Introduction

A distributed system is:

- One in which components located at networked computers communicate and coordinate their actions only by passing messages. –G. Coulouris
- A collection of independent computers that appears to its users as a **single coherent system**. –S. Tanenbaum
- A distributed system organized as **middleware**.

Characteristics of Distributed Systems:

- Concurrency of components
    - In a network of computers, concurrent program execution is the norm, sharing resources
- No global clock
    - When programs need to cooperate, they coordinate their actions by exchanging messages
    - Close coordination depends on shared time
    - There are limits to the accuracy with which the computers in a network can synchronize their clocks – there is no single global notion of the correct time [Section 6.1.1 of Tanenbaum]
    - This is a consequence of the fact that the only communication is by sending messages through the network
- Independent failures of components
    - Each component of a distributed system can fail independently, leaving the others still running

Service, Server, Client

- Request, respond
- A complete interaction between a client and a server, is called a remote invocation

Challenges in building distributed systems

- Heterogeneity of components
    - 多样性
    - Varying software and hardware
    - Virtual machines
- Openness

    - Extensibility
    - Independence of vendors
    - Publishable key interfaces
        - CORBA
    - Publishable communication mechanisms
        - Java RMI
- Security
    - Confidentiality
    - Integrity
- Scalability
    - A system is scalable if it will remain effective when there is a significant increase in the number of resources and users
    - Controlling the cost of physical resources
    - Controlling the performance loss
    - Preventing software resources running out
    - Avoiding performance bottleneck
- Failure handling
    - Detecting failures
    - Masking(Hiding) failures
    - Tolerate failures
    - Recovery from failures
    - Redundancy
- Concurrency
    - Synchronization
    - IPC: Inter-process communication
- Transparency
    - Concealing from the user and the application programmer of the nature of separation of components in a distributed system
    - (Persistence transparency)
    - Access transparency
    - Location transparency
    - Concurrency transparency
    - (Migration transparency)
    - Replication transparency
    - Mobility(Relocation) transparency
    - Failure transparency
    - Performance transparency
    - Scaling transparency

The WWW

The World Wide Web (WWW), commonly known as the Web, is an information space where documents and other web resources are identified by Uniform Resource Locators (URLs, such as https://www.example.com/), which may be interlinked by hypertext, and are accessible over the Internet.

- HTML: Hypertext Markup Language
- URL: Uniform Resource Locator
- Client-Server architecture, with HTTP