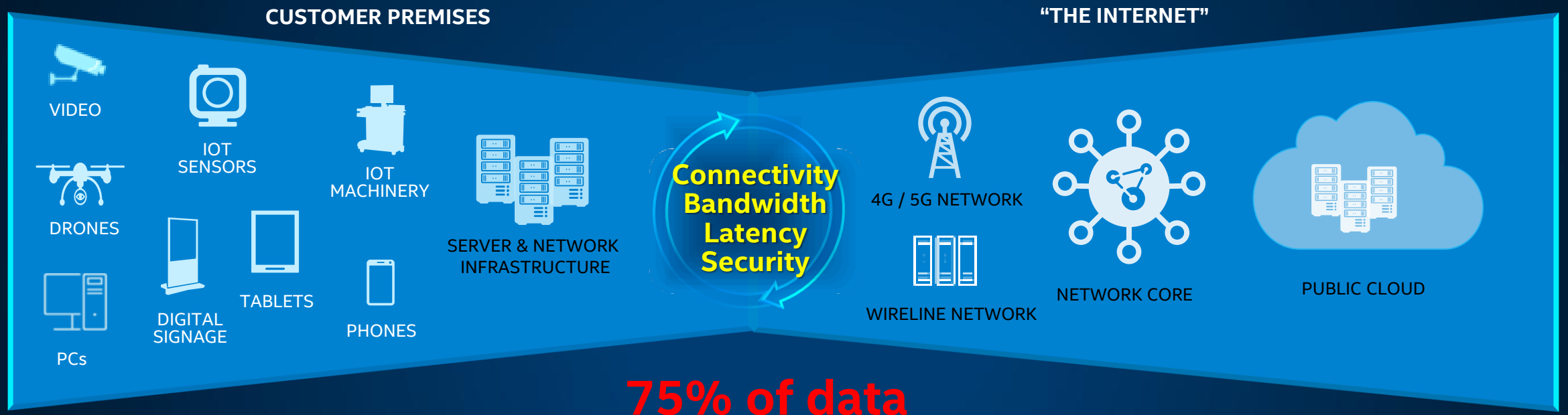
Intel边缘云解决方案与最佳实践

许渊 Intel资深解决方案架构师

EXPANDING IOT/IT/CLOUD TRANSFORMATION INCREASED EDGE PAIN POINTS



75% of data

will be created outside of data centers by 2025*

Connectivity

Insufficient coverage
Per-device cost structure

Bandwidth

Data growth → higher costs

Latency

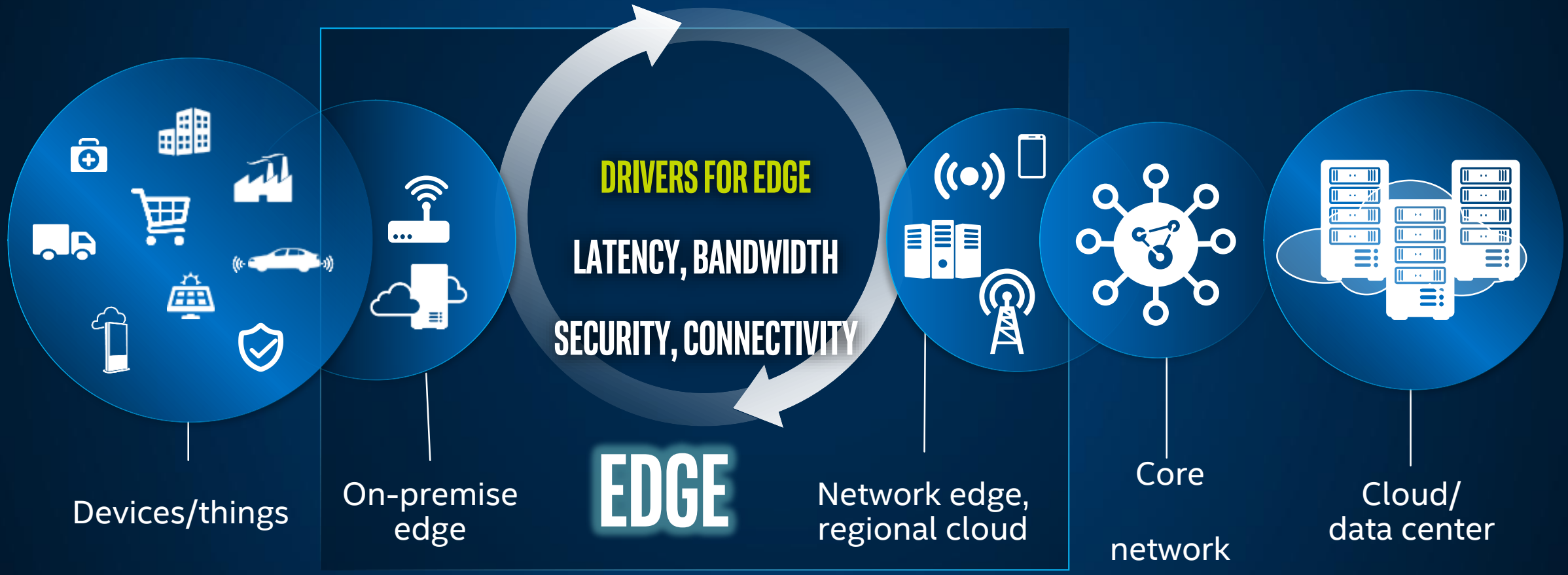
Long routes limit real-time processing

Security

Data traverses' multiple networks and resides "in the cloud"

*Gartner: https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders

THE EDGE: PLACEMENT OF COMPUTE CLOSER TO THE DATA SOURCE OR POINT OF SERVICE DELIVERY



AI and 5G are major accelerants for edge computing

EDGE COMPUTING SCENARIO

- **V** Volume and Cost of Data Transmission
- **L** Latency to/from Cloud
- **C** Lack of Persistent Connectivity
- **P** Privacy: Secure the Data Locally, Isolation from public Data Center
- **W** Workload Orchestration
- **A** Analytics and AI

EDGE CLOUD

MARKET SEGMENTS AND USE CASES

VISION INSPECTION

- L V A Defect Detection, Anomaly Detection
- L A P Instruments Reading, Object Identification

INDUSTRIAL

- L A Predictive Maintenance
- W L P Factory Automation

HEALTHCARE

- V P Clinical Fusion

SMART CITY

- W C Traffic Monitoring

RETAIL

- W A Personalized

TELCO

- L W A MEC

INTEL DIVERSE PRODUCT PORTFOLIO

BUILDING FOR THE EDGE

COMPUTER VISION | MANAGEA-BILITY | FUNCTIONAL SAFETY | REAL TIME | RAN | CORE | SECURITY | NETWORK | RELIABILITY | EXTENDED TEMP | LONG LIFE SUPPORT | VIRTUAL-IZATION

PROCESS

eASIC | FPGA | Xe | IRIS PLUS GRAPHICS | MOVIDIUS | HABANA

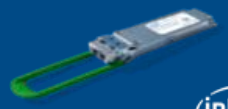
ATOM | CORE | XEON

MOVE

ETHERNET

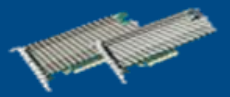
BAREFOOT NETWORKS | an Intel company

Intel® Wi-Fi 6



SILICON PHOTONICS

QUICKASSIST



STORE

OPTANE™ DC SOLID STATE DRIVE

OPTANE™ DC PERSISTENT MEMORY

Delivering AI compute from cloud to edge

oneAPI | OpenVINO | Caffe | GLOW | mxnet | ONNX RUNTIME | PaddlePaddle | PYTORCH | TensorFlow

CPU | SCALAR

Intel Xeon Scalable processors

GPU | VECTOR

Intel Discrete Graphics

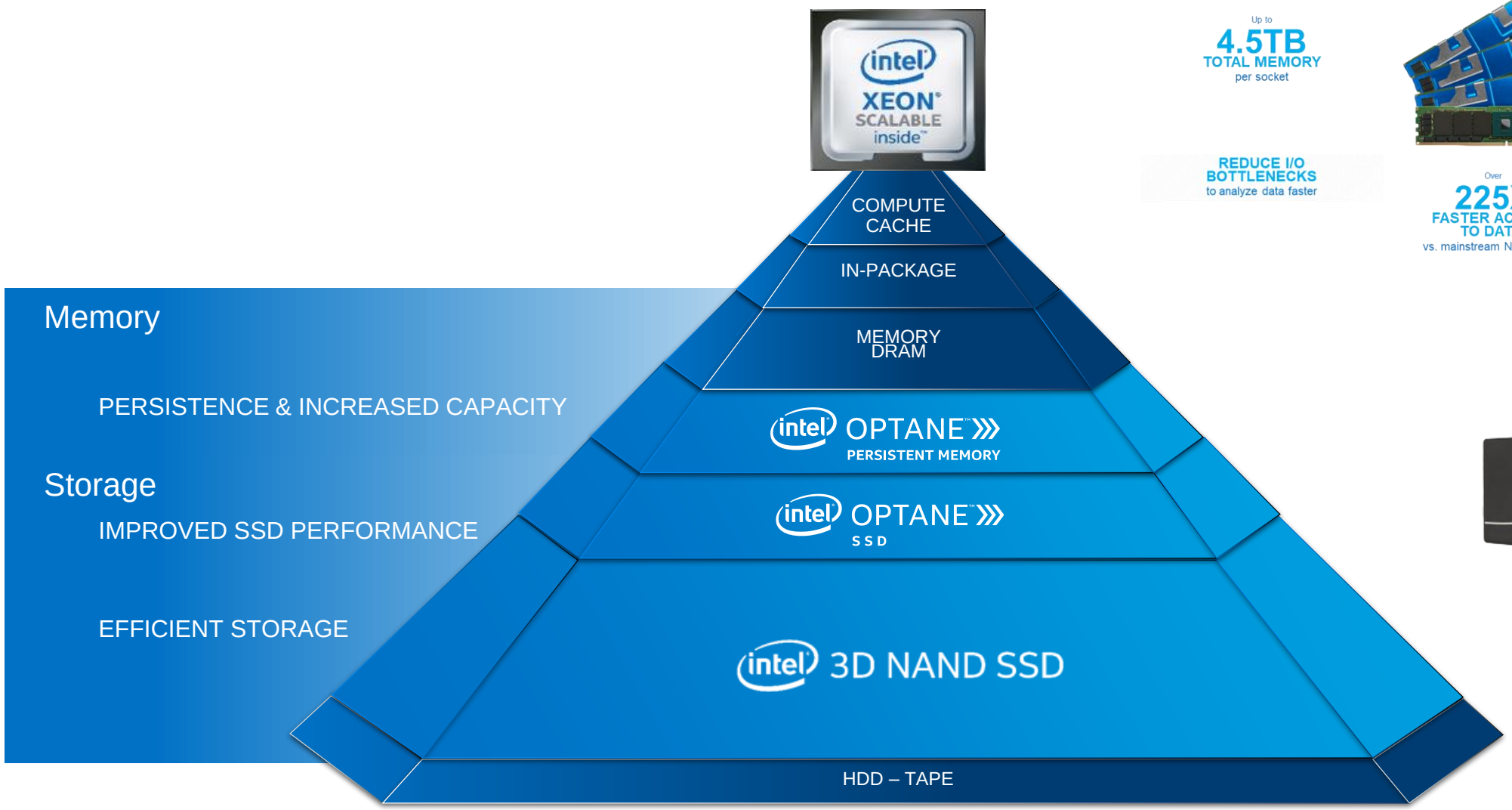
FPGA | SPATIAL

Intel FPGA

SPECIALIZED | MATRIX

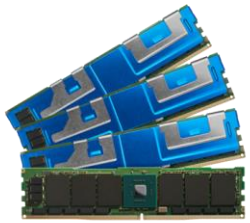
Habana
Intel Movidius Myriad
Intel Mobileye EyeQ

Transforming memory & storage



Up to
4.5TB
TOTAL MEMORY
per socket

REDUCE I/O BOTTLENECKS
to analyze data faster



Over
225X
FASTER ACCESS TO DATA[8]
vs. mainstream NAND SSD

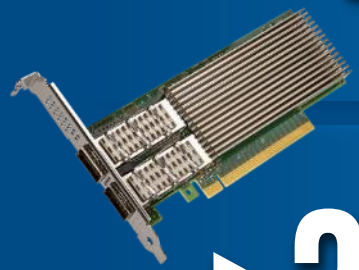
Average
25%
HIGHER MEMORY BANDWIDTH[7]
vs. prior gen

BOOST APPLICATION PERFORMANCE

INTEL® ETHERNET 800 SERIES
WITH APPLICATION DEVICE QUEUES (ADQ)

Application Device Queues

– An application-specific queuing and steering technology
– Filters application traffic to a dedicated set of queues



>45% LATENCY REDUCTION[11]
WITH OPEN SOURCE REDIS USING 2ND GEN INTEL® XEON® SCALABLE PROCESSORS AND INTEL® ETHERNET 800 SERIES WITH ADQ

>30% THROUGHPUT IMPROVEMENT[12]
WITH OPEN SOURCE REDIS USING 2ND GEN INTEL® XEON® SCALABLE PROCESSORS AND INTEL® ETHERNET 800 SERIES WITH ADQ

BETTER TOGETHER

CONSISTENTLY IMPROVED SERVICES AND SLA
WITH 2ND GEN INTEL® XEON® SCALABLE PROCESSORS AND INTEL® ETHERNET 800 SERIES

- INCREASES APPLICATION PREDICTABILITY
- REDUCES APPLICATION LATENCY
- IMPROVES APPLICATION THROUGHPUT
- CONSISTENT APPLICATION RESPONSE TIME WITH ADQ

For more complete information about performance and benchmark results, visit www.intel.com/benchmarks. See configuration slide 50 for details. For more information regarding performance and optimization choices in Intel software products., please visit https://software.intel.com/en-us/articles/optimization-notice.

EDGE CLOUD SOLUTION ARCHITECT

DEVICE

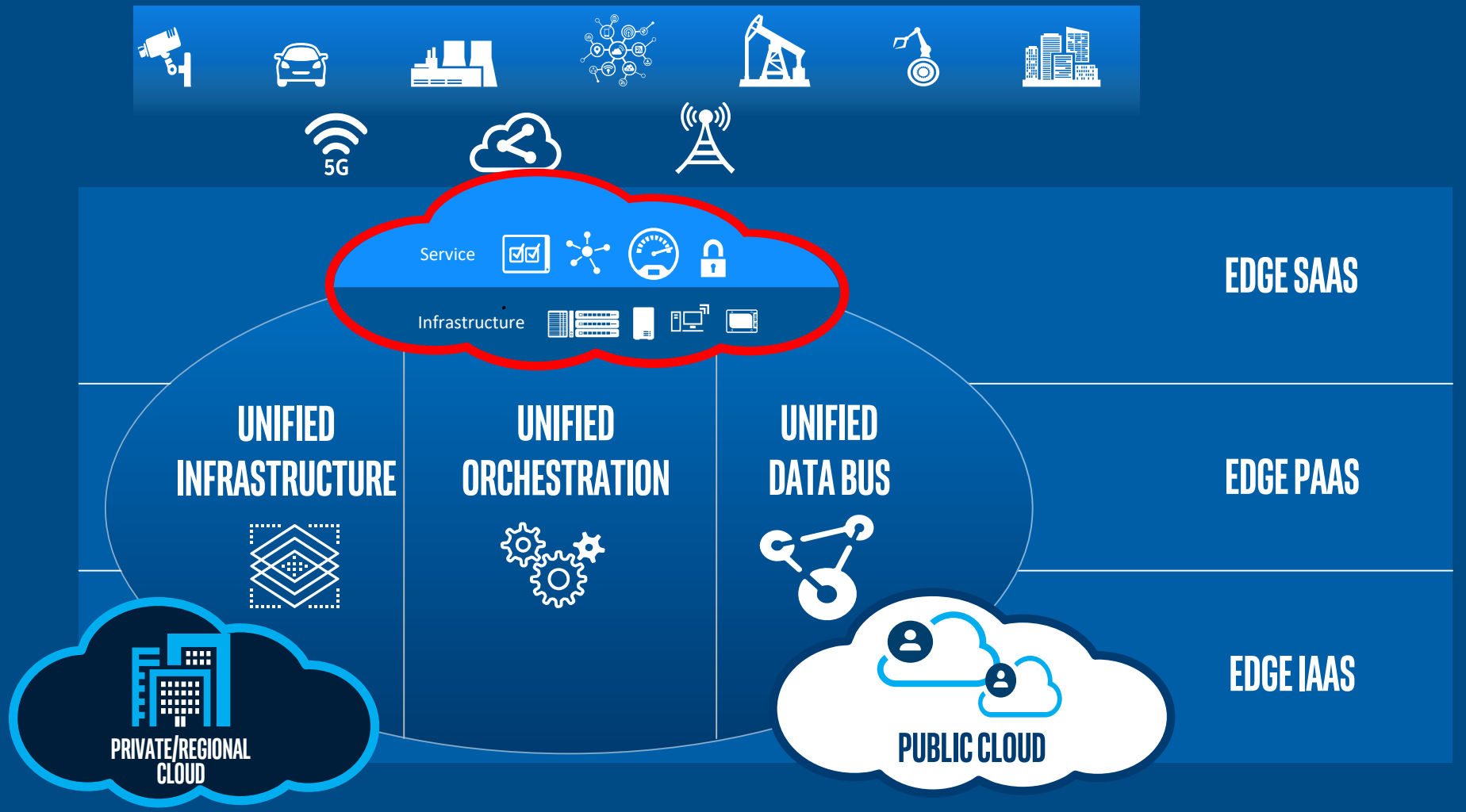
Generate Data

EDGE

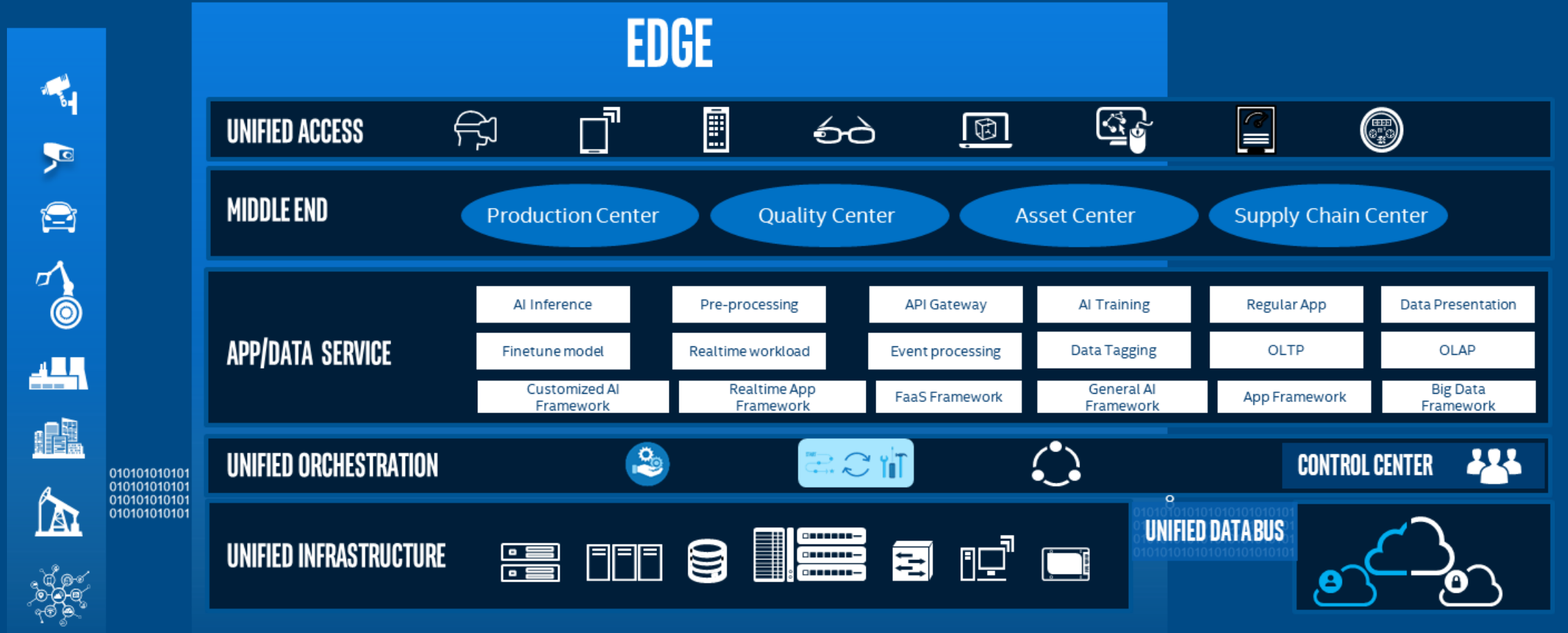
First tier Data Governance

CLOUD

Full Data Process/Store Control Center

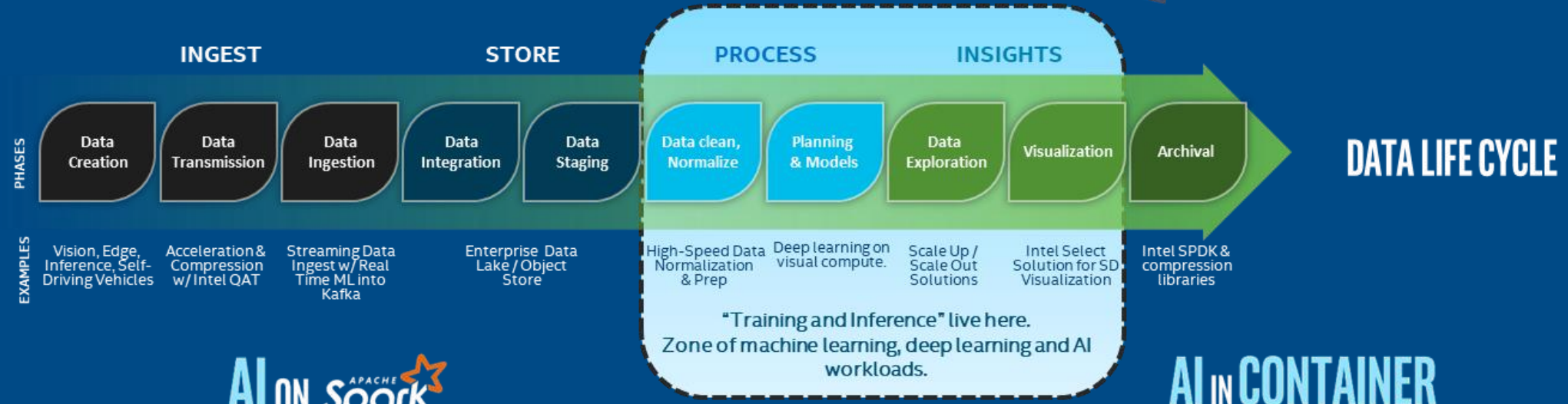


EDGE SIDE ARCHITECT



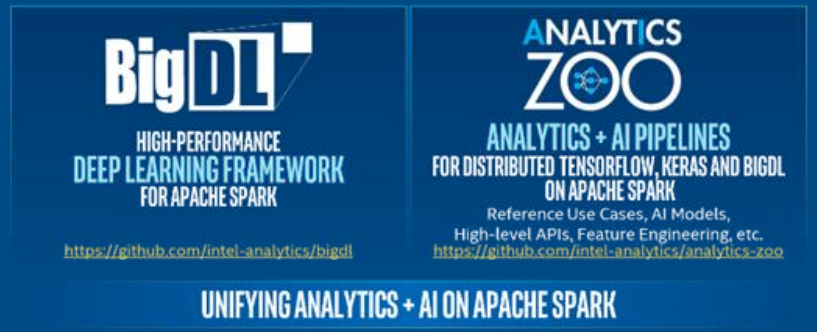
EDGE CLOUD AI PLATFORM

INFERENCE

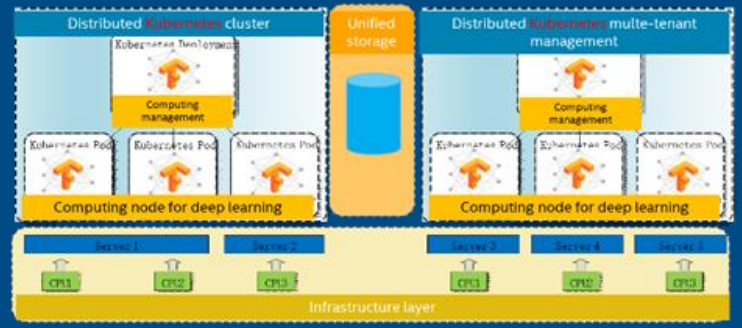


TRAINING

AI ON Apache Spark

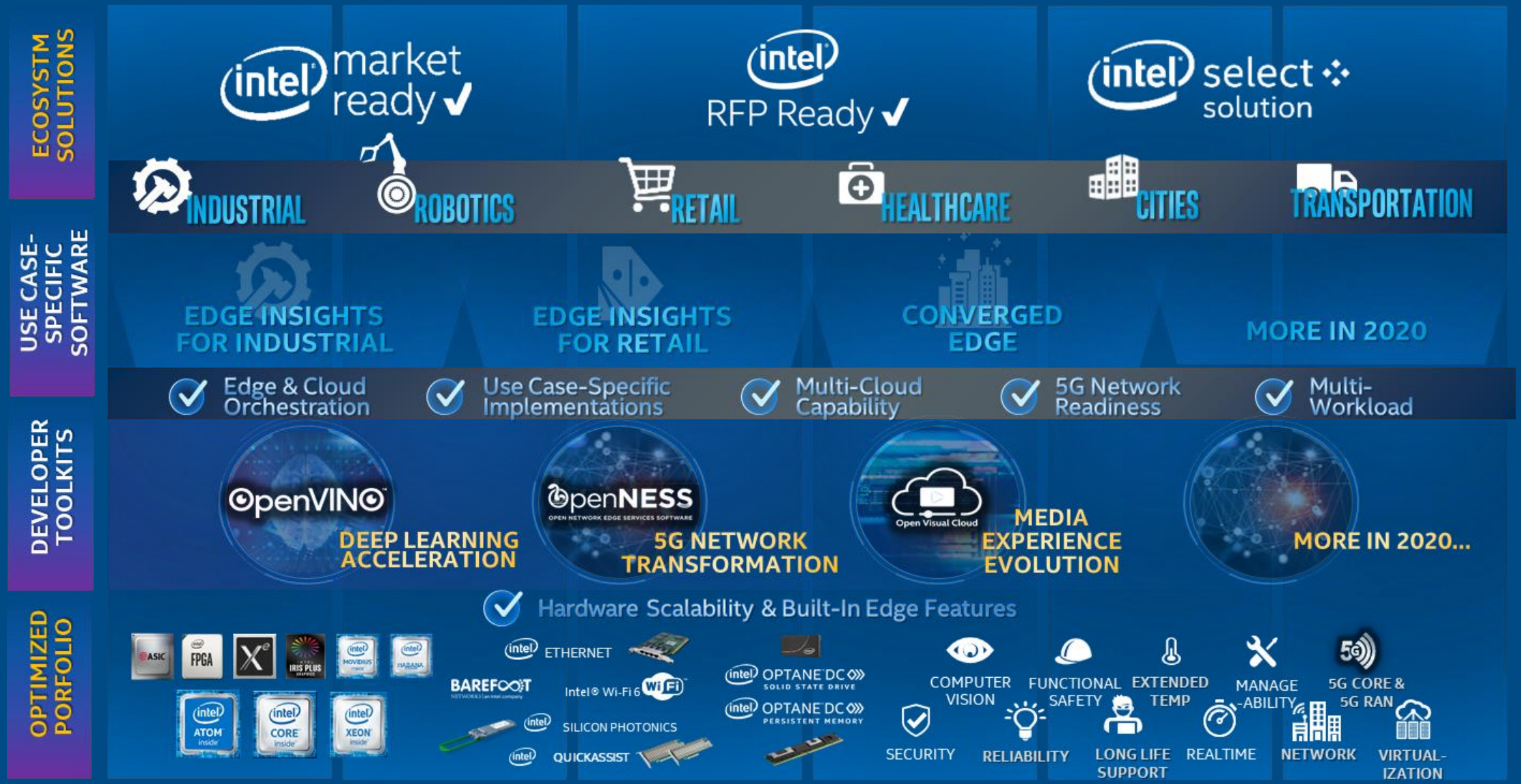


AI IN CONTAINER



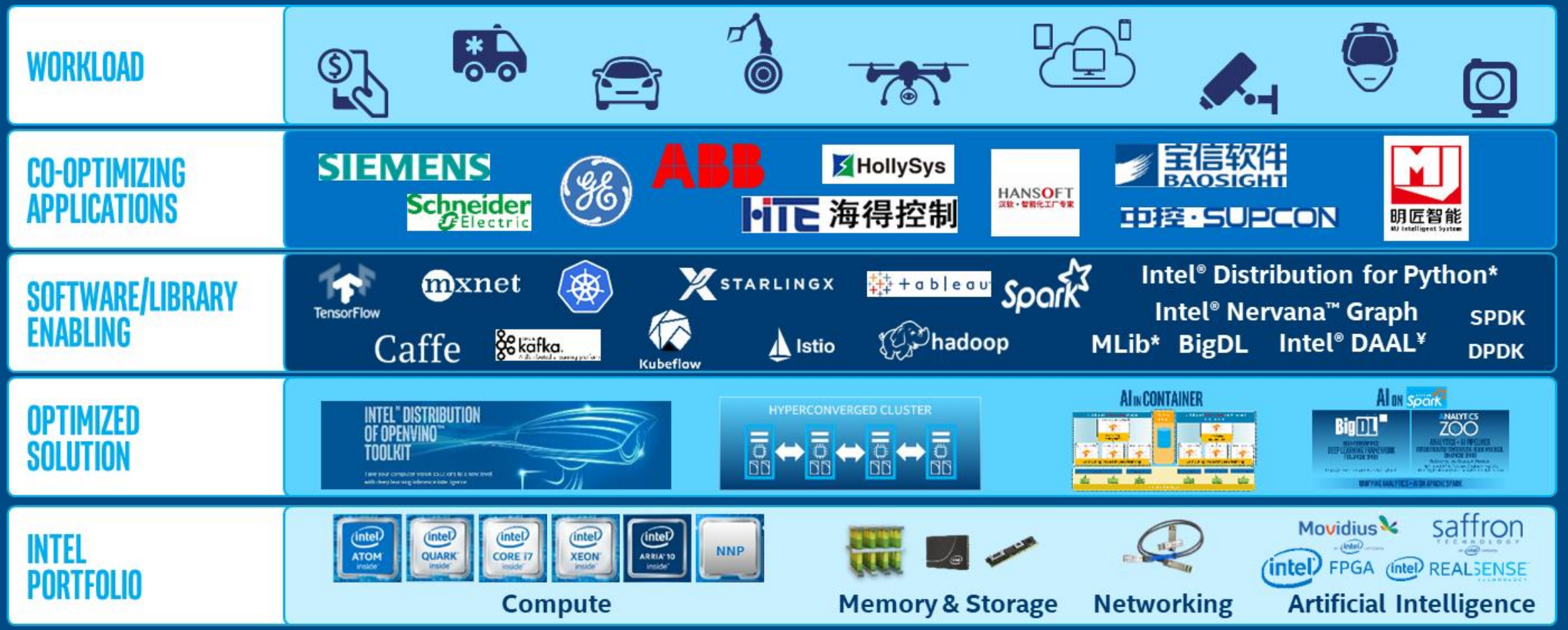
*Other names and brands names may be claimed as the property of others.

EDGE SOFTWARE OFFERINGS



*Other names and brands names may be claimed as the property of others.

EDGE CLOUD SOLUTION LANDSCAPE



¥Note: Intel® Data Analytics Acceleration Library, Intel® Math Kernel Library, Intel® Math Kernel Library for Deep Neural Networks, BigDL: Distributed Deep Learning on Apache Spark*, MLib: Apache Spark's Scalable Machine Learning Library
*Other names and brands may be claimed as the property of others.


intel