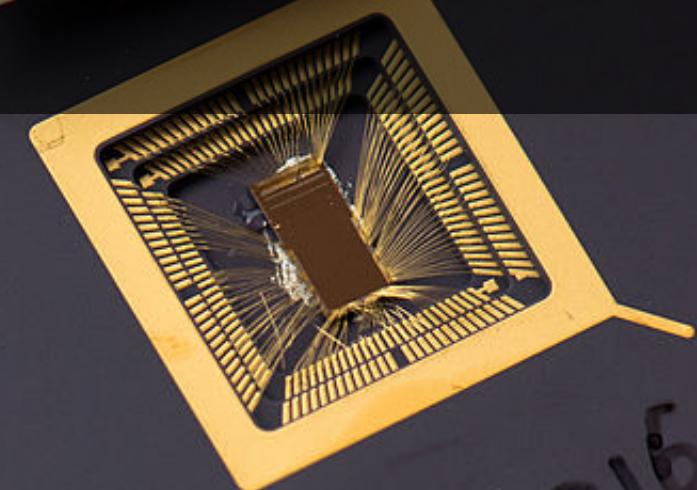


삼성 오픈소스 컨퍼런스 (SOSCON) 2020 오픈소스 하드웨어를 넘어서 오픈소스 프로세서로

한양대 박사과정 / 박상수 (Constant Park)

2020. 10. 14



About Me

■ 한양대학교 박사과정 박상수

- 연구분야: 신경망 가속기 (하드웨어) 설계, 신경망 추론 최적화 (CPU/GPU/NPU)



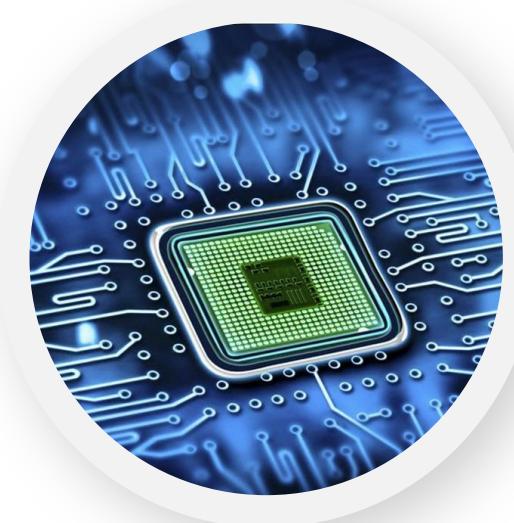
음성인식 AI 가속기 설계

MFCC, LM/AM Acceleration



가전제품용 AI 프레임워크

로봇청소기 및 시스템 에어컨



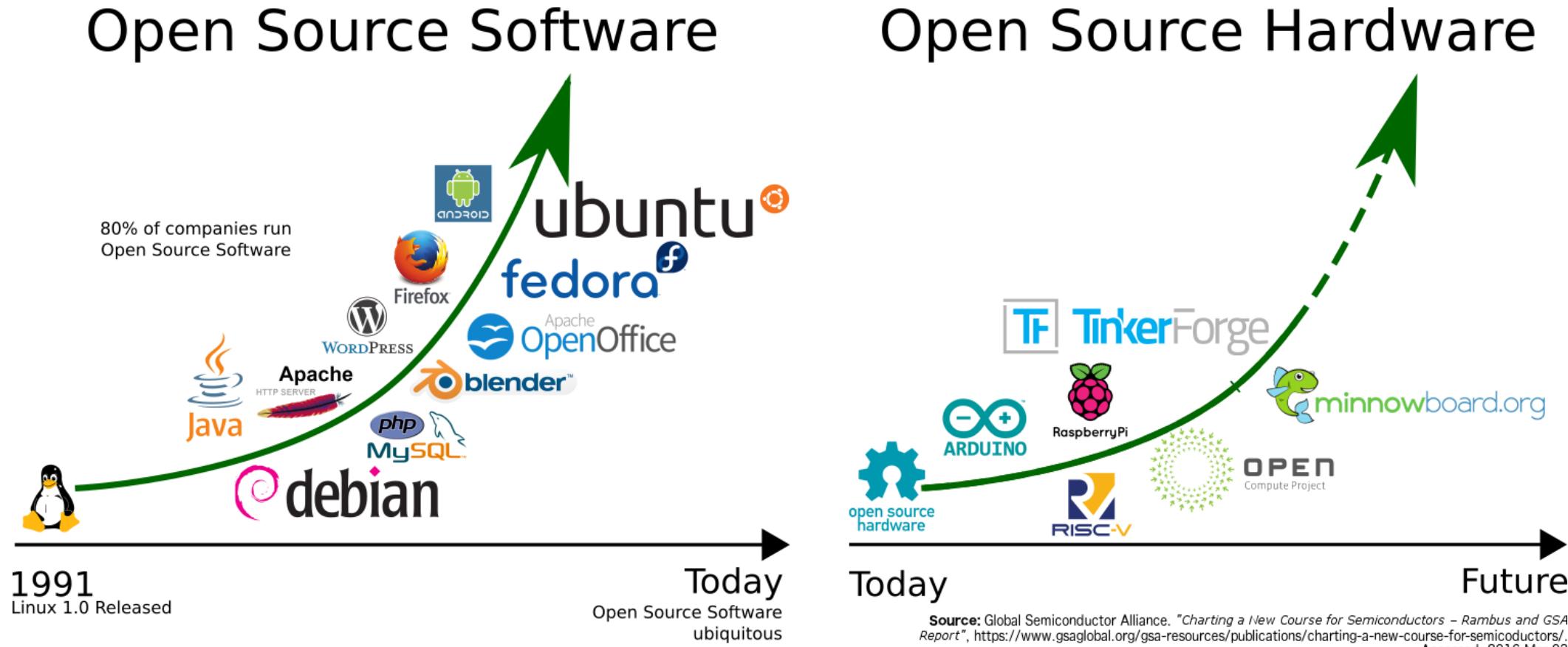
신경망 가속기 설계

CNN, RNN 가속을 위한 IP

지금은 오픈소스의 시대

■ 오픈소스

- 개발하는 과정에 필요한 소스 코드나 설계도를 누구나 접근해서 열람할 수 있도록 공개한 대상^[1]



오픈소스 하드웨어 #1

- 해당 제품과 똑같은 모양 및 기능을 가진 제품을 만드는데 필요한 것: Arduino, Raspberry Pi
- 하드웨어 기술 언어 (HDL)가 대중에게 공개된 하드웨어 IP: CPU/GPU 설계 코드

Not logged in 1

Article Talk Read Edit View history Search Wikipedia

Open-source hardware

From Wikipedia, the free encyclopedia

Open-source hardware (OSH) consists of physical artifacts of technology designed and offered by the open-design movement. Both free and open-source software (FOSS) and open-source hardware are created by this open-source culture movement and apply a like concept to a variety of components. It is sometimes, thus, referred to as FOSH (free and open-source hardware). The term usually means that information about the hardware is easily discerned so that others can make it – coupling it closely to the maker movement.^[1] Hardware design (i.e. mechanical drawings, schematics, bills of material, PCB layout data, HDL source code^[2] and integrated circuit layout data), in addition to the software that drives the hardware, are all released under free/libre terms. The original sharer gains feedback and potentially improvements on the design from the FOSH community. There is now significant evidence that such sharing can drive a high return on investment for the scientific community.^[3]

Since the rise of reconfigurable programmable logic devices, sharing of logic designs has been a form of open-source hardware. Instead of the schematics, hardware description language (HDL) code is shared. HDL descriptions are commonly used to set up system-on-a-chip systems either in field-programmable gate arrays (FPGA) or directly in application-specific integrated circuit (ASIC) designs. HDL modules, when distributed, are called semiconductor intellectual property cores, also known as IP cores.

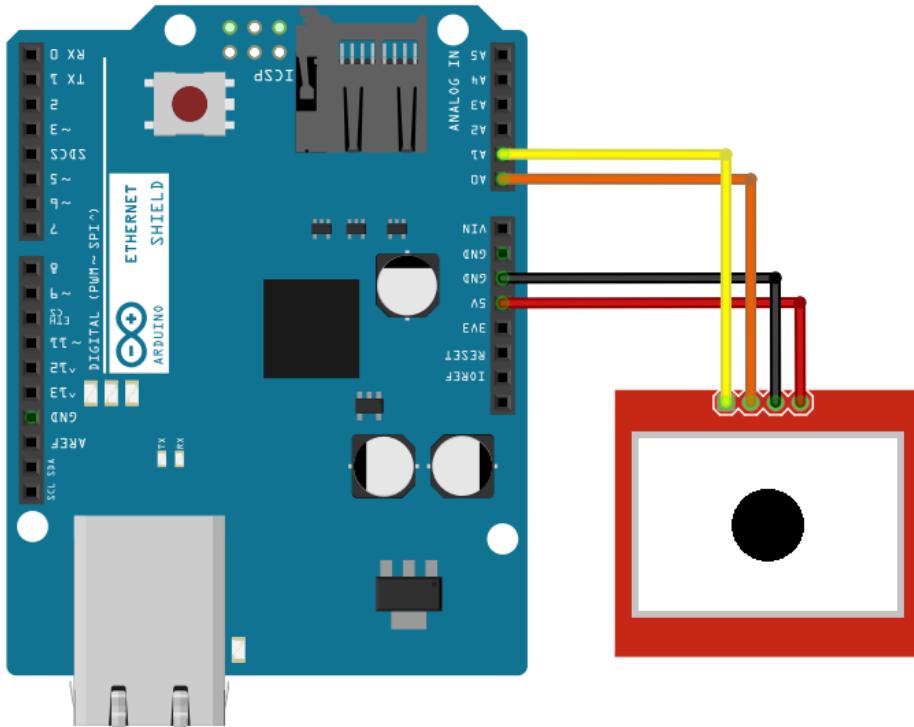
Open-source hardware also helps alleviate the issue of proprietary device drivers for the free and open-source software community, however, it is not a pre-requisite for it, and should not be confused with the concept of open documentation for proprietary hardware, which is already sufficient for writing FLOSS device drivers and complete operating systems.^{[4][5]} The difference between the two concepts is that OSH includes both the instructions on how to replicate the hardware itself as well as the information on communication protocols that the software (usually in the form of device drivers) must use in order to communicate with the hardware (often called register documentation, or open documentation for hardware^[4]), whereas open-source-friendly proprietary hardware would only include the latter without including the former.

Contents [hide]

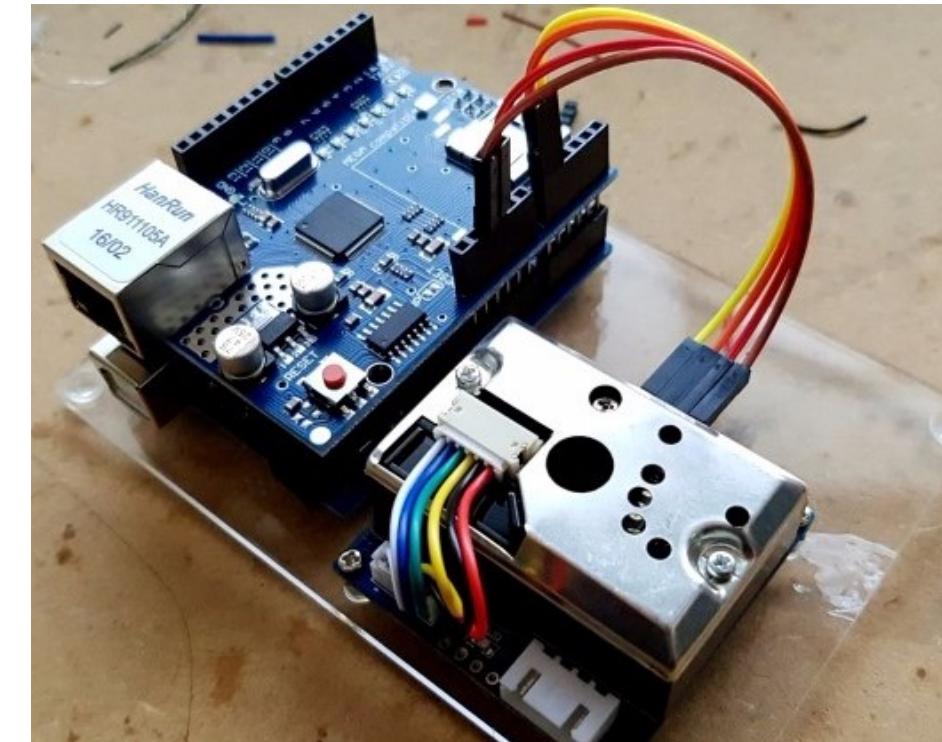
- 1 History
- 2 Forms of open-source hardware
 - 2.1 Computers

오픈소스 하드웨어 #2

- 해당 제품과 똑같은 모양 및 기능을 가진 제품을 만드는데 필요한 것: Arduino, Raspberry Pi
 - 회로도를 사용하여 Arduino와 센서/액추에이터 연결



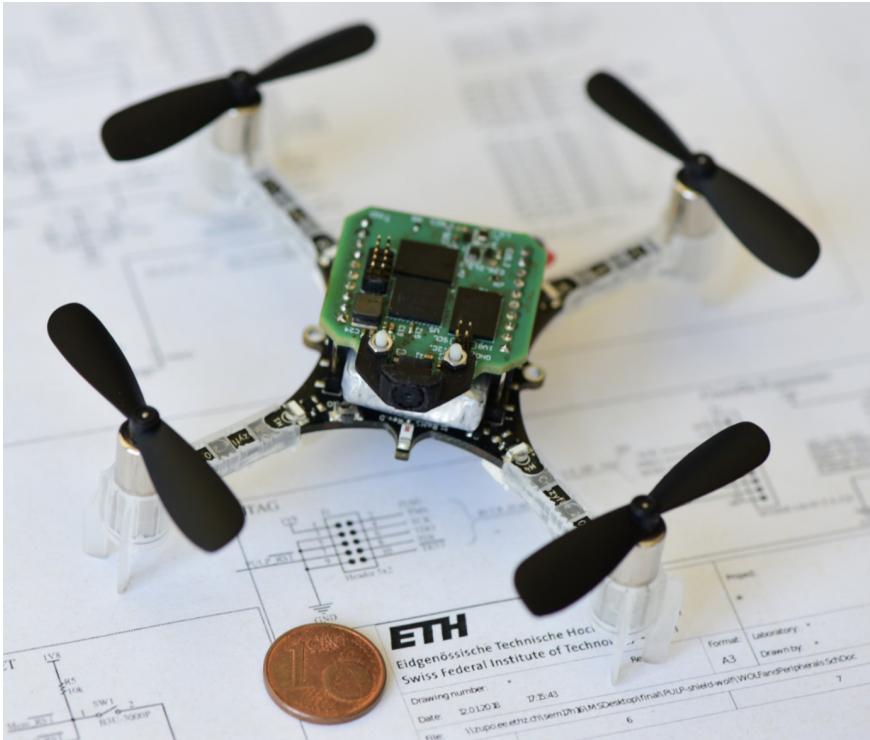
Arduino와 주변기기 연결 회로도



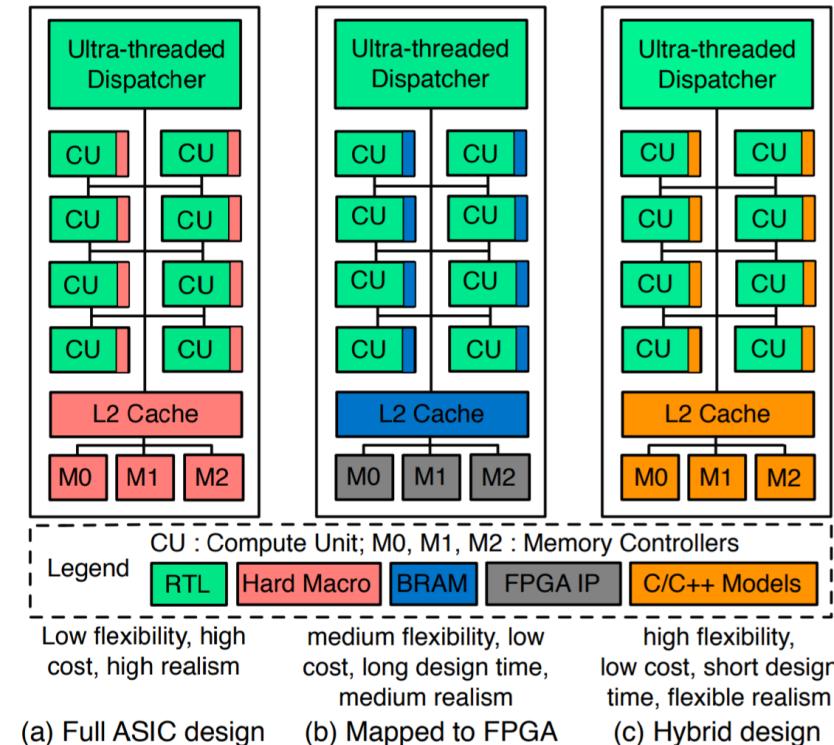
Arduino와 주변기기 연결

오픈소스 하드웨어 #3

- 하드웨어 기술 언어 (HDL)가 대중에게 공개된 하드웨어 IP: CPU/GPU/NPU 설계 코드
 - RISC-V (CPU)^[2], MIAOW (GPU)^[3], NVDLA (NPU)^[4]



RISC-V를 사용하는 Drone



오픈소스 GPU: MIAOW

오늘 하려는 이야기는 ?

- 오픈소스 프로세서 -

목차

- 프로세서 소개
- 오픈소스 프로세서 (RISC-V, NVDLA)
- 오픈소스 프로세서의 미래

프로세서 정의

■ 컴퓨터 내에서 프로그램 (연산)을 수행하는 하드웨어 장치

- CPU/GPU, DSP, NPU (Neural Processing Unit)



WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute
Help
Community portal
Recent changes
Upload file

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Cite this page
Wikidata item

Article

Talk

Processor

From Wikipedia, the free encyclopedia

Processor may refer to:

Computing [edit]

Hardware [edit]

- Processor (computing)
 - Central processing unit (CPU), the hardware within a computer that executes a program
 - Microprocessor, a central processing unit contained on a single integrated circuit (IC)
 - Application-specific instruction set processor (ASIP), a component used in system-on-a-chip design
 - Graphics processing unit (GPU), a processor designed for doing dedicated graphics-rendering computations
 - Physics processing unit (PPU), a dedicated microprocessor designed to handle the calculations of physics
 - Digital signal processor (DSP), a specialized microprocessor designed specifically for digital signal processing
 - Image processor, a specialized DSP used for image processing in digital cameras, mobile phones or other devices
 - Coprocessor
 - Floating-point unit
 - Network processor, a microprocessor specifically targeted at the networking application domain
 - Multi-core processor, single component with two or more independent CPUs (called "cores") on the same chip carrier or on the same die
 - Front-end processor, a helper processor for communication between a host computer and other devices

프로세서 종류

■ 컴퓨터 내에서 프로그램 (연산)을 수행하는 하드웨어 장치

- CPU (x86, ARM, RISC-V, SPARK, PowerPC), GPU (NVIDIA, AMD)
- NPU (ASIC/FPGA based accelerator): TPU, Tensor Core



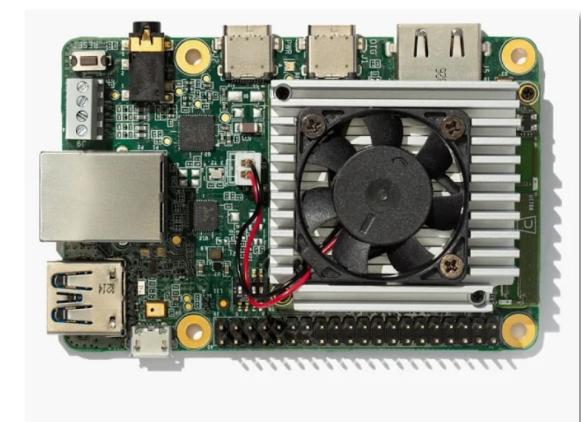
CPU: Ryzen Threadripper



GPU: VOTLA, TURING



NPU: ALVEO (FPGA)



NPU: Edge TPU (ASIC)

오픈소스 프로세서가 주목받는 이유 #1

■ “화웨이 사태”

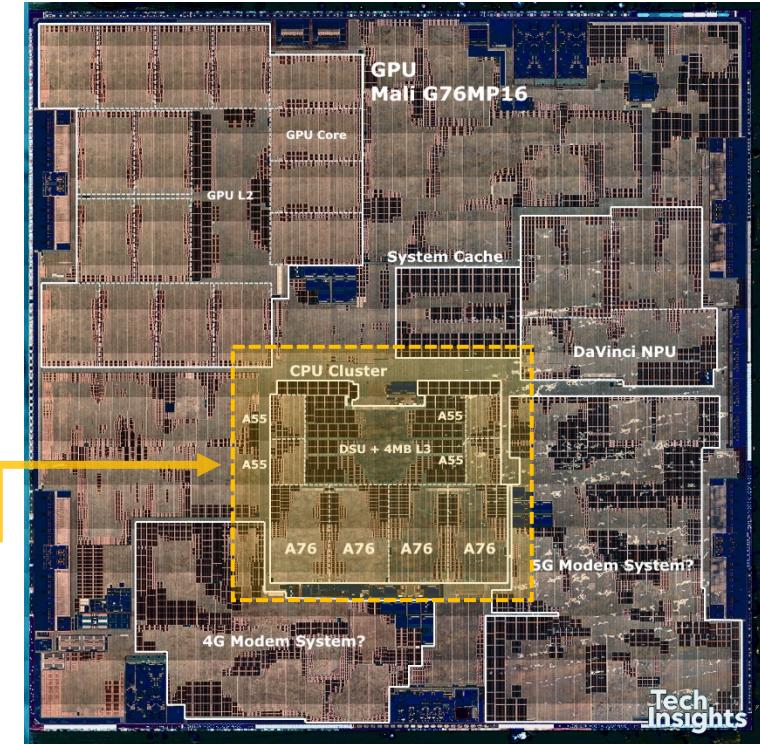
- 화웨이는 자사 스마트폰을 위한 SoC (System On a Chip) 설계 및 사용
- ARM 라이선스 사용불가, 칩 설계 자체 문제 발생, MediaTek 칩 사용

미국 상무부는 화웨이(Huawei)와 70개 계열사를 수출 제한 국가인 엔티티리스트(Entity List)에 추가한다는 내용과 반도체 기업이 화웨이에 제품을 납품할 때 미국 정부의 허가를 받도록 하는 내용을 포함하는 무역제재 방침을 발표했고, 이와 동시에 영국의 반도체 설계업체인 ARM은 화웨이와의 거래를 전면적으로 중단한다고 밝혔다.

이에 따라 화웨이는 ARM, 퀄컴 등 시장을 주도하는 반도체 기업으로부터 애플리케이션프로세서(AP)를 포함하는 핵심 반도체 부품을 조달 받기가 매우 어려워졌다.

이러한 ‘화웨이 사태’를 통해서 알 수 있듯이, 반도체 설계 시장에서의 독점적인 우위를 점하고 있는 ARM이 IP 공급을 중단하게 될 경우 IP를 공급받는 기업은 치명적인 피해를 입게 되는데, CPU 코어의 IP 기술력을 자체적으로 확보하지 못한다면 칩 설계 자체에 문제가 생기기 때문이다.

이번 미국과 ARM의 화웨이 제재 정책으로 인하여 삼성, 퀄컴 등 많은 기업들이 반도체 생산과 관련하여 직접적으로 위기 의식을 느끼게 되었고, 자체 CPU 코어 개발에 대한 수요가 커지게 되었다.



HiKey의 Kirin990 프로세서

오픈소스 프로세서가 주목받는 이유 #2

■ 라이선스 및 로열티를 지불할 필요가 없는 프로세서

- 칩 기능 업그레이드: 명령어 (ISA, Instruction Set Architecture) 수정 및 추가, 가속기 (GPU/NPU) 연동
- ARM사의 CPU/버스/GPU를 사용하기 위해서는 막대한 비용 필요

ARM IP와 비교해 RISC-V가 가진 최대 장점은 '오픈소스'다. 높은 사용료를 지불해야 하는 ARM의 IP와 달리, 인터넷이나 RISC-V 재단 웹사이트를 검색하면 금방 등장하는 소스를 활용하면 누구나 저렴한 비용으로 칩을 설계할 수 있다.

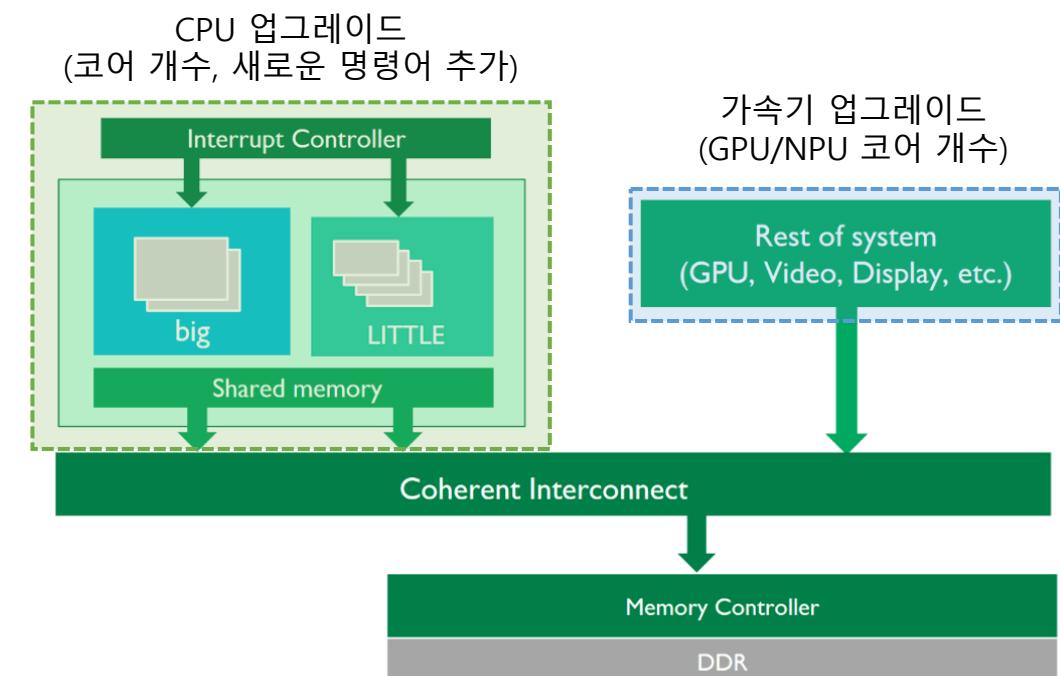
RISC-V의 또 다른 장점은 칩 기능 확장을 상대적으로 쉽게 할 수 있다는 것이다.

통상 ARM IP를 기반으로 칩을 설계할 경우, 칩 기능을 업그레이드하는 추가 명령어를 더 넣고 싶을 때 반드시 ARM의 아키텍처 라이선스를 구매해야만 한다.

그러나 RISC-V는 무료로 사용할 수 있는 기본 ISA에 사용자가 원하는 명령어를 마음껏 추가할 수 있다. 저렴한 비용으로도 칩 확대 가능성이 무궁무진하다.

게다가 저마다 계획한 '맞춤 칩'을 설계할 수 있어 고사양·저전력 칩을 개발하는 데도 상당히 용이하다. RISC-V 기반으로 개발된 칩이 ARM 기반 칩과 비교해 성능이 큰 차이를 보이지 않는다는 것도 매력이다.

전문가들은 RISC-V의 ISA 수준이 ARM 최상위 제품군인 ARM 코어텍스-A55 시리즈까지 도달했다고 평가한다.

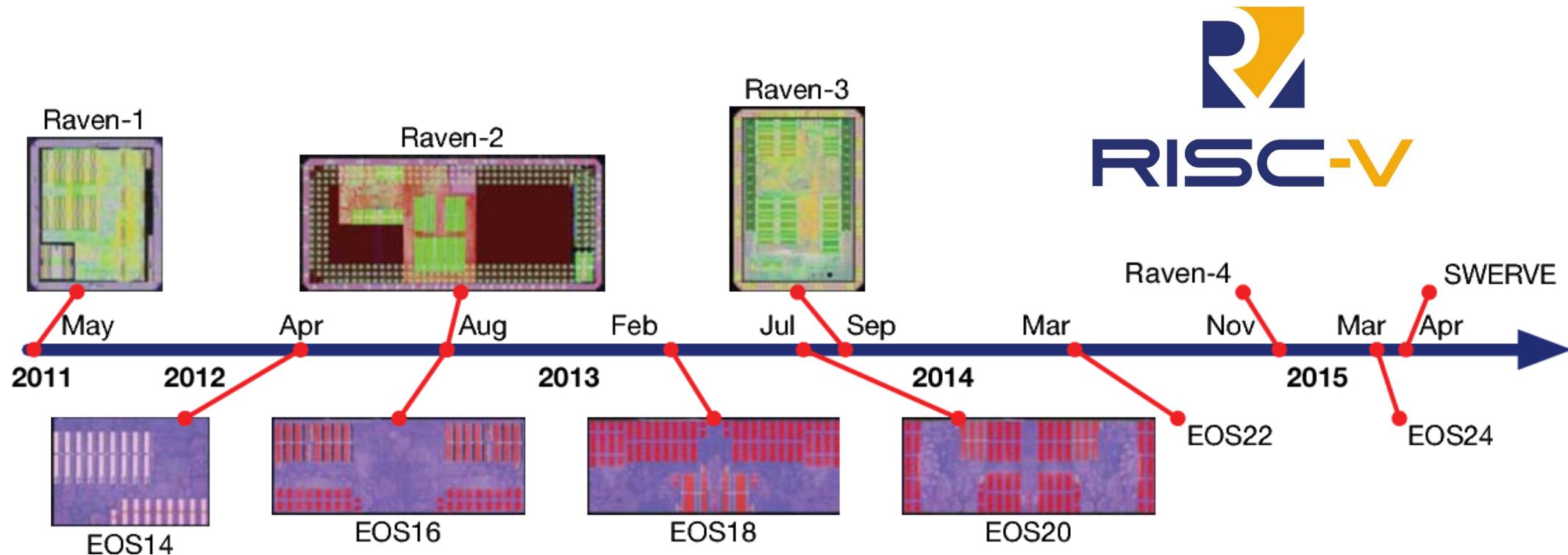


ARM기반의 SoC (예: Exynos, Kirin)

오픈소스 프로세서: RISC-V 소개

■ 오픈소스 프로세서의 대표적인 모델 RISC-V (risc-five)

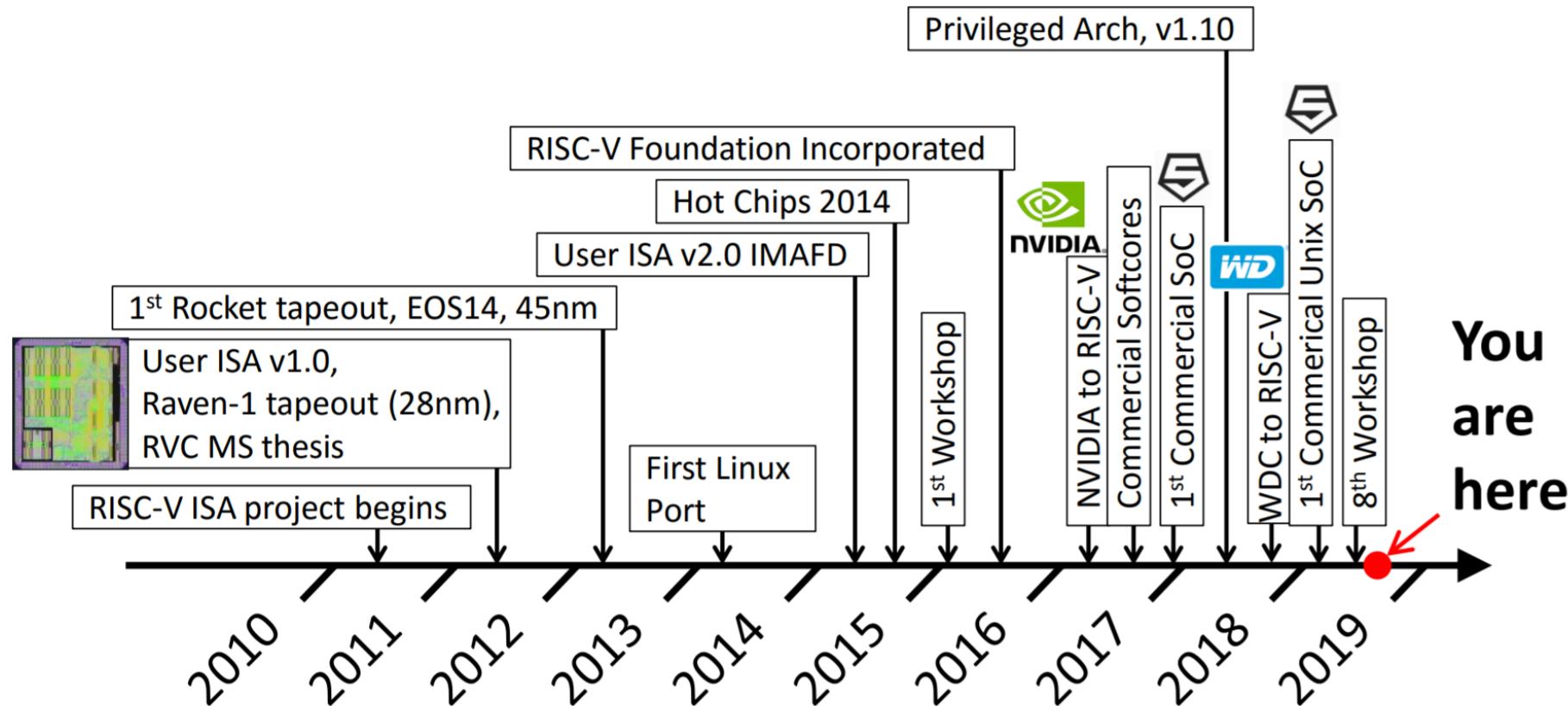
- 2010년부터 UC Berkeley의 Computer Science 학과에서 연구 시작
- 2013년 1월 프로토 타입 발표
- 구글, IBM, 화웨이, 삼성전자, SK하이닉스 등 100여개 기업이 RISC-V 재단에 멤버로 참여



오픈소스 프로세서: RISC-V 소개

■ 오픈소스 프로세서의 대표적인 모델 RISC-V (risc-five)

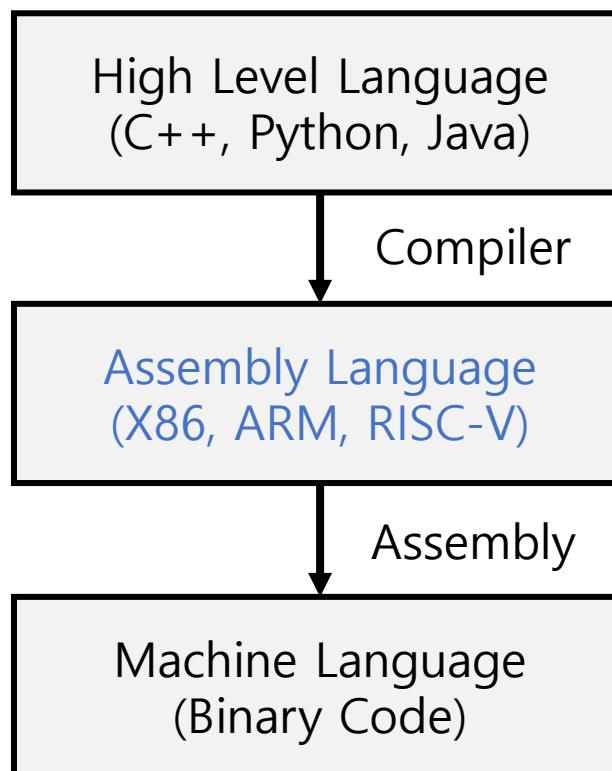
- 2010년부터 UC Berkeley의 Computer Science 학과에서 연구 시작
- 2013년 1월 프로토 타입 발표
- 구글, IBM, 화웨이, 삼성전자, SK하이닉스 등 100여개 기업이 RISC-V 재단에 멤버로 참여



오픈소스 프로세서: RISC-V 특징 #1

간결한 명령어 (Simple ISA)

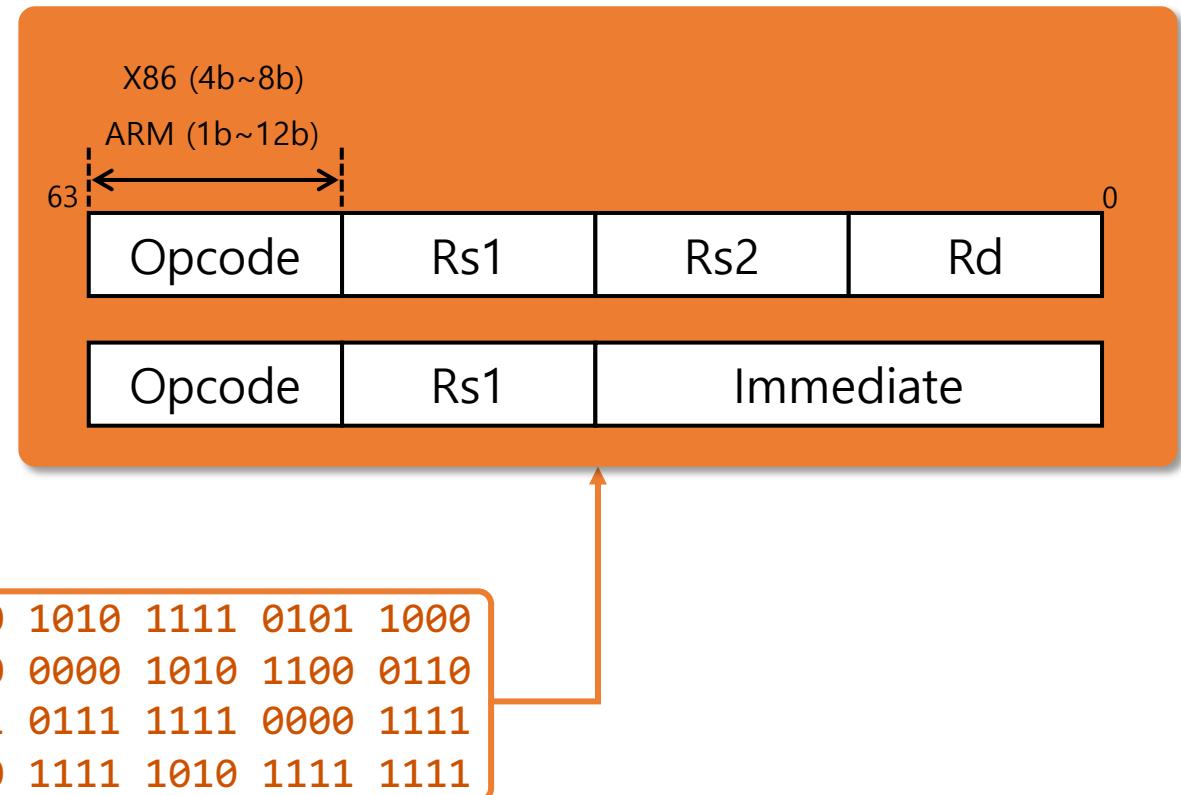
- RISC-V는 고정된 길이의 명령어를 사용 (하드웨어 구현 비용에 효과적)
- X86은 가변형 길이의 명령어 (수천개의 명령어)



temp=v[k];
v[k]=v[k+1];
v[K+1]=temp;

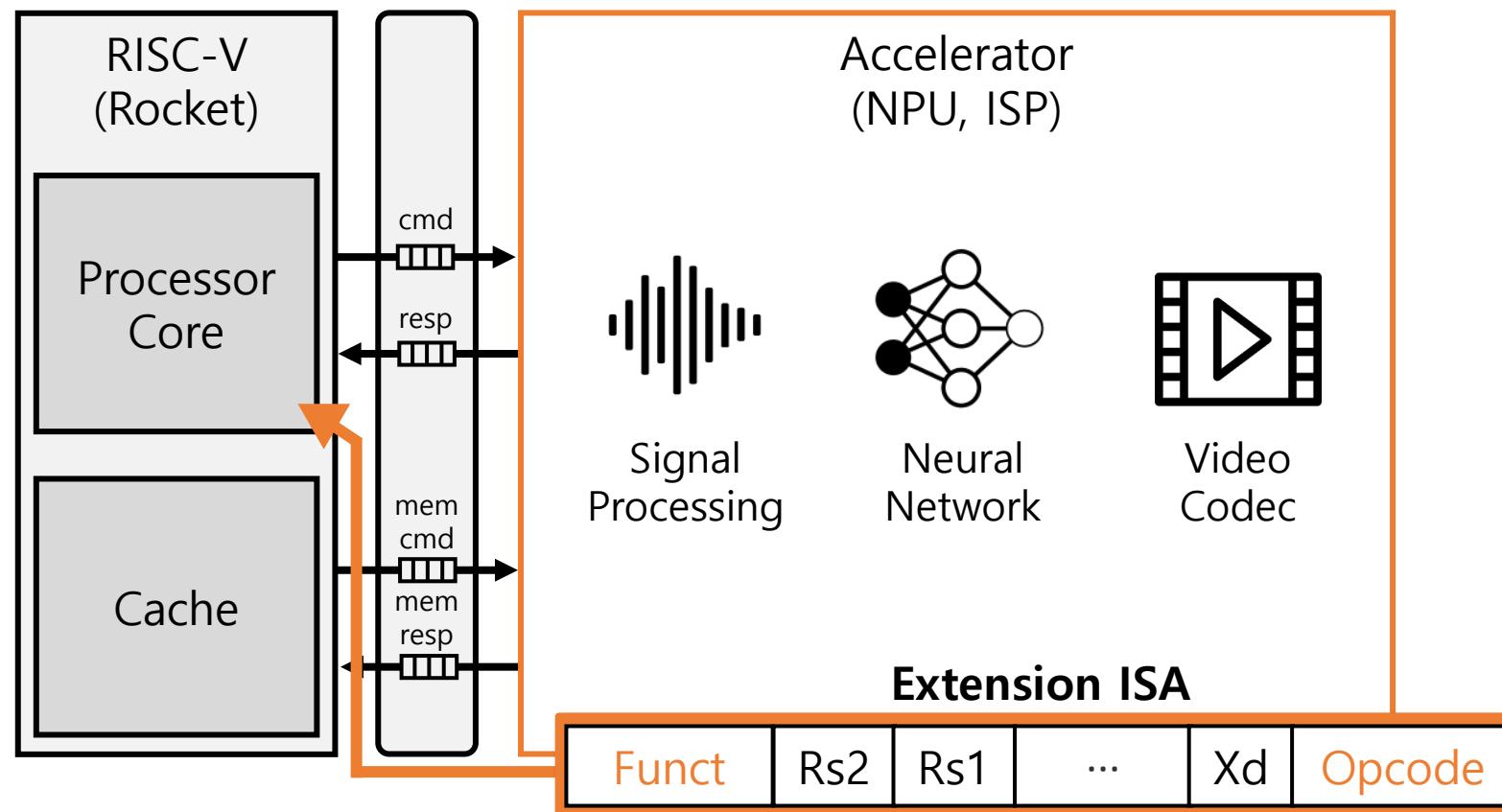
lw \$t0, 0(\$2)
lw \$t1, 4(\$2)
sw \$t1, 0(\$2)
sw \$t1, 4(\$2)

Instruction Set Architecture (ISA)



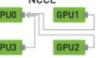
오픈소스 프로세서: RISC-V 특징 #2

- 확장/특화된 기능 추가를 위한 ISA (Modular ISA for extensibility/specialization)
 - 새로운 명령어 (ISA)를 추가, 새로운 기능 구현 가능
 - 예) 딥러닝 추론을 더 빠르게 가속하기 위한 행렬 연산 명령어



오픈소스 프로세서: NVDLA 특징 #1

- NVDLA: NVIDIA에서 DLA 표준 확립을 위해 공개한 오픈 아키텍처
 - 하드웨어 설계코드 (Verilog, SystemC), 소프트웨어 (디바이스 드라이버, API) 포함
 - NVDLA를 사용하여 신경망을 빠르고 저전력으로 가속 가능

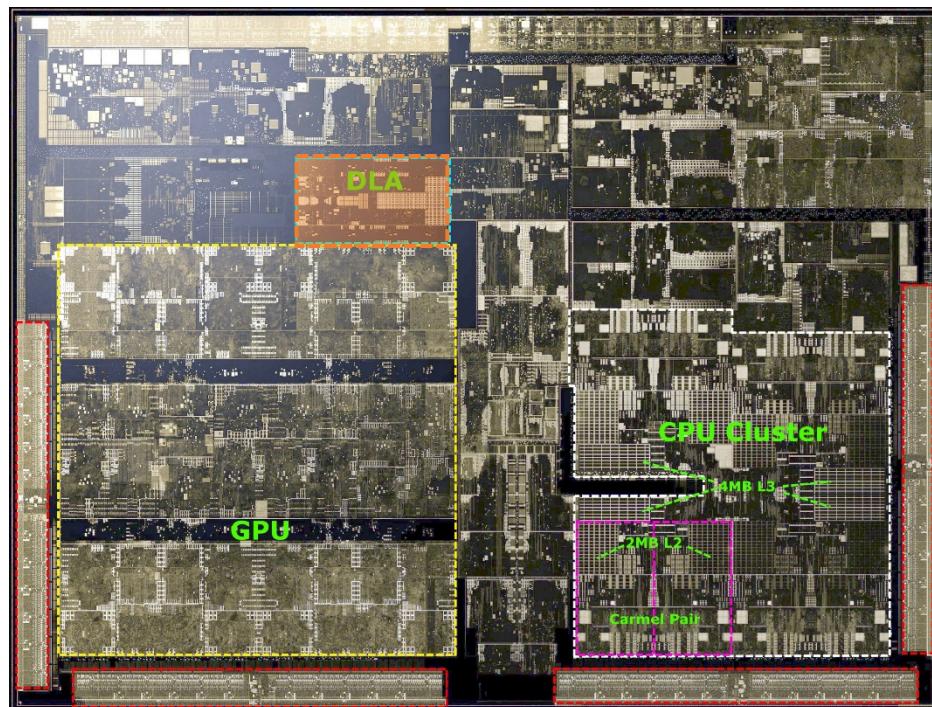
RAPIDS Open GPU data science Website > GitHub	TensorRT High-performance platform for deep learning inference Website > GitHub	 NCCL Optimized primitives for collective multi-gpu communication Website > GitHub	
TensorRT Inference Server Inference microservice for data center production that maximizes GPU utilization Website > GitHub	 DALI Data pre-processing in deep learning applications GitHub	 Apex A PyTorch Extension: Tools for easy mixed precision and distributed training in Pytorch Website > GitHub	
DIGITS Deep Learning GPU Training System Website > GitHub	 TensorFlow Examples that show how to use TF-TRT Website > GitHub	 ONNX TensorRT backend for ONNX Website > GitHub	NVDLA Open source Deep Learning Inference Accelerator Website > GitHub

NVIDIA 오픈소스 기여 (Contribution)

오픈소스 프로세서: NVDLA 특징 #2

■ 다양한 플랫폼에서 사용되는 NVDLA

- Jetson AGX (SoC): TensorRT를 사용하여 사용 가능
- FireSim (FPGA): RISC-V에 연동 가능한 에코시스템



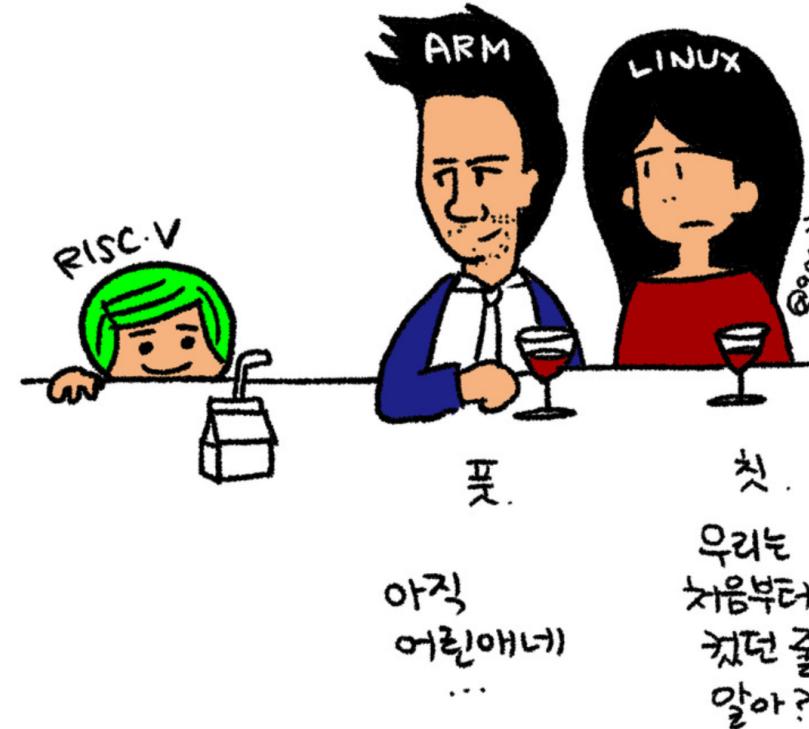
Jetson AGX Die Photo



FireSim (FPGA based RISC-V + NVDLA)

오픈소스 프로세서의 미래

- 제한적인 분야에서만 사용되고 있는 오픈소스 프로세서
 - 제한적인 분야: IoT 디바이스, SSD 컨트롤러, NPU, 5G 무선통신 칩
 - 재단에 더 많은 회사들이 참여하면서 에코시스템도 확대되고 있는 상태
 - 추후 스마트폰의 AP (Application Processor)에도 사용될 것으로 예상 !



RISC-V vs ARM/Linux^[5]

참조문헌 (Reference)

- [1] 오픈 소스, 나무위키 (<https://namu.wiki/w/오픈%20소스>), 2020.
- [2] A 45nm 1.3GHz 16.7 Double-Precision GFLOPS/W RISC-V Processor with Vector Accelerators, *ESSCIR*, 2014.
- [3] Enabling GPGPU low-level hardware explorations with MIAOW: An open-source RTL implementation of a GPGPU, *TACO*, 2014.
- [4] Integrating NVIDIA deep learning accelerator (NVDLA) with RISC-V soc on firesim, *EMC2*, 2019.
- [5] <https://hub.zum.com/goodhyun/38251>

감사합니다
