AWS re:Invent

GPSCT402

Keeping the Pace with Data Ingestion

Paul Sears
Partner Solutions Architect
Amazon Web Services

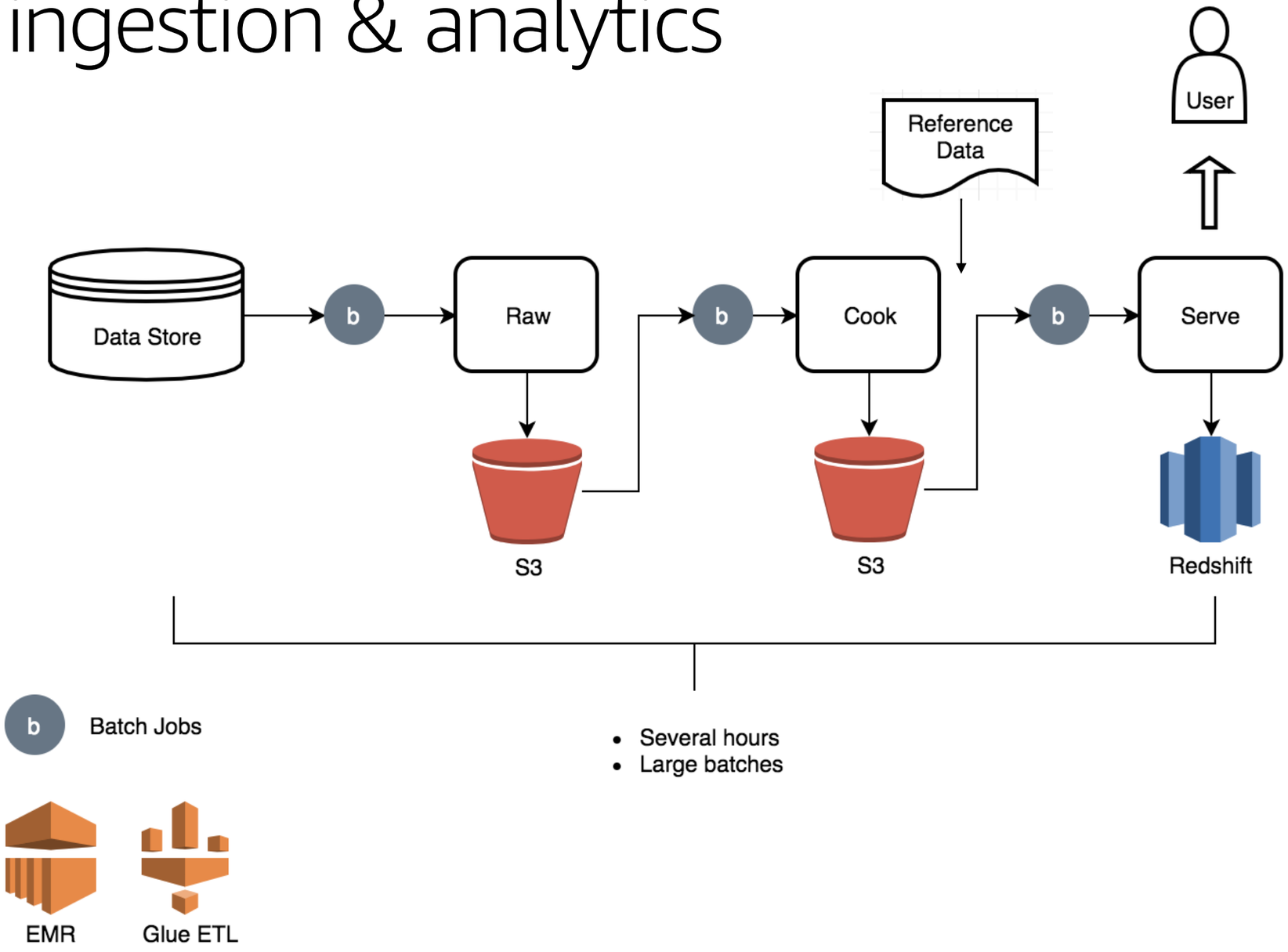
Rohit Pujari
Partner Solutions Architect
Amazon Web Services

AWS re:Invent

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

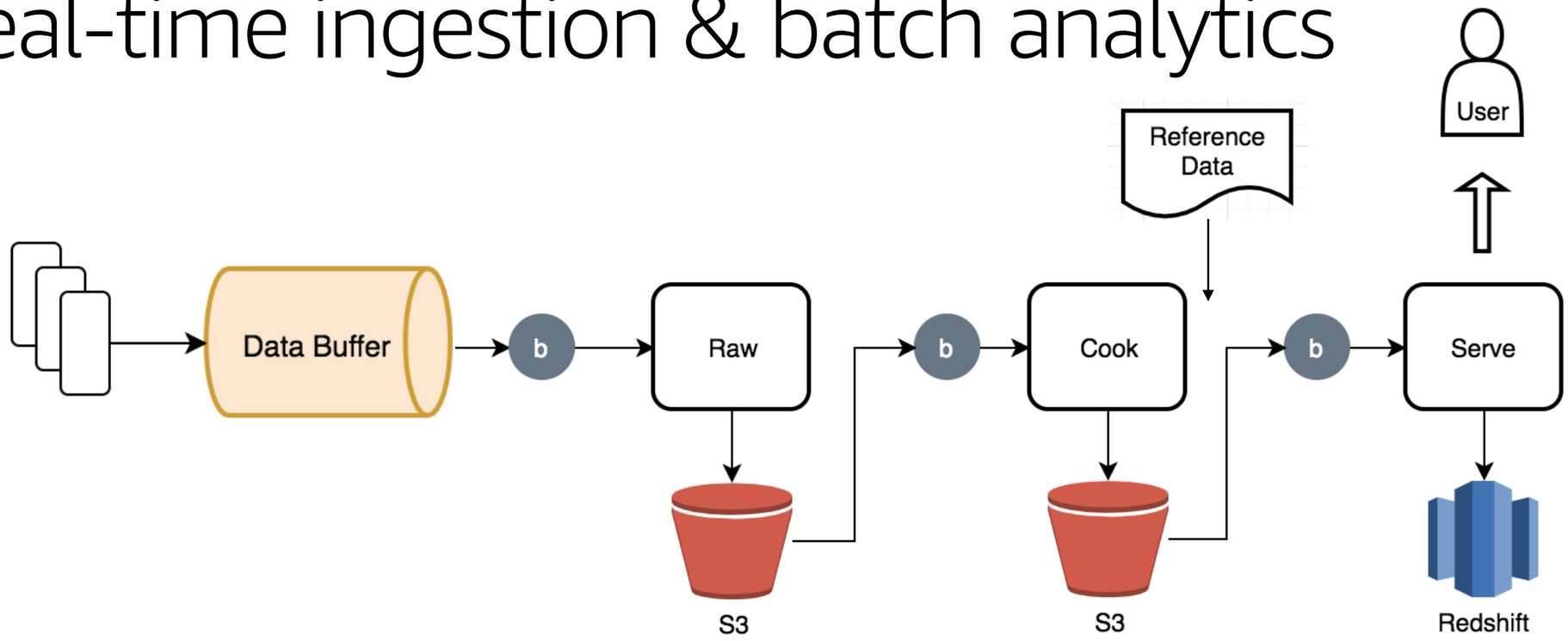
aws

Batch ingestion & analytics



© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Real-time ingestion & batch analytics



b Batch Jobs

- Several hours
- Large batches
  - Few hours
  - Mini batches

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

value = data

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

$$value = \frac{data}{time}$$

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Decision windows are shrinking

Hourly server logs → Real-time metrics

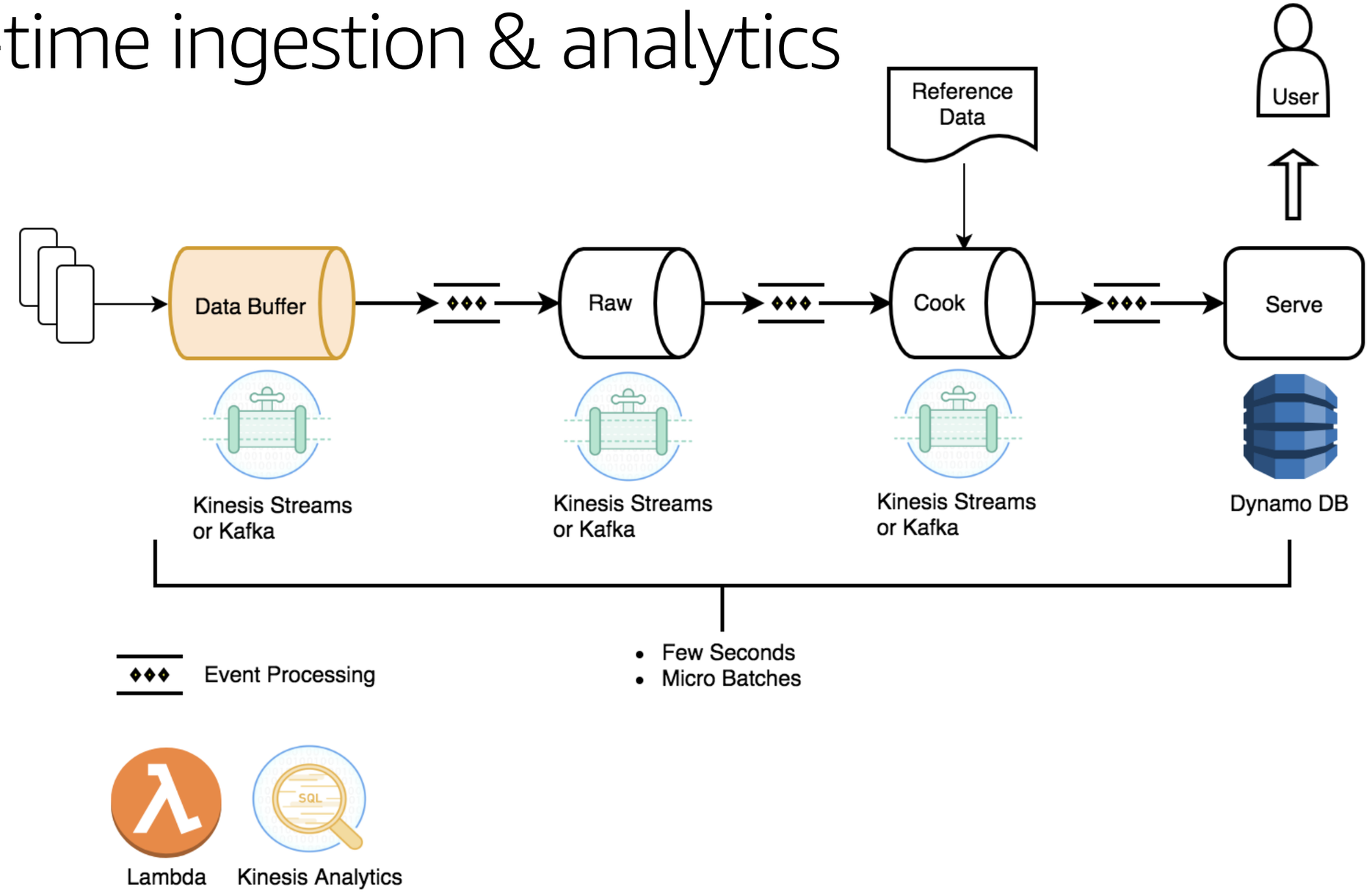
Weekly or monthly bills → Real-time spending alerts

Daily website clickstream → Real-time clickstream analysis

Daily fraud reports → Real-time detection

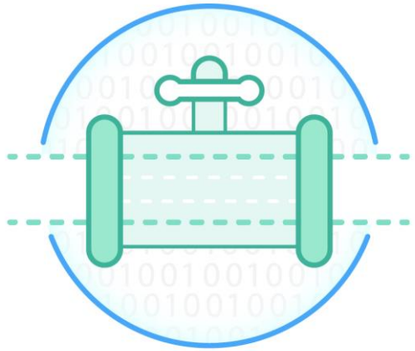
© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Real-time ingestion & analytics



© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Kinesis

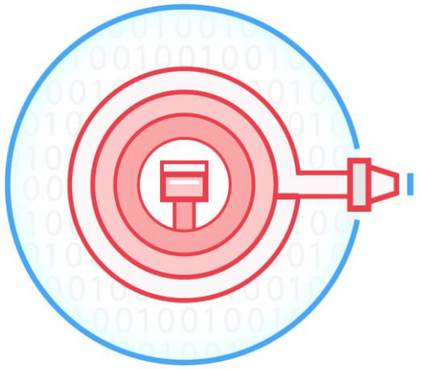


Amazon Kinesis Data Streams

Build custom applications that process and analyze streaming data

Amazon Kinesis Data Analytics

Easily process and analyze streaming data with standard SQL



Amazon Kinesis Data Firehose

Easily load streaming data into AWS

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Simple pattern for streaming data

Data producer

Continuously creates data

Continuously writes data to a stream

Can be almost anything

Streaming service

Durably stores data

Provides temporary buffer that preps data

Supports very high-throughput

Data consumer

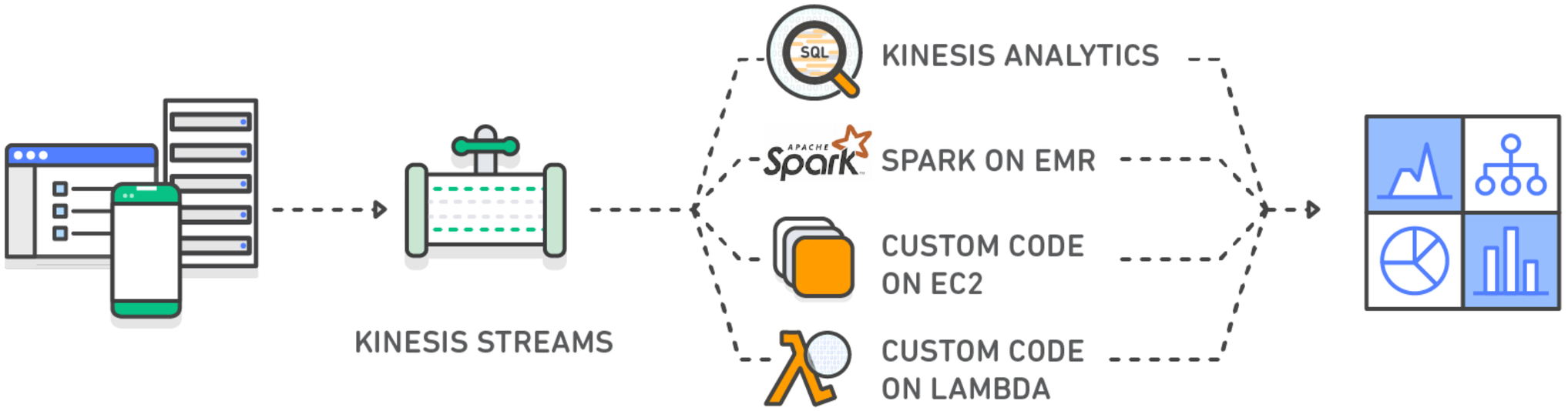
Continuously processes data

Cleans, prepares, & aggregates

Transforms data to information

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Kinesis Data Streams



Capture and send data to Kinesis Streams

Build custom, real-time applications using Kinesis Analytics, stream processing frameworks like Apache Spark, or your code running on Amazon EC2 or AWS Lambda

Load processed data to any data store, send real-time alerts, feed live dashboards, and more

- Easy administration and low cost
- Build real-time applications with framework of choice
- Secure, durable storage

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Kinesis Data Analytics

Capture streaming data with Kinesis Streams or Kinesis Firehose

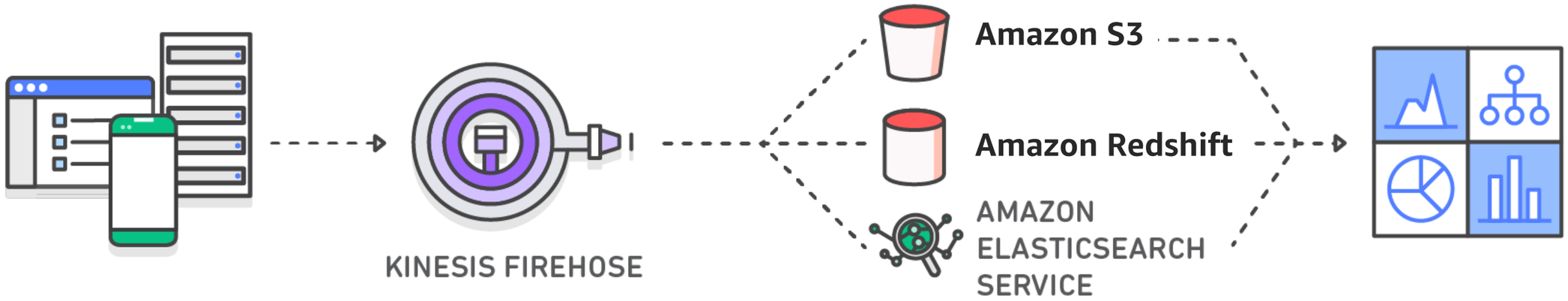
Run standard SQL queries against data streams

Kinesis Analytics can send processed data to analytics tools so you can create alerts and respond in real-time

- Powerful real-time applications
- Easy to use, fully managed
- Automatic elasticity

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Kinesis Data Firehose



Capture and send data to Kinesis Firehose

Kinesis Firehose prepares and loads the data continuously to the destinations you chose from among S3, Redshift, Amazon Elasticsearch Service, and Kinesis Analytics

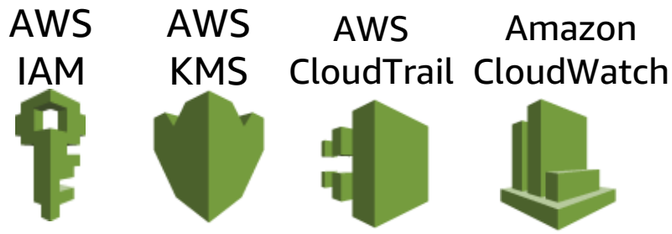
Analyze streaming data using your favorite BI tools

- Zero administration and seamless elasticity
- Direct-to-data store integration
- Serverless, continuous data transformations

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Modern data architecture (AWS)

Insights to enhance business applications, new digital services



Data sources

Transactions

ERP

Web logs / cookies

Connected devices

Social media

Ingest

Internet interfaces

AWS Direct Connect

AWS Database Migration Service

Amazon Kinesis

Scale (Batch)

Raw Data
Amazon S3

ETL
Amazon EMR

Advanced analytics
Apache Spark
MLlib

Staged data (Data lake)
Amazon S3

Speed (Real-time)

Stream Analysis
Amazon EMR
Apache Spark
Scala

Event Capture
Amazon Kinesis

Serving

Direct query
Amazon Athena

Schemaless
Amazon ES

Semi/Unstructured
Amazon EMR

Data Warehouse
Amazon Redshift

Legacy Apps
Amazon RDS

Near-zero latency
Amazon DynamoDB

Data scientists

Data analysts

Business users

Engagement platforms

Automation / events

Thank you!

AWS re:Invent

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Please complete the session survey in the mobile app.

AWS re:Invent

© 2018, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws