Flink入门-4

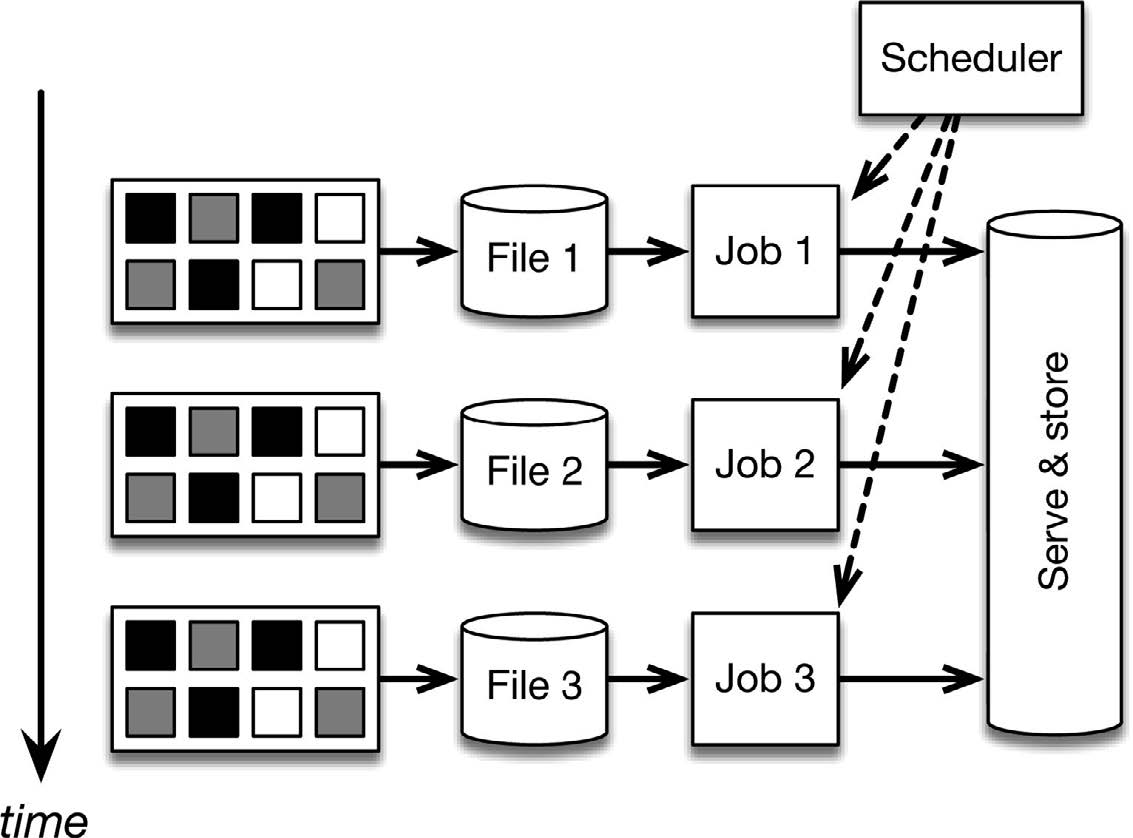
事件时间

处理时间 事件时间

为了将事件指定给某一个窗口，程序员很有可能会选择采用事件时间，即事件

实际发生的时间。另一种方式是采用处理时间，即事件流数据开始被程

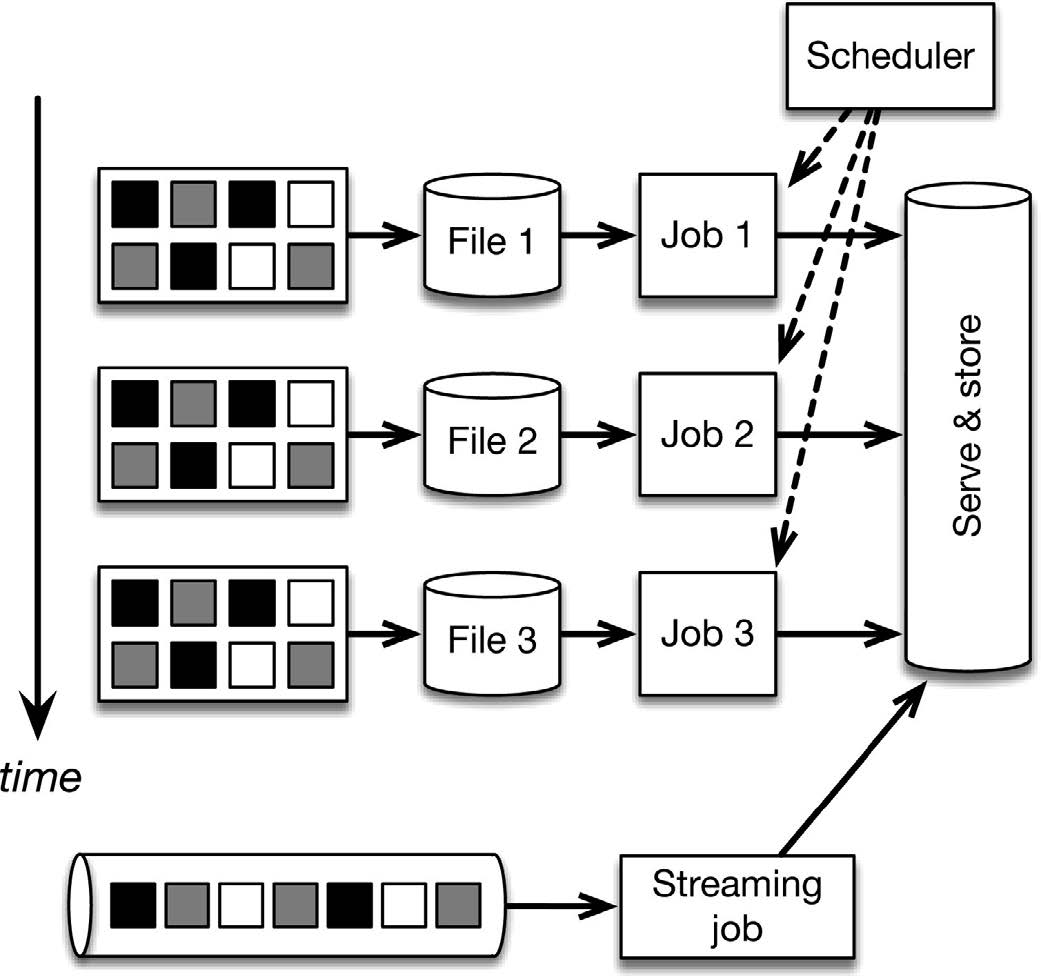
序处理的时间。



太多独立的部分

对时间的处理方法不明确

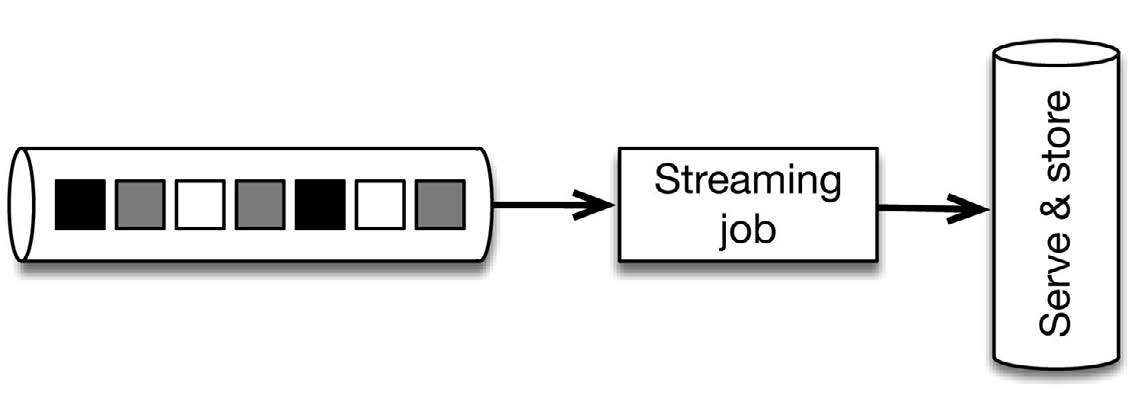
预警



乱序事件流

批处理作业的界限不清晰

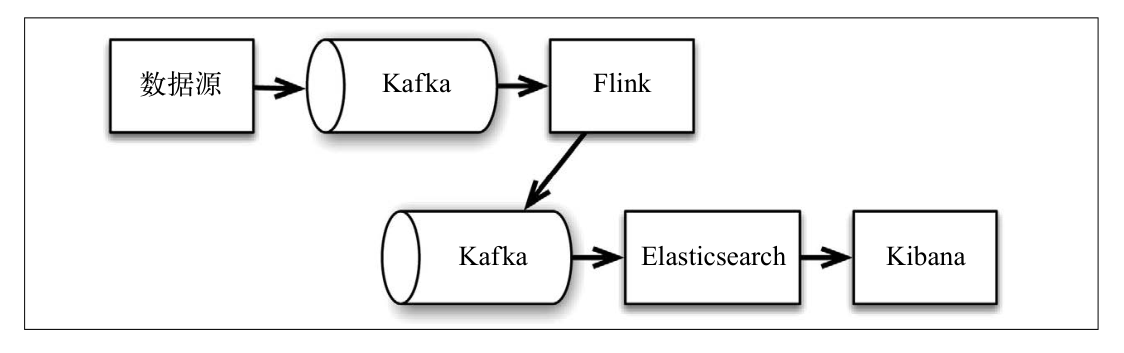
批处理作业的界限不清晰



摄取时间

时空穿梭

水印



有助于准确地识别异常

有助于采用流处理架构

**流式计算分为无状态和有状态两种情况**

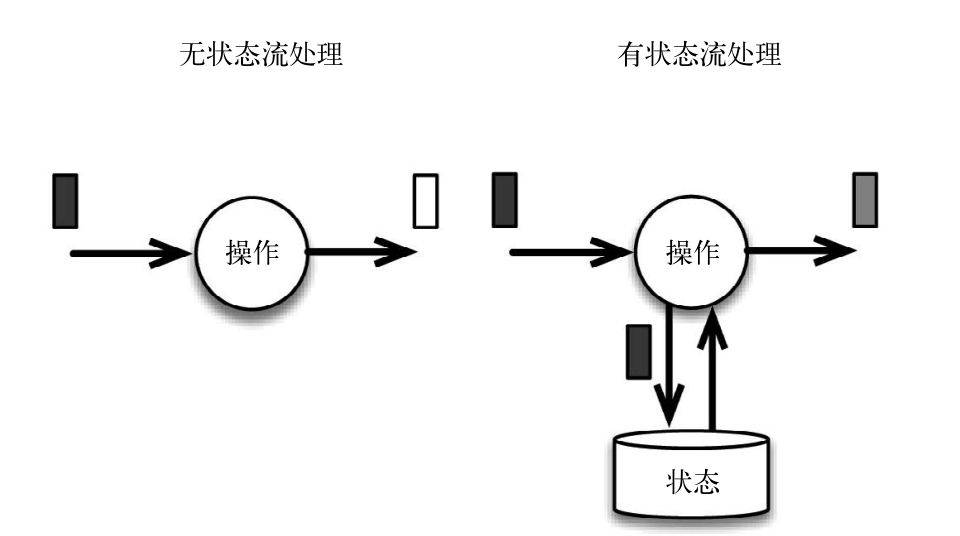
**有状态**

计算过去一小时的平均温度

若在一分钟内收到两个相差 20度以上的温度读数，则发出警告

**无状态**

温度超过 90 度时发出警告



一致性

一致性分为 3 个级别

at-most-once: 故障发生之后，计数结果可能丢失

at-least-once: 这表示计数结果可能大于正确值，但绝不会小于正确值

exactly-once: 发生故障后得到的计数结果与正确值一致

检查点 保证exactly-once

保存点savepoint

应用程序代码升级

Flink 版本更新

维护和迁移

假设模拟与恢复

A/B 测试

端到端的一致性和作为数据库的流处理器

