

50 YEARS OF COMPUTER ARCHITECTURE

IBM'in 4 uyumsuz bilgisayarını tek ISA hattında birleştirmeye çalışması ile başlayan serüven, 64 bit sürümün keşfini doğurdu. O zamanlar ROM, RAM den çok daha ucuzu ve Wilkes, kontrolün ROMdan yapılmasını istedi ve bu sayede kontrolü daha rahat ele alacaklardı. Bu sayede 7 Nisanda 1964'de IBM, eski mimariden uzaklaşıp herşeyin tek bir ISA hattında birleşeceğini açıkladı. ROM ve RAM aynı transistörlerden yapılmasına rağmen Moore yasasına dayanarak daha fazla işlemi yapan mikroyapılar üretilecekti. Fakat bu şekilde mikro programlar bazı buglara sahip olacaktı ve bu nedenle RAM üzerinden kontrol edilmesi gerekiydi. Moore mikroişemicilerin mini bilgisayarlarla aynı işlemi düşük maliyette yapabileceğine inandığı için yapılan araştırmalar sonucu yeni program olan ADA'yı ürettiler. Fakat Intel 8080 ile uyumlu olmayacağı, yavaş ve maliyetli olacağını öğrenen Moore 3 haftalık derin bir çalışmasıyla isteklere uygun bir mikroişlemci tasarladı. Uyumluluk için Intel 8080 sürümüne 20 bit ekleme yapıldı. Bir yandan da yeni müşteri kitlesi oluşturmak için IBM çalışmalar gerçekleştirdi. Motorola 68000 modelini uygun bulmayan IBM tekrardan Intel 8086 sürümü ile yola devam edecekti. Beklenen satış miktarının çok çok üzerinde satış yapan IBM, Intel 8086 sayesinde inanılmaz kar etti. Moore mikromimari için öngörülen tahminlerde bulunsa da kalıcı olmak için Intel 8086 sürümünü 32 bit boyutuna geliştirdi ve bu sayede yaklaşık 15 yıl boyunca sabit kalacak mimariyi ilk gerçekleyen isim oldu. DEC mühendislerinin yaptığı çalışmaları değerlendiren John Cocke sayesinde yapılan denemeler ile eskiye oranla daha hızlı işlem yapılması ile yeni hedef kontrol deposunun hızlı RAM'ini kullanıcı tarafından görülebilen talimatların bir talimat önbelleğine dönüştürmekte. Moore yasasına dayanarak 32 bit veri yolunun tek yongada birleşebileceğini öngören araştırmacılar, karşılaşlıklarını ikileme özgü bir formul ile sorunu ortadan kaldırdılar. Devre tasarıımı göz önüne alındığında, 80x86 sonunda en hızlı işlemcilere sahipti ve PC pazarına hakimdi. 2007 yılında Apple şirketinin kurulmasıyla Intel ve AMD'nin pazar payları gerilerken yapılan üretim sayıları ve pazar payları azaldı. Bu durumda zaten 90'lı yıllarda 32bit işlemenin yetersiz kalacağını öngören Intel, AMD ile ortaklık yapmak isteyecekti. Fakat AMD bu ortaklıkta sempozyuma katılmaya davet edilmediği için 64 bit işlemcisini kendi başına çıkarmak zorunda kalacaktı. Buna karşılık Intel kendi yeni 64 bit sürümünü tanıtsa bile AMD64 serisini benimseyecekti. Alan mimarisi için kullanılan 3 kanununun da mikroişlemciler üzerindeki etkisi gözükmektedir ve ilk olarak artık Moore yasası sona ermiş ve transistörler artık alan olarak geliştirilemiyor. İkinci olarak mm2 bbaşına düşen tepe değeri artıyordu ve son oalrak çok çekirdekli kartlar. Bunların hepsi göz önüne alınınca artık tek yolun DSA olduğu kanaatine varılmıştır. TPU'nun blok diyagramı, ana bilgisayarın PC ile veri yolu üzerinden bir talimat arabelleğine talimatlar gönderdiğini gösteriyor ve dahili bloklar tipik olarak 256 bayt genişliğindeki yollarla birbirine bağlanır. Matris biriminin ağırlıkları, Ağırlık Belleği adı verilen bir yonga dışı 8GiB DRAM'den okunan bir yonga üzerinde Ağırlık FIFO ile hazırlanır. FIFO'nun ağırlığı 8 bit ağırlığa kadar dört 256×256 karo tumaktaydı. TPU mikro mimarisinin felsefesi, büyük matris ünitesini meşgul etmekti ve bu amaçla, ağırlıkları okuyan talimat, adresini gönderdikten sonra ancak Ağırlık Belleğinden alınmadan önce tamamlanabilmesi için ayrıstırılmış erişim / yürütme felsefesini takip ederdi. Haswell ve K80, yanıt süresi sınırı gevsetilirse elde edilebilecek en yüksek verimin sırasıyla sadece %42 ve %37'sinde çalışır. Bu sınırlar TPU'yu da etkiler, ancak %80'de en yüksek uygulama verimine çok daha yakın çalışır. Ortalama olarak, TPU CPU ve GPU'DAN 29 ve 15 kat daha hızlıdır. Bu sayede oluşan birçok performans avantajı vardır ve bence bunlardan en önemli TPU, Dnns'nin kullanmadığı CPU'lar ve GPU'ların gerektirdiği özellikleri düşürür, bu da TPU'yu daha ucuz hale getirir, enerji tasarrufu sağlar ve transistörlerin etki alanına özgü optimizasyonlar için yeniden kullanılmasını sağlar. Bu ISA konsensüsü artı Dsa'lara geçiş, RISC-V olarak adlandırılan Isa'lara yeni bir girişime yol açmış ve RISC-V Vakfı'nın amacı, RISC-V'nin istikrarını korumak, teknolojik gereksinimler için yavaş ve dikkatli bir şekilde geliştirmek ve Linux işletim sistemleri için olduğu kadar işlemciler için de popüler hale getirmeye çalışmaktadır ve ayrıca RISC-V'nin amacı, en küçükten en

hızlıya kadar tüm bilgi işlem cihazları için etkili olmakla beraber kar amacı gütmeyen bir temele sahip olarak uzun ömürlü olmak; ve Dsa'ları desteklemek için opcode alanını korumaktır.

Muhammet Fatih KESKİN