System Administration

Processing power and Bandwidth

http://courses.cs.ut.ee/2012/syshald/
Outline

- Processing power
- Bandwidth
- Course project
- About test
Processing Power

- systems ability/capacity to process data
  - data processing – performing mathematical operations with stored numeric data and storing results for later use
- centered around the CPU
  - Central Processing Unit
Processing Power (2)

- CPUs historically had logical circuits for every supported instruction, for example addition, multiplication etc.
- Modern CPUs interpretate incoming instructions with built-in `microcode` and executes some directly on hardware and splits some to set of simpler instructions
Processing Power (3)

- CPUs processing power is described by
  - clock rate
    - Describes how long will take to compute one simple circuit level instruction
    - More complex instructions and instructions performed by microcode take several machine cycles
  - features of internal instruction set
    - which instructions are executed directly on circuit, which are interpreted by microcode etc.
  - count of processing cores
Processing Power (4)

- Different CPUs may support different set of instructions
  - internal vs. external instruction set
  - external instruction set is used by programs
  - internal instruction set is used internally by processor
- For example, Intel and AMD processors implement identical \textbf{x86} external instruction set, but have completely different internal instruction sets
Processing Power (5)

Processor Internal Architectures

PC/microcomputer
- Intel's x86, IA32, x86-64, AMD64, Intel64
- Power Architecture (PPC)

Workstations/Servers
- MIPS
- Alpha
- Itanium
- Sparc

Embedded
- ARM
- PIC
What is currently most widely used processor architecture?
Processing Power (6)

Because different CPUs implement different external and internal instruction sets, one cannot compare CPUs processing power only by clock rate:

- Actually, there's no reliable single figure to describe CPUs processing power
- MIPS – Million Instructions Per Second
- FLOPS – Floating-point operations Per Second
Processing Power (7)

- Reliable comparisons between CPUs can and should be done by benchmarking software which mimics expected workload
  - Linpack: http://www.netlib.org/linpack
  - Dhrystone: http://www.ct.se/dhrystone
  - PCMark: http://www.futuremark.com
  - etc...
Processing Power (8)

- Different external instruction sets are hidden from user/sysadmin by compilers and standardized APIs
- For example, Debian/GNU Linux is available on at least 9 architectures
  - amd64
  - ia64
  - powerpc
  - armel
  - mips
  - s390
  - i386
  - mipsel
  - sparc

http://www.debian.org/ports
Processing Power (9)

- Computing power can not be stored
  - CPU can make certain amount of calculations on any given time. If those calculations are not used then they are irrecoverably lost
  - Compare with disk space: when disk space is not used now, it can be used later
Processing Power (10)

- One way to increase systems processing power is to add additional CPUs to the system
- Intel Core i7 has $731 \times 10^6$ transistors
- Intel Pentium II has $7.5 \times 10^6$ transistors
- So, we could have 100-core Pentium II with same amount of transistors as i7
Still, technology has not developed towards massively multicore CPUs. Why?
Processing Power (11)

Processing Power (12)

• Amdahl's law
  
  - Let $P$ is proportion of program that can made parallel. Thus $(1-P)$ is proportion which remain serial. Then maximum speedup that can achieved with $N$ CPUs is

  $$S(N) = \frac{1}{(1-P) + \frac{P}{N}}.$$
Processing Power (13)

Processing Power (14)

- Parallelization has fundamental limits
- Parallel / multithreaded software is hard to write
- Some C/C++ compilers can automatically parallelize serially written code
Bandwidth

- capacity for data transfer in a given period of time
- many faces of bandwidth:
  - width / parallelism
    - how many bits/bytes are transferred simultaneously
  - frequency / rate
    - how often can new data read or sent
  - physical properties, protocol
    - bus, data channel / data path
Bandwidth (2)

- bus – communication channel in which set of clients are connected via shared communication line
- data channel / data path – specialized communications channel between two components
Bandwidth (3)

- buses
  - standardized
  - connect different devices
  - multiple manufacturers

- datapaths
  - little standardization
  - manufacturer specific
  - faster
Bandwidth (4)

• buses
  - Ethernet
  - IDE
  - SCSI
  - PCI
  - PCIe

• datapaths
  - between *level 1 cache* and CPU
  - between read head and electronic controller of HDD
Bandwidth (5)

- Bandwidth vs Latency
  - bandwidth $\neq$ fastness
- perception of fastness depends on:
  - bandwidth, throughput (speed)
  - latency, the delay between sending and receiving data
Consumers

- Processing power and bandwidth is consumed/used by
  - generally speaking: software
  - application software
  - operating system
    - kernel, user space
  - hardware consumers
Resource Problems

- Computers work-capacity is bounded by processing power and different bandwidths
  - Why amount of memory is not mentioned?
- Different workloads require different amount of different resources from system
  - For example, typical web-server is bounded by network and/or hard disk bandwidth and consumes little CPU power
  - Media encoding server is bounded by CPU power, network and maybe even by PCIe bus
Resource Problems (2)

- When available resources are utilized unevenly, bottleneck effect appears
- It means that systems component, which is utilized to full capacity, prevents other components from working with their full capacity
  - Web server has plenty of network bandwidth but can serve content as fast as it gets files from hard disk
  - What can be done?
Resource Problems (3)

• In general, resource problems are caused by
  - shortage of resources
  - impractical use of resources

• Solutions to resource problems could be
  - reducing the systems load
  - increasing the capacity
  - using alternative paths (redirecting)
Resource Problems (4)

- Amount of needed computing resources is hard to plan and predict
- Constantly monitor performance of Your systems and act early as possible
- Rule of thumb: (slight) over-dimensioning early is generally cheaper then adding resources later
  - This rule of thumb has its limits of application
Course Project

• Course project
  - Individual work or work of groups of two
  - Demonstrates students system administration skills and knowledge
  - Will be graded at the end of course (20% of total grade)
  - Every group will review and grade some other groups project
  - Should we have a lecture for project presentations and discussion?
Course Project (2)

• Course project
  - Set up at least two Virtualbox virtual machines which work together
  - At least one server machine and one client machine
  - Use of some other operating system then Debian/GNU Linux is encouraged (extra points)
  - At least 10 demonstrable tasks per group member
Course Project (3)

- Course project. Example tasks
  - Firewalled machines, only necessary services are passed through
  - Dhcp server is set up, client gets its network configuration
  - NFS/CIFS server is set up, client can access
  - Fileserver serves files from RAID/LVM volumes
  - Web server is set up, client can access web pages
Course Project (4)

- Course project
  - Project work is presented as Virtualbox images to the reviewers
  - We will arrange file server for image uploads later
  - To surely get maximum points from project
    - Use at least 2 different op-systems per group member
    - Build a fancy redundant client-server system
Test!

- 08.03.2012 12:15-12:40 in Tartu, Liivi 2-111 (this room, lecture time)
- 06.03.2012 in Tallinn, Akadeemia tee 15A. Building near KBFI, with Nordea bank logos and cranes.
  - 16:15, 17:15, 18:15? Vote!?
  - I'll wait You in lobby. Don't be late, as i don't know in which room we will be.
  - If You are late then just wait in a lobby until someone who finished the test comes
Questions?