阿里云
日志服务

快速入门

文档版本：20190813

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

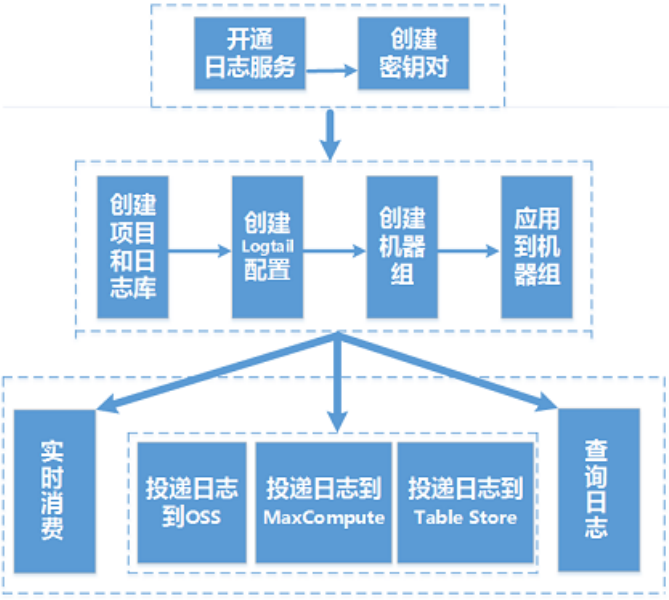
1 五分钟快速入门

本快速入门在Windows环境下，演示配置Logtail采集阿里云ECS日志并投递日志到MaxCompute的基本流程。本快速入门涉及采集日志、实时查询、投递日志等日志服务基本功能，是日志服务的入门级操作指南。

日志服务（Log Service）是阿里云提供的、针对海量日志收集、存储、查询的平台化服务。您可以使用日志服务来集中收集服务集群中所有的日志，并支持实时消费，实时查询和投递到OSS、MaxCompute等其他云产品做进一步分析。

日志服务操作流程

图 1-1: 操作流程



视频介绍：阿里云日志服务快速指南

准备开始

1. 开通日志服务。

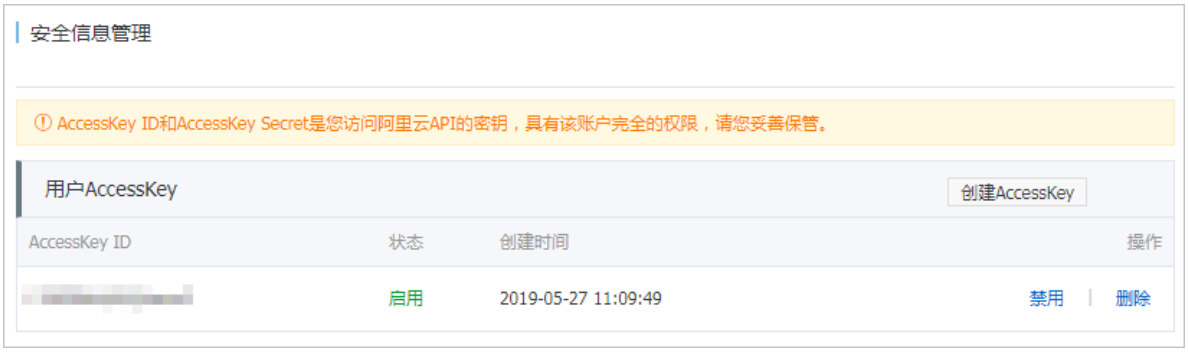
   使用注册成功的阿里云账号登录 日志服务产品页，单击 管理控制台。

2. （可选） 创建密钥对 。

   说明:

如您需要通过SDK写入数据，请创建主账号或子账号密钥对。通过Logtail采集日志不需要创建密钥对。

在日志服务管理控制台，将鼠标移至页面右上角您的用户名上方，在显示的菜单中单击AccessKey管理。创建密钥对（AccessKey），确认状态已设置为启用。



3. 创建项目。

当您第一次进入日志服务管理控制台，系统会提示您创建一个项目（Project）。您也可以通过单击创建Project进行操作。

创建Project需要指定Project名称与所属区域，请根据您的实际需求进行创建。

4. 创建日志库。

在Project创建完成的同时，系统会提示您创建一个日志库（以下称为Logstore）。您也可以进入该Project，通过单击页面左侧搜索框后的 + 进行操作。创建Logstore需要指定如何使用这些日志。

配置数据接入向导

在日志服务管理控制台单击目标Project进入概览页面。单击页面右上方的接入数据按钮。

日志服务支持多种云产品、自建软件和自定义数据的日志收集。本文以采集文本日志为例，详细步骤及说明请参考采集文本日志。

1. 选择数据类型。

   单击接入数据中的单行-文本日志。

2. 选择日志空间

   可以选择已有的Logstore，也可以新建Project和Logstore。

3. 创建机器组。

   在创建机器组之前，您需要首先确认已经安装了Logtail。

   - 集团内部机器：默认自动安装，如果没有安装，请根据界面提示进行咨询。
   - ECS机器： 勾选实例后单击安装进行一键式安装。Windows系统不支持一键式安装，请参考安装Logtail（Windows系统）手动安装。
   - 自建机器：请根据界面提示进行安装。或者参考安装Logtail（Linux系统）或安装Logtail（Windows系统）文档进行安装。

   安装完Logtail后单击确认安装完毕创建机器组。如果您之前已经创建好机器组，请直接单击使用现有机器组。

4. 机器组配置

   将源机器组中的机器组移动到应用机器组中。

5. Logtail配置。

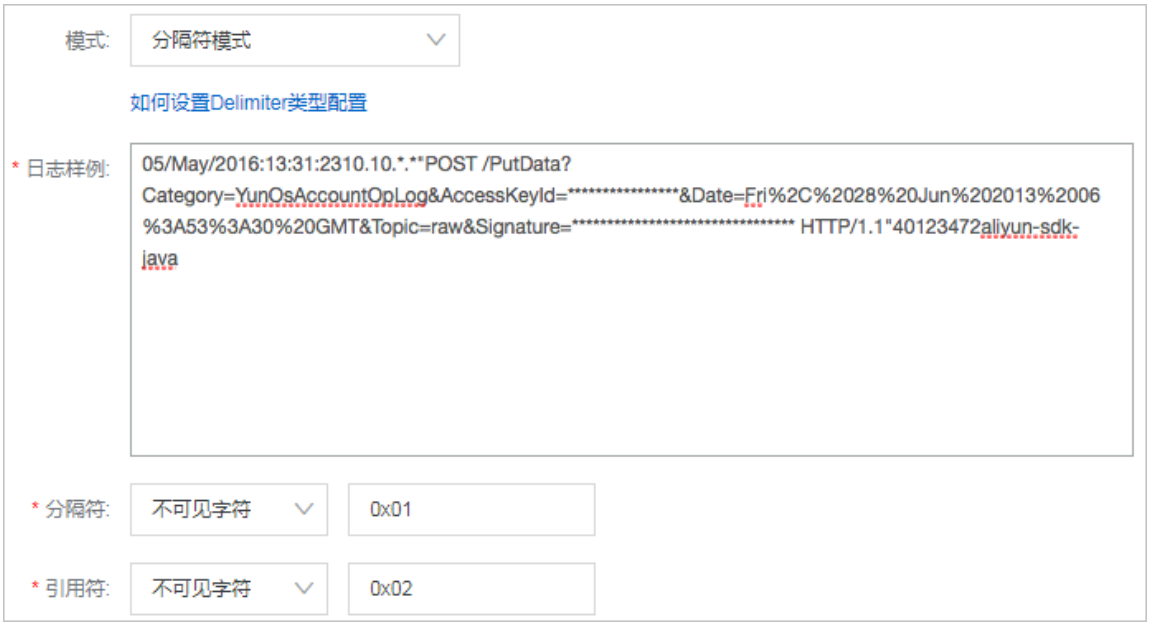
- 指定配置名称和日志路径。

  请按照页面提示填写配置名称、日志路径和日志文件名称。文件名称可以填写完整名称，也支持通配符模式匹配。

- 指定日志收集模式。

  日志服务目前支持极简模式、分隔符模式、JSON模式、完整正则模式等方式解析日志。本文以分隔符模式为例，关于收集模式的详细说明请参考采集文本日志和配置解析。

  图 1-2: 设置数据源



- 填写日志样例。

  指定日志收集模式为分隔符模式或完整正则模式时，需要您填写日志样例。日志服务支持在配置Logtail的同时，根据您选择的配置对日志样例尝试解析。如果解析失败，需要您修改分隔符配置或者正则表达式。请将需要解析的日志样例填写到对应位置。

- 指定分隔符和引用符。

  您可以指定分隔符为制表符、竖线、空格，也可以自定义分隔符。请根据您的日志格式选择正确的分隔符，否则日志数据会解析失败。

- 指定日志抽取结果中的Key。

填写日志样例并选择分隔符后，日志服务会按照您选择的分隔符提取日志字段，并将其定义为Value，您需要分别为Value指定对应的Key。

图 1-3: 日志内容抽取结果

- 酌情选择高级选项。

  一般情况下，请保持高级选项的默认配置。如您有其他需求，请参考采集文本日志中的高级配置相关说明。

6. 查询分析配置。

- 全文索引

  您可以选择开启全文索引，中文分词，并确认是否启用大小写敏感、确认分词符内容。

- 字段索引属性

  单击字段名称下方的加号新增一行，填写字段名称、类型、别名、大小写敏感、分词符、选择是否中文分词，开启统计。

> 说明:
>
> a. 全文索引和字段索引属性必须至少启用一种。同时启用时，以字段索引属性为准。
>
> b. 索引类型为long/double时，大小写敏感和分词符属性无效。
>
> c. 如何设置索引请参考开启并配置索引。个别字段为日志服务保留字段，请知悉。

d. 如需使用Nginx模板或消息服务模板，请在查询界面的查询分析属性中配置。

图 1-4: 查询分析

完成以上步骤后，日志服务开始收集阿里云ECS上的日志，您可以通过控制台和API/SDK对已收集的日志进行实时消费。

说明:

- Logtail配置推送生效时间最长需要3分钟，请耐心等待。
- Logtail收集错误可以参见诊断采集错误。

7. 日志投递。

   日志服务不仅支持对多种来源、格式的数据进行批量采集和管理维护，还支持将日志数据投递到OSS、MaxCompute等云产品进行计算分析。

   如您需要投递日志到OSS、MaxCompute，请单击对应的立即尝试按钮。

   本文档以投递MaxCompute为例。为把日志服务内的日志投递到MaxCompute，您需要首先准备好相应的MaxCompute环境。

   - 准备步骤

     a. 开通MaxCompute服务。您需要在阿里云管理控制台上启用MaxCompute服务。
     b. 创建存储投递日志的MaxCompute表。请参考 MaxCompute了解表的结构和相关注意事项。

   - 操作步骤

单击开始投递后跳转至LogHub —— 数据投递页面，在该页面需要配置投递大数据计算服务MaxCompute（原 ODPS）的相关内容。

图 1-5: 开启投递

> **说明：**
> `__source__`、`__time__`、`__topic__`、`__extract_others__`和`__partition_time__`是日志服务的系统保留字段，建议使用。对于映射配置的限制详情请参见通过日志服务投递日志到MaxCompute。

填写完成后，单击确认即可开始投递。

查询日志

在配置数据接入向导中配置了索引，可以对采集到的日志数据进行实时查询与SQL分析。

请返回至概览页面，单击日志库名称后的图标，选择查询分析后进入查询界面。输入关键字、Topic或者查询分析语句，并设定日志时间范围、单击查询/分析来查询日志，详细说明请参考查询日志。

另外，日志服务支持上下文查询、快速查询、快速分析等查询功能，还支持通过SQL语句获取多种格式的分析图表、通过仪表盘自制数据分析大盘等多种方式的可视化分析手段，更多功能请参考其他功能。

例如，查询指定时间段的PV，并以表格形式表示。

2 采集ECS日志

您可以通过Logtail采集ECS日志，本文为您介绍控制台方式的采集流程。

前提条件

- 已开通ECS和日志服务。
- 已创建了日志服务Project和Logstore。详细步骤请参见准备流程。

> 说明:
> 如果您的ECS为经典网络或VPC专有网络，请保证日志服务Project和ECS位于同一地域。

- 如果您的日志服务和ECS不在同一账号名下，则不能自动获取ECS对应的owner信息，需要配置主账号AliUid。

入门教程

集ECS日志的视频教程请单击视频教程查看。

配置流程

选择文件类型后进入数据采集配置流程：

采集步骤

1. 登录日志服务控制台，单击Project名称。
2. 选择数据类型。

   在概览页面，单击接入数据按钮，在接入数据页面中选择单行-文本文件。
3. 选择日志空间。

   可以选择已有Logstore，也可以新建Project和Logstore。

> 说明:
> 如果您的ECS为经典网络或VPC专有网络，请保证日志服务Project和ECS位于同一地域。

4. 创建机器组。

   在创建机器组之前，您需要首先确认已经安装了Logtail。

   - 集团内部机器：默认自动安装，如果没有安装，请根据界面提示进行咨询。
   - ECS机器： 勾选实例后单击安装进行一键式安装。Windows系统不支持一键式安装，请参见安装Logtail（Windows系统）手动安装。
   - 自建机器：请根据界面提示进行安装。或者参见安装Logtail（Linux系统）或安装Logtail（Windows系统）文档进行安装。

   安装完Logtail后单击确认安装完毕创建机器组。如果您之前已经创建好机器组，请直接单击使用现有机器组。

5. 配置机器组。

   将源机器组中的机器组移动到应用机器组中。

6. Logtail配置。

   详细数据源配置请参见文本日志，本文档以极简模式为例说明。

   填写ECS日志所在的文件路径，并单击下一步。

   图 2-1: 极简模式

   说明:

   - Logtail配置推送生效时间最长需要3分钟，请耐心等待。
   - Logtail收集错误可以参见诊断采集错误。

7. 查询分析配置。

   - 全文索引

     您可以选择开启全文索引，中文分词，并确认是否启用大小写敏感、确认分词符内容。

   - 字段索引属性

     单击字段名称下方的加号新增一行，填写字段名称、类型、别名、大小写敏感、分词符、选择是否中文分词，开启统计。

   说明:

- 全文索引和字段索引属性必须至少启用一种。同时启用时，以字段索引属性为准。
- 索引类型为long/double时，大小写敏感和分词符属性无效。
- 如何设置索引请参见开启并配置索引。个别字段为日志服务保留字段，请知悉。
- 如需使用Nginx模板或消息服务模板，请在查询界面的查询分析属性中配置。

完成以上步骤后，您已成功配置Logtail方式采集ECS日志。如果您需要为采集到的日志配置日志投递，请根据页面提示，在后续步骤中完成。

查看日志

成功配置后，您可以尝试重新登录ECS，或输入命令`echo "test message" >> /var/log/message`，本地`/var/log/message`文件会有新的日志产生，这部分日志会被Logtail采集到日志服务中。

请返回至概览页面，单击日志库名称后的 图标，选择查询分析或消费预览，可以查看Logtail采集到的日志数据。

图 2-2: 日志查看方式

图 2-3: 预览日志

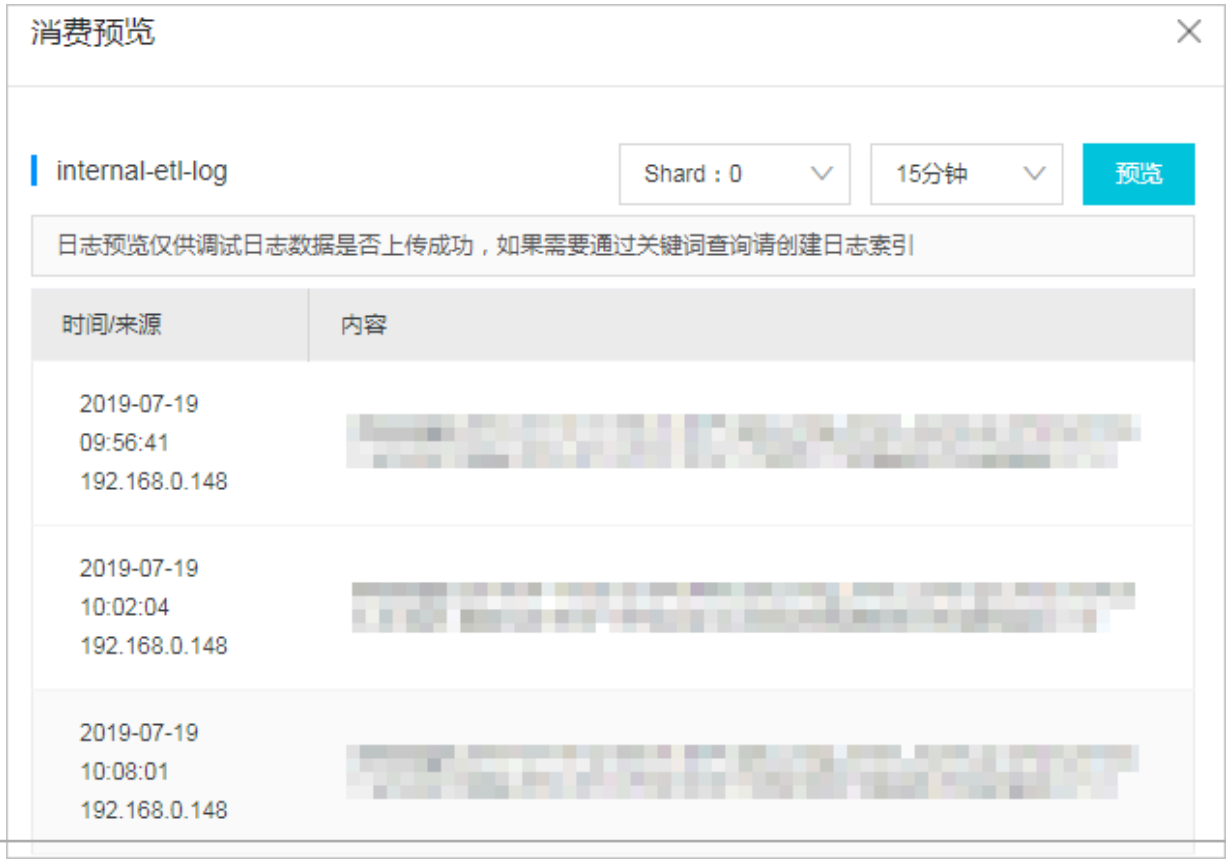


图 2-4: 检索日志

3 采集Kubernetes日志

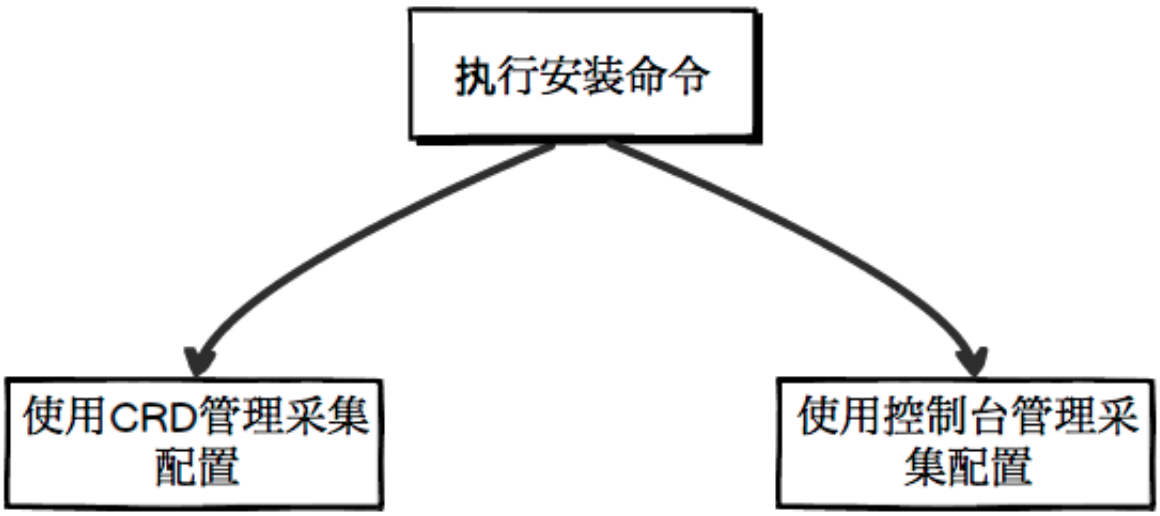
日志服务支持通过Logtail采集Kubernetes集群日志，并支持CRD（CustomResourceDefinition）进行采集配置管理。本文主要介绍如何安装并使用Logtail采集Kubernetes集群日志。

采集流程

1. 安装alibaba-log-controller Helm包。
2. 创建采集配置。

   根据您的需求选择使用控制台或CRD（CustomResourceDefinition）创建采集配置，本文以控制台为例。

图 3-1: 采集流程



视频教程：http://cloud.video.taobao.com/play/u/3220778205/p/1/e/6/t/1/50145086346.mp4

步骤1 安装软件包

1. 登录您的阿里云容器服务Kubernetes的Master节点。

   如何登录参考SSH密钥对访问Kubernetes集群。

2. 替换参数后执行以下安装命令。

将下述命令中的`${your_k8s_cluster_id}`替换为您的Kubernetes集群ID，并执行此命令。

```
 wget http://logtail-release-cn-hangzhou.oss-cn-hangzhou.aliyuncs
.com/kubernetes/alicloud-log-k8s-install.sh -O alicloud-log-k8s-
install.sh; chmod 744 ./alicloud-log-k8s-install.sh; sh ./alicloud-
log-k8s-install.sh ${your_k8s_cluster_id}
```

安装示例

执行安装命令，回显信息如下：

```
[root@iZbp******biaZ ~]# wget http://logtail-release-cn-hangzhou.oss
-cn-hangzhou.aliyuncs.com/kubernetes/alicloud-log-k8s-install.sh -O
alicloud-log-k8s-install.sh; chmod 744 ./alicloud-log-k8s-install.sh;
sh ./alicloud-log-k8s-install.sh c12ba20**************86939f0b
....
....
....
alibaba-cloud-log/Chart.yaml
alibaba-cloud-log/templates/
alibaba-cloud-log/templates/_helpers.tpl
alibaba-cloud-log/templates/alicloud-log-crd.yaml
alibaba-cloud-log/templates/logtail-daemonset.yaml
alibaba-cloud-log/templates/NOTES.txt
alibaba-cloud-log/values.yaml
NAME: alibaba-log-controller
LAST DEPLOYED: Wed May 16 18:43:06 2018
NAMESPACE: default
STATUS: DEPLOYED


RESOURCES:
==> v1beta1/ClusterRoleBinding
NAME AGE
alibaba-log-controller 0s


==> v1beta1/DaemonSet
NAME DESIRED CURRENT READY UP-TO-DATE AVAILABLE NODE SELECTOR AGE
logtail 2 2 0 2 0  0s


==> v1beta1/Deployment
NAME DESIRED CURRENT UP-TO-DATE AVAILABLE AGE
alibaba-log-controller 1 1 1 0 0s


==> v1/Pod(related)
NAME READY STATUS RESTARTS AGE
logtail-ff6rf 0/1 ContainerCreating 0 0s
logtail-q5s87 0/1 ContainerCreating 0 0s
alibaba-log-controller-7cf6d7dbb5-qvn6w 0/1 ContainerCreating 0 0s


==> v1/ServiceAccount
NAME SECRETS AGE
```

```
alibaba-log-controller 1 0s


==> v1beta1/CustomResourceDefinition
NAME AGE
aliyunlogconfigs.log.alibabacloud.com 0s


==> v1beta1/ClusterRole
alibaba-log-controller 0s




[SUCCESS] install helm package : alibaba-log-controller success.
```

您可以使用`helm status alibaba-log-controller`查看Pod当前状态，若状态全部成功后，表示安装成功。

安装成功后，日志服务会自动为您创建k8s-log开头的Project，在日志服务控制台搜索k8s-log关键字即可查看。

步骤2 创建采集配置

在该步骤中为您演示如何使用控制台创建Logstore并采集K8S所有容器的stdout标准输出。

1. 进入概览。

单击步骤1中自动创建的Project，进入概览页面。

2. 创建Logstore。

单击页面左侧搜索框后的 ，在弹出的页面中创建一个Logstore。

图 3-2: 创建Logstore

3. 创建采集配置。

a. 创建Logstore后，根据提示进入数据接入向导。

b. 在接入数据页面中选择Docker标准输出-容器。

c. 创建机器组

A. 在ECS上安装Logtail客户端

在ECS机器页面勾选实例后单击安装。Windows系统请参考安装Logtail（Windows系统）。

B. 创建机器组。

在创建机器组之前，您需要首先确认已经安装了Logtail。

- 集团内部机器：默认自动安装，如果没有安装，请根据界面提示进行咨询。
- ECS机器： 勾选实例后单击安装进行一键式安装。Windows系统不支持一键式安装，请参考安装Logtail（Windows系统）手动安装。
- 自建机器：请根据界面提示进行安装。或者参考安装Logtail（Linux系统）或安装Logtail（Windows系统）文档进行安装。

安装完Logtail后单击确认安装完毕创建机器组。如果您之前已经创建好机器组，请直接单击使用现有机器组。

d. 机器组配置。

将源机器组中的机器组移动到应用机器组中。

e. 数据源设置

在配置页面中直接单击下一步。无需做任何修改，即可实现采集所有容器的stdout文件。

图 3-3: Docker标准输出

f. 查询分析配置

默认开启全文索引，您也可以手动增加字段索引或者自动生成索引。

数据采集配置已经完成，如您需要配置数据投递，请根据页面提示在后续步骤中填写配置。

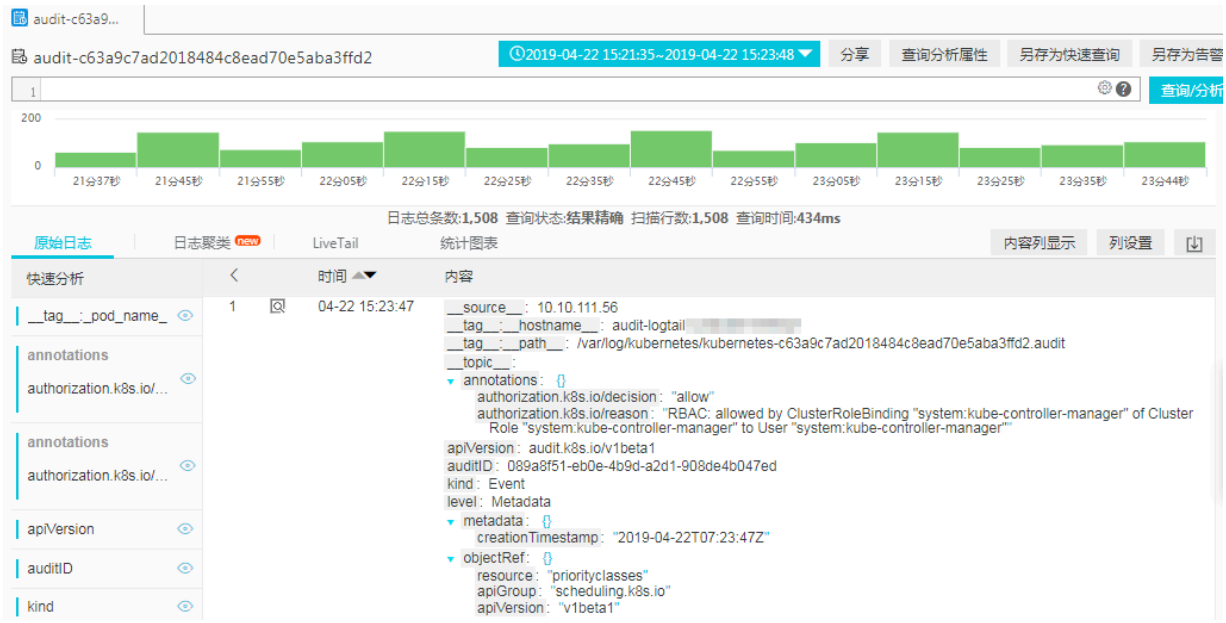
查看采集到的日志数据

完成数据采集配置后，若您的集群中有容器在输入stdout，则一分钟后即可采集到stdout日志。请返回至概览页面，单击日志库名称后的图标，选择消费预览可以快速预览当前采集到的日志数据；选择查询分析对采集到的日志数据进行自定义查询和分析。

图 3-4: 预览和查询

如下图所示，查询分析页面中可直接单击日志中的关键字进行快速查询，也可以再查询输入框中输入指定的关键字进行查询。

图 3-5: 查询日志



其他采集配置方式

以上示例为控制台方式配置标准输出的采集方式，其他采集方式及详细配置请参考：

控制台配置方式

控制台配置方式请参见：

- 容器内文本文件（推荐）
- 容器标准输出（推荐）
- 宿主机文本文件

  默认宿主机根目录挂载到Logtail容器的`/logtail_host`目录，配置路径时，您需要加上此前缀。例如需要采集宿主机上`/home/logs/app_log/`目录下的数据，配置页面中日志路径设置为`/logtail_host/home/logs/app_log/`。
- 自定义插件

CRD配置方式

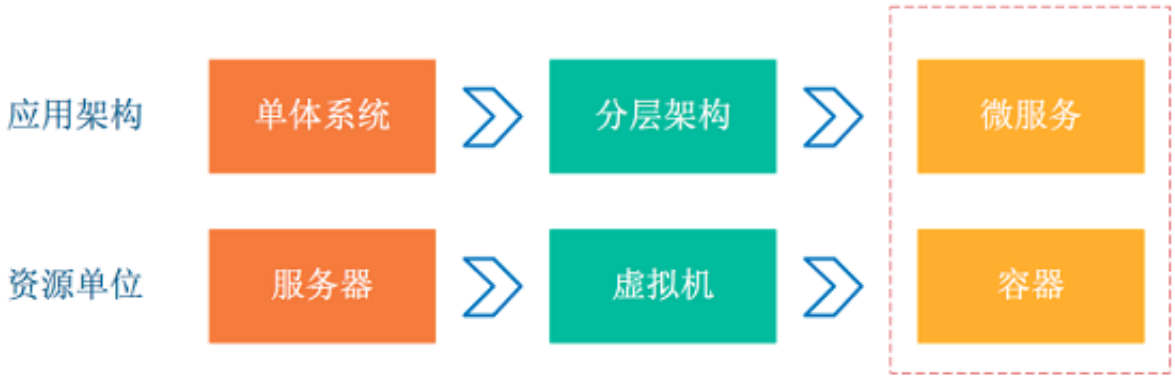
CRD（CustomResourceDefinition）配置方式请参考Kubernetes-CRD配置日志采集。

4 分析Log4j日志

近几年来，无状态编程、容器、Serverless编程方式的诞生极大提升了软件交付与部署的效率。在架构的演化过程中，可以看到两个变化：

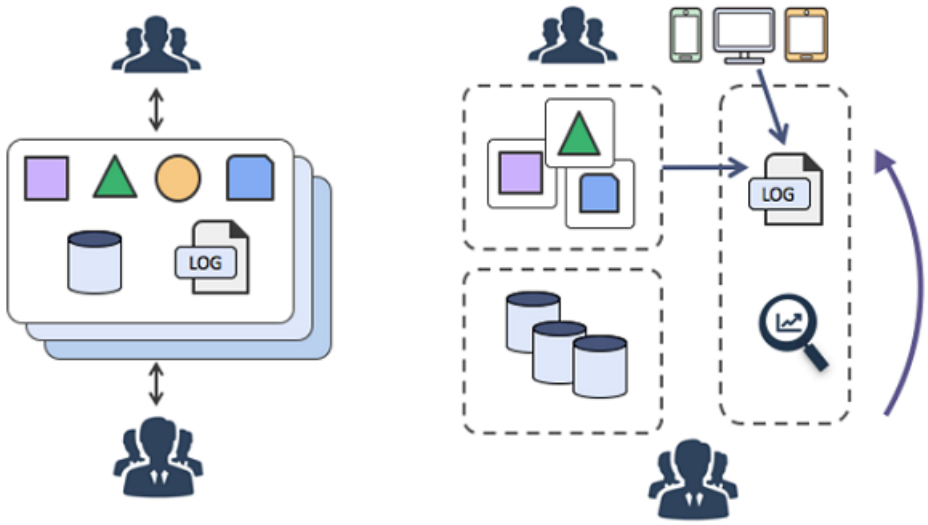
- 应用架构开始从单体系统逐步转变微服务，其中的业务逻辑随之就会变成微服务之间调用与请求。
- 资源角度来看，传统服务器这个物理单位也逐渐淡化，变成了看不见摸不到的虚拟资源模式。

图 4-1: 架构演化



从以上两个变化可以看到这种弹性、标准化架构背后，原先运维与诊断的需求也变得越来越复杂。在10年前我们可以快速登录到服务器上获取日志，Attach进程的模式再也不存在，面对我们的更多是一个标准化的“黑盒”。

图 4-2: 变化趋势



为了应对这种变化趋势，诞生一系列面向DevOps诊断与分析的工具。例如集中式监控、集中式日志系统、以及SaaS化的各种部署、监控等服务。

日志中心化解决的是以上这个问题，既应用产生日志后实时（或准实时）传输到中心化的节点服务器，例如Syslog，Kafka，ELK，Hbase进行集中式存储是一些常见的模式。

中心化的优势

- 使用方便：在无状态应用中最麻烦的要属Grep日志了，集中式存储只要运行一个Search命令就能够替代原先漫长的过程。
- 存储与计算分离：定制机器硬件时无需为日志存储考虑空间大小。
- 更低成本：集中式日志存储可以削峰填谷，预留更高水位。
- 安全：当发生黑客入侵以及灾难时，关键数据留作取证。

图 4-3: 中心化的优势

采集端 （Java系列）

日志服务提供30+种数据采集方式，针对服务器、移动端、嵌入式设备及各种开发语言都提供完整的接入方案。对Java 开发者而言，需要熟悉的日志框架 Log4j、Log4j2、Logback Appender。

对Java应用而言，目前有两种主流的日志采集方案：

- Java程序将日志落盘，通过Logtail进行实时采集。
- Java程序直接配置日志服务提供的Appender，当程序运行时，实时将日志发往服务端。

两者的差别如下：

-	日志落盘+Logtai采集	Appender 直接发送
实时性	日志落文件，通过Logtail采集	直接发送
吞吐量	大	大
断点续传	支持，取决于Logtail 配置大小	支持，取决于内存大小
关心应用位置	需要，配置采集机器组时	不需要，主动发送

-	日志落盘+Logtai采集	Appender 直接发送
本地日志	支持	支持
关闭采集	Logtail移除配置	修改Appender配置，重启应用

通过Appender可以在不改任何代码情况下，通过Config就能够非常容易完成日志实时采集工作，日志服务提供的Java系列Appender有如下几大优势：

- 无需修改程序，修改配置即生效。
- 异步化 + 断点续传：I/O不影响主线程，支持一定网络和服务容错。
- 高并发设计：满足海量日志写入需求。
- 支持上下文查询功能：写入支持在服务端还原原始进程中精确上下文（前后N条日志）。

Appender 简介与使用

目前提供Appender如下，底层均使用 aliyun-log-producer-java 完成数据的写入。

- aliyun-log-log4j-appender
- aliyun-log-log4j2-appender
- aliyun-log-logback-appender

这四者的差别如下：

Appender名称	说明
aliyun-log-log4j-appender	针对 log4j 1.x 开发的 appender，如果您的应用程序正在使用 log4j 1.x 日志框架，建议您选择该 appender。
aliyun-log-log4j2-appender	针对 log4j 2.x 开发的 appender，如果您的应用程序正在使用 log4j 2.x 日志框架，建议您选择该 appender。
aliyun-log-logback-appender	针对 logback 开发的Appender，如果您的应用程序正在使用 logback 日志框架，建议您选择该Appender。
aliyun-log-producer-java	针对 Java 应用程序编写的高并发写 LogHub 类库，目前提供的Appender底层均使用它完成数据写入。它具有很高的灵活性，可以让您指定写入 Loghub 中数据的字段和格式，如果您发现我们提供的Appender无法满足您的业务需求，您可以基于它开发适合您的日志采集程序。

步骤1 接入 Appender

可参考 aliyun-log-log4j-appender 的配置步骤部分接入Appender。

配置文件`log4j.properties`的内容如下：

```
log4j.rootLogger=WARN,loghub
log4j.appender.loghub=com.aliyun.openservices.log.log4j.LoghubAppender
#日志服务的project名，必选参数
log4j.appender.loghub.projectName=[your project]
#日志服务的logstore名，必选参数
log4j.appender.loghub.logstore=[your logstore]
#日志服务的http地址，必选参数
log4j.appender.loghub.endpoint=[your project endpoint]
#用户身份标识，必选参数
log4j.appender.loghub.accessKeyId=[your accesskey id]
log4j.appender.loghub.accessKey=[your accesskey]
```

步骤2 查询与分析

通过上述方式配置好appender 后，Java 应用产生的日志会被自动发往日志服务。可以通过 LogSearch/Analytics 对这些日志实时查询和分析。本文提供的样例的日志格式如下：

- 记录用户登录行为的日志：

```
level:  INFO
location:  com.aliyun.log4jappendertest.Log4jAppenderBizDemo.login(
Log4jAppenderBizDemo.java:38)
message:  User login successfully. requestID=id4 userID=user8
thread:  main
time:  2018-01-26T15:31+0000
```

- 记录用户购买行为的日志：

```
level:  INFO
location:  com.aliyun.log4jappendertest.Log4jAppenderBizDemo.order(
Log4jAppenderBizDemo.java:46)
message:  Place an order successfully. requestID=id44 userID=user8
itemID=item3 amount=9
thread:  main
time:  2018-01-26T15:31+0000
```

步骤3 开启查询分析

若要对数据进行查询和分析，需要首先开启查询分析功能。开启步骤如下：

1. 登录日志服务控制台，单击Project名称。
2. 单击目标日志库名称后的  图标，选择查询分析进入查询分析界面。
3. 单击右上角的查询分析属性 > 设置。

4. 进入查询分析页面，为下列字段开启查询。

图 4-4: 指定字段查询

步骤4 分析日志

视频教程

1. 统计过去1小时发生Error最多的3个位置。

```
level: ERROR | select location ,count(*) as count GROUP BY  location
  ORDER BY count DESC LIMIT 3
```

2. 统计过去15分钟各种日志级别产生的日志条数。

```
| select level ,count(*) as count GROUP BY level ORDER BY count DESC
```

3. 日志上下文查询。

   对于任意一条日志，能够精确还原原始日志文件上下文日志信息。详情请参考上下文查询。

4. 统计过去1小时，登录次数最多的三个用户。

```
login | SELECT regexp_extract(message, 'userID=(?<userID>[a-zA-Z\d
]+)', 1) AS userID, count(*) as count GROUP BY userID ORDER BY count
 DESC LIMIT 3
```

5. 统计过去15分钟，每个用户的付款总额。

```
order | SELECT regexp_extract(message, 'userID=(?<userID>[a-zA-Z\
d]+)', 1) AS userID, sum(cast(regexp_extract(message, 'amount=(?<
amount>[a-zA-Z\d]+)', 1) AS double)) AS amount GROUP BY userID
```

5 采集并分析Nginx访问日志

日志服务支持通过数据接入向导配置采集Nginx日志，并自动创建索引和Nginx日志仪表盘，帮助您快速采集并分析Nginx日志。

许多个人站长选取了Nginx作为服务器搭建网站，在对网站访问情况进行分析时，需要对Nginx访问日志统计分析，从中获取网站的访问量、访问时段等访问情况。传统模式下利用CNZZ等方式，在前端页面插入js，用户访问的时候触发js，但仅能记录访问请求。或者利用流计算、离线统计分析Nginx访问日志，但需要搭建一套环境，并且在实时性以及分析灵活性上难以平衡。

日志服务在支持查询分析实时日志功能，同时提供Nginx日志仪表盘（Dashboard），极大的降低了Nginx访问日志的分析复杂度，可以用于便捷统计网站的访问数据。本文档以分析Nginx访问日志为例，介绍日志分析功能在分析Nginx访问日志场景下的详细步骤。

应用场景

个人站长选取Nginx作为服务器搭建了个人网站，需要通过分析Nginx访问日志来分析网站的PV、UV、热点页面、热点方法、错误请求、客户端类型和来源页面制表，以分析、评估网站的访问情况。

Nginx日志格式

为了更好满足分析场景，推荐Nginx日志格式采用如下`log_format`配置。

```
    log_format  main  '$remote_addr - $remote_user [$time_local] "$
request" $http_host '
                      '$status $request_length $body_bytes_sent "$
http_referer" '
                      '"$http_user_agent"  $request_time $upstream_r
esponse_time';
```

各字段含义如下：

字段	含义
remote_addr	客户端地址
remote_user	客户端用户名
time_local	服务器时间
request	请求内容，包括方法名、地址和http协议
http_host	用户请求时使用的http地址
status	返回的http状态码
request_length	请求大小

字段	含义
body_bytes_sent	返回的大小
http_referer	来源页
http_user_agent	客户端名称
request_time	整体请求延时
upstream_response_time	上游服务的处理延时

采集Nginx日志

采集日志前，请确认您已创建Project和创建Logstore。

1. 登录日志服务控制台，单击Project名称。
2. 在对应日志库下，单击数据接入后的加号。

   您也可以直接单击概览页面右上角的接入数据按钮，在后续步骤中再选择对应Logstore。

   图 5-1: 接入数据

3. 选择数据类型为Nginx-正则。

   日志服务提供多种数据类型接入（云产品、自定义代码、自建开源/商业软件等），分析NGINX访问日志请选择自建开源/商业软件 > Nginx-文本日志。

4. 创建机器组。

   在创建机器组之前，您需要首先确认已经安装了Logtail。

   - 集团内部机器：默认自动安装，如果没有安装，请根据界面提示进行咨询。
   - ECS机器： 勾选实例后单击安装进行一键式安装。Windows系统不支持一键式安装，请参考安装Logtail（Windows系统）手动安装。
   - 自建机器：请根据界面提示进行安装。或者参考安装Logtail（Linux系统）或安装Logtail（Windows系统）文档进行安装。

   安装完Logtail后单击确认安装完毕创建机器组。如果您之前已经创建好机器组，请直接单击使用现有机器组。

5. 机器组配置。

   将源机器组中的机器组移动到应用机器组中。

6. Logtail配置。

   a. 指定配置名称和日志路径。

   b. 将推荐的`log_format`日志格式填写到NGINX日志配置中。

   图 5-2: 配置数据源

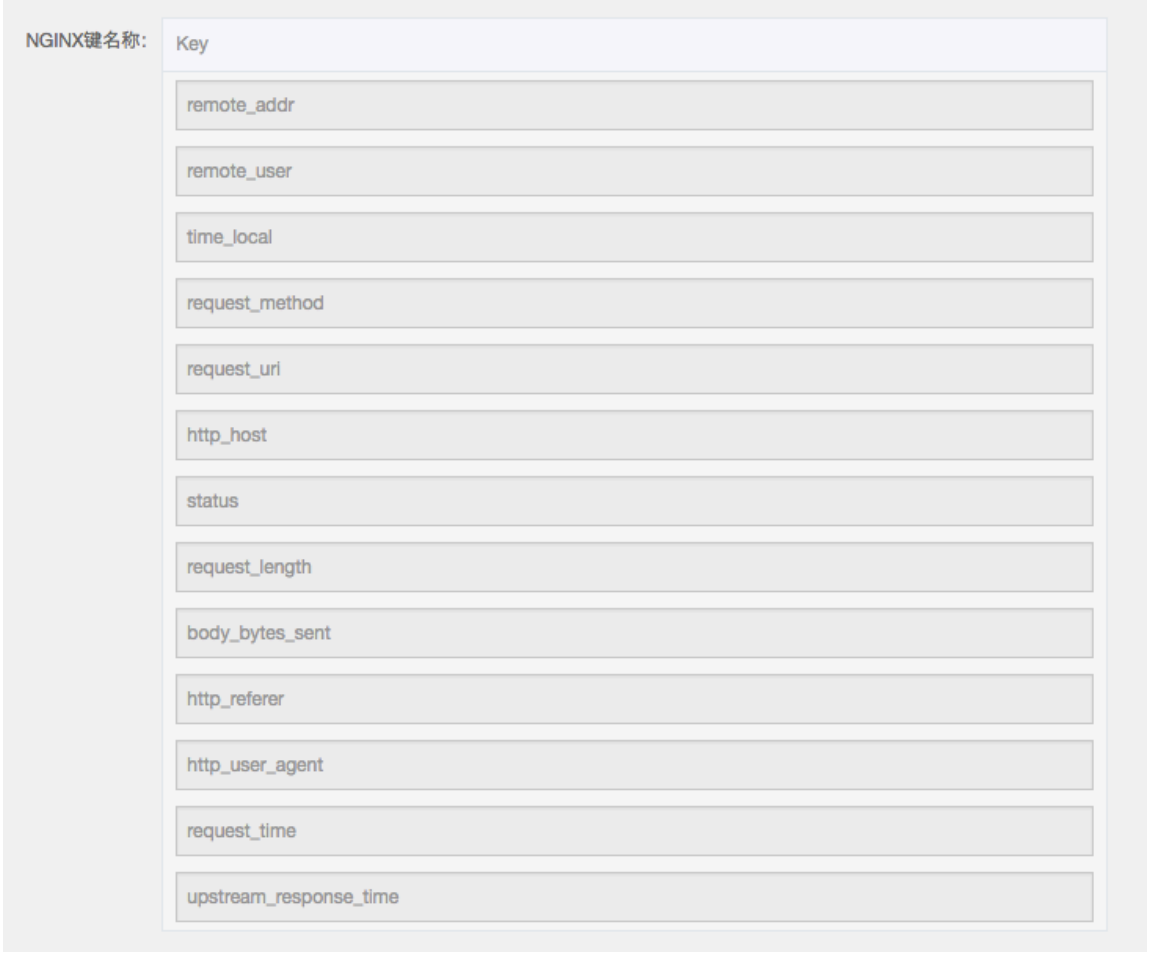
   c. 确认Nginx键名称。

   日志服务会自动提取出Nginx日志中的键名称，请确认是否正确。

   说明:

Nginx日志格式中的`$request`会被提取为`request_method`和`request_uri`两个键。

图 5-3: NGINX键名称



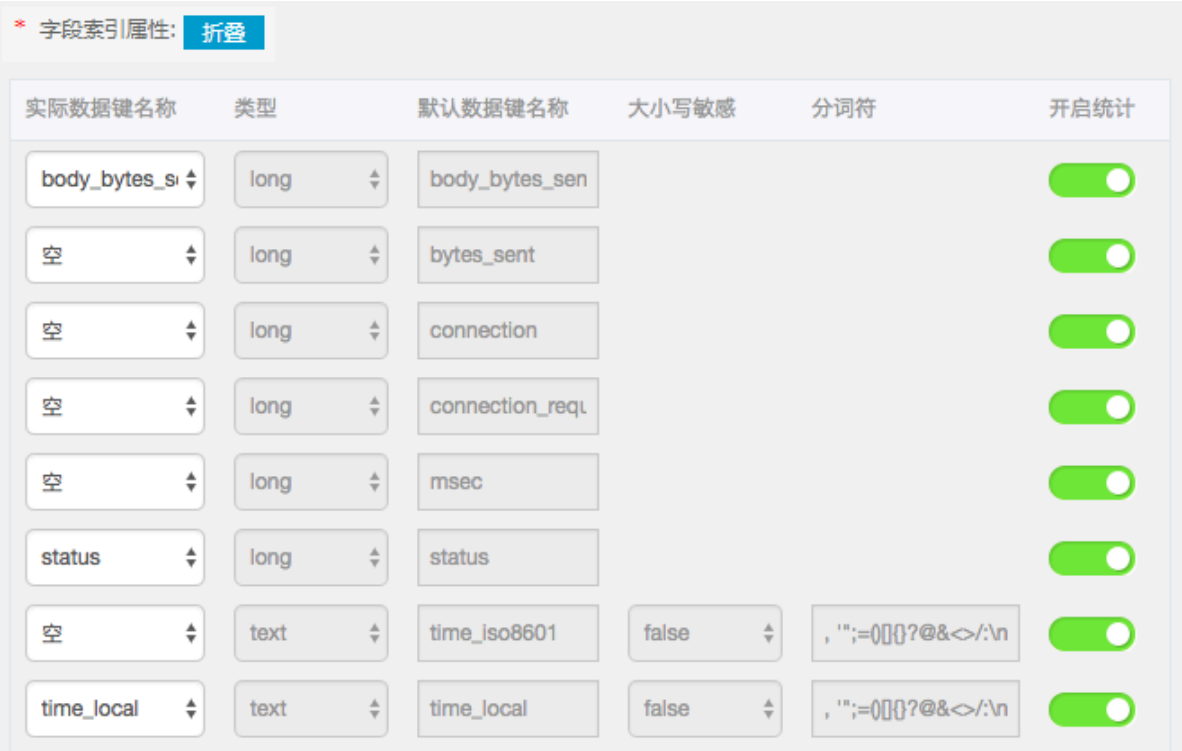
d. 确认是否丢弃解析失败日志，并单击下一步。

开启该功能后，解析失败的Nginx访问日志不上传到日志服务；关闭后，Nginx访问日志解析失败时上传原始Nginx访问日志。

7. 查询分析配置

默认已经设置好索引，如果您需要重新设置索引，请在查询分析页面选择查询分析属性 > 设置进行修改。

图 5-4: 字段索引属性



确保日志机器组心跳正常的情况下，可以预览采集上来的数据。

图 5-5: 预览

单击下一步，日志服务会为您设置好Nginx访问日志的索引属性并创建`nginx-dashboard`仪表盘以供分析使用。

说明:

Nginx日志采集配置生效时间最长需要3分钟，请耐心等待。

分析Nginx访问日志

开启索引后，日志服务默认生成Nginx访问日志的索引和仪表盘，可以通过以下方式分析Nginx访问日志。

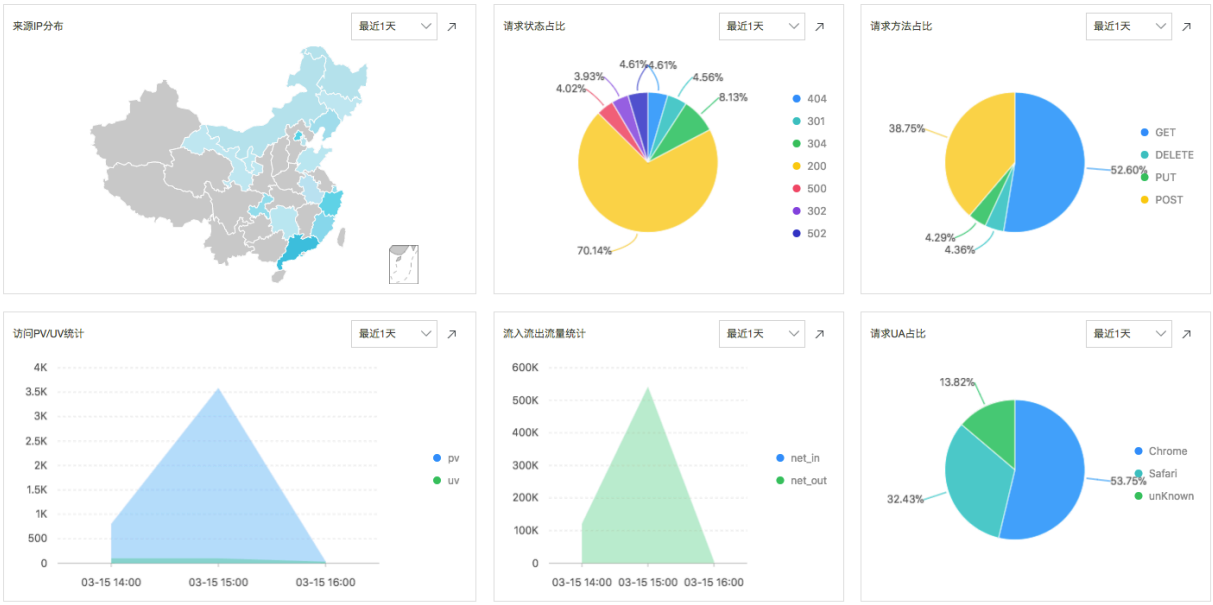
- 使用SQL语句分析Nginx访问日志：

  在日志服务查询分析页面输入查询分析语句，可以查看符合条件的Nginx原始日志，或查看可视化的分析结果。另外，查询分析页面还提供快速分析、快速查询等功能，详细说明请查看查询日志和Nginx访问日志诊断及优化。

- 查看预设仪表盘的分析数据，分析Nginx访问日志：

  日志服务预设的Nginx访问日志仪表盘中展示了各个分析指标的详细数据大盘，例如PV/UV统计等数据。关于如何使用仪表盘，请参考创建和删除仪表盘。

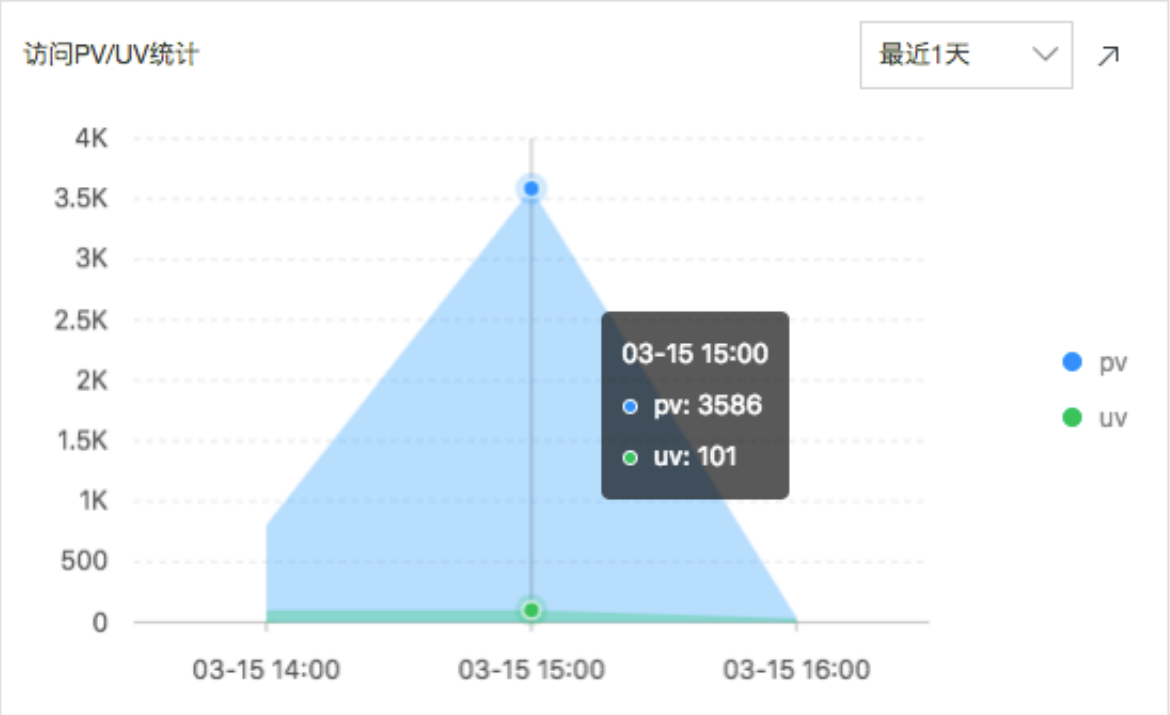
图 5-6: Nginx访问日志仪表盘



- **PV/UV统计(pv_uv)**

  统计最近一天的PV数和UV数。

  图 5-7: PV/UV统计



统计语句:

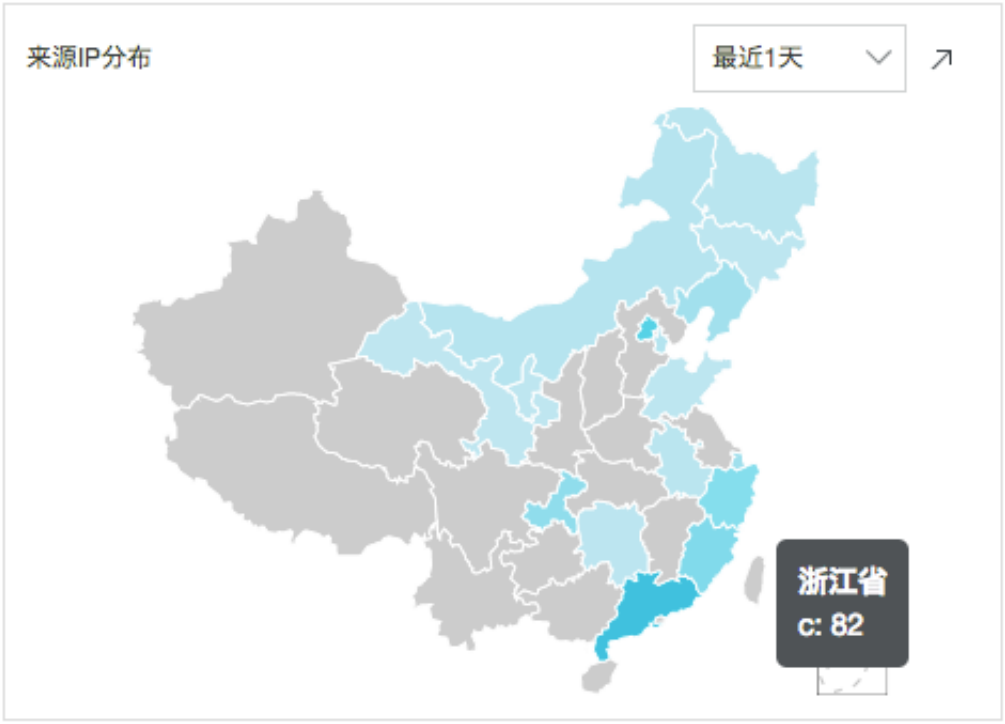
```
  * | select approx_distinct(remote_addr) as uv ,
          count(1) as pv ,
          date_format(date_trunc('hour', __time__), '%m-%d %H:%i')
as time
          group by date_format(date_trunc('hour', __time__), '%m-%d %
H:%i')
          order by time
```

```
        limit 1000
```

- **访问地域分析(ip_distribution)**

  统计访问IP来源情况。

  图 5-8: 访问地域分析



  统计语句：

```
* | select count(1) as c,
    ip_to_province(remote_addr) as address
```

```
    group by ip_to_province(remote_addr) limit 100
```

- **访问前十地址(top_page)**

  统计最近一天访问PV前十的地址。

  图 5-9: 统计访问

访问前十地址 最近1天

path	pv
/url7	157
/url1	156
/url10	153
/url4	149
/url2	143
/url9	143
/url6	141
/url5	140
/url3	136
/url8	117

统计语句:

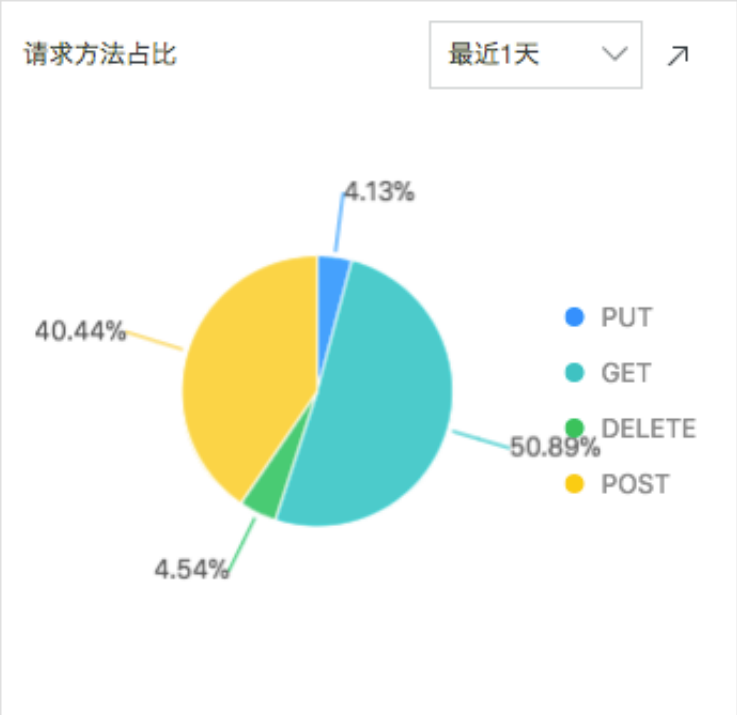
```
* | select split_part(request_uri,'?',1) as path,
    count(1) as pv
    group by split_part(request_uri,'?',1)
```

```
    order by pv desc limit 10
```

- **请求方法占比(http_method_percentage)**

  统计最近一天各种请求方法的占比。

  图 5-10: 请求方法占比



统计语句：

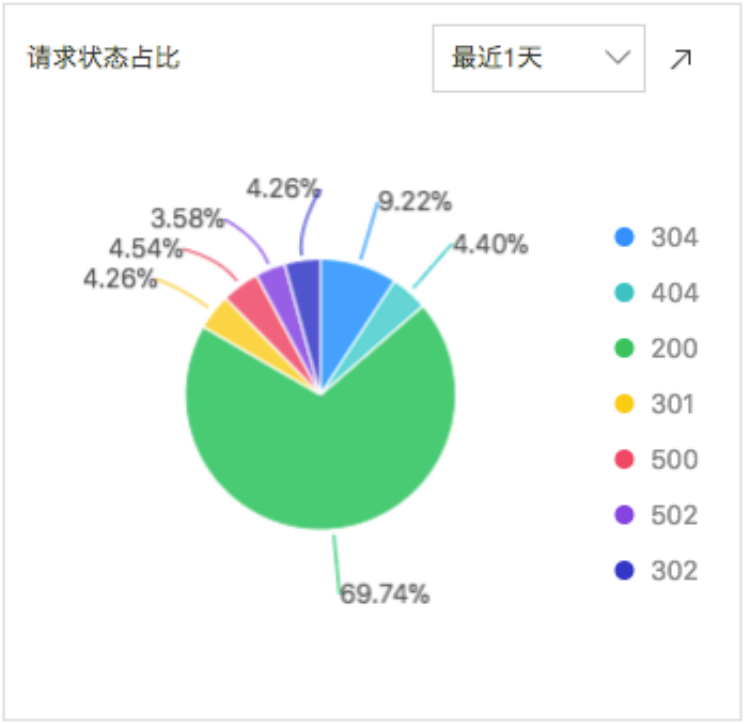
```
* | select count(1) as pv,
        request_method
```

```
        group by request_method
```

- **请求状态占比(http_status_percentage)**

  统计最近一天各种http状态码的占比。

  图 5-11: 请求状态占比



  统计语句:

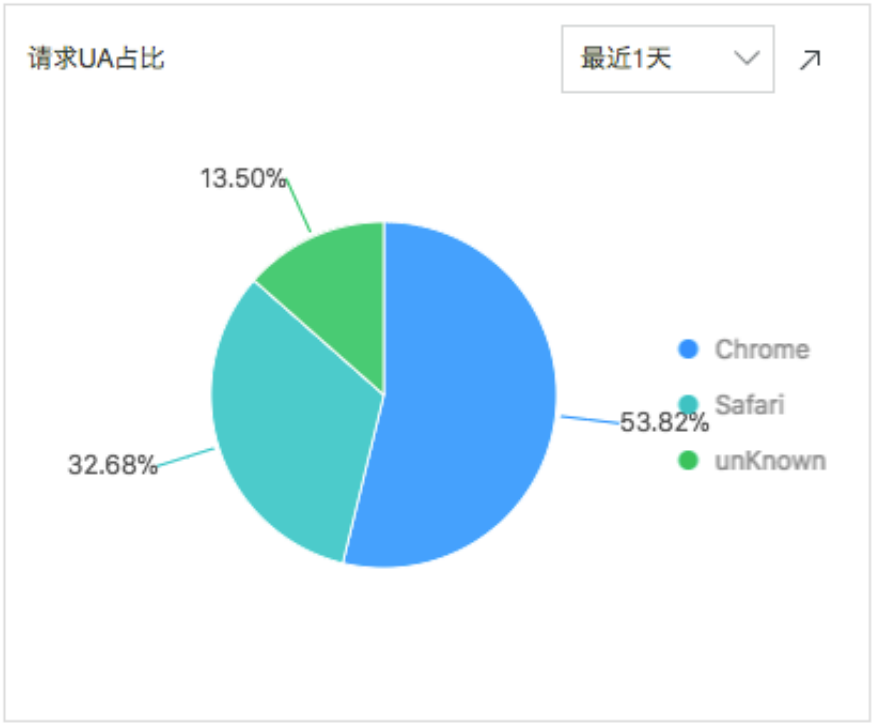
```
* | select count(1) as pv,
        status
```

```
        group by status
```

- **请求UA占比(user_agent)**

  统计最近一天各种浏览器的占比。

  图 5-12: 请求UA占比



统计语句：

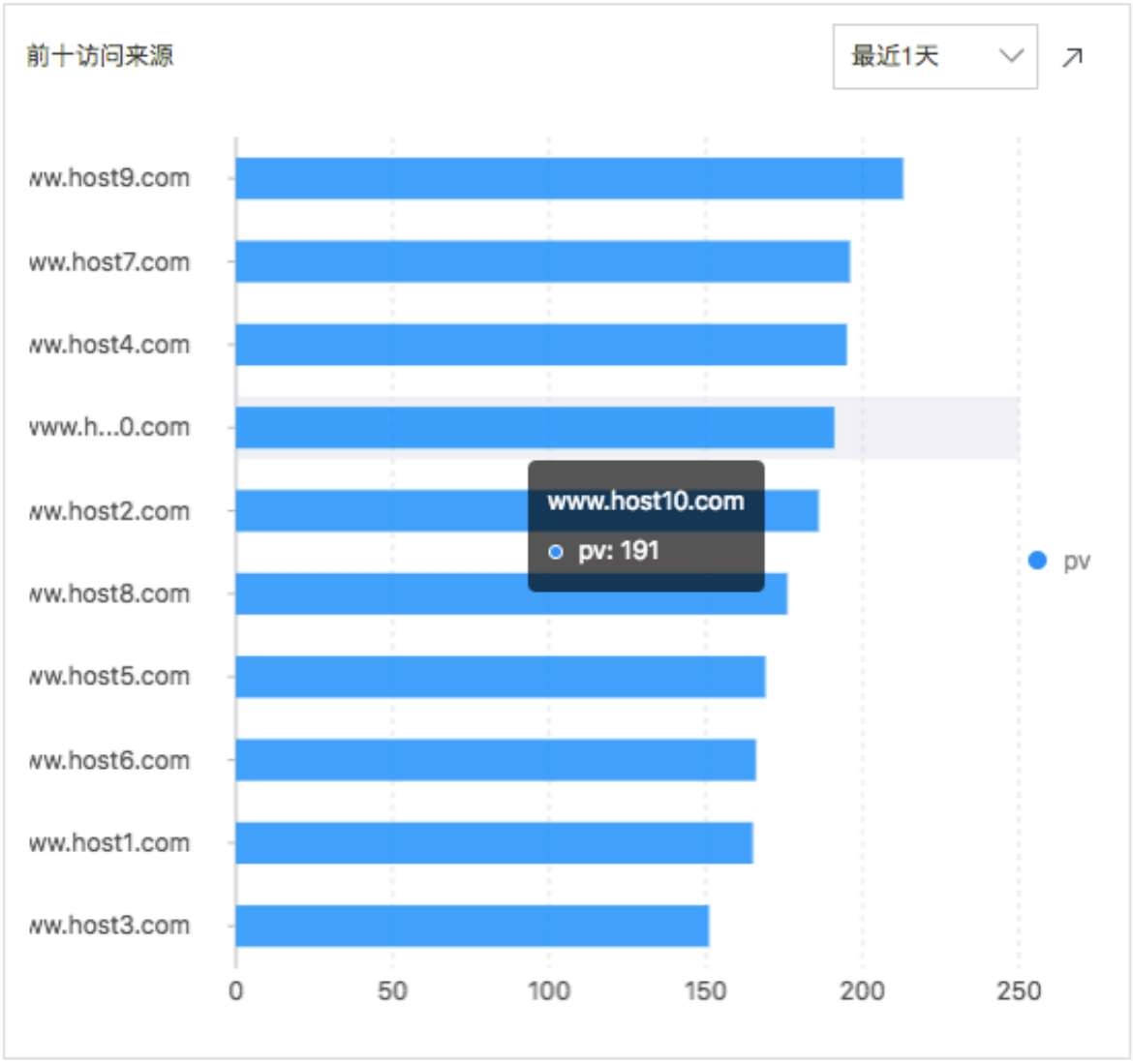
```
* | select count(1) as pv,
    case when http_user_agent like '%Chrome%' then 'Chrome'
    when http_user_agent like '%Firefox%' then 'Firefox'
    when http_user_agent like '%Safari%' then 'Safari'
    else 'unKnown' end as http_user_agent
    group by  http_user_agent
    order by pv desc
```

```
    limit 10
```

- 前十访问来源(top_10_referer)

  统计最近一天访问前十的来源信息。

  图 5-13: 前十访问来源



  统计语句:

```
* | select count(1) as pv,
        http_referer
        group by http_referer
        order by pv desc limit 10
```

Nginx访问日志诊断及优化

除了一些默认的访问指标外，站长常常还需要对一些访问请求进行诊断，分析Nginx访问日志中记录的处理请求的延时如何、有哪些比较大的延时、哪些页面的延时比较大。此时可以进入查询页面进行快速分析。

- 统计平均延时和最大延时

  通过每5分钟的平均延时和最大延时，从整体上了解延时情况。

  统计语句：

  ```
  * | select from_unixtime(__time__ -__time__% 300) as time,
          avg(request_time) as avg_latency ,
          max(request_time) as max_latency
          group by __time__ -__time__% 300
  ```

- 统计最大延时对应的请求页面

  知道了最大延时之后，需要明确最大延时对应的请求页面，以方便进一步优化页面响应。

  统计语句：

  ```
  * | select from_unixtime(__time__ - __time__% 60) ,
          max_by(request_uri,request_time)
          group by __time__ - __time__%60
  ```

- 统计请求延时的分布

  统计网站的所有请求的延时的分布，把延时分布在十个桶里面，看每个延时区间的请求个数。

  统计语句：

  ```
  * |select numeric_histogram(10,request_time)
  ```

- 统计最大的十个延时

  除最大的延时之外，还需要统计最大的十个延时及其对应值。

  统计语句：

  ```
  * | select max(request_time,10)
  ```

- 对延时最大的页面调优

  假如/url2这个页面的访问延时最大，为了对/url2页面进行调优，接下来需要统计/url2这个页面的访问PV、UV、各种method次数、各种status次数、各种浏览器次数、平均延时和最大延时。

  统计语句：

  ```
  request_uri:"/url2" | select count(1) as pv,
          approx_distinct(remote_addr) as uv,
          histogram(method) as method_pv,
          histogram(status) as status_pv,
          histogram(user_agent) as user_agent_pv,
          avg(request_time) as avg_latency,
          max(request_time) as max_latency
  ```

得到以上数据后，就可以对网站的访问情况进行有针对性的详细评估。

6 分析Apache日志

日志服务支持通过数据接入向导一站式配置采集Apache日志与设置索引。您可以通过默认仪表盘与查询分析语句实时分析网站访问情况。

前提条件

- 已开启日志服务。
- 已创建了Project和Logstore。

背景信息

个人站长选用Apache作为服务器搭建网站，需要通过分析Apache访问日志来获取PV、UV、IP区域分布、错误请求、客户端类型和来源页面等，以评估网站访问情况。

日志服务支持通过数据接入向导一站式配置采集Apache日志与设置索引，并为Apache日志默认创建访问分析仪表盘。

为了贴合分析场景，推荐您对Apache日志采用以下自定义配置。

```
LogFormat "%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i
\" %D %f %k %p %q %R %T %I %O" customized
```

说明:

请根据您的日志内容判断是否有部分字段对应的日志内容中存在空格，例如`%t`、`%{User-Agent}i`、`%{Referer}i`等。若有存在空格的字段，请在配置信息中用`\"`包裹该字段，以免影响日志解析。

各个字段含义如下：

字段	字段名	说明
%h	remote_addr	客户端IP地址。
%l	remote_ident	客户端日志名称，来自identd。
%u	remote_user	客户端用户名。
%t	time_local	服务器时间。
%r	request	请求内容，包括方法名、地址和http协议。
%>s	status	返回的http状态码。
%b	response_size_bytes	返回的大小。

字段	字段名	说明
%{Rererer}i	http\u0008_referer	来源页。
%{User-Agent}i	http_user_agent	客户端信息。
%D	request_time_msec	请求时间，单位为毫秒。
%f	filename	带路径的请求文件名。
%k	keep_alive	keep-alive请求数。
%p	remote_port	服务器端口号。
%q	request_query	查询字符串，如果不存在则为空字符串。
%R	response_handler	服务器响应的处理程序。
%T	request_time_sec	请求时间，单位为秒。
%I	bytes_received	服务器接收的字节数，需要启用mod_logio模块。
%O	bytes_sent	服务器发送的字节数，需要启用mod_logio模块。

操作步骤

1. 登录日志服务控制台，单击Project名称。
2. 在对应日志库下，单击数据接入后的加号。
3. 选择数据类型APACHE-文本日志。
4. 选择日志库

   可以选择已有的Logstore，也可以新建Project和Logstore。
5. 创建机器组。

   在创建机器组之前，您需要首先确认已经安装了Logtail。

   - 集团内部机器：默认自动安装，如果没有安装，请根据界面提示进行咨询。
   - ECS机器： 勾选实例后单击安装进行一键式安装。Windows系统不支持一键式安装，请参见安装Logtail（Windows系统）手动安装。
   - 自建机器：请根据界面提示进行安装。或者参见安装Logtail（Linux系统）或安装Logtail（Windows系统）文档进行安装。

   安装完Logtail后单击确认安装完毕创建机器组。如果您之前已经创建好机器组，请直接单击使用现有机器组。
6. 应用机器组。

7. Logtail配置。

a) 填写配置名称。

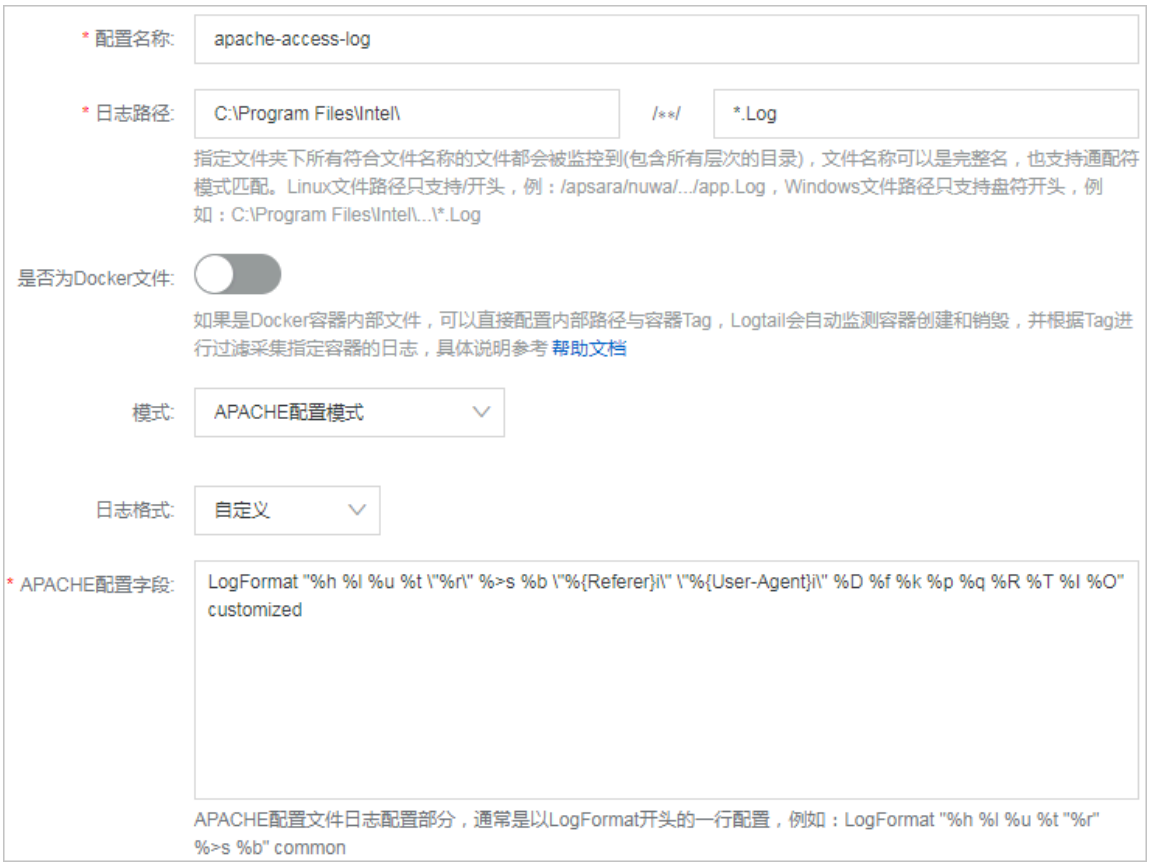
b) 填写日志路径。

c) 选择日志格式。

请按照您的Apache日志配置文件中声明的格式选择日志格式。为了便于日志数据的查询分析，日志服务推荐您使用自定义的Apache日志格式。

d) 填写APACHE配置字段。

若您的日志格式为common或combined，此处会自动填写对应的配置。若您选择了自定义日志格式，请在此处填写您的自定义配置。建议填写上文中日志服务的推荐配置。



e) 确认APACHE键名称。

日志服务会自动解析您的APACHE键名称，您可以在页面中确认。

说明:

%r会被提取为request_method、request_uri和request_protocol三个键。

f) （可选）配置高级选项，并单击下一步。

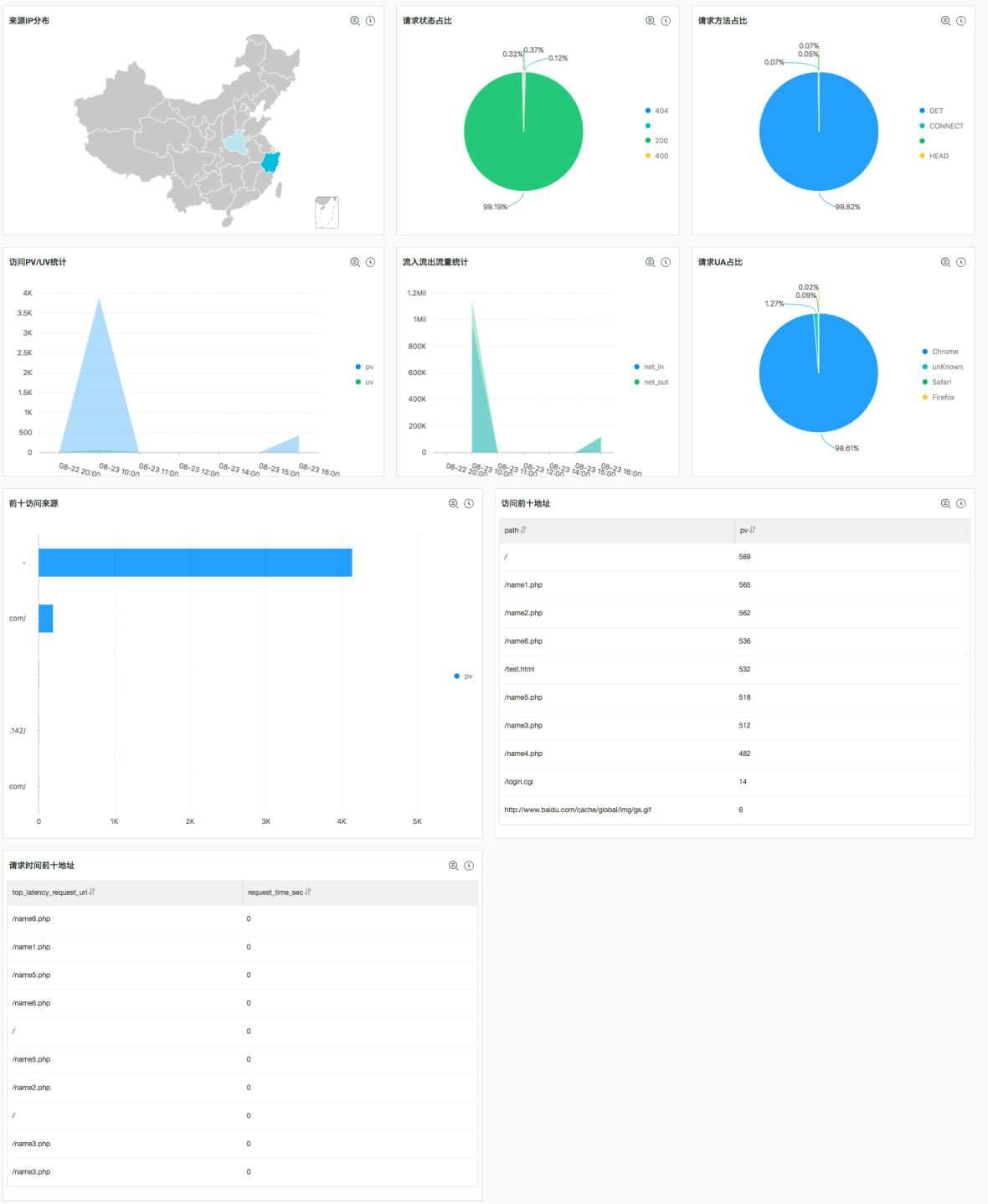
g) 查询分析配置

默认已经设置好索引，如果您需要重新设置索引，请在查询分析页面选择查询分析属性 > 设置进行修改。

确保日志机器组心跳正常的情况下，可以预览采集上来的数据。

8. （可选）可视化分析。

系统已为您预设了名为*LogstoreName-apache-dashboard*的仪表盘。配置完成后，您可以在仪表盘页面中查看来源IP分布、请求状态占比等实时动态。



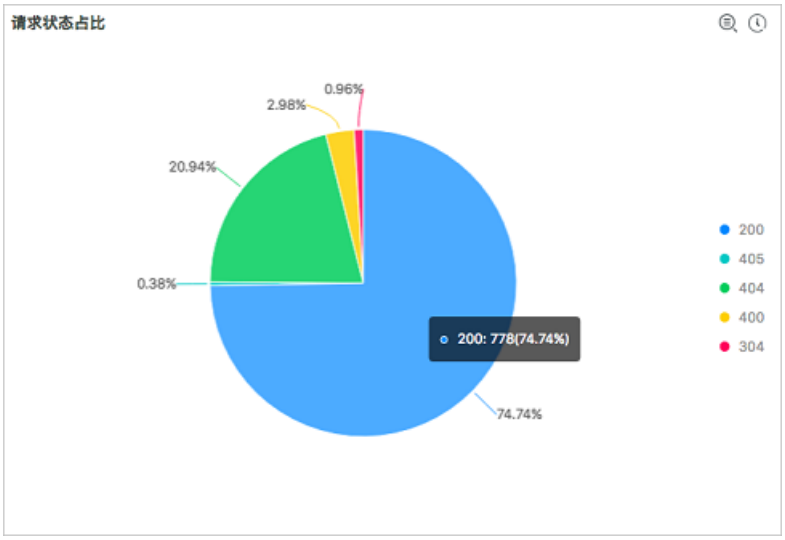
- 访问地域分析（ip_distribution）：统计访问IP来源情况，统计语句如下：

```
* | select ip_to_province(remote_addr) as address,
          count(1) as c
```

```
            group by address limit 100
```

- 请求状态占比（http_status_percentage）：统计最近一天各种http状态码的占比，统计语句如下：

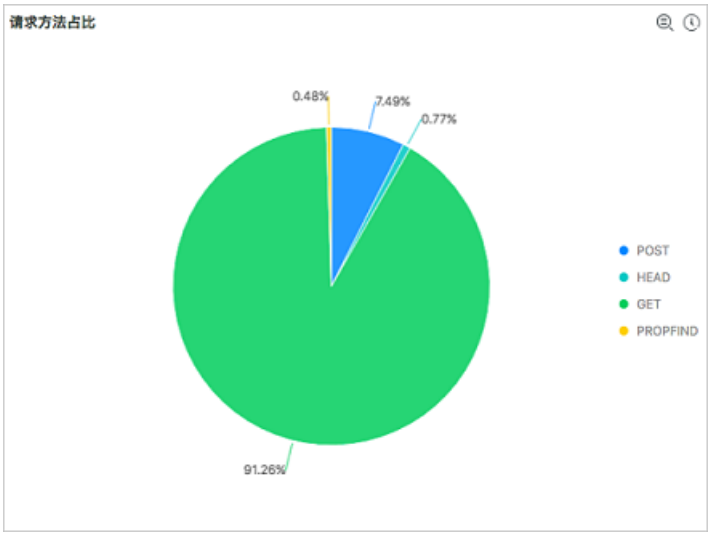
```
status>0 | select status,
                count(1) as pv
                group by status
```



- 请求方法占比（http_method_percentage）：统计最近一天各种请求方法的占比，统计语句如下：

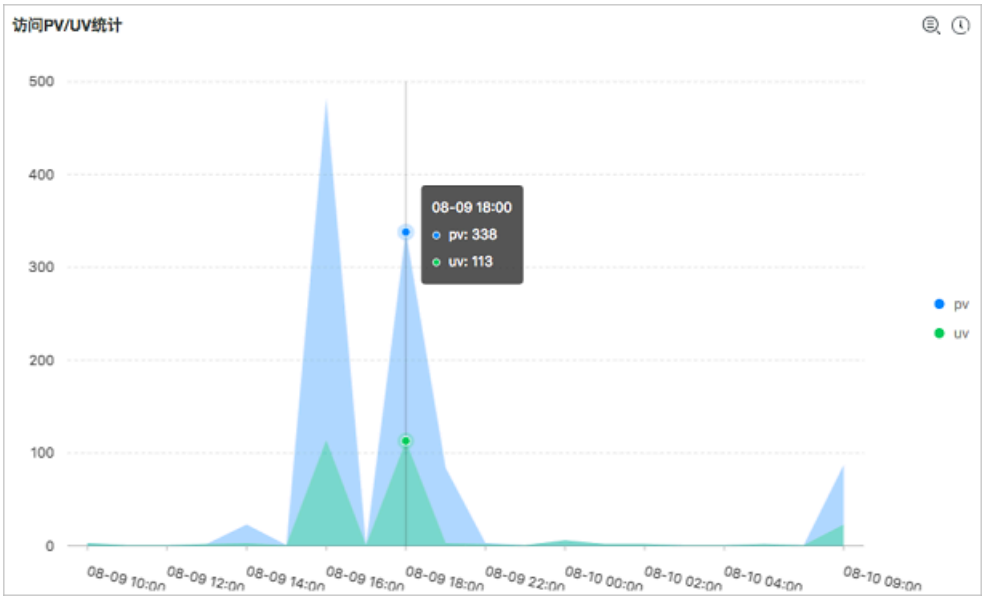
```
* | select request_method,
            count(1) as pv
```

```
            group by request_method
```



- **PV/UV统计（pv_uv）：统计最近的PV数和UV数，统计语句如下：**

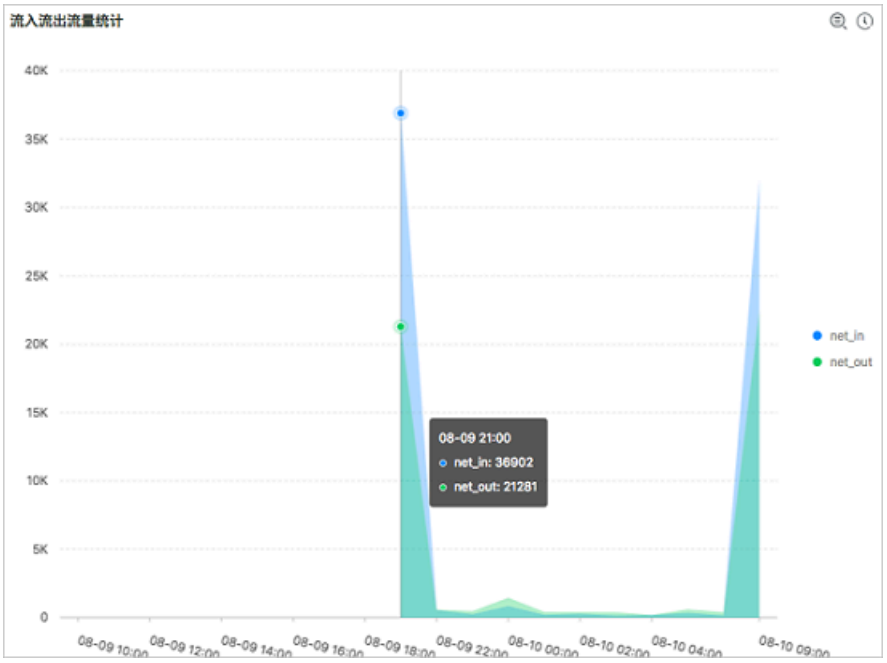
```
* | select date_format(date_trunc('hour', __time__), '%m-%d %H:%i
')  as time,
            count(1) as pv,
            approx_distinct(remote_addr) as uv
            group by time
            order by time
            limit 1000
```



- **出入流量统计（net_in_net_out）：统计流量的流入和流出情况，统计语句如下：**

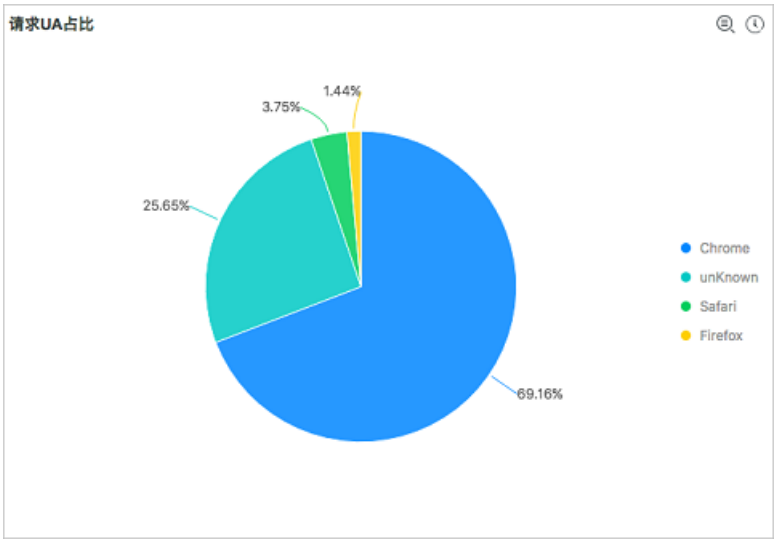
```
* | select date_format(date_trunc('hour', __time__), '%m-%d %H:%i
') as time,
            sum(bytes_sent) as net_out,
            sum(bytes_received) as net_in
            group by time
            order by time
```

```
            limit 10000mit 10
```



- **请求UA占比（http_user_agent_percentage）：统计最近一天各种浏览器的占比，统计语句如下：**

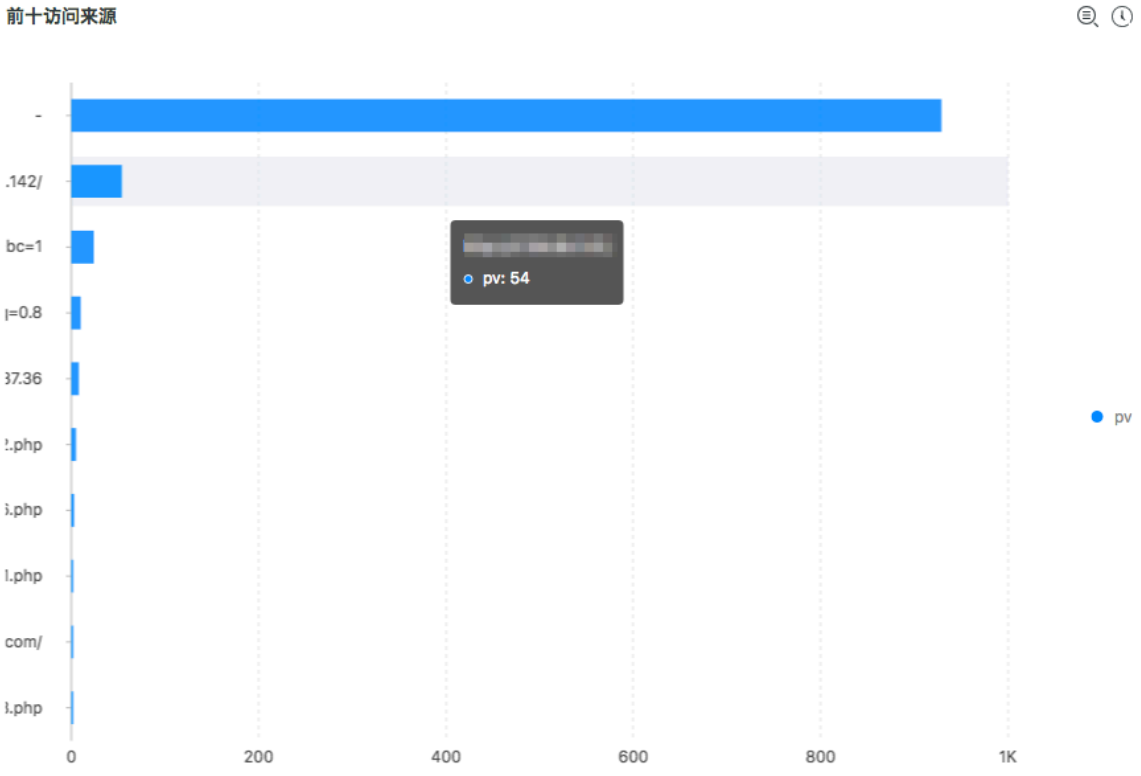
```
* | select  case when http_user_agent like '%Chrome%' then 'Chrome
'
            when http_user_agent like '%Firefox%' then 'Firefox'
            when http_user_agent like '%Safari%' then 'Safari'
            else 'unKnown' end as http_user_agent,count(1) as pv
            group by  http_user_agent
            order by pv desc
            limit 10
```



- **前十访问来源（top_10_referer）：统计最近一天访问PV前十的访问来源页面，统计语句如下：**

```
* | select  http_referer,
```

```
            count(1) as pv
            group by http_referer
            order by pv desc limit 10
```



- 访问前十地址（top_page）：统计最近一天访问pv前十的地址，统计语句如下：

```
* | select split_part(request_uri,'?',1) as path,
            count(1) as pv
            group by path
            order by pv desc limit 10
```

- 请求时间前十地址（top_10_latency_request_uri）：统计最近一天请求响应延时最长的前十个地址，统计语句如下：

```
* | select request_uri as top_latency_request_uri,
            request_time_sec
```

```
order by request_time_sec desc limit 10 10
```

请求时间前十地址

top_latency_request_uri	request_time_sec
/name6.php	0
/name1.php	0
/name5.php	0
/name6.php	0
/	0
/name5.php	0
/name2.php	0
/	0
/name3.php	0
/name3.php	0

7 分析IIS日志

个人站长选用IIS作为服务器搭建网站，需要通过分析IIS访问日志来获取PV、UV、IP区域分布、错误请求、流量流入流出，以评估网站访问情况。

前提条件

- 已开启日志服务。
- 已创建了Project和Logstore。详细步骤请参见准备流程。
- IIS日志采用W3C日志格式。

  为了更好满足分析场景，推荐选用W3C日志格式，在IIS管理器中单击选择字段按钮，勾选发送的字节数和和接收的字节数。

  图 7-1: 选择字段

背景信息

日志格式

W3C配置格式如下：

```
logExtFileFlags="Date, Time, ClientIP, UserName, SiteName, ComputerNa
me, ServerIP, Method, UriStem, UriQuery, HttpStatus, Win32Status
, BytesSent, BytesRecv, TimeTaken, ServerPort, UserAgent, Cookie,
Referer, ProtocolVersion, Host, HttpSubStatus"
```

- 字段前缀说明

前缀	说明
s-	服务器操作
c-	客户端操作
cs-	客户端到服务器的操作
sc-	服务器到客户端的操作

· 各个字段说明

字段	说明
date	日期，表示活动发生的日期。
time	时间，表示活动发生的时间。
s-sitename	服务名，表示客户端所访问的该站点的 Internet 服务和实例的号码。
s-computername	服务器名，表示生成日志项的服务器名称。
s-ip	服务器IP，表示生成日志项的服务器的IP地址。
cs-method	方法，例如GET或POST。
cs-uri-stem	URI资源，表示请求访问的地址。
cs-uri-query	URI查询，表示查询HTTP请求中问号（?）后的信息。
s-port	服务器端口，表示客户端连接的服务器端口号。
cs-username	通过验证的域或用户名，对于通过身份验证的用户，格式是域\用户名；对于匿名用户，是一个连字符 (-)。
c-ip	客户端IP，表示访问服务器的客户端真实IP 地址。
cs-version	协议版本，例如 HTTP 1.0 或 HTTP 1.1。
user-agent	用户代理，表示在客户端使用的浏览器。
Cookie	Cookie，表示发送或接受的Cookie内容，如果没有Cookie，则显示连字符（-）。
referer	引用站点，表示用户访问的前一个站点。此站点提供到当前站点到链接。
cs-host	主机，表示主机头内容。
sc-status	协议返回状态，表示HTTP或FTP的操作状态。
sc-substatus	HTTP子协议的状态。
sc-win32-status	win32状态，即用 Windows使用的术语表示的操作的状态。
sc-bytes	服务器发送字节。
cs-bytes	服务器接收字节。
time-taken	所用时间，即操作所花时间长短，单位为毫秒。

操作步骤

1. 进入数据接入向导。

   a) 在日志服务控制台首页单击Project名称。

   b) 在对应日志库下，单击数据接入后的加号。

2. 在接入数据中选择IIS-文本日志。
3. 选择日志库。

   可以选择已有的Logstore，也可以新建Project和Logstore。
4. 创建机器组。

   在创建机器组之前，您需要首先确认已经安装了Logtail。

   - 集团内部机器：默认自动安装，如果没有安装，请根据界面提示进行咨询。
   - ECS机器： 勾选实例后单击安装进行一键式安装。Windows系统不支持一键式安装，请参考安装Logtail（Windows系统）手动安装。
   - 自建机器：请根据界面提示进行安装。或者参考安装Logtail（Linux系统）或安装Logtail（Windows系统）文档进行安装。

   安装完Logtail后单击确认安装完毕创建机器组。如果您之前已经创建好机器组，请直接单击使用现有机器组。
5. 应用机器组。

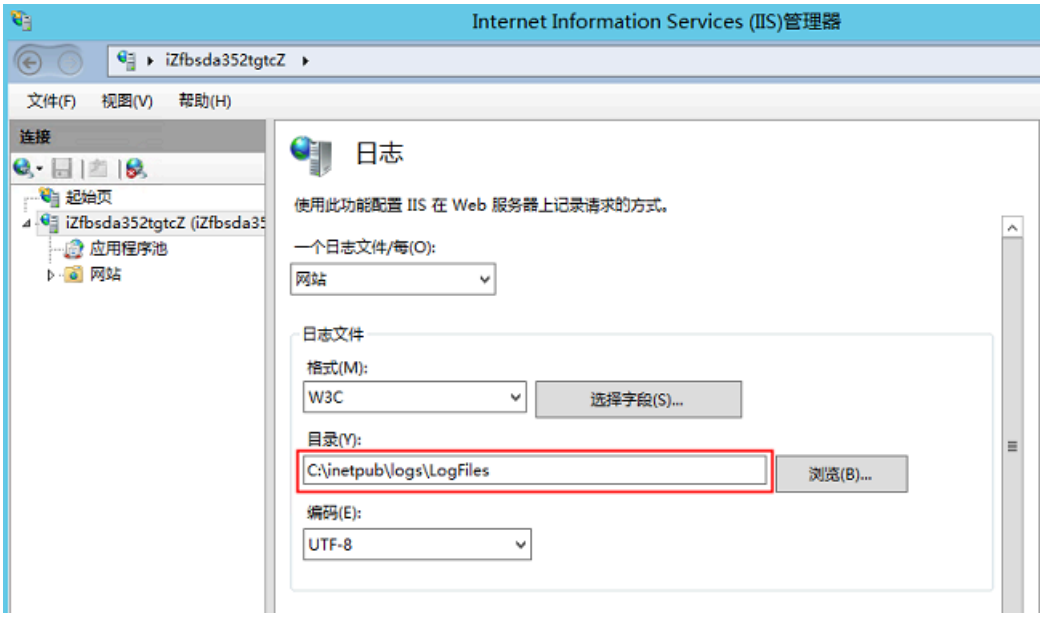
   选择一个机器组，将该机器组从源机器组移动到应用机器组。

6. Logtail配置。

a) 填写配置名称和日志路径。

您可以在IIS管理器中查看日志路径。

图 7-2: 查看日志路径



b) 选择日志格式。

选择您的IIS服务器日志采用的日志格式。

- `IIS`：Microsoft IIS日志文件格式。
- `NCSA`：NCSA公用日志文件格式。
- `W3C`：W3C扩展日志文件格式。

c) 填写IIS配置字段。

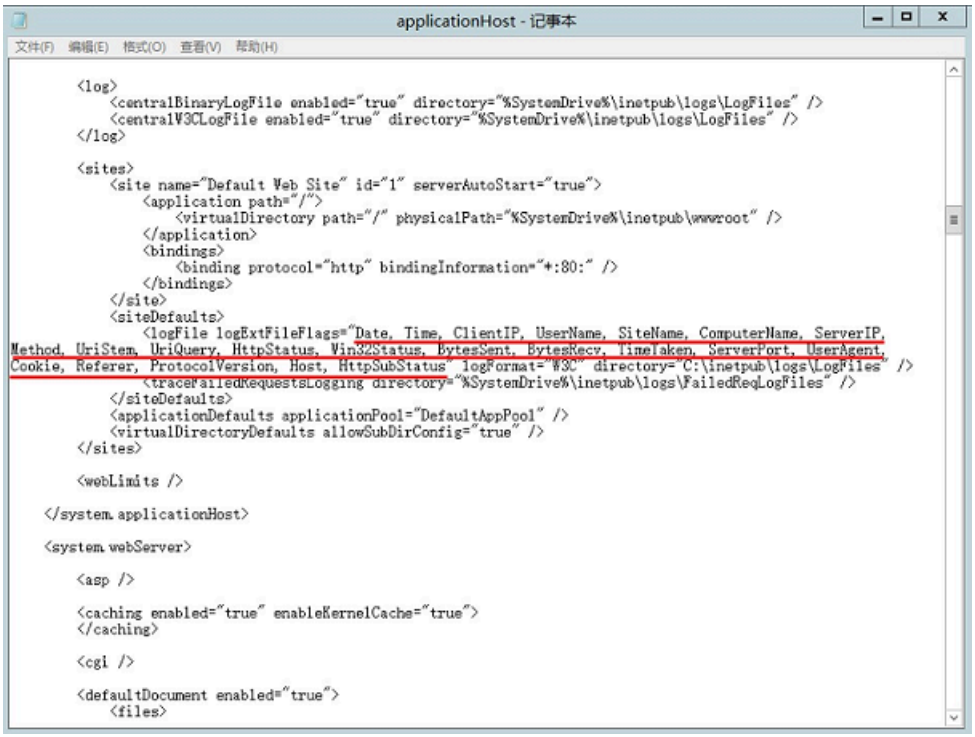
- IIS或NCSA格式：配置字段已预设。
- W3C日志：请按照以下步骤配置IIS配置字段。

A. 打开IIS配置文件。

- IIS5配置文件默认路径：*C:\WINNT\system32\inetsrv\MetaBase.bin*
- IIS6配置文件默认路径：*C:\WINDOWS\system32\inetsrv\MetaBase.xml*

- **IIS7配置文件默认路径：***C:\Windows\System32\inetsrv\config\applicationHost.config*

图 7-3: 查看配置文件



B. 找到`logFile logExtFileFlags`字段，并复制引号内的字段内容。

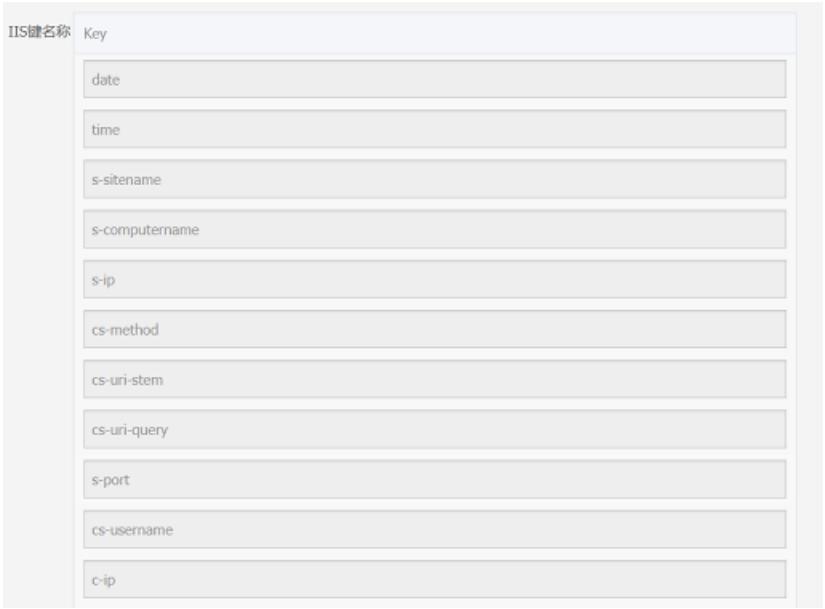
C. 粘贴字段内容到IIS配置字段输入框中的引号内。

图 7-4: 配置数据源

d) 确认IIS键名称。

IIS日志服务会自动提取出相应的键名称。

图 7-5: IIS键名称



e) 选择是否丢弃解析失败日志。

请选择解析失败的日志是否上传到日志服务。

开启后，解析失败的日志不上传到日志服务；关闭后，日志解析失败时上传原始日志，其中Key为`__raw_log__`、Value为日志内容。

f) 酌情配置高级选项（可选）。

确认配置后单击下一步。

7. 查询分析配置。

日志服务默认提供数据键名称以便分析使用，可以设置实际数据键名称（根据预览数据生成），和默认数据键名称形成映射关系。如果您需要重新设置索引，请在查询分析页面选择查询分析属性 > 设置进行修改。

确保日志机器组心跳正常的情况下，浏览采集上来的数据。

图 7-6: 预览日志

系统已为您预设了名为*LogstoreName-iis-dashboard*的仪表盘。配置完成后，您可以在仪表盘页面中查看来源IP分布、请求状态占比等实时动态。

图 7-7: 仪表盘

- 访问地域分析(ip_distribution)：统计IP来源情况，统计语句如下：

```
| select ip_to_geo("c-ip") as country, count(1) as c group by
ip_to_geo("c-ip") limit 100
```

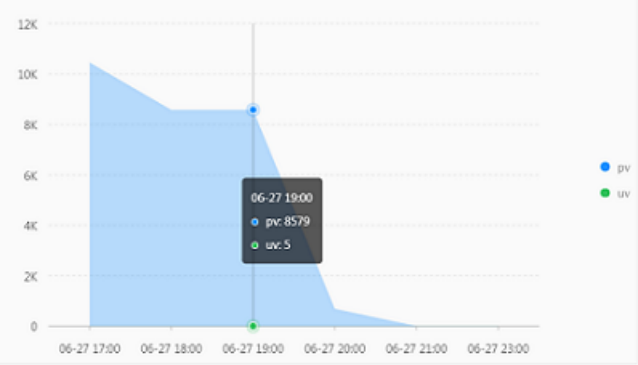
图 7-8: 访问地域分析

- PV/UV统计(pv_uv)：统计最近的PV数和UV数，统计语句如下：

```
*| select approx_distinct("c-ip") as uv ,count(1) as pv ,
date_format(date_trunc('hour', __time__), '%m-%d %H:%i') as time
```

```
group by date_format(date_trunc('hour', __time__), '%m-%d %H:%i')
order by time limit 1000
```

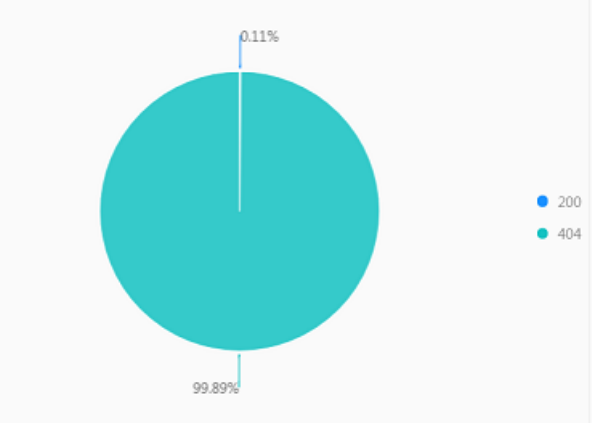
图 7-9: PV/UV统计



- 请求状态占比(http_status_percentage)：统计http请求状态码的占比，统计语句如下：

```
*| select count(1) as pv ,"sc-status" group by "sc-status"
```

图 7-10: 请求状态占比

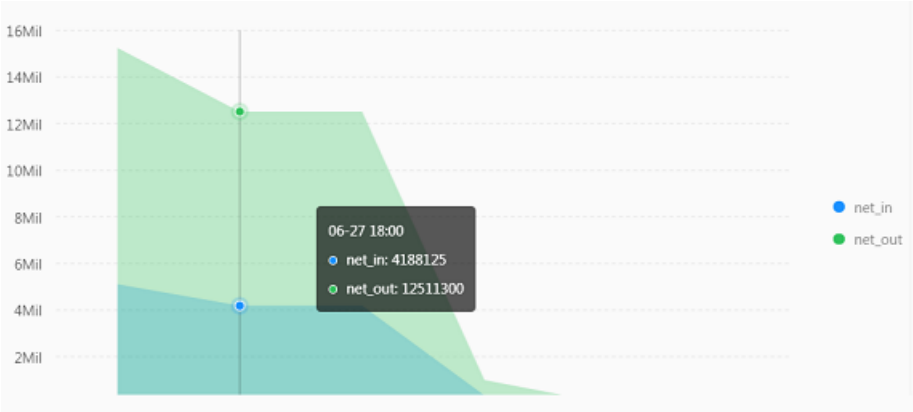


- 浏览流入流出统计（net_in_net_out）：统计流量的流入和流出情况，统计语句如下：

```
*| select sum("sc-bytes") as net_out, sum("cs-bytes") as net_in ,
date_format(date_trunc('hour', time), '%m-%d %H:%i') as time group
```

```
 by date_format(date_trunc('hour', time), '%m-%d %H:%i') order by
time limit 10000
```

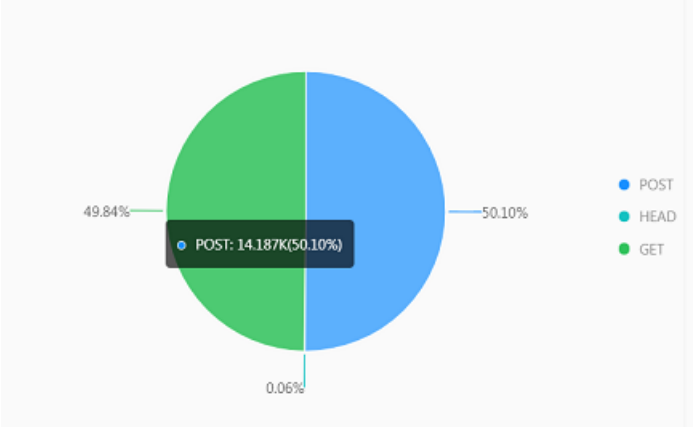
图 7-11: 出入流量统计



- 请求方法占比(http_method_percentage)：统计各种请求方法的占比，统计语句如下：

```
*| select count(1) as pv ,"cs-method" group by "cs-method"
```

图 7-12: 请求方法占比

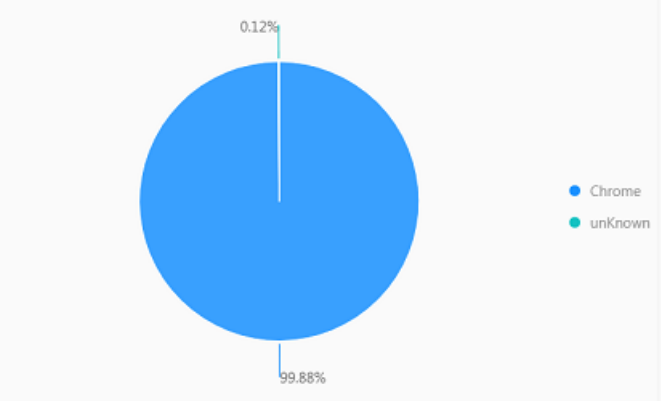


- 请求UA占比(user_agent)：统计各种浏览器的占比，统计语句如下：

```
*| select count(1) as pv, case when "user-agent" like '%Chrome%'
then 'Chrome' when "user-agent" like '%Firefox%' then 'Firefox'
when "user-agent" like '%Safari%' then 'Safari' else 'unKnown' end
 as "user-agent" group by case when "user-agent" like '%Chrome%'
then 'Chrome' when "user-agent" like '%Firefox%' then 'Firefox'
```

```
when "user-agent" like '%Safari%' then 'Safari' else 'unKnown' end
 order by pv desc limit 10
```

图 7-13: 请求UA占比



- 访问前十地址（top_10_page）：统计访问数量前十的地址，统计语句如下：

```
*| select count(1) as pv, split_part("cs-uri-stem",'?',1) as path
group by split_part("cs-uri-stem",'?',1) order by pv desc limit 10
```

图 7-14: 访问前十地址

pv	path
201	/bucketname3
186	/bucketname15
182	/bucketname11
168	/bucketname0