

Batch Computation

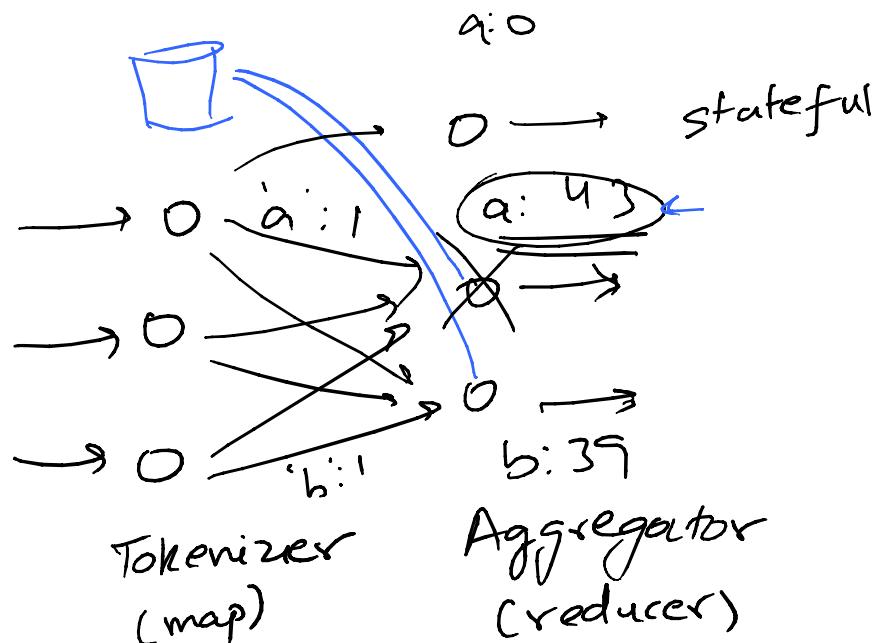
Note Title

07-Dec-22

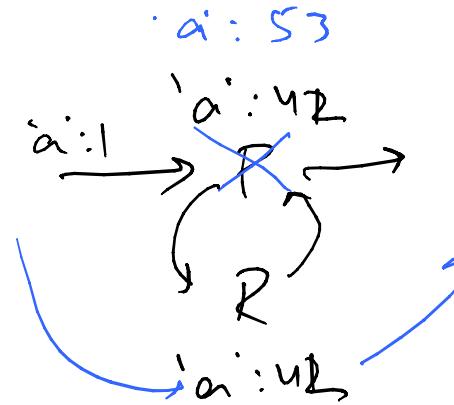
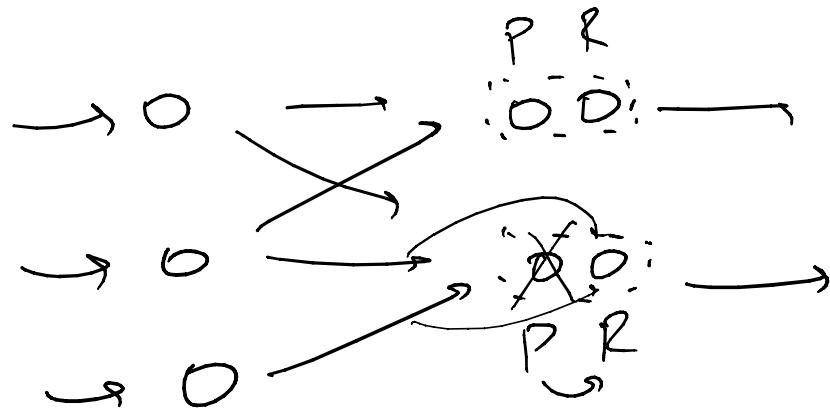
Streaming computation

- Real time decision making
- Live analytics

Freshness



Replicated state machines



uncorrelated
failure

Spark streaming

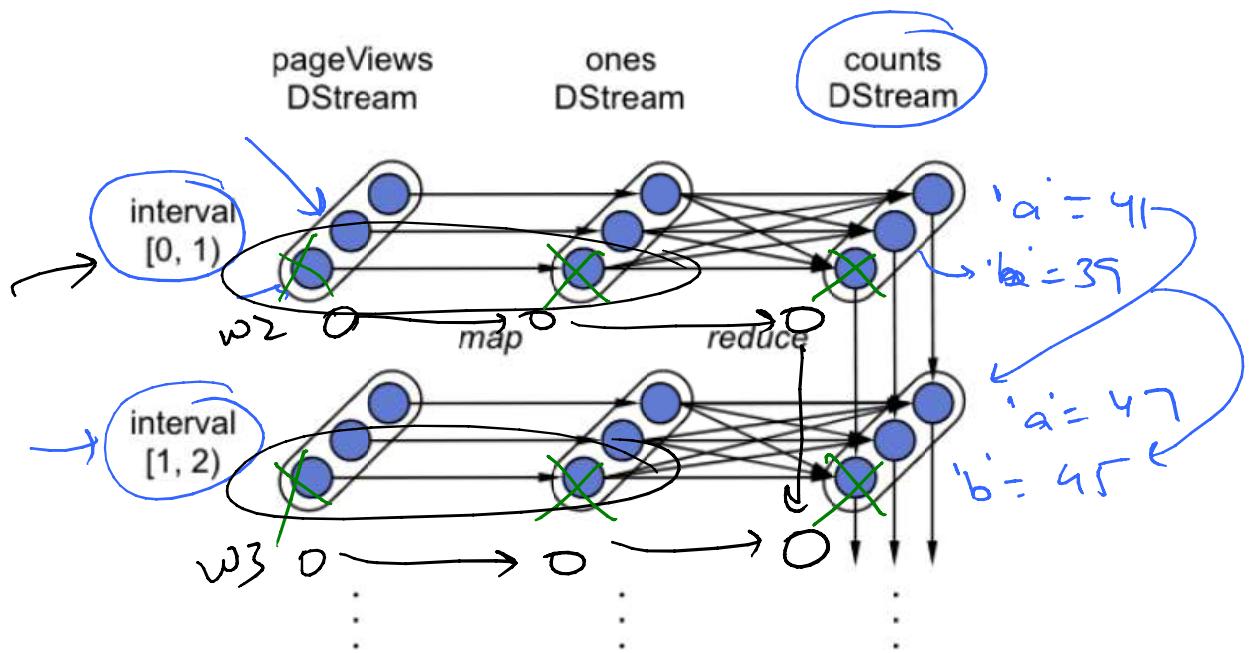
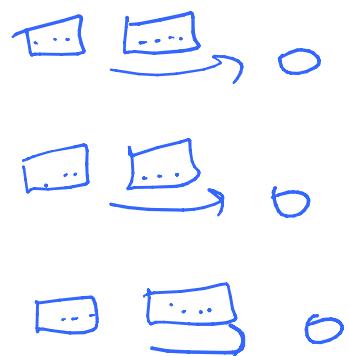
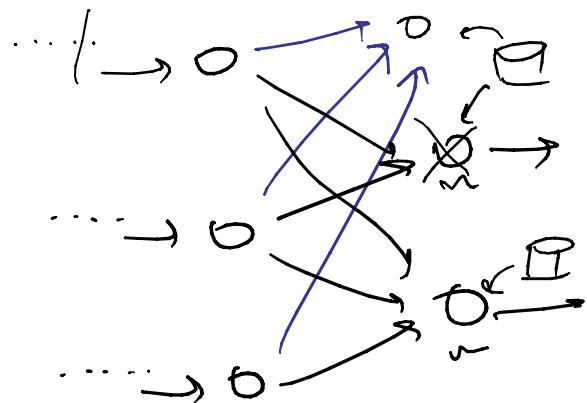
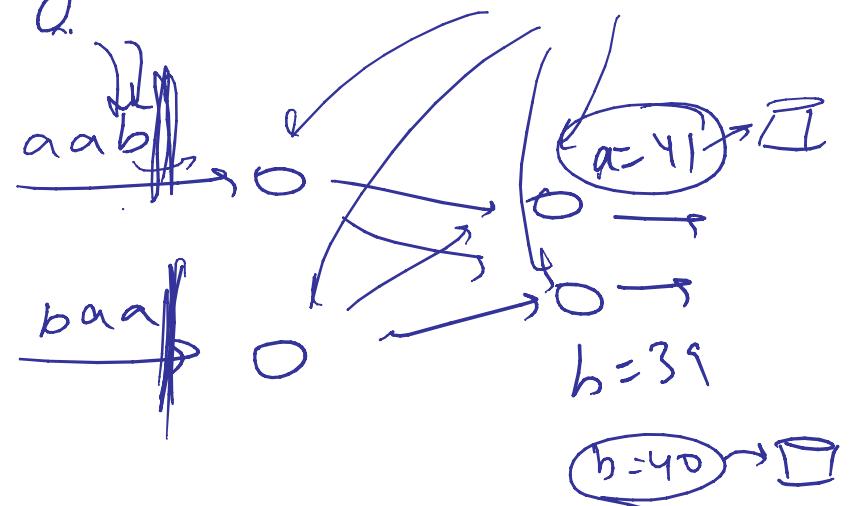


Figure 3: Lineage graph for RDDs in the view count program. Each oval is an RDD, with partitions shown as circles. Each sequence of RDDs is a D-Stream.

Checkpoints



Synchronous Global checkpoint



Flink

