به نام خدا

LaTeX

فاطمه علی‌ملکی
امیررضا جهانگیری
محمدحسین چهکندی
مهدی حق‌وردی
خدیجه نظری

دانشگاه اصفهان

فهرست مطالب

مقدمه

مقدمه

- در این ارائه به بررسی معماری کامپیوتر می‌پردازیم
- ابتدا سرگذشت و روند تکاملی معماری را بررسی می‌کنیم،
- سپس به معرفی اجزای اصلی یک کامپیوتر می‌پردازیم،
- پس از آن به داخل CPU می‌رویم و معماری‌های متفاوت آن را می‌بینیم،
- سپس در مورد آینده‌ی معماری کامپیوتر صحبت می‌کنیم
- و در آخر، تاثیر هوش مصنوعی به روی معماری کامپیوتر را بررسی می‌کنیم.

معماری کامپیوتر - دیروز تا امروز

تکامل معماری کامپیوتر

- در دنیای امروزی کامپیوترها برای اهداف زیاد و توسط افراد زیادی استفاده می‌شوند،
- کارها و اتفاقاتی که زمانی غیر قابل تصور بود، برای جامعه‌ی ما بسیار بدیهی و مرسوم است،
- تکنولوژی معماری کامپیوتر در طول سالیان متمادی، عمدتا به دلیل پیشرفت‌های تکنولوژی ساخت قطعات الکترونیکی، پیشرفت علوم کامپیوتر و نیازهای افراد پیشرفت کرده است.

نسل اول کامپیوترها

- در سال ۱۹۳۷، اولین کامپیوتر با استفاده از لامپ‌های خلاء توسط پروفسور ایکن اختراع شد.
- در سال ۱۹۴۷، دانشگاه پنسیلوانیا کامپیوتری به نام ENIAC را طراحی کرد که از مبنای دودویی برای نمایش اطلاعات استفاده می‌کرد.
- معماری کامپیوترهای این دوره (و تمام دوره‌ها،) بر اساس مدل Von Neumann بود (و هست،) که شامل
  ۱. واحد حافظه،
  ۲. واحد پردازش،
  ۳. واحد کنترل و
  ۴. واحد ورودی/خروجی

  می‌شود.

نسل دوم کامپیوترها

- در دهه‌ي ۱۹۵۰، ترانزیستورها به جای لامپ‌های خلاء در کامپیوترها استفاده شدند،
- این باعث کاهش حجم و افزایش سرعت کامپیوترها شد.
- در این دوره کامپیوترهای دیجیتال و مینی‌کامپیوترها شروع به ظهور کردند

نسل سوم کامپیوترها

- در دهه‌ی ۱۹۶۰، مدارهای مجتمع (IC) جایگزین ترانزیستورها شدند.
- استفاده از ICها باعث افزایش قابلیت پیچیدگی و کارای کامپیوترها شد.
- این به این معنی‌ست که تعداد بیشتری ترانزیستورها را در یک تراشه کوچک‌تر قرار دادند و این امر به کامپیوتر امکان انجام محاسبات پیچیده‌تر و سریع‌تر را می‌داد.
- کامپیوترهای این دوره (و دوره‌های بعدی) از معماری مجموعه دستورات (Instruction Set Architecture) استفاده می‌کردند.
- معماری کامپیوتر IBM 360 از معماری‌های مشهور این دوره است.

نسل چهارم کامپیوترها

- در دهه ۱۹۷۰، ریزپردازنده‌ها به جای ICها استفاده شدند.
- این باعث افزایش قابلیت انعطاف پذیری و کاهش هزینه‌ی ساخت کامپیوترها شد.
- در این دوره معماری کامپیوترها شخصی و کامپیوترهای قابل حمل توسعه یافت.

نسل پنجم کامپیوترها

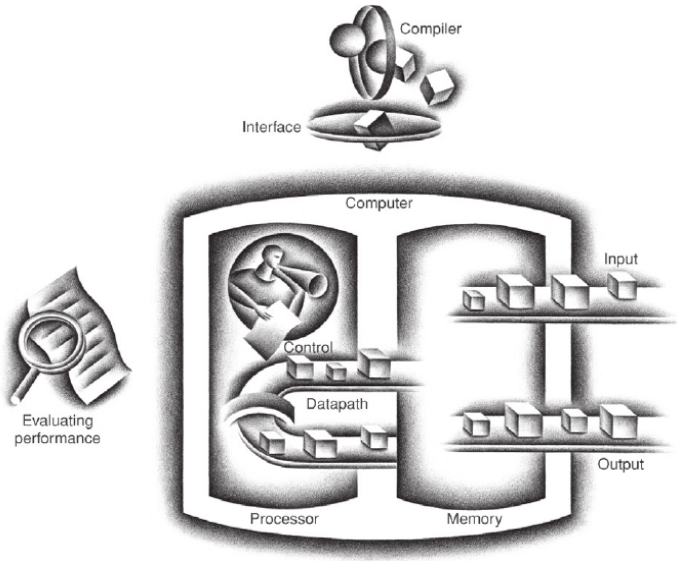
- در دوره‌ی نسل پنجم کامپیوترها که از دهه‌ی ۱۹۸۰ شروع شد، تحولات مهمی در معماری کامپیوتر رخ داد.
- در این دوره کامپیوترهای موازی که قدرت پردازش با با استفاده از چندین واحد پردازشگر به صورت همزمان را داشتند، طراحی و ساخته شدند.
- کامپیوترهای برداری، یکی دیگر از پیشرفت‌ها این دوره بود. این کامپیوترها مجهز به پردازنده‌هایی بودند که مخصوص انجام عملیات به روی بردارها و ماتریس‌ها بودند و برای برنامه‌های علمی و مهندسی که با این داده‌ها سر و کار داشتند بسیار مناسب بودند.
- در این دوره استفاده از ICهای فوق بزرگ (VLSI) و ICهای فوق فوق بزرگ (ULSI) نیز رایج شد.
- در این دوره شاهد ظهور کامپیوترهای شخصی (Personal Computer) و سیستم‌های توزیع شده (Distributed Systems) هستیم.

اجزا

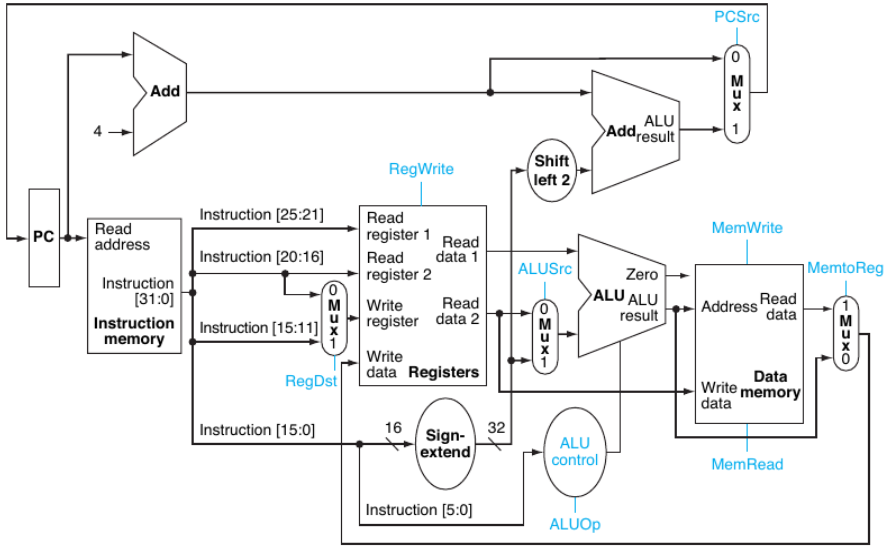
- در سطح بالا CPU از دو قسمت اصلی تشکیل شده که خود به قسمت‌های دیگری تقسیم می‌شوند:
  ۱. Data Path
  این قسمت عملیات‌های ریاضی و محاسبات را انجام می‌دهد.
  Data Path از قسمت‌هایی همچون:
  ۱.۱ Register File
  ۲.۱ ALU
  ۳.۱ چندین Multiplexer و واحدهای جمع و extend تشکیل شده است.

  ۲. Control Unit
  واحد کنترل پردازنده، به data path، مموری و دستگاه‌های I/O دستورات لازم برای اینکه چه کاری را باید انجام دهند، می‌دهد.

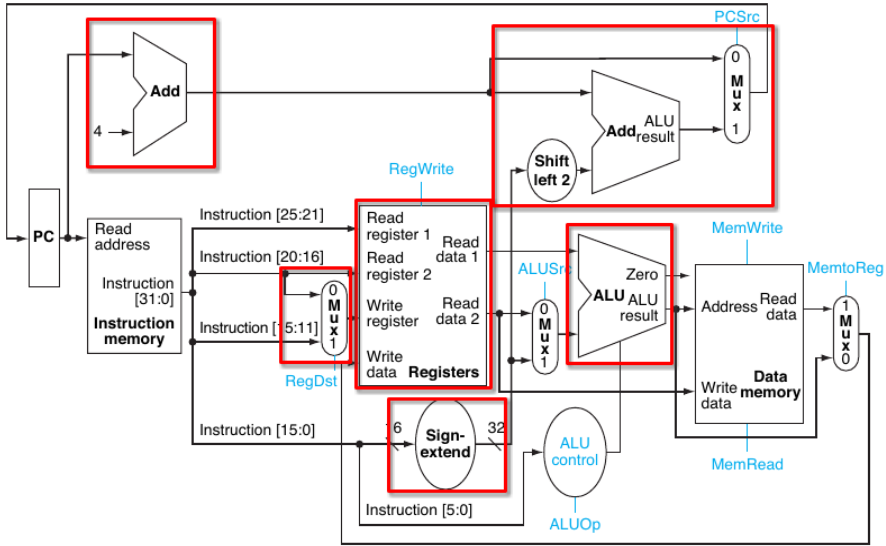
CPU از نگاه بالا



Data Path


PCSrc
Add
4
M
u
x
0
1
Add
ALU result
Shift left 2
RegWrite
PC
Read address
Instruction [31:0]
Instruction memory
Instruction [25:21]
Instruction [20:16]
Instruction [15:11]
Instruction [15:0]
Instruction [5:0]
Read register 1
Read register 2
Write register
Write data
Read data 1
Read data 2
Registers
RegDst
ALUSrc
Zero
ALU
ALU result
MemWrite
MemtoReg
Address
Read data
Write data
Data memory
MemRead
16
Sign-extend
32
ALU control
ALUOp

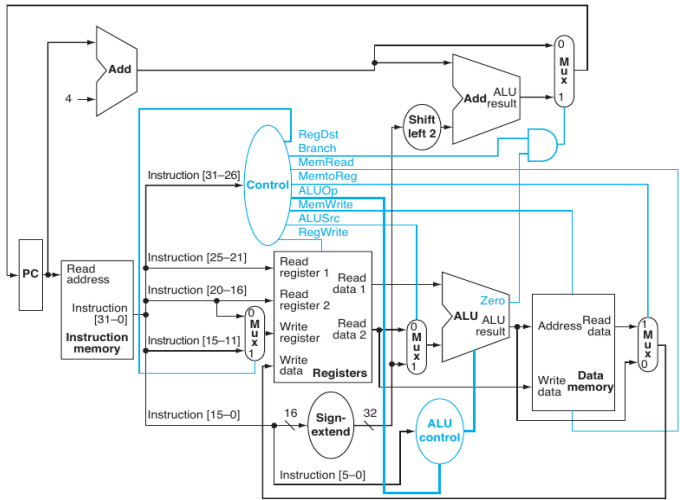
قسمت‌های Data Path



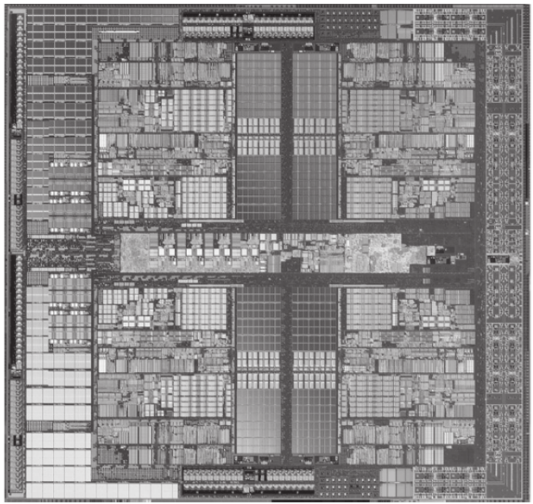
Control Unit

- وقتی برای یک معماری، Instruction Set نوشته می‌شود، به این معنی‌ست که هر یک از instruction ها یک معنی می‌دهد، یک سری ریجستر خاص را نیاز دارد و باید از مسیر متفاوتی از داخل data path رد بشود،
- کدگشایی و کنترل کردن مسیر گذر یک instruction و داده‌هایش به عهده‌ی Control Unit است.

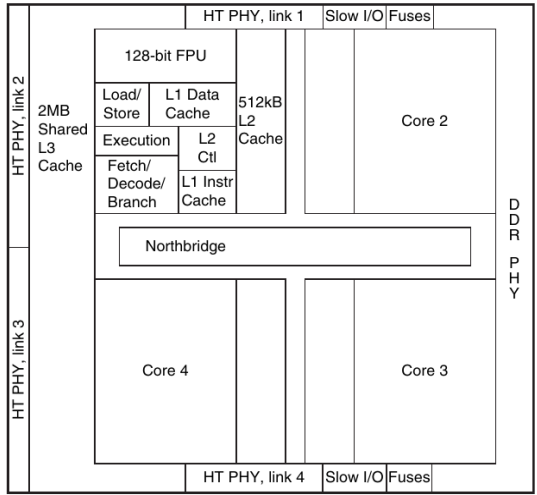
Data Path and Control Unit



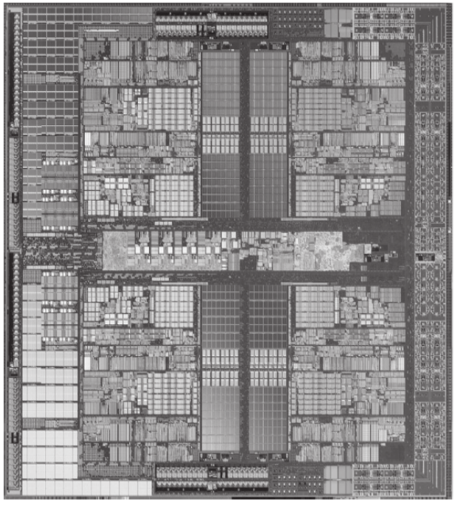
AMD Barceloca Microprocessor

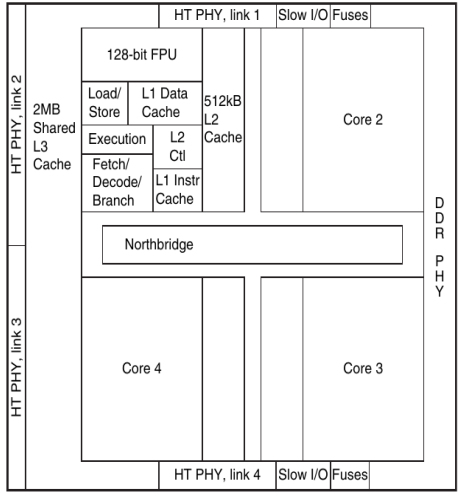


AMD Barceloca Microprocessor Sketch



AMD Barceloca Microprocessor





معماری‌های مختلف

مقدمه

- اگر شما الان بخواهید یک کامپیوتر بخرید، از بین معماری‌های مختلف دو معماری اصلی پیش روی شما هستند:
  ۱. x86
  ۲. ARM

معماری x86

- این معماری بر پایه‌ی معماری Inter 8008 که در سال ۱۹۷۲ معرفی شد، است.
- در واقع کدهایی که برای این معماری نوشته شده‌اند را می‌توان برای آخرین CPUهای Intel یا AMD اسمبل و اجرا کرد.
- پس از Intel 8008 معماری‌های Intel 8088، 16-bit 8086 و سپس 80186، 80286 و... معرفی شدند و در کل x86 نام گرفتند.
- پردازنده‌های شرکت‌های Intel و AMD همگی بر پایه این معماری هستند.

معماری ARM

- نام این معماری برگرفته از **A**dvanced **R**ISC **M**achines که قبل‌تر از **A**corn **R**ISC **M**achine گرفته شده بود، است.
- پردازنده‌های آرم، بخاطر
  - قیمت ارزان،
  - مصرف انرژی کم و
  - تولید گرمای کم

  برای دستگاه‌های سبک و دارای باتری، مثل تلفن‌های هوشمند و لپ‌تاپ‌ها بسیار مناسب هستند.
- بین سال‌های ۲۰۲۰ تا ۲۰۲۲ سریع‌ترین سوپرکامپیوتر دنیا (Fugako) هم از پردازنده‌های معماری آرم استفاده می‌کرد.
- چیپ‌های سری M شرکت اپل هم از معماری آرم استفاده می‌کنند.

معماری کامپیوتر در آینده

آینده‌ی معماری کامپیوتر ← رفع نیازهای جدید

- هوش مصنوعی و یادگیری ماشین
- پردازش‌های داده‌های کلان
- محاسبات کوانتومی
- محاسبات نورومورفیک

معماری کوانتومی

Qubits ـکیوبیت بجای بیت‌های صفر و یک.
استفاده از گیت‌ها و ساختارهای جدید بر اساس کیوبیت‌ها

نتایج:

- سرعت بالا در حل مسائل پیچیده
- توانایی محاسبات مختلف به صورت موازی و همزمان
- دسترسی به رمزنگاری و امنیت بالاتر
- فناوری‌های نورومورفیک

هوش مصنوعی و معماری کامپیوتر

مقدمه

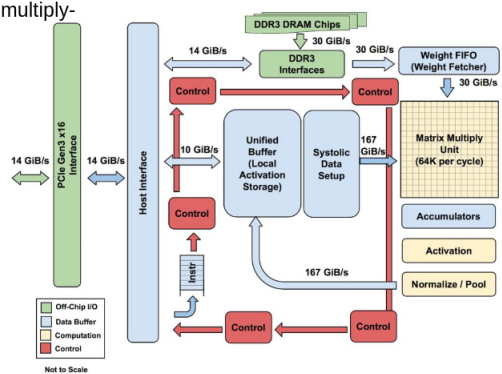
- در دهه‌های ۸۰ و ۹۰ میلادی، کامپیوترها هر ۱۸ تا ۲۴ ماه (قانون مور) سریع‌تر می‌شدند.
- این یعنی اگر شما امسال یک کامپیوتر می‌خریدید و دوستان شما یک سال بعد از شما کامپیوتر جدیدی می‌خریدند، کامپیوتر آنها بسیار سریع‌تر می‌بود
- اما امروزه، تنها راه پیشرفت در معماری کامپیوتر ساخت سخت‌افزار برای یک کاربرد خاص است.
- برای مثال پردازنده‌ها گرافیکی (GPU) برای انجام محاسبات گرافیکی بسیار کارامد هستند. آنها می‌توانند میلیون‌ها ضرب ماتریس در یک هر ثانیه انجام بدهند.

پردازنده‌ی TPU

- با رخ دادن انقلابی در هوش مصنوعی به نام **یادگیری ماشین** که به ضرب ماتریسی متکی بود، نیاز به پردازنده‌های مخصوص ضرب تانسورها برای اجرا سریع‌تر و دقیق‌تر الگوریتم‌ها یادگیری ماشین احساس می‌شد،
- واحد پردازش تانسور (Tensor Processing Unit (TPU)) شتاب‌دهنده‌ی یادگیری ماشین است که توسط گوگل طراحی شده است. TPUها برای ضرب ماتریس‌ها بسیار کارامند هستند که برای آموزش شبکه‌های عصبی ANN ضروریست.

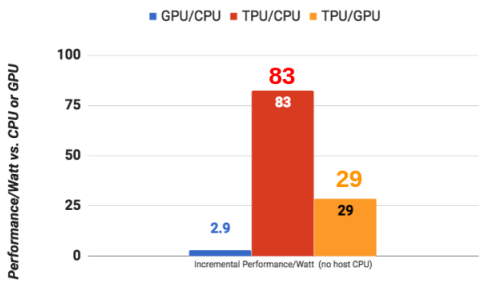
TPU: High-level Chip Architecture

- The Matrix Unit: 65,536 (256x256) 8-bit multiply-accumulate units
- 700 MHz clock rate
- Peak: 92T operations/second
  - 65,536 * 2 * 700M
- >25X as many MACs vs GPU
- >100X as many MACs vs CPU
- 4 MiB of on-chip Accumulator memory + 24 MiB of on-chip Unified Buffer (activation memory)
- 3.5X as much on-chip memory vs GPU
- 8 GiB of off-chip weight DRAM memory



Perf/Watt TPU vs CPU & GPU

Using production applications vs contemporary CPU and GPU



Measure performance of Machine Learning?

See MLPerf.org (“SPEC for ML”)

- Benchmark suite being developed by 23 companies and 7 universities
- 1st Results November 2018