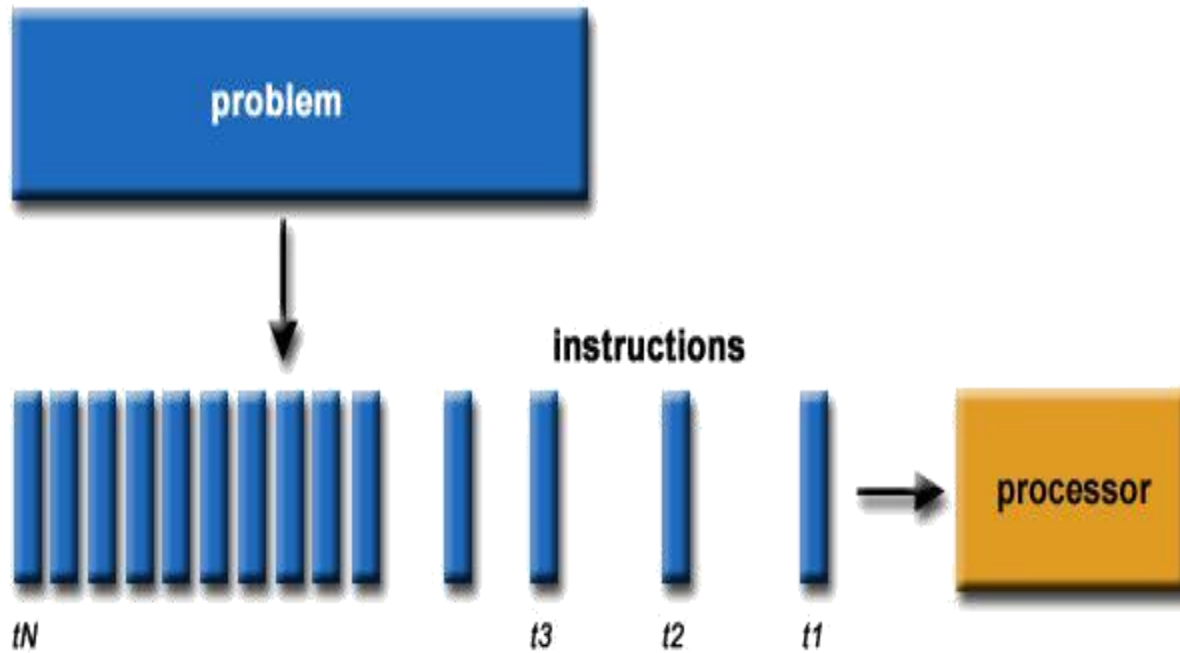# Cloud Computing
Unit 1

---NotesHub---

- **Serial Computing:**
- Traditionally, software has been written for *serial* computation:
  - A problem is broken into a discrete series of instructions
  - Instructions are executed sequentially one after another
  - Executed on a single processor
 - Only one instruction may execute at any moment in time

- **Parallel Computing:**
- In the simplest sense, *parallel computing* is the simultaneous use of multiple compute resources to solve a computational problem:
  - A problem is broken into discrete parts that can be solved concurrently
  - Each part is further broken down to a series of instructions
  - Instructions from each part execute simultaneously on different processors
  - An overall control/coordination mechanism is employed
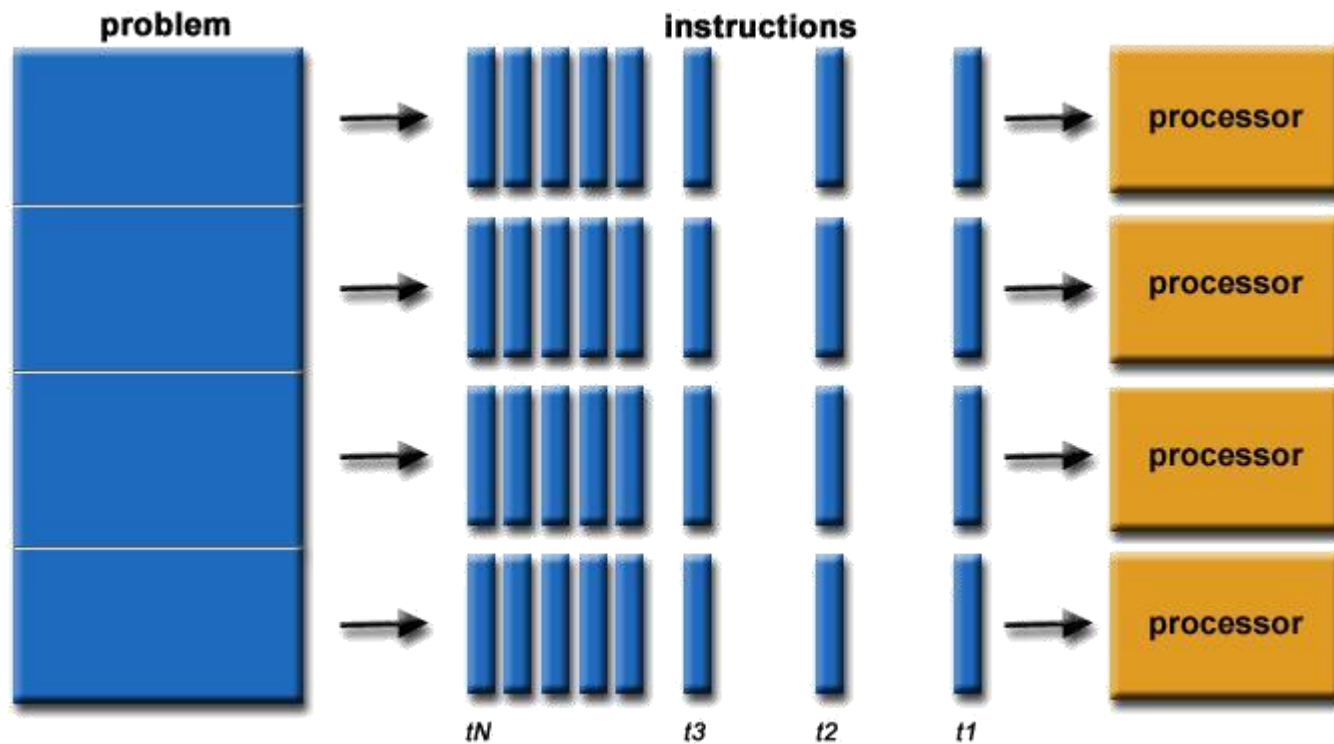
- The computational problem should be able to:

  - Be broken apart into discrete pieces of work that can be solved simultaneously;
  - Execute multiple program instructions at any moment in time;
  - Be solved in less time with multiple compute resources than with a single compute resource.

- The compute resources are typically:
  - A single computer with multiple processors/cores
  - An arbitrary number of such computers connected by a network

• Networks connect multiple stand-alone computers (nodes) to make larger parallel computer clusters.
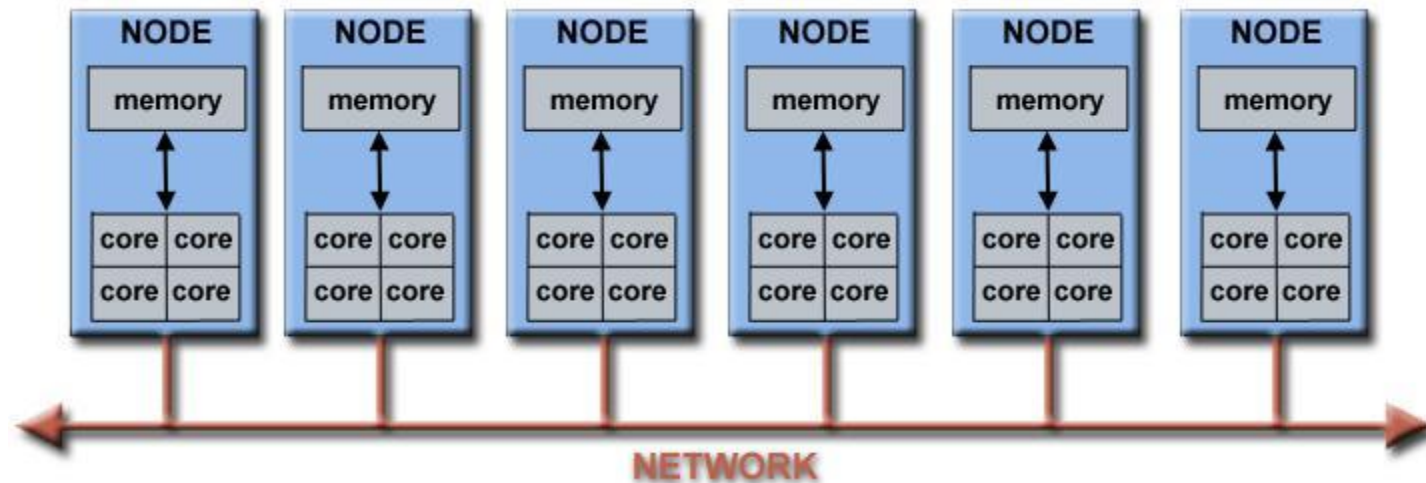
NEED OF PARALLEL COMPUTING

1.Save time and save large problems:

  With the progress of computer science, computational speed of the processors has also increased many a timeSo we have to look for alternatives. The answer lies in parallel computing. It is obvious that with the increase in number of processors working in parallel, computation time is bound to reduce. Also, they're some scientific problems that even the fastest processor to takes months or even years to solve. However, with the application of parallel computing these problems may be solved in a few hours

2. Cost savings:
   We can use multiple cheap computing resources instead of paying ii) heavily for a supercomputer.

3. Overcoming memory constraints:
   Single computers have very finite memory resources. For large problems, using the memories of multiple computers may overcome this obstacle. So if we combine the memory resources of multiple computers then we can easily fulfill the memory requirements of the large-size problems.

4. Limits to serial computing: Both physical and practical factors pose significant constraints to simply building ever faster serial computers. The speed of a serial computer is directly dependent upon how fast data can move through hardware. Increasing speeds necessitate increasing proximity of processing elements. Secondly, processor technology is allowing an increasing number of transistors to be placed on a chip. However, even with molecular or atomic-level components, a limit will be reached on how small components can be made. It is increasingly expensive to make a single processor faster. Using a larger number of moderately fast commodity processors to achieve the same (or better) performance is less expensive

The Grid

"Resource sharing & coordinated problem solving in dynamic, multi-institutional virtual organizations"
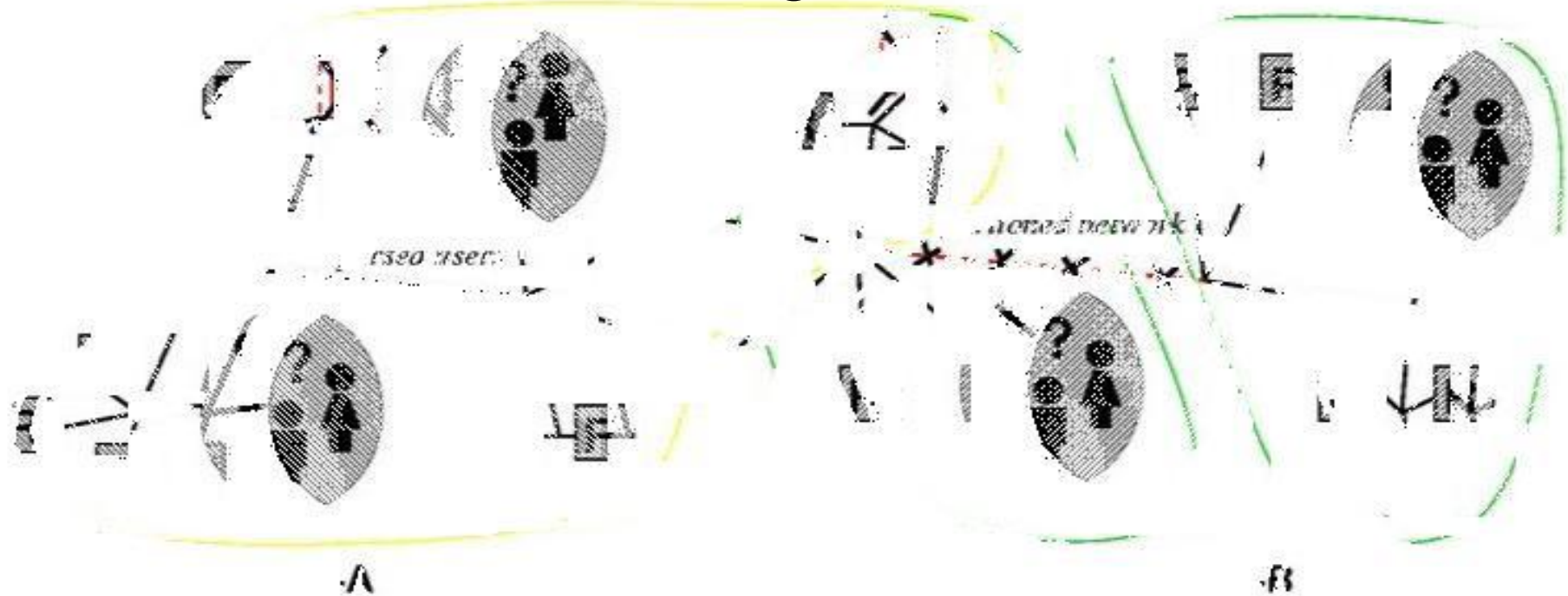
A Comparison

SERIAL

PARALLEL

GRID

Fetch/Store	**Fetch/Store**	**Fetch/Store**
Compute	**Compute/ communicate**	**Discovery of Resources**
	Cooperative game	**Interaction with remote application**
		Authentication / Authorization
		Security
		Compute/Communicate
		Etc

Web vs. GRID

- Web
  - **Uniform naming access to documents**

http://

http://
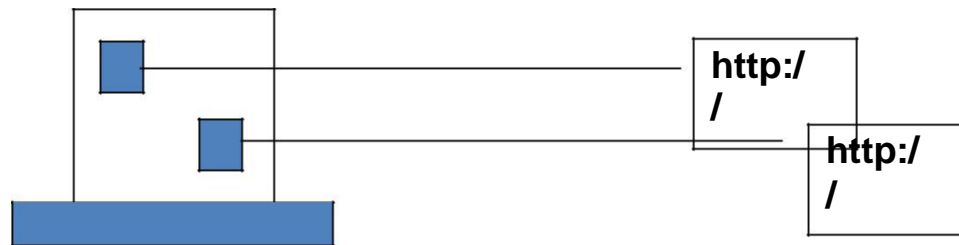
- Grid

Software Uniform, high performance access to computational resources
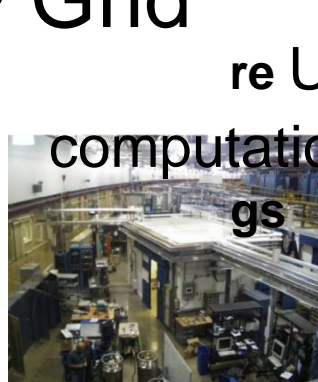
Catalo gs

Colleges/R&D Labs

Sensor nets

Why Grids?

- Large-scale science and engineering are done through the interaction of people, heterogeneous computing resources, information systems, and instruments, all of which are geographically and organizationally dispersed.
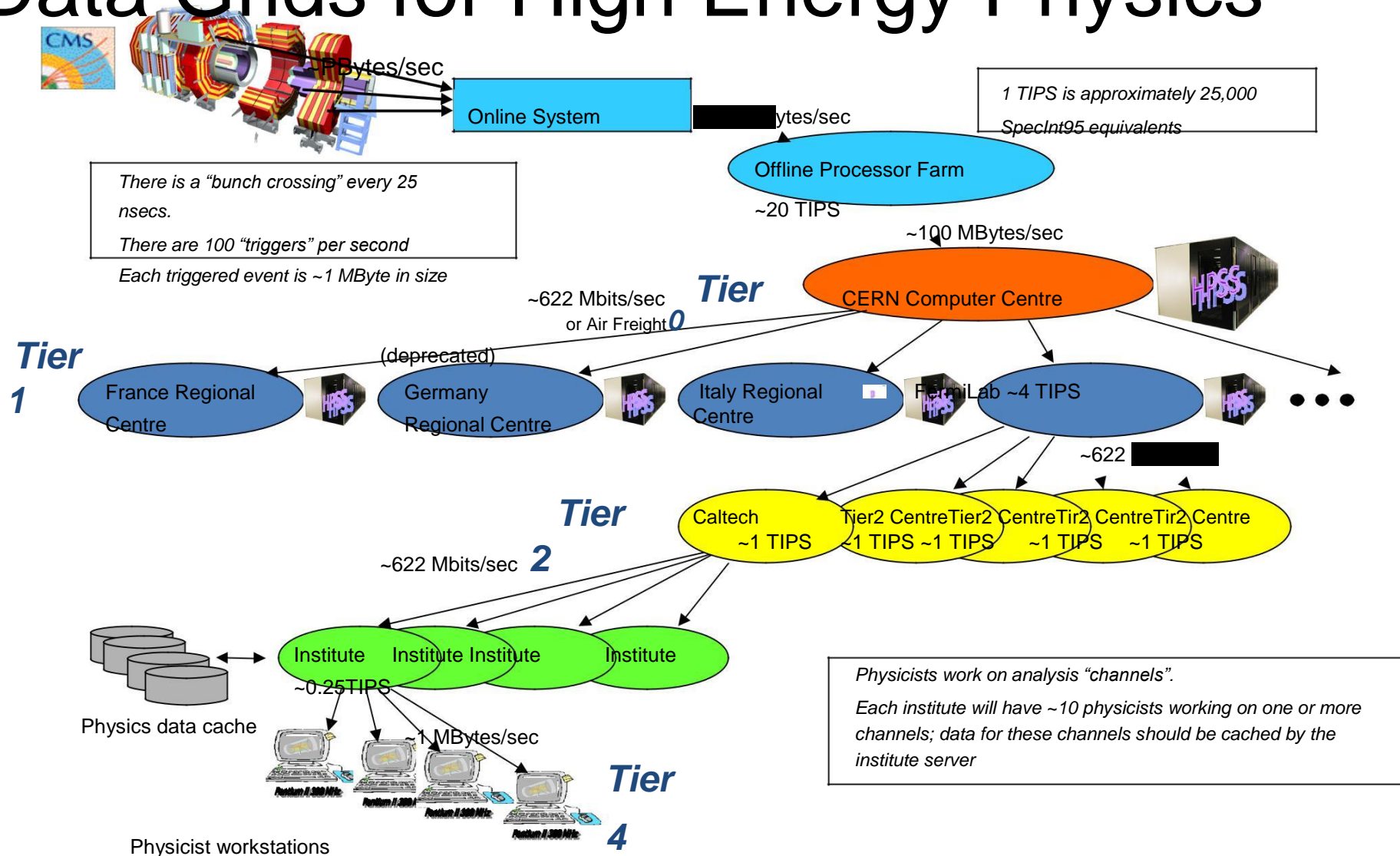

- The overall motivation for "Grids" is to facilitate the routine interactions of these resources in order to support large-scale science and Engineering.

Grid Communities & Applications:
Data Grids for High Energy Physics

CMS

~PBytes/sec

Online System

~ytes/sec

1 TIPS is approximately 25,000 SpecInt95 equivalents

There is a "bunch crossing" every 25 nsecs.

There are 100 "triggers" per second

Each triggered event is ~1 MByte in size

Offline Processor Farm

~20 TIPS

~100 MBytes/sec

Tier 0

~622 Mbits/sec
or Air Freight (deprecated)

CERN Computer Centre

HPSS

Tier 1

France Regional Centre — HPSS

Germany Regional Centre — HPSS

Italy Regional Centre — HPSS

FermiLab ~4 TIPS — HPSS

• • •

~622

Tier 2

~622 Mbits/sec

Caltech ~1 TIPS

Tier2 Centre ~1 TIPS

Tier2 Centre ~1 TIPS

Tier2 Centre ~1 TIPS

Tier2 Centre ~1 TIPS

Institute ~0.25TIPS

Institute

Institute

Institute

Physics data cache

~1 MBytes/sec

Physicists work on analysis "channels".

Each institute will have ~10 physicists working on one or more channels; data for these channels should be cached by the institute server

Tier 4

Physicist workstations

www.griphyn.org   www.ppdg.net   www.eu-datagrid.org

The Grid Opportunity: eScience and eBusiness

- Physicists worldwide pool resources for peta-op analyses of petabytes of data
- Civil engineers collaborate to design, execute, & analyze shake table experiments
- An insurance company mines data from partner hospitals for fraud detection
- An application service provider offloads excess load to a compute cycle provider
- An enterprise configures internal & external resources to support eBusiness workload

What is Distributed System?

A *distributed system* is one in which **components** located at networked computers communicate and **coordinate** their actions only by passing **messages**. This definition leads to the following characteristics of distributed systems:

- Concurrency of components
- Lack of a global 'clock'
- Independent failures of components

Centralized System Features

- One component with non-autonomous parts

- Component shared by users all the time

- All resources accessible

- Software runs in a single process

- Single point of control

- Single point of failure

Distributed System Features

- Multiple autonomous components

- Components are not shared by all users

- Resources may not be accessible

- Software runs in concurrent processes on different processors

- Multiple points of control

- Multiple points of failure

Examples Of Distributed System

- Local Area Network and Intranet

- Database Management System

- Automatic Teller Machine Network

- Internet/World-Wide Web

- Mobile and Ubiquitous Computing

Advantages and Disadvantages of Distributed Systems

- Advantages
  - Shareability
  - Expandability
  - Local autonomy
  - Improved performance
  - Improved reliability and availability
  - Potential cost reductions
- Disadvantages
  - Network reliance
  - Complexities
  - Security
  - Multiple point of failure

- Grid Computing
  1: Loosely coupled(Decentralization)
  2: Diversity and Dynamism
  3: Distributed Job Management&
  scheduling
- Distributed Computing
  Is to solve a single large problem by breaking
  it down into several tasks where each task is
  computed in the individual computers of the
  distributed system.

DISTRIBUTED COMPUTING VARIANTS

- Client Server
- Proxy Sever
- Peer to Peer
- Services provided by multiple servers
- Mobile code
- Mobile agents
- Network computers
- Thin clients
- Mobile devices
- MANETS

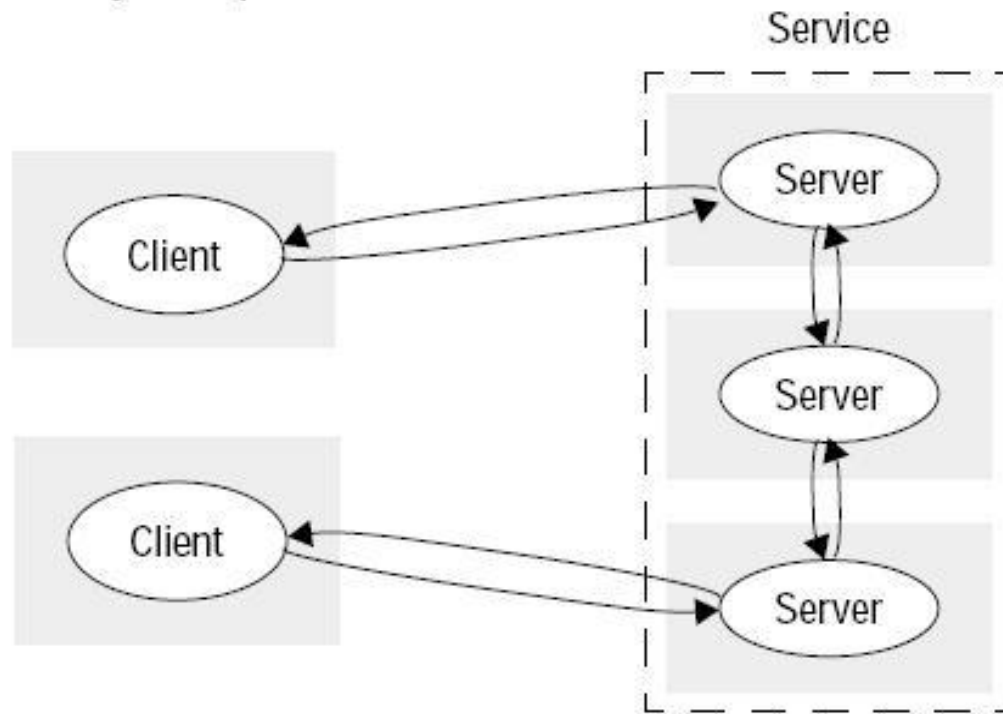# 1.Client-Server Architecture

Clients invoke individual servers
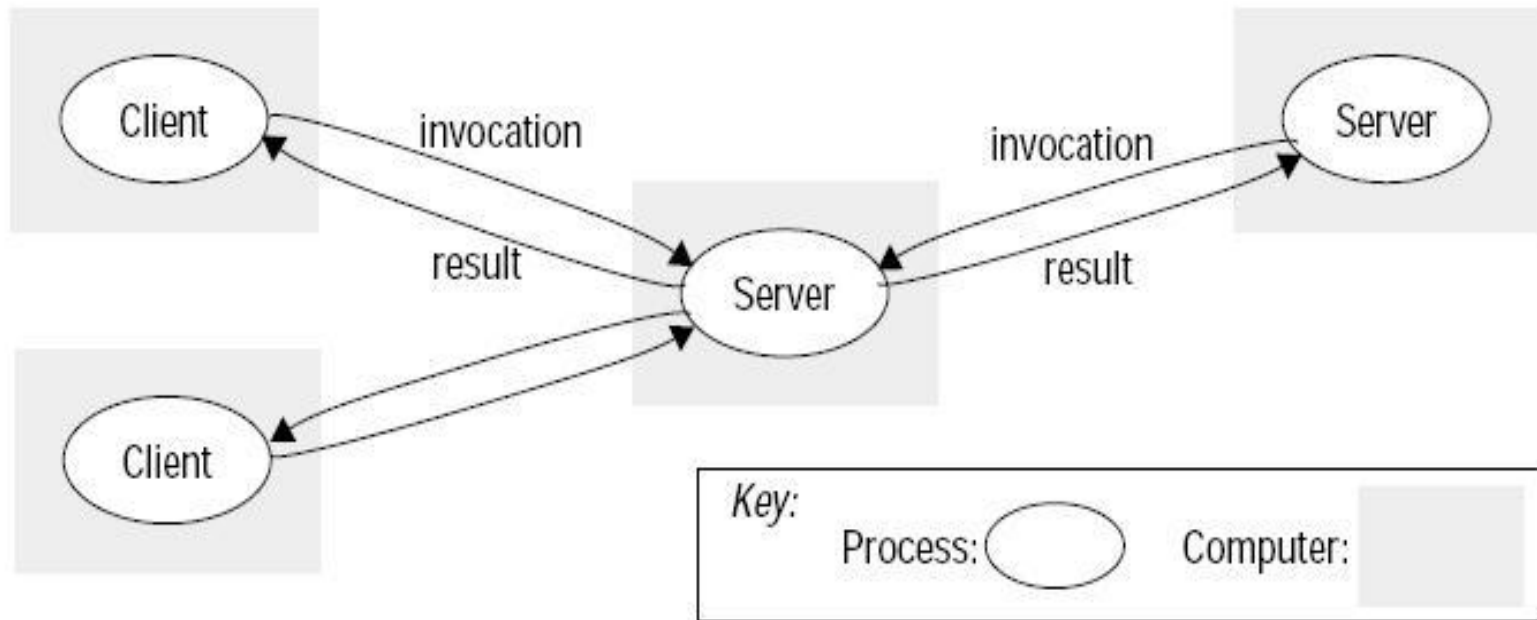
2.Peer-to-Peer Architecture

- All processes play similar roles – i.e., they interact as *peers*
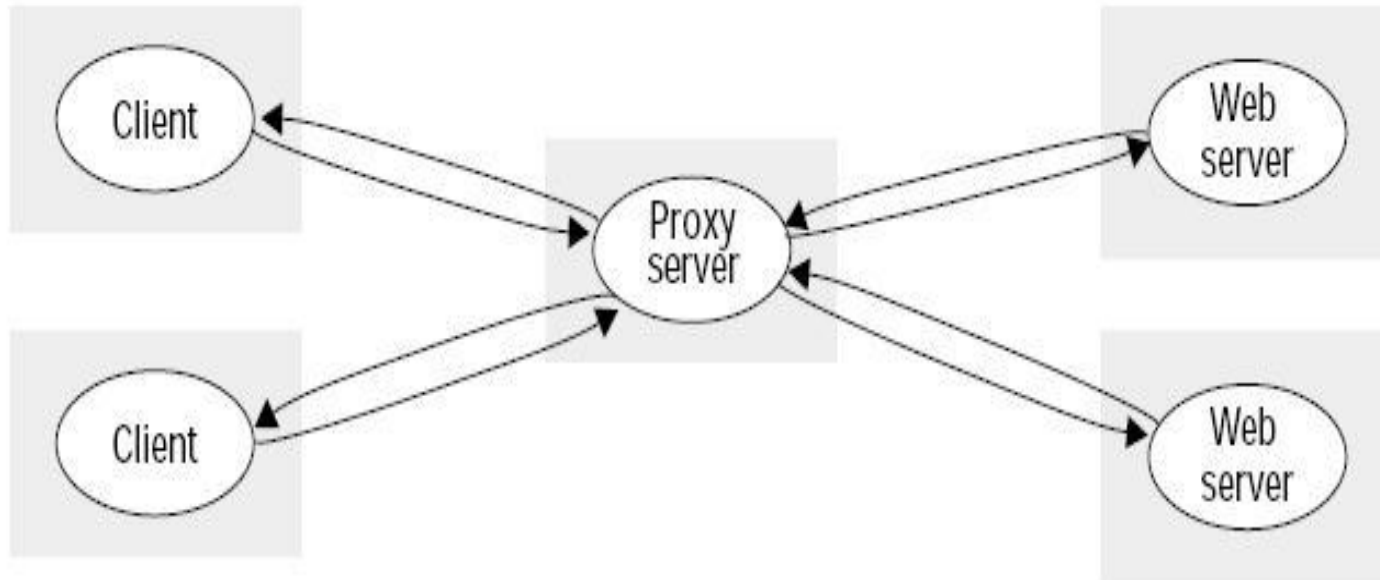- No central component – potentially better scalability and resiliency to failures

Services by multiple servers

A service provided by multiple servers
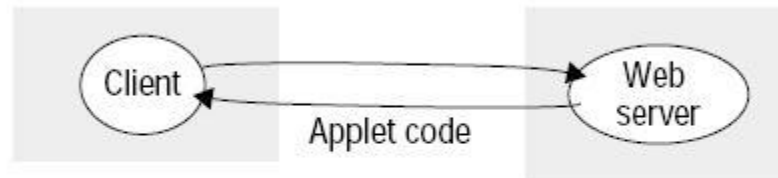
Proxy Servers and Caches

Web proxy server

Mobile Code

Web applets

a) client request results in the downloading of applet code

b) client interacts with the applet

Mobile Agents

- A running program (both code and data) that travels from one computer to another

- Example: a worm

Network Computers

- Does not rely on locally installed software
- Downloads operating system and applications from a remote computer
- Applications are run locally, but files are managed on a remote server
- Users can migrate from one network computer to another

Thin Clients

- Similar to a network computer
- Instead of downloading code to the user computer, it runs it on a compute server
- Software layer provides a window-based interface to the client (X Windows)

MANETS

- A **mobile ad hoc network** (**MANET**) is a continuously self-configuring, infrastructure-less network of mobile devices connected wirelessly.
- Each device in a MANET is free to move independently in any direction, and will therefore change its links to other devices frequently. Each must forward traffic unrelated to its own use, and therefore be a router. The primary challenge in building a MANET is equipping each device to continuously maintain the information required to properly route traffic. Such networks may operate by themselves or may be connected to the larger Internet. They may contain one or multiple and different transceivers between nodes. This results in a highly dynamic, autonomous topology.

AUTOMATIC COMPUTING

- **Autonomic computing** refers to the self-managing characteristics of distributed computing resources, adapting to unpredictable changes while hiding intrinsic complexity to operators and users.

CHARACTERSTICS OF AUTOMATIC COMPUTING

- 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 the defined requirements
- Self-protection: Proactive identification and protection from arbitrary attacks.

MODEL OF AUTOMATIC COMPUTING

- A fundamental building block of an autonomic system is the sensing capability (*Sensors S_i*), which enables the system to observe its external operational context. Inherent to an autonomic system is the knowledge of the *Purpose* (intention) and the *Know-how* to operate itself (without external intervention.

- The actual operation of the autonomic system is dictated by the *Logic*, which is responsible for making the right decisions to serve its *Purpose*, and influence by the observation of the operational context (based on the sensor input).

- This model highlights the fact that the operation of an autonomic system is purpose-driven. This includes its mission (e.g., the service it is supposed to offer), the policies (e.g., that define the basic behaviour).

-
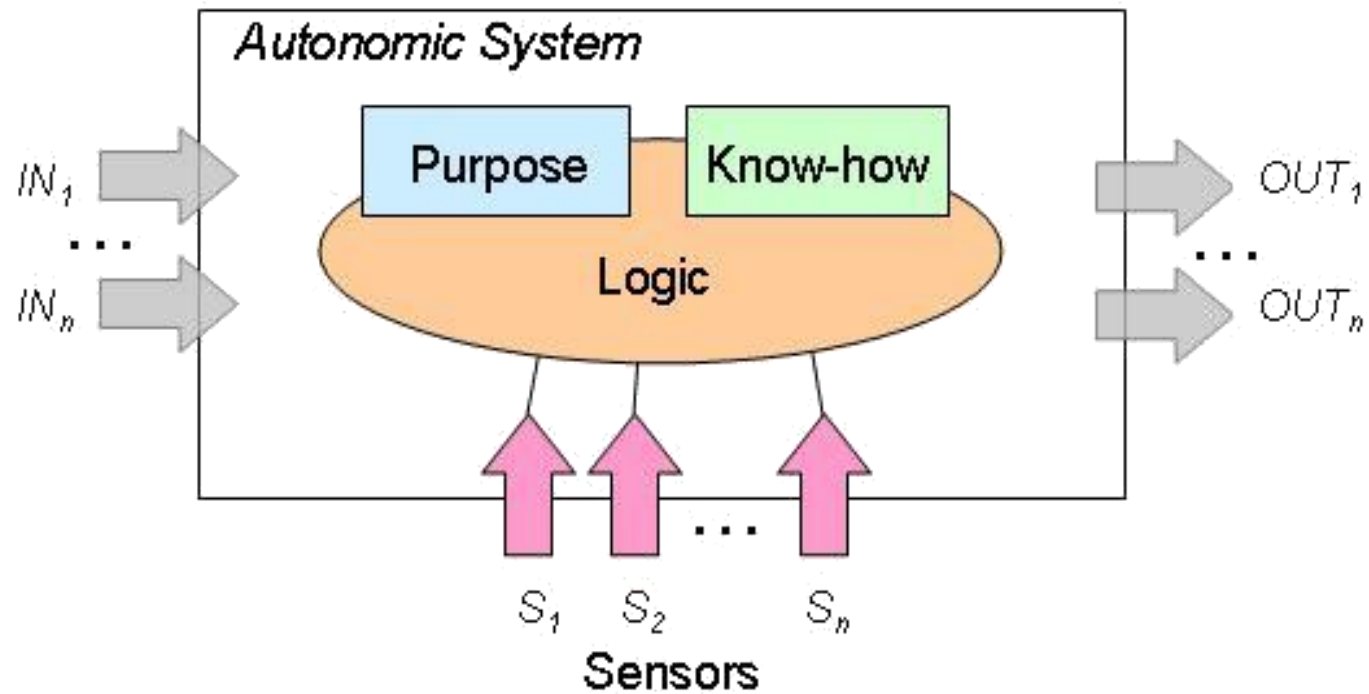
Cloud Computing

- Cloud Computing provides us a means by which we can access the applications as utilities, over the Internet. It allows us to create, configure, and customize applications online
- The term Cloud refers to a Network or Internet. In other words, we can say that Cloud is something, which is present at remote location. Cloud can provide services over network, i.e., on public networks or on private networks, i.e., WAN, LAN or VPN. Applications such as e-mail, web conferencing, customer relationship management (CRM),all run in cloud.

- Buyya's definition
  - "A Cloud is a type of parallel and distributed system consisting of a collection of inter-connected and **virtualised** computers that are **dynamically** **provisioned** and presented as one or more unified computing resources based on **service-level agreements** established through **negotiation** between the service provider and consumers."

- Cloud Computing refers to manipulating, configuring, and accessing the applications online. It offers online data storage, infrastructure and application
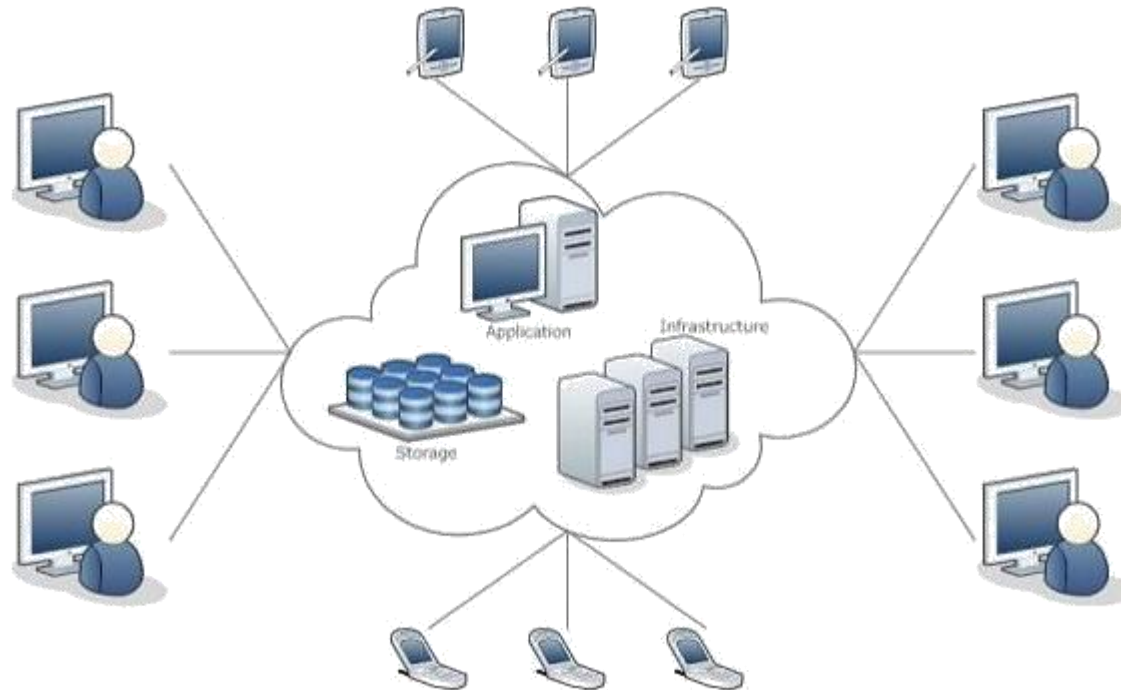
What is Cloud Computing
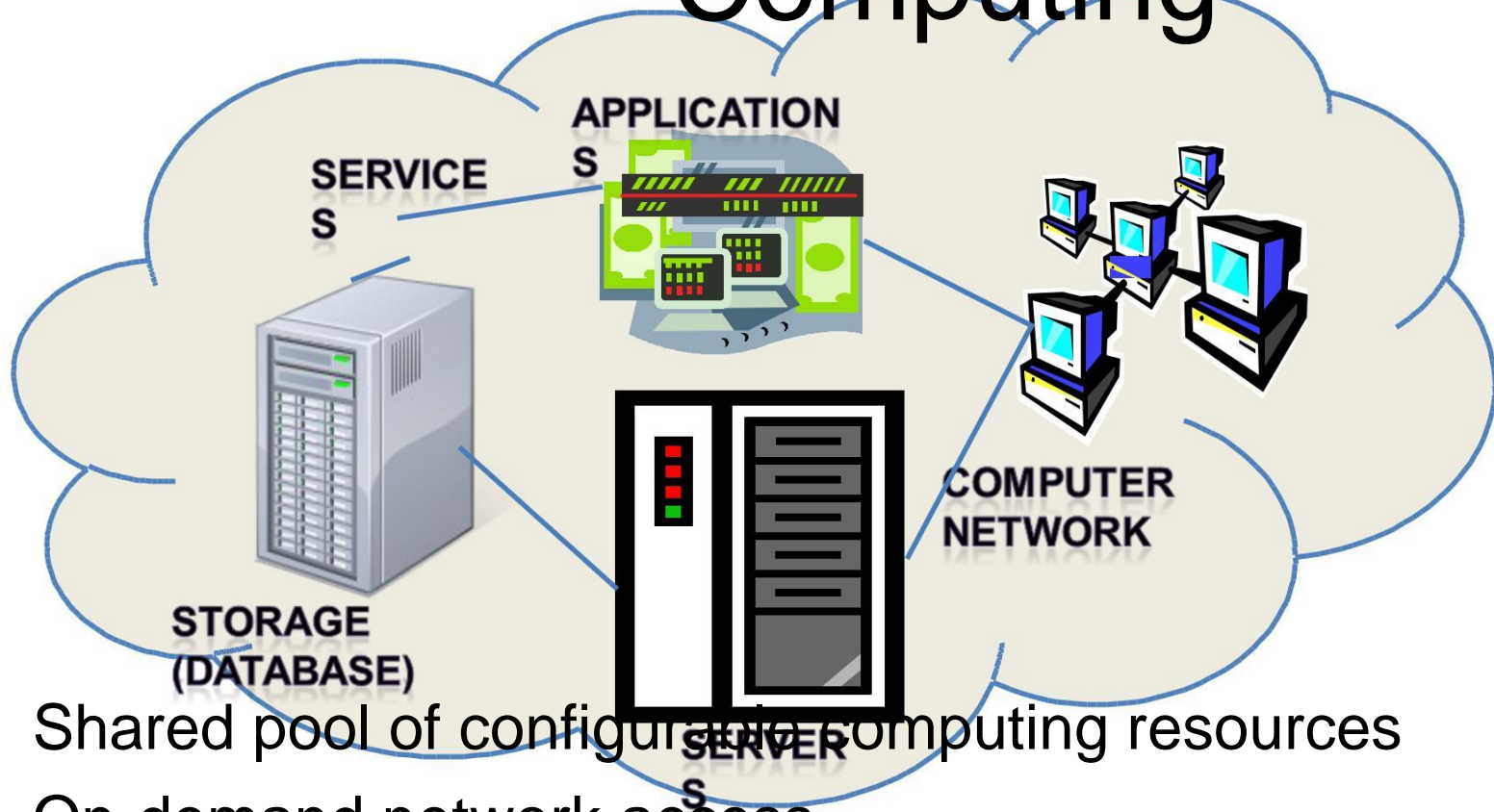
- Shared pool of configurable computing resources
- On-demand network access
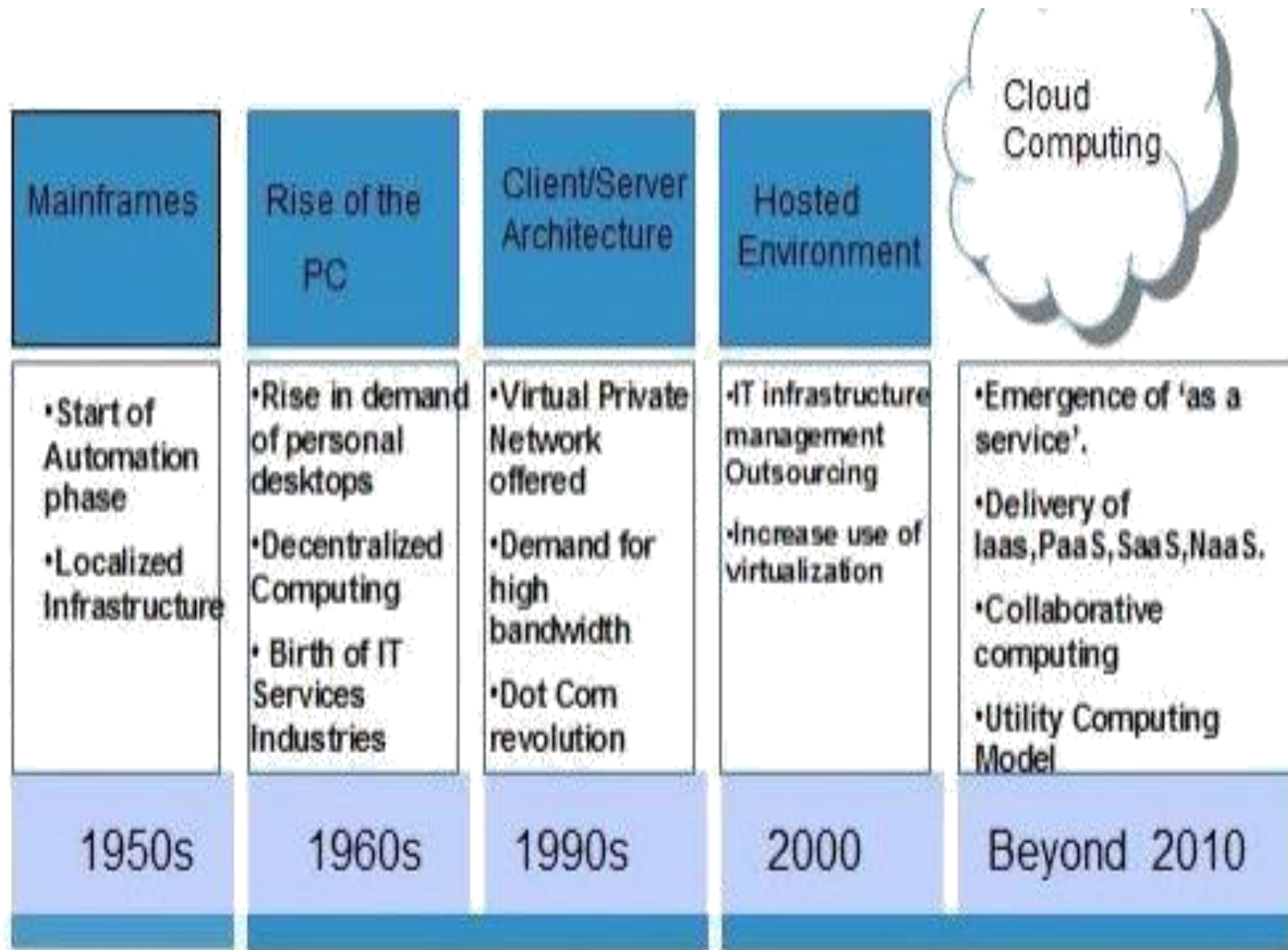- Provisioned by the Service Provider

40

EVOLUTION OF CLOUD COMPUTING

- he concept of Cloud Computing came into existence in 1950 with implementation of mainframe computers, accessible via thin/static clients. Since then, cloud computing has been evolved from static clients to dynamic ones from software to services.

| Mainframes	Rise of the PC	Client/Server Architecture	Hosted Environment	Cloud Computing
•Start of Automation phase •Localized Infrastructure	•Rise in demand of personal desktops •Decentralized Computing • Birth of IT Services Industries	•Virtual Private Network offered •Demand for high bandwidth •Dot Com revolution	•IT infrastructure management Outsourcing •Increase use of virtualization	•Emergence of 'as a service'. •Delivery of Iaas,PaaS,SaaS,NaaS. •Collaborative computing •Utility Computing Model
1950s	1960s	1990s	2000	Beyond 2010

BENEFITS OF CLOUD

- One can access applications as utilities, over the Internet.
- Manipulate and configure the application online at any time.
- It does not require to install a specific piece of software to access or manipulate cloud application.
- Cloud Computing offers online development and deployment tools, programming runtime environment through Platform as a Service model.
- Cloud resources are available over the network in a manner that provides platform independent access to any type of clients.

- Cloud Computing offers on-demand self-service. The resources can be used without interaction with cloud service provider.
- Cloud Computing is highly cost effective because it operates at higher efficiencies with greater utilization. It just requires an Internet connection.
- Cloud Computing offers load balancing that makes it more reliable

Vision of Cloud Computing

→ Cloud computing provides the facility to provision virtual hardware, runtime environment and services to a person having money.

→ These all things can be used as long as they are needed by the user, there is no requirement for the upfront commitment.

→ The whole collection of computing system is transformed into a collection of utilities, which can be provisioned and composed together to deploy systems in hours rather than days, with no maintenance costs.

→ The long term vision of a cloud computing is that IT services are traded as utilities in an open market without technological and legal barriers.

Vision of Cloud Computing
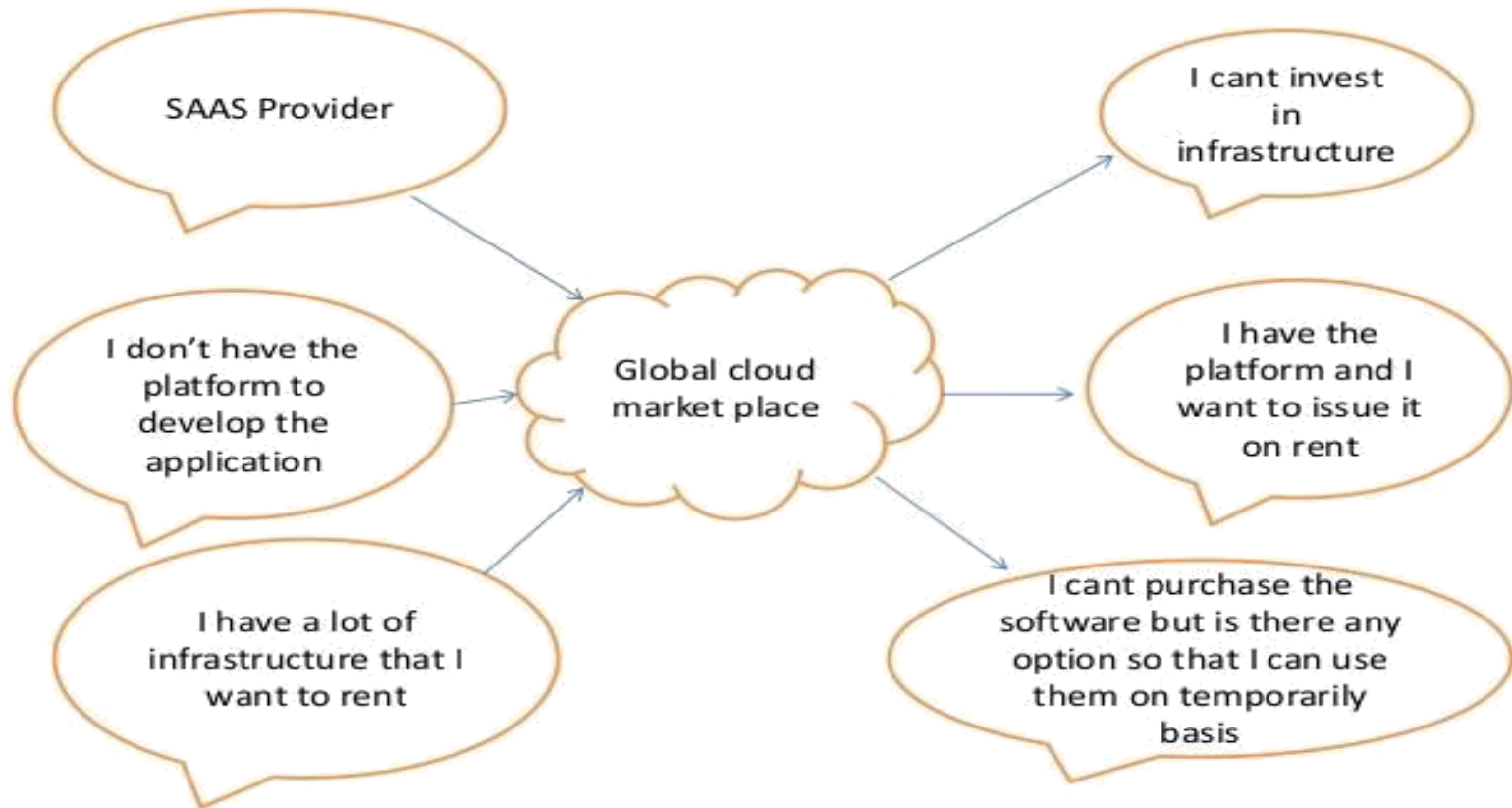
Challenges of cloud computing

- **Security and Privacy**

  The main challenge to cloud computing is how it addresses the security and privacy concerns of businesses thinking of adopting it. The fact that the valuable enterprise data will reside outside the corporate firewall raises serious concerns. **Hacking and various attacks** to cloud infrastructure would affect multiple clients even if only one site is attacked. These risks can be mitigated by using security applications, encrypted file systems, data loss software, and buying security hardware to track unusual behavior across servers.

2. Service Delivery and Billing

- It is difficult to assess the costs involved due to the on-demand nature of the services. Budgeting and assessment of the cost will be very difficult unless the provider has some good and comparable benchmarks to offer. The service-level agreements (SLAs) of the provider are not adequate to guarantee the availability and scalability. Businesses will be reluctant to switch to cloud without a strong service quality guarantee.

3. Interoperability and Portability

- Businesses should have the leverage of migrating in and out of the cloud and switching providers whenever they want, and there should be no lock-in period. Cloud computing services should have the capability to integrate smoothly with the on-premise IT.

4. Reliability and Availability

- Cloud providers still lack round-the-clock service; this results in frequent outages. It is important to monitor the service being provided using internal or third-party tools. It is vital to have plans to supervise usage, SLAs, performance, robustness, and business dependency of these services.

5. Performance and Bandwidth Cost

- Businesses can save money on hardware but they have to spend more for the bandwidth. This can be a low cost for smaller applications but can be significantly high for the data-intensive applications. Delivering intensive and complex data over the network requires sufficient bandwidth. Because of this, many businesses are waiting for a reduced cost before switching to the cloud.

Applications Of Cloud Computing

- Email Communication
- No Need of local data storage-Data stored on your home or business computer suffers from many of the same restrictions as email and, as with email, the cloud offers a solution. Storing your MP3's, video, photos and documents online instead of at home gives you the freedom to access them wherever you can find the means to get online.

- Are You Collaborator –Some occasion you may find yourself in need of the opinion of your peers. Downloading files onto flash memory, emailing documents to friends or family or colleagues or sending submissions by snail mail is so last century. Last year Google launched a service that allowed groups of people to work on the same document, idea or proposal in real time or whenever convenient to each participant. Using Google Wave you can create a document and then invite others to comment, amend, offer opinion, or otherwise join in with the creation of the final draft.

- Working in Virtual Office -Google's online suite of office applications is probably the best known but by no means the only solution on offer. Rather than having a system and space hogging suite of applications like a word processor, a spreadsheet creator and a presentation or publishing platform sitting on your computer, you could opt to work online instead. Accessibility, potential for collaboration and perhaps even online storage are just some of the benefits of satisfying your office suite needs by working online.

- Need extra processing

  Rather than purchasing servers, software, network equipment and so on, users would buy into a fully outsourced set of online services instead.

  Most cloud environments on offer can customize the kind of service provided to exactly suit the needs of the user. If you need more processing power from time to time, a cloud-based infrastructure, being scalable, negates the need for up-front investment in client-owned resources.

  Other service providers include the open source AbiCloud, Elastichosts and NASA's Nebula platform.

Advantages of Cloud Computing

- Lower computer costs:
  - You do not need a high-powered and high-priced computer to run cloud computing's web-based applications.
  - Since applications run in the cloud, not on the desktop PC, your desktop PC does not need the processing power or hard disk space demanded by traditional desktop software.

  - When you are using web-based applications, your PC can be less expensive, with a smaller hard disk, less memory, more efficient processor...
  - In fact, your PC in this scenario does not even need a CD or DVD drive, as no software programs have to be loaded and no document files need to be saved.

55

Advantages of Cloud Computing

- Improved performance:
  - With few large programs hogging your computer's memory, you will see better performance from your PC.
  - Computers in a cloud computing system boot and run faster because they have fewer programs and processes loaded into memory…
- Reduced software costs:
  - Instead of purchasing expensive software applications, you can get most of what you need for free-ish!
    - most cloud computing applications today, such as the Google Docs suite.
  - better than paying for similar commercial software
    - which alone may be justification for switching to cloud applications.

56

Advantages of Cloud Computing

- Instant software updates:
  - Another advantage to cloud computing is that you are no longer faced with choosing between obsolete software and high upgrade costs.

  - When the application is web-based, updates happen automatically
    - available the next time you log into the cloud.
  - When you access a web-based application, you get the latest version

    - without needing to pay for or download an upgrade.

- Improved document format compatibility.
  - You do not have to worry about the documents you create on your machine being compatible with other users' applications or OSes
  - There are potentially no format incompatibilities when everyone is sharing documents and applications in the cloud.

57

Advantages of Cloud Computing

- Unlimited storage capacity:
  - Cloud computing offers virtually limitless storage.
  - Your computer's current 1 Tbyte hard drive is small compared to the hundreds of Pbytes available in the cloud.

- Increased data reliability:
  - Unlike desktop computing, in which if a hard disk crashes and destroy all your valuable data, a computer crashing in the cloud should not affect the storage of your data.
    - if your personal computer crashes, all your data is still out there in the cloud, still accessible
  - In a world where few individual desktop PC users back up their data on a regular basis, cloud computing is a data-safe computing platform!

58

Advantages of Cloud Computing

- Universal document access:
  - That is not a problem with cloud computing, because you do not take your documents with you.
  - Instead, they stay in the cloud, and you can access them whenever you have a computer and an Internet connection

  - Documents are instantly available from wherever you are
- Latest version availability:
  - When you edit a document at home, that edited version is what you see when you access the document at work.
  - The cloud always hosts the latest version of your documents
    - as long as you are connected, you are not in danger of having an outdated version

59

Advantages of Cloud Computing

- Easier group collaboration:
  - Sharing documents leads directly to better collaboration.
  - Many users do this as it is an important advantages of cloud computing
    - multiple users can collaborate easily on documents and projects

- Device independence.
  - You are no longer tethered to a single computer or network.
  - Changes to computers, applications and documents follow you through the cloud.
  - Move to a portable device, and your applications and documents are still available.

60

Disadvantages of Cloud Computing

- Requires a constant Internet connection:
  - Cloud computing is impossible if you cannot connect to the Internet.
  - Since you use the Internet to connect to both your applications and documents, if you do not have an Internet connection you cannot access anything, even your own documents.
  - A dead Internet connection means no work and in areas where Internet connections are few or inherently unreliable, this could be a deal-breaker.

61

Disadvantages of Cloud Computing

- Does not work well with low-speed connections:
  - Similarly, a low-speed Internet connection, such as that found with dial-up services, makes cloud computing painful at best and often impossible.
  - Web-based applications require a lot of bandwidth to download, as do large documents.
- Features might be limited:
  - This situation is bound to change, but today many web-based applications simply are not as full-featured as their desktop-based applications.
    - For example, you can do a lot more with Microsoft PowerPoint than with Google Presentation's web-based offering

62

Disadvantages of Cloud Computing

- Can be slow:
  - Even with a fast connection, web-based applications can sometimes be slower than accessing a similar software program on your desktop PC.
  - Everything about the program, from the interface to the current document, has to be sent back and forth from your computer to the computers in the cloud.

  - If the cloud servers happen to be backed up at that moment, or if the Internet is having a slow day, you would not get the instantaneous access you might expect from desktop applications.

63

Disadvantages of Cloud Computing

- Stored data might not be secure:
  - With cloud computing, all your data is stored on the cloud.

    - The questions is How secure is the cloud?
  - Can unauthorised users gain access to your confidential data?
- Stored data can be lost:
  - Theoretically, data stored in the cloud is safe, replicated across multiple machines.
  - But on the off chance that your data goes missing, you have no physical or local backup.
    - Put simply, relying on the cloud puts you at risk if the cloud lets you down.

64

Disadvantages of Cloud Computing

- HPC Systems:
  - Not clear that you can run compute-intensive HPC applications that use MPI/OpenMP!
  - Scheduling is important with this type of application
    - as you want all the VM to be co-located to minimize communication latency!

- General Concerns:
  - Each cloud systems uses different protocols and different APIs
    - may not be possible to run applications between cloud based systems
  - Amazon has created its own DB system (not SQL 92), and workflow system (many popular workflow systems out there)
    - so your normal applications will have to be adapted to execute on these platforms.

65

The Future

- Many of the activities loosely grouped together under cloud computing have already been happening and centralised computing activity is not a new phenomena

- Grid Computing was the last research-led centralised approach
- However there are concerns that the mainstream adoption of cloud computing could cause many problems for users
- Many new open source systems appearing that you can install and run on your local cluster
  - should be able to run a variety of applications on these systems

66