Elasticsearch Service

Logstash 指南

腾讯云

文档目录

Logstash 指南
- 腾讯云 Logstash 概述
- 快速入门
  - 创建实例
  - 创建管道
  - 查看数据同步结果
- 实例管理
  - 实例列表
  - 重启实例
  - 销毁实例
- 实例扩缩容
- 管道管理
- YML 文件配置
- 监控与告警
  - 查看监控
  - 配置 X-Pack 监控
  - 配置告警
- 查询日志
- 实践教程
  - 接收 Filebeat 发送的数据并写入到 Elasticsearch
  - 同步 MySQL 中的数据到 Elasticsearch
  - 同步两个 Elasticsearch 集群中的数据
  - 消费 kafka 数据并写入到 Elasticsearch
  - 读取 COS 中的日志文件并写入到 Elasticsearch

Logstash 指南
腾讯云 Logstash 概述

最近更新时间：2024-10-16 10:03:41

腾讯云 Logstash（简称 Logstash）是基于 开源数据收集引擎 Logstash 构建的云端托管服务，它是一个服务器端的数据处理管道，支持动态的从不同来源采集和转换数据，并将数据标准化到目标位置。Logstash 常和 Elasticsearch 配合，通过输入、过滤和输出插件，加工和转换任何类型的事件，将数据加载到 Elasticsearch。

Logstash 的工作方式

- 数据输入：支持多样的数据来源，通过输入插件方便的采集日志、指标、Web 应用、数据库、消息队列、传感器等来源数据。
- 数据过滤：通过过滤插件清理和转换数据，如将非结构化数据解析导出结构、解析 IP 地址、标准化日期、通过编解码器简化常见格式等。
- 数据输出：通过输出插件将数据传输到需要的地方，例如 Elasticsearch、数据库等，以便对数据做进一步的分析和处理。

特点与优势

- 易于部署和管理，简化运维操作。
- 支持弹性扩展节点数量。
- 集成官方所有 Input、Output、Filter 插件。
- 支持 CKafka、MySQL、PostgreSQL、COS 等腾讯云产品的输入或输出插件。

快速入门
创建实例

最近更新时间：2024-10-16 10:03:41

本文为您介绍通过腾讯云官网控制台快速创建 Logstash 实例。

前提条件

已创建腾讯云账号，创建账号可参考 注册腾讯云 。

操作步骤

登录控制台

登录 Elasticsearch Service 控制台，在左侧导航栏单击 **Logstash 管理**，在 Logstash 实例列表页单击**新建**进入购买页。

创建 Logstash 实例

一、选择实例配置

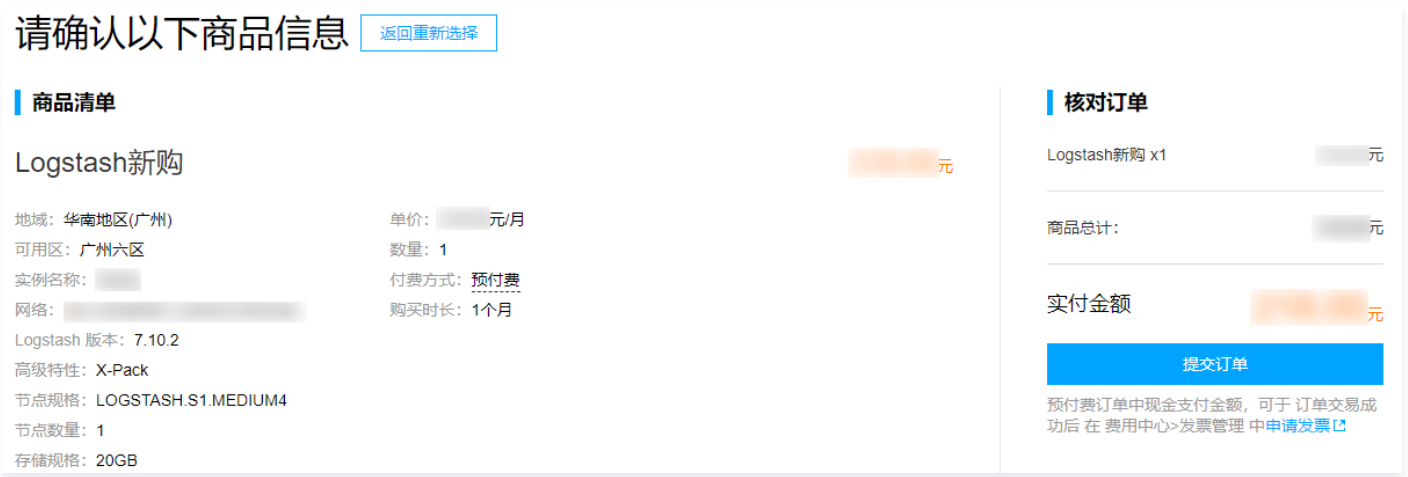
- 计费模式：支持包年包月和按量计费。
- 地域：支持国内和境外多个地域，其中上海金融、深圳金融、北京金融需要联系 售前咨询 申请。
- 可用区：指腾讯云在同一地域内电力和网络互相独立的物理数据中心，请根据实际需求进行选择。
- 网络及子网：若需要将 Logstash 实例与 Elasticsearch 集群搭配使用，请选择 ES 集群所在的 VPC，或者使用 云联网 打通不同 VPC 的 ES 集群。

> **说明**
> Logstash 实例创建完成后，不支持修改调整 VPC。

- Logstash版本：支持7.14.2、7.10.2、6.8.13版本。
- 高级特性：支持开源版和 X-Pack 版，其中开源版搭配 ES 开源版，X-Pack 版搭配 ES 基础版和白金版使用。
- 实例名称：自定义实例名称，不作为全局唯一标示，可以设置为业务相关描述。
- 节点数量：希望购买的节点数量。
- 节点规格：每个节点的机型和规格，不同规格包含不同的 CPU 核数和内存。
- 节点存储：每个节点配置的磁盘类型和容量，整个实例的存储量 = 单个节点存储 × 节点个数。

二、包年包月确认订单及支付

若选择包年包月，会出现订单支付的确认页面。



三、创建完成

支付完成，实例创建成功后，即可跳转到 Elasticsearch Service 控制台查看刚才创建的实例。

创建管道

最近更新时间：2024-10-16 10:03:41

创建 Logstash 实例后，可以创建管道进行数据同步。本文为您介绍如何创建一个 Logstash 管道，并将此管道用于将一个 Elasticsearch 集群中的索引同步到另一个 Elasticsearch 集群中。

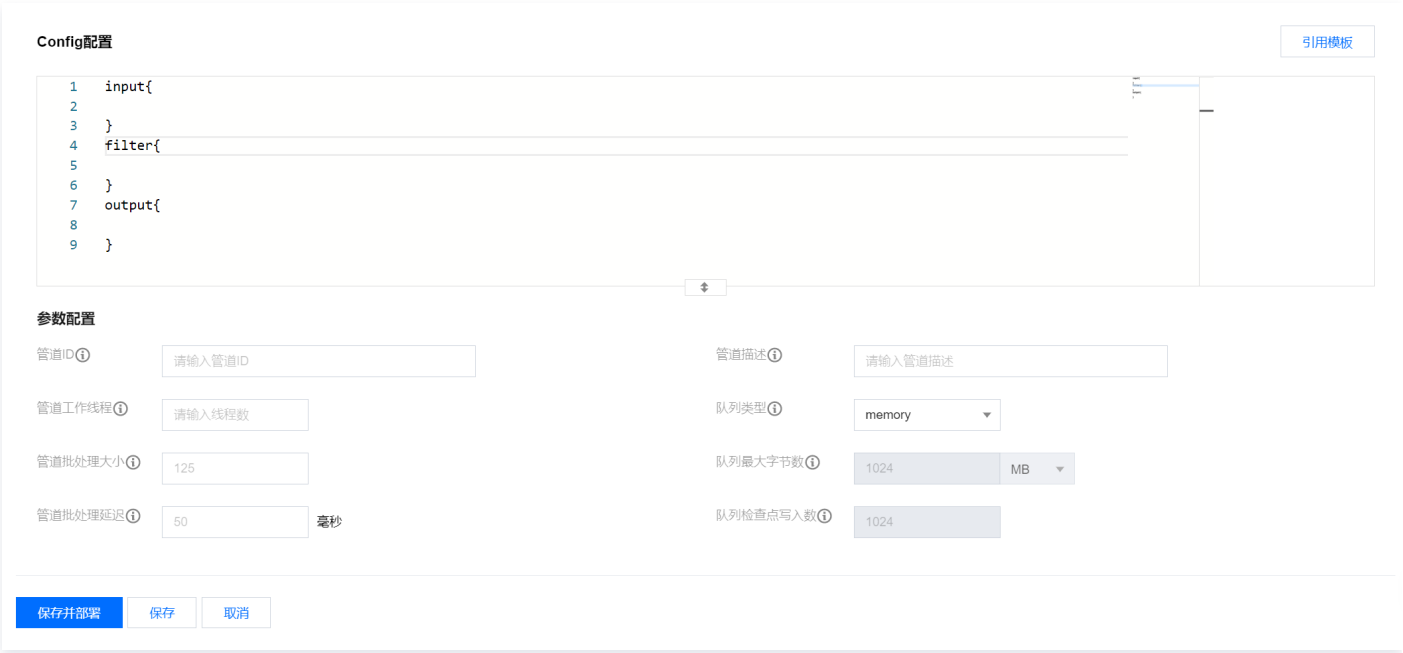
操作步骤

登录控制台

登录 Elasticsearch Service 控制台，在左侧导航栏单击 **Logstash 实例**，进入 Logstash 实例列表页。

创建管道

1. 在实例列表页，单击实例 **ID/名称**进入实例基本信息页，然后进入**管道管理**页签，单击**新建管道**，进入新建管道页面。



2. 在新建管道页面，输入 **Config 配置**，示例如下：

```
input {
 elasticsearch {
   hosts => ["http://x.x.x.x:9200"]
   user => "elastic"
   password => "password"
   index => "test1"
   docinfo => true
 }
}

output {
  elasticsearch {
    hosts => ["http://x.x.x.x:9200"]
    user => "elastic"
    password => "password"
    index => "index_a"
    document_id => "doc_id_1"
  }
}
```

参数说明：

- hosts：elasticsearch 集群地址列表，input 中的 hosts 为源 elasticsearch 集群地址，output 中的 host 为目标 elasticsearch 集群的地址。
- user：elasticsearch 集群账号。
- password：elasticsearch 集群密码。
- index：索引名称。
- docinfo：是否在 event 中填充索引名称，type 以及 id 等文档元信息，默认为 false。
- document_type：索引 type，若目标 elasticsearch 集群为7.x及以上的版本，不必设置该字段。
- document_id：文档 ID。

3. 在新建管道页面，输入**参数配置**（参数说明请参见 管道管理），示例如下：

4. 配置完成后，单击**保存**或者**保存并部署**。
5. 单击**保存**：保存管道信息到 Logstash 并触发实例变更，配置不会生效。保存后返回**管道管理**页。可在管道列表中选择**操作** > **部署**，触发实例重启生效。
6. 单击**保存并部署**：保存并且部署后，触发实例重启生效。

查看数据同步结果

最近更新时间：2024-10-16 10:03:41

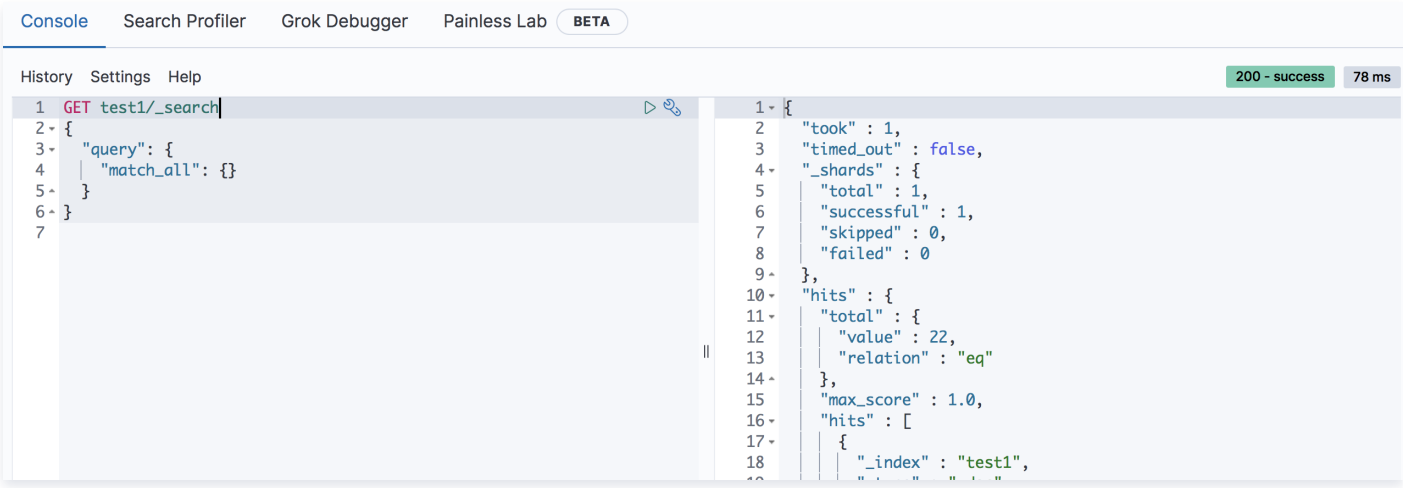
管道任务配置完成并运行后，可以在目标 ES 集群的 Kibana 控制台上查看数据同步结果。

操作步骤

- 访问目标 ES 集群的 Kibana 控制台，在左侧导航栏，单击 **Dev Tools** 开发工具。
- 在 **Console** 中执行如下命令，查看数据是否已经成功写入。

```
GET /test1/_search
{
      "query": {
          "match_all": {}
      }
}
```

上述命令如果执行成功，则会返回如下结果：



实例管理
实例列表

最近更新时间：2024-10-16 10:03:41

登录 Elasticsearch Service 控制台，单击左侧导航栏 **Logstash 管理**进入实例列表页，实例列表展示了账号当前区域下所有 Logstash 实例的基本信息，并提供了操作入口方便对实例进行管理，详情如下：

功能	说明
实例列表信息	包括 ID/名称、运行状态（与 CVM 状态相同）、节点数量、实例配置、可用区、网络、版本、付费类型等。
ID/名称	可单击进入实例基本信息页，实例名称可就地编辑。
网络	可单击进入所属私有网络的基本信息页。
管道管理	跳转此实例的管道管理页。
重启	重启实例。
调整配置	跳转到调整配置页，对实例的基本配置进行调整。
按量转包年包月	仅适用于按量付费类型的实例，可跳转到相应页面变更付费类型。
续费	仅适用于包年包月类型的实例，可跳转到相应页面进行续费操作。
销毁	执行销毁实例操作。

重启实例

最近更新时间：2024-10-16 10:03:41

用户可以根据业务需要对实例进行重启操作。建议在实例的状态为正常时进行重启操作。

操作步骤

1. 登录 Elasticsearch Service 控制台，在左侧导航栏单击 **Logstash 管理**，进入 Logstash 实例列表页。

2. 在实例列表页，选择需要重启的实例，选择**操作 > 更多 > 重启**进行重启；或单击实例 **ID/名称**进入实例基本信息页，选择右上角**更多操作 > 重启**进行重启。

3. 单击**重启**后，在弹出的“确认重启实例？”页面中，选择重启方式，选择完成后，单击**确定**即可重启实例。正常实例重启时，运行状态变为处理中，等待状态恢复为正常，重启操作即完成。

销毁实例

最近更新时间：2024-10-16 10:03:41

操作场景

当 Logstash 实例无法满足您的需求，需要退货时，您可以在 Elasticsearch Service 控制台对实例进行销毁，以避免服务继续运行而产生费用。如果是实例配置无法满足需求，您也可以通过调整实例配置把实例调整到合适的规格，详情可参见 实例扩缩容 。

不同计费模式退费说明

不同计费模式下的实例，销毁实例的条件如下：

- 预付费包年包月的实例，如果实例还未到期，需要提前销毁时，可参见 包年包月退费 。
- 后付费按量计费的实例，根据使用量计费，可以随时销毁实例，销毁后，就不再产生费用。

> **注意**
> 实例被销毁后，数据无法恢复，请谨慎操作。

操作步骤

1. 登录 Elasticsearch Service 控制台 ，在左侧导航栏单击 **Logstash 管理**，进入 Logstash 实例列表页。

2. 在实例列表页，选择需要销毁的实例，选择**操作** > **更多** > **销毁**进行销毁；或单击实例 **ID/名称**进入实例基本信息页，选择右上角**更多操作** > **销毁**进行销毁。

3. 在销毁实例页面中，单击**确定**，系统将清空实例数据，并回收资源，**数据清空后，无法恢复**。包年包月的费用退还方式，可参见 包年包月退费 。

实例扩缩容

最近更新时间：2024-10-16 10:03:41

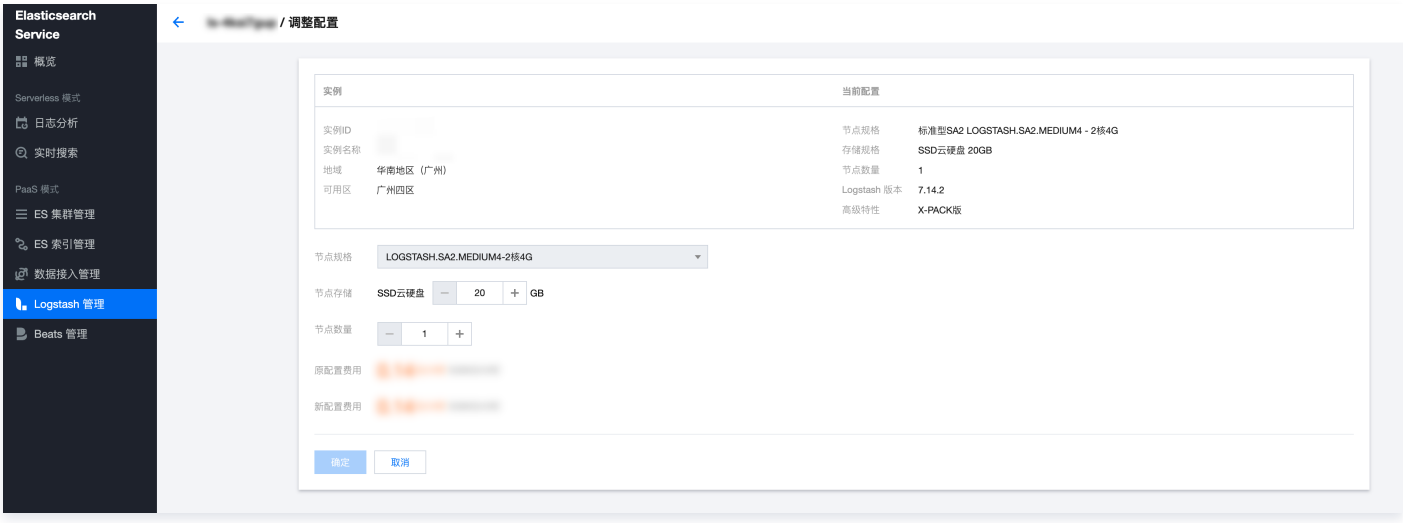
随着同步数据的流量变化，当实例规模跟实际业务需求不太匹配时，可以动态调整实例的配置，目前仅支持对节点个数进行扩容。

操作步骤

1. 登录 Elasticsearch Service 控制台，在左侧导航栏单击 **Logstash 管理**，进入 Logstash 管理列表页。
2. 在列表页选择需要调整配置的实例，然后选择**操作 > 更多 > 调整配置**，进入调整配置页面。

或者直接单击实例 ID/名称进入实例基本信息页，然后选择右上角**更多操作 > 调整配置**，进入调整配置页面。

3. 在**调整配置**页面，根据业务需求对实例节点数量进行调整，然后单击**确定**。配置调整开始后，实例状态为**处理中**，待实例状态变为**正常**，即可正常使用。



4. 可在实例详情页的**变更记录**页签，查看实例的变更进度。

调整配置费用说明

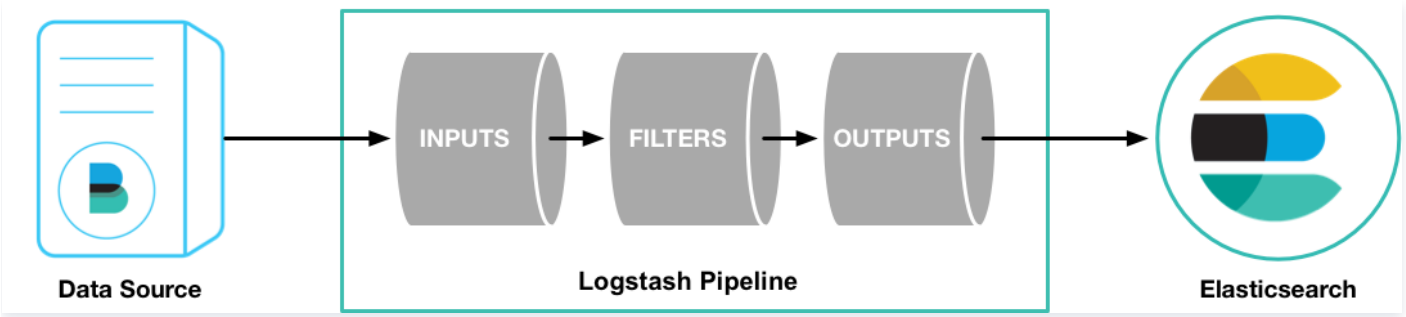
不同计费模式下的实例，调整配置时，费用结算方式会有所不同。

- 预付费包年包月的实例，在调整配置时，会根据实例剩余有效期以及新的配置的定价，计算需要的费用，具体可以参见 调整配置费用说明 。
- 后付费按量计费的实例，计费周期为秒，当配置调整完成后，下一秒费用按新的配置定价进行结算。

管道管理

最近更新时间：2024-10-16 10:03:41

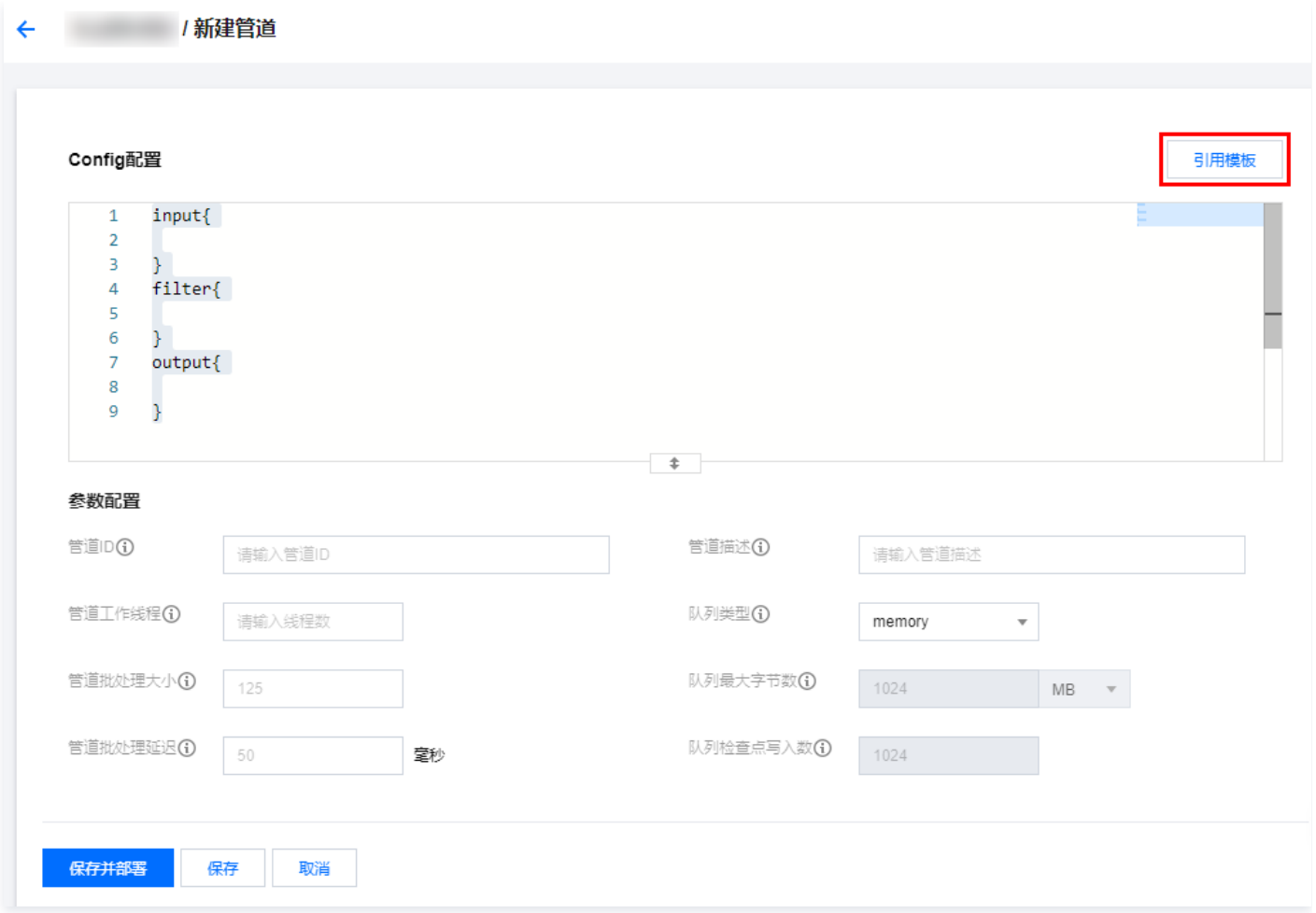
Logstash 通过管道来实现数据的采集处理，它包含必选的 input 和 output 插件，以及可选的 filter 插件，并支持多管道并行运行，目前支持并行的上限为10个。本文介绍如何通过配置文件管理管道，包括创建管道、修改管道、复制管道和删除管道。



创建管道

1. 登录 Elasticsearch Service 控制台，在左侧导航栏单击 **Logstash 实例**，进入 Logstash 实例列表页。
2. 在实例列表页，单击**实例 ID/名称**进入实例基本信息页，然后进入**管道管理**页签，单击**新建管道**。

3. 进入新建管道页面，单击**引用模板**进入选择模板页面。



4. 在选择模板页面，勾选默认提供的 input 和 output 模板，然后单击**引用**将模板引入到 Config 配置中。

5. **Config 配置**

   用户需根据实际需要修改 Config 配置。

```
input {
  ...
```

```
}
filter {
    ...
}
output {
    ...
}
```

- **参数说明**

参数	说明
input	输入数据源配置。Logstash 支持的输入数据源类型，可参见 Input plugins
filter	对数据进行过滤或者预处理的配置。Logstash 支持的 filter 插件类型，可参见 Filter plugins
output	输出数据源配置。Logstash 支持的输出数据源类型，可参见 Output plugins

管道配置详情可参考 配置文件结构 。

- **修改参数配置**

参数	说明	默认值
管道 ID	pipeline.id，管道的唯一标识	-
管道工作线程	pipeline.workers，管道的工作线程数量，也是并行执行管道的 filter 和 output 的工作线程数量	实例单节点的 CPU 核数
管道批处理大小	pipeline.batch.size，每个批次处理的最大事件数量	125
管道批处理延迟	pipeline.batch.delay，当管道批处理大小不满足时，每个批次最大的等待时间，单位为毫秒	50ms
队列类型	queue.type，用于事件缓冲的排队模型，可选值为 memory（基于内存的内列）或者 persisted（基于磁盘的持久化队列）	memory
队列最大字节数	queue.max_bytes，当选择 persisted 队列类型时，队列中可存放的最大字节数量，需确保该值小于实例单节点的磁盘容量	1024MB
队列检查点写入数	queue.checkpoint.acks，当选择 persisted 队列类型时，在强制执行检查点时已写入的最大的事件数量，若设置为0，则表示无限制	1024

6. 单击**保存**或者**保存并部署**，即可新建管道。**新建的管道需要保存并部署才能生效。**
   - 单击**保存**：保存管道信息到 Logstash 并触发实例变更，配置不会生效。保存后返回**管道管理**页，可在管道列表中选择**操作 > 部署**，触发实例自动加载管道配置并生效。
   - 单击**保存并部署**：保存并且部署后，触发实例自动加载管道配置并生效。

修改管道

修改管道后，需要**保存并部署**才能生效，此操作会触发实例自动加载管道配置并生效。

1. 在管道列表中，单击要修改的管道 ID，可以进入管道修改页。
2. 在管道修改页，单击**修改**，修改管道的Config 配置和参数配置。
3. 单击**保存并部署**，待实例自动加载管道配置后完成管道修改。

复制管道

复制管道后，需要**保存并部署**才能生效，此操作会触发实例自动加载管道配置并生效。

1. 在管道列表中，找到需要复制的管道，在**操作**列中单击**复制**。
2. 在复制管道页，修改管道 ID。
3. 单击**保存并部署**，待实例自动加载管道配置后完成管道修改。

删除管道

> **注意**
> - 管道删除后无法恢复，正在运行的管道会被中断，请确认后操作。
> - 管道删除会触发实例变更，请在不影响业务的情况下操作。

1. 在管道列表中，找到需要删除的管道，在**操作**列中单击**删除**。
2. 在删除管道对话框中，单击**确定**删除管道。

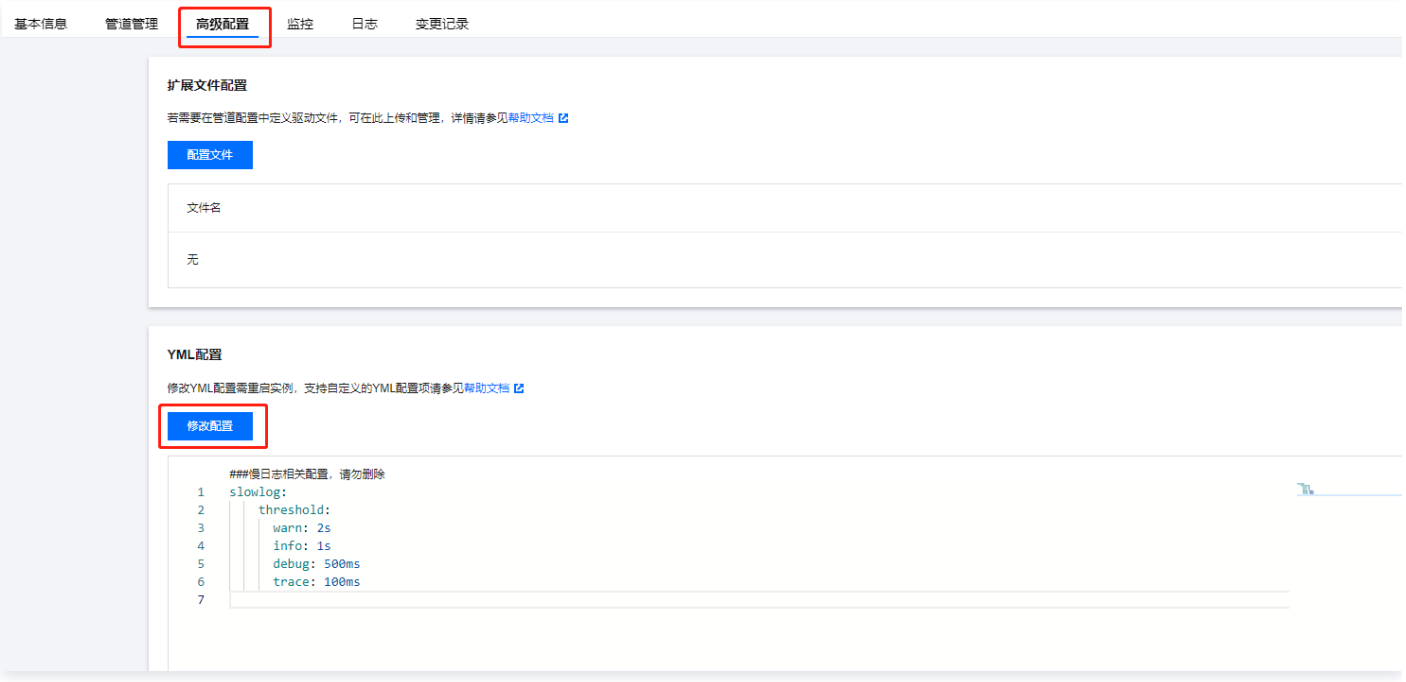
YML 文件配置

最近更新时间：2024-10-16 10:03:41

本文为您介绍如何通过 Elasticsearch Service 控制台配置腾讯云 Logstash 实例的 YML 参数。

操作步骤

1. 登录 Elasticsearch Service 控制台，在左侧导航栏单击 **Logstash 管理**，进入 Logstash 管理列表页。
2. 选择要修改 YML 参数配置的实例，单击 **ID/名称**，进入实例基本信息页。
3. 在实例基本信息页面，切换到**高级配置**页签，单击**修改配置**，根据业务需求修改 YML 参数。详细参数说明，可参见 Logstash 配置文件 。



4. YML 参数配置完成后，单击**保存**。将提示您是否重启实例，因修改 YML 参数需要重启 Logstash 实例才能生效，所以确认后 Logstash 实例将会重启，重启进度可查看**变更记录**。

监控与告警
查看监控

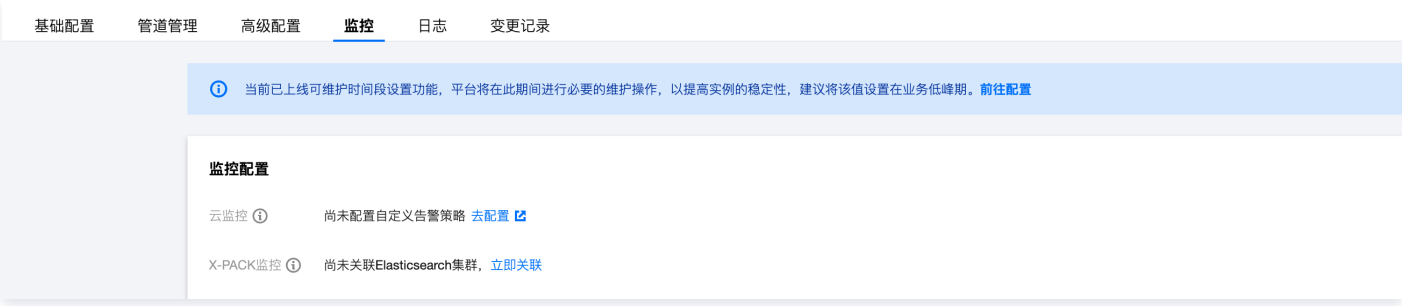
最近更新时间：2024-10-16 10:03:41

操作场景

腾讯云对运行中的 Logstash 实例，提供了多项监控指标，用以监测实例的运行情况，如 CPU、JVM、磁盘使用率等。您可以根据这些指标实时了解实例的运行状况，针对可能存在的风险及时处理，保障实例的稳定运行。本文为您介绍通过控制台查看实例监控的操作。

操作步骤

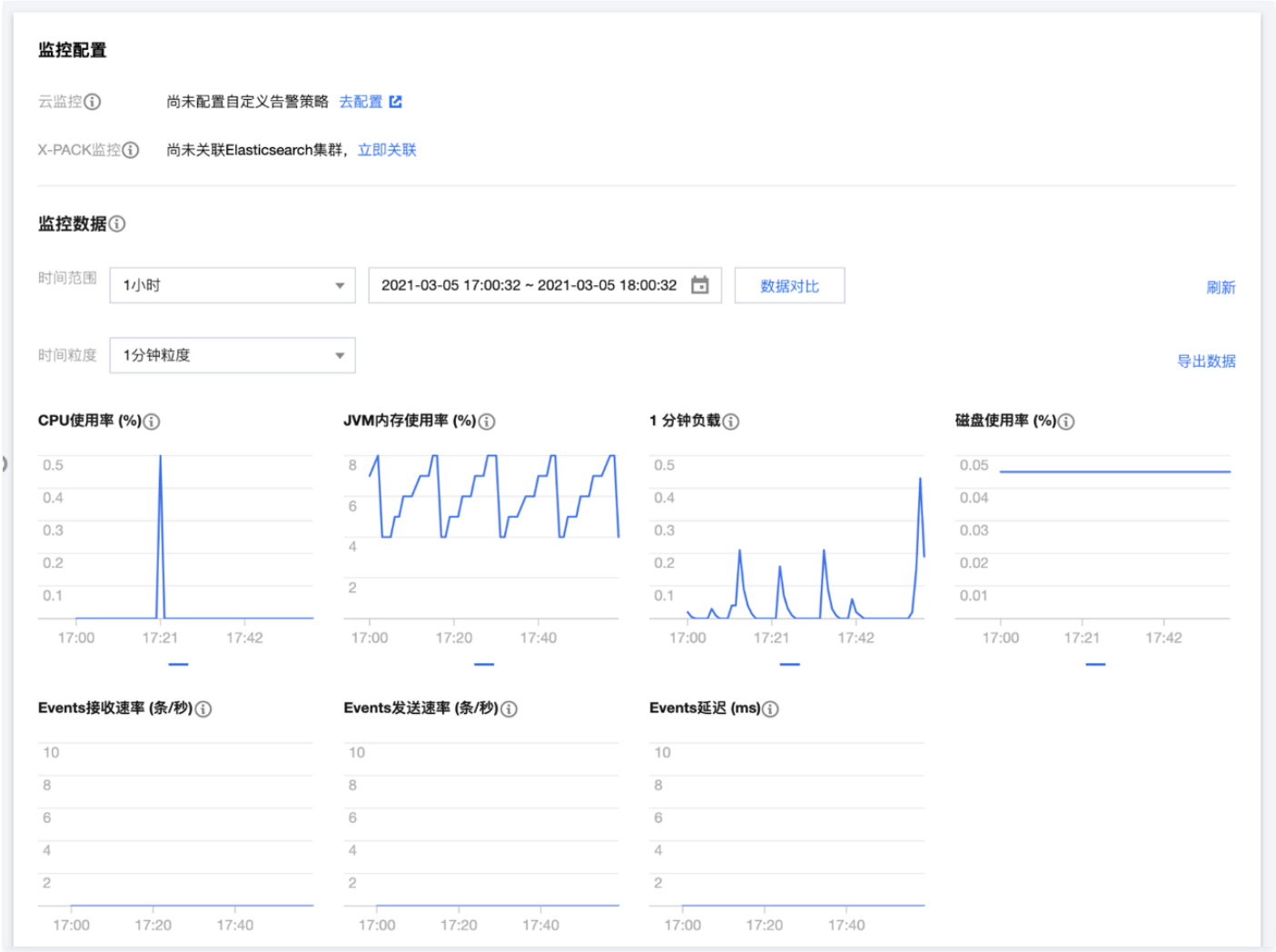
1. 登录 Elasticsearch Service 控制台，在左侧导航栏单击 **Logstash 管理**，进入 Logstash 管理列表页。在实例列表中，选择需要查看监控的实例，单击实例 **ID/名称**，进入实例基本信息页。
2. 在实例基本信息页面，切换到**监控**页签，即可查看实例的运行情况。



监控状态

页面展示了实例详细的指标和随时间变化的指标，可了解实例过去一段时间内的运行情况。

> **说明**
> Logstash 实例完整的监控指标也可通过 腾讯云可观测平台控制台 查看。



指标含义及说明

Logstash 实例一般由多个节点构成，所有指标的统计周期均为1分钟，即每1分钟对实例的指标采集1次。具体各指标含义说明如下：

监控指标	统计方式	详情
CPU 使用率	每单位统计周期内（1分钟），实例各个节点的 CPU 使用率的平均值	当实例各节点处理的读写任务超出节点 CPU 的负载能力时，该指标就会过高，CPU 使用率过高会导致实例节点处理能力下降，甚至宕机。您可观察该指标是持续性较高，还是临时飙升。若是临时飙升，确定是否有临时性复杂任务正在执行。
JVM 内存使用率	每单位统计周期内（1分钟），实例各个节点的 JVM 内存使用率的平均值	该值过高会导致实例节点 GC 频繁，甚至有出现 OOM。导致该值过高的原因，一般是节点上管道处理任务超出节点 JVM 的负载能力。您需要注意观察实例正在执行的任务，或调整实例的配置。
1分钟负载	实例1分钟所有节点的平均负载 load_1m，指标来源：Logstash 节点监控 api：_node/stats/process?pretty	load_1m 过高时，建议调大实例节点规格。
磁盘使用率	每单位统计周期内（1分钟），实例各个节点的磁盘使用率的平均值	磁盘使用率过高会导致 Logstash 无法正常工作。可对实例进行扩容，增加单节点的磁盘容量。
Events 接收速率	Logstash 实例在统计周期内各节点 Events 接收速率的总和	Logstash 各个节点上的所有管道每秒接收 Events 数量的总和。
Events 发送速率	Logstash实例在统计周期内各节点 Events 发送速率的总和	Logstash 各个节点上的所有管道每秒发送 Events 数量的总和。

Events 延迟	Logstash 实例在统计周期内各节点 Events 处理延迟的平均值	Logstash 节点 Events 处理延迟的平均值。

配置 X-Pack 监控

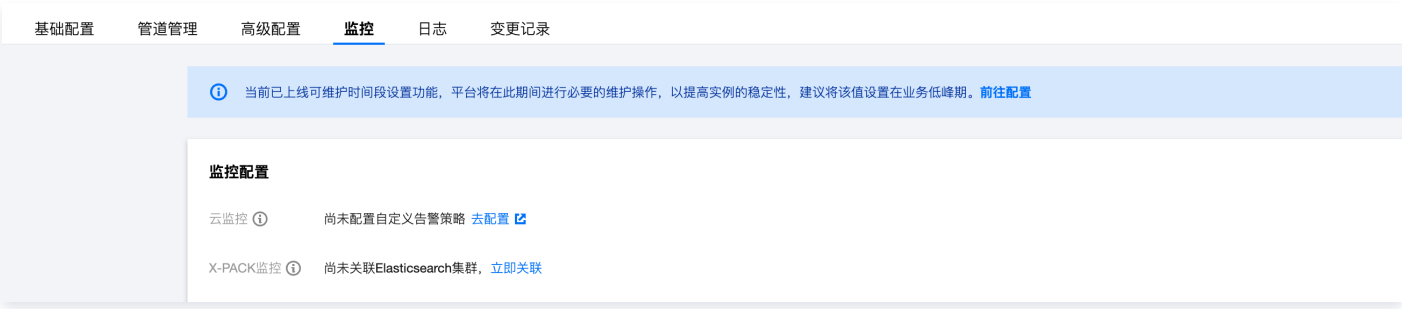
最近更新时间：2024-10-16 10:03:41

本文主要介绍通过配置 X-Pack，来通过 Kibana 监控腾讯云 Logstash 服务。

> **说明**
> - 对于 X-Pack 版 Logstash，关联基础版或白金版腾讯云 ES 实例后，可以在 Kibana 中监控 Logstash 服务，开源版 Logstash 不支持此能力。
> - Logstash 实例需要和 ES 实例在同一个 VPC 内，且大版本相同。

操作步骤

1. 登录 Elasticsearch Service 控制台，在左侧导航栏单击 **Logstash 管理**，进入 Logstash 管理列表页。
2. 单击实例列表中要操作的实例的 **ID/名称**，进入实例基本信息页，然后切换到**监控**页签。在“监控配置”中，单击“X-Pack 监控”中的**立即关联**。

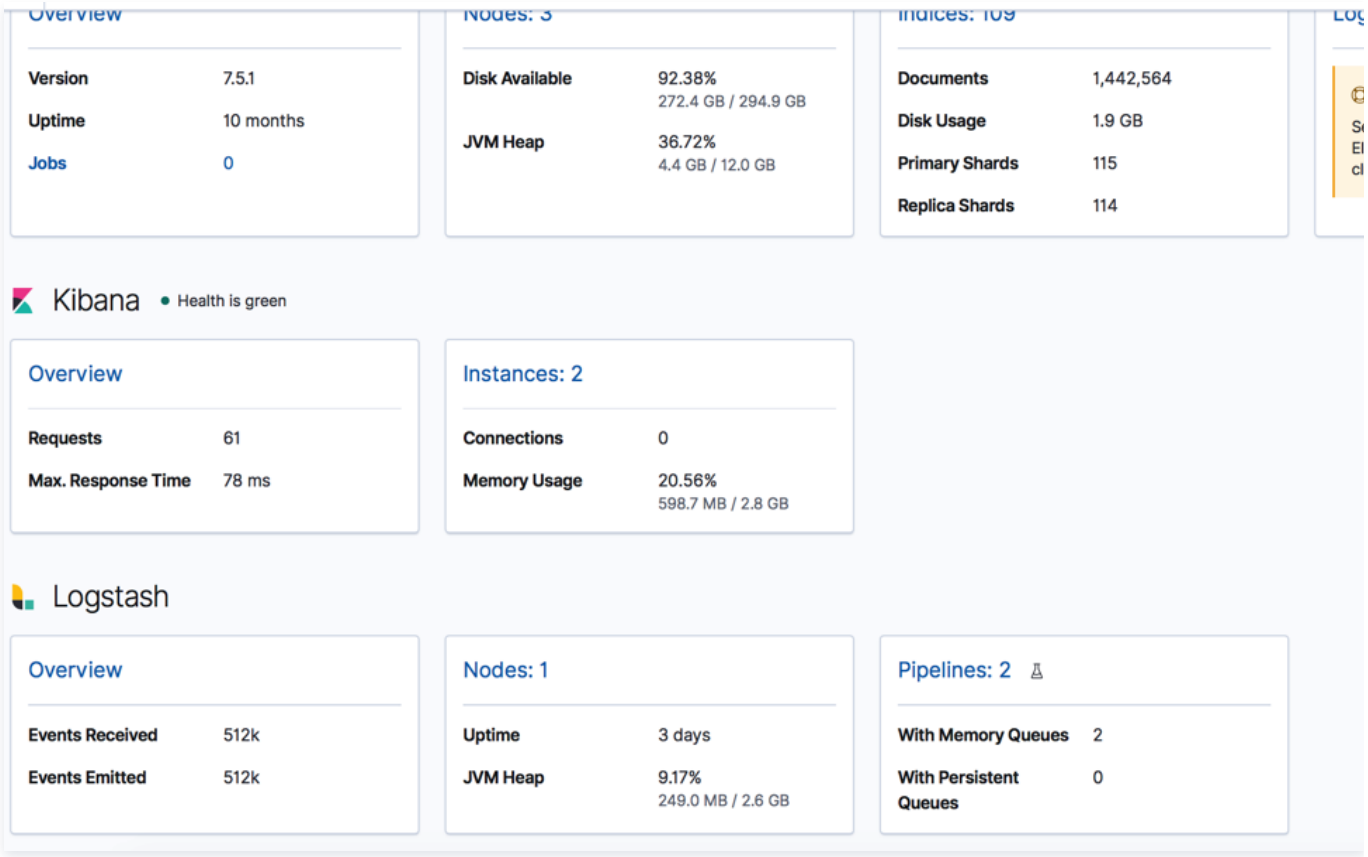


3. 在弹窗中选择要关联的腾讯云 Elasticsearch 实例，单击**确定**。
   Elasticsearch 集群：选择要关联的腾讯云 Elasticsearch 集群，需要与 Logstash 实例在相同 VPC，且大版本相同。

> **注意**
> 关联操作涉及修改 X-Pack 配置，会触发实例重启。

4. 查看 Logstash 监控信息。
   实例重启完成后，“X-Pack 监控”状态变为开启，同时会显示当前关联的腾讯云 Elasticsearch 实例。
   - 在**监控**页签，单击**前往 Kibana 控制台**，跳转到 Kibana 控制台。

- 登录 Kibana 控制台后，在左侧导航栏单击 **Stack Monitoring** 切换到监控页面，在 **Logstash** 区域就可以相应的监控信息。



配置告警

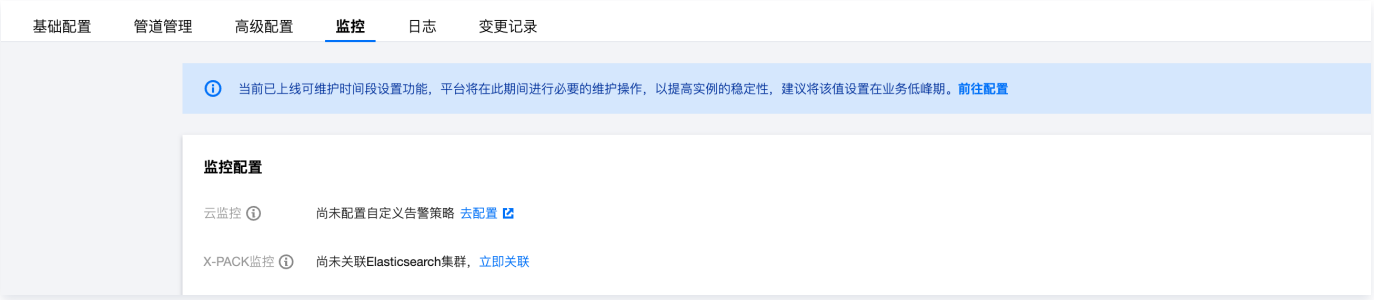
最近更新时间：2025-04-01 16:02:21

腾讯云 Logstash 提供一些关键指标的配置告警功能，配置告警可帮助您及时发现实例问题并进行处理。本文为您介绍通过控制台配置告警的操作。

查看实例是否已配置告警

1. 登录 Elasticsearch Service 控制台，在左侧导航栏单击 **Logstash 管理**，进入 Logstash 管理列表页。
2. 单击实例列表中要操作的实例的 **ID/名称**，进入实例基本信息页，然后切换到**监控**页签。在“监控配置”中，可查看实例是否已经配置了告警。若未配置告警策略，强烈建议您配置告警策略，以便及时获取并处理实例运行的状况及风险，保障服务的稳定。

> **说明**
> 也可登录 腾讯云可观测平台控制台，在**告警配置 > 告警策略**，通过筛选策略和产品，查询某个实例是否已经配置了告警策略。



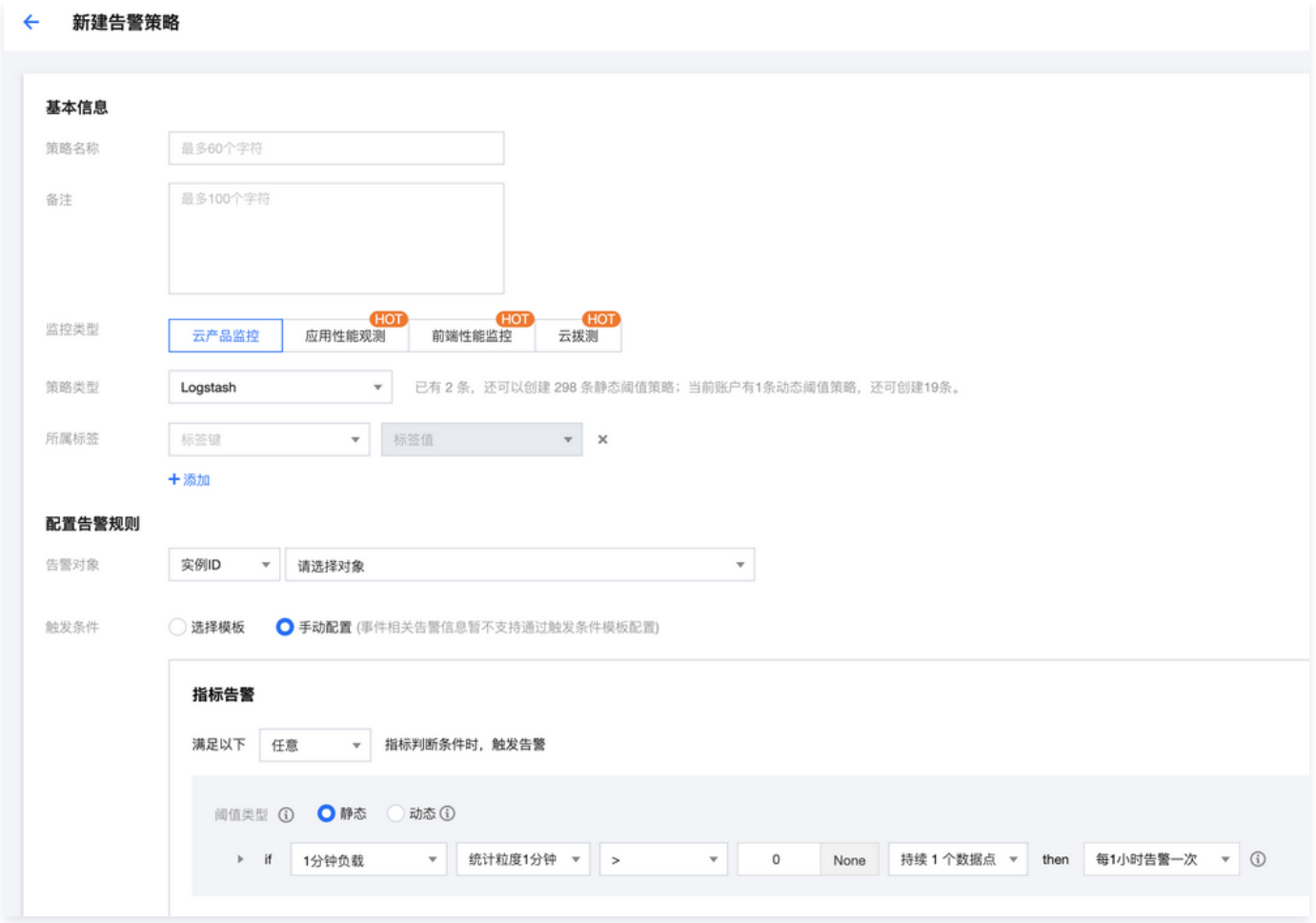
自定义告警配置

1. 登录 腾讯云可观测平台控制台，在**告警管理 > 策略管理**页面，单击**新建策略**。
2. 在新建策略页面，配置策略参数。
   - **策略类型**：选择 Logstash。
   - **告警对象**：选择需要配置告警策略的实例。
   - **触发条件**：支持**选择模板**和**手动配置**，默认选择手动配置，详细配置如下。新建模板可参见 新建触发条件模板 。

> **说明**
> - 指标：例如“CPU 使用率”，统计周期为1分钟或5分钟。因 Logstash 实例的各项指标都是1分钟采集1次，所以选择统计周期是1分钟时，当实例出现一次超过阈值就会触发告警，如果选择5分钟，则5分钟内，连续超过阈值才会触发告警。
> - 告警频次：例如“每30分钟警告一次”，指每30分钟内，连续多个统计周期指标都超过了阈值，如果有一次告警，30分钟内就不会再次进行告警，直到下一个30分钟，如果指标依然超过阈值，才会再次告警。

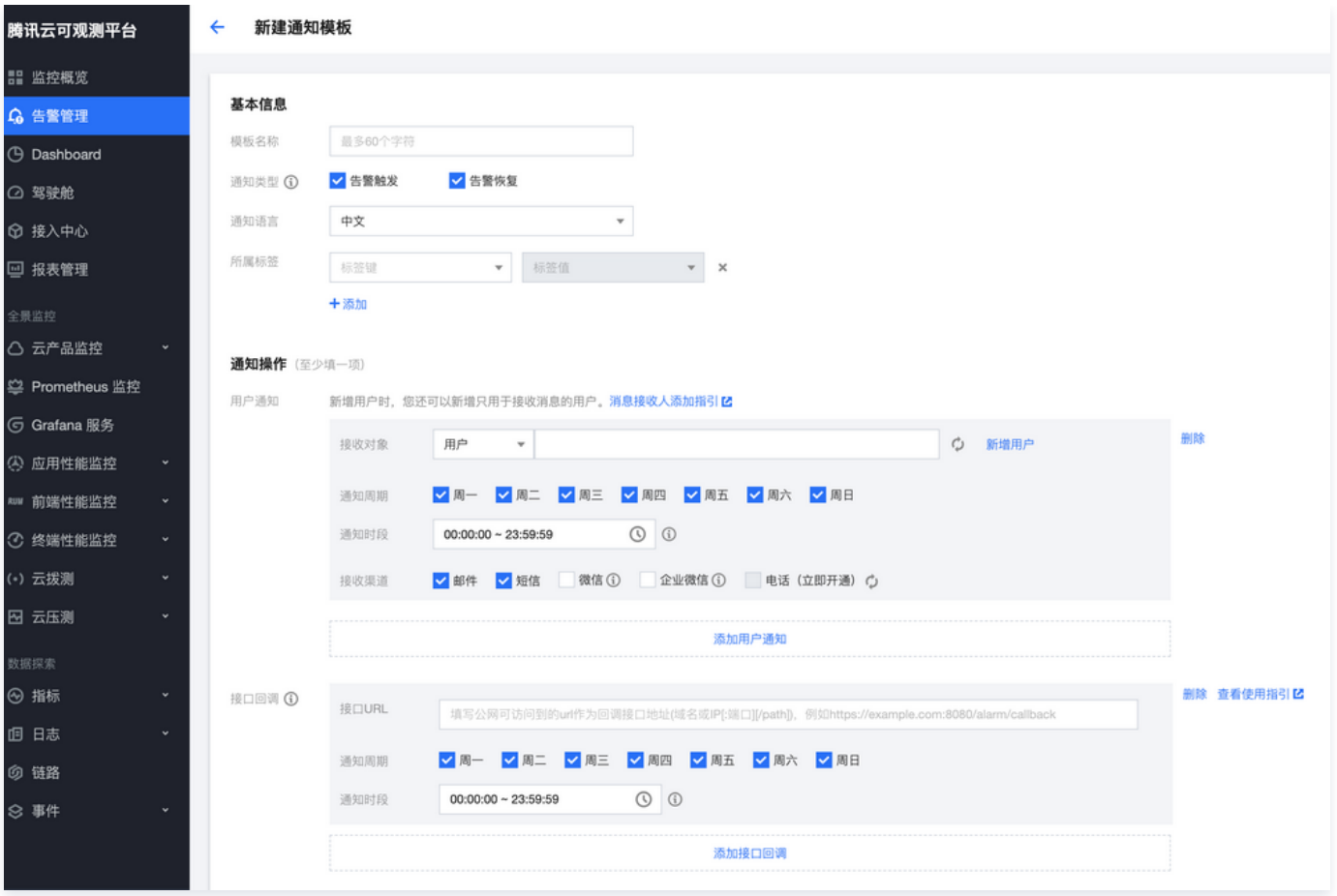
   - **告警渠道**：选择接收组、有效时段、接收渠道。配置方法可参见 新建接收人（组） 。
3. 配置完成后，单击**完成**，跳转到**策略管理**列表，即可看到刚配置的告警策略。

> **说明**
> 告警策略更详细配置教程可参见 告警配置 。



新建触发条件模板

1. 登录 腾讯云可观测平台控制台，在**策略管理**页面，单击**新建策略**。
2. 在新建策略页面，**配置告警规则** > **触发条件**选中**选择模板**，然后单击**新增触发条件模板**，进入触发条件模板列表页。
3. 在触发条件模板列表页单击**新建**，进入在新建触发条件模板页，配置策略类型。
   - **策略类型**：选择 Logstash。
   - **触发条件**：配置需要的触发条件。



4. 确认无误后，单击**保存**。返回新建告警策略页，刷新页面，即出现刚配置的告警策略模板。

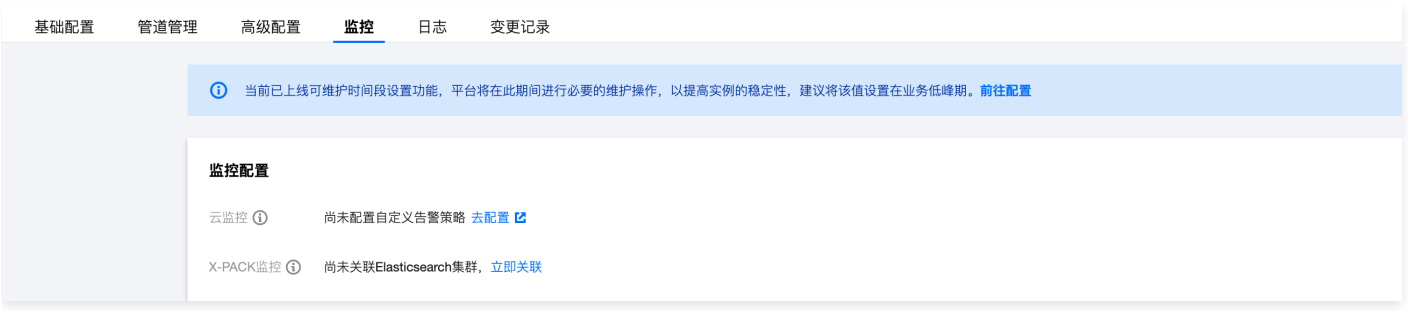
查询日志

最近更新时间：2024-10-16 10:03:41

本文为您介绍 Logstash 实例运行日志的使用说明。用户可以通过实例的运行日志，可以了解实例的运行状况、定位问题、辅助实例的应用开发和运维。

查询实例日志

1. 登录 Elasticsearch Service 控制台，在左侧导航栏单击 **Logstash 管理**，进入 Logstash 管理列表页。
2. 选择要查询实例日志的实例，单击 **ID/名称**，进入实例基本信息页。
3. 在实例基本信息页，切换到**日志**页签，即可查看实例的运行日志。
   - 日志类型：主日志、慢日志和 GC 日志，日志内容包括日志时间、日志级别，以及具体信息等。
   - 默认提供实例7天内的运行日志，按时间倒序展示，用户可以按时间和关键字进行查询。
4. 在日志页面的搜索框，可以按照时间范围和关键字查询相关日志，关键字查询语法同 lucene 查询语法一致。
   - 输入关键词查询，例如：“logstash”。
   - 指定字段设置关键词，例如：`message:logstash` 。
   - 多个条件组合：`level:INFO AND ip:x.x.x.x`，可以查询相关日志。



日志说明

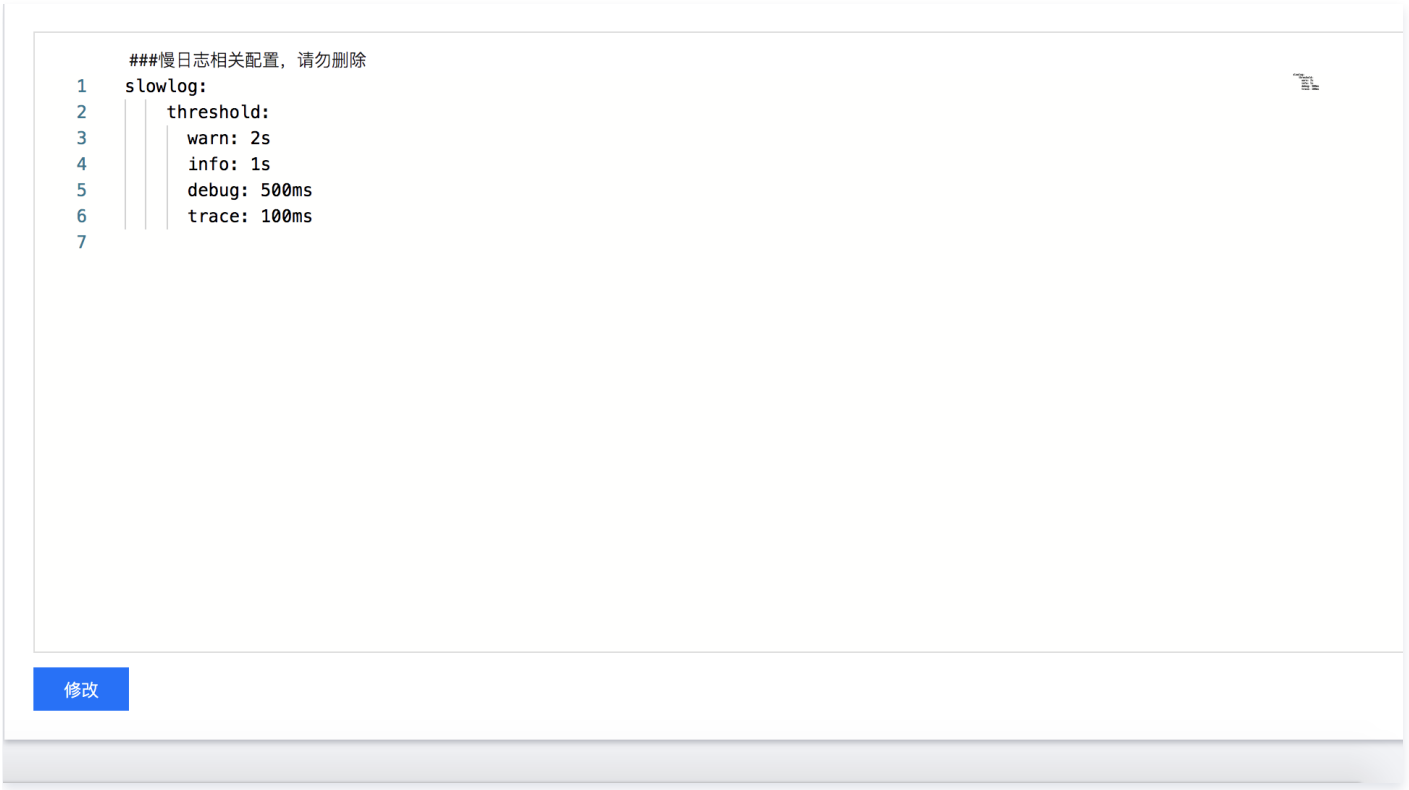
主日志

展示实例运行产生日志的时间、级别、信息等，有 INFO、WARN、DEBUG 等不同级别。

```
[2021-04-20T16:49:21,909][INFO ][logstash.setting.writabledirectory] Creating directory
{:setting=>"path.queue", :path=>"/usr/local/service/logstash/data/queue"}
[2021-04-20T16:49:21,927][INFO ][logstash.setting.writabledirectory] Creating directory
{:setting=>"path.dead_letter_queue", :path=>"/usr/local/service/logstash/data/dead_letter_queue"}
[2021-04-20T16:49:22,316][INFO ][logstash.runner          ] Starting Logstash
{"logstash.version"=>"6.8.13"}
[2021-04-20T16:49:22,342][INFO ][logstash.agent           ] No persistent UUID file found. Generating new
UUID {:uuid=>"2233f3e8-369c-4252-9322-8e2ee8b3371c", :path=>"/usr/local/service/logstash/data/uuid"}
[2021-04-20T16:49:29,260][INFO ][logstash.pipeline        ] Starting pipeline {:pipeline_id=>"ls-ngc79myh-
temp-pipeline", "pipeline.workers"=>1, "pipeline.batch.size"=>125, "pipeline.batch.delay"=>50}
[2021-04-20T16:49:29,367][INFO ][logstash.pipeline        ] Pipeline started successfully
{:pipeline_id=>"ls-ngc79myh-temp-pipeline", :thread=>"#<Thread:0x4f68fc17 run>"}
[2021-04-20T16:49:29,430][INFO ][logstash.agent           ] Pipelines running {:count=>1,
:running_pipelines=>[:"ls-ngc79myh-temp-pipeline"], :non_running_pipelines=>[]}
[2021-04-20T16:49:29,628][INFO ][logstash.pipeline        ] Pipeline has terminated {:pipeline_id=>"ls-
ngc79myh-temp-pipeline", :thread=>"#<Thread:0x4f68fc17 run>"}
[2021-04-20T16:49:29,730][INFO ][logstash.agent           ] Successfully started Logstash API endpoint
{:port=>9600}
```

慢日志

通过慢日志可以查看哪些 event 在管道传输过程中耗时比较久。默认情况下，慢日志不开启。开启 Logstash 实例的慢日志，可通过在控制台中修改 YML 配置完成。YML 配置修改完成后，单击**修改**，在弹出页面单击**保存**，确认可以重启实例后，即可开启慢日志。



GC 日志

Logstash 默认开启 GC 日志，典型的 GC 日志内容如下：

```
Java HotSpot(TM) 64-Bit Server VM (25.181-b13) for linux-amd64 JRE (1.8.0_181-b13), built on Jul  7 2018
00:56:38 by "java_re" with gcc 4.3.0 20080428 (Red Hat 4.3.0-8)
Memory: 4k page, physical 16092620k(14287236k free), swap 0k(0k free)
CommandLine flags: -XX:CMSInitiatingOccupancyFraction=75 -
XX:ErrorFile=/usr/local/service/logstash/temp/hs_err_pid%p.log -XX:GCLogFileSize=67108864 -
XX:+HeapDumpOnOutOfMemoryError -XX:HeapDumpPath=/usr/local/service/logstash/temp -
XX:InitialHeapSize=12884901888 -XX:MaxHeapSize=12884901888 -XX:MaxNewSize=348966912 -
XX:MaxTenuringThreshold=6 -XX:NewSize=348966912 -XX:NumberOfGCLogFiles=32 -XX:OldPLABSize=16 -
XX:OldSize=697933824 -XX:+PrintGC -XX:+PrintGCApplicationStoppedTime -XX:+PrintGCDateStamps -
XX:+PrintGCDetails -XX:+PrintGCTimeStamps -XX:+PrintTenuringDistribution -
XX:+UseCMSInitiatingOccupancyOnly -XX:+UseCompressedClassPointers -XX:+UseCompressedOops -
XX:+UseConcMarkSweepGC -XX:+UseGCLogFileRotation -XX:+UseParNewGC
2021-04-20T16:49:04.939+0800: 0.361: Total time for which application threads were stopped: 0.0001580
seconds, Stopping threads took: 0.0000350 seconds
2021-04-20T16:49:04.954+0800: 0.376: Total time for which application threads were stopped: 0.0000788
seconds, Stopping threads took: 0.0000196 seconds
2021-04-20T16:49:05.005+0800: 0.427: Total time for which application threads were stopped: 0.0000990
seconds, Stopping threads took: 0.0000332 seconds
2021-04-20T16:49:05.408+0800: 0.829: Total time for which application threads were stopped: 0.0002044
seconds, Stopping threads took: 0.0000273 seconds
2021-04-20T16:49:05.787+0800: 1.209: Total time for which application threads were stopped: 0.0002146
seconds, Stopping threads took: 0.0000266 seconds
2021-04-20T16:49:06.262+0800: 1.684: Total time for which application threads were stopped: 0.0001038
seconds, Stopping threads took: 0.0000253 seconds
2021-04-20T16:49:06.285+0800: 1.707: Total time for which application threads were stopped: 0.0000858
seconds, Stopping threads took: 0.0000486 seconds
```

实践教程
接收 Filebeat 发送的数据并写入到 Elasticsearch

最近更新时间：2024-10-16 10:03:41

Logstash 的一个典型应用场景，就是接收 filebeat 发送过来的数据然后写入到 Elasticsearch，使用腾讯云的 Logstash 产品，可以通过简单的配置快速地完成这一过程。

创建管道

1. 登录 Elasticsearch Service 控制台，选择需要操作的实例，单击实例 ID/名称，进入实例基本信息页面。切换到“管道管理”页签，单击**新建管道**，创建一个管道。

2. 进入新建管道页面，单击**引用模板**，同时引用“input-beats”和“output-elasticsearch”两个模板：

3. 在管道配置中，分别针对“input-beats”和“output-elasticsearch”进行配置，一些关键的配置参数说明如下：

input-beats

- host：logstash 要监听的 IP 地址，可设置为节点的 IP，默认为0.0.0.0。
- port：logstash 要监听的端口号，默认为5044。
- type：标识字段

查看更多参数，详情可参见 input-beats 。

output-elasticsearch

- hosts：elasticsearch 集群地址列表
- user：elasticsearch 集群账号
- password：elasticsearch 集群密码
- index：索引名称
- document_type：索引 type，对于不同版本的 ES 集群，该字段有不同的默认值，5.x及以下版本的集群，默认会使用 input 中指定的 type 字段。如果 type 字段不存在，则该字段的值为 doc；6.x版本的集群，该字段默认值为 doc；7.x版本的集群，该字段默认值为_doc；8.x版本的集群，不会使用该字段。

- document_id：文档 ID

查看更多参数，详情可参见 output-elasticsearch 。

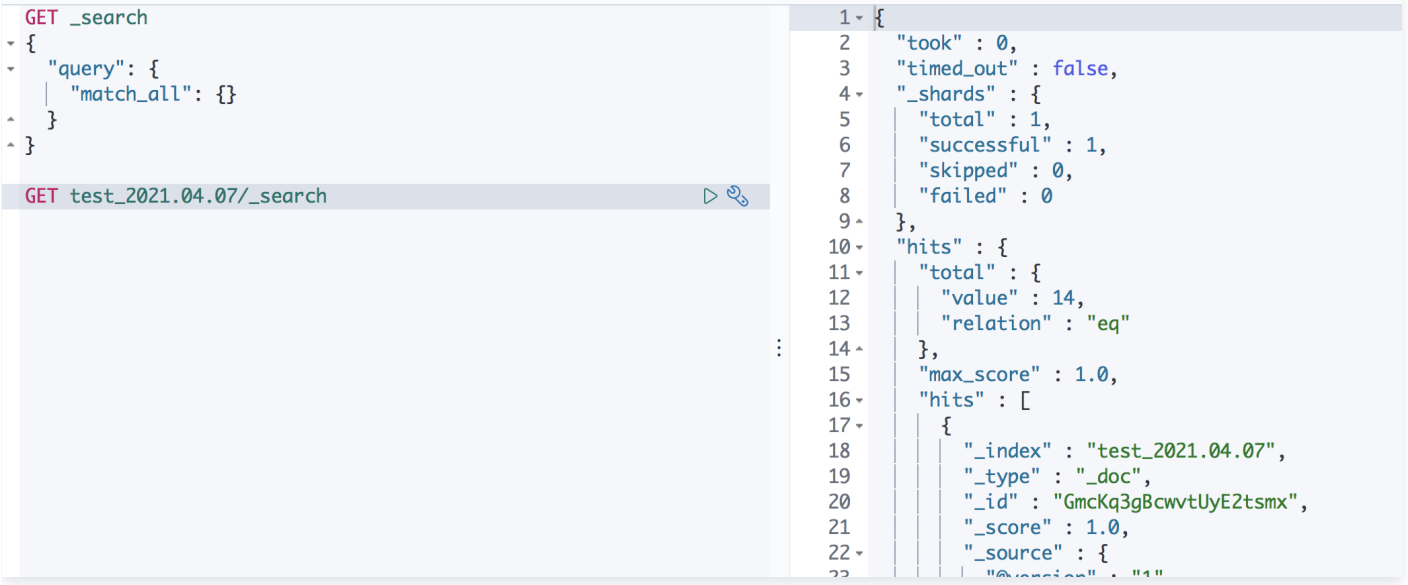
在配置完管道后，单击**保存并部署**即可创建一个管道并自动部署。

查看日志

在控制台查看 Logstash 的运行日志，如果没有 ERROR 级别的日志，则说明管道运行正常。

查看数据写入情况

进入到 output-elasticsearch 中定义的输出端的 ES 集群对应的 kibana 页面，在 Dev tools 工具栏里查看索引是否存在，以及索引的文档数量是否正确。



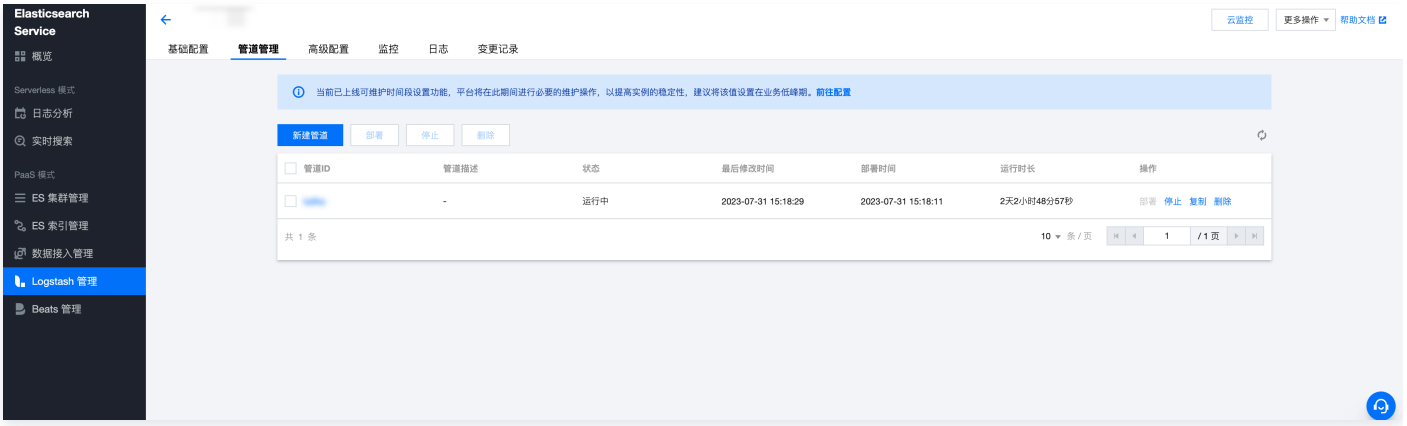
同步 MySQL 中的数据到 Elasticsearch

最近更新时间：2024-10-16 10:03:41

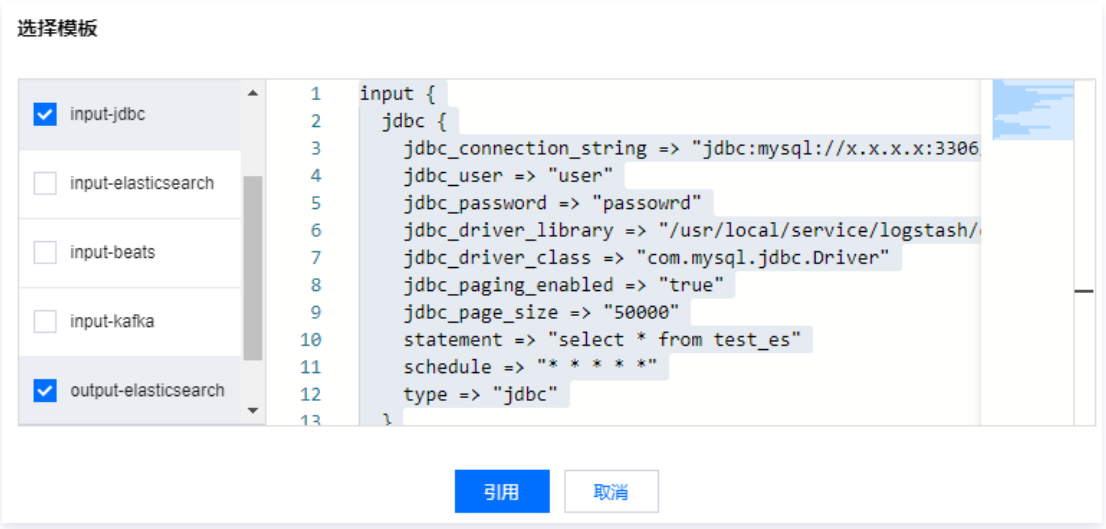
使用 Logstash 可以把关系型数据库如 mysql、postgresql 中的数据同步到其它存储介质，下面介绍如何使用腾讯云 Logstash 同步 mysql 中的数据到 Elasticsearch。

创建管道

1. 登录 Elasticsearch Service 控制台，选择需要操作的实例，单击实例 ID/名称，进入实例基本信息页面。切换到“管道管理”页签，单击**新建管道**，创建一个管道。



2. 进入新建管道页面，单击**引用模板**，同时引用“input-jdbc”和“output-elasticsearch”两个模板：



3. 在管道配置中，分别针对“input-jdbc”和“output-elasticsearch”进行配置，一些关键的配置参数说明如下：

input-jdbc

- jdbc_connection_string：数据库连接地址
- jdbc_user：数据库账号
- jdbc_password：数据库账号密码
- jdbc_driver_library：jdbc 驱动 jar 包，在 Logstash 节点的 `/usr/local/service/logstash/extended-files` 目录下，有大多数版本的 mysql 以及 postgresql 数据库的 jdbc 驱动 jar 包，可根据需要直接引用，可用的驱动 jar 包列表如下：
  - mysql-connector-java-5.1.27.jar
  - mysql-connector-java-5.1.35.jar
  - mysql-connector-java-5.1.39-bin.jar
  - mysql-connector-java-5.1.39.jar
  - mysql-connector-java-5.1.40.jar

  - mysql-connector-java-5.1.43.jar
  - mysql-connector-java-5.1.47.jar
  - mysql-connector-java-5.1.48.jar
  - mysql-connector-java-5.1.9.jar
  - mysql-connector-java-6.0.2.jar
  - mysql-connector-java-6.0.6.jar
  - mysql-connector-java-8.0.11.jar
  - mysql-connector-java-8.0.17.jar
  - mysql-connector-java-8.0.18.jar
  - postgresql-42.0.0.jar
  - postgresql-42.1.4.jar
  - postgresql-42.2.0.jar
  - postgresql-42.2.10.jar
  - postgresql-42.2.13.jar
  - postgresql-42.2.1.jar
  - postgresql-42.2.8.jar
- jdbc_driver_class：jdbc 驱动类，对于 mysql 可填写“com.mysql.jdbc.Driver”，postgresql 可填写“org.postgresql.Driver”
- jdbc_paging_enabled：从数据库批量拉取数据时是否开启分页，可选值"true"或者"false"
- jdbc_page_size：jdbc 分页大小
- statement：用于拉取数据的 sql 语句
- tracking_column：当在 statement 中指定了 sql_last_value 用于记录读取数据的 offset 时，使用数据库表中的哪个字段的值来记录 offset。
- use_column_value：当在 statement 中指定了 sql_last_value 用于记录读取数据的 offset 时，是否使用数据库表中的字段；设置为 true 则使用 tracking_column 定义的字段，否则使用前一次 sql 语句执行时的时间戳。
- schedule：是否开启定时任务持续执行 sql 语句，不设置的话则只会执行一次 sql 语句，执行结束后管道自动结束。
- type：标识字段

查看更多参数的具体含义，详情可参见 logstash-input-jdbc 。

output-elasticsearch

- hosts：elasticsearch 集群地址列表
- user：elasticsearch 集群账号
- password：elasticsearch 集群密码
- index：索引名称
- document_type：索引 type，对于不同版本的 ES 集群，该字段有不同的默认值，5.x及以下版本的集群，默认会使用 input 中指定的 type 字段。如果 type 字段不存在，则该字段的值为 doc；6.x版本的集群，该字段默认值为 doc；7.x版本的集群，该字段默认值为_doc；8.x版本的集群，不会使用该字段。
- document_id：文档 ID

查看更多参数，详情可参见 output-elasticsearch 。

在配置完管道后，单击**保存并部署**即可创建一个管道并自动部署。

实战案例

全量同步 mysql 表中的数据到 Elasticsearch

当 mysql 的某张表不再进行写入时，可使用如下配置全量地把数据同步到 Elasticsearch 集群中，管道配置如下：

```
input {
  jdbc {
    jdbc_connection_string => "jdbc:mysql://x.x.x.x:3306/logstash_test"
    jdbc_user => "user"
    jdbc_password => "xxxxx"
    jdbc_driver_library => "/usr/local/service/logstash/extended-files/mysql-connector-java-5.1.40.jar"
    jdbc_driver_class => "com.mysql.jdbc.Driver"
    jdbc_paging_enabled => "true"
    jdbc_page_size => "5000"
    statement => "select * from newTable0"
    type => "jdbc"
  }
}
output {
    elasticsearch {
        hosts => ["http://x.x.x.x:9200"]
        user => "elastic"
        password => "xxxxx"
        index => "newTable0"
    }
}
```

增量同步 mysql 表中的数据到 Elasticsearch

当 mysql 的某张表在持续写入时，可使用如下配置，通过 sql_last_value 记录 offset，把数据增量地同步到 Elasticsearch 集群中，管道配置如下：

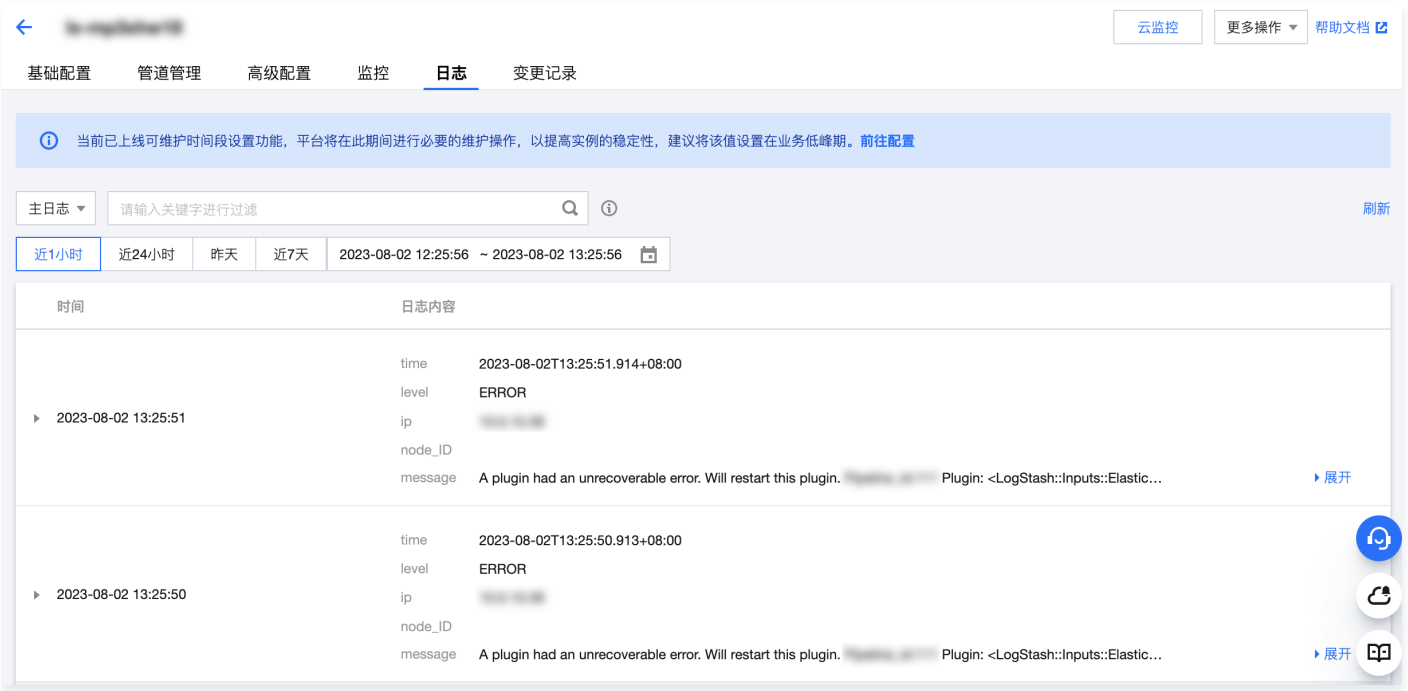
```
input {
  jdbc {
    jdbc_connection_string => "jdbc:mysql://x.x.x.x:3306/logstash_test"
    jdbc_user => "user"
    jdbc_password => "xxxxx"
    jdbc_driver_library => "/usr/local/service/logstash/extended-files/mysql-connector-java-5.1.40.jar"
    jdbc_driver_class => "com.mysql.jdbc.Driver"
    jdbc_paging_enabled => "true"
    jdbc_page_size => "5000"
    statement => "select * from newTable0 where id > :sql_last_value"
```

```
    use_column_value => true
    tracking_column => "id"
    schedule => "* * * * *"
    last_run_metadata_path => "/usr/local/service/logstash/temp/jdbc-sql_last_value.yml"
    type => "jdbc"
  }
}
output {
    elasticsearch {
        hosts => ["http://x.x.x.x:9200"]
        user => "elastic"
        password => "xxxxx"
        index => "newTable0"
    }
}
```

上述配置中指定了 tracking_column 为字段"id"，需要数据表中包含一个自增的"id"字段，当然可以根据实际情况使用不同的字段。

查看日志

在控制台中查看日志，如果没有 ERROR 级别的日志，则说明管道配置没有问题。



查看数据写入情况

进入 output-elasticsearch 中定义的输出端的 ES 集群对应的 kibana 页面，在 Dev tools 工具栏里查看索引是否存在，以及索引的文档数量是否正确。

```
GET _search
{
  "query": {
    "match_all": {}
  }
}

GET test_2021.04.07/_search
```

```
{
  "took" : 0,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 14,
      "relation" : "eq"
    },
    "max_score" : 1.0,
    "hits" : [
      {
        "_index" : "test_2021.04.07",
        "_type" : "_doc",
        "_id" : "GmcKq3gBcwvtUyE2tsmx",
        "_score" : 1.0,
        "_source" : {
```

同步两个 Elasticsearch 集群中的数据

最近更新时间：2024-10-16 10:03:41

使用 Logstash 可以完成同步两个 Elasticsearch 集群中的数据，例如把数据从自建的 Elasticsearch 集群同步到腾讯云上的 Elasticsearch 集群，或者同步两个腾讯云上的 Elasticsearch 集群中的数据。下面介绍如何使用腾讯云 Logstash 同步两个 Elasticsearch 集群中的数据。

创建管道

登录 Elasticsearch Service 控制台，选择需要操作的实例，单击实例 ID/名称，进入实例基本信息页面。切换到“管道管理”页签，单击**新建管道**，创建一个管道。

进入新建管道页面，单击**引用模板**，同时引用“input-elasticsearch”和“output-elasticsearch”两个模板：

在管道配置中，分别针对“input-elasticsearch”和“output-elasticsearch”进行配置，一些关键的配置参数说明如下：

input-elasticsearch

- hosts：elasticsearch 集群地址列表
- user：elasticsearch 集群账号
- password：elasticsearch 集群密码
- index：索引名称
- query：es 查询语句，用于查询某一部分的数据。
- schedule：是否开启定时任务持续从 elasticsearch 集群中拉取数据，如果不配置，则只会拉取一次。
- scroll：批量从 elasticsearch 集群中拉取数据时，用于保持 scroll context 的时间，默认为"1m"
- size：批量从 elasticsearch 集群中拉取数据时，每个批次拉取多少条数据，默认为1000。
- type：标识字段
- docinfo：是否在 event 中填充索引名称，type 以及 id 等文档元信息，默认为 false。

查看更多参数，详情可参见 input-elasticsearch 。

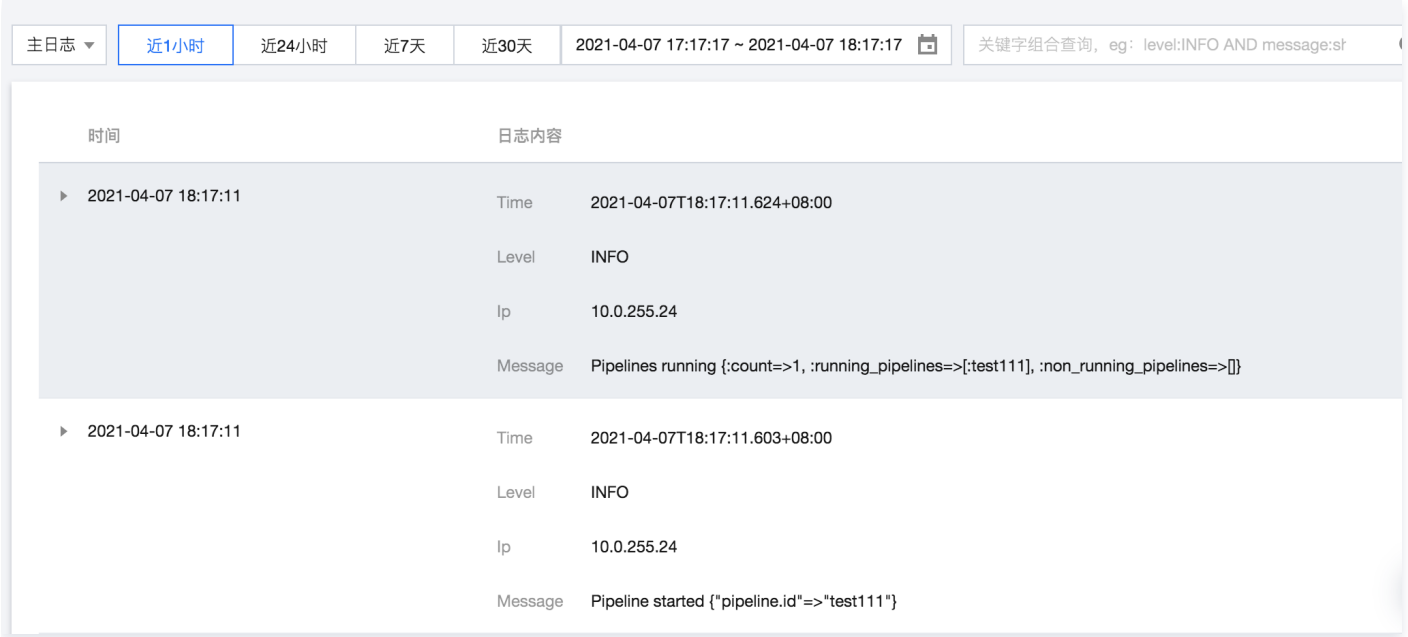
output-elasticsearch

- hosts：elasticsearch 集群地址列表
- user：elasticsearch 集群账号
- password：elasticsearch 集群密码
- index：索引名称
- document_type：索引 type，对于不同版本的 ES 集群，该字段有不同的默认值，5.x及以下版本的集群，默认会使用 input 中指定的 type 字段。如果 type 字段不存在，则该字段的值为 doc；6.x版本的集群，该字段默认值为 doc；7.x版本的集群，该字段默认值为_doc；8.x版本的集群，不会使用该字段。
- document_id：文档 ID

查看更多参数，详情可参见 output-elasticsearch 。
在配置完管道后，单击**保存并部署**即可创建一个管道并自动部署。

查看日志

在控制台查看 Logstash 的运行日志，如果没有 ERROR 级别的日志，则说明管道运行正常。



查看数据写入情况

进入到 output-elasticsearch 中定义的输出端的 ES 集群对应的 kibana 页面，在 Dev tools 工具栏里查看索引是否存在，以及索引的文档数量是否正确，在下图框中写入索引：

```
GET _search
{
  "query": {
    "match_all": {}
  }
}

GET [redacted]/_search
```


```
{
  "took" : 0,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 14,
      "relation" : "eq"
    },
    "max_score" : 1.0,
    "hits" : [
      {
        "_index" : "test_2021.04.07",
        "_type" : "_doc",
        "_id" : "GmcKq3gBcwvtUyE2tsmx",
        "_score" : 1.0,
        "_source" : {
```

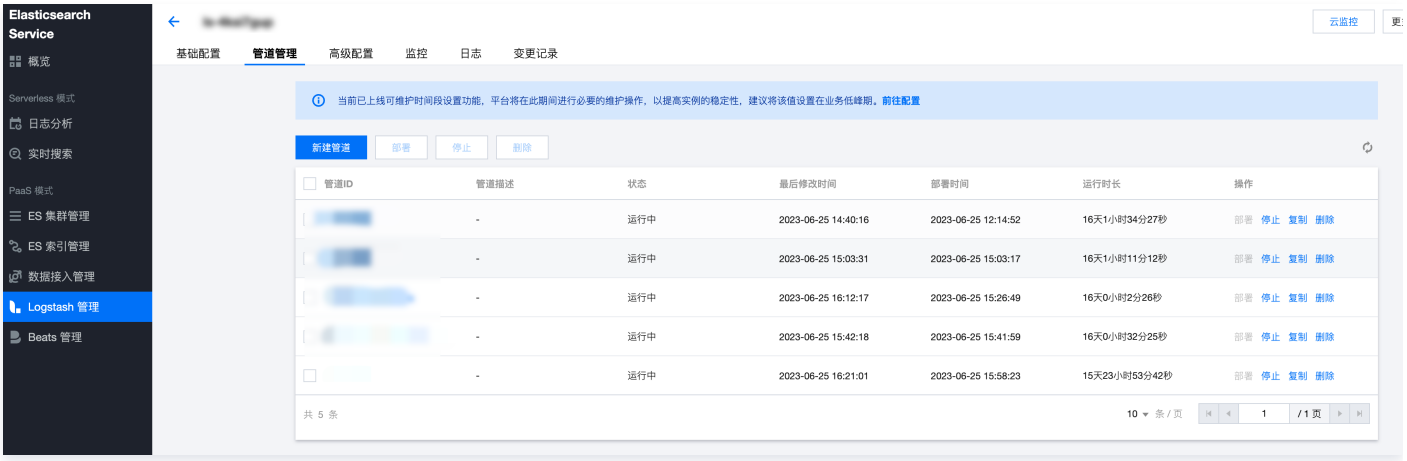
消费 kafka 数据并写入到 Elasticsearch

最近更新时间：2024-10-15 22:22:01

Logstash 的一个典型应用场景，就是消费 kafka 中的数据并且写入到 Elasticsearch，使用腾讯云的 Logstash 产品，可以通过简单的配置快速地完成这一过程。

创建管道

1. 登录 Elasticsearch Service 控制台，单击 **Logstash 管理**，选择需要操作的实例，单击实例 **ID/名称**，进入实例基本信息页面。
2. 切换到“管道管理”页签，单击**新建管道**，创建一个管道。



3. 进入新建管道页面，单击**引用模板**。

同时引用“input-kafka”和“output-elasticsearch”两个模板：

在管道配置中，分别针对“input-kafka”和“output-elasticsearch”进行配置，一些关键的配置参数说明如下：

input-kafka

- bootstrap_servers：kafka 服务端地址列表
- client_id：客户端 ID
- group_id：消费组 ID
- consumer_threads：消费线程数量，建议保持：该参数 × Logstash 实例节点的数量 = topic 的 partitions 数量
- topics：topic 列表
- auto_offset_reset：当 kafka 中 topic 没有初始的 offset 时，如何重置 offset，常用可选值为 earliest（最早）、latest（最新）
- type：标识字段

更多参数详情可参见 input-kafka 。

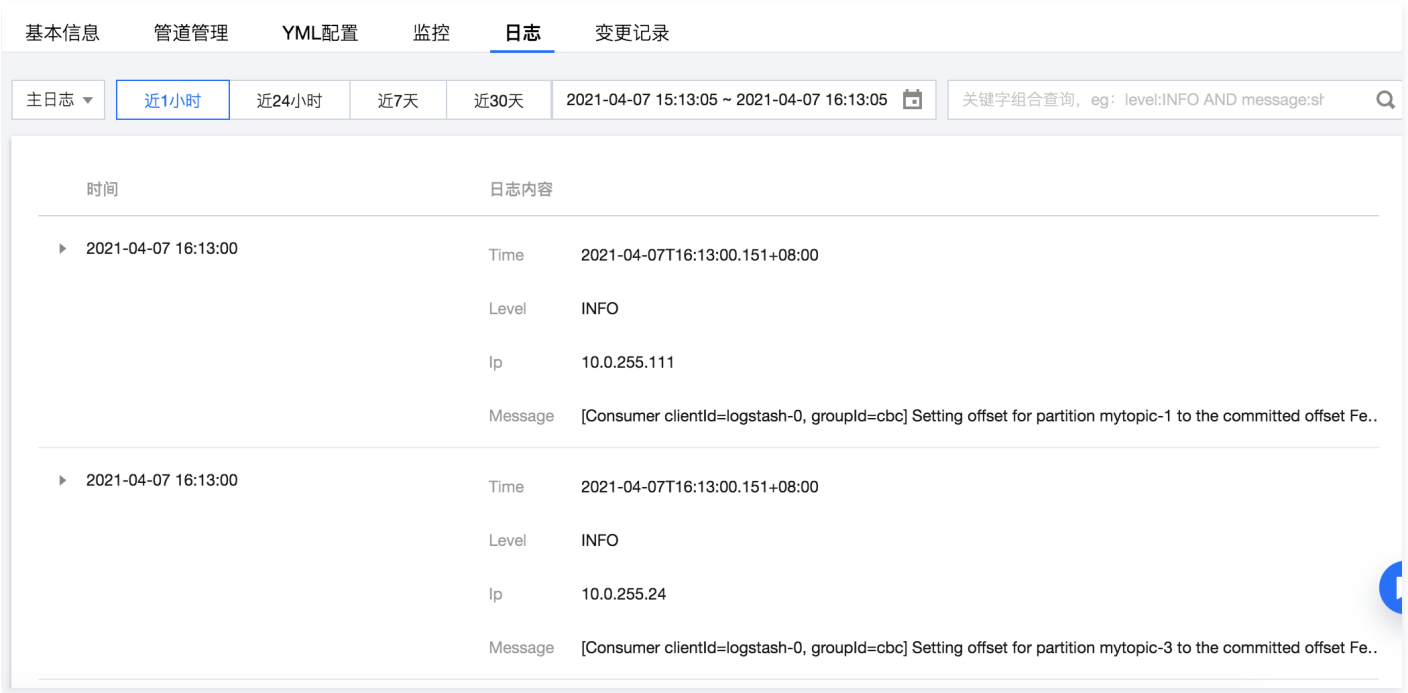
output-elasticsearch

- hosts：elasticsearch 集群地址列表
- user：elasticsearch 集群账号
- password：elasticsearch 集群密码
- index：索引名称
- document_type：索引 type，对于不同版本的 ES 集群，该字段有不同的默认值，5.x及以下版本的集群，默认会使用 input 中指定的 type 字段。如果 type 字段不存在，则该字段的值为 doc；6.x版本的集群，该字段默认值为 doc；7.x版本的集群，该字段默认值为_doc；8.x版本的集群，不会使用该字段。
- document_id：文档 ID

查看更多参数，详情可参见 output-elasticsearch 。
在配置完管道后，单击**保存并部署**即可创建一个管道并自动部署。

查看日志

在控制台查看 Logstash 的运行日志，如果没有 ERROR 级别的日志，则说明管道运行正常。



查看数据写入情况

进入到 output-elasticsearch 中定义的输出端的 ES 集群对应的 kibana 页面，在 Dev tools 工具栏里查看索引是否存在，以及索引的文档数量是否正确。

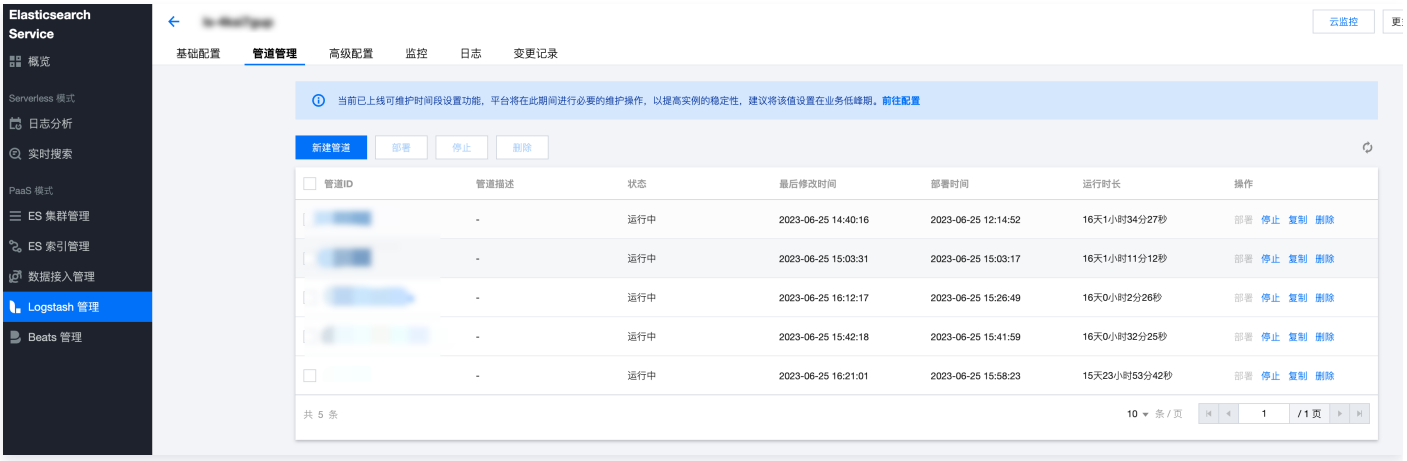
读取 COS 中的日志文件并写入到 Elasticsearch

最近更新时间：2024-10-15 22:03:51

在某些场景中，业务服务端或云上组件的日志会归档存储到对象存储 COS 中，在需要进行查询时，需要从 COS 中获取并查询日志。此时，可借助 Logstash 自动地读取 COS 中指定 bucket 的日志文件，然后写入到 Elasticsearch 中，再使用 Kibana 可视化组件进行查询和分析。

创建管道

1. 登录 Elasticsearch Service 控制台，单击 **Logstash 管理**，选择需要操作的实例，单击实例 **ID/名称**，进入实例基本信息页面。
2. 切换到“管道管理”页签，单击**新建管道**，创建一个管道。



3. 进入新建管道页面，单击**引用模板**。

同时引用“input-s3”和“output-elasticsearch”两个模板：

在管道配置中，分别针对“input-s3”和“output-elasticsearch”进行配置，一些关键的配置参数说明如下：

input-s3

- access_key_id：腾讯云账号的 API 密钥 ID。
- secret_access_key：腾讯云账号的 API 密钥 KEY。
- endpoint：COS 对象存储的访问地址，不同地域的地址不同，如广州地域为 `https://cos.ap-guangzhou.myqcloud.com`
- bucket：COS 对象存储的 bucket。
- region：COS 对象存储 bucket 所在的地域，如 ap-guangzhou。
- prefix：要读取的日志文件名称前缀。

查看更多参数，详情可参见 input-s3 。

output-elasticsearch

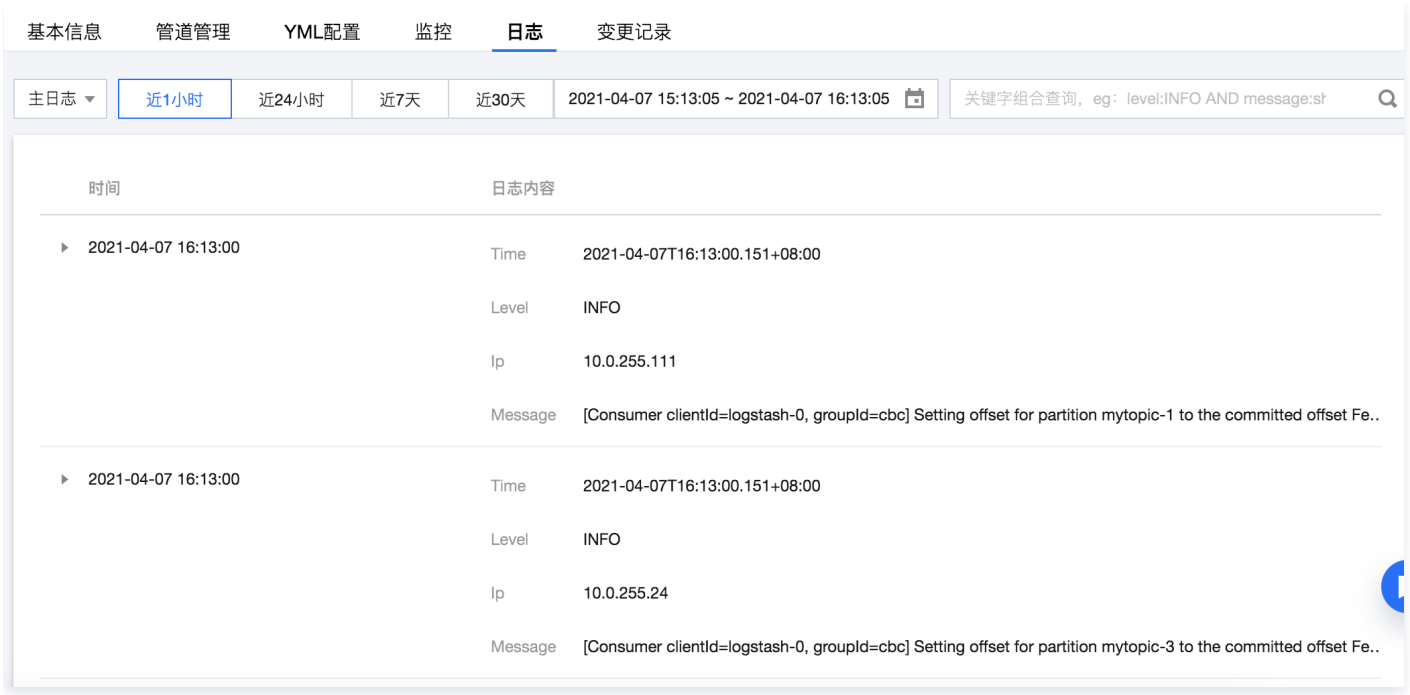
- hosts：elasticsearch 集群地址列表。
- user：elasticsearch 集群账号。
- password：elasticsearch 集群密码。
- index：索引名称。
- document_type：索引 type，对于不同版本的 ES 集群，该字段有不同的默认值，5.x及以下的集群，默认会使用 input 中指定的 type 字段，如果 type 字段不存在，则该字段的值为 doc；6.x的集群，该字段默认值为 doc；7.x的集群，该字段默认值为_doc；8.x的集群，不会使用该字段。
- document_id：文档 ID。

查看更多参数，详情可参见 output-elasticsearch 。

在配置完管道后，单击**保存并部署**即可创建一个管道并自动部署。

查看日志

在控制台查看 Logstash 的运行日志，如果没有 ERROR 级别的日志，则说明管道运行正常。



查看数据写入情况

进入到 output-elasticsearch 中定义的输出端的 ES 集群对应的 kibana 页面，在 Dev tools 工具栏里查看索引是否存在，以及索引的文档数量是否正确。