AWS re:Invent

OPN202-S

Will your infrastructure support your ML initiatives?

Abram Van Der Geest
Machine Learning Engineer
TIBCO Software, Inc.

Matt Ellis
Director, Product Management & Strategy
TIBCO Software, Inc.

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Agenda

The rise of data and the event

DL, AI, ML, huh?

Wait, say that part about that thing again…

Events are everywhere: Project Flogo

Event-driven ML-powered applications need to scale

The rise of data and the event

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

All hail the king… Data

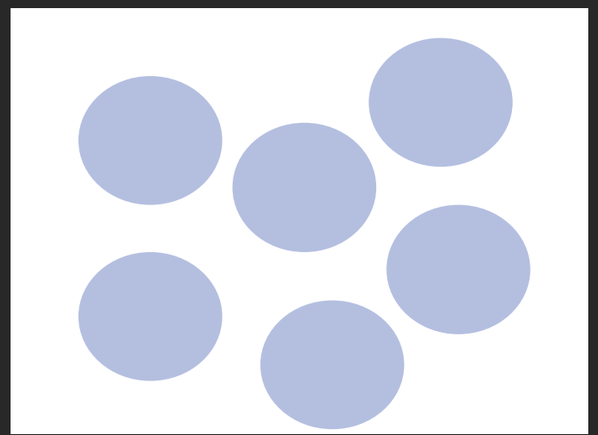


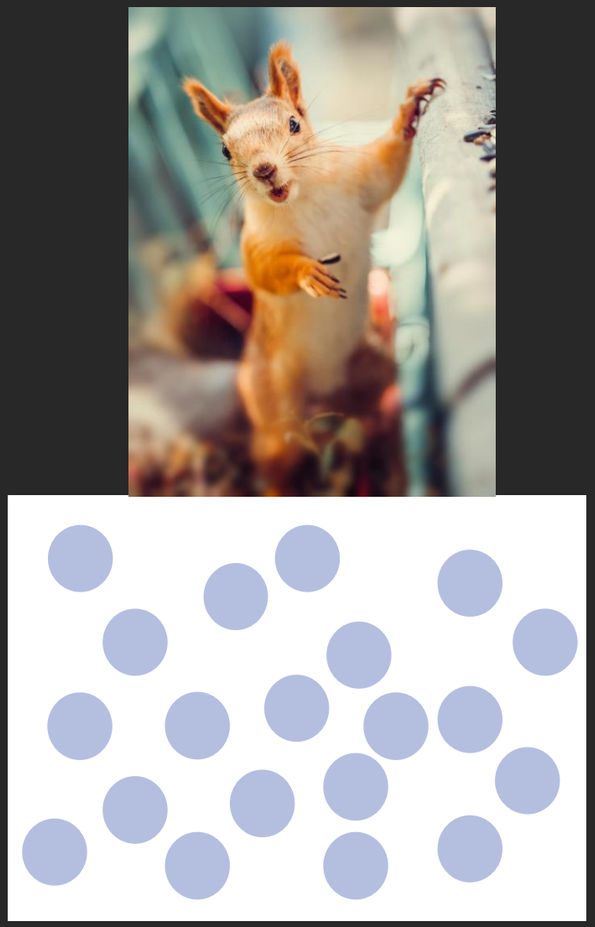
New application architectures

Hard to change

Easier







Monolith

Microservices

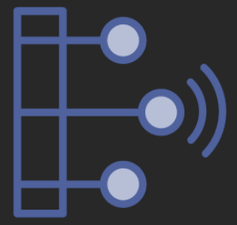
Functions

Goodbye software?!

Or so they say...



Does this mean that machine learning (ML) is always the answer?



Streaming: Data aggregation

- Median, mean, time-weighted averages, variability/robustness
- *Sometimes streaming data analytics in real time is sufficient for your problem!*



Machine learning

- Classify large quantities of data, such as images, text, etc.
- Broad set of patterns need to be detected
- Sufficient data must be available

DL, AI, ML, huh?

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

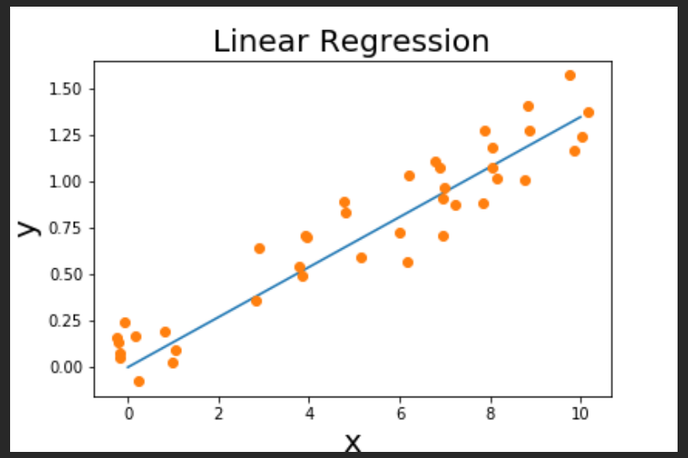
ML: Unsupervised vs. supervised

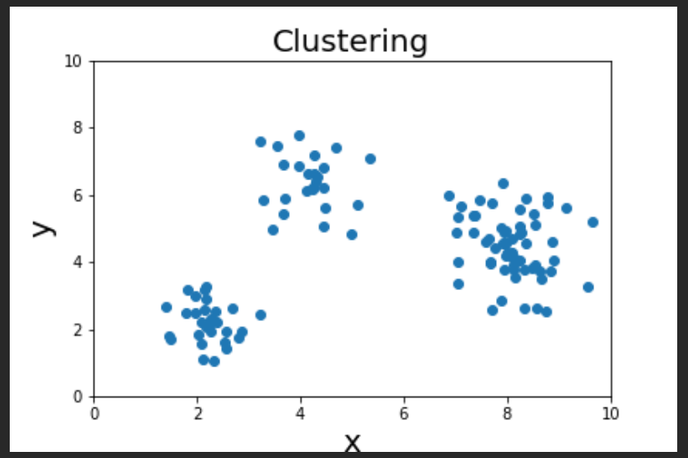
Supervised

- Model known problems: y = f(x)
- Predicts an observed condition; i.e., what factors are driving failures?
- Requires lots of *labeled* data
- Subsets of supervised learning
  - Semi-supervised learning: Using an incomplete training signal
  - Active learning: Optimizes choice of objects requiring labels
  - Reinforcement learning: Based on rewards and punishment

Unsupervised

- Explore observed data: x only
- Understand structure, detect anomalies; e.g., are there new patterns/failure modes emerging?
- Often used to uncover new phenomenon
- Feature choices drastically change information extracted
- Can be a goal in itself (new patterns) or can be used to discover features for supervised learning





ML: Applications

Classification

- Spam detection
- Activity detection
- Fraud detection
- Bucketing
- Recommendation engines
- And more

Clustering/pattern recognition

- Class discovery
- Feature discovery
- And more

Regression

- Any graph fitting
- Numeric prediction
- Value estimation
- And more

Dimensionality reduction

- Feature importance
- Feature reduction/selection
- Optimization of other ML (fewer dimensions = faster)
- Noise reduction
- Word embeddings
- And more

Information filtering

- Recommendation engines
- Collaborative filters
- Security investigation
- Improve signal-to-noise ratio of data
- And more

Outlier/anomaly detection

- Fraud detection
- Network intrusion detection
- Failure mode detection
- Discover trending topics
- And more

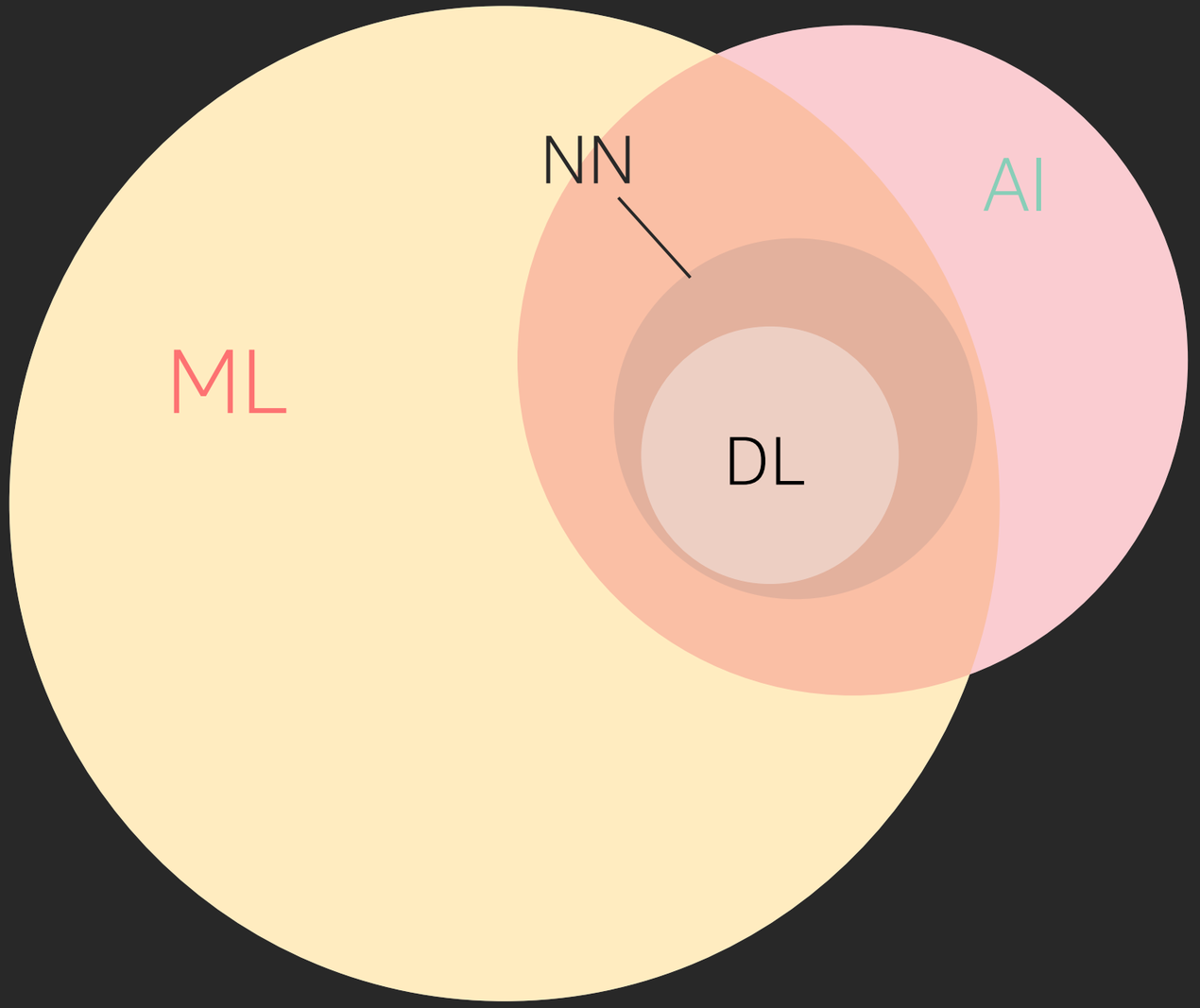
ML, AI, neural networks, and deep learning

Definitions

- **Machine learning (ML): A**lgorithms focused on learning from data to provide insights
- **Artificial intelligence (AI):** Computer systems that perform tasks that replicate human intelligence and activities
- **Neural networks (NN):** Class of algorithms originally modeled on the human brain that uses networks of linear algebra operations to perform human-like ML tasks
- **Deep learning (CL):** Hierarchical form of neural networks (read multi-layer NN) to learn data representations successful in computer vision, speech recognition, and NLP

Notes

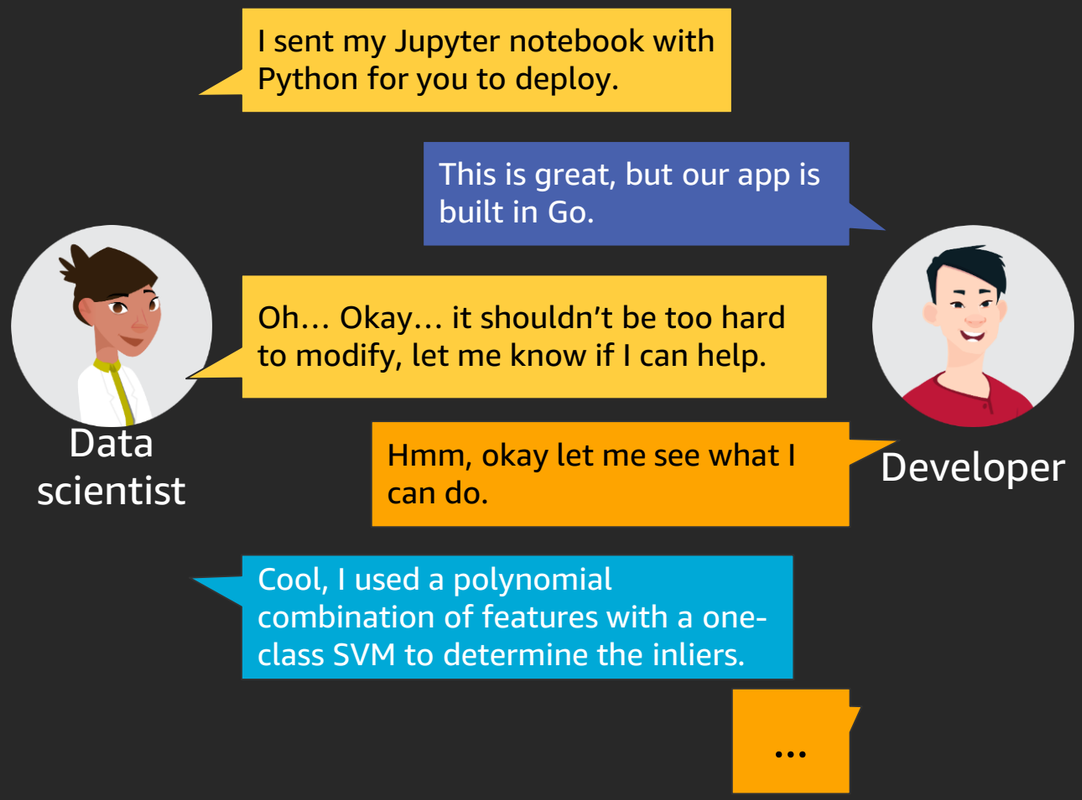
- All four terms have incorrectly been used interchangeably
- Some consider ML a subset of AI, others vice versa
- Forms of AI exist that do not rely on ML (i.e., rules based)
- NN and DL consist of algorithms (loosely) modeled on the human brain to perform human actions



Wait, say that part about that thing again…

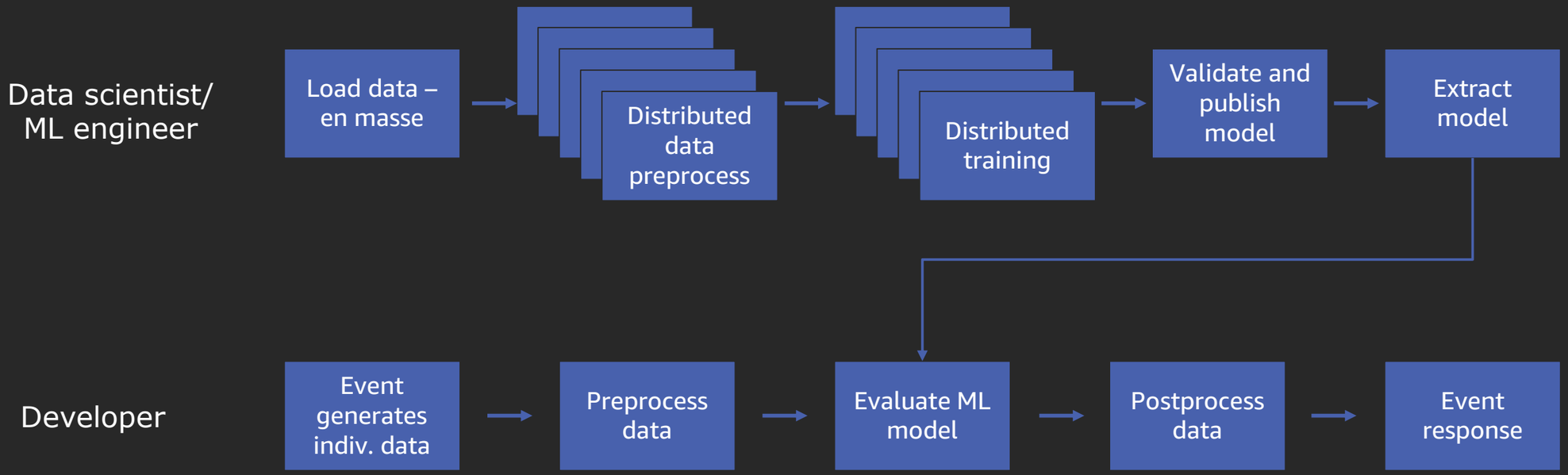
© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What's a dev gonna do?



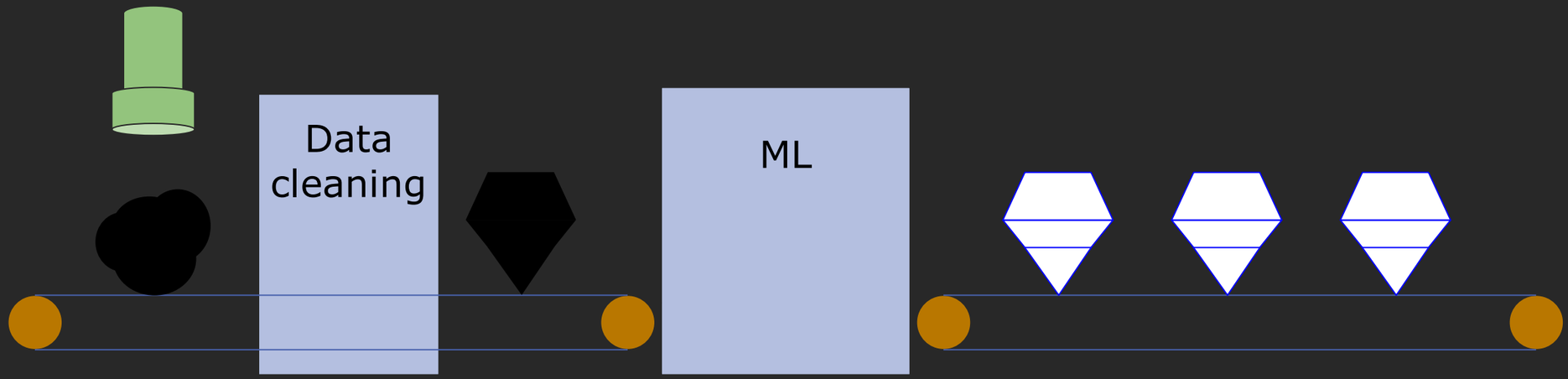
- Different tooling (language, framework, libraries, etc.)
- Model transferability
- Data preparation reproduced
- Communication of deployment needs
- Different success metrics

Similar process, but different targets



Real-time data transformations should be easy

- 80% of data science is data cleaning – production has to replicate much of this
- Data scientist and developers receive different starting data
- Developers aren’t trained for feature creation
- Tools for pre- and post-processing data in production unaddressed



CatalystML

Real-time data transformations for ML

Over 85 mathematical transformations

100% open-source specification & implementation

Language-agnostic

Embed into your real-time applications



Support for 1:1 mapping into ML models

Demo

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Events everywhere

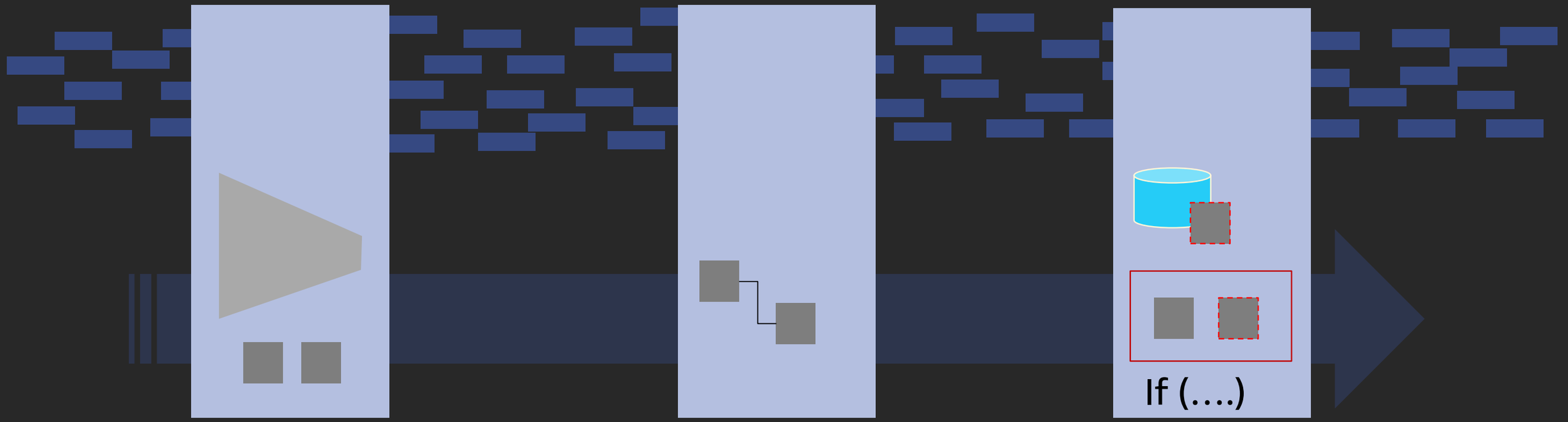
© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Sensors everywhere



Event-processing premise

The simple notion is that an event is just that – an event. Whether it be one or millions, its how you process them that differs.



Project Flogo



Project Flogo

Open-source ecosystem for event-driven applications



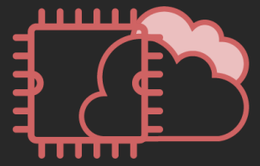
10–50x lighter than
Java, .NET, or Node.js



100% open-source stack
for all things event-driven



Common core for all
event-driven capabilities



Deploy as serverless
functions, containers, or to
IoT edge devices



Machine learning

Flogo Core

Flogo Core



Event-driven app kernel for serverless & IoT edge



Powerful event-driven programming model based on **triggers** and **actions**



Internal channel for action chaining; bundle all application needs into a single bin



Common contribution model enabling you to leverage 100s of activities and triggers

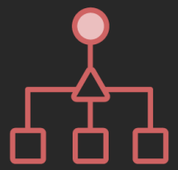


Simple interface to extend event processing capabilities



Integration Flows

Ultra-lightweight application integration for serverless and edge



Ultra-light process engine for application integration



Conditional control flow



Powerful step-back debugging capabilities in web browser

Flogo Streams

Stream pipeline for edge & cloud-native



Lightweight stream process for edge devices

f(x)

Aggregation capabilities

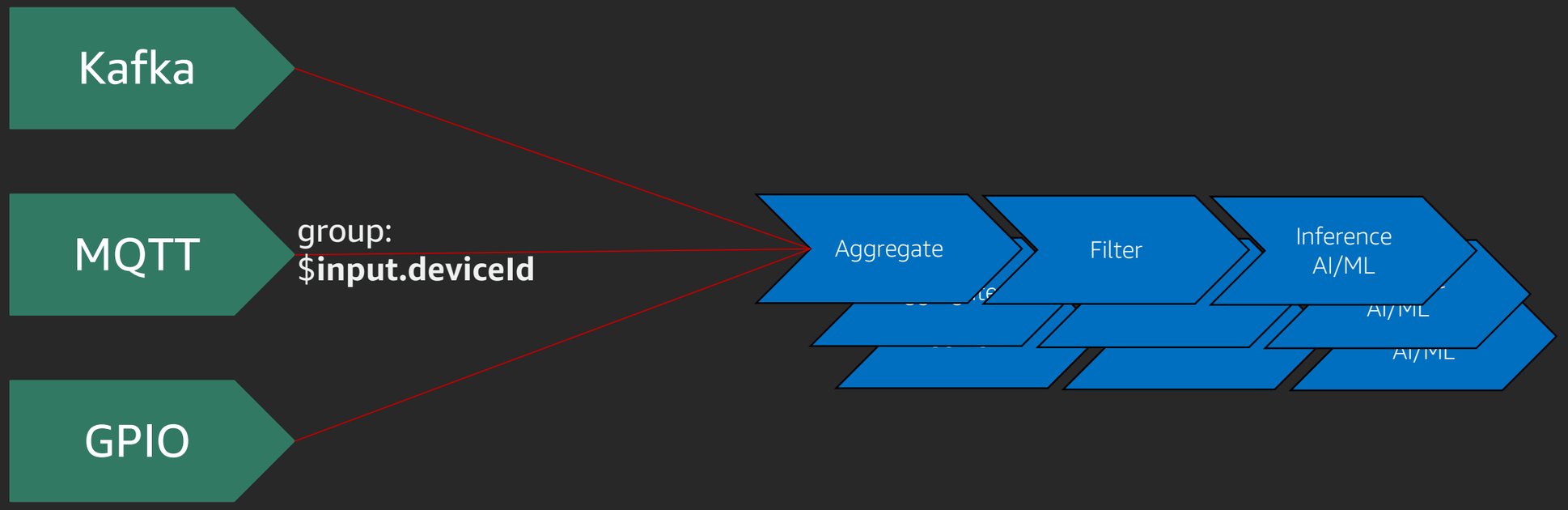
Join streams from multiple event sources

Filter out the noise

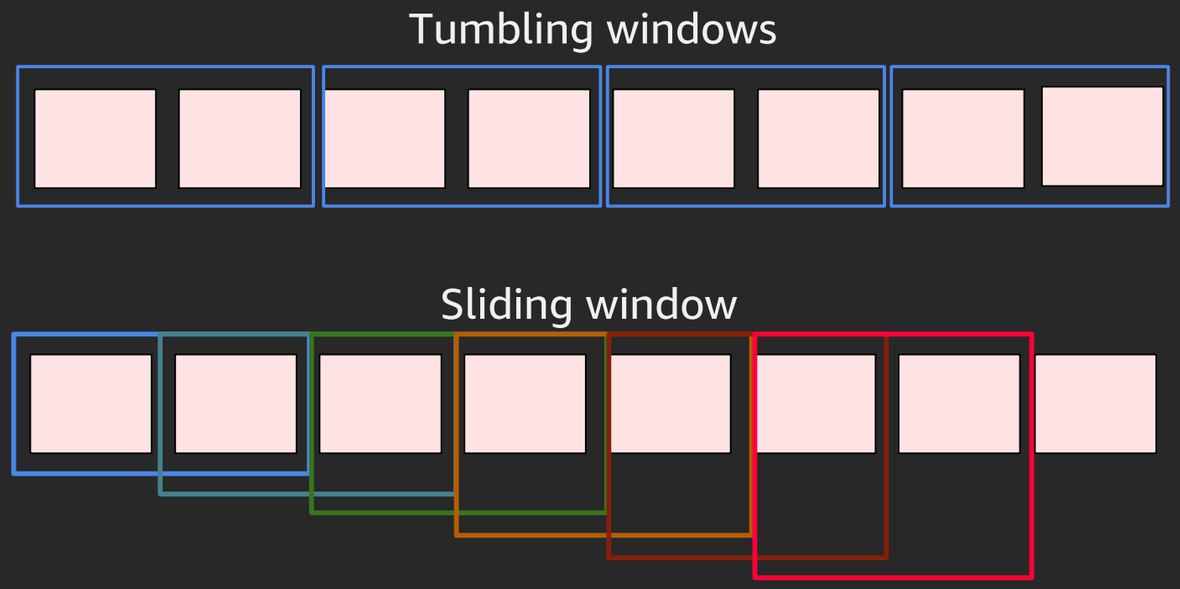
Grouping & joining events across triggers

With Flogo Streams, you can group and join events across multiple different streams of data

Joins occur against pipeline input/output



Aggregation



Operations

- Tumbling
- Time tumbling
- Sliding
- Time sliding

Functions

Avg, sum, min, max, count, accumulate

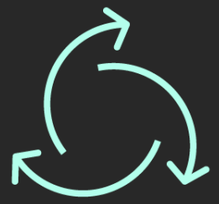


Rules Action for Flogo

Real-time contextual decisions



Real-time decisions to determine next best action/outcome and experience



Open-source/core complete with extensions, APIs, and interfaces

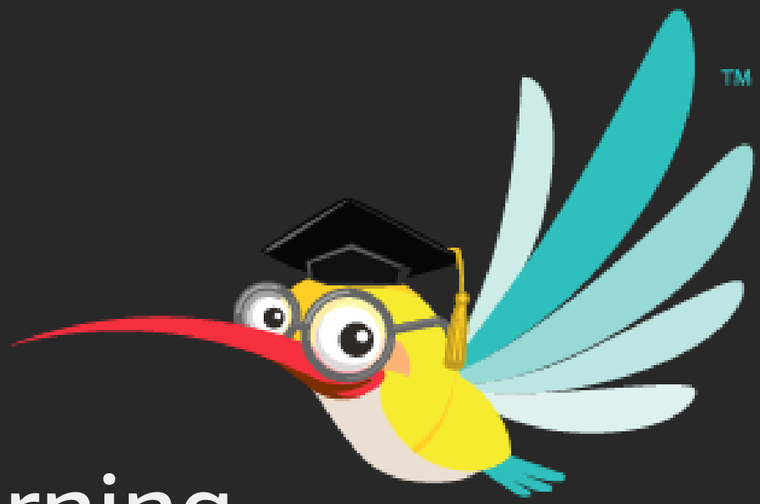


Declarative rules/rapid application development



Stateful/contextual reasoning across time and space

ML capabilities



Native Machine Learning

Execute TensorFlow models

100% open-source with zero lock-in

Serverless means…



No servers to provision or manage



Scales with usage



Pay for value



Availability and fault tolerance built in

Scale to meet demand

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Serverless means…

Serverless means…



No servers to provision or manage



Scales with usage



Pay for value



Availability and fault tolerance built in

Serverless means…

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Serverless means…

Greater agility

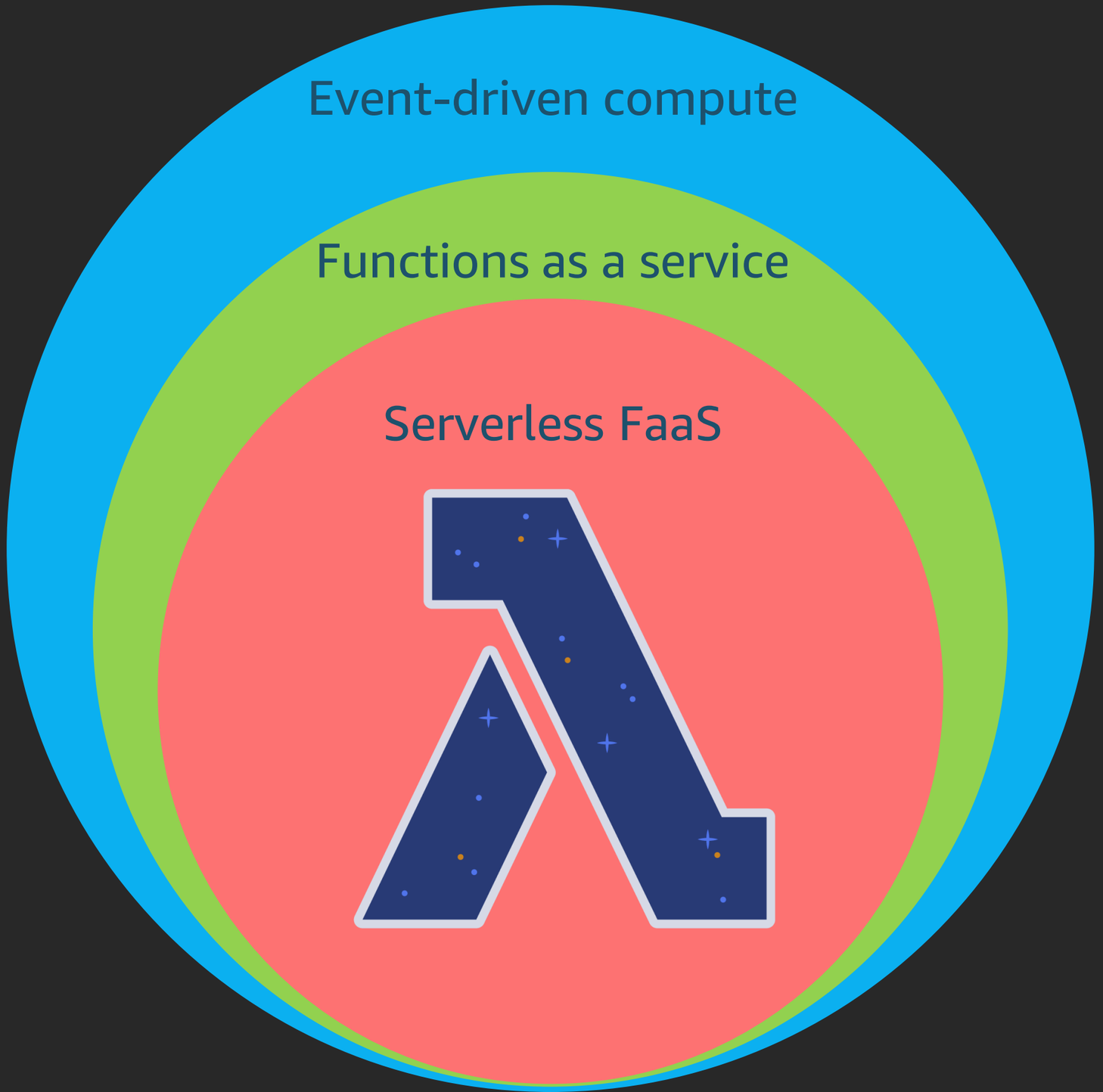
Less overhead

Better focus

Increased scale

More flexibility

Faster time to market


Event-driven compute
Functions as a service
Serverless FaaS

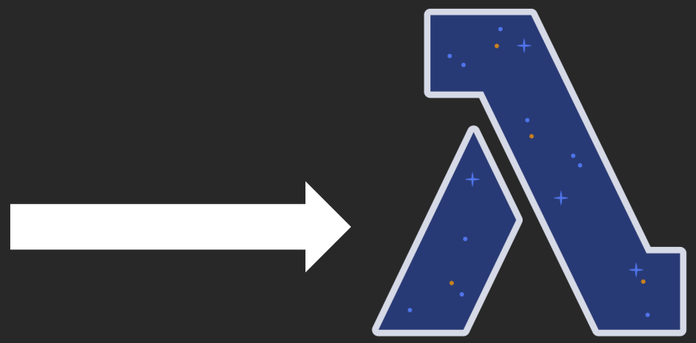
Serverless applications

Event source

Function

Services (anything)









Changes in data state



Requests to endpoints

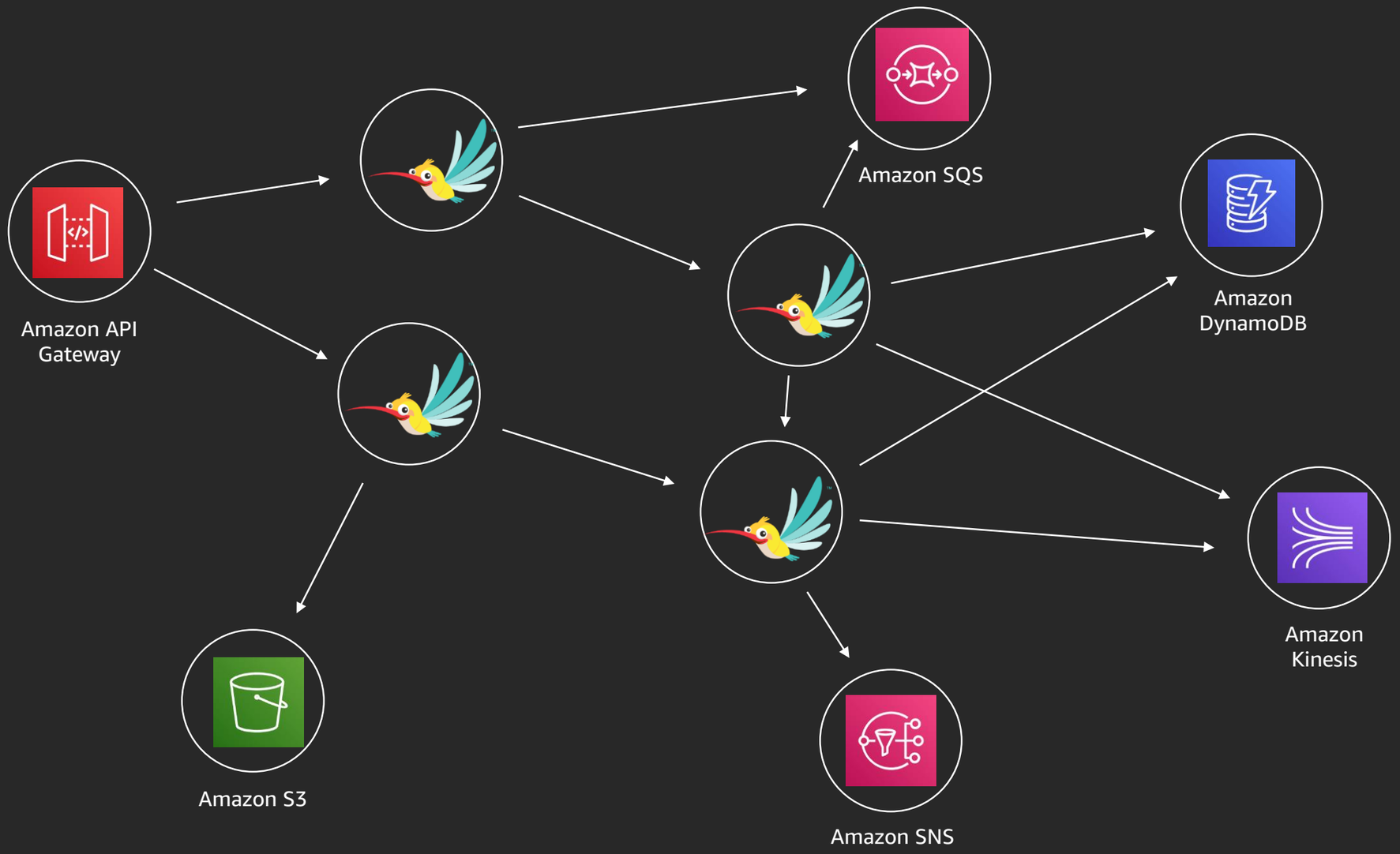


Changes in resource state



Node.js
Python
Java
C#
Go
Ruby
Runtime API
Project Flogo

This is where Project Flogo fits in perfectly



Demo

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Abram Van Der Geest
@AbramGeest

Matt Ellis
@mellismi

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws



Please complete the session survey in the mobile app.

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.