

Flume 本身不限制 Agent 中 Source、Channel 和 Sink 的数量。因此 Flume Source 可以接收事件，并可以通过配置将事件复制到多个目的地。这使得 Source 通过 Channel 处理器、拦截器和 Channel 选择器，写入数据到 Channel 成为可能。

每个 Source 都有自己的 Channel 处理器。每次 Source 将数据写入 Channel，它是通过委派该任务到其 Channel 处理器来完成的。然后，Channel 处理器将这些事件传到一个或多个 Source 配置的拦截器中。

拦截器是一段代码，可以基于某些它完成的处理来读取事件和修改或删除事件。基于某些标准，如正则表达式，拦截器可以用来删除事件，为事件添加新报头或移除现有的报头等。每个 Source 可以配置成使用多个拦截器，按照配置中定义的顺序被调用，将拦截器的结果传递给链的下一个单元。这就是所谓的责任链 (*chain-of-responsibility*) 的设计模式。一旦拦截器处理完事件，拦截器链返回的事件列表传递到 Channel 列表，即通过 Channel 选择器为每个事件选择的 Channel。

Source 可以通过处理器 - 拦截器 - 选择器路由写入多个 Channel。Channel 选择器是决定每个事件必须写入到 Source 附带的哪个 Channel 的组件。因此拦截器可以用来插入或删除事件中的数据，这样 Channel 选择器可以应用一些条件在这些事件上，来决定事件必须写入哪些 Channel。Channel 选择器可以对事件应用任意过滤条件，来决定每个事件必须写入哪些 Channel，以及哪些 Channel 是必需的或可选的。

◀ 12

写入到必需的 Channel 失败将会导致 Channel 处理器抛出 `ChannelException`，表明 Source 必须重试该事件（实际上，所有的事件都在那个事务中），而未能写入可选 Channel 失败仅仅忽略它。一旦写出事件，处理器会对 Source 指示成功状态，可能发送确认 (ACK) 给发送该事件的系统，并继续接受更多的事件。图 2-2 展示了该流程。