Radisys

Reducing Complexities to Integrate OpenVINO™ Toolkit Audio-Video Analytics in 5G & Edge IoT Applications

April 24, 8 a.m. PT

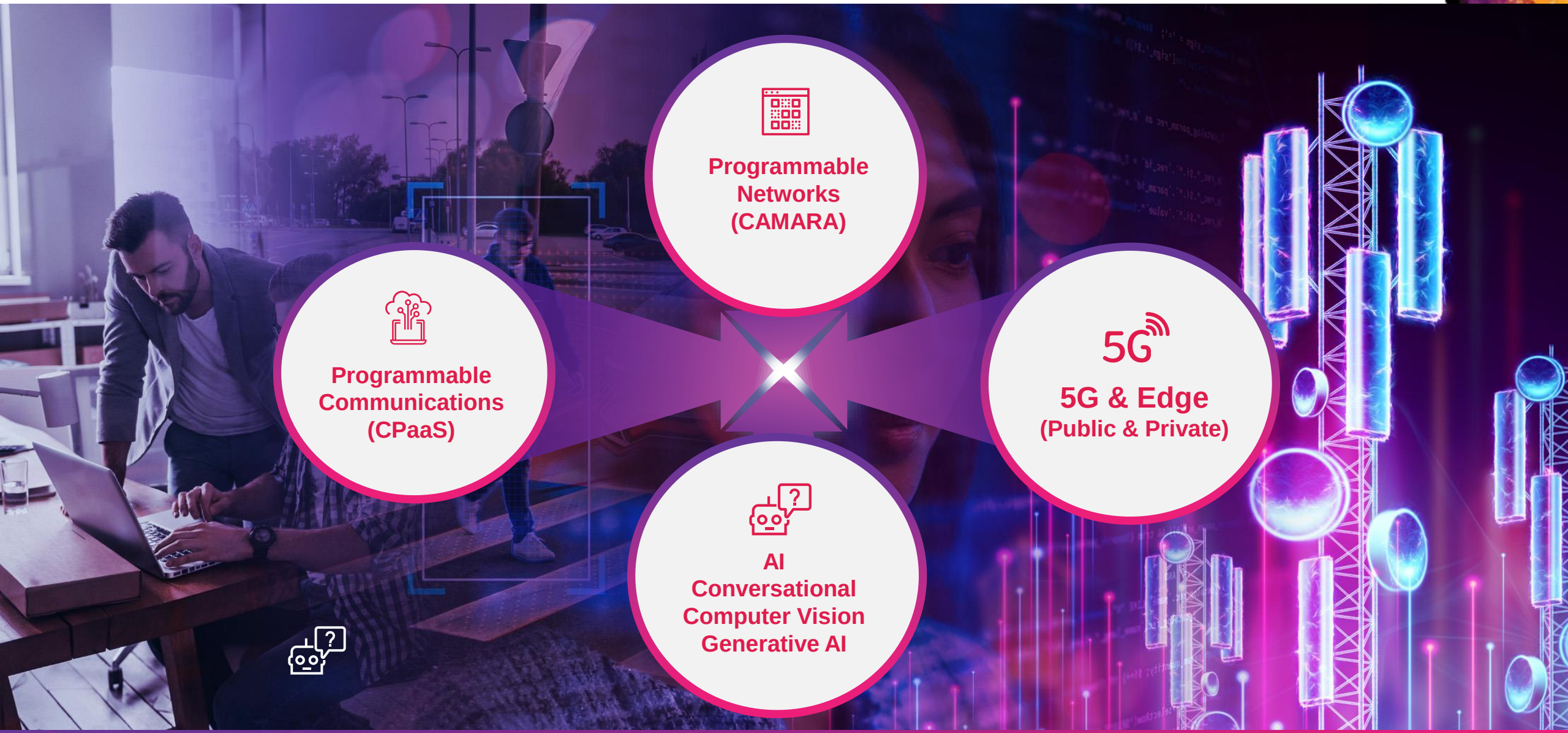
Presenters:

Adnan Saleem, CTO, Cloud and Software Solutions, Radisys

Adrian Boguszewski, AI Software Evangelist, Intel

intel partner
Titanium

The New Era: Fostered by the Convergence of Diverse & Mature Forces



The Complexity of Integrations Remains a Barrier

Convergence of Technologies

Reducing Integration Complexities Beyond Voice & Messaging

Key Technologies

- Programmable Communications
- Conversational and Computer Vision AI
- Programmable 5G Networks

Enables New Monetization

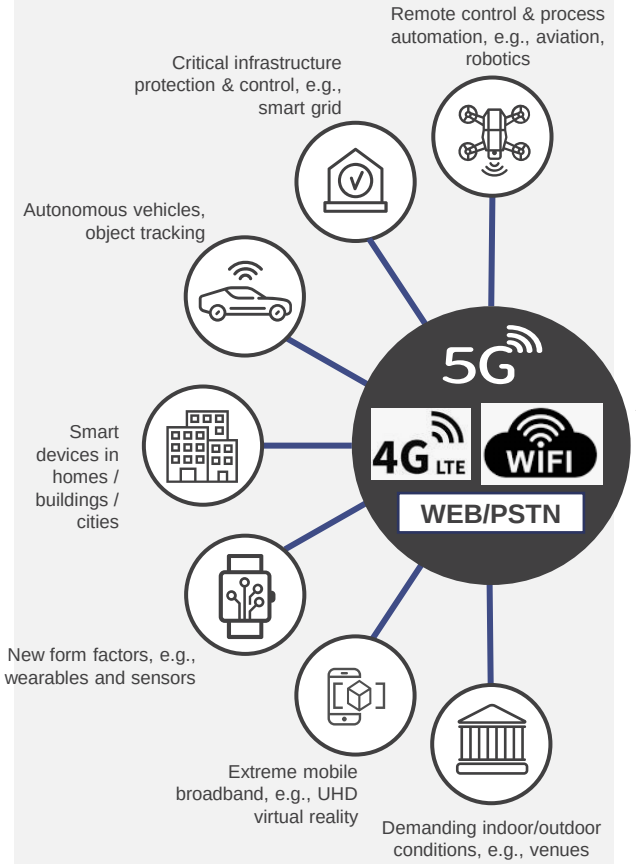
- ✓ Public and Private 4G/5G applications
- ✓ Industry 4.0 edge cloud applications
- ✓ OTT digital engagement services
- ✓ Inclusivity & Accessibility apps
- ✓ AI-driven customer service
- ✓ UC & Collaboration value-adds

Accelerates Innovation Beyond Voice & Messaging

Engage Digital

Programmable Communication & Digital Engagement Platform for Performance Intensive Apps

Digital Devices & Use Cases



Simple APIs, Plug & Play SDKs, Low/No-code UI, and Programmable Applications

Built on carrier-proven underlying technologies deployed by 200+ mobile operators globally

Enabling Digital Experiences

Communication & Collaboration

Face / Object / Video Recognition

Augmented Reality Overlays

Speech Rec & NLU Analytics

IoT Triggered Real Time Communication

Accelerates Time to Revenue for Service Providers to Offer Innovative Digital Services

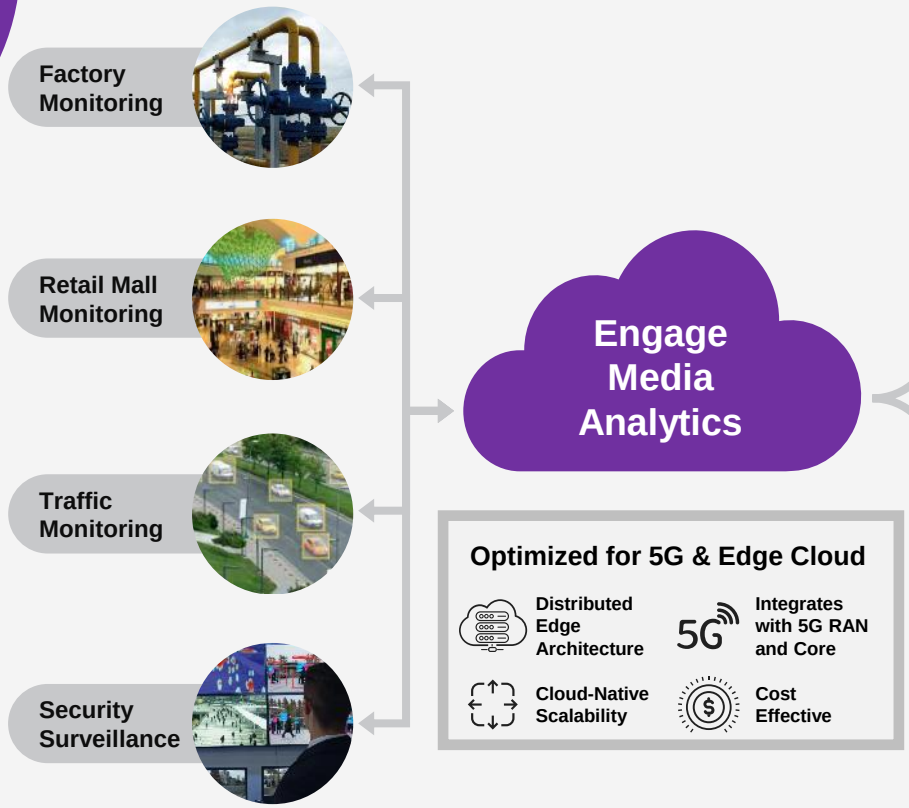
Accelerate Industry 4.0 Application Innovation

ENGAGE MEDIA ANALYTICS

Monetize 5G & Edge Cloud Investments

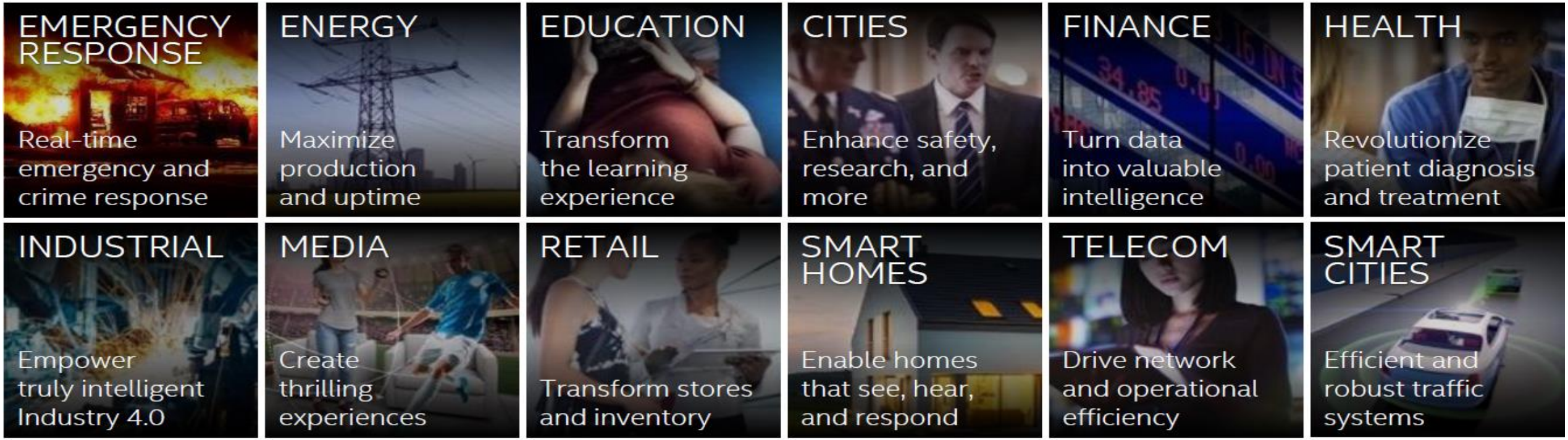
- Integrated Programmable Audio-Video Analytics & Communications
- Intuitive Low-Code/No-Code Visual Design Tool
- Open & Extensible Platform
- Data Privacy & Control

Programmable Computer Vision Analytics



Removes Complexities for Rapid Creation and Scaling of Industry 4.0 Applications

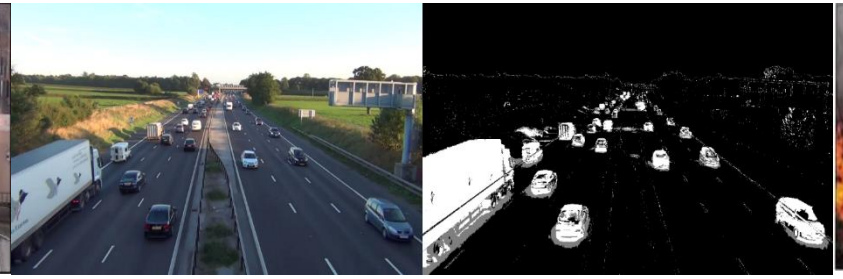
Engage Digital Programmable Video Analytics Use Cases



source: Intel

Facial recognition

People Detection & Counting



Vehicle Detection & Counting

Heat Map Fire Detection

Source: Intel

Monetizing 5G & Edge Beyond Connectivity

Communications + Computer Vision

Vertical Specific Digital Automation Applications

Expedite Airport
or stadium entry based on Face Match

Monitoring Drive
through lines for wait times and cars leaving for revenue loss analysis

Monitoring emotions
of people in retail shop, classroom, stadiums, public meetings

Demographic analysis
of people (age, gender, sex) in crowded places

Notifying concierge
when a guest in a wheelchair needs assistance

Notifying assistant
when an elderly at-home patient has fallen from the bed

Augment emergency
call live agent discussion by identifying critical keywords or background sounds

Detecting a fire
in a factory, campus building and alerting emergency services personnel

Detecting a gun shot sound
in a building and immediate alert law enforcement services

Monitoring warehouse
shelves for inventory management

Identify Shopper
traffic volume in retail aisles and send special deals/ coupons

Programmable Low-Code/ No-Code Video Analytics + CPaaS = Many Killer 5G Apps

Endless Possibilities: Programmable Computer Vision Analytics + Communications

100+ Pre-Trained Models in EDP

- Age & Gender
- Facial landmarks
- Face detection – standard & enhanced
- Face re-identification
- Head position
- Human detection – eye level & high-angle detection
- Human pose estimation
- Person detection & action recognition
- Person attributes recognition crossroad
- Person re-identification – ultra small/ ultra fast
- Identify someone from different videos – standard & enhanced
- Emotion recognition
- Identify roadside objects
- Vehicle detection
- Vehicle metadata
- License plate detection: small & front facing
- Pedestrian detection
- Pedestrian & Vehicle detection
- Detect People, Vehicles & Bikes
- Advanced roadside identification
- Landmarks regression
- Single image super resolution
- Text detection
- Retail environment
- Smart classroom use cases

Pre-Trained Models + Models Trained w/ Custom Data

Expedites solution development for monetizing video content

Expedite Airport or Stadium entry based on Face Match

Automated attendance system in offices, schools, colleges, shops, etc.

Monitoring Drive through lines for wait times and cars leaving for revenue loss analysis

Monitoring emotions of people in retail shop, classroom, stadiums, public meetings

Demographic analysis of people (age, gender, sex) in crowded places

Notifying concierge when a guest in a wheelchair needs assistance

Notifying assistant when an elderly at-home patient has fallen from the bed

Augment emergency call live agent discussion by identifying critical keywords or background sounds

Detecting a fire in a factory, campus building and alerting emergency services personnel

Detecting a gun shot in a building and immediate alert law enforcement services

Monitoring warehouse shelves for inventory management

Identify Shopper Traffic Volume in retail aisles and send special deals/ coupons

Radisys

Endless Possibilities: Programmable Computer Vision Analytics + Communications

- Face Recognition
- Object Detection & Counting
- Image Recognition
- Motion Detection
- Intrusion Detection
- Loitering Detection
- Emotion Recognition
- Age & Gender Recognition
- Audio detection & Classification
- Virtual Reality
- Augmented Reality
- And more …

Edge Computer Vision Inferencing + Communications

5G and AI Powered Worker Safety

engagedigital.ai

Programmable Computer Vision Analytic

- A worker safety solution uses Engage video analytics to perform anomaly detection on streaming video
- The Engage video analytics software uses QoD to manage bandwidth across cameras/devices; i.e. switching from standard resolution feeds to high-resolution, interaction/real-time content when anomalies are detected
- When a safety or security issue is confirmed, real-time notifications are sent to the worker safety platform and affected workers.

ALERT — *Message/ Call* — Affected Workers

Locate Camera — *Location API*

Boost Camera Resolution — *QoD API*

Programmable 5G Network — Operator 5G Network

STREAM — Worker Safety Platform — Enterprise Remote Monitoring Center

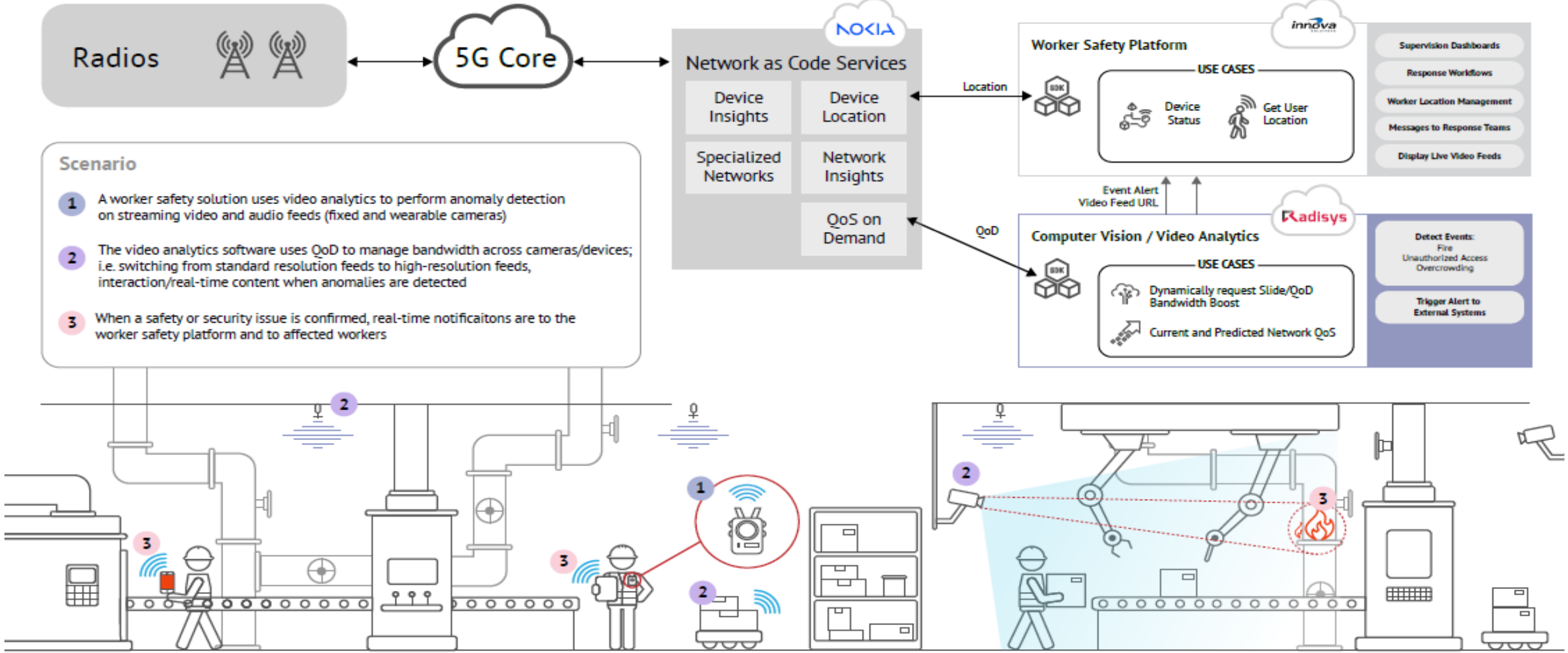
Send the location of the camera and stream high-resolution video

Eliminate Complexities to Implement Industry 4.0 Applications with 5G Network APIs

Live Programmable Network Demo: Incidence Response

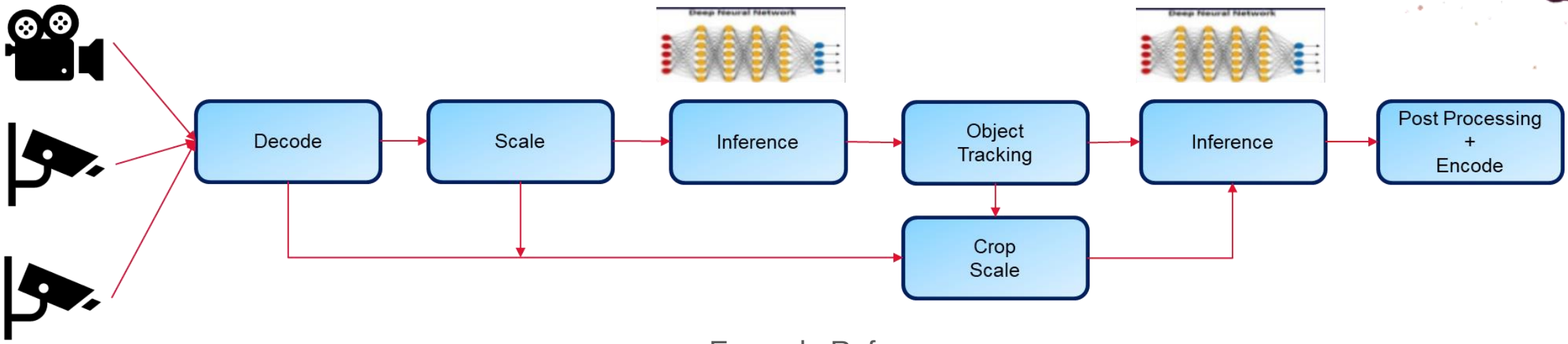
Orange

Open Tech 2024



Demo Presentation Video:
https://www.youtube.com/watch?v=xFrw7vPfcG0

Computer Vision Media Pipeline - Workflow



Example Reference:

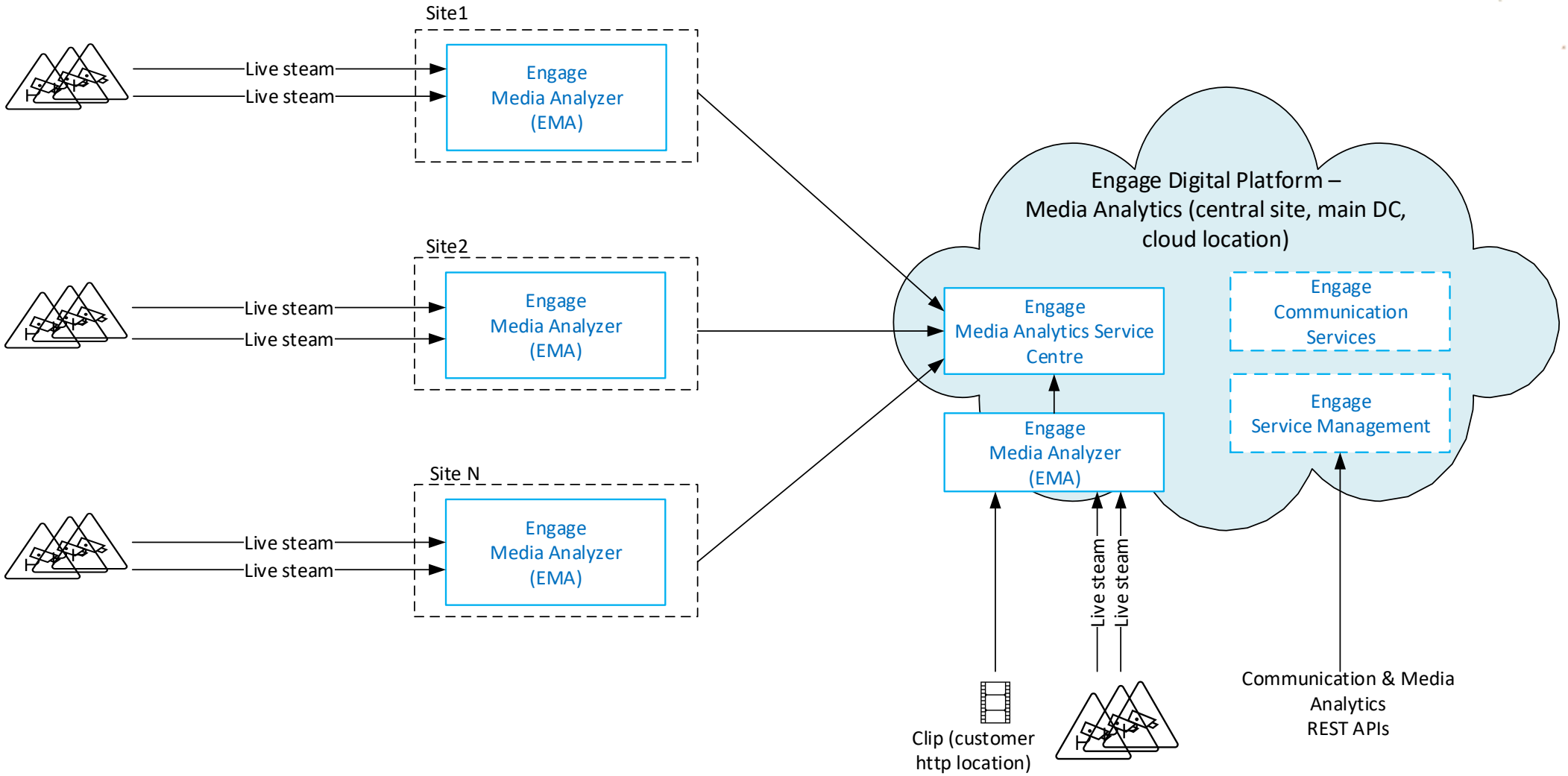
Ingress Video & Audio – H.264 Video Codec, 720p, 30fps

AI/ML Inference @ 15 fps (Age, Gender, and Emotion Identification) Keyword Detection like “Police”, “Fire”, “Help” in English or Local Language

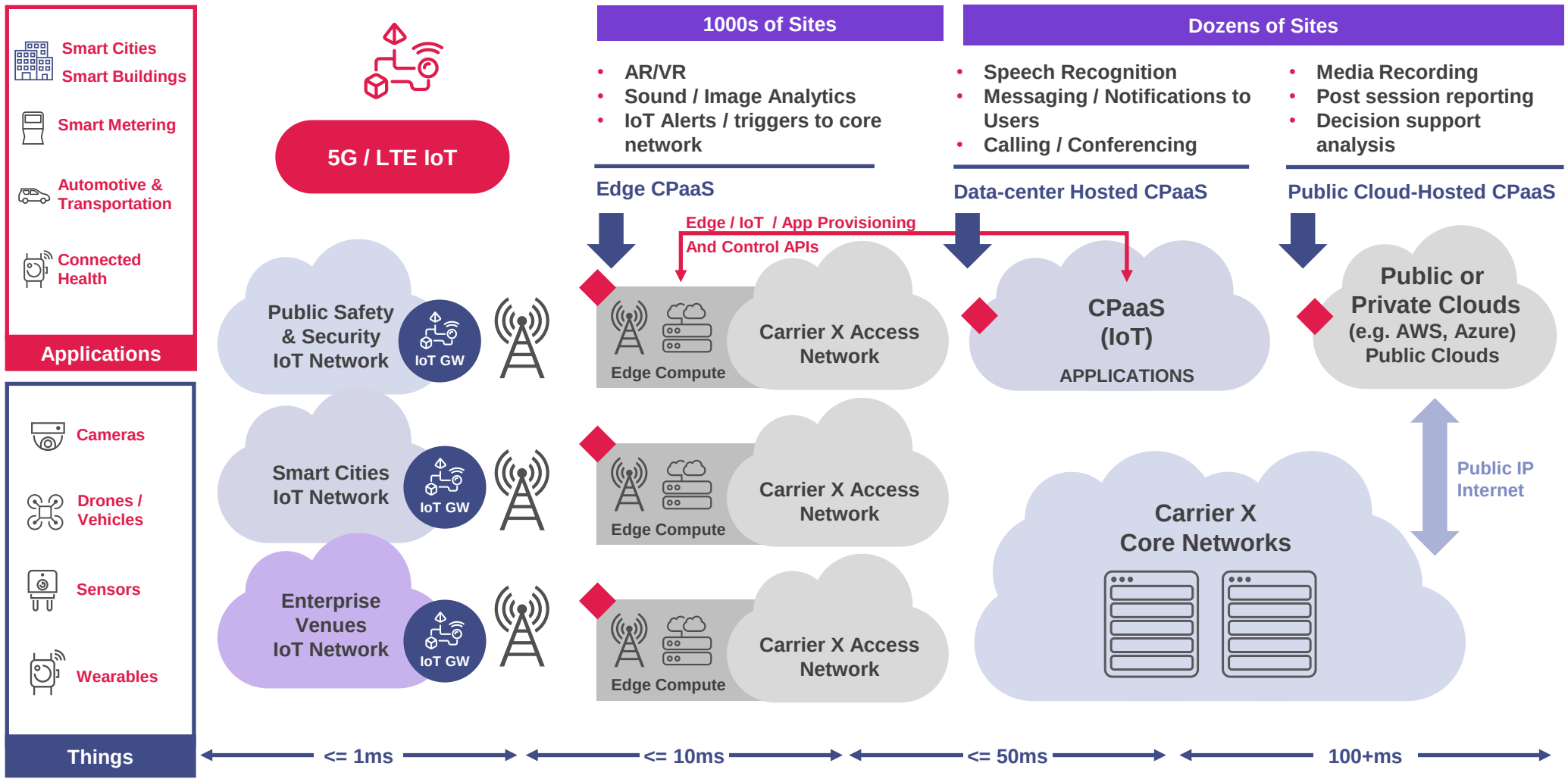
AR/Text Image Overlay with Inference determined identification Video Encoder @ 720p 30 fps

Egress Video & Audio – H.264 Video RTP stream, 720p, 15 fps

Distributed Edge Media Analytics Architecture



Edge, Near Edge Data Center, Public/Private Cloud Optimizing Computer Vision and Network Resources



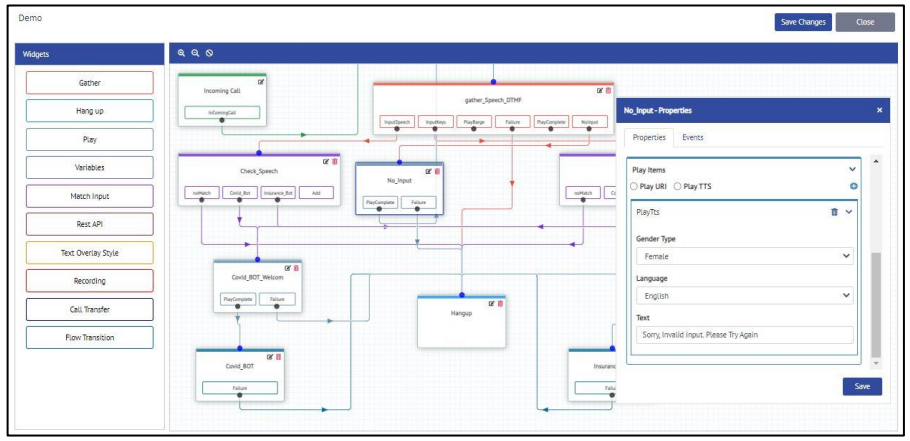
Benefits

- Execute application in optimal location(s) based on proximity, context, QoS / latency requirements, etc.
- Future proof your application
- Backhaul savings
- Reduce public cloud processing and costs
- Enterprise level SLAs
- Regulatory compliance
- Greater control over data and security

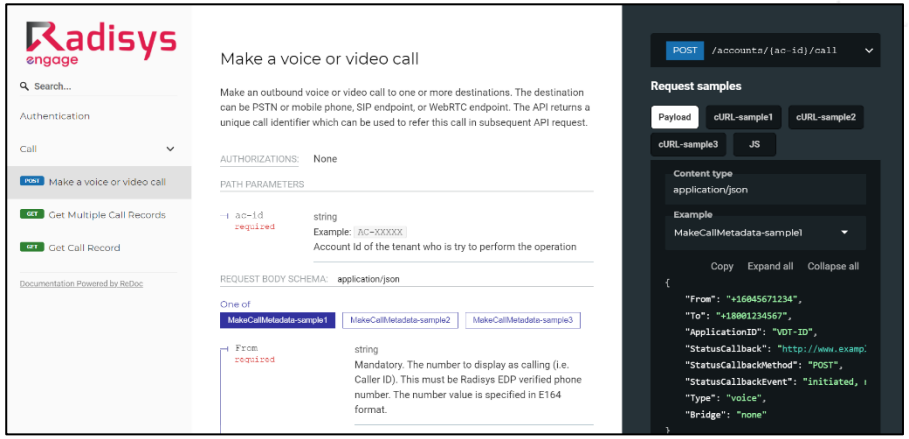
Representative Distributed CPaaS Applications / Capabilities

Reducing Complexities via APIs, SDKs, VDT

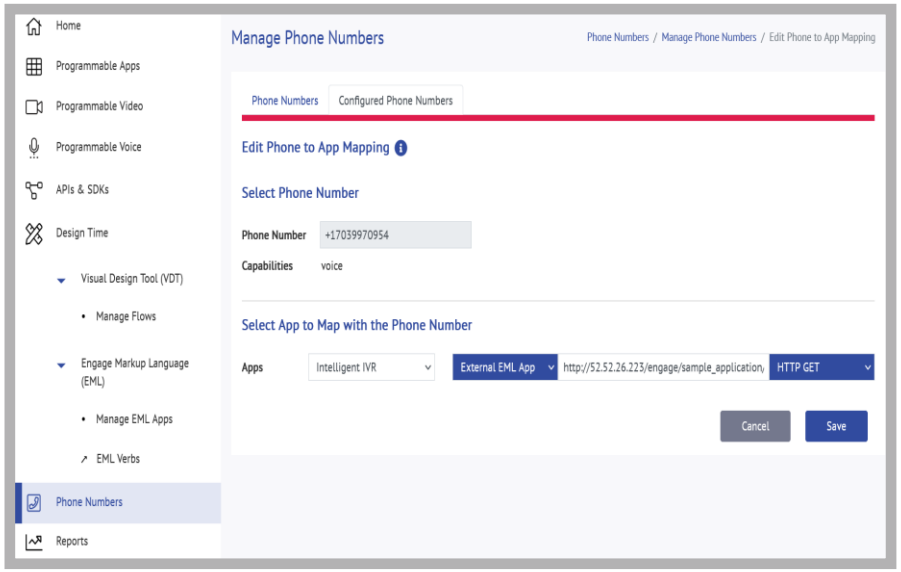
No Code: Visual Design Tool



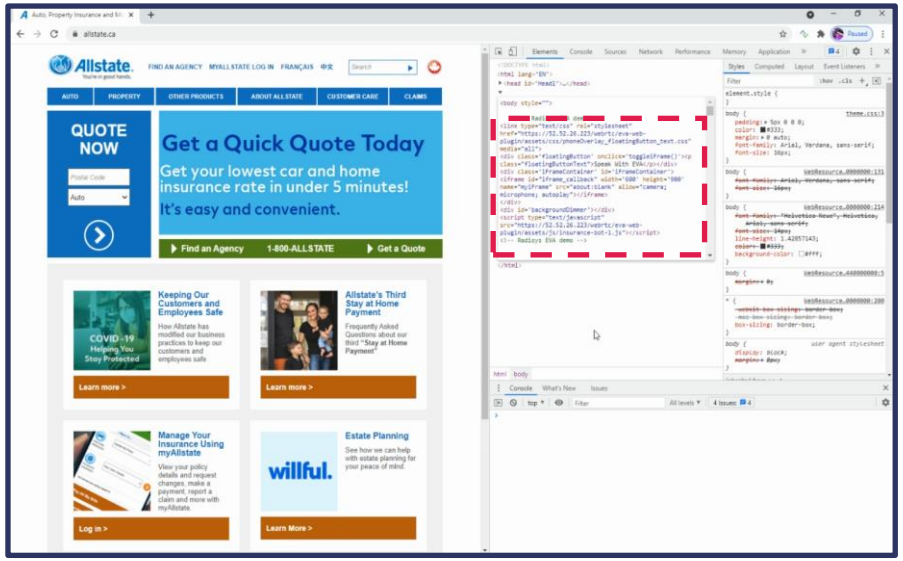
Develop with Rest API



Low Code: Engage Markup Language (EML)



Develop with WebRTC SDK



Enabling Service Providers to Accelerate and Differentiate

Monetize digital services beyond voice & messaging

In-Network, Edge & Cloud to ensure data privacy and quality of experience

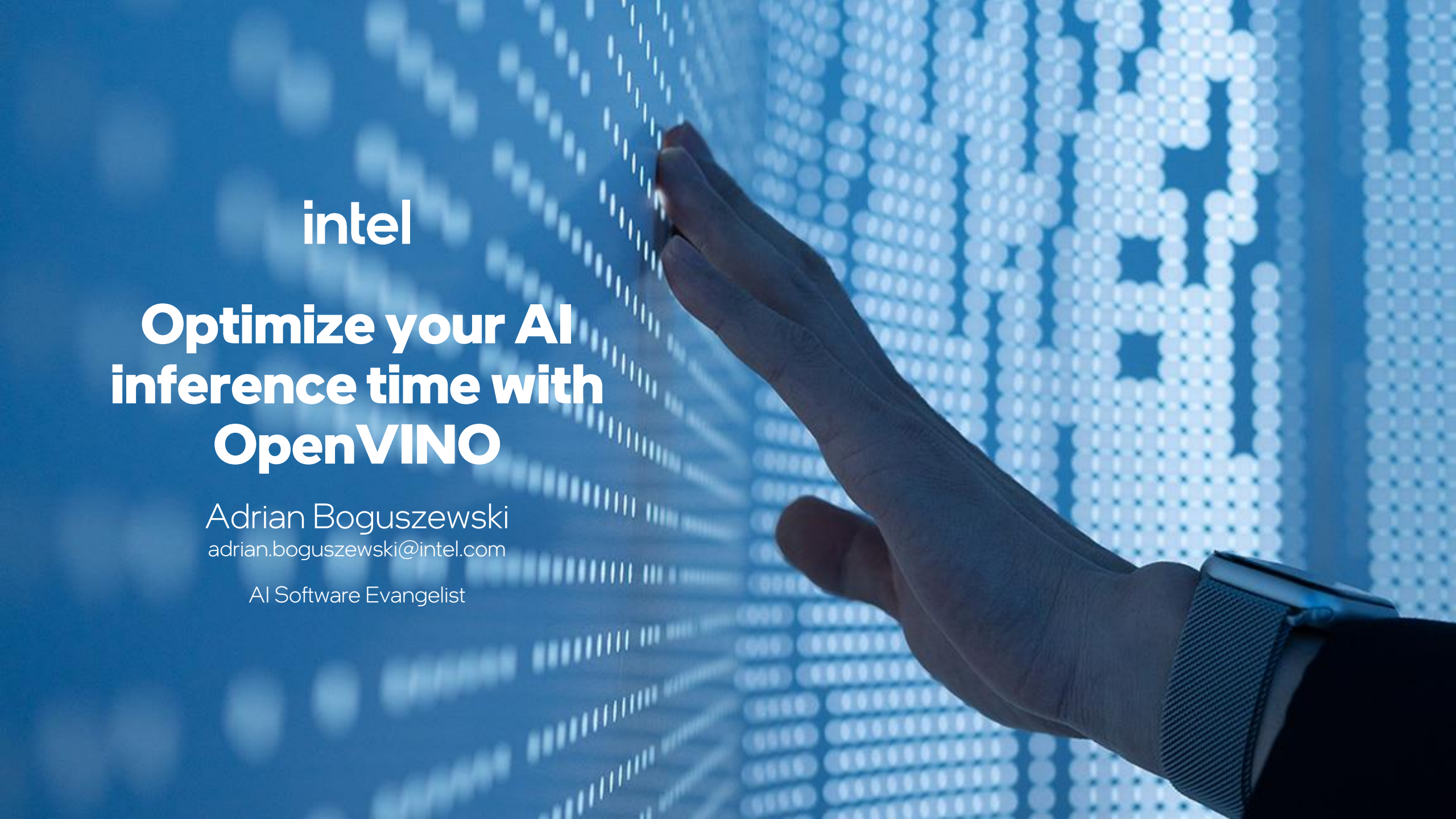
Flexibility to bring your own components

Deliver on SLAs not available through public cloud-based offers

Enjoy significantly higher margins

Promote and build your brand

Accelerate Time to Market & Revenue



Optimize your AI inference time with OpenVINO

Adrian Boguszewski
adrian.boguszewski@intel.com

AI Software Evangelist

The Challenge

How to get the best
AI application performance
out of the available hardware?

How to get the most
popular use cases and
cutting-edge AI
running?

How to maintain code
across different
architectures?

The Solution

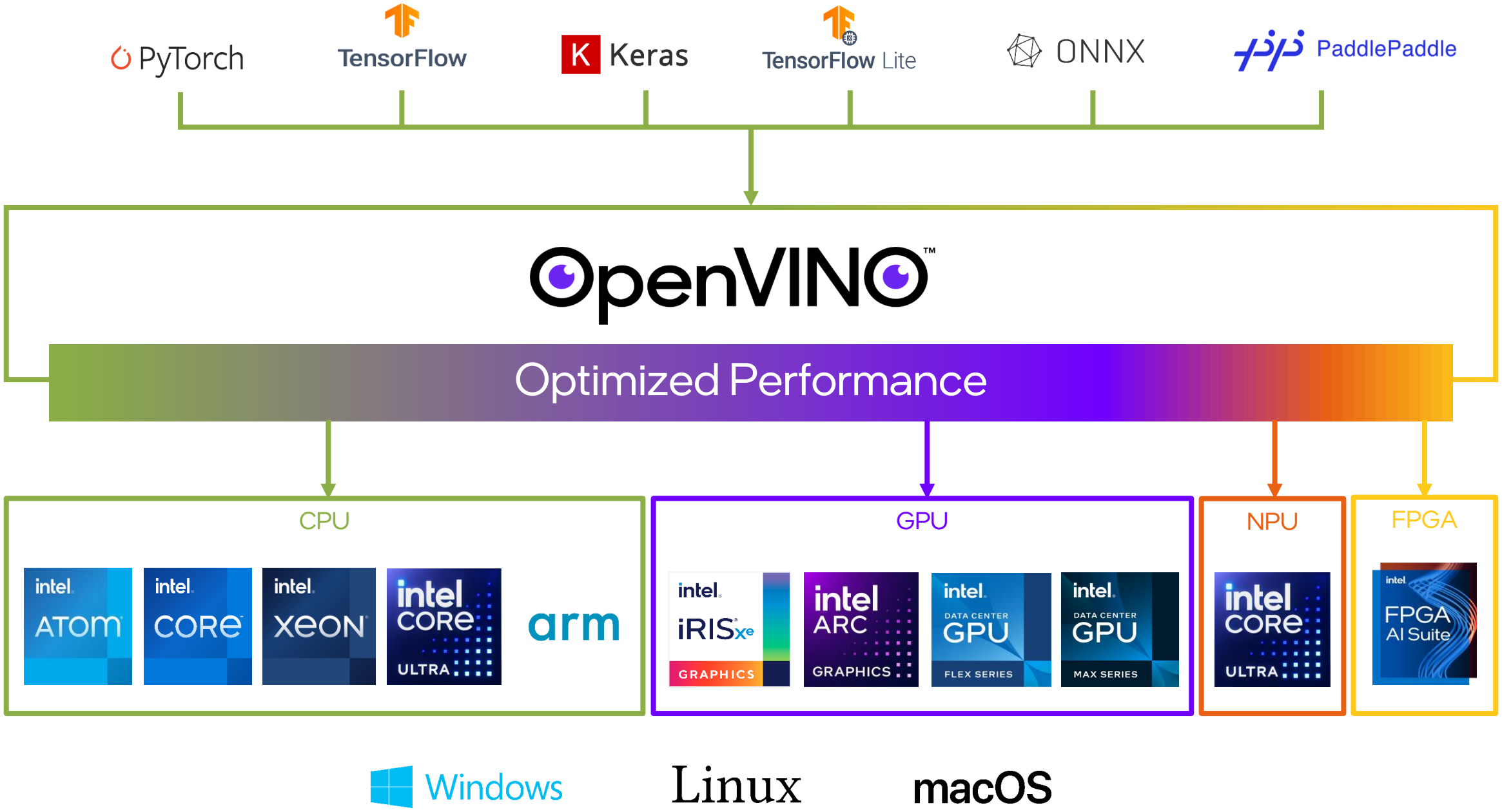
Google Summer of Code

open source initiative® Apache 2.0

OpenVINO™

- Visual+ NLP, Audio, LLM, GenAI, Transformers...
- Inference
- and Neural Network
- Optimization

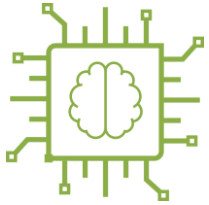
Powered by oneAPI


PyTorch
TensorFlow
Keras
TensorFlow Lite
ONNX
PaddlePaddle
OpenVINO™
Optimized Performance
CPU
intel ATOM
intel CORE
intel XEON
intel CORE ULTRA
arm
GPU
intel iRIS Xe GRAPHICS
intel ARC GRAPHICS
intel DATA CENTER GPU FLEX SERIES
intel DATA CENTER GPU MAX SERIES
NPU
intel CORE ULTRA
FPGA
intel FPGA AI Suite
Windows
Linux
macOS

What's New in the 2025.0 Release?

1 Model

- Support for new models: Qwen 2.5, Deepseek-R1-Distill-Llama-8B, DeepSeek-R1-Distill-Qwen-7B, DeepSeek-R1-Distill-Qwen-1.5B, FLUX.1 Schnell
- Enhanced Whisper performance with OpenVINO GenAI
- Preview: Intel® NPU support via torch.compile



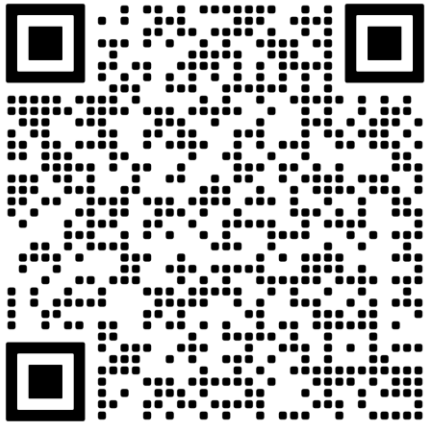
2 Optimize

- Preview: Prompt Lookup in OpenVINO GenAI optimizes 2nd token latency by using predefined prompts
- Preivew: Image-to-image inpainting added to OpenVINO GenAI
- Asymmetric KV Cache compression enabled by default for INT8 on CPUs

3 Deploy

- Support for Intel® Core™ Ultra 200H processors (Arrow Lake-H)
- Triton Inference Server integration for optimized AI model serving on Intel CPUs
- Preview: OpenVINO backend for Keras 3.8
- OpenVINO Model Server now supports native Windows Server deployments

Installation



ubuntu

Red Hat

docker

ANACONDA

CONDA-FORGE

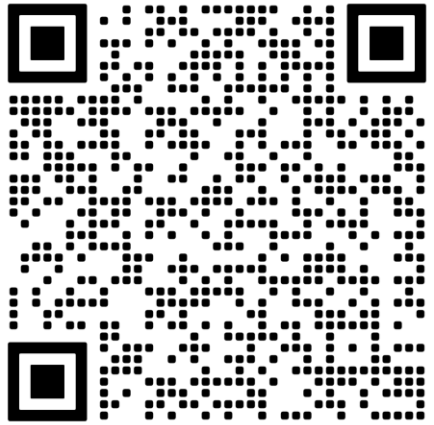
CONAN
C/C++ Package Manager

npm

www.openvino.ai

Installation

```
pip install openvino
```



ubuntu

Red Hat

docker

ANACONDA

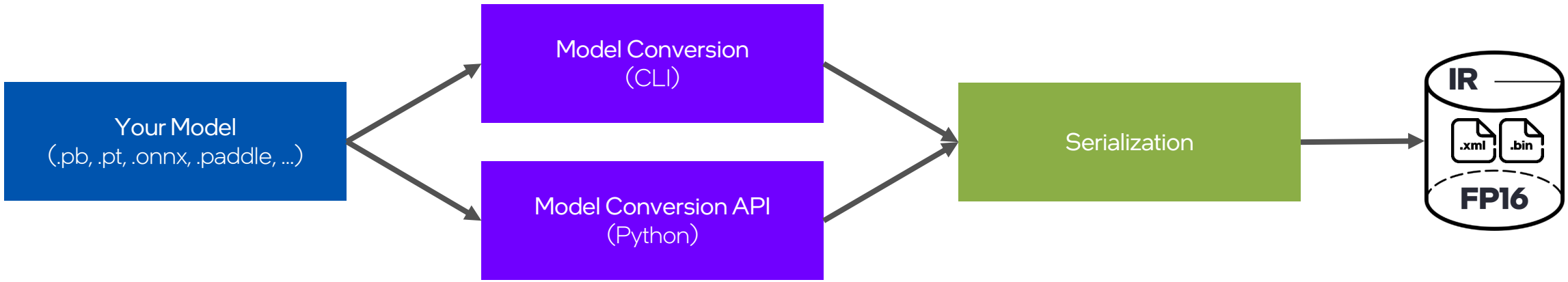
CONDA-FORGE

CONAN
C/C++ Package Manager

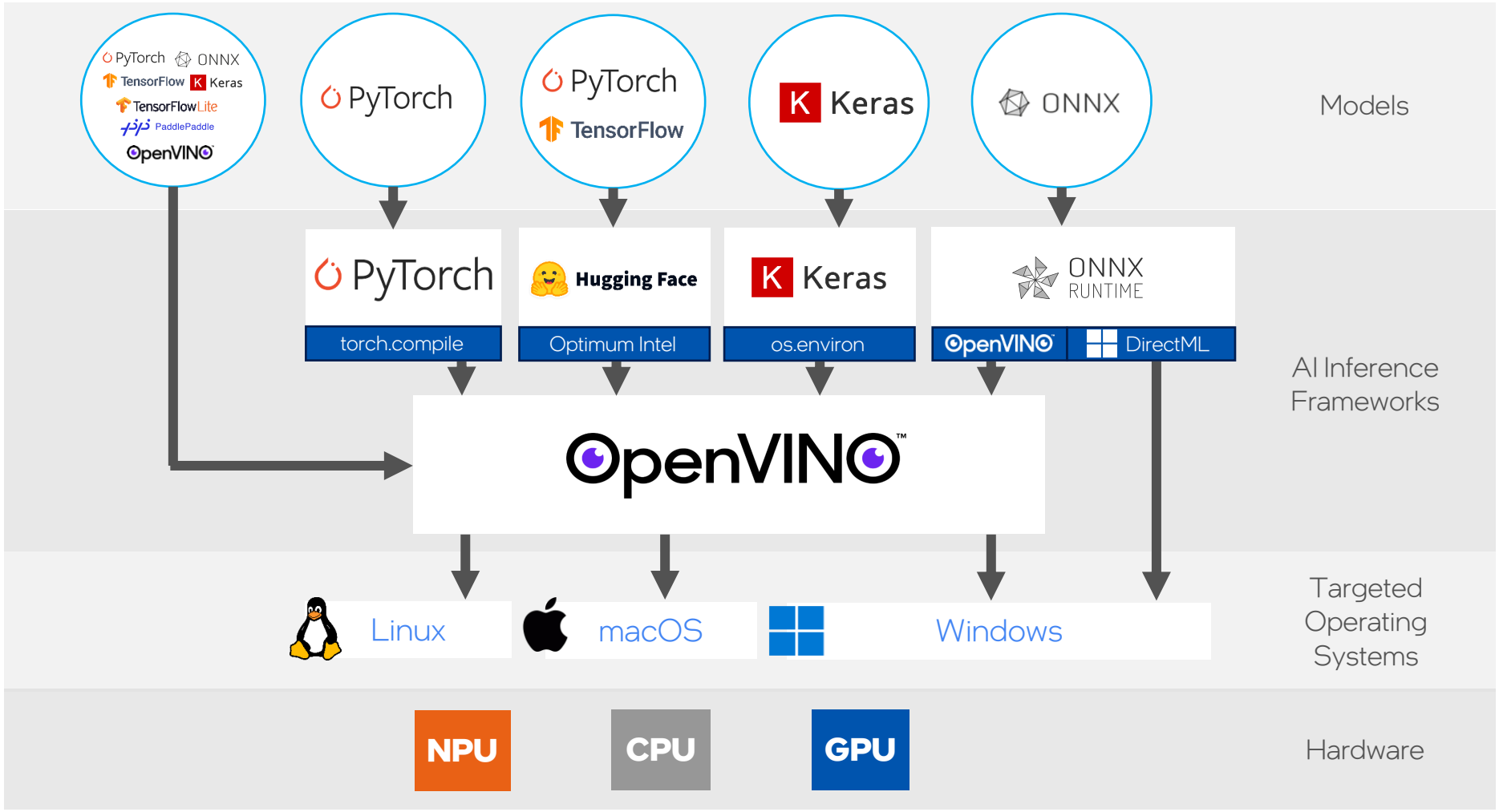
npm

www.openvino.ai

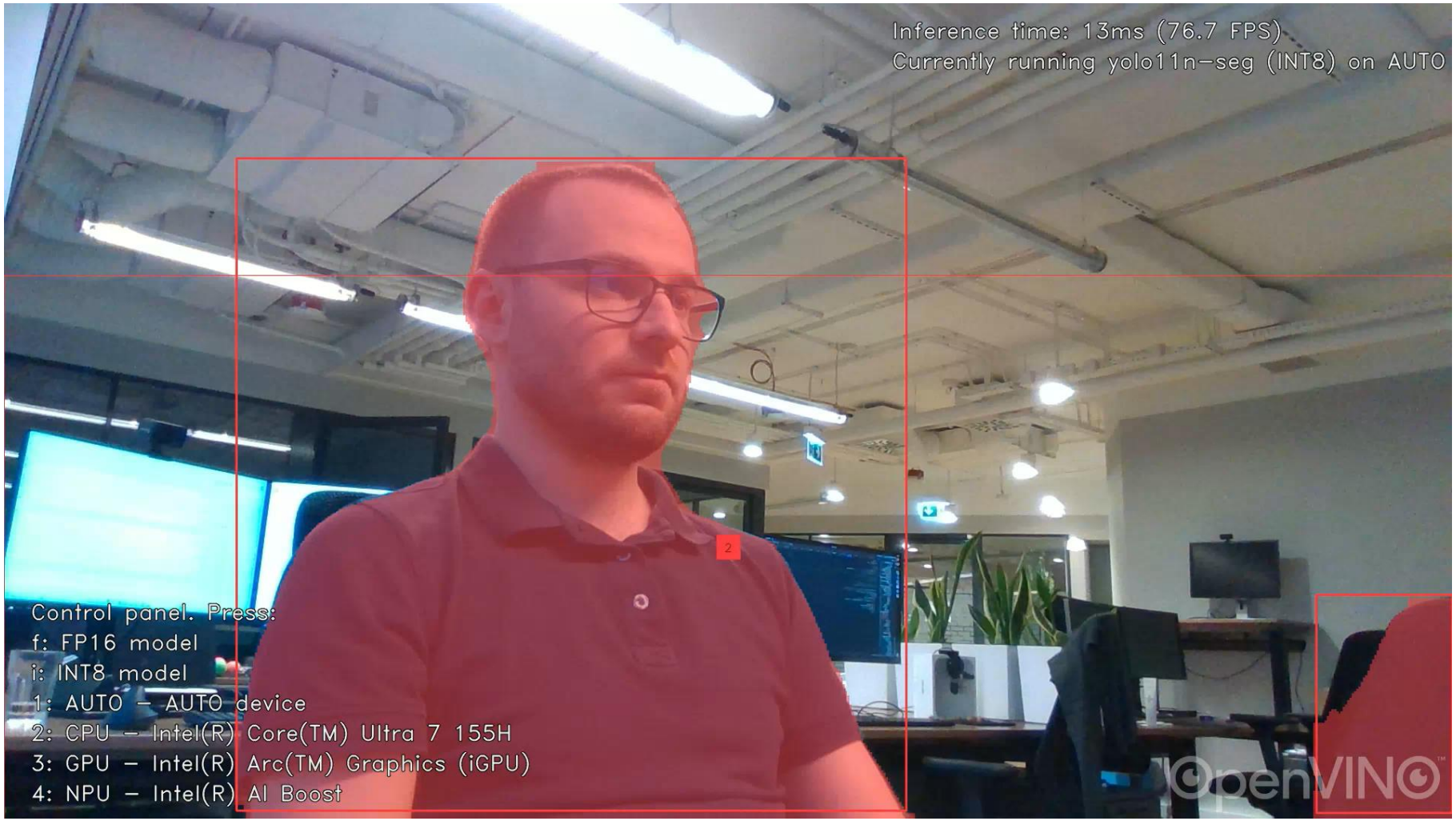
OpenVINO™ Model Converter



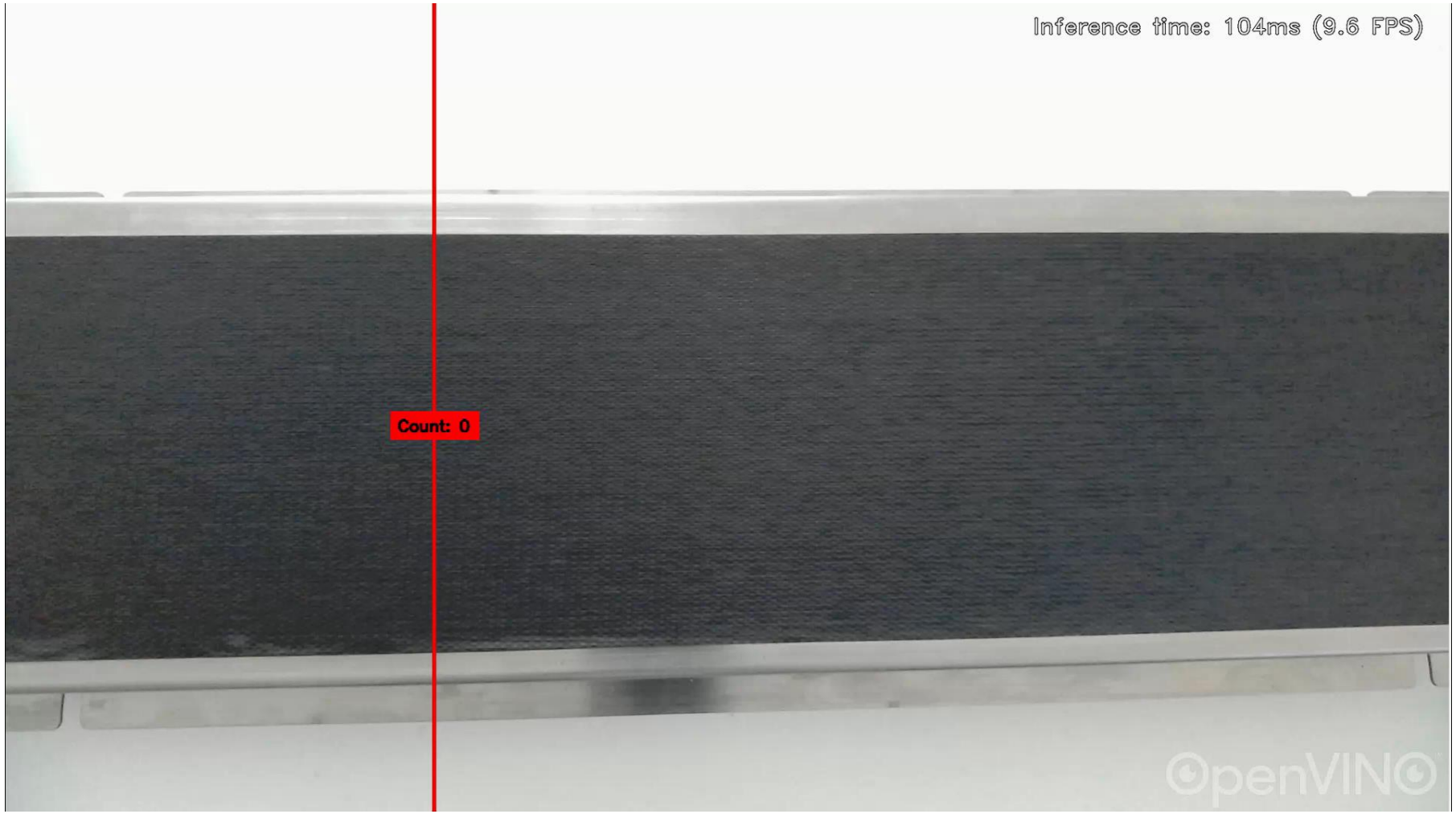
OpenVINO™ as Backend



People Detection/Segmentation/Counting

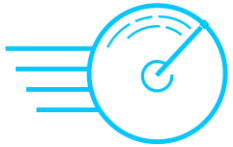


Object Tracking/Counting



Benefits of Building Applications with OpenVINO™

Build and deploy AI applications in simple steps

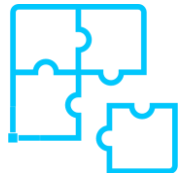


Faster inference speed

Maximize AI performance across CPU, GPU, NPU

Smaller model and binary size

Reduce memory footprint



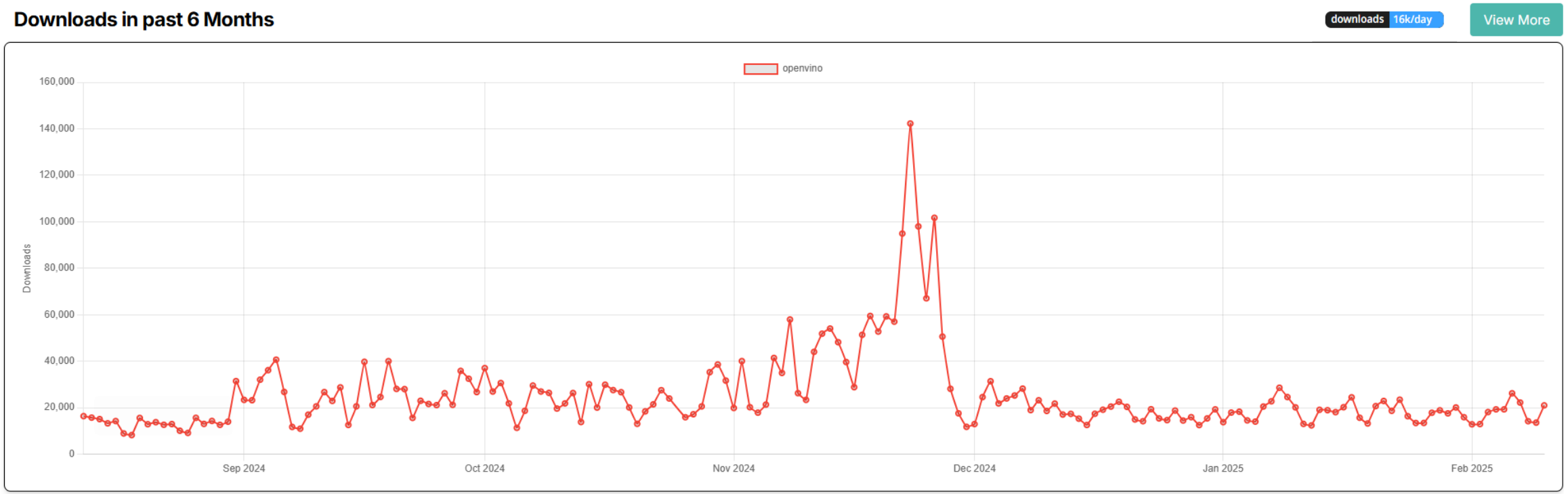
Ability to scale to many nodes with serving

OpenVINO™ Ecosystem Adoption

OpenVINO

intel ai

OpenVINO™ Ecosystem Adoption



Over 12,500,000 downloads!

Source: https://piptrends.com/package/openvino

Adrian Boguszewski
adrian.boguszewski@intel.com

Thank You

Visit: https://www.engagedigital.ai/

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No product or component can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com.

Intel, the Intel logo, are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries.
*Other names and brands may be claimed as the property of others.

© Intel Corporation