amazon web services

AWS INNOVATE

ONLINE CONFERENCE 2017

Serverless Data Processing on AWS

Daniel Zoltak

What to Expect From the Session

Together, we will:

- Explore a real use cases of **IoT Analytics** using the **Amazon Kinesis** family of services.
- See a demo of IoT and Amazon Kinesis in action.
- Take a deep dive into underlying **reference architectures** and **implementation.**

amazon web services
intel

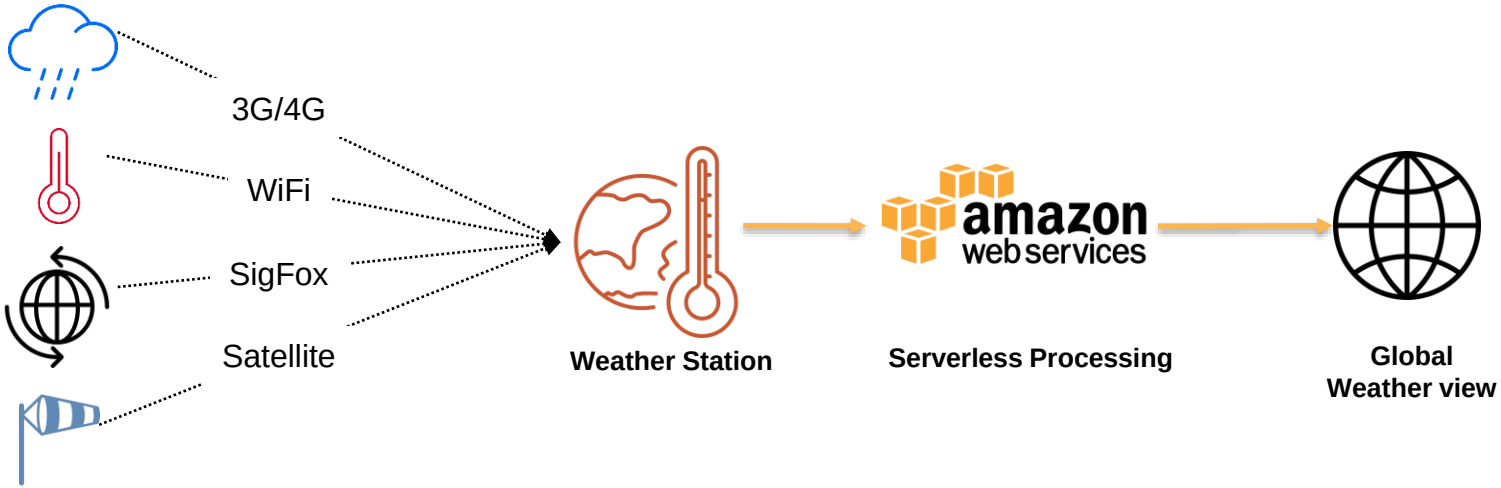
What to Expect From the Session

By the end of this session, you will:

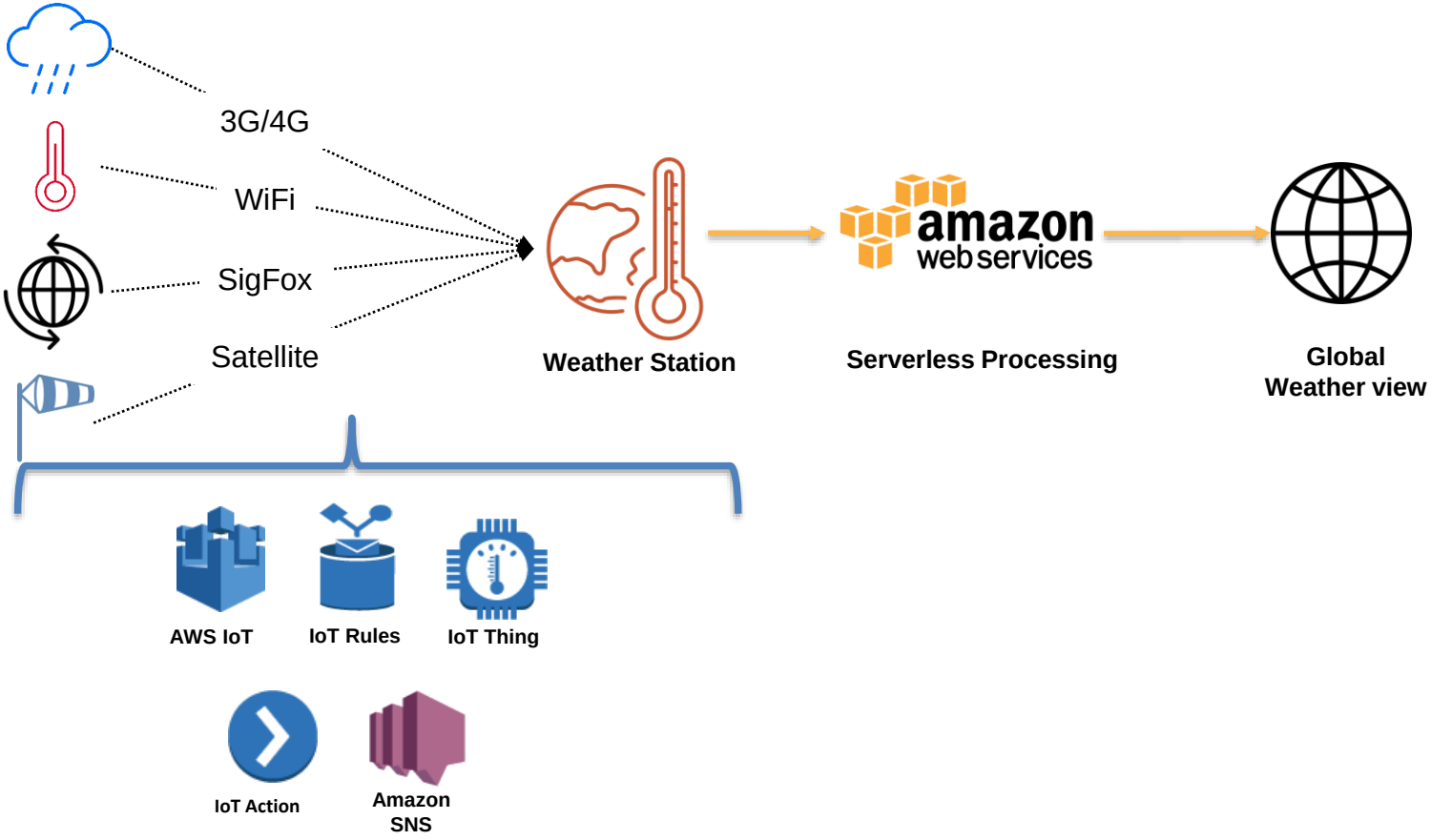
- Have an appreciation of the AWS services required to build a **serverless** IoT **analytics** platform.
- Be able to describe the role and functionality of Amazon Kinesis **Firehose**, Amazon Kinesis **Streams**, and Amazon Kinesis **Analytics**.
- Understand how to **acquire**, **process**, and **store** IoT data.

amazon web services
intel

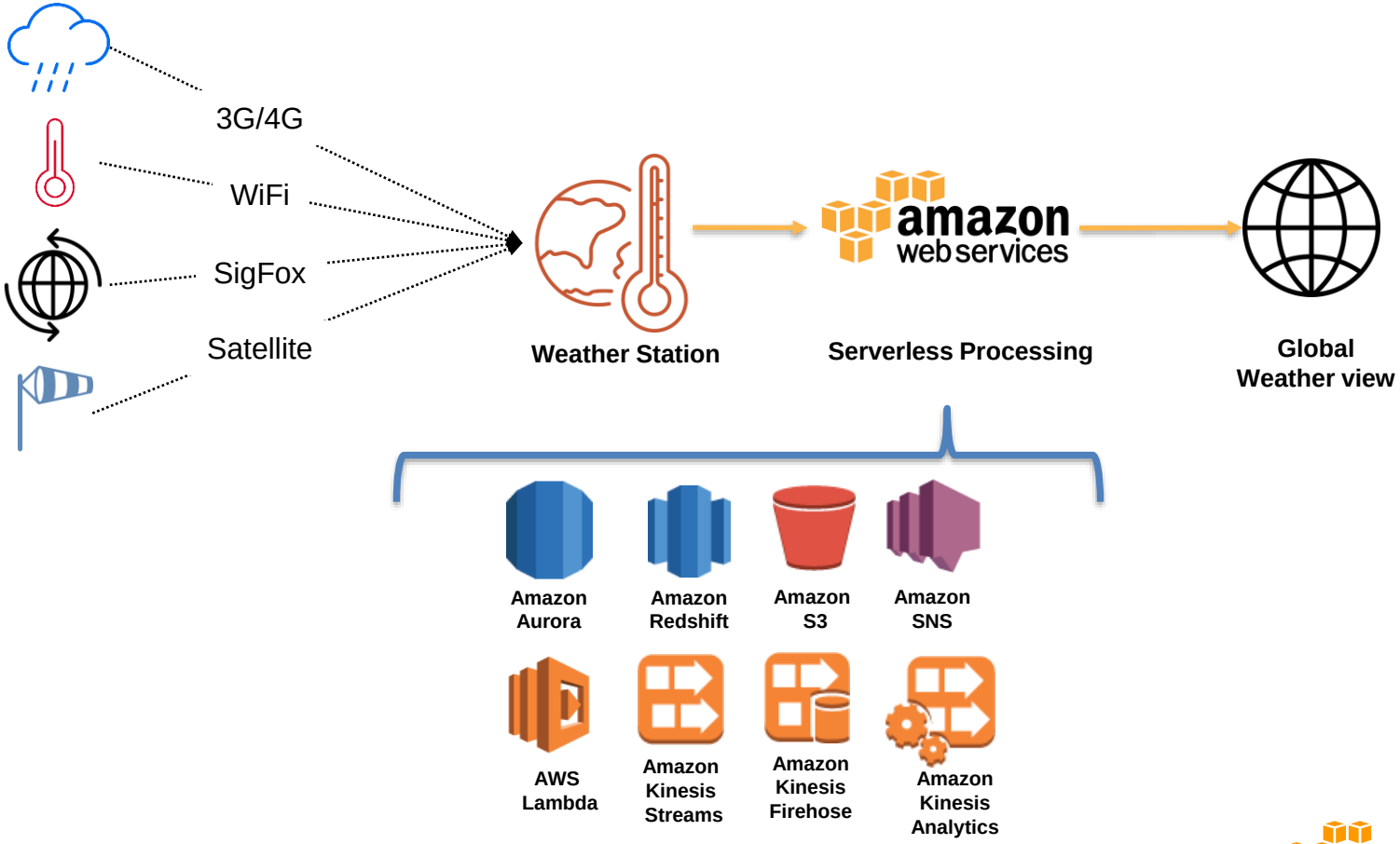
What You Are About to See



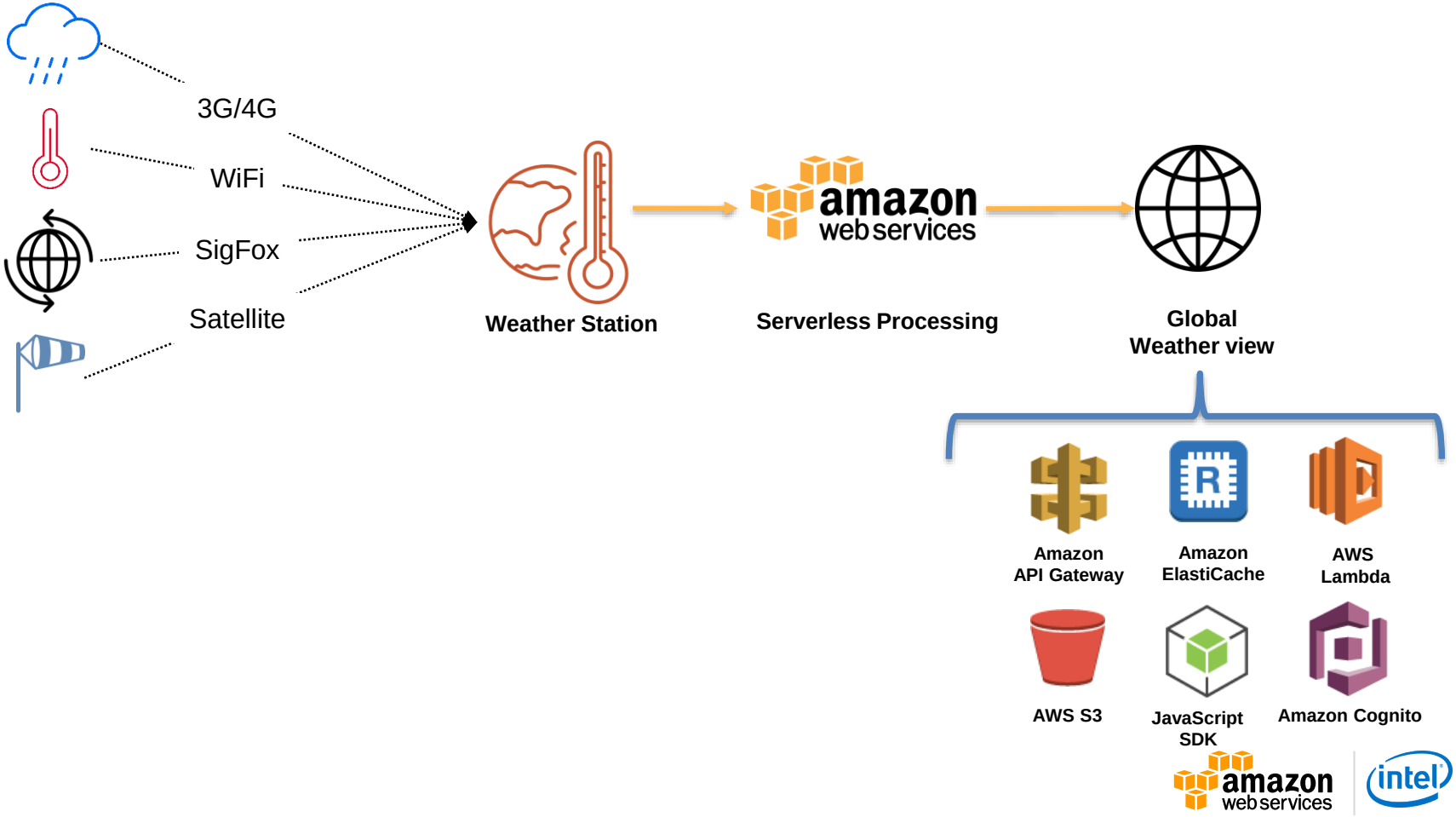
What You Are About to See



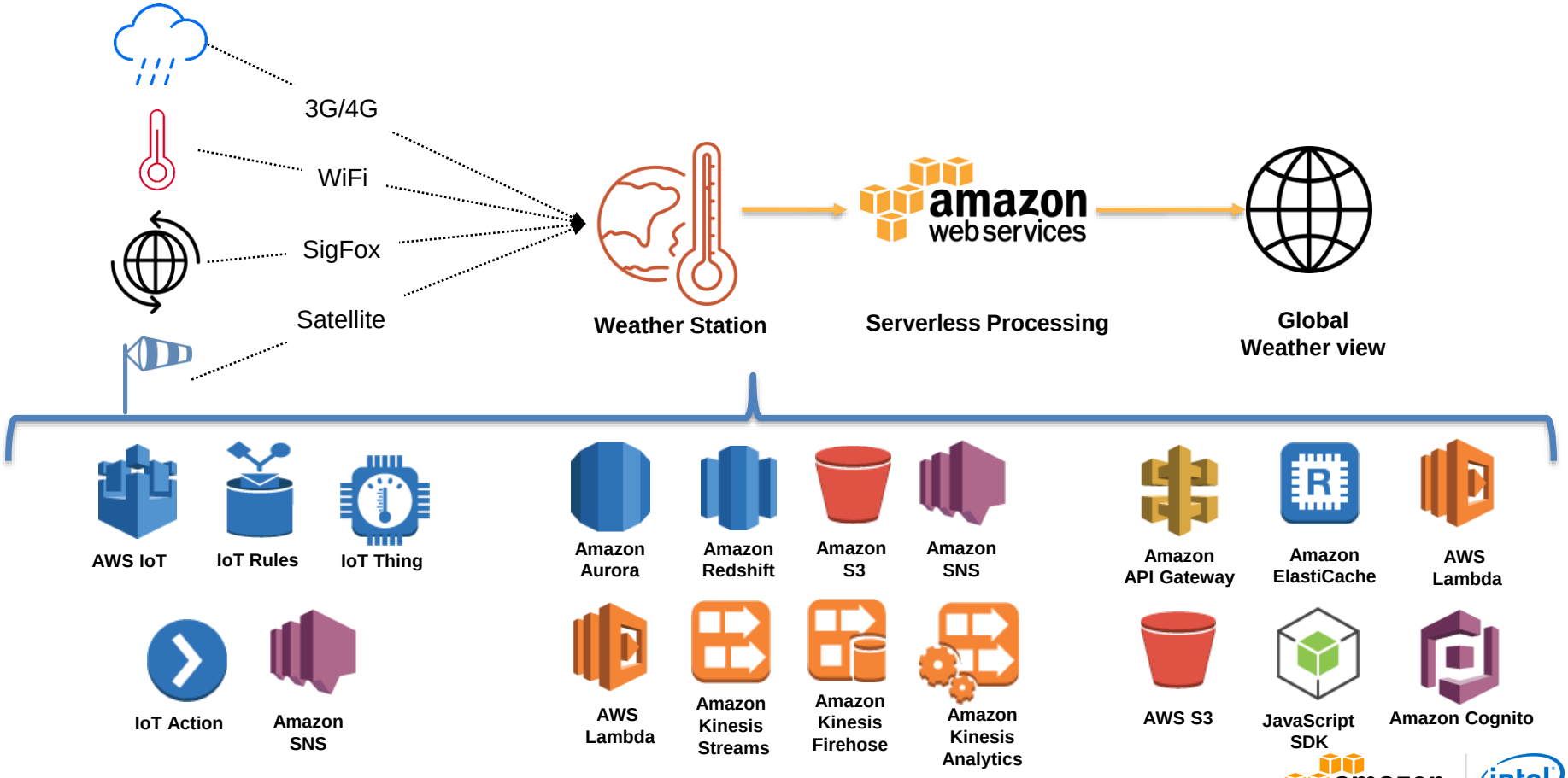
What You Are About to See



What You Are About to See



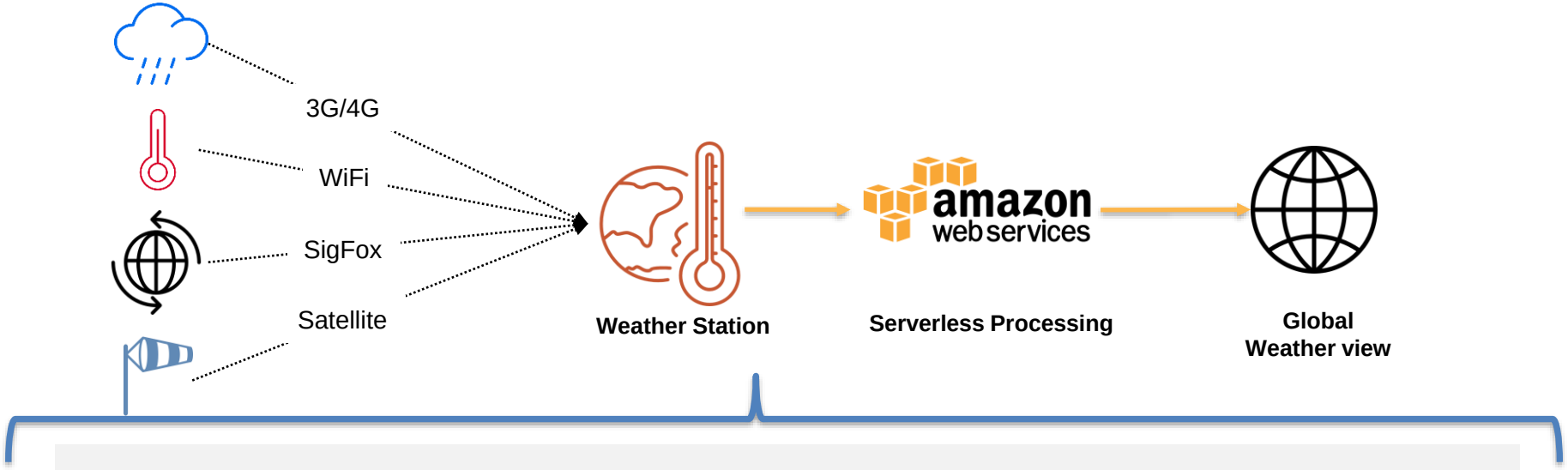
What You Are About to See



amazon web services

intel

What You Are About to See



10 AWS features and services
&
0 servers to manage

amazon web services
intel

Demo

View Demo: https://s3-ap-southeast-1.amazonaws.com/mktg-apac/Final+AWS+Innovate+Videos/AWS+Innovate+-+Demos/Serverless+Data+Processing+on+AWS+(Level+300)+-+Demo+1.mp4

amazon web services
intel

What Do Our Customers Ask For?

Our customers ask us to help them

- Ingest **large volumes** of real-time data from a large fleet of distributed IoT devices at scale.
- Perform advanced **analytics** of streaming data in real-time.
- **Process** and store large volumes of data.
- Eliminate capacity planning, **scaling**, and the management of infrastructure.

amazon web services | intel

What Do Our Customers Ask For?

Designing for failure in global, real-time, distributed systems is hard.

amazon web services
intel

What Do Our Customers Ask For?

Designing for failure in global, real-time, distributed systems is hard.

Infrastructure required to process billions of devices sending trillions of messages is expensive.

amazon web services
intel

What Do Our Customers Ask For?

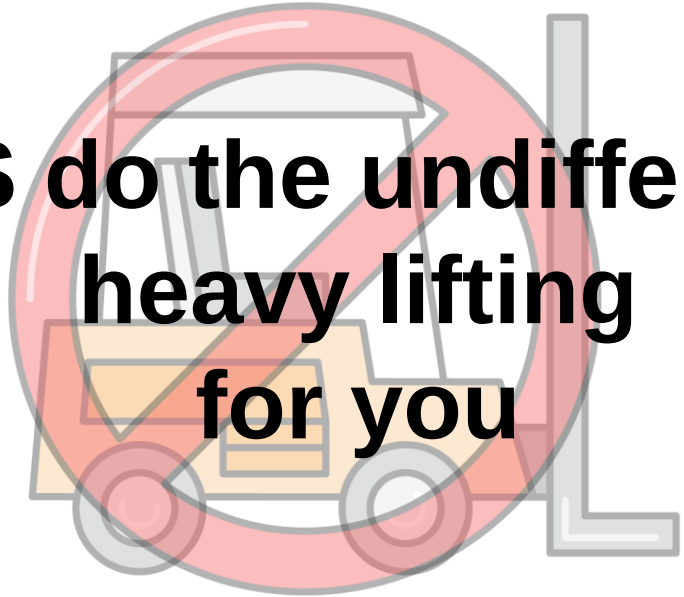
Designing for failure in global, real-time, distributed systems is hard.

Infrastructure required to process billions of devices sending trillions of messages is expensive.

Management overhead and scale limitations impede innovation.

amazon web services | intel

What Do Our Customers Ask For?



Let AWS do the undifferentiated heavy lifting for you

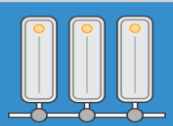
amazon web services
intel

Reference Model

amazon web services
intel

Reference Model

SECURITY



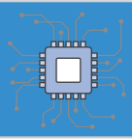
NETWORKING

amazon web services

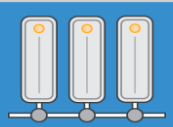
intel

Reference Model

SECURITY



COMPUTE



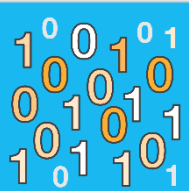
NETWORKING

amazon web services

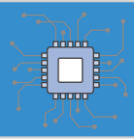
intel

Reference Model

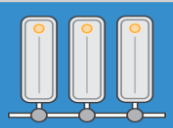
SECURITY



DATA SOURCE



COMPUTE

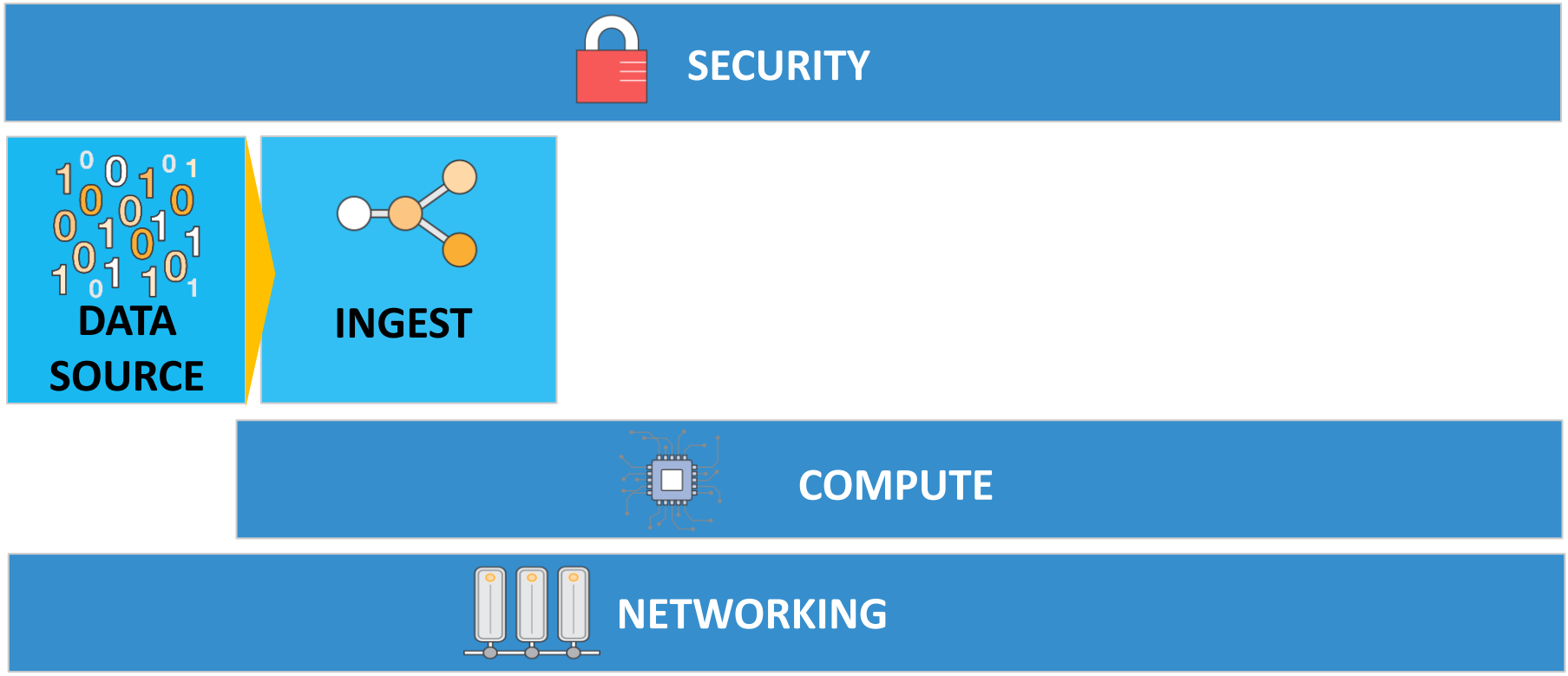


NETWORKING

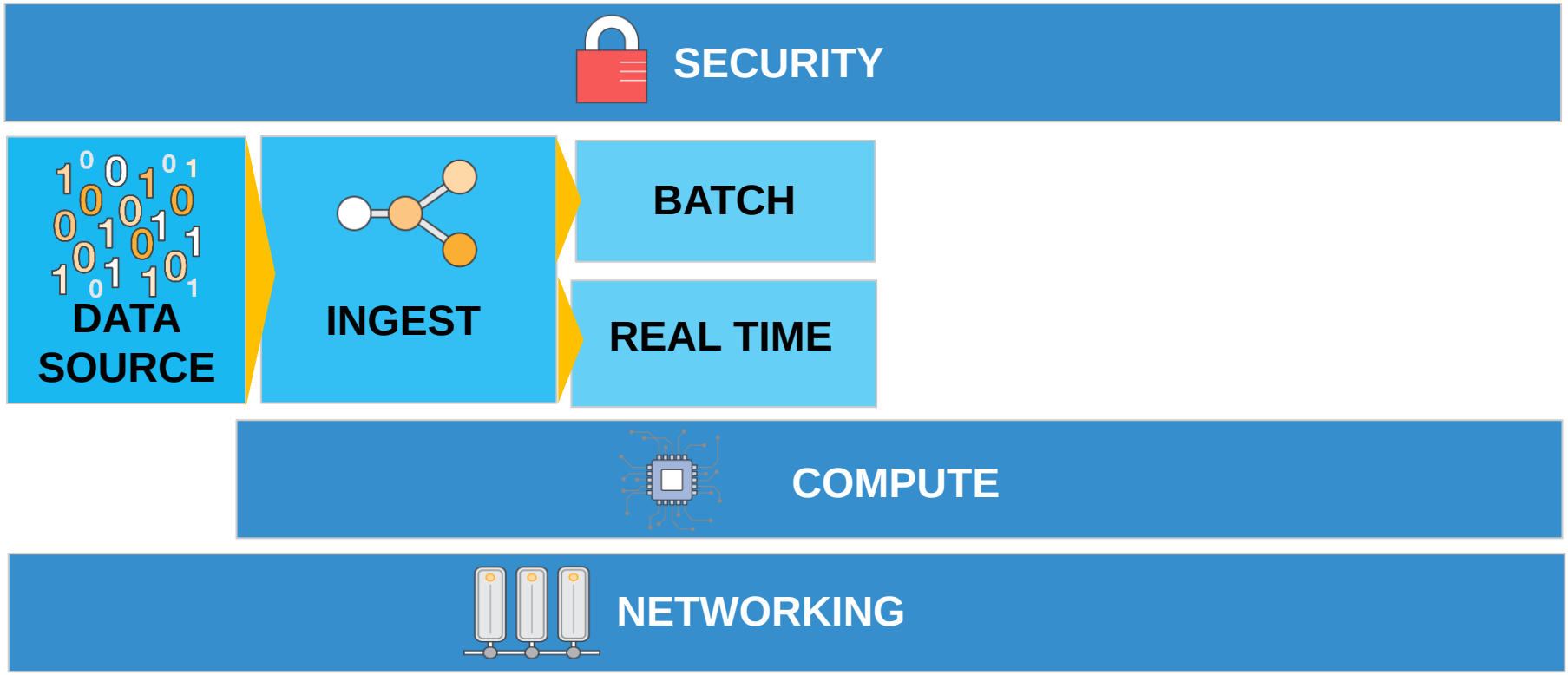
amazon web services

intel

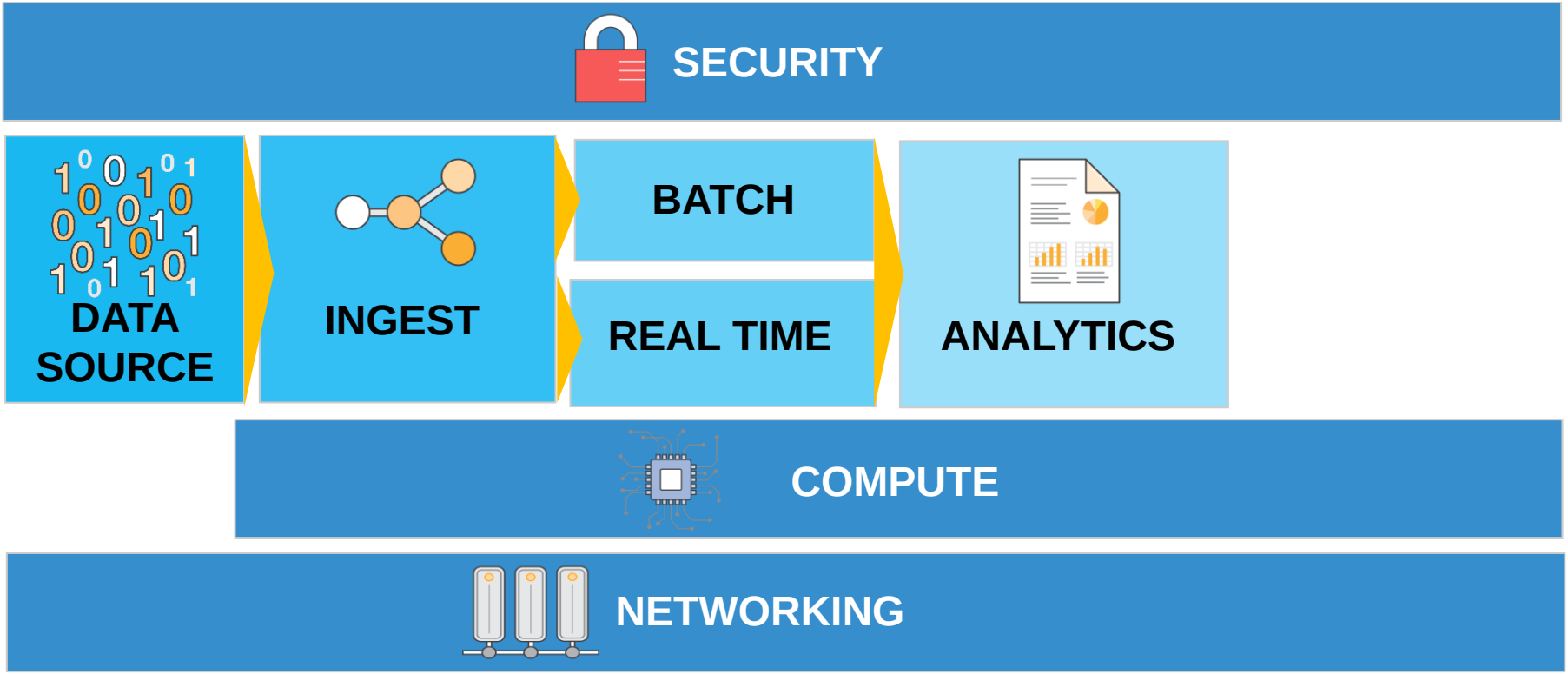
Reference Model



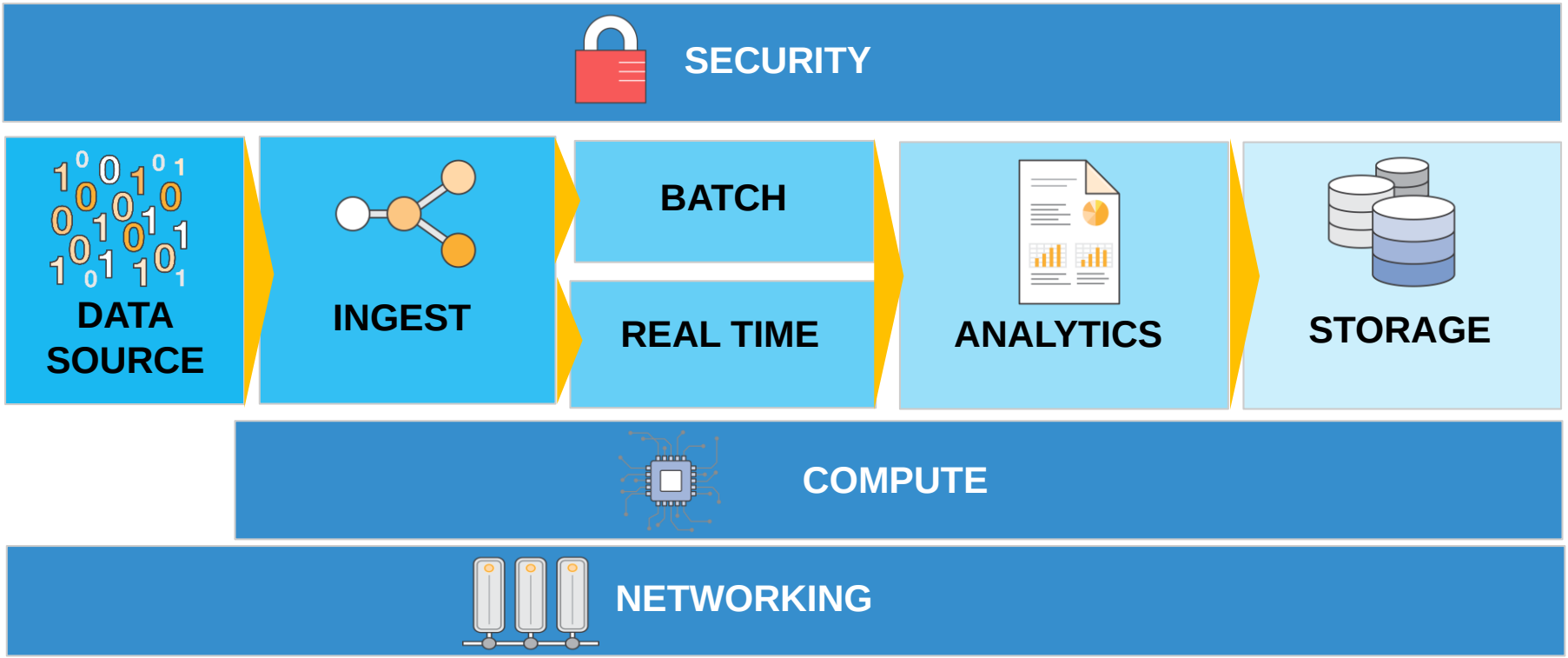
Reference Model



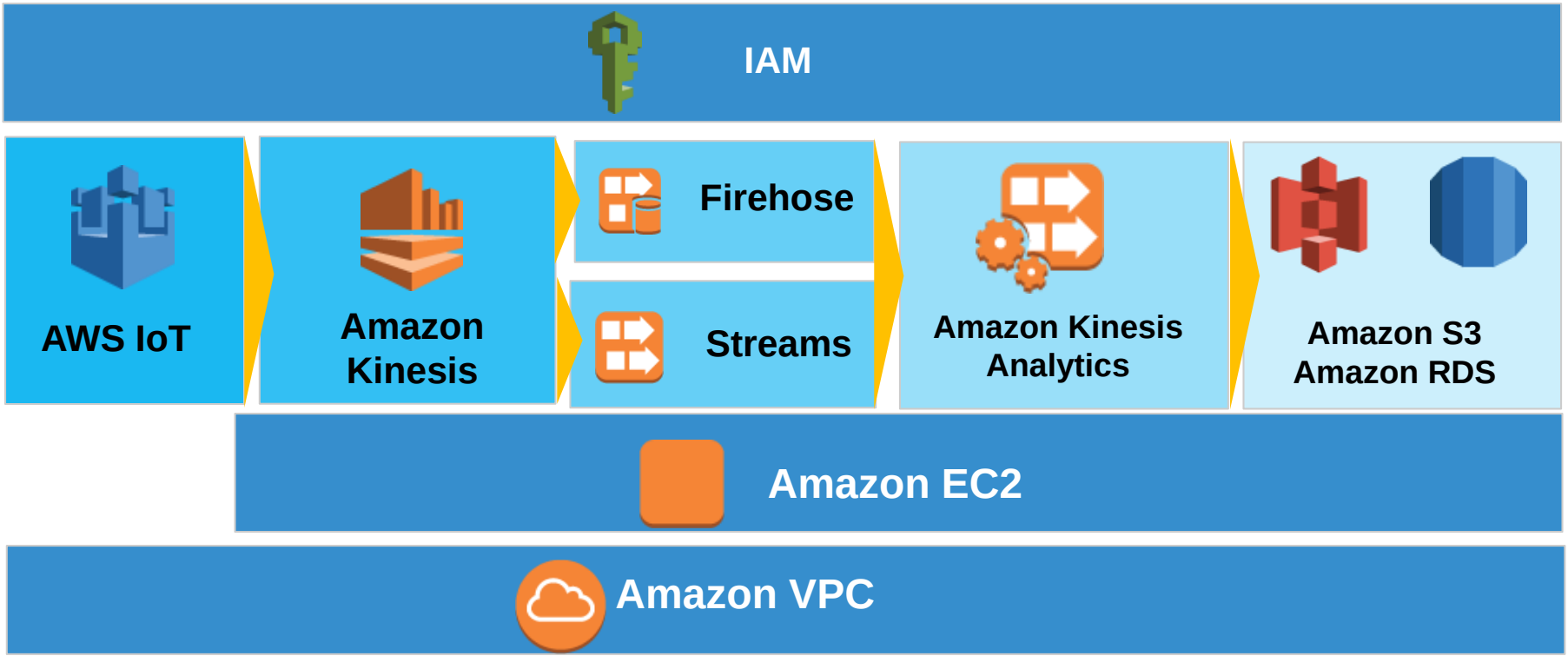
Reference Model



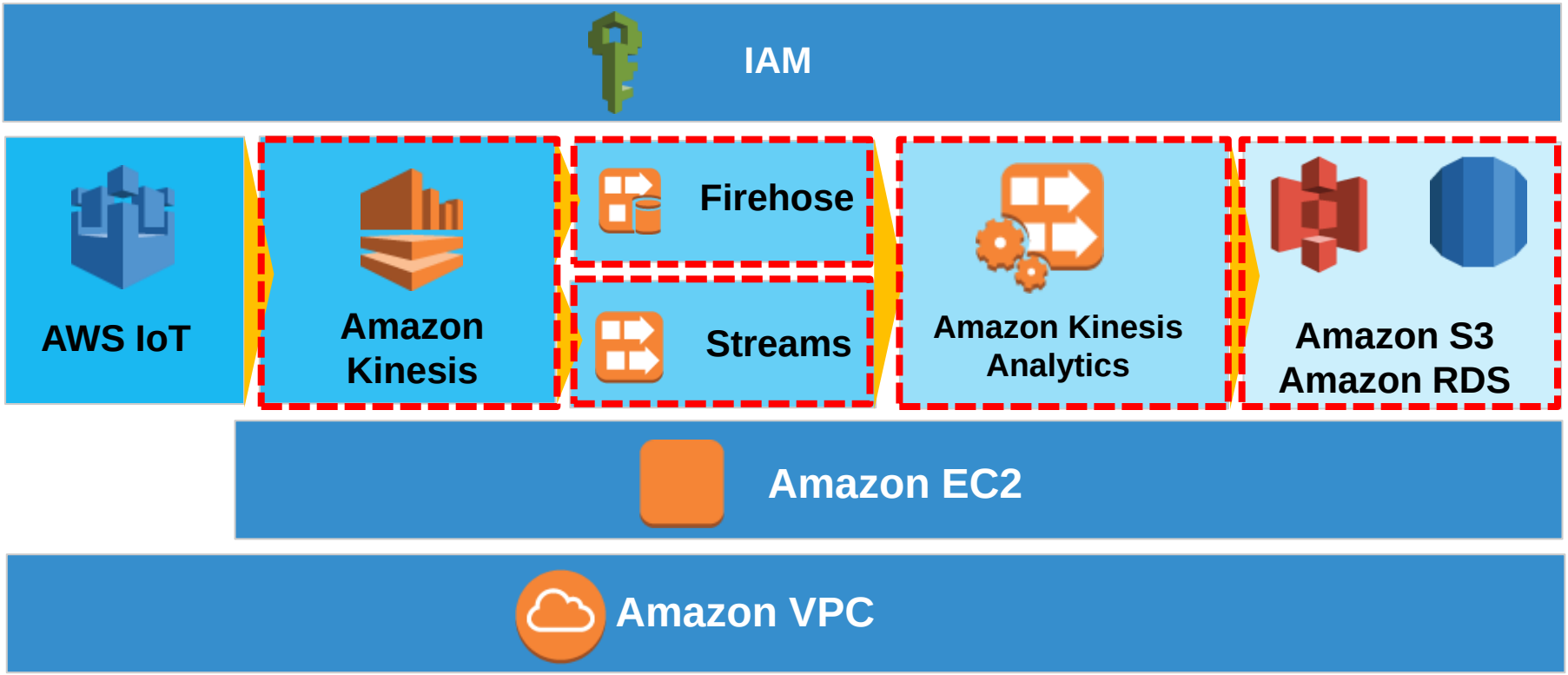
Reference Model



Reference Model

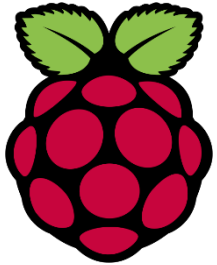


Reference Model - Focus Today

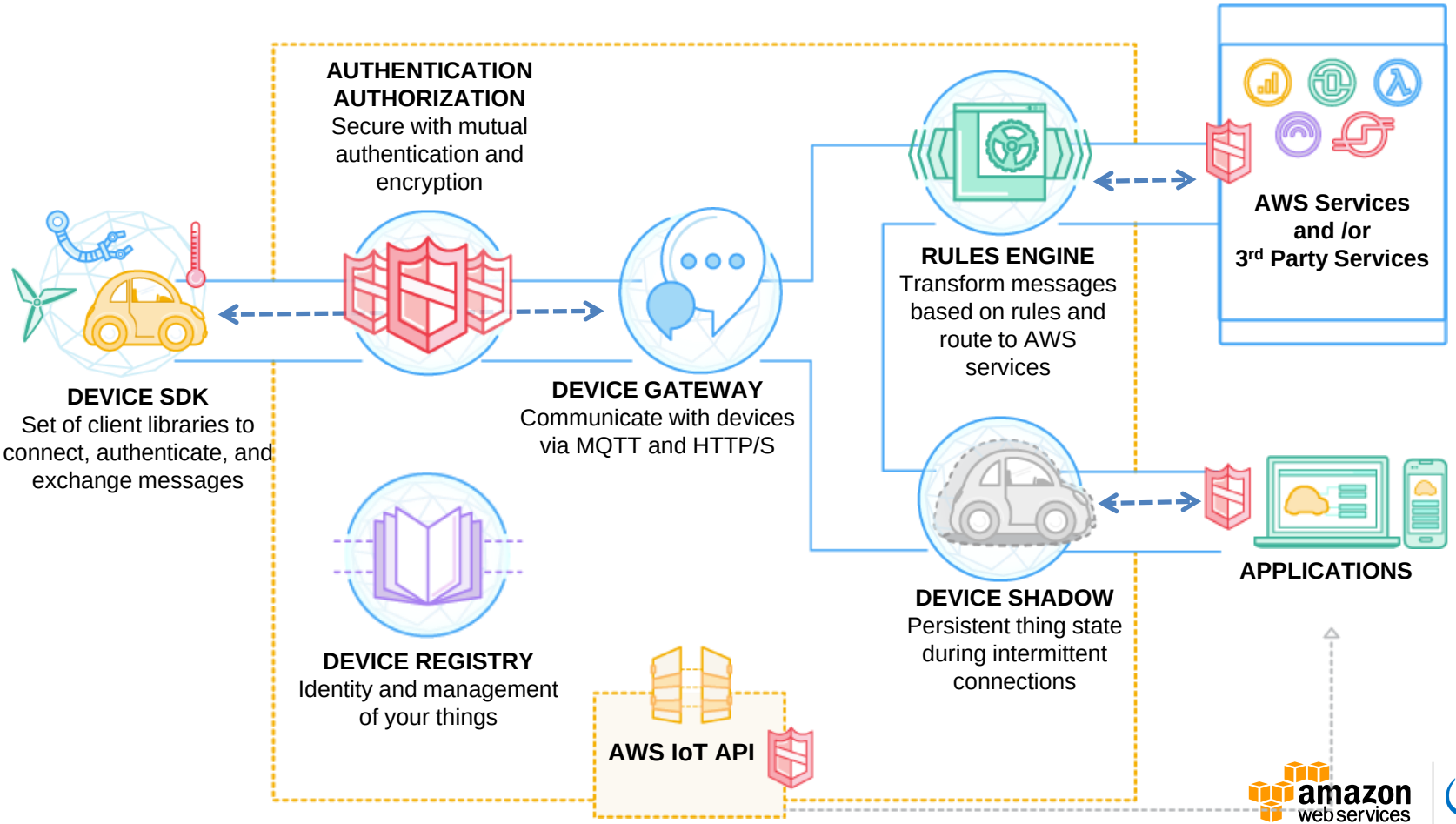


What Is An IoT “Thing”?

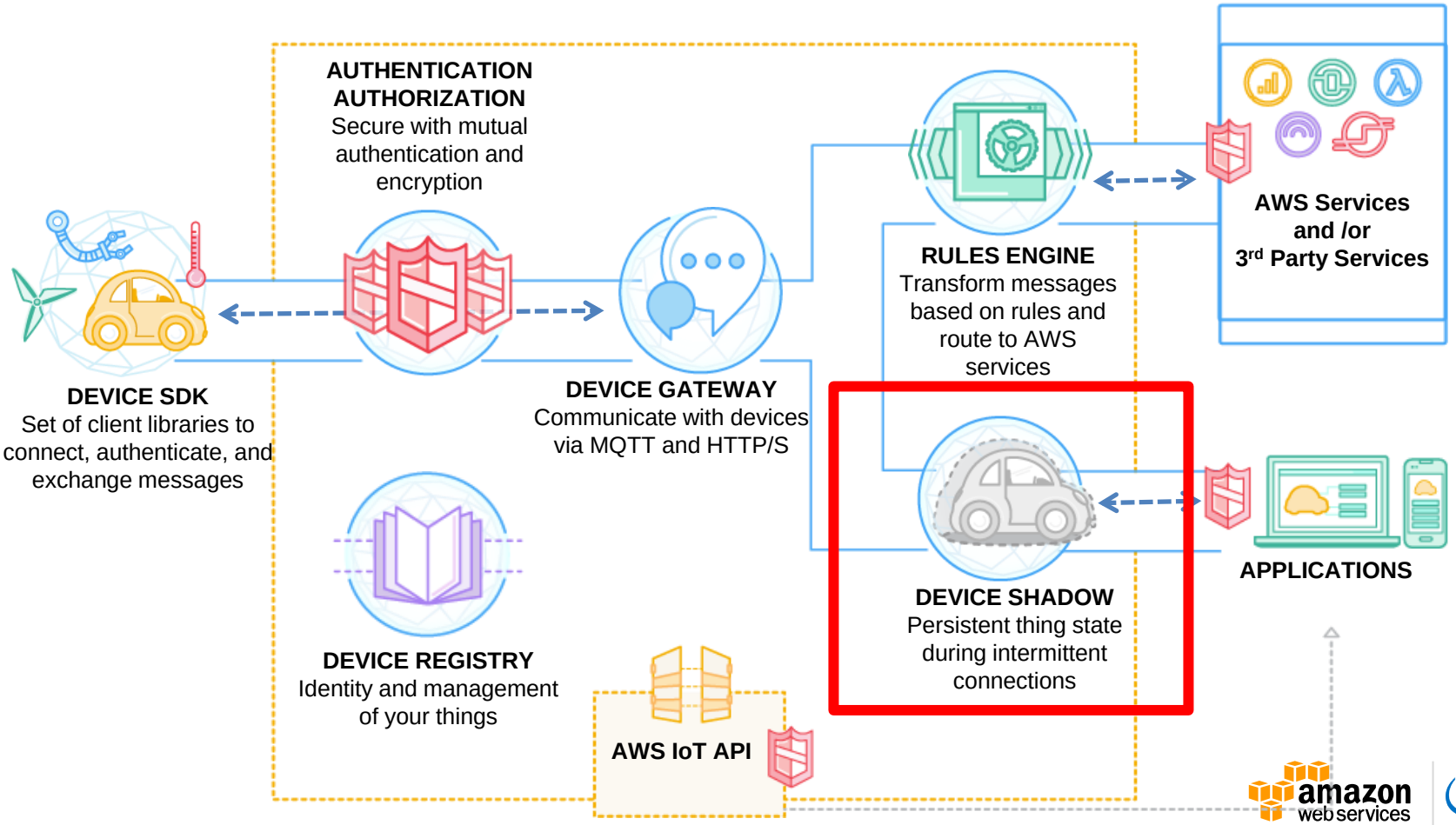
- Mobile Devices
  - IOS, Android, Kindle, Tablets.
- Maker Devices
  - Arduino, Raspberry Pi, Intel Edison.
- Embedded devices and wearables
  - Health and fitness management; safety and tracking.
- Smart Home
  - Smoke alarms, temperature sensors, light globes, and switches.



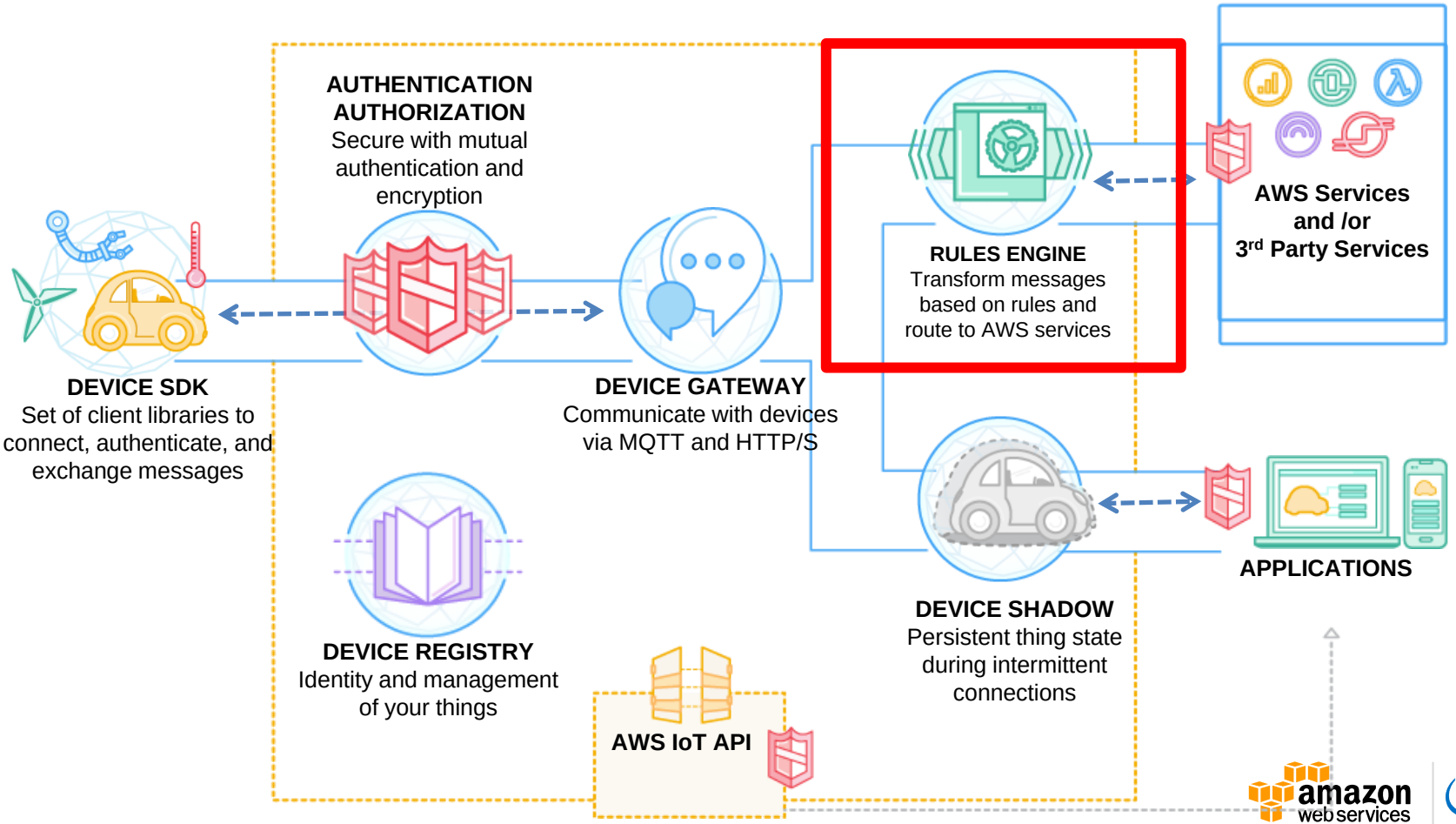
AWS IoT Framework



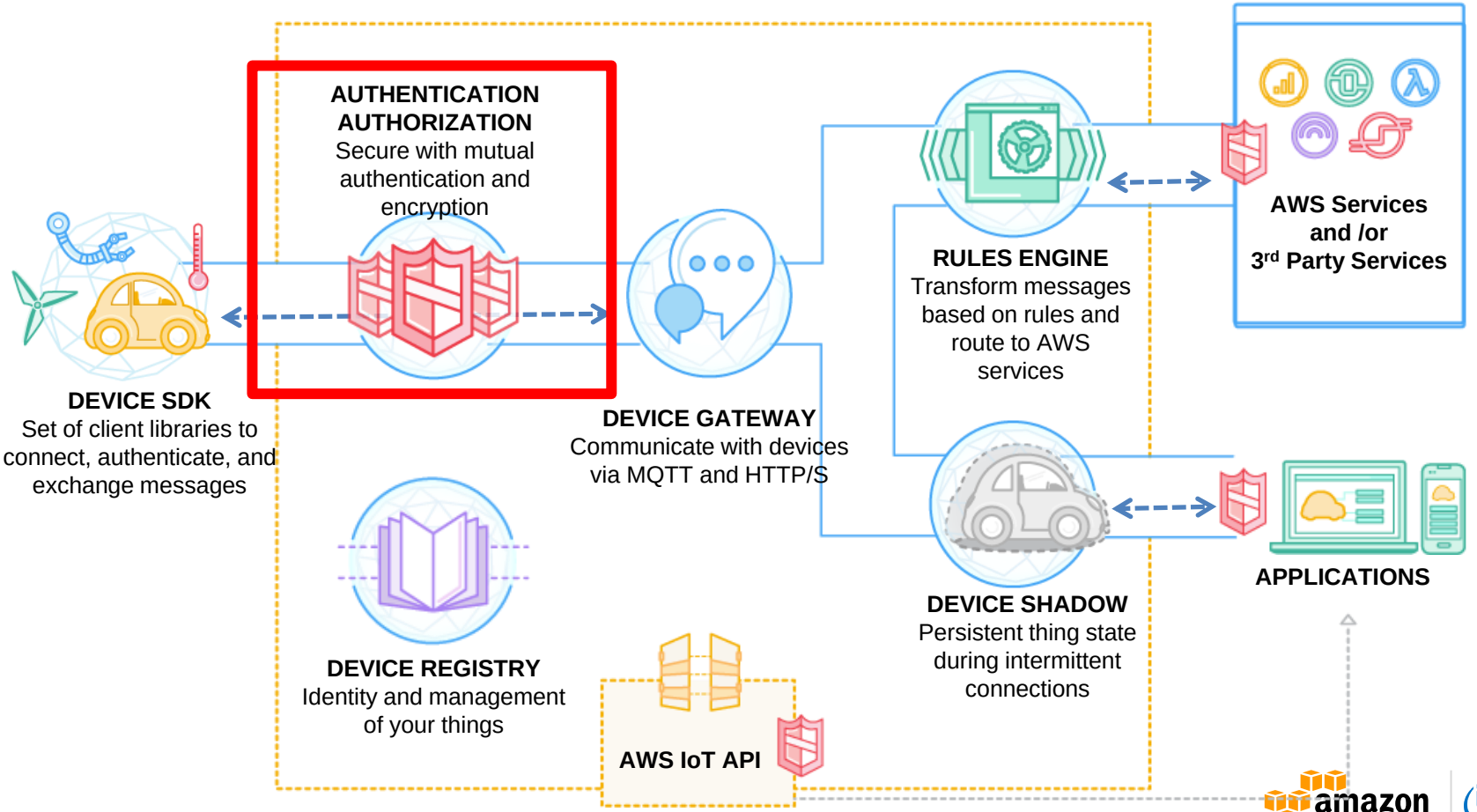
AWS IoT Framework



AWS IoT Framework



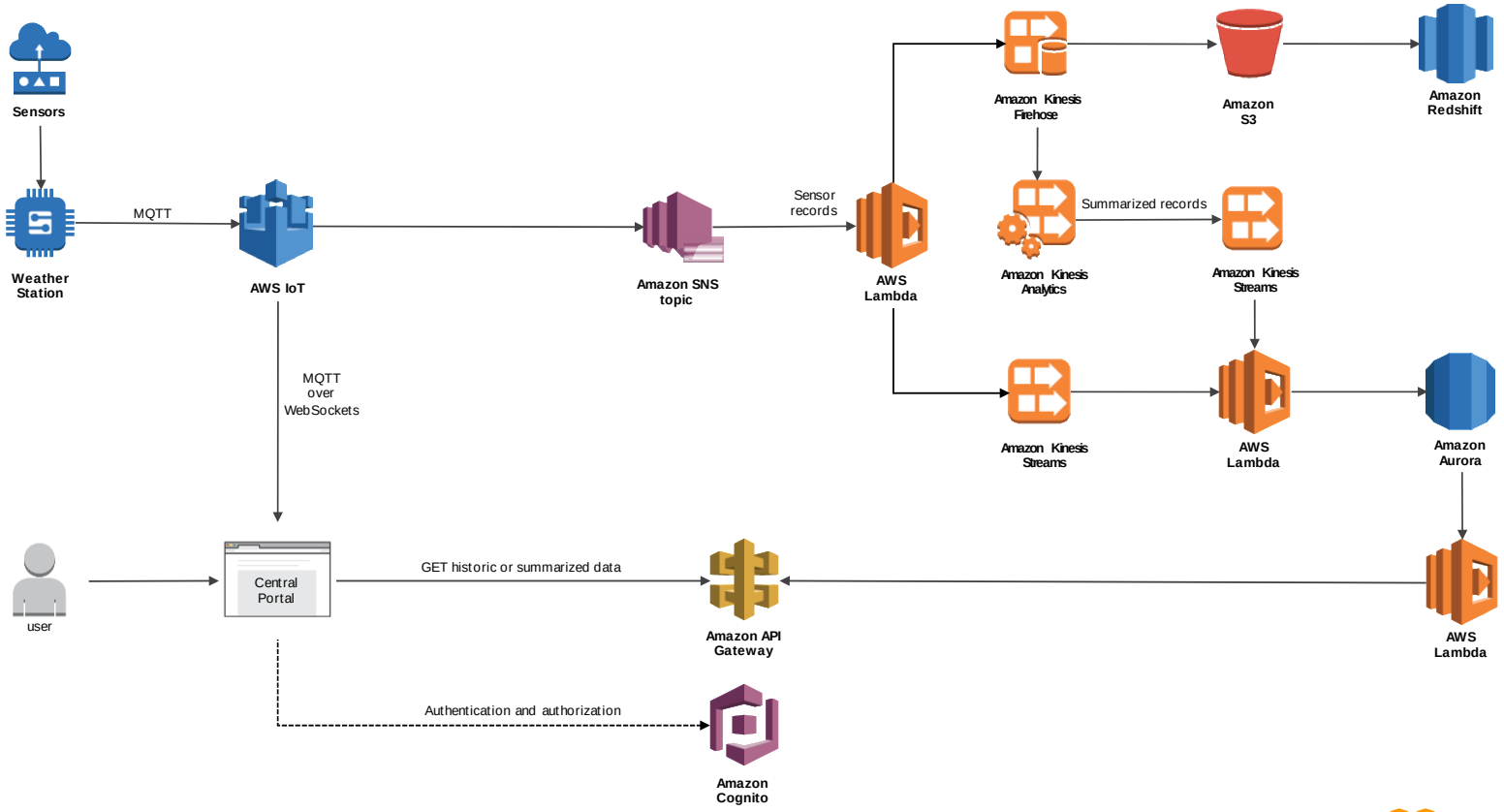
AWS IoT Framework



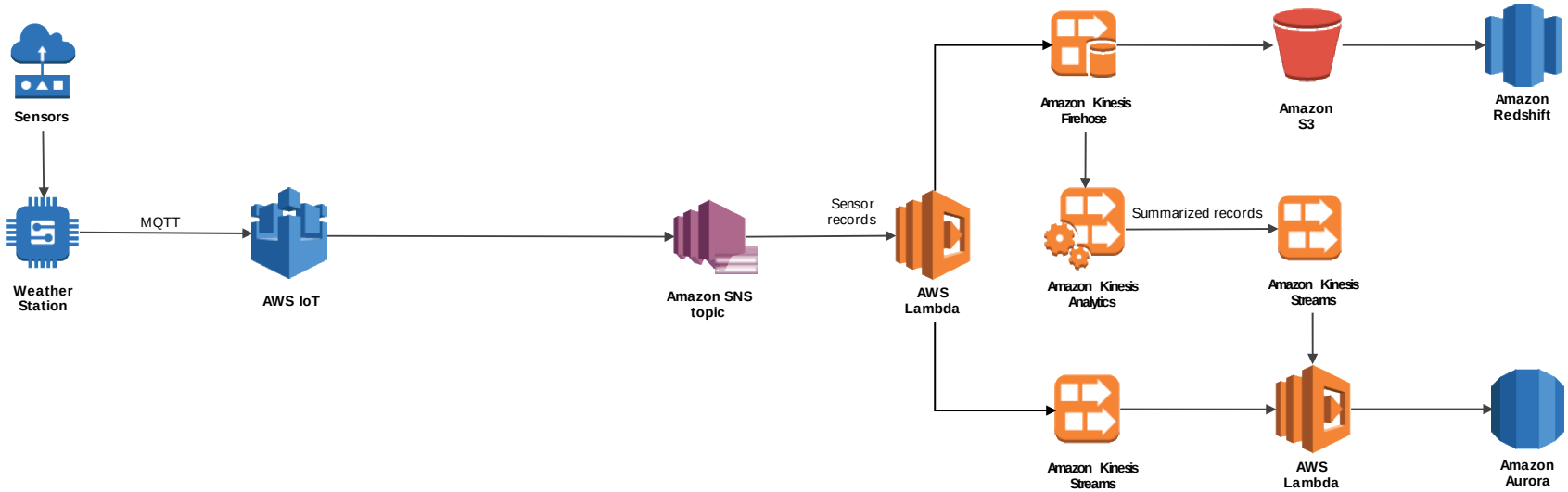
Global Weather Service Architecture

amazon web services
intel

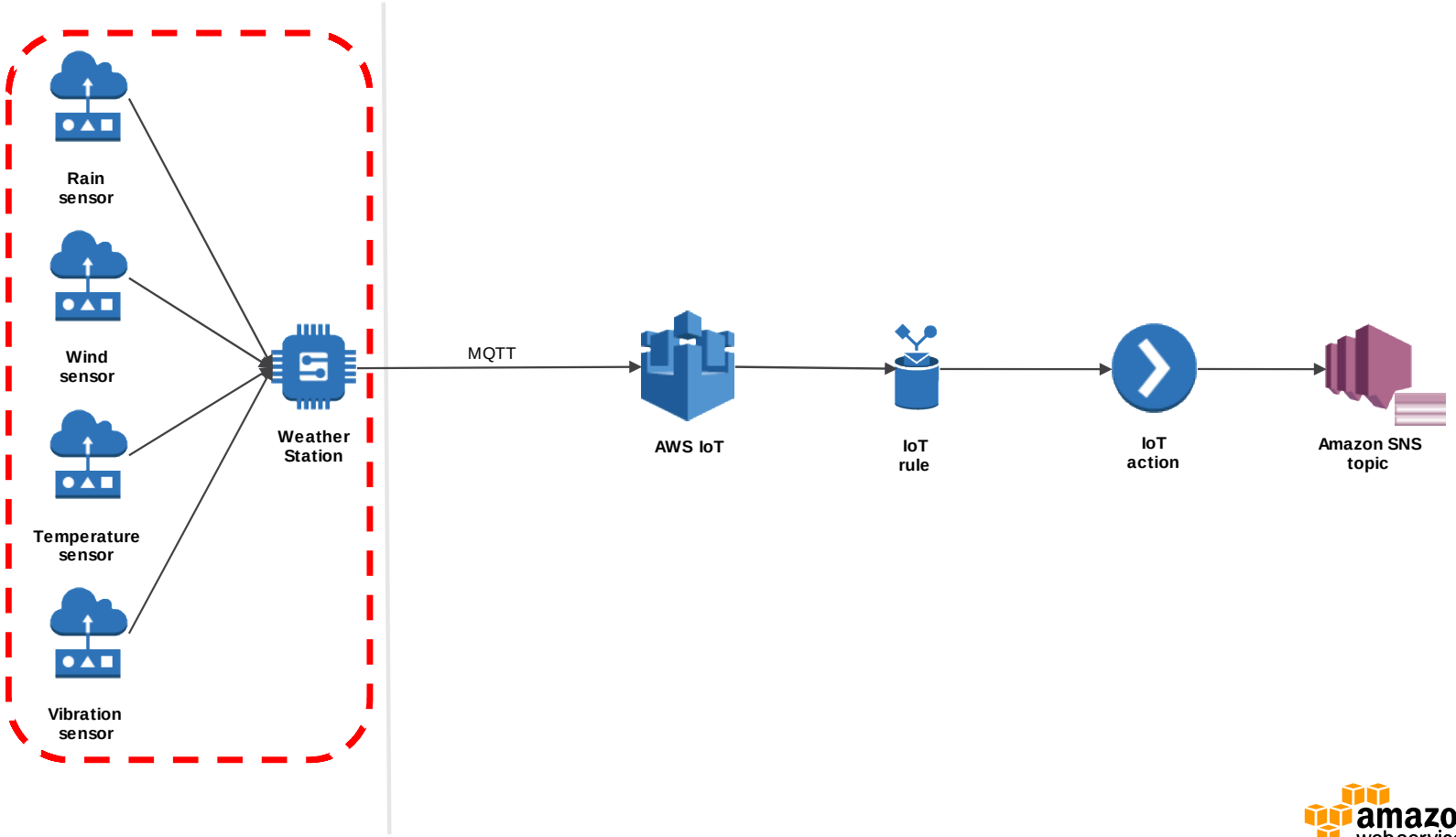
Global Weather Service Architecture



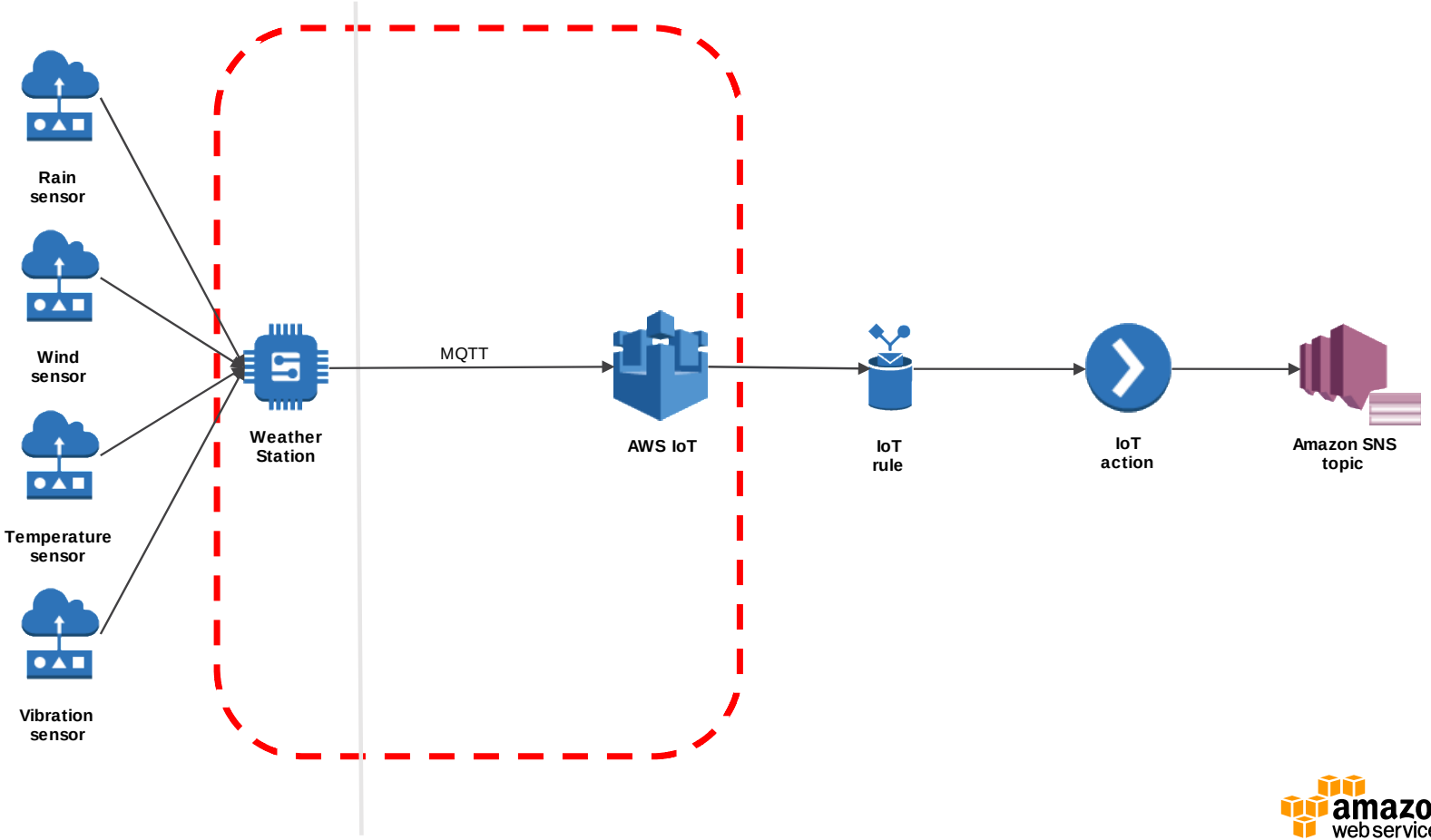
Global Weather Service Architecture



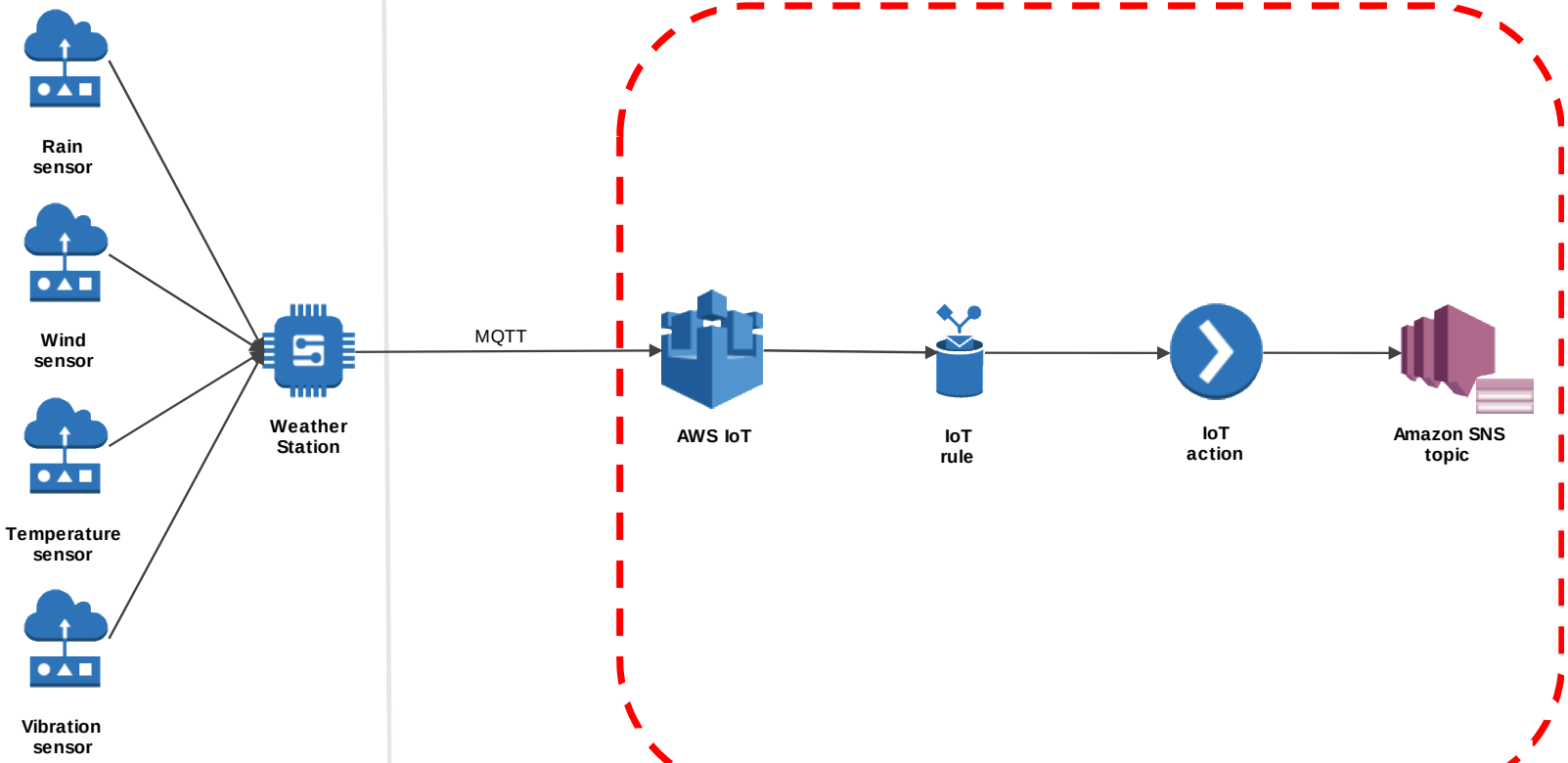
Acquisition Architecture



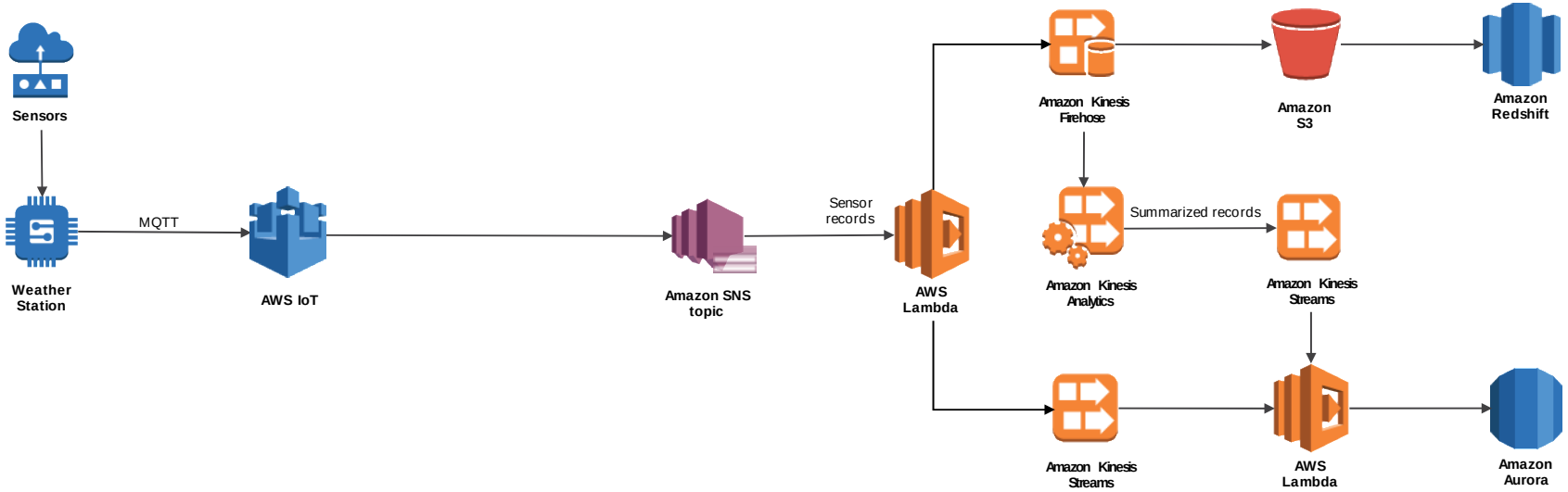
Acquisition Architecture



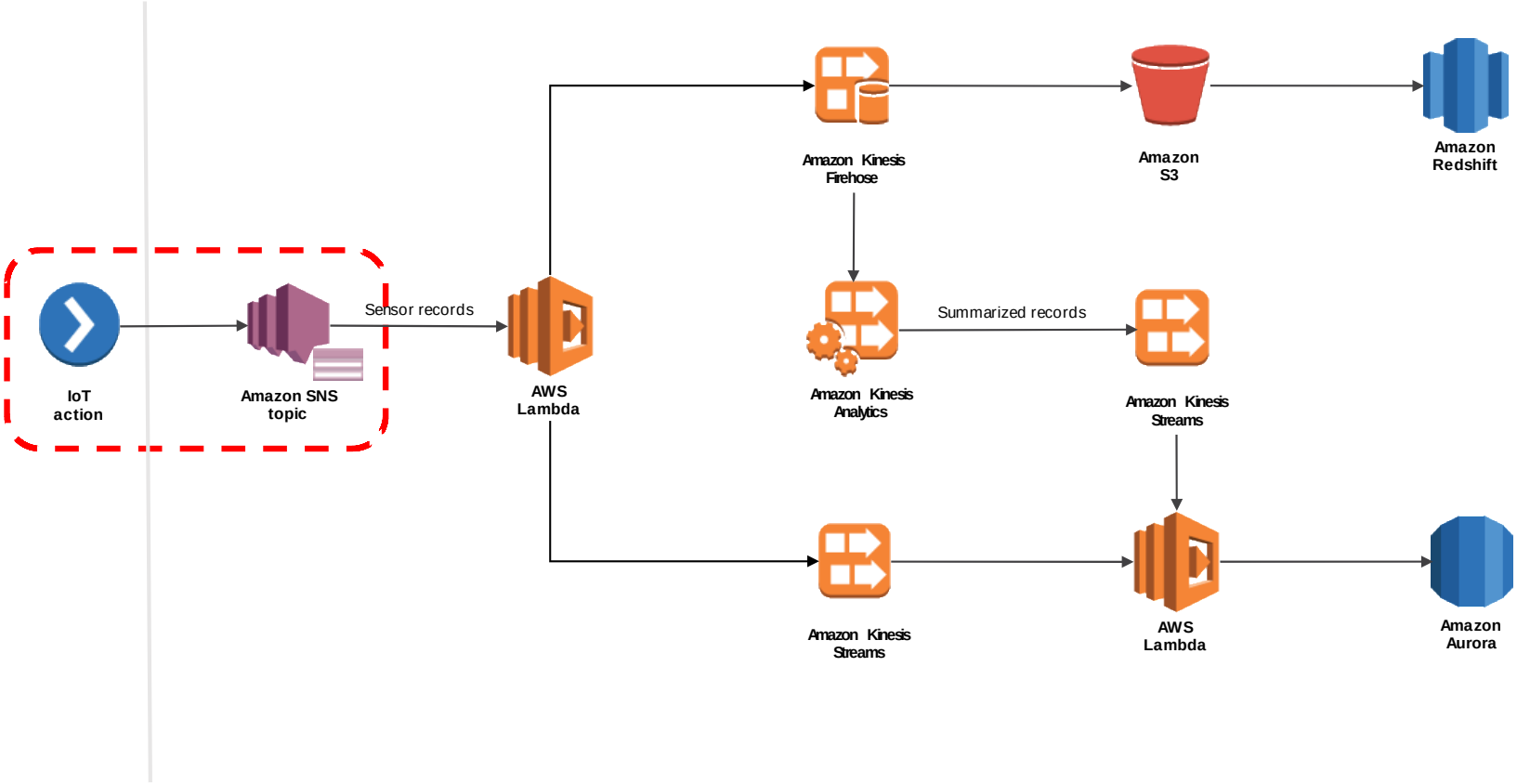
Acquisition Architecture



Global Weather Service Architecture



Processing Architecture



AWS IoT – Rule Setup

Incoming MQTT Topic structure

```
weather/<state>/<city>/<station_id>/<sensor_type>/<sensor_id>
```

AWS IoT – Rule Setup

```
SELECT * FROM
topic(6) AS sensor_id, topic(4) AS station_id,
topic(5) AS sensor, sensor_timestamp,
cast(sensor_value as float) AS sensor_value,
cast(sensor_value_smoothed as float) AS sensor_value_smoothed,
cast(direction as int) AS direction
```

amazon web services
intel

AWS IoT – Rule Setup

```
SELECT * FROM
topic(6) AS sensor_id, topic(4) AS station_id,
topic(5) AS sensor, sensor_timestamp,
cast(sensor_value as float) AS sensor_value,
cast(sensor_value_smoothed as float) AS sensor_value_smoothed,
cast(direction as int) AS direction
```

References the AWS IoT MQTT topic segment

<topic 1>/<topic 2>/…/<topic n>

amazon web services

intel

AWS IoT – Rule Result


```
{
  "value": 0.610802791886758,
  "direction": -1,
  "smoothed": 0.9843152123890655,
  "timestamp": 1472611226005
}
```

Incoming payload

amazon web services | intel

AWS IoT – Rule Result

```
{
  "value": 0.610802791886758,
  "direction": -1,
  "smoothed": 0.9843152123890655,
  "timestamp": 1472611226005
}
```

Incoming payload

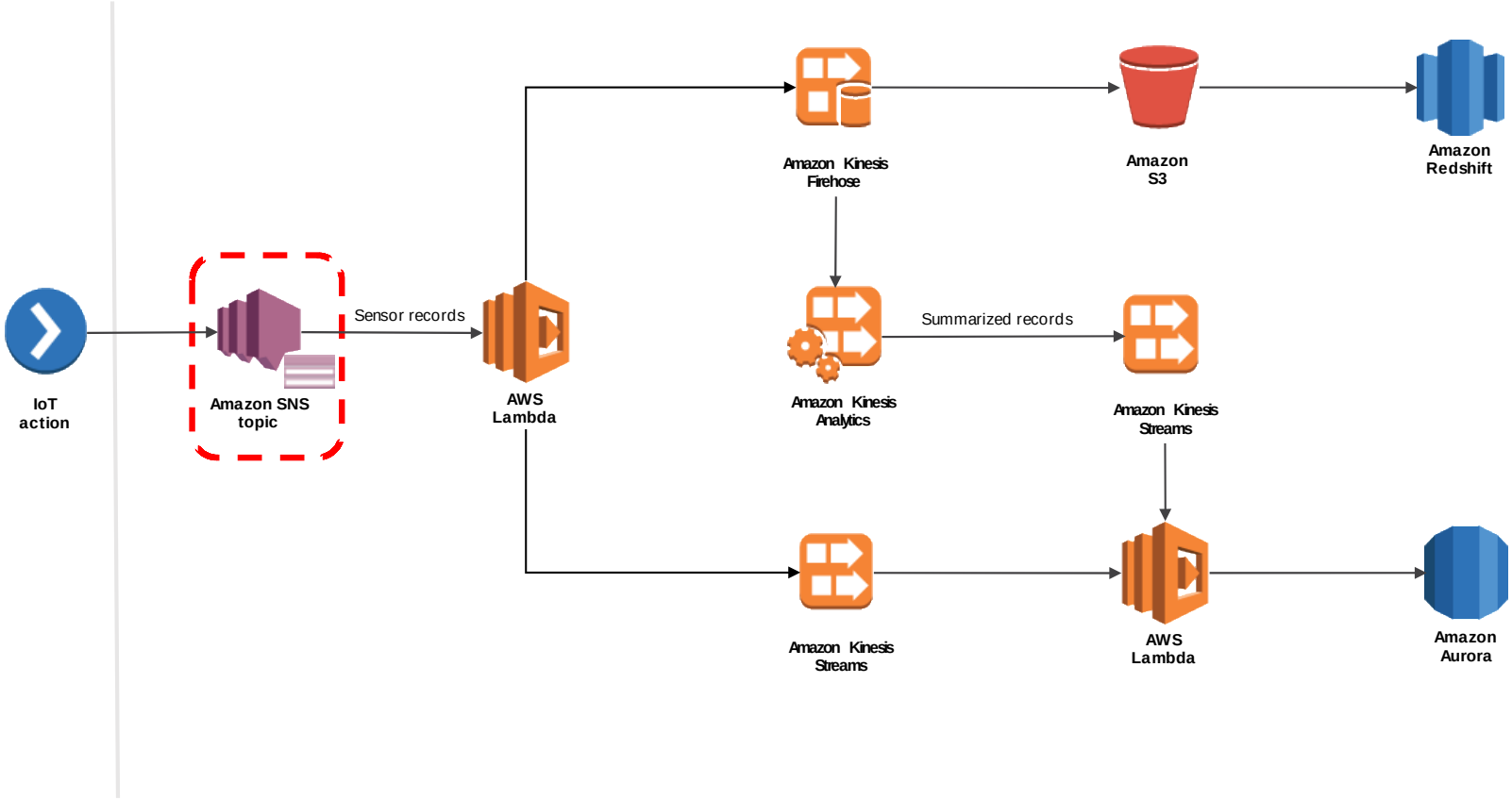
```
{
  "sensor_id": "bQ7KcaMEas",
  "station_id": "vzqHb8vghO",
  "sensor": "vib",
  "timestamp": 1472611226005,
  "value": 0.610802791886758,
  "value_smoothed": 0.9843152123890655,
  "direction": -1
}
```

Transformed payload

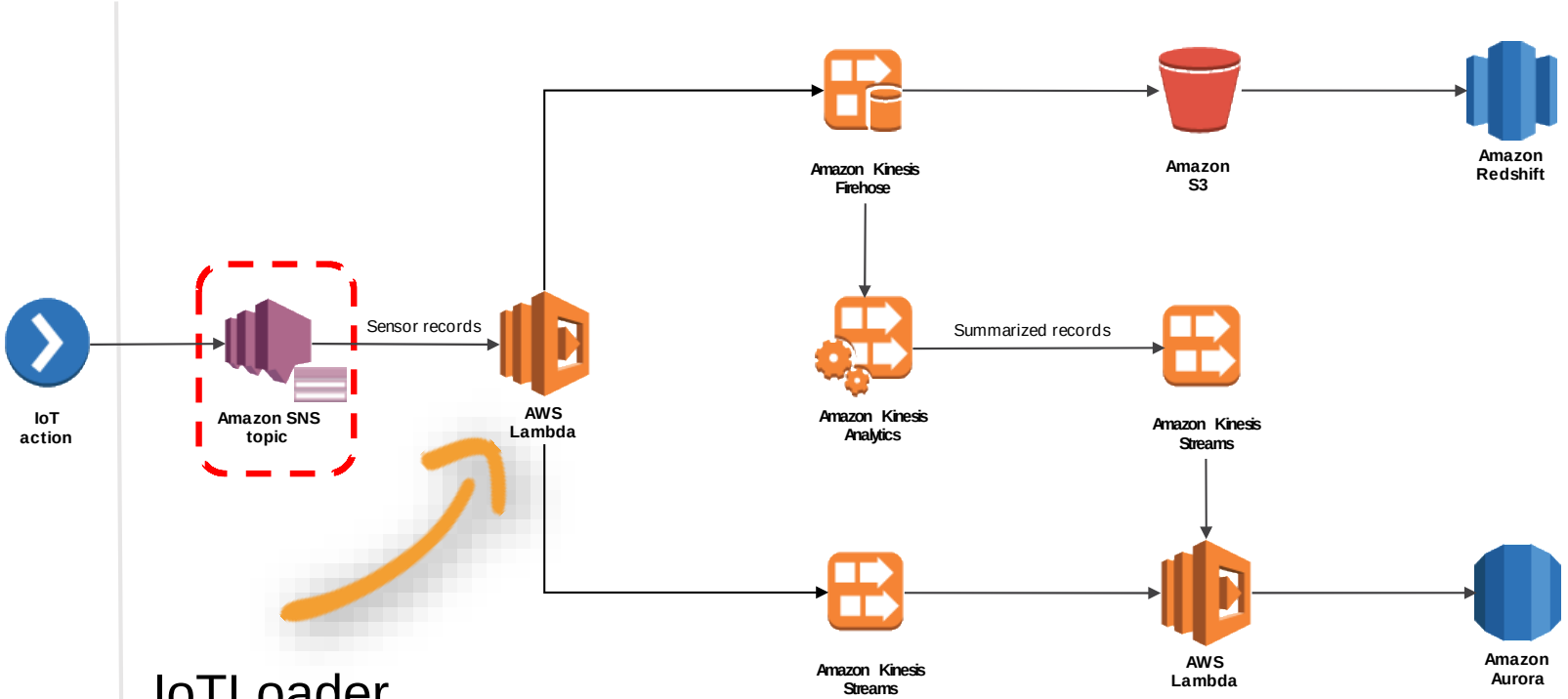
amazon web services

intel

Processing Architecture



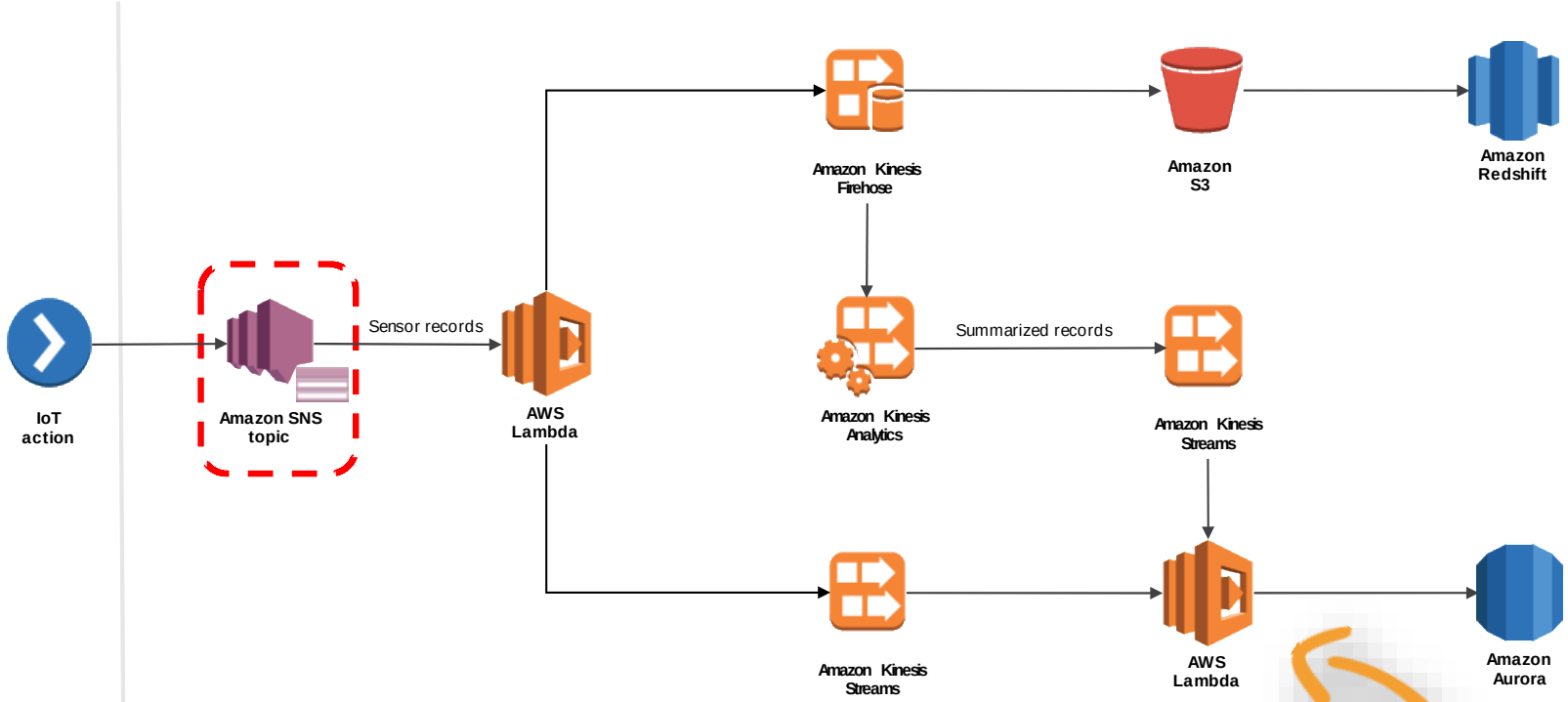
Processing Architecture



IoTLoader

- Process sensor data records from an AWS IoT action and injects them into an Amazon Kinesis stream and Amazon Kinesis Firehose delivery stream.

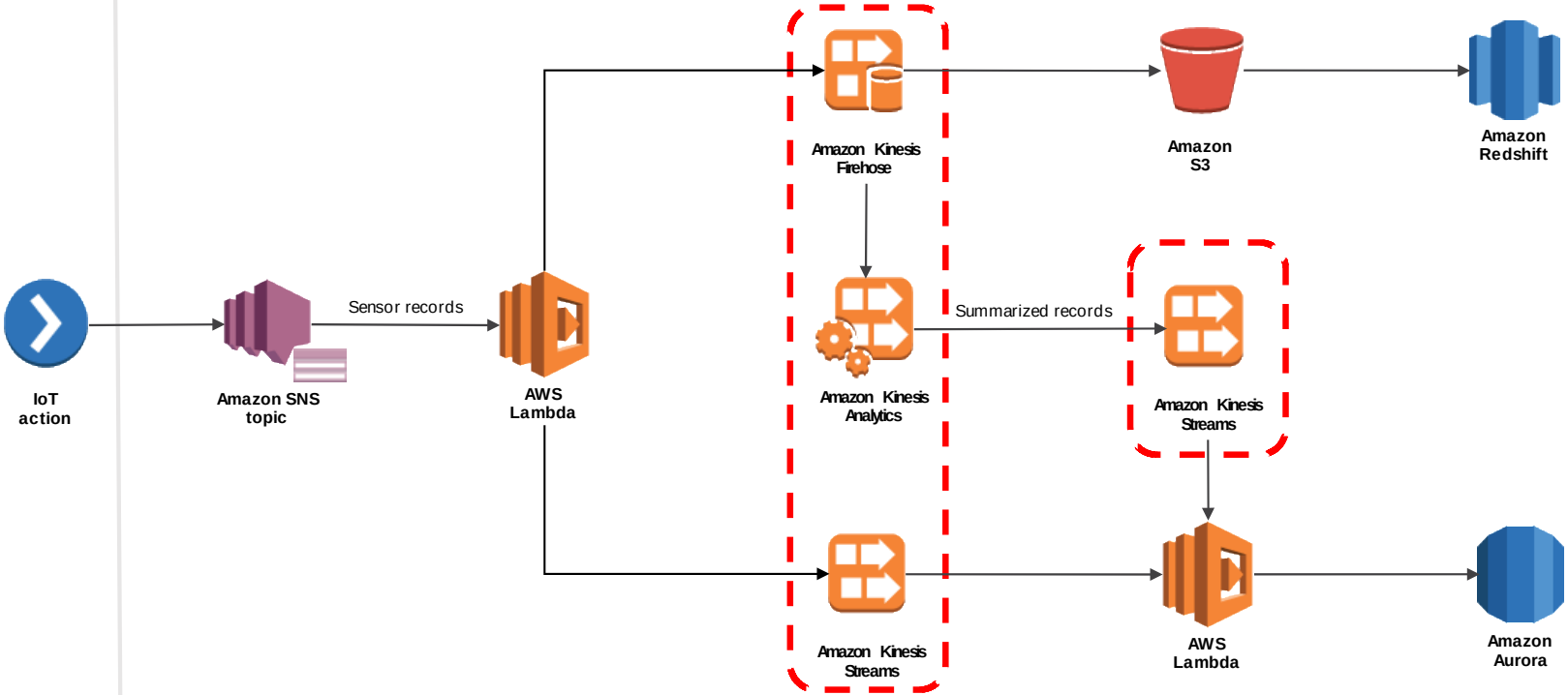
Processing Architecture



RdsLoader

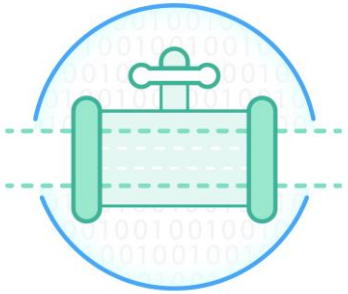
- Process sensor data records from an Amazon Kinesis stream and inserts them into an Amazon Aurora RDS database.

Processing Architecture



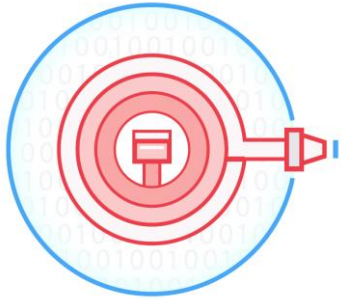
Amazon Kinesis: Streaming Data Made Easy

Services make it easy to capture, deliver, process streams on AWS



Amazon Kinesis Streams

- For technical developers
- Build your own custom applications that process or analyze streaming data



Amazon Kinesis Firehose

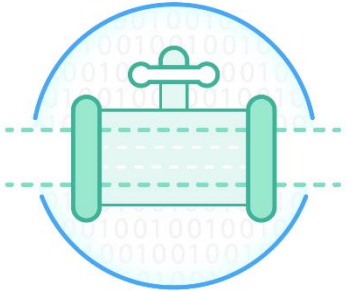
- For ETL, data engineer
- Easily load massive volumes of streaming data into S3, Amazon Redshift and Amazon Elasticsearch Service

Amazon Kinesis Analytics

- For all developers, data scientists
- Easily analyze data streams using standard SQL queries

amazon web services | intel

Amazon Kinesis - Streaming Data Made Easy

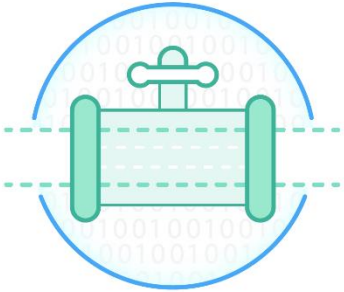


Amazon Kinesis Streams

Low latency streaming ingest at scale

amazon web services
intel

Amazon Kinesis - Streaming Data Made Easy



Amazon Kinesis Streams

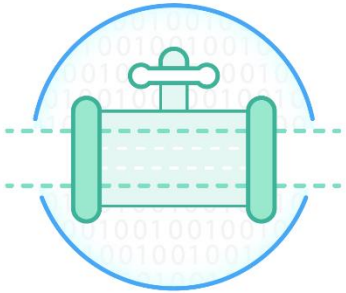
Low latency streaming ingest at scale

Amazon Kinesis Analytics

Streaming analytics in near real-time

amazon web services
intel

Amazon Kinesis - Streaming Data Made Easy

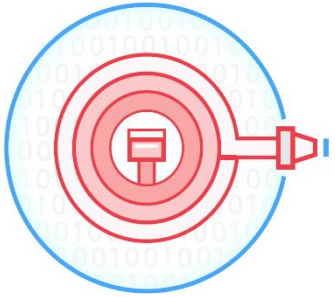


Amazon Kinesis Streams

Low latency streaming ingest at scale

Amazon Kinesis Analytics

Streaming analytics in near real-time



Amazon Kinesis Firehose

Batch data delivery based on time/size into S3

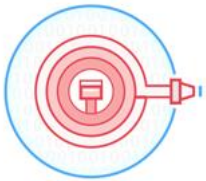
amazon web services

intel

Amazon Kinesis Firehose vs. Amazon Kinesis Streams

Amazon Kinesis Streams

Amazon Kinesis Streams is for use cases that require **custom processing**, per incoming record, with sub-1 second processing latency, and a choice of stream processing frameworks.



Amazon Kinesis Firehose

Amazon Kinesis Firehose is for use cases that require zero administration, ability to **use existing analytics tools based on Amazon S3, Amazon Redshift, and Amazon Elasticsearch Service** and a data latency of 60 seconds or higher.

amazon web services | intel

Use SQL To Build Real-Time Applications

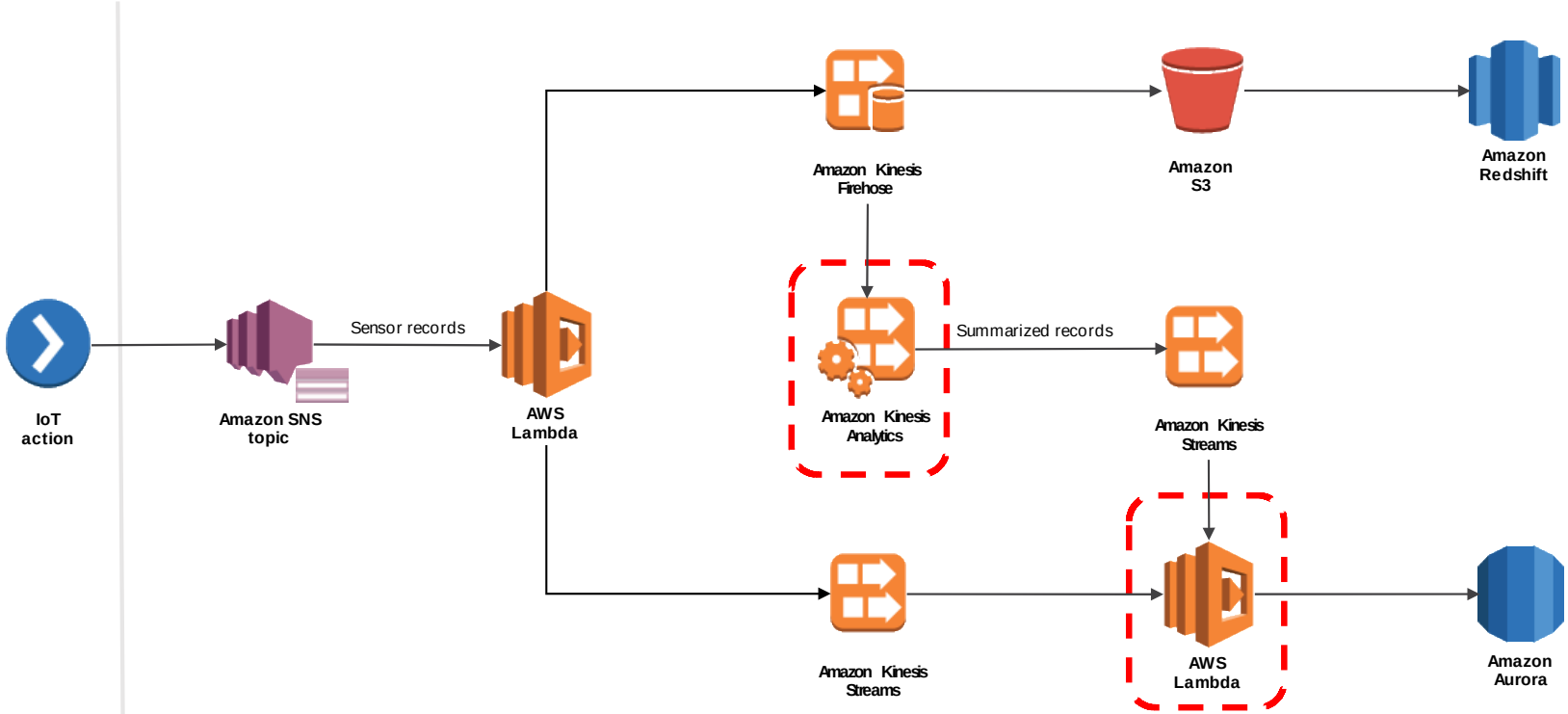
Connect to streaming source

Easily write SQL code to process streaming data

Continuously deliver SQL results

amazon web services | intel

Processing Architecture



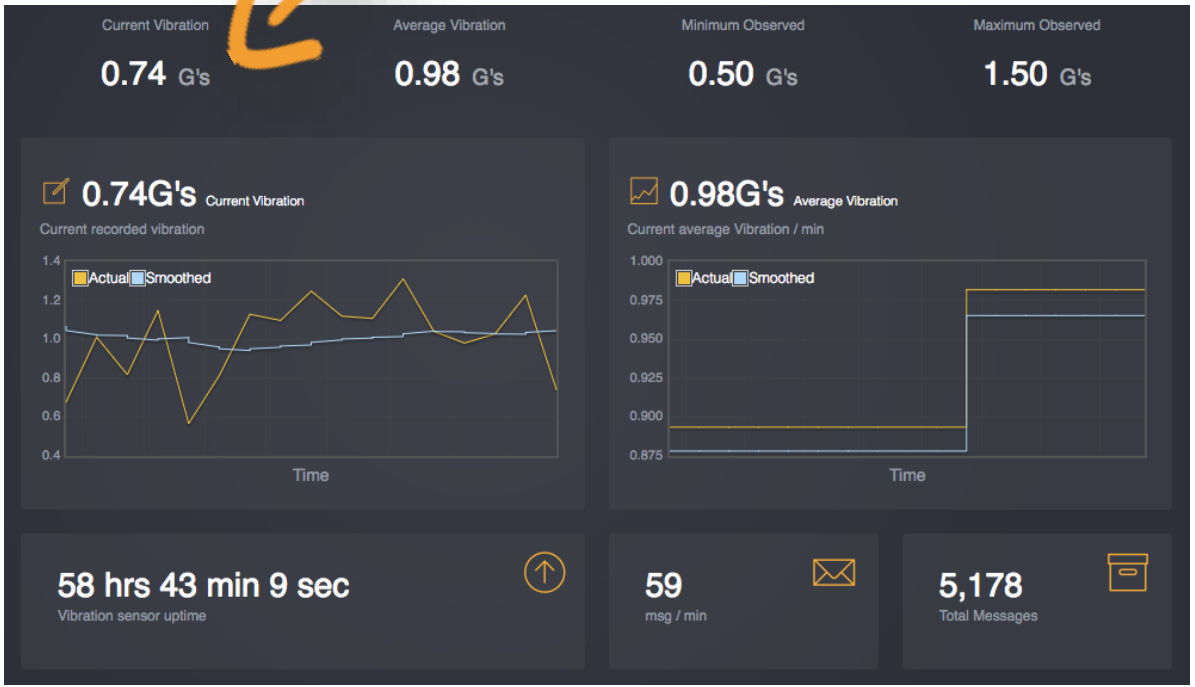
Amazon Kinesis Analytics – Answering Questions

amazon web services

intel

Amazon Kinesis Analytics – Answering Questions

What is the current value ?



amazon web services
intel

Amazon Kinesis Analytics – Answering Questions

What is the average value ?

Amazon Kinesis Analytics – Answering Questions

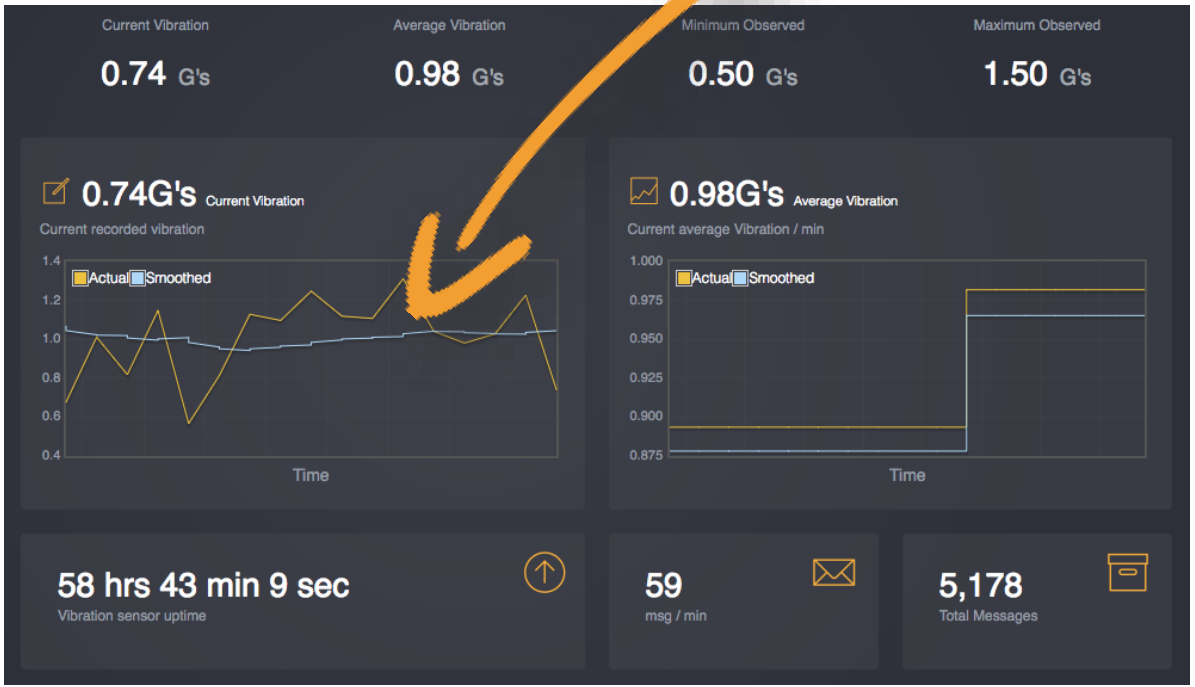
What is the minimum value ?

Amazon Kinesis Analytics – Answering Questions

What is the maximum value ?

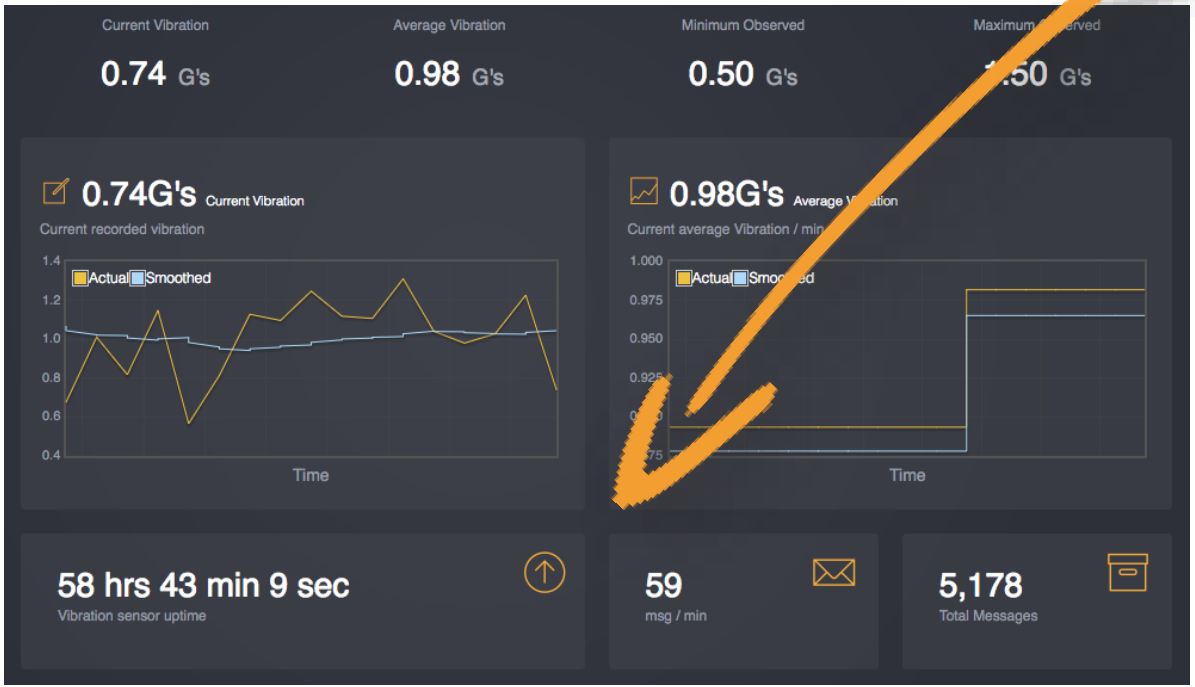
Amazon Kinesis Analytics – Answering Questions

Visual graphs for short term trending



Amazon Kinesis Analytics – Answering Questions

Service performance statistics



Amazon Kinesis Analytics – Processing Setup

```
1  CREATE OR REPLACE STREAM "DESTINATION_SQL_STREAM" ("sensor_id" VARCHAR(32), "sensor" VARCHAR(15),
2      "station_id" VARCHAR(32), "sensor_avg_value" double, "sensor_smooth_avg_value" double,
3      "60sec_sum_of_sensor_value" double, "60sec_number_of_msg" int, "record_timestamp" TIMESTAMP);
4
5  CREATE OR REPLACE PUMP "STREAM_PUMP" AS INSERT INTO "DESTINATION_SQL_STREAM"
6  SELECT STREAM "sensor_id", "sensor", "station_id",
7      AVG("sensor_value"), AVG("sensor_value_smoothed"),
8      SUM("sensor_value") AS "60sec_sum_of_sensor_value",
9      count(*) AS "60sec_number_of_msg", ROWTIME AS "record_timestamp"
10 FROM "SOURCE_SQL_STREAM_001"
11 GROUP BY "sensor_id","sensor","station_id",
12     FLOOR(("SOURCE_SQL_STREAM_001".ROWTIME - TIMESTAMP '1970-01-01 00:00:00') SECOND / 60 TO SECOND);
```

Cancel

Source data | Real-time analytics | Destination

Applicati

In-application streams:

DESTINATION_SQL_STREAM

error_stream

Pause results New results will be added every 2-10 seconds

Scroll to bottom when new results arrive.

Filter by column name

ROWTIME	sensor_id	sensor	station_id	sensor_avg_value	sensor_sm
2016-11-11 01:44:00.0	813c174428fc4843	temp	qwbKAMlbZW	63.23000000000001	63.230000
2016-11-11 01:44:00.0	caaed84a552deebf	rain	qwbKAMlbZW	0.0	0.0

Web Services

intel

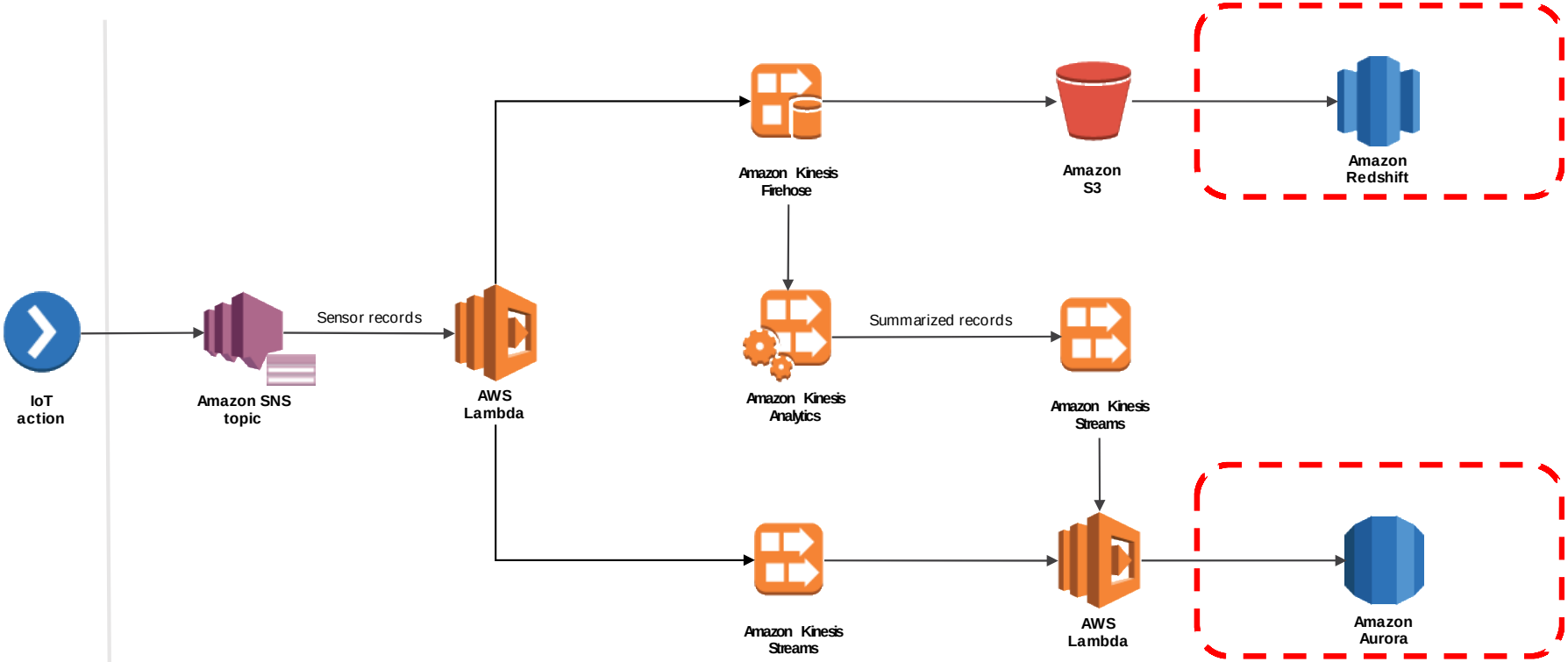
Amazon Kinesis Analytics – Processing Result

Emitted payload


```
{
  "sensor_id": "dc2b8383eb79fe49",
  "sensor": "vib",
  "station_id": "qwbKAMlbZW",
  "sensor_avg_value": 1.072153418386984,
  "sensor_smooth_avg_value": 1.015843804467917 2,
  "60sec_sum_of_sensor_value": 64.32920510321904,
  "60sec_number_of_msg": 60,
  "record_timestamp": "2016-11-09 06:29:00.0"
}
```


amazon web services | intel

Processing Architecture



Data Store Summary

Amazon S3

- Raw long term storage for warm data
- Lifecycle management
- Reprocess and reload data

amazon web services
intel

Data Store Summary

Amazon S3

- Raw long term storage for warm data
- Lifecycle management
- Reprocess and reload data

Amazon Redshift

- Optimized for data warehousing and analytics
- Query large amounts of data fast
- Scale to increase performance

amazon web services
intel

Data Store Summary

Amazon S3

- Raw long term storage for warm data
- Lifecycle management
- Reprocess and reload data

Amazon Redshift

- Optimized for data warehousing and analytics
- Query large amounts of data fast
- Scale to increase performance

Amazon Aurora

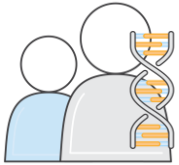
- Optimized for distributed data access
- Scale read throughput
- Fault tolerant

amazon web services | intel

Summary

IoT with real-time analytics provides meaningful information, not just data

Scale without intervention or cost



Remove management and scaling overhead to accelerate innovation

amazon web services | intel

New to AWS

Introductory labs and videos can help you ramp up

Start learning

Online Labs

Take an online Self-Paced Lab to get hands-on-practice with AWS services

Start practicing

Take a Class

Build technical skills and learn best practices from an accredited instructor

Find a class

AWS Certification

Validate knowledge and show expertise with industry recognized certifications

Get Certified

Learn more: aws.amazon.com/training

amazon web services | intel

Thank You for Attending AWS Innovate

We hope you found it interesting! A kind reminder to **complete the survey.** Let us know what you thought of today's event and how we can improve the event experience for you in the future.

aws-apac-marketing@amazon.com

twitter.com/AWSCloud

facbook.com/AmazonWebServices

youtube.com/user/AmazonWebServices

slideshare.net/AmazonWebServices

twitch.tv/aws

amazon web services | intel