CS 744: FLINK

Shivaram Venkataraman

Spring 2024

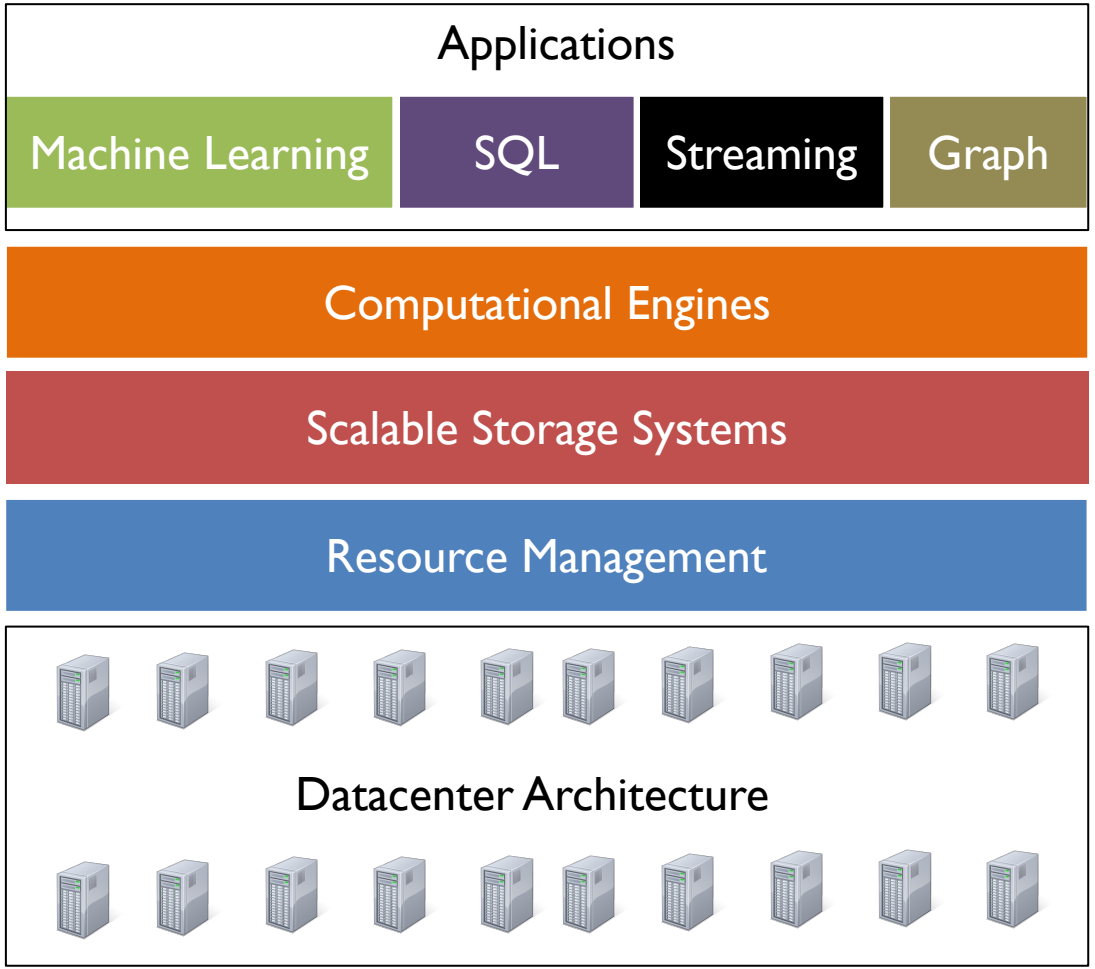
ADMINISTRIVIA

Grading

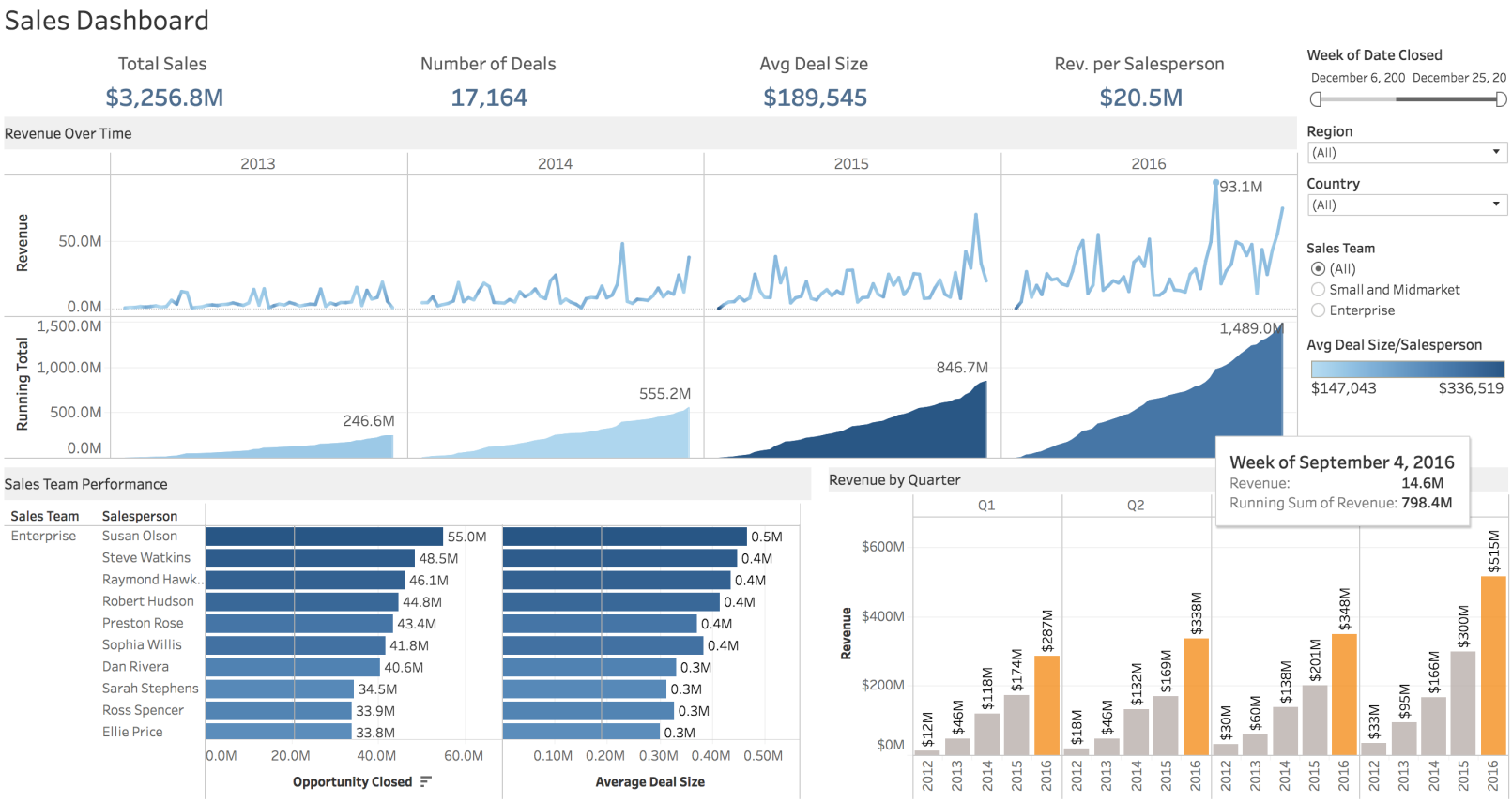
- Assignment 2 grading
- Course Project Proposal feedback
- Midterm

Resources for Course Projects

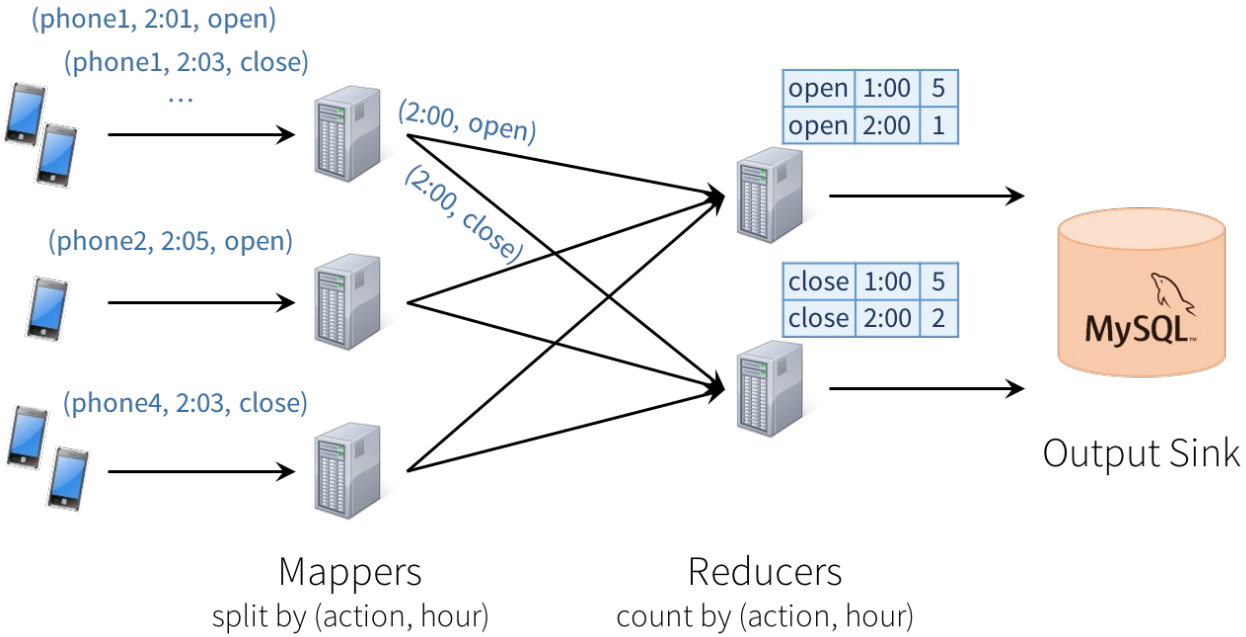
- Cloudlab (Reservations?)
- GCP credits (Email Tzu-Tao and me)


Applications
Machine Learning
SQL
Streaming
Graph
Computational Engines
Scalable Storage Systems
Resource Management
Datacenter Architecture

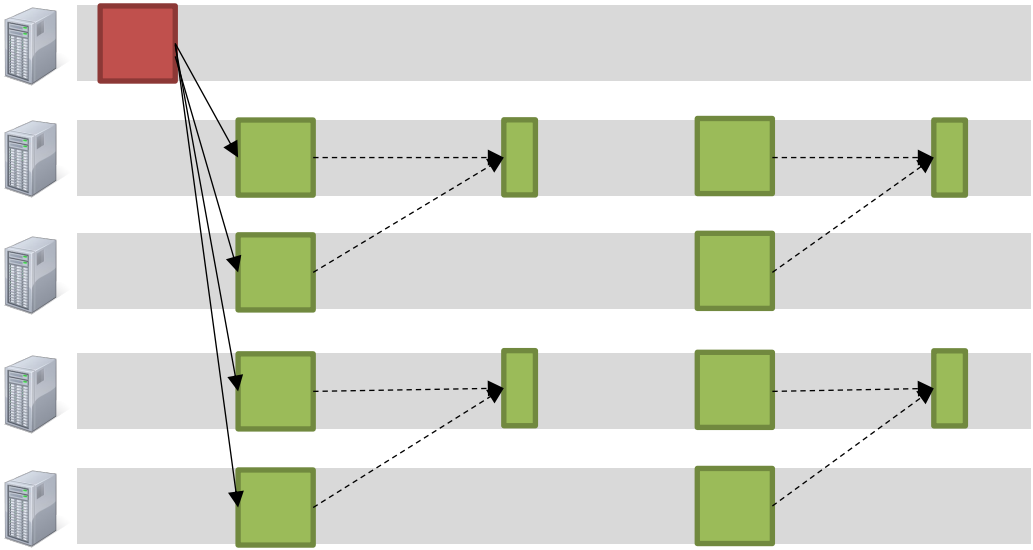
DASHBOARDS


Sales Dashboard
Total Sales
$3,256.8M
Number of Deals
17,164
Avg Deal Size
$189,545
Rev. per Salesperson
$20.5M
Week of Date Closed
December 6, 200 December 25, 20
Revenue Over Time
2013
2014
2015
2016
Revenue
50.0M
0.0M
93.1M
Running Total
1,500.0M
1,000.0M
500.0M
0.0M
246.6M
555.2M
846.7M
1,489.0M
Region
(All)
Country
(All)
Sales Team
(All)
Small and Midmarket
Enterprise
Avg Deal Size/Salesperson
$147,043
$336,519
Week of September 4, 2016
Revenue: 14.6M
Running Sum of Revenue: 798.4M
Sales Team Performance
Sales Team
Salesperson
Enterprise
Susan Olson 55.0M 0.5M
Steve Watkins 48.5M 0.4M
Raymond Hawk.. 46.1M 0.4M
Robert Hudson 44.8M 0.4M
Preston Rose 43.4M 0.4M
Sophia Willis 41.8M 0.4M
Dan Rivera 40.6M 0.3M
Sarah Stephens 34.5M 0.3M
Ross Spencer 33.9M 0.3M
Ellie Price 33.8M 0.3M
0.0M 20.0M 40.0M 60.0M
Opportunity Closed
0.10M 0.20M 0.30M 0.40M 0.50M
Average Deal Size
Revenue by Quarter
Q1
Q2
Revenue
$600M
$400M
$200M
$0M
$12M $46M $118M $174M $287M
$18M $46M $132M $169M $338M
$30M $60M $138M $201M $348M
$33M $95M $166M $300M $515M
2012 2013 2014 2015 2016

STREAMING COMPUTATION



FLINK: COMPUTATION MODEL



Long-lived operators

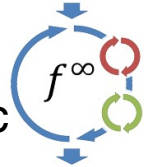
Mutable State

Google MillWheel

Manager/Driver → Control Message

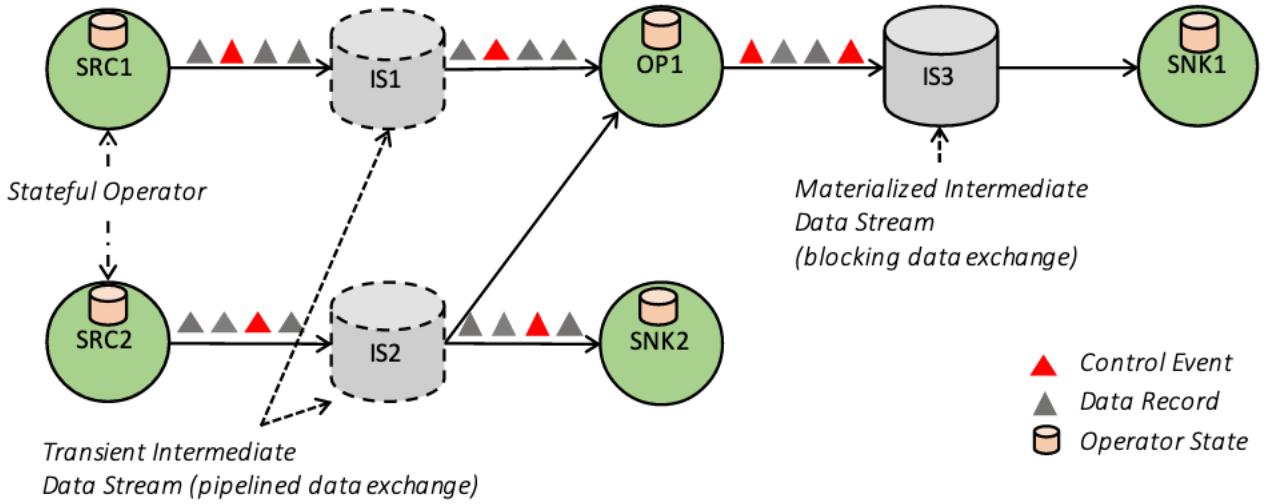
Task/Computation ⇢ Network Transfer

Streaming DBs: Borealis, Flux etc



Naiad

INTERMEDIATE DATA STREAMS



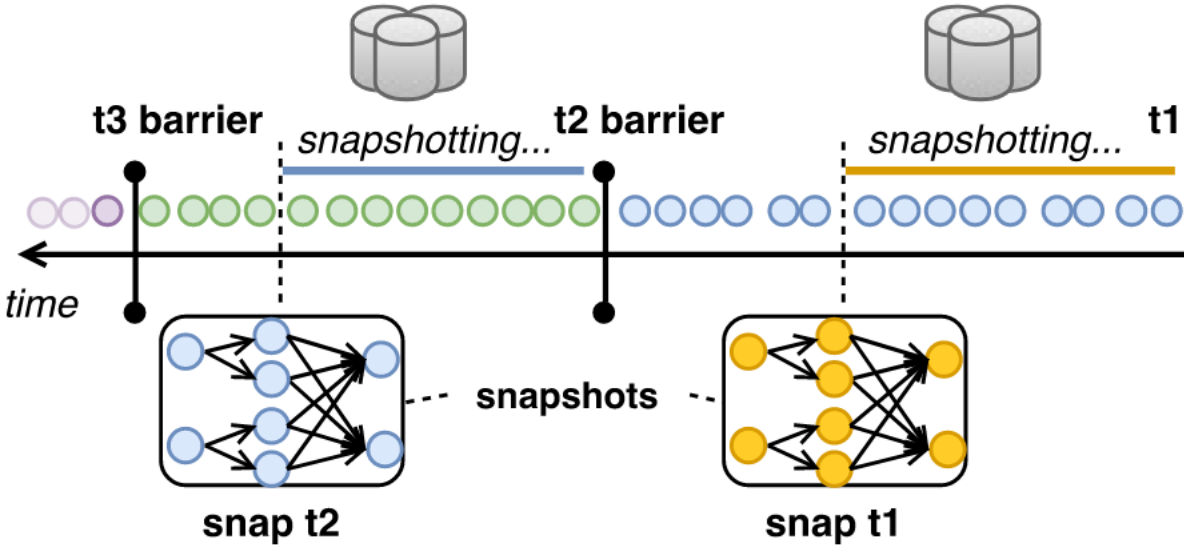
STATEFUL OPERATORS

Examples?

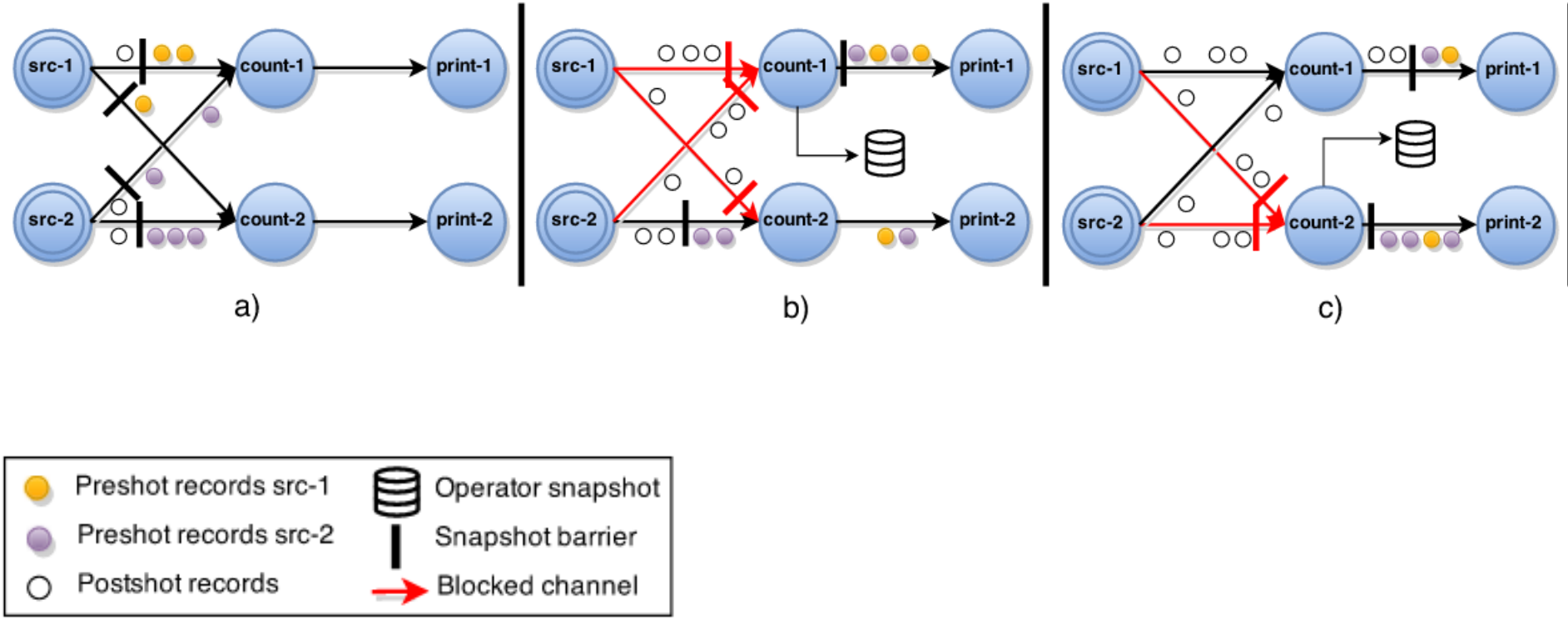
Challenge

- How to ensure fault tolerance?
- Explicitly register local variables
- StateBackends that are automatically saved/recovered

FAULT TOLERANCE: CHECKPOINTING



ASYNCHRONOUS BARRIER SNAPSHOTTING



WATERMARKS, WINDOWS

Implements similar model as Dataflow

"Watermarks originate at the sources of a topology"

Propagate through the other operators of dataflow

Windows based on event-time, processing time, ingest time(?)

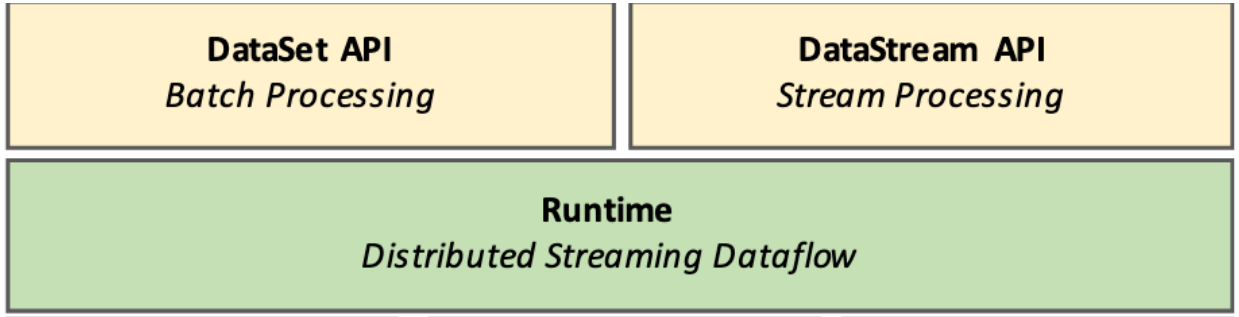
```
stream
.window(SlidingTimeWindows.of(
        Time.of(6, SECONDS), Time.of(2, SECONDS))
.trigger(EventTimeTrigger.create())
```

COMBINING BATCH, STREAMING

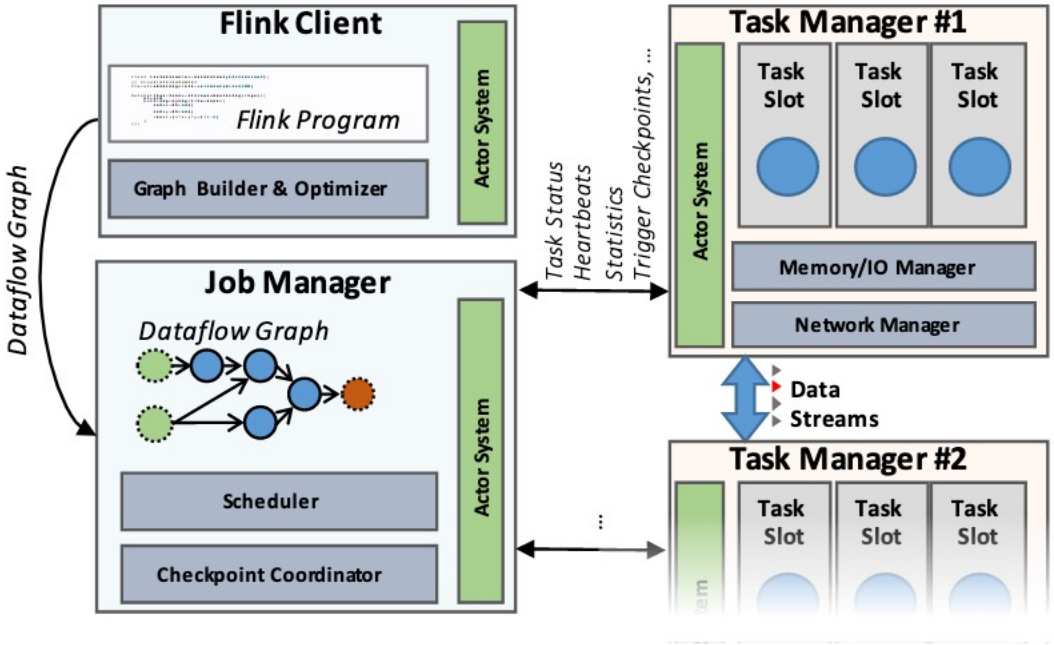
Blocked DataStreams

Turn off periodic snapshots

Blocking operators (e.g., sort)



OVERALL ARCHITECTURE



SUMMARY

Stream processing → Increasingly important workload trend

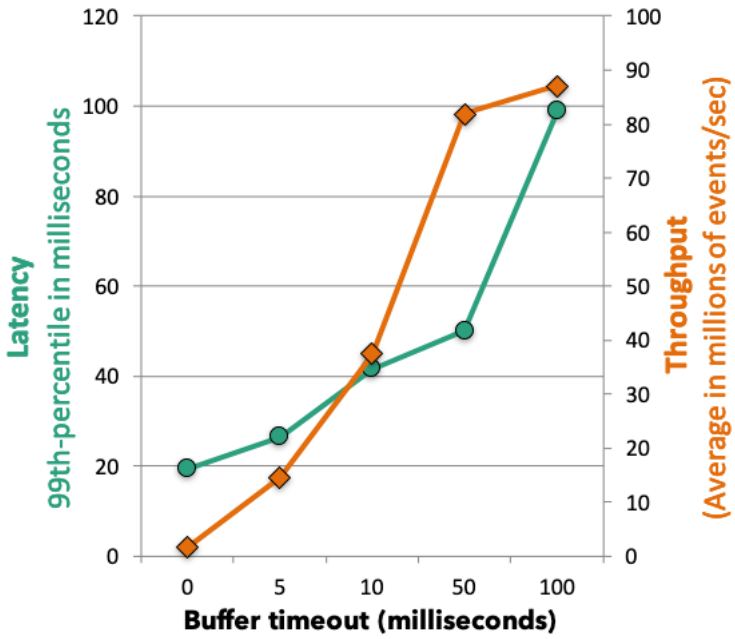
Flink: Distribtuted streaming dataflow to run streaming, batch, iterative

Distribtuted streaming dataflow

- Stateful operators
- Checkpointing based FT

DISCUSSION

https://forms.gle/j9Z7rm4qQpogbz5W8


Latency
99th-percentile in milliseconds
Throughput
(Average in millions of events/sec)
Buffer timeout (milliseconds)
0
5
10
50
100
0
20
40
60
80
100
120
0
10
20
30
40
50
60
70
80
90
100

Consider you are implementing a micro-batch streaming API on top of Apache Spark. What are some of the bottlenecks/challenges you might have in building such a system?

SUMMARY

Next week: Spring break!!

Next class: Spark Streaming