

Modern Computer Architecture Research

2017320114 컴퓨터학과 최재원

이번 수업에서는 구건재 교수님의 최신 컴퓨터 구조 연구에 대한 세미나를 들었다. 최초의 컴퓨터는 instruction에 대한 개념이 많지 않았다. 보통 소프트웨어를 하드웨어 맞춰서 설계를 했다. 그러다 IBM에서 최초로 instruction set architecture라는 개념이 나오기 시작했다. Instruction set architecture란 소프트웨어한테 어떤 instruction은 "이 하드웨어가 실행할 수 있어" 라고 알려주는 어떤 하나의 약속이라고 할 수 있다. 소프트웨어가 하드웨어가 이해할 수 있는 instruction으로 컴파일이 되고, 이 translation이 되면 하드웨어는 소프트웨어가 원하는 결과를 내어준다. ISA의 개념으로 컴퓨터를 얘기한다고 하면, ISA로 정의가 된 instruction을 processor에 넣어줘서 이 컴퓨터 프로세서가 정의된 결과를 만들어 내는 기계라고 보면 된다. 즉 컴퓨터가 가지고 있는 상태에 변화를 준다. 상태(state)란 데이터를 의미한다. 우리가 일상생활에서는 빠르고 cost-effective한 컴퓨터를 원하게 된다. 이러한 연구는 microarchitecture에 대한 연구라고 한다. Computer Architecture에 대한 연구는 microarchitecture도 포함을 하고 있다.

우리가 좋은 processor를 만들려면 transistor가 많이 들어가야 하며, 프로세서 칩이 전기를 적게 소모해야 한다. 현재는 우리가 점점 transistor를 작게 만들어도 power는 점점 더 많이 쓰게 된다. 그만큼 더 좋은 transistor를 만드는 것은 힘들다. 또한 single core processor의 성능 역시 포화상태이다. 그렇다면 어떻게 더 좋은 컴퓨터 시스템을 만들 것인가? Moore's Law가 끝나는 것에 대한 해결 방법으로 Domain Specific Architecture가 나온다. 현재 transistor의 technology는 한계에 부딪히고 있고, 어플리케이션은 점점 더 높은 성능을 요구하고 있다. 이것을 해결하려면 구조적인 변화, Domain Specific Architecture라는 새로운 구조를 통해 해결하려고 한다. DSA의 예로는 TPU, GPU 등이 있

다. 프로세서라는 것은 instruction을 받아서 instruction에 맞게 데이터를 불러와서 데이터를 isa에서 정리한 그런 결과에 맞게 만들어 내는 과정을 하고 있다. Instruction을 거기에 맞도록 계산하는 것은 오버헤드가 크다. 그렇지만 프로세싱에만 집중을 하게 된다면 오버헤드를 줄일 수 있다. TPU는 matrix multiply unit이 있다. CPU는 하나하나 계산을 해야 되는데, TPU는 한 번에 계산을 한다. Matrix Multiplication을 요구하는 어플리케이션들이 점점 늘어나고 있기 때문에 TPU가 필요로 한다.

빅데이터를 위한 컴퓨터 시스템은 data processing을 위한 높은 성능을 필요로 하고, 많은 데이터를 저장을 했다가 컴퓨터가 처리를 해야 되기 때문에 좀 더 발전된 메모리 시스템을 요구를 하고 있다. 메모리 hierarchy가 가능했던 이유는 우리가 적은 양의 데이터를 반복적으로 사용했기 때문에 캐시 등에 넣고 storage까지 갈 필요가 없었다. 문제는 데이터의 사이즈가 너무 커지면서 이런 memory hierarchy의 장점이 사라지고 있다. In-Storage Computing이란 computing unit을 데이터 근처로 이동시키는 것이다. 데이터를 옮기는 것은 오래 걸리기 때문에 computer unit을 옮기는 것이다. 지금까지 conventional한 computing 구조는 밑에 있는 데이터를 processor 근처까지 어떻게 효율적으로 옮기는 것이었다면 이 방법은 예전과는 다른 것이다.

이처럼 현재 컴퓨터 하드웨어가 점점 발전하는 소프트웨어에 맞춰서 요구하는 성능을 높이는 연구와 방법에 대해서 배웠는데, 컴퓨터 구조 및 논리 설계를 인상 깊게 들은 학생으로서 여태까지 세미나에서 ai에 관련된 얘기를 했다면 하드웨어 측면에 대해서 들을 수 있어서 매우 흥미로운 시간이었다. 고성능 저전력을 어떻게 하면 극대화시킬 수 있는지 앞으로 기회가 된다면 연구해 보고 싶다는 생각도 가지게 되는 유익한 세미나였다.