ALIBABA CLOUD

阿里云

阿里云Elasticsearch

高级监控报警

文档版本：20220701

阿里云

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
危险	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。	危险 重置操作将丢失用户配置数据。
警告	该类警示信息可能会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，恢复业务时间约十分钟。
注意	用于警示信息、补充说明等，是用户必须了解的内容。	注意 权重设置为0，该服务器不会再接受新请求。
说明	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。	单击**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。	在**结果确认**页面，单击**确定**。
Courier字体	命令或代码。	执行 `cd /d C:/window` 命令，进入Windows系统文件夹。
斜体	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。	`switch {active|stand}`

目录

1.高级监控报警概述

高级监控报警服务是基于Elasticsearch开发的一种SAAS服务，具备对集群指标和日志数据的采集、加工、监控、检索、可视化和报警等多种能力，为云上用户提供了一种开箱即用的一站式监控报警解决方案。通过使用高级监控报警服务，您可以实现对所有区域的Elasticsearch集群集中管理，查看或根据业务需要配置监控大屏，自由定制指标及日志报警规则等。此服务能够帮助您更加方便地监控Elasticsearch集群下各维度的信息，实时了解集群状况，及时定位并解决问题。

功能特性

高级监控报警服务支持的功能特性如下表所示。

类别		功能	说明	相关文档
服务		默认自动开通高级监控服务	阿里云Elasticsearch会为您自动开通高级监控报警服务，并将您账号下的存量实例和新购实例接入监控报警服务。	无
监控		指标监控	您可以在指标监控页面监控所有集群的基础指标和引擎侧指标，也可以根据实例、索引和节点等筛选数据，精确掌握实时信息。	• 查看监控指标 • 基础指标 • 引擎指标 • 日志监控 • 配置自定义监控大屏
		日志监控	您可以在日志监控页面查看所有集群的日志概况，也可以根据实例、索引、节点和检索条件等查询各类日志，快速发现和定位问题。	
		自定义监控	您可以在自定义监控页面根据业务需要配置监控大屏，并查看您自定义的监控大屏和日志报警大盘。	
	日志报警	日志报警配置	您可以通过日志监控页面的入口进入到日志报警配置界面，可参考系统为您提供的日志报警模板，自定义日志报警规则并添加报警联系人。	日志报警
		日志报警管理	您可以在Grafana报警页面，查看所有已配置的日志报警规则、修改配置、管理报警规则的生效状态等。	

类别		功能	说明	相关文档
报警	指标报警	指标报警组和报警规则	一个指标报警组可以包含一个或多个报警规则，同一个报警规则可以加入多个报警组。 通过指标报警规则配置，您可以设置多维度的监控指标和Tags，帮助您快速定位Elasticsearch的性能问题，提高运维排查效率。	• 管理报警组 • 配置报警规则 • 查看报警通知记录和事件
		指标报警联系人和联系人组	指标报警联系人组可以包含一个或多个报警联系人。同一个报警联系人，也可以被加入到多个报警联系人组中。 在指标报警规则设置中，您可以添加报警联系人组或报警联系人，将报警通知发送给该组下所有联系人或某个指定的联系人。	• 管理报警联系人 • 管理报警联系人组
		指标报警通知和报警事件	您可以在概览页面，查看所有指标报警组的通知记录和报警事件；也可以在报警组列表页面，查看单个报警组的通知记录和报警事件。	• 通过钉钉群接收报警通知 • 查看报警通知记录和事件 • 查看通知记录 • 查看报警事件

优势

- 海量指标全覆盖

  基于阿里云自身丰富的运维经验，对集群指标和日志的采集实现了全方位覆盖，特别是提供了全面的自研引擎侧指标数据。

- 集中管理

  数据汇总，操作集中，易于管理，方便用户随时掌握集群整体情况；且默认为您账号下的存量实例和新购实例提供高级监控报警服务，便捷省力。

- 数据可视化，监控自定义

  内嵌多组图表大盘，将复杂信息清晰展示，帮助用户快速了解集群状况和变化趋势；支持用户根据自身业务需要自由定制监控大盘和各类报警规则。

2.快速入门

当您使用阿里云Elasticsearch时，系统会为您自动开通高级监控报警服务，并将您账号下的存量和新购Elasticsearch实例接入监控报警服务。本文为您介绍如何查看与配置可视化监控，以及如何配置日志报警规则和指标报警规则。

背景信息

高级监控报警服务能够为您所有地域下的Elasticsearch集群提供全维度指标和日志监控分析服务。您可以在平台为您提供的Grafana中查看集群、节点、索引和机器资源等维度的可视化监控数据，进行集群的异常日志分析，并可以根据业务需求自定义监控大屏和报警规则。关于高级监控报警的更多信息，请参见高级监控报警概述。

前提条件

- 已在支持高级监控报警服务的地域下创建阿里云Elasticsearch实例：
  - 目前高级监控报警服务支持的地域包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、中国香港。
  - 创建实例的具体操作，请参见创建阿里云Elasticsearch实例。
- 熟悉Grafana监控大屏的使用方法。详细信息，请参见Grafana Dashboard。

使用限制

- 高级监控报警功能提供了基础指标、引擎指标和日志数据的监控和报警。阿里云Elasticsearch所有版本都支持对实例的基础指标和日志数据监控，仅内核版本大于1.2.0的6.7.0或7.10.0版本支持引擎指标监控。如果内核版本低于1.2.0，可升级内核版本。具体操作，请参见升级版本。
- 高级监控报警服务存在地域限制，支持的地域仅包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、香港。

操作流程

1. 步骤一：查看和配置可视化监控
2. （可选）步骤二：配置日志报警规则
3. （可选）步骤三：配置指标报警规则

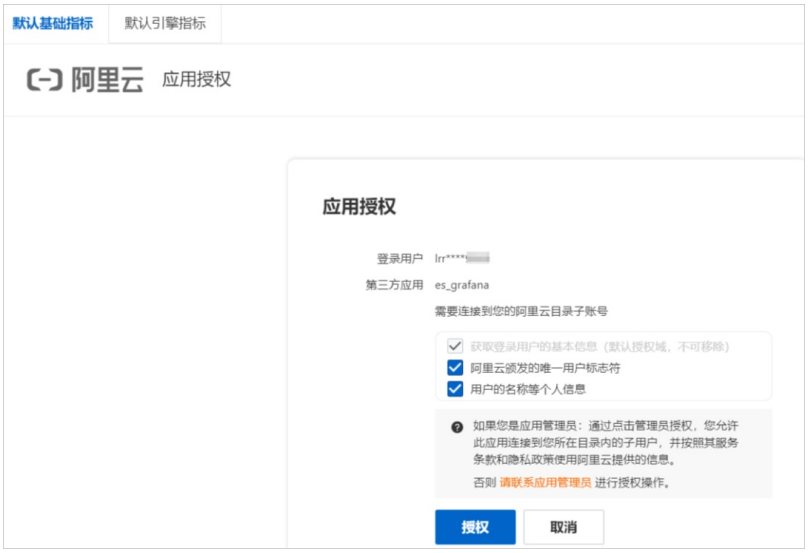
步骤一：查看和配置可视化监控

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。

   系统默认将您账号下的存量和新购Elasticsearch实例全部接入监控报警服务。
3. 在**高级监控报警**页面，查看默认监控。

   i. （可选）应用授权。

   如果您是首次使用监控大盘，则需要进行应用授权。为保证您能正常使用高级监控报警功能，请确保获以下三项授权同时选中。



选项	说明
获取登录用户的基本信息（默认授权域，不可移除）	系统默认已经选中。从当前阿里云账号获取登录用户的基本信息，例如令牌过期时间戳、令牌主体、令牌接收者以及颁发者等信息。
阿里云颁发的唯一用户标志符	需要手动选中。获取当前阿里云账号的UID，以避免多个RAM用户重复授权。
用户的名称等个人信息	需要手动选中。获取当前云账号（可以是阿里云账号，也可以是RAM用户）登录用户名的相关信息，例如用户的显示名称、登录名称，授权之后用户能看到当前登录的用户账号名称。

> 说明
>
> - 如果您使用的是阿里云账号，则按照以上说明同时勾选三项授权即可登录；如果您使用的是RAM用户身份，那么您需要由阿里云账号授权，或者由阿里云账号完成首次的登录授权后，您才可以正常登录。如果由阿里云账号授权，您需要参见为RAM用户授权，将**策略内容**中的Action和Resource替换为以下信息：
>
>   ```
>   Action:ims:*
>   Resource:acs:ims::<yourAccountId>:application/*
>   ```
>
>   其中，<yourAccountId>需要替换成您自己的RAM用户身份ID。
> - 使用RAM角色单点登录阿里云控制台时，不支持访问高级监控报警服务。如果需要访问，可使用RAM用户单点登录阿里云控制台。
> - 如果您的授权出现问题，请通过重新授权进行处理。

ii. 在左侧导航栏，选择**监控可视化 > 指标监控**，查看已接入实例的指标监控数据。

指标监控的详细信息，请参见基础指标和引擎指标。

iii. 在左侧导航栏，选择**监控可视化 > 日志监控**，查看已接入实例的日志监控数据。

日志监控的详细信息，请参见日志监控。

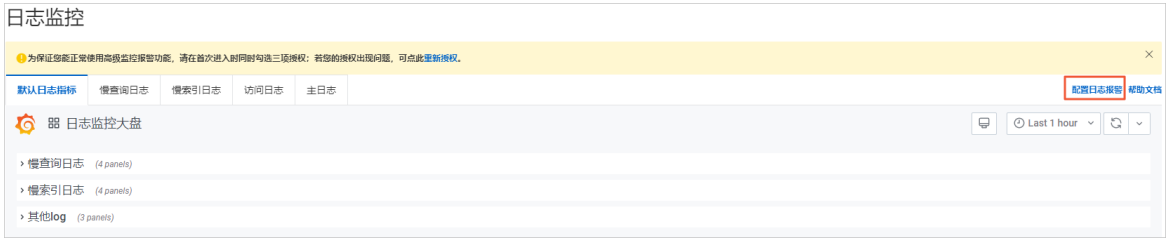
4. 在**高级监控报警**页面的左侧导航栏，选择**监控可视化 > 自定义监控**，配置并查看自定义监控。

具体操作步骤，请参见配置自定义监控大屏。

（可选）步骤二：配置日志报警规则

如果您需要通过监控日志进行报警通知，请执行以下操作：

1. 在**高级监控报警**页面的左侧导航栏中，选择**监控可视化 > 日志监控**。
2. 在**默认日志指标**页签右侧，单击**配置日志报警**。



3. 参考系统为您提供的报警模板，配置日志报警规则或自由定制日志报警规则。

   详细操作步骤，请参见日志报警。

（可选）步骤三：配置指标报警规则

如果您需要通过监控指标进行报警通知，请在**高级监控报警**页面的左侧导航栏中，执行以下操作：

1. 选择**指标报警模块 > 报警组列表**，创建报警组并添加报警规则。

   具体操作步骤，请参见创建报警组和配置报警规则。

2. 选择**指标报警模块 > 联系人管理**，添加指标报警通知人或联系人组。

   具体操作步骤，请参见新增联系人和新增联系人组。

3. 查看指标报警通知记录和报警事件。
   - 在**概览页**页面，查看所有报警组的通知记录和报警事件，详细信息请参见查看报警通知记录和事件。
   - 在**报警组列表**页面，查看单个报警组的通知记录和报警事件，详细信息请参见查看通知记录和查看报警事件。

常见问题

Q：同一时段内监控同一实例，为什么高级监控和Kibana监控的数据不一致？

A：阿里云Elasticsearch的高级监控是内部自研监控，在使用时会和其他监控服务的数据存在差异，具体如下：

- 采样周期差异性：采集周期和Kibana或第三方监控存在差异，采集到的数据不同，因此会存在差异。
- 查询算法差异性：例如，高级监控和Kibana监控采集数据时都会受集群稳定性的影响，高级监控QPS指标会因集群的抖动会出现监控突增、负值或无监控等状况，而Kibana监控可能显示为空。

  > **说明** 如果高级监控提供的指标比Kibana监控多，在实际使用时，建议将高级监控和Kibana监控结合起来分析集群监控详情。

- 采集接口差异性：Kibana监控指标依赖于Elasticsearch API，而高级监控部分节点级别的指标（例如CPU使用率、load_1m、磁盘使用率等），调用的是阿里云Elasticsearch底层系统接口，因此监控中除了Elasticsearch进程外还包含了系统级别资源的占用情况。

3.事件中心

通过阿里云Elasticsearch的事件中心功能，您可以查看对应的系统运维事件，并通过手动运维机制完成事件的追溯与处理。本文为您介绍如何查看事件并进行对应操作。

前提条件

已在支持事件中心功能的地域下创建阿里云Elasticsearch实例：

- 支持事件中心功能的地域包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、中国香港。
- 创建实例的具体操作，请参见创建阿里云Elasticsearch实例。

使用限制

事件中心功能存在地域限制，支持的地域仅包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、香港。

注意事项

为保障云服务的可持续性，当探测到集群资源存在异常或风险，系统会自动触发硬件运维事件，从而最大程度减少对集群的影响，运维事件执行期间可能会造成集群短时间的抖动，但正常的集群访问不会受到影响。当系统无法自动执行或自动执行失败后，您可以在事件中心页面手动触发节点重启操作，人工可干预的窗口期为48小时。

查看事件

1. 登录阿里云Elasticsearch控制台。
2. 进入**事件中心**页面。

   您可以通过以下两种方式进入**事件中心**页面：

   - 在概览页面的**事件中心**区域，单击**查看详情**。
   - 在左侧导航栏，单击**高级监控报警**。再在**高级监控报警**页面的左侧导航栏，单击**事件中心**。
3. 选择地域，查看对应地域下的事件。

   您可以按照实例ID或节点IP查找事件，也可以按照**事件创建时间**、**系统执行时间**或**系统完成时间**筛选事件。

通过事件中心，您可以查看事件的相关信息或根据事件状态进行相应操作，具体说明如下。

事件信息	说明
实例ID/实例名称	触发事件的目标实例ID和名称。单击实例ID，可进入实例管理页面查看实例的详细信息。
事件等级	事件的严重程度，包含：**严重**、**警告**。
节点IP	触发事件的目标节点的IP地址。

事件信息	说明
事件状态	事件的执行状态，包含：**待执行**、**执行中**、**已完成**、**执行失败**、**已取消**。
事件类型	事件的类型，包含：**因探测节点失联触发的节点重启**、**因底层资源运维触发的节点重启**。
事件创建时间	系统探测到事件的时间。
系统执行时间	系统自动运维动作的开始时间。
系统完成时间	系统自动运维动作的结束时间，不受事件状态（成功/失败）影响。
操作	当事件状态为**执行失败**时，在系统完成时间后的48小时窗口期内，您可以在**操作**列下单击**重启节点**，手动重启对应节点。 **注意** 重启操作会触发底层资源重启，为了您的集群稳定性，请不要在集群变更期间重启，并在重启后30分钟内避免对集群进行其他变更。若重启未能生效，系统会在下一次探测到异常后为您生成新事件。

4.可视化监控

4.1. 指标监控

4.1.1. 查看监控指标

高级监控报警服务的指标监控功能提供基础指标和引擎指标等指标监控能力，方便您实时获取Elasticsearch集群侧和引擎侧指标数据，帮助您快速了解集群状况，更好地排查Elasticsearch集群引擎性能及稳定性问题。本文主要介绍如何通过指标监控功能获取监控数据。

前提条件

- 已在支持高级监控报警服务的地域下创建阿里云Elasticsearch实例：
  - 目前高级监控报警服务支持的地域包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、中国香港。
  - 创建实例的具体操作，请参见创建阿里云Elasticsearch实例。
- 熟悉Grafana监控大屏的使用方法。详细信息，请参见Grafana Dashboard。

使用限制

- 高级监控报警功能提供了基础指标、引擎指标和日志数据的监控和报警。阿里云Elasticsearch所有版本都支持对实例的基础指标和日志数据监控，仅内核版本大于1.2.0的6.7.0或7.10.0版本支持引擎指标监控。如果内核版本低于1.2.0，可升级内核版本。具体操作，请参见升级版本。
- 高级监控报警服务存在地域限制，支持的地域仅包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、香港。

操作步骤

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。

   > **说明** 高级监控报警服务默认展示您账号下所有区域接入的实例数据，与您在控制台选择的可用区无关。例如，您在控制台选择北京区域，进入高级监控报警页面后，仍可以看到杭州区域接入的实例数据。

3. 在**高级监控报警**页面，选择**监控可视化 > 指标监控**，即可看到所有接入实例的指标监控数据。

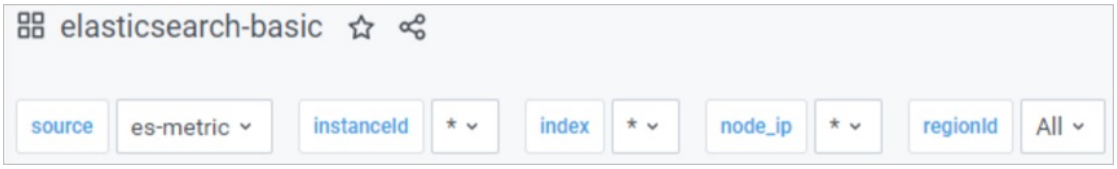
   高级监控报警服务在指标维度提供基础指标和引擎指标监控，两者主要区别在于支持的监控对象不同，详情请参见下表。

指标维度	说明
基础指标	偏向粗粒度的资源监控，帮助您一站式获取集群整体资源状况，支持cluster、index、index Resource、Node Network、Node Disk、Node JVM和Thread_pool相关监控项。
引擎指标	偏向细粒度的资源监控，帮助您快速获取多维度数据处理情况，支持search、bulk（shard）、时序写入Serverless、cache、refresh、merge、cluster state、segment replication和isolator相关监控项。

> 说明
> - 高级监控报警服务中的Grafana监控大盘，使用方式与开源Grafana一致。更多信息，请参见Grafana documentation。
> - 高级监控报警服务提供的所有默认监控大盘，均不支持任何修改。如需修改，您可通过配置自定义监控大屏定制更贴合业务需求的监控大盘。
> - 如果您需要获取更详细的指标监控项说明，请参见基础指标和引擎指标。

4. 查看指定实例、节点或索引等的监控数据。
   i. 鼠标停留在监控窗口，按键盘Esc键，将跳出Grafana菜单页及过滤栏。
   ii. 在过滤栏中，根据需求输入指定的相关信息，即可查看所需的监控数据。参数详细说明，请参见下表。

模块	标签	说明
过滤栏	source	指标监控数据源，默认值为*es-metric*。 **说明** 下拉列表中的default源数据和es-metric源数据一致。
	instanceId	通过实例ID过滤监控数据，默认*，表示无实例限制，将展示所有实例监控。
	IP	通过集群节点IP过滤监控数据，默认*，表示无IP限制，将展示实例下所有节点。
	index	通过索引名过滤监控数据，默认*，表示无索引限制，将显示所有索引监控。
	shardId	通过shardId过滤监控数据，默认*，表示无shardId限制，将显示所有shard监控。
	regionId	通过区域过滤监控数据，默认*ALL*，表示无地域限制，将显示所有区域下开启高级监控报警功能的实例监控。



4.1.2. 基础指标

高级监控报警服务能够为您提供丰富的Elasticsearch指标，其中基础指标不仅包含集群状态、节点及索引数量等资源使用指标和集群或节点的写入与读取QPS等并发性能指标，还包括资源使用情况和网络监控指标等，能够帮助您更好地掌握Elasticsearch集群的使用情况。通过使用高级监控报警服务，您不仅可以查看集群基础指标大盘，还可以自定义相关报警规则，实时监控集群性能并发送报警通知。本文为您介绍默认基础指标大屏中各监控项中的指标含义。

阿里云Elasticsearch实例的版本不同，支持的高级监控指标也不同，具体如下：

- 通用商业版6.7或7.10：支持index写入和查询QPS相关指标，不支持耗时相关指标。
- 通用商业版非6.7和7.10：不支持index写入和查询QPS，以及耗时相关指标。
- 日志增强版6.7：不支持index写入和查询QPS，以及磁盘使用率相关指标。
- 日志增强版7.10：支持磁盘使用率相关指标，不支持index写入和查询QPS相关指标。

基础指标及含义

> **说明**
>
> - cluster、index、Node JVM、Thread_pool维度涉及到的指标均由Elasticsearch模块自身提供，具体请参见Elasticsearch Fields。
> - 在监控集群级别的QPS相关指标时，可能因集群抖动出现不稳定的情况，推荐参考Kibana监控相关指标。高级监控和Kibana监控都会受集群稳定性影响，只是高级监控QPS指标因集群抖动出现的是监控突增、负值或无监控等状况，而Kibana更多的是出现无监控的状况。

类别	指标	含义
cluster	aliyunes.elasticsearch.index.summary.total.indexing.index_total_qps	集群总体写入QPS。
	aliyunes.elasticsearch.index.summary.total.search.query_total_qps	集群总体读取QPS。
	aliyunes.elasticsearch.cluster.stats.status	集群状态，支持以下三种状态： • 0：green • 1：yellow • 2：red
	aliyunes.elasticsearch.cluster.stats.indices.shards.count	shard数目。
	aliyunes.elasticsearch.cluster.stats.indices.total	index数目。
	aliyunes.elasticsearch.cluster.stats.nodes.count	节点数目。
	aliyunes.elasticsearch.aliyun_auto_snapshot.latest_duration.ms	最新快照持续时长，单位：ms。
	aliyunes.elasticsearch.cluster.stats.indices.fielddata.memory.bytes	fielddata内存使用情况，单位：Byte。
	aliyunes.elasticsearch.cluster.stats.indices.shards.primaries	主shard数目。
	aliyunes.elasticsearch.index.indexing.index_total	index写入QPS。

类别	指标	含义
index	aliyunes.elasticsearch.index.search.query_total	index查询QPS。
	aliyunes.elasticsearch.index.indexing.index_time.ms	index耗时，单位：ms。
	aliyunes.elasticsearch.index.search.query_time.ms	查询耗时，单位：ms。
	aliyunes.elasticsearch.index.segments.memory.bytes	index segments内存使用情况，单位：Byte。
	aliyunes.elasticsearch.index.store.size.bytes	索引存储大小，单位：Byte。
	aliyunes.elasticsearch.index.segments.stored_fields_memory.bytes	segments stored fields的内存大小，单位：Byte。
	aliyunes.elasticsearch.index.segments.count	index segments数目。
Node Resource	aliyunes.ecs.node_stats_process_cpu_percent_raw	节点的CPU平均使用率。
	aliyunes.ecs.node_stats_os_cpu_load_average_1m_raw	节点每分钟负载。
	aliyunes.ecs.node_stats_os_per_cpu_load_average_1m_raw	节点单CPU每分钟负载。
	aliyunes.elasticsearch.node.stats.jvm.mem.heap_used_percent	JVM堆内存使用率。
	aliyunes.ecs.node_stats_system_disk_space_usage	系统磁盘使用率。
	aliyunes.ecs.node_stats_fs_data_disk_total_usage	节点磁盘使用率。
Node Network	aliyunes.ecs.node_stats_networkin_packages	节点网络流入包。
	aliyunes.ecs.node_stats_networkout_packages	节点网络流出包。
	aliyunes.ecs.node_stats_networkin_rate	数据流入率。
	aliyunes.ecs.node_stats_networkout_rate	节点网络流出率。

类别	指标	含义
	aliyunes.ecs.node_stats_tcp_established	节点TCP链接数。
Node Disk	aliyunes.ecs.node_stats_data_disk_r	每秒完成的读请求数量。
	aliyunes.ecs.node_stats_data_disk_rm	每秒钟读取的大小，单位：MB。
	aliyunes.ecs.node_stats_data_disk_w	每秒完成的写请求数量。
	aliyunes.ecs.node_stats_data_disk_wm	每秒钟写入的大小，单位：MB。
	aliyunes.ecs.node_stats_data_disk_r_await	平均每次读请求的等待时间，单位：ms。
	aliyunes.ecs.node_stats_data_disk_w_await	平均每次写请求的等待时间，单位：ms。
	aliyunes.ecs.node_stats_data_disk_svctm	平均每次请求的服务时间，单位：ms。
	aliyunes.ecs.node_stats_data_disk_util	设备的利用率。
	aliyunes.ecs.node_stats_data_disk_avgqu_sz	平均请求队列的长度。
Node JVM	aliyunes.elasticsearch.node.stats.jvm.mem.heap_used_percent	heap使用率。
	aliyunes.elasticsearch.node.stats.jvm.mem.pools.old.used.bytes	old区使用情况，单位：Byte。
	aliyunes.elasticsearch.node.stats.jvm.gc.collectors.old.collection.ms	old GC耗时，单位：ms。
	aliyunes.elasticsearch.node.stats.jvm.gc.collectors.young.collection.ms	young GC耗时，单位：ms。
	aliyunes.elasticsearch.node.stats.jvm.gc.collectors.old.collection.count	old GC频次。
	aliyunes.elasticsearch.node.stats.jvm.gc.collectors.young.collection.count	young GC频次。

类别	指标	含义
	aliyunes.elasticsearch.node.stats.jvm.mem.pools.survivor.used.bytes	survivor空间当前使用的内存量，单位：Byte。
	aliyunes.elasticsearch.node.stats.jvm.mem.pools.survivor.max.bytes	survivor空间使用的最大内存量，单位：Byte。
	aliyunes.elasticsearch.node.stats.jvm.mem.pools.old.peak.bytes	JVM老年代空间使用的最大内存，单位：Byte。
	aliyunes.elasticsearch.node.jvm.memory.nonheap.init.bytes	JVM初始化堆外内存，单位：Byte。
	aliyunes.elasticsearch.node.jvm.memory.nonheap.max.bytes	堆外内存最大使用量，单位：Byte。
Thread_pool	aliyunes.elasticsearch.node.stats.thread_pool.search.threads	线程池中的线程总数。
	aliyunes.elasticsearch.node.stats.thread_pool.search.rejected	查询线程池中被拒绝的请求数。
	aliyunes.elasticsearch.node.stats.thread_pool.search.queue	查询线程池中排队的请求数。
	aliyunes.elasticsearch.node.stats.thread_pool.generic.queue	通用线程池中排队的请求数。
	aliyunes.elasticsearch.node.stats.thread_pool.generic.threads	通用池中的线程总数。
	aliyunes.elasticsearch.node.stats.thread_pool.generic.rejected	通用线程池中被拒绝的请求数。

4.1.3. 引擎指标

高级监控报警服务能够为您提供丰富的Elasticsearch指标，其中引擎指标是基于阿里云工程师丰富的运维经验，自主研发和采集的包括集群状态、查询、写入和缓存等方面的各项指标，能够帮助您排查Elasticsearch集群引擎性能及稳定性问题。使用高级监控报警服务，您不仅可以查看集群引擎指标大盘，还可以自定义相关报警规则，实时监控集群性能并发送报警通知。本文为您介绍阿里云Elasticsearch各引擎指标的含义。

概览

高级监控报警服务为您提供以下类别的引擎指标：

- overview（概况）
- search（查询）
- bulk（写入）
- 时序写入Serverless
- cache（缓存）

- refresh（可见性）
- merge（合并）
- cluster state（集群元数据）
- segment replication（物理复制）
- isolator（隔离池）

标签（表头）说明

- 指标：用于展示高级监控报警可供配置的各引擎指标。配置报警规则时需要填写指标，您可以复制此指标并粘贴到搜索框内，系统会自动为您匹配对应指标，详细信息请参见配置报警规则。
- 指标含义：控制台中显示的指标含义。
- 说明：指标的详细说明。
- Tags：配置报警规则时，各监控项支持包含哪些属性标签。

> 注意
> - 不同的指标支持不同粒度的Tags。通过配置Tags，您可以进一步过滤指标数据。
> - 以下Tags在通用Tags属性（instanceId、ip）的基础上，进行了更细粒度的划分。未提到的Tags不在Elasticsearch的监控范围内，例如hostname、kmon_tenant_name、kmon_service_name。

- 聚合算子：
  - 指标聚合：所选Tags内的指标值采用的聚合方式。
  - 采样聚合：对采样周期内的数据采用的聚合方式。

overview（概况）

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.search_total	端到端查询QPS	每秒端到端查询次数。 例如客户端每秒发送两个查询index的请求，则search_total为2。	• instanceId • es_region	• 指标聚合：sum() • 采样聚合：avg()
elasticsearch-server.search_time_in_millis.max	端到端查询延迟max	端到端查询延迟时间。	• instanceId • es_region	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.bulk_total_operations	bulk请求tps	shard维度，每秒bulk操作的次数。	• instanceId • es_region	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.bulk_total_time_in_millis.max	bulk请求延迟max	shard维度，bulk操作总耗时。	• instanceId • es_region	• 指标聚合：max() • 采样聚合：avg()

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.search_aggregation_total	端到端agg查询QPS	每秒端到端聚合查询的次数。 例如客户端每秒发送两个聚合查询请求，则aggregation_total为2。	• instanceId • es_region	• 指标聚合：sum() • 采样聚合：avg()

search（查询）

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.search_total	索引端到端查询QPS	索引间每秒端到端查询次数。 例如客户端每秒发送两个查询index的请求，则search_total为2。	• instanceId • index • es_region	• 指标聚合：sum() • 采样聚合：avg()
elasticsearch-server.search_time_in_millis.max	索引端到端查询延迟_max	索引间端到端查询延迟时间。	• instanceId • index • es_region	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.search_aggregation_total	索引端到端agg查询QPS	索引间每秒端到端聚合查询的次数。 例如客户端每秒发送两个聚合查询请求，则aggregation_total为2。	• instanceId • index • es_region	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.search_total	协调节点查询QPS	协调节点每秒查询次数。 例如客户端每秒发送两个查询index的请求，则通过协调节点search_total为2。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.search_time_in_millis.max	协调节点查询延迟_max	协调节点查询延迟时间。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.search_aggregation_total	协调节点agg查询QPS	协调节点每秒聚合查询的次数。 例如客户端每秒发送两个聚合查询请求，则通过协调节点aggregation_total为2。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.allocated_bytes.max	node聚合查询大对象分配速度_max	聚合查询分配的内存大小。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.query_total	node维度query阶段QPS	node维度整合整个节点上所有shard每秒执行查询的次数，主要与每个shard的个数有关。 例如，每个shard上，您需要查询的索引有5个主shard，则每秒执行shard查询的次数为5。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_took_in_millis.max	node维度query阶段延迟max	nodes维度shard查询阶段的延迟。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.fetch_total	node维度fetch阶段QPS	node维度shard召回阶段每秒的查询次数。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.fetch_took_in_millis.max	node维度fetch阶段延迟max	node维度shard召回阶段总耗时。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_total	shard维度query阶段QPS	shard维度每秒执行shard查询的次数，主要与shard个数有关。 例如，您需要查询的索引有5个主shard，则每秒执行shard查询的次数为5。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_took_in_millis.max	shard维度query阶段延迟max	shard维度shard查询阶段的延迟时间。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.fetch_total	shard维度fetch阶段QPS	shard维度shard召回阶段每秒的查询次数。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.fetch_took_in_millis.max	shard维度fetch阶段延迟max	shard维度shard召回阶段总耗时。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()

bulk（写入）

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.bulk_total_operations	索引维度bulk请求tps	索引维度，每秒bulk操作的次数。	• instanceId • index	• 指标聚合：sum() • 采样聚合：avg()
elasticsearch-server.bulk_total_time_in_millis.max	索引维度bulk请求延迟max	索引维度，bulk操作总耗时。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.bulk_avg_size_in_bytes	索引维度单条bulk平均大小	索引维度，单条bulk命令包含的请求平均大小。	• instanceId • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.bulk_total_operations	node维度bulk请求tps	node维度，每秒bulk操作的次数。	• instanceId • ip	• 指标聚合：sum() • 采样聚合：avg()
elasticsearch-server.bulk_total_time_in_millis.max	node维度bulk请求延迟max	node维度，bulk操作总耗时。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.bulk_avg_size_in_bytes	node维度单条bulk平均大小	node维度，单条bulk命令包含的请求平均大小。	• instanceId • ip	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.bulk_total_operations	shard维度bulk请求tps	shard维度，每秒bulk操作的次数。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.bulk_total_time_in_millis.max	shard维度bulk请求延迟max	shard维度，bulk操作总耗时。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.bulk_avg_size_in_bytes	shard维度单条bulk平均大小	shard维度，单条bulk命令包含的请求平均大小。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()

时序写入Serverless

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.cube.follower_indices_throughput_in_bytes	时序写入Serverless流量	通过时序写入到Elasticsearch时，写入索引的流量大小。	• indexName • es_region	• 指标聚合：sum() • 采样聚合：avg()
elasticsearch-server.cube.follower_indices_store_size_in_bytes	时序写入Serverless索引空间大小	通过时序写入到Elasticsearch上，写入索引所占空间内存大小。	• indexName • es_region	• 指标聚合：sum() • 采样聚合：avg()

cache（缓存）

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.query_cache_shard_hit_total	索引维度query_cache命中QPS	从索引维度观察shard查询时，每秒命中节点缓存的查询次数。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_cached_total	索引维度query_cache缓存QPS	从索引维度观察shard查询时，每秒在节点缓存中新增的查询次数。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_miss_total	索引维度query_cache miss QPS	从索引维度观察shard查询时，每秒未命中节点缓存的查询次数。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_evictions_total	索引维度query_cache踢出QPS	从索引维度观察shard查询时，每秒从节点缓存中踢出的查询次数。 例如，当缓存已满时，将最近使用最少的查询结果踢出，以留出空间来存放新数据。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_hit_total	node维度query_cache命中QPS	从node维度观察shard查询时，每秒命中节点缓存的查询次数。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_cached_total	node维度query_cache缓存QPS	从node维度观察shard查询时，每秒在节点缓存中新增的查询次数。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.query_cache_shard_miss_total	node维度query_cache miss QPS	从node维度观察shard查询时，每秒未命中节点缓存的查询次数。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_evictions_total	node维度query_cache踢出QPS	从node维度观察shard查询时，每秒从节点缓存中踢出的查询次数。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_cached_size_in_bytes.max	node维度query_cache缓存大小max	shard查询时，从node维度观察节点缓存新增数据的总大小。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_hit_total	shard维度query_cache命中QPS	shard查询时，每秒命中节点缓存的查询次数。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_cached_total	shard维度query_cache缓存QPS	shard查询时，每秒在节点缓存中新增的查询次数。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_miss_total	shard维度query_cache miss QPS	shard查询时，每秒未命中节点缓存的查询次数。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_evictions_total	shard维度query_cache踢出QPS	shard查询时，每秒从节点缓存中踢出的查询次数。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.query_cache_shard_cached_size_in_bytes.max	shard维度query_cache缓存大小max	shard查询时，缓存新增数据的总大小。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()

refresh（可见性）

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.refresh_total	索引维度refresh_qps	刷新动作落在索引上每秒的查询次数。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.refresh_interval_in_millis.max	索引维度refresh间隔max	每次刷新动作落在索引之间的间隔。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.refresh_took_in_millis.max	索引维度refresh动作耗时max	刷新动作落在每条索引所占用时间。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.refresh_total	node维度refresh_qps	刷新动作落在节点上每秒的查询次数。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.refresh_interval_in_millis.max	node维度refresh间隔max	每次刷新动作落在节点之间的间隔。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.refresh_took_in_millis.max	node维度refresh动作耗时max	刷新动作落在每个节点所占用时间。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.refresh_total	shard维度refresh_qps	刷新动作落在索引shard上每秒的查询次数。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.refresh_interval_in_millis.max	shard维度refresh间隔max	每次刷新动作落在索引shard之间的间隔。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.refresh_took_in_millis.max	shard维度refresh动作耗时max	刷新动作落在每个索引shard所占用时间。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()

merge（合并）

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.merge_total	索引维度merge_qps	刷索引merge阶段每秒的查询次数。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.merge_took_in_millis.max	索引维度merge耗时max	索引merge数据时所用时间。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.merge_size_in_bytes.max	索引维度merge大小max	索引merge数据后占用的内存大小。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.merge_total	node维度merge_qps	各节点merge阶段每秒的查询次数。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.merge_took_in_millis.max	node维度merge耗时max	各节点merge数据时所用时间。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.merge_size_in_bytes.max	node维度merge大小max	各节点merge数据后占用的内存大小。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.merge_total	shard维度merge_qps	索引shard在merge阶段每秒的查询次数。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.merge_took_in_millis.max	shard维度merge耗时max	索引shard在merge数据时所用时间。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.merge_size_in_bytes.max	shard维度merge大小max	索引shard在merge数据后占用的内存大小。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()

cluster state（集群元数据）

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.applied_cluster_state_count	cluster_state本地应用QPS	Master节点同步集群状态给其他节点，其他节点接收成功的次数。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.applied_cluster_state_took_in_millis.max	cluster_state本地应用耗时max	Master节点同步集群状态给其他节点，其他节点接收成功所消耗的时间。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.publish_time_in_millis.max	cluster state广播耗时max	集群状态广播耗时。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.failed_cluster_state_count	cluster_state本地应用失败QPS	Master节点同步集群状态给其他节点，其他节点接收失败的次数。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.failed_cluster_state_took_in_millis.max	cluster_state本地应用失败耗时max	Master节点同步集群状态给其他节点，其他节点接收失败所消耗的时间。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.task_execution_count	master处理task QPS	集群状态变化次数。 例如当集群中存在频繁的节点变更、频繁的设置索引Mapping和Setting等操作时，Master节点都会重新计算集群状态变化次数。 注意 状态变化次数越大，说明集群或索引存在频繁的变更，可能会影响集群的稳定性。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.task_execution_time_in_millis.max	master处理task耗时max	Master节点获取集群状态所消耗的时间。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.task_wait_time_in_millis.max	task队列等待时间max	每个获取集群状态的任务在Master节点的任务队列中的等待时间。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()

segment replication（物理复制）

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.segment_replication.refresh_copy_file_size	索引维度增量拷贝平均大小	使用物理复制功能时，每次索引维度执行refresh操作，拷贝主副本增量数据的大小。	• instanceId • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.refresh_copy_file_size	node维度增量拷贝平均大小	使用物理复制功能时，每次node维度执行refresh操作，拷贝主副本增量数据的大小。	• instanceId • ip	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.refresh_copy_file_size	shard维度增量拷贝大小	使用物理复制功能时，每次shard维度执行refresh操作，拷贝主副本增量数据的大小。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.refresh_latency_time	索引维度增量拷贝延迟-avg	使用物理复制功能时，每次索引维度执行refresh操作所消耗的平均值时间。	• instanceId • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.refresh_latency_time	索引维度增量拷贝延迟-max	使用物理复制功能时，每次索引维度执行refresh操作所消耗的最大值时间。	• instanceId • index	• 指标聚合：max() • 采样聚合：max()
elasticsearch-server.segment_replication.refresh_latency_time	node维度增量拷贝延迟-avg	使用物理复制功能时，每次node维度执行refresh操作所消耗的平均值时间。	• instanceId • ip	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.refresh_latency_time	shard维度增量拷贝延迟-avg	使用物理复制功能时，每次执行refresh操作所消耗的平均值时间。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.refresh_latency_time	shard维度增量拷贝延迟-max	使用物理复制功能时，每次shard维度执行refresh操作所消耗的最大值时间。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：max()

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.segment_replication.merge_copy_file_size	索引维度merge预拷贝平均大小	使用物理复制功能时，每次索引维度merge阶段结束后，拷贝到副本的平均值数据大小。	• instanceId • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.merge_copy_file_size	node维度merge预拷贝平均大小	使用物理复制功能时，每次node维度merge阶段结束后，拷贝到副本的平均值数据大小。	• instanceId • ip	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.merge_copy_file_size	shard维度merge预拷贝大小	使用物理复制功能时，每次shard维度merge阶段结束后，拷贝到副本的平均值数据大小。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.merge_latencc	索引维度merge预拷贝延迟-avg	使用物理复制功能时，每次索引维度merge阶段结束后，数据拷贝到副本所消耗的平均值时间。	• instanceId • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.merge_latencc	索引维度merge预拷贝延迟-max	使用物理复制功能时，每次索引维度merge阶段结束后，数据拷贝到副本所消耗的最大值时间。	• instanceId • index	• 指标聚合：max() • 采样聚合：max()
elasticsearch-server.segment_replication.merge_latencc	node维度merge预拷贝延迟-avg	使用物理复制功能时，每次node维度merge阶段结束后，数据拷贝到副本所消耗的平均值时间。	• instanceId • ip	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.merge_latencc	shard维度merge预拷贝延迟-avg	使用物理复制功能时，每次shard维度merge阶段结束后，数据拷贝到副本所消耗的平均值时间。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.merge_latencc	shard维度merge预拷贝延迟-max	使用物理复制功能时，每次shard维度merge阶段结束后，数据拷贝到副本所消耗的最大值时间。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：max()

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.segment_replication.replica_checkpoint_gap	索引维度replica和复制位点的gap	使用物理复制功能时，每次索引checkpoint阶段中复制位点的间隙数.	• instanceId • index	• 指标聚合：max() • 采样聚合：max()
elasticsearch-server.segment_replication.replica_checkpoint_gap	node维度replica和复制位点的gap	使用物理复制功能时，每次node checkpoint阶段中复制位点的间隙数。	• instanceId • ip	• 指标聚合：max() • 采样聚合：max()
elasticsearch-server.segment_replication.replica_checkpoint_gap	shard维度replica和复制位点的gap	使用物理复制功能时，每次shard checkpoint阶段中复制位点的间隙数。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：max()
elasticsearch-server.segment_replication.refresh_count	shard维度增量拷贝QPS	使用物理复制功能时，每次执行refresh操作进行增量数据拷贝的QPS。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.merge_error_count	shard维度增量拷贝失败QPS	使用物理复制功能，数据拷贝异常的QPS。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.merge_error_count	shard维度merge预拷贝失败QPS	使用物理复制功能，在merge阶段，数据拷贝异常的QPS。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.merge_count	shard维度merge预拷贝QPS	merge合并次数。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()
elasticsearch-server.segment_replication.checkpoint_gap_count	shard维度gap汇报的QPS	使用物理复制功能，在checkpoint阶段的间隙数QPS。	• instanceId • shard_id • ip • index	• 指标聚合：avg() • 采样聚合：avg()

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.segment_replication.primary_checkpoint_gap.max	shard维度primary和复制位点的gap	使用物理复制功能，在副本的checkpoint阶段间隙数大小。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()

isolator（隔离池）

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.isolator_tasks_isolated_total	索引维度query isolated_total	索引维度每秒慢查询隔离池中索引的查询数量。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.isolator_tasks_killed_total	索引维度query cancel QPS	索引维度每秒慢查询隔离池中索引触发熔断的查询数量。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.isolator_tasks_killed_mem_size_in_bytes.max	索引维度query cancel mem_size_in_bytes max	慢查询隔离池中索引触发熔断的查询消耗内存大小。	• instanceId • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.isolator_tasks_isolated_total	node维度query isolated_total	每秒慢查询隔离池中节点的查询数量。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.isolator_tasks_killed_total	node维度query cancel QPS	每秒慢查询隔离池中节点触发熔断的查询数量。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.isolator_tasks_killed_mem_size_in_bytes.max	node维度query cancel mem_size_in_bytes max	慢查询隔离池中节点触发熔断的查询消耗内存大小。	• instanceId • ip	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.isolator_tasks_isolated_total	shard维度query isolated_total	每秒慢查询隔离池中索引shard的查询数量。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()

指标	指标含义	说明	Tags	聚合算子
elasticsearch-server.isolator_tasks_killed_total	shard维度query cancel QPS	每秒慢查询隔离池中索引shard触发熔断的查询数量。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()
elasticsearch-server.isolator_tasks_killed_mem_size_in_bytes.max	shard维度query cancel mem_size_in_bytes max	慢查询隔离池中索引shard触发熔断的查询消耗内存大小。	• instanceId • shard_id • ip • index	• 指标聚合：max() • 采样聚合：avg()

4.2. 日志监控

高级监控报警的日志监控功能提供了慢查询日志、慢索引日志、访问日志和主日志等日志监控能力，方便您实时获取集群日志情况，从运维角度，能够帮助您快速排查和定位问题。本文主要介绍如何通过日志监控功能获取监控数据以及如何快速过滤日志数据。

前提条件

- 已在支持高级监控报警服务的地域下创建阿里云Elasticsearch实例：
  - 目前高级监控报警服务支持的地域包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、中国香港。
  - 创建实例的具体操作，请参见创建阿里云Elasticsearch实例。
- 熟悉Grafana监控大屏的使用方法。详细信息，请参见Grafana Dashboard。

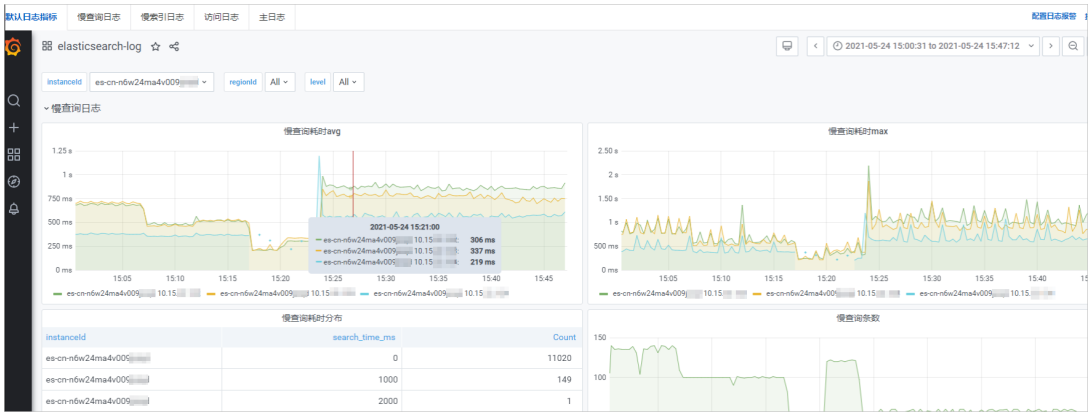
使用限制

- 高级监控报警功能提供了基础指标、引擎指标和日志数据的监控和报警。阿里云Elasticsearch所有版本都支持对实例的基础指标和日志数据监控，仅内核版本大于1.2.0的6.7.0或7.10.0版本支持引擎指标监控。如果内核版本低于1.2.0，可升级内核版本。具体操作，请参见升级版本。
- 高级监控报警服务存在地域限制，支持的地域仅包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、香港。

查看默认日志指标

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。
3. 选择**监控可视化 > 日志监控**，即可看到所有接入实例的日志监控数据。
4. 查看特定实例的监控数据。
   - 方法一：通过过滤栏筛选instanceID查看监控数据
     a. 鼠标停留在监控窗口，按键盘**Esc**键，将跳出Grafana菜单页及过滤栏。

b. 输入instanceID在过滤栏中，选择**instanceId**、**regionId**和**level**，即可查看该实例慢查询日志、慢索引日志、访问日志和主日志等相关监控数据。



- 方法二：从实例列表入口跳转
  a. 在**日志监控**页面左上角，单击**阿里云Elasticsearch**。
  b. 在左侧导航栏，单击**Elasticsearch实例**。
  c. 在**Elasticsearch实例**列表中，单击目标实例ID。
  d. 在左侧导航栏，选择**监控与日志 > 日志查询**。
  e. 单击**高级日志监控**，即可查看当前实例的相关数据。

关键词	标签示例	说明
avg	慢查询耗时avg	数据节点慢查询平均耗时。
max	慢查询耗时max	数据节点慢查询最大耗时。
分布	慢查询耗时分布	秒间隔时间内，慢查询数量分布。例如： - 0 ms ≤ search_time_ms（慢查询耗时）< 1000 ms，此区间分布了11020条慢查询日志。 - 1000 ms≤ search_time_ms（慢查询耗时）< 2000 ms，此区间分布了149条慢查询日志。 - search_time_ms（慢查询耗时）≥2000 ms，此区间分布了1条慢查询日志。
条数	慢查询条数	集群中慢查询日志总条数。

> 说明
>
> - 高级监控报警服务中的Grafana监控大盘，使用方式与开源Grafana一致。更多信息，请参见Grafana documentation。
> - 高级监控报警服务提供的所有默认监控大盘，均不支持任何修改。如需修改，您可通过配置自定义监控大屏定制更贴合业务需求的监控大盘。

查询日志

日志监控支持对慢查询日志、慢索引日志、访问日志和主日志进行过滤查询，不同的日志类型来自不同的数据源。具体信息见下表。

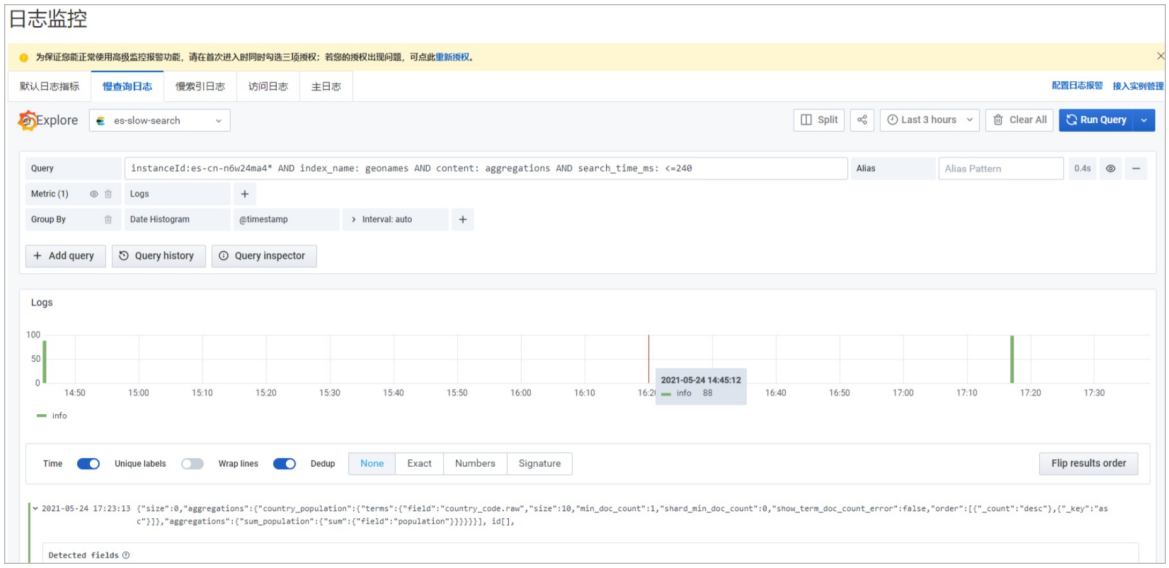
数据源	说明
es-slow-search	提供慢查询日志数据。
es-slow-index	提供慢索引日志数据。
es-access-log	提供访问日志数据（当前仅支持6.7.0和7.10.0版本的实例）。
es-instance-search	提供主日志数据。

由于各类日志的查询流程一致，所以本文以慢查询日志为例，介绍日志查询的操作步骤。

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。
3. 选择**监控可视化 > 日志监控**。
4. 在**日志监控**页面，单击**慢查询日志**页签。
5. 设置**Query**。

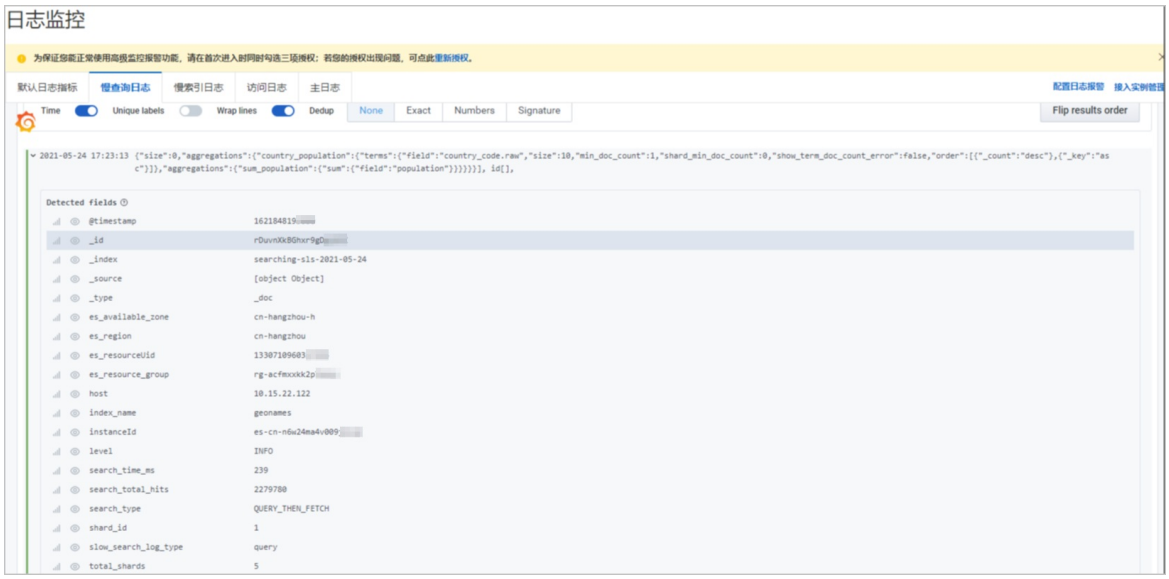
   例如：如果您需要过滤出实例以es-cn-n6w24ma4开头、索引名为geonames、content中包含aggregations并且查询耗时小于等于240ms的慢查询日志，需要设置Query语句为：

   ```
   instanceId:es-cn-n6w24ma4* AND index_name: geonames AND content: aggregations AND search_time_ms: <=240
   ```



> 说明
>
> - Query支持 `>` 、 `<` 、 `=` 、 `>=` 、 `<=` 、 `AND` 和 `OR` 等条件符，具体请参考Query string。
> - 不同的日志属性支持的数据源不一样，不同的数据源支持的query字段存在部分不同。例如上面的慢查询日志页面中 `es-slow-search` 表示慢查询数据源，提供的 `search_time_ms` 表示查询时间。具体参考日志查询内置字段。
> - 日志监控页面中仅Query可用，且仅支持对日志数据的检索，其他检索项修改后无法生效。例如将检索范围的默认*logs*修改为*sum*或其他值，则不会生效。

6. 单击Run Query。
7. 单击其中一条日志，即可在Detected fields下查看日志内容及Query可查询的具体字段。



> 说明
>
> - Detected fields key部分为Query支持的过滤字段，例如es_available_zone表示可用区、node表示节点，具体请参见日志查询内置字段。
> - 不同的日志属性支持的数据源不一样，不同的数据源支持的query字段存在部分不同。例如：仅慢查询数据源提供 `search_time_ms` ，而慢写入中未提供。具体请参见日志查询内置字段。

日志查询内置字段列表

- 独立字段

类型	独立字段	说明
	search_time_ms	查询耗时时长。
	search_total_hits	查询命中的文档数。
	search_type	查询类型。
	shard_id	执行该条查询的shard编号。

慢查询	独立字段	说明
	slow_search_log_type	慢日志类型。
	total_shards	总shard数。
	content	query查询体。
慢写入	index_time_ms	写入耗时时长。
	content	query查询体。
主日志	content	query查询体。
访问日志	node	产生访问日志的Elasticsearch节点。
	query	执行的查询体，过滤时请使用**source**代替query字段做查询。
	remote	远程服务器IP地址。
	bodySize	请求大小。
	uri	访问路径。

说明 独立字段仅说明各个日志类型支持的不同的字段部分，相同字段部分请参考下表通用字段。

- 通用字段

通过字段	说明
es_available_zone	实例可用区。
es_region	实例所在地域。
es_resourceUid	实例uid。
es_resource_group	实例所在资源组。
host	节点ip。
instanceId	实例id。
level	日志级别。

说明 以`_`开头的字段均为Elasticsearch元数据自带的。

4.3. 配置自定义监控大屏

高级监控报警服务支持您根据业务需要自定义监控大屏，帮助您更加灵活地监控Elasticsearch集群，作为默认监控能力的补充，确保满足您在不同场景下的监控需求。本文主要介绍如何配置自定义监控大屏。

背景信息

您可以通过以下两种方式配置自定义监控大屏：

- 方式一：通过修改导入的模板自定义监控大屏
- 方式二：新建自定义大屏

前提条件

- 已在支持高级监控报警服务的地域下创建阿里云Elasticsearch实例：
  - 目前高级监控报警服务支持的地域包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、中国香港。
  - 创建实例的具体操作，请参见创建阿里云Elasticsearch实例。
- 熟悉Grafana监控大屏的使用方法。详细信息，请参见Grafana Dashboard。

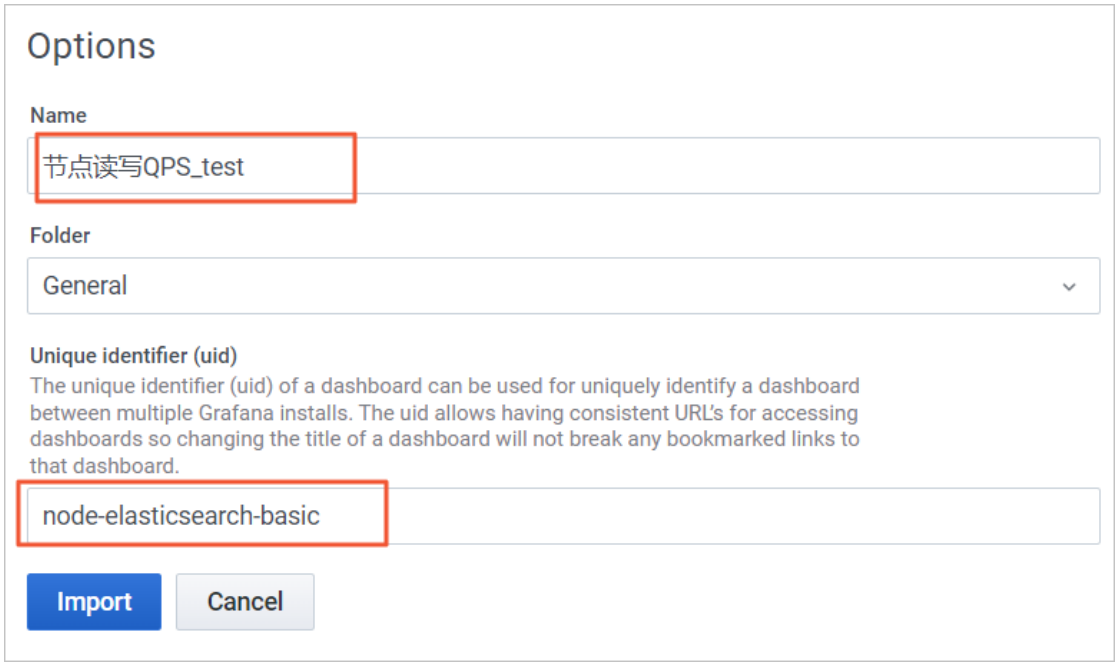
使用限制

- 高级监控报警功能提供了基础指标、引擎指标和日志数据的监控和报警。阿里云Elasticsearch所有版本都支持对实例的基础指标和日志数据监控，仅内核版本大于1.2.0的6.7.0或7.10.0版本支持引擎指标监控。如果内核版本低于1.2.0，可升级内核版本。具体操作，请参见升级版本。
- 高级监控报警服务存在地域限制，支持的地域仅包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、香港。

方式一：通过修改导入的模板自定义监控大屏

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。
3. 在**高级监控报警**页面，复制默认监控规则的JSON模板。
   i. 选择**监控可视化 > 指标监控**。
   ii. 在**默认基础指标**页签，鼠标左键单击监控窗口的任意空白处，然后按下键盘中的**Esc**键。
   操作成功后，当前页面会弹出Grafana菜单页及过滤栏。
   iii. 在Grafana页面，单击**基础指标大盘**右侧的  图标。
   iv. 在对话框中，单击**Export**页签。
   v. 单击**View JSON**。
   vi. 单击**Copy to Clipboard**，复制JSON模板。
4. 导入模板。
   i. 在左侧Grafana菜单栏中，单击  图标，选择**Import**。
   ii. 在**Import via panel json**输入框中，粘贴已复制的JSON模板，单击**Load**。

iii. 修改Name，并重新定义Unique identifier (uid)。



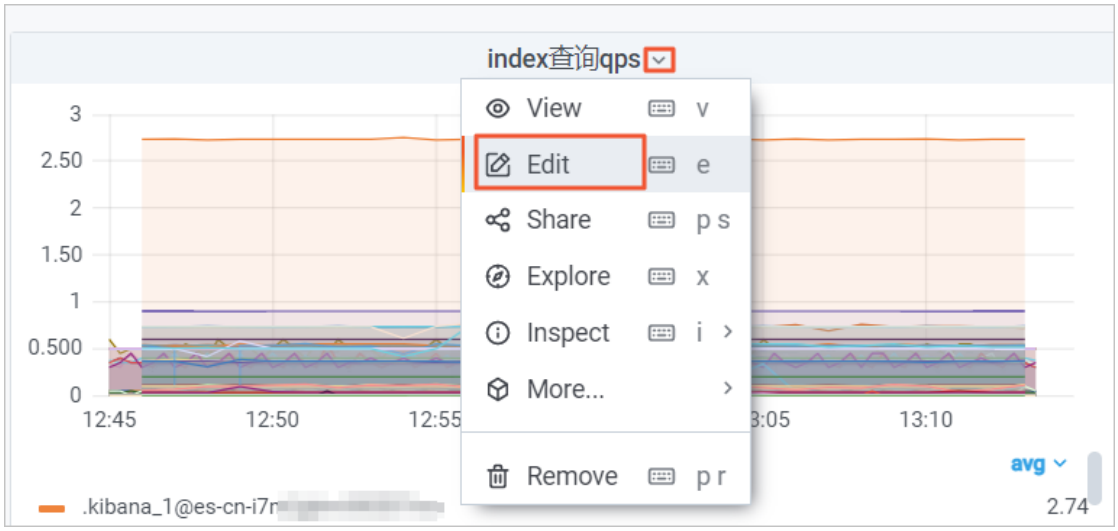
iv. 单击Import，即可完成模板的导入。

5. 在**自定义监控**页面，修改导入的监控模板。

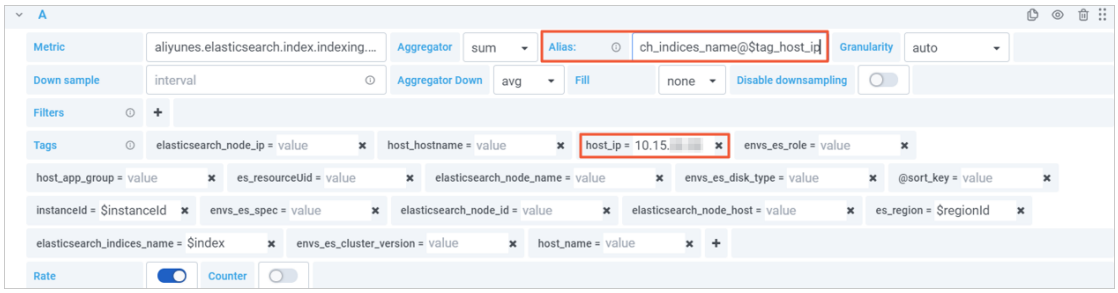
i. 在左侧导航栏，选择**监控可视化 > 自定义监控**。

ii. 在页面上方的自定义监控列表中，单击您自定义的监控模块页签。

iii. 展开Index(索引)模块，将鼠标悬浮至目标监控指标名称上，单击右侧的图标，选择Edit。

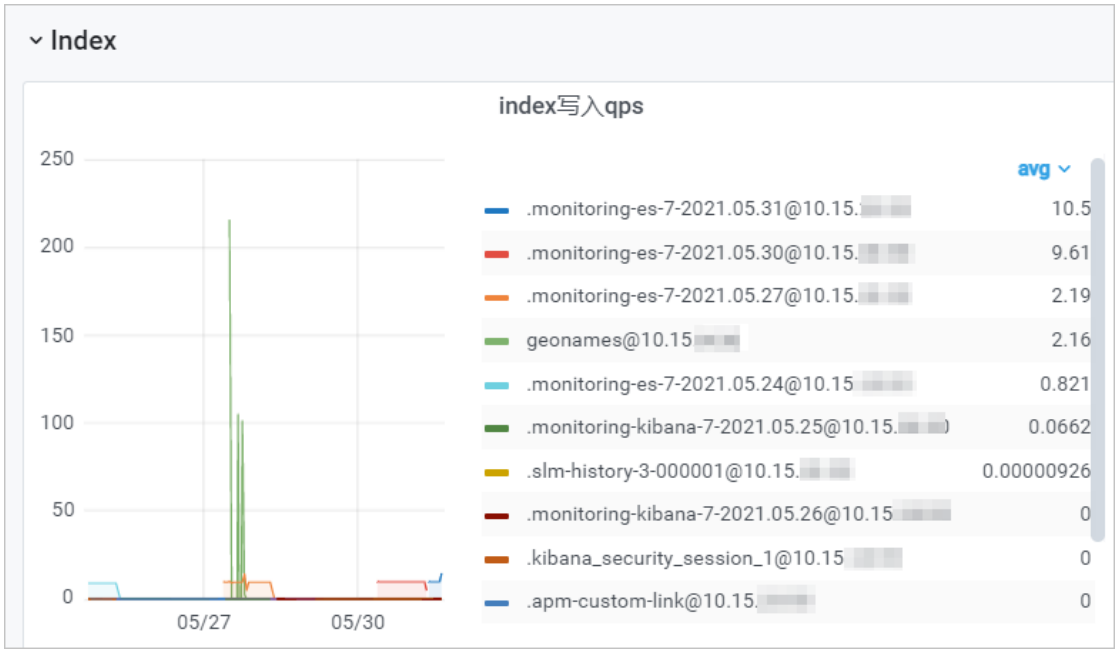


iv. 在host_ip中输入具体的节点IP，在Alias中调用host_ip变量。



v. 单击右上角的Save，按照页面提示保存配置。

vi. 单击Apply应用配置，即可展示单个节点index写入QPS的情况。



方式二：新建自定义大屏

1. 进入自定义监控页面。
   i. 登录阿里云Elasticsearch控制台。
   ii. 在左侧导航栏，单击**高级监控报警**。
   iii. 选择**监控可视化 > 自定义监控**。
2. 在**自定义监控**页面，单击页面上方的**+自定义监控**页签。
3. 在Grafana页面，单击左侧菜单栏的+图标，选择**Create**。
4. 在**New dashboard**页面中，单击**Add new panel**。
5. 参见Grafana documentation，配置监控大屏。

   高级监控报警服务中自定义监控大屏，配置方法与开源Grafana一致。配置时需要设置监控指标，各监控指标名称及详细说明请参见基础指标和引擎指标。

5.指标报警

5.1. 基本概念

本文介绍使用高级监控报警服务配置指标报警时，遇到的常用名词的基本概念和简要描述。

报警规则

报警的触发条件和通知方式。

报警组

一个报警组包含多条报警规则。

报警事件

系统每隔1分钟，就会根据报警规则中设置的报警触发条件，判断指标是否触发报警。如果触发，则会生成一个报警事件记录。

通知记录

报警事件生成之后，系统会根据报警规则中设置的报警生效时段和报警间隔，判断是否发送报警通知（电话、短信、钉钉群机器人）给您。如果发送，则会生成一个通知记录。

指标

表示事物的状态大小。例如disk.io.util表示节点磁盘使用率、load.1min表示1分钟内节点的负载。

tags

指标的属性标签，能够进一步对指标进行过滤，取值是一组键值对。例如指标disk.io.util通常带有属性host=localhost、dev=/ssd/1，表示localhost主机中/ssd/1磁盘的使用率。

指标聚合

如果指标有多条曲线（指标的所有tags取值的组合表示曲线的个数），多条曲线聚合成一条曲线的算法。

采样聚合

指标的单条曲线在检测周期内（默认1分钟），多个数据点聚合成一个点的算法。

阈值报警

当前指标的值和阈值实时比较，如果符合设定的阈值条件，则触发报警。

波动报警

假设当前指标的值为a，一段时间前指标的值为b。对两者计算差值(a-b)，或者变化率(a-b)/b，然后将计算结果和阈值条件进行比较，如果符合条件，则触发报警。

无数据校验

如果系统连续一段时间（默认1分钟）没有检测到任何数据，则触发报警。

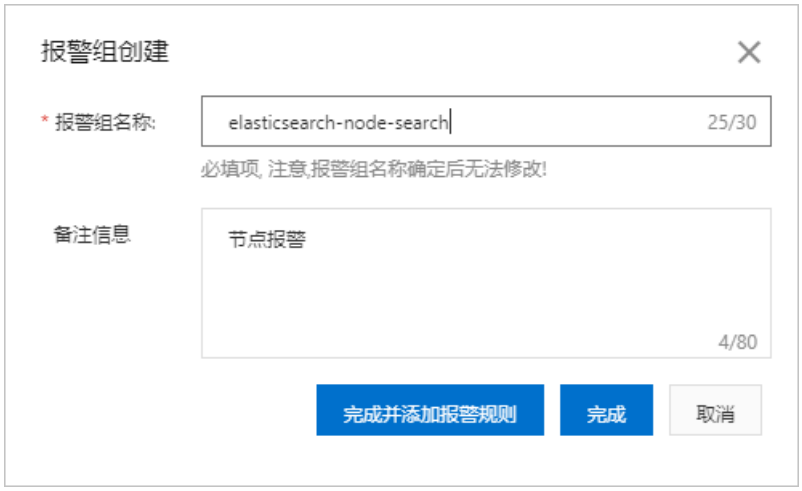
5.2. 报警组和报警规则

5.2.1. 管理报警组

一个报警组可以包含一个或多个报警规则，同一个报警规则可以加入多个报警组。本文为您介绍如何创建报警组、新建和管理报警规则、查看通知记录和报警事件、删除报警组。

创建报警组

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。
3. 在左侧导航栏，选择**指标报警模块 > 报警组列表**。
4. 在**报警组列表**页面，单击**创建报警组**。
5. 在**报警组创建**对话框中，填写报警组信息。



参数	说明
报警组名称	长度为0~30个字符，以大小写字母、数字或中文开头，可以包含下划线（_）或连接符（-）。**报警组名称**不可重复，且确定后无法修改。
备注信息	填写报警规则说明，方便您快速查找定位。

6. 单击**完成并添加报警规则**或**完成**。
   - **完成并添加报警规则**：完成报警组创建后，将直接进入**新建报警规则**页面，添加报警规则。
   - **完成**：仅完成报警组创建，如果需要创建报警规则，还要在**报警组列表**右侧**操作**列下，单击**新建报警规则**，进入**新建报警规则**页面，添加报警规则，详情请参见配置报警规则。

   报警组创建完成后，您还可以完成以下操作。

   - 新建报警规则
   - 查看报警规则

- 查看通知记录
- 查看报警事件
- 删除报警组

新建报警规则

1. 在**报警组列表**页面，单击目标报警组右侧**操作**列下的**新建报警规则**。
2. 在**新建报警规则**页面，输入规则配置，单击**完成创建**。

   规则配置的详细参数说明请参见配置报警规则。

查看报警规则

1. 在**报警组列表**页面，单击目标报警组右侧**操作**列下的**报警规则列表**。
2. 在**报警规则列表**页面，查看您已创建的报警规则。

   > **说明** 如果没有报警规则，可单击**新建规则**创建规则。

   一个报警组中可以包含多个报警规则，报警规则之间无直接影响。您可以在**报警规则列表**中管理报警规则。

部分参数说明如下。

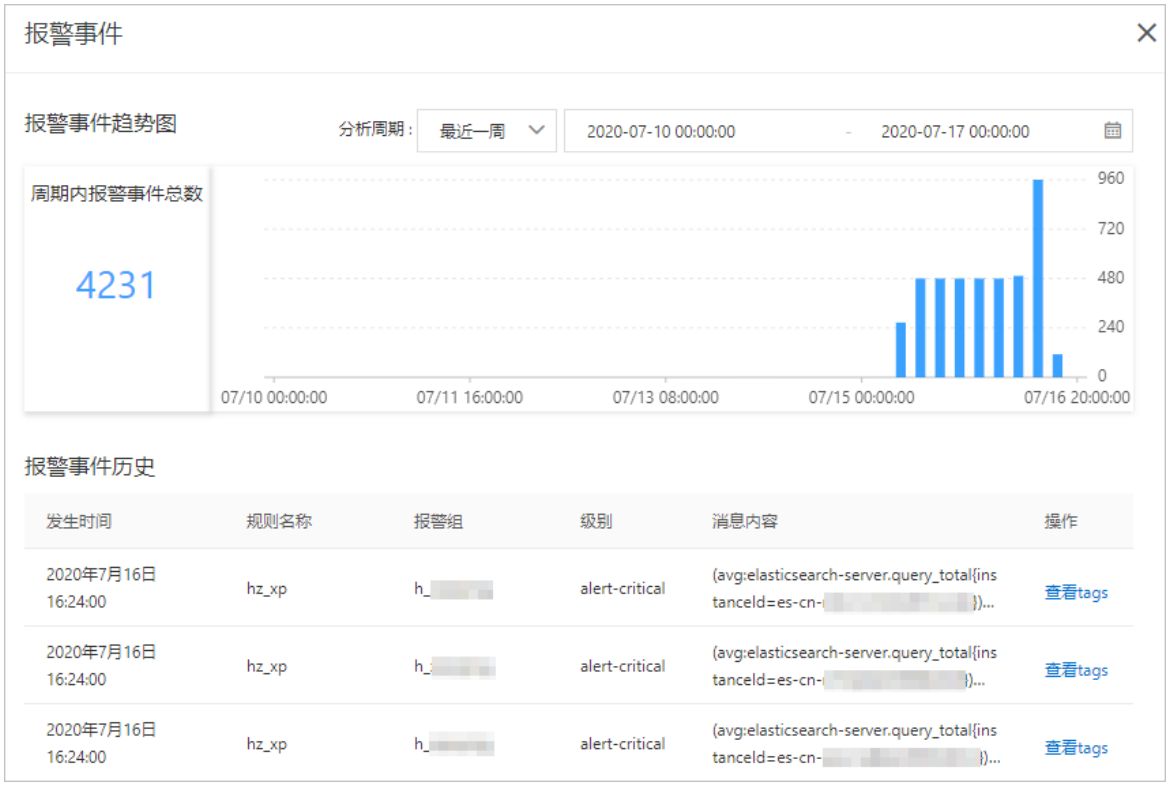
参数	说明
状态	对于新建成功的规则，默认状态为**已开启**，支持关闭状态。
操作	○ **查看/修改**：查看并修改已创建的报警规则。 ○ **复制**：复制源报警规则生成新规则，新规则默认命名为：源规则名_copy。 ○ **删除**：删除已创建的报警规则。 ○ **关闭**：支持暂时关闭及永久关闭报警规则。单击关闭后，系统会立即关闭规则报警。关闭后，**状态**会显示**已关闭**。

查看通知记录

1. 在**报警组列表**页面，单击目标报警组右侧**操作**列下的**查看通知记录**。
2. 在**通知记录**页面，查看所选周期内（默认为最近一周）的通知记录和明细。

查看报警事件

1. 在**报警组列表**页面，单击目标报警组右侧**操作**列下的**查看报警事件**。
2. 在**报警事件**页面，查看所选周期内（默认为最近一周）系统定时检测触发的报警事件。



删除报警组

> **警告** 删除报警组后，该报警组中的报警规则也会被删除，且不可恢复，请谨慎操作。

1. 在**报警组列表**页面，单击目标报警组右侧**操作**列下的 > **删除报警组**。
2. 在弹出框中，单击**确定**。

5.2.2. 配置报警规则

高级监控报警能够为阿里云Elasticsearch实例设置更细粒度的指标报警规则。例如某个分片的QPS达到某个量级，就会触发报警，并且第一时间通知您。通过报警规则配置，您可以设置多维度的监控指标和Tags，帮助您快速定位Elasticsearch的性能问题，提高运维排查效率。本文为您介绍如何配置报警规则，并提供详细的参数说明。

前提条件

已创建阿里云Elasticsearch实例。阿里云Elasticsearch所有版本都支持接入高级监控报警服务，仅内核版本大于1.2.0的6.7.0或7.10.0版本支持引擎指标监控。

- 创建实例的具体操作，具体操作请参见创建阿里云Elasticsearch实例。
- 如果内核版本低于1.2.0，可升级内核版本。具体操作请参见升级版本。

创建报警规则

参见创建报警组或新建报警规则，创建报警规则。报警规则中需要配置的参数如下：

- 规则类型
- 基本信息

- 指标
- tags（可选）
- 触发条件
- 无数据校验（可选）
- 规则触发后动作

规则类型

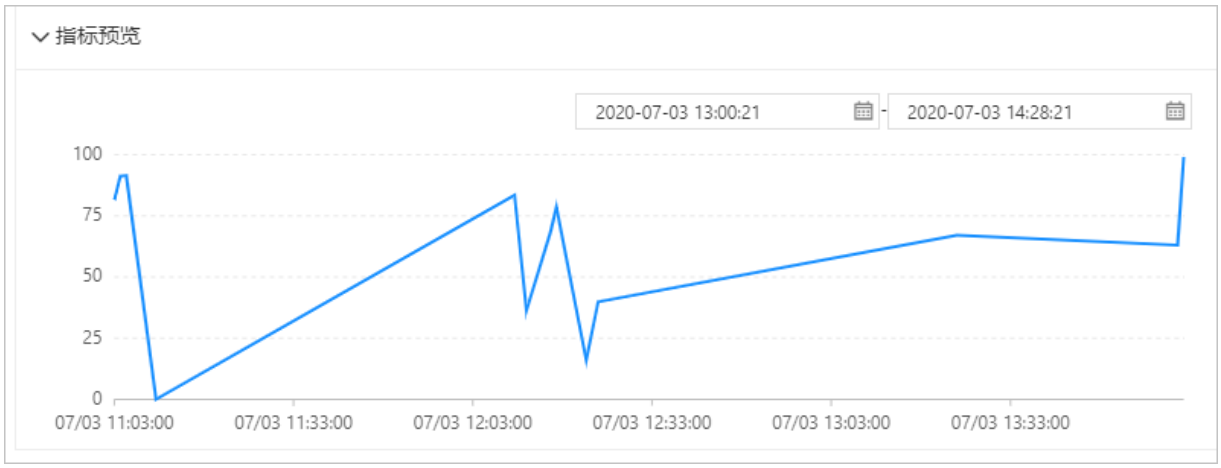
报警规则类型固定为**指标报警**，表示对指定指标（metric）设置报警阈值。

基本信息

参数	说明
规则名称	无限制，支持中英文任何字符。
描述备注	长度为0~100个字符，请填入规则的简单描述，便于快速排查定位。

指标预览

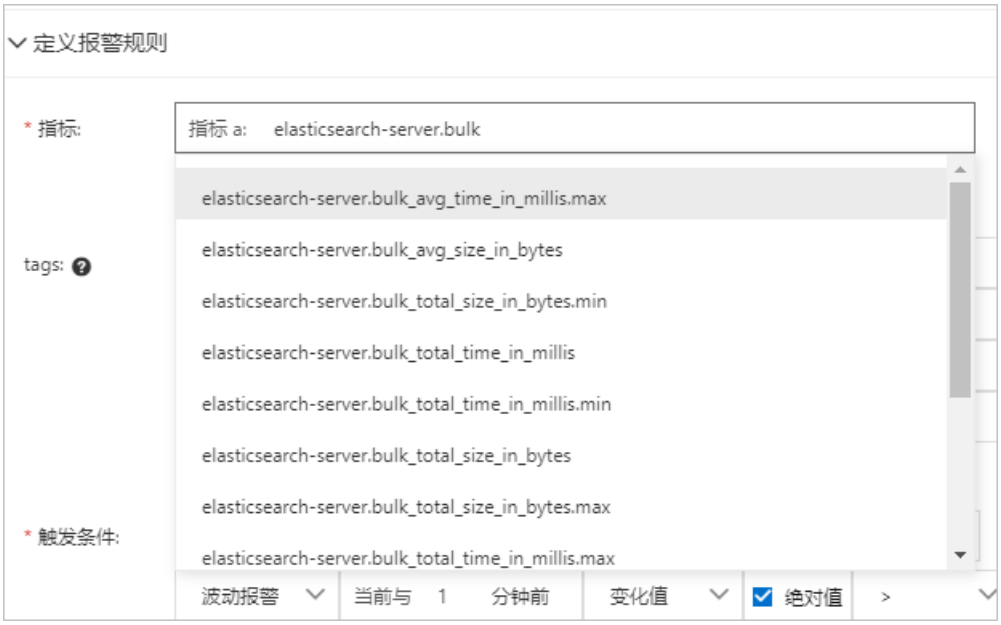
定义了报警指标及tags后，系统会自动生成指标预览图。默认情况下，组成指标预览图中的每个点的间隔是1分钟。



> 说明
>
> - 由于指标预览图默认采样周期为1分钟，而底层默认采样周期为5s，因此系统会通过采样聚合算法，将1分钟内多个数据点聚合成一个点。
> - 由于索引中包含多个shard，而每个shard会产生一条曲线，因此系统会通过指标聚合算法，将多个曲线合成一条曲线，形成索引的监控曲线图。

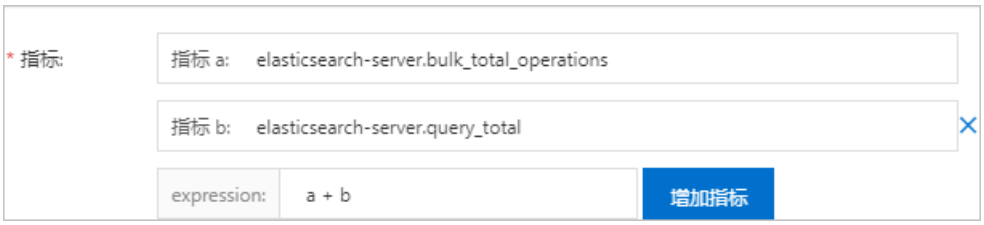
指标

- 单指标



从**指标**列表中，选择报警指标。或在输入框中输入指标前缀，例如输入elasticsearch-server.bulk，系统将匹配以此前缀开头的所有指标供您选择。指标说明请参见引擎指标或者基础指标。

- 多指标



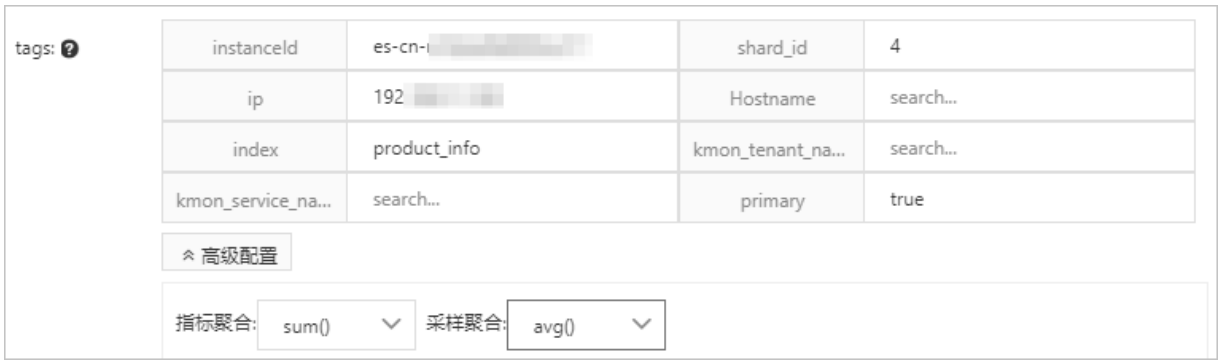
单击**增加指标**，可添加多个指标。添加后，系统会根据多指标运算结果，判断是否触发报警。

- 每个指标都会对应一个标签名，例如上图中的指标a、指标b。
- 必须添加同一类型的指标，例如添加多个QPS监控类指标。
- **expression**：多指标间的计算表达式，运算符支持+、-、*、/、&&、||、>、<，默认为+。例如上图中生成的指标图为：在各个时刻，指标a的值与指标b的值进行求和，其结果随时间变化的曲线图。

  例如**expression**为(a>1200) && (b<1500) && (c<1)，表示系统将绘制这个表达式在各个时刻的计算结果。由于该表达式为布尔表达式，因此这个表达式的指标预览图中曲线的取值是0或1。

tags（可选）

定义指标属性标签，即进一步对指标进行过滤。取值是一组键值对。



- 属性说明

  根据下表说明，填写需要进行指标数据采集的属性值。

参数	说明
instanceId	实例ID。
shard_id	分片ID。
ip	集群中节点的IP地址。
index	索引的名称。
primary	分片的属性，取值如下： ○ true：主分片 ○ false：副本分片 ○ 空：主分片和副本分片

- 高级配置

参数	说明
指标聚合	如果**tags**中存在多个取值，系统将生成多条曲线。**指标聚合**用来定义多个曲线合成一条曲线的算法。支持算法：sum()、avg()、max()、min()、count()。
采样聚合	由于指标预览图默认采样周期为1分钟，而底层默认采样周期为5s，因此系统会通过指标聚合算法，将1分钟内多个数据点聚合成一个点。支持算法：sum()、avg()、max()、min()。

- tags语法

  tags支持根据多个属性值进行过滤。例如同时对a集群和b集群的查询QPS进行监控报警，则**instanceId**设置为literal_or(a|b)，详细语法如下。

名称	说明	示例
literal_or	过滤出满足一个或多个属性值的数据。	host=literal_or(web01|web02|web03)：过滤出host为web01、web02或web03的数据。

名称	说明	示例
not_literal_or	过滤出不包含一个或多个属性值的数据。	host=not_literal_or(web01|web02|web03)：过滤出host不为web01、web02或web03的数据。
wildcard	过滤出满足通配符的属性值的数据。	host=wildcard(web*)：过滤出host以Web开头的数据。

触发条件

定义报警条件。即当监控指标项满足您定义的报警触发条件后，系统将通知您。

参数	说明
阈值报警	当监控指标到达或超过设置的阈值时，系统会触发对应的**WARNING**或**CRITICAL**报警。
波动报警	**波动报警**支持对波动变化率或变化值进行监控。例如当前指标的值为a，某个指定时间点前的指标的值为b，系统会计算差值(a-b)或者变化率(a-b)/b，并与设定的阈值进行比较，如果符合条件，则触发报警。
高级配置	• **多条件判断关系**：可选值为**AND**、**OR**。当您添加了多个触发条件时，设置为**AND**表示指标必须同时满足这些条件，才会触发报警；设置为**OR**，表示只要满足一个条件，就会报警。 • **连续触发几次**：连续触发几次报警后，通知报警人。默认为1，您可以按需修改。

无数据校验（可选）

当指标数据为空时，是否触发报警，默认为**忽略**。如果指定为**CRITICAL报警**，当连续一段时间（默认1分钟）没有监控到数据时，系统将进行无数据报警。

> **说明** 建议您选择**忽略**。如果遇到监控自身原因，导致采集的数据为空，也会触发报警。

规则触发后动作

参数	说明
生效时段	接收报警消息通知的时间段。默认每天24小时都接收通知，每隔5分钟发送一次。
通知人	发生报警时，需要通知的对象。支持选择联系人和联系组，如果选择联系组，系统会为该组中的所有成员发送消息。
通知方式	报警通知的方式。您可以为不同等级的报警指定不同的通知方式。

说明 完成配置后，您可以单击**校验**，校验**通知人**是否已配置对应的联系方式。校验结果对创建规则无影响。

5.2.3. 查看报警通知记录和事件

当您需要了解近一段时间内系统中所有报警组的通知记录、报警渠道资源占比、报警趋势等信息时，可在高级监控报警概览页面，获取这些信息的趋势图及历史记录。本文介绍具体的操作方法。

前提条件

创建报警组和报警规则。具体操作步骤请参见创建报警组和配置报警规则。

背景信息

概览页面展示了高级监控报警系统中所有报警组的通知记录和报警事件，如果您需要查看单个报警组的信息，可在**报警组列表**中获取，获取方法请参见查看通知记录和查看报警事件。

操作步骤

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。
3. 在左侧导航栏，单击**概览页**。

   您可以在**概览页**查看通知记录和报警事件，两者区别如下：

   - 通知记录：当报警事件生成后，系统会根据报警规则中设置的报警生效时段和报警间隔，判断是否给

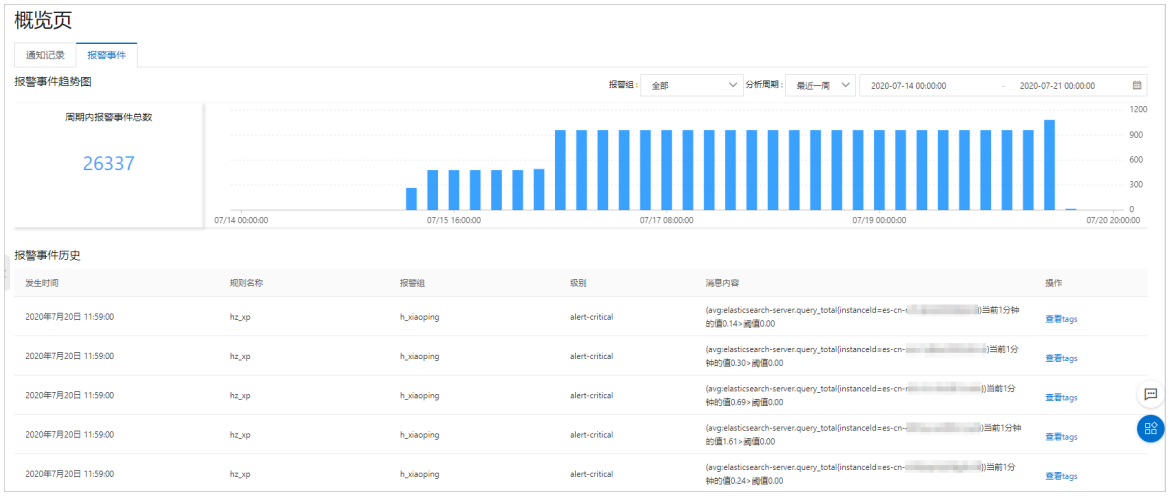
您发送报警通知（电话、短信、钉钉群机器人）。如果发送，则会生成一个通知记录。

- 报警事件：系统每隔1分钟，就会根据报警规则中设置的触发条件，判断指标是否触发报警。如果触发，则会生成一个报警事件记录。

4. 在**通知记录**页签，查看通知记录趋势图和明细。

- **通知记录趋势图**：默认展示最近一周的通知记录总数、时间分布、报警渠道资源用量、报警组分布、报警接收人分布及报警等级分布等。
- **通知记录明细**：默认展示最近一周的报警发生时间、规则名称、渠道、级别、通知人及消息内容等。

5. 单击**报警事件**，查看报警事件趋势图和历史。



5.3. 报警联系人

5.3.1. 管理报警联系人

报警联系人是指在使用高级监控报警功能，所选指标触发报警时，报警通知的接收对象。在设置报警规则时，您可以添加报警联系人，将报警通知发送给该联系人。本文介绍如何创建、查看或修改、停用或启用报警联系人。

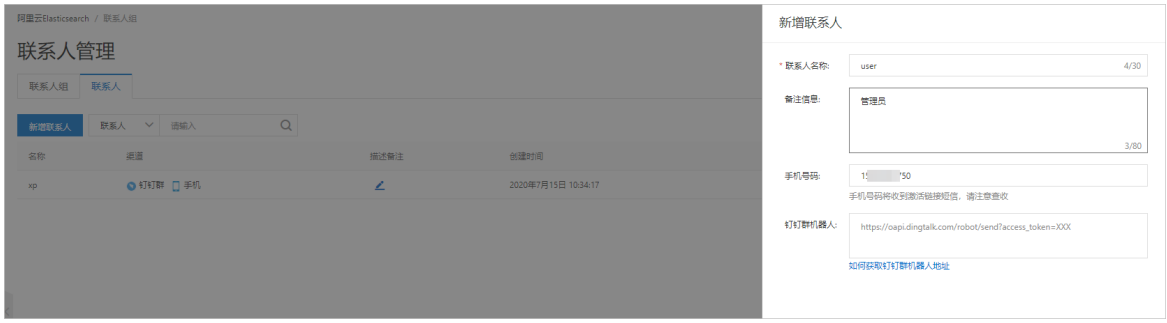
背景信息

阿里云监控报警服务支持通过以下方式，将报警通知发送给您：

- 短信
- 手机语音提示
- 钉钉群消息

新增联系人

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。
3. 在左侧导航栏，选择**指标报警模块 > 联系人管理**。
4. 在**联系人管理**页面，单击**联系人**页签。
5. 单击**新增联系人**。
6. 在**新增联系人**弹框中，输入联系人信息。



参数	说明
联系人名称	长度为0~30个字符，以大小写字母、数字或中文开头，可以包含下划线（_）或连字符（-）。**联系人名称**不可重复，且确定后无法修改。
备注信息	请填写有意义的备注信息，方便您快速查找定位。
手机号码	填写后，您的手机号码将收到激活链接短信。根据提示激活后，才可启用联系人。如果未激活，该联系人将无法接收到报警通知。
钉钉群机器人	填入钉钉群机器人的地址，获取方式请参见通过钉钉群接收报警通知。

7. 单击**确定**。

查看或修改联系人信息

1. 在**联系人**页签，单击目标联系人右侧**操作**列下的**查看/修改**。
2. 在**更新联系人**页面，查看或修改联系人信息。

> 注意 修改手机号码后，需要通过短信重新激活，才可接收到报警通知。

3. 单击**确定**。

停用或启用联系人

1. 在**联系人**页签，单击目标联系人右侧**操作**列下的**停用**。

   > 注意
   > - 执行**停用**操作时，目标联系人的状态必须为**已启用**或**未激活**。
   > - 系统不支持删除联系人。如果您不希望联系人收到报警通知，可停用该联系人。

   停用成功后，联系人的**状态**变为**已停用**，且不会再接收到报警通知。如果您需要再次启用联系人，可继续执行以下步骤。

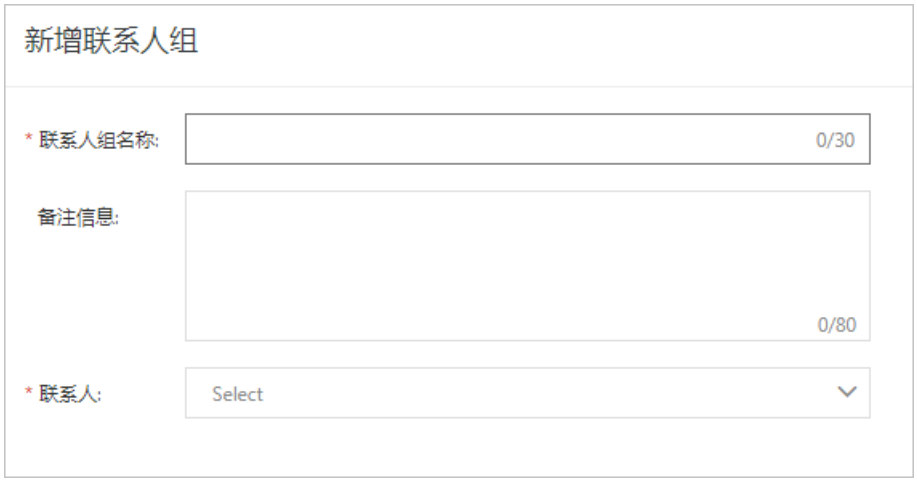
2. 单击已停用的联系人右侧**操作**列下的**启用**。
   启用成功后，联系人的**状态**变为**已启用**，且会重新接收到报警通知。

5.3.2. 管理报警联系人组

报警联系人组可以包含一个或多个报警联系人。同一个报警联系人，可以加入多个报警联系人组。在报警规则设置中，您可以添加报警联系人组，将报警通知发送给该组下所有联系人。本文为您介绍如何新增、查看或修改、停用或启用报警联系人组。

新增联系人组

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。
3. 在左侧导航栏，选择**指标报警模块 > 联系人管理**。
4. 在**联系人管理**页面，单击**联系人组**页签。
5. 单击**新增联系人组**。
6. 在**新增联系人组**弹框中，输入联系人组信息。



参数	说明
联系人组名称	长度为0~30个字符，以大小写字母、数字或中文开头，可以包含下划线（_）或连字符（-）。**联系人组名称**不可重复，且确定后无法修改。
备注信息	请填写有意义的备注信息，方便您快速查找定位。

参数	说明
联系人	选择一个或多个联系人，将联系人加入联系人组。如果还没有联系人，请先创建联系人，详情请参见新增联系人。

7. 单击**确定**。

查看或修改联系人组信息

1. 在**联系人组**页签，单击目标联系人组右侧**操作**列下的**查看/修改**。
2. 在**更新联系人组**页面，查看或修改联系人组信息。

   > **说明** 如果您不希望联系人组中的某一个或多个联系人收到报警通知，可在**联系人**列表中，删除对应联系人。

3. 单击**确定**。

停用或启用联系人组

1. 在**联系人组**页签，单击目标联系人组右侧**操作**列下的**停用**。

   > **注意**
   > - 执行停用操作时，目标联系人组的状态必须为**已启用**。
   > - 系统不支持删除联系人组。如果您不希望联系人组中的所有联系人收到报警通知，可停用该联系人组。如果您不希望该组下的某一个或多个联系人收到报警通知，可修改联系人组信息。

   停用成功后，联系人组的**状态**变为**已停用**，且该组中的所有联系人不会再接收到报警通知。如果您需要再次启用联系人组，可继续执行以下步骤。
2. 单击已停用的联系人组右侧**操作**列下的**启用**。
   启用成功后，联系人组的**状态**变为**已启用**，且该组中的所有联系人会重新接收到报警通知。

5.3.3. 通过钉钉群接收报警通知

阿里云Elasticsearch支持通过钉钉群接收报警通知。您只需要在联系人信息中添加钉钉群机器人的Webhood地址，就可以在钉钉群中收到报警通知，无需修改报警规则。本文为您介绍如何通过钉钉群接收报警通知。

前提条件

在PC中安装钉钉，并且创建用来接收报警通知的钉钉群。具体操作，请参见钉钉使用手册。

创建钉钉机器人（PC版）

1. 在PC中，打开您要接收报警通知的钉钉群，单击右上角的图标。
2. 在**群设置**页面中，单击**智能群助手**。
3. 在**智能群助手**面板，单击**添加机器人**。
4. 在**群机器人**对话框中，选择一个**自定义**机器人，单击**添加**。
5. 输入机器人名字，例如阿里云ES监控。

6. 在**安全设置**中，选中**自定义关键词**，并添加阿里云EMON关键词。
7. 勾选服务协议，单击**完成**。
8. 在**Webhook**右侧，单击**复制**，复制**Webhook**地址备用。

9. 单击**完成**。

在报警联系人中添加钉钉机器人

参见管理报警联系人，修改联系人信息，在**钉钉群机器人**中，填入您在创建钉钉机器人（PC版）时，获取的**Webhook**地址。

6.配置事件报警

通过配置事件报警，您可以及时获取控制台事件中心中的Elasticsearch集群的底层硬件运维事件，便于您及时查看和处理问题。本文介绍如何配置事件报警以及查看和处理事件。

前提条件

已在支持事件报警功能的地域下创建阿里云Elasticsearch实例：

- 支持事件报警功能的地域包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、中国香港。
- 创建实例的具体操作，请参见创建阿里云Elasticsearch实例。

使用限制

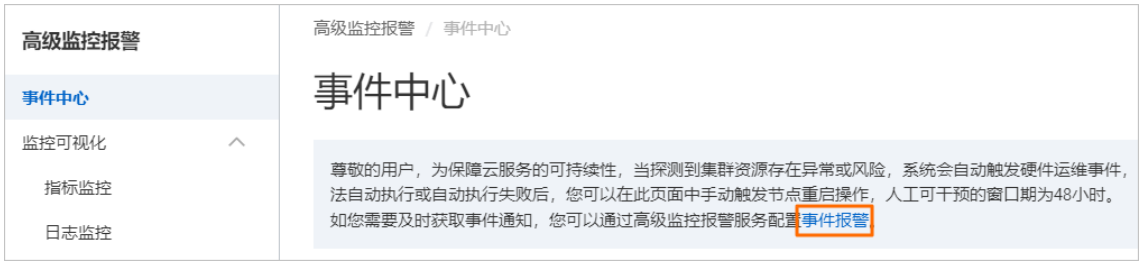
事件报警功能存在地域限制，支持的地域仅包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、中国香港。

创建报警规则

1. 登录阿里云Elasticsearch控制台。
2. 进入**报警组列表**页面。

   您可以通过两种方式进入：

   - 在左侧导航栏，单击**高级监控报警**。再在高级监控报警页面的左侧导航栏，选择**指标报警模块 > 报警组列表**。
   - 在**概览**页面的**事件中心**区域，单击**查看详情**。再在高级监控报警的**事件中心**页面，单击**事件报警**。



3. 创建事件报警。

   具体操作请参见创建报警组和创建报警规则。对应的报警规则中配置的参数如下：

   - 因探测节点失联触发的节点重启事件，配置如下图所示。

- 因底层资源运维触发的节点重启事件，配置如下图所示。

规则配置的详细参数说明，请参见创建报警规则。本示例的部分参数说明如下。

参数		说明
报警规则类型		选择**事件报警**。
	事件类型	**事件报警**支持两种事件类型： ○ 因探测节点失联触发的节点重启 ○ 因底层资源运维触发的节点重启

参数		说明
定义报警规则	报警范围	选择报警的目标实例。默认为**全区域下所有实例**，您也可以选择**自定义设置**。选择**自定义设置**后，需要选择region和instanceID： ○ **region**：目标实例所在地域，可选择一个或多个。 ○ **instanceID**：目标实例ID，可选择一个或多个。
	事件状态	事件的状态，支持3种：**已完成**、**执行失败**和**执行中**。 **说明** ○ 对于**因探测节点失联触发的节点重启**事件，系统会自动执行至完成，目前仅支持对执行结果配置报警，暂无**执行中**的事件状态。 ○ 对于**因底层资源运维触发的节点重启**事件，支持以上3种事件状态。

说明

- 目前已接入的底层运维事件，事件等级均为严重（CRITICAL）。
- 阿里云Elasticsearch不支持在同一个报警规则中同时选择多种事件类型或多种事件状态，您可以通过在同一报警组中配置多条报警规则的方式实现。

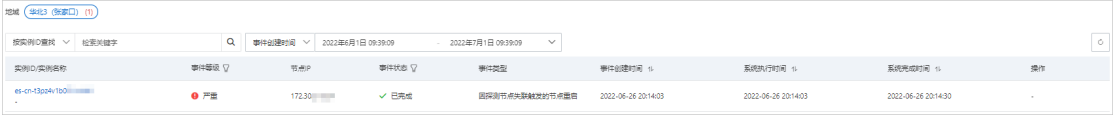
4. 配置接收报警通知。

   报警配置成功后，当您配置的事件发生时，您指定的报警通知人就可以通过配置的通知方式接收到报警通知，详细信息请参见通过钉钉群接收报警通知。

查看并处理事件

1. 查看事件。
   i. 登录阿里云Elasticsearch控制台。
   ii. 在**概览**页面的**事件中心**区域，查看近48小时内新增的事件中，执行失败和执行完成的数量。
   iii. 单击**查看详情**，进入**事件中心**页面，选择地域，查看对应地域下的事件。

   您可以按照实例ID或节点IP查找事件，也可以按照**事件创建时间**、**系统执行时间**或**系统完成时间**筛选事件。事件相关信息的详细说明，请参见事件中心。



2. 处理事件。

   对于执行失败的事件，如果**事件类型**为**因探测节点失联触发的节点重启**，且**事件状态**为**执行失败**时，支持用户在控制台进行手动重启节点进行异常干预。

> **说明** 重启节点仅需执行一次，如果问题未修复，系统会在下一次探测到异常时再次通知您。

7.日志报警

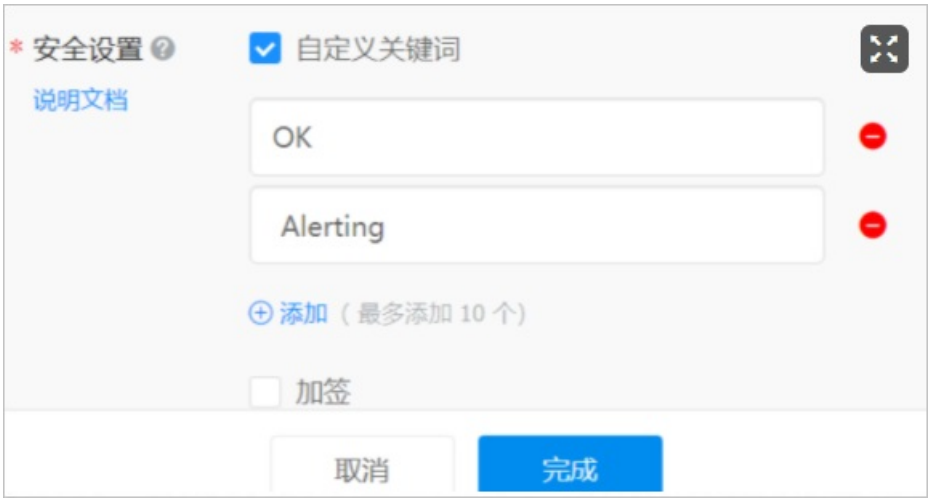
高级监控报警服务为您所有区域的Elasticsearch集群提供全维度指标和日志监控分析服务，不仅为您提供了多个维度下的可视化监控数据，还支持您根据业务需要自定义监控大屏和报警规则。本文为您介绍如何使用日志报警功能配置报警。

前提条件

- 已创建钉钉机器人，针对机器人配置了**OK**和**Alerting**关键词，并且获取到机器人**webhook**，具体操作参考通过钉钉群接收报警通知。
- 熟悉Grafana监控大屏的使用方法。详细信息，请参见Grafana Dashboard。

使用限制

- 高级监控报警服务存在地域限制，支持的地域仅包括：杭州、北京、上海、深圳、青岛、张家口、美国东部、美国西部、日本、印度、印度尼西亚、香港。
- 日志报警通知仅支持DingDing和Webhook两种方式，其他方式不支持。
- 钉钉机器人中**自定义关键词**必须是**OK**和**Alerting**，否则接收不到报警，如下图：



操作流程

1. 步骤一：进入高级监控报警页面
2. 步骤二：配置日志报警联系人相关信息
3. 步骤三：配置日志报警规则
4. 步骤四：查看日志报警规则

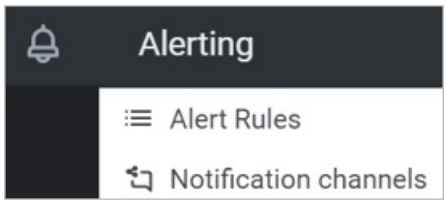
步骤一：进入高级监控报警页面

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。

步骤二：配置日志报警联系人相关信息

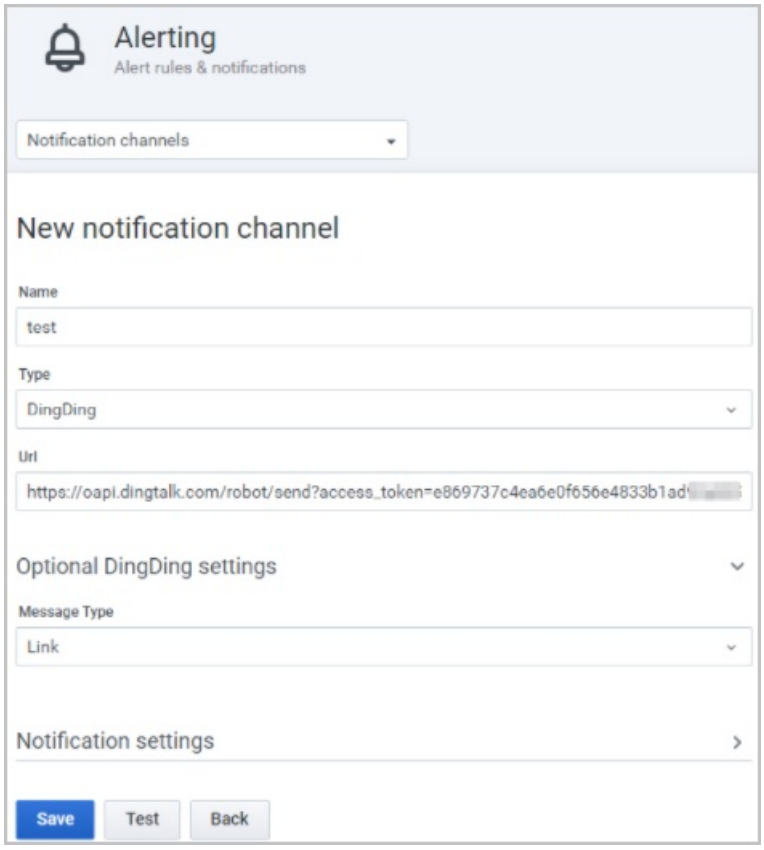
1. 在**高级监控报警**页面的左侧导航栏中，选择**监控可视化 > 日志监控**。
2. 单击**配置日志报警**，进入Grafana页面。

3. 单击左侧图标，选择**Notification channels**。



4. 配置通知渠道。填写信息可参考如下内容：

参数	说明
Name	自定义报警名称。本操作中使用的样例值为**test**。
Type	仅支持DingDing和Webhook提醒，其他不支持。本操作中使用的样例值为**DingDing**。
Url	钉钉机器人webhook路径。
Message Type	消息类型，仅支持**Link**。



5. 单击**Save**，保存配置信息。

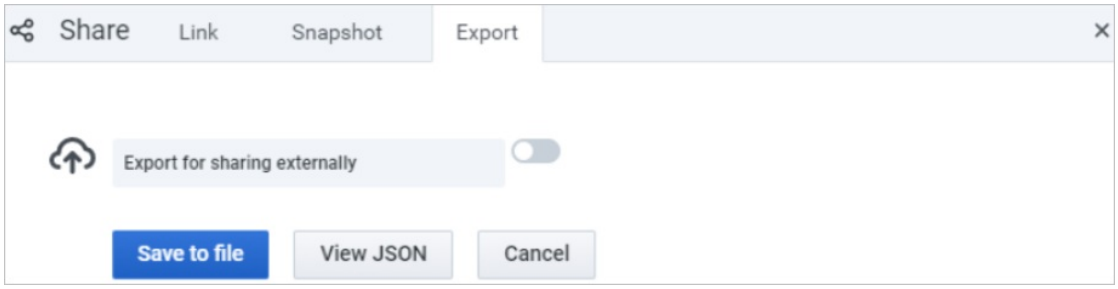
步骤三：配置日志报警规则

1. 在**高级监控报警**页面的左侧导航栏中，选择**监控可视化 > 日志监控**。
2. 单击**配置日志报警**，进入Grafana页面。

3. 使用报警模板配置日志报警规则（日志报警模板不支持直接编辑）。

   i. 在Grafana页面，单击监控模板上的图标。

   ii. 在对话框中，单击**Export**页签。

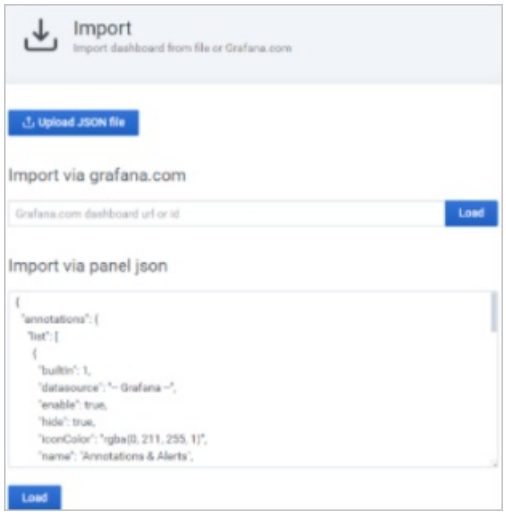


   iii. 单击**View JSON**。

   iv. 单击**Copy to Clipboard**，复制JSON模板。

   v. 单击左侧图标，选择**Import**。

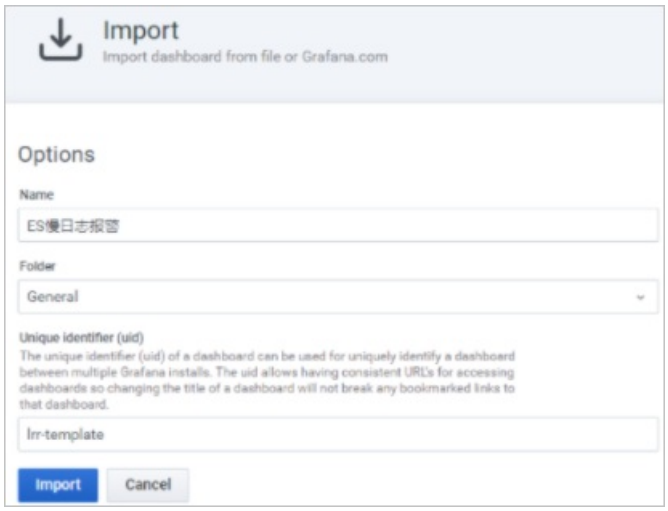
   vi. 在**Import via panel json**中，粘贴JSON模板中默认的报警规则。



   vii. 单击**Load**。

   viii. 修改**Name**，并重新定义**Unique identifier (uid)**。

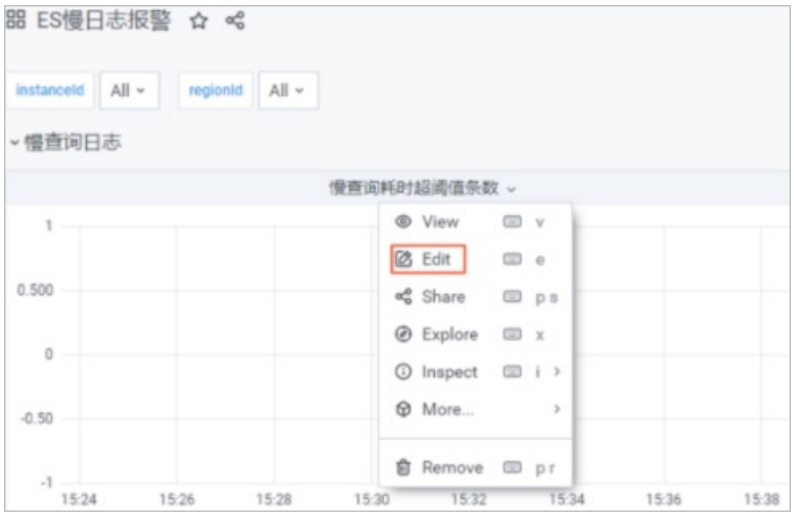
ix. 单击**Import**，即可完成模板的导入。



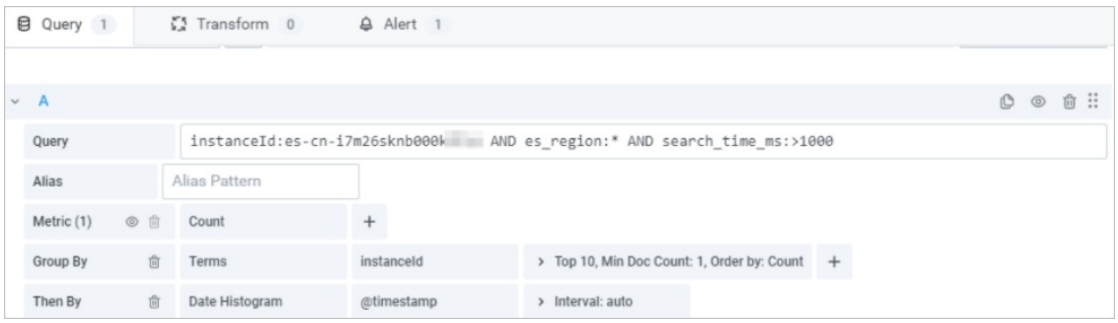
4. 调整报警规则。

   以配置**慢查询耗时超阈值条数**报警为例。

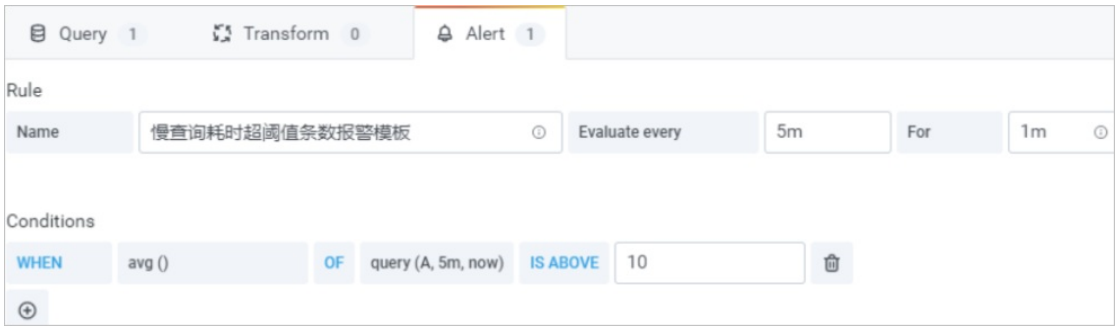
   i. 单击**慢查询耗时超阈值条数**，从下拉列表中选择**Edit**。



   ii. 单击**Query**页签，设置查询条件。当满足该查询条件时，将触发报警检测机制。

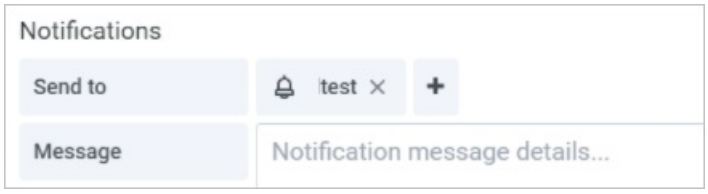


iii. 单击Alert页签，设置Rule和Conditions。



> **说明**
>
> - 系统默认报警规则为：每5分钟执行一次Query条件，在每一个过去的5分钟内，各时间点下慢查询耗时超过1000ms的日志条数平均值大于10条，且该状态持续超过1分钟，则系统上报告警。
> - Evaluate every为检测频率，最小值可配置1分钟，其他配置项无限制。

iv. 单击Send to后的+图标，添加步骤二中已经配置好的日志报警联系人。



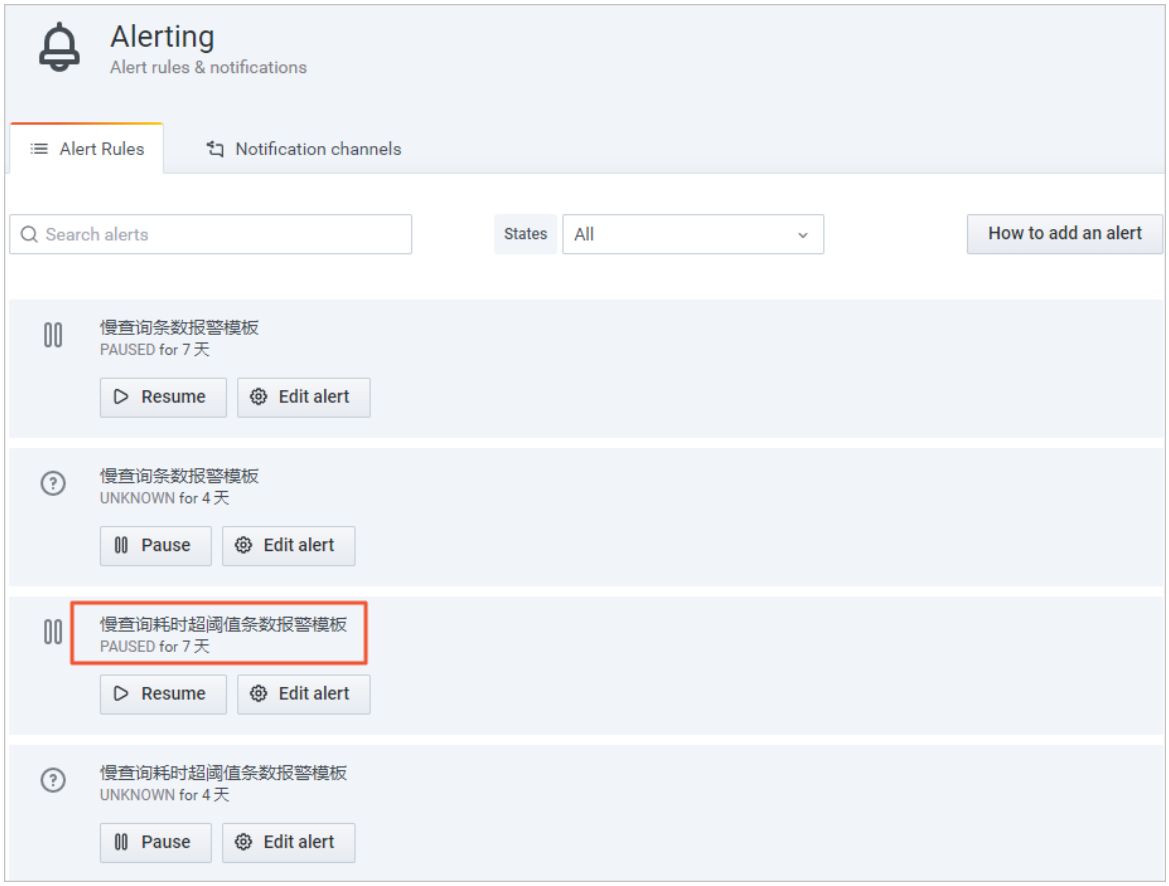
v. 单击Save。

vi. 单击Apply。

> **说明** 配置中未提及的项，可使用默认值，也可按需配置，更高阶的报警配置可以参考Grafana官方文档

步骤四：查看日志报警规则

1. 在**高级监控报警**页面的左侧导航栏中，选择**监控可视化 > 日志监控**。
2. 单击**配置日志报警**，进入Grafana页面。
3. 单击左侧图标，选择**Alert Rules**，获取报警规则列表。



说明 您还可以通过**监控可视化 > 自定义监控**查看成功配置的报警大盘。

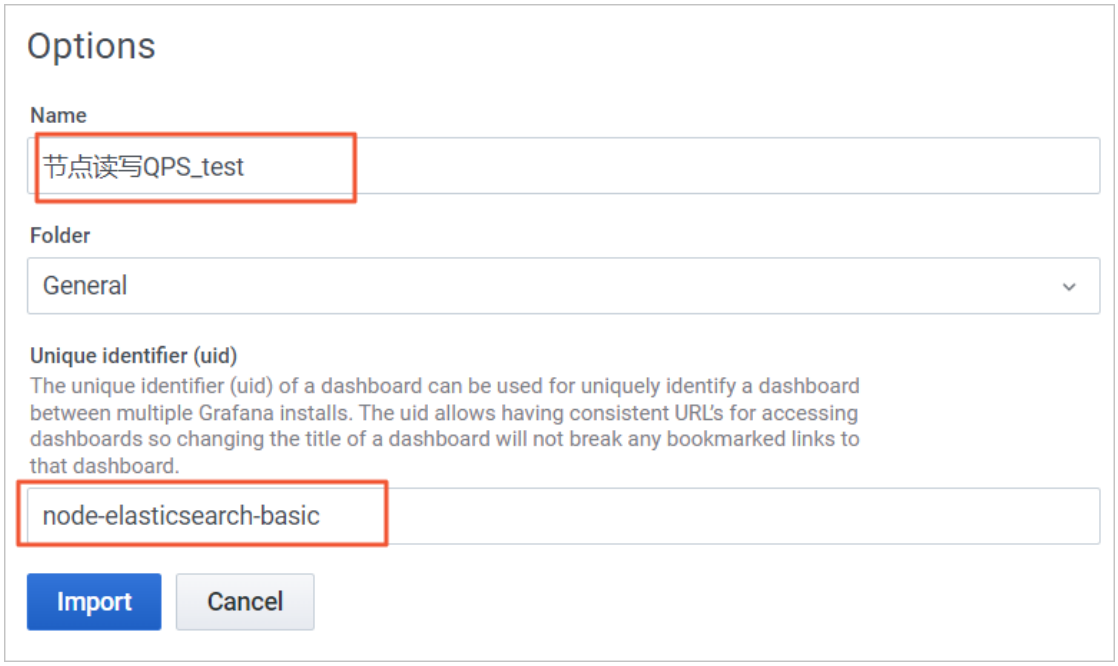
8.最佳实践

8.1. 自定义高级监控实战

高级监控报警服务支持您根据业务自定义监控大屏，帮助您更加灵活地监控Elasticsearch集群。作为默认监控能力的补充，确保满足您在不同场景下的监控需求。本文以配置节点维度的QPS监控、索引文档数监控以及删除文档数监控为例，为您介绍自定义监控的配置方法。

配置节点维度的QPS监控

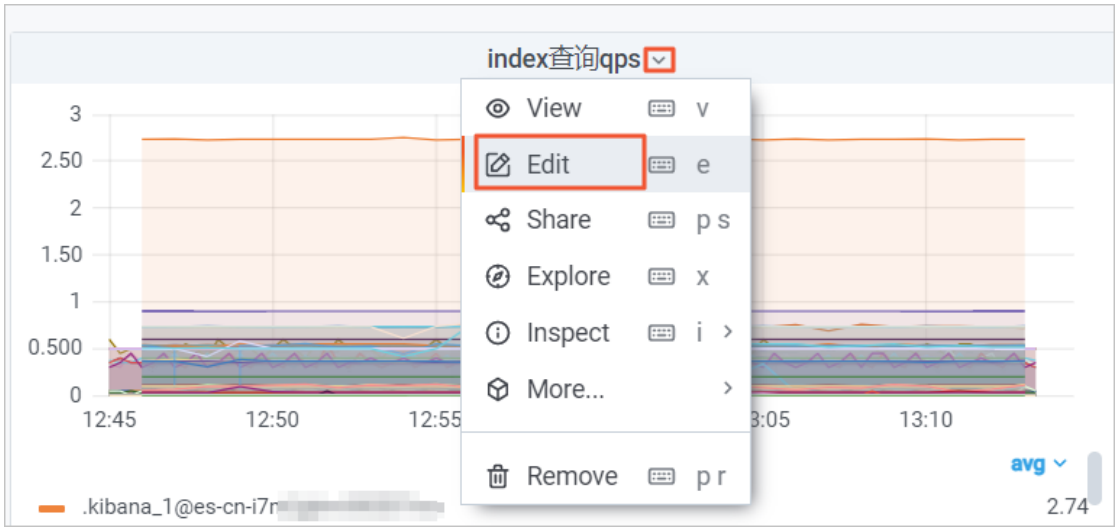
1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。
3. 在**高级监控报警**页面，复制默认索引QPS监控的JSON模板。
   i. 选择**监控可视化 > 指标监控**。
   ii. 在**默认基础指标**页签，鼠标左键单击监控窗口的任意空白处，然后按下键盘中的**Esc**键。
   操作成功后，当前页面会弹出Grafana菜单页及过滤栏。
   iii. 在Grafana页面，单击**基础指标大盘**右侧的图标。
   iv. 在对话框中，单击**Export**页签。
   v. 单击**View JSON**。
   vi. 单击**Copy to Clipboard**，复制JSON模板。
4. 导入模板。
   i. 在左侧Grafana菜单栏中，单击图标，选择**Import**。
   ii. 在**Import via panel json**输入框中，粘贴已复制的JSON模板，单击**Load**。
   iii. 修改**Name**，并重新定义**Unique identifier (uid)**。



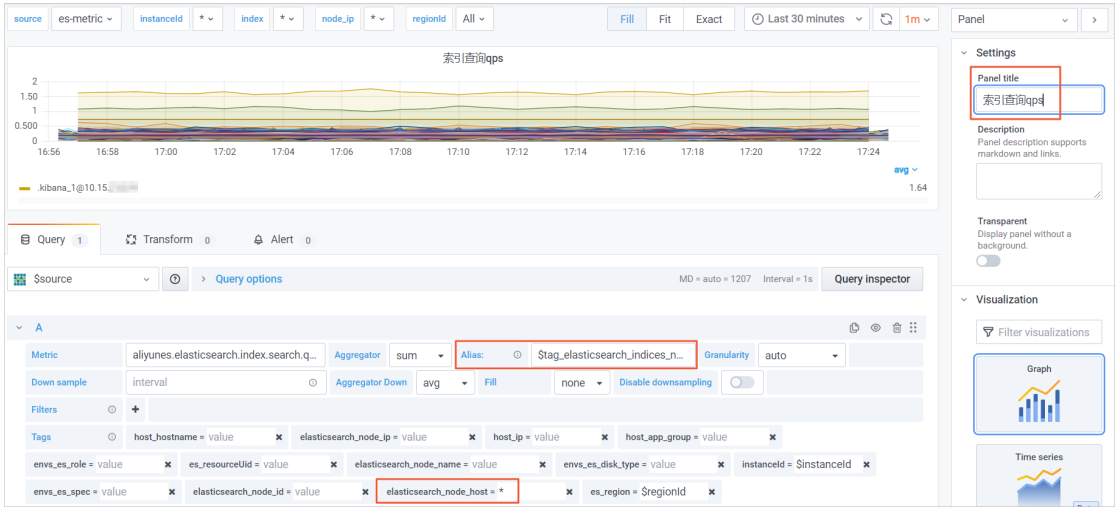
   iv. 单击**Import**，即可完成模板的导入。

5. 在**自定义监控**页面，配置索引写入或查询QPS监控。
   i. 在左侧导航栏，选择**监控可视化 > 自定义监控**。
   ii. 在页面上方的自定义监控列表中，单击您自定义的监控模块页签。

   iii. 展开**Index(索引)**模块，将鼠标悬浮至目标监控指标名称上，单击右侧的图标，选择**Edit**。



iv. 按照以下说明配置索引查询QPS监控。



参数	说明
Panel title	设置监控面板的标题。本示例设置为**节点查询qps**。
Alias	设置监控面板中节点信息的显示格式。本示例设置为**$tag_elasticsearch_indices_name@$tag_elasticsearch_node_host**，表示节点信息的显示格式为索引名称@节点IP地址，例如.kibana_1@10.15.xx.xx。
elasticsearch_node_host	通过输入节点IP地址，指定需要监控的节点。本示例设置为*，表示监控集群中所有节点的索引查询QPS。

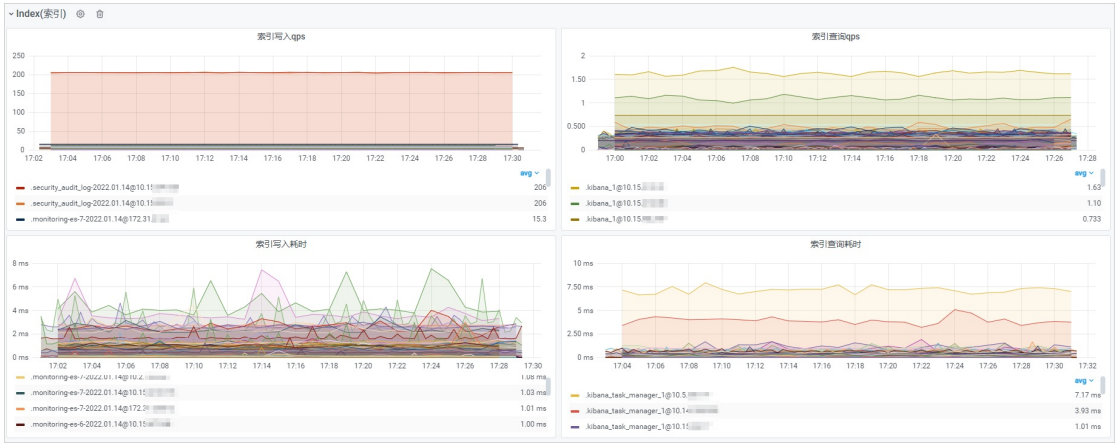
配置完成后，您可在页面上方预览配置效果。

v. 单击右上角的**Save**，按照页面提示保存配置。

vi. 单击**Apply**，应用配置。

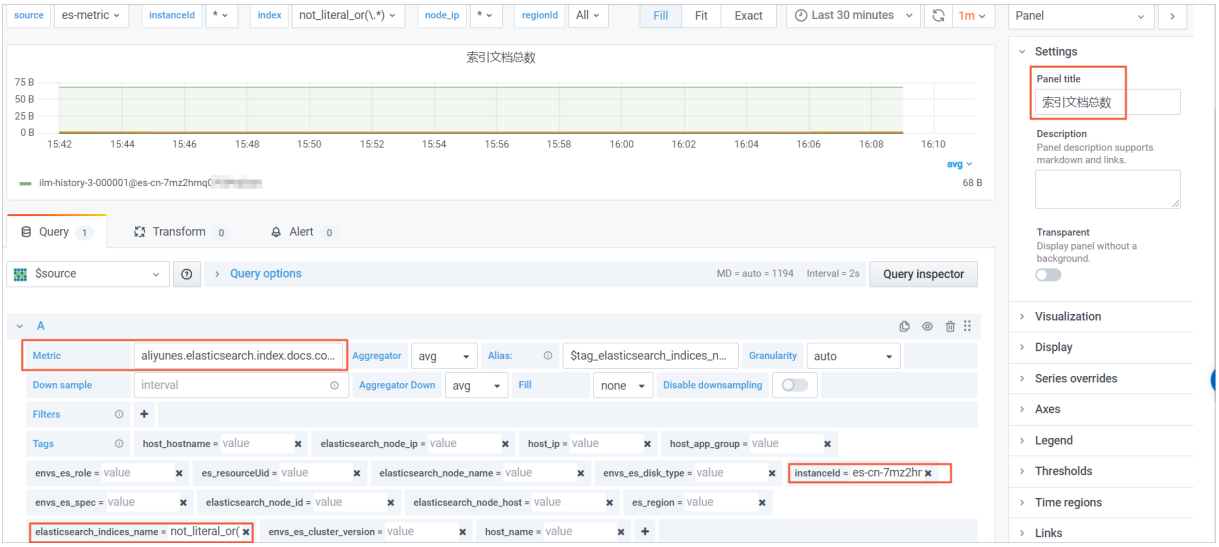
vii. 使用同样的方式配置索引写入QPS、索引写入耗时和索引查询耗时监控。

本文的配置效果如下。



配置索引文档总数监控

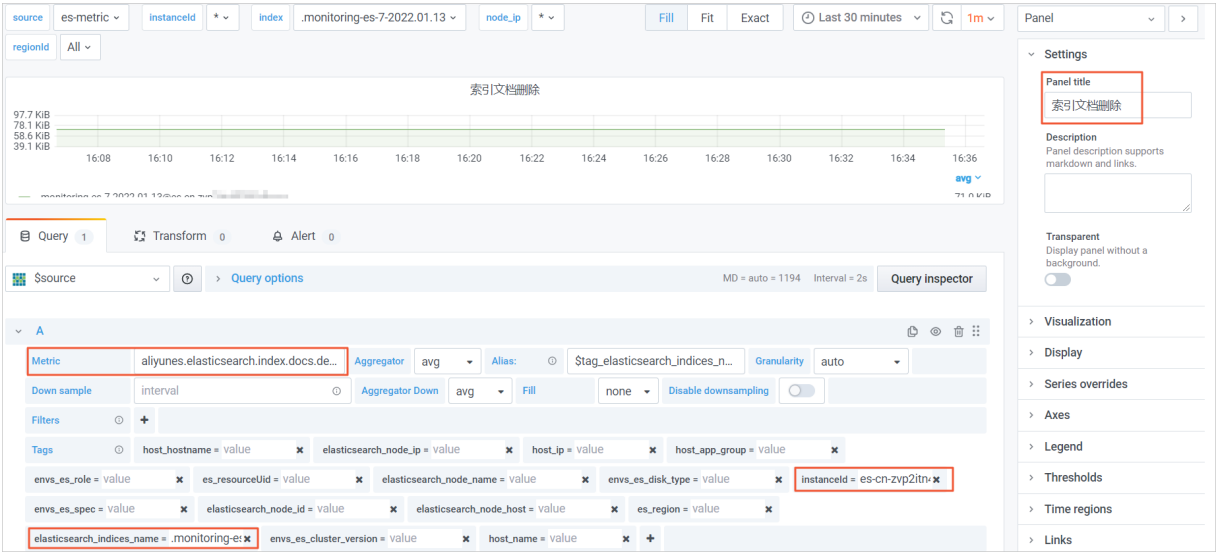
参见配置节点维度的QPS监控，编辑任意Index(索引)模块中的监控指标，在其基础上配置索引文档总数监控。本文需要修改的配置如下。



参数	说明
Panel title	设置监控面板的标题。本示例设置为**索引文档总数**。
Metric	设置需要监控的指标名称。本示例设置为**aliyunes.elasticsearch.index.docs.count**，表示需要监控的指标为索引中文档的总数。
instanceId	通过输入实例ID，指定需要监控的实例。
elasticsearch_indices_name	通过输入索引名称，指定需要监控的索引。本示例设置为**not_literal_or(\.*)**，表示监控除了.开头的系统索引外的所有索引的文档总数。

配置索引文档删除情况监控

参见配置节点维度的QPS监控，编辑任意Index(索引)模块中的监控指标，在其基础上配置索引文档删除情况监控。本文需要修改的配置如下。



参数	说明
Panel title	设置监控面板的标题。本示例将其设置为**索引文档删除**。
Metric	设置需要监控的指标名称。本示例设置为**aliyunes.elasticsearch.index.docs.deleted**，表示需要监控的指标为索引文档的删除情况。
instanceId	通过输入实例ID，指定需要监控的实例。
elasticsearch_indices_name	通过输入索引名称，指定需要监控的索引。本示例设置为**.monitoring-es-7-2022.01.13**，表示监控**.monitoring-es-7-2022.01.13**索引中文档的删除情况。

8.2. 指标报警配置最佳实践

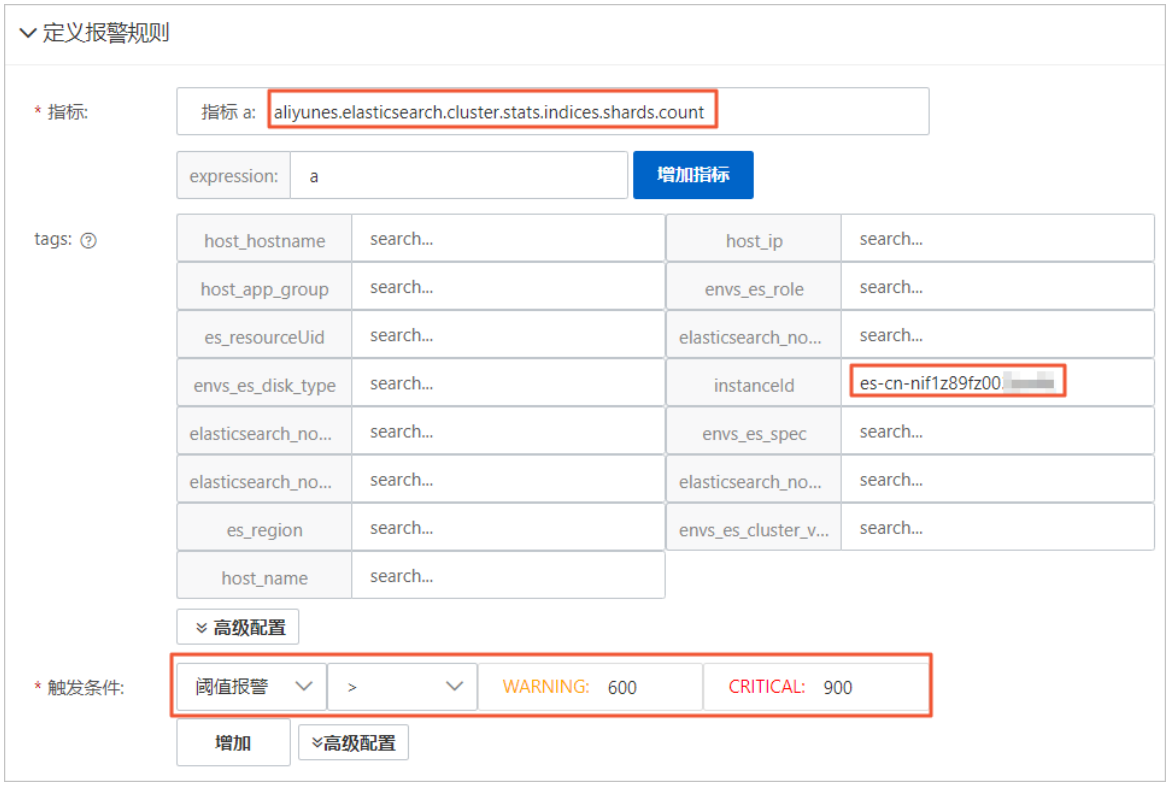
通过指标报警，您可以设置多维度的监控指标和Tags，帮助您快速定位Elasticsearch的性能问题，提高运维排查效率。本文以配置集群shard数监控报警、节点个数监控报警和查询队列监控报警为例，为您介绍如何将指标报警配置应用到具体的业务中。

集群shard数监控报警

Elasticsearch 7.x版本开始对单机分片数进行了限制，默认单机分片数不能超过1000。高级监控报警提供的集群分片数监控报警能力，可以对单实例分片总数进行报警。您可以参见Shard评估规划单机分片数，当单机分片总数达到阈值建议优化索引。

1. 登录阿里云Elasticsearch控制台。
2. 在左侧导航栏，单击**高级监控报警**。
3. 在**高级监控报警**页面的**指标报警模块**，参见管理报警组、配置报警规则和管理报警联系人，定义集群shard数监控报警规则及报警联系人。

   以2核4 GB，3个数据节点的集群为例。按照shard评估，建议单节点的shard数在120~200之间，三个节点的总shard数在360~600之间。当集群shard数大于600时进行WARNING报警，超过900时进行CRITICAL报警。对应的报警规则配置如下。



规则配置的详细参数说明，请参见配置报警规则。本示例的部分参数配置如下。

参数	配置
指标	选择**aliyunes.elasticsearch.cluster.stats.indices.shards.count**。
tags	**instanceId**设置为待监控的实例ID。
触发条件	选择**阈值报警**。设置集群shard数>600进行WARNING报警，>900进行CRITICAL报警。

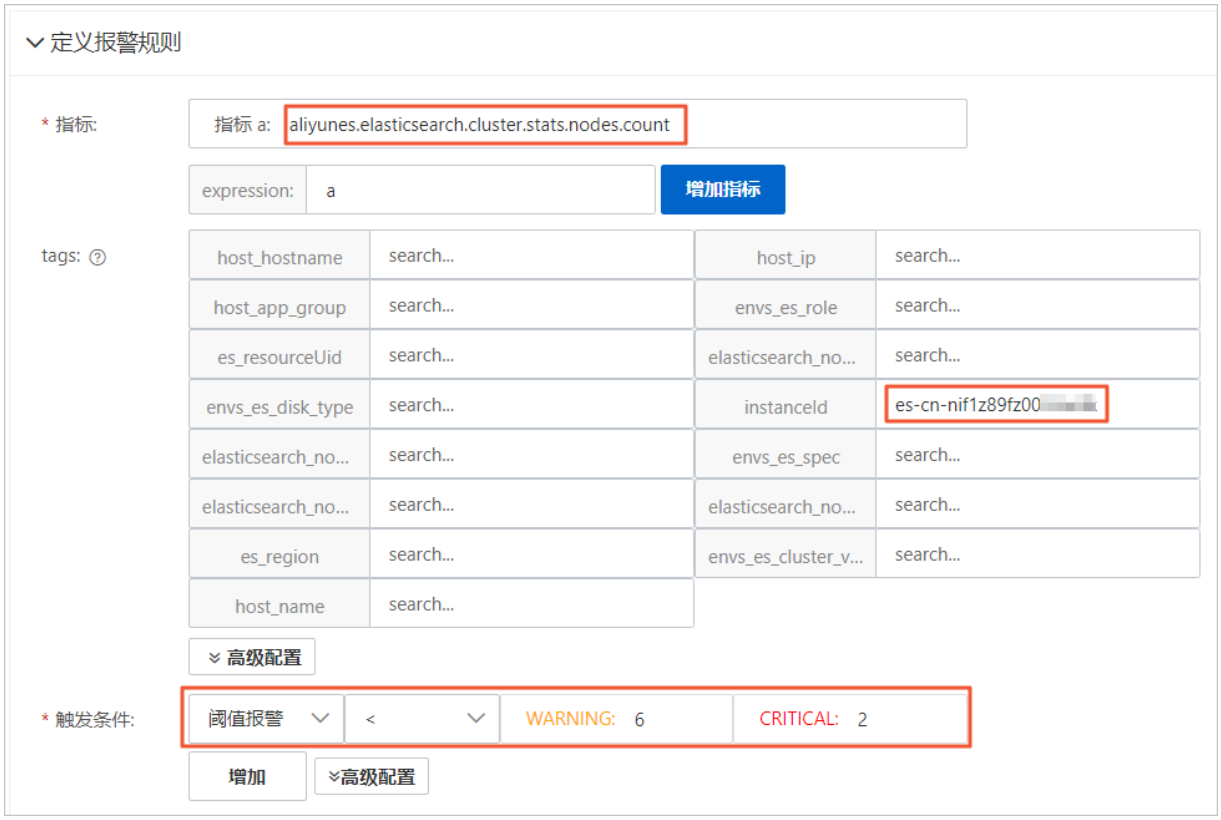
4. 验证结果。

   报警配置成功后，当集群shard数超过设定阈值时，您指定的报警通知人就可以通过钉钉群接收到报警通知，详细信息请参见通过钉钉群接收报警通知。

节点个数监控报警

节点脱离集群后不易被发现，脱离时间太久节点会自动从集群隔离。为解决此问题，您可以配置高级监控报警，对集群中的节点个数进行监控。

在**高级监控报警**页面的**指标报警模块**，参见管理报警组、配置报警规则和管理报警联系人，定义集群节点个数监控报警规则及报警联系人。本示例的报警规则配置如下。



规则配置的详细参数说明，请参见配置报警规则。本示例的部分参数配置如下。

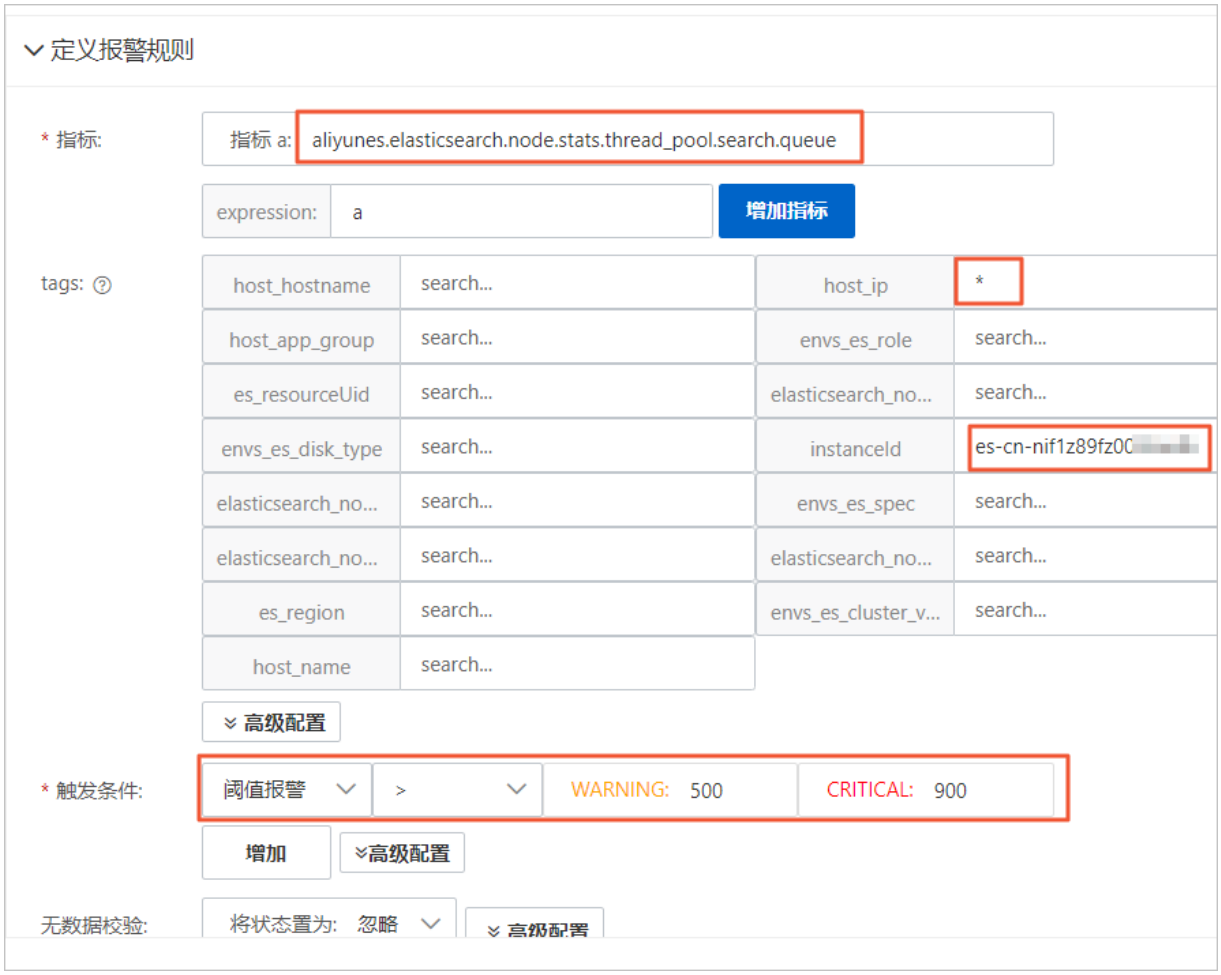
参数	配置
指标	选择**aliyunes.elasticsearch.cluster.stats.nodes.count**。
tags	**instanceId**设置为待监控的实例ID。
触发条件	选择**阈值报警**。设置集群节点个数<6进行WARNING报警，<2进行CRITICAL报警。

报警配置成功后，当集群中的节点个数小于设定阈值时，您指定的报警通知人就可以通过钉钉群接收到报警通知，详细信息请参见通过钉钉群接收报警通知。

查询队列监控报警

Elasticsearch的查询队列大小默认为1000，当队列堆积严重时，新的请求将被中止。您可以通过阿里云Elasticsearch的高级监控报警功能，对数据节点查询队列的等待任务数进行监控报警。

在**高级监控报警**页面的**指标报警模块**，参见管理报警组、配置报警规则和管理报警联系人，定义数据节点查询队列等待任务数监控报警规则及报警联系人。本示例的报警规则配置如下。



规则配置的详细参数说明，请参见配置报警规则。本示例的部分参数配置如下。

参数	配置
指标	选择**aliyunes.elasticsearch.node.stats.thread_pool.search.queue**。
tags	• **host_ip**：设置为*。 • **instanceId**：设置为待监控的实例ID。
触发条件	选择**阈值报警**。设置数据节点查询队列等待任务数>500进行WARNING报警，>900进行CRITICAL报警。

报警配置成功后，当数据节点查询队列等待任务数大于设定阈值时，报警通知人就可以通过钉钉群接收到报警通知，详细信息请参见通过钉钉群接收报警通知。