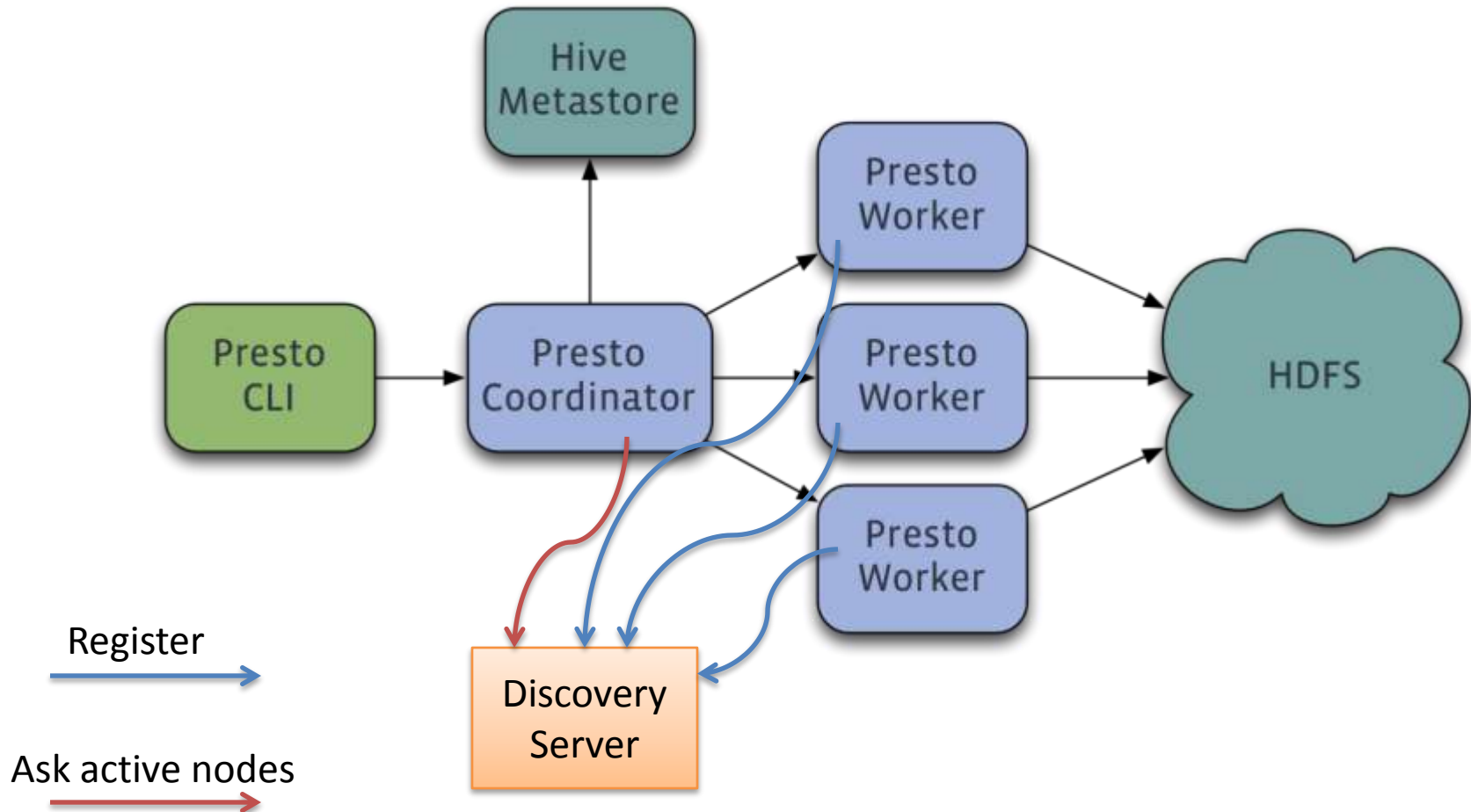


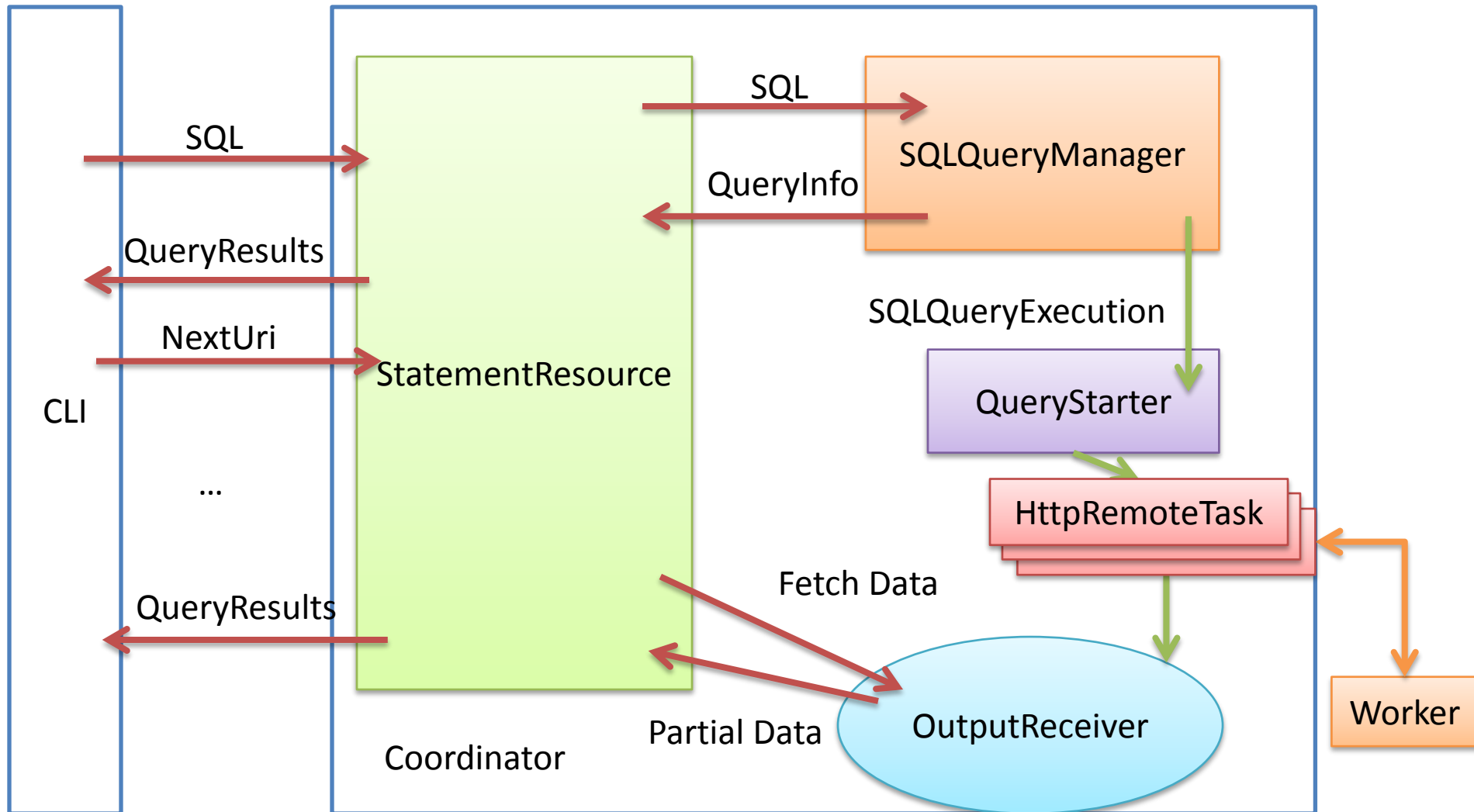
Presto Overview

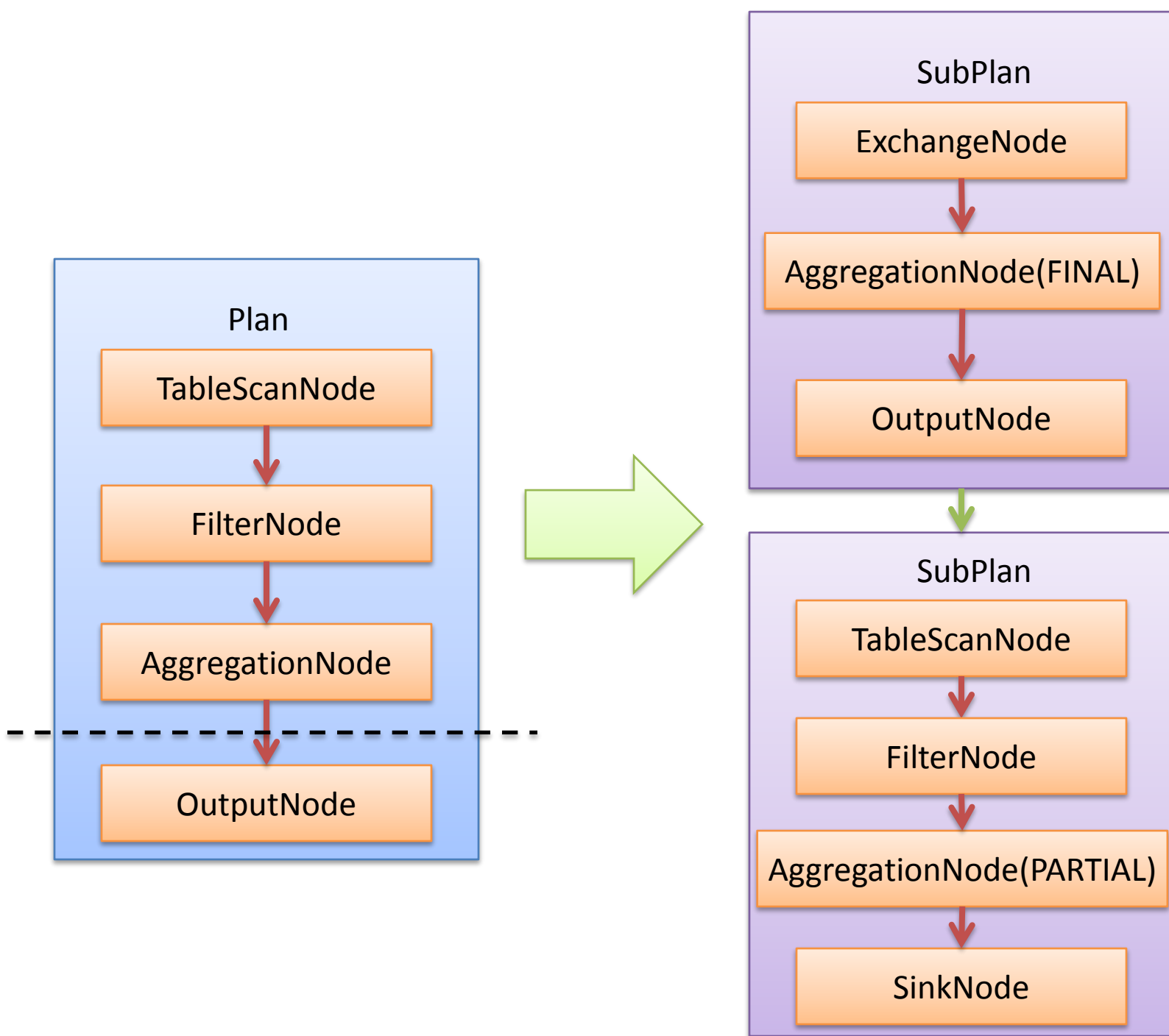
Shixiong Zhu

Overview

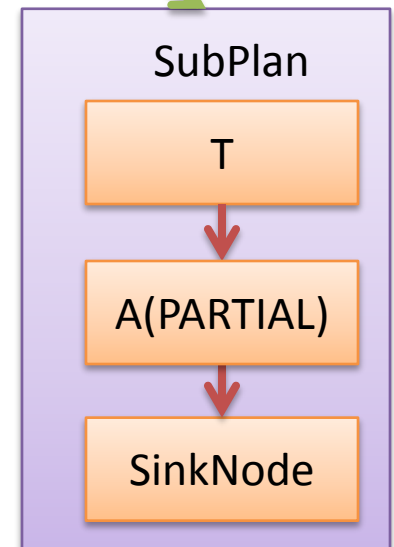
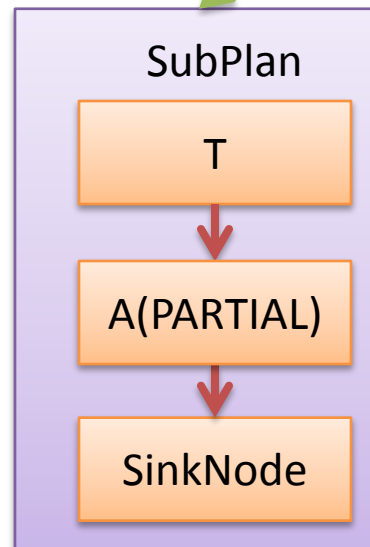
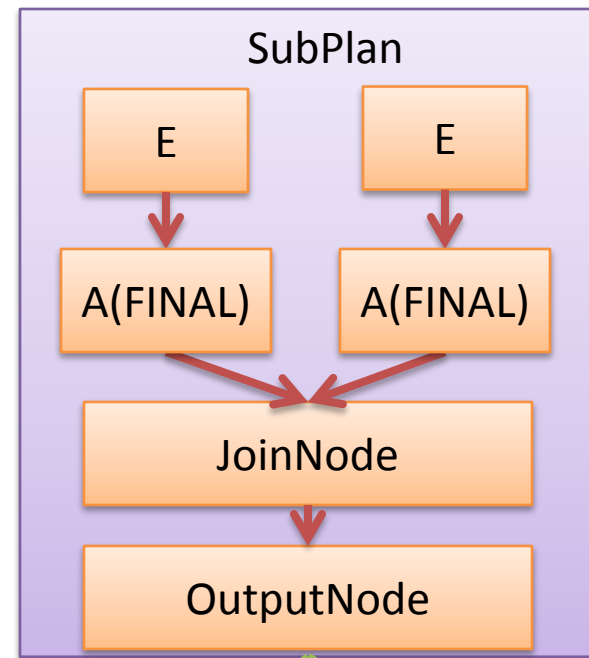
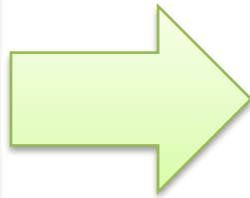
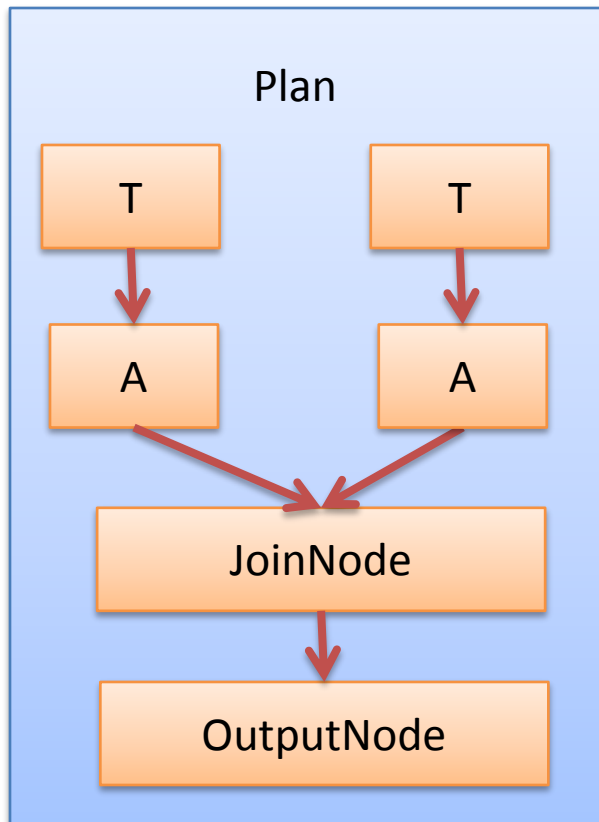


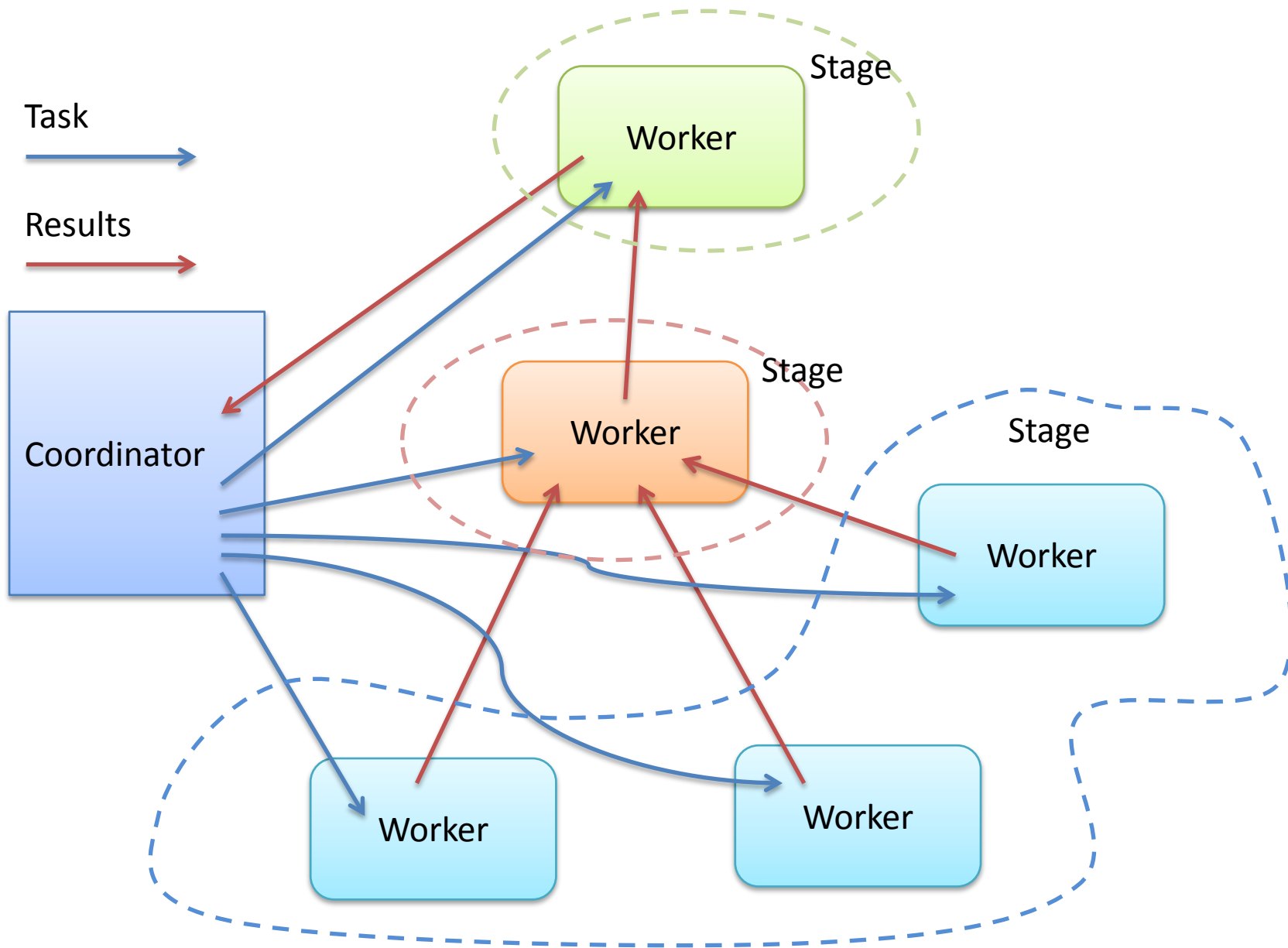
Coordinator



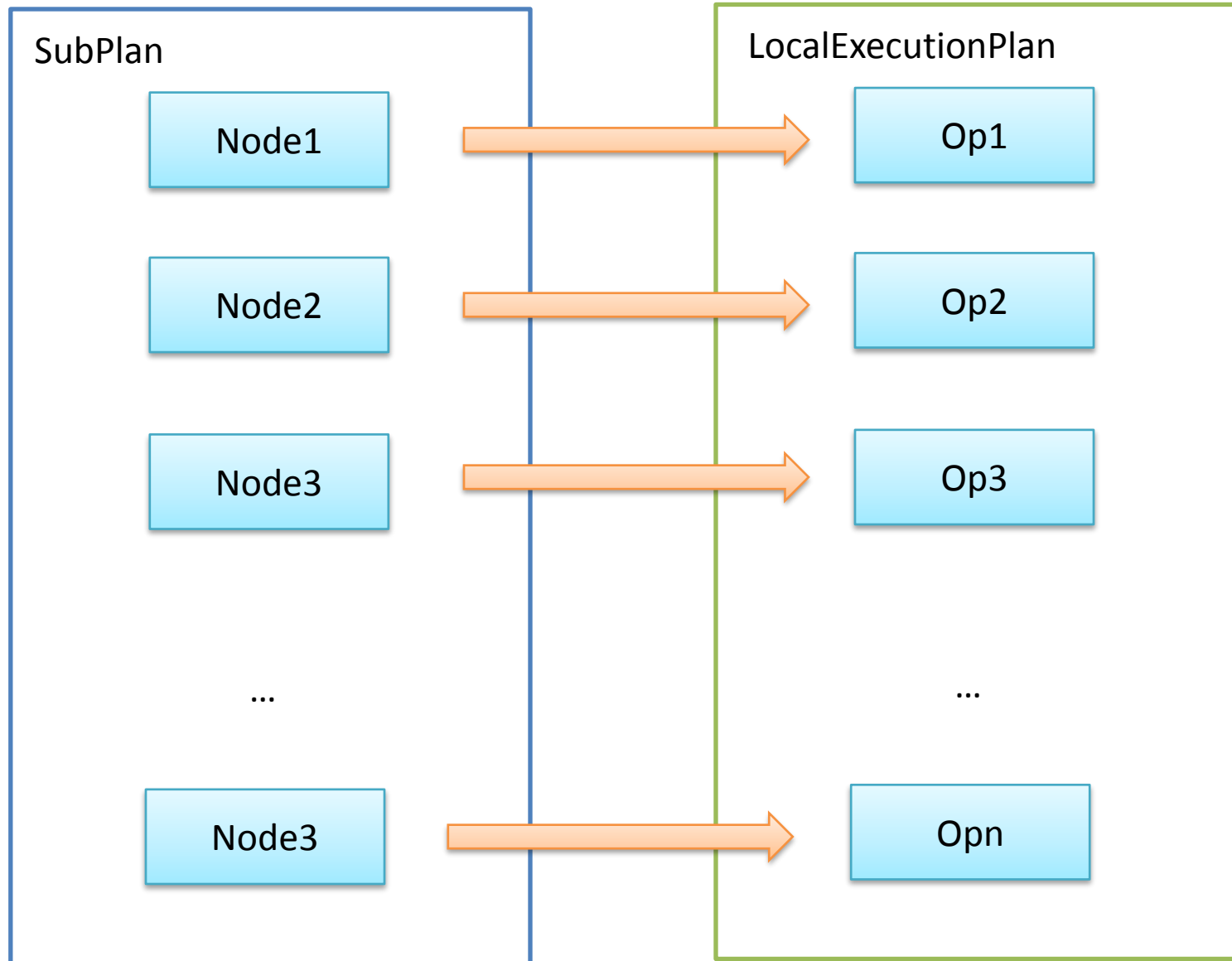


T: TableScanNode
A: AggregationNode
E: ExchangeNode





Worker



SubPlan

Node1

Node2

JoinNode

Node3

LocalExecutionPlan

Op1

HashJoinOperator

Op3

LocalExecutionPlan

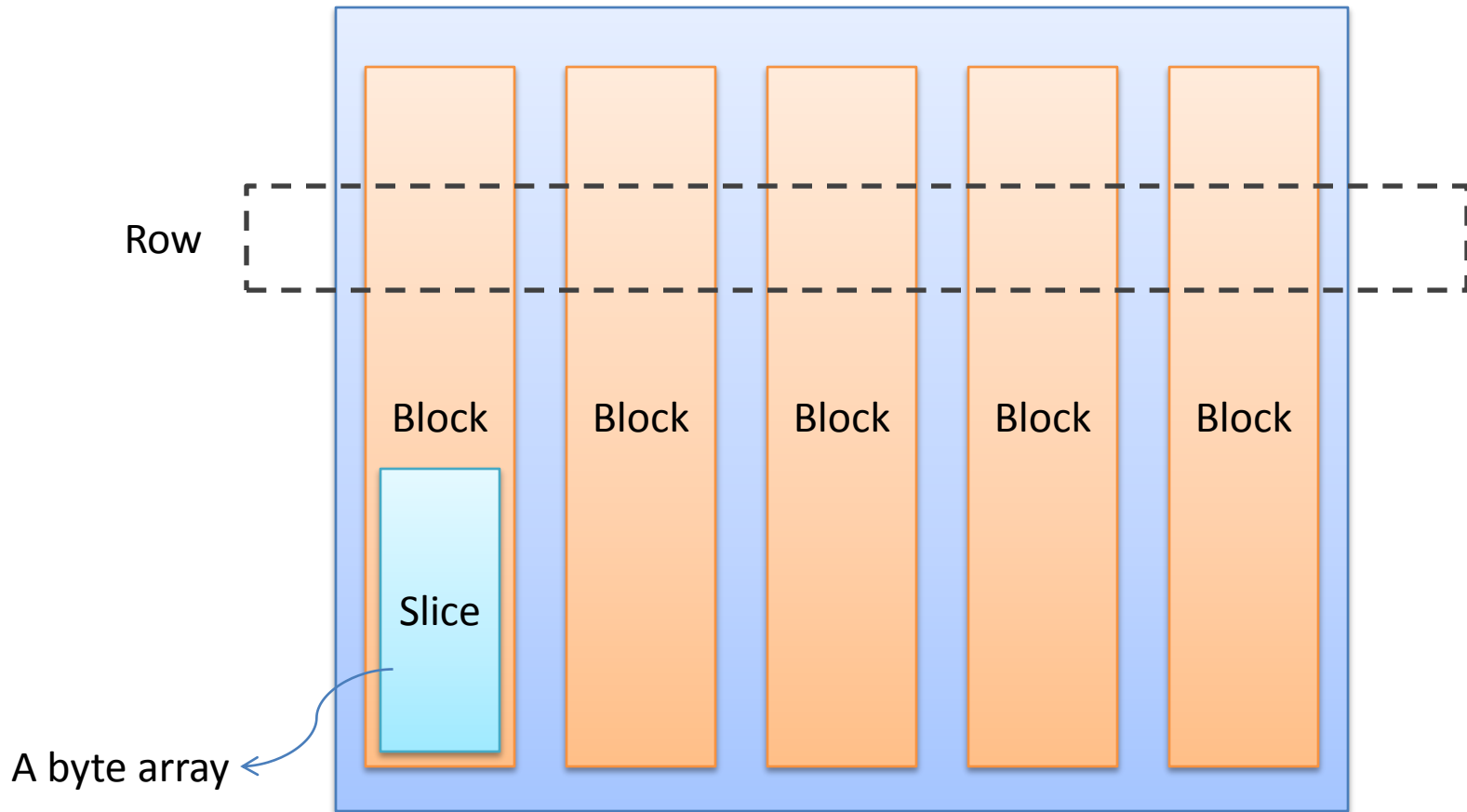
Op2

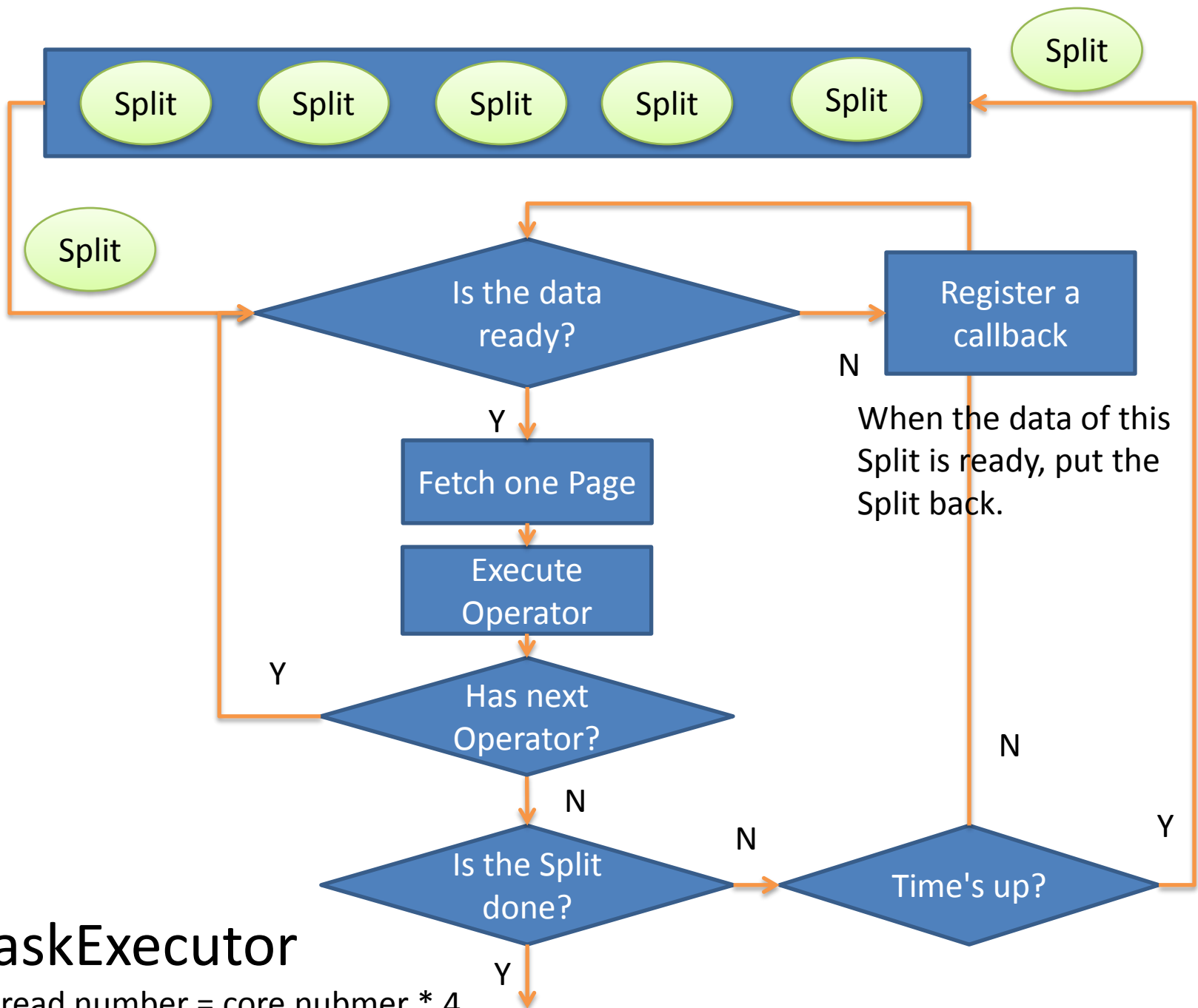
HashBuilderOperator

SourceHash



Page(max page size: 1MB, max rows:
16 * 1024)

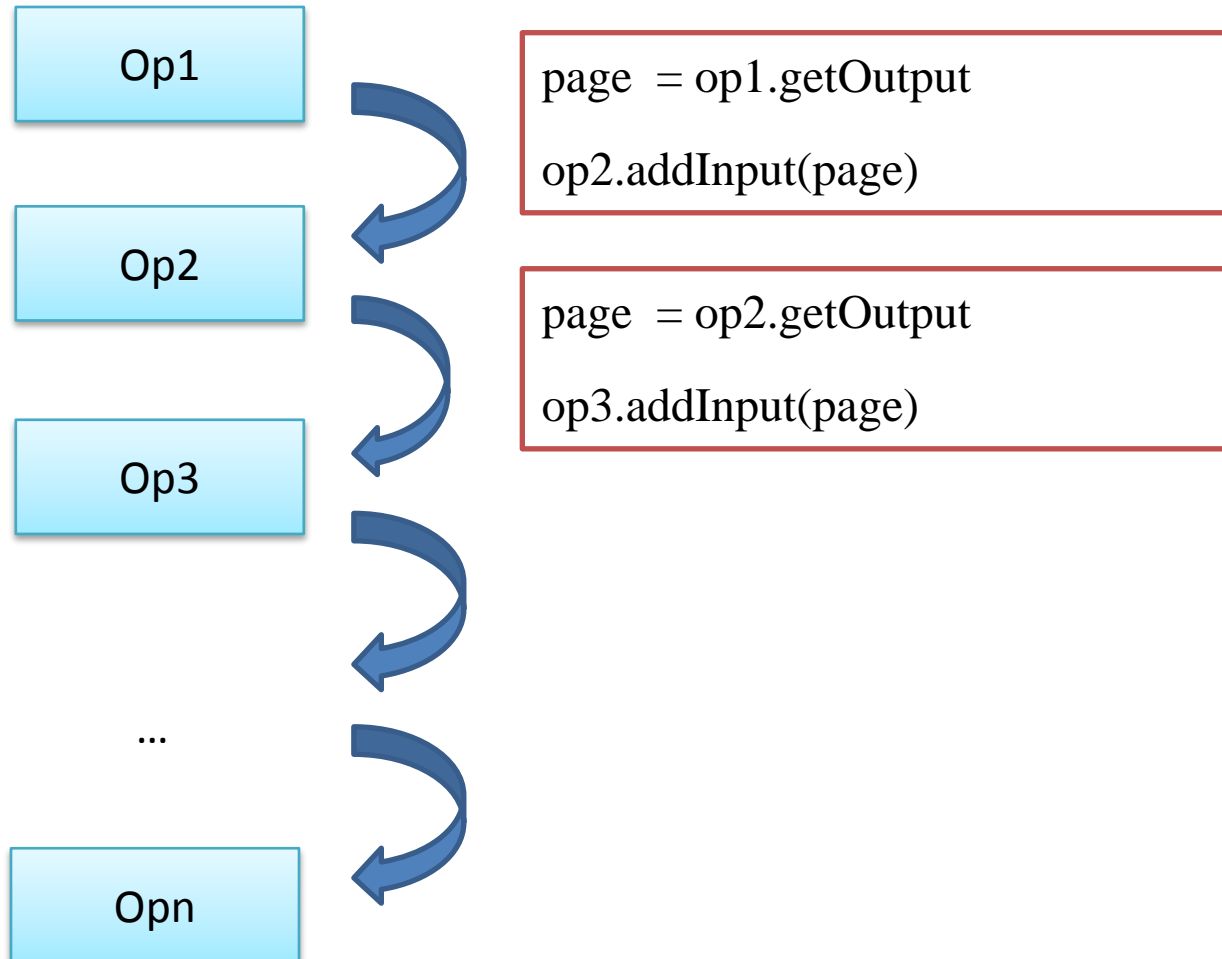




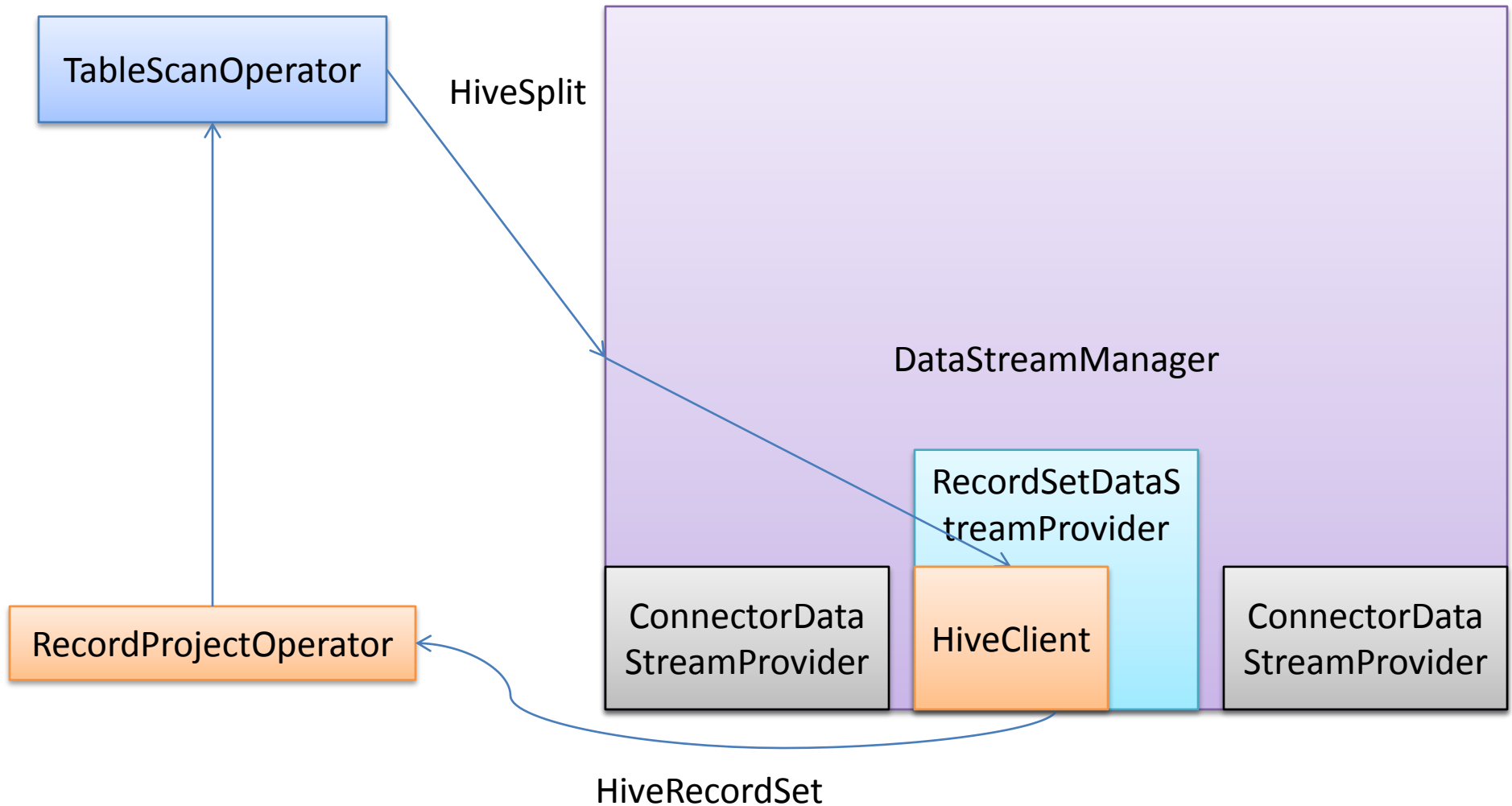
TaskExecutor

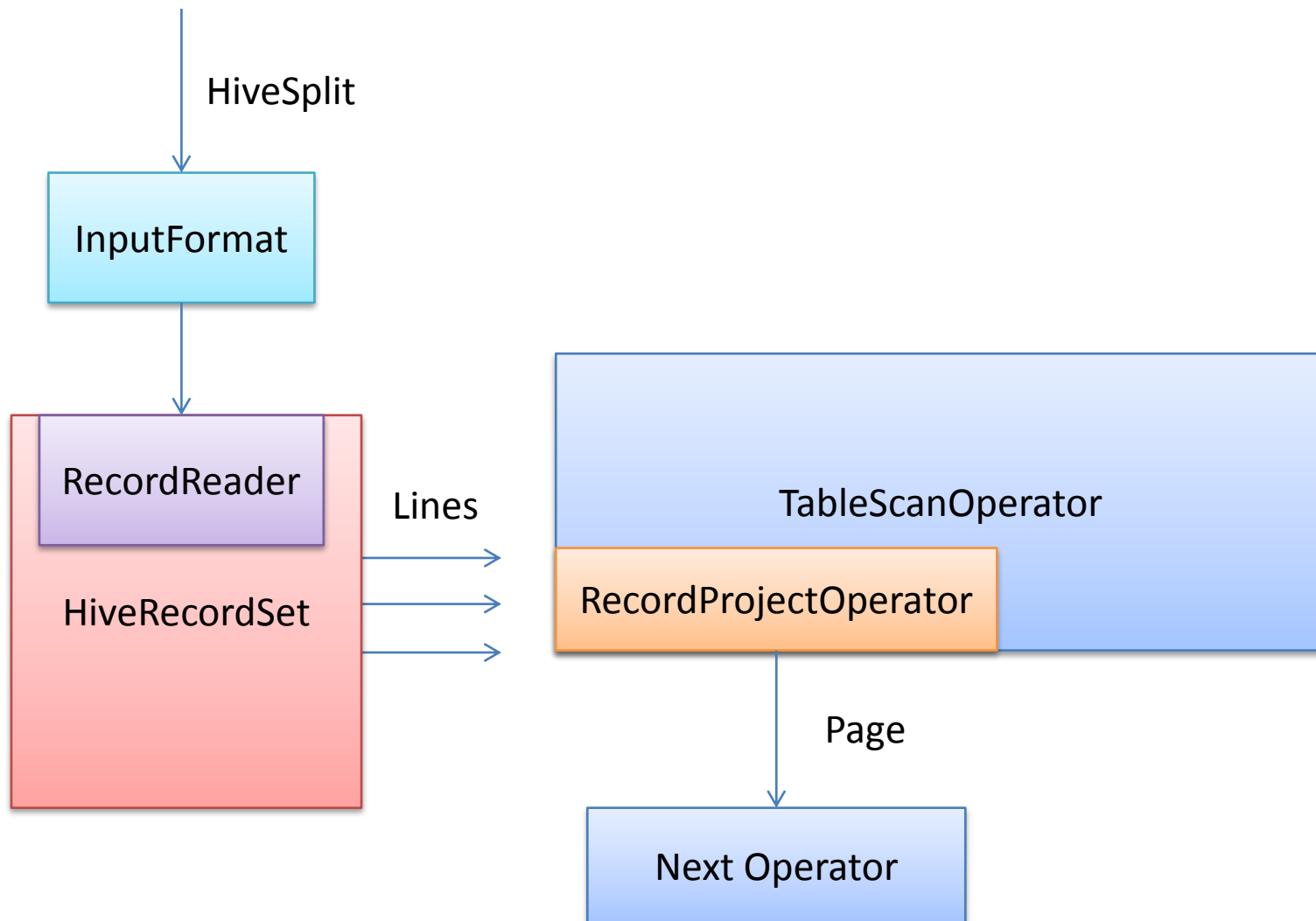
Thread number = core number * 4

Execution Operators

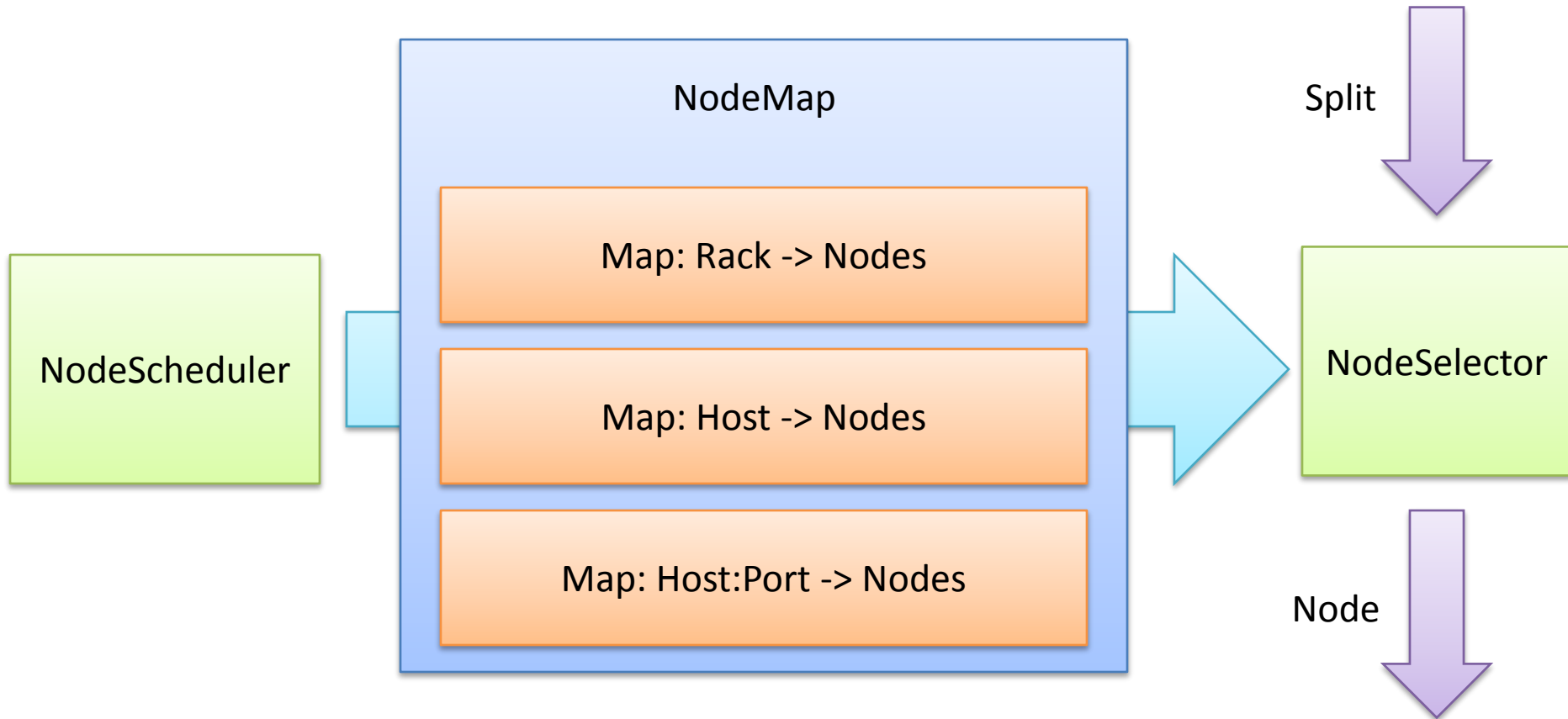


Input





Load Balance



NodeSelector.selectNode

- Select acceptable nodes (as least 10 nodes by default)
 - Nodes has the same address
 - If not enough, add nodes in the same rack
 - If not enough, randomly select nodes in other racks
- Select the node with the smallest number of assignments (pending tasks)

Output

- Only has **SELETE** statement
 - Currently query results are streamed to the client

Communication

- Protocol: HTTP
- Data Format: JSON
- Every instance has one server and one client

Q&A