PRESIDIO® | Case Study

Manufacturing & Industrial

# Building the AI brain for autonomous vehicles - on a startup budget, at fleet scale

## The Outcome

Halo, a UK autonomous vehicle innovator, needed a cloud architecture capable of ingesting terabytes of real-world fleet data daily, training vehicle-aware AI models, and doing it all cost-effectively on a startup budget. Presidio designed and deployed an AWS-based platform that handles the full data lifecycle - ingest, categorize, train, delete - and hosts Halo's V2X platform for global fleet deployment.

## How They Got There

Presidio worked closely with Halo to map data ingestion requirements across a growing multi-vehicle fleet, then designed a tailored AWS architecture using S3, IoT Core, Lambda, SageMaker, Glue, and ECR. Data is stored only for the duration of training, then automatically and securely deleted. The Cloud Management Portal adds real-time visibility into compute, storage, and pipeline performance, with granular billing transparency.

## Why It Worked

Halo's AI approach incorporates surrounding infrastructure, other vehicles, and human behavior - not just the vehicle itself. Training those models requires diverse, real-world fleet data at scale. The architecture Presidio built gives Halo the throughput to train continuously on live data while keeping cloud costs lean enough for a scaling startup to sustain.

### The Situation

Halo takes a vehicle-aware approach to autonomous driving AI - one that incorporates connected infrastructure, surrounding vehicles, and human inputs, not just the vehicle itself. Halo had secured data collection agreements with existing fleet operators to access real-world driving data. The challenge was building a scalable cloud architecture that could ingest that data from hundreds of simultaneous vehicles, manage the orchestration and AI training workflow, and do it cost-effectively on a startup budget.

### The Approach

Presidio designed and implemented a tailored AWS cloud architecture built to handle Halo's high-throughput, multi-source data environment. The platform ingests and categorises HD video and LiDAR point cloud data captured during routine fleet operations. These multi-modal datasets feed Halo's AI training pipeline, producing models designed to interpret dynamic hazards, adverse weather, and human behaviour. Data is automatically and securely deleted after training, keeping storage costs low and compliance clean.

### The Result

Halo now operates a scalable, intelligent data management platform that supports continuous AI training at fleet scale. Storage costs stay low because data lifecycle management is automated. Cloud spend is predictable and transparent. Halo's ViewPort V2X platform runs on the same AWS architecture, providing real-time fleet routing optimisation and scalability for global operator deployment.

*"The new cloud architecture that Presidio delivered drives our mission to accelerate the development of autonomous vehicle technology in the UK." — Terry Sunny, Founder, Halo*

**Terabytes/day**

Of raw sensor data ingested, processed, and auto-deleted post-training

**Multi-modal**

LiDAR + HD video combined into a single AI training pipeline

**Real-time**

Fleet-scale data ingestion from hundreds of simultaneous vehicles

AI Training | AWS | Autonomous Vehicles | IoT | SageMaker | Cloud Architecture | Startup

PRESIDIO®