# CENG510
# Distributed Systems

Instructor : Elif Haytaoğlu

Course : F 13:30-16:30

Office Hour : W 15:00-16:30

http : //pau.edu.tr/eacar/CENG510

Set of computers connected by a communication network.

Gives the user an illusion of a single computer

Old platform : Usually a number of WSs over a LAN

Now, ranges from a LAN to a sensor network to a mobile network

Each node in a DS :

- is autonomous
- communicates by messages
- needs to synchronize with others

  To achieve a common goal (load balancing, fault tolerance, an application..)

- There are a number of benefits to be gained by utilizing distributed systems:
- One of the obvious advantages of using a distributed system is resource sharing . Access to a central resource has two disadvantages as this central site becomes a bottleneck for communications and also is a single point of failure.
- Distributing the resources such as the database and peripherals over a network overcomes these problems.

- Resources and computation can be replicated at various sites providing fault tolerance as a replica may be substituted in the case of the disfunctioning of a node.
- This type of fault tolerance is an important reason to employ distributed systems.

- It is also possible for the application to be inherently distributed such as bank transaction systems and airline reservation systems where employment of distributed systems is inevitable.

- A distributed system is a collection of entities, each of which is autonomous, programmable, asynchronous and failure-prone, and communicating through an unreliable communication medium.

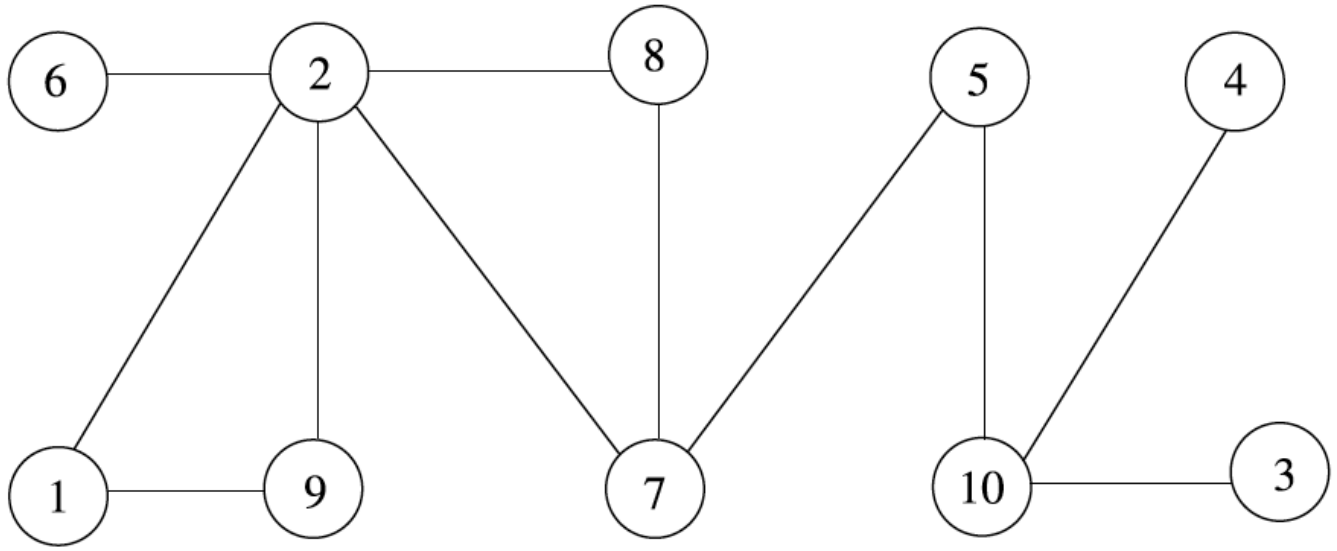
- Our interest in distributed systems in this course is about

  - Study of algorithms rather than systems
  - We will model the distributed system as a graph G(V,E) where

    - V is the set of nodes (processors, process on a device)
    - E is the set of edges (communication links, wired or wireless)

# A distributed system can be modeled as a graph G(V,E).

**Fig. 1.1** A graph representing a distributed system



# Distributed Computing Platforms

- Due to the recent technological advancements, in the last few decades, we have witnessed diverse distributed system platforms such as **The Grid**, **The Cloud**, **mobile ad hoc networks**, and **wireless sensor networks** that are described below.

# The Grid

- The Grid consists of loosely coupled, heterogeneous, and geographically dispersed computing elements that are connected by a network acting together to perform large tasks [3 ].

- These computationally intensive scientific tasks may include various applications such as **seismic analysis**, **drug discovery**, and **bioinformatics problems**.
- Grid computing provides effective usage of the unused processing power and results in decreased completion time for a task due to parallelization.

# The Grid

- Middleware
- Resource Discovery
- Fault Tolerance
- Security

# The Grid

- The term *grid computing* originated in the early 1990s as a metaphor for making computer power as easy to access as an electric power grid.
- The power grid metaphor for accessible computing quickly became canonical when Ian Foster and Carl Kesselman published their seminal work, "The Grid: Blueprint for a new computing infrastructure" (1999)
- CPU scavenging and volunteer computing were popularized beginning in 1997 by distributed.net and later in 1999 by SETI@home to harness the power of networked PCs worldwide, in order to solve CPU-intensive research problems.
- BOINC brings together about 311,742 active participants and 834,343 active computers (hosts) worldwide processing on average 16.912 PetaFLOPS as of 13 January 2017.[

# The Grid

- The European Grid Infrastructure (EGI)
  - High-energy physics,
  - Earth observation,
  - Biology applications

- The United States, the National Grid (USNG)
  - Access grid for people.

# The Cloud

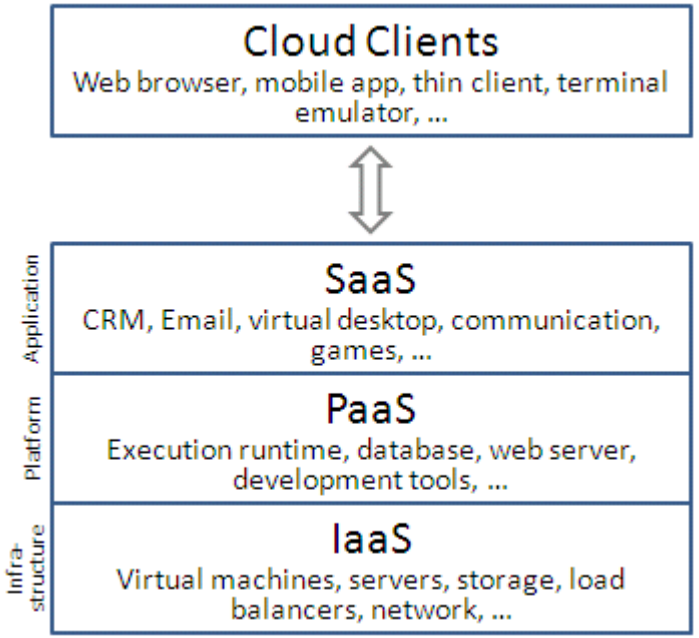
- The cloud computing evolved from grid computing

- Cloud computing provides computation, software applications, data access, data management, and storage for resources without requiring cloud users to know the location and other details of the computing infrastructure. Grid computing may be included in the cloud or not depending on the type of application.

- Grid computing may be included in the cloud or not depending on the type of application and users.

# Cloud Flavors?

- SaaS – Software as a Service
- PaaS – Platform as a Service
- IaaS – Infrastructure as a Service



# SaaS Examples

salesforce.com
Success On Demand.

NETSUITE
ONE SYSTEM. NO LIMITS.


postini


facebook

# Platform as a Service (PaaS)

- PaaS provides all of the facilities required to support the complete life cycle of building and delivering web applications and services entirely from the Internet.
  - Typically applications must be developed with a particular platform in mind
  - Highly scalable multi tier architecture

# PaaS Examples






# Infrastructure as a Service (IaaS)

- IaaS is the delivery of technology infrastructure as an on demand scalable service
  - Usually billed based on usage
  - Can be coupled with Managed Services for OS and application support

# IaaS Examples






# The Cloud

- In 2013, it was reported that cloud computing had become a highly demanded service or utility due to the advantages of high computing power, cheap cost of services, high performance, scalability, accessibility as well as availability. Some cloud vendors are experiencing growth rates of 50% per year, but being still in a stage of infancy, it has pitfalls that need to be addressed to make cloud computing services more reliable and user friendly.

# The Cloud

- The origin of the term *cloud computing* is unclear. The word "cloud" is commonly used in science to describe a large agglomeration of objects that visually appear from a distance as a cloud and describes any set of things whose details are not further inspected in a given context.

- Another explanation is that the old programs that drew network schematics surrounded the icons for servers with a circle, and a cluster of servers in a network diagram had several overlapping circles, which resembled a cloud.

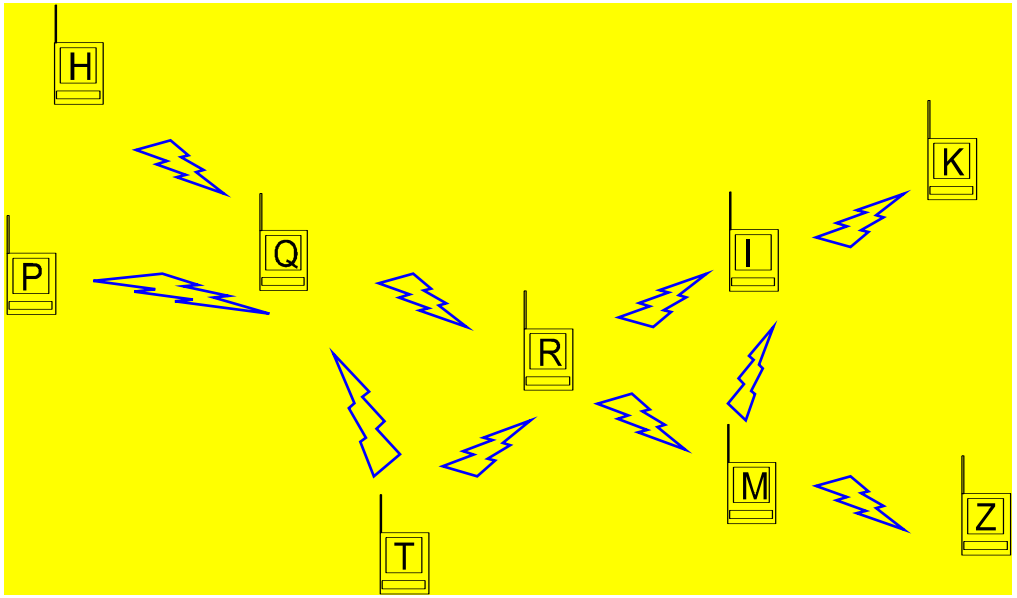
# Mobile Ad Hoc Networks (MANET)

- A wireless ad hoc network is a decentralized network consisting of wireless nodes that do not rely on a predefined infrastructure such as routers or access points.

- Instead, each node participates in routing by forwarding data to other nodes regarding dynamically changing network topology

# MANET

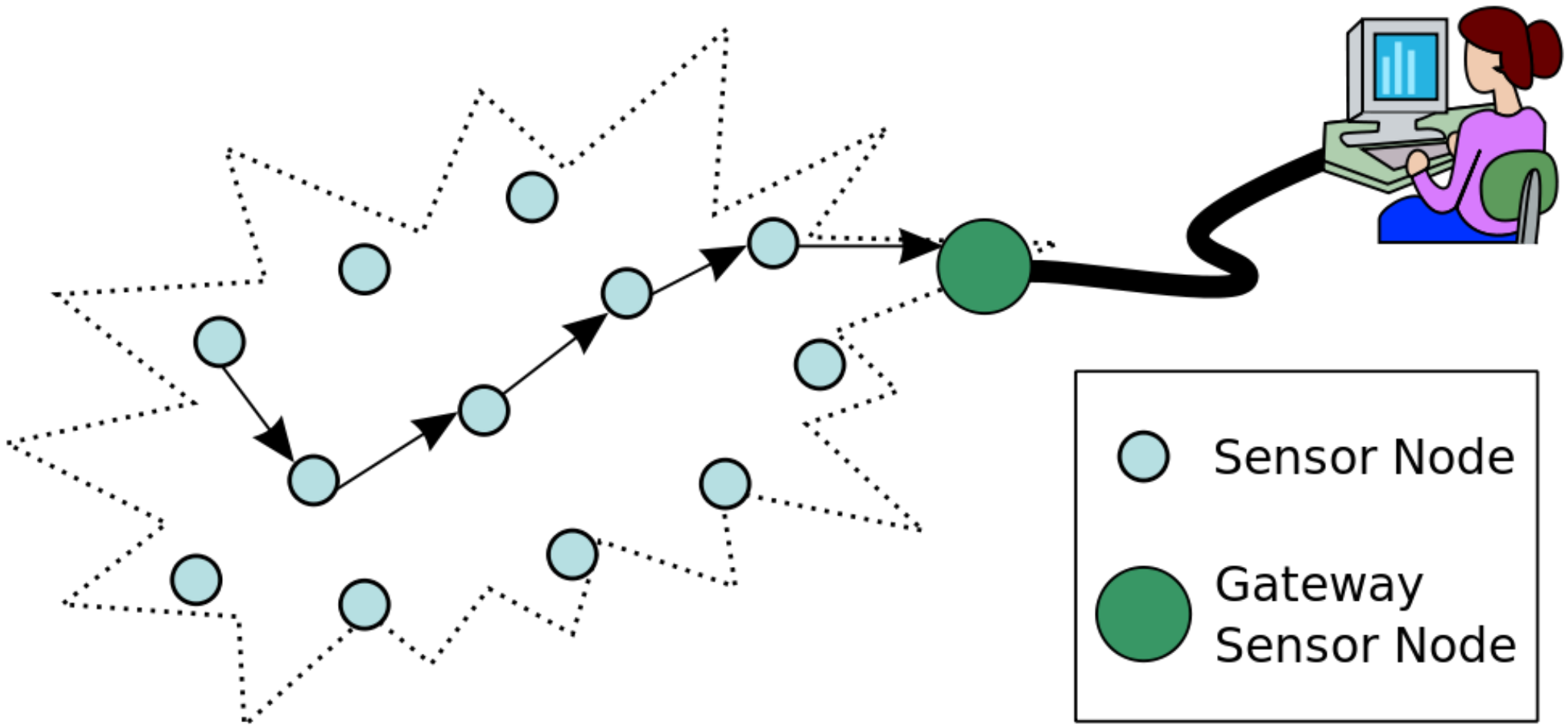
- MANET = Mobile Ad Hoc Networks
  - multi-hop communication
  - needs support of dynamic routing protocols




# Wireless Sensor Networks

- A wireless sensor network (WSN ) consists of many small nodes of computing elements, each equipped with sensing and wireless communication capabilities.
  - habitat monitoring,
  - military surveillance, and target
  - tracking
- WSNs form a large-scale distributed system and require scalable distributed algorithms to solve problems such as data aggregation, topology control and routing.

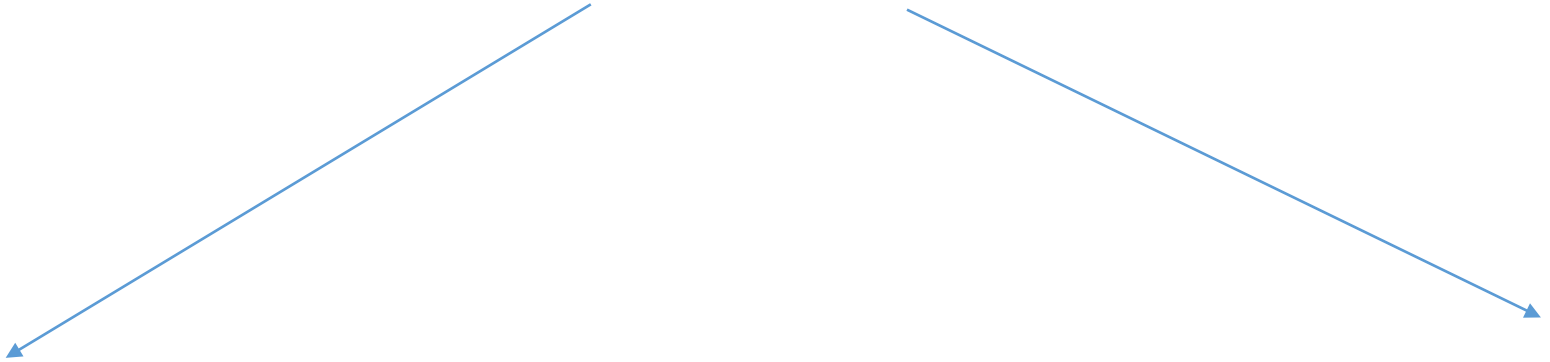
# Wireless Sensor Networks



# Distributed System Communication Models

- Message Passing
  - Synchronous
  - Asyncronous
- Shared Memory
  - Synchronous

# Distributed Systems Design Issues



- System Software
  - Communication
  - Synchronization
  - Security

Distributed Algorithms

- Load Balancing
- Leader Election
- Mutual Exclusion
- Deadlock
- Fault Tolerance Algorithms

# The Aim of the Course

- The aim of this course is the design of such distributed approximation graph algorithms that may be of use in distributed applications.
- As a concrete example, finding a minimum connected dominating set that is the subset V’ of vertices of a graph G with minimum size such that every vertex of the graph is either in V’ ora neighbor of V’ and all of the vertices in V are connected is NP-hard for general graphs.
- Therefore finding an approximation algorithm that has a better approximation than the best known algorithm is clearly a contribution on its own.

# Course Overview

- Graphs
- Models
- Spanning Tree Construction
- Graph Traversal
- Routing
- Vertex Coloring
- Dominating Sets
- Time Clocks
- Distributed Mutual Exclusion