

CPU 기술과 미래 반도체 산업 (I)

CPU Technology and Future Semiconductor Industry (I)

박상기 (Sahnggi Park, sahggi@etri.re.kr) 소재부품원천연구본부 책임연구원

ABSTRACT

Knowledge of the technology, characteristics, and market trends of the latest CPUs used in smartphones, computers, and supercomputers and the research trends of leading US university experts gives an edge to policy-makers, business executives, large investors, etc. To this end, we describe three topics in detail at a level that can help educate the non-majors to the extent possible. Topic 1 comprises the design and manufacture of a CPU and the technology and trends of the smartphone SoC. Topic 2 comprises the technology and trends of the x86 CPU and supercomputer, and Topic 3 involves an optical network chip that has the potential to emerge as a major semiconductor chip. We also describe three techniques and experiments that can be used to implement the optical network chip.

KEYWORDS Photonic and optical interconnect, Optical network-on-chip, Optically interconnected CPU, Supercomputer architecture, CPU design and fabrication, Smartphone SoC, x86 CPU technology, Package on package, 2.5D package, 3D package

I. 서론

컴퓨터 CPU는 인간이 도달할 수 있는 기술의 최고 수준을 나타내는 척도라고 할 수 있다. 인간이 상상하는 어떠한 종류의 지능형 모델과 프로그램도 CPU 성능에 구애 받지 않을 수 없다. 4차 산업 뿐만 아니라 전체 IT 산업의 발전을 이끌어 가는 정점에 CPU가 있다. 2019년 출시된 스마트폰과

컴퓨터, 슈퍼컴퓨터에 사용된 최신 CPU의 기술과 특성 및 시장동향, 그리고 미국 주요기업과 대학교 전문가들의 연구방향, CPU 광인터커넥션 신기술을 3부에 걸쳐 기술하였다.

먼저 I 부에 CPU의 설계와 제작 과정에 대한 개략적인 설명과 2019년형 스마트폰 SoC의 기술 및 동향에 대해 기술하였다. CPU는 프로그램을 실행하는 소자이고 프로그램은 수많은 Instructions을 나

* DOI: <https://doi.org/10.22648/ETRI.2020.J.350208>

열하고 있다. CPU 설계는 Instruction set을 만들고 Instruction에 포함된 모든 함수를 트랜지스터로 변환하는 것에 해당한다. CPU 제작은 모든 트랜지스터와 통신라인을 포토마스크 도면에 그리고 도면에 따라 CMOS 공정을 수행하는 작업이다[1-3].

스마트폰의 CPU가 포함된 SoC(System on Chip)에는 6개 또는 8개 core로 구성된 CPU, GPU, AI engine, ISP, 4G 또는 5G modem, memory controller가 약 1cm² 크기의 칩에 집적되어 있다. 2019년 현재 ARM CPU의 IP license를 구매하여 스마트폰 CPU가 포함된 SoC 공급 기업 중 기술과 시장점유율이 가장 높은 기업은 Qualcomm, Samsung, Apple, Huawei의 4개사라고 할 수 있다. 통계 업체(Statista) 자료에 의하면 스마트폰 SoC에 대한 2019년 2분기 시장점유율은 Qualcomm 40%, Samsung 13%, Apple 20%, 기타 27%이다. 기타에는 HiSilicon과 MediaTek이 가장 큰 점유율을 차지하고 있다. 스마트폰 SoC 기술과 내부 구성, ARM사의 license 계약 유형에 대해 설명하고 2019년 최신 SoC 성능과 특성, 생산기업, 시장점유율, 스마트폰 내부 분해 사진 등을 나타내었다[4-7].

II. CPU 설계 및 제작 개요

CPU 설계를 위해서는 우선 컴퓨터가 인식하는 언어 구조를 알아야 한다. 그림 1과 같이 사람이 코딩할 때 사용하는 C++, Python, Java 등이 있고, compile 후 assembly 언어로 변환된다. Assembly어는 ARM, x86, MIPS, RISC-V 등과 같이 instruction set architecture(ISA)을 나타내는 언어이고, assembler에 의해 10101와 같은 기계어(machine language)로 변환된다. 프로그램을 실행하면 CPU는 기계어만 인식하지만 실제 CPU 설계에 있어서는 assembly 언어로 나타내어진 instruction set이 중요한 역할을

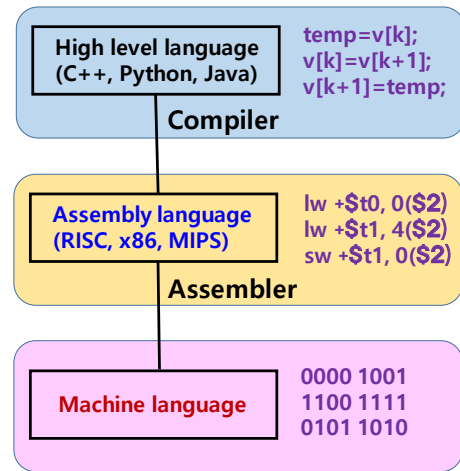


그림 1 컴퓨터 언어 레벨

한다.

Instruction set은 현재 두 종류가 주로 사용되고 있다. 하나는 ARM사에서 사용하는 RISC(Reduced Instruction Set Computer)이고 다른 하나는 Intel, AMD 등에서 사용하는 x86 instruction set이다. x86은 RISC의 반대 개념인 CISC(Complex Instruction Set Computer)에 속하지만 Intel사에서 많이 발전시켜 x86 instruction set으로 불리곤 한다. Reduced instruction set은 각 instruction이 고정된 길이(fixed length instructions, typically 4bytes)를 가지고 있어서 명령어 표현에 한계가 있고 총 instructions의 수가 많지 않다. ARM사의 instruction set은 기업의 기밀로 취급되어 자세히 알 수 없지만 유사 계열의 RISC-V은 무료로 공개되어 있으며 총 instructions의 수가 100여 개이다. x86 instruction set은 각 instruction이 변화 가능한 길이(variable length instructions, typically 1~16bytes)를 가지므로 명령어 표현이 자유롭고 총 instructions의 수가 전자에 비해 훨씬 많다. x86 instruction set 역시 기업 비밀이어서 공개되어 있지 않지만 약 수천 개인 것으로 알려져 있다.

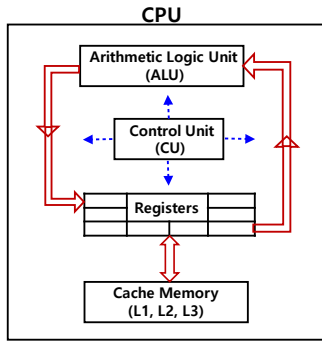


그림 2 Processor 내부 구성 개략도

Processor는 크게 logic part와 register part, cache part로 구성되어 있다. Logic part는 주어진 함수와 데이터를 이용하여 산술적 연산을 하는 부분이고 register와 cache part는 instructions와 데이터를 저장하는 부분이다. Register는 현재 연산을 수행하고 있는 instructions와 데이터를 저장하고, cache는 연산이 끝나거나 대기 중인 instructions와 데이터 중 긴급하게 재사용하게 될 우선순위에 따라 L1, L2, L3 cache에 저장한다. Register에 올라온 instruction에는 처리할 함수와 필요한 데이터가 지정되어 있고, 각 함수에 해당하는 logic 부분에서 계산을 수행하게 한 후 처리 결과를 다음 계산을 위해 register에 두거나 cache에 저장하며 우선순위가 아주 낮을 경우엔 메인 메모리(DRAM)로 보내어 저장한다.

Logic part는 AND, OR, NOT, NAND, NOR, 등과 같은 logic 함수의 transistor로 구성되어 있다. Transistor는 nMOS와 pMOS로 구성되고 gate에 가해진 전압에 따라 source와 drain 간 전류를 on/off 하는 switch 기능을 갖는다. 그림 3은 1bit NOT 함수를 실리콘 웨이퍼상에 형성하는 개념도이다. 각 instruction에는 함수들이 포함되어 있고 각 함수는 수많은 logic 함수를 포함하고 있으며, 복잡한 연결과 배치를 통해 64bit 계산을 수행하고 그 결과를 register로 보내준다. x86 instruction set의 경우 수

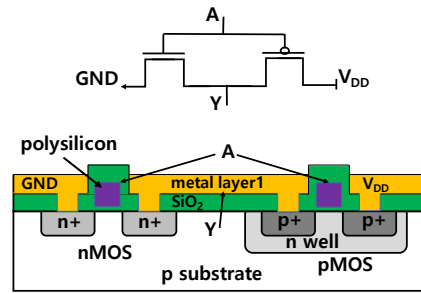


그림 3 NOT 함수 구조도

천 개의 instruction이 있으므로 logic part는 각각의 instruction에 해당하는 transistor 배치 영역이 있고 복잡한 연결 구조를 통해 register와 input/output 라인들이 연결되어 있다. Instruction set 전체를 배치하기 위해 CPU는 약 10~100억 개의 transistors를 포함하고 있다. Register와 logic part 및 cache part 사이에는 instruction 및 data 전송을 컨트롤하는 CU (Control Unit)가 있고 모든 전송 신호는 이를 거치게 되어 있다.

Register와 cache part는 instructions과 data를 저장하기 위해 SRAM 구조의 transistor로 구성되어 있다. 그림 4는 1bit 신호를 저장하는 SRAM의 기본 구조를 나타낸다. 두 개의 NOT 함수가 서로 연결되어 output은 input으로, input은 output으로 연결된 구조이다. WL(Word Line)의 신호가 on 상태에서 BL(Bit Line)으로 1bit 신호가 입력된 후 WL이 off 상태가 되면 신호는 오른쪽 NOT 함수의 input에 가서 신호가 반전되어 output으로 나오고 신호는 다시 왼쪽 NOT 함수의 input으로 가서 원래 상태로 반전 되어 오른쪽 NOT 함수의 input으로 가는 방식으로 반복적으로 순환되어 저장 상태를 유지한다. Register와 L1, L2, L3 cache는 CPU 설계에 따라 정해진 크기만큼의 byte 수를 갖고 해당 위치에 배치되며 CPU 내 64bit bus line과 연결된다.

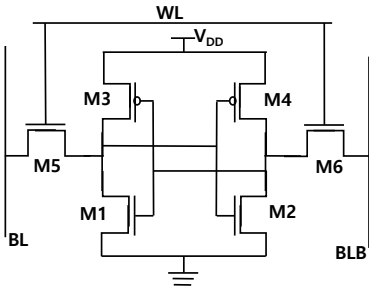
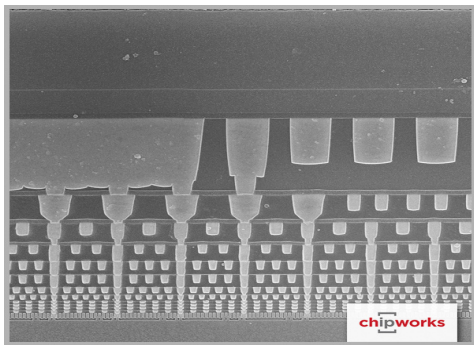


그림 4 1bit 저장 SRAM 회로



출처 <https://www.chipworks.co.uk>

그림 5 Broadwell CPU chip의 단면도

CPU 설계와 제작을 개략적으로 설명하면 다음과 같다. 기술이 진보하면 기존의 instruction set에 새로운 기능을 갖는 instruction을 추가하거나 변경하게 된다. 최종 instruction set이 결정되면 각 instruction의 함수를 계산하는 logic 함수를 배치하고 해당 transistor와 연결 line을 그리게 된다. CPU의 logic part는 전체 instructions이 다 수행될 수 있어야 하므로 x86의 경우 수천 개의 instructions에 해당하는 logics 함수들과 transistors들이 배치되고 연결되어야 한다. 정해진 specification에 따라 register, caches, control unit, bus line 등도 그려져야 한다. 이와 같은 작업은 각 기업이 보유하고 있는 설계 프로그램(high-level synthesis tools)을 이용하여 각 components들을 그리게 되며 많은 기술자들이 나누어 작업을 한다. 특히 제작에 들어간 뒤 bug

가 확인되면 수천억~수조 원의 손실이 발생하므로 simulation과 verification 작업에 가장 많은 인력과 시간이 소요된다고 한다. 이 작업이 완료되면 실제 칩 제작을 위해 mask layer를 그리는 작업을 하게 된다. 이때 TSMC와 같은 foundry 회사에 제작을 의뢰할 경우 foundry 회사에서 제공하는 기술 규격을 따라 mask layer를 그려야 한다. 예를 들면, foundry 회사의 7nm technology를 사용할 경우 해당 transistor의 gate width가 이 규격에 맞아야 한다. 기본적인 규격이 정해지면 mask layer 작업은 컴퓨터 프로그램을 이용하여 자동적으로 옮겨지게 된다. Mask layer 도면 작업이 완성되면 foundry 회사로 보내어 칩을 제작하게 된다.

그림 5는 2013년 Intel이 제작한 Broadwell CPU chip의 단면도이다. 14nm 공정을 사용하였고 fin-FET transistor 기술을 적용하였다. finFET transistor는 가장 아래에 있는 실리콘 층에 형성되어 있다. Gate width가 14nm이지만 면적당 transistor 수에 직접적인 영향을 주는 gate pitch는 70nm, fin pitch는 42nm, 메탈 라인의 최소 pitch는 52nm인 것으로 알려져 있다. Technology node에 따라 면적당 집적되는 transistor의 수는 점점 증가하지만 신호 전송선의 전기적 특성을 유지하기 위해 metal line의 width는 비례하여 작아질 수 없다. 따라서 3차원적으로 metal 층의 개수를 증가시켜 오고 있으며 그림 5는 최초로 13개 층이 적용된 단면도를 보여주고 있다. 2019년 현재 Intel은 최신 CPU 칩 제작을 위해 10nm 기술을 적용하고 있고 메탈 층은 역시 13개인 것으로 알려져 있다. AMD, Apple, Huawei, Qualcomm 등의 최신 CPU 칩을 7nm 공정으로 제작하고 있는 TSMC는 곧 5nm 공정을 도입할 예정이고 15개 메탈 층을 사용할 예정임을 발표한 바 있다. 삼성은 2019년 판매되는 Exynos칩 제작에 8nm 공정을 사용하였고 2019년 하반기에 7nm 공정을

도입한 것으로 알려져 있다. 최근 CPU 및 메모리 칩에 적용되는 finFET transistor 기술에는 한국 연구진의 기여도가 상당히 높은 것으로 알려져 있다.

사용자가 Window 또는 Android OS를 통해 프로그램을 실행하면 hard disk에 저장된 프로그램이 메인 메모리(DRAM)로 불러오게 된다. 프로그램은 많은 instructions 들이 순차적으로 나열된 구조이고 101010와 같은 기계어 신호가 전송되어 온다. 각 instruction은 순차적으로 CPU로 불러온 후 L1 cache에 저장되어 대기 상태가 된다. 순서가 되면 decoder로 전달되어 instruction의 내용이 해석된 후 register로 이동한다. 또한 instruction에 포함된 함수가 data를 필요로 할 경우 DRAM이나 hard disk에서 CPU register나 cache로 data가 이동된다. Instruction과 data가 준비가 되면 해당 logic part로 보내어 계산을 수행하게 한 후 결과를 register에 두거나 cache에 저장한다. 이와 같은 과정을 그림으로 나타내면 그림 6과 같다.

하나의 instruction을 실행하고 결과 data를 저장한 후 다음 instruction을 실행하는 것이 아니라 실행과정을 단계별로 세분화하여 instruction을 불러오는 단계(Fetch), decoding 단계, data를 logic part로 보내어 함수를 계산하는 단계(Execute), 계산 결과를 저장하는 단계(Write back)로 나누어 마치 컨베이어 벨트로 이동하며 조립되는 기계와 같이 실행된다. 다수의 instructions이 대기상태에서 첫 번째 instruction이 fetch 단계 후 decode 단계로 들어가면, 두 번째 instruction은 fetch 단계로 들어가고 다시 첫 번째 instruction이 execute 단계로 들어가면 순차적으로 다음 instruction이 하나씩 다음 단계에 진입하게 된다. 이와 같은 실행과정을 pipeline이라 한다. CPU 내 하나의 core가 두 개 이상의 pipe lines을 운영하는 기술을 superscalar technology라 하고 그러한 CPU를 superscalar processor라고 한다. Pipeline을

운영하는 구성요소를 thread라 하며 두 개 이상의 thread을 운영하기 위해서 registers와 logic part, 그리고 cache는 공유하고 decoder는 thread 수만큼 비치되어 개별적으로 운영하게 된다. 하나의 core가 동시에 두 개의 instructions을 실행할 경우 logic part를 공유하므로 함수가 다르거나 동일한 함수일 경우 하나는 실수 계산, 다른 하나는 정수 계산으로 분리되어야 가능하게 된다. 최근 판매되는 Intel과 AMD의 모든 CPU는 각 core당 2개의 threads를 가지고 있으며 이를 hyper thread라 부른다. Quad core CPU의 경우 물리적인 core는 4개지만 OS는 8개로 인식한다.

CPU 설계에 핵심 기술 중 하나가 병렬 컴퓨팅(parallel computing) 기술이다. 병렬 컴퓨팅에는 다양한 병렬 연산 방법이 있고 통상 bit level, instruction level, task level, data level 병렬컴퓨팅으로 구분하는 경우가 많다. 최근 스마트폰 및 PC에 사용되는 CPU는 모두 64bit를 지원하고 있다. CPU에서 64bit의 의미는 register와 logic 또는 cache 사이에 최소 64개의 신호선이 있고 하나의 clock cycle 동안 각 선마다 1bit씩 64bit 신호가 동시에 전달되어 계산되거나 저장되는 것을 의미한다. 즉, bit level 병렬 컴퓨팅은 64bit의 data가 동시에 병렬적으로 처리되는 것을 나타낸다. 32bit 컴퓨터의 경우 64bit

clock cycle	0	1	2	3	4	5	6	7	8
waiting instructions (L1 cache)	Yellow	Green	Blue	Red					
	Green	Blue	Red	Yellow					
pipeline	1. Fetch			Red	Blue	Green	Yellow		
	2. Decode				Red	Blue	Green	Yellow	
	3. Execute					Red	Blue	Green	Yellow
	4. Write						Red	Blue	Green
completed instructions							Red	Blue	Green

그림 6 pipeline 개념도

data에 대해 2회로 나누어 계산하고 저장하는 작업을 하게 되므로 64bit 컴퓨터가 bit level 병렬 컴퓨팅을 통해 32bit 컴퓨터에 비해 2배 빠른 처리속도를 갖게 된다. 프로그램은 많은 instructions들이 순차적으로 나열되어 있고 복수의 instructions이 여러 개의 threads를 갖고 있는 하나의 core에서 또는 다수의 core에서 동시에 병렬적으로 실행하게 되며, 이를 instruction level 병렬 컴퓨팅이라 한다. 규모가 큰 프로그램은 다수의 task로 구성되어 있고 각 task는 다수의 instructions을 가지고 있다. CPU는 각 task를 여러 개의 core에 분산시켜 병렬적으로 진행하며 이를 task level 병렬 컴퓨팅이라 한다. 특수한 프로그램의 경우 하나의 instruction으로 많은 data를 처리하는 경우가 있다. 벡터 $A=[1, 2, \dots, 10000]$ 와 $B=[10001, 10002, \dots, 20000]$ 의 합($A+B=C$)을 구하는 프로그램이 있다고 하자. 1만 개의 성분을 1천 개씩 나누어 10개의 core에 분산하여 계산한 후 최종 합을 구한다고 하자. 10개의 core는 동일한 1개의 instruction을 수행하고 모든 data는 배열을 갖고 규칙적으로 입력된다(SIMD: Single Instruction Multiple Data). 이러한 연산을 data level 병렬 컴퓨팅이라 하고, instruction을 vector instruction, processor를 vector processor라 한다[1,2].

ARM CPU 또는 x86 CPU는 일반적인 함수 계산을 위한 processor이고 수백 개 또는 수천 개의 instructions에 대응하는 logic 함수들이 logic part에 배열되어 있다. 만약 data level 병렬 컴퓨팅과 같이 소수의 instructions으로 대단위 data를 처리하는 작업을 일반적인 목적의 CPU로 실행할 경우 나머지 다른 instructions에 대응하는 logic 함수들은 하는 일이 없게 되고 CPU 성능을 제대로 활용하지 못하게 된다. 이를 해결하기 위해 다양한 가속 프로세서(accelerator processor)들이 개발되고 있다. GPU(graphics processing unit), AI processor 또

는 NPU(neural processing unit), IPU(image processing unit) 또는 ISP(image signal processor) 또는 DSP(digital signal processor) 등이 대표적이다. 이 중 서버용 GPU를 예로 들면 GPU 내에 수백 개의 core가 있고 각 core는 독립적인 registers와 소수의 instructions을 처리할 수 있는 독립적인 logic part로 구성되어 있다. 대단위 data를 적절히 분산하여 수백 개의 core들은 동시에 병렬적으로 instructions을 처리한다. 이러한 가속 프로세서는 설계와 제작과정이 일반적인 목적의 CPU와 동일하기 때문에 스마트폰과 PC용 CPU는 SoC(System on Chip) 형태로 CPU, GPU, IPU(NPU)가 하나의 Die로 제작되고 동일한 칩에 집적되어 있다.

CPU의 clock 주파수는 1.5~4.0GHz의 범위에서 고정된 반면 CPU의 instruction 처리 속도는 지속적으로 증가되고 있고, 이는 병렬 컴퓨팅 기술이 발전하기 때문이다. 즉, 다수의 instructions이 동시에 병렬적으로 실행되므로 clock cycle당 instructions(IPC: Instructions Per Cycle)가 지속적으로 증가한다. CPU의 속도를 나타내는 지표로 IPC 또는 CPI(Cycles Per Instructions)를 사용하기도 하지만 x86 CPU의 경우 instruction의 길이가 서로 다르므로 최근 ARM CPU와 x86 CPU도 슈퍼컴퓨터에 사용하는 FLOPS(floating point operations per second)를 사용하여 속도를 나타낸다. Floating point는 실수 값을 $+1.12345 \times 10^8$ 과 같은 형식으로 나타내고 소수점이 자유롭게 옮겨 다닐 수 있는 데서 붙여진 이름이다. 이 숫자를 2진수 형식으로 표현하면 $+10101010 \times 2^{10101}$ (전자와 동일한 숫자 값이 아님)와 같은 형식이 된다. 64bit 데이터를 double precision(소수점 아래 정확도를 나타냄)이라 하고 +/- 자리에 1bit, 10101010자리에 52bit, 지수 10101자리에 11bit을 할당하여 숫자를 나타낸다. 32bit 데이터는 single precision이라 하며 각 자리에

1bit, 23bit, 8bit를 할당한다. FLOPS 계산에 사용하는 operation을 FMA(fused multiply-add)라 하며 $(A \times B) + C$ 와 같이 두 개의 숫자를 곱하고, 그 결과를 다른 하나의 숫자와 더한 후 마지막으로 반올림하여 소수점 끝자리 숫자를 확정한다. 곱과 합으로 된 2개의 operations이 포함되어 있다. 하나의 clock cycle 동안 cache에서 64bit 데이터 A와 B를 logic 함수로 보내어 곱을 한 후 register에 저장된 C값과 전자의 결과를 더하기 함수로 보내어 합을 구한다. 하나의 core가 이와 같은 operation을 완료하면 2 FLOPS이 되고 clock 주파수가 3GHz이면 core의 속도는 6Giga FLOPS가 된다. CPU에 4개의 core가 있으면 CPU는 대략 24Giga FLOPS의 속도 지표를 갖게 된다. 또한 동일한 CPU에 대해 single precision의 data 계산 기준 48Giga FLOPS의 속도 지표를 갖는다. 실제로 FLOPS 지표가 core의 수와 정비례하지 않고 CPU 설계에 의존하며 FLOPS를 측정하는 benchmark program을 실행하여 구한다.

CPU의 처리속도를 높이기 위해 CPU 내에 많은 하드웨어 요소들이 매 시간 늘지 않고 최대한 작업 수행에 참여하도록 하는 기술들이 발전하고 있다. 프로그램에 짜여진 instructions의 순서가 A-B-C라고 할 때 B는 A가 실행된 후 그 결과를 이용하는 데 반해 C는 A의 결과와 상관없이 cache data로 실행 가능할 경우 CPU는 프로그램 순서와 달리 A-C-B 순서로 진행한다. 즉, 프로그램 전체에 대해 최적의 순서를 찾아서 실행하게 되며 이러한 기능을 가진 CPU는 성능 표시에 out-of-order라고 표시한다. 프로그램상에 if 함수와 같이 어떤 조건에 따라 1번 task 또는 2번 task 등을 선택하여 진행하는 경우가 있다. 이와 같은 함수의 instruction을 branch instruction이라 하며 프로그램의 약 20%를 차지한다. 처리속도를 높이기 위해 CPU는 함수의 선택 결과가 나오기 전에 미리 예측을 하고 예측된

task를 수행할 thread 또는 core를 지정하여 작업을 수행하기 시작한다. 예측이 맞으면 그대로 진행하고 틀리면 수행한 작업을 폐기하고 선택된 task를 다시 시작한다. 이와 같은 기능을 branch prediction이라 하고 CPU에는 이를 수행하기 위한 하드웨어가 설계되어야 하고 이를 branch predictor라 한다. 이와 같은 예측은 프로그램을 통해 학습하는 방법으로 적중률을 높이게 되어 있고 약 90% 이상의 적중률을 갖는다고 한다. 프로그램이 실행되면 사정(speculation)을 하고 실행에 필요한 모든 요소들을 최적의 조건으로 활용하기 위해 scheduling 작업을 하게 되며 이에 필요한 scheduler가 CPU상에 설계되어야 한다.

병렬 컴퓨팅, out-of-order, branch prediction, scheduling 등을 가능하게 하는 가장 필요한 기술 중 하나가 cache coherence이다. Coherence란 각자의 작업이 서로 연결되어 움직이는 구성요소 간 불협화음이나 불일치가 전혀 없는 완벽한 조화 상태를 말한다. 4개의 core를 가진 CPU를 예로 들자. 각 core는 L1, L2 cache를 사적으로 소유하고 L3 cache는 공동으로 소유한다. Core 1에 소속된 L1, L2 cache 중 L1은 core 1만 쓰고 읽기 가능하고, L2는 core 1에게 우선권이 있지만 다른 core 2~core 4도 읽기가 가능하고 쓰기도 가능한 경우가 있다. L3는 특별한 우선권이 없이 같은 자격으로 읽고 쓰기가 가능하다. 표 1은 저장 위치별로 읽고 쓰는데 걸리는 시간(latency)과 저장 용량(capacity)을 나타낸다. L3의 특정한 주소에 A라는 data가 저장되어 있는 상태에서 core 1, core 2, core 3가 A 값을 읽어온 후 각 L1 cache에 저장하였다고 하자. 이후 core 1이 자신의 L1 cache data를 A에서 B로 변경하면 변경된 사실을 모든 core에게 통보하고 data A는 더 이상 유효하지 않다는 사실을 공유하여야 한다. L3 해당 주소의 데이터도 A에서 B로 변경하여야 한

표 1 데이터 저장 위치별 Latency와 용량

Component	Example Latency	Example Capacity
L1 Cache	1ns	64KB per core
L2 Cache	3ns	256KB per core
L3 Cache	10ns	20,480KB
Memory	100ns	64GB
SSD Storage	30,000ns	100GB
HDD Storage	15,000,000ns	1TB

다. 만약 core 간 cache data가 정확히 조율되지 않으면 일부 core는 data B가 들어가는 수식에 A가 들어가거나 또는 그 반대가 되어 프로그램 전체는 오류에 빠지게 된다. Cache coherence는 모든 core 또는 threads가 수 없이 많이 쓰고 읽기를 하여도 cache data가 최대한 빠르고 정확히 조율되어 오류가 발생하지 않는 상태를 나타낸다. Cache coherence에는 core 간 통신속도가 절대적으로 중요한 역할을 하게 된다. 코어의 수가 증가함에 따라 전기 신호를 이용한 CPU 내 통신은 곧 한계에 이를 것으로 예측하고 광 신호를 도입하기 위한 연구가 2000년대 중반부터 silicon photonics라는 명칭으로 세계 최고의 기업과 연구기관이 참여하는 가운데 진행되었으나 아직 성공을 거두지 못하고 있다.

Cache coherence를 위해 다양한 protocol이 개발되어 사용되고 있으며 대표적으로 Intel에서 사용하고 있는 MESI(modified, exclusive, shared, invalid) protocol을 들 수 있다(표 2 참조). Core 1이 L3에서 data A를 읽거나 쓰면 Core 1은 exclusive 상태가 되어 L3의 해당 주소는 단독으로 만 사용하게 되고 나머지 core에 있는 기존 data는 모두 invalid 상태가 되어 사용할 수 없게 된다. Core 2와 core 3가 data A을 읽기 위한 신호(snoopy 신호라 함)를 snoopy bus에 올리면 core 1의 cache data를 data bus를 통해 전달 받게 된다. 이때 모든 core는 shared 상태가 된다. 그 후 core 3가 data A를 B로 변경하면 core 3는

표 2 MESI protocol

	Core1	Core2	Core3	Core4
Read/Write	E	I	I	I
	A	-	-	-
Read	S	S	S	S
	A	A	A	-
Modified	I	I	M	I
	A	A	B	-
Write	I	I	E	I
	A	A	B	-
Read	S	S	S	S
	-	-	B	B

modified 상태가 되고 다른 core들은 invalid 상태가 되어 기존 data A를 사용할 수 없게 된다. Core 3가 L3의 해당 주소에 data B를 쓰면 core 3는 exclusive가 되어 L3의 해당 주소는 core 3만 사용할 수 있게 된다. 그리고 나머지 core는 Invalid 상태가 된다. Core 4가 read 신호를 snoopy bus에 올리면 core 3의 data B을 data bus를 통해 전달 받고 이때 모든 core는 shared 상태가 된다. Data bus에 올라온 data는 모든 core에게 전달되므로 Core 1과 core 2도 필요하면 data B를 각 cache에 저장할 수 있다. Cache coherence가 완전하게 이루어지면 다수의 core를 가진 CPU는 마치 하나의 큰 core 만 가진 것처럼 프로그램을 실행하게 된다[1-3].

2018년 판 통계(2018 edition of IC Insights' Mc-Clean Report)에 의하면 ARM CPU와 x86 CPU의 시장규모 및 점유율은 그림 7과 같다. Tablet, Cell-phone App, Embedded 부분에 x86 CPU가 일부 사용되고 있지만 대부분 ARM CPU가 사용되고 있으므로 3개 부분 합을 ARM CPU 점유율로 추정할 수 있다. ARM CPU의 점유율이 50%에 근접해 가고 있음을 알 수 있다. 같은 기관의 2012년도 통계에 의하면 순수 x86 CPU 점유율이 64%에서 2018년도에 51%로 축소되었다. Intel이 Atom CPU 생산

을 중단하였고 Apple은 2020년 이후 자사의 Mac-Book 컴퓨터에 Intel CPU 대신 ARM CPU를 사용하겠다고 발표한 사실을 참고하면 ARM CPU의 시장점유율 확대는 더욱 가속화될 것으로 예상된다. 총 시장 규모는 약 80조 원이다.

III. 스마트폰 SoC 기술과 동향

ARM사의 홈페이지에 공개된 최신 processor 모델은 Cortex-A77, Cortex-A76AE, Cortex-A76, 등이 있다. 이중 Cortex-A77은 2019년에 공개되어 아직 채택된 스마트폰이 없는 상태이고 Cortex-A76는 2018년에 공개되어 2019년 최고 사양 스마트폰에 채택되었다. Cortex-A76AE는 Cortex-A76에서 자동차 (Automotive Enhanced) 분야 사용 기능을 향상하였다고 한다. 홈페이지에 공개된 Cortex-A76의 특성은 표 3과 같다. Instruction set은 ARMv8.5-A version까지 나와 있고 이중 ARMv8.2-A 을 제외한 후속 판은 특수한 기능을 추가한 것이다. Cortex-A76 processor는 표 3에 나열된 instruction set 과 extension set을 포함하도록 설계되어 있다. 제조 공정과 technology node에 따라 최대 3.0~3.3GHz 의 clock 주파수를 가질 수 있다. 각 core에는 일반적인 registers와 logics 부분 외에 vector 연산을 위한

표 3 Cortex-A76의 특성

ISA version	ARMv8.2-A
Max, CPU clock	3.0GHz/3.3GHz
Microarch itecture	Out of order pipeline, Superscalar, Neon/Floating Point Unit, Cryptography Unit
Registers	1. 64bit 프로그램용 registers 64bit registers 31개, 128bit registers 32개 2. 32bit 프로그램용 registers 32bit 13개, 64bit 32개
Memory system	L1 cache-instruction 64KB and data 64KB L2 cache-128~512KB L3 cache-optional, 512KB~4MB

Neon/Floating point unit과 보안 기능을 강화하는 cryptography unit이 추가되어 있다. Big-little로 구성되는 core의 배치는 big 4개, little 4개까지 가능하다. 그 외 registers와 cache memory는 표 3에 나타난 바와 같다[4-6].

ARM processor는 x86과 다른 특징이 있다. Core의 수가 8개인 octa-core CPU의 경우 연산 속도가 빠르고 전력소모가 큰 최신 사양의 big core 4개와 연산 속도는 상대적으로 느리지만 전력 효율이 높은 little core 4개를 배치하되 각각 1:1로 짝을 지어놓는다. CPU의 높은 성능이 요구될 때는 big core들이 작업에 동원되고 낮은 성능으로 작업이 가능하고 전력 소모를 줄이고자 할 때는 little core를 사용한다. 따라서 실제 사용되는 core 수는 quad-core CPU에 해당한다. 2017년 ARM사는 big-little core뿐만 아니라 SoC 내부의 accelerator processor 간 work load 및 data sharing 을 최대한 효율적으로 운영하기 위한 기술로 자체 개발한 DynamIQ technology를 발표하였다. 계약된 license 유형에 따라 Huawei와 Qualcomm의 최신 SoC에는 이 기술이 적용되어 있고 Samsung과 Apple의 SoC에는 적용되어 있지 않다. 이 경우 자체 개발한 기

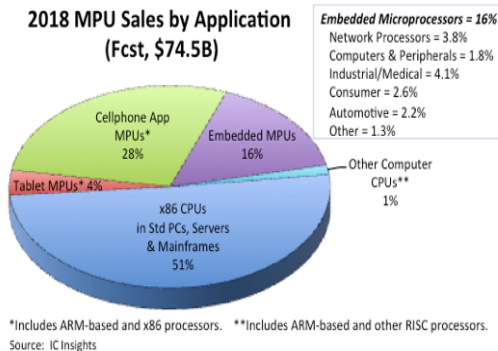


그림 7 ARM CPU와 x86 CPU의 시장점유율

술을 사용하거나 ARM사의 과거 기술을 사용하여야 한다. DynamIQ 기술은 big-little뿐만 아니라 big-middle-little 등 다양한 구성이 가능하도록 하였고, Qualcomm의 Snapdragon 855의 경우 1-3-4로 구성되어 있다. ARM사의 processor 모델 중 DynamIQ 기술과 호환되는 모델은 Cortex-A55, Cortex-A75, Cortex-A76, Cortex-A77이고 Cortex-A55를 little로 구성하고 나머지를 middle 또는 big으로 사용한다[6].

ARM사 license는 계약하는 기업에 따라 다양한 options이 적용되고 제3자가 파악하기 어렵지만 대략 3개 유형으로 구분된다고 한다(표 4 참조). Cortex license는 ARM사가 instruction set과 CPU core에 대한 회로설계, simulation과 verification 작업을 완료한 상태까지 제공한다. Partner사는 SoC에 포함된 다른 구성 성분(예, GPU, NPU 등)에 대한 설계와 위 CPU 설계를 포함하여 제작에 필요한 mask 작업을 하게 된다. Huawei CPU가 이

유형에 속하는 것으로 알려져 있다. Architecture license는 ARM사로부터 instruction set만 제공받고 core 설계 및 검증과 제작에 필요한 후속 작업은 partner사가 한다. 그리고 2016년부터 제공하는 'Built on ARM Cortex' license는 partner사가 필요한 요구사항을 instruction set에 반영하도록 요구할 수 있고, 이렇게 만들어진 customer 맞춤형 ISA는 타사에 제공하지 못하는 것으로 알려져 있다. ARM사의 기술제공 범위는 cortex license처럼 회로 설계와 검증이 완료된 단계까지 정도이다. 이와 같은 단순 분류 외 각 유형별 또는 partner별 많은 세부사항이 있는 것으로 알려져 있다. Apple, Samsung, Qualcomm은 Architecture license 유형에 속하고 Qualcomm은 그 후 'Built on ARM Cortex' license로 변경한 것으로 알려져 있다. 그리고 Architecture license의 경우 ARM사가 제공하는 DynamIQ 기술을 사용할 수 없다. Cortex license를 제외한 나머지 두 경우는 processor에 partner사 자체 brand name을 사용할 수 있다.

2019년 현재 ARM CPU의 IP license를 구매하여 스마트폰 CPU가 포함된 SoC 공급 기업 중 기술과 시장점유율이 가장 높은 기업은 Qualcomm, Samsung, Apple, Huawei 4개사라 할 수 있다. 통계 업체(Statista) 자료에 의하면 스마트폰 SoC에 대한 2019년 2분기 시장점유율은 그림 8과 같다.

표 4 ARM사의 license 유형

Cortex License	<ul style="list-style-type: none"> • Partner licenses complete microarchitecture design • CPU differentiation through -flexible configuration options -wide implementation envelope with different process technologies
Architecture License	<ul style="list-style-type: none"> • Partner licenses complete CPU microarchitecture from scratch • Freedom to develop any design -must conform to the rules and programmers model of a given architecture -Must pass ARM architecture validation to preserve software compatibility
'Built on ARM Cortex' license	<ul style="list-style-type: none"> • A new engagement model for Cortex-A partners, announced in Feb. 2016 • Partners work with ARM to develop a custom RTI, databases

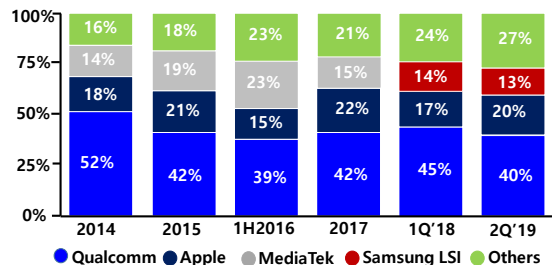


그림 8 ARM CPU 기업별 시장점유율

표 5 2019년형 SoC 특징

	Qualcomm		Samsung
2019년 SoC	Snapdragon 855, Snapdragon 855+	2019년 SoC	Exynos9820, Exynos9825
2019년 출시 Top device	Galaxy S10+, Galaxy note 10+, LG G8 ThinkQ, Oneplus 7 pro	2019년 출시 Top device	Galaxy S10+, Galaxy note 10+
주요 특징	CPU 1개 Kryo 485 Prime (A76-based), 2.96GHz, L2 cache 512KB 3개 Kryo 485 (A76-based), 2.4GHz 각 L2 cache 256KB 4개 Kryo 485 (A55-based), 1.8GHz 각 L2 cache 128KB L3 cache 2MB+system cache 3MB GPU Adreno 640(~672MHz) DSP Hexagon 690(AI engine) ISP 48MP at 30fps single camera, 22MP at 30fps Dual Camera Modem Internal X24 LTE Modem(4G): Download: 2000Mbit/s, Upload: 316Mbit/s External Snapdragon X50(5G): Download: 5000Mbit/s Memory controller LPDDR4X Quad-channel 16-bit (64-bit) 2133MHz(34.13GB/s) TSMC 7nm FinFET, Die size: 73mm ²	주요 특징	CPU 2개 M4(A76-based), 2.73GHz 각 L2 cache 512KB 2개 Cortex A75, 2.4GHz 각 L2 cache 256KB 4개 Cortex A55, 1.95GHz L3 cache 4MB GPU Mali G76MP12, 702MHz NPU Dual-core 1024 MAC units @933MHz ISP 8K @30fps & 4K @150fps Modem Internal Shannon 5000 LTE(4G) Download: 2000Mbps, Upload: 316Mbps External 5G Exynos Modem 5100. Memory controller LPDDR4X Quad-channel 16-bit (64-bit) 2093MHz(33.53GB/s) Samsung 7nm FinFET
	Huawei		Apple
2019년 SoC	Kirin990-5G	2019년 SoC	A13 Bionic
2019년 출시 Top device	Mate 30 Pro, Mate 30	2019년 출시 Top device	iPhone 11, iPhone 11 Pro, iPhone 11 Pro Max
주요 특징	CPU 2개 Cortex-A76, 2.86GHz, 각 L2 cache 512KB 2개 Cortex-A76, 2.36GHz, 각 L2 cache 512KB 4개 Cortex-A55, 1.95GHz, 각 L2 cache 128KB L3 cache: No information available GPU Mali G76MP16, 600MHz NPU 2+1 Da Vinci NPU ISP 8K @30fps, 4K @60fps Modem Integrated Balong 5G Modem, Download: 4.6Gbps(최대 7.5Gbps), Upload 2.5Gbps Memory controller LPDDR4X Quad-channel 16-bit 2133MHz(34.13GB/s) TSMC 7nm FinFET, Die size: 100mm ²	주요 특징	CPU 2개 Lightning(A76-based), 2.65GHz 4개 Thunder(A55-based), 1.82GHz L2, L3 cache: No information available GPU Apple Custom GPU(quad-core) NPU Neural Engine(octa-core) + AMX Blocks(dual-core) ISP 4K @60fps Modem Internal: Intel Gigabit-Class LTE(4G) No 5G Modem supported Memory controller LPDDR4X Dual-channel 64-bit 2133MHz(68.2GB/s) TSMC 7 nm FinFET, Die size: 98.48mm ² OS: iOS 13.1.1

Qualcomm 40%, Samsung 13%, Apple 20%, 기타 27%이다. 기타에는 HiSilicon과 MediaTek이 가장 큰 점유율을 가지고 있고 HiSilicon은 Huawei의 설계 전문 자회사이고 MediaTek은 대만의 SoC 설계 전문회사이며 중저가 스마트폰을 대상으로 SoC를 공급하고 있다. 2019년 기준 가장 높은 기술을 파악하기 위해서는 위 4개사의 SoC를 분석할 필요가 있다. 표 5는 2019년 출시한 위 4개사의 SoC 모델명 및 특징 그리고 SoC가 적용되어

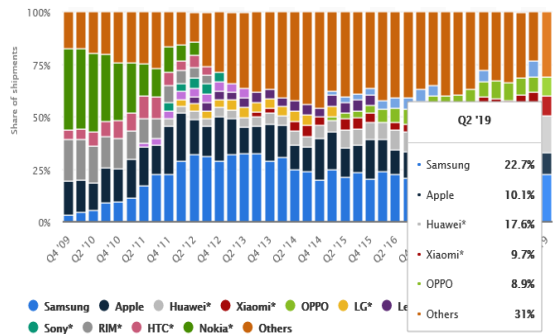


그림 9 분기별 스마트폰 시장점유율

출시된 2019년 스마트폰 모델명을 나타낸다. 각 SoC에는 공통적으로 6개 또는 8개 core로 구성된 CPU, GPU, AI engine, ISP, 4G 또는 5G modem, memory controller가 약 1cm² 크기의 칩에 집적되어 있고 총 transistor 수는 80~100억 개다.

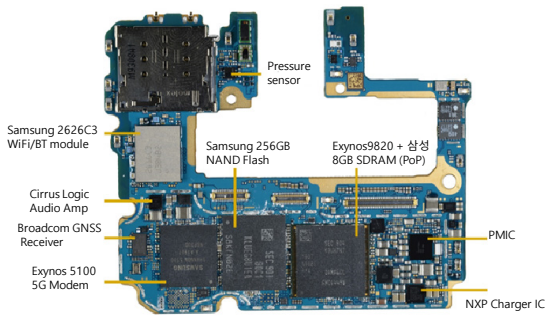
4개사 중 자체 기술력이 가장 높은 회사는 Qualcomm이라 할 수 있다. 자사 모델의 스마트폰은 생산하고 있지 않지만 LG를 포함한 다수의 기업이 Qualcomm SoC를 사용하여 2019년 최신 스마트폰을 출시하였다. 삼성 Galaxy S10과 note10의 경우 미국, 일본, 중국에 출시하는 제품에 Qualcomm의 SoC를 사용하고 유럽, 한국, 기타 시장에는 삼성 SoC인 Exynos를 사용한다. Qualcomm은 CPU를 제외한 나머지 소자는 모두 자사가 보유한 기술로 설계하였고 CPU는 ARM사의 'Built on ARM Cortex' license 계약으로 자사 brand인 Kryo485, 8개 core를 사용하였다. Qualcomm SoC는 2018년 10nm finFET까지 Samsung foundry를 이용하여 생산하였으나 2019년 7nm finFET부터 대만의 TSMC foundry를 사용하고 있다.

Samsung은 2019년 초반까지 8nm finFET이 가능하였고 2019년 후반부터 최초로 EUV(extreme ultra violet) lithography를 적용하여 7nm finFET 기술을 확보한 것으로 알려져 있다.

삼성은 2010년 미국 텍사스 Austin 시에 SARC (Samsung's Austin R&D Center)을 설립하고 전직 AMD, Intel, ARM 출신 엔지니어들을 영입하여 SoC설계를 하는 것으로 알려져 있다. 또한 2017년 캘리포니아 San Jose에 ACL(Advanced Computing Lab)을 설립하고 MediaTek의 GPU 기술을 도입하여 GPU를 개발하는 것으로 알려져 있다. Exynos 9820, 9825에는 자체 brand명을 붙인 M4 core를 사용하고 있다. Exynos의 일부 core와 GPU는 ARM사의 설계를 그대로 사용하고 있지만 그 외 소자들

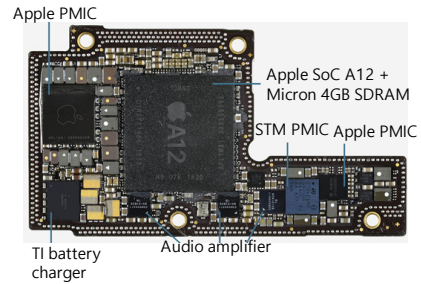
은 자체 기술로 개발한 것으로 알려져 있다. 그림 9는 스마트폰 생산업체의 시장점유율 통계(Statista 2019)를 나타낸다. 2019년 2분기 기준 삼성은 22.7%로 세계최대 스마트폰 생산업체이다. Huawei는 2018년 Apple를 제치고 점유율 2위의 스마트폰 생산업체로 올랐고 기술개발을 위해 현재 가장 과감한 투자를 하고 있다. HiSilicon은 Huawei가 전적으로 소유하고 있는 SoC설계 전문 자회사이고 Kirin 980, 990 등 Huawei의 SoC는 HiSilicon에서 개발되었다. CPU와 GPU는 ARM사 설계를 그대로 사용하고 있지만 나머지 소자는 자체 기술인 것으로 알려져 있다. 특히 Da Vinci NPU와 5G modem은 Qualcomm이나 삼성보다 우수한 것으로 평가받고 있다. 5G modem의 경우 삼성과 Qualcomm은 SoC칩과 별개의 external 칩으로 PCB에 올려져 있지만 Kirin 990 5G는 최초로 SoC 칩 내부에 집적되어 있다. 또한 삼성과 Qualcomm의 external 5G 칩은 이전 기술을 사용하지만 Huawei는 SoC에 적용된 TSMC 7nm 기술을 그대로 사용하여 download/upload 속도에서 상당한 격차가 있는 것으로 평가되고 있다.

Android OS 스마트폰이 약 90%의 시장점유율을 차지한 가운데 iOS만 독자적으로 사용하는 Apple의 시장점유율은 급속히 위축되고 있다. iOS개발, SoC설계, 스마트폰 생산을 하나의 기업이 전담하는 상태에서 경쟁력을 지속적으로 유지하기는 어렵다. SoC의 주요 소자인 CPU, GPU(PowerVR사 license), modem에 자체 기술이 낮은 편이고 스마트폰 생산 역시 위탁 조립하고 있으므로 iOS를 제외한 나머지 부분에서 경쟁력이 떨어진다고 볼 수 있다. 5G modem의 경우 경쟁기업 3사가 2019년 전반기에 개발을 완료하고 하반기에 이미 장착된 스마트폰을 출시한 반면, Intel에서 4G modem를 공급 받아온 Apple은 Intel이 2019년 하반기까지 5G



출처 Reprinted with permission from Techinsights(<https://www.techinsights.com/ko>).

그림 10 Samsung Galaxy S10 5G

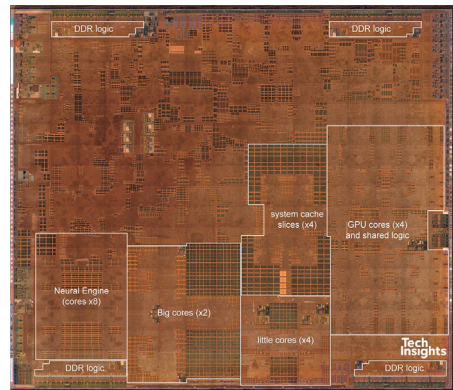


출처 Reprinted with permission from Techinsights(<https://www.techinsights.com/ko>).

그림 11 Apple iPhone Xs Max

칩 개발을 완료하겠다는 약속을 지키지 못함에 따라 Intel과 공급계약을 해지한 것으로 알려졌다. Apple은 제3의 5G 칩 공급 회사를 물색하고 Intel은 5G 칩 개발을 포기한 상태에서 결국 2019년 7월 Apple은 Intel의 modem 사업을 약 1.2조 원에 사들여 하반기부터 자체 개발에 나선 것으로 알려졌다. 2020년 출시 Apple 폰에 5G 장착이 어려울 경우 Apple의 시장점유율은 더욱 위축될 전망이다. 2019년 9월 기준 5G modem 칩 개발이 완료된 회사는 위 3사 외 MediaTek이 있고, Spreadtrum이 이름을 바꾼 Unisoc는 저가 폰에도 5G 칩 장착을 준비 중인 것으로 알려져 있다. Huawei와 Unisoc는 사실상 중국 국가 소유로써 기술이 공유된다고 볼 수 있다.

그림 10은 2019년 출시된 Samsung Galaxy S10 5G의 메인 PCB사진을 나타낸다. 미국의 주요 기술 분석 회사들은 신제품 출시 후 몇 주일 이내에 분해(teardown) 사진과 함께 주요부품, 각 부품의 원가를 분석하여 인터넷에 올린다. 이 사진은 Techinsights 사의 자료를 인용한 것이다. PoP(pack-on package)로 서로 맞붙여 놓은 삼성 SoC Exynos9820과 메인 메모리인 8GB SDRAM이 있다. 그리고 삼성 Exynos5100 external 5G modem과 256GB flash memory, WiFi/Bluetooth 모듈이 있다.



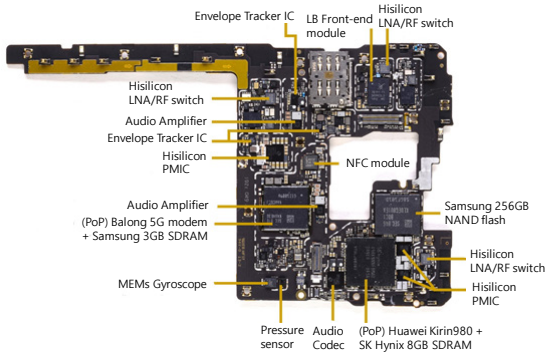
출처 Reprinted with permission from Techinsights(<https://www.techinsights.com/ko>).

그림 12 Apple SoC A12 Bionic

표 6 2019년 부품 원가

Samsung		Apple	
부품	원가	부품	원가
Exynos9820	\$70.5	SoC A12	\$72
Camera/Image sensor	\$62.5	Camera	\$44
Touchscreen	\$90	Touchscreen	\$90.5
SDRAM	\$37.5	SDRAM+ 256GB Flash	\$64.5
RF modules	\$46	Mixed signal/ RF	\$23
5G Modem	\$32.08	Mechniclas/ Housing	\$58
256GB Flash	\$18	기타	\$101
기타	\$133.5	원가 총액	\$453.00
원가 총액	\$490.08		

출처 Reprinted with permission from Techinsights(<https://www.techinsights.com/ko>).



출처 Reprinted with permission from Techinsights(<https://www.techinsights.com/ko>).

그림 13 Huawei mate 20X(5G)

PCB는 큰 것과 작은 것 2개로 이루어져 있고 큰 메인 PCB는 전면과 후면에 소자가 부착되어 있으며 그림 10에는 메인 PCB 전면만 나타내었다.

그림 11은 2018년 출시된 Apple iPhone Xs Max의 메인 PCB 전면 사진이다. 삼성 제품처럼 PCB는 2개로 이루어져 있고 그림에 나타내지 않았으나 Intel의 4G LTE modem은 작은 PCB에 있다. Apple SoC A12 Bionic와 Micron사의 4GB SDRAM이 PoP 기술로 접합되어 있고 Sandisk사의 256GB flash 메모리가 있다. 각 부품 원가는 표 6에 나타나 있다. 그림 12에는 2018년 출시된 Apple SoC A12 Bionic에서 SDRAM을 떼어낸 후 모습을 나타내고 있다. Big cores, little cores, caches, GPU, neural engine, 등의 위치가 표시되어 있다.

그림 13은 2019년 출시된 Huawei mate 20X(5G)의 메인 PCB 전면 사진이다. 2019년 하반기 출시된 mate 30Pro(5G)제품에 5G modem이 SoC에 집적되어 있지만 이 제품에는 external 5G가 삽입되어 있다. Huawei SoC Kirin980 과 SK Hynix 8GB SDRAM이 PoP 기술로 접합되어 있고 삼성 256GB flash memory가 있다. 이 제품에 대한 부품 원가는 공개되어 있지 않다.

IV. 결론

CPU의 성능을 높이기 위해 many cores, multi-threads, 병렬 컴퓨팅, superscalar, vector, out-of-order, branch prediction, scheduling 등 다양한 기술이 설계와 제작에 적용되고 있다. 특히 병렬 컴퓨팅을 가능하게 하는 가장 필요한 기술 중 하나가 cache coherence이다. Core 간 통신속도가 중요한 역할을 하며 CPU 내 광신호를 도입하기 위한 연구가 2000년대 중반부터 활발히 진행되어 왔으나 아직 성과가 나타나지 않고 있다.

2019년 2분기 기준 삼성은 22.7%로 세계 최대 스마트폰 생산업체이고, Huawei는 2018년 Apple(10.1%)를 제치고 점유율 2위(17.6%)로 올랐고 기술개발을 위해 현재 가장 과감한 투자를 하고 있다. 특히 Da Vinci NPU와 5G modem은 Qualcomm이나 삼성보다 우수한 것으로 평가받고 있다.

Qualcomm SoC는 2018년 10nm finFET까지 Samsung foundry를 이용하여 생산하였으나 2019년 7nm finFET 부터 대만의 TSMC foundry를 사용하고 있다. Samsung은 2019년 초반까지 8nm finFET이 가능하였고, 2019년 후반부터 최초로 EUV(extreme ultra violet) lithography를 적용하여 7nm finFET 기술을 확보한 것으로 알려져 있다. Huawei는 TSMC 7nm기술을 적용하여 2019년 후반에 Kirin 990 5G을 생산하였고 최초로 5G 모뎀을 SoC 칩 내부에 집적하였다.

참고문헌

- [1] Daniel Page, "A Practical Introduction to Computer Architecture," Springer, ISBN 978-1-84882-255-9, 2009.
- [2] Mats Brorsson, "Chapter2 Multi-core and Many-core Processor Architectures," in the book 'Programming Many-Core Chips,' ed. by A. Vajda, Springer Science+Business Media, LLC 2011.

- [3] Aad J. van der Stee, "Overview of recent supercomputers," NCF/HPC Research, Netherlands, August 2008.
- [4] Arm Company, "Arm Architecture Reference Manual; Armv8, for Armv8-A architecture profile," Copyright © 2013-2019 Arm Limited or its affiliates. All rights reserved. Non-Confidential.
- [5] Arm Company, "ARM Compiler version 6.6; User Guide" Copyright © 2013-2016 Arm Limited or its affiliates. All rights reserved. Non-Confidential.
- [6] Emily Blemd et al., "A Detailed Analysis of Contemporary ARM and x86 Architecture" 19th IEEE Intl. Symposium on High Performance Computer Architecture, HPCA 2013.
- [7] Antonio La Manna, "Outlook for 3D and 2.5D chips in smartphones," Chip Scale Review, vol. 19, no. 2, March 2015. pp. 9-11.