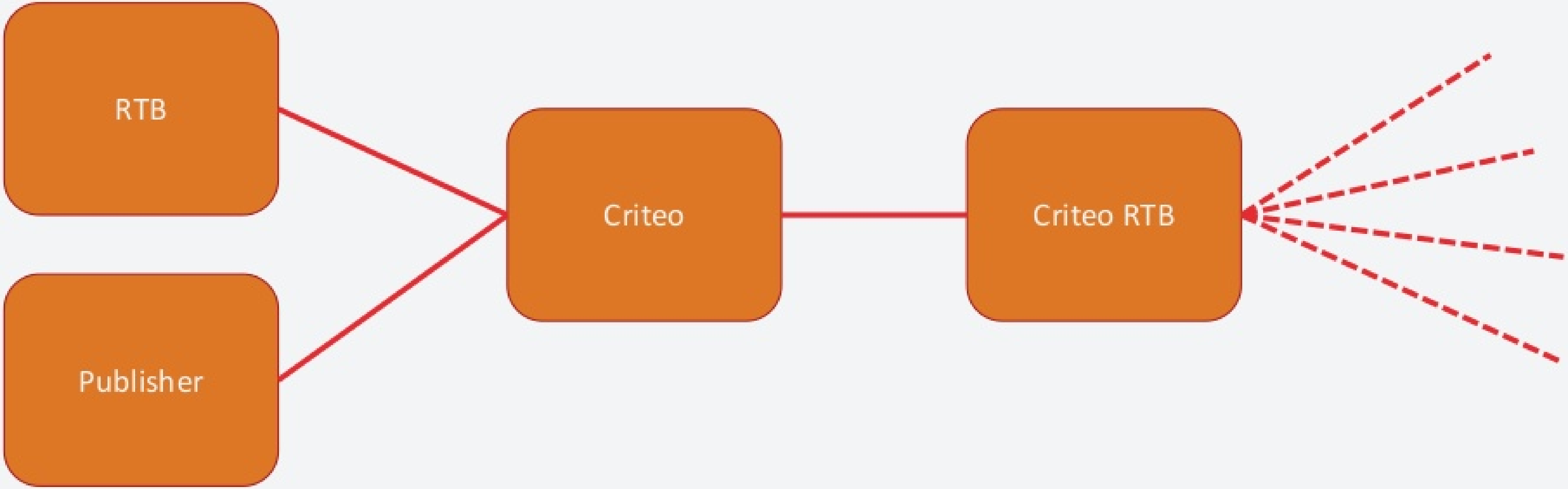
Data LossLess streaming processing

Oleksandr Nitavskyi, Criteo

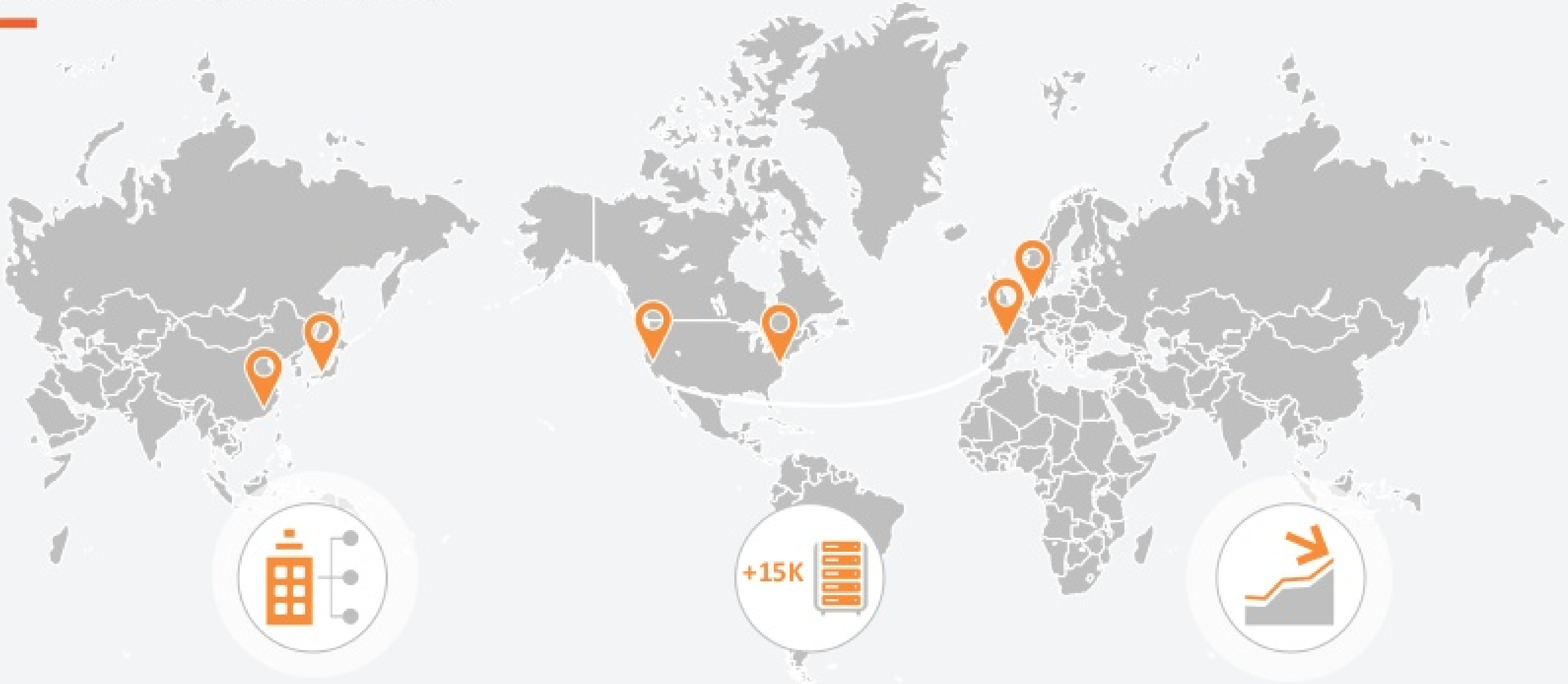
criteo

“Connecting shoppers to the things they need and love.”

Criteo: business model



Criteo: in numbers



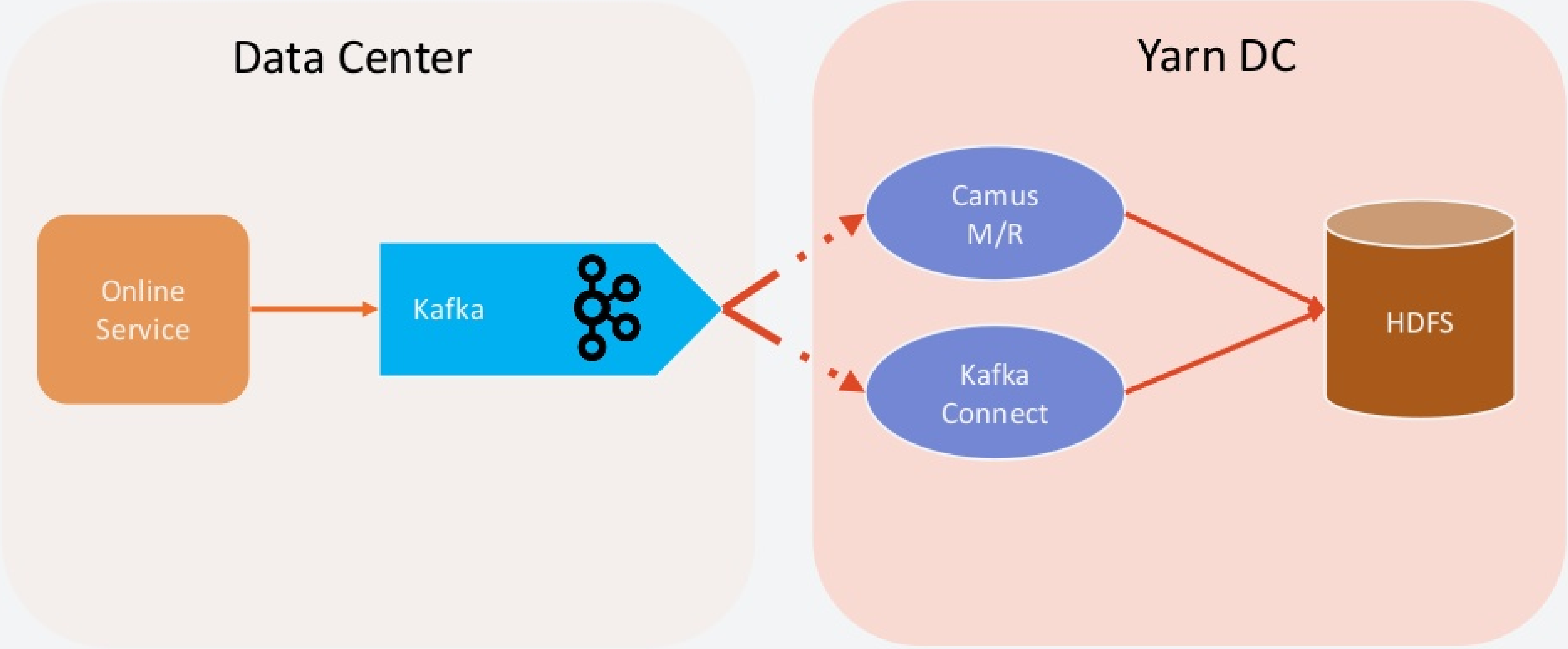
16 DATA CENTERS

MORE THAN
25 000 SERVERS

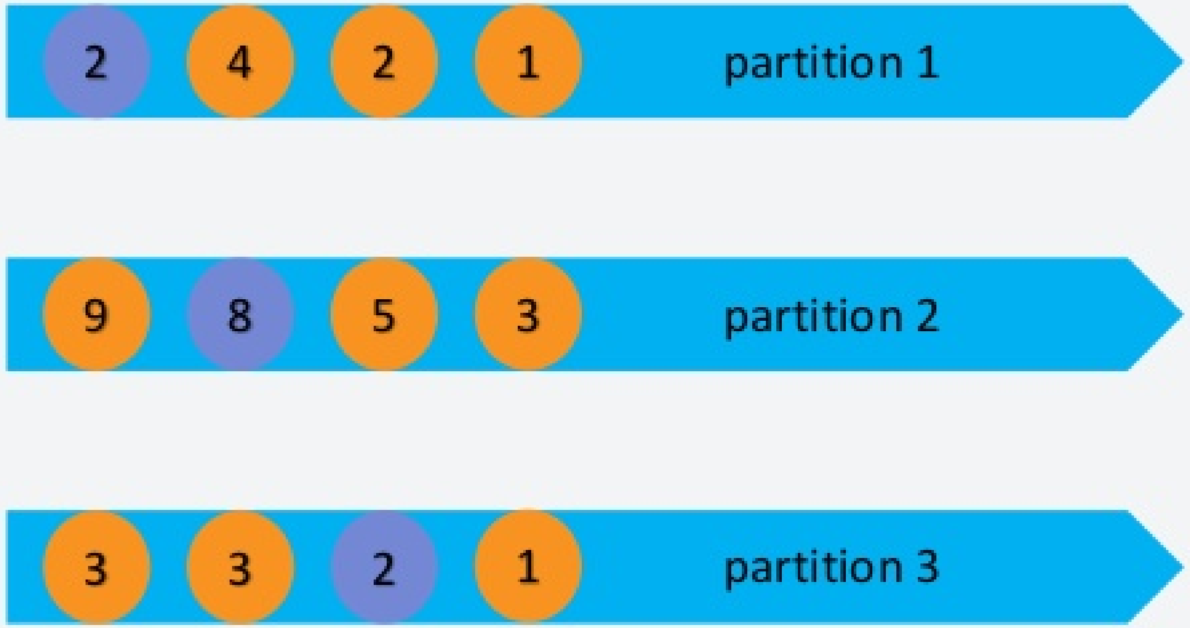
PEAK TRAFFIC > 4 mln QPS

2 CDH5 CLUSTER OF 4000 NODES (92TB, 200 GB RAM, 48 cores)

Gathering all data together

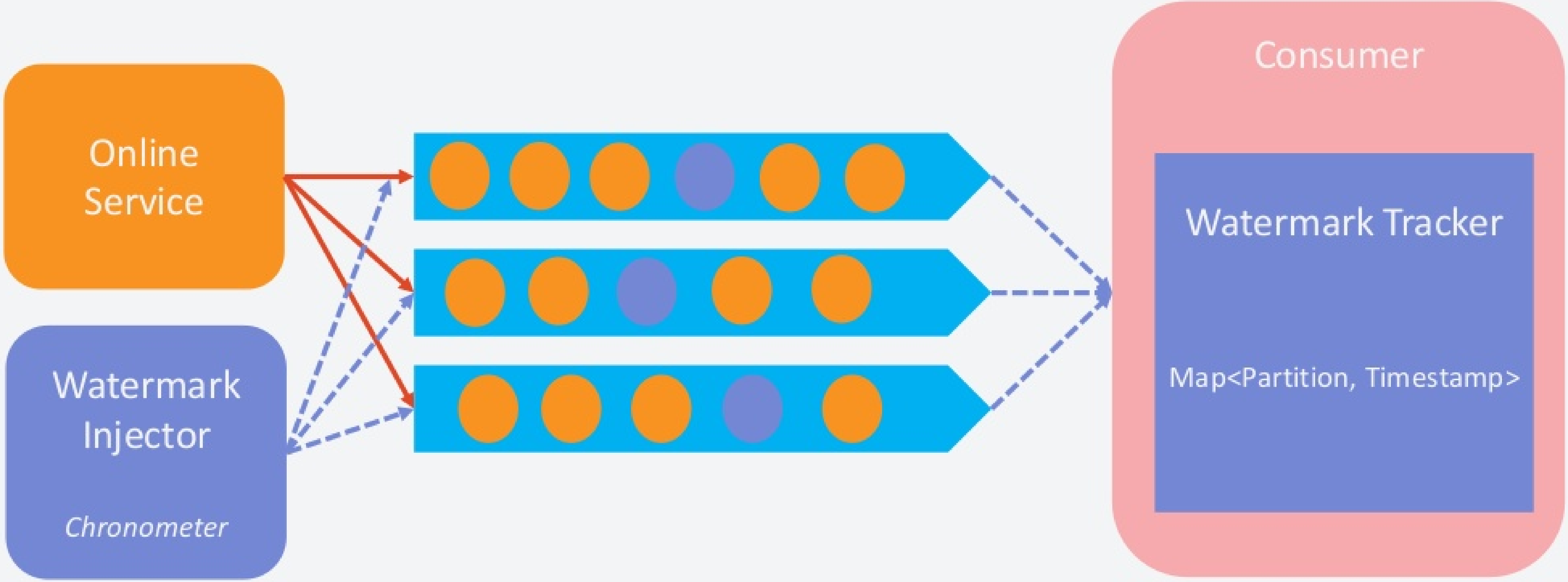


Time series management

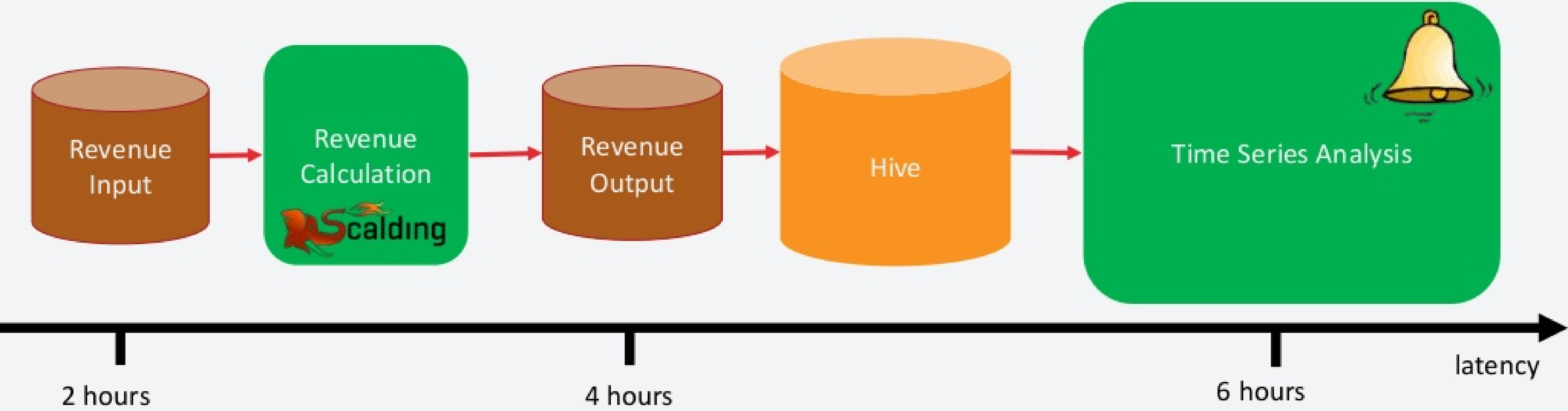


- Single Kafka partition is ordered
- Time of partition **max(watermark)**
- Global time **min(partitions)**

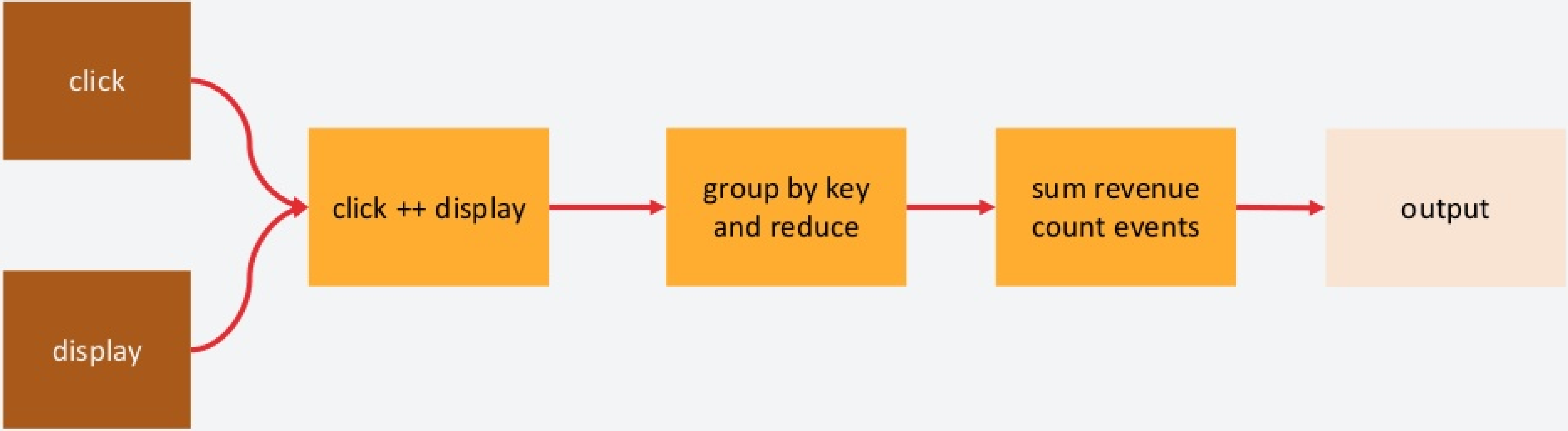
Time series management



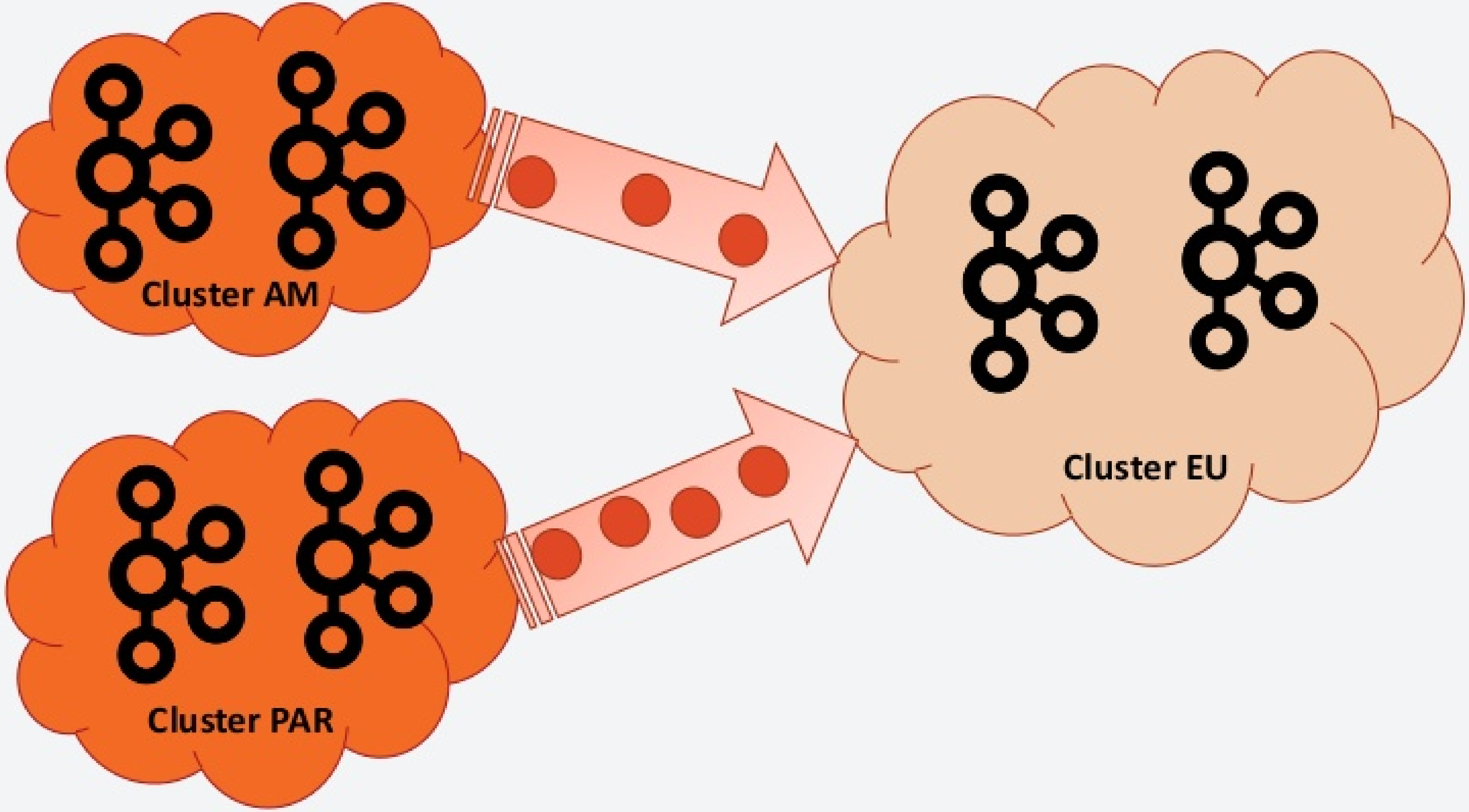
Revenue anomalies offline



Business logic



Stream processing infrastructure



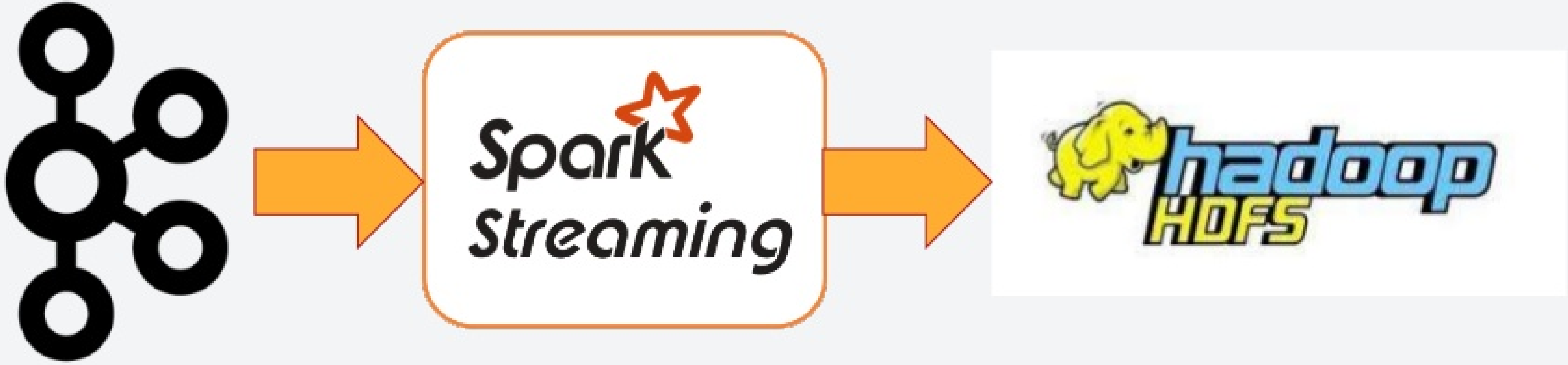
Stream processing infrastructure

- A rogue Kafka client can significantly reduce Kafka cluster performance

- We need to shuffle the data uniformly

- A streaming pipeline needs to have platform level (EU, AS, …) data

Stream processing first try

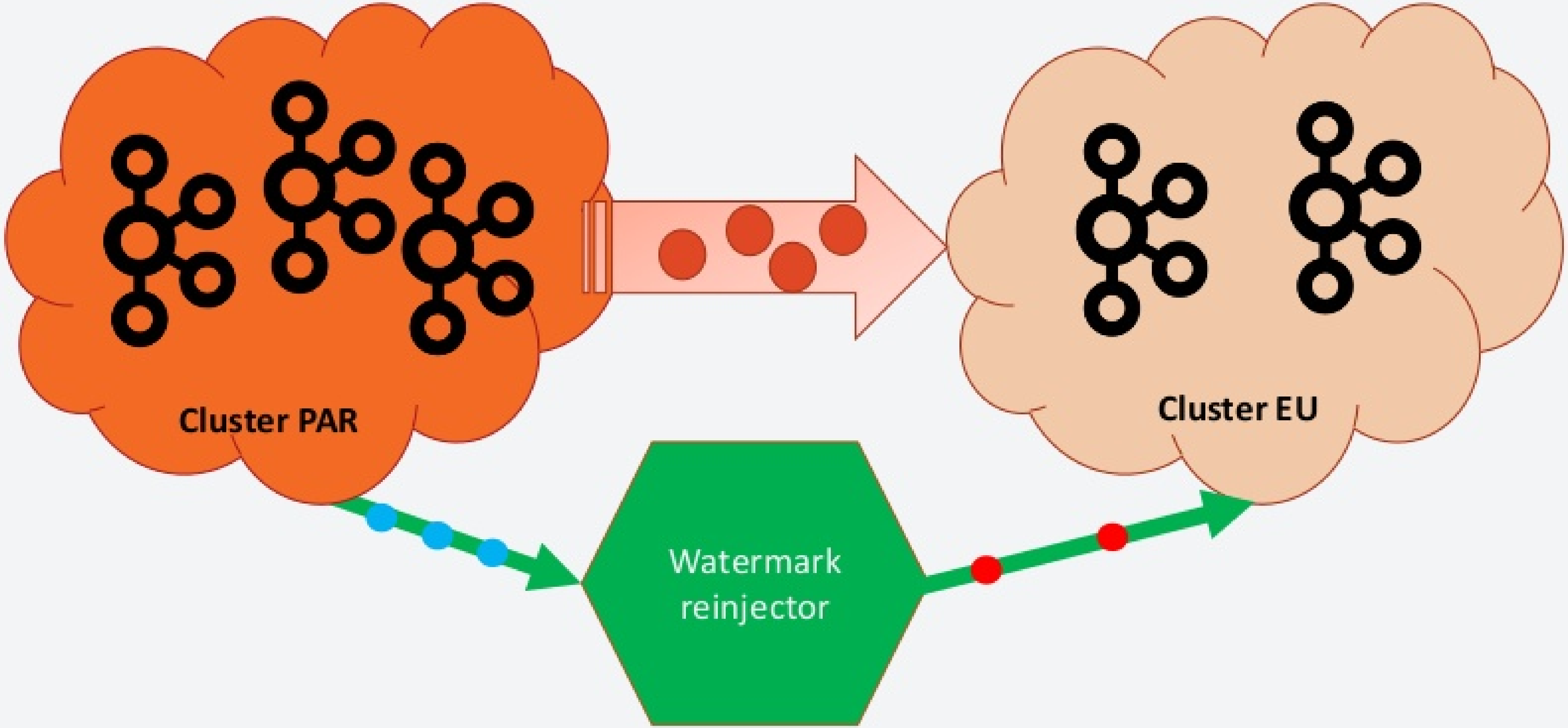


Streaming done right

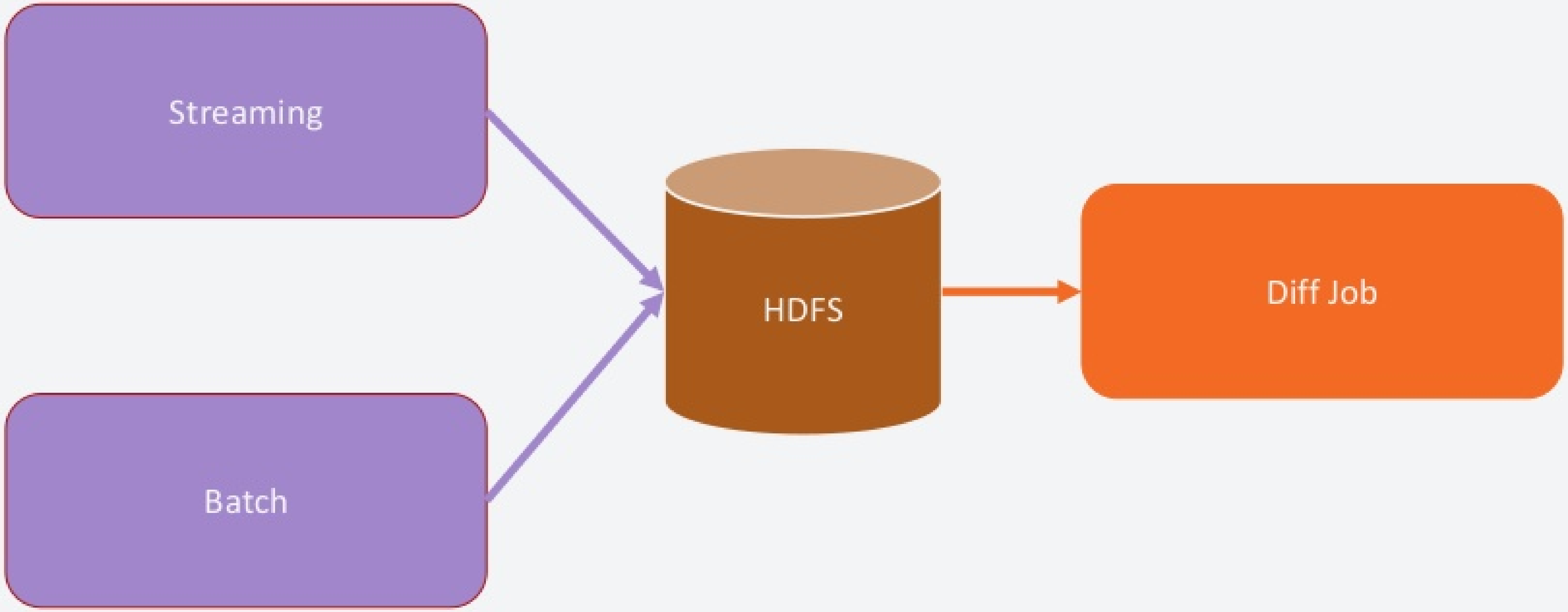
Streaming done right

- Flexible event time support
- Robust state management
- Short development cycle

Accurate event time processing



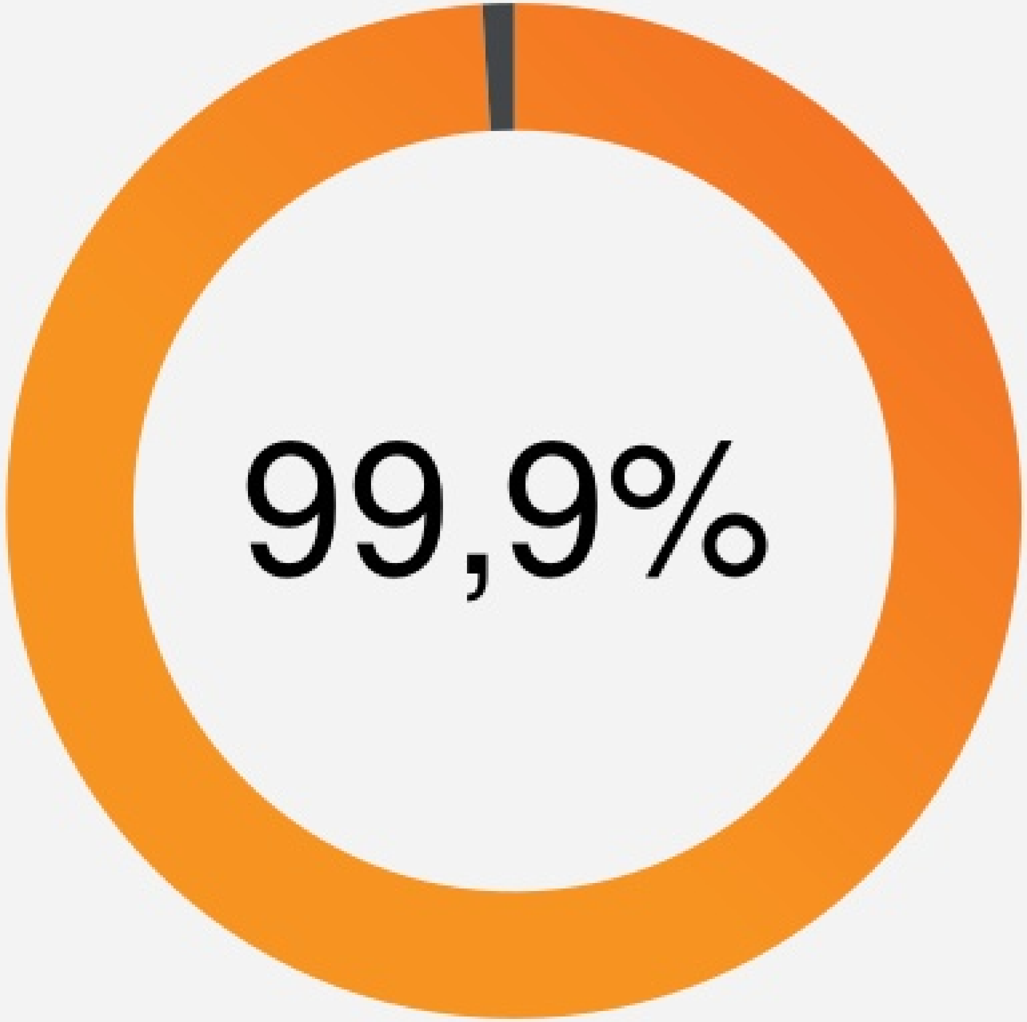
Result verification



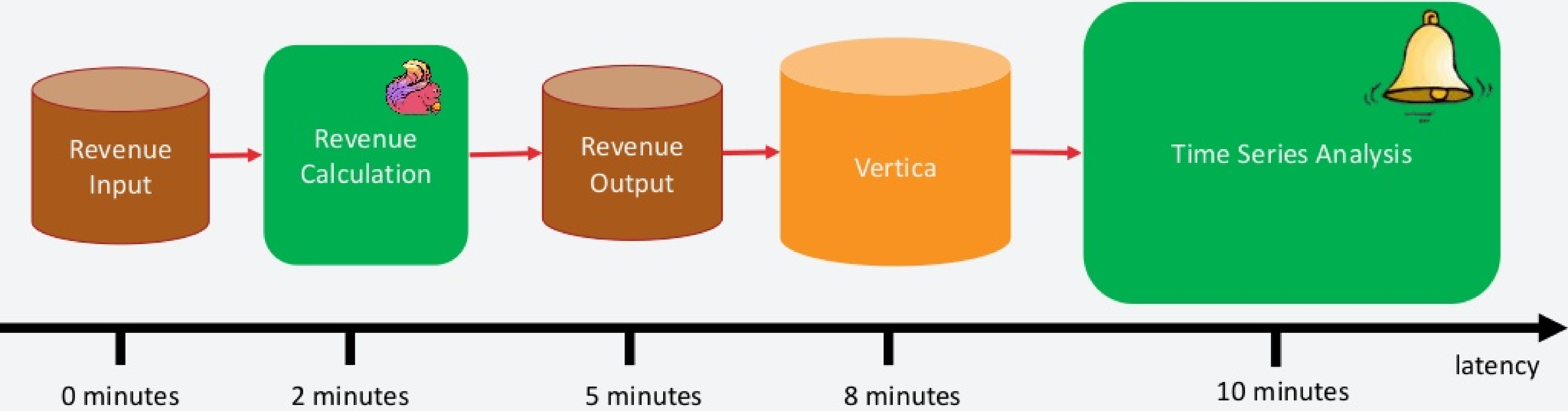
Correctness results

Happy path looked great

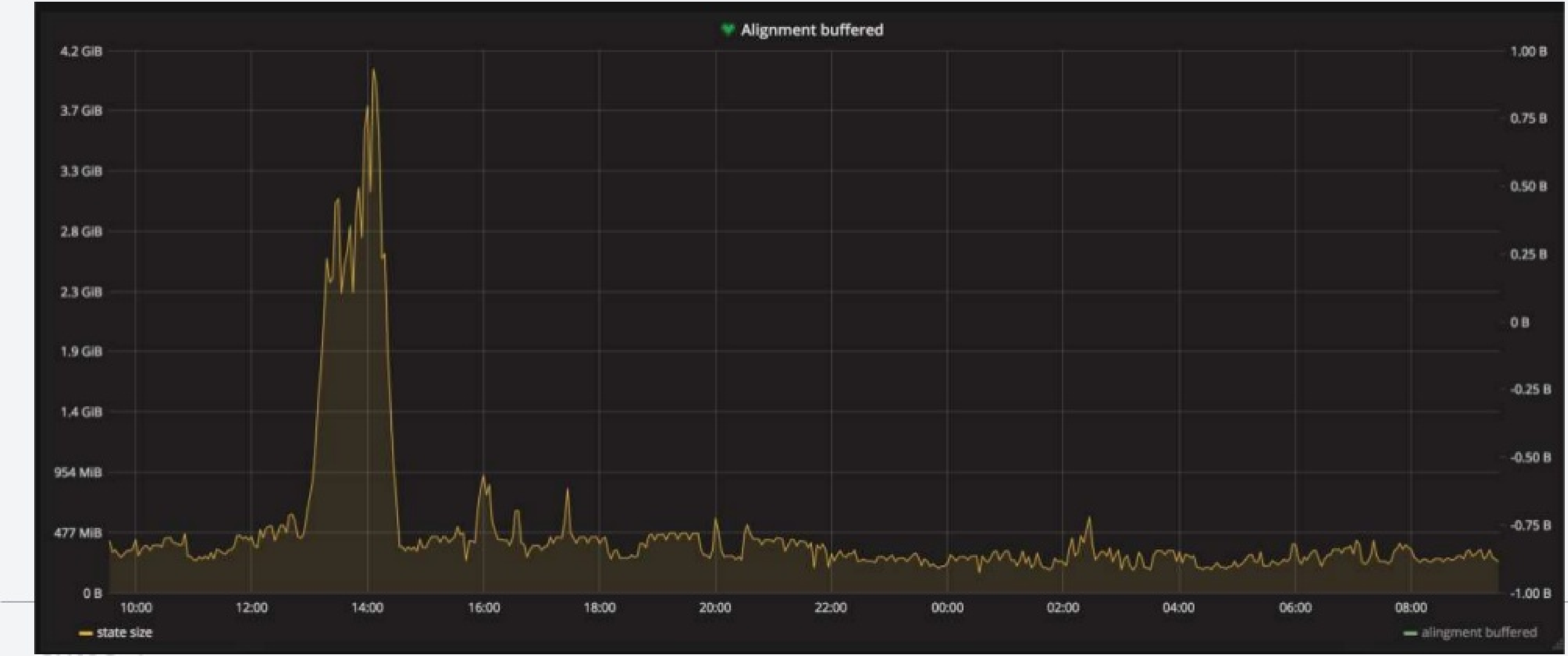
- no watermark issues
- small duplication rate
- no lost state of the Flink
- no big latency
- no spike of late events



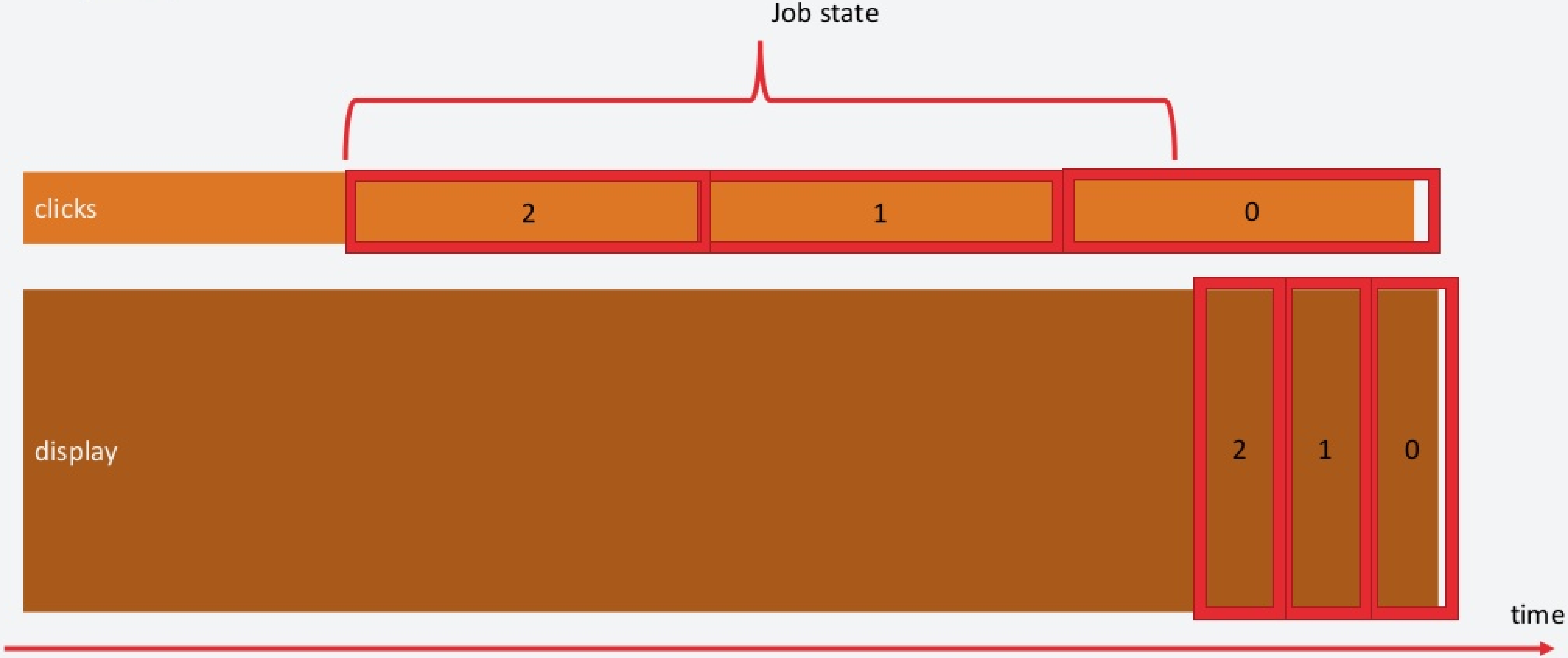
Revenue anomalies online



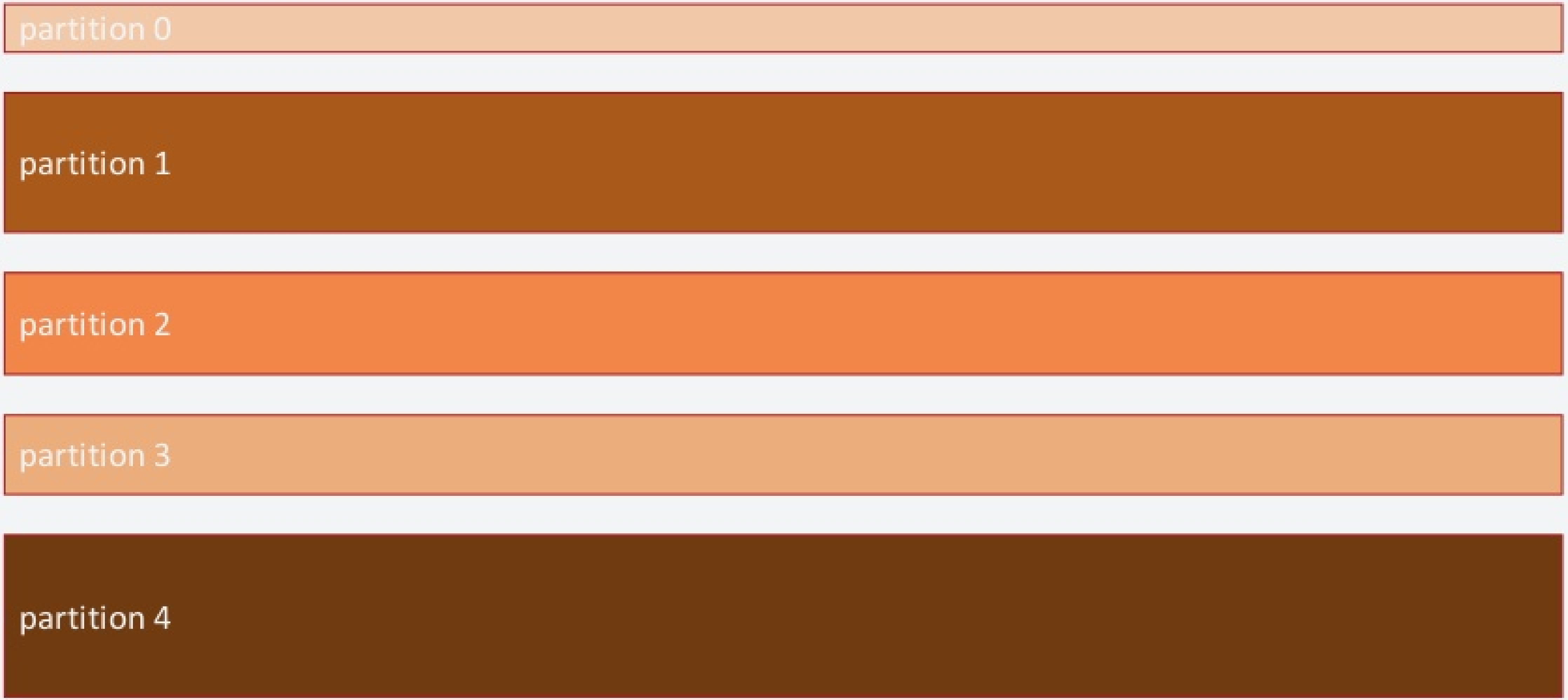
Job state


Alignment buffered
4.2 GiB
3.7 GiB
3.3 GiB
2.8 GiB
2.3 GiB
1.9 GiB
1.4 GiB
954 MiB
477 MiB
0 B
1.00 B
0.75 B
0.50 B
0.25 B
0 B
-0.25 B
-0.50 B
-0.75 B
-1.00 B
10:00
12:00
14:00
16:00
18:00
20:00
22:00
00:00
02:00
04:00
06:00
08:00
state size
alingment buffered

Problems with a Kappa I



More generic problems with a Kappa



partition 0

partition 1

partition 2

partition 3

partition 4

Streaming enables use-cases

- Our TSE can monitor any ads campaign changes in ”realtime”
- We have enabled a short term campaigns which takes only several hours to run
- We efficiently cut campaigns spending when campaign is over

Notes to take

- Simple business use-case can allow to focus on the infrastructure
- Store watermarks within your data for catching up
- Flink is an enabler

Thank you

we are hiring

criteo