2020 embedded VISION summit

Modern SoCs for Consumer Robotics and AIoT

Zhe Zhang
Trifo
September 2020

trifo

Trifo – an AI home robot company

2020 embedded VISION summit

2016

Foundation

- Founded in Silicon Valley, CA
- A genuine AI home robot company

2017

Ironsides

B2B Model

- Visual Inertial Computing Module for various robotics applications

2018

Ironpie

1st Robot Vacuum

- Smart navigation system

Jan 2020

Max

Home Surveillance Robot Vacuum

- Advanced surveillance feature with motion and audio detection
- 2-way speaker/mic to enhance human-robot interaction
- VSLAM & multiple sensor fusion

Mar 2020

Emma

Entry-level Robot Vacuum

- Smart navigation
- More powerful suction and battery life

May 2020

Lucy

1st generation AI home robot

- Obstacle recognition and avoidance
- Smart room recognition & segmentation
- Day/night surveillance capabilities

Q4 2020

Trifo Home+

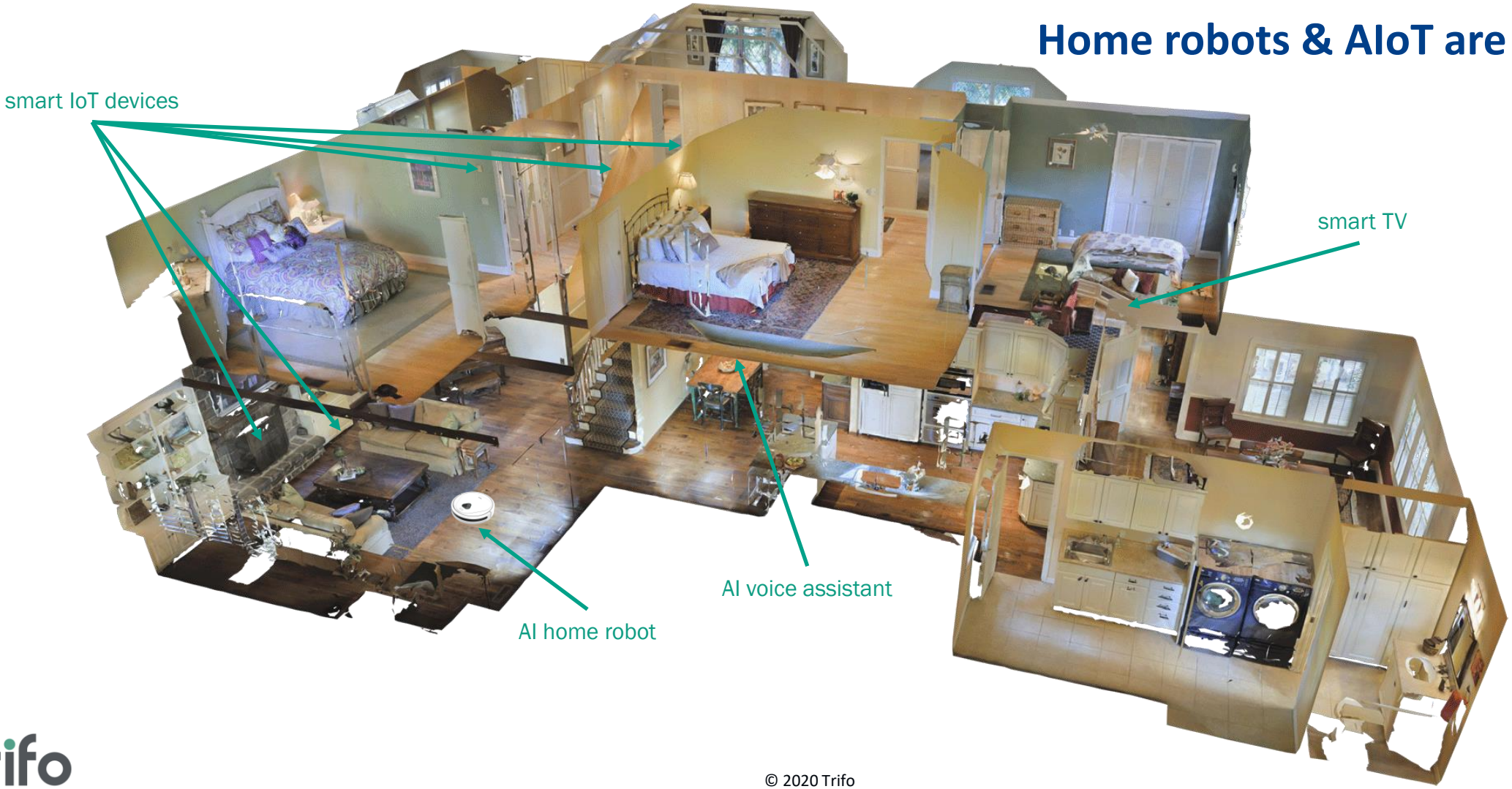
AI home robot ecosystem

- Rich hardware/software extensions based on Lucy

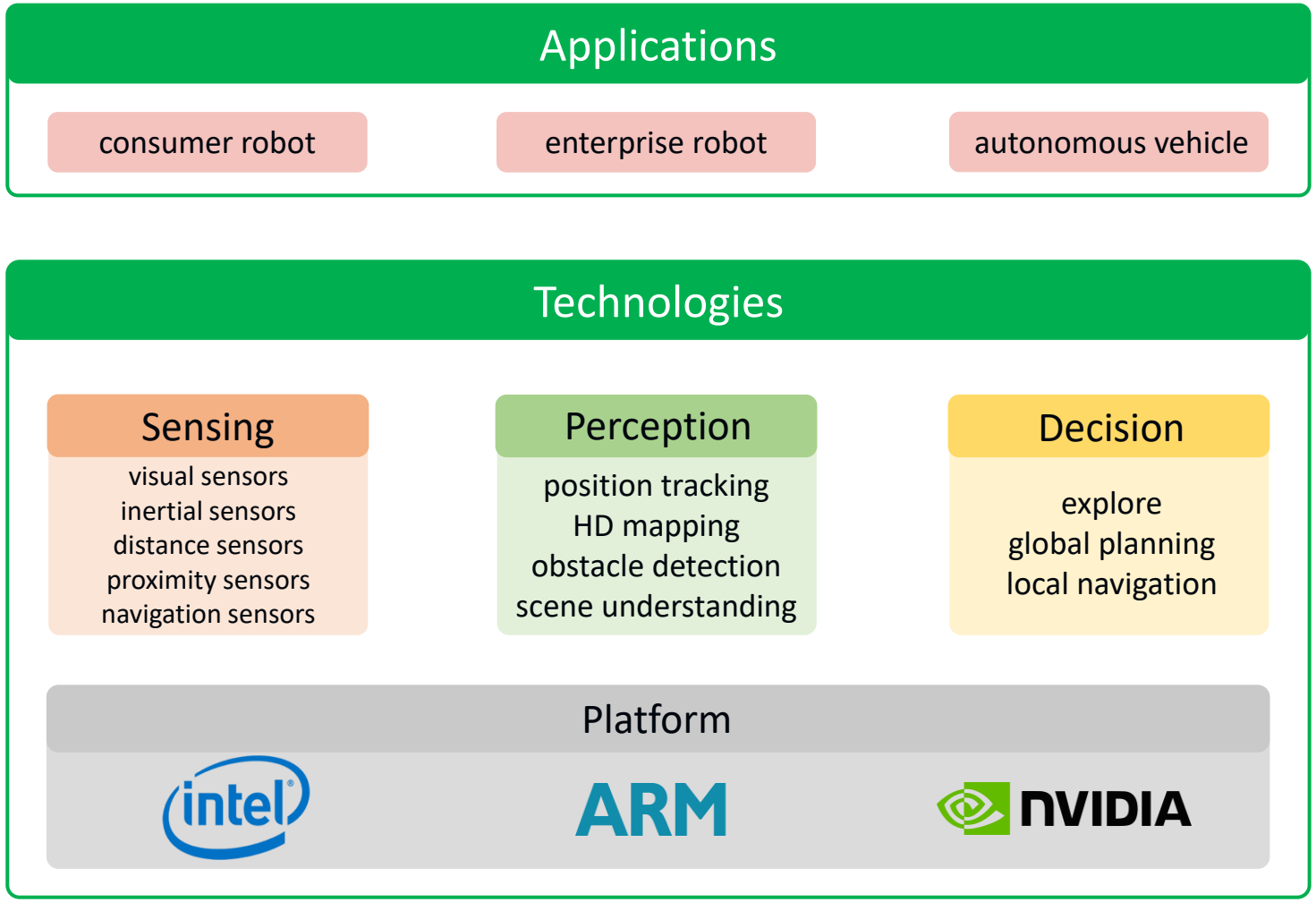
trifo

Market

Home robots & AIoT are merging.

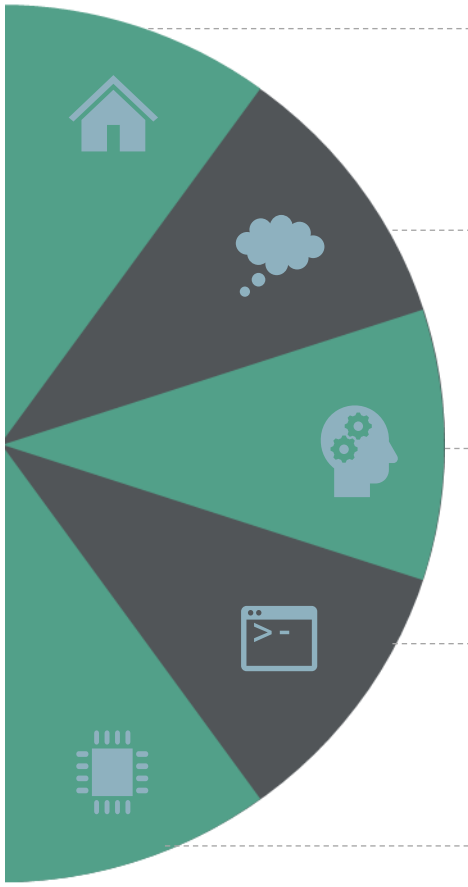


Robotics Technologies & Applications



- Flexible core functionalities for fast product integration
  - The processor needs to provide the computing power in a flexible way for the core functionalities.
- Fully optimized for each application
  - The processor needs to have the flexibility for various applications.
- Hardware/software co-design
  - The processor design needs to consider the future software running on it.
- Accurate factory calibration
  - hardware scalability
- State-of-the-art proprietary algorithms
  - Algorithm-driven chip design sounds crazy but it might make sense.
- Highly optimized implementation
  - SSE / NEON / CUDA
  - accelerate deep learning on HAL

Robotics Architecture



Extension & Communication

- peripheral products for different family members: elder, men/women, children, pets
- customized applications for specific functions

Cloud AI

- deep learning on cloud: advanced training and inference services
- user management: basic information, home data, customized service access/integration

Edge AI (edge computing power for the processor)

- 3D geometry: SLAM, room reconstruction, obstacle avoidance
- scene understanding: obstacle/object/room/human recognition/classification
- decision: unknown environment exploration, global planning, local navigation

Run-Time (RTOS or not, is a question for the processor)

- run-time: low latency, smart dynamic resource allocation
- deep optimization: instruction set level optimization, hardware acceleration

Hardware

- specially made chassis
- customized “eyes”
- customized “brain”

Extension

Decision

Sensing

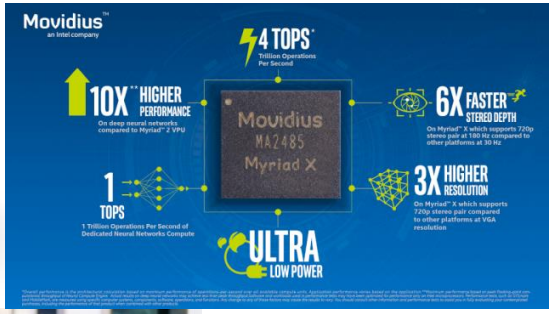
Communication

Connection

Movement

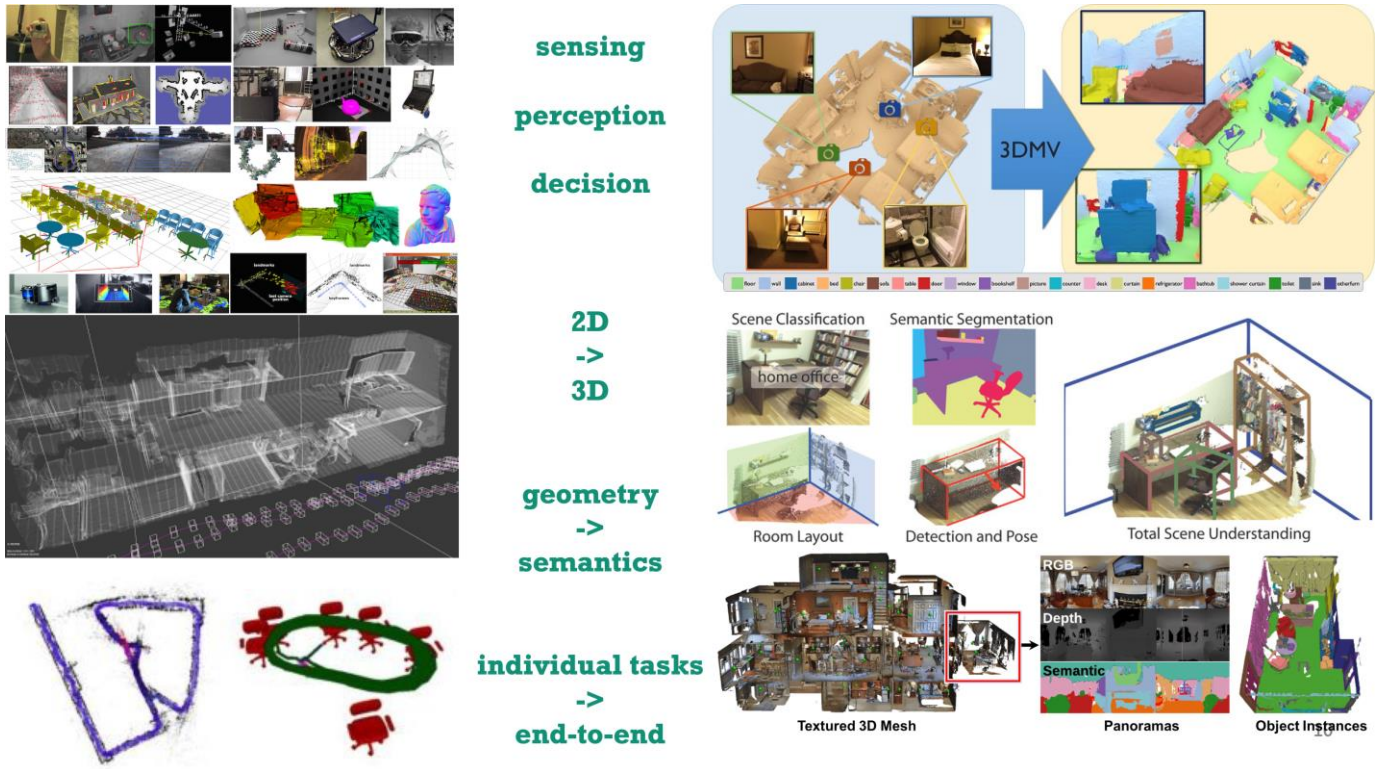
“Robotics SoC” is …?

- Microcontroller: ARM Cortex-M
- Real-time: ARM Cortex-R (depending on the need of real-time)
- Application: ARM Cortex-A (32-bit and 64-bit)
- High-end SoC with “NPU” or “edge AI chip”
  - Nvidia Jetson AGX Xavier
  - Qualcomm Robotics RB3
  - Intel RealSense + Movidius



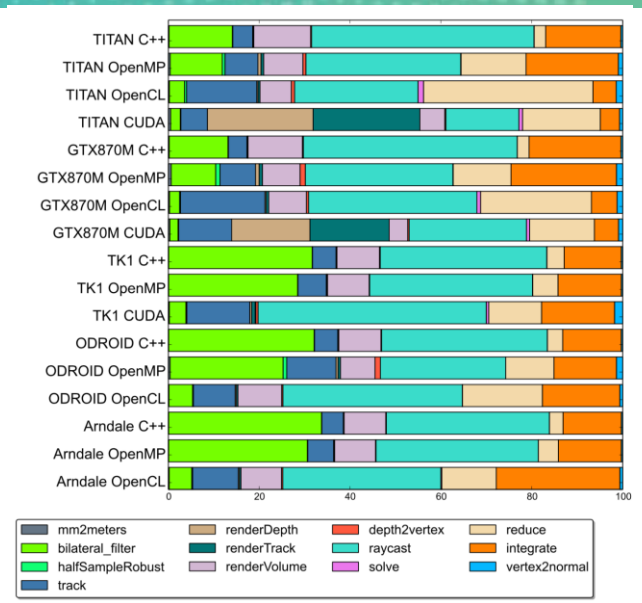
“Robotics SoC” does …?

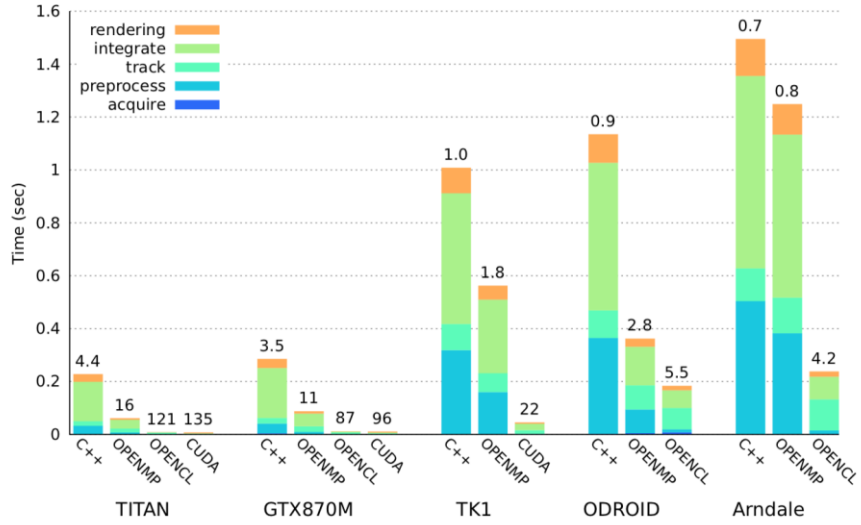
- Microcontroller: the most energy-efficient embedded devices
- Real-time: reliable mission-critical performance
- Application: supreme performance at optimal power
- High-end SoC with “NPU” or “edge AI chip”: AI specific computing needs



Public Benchmark - geometry

- SLAM global optimization is hard to parallelize so GPU is not quite useful.
  - SLAM has a lot of “if-else” logic in real product.
  - The sparse matrix operations in SLAM optimization can’t be easily parallelized.
- Run-Time performance is super important and impacts algorithm performance.
- Real product’s SLAM system always has hardware dependency.
  - In a robotic system, sensing highly depends on sensors.
  - The perception/decision in a robotic system will adapt accordingly.



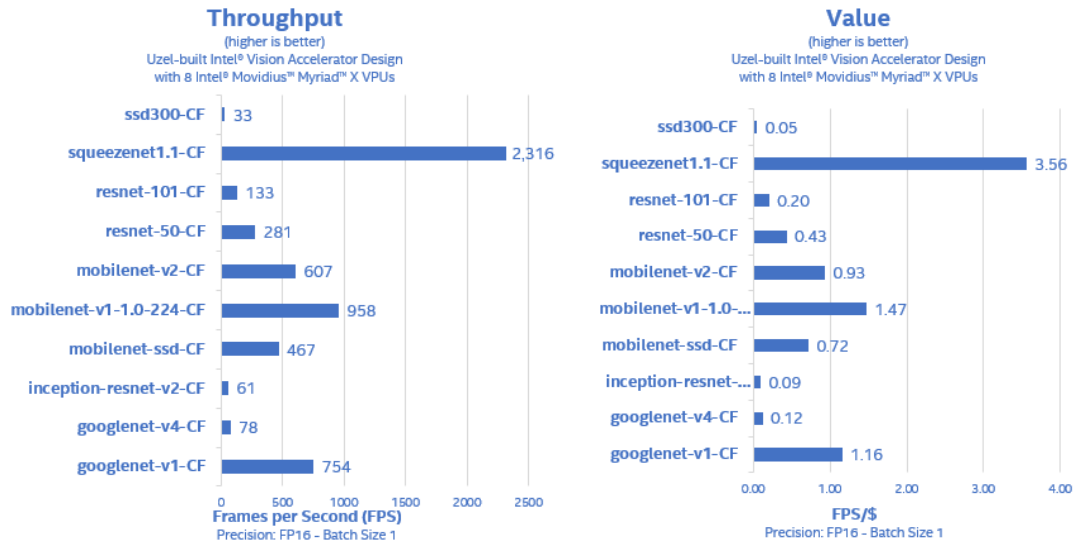


Public Benchmark - geometry

- Take a robotic system with SLAM functionality as an example
  - Is it a demo or a product? (no joking at all)
  - User experience & application scenario decide your SLAM & chip choice.
  - Which is more important in SLAM, “L” (localization) or “M” (mapping)?
  - What is your BOM budget for sensors and chips?
  - What is more important, perfect performance or robustness with some perf sacrifice?
  - Does the robotic system need deep learning capabilities?
  - Safety? Privacy? Security? Localization? Apps?
- “Hey we are talking about chips, why do you have so many questions for those other things?”
  - “Because it matters.”

Public Benchmark - semantics

- GPU helps A LOT!
- A lot of work has been done by framework provider.



Intel Vision Accelerator with Movidius Myriad X VPUs

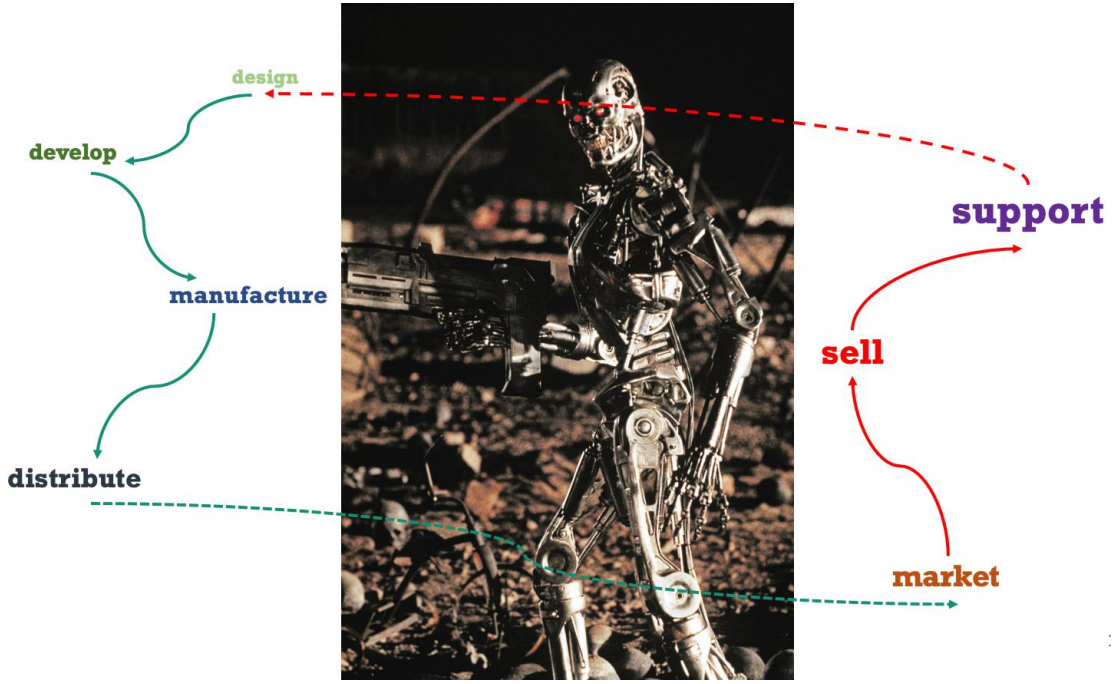
Nvidia Jetson AGX Xavier DL Inference

15W Mode

NETWORK	BATCH SIZE	PERF (img/sec)	LATENCY (ms)	MODULE POWER (watts)	MODULE PERFORMANCE / watt	FUTURE PERF (img/sec)	FUTURE MODULE POWER (watts)	FUTURE MODULE PERFORMANCE / watt
ResNet-50	1	358	2.8	11.5	31.2	800	12	67
ResNet-50	2	508	3.9	12.8	39.7	1090	14	78
ResNet-50	4	634	6.3	13.6	46.5	1280	14	91
ResNet-50	8	717	11.2	14.4	49.8	1360	14	97
ResNet-50	16	767	20.9	14.9	51.3	1410	15	94
ResNet-50	32	841	38.0	15.1	55.7	1430	15	95
ResNet-50	64	869	73.6	15.1	57.6	1430	15	95
ResNet-50	128	879	145.7	15.2	57.7	1430	15	95
VGG19	1	84	11.9	14.2	5.9	230	12	19
VGG19	2	132	15.2	14.4	9.1	290	13	22
VGG19	4	174	22.9	14.6	11.9	320	13	25
VGG19	8	191	41.8	14.9	12.8	340	13	26
VGG19	16	231	69.4	15.0	15.3	350	13	27
VGG19	32	260	123.1	15.2	17.1	350	13	27
VGG19	64	269	238.0	15.3	17.6	350	13	27
VGG19	128	274	467.8	15.4	17.8	350	13	27
GoogleNet	1	542	1.8	9.8	55.0	1310	11	119
GoogleNet	2	684	2.9	10.4	65.8	1670	13	128
GoogleNet	4	890	4.5	11.4	78.1	1920	15	128
GoogleNet	8	1015	7.9	12.0	84.4	1940	15	129
GoogleNet	16	1121	14.3	12.8	87.6	1950	15	130
GoogleNet	32	1184	27.0	13.2	90.0	1980	15	132
GoogleNet	64	1235	51.8	13.2	93.6	1980	15	132
GoogleNet	128	1255	102.0	13.3	94.3	1980	15	132
AlexNet	1	299	3.3	14.0	21.3	1090	12	91
AlexNet	2	466	4.3	14.3	32.6	1790	12	149
AlexNet	4	721	5.5	14.9	48.5	2650	13	204
AlexNet	8	990	8.1	13.5	73.4	3510	13	270
AlexNet	16	1291	12.4	14.2	90.8	4200	14	300
AlexNet	32	1713	18.7	14.4	119.0	4670	14	334
AlexNet	64	2087	30.7	14.8	141.3	4670	14	334
AlexNet	128	2270	56.4	14.9	152.5	4670	14	334

Challenges!

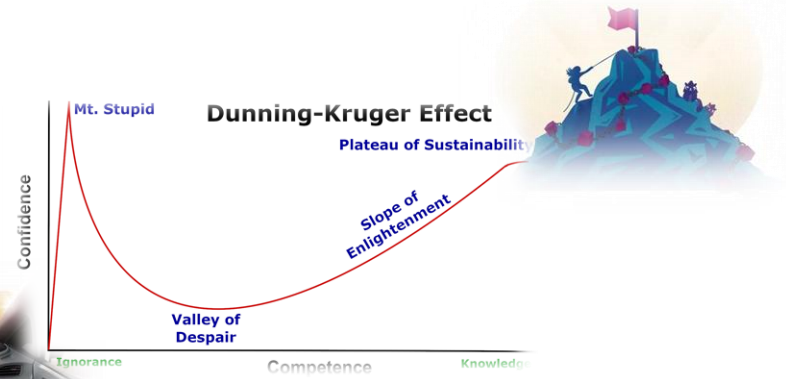
- Not designed for robotics
  - real-time performance: Real-time is not fastness, but guaranteed timing.
  - hardware synchronization: It is key to sensor fusion, but few SoCs have such design.
  - computation resource: Parallel computing and deep network acceleration is not the whole thing.
- Not enough tech support
  - To support sensors (sensor itself, driver, perf tuning) need SoC makers' support.
  - Run-time performance is equally important as innovative algorithms themselves.



- End-to-end is soooo hard: technology->product & product cycle

Hopes!

- Advanced integrated SoCs are showing up.
  - Nvidia Jetson AGX Xavier
  - Qualcomm Robotics RB3
  - Intel RealSense + Movidiu
  - NXP, Rockchip, Allwinner



- High-tech commoditization is accelerating.
  - Smartphone supply chain has been benefiting other smart hardware.
  - Consumer electronics drives the mainstreaming of technologies.
- Consumer Robotics and AIoT is happening.
  - SoC chips are essential to consumer robotics and AIoT. We are still in quite an early stage.