

目录

目录	1
步骤一：规格容量评估	2
存储容量评估	2
Shard数评估	2
集群规格评估	2
机型选择建议	2
步骤二：创建集群	2
步骤三：访问集群	3
通过 Kibana 访问集群	3
步骤四：业务查询	3
创建索引	3
创建文档并插入数据	4
查询数据	4

步骤一：规格容量评估

存储容量评估

影响金山云托管ES 服务磁盘存储的主要因素如下：

- 副本数量：默认和建议的副本数量为1，对于可以承受异常情况导致数据丢失的场景，可考虑设置副本数量为0。
- 索引开销：除原始数据外，ES 需要存储索引、列存数据等，通常比源数据大10%（_all等未计算）。
- ES内部任务开销：segment 合并、ES Translog、日志等，占用约20%的磁盘空间。
- 操作系统预留：Linux 操作系统默认为 root 用户预留5%的磁盘空间，用于关键流程处理、系统恢复、防止磁盘碎片化等问题。
- 安全阈值。预留20%的安全阈值。

根据以上因素得到最小磁盘总大小 = 源数据大小 * 3.6。计算方式如下：

磁盘总大小 = 源数据 * (1 + 副本数量) * (1 + 索引开销) / (1 - Linux预留空间) / (1 - ES内部任务开销) / (1 - 安全阈值) = 源数据 * (1 + 副本数量) * 1.8 = 源数据 * 3.6

Shard数评估

shard大小和数量是影响ES集群稳定性和性能的重要因素之一。ES集群中任何一个索引都需要有一个合理的shard规划（默认为5个）。

- 单个shard大小不要超过50G，建议单个分片大小在10G-50G之间。
- shard的个数（包括副本）要尽可能等于节点数，或者是节点数的整数倍。
- 过多的分片数影响集群状态管理，建议单个es实例每1G内存对应的分片数不要超过20个（假设es实例内存是10G，那么这个es实例上的分片数不要超过200个）

集群规格评估

ES 的计算资源主要消耗在写入和查询过程，而不同业务场景在写入和查询方面的复杂度不同、比重不同，导致计算资源相比存储资源较难评估。因此建议您优先评估存储资源量，然后初步选择计算资源，在测试过程中确认计算资源是否足够。

我们建议您首先至少选择3个节点，避免 ES 实例出现脑裂问题，保证 ES 实例具有较高的节点故障容错能力。

当完成实例类型的初步选择后，您可以使用真实数据进行测试，通过观察 CPU 使用率、写入指标（性能、拒绝率）、查询指标（QPS、拒绝率）等监控信息，进一步确认实例类型是否合适。另外，建议针对上述监控信息配置告警，方便在线上使用时，及时发现资源不足等问题。

机型选择建议

- 本地SSD型，适合存储要求较小，性能优越，数据量较小客户，成本较低；
- 云盘型，适合对存储容量要求较高，且兼顾性能，对数据高可用要求较高，如审计日志、金融数据等场景，成本偏高。
- 云物理机，适合对性能有极致的要求，且业务规模庞大的客户。

步骤二：创建集群

集群是 KES 提供托管 Elasticsearch 服务的基本单元，也是用户使用和管理 Elasticsearch 服务的主要对象。本文为您介绍通过金山云官网控制台，快速创建 Elasticsearch 集群。

The screenshot displays the 'ElasticSearch > 新建集群' (New Cluster) configuration interface. It is organized into three main sections:

- 计费信息 (Billing Information):** Offers four payment methods: '预付费(包年包月)' (Prepaid Annual Package), '按日配置付费(月结)' (Daily configuration pay monthly), '按小时配置实时付费' (Hourly configuration real-time pay), and '按日配置付费(试用)' (Daily configuration pay trial). The '1月' (1 month) purchase duration is selected.
- 基本信息 (Basic Information):** Includes fields for '集群名称' (Cluster Name) set to 'es-20200228233654', '数据中心' (Data Center) set to '广州1区(VPC)', '可用区' (Availability Zone) set to '可用区A', '产品版本' (Product Version) set to '开源版' (Open Source), and '版本号' (Version Number) set to '6.8.4'.
- 节点组配置 (Node Group Configuration):** Shows '配额信息' (Quota Information) as 'CPU: 1800核 (1296核 可用); 内存: 6000GB (4260GB 可用); 磁盘: 300000GB (285650GB 可用)'. Under '数据节点' (Data Node), the '机型' (Instance Type) is '云盘型' (Cloud Disk), '规格' (Specification) is 'CPU: 2核, 内存: 4G', '云盘配置' (Cloud Disk Configuration) is 'SSD云盘', '云盘大小' (Cloud Disk Size) is '20 GB', and '节点数量' (Number of Nodes) is '3'.

A '配置详情' (Configuration Details) sidebar on the right lists the selected settings: 集群名称: es, 数据中心: 广, 计费方式: 预, 购买时长: 1j, 可用区: 广, 产品版本: 开, 版本号: 6, 数据节点: (1), 开启公网: 关, 项目组: 默.

操作步骤

1、登录金山云控制台，选择大数据 -> Elasticsearch 服务，进入KES控制台

2、在列表页，点击“新建集群”进入创建集群页面。

3、填写集群配置信息

- 计费方式：支持【预付费（包年包月）】、【按小时实时付费】、【按日配置付费（试用）】、【按日配置付费（月结）】
- 集群名称：创建集群时，会根据系统时间戳生成一个默认名称。您也可以为KES集群输入自定义名称。长度限制为1-25个字符，支持数字、大小写字母、减号和下划线。
- 数据中心：支持北京6区，上海2区和广州1区
- 可用区：北京支持A、B、C三个可用区；上海的A、B可用区；广州的A可用区
- 产品版本：目前仅开放开源版
- 版本号：支持ElasticSearch5.6.16；ElasticSearch6.8.4；ElasticSearch7.4.2

- 数据节点：机型可选云盘型、本地SSD型、云物理机；节点数量不少于3个，如集群没开启专有主节点，系统会自动选取3个数据节点作为主节点。物理机支持一机多实例，默认实例个数为内存/64，Raid方式支持raid1和raid0。
- 专有主节点：默认机型是云盘型、节点数量3个。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：默认机型是云盘型，如果开启，节点数量不少于2个。在大规模集群中，开启协调节点，可缓解数据节点压力，避免因CPU、I/O和内存等资源竞争，导致服务不可用。
- 冷数据节点：机型可选云盘型、本地SSD型。如果开启，节点数量不少于2个，如果集群中有大量不经常被访问的历史索引，可选择冷数据节点，实现冷热数据分离。
- 开启公网：如果您需要使用Kibana或者使用elasticsearch-SQL插件，需开启公网，绑定EIP。如果是【按日配置付费（试用）】的计费方式，创建时不支持开启公网，创建后，请到集群详情页绑定EIP。
- 绑定EIP：您可以选择已创建的EIP或者去EIP控制台先创建EIP再绑定，也可以选择“默认创建1M”EIP，我们会在创建集群的时候为您创建一个1M的EIP。所选EIP的数据中心与集群数据中心需一致。
- VPC网络：您可以使用同一数据中心内任一VPC来创建集群，若该数据中心无VPC，请您到VPC控制台创建，同时创建该VPC下EndPoint子网及与集群同一可用区内VPC子网。
- VPC子网：VPC子网是VPC中用于管理云主机的网络单元，您可以使用该VPC下与集群同一可用区内任一VPC子网来创建集群。
- EndPoint子网：EndPoint可以在您的VPC和其他金山云服务之间创建私有连接，使用KES服务必须指定EndPoint。您可以使用该VPC下任一EndPoint子网来创建集群。
- 所属项目：选择集群的归属项目。集群创建后，可在集群列表页【分配至项目】选择集群所属项目

4、确认及购买

确认配置详情后，点击【购买】，进入确认订单页面，然后预览集群配置。配置不符合预期时，可【返回】修改。

5、勾选我已阅读并同意《服务使用协议》

自动续费是可选项，包年包月的集群有使用期限，到期后会影响到服务使用，建议包年包月购买的集群勾选【账户余额足够时，资源到期后按月自动续费】。

6、付款并且订单创建成功后，在左侧导航栏选择Elasticsearch服务，进入集群列表页面，等待集群状态变为【运行中】，即可开始使用kes服务。

步骤三：访问集群

通过 Kibana 访问集群

KES 包含 Kibana 模块，用户可以通过公网访问集群对应的 Kibana 页面，进行数据的可视化查询分析，以及数据的管理操作。

入口

kibana入口位于集群详情页下面的可视化模块，使用kibana前请确认该集群已绑定EIP。

The screenshot shows the ElasticSearch cluster details page. The cluster ID is es-20200228133406. The basic information section includes cluster ID, data center (Guangzhou 1 Zone), version (6.8.4), and creation time. The network information section shows the VPC network and endpoint subnet. The configuration information section shows the data node configuration. In the visualization section, the kibana icon is highlighted with a red box.

访问

在Kibana控制台中，单击左侧导航栏的Dev Tools（开发工具），在Console中执行如下命令访问KES实例。

GET /

访问成功后，显示结果如下。

```
{
  "name": "10.0.0.248_0",
  "cluster_name": "es-20200228133406",
  "cluster_uuid": "F6F2wbGJRl1jqCetPqoSdg",
  "version": {
    "number": "6.8.4",
    "build_flavor": "default",
    "build_type": "tar",
    "build_hash": "bca0c8d",
    "build_date": "2019-10-16T06:19:49.319352Z",
    "build_snapshot": false,
    "lucene_version": "7.7.2",
    "minimum_wire_compatibility_version": "5.6.0",
    "minimum_index_compatibility_version": "5.0.0"
  },
  "tagline": "You Know, for Search"
}
```

步骤四：业务查询

创建索引

在Kibana控制台中，创建一个名称为school的索引，类型为student，包含4个字段，分别是name、address、age和date。关于字段类型说明，可查看官方文档 [Field Datatypes](#)

```

1 PUT /school
2 {
3   "settings":{
4     "number_of_shards": 3,
5     "number_of_replicas": 1
6   },
7   "mappings":{
8     "student":{
9       "properties":{
10        "name":{
11          "type":"text"
12        },
13        "address":{
14          "type":"keyword"
15        },
16        "age":{
17          "type":"integer"
18        },
19        "date":{
20          "type":"date",
21          "format":"yyyy-MM-dd HH:mm:ss"
22        }
23      }
24    }
25  }
26 }

```

```

PUT /school
{
  "settings":{
    "number_of_shards": 3,
    "number_of_replicas": 1
  },
  "mappings":{
    "student":{
      "properties":{
        "name":{
          "type":"text"
        },
        "address":{
          "type":"keyword"
        },
        "age":{
          "type":"integer"
        },
        "date":{
          "type":"date",
          "format":"yyyy-MM-dd HH:mm:ss"
        }
      }
    }
  }
}

```

创建文档并插入数据

如果返回显示"errors" : false, 说明数据插入成功。

```

1 POST /school/student/_bulk
2 [{"index":{}}]
3 [{"name":"张三","address":"山东济南","age":"22","date":"2020-01-19 08:10:00"}]
4 [{"index":{}}]
5 [{"name":"李四","address":"北京海淀","age":"21","date":"2020-02-19 09:00:00"}]
6 [{"index":{}}]
7 [{"name":"王五","address":"吉林长春","age":"23","date":"2020-02-19 10:00:00"}]
8

```

```

1 {
2   "errors": false,
3   "items": [
4     {
5       "index": {
6         "_index": "school",
7         "_type": "student",
8         "_id": "PtV813A8Z40je_1_1jw7",
9         "_version": 1,
10        "result": "created",
11        "_shards": {
12          "total": 2,
13          "successful": 2,
14          "failed": 0
15        }
16      },
17      "_seq_no": 0,
18      "_primary_term": 1,
19      "status": 201
20    }
21  ],
22  {
23    "index": {
24      "_index": "school",
25      "_type": "student",
26      "_id": "P9V813A8Z40je_1_1jw7",
27      "_version": 1,
28      "result": "created",
29      "_shards": {
30        "total": 2,
31        "successful": 2,

```

```

POST /school/student/_bulk
{"index":{}}
{"name":"张三","address":"山东济南","age":"22","date":"2020-01-19 08:10:00"}
{"index":{}}
{"name":"李四","address":"北京海淀","age":"21","date":"2020-02-19 09:00:00"}
{"index":{}}
{"name":"王五","address":"吉林长春","age":"23","date":"2020-02-19 10:00:00"}

```

查询数据

```
Console Search Profiler Grok Debugger

1 GET /school/_search
2 {
3   "query": {
4     "match": {
5       "address": "北京海淀"
6     }
7   }
8 }

1- {
2   "took": 3,
3   "timed_out": false,
4   "_shards": {
5     "total": 3,
6     "successful": 3,
7     "skipped": 0,
8     "failed": 0
9   },
10  "hits": {
11    "total": 1,
12    "max_score": 0.6931472,
13    "hits": [
14      {
15        "_index": "school",
16        "_type": "student",
17        "_id": "P9V8i3ABZ40je_1_ljw7",
18        "_score": 0.6931472,
19        "_source": {
20          "name": "李四",
21          "address": "北京海淀",
22          "age": "21",
23          "date": "2020-02-19 09:00:00"
24        }
25      }
26    ]
27  }
28 }
29

GET /school/_search
{
  "query": {
    "match": {
      "address": "北京海淀"
    }
  }
}
```