日志服务

产品简介

腾讯云

【联系我们】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或95716。

文档目录

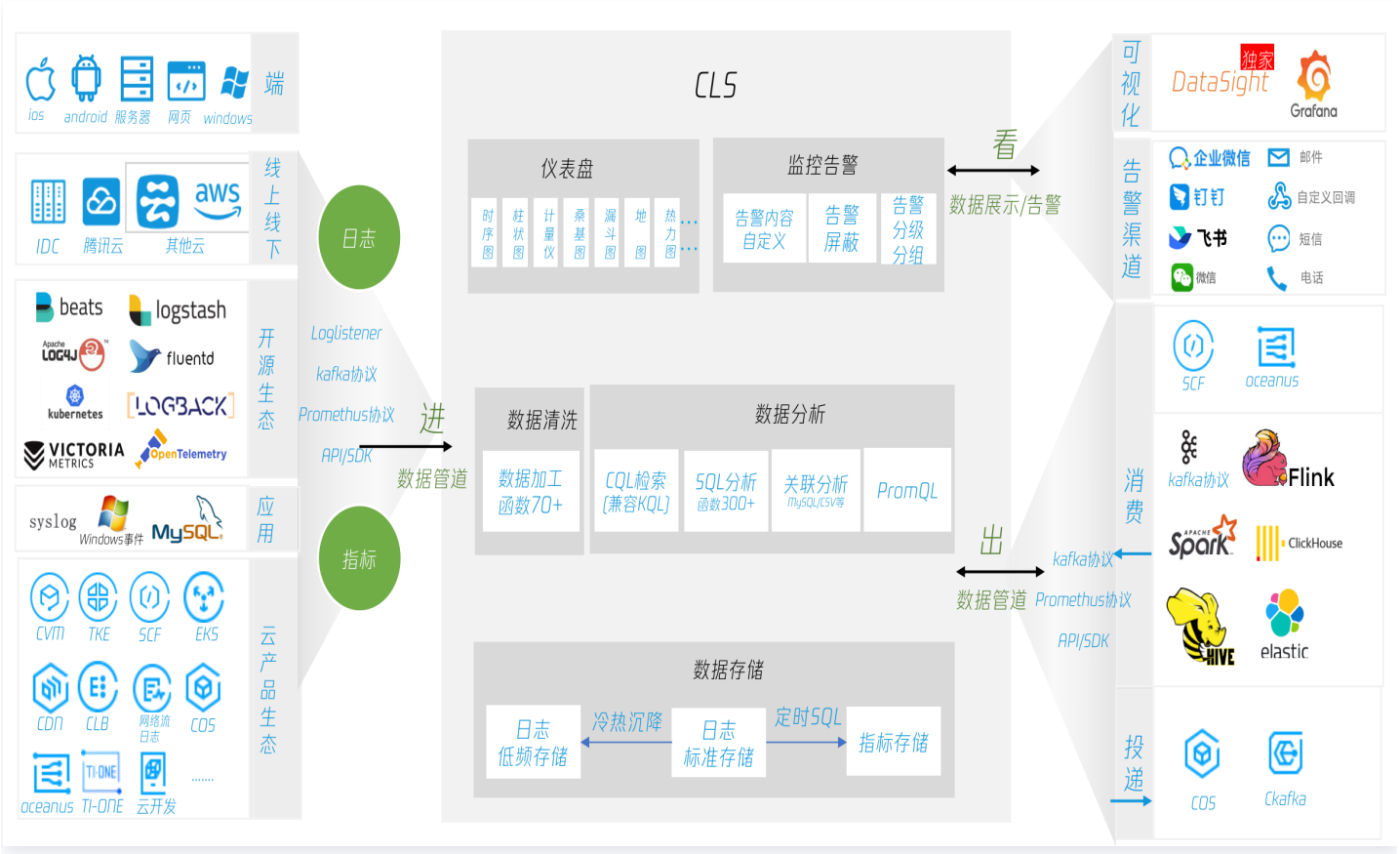
产品简介
产品概述

最近更新时间：2024-11-01 11:45:22

日志服务（Cloud Log Service，CLS）是一款支持日志（Log）、指标（Metric）数据的采集、ETL、检索分析、可视化及告警的一体化可观测 SaaS 服务。

功能概览

日志服务主要提供以下功能：



日志采集

日志服务目前支持 LogListener、API、Kafka 协议、对象存储（Cloud Object Storage，COS）导入等多种采集方式：

- Loglistener 实时采集：使用腾讯云 LogListener 采集日志，便捷安装，服务稳定可靠安全、支持大部分主流 Linux 操作系统，兼顾高性能与低资源占用。
- 通过 API：无需安装 LogListener，直接调用 API 上传日志，支持多种语言。

- Kafka 协议：支持使用 Kafka Producer SDK 上传日志到 CLS。
- COS 导入：将腾讯云对象存储中的数据导入到日志服务。

日志存储

根据用户对日志不同的检索时延性及日志处理能力需求，提供两种存储：

- 实时存储：适用于用户对日志有统计分析的需求，提供日志的秒级检索，实时统计分析，实时监控，流式消费等应用能力。
- 低频存储：适用于访问频率较低且无统计分析需求的日志，如审计归档日志。低频存储提供全文检索日志的能力，满足用户对历史日志的回溯检索，归档存储等诉求。长期存储下，整体使用成本相比实时存储降低80%。

日志检索分析

针对存储在 CLS 的日志进行实时检索分析，帮助用户快速定位异常日志、统计系统及业务指标。

- 日志检索：使用关键词按全文或字段检索日志。
- 统计分析：使用 SQL 灵活统计日志内的系统及业务指标并通过图表进行展示，兼容 SQL 92标准，支持200+ SQL 函数。例如按省份统计业务请求数，按时间查询请求错误率变化趋势。
- 性能优异：秒级返回查询结果，支持亿级别日志检索分析。

指标采集与存储

兼容 Prometheus 生态，提供如下指标采集与存储功能：

- 指标采集：支持 Prometheus Remote Write 协议，可使用兼容该协议的各类采集器采集指标并上报至指标主题，例如 VMAgent 及Telegraf。
- 日志转指标：原始日志可通过执行定时 SQL 统计，获取其中的关键指标，例如基于日志统计接口吞吐量及耗时。
- 指标存储：兼容 Prometheus 指标数据模型，可使用 PromQL 对指标进行查询和统计计算。

仪表盘

仪表盘提供丰富的图表类型，支持用户统计分析日志数据与指标数据，并通过图表进行展示。

- 多种图表类型：支持时序图、饼图、单值图、桑基图、地图、词云、全局过滤与变量配置等。
- 订阅仪表盘：支持订阅仪表盘，可以将仪表盘以图片的形式导出。用邮件、企业微信等方式定期发送给指定的对象。
- 交互事件：图表自定义交互事件功能，支持单击图表内容触发交互打开检索分析页面、仪表盘页面、第三方 URL 等。

DataSight

DataSight 是 CLS 日志服务提供的独立控制台，支持免腾讯云账号使用日志服务，支持自定义的账号密码登录或免登录。适用于多人多团队共用 CLS 日志服务的场景。

告警

日志内出现较多错误日志或系统及业务指标超出阈值时秒级告警，主动发现系统及业务异常问题。

- 通知渠道：支持电话、短信、邮件、微信、企业微信、钉钉、飞书和自定义接口回调。
- 多维分析：触发告警时可针对原始日志进行额外的检索分析，将结果附加在告警通知中，辅助定位告警原因。

日志数据处理

- 数据加工：日志过滤清洗、脱敏、结构化、分发。 对非结构化的日志而言，通过数据加工处理成结构化数据，可用于后续的 SQL 分析、仪表盘、告警等。数据加工支持按照指定的规则，将原始日志分发到不同的日志主题。
- 定时 SQL 分析：将日志转为指标（Metric），然后使用 Metric 做仪表盘、告警，查询性能更高，存储量更低。日志 OLAP，按照指定时间和窗口调度离线 SQL 分析任务，生成结果数据，例如日报表、周报表。

日志投递与消费

您可以将日志投递至腾讯云产品中，如对象存储 COS、消息中间件 Ckafka，也可以将 CLS 日志消费到其它大数据组件。

- 日志投递至 COS：日志深度归档、大数据计算。
- 日志投递至 Ckafka：日志的实时处理和计算。
- Kafka 协议消费：支持原生 Kafka 协议，轻松对接大数据组件，如 Flink、Logstash、Flume，也可使用 Python/Java SDK 消费。

产品优势

最近更新时间：2024-05-29 14:28:51

功能丰富

- 提供采集、存储、检索、转存投递等功能一站式日志服务。
- 采集客户端 LogListener 提供单行/多行全文、分隔符、JSON、正则等日志结构化解析方式。
- 提供多种数据接入方式，用户可根据业务情况选择适合的接入方式，详情请参见 采集概述 。
- 提供丰富的检索语法，方便用户进行关键词查询、模糊查询、范围查询等日志查询操作。

稳定可靠

- 日志服务采用高可扩展性的分布式存储架构，支持横向水平扩容，服务弹性伸缩，轻松存储管理海量日志数据。
- 日志服务后端存储采用多副本机制管理存储日志数据，为数据安全提供可靠性保障。

简单高效

- 采集端 LogListener 提供界面式的配置方式，配置简单直观，使用 LogListener 可快速接入日志服务。
- 数据写入 CLS 即可被消费，亿级数据查询支持秒级返回结果。
- 服务按实际用量收费，无需单独搭建和运维日志系统，避免了资源闲置浪费问题。

生态扩展

- 部分云产品日志已接入 CLS，详情请参见 接入列表 。
- 日志数据投递 COS，满足对日志数据长时间归档存储的需求。
- 日志数据投递 Ckafka，满足对日志数据实时消费的需求，便于进一步处理分析。

地域和访问域名

最近更新时间：2025-04-24 10:34:12

简介

地域（Region）是指数据中心所在的区域，不同地域之间无法通过内网进行数据流转。日志服务（Cloud Log Service，CLS）基于数据中心提供服务。您可根据自己的业务场景以及日志数据所在的地理位置选择就近的地域，以降低日志数据上报及访问延时。
例如，当您的业务分布在华南地区，那么选择在广州地域创建日志主题及相关应用。

可用地域

地域	地域简称
北京	ap-beijing
广州	ap-guangzhou
上海	ap-shanghai
成都	ap-chengdu
南京	ap-nanjing
重庆	ap-chongqing
中国香港	ap-hongkong
硅谷	na-siliconvalley
弗吉尼亚	na-ashburn
新加坡	ap-singapore
曼谷	ap-bangkok
法兰克福	eu-frankfurt
东京	ap-tokyo
首尔	ap-seoul
雅加达	ap-jakarta
圣保罗	sa-saopaulo

深圳金融	ap-shenzhen-fsi
上海金融	ap-shanghai-fsi
北京金融	ap-beijing-fsi
上海自动驾驶云	ap-shanghai-adc

说明：

- **针对金融行业监管要求定制的合规专区，具有高安全，高隔离性的特点。目前日志服务已支持深圳、上海、北京金融专区。已认证通过的金融行业客户可 提交工单 申请使用专区，详情请参见 金融专区介绍 。**
- **如果日志服务中接入了其他云产品，请您尽量选择与其他云产品相同的地域。相同地域的云产品之间通过内网读写数据，能有效降低延迟、提高访问速度。**

域名

日志服务在不同模块使用的域名有所区别，具体如下：

LogListener

LogListener 是 CLS 所提供的日志采集客户端，可将机器本地的日志上报至 CLS，其所使用的域名如下：

地域	简称	内网域名	外网域名
北京	ap-beijing	ap-beijing.cls.tencentyun.com	ap-beijing.cls.tencentcs.com
广州	ap-guangzhou	ap-guangzhou.cls.tencentyun.com	ap-guangzhou.cls.tencentcs.com
上海	ap-shanghai	ap-shanghai.cls.tencentyun.com	ap-shanghai.cls.tencentcs.com
成都	ap-chengdu	ap-chengdu.cls.tencentyun.com	ap-chengdu.cls.tencentcs.com
南京	ap-nanjing	ap-nanjing.cls.tencentyun.co	ap-nanjing.cls.tencentcs.com

		m	
重庆	ap-chongqing	ap-chongqing.cls.tencentyun.com	ap-chongqing.cls.tencentcs.com
中国香港	ap-hongkong	ap-hongkong.cls.tencentyun.com	ap-hongkong.cls.tencentcs.com
硅谷	na-siliconvalley	na-siliconvalley.cls.tencentyun.com	na-siliconvalley.cls.tencentcs.com
弗吉尼亚	na-ashburn	na-ashburn.cls.tencentyun.com	na-ashburn.cls.tencentcs.com
新加坡	ap-singapore	ap-singapore.cls.tencentyun.com	ap-singapore.cls.tencentcs.com
曼谷	ap-bangkok	ap-bangkok.cls.tencentyun.com	ap-bangkok.cls.tencentcs.com
法兰克福	eu-frankfurt	eu-frankfurt.cls.tencentyun.com	eu-frankfurt.cls.tencentcs.com
东京	ap-tokyo	ap-tokyo.cls.tencentyun.com	ap-tokyo.cls.tencentcs.com
首尔	ap-seoul	ap-seoul.cls.tencentyun.com	ap-seoul.cls.tencentcs.com
雅加达	ap-jakarta	ap-jakarta.cls.tencentyun.com	ap-jakarta.cls.tencentcs.com
圣保罗	sa-saopaulo	sa-saopaulo.cls.tencentyun.com	sa-saopaulo.cls.tencentcs.com
深圳金融	ap-shenzhen-fsi	ap-shenzhen-fsi.cls.tencentyun.com	ap-shenzhen-fsi.cls.tencentcs.com

上海金融	ap-shanghai-fsi	ap-shanghai-fsi.cls.tencentyun.com	ap-shanghai-fsi.cls.tencentcs.com
北京金融	ap-beijing-fsi	ap-beijing-fsi.cls.tencentyun.com	ap-beijing-fsi.cls.tencentcs.com
上海自动驾驶云	ap-shanghai-adc	ap-shanghai-adc.cls.tencentyun.com	-

日志服务 API 3.0

日志服务 API 3.0 是 CLS 最新版本的 API，符合腾讯云统一的 API 规范，可通过 API 管理日志主题和告警策略等资源，其所使用的域名如下：

> **说明**
> 通过外网访问时，也可使用统一域名 cls.tencentcloudapi.com（仅支持非金融区），将根据调用接口时客户端所在位置，自动解析到最近的某个具体地域的服务器，对时延敏感的业务，建议指定带地域的域名。

地域	简称	内网域名	外网域名
北京	ap-beijing	cls.internal.tencentcloudapi.com	cls.ap-beijing.tencentcloudapi.com
广州	ap-guangzhou	cls.internal.tencentcloudapi.com	cls.ap-guangzhou.tencentcloudapi.com
上海	ap-shanghai	cls.internal.tencentcloudapi.com	cls.ap-shanghai.tencentcloudapi.com
成都	ap-chengdu	cls.internal.tencentcloudapi.com	cls.ap-chengdu.tencentcloudapi.com
南京	ap-nanjing	cls.internal.tencentcloudapi.com	cls.ap-nanjing.tencentcloudapi.com

重庆	ap-chongqing	cls.internal.tencentcloudapi.com	cls.ap-chongqing.tencentcloudapi.com
中国香港	ap-hongkong	cls.internal.tencentcloudapi.com	cls.ap-hongkong.tencentcloudapi.com
硅谷	na-siliconvalley	cls.internal.tencentcloudapi.com	cls.na-siliconvalley.tencentcloudapi.com
弗吉尼亚	na-ashburn	cls.internal.tencentcloudapi.com	cls.na-ashburn.tencentcloudapi.com
新加坡	ap-singapore	cls.internal.tencentcloudapi.com	cls.ap-singapore.tencentcloudapi.com
曼谷	ap-bangkok	cls.internal.tencentcloudapi.com	cls.ap-bangkok.tencentcloudapi.com
法兰克福	eu-frankfurt	cls.internal.tencentcloudapi.com	cls.eu-frankfurt.tencentcloudapi.com
东京	ap-tokyo	cls.internal.tencentcloudapi.com	cls.ap-tokyo.tencentcloudapi.com
首尔	ap-seoul	cls.internal.tencentcloudapi.com	cls.ap-seoul.tencentcloudapi.com
雅加达	ap-jakarta	cls.internal.tencentcloudapi.com	cls.ap-jakarta.tencentcloudapi.com
圣保罗	sa-saopaulo	cls.internal.tencentcloudapi.com	cls.sa-saopaulo.tencentcloudapi.com
深圳金融	ap-shenzhen-fsi	cls.internal.tencentcloudapi.com	cls.ap-shenzhen-fsi.tencentcloudapi.com
上海金融	ap-shanghai-fsi	cls.internal.tencentcloudapi.com	cls.ap-shanghai-fsi.tencentcloudapi.com

北京金融	ap-beijing-fsi	cls.internal.tencentcloudapi.com	cls.ap-beijing-fsi.tencentcloudapi.com
上海自动驾驶云	ap-shanghai-adc	cls.internal.tencentcloudapi.com	-

API 上传日志

以下域名用于 API 上传日志，具体域名如下：（除上传日志外的其他接口，请更新至日志服务 API 3.0。）

地域	简称	内网域名	外网域名
北京	ap-beijing	ap-beijing.cls.tencentyun.com	ap-beijing.cls.tencentcs.com
广州	ap-guangzhou	ap-guangzhou.cls.tencentyun.com	ap-guangzhou.cls.tencentcs.com
上海	ap-shanghai	ap-shanghai.cls.tencentyun.com	ap-shanghai.cls.tencentcs.com
成都	ap-chengdu	ap-chengdu.cls.tencentyun.com	ap-chengdu.cls.tencentcs.com
南京	ap-nanjing	ap-nanjing.cls.tencentyun.com	ap-nanjing.cls.tencentcs.com
重庆	ap-chongqing	ap-chongqing.cls.tencentyun.com	ap-chongqing.cls.tencentcs.com
中国香港	ap-hongkong	ap-hongkong.cls.tencentyun.com	ap-hongkong.cls.tencentcs.com
硅谷	na-siliconvalley	na-siliconvalley.cls.tencentyun.com	na-siliconvalley.cls.tencentcs.com

弗吉尼亚	na-ashburn	na-ashburn.cls.tencentyun.com	na-ashburn.cls.tencentcs.com
新加坡	ap-singapore	ap-singapore.cls.tencentyun.com	ap-singapore.cls.tencentcs.com
曼谷	ap-bangkok	ap-bangkok.cls.tencentyun.com	ap-bangkok.cls.tencentcs.com
法兰克福	eu-frankfurt	eu-frankfurt.cls.tencentyun.com	eu-frankfurt.cls.tencentcs.com
东京	ap-tokyo	ap-tokyo.cls.tencentyun.com	ap-tokyo.cls.tencentcs.com
首尔	ap-seoul	ap-seoul.cls.tencentyun.com	ap-seoul.cls.tencentcs.com
雅加达	ap-jakarta	ap-jakarta.cls.tencentyun.com	ap-jakarta.cls.tencentcs.com
圣保罗	sa-saopaulo	sa-saopaulo.cls.tencentyun.com	sa-saopaulo.cls.tencentcs.com
深圳金融	ap-shenzhen-fsi	ap-shenzhen-fsi.cls.tencentyun.com	ap-shenzhen-fsi.cls.tencentcs.com
上海金融	ap-shanghai-fsi	ap-shanghai-fsi.cls.tencentyun.com	ap-shanghai-fsi.cls.tencentcs.com
北京金融	ap-beijing-fsi	ap-beijing-fsi.cls.tencentyun.com	ap-beijing-fsi.cls.tencentcs.com
上海自动驾驶云	ap-shanghai-adc	ap-shanghai-adc.cls.tencentyun.com	-

Kafka 上传日志

使用 Kafka 协议上传日志 支持使用 Kafka Producer SDK 和其他 Kafka 相关 agent 上传日志到日志服务，其所使用的域名如下：

地域	简称	内网域名	外网域名
北京	ap-beijing	bj-producer.cls.tencentyun.com	bj-producer.cls.tencentcs.com
广州	ap-guangzhou	gz-producer.cls.tencentyun.com	gz-producer.cls.tencentcs.com
上海	ap-shanghai	sh-producer.cls.tencentyun.com	sh-producer.cls.tencentcs.com
成都	ap-chengdu	cd-producer.cls.tencentyun.com	cd-producer.cls.tencentcs.com
南京	ap-nanjing	nj-producer.cls.tencentyun.com	nj-producer.cls.tencentcs.com
重庆	ap-chongqing	cq-producer.cls.tencentyun.com	cq-producer.cls.tencentcs.com
中国香港	ap-hongkong	hk-producer.cls.tencentyun.com	hk-producer.cls.tencentcs.com
硅谷	na-siliconvalley	usw-producer.cls.tencentyun.com	usw-producer.cls.tencentcs.com
弗吉尼亚	na-ashburn	use-producer.cls.tencentyun.com	use-producer.cls.tencentcs.com
新加坡	ap-singapore	sg-producer.cls.tencentyun.com	sg-producer.cls.tencentcs.com

曼谷	ap-bangkok	th-producer.cls.tencentyun.com	th-producer.cls.tencentcs.com
法兰克福	eu-frankfurt	de-producer.cls.tencentyun.com	de-producer.cls.tencentcs.com
东京	ap-tokyo	jp-producer.cls.tencentyun.com	jp-producer.cls.tencentcs.com
首尔	ap-seoul	kr-producer.cls.tencentyun.com	kr-producer.cls.tencentcs.com
雅加达	ap-jakarta	jkt-producer.cls.tencentyun.com	jkt-producer.cls.tencentcs.com
深圳金融	ap-shenzhen-fsi	szjr-producer.cls.tencentyun.com	szjr-producer.cls.tencentcs.com
上海金融	ap-shanghai-fsi	shjr-producer.cls.tencentyun.com	shjr-producer.cls.tencentcs.com
北京金融	ap-beijing-fsi	bjjr-producer.cls.tencentyun.com	bjjr-producer.cls.tencentcs.com
上海自动驾驶云	ap-shanghai-adc	ap-shanghai-adc.cls.tencentyun.com	-

Kafka 消费日志

使用 Kafka 协议消费日志 支持使用 Kafka Consumer SDK 和其他大数据组件消费到用户的数据仓库，其所使用的域名如下：

地域	简称	内网域名	外网域名

北京	ap-beijing	kafkaconsumer-ap-beijing.cls.tencentyun.com	kafkaconsumer-ap-beijing.cls.tencentcs.com
广州	ap-guangzhou	kafkaconsumer-ap-guangzhou.cls.tencentyun.com	kafkaconsumer-ap-guangzhou.cls.tencentcs.com
上海	ap-shanghai	kafkaconsumer-ap-shanghai.cls.tencentyun.com	kafkaconsumer-ap-shanghai.cls.tencentcs.com
成都	ap-chengdu	kafkaconsumer-ap-chengdu.cls.tencentyun.com	kafkaconsumer-ap-chengdu.cls.tencentcs.com
南京	ap-nanjing	kafkaconsumer-ap-nanjing.cls.tencentyun.com	kafkaconsumer-ap-nanjing.cls.tencentcs.com
重庆	ap-chongqing	kafkaconsumer-ap-chongqing.cls.tencentyun.com	kafkaconsumer-ap-chongqing.cls.tencentcs.com
中国香港	ap-hongkong	kafkaconsumer-ap-hongkong.cls.tencentyun.com	kafkaconsumer-ap-hongkong.cls.tencentcs.com
硅谷	na-siliconvalley	kafkaconsumer-na-siliconvalley.cls.tencentyun.com	kafkaconsumer-na-siliconvalley.cls.tencentcs.com
弗吉尼亚	na-ashburn	kafkaconsumer-na-ashburn.cls.tencentyun.com	kafkaconsumer-na-ashburn.cls.tencentcs.com
新加坡	ap-singapore	kafkaconsumer-ap-singapore.cls.tencentyun.com	kafkaconsumer-ap-singapore.cls.tencentcs.com
曼谷	ap-bangkok	kafkaconsumer-ap-bangkok.cls.tencentyun.com	kafkaconsumer-ap-bangkok.cls.tencentcs.com
法兰克福	eu-frankfurt	kafkaconsumer-eu-frankfurt.cls.tencentyun.com	kafkaconsumer-eu-frankfurt.cls.tencentcs.com

东京	ap-tokyo	kafkaconsumer-ap-tokyo.cls.tencentyun.com	kafkaconsumer-ap-tokyo.cls.tencentcs.com
首尔	ap-seoul	kafkaconsumer-ap-seoul.cls.tencentyun.com	kafkaconsumer-ap-seoul.cls.tencentcs.com
雅加达	ap-jakarta	kafkaconsumer-ap-jakarta.cls.tencentyun.com	kafkaconsumer-ap-jakarta.cls.tencentcs.com
深圳金融	ap-shenzhen-fsi	kafkaconsumer-ap-shenzhen-fsi.cls.tencentyun.com	kafkaconsumer-ap-shenzhen-fsi.cls.tencentcs.com
上海金融	ap-shanghai-fsi	kafkaconsumer-ap-shanghai-fsi.cls.tencentyun.com	kafkaconsumer-ap-shanghai-fsi.cls.tencentcs.com
北京金融	ap-beijing-fsi	kafkaconsumer-ap-beijing-fsi.cls.tencentyun.com	kafkaconsumer-ap-beijing-fsi.cls.tencentcs.com
上海自动驾驶云	ap-shanghai-adc	kafkaconsumer-ap-shanghai-adc.cls.tencentyun.com	-

规格与限制

基础资源

最近更新时间：2024-12-13 11:59:42

配额限制

> **说明：**
> 如当前配额无法满足用户需求，可 提交工单 申请放开目标配额。

相关词汇及说明如下表所示：

相关词汇	说明
日志集	每个账号每个地域下最多可以创建100个日志集
	每个日志集下最多可以创建500个日志主题及指标主题
	名称长度1至255个字符，允许的字符为 a-z、A-Z、0-9、_、-
日志主题	主题名称长度为1至255个字符，允许的字符为 a-z、A-Z、0-9、_、-
	每个日志主题下最多50个分区，新增日志主题时最多能创建10个分区
	每个主题最多可以绑定100个机器组
	标准存储类型的日志主题，保存时间可为1 - 3600天，或永久保存
	低频存储类型的日志主题，保存时间可为7 - 3600天，或永久保存
	标准存储类型的日志主题，保存时间大于7天时，才能开启日志沉降功能
指标主题	主题名称长度为1至255个字符，允许的字符为 a-z、A-Z、0-9、_、-
机器组	每个账号在不同地域下最多可以创建200个机器组
	每个 IP 机器组最多可以绑定200个机器 IP
	每个标识机器组最多可以绑定50个 Label
仪表盘	每个账号最多创建100个仪表盘订阅任务

日志采集
Kafka 数据订阅

最近更新时间：2024-06-24 11:35:42

配置限制

限制	说明
Kafka 数据订阅任务	单个日志主题最多可以创建100个 Kafka 数据订阅任务。

采集限制

限制	说明
Kafka Topic	单个 Kafka 数据订阅任务最多可以消费100个 Kafka Topic。
单条日志大小	单条日志不可超过1MB。单条日志大小超过限制时，该日志会被丢弃。
单条 Kafka 消息	单条 Kafka 消息中最多只可包含一条日志，若超出整条消息将被作为解析失败日志处理。
起始位置	仅支持指定最早和最晚位置。不支持从指定时间的位置开始导入。

性能限制

限制	说明
单个任务最大吞吐	单个任务最大吞吐不超过240MB/s。

Kafka 协议上传

最近更新时间：2024-08-12 10:22:31

写入限制

日志写入限制及说明如下表所示：

限制	说明
上传日志	单次上传请求的日志压缩前不能超过5MB，建议控制在3MB以内。

LogListener 限制说明

最近更新时间：2025-04-28 16:29:12

本文简介 Loglistener 采集数据时在文件采集、配置、资源、性能、错误处理等方面的能力与限制，以及相关使用说明。

文件采集限制

项目	能力与限制
文件编码	支持 UTF8、GBK 格式编码。 **⚠ 注意：** GBK 编码格式需要 Loglistener 2.6.2及以上版本。
软链接	支持。
单条日志大小	• 单行日志大小限制为512KB，若日志超过512KB后，会截断只保留前512KB。 • 多行日志按行首正则表达式划分后（判断首行时会读取1MB数据进行判断），单条日志最大限制为512KB。若日志超过512KB后，会截断只保留前512KB。
正则表达式	正则表达式类型支持 Perl 兼容正则表达式。
首次日志采集行为	Loglistener 支持全量采集/增量采集策略： • 全量采集：首次安装启动 Loglistener 后，会采集所有符合条件的日志，包括已经没有写入的文件。 • 增量采集：首次安装启动 Loglistener 后，存量文件会从最新位置开始采集。 **⚠ 注意：** 增量/全量采集需要 Loglistener 2.6.2及以上版本。
日志文件轮转	支持。（推荐轮转后的文件名，不要被采集通配路径覆盖到） **⚠ 注意：** 文件轮转后60秒内，若无新增日志写入轮转后的日志文件，轮转后的日志文件将不会再被采集，否则将持续采集轮转后的日志文件。
日志解析失败时采集行为	推荐开启解析失败上传功能，开启后将会把解析失败的日志按照单行全文格式上传至预设索引中。否则，解析失败日志将被丢弃。

文件打开行为	Loglistener 会在采集读取文件时打开，读完关闭。
过滤规则	单、多行全文采集最多可配置1条过滤规则。
	分隔符、正则提取模式、JSON 格式最多可以配置5条过滤规则。
Key 字段长度	分隔符、正则提取模式 Key 字段长度最多为128Byte。
Key-Value 组个数	所有解析模式下，提取 Key-Value 组的个数最多为1000个，超过部分将会被丢弃。
LogListener 最大连接数	LogListener 最大连接数限制：1024。
内存占用	LogListener 的内存占用，正常情况最多50MB。
CPU占用	LogListener 的 CPU 使用在5MB/s日志量下，业务进程合计不超过单核20%。
采集配置生效延时	更新采集配置之后会有1分钟的生效延时。
日志文件	LogListener 运行的采集目标日志文件大小建议小于500MB。

Checkpoint 管理

项目	能力与限制
Checkpoint 保存位置	保存路径默认为 Loglistener 安装目录下 data 目录下。若需修改，可参见 配置 LogListener，修改 LogListener 配置文件中的 checkpoint_cache_file，修改 Checkpoint 保存位置。
Checkpoint 保存策略	Loglistener 存有两份 checkpoint 元数据： • 一份只记录已上传完成的位点信息，实时持久化到磁盘上。 • 一份同时记录了已完成位点信息与已采集但尚未完成的位点信息，周期性持久化到磁盘上。程序退出时也会优先进行持久化。

资源、性能限制

项目	能力与限制
默认 CPU 核数限制	LogListener 仅支持单核运行。
默认 CPU 资源限制	LogListener 默认未做 CPU 资源限制，若需限制，可参见 配置 LogListener，修改 LogListener 配置文件中的 cpu_usage_thres 最大 CPU 利用率。目前

	Loglistener 的实现架构，如未作 CPU 资源限制，其最高能够达到的单核 CPU 使用率为110%左右（业务线程最大100%， 管控线程10%左右）。
默认线程数限制	LogListener 2.x.x版本仅支持单线程运行。多线程运行需 提交申请 使用 LogListener 3.x.x版本
默认内存资源限制	Loglistener 默认内存阈值设置为2G，若需限制，可参见 配置 LogListener，修改 LogListener 配置文件中的 max_mem 修改最大内存占用。建议不低于300M。
默认带宽资源限制	Loglistener 默认未限制带宽资源，若需限制，可参见 配置 LogListener，修改 LogListener 配置文件中的 max_send_rate，对程序使用的网络带宽进行限制。
资源超限处理策略	若 Loglistener 占用相关 CPU 和内存资源超过最大限制的时间超过5分钟，则采集程序会强制自动重启。
日志压缩	采集日志默认会压缩后上传，若不需要压缩，可参见 配置 LogListener，修改 LogListener 配置文件中的 request_compression 关闭压缩上传。
监控目录数	默认最大监控目录数5000，如超出可能会引起采集失败。
监控文件数	推荐最大监控文件数10000，如超出可能会引起采集失败。
监控事件数	Polling 模式下，推荐最大事件数10000，如超出可能会引起采集失败。

错误处理

项目	能力与限制
网络错误处理	非需要特殊处理的异常（如日志主题删除），其它错误都会进行重试（网络异常、超时、频控、欠费等）。
超时最大重试时间	若数据持续发送失败超过1小时，则丢弃该数据。 默认行为是间隔重试，且重试间隔时间会越来越长，直至超过超时最大尝试时间。
重试次数	可通过 loglistener.conf 配置文件设置，默认不配置。此时默认会一直重试，直至超过最大重试时间，之后丢弃。 如果配置了重试次数，则按照重试次数进行重试，超过最大次数则丢弃。
换行符缺失	采集时，若单条日志结尾缺少换行符， 采集器将等待五分钟， 若五分钟内没有检测到换行符， 将自动填充换行符，并继续往下采集

文件采集规则

项目	能力与限制

日志上传策略	Loglistener 会将同一文件的日志自动聚合上传，聚合条件为：10000条日志、日志集合总大小达到1M或者日志采集时间超过3秒，任一条件满足则触发上传行为。
文件采集的处理策略	单个目标文件（采集路径所能匹配到的一个文件）只支持上传日志到一个日志主题中，不支持多个 topic 的采集路径覆盖到同一个文件。如果想要将一个文件上传到多个日志主题中，可以通过软链接进行；对同一个目标文件创建多个软链接，不同的日志主题采集不同的软链接。
日志采集延迟	实时采集情况下，会在1分钟之内完成数据采集、传输、存储落盘，达到控制台可检索的效果。 如果日志生产量巨大，或者将采集程序使用的资源限制的较小，会有一定的采集延迟。

日志主题相关规则

项目	能力与限制
采集配置数量限制	一个日志主题可关联的采集配置数量上限为100个

机器组相关规则

项目	能力与限制
机器组逻辑	目前机器组分为两类，其使用方法是相互独立的，且两种用法是不兼容的，如果混合使用，采集机器将拉取不到正确的采集配置，造成不采集的现象。 • IP 机器组，机器 IP 需要在控制台上手动加入机器组，对应机器上 loglistener.conf 的 group_label 需为空。 • 标签机器组，控制台设置机器组标签，对应机器上 loglistener.conf 的 group_label 需要设置为相同的标签。
机器组与日志主题的关系	• 单个日志主题，可以绑定多个机器组。 • 单个机器组，可以绑定到多个日志主题上。
机器组与采集机器的关系	• 单个采集机器，可以加入到多个机器组中。 • 对于 IP 机器组，采集机器加入的机器组数量不受限制。 • 对于标签机器组，采集机器加入的机器组数量上限为20个。
标签机器组 label 限制	• 标签机器组的 label，目前限制为32字符。 • 标签机器组，单个机器组最多可设置20个 label。

采集路径/采集黑名单用法

项目	能力与限制

采集黑名单	此功能是用来指定采集路径下，需要忽略采集的内容，目前采集黑名单分为两类： • 文件路径模式：采集路径下，需要忽略采集的完整文件路径，支持通配模式。 • 目录路径模式：采集路径下，需要忽略采集的目录前缀，支持通配模式。 **⚠ 注意：** • FILE/PATH 模式可以同时使用。 • 采集黑名单是在采集路径下进行排除，因此无论是文件路径模式，还是目录路径模式，其指定路径要求为采集路径的子集。

API/SDK 日志上传限制

最近更新时间：2024-10-24 20:59:22

写入限制

日志写入限制及说明如下表所示：

限制项	说明
上传日志	一个 pb 包里 logGroup 不能超过10个。
	一个 pb 包里 logGroup 包含日志条数最多为10000条，至少包含1条。
	log 中单个 value 最大为1MB。
	一个 pb 包里 logGroup 部分所有 Value 大小最大总和为5MB。
	Log 中的 Key 不能是以下保留字段，且不能以 _ 开头 "__FILENAME__"、"__TIMESTAMP__"、"__LOGSETID__"、"__TOPICID__"。
	单次上传请求的包体压缩前不能超过6MB。
频控	单个主题分区写请求限制：500QPS。超限返回状态码429，提示错误 SpeedQuotaExceed 。写频率较大时，建议 分裂主题分区 。 **说明：** 此处500QPS不代表500条日志。 一次请求中的 pb 包可以包含多条日志批量上传。
流控	单个主题分区压缩后写流量限制：5MB/s。超限返回状态码429，提示错误 SpeedQuotaExceed。写流量较大时，建议 分裂主题分区 。

COS 数据导入规格限制

最近更新时间：2024-08-30 11:44:41

控制台限制

限制项	说明
文件预览限制	仅支持预览小于1G 的文件

导入限制

限制项	说明
单个文件大小	单个文件大小不可超过5G。若超过限制，导入任务将忽略整个文件
单条日志大小	• 单行日志大小限制为512KB，若日志超过512KB 后，会截断只保留前512KB。 • 多行日志按行首正则表达式划分后，单条日志最大限制为1M。
正则表达式	正则表达式类型支持 Perl 兼容正则表达式。
日志解析失败时采集行为	推荐开启解析失败上传功能，开启后将会把解析失败的日志按照单行全文格式上传至预设索引中。否则，解析失败日志将被丢弃。
Key 字段长度	分隔符、正则提取模式 Key 字段长度最多为128Byte。
Key-Value 组个数	所有解析模式下，提取 Key-Value 组的个数最多为1000个，超过部分将会被丢弃。

指标采集

最近更新时间：2024-10-18 10:20:11

使用指标主题采集指标时，规格与限制如下：

限制项	说明
指标名称（metric name）	支持英文字母、数字、下划线及冒号，即需符合正则表达式 `[a-zA-Z_:][a-zA-Z0-9_:]*`
维度名称（label name）	支持英文字母、数字、下划线，即需符合正则表达式 `[a-zA-Z_][a-zA-Z0-9_]*`
维度值（label value）	无特殊限制，支持各类 Unicode 字符
样本值（sample value）	float64类型数值
样本时间戳（sample timestamp）	毫秒精度
指标上传频控	与分区数量有关，最大分区数量为50，单个主题分区写请求限制：500QPS。建议在 创建指标主题 时启用分区自动分裂。
指标上传流控	与分区数量有关，最大分区数量为50，单个主题分区写流量限制：5MB/s。建议在 创建指标主题 时启用分区自动分裂。

检索分析

最近更新时间：2024-08-26 10:35:11

日志主题

日志主题检索分析规格限制如下表：

分类	限制项	说明
索引	字段数量	单个日志主题键值索引最多可添加300个字段
	字段名称	• 支持除 `*\",` 外的所有英文字母、符号及数字，不能以 `_` 开头(`__CONTENT__` 字段例外) • 不允许同时包含 JSON 父子级字段，例如 `a` 及 `a.b`
	字段层级	为多层级 JSON 配置键值索引时，Key 的层级不能超过10层，例如 a.b.c.d.e.f.g.h.j.k
	分词符	仅支持英文符号、\n\t\r及转义符\
	字段长度	字段开启统计后，仅前32766个字节参与 SQL 运算，超出部分无法执行 SQL，但仍会完整存储日志
	分词长度	分词后，单个词仅前10000个字符参与检索，超出部分无法执行检索，但仍会完整存储日志
	数值字段精度及范围	• long 类型字段支持数据范围为-1E15 ~ 1E15，超出该范围的数据可能会丢失精度或不支持检索 • double 类型字段支持数据范围为-1.79E+308 ~ +1.79E+308，如果浮点数编码位数超过64位，会造成精度丢失 超长数值字段索引配置建议： • 如果不需要通过数值范围比较来检索该字段，可将该字段存储为 text 类型 • 如果需要通过数值范围比较来检索该字段，可将该字段存储为 double 类型，可能会丢失部分精度
	生效机制	索引配置仅针对新采集的数据生效，索引规则编辑后仅对新写入的日志生效，已有数据不会更新。如需更新已有数据，需重建索引
	修改索引配置	创建、修改及删除索引配置时，单个用户最大同时拥有10个执行中的任务，超出后需等待之前的任务执行完毕，单个任务执行一般不超过1分钟
	重建索引	• 单个日志主题同时仅允许运行一个重建索引任务，单个日志主题最多同时拥有10个重建索引任务记录，需删除不再需要的任务记录后才能新建索引任务

		• 同一时间范围内的日志，仅允许重建一次索引，需删除之前的任务记录后才能再次重建 • 所选时间范围对应日志写流量不能超出5TB • 重建索引时间范围以日志时间为准，日志上传时间与重建索引时间范围有超过1小时的偏差时（例如16:00上传了一条02:00的日志到 CLS，重建00:00~12:00的日志索引）不会被重建且后续无法进行检索。新上报一条日志到已经被重建的日志时间范围时，也不会被重建且后续无法进行检索
查询	语句长度	检索分析语句最长支持12000字符
	查询并发	单个日志主题支持15并发，包括检索及分析
	模糊检索	不支持前缀模糊检索，例如通过 `*rror` 检索 error
	短语检索	短语检索中的通配符仅能匹配到符合条件的128个词，返回包含这128个词的所有日志，指定的词越精确，查询结果越精确
	逻辑分组嵌套深度	CQL 语法规则下使用括号对检索条件进行逻辑分组时，最多可嵌套10层，Lucene 语法规则无该限制 例如 `(level:ERROR AND pid:1234) AND service:test` 为2层嵌套，可正常执行检索；而下面的语句为11层嵌套，执行检索将会报错 `status:"499" AND ("0.000" AND (request_length:"528" AND ("https" AND (url:"/api" AND (version:"HTTP/1.1" AND ("2021" AND ("0" AND (upstream_addr:"169.254.128.14" AND (method:"GET" AND (remote_addr:"114.86.92.100"))))))))))`
	内存占用量（分析）	• 统计分析每次所占用的服务端内存不能超过3GB • 通常在使用 group by、distinct()、count(distinct()) 时可能触发该限制，是由于被统计的字段在通过 group by 或 distinct() 去重后值过多导致的。建议优化查询语句，使用值更少的字段对数据进行分组统计；或使用 approx_distinct() 替代 count(distinct())
	查询结果	查询结果为原始日志时，单次最大返回1000条原始日志
		查询结果为统计分析结果时，单次默认返回100条结果，使用 SQL LIMIT 语法 单次可最大返回100万条结果
		查询结果返回数据包最大49MB，使用 API 时可启用 gzip 压缩（Header Accept-Encoding:gzip）
	超时时间	单次查询超时时间为55秒，包括检索及分析
	查询延迟	从日志上报到可被检索分析的延迟小于1分钟
下载	日志条数	单次下载日志条数最多为5000万条

	任务数量	• 单个日志主题最多2个任务处于“文件生成中”状态，其它未完成的任务处于“等待中”状态，排队等待执行 • 单个日志主题最多同时存在1000个任务，包含处于“文件已生成”状态的已完成的任务
	文件保存时长	生成的日志文件仅保留3天
关联外部数据	数量限制	单个日志主题最多可关联20个外部数据
	查询超时	查询外部数据库时，超时时间为50秒
	MySQL 版本	兼容 MySQL 5.7, 8.0 及更高版本。MySQL 5.6 未经过完整兼容性测试，需实际试验 SQL 语句执行是否正常
	CSV 文件大小	不能超过50MB，不支持压缩

指标主题

指标主题检索分析规格限制如下表：

限制项	说明
查询并发	单个指标主题支持15并发
查询数据量	单次查询涉及 time series 不超过20万，查询结果中单个 time series 最大1.1万个时间点
超时时间	单次查询超时时间为55秒，包括检索及分析

监控告警

最近更新时间：2025-04-15 10:48:44

监控告警规格限制如下表所示：

分类	限制项	说明
告警策略	监控对象	单个告警策略最多可选择20个日志主题，不支持跨地域。
	执行语句	限制同 检索分析 ，包括语句长度、并发限制等。
	执行语句数量	• 各个执行语句使用相同日志主题时，单个告警策略最多可添加3个执行语句。 • 各个执行语句单独选择日志主题时，单个告警策略最多可添加10个执行语句。
	执行语句查询时间范围	最大时间范围为最近24小时。
	执行周期	最小执行周期为1分钟，最大执行周期为1440分钟。
	分组触发	最多可将执行语句结果分为1000个组，对这1000个组分别判断是否满足触发条件，超出范围的组将不执行判断。 **说明：** 分组触发为可选功能，未启用时不涉及该限制。
	触发条件判断次数	当执行语句结果为多个值时（分组后），会根据结果数据依次判断是否满足触发条件，当满足触发条件或超过1000次后停止判断。
告警渠道	电话告警	• 同一号码30秒内最多接收1条电话告警，10分钟内最多接收3条电话告警，1天内最多接收50条电话告警。 • 同一号码22:00至次日08:00最多接收3条电话告警，如需去除该限制，请在 消息中心 中开启**夜间接收状态**开关。 • 以上限制不局限于日志服务产生的电话告警通知，腾讯云可观测平台及消息中心产生的电话通知会共享该限制。 • 使用电话接收告警时，建议您将以下电话号码添加至手机及运营商骚扰电话拦截白名单中，以免被拦截：

		02155931000,02155931001,02155931002,02155931003,02155931004,02155931005,02155931006,02155931007,02155931008,02155931009,02155931010,02155931011,02155931012,02155931013,02155931014,02155931015,02155931016,02155931017,02155931018,02155931019,02155931020,02155931021,02155931022,02155931023,02155931024,02155931025,02155931026,02155931027,02155931028,02155931029,02155931030,02155931031,02155931032,02155931033,02155931034,02155931035,02155931036,02155931037,02155931038,02155931039,02155931040,02155931041,02155931042,02155931043,02155931044,02155931045,02155931046,02155931047,02155931048,02155931049,02155931050,02155931051,02155931052,02155931053,02155931054,02155931055,02155931056,02155931057,02155931058,02155931059,01021364689,01021364720
	企业微信、钉钉、飞书	每个机器人发送的消息不能超过20条/分钟。
	自定义接口回调	• 须支持公网访问，不支持内网地址。 • 接口返回状态码为200时表示回调成功。
告警历史	存储时长	30天

告警通知	告警通知变量	以下 告警通知变量 存在数据条数或大小限制，超出限制后将自动截断：
		• {{.Message}}：最大10KB。
		• {{.QueryResult}}：每个执行语句最多1000条数据、最大10KB，最多10个执行语句，总计最大100KB。
		• {{.TriggerResult}}：每个语句最多1000 条数据、最大10KB，最多10个执行语句，总计最大100KB。
		• {{.QueryLog}}：每个执行语句最多10条数据、最大50KB，最多10个执行语句，总计最大150KB。
		• {{.AnalysisResult}}：每个多维分析最多10KB，最多5个多维分析，总计最大50KB。
		  ○ 相关原始日志：每个语句最多10条数据，以用户配置为准。
		  ○ 字段Top5及占比统计：每个语句最多5条数据。
		  ○ 自定义检索分析（SQL）：每个语句最多1000条数据。

		○ 自定义检索分析（检索）：每个语句最多10条数据。

数据处理

最近更新时间：2024-12-13 11:59:42

数据处理规格限制如下表所示：

分类	限制项	说明
数据加工	数据加工任务数	1个日志主题可以创建50个数据加工任务。
	目标主题个数	1个数据加工任务最多可以有1000个目标主题。
	调试数据条数	默认加载100条源日志主题中的日志，用于您的数据加工语句调试。
	处理能力	源日志主题的每个分区可处理 10 – 20MB(非压缩)/秒的数据。
	加工任务时延	99.9%的数据将在1秒内在数据加工模块内处理完成。
	加工任务启停对数据的影响	加工任务暂停后12小时内重启，从暂停时的数据开始加工。超出12小时重启，从最新数据开始加工。
	跨地域/跨账号	暂不支持。
	加工历史数据	暂不支持。
定时 SQL 分析	定时 SQL 任务数	1个日志主题可以和10个定时 SQL 任务相关（该日志主题可以是任务的源日志主题，或者目标日志主题，都计算在内）。
	查询并发	单个日志主题支持15并发（并发包含 定时 SQL 、检索分析、仪表盘、告警）。
	查询结果	查询结果为统计分析结果时，单次默认返回100条结果，使用 SQL LIMIT 语法，单次可最大返回100万条结果。
		查询结果返回数据包最大49MB，使用 API 时可启用 gzip 压缩。
	查询超时时间	单次查询超时时间为55秒。
	SQL 时间窗口	1分钟 – 7天。
	调度周期	1分钟 – 24小时。
	自定义时间戳	最小自定义时间戳的间隔为15秒，例如15:00:00，15:00:15，15:00:30。
	延迟执行 SQL	默认60秒，最大为120秒。99.9%的日志索引需要5秒左右生成，极端情况60秒，因此配置延迟执行，确保该查询可以成功获取索引

		数据。
	跨地域/跨账号	暂不支持。

投递与消费

最近更新时间：2024-12-13 11:59:42

相关限制说明如下表所示：

分类	限制项	说明
投递 COS	投递任务数	每个日志主题投递 COS 任务配置个数最多为10个。
	可投递的数据范围	日志主题的生命周期内的数据（包含历史数据和实时数据）。
	投递任务时延	投递 COS 的时延在60秒左右。 **说明：** 假如您配置投递时间间隔为5分钟，那么12:00 - 12:05的日志，满足了5分钟间隔的条件，这部分日志会作为一个批次，整体投递到 COS，时延为60秒，那么这个批次的日志将在12:06左右到达 COS。
	投递触发	单个分区满足以下条件任意一个时，触发一次投递： • 投递日志大小（累计日志数据），您可配置该参数，5-256MB（未压缩）。 • 投递时间间隔，5-15分钟，您可配置该参数，5-15分钟。
	跨账号/跨地域	• 跨账号：支持 • 跨地域：暂不支持
投递 CKafka	投递任务数	每个日志主题可配置一个 CKafka 投递任务。
	可投递的数据范围	实时日志数据，暂不能投递历史数据。
	投递任务时延	99.9%的日志投递 CKafka 的时延在5秒内。（目标 CKafka 状态正常，且吞吐能力足够处理来自 CLS 的消息，不存在磁盘写满或者限流的情况）。 **说明：** 某一条日志采集到 CLS 的时间戳（_TIMESTAMP_字段的值）为12:00:00，该条日志将在12:00:05秒之前到达 CKafka。

	跨账号/跨地域	• 跨账号：支持 • 跨地域：暂不支持
Kafka 协议消费	版本	支持 Kafka 协议版本为：1.1.1及更早的版本。
	消费组个数	每个日志主题可配置20个消费组。
	认证方式	仅支持 SASL_PLAINTEXT。
	可消费的数据范围	开启该功能后，默认将从最新数据开始消费。消息保留时间2小时，超过2小时还未消费的数据将不会被保存。 **说明：** 例如您12:00打开 Kafka 协议消费，那么您将从12:00的数据开始消费，到15:00，您可以消费13:00-15:00的数据，13:00之前的数据不能消费。
	消费时延	99.9%的日志消费的时延在5秒内。 **说明：** 某一条日志采集到 CLS 的时间戳（_TIMESTAMP_字段的值）为12:00:00，该条日志将在12:00:05秒之前到达消费端。
	跨账号/跨地域	• 跨账号：A 账号可以作为 B 账号的协作者，消费 B 账号的日志主题。需对协作者设置权限，详情请参见 权限配置 。 • 跨地域：通过外网消费实现跨地域消费。

基本概念

日志

最近更新时间：2024-10-18 10:20:11

日志

日志（Log）是应用系统运行过程中产生的记录数据，如用户操作日志、接口访问日志、系统错误日志等。日志通常以文本的形式存储在应用系统所在的机器上，一条系统运行记录对应的日志可能为一行文本（单行日志），也可能为多行文本（多行日志）。
日志可通过 LogListener 上传至日志服务（Cloud Log Service，CLS），也可以通过 API、SDK 等其他方式上传。
单行日志示例：

```
59.x.x.x - - [06/Aug/2019:12:12:19 +0800] "GET /nginx-logo.png HTTP/1.1"
200 368 "http://119.x.x.x/" "Mozilla/5.0 (Macintosh; Intel Mac OS X
10_14_5) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/75.0.3770.142
Safari/537.36" "-"
```

多行日志示例：

```
  java.net.SocketTimeoutException:Receive timed out
    at j.n.PlainDatagramSocketImpl.receive0(Native Method)[na:1.8.0_151]
    at
j.n.AbstractPlainDatagramSocketImpl.receive(AbstractPlainDatagramSocketI
mpl.java:143)[^]
    at j.n.DatagramSocket.receive(DatagramSocket.java:812)[^]
    at o.s.n.SntpClient.requestTime(SntpClient.java:213)[classes/]
    at o.s.n.SntpClient$1.call(^:145)[^]
    at ^.call(^:134)[^]
    at o.s.f.SyncRetryExecutor.call(SyncRetryExecutor.java:124)[^]
    at o.s.f.RetryPolicy.call(RetryPolicy.java:105)[^]
    at o.s.f.SyncRetryExecutor.call(SyncRetryExecutor.java:59)[^]
    at o.s.n.SntpClient.requestTimeHA(SntpClient.java:134)[^]
    at ^.requestTimeHA(^:122)[^]
    at o.s.n.SntpClientTest.test2h(SntpClientTest.java:89)[test-
classes/]
    at s.r.NativeMethodAccessorImpl.invoke0(Native Method)[na:1.8.0_151]
```

对于一条日志，其主要组成部分如下：

分类	字段	含义	示例
内置保留字段	__TIMESTAMP__	日志时间戳（毫秒级别 Unix 时间戳），按时间范围检索日志时，将自动使用该时间对日志进行检索，在控制台显示为“日志时间”	1640005601188
	__FILENAME__	日志采集的文件名	/data/log/nginx/access.log
	__SOURCE__	日志采集的源 IP	10.0.1.2
	__HOSTNAME__	日志来源机器名称，需使用2.7.4及以上版本的 Loglistener 才会采集该字段	localhost
	__RAWLOG__	日志创建索引过程中，如果原始日志格式异常或索引配置与原始日志不匹配，可能会导致索引创建异常，此时 CLS 为了确保日志不丢失，会将原始日志存储在该字段中，作为兜底的异常处理方式	{"ip":"10.20.20.10","request":"GET /online/sample HTTP/1.1","status":200,"time":"[2018-07-16 13:12:57]"}
	__INDEX_STATUS__	日志创建索引异常原因，值非空时代表当前日志创建索引出现异常，原始日志被存储在__RAWLOG__中，__INDEX_STATUS__为异常原因	IndexPartFail:part of fields do not match the index configuration
日志正文	__CONTENT__	日志提取模式为单行全文或多行全文（即原始日志不经切分，直接将整条日志上报）时，整条日志存储在该字段中	10.20.20.10;[2018-07-16 13:12:57];GET /online/sample HTTP/1.1;200
	普通字段	日志采集时对日志进行 结构化 后，原始日志以 `key:value` 的形式存储	IP: 10.20.20.10 request: GET /online/sample HTTP/1.1 status: 200 time: [2018-07-16 13:12:57]
元数据	元数据字段	对日志的简单描述或归类，例如 TKE 日志中，某条日志所属的集群或容器，以 `key:value` 的形式存储，key 以 __TAG__.开头	__TAG__.clusterId:1skzv59c

日志组

日志组（LogGroup）是一个包含多条日志的集合。在上传日志的过程中，为提高数据读写效率，将多条日志打包成一个日志组，并以日志组为单位上传到 CLS。
一个日志组里的日志具有相同的基本信息（__TIMESTAMP__、__FILENAME__、__SOURCE__、__HOSTNAME__ 和元数据）。

指标

最近更新时间：2024-08-07 16:59:31

指标（Metric）是用来衡量系统和应用程序性能及运行情况的度量值，例如 CPU 利用率、内存使用率、访问吞吐量、响应耗时和响应成功率等。指标一般会定时产生，每一个时刻都会有一个值，随着时间变化形成一个序列，这个序列一般被称为时间序列（time series），简称时序。
日志服务（Cloud Log Service，CLS）兼容 Prometheus 指标数据模型，将相同指标名称（metric name）、相同维度（labels）的包含时间戳的指标数据保存为时间序列（time series）。在时间序列中每个数据点称为样本（sample），样本则由时间戳及样本值构成。
示例
例如系统在2020/12/30 15:35:23.123时某一接口的请求总量就是一个样本，其数据如下：

```
requests_total{method="POST", handler="/messages"} 217
```

其由如下几部分构成：

- 指标名称：requests_total
- 维度：{method="POST", handler="/messages"}，即接口名称为 messages，请求方式为 POST
- 时间戳：2020/12/30 15:35.123
- 样本值：217

被监控系统实际上往往同时有多个指标，其在某一时刻具备大量不同的指标名称和维度，例如 Nginx 监控指标如下：

```
# HELP nginx_http_requests_total The total number of HTTP requests
# TYPE nginx_http_requests_total counter
nginx_http_requests_total 10234

# HELP nginx_http_requests_duration_seconds The HTTP request duration in
seconds
# TYPE nginx_http_requests_duration_seconds histogram
nginx_http_requests_duration_seconds_bucket{le="0.005"} 2405
nginx_http_requests_duration_seconds_bucket{le="0.01"} 5643
nginx_http_requests_duration_seconds_bucket{le="0.025"} 7890
nginx_http_requests_duration_seconds_bucket{le="0.05"} 9234
nginx_http_requests_duration_seconds_bucket{le="0.1"} 10021
nginx_http_requests_duration_seconds_bucket{le="0.25"} 10234
nginx_http_requests_duration_seconds_bucket{le="0.5"} 10234
nginx_http_requests_duration_seconds_bucket{le="1"} 10234
nginx_http_requests_duration_seconds_bucket{le="2.5"} 10234
nginx_http_requests_duration_seconds_bucket{le="5"} 10234
```

```
nginx_http_requests_duration_seconds_bucket{le="10"} 10234
nginx_http_requests_duration_seconds_bucket{le="+Inf"} 10234
nginx_http_requests_duration_seconds_sum 243.56
nginx_http_requests_duration_seconds_count 10234

# HELP nginx_http_connections Number of HTTP connections
# TYPE nginx_http_connections gauge
nginx_http_connections{state="active"} 23
nginx_http_connections{state="reading"} 5
nginx_http_connections{state="writing"} 7
nginx_http_connections{state="waiting"} 11

# HELP nginx_http_response_count_total The total number of HTTP
responses sent
# TYPE nginx_http_response_count_total counter
nginx_http_response_count_total{status="1xx"} 123
nginx_http_response_count_total{status="2xx"} 9123
nginx_http_response_count_total{status="3xx"} 456
nginx_http_response_count_total{status="4xx"} 567
nginx_http_response_count_total{status="5xx"} 65

# HELP nginx_up Is the Nginx server up
# TYPE nginx_up gauge
nginx_up 1
```

其中指标含义如下：

- nginx_http_requests_total：Nginx 处理的 HTTP 请求总数。
- nginx_http_requests_duration_seconds：HTTP 请求的持续时间，使用 Histogram 类型提供不同区间的请求数量。
- nginx_http_connections：当前 Nginx 的 HTTP 连接数，分为 active、reading、writing 和 waiting 状态。
- nginx_http_response_count_total：Nginx 返回的 HTTP 响应总数，按状态码分类。
- nginx_up：Nginx 服务器的运行状态，1表示运行中，0表示未运行。

日志（Log）和指标（Metric）都是系统和应用程序中收集和记录数据的方法，但它们在数据的类型和用途上有所不同，在 CLS 中区别主要如下：

数据类型	日志（Log）	指标（Metric）
数据格式	无严格限制，普通文本及采用 json 格式记录的结构化数据均可作为日志上报	需符合 Prometheus 指标数据模型

数据采集	支持使用 LogListener， API 和 SDK 采集日志	需通过定时 SQL 分析将原始日志转换为指标，详见 日志转指标
数据存储	支持 标准存储、低频存储 及 日志沉降，不同存储类型功能、性能及价格不同	不区分存储类型
索引配置	需 配置索引，包括字段名称、类型、分词符等	不需要配置索引
数据查询	使用 检索条件 过滤原始日志，使用 SQL 对原始日志进行统计分析	使用 PromQL 查询指标

日志/指标主题与日志集

最近更新时间：2024-10-16 16:03:11

日志主题

日志主题（Log Topic）是日志数据在日志服务（Cloud Log Service，CLS）平台进行采集、存储、检索和分析的基本单元，采集到的海量日志以日志主题为单元进行管理，包括采集规则配置、保存时间配置、日志检索分析以及日志下载/消费/投递等。
一个日志主题通常对应某一个应用/服务，建议将同一个应用/服务在不同机器上的同类日志采集到同一个日志主题。例如，某支付服务（payService）部署在数十台机器上，包含访问日志（access_log）和错误日志（error_log）两类日志。可创建 payService_access_log_topic 和 payService_error_log_topic 两个日志主题，分别对应这数十台机器上的两类日志，通过这两个日志主题即可完成数十台机器上的所有日志的集中检索和分析。
日志主题与应用/服务之间并非严格的一一对应关系，如果两个服务之间的日志结构相似度较高，且经常需要集中分析日志，也可以将这两个服务的日志上报至同一日志主题下。

指标主题

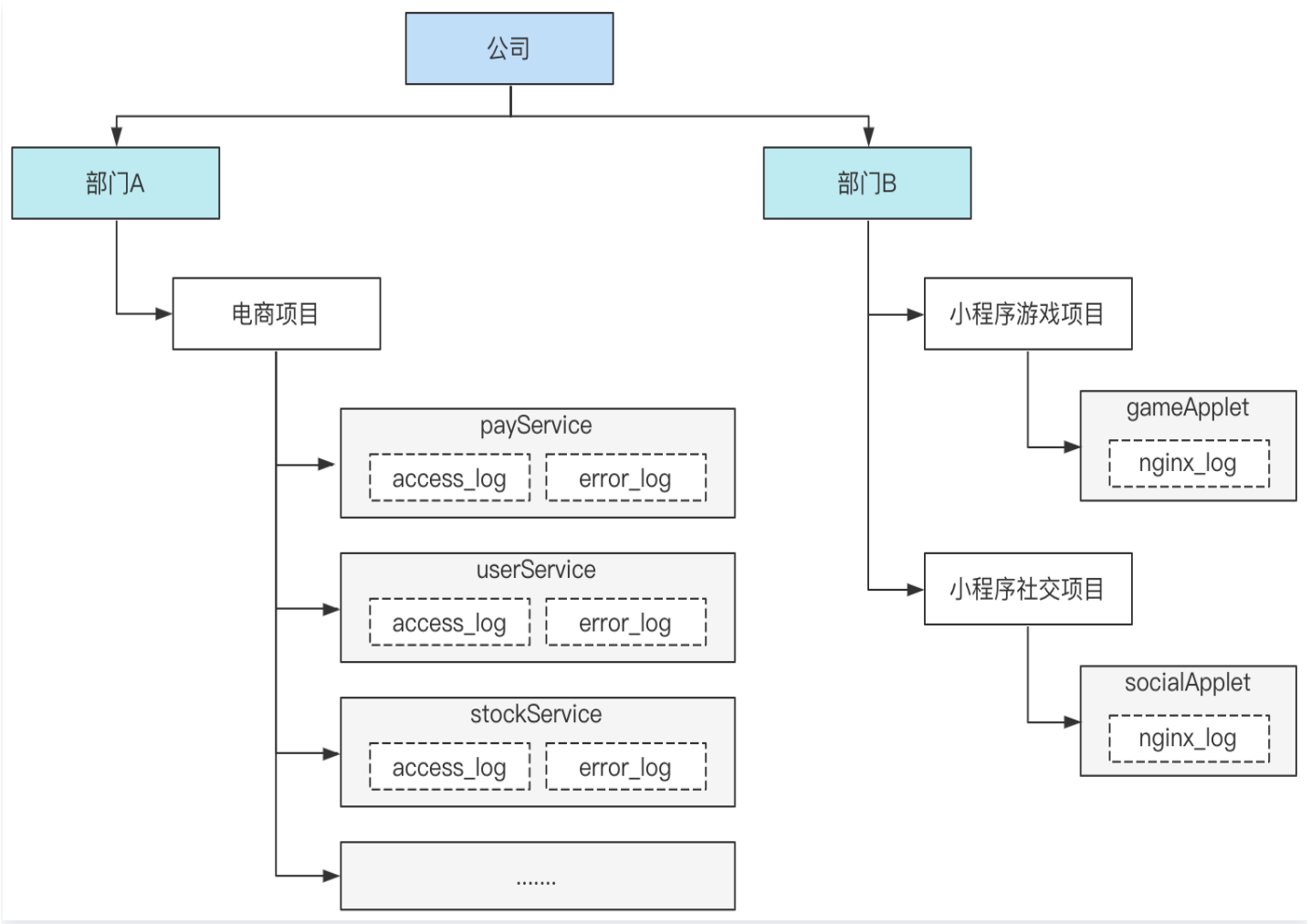
指标主题（Metric Topic）是指标数据在日志服务（Cloud Log Service，CLS）平台进行采集、存储、检索和分析的基本单元，采集到的指标数据以指标主题为单元进行管理，包括保存时间配置及检索分析等。
指标主题兼容 Prometheus 指标数据模型及指标查询接口，相当于一个 Prometheus 实例，不同应用/服务的指标只要指标名称不发生冲突，且指标上传未超过 25000QPS 及 250MB/s 的规格限制，就可以存储在一个指标主题中。实际应用中，一般将业务系统生产环境、测试环境、开发环境的指标分别存储在不同的指标主题中。

日志集

日志集（Logset）是对日志主题及指标主题的分类，一个日志集可包含多个日志主题和指标主题。日志集本身不存储任何日志数据，仅方便用户管理主题。
一个日志集通常对应公司内的某一个项目/业务，建议将某个项目/业务下的多个应用/服务的日志主题归属到同一个日志集下。例如，公司某电商项目下包含多个服务（支付服务 payService、用户服务 userService、库存管理服务 stockService 等），可创建一个日志集 e_commerce_logset，将这些服务的日志主题均归属至该日志集下。这样当公司有多个项目时，具体的项目人员只需要查看所属项目对应的日志集下的日志主题即可，其它项目的日志主题不会对其产生干扰。

> **注意**
> 新建主题时可指定其归属的日志集，保存后不能再变更。

场景示例



如上图，该公司有两个部门：

- 部门 A 有一个电商项目，采用微服务架构，每个服务均包含访问日志（access_log）和错误日志（error_log）两类日志。
- 部门 B 有两个项目，分别为小程序游戏项目和小程序社交项目，技术架构比较简单，均各有一个 Nginx 的访问日志（nginx_log）。

使用 CLS 监控上述这些应用日志时，可创建如下的日志集及日志主题：

日志集	日志主题	标签
e_commerce_logset	payService_access_log_topic	部门:部门 A
e_commerce_logset	payService_error_log_topic	部门:部门 A
e_commerce_logset	userService_access_log_topic	部门:部门 A
e_commerce_logset	userService_error_log_topic	部门:部门 A

e_commerce_logset	stockService_access_log_topic	部门:部门 A
e_commerce_logset	stockService_error_log_topic	部门:部门 A
e_commerce_logset	.....	部门:部门 A
gameApplet_logset	gameApplet_nginx_log_topic	部门:部门 B
socialApplet_logset	socialApplet_nginx_log_topic	部门:部门 B

其中的标签用来区分日志集及日志主题归属的部门，结合权限策略可以控制每个部门的人员仅可查看所属部门的数据。
对于部门 A，e_commerce_logset 日志集涵盖了其电商业务下的所有服务的日志主题，后续如果新增其他的项目，新建一个日志集即可。
对于部门 B，虽然目前的技术架构比较简单，总共只有两类日志，但却创建了两个日志集，每个日志集只有一个日志主题，是出于如下目的：

- 支持后续架构扩展：如果业务规模上升，技术架构也演变为微服务架构后，可继续沿用当前的日志集，在当前日志集下新增日志主题即可，不同的项目之间互不影响。
- 灵活应对项目调整：如果把整个部门作为一个日志集，当其中的某个项目需要独立为一个部门，或需调整至另一个部门时，由于日志主题不能直接变更其归属的日志集，调整将变得非常麻烦，可能需要重新采集日志。而每个项目分别对应一个日志集时，则不存在该情况，只需调整日志集和日志主题对应的标签即可。

机器组

最近更新时间：2024-10-16 16:03:11

机器组

机器组（MachineGroup）是一组需要采集日志的机器列表，日志服务通过机器组来管理所有需要通过 LogListener 采集日志的机器。
一个机器组可以包含多台机器。当应用/服务部署在多台机器上，且这些机器的日志文件路径相同时，便可将其归为一个机器组。这样在控制台只需要配置一次日志数据采集规则，便可批量生效至机器组内的所有机器。
机器组可关联至一个日志主题，即机器组内所有的日志均上报至同一个日志主题；也可以关联至多个日志主题，即机器组内不同路径的日志分别上报至不同的日志主题。
机器组具备两种定义方式：

- IP 地址：添加 IP 地址列表到机器组，机器组将自动添加这些IP对应的机器。
- 标识（Label）：安装 LogListener 时为所在机器添加标识，机器组将自动添加包含这些标识的机器。

场景示例

某电商项目，共有6台机器及三个服务。其部署方式如下表：
payService 部署在2台机器上，有2个日志文件路径。userService和stockService 混部在4台机器上，共有4个日志文件路径。每个服务的每类日志（访问日志 access_log 和错误日志 error_log）均需上报至单独的日志主题。

机器	部署服务	日志文件路径
192.168.1.1	payService	/data/log/payService/access_log.log /data/log/payService/error_log.log
192.168.1.2	payService	/data/log/payService/access_log.log /data/log/payService/error_log.log
192.168.1.3	userService,stockService	/data/log/userService/access_log.log /data/log/userService/error_log.log /data/log/stockService/access_log.log /data/log/stockService/error_log.log
192.168.1.4	userService,stockService	/data/log/userService/access_log.log /data/log/userService/error_log.log /data/log/stockService/access_log.log /data/log/stockService/error_log.log

192.168.1.5	userService,stockService	/data/log/userService/access_log.log /data/log/userService/error_log.log /data/log/stockService/access_log.log /data/log/stockService/error_log.log
192.168.1.6	userService,stockService	/data/log/userService/access_log.log /data/log/userService/error_log.log /data/log/stockService/access_log.log /data/log/stockService/error_log.log

在机器上部署 LogListener 时，可为每台机器按其上运行的服务添加标识，具体如下：

机器	部署服务	标识
192.168.1.1	payService	payService
192.168.1.2	payService	payService
192.168.1.3	userService,stockService	userService,stockService
192.168.1.4	userService,stockService	userService,stockService
192.168.1.5	userService,stockService	userService,stockService
192.168.1.6	userService,stockService	userService,stockService

然后在控制台创建3个机器组，采用标识方式进行定义，分别为 payService、userService 和 stockService，再在日志主题中关联该机器组并添加对应的采集配置即可完成日志上报。

日志主题	关联机器组	采集路径
payService_access_log_topic	payService	/data/log/payService/access_log.log
payService_error_log_topic	payService	/data/log/payService/error_log.log
userService_access_log_topic	userService	/data/log/userService/access_log.log
userService_error_log_topic	userService	/data/log/userService/error_log.log

stockService_access_log_topic	stockService	/data/log/stockService/access_log.log
stockService_error_log_topic	stockService	/data/log/stockService/error_log.log

后续服务需要扩容时，只需要在新增的机器上添加部署的服务作为标识，即可自动将新增的机器添加至对应的机器组并采集日志，极大的提升运维部署效率。

分词与索引

最近更新时间：2024-10-24 20:59:22

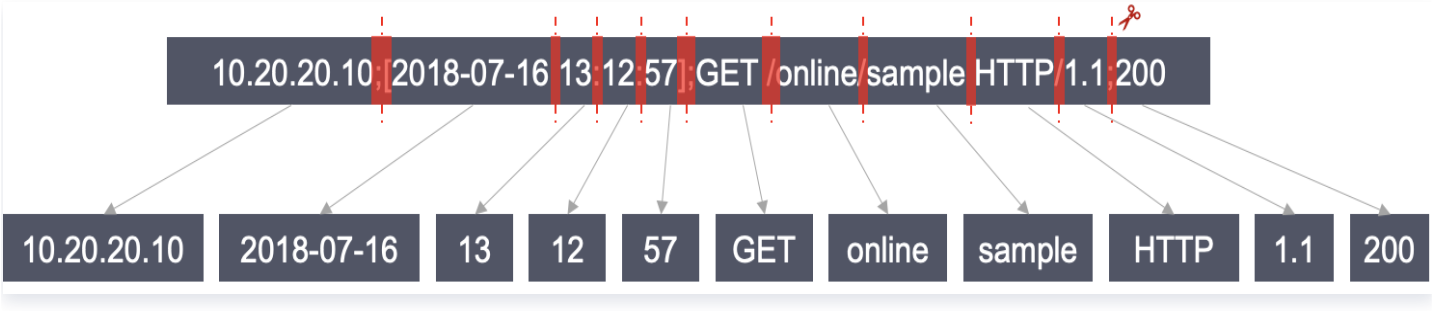
分词

分词示意

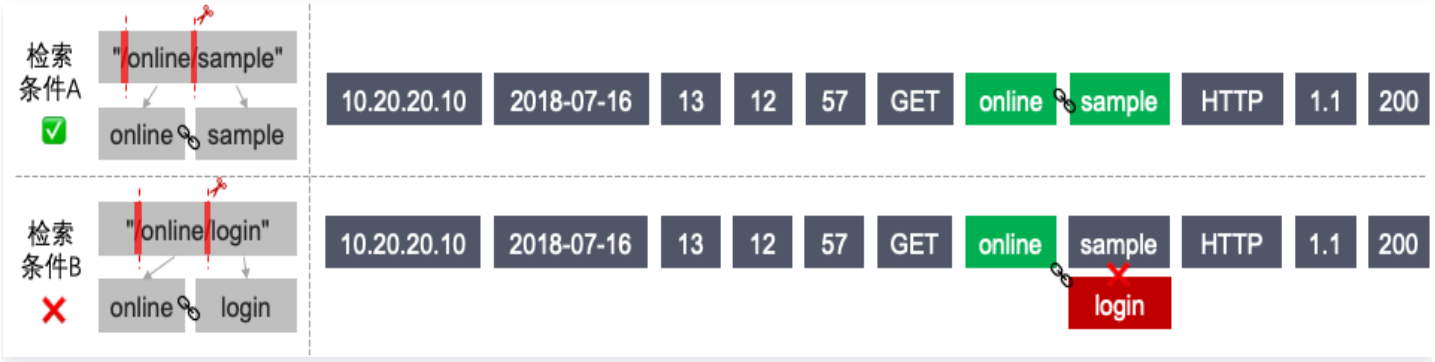
检索一个较长的日志时，通常只使用其中一部分内容进行检索。例如：需要检索出包含 `sample` 的如下完整日志：

```
10.20.20.10;[2018-07-16 13:12:57];GET /online/sample HTTP/1.1;200
```

由于日志全文很长，除 `sample` 外还有很多其他内容，其与检索条件并不是直接相等，因此不能直接使用 `sample` 作为检索条件来检索该日志。为满足该检索需求，需要将日志全文切分为多个片段，每个片段称之为一个“词”，而这个过程称之为“分词”。
例如：按符号对上述示例日志进行切分，只要出现了 `@&()='",;:<>[]{}/ \n\t\r\\` 范围内的符号，就切分日志，将得到如下的词：



检索条件为 `sample` 时，上述分词后的日志包含 `sample` ，即认为符合检索条件。
检索条件本身也会进行分词，例如下图2种检索条件：



- 检索条件A：`"/online/sample"`
  - 双引号表示日志需**同时存在**这两个词，且两个词**顺序严格一致**才符合检索条件。

- 上述示例日志同时包含 `online` 和 `sample` ，且词的顺序与检索条件一致，符合该检索条件。

- 检索条件B： `"/online/login"`

  上述示例日志不包含 `login` ，**不符合**该检索条件。

分词设置

日志分词依据包含两类，可在 索引配置 中设置。

- 分词符：可自定义需要按照哪些符号对日志进行切分，支持英文符号及\n\t\r\\。在上面的例子中，`@&()='",;:<>[]{}/ \n\t\r\\` 即为分词符。
- 是否包含中文：中文较为特殊，不能使用中文符号作为分词符，而且仅按照符号对中文进行分词也往往达不到预期效果。例如，日志为 `用户登录失败，密码错误` ，需要使用 `"登录失败"` 进行检索，是不能通过符号对日志进行切分来满足检索需求的。此时可在 索引配置 中设置该日志“包含中文”，日志服务（Cloud Log Service，CLS）将自动将日志中的每一个汉字及汉字符号切分为独立的词。

索引

CLS 对日志进行分词后，采用倒排索引来存储日志中每个“词”在文档中的位置，用户快速根据检索条件匹配具体的原始日志。倒排索引是文档检索系统中最常用的数据结构。
索引决定了日志能够以什么样的条件来进行检索和分析，开启索引后才能进行检索分析。因此在上传日志数据前，需要为日志主题设置一个合理的索引规则，配置方式参见 索引配置 。

主题分区

最近更新时间：2024-06-17 10:00:41

主题分区（Partition）是日志服务（Cloud Log Service，CLS）的最小读写单元，一个日志主题可以包含多个分区，分区数量决定了单个日志主题的最大数据读写能力。单个分区的读写能力如下表：

频控	单个主题分区读请求限制：200QPS。超限返回状态码429，提示错误 SpeedQuotaExceed。
	单个主题分区写请求限制：500QPS。超限返回状态码429，提示错误 SpeedQuotaExceed。 **说明：** 此处500QPS不代表500条日志。 一次请求中的 pb 包可以包含多条日志批量上传。
流控	单个主题分区读流量无限制，建议不超过5MB/s。
	单个主题分区压缩后写流量限制：5MB/s。超限返回状态码429，提示错误 SpeedQuotaExceed。

假设某个日志主题有10个分区，则该日志主题最大可承载的压缩后日志写流量为 `5MB/s * 10 = 50MB/s` 。单个日志主题最大可拥有50个分区，对应的最大压缩后日志写流量为 `5MB/s * 50 = 250MB/s` 。

说明：
实际使用过程中建议在创建主题时开启分区自动分裂功能，CLS 将根据实际数据量自动扩增主题分区数量，无需手动管理分区。

数据处理

最近更新时间：2024-10-24 20:59:22

数据加工

数据加工指的是对日志数据进行过滤、清洗、脱敏、富化、分发至目标日志主题的过程，可理解为日志 ETL（Extract-Transform-Load）。

- **源日志主题**：数据加工任务的输入。
- **目标日志主题**：数据加工任务的输出。
- **目标名称**：自定义目标主题名称，一是提高目标主题的可读性（业务属性），二是将日志输出到指定目标主题时，调用的函数 log_output("别名") 中使用，数据加工任务必须要有输出的目标主题，否则任务不能创建。
- **DSL 加工函数**：DSL（Domain Specific Language）是 CLS 针对日志 ETL 的需求，开发的日志数据处理函数。函数简单易用，处理性能高，底层基于 Flink 实现，可实时地处理日志。

定时 SQL 分析

定时 SQL 分析指的是根据指定的时间窗口，周期性对日志数据进行查询（支持检索和 SQL），并将查询结果保存至目标日志主题的过程。

- **源日志主题**：定时 SQL 任务的输入。
- **目标日志主题**：定时 SQL 任务的输出。
- **调度范围**：查询日志的时间范围，例如需要查询2023年1月1日 00:00:00-2023年3月31日00:00:00的日志数据。
- **调度周期**：周期性查询，取值范围1-1440分钟，如需生成日报表，可配置为1440分钟。
- **SQL时间窗口**：指定查询语句的时间窗口。配合调度周期，可以实现滚动窗口、滑动窗口。
- **滚动窗口**：没有重叠的查询窗口。如调度周期是60分钟，SQL时间窗口是60分钟。典型场景：小时报表。
- **滑动窗口**：有重叠的查询窗口。例如调度周期是1分钟，SQL时间窗口是60分钟。典型场景：绘制【1小时内的活跃用户】时序图，时间轴的粒度是1分钟。
- **延迟执行**：查询延迟的时间，在控制台【高级设置】中，取值范围60-120秒。日志生成索引一般会有延迟，在索引生成之前，不可查询，因此设置60秒延迟查询，此时索引已生成（99.9%的索引数据将在5秒内生成）。