SpringOne

Getting Started with Spring Cloud Data Flow

Sabby Anandan (@sabbyanandan)
Product Manager

Covering today ..

Spring Cloud Data Flow

- Overview
- Features / Use-cases

Hands-on Labs

- SCDF on Kubernetes
- Event Streaming
- Cloud-native Batch

Housekeeping ..

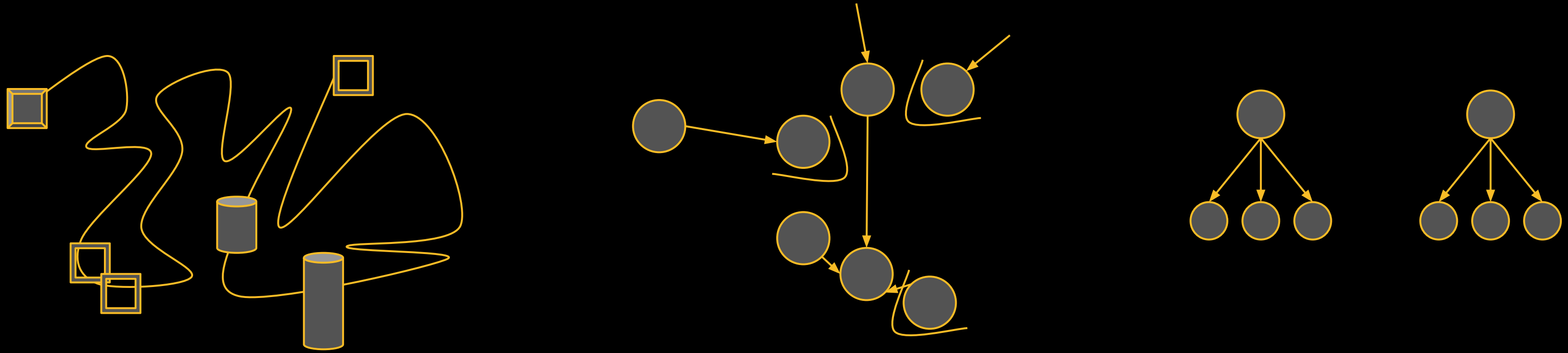
- Mix of presentation + walkthrough (30 mins)

- 3 hands-on labs (~15 mins each)

- Questions? raise hand; ask questions in chat

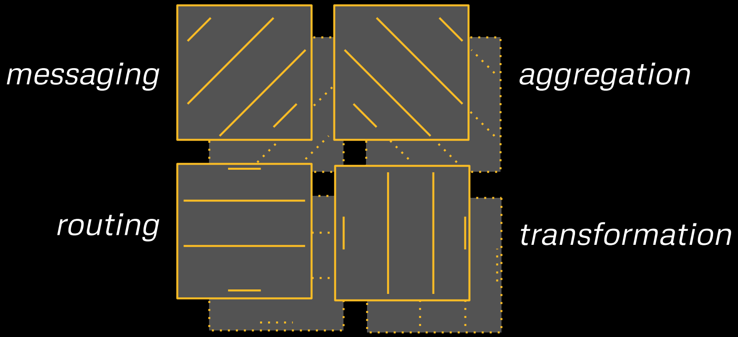
Enterprise Integration

“communication between people, machines, computers, and devices and their efficient coordination”



Enterprise Integration Patterns

“communication between people, machines, computers, and devices and their efficient coordination”

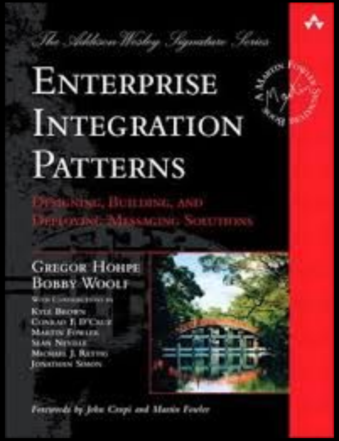


messaging

aggregation

routing

transformation



The Addison-Wesley Signature Series

ENTERPRISE INTEGRATION PATTERNS

DESIGNING, BUILDING, AND DEPLOYING MESSAGING SOLUTIONS

GREGOR HOHPE
BOBBY WOOLF

WITH CONTRIBUTIONS BY
KYLE BROWN
CONRAD F. D'CRUZ
MARTIN FOWLER
SEAN NEVILLE
MICHAEL J. RETTIG
JONATHAN SIMON

Forewords by John Crupi and Martin Fowler

Enterprise Integration in Reality

"communication between people, machines, computers, and devices and their efficient coordination"

Unpredictable Traffic Patterns

Slow Dependencies

Schema Changes

Ordering Guarantees

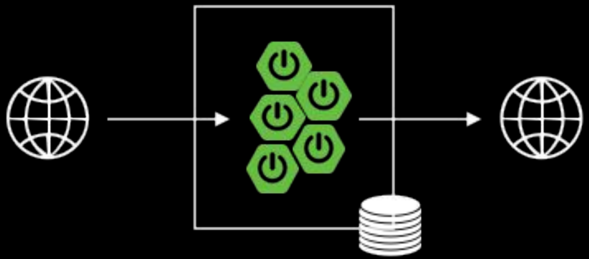
Data Loss

Spring to Rescue!

Cloud-native Enterprise Integration

Spring Cloud Task

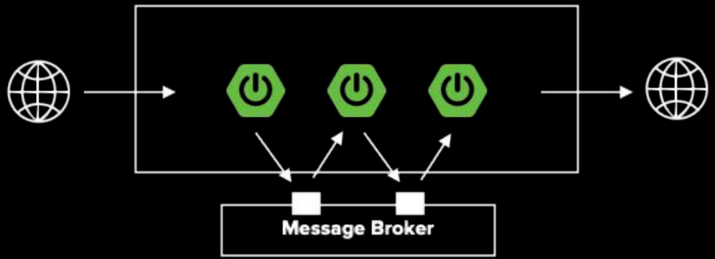
Short-lived Spring Boot microservices for batch data processing



<u>Use-cases</u>:
Scheduled data migration jobs
Extract, Transform, and Load (ETL)
Offline machine learning and model training

Spring Cloud Stream

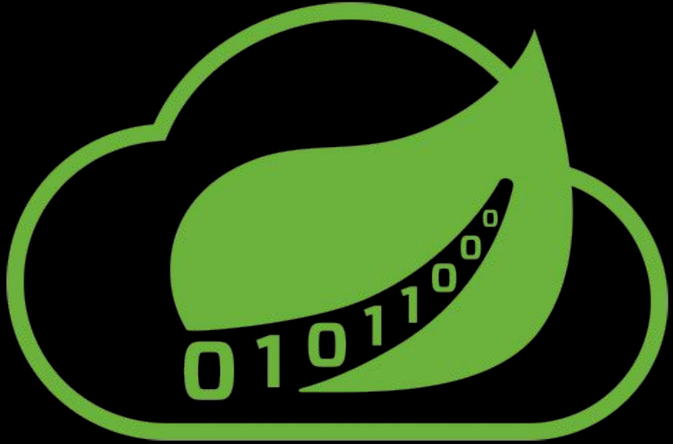
Event-driven Spring Boot microservices for real-time data processing



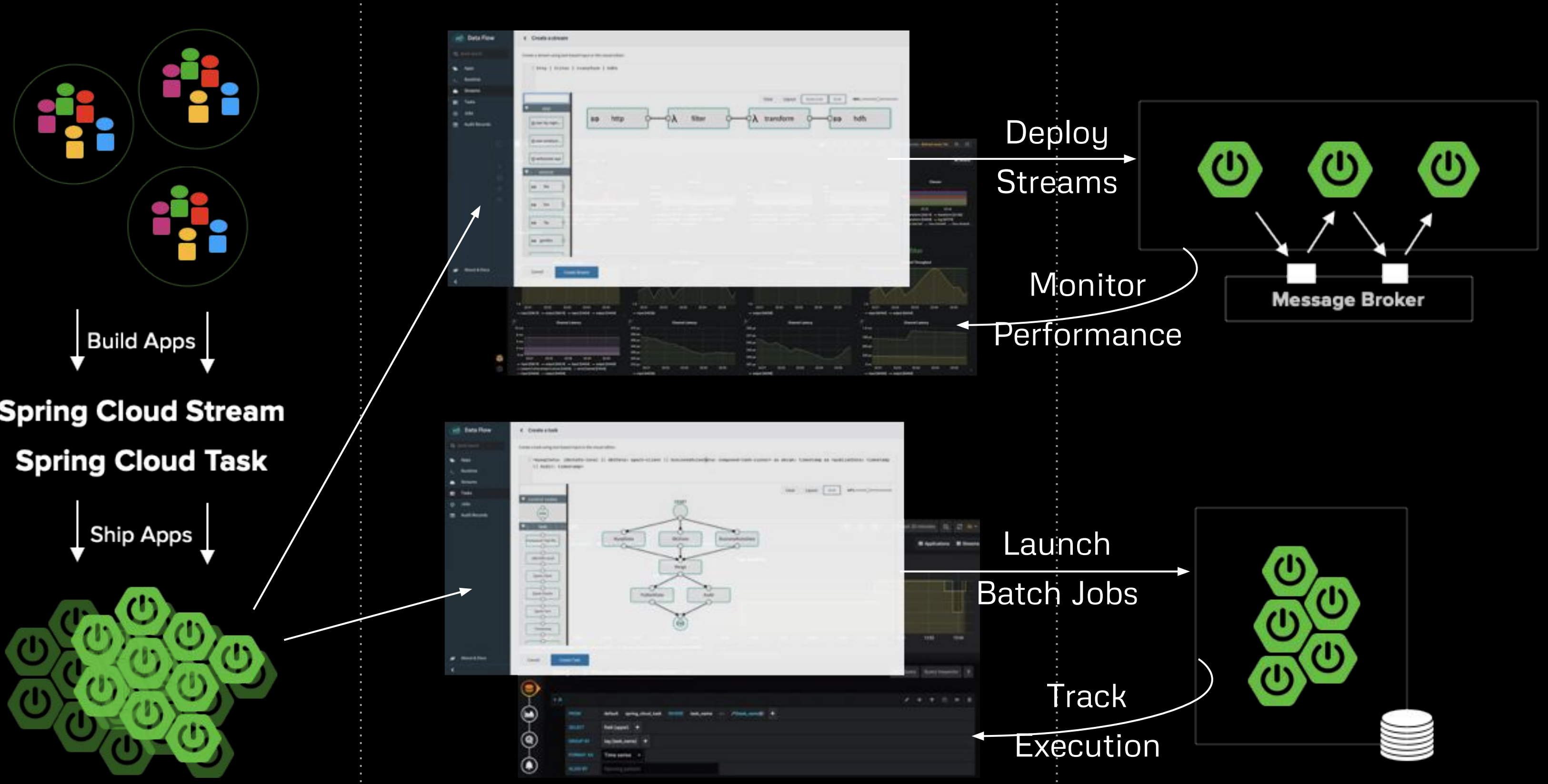
<u>Use-cases</u>:
Enterprise data integration (EAI/EIP)
Event-driven architectures
Real-time predictive analytics (e.g., IoT)

Spring Cloud Data Flow

A microservices based
Streaming and ***Batch***
data processing
in the cloud



Cloud-native Data Pipelines



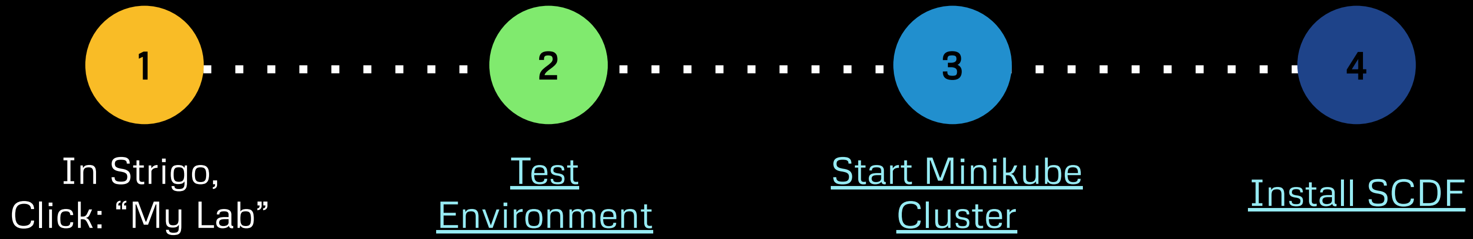
Spring Cloud Data Flow

Kubernetes

SpringOne

15
Mins

Lab #1: Set up SCDF on Kubernetes



1 In Strigo, Click: “My Lab”

2 Test Environment

3 Start Minikube Cluster

4 Install SCDF

Cloud-native Data Pipelines

- Extend cloud-native patterns to data processing use-cases
- CI/CD-centered microservices architecture
- Elastically scale Event Streaming and Batch, on-demand

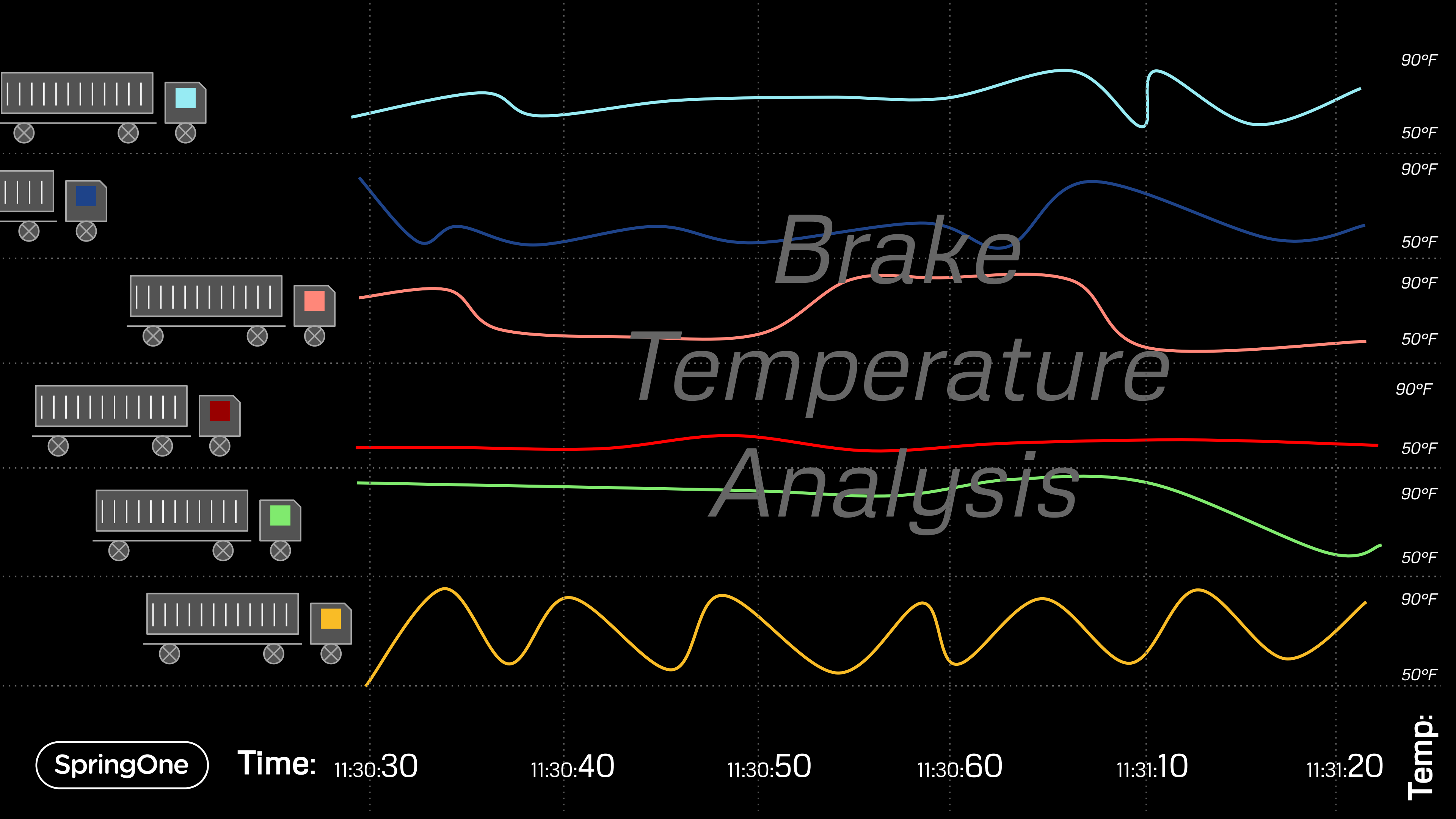
Spring Cloud Data Flow

See: Features

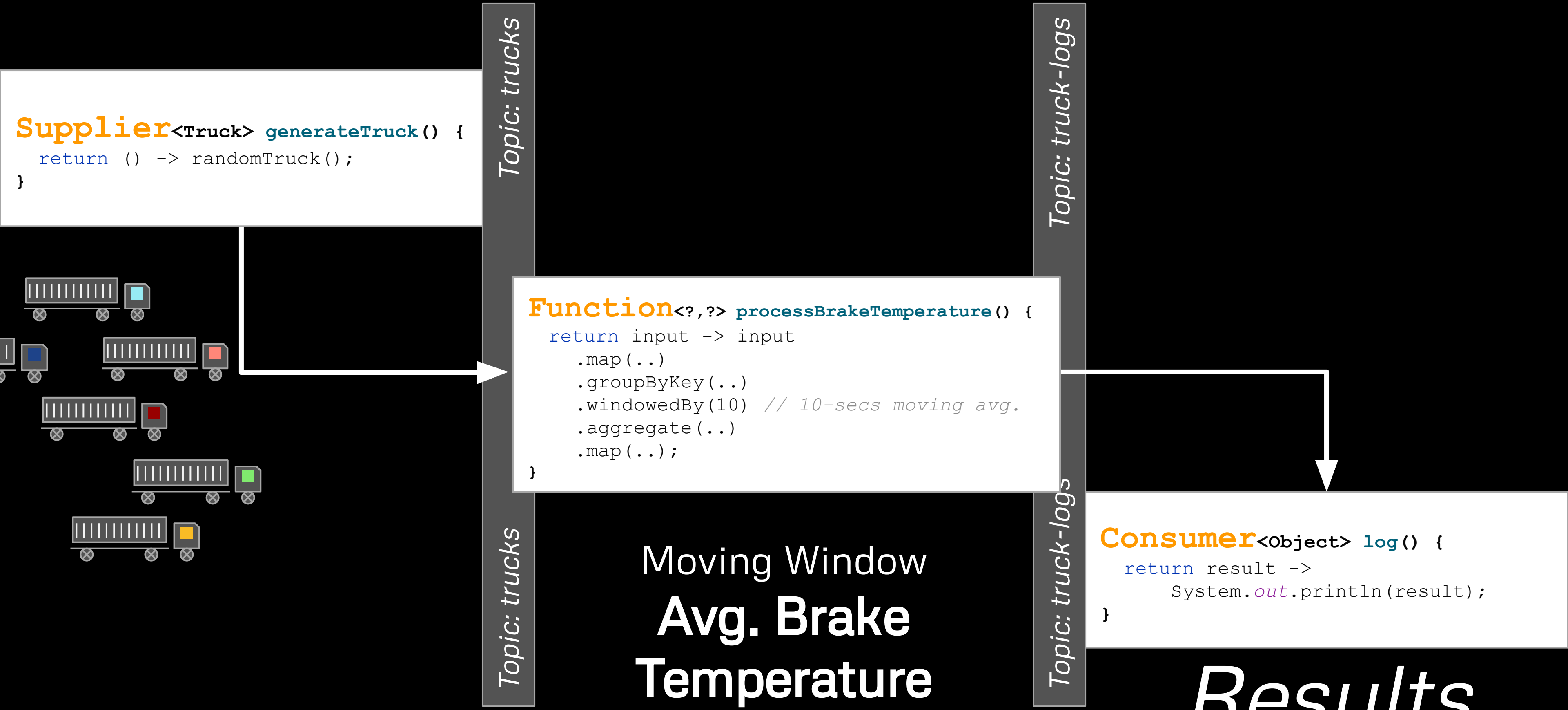
15
Mins

Lab #2: Event Streaming

Brake Temperature Analysis



SpringOne



```
Supplier<Truck> generateTruck() {
  return () -> randomTruck();
}
```

```
Function<?,?> processBrakeTemperature() {
  return input -> input
    .map(..)
    .groupByKey(..)
    .windowedBy(10) // 10-secs moving avg.
    .aggregate(..)
    .map(..);
}
```

```
Consumer<Object> log() {
  return result ->
      System.out.println(result);
}
```

Moving Window

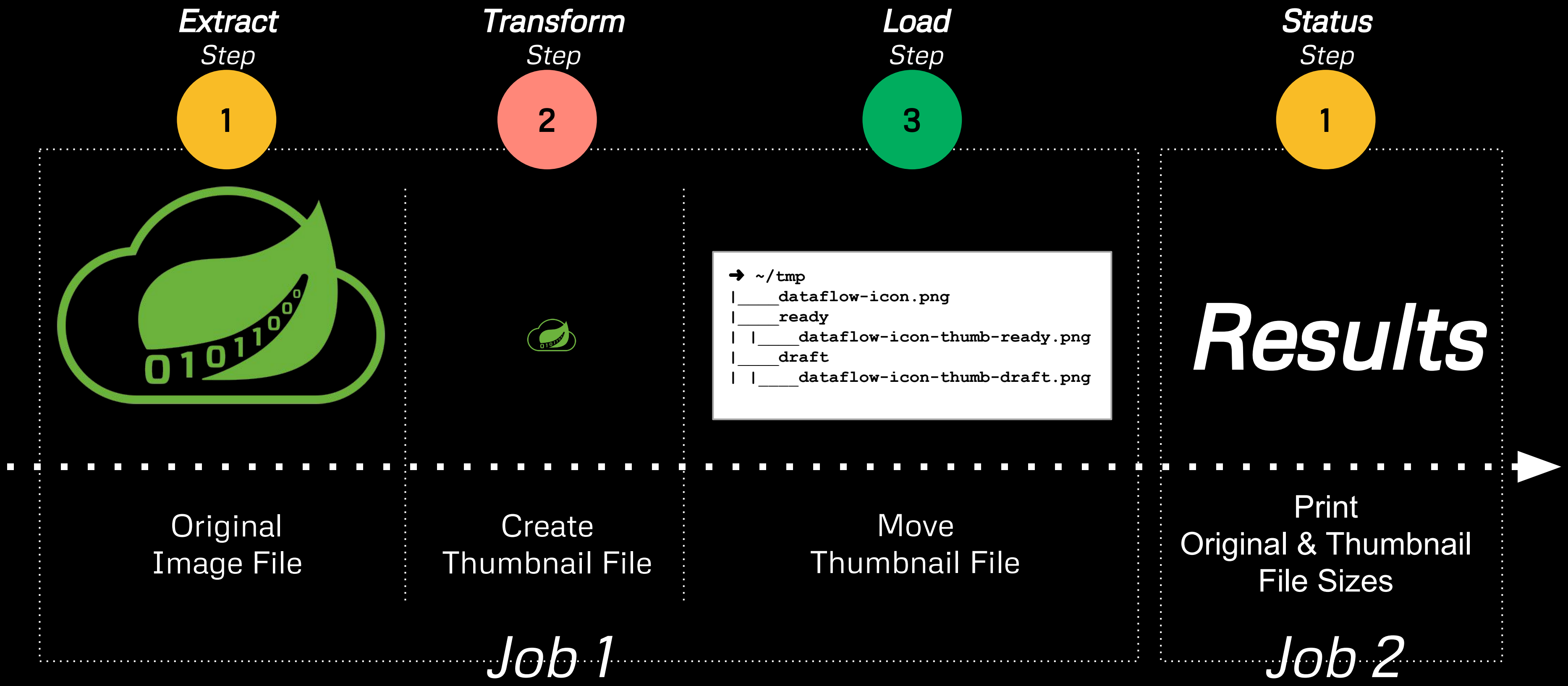
Avg. Brake Temperature

Results

SpringOne

15
Mins

Lab #2: Cloud-native ETL



Extract
Step
1

Original
Image File

Transform
Step
2

Create
Thumbnail File

Load
Step
3

```
➜ ~/tmp
|____dataflow-icon.png
|____ready
| |____dataflow-icon-thumb-ready.png
|____draft
| |____dataflow-icon-thumb-draft.png
```

Move
Thumbnail File

Job 1

Status
Step
1

Results

Print
Original & Thumbnail
File Sizes

Job 2

Resources

Spring Cloud Data Flow: https://dataflow.spring.io/

Source Code: https://github.com/sabbyanandan/SpringOne2020

Lab Instructions: https://hackmd.io/@sabbyanandan/B1bDf74fv

Slides: Speaker Deck

Batching for the Modern Enterprise

Michael Minella, VMware

Sep 2nd

Game of Streams 🐉: How to Tame and Get the Most from Your Messaging Platforms

Mark Heckler, VMware

Sep 2nd

Walking Through Spring Cloud Data Flow

Glenn Renfro, VMware
Ilayaperumal Gopinathan, VMware

Sep 3rd

IoT Scale Event-Stream Processing for Connected Fleet at Penske

Krishna Gogineni, VMware
Shruti Modi, Penske

Sep 3rd