



PV109: Historie a vývojové trendy ve VT

Vývojové trendy

Luděk Matyska

podzim 2022

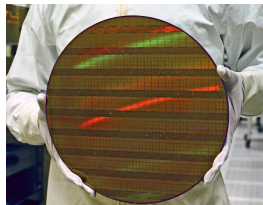


Vývojové trendy ve výpočetní technice

- Procesory
- Operační paměti
- Internet
- Superpočítače

Processory

- Pojem *procesor* je používán v počítačovém průmyslu již od 60. let 20. století
- „Mozek počítače“
- Zpracování sledu instrukcí programů
- Provádí aritmetické a logické operace spolu s operacemi vstupu a výstupu



Zdroj: <http://intel.com>

Terminologie

- **Instrukce** – posloupnost bitů reprezentujících příkaz pro provedení jedné atomické operace
- **Program** – posloupnost instrukcí
- **Takt procesoru** – frekvence krystalového oscilátoru
- **Délka slova** – vyjadřuje počet bitů, který je procesor schopen zpracovat v jednom kroku
- **Počet tranzistorů** – udává, kolik tranzistorů je na procesoru; míra složitosti procesoru
- **Výrobní technologie** – značí zpravidla velikost nejmenší součástky, kterou je možné vyrobit; jednotky μm , nm

Mooreův zákon

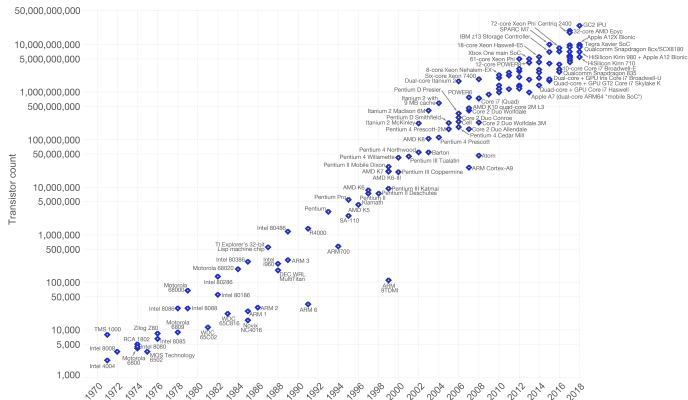
- Autorem je Gordon Moore, spoluzakladatel a bývalý ředitel Intel Corp.
- „Počet tranzistorů, které mohou být umístěny na integrovaný okruh se při zachování stejné ceny zhruba každých 18 měsíců zdvojnásobí“.
- Jedná se spíše o empirické pravidlo, vyslovené roku 1965.
- Předpokládá se, že tento trend bude pokračovat minimálně do roku 2015, možná i déle.
- Další parametry korespondující s Mooreovým zákonem: výkon procesoru, kapacita pamětí, počet a velikost pixelů v digitálních fotoaparátech, ...

Mooreův zákon pro procesory a paměti

Moore's Law – The number of transistors on integrated circuit chips (1971-2018)

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are linked to Moore's law.

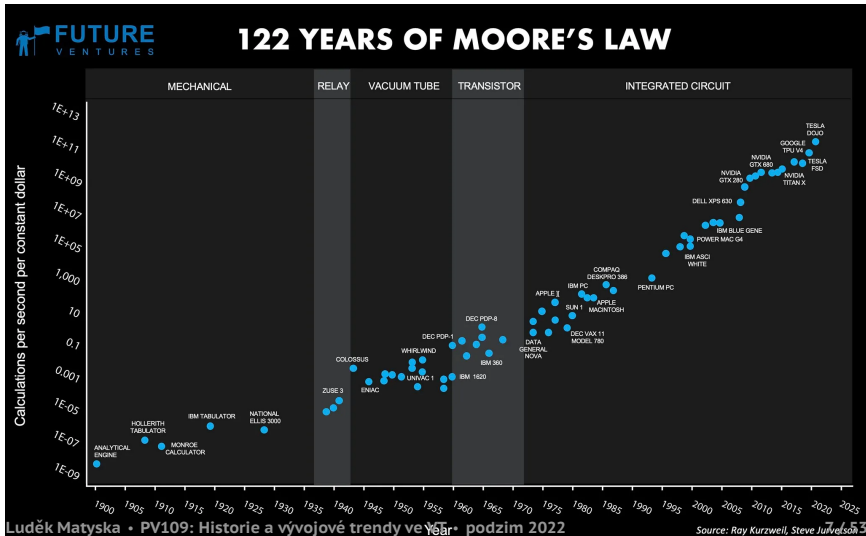
OurWorld
in Data



Data source: Wikipedia (https://en.wikipedia.org/wiki/Transistor_count)
The data visualization is available at OurWorldInData.org. There you find more visualizations and research on this topic.

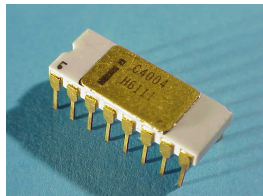
Licensed under CC-BY-SA by the author Max Roser.

Mooreův zákon ještě jednou



Intel 4004

- První jednočipový mikroprocesor
- Uveden 15. listopadu 1971
- Frekvence 740 kHz, 0,07 MIPS
- Instrukční sada čítala 46 instrukcí
- 2300 tranzistorů vyrobených 10 μ m technologií
- Adresovatelná paměť 640 bytů
- Šířka sběrnice 4b (multiplex adresová/datová kvůli malému počtu pinů)
- Původně určená pro kalkulačtor Busicom 141-PF



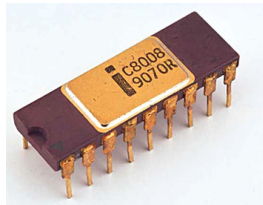
Zdroj:
<http://computermuseum.li>



Busicom 141-PF

Intel 8008

- Uveden 1. dubna 1972
- První 8bitový procesor
- Frekvence 800 kHz
- Instrukční sada čítala 48 instrukcí
- 3500 tranzistorů vyrobených 10 μ m technologií
- Adresovatelná paměť 16 KB
- Určen pro mikropočítač Datapoint 2200



Zdroj:
<http://old-computers.com>



Zdroj: <http://history-computer.com>

Motorola MC6800

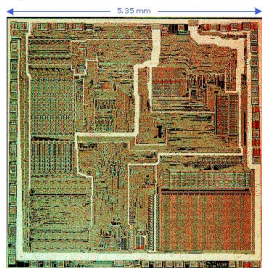
- 8bitový procesor představený v roce 1974
- Snadnější integrace než u Intel 8080 – nevyžadoval další podpůrné čipy na základní desce
- 16bitová adresová sběrnice, 8bitová (obousměrná) datová sběrnice
- Instrukční sada čítala 72 instrukcí
- Frekvence 1 MHz, poslední generace až 2 MHz
- Uplatnění vedle osobních počítačů (např. SWTPC 6800, Tektronix 4051) také v průmyslu



Zdroj:
<http://en.wikipedia.org>

Intel 8086

- Uveden 8. června 1978, první procesor architektury x86
- Frekvence 4,77 – 10 MHz
- 16bitová datová sběrnice, 20bitová adresová
- 29 000 tranzistorů vyrobených $3\mu\text{m}$ technologií
- Adresovatelná paměť až 1 MB
- Používán v přenosných počítačích (např. Compaq Portable) a v IBM PS/2

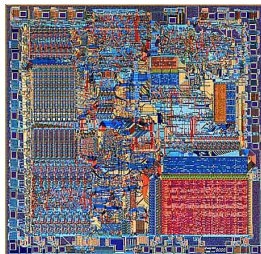


8086 internal structure

Zdroj:
<http://old-computers.com>

Intel 8088 – mozek IBM-PC

- Uveden 1. června 1979
- Zpětně kompatibilní s 8086
- Vnitřně 16bitová architektura, navenek pouze 8bitová sběrnice
- 20bitová adresová sběrnice (stejně jako 8088)
- Velmi úspěšný, zejména kvůli IBM-PC

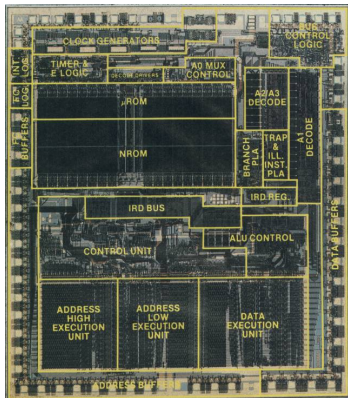


Zdroj:

<http://micro.magnet.fsu.edu>

Motorola 68000

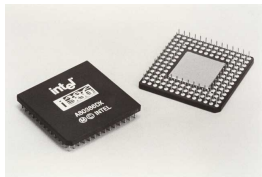
- 32bitový CISC mikroprocesor uvedený v roce 1979
- Takt procesoru postupně 4–16,67 MHz
- Vnitřní 32bitová adresová sběrnice, vnější 16bitová
- Umožňoval adresovat až 16 MB paměti
- Používán ve víceuživatelských mikropočítačích (např. HP9000, systémy SUN Microsystems), po snížení cen ve 2. polovině 80. let i v osobních počítačích (např. Apple Macintosh, Commodore Amiga)
- Architektura se používá při návrhu nových procesorů dodnes



Zdroj: <http://www.hacking-cult.org/>

Intel 80386

- Uveden 17. října 1985
- Taktovací frekvence 16 – 33 MHz
- Plně 32bitový procesor (datová i adresová sběrnice)
- Adresovatelná paměť až 4 GB
- Virtuální paměť 64 TB
- 275 000 tranzistorů (1 μ m technologie)
- Zpětně kompatibilní s x86 (16bit)
- Různé varianty: i386 SX/CX/DX/SL



Zdroj: <http://pipux.net>

Sun SPARC

- Scalable Processor ARChitecture
- 32bitová RISC architektura navržená firmou Sun Microsystems, uvedená v roce 1986
- V následujících letech se dočkala několika revizí
- Takt procesoru 14,28 – 40 MHz, výrobní technologie $1,3\mu m$,
- Až 128 KB L1 cache
- Výkon 10 MIPS, resp. 1,6 MFLOPS
- Vzniklo několik open source implementací SPARC
- Koncept dopředné binární kompatibility

IBM POWER 1

- Performance Optimization With Enhanced RISC
- Více-čipový procesor vyvinutý v IBM a uveden v roce 1990
- Procesory s frekvencí 20, 25 nebo 30 MHz
- Modulární návrh
 - ICU – instruction cache unit
 - FXU – fixed-point unit
 - FPU – floating-point unit
 - DCU – data-cache unit (několik)
 - SCU – storage-control unit
 - I/O unit
- Určen primárně pro servery a výkonné pracovní stanice

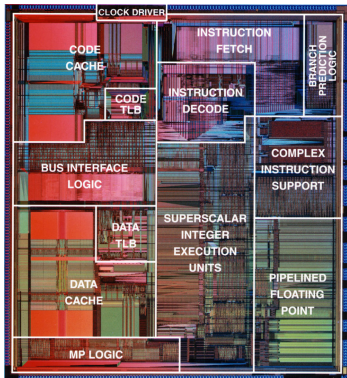


Zdroj: <http://en.wikipedia.org>

//en.wikipedia.org

Intel Pentium

- 5. generace procesorů Intel (P5), uveden 22. 3. 1993
- Označení názvem kvůli patentům (nelze patentovat číslo)
- Frekvence 60–66 MHz, 16 KB L1 cache
- 64bitová datová a 32bitová adresová sběrnice
- 4 GB adresovatelné paměti, až 64 TB virtuální
- 3,1 milionu tranzistorů vyrobených 0,8 μ m technologií
- Superskalární architektura
- PGA pouzdro (Pin Grid Array), 273 pinů



Zdroj: <http://computerhistory.org>

Pentium MMX – Multi-Media eXtension

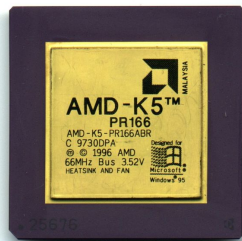
- Uveden 8. ledna 1997
- Frekvence sběrnice 66 MHz, 32 KB L1 cache
- 296/321pinové PGA pouzdro
- 4,5 milionu tranzistorů ($0,35\mu m$)
- Podpora pro multimédia **Intel MMX** – SIMD instrukční sada



Zdroj:
<http://it.wikipedia.org>

AMD K5

- První x86 procesor vyvinutý zcela v AMD
- Hlavní konkurent procesorové řadě Intel Pentium
- Představen se zpožděním v roce 1995
- Takt procesoru 75 – 133 MHz
- 4,3 milionu tranzistorů
- 5 celočíselných jednotek, 1 pro práci s čísly s pohyblivou desetinnou čárkou
- Označení **PR** (Pentium Rate) pro srovnání s procesory Pentium (např. AMD K5 PR166 běžel na frekvenci 116 MHz)



Zdroj: <http://pipux.net>

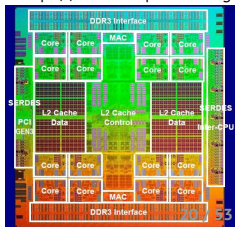
Sun UltraSPARC

- Představen v roce 1995
- Procesor implementuje architekturu SPARC V9
- Takt procesoru 143–200 MHz
- 3,8 milionu tranzistorů, 0,47 μ m technologie
- Superskalární procesor s *in-order* vykonáváním instrukcí
- 2 \times L1 cache o velikosti 16 KB
- Volitelná L2 cache 512 KB – 4 MB
- Další procesor architektury SPARC V9 – Fujitsu SPARC64



Zdroj:

<http://en.wikipedia.org>

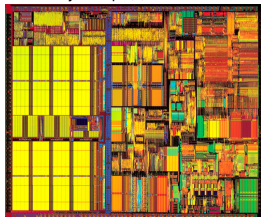


Procesory 6. generace (P6)

- Pentium Pro (1995)
- Pentium II (1997)
- Pentium II Xeon (1998) – serverová řada
- Pentium III (1999)
 - Zavedení L2 cache (512 KB)
 - Takt procesoru až 600 MHz
 - Internet Streaming SIMD instrukce
 - 9,5 milionu tranzistorů (0,25 μ m)



Zdroj: <http://intel.com>



Zdroj: <http://www.tayloredge.com/museum/>

Další Intel

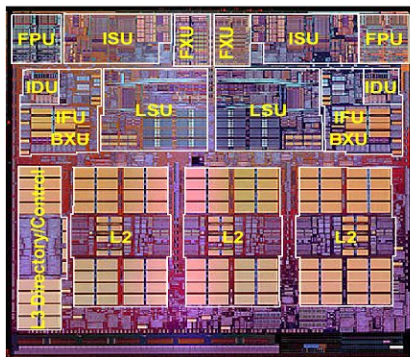
- 2000: Pentium 4 a architektura Netburst
- 2005: Pentium D a dvě jádra
- Nehalem: i7 core (4-12 MB level 3 cache)
- Westmere: Grafika v CPU
- Sandy Bridge: zkrácená pipeline (14–19 kroků)
- Ivy Bridge: energie začíná být cílem
- Haswell: nižší spotřeba, vyšší výkon, GPU on a chip
- 2014: Broadwell, mobilní procesory
- Skylake: DDR4 paměť, lepší grafika

Další procesory řady IBM POWER

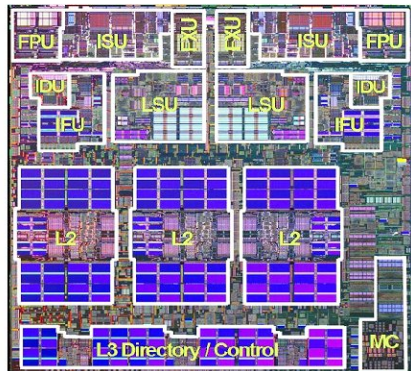
- PowerPC (1993)
 - Upravená RISC architektura, vyvinut ve spolupráci s Apple a Motorola
 - Určen pro osobní počítače, 32bitová architektura později rozšířená na 64bitovou
- POWER3 (1998)
 - Plně 64bitový procesor
 - 15 milionů tranzistorů, zdvojená L1 cache navýšena na 64 KB
- POWER5 (2004)
 - Dvujádrový procesor s podporou SMT (2 vlákna/jádro)
- Současná verze POWER9

Srovnání čipů POWER4 a POWER5

POWER4 Core

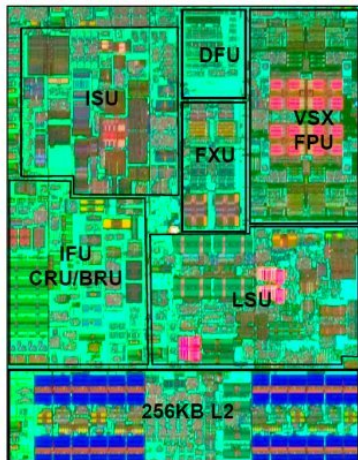


POWER5 Core

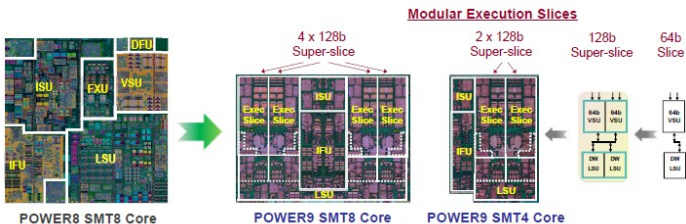


Zdroj: <http://ibm.com>

POWER7



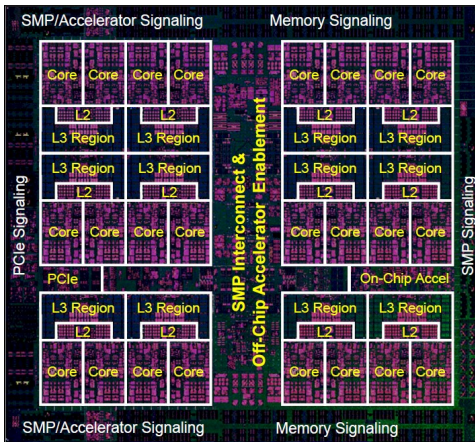
POWER9



Re-factored Core Provides Improved Efficiency & Workload Alignment

- Enhanced pipeline efficiency with modular execution and intelligent pipeline control
- Increased pipeline utilization with symmetric data-type engines: Fixed, Float, 128b, SIMD
- Shared compute resource optimizes data-type interchange

POWER9



Viz <https://www.extremetech.com/computing/266031-ibms-power9-dent-x86-server-market-oems-prep-new-systems-emphasize-gpu-compute>

POWER9

New Core Microarchitecture

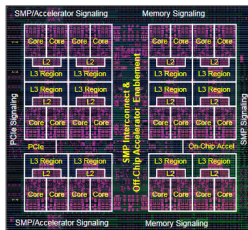
- Stronger thread performance
- Efficient agile pipeline
- POWER ISA v3.0

Enhanced Cache Hierarchy

- 120MB NUCA L3 architecture
- 12 x 20-way associative regions
- Advanced replacement policies
- Fed by 7 TB/s on-chip bandwidth

Cloud + Virtualization Innovation

- Quality of service assists
- New interrupt architecture
- Workload optimized frequency
- Hardware enforced trusted execution



14nm finFET Semiconductor Process

- Improved device performance and reduced energy
- 17 layer metal stack and eDRAM
- 8.0 billion transistors

Leadership

Hardware Acceleration Platform

- Enhanced on-chip acceleration
- Nvidia NVLink 2.0: High bandwidth and advanced new features (25G)
- CAPI 2.0: Coherent accelerator and storage attach (PCIe G4)
- New CAPI: Improved latency and bandwidth, open interface (25G)

State of the Art I/O Subsystem

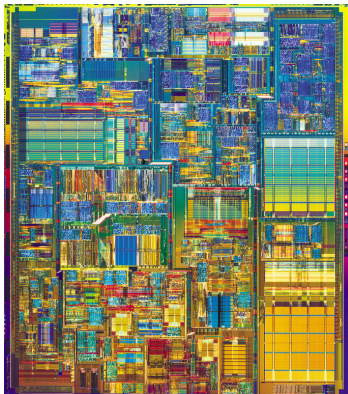
- PCIe Gen4 – 48 lanes

High Bandwidth Signaling Technology

- 16 Gb/s interface
 - Local SMP
- 25 Gb/s Common Link interface
 - Accelerator, remote SMP

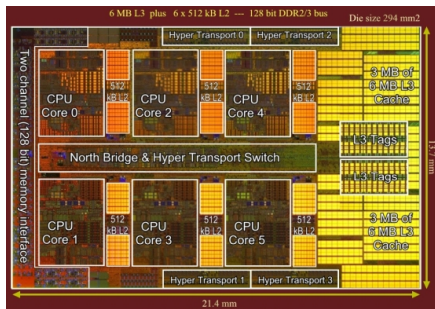
Intel Pentium 4 – 7. generace (P7)

- 2000–2008 – procesory Intel Pentium 4
- Takt procesoru až 3,8 GHz, L2 cache až 512 KB
- Instrukce SSE2 SIMD
- 42 milionu tranzistorů ($0,18 \mu\text{m}$)
- 2004: Přejechod na 64bitovou architekturu (x86→x86-64)



64bitové procesory rodiny AMD

- Jádro AMD K8
 - Athlon 64 X2 (2005) – první dvoujádrový 64bitový, plně x86 kompatibilní procesor firmy AMD určený pro osobní počítače
 - Další větve – Opteron (serverové), Athlon 64 FX
- Jádro AMD K10
 - Jádro bylo představeno v roce 2007
 - Vychází z AMD K8
 - Hlavní zástupci – AMD Opteron, AMD Phenom

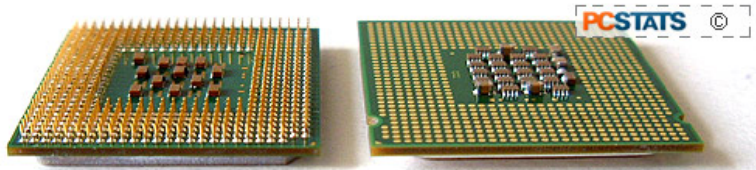


AMD Opteron (jádro Istanbul)

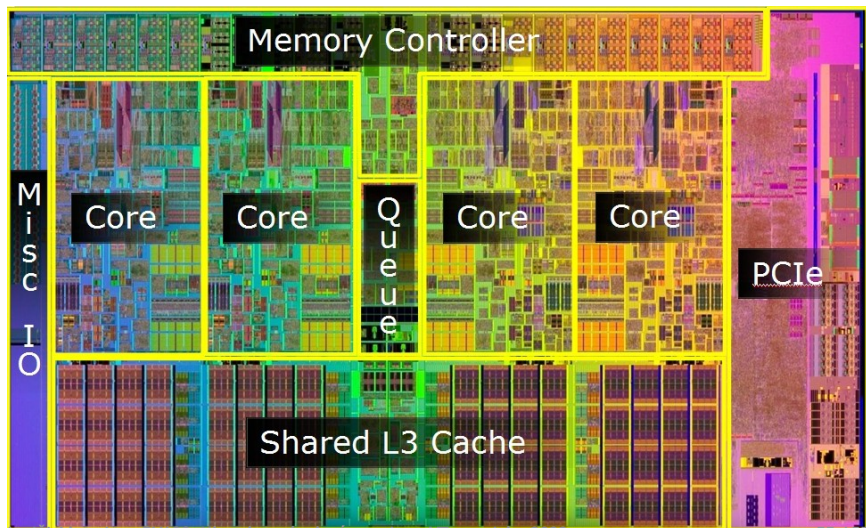
Zdroj: <http://www.generation-gpu.fr>

Procesory Intel Core

- Dvě produktové řady
 - Core 2 – Core 2 Duo, Core 2 Quad, Core 2 Extreme
 - Core i – Core i3, Core i5, core i7, Core i7 Extreme
- Čistě 64bitová architektura
- SSSE3 SIMD instrukce (4. generace)
- 2 a více jader v jednom pouzdře procesoru
- Další zvětšování L2 paměti – 2–12 MB
- LGA pouzdro (Land Grid Array) – piny jsou v patici na základní desce



Intel Core i7 „Nehalem”



Používané technologie

Hyper-Threading

- Vylepšení paralelizace pomocí virtualizace
- Každé fyzické jádro je reprezentováno v OS jako dvě virtuální
- Jedno jádro tak může zpracovávat dvě vlákna zároveň

Turbo-Boost

- Dynamické zvýšení výkonu na žádost (tzv. dynamické přetaktování)
- Při běžné práci procesor běží na nižší výkon (úspora energie)
- Implementováno v procesorech Nehalem, Sandy-Bridge a Ivy-Bridge

Současné trendy ve vývoji procesorů

- Takt procesoru stagnuje kolem 4 GHz
- Výrobní technologie 32 nm umožňuje uspořít prostor i výkon
- Zavádění L3 cache; až 12 MB
- Až 16 jader na jednom čipu (AMD Opteron 6200)
- Řadič paměti integrován na čipu
- Budoucnost – *manycore* architektury ($\times 100$ -1000 jader), např. Intel MIC¹

¹Intel Many Integrated Core Architecture

Výpočty na grafických akcelerátorech

- od 80. let – vývoj grafických akceleratorů, motivováno herním průmyslem
- 90. léta – DirectX, minimální možnosti programování; 3dfx Voodoo, Nvidia GeForce 256
- 2000–2006 – OpenGL, programování vertex/pixel shader jednotek (jazyk Cg: operace v plovoucí čárce, smyčky), položeny základy pro GPGPU; ATI Radeon 9700, Nvidia GeForce 3
- od 2007 – CUDA, OpenCL, období GP-GPU² – SIMT model, desítky multiprocessorů ($= \times 1000$ jader); řada GeForce 8 = první GP-GPU karty.
- Další vývoj – kombinace CPU a GPU = hybridní procesory, výzva = efektivní algoritmy

²General Purpose Computing on GPU

Vývoj operačních pamětí

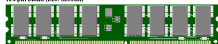
- Hlavními parametry operačních pamětí jsou
 - propustnost (maximalizace)
 - pracovní frekvence (maximalizace)
 - kapacita (zvyšování)
 - přístupová doba (snižování)
 - napájecí napětí
- Synchronní vs. asynchronní paměti
- ECC vs. non-ECC paměti (Error-correcting code)
- JEDEC SSTA – organizace schvalující technologické standardy, mj. právě v oblasti operačních pamětí

From Computer Desktop Encyclopedia
© 2007 The Computer Language Co., Inc.

740-pin DIMM (DDR2 and DDR3 SDRAM)



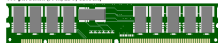
184-pin DIMM (DDR SDRAM)



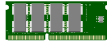
184-pin SDRAM (Rambus) Chips are covered with metal heat sink.



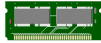
168-pin DIMM (FPM, EDO, SDRAM)



144-pin SODIMM (FPM, EDO, SDRAM)



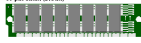
72-pin SODIMM (FPM, EDO)



260-pin SODIMM (DDR SDRAM)



39-pin SIMM (DRAM)



72-pin SIMM (FPM)



Přehled technologií operačních pamětí I.

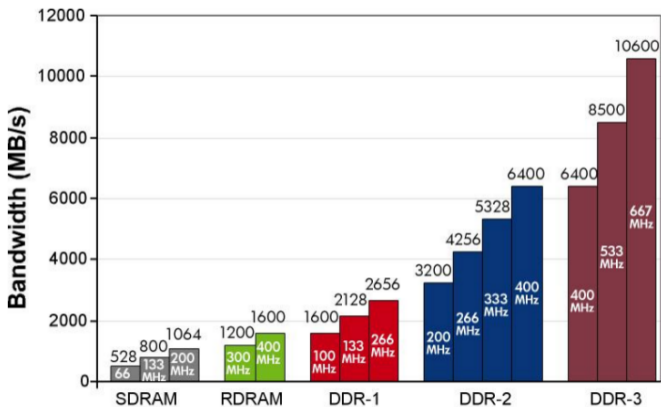
- FPM – Fast Page Mode DRAM (1987)
 - asynchronní paměť
 - přístupová doba 60–80 ns
- EDO – Enhanced Data output DRAM (1995)
 - též označována jako Hyper Page Mode DRAM
 - asi o 5% rychlejší než FPM, přístupová doba cca 60 ns
 - kapacita modulu až 32 MB
- SDRAM – Synchronized Dynamic Random Access Memory (1996)
 - synchronní paměť (pracuje synchronně podle externího taktu)
 - kapacita 64–512 MB, frekvence sběrnice 66–133 MHz
- RDRAM – Direct Rambus DRAM (1999)
 - vysoká propustnost (až 10 GB/s) za cenu vyšších latencí
 - postupně vytlačeny DDR paměťmi

Přehled technologií operačních pamětí II.

- DDR – Double Data Rate SDRAM (2000)
 - efektivní pracovní frekvence až 400 MHz
 - propustnost až 3,2 GB/s
 - kapacita modulu 64 MB až 2 GB
 - 2bit prefetch
- DDR2 SDRAM
 - efektivní pracovní frekvence až 800 MHz
 - propustnost 3,2–6,4 GB/s
 - 4bit prefetch
 - kapacita modulu 128 MB až 8 GB
- DDR3 SDRAM
 - zvýšení rychlosti paměti (až 2133 MHz)
 - nižší spotřeba energie
- GDDR3/4/5 – paměti pro grafické akcelerátory

Srovnání paměti dle šířky pásma

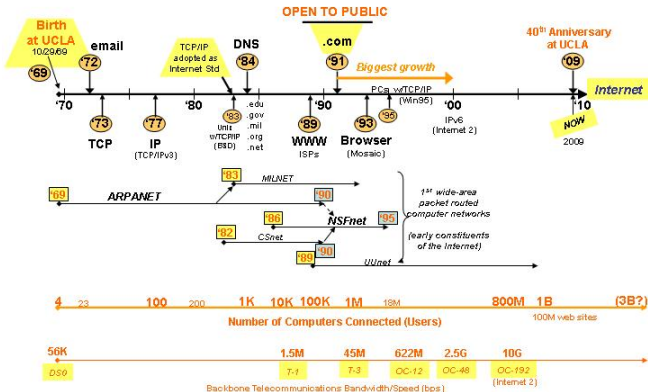
Figure 8. Peak bandwidth comparison of SDRAM and advanced SDRAM technologies



Zdroj: Memory technology evolution: an overview of system memory technologies, Technology brief, 9th edition, Hewlett-Packard Development Company, 2010.

Vývoj Internetu

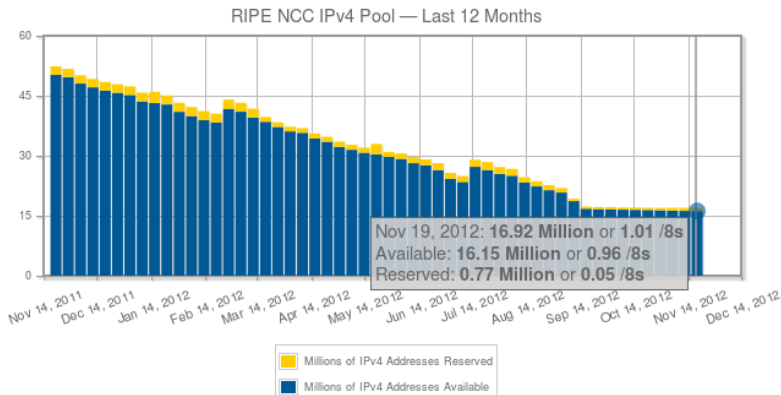
INTERNET TIMELINE



© 2008 Jeffrey H. Drobman

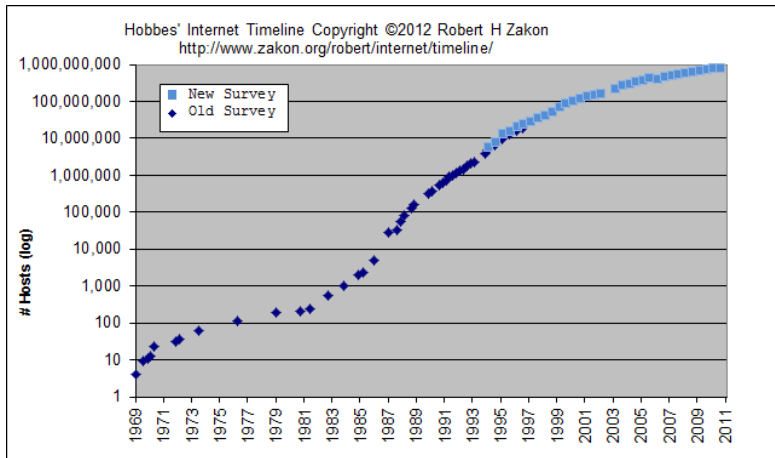
Zdroj: <http://www.drjeffsoftware.com/history.html>

Dostupný počet IPv4 adres



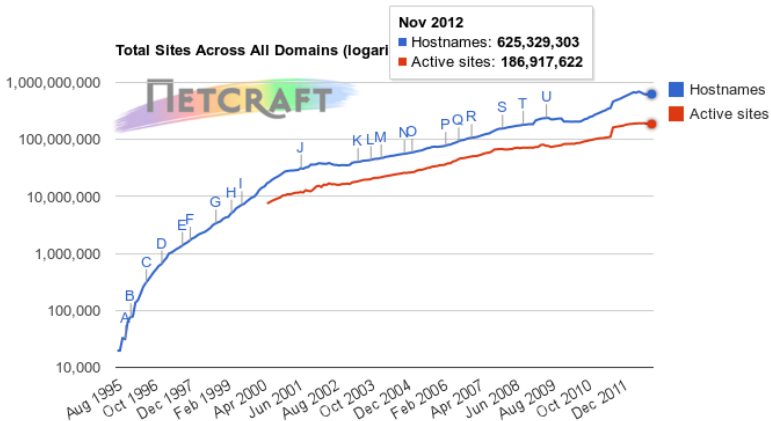
This graph shows the number of available and reserved IPv4 addresses managed by the RIPE NCC over the past 12 months.
 Zdroj: <https://www.ripe.net/internet-coordination/ipv4-exhaustion/ipv4-available-pool-graph>

Vývoj počtu stanic připojených k Internetu



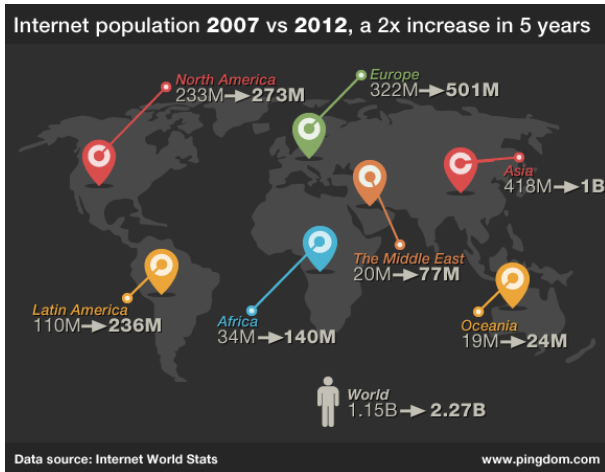
Zdroj: <http://www.zakon.org/robert/internet/timeline/>

Růst počtu webových stránek



Zdroj: <http://www.zakon.org/robert/internet/timeline/>

Internetová populace mezi 2007 a 2012

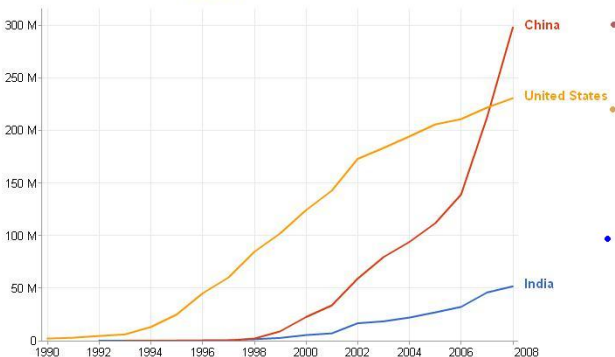


Zdroj: <http://royal.pingdom.com>

Internetoví uživatelé ve vybraných zemích

Internet users

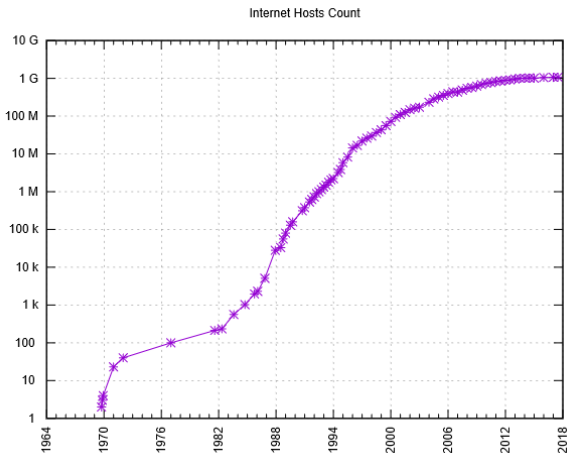
People with access to the Internet. [More info >](#)



Data source: [World Bank, World Development Indicators](#) - Last updated December 13, 2010

Zdroj: <http://royal.pingdom.com>

Internetové uzly

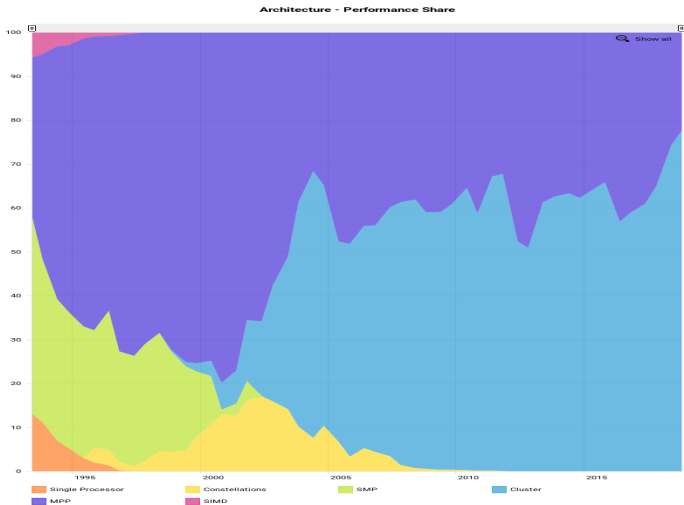


Zdroj: https://commons.wikimedia.org/wiki/File:Internet_Hosts_Count_log.svg

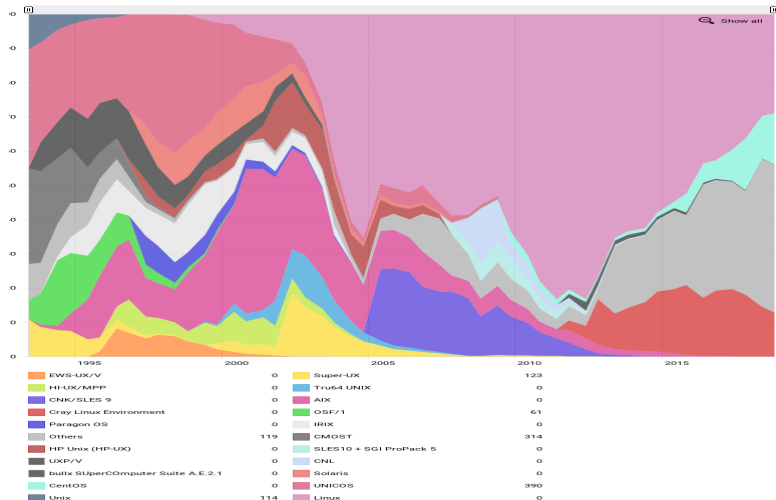
Superpočítače

- Data z TOP500
 - <http://top500.org/statistics/overtime>
 - 2× do roka aktualizované údaje
- Vybrané statistiky:
 - Architektury superpočítačů
 - Používané operační systémy
 - Výrobci
 - Zastoupení zemí v TOP500
- Vývoj v oblasti od roku 1993

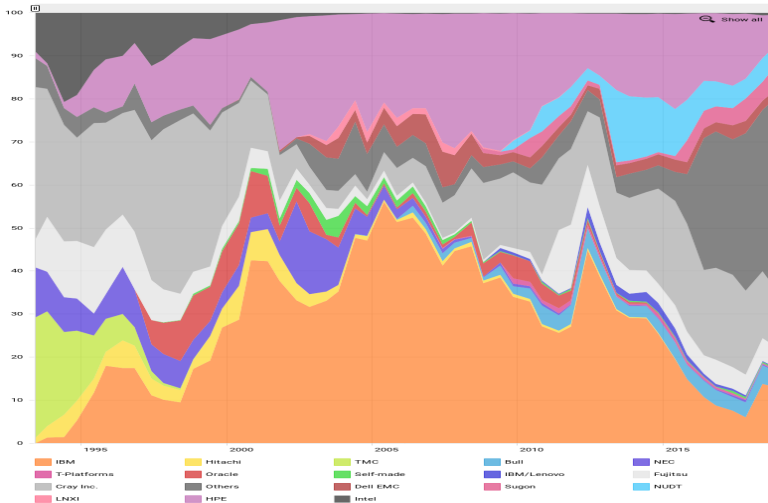
TOP500 – Architektury superpočítačů



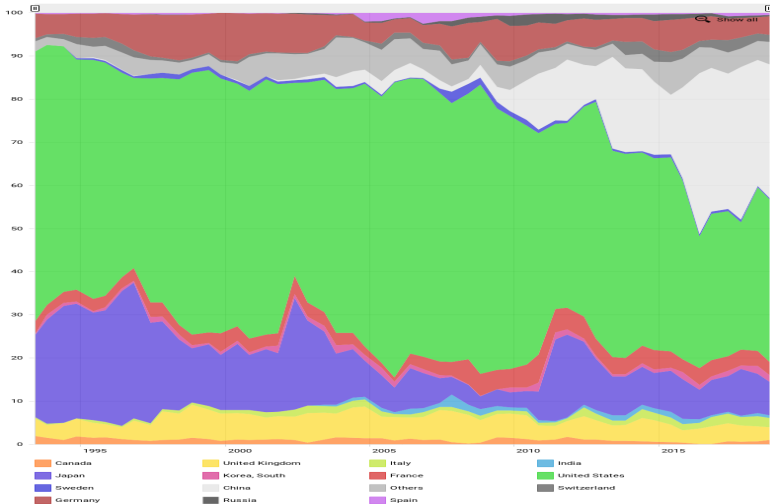
TOP500 – Operační systémy



TOP500 – Výrobci



TOP500 – Zastoupení zemí



Shrnutí

- Stále roste složitost, ale i význam IT systémů
- V mnoha ohledech exponenciální růst se v poslední době „zastavuje“
- Budeme potřebovat nové přístupy
 - Stále rostoucí závislost na software
 - Skutečně masivně paralelní systémy (10^6 až 10^9 jader/vláken)