The AMD and Intel Architectures

COMP311 2008

Jamie Curtis

Intel History

- 4 & 8 bit Micro's introduced in 1971 & 1972
  - 740kHz, 2300 transistors
- First x86 arrives in June, 1978 – the 8086
  - 16bit data, 20bit address
  - 8088 variant used in the IBM PC
    - Identical internal to 8086 with 8bit external bus
  - 29,000 transistors
- 80286 introduced in 1982
  - Still 16bit. Introduced memory protection and protected mode
  - 6 – 12MHz, 134,000 transistors

Intel History cont.

- 80386 introduced in October 1985
  - First 32bit processor
  - Re-worked protection allows paged virtual memory
  - Codenamed P3, called the i386
  - Floating point via the 80387 co-processor
  - 275,000 Transistors
  - 4 GB Addressable memory

Intel History cont.

- 80486 introduced in 1989
  - With the arrival of the 80486DX2 in 1992 for the first time the external bus no longer runs at the CPU frequency
  - L1 Cache now on chip
- Pentium introduced in 1993
  - Intel have one of their first big public recalls, the FDIV bug in early Pentiums
  - Pentium MMX follows in 1997

Intel History cont.

- Pentium Pro introduced in 1995
  - First mainstream processor that translates instructions into RISC like microinstructions before executing them
  - Integrated L1 & L2 cache
  - First product based on the P6 architecture
  - Highly optimised for 32bit code made it a poor choice for the 16bit Windows 3.11

Intel History cont.

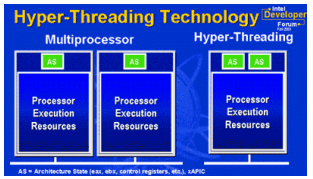
- P6 Core continues into many designs
- Pentium II adds MMX (1997)
  - L2 cache ½ speed external
- Pentium III adds SSE (1999)
  - L2 cache becomes integrated again

Intel NetBurst

- Intel introduces the Pentium 4 in 2000 based on the all new NetBurst (P7) architecture.
  - all about increasing clock speed
    - when released > 10GHz promised
  - To achieve the high clock speeds, a ***very*** deep pipeline is required
    - 20 stages originally, 31 stages in the Prescott core
    - Avoiding stalls requires the “Rapid-Execution-Engine”
      - ALU runs at twice the core frequency
    - Branch prediction becomes important
      - > 80% correctly predicted by the P4

Intel NetBurst

- Deep, wide CPU has problem keeping it’s ALU’s busy
  - “According to Intel, most IA-32 x86 code uses only 35% of the Pentium 4's execution units”
- Hyper-Threading allows one CPU to act like SMP system

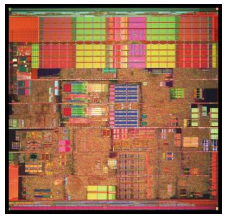


Intel NetBurst

- Works well for traditional “Enterprise” applications that can be parallelised well
- Trouble is, most code is very similar
  - Causes collisions on execution units
- Some code actually runs slower
- NetBurst does allow execution units to be added relatively easily
  - Fits Hyper-Threading

Intel NetBurst

- Well, what went wrong ?
  - **HEAT !**
  - While you can scale frequency with a deeper pipeline, the heat dramatically rises with this rise in frequency.
  - Increasing the supporting logic to allow the pipeline to work effectively also increases transistor count, all creating more heat.
  - Prescotts contain over 125 million transistors



Intel Core Microarchitecture

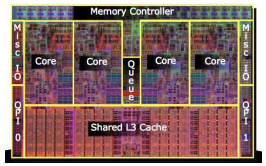
- Pentium M resurrects P6 architecture again in 2003
  - Adds SIMD
  - Adds NetBurst (P7) FSB
- Further development from the P6 leads to the Core architecture.
  - Designed from the ground up, but the design follows the P6 much more than the NetBurst

Intel Core cont.

- Much wider design than previously
  - Pipeline down from 30+ to 14 stages
  - Up to 5 x86 instructions can be decoded per cycle
    - NetBurst could only decode one instruction
- Massive increase in re-ordering ability
  - Also added operation fusing
- L2 cache up to 6Mb
  - 4 smarter prefetch units help keep this utilised

Intel Core i7

- Code named “Nehalem”
  - Being released over the next couple of months
- An incremental update to the Core / Core 2
- 750 million transistors
  - 45nm process

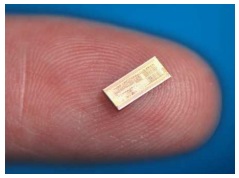


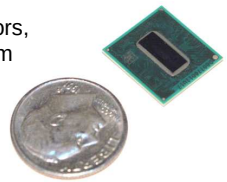
Intel Core i7 cont.

- Focus on power efficiency
  - Each feature added must add 2% performance for every 1% of increased consumption
- Incremental improvements
  - Better loop detection
  - Increased operation fusing
- Sees the return of Hyper Threading
  - Improvements in latency and memory throughput from the P4 gives much better performance gains

Intel Atom

- Originally designed as a small core for dense multi-core designs
  - 10-100 cores per chip
- Much simpler architecture
  - keep power usage as low as possible
- 47 million transistors
  - One of Intels smallest ever processors, die of 25mm^2, package 13mmx14mm





Intel Atom cont.

- First in-order CPU since the Pentium (15 years ago)
  - On a 45nm process the Pentium would be 3mm^2
- Architecture is designed with handhelds in mind
  - Currently still uses to much power mainly due to chipset support
- Future SoC designs may lead to an Atom in your cellphone!

AMD History

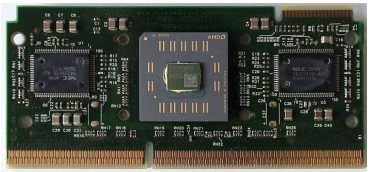
- From 1979 – 1991 AMD was second sourcing Intel processors
  - A requirement to supply to IBM
- Intel attempted to stop this for the 386, so AMD cloned their own version, the Am386
- AMD introduce the K5 in 1995
  - compete with the Pentium
  - ultimately it becomes a failure


AMD History cont.

- The first of 3 K6 variants is introduced in 1997
  - Backward compatible with Intel Pentium motherboards
- K6-III introduces on chip full speed cache, topping Intel's ½ speed external L2 cache in the PII
  - K6-III production is halted due to the demand for the new K7 – Athlon CPU.

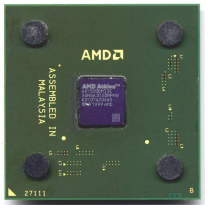
AMD History cont.

- Introduced in 1999, the K7 core was renamed to the Athlon
  - First time AMD required an incompatible motherboard from Intel, although the Athlon was introduced using a CPU SEC cartridge that was mechanically identical to Intel's P-II cartridge.



AMD History cont.

- AMD have used "PR" ratings for their CPU's since the K5.
  - "Performance Ratings" are a way to combat the higher frequency of the Intel products vs the higher IPC of the AMD products.
  - Re-introduced for the Athlon XP because of the much higher P4 clockrate.



AMD K8

- First released as the “Opteron” for the server market in 2003 and later as the “Athlon 64” for the desktop market.
- First 64bit CPU that could run 32bit x86 code without a performance hit
- The K8 is the first x86 CPU to bring the memory controller onto the CPU die
  - Much lower latency
  - Less dependence on chipsets
  - Runs at core speed

AMD K8

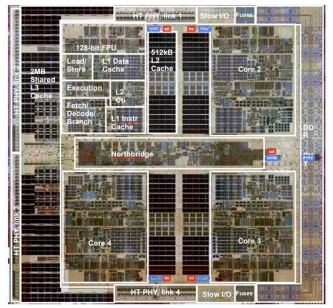
- Based heavily off of the K7 design
  - Better branch prediction allows a slightly deeper pipeline (therefore higher clocks)
    - 10 stages in K7, 12 in K8
  - Increased TLBs
    - Allows better cache performance for large memory
- New FSB system
  - HyperTransport
  - Used to also allow much more scalable SMP

AMD 10h Family (K10)

- Plug in replacements for K8
  - AM2+ for Desktop, Socket F+ for Servers
  - Multi core design
    - Additional shared L3, individual L2 and L1 (split I & D)
- 128 bit FPU units
  - Significant improvement of FP over K8
- Similar enhancements to fetch, decode and prediction logic to the Intel Core

AMD 10h Family cont.

- Known as “Phenom” for the desktop, Opteron for servers
  - First 10h Opteron was codenamed Barcelona
- 463M Transistors



AMD64 + EM64T

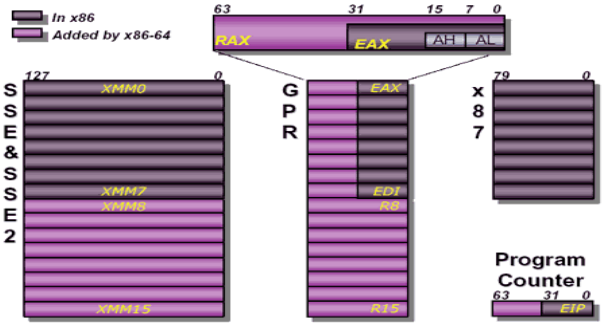
- Designed by AMD
  - Intel focused 64 bit development on Itainum (IA-64).
    - Server focused
  - EM64T reverse engineered when AMD64 became popular for entry level servers and desktops and added to P4
- First major extension to the x86 ISA since i386

x86-64

- i386 has 2 major modes
  - Real mode (8086 emulation)
  - Protected mode (32bit)
- x86-64 bundles the above into *Legacy Mode*
  - Legacy Mode works with all existing code
- Adds *Long Mode*
  - Split into full "*64 bit mode*" and 32bit "*Compatibility Mode*"
  - Requires OS support
  - Processes running in compatibility mode require no changes

x86-64 Enhancements

- ***REGISTERS !***



x86-64 Enhancements

- 64bit addressing allows memory accesses above 4G without nasty hacks.
- "NX" (No Execute) bit on a per page basis.
- Improved support for Position Independent Code (by IP relative addressing modes).
- Adds some extra opcodes to improve common operations.
- Adds virtualisation features.
- Individual processes can be 32 or 64bit