



**CPU의 성능 향상이 둔화된 이유와 기술적 대안**

**8기-23 정승민**

## 서론: 무어의 법칙과 CPU 성능 둔화

반도체 업계는 오랫동안 무어의 법칙(Moore's Law)에 따라 CPU 성능이 기하급수적으로 향상되어 왔다. 무어의 법칙이란 집적회로에 들어가는 트랜지스터 수가 약 2년마다 두 배로 증가한다는 관찰로, 1965년 인텔의 공동 창업자 고든 무어(Gordon Moore)가 제시한 개념이다. 실제로 1970~90년대에는 반도체 공정의 미세화가 마이크로프로세서의 성능과 에너지 효율을 꾸준히 향상시켰으며, 비용 대비 성능도 크게 개선되었다. **하지만 2000년대 중반 이후 이러한 성능 향상의 지속적인 추세에 제동이 걸렸다. 트랜지스터 수가 증가함에도 불구하고 CPU의 실제 성능 개선 속도가 크게 둔화된 것이다.**

CPU 성능 둔화의 주된 원인은 **데나드 스케일링(Dennard Scaling)의 한계**에서 찾을 수 있다. 데나드 스케일링이란 트랜지스터가 작아질수록 전력 밀도가 일정하게 유지되어 칩의 소비 전력이 면적에 비례해 줄어든다는 이론으로, 과거에는 미세 공정으로 클럭 속도를 높여도 전력 문제가 심각하지 않았다. 그러나 트랜지스터 선폭이 나노미터 수준으로 줄어들면서 누설 전류 증가와 발열 문제로 인해 2007년경 데나드 스케일링이 사실상 종료되었다. 이에 따라 더 이상 클럭 주파수를 높이거나 전압을 올리는 방법으로 성능을 향상시키는 것이 어려워졌다.

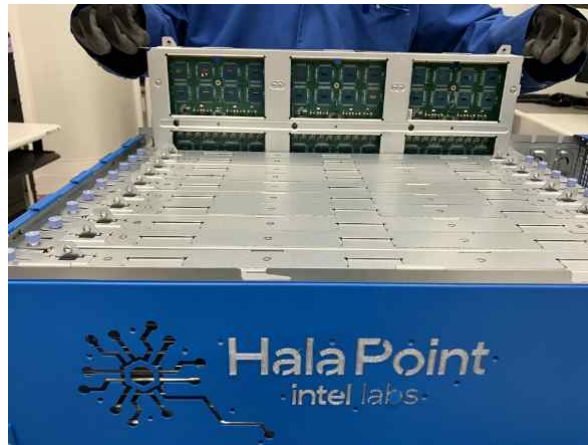
또한 **메모리 장벽(Memory Wall)**과 **명령어 수준 병렬성 장벽(ILP Wall)**도 성능 향상을 가로막는 주요 요인이 되었다. CPU의 연산 속도는 급속히 빨라졌으나 메모리 접근 속도는 상대적으로 느리게 발전하면서 CPU와 메모리 간 속도 격차가 심화되었다. 이로 인해 메모리 접근 시 발생하는 대역폭 한계와 지연이 병목으로 작용하고 있다. 아울러 명령어 수준 병렬성을 극대화하기 위해 복잡한 기법들이 도입되었지만, 추가적인 성능 향상 효과는 갈수록 미미해지고 있다.

## 본론: 기술적 대안

이러한 기술적 한계 속에서, 과거 **증기기관이 산업 발전에 크게 기여했으나 현재는 더 이상 주력 기술로 활용되지 않고 기반 기술로 남아있듯이, 트랜지스터 증가를 기반으로 한 무어의 법칙 또한 한계에 다다랐다.** 따라서 기존 폰 노이만(Von Neumann) 구조에서 벗어나 새로운 방식의 아키텍처를 도입해야 한다는 필요성이 대두되고 있다.

## 비폰 노이만(Non-Von Neumann) 구조

비폰 노이만 구조란 전통적인 컴퓨터 구조를 탈피한 모든 새로운 접근 방식을 의미한다. 대표적인 예로 메모리와 연산의 물리적 분리를 최소화하여 데이터 이동을 줄이고 병렬성을 극대화하는 **인-메모리 컴퓨팅(In-Memory Computing)**과 인간 두뇌의 신경망 구조에서 영감을 얻은 **뉴로모픽(Neuromorphic)** 컴퓨팅이 있다.



인텔이 구축한 대규모 뉴로모픽 시스템 Hala Point의 모습.

뉴로모픽 컴퓨팅은 신경망의 동작 원리를 모방한 구조로, 수많은 연산 코어(뉴런)와 메모리 소자(시냅스)를 네트워크 형태로 연결하여 중앙 클럭 없이 이벤트 기반으로 병렬 연산을 수행한다. 이러한 방식은 기존 **CPU 대비 매우 적은 전력 소모로 대규모 병렬 연산이 가능하며 에너지 효율적이다**. 실제로 인텔의 대규모 뉴로모픽 시스템인 Hala Point는 기존 디지털 컴퓨터 대비 훨씬 뛰어난 에너지 효율과 병렬 처리 능력을 보여주고 있다.

이외에도 메모리 중심 아키텍처나 아날로그 계산 기법 등이 비폰 노이만 구조에 포함된다. 예컨대 메모리 내 연산 기술에서는 저항 변화 메모리(memristor)나 상변화 메모리(PCM) 등의 신소자를 이용하여 **메모리 셀 어레이에서 곧바로 행렬-벡터 연산을 병렬로 수행하는 연구가 활발히 진행 중이다**.

## 양자 컴퓨팅

또 다른 유망한 기술적 대안으로 양자 컴퓨팅이 있다. 양자 컴퓨팅은 기존의 디지털 컴퓨팅과 근본적으로 다른 양자 역학의 원리를 활용한 기술이다. 고전 컴퓨터는 0과 1의 비트(bit)만을 사용하지만, 양자 컴퓨터는 0과 1이 동시에 존재할 수 있는 중첩(superposition) 상태를 갖는 큐비트(qubit)를 사용한다. 나아가 양자 얽힘(entanglement)을 통해 멀리 떨어진 큐비트 간에도 강한 상관관계를 형성하여, 고전 컴퓨터로 해결하기 어려운 문제를 효율적으로 처리할 수 있다. 이러한 특성을 이용하면 고전 알고리즘에서는 지수적으로 긴 시간이 필요한 문제를 다항 시간 내에 해결

할 가능성이 제시된다. 현재 양자 컴퓨팅은 아직 초기 단계이지만, IBM, 구글, 인텔 등 주요 산업체에서 적극적으로 투자하며 실용화를 앞당기기 위해 노력하고 있다.

## **결론 : 포스트 무어 시대로의 전환**

CPU 성능 향상의 기반이었던 무어의 법칙과 데나드 스케일링은 이제 명확한 한계에 직면했다. 트랜지스터를 더 작고 빠르게 만드는 것이 점점 어려워짐에 따라 범용 프로세서의 성능 향상 곡선은 완만해지고 있다. 이를 보완하기 위해 병렬화와 특화된 컴퓨팅 기술이 주목받고 있으며, 비폰노이만 구조와 양자 컴퓨팅과 같은 혁신적인 접근 방식이 향후 컴퓨팅 기술의 주요 발전 방향으로 자리 잡을 것으로 기대된다.