

THE UNIVERSITY *of York*

High Performance Computing - History of the Supercomputer

Prof Matt Probert

<http://www-users.york.ac.uk/~mijp1>

Autumn Term 2022

Overview

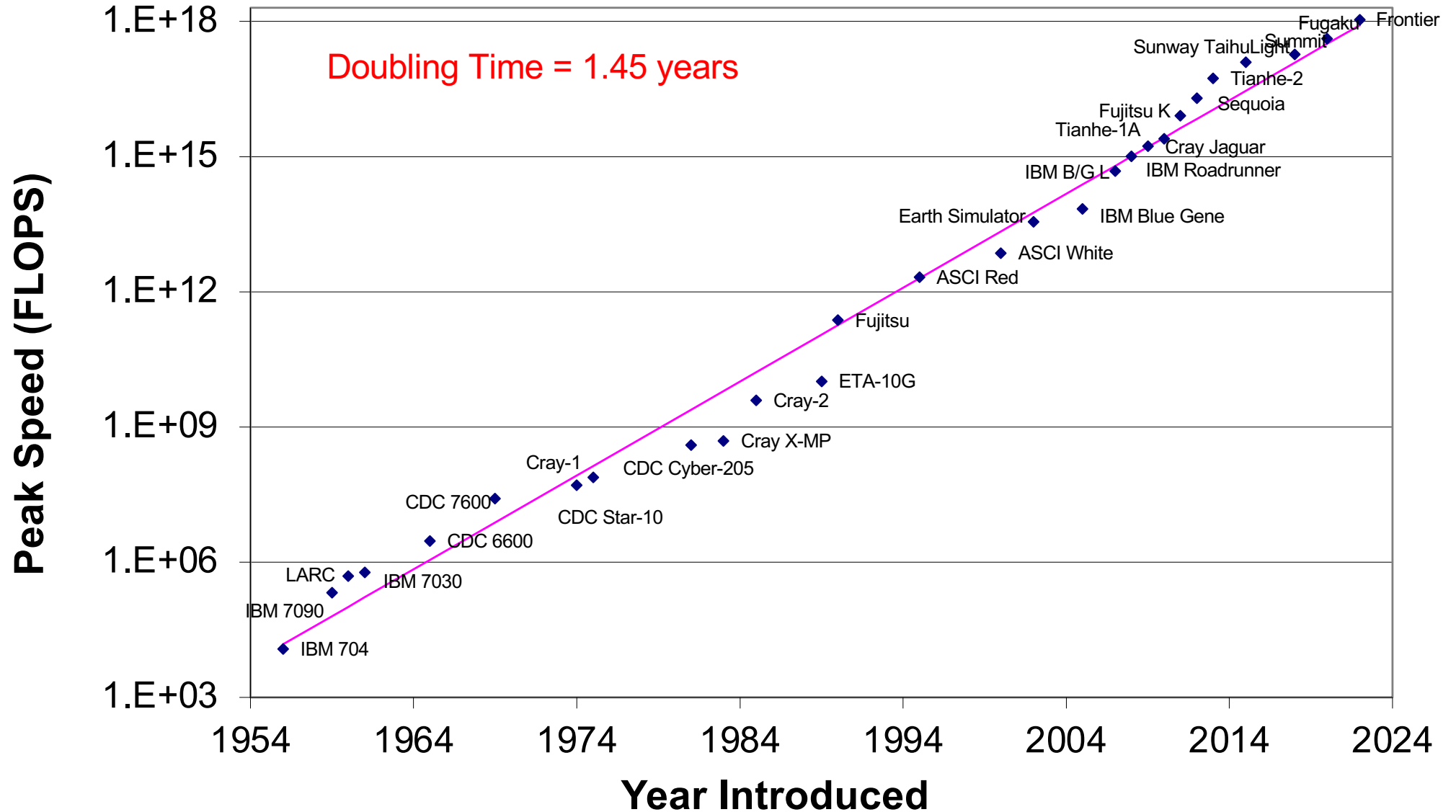
- Early systems with proprietary components, operating systems and tools
- Development of vector computing
- Development of parallel computing
- Development of cluster computing
- Latest situation in world / UK / York

Modern Definitions

Supercomputer

- class of fastest computers currently available
- national research resource or multi-national company with cost ~ £1m+
- MFLOPS
 - 10^6 Floating-point Operations Per Second
- HPC
 - “At least 10 times more compute power than is available on one’s desktop” JISC definition
- Performance typically determined using the LINPACK benchmark and results collated by the top500 organisation

Growth of Supercomputing



Early History (-1969)

Year	Name	Peak speed	Location
1943	Colossus	5000 char/sec	Bletchley Park, England
1945	Manchester Mark I	500 inst/sec	Manchester, England
1950	MIT Whirlwind	20 kIPS	MIT, USA
1956	IBM 704	20 kIPS 12 kFLOPS	
1959	IBM 7090	210 kFLOPS	USAF, USA
1960	LARC	500 kFLOPS (2 CPUs)	Lawrence Livermore Lab. USA
1961	IBM 7030	1.2 MIPS 600 kFLOPS	Los Alamos Lab. USA
1965	CDC 6600	10 MIPS 3 MFLOPS	Lawrence Livermore Lab. USA
1969	CDC 7600	36 MFLOPS	Lawrence Livermore Lab. USA

The Early Days - 1943

- Colossus was the first *programmable, digital, electronic* computer
 - Used *vacuum valves*
 - Designed for a single task – code breaking – for which it was only overtaken by general purpose PC chips in mid-1990s!
 - Top secret => little impact on subsequent computer designs
 - Churchill ordered everything be destroyed at end of WWII

The Early Days - 1945

- Manchester Mark I is significant as it was the first machine to use *index registers* for modifying base addresses
 - Had a 40-bit *accumulator* and 20-bit *address registers* and could perform 40-bit serial arithmetic, with hardware add, subtract and multiply and logical instructions
 - Used two *Williams tubes* (CRTs) for memory – based upon charge measurement at each pixel (lit or unlit). Each tube could store 64 rows of 40 points = 32 “words” = 2 “pages”
 - Also used two *magnetic drums* for permanent storage, each of which could store 64 pages

The Early Days - 1950

- MIT Whirlwind was the first computer that operated in *real time*, used video displays for output and was the first digital flight simulator!
 - All previous computers ran in *batch mode*, i.e. a series of paper tapes/cards were set up as input in advance, fed into the computer to calculate and print results.
 - For simulating an aircraft control panel, Whirlwind need to operate continually on an ever-changing series of inputs => need high-speed stored-program computer.
 - Original design was too slow due to the Williams tubes and so *core memory* (ferrite rings that store data in polarity of magnetic field) was created => doubled speed => design successfully mass-produced by IBM

The Early Days - 1956

- IBM (originated with mechanical analogue calculators) used the core memory developed for the Whirlwind to make the IBM 704, which was also the first mass-produced computer to have *floating-point* hardware.
- The 709 improved the 704 by adding *overlapped I/O, indirect addressing* and *decimal instructions*.
- The 7090 improved the 709 by the use of *transistors* and not valves

IBM 704



The Early Days - 1960

- The LARC was the first true supercomputer – it was designed for *multiprocessing*, with 2 CPUs and a separate I/O processor.
 - Used 48 bits per word with a special form of decimal arithmetic => 11 digit signed numbers. Had 26 general purpose registers with an access time of 1 microsecond.
 - The I/O processor controlled 12 magnetic drums, 4 tape drives, a printer and a punched card reader.
 - Had 8 banks of core memory (20000 words) with an access time of 8 microseconds and a cycle time of 4 microseconds.

The Early Days - 1961

- IBM feared the success of the LARC so designed the 7030 to beat it.
 - Initially designed to be 100x faster than the 7090, it was only 30x faster once built. An embarrassment for IBM => big price drops and in the end only 9 were sold (but only 2 LARCs were built so not all bad!)
 - Many of the ideas developed for the 7030 were used elsewhere, e.g. *multiprogramming*, *memory protection*, *generalized interrupts*, the 8-bit *byte*, *instruction pipelining*, *prefetch* and *decoding*, and *memory interleaving* were used in many later supercomputer designs.
 - Ideas also used in modern commodity CPUs!

The Early Days - 1965

- CDC (Control Data Corporation) employed Seymour Cray to design the CDC 6600 which was 10x faster than any other computer when built.
 - First ever *RISC* system – simple instruction set – which simplified timing within the CPU and allowed for *instruction pipelining* leading to higher throughput and a higher clock speed, 10 MHz.
 - Used *logical address translation* to map addresses in user programs and restrict to using only a portion of contiguous core memory. Hence user program can be moved around in core memory by the operating system.
 - System contained 10 other “Peripheral Processors” to handle I/O and run the operating system.

CDC 6600



History (1970-1990)

Year	Name	Peak speed	Location
1974	CDC Star-100	100 MFLOPS (vector) ~2 MFLOPS (scalar)	Lawrence Livermore Lab. USA
1975	Cray-1	80 MFLOPS (vector) 72 MFLOPS (scalar)	Los Alamos Lab. USA
1981	CDC Cyber- 205	400 MFLOPS (vector) peak, avg much lower	
1983	Cray X-MP	500 MFLOPS (4 CPUs)	Los Alamos Lab. USA
1985	Cray-2	1.95 GFLOPS (4 CPUs) 3.9 GFLOPS (8 CPUs)	Lawrence Livermore Lab. USA
1989	ETA-10G	10.3 GFLOPS (vector) peak, avg much lower (8 CPUs)	
1990	Fujitsu Numerical Wind Tunnel	236 GFLOPS	National Aerospace Lab, Japan

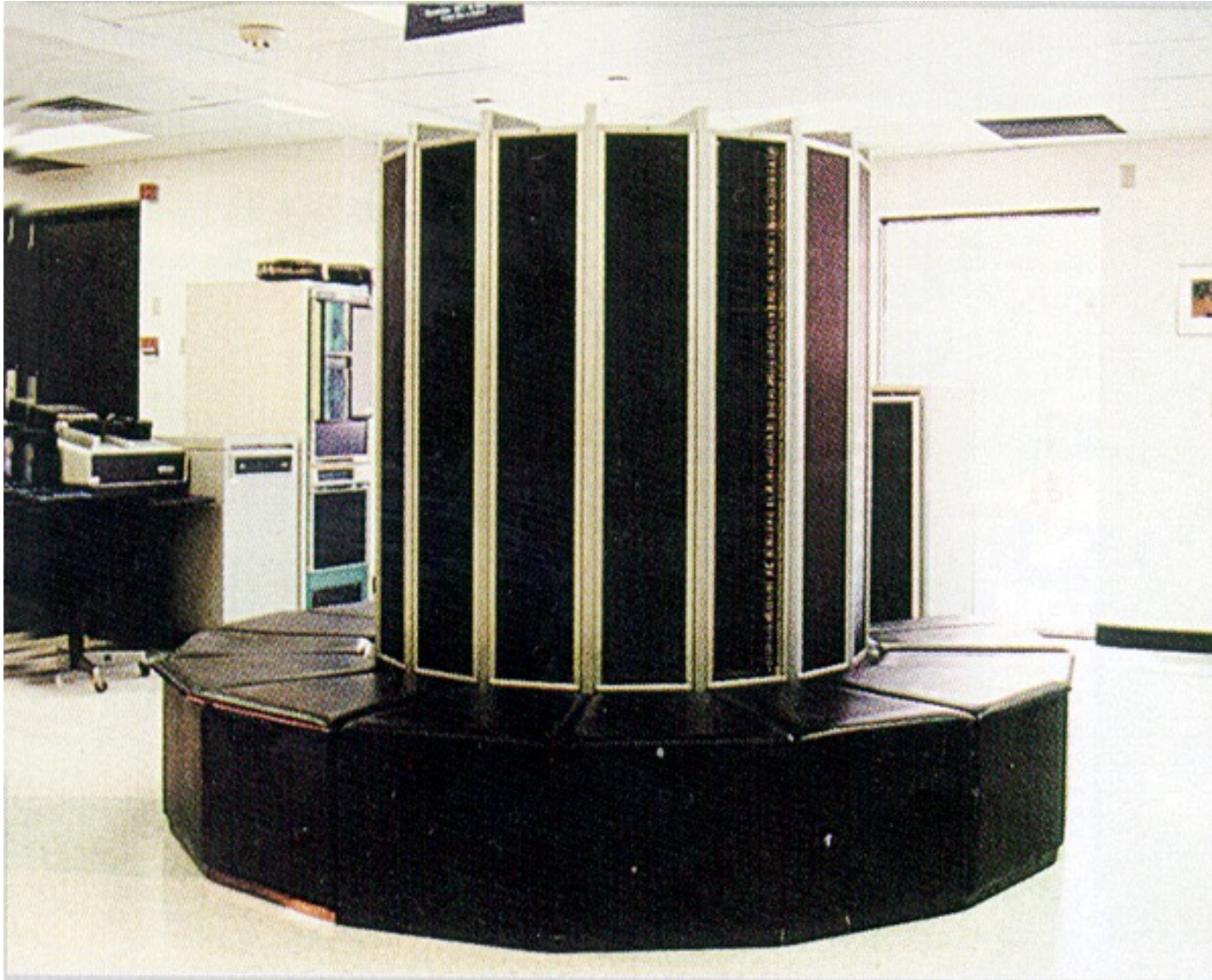
The Vector Years - 1974

- The CDC Star-100 was one of the first machines to use a *vector processor*
 - Used “deep” pipelines (25 vs. 8 on 7600) which need to be "filled" with data constantly and had high setup cost. The vector pipeline only broke even with >50 data points in each set.
 - But the number of algorithms that can be effectively vectorised is very low and need careful coding otherwise the high vector setup cost dominates.
 - And the basic scalar performance had been sacrificed in order to improve vector performance => machine was generally considered a failure.
 - Today almost all high-performance commodity CPU designs include vector processing instructions, e.g. SIMD.

The Vector Years - 1975

- Seymour Cray left CDC to form Cray Research to make the Cray-1.
 - A vector processor without compromising the scalar performance using *vector registers* not pipelined memory operations
 - Uses ECL transistors
 - No wires more than 4' long
 - 8 MB RAM and 80 MHz clock speed
 - Cost \$5-\$8m, with ~80 sold worldwide
 - Ships with Cray OS, Cray Assembler and Cray FORTRAN – the world's first auto-vectorising FORTRAN compiler

Cray 1



The Vector Years - 1981

- CDC Cyber-205
 - CDC put right the mistakes made with the Star-100
 - 1-4 separate vector units
 - Rarely got anywhere near peak speed except with hand-crafted assembly code
 - Used semiconductor memory and *virtual memory* concept

The Vector Years - 1983

- Cray X-MP
 - A parallel (1-4) vector processor machine with 120 MHz clock speed for ~125 MFLOPS/CPU with 8-128 MB of RAM main memory
 - Better chaining support, parallel arithmetic pipelines and *shared memory access* with multiple pipelines per processor.
 - Switched from Cray OS to UniCOS (a UNIX variant) in 1984
 - Typical cost ~\$15m plus disks!

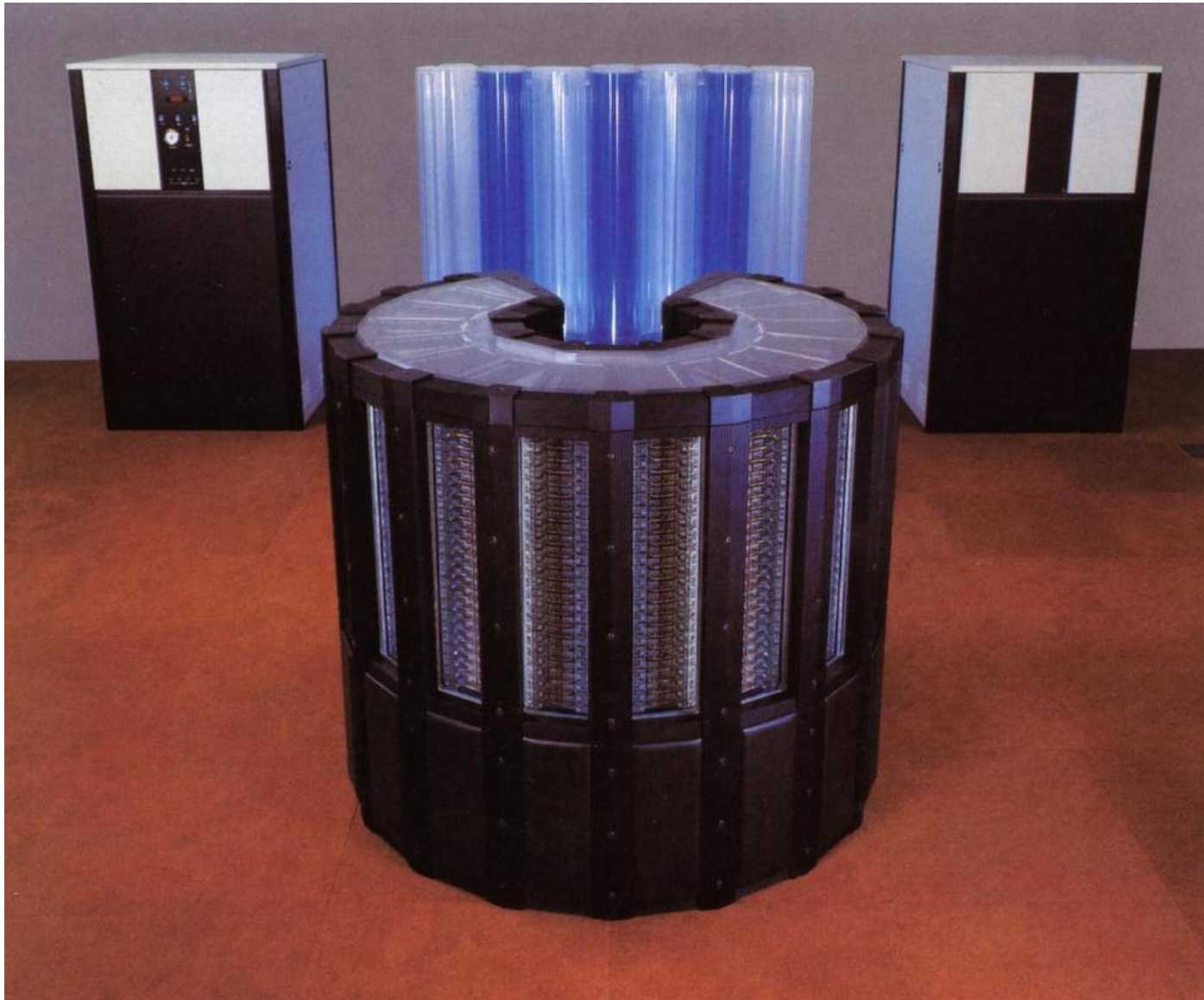
Cray XMP



The Vector Years – 1985 - 1989

- Cray-2 = 1985
 - A completely new, compact 4-8 processor design but higher memory latency than X-MP
 - Hence X-MP faster than Cray-2 on certain problems – impact of memory architecture ...
- ETA = 1989
 - Spin-off from CDC - commercial failure
 - Shared memory multiprocessor with mix of private memory/CPU plus shared memory
 - liquid nitrogen cooling variant

Cray 2



The Vector Years - 1990

- Fujitsu Numerical Wind Tunnel
 - Another vector parallel architecture with advanced Ga-As CPUs for lowest gate delay
 - Each CPU had four independent pipelines with a peak speed of 1.7 GFLOPS/CPU and 256 MB main memory.
 - Had sustained performance ~100 GFLOPS for CFD codes c.f. peak=236 GFLOPS!

Recent History (1991-2010)

Year	Name	Peak speed	Location
1995	Intel ASCI Red	2.15 TFLOPS	Sandia National Lab. USA
2000	IBM ASCI White	7.226 TFLOPS	Lawrence Livermore Lab. USA
2002	Earth Simulator	35.86 TFLOPS	Yokohama Institute for Earth Sciences, Japan
2005	IBM ASCI Blue Gene	70 - 478 TFLOPS	Lawrence Livermore Lab. USA
2008	IBM Roadrunner	1.105 PFLOPS	Los Alamos Lab. USA
2009	Cray Jaguar	1.75 PFLOPS	Oak Ridge Lab. USA

The Conventional Era - 1995

- Intel ASCI-Red
 - Developed under ASCI to build nuclear weapon simulators
 - Used commodity components for low-cost & very scalable
 - *A massively-parallel processing machine with Pentium II Xeons and MIMD (multiple instruction, multiple data) paradigm*
 - See MPI later



The Conventional Era - 2000

- IBM ASCI-White
 - A cluster computer based upon the commercial IBM RS/6000 SP computer with 512 machines
 - 3 MW electricity to run and additional 3 MW to cool
 - Cost \$110 million



The Conventional Era - 2002

- Earth Simulator

- A cluster computer based upon the NEC SX-6, which comprised 8 **vector processors**.
- Codes written using HPF with optimised NEC compiler
- Cost ¥7.2b (~£36m)



The Conventional Era – 2005

- IBM ASCI-Blue Gene/L
 - Had 65,536 power4 CPUs (dual core) with 3 integrated networks (different underlying topologies)
 - A *constellation* computer – made of an integrated collection of smaller parallel nodes
 - Cost \$100m



The Conventional Era – 2008

- IBM Roadrunner
 - AMD Opteron CPUs to handle O/S, interconnect, scheduling etc. and PowerXCell CPUs for computation
 - A unique **hybrid** architecture that required all software to be specially written - BIG challenge to program
 - Cost \$133m and 2.35 MW to operate



The Conventional Era - 2009

- Cray Jaguar
 - Conventional architecture with AMD 6-core CPUs
 - Required 6.9MW to operate



The GPU Era (2010-)

Year	Name	Peak speed	Location
2010	Tianhe-1A	2.57 PFLOPS	National Supercomputer Centre, Tianjin, China
2011	Fujitsu K computer	8.2 – 10.5 PFLOPS	RIKEN Advanced Institute for Computational Science, Japan
2012	IBM Sequoia	20.1 PFLOPS	Lawrence Livermore Lab. USA
2013	Tianhe-2	54.9 PFLOPS	National Super Computer Center in Guangzhou, China
2015	Sunway TaihuLight	125.4 PFLOPS	National Supercomputing Center in Wuxi, China
2018	Summit	187.6 PFLOPS	Oak Ridge National Lab, USA
2020	Fugaku	415.5 PFLOPS	RIKEN Center for Computational Science, Japan
2022	Frontier	1102 PFLOPS	Oak Ridge National Lab, USA

The GPU Era - 2010

- Tianhe-1A
 - Hybrid design – mix of CPU and GPU
 - Intel Xeon X5670 CPUs and Nvidia Tesla GPUs
 - Cost \$88m
 - 4MW to operate
 - GPU lecture later



The GPU Era – 2011-12

- Fujitsu K = 2011
 - Conventional design – no accelerators
 - SPARC64 VIIIfx 2.0 GHz 8-core CPUs
 - Required 12.6 MW to operate
 - Cost \$1.0bn plus \$10m/year to run ...
- IBM Sequoia = 2012
 - Conventional design = 98,304 IBM Power CPUs
 - Very energy efficient = only 7.9 MW
 - LAPACK 81% peak
 - Cost \$97m

The GPU Era – 2013-15

Tianhe-2 = 2013

- Intel Xeon Ivy Bridge + Intel Xeon-Phi
- High power = 17.8 MW and only 62% of peak
- Cost \$390m

Sunway TaihuLight = 2015

- Made entirely out of Chinese chips!
- 2x speed and 3x efficient as Tianhe-2
- Less power = 15.37 MW and 75% peak

The GPU Era – 2018-20

Summit – 2018

- IBM Power9 system with nVidia Volta GPUs
- Energy efficient = 8.8 MW and 65% peak
- 13.889 GFLOP/ W
- Cost ~\$162m

Fugaku – 2020

- Fujitsu with ARM chips – no accelerators
- Each node = 1x 48-core A64FX-48 CPU
- LAPACK = 80% of peak but energy = 28.3 MW!
- Cost ~\$1b

Current Best – 2022

Frontier

- HPE Cray with AMD CPU & AMD GPU
 - AMD EPYC3 Milan CPU (64-core) + 4 GPU per node, with 9,472 nodes in 74 cabinets
 - AMD Radeon Instinct GPUs – 1st time #1
 - Theoretical peak = 2 exaflop but LAPACK= 55%
 - Gigabit ethernet with Slingshot-11 interconnect
 - 62.68 GFLOP/ W so #1 in Green500 table
 - Cost ~\$600m
- <https://www.olcf.ornl.gov/frontier/>

Frontier



ARCHER2 – the UK national academic supercomputer

- Academic Research Computer High End Resource
 - Delayed from May 2020 to Dec 21 due to COVID
 - HPE Cray EX system with 25 PFLOPS (#25)
 - 2*AMD EPYC2 Rome 64-core CPU per node
 - 716,800 cores in total
 - No GPUs
 - Cray Slingshot interconnect
 - £48m for system + support + training etc
 - 6 MW power limit of the EPCC computer centre!
 - <https://www.archer2.ac.uk/>

And for York ...

- First came Erik – a research cluster with 32 dual Intel P4 Xeon nodes – 2002 – for ~0.1 TFLOPS
- 2008 Edred – a research cluster with 32 dual nodes with 4-core AMD Barcelona CPUs – (1.2 TFLOPs)
- 2010 Jorvik – a teaching cluster with 5 hex-core AMD CPUs in 2010 (~0.16 TFLOPs)
- 2015 Ebor with 16 dual nodes with 8-core Intel Xeon E5-2630 v3 (2015, ~12 TFLOPs)
- 2018 Viking – a £1.8m supercomputer – 170 nodes each with 2x 20-core Xeon 6138 (~150 TFLOPs)
<https://wiki.york.ac.uk/display/RCS/Viking+-+University+of+York+Research+Computing+Cluster>

Trends

- Vector processing has come, gone and is now back
- Parallel processing is here to stay
- Hybrid architectures, with CPU + GPU accelerator, becoming more common
- Cluster computing gives high performance with marginal cost using commodity components – big shift away from niche/proprietary components since 1990s
- Difficult to exploit many processors in a single task – need parallel programming concepts
- Many hardware & software concepts developed for supercomputers are now being used in latest commodity high-performance CPUs (Intel, AMD, ARM)
- Electrical power requirement now non-trivial – rise of “green computing”

Further Reading

- Wikipedia at <http://www.wikipedia.org>
- The History of Computing at <http://www.thocp.net>
- Top500 and Green500 at <http://www.top500.org>
- “Computer Architecture – A Quantitative Approach (6th edition)”, John L Hennessy and David A Patterson, Morgan Kaufmann Pub. Inc. (2017).