

# 자바쌤~!

<https://cafe.naver.com/javassem>

검색

63기-자바 &amp; 빅데이터 &gt;

## [ 플룸 ] 설치 및 개념



자바쌤 1:1 채팅

2020.06.22. 22:56 조회 25

댓글 0 URL 복사



### 1. 플럼의 개요

: HDFS의 수집 전용 소프트웨어

연속적으로  
생성되는  
데이터 스트림

수집 및 전송하고  
HDFS에 저장할 수  
있는 도구

▶ 로그 파일, 소셜  
미디어 데이터,  
이메일 메시지 등

분산환경에서 대량의 로그 데이터를 효과적으로 수집하여,  
합친 후 다른 곳으로 전송할 수 있는 신뢰할 수 있는 서비스

### 2. 플럼의 구성요소

소스(Source)

외부 데이터 소스에 설치되는 에이전트

싱크(Sink)

데이터 목적지에 설치되는 에이전트

#### 카페정보 나의활동



매니저 자바쌤  
since 2013.03.15.  
카페소개

새싹2단계

437

초대하기

즐거찾는 멤버

64명

게시판 구독수

3회

우리카페엽수

0회

카페 글쓰기

카페 채팅

#### 즐거찾는 게시판

전체글보기

4,034

#### KOSMO

63기-자바 &amp; 빅데이터

#### 팁

- 심심문제
- 기타등등
- JAVA
- Oracle/Mysql
- JSP
- WebUI
- 파이썬
- Spring
- 리눅스
- 하둡+eco
- 프로젝트

#### 공유해요

- 알고리즘도전
- 기술면접 준비하기
- 실존 면접
- 기회는 준비된자에게
- 프로젝트 팁

#### 쉬어가기

#### 청년아카데미

#### 전문가과정

공지사항  
카페회칙

카페탈퇴하기

## 채널(Channel)

소스와 싱크 간에 데이터를 받는 통로



Source : 원본  
Sink : 목적지로 대표적으로 HDFS  
Channel : 연결통로

## 2. 플럼의 구성요소

## (1) 소스

- 데이터로 수신한 후 채널로 전달하는 모듈
- 하나의 소스가 하나 이상의 채널로 데이터 전달 가능

## (2) 싱크

- 데이터를 HDFS, 로컬파일, 혹은 다른 플럼 에이전트 등에 전달하는 모듈
- 하나의 싱크는 오직 한 채널에서만 데이터를 전달 받을 수 있음



## 2. 플럼의 구성요소

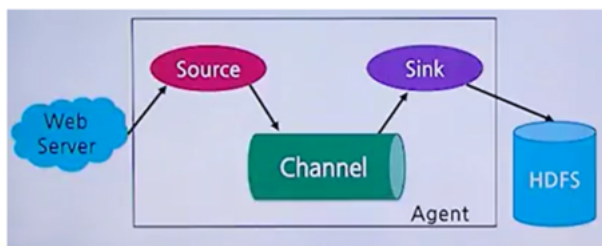
## (3) 채널

- 소스 데이터를 싱크로 전달하는 데이터 큐
- 소스의 속도와 싱크의 속도를 조절하는 일종의 버퍼
- 채널의 데이터는 기본적으로 메모리에 저장
- 네트워크 장애 시 데이터 유실을 방지하기 위해 디스크에 저장하는 설정 옵션 가능



## 3. 플럼의 구성방법

## (1) 기본적인 구성

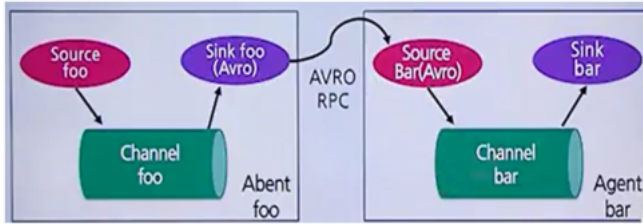


기본적인 에이전시는 하나의 소스에서 하나의 채널을 타고 하나의 싱크로 가는 것이다.



## 3. 플럼의 구성방법

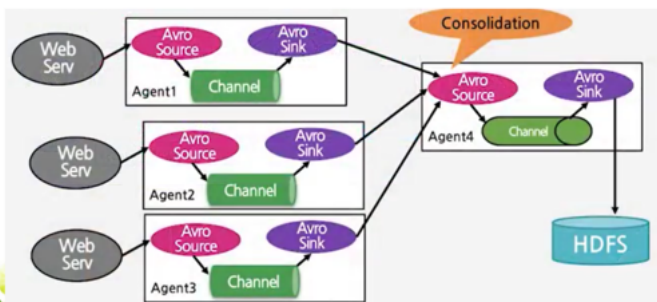
## (2) MultiAgent Flow 구성



하나의 싱크로 간 결과가 다른 곳의 데이터로 사용될 때 다른 에이전트의 소스로 보내지는 것이다.  
이 때 물리적으로 두 개의 에이전트는 네트워크 통신을 하게 되니깐 데이터 직렬화로 전송하게 되어 Avro 같은 역할을 하는 것이 필요하다.

## 3. 플럼의 구성방법

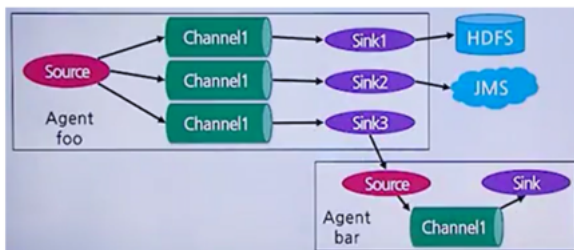
### (3) MultiAgent 통합구성



예를 들어 여러 웹 서버에서 로그데이터를 수집한다면 각각의 웹서버의 데이터를 받는 에이전트가 필요하고 이것들을 통합하는 에이전트가 필요하다.

## 3. 플럼의 구성방법

### (4) Multiplexing Flow 구성



소스 자체가 너무 중요한 자료가거나 활용도가 높은 자료인 경우 여러 곳에서 사용되기 위해 또는 HDFS로 보내면서 자료 자체를 백업한다면가 하여 2가지 이상의 작업을 하게 되면 채널을 여러개 두어 목적지가 되는 싱크를 여러개 하는 것이다.

## 4. 설치

### (1) dn01 노드의 root 계정에서 설치

```
[root@dn01 ~]# cd /tmp
[root@dn01 tmp]# wget
http://apache.mirror.cdnetworks.com/flume/stable/apache-
flume-1.9.0-bin.tar.gz
[root@dn01 tmp]# tar xzvf apache-flume-1.9.0-bin.tar.gz
[root@dn01 tmp]# mkdir -p /opt/flume/1.9.0
```



## 4. 설치

## (2) dn01 노드의 hadoop 계정에서 환경변수 설정

```
[hadoop@dn01 ~]$ vi ~/.bash_profile
```

```
#### FLUME 1.9.0 #####
export FLUME_HOME=/opt/flume/current
export PATH=$PATH:$FLUME_HOME/bin
#### FLUME 1.9.0 #####
```



## 4. 설치

### (3) dn01 노드의 hadoop 계정에서 Flume 설정

```
[hadoop@dn01 ~]$ source ~/.bash_profile
[hadoop@dn01 ~]$ cd $FLUME_HOME/conf
[hadoop@dn01 conf]$ cp flume-conf.properties.template
flume-conf.properties
```

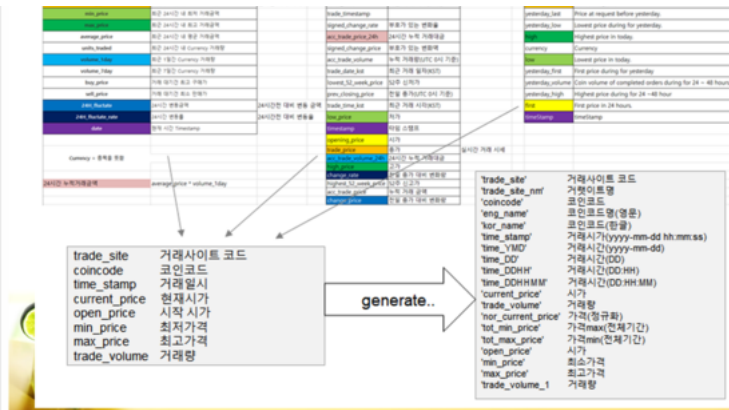
\* flume-conf.properties 파일의 기본값으로 사용하기 수정없음

## 5. 예제

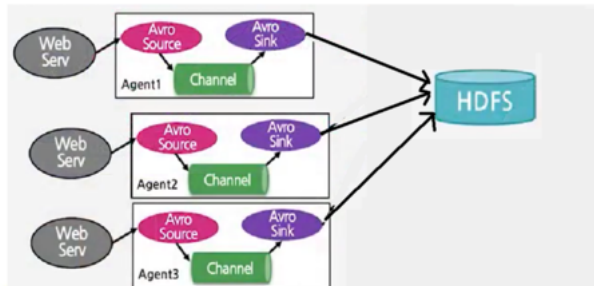


## 5. 예제

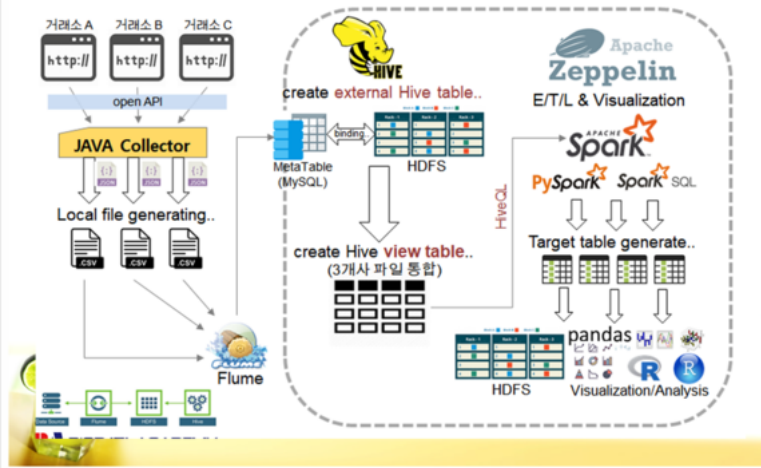
=비밀=		=일반비밀=		=공개정보=	
Key Name	Description	Key Name	Description	Key Name	Description
password	비밀번호	password	비밀번호	password	비밀번호
spring_priv	스프링 16진간 내 이력 가져올때	track_volume	가장 최근 가져올	volume	1천 volume of completed orders in 24 hours



## 5. 예제



## 5. 예제



작성된 태그가 없습니다.

태그달기

<input title="태그를 입력하세요." class="input\_text pcol2\_tagInputArea" type="text" value="" /> 취소 확인

. . . . .

댓글

댓글알림

김용선

댓글을 남겨보세요



등록



자바쌤~!    <https://cafe.naver.com/javassem>