

## < CHIP 4주차 선각 레포트 >

### - 더 빠른 CPU는 없다: 아키텍처의 전환점에서

KJ-8-301-05 김재현

한때 CPU는 ‘무어의 법칙’이란 마법 아래에서 매년 성능을 거침없이 끌어올렸다. 새로운 세대가 등장할 때마다 두 배에 가까운 속도 향상은 거의 당연한 일이었고, ‘트랜지스터 수의 증가 = 성능의 증가’라는 공식을 부정하는 이도 드물었다. 하지만 오늘날, 우리는 그 공식을 더 이상 신뢰하지 않는다. 단순히 ‘클럭 속도 몇 MHz 증가’로는 체감 가능한 성능 향상을 기대하기 어려운 시대가 왔다.

그렇다면, CPU 성능 향상의 둔화는 단순한 기술 정체가 아니라, 구조적이고 물리적인 한계의 결과라고 보는 것이 타당하다. 첫 번째 벽은 바로 미세 공정의 한계다. 오늘날 5nm 이하의 공정은 양자역학적 현상, 특히 터널링 효과와 누설 전류의 증가라는 치명적 문제를 야기한다. 전자는 더 이상 우리가 설계한 트랜지스터 안에서만 움직이지 않으며, 이로 인해 전력 손실과 발열이 심화된다. 이는 곧 안정적인 고클럭 구현을 어렵게 만든다.

두 번째는 전력과 열의 문제다. 전력 밀도가 높아질수록 열도 기하급수적으로 상승하며, 이는 결국 클럭 속도의 증가를 물리적으로 제한하는 요인이 된다. 전력이 충분하더라도, 그것을 감당할 열 해소 장치가 없다면 설계는 무의미해진다. 과거와 달리 클럭 속도가 5GHz를 넘어가는 순간, 우리는 냉각 솔루션에 대해 심각하게 고민해야 한다.

세 번째는 아키텍처적 병목이다. 현대 CPU는 이미 명령어 수준 병렬성(ILP)의 한계에 도달했다. 슈퍼스칼라, 파이프라이닝, 아웃오브오더 실행 등 다양한 기법들이 총동원되었지만, 결국 한 스레드 내에서의 병렬 처리에는 물리적 한계가 존재한다. 더 이상 단일 스레드 기반의 성능 향상은 기존 방식으로 달성되기 어렵다.

이러한 한계들을 극복하기 위한 대안으로, 우리는 멀티코어 아키텍처와 이기종 컴퓨팅(Heterogeneous Computing)에 주목하고 있다. 멀티코어는 하나의 칩에 여러 개의 코어를 집적함으로써 전체 처리량을 증가시키는 접근이다. 하지만 이것 또한 소프트웨어 병렬화가 수반되지 않으면 실질적인 성능 향상을 기대하기 어렵다. 이는 결국 소프트웨어-하드웨어 간의 공동 설계(co-design)가 필수적인 시대로의 전환을 의미한다.

또한 GPU, TPU, NPU 등 특화된 연산 유닛의 부상이 중요한 흐름이다. 이는 범용 연산 장치로서의 CPU가 특정 목적에 한정된 연산에서는 비효율적이라는 판단에서 비롯된 진화이며, AI 및 머신러닝과 같이 대규모 행렬 연산이 반복되는 분야에서는 가속기의 역할이 점점 더 중요해지고 있다. Google TPU, Apple Neural Engine, AMD XDNA AI Engine 등의 사례는 이 흐름을 잘 보여준다.

하드웨어 설계 방식도 변화하고 있다. 칩렛(Chiplet) 아키텍처는 대형 단일 다이 대신, 여러 개의 소형 칩을 패키지 수준에서 통합함으로써 수율과 비용 측면의 장점을 얻는다. 여기에 3D 집적(3D IC), TSV(Through-Silicon Via) 기술이 결합되면, 면적 한계를 넘어서 수직적 확장을 가능하게 한다. AMD의 3D V-Cache, Intel의 Foveros 등이 그 대표 사례다.

결론적으로, CPU 성능 향상의 둔화는 기술이 정체되어서가 아니라, 기존의 성장 공식을 지속할 수 없게 된 구조적 한계의 반영이다. 이 변화는 종말이 아니라 전환이며, 성능의 정의 자체가 ‘클럭 속도’나 ‘트랜지스터 수’가 아닌, ‘작업 적합성’과 ‘시스템 최적화’로 이동하고 있음을 보여준다. CPU는 더 이상 홀로 성장하지 않는다. 그 곁에는 다양한 가속기, 유연한 패키징 기술, 정교한 소프트웨어 생태계가 함께 서 있다. 그리고 우리는 이제 ‘더 빠른 칩’이 아닌, ‘더 똑똑한 플랫폼’을 설계해야 할 시점에 와 있다.

---

## 참고문헌

1. Hennessy, J. L., & Patterson, D. A. (2019). *Computer Architecture: A Quantitative Approach* (6th ed.). Morgan Kaufmann.
2. Yazdanbakhsh, A. et al. (2023). *Beyond Moore's Law: A Call for Hardware-Software Co-Design*, arXiv:2304.08931.
3. Sze, V. et al. (2017). *Efficient Processing of Deep Neural Networks: A Tutorial and Survey*, Proceedings of the IEEE.
4. TSMC (2022). *SoIC and Chiplet Integration*, Technology White Paper.
5. Borkar, S., & Chien, A. A. (2011). *The Future of Microprocessors*, Communications of the ACM, 54(5), 67–77.