# 大数据

## Storm

<https://blog.csdn.net/u011082453/article/details/82417259>

<https://blog.csdn.net/weixin_34116110/article/details/93349980?utm_medium=distribute.pc_relevant.none-task-blog-BlogCommendFromMachineLearnPai2-1.channel_param&depth_1-utm_source=distribute.pc_relevant.none-task-blog-BlogCommendFromMachineLearnPai2-1.channel_param>

## Flink

<https://www.jianshu.com/p/cdcdec05a249>

借助先进的流处理引擎，还可以实时地执行分析。流式查询或应用程序不是读取有限数据集，而是摄取实时事件流，并在消耗事件时不断生成和更新结果。结果要么写入外部数据库，要么保持为内部状态。仪表板应用程序可以从外部数据库读取最新结果或直接查询应用程序的内部状态。Apache Flink支持流式和批量分析应用程序。  
与批量分析相比，连续流分析的优势不仅限于低延迟。与批量查询相比，流式查询不必处理输入数据中的人为边界，这些边界是由定期导入和输入的有界性质引起的。  
另一方面是更简单的应用程序架构。批量分析管道由若干独立组件组成，以定期调度数据提取和查询执行。可靠地操作这样的管道并非易事，因为一个组件的故障会影响管道的后续步骤。相比之下，在像Flink这样的复杂流处理器上运行的流分析应用程序包含从数据摄取到连续结果计算的所有步骤。