目录

产品概述

Elasticsearch 是一款基于Lucene构建的，支持RESTful API的分布式搜索和分析引擎。具有内存消耗小、搜索快、扩展性强等优点，可提供近乎实时的存储、搜索和分析超大数据集的能力。

金山云Elasticsearch服务（简称KES）是基于开源Elasticsearch打造的低成本、高可用的云端全托管Elasticsearch服务。内置Elasticsearch 5.6.16、6.8.4和7.4.2版本、Kibana及常用插件。保持了 Elasticsearch 本身兼容与开放性的同时，在开源基础上，提供企业级安全管控、监控告警以及丰富的集群管理功能，多种计算规格和存储介质供选择，使用金山云Elasticsearch服务您可一键部署、按需升配、轻松管理、安全使用Elasticsearch服务，快速构建搜索中台、日志分析、性能监控等业务场景。主要功能如下：

数据采集与同步

- 用户通过 Beats 功能，可以把数据传输到 KES 中进行存储，也可以传输到 Logstash 中进行自定义转换和解析后，再传输到KES中。
- KES提供了易用的 RESTful API，用户可以自行开发客户端，调用数据存储 API，存储数据到KES集群中。

数据存储

- 多种机型和存储介质供选择，灵活应对各种业务场景，在有效保障数据读写性能的同时降低成本,实现数据冷热分离，读写分离。
- 支持弹性扩容，扩展到上百个节点，能达到 PB 级数据的存储。
- 支持故障节点探测及替换，保障集群高可用性。

数据可视化搜索分析

- KES拥有全文检索、结构化搜索、数据过滤和指标统计等搜索功能，可以应用于信息搜索和数据分析等多种场景。
- KES提供了简单易用的 RESTful API 以及各种语言的客户端，用户可以很方便地构建自己的搜索服务。
- 使用 Kibana，用户可以方便地在浏览器里对集群的数据进行搜索和统计分析。

基本概念

集群(Cluster)

集群是一个或多个节点（服务器）的集合，这些节点共同保存整个数据，并在所有节点上提供联合索引和搜索功能。

节点(Node)

一个节点是集群中的一个服务器，用来存储数据并参与集群的索引和搜索。KES区分了四种节点类型，如下：

- 数据节点（Data Node），存储索引数据的节点，主要对文档进行增删改查操作，聚合操作等。数据节点对cpu，内存，io要求较高，在优化的时候需要监控数据节点的状态，当资源不够的时候，需要在集群中添加新的节点。如果集群中只有数据节点，系统会默认3台作为主节点。

- 专有主节点（Master Node），为确保一个集群的稳定，分离主节点和数据节点，独立出3台机器担当主节点。主要职责是和集群操作相关的内容，如创建或删除索引，跟踪哪些节点是集群的一部分，并决定哪些分片分配给相关的节点。在大规模集群中，开启专有主节点，可增强集群的稳定性。

- 协调节点（Coordinator Node），该节点只处理路由请求，处理搜索，分发索引操作，相当于一个智能负载平衡器，协调节点将请求转发给存储数据的Data Node。每个Data Node在本地执行请求，并将结果返回协调节点。协调节点收集完数据后，将每个数据节点的结果合并为单个全局结果。

- 冷数据节点（Warm Node），这种类型的节点被用于存放只读并且很少被查询的索引。与之区别，热数据节点存放查询频率高，时间较近的数据索引，写入压力大，这样实现冷热分离，在提升查询和写入效率的同时降低存储成本。

索引（Index）

索引是有相同特性的文档集合，(相当于关系型数据库里的一个数据库)。它是我们存储和索引关联数据的地方。一个索引通常使用一个名称（所有字母必须小写，不能以下划线开头，不能包含逗号）来标识，当针对这个索引的文档执行索引、搜索、更新和删除操作的时候，这个名称被用来指向索引。

类型（type）

一个类型通常是一个索引的一个逻辑分类/分区，允许在一个索引下存储不同类型的文档（相当于关系型数据库中的一张表），在Elastic 6.x 版本，只允许每个 Index 包含一个 Type，在Elastic 7.x 及以后版本中Type概念被删除。

文档（document）

一个文档是可以被索引的基本信息单元（相当于关系型数据库中的一行数据）。文档可以用JSON格式来表示。在一个索引

中，您可以存储任意多的文档，且文档必须被索引。

映射（Mapping）

模式映射（相当于关系型数据库中的schema）用于定义索引结构。Elasticsearch在映射中存储有关字段的信息。映射在文件中以JSON对象传送。

字段（Field）

字段是ElasticSearch里的最小单元，相当于数据的某一列，类似于json里一个键。

分片（Shards）

索引分片,当有大量的文档时，由于内存的限制、硬盘能力、处理能力不足、无法足够快地响应客户端请求等，一个节点不够的情况下， Elasticsearch可以把一个完整的索引分成多个分片，分布到不同的节点上，构成分布式搜索。分片的数量只能在索引创建前指定，并且索引创建后不能更改。

副本(Replica)

Elasticsearch可以设置多个索引分片的副本，副本具有以下作用：

- 提高系统的容错性，当某个节点某个分片损坏或丢失时可以从副本中恢复。
- 提高Elasticsearch的查询效率，Elasticsearch会自动对搜索请求进行负载均衡。

应用场景

搜索服务

广泛应用于应用程序搜索、网站搜索、企业搜索等场景，我们每天工作生活，都需要用“搜索”快速锁定要找的内容，如搜索商品、搜索位置、搜索视频、搜索文档、搜索酒店航班等。ES为您提供高可用、高并发、低延时的搜索体验，仅需要几毫秒的时间，即可帮您在 PB 级的数据中搜索到匹配信息。

日志实时分析

应用于业务日志（如用户行为日志）、状态日志（如慢日志、异常日志）、系统日志、审计日志等日志实时分析的场景。日志在异常情况下可以帮我们定位问题，但它存储分散、种类多样、规模庞大，ES提供了全套的解决方案，使用Beats、Logstash可快速对接各种数据源，数据统一存储在ES中，利用全文搜索功能，以秒级交互式分析响应性能对海量日志进行分析，并借助Kibana默认的仪表板可视化展示分析结果。让日志快速产生价值而不至于成为累赘。

时序数据分析

应用于基础设施指标、容器监测、应用程序性能监测、IOT设备监控等场景。ES支持每秒千万级的时序写入，具备多维度、灵活可扩展的统计分析能力和秒级数据响应性能，您可以快速感知系统中实时发生的事件，也可以观察其历史趋势性、规律性、异常性，并做出预测和预警。

商业智能分析

在数据驱动决策的趋势下，越来越多的企业借助数据分析和挖掘来辅助商业决策比如一个大型商场，想分析一下某区域最近3年的用户消费金额的趋势以及用户群体的构成，并产出相关报表，面对数据体量越来越大，分析时效越来越高的场景，ES拥有结构化查询的能力，支持复杂的过滤和聚合统计功能，辅助决策。

产品优势

易于部署和管理

即开即用，您只需简单操作，即可在线开通并创建一个 KES 集群，免去软硬件部署调试的复杂流程。通过KES的控制台您可以对集群扩缩容、升配、插件安装卸载、集群配置管理、集群和节点各项指标监控等操作，简化您日常的集群管理工作。您可以使用 Kibana 对 Elasticsearch 索引中的数据进行搜索、查看、交互操作。您可以很方便的利用图表、表格及地图对数据进行多元化的分析和呈现。

冷热分离，弹性扩缩容

KES区分数据节点、专有主节点、协调节点和冷数据节点，各角色节点各尽其责，实现冷热分离，读写分离，您可以根据需求选择各角色节点的规格和存储介质，并动态调整集群的配置、扩容缩容或升配，充分发挥分布式集群的优势，灵活应对各种业务场景，实时保障业务发展并有效控制成本。

安全保障

通过部署在逻辑隔离的私有网络 VPC，客户可以完全掌控自己的环境配置，自定义网络访问控制列表，配置安全组策略。集成

了x-pack加密通信、基于角色的访问控制、文件和原生身份验证、Kibana Spaces、Kibana 功能控制、API 密钥管理等安全特性，切实保障您云上资源的安全性。

高可靠

通过快照自动备份与恢复、快照存储与恢复以及负载均衡，确保了集群的高可靠性。您可以在金山云Elasticsearch控制台中的数据备份页面开启自动备份功能，并根据实际业务需要设置自动备份周期，实现每天定时自动备份数据快照。快照备份存储在KS3(Kingsoft Standard Storage Service)中，您可以手动执行恢复快照命令来恢复对应的索引数据，该方案可实现一定程度的数据容灾。