

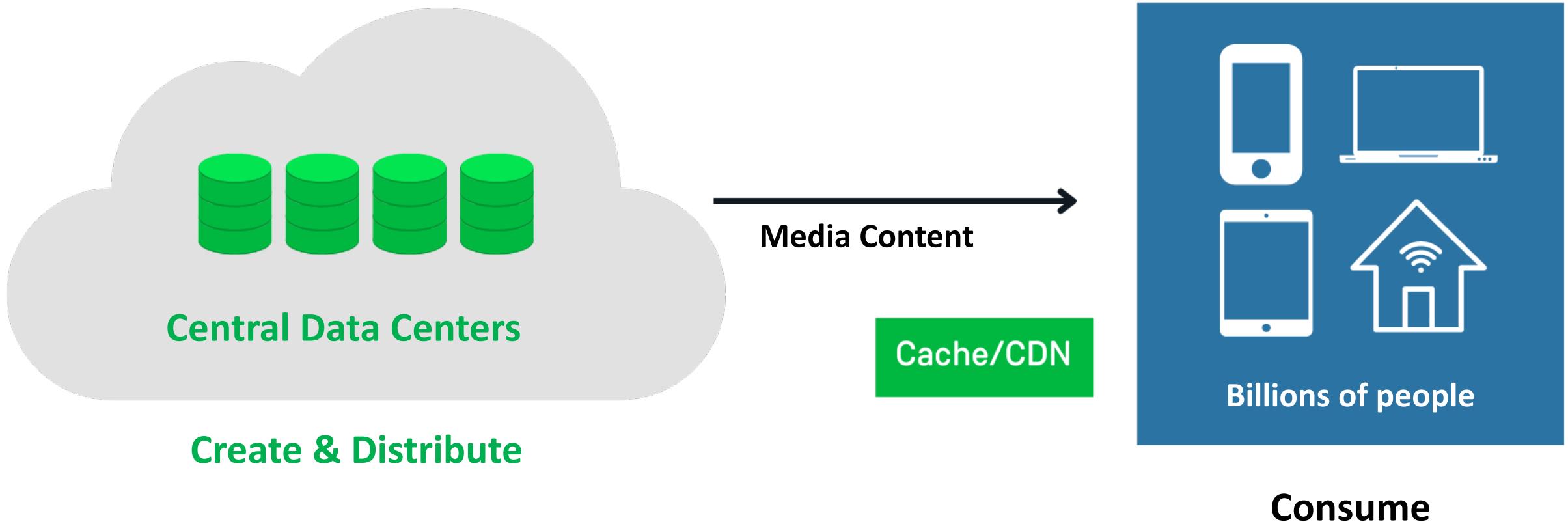


Hardware at the Edge: Selecting Optimal Technology for AI Applications at the Edge

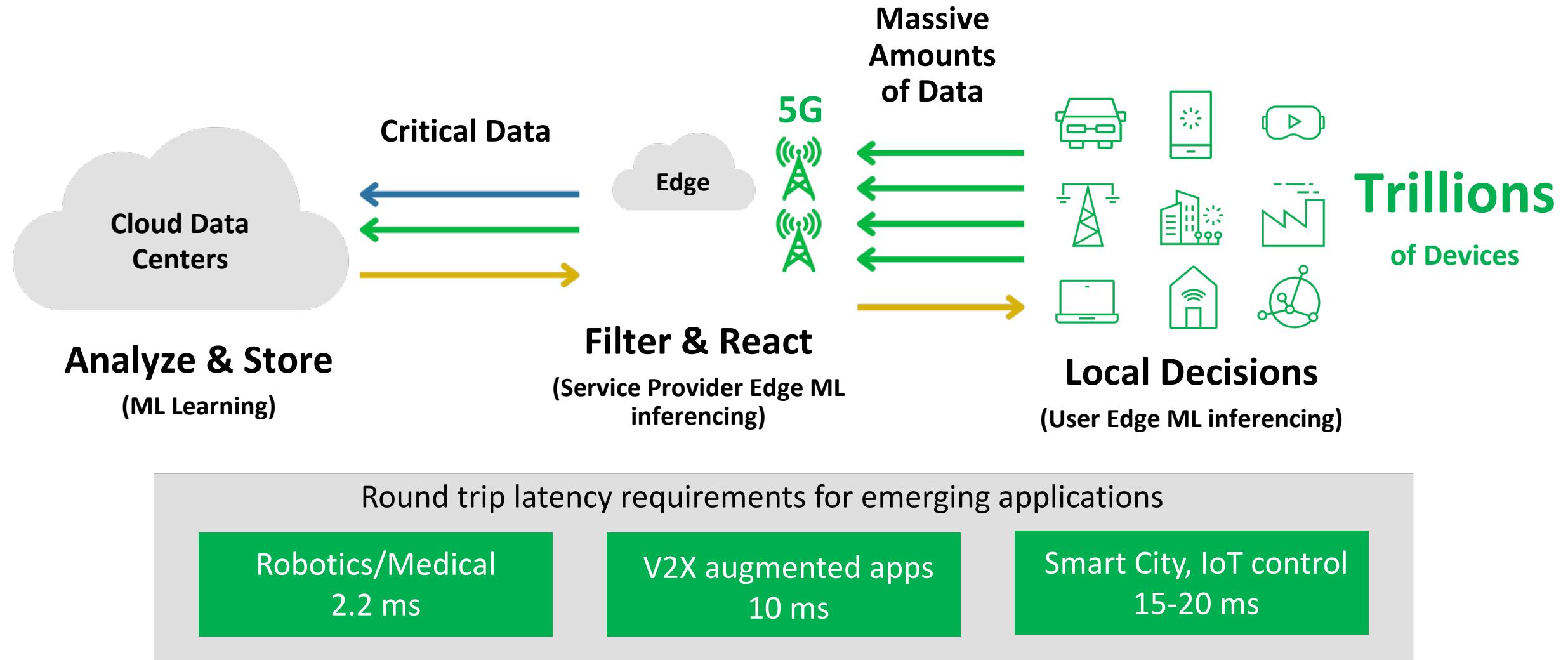
Colin Alexander,
Director of Product Marketing



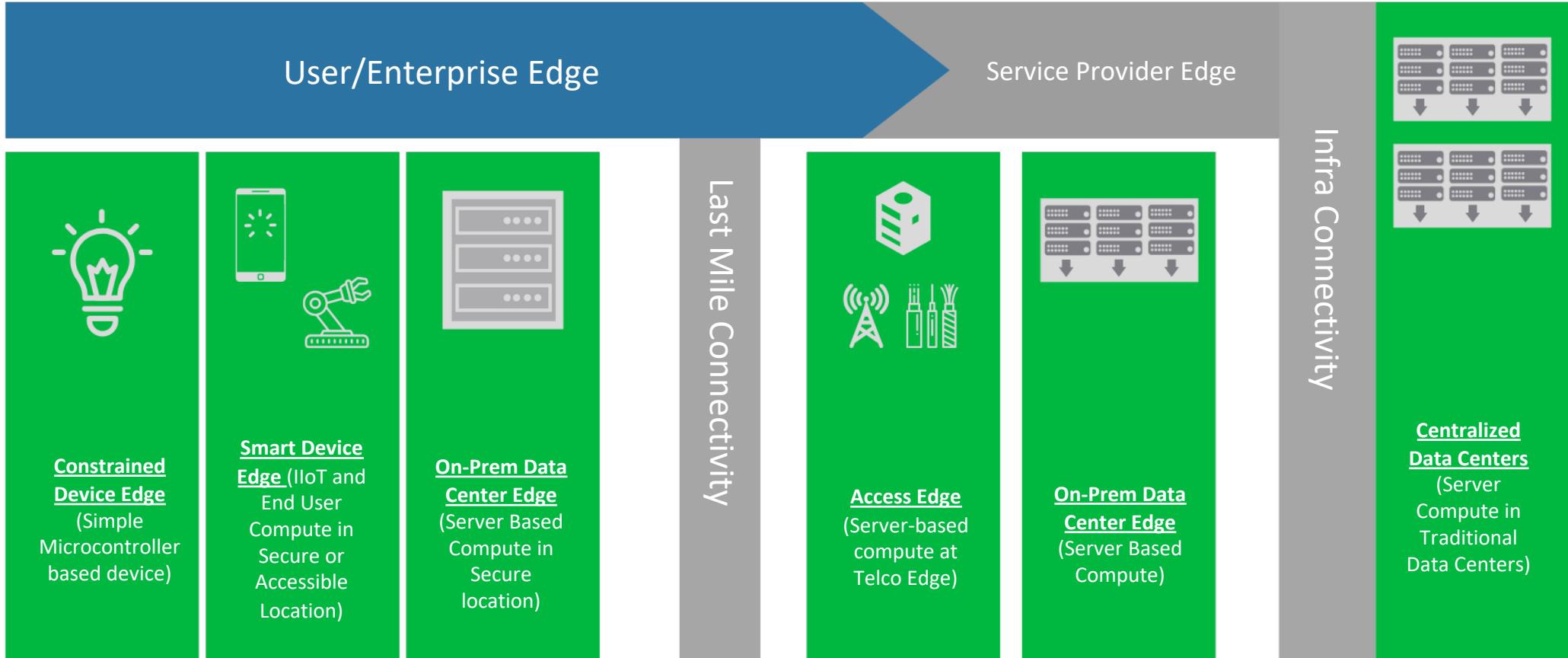
Today's Compute Model



Servicing New Use Cases at the “Edge”

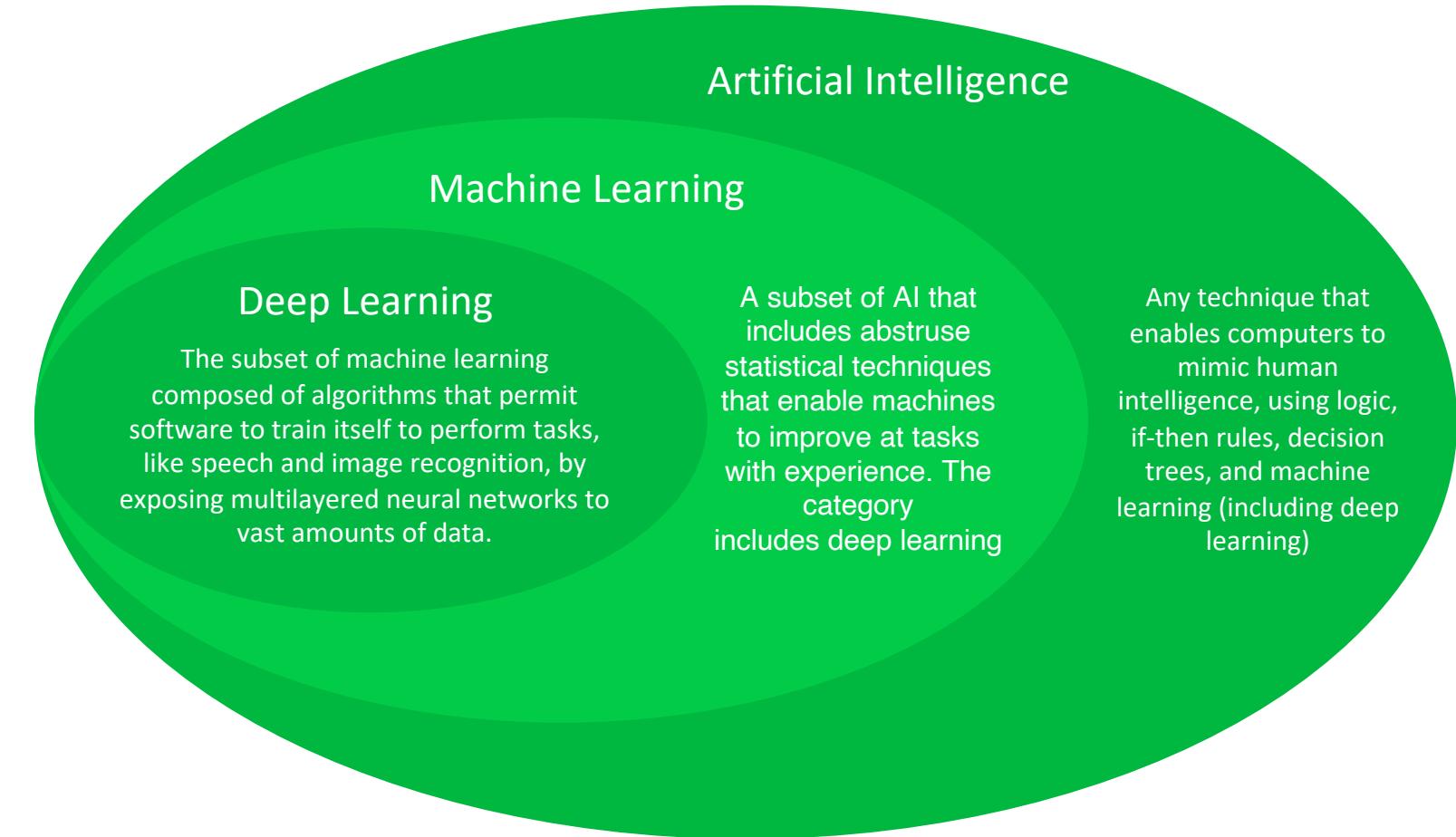


Pressures on the Infrastructure – the New “Edge”



Source: Linux Foundation Edge

Use Cases Driving of Artificial Intelligence, Machine & Deep Learning



Edge Infrastructure Use Case Drivers by highest growth rate:

- Manufacturing
- Smart Grid
- Automotive
- Smart Cities
- Commercial UAV's
- Health Care

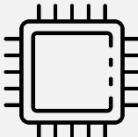
Other major areas:

- Service Provider networks
- Mobile consumer
- Retail
- Residential/Home

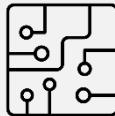
Acceleration for AI/ML Edge processing



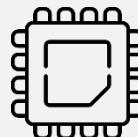
Compute/ ML Processing Platform



FPGA



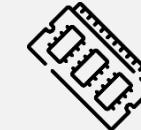
ASIC



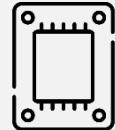
CPU



GPU



RAM



SSD

Link with ML Inferencing

Transport



Voice



Image



Video



GPS



Search



Text



Medical



Weather



Robotics



Auto



Factory



Traffic

Dis-aggregation of platforms to support inferencing at the edge



Data Transport



Memory Bandwidth



Compute Power

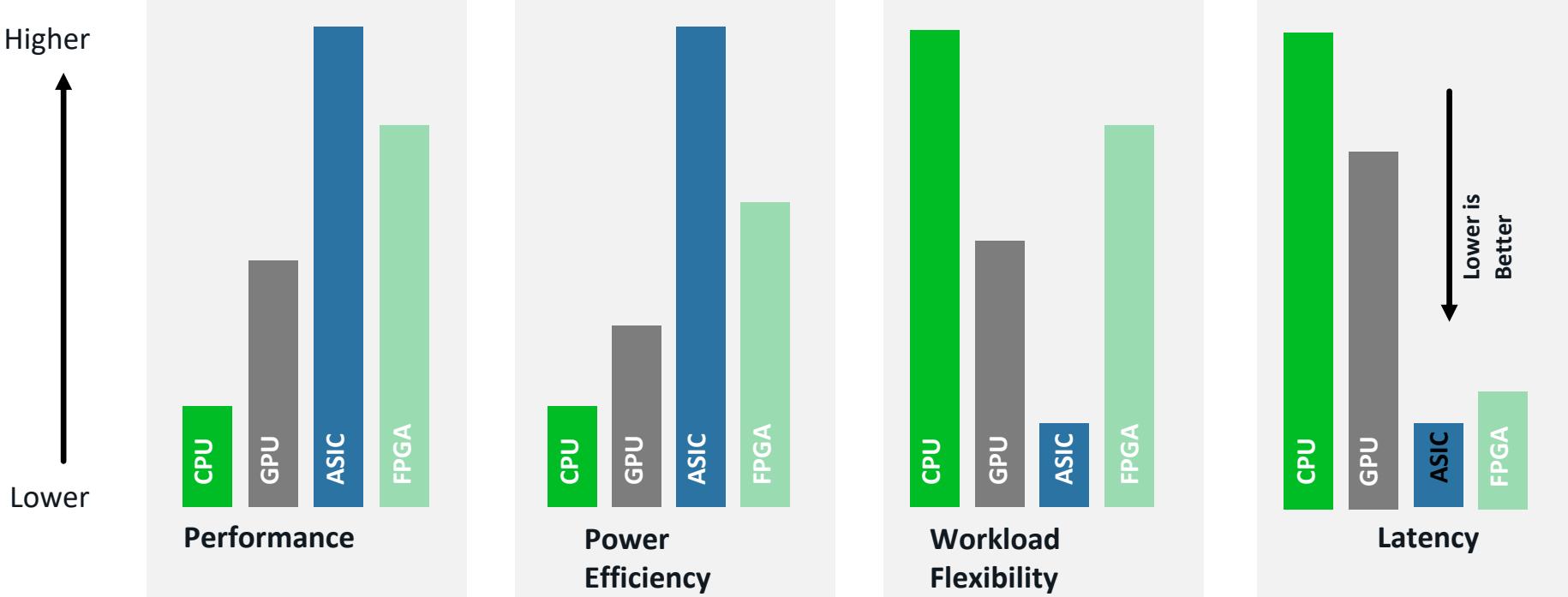


Programmability



Connectivity

Technology Choices for implementing AI/ML at the Edge



CPU GPU ASIC FPGA

Flexibility

- easily modify ML algorithms

Low latency for mission-critical applications

- Support for fine-grained, bit-level operations

Power savings

- A few clock cycles on an FPGA can equal 1000s of clock cycles on GPUs

Data flow

- No limitations on how data flows through device unlike CPUs and GPUs

Achronix FPGA Machine Learning Solutions

ML Data Acceleration Requirements

1 COMPUTE POWER

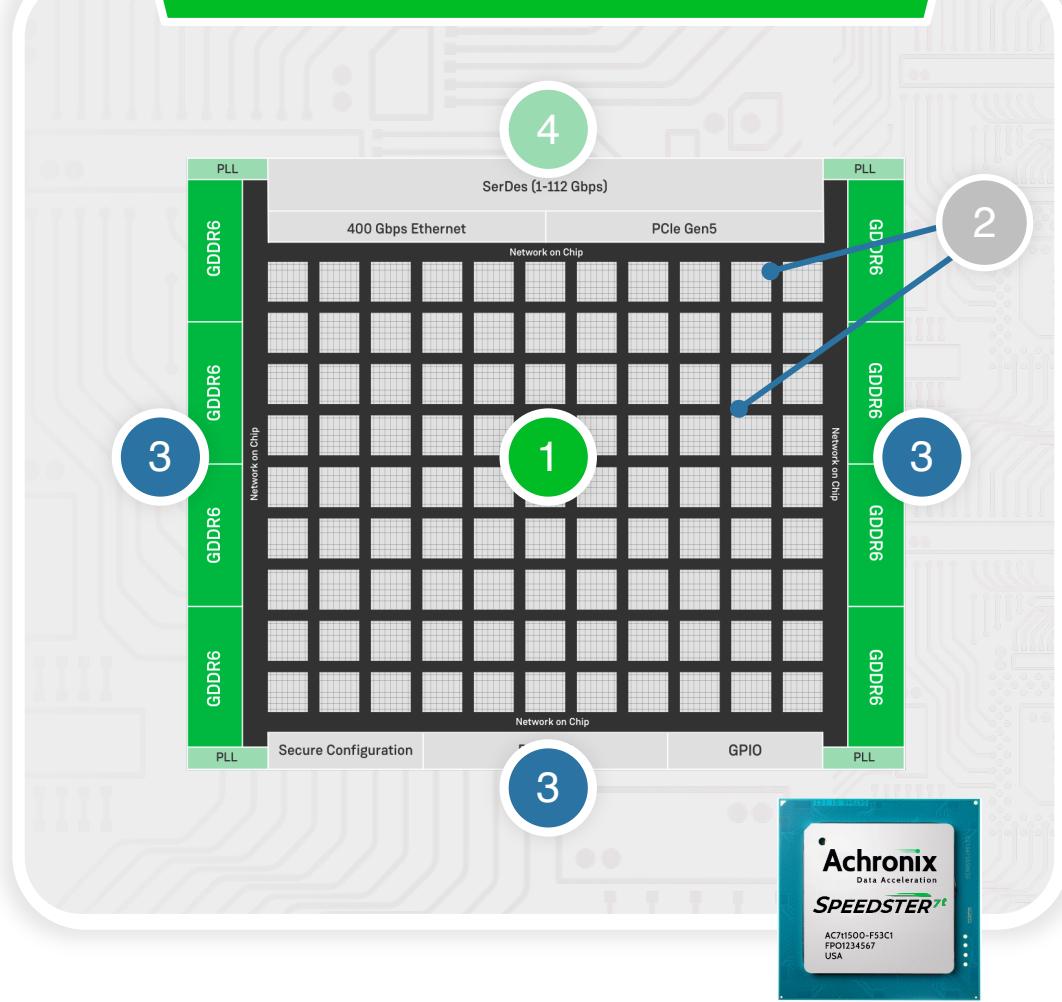
2 DATA TRANSPORT

3 HIGH BANDWIDTH MEMORY INTERFACES

4 Fast I/O

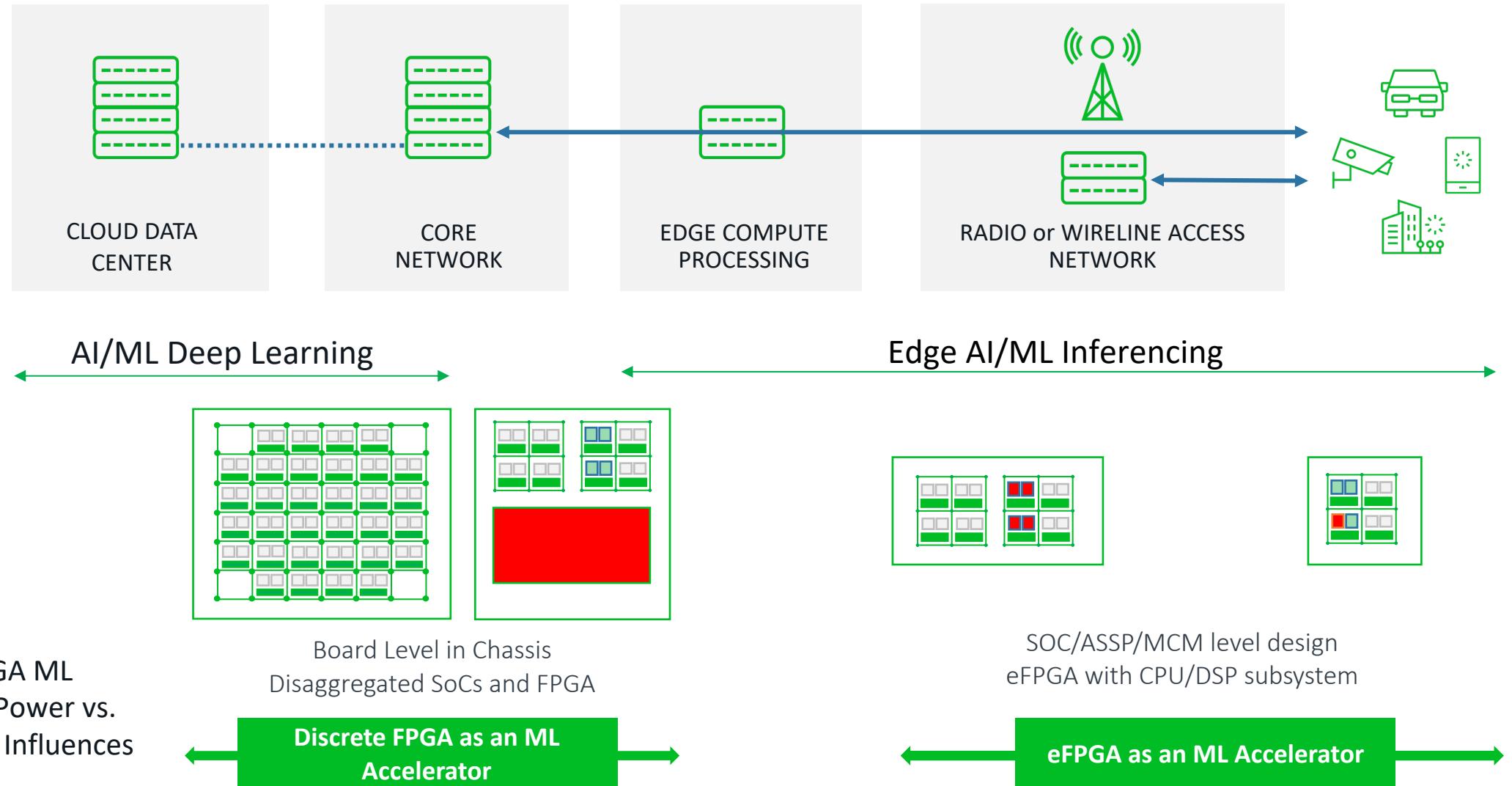
5 SOFTWARE

Standalone FPGA - Speedster7t overview



- 1 - Matrix/Vector Manipulation processors
- Tightly coupled memory
- FPGA fabric
- 2 - 2D NOC interconnect between I/O, memory and fabric
- 3 - High Speed Flexible memory interfaces
- 4 - Multiple 10/100/400Gbps ethernet interfaces
- Multiple Serdes
- PCIe Gen 5
- 5 - Flexible software
- ML Embedded and Turnkey applications

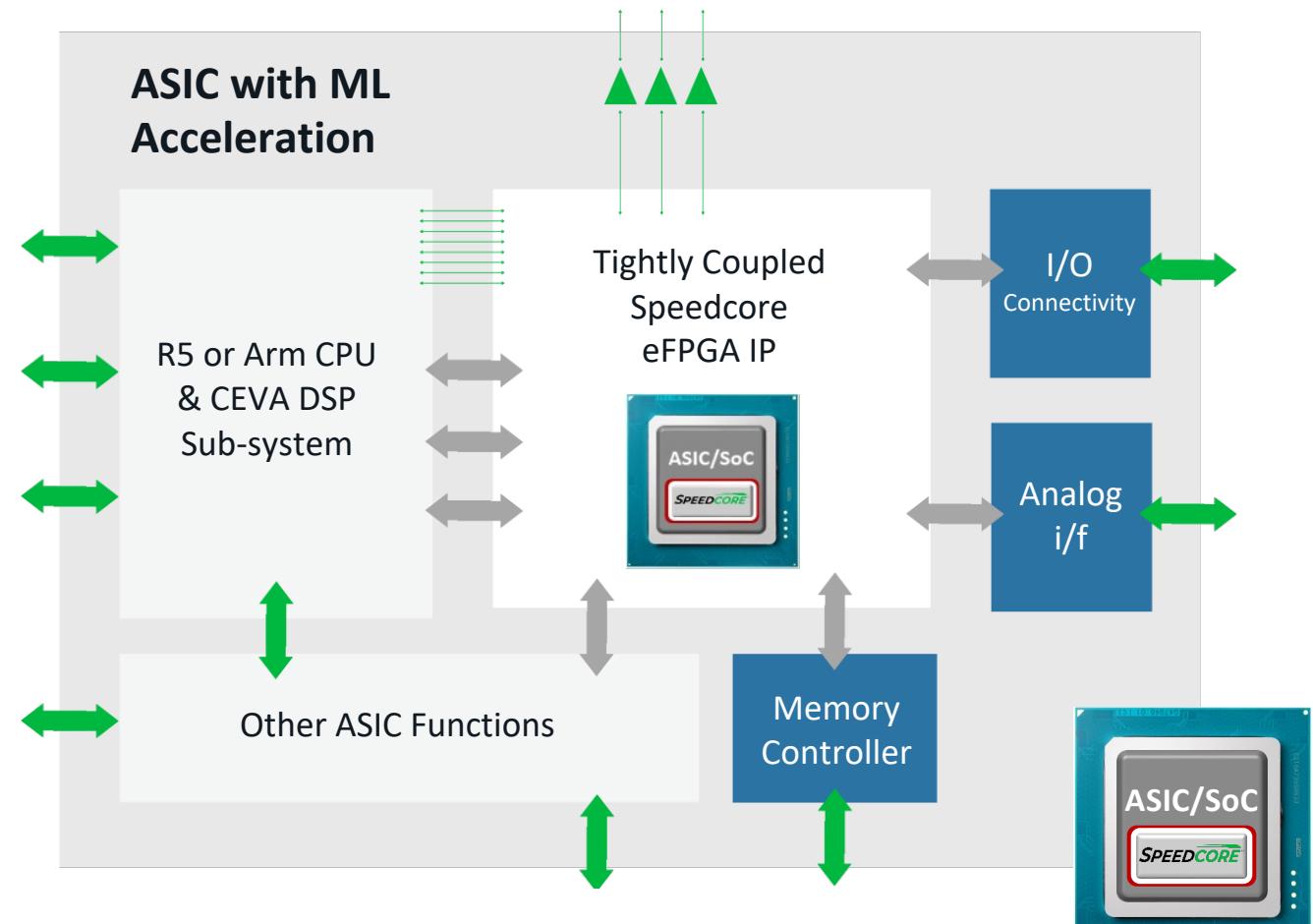
Diverse workloads influence ML platform choice



FPGA/eFPGA usage for Edge ML Inferencing



Stand alone FPGA as the basis of a flexible
“Smart NIC” type platform for Service Provider
Edge and Cloud ML Inferencing

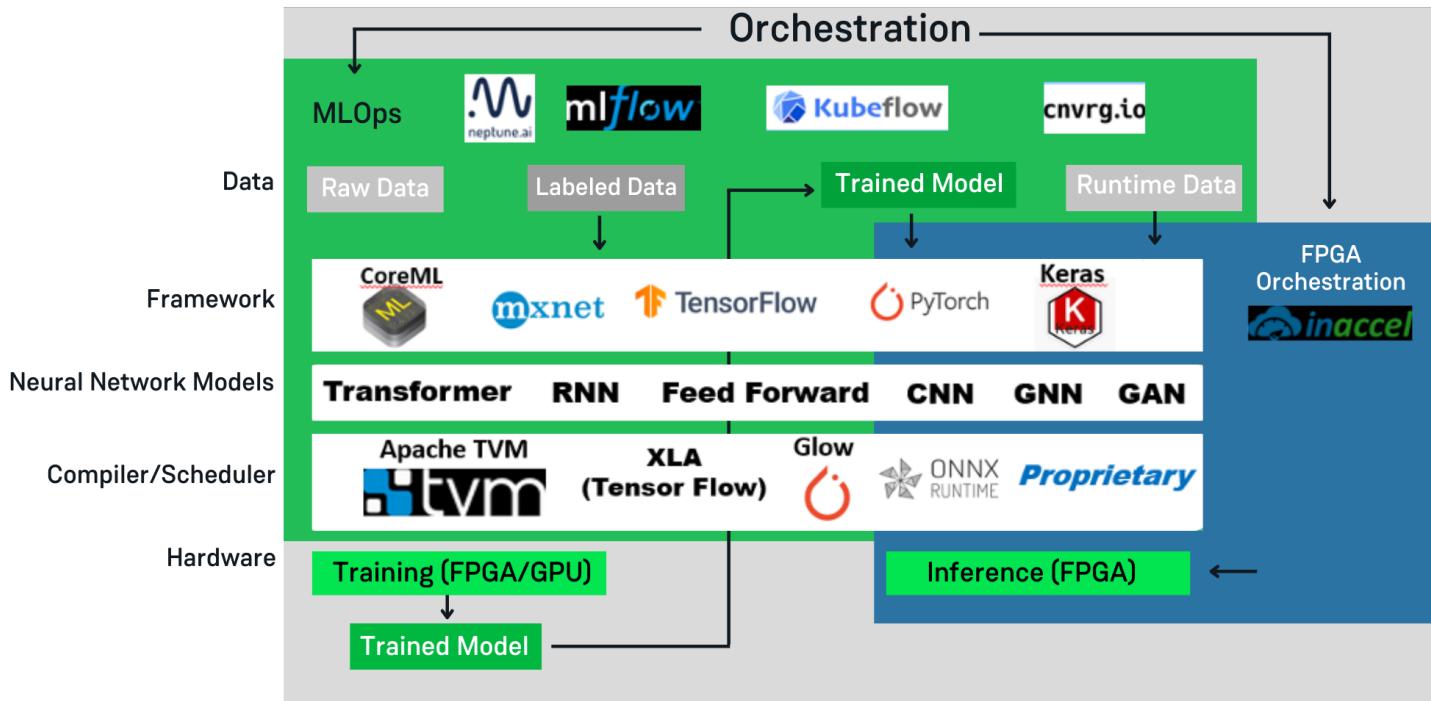


SOC ASIC or MCM based Embedded FPGA Tailored to embedded User/Enterprise Edge ML workloads

Achronix FPGA Machine Learning Solutions



- Efficient and powerful hardware optimized for machine learning workloads
 - Flexible and Embedded Models
- Complete solutions targeting
 - **CNNs** – Classification, Object Detection, Pose Estimation
 - **RNNs** – Automatic Speech Recognition (ASR)
 - **Transformer Networks** – Natural Language Processing
 - **Feed Forward** – Speech Synthesis
- Achronix working closely with ecosystem partners leading to complete solutions





Assemble all the pieces needed
for AI at the edge of IoT



THANK YOU.

Colin Alexander, Director of Product Marketing
Achronix Semiconductor

www.achronix.com