**[Flume]**

1. Introduction

Apache Flume is a distributed, reliable, and available system

for efficiently collecting, aggregating and moving large amounts of log data

from many different sources to a centralized data store.

基本构件：Source——Channel——Sink

配置：属性名称.分级结构

|顶级：代理名称.不同组件名称.组件中不同组件名称.组件的类型和属性

（属性种类由类型决定）

2. 事务与可靠性

Flume通过两个独立的事务负责从Source到Channel和从Channel到Sink的事件传递。

只有事务中的所有事件全部传递到Channel且提交成功，Source才将文件标记为完成。

Channel到Sink的事务传递过程也类似。

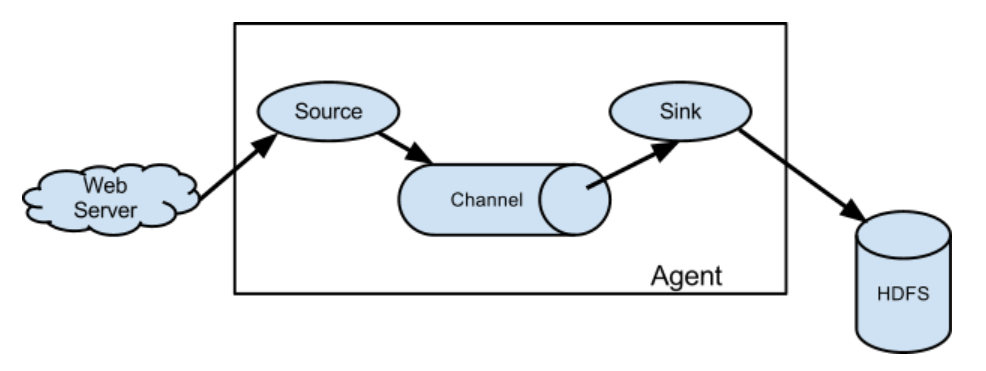
3. 扇出

扇出是指从一个Source向多个Channel和多个Sink传递事件。

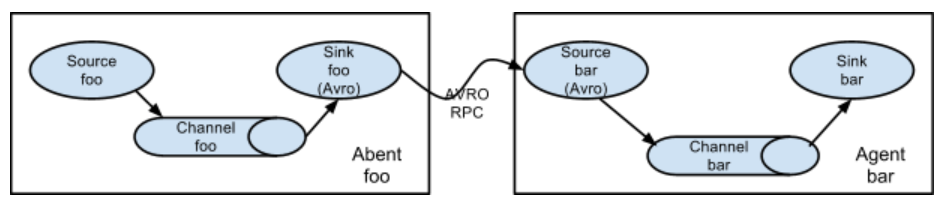
正常扇出流是向所有Channel复制事件，

但可以通过在Source上设置一个复用选择器实现不同事件不同Channel。

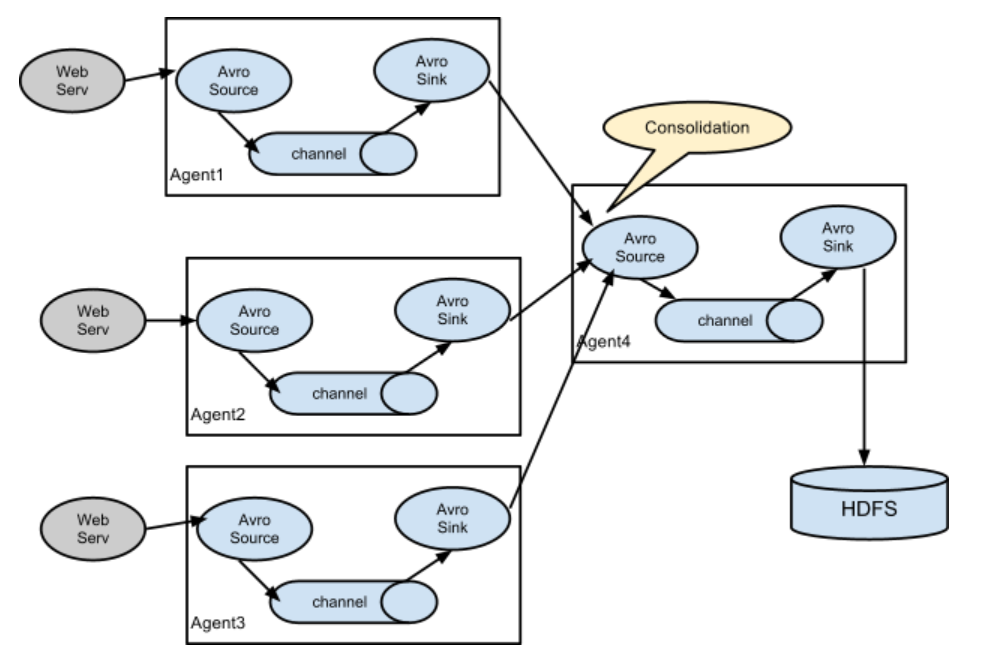
4. Data flow model



5. Mmulti-agent flow



6. Consolidation



7. Multiplexing the flow

