



Formação Engenheiro de Dados

Estudo de Caso

Flume



Etapas

- ◆ Criar Diretório no Sistema de Arquivo do Linux

```
mkdir /home/cloudera/flume
```

- ◆ Criar Diretório no HDFS

```
hdfs dfs -ls /user/flume/
```

Criar Arquivo de Configuração

```
# componentes do agente
```

```
a1.sources = r1
```

```
a1.sinks = k1
```

```
a1.channels = c1
```

```
# Source
```

```
a1.sources.r1.type = spooldir
```

```
a1.sources.r1.channels = c1
```

```
a1.sources.r1.spoolDir = /home/cloudera/flume
```

```
#a1.sources.r1.fileHeader = true
```

```
a1.sources.r1.pollDelay = 5000
```

```
# Sink
```

```
a1.sinks = k1
```

```
a1.sinks.k1.type = hdfs
```

```
a1.sinks.k1.channel = c1
```

```
a1.sinks.k1.hdfs.path = /user/flume/
```

```
a1.sinks.k1.hdfs.fileType = DataStream
```

```
# Channel
```

```
a1.channels.c1.type = memory
```

```
a1.channels.c1.capacity = 1000
```

```
a1.channels.c1.transactionCapacity = 100
```

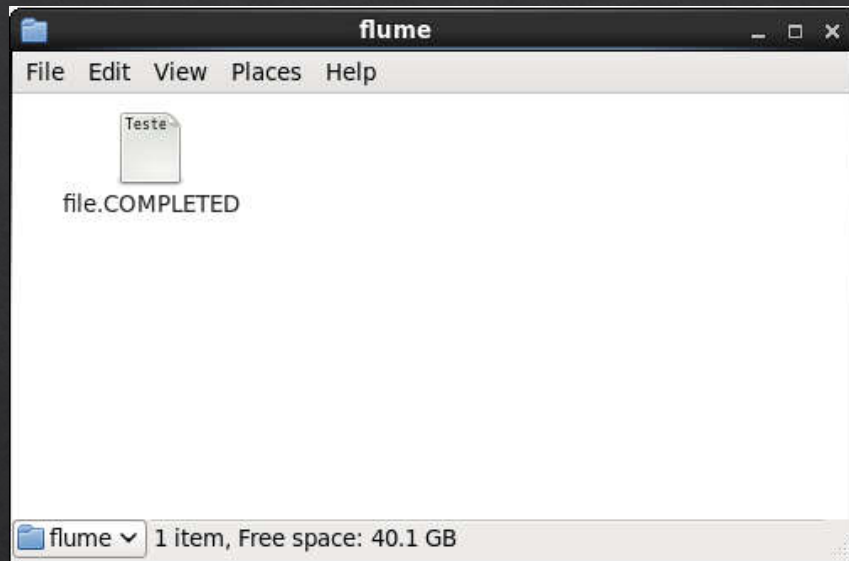
```
# Vincula source e sink com channel
```

```
a1.sources.r1.channels = c1
```

```
a1.sinks.k1.channel = c1
```


Executar o Agente

```
sudo flume-ng agent --conf conf --  
conf-file /home/cloudera/example.conf  
--name a1 -  
Dflume.root.logger=INFO,console
```

A screenshot of a terminal window titled "cloudera@quickstart:~". The terminal shows the command "hdfs dfs -ls /user/flume/" being executed. The output is "Found 1 items" followed by a line of file details: "-rw-r--r-- 1 root supergroup 6 2019-07-09 14:32 /user/flume/FlumeData.1562707910764". The prompt "[cloudera@quickstart ~]" is visible at the bottom.

Execução

As salvamos um arquivo em source, ele é renomeado e transferido para o sink