

Build a serverless data streaming workload with Amazon Kinesis Services

Nihar Sheth (he/him)

Senior Product Manager
Amazon Kinesis
AWS

Pratik Patel (he/him)

Senior Technical Account Manager
AWS



Agenda

What is real-time analytics?

Amazon Kinesis Services

Serverless Streaming Architecture



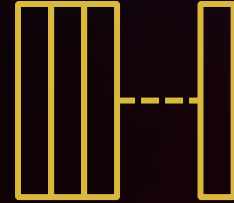
What is Streaming Data?



High volume



Continuous



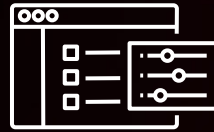
**Ordered,
incremental**



Low-latency



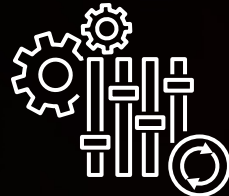
Mobile apps



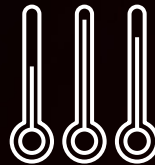
Web clickstream



Application logs



Metering records

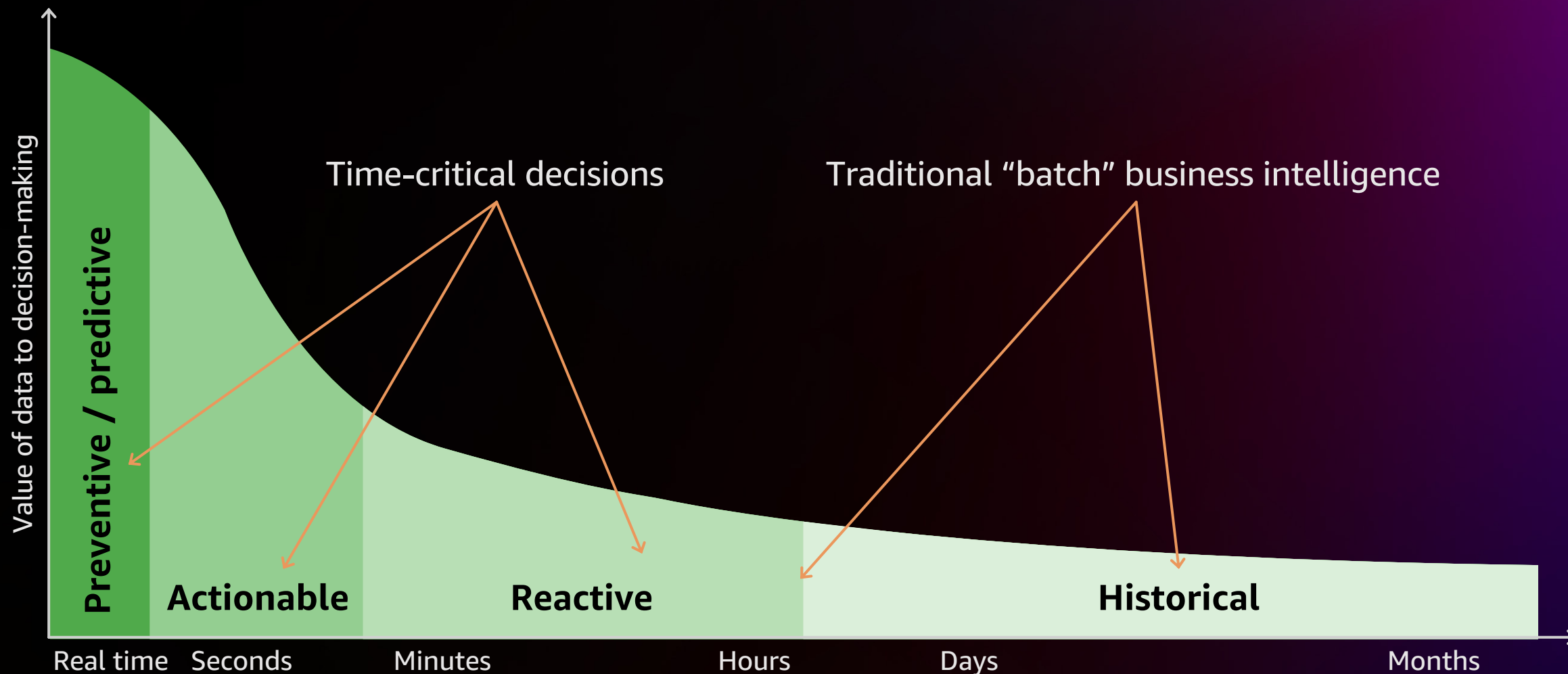


IoT sensors



Smart buildings

Why real-time analytics?



Source: Mike Gaultieri, Forrester, *Perishable Insights*



© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

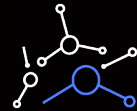
Common real-time analytics use-cases



Anomaly and fraud detection



Log analytics



Empowering IoT analytics



Nourishing marketing campaigns



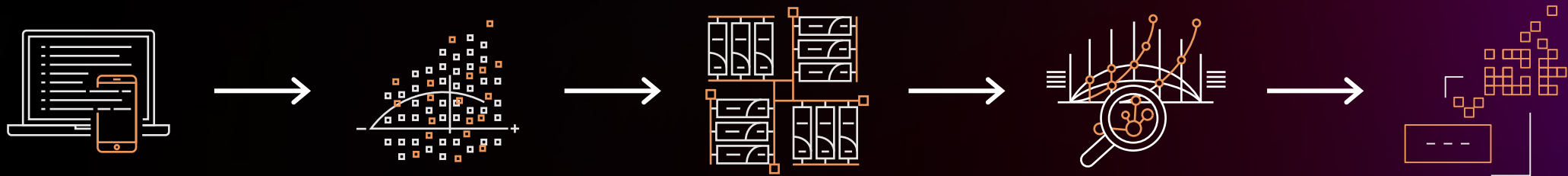
Real-time personalization



Supporting healthcare and emergency services

Data Streaming Workflow

INGEST, PROCESS, AND ANALYZE HIGH VOLUMES OF HIGH-VELOCITY DATA FROM A VARIETY OF SOURCES IN REAL TIME



Source

Devices and/or applications that produce real-time data at high velocity

Stream ingestion

Data from tens of thousands of data sources can be collected and ingested in real time

Stream storage

Data is stored in the order it was received for a set duration of time, and can be replayed indefinitely during that time

Stream processing

Records are read in the order they're produced, allowing for real-time analytics or streaming ETL

Destination

Data lake
Data warehouse (most common)
Database (least common)

Challenges of self managed data streaming



Apache
Kafka



Apache
Flink



Difficult to set up



Tricky to scale



Hard to achieve
high availability



Integration requires
development

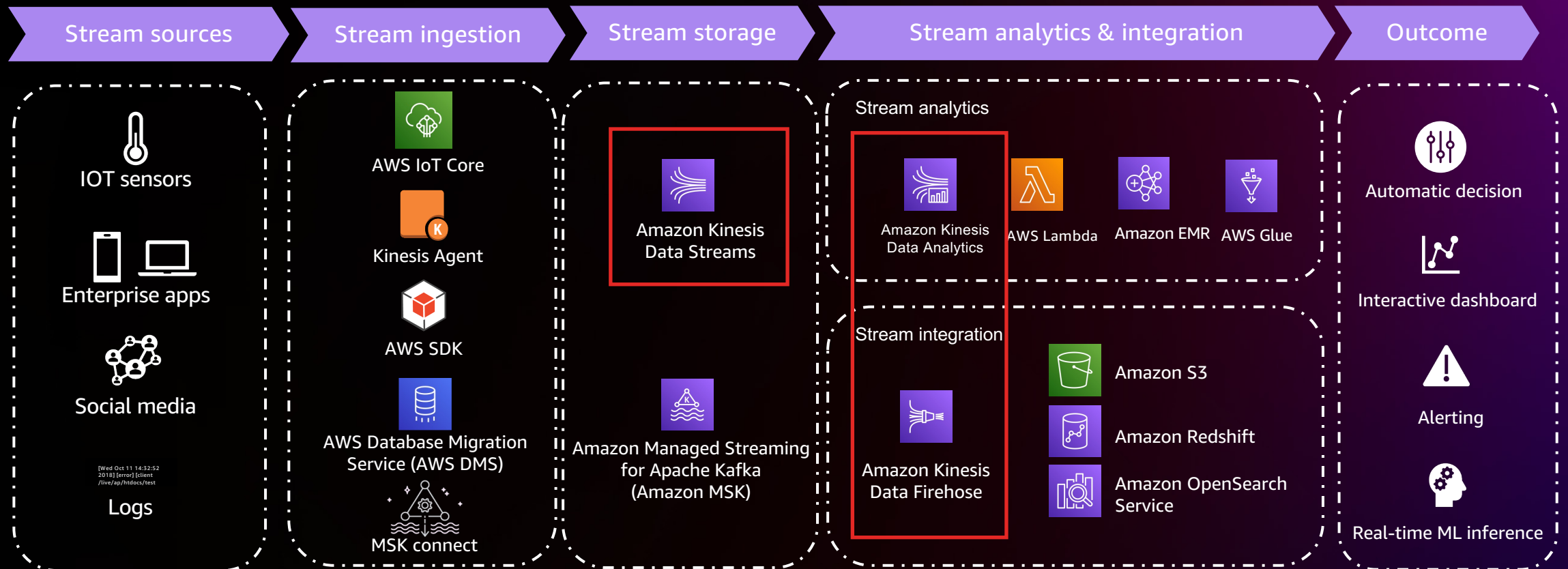


Error prone and
complex to manage

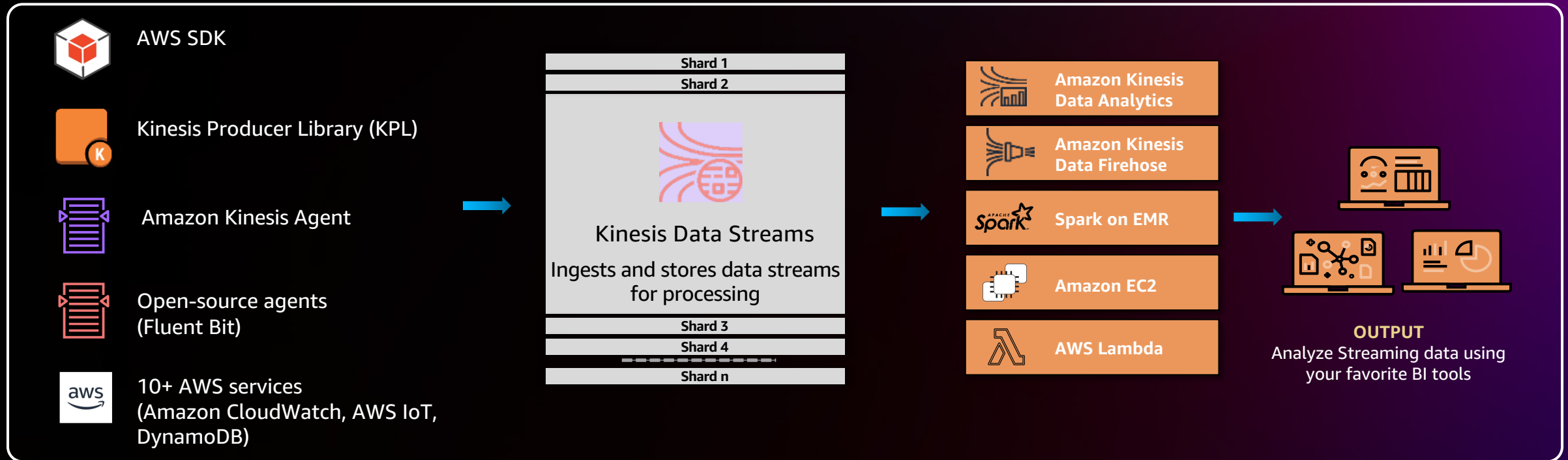
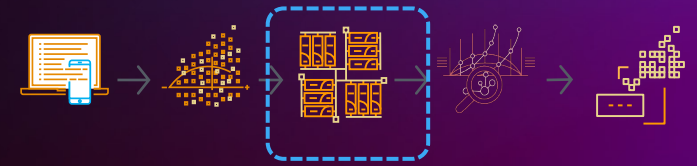


Expensive
to maintain

AWS services for streaming workloads



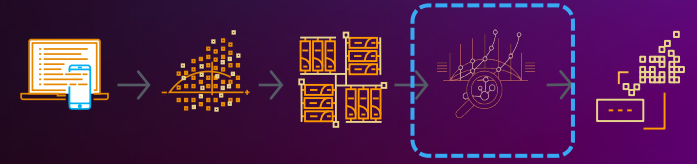
Amazon Kinesis Data Streams



- Easy administration and low cost
- Serverless Scaling
- Security, durability and availability out of the box

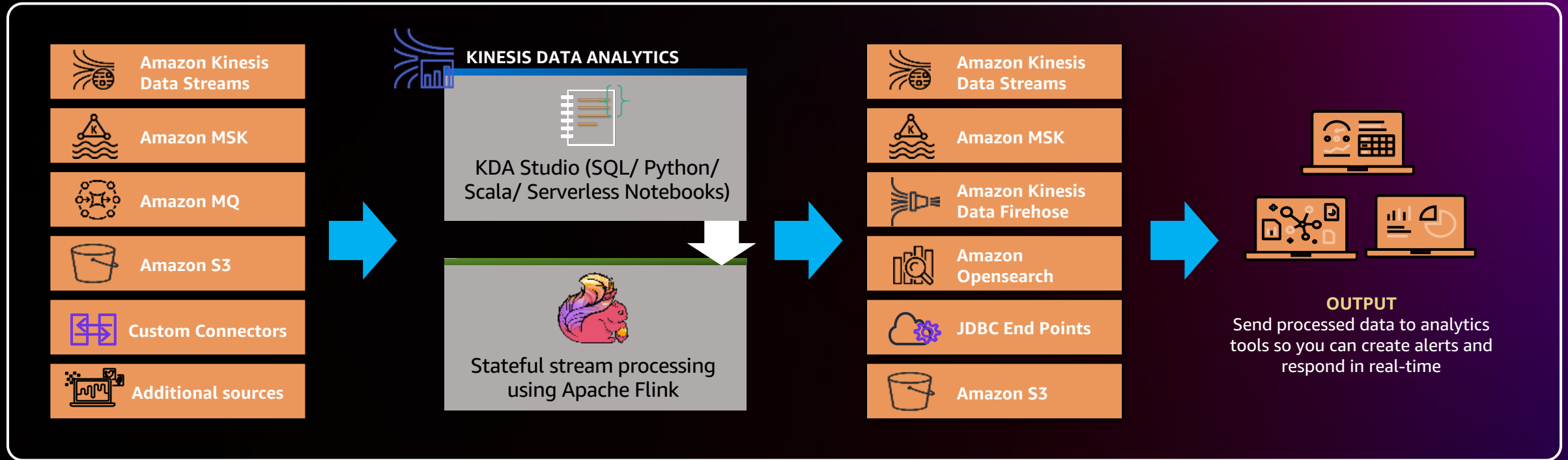
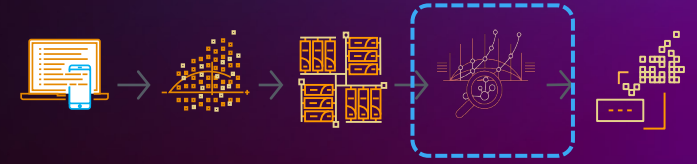
- Performance at Scale
- Concurrent consumers at low latency
- Data retention up to 1 year

Amazon Kinesis Data Firehose



- Zero administration and seamless scaling
- Direct-to-data store integration
- Serverless data transformations
- Buffer and batching flexibility
- Data format conversion to Parquet/ ORC
- Dynamic partitioning to S3
- Deliver data directly to 15+ destinations (Datadog, Sumo Logic, New Relic and MongoDB..)

Amazon Kinesis Data Analytics



- Interact with streaming data in real time using SQL, Python, Scala and Java or integrated Apache Flink applications
- Deploy KDA studio adhoc analysis as a durable state application with in KDA for Apache Flink
- Build fully managed and scalable stream processing applications

Customers References



Billions of events per day from TVs and connected devices



Migrated data bus from Self-Managed Apache Kafka to Kinesis

SONOS

1 billion events per day from connected devices



Near-real-time home valuation (Zestimates)



Live clickstream dashboards refreshed under 10s



IoT predictive analytics



10 TB/day clickstreams from 250+ sites



50 billion daily ad impressions, sub-50 ms responses

NORDSTROM

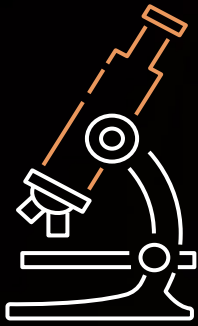
Online stylist processing 100 million events/day



Facilitate communications between 100+ microservices



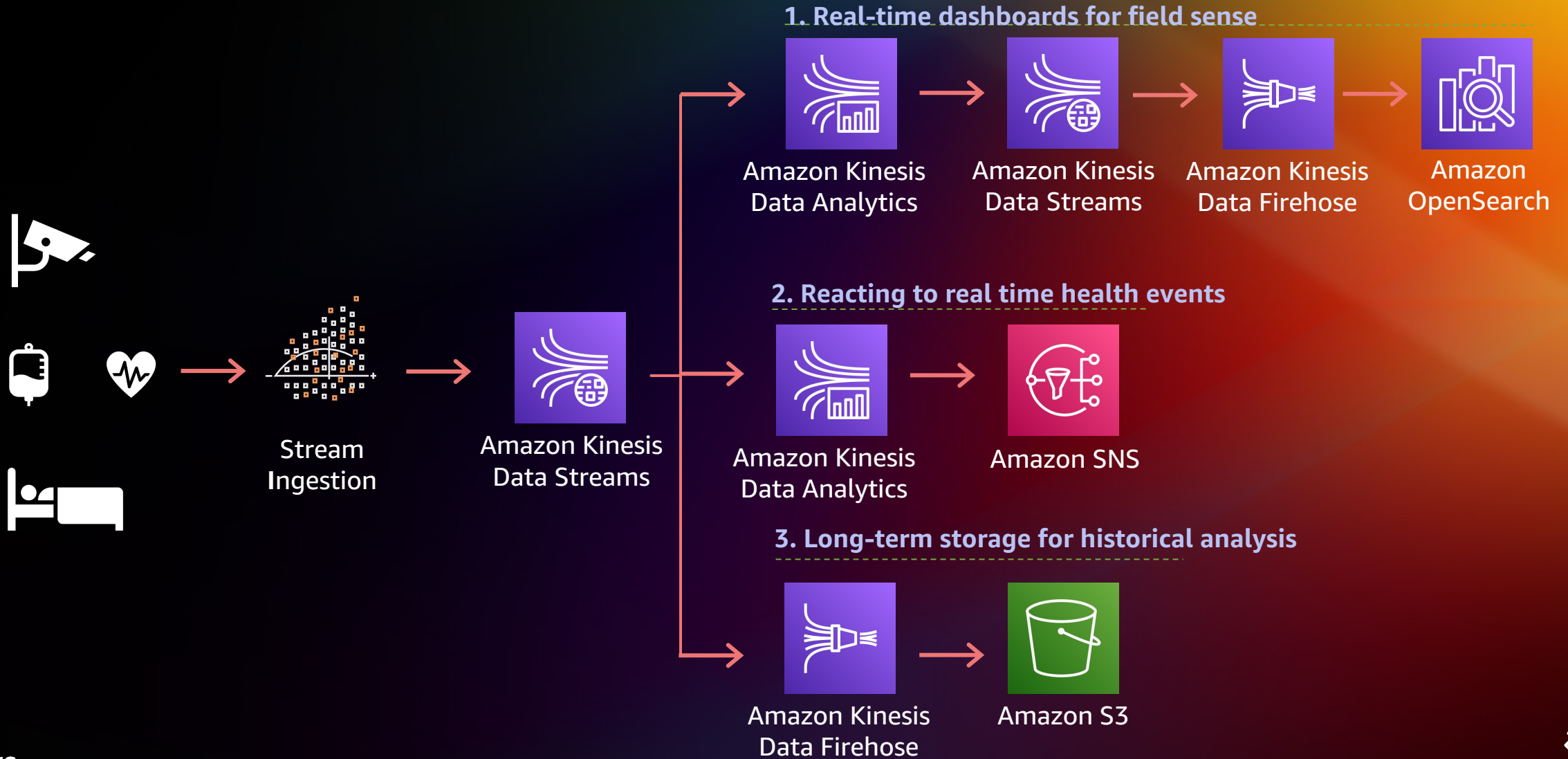
Customer: Medical device company



Supporting healthcare
and emergency services

1. Real-time dashboards for field sense
2. Reacting to real time health events
3. Long-term storage for historical analysis

Serverless Streaming Architecture



Getting started

Spin-up a Kinesis application within minutes

<https://aws.amazon.com/solutions/implementations/aws-streaming-data-solution-for-amazon-kinesis/>

Kinesis developer documentation

<https://aws.amazon.com/kinesis/getting-started/>

Kinesis Blogs

<https://aws.amazon.com/blogs/big-data/tag/amazon-kinesis/>

Thank you!

Nihar Sheth

niharsheth-aws (LinkedIn)

Pratik Patel

pratikpatel-aws (LinkedIn)

