Try FREE

confluent.io/cloud

$200 Free each month

3 Months from signup

CONFLUENT

Stream Processing Fundamentals

How Stream Processing Works: Basic Concepts

Mark Fei — Senior Instructor
Confluent

Session Schedule

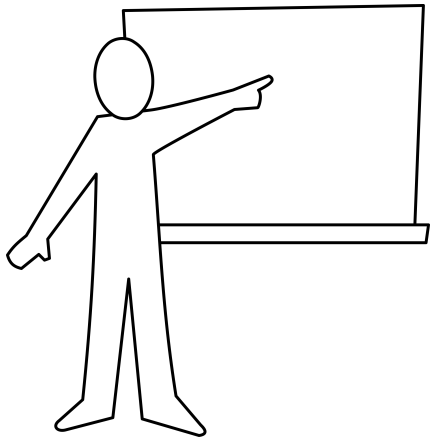
- **Session 1 - How Stream Processing Works: Basic Concepts**
- Session 2 - Stream Processing with Kafka Streams
- Session 3 - Introduction to ksqlDB

Learning Objectives

After this module you will be able to:

- List at least 3 advantages of **stream processing** over **batch processing**
- Compare and contrast **Kafka Streams** and **ksqlDB**
- Describe the main **characteristics** of a **stream**
- Explain dualism between **stream** and **table**
- Elaborate on a **stream processing job**

Module Map



- Motivation and Use Cases for Real-Time Streaming …
- High Level Comparison of Kafka Streams and ksqlDB
- Introduce Stream Processing Concepts

Motivation (1)

The Shift to Event-driven Systems has Already Begun…

From a static snapshot...

Occasional call to a friend

...to a continuous stream of events

A constant feed about the activities of all your friends

Daily news reports

Real time news feeds, accessible online anytime, anywhere

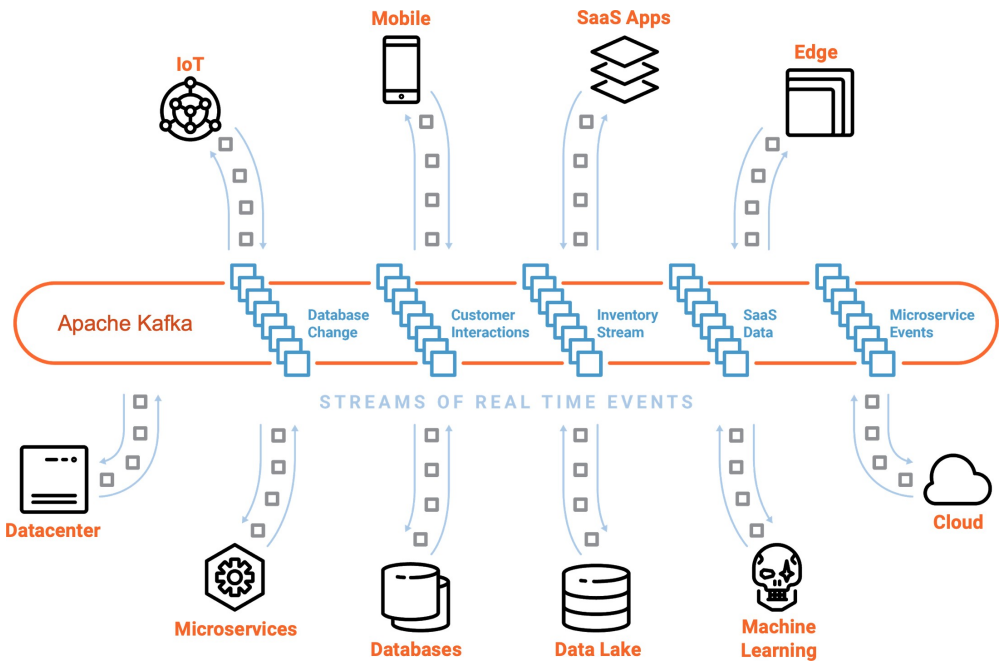
Motivation (2)

Single platform to connect every system to every event

Real-time stream of events

All events stored for historical view

Motivation (3)

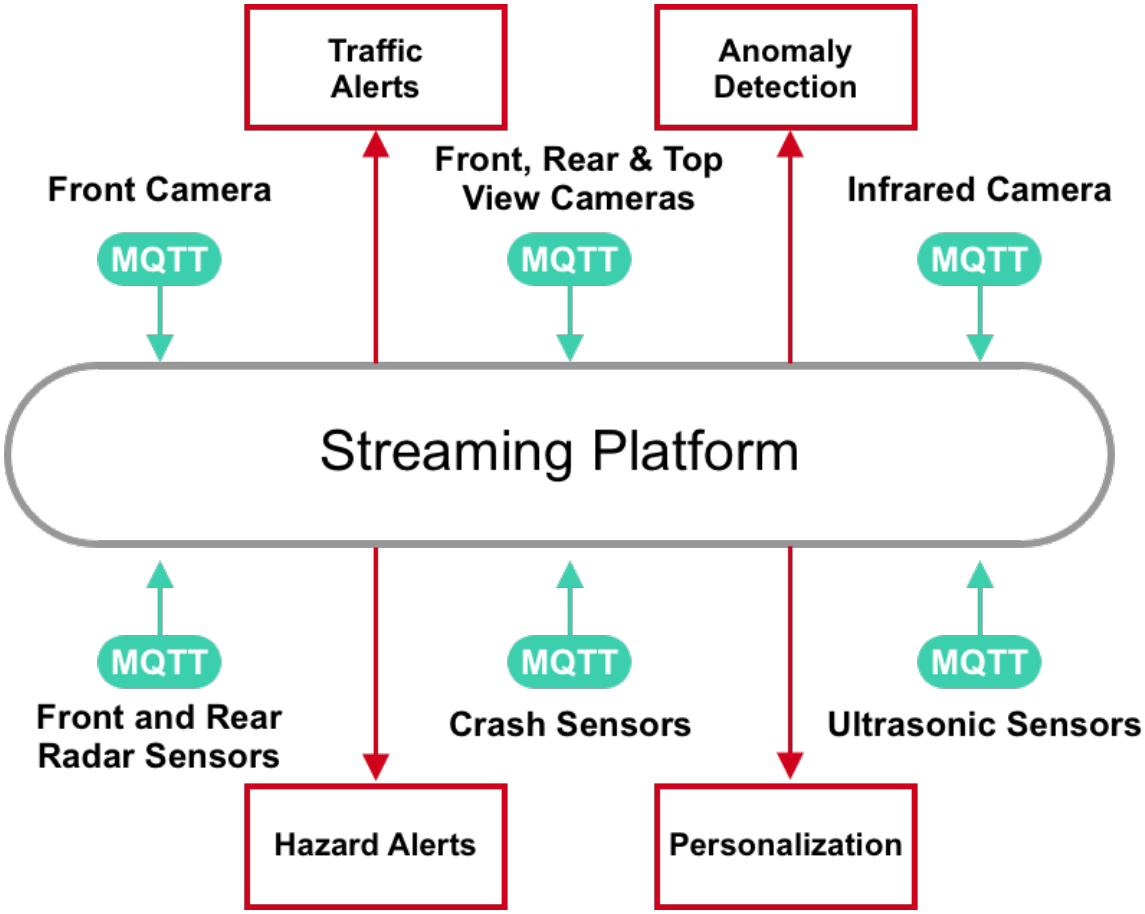


Apache Kafka®: the Defacto Standard for Real-Time Event Streaming

- Global-scale
- Real-time
- Persistent Storage
- Stream Processing

Use Case: Automotive Internet of Things

The Future of the Automotive Industry is a Real Time Data Cluster

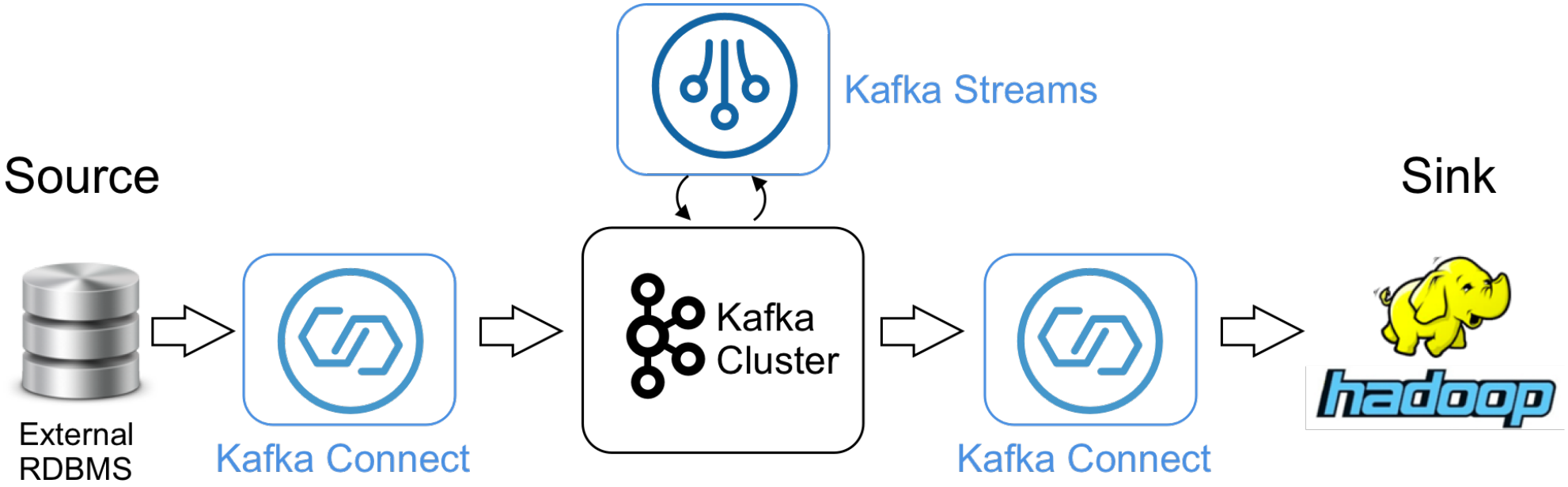


Use Case: Online Gaming

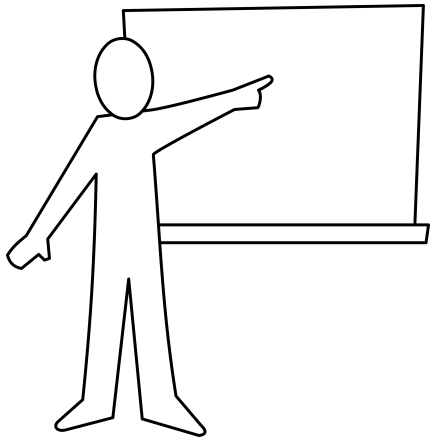
Real-Time Gaming

- In-game location data
- Analyze player behavior
- Analyze server performance
- Process data at scale

Real-time Data Pipeline



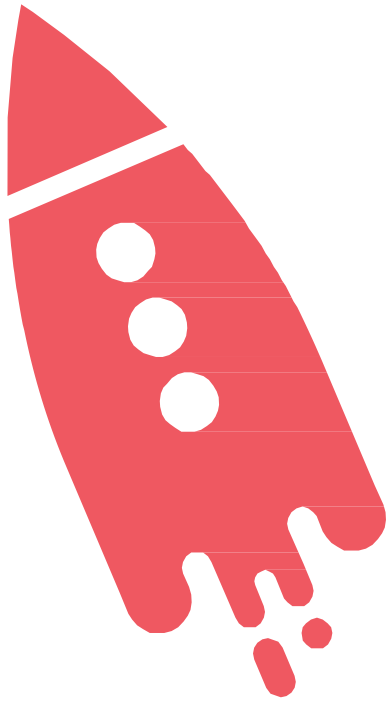
Module Map



- Motivation and Use Cases for Real-Time Streaming
- High Level Comparison of Kafka Streams and ksqlDB ...
- Introduce Stream Processing Concepts

Real-Time Processing with Kafka Streams and ksqlDB

Kafka Streams



Confluent ksqlDB

ksqlDB vs. KStreams

Start with ksqlDB when...

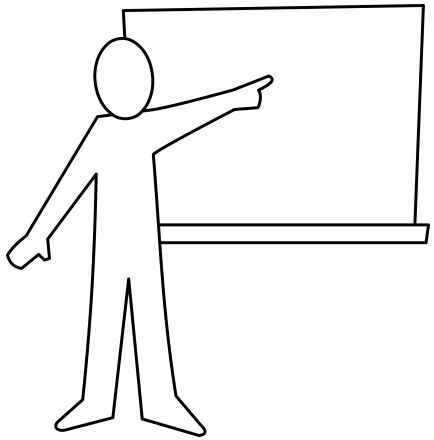
- You don't use Java/Scala
- You are new to streaming or Kafka
- You prefer a UI or REST API
- You can achieve stream processing and state querying use case with KSQL syntax

Start with Kafka Streams when ...

- You already use Java/Scala
- You need tight control over performance
- You prefer the deployment flexibility of using a Java Library
- You need custom logic that can't be described with KSQL syntax

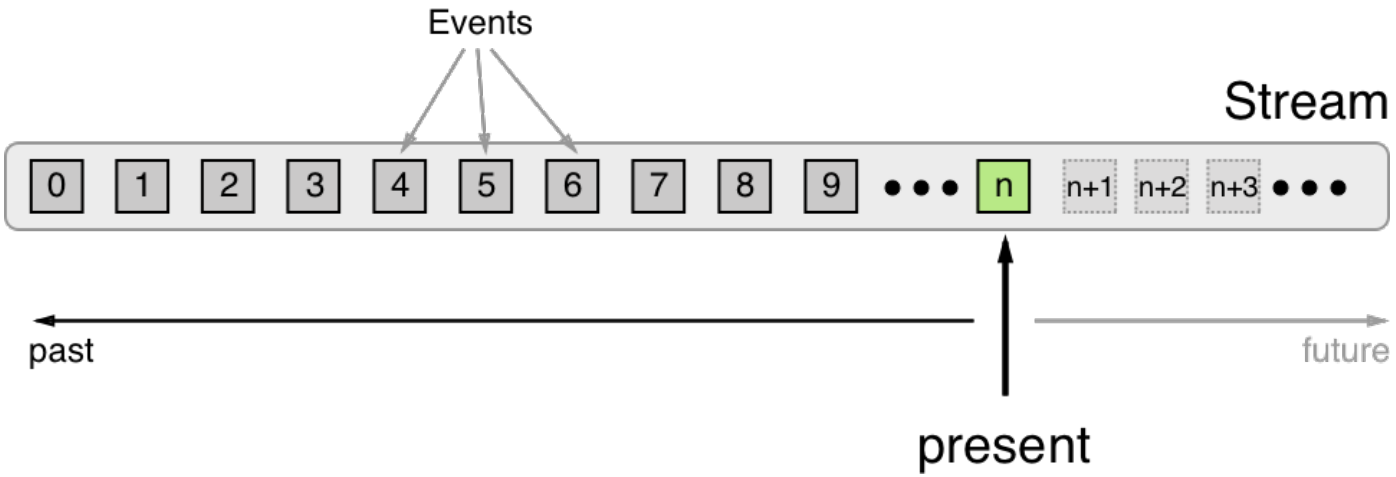
You can start with ksqlDB and expand functionality later with User Defined Functions (UDFs) in Java or Scala!

Module Map

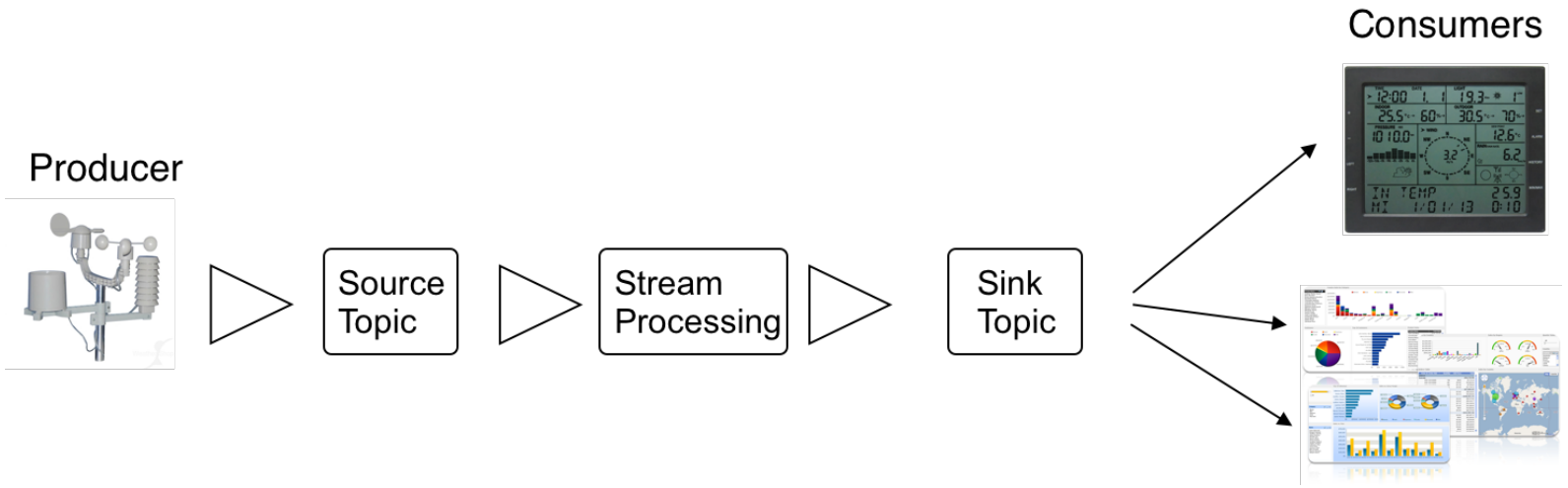


- Motivation and Use Cases for Real-Time Streaming
- High Level Comparison of Kafka Streams and ksqlDB
- Introduce Stream Processing Concepts ...

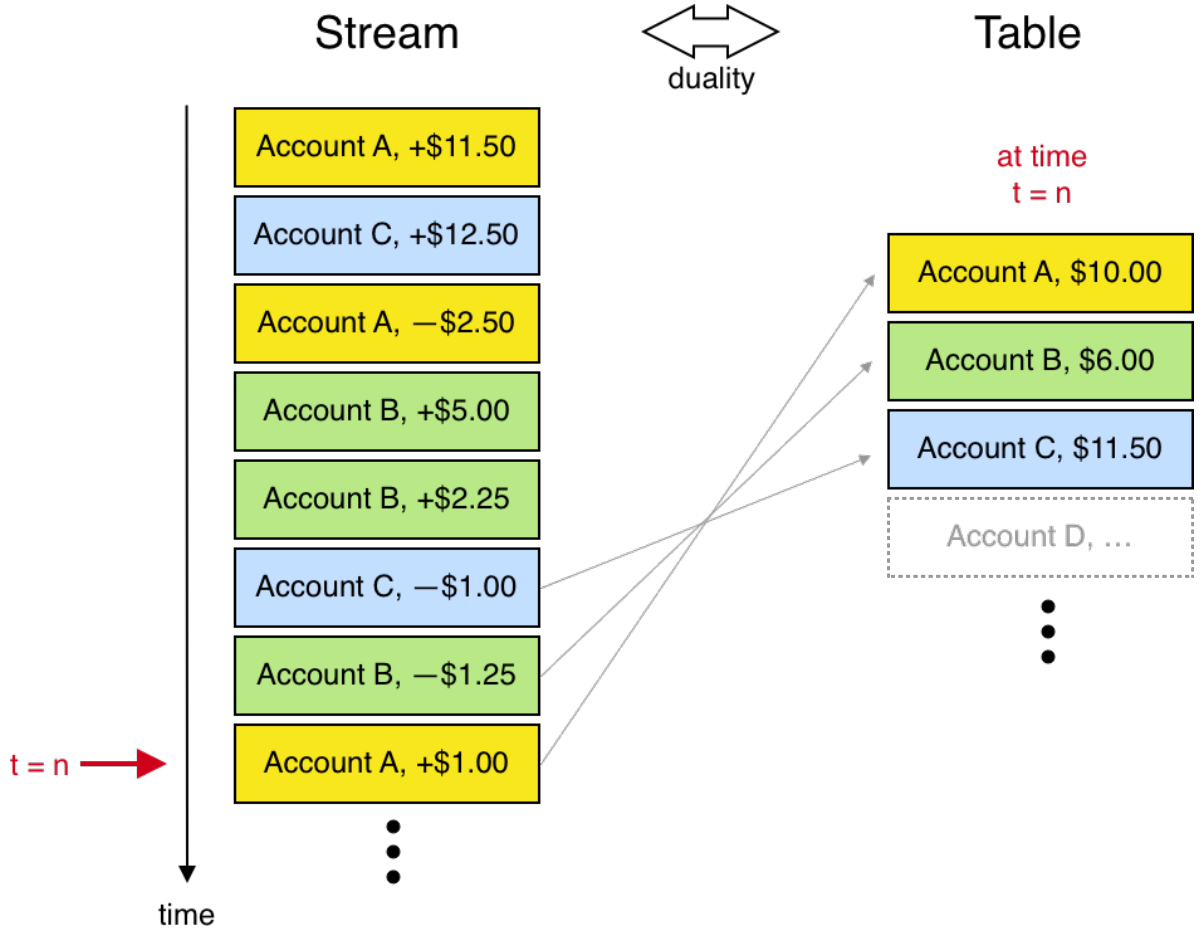
The Stream



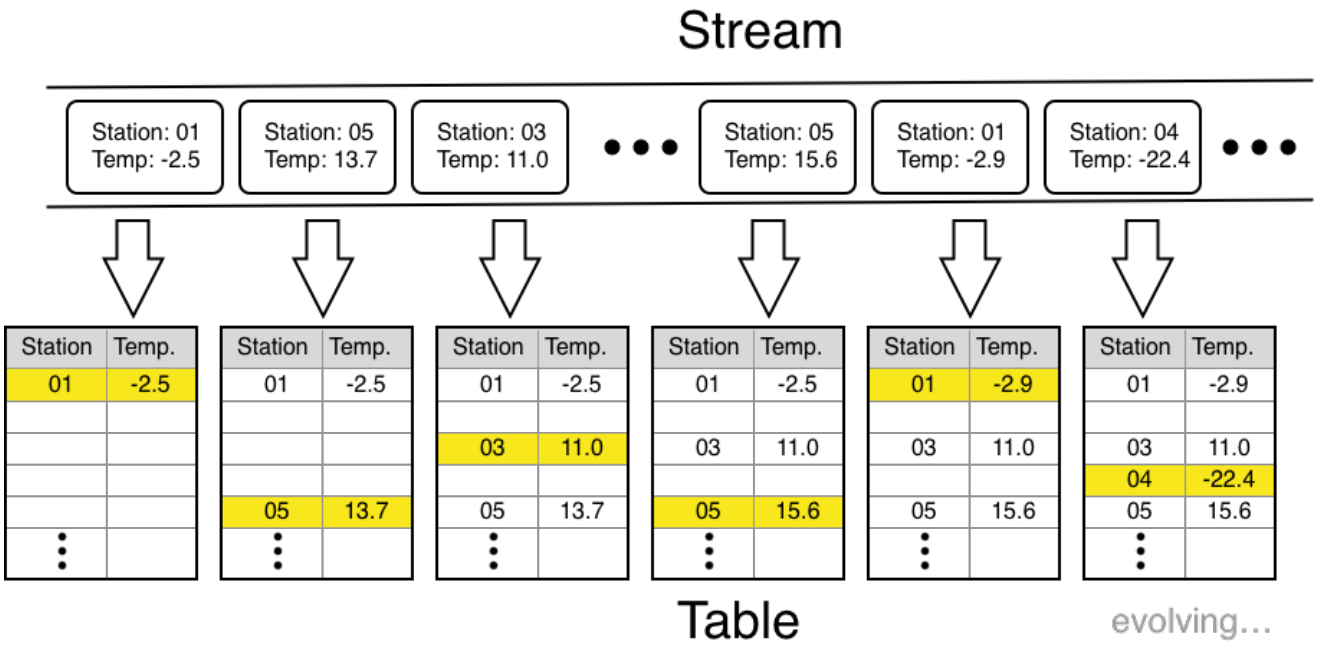
Streaming Pipeline



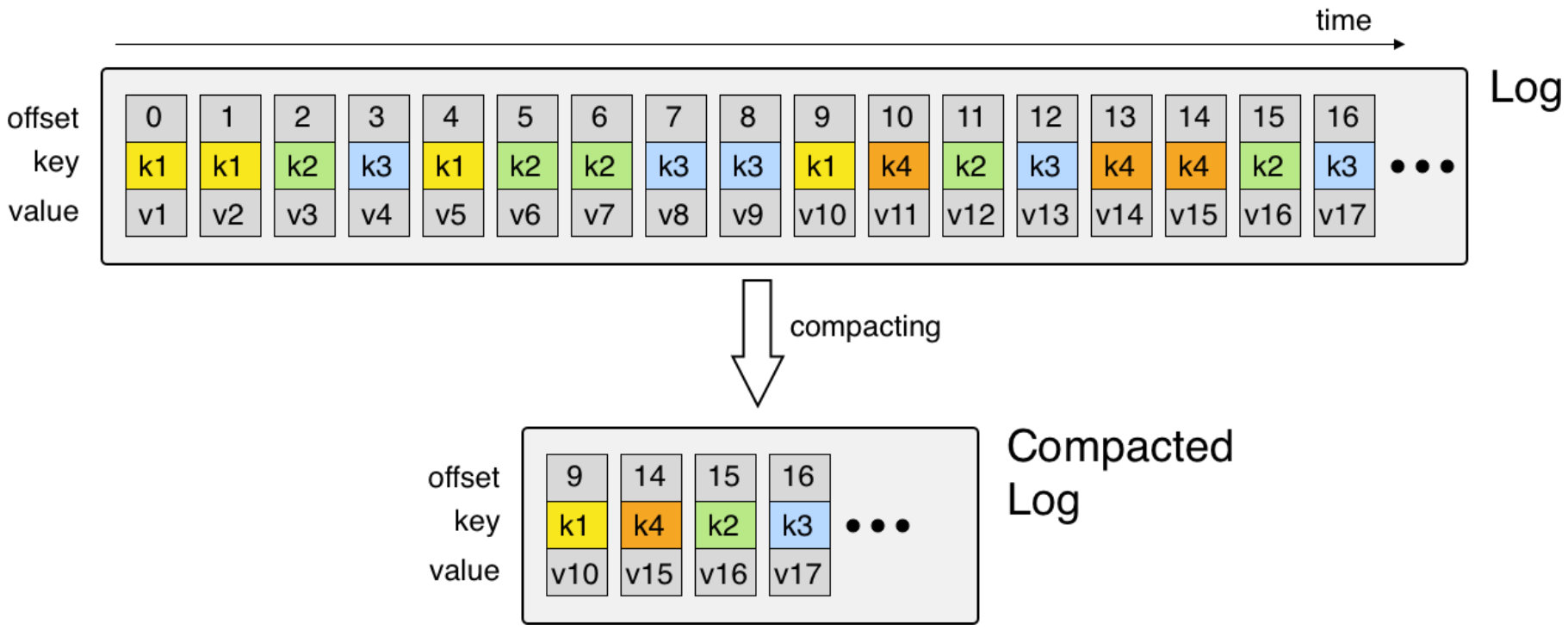
Stream - Table Duality



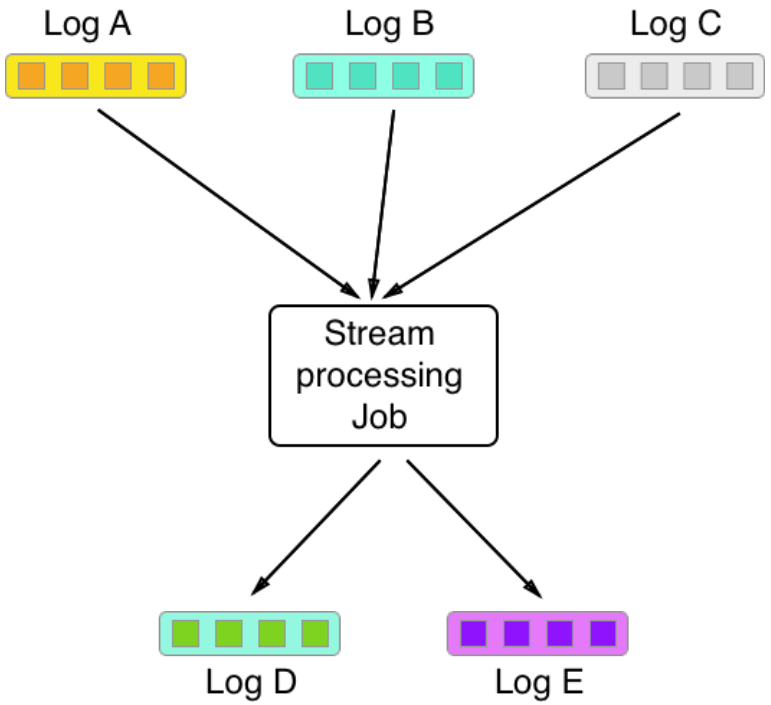
The Table



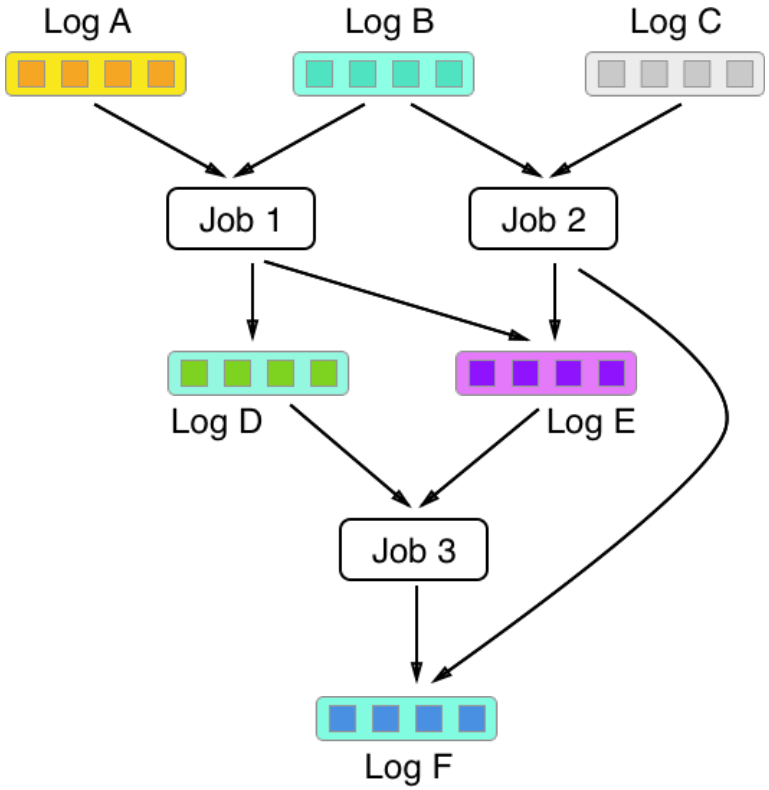
Log Compaction



Stream Processing Job



Multi Job DataFlow



Review

Question: Using a **stream** as the backing structure for a system like Kafka has many advantages. Yet a stream-based data store is not meant to be queryable. What can you do to satisfy the needs of business to query the data and still use a log based system such as Kafka?

Continue your Apache Kafka Education!

- Confluent Operations for Apache Kafka
- Confluent Developer Skills for Building Apache Kafka
- Confluent Stream Processing using Apache Kafka Streams and KSQL
- Confluent Advanced Skills for Optimizing Apache Kafka

For more details, see http://confluent.io/training

Certifications

Confluent Certified Developer for Apache Kafka

(aligns to Confluent Developer Skills for Building Apache Kafka course)

Confluent Certified Administrator for Apache Kafka

(aligns to Apache Kafka Administration by Confluent course)

What you Need to Know

- 6-to-9 months hands-on experience
- 90 mins, online 24/7
- Orderable/purchasable:
  - Confluent Order Form
  - Website – self-transact
- Single seat - $150 USD
- 5 Vouchers (10% off) – $675
- 10 Vouchers (20% off) - $1200

Stay in touch!

Online Talks

cnfl.io/online-talks