Understanding Advanced Stream Processing

Eugene Meidinger
BI CONSULTANT

@sqlgene www.sqlgene.com

Overview

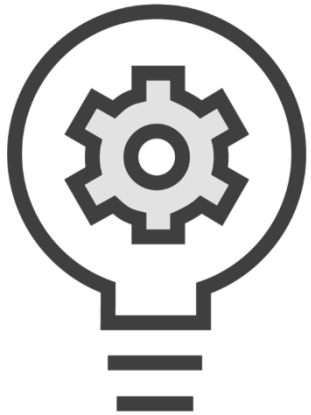
State

Time

Consistency

What Is State?

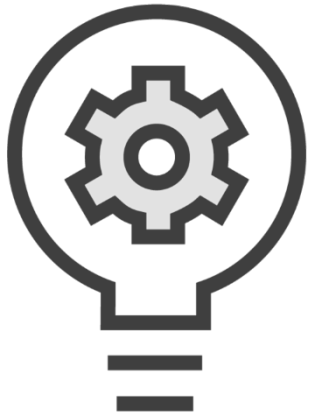
On



Off

What Is State?

1



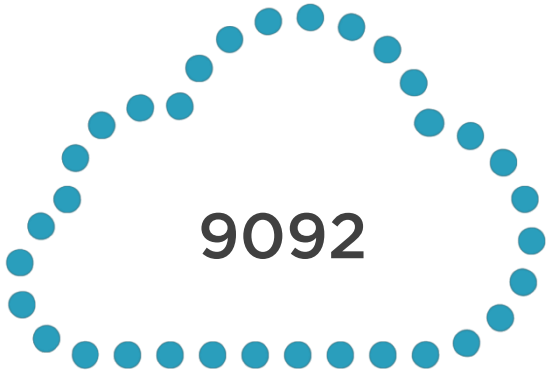
0

What Is State?

What Is State?

Stateless applications do work as it comes in and then forget it.

Restoring State



Restoring State

Recap

State

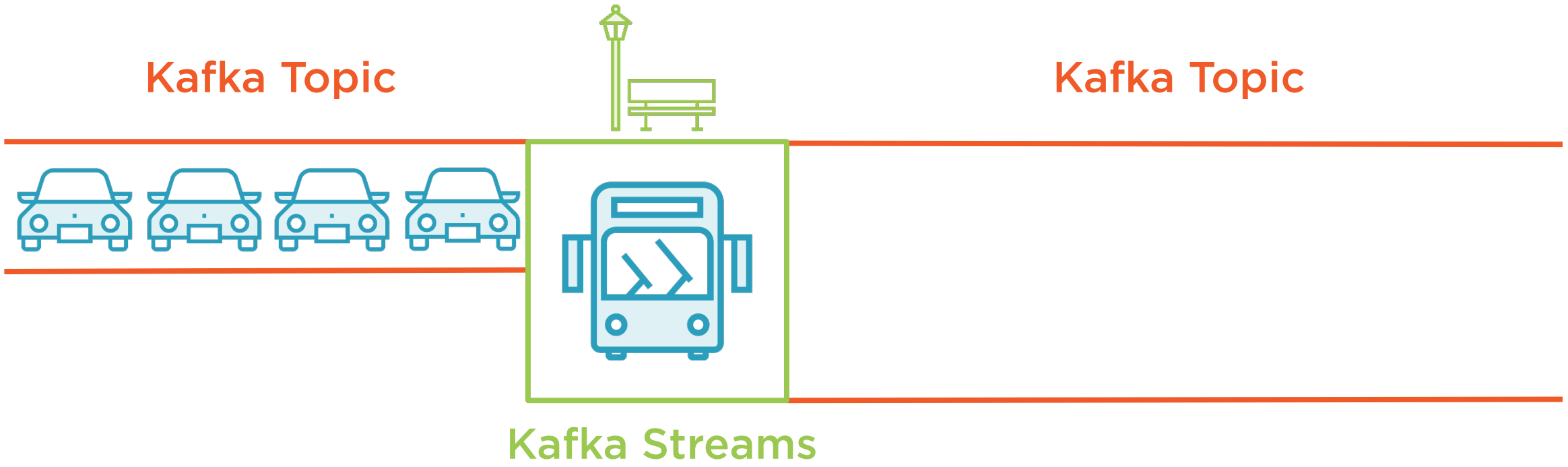
Recovery

State store

Stateful versus Stateless

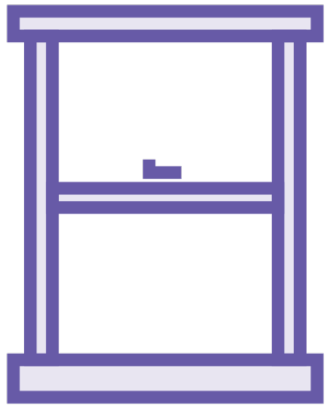
Stateless Applications

Stateful Applications

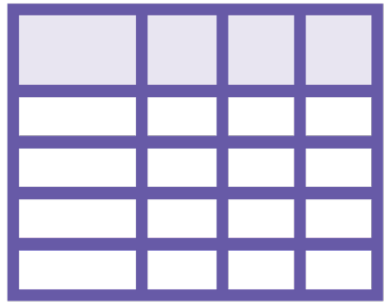


What Requires State?

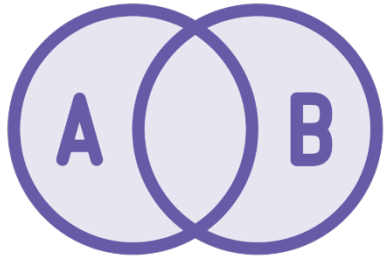
Aggregations



Windowing



Tables



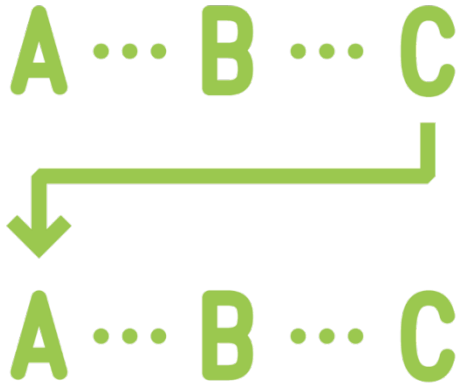
Joins

Managing Time

Streaming technology appreciates the value of time

Types of Timestamps

Event Time



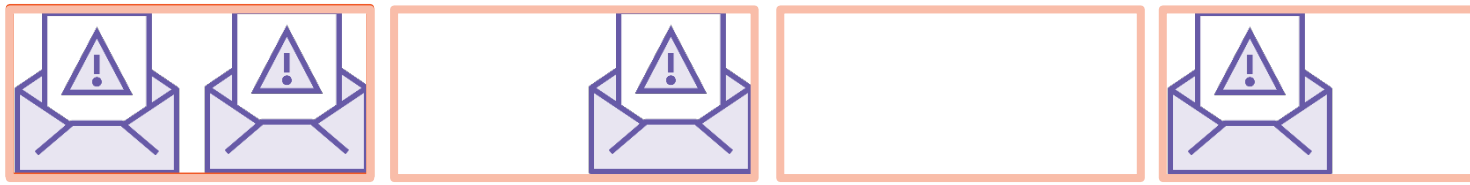
Ingestion Time

Processing Time

Windowing

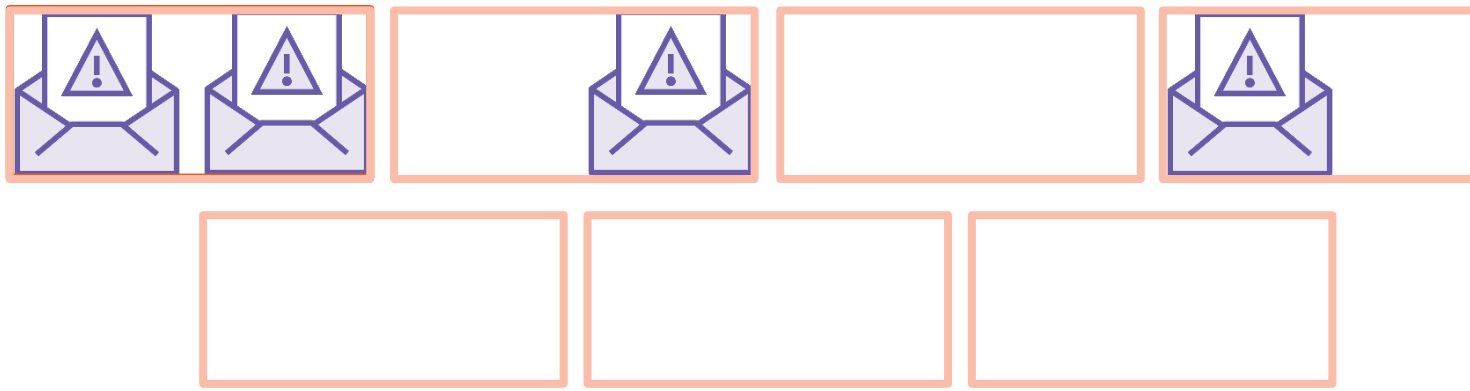
The process of is grouping records based on event time.

Tumbling



Tumbling

Hopping



Tumbling

Hopping

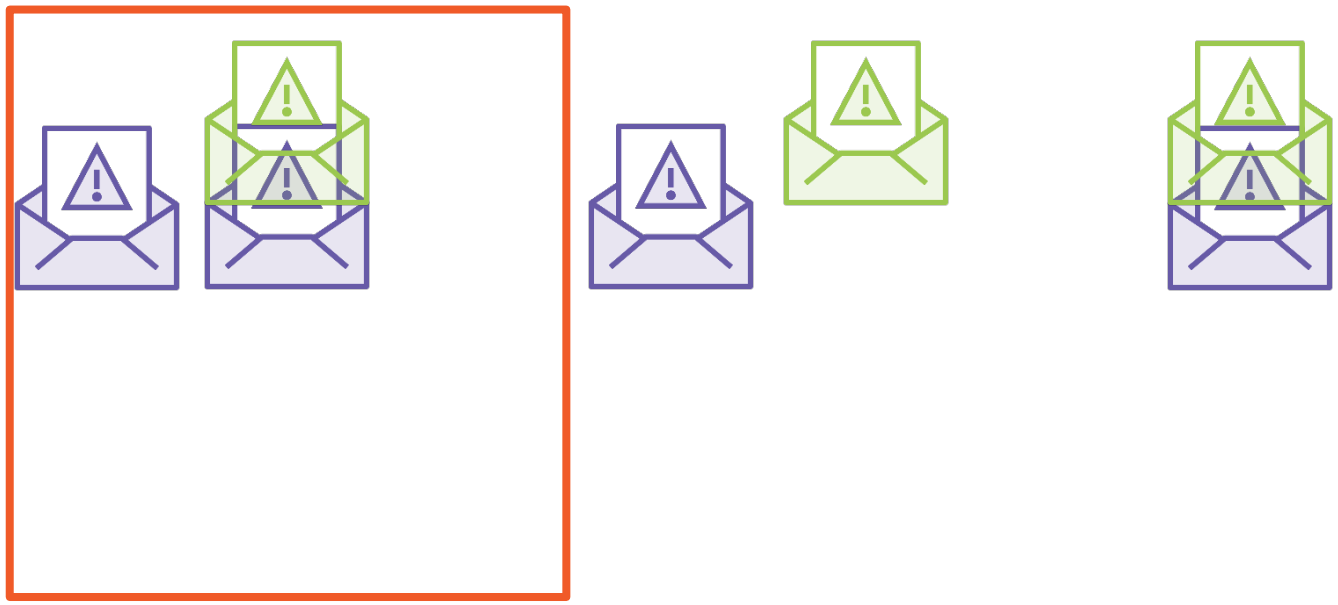
Session

Tumbling

Hopping

Session

Sliding



Grace Period

The period of time a system will wait for late and out of order events.

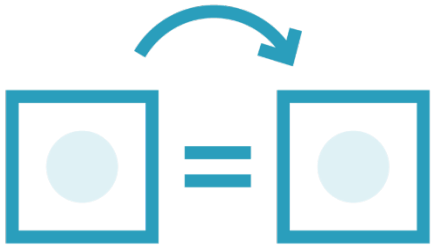
Retention Period

The period of time a system will keep, or retain, data.

Managing Consistency

Consistency Failures

Read Never



Read Once

Read Multiple

Avoiding Read Never

State Store

Recovery Process

Idempotent

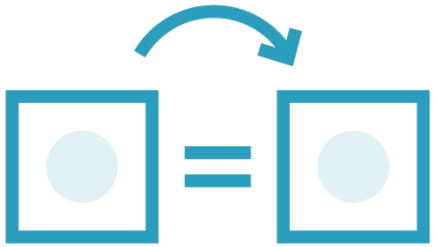
A mathematical operation that produces the same result if repeated.

Idempotent

An action that is safe to repeat.

Consistency Failures

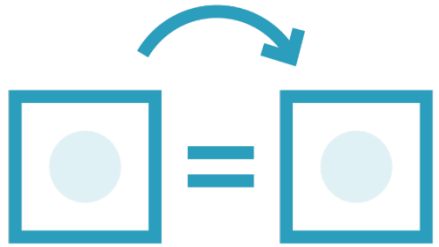
Read Never



Read Once

Read Multiple

Processing Guarantees



Exactly-once

At-least-once

Exactly-once semantics
aren't free.

Summary

Some apps need state

Window over event time

Consistency isn’t free