

08-Flume

笔记本： 模块一配置笔记

创建时间： 2020/10/5 11:00

更新时间： 2021/7/18 22:59

作者： 封

八、Flume部署

0.软件包

涉及到的软件包，都存放在/opt/soft/

1.前置工作

请参照一前置工作操作

2.解压flume包

在master中，把apache-flume-1.9.0-bin.tar.gz包解压到指定位置,按照要求是否需要重命名

解压：tar -zxvf /opt/soft/apache-flume-1.9.0-bin.tar.gz -C /opt/soft/

从命名：mv /opt/soft/apache-flume-1.9.0-bin /opt/soft/flume

3.配置环境变量

在master中，vi /etc/profile

输入：

```
export FLUME_HOME=/opt/soft/flume
```

```
export PATH=$FLUME_HOME/bin:$PATH
```

```
source /etc/profile 更新环境变量
```

4.修改配置文件

需要修改flume-conf.properties，但在文件夹中没有该文件，只能复制一份。
进入/opt/soft/flume/conf 的路径并找到flume-conf.properties.template
执行cp ./flume-conf.properties.template ./flume-conf.properties

编辑flume-conf.properties

指令：vi ./flume-conf.properties

```
#定义agent里面的sources, channels, sinks
a1.sources = s1
a1.channels = c1
a1.sinks = k1
#配置sources数据源
a1.sources.s1.type = spooldir
a1.sources.s1.spoolDir=/opt/soft/hadoop/logs
#配置channels，做sources和channels绑定，指定channels用什么类型做传输。
a1.sources.s1.channels = c1
a1.channels.c1.type=memory
#配置sinks，落地。存在hdfs。
a1.sinks.k1.type = hdfs
a1.sinks.k1.hdfs.path=hdfs://master:9000/flume/%Y%m%d
a1.sinks.k1.hdfs.filePrefix=%Y%m%d-
a1.sinks.k1.hdfs.fileType=DataStream
a1.sinks.k1.channel=c1
a1.sinks.k1.hdfs.useLocalTimeStamp = true
```

该配置是把hadoop集群的日志采集并存储在HDFS上。

注意：不同的场景，需要对这个配置进行修改。

5.HDFS建立文件夹

启动hadoop集群，

在master执行hdfs dfs -mkdir /flume 是在HDFS建立文件夹

在master执行：flume-ng agent -n a1 -c conf -f /opt/soft/flume/conf/flume-conf.properties

启动成功后，就会看到采集过程。

6.查看HDFS内容

在web页面查看HDFS的文件系统。