恩智浦处理器用于AGV应用

弋 方
EMAIL: FANG.YI@NXP.COM
恩智浦半导体 微处理器高级市场经理

2021年10月28日

SECURE CONNECTIONS
FOR A SMARTER WORLD

恩智浦完整的软硬件产品和生态

EDGEVerse

Industrial Edge	IoT Edge	Network Edge

垂直应用领域

CLOUD PLATFORM EXAMPLES

aws · Microsoft Azure · Alibaba Cloud · Google Cloud · Baidu 百度 · salesforce · GE PREDIX · IBM Cloud

云服务

NXP eIQ™ Machine Learning	Audio Solutions	Key injection EdgeLock-2-Go	Turnkey solutions & SW services	Partners
Edge Intelligence		Security & Provisioning		Device Management

软件赋能与服务

SECURITY | CONNECTIVITY | GRAPHICS | AUDIO / VIDEO | NETWORKING

中间件与参考代码

ubuntu · yocto PROJECT · CentOS · Open Industrial Linux

MCUXpresso

Android	Common Linux Platform	Zephyr	FreeRTOS	Azure RTOS

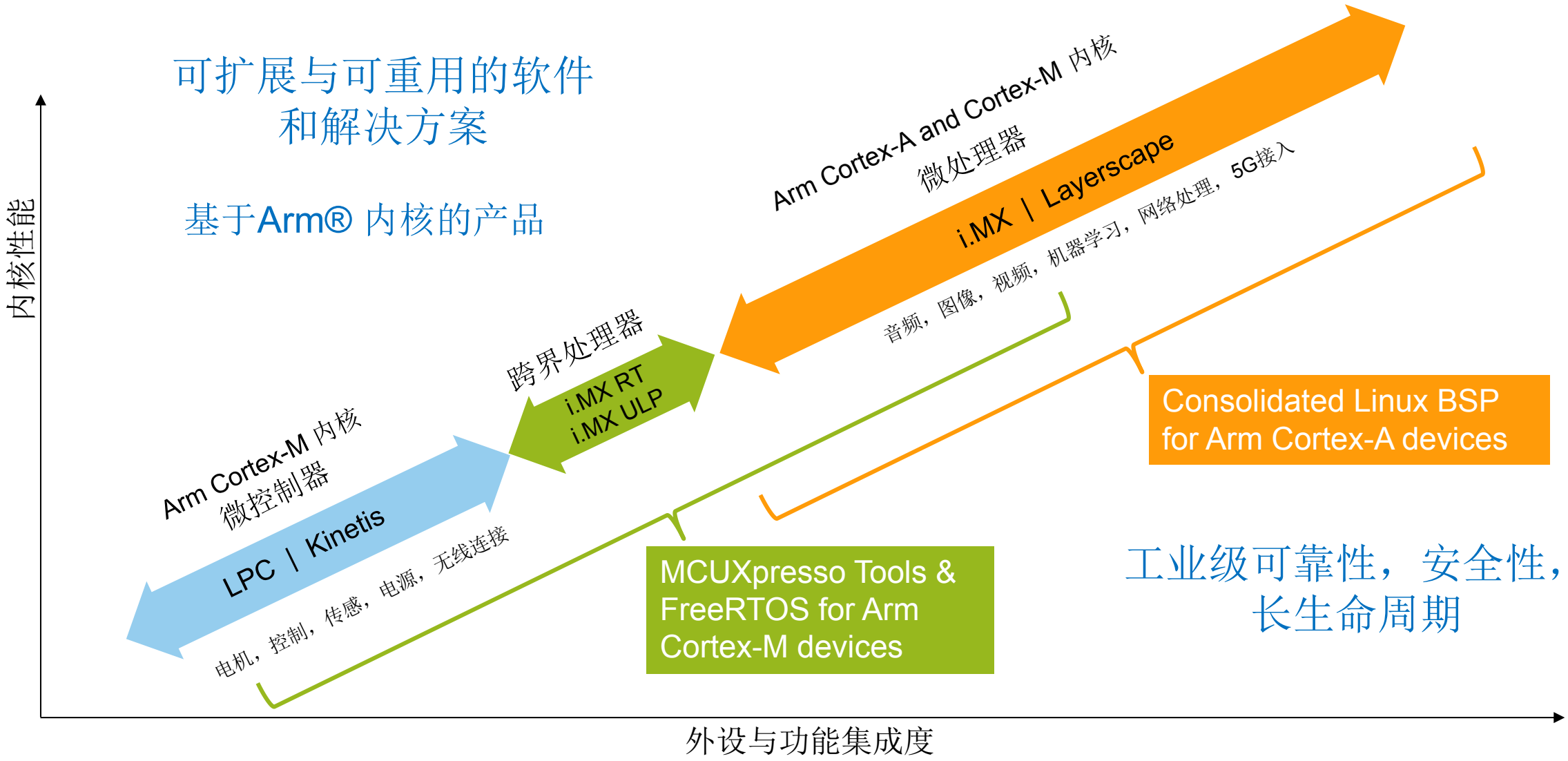
操作系统，**BSP**

i.MX	MPUs	Crossover	MCUs

芯片

NXP

恩智浦提供广泛的微控制器和微处理器产品用于边缘计算

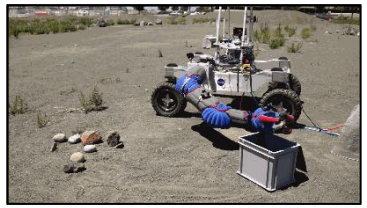


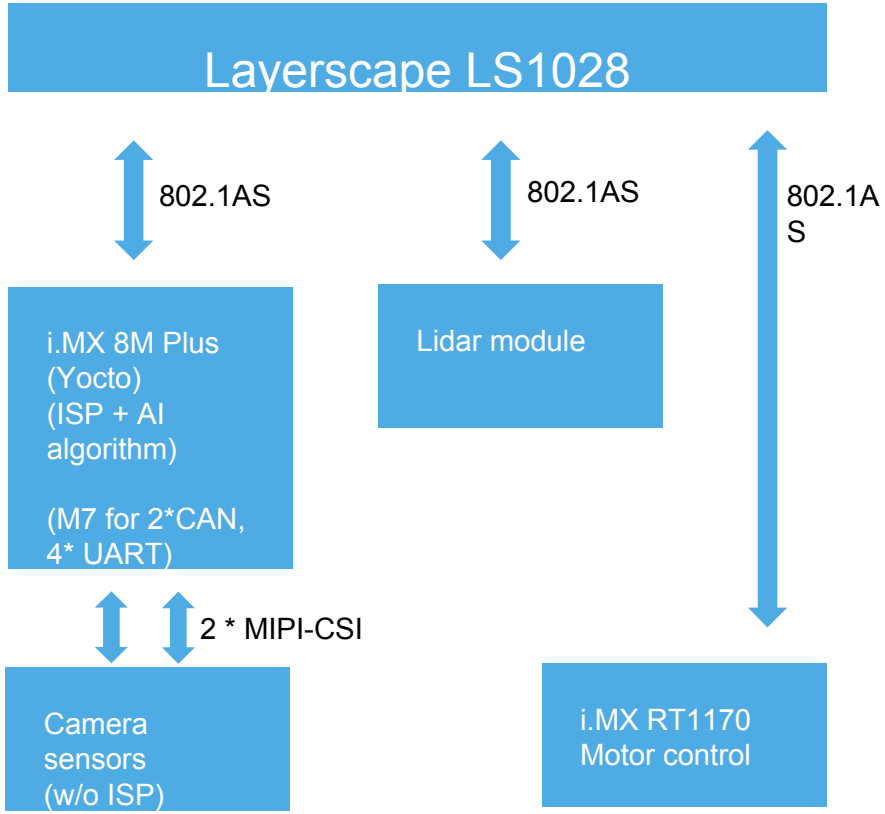
自动引导车（AUTOMATIC GUIDED VEHICLE，AGV） 分类

- 入门级
  - 基础避障及路径引导算法，采用标签等方式循迹
  - 或基于单目的简单 vSLAM 算法
  - NXP i.MX RT1170

- 中阶
  - 采用双目视觉避障以及路径引导算法
  - 具有完善的 vSLAM 算法能力，具有视觉AI能力
  - NXP i.MX 8M Plus

- 高阶
  - 复杂系统项目，板载高速网关，用于多模块数据交互
  - 工业级相机，激光雷达，毫米波雷达等多种传感器通过TSN技术接入网关
  - 驱动模块通过TSN技术接入网关
  - 强大的片上数据处理能力
  - NXP LS1028/LS1046 + i.MX 8M Plus + i.MX RT1170





i.MX RT 系列产品用于入门级AGV应用的技术特点

High Performance Real-time Processing

- Cortex-M7 up to 1GHz (RT11xx) (50% faster than current existing M7 products)
- Cortex-M33 up to 300MHz with Trust Zone (RT6xx/5xx)
- Up to 600MHz DSP and other co-processor accelerators (RT6xx/5xx)
- Up to 5MB On Chip SRAM

Low BOM Cost

- Competitive Pricing – starting @ $0.99 100K RSL
- Fully integrated PMIC with DC-DC
- Low cost packages enabling 4 layer (BGA) & 2 layer (LQFP) PCB design
- Quad/Octal Serial and SDRAM interface

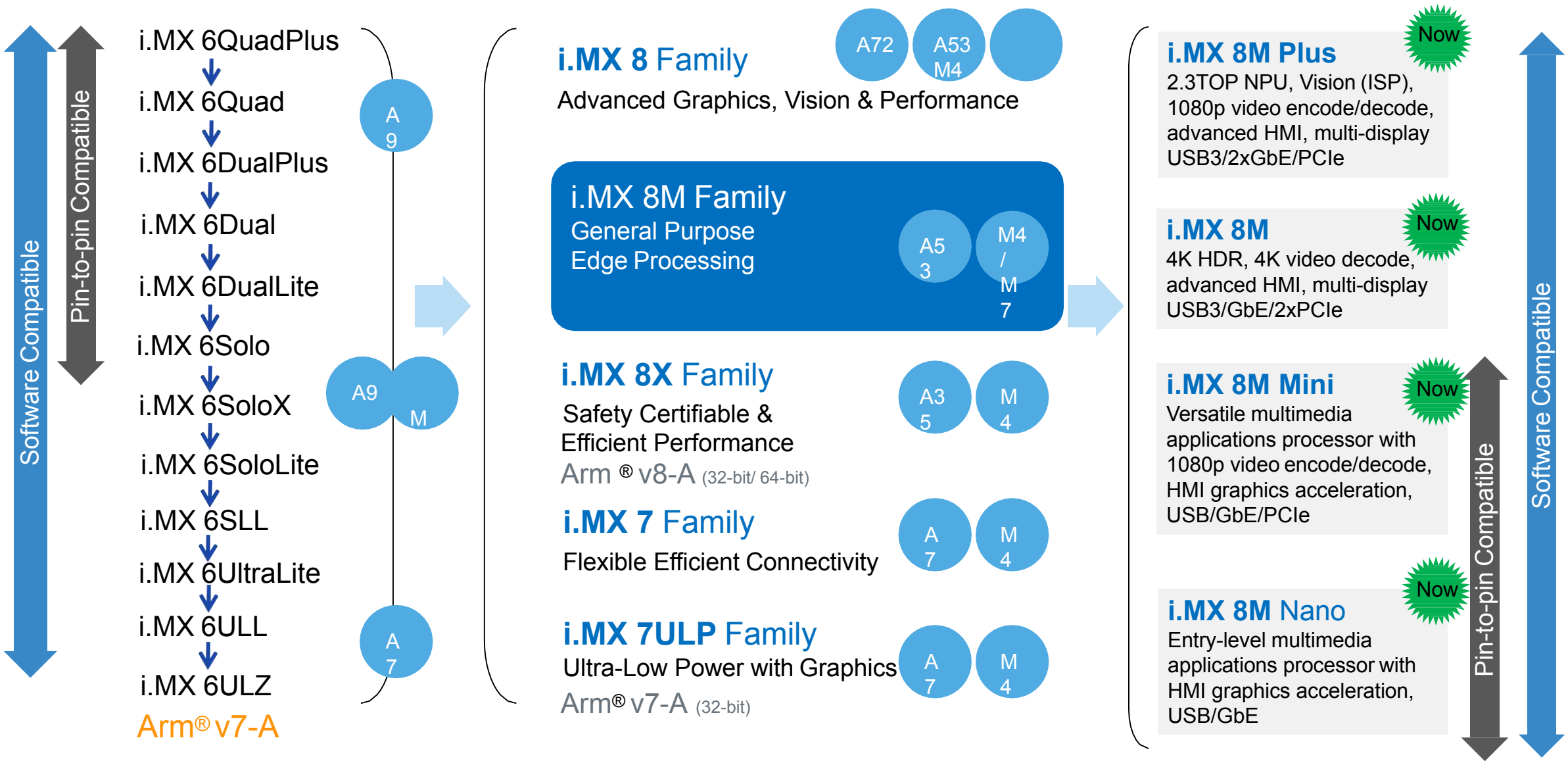
High Level of Integration

- High Security enabled by AES-256, HAB and On-the-fly QSPI Flash Decryption
- GPU2D, 2D graphics acceleration engine with Parallel & MIPI DSI/CSI
- LCD display controller up to WXGA (1366x768)
- Digital microphone and I2S interfaces for multichannel, high performance audio and voice
- Up to 3x Ethernet w/ AVB, TSN, and 3x CANFD

Easy to Use

- MCU customers can leveraging their current toolchain (MCUXpresso, IAR, Keil)
- Rapid and easy prototyping and development with MCUXpresso software and tools, partner solutions and the global ARM ecosystem
- Single voltage input simplifies power circuit design
- Scalability to LPC, Kinetis & i.MX RT products

i.MX 8M 系列可应用于中阶AGV应用场景



i.MX 8M 系列产品对比表

Product	i.MX 8M Quad / Quad Lite	i.MX 8M Mini / Mini Lite	i.MX 8M Nano / Nano Lite	i.MX 8M Nano UltraLite	i.MX 8M Plus
Production Status	Mass Production	Mass Production	Mass Production	Mass Production	Mass Production
Main CPU	2-4xA53 1.5GHz, 1MB L2	**1-4xA53 1.8GHz, 512KB L2**	1-4xA53 1.5GHz, 512KB L2	1-4xA53 1.4GHz, 512KB L2	**2-4xA53 1.8GHz, 512KB L2**
MCU/DSP	M4 266MHz	M4 400MHz	M7 up to 750MHz	M7 up to 750MHz	**M7 800MHz, HiF i4 800 MHz**
DDR	x16 or x32 LPDDR4/DDR4/DDR3L	x16 or x32 LPDDR4/DDR4/DDR3L	x16 LPDDR4/DDR4/DDR3L	x16 LPDDR4/DDR4/DDR3L	**x32 LPDDR4/DDR4 Inline ECC**
GPU	3D – GC7000L (OpenGL® ES 2.1/3.0/3.1, OpenCL™ 1.2, Vulkan®)	2D – GC320L 3D – GC NanoULTRA (OpenGL ES 2.1)	GC7000UL (OpenGL ES 2.1/3.0/3.1, OpenCL 1.2, Vulkan)	-	2D - GC520L 3D – GC7000UL (OpenGL ES 2.1/3.0/3.1, OpenCL 1.2, Vulkan)
Security	CAAM, RDC, Arm® TrustZone®	CAAM, RDC, Arm TrustZone	CAAM, RDC, Arm TrustZone	CAAM, RDC, Arm TrustZone	CAAM, RDC, Arm TrustZone
AI/ML	OpenCL CPU: 32 GOPS	A53	OpenCL CPU, GPU: 32 GOPS	A53	**ML Accel 2+ TOPS**
SRAM	128KiB + 32KiB	256KiB + 32KiB	512KiB + 32KiB	512KiB + 32KiB	868KiB
Camera	2x MIPI CSI (4-lane)	1x MIPI CSI (4-lane)	1x MIPI CSI (4-lane)	1x MIPI CSI (4-lane)	**2x MIPI CSI (4-lane), 2 ISPs (3-exposure HDR)**
Display	**HDMI 2.0a Tx, MIPI DSI (4-lane), eDP; HDR10, HLG, Dolby Vision**	1x MIPI DSI (4-lane)	1x MIPI DSI (4-lane)	-	**HDMI 2.0a Tx (eARC), MIPI DSI (4-lane), 1x LVDS (8-lane)**
OSD Overlay	**4Kp60**	1080p60	1080p60	-	**4Kp30**
Video Decode	**4Kp60 HEVC, VP9, 4Kp30 H.264, legacy codecs**	**1080p60 HEVC, H.264, VP9, VP8**	None	None	**1080p60 H.265, H.264, VP9, VP8**
Video Encode	None	1080p60 H.264, VP8	None	None	1080p60 H.265, H.264
Connectivity	**PCIe, SDIO, USB**	**PCIe, SDIO, USB**	SDIO, USB	SDIO, USB	**PCIe, SDIO, USB**
Audio	20x I2S TDM (32b @384KHz), S/PDIF Tx+Rx	20x I2S TDM (32b @384KHz), 8ch PDM DMIC input, S/PDIF Tx+Rx	**12x I2S TDM (32b @384KHz), ASRC, 8ch PDM DMIC input, S/PDIF Tx+Rx**	**12x I2S TDM (32b @384KHz), ASRC, 8ch PDM DMIC input, S/PDIF Tx+Rx**	**18x I2S TDM (32b @384KHz), ASRC, 8ch PDM DMIC input), S/PDIF Tx+Rx**
Expansion I/O	**2x USB3.0, 2x PCIe Gen 2**	**2x USB2.0, 1xPCIe Gen 2**	1x USB2.0	1x USB2.0	**2x USB 3.0, 1x PCIe Gen 3**
Network, Storage	1x GbE, 2x SD/eMMC, MLC/SLC NAND	1x GbE, 3xSD/eMMC, MLC/SLC NAND	1x GbE, 3xSD/eMMC, MLC/SLC NAND	1x GbE, 3x SD/eMMC, MLC/SLC NAND	**2x GbE (1x TSN), 2x CAN-FD, 3x SD/eMMC, MLC/SLC NAND**
Process	28nm	**14nm FinFET**	**14nm FinFET**	**14nm FinFET**	**14nm FinFET**
Package	17x17mm, 0.65mm	**14x14mm, 0.5mm de-pop**	**14x14mm, 0.5mm de-pop**	11x11mm, 0.5mm de-pop	15x15mm, 0.5mm de-pop

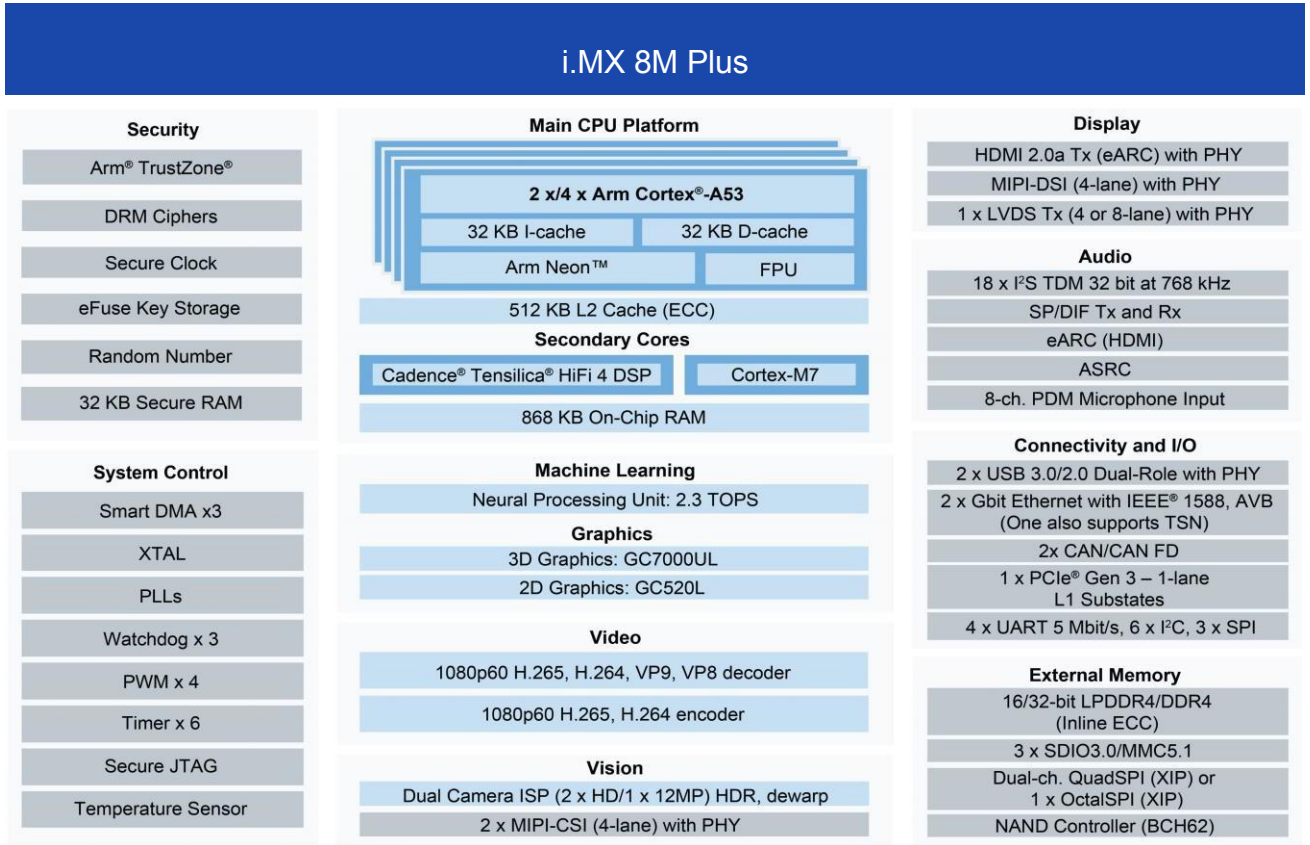
Highlighted Feature

Pin Compatible

i.MX 8M PLUS 处理器适合于中阶AGV应用，单芯片解决方案

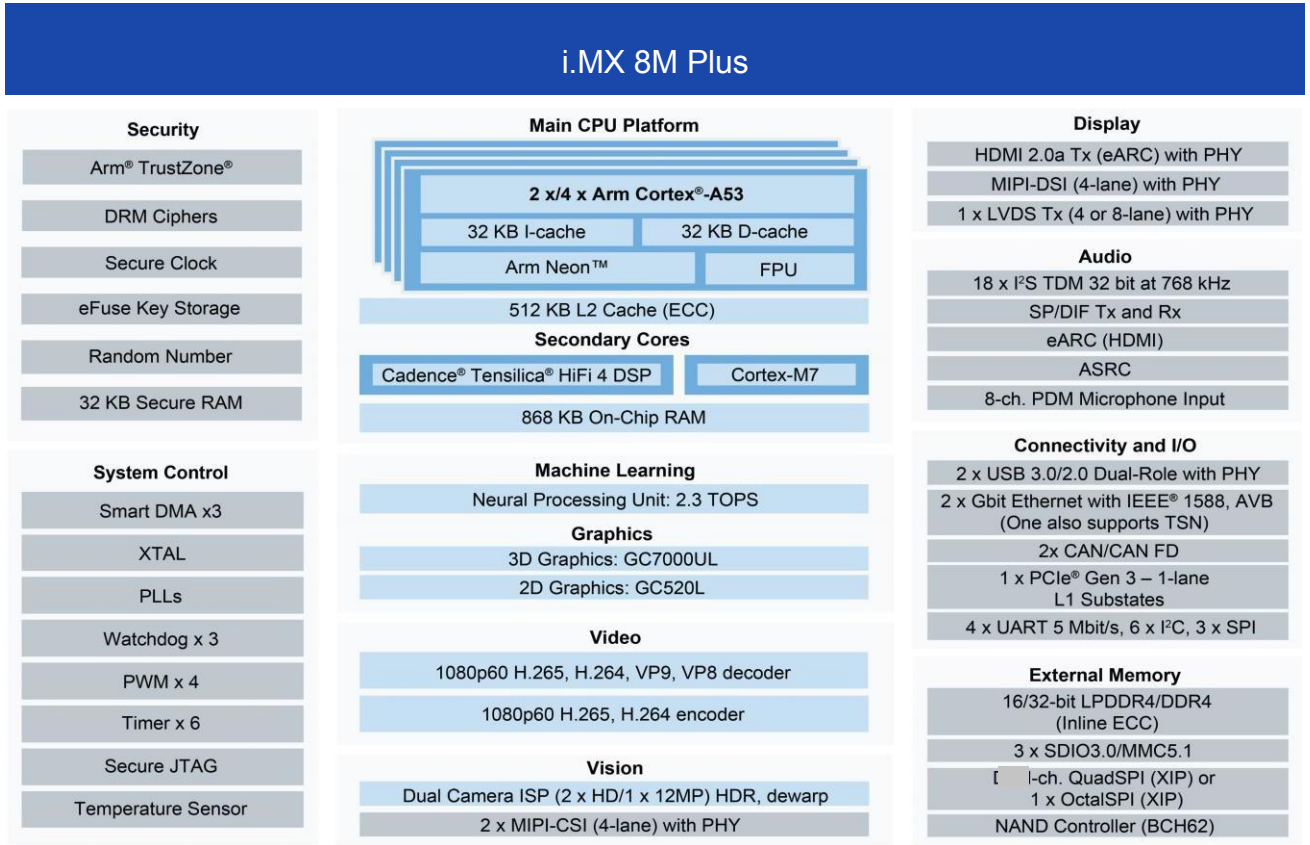
Feature Highlights:

- Quad Arm® Cortex-A53 up to 1.8 GHz
- Arm v8 fully 64-bit capable, 512KB L2 cache (ECC)
- Arm Cortex-M7 up to 800MHz with 512KB RAM (ECC)
- Voice Acceleration co-processor: HiFi 4 DSP up to 800MHz with 256KB RAM (ECC)
- Machine Learning accelerator: Neural Processing Unit (NPU) 2.3 TOPS performance
- Package: FCBGA 15x15mm, 0.5mm pitch, depop (consumer and industrial)
- Operating System targets: Linux OS, Android OS, HiFi4 SDK, FreeRTOS
- Qualification: Consumer (0C to +95C); Industrial (-40C to +105C)
- External memory:
  - 32-bit LPDDR4/DDR4 (Inline ECC)
  - 3x SDIO3.0/eMMC5.1
  - Dual-channel QuadSPI (XIP) or 1x OctalSPI
  - NAND Controller (BCH62)
  - SPI NAND
- Graphics Processing Unit (GPU):
  - GC7000UL (3D GPU, 2-shaders, OpenGL® ES 3.1, Vulkan®, Open CL™ 1.2 FP)
  - GC520L (2D GPU, OpenVG™ 1.1)



i.MX 8M PLUS 处理器适合于中阶AGV应用，单芯片解决方案

- Video Processing Unit (VPU):
  - Decode: 1080p60 H.265, H.264, VP9, VP8
  - Encode: 1080p60 H.265, H.264
- Display controllers (up to 3 simultaneous displays):
  - 1x HDMI 2.0a Tx (eARC) with PHY
  - 1x LVDS Tx (4 or 8-lane) with PHY
  - 1x MIPI-DSI (4-lane) with PHY
- Vision:
  - Camera (up to 2 cameras): 2x MIPI-CSI (4-lane) with PHY
  - Image Signal Processor (ISP): 12MP resolution, 2x187MP or 1x375MP input rate, 3-exposure HDR, Dewarp
- Audio:
  - 18x I2S TDM (32-bit @ 768KHz), DSD512, SP/DIF Tx + Rx
  - 8 channel PDM microphone input
  - eARC, ASRC
- Connectivity and I/O:
  - 2x USB 3.0/2.0 Type C with PHY
  - 1x PCIe Gen 3.0 (1 lane) with L1 Substates for fast wake from low power mode
  - 2x Gigabit Ethernet with IEEE 1588, EEE and AVB (one with TSN, but no EEE)
  - 2x CAN/CAN FD



高阶AGV系统核心技术点之一：实时性

REAL-TIME EDGE SOFTWARE OVERVIEW (FORMERLY KNOWN AS OPEN IL)

Real-Time System

- Real-time Linux enabled with PREEMPT_RT
- U-Boot based baremetal framework low latency scenarios
- Real-time operating system on Cortex-A cores and Cortex-M core for low latency
- Jailhouse as partitioning hypervisor for hardware resource partitioning

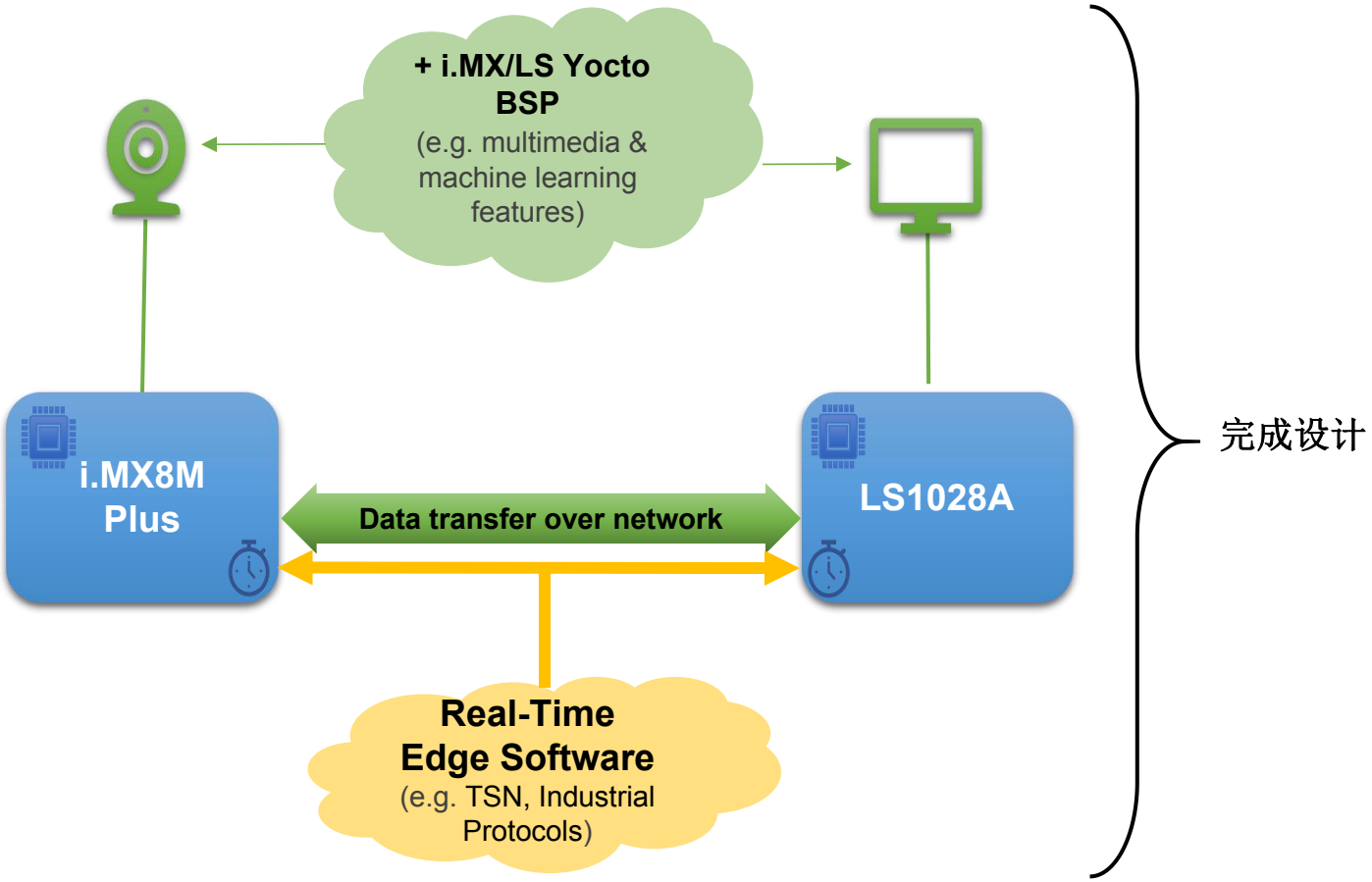
Real-Time Networking

- IEEE802.1AS clock synchronization
- TSN standards implementation
- TSN Configuration with Linux tc command and NETCONF/YANG
- GenAVB/TSN stack
- Network redundancy for seamless recovery with 802.1CB, HSR and ERPS

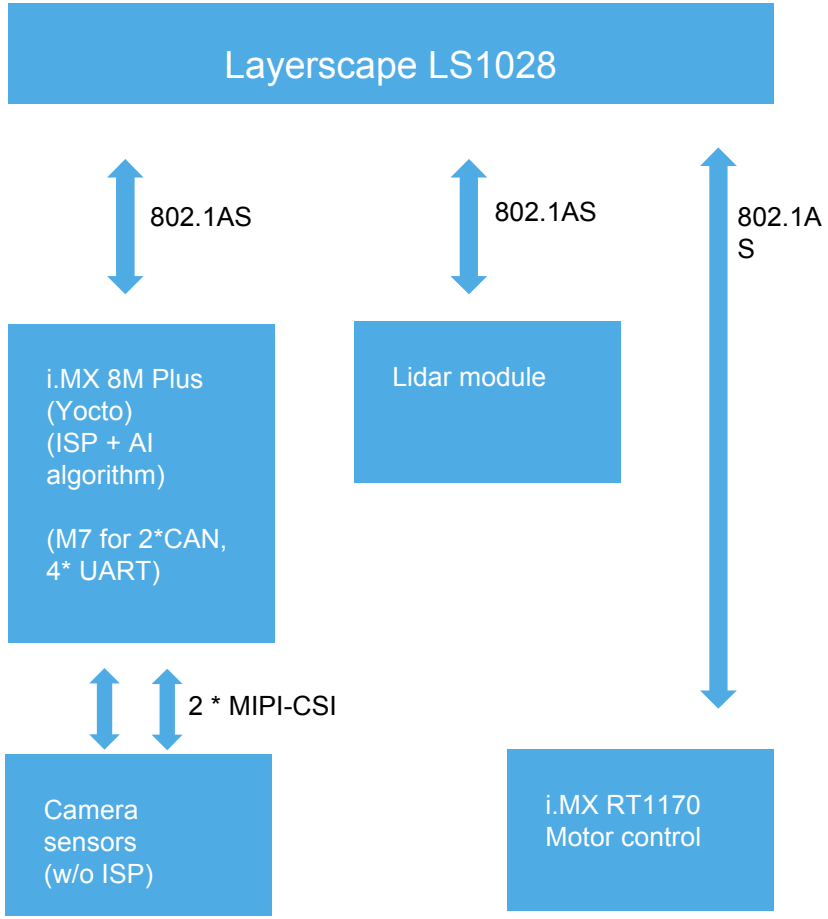
Protocols

- IGH EtherCAT master stack and EtherCAT-capable native drivers
- CANopen over EtherCAT framework based on IGE CoE interface and multiples axes control
- OPC-UA and OPC-UA pub/sub
- FlexCAN and CANopen

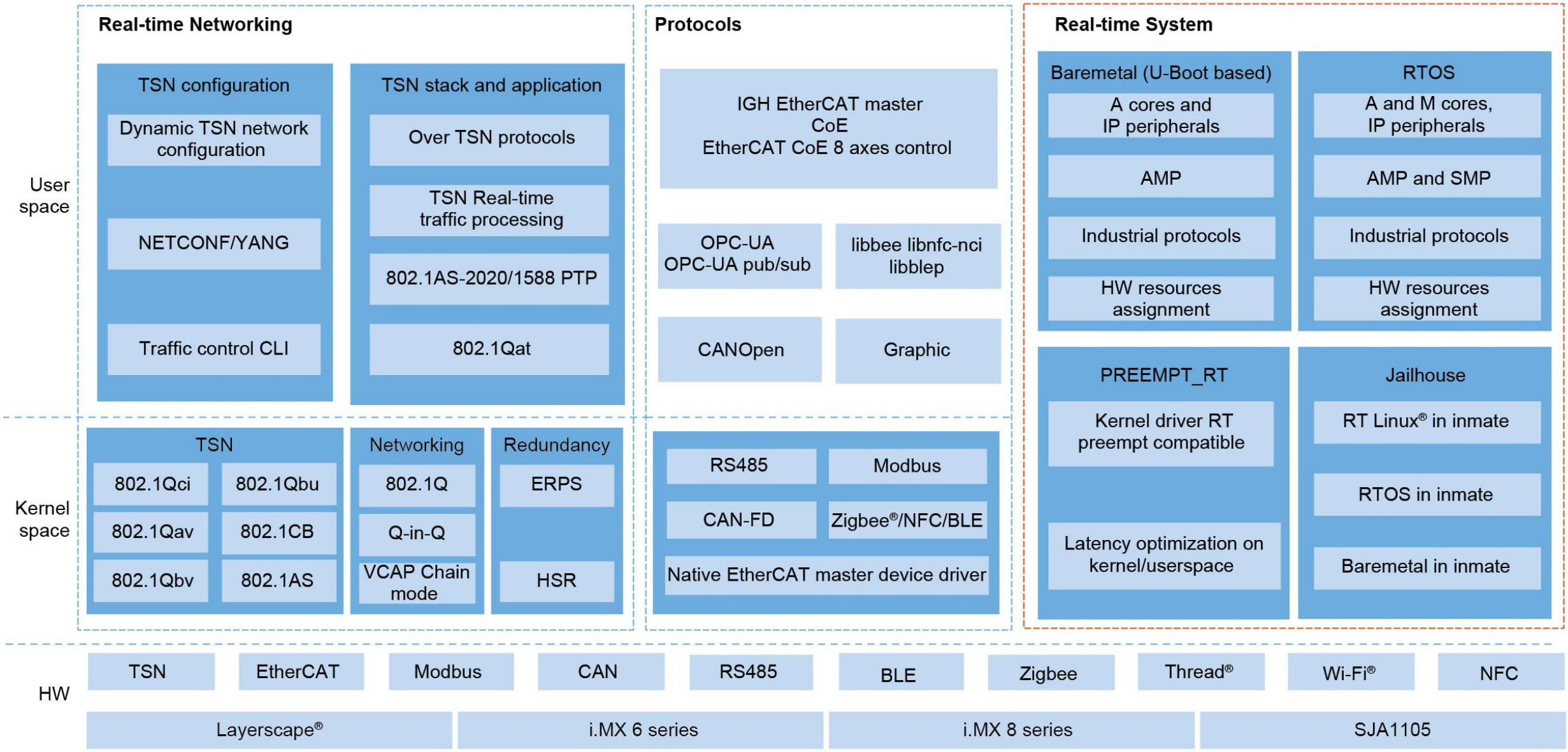
AGV系统板卡级TSN通讯



完成设计

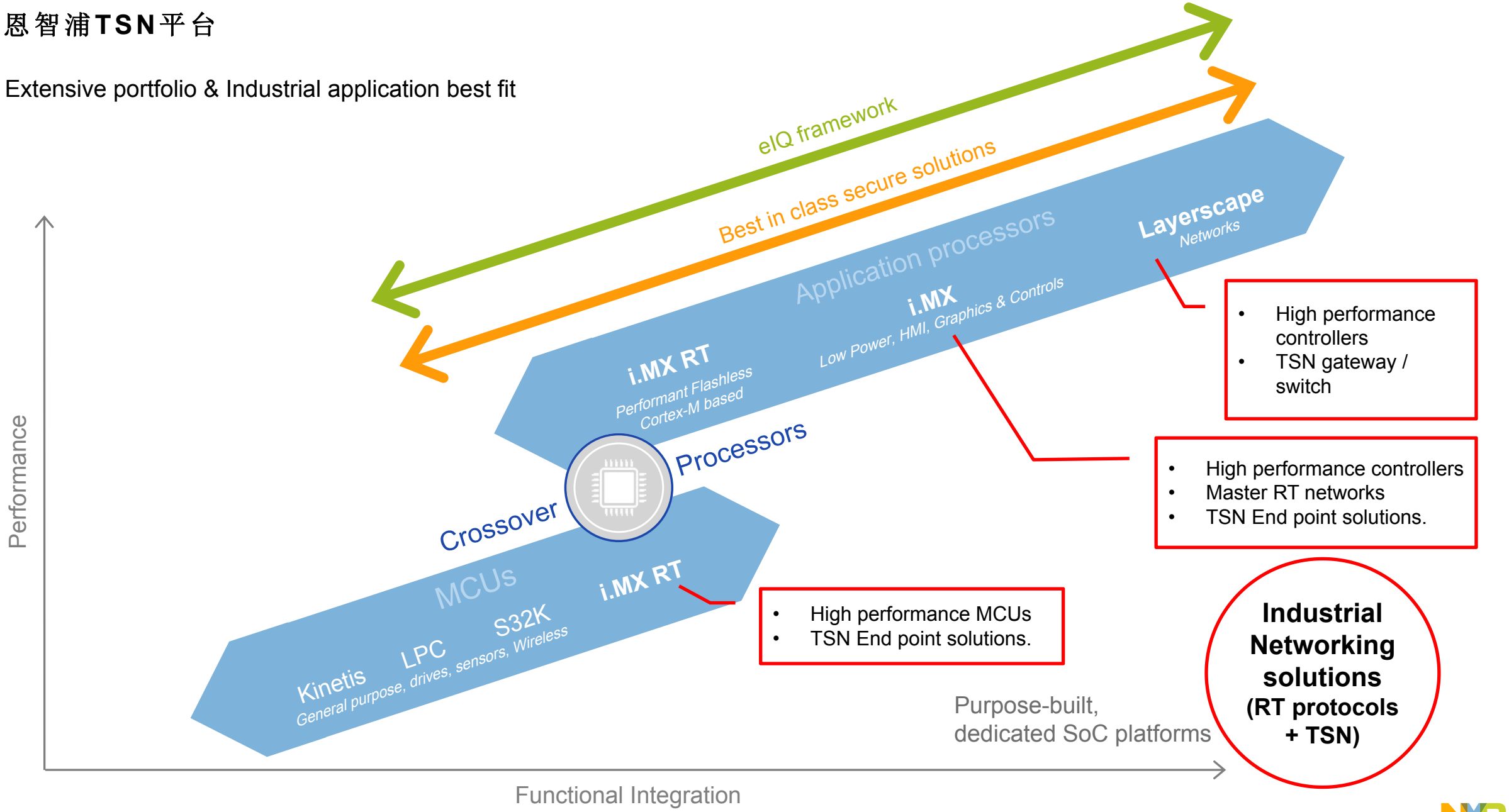


恩智浦边缘实时软件架构：REAL-TIME EDGE

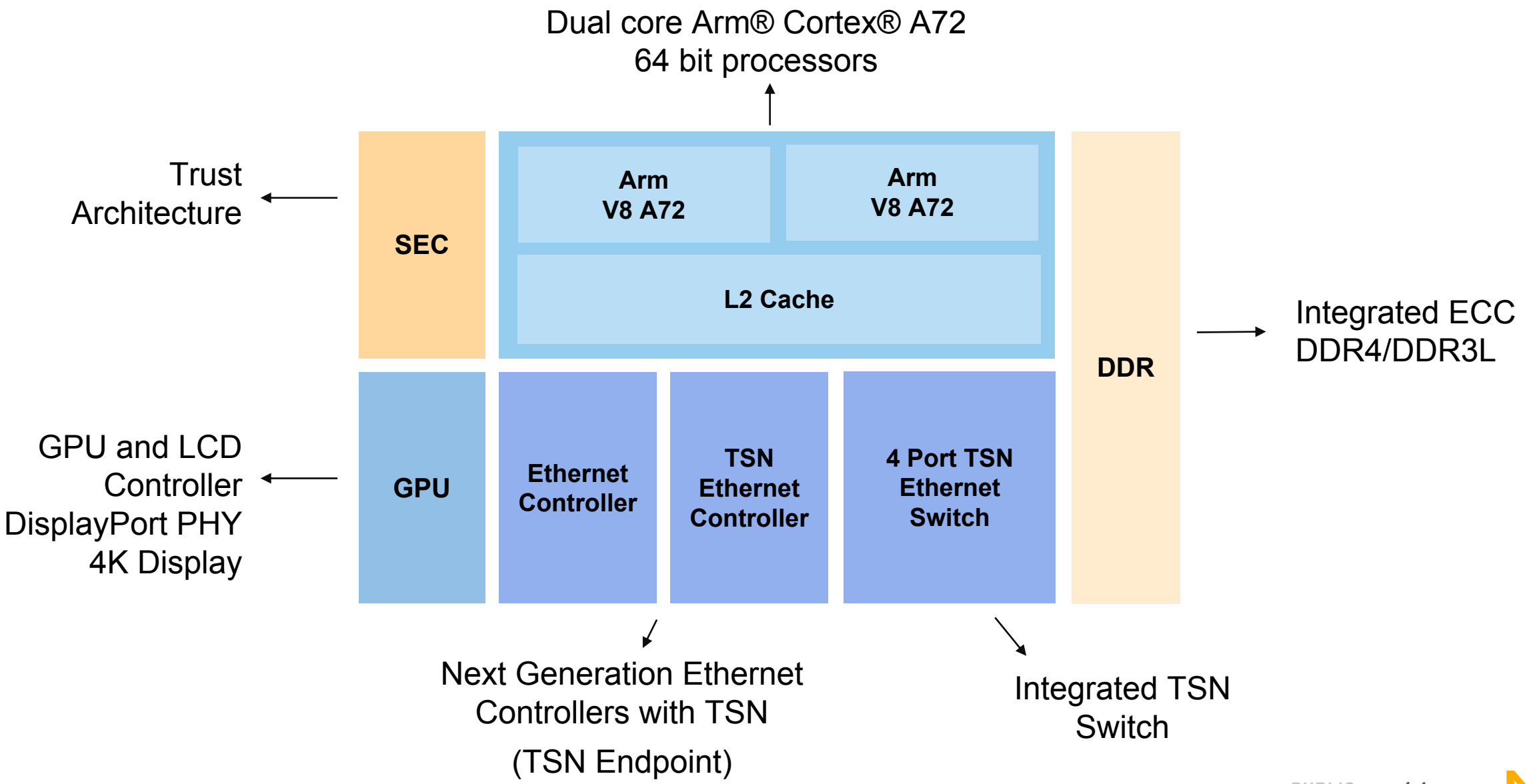


恩智浦TSN平台

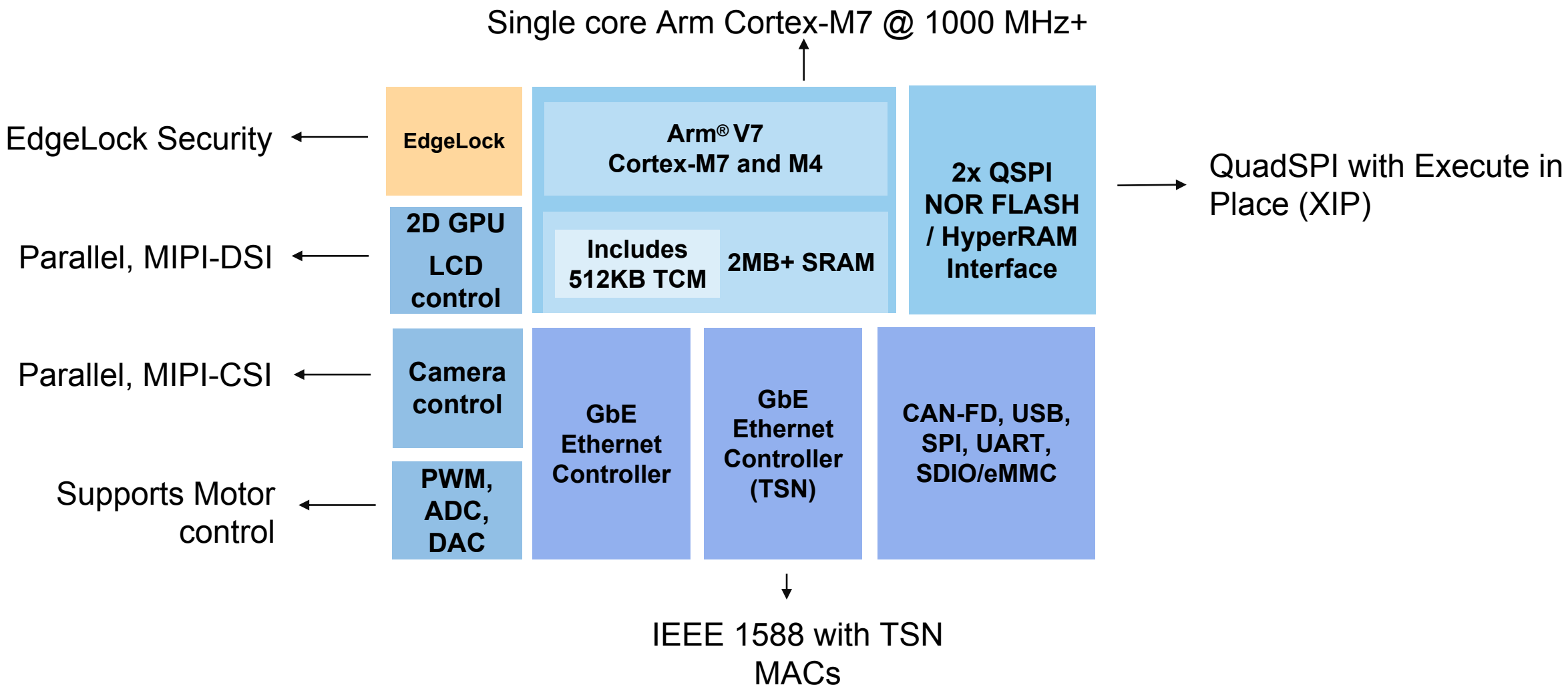
Extensive portfolio & Industrial application best fit



支持TSN网关芯片：LAYERSCAPE LS1028A



支持TSN的控制节点：i.MX RT1170 CROSSOVER PROCESSOR



支持TSN的AI算力节点 i.MX 8M Plus

Machine Learning Accelerator (1GHz)

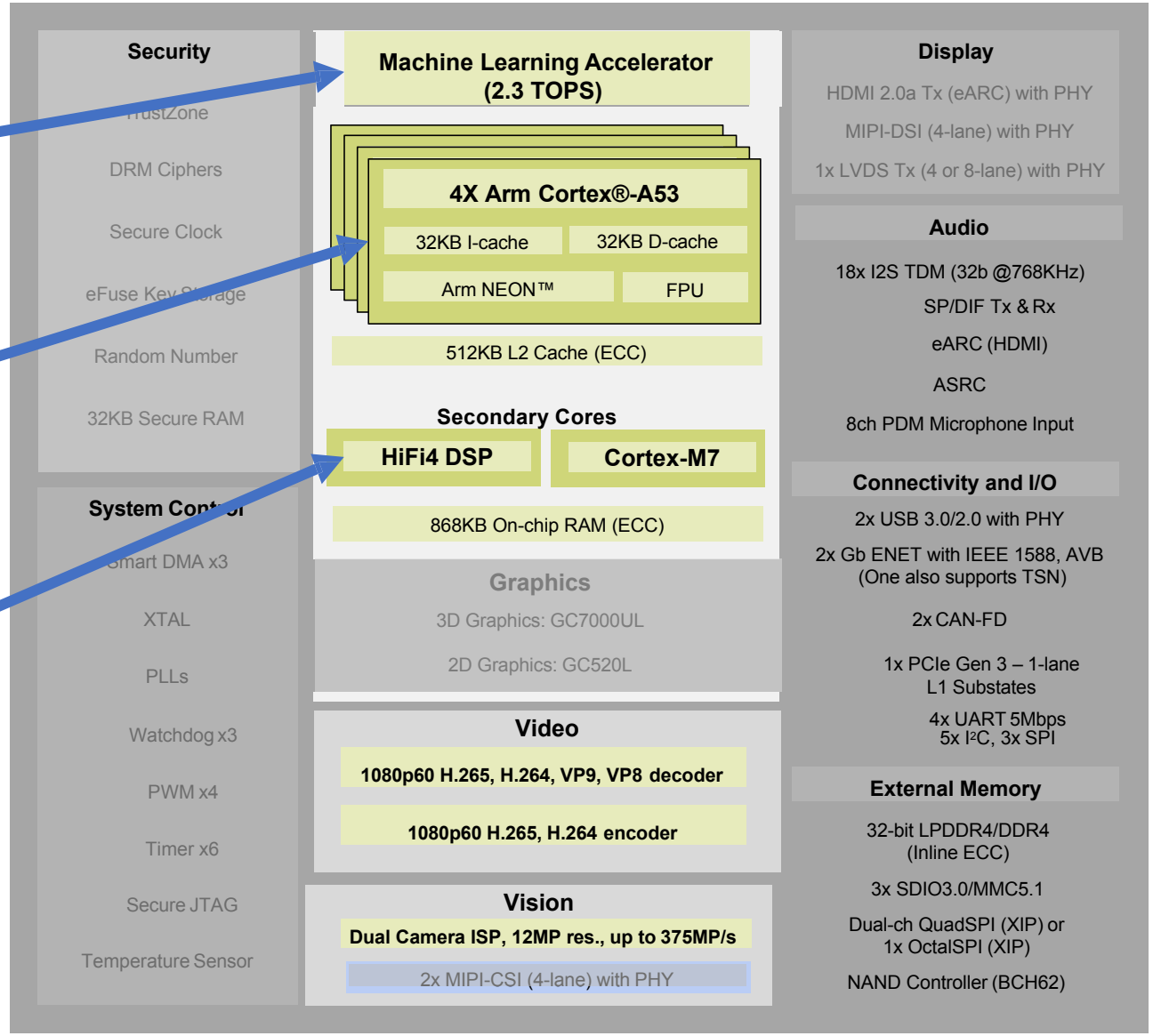
- Primary Use: Multi-camera classification/detection

Quad Arm® Cortex-A53 (up to 1.8GHz)

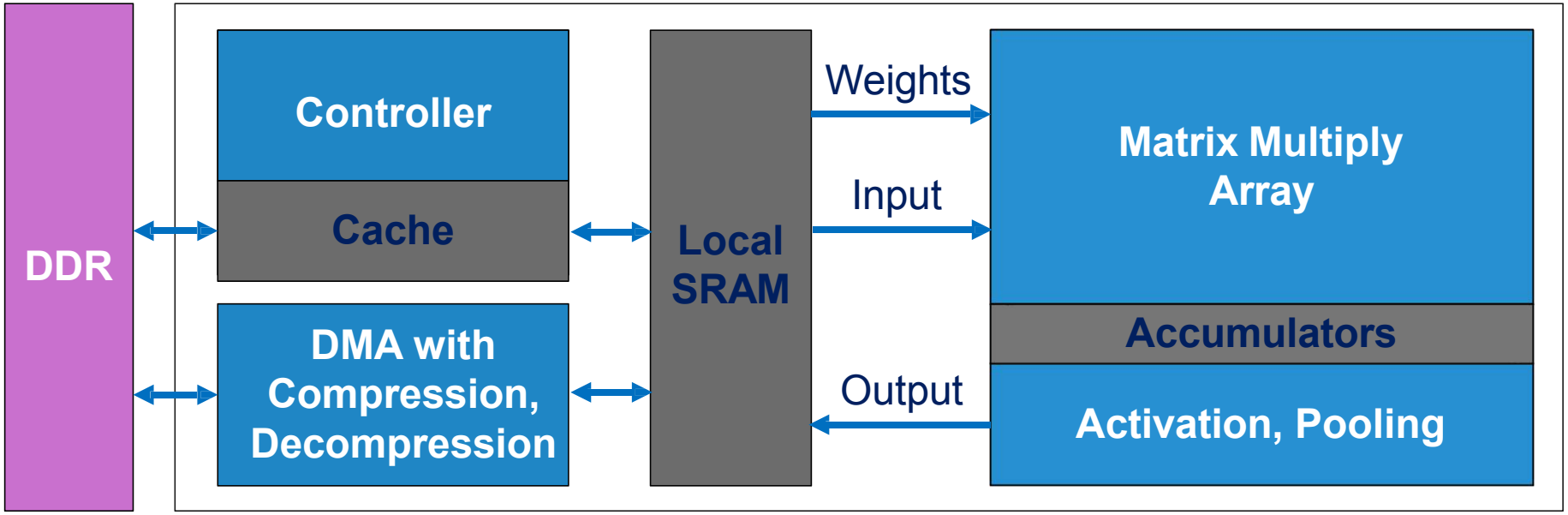
- Primary Use: Speech command recognition, object detect/classification

Cortex-M7+HiFi4 DSP (800MHz)

- Primary Use: Keyword detection, sensor fusion



i.MX 8M PLUS 神经网络处理单元



- Scalable 8- and 16-bit integer multiply-accumulate (MAC) engine for tensor operations
- CPU for non-convolutional layers and custom operations
- Data compression: Reduces bandwidth between NPU and system DRAM
- Specialized NN hardware supports activation and pooling Supports variety of NN topologies:
  - Convolutional (CNN): MobileNet, YOLO, etc.
  - Recurrent (RNN, GRU, LSTM): Deep Speech 2, etc.

i.MX 8M PLUS ISP（图像处理单元）

- **Image Signal Processor (ISP) basic function**
  - Converts the image color code from raw Bayer (output of the image sensor) to YUV so it can be processed by the SoC
  - Also provides additional processing to improve the image quality:
    HDR extracts maximum image detail in high contrast scenes
    De-Warp: Fisheye lens or low-cost lens geometry correction
    Image Enhancement
- **ISP Benefits:**
  - i.MX 8M Plus's ISP architecture enables low latency and high performance. ISP receives pixels directly from camera input.
  - ISP embedded in the Applications Processor provides a lower cost vision system. No need for an external ISP IC.
  - Integrated ISPs in the processor support a higher product longevity. i.MX 8M Plus longevity is 15 years.
  - High-definition camera modules rarely come with embedded ISP, rely on a processor ISP.

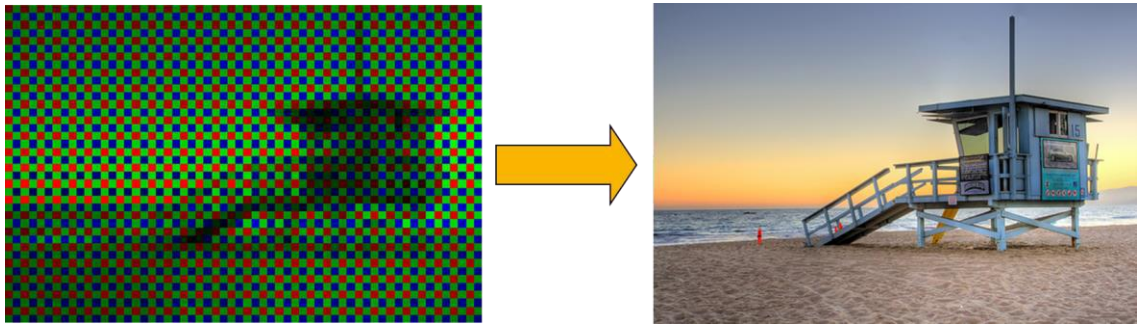
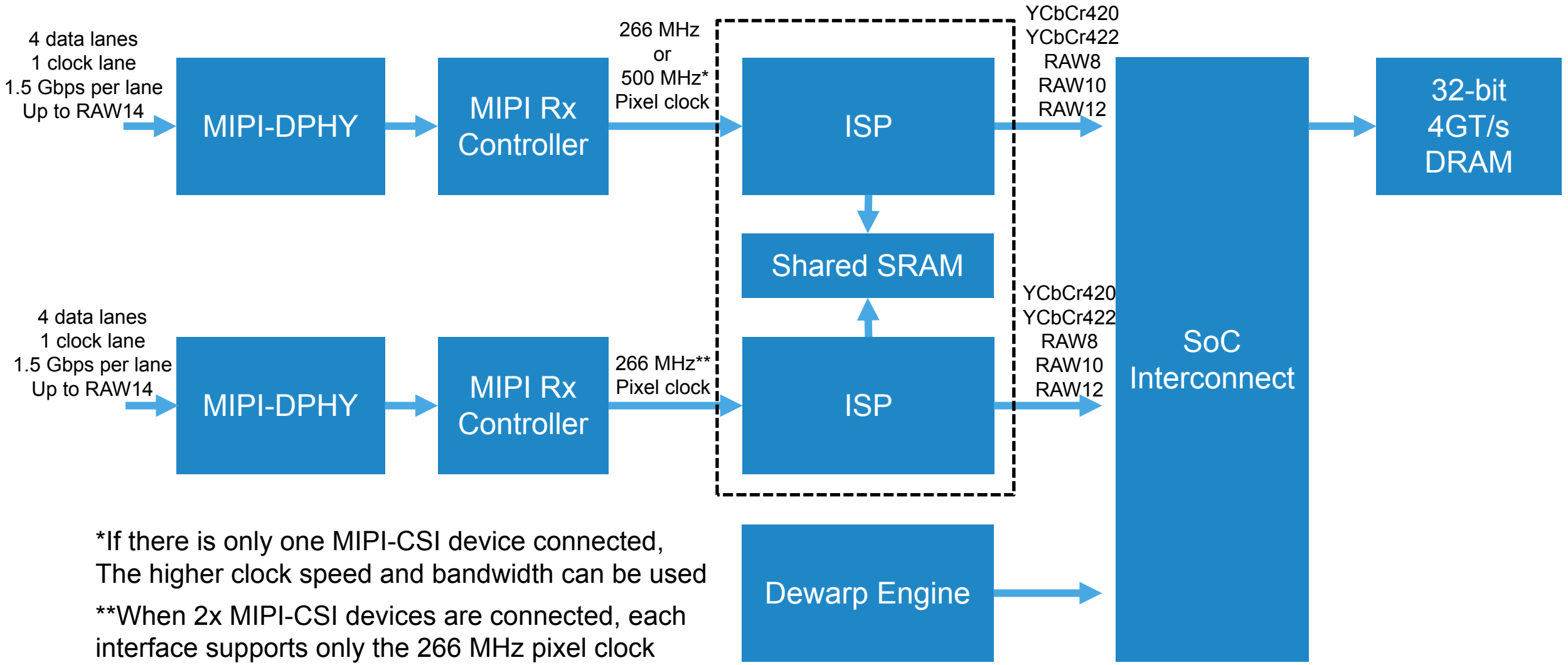


Image Sensor Output　　ISP Output

i.MX 8M PLUS ISP（图像处理单元）应用框图



*If there is only one MIPI-CSI device connected, The higher clock speed and bandwidth can be used

**When 2x MIPI-CSI devices are connected, each interface supports only the 266 MHz pixel clock

i.MX 8M PLUS 神经网络处理单元

Programmable Engine Unit
(1 instance on i.MX8M Plus)

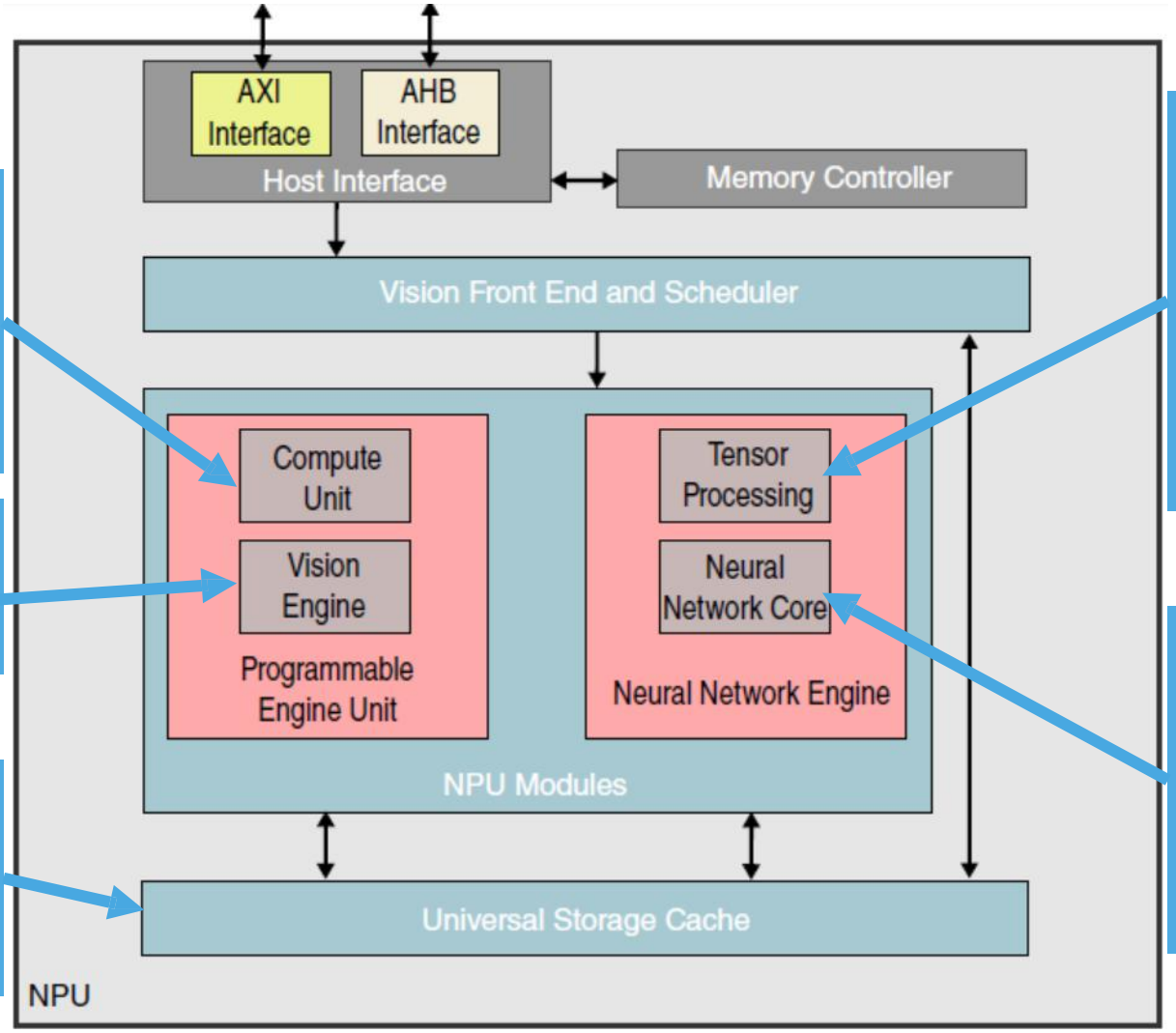
128-bit vector processing.
INT 8/16/32b, FLOAT 16/32b.
Most flexible programming unit to handle everything else

Vision Engine

Provides advanced image processing functions.

Universal Storage Cache
(1 instance on i.MX8M Plus)

Local memory and L1 cache to pass data amongst NPU modules



Tensor Processing Core
(3 instances on i.MX8M Plus)
INT 8/16b, FLOAT 16b

Non-convolution layers.
Multi-lane processing for data shuffling, normalization, pooling/unpooling, LUT, etc.
Network pruning support, zero skipping, compression

Neural Network Core
(6 instances on i.MX8M Plus)
2.3 TOPs INT8

Handle Convolution Layer + RELU + Max Pooling, and Compute Bounded Fully Connected Layers

EIQ™机器学习软件开发环境

嵌入式开发人员

数据分析师

机器学习专家

PyTorch Caffe2 ONNX

eIQ 工具集

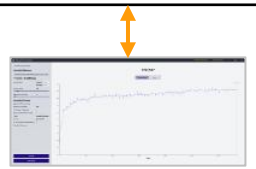
eIQ 用户接口

模型优化工具

提供用户自己的数据处理流程

数据整理

模型选择



模型训练、量化、优化

模型验证、部署

提供用户自己的模型工作流程

PyTorch Caffe2 ONNX

eIQ inference

Applications running on EdgeVerse processors

异常检测 | 声音识别 | 人脸识别 | 多目标分类 | 复杂情况诊断

eIQ inference with…

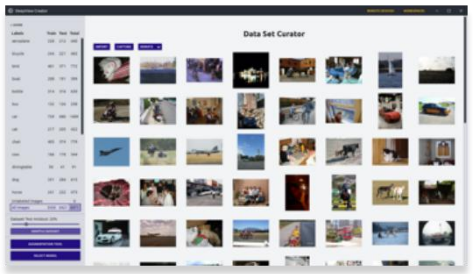
TFLite | TFMicro | Glow | DeepView™

Arm® Cortex®-M, DSP

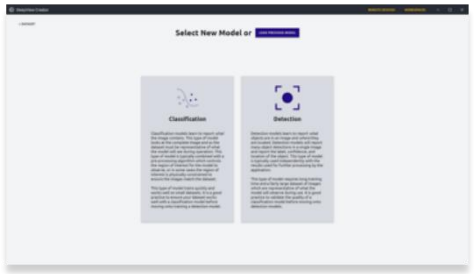
TFLite | Arm NN | ONNX Runtime | DeepView™

Arm Cortex-A, GPU, DSP, NPU, microNPU

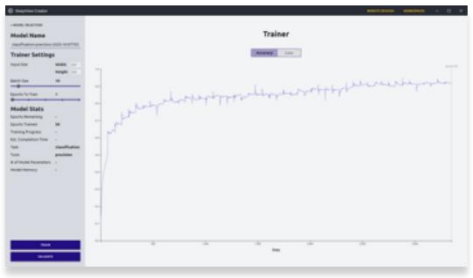
基于原始数据： DeepView™ **Bring Your Own Data** Workflow **(BYOD)**



数据导入与标定



模型选择，硬件平台选择



模型训练，优化，量化

模型验证与测试

基于已有模型： DeepView™ **Bring Your Own Model** Workflow **(BYOM)**

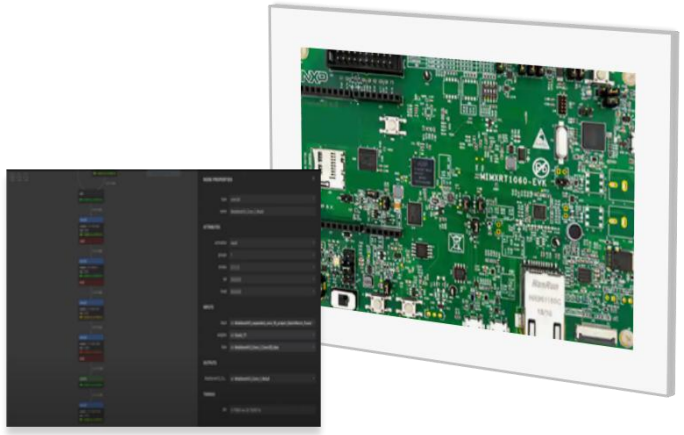
PyTorch

ONNX

Caffe2

公有或私有模型

模型格式转换，优化，量化



嵌入式板级推理实现

基于已有模型: DeepView™ **Bring Your Own Model** Workflow **(BYOM)**

Source/ Destination	RTM	TFLite (MLIR/TOCO)	ONNX	RTM q	TF:OTE q	ONNX q
1.x Pb	Y	Y	Y	N	Y	Y
Saved Model (Folder/TAR)	Y	Y	Y	Y	Y	Y
Keras (h5)	Y	Y	Y	Y	Y	Y
RTM		N	N		N	N
TFLite (MLIR/TOCO)	Y		Y	N	N	Y
ONNX	Y	Y		N	Y	Y
TFLite q (MLIR/TOCO)	Y		Y	N		
ONNX q	Y	Y		N		

Notes: (1) ***RTM*** *is for DeepViewRT*
(2) ***q*** *is for quantized*

MPU硬件平台： i.MX 8M Plus 采用 TensorFlow Lite INT8 模型性能分析

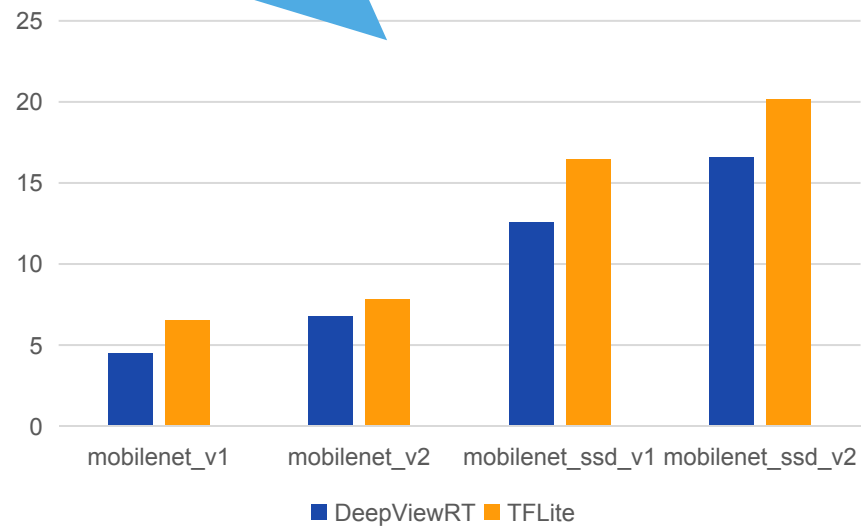
Hardware: i.MX 8M Plus EVK
Software: Linux_IMX_5.4.70-2.3.2, TFLite 2.4.0, DeepViewRT 2.4.24

Models INT8 model performance on i.MX 8M Plus CPU and NPU	Time per inference (ms)	
	1x A53 (1.8 GHz) CPU	NPU
inception_v1_INT8	343.48	5.80
inception_v2_INT8	529.75	7.35
inception_v3_INT8	1266.28	17.12
inception_v4_INT8	2658.42	33.17
mobilenet_v1_100_224_INT8	153.49	2.72
mobilenet_v2_100_224_INT8	119.68	3.02
deeplabv3_mnv2_pascal_quant	1089.63	200.08
ssd_mobilenet	210.42	12.76
ssd_mobilenet_v2_coco_quant	273.32	16.29
ssd_mobilenet_v2_face_quant	269.87	7.31

Models INT8 model performance on i.MX 8M Plus NPU	Time per inference (ms)	
	DeepViewRT	TFLite
mobilenet_v1_1.0_224 (Per channel)	4.51	6.53
mobilenet_v1_1.0_224 (Per tensor)	3.38	-
mobilenet_v2_1.0_224 (Per channel)	6.75	7.84
mobilenet_v2_1.0_224 (Per tensor)	4.01	-
mobilenet_ssd_v1_300(Per channel)	12.59	16.46
mobilenet_ssd_v1_300(Per tensor)	8.67	-
mobilenet_ssd_v2_300(Per channel)	16.57	20.15
mobilenet_ssd_v2_300(Per tensor)	10.92	-

专用硬件加速器NPU可提高约 50倍 的推理性能！

eIQ DeepViewRT 相对于 TFlite，可以提高约 20%



总结：恩智浦处理器适用于各类AGV场景

- 入门级
  - 基础避障及路径引导算法，采用标签等方式循迹
  - 或基于单目的简单 vSLAM 算法
  - NXP i.MX RT1170

- 中阶
  - 采用双目视觉避障以及路径引导算法
  - 具有完善的 vSLAM 算法能力，具有视觉AI能力
  - NXP i.MX 8M Plus

- 高阶
  - 复杂系统项目，板载高速网关，用于多模块数据交互
  - 工业级相机，激光雷达，毫米波雷达等多种传感器通过TSN技术接入网关
  - 驱动模块通过TSN技术接入网关
  - 强大的片上数据处理能力
  - NXP LS1028/LS1046 + i.MX 8M Plus + i.MX RT1170

SECURE CONNECTIONS
FOR A SMARTER WORLD