JASMINE

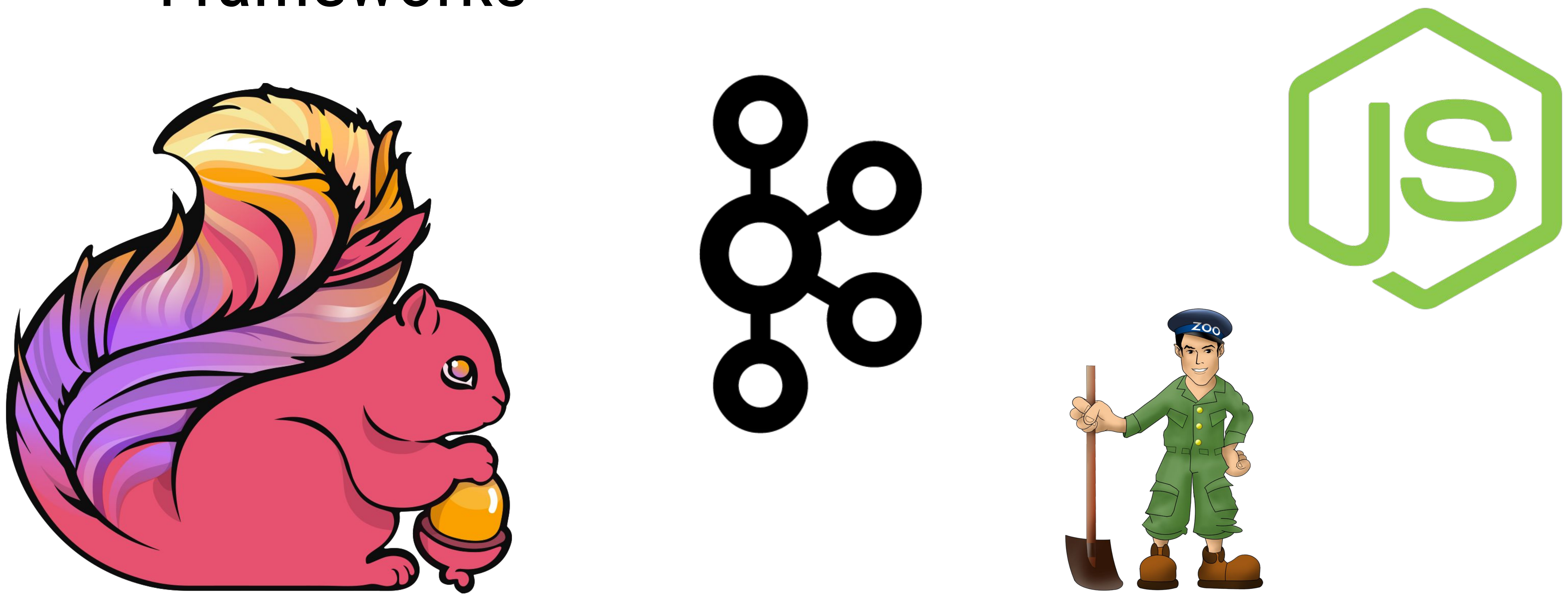
Data Stream Processing

Simone Mancini   Francesco Ottaviano   Andrea Silvi

JASMINE

Design Choices

Frameworks

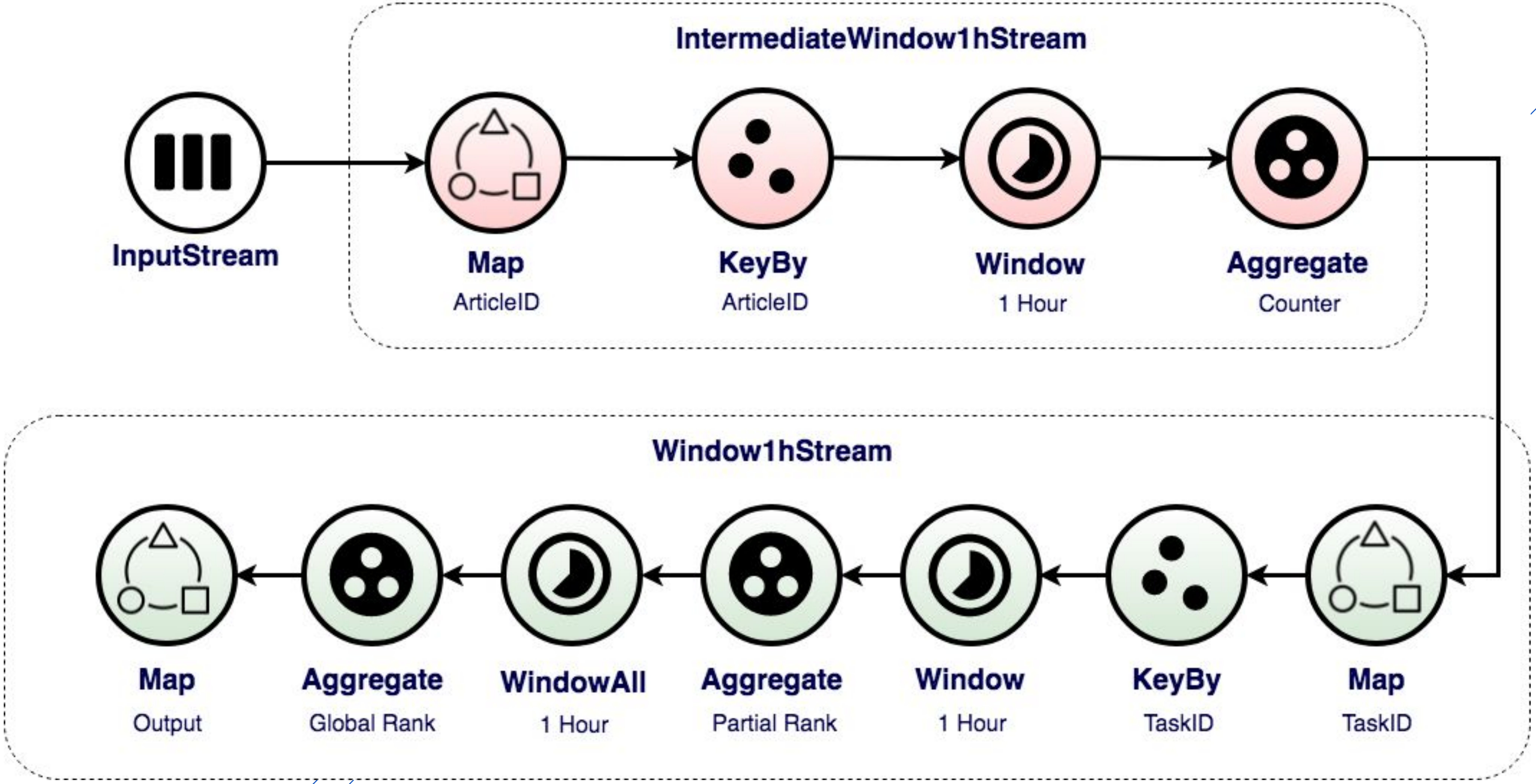


Programming Languages

Store

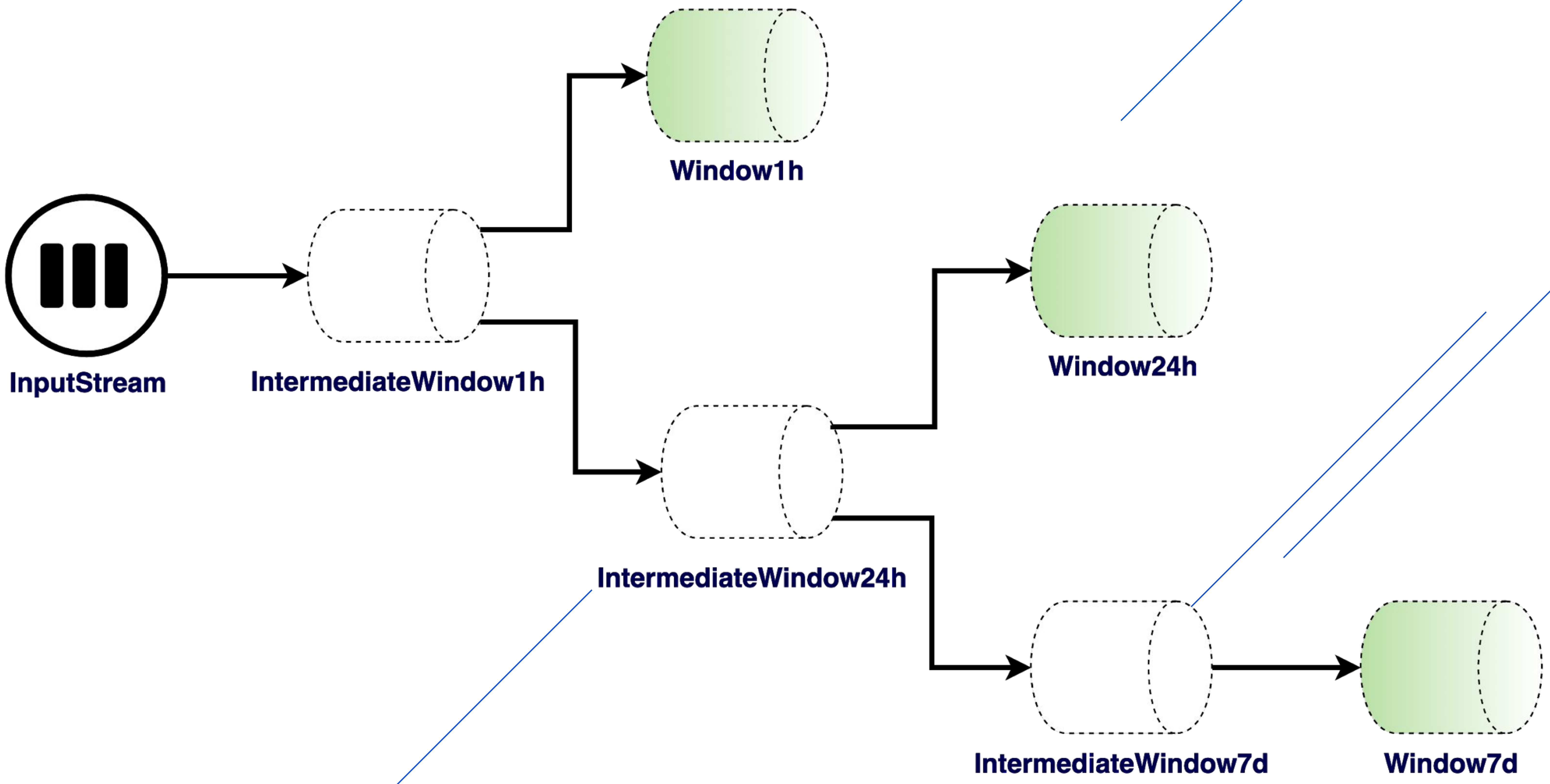
Queries

Query 1 - DAG


IntermediateWindow1hStream
InputStream
Map
ArticleID
KeyBy
ArticleID
Window
1 Hour
Aggregate
Counter
Window1hStream
Map
Output
Aggregate
Global Rank
WindowAll
1 Hour
Aggregate
Partial Rank
Window
1 Hour
KeyBy
TaskID
Map
TaskID

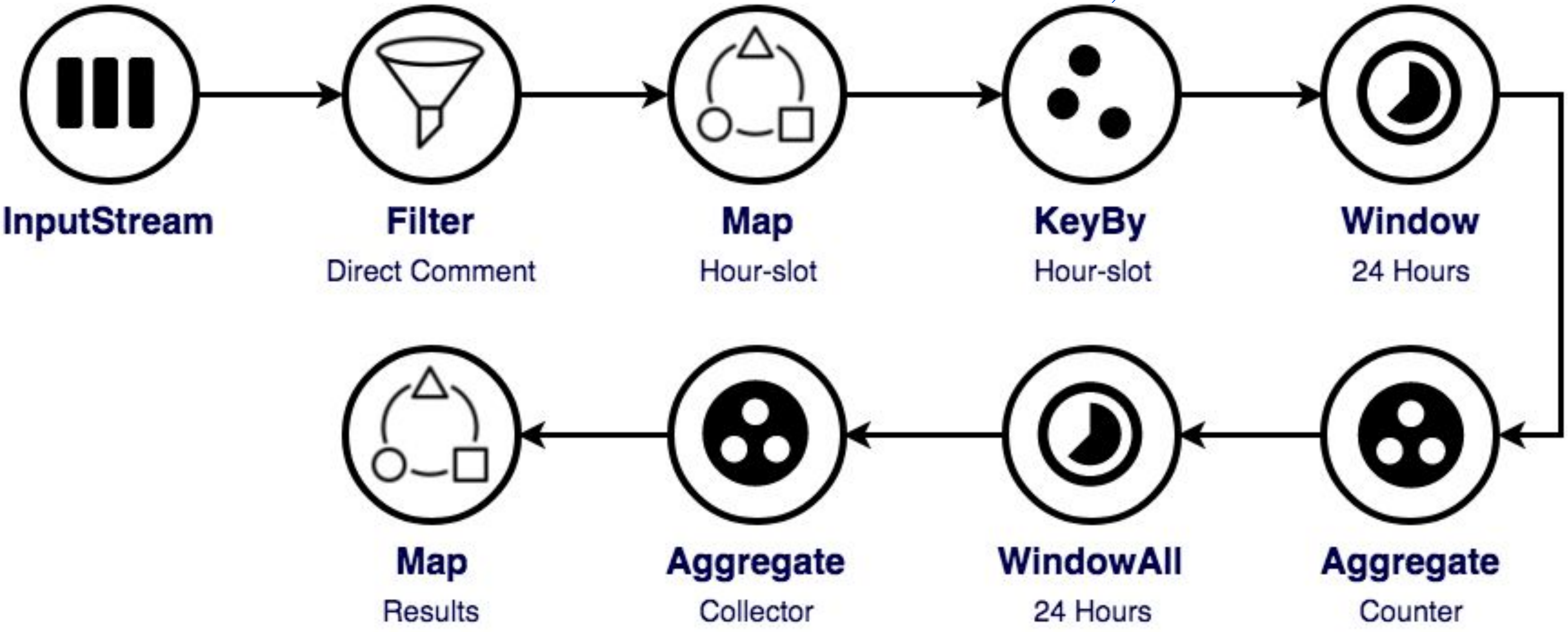
JASMINE

Query 1 - Chaining

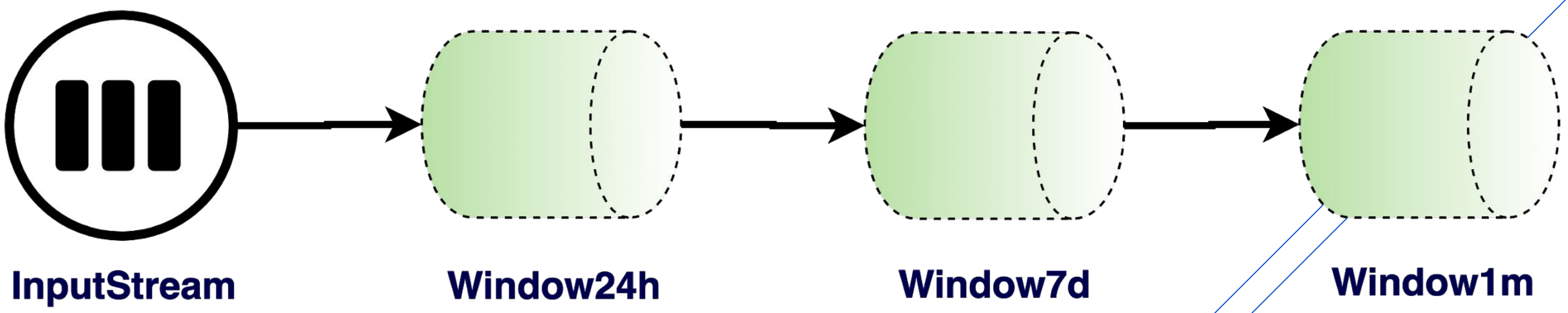


JASMINE

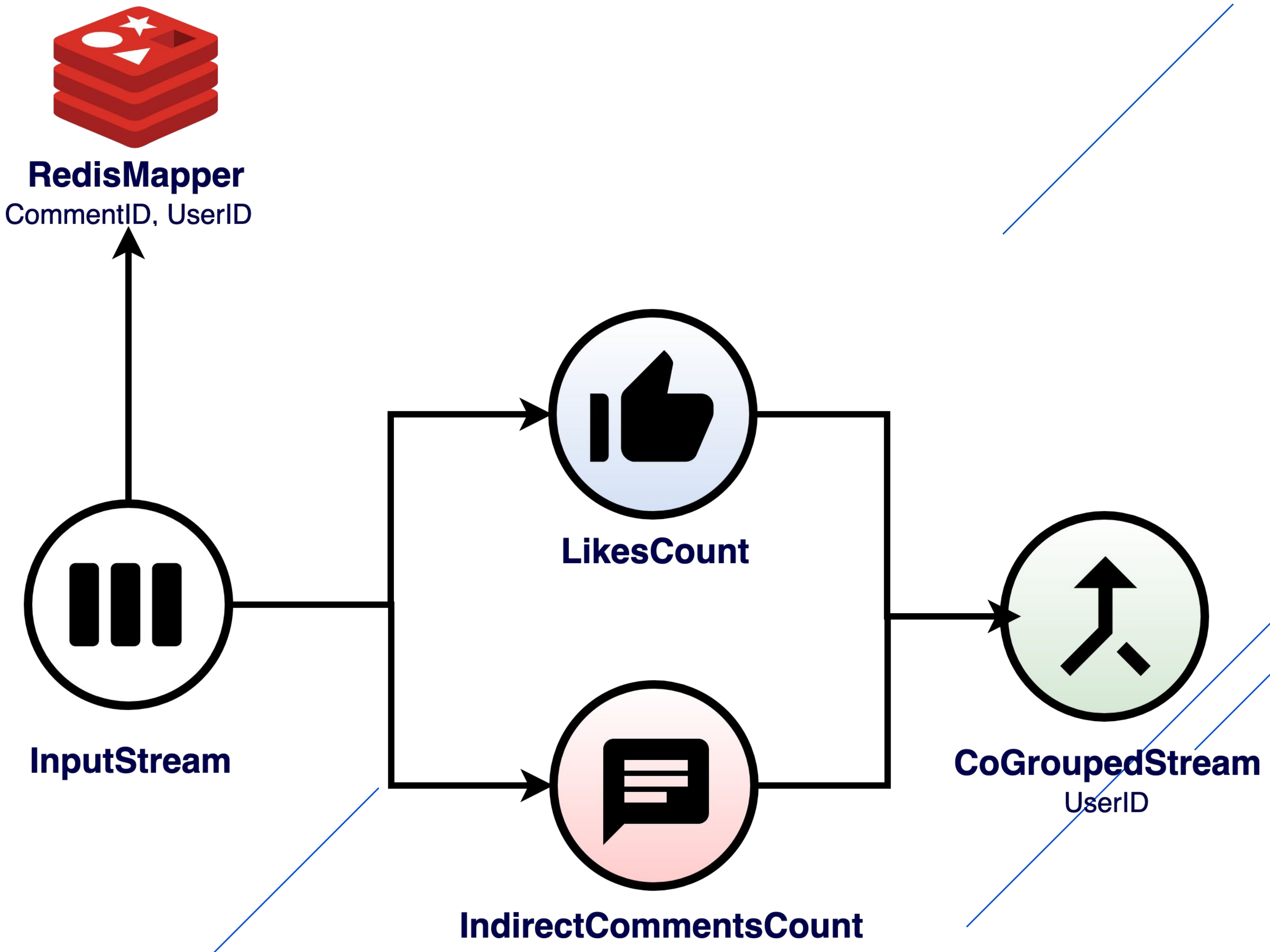
Query 2 - DAG



Query 2 - Chaining

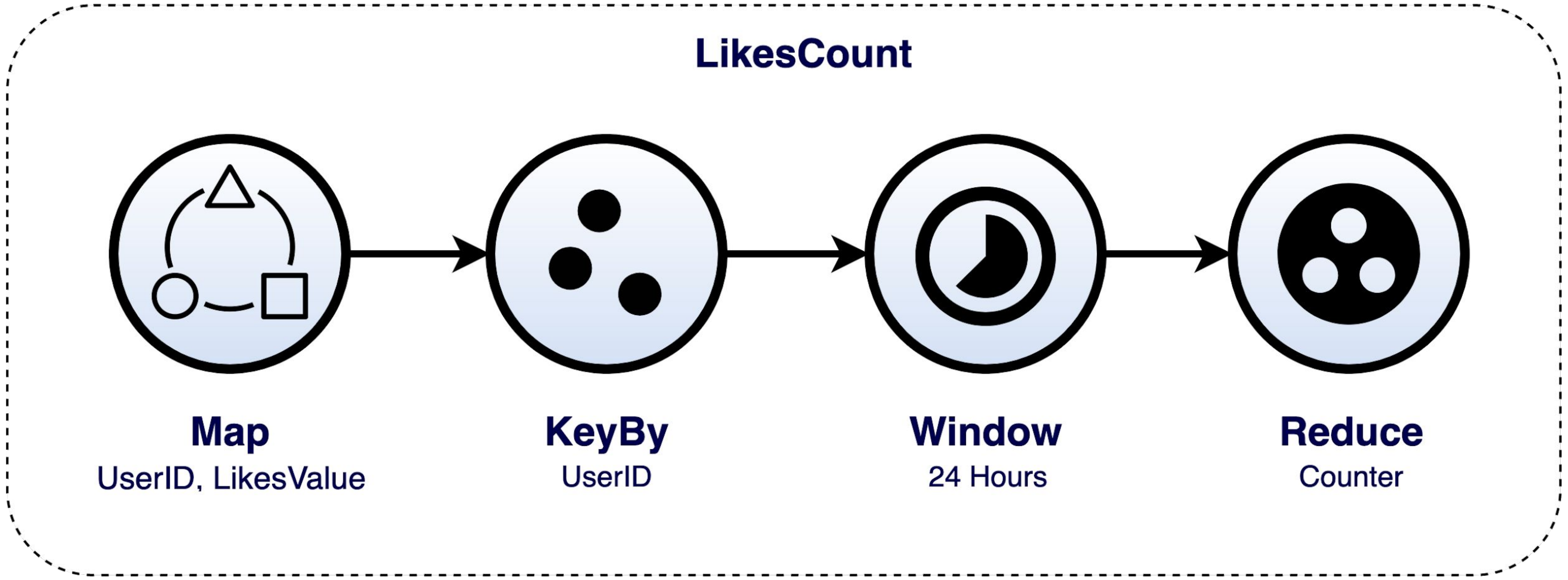


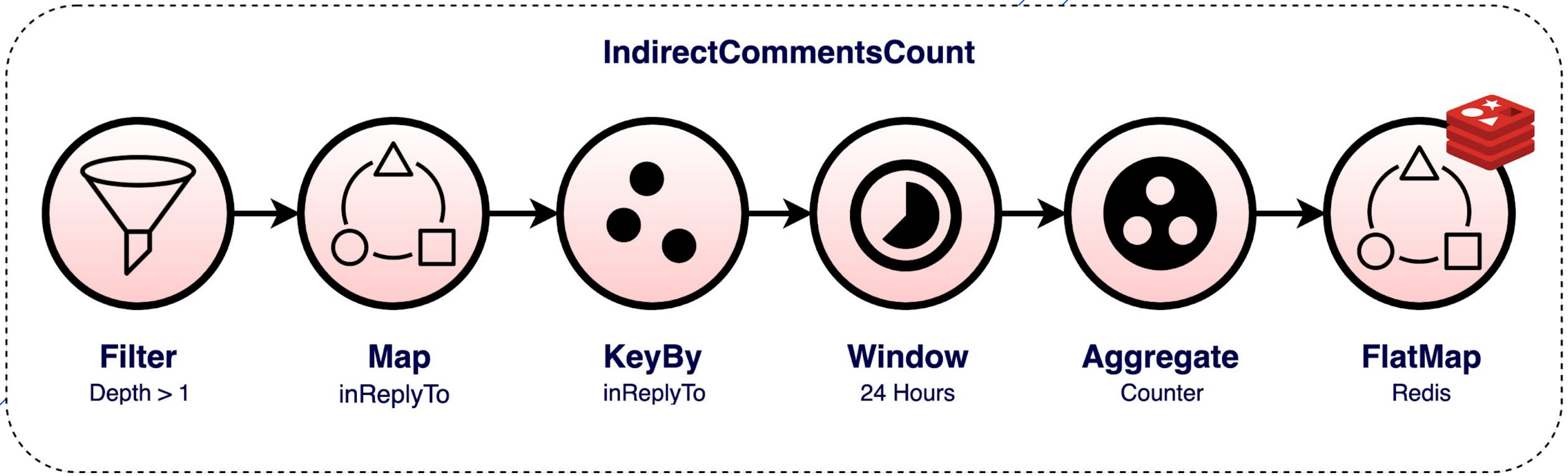
Query 3 - DAG


RedisMapper
CommentID, UserID
InputStream
LikesCount
IndirectCommentsCount
CoGroupedStream
UserID

JASMINE

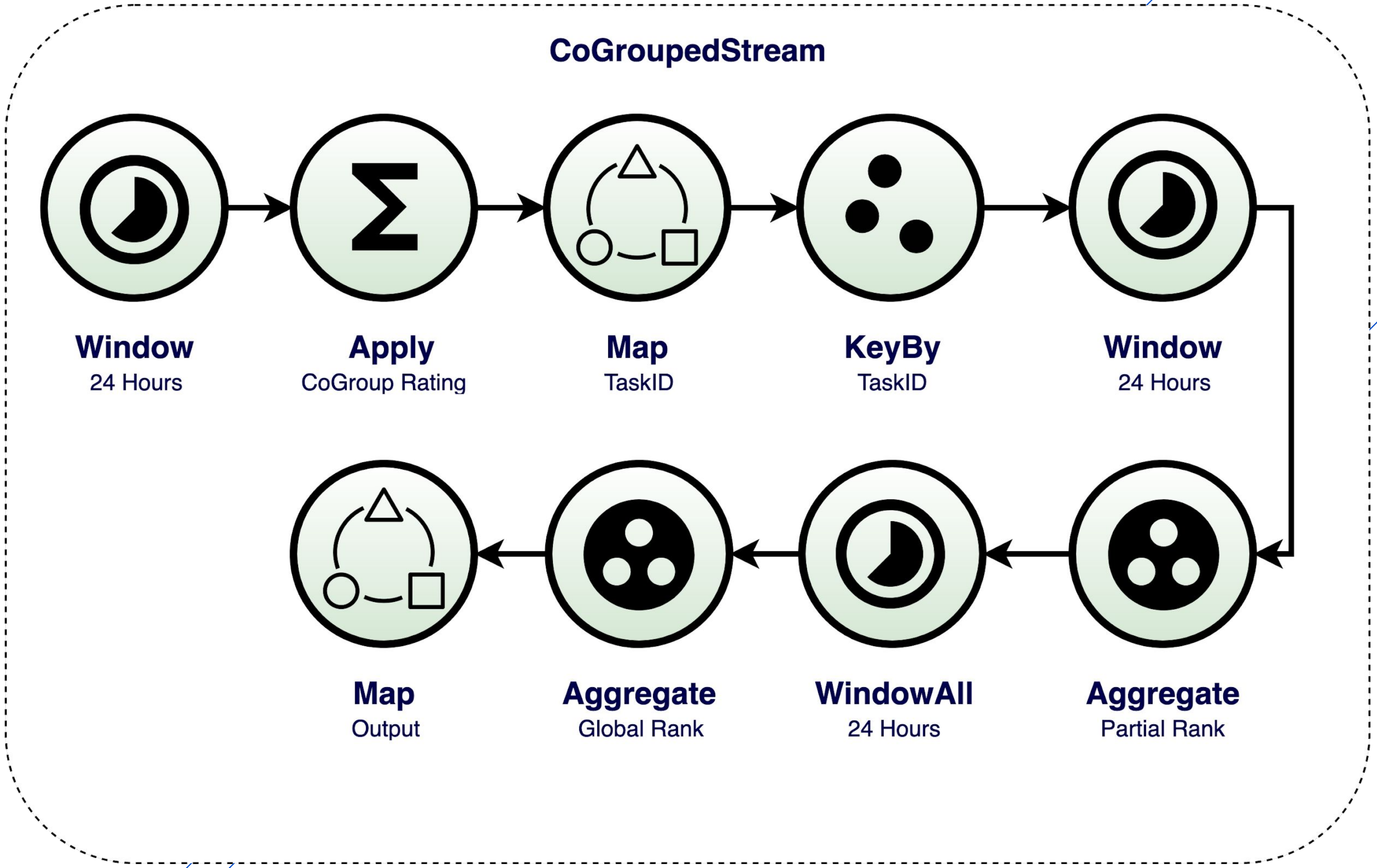
Query 3 - DAG





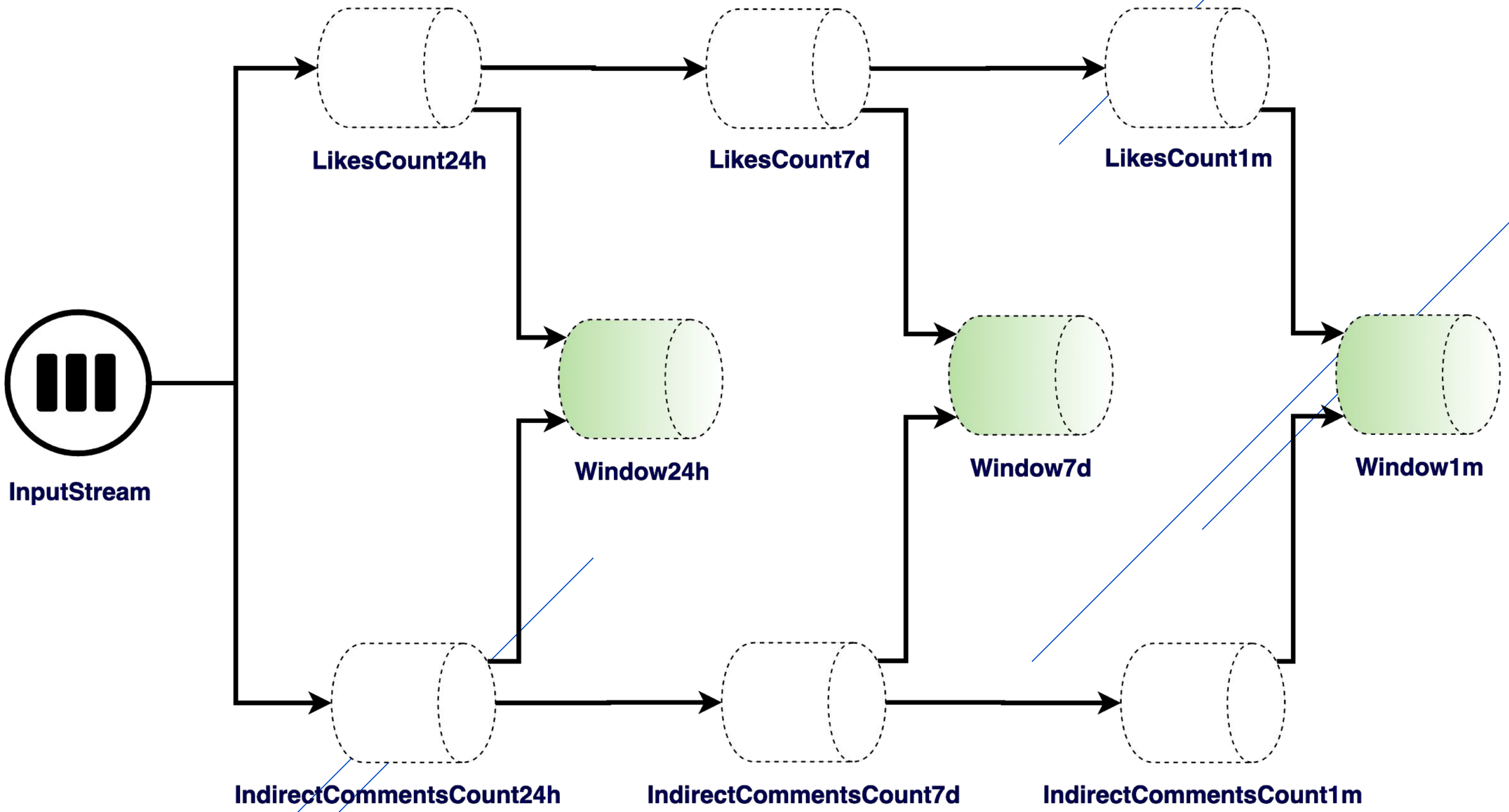
JASMINE

Query 3 - DAG



JASMINE

Query 3 - Chaining



Kafka Streams

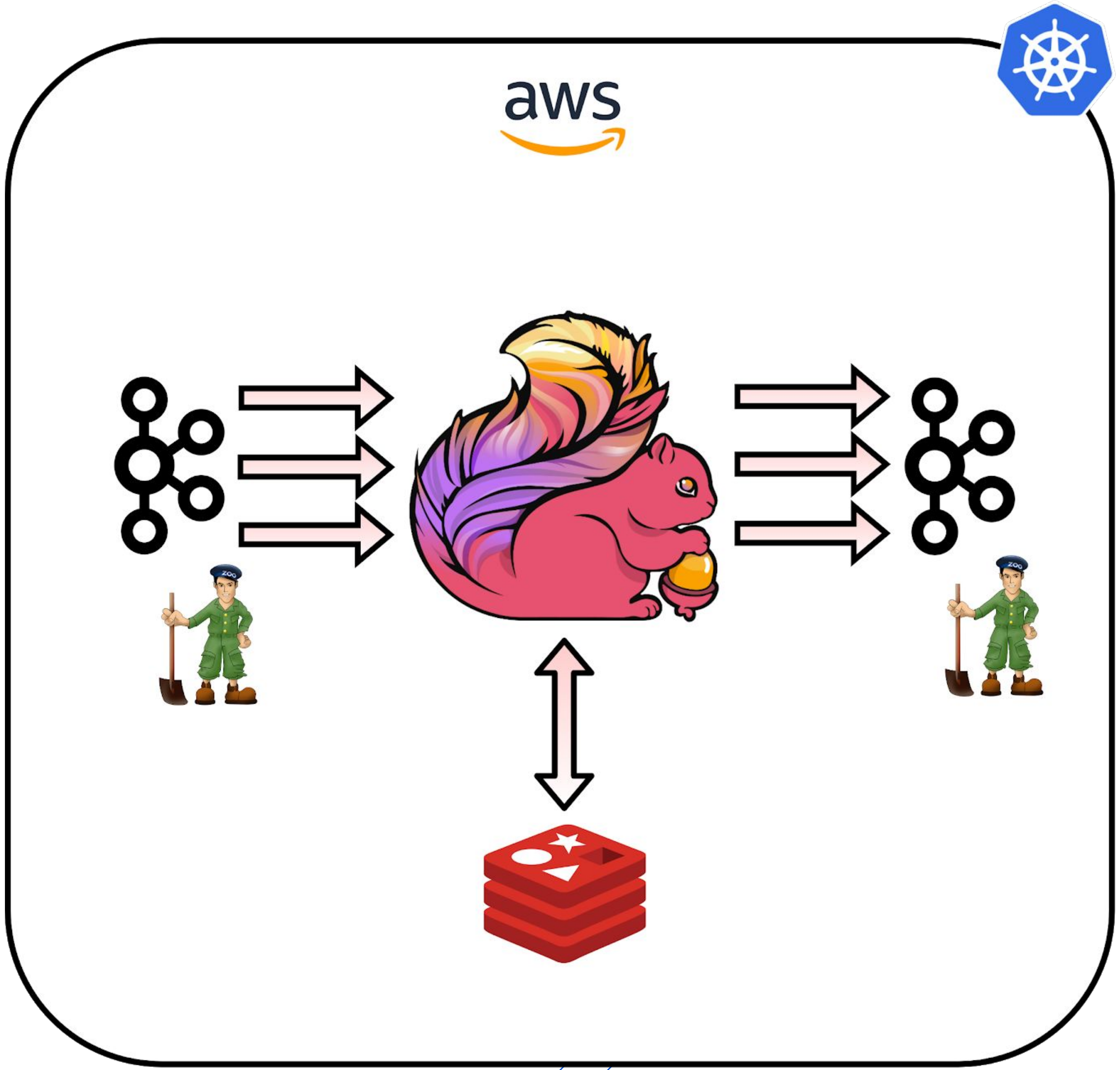
JASMINE

KAFKA STREAMS

- Different approach: *Kstreams* and *Ktable*
- Different ranking solution
- Parallelism
- *Suppress* in chaining mode limitations
- Custom Serde

Architecture

Architecture



- *m4.large EC2 instances*
  - 2,4 GHz Intel Xeon E5-2676 v3 Processor
  - 2 vCPU, 8GiB Mem
- *Kubernetes*
  - *Kafka*: 1 node per machine, 3 nodes with replication factor 3 on topics.
  - *Zookeeper*: 1 node per machine, 3 nodes.
  - *Flink*: 1 job manager, 8 task manager.

JASMINE

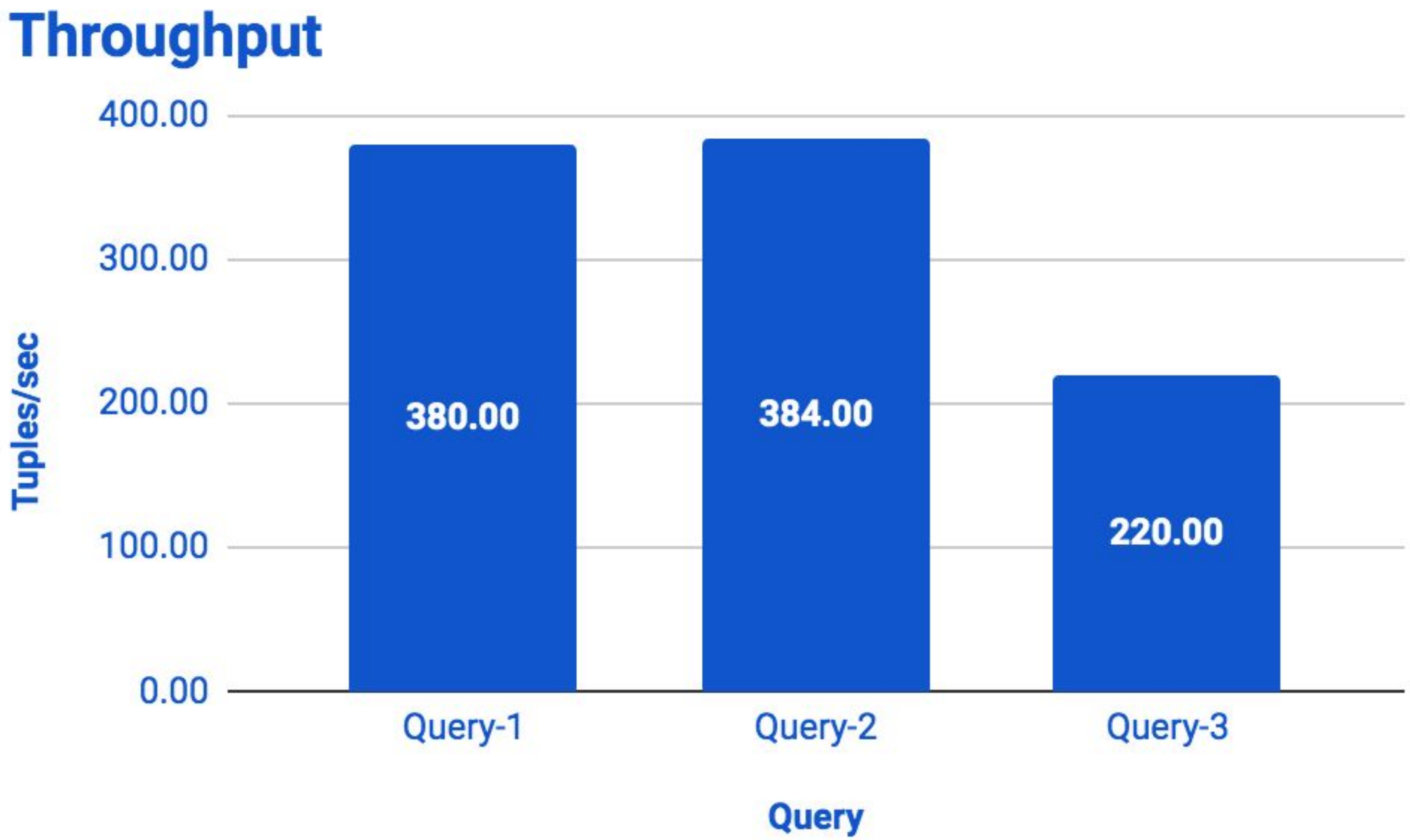
Simulator

- Emulates data stream processing compressing time
- *NodeJS*
- *CreateDate* as tuple creation time

Evaluations

JASMINE

Apache Flink - Throughput

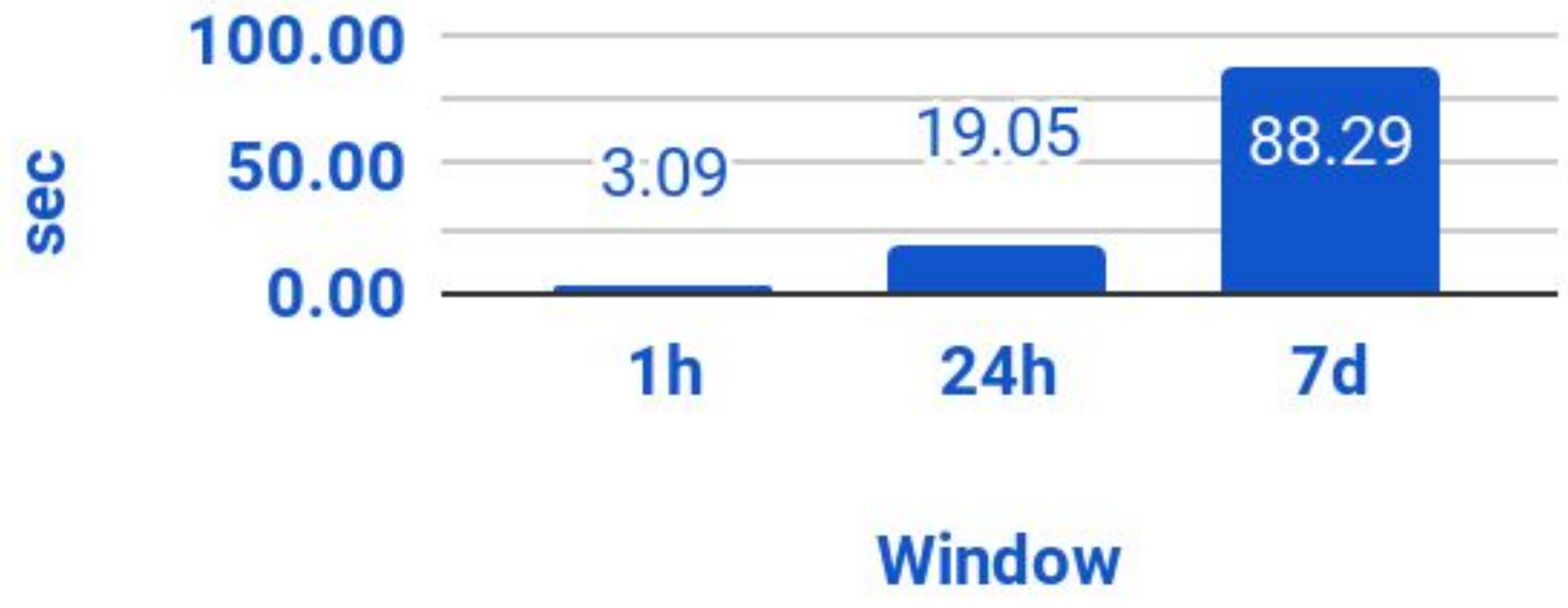


JASMINE

Apache Flink - Latency

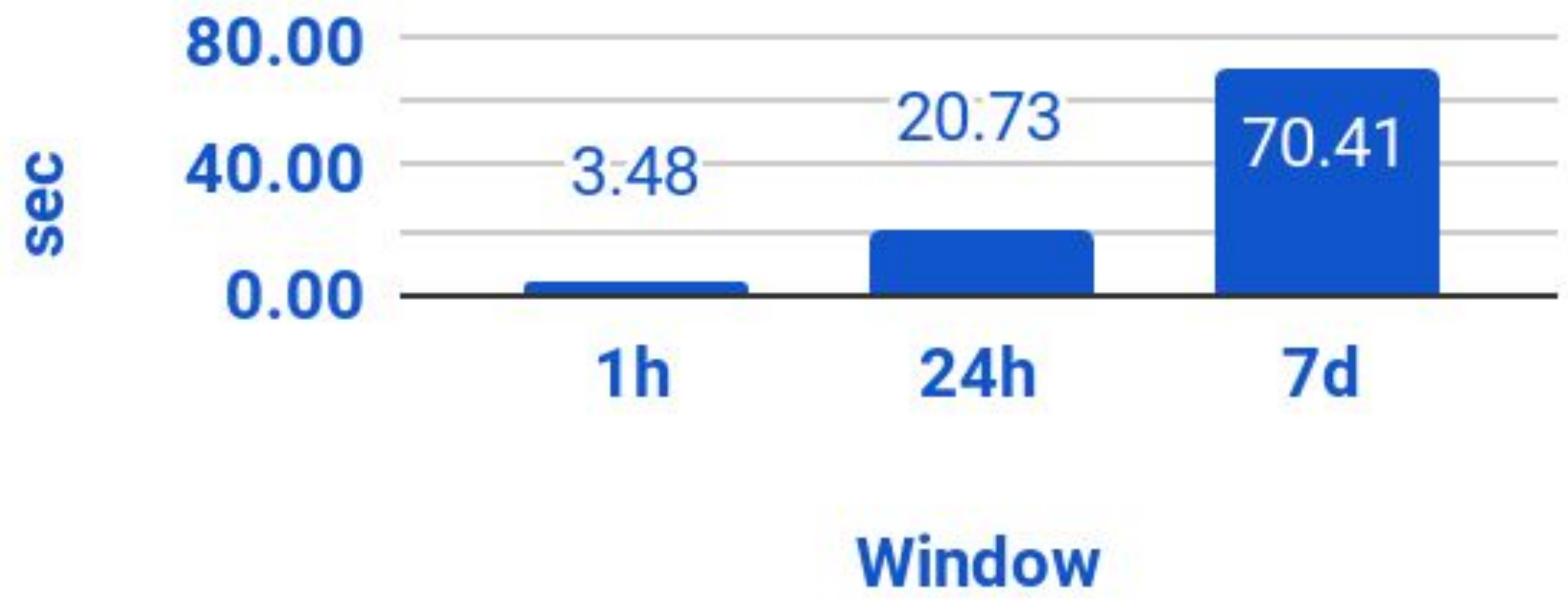
Query 1

10.000 compression



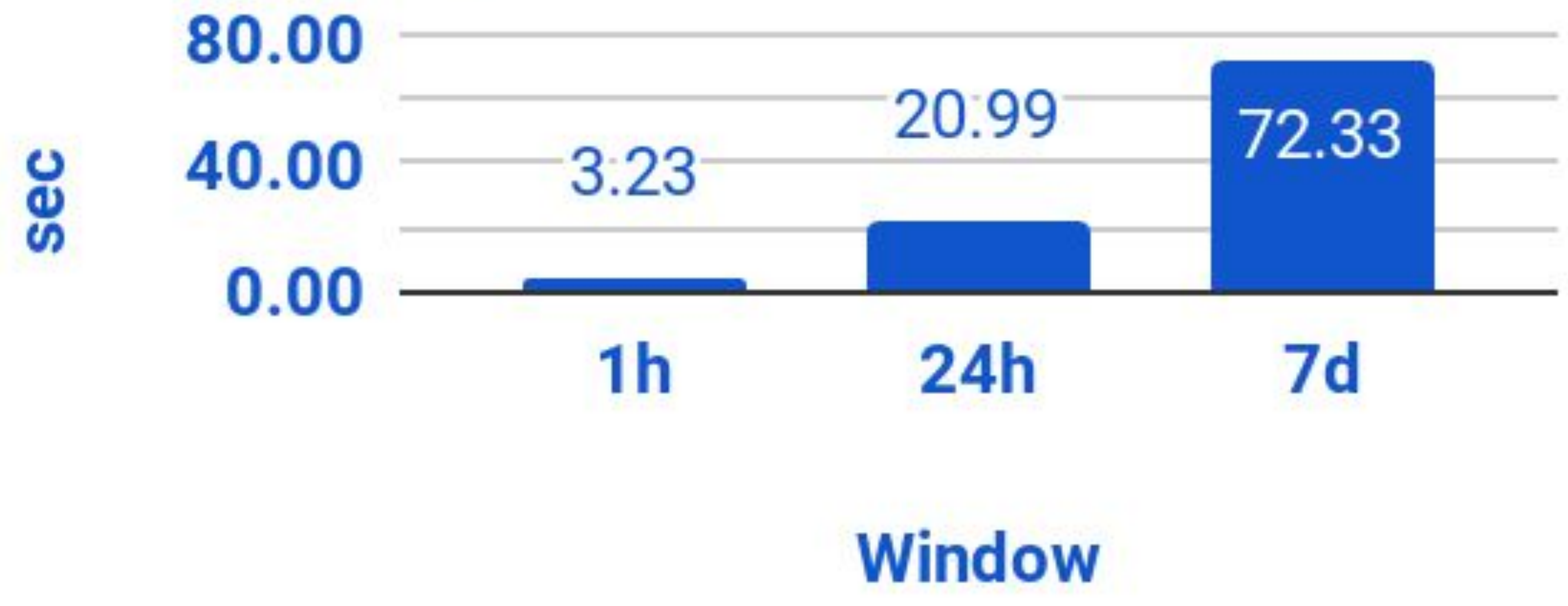
Query 1

50.000 compression



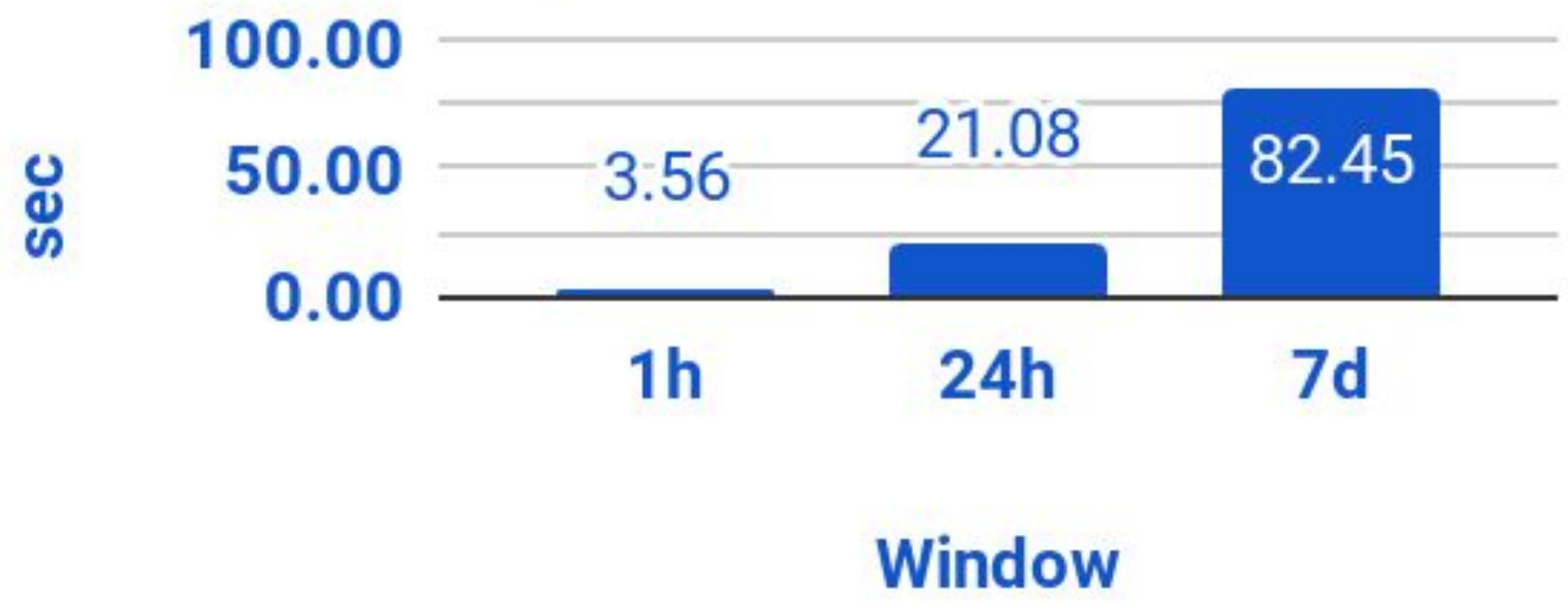
Query 1

100.000 compression

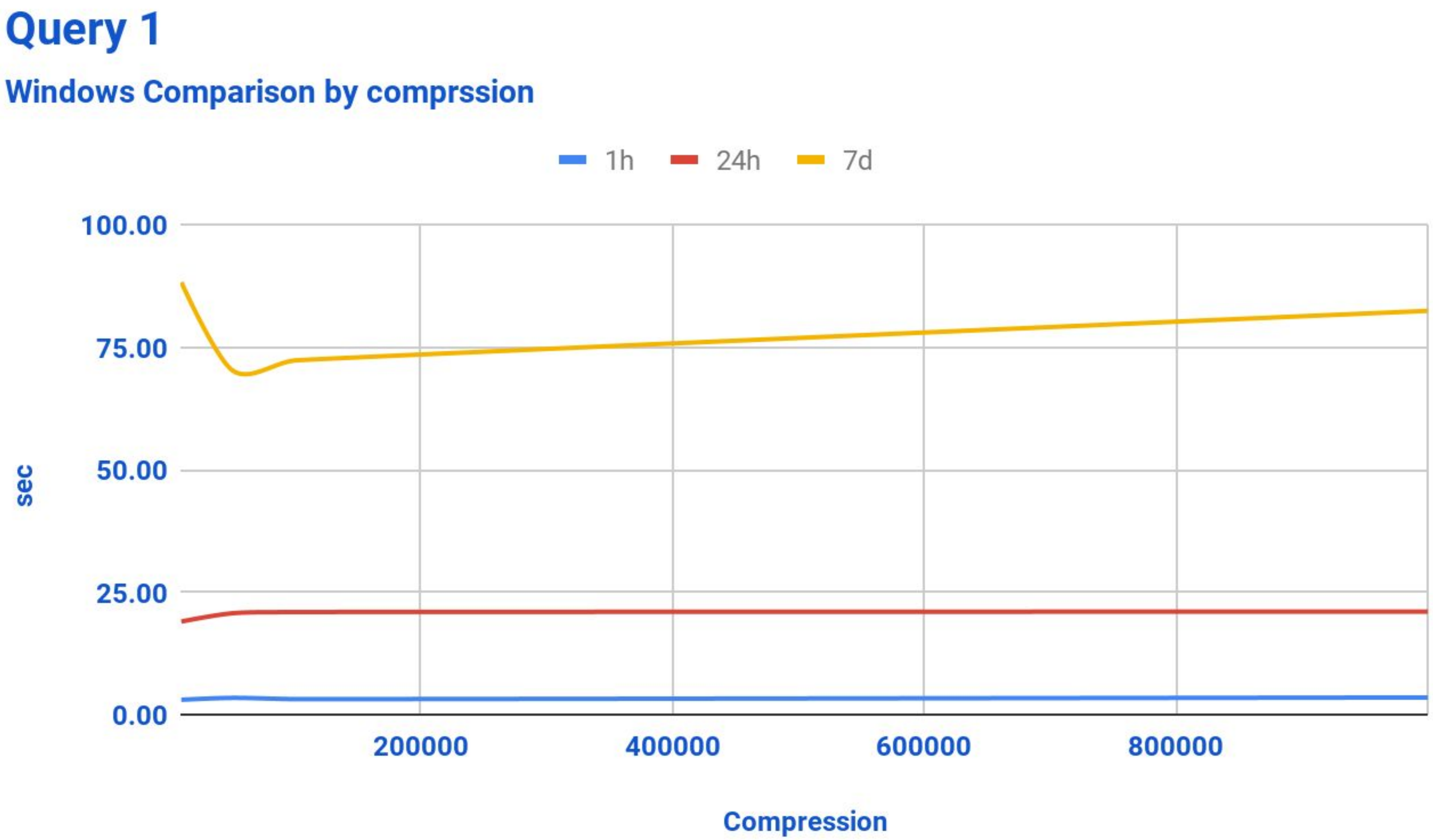


Query 1

1.000.000 compression



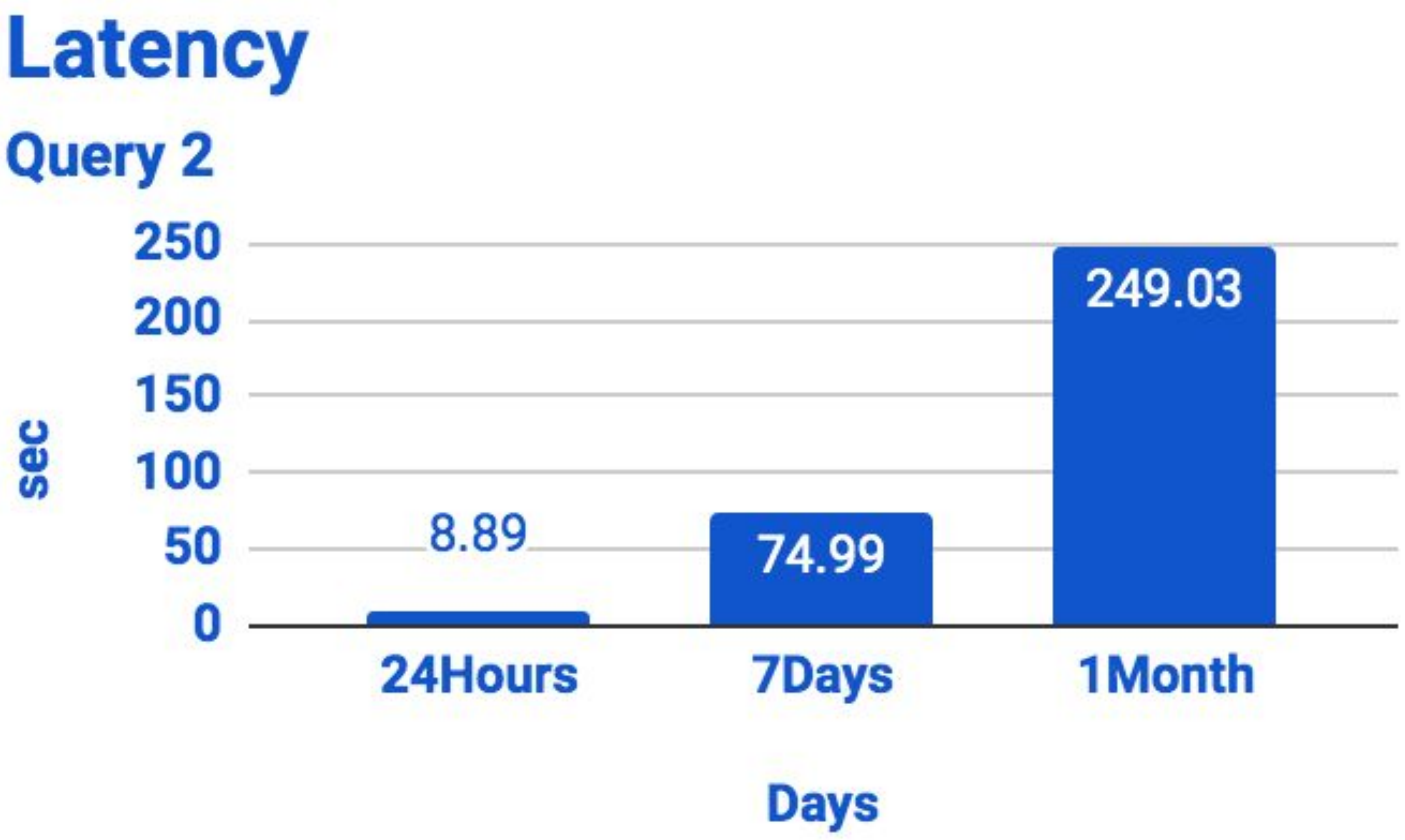
Apache Flink - Latency

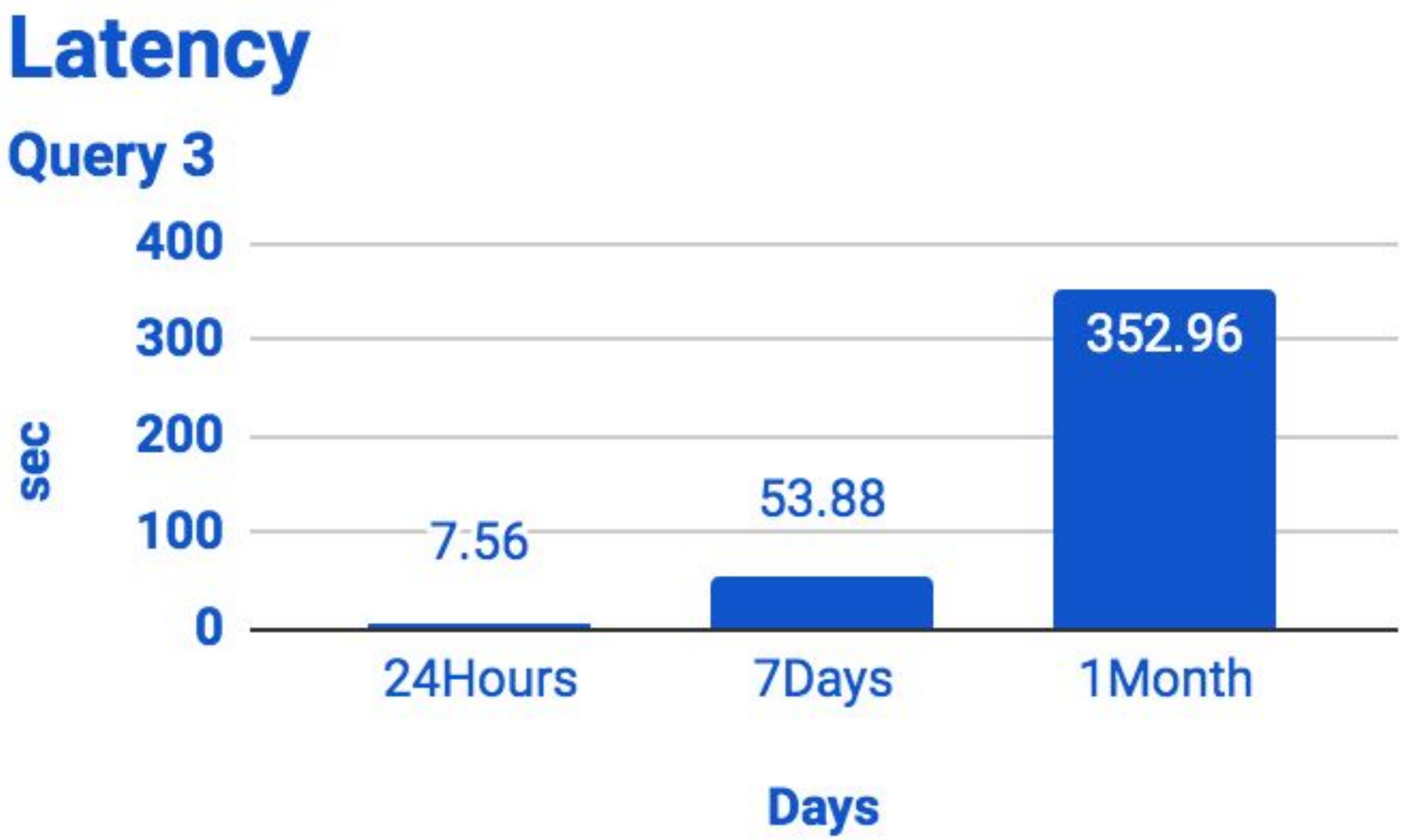


Query 1

Windows Comparison by comprssion

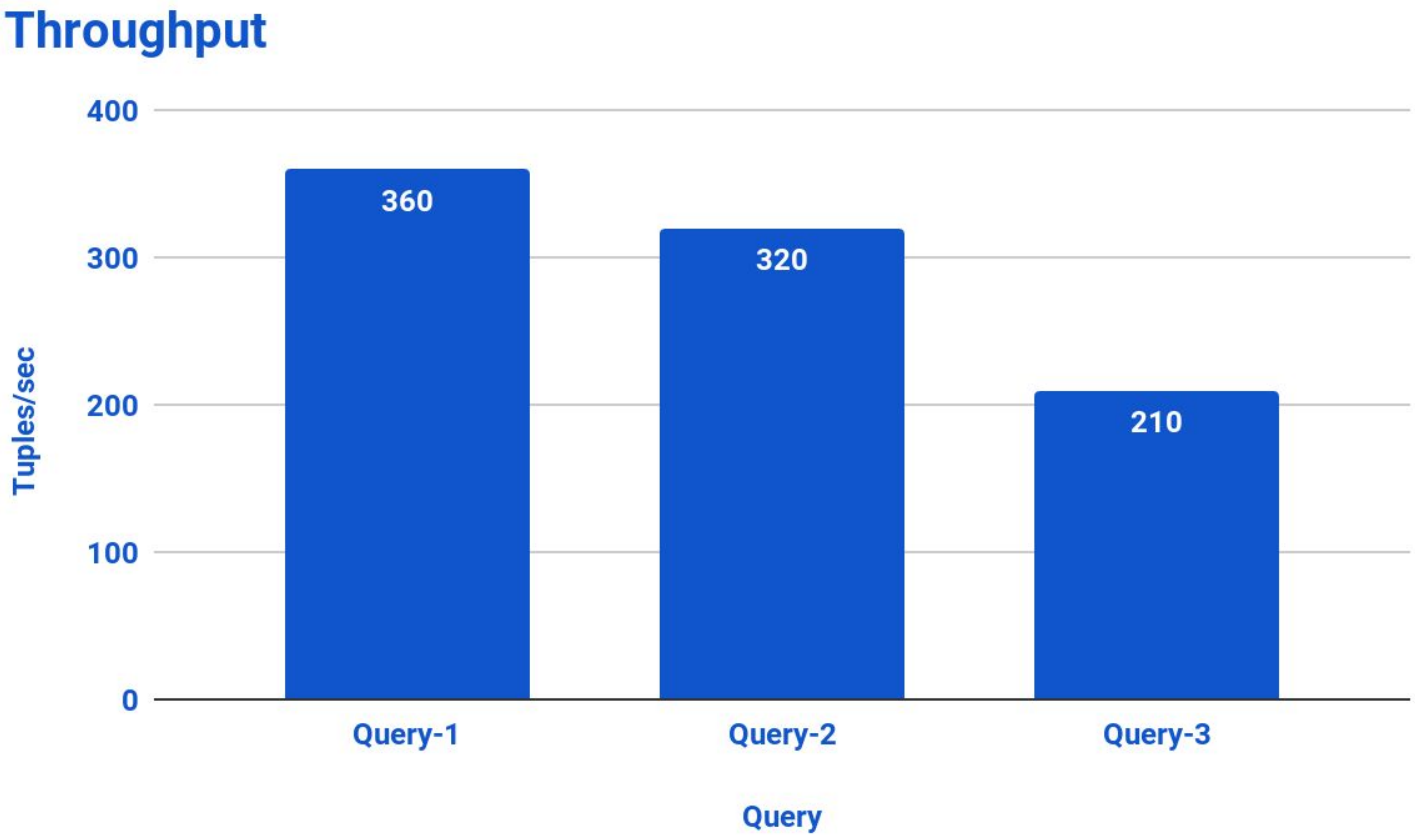
Apache Flink - Latency





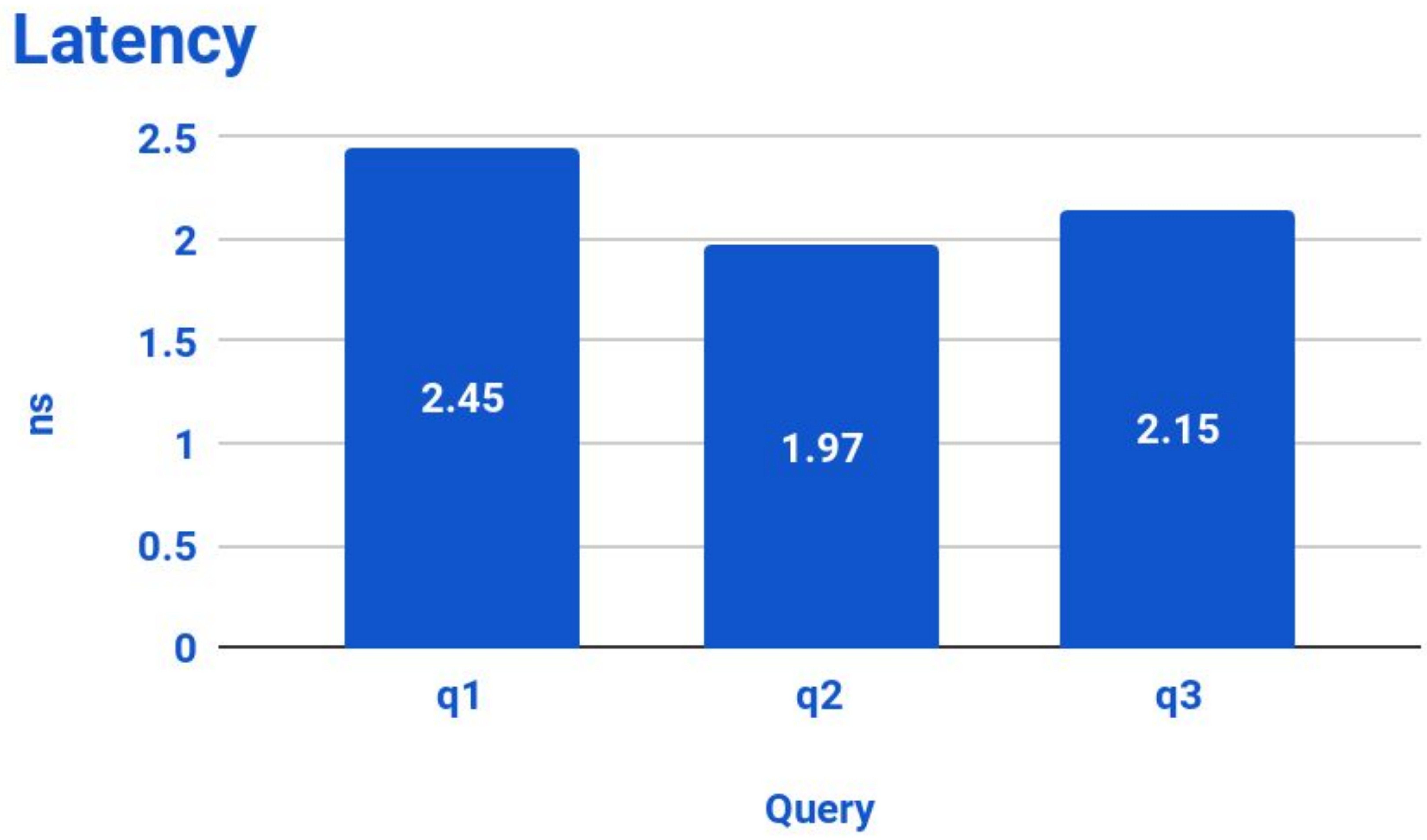
JASMINE

KafkaStreams - Throughput

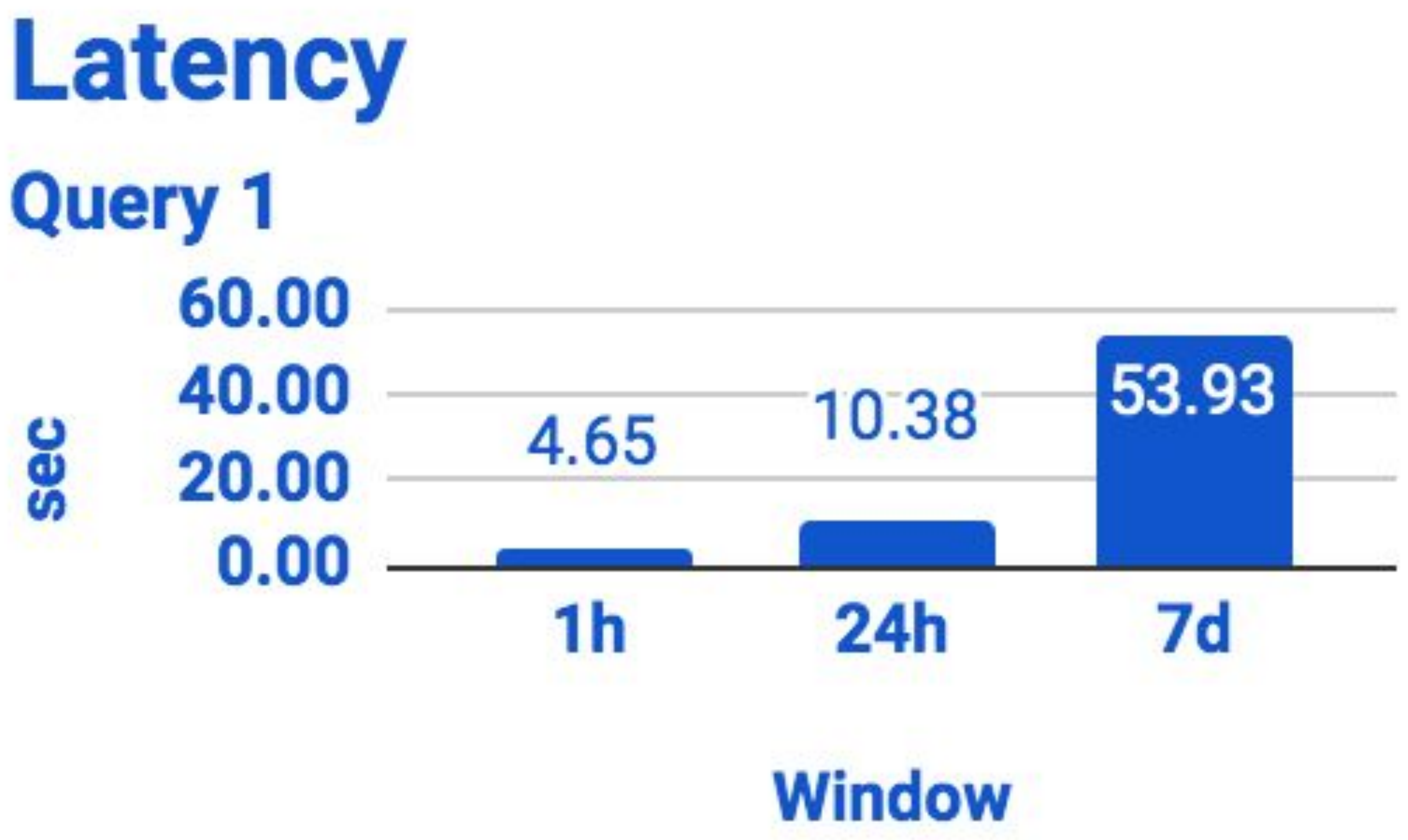


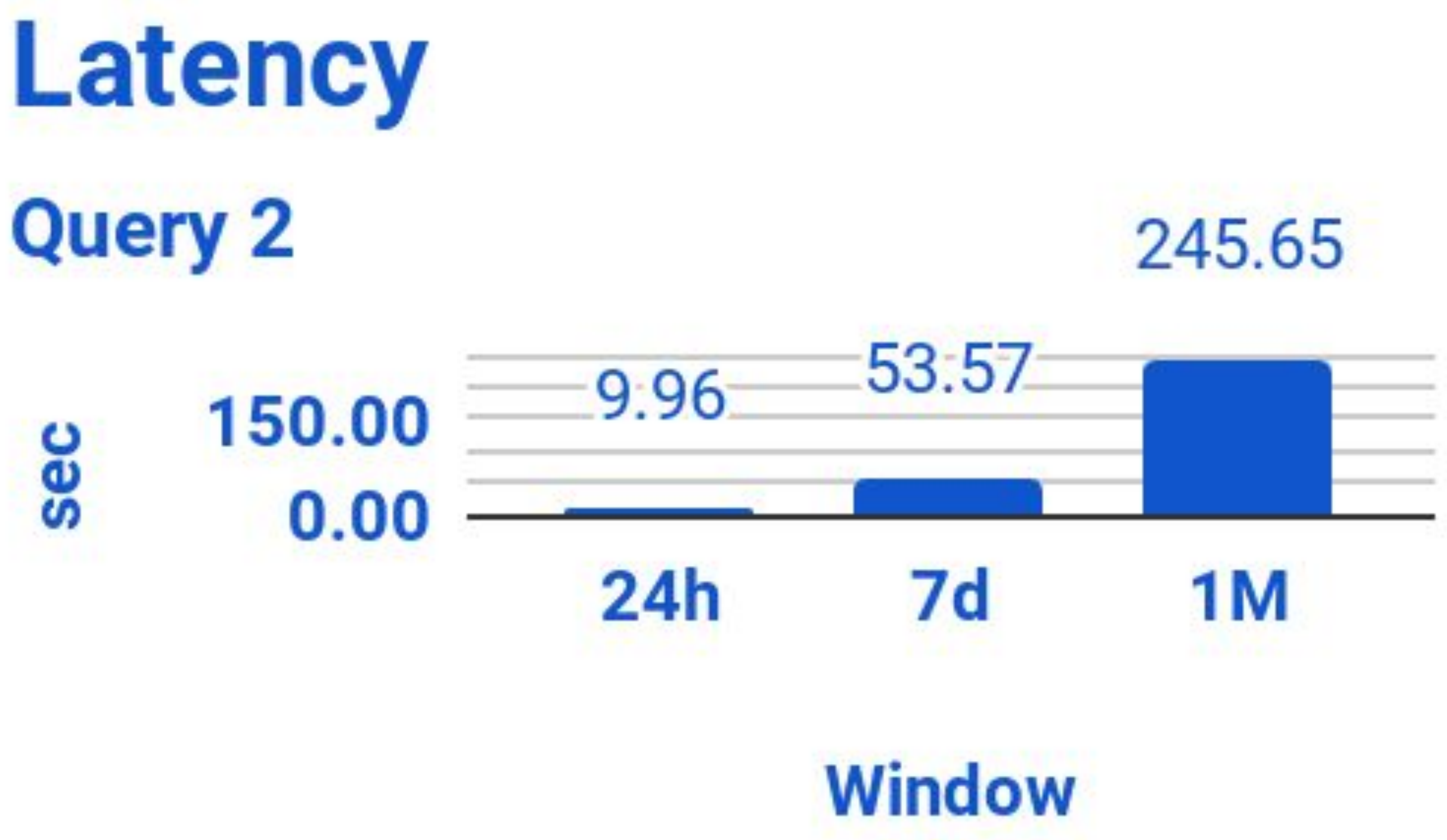
JASMINE

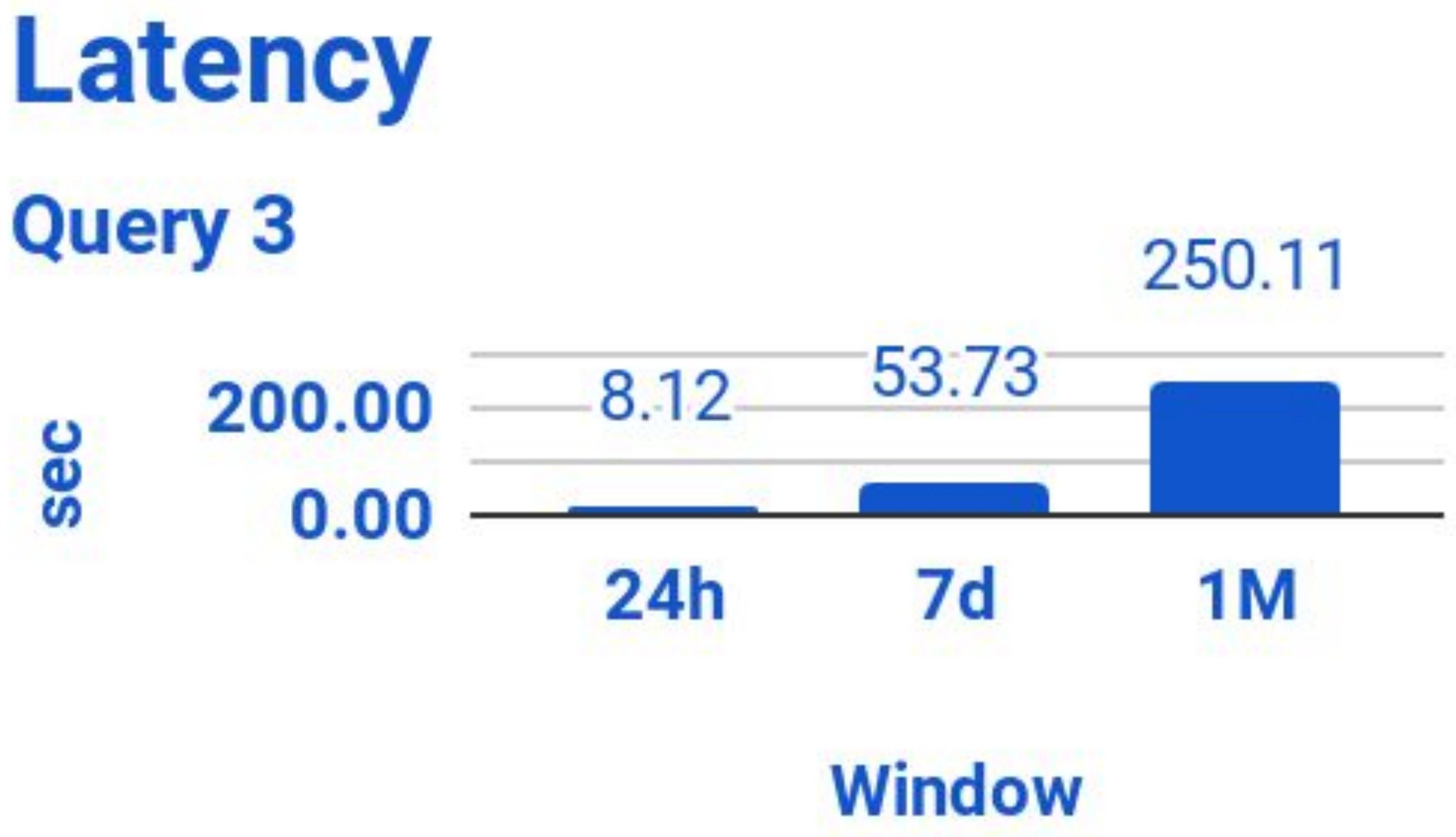
KafkaStreams - Latency

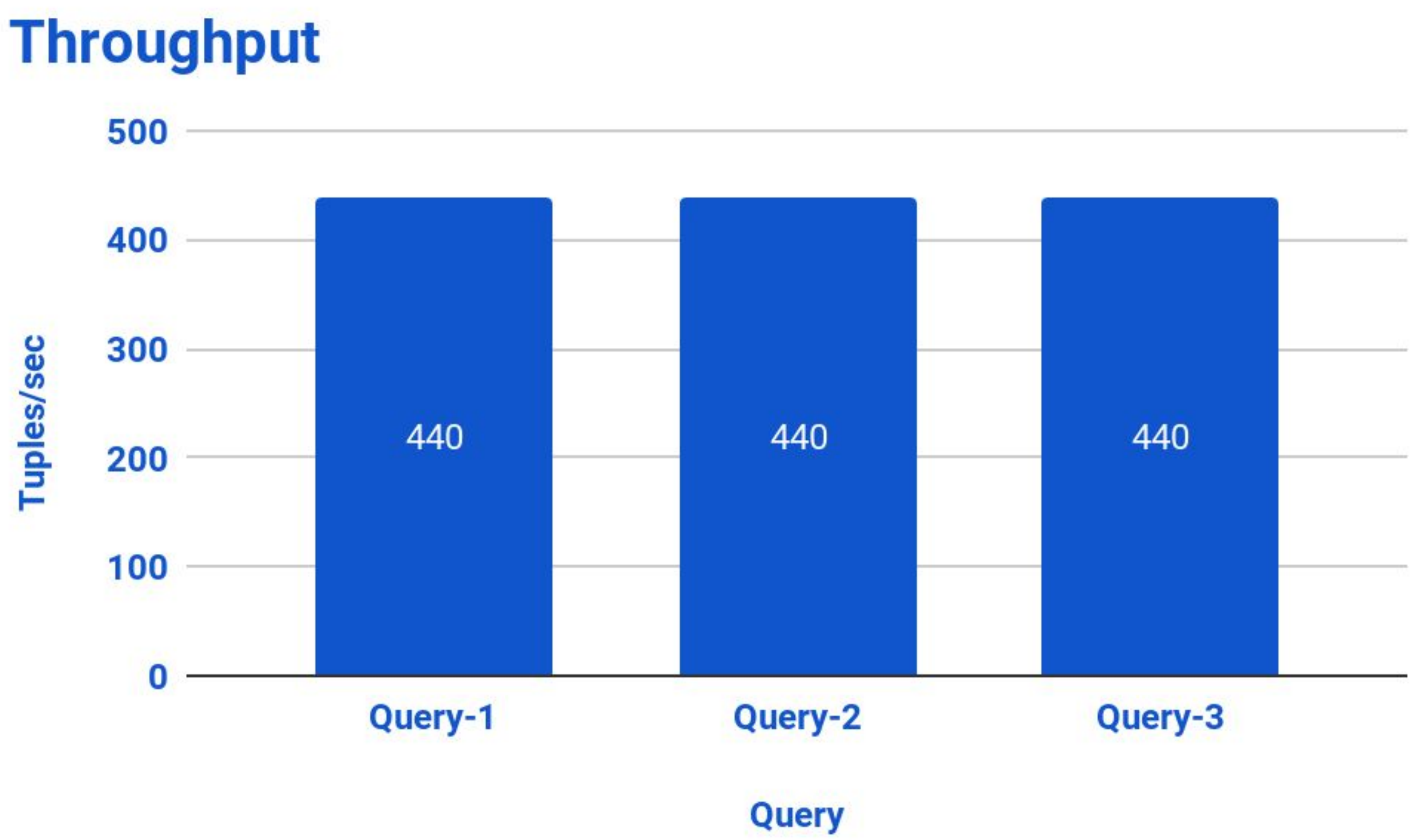


Apache Flink - Cluster evaluation









JASMINE

Thank You!