# BITS Pilani presentation


BITS Pilani
Hyderabad Campus

D. Powar
Lecturer,
BITS-Pilani, Hyderabad Campus

innovate achieve lead

BITS Pilani
Hyderabad Campus

# SS ZG527

# Cloud Computing

## Introduction to Course and Evaluation

## Lecture -1

## Text Books:

1. Dinkar Sitaram and Geetha Manjunath. Moving to the Cloud. Syngress (Elsevier) Pub, 2011
2. Rajkumar Buyya, James Broburg & anderzej M.G, Cloud Computing – Principles and Paradigms. John Wiley Pub, 2011

## Reference:

1. Dan C. Marinescu, Cloud Computing Theory and practice, Elsevier, 2013
2. Cloud Computing bible by Barrie Sosinsky, Wiley Publishers, 2010
3. Virtualization A Beginner's guide, Danielle Ruest, Nelson Ruest, TMH, 2009

# Evaluation Components

| EC No. | Evaluation Component & Type of Examination | Duration | Weigh-tage | Day, Date, Session,Time |
|---|---|---|---|---|
| **EC-1** | Online Quizzes (Multiple choice) | *Details will be announced on Taxila website* | 15% | August 21-30, 2015<br>September 20-30, 2015<br>October 20-30, 2015 |
| **EC-2** | Mid-Semester Test **(Closed Book)** | 2 Hours | 35% | 20/09/2015 (AN) 2 PM TO 4 PM |
| **EC-3** | Comprehensive Exam **(Open Book)** | 3 Hours | 50% | 08/11/2015 (AN) 2 PM TO 5 PM |

# Course Components

## What are we trying to answer?

What is Cloud Computing?

Why Cloud Computing?

What are its challenges and opportunities?

How does Cloud Computing work?

# Course Review

- Networked and Distributed Computing – Models;
- Cloud Computing, services, deployment models
- Virtualization Techniques and Types
- Introduction to IaaS, case study
- Managing Virtual Resources on the Cloud: Provisioning of Resources; Migration of Virtual Machines
- SaaS and PaaS: Case Studies
- Distributed File systems (GFS, HDFS)
- Storage as a Service (RAID levels)
- Map-Reduce Programming model
- Issues and Challenges : Availability, Multi-Tenancy, and Security
- Cloud Operating systems
  - ✓ Eucalyptus
  - ✓ OpenStack
  - ✓ Amazon
  - ✓ Window Azure
  - ✓ Vmware
  - ✓ HP

# What Computing Paradigms Are There?

Personal Computing

Grid Computing

Reconfigurable Computing

Cluster Computing

Parallel Computing

Utility Computing

Distributed Computing

Pervasive Computing

Ubiquitous Computing

Mobile Computing

Super Computing

Autonomic Computing

Cloud Computing

# Network-centric computing



- Information processing can be done more efficiently on large farms of computing and storage systems accessible via the Internet
  - Grid computing – initiated by the National Labs in the early 1990s; targeted primarily at scientific computing
  - P2P systems computing
  - Utility computing – initiated in 2005-2006 by IT companies and targeted at enterprise computing
- The focus of utility computing is on the business model for providing computing services; it often requires a cloud-like infrastructure
- Cloud computing is a path to utility computing embraced by major IT companies including: Amazon, HP, IBM, Microsoft, Oracle, and others

# Computing is embedded everywhere in the environment



Began during 1985+

# Ubiquitous/ Pervasive computing

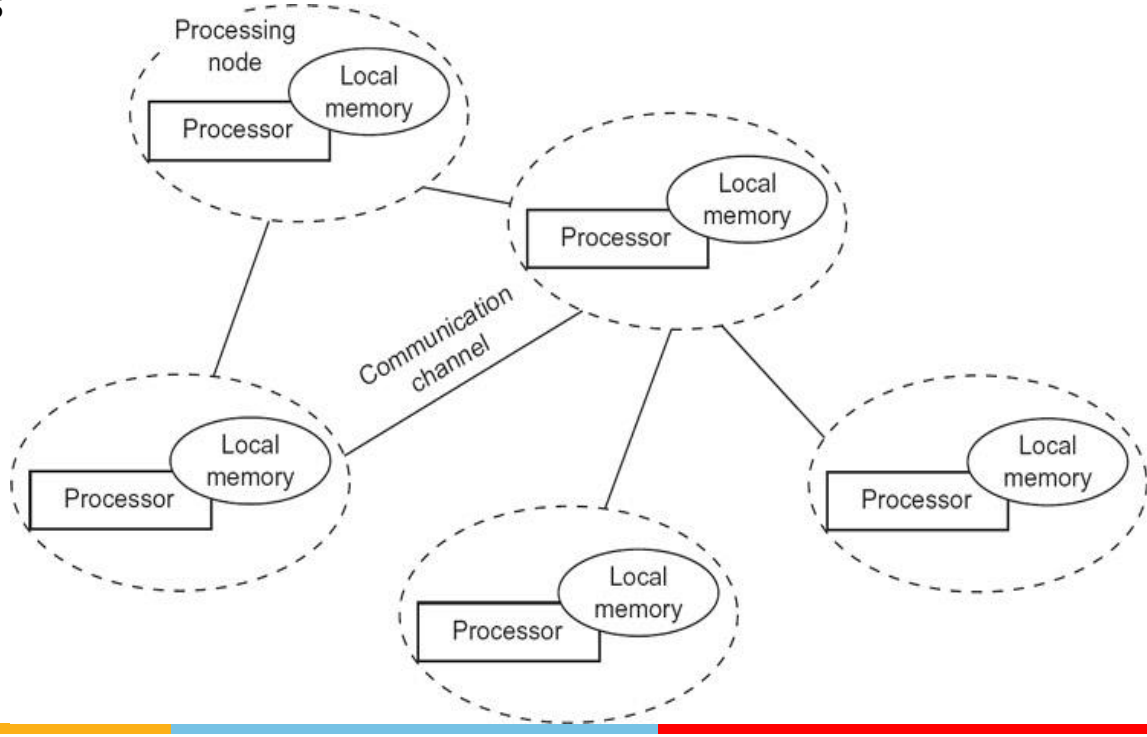
- It is an advanced computing concept where computing is made to appear everywhere and anywhere.
- Ubiquitous computing can occur using any device, in any location, and in any format. A user interacts with the computer, which can exist in many different forms - laptop, tablets, terminals, phones, etc.
- Move beyond desktop machine
- **Ex**: digital audio players, radio-frequency identification tags, PDAs, smartphones, GPS, and interactive whiteboards disturb

# Ubiquitous/ Pervasive computing (contd..)

- Present or noticeable in every part of a thing or place”
- Information processing engaged in every day’s activities and objects

# Distributed Computing

- A distributed computing system is basically a collection of processors interconnected by a communication network in which each processor has its own local memory and other peripherals, and the communication between any two processors of the system takes place by message passing over the communication network
  - ❖ Loosely coupled systems
- Examples
  - Cluster
  - Grid
  - P2P
  - Cloud computing



# Distributed Computing (contd..)

- Uses **middleware**, which enables computers to coordinate their activities and to share the resources of the system.
- This distribution software should support a set of desirable properties of a distributed systems:
- Transparency?
  - Location transparency
  - Replication transparency
  - Migration transparency
  - Scaling transparency
  - Access transparency
  - Concurrency transparency
  - Failure transparency
  - Performance transparency  etc.
- Single integrated computing facility from user perspective

# Distributed Computing (contd..)



- Can include heterogeneous computations where some nodes may perform a lot more computation, some perform very little computation and a few others may perform specialized functionality (like processing visual graphics)
- Using which efficient scalable programs can be designed so that independent processes are scheduled on different nodes and they communicate only occasionally to exchange results

- Cloud computing is also a specialized form of distributed computing, where distributed SaaS applications utilize thin clients (such as browsers) which offload computation to cloud-hosted servers (and services).
- Additionally, cloud-computing vendors providing (IaaS and PaaS) solutions may internally use distributed computing to provide highly scalable cost-effective infrastructure and platform.

# Distributed Computing (contd..)



Cluster Computing:

- A computer cluster is a group of loosely or tightly coupled computers that work together closely so that in many respects it can be viewed as though it were a single computer
- Better performance and availability and more cost- effectiveness over single computer with same capabilities
- Characteristics:
  - Loosely / tightly coupled computers
  - Centralized Job management & scheduling
  - Coined in 1987

Grid Computing:

- Grid is a collection of a large number of loosely coupled, heterogeneous, and geographically dispersed systems in different administrative domains
- Generally owned by multiple organizations that is coordinated to allow them to solve a common problem
- Characteristics
  - Loosely coupled computers
  - Distributed Job management & scheduling
  - Originated (early 1990s)

# Grid Computing (contd..)

Vision: To enable computing to be delivered as a utility

- This vision is most often presented with an analogy to electrical power grids, from which it derives the name “grid”
- The key emphasis of grid computing was to enable sharing of computing resources or forming a pool of shared resources that can then be delivered to users.
- Focus of grid computing was limited to enabling shared use of resources with common protocols for access
- a particular emphasis was given to handle heterogeneous infrastructure, which was typical of a university datacenter.

# Grid Computing (contd..)



- *"coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations."*
  - Ian Foster & Steve Tucker -> "Anatomy of Grid"

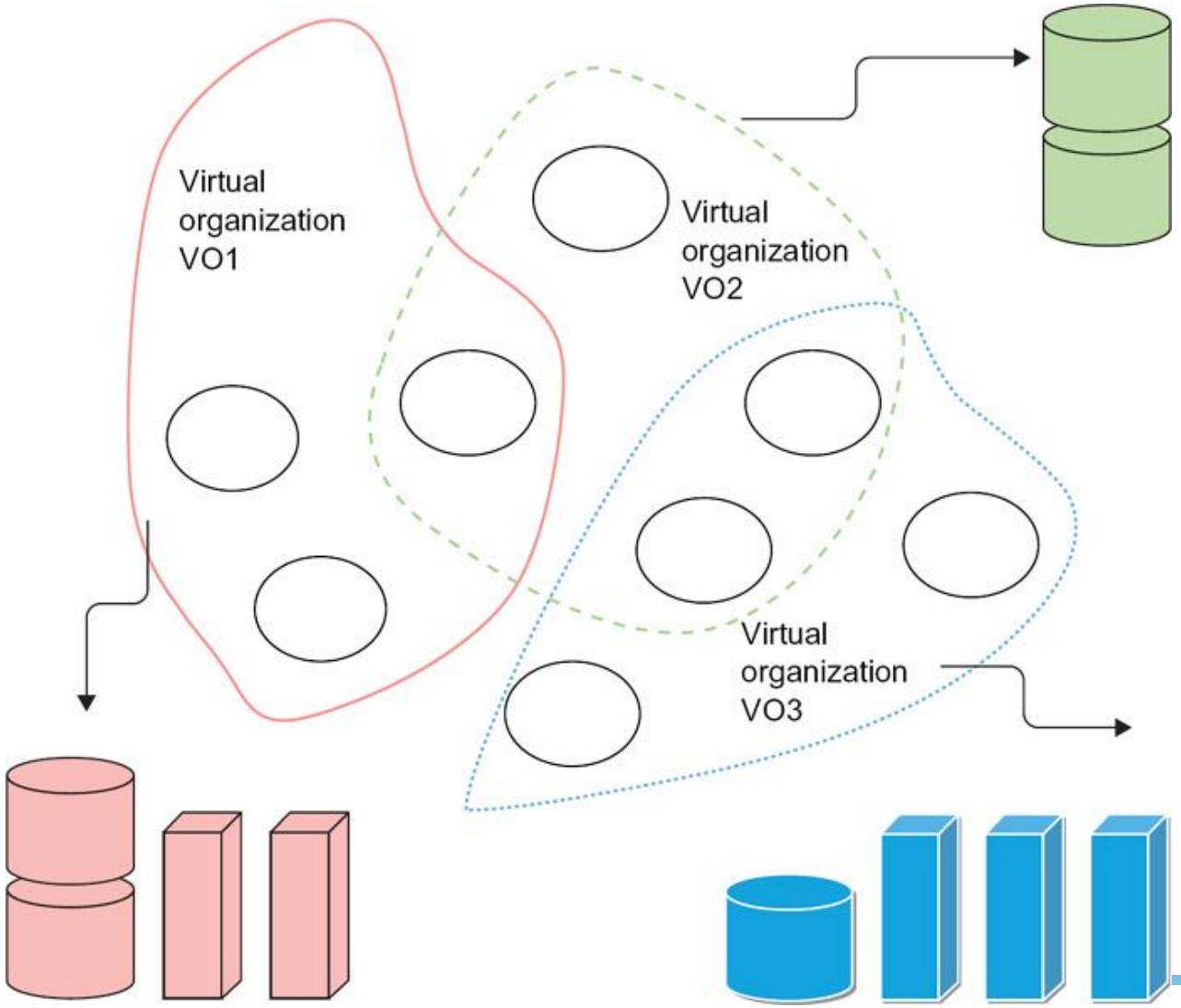
Grid (three-point checklist)?

1. Co-ordinates resources that are not subject to centralized control
2. Using standard, open, general purpose protocols and interfaces
3. To deliver nontrivial quality of service

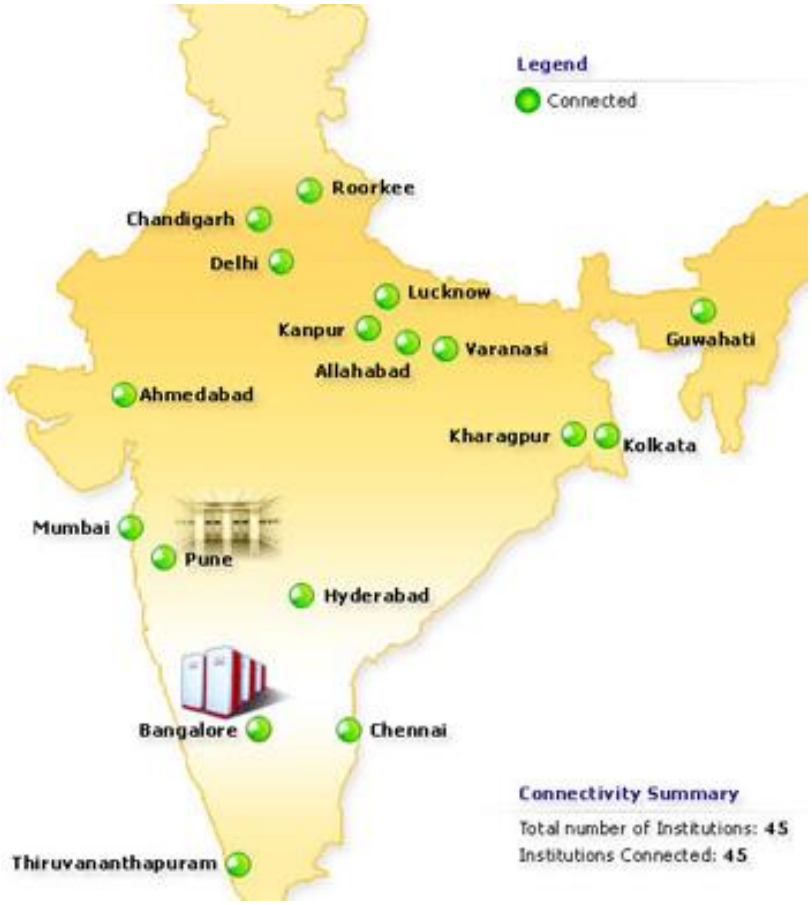
Note:

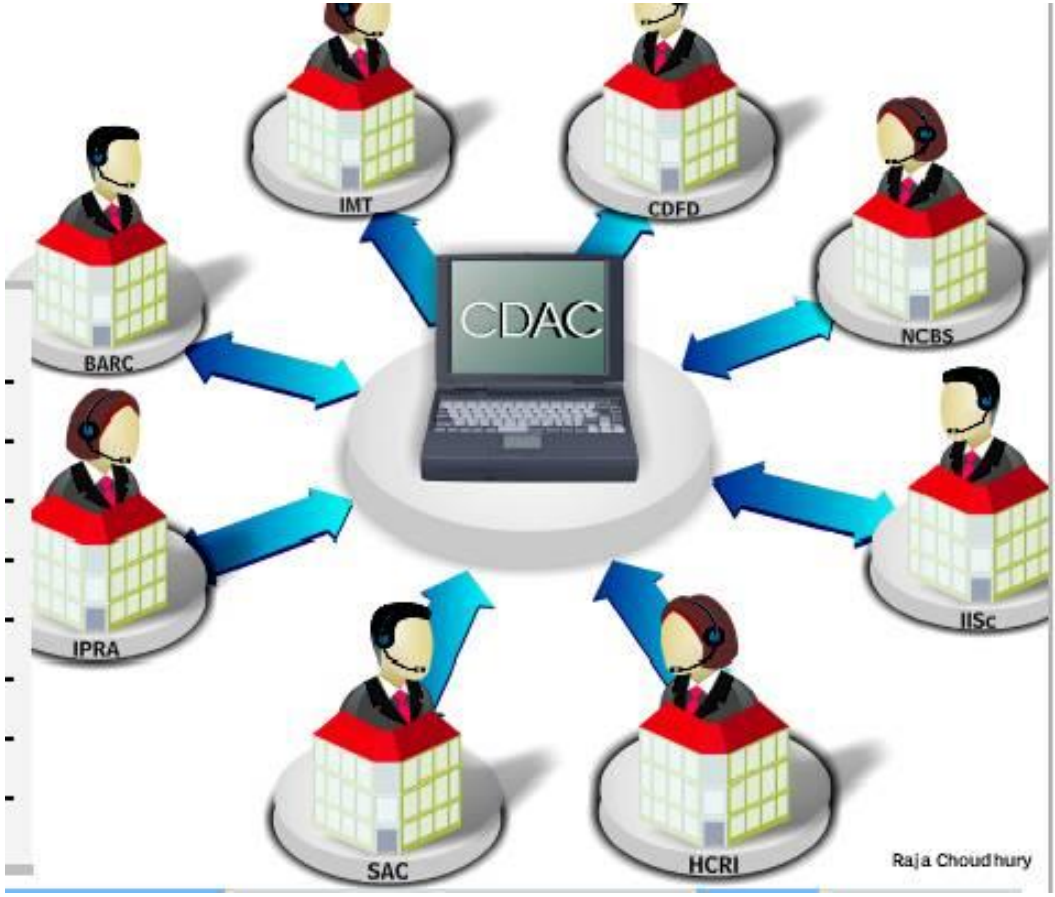
- There are also very specific differences between a grid computing infrastructure and the features one should expect from a cloud computing infrastructure

# Virtual organizations in the grid



# GARUDA grid project (wiki)




**GARUDA** is India's Grid Computing initiative connecting 17 cities across the country. The 45 participating institutes in this nationwide project include all the IITs and C-DAC centers and other major institutes in India

# Advantages - Distributed computing

**Inherently Distributed applications:**

Several applications are inherently distributed in nature and require distributed computing system for their realization

**Information Sharing among Distributed Users:**

In a distributed computing system, information generated by one of the users can be easily and efficiently shared by the users working at other nodes of the system .

**Resource Sharing:**

Sharing of software resources such as software libraries and databases as well as hardware resources such as printers, hard disks, etc can also be done in a very effective way among all the computers and the users of a single distributed computing system.

## Extensibility and Incremental Growth:

- It is possible to gradually extend the power and functionality of a distributed computing system by simply adding additional resources (both hardware and software) to the system as and when the need arises .
- Incremental growth is very attractive feature because for most existing and proposed applications it is practically impossible to predict future demands of the system.
- Addition of new resources to an existing system can be performed without significant disruption of the normal functioning of the system.

## Shorter Response Times and Higher Throughput:

- The multiple processors of the distributed computing system can be utilized properly for providing shorter response times and higher throughput than a single processor centralized system.
- Another method often used in distributed computing systems for achieving better overall performance is to distribute the load more evenly among the multiple processors by moving the jobs from currently overloaded processors to lightly loaded ones

## Higher Reliability:

- Reliability refers to the degree of tolerance against errors and component failures in a system.
- A reliable system prevents loss of information even in the event of component failures
- An important aspect of reliability is availability, which refers to the fraction of time for which a system is available for use.

## Better Flexibility in Meeting User's Needs:

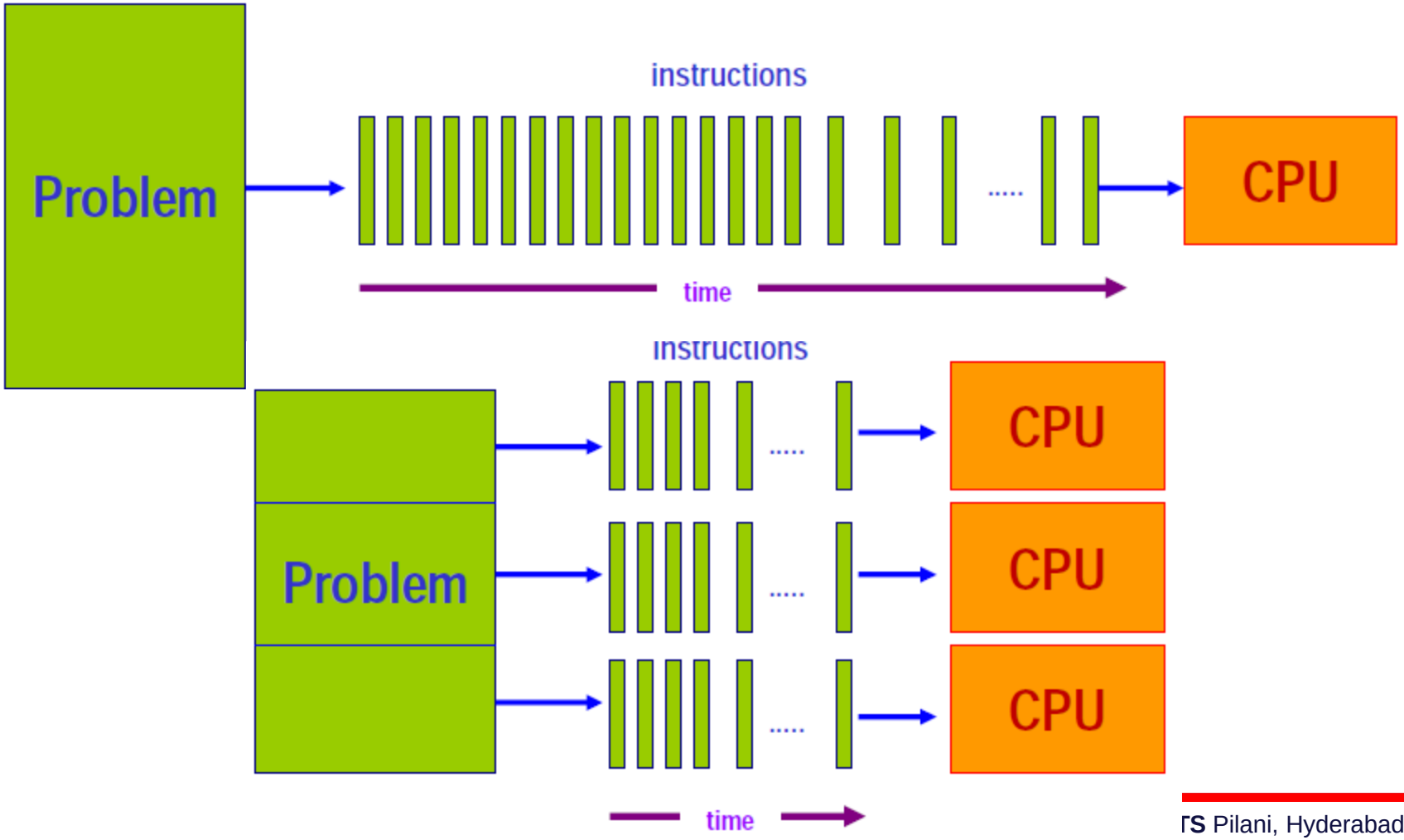
- A distributed computing system may have a pool of different types of computers, in which case the most appropriate one can be selected for processing a user's job depending on the nature of the job.

## Better Price-Performance Ratio:

- With the rapidly increasing power and reduction in prize of the microprocessors, combined with the increasing speed of communication network, distributed computing systems potentially have a much better price performance ratio than a single large centralized system.

# Parallel Computing

Calculations of large problems are divided into smaller parts and carried out simultaneously/concurrently on different processors



# Super Computing

- Thousands of processors
- Used for compute-intensive problems
  - Days instead of Years!!!
  - Introduced in the 1960s


Nasa super computer

Blue Gene

- Tianhe-2, a supercomputer developed by China (33.86 petaflops)
- Titan by U.S (17.59 petaflops)
- PARAM Yuva II by India(524 teraflops)
  - PARAM 8000 (India's first supercomputer by CDAC in 1990)

# Utility Computing

- John McCarthy - *'computers of the future'*
- **Utility computing** is the packaging of computing resources, such as computation, storage and applications, as a metered service similar to traditional public utility (such as electricity, water, natural gas, or the telephone network).
- This model has the advantage of a low or no initial cost to acquire computer resources; instead, computational resources are essentially rented
  - ✓ You get connected to the utility companies' "public" infrastructure
  - ✓ You get these utility services on-demand
  - ✓ And you pay-as-you use (metered service)

** Cloud computing is the most recent technology innovation which has made utility computing a reality!

# Autonomic Computing

by Paul Horn of IBM in 2001

- Vision
  - making all computing systems manage themselves automatically
  - make computing systems self-configuring, self-optimizing, and self-protecting – as well as self-healing
- It refers to self-managing characteristics of distributed computing resources, which recognize and understand changes in the system, take appropriate corrective actions completely automatically, with close to zero human intervention.

- Cloud computing shares the vision of autonomic computing and more

# Summary:

- Evaluation component details
- Course review
- Computing Paradigms
  - Cluster Computing
  - Grid Computing
  - Parallel Computing
  - Super Computing
  - Utility Computing
  - Automatic Computing, etc.