AI Systems in IoT-Scale Deployments

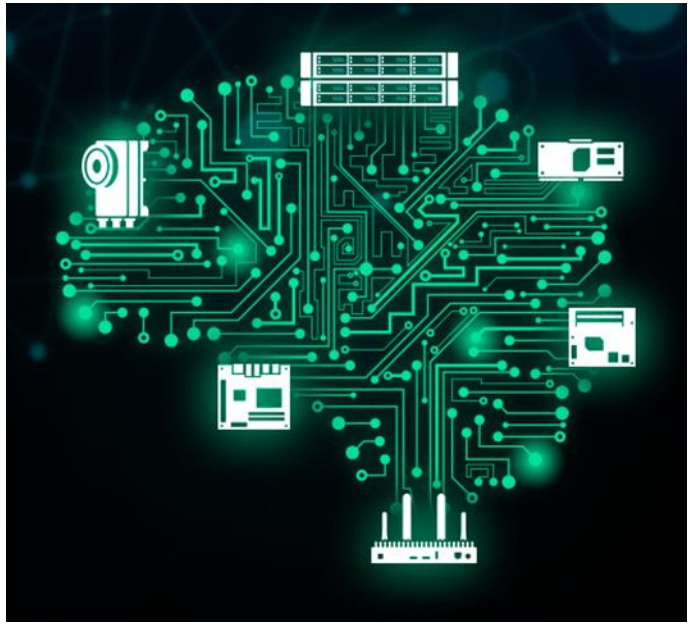
Simon Collins
Senior Product Manager

GTC Europe
Thursday, 11 October 2018

ADLINK
Leading EDGE COMPUTING

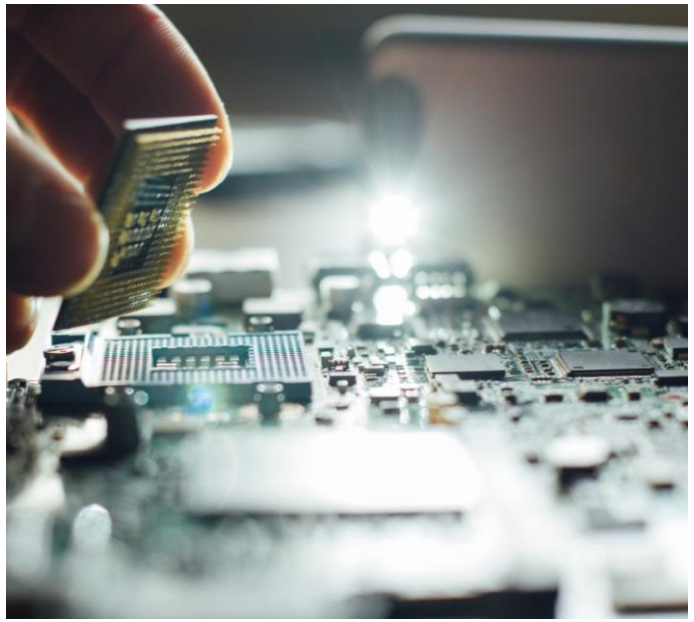
Massive global growth driven by data

Enabled by convergence of key technologies



Artificial Intelligence to create $3.5 trillion - $5.8 trillion of value globally

Internet of Things to have 24 billion connected devices, 163 zetabytes of data by 2025



Edge Computing set to reach $3.24 billon by 2025

How do we manage data across this vast ecosystem?

AI Systems in IoT-Scale Deployments

Problem space

As IoT grows, the “Things” become more widely distributed

☺ Generates more value

😐 Generates more data

☹ Increases total system complexity
☹ Increases management costs
☹ Increases data connectivity costs

AI at the Edge allows “intelligent” machines to make sense of the data and transmit only valuable information

→ Lowers latency
→ Keeps data on-premise = privacy
→ Increases reliability

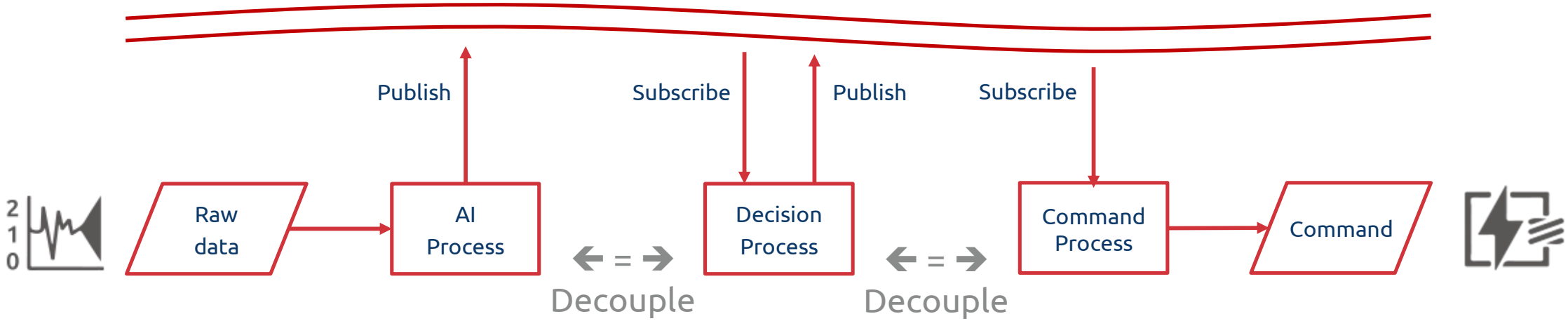
Increasingly capable hardware allow us to deploy more sophisticated tools at the Edge

We need to be able to easily and securely manage the distribution of intelligence, while enabling continuous upgrades in capability

Solution – Data-centric architecture

Focus on the data, not on the connectivity

Easy to Integrate - Data-centric architecture that hides topology details, enabling **true plug-and-play**
Scalable - Scaling across thousands or even millions of participants
Open – **Mature, proven and open standard** with future-proof APIs plus wire-protocol
Secure – **End-to-end secure data connectivity** with authentication, encryption & access control capabilities
Fault-tolerant – Peer-to-peer communication **without message brokers or servers**
Widely applicable - Available for embedded, mobile, web, enterprise and cloud applications
QoS-enabled – Full control over data distribution: **timeliness, prioritization, reliability and resource usage**
High-performing - Latencies as low as 30 µsec, throughput of millions of messages/sec

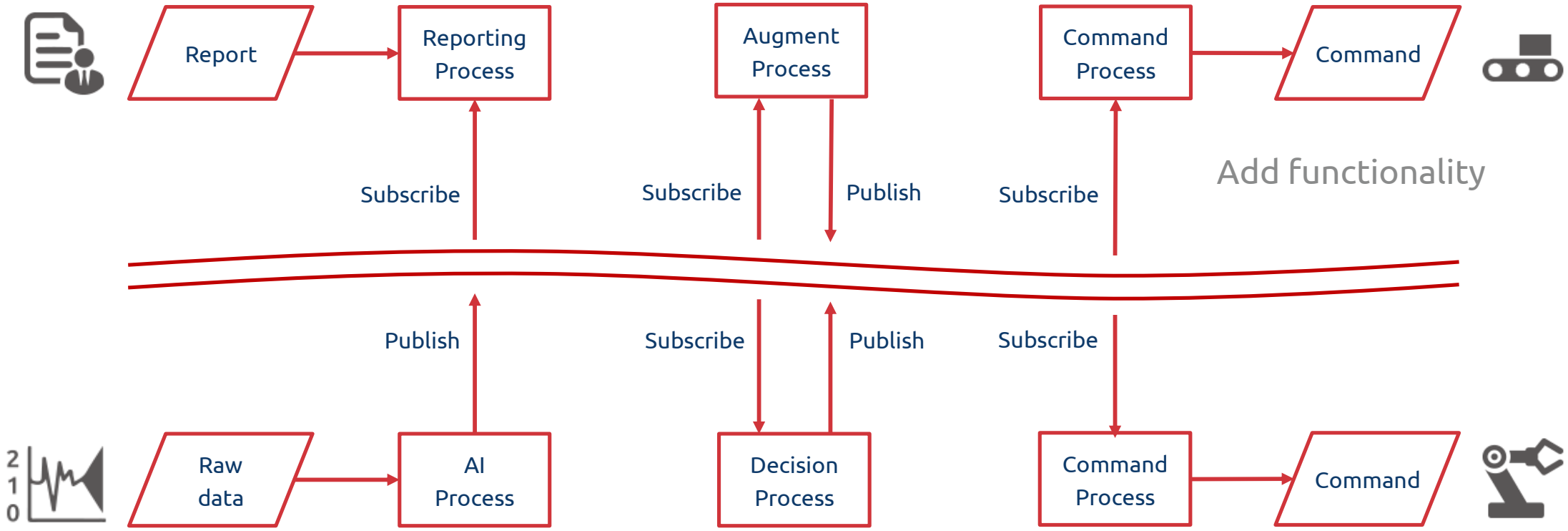


Solution – Data-centric architecture

Rapid scalability and re-configuration

Allows a microservice architecture, each component focussed on its task

→ Move towards *Software Defined Machines*



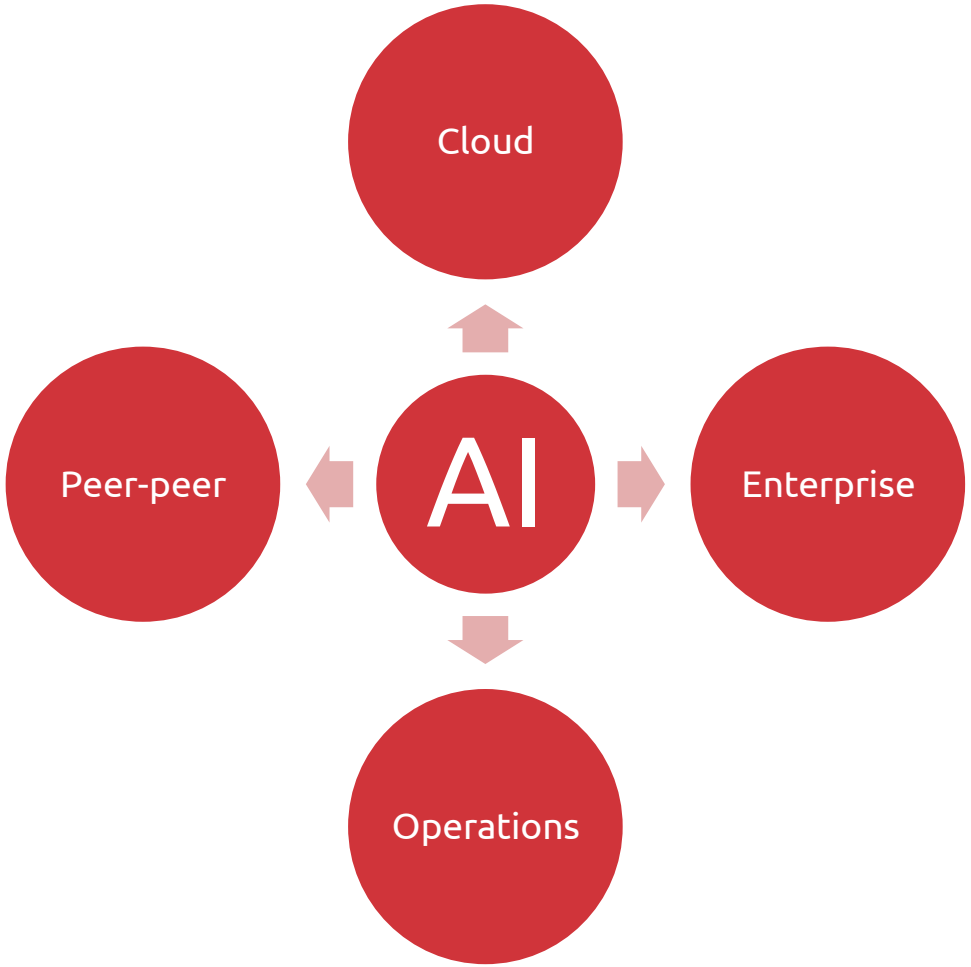
IoT needs a lot of other connectivity

Decouple to achieve focus

AI developers focus on their core task, turning the input data into valuable information

Decouple from other aspects of the system; let others deal with:

- Cloud connectivity – *northbound*
- Peer-peer communications – *east-west*
- Operations & control – southbound
  - Modbus, PROFINET, CANbus, …
- Enterprise connectivity

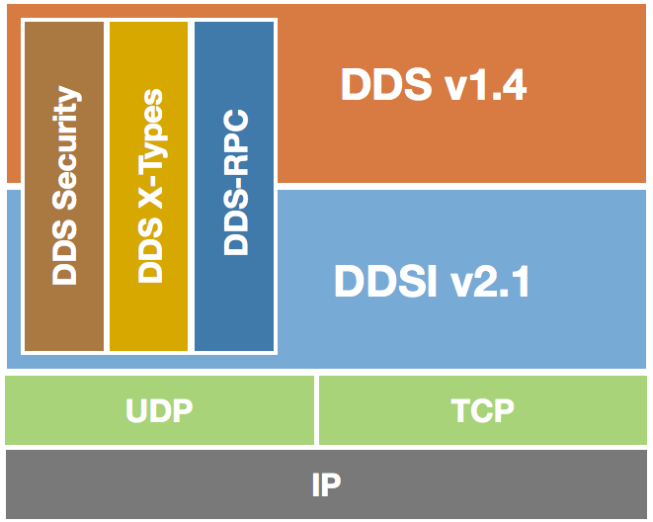


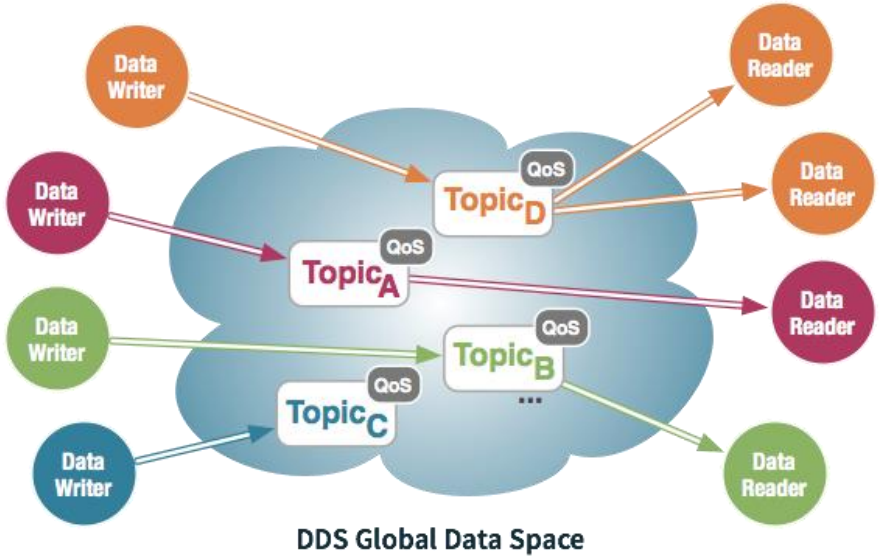
Underlying technology

Data Distribution Service

Addresses the needs of large-scale mission and business-critical applications, scaling to ultra-large-scale deployments

Open standard for efficient, secure, interoperable, platform- and programming-language independent data sharing between networked devices





Applications can autonomously and asynchronously read and write data enjoying spatial and temporal decoupling

Built-in dynamic discovery isolates applications from network topology and connectivity details

Proven in Defense & Aerospace

Widely deployed by early adopter market

Integrated Modular Vetronics

Training & Simulation Systems

Naval Combat Systems

Air Traffic Control & Management

Unmanned Air Vehicles

Aerospace Applications

Broad Commercial Applications

Autonomous Vehicles

Industrial Automation

Smart Cities

Transport Management Systems

Complex Medical Devices

Smart Energy Grid

Incorporated into further standards

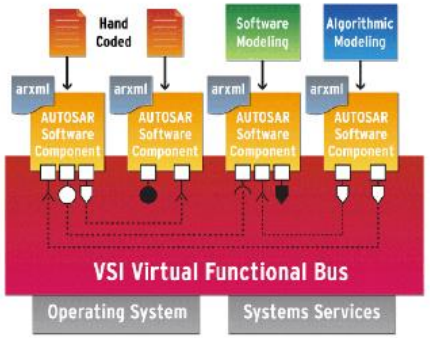
ADLINK

ROS industrial

Ministry of Defence

Generic Vehicle Architecture

AUTOSAR

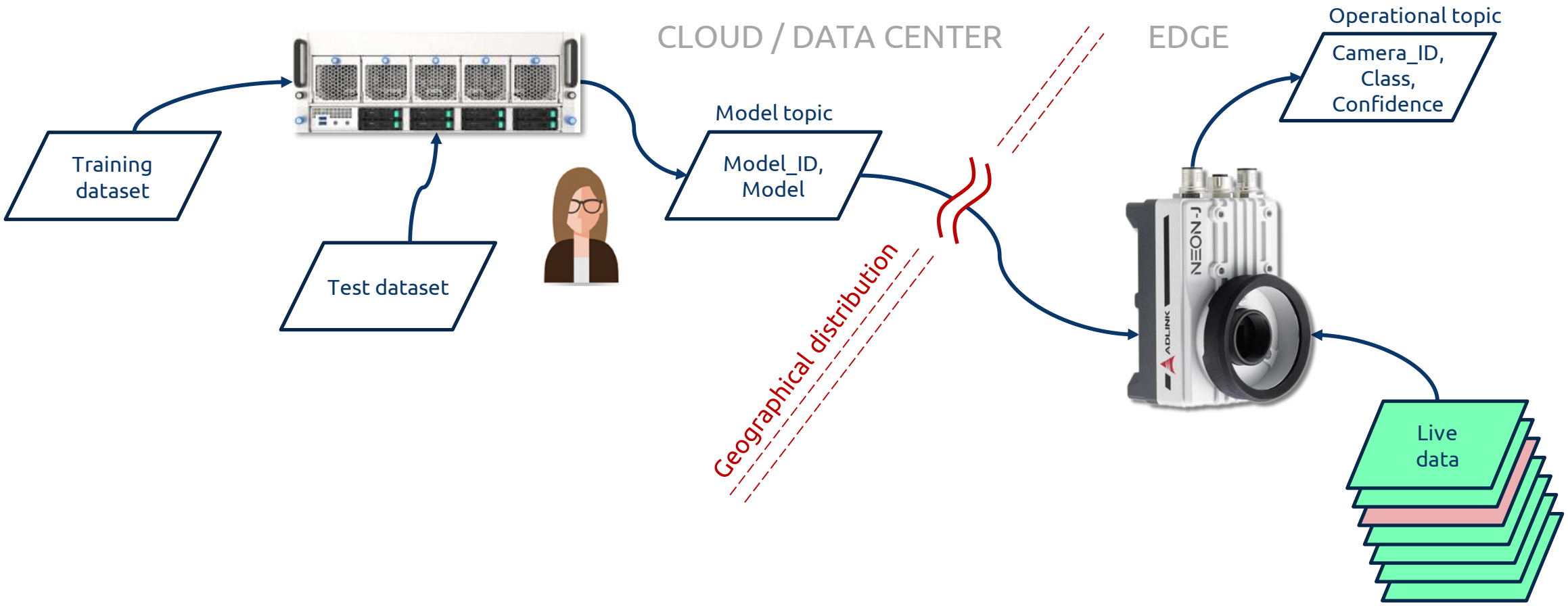


BOSCH

FACE™
Future Airborne Capability Environment

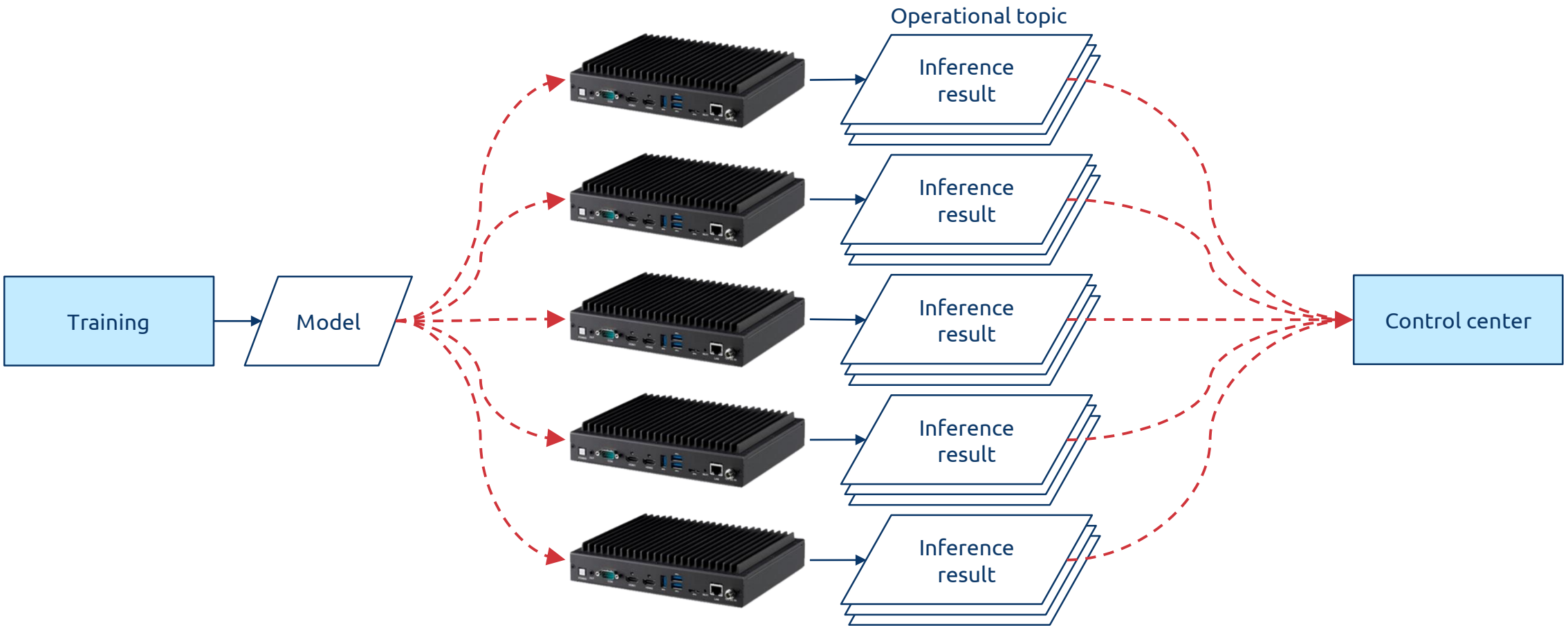
AI model training & deployment use case

Standard deployment



Distributed AI – homogeneous system

Many identical nodes collaborating



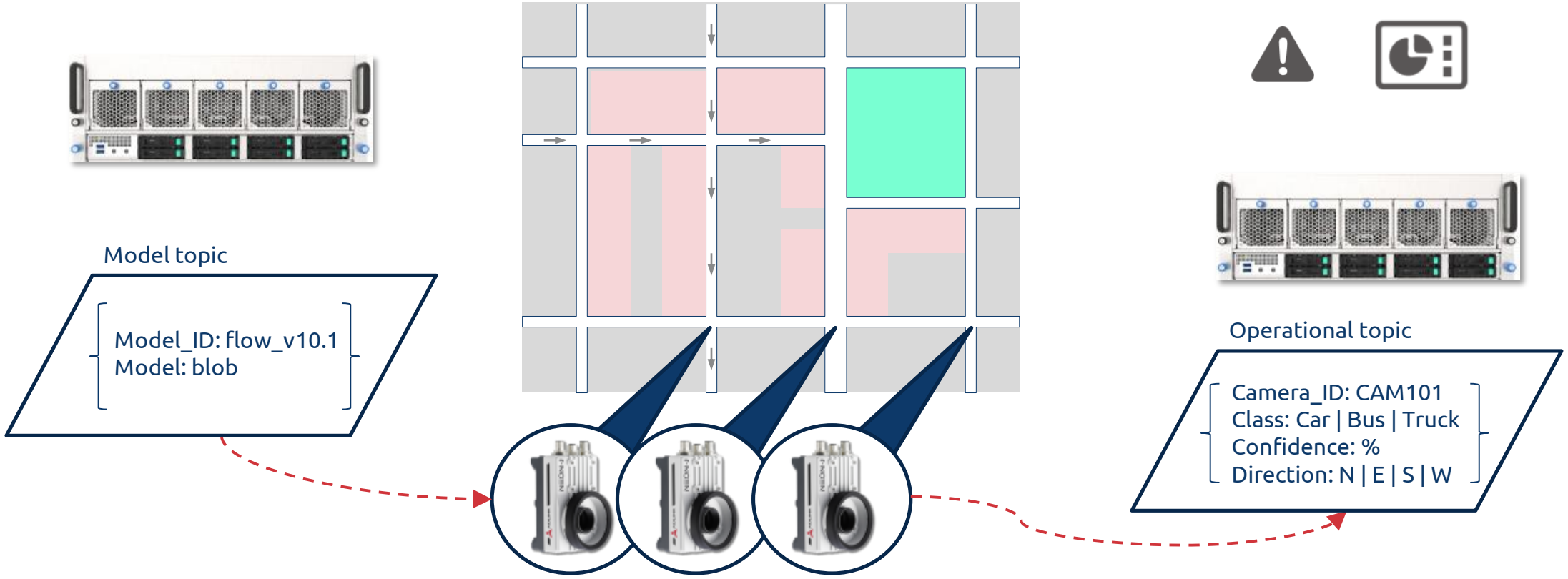
DATA CENTER

EDGE

CONTROL CENTER

Traffic flow example

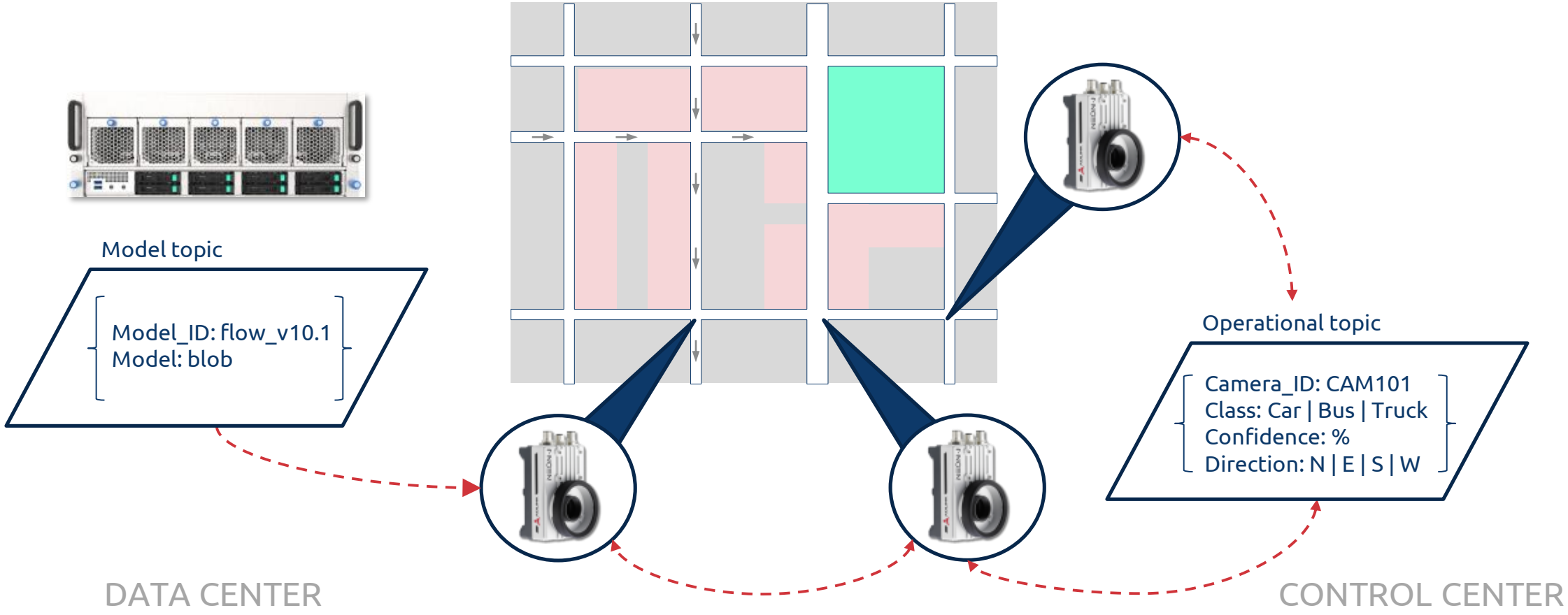
Central command center analysing traffic flow across a city



Traffic flow example

Mesh of intelligent co-operating nodes, “east-west” communication

Data produced by the app is the same; deploy new apps to consume locally

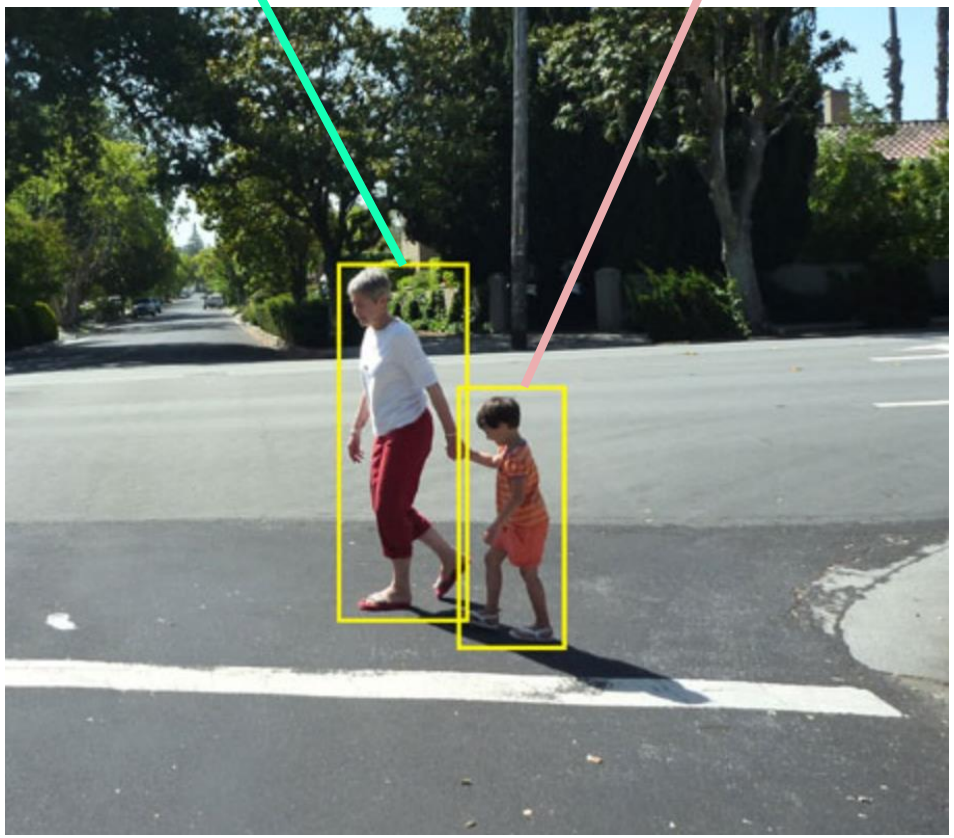


AI in the real world

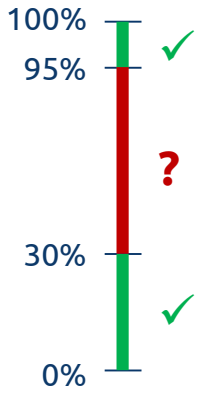
The real world is messy!

Class: **Pedestrian**
Confidence: **98%**

Class: **Pedestrian**
Confidence: **71%**



Confidence



I'm confident of a positive match in the data

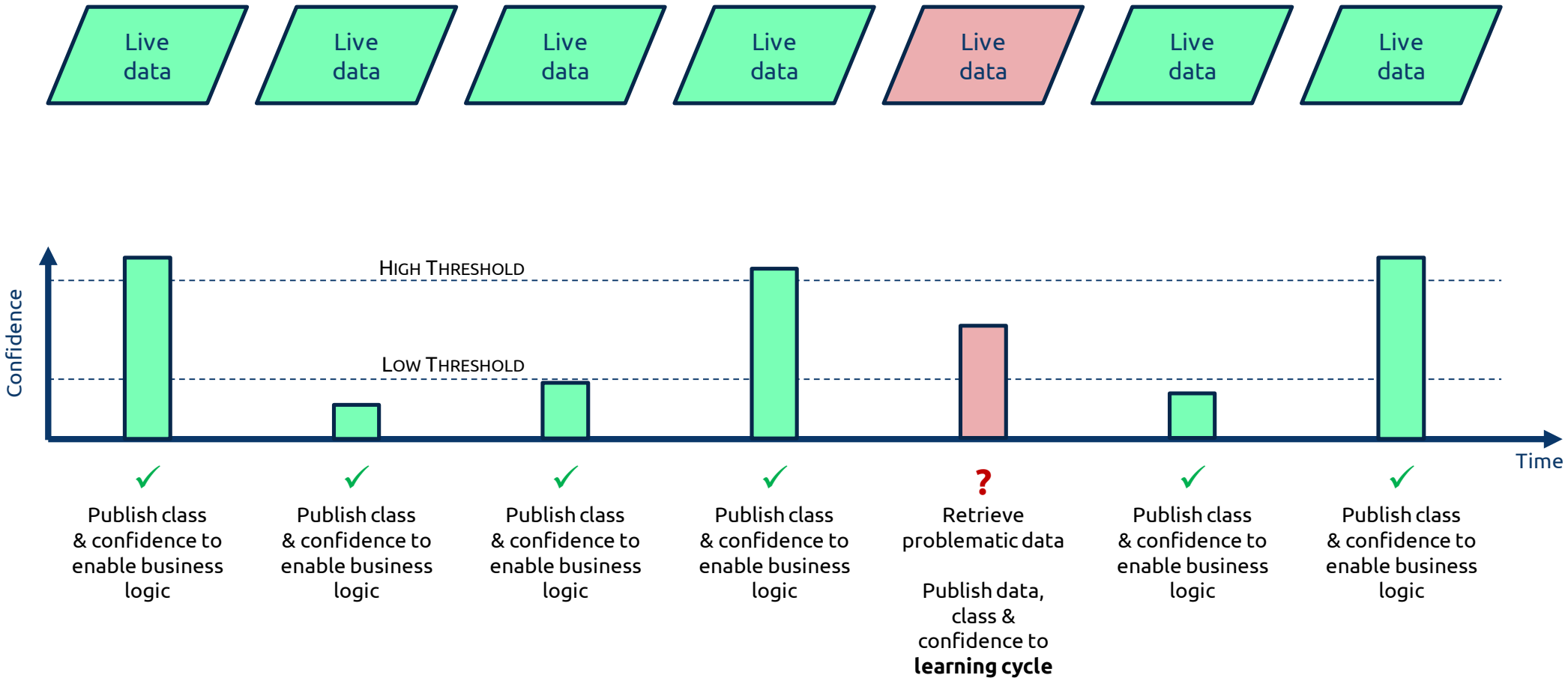
I'm not really sure whether there is a match in the data

I'm confident of no match in the data

Note: high and low threshold points will be specific to each application, and a relevant risk assessment (probability of occurrence / severity of occurrence)

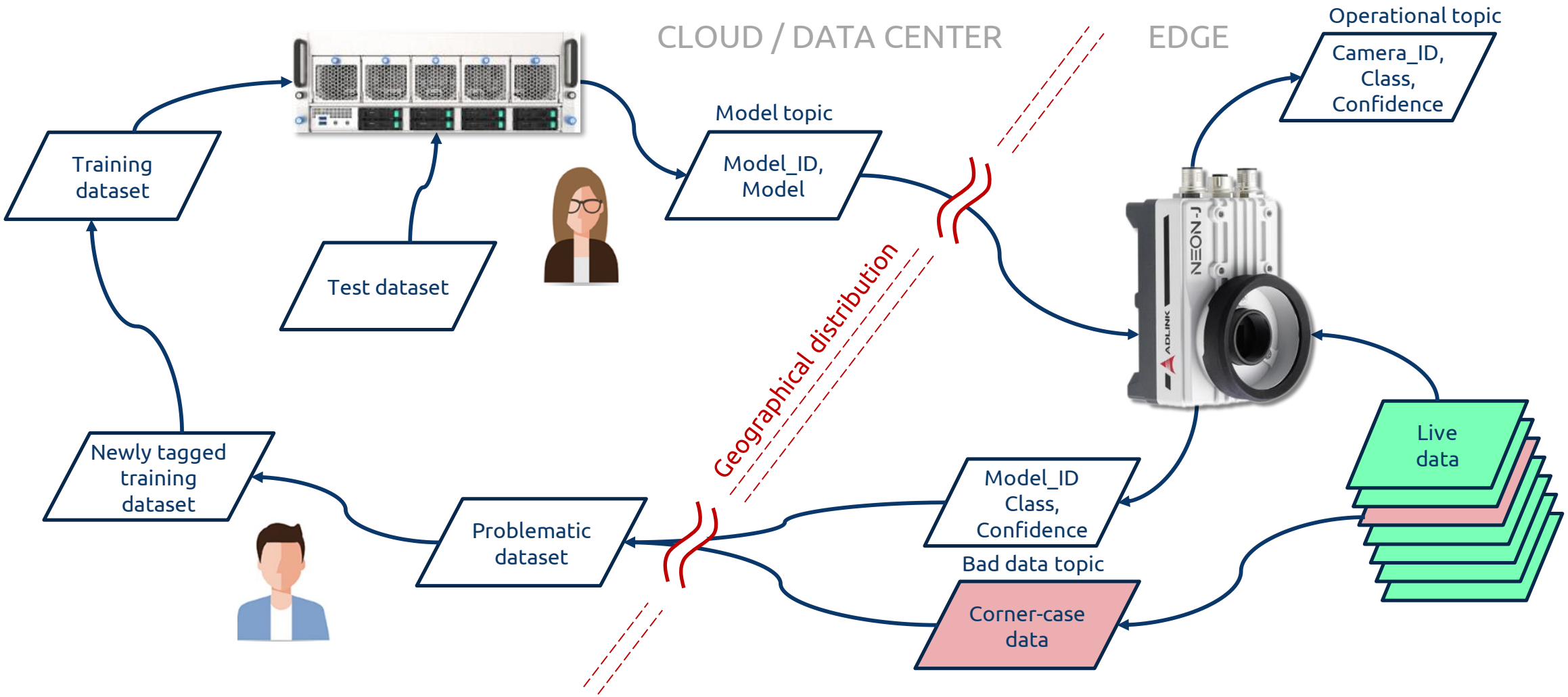
Confidence matters!

For each frame, are we confident to act?



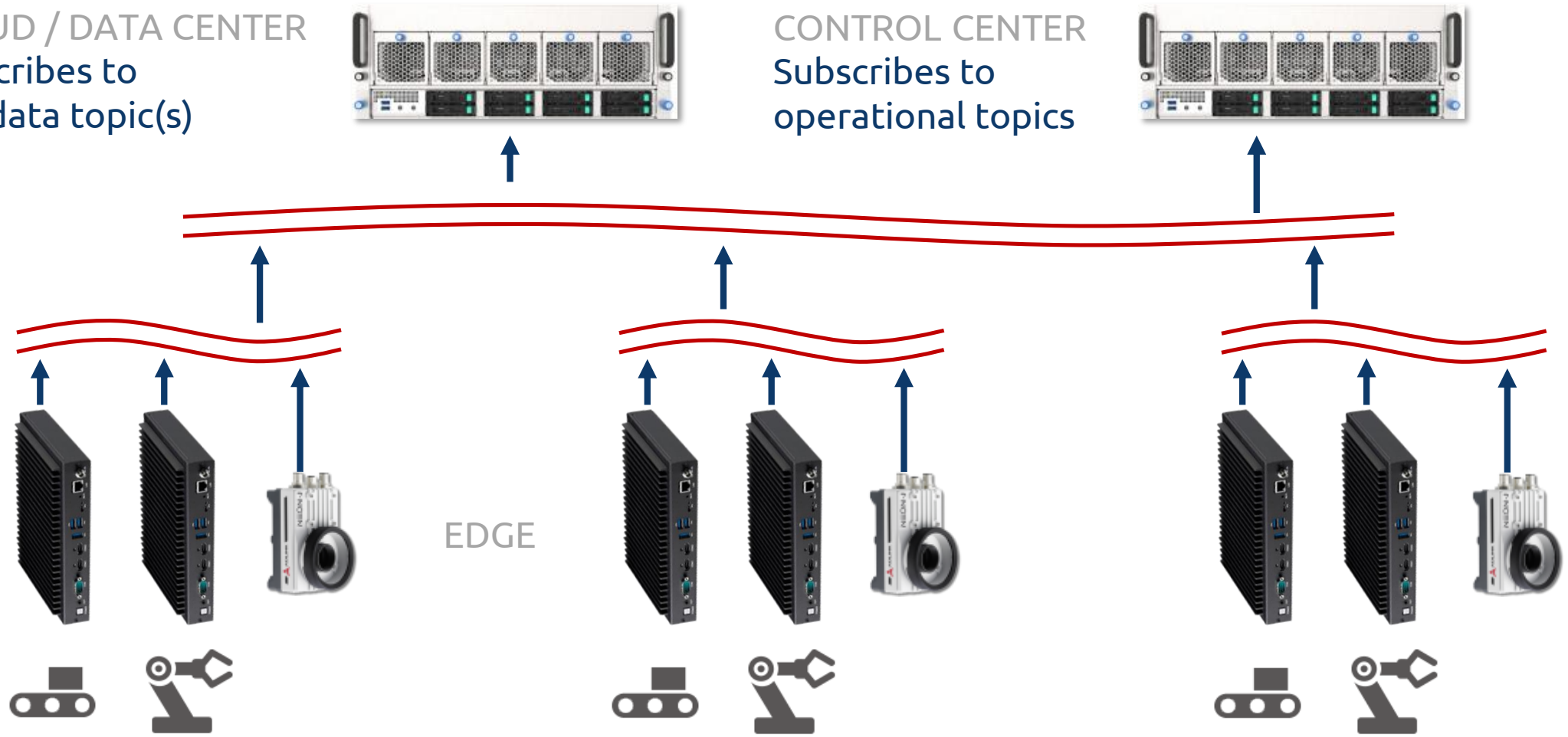
Iterative AI model training use case

Example: supervised learning of classification



Operations & training systems in parallel

Multi-layered network implementation



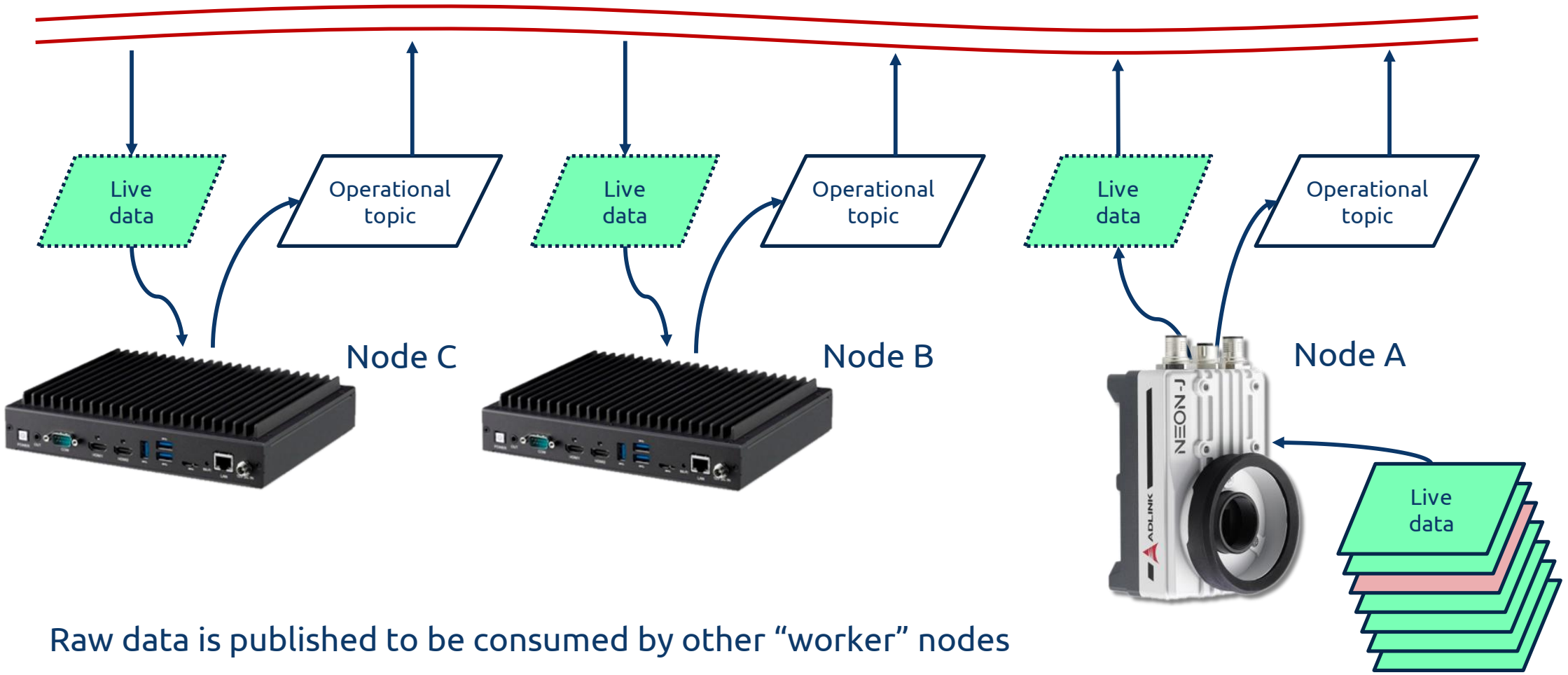
Distributed AI – load balancing

Processing cycle time longer than frame period

Data is dropped & lost if processing node is not available

Distributed AI – load-balancing

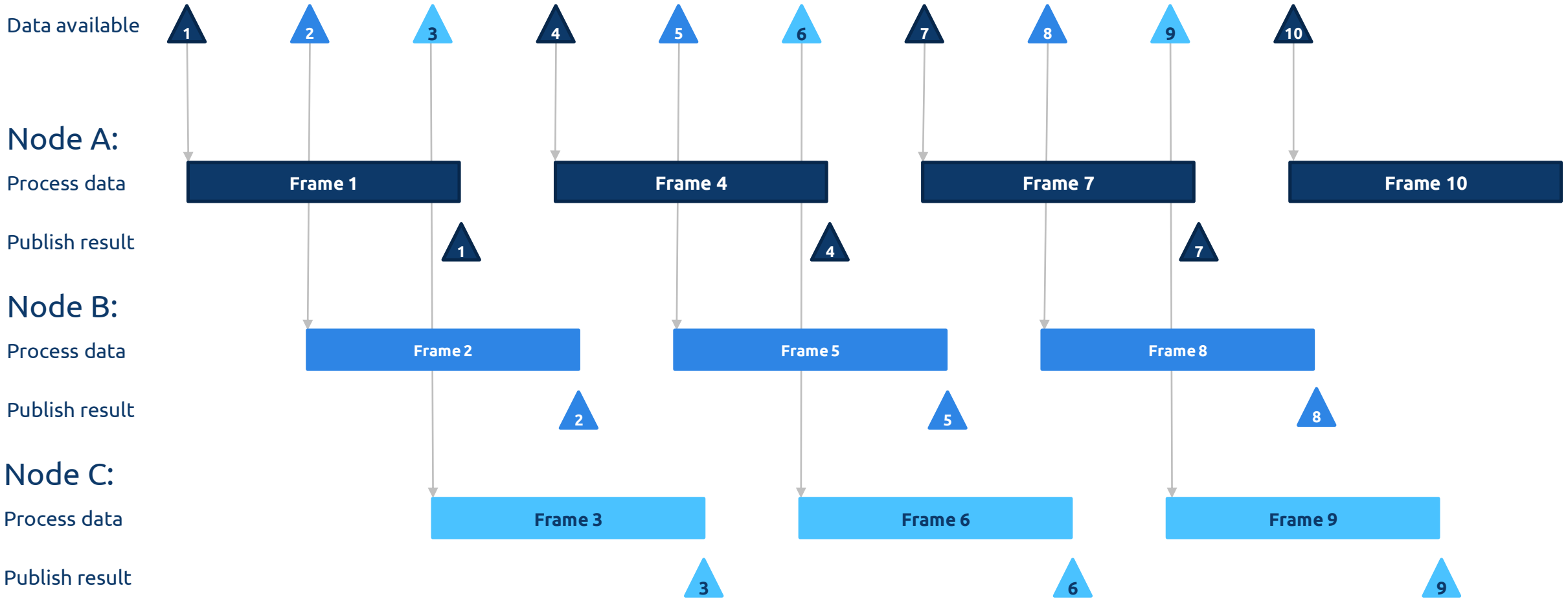
Example: Introduce “worker pattern” to share load



Raw data is published to be consumed by other “worker” nodes

Distributed AI – load balancing

Data processing is interleaved across worker nodes



AI Systems in IoT-Scale Deployments

Summary

Enormous growth fuelled by convergence of technologies

A data-centric solution allows easy and robust scaling to IoT volumes

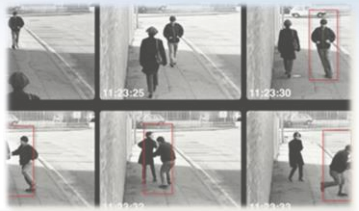
Developers focus on how to best exploit the data, rather than how to connect everything together

Leads to Software Defined Machines, which can be flexibly reconfigured as requirements change over time, supporting an iterative system evolution

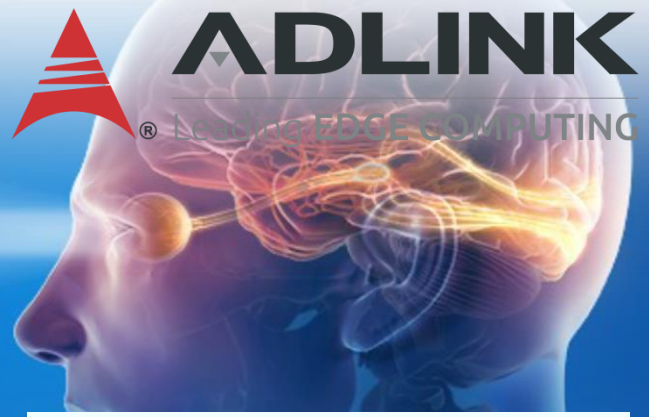
AI at ADLINK

Innovative E2E Solutions powered by Deep-Learning Technology

Digital Surveillance & Security



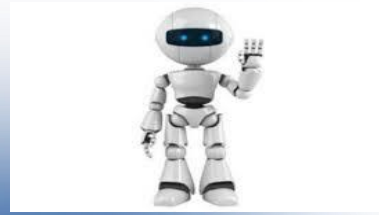
Intelligent Network Video Recorder

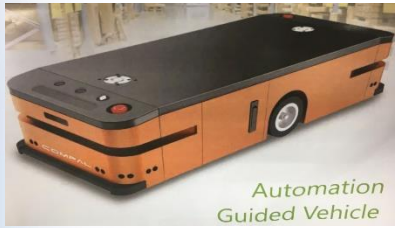


Microsoft AI Factory

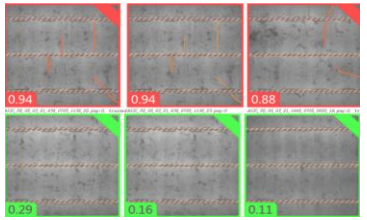
VORTEX EDGE

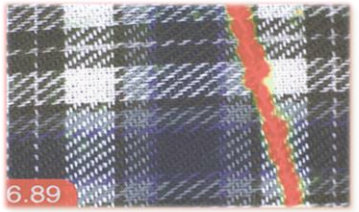
Robotics & AGV





Smart Camera for AOI





Smart Manufacturing

ADLINK®

Leading EDGE COMPUTING

Additional Use Cases

ADLINK
Leading EDGE COMPUTING

Coflight eFDP

European Air Traffic Control

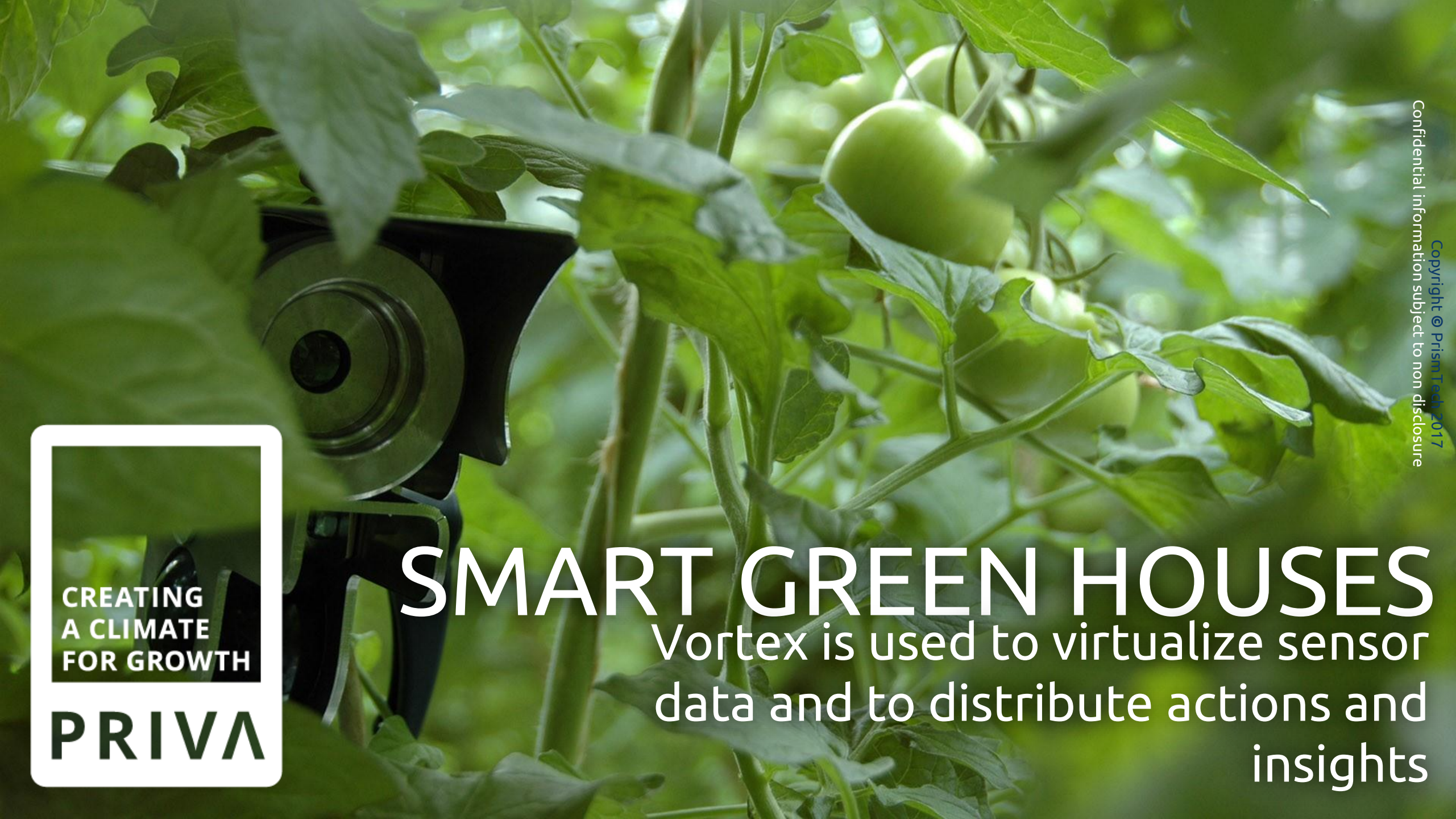
6GB of data in movement
~400 Mbps average Throughput
Pan-European Deployment

"Vortex is at the heart of the EU next-generation Air Traffic Control System"
– CoFlight

DDS is the standard recommended by Eurocontrol/Eurocare

World leader in automated agricultural systems have implemented DDS architecture for real-time control and monitoring of their milking robots


THALES
COMBAT MANAGEMENT
Confidential information subject to non disclosure
Vortex guarantees mission critical reliability and real-time behaviour in large-scale naval combat management systems



SMART GREEN HOUSES

Vortex is used to virtualize sensor data and to distribute actions and insights

LAUNCH SYSTEM

80K+ data points with aggregate updates rate of ~400K msgs/sec