service*.

1.2.5 Advanced functions

Logstash provides *multiple plug-ins* to meet personalized requirements. For example:

- *grok*: Structurally parses logs into multiple fields by using regular expressions.
- *json_lines* and *json*: Structurally parses JSON logs.

- *date*: Parses and converts the date and time fields of logs.
- *multiline*: Customizes complex types of multiline logs.
- *kv*: Structurally parses logs of key-value pair type.

1.2.6 Logstash error processing

If you encounter the following collection errors when using Logstash to collect logs, follow the corresponding suggestions and process the errors.

If you encounter the following collection errors when using Logstash to collect logs, follow the corresponding suggestions and process the errors.

- Data with garbled characters in Log Service

  Logstash supports UTF-8 file encoding by default. Check whether input files are correctly encoded or not.

- Error message in the console

  The error `io/console not supported; tty will not be manipulated` is prompted in the console. However, the error does not affect the functions and can be ignored.

  If other errors occur, we recommend that you search Google or Logstash forums for help.

1.3 SDK collection

1.3.1 Producer Library

LogHub Producer Library is a LogHub class library written for high-concurrency Java applications. Producer Library and *Consumer Library* are the read and write packaging for LogHub to lower the threshold for data collection and consumption.

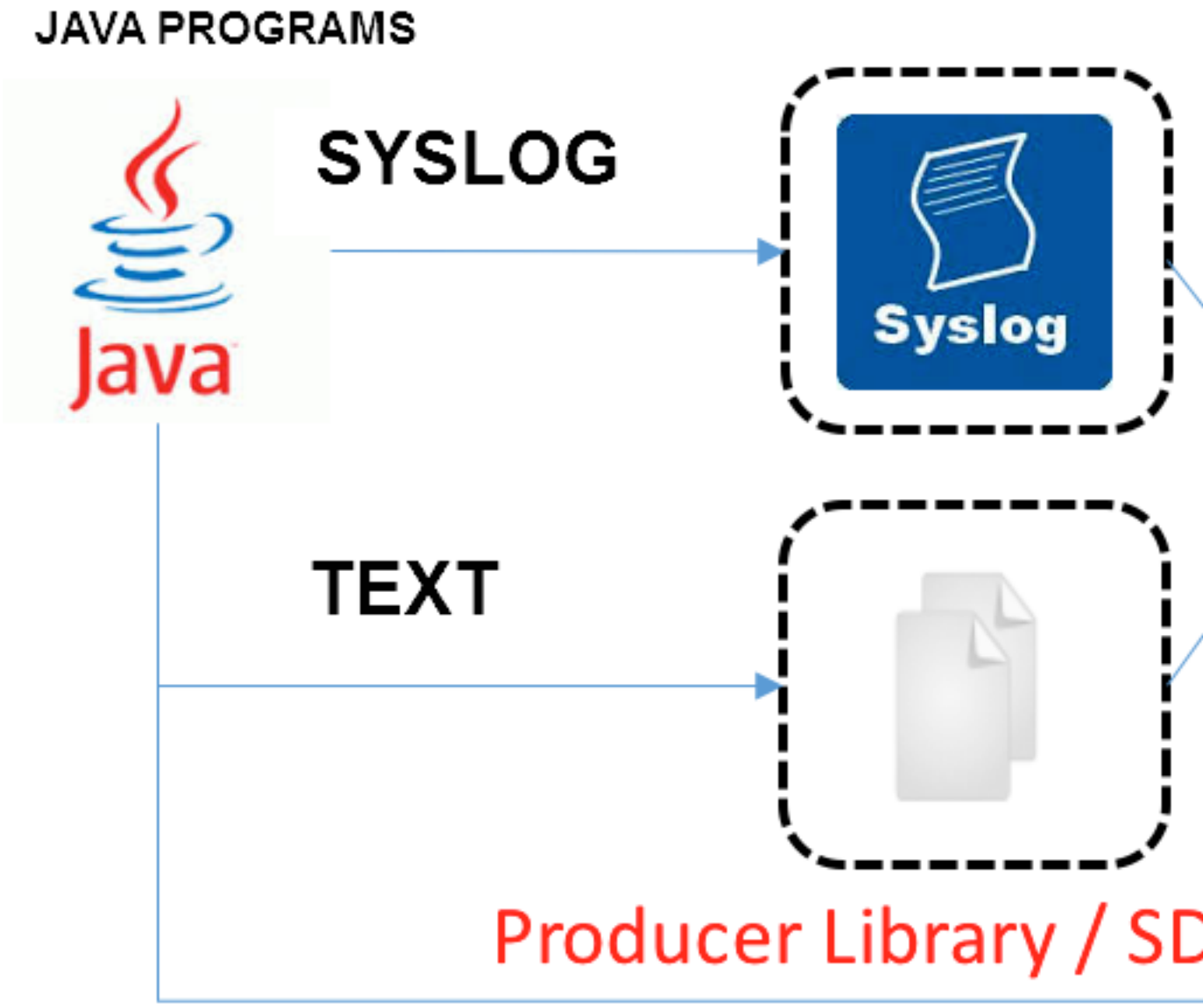
Function features

- Provides an asynchronous send interface to guarantee the thread security.
- Configurations of multiple projects can be added.
- The number of network I/O threads used for sending logs can be configured.
- The number and size of logs of a merged package can be configured.
- The memory usage is controllable. When the memory usage reaches your configured threshold value, the send interface of producer is blocked until idle memory is available.

Function advantages

- Logs collected from the client are not flushed into the disk. Data is directly sent to Log Service by using the network after being generated.
- High concurrency write operations on the client. For example, more than one hundred write operations are performed in one second.
- Client computing logically separated from I/O. Printing logs does not affect the computing time used.

In the preceding scenarios, Producer Library simplifies your program development steps, aggregates write requests in batches, and sends the requests to the LogHub server asynchrono usly. During the process, you can configure the parameters for aggregation in batches and the logic to process server exception.



Compare the preceding access methods:

Access method	Advantages/disadvantages	Scenario
Log flushed into the disk + Logtail	Log collection decoupled from logging, no need to modify the code.	Common scenarios

Access method	Advantages/disadvantages	Scenario
Syslog + Logtail	Good performance (80 MB/s). Logs are not flushed into the disk. The syslog protocol must be supported.	Syslog scenarios.
SDK direct transmission	Not flushed into the disk, and directly sent to the server. Switching between the network I/O and program I/O must be properly processed.	Logs are not flushed into the disk.
Producer Library	Not flushed into the disk, asynchronously merged and sent to the server, with good throughput.	Logs are not flushed into the disk and the client QPS is high.

Procedure

- *Java Producer*
- *Log4J1. Log4J1.XAppender (based on Java Producer)*
- *Log4J2. XAppender (based on Java Producer)*
- *LogBack Appender (based on Java Producer)*
- *C Producer*
- *C Producer Lite*

1.3.2 Log4j Appender

Log4j is an open-source project of Apache, which allows you to set the log output destination to console, file, GUI component, socket server, NT event recorder, or UNIX Syslog daemon. You can also set the output format and level of each log to control log generation with a finer granularity. These configurations can be performed flexibly by using a configuration file without modifying application codes.

Alibaba Cloud Log4j Appender allows you to set the log output destination to Alibaba Cloud Log Service. For more information about download link and user guide, refer to *Github*.

1.3.3 C Producer Library

Besides the Producer Library of Java version, LogHub also supports the Producer Library and Producer Lite Library of the C version, which provides you with a simple and high-performance one-stop log collection solution across platforms and with low consumption of resources.

For the GitHub project address, see:

- *C Producer Library (recommended for servers)*

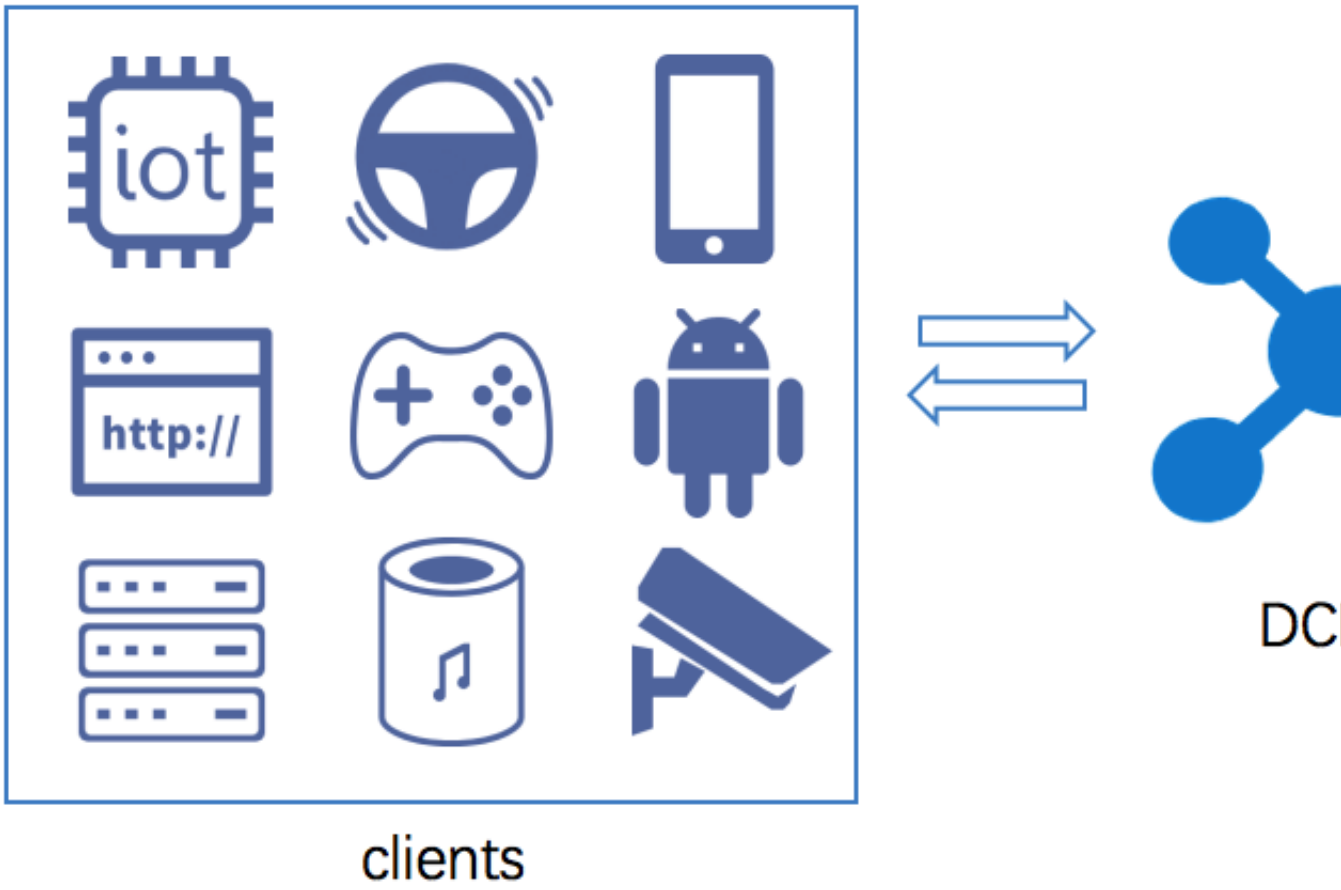
- *C Producer Lite Library (recommended for IOT and smart devices)*

1.4 Collection acceleration

1.4.1 Overview

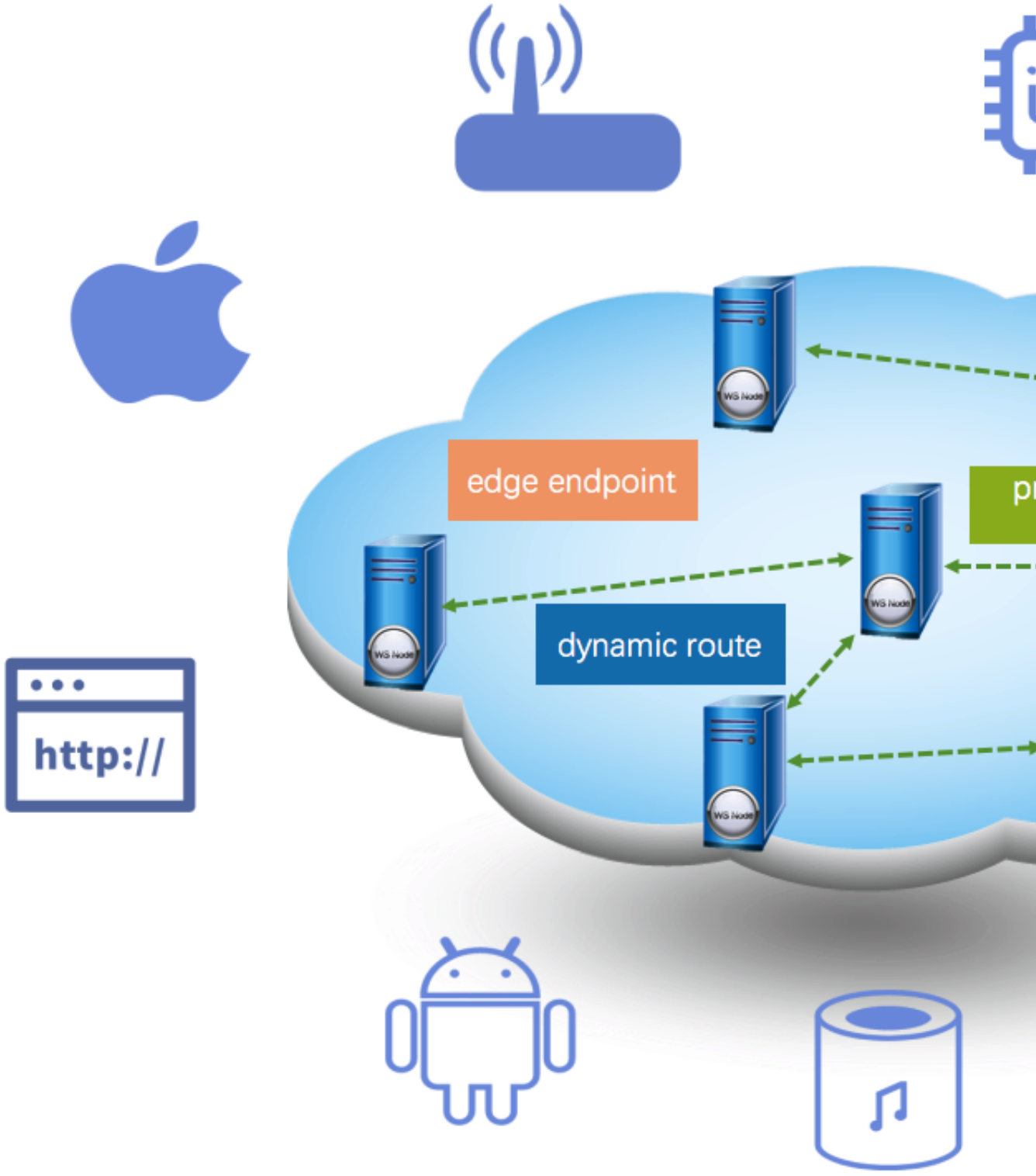
Log Service adds a network type of **Global Acceleration Public Network** on the basis of Virtual Private Cloud (VPC) and public network. Compared with the ordinary public network access, Global Acceleration Public Network has significant advantages in terms of delay and stability, and is suitable for scenarios with high demands for data collection, low consumption delay, and reliability. Global Acceleration for Log Service depends on the acceleration environment provided by Alibaba Cloud *Dynamic Route for CDN* products. This function improves overall site performance and user experience by solving problems of slow response, packet loss, and unstable services. These problems are caused by factors such as cross-carriers access, network instability, traffic spikes, and network congestion.

Global Acceleration for Log Service is based on Alibaba Cloud Content Delivery Network (CDN) hardware resources, and optimizes the stability of log collection and data transmission from various forms of data sources such as mobile phones, Internet of Things (IoT) devices, smart devices, self-built Internet Data Centers (IDCs), and other cloud servers.



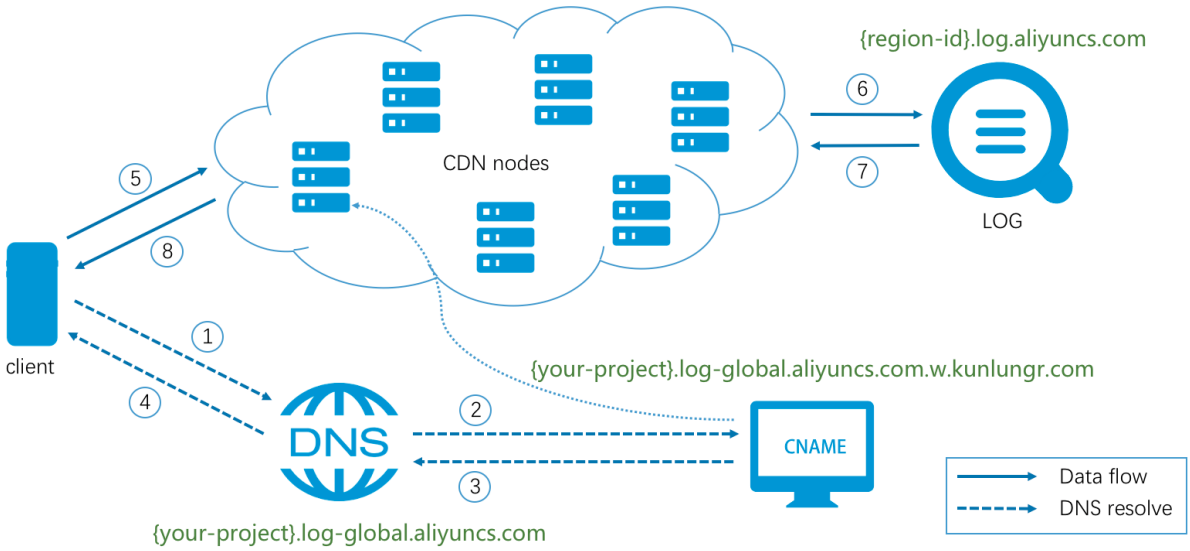
Technical principles

Global Acceleration for Log Service is based on Alibaba Cloud CDN hardware resources. Your global access terminals (such as mobile phones, IOT devices, smart devices, self-built IDCs, and other cloud servers), access the nearest edge node of Alibaba Cloud CDN all over the world and route to Log Service through CDN inner high-speed channels. Compared with common public network transmission, network delay and jitter can be reduced greatly in this method.



The processing flow of Global Acceleration requests for Log Service is shown in the preceding figure. The overall flow is detailed as follows:

1. The client needs to send a domain name resolution request to the public DNS before sending requests of log upload or log download to the Log Service acceleration domain name *your-project*.log-global.aliyuncs.com.
2. The domain name at the public DNS *your-project*.log-global.aliyuncs.com points to the CNAME address *your-project*.log-global.aliyuncs.com.w.kunlungr.com. The domain name resolution is forwarded to the CNAME nodes of Alibaba Cloud CDN.
3. Based on Alibaba Cloud CDN smart scheduling system, CNAME nodes return the IP address of the optimal CDN edge node to the public DNS.
4. The public DNS returns the IP address finally resolved to the client.
5. The client sends a request to the server based on the obtained IP address.
6. After receiving the request, the CDN edge node routes the request to the node closest to the Log Service server based on the dynamic route lookup and private transport protocol. Finally, the request is forwarded to Log Service.
7. After receiving the request from the CDN node, the server of Log Service returns the result of the request to the CDN node.
8. CDN transparently transmits the result or data returned by Log Service to the client.



Billing method

Global Acceleration costs for Log Service include:

- Costs for accessing Log Service

Costs for accessing Log Service is the same as that in common public network. Log Service supports **Pay-As-You-Go** billing method, and provides **FreeTier quota**. For more information, see *Billing method*.

- Service costs for Dynamic Route for CDN

  For information about cloud product costs of Dynamic Route for CDN, see *Billing Method of Dynamic Route for CDN*.

Scenarios

- Advertisement

  Log data about advertising browsing and clicking are extremely important for advertising billing . Advertising carriers include mobile terminal embedding, H5 pages, PC ends, and more all over the world. In some remote areas, the public network data transmission is less stable and risks of log loss exist. A more stable and reliable log upload channel can be obtained through Global Acceleration.

- Online game

  The online game industry has high requirements on the performance and stability of data collection in the official website, logon service, sales service, game service, and other services . The timeliness and stability of data collection are hard to be guaranteed in the case of mobile game data collection and data back transmission from globalized games. We recommend that you use Global Acceleration for Log Service to solve the preceding issues.

- Finance

  Financial-related applications require high availability and high security for network. Audit logs of each transaction and each user action must be collected securely and reliably to the server. At present, mobile transactions have become mainstream. For example, online banking, credit card malls, mobile securities, and other types of transactions can achieve secure, fast, and stable log collection by using HTTPS Global Acceleration for Log Service.

- Internet of Things

  IoT devices and smart devices (for example, smart speakers and smart watches) collect sensor data, operation logs, critical system logs, and other data to the server for data analysis. These devices are usually distributed all over the world and the surrounding network is not always reliably. To achieve stable and reliable log collection, we recommend using Global Acceleration for Log Service.

Acceleration effect

Region	Delay ms (common public network)	Delay ms (acceleration)	Time-out ratio % (common public network)	Time-out ratio % (acceleration)
Hangzhou	152.881	128.501	0.0	0.0
Europe	1750.738	614.227	0.5908	0.0
USA	736.614	458.340	0.0010	0.0
Singapore	567.287	277.897	0.0024	0.0
Middle East	2849.070	444.523	1.0168	0.0
Australia	1491.864	538.403	0.014	0.0

The test environment is as follows:

- Region of Log service: North China 5 (Hohhot)
- Average upload packet size: 10KB
- Test time range: one day (average)
- Request type: HTTPS
- Request server: Alibaba Cloud ECS (Specification 1C1GB)

Note:
The acceleration effect is for reference only.

1.4.2 Enable Global Acceleration

To enable Global Acceleration for Log Service, see the following steps.

Prerequisite

- You have enabled Log Service and created the project and Logstore.
- You have enabled *Dynamic Route for CDN*.
- To *Enable HTTPS acceleration*, *Enable HTTP acceleration* first.

Configuration

After HTTP Global Acceleration is enabled for the project, you can also configure Global Acceleration of Logtail, SDK, and other methods according to your needs.

1. *Enable HTTP acceleration*.
2. Enable Global Acceleration of Logtail, SDK, and other methods.

- HTTPS

  If you use HTTPS to access Log Service, make sure that HTTPS acceleration is enabled. To configure HTTPS acceleration, see *Enable HTTPS acceleration*.

- Logtail log collection

  When you install Logtail, select the **Global Acceleration** network type at the page prompt. Then you can obtain global acceleration when you collect logs by using Logtail.

- SDK, Producer, and Consumer

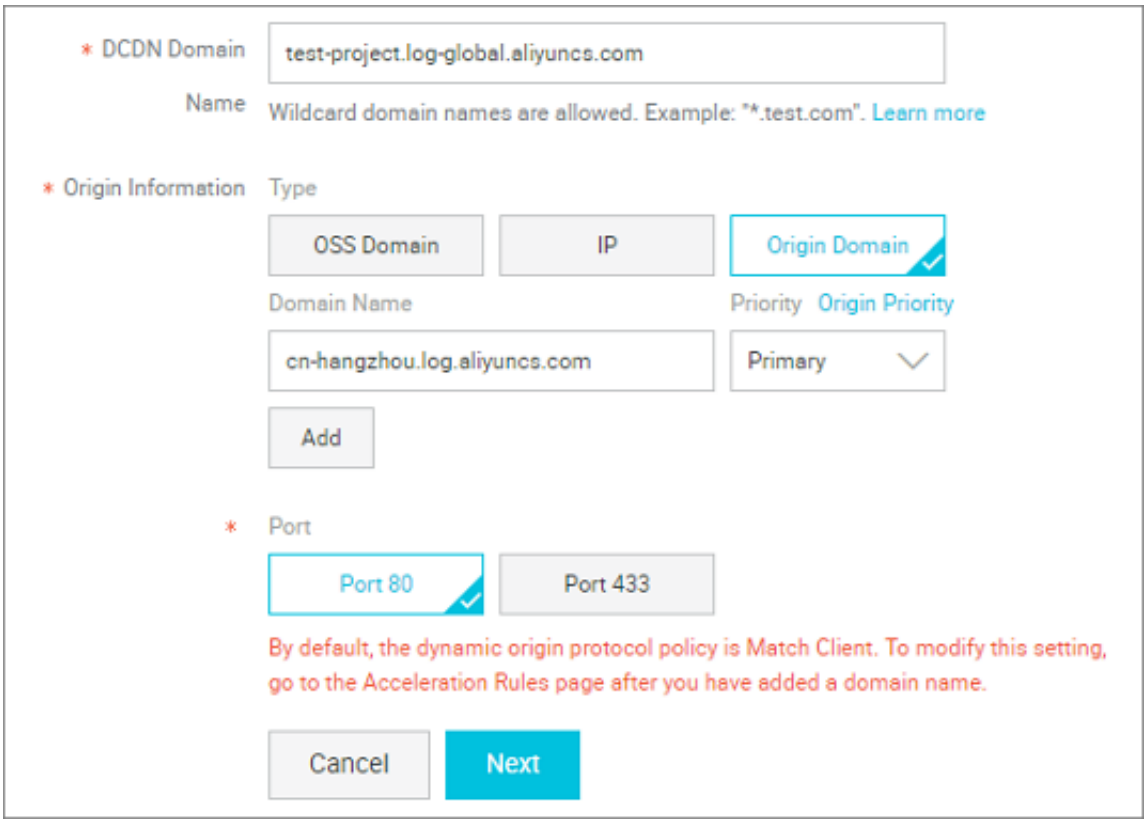
  Other ways to access Log Service such as SDK, Producer, and Consumer, can be accelerated by replacing the endpoint with `log-global.aliyuncs.com`.

Enable HTTP acceleration

1. Log on to the *Dynamic Route for CDN Console*. Click **Domain Names** in the left-side navigation pane to enter the **Domain Names** page.
2. Click **Add Domain Name** in the upper left corner to enter the **Add Domain Name** page.
3. Enter the **`DCDN Domain`** and other information, and click **Next**.

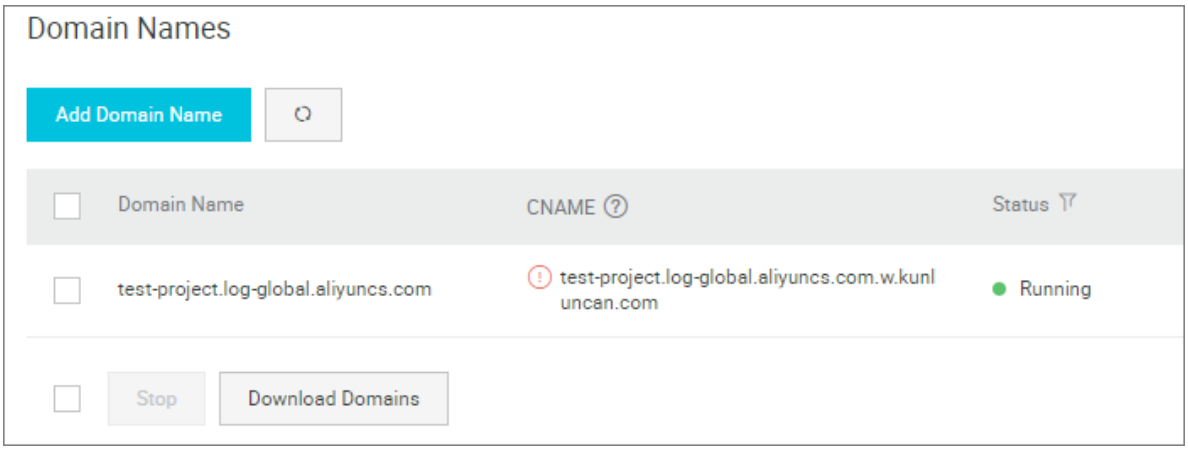
Configuration	Description
Accelerated domain name	*`project_name`*`.log-global.aliyuncs.com` Replace *`project_name`* with your project name.
Origin site type	Select `Origin Domain`.
Domain name	Enter the public network endpoint for the region to which your project belongs. For information about endpoints, see *Service endpoint*.
Port	Please select port 80. If you have an HTTPS acceleration requirement, see **Enable HTTPS acceleration**.
Accelerated area	By default, this configuration item is not displayed and the acceleration area is `Domesticate acceleration`. If you have a demand for Global Acceleration, open a ticket for the Dynamic Route for CDN product to apply for a whitelist . After your application is approved, you can select an acceleration region based on your needs.

For more information about adding domain names, see 8.

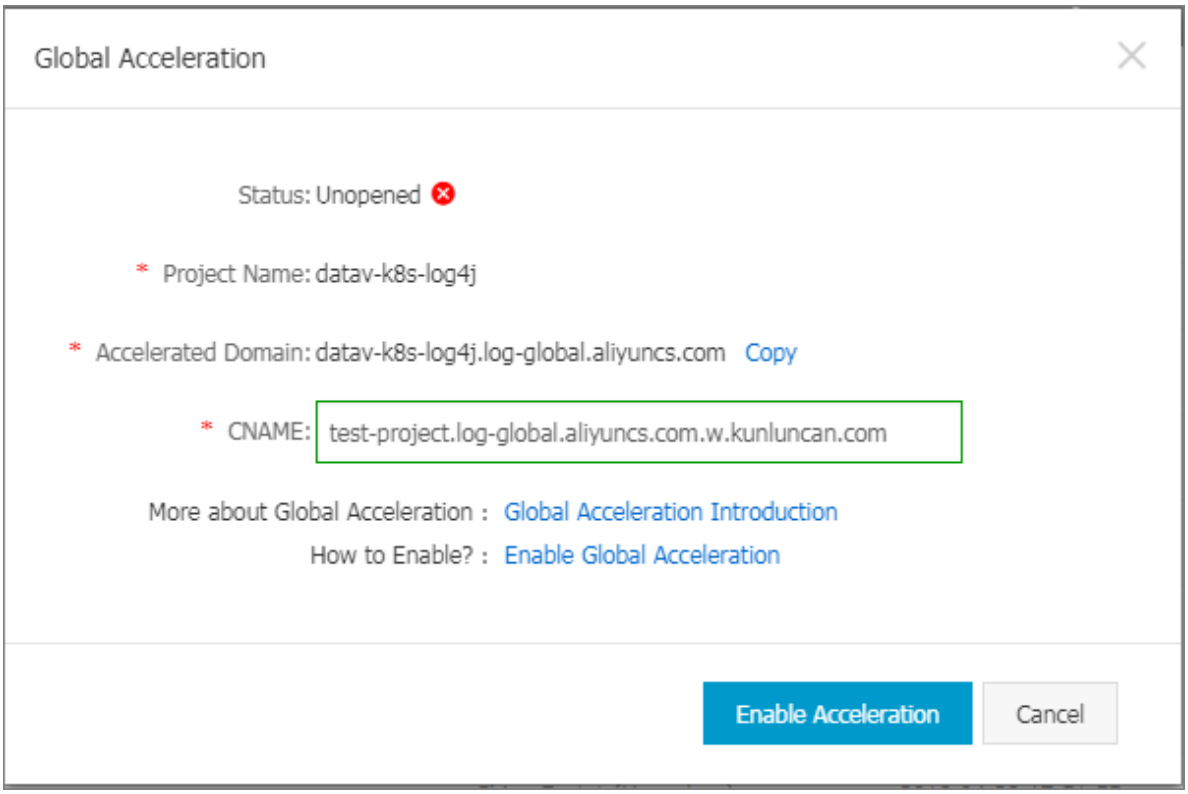


4. Go to the **Domain management** page as prompted.

   You can view the **CNAME** of each corresponding domain name in the **Domain name management** page.



5. Log on to the Log Service console and click **Global Acceleration** at the right of a specified project in the **Project list**.
6. Enter the **CNAME** corresponding to the accelerated domain name in the dialog box. Click **Enable acceleration**.

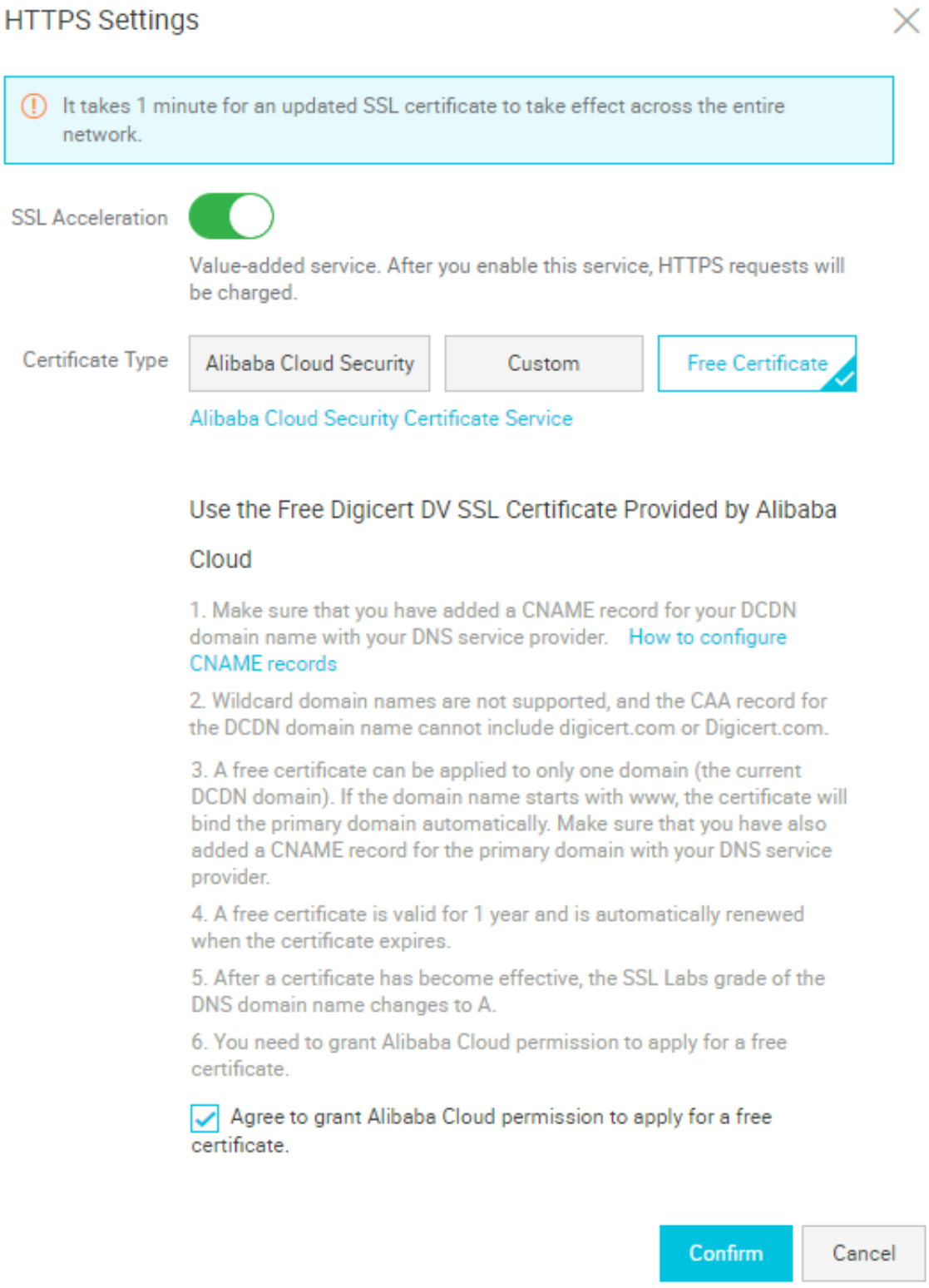


After you complete the preceding steps, Global Acceleration for Log Service is enabled.

Enable HTTPS acceleration

After enabling HTTP acceleration, if you have HTTPS access requirements, you can use the following steps to enable HTTPS acceleration.

1. Log on to the *Dynamic Route for CDN Console*. Click **Domain Names** in the left-side navigation pane to enter the **Domain Names** page.
2. Click **Configure** to the right of a specified domain name.
3. Click **HTTPS Settings** in the left-side navigation pane and click **Modify** in the column of **SSL Certificate** to enter the **HTTPS Settings** page.
4. Configure **`SSL Acceleration`** and **`Certificate Type`**.
   - Enable **`SSL Acceleration`**.
   - Select `Free Certificate` for **`Certificate Type`**.



After the configuration is completed, select **Agree to grant Alibaba Cloud permission to apply for a free certificate.**, and click **Confirm**.

For more information about HTTPS settings, see *HTTPS##*.

Verify if the acceleration configuration takes effect

FAQ

- **How to verify if the acceleration configuration takes effect** ?

  After the configuration is completed, you can verify if the acceleration takes effect by accessing your accelerated domain name.

  For example, if Global Acceleration is enabled for the *test-project* project, you can use `curl` to send a request to the accelerated domain name. If the following type of output is returned, the acceleration takes effect.

  ```
  $curl test-project.log-global.aliyuncs.com
  {"Error":{"Code":"OLSInvalidMethod","Message":"The script name is
  invalid : /","RequestId":"5B55386A2CE41D1F4FBCF7E7"}}
  ```

  For more information about checking methods, see *How to verify if the acceleration takes effect* .

- **How to handle the error of** `project not exist` reported in accessing accelerated domain name?

  This problem is caused usually by an invalid source site address. Log on to the Dynamic Route for CDN console and change the source site address to the public network address of the region to which your project belongs. For information about address list, see *Service endpoint*.

  Note:
  Changing the source site address has a synchronization delay of several minutes.

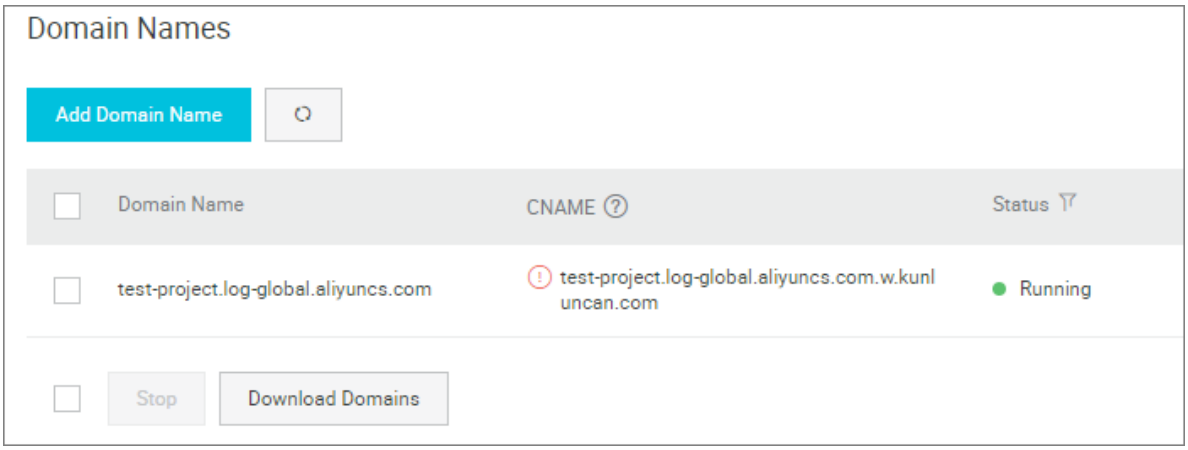
1.4.3 Disable Global Acceleration

To disable Global Acceleration for Log Service, perform the following operations.

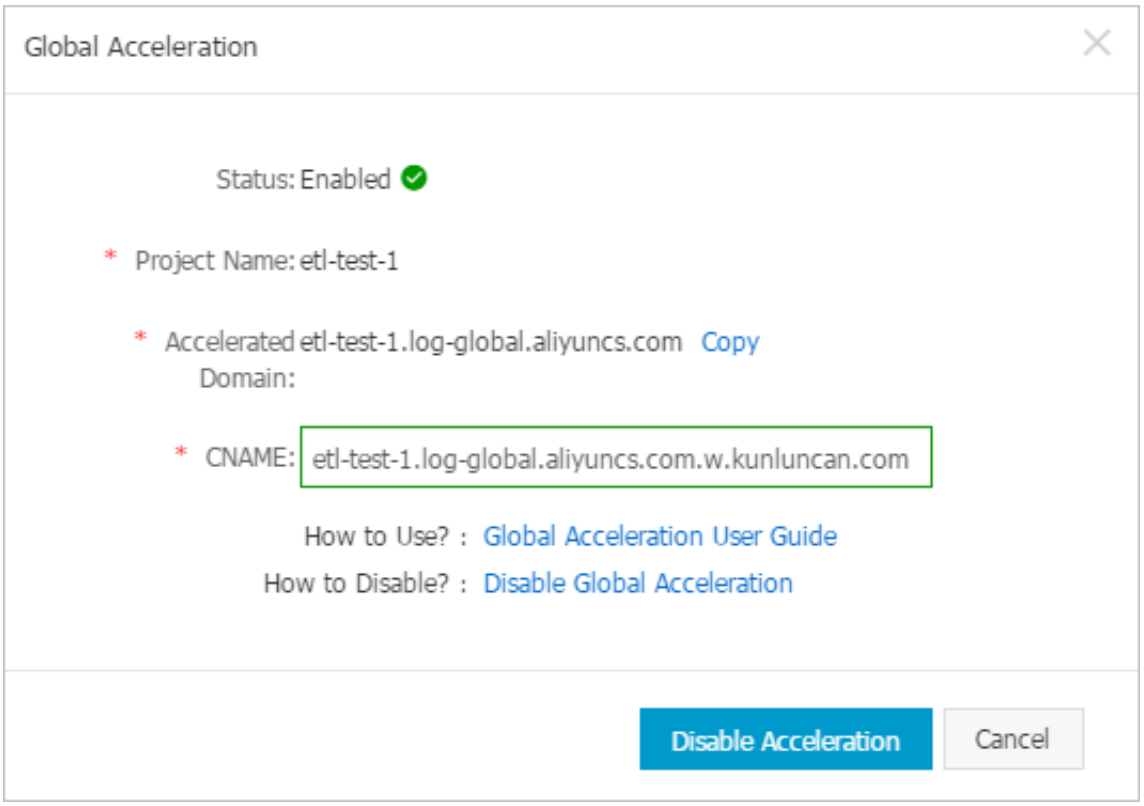
Note:
When you disable Global Acceleration, the accelerated domain name configured during provisioning becomes unavailable. Make sure that all of your clients do not upload or request data through the domain name before you disable Global Acceleration.

Disable Global Acceleration

1. Log on to the *Dynamic Route for CDN Console*. Click **Domain name management** in the left-side navigation pane to enter the **Domain name management** page.
2. View the **CNAME** corresponding to the domain name that is to be disabled .



3. Log on to the Log Service console. On the **Project list** page, click **Global Acceleration** at the right of a specified project.
4. Enter **CNAME** and click **Disable acceleration**.



2 Logtail collection

2.1 Limits

Table 2-1: Limits on file collection

Item	Capabilities and limits
File encoding	Log files encoded in UTF-8 and GBK are supported. Log files encoded in other formats result in undefined behaviors such as gibberish and data loss. We recommend that you use UTF-8 encoding for better processing performance.
Log file size	Unlimited.
Log file rotation	Both `.log*` and `.log` are supported.
Log collection behavior upon log parsing block	When block occurs in log parsing, Logtail keeps the open status of the log file FD. If log file rotation occurs multiple times during the block, Logtail attempts to keep the log parsing sequence of each rotation. If the number of unparsed log rotations is more than 20, Logtail does not process subsequent log files. Soft link support More information, see here.
Single log size	Monitored directories can be soft links.
Single log size	The size of a single log cannot exceed 512 KB . If multiple-line logs are divided by a regular expression, the maximum size of each log is still 512 KB. If the log size exceeds 512 KB , the log is forced to be divided into multiple parts for collection. For example, a log is 1025 KB. The first 512 KB is processed for the first time, the subsequent 512 KB is processed for the second time, and the last 1 KB is processed for the third time.
Regular expression type	Use regular expressions that are compatible with Perl.
Multiple collection configurations for the same file	Not supported. We recommend that you collect log files to a Logstore and configure multiple subscriptions. If this function is required,

Item	Capabilities and limits
	configure a soft link for the log file to bypass this limit.
File opening behavior	Logtail keeps a file to be collected in the open status. Logtail closes the file if the file does not have any modification within five minutes.
First log collection behavior	Logtail only collects incremental log files. If modifications are found in a file for the first time and the file size exceeds 1 MB, Logtail collects the logs from the last 1 MB. Otherwise, Logtail collects logs from the beginning. If a log file is not modified after the configuration is issued, Logtail does not collect this file.
Non-standard text log	For a row containing '\0' in the log. The log is truncated to the first '\0'.

Table 2-2: Checkpoint management

Item	Capabilities and limits
Checkpoint timeout period	If the file has not been modified for more than 30 days, the Checkpoint is deleted.
Checkpoint storage policy	Regular save every 15 minutes, automatically saved when the program exits.
Checkpoint save path	The default save path is `/tmp/logtail_checkpoint` , you can modify the parameters according to *Configure startup parameters*.

Table 2-3: Limits on configuration

Item	Capabilities and limits
Configuration update	Your updated configuration takes effect with a delay of about 30 seconds.
Dynamic configuration loading	Supported. The configuration update does not affect other collections.
Number of configurations	Theoretically unlimited. We recommend that the number of collection configurations for a server is no more than 100.
Multi-tenant isolation	The isolation between collection configurations.

Table 2-4: Limits on resources and performance

Item	Capabilities and limits
Log processing throughput	The default limit to raw log traffic is 2 MB/s. Data is uploaded after being encoded and compressed, generally with a compression ratio of 5–10 times. Logs may be lost if the log traffic exceeds the limit. To adjust the parameter, see *Configure startup parameters* Configure startup parameters.
Maximum performance	In case of single core, the maximum processing capability is 100 MB /s for logs in simple mode, 20 MB/s by default for logs in full mode (depending on the complexity of the regular expression), 40 MB/s for logs in delimiter mode, and 30 MB/s for logs in JSON mode. Enabling multiple log processing threads improves the performance by 1.5–3 times.
Number of monitored directories	Logtail actively limits the depth of monitored directories to conserve your resources. If the upper limit is reached, Logtail stops monitoring more directories and log files. Logtail monitors at most 3,000 directories (including subdirectories).
Default resource limit	By default, Logtail occupies up to 40% of CPU usage and 256 MB of memory usage. If logs are generated at a high speed, you can adjust the parameter by following the *Configure startup parameters* Configure startup parameters.
Processing policy for resource limit exceeding	If the resources occupied by Logtail in 3 minutes exceed the upper limit, Logtail is forced to restart, which may cause loss or duplication of data.

Table 2-5: Limits on error handling

Item	Capabilities and limits
Network error handling	If the network connection is abnormal, Logtail actively retries and automatically adjusts the retry interval.
Handling of resource quota exceeding	If the data transmission rate exceeds the maximum quota of Logstore, Logtail blocks log collection and automatically retries.
Maximum retry period for timeout	If data transmission fails for more than 6 successive hours, Logtail discards the data.
Status self-check	Logtail automatically restarts in the case of an exception, for example, abnormal exit of a program or resource limit exceeding.

Table 2-6: Other limits

Item	Capabilities and limits
Log collection delay	Except for block status, the delay in log collection by Logtail does not exceed one second after logs are flushed to a disk.
Log uploading policy	Logtail automatically aggregates logs in the same file before uploading them. Log uploading is triggered in the condition that more than 2,000 logs are generated, the log file exceeds 2 MB, or the log collection exceeds 3 seconds.

3 Index and query

3.1 Overview

Log Service provides the LogSearch/Analytics function to query and analyze large amounts of logs in real time. You can use this function by enabling the index and field statistics.

Functional advantages

- Real-time: Logs can be analyzed immediately after they are written.
- Fast:
  - Query: Billions of data can be processed and queried within one second (with five conditions).
  - Analysis: Hundreds of millions of data can be aggregated and analyzed within one second (with aggregation by five dimensions and the GroupBy condition).
- Flexible: Query and analysis conditions can be changed as required to obtain results in real time.
- Ecologic: Besides functions such as reports, dashboards, and quick analysis provided in the console, Log Service seamlessly interconnects with products such as Grafana, DataV, and Jaeger, and supports protocols such as RESTful API and JDBC.

Basic concepts

Without enabling the LogSearch/Analytics (index) function, raw data is consumed according to the sequence in the shard, which is similar to Kafka. With the LogSearch/Analytics (index) function enabled, besides the consumption in sequence, you can also count and query the logs. For the difference between log consumption and log query, see Differences between log consumption and log query.

Enable an index

1. Log on to the Log Service console. On the Project List page, click the project name.
2. Select the Logstore, and click **Search**. Then, click **Enable Index** in the upper-right corner. If you have enabled the index before, click **Index Attributes** > **Modify.**.
   - After enabling the query and statistics, data is indexed in the backend. Traffic and storage space for the index are required.
   - ◦If this function is not required, click **Disable** to disable it.

3. Enter the Settings menu to complete configuration.

Data types

You can configure the type of each key in a log (full text index is a special key, whose value is the log). Currently, Log Service supports the following data types.

Category	Type	Description	Query example
Basic	*TEXT*	The text type that supports keyword and fuzzy match.	`uri:"login*" method:"post"`
Basic	*Long*	The value type that supports interval query.	`status>200, status in [200, 500]`
Basic	*Double*	The value type with a float.	`price>28.95, t in [20.0, 37]`
Combination	*JSON*	The content is a JSON field, which is of the text type by default and supports the nested model. You can configure indexes of text, long, and double type for element b under a by using the path format such as a.b . The field type after the configuration is subject to the configuration.	`level0.key>29.95 level0.key2:"action"`
Combination	*Full text*	Use a log as the text for query.	`error and "login fail"`

Query and analysis syntax

Real-time query and analysis is composed of Search and Analytics, which are separated with a vertical line (|):

`$Search |$Analytics`

- •Search: The query condition, which is generated by using keywords, fuzzy match conditions, values, ranges, and combination conditions. If Search is empty or an asterisk (*), all data is queried.
- Analytics: Calculate and count the query results or the full data.

 **Note:**

Both Search and Analytics are optional. If Search is empty, all the data in the specified period is not filtered and the results are counted directly. If Analytics is empty, the query results are returned and no statistics are collected.

Note:

For more information, see *Query syntax*, *Syntax description*.

Query examples

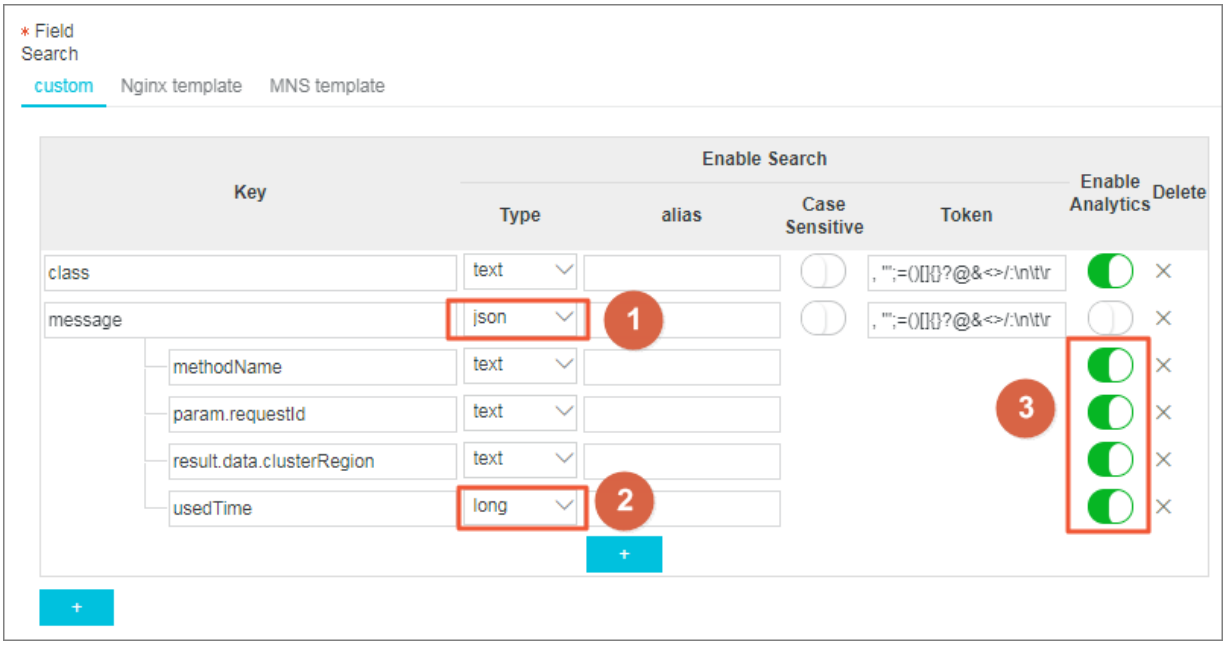
Besides time, the following log also contains four key values.

Sequence number	Key	Type
0	time	-
1	class	text
2	status	Long
3	Latency	double
4	message	json

```
0. time:2018-01-01 12:00:00
  1. class:central-log
  2. status:200
  3. latency:68.75
  4. message:

      "methodName": "getProjectInfo",
      "success": true,
      "remoteAddress": "1.1.1.1:11111",
      "usedTime": 48,
      "param": {
              "projectName": "ali-log-test-project",
              "requestId": "d3f0c96a-51b0-4166-a850-f4175dde7323"

      "result": {
          "message": "successful",
          "code": "200",
          "data": {
              "clusterRegion": "ap-southeast-1",
              "ProjectName": "ali-log-test-project",
              "CreateTime": "2017-06-08 20:22:41"

          "success": true
      }
```

Configuration is as follows:

Figure 3-1: Index settings



Where:

- ① indicates that all the data of the string type and bool type in the JSON field can be queried.
- ② indicates that data of the long type can be queried.
- ③ indicates that you can analyze the configured field by using SQL statements.

Example 1: Query string, bool type

```
class : cental*
message.traceInfo.requestId : 92.137_1518139699935_5599
message.param.projectName : ali-log-test-project
message.success : true
```

Note:

- No configurations in the JSON field are needed.
- JSON Map and Array are auto scaling and support multi-level nesting. Each layer is separated with a period (.).

Example 2: Query double、long type

```
latency>40
message.usedTime > 40
```

Note:

You configure JSON fields independently. The fields must not be in array.

Example 3: Combined query

```
class : cental* and message.usedTime > 40 not message.param.projectNam
e:ali-log-test-project
```

Other information

If you query a large amount of log data (such as a long query time span, where the data volume is over 10 billion), one request cannot query all the data. In this case, Log Service returns the existing data and notifies you that the query result is incomplete.

At the same time, the server caches the results of the query within 15 minutes. When the query result is partially cached, the server continues to scan log data that has not been cached. To reduce the workload of merging multiple query results, Log Service merges the result of the cache hit with the result of the new query and returns it to you.

Therefore, Log Service enables you to get the final result by calling the interface repeatedly with the same parameters.

3.2 Data type of index

3.2.1 Text type

Similar to search engines, text data is queried based on terms. Therefore, you must configure word segmentation, case sensitivity, including options.

Instructions

Case sensitivity

Determine whether to support case sensitivity when querying raw logs. For example, the raw log is internalError .

- After turning off the Case Sensitive switch, the sample log can be queried based on the keyword INTERNALERROR or internalerror .
- After turning on the Case Sensitive switch, the sample log can only be queried based on the keyword internalError .

Token

You can separate the contents of a raw log into several keywords by using a token.

For example, the raw log is

```
/url/pic/abc.gif
```

- If no token is set, the string is considered as an individual word `/url/pic/abc.gif`. You can only query this log by using the complete string or fuzzy match such as `/url/pic/*` .
- If `/` is set as the token, the raw log is separated into three words: `url` , `pic` , and `abc.gif` . You can query this log by using any of the three words or fuzzy match, for example, `url` , `abc.gif` , or `pi*` . You can also use `/url/pic/abc.gif` to query this log (`url and pic and abc.gif` is separated into the following three conditions during the query: url , pic , and abc.gif).
- If `/.` is set as the token, the raw log is separated into four words: `url` , `pic` , `abc` , and `gif` .

Note:
You can broaden the query range by setting appropriate tokens.

Full text index

By default, full text query (index) considers all the fields and keys of a log, except the time field, as text data, and does not need to specify keys. For example, the following log is composed of four fields (time/status/level/message):

```
[20180102 12:00:00] 200,error,some thing is error in this field
```

- time:2018-01-02 12:00:00
- level:”error”
- status:200
- message:”some thing is error in this field”

After enabling full text index, the following text data is assembled in the “key:value + space” mode.

```
status:200 level:error message:"some thing is error in this field"
```

Note:

- Prefix is not required for full text query. Enter error as the keyword, both level field and message field meet the query condition.
- You must set a token for the full text query. If a space is set as the token, status:200 is considered as a phrase. If : is set as the token, status and 200 are considered as two independent phrases.

- Numbers are processed as texts. For example, you can use the keyword 200 to query this log . The time field is not processed as a text.
- You can query this log if you enter a key such as ”status" .

3.2.2 Value type

When configuring indexes, you can configure a field as the value type and query the key by using a value range.

Instructions

Supported types: `long` (long integer) and `double` (decimal). After configuring a field as the value type, you can only query the key by using a value range.

Example

To query the longkey whose key range is (1000 2000], use the following methods.

- Use values to query the longkey:

```
longKey > 1000 and longKey <= 2000
```

- Use an interval to query the longkey:

```
longKey in (1000 2000]
```

For more syntaxes, see *Query syntax*.

3.3 Query syntax

To help you query logs more effectively, Log Service provides a set of query syntax to express query conditions. You can specify query conditions by using the *GetLogs* and *GetHistograms* interfaces in Log Service API or on the query page of the Log Service console. This document introduces the syntax of query conditions in details.

Index types

Log Service supports creating an index for the LogStore in the following methods:

- Full text index: Query the entire line of logs as a whole without differentiating key and value.
- Key/value index: Query logs after specifying a key. For example, *`FILE:app`* and *`Type: action`*. All the strings with the specified key are queried.

Syntax keywords

LogSearch query conditions support the following keywords.

Name	Meaning
and	Binary operator. Format: query1 and query2. Indicates the intersection of the query results of query1 and query2. With no syntax keyword among multiple words, the relation is and by default.
or	Binary operator. Format:`query1 or query2`. Indicates the union of the query results of `query1` and `query2` .
not	Binary operator. Format: `query1 not query2`. Indicates a result that matches `query1` and does not match `query2`, which is equivalent to `query1-query2`. If only `not query1` exists, it indicates to select the results excluding `query1` from all the logs.
(,)	Parentheses () are used to merge one or more sub-queries into one query to increase the priority of the query in the parentheses ().
:	Used to query the key-value pairs. `term1:term2` makes up a key-value pair. If the key or value contains reserved characters such as spaces and colons (:), use quotation marks (“) to enclose the entire key or value.
“	Converts a keyword to a common query character. Each term enclosed in quotation marks (“) can be queried and is not be considered as a syntax keyword. Or all the terms enclosed in quotation marks (“) are regarded as a whole in the key-value query.
\	Escape character. Used to escape quotation marks. The escaped quotation marks indicate the symbols themselves, and they cannot be used as escape characters, such as `"\""`.
|	The pipeline operator indicates more calculations based on the previous calculation, such as query1 | timeslice 1h | count.
timeslice	The time-slice operator indicates how long the data is calculated as a whole. Timeslice 1h, 1m, 1s indicates 1 hour, 1 minute, and 1 second respectively. For example, query1 | timeslice 1h | count represents the query query condition, and returns to the total number of hours divided by 1 hour.
count	The count operator indicates the number of log lines.
*	Fuzzy query keyword. Used to replace zero or multiple characters. For example, the query results of `que*` start with `que`.  **Note:**

Name	Meaning
	At most 100 query results can be returned.
?	Fuzzy query keyword. Used to replace one character. For example, the query results of `qu? ry` start with `qu`, end with `ry`, and have a character in the middle.
`__topic__`	Topic data query. With the new syntax, you can query the data of zero or multiple topics in the query. For example, `__topic__: mytopicname`.
`__tag__`	Query a tag value in a tag key. For example, `__tag__:tagkey: tagvalue`.
Source	Query the data of an IP. For example, `source:127.0.0.1`.
>	Query the logs with a field value greater than a specific number. For example, `latency > 100`.
>=	Query the logs with a field value greater than or equal to a specific number. For example, `latency >= 100`.
<	Query the logs with a field value less than a specific number. For example, `latency < 100`.
<=	Query the logs with a field value less than or equal to a specific number. For example, `latency <= 100`.
=	Query the logs with a field value equal to a specific number. For example, `latency = 100`.
in	Query the logs with a field staying within a specific range. Braces ([]) are used to indicate closed intervals and parentheses (()) are used to indicate open intervals. Enclose two numbers in braces ([]) or parentheses (()) and separate the numbers with several spaces. For example, `latency in [100 200]` or `latency in (100 200]]`.

Note:

- Syntax keywords are case-insensitive.
- Priorities of syntax keywords are sorted in descending order as follows: `: > " > ( ) > and not > or`.
- Log Service reserves the right to use the following keywords: `sort asc desc group by avg sum min max limit`. To use these keywords, enclose them in quotation marks (“).

- If both the full text index and key/value index are configured and have different word segmentation characters, data cannot be queried using the full text query method.
- Set the column type as double or long before performing a numeric query. If the column type is not set or the syntax used for the numeric range query is incorrect, Log Service translates the query condition as a full text index, which may lead to an unexpected result.
- If you change the column type from text to numeric, only the = query is supported for the data before this change.

Query examples

1. Logs that contain a and b at the same time: `a and b` or `a b`.
2. Logs that contain a or b: `a or b`.
3. Logs that contain a but do not contain b: `a not b`.
4. All the logs that do not contain a: `not a`.
5. Query the logs that contain a and b, but do not contain c: `a and b not c`.
6. Logs that contain a or b and must contain c: `(a or b ) and c`.
7. Logs that contain a or b, but do not contain c: `(a or b ) not c`.
8. Logs that contain a and b and may contain c: `a and b or c`.
9. Logs whose FILE field contains apsara: `FILE:apsara`.
10. Logs whose FILE field contains apsara and shennong: `FILE:"apsara shennong"`, `FILE: apsara FILE: shennong` or `FILE:apsara and FILE:shennong.`
11. Logs containing and: `and`.
12. Logs with the FILE field containing apsara or shennong: `FILE:apsara or FILE:shennong`.
13. Logs with the file info field containing apsara: `"file info":apsara`.
14. Logs that contain quotation marks (“): `\"`.
15. Query all the logs starting with shen: `shen*`.
16. Query all the logs starting with shen in the FILE field: `FILE:shen*`.
17. Query all the logs starting with shen, ending with ong, and having a character in the middle: `shen? ong.`
18. Query the logs starting with shen and aps: `shen* and aps*`.
19. Query the logs starting with shen every 20 minutes: `shen*| timeslice 20m | count` .
20. Query all the data in the topic1 and topic2: `__topic__:topic1 or __topic__ : topic2`.
21. Query all the data of the tagvalue2 in the tagkey1: `__tag__ : tagkey1 : tagvalue2`.

22.Query all the data with a latency greater than or equal to 100 and less than 200: `latency >= 100 and latency < 200` or `latency in [100 200)`.

23.Query all the requests with a latency greater than 100: `latency > 100`.

24.Query the logs that do not contain spider and do not contain opx in http_referer: `not spider not bot not http_referer:opx`.

25.Query logs with the empty cdnIP field: `cdnIP:""`.

26.Query logs without cdnIP field:`not cdnIP:*`.

27.Query logs with the cdnIP field: `cdnIP:*`.

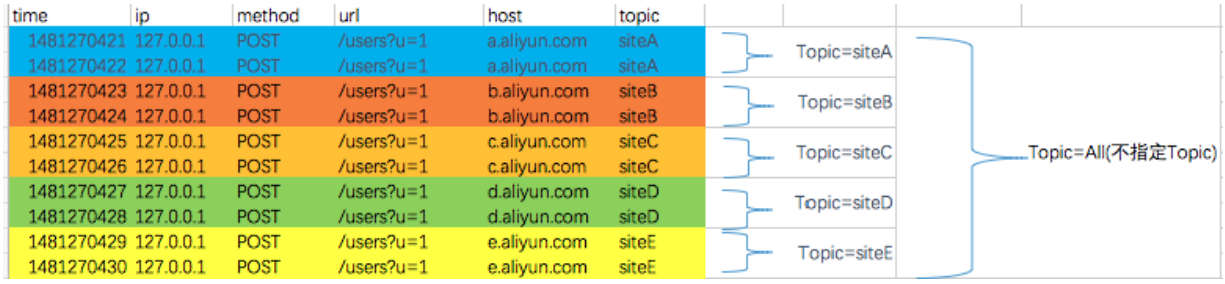
Specified or cross-topic query

Each LogStore can be divided into one or more subspaces by the topic. During therfhfrg query, specifying topics can limit the query range so as to increase the speed. Therefore, we recommend that you use topic to divide the LogStore if you have a secondary classification requirement for the LogStore.

With one or more topics specified, the query is only performed in the topics that meet the conditions. However, if no topic is specified, data of all the topics is queried by default.

For example, use topic to classify logs with the different domain names:

Figure 3-2: Log topic



Topic query syntax:

- Data of all the topics can be queried. If no topic is specified in the query syntax and parameter, data of all the topics is queried.
- Supports query by topic. The query syntax is `__topic__:topicName`. The old mode (specify the topic in the URL parameter) is still supported.
- Multiple topics can be queried. For example, `__topic__:topic1 or __topic__:topic2` indicates the union query of data from Topic1 and Topic2 .

Fuzzy search

Log Service support fuzzy search. Specify a word within 64 characters, and add fuzzy search keywords such as `*` and `?` in the middle or in the end of the word. 100 eligible words will be searched out, in the meantime, all the logs eligible and contain the 100 words will be returned.

Limits：

- Prefix must be specified when query logs, that is, the word can not begin with `*` and `?` .
- Precise the specified word, you will get a more accurate result.
- Fuzzy search cannot be used to search for words that exceeds 64 characters. It is recommended that you specified a word under 64 characters.

3.4 Context query

When you expand a log file, each log records an event. Generally, logs are not independent from each other. Several consecutive logs allow you to view the process of a whole event in sequence.

Log context query specifies the log source (machine + files) and a log in the log source. It also queries several logs before and after the log in the original log file, providing a helpful method for troubleshooting the problem in the DevOps scenario.

The Log Service console provides a query page, you can view the context information of the specified log in the original file in the console. It is similar to paging up and down in the original log file. By viewing the context information of a specified log, you can quickly locate the problem.

Scenarios

For example, the O2O take-out website will record the transaction track of a order in the program log on the server:

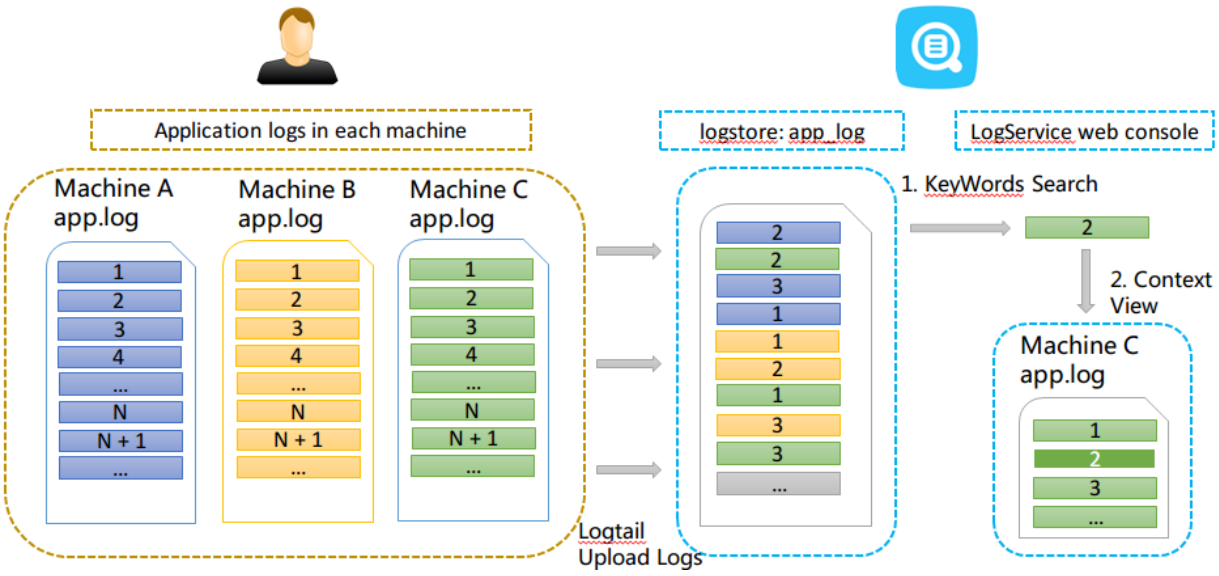
User logon > **Browse products** > **Click items** > **Add to shopping cart** > **Place an order** > **Pay for the order** > **Deduct payment** > **Generate an order**

If the order cannot be placed, the Operation & Maintenance (O&M) personnel must quickly locate the cause of the problem. In the conventional context query, the administrator grants the machine logon permission to related members, and then the investigator logs on to each machine where applications are deployed in turn, uses the order ID as the keyword to search application log files , and determines what causes the failure.

In Log Service, you can troubleshoot the problem by following these steps:

1. Install the log collection client Logtail on the server, and add the machine group and log collection configuration in the console. Then, Logtail starts to upload the incremental logs. You can also use producer-related SDK uploads, such as Log4J, LogBack, C-Producer
2. On the log query page in the Log Service console, specify the time range, and find the order failure log according to the order ID.
3. Based on the found error log, page up until other related logs are found (for example, the deduction failure of credit card).

Figure 3-3: Scenarios



Benefits

- No intrusion into the application. No need to modify the log file format.
- You can view the log context information of any machine or file in the Log Service console, without logging on to each machine to view the log file.
- Combined with the time when the event occurred, you can specify the time range to quickly locate the suspicious log and then query its context information in the Log Service console to improve the efficiency.
- No need to worry about the data loss caused by insufficient server storage space or log file rotation. You can view historical data in the Log Service console at any time.

Prerequisites

- *Use Logtail to collect logs* . Upload data to the Logstore. Create the machine groups and collection configuration. No other configurations are needed. You can also use producer-related SDK upload, such as Producer Library.
- Enable the Query logs function.

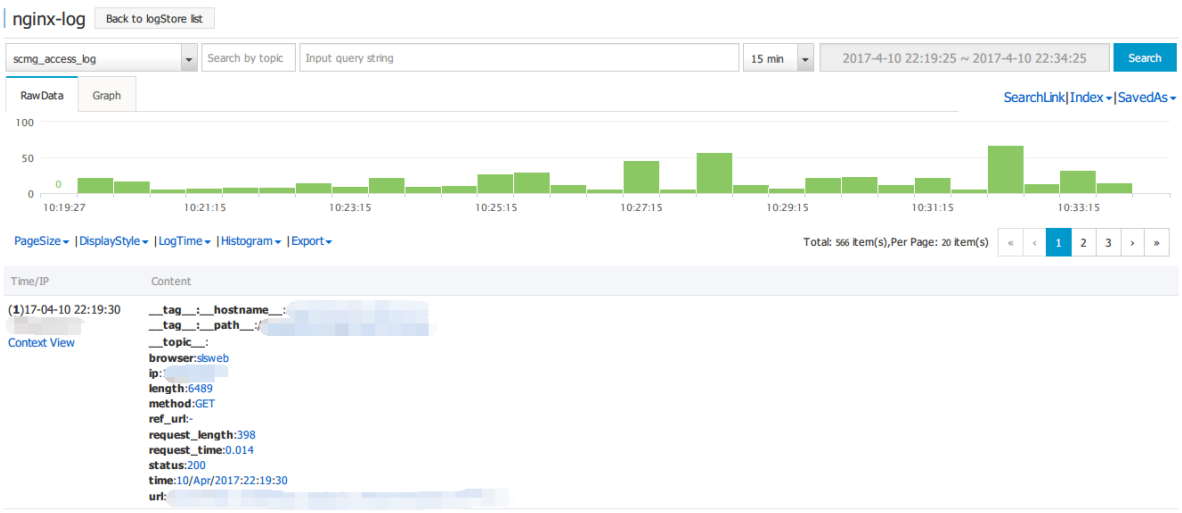
Note:
Currently, you cannot query the context information of syslog data.

Procedure

1. Log on to the Log Service console.
2. On the Project List page, click the project name.
3. On the **Logstore List** page, click **Query** at the right of the Logstore to enter the query interface.
4. Enter your query and analysis statement and select the time range. Then, click **Search**.

   Click **Context View** at the left side of the log, and the window with the context information of the target log is displayed on the right.

   Figure 3-4: Query log



5. Select a log and click **Context View**. View the context log for the target log on the right pop-up page.
6. Scroll with the mouse on the page to view the context information of the selected log. To view more context logs, click **Earlier** or **Later**.

3.5 Other functions

In addition to the statement-based query capability, the query and analysis function of Log Service provides the following extended functions for the query optimization:

- *Raw logs*
- *Graph*
- *Contextual Query*
- *Quick analysis*
- *Quick query*
- *Tag*
- *Dashboard*
- *Save as an alarm*

Raw logs

After the index is enabled, enter the keywords in the search box and select the search time range. Then, click **Search** to view the histogram of the log quantity, the raw logs, and the statistical graph.

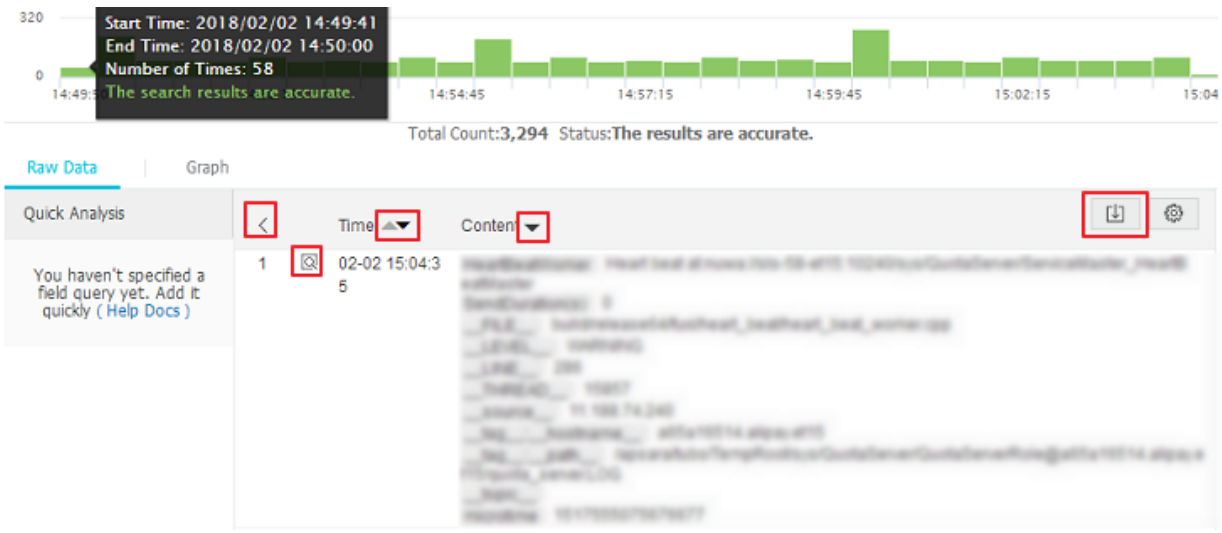
The histogram of the log quantity displays the time-based distribution of log search hit counts . With the histogram, you can view the log quantity changes over a certain period of time. By clicking the rectangular area to narrow down the time range, you can view the information about the log hits within the specified time range to refine the display of the log search results.

On the Raw Data tab, you can view the hit logs in chronological order.

- By clicking the triangle symbol next to **Time**, you can switch between the **chronological** and reverse **chronological** orders.
- By clicking the triangle symbol next to **Content**, you can switch between **Display with Line Breaks** and **Display in One Line**.
- By clicking the value keyword in the log content, you can view all logs containing this keyword.
- By clicking the **Download**button in the upper-right corner of the Raw Data tab, you can download the query results in CSV format. By clicking the **Config**button, you can add fields as displayed columns in the display results of raw logs so that you can view the target field content of each raw log in the new columns in a more intuitive way.
- By clicking **Context**, you can view 15 logs before and after the current log entry. For more information, see *Context query*.

> **Note:**
> Currently, the context query function supports only the data uploaded with Logtail.

Figure 3-5: Raw logs



Graph

After enabling the index and entering a statement for query and analysis, you can view the statistics of logs under the **Graph** tab.

- Data can be displayed in the following ways: tables, line charts, column charts, bar charts, pie charts, numeric values, area charts, and maps.

  You can select an appropriate statistical graph type based on the actual statistical analysis needs.
- You can adjust the display content of axes X and Y to obtain the display results that meet your needs.
- Add the analysis results to **Dashboard**. For more information, see *Dashboard*.

Figure 3-6: Dashboard

Contextual Query

The Log Service console provides a query page, you can view the context information of the specified log in the original file in the console. It is similar to paging up and down in the original log file. By viewing the context information of the specified log, you can quickly locate the failure information during the business troubleshooting. For more information, see *Context query*.

Quick analysis

The quick analysis function of Log Service supports an interactive query with only one click, allowing you to quickly analyze the distribution of a field over a period of time and reduce the cost of indexing key data. For more information, see *Quick analysis*.

Quick query

By clicking **Saved Search** in the upper-right corner of the query page, you can save the current query action as a quick query. To perform this query again, you can quickly complete it on the **Saved Search** tab on the left without manually entering the query statement.

You can also use this quick query condition in alarm rules. If you have added this quick query to **Tag**, you can directly access it in tags.

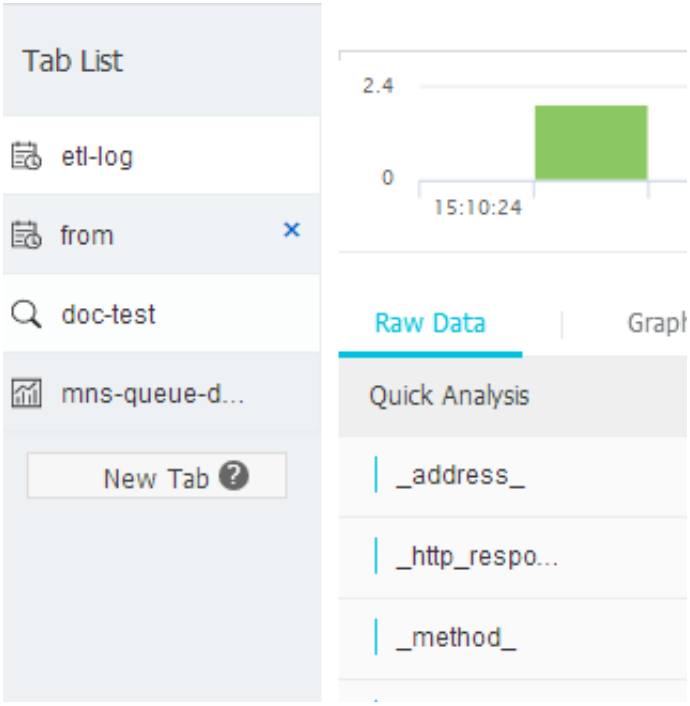
Tag

You can add the following three types of data pages to the tag list on the left of the Log Service query homepage:

- Logstore
- Quick query
- Dashboard

The tag list allows you to open pages easily and quickly. You can directly click to open the Logstore, saved quick queries, and dashboards in the tag list. To add the Logstore, quick query, or dashboard as tags, click **Add Tag** in the tag list and select the Logstore, quick query, or dashboard you want to add in the menu appeared in the right side of the page. To delete a tag, click the remove (X) button at the right of the tag name to be deleted in the tag list.

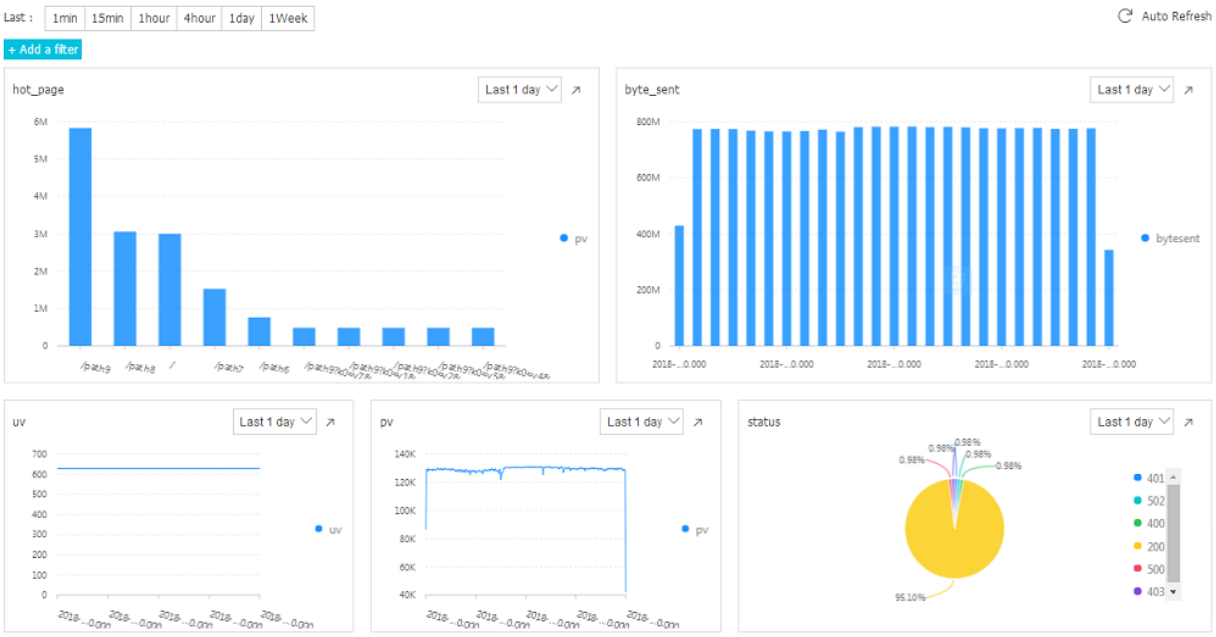
Figure 3-7: Tag



Dashboard

Log Service provides the dashboard function, which can visualize the query and analysis statements. For more information, see *Dashboard*.

Figure 3-8: Dashboard



Save as an alarm

Log Service can generate an alarm based on your **LogSearch Results**. You can configure the alarm rules so that specific alarm content can be sent to you by using in-site notifications or DingTalk messages.

The basic process is as follows:

1. Configure the quick query
2. Configure the alarm rules.
3. Configure notification type.
4. The system sends an SMS/email so that you can view the alarm result.

For more information, see *Set alarms*.