



图2-6 一个支持大量应用服务器、提供大量缓冲的巨大的复杂的拓扑结构

在大多数情况下，应用程序和 Flume Agent 之间、Flume Agent 本身之间的通信，对于 Agent 或机器失败必须是弹性的。即使当机器宕机，失败机器上的数据是不可用的，如果容量尚未耗尽应该不会引起积压。Flume 具备针对这样一类应用程序的特性，该类应用程序允许使用 Flume API 自动在多个 Flume Agent 之间负载均衡（这将是 Flume Agent 最外层的一个子集），也允许 Sink 通过 Sink 处理器，在下一层的多个 Agent 之间负载均衡。如果仍有容量存在于失败的 Agent 上，这两个组合能确保数据流持续。

## 复制数据到不同目的地

通常，在 HBase 中将事件计数器和一些元数据聚合在一起，来用于用户应用程序的实时查询，同时实际数据写入 HDFS 做详细的处理和分析。Flume 也允许这样的拓扑。图 2-7 展示了这样一个例子。