

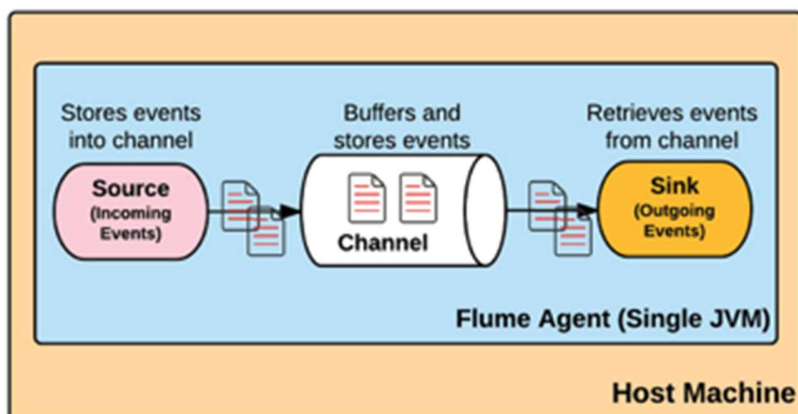


Data Lake – Design, Projeto e Integração 2.0

Data Lake - Design, Projeto e Integração Versão 2.0

Funcionamento e Configuração do Agente Flume

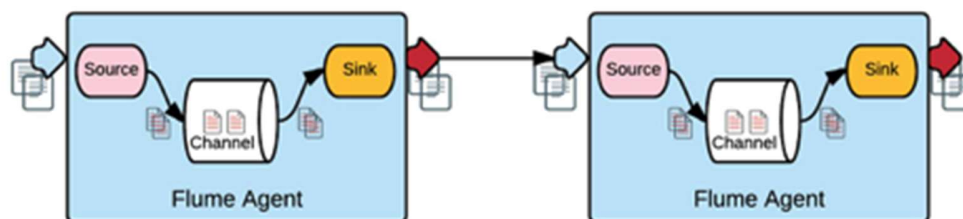
O Agente Flume é a menor implantação possível que compreende Source, Channel e Sink como seus principais componentes. A figura a seguir mostra uma implementação típica do agente Flume:



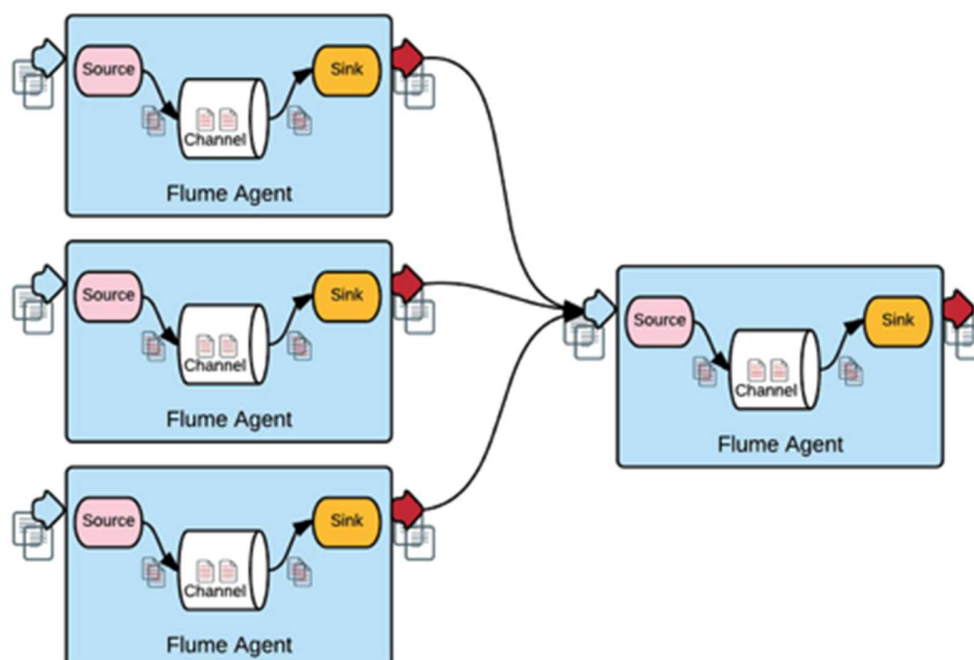
O Agente Flume é um daemon Java que recebe o evento de uma fonte e depois passa para um canal, onde geralmente é gravado no disco (de acordo com o nível de confiabilidade definido) e, em seguida, move o evento para o coletor. Quando o coletor recebe o evento, ele envia o reconhecimento de volta ao canal e o canal apaga o evento do seu armazenamento. O agente tem um tamanho de memória muito pequeno (-Xmx20m) e pode ser controlado declarativamente usando configurações.

A seguir estão as principais configurações como os agentes podem ser organizados:

Multi-hop (multi-agent flow): Semelhante ao Distributed Pipeline apresentado nas aulas anteriores. Um agente alimenta o evento para outro (chamado hop) e este continua de acordo com caso de uso demanda. A imagem abaixo mostra esse tipo de configuração:



Consolidation (converging flow): Semelhante ao design de 3 camadas apresentado nas aulas anteriores. Vários agentes consolidam os eventos e os enviam para outro conjunto de agentes. A imagem abaixo mostra esse tipo de configuração:



Multiplexing: Dentro de um agente, uma fonte se espalha por vários canais e um canal se alimenta da origem de outro agente. Essa é uma configuração flexível e vários tipos de configurações são possíveis, conforme exigido pelo caso de uso.