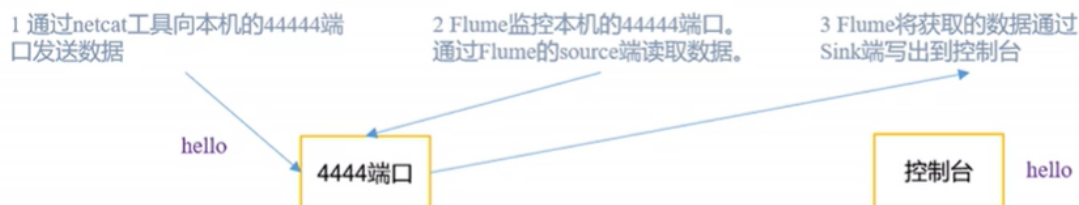


Flume3 监控端口数据

案例需求

使用Flume 监听一个端口，收集该端口数据，并打印到控制台

需求分析



实现步骤

- 安装 telnet 工具

将 rpm 软件包 (xinetd-2.3.14-40.el6.x86_64.rpm 、 telnet-0.17-48.el6.x86_64.rpm 和telnet-server-0.17-48.el6.x86_64.rpm)入/opt/software 文件夹下面。执行 RPM 软件包安装命令：

```
[atguigu@hadoop102 software]$ sudo rpm -ivh
xinetd-2.3.14-40.el6.x86_64.rpm
[atguigu@hadoop102 software]$ sudo rpm -ivh
telnet-0.17-48.el6.x86_64.rpm
[atguigu@hadoop102 software]$ sudo rpm -ivh
telnet-server-0.17-48.el6.x86_64.rpm
```

- 判断 4444端口是否被占用

```
[atguigu@hadoop102 flume-telnet]$ sudo netstat -tunlp | grep 44444
```

功能描述：netstat 命令是一个监控 TCP/IP 网络的非常有用的工具，它可以显示路由表、实际的网络连接以及每一个网络接口设备的状态信息。

基本语法：netstat [选项]

选项参数：

- -t 或--tcp：显示 TCP 传输协议的连线状况；
- -u 或--udp：显示 UDP 传输协议的连线状况；
- -n 或--numeric：直接使用 ip 地址，而不通过域名服务器；
- -l 或--listening：显示监控中的服务器的 Socket；
- -p 或--programs：显示正在使用 Socket 的程序识别码和程序名称；

- 创建 Flume Agent 配置文件 flume-telnet-logger.conf

在 flume 目录下创建 job 文件夹并进入 job 文件夹。

```
[atguigu@hadoop102 flume]$ mkdir job
[atguigu@hadoop102 flume]$ cd job/
```

在 job 文件夹下创建 Flume Agent 配置文件 flume-telnet-logger.conf。

```
[atguigu@hadoop102 job]$ touch flume-telnet-logger.conf
```

在 flume-telnet-logger.conf 文件中添加如下内容。

```
[atguigu@hadoop102 job]$ vim flume-telnet-logger.conf
```

添加内容如下：

```
# Name the components on this agent
a1.sources = r1
a1.sinks = k1
a1.channels = c1
# Describe/configure the source
a1.sources.r1.type = netcat
a1.sources.r1.bind = localhost
a1.sources.r1.port = 44444
# Describe the sink
a1.sinks.k1.type = logger
# Use a channel which buffers events in memory
a1.channels.c1.type = memory
a1.channels.c1.capacity = 1000
a1.channels.c1.transactionCapacity = 100
# Bind the source and sink to the channel
a1.sources.r1.channels = c1
a1.sinks.k1.channel = c1
```

配置文档解释：

```
# Name the components on this agent    #a1:表示agent的名称
a1.sources = r1        #r1: 表示a1的输入源
a1.sinks = k1          #k1: 表示a1的输出目的地
a1.channels = c1       #c1: 表示a1的缓冲区
# Describe/configure the source
a1.sources.r1.type = netcat  #表示a1的输入类型元类型为netcat端口类型
a1.sources.r1.bind = localhost #表示a1的监听的主机
a1.sources.r1.port = 44444  #表示a1的监听的端口号
# Describe the sink
a1.sinks.k1.type = logger  #表示a1的输出目的地是控制台logger类型
# Use a channel which buffers events in memory
a1.channels.c1.type = memory # 表示a1的channel类型是memory内存型
a1.channels.c1.capacity = 1000 # 表示a1的channel总容量1000个event
a1.channels.c1.transactionCapacity = 100 # 表示a1的channel传输的时候收集到了100
条event再提交
# Bind the source and sink to the channel
a1.sources.r1.channels = c1 #表示r1和c1连接起来
a1.sinks.k1.channel = c1  #表示将k1和c1连接起来
```

注：配置文件来源于官方手册 <http://flume.apache.org/FlumeUserGuide.html>

- 先开启 flume 监听端口

```
[atguigu@hadoop102 flume]$ bin/flume-ng agent --conf conf/ --name  
a1 --conf-file job/flume-telnet-logger.conf  
-Dflume.root.logger=INFO,console
```

参数说明:

- o `--conf conf/`: 表示配置文件存储在 `conf/` 目录
- o `--name a1`: 表示给 `agent` 起名为 `a1`
- o `--conf-file job/flume-telnet.conf`: `flume` 本次启动读取的配置文件是在 `job` 文件夹下的 `flume`
- o `-telnet.conf` 文件。
- o `-Dflume.root.logger==INFO,console`: `-D` 表示 `flume` 运行时动态修改 `flume.root.logger` 参数属性值, 并将控制台日志打印级别设置为 `INFO` 级别。日志级别包括: `log`、`info`、`warn`、`error`。

可以更改为:

```
bin/flume-ng agent -n $agent_name -c conf -f conf/flume-  
conf.properties.template
```

- 使用 `telnet` 工具向本机的 44444 端口发送内容

```
[atguigu@hadoop102 ~]$ telnet localhost 44444
```

```
[atguigu@hadoop102 ~]$ telnet localhost 44444  
Trying ::1...  
telnet: connect to address ::1: Connection refused  
Trying 127.0.0.1...  
Connected to localhost.  
Escape character is '^]'.  
hello  
OK
```

- 在 Flume 监听页面观察接收数据情况

```
2018-09-07 21:17:48,974 (SinkRunner-PollingRunner-DefaultSinkProcessor) [INFO - org.apache.flume.sink.  
Sink.process(LoggerSink.java:95)] Event: { headers:{} body: 68 65 6C 6C 6F 0D  
hello. }
```