IT Infrastructure

Reading:

Laudon & Laudon

chapter 4

Additional Reading:

Brien & Marakas

chapter 3, 4

Outline

- Computer Hardware
- Computer Software
- Case Study
  - Technology Cure at UPMC
  - Computing Goes Green
- Managing Hardware and Software Technology

Technology Cure at UPMC

- Problem
  - Large, complicated infrastructure was difficult to manage
  - Storage needs increased quickly, ↑ 20% year

- Solution
  - Use IBM storage technology to reduce infrastructure spending and conserve space
  - IBM → Use Virtualization to reduce number of servers
    - Virtualization – Several applications on single physical layer on different OS
  - UNIX OS on IBM System p5 595 servers and *VMware*
    - Server Utilization ↑ from 3 to 80%
    - Server Capacity ↑ 150%; Space required for Servers ↓ 40%
    - *Server Virtualization Project* can save 18-22 million $ (hardware, floor space, staff)

Infrastructure Components

Computing Requirements

- Google – 80 million queries/day (most within one second) in USA
  - Network of 450,000 PC Servers linked together to spread workload

IT Infrastructure

- Provides platform for supporting all information systems in the business
- Computer hardware
- Computer software
- Data management technology
  - Organizes, manages, and processes business data concerned with inventory, customers, and vendors
- Networking and telecommunications technology
- Technology services
  - Example - Consultants for systems integration with *legacy systems*

Types of Computers

- Computers come in different sizes with varying capabilities for processing information
  - FLOPS (Floating point operations per second)
  - 500 (Handheld) – 1 Trillion (Supercomputer)
- PDAs, handheld mobile devices
- PCs
- Workstation
  - More powerful mathematical and graphics-processing capabilities than a PC
- Servers
  - Type of midrange computer
  - Support computer network, sharing files and resources
  - Provide hardware platform for e-commerce
  - PC

Types of Computers

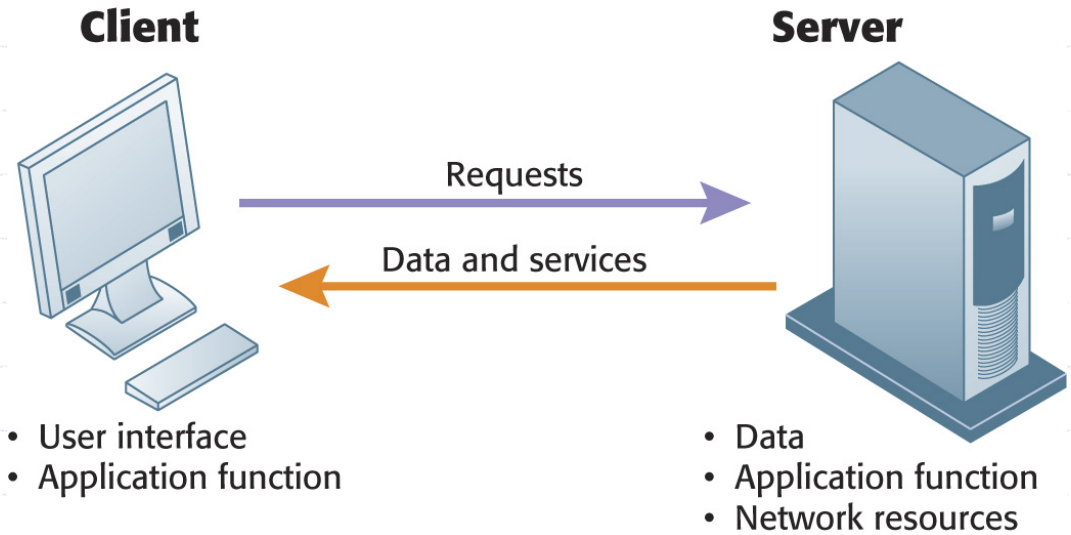
- Mainframes
  - Large-capacity, high-performance computer that can process large amounts of data very rapidly
  - Example - Airlines usage for 3000+ reservations per second
- Supercomputer
  - More sophisticated computer used for tasks requiring extremely rapid and complex calculations with thousands of variables, millions of measurements
  - Used in engineering, scientific simulations, military/weapons research, weather forecasting
- Grid Computing
  - Power of geographically remote computers connected into single network to act as “virtual supercomputer”
  - Combines computational power of all computers on grid

Types of Computers

Computer	Processor/Speed	Performance	Comment
Personal digital assistant (PDA) Palm handheld	Intel™ XScale/ 312 MHz	~500 FLOPS	PDAs are generally asked to perform one task at a time by the operator. Most of the processing power is used to draw the screen and handle voice messages.
Personal computer Dell XPS 720 H2C	Intel Core 2 Extreme (quad-core) processor/3.67 GHz,	4 Giga FLOPS	High-end game machine. Most PCs used in business are 1–3 GHz, with 2 GFLOPS performance, plenty for word processing, Web surfing, and spreadsheets.
Server computer (midrange computer) SUN Sun Fire E6900 Server	UltraSPARC IV+/1.8 GHz	~48 Giga FLOPS	Up to 24 processors can be used with this powerful server.
Mainframe computer IBM System z9 Enterprise Class	System z9 Integrated Information Processor/Equal to 100 or more distributed processors	~1 Tera FLOPS	Up to 60 logical partitions, each with 64-bit central memory addressability.
SuperComputer IBM Blue Gene/P	4 PowerPC 450 processors per chip/850 MHz	~1 Peta FLOPS	Configured with 294,912 PowerPC processors on 72 racks. A chip of 4 processors capable of 13.6 billion operations per second.
Distributed Computing Grid Folding@home	Various PC processors, whatever is available on the Internet.	~125 Peta FLOPS	A volunteer program with approximately 250,000 CPUs online; the largest and fastest online distributed computing project devoted to study protein folding.

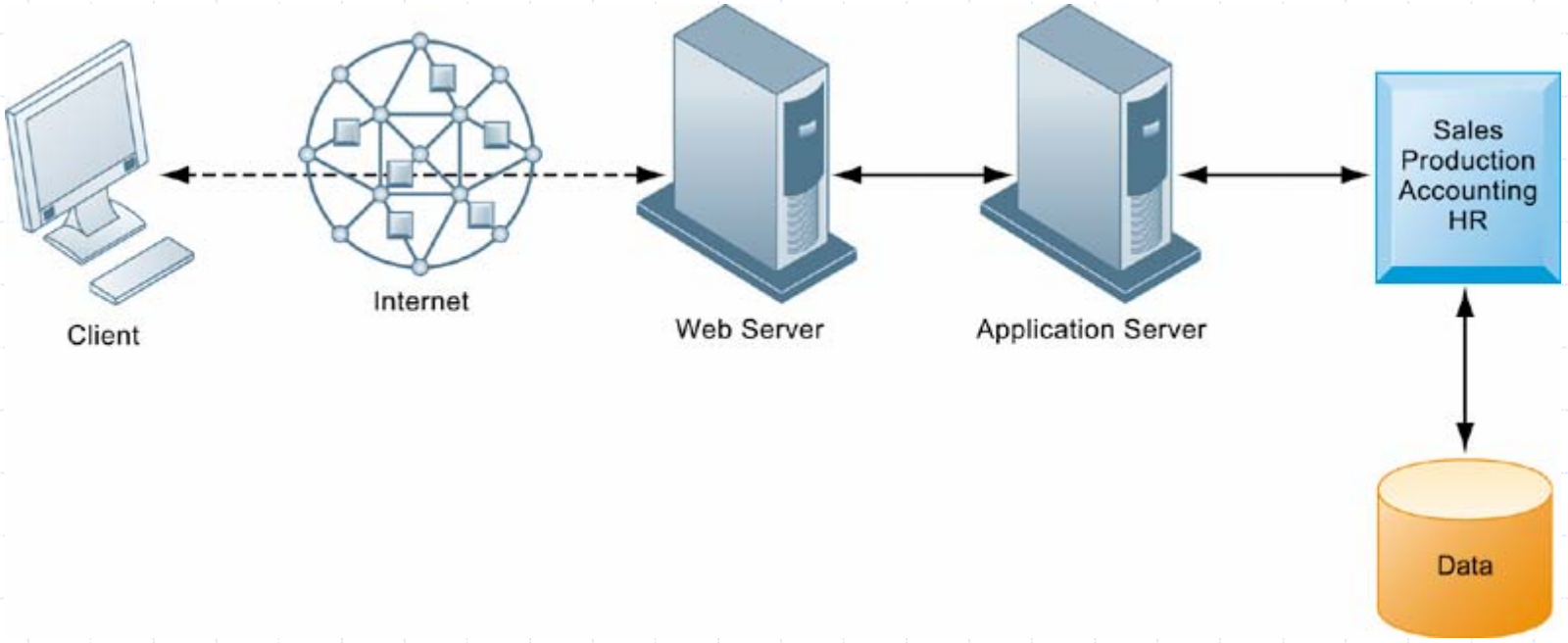
Types of Computers

- Client/Server Computing
  - Widely used form of *distributed computing*
  - Splits processing between “clients” and “servers”
  - Clients → User point of entry
  - Servers → Store and process shared data and perform network management activities
  - Two tiered **Two-tiered client/server architecture**
    - Uses two types of machines



Types of Computers

- Client/Server Computing
  - Multitiered client/server architecture (N-tier)
    - Balances load of network over several levels of servers
    - Example - Web servers and application servers

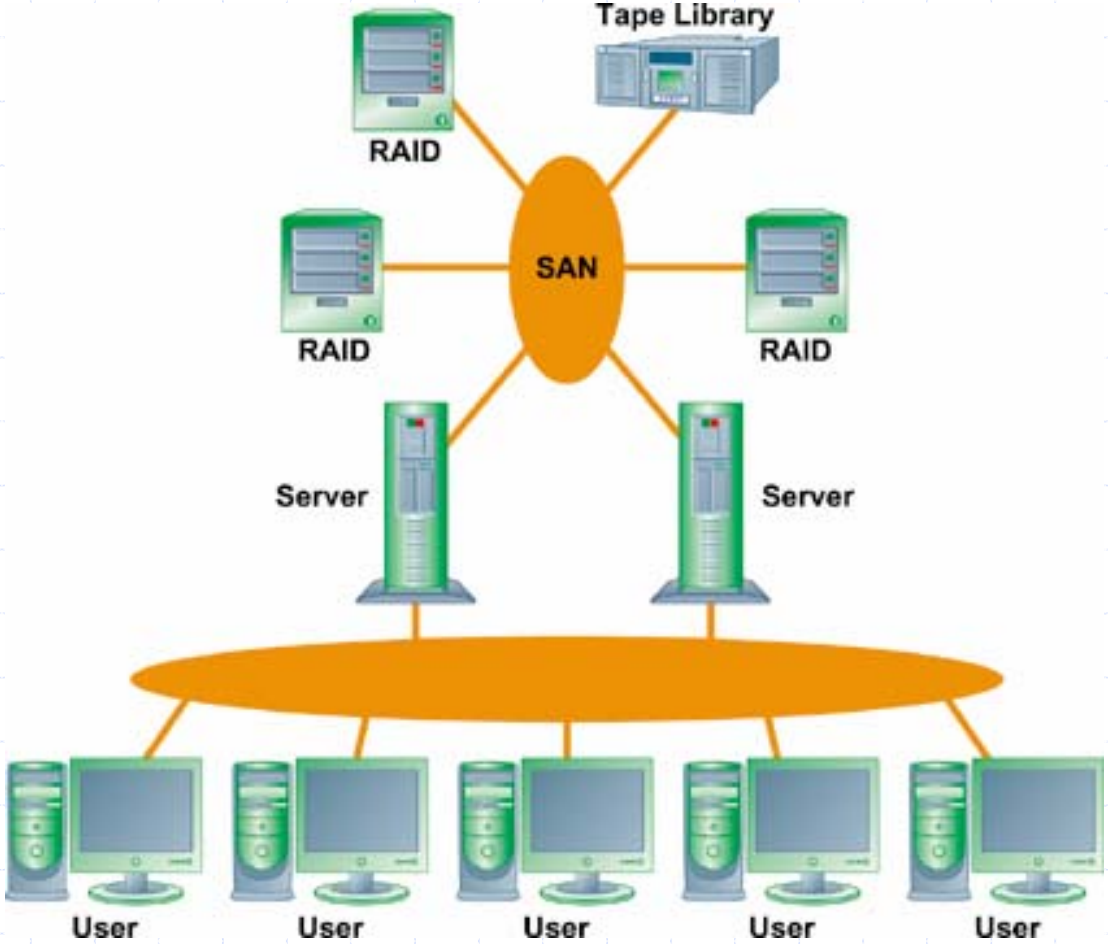


Storage, Input and Output

- Primary secondary storage technologies
  - Magnetic disk
    - Hard drives, USB flash drives
    - RAID → Redundant Array of Inexpensive disks
      - Can package hundreds of drives for data storage requirements
      - Performance and reliability issues
      - Reliability → Introduce Redundancy (Mirroring)
  - Optical disks
    - CD-ROM, CD-RW, DVD
  - Magnetic tape
  - Storage networking: SANs
    - Connect multiple storage devices on a separate high-speed network dedicated to storage

Storage, Input and Output

- Primary secondary storage technologies
  - Storage networking: SANs



Storage, Input and Output

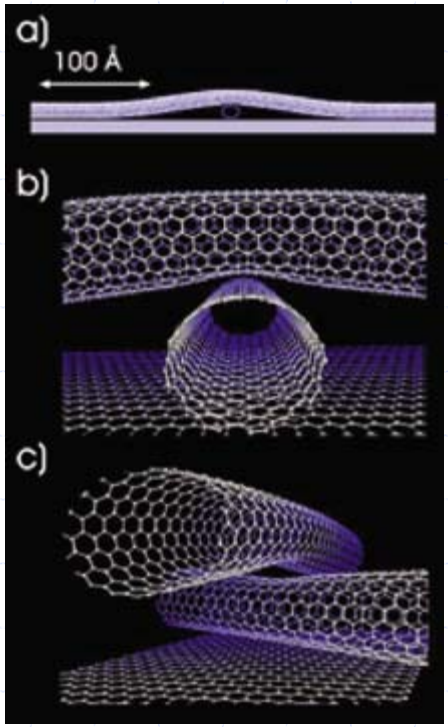
- Input devices
  - Gather data and convert them into electronic form
    - Keyboard
    - Computer mouse
    - Touch screen
    - Optical character recognition
    - Magnetic ink character recognition
    - Pen-based input
    - Digital scanner
    - Audio input, Sensors
- Output devices
  - Display data after they have been processed
    - Monitor
    - Printer
    - Audio output

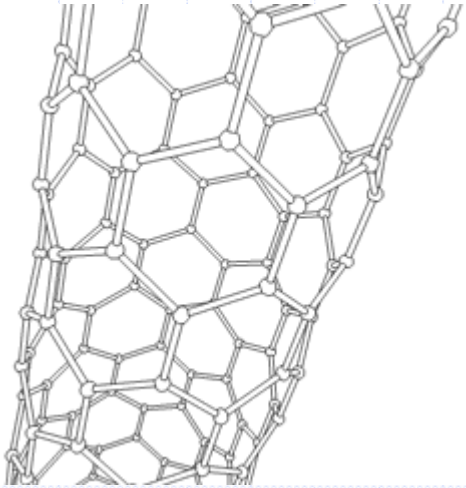
Contemporary Hardware Trends

- Integration → Computing and Telecom
  - Cell phones merging with handhelds
  - Growth of Internet telephony
- Nanotechnology
  - Deals with the study, development and practical application of structures and devices at molecular scale (between 1 and 100 nanometers)
  - Using individual atoms and molecules to create computer chips and other devices
  - Thousands of times smaller through manipulating individual atoms, molecules
  - IBM & Other Labs → Created transistors from *nonotube*

Contemporary Hardware Trends

- Nanotube

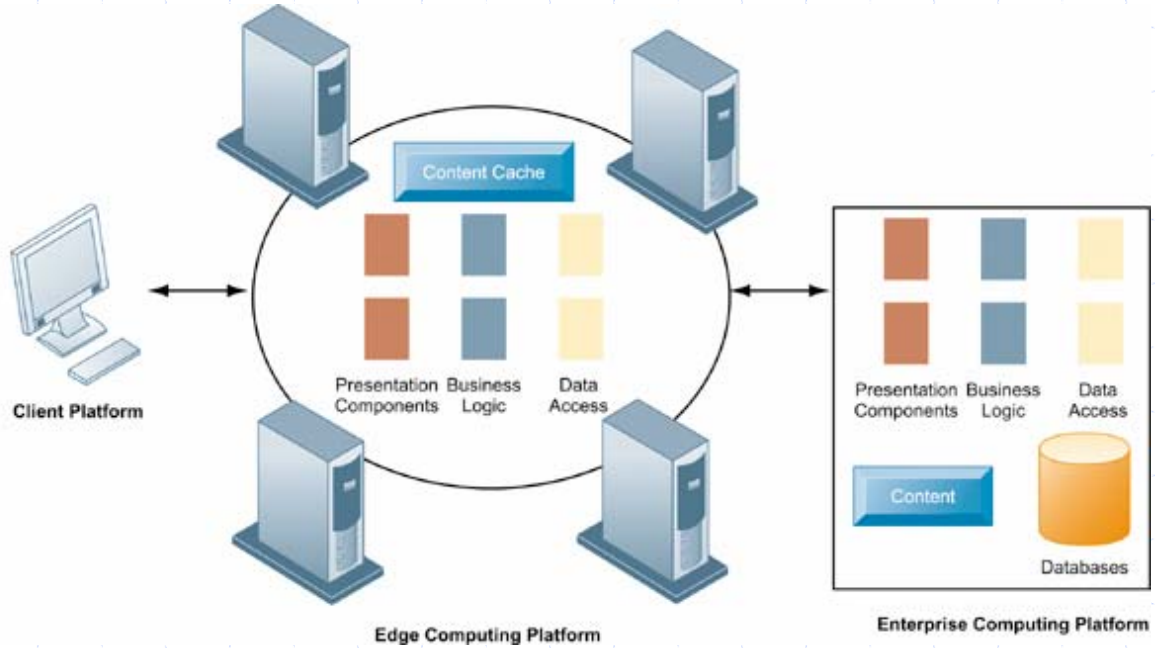




- Rolled up sheets of carbon hexagons
- High efficiency of heat transfer - Chip and its Heatsink
- High Conductivity – Ultra small electronic devices

Contemporary Hardware Trends

- Edge Computing
  - Multitier, load-balancing scheme for Web-based applications
  - Significant parts of Web site content, logic, and processing performed by smaller, less costly servers located nearby the user
  - Increases response time and lowers technology costs
  - Some sense like grid computing



Contemporary Hardware Trends

➢ Autonomic Computing

- Initiated by IBM in 2001, *Autonomic Nervous System*
- Systems that can configure or heal themselves
  - Example - Self-updating antivirus software
- IBM has defined four functional areas
  - **Self-Configuration** → Automatic configuration of components
  - **Self-Healing** → Automatic discovery and correction of faults
  - **Self-Optimization** → Automatic monitoring and control of resources to ensure the optimal functioning with respect to defined requirements
  - **Self-Protection** → Proactive identification and protection from arbitrary attacks
- Five Levels
  - Level 1 → Manual Management
  - Level 5 → Ultimate Goal of autonomic self managing systems

Contemporary Hardware Trends

- Virtualization
  - *Abstraction of computer resources*
    - Most Servers run 10-15% of their capacity
    - Utilization ↑ by 70% or higher, Fewer computers to process same work
  - Makes it possible to run multiple OS and multiple applications on the same computer at the same time, increasing the utilization and flexibility of hardware

  - Advantages
    - Server Consolidation and Infrastructure Optimization
    - Physical Infrastructure Cost Reduction
    - Improved Operational Flexibility & Responsiveness
    - Increased Application Availability & Improved Business Continuity
    - Improved Desktop Manageability & Security
  - Reference - http://www.vmware.com/technology/virtual-infrastructure.html

Contemporary Hardware Trends

➢ Multi-core Processors

- Chip-level Processor (CMP) combines two or more independent cores into a single package composed of a single IC/die
- *Dual-Core, Quad-Core*
- Enhanced performance, reduced power consumption and more efficient simultaneous processing of multiple tasks {resource hungry chip *if* single processor on core}
- Cores may share a single coherent cache at highest on-device cache level
  - Example - L2 for the *Intel* Core 2
- Cores may have separate caches
  - Example – Current *AMD* dual-core processors
- The ability of multi-core processors to increase application performance depends on the *use of multiple threads within applications*

Contemporary Hardware Trends

- RFID
  - *Radio Frequency IDentification*
    - Tag and Identify Mobile Objects
    - Communication → Transference of data over electromagnetic waves
    - Generates specific electromagnetic wave at the source, its effect can be noticed at the receiver far from the source, which then identifies it and thus the information
    - Demonstration of reflected power (modulated backscatter) 12 bit RFID tags - 1973
  - RFID Tag (Transponder)
    - Antenna, a wireless transducer and an encapsulating material
    - Active or Passive
    - Passive tags → Cheaper, Lower Range (<10mts) and more sensitive to regulatory and environmental constraints, as compared to active tags
  - RFID Reader (Transceiver)
    - Antenna, Transceiver and Decoder, which sends periodic signals to inquire about any tag in vicinity
  - Data processing subsystem
    - Means of processing and storing the data
  - Frequency Ranges
    - Low-Frequency (LF: 125 - 134.2 kHz and 140 - 148.5 kHz)
      - Short reading ranges, Lower system costs
      - Commonly used in security access, asset tracking, and animal identification applications
    - High-Frequency (13.56 MHz) and Ultra-High-Frequency (868 MHz - 928 MHz)
      - Long read ranges (greater than 90 feet), High reading speeds, Higher System Costs
      - Railroad car tracking and automated toll collection

Contemporary Hardware Trends

- RFID
  - Frequency Range

Frequency Range	Description	Typical Applications
<135KHz	Low Frequency, Inductive coupling	Access Control & Security Widgets identification through manufacturing processes Ranch animal identification OEM applications
13.56 MHz	High Frequency, Inductive coupling	Access Control Library books Laundry identification OEM applications
868 to 870 MHz 902 to 928 MHz	Ultra High Frequencies (UHF), Backscatter coupling	Supply chain tracking
2.400 to 2.483 GHz	SHF, Backscatter coupling	Asset tracking Highway toll tags Vehicle tracking

  - Examples
    - High-Frequency RFID systems → Hospitals to track a patient's location, and to provide real-time tracking of the location of doctors and nurses in the hospital
    - RFID chips for animals → Extremely small devices injected via syringe under skin. Under a government initiative to control rabies, all Portuguese dogs must be RFID tagged
    - RFID based traffic-monitoring system, which uses roadside RFID readers to collect signals from transponders that are installed in about 1 million E-Pass and SunPass customer vehicles

Contemporary Hardware Trends

- RFID
  - Active Vs Passive

	Passive RFID	Active RFID
Power Source	External (Reader provided)	Internal (Battery)
Tag Readability	Only within the area covered by the reader, typically up to 3 meters.	Can provide signals over an extended range, typically up to 100 meters..
Energization	A passive tag is energized only when there is a reader present.	An active tag is always energized.
Magnetic Field Strength	High, since the tag draws power from the electromagnetic field provided by the reader.	Low, since the tag emits signals using internal battery source.
Shelf Life	Very high, ideally does not expire over a life time.	Limited to about 5 years, the life of a battery.
Data storage	Limited data storage, typically 128 bytes.	Can store larger amounts of data.
Cost	Cheap	Expensive
Size	Smaller	Slightly bulky (due to battery)

  - Hitachi μ-chip → 0.05 x 0.05 mm, embedded in a sheet of paper, much more data, data can be extracted from few hundred meters
    - Size of Antenna – about 80 times than chip (in best version)
  - E-Passports
    - First RFID passports (E-Passport) → Malaysia in 1998, Standards - ICAO
    - Portugal, Poland, Japan, UK, Australia, USA (2007) – 10 cm; Thin metal lining; now PIN
  - HSBC (HK) → Automatically track customers (>1million HK$) using RFID

Contemporary Hardware Trends

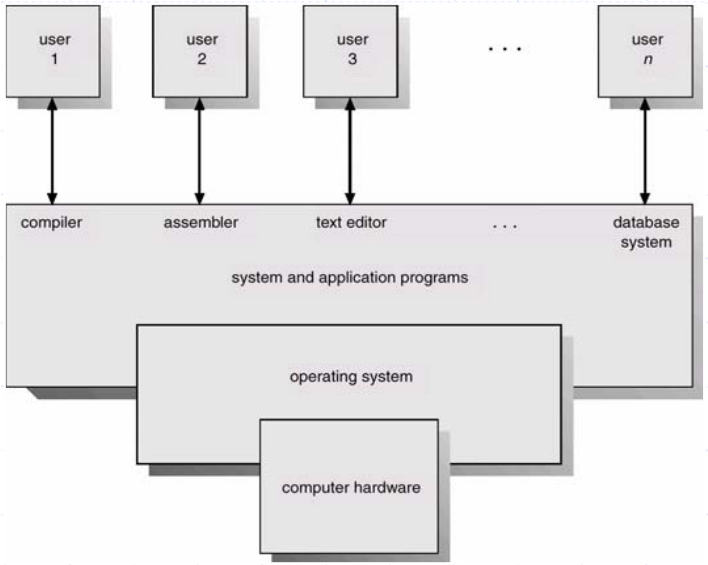
- RFID
  - Bar Code Vs RFID

Parameter	Bar Code	RFID
Frequencies used for tag reading	Optical frequencies	Radio frequencies
Type of communication	Line of sight communication	Non-line of sight communication
Data Volume	Physical limitation exists. It is very difficult to read a very long barcode.	Can carry relatively large volume of data.
Range of data readability	Very limited range, less than a feet or two.	Can be read up to several feet.
Cost	Cheap	Expensive, but likely to cost less as more industries adopt the technology.

- Privacy Concerns → Invisible Nature
- RFIDsec (2007, Danish Company) → Passive RFID with built in Firewall, access control, communication encryption, silent mode (exclusive control)
  - No response from RFID unless consumer authorizes
  - Consumer validates RFID without leaking Identifier
  - RFID without being trackable

Contemporary Software Trends

- Operating System
  - The software that manages and controls the computer's activities



- OS and Graphical User Interface
  - GUIs
  - Windows XP, Windows Vista, and Windows Server 2003
  - UNIX
  - Linux
    - **Open-source software**

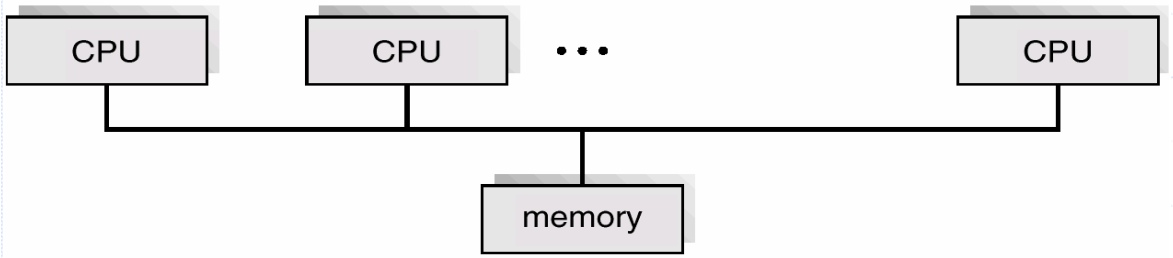
Parallel Systems

- Also known as *multiprocessor systems* or *tightly coupled systems*
- Operate with more than one CPU in close communication
- Processors share memory and a clock; communication usually takes place through the shared memory

- Advantages of Parallel Systems:
  - Higher throughput – increases but less than *N*
  - Economical – *Sharing*
  - Higher reliability - graceful degradation, fault tolerant

Parallel Systems

- Symmetric multiprocessing (SMP)
  - Each processor runs and identical copy of the operating system
  - Many processes can run at once without performance deterioration
  - Most modern operating systems support SMP (*Windows NT, XP, Solaris, Digital UNIX, OS/2, Linux*)



Architecture of general SMP system

Parallel Systems

- Asymmetric multiprocessing
  - Each processor is assigned a specific task
  - Master processor schedules and allocated work to slave processors
  - More common in extremely large systems

Clustered Systems

- Unlike multiprocessor systems, two or more individual systems coupled together
- Clustering allows two or more systems to share storage, *definitions*
- Service even if one or more cluster fails
- Higher *reliability* by adding *redundancy*

- Asymmetric clustering
  - One machine is in *hot-standby* while other running applications
  - *Hot-standby* machine only monitors the active server
  - If server fails, *hot-standby* machine $\rightarrow$ active server

- Symmetric clustering
  - Two or more hosts are running applications and *monitoring each other*
  - More efficient

- Rapid advances in Cluster Technology, clustering over *WANs*, *SANs*

Real-Time Systems

- Special purpose OS with rigid time constraints
- Used in dedicated application, *e.g*. control systems, imaging systems, display systems
- Fixed time constraints, returns the correct results within its time-constraints
- Time-sharing systems, desirable but NOT mandatory
- Real-Time systems may be either *hard* or *soft* real-time

Real-Time Systems

- Hard real-time
  - *Guarantees* completion of critical task in time
  - Secondary storage limited or absent, data stored in short term memory, or read-only memory (ROM)
  - Conflicts with time-sharing systems, not supported by general-purpose operating systems
- Soft real-time
  - Less restrictive, critical real-time task has priority
  - Retains this priority until its completion
  - Lack of deadline support, risky for robotics/control
  - Useful in applications (multimedia, virtual reality) requiring advanced OS features (not supported in hard real-time systems)

Programming Languages

- Example of Four Levels of Programming Languages
  - Sum of two numbers → X = Y + Z

Four Levels of Programming Languages	
• **Machine Languages:** Use binary coded instructions 1010 11001 1011 11010 1100 11011	• **High-Level Languages:** Use brief statements or arithmetic notations BASIC: X = Y + Z COBOL: COMPUTE X = Y + Z
• **Assembler Languages:** Use symbolic coded instructions LOD Y ADD Z STR X	• **Fourth-Generation Languages:** Use natural and nonprocedural statements SUM THE FOLLOWING NUMBERS

  - Natural Languages → 5th generation languages
    - Close to English, INTELLECT
    - 'What is the average exam score in COMP 5131?'

Programming Languages

- Application programming languages for business
  - COBOL
  - C, C++
  - Visual Basic: Visual programming language

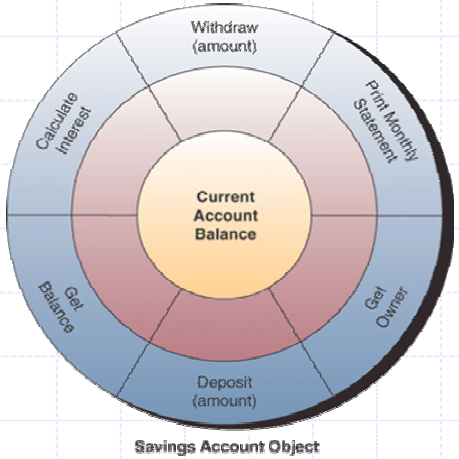
- Summary – 4th generation languages
  - Software tools that enable end-users to develop software applications with minimum or no technical help
  - Tend to be nonprocedural (*what*) or less procedural than conventional programming (*how*) language
  - Some of these use natural languages (speech)
  - Seven Categories – Top (end user) to Bottom (IS Pros)

Programming Languages

Tool	Description	Example
PC software tools	General-purpose software packages for PCs	WordPerfect Microsoft Access
Query language	Languages for retrieving data stored in databases or files	SQL
Report generator	Specialized tools for creating highly customized reports	Crystal Reports
Graphics language	Display data from databases in graphic format	SAS Graph Systat
Application generator	Preprogrammed modules to generate entire applications	FOCUS Microsoft FrontPage
Application software package	Software programs that eliminate need for custom, in-house software	Oracle PeopleSoft HCM mySAP ERP
Very high-level programming language	Generate program code with fewer instructions than conventional languages	APL Nomad2

Programming Languages

- Object Oriented Languages
  - Also considered as 5th generation languages (Visual Basic, C++, Java)
  - Most Programming Languages
    - Separate data elements from procedures/action that will be performed on them
  - Object oriented languages tie them together into objects
    - Object → Data and Actions
    - Example → An object could be data about a *bank account* and the procedures performed on it such as *interest calculations*



  - Most widely used software development languages today
  - Easier to use and more efficient for *graphics-oriented user interfaces*
  - Reusable: can use an object from one application in another application
  - Example → Visual Basic, C++, Java

Programming Languages

- Java
  - Operating system-independent, processor-independent, object-oriented programming language
  - JVM → Built for each type of computer and OS
  - Nearly all web browser software has Java platform built in
  - Applets → Miniature java programs, resides on centralized server
- AJAX (Asynchronous JavaScript and XML)
  - Allows a client and server to exchange data behind the scenes to avoid reloading a Web page after each change, *Google Maps*
  - Uses JavaScript programs downloaded to clients → Continuity
- Hypertext markup language (HTML)
  - Page description language for specifying how elements are placed on a Web page and for creating links to other pages and objects
  - Custom written or created using HTML authorizing capabilities

Software for Web

- Web Services
  - Set of loosely coupled software components that exchange information regardless of OS or programming language

- XML (eXtensible Markup Language)
  - Designed to transport and store data, with focus on what data is
  - HTML was designed to display data, with focus on how data looks

```
<?xml version="1.0"?>
<person>
 <name>
    <firstname>Ajay</firstname>
    <lastname>Kumar</lastname>
 </name>
 <job> Professor</job>
 <gender>Male</gender>
</person>
```

  - XML standard → Created by W3C
    - Easy to use and standardized way to store self-describing data
    - Data that describes both its content and its structure
  - Web services communicates through XML messages → web protocols
    - SOAP (simple object access protocol)
    - WSDL (web services description language)
    - UDDI (universal description, discovery, and integration)
  - Service Oriented Architecture (SOA)

Software Trends

- Open Source Software
  - Mandriva, Ubuntu
- Cloud Computing
  - Web-based applications that are stored and accessed via 'internet'
  - Software and Data → Hosted on powerful servers, internet access
  - Comparison/Confusion
    - Grid Computing → Form of distributed computing, whereby a *super and virtual computer* is composed of a cluster of networked, loosely-coupled computers, acting in concert to perform very large tasks
    - Utility Computing → Packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility as electricity
  - Many cloud computing deployments are today powered by grids, have autonomic characteristics and are billed like utilities
  - Cloud Computing → Seen as a *natural next step* from the *grid-utility model*
  - Companies
    - Google - Google Apps desktop productivity tools
    - Microsoft – Windows Live, Office Live
    - Salesforce, Yahoo!, IBM, SAP
  - Gartner, Aug. 2008
    - *"Organizations are switching from company-owned hardware and software assets to per-use service-based models"*
    - *"Projected shift to cloud computing will result in dramatic growth in IT products in some areas and in significant reductions in other areas"*

Software Trends

- Mashups
  - Web application that combines data from more than one source into single application, web 2.0
  - Content for mashups → Typically sourced from a third party via a public interface or API (web services)
  - Not simple embedding, mashup site should itself access third party data using an API, and process that data in some way to increase its value
  - Examples
    - *youMashTube* → End-users make video mashups using YouTube's API
    - ChicagoCrime.org combines Google Maps map with crime data

- Widgets
  - Small software programs that can be added into any html based webpage or desktop to provide additional functionality
  - web widgets → gadget, badge, module, webjit, capsule, snippet, mini, flake
  - Example → Flixter widget on Facebook profiles transports users to place where they can list films, ratings, reviews, etc.
  - Desktop widgets integrate contents from an external source into user's desktop to provide services (calculator, current weather conditions, etc.)
  - Microsoft Windows Sidebar (Vista), Google Desktop Gadgets provide widgets

Managing Technology

Capacity Planning

- Process of predicting when hardware system becomes saturated
- Ensuring firm has enough computing power for current and future needs
- Factors include:
  - Maximum number of users at a time
  - Impact of current, future software applications
  - Performance measures (e.g. minimum response time)

Scalability

- Ability of system to expand to serve large number of users without breaking down

Managing Technology

- Total Cost of Ownership (TCO) model
  - Used to analyze direct and indirect costs to help determine actual cost of owning a specific technology
    - Direct costs: Hardware, software purchase costs
    - Indirect costs: Ongoing administration costs, upgrades, maintenance, technical support, training, utility and real estate costs
    - Hidden costs: Support staff, downtime, additional network management
  - TCO can be reduced through increased centralization, standardization of hardware and software resources

Using Technology Service Providers

- Outsourcing
  - Using external service provider to
    - Run networks
    - Host, manage Web site(s)
    - Develop software (**offshore software outsourcing**)
    - Manage IT infrastructures
  - Requires **Service Level Agreements (SLAs)**
- On-demand computing (utility computing)
  - Firms off-loading peak demand for computing power to remote, large-scale data processing centers
  - Example → Amazon.com {also from IBM, HP, Sun microsystems)
    - Elastic Compute Cloud (EC2)
      - 1.7 GHz x86 Processor, 1.75GB, 250 GB, 250 Mbps BW → 10 cents per instance/hr
    - Simple Storage Service (S3)
      - 15 cents/GB storage per month, 20 cents/GB data transfer, 99.9% availability
- Software as a Service (SaaS)
  - Firms rent software functions from Web-based services, with users paying, subscription or per-transaction basis
  - Example → Salesforce.com, on-demand service CRM, automation

Using Technology Service Providers

- Managing software localization for global business
  - Local language interfaces
    - English not typically standard at middle, lower levels
    - Interfaces are complex: Menu bars, error messages, online forms, search results, etc.
  - Differences in local cultures
  - Differences in business processes

- All of these factors add to TCO of using technology service providers