Log Service

FAQ

FAQ

Basic questions

Question list

1. What is Log Service?

Log Service (abbreviated to LOG) is a platform service used to collect, store, and subscribe to logs. Various types of logs can be collected in real time, managed in a centralized way, and consumed by using Log Service.

2. What functions does Log Service provide?

- Provide multiple ways of writing logs (use APIs, SDKs, and Logtail to access Log Service).
- Use Logtail to define the log collection and parsing methods.
- Use machine groups to manage log collection on thousands of machines.
- Consume and subscribe to logs in real time.
- Provide the simple configuration in the console, which allows you to perform all the operations on the Web.
- Seamlessly interconnect with multiple cloud products of Alibaba Cloud in the backend.

3. What are the basic concepts of Log Service?

- Core concepts: Project (the basic unit used to manage logs), Logstore, shard, topic (used for secondary classification of Logstores), log (number of logs), and log group.
- Concepts about log collection: Logtail configuration (used to define how to collect logs) and

machine group (used to manage machines in groups).

4. What are the components of Log Service?

Log Service consists of a log collection client, a server, and other systems. Currently, the client is a log collection agent (Logtail) that is compatible with Windows and Linux. The server is responsible for reading, writing, and configuring Log Service APIs. Other systems include Alibaba Cloud products such as Object Storage Service, that is, Log Service supports synchronizing logs to cloud products such as OSS.

5. How is a log defined in Log Service?

A log contains three parts: time (required), log content (composed of key-value pairs), and metadata (the source IP address of logs).

Manage logs

Question list

1. How does Log Service store and manage your logs?
2. Are logs lost after I delete a Logstore?
3. What is the log storage period of Log Service? Can I modify this period?

1. How does Log Service store and manage your logs?

The Logstore is the basic unit for storing and querying logs in Log Service. Generally, a Logstore is used to store a specific type of logs. Currently, you can add, delete, modify, and query the Logstores in the console or by using APIs. You can write logs to a created Logstore by using APIs/SDKs. To collect logs from Alibaba Cloud Elastic Compute Service (ECS) instances, you can use the Logtail, a log collection service provided by Log Service.

2. Are logs lost after I delete a Logstore?

Yes. So proceed with caution.

3. What is the log storage period of Log Service? Can I

modify this period?

The following three functions of Log Service are related to the log storage period:

- LogHub and LogSearch/Analytics: Configure the log storage period as per your needs.
- LogShipper: After shipping logs to Object Storage Service (OSS), configure the log storage period in OSS.

Log Collection

Basic questions for log collection

What types of logs does the Log Service collect?

The Log Service supports timestamped text logs encoded in UTF-8 that are generated within the past seven days and are no more than 15 minutes later than the current time.

In what ways does the Log Service collect logs? How do I choose among them?

The Log Service supports direct data writing using APIs (SDKs are currently available in four languages: Java, Python, PHP, and C#). It provides Linux- and Windows-compatible Logtail used to collect real-time updated logs from disk files.

1. If the logs generated by application programs are not flushed into disk, those logs can be written directly to the Log Service through APIs.
2. Logs that are written into disks in real time can be collected by Logtail.

How does the Log Service collect logs from ECS?

You can use Logtail to collect the ECS logs that are flushed into disk as follows:

1. Create a Logstore on the Log Service Console.
2. Perform Logtail configuration.
3. Create a machine group.
4. Execute the installation script to install the Logtail agent.

5. Apply the Logtail configurations to the machine group.

Does the Log Service collect historical logs?

You can only write data generated during the past seven days using an API. However, Logtail does not support historical data collection for the moment.

What data collection capability does the Log Service provide? Does it have any limitation?

You can adjust the number of shards in a Logstore as needed. Logtail collects data at a maximum rate of 1 MB/s on ECS.

What should I pay attention to when using Logtail to collect logs on NAS?

For collection of Nginx access logs, the Nginx configurations of web servers are the same. Logs are written into files with the same name on different machines. (In this case, Logtail collects logs properly.) When NAS is used, Logtail may have missing logs or encounter a collection error if the Nginx logs on multiple machines are written into the same file on NAS (concurrent write to the same file).To avoid this problem, ensure that the logs on different web servers are written into different files on NAS.

Basic questions for Logtail

What is Logtail?

Logtail is a log collection agent provided by the Log Service. Once installed on your machine, Logtail monitors specified log files and automatically uploads the new logs written into these files to your designated LogStore.

Does Logtail collect static log files?

Logtail monitors file changes based on change events in the file system and sends logs generated in real time to the Log Service.Logtail does not collect the content of unchanged logs.

What platforms does Logtail support?

Currently, Logtail supports 64-bit Linux and 32/64-bit Windows Server 2003-2012 system.

Linux:

- Alibaba Cloud Linux
- Ubuntu
- Debian
- CentOS
- OpenSUSE

Windows:

- Windows 7 (Client) 32bit
- Windows 7 (Client) 64bit
- Windows Server 2003 32bit
- Windows Server 2003 64bit
- Windows Server 2008 32bit
- Windows Server 2008 64bit
- Windows Server 2012 64bit

How do I install and upgrade the Logtail agent?

Installation: Install the Logtail agent using an installation script. Upgrade: The Log Service regularly upgrades the Logtail agent without interrupting the data collection process.

How do I configure the Logtail agent?

Refer to Logtail collection configuration on the console.

How does Logtail work?

1. On the console, configure the directory you want to monitor, the name of the log file, and the related parsing rule (regular expression).
2. When a log file is changed on your machine, Logtail receives an event from the file system and reads the new log.
3. Logtail parses the log format based on the regular expression and sends the log to the Log Service.

Does Logtail support log rotation?

When the log file a.LOG reaches a given size or lasts for a given period of time since created, a.LOG is renamed a.LOG.1 (or another name). A new a.LOG file is created for writing new logs. This process is called rotation.Logtail automatically rotates logs based on event notifications from the file system.

How does Logtail handle network exceptions?

In the case of a network exception or write quota overrun, Logtail caches collected logs to the local

disk and resends those logs later.The maximum disk cache capacity is 500 MB. Newly cached data overwrites the old one when the 500-MB limit is exceeded. Cached files that fail to be sent to the Log Service within 24 hours are automatically deleted.

What is the log collection delay of Logtail?

Logtail collects logs based on events and sends collected logs to the Log Service within 3s.

How does Logtail process historical logs?

Logtail only collects real-time logs. If the logging time is more than 5 minutes different from the system time at which Logtail processes the log, the log is regarded as a historical log.

How long does a change in log collection configuration take effect for the Log Service?

After you apply configurations to a machine group on the console, Logtail loads and applies the configurations in 3 minutes or less.

How do I locate any log collection problems of Logtail?

1. Check whether the Logtail heartbeat is normal. If it is abnormal, reinstall Logtail.
2. Check whether the log files in log collection configuration are generated in real time.
3. Check whether the regular expression in log collection configuration matches the log content. If the regular expression does not match, view the error in the Logtail run log (Linux:/usr/local/ilogtail/ilogtail.LOG).

Why is the Logtail heartbeat abnormal?

1. Currently, the Logtail agent only supports 64-bit Linux operating systems.
2. Use LogStash to collect logs in a Windows system.

If the Logtail heartbeat is abnormal, follow these steps below to perform diagnosis.

- Check whether the Logtail process exists by running the following command. If the process does not exist, reinstall Logtail. If it exists, go to the next step.

```
sudo /etc/init.d/ilogtaild status
```

Run the following commands to check network connectivity:

Classic network

```
telnet logtail.cn-<region>-intranet.log.aliyuncs.com 80
```

VPC

```
telnet logtail.cn-<region>-vpc.log.aliyuncs.com 80
```

If your machine is not connected, perform the following check:

1. If the machine is configured with host name binding (run the hostname command to view the host name; the related file is /etc/hosts), check whether the bound IP address is the same as that in the Log Service machine group.
2. If no host name is bound, check whether the IP address of the machine’ s first network adapter is the same as that in the Log Service machine group.

If the machine is not connected, the Log Service cannot receive heartbeat packets from the machine. In this case, contact the Log Service technical support team for troubleshooting.

If the problem persists, submit a ticket in the ticket system. The Log Service technical support team will look into the problem.

Logtail heartbeat error

If Logtail machine group heartbeat is abnormal when collect logs using Logtail, you can identify a problem using Logtail checking tool or manual diagnosis.

Automatic diagnosis

Log Service supports Logtail checking tool.

If the checking result is normal, check step 3–6 of **Manual diagnose** based on the echo message of checking result.

Manual diagnose

Logtail machine group heartbeat failed is usually caused by the following reasons, please inspect one by one.

1. Network disconnected

Execute the following command to check the network connectivity, and make sure the network is running normally.

Classic network

telnet logtail.cn-<region>-intranet.log.aliyuncs.com 80

VPC

telnet logtail.cn-<region>-vpc.log.aliyuncs.com 80

Internet

telnet logtail.cn-<region>.log.aliyuncs.com 80

2. Logtail process does not exists

Check whether the Logtail process exists by running the following command. If it does not exist, install Logtail. Identify the region where your project is located and check whether the region matches that in config. If Logtail process exists, go to the next step.

Linux:

```
sudo /etc/init.d/ilogtaild status
```

Windows:

```
Control Panel -> Management Tool -> Service
Check LogtailDaemon, LogtailWorker running status.
```

3. Parameters error

You need to specify the right endpoint for your Logtail. Check the parameters you have already set:

- Linux: /usr/local/ilogtail/ilogtail_config.json
- Windows x64: C:\Program Files (x86)\Alibaba\Logtail\ilogtail_config.json
- Windows x32: C:\Program Files\Alibaba\Logtail\ilogtail_config.json

Make sure that:

- The project is located in the same region with your Logtail Endpoint.
- You hava already choose the same domain name based on network environment of your servers. An internal domain name selected in VPC environment may fail to be connected. Telnet to the domain name that you configured in ilogtail_config.json, such as telnet logtail.cn-hangzhou-intranet.log.aliyuncs.com 80.

4. Logtail was configured Wrong IP or user ID

Usually:

- The IP address is only a tag and does not affect network access. The method of setting this IP address is as follows: Obtain the bound IP address in /etc/hosts. Run the hostname command to set machinename.
- If no IP address is bound, obtain the IP address of the first network adapter (ifconfig eth0).

Check IP address on servers:

- Linux: /usr/local/ilogtail/app_info.json
- Windows x64: C:\Program Files (x86)\Alibaba\Logtail\app_info.json
- Windows x32: C:\Program Files\Alibaba\Logtail\app_info.json

If the IP address filled in machine group console is diffrent with that of Logtail obtained, modify it depend on these circumstances:

- If the IP address filled in machine group console is wrong, please correct it and save the configuration.
- If you changed the network configuration (such as /etc/hosts), restart Logtail to obtain the new IP address.

If necessary, you can execute the following command to restart Logtail.

- Linux: sudo /etc/init.d/ilogtaild stop; sudo /etc/init.d/ilogtaild start
- Windows: **Control Panel** -> **Management Tool** -> **Service** -> **Restart LogtailWorker**

5. AccessKey is not configured

Check /usr/local/ilogtail/ilogtail.LOG if there is a mistake:Unauthorized ErrorMessage:no authority, denied by ACL

If the above error occurs, your main account is not configured AccessKey, so Logtail does not work.Refer to 5 Minute Quick Start to configure AccessKey steps to properly configure AccessKey.

6. The server is non-Alibaba Cloud ECS or not belong to the same account with the current Project of Log Service

As shown below, these are two types of situations that servers installed Logtail must be authenticated to collect logs. For more information, refer to aliuid.

1. The server is non-Alibaba Cloud ECS
2. The server is not belong to the same account with the current Project of Log Service

If the problem persists, submit a ticket in the ticket system. The Log Service technical support team will look into the problem. Please provide your Project name, Logstore name , machine group,

app_info.json, ilogtail_config.json and the checking results of Logtail checking tool.

Logtail quick diagnostic tool

When any exception occurs during log collection, you can use the Logtail automatic checking tool to check the client for exceptions, and quickly locate and solve the problems as instructed by the tool.

If the machine group heartbeat status shows Fail, see What If the Logtail Machine Has No Heartbeat for a solution.

Preparations

Download the checking tool script

wget http://logtail-release.oss-cn-hangzhou.aliyuncs.com/linux64/checkingtool.sh

wget http://logtail-corp.oss-cn-hangzhou-zmf.aliyuncs.com/linux64/checkingtool.sh

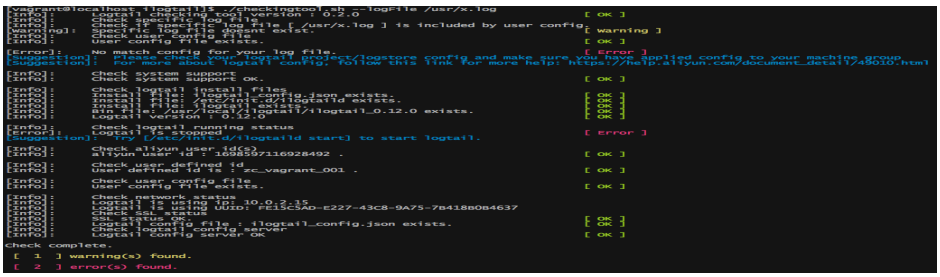
Common parameters for the checking tool

- - help Views help documentation
- - logFile [LogFileFullPath] Verifies whether Logtail collectslogs with a path of LogFileFullPath and also checks the basic running environment of Logtail (such as installation file integrity, running status, Alibaba Cloud userID, and network connectivity)
- - logFileOnly [LogFileFullPath] Only checks whether Logtail collects logs with a path of LogFileFullPath
- - envOnly Only checks the running environment of Logtail

Usage

Run the script ./checkingtool.sh --logFile [LogFileFullPath] to perform the check. If any exception is encountered during the check, proceed as instructed by script.

Note: If the specific log file passes the check and the Logtail running environment is normal, we recommend that you log on to the Alibaba Cloud console to view the exception logs of configuration items of Log Service. For details, see Log Collection Error Query.



Common Logtail collection exceptions

Running the Logtail quick diagnostic tool helps identify the causes of Logtail collection exceptions. You can find an appropriate solution to your problem based on the specific cause. Common causes of Logtail collection problems and their solutions are as follows.

Problem	Solution
Installation file is missing	Reinstall Logtail.
Logtail is not running	Start Logtail with the command /etc/init.d/ilogtaild start.
Multiple Logtail processes	Stop Logtail with the command /etc/init.d/ilogtaild stop and start it with the command /etc/init.d/ilogtaild start.
Port 443 is disabled	Open the firewall port 443.
Config server cannot be found.	Check installation correctness. If the installation is incorrect, uninstall and reinstall Logtail.
User configuration does not exist.	Confirm that Logtail configuration has been created in the console, the client is contained in the machine group, and the configuration has been applied to the machine group.
No specified log file is matched.	Check Logtail configuration correctness.
Specified log files are matched more than once.	Logtail selects one configuration randomly when multiple matches exist, so deduplication is recommended.

Note

- The **Alibaba Cloud ID** and **dynamic machine group/custom ID** configured on the client side are outputted when running the checking tool. No warning is triggered if no such configuration exists. If configuration of Alibaba Cloud ID or dynamic machine group/custom ID is required on the client side, check whether the output of the tool is the same as the one you configured. If not, reconfigure using the methods in Alibaba Cloud UserId Configuration and Dynamic Machine Group Configuration.
- The checking tool needs to use curl to check network connectivity. Make sure the machine has the curl tool installed.

Compare Log Service LogHub and Kafka

Kafka is a distributed messaging system with high throughput and horizontal scaling and is widely used for message publishing and subscription. It is available as an open-source software. You can build a Kafka cluster as needed.

Log Service is a log-specific platform service built based on Apsara Pangu, supports the real-time collection, storage, distribution, and real-time query of various types of logs, and provides services by using standard RESTful APIs.

The Log Service LogHub provides public channels of log collection and distribution. To not build and maintain the Kafka cluster on your own, you can use the Log Service LogHub.

Concept mapping between Log Service LogHub and Kafka

Concept	Kafka	LogHub
Storage object	Topic	Logstore
Horizontal partitioning	Partition	Shard
Data consumption location	Offset	Cursor

Function comparison between LogHub and Kafka

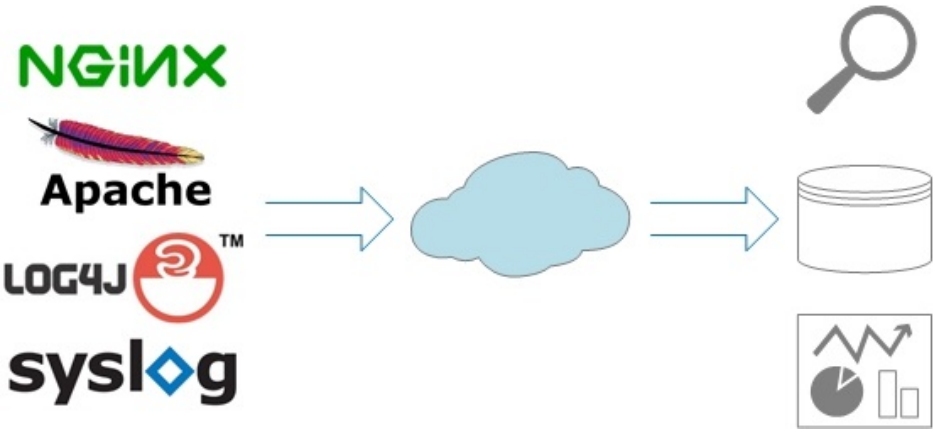
Function	Kafka	LogHub
Usage dependency	Self-built or shared Kafka cluster	Log Service
Communication protocol	Network interconnection by using TCP	HTTP (RESTful API), port 80
Access control	None	Signature authentication + access control based on an Alibaba Cloud account
Dynamic expansion	None	Auto scaling (merge/split) of shard quantities in a dynamic manner without affecting users
Multi-tenant QoS	None	Shard-based standard throttling

Number of data copies	Custom	Three copies by default and cannot be modified
Failover/replication	Completed by using tools	Completed in an automatic and perception-free manner
Expansion/upgrade	Completed by using tools with service impact	Perception-free
Write mode	Round robin/key hash	Currently, only round robin and key hash are supported
Current consumption location	Stored in the zookeeper of the Kafka cluster	Maintained in Log Service without user intervention
Storage period	Based on the configuration	Adjusted dynamically based on the requirement

Comparison among log collection tools: Logstash, fluentd, and Logtail

Assessments of log collection clients

Hundreds of millions of servers, mobile terminals, and network devices generate massive logs every day in the DT era. Centralized log processing solution effectively supports log consumption in the entire lifecycle. The first step is to collect logs from devices to the cloud.



Three log collection tools

Logstash

- Logstash is the “L” of the ELK stack, which is famous in the open source community. It plays an active role in the community and supports many plug-ins in the ecosystem.
- Logstash is implemented based on JRuby and can be run on JVM across platforms.
- Its modular design delivers high scalability and interoperability.

Fluentd

- Fluentd is a popular log collection tool in the open source community. td-agent, the commercial version of fluentd, is maintained by Treasure Data and is assessed in this document.
- Fluentd is implemented based on CRuby and re-implements the components essential for performance by using the C language. The overall performance is good.
- Fluentd features concise design and provides high reliability for the data transfer in the pipeline.
- Compared with Logstash, fluentd supports fewer plug-ins.

- Logtail

- Logtail is the producer of Alibaba Cloud Log Service and has been widely applied in massive big data scenarios of Alibaba Group for more than three years.
- Logtail is implemented by using the C++ language and delivers good performance after great efforts made to improve its stability, resource control capability, and management.
- Compared with the community support of Logstash and fluentd, Logtail is dedicated to log collection with lower functional variety.

Function comparison

Function	Logstash	Fluentd	Logtail
Log reading	Polling	Polling	Event triggered
File rotation	Supported	Supported	Supported
Failover (local checkpoint)	Supported	Supported	Supported
General log parsing	Grok parsing (based on a regular expression)	Parsing based on a regular expression	Parsing based on a regular expression
Specific log type	Supports mainstream formats such as delimiter, key-value, and JSON	Supports mainstream formats such as delimiter, key-value, and JSON	Supports mainstream formats such as delimiter, key-value, and JSON
Data compression before being sent	Supported by using plug-ins	Supported by using plug-ins	LZ4
Data filter	Supported	Supported	Supported

Buffer-based data transfer	Supported by using plug-ins	Supported by using plug-ins	Supported
Transfer exception handling	Supported by using plug-ins	Supported by using plug-ins	Supported
Running environment	JRuby implementation with JVM environment dependency	CRuby and C implementation with Ruby environment dependency	C++ implementation without special requirements
Thread support	Multiple threads	Multiple threads restricted by GIL	Multiple threads
Hot upgrade	Not supported	Not supported	Supported
Centralized configuration management	Not supported	Not supported	Supported
Self-detection of the running status	Not supported	Not supported	Supports CPU/memory threshold protection

Log file collection – Performance comparison

Log sample: Take the Nginx access logs as an example. The following log is a 365-byte Nginx access log with 14 structured fields.

The following test repeatedly writes the log to a file at different simulated pressures. The time field of each log is set to the system time when the log is written, and the other 13 fields are the same for all logs.

The log parsing process in the simulated scenario is the same as that in the actual scenario, except that the network traffic generated by the write operation is reduced due to a relatively high data compression rate.

Logstash

logstash-2.0.0 parses logs by using grok and writes parsed logs to Kafka (which has built-in plug-ins and enables Gzip compression).

Log parsing configuration:

```
grok {
```

```
patterns_dir=>"/home/admin/workspace/survey/logstash/patterns"
match=>{ "message"=>"%{IPORHOST:ip} %{USERNAME:rt} - \[%{HTTPDATE:time}\] \"%{WORD:method}
%{DATA:url}\" %{NUMBER:status} %{NUMBER:size} \"%{DATA:ref}\" \"%{DATA:agent}\" \"%{DATA:cookie_unb}\"
\"%{DATA:cookie_cookie2}\" \"%{DATA:monitor_traceid}\" %{WORD:cell} %{WORD:ups}
%{BASE10NUM:remote_port}" }
remove_field=>["message"]
}
```

Test results

Write TPS	Write traffic (KB/s)	CPU usage (%)	Memory usage (MB)
500	178.22	22.4	427
1000	356.45	46.6	431
5000	1782.23	221.1	440
10000	3564.45	483.7	450

Fluentd

td-agent-2.2.1 parses logs by using a regular expression and writes parsed logs to Kafka (which has the third-party plug-in fluent-plugin-kafka and enables Gzip compression).

Log parsing configuration:

```
<source>
type tail
format /^(?<ip>\S+)\s(?<rt>\d+)\s-
\s\[(?<time>[^\]]*)\]\s"(?<url>[^\"]+)"\s(?<status>\d+)\s(?<size>\d+)\s"(?<ref>[^\"]+)"\s"(?<agent>[^\"]+)"\s"(?<
cookie_unb>\d+)"\s"(?<cookie_cookie2>\w+)"\s"(?
<monitor_traceid>\w+)"\s(?<cell>\w+)\s(?<ups>\w+)\s(?<remote_port>\d+).*$/
time_format %d/%b/%Y:%H:%M:%S %z
path /home/admin/workspace/temp/mock_log/access.log
pos_file /home/admin/workspace/temp/mock_log/nginx_access.pos
tag nginx.access
</source>
```

Test results

Write TPS	Write traffic (KB/s)	CPU usage (%)	Memory usage (MB)
500	178.22	13.5	61
1000	356.45	23.4	61
5000	1782.23	94.3	103

Note: A single process of fluentd uses at most one CPU core due to GIL limits. You can use the multiprocess plug-in to support higher log throughput in multiple processes.

Logtail

Logtail 0.9.4 performs log structuring by using a regular expression and writes LZ4-compressed data to Alibaba Cloud Log Service in the HTTP protocol. The batch_size is set to 4000.

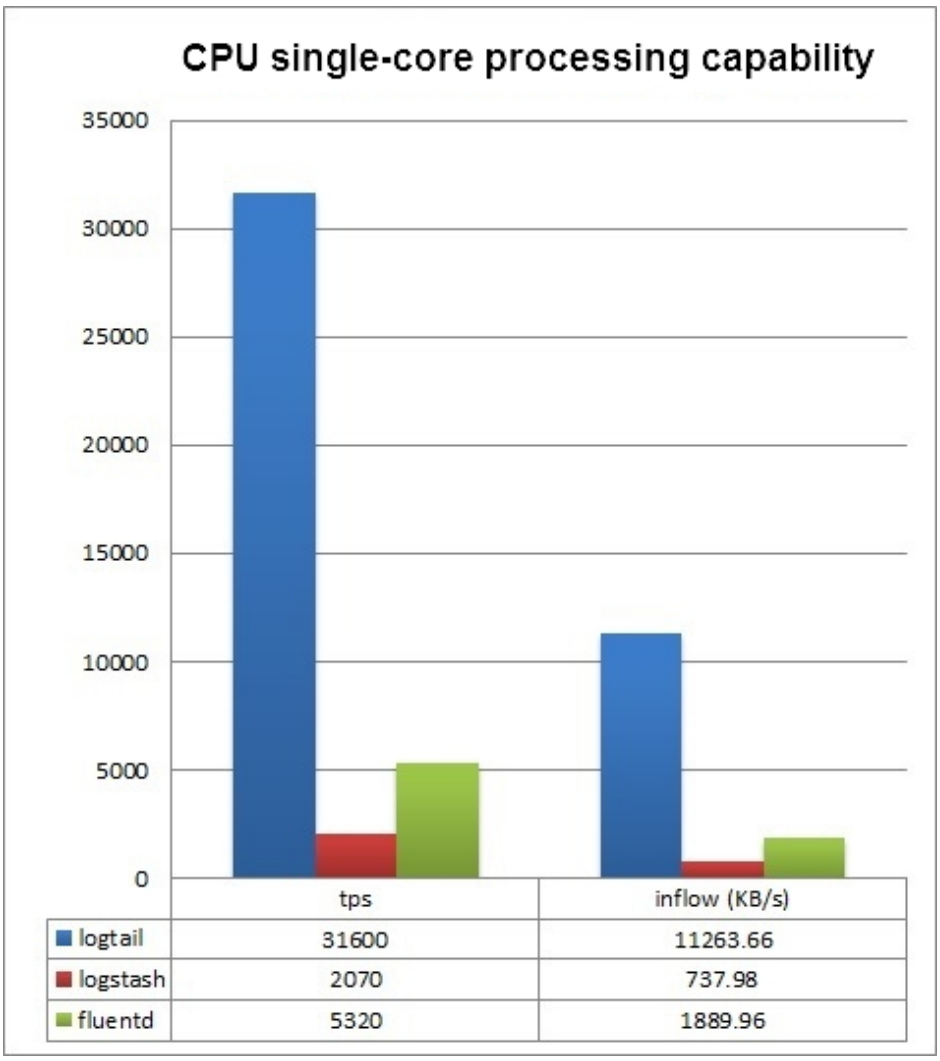
Log parsing configuration:

```
logRegex : (\S+)\s(\d+)\s-
\s\[([^]]+)]\s"([^"]+)"\s(\d+)\s(\d+)\s"([^"]+)"\s"([^"]+)"\s"(\d+)"\s"(\w+)"\s"(\w+)\s(\w+)\s(\w+)\s(\d+).*
keys : ip,rt,time,url,status,size,ref,agent,cookie_unb,cookie_cookie2,monitor_traceid,cell,ups,remote_port
timeformat : %d/%b/%Y:%H:%M:%S
```

Test results

Write TPS	Write traffic (KB/s)	CPU usage (%)	Memory usage (MB)
500	178.22	1.7	13
1000	356.45	3	15
5000	1782.23	15.3	23
10000	3564.45	31.6	25

Single-core processing capability comparison



	tps	inflow (KB/s)
logtail	31600	11263.66
logstash	2070	737.98
fluentd	5320	1889.96

Conclusion

Logstash, fluentd, and Logtail have their own features as follows:

- Logstash supports all the mainstream log types, diverse plug-ins, and flexible customization, but has relatively low performance and is prone to high memory usage because of JVM.
- Fluentd supports all the mainstream log types and many plug-ins, and delivers good performance.
- Logtail occupies the least CPU and memory resources of the machine, delivers good performance throughput, and provides full support for common log collection scenarios. However, it has no plug-in support and delivers lower flexibility and scalability than Logstash and fluentd.

Log query

FAQs about log query

Question list

1. How to query an IP address in logs?
2. How to query a keyword containing a space in logs?
3. How to query logs based on two query conditions?
4. What methods does Log Service provide to query collected logs?
5. What query capabilities does Log Service provide?
6. What are the limits of log query?

Solutions

1. How to query an IP address in logs?

Log Service supports querying IP addresses in logs in the full match way. You can directly query logs related to an IP address, such as the logs containing or excluding the specified IP address. However, partial match is not supported, that is, you cannot query a part of an IP address directly because a dot (.) is not the default token of Log Service. You can filter the IP addresses on your own if needed. For example, download data by using SDKs first and then use a regular expression or string.indexof in the codes to determine the IP addresses.

For example, the query condition in a Log Service project is not ip:121.42.0 not status:200 not 360jk not DNSPod-Monitor not status:302 not jiankongbao not 301 and status:403.

The 121.42.0 Classless Inter-Domain Routing (CIDR) block addresses are still in the query results. This is because Log Service considers 121.42.0.x as a word and only queries the correct results if the query condition is 121.42.0.x. Therefore, the IP address 121.42.0 is not filtered in the query results if the query condition is not 121.42.0.

2. How to query a keyword containing a space in logs?

Logs containing the keyword at the left or right of the space can also be queried if you directly enter the keyword containing a space as the query condition. Therefore, we recommend that you enclose the keyword containing a space in quotation marks (“”) and use the contents in the quotation marks (“”) as a keyword to query the logs. Then, only logs with the keyword containing a space can be queried.

For example, query the logs containing the keyword POS version in the following logs:

```
post():351]:&amp;nbsp;device_id:&amp;nbsp;BTAddr&amp;nbsp;:&amp;nbsp;B6:xF:xx:65:xx:A1&amp;nbsp;IMEI&amp;nbsp;:&amp;nbsp;35847xx22xx81x9&amp;nbsp;WifiAddr&amp;nbsp;:&amp;nbsp;4c:xx:0e:xx:4e:xx&amp;nbsp;|&amp;nbsp;user_id:&amp;nbsp;bb07263xxd2axx43xx9exxea26e39e5f&amp;nbsp;POS&amp;nbsp;version:903
```

Logs containing POS or version are also included in the query results if you directly use POS version as the query condition, which does not meet the query requirement. Use "POS version" as the query condition and then all the logs containing the keyword POS version can be queried.

3. How to query logs based on two query conditions?

Enter two statements at the same time if you have two query conditions.

For example, to query logs whose status is not OK or Unknown in a Logstore, directly use not OK not Unknown as the query condition to query the logs.

4. What methods does Log Service provide to query collected logs?

Log Service provides three methods to query logs:

1. Query logs in the Log Service console. For more information, see Query logs.
2. Use SDKs to query logs. For more information, see SDK.
3. Use RESTful APIs to query logs. For more information, see API.

5. What query capabilities does Log Service provide?

- Supports filtering and querying logs by using a combined condition. For more information about the query syntax, see Query syntax.
- Supports querying one billion logs in one second for a single query. You can query logs based on specified query conditions, read the time-based distribution of the query results, and obtain the raw logs.
- Supports caching logs, allowing you to obtain more complete query results for a second query with the same query condition.

6. What are the limits of log query?

- Supports querying logs based on a combined condition composed of at most 30 words.
- Supports obtaining at most 100 lines of raw logs for a single query. You can download more logs by turning the page.
- Supports processing one billion lines of logs within one second for a single query.

Differences between log consumption and log query

Log Service provides two functions related to the read operation.

Log collection and consumption (LogHub): Provides public channels for log collection and distribution, sequential (first in, first out (FIFO)) read and write of full data, and functions similar to Kafka.

- Each Logstore has one or more shards. Data is written to a shard at random.
- You can read logs in batches from a specified shard according to the sequence that logs are written to the shard.
- You can set the start point (cursor) to pull logs from shards in batches based on the time when Log Service receives logs.
- By default, logs are retained in LogHub for two days, during which logs can be consumed.

Log query (index): Log Service supports querying massive logs based on LogHub. You can query logs by using keywords.

- Use the keyword to query logs that meet your requirement.
- Supports using the boolean combination of AND, NOT, and OR to query logs based on keywords.
- You can only query logs in all shards, but not a specified shard.

Differences

Function	Log query (LogSearch)	Log collection and consumption (LogHub)
Query logs by using keywords	Supported	Not supported
Read small amounts of data	Fast	Fast
Read full data	Slow (100 logs every 100ms, not recommended)	Fast (1 MB logs every 10ms, recommended)
Read logs by topic	Yes	No. Logs are read by shard
Read logs by shard	No. You can only query logs in all shards	Yes. You must specify a shard for reading logs
Cost	Relatively high	Low
Scenario	Scenarios that need to filter data such as monitoring and troubleshooting	Full processing scenarios such as stream computing and batch processing

Common errors for log query and analysis

This document describes the common errors for log query and analysis. For more information about the basic syntaxes, see Syntax description.

List of common errors

1. line 1:44: Column ‘my_key_field’ cannot be resolved;please add the column in the index attribute
2. Column ‘xxxx__line__ ‘ not in GROUP BY clause;please add the column in the index attribute
3. sql query must follow search query,please read syntex doc
4. key word(where) is not supported,please read query syntex # | *select apiName,count()* as count where apiName=” ” group by apiName order by count desc limit 10
5. please read syntex document,and make sure all related fields are indexed. error after select .error detail:line 1:10: identifiers must not start with a digit; surround the identifier with double quotes
6. please read syntex document,and make sure all related fields are indexed. error after select .error detail:line 1:9: extraneous input ‘’ expecting

1. line 1:44: Column ‘my_key_field’ cannot be resolved;please add the column in the index attribute

Error cause: The key my_key_field does not exist. Therefore, you cannot reference the key for query.

Solution: On the query page, add this field as a key/value index and enable the analytics in the index attributes.

2. Column ‘xxxx__line__ ‘ not in GROUP BY clause;please add the column in the index attribute

Error cause: You use the GROUP BY syntax for query, but a non-agg field which is not contained in GROUP BY is referenced in SELECT. For example, in select key1, avg(latency) group by key2, key1 is not contained in GROUP BY.

Solution: The correct syntax is select key1,avg(latency) group by key1,key2.

3. sql query must follow search query,please read syntex doc

Error cause: The filter condition is not specified. For example, select ip,count(*) group by ip.

Solution: The correct syntax is *|select ip,count(*) group by ip.

4. key word(where) is not supported,please read query syntex # | *select apiName,count()* as count where apiName=” ” group by apiName order by count desc limit 10

Error cause: The SQL syntax contains a WHERE condition, which is not allowed.

Solution: Add WHERE in the filter condition. The correct syntax is apiName:"" | select apiName,count(*) as count group by apiName order by count desc limit 10.

5. please read syntex document,and make sure all related fields are indexed. error after select .error detail:line 1:10: identifiers must not start with a digit; surround the identifier with double quotes

Error cause: The column name or variable name referenced in SQL starts with a number, which is not allowed.

Solution: Modify the name to start with a letter.

6. please read syntex document,and make sure all related fields are indexed. error after select .error detail:line 1:9: extraneous input ‘’ expecting

Error cause: The word spelling is incorrect.

Solution: Find the incorrect word spelling according to the location specified in the error and correct the spelling.

Common Errors of Log Service

Errors of Log Service:

1. illegal param! [LogContent] is null
2. send data fail, error_code:WriteQuotaExceed error_message:Write quota exceed projectName:project_name
3. WARN Aliyun.SLS.Logtail.LogFileReader - logfile xxx.log has n old enough logs with first fail case xxx

Solutions

1. illegal param! [LogContent] is null

Check the following configurations:

Check whether you have entered all the required information of the sample log.

Check whether the regular expressions in the first line are correct.

If your issue is still not resolved after the preceding troubleshooting process, send us the sample log and the regular expressions, so that we can reproduce the issue for solutions.

If this issue continues, contact our after-sales technical support.

2. send data fail, error_code:WriteQuotaExceed error_message:Write quota exceed projectName:project_name

If a similar error occurs in ilogtail.log when you use Log Service:

send data fail, error_code:WriteQuotaExceed error_message:Write quota exceed projectName:project_name

there are insufficient quota for writing. The volume of logs you write into it exceeds the predefined threshold, or the write speed is higher than the limit.

Currently, the processing capacity of each shard is:Write: 5 MB/s, 2,000 times/sRead: 10 MB/s, 100 times/s

If your data volume is beyond what the shard can process, you can split the shard. For more information, see Split the Shard.The maximum number of write requests per minute at the project

level is 300 thousand. If you write logs using a program, and some requests may exceed the quota limit, we recommend that you write logs in bulk or use Producer-lib to limit the maximum packet size for each upload to 3 MB and the maximum entry counts to 4,096.

3. WARN Aliyun.SLS.Logtail.LogFileReader - logfile xxx.log has n old enough logs with first fail case xxx

The following error messages may appear in the ilogtail.log:

WARN Aliyun.SLS.Logtail.LogFileReader - logfile xxx.log has n old enough logs with first fail case xxx

This is caused by the old logs.

Log collection rules for Logtail:

Process the historical data separately.

Reduce the cache time for data flushing or even perform real-time flushing.

Pay attention to the time zone when you change the log entries.

Troubleshooting solutions:

When a new log file is monitored by Logtail, the logs written in the log file in the past one minute are considered as old data and discarded.

The new data written in the past five minutes is also considered as old data and discarded. The error indicates that your logs are cached in the memory and has timed out when you actually write them into files.

For the log time beyond the range of -7 days to 360s, such logs are discarded by the server. The error indicates that you do not correctly set the time zone of the log.

If the time zone setting does not conform to rule 1, the historical data is discarded. But if there is no violation of other rules, no error occurs in the subsequent logs.

If rule 2 is violated, an error may occur occasionally, and some logs can be viewed in the console.

If rule 3 is violated, logs are not collected, so they cannot be viewed in the console.

If this issue continues, contact our after-sales technical support.