Batch Processing Systems

DISTRIBUTED FRAMEWORKS

idea: specialized framework(P21)

Challenges(P34-37)

MapReduce: Simplified Data Processing on Large Clusters

- 定义概述（P40）
- **MapReduce abstraction models 怎么满足了 our requirements （P41）**
- **MapReduce Programming Interface（P48）**
- **MapReduce Execution Flow（P50）**
- **MapReduce 流程总述: The Complete Picture （P51-59）**
- MapReduce如何解决前面的**Challenges(P65-71)**
- 优缺点（P76）
- Mapreduce无法适用于所有问题的原因（P78）
- **MapReduce is not optimized for multiple-sages execution（P82）**

Dryad: Distributed Data-Parallel Programs from Sequential Building Blocks

- **Programming model of Dryad（P86）**
- **The runtime of Dryad（P89-92）**
- **Advantages of Dryad（P93）**