阿里云
日志服务

服务监控

文档版本：20190916

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的”现状“、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含”阿里云”、Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 **禁止:** 重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	 **警告:** 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 **说明:** 您也可以通过按Ctrl + A选中全部文件。
>	多级菜单递进。	设置 > 网络 > 设置网络类型
粗体	表示按键、菜单、页面名称等UI元素。	单击 确定。
`courier` 字体	命令。	执行 `cd /d C:/windows` 命令，进入Windows系统文件夹。
*## *	表示参数、变量。	`bae log list --instanceid` *`Instance_ID`*
[]或者[a|b]	表示可选项，至多选择一个。	`ipconfig` *`[-all|-t]`*
{}或者{a|b}	表示必选项，至多选择一个。	`swich` *`{stand | slave}`*

目录

1 监控日志服务

日志服务支持通过以下方式实现日志监控：

- 通过云监控
  - 日志库（Logstore）级读写
  - Agent（Logtail）日志收集
- 控制台
  - 实时订阅消费（Spark Streaming、Storm、Consumer Library）当前点位
  - 投递 OSS/MaxCompute 状态

其中，云监控方式即通过阿里云云产品云监控完成，控制台方式监控日志请参见控制台#unique_4、#unique_5和#unique_6。

云监控配置步骤

> 说明：
> 使用子账号配置云监控需要先授权。详细内容请参考 #unique_7。

1. 登录日志服务管理控制台。
2. 选择所需的项目，单击项目名称。

3. 选择所需的日志库并单击监控列下的图标打开云监控。

   您也可以通过云监控管理控制台左侧导航栏中的云服务监控 > 日志服务进入监控配置页面。

   在云监控中对日志数据进行监控，详细信息请参考#unique_8。

   图 1-1: 云监控中的日志监控

监控项含义

请参考#unique_9。

设置报警规则

在监控图表页面右上角单击创建报警规则，设置关联资源、报警规则和通知方式。详细步骤请参考#unique_10。

2 服务日志

2.1 简介

阿里云日志服务提供服务日志功能，支持记录Project内的用户操作日志等多种日志数据，并提供多种分析维度的仪表盘。您可以通过多种方式实时掌握日志服务的使用状况、提高运维效率。

限制说明

- 专属日志库用于存入日志服务产生的日志，因此不支持写入其他数据。查询、统计、报警、流式消费等其他功能没有限制。
- RAM用户开通服务日志，需要主账户为其授权。
- Project产生的日志只能保存至同一地域的Project中，不支持跨地域操作。
- 日志服务的服务日志功能产生的所有日志数据遵循日志服务的计费策略，即按量计费，且按月提供免费额度。计费说明请查看#unique_13。
- 开启服务日志功能后，您可以取消勾选服务日志中的日志类型，以关闭该功能。关闭功能后停止写入服务日志，但不删除历史日志，可能会产生少量费用。如果您需要删除历史日志，可以直接删除保存日志的Logstore。

默认配置

表 2-1: 默认配置

默认配置项	配置内容
Logstore	默认为您创建以下2个Logstore。当前Project产生的所有日志数据都会被分类保存到特定的Logstore中。 - `internal-operation_log`：记录操作日志，每条日志对应一次请求。默认保存30天，计费方式与普通Logstore一致。 - `internal-diagnostic_log`：记录计量日志，消费组延时和Logtail相关的日志，根据topic进行区分。默认保存30天。不产生费用。 关于日志类型和字段，请参见日志类型。
地域	- 选择自动创建（推荐）时，将在相同地域创建Project。 - 日志服务的服务日志仅支持将产生的日志信息保存至当前地域的Project中。

默认配置项	配置内容
Shard	每个Logstore默认创建2个Shard，并开启自动分裂Shard功能。
日志存储时间	默认保存30天，支持修改保存时间。详细步骤请参见#unique_16。
索引	默认为采集到的所有日志数据开启索引。如果没有查询分析和设置告警等需求，可以在查询页面关闭索引。
仪表盘	默认创建如下5个仪表盘： · 用户操作统计 · 计量统计 · Logtail日志采集统计 · Logtail异常监控 · 消费组监控 关于仪表盘的更多信息，请参见#unique_17。

应用场景

· 查看计量数据

  用户开通日志服务后，日志服务每小时统计一次每个Logstore日志和索引占用的存储空间、上一次统计之后用户的读写次数和索引流量等和计费相关的数据。开通服务日志后，计量日志将会实时保存到一个单独的Logstore中，用户可以了解账户的存储量和消费等信息。

· 判断Shard写入和消费是否均衡

  您可以通过预定义的仪表盘对比Shard数据写入和消费变化趋势来判断Shard写入和消费是否均衡。

  如果Project中存在多个Logstore，不同的Logstore可能存在重复的Shard。因此如果需要查看某个Logstore的Shard写入分布，请在仪表盘左上角的过滤条件增加Project和Logstore作为过滤条件。

· API请求状态监控

  用户的所有操作如日志写入、消费、创建Project和Logstore等，都是通过API请求的方式。用户的每次操作都对应internal-operation_log这个Logstore中的一条日志。如果请求失败，则日志的Status字段为一个大于200的整数如404。因此，用户可以通过监控Status>200的日志条数来监控API请求是否正常。

- 查看Logtail的状态

  服务日志开通后默认创建两个Logtail相关的仪表盘，分别为异常监控和采集的数据统计。通过异常监控这个仪表盘，用户可以看到Logtail有哪些异常状况出现，如日志解析失败，正则表达式不匹配等。

2.2 开通、修改和关闭服务日志

您可以在Project列表中开启服务日志功能，日志服务会将当前Project产生的所有日志保存到您指定的Project中。该功能默认为关闭状态。

背景信息

日志服务提供记录服务日志的功能，可以记录指定Project的操作日志、Logtail告警日志等日志数据，支持保存在新建的Project或已有的Project中，并在指定保存位置自动创建Logstore，分别用来保存操作日志和其他日志。同时针对各种日志场景提供5个仪表盘，实时查看和监控运行状态。

说明:

- 开通服务日志会在您选择的存储位置创建对应的Logstore和仪表盘，用于存储操作日志的Logstore按照正常Logstore计费，存放其他日志的Logstore不产生费用。
- 建议将同一地域的日志都保存在同一个日志服务自动创建的Project中。
- 仅记录开通之后的服务日志，开通之前的服务日志不会被记录。

前提条件

1. 已创建Project。
2. 如果是子账号身份登录，必须先由主账号授予对应权限。

开启服务日志

1. 登录日志服务控制台首页，找到需要开启服务日志的Project。
2. 在概览页面中，展开服务日志节点并单击开通服务日志。
3. 勾选需要记录的日志类型。

   日志服务提供详细日志和重要日志两种日志类型。

   - 详细日志：记录Project内所有资源创建、修改、更新、删除等操作日志和数据读写日志。将保存在指定Project的Logstore internal-operation_log中。
   - 重要日志：包括Logstore粒度的计量日志、消费组消费延时日志、Logtail相关的错误、心跳和统计日志。将保存在指定Project的Logstore internal-diagnostic_log中。

4. 选择日志存储位置。
   - 自动创建（推荐）：日志服务会自动在相同地域创建一个Project，名称为`log-service-{用户ID}-{region}`，建议将同一地域的日志都保存到该Project中。
   - 其他已存在的Project：将服务日志存储在其他已存在的Project中。
5. 单击确认。

   已成功开启服务日志功能，功能开启期间，该Project产生的日志会实时记录在指定位置。

   图 2-1: 开通服务日志

修改记录的日志类型和存储位置

1. 主账号登录日志服务控制台首页，单击需要修改服务日志的Project。
2. 在概览页面服务日志区域右侧单击操作按钮。
3. 修改记录的日志类型。

   勾选需要记录的日志类型、取消勾选不需要记录的日志类型即可。
4. 修改日志存储位置。

   在日志存储位置中选择其他Project。

   说明:

- 建议将服务日志保存在自动创建的Project中。同一个地域的Project日志可以保存在同一个Project中。
- 修改日志存储位置后，新的日志数据会保存在指定Project中。原Project中保存的日志数据和仪表盘不会同步迁移或删除，若您不再需要该部分数据，可以手动删除。

图 2-2: 修改记录的日志类型和存储位置

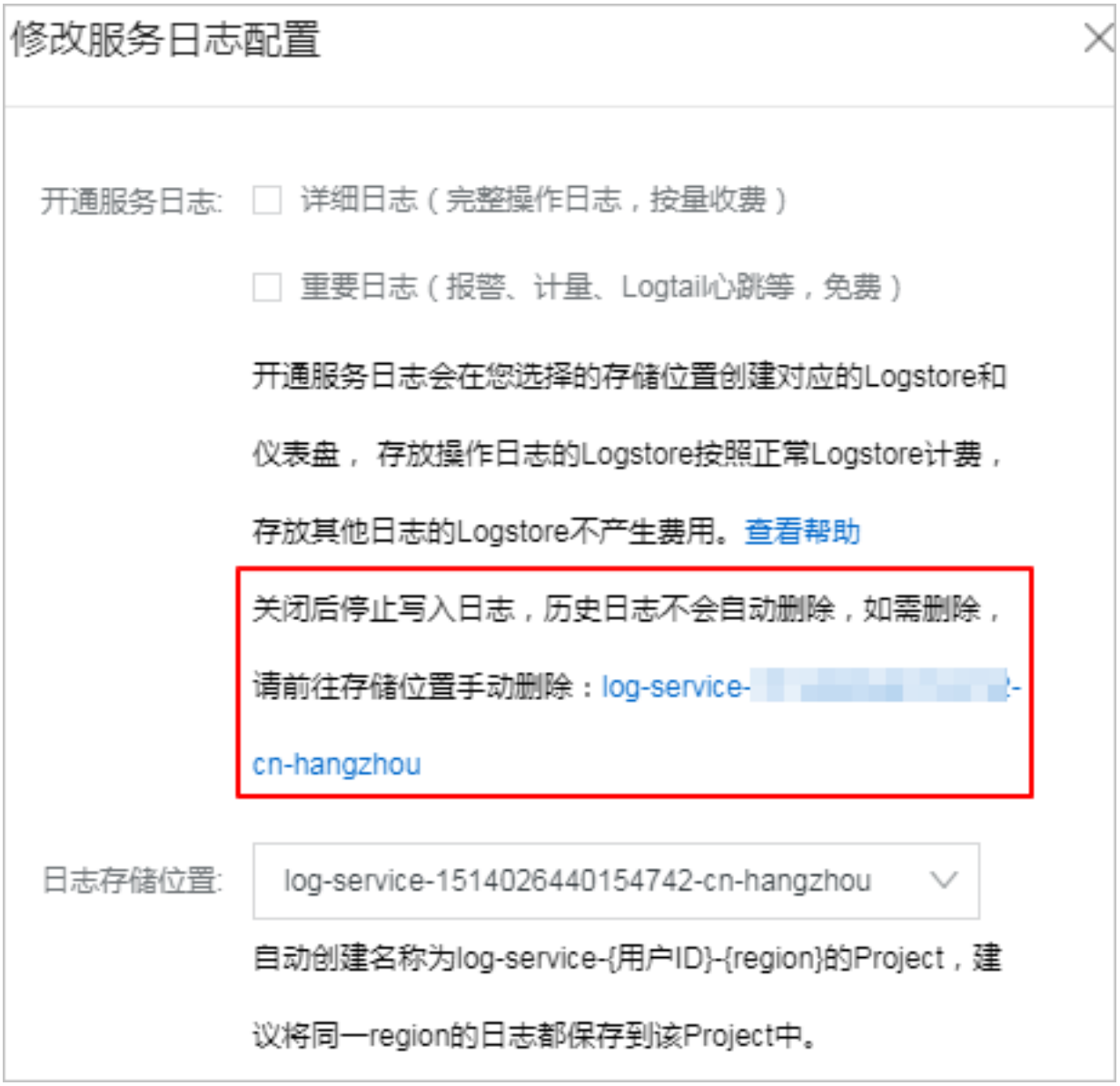
关闭服务日志

说明:

关闭服务日志后，Project中保存的日志数据和仪表盘不会自动删除，如果不需要继续保存这部分日志数据，可以选择手动删除保存日志的Project或Logstore。

1. 主账号登录日志服务控制台首页，单击需要关闭服务日志的Project。
2. 在概览页面服务日志区域右侧单击操作按钮。

3. 取消勾选所有类型的日志即可关闭服务日志功能。



4. 单击确认。

为RAM用户授权

使用子账号操作服务日志相关功能之前，必须由主账号为其授予相关权限。授权操作请参考#unique_20。对应权限内容如下：

```
{
  "Version": "1",
  "Statement": [
    {
      "Action": [
        "log:CreateDashboard",
        "log:UpdateDashboard"
      ],
      "Resource": "acs:log:*:*:project/{存储日志的Project}/dashboard/*",
      "Effect": "Allow"
    },
    {
```

```
      "Action": [
        "log:GetProject",
        "log:CreateProject",
        "log:ListProject"
      ],
      "Resource": "acs:log:*:*:project/*",
      "Effect": "Allow"
    },
    {
      "Action": [
        "log:List*",
        "log:Create*"
        "log:Get*",
        "log:Update*",
      ],
      "Resource": "acs:log:*:*:project/{存储日志的Project}/logstore/*",
      "Effect": "Allow"
    },
    {
      "Action": [
        "log:*"
      ],
      "Resource": "acs:log:*:*:project/{开通日志的Project}/logging",
      "Effect": "Allow"
    }
  ]
}
```

2.3 日志类型

服务日志功能记录多种日志类型，本文档详细介绍各种日志类型的日志字段。

日志类型

开启服务日志时，可以选择记录的日志类型，包括：

- 操作日志：记录Project内所有资源的创建、修改、更新、删除操作日志和数据读写日志。将保存在指定Project的Logstore internal-operation_log中。
- 其他日志：包括Logstore粒度的计量日志、消费组消费延时日志、Logtail相关的错误、心跳和统计日志。将保存在指定Project的Logstore internal-diagnostic_log中。

日志来源	Logstore	日志来源	说明
操作日志	internal-operation_log	用户操作日志	所有API请求和操作日志，包括控制台，消费组，SDK等所有客户端发送的请求。
其他日志	internal-diagnostic_log	消费组快照日志	消费组的消费延时日志，2分钟上报一次。指定查询消费组快照日志时，需要在查询语句中指定`__topic__: consumergroup_log`。

日志来源	Logstore	日志来源	说明
		Logtail告警日志	Logtail的错误日志。 每30秒记录一次，30秒内重复出现的错误类型只记录错误总和，错误Message随机选择一条。指定查询Logtail告警日志时，需要在查询语句中指定`__topic__: logtail_alarm`。
		Logtail采集日志	Logtail文件采集统计信息。 10分钟记录一次。指定查询Logtail采集日志时，需要在查询语句中指定`__topic__: logtail_profile`。
		计量日志	用户计量相关统计。 Logstore粒度的存储空间，读写流量，索引流量，请求次数等统计信息。每小时统计一次。指定查询计量日志时，需要在查询语句中指定`__topic__: metering`。
		Logtail状态日志	Logtail定时上报的状态日志。 每分钟记录一次。指定查询Logtail状态日志时，需要在查询语句中指定`__topic__: logtail_status`。

用户操作日志

操作日志可以根据Method字段分为读数据、写数据和资源操作三种操作。

分类	方法
读数据	· GetLogStoreHistogram · GetLogStoreLogs · PullData · GetCursor · GetCursorTime

分类	方法
写数据	· PostLogStoreLogs · WebTracking
资源操作	CreateProject、DeleteProject等其他方法。

公共字段

所有操作的公共字段如下表所示。

字段	描述	示例
APIVersion	API 版本。	0.6.0
InvokerUid	执行操作的用户的账户id。	1759218115323050
NetworkOut	通过公网入口读取的流量，单位为字节。	10
Latency	请求延时，单位为微秒。	123279
LogStore	Logstore的名称。	logstore-1
Method	执行方法。	GetLogStoreLogs
Project	Project名称。	project-1
NetOutFlow	读取的流量，单位为字节。	120
RequestId	请求ID。	8AEADC8B0AF2FA2592C9509E
SourceIP	发送请求的客户端IP。	1.2.3.4
Status	请求响应状态码。	200
UserAgent	客户端用户代理。	sls-java-sdk-v-0.6.1

读操作字段

读请求特有的字段如下表所示：

字段	描述	示例
BeginTime	请求开始时间，Unix时间戳。	1523868463
DataStatus	请求响应数据状态。包括Complete、OK、Unknown等。	OK
EndTime	请求结束时间，Unix时间戳。	1523869363
Offset	GetLog请求偏移行数。	20

字段	描述	示例
Query	原始查询语句。	UserAgent: [consumer-group-java]*
RequestLines	期望返回行数。	100
ResponseLines	返回结果行数。	100
Reverse	是否按日志时间戳逆序返回日志，其中： · 1 表示逆序。 · 0表示顺序。 默认值为 0。	0
TermUnit	搜索语句中包含的词项个数。	0
Topic	读取的主题名称。	topic-1

写操作字段

写请求特有的字段如下表所示：

字段	描述	示例
InFlow	原始写入字节数。	200
InputLines	请求写入行数。	10
NetInflow	压缩之后的写入字节数。	100
Shard	写入的Shard Id。	1
Topic	数据写入的topic。	topic-1

消费组快照日志

字段	描述	示例
consumer_group	消费组名称。	consumer-group-1
fallbehind	当前消费位置距离最新写入日志的落后时间，单位为秒。	12345
logstore	日志库名称。	logstore-1
project	项目名称。	project-1
shard	Shard Id。	1

Logtail告警日志

字段	描述	示例
alarm_count	采样窗口内告警次数。	10
alarm_message	触发告警的原始日志采样。	M_INFO_COL,all_status_monitor,T22380,0,2018-04-17 10:48:25.0,AY66K,AM5,2018-04-17 10:48:25.0,2018-04-17 10:48:30.561,i-23xebl5ni.1569395.715455,901,00789b
alarm_type	告警类型。	REGISTER_INOTIFY_FAIL_ALARM
logstore	日志库名称。	logstore-1
source_ip	Logtail运行的机器IP地址。	1.2.3.4
os	操作系统，Linux或Windows等。	Linux
project	项目名称。	project-1
version	Logtail版本。	0.14.2

Logtail采集日志

Logtail采集日志可以根据字段`file_name`分为两类，一类是针对单个文件的采集统计信息，一类是Logstore级别的统计信息，即`file_name`为`logstore_statistics`的部分。当`file_name`为`logstore_statistics`时，文件相关的字段如`file_dev`、`file_inode`无意义。字段说明如下表所示。

字段	描述	示例
logstore	日志库名称。	logstore-1
config_name	采集配置名。由`#########projectName$####`组成的全局唯一配置名。	##1.0##project-1$logstore-1
error_line	引起错误的原始日志。	M_INFO_COL,all_status_monitor,T22380,0,2018-04-17 10:48:25.0,AY66K,AM5,2018-04-17 10:48:25.0,2018-04-17 10:48:30.561,i-23xebl5ni.1569395.715455,901,00789b

字段	描述	示例
file_dev	该日志文件的device ID。	123
file_inode	该日志文件的inode。	124
file_name	日志文件完整路径或者或者`logstore_statistics`。	/abc/file_1
file_size	当前文件大小，单位为字节。	12345
history_data_failures	历史处理失败次数。	0
last_read_time	窗口内最近的读取时间，Unix时间戳。	1525346677
project	项目名称。	project-1
logtail_version	Logtail版本。	0.14.2
os	操作系统。	Windows
parse_failures	窗口日志解析失败的行数。	12
read_avg_delay	窗口内平均每次读取时当前偏移量与文件大小差值的平均值。	65
read_count	窗口内日志读取次数。	10
read_offset	当前读取到文件偏移位置，单位为字节。	12345
regex_match_failures	正则表达式匹配失败次数。	1
send_failures	窗口内发送失败的次数。	12
source_ip	Logtail运行机器IP地址。	1.2.3.4
succeed_lines	处理成功的Log行数。	123
time_format_failures	日志时间匹配失败次数。	122
total_bytes	读取的总字节数。	12345

如下字段只有在`file_name`为`logstore_statistics`时才会存在：

字段	描述	示例
send_block_flag	窗口结束时发送队列是否阻塞。	false
send_discard_error	窗口内因数据异常或无权限导致丢弃数据包的个数。	0

字段	描述	示例
send_network_error	窗口内因网络错误导致发送失败的数据包个数。	12
send_queue_size	窗口结束时当前发送队列中未发送数据包数。	3
send_quota_error	窗口内因quota超限导致发送失败的数据包个数。	0
send_success_count	窗口内发送成功的数据包个数	12345
sender_valid_flag	窗口结束时该Logstore的发送标志位是否有效，其中： · true表示正常。 · false表示可能因为网络错误或quota错误而被禁用。	true
max_send_success_time	窗口内发送成功数据的最大时间。Unix时间戳。	1525342763
max_unsend_time	窗口结束时发送队列中未发送数据包的最大时间，队列空时为0。Unix时间戳。	1525342764
min_unsend_time	窗口结束时发送队列中未发送数据包的最小时间，队列空时为0。Unix时间戳。	1525342764

计量日志

字段	描述	示例
begin_time	统计窗口开始时间，Unix时间戳。	1525341600
index_flow	统计窗口内索引流量，单位为字节。	12312
inflow	统计窗口内写入流量，单位为字节。	12345
logstore	Logstore名称。	logstore-1
network_out	统计窗口内流出到公网流量，单位为字节。	12345
outflow	统计窗口内读取流量，单位为字节。	23456
project	项目名称。	project-1

字段	描述	示例
read_count	统计窗口内读操作次数。	100
shard	统计窗口内平均使用shard个数。	10.0
storage_index	统计时间点Logstore内索引总存储用量，单位为字节。	10000000
storage_raw	统计时间点Logstore内日志总存储用量，单位为字节。	20000000
write_count	统计窗口内写操作次数。	199

Logtail状态日志

字段	描述	示例
cpu	进程CPU负载。	0.001333156
hostname	主机名。	abc2.et12
instance_id	实例ID，是一个随机ID。	05AFE618-0701-11E8-A95B-00163E025256_10.11.12.13_1517456122
ip	IP地址。	1.0.1.0
load	系统平均负载。	0.01 0.04 0.05 2/376 5277
memory	Logtail进程占用的内存大小，单位为MB。	12
detail_metric	各项计量值，JSON格式，详细内容请参考detail_metric。	detail_metric
os	操作系统。	Linux
os_cpu	系统整体的CPU使用率。	0.004120005
os_detail	操作系统详细信息。	2.6.32-220.23.8.tcp1.34.el6.x86_64

字段	描述	示例
status	客户端状态，包括： · ok · busy · many_log_files · process_block · send_block · send_error 详细说明请查看Logtail运行状态。	busy
user	用户名。	root
user_defined_id	用户定义的ID。	aliyun-log-id
uuid	机器的uuid。	64F28D10-D100-492C-8FDC-0C62907F1234
version	Logtail版本。	0.14.2
project	Logtail配置所属的Project。	my-project

其中，字段 detail_metric 包含如下属性：

字段	描述	示例
config_count	Logtail配置数量。	1
config_get_last_time	上一次获取配置的时间。	1525686673
config_update_count	Logtail启动之后配置的更新次数。	1
config_update_item_count	Logtail启动之后配置项的更新总和。	1
config_update_last_time	Logtail启动之后配置的最后一次更新时间。	1
event_tps	事件数TPS。	1
last_read_event_time	上一次获取事件的时间。	1525686663
last_send_time	上一次发送数据的时间。	1525686663
open_fd	目前打开的文件数量。	1
poll_modify_size	监听修改事件的文件数量。	1
polling_dir_cache	扫描的文件夹数量。	1
polling_file_cache	扫描的文件数量。	1

字段	描述	示例
process_byte_ps	每秒处理的日志量（字节）。	1000
process_lines_ps	每秒处理的日志条数。	1000
process_queue_full	达到最大长度限制的发送队列个数。	1
process_queue_total	处理队列数量。	10
process_tps	处理TPS。	0
reader_count	正在处理的文件数。	1
region	Logtail所在的地域。	cn-hangzhou,cn-shanghai
register_handler	注册监听的文件夹数量。	1
send_byte_ps	每秒发送的原始日志量（字节）。	11111
send_line_ps	每秒发送的日志条数。	1000
send_net_bytes_ps	每秒发送的网络数据量（字节）。	1000
send_queue_full	达到最大长度限制的发送队列个数。	1
send_queue_total	发送队列数量。	12
send_tps	发送TPS。	0.075
sender_invalid	异常的发送队列数。	0

2.4 服务日志仪表盘

开通服务日志后，自动创建5个仪表盘，展示用户操作统计、计量统计、Logtail日志采集统计、Logtail异常监控和消费组监控数据。

仪表盘说明

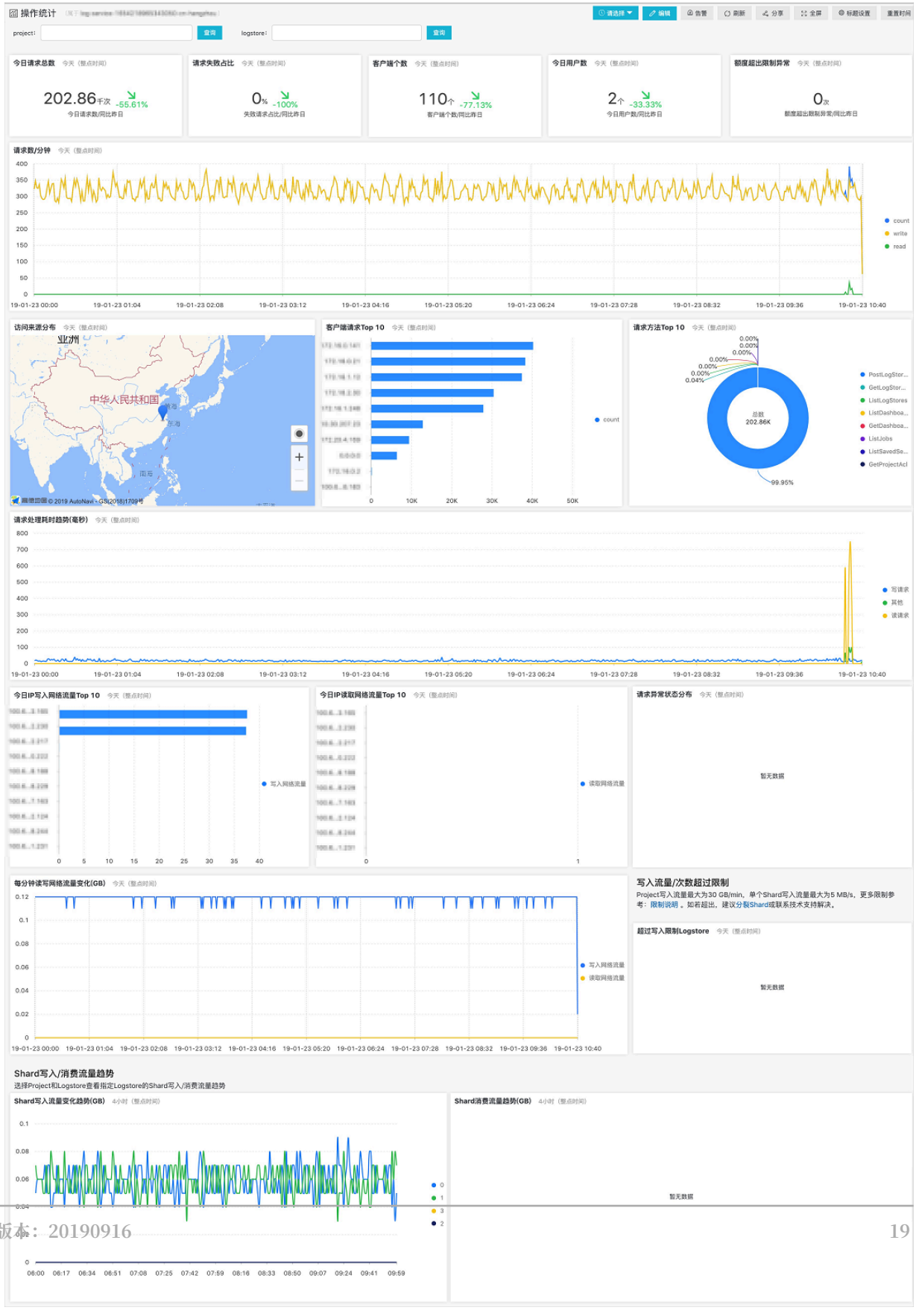
开启服务日志时，可以选择记录的日志类型，其中

- 开启操作日志后，日志服务提供操作统计仪表盘，展示今日请求统计、请求失败占比等数据。
- 开启其他日志后，日志服务提供计量统计、Logtail日志采集统计、Logtail运行监控和消费组监控仪表盘。

操作统计

展示所有的用户访问和操作信息，如QPS、请求延时等信息。包括API请求、Project的所有资源的操作等所有请求。

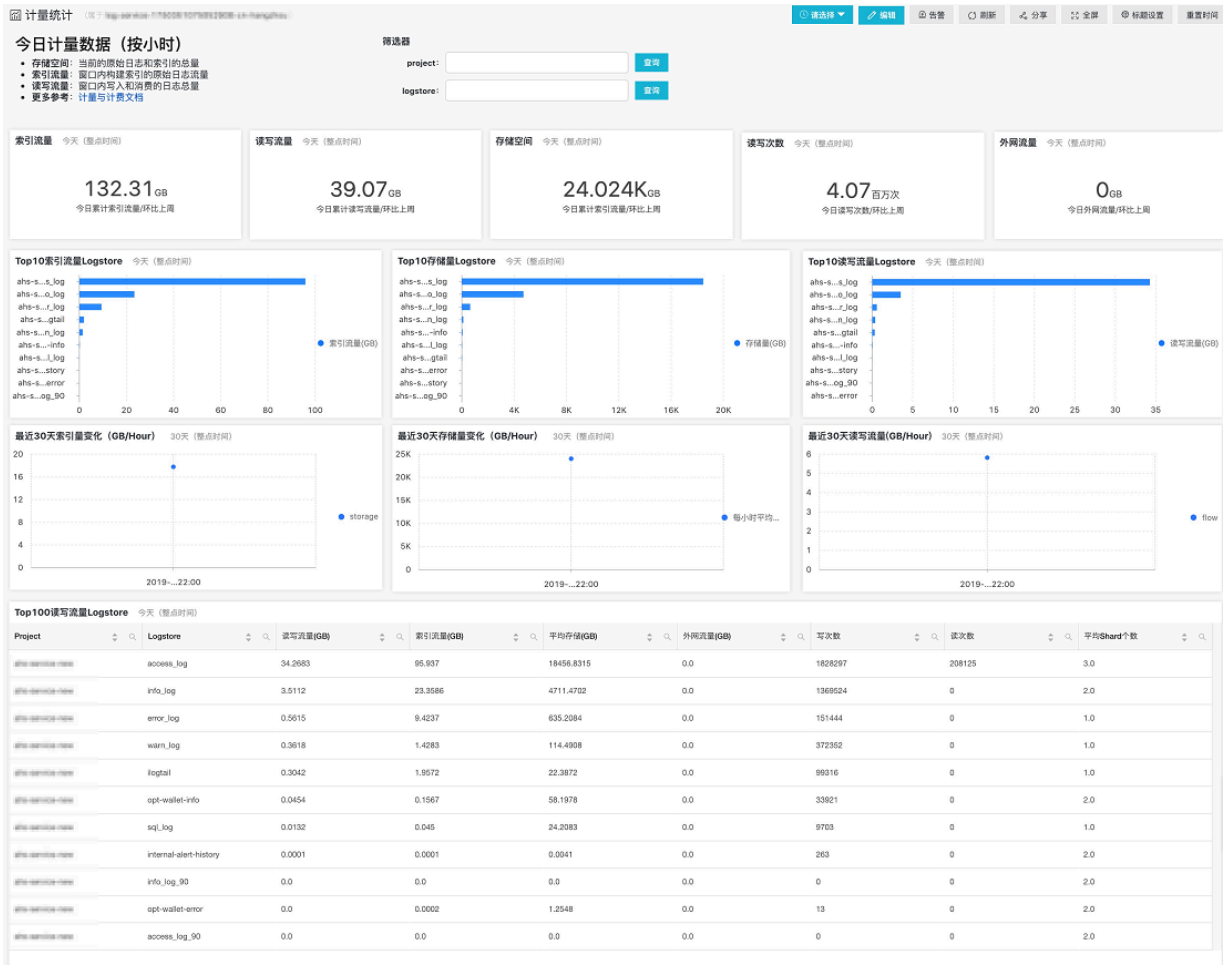
图 2-3: 用户操作统计



计量统计

展示用户计量相关的数据，如使用的存储量、索引流量、读写次数、读写流量等。

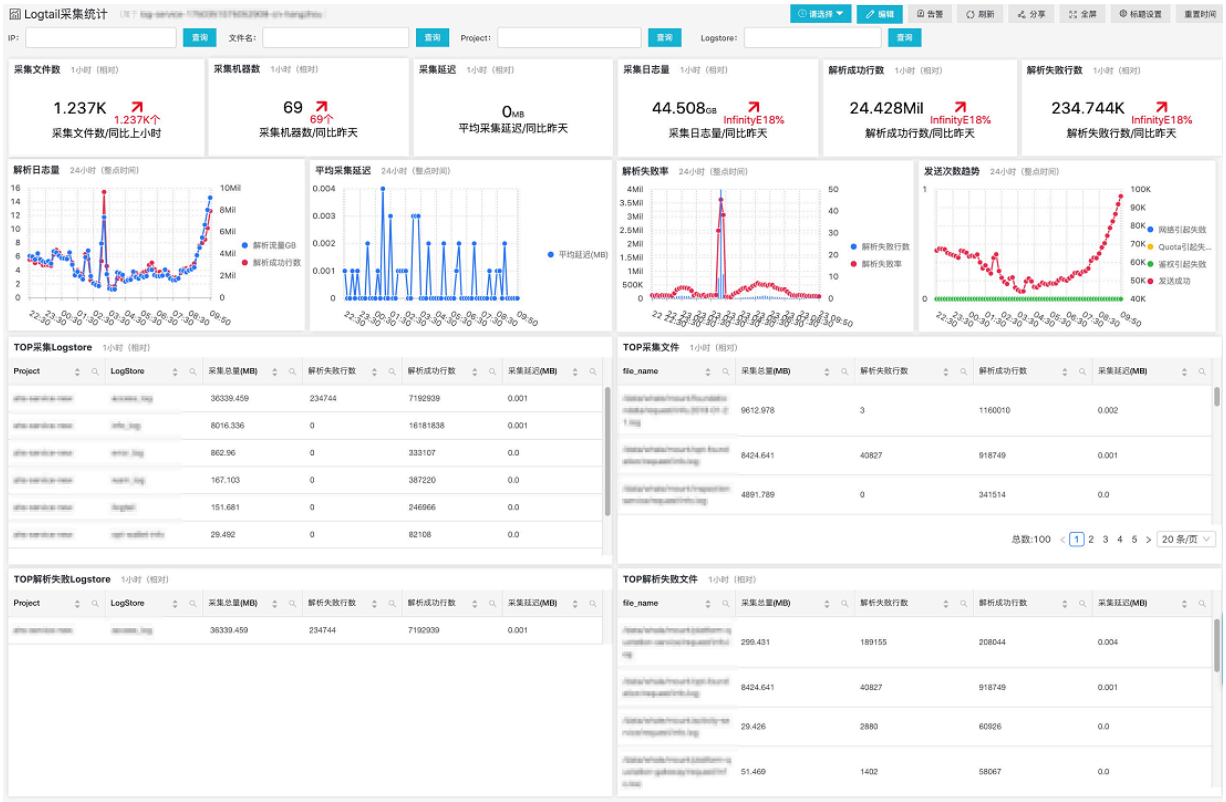
图 2-4: 计量统计



Logtail日志采集统计

展示了Logtail日志采集相关的统计信息。

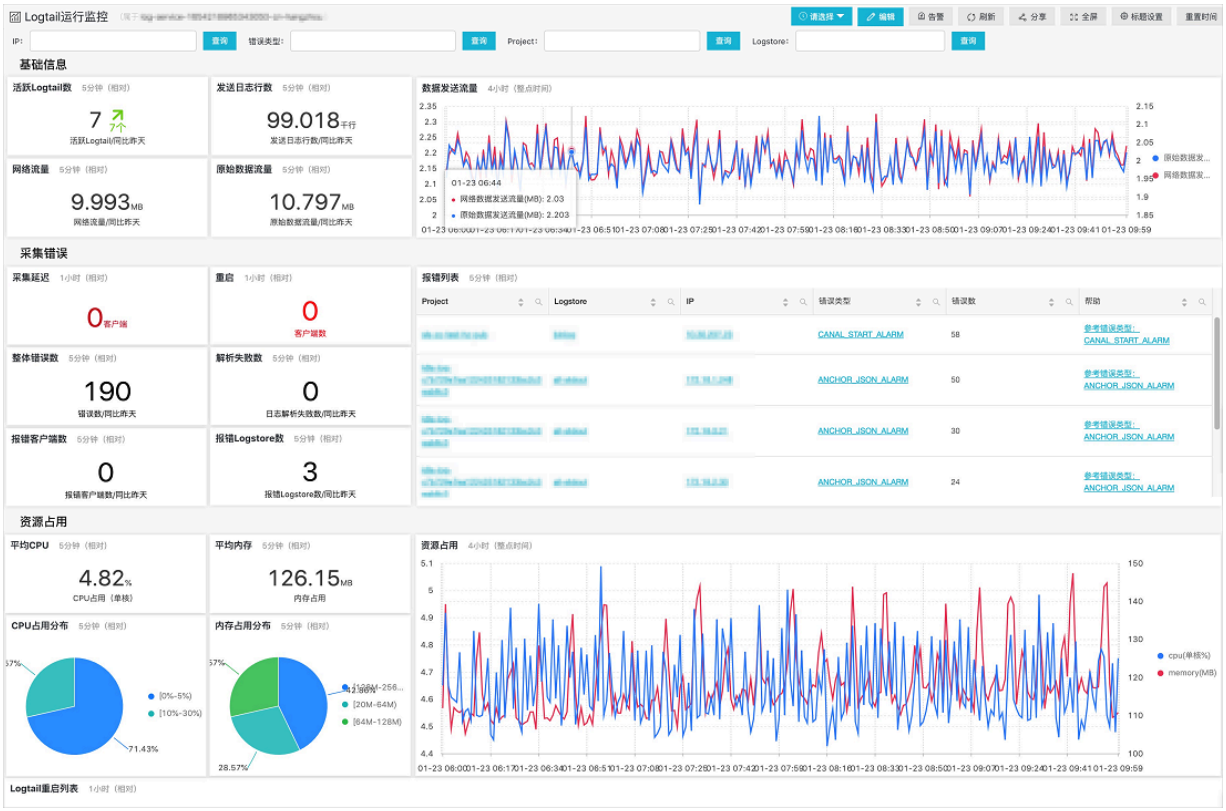
图 2-5: Logtail日志采集统计



Logtail运行监控

展示Logtail所有的错误告警信息，便于用户实时监控Logtail的健康状态。

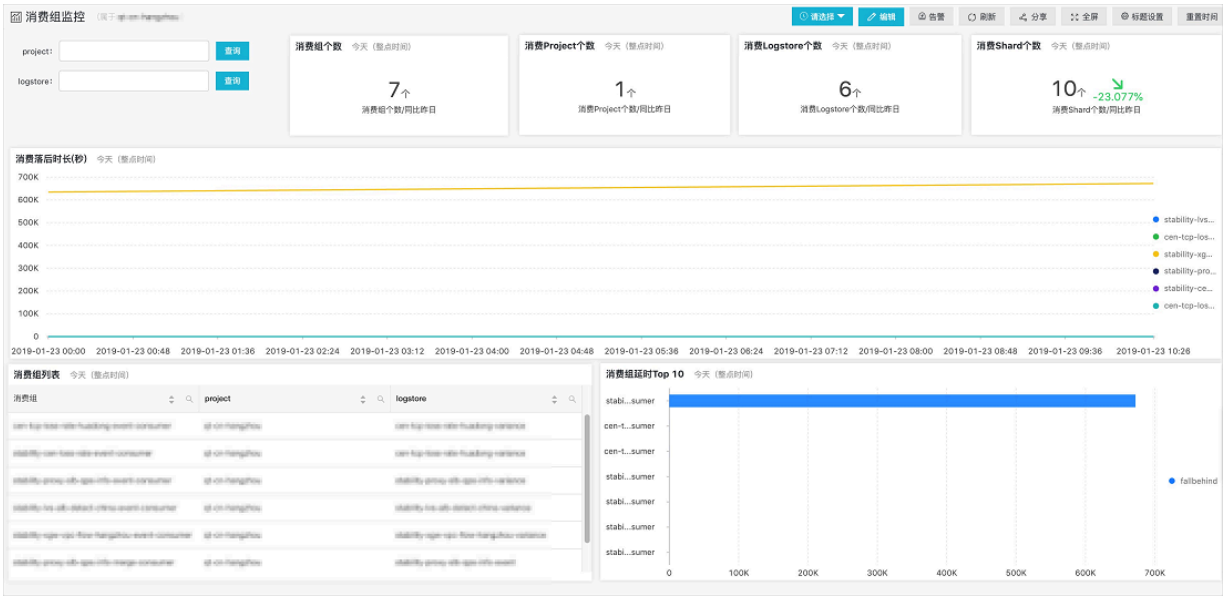
图 2-6: Logtail运行监控



消费组监控

展示和消费组相关的统计，包括分Shard的消费数据、消费的落后时间、消费组列表等。

图 2-7: 消费组监控



3 云监控方式

3.1 云监控指标

监控数据入口请参考#unique_32。

1. 写入/读取流量
   - 含义：每个日志库（Logstore）写入、以及读取实时情况。统计该Logstore通过ilogtail和SDK、API等读写实时流量，大小为传输大小（压缩情况下为压缩后），每分钟统计一个点。
   - 单位：Bytes/min
2. 原始数据大小
   - 含义：每个Logstore写入数据原始大小，即压缩前的大小。
   - 单位：Byte/min
3. 总体QPS
   - 含义：所有操作QPS，每分钟统计一个点。
   - 单位：Count/Min

4. 操作次数

- 含义：统计用户的各种操作对应的QPS，每分钟统计一个点。
- 单位：次/分钟（Count/Min）
- 所有的操作包括：
  - 写入操作：
    - PostLogStoreLogs： 0.5API以后版本接口。
    - PutData： 0.4 API以前版本接口。
  - 根据关键字查询：
    - GetLogStoreHistogram：查询关键字分布情况，0.5API以后版本接口。
    - GetLogStoreLogs：查询关键字命中日志，0.5API以后版本接口。
    - GetDataMeta： 同GetLogStoreHistogram，为0.4API以前版本接口。
    - GetData : 同GetLogStoreLogs，为0.4API以前版本接口。
  - 批量获取数据：
    - GetCursorOrData：该操作包含了获取Cursor和批量获取数据两种方法。
    - ListShards：获取一个Logstore下所有的Shard。
  - List操作：
    - ListCategory：同ListLogStores，为0.4API以前版本接口。
    - ListTopics：遍历一个Logstore下所有的Topic。

5. 服务状态

- 含义：该视图统计用户各种操作返回的HTTP状态码对应的QPS，方便用户根据返回的状态码来判断操作是否异常，及时调整程序。
- 各状态码:
  - 200：为正常的返回码，表示操作成功。
  - 400：参数错误，包括Host、Content-length、APIVersion、RequestTimeExpired、查询时间范围，Reverse，AcceptEncoding、AcceptContentType、Shard、Cursor、PostBody、Paramter和ContentType等方面的错误。
  - 401：鉴权失败，包括AccessKeyId不存在、签名不匹配或者签名账户没有操作权限，请到日志服务控制台上查看Project权限列表，是否包含了该AK。
  - 403：超过预定Quota，包括能够创建的Logstore个数、Shard总数、以及读写操作的每分钟限额，请根据返回的Message判断发生了哪种错误。
  - 404：请求的资源不存在，包括Project、 Logstore、Topic 、User等资源。
  - 405：错误的操作方法，请检查请求的URL路径。
  - 500：服务端错误，请重试。
  - 502：服务端错误，请重试。

6. 客户端解析成功流量

- 含义：Logtail收集成功的日志大小，为原始数据大小。
- 单位：字节

7. 客户端（Logtail）解析成功行数

- 含义：Logtail收集成功的日志的行数。
- 单位: 行

8. 客户端解析失败行数

- 含义：Logtail收集日志过程中，采集出错的行数大小，如果该视图有数据，表示有错误发生。
- 单位：行

9. 客户端错误次数

- 含义：Logtail收集日志过程中，出现所有收集错误的IP总数。
- 单位：次

10.发生客户端错误机器数

- 含义：Logtail收集日志过程中，出现收集错误的告警客户端数目。
- 单位：个

11.错误IP统计（Count/5min)

- 含义：分类别展示各种采集错误发生的IP数，各种错误包括：
  - LOGFILE_PERMINSSION_ALARM：没有权限打开日志文件。
  - SENDER_BUFFER_FULL_ALARM：数据采集速度超过了网络发送速度，数据被丢弃。
  - INOTIFY_DIR_NUM_LIMIT_ALARM（INOTIFY_DIR_QUOTA_ALARM）：监控的目录个数超过了3000个，请把监控的根目录设置成更低层目录。
  - DISCARD_DATA_ALARM：数据丢失，因为数据时间在系统时间之前15分钟，请保证新写入日志文件的数据是在15分钟之内的。
  - MULTI_CONFIG_MATCH_ALARM：有多个配置在收集同一个文件，Logtail会随机选择一个配置进行收集，另一个配置则收集不到数据。
  - REGISTER_INOTIFY_FAIL_ALARM：注册inotify事件失败，具体原因请查看Logtail日志。
  - LOGDIR_PERMINSSION_ALARM：没有权限打开监控目录。
  - REGEX_MATCH_ALARM：正则式匹配错误，请调整正则式。
  - ENCODING_CONVERT_ALARM：转换日志编码格式时出现错误，具体原因请查看Logtail日志。
  - PARSE_LOG_FAIL_ALARM：解析日志错误，一般是行首正则表达式错误或单条日志超过512KB导致的日志分行错误，请查看Logtail日志确定原因，如行首正则表达式错误请调整配置。
  - DISCARD_DATA_ALARM：丢弃数据，Logtail发送数据到日志服务失败且写本地缓存文件失败导致，可能的原因是日志文件产生较快但写磁盘缓存文件较慢。
  - SEND_DATA_FAIL_ALARM：解析完成的日志数据发送到日志服务失败，请查看Logtail日志发送数据失败相关ErrorCode和ErrorMessage，常见的错误有服务端Quota超限、客户端网络异常等。
  - PARSE_TIME_FAIL_ALARM：解析日志time字段出错，Logtail根据正则表达式解析出来的time字段按照时间格式配置无法解析成功，请修改配置。
  - OUTDATED_LOG_ALARM：Logtail丢弃历史数据，请保证当前写入日志数据的时间与系统时间相差在5分钟以内。
- 请根据具体错误找到出错IP，登录机器查看/usr/logtail/ilogtail.LOG，分析错误原因。

3.2 设置云监控告警规则

日志服务支持通过云监控设置报警，当服务状态符合设置的报警规则时发送报警短信或邮件。您可以通过配置云监控中的日志监控报警规则，对日志收集状态、Shard资源使用状态等异常状态进行监控。

操作步骤

在云监控控制台的云服务监控 > 日志服务页面，找到需要设置监控报警的Logstore，单击其右侧的报警规则。在报警规则右上角单击创建报警规则。

1. 关联资源。

   a. 选择产品。此处请选择日志服务。

   b. 选择资源范围。

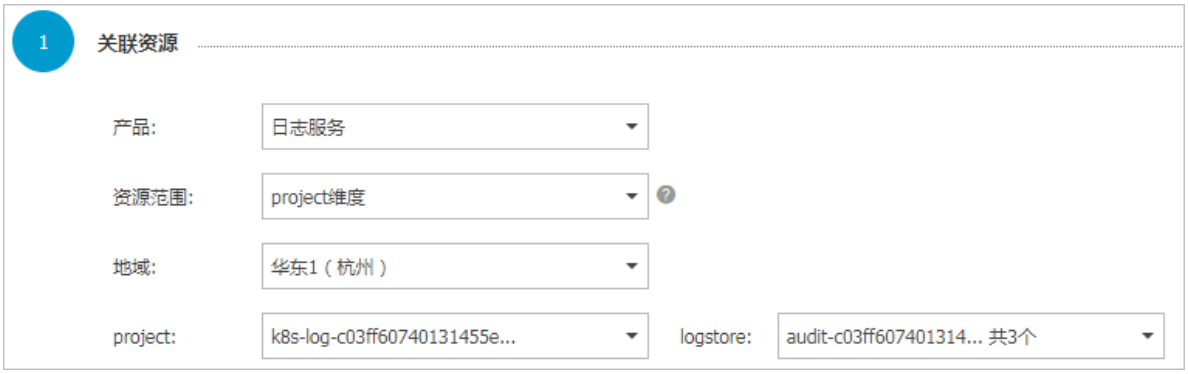
      您可以选择全部资源和Project维度。

      - 资源范围选择全部资源，则产品下任何实例满足报警规则描述时，都会发送报警通知。
      - 选择Project维度，则选中的实例满足报警规则描述时，才会发送报警通知。

   c. 选择地域。

   d. 选择Project和Logstore。您可以选择一个及以上的Project和Logstore。

   图 3-1: 关联资源



2. 设置报警规则。

   您可以设置一条或多条报警规则。

   a. 填写规则名称。

   b. 填写规则描述。

      您需要在此处定义您的监控策略，选择需要的监控项目，并为其设定阈值。超出该值后云监控会发送报警通知。

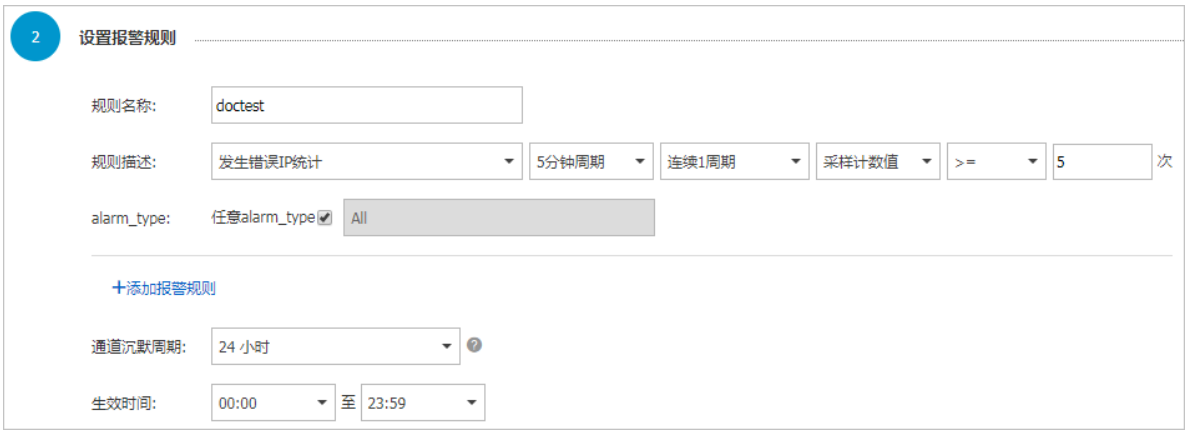
      各个监控项的含义请参考#unique_9，统计方法请参考#unique_32。

   c. 选择alarm_type。默认为任意alarm_type。

   d. 设置通道沉默时间。指报警发生后如果未恢复正常，间隔多久重复发送一次报警通知。

   e. 生效时间。为您的监控策略选择生效时间，设定后仅在该时段内执行监控报警策略。

   图 3-2: 设置报警规则



3. 配置通知方式。

   a. 通知对象。以联系人通知组级别发送通知。

   b. 报警级别。您可以按照需要选择Warning和Info级别，不同级别通知方式不同。

   c. 邮件主题和邮件备注。邮件主题默认为产品名称+监控项名称+实例ID。

   d. 报警回调。填写公网可访问的URL，云监控会将报警信息通过POST请求推送到该地址，目前仅支持HTTP协议。

图 3-3: 通知方式

配置完成后单击确认，完成监控策略配置。

示例

- 监控Logtail日志收集状态

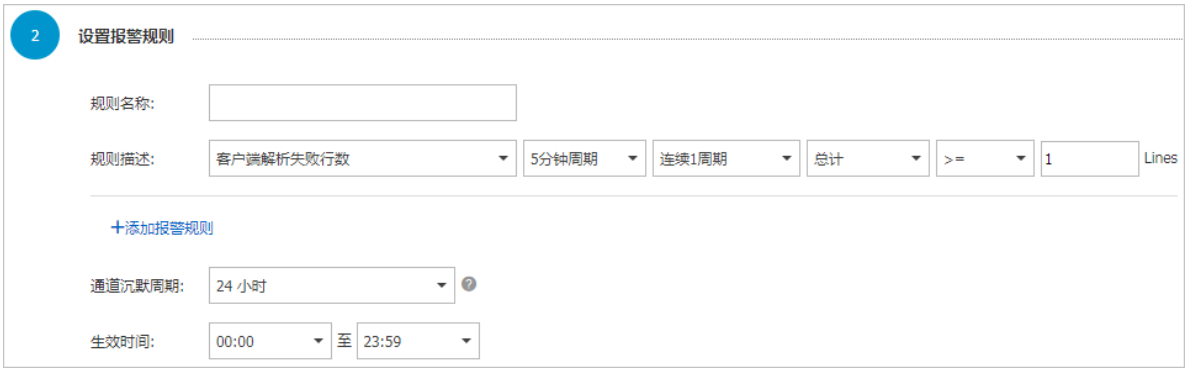
  Logtail客户端在运行过程中，可能会因设置不正确产生错误，例如某些日志格式不匹配、一个日志文件被重复收集等。为了及时发现这种情况，您可以对客户端解析失败行数、客户端错误次数等指标进行监控，以便及时发现这类问题。

  该监控的监控规则设置如下：

  您可以根据需要选择客户端解析失败行数或客户端错误次数选项，并配置统计周期、统计方法等规则项。除此之外，还可以根据Logtail其他错误项进行报警，第一时间发现各类日志收集过程中发现的问题。

  下图示例表示：五分钟内客户端解析失败行数达到一行以上即发送报警，持续24H监控。

  图 3-4: 监控Logtail日志收集状态



- 监控Shard资源使用状态

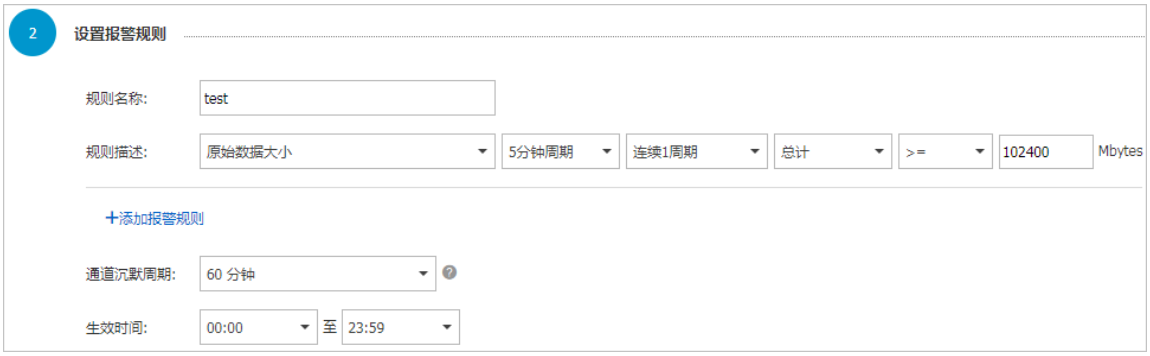
  Logstore下每个Shard提供5MB/s（500次/s）写入能力，这个数值对于大部分用户而言都是足够的，在超过时日志服务会尽可能去服务（非拒绝）您的请求，但在高峰期间不保证超出部分

的可用性。您可以设置Logstore出入流量报警以检测该情况。如果您的日志量非常大，需要添加更多Shard，请及时在控制台中进行调整。

- 设置流量预警

  设置规则名称，并配置规则描述为原始数据大小。您可以在此处设置统计周期和统计方法，如需超过100GB/5Min后进行报警，请设置5分钟、总计、>=、102400，表示5分钟内总计流量超出102400Mbytes时进行报警。

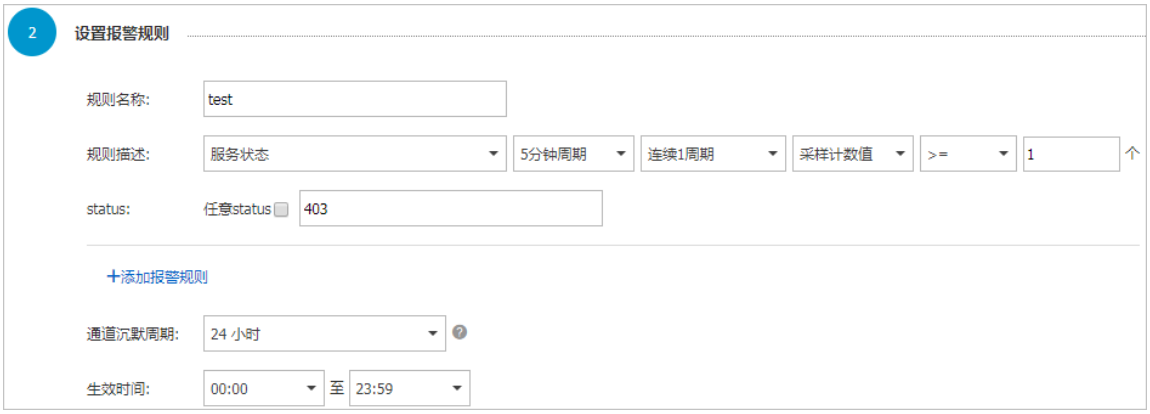
  图 3-5: 设置流量预警



- 设置服务状态报警

  设置规则名称，并配置规则描述为服务状态。您可以在此处设置统计周期和统计方法，如您需要在5分钟内出现1个以上403服务状态时收到报警，请参考下图配置。

  图 3-6: 设置服务状态报警



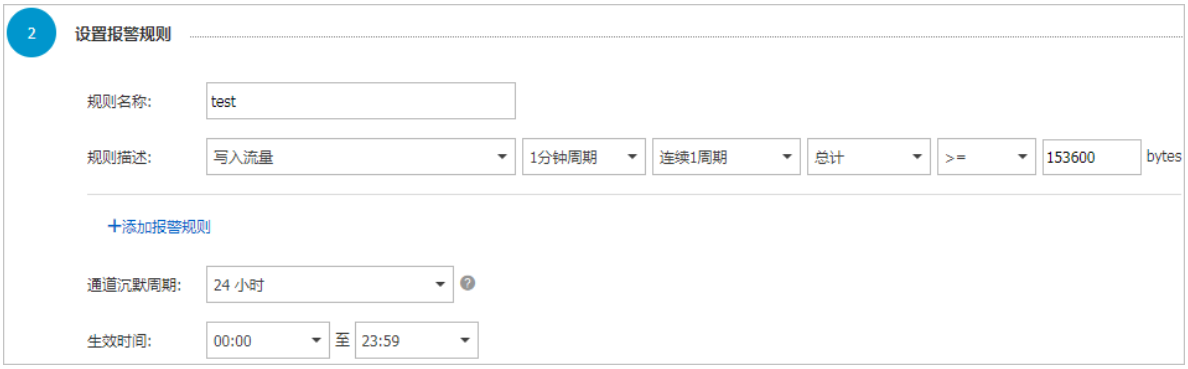
- 监控Project的写入流量

  每个Project默认写入限制为30 GB/分钟（原始数据大小），这个数值主要目的是为了保护用户因程序错误产生大量日志，在一般场景中对于大部分用户都是足够的。如果您的日志量非常大，可能会超过限制，可以通过工单联系我们调整大这个数值。

  您可以按照以下示例设置Project Quota的监控策略。

  该示例表示：当您的五分钟写入流量超过了150 GB时，为您发送提醒。

  图 3-7: 监控Project的写入流量



4 最佳实践

4.1 开通、监控和消费服务日志

日志服务提供服务日志功能，帮助您实时掌握日志服务的使用状况、提高运维效率。

开通服务日志

服务日志分为详细日志和重要日志（包括Logtail相关日志，消费组延时日志和计量日志等），分别保存在internal-operation_log和internal-diagnostic_log中。internal-diagnostic_log不计费，internal-operation_log和普通Logstore一样计费。详细日志对应用户的每次操作或者API请求，当读写请求较多时会产生较多的操作日志。

用户可以根据需要开通服务日志：

日志存储位置建议使用推荐配置，将同一个地域的服务日志保存到相同的Project中，便于管理和统计。

监控Logtail心跳

Logtail 安装好之后，可以通过服务日志Logtail的状态日志，来检查Logtail的工作状态。

Logtail状态日志保存在internal-diagnostic_log 中，进入internal-diagnostic_log搜索：`__topic__: logtail_status`，即可搜索Logtail状态日志。我们可以统计最近一段时间内的机器个数，用于和Logtail应用机器组的总机器数比较，并配置告警规则，如果低于总机器数则触发告警。

- **查询语句**

```
__topic__: logtail_status | SELECT COUNT(DISTINCT ip) as ip_count
```

- **查询截图**

- 配置告警规则（此处假设总机器数为100）：

如果产生告警，则用户可以前往控制台查看机器组的状态，检查哪些机器心跳异常。

查看存储空间等计量数据

用户将日志写入日志服务后，通过日志流量，读写次数，存储空间等计量信息，可以看到日志服务的使用状况和各个计费项的详细内容。

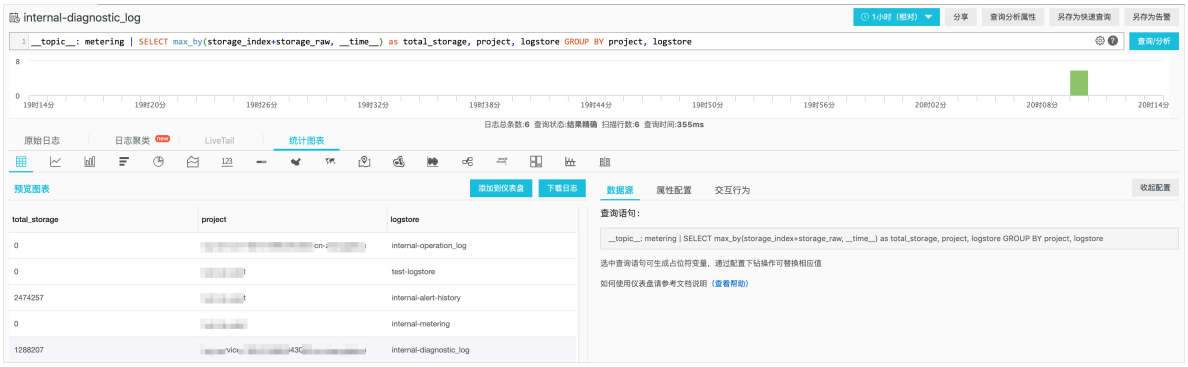
每个Logstore每小时产生一条计量日志，包含统计时间段内的读写流量，读写次数以及当前时间点的原始日志和索引的存储空间大小，各个字段解释可以参考计量日志字段解释。日志服务的计量日志保存在服务日志的internal-diagnostic_log中，通过搜索：`__topic__: metering`，即可搜索计量日志。

使用max_by函数计算各个Logstore的存储空间：

- 查询语句

```
__topic__: metering | SELECT max_by(storage_index+storage_raw,
__time__) as total_storage, project, logstore GROUP BY project,
logstore
```

- 查询截图



服务日志的默认仪表盘提供了丰富的计量日志相关的图表，参考计量日志报表。

查看消费组延时

日志写入日志服务后，除了普通的查询分析之外，可能还需要对日志进行消费。日志服务提供了多种语言实现的消费组协同消费库。

使用消费组消费时，日志的消费延时是用户比较关心的问题。通过观察消费延时，我们可以知道当前的消费进度，当延时较大时我们可以通过调整消费者个数等方式来改进消费速度。

消费组延时日志作为服务日志的一种，同样保存在internal-diagnostic_log中，每2分钟生成一次。通过搜索：`__topic__: consumergroup_log`，就可以搜索所有的消费组延时日志。

查询test-consumer-group这个消费组的消费延时：

- 查询语句

```
__topic__: consumergroup_log and consumer_group:  test-consumer-
group | SELECT max_by(fallbehind, __time__) as fallbehind
```

- 查询截图：

Logtail异常监控

Logtail是否正常运行关系到日志的完整性，如果用户能够及时发现Logtail的异常状况，就能尽快调整Logtail配置，使日志不会丢失。

Logtail的异常日志可以通过如下方式搜索：`__topic__: logtail_alarm`。

查询15分钟各种异常类型的次数：

- 查询语句

```
__topic__: logtail_alarm | select sum(alarm_count)as errorCount,
alarm_type  GROUP BY alarm_type
```

- 查询截图

Logstore写入流量监控

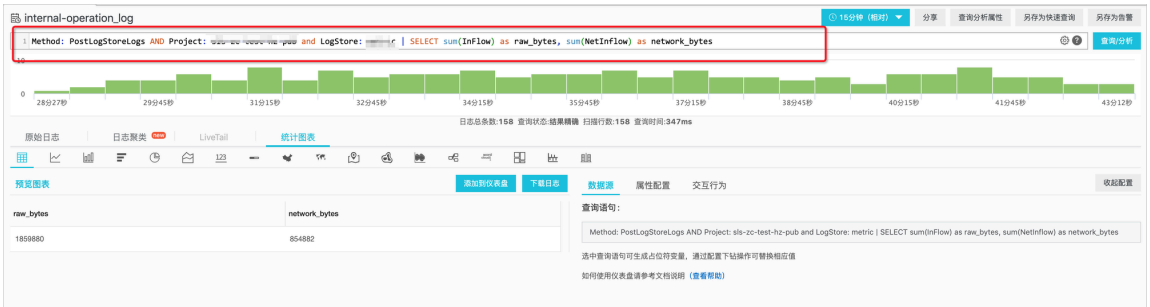
通过计量日志，我们可以统计一个小时内的读写流量，但是如果需要统计更精确的时间范围，如15分钟内的读写流量，则需要使用操作日志。用户的每次操作都对应一次请求，每次请求都会产生一条操作日志。以15分钟为例，我们只需要统计所有写入操作的流量总和即可。

1. 统计15分钟内Logstore写入的原始日志流量和压缩流量：

- 查询语句

```
Method: PostLogStoreLogs AND Project: my-project and LogStore: my
-logstore | SELECT sum(InFlow) as raw_bytes, sum(NetInflow) as
network_bytes
```

- 查询截图



2. 查询15分钟内写入网络流量的下降百分比：

- 查询语句

```
Method: PostLogStoreLogs AND Project: my-project and LogStore:
my-logstore | select round((diff[1]-diff[2])/diff[1],2) as rate
from (select compare(network_bytes, 900) as diff from (select sum(
NetInflow) as network_bytes from log))
```

- 查询截图

3. 创建告警：

设置大于50%时告警

操作审计

Project下所有资源的操作日志保存在internal-operation_log中，每行日志不仅会记录操作相关的信息，如创建机器组会记录机器组名称，操作Logstore会记录Logstore名称等，还会记录对应操作的用户信息。目前，用户信息可以分为如下类型：

类型	字段
主账户	· InvokerUid：主账户Aliuid。 · CallerType：Parent。
子账户	· InvokerUid：子账户Aliuid。 · CallerType：Subuser。

类型	字段
Sts	· InvokerUid：主账户Aliuid。 · CallerType：Sts。 · RoleSessionName：<session名称>。

通过上述信息就能知道不同操作日志对应的用户信息。