实训任务五：基于 Linux Syslog 数据分析洞察操作系统

一、目的

1、了解 Linux Syslog；

2、实现对 Linux Syslog 的数据分析。

二、学时

4 学时

三、类型

综合型

扫码看实训演示

扫码看任务步骤

四、需求

1、硬件

每人配备计算机 1 台，不低于双核 CPU、8G 内存、500GB 硬盘。

每小组配备服务器 1 台。

2、软件

Windows 操作系统，安装 puTTY 管理终端软件，安装 VMware ESXi 控制台软件。

服务器安装 VMware ESXi 7.0。

3、网络

计算机、服务器、虚拟主机使用固定 IP 地址接入局域网，并支持对互联网的访问。

4、工具

无。

五、任务要求

1、完成 Filebeat 的安装与配置；

2、实现 Linux Syslog 分析以达到对 Linux 操作系统洞察。

六、考核要求

1、提交《Linux Syslog 日志分析报告》；

2、提交 Linux Syslog 日志可视化分析成果截图/演示视频。

七、任务步骤

本任务需采集如表 1-1 所示服务器的 Linux Syslog 数据。

表 5-1 虚拟机列表

序号	虚拟机名称
1	YWSX-10.10.2.162-CentOS8X64-Cacti
2	YWSX-10.10.2.163-CentOS8X64-DNS

任务 1：部署 Syslog

Syslog 是 Linux 系统默认的日志守护进程，接受来自系统的各种功能的信息，帮助洞察系统运行情况。

步骤 01：CentOS 系统已默认安装 Syslog，其对应服务为 rsyslog，若未安装执行如下命令即可。

参考命令：

```
yum install syslog-ng -y
```

步骤 02：修改 Syslog 配置文件

syslog 配置路径为/etc/rsyslog.conf，配置内容如下。

配置文件：

```
*.info;mail.none;authpriv.none;cron.none        /var/log/messages
authpriv.*                                      /var/log/secure
mail.*                                          /var/log/maillog
cron.*                                          /var/log/cron
```

步骤 03：启动 Syslog

参考命令：

```
systemctl reload rsyslog
```

任务 2：使用 Filebeat 采集 Linux Syslog 数据

Filebeat 是用于转发和集中日志数据的轻量级采集器。Filebeat 作为采集器安装在服务器上，监视指定的日志文件或位置，收集日志事件，并将它们转发到 Elasticsearch 或 Log stash 进行处理。

步骤 01：获取 Filebeat 程序

从 https://www.elastic.co/cn/downloads/beats/filebeat 获取安装的 RPM 64-BIT 文件，并上传至 YWSX-10.10.2.162-CentOS8X64-Cacti 虚拟机的/opt 目录。

步骤 02：安装配置 Filebeat

参考命令：

```
rpm -ivh filebeat-7.9.0-x86_64.rpm
```

步骤 03：修改 Filebeat 配置文件

Filebeat 配置文件为/etc/filebeat/filebeat.yml。

参考命令：

```
rpm -ivh filebeat-7.9.0-x86_64.rpm
```

配置文件：

```
filebeat.inputs:
- type: log
  enabled: true
  paths:
    - /var/log/cron
    - /var/log/messages
    - /var/log/secure
    - /var/log/maillog
fields:
    indextype: syslog-cacti-10.10.2.162
  fields_under_root: true

output.Logstash:
  enabled: true
  hosts: ["10.10.2.155:5044"]
  topic: syslog-10.12.24.246
```

步骤 04：启动 Filebeat

参考命令：

```
filebeat -e -c /etc/filebeat/filebeat.yml
```

任务 3：使用 Logstash 清洗与格式化数据

步骤 01：配置 Logstash

创建 Logstash 配置文件，配置文件路径为/etc/logstash/conf.d/logstash_syslog.conf，配置文件内容如下。

配置文件：

```
input {
      beats {
      port => 5044
   }
}

filter {
      grok {
      # 正则匹配
            match => { "message" => "%{SYSLOGTIMESTAMP:message_timestamp} %{SYSLOG
HOST:hostname} %{DATA:message_program}(?:\[%{POSINT:message_pid}\])?: %{GREEDYDATA:m
essage_content}" }
            # 如果成功匹配，则加一个匹配标签
            add_tag => [ "matched" ]
      }
      # 如果标签不存在则丢弃这条消息
      if ("matched" not in  [tags]) {
            drop {}
      }
      mutate {
            # 移除不需要的字段
            remove_field => "agent"
            remove_field => "host"
            remove_field => "@version"
            remove_field => "ecs"
            remove_field => "input"
      }
      ruby {
```

```
            path => "/etc/logstash/script/syslog.rb"
        }
}

output {
        if [indextype]== "syslog-cacti-10.10.2.162" {
            elasticsearch{
                hosts => ["10.10.2.152:9200"]
                index => "syslog-cacti-10.10.2.162-%{+YYYY.MM.dd}"
            }
        }
}
```

步骤 02：撰写清洗脚本

创建/etc/logstash/script/syslog.rb，并编辑内容。

配置文件：

```
require 'rubygems'
require 'json'
require 'time'
def register(params)

end

def filter(event)
        begin
            log = event.get('log')
            source = log["file"]["path"]
            index_type=event.get("indextype").split("-")
            #增加索引
            if source.include?("message")
                index_names = ["slg","msg",index_type[1]]
                event.set('index_array',index_names)
            elsif source.include?("cron")
                index_names = ["slg","cro",index_type[1]]
                event.set('index_array',index_names)
            elsif source.include?("maillog")
                index_names = ["slg","mai",index_type[1]]
                event.set('index_array',index_names)
            else
                index_names = ["slg","sec",index_type[1]]
                event.set('index_array',index_names)
            end
            #时间处理
            local_time_e = event.get('message_timestamp')
            local_times = local_time_e.split(/[A-Z]/)
            local_time_a = local_times[0]+" "+local_times[1]
            local_time_b = local_time_a.split(/\./)
            local_time_c = local_time_b[0]
            local_time_d = local_time_c.to_s
            local_time = Time.parse(local_time_d)
            #local_timesss = DateTime.strftime(local_times,"%Y-%m-%d %H:%M:%S")
            event.set("local_time", local_time)
            event.set('[local_time_arr][year]', local_time.year)
            if(local_time.month.to_s.length==1)
                event.set('[local_time_arr][month]', "0"+local_time.month.to_s)
            else
```

```
            event.set('[local_time_arr][month]', local_time.month.to_s)
        end
        if(local_time.day.to_s.length==1)
            event.set('[local_time_arr][day]', "0"+local_time.day.to_s)
        else
            event.set('[local_time_arr][day]', local_time.day.to_s)
        end
        event.set('[local_time_arr][wday]', local_time.wday)
        event.set('[local_time_arr][yday]', local_time.yday)

        if(local_time.hour.to_s.length==1)
            event.set('[local_time_arr][hour]', "0"+local_time.hour.to_s)
        else
            event.set('[local_time_arr][hour]', local_time.hour.to_s)
        end
        event.set('[local_time_arr][min]', local_time.min)
        event.set('[local_time_arr][sec]', local_time.sec)
    end
    return [event]
end
```

步骤 03：重启 Logstash

参考命令：

```
systemctl restart logstash
```

依照以上步骤，基于 YWSX-10.10.2.163-CentOS8X64-Bind 虚拟机实现 Syslog 分析。

任务 4：在 Kibana 上创建分析模型并进行分析

步骤 01：设计分析模型

表 5-2 Syslog 分析模型

序号	分析字段	分析模型	图表类型
1	log.file.path	日志来源总数	数字
2	message_program+llog.file.path	message 日志程序占比	饼状图
3	message_program+ log.file.path	cron 日志程序占比	饼状图
4	message_program+ log.file.path	security 日志程序占比	饼状图
5	local_time_arr.hour+log.file.path	message 日志增量时间变化趋势	折线图
6	local_time_arr.hour+log.file.path	cron 日志增量时间变化趋势	折线图
7	local_time_arr.hour+log.file.path	security 日志增量时间变化趋势	折线图
8	local_time_arr.hour+message_program	sshd 程序增量时间变化趋势	折线图
9	local_time_arr.hour+message_program	systemd 程序增量时间变化趋势	折线图
10	message_content	程序内容标签云	标签云

步骤 02：可视化分析

依据表 5-2 分型模型实现图表可视化，如图 5-1 所示。

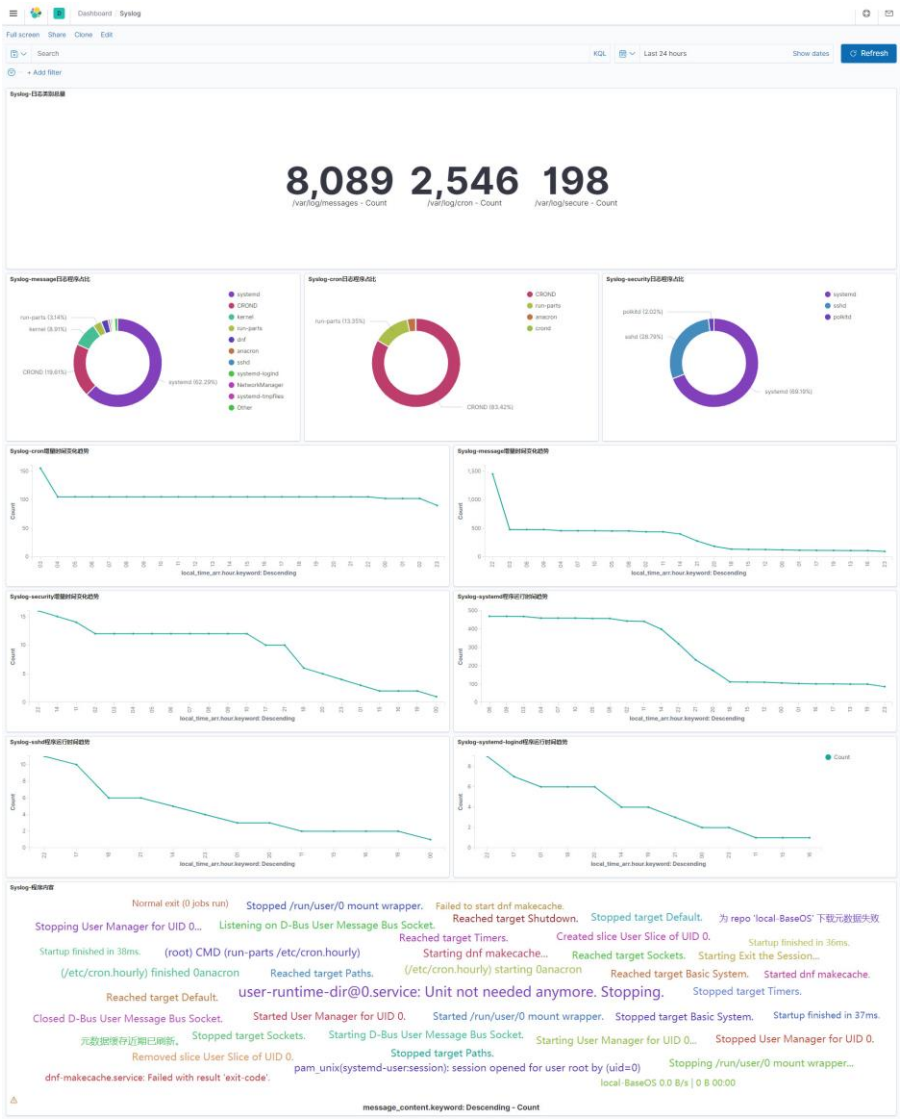

图 5-1 Syslog 数据分析

八、实验思考

1、了解 Filebeat。

（1）FileBeat 工作原理是什么？

（2）FileBeat 配置文件？

2、深入理解 Linux Syslog。

（1）什么是 Syslog?？

（2）Syslog 数据格式是什么？

（3）Syslog 配置函数有哪些？

3、深入洞察 Linux 操作系统。

（1）Linux 操作系统运行状态评价指标有哪些？

（2）Linux 操作系统可从哪些角度进行优化？