Computer Structure and Language

Hamid Sarbazi-Azad
Department of Computer Engineering
Sharif University of Technology (SUT)
Tehran, Iran

REMINDER OF LAST SESSION

Computers Then vs Now

IBM System 390

- Introduced in 1990s.
- Top model had:
  - x12 637 MHz CPUs
  - 32 GB Memory

https://en.wikipedia.org/wiki/IBM_System/390

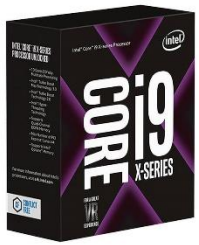
IBM Z16

- Introduced in 2023.
- Fully fitted model has:
  - x200 5.2 GHz CPUs
  - 40 TB Memory

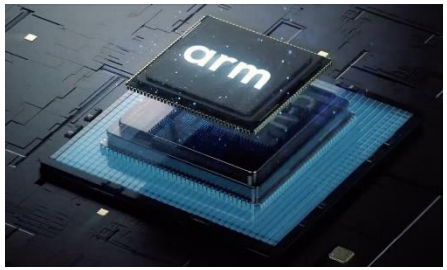
https://en.wikipedia.org/wiki/IBM_Z

Last Session ISA Examples

How many of these are in use today?



x86



ARM

RISC-V

IBM Z

IBM Power

WHAT IS A SERVER?

What is A Server?

- A computer that provides service to others.
  - Often happens over network
- Any computer can act as a server
  - Only needs an interface for providing the service
- Servers don’t require special hardware.
  - Your home PCs can become servers (e.g. Web Servers, Storage Server)

HPCAN Monitoring Server

So What is Server Grade (Enterprise) Hardware?

- More Expensive than Desktop
- Often Slower than Desktop
- Higher Build Quality and More thorough Testing
- Often More Stable and More Reliable
- Optimized for Different Workloads
- CPUs
  - support less clock variation (or no variation)
  - Inter-compatible with other server components (ECC Memory)
  - Some doesn’t have dedicated support for graphical operations
  - Some come with specific accelerators
  - Support Larger amounts of memory
  - Support for multi CPUs on a single board
- More Energy Efficient

Server Form Factors

Tower Server

Rack Server

Blade Server

Example

SuperMirco X13/H13

- CPU: 4th Gen Intel® Xeon® Scalable Processors or AMD EPYC™ 9004 Series Processors
- Memory: Up to 32 DIMMs, 8TB DDR5, support for Intel® Optane® Persistent Memory
- Drives: Up to 24 hot-swap 2.5” NVMe/SATA or 12 hot-swap 3.5” NVMe/SATA
- PCIe: Up to 4 PCIe 5.0 x16 or 8 PCIe 5.0 x8 slots

GOING BIGGER

Super Computers

Super Computers

- Massively Parallel General Purpose Computers
- Generally Run Long Running Tasks
  - e.g. simulations, AI training
- Generally are not interactive & get tasks from a job queue

- Cluster (Distributed System) or MPP (Single Machine)
- CPUs from all architectures

- Check https://www.top500.org/ for rankings!!!

Top Super Computers – June. 2023

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	**Frontier** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	8,699,904	1,194.00	1,679.82	22,703
2	**Supercomputer Fugaku** - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
3	**LUMI** - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,220,288	309.10	428.70	6,016
4	**Leonardo** - BullSequana XH2000, Xeon Platinum 8358 32C 2.6GHz, NVIDIA A100 SXM4 64 GB, Quad-rail NVIDIA HDR100 Infiniband, Atos EuroHPC/CINECA Italy	1,824,768	238.70	304.47	7,404
5	**Summit** - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148.60	200.79	10,096
6	**Sierra** - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94.64	125.71	7,438
7	**Sunway TaihuLight** - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi	10,649,600	93.01	125.44	15,371

Mainframes

Mainframes

- High Performance Computers With large amounts of memory and processing power
- Highly resilient and secure
- Capable of handling very large number of transactions in real time
- Often feature highly redundant hardware
- Some models support live maintenance operations
- Made By
  - IBM (IBM Z, Linux One)
  - HP (HPE Superdome)
  - FUJITSU (Server GS21, BS2000)
  - …

Mainframes

- Used by 71% of Fortune 500 Companies
- Handle 90% of all credit card transactions
- Handle 68% of world's production IT workload, account for 6% of the costs
- 92 of world's top 100 banks and all top 10 insurance companies use mainframes (as of 2020)
- The median mainframe programmer salary has skyrocketed to $111,116 according to ZipRecruiter.
  - $40,000 higher than the median salary for computer programmers in general, according to the same site.

- According to: https://www.precisely.com/blog/mainframe/9-mainframe-statistics

IBM Mainframes

- Dominates Market with over 90% market share.
- Highly redundant (even in CPU cores)
- Multi frame version has 200 active 5.2 GHz CPU cores and 40 TB (IBM/Z 16) or 48 TB (LinuxOne Emperor 4) RAIM memory in 4 compute drawers
- 250 Gb/s Memory Access (For each drawer)
- Built in backup battery, redundant cooling, etc.
- One Compute Drawer could be removed or go down while the server is running without affecting the applications.
- Can be partitioned between up to 30 operating systems without virtualization.
- Has many more processors (e.g. on each IO Card) other than CPUs.
- Chip Accelerators for AI, Compression and Quantum Safe Cryptography
- Memory could be encrypted

Further Study

- What are Mainframes?:
  - https://youtu.be/ximv-PwAKnc?si=Hg4LuRxSYPYOSZLG
- I Tried to Break a Million Dollar Computer - IBM Z16 Facility Tour!:
  - https://youtu.be/ZDtaanCENbc?si=CbMUNhrDbBiSzbKb
- Why Do Mainframes Still Exist? What's Inside One? 40TB, 200+ Cores, AI, and more! :
  - https://youtu.be/ouAG4vXFORc?si=vaD3jOGPKI7lLIEs
- Supercomputers as Fast As Possible:
  - https://youtu.be/unDqlQA9rU4?si=Np0wGRqgrFENO1_0
- Preparing for Frontier:
  - https://youtu.be/HmQjMEPhfDM?si=NdX82s4HJTHolquC
- The Journey to Frontier:
  - https://youtu.be/Ny_NvpuaAiQ?si=uq6KpYusU0n4rB_1

END OF SLIDES