云搜索服务日志分析场景实战

1 任务介绍

本次任务,通过对网站 <u>apache 日志</u>的分析,介绍华为<u>云搜索服务</u>集群创建、使用,并向大家展示内置图形化工具 Kibana 的效果。

2 任务执行

- 2.1 准备环境
- 2.1.1 创建虚拟私有云

步骤 1 从华为云官网进入虚拟私有云产品页。



步骤 2 点击"立即购买"登录华为云控制台。



步骤 3 选择左上角区域为"华北-北京一"。



步骤 4 在 "网络控制台"页面右上角单击"创建虚拟私有云",已创建过"vpc-css-demo"的可以跳到 2.1.2。



步骤 5 根据界面提示配置虚拟私有云参数后,单击"立即创建"。

- "区域"保持默认为"华北-北京一"。
- "名称"配置为"vpc-css-demo",也可以按规范命名。
- "网段"保持默认值。
- "标签"保持默认值,不填写。
- "可用分区"保持默认值"可用区 2"。
- "子网名称"配置为"subnet-css-demo",也可以按规范命名。
- "子网网段"保持默认值。
- "高级配置"选择默认配置。

参数配置完成后,如图所示。



步骤 6 确认安全组 Sys-default 的出入方向都允许所有协议和所有端口访问。





步骤 7 若没有上述规则,请添加。否则,略过此步骤。



2.1.2 购买弹性云服务器

步骤 1 在从华为云官网进入弹性云服务器产品页。



步骤 2 点击"立即购买"进入购买页面。

弹性云服务器 ECS 🕨

弹性云服务器(Elastic Cloud Server)是一种可随时自助获取、可弹性伸缩的云服务器,帮助用户打造可靠、安全、灵活、高效的应用环境

三年低至5折,多种配置可选 了解价格详情→



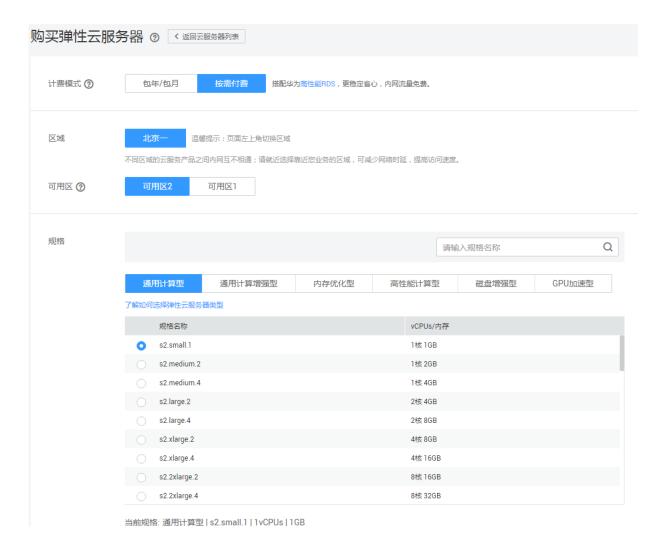
价格计算器

帮助文档 最佳实践

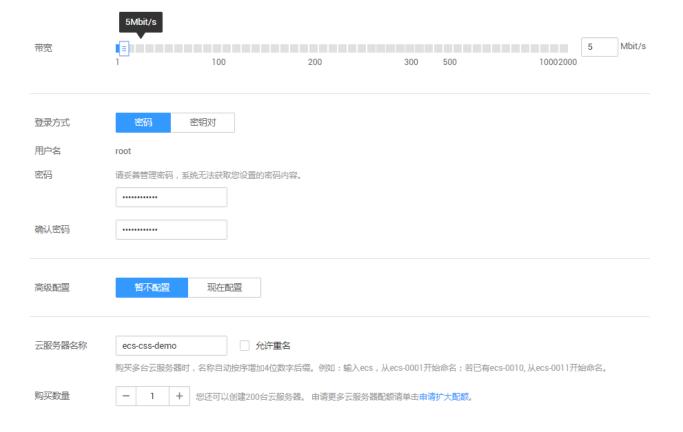
步骤 3 根据界面提示配置参数后,单击"立即购买"。已购买过可以跳到 2.1.3。

- "计费模式"选择"按需付费"。
- "区域"保持默认值"北京一"。
- "可用区"保持默认值"可用区 2"
- "规格" 这里选择"诵用计算性 s2.small.1"。
- "镜像"选择"公共镜像 EulerOS EulerOS 2.2 64bit(40GB)"。
- "磁盘 系统盘"选择"普通 IO 50GB",其他保持默认。
- "自动备份"保持默认不勾选。
- "虚拟私有云"选择我们之前创建的"vpc-css-demo"。
- "安全组"选择"Sys-default"。
- "网卡 主网卡"选择"subnet-css-demo",其他保持默认。
- "弹性 IP" 选择"现在购买"。
- "规格"保持默认选择"静态 BGP"。
- "带宽类型"保持默认选择"独享带宽"。
- "计费方式"保持默认选择"按带宽计费"。
- "带宽"选择"5Mbit/s"。
- "登录方式"选择"密码",并输入登录弹性云服务器的 root 账户密码。请牢记密码,后面会使用。
- "高级配置"保持默认"暂不配置"。
- "云服务器名称"输入"ecs-css-demo",不勾选允许重名。
- "购买数量"选择"1"。

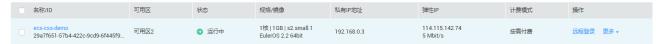
参数配置后,如下图:





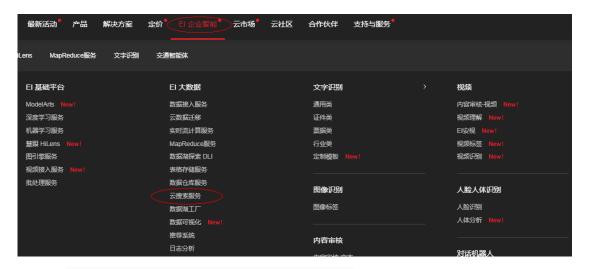


步骤 3点击"立即购买",在"规格确认"确认无误后,勾选免责声明,点击"提交申请",等待 3 分钟左后,弹性云服务器创建成功。



2.1.3 创建云搜索服务集群

步骤1 从华为云官网进入云搜索服务产品页。



步骤 2 点击"立即使用"按钮进入控制台页面。



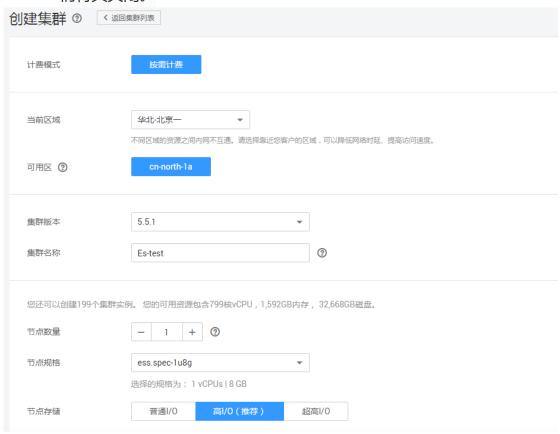
步骤 3 在"云搜索服务"页面右上角单击"创建集群"。之前创建过"Es-test"集群的,可以跳到 2.1.4。



步骤 4 根据界面提示配置参数后,单击"立即申请"。

- "当前区域"从下拉框选择"华北-北京一"
- "集群版本"从下拉框选择,是 Elasticsearch 的软件版本,选择 5.5.1。
- "集群名称"按规范命名配置,可以输入 Es-test。
- "节点数量"自由选择配置,生产集群一般建议3个以上。这里选择1。
- "节点规格"自由选择,生产集群一般建议 ess.spec-2u16g 以上。这里我们选择 ess.spec-1u8g。
- "节点存储"自由选择,这里选择"高 I/O(推荐)"。

- "节点存储容量"可以在允许范围内自由选择。这里输入 40GB。
- "虚拟私有云" VPC 即虚拟私有云,是通过逻辑方式进行网络隔离,提供安全、隔离的网络环境。选择刚才创建的"vpc-css-demo"。
- "子网" : 通过子网提供与其他网络隔离的、可以独享的网络资源,以提高网络安全。选择"subnet-css-demo"。
- "安全组":安全组是一个逻辑上的分组,为同一个 VPC 内具有相同安全保护需求 并相互信任的弹性云服务器提供访问策略。选择"Sys-default"。
- "集群快照":为了增强数据可靠性保障,系统默认开启自动快照功能。这里,我们将其关闭。



节点存储容量	_ 40 + GB ②	
虚拟私有云 ②	vpc-css-demo ▼ C 查看虚拟私有云	
子网 ②	subnet-css-demo(192.168.0.0/24) ▼	
安全组 ⑦	Sys-default ▼ C 查看安全组	
集群快照	启用自动快照 ————————————————————————————————————	

- 步骤 5 在规格确认界面确认完成后,单击"提交申请"开始创建集群。
- 步骤 6 单击"返回集群列表",系统将跳转到"集群管理"页面,您创建的集群将展现在集群列表中。且集群状态为"创建中"。

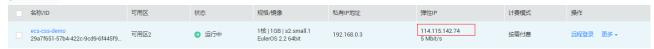


步骤 7 耐心等待十几分钟,创建成功后集群状态会变为"可用"。



2.1.4 安装 java 和 logstash

步骤 1 使用 ssh 工具(比如 putty)登录弹性云服务器,连接 IP 为弹性云服务器列表中的弹性 IP。



步骤 2 在弹性云服务器 ssh 命令行,运行下列命令。安装 java 环境,java -version 测试是 否安装成功。输出如下显示安装成功。

[root@ecs-css-demo ~]# /usr/bin/yum install java [root@ecs-css-demo ~]# java -version openjdk version "1.8.0_171" OpenJDK Runtime Environment (build 1.8.0_171-b10) OpenJDK 64-Bit Server VM (build 25.171-b10, mixed mode)

步骤 3 下载软件包 css_apache_demo.tar.gz 到本地,并通过 sftp 工具(比如 winscp)上传到弹性云服务器某个目录,这里以/root 目录为例。

步骤 4 在弹性云服务器 ssh 命令行,运行下列命令。解压软件包 css_apache_demo.tar.gz,并进入 css_apache_demo 目录,得到目录 demo_config、demo_logs 和压缩包 logstash-5.5.1.tar.gz。

[root@ecs-css-demo ~]# tar -zxf css_apache_demo.tar.gz [root@ecs-css-demo ~]# cd /root/css_apache_demo/;ll total 91164 drwx----- 2 3397668 1049089 4096 Jul 19 13:44 demo_config drwx----- 2 3397668 1049089 4096 Jul 19 14:00 demo_logs -rw----- 1 3397668 1049089 93242602 Jul 17 10:14 logstash-5.5.1.tar.gz

步骤 5 在弹性云服务器 ssh 命令行,运行下列命令。解压 logstash-5.5.1.tar.gz 得到目录 logstash-5.5.1,并更改用户组为 root。

[root@ecs-css-demo css_apache_demo]# tar -zxf logstash-5.5.1.tar.gz; chown -R root.root/root/css_apache_demo

[root@ecs-css-demo css_apache_demo]# II

total 91164

 drwx----- 2 root root
 4096 Jul 19 13:44 demo_config

 drwx----- 2 root root
 4096 Jul 19 14:00 demo_logs

drwx----- 11 root root 4096 Jul 19 10:39 logstash-5.5.1

-rw-r--r-- 1 root root 93242602 Jul 17 10:14 logstash-5.5.1.tar.gz

2.2 准备数据

2.2.1 理解数据

解压 css_apache_demo.tar.gz 后,在 demo_logs/apache_logs 文件中,存放的是某网站产生的 apache 格式日志。一条日志中,包含 ip 信息、timestamp 信息、request 信息、status 返回码信息、bytes 请求流量信息、referer 信息、agentinfo 访问客户端信息。如下图:



2.2.2 导入数据

步骤 1 在弹性云服务器 ssh 命令行,运行下列命令。输入编辑 demo_config/apache_logstash.conf 文件,修改<mark>第 3 行</mark>日志文件路径、<mark>第 37 行</mark> Elasticsearch 地址(云搜索服务集群列表中的内网访问地址)、 第 39 行 template 路径。 [root@ecs-css-demo css_apache_demo]# vi /root/css_apache_demo/demo_config/apache_logstash.conf

修改后如图:

```
1 input {
2  file {
3    path => "/root/css_apache_demo/demo_logs/apache_logs"
4    start_position => "beginning"
5  }
6 }
```

```
elasticsearch {

hosts => "192.168.0.3:9200"

index => "apache_elastic_example"

template => "/root/css_apache_demo/demo_config/apache_template.json"

template_name => "apache_elastic_example"

template_overwrite => true

template_overwrite => true

}
```

步骤 2 在弹性云服务器 ssh 命令行,运行下列命令。启动 logstash 开始导入数据,至小圆点不再输出后,数据导入完成。但 logstash 仍然运行监听日志文件。

[root@ecs-css-demo css_apache_demo]# /root/css_apache_demo/logstash-5.5.1/bin/logstash -f /root/css_apache_demo/demo_config/apache_logstash.conf如下图:

```
[2018-07-19115:25:13,545][INFO ][logstash.agent ] Pipetine main Started
[2018-07-19T15:25:13,711][INFO ][logstash.agent ] Successfully started Logstash API endpoint {:port=>9601}
```

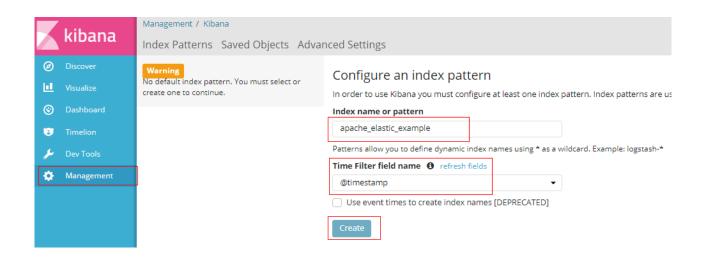
2.3 搜索分析

2.3.1 日志搜索

步骤 1 从云搜索服务控制台=>集群列表的 "Kibana" 链接进入可视化界面。



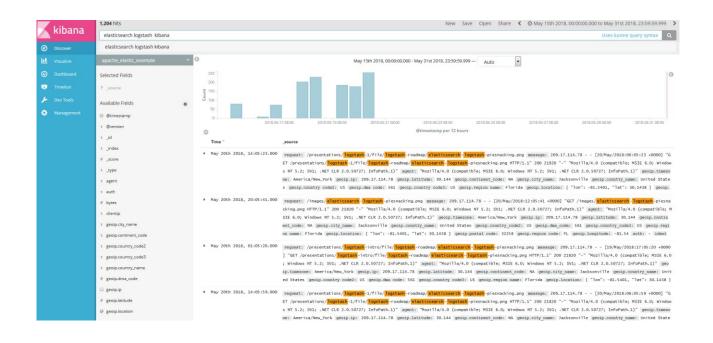
步骤 2 在可视化界面 "Management" => "Index Patterns" 页签 ,输入 "apache_elastic_example" 作为 index name , 选择@timestamp 作为刷新字段 , 点击 create 按钮创建 index patten。



步骤 3 可视化界面 "Discover" 页签,点击右上角的时间窗口选择器,将时间窗口定义在 "2018-05-15" 到 "2018-05-31"。

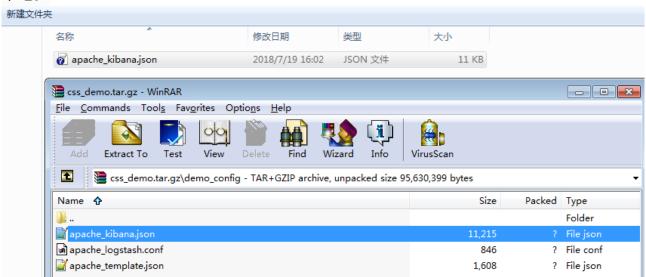


步骤 4 现在你可以像普通搜索引擎一样进行搜索。比如:在搜索框中输入搜索字符串 "elasticsearch logstash kibana",在任何字段信息中,只要任意有匹配的日志信息都找 出来了。这样,就可以看到,有没有用户在访问和"elasticsearch logstash kibana"相关的日志信息。



2.3.2 日志分析

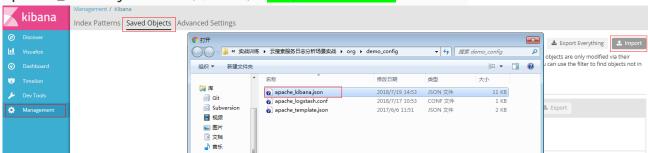
步骤 1 在本机上,从软件包 css_apache_demo.tar.gz 中解压 apache_kibana.json 文件到本地。



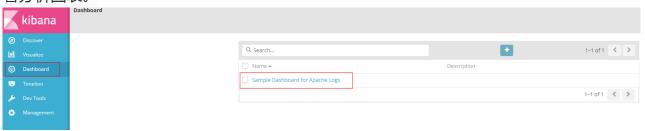
步骤 2 在本机用文本编辑器打开 apache_kibana.json 文件,替换 37 行 "UserName" 为自己的华为云用户名,保存。

```
📙 apache_kibana. json
31
     · } ,
32
     · {
33
     "_id": "Apache-Dashboard",
      "_type": "visualization",
34
35
        source": {
36
        "title": "Overview",
        "visState": "{\"type\":\"markdown\",\"params\":{\"markdown\":\"####
37
        Apache日志可视化分析仪表盘样例 [JserName]\\n在网站的运营过程中,网站访问者的各种信息
        对于网站运营者来说,是非常有价值的信息。而网站日志则是这些信息的来源,其记录着web服
        务器接收处理请求等各种原始信息。\\n\\n本实战场景,以一个网站的apache日志作为举例,
        体验日志可视化分析场景中[云搜案服务](https://www.huaweisloud.com/product/eg.html)的
        应用。\\n\\n**分析包括:**\\n* 网站独立访问者的数量 \\n* 访问热度地图\\n*
        不同国家的访问设备比例关系 \\n*
        不回国家的访问操作系统比例关系\\n\\n\"},\"aggs\":[],\"listeners\":{}}",
```

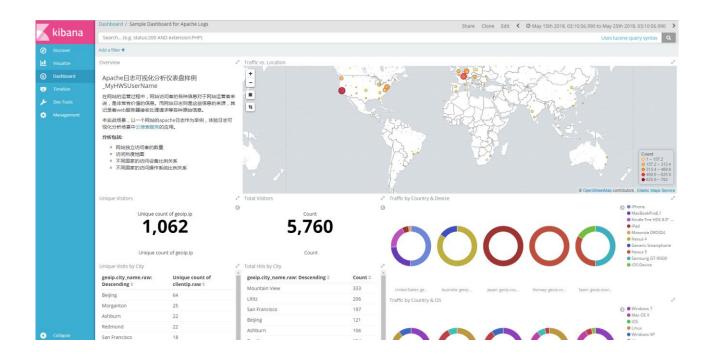
步骤 3 在可视化界面 "Management – Saved Objects" 页签,点击 import 按钮,选择 apache_kibana.json 文件,并点击 "Yes, overwrite all" 导入。



步骤 4 在可视化界面 "Dashboard" , 点击 "Sample Dashboard for Apache Logs" 查看分析图表。



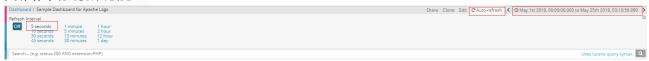
如图:



至此,关于网站的访问情况,已经从多个维度,图形化展现了分析结果。在图形化界面创建这里,我们直接使用了模板导入,您也可以在了解了 Kibana 后,采用托拉拽的方式,自己创建适合自己的图形化界面。

2.3.3 近实时分析

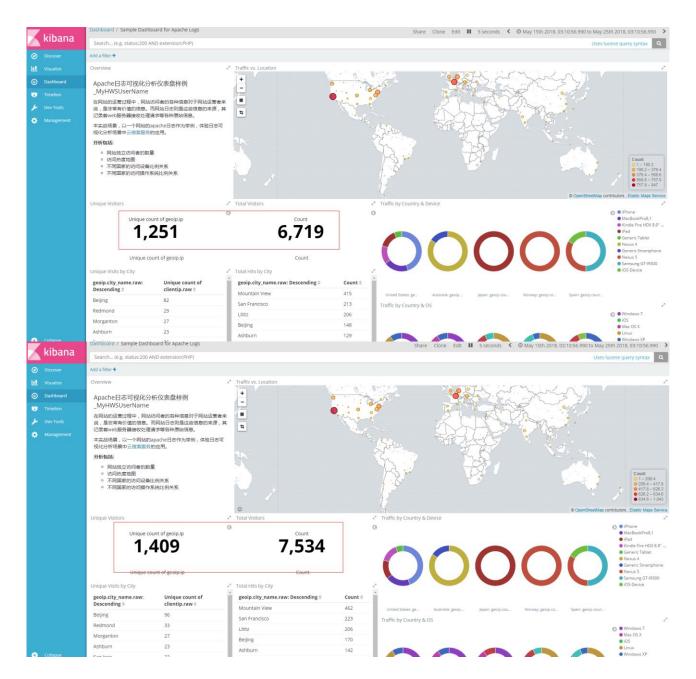
步骤 1 点击 Dashboard 界面右上角的时间窗口 单击 Auto-refresh ,并选择刷新间隔(5s) , 开启自动刷新功能。



步骤 2 ssh 登录一个新的云服务器命令行,运行下面命令,模拟写入增量日志到 demo_logs/apache_logs 文件。

[root@ecs-css-demo css_apache_demo]# while read line; do echo \$line >> /root/css_apache_demo/demo_logs/apache_logs; usleep 20000;done < /root/css_apache_demo/demo_logs/delta_apache_logs

回到可视化界面,可以看到可视化界面的数据开始动态刷新:



3 任务打卡

在 kibana 图形化页面, 截图带有自己华为云用户名的分析结果, 完成打卡。

