
AWS
re:Invent

Paving the way toward automated driving with BMW Group

Marc Neumann

Product Owner Data Engineering
Automated Driving, BMW AG

Junjie Tang

Principal Consultant, AWS

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Paving the way toward automated driving with BMW

- Customer vehicle fleet analytics at BMW Group
- Autonomous driving data lake



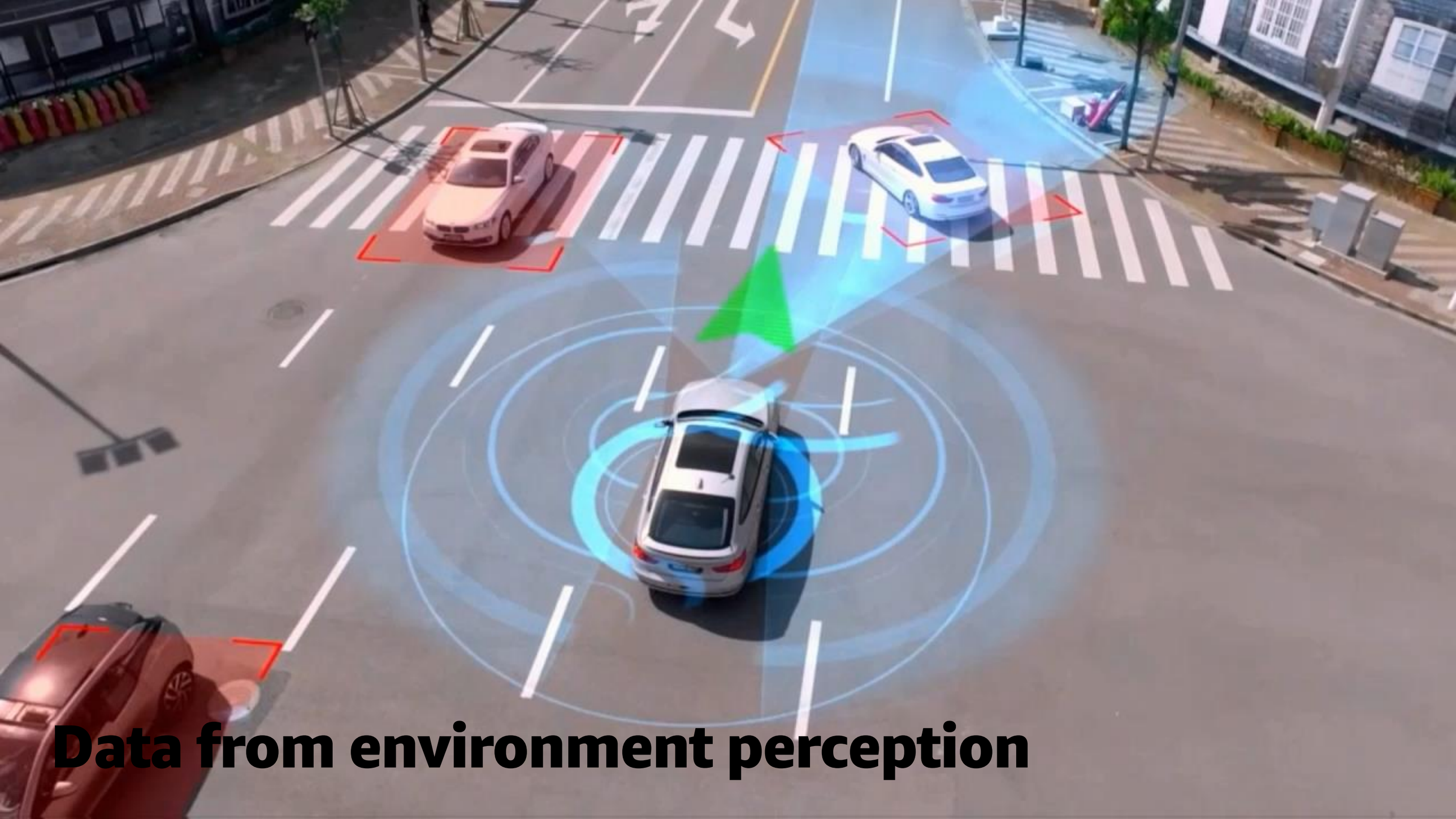
AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Customer vehicle fleet analytics at BMW group

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.


Data from environment perception

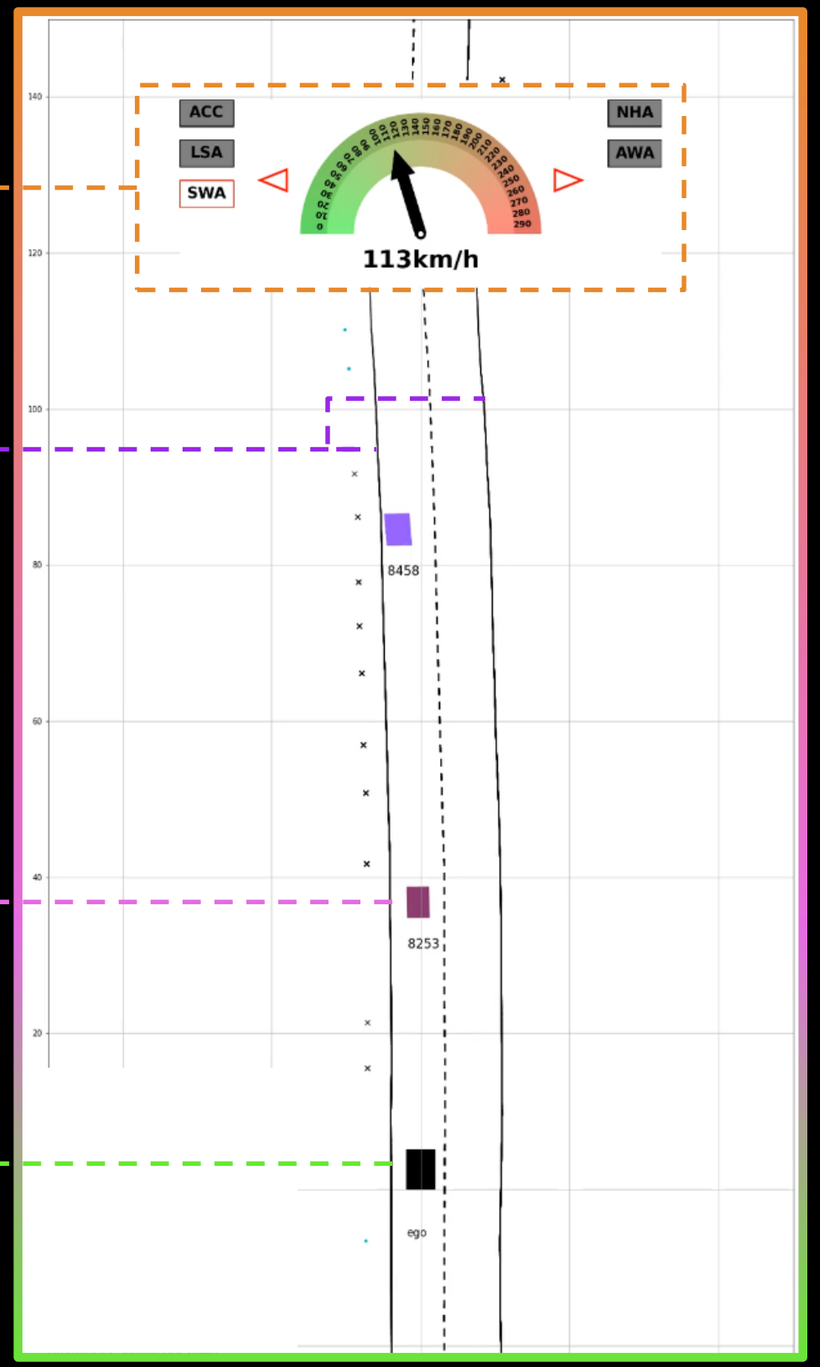
Data we collect

ADAS functions

Lane markings

Detected vehicle attributes

Customer vehicle dynamics



Data privacy and security compliant, anonymized fleet data

Motivation

Validation of automated driving safety and performance

Vehicle function development

Product cost reduction

Customer orientation

Quality monitoring

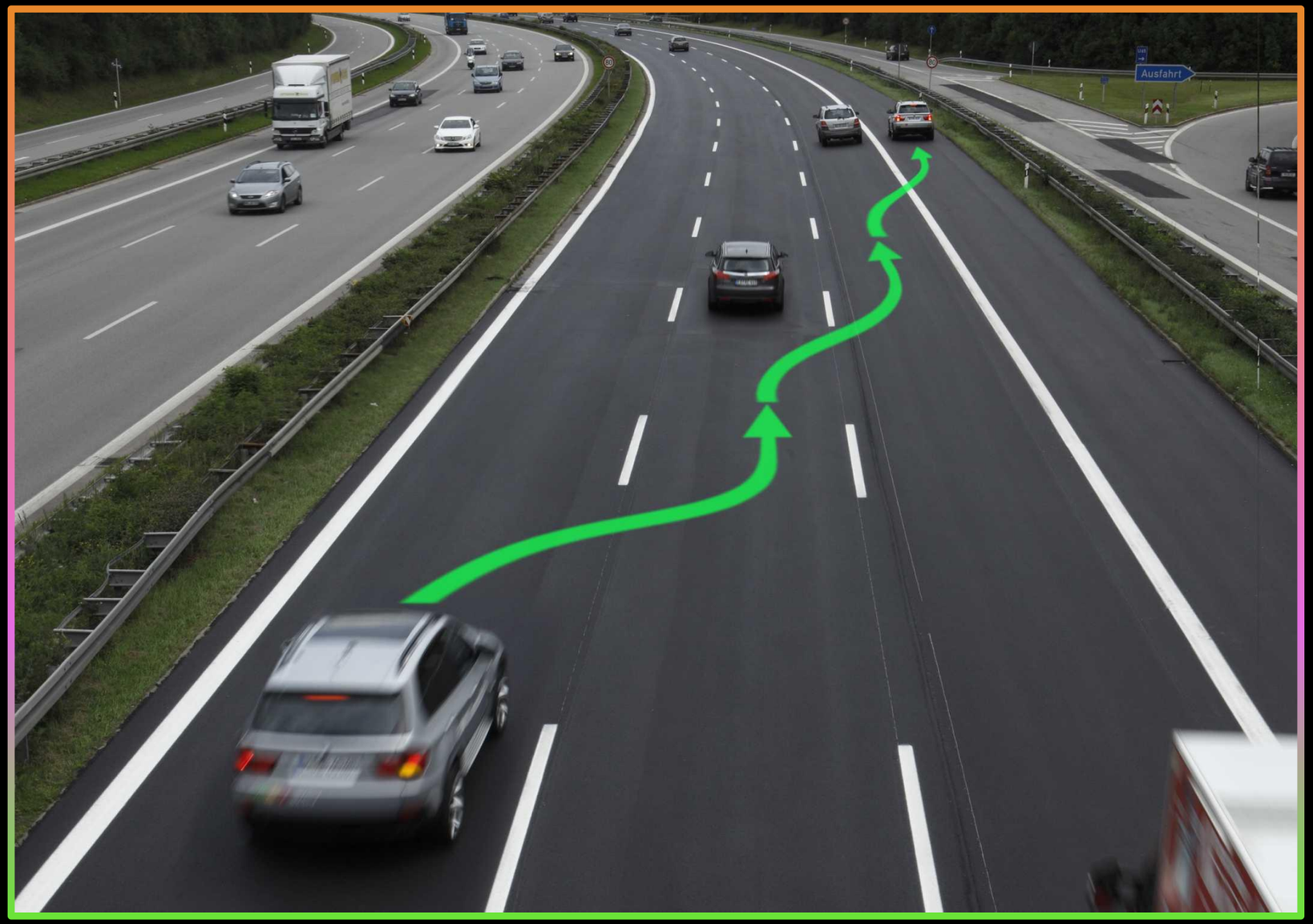


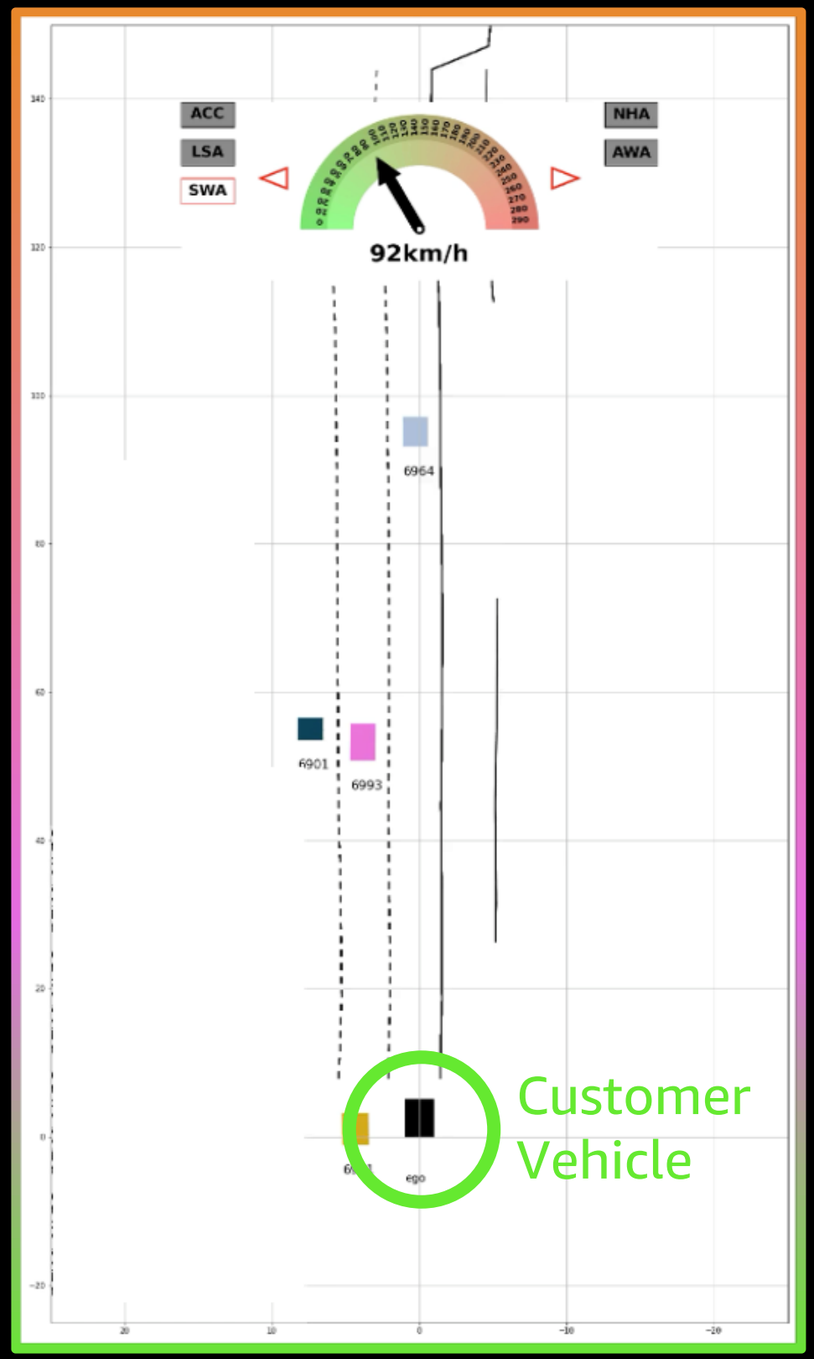
DATA FROM CUSTOMER FLEET

Data volume > 1 billion kilometers

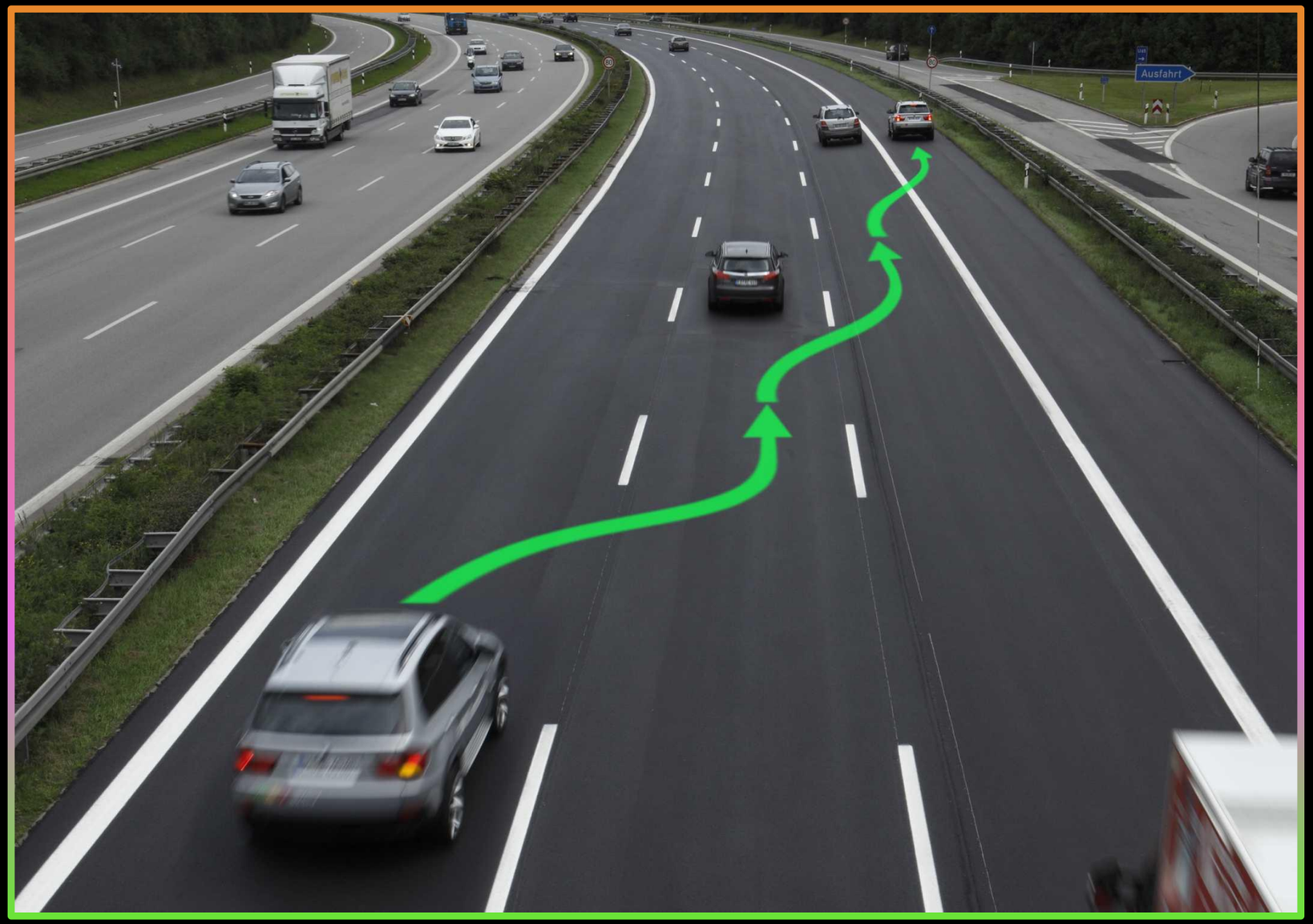


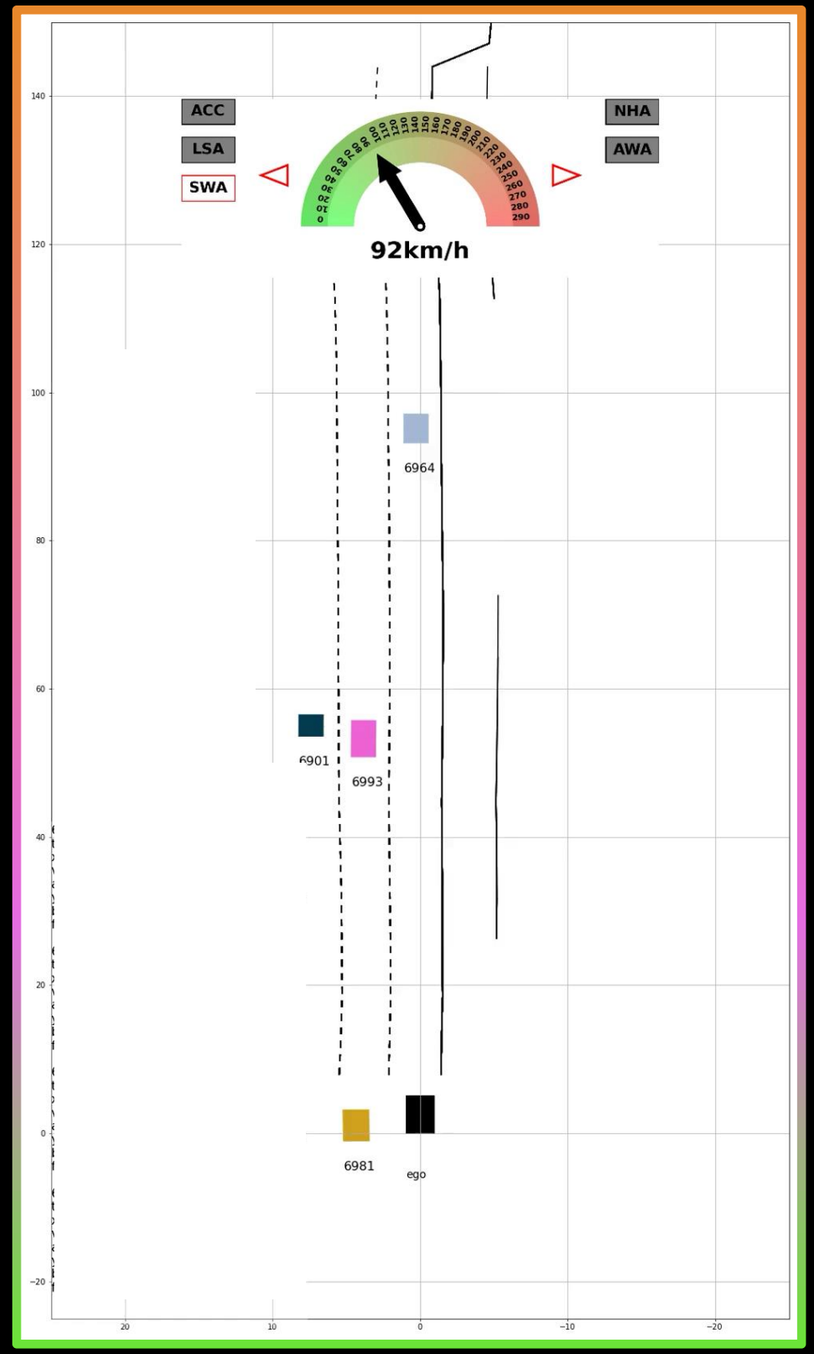
Collect safety critical situations





Collect safety critical situations

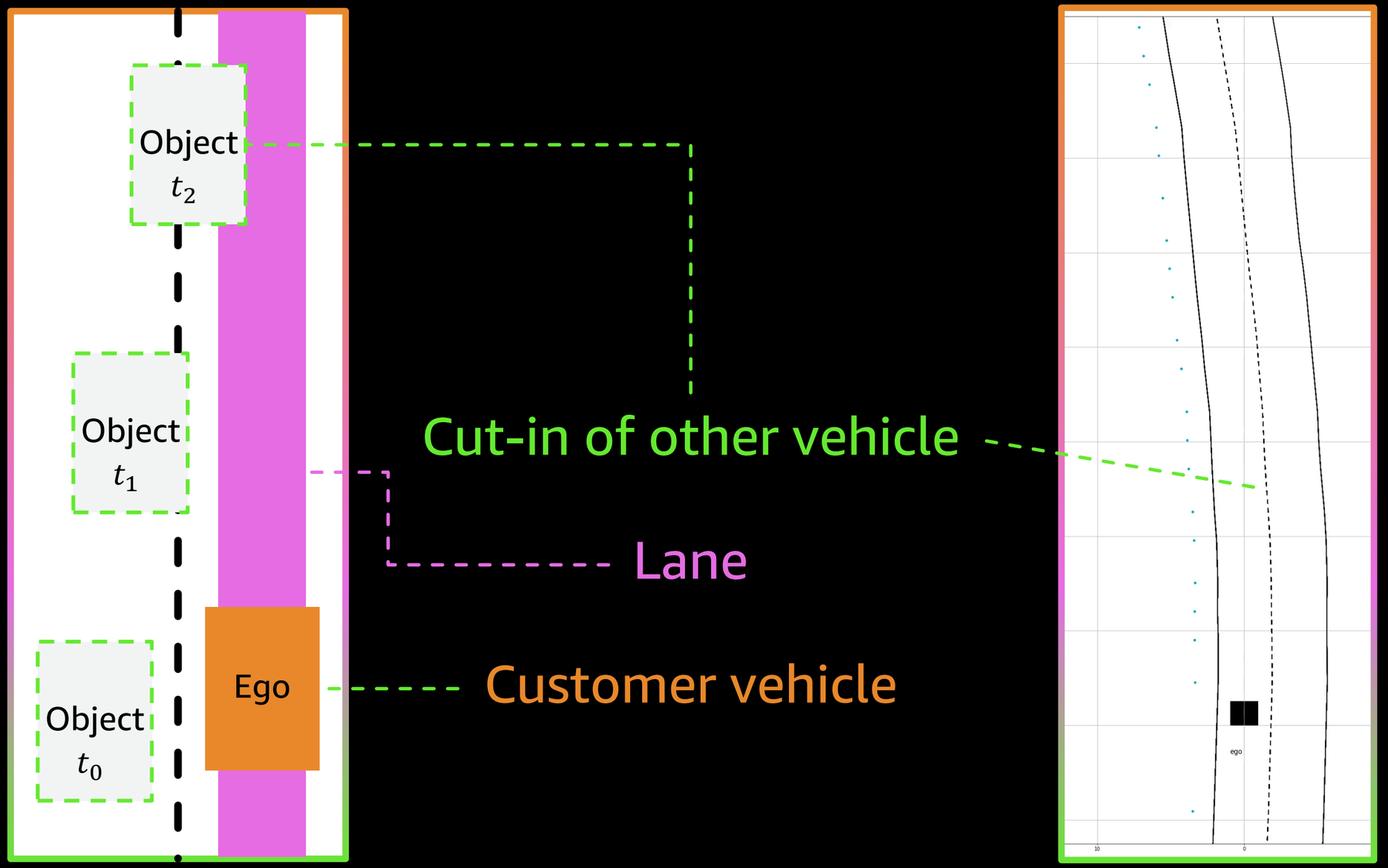




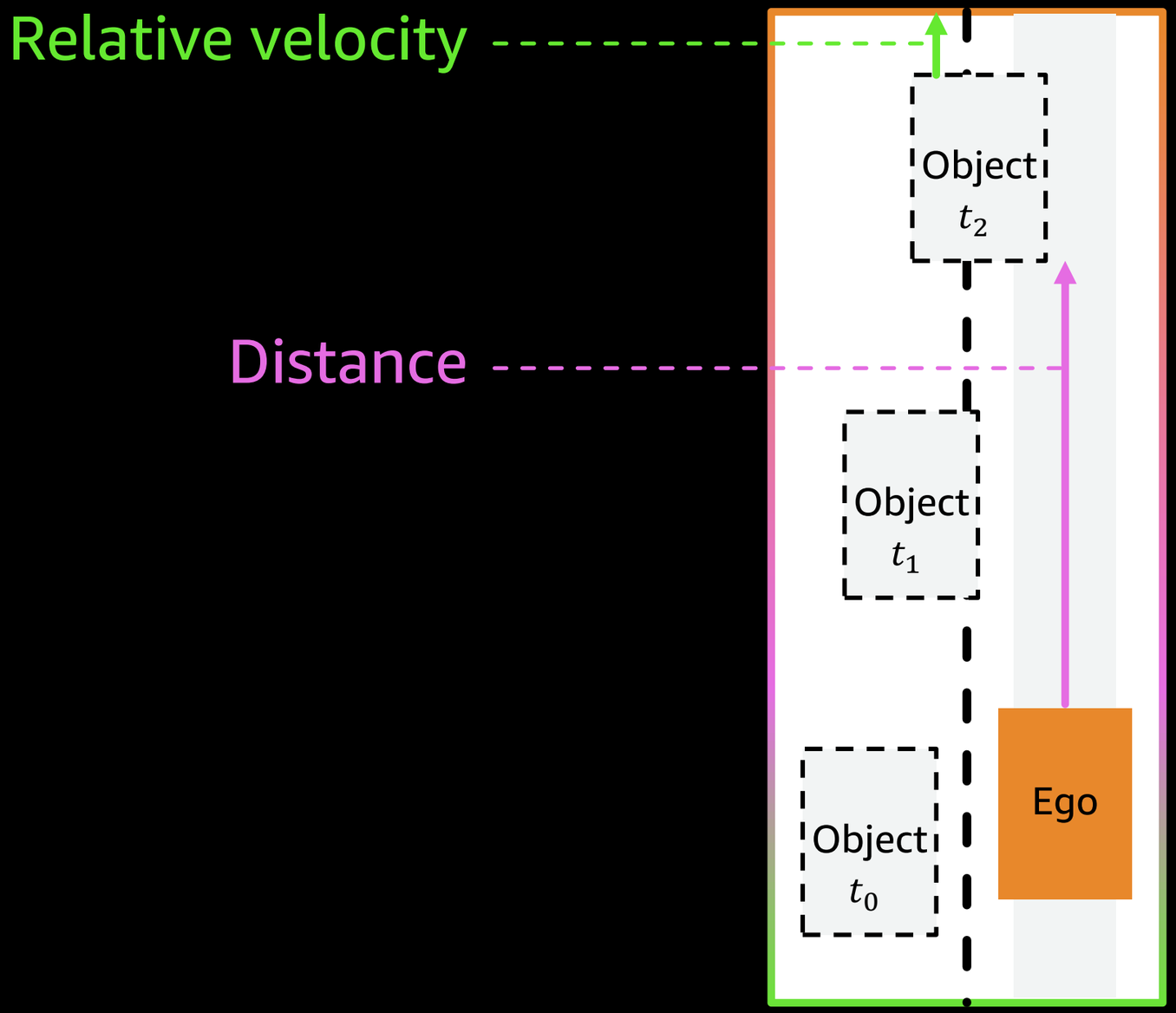
Cut-in analysis example

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cut-in


Object
t_2
Object
t_1
Object
t_0
Ego
Cut-in of other vehicle
Lane
Customer vehicle
ego
10
0

Parameters of a cut-in

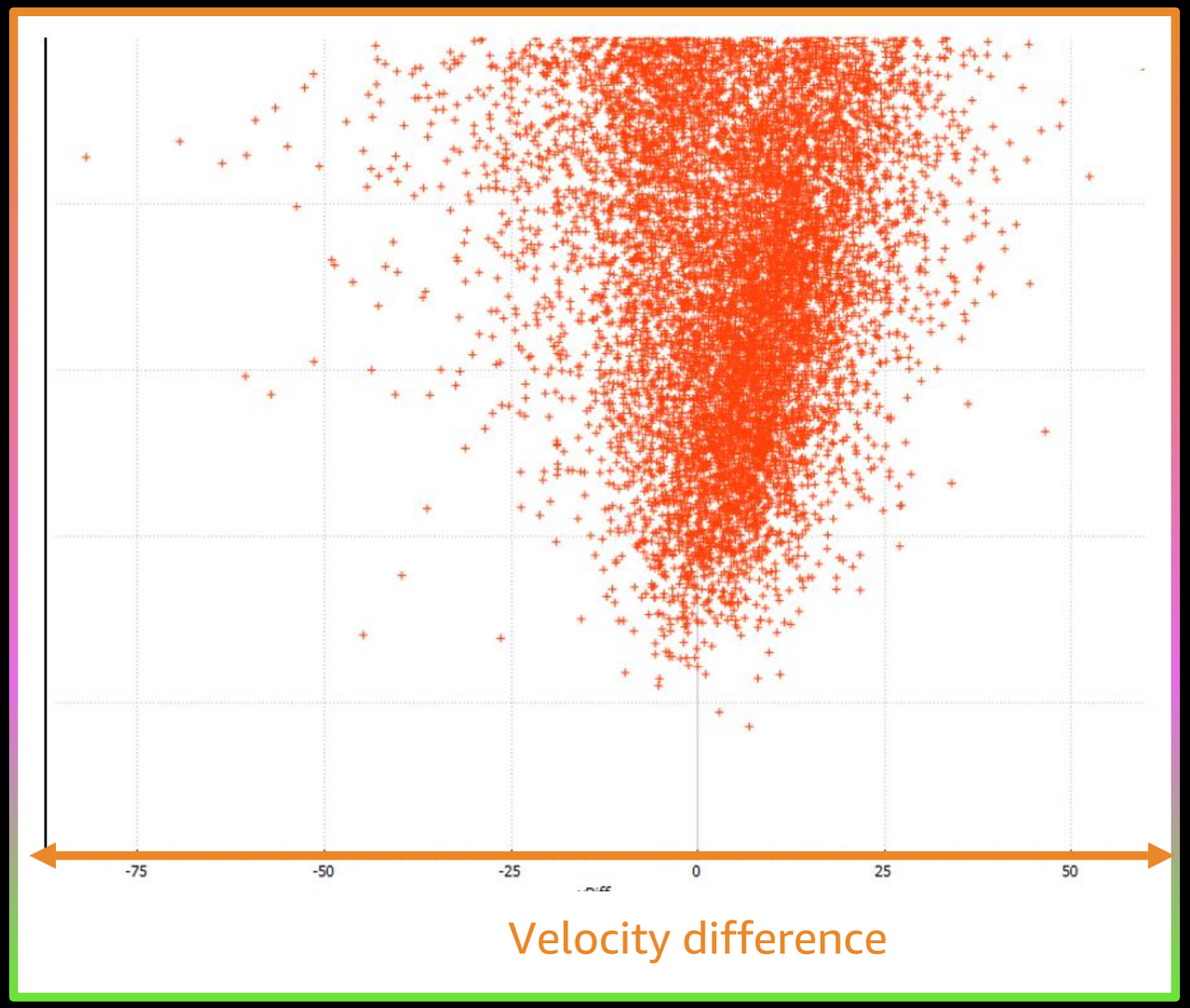


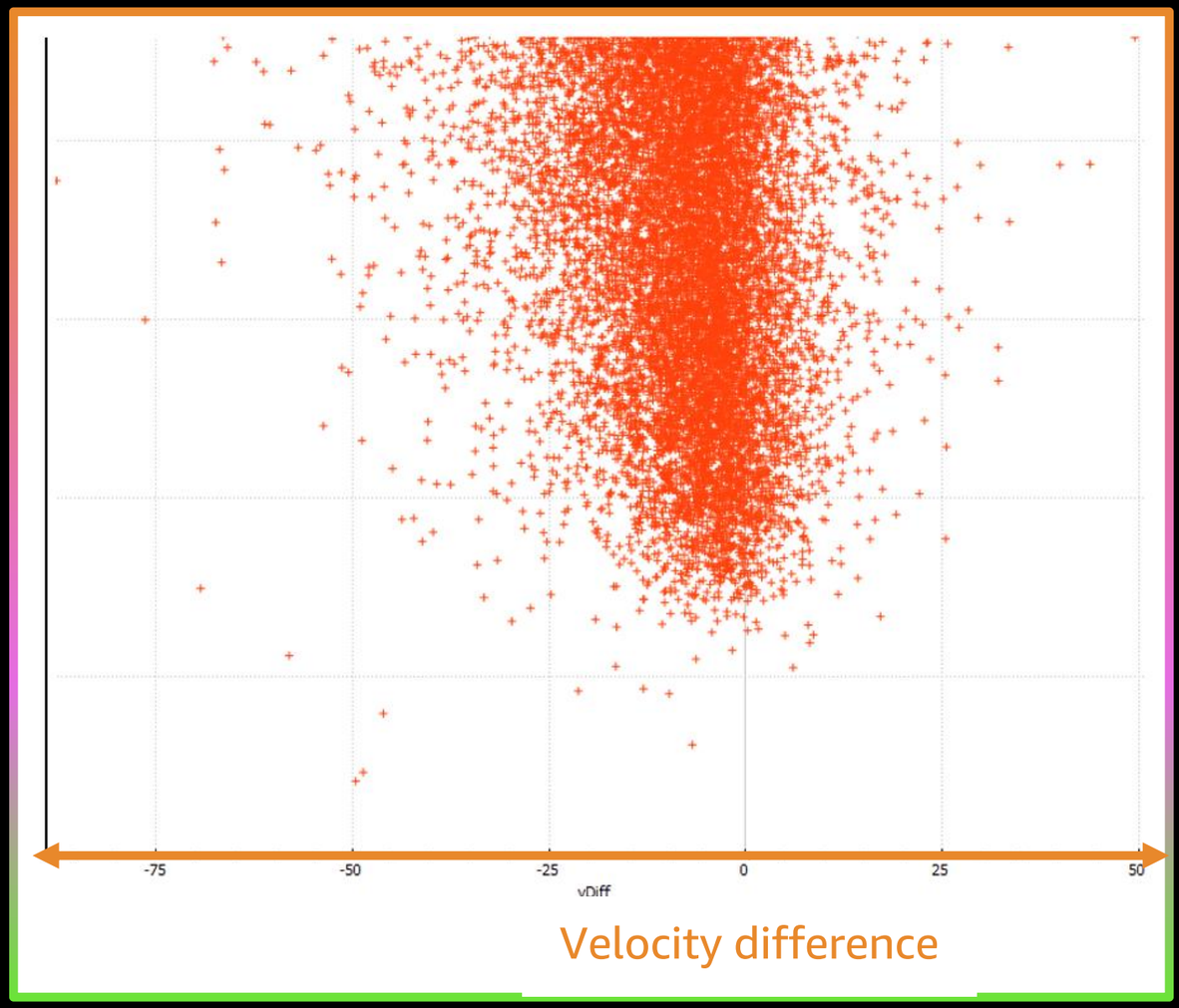
Sample distribution of cut-ins

From left lane

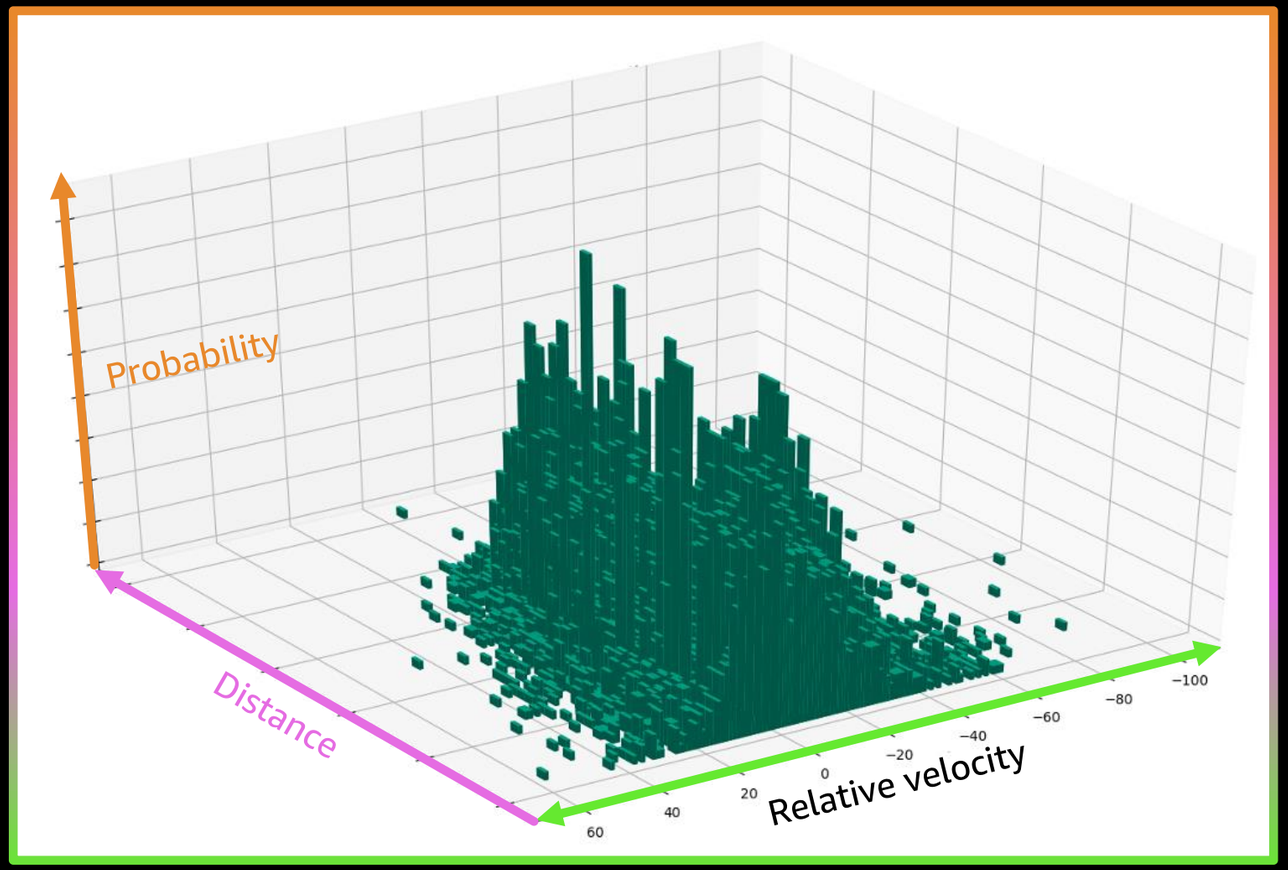
Distance

From right lane





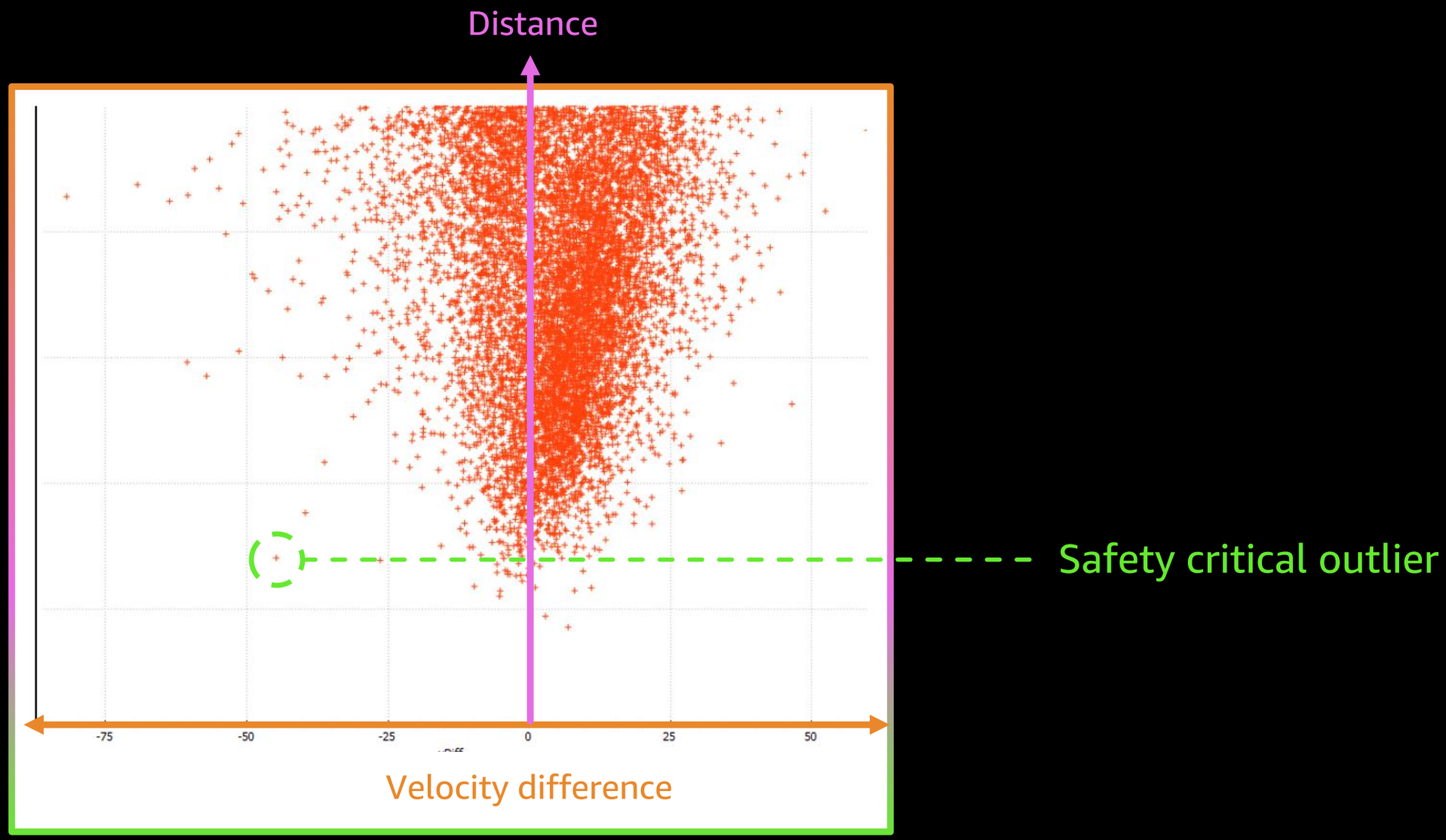
Distribution modeling



Generate

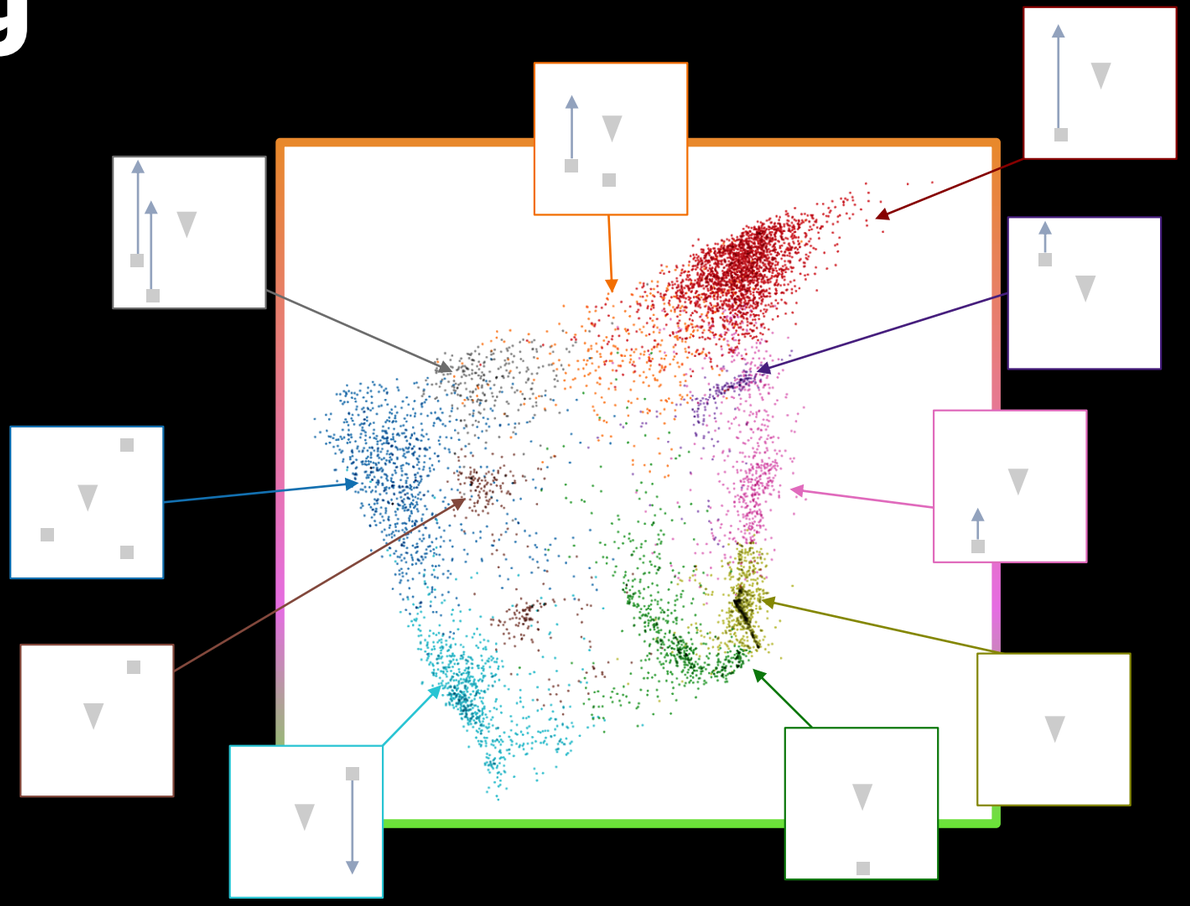
Simulation scenario

Outliers of cut-ins



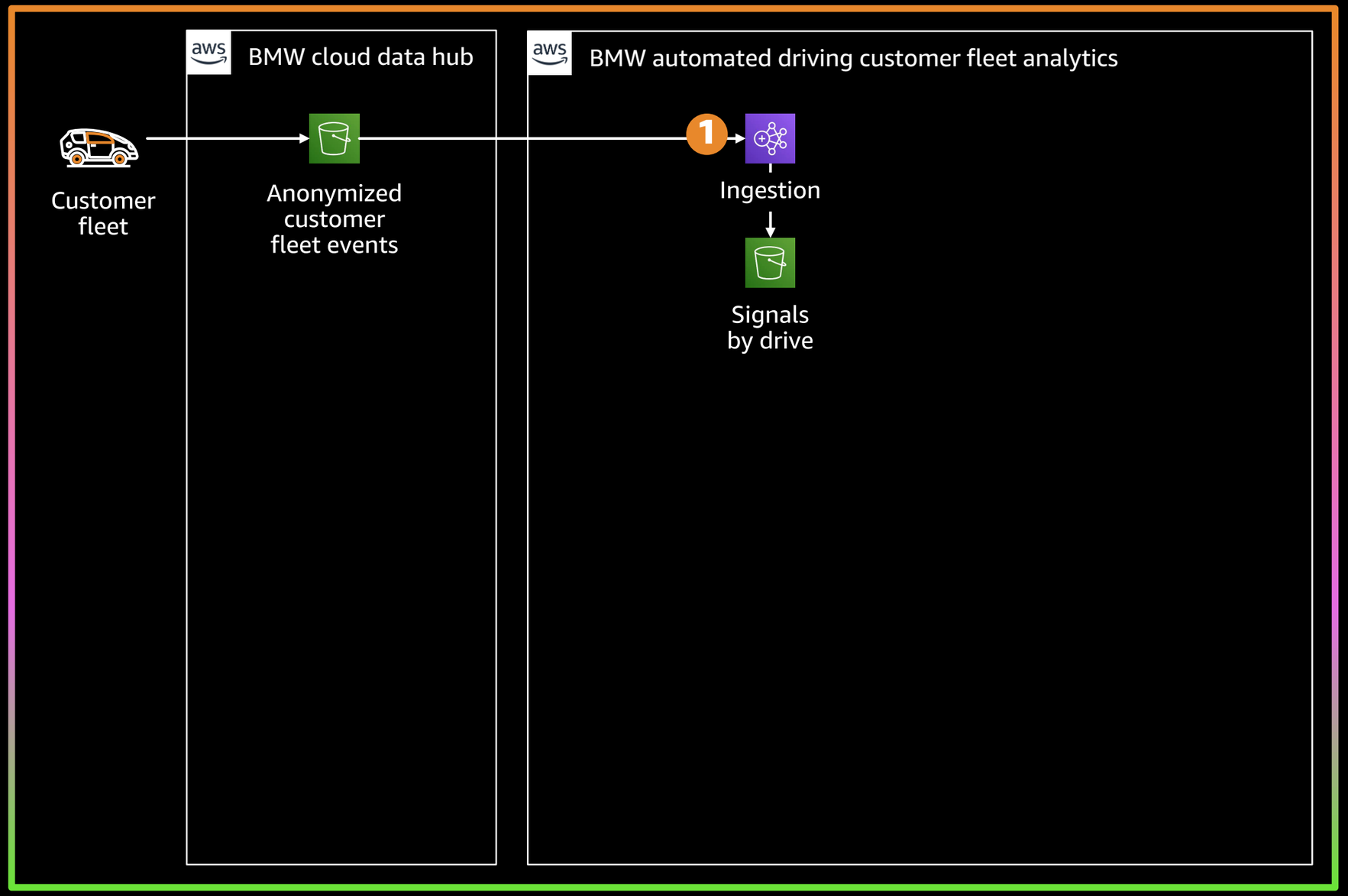
Deep representation learning and clustering of traffic scenarios

Traffic scenario categorization



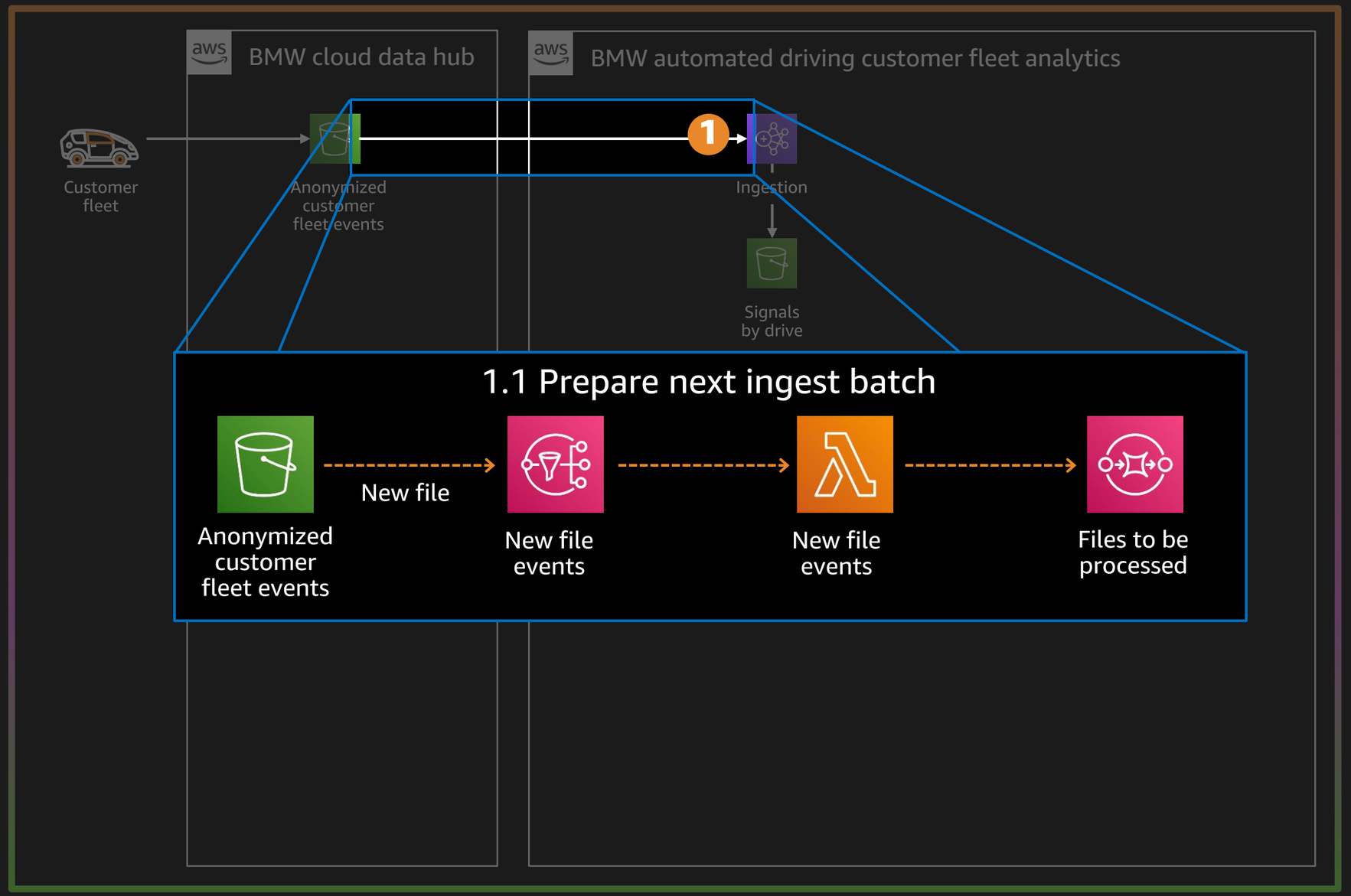
*Presented at ICML AI/AD 2020 (Nick Harmening et al, best paper award)

Data pipeline



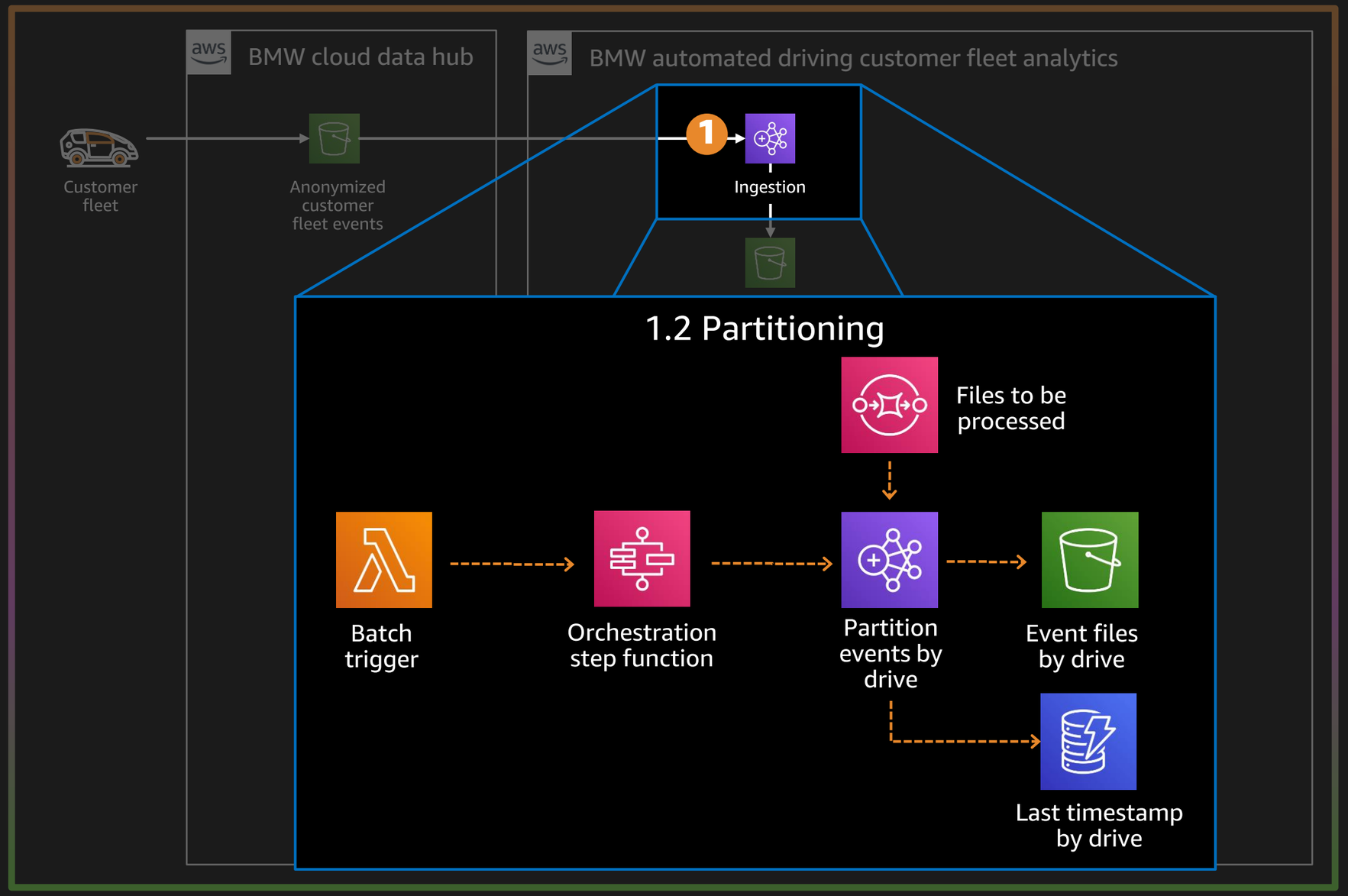
1 Ingestion

Data pipeline



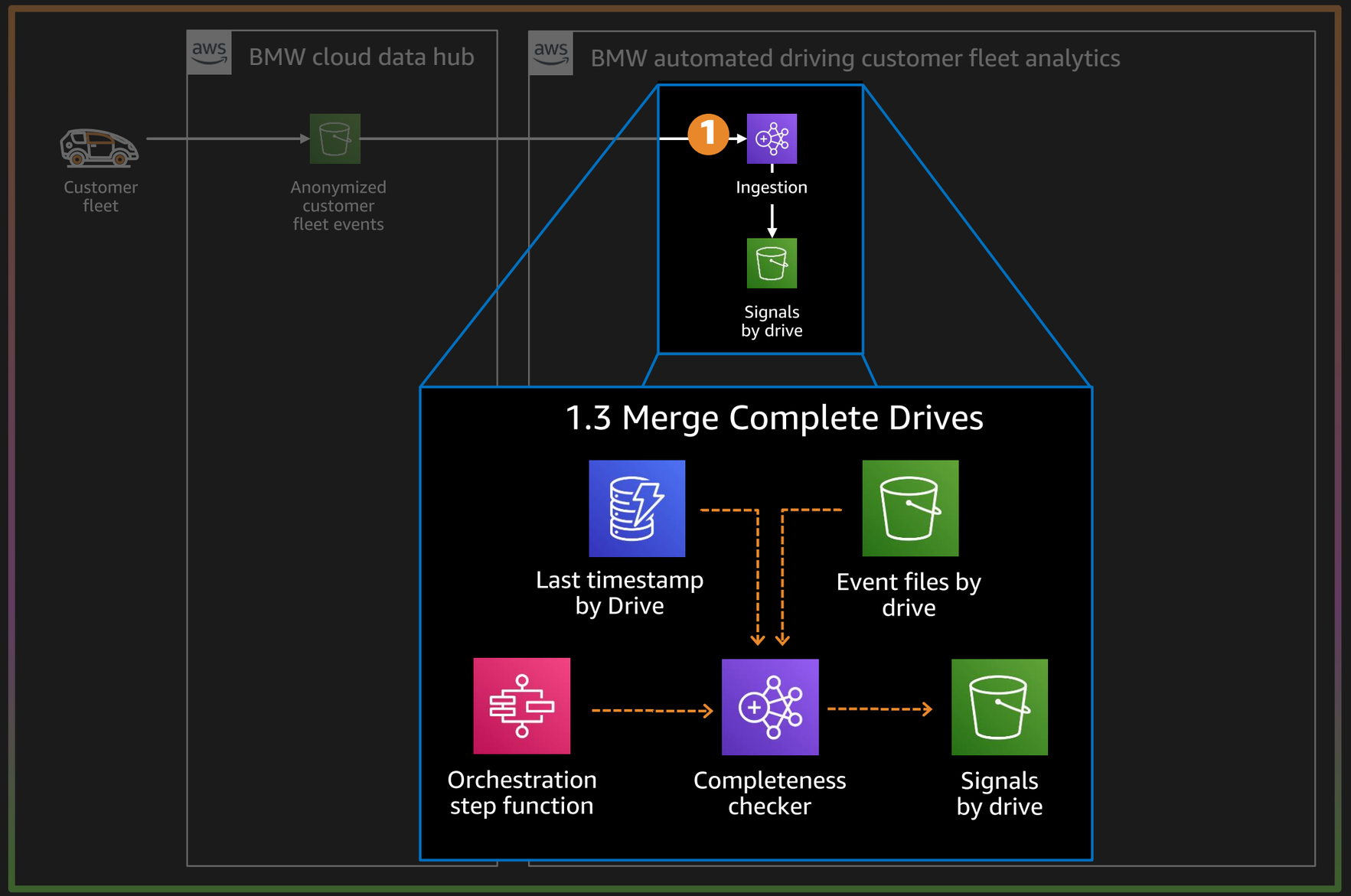
1 Ingestion

Data pipeline



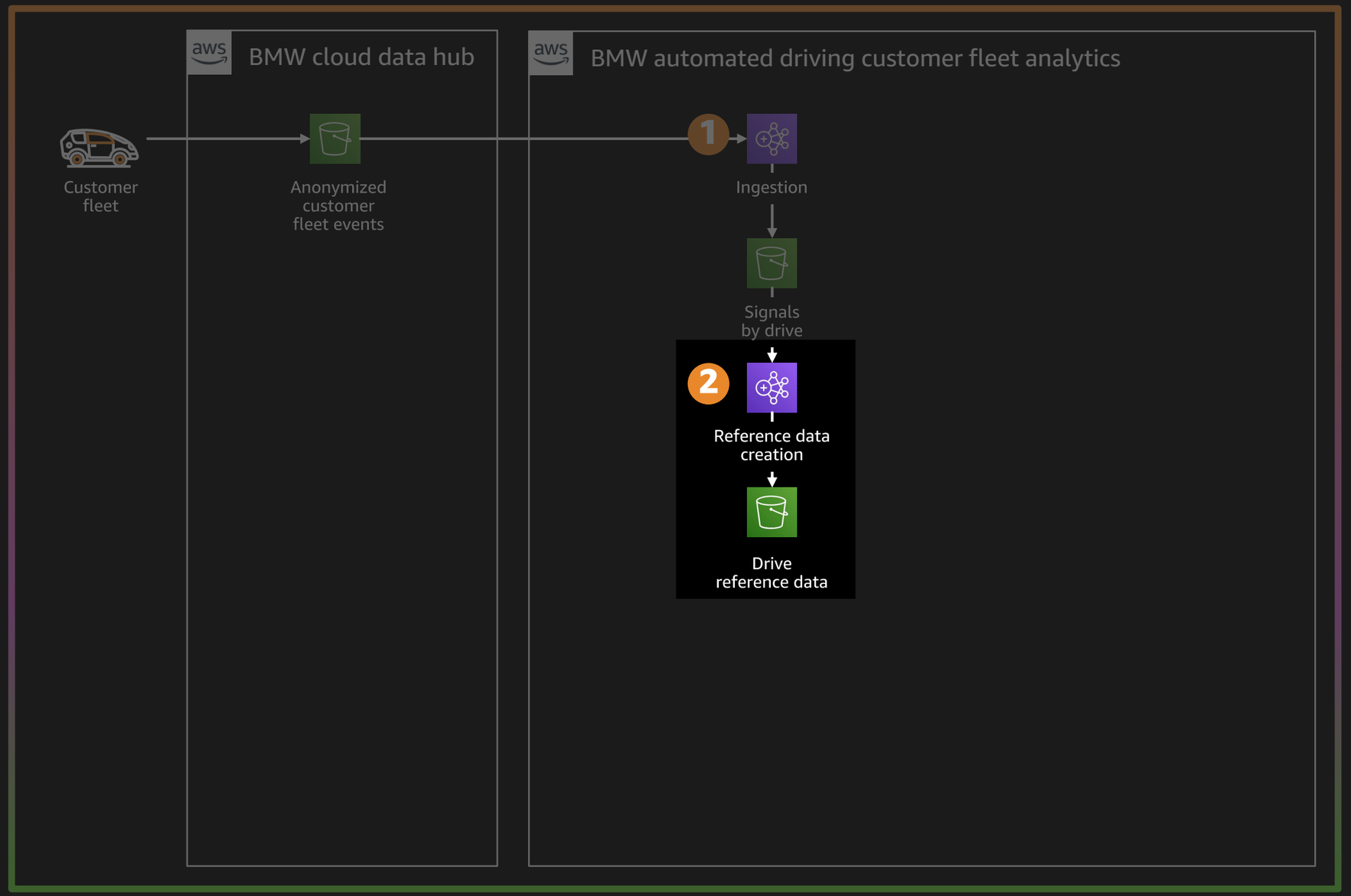
1 Ingestion

Data pipeline



1 Ingestion

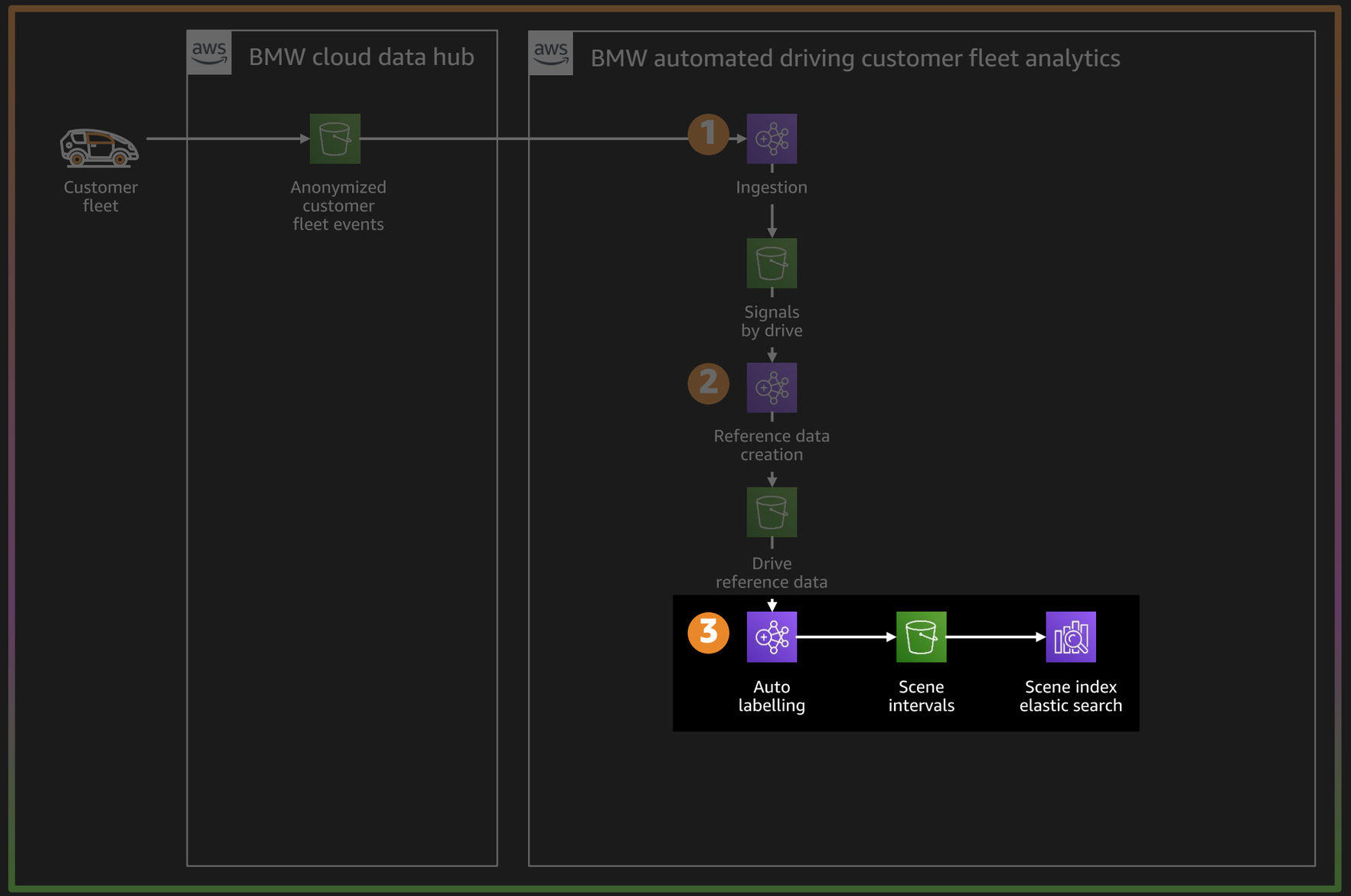
Data pipeline



1 Ingestion

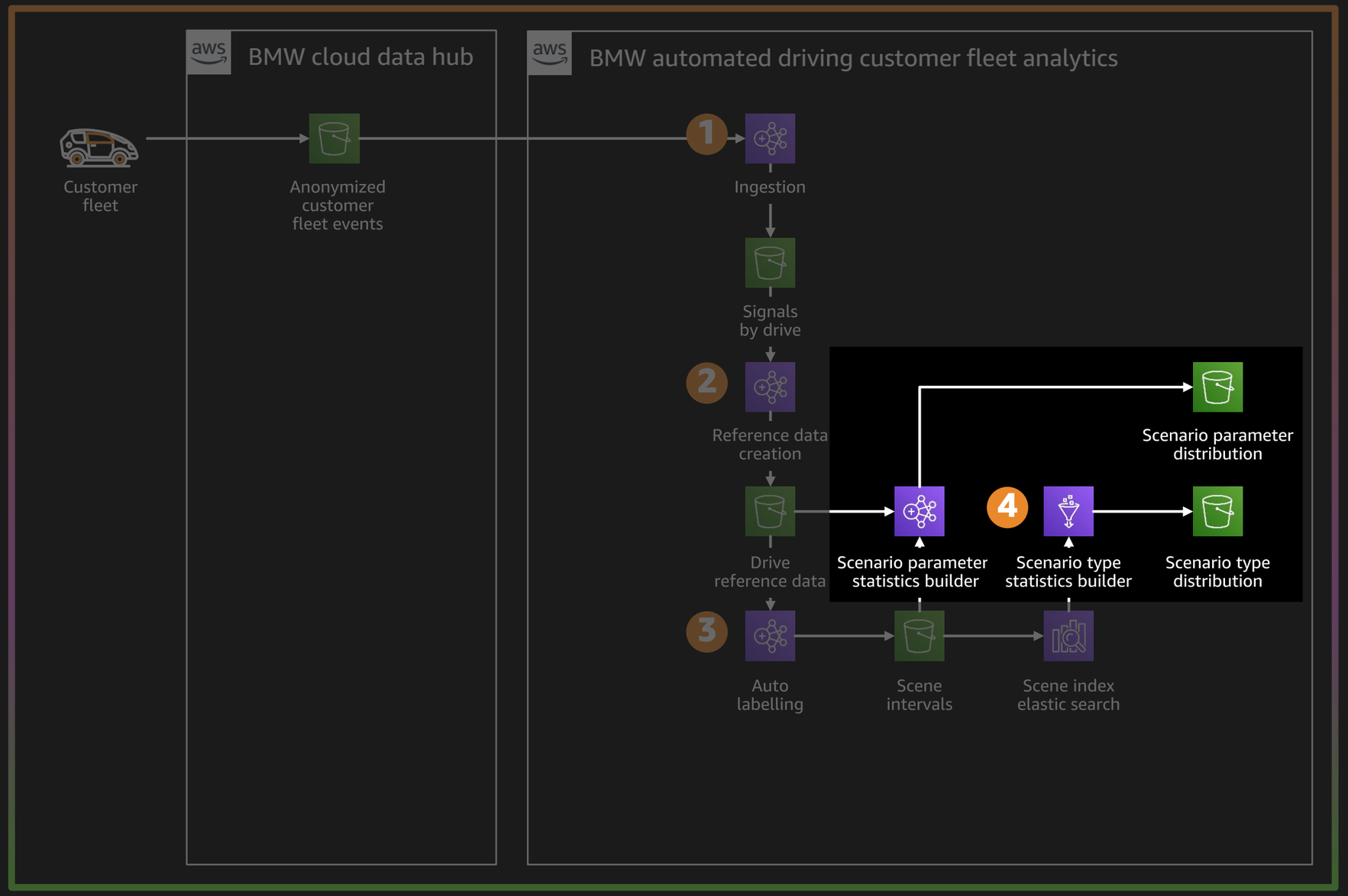
2 Reference data creation

Data pipeline



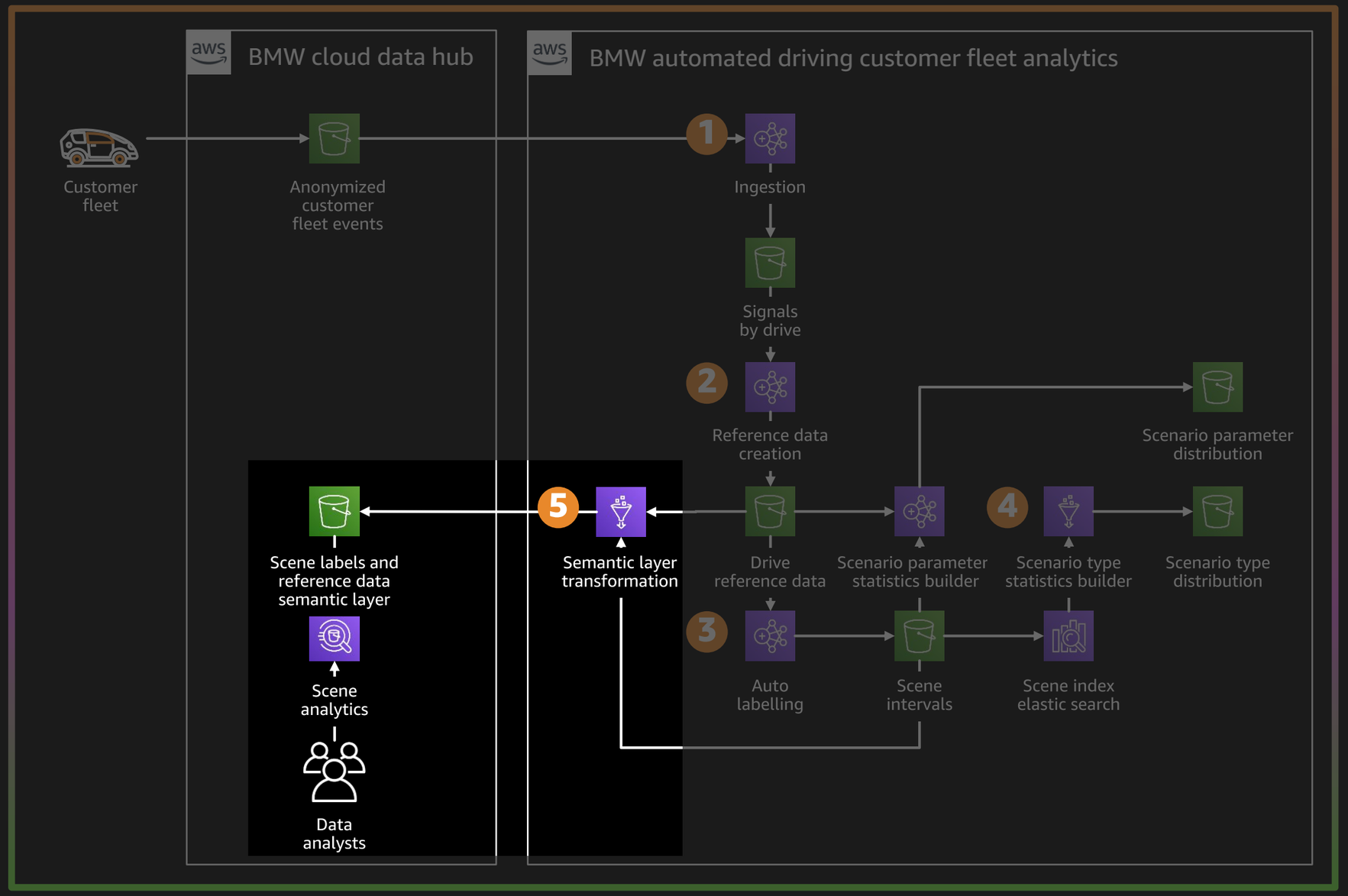
1. Ingestion
2. Reference data creation
3. **Scenario detection**

Data pipeline



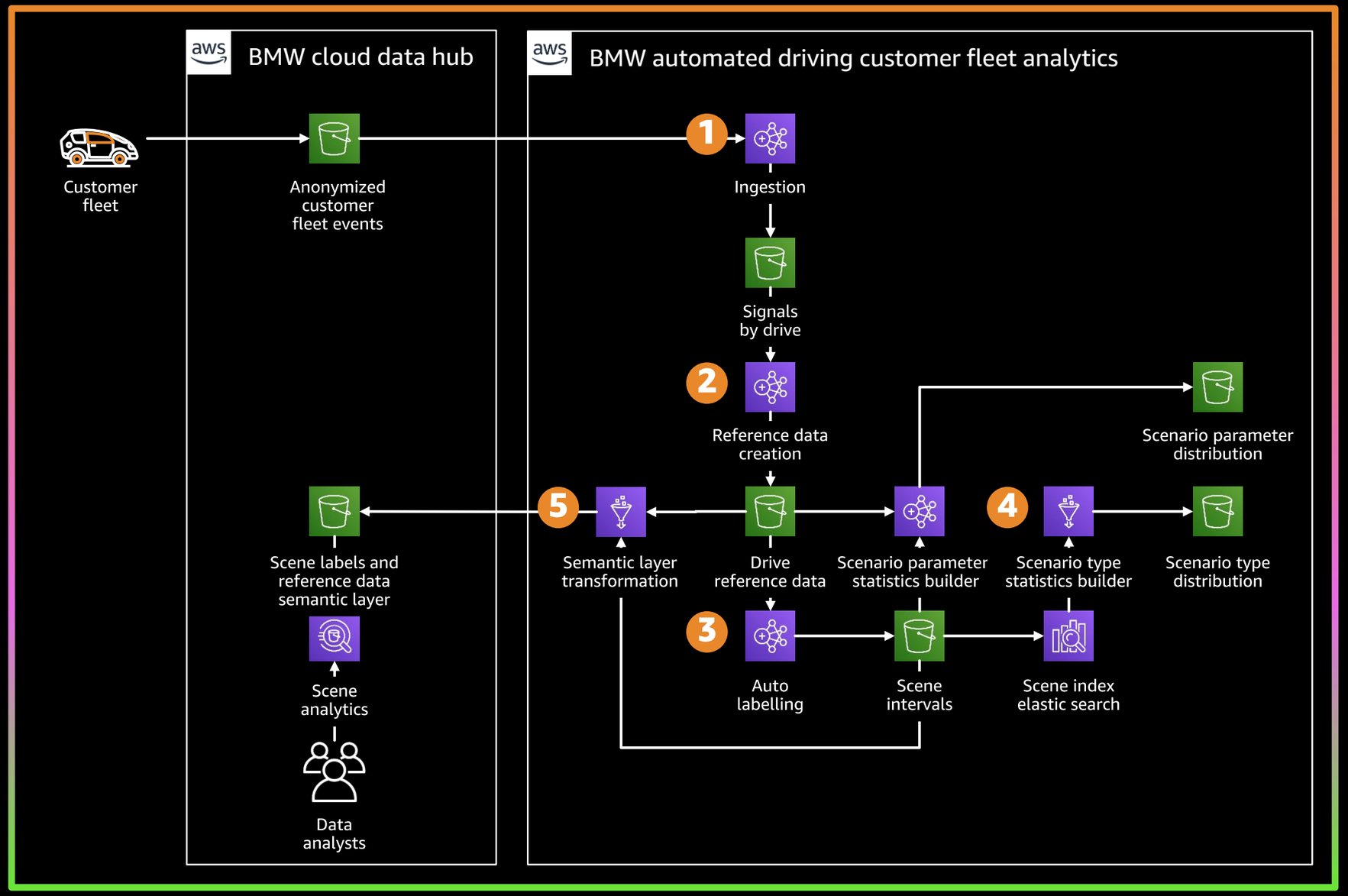
1. Ingestion
2. Reference data creation
3. Scenario detection
4. **Derive distributions**

Data pipeline



1. Ingestion
2. Reference data creation
3. Scenario detection
4. Derive distributions
5. Analytics enabling

Data pipeline



1. Ingestion
2. Reference data creation
3. Scenario detection
4. Derive distributions
5. Analytics enabling

Autonomous vehicle data lake

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Autonomous driving development workflow

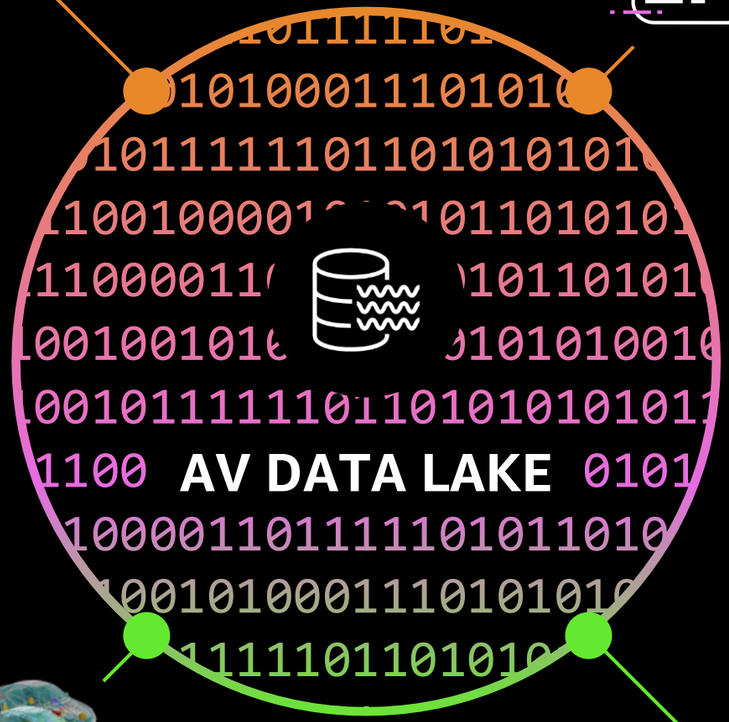
Data processing and analytics

Labeling

Map development

Model and algorithm development

Simulation and validation



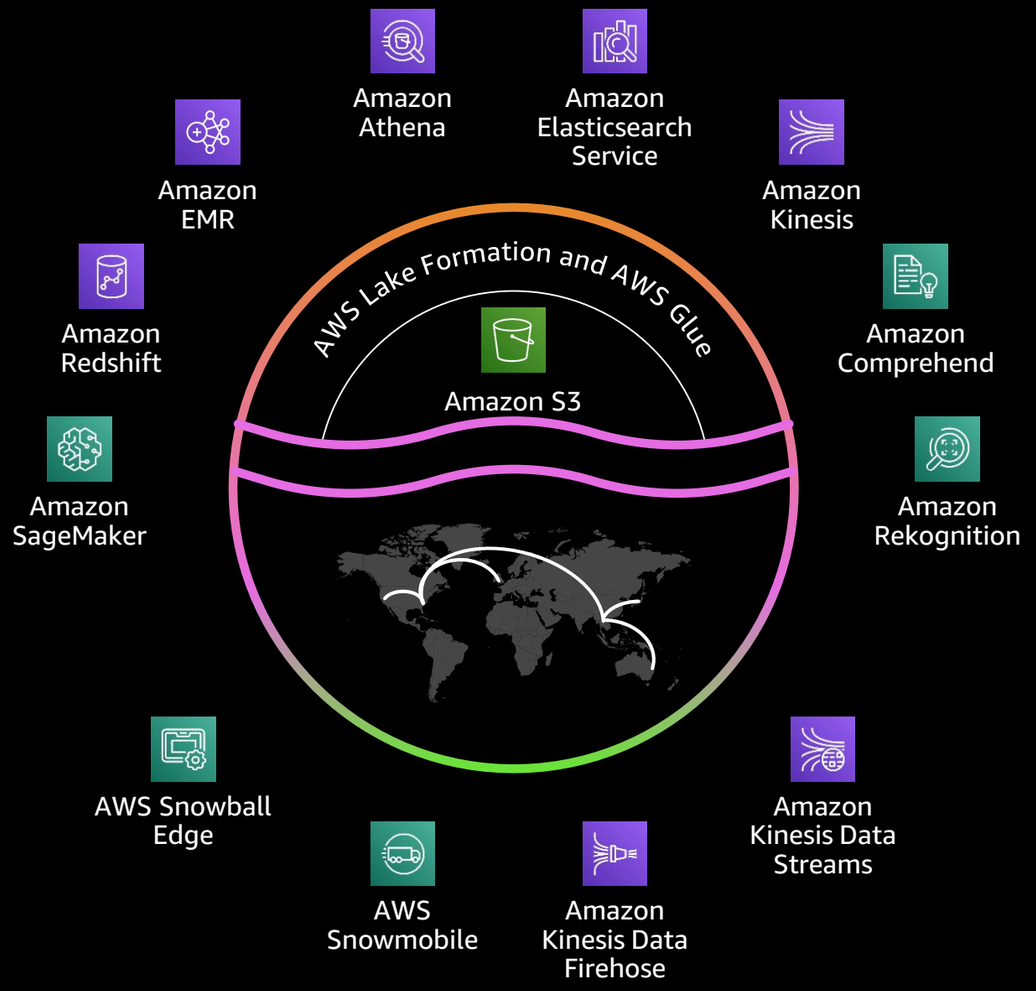
Data acquisition and ingest

Orchestration and deployment

Autonomous vehicle data lake challenges

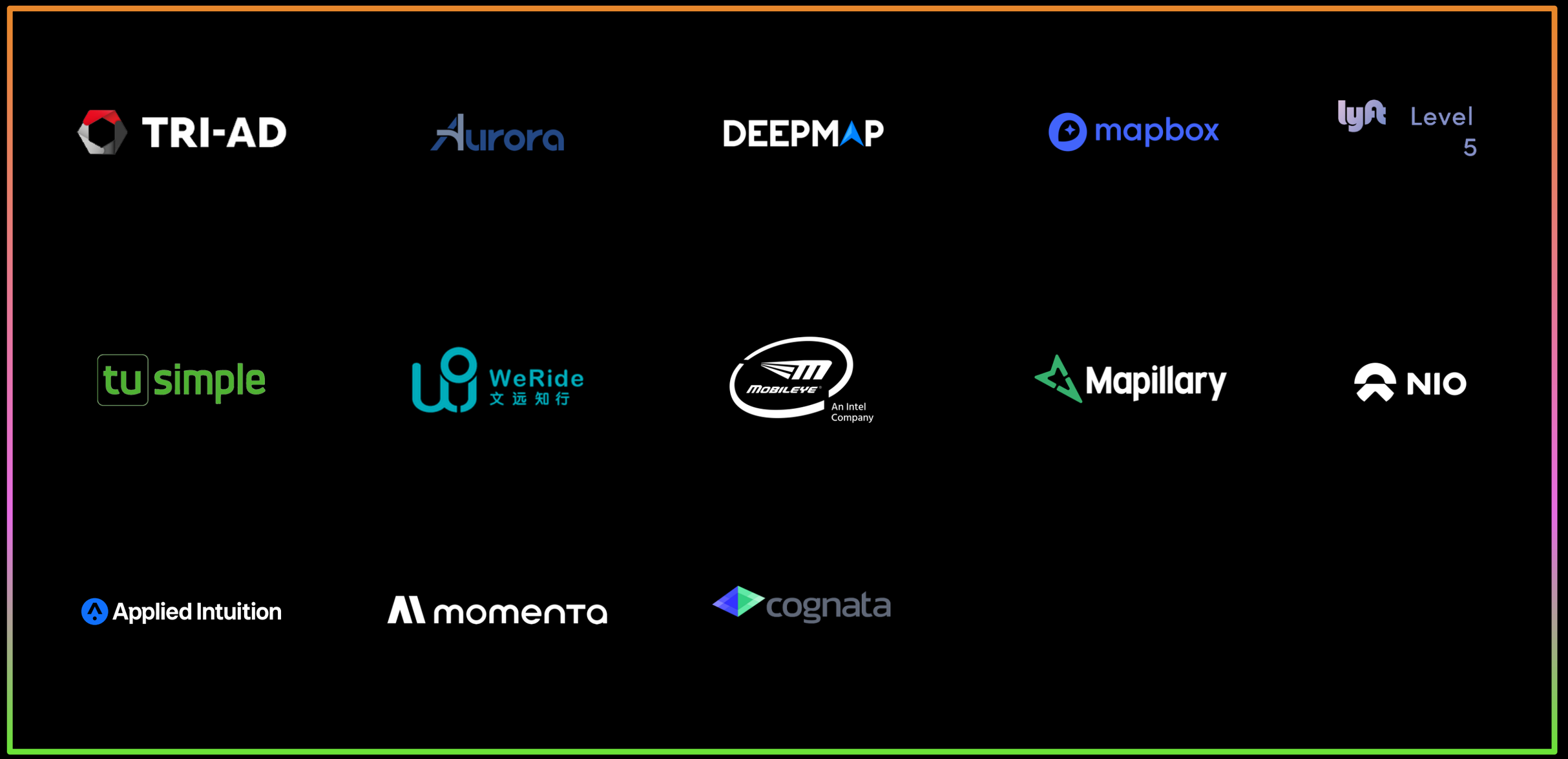
- Big data repository for drive logs (raw/processed) used for AV development
- Enable diverse processing pipelines for many approaches including globally distributed compute resources
- Standards based search and analysis by developers, data scientists, researchers, and even legal departments
- Enable high performance throughput and latency needs of the downstream workloads

Amazon S3 as the foundation for AV data lake



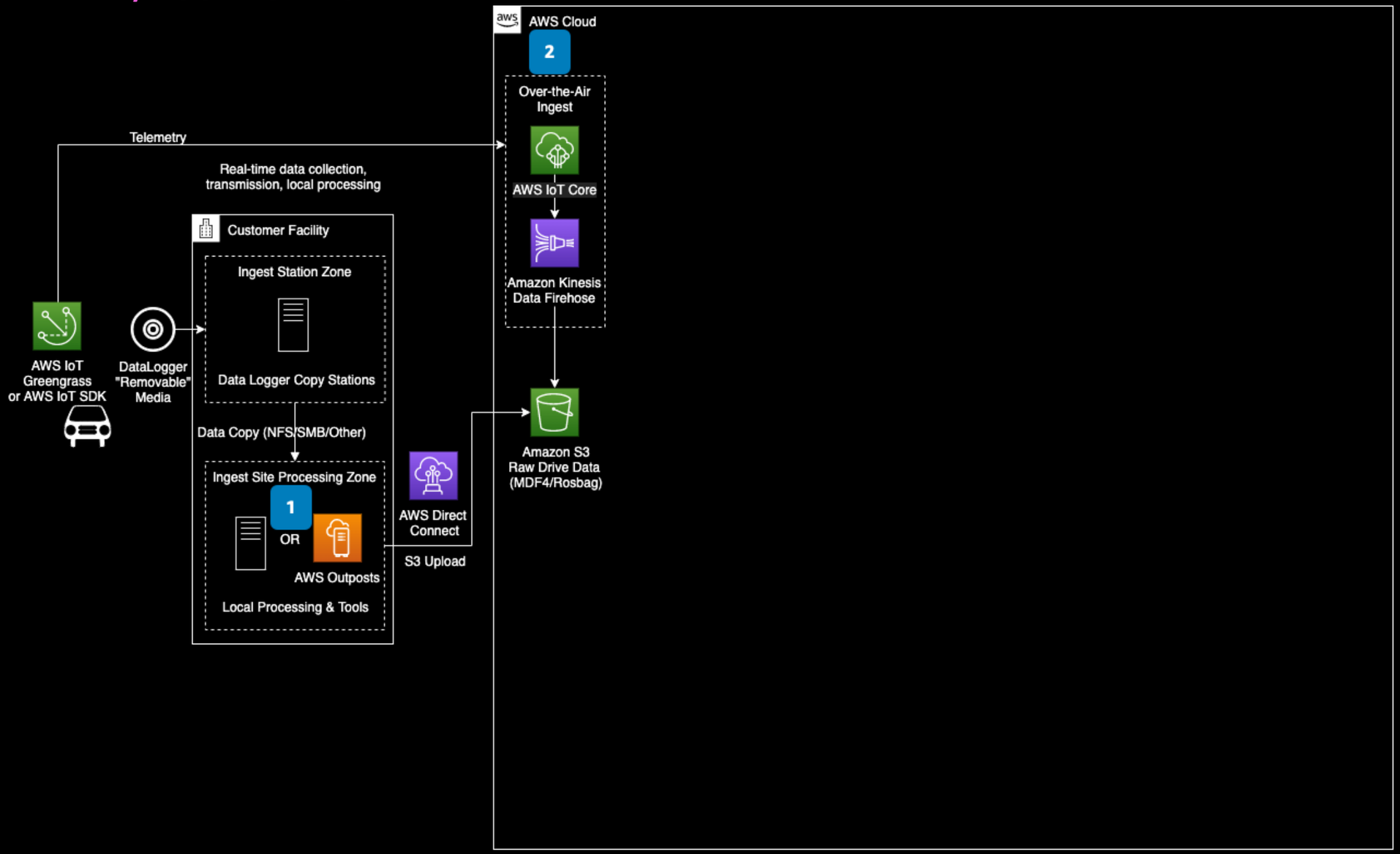
- Durable, available, exabyte-scalable
- Secure, compliant, auditable
- High performance
- Low-cost storage and analytics
- Broad network integration

AWS autonomous vehicle customer references



AV Data lake reference architecture

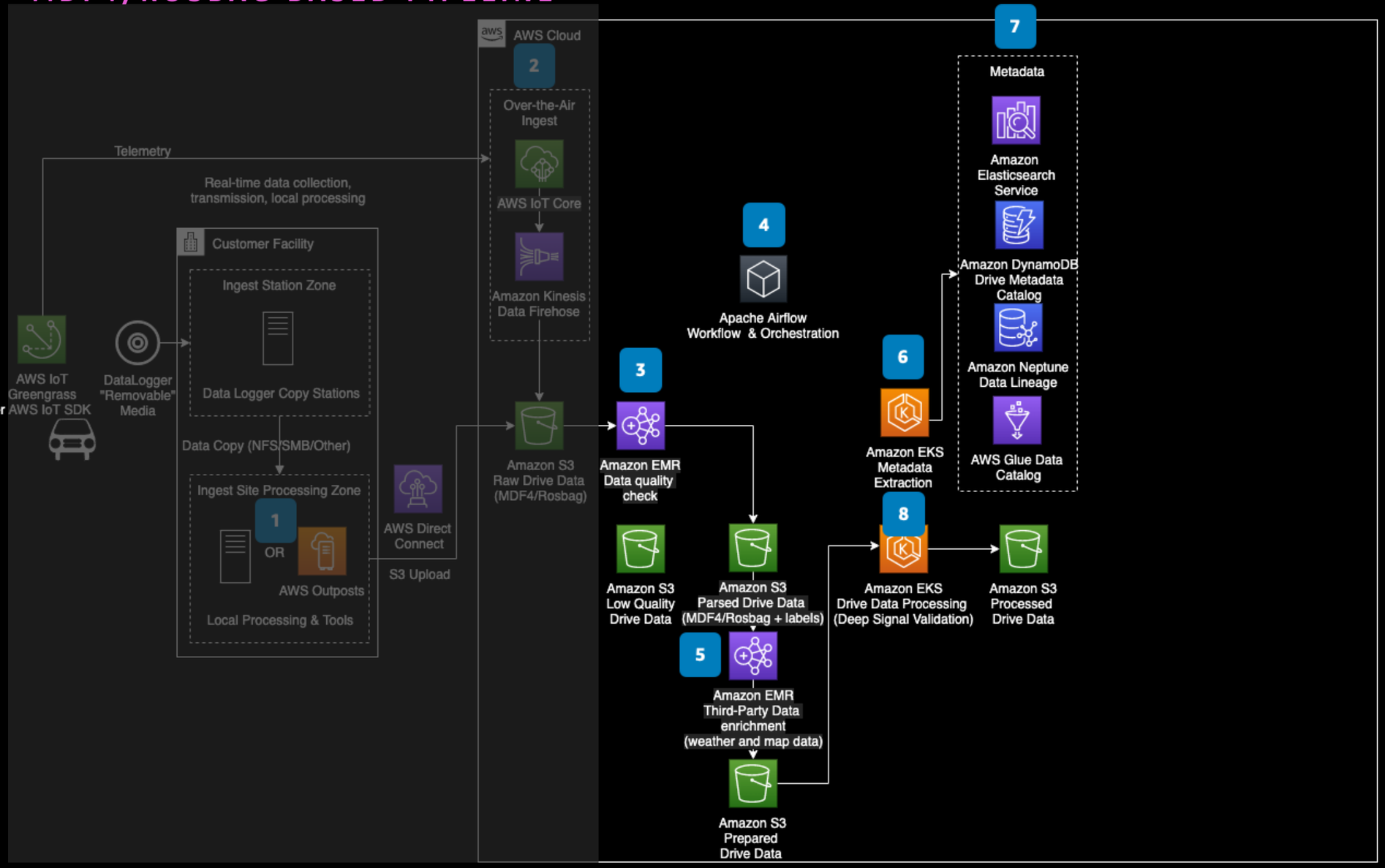
MDF4/ROSBAG BASED PIPELINE



1. Data Ingestion of AV development fleet using AWS Outpost for local data processing
2. Real-time ingestion of Fleet Data using AWS IoT Core and Kinesis Data Firehose

AV Data lake reference architecture

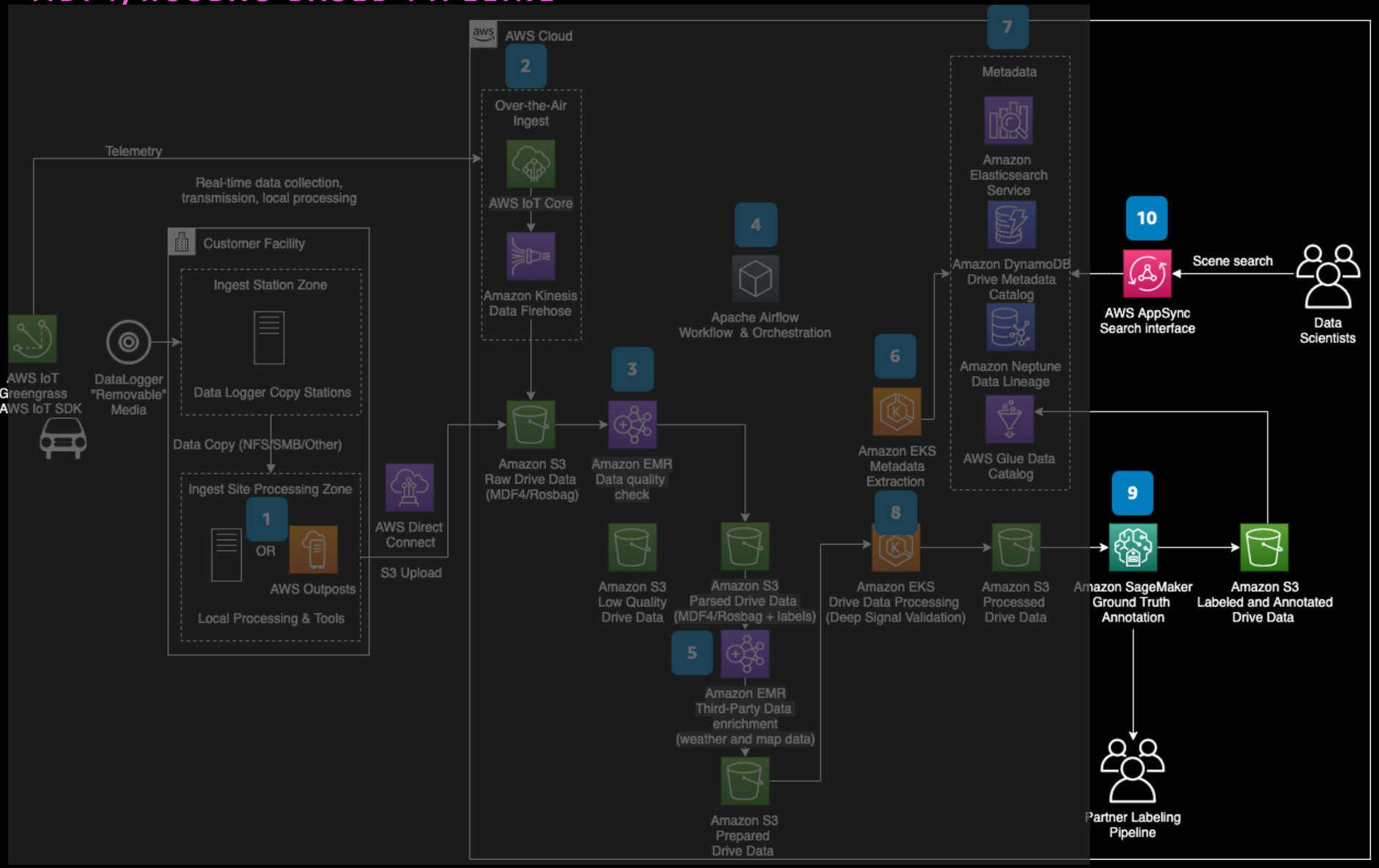
MDF4/ROSBAG BASED PIPELINE



1. Data Ingestion of AV development fleet using AWS Outpost for local data processing
2. Real-time ingestion of Fleet Data using AWS IoT Core and Kinesis Data Firehose
3. Data quality sanity check to filter data with low quality
4. Parse the data with right labels and components
5. Enrichment with Weather Data based on GPS location and timestamp. Enrichment with Map Data based on GPS location from e.g. HERE Open Location Platform
6. Extract Metadata using ASAM OpenSCENARIO and store them into AWS DynamoDB and Amazon Elasticsearch
7. Store data lineage in Amazon Neptune and data catalog in Glue Data Catalog
8. Process drive data using EKS like Deep Signal Validation

AV Data lake reference architecture

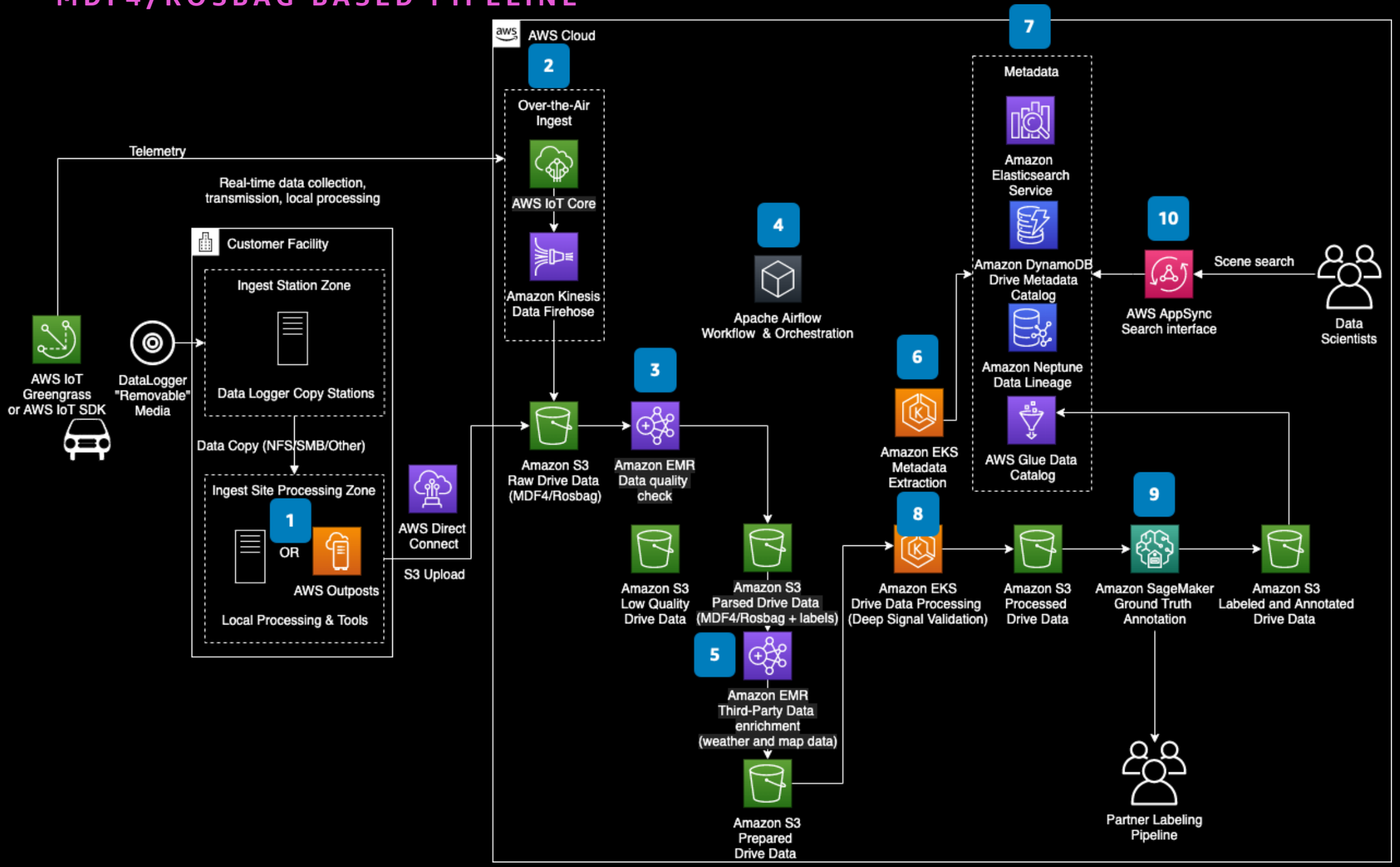
MDF4/ROSBAG BASED PIPELINE



1. Data Ingestion of AV development fleet using AWS Outpost for local data processing
2. Real-time ingestion of Fleet Data using AWS IoT Core and Kinesis Data Firehose
3. Data quality sanity check to filter data with low quality
4. Parse the data with right labels and components
5. Enrichment with Weather Data based on GPS location and timestamp. Enrichment with Map Data based on GPS location from e.g. HERE Open Location Platform
6. Extract Metadata using ASAM OpenSCENARIO and store them into AWS DynamoDB and Amazon Elasticsearch
7. Store data lineage in Amazon Neptune and data catalog in Glue Data Catalog
8. Process drive data using EKS like Deep Signal Validation
9. Auto labeling using Amazon SageMaker Ground Truth
10. Provide search function for particular scene using AWS AppSync

AV Data lake reference architecture

MDF4/ROSBAG BASED PIPELINE



1. Data Ingestion of AV development fleet using AWS Outpost for local data processing
2. Real-time ingestion of Fleet Data using AWS IoT Core and Kinesis Data Firehose
3. Data quality sanity check to filter data with low quality
4. Parse the data with right labels and components
5. Enrichment with Weather Data based on GPS location and timestamp. Enrichment with Map Data based on GPS location from e.g. HERE Open Location Platform
6. Extract Metadata using ASAM OpenSCENARIO and store them into AWS DynamoDB and Amazon Elasticsearch
7. Store data lineage in Amazon Neptune and data catalog in Glue Data Catalog
8. Process drive data using EKS like Deep Signal Validation
9. Auto labeling using Amazon SageMaker Ground Truth
10. Provide search function for particular scene using AWS AppSync

Thank you!

Marc Neumann

Product Owner Data Engineering Automated Driving,
BMW AG

Junjie Tang

Principal Consultant,
AWS

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Please complete the session survey

AWS re:Invent

© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.