

Delivering Media at Scale: The Fast Data Perspective

Dean Wampler, Ph.D.
dean@lightbend.com
[@deanwampler](https://twitter.com/deanwampler)
polyglotprogramming.com/talks



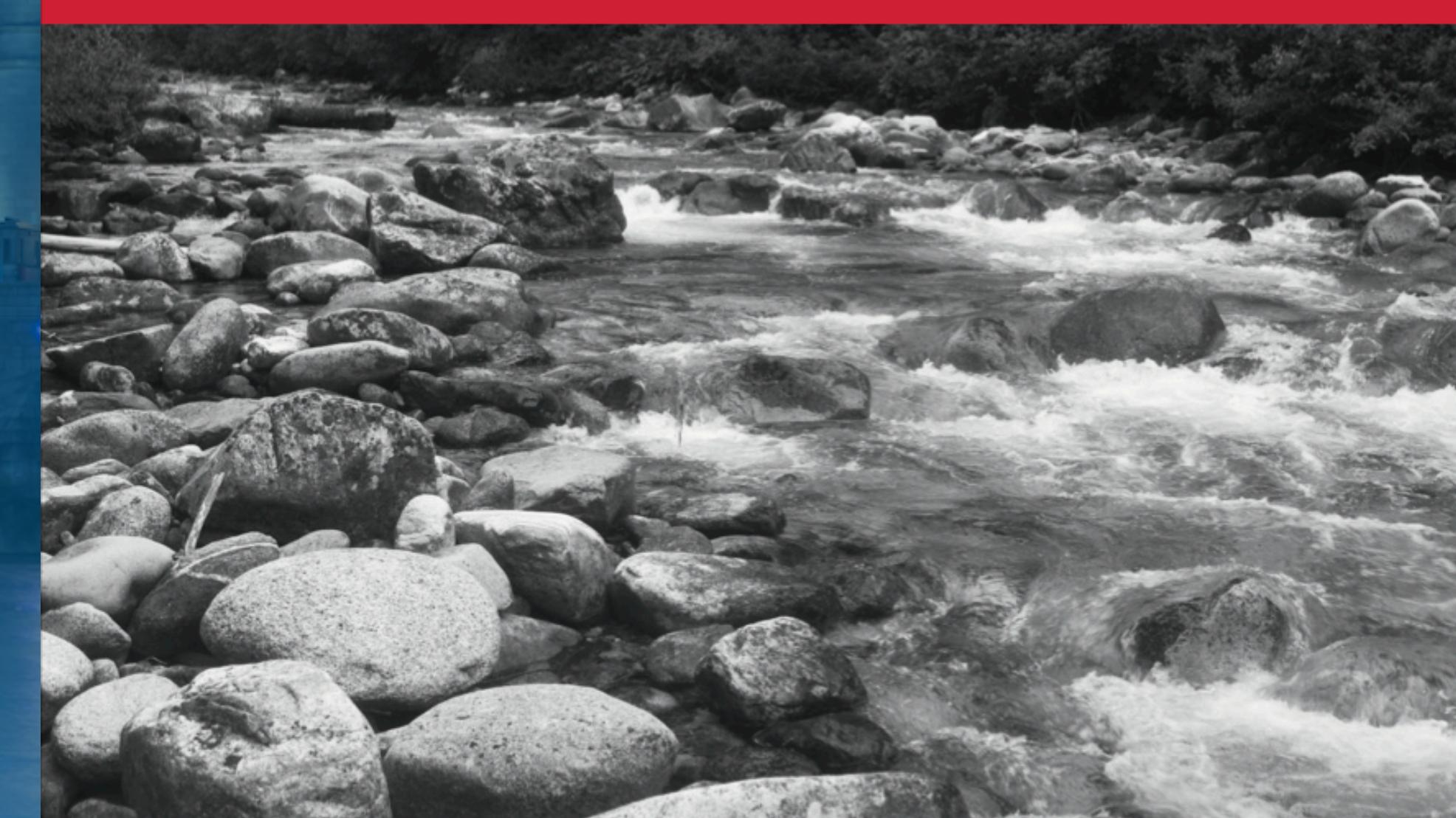


Ideas adapted
from this report

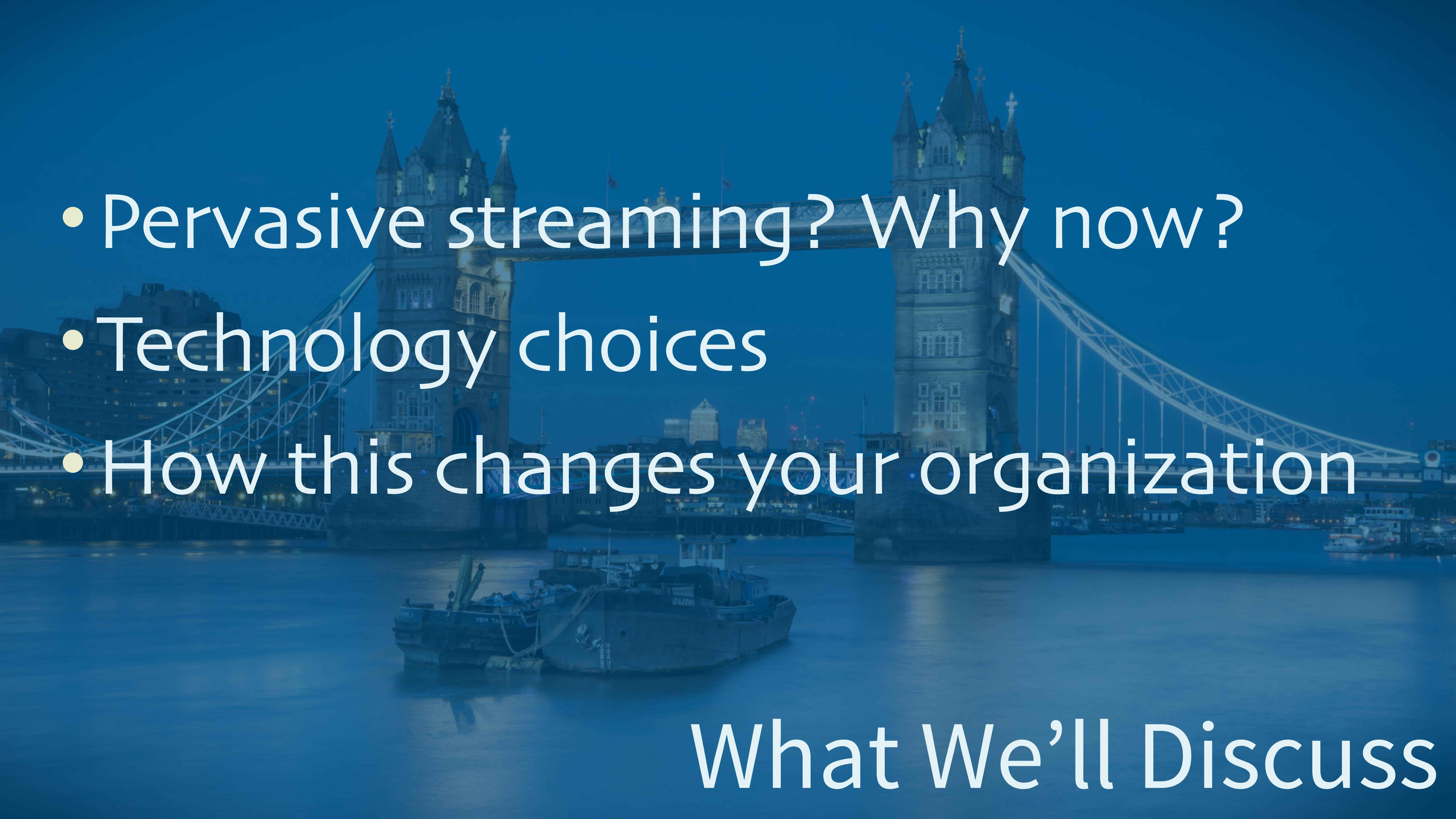
lightbend.com/fast-data-platform
2nd edition coming in October!

Fast Data Architectures for Streaming Applications

Getting Answers Now from
Data Sets that Never End



Dean Wampler

- 
- Pervasive streaming? Why now?
 - Technology choices
 - How this changes your organization

What We'll Discuss

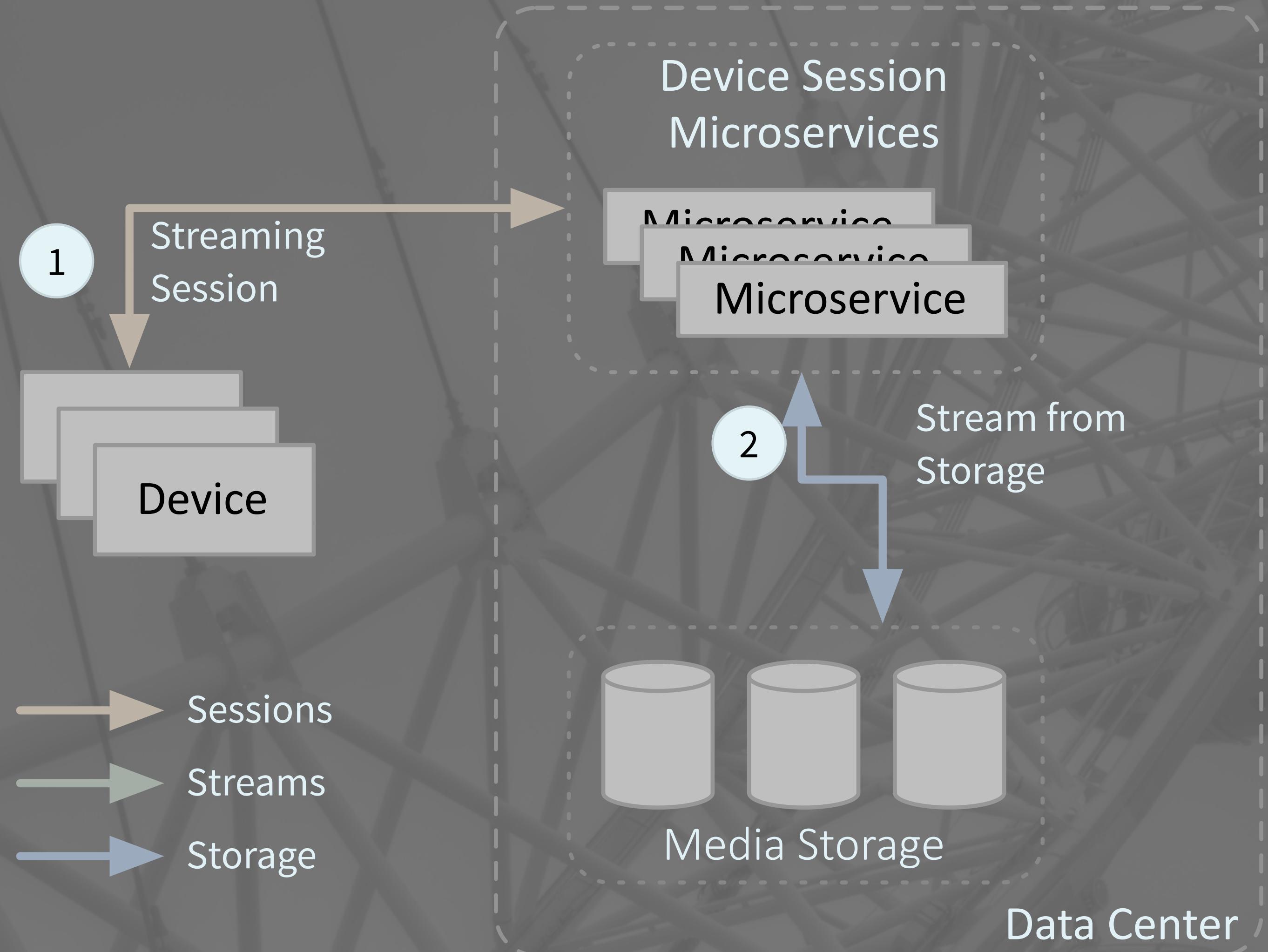
Pervasive Streaming



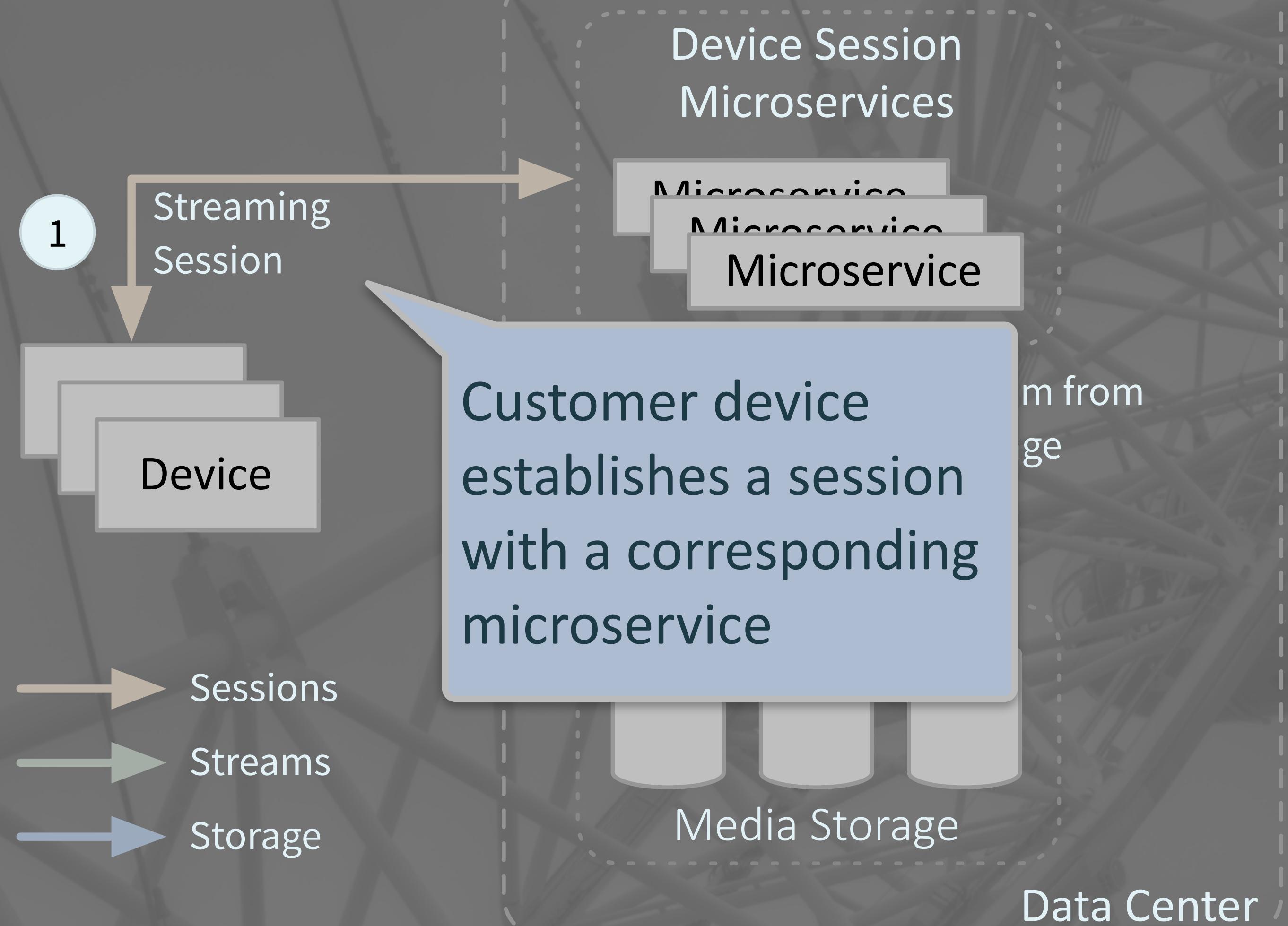
Pervasive Streaming

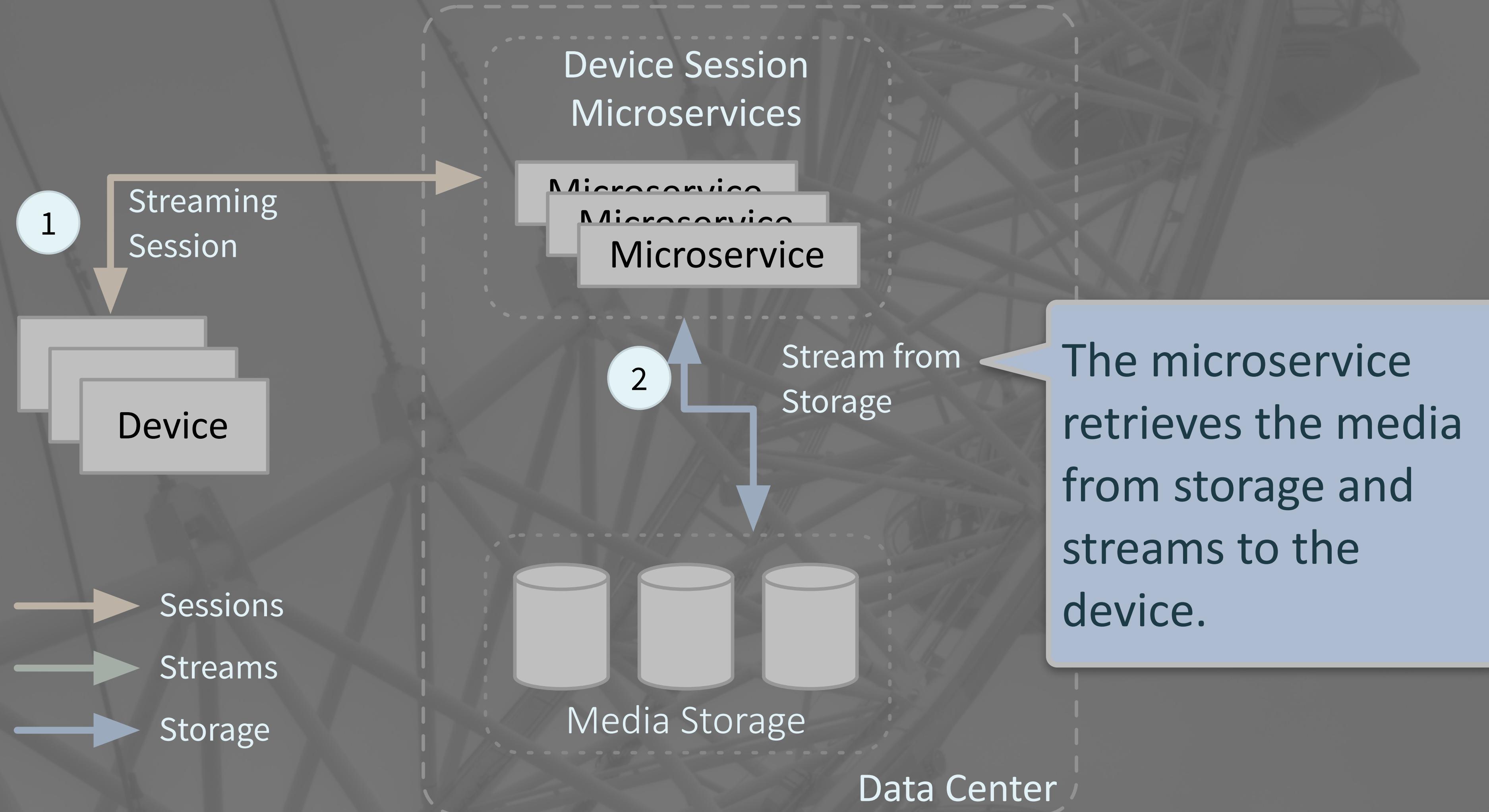
- Media streaming is obvious ;)
- How should we implement it?
- Other uses for streaming?





Media Streaming - V1

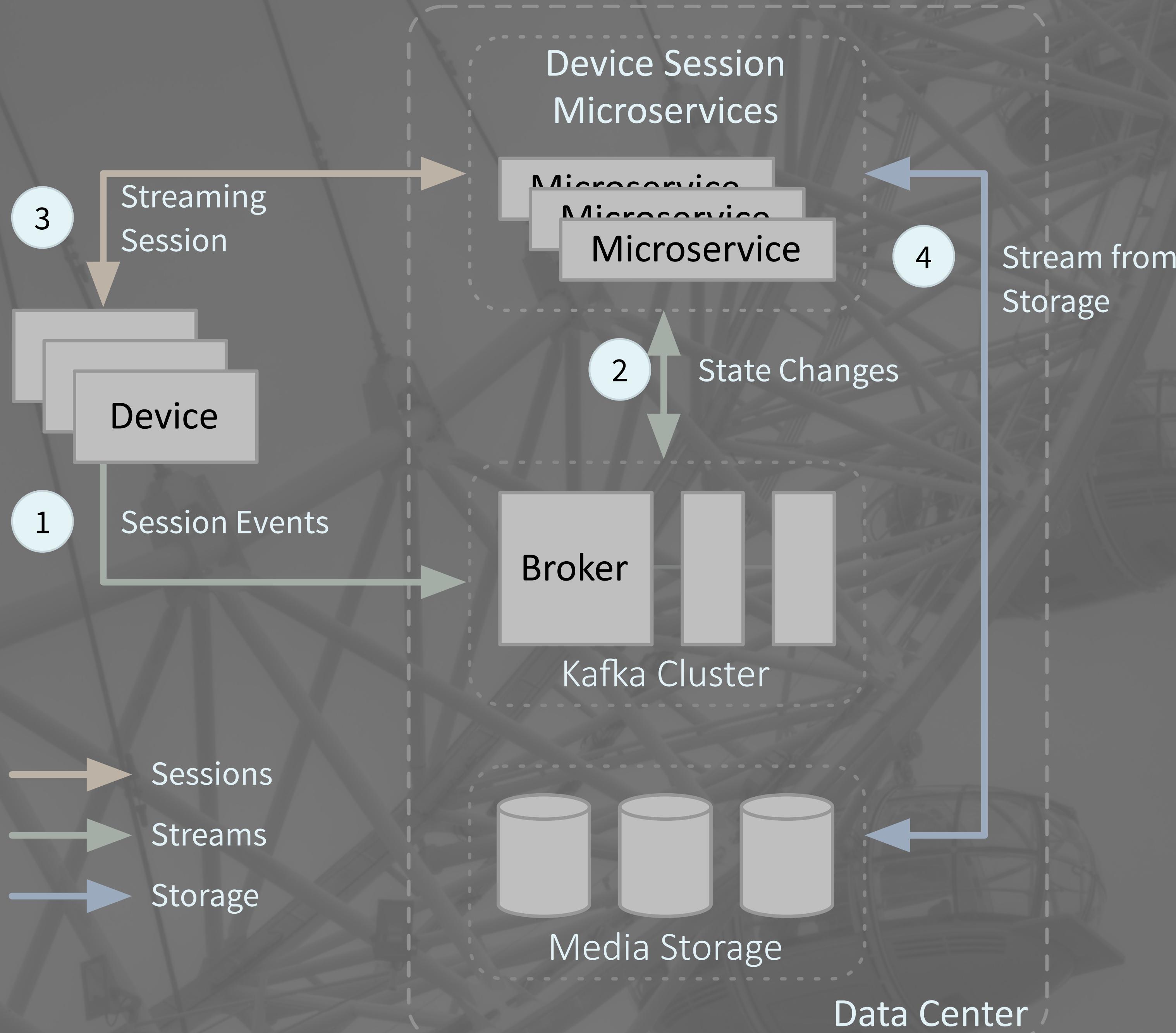




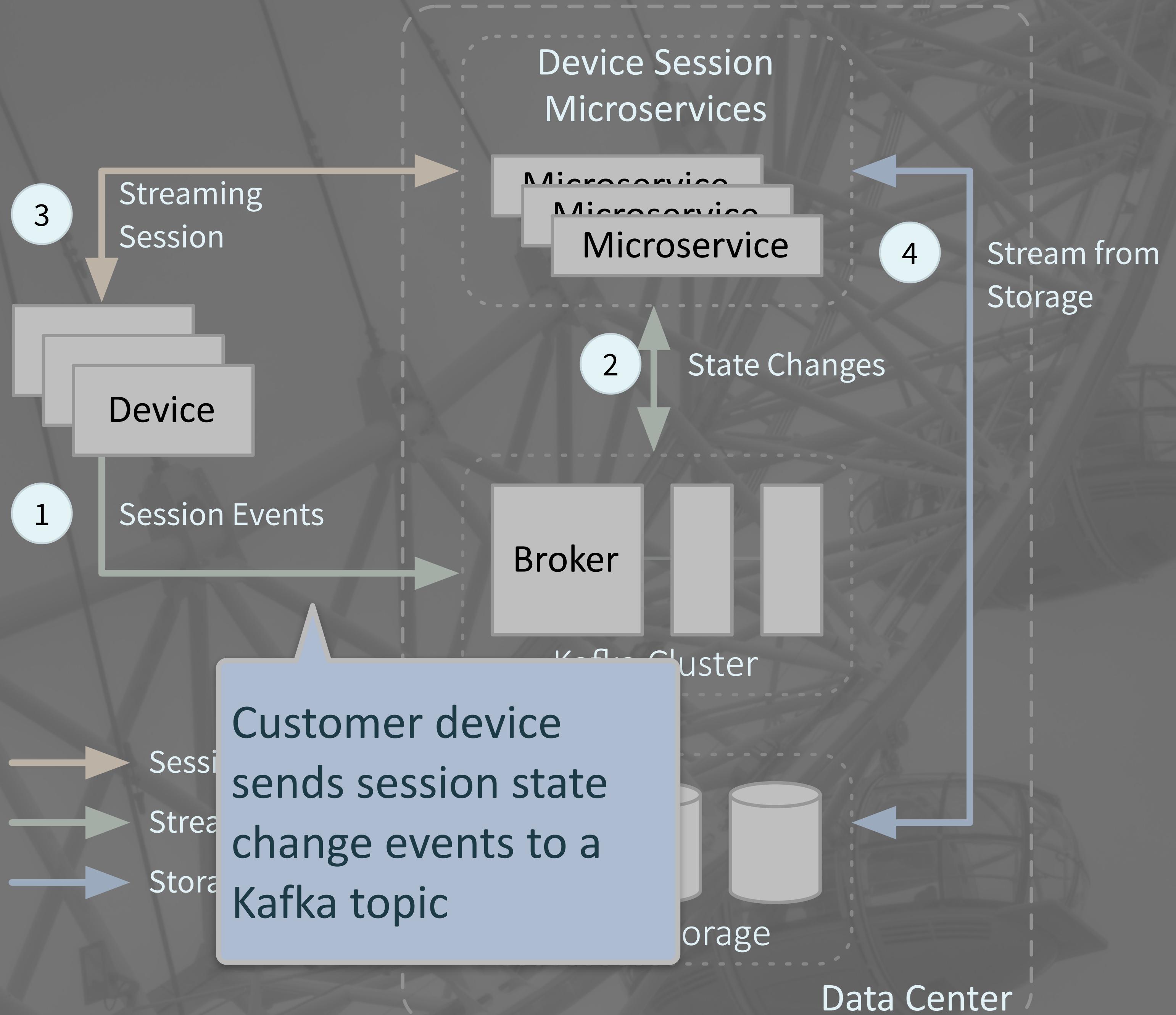
The microservice retrieves the media from storage and streams to the device.

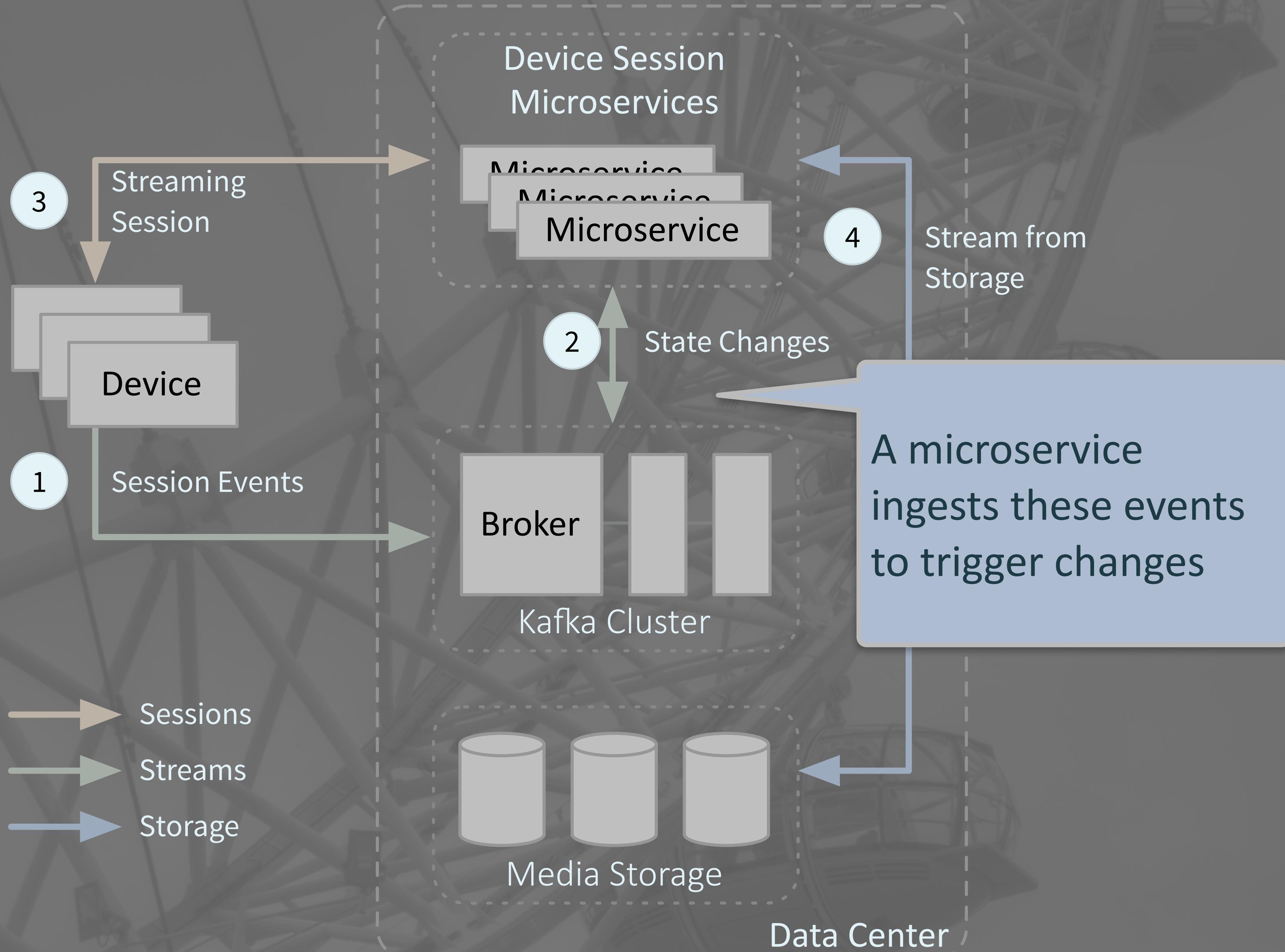
Challenges

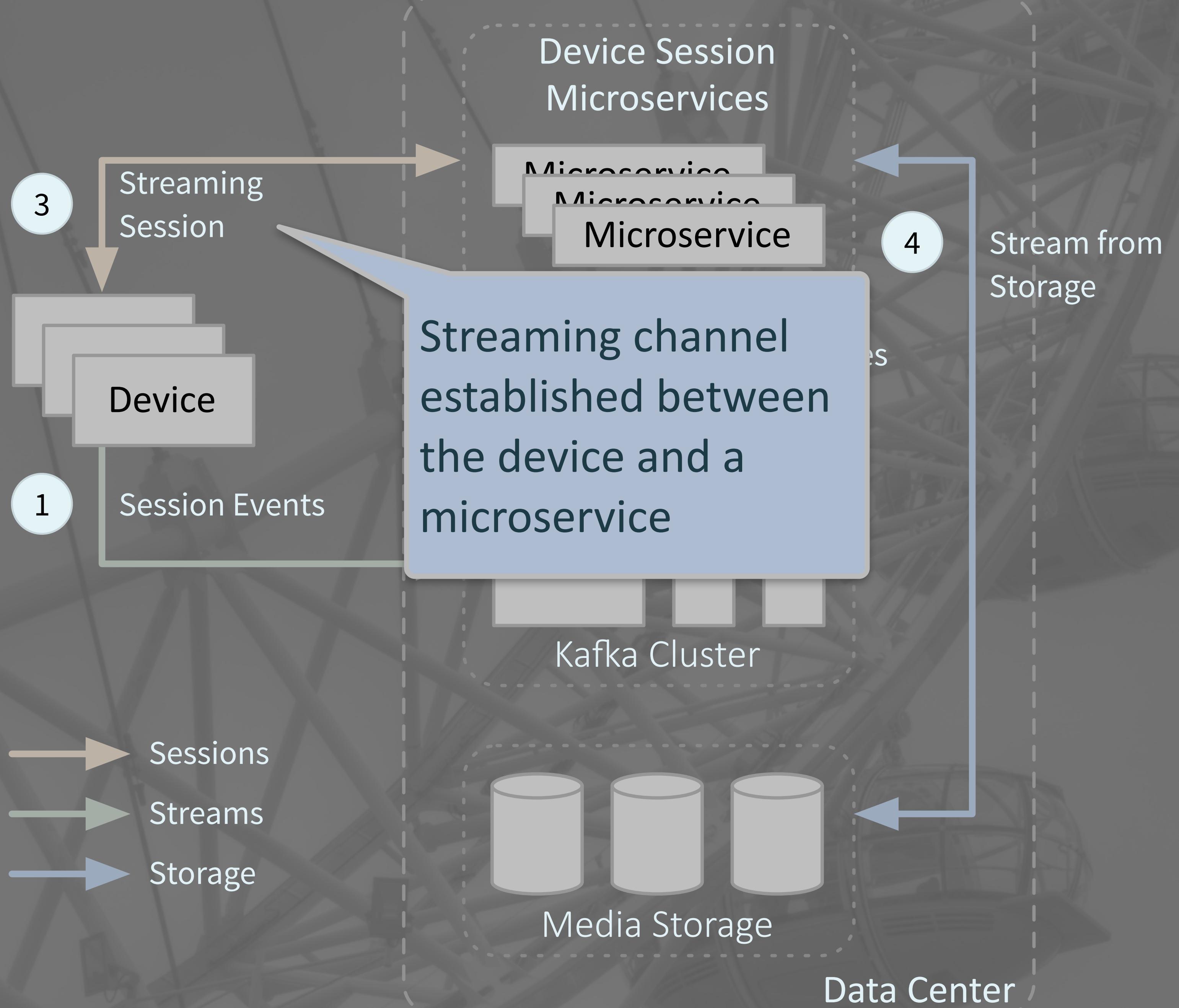
- Does this really scale to lots of devices?
- What if a connection or microservice dies?

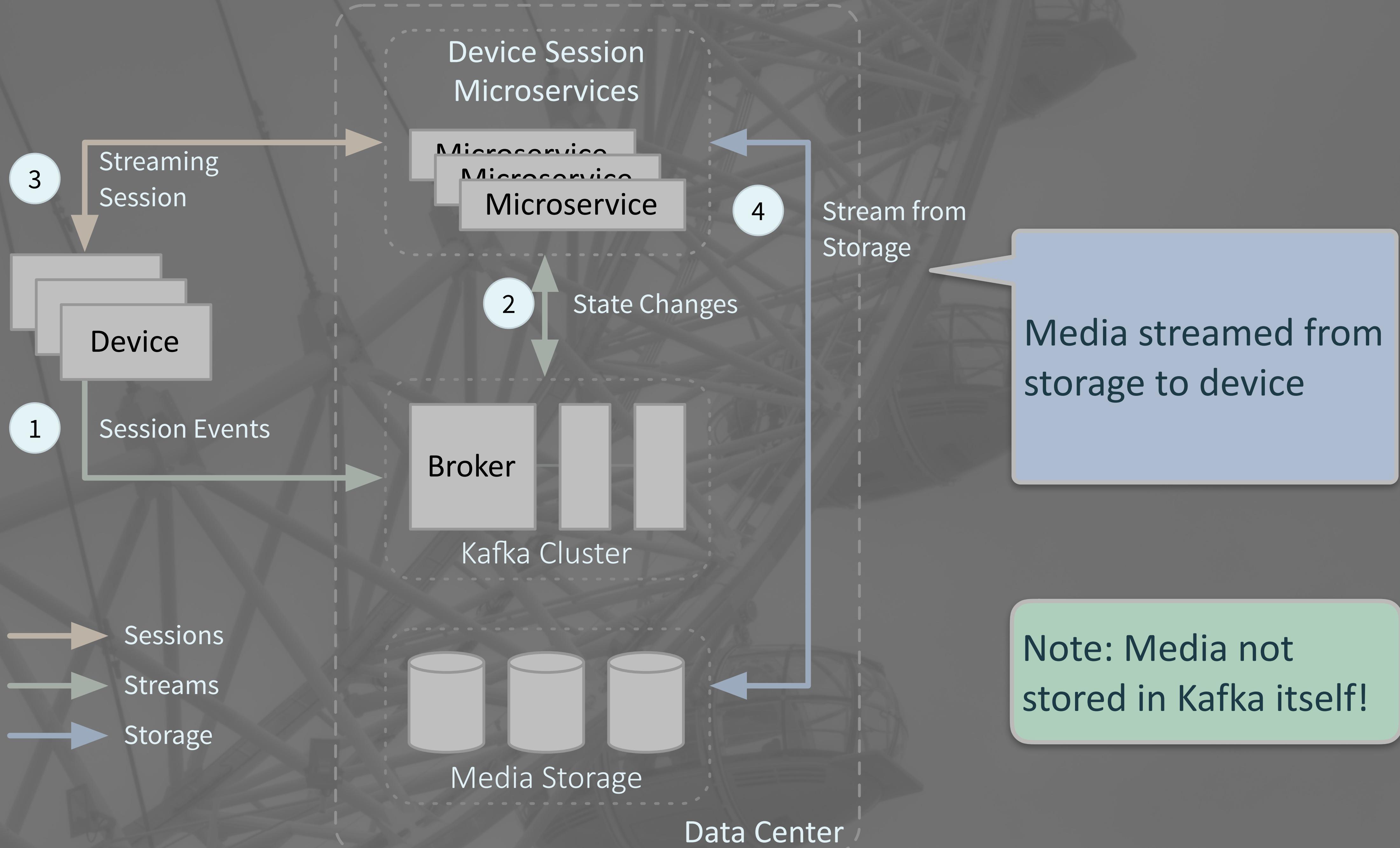


Media Streaming - V2









Advantages

- Fully asynchronous - no blocking REST...
- Recover from outages with “management” microservice that reads the same Kafka topics

Other Uses for Streaming

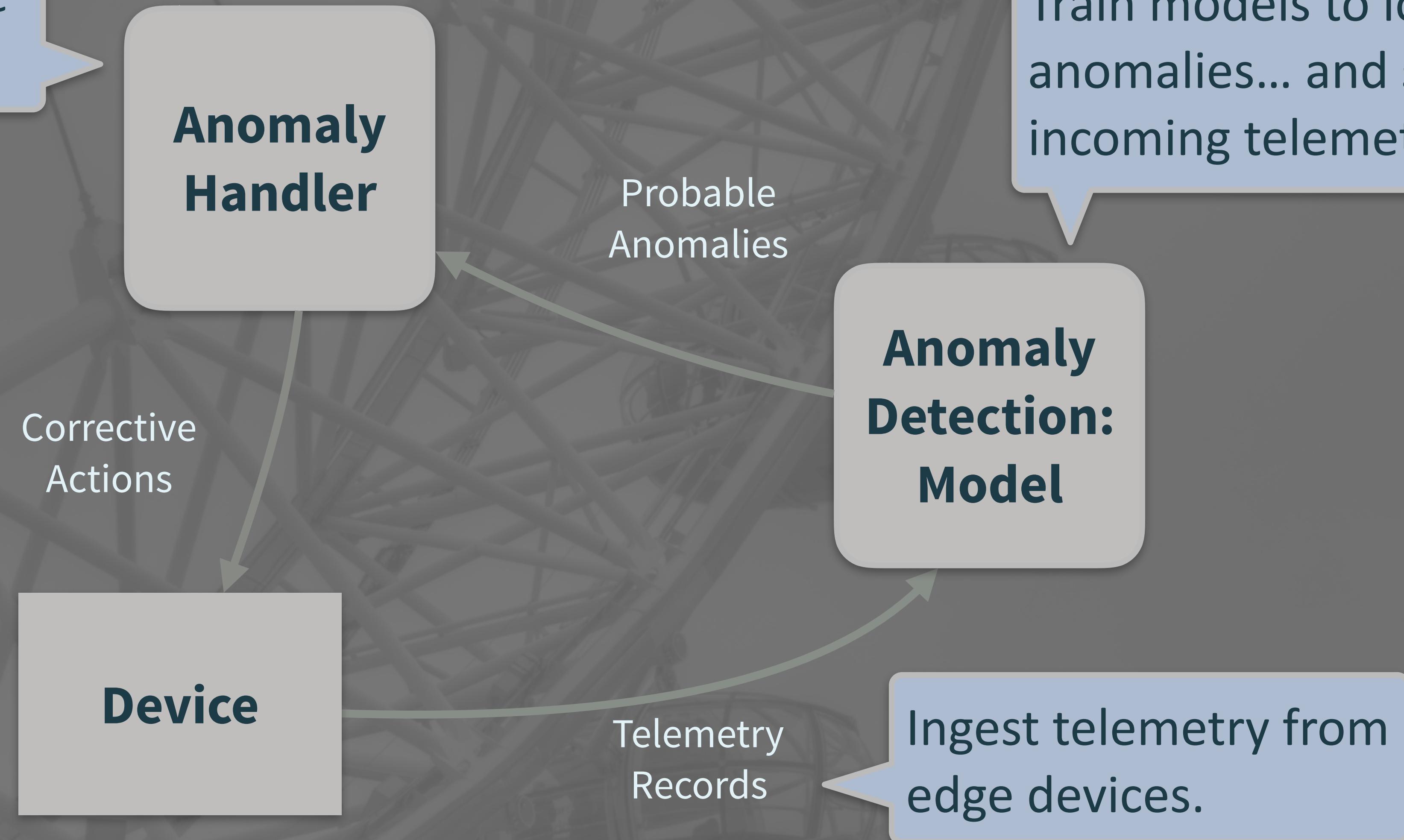


Predictive Analytics

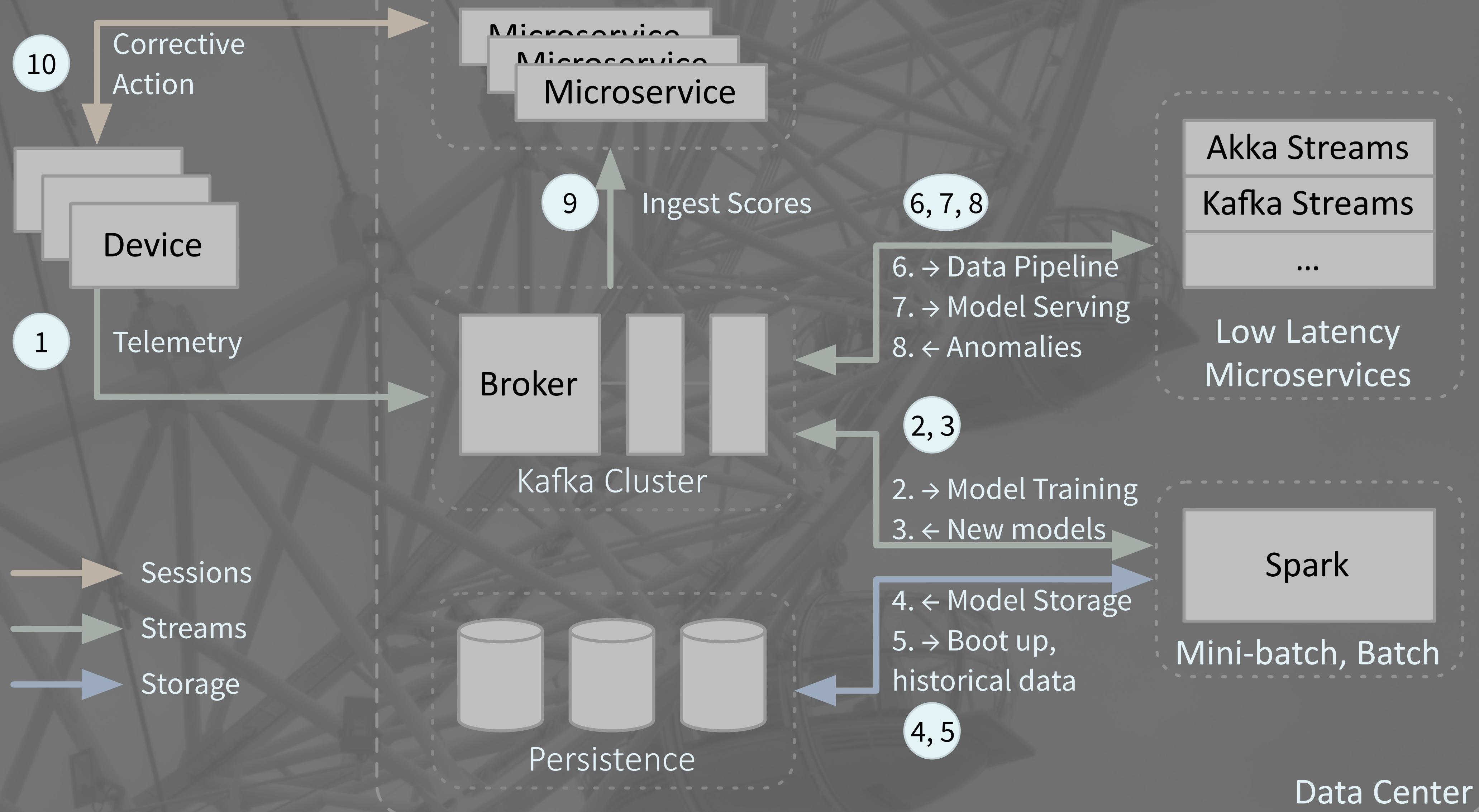
- ML models applied to device telemetry to detect anomalies
- Preemptive maintenance prevents potential failures that would impact users

Predictive Analytics - Core Idea

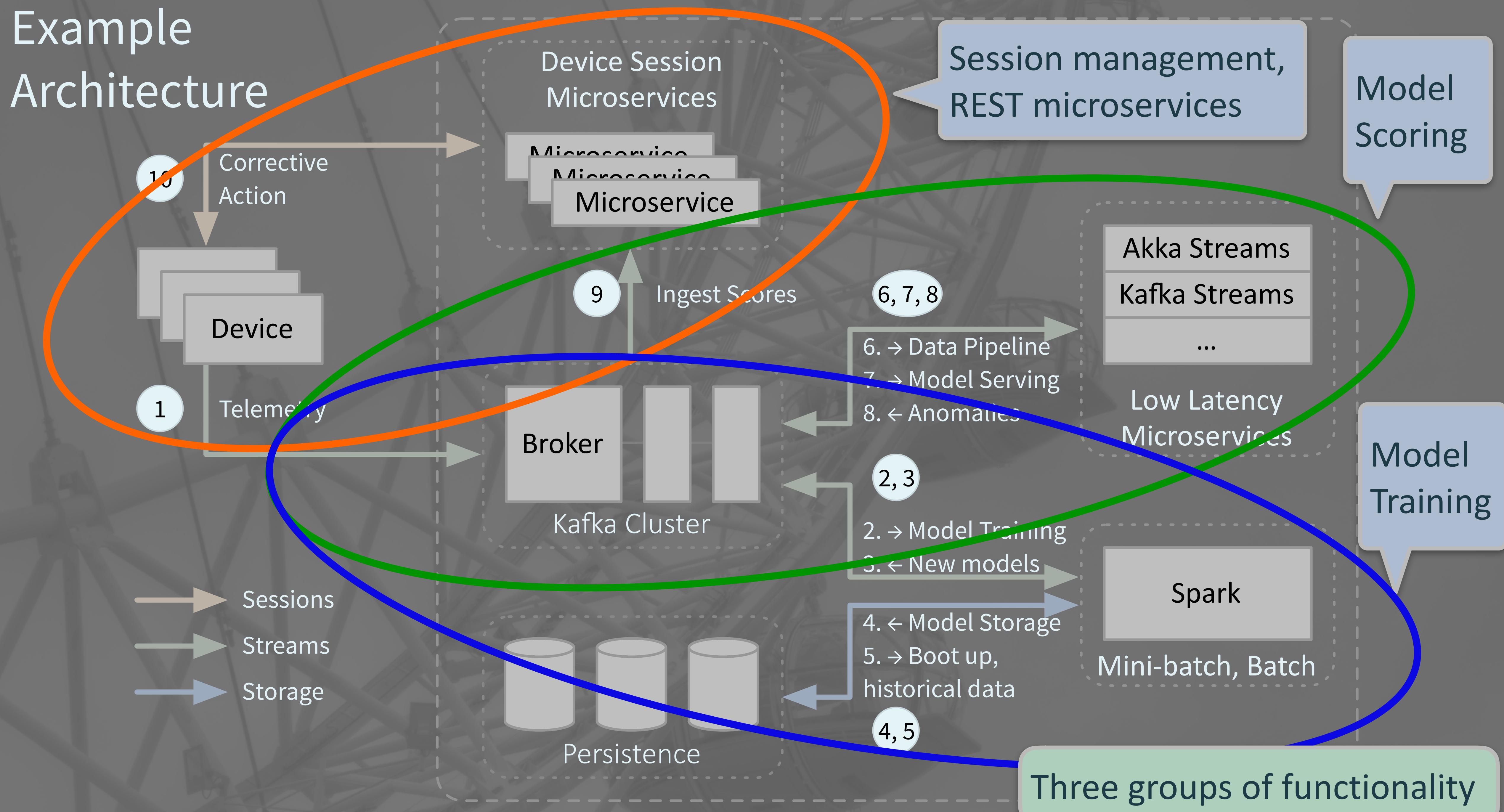
Handle anomaly: move activity off component, schedule maintenance window to replace it.



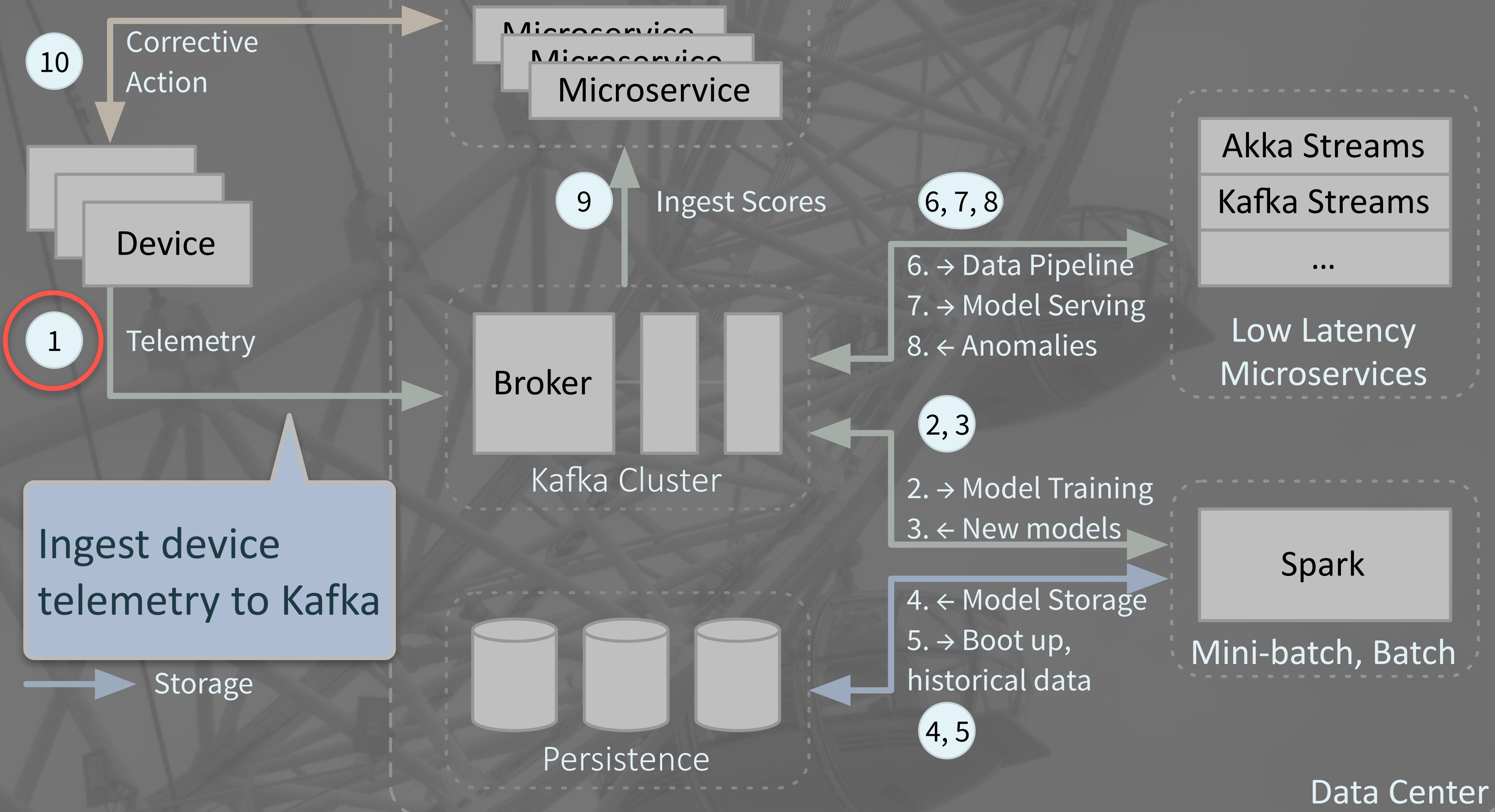
Example Architecture



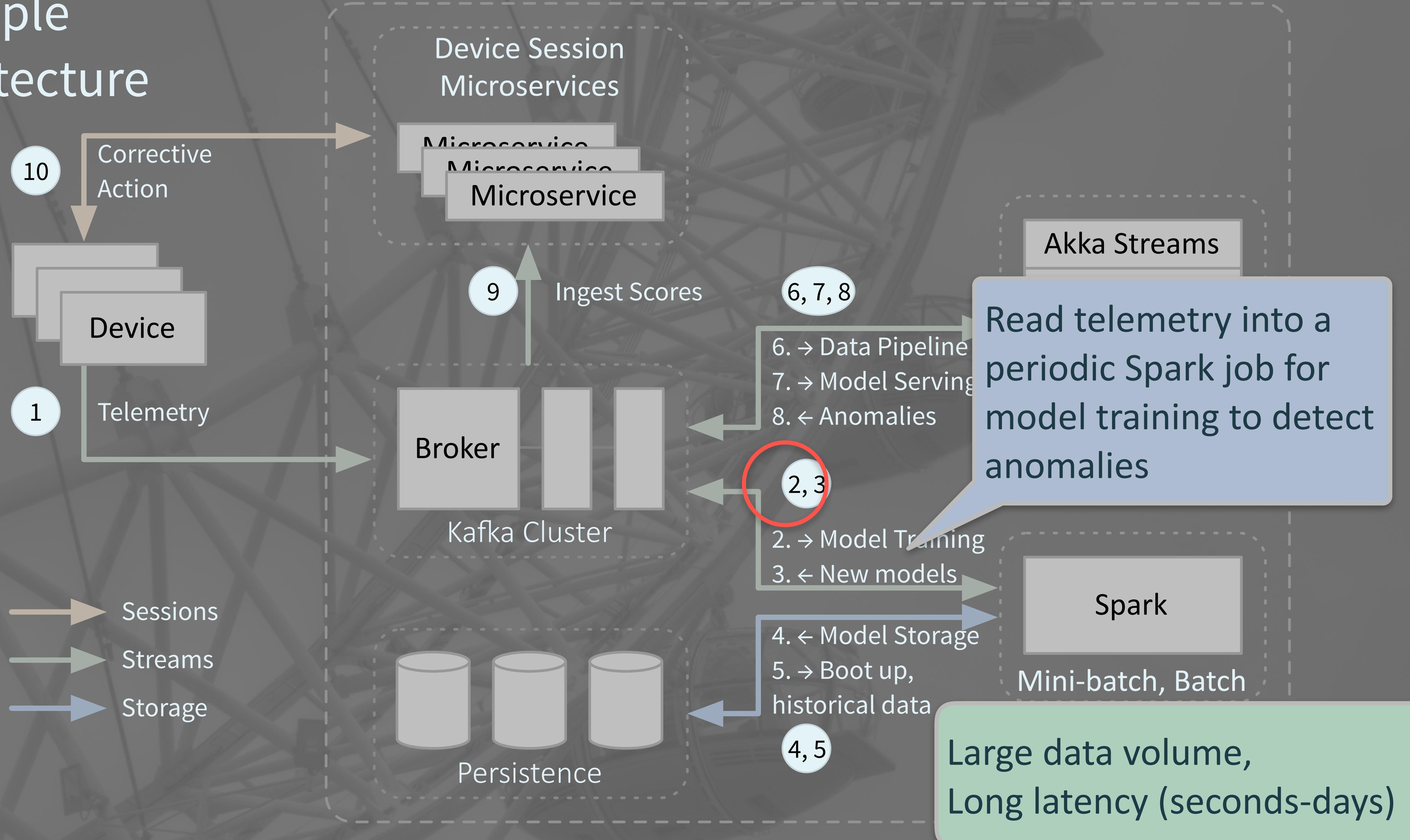
Example Architecture



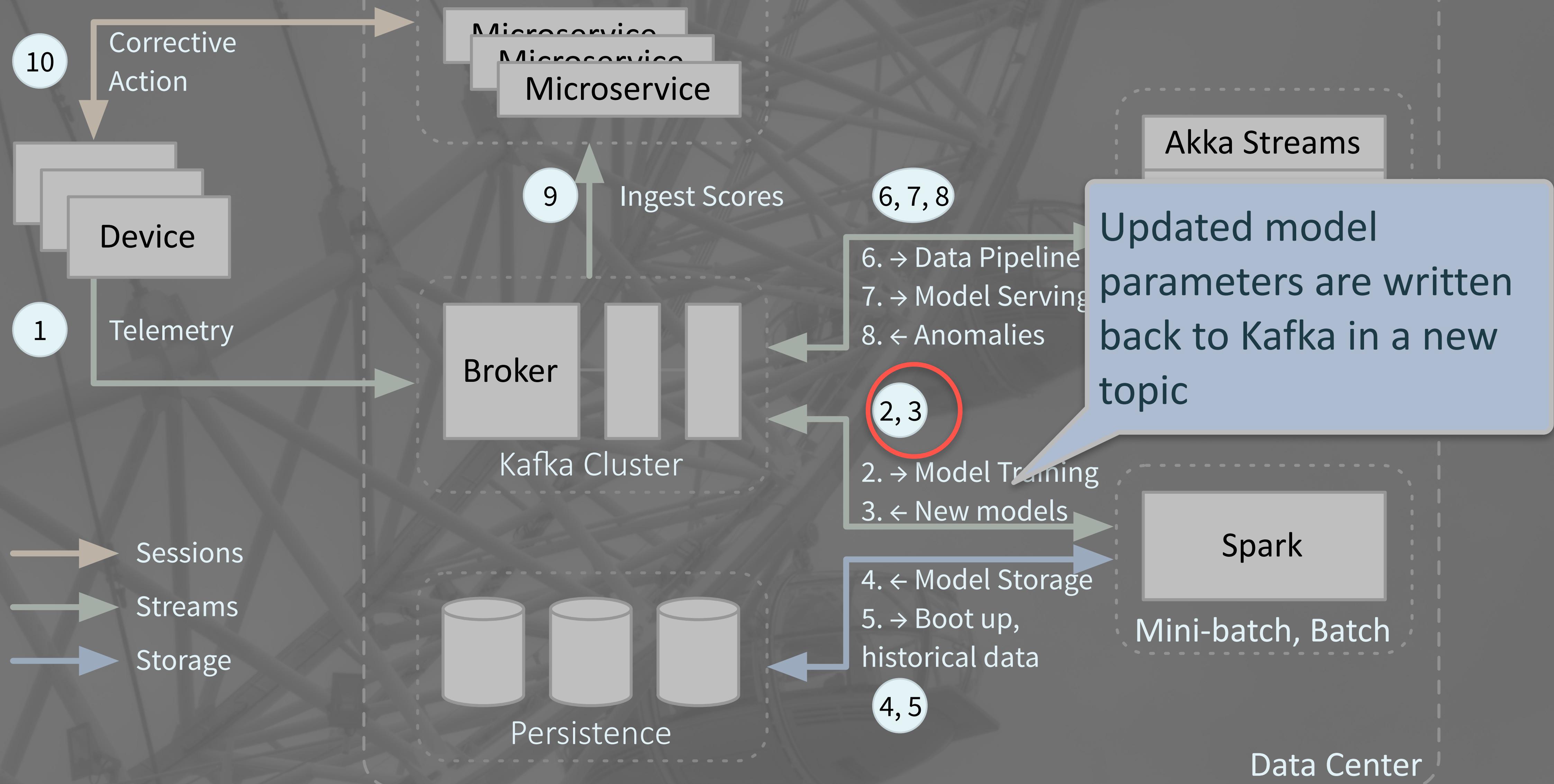
Example Architecture



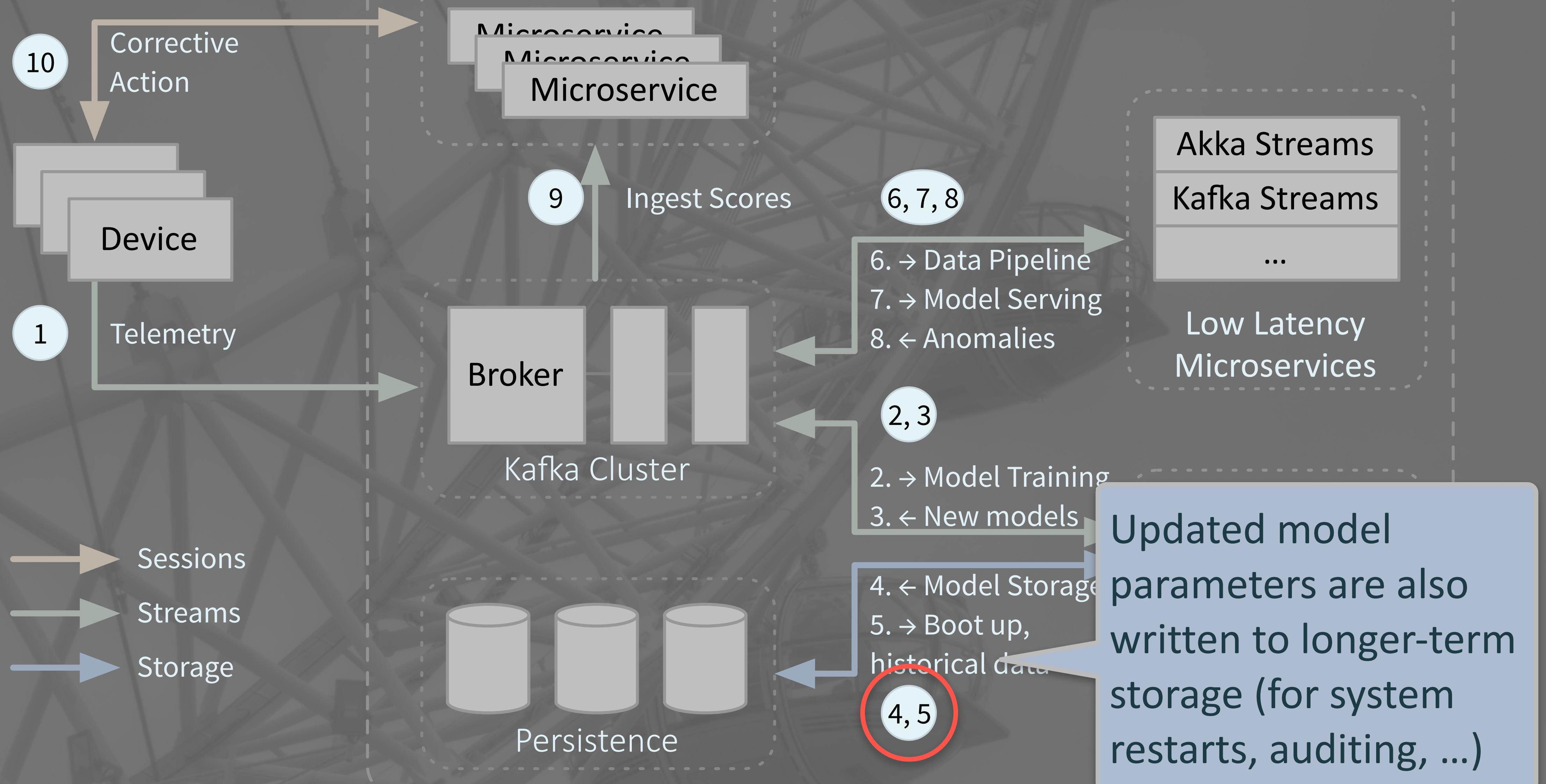
Example Architecture



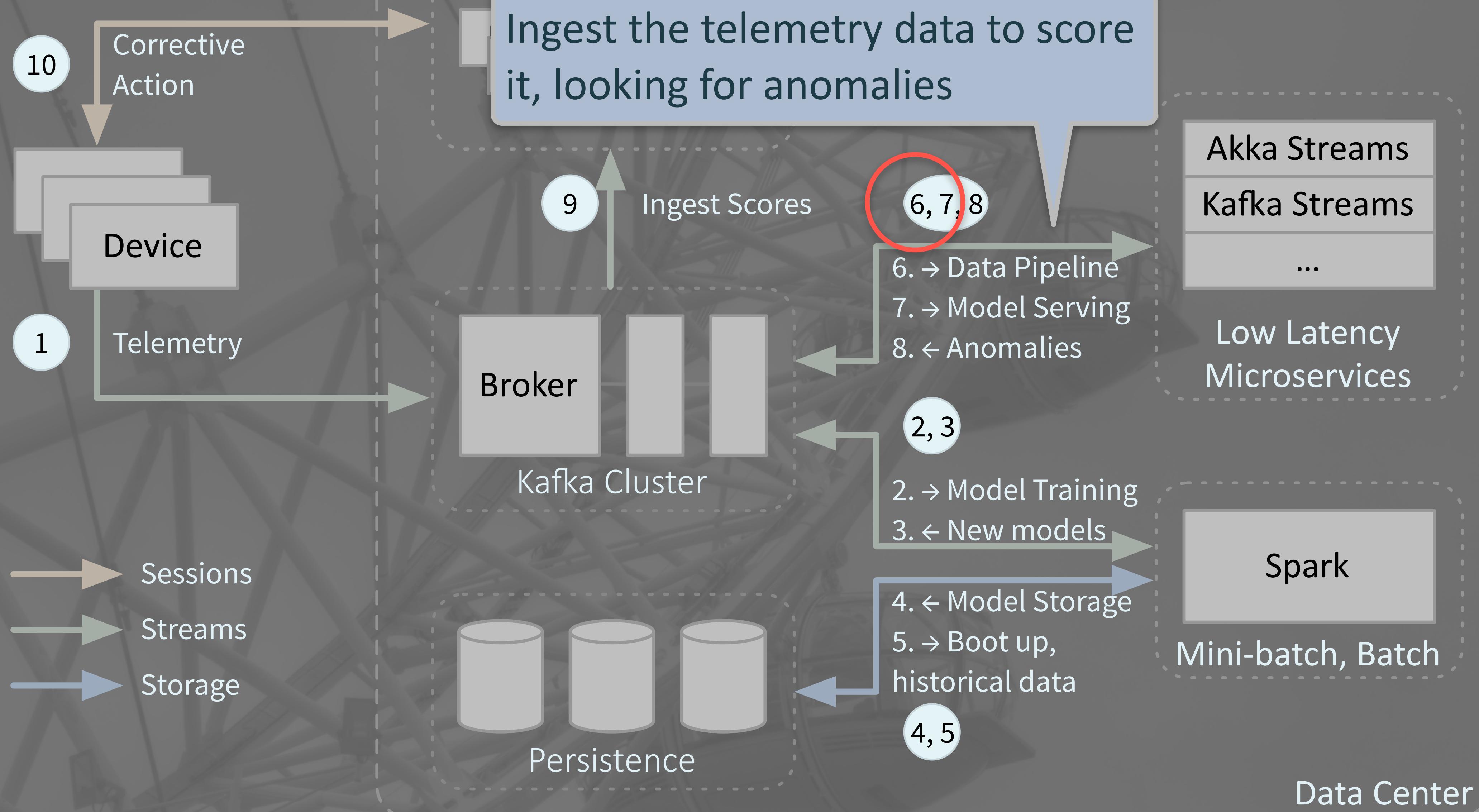
Example Architecture



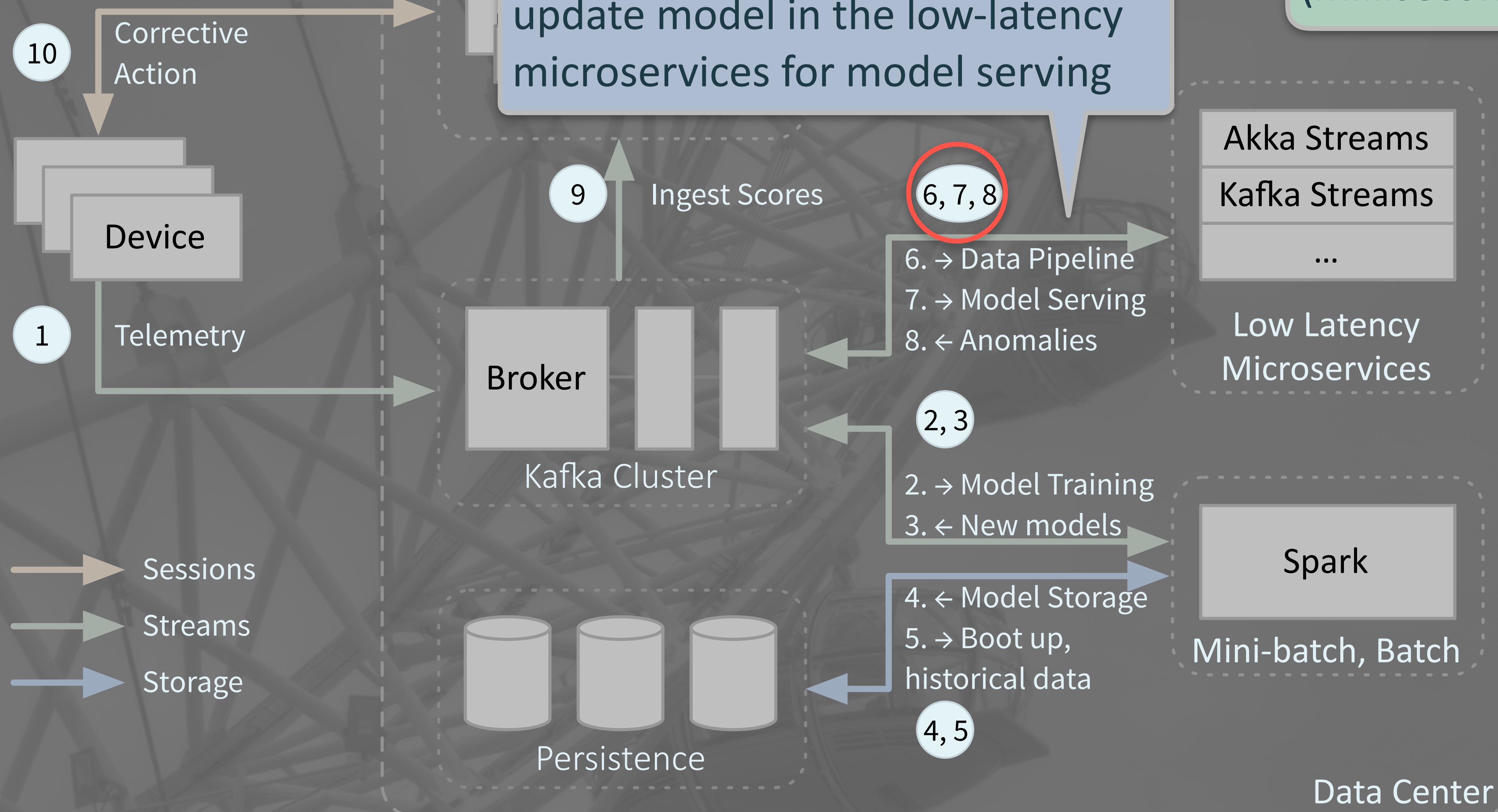
Example Architecture



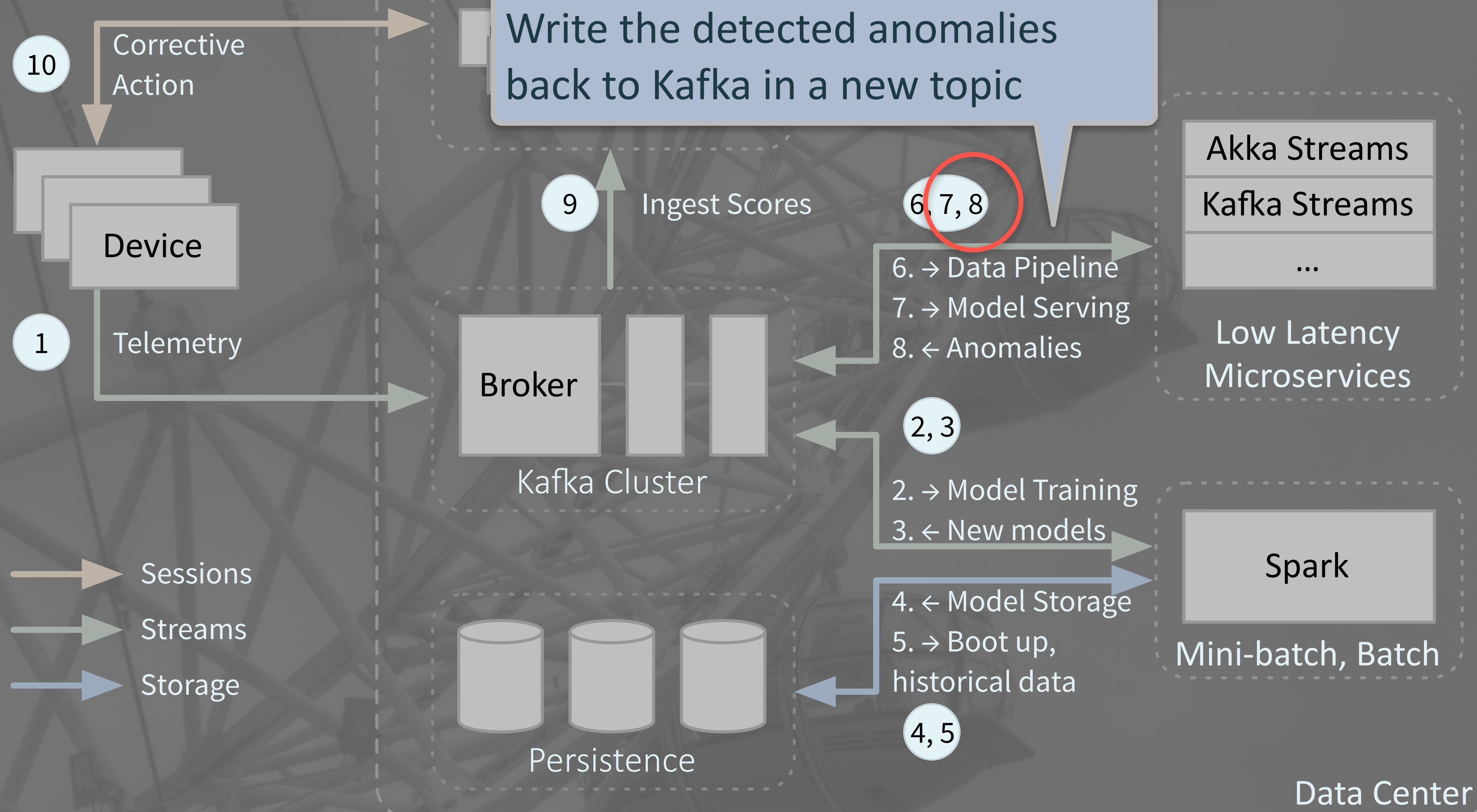
Example Architecture



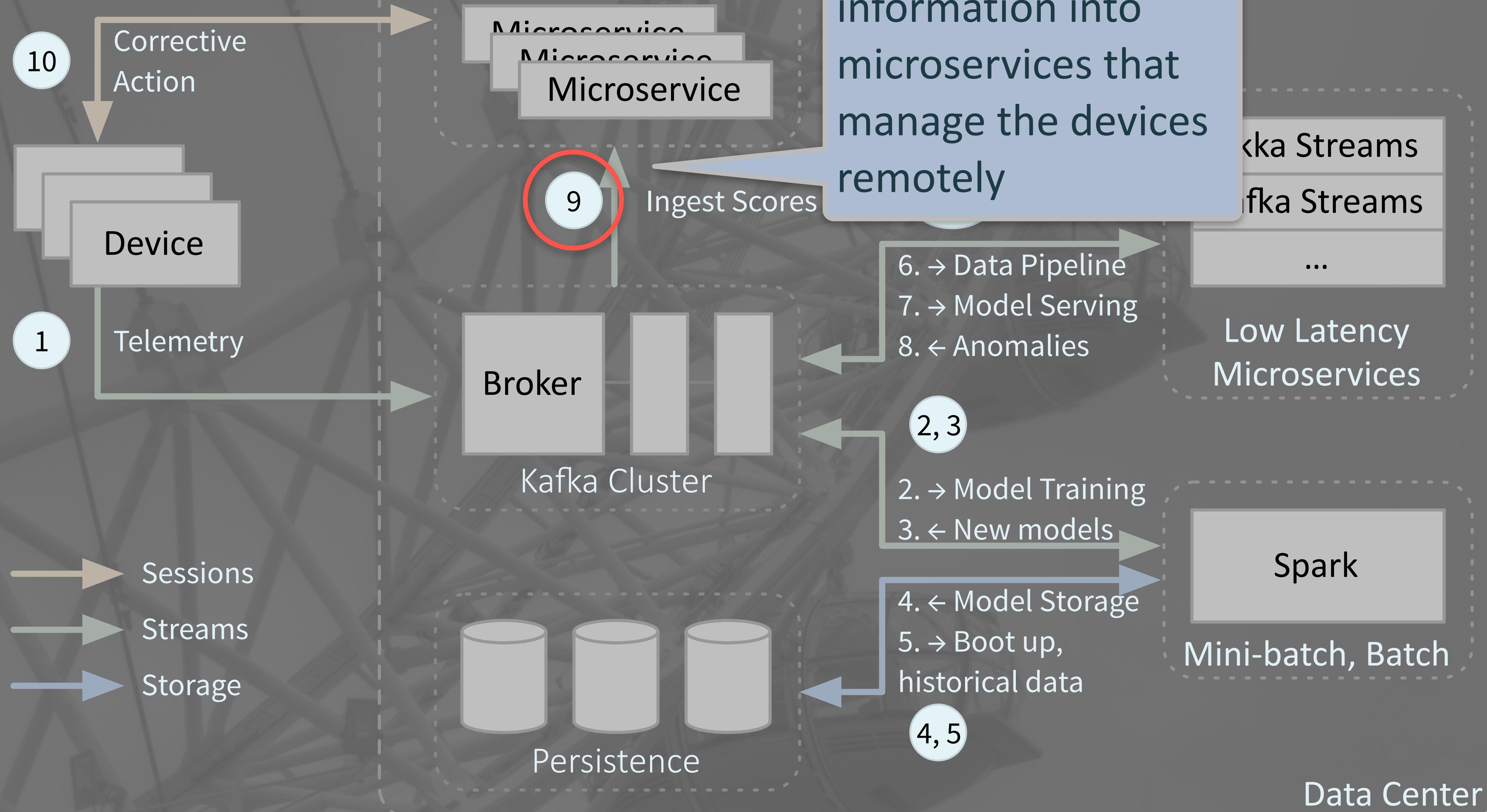
Example Architecture



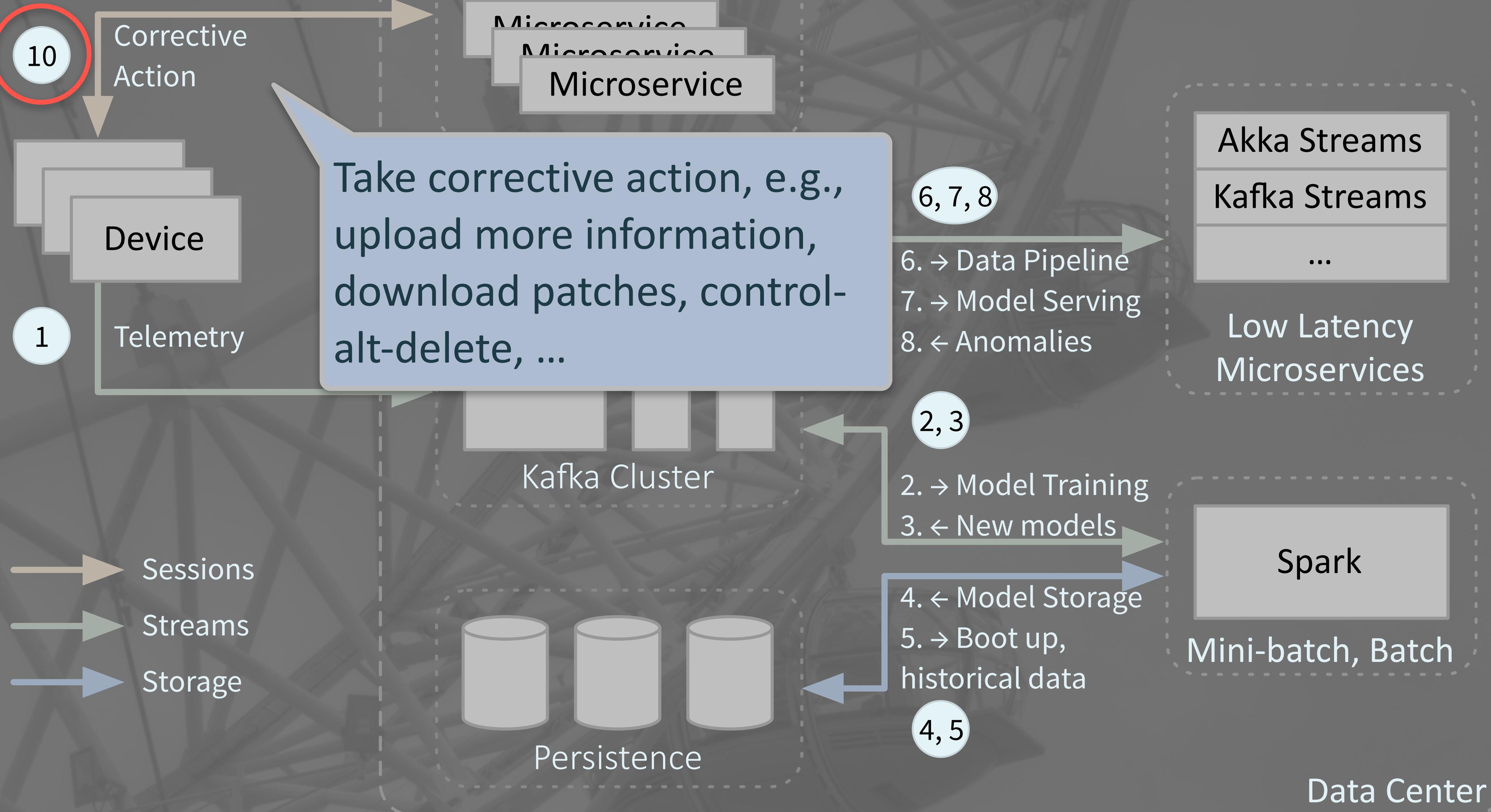
Example Architecture



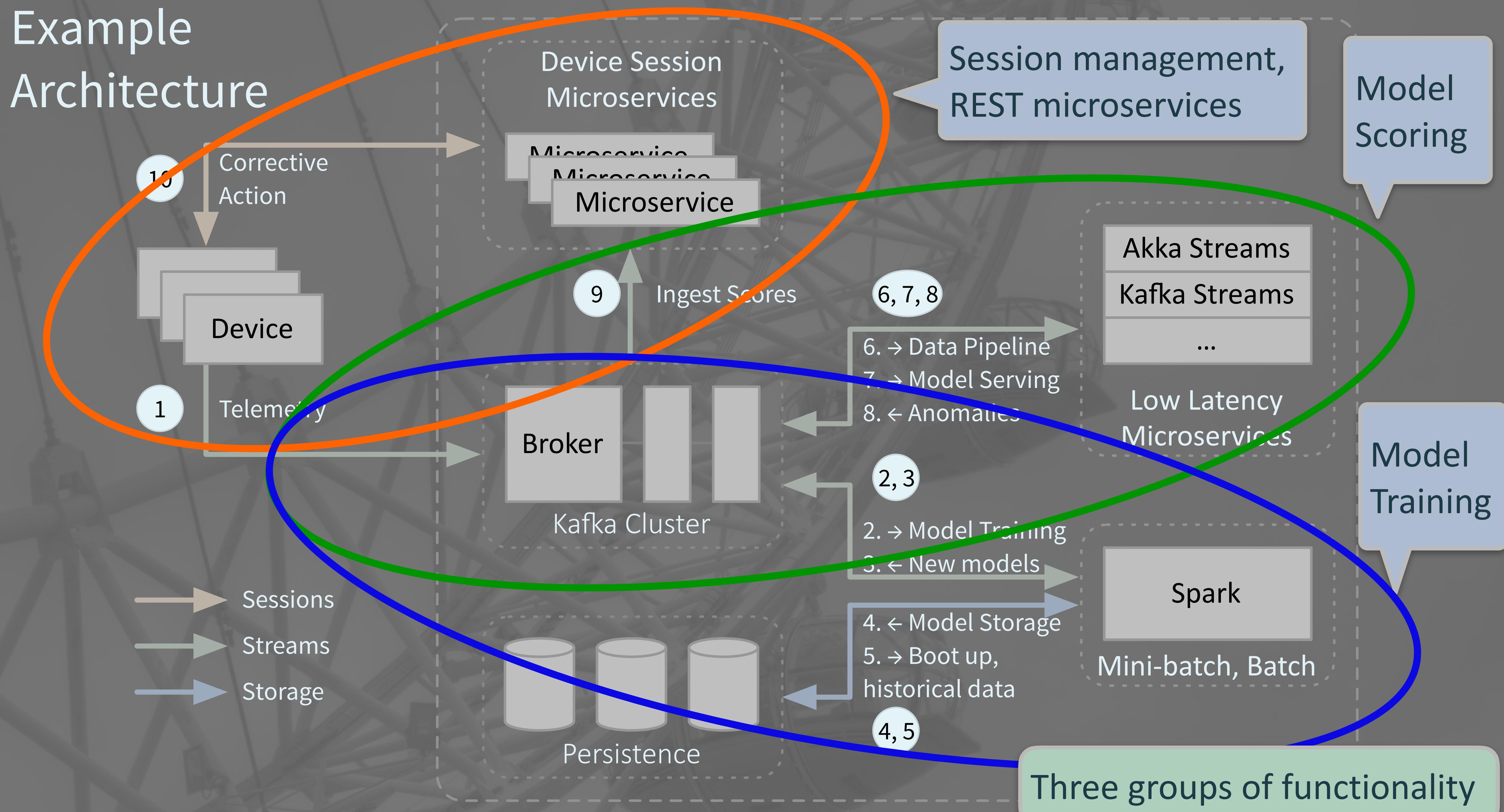
Example Architecture



Example Architecture

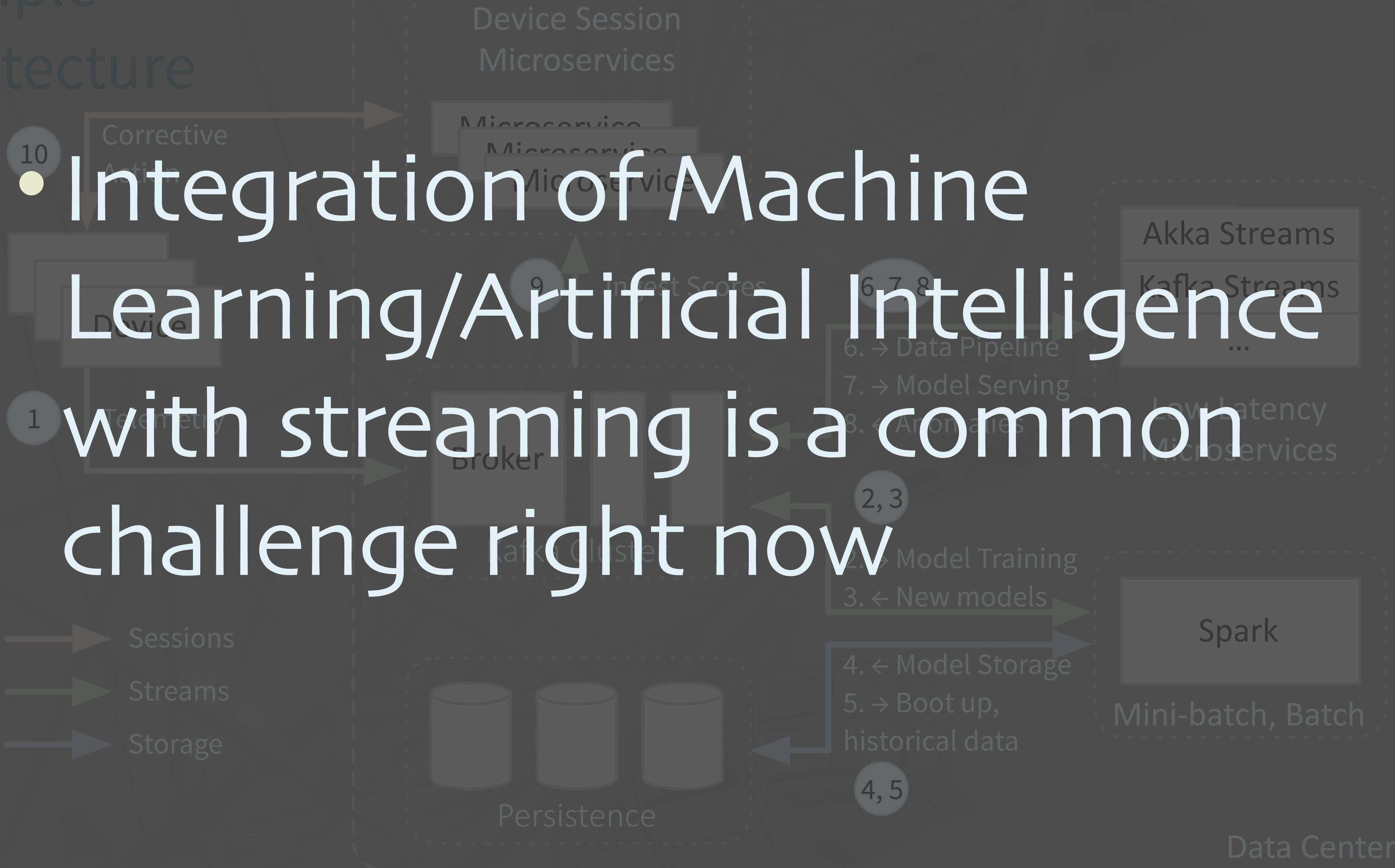


Example Architecture



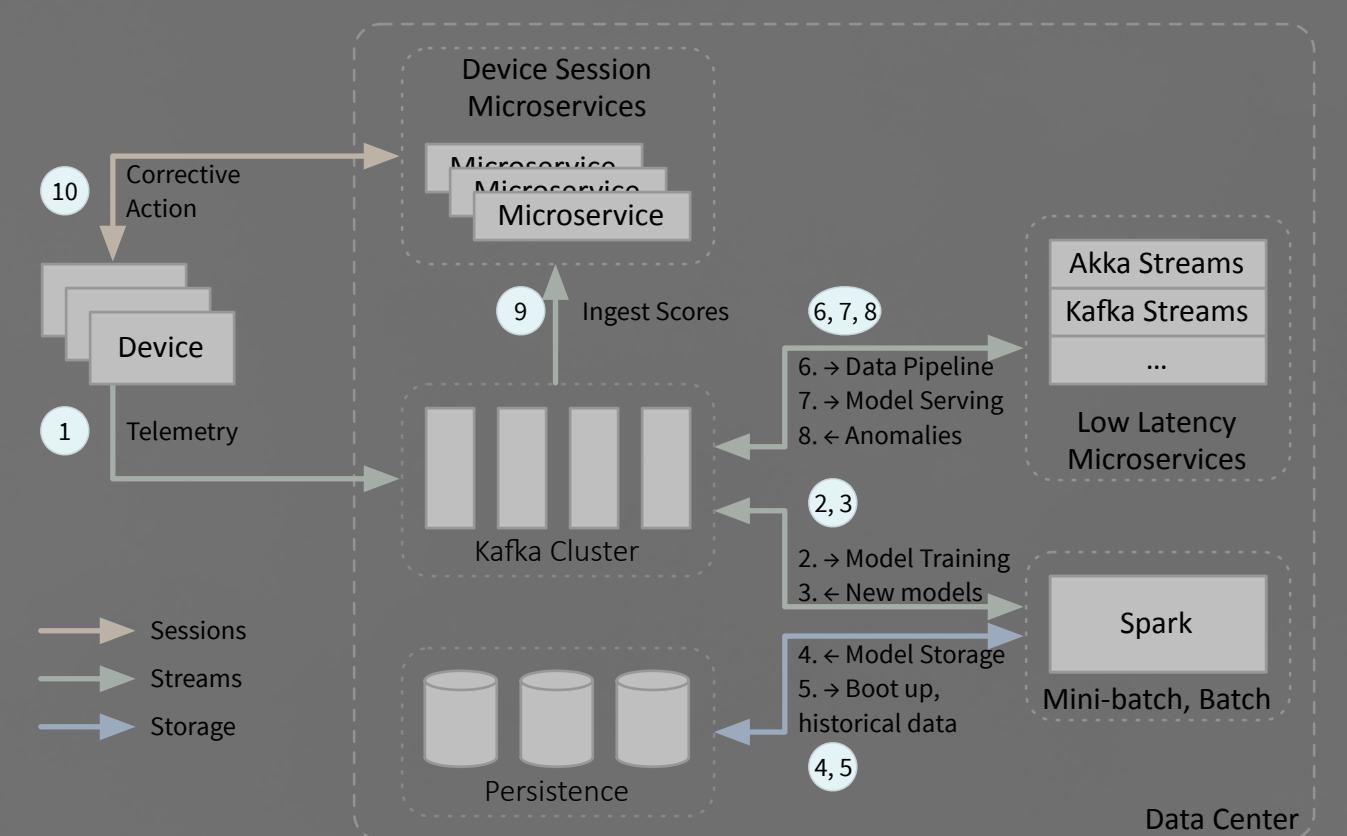
Example Architecture

• Integration of Machine Learning/Artificial Intelligence with streaming is a common challenge right now

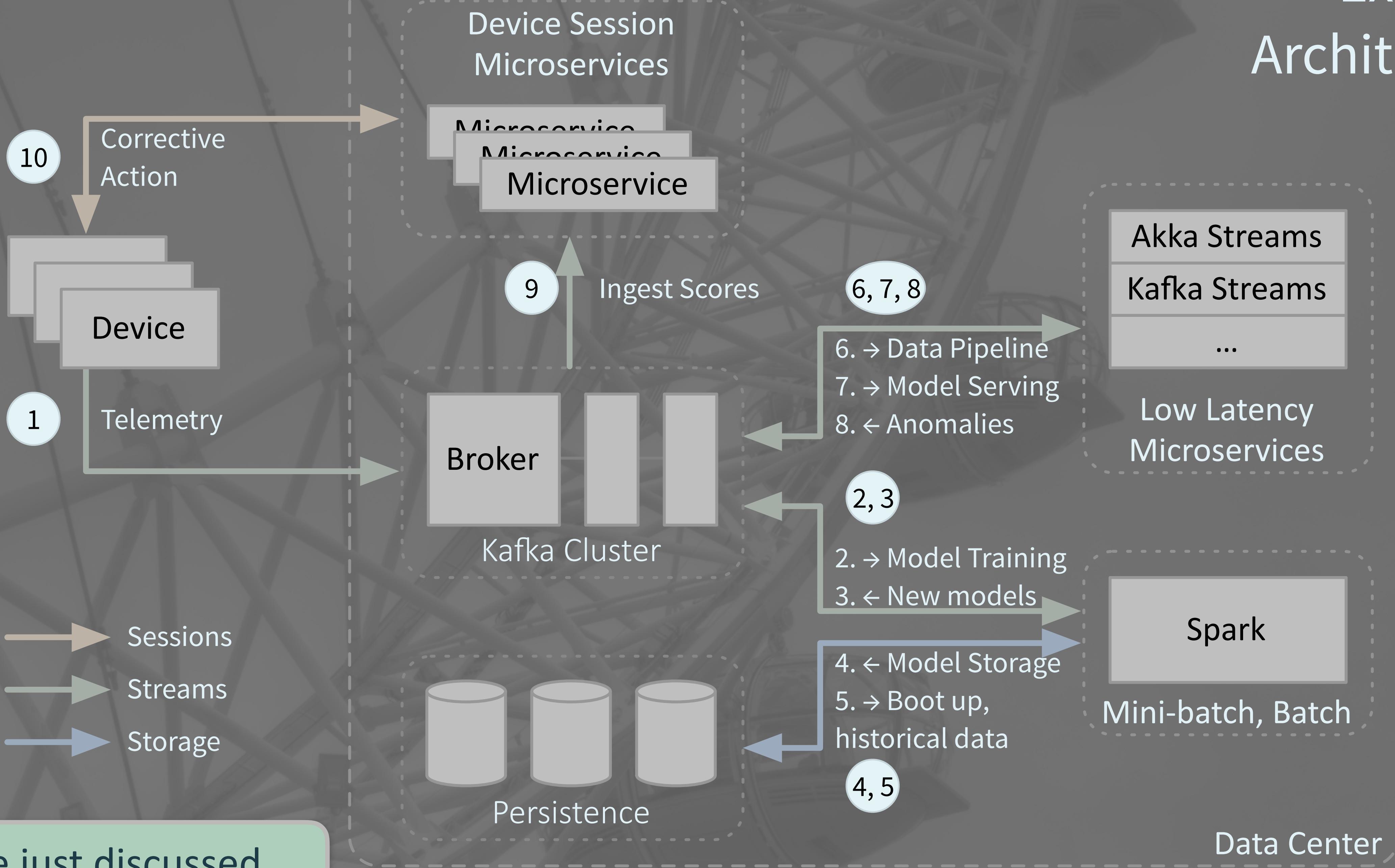


Challenges

- Network overhead for telemetry ingestion too high?
- Model serving latency too long?
- Datacenter unavailable?
- Idea: Serve models on the device!

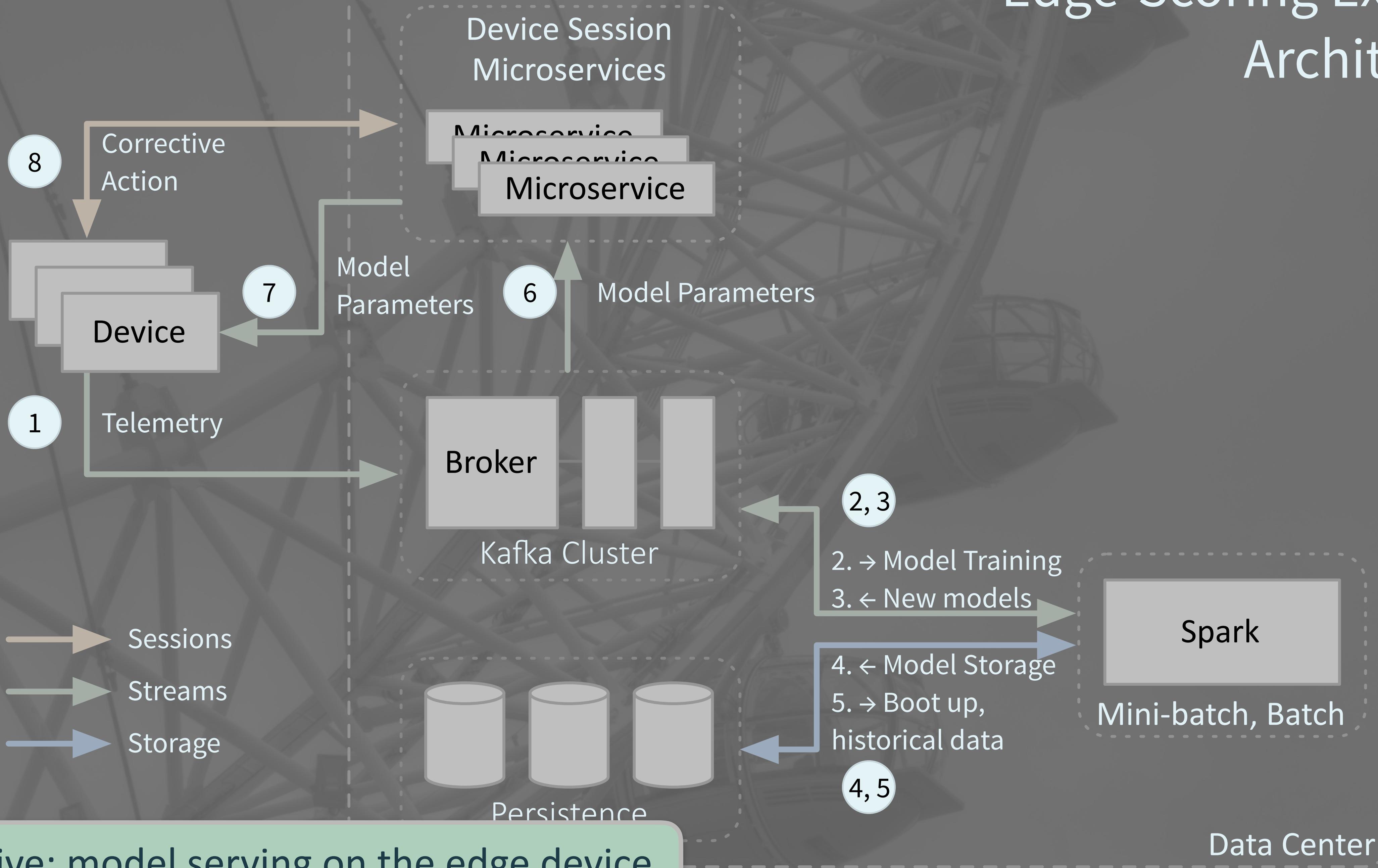


Example Architecture

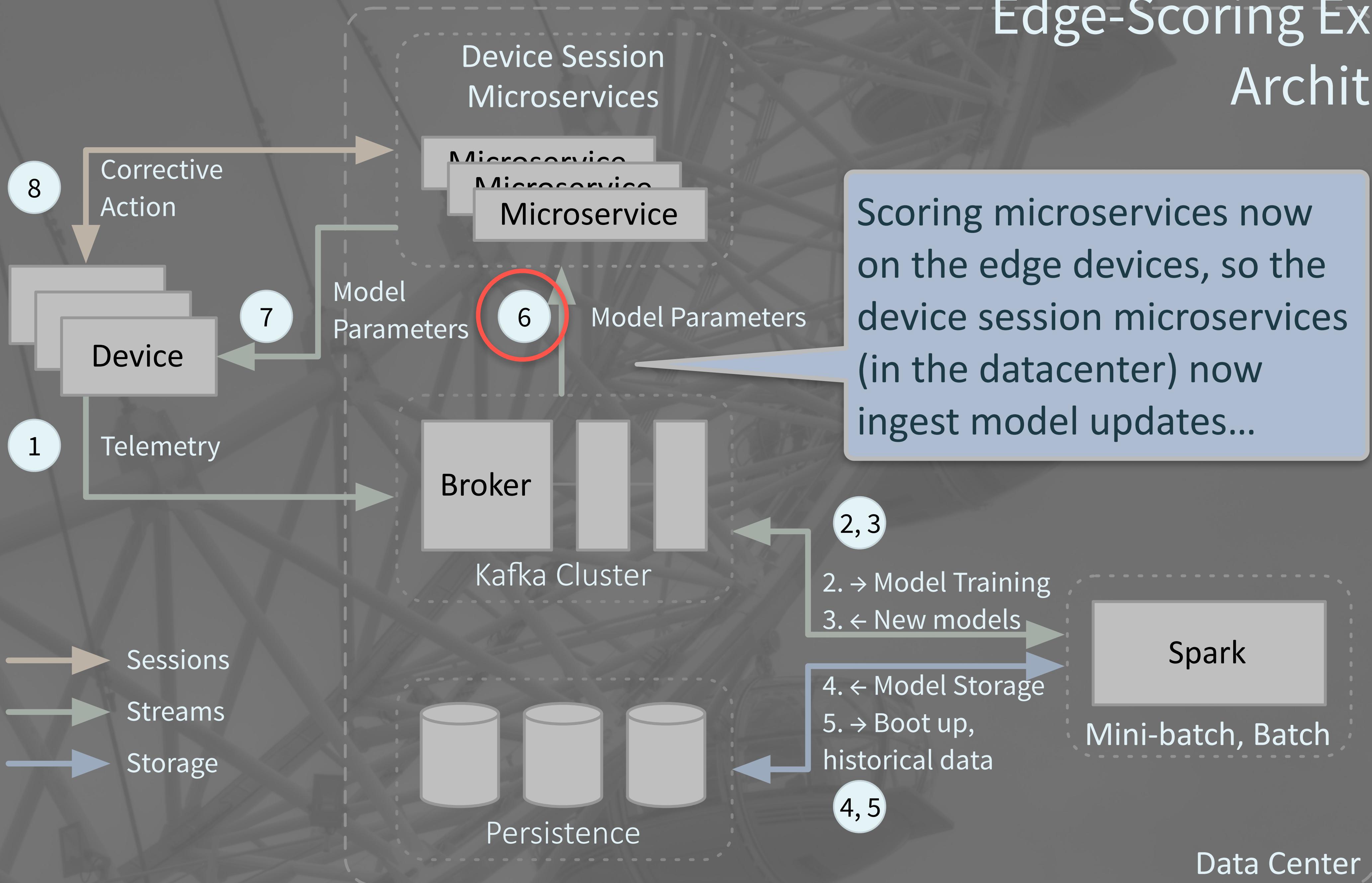


What we just discussed...

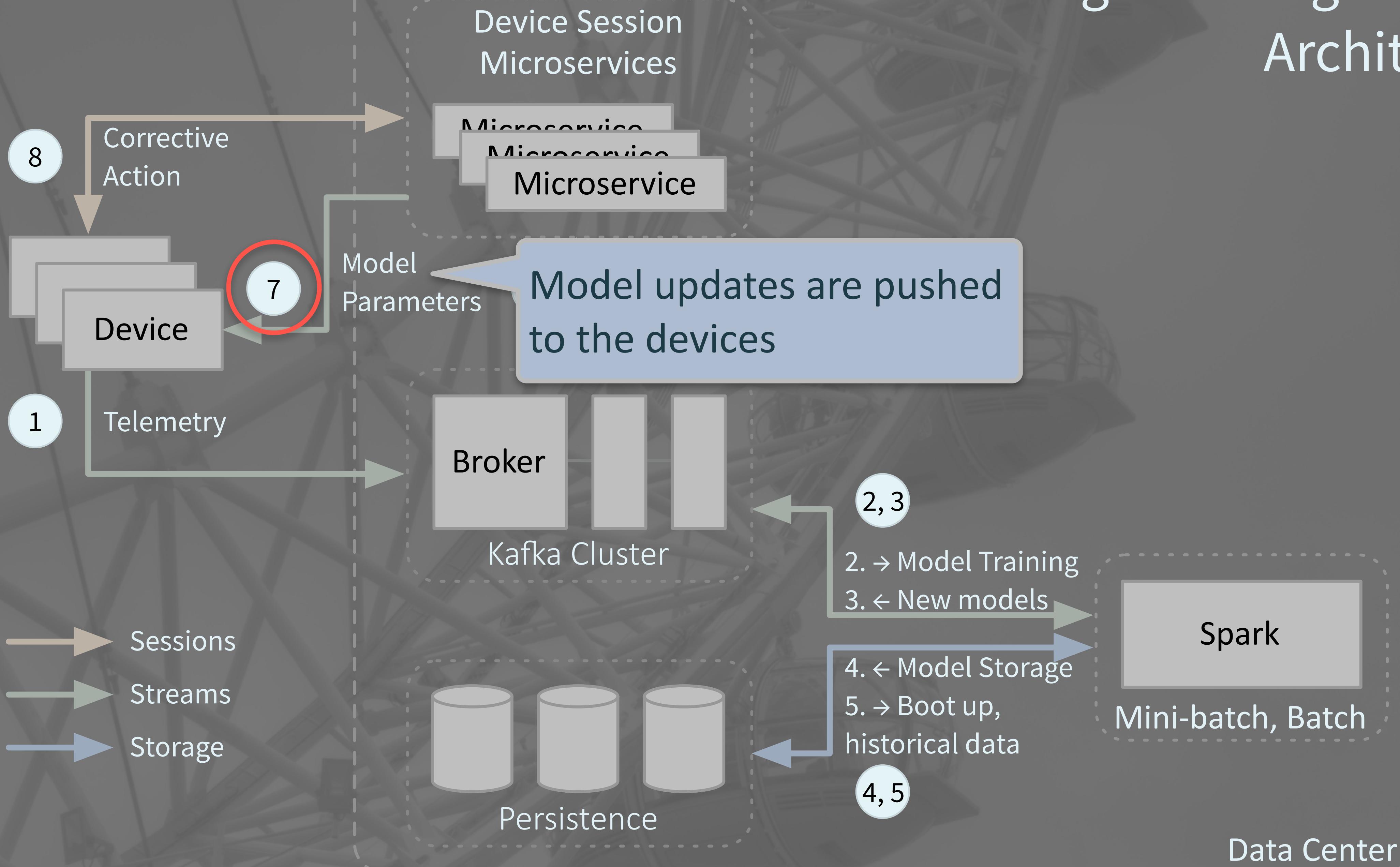
Edge-Scoring Example Architecture



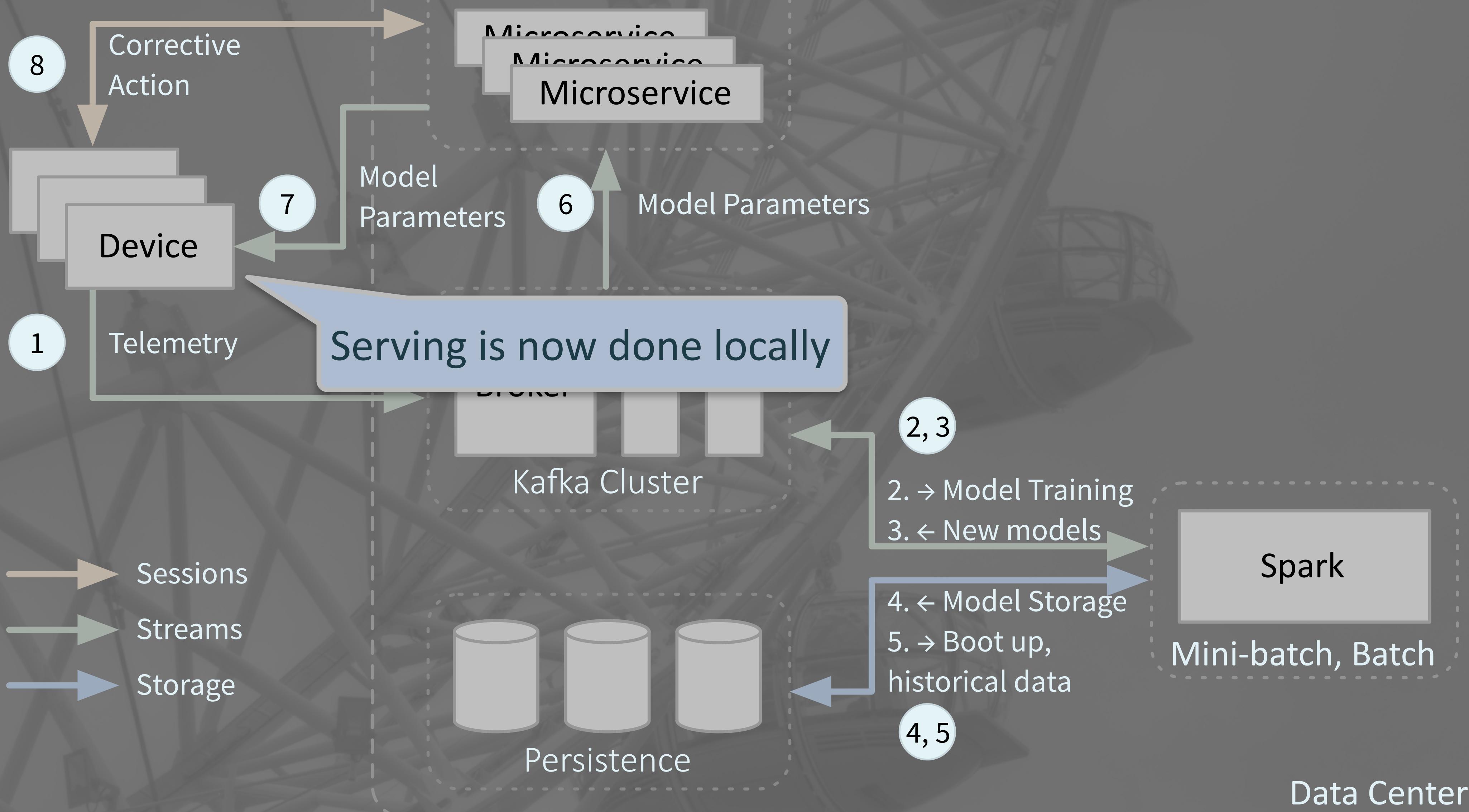
Edge-Scoring Example Architecture



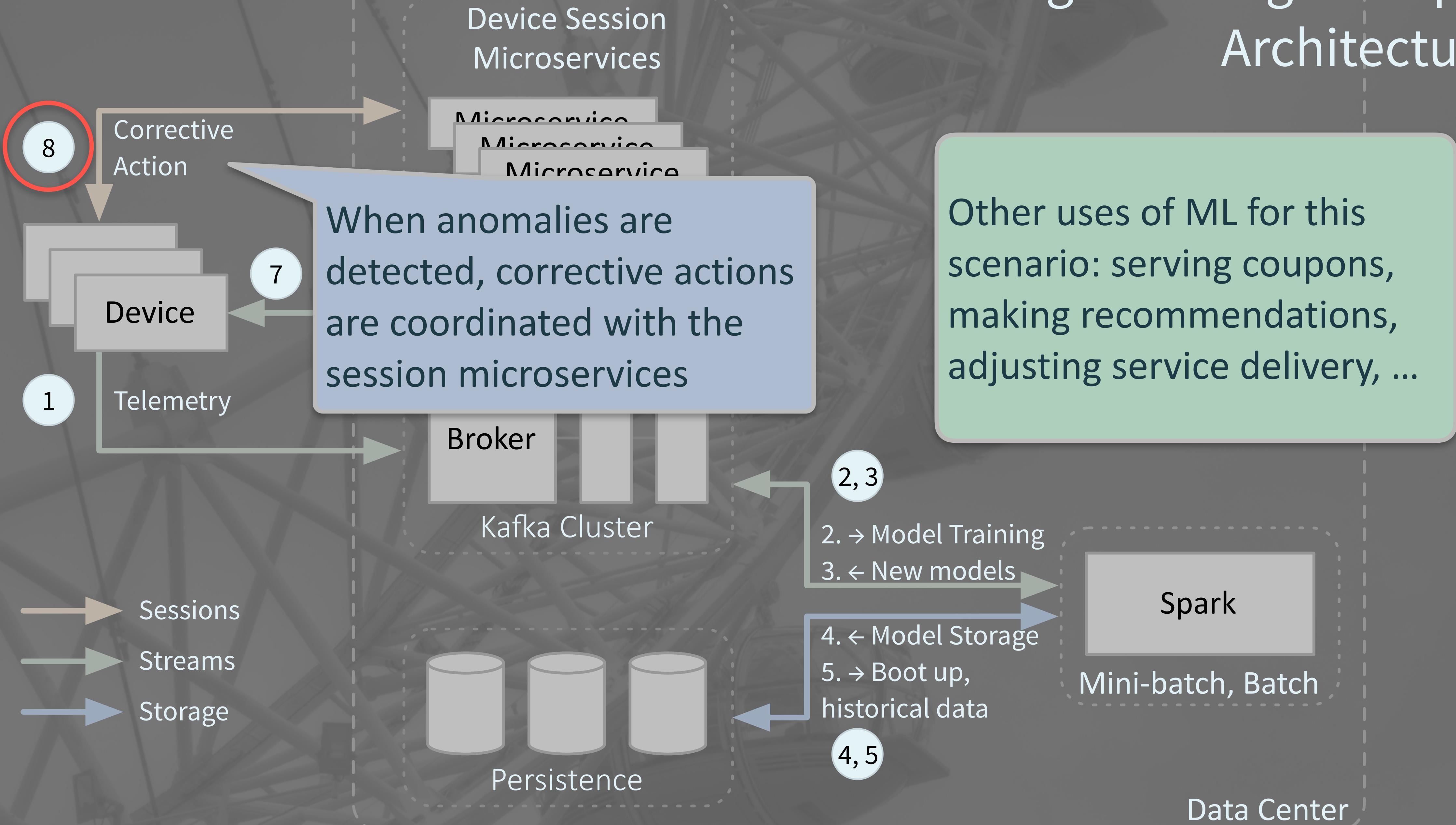
Edge-Scoring Example Architecture



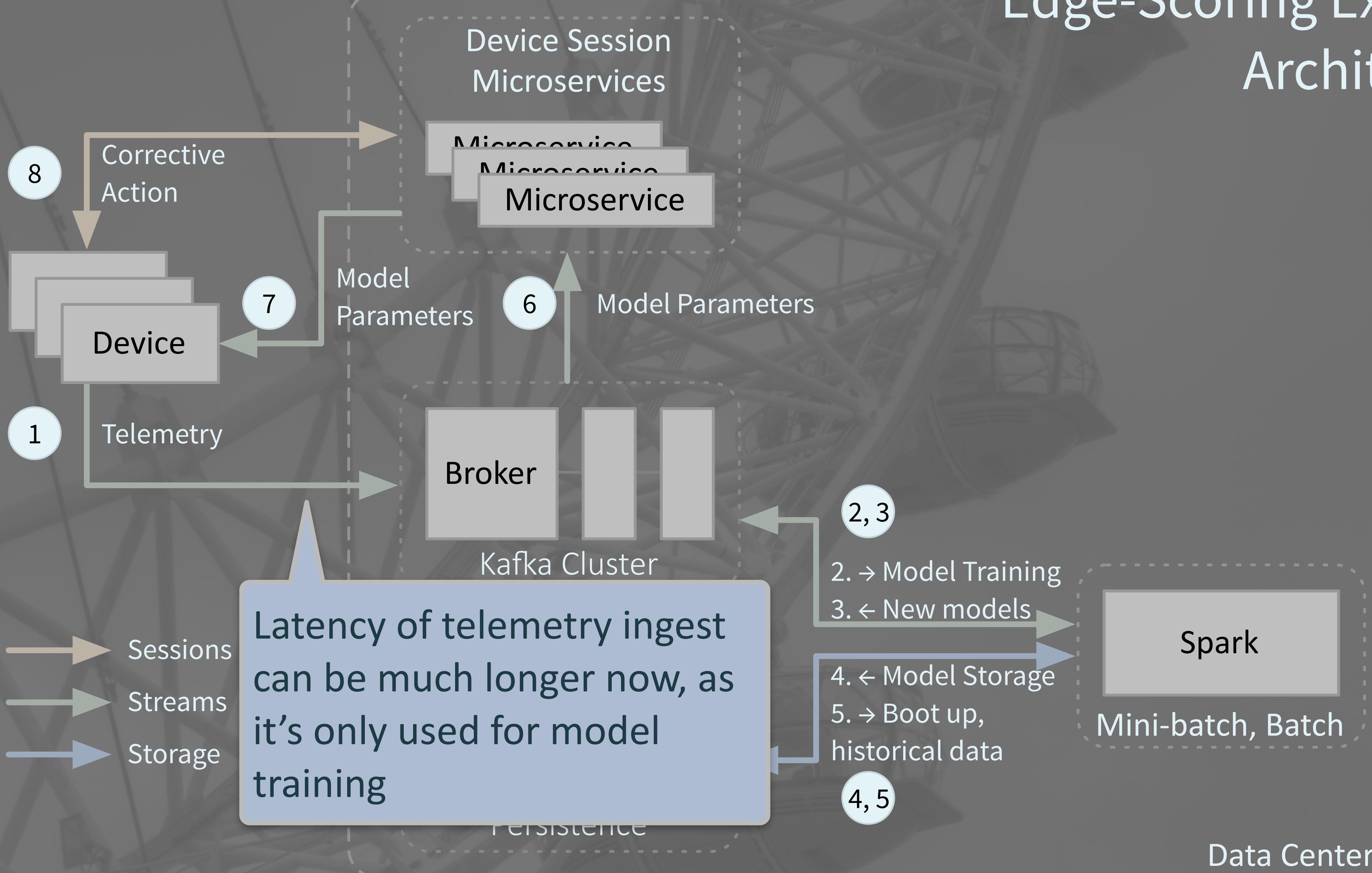
Edge-Scoring Example Architecture



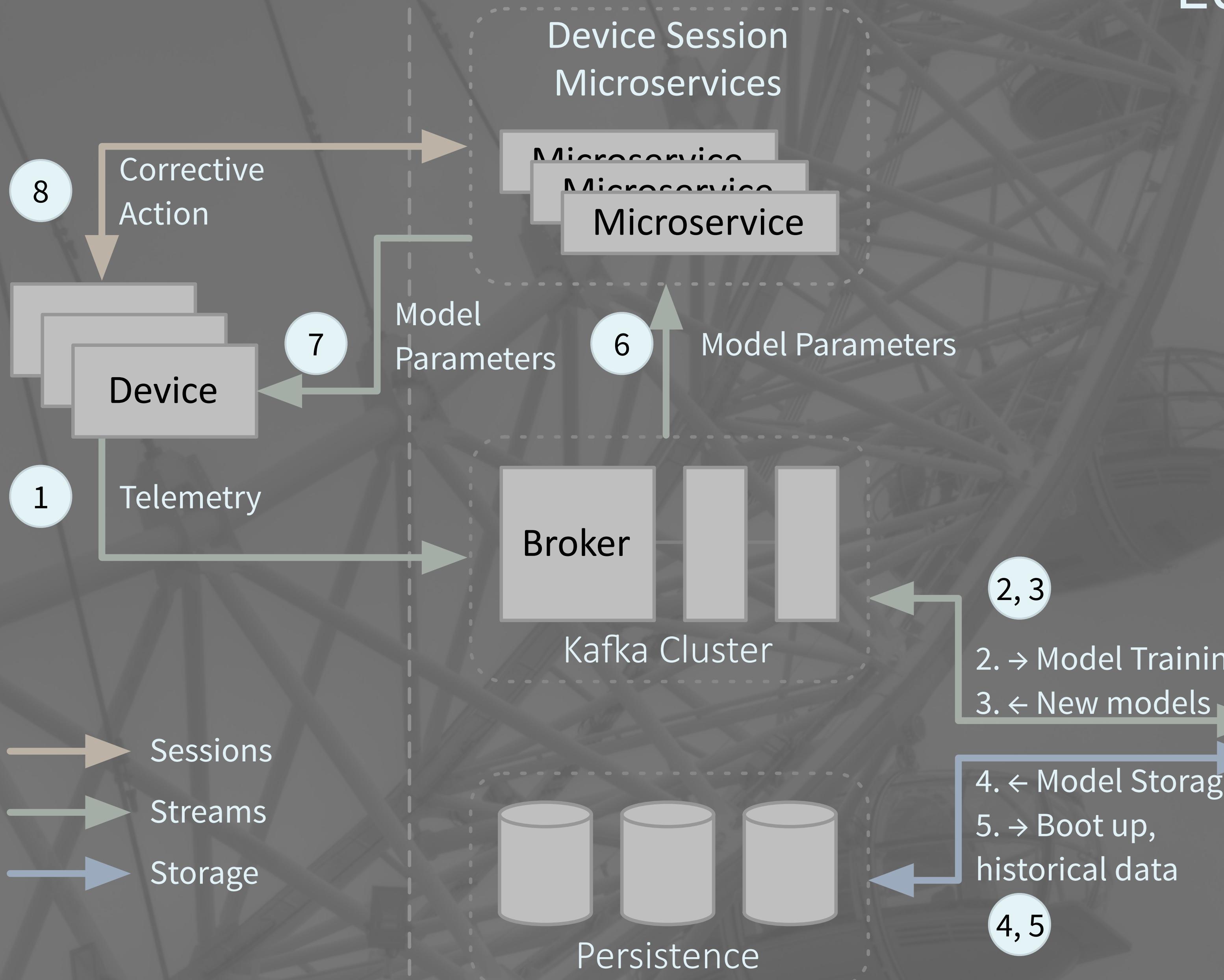
Edge-Scoring Example Architecture



Edge-Scoring Example Architecture



Edge-Scoring Example Architecture



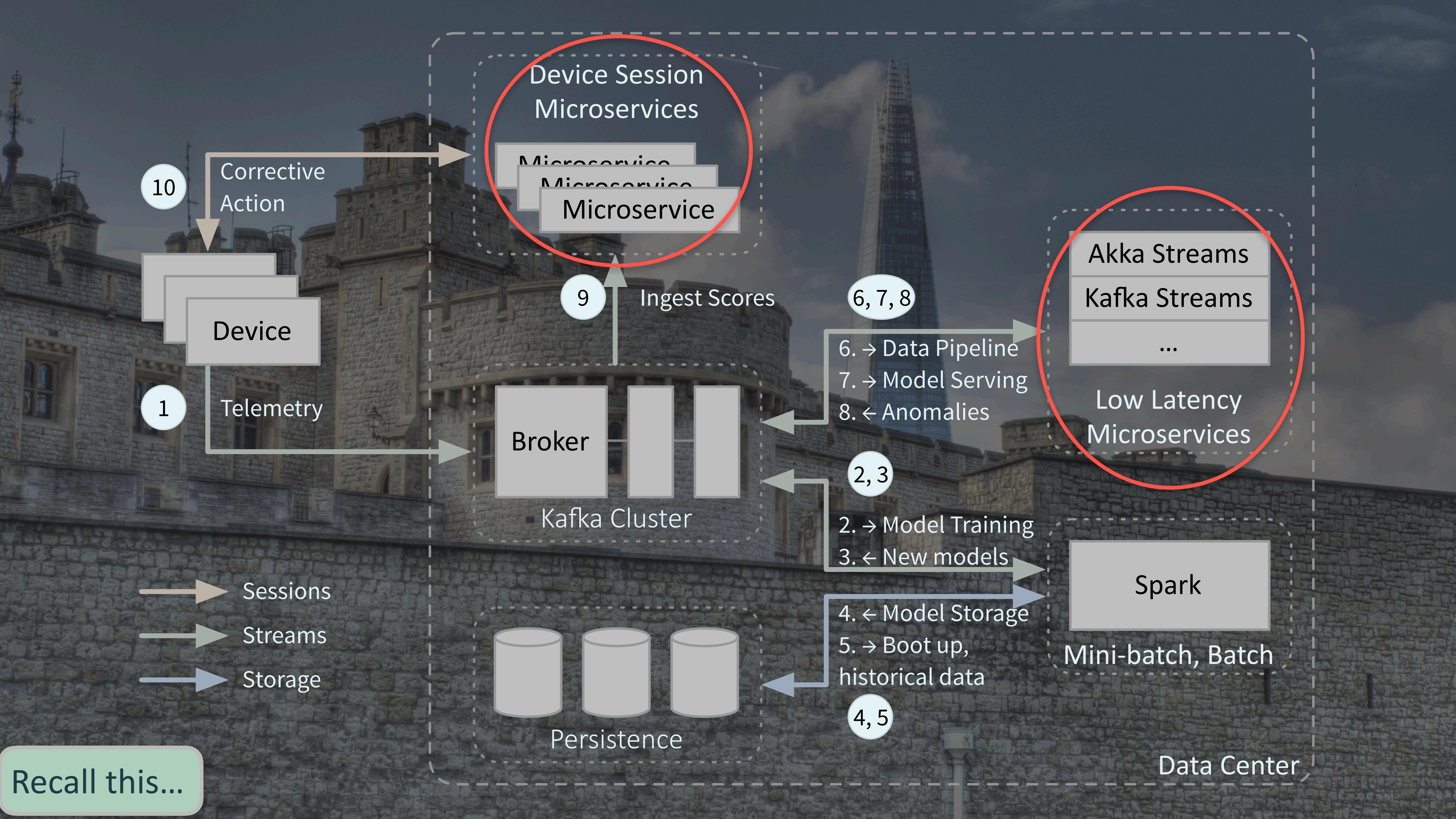
Recap: Edge Serving

More Challenges...



More Challenges...

How do you keep all this healthy for months and months??



Why Microservices in Fast Data?

1. The trend is to run everything in big clusters using Kubernetes or Mesos
 - In the cloud or on-premise

Why Microservices in Fast Data?

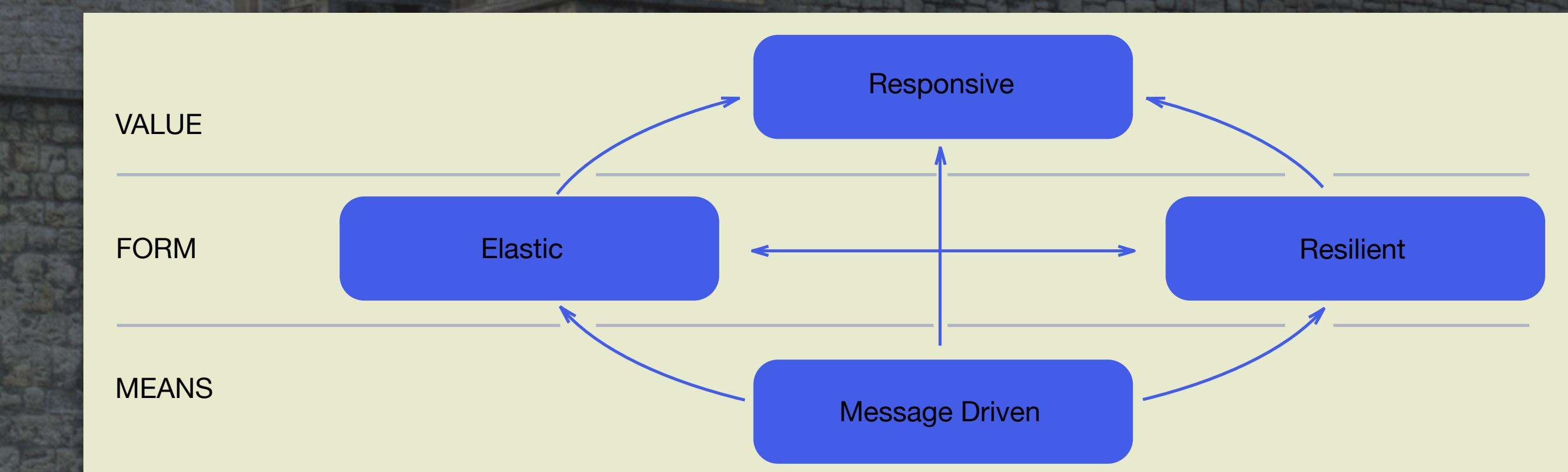
2. If streaming gives you information faster...

- ... you'll want quick access to it in your other services!

Why Microservices in Fast Data?

3. Streaming raises the bar on data services

- Compared to batch services, long-running streaming services must be more:
- Scalable
- Resilient
- Flexible



Why Microservices in Fast Data?

4. This leads to our last major point...



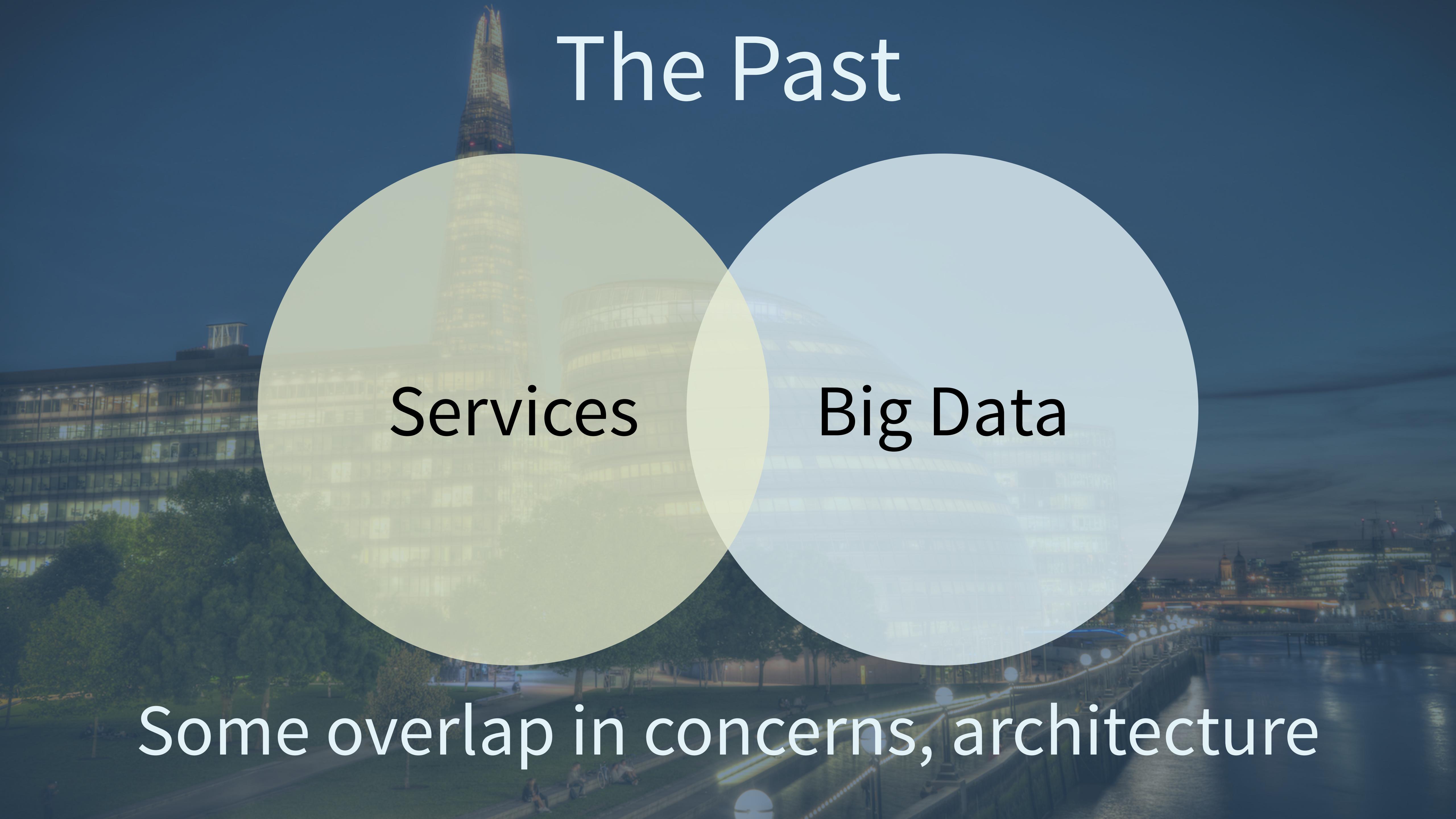
Organizational Impact



Organizational Impact

- Data scientists have to understand production issues
- Data engineers have to become good at highly-available microservices
- Microservice engineers have to become good at data

The Past

A Venn diagram with two overlapping circles is overlaid on a photograph of a city skyline at dusk. The left circle is yellow and contains the word 'Services'. The right circle is light blue and contains the words 'Big Data'. The background shows the Shard skyscraper, the London Eye, and other buildings along the River Thames.

Services

Big Data

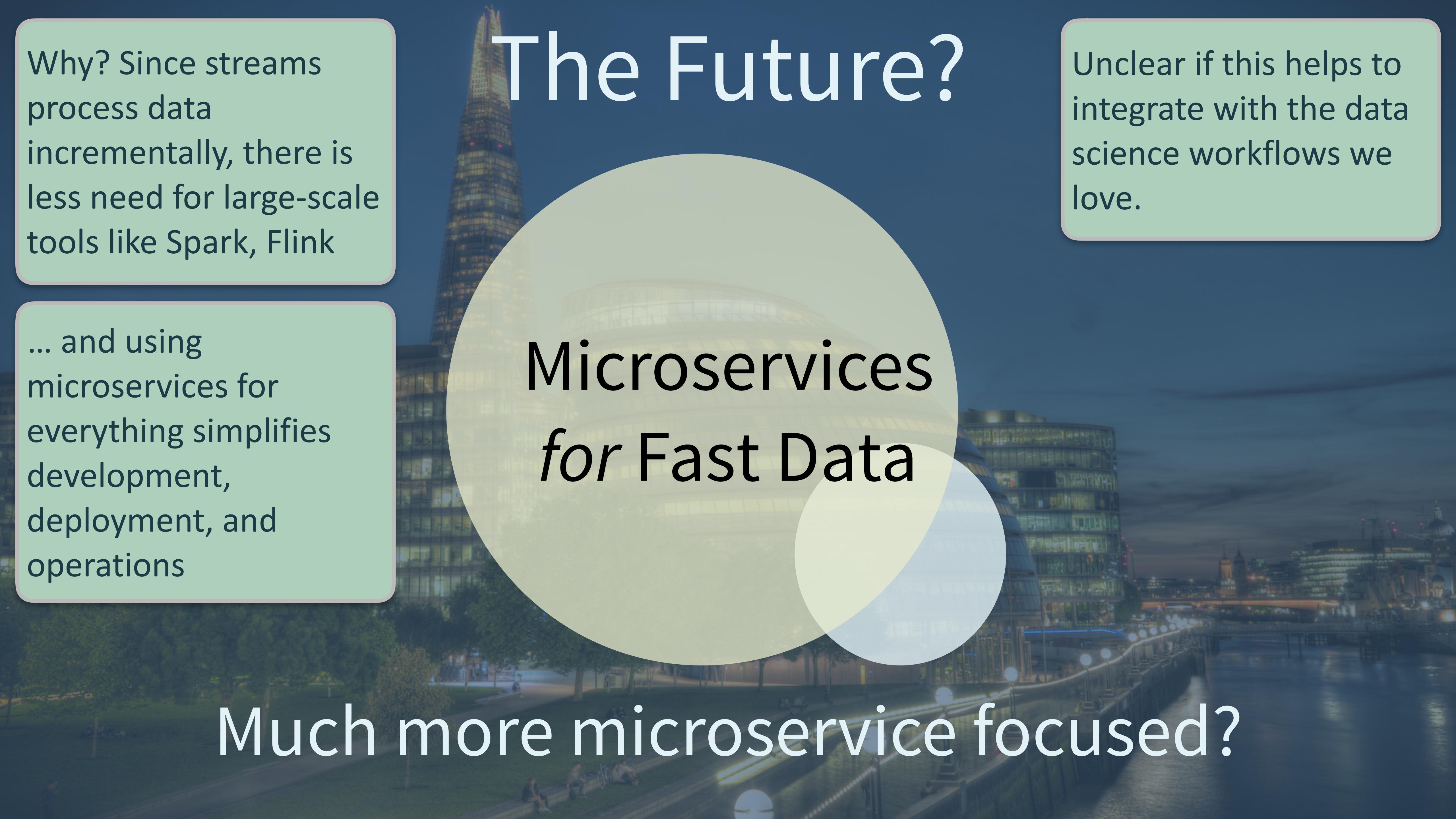
Some overlap in concerns, architecture

The Present



Microservices
& Fast Data

Much more overlap



Why? Since streams process data incrementally, there is less need for large-scale tools like Spark, Flink

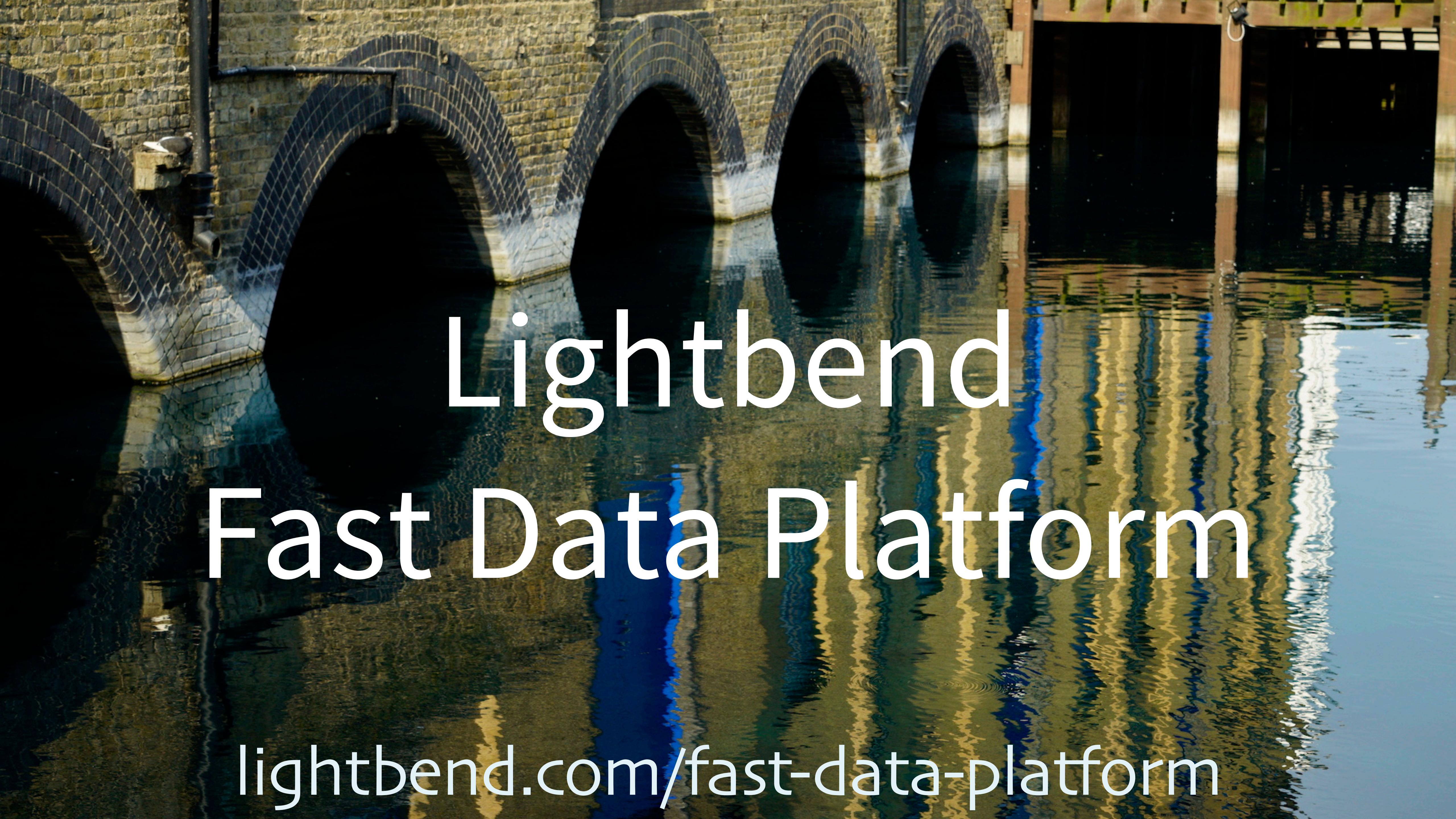
... and using microservices for everything simplifies development, deployment, and operations

The Future?

Microservices for Fast Data

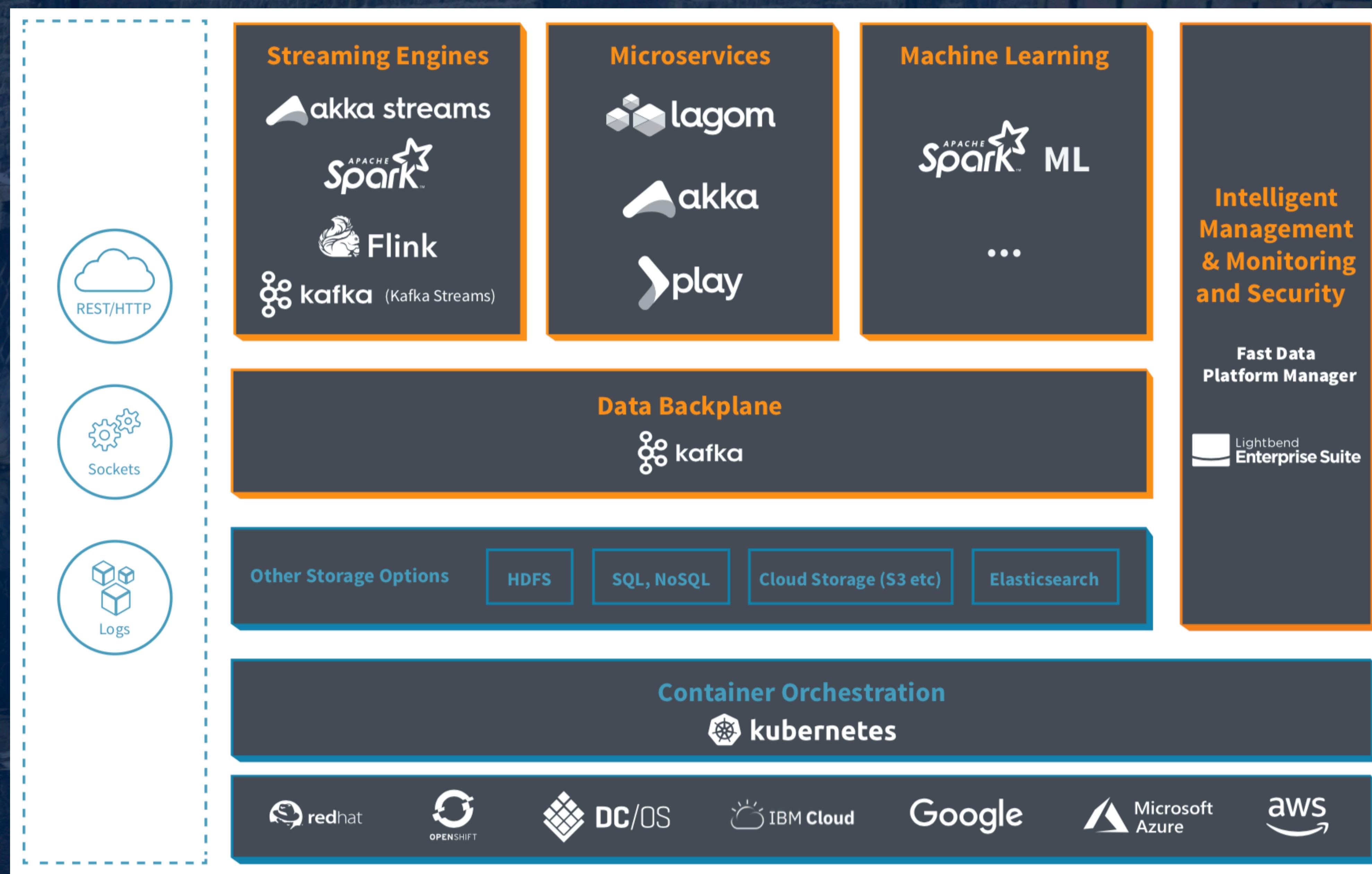
Much more microservice focused?

Unclear if this helps to integrate with the data science workflows we love.

A photograph of a multi-arched brick bridge reflected perfectly in the still water below. The bridge's structure is composed of several large, semi-circular arches made of dark, weathered brick. The reflection in the water is sharp and clear, mirroring the bridge's form and the surrounding environment. The overall scene is peaceful and symmetrical.

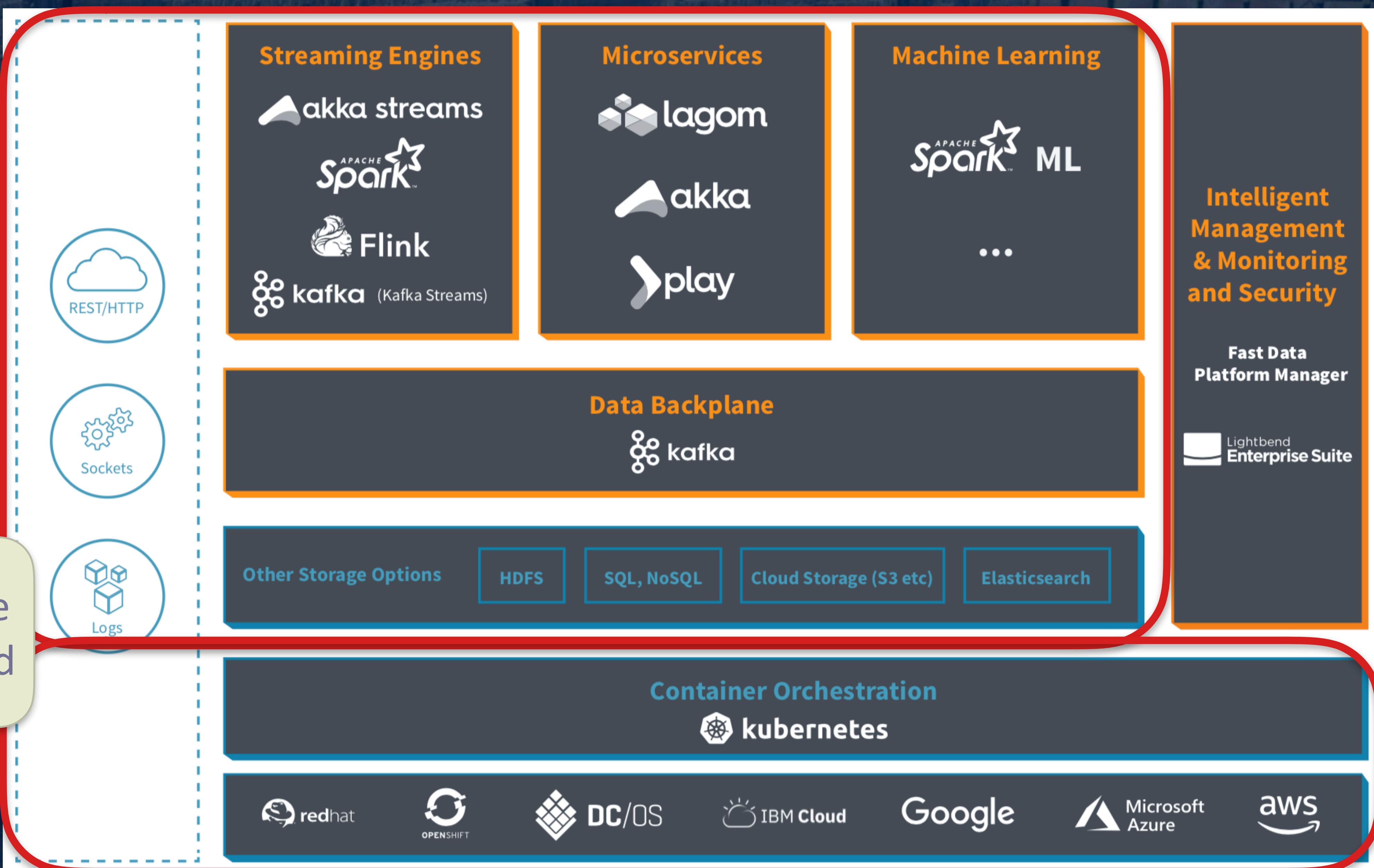
Lightbend Fast Data Platform

lightbend.com/fast-data-platform

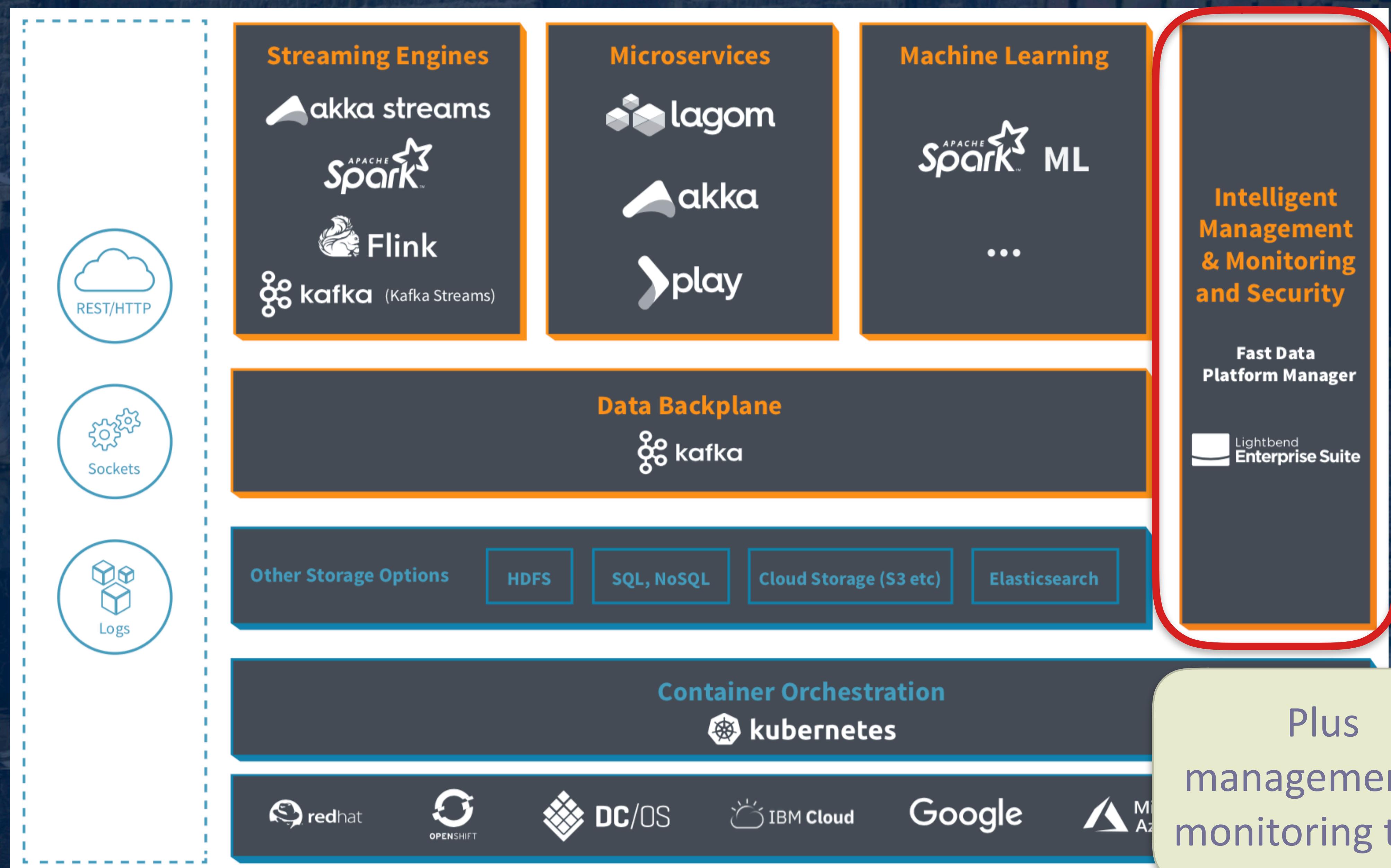


lightbend.com/fast-data-platform

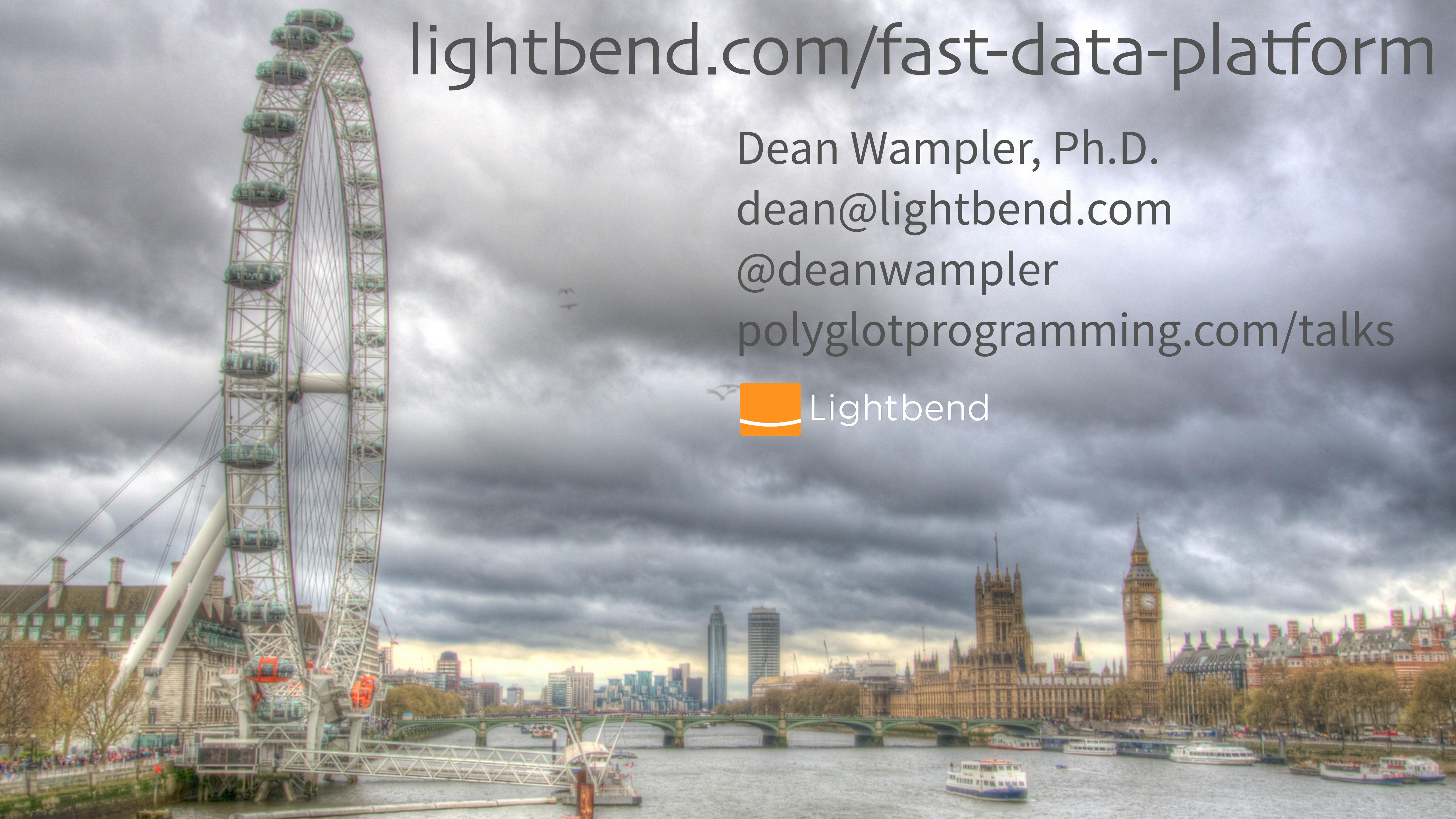
What we
discussed



lightbend.com/fast-data-platform



lightbend.com/fast-data-platform

A photograph of the London skyline featuring the London Eye (Millennium Wheel) on the left and the Palace of Westminster with Big Ben on the right. The sky is overcast with dramatic clouds. In the foreground, the River Thames is visible with several boats, and a green bridge spans the river.

lightbend.com/fast-data-platform

Dean Wampler, Ph.D.
dean@lightbend.com
[@deanwampler](https://twitter.com/deanwampler)
polyglotprogramming.com/talks

