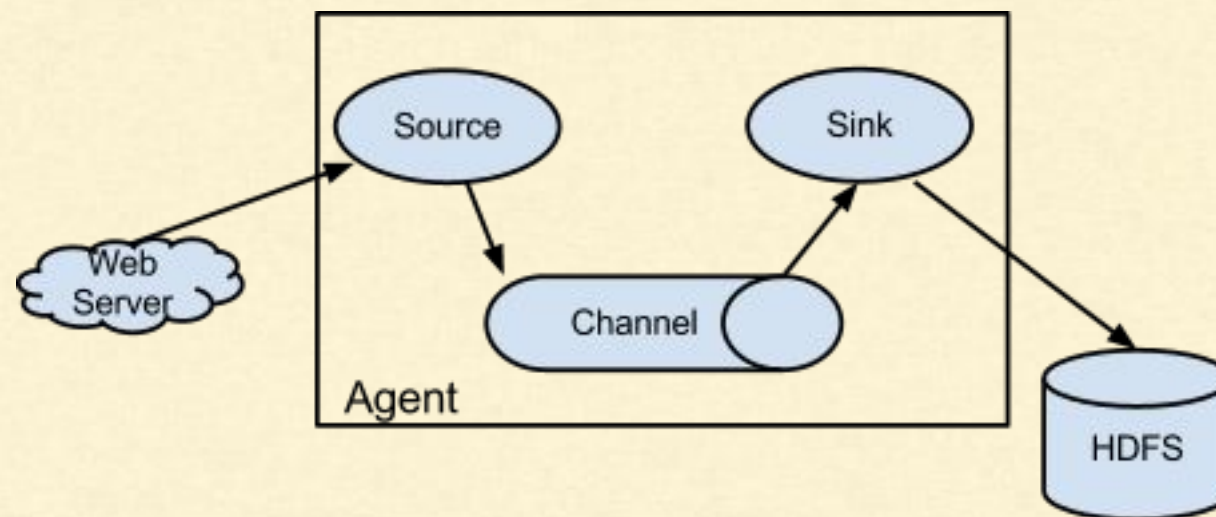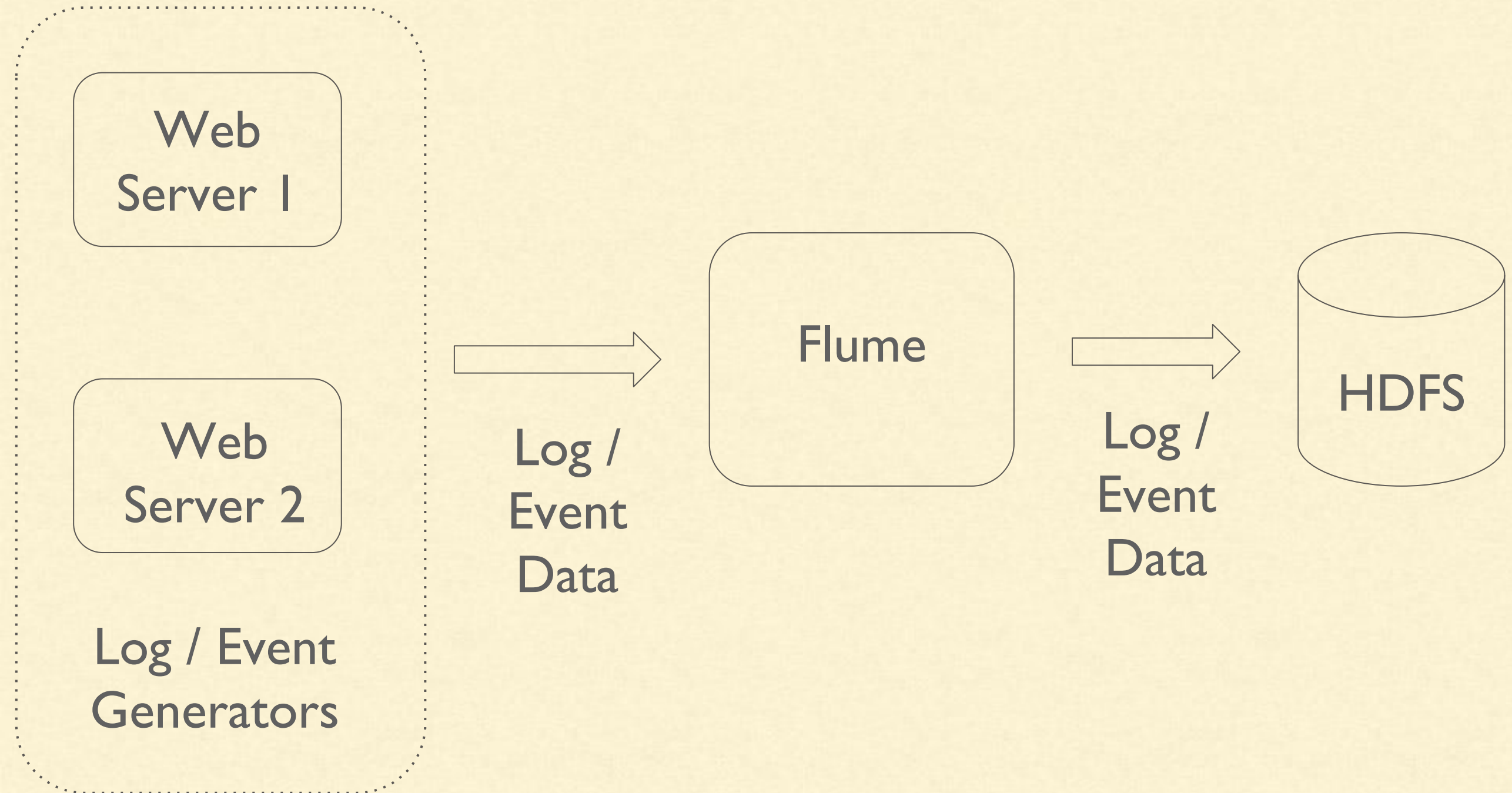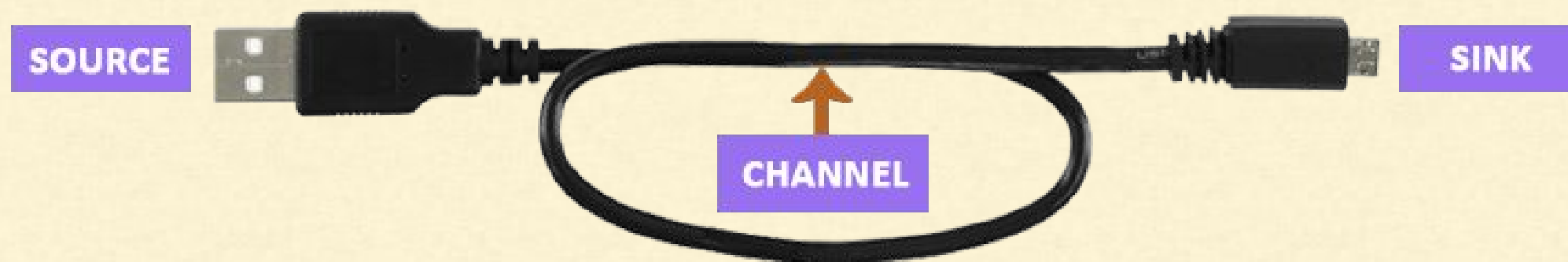# Welcome to Flume

Flume - Introduction

A distributed, reliable, and available system for
efficiently collecting, aggregating & moving large data from
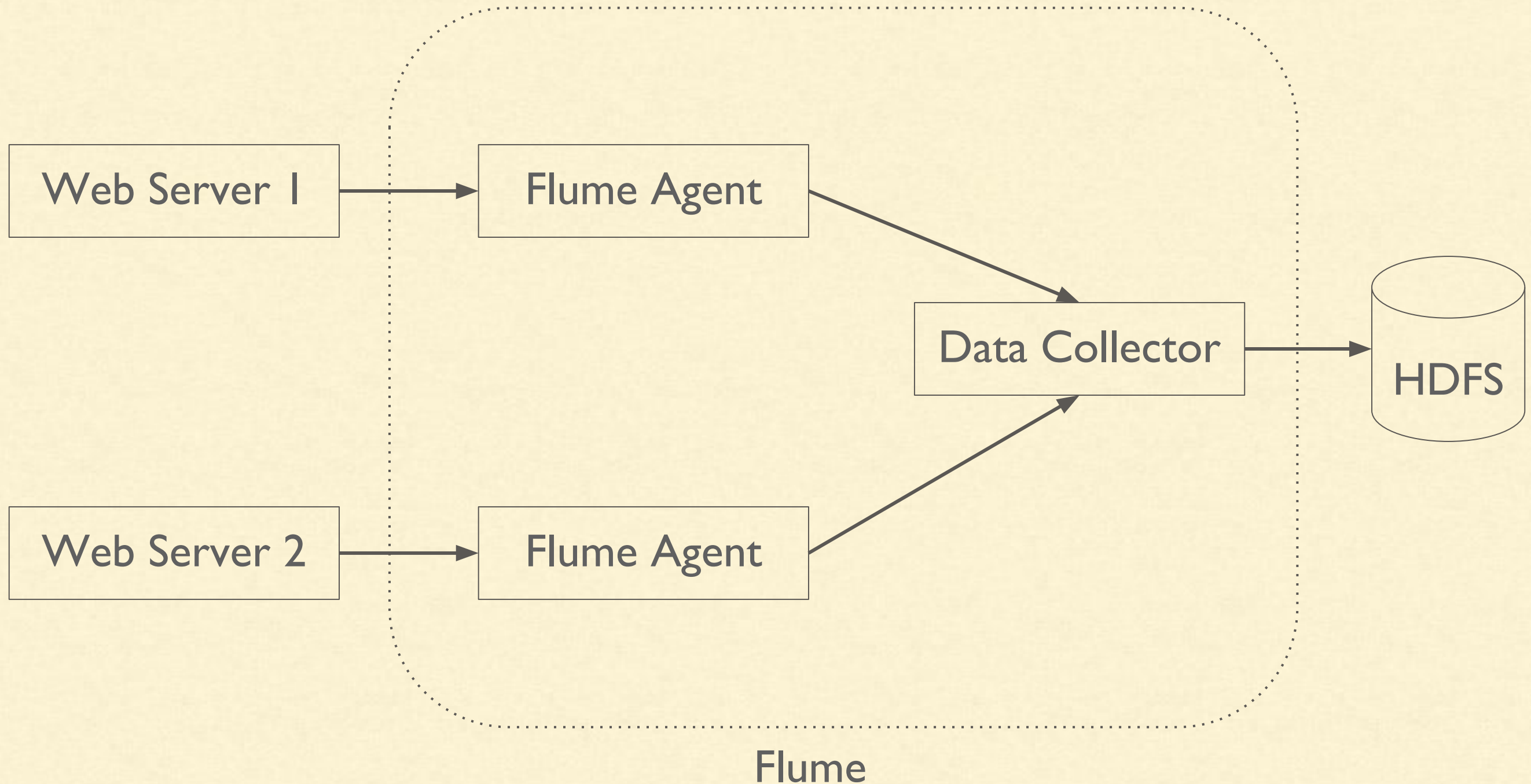many different sources to a centralized data store.
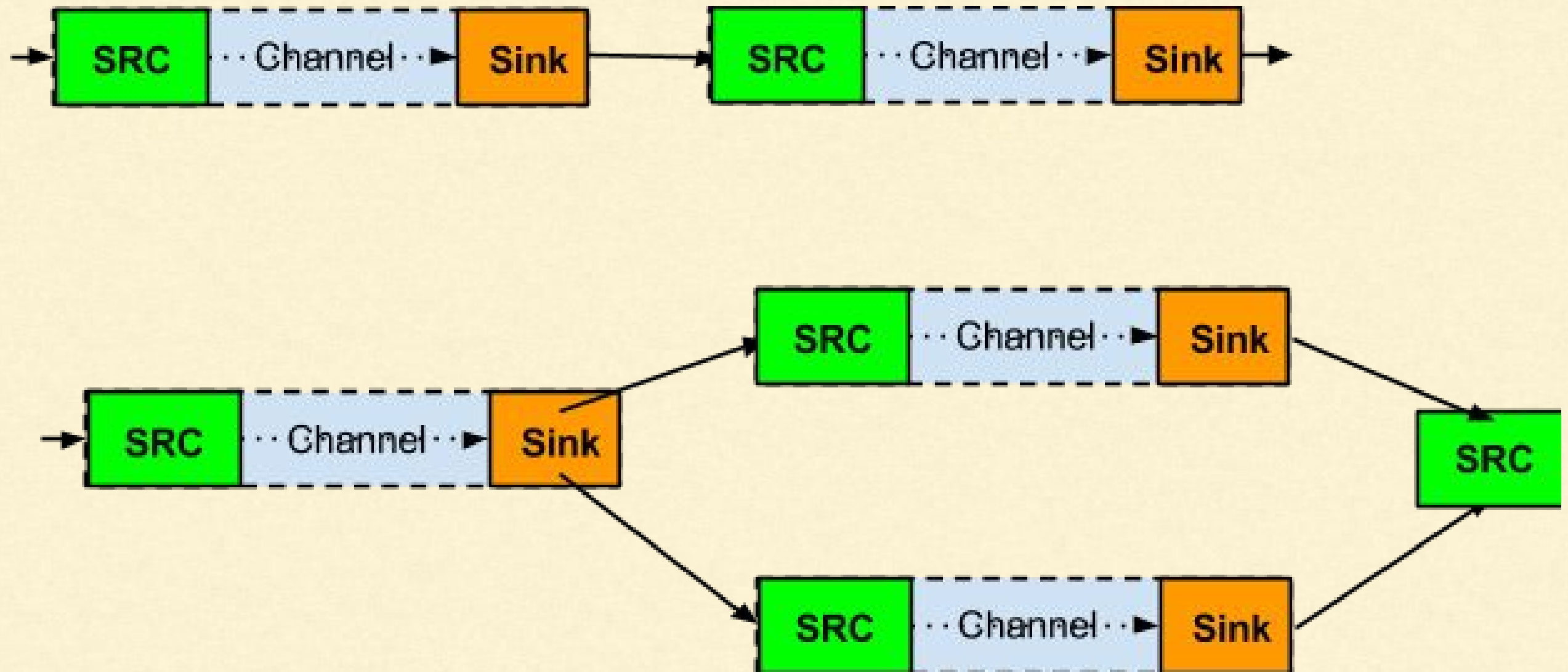
Flume - Use Case

Flume - Agents

Flume - Use Case - Agents

Flume - Multiple Agents

CLOUD x LAB

Flume - Sources

Supports a large variety of sources Including:

- tail (like unix tail -f)

- Syslog

- log4j

CLOUD x LAB

Flume - Delivery Reliability

- best-effort delivery - doesn't tolerate any node failures

- end-to-end - which guarantees delivery in node failures

Flume - Hands-on

Read the data at a port and push to HDFS

CLOUD x LAB

1. Start The Agent

flume-ng agent --conf conf --conf-file
*conf/flume.properties *
--name a1 Dflume.root.logger=INFO,console

2. Generate Some Data

nc localhost 44443

CLOUD x LAB

Flume - Summary

- Introduction & Use case

- Agents

- Hands-on

Thank you!