

2.4 运行时

2.4.1 运行时结构

1.Task线程

线程是程序运行时的最小单元，是进程中的一个实体，是被系统独立调度的基本单位，同属一个进程的所有线程共享进程所拥有的全部资源。

Flink的每个Operator称为一个任务（task），Operator的每个实例称为子任务，每个任务（包括子任务）在一个JVM线程中执行。可以将多个子任务链接（chain）成一个任务，在一个线程中执行，这会降低线程上下文切换产生的开销，减小缓存容量，提高系统吞吐量的同时降低延迟。此外，这种链接机制是可配置的，这增强了数据处理应用程序的灵活性。

机制与策略：理论指导引擎架构实现，但是实现通常考虑得更多。架构通常要权衡复杂度和灵活性，即引擎在简化复杂度的情况下提供什么机制（提供什么能力）、应用程序将会获得更丰富的实现策略（如何使用这些能力）。

在Programm 2.1中，Source[1]和map（）[1]、Source[2]和map（）[2]分别链接成一个任务，这种编排是因为Source和map的两个实例之间采用直连模式，它们之间的数据传输可以通过缓存而不是网络通信，正是这种针对性的优化提升了Flink的执行效率。整个应用程序由5个线程构成，如图2-8所示。

