IT4371: Distributed Systems

Spring 2016

Introduction to Distributed Systems

Dr. Nguyen Binh Minh

Department of Information Systems
School of Information and Communication Technology
Hanoi University of Science and Technology

Today…

- Last Session:
  - Course
  - Why distributed systems?
  - Main requirements for building distributed systems
- Today's session:
  - Part I: Some trends in distributed systems
  - Part II: General challenges in building distributed systems

Today…

Part I

Some trends in distributed systems

Some Trends in Distributed Systems

- Distributed systems are undergoing a period of significant change and this can be traced back to a number of influential trends:
  - ✓ The emergence of pervasive networking technology.
  - ✓ The emergence of ubiquitous computing coupled with the desire to support user mobility in distributed systems.
  - ✓ The increasing demand for multimedia services.
  - ✓ The view of distributed systems as utility.
  - ✓ Others…

Trends in Distributed Systems

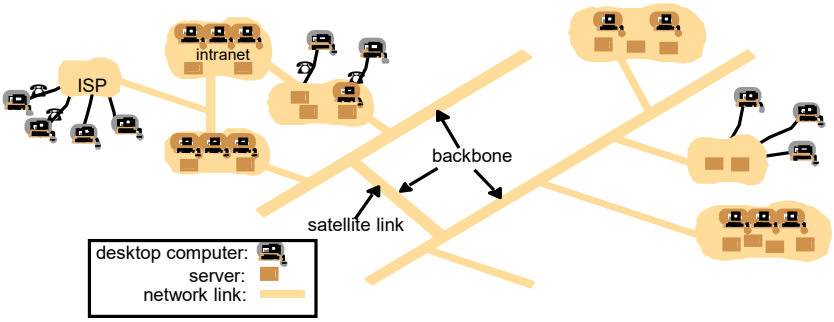
- Distributed systems are undergoing a period of significant change and this can be traced back to a number of influential trends:
  - ✓ The emergence of pervasive networking technology.
  - ✓ The emergence of ubiquitous computing coupled with the desire to support user mobility in distributed systems.
  - ✓ The increasing demand for multimedia services.
  - ✓ The view of distributed systems as utility.

Pervasive Networking

- Technological advances in computer networking have led to different types of networks including but not limited to:
  - WiFi
  - WiMAX
  - Bluetooth
  - Overlay networks (e.g., BitTorrent and peer-to-peer networks)
- Networking has become a *pervasive* resource and devices can mostly be connected at any time and in any place.

The Modern Internet

- The modern Internet is a vast interconnected collection of computer networks of many different types.
- The modern Internet allows users wherever they are, to make use of services such as World Wide Web, email and file transfer.



Pervasive Networking and the Modern Internet

- Some important questions:
  - How do programs running on networked heterogeneous computers interact?
  - Can the set of services provided by the Internet be extended?
  - Can programs interact reliably and securely?

Trends in Distributed Systems

- Distributed systems are undergoing a period of significant change and this can be traced back to a number of influential trends:
  - ✓ The emergence of pervasive networking technology.
  - ✓ The emergence of ubiquitous computing coupled with the desire to support user mobility in distributed systems.
  - ✓ The increasing demand for multimedia services.
  - ✓ The view of distributed systems as utility.

Mobile and Ubiquitous Computing

- Advances in device miniaturization and wireless networking have led to the integration of small and portable computing devices into distributed systems. These devices include:
  - Laptop computers
  - Handheld devices (e.g., GPS-enabled devices and PDAs)
  - Wearable devices (e.g., smart watches)
  - Devices embedded in appliances (e.g., washing machines and cars)
- Mobile and ubiquitous computing enablers:
  - Device portability
  - Ease of connection

Mobile and Ubiquitous Computing

Mobile computing is the performance of computing tasks while a user is on the move, or visiting places other than her/his usual environment.

Ubiquitous computing is the availability of computational power on demand in your vicinity to such a degree that information processing becomes integrated into everyday objects and activities.

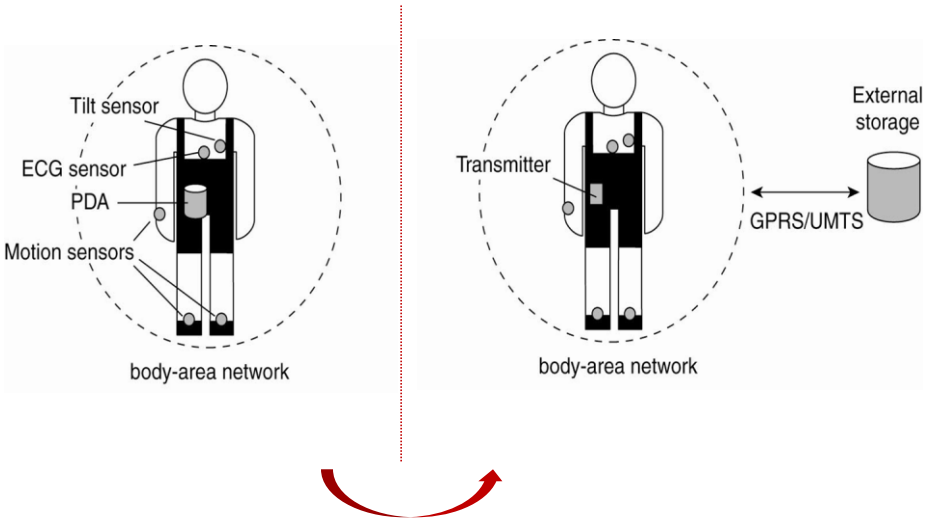
Mobile and Ubiquitous Computing

- Ubiquitous computing is only realized when mobile computing matures.
- Mobile and ubiquitous computing introduces (generally) a couple of challenges for distributed systems:
  - Dealing with variable connectivity and disconnections.
  - Maintaining operation in the face of device mobility.

Example

- Health Care Systems:
  - New devices are being developed to monitor the well-being of individuals and to automatically contact physicians when needed.
  - Personal health care systems are often equipped with various sensors organized in a (preferably wireless) body-area network (BAN).
  - BAN should be able to operate while a person is moving, with no strings (i.e., wires) attached to immobile devices.

Health Care Systems (HCS)



Issues for HCS

- Where and how should monitored data be stored?
- How can we prevent loss of crucial data?
- How can physicians provide online feedback?
- What are the security issues and how can the proper policies be enforced?

Trends in Distributed Systems

- Distributed systems are undergoing a period of significant change and this can be traced back to a number of influential trends:
  - ✓ The emergence of pervasive networking technology.
  - ✓ The emergence of ubiquitous computing coupled with the desire to support user mobility in distributed systems.
  - ✓ The increasing demand for multimedia services.
  - ✓ The view of distributed systems as utility.

Distributed Multimedia Systems

- Supporting multimedia services can be defined as the ability to support a range of discrete and continuous media.

- Benefits of distributed multimedia computing:
  - Access to live or pre-recorded TV broadcasts
  - Access to music libraries
  - Webcasting
  - Access to film libraries offering video-on-demand services
  - The provision of audio and video conferencing facilities and integrated telephony features

Demands of a Distributed Multimedia Systems

- Distributed multimedia applications place considerable demands on the underlying distributed infrastructure in terms of:
  - Providing support for an (extensible) range of encoding and encryption formats (e.g., MPEG series of standards).
  - Providing a range of mechanisms to ensure that the desired *quality of service* (QoS) can be met.

Trends in Distributed Systems

- Distributed systems are undergoing a period of significant change and this can be traced back to a number of influential trends:
  - ✓ The emergence of pervasive networking technology.
  - ✓ The emergence of ubiquitous computing coupled with the desire to support user mobility in distributed systems.
  - ✓ The increasing demand for multimedia services.
  - ✓ The view of distributed systems as utility.

Distributed Computing As Utility

- Distributed resources can be viewed as a commodity or utility similar to water or electricity.
- Resources are provided by appropriate service suppliers and rented rather than owned by an end user.
- The term *cloud computing* is used to capture this vision.

Enablers and Advantages

- Clouds are generally implemented on commodity computers to provide necessary scale and performance.
- *Virtualization* is deemed a key enabler for resource sharing, customization and elasticity.

- Advantages:
  - Promotes a view of everything as a service.
  - Allows very simple desktop or portable devices to access a potentially wide range of resources and services.
  - Scalability and elasticity

Open Challenges in Cloud Computing

- Cloud Security
- Cloud Applications
- Cloud Programming Models
- QoS in Clouds
- Energy-efficient Clouds
- Cloud Storage Systems

Today…

Part II

Challenges When Designing Distributed Systems

Pitfalls when Developing Distributed Systems

False assumptions made by first time developer:

- The network is reliable.
- The network is secure.
- The network is homogeneous.
- The topology does not change.
- Latency is zero.
- Bandwidth is infinite.
- Transport cost is zero.
- There is one administrator.

Challenges When Designing Distributed Systems

- Many issues arise when designing distributed systems:

1. Heterogeneity
2. Openness
3. Security
4. Scalability
5. Failure Handling
6. Concurrency
7. Transparency
8. Quality of Service

Challenges When Designing Distributed Systems

- Many issues arise when designing distributed systems:

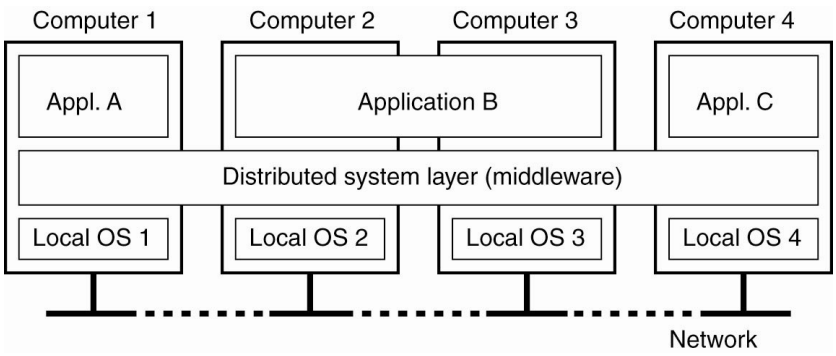
1. Heterogeneity
2. Openness
3. Security
4. Scalability
5. Failure Handling
6. Concurrency
7. Transparency
8. Quality of Service

Challenge 1: Heterogeneity

- Heterogeneity applies to:
  a. Networks
  b. Computer Hardware
  c. Operating Systems.
  d. Programming Languages.

A Solution: Middleware

- *Middleware* is a software layer that provides a programming abstraction and masks the heterogeneity of the underlying networks, hardware, operating systems and programming languages.



Challenges When Designing Distributed Systems

- Many issues arise when designing distributed systems:

1. Heterogeneity
2. Openness
3. Security
4. Scalability
5. Failure Handling
6. Concurrency
7. Transparency
8. Quality of Service

Challenge 2: Openness

- The openness of a distributed system is the degree to which the system can be extended and re-implemented in various ways.

- Open distributed systems are:
  - Characterized by the fact that their key interfaces are published.
  - Based on the provision of a uniform communication mechanism for access to shared resources.
  - (or can be) Constructed from heterogeneous hardware and software components with tested conformance to published standards.

Challenges When Designing Distributed Systems

- Many issues arise when designing distributed systems:

1. Heterogeneity
2. Openness
3. Security
4. Scalability
5. Failure Handling
6. Concurrency
7. Transparency
8. Quality of Service

Challenge 3: Security

- Many of the information resources maintained in distributed systems have a high intrinsic value to their users.

- Requirements are:
  - To send sensitive information in a message over a network in a secure manner
  - To recognize the identity of the user on whose behalf a message was sent.
  - To ensure that a process gets only those access rights it is entitled to.

Challenges When Designing Distributed Systems

- Many issues arise when designing distributed systems:

1. Heterogeneity
2. Openness
3. Security
4. Scalability
5. Failure Handling
6. Concurrency
7. Transparency
8. Quality of Service

Challenge 4: Scalability

- A distributed system is said to be scalable if it will remain effective when the number of resources and users is significantly increased.

Date	Computers	Web servers
1993, July	1,776,000	130
1995, July	6,642,000	23,500
1997, July	19,540,000	1,203,096
1999, July	56,218,000	6,598,697
2001, July	125,888,197	31,299,592
2003, July	~200,000,000	42,298,371
2005, July	353,284,187	67,571,581

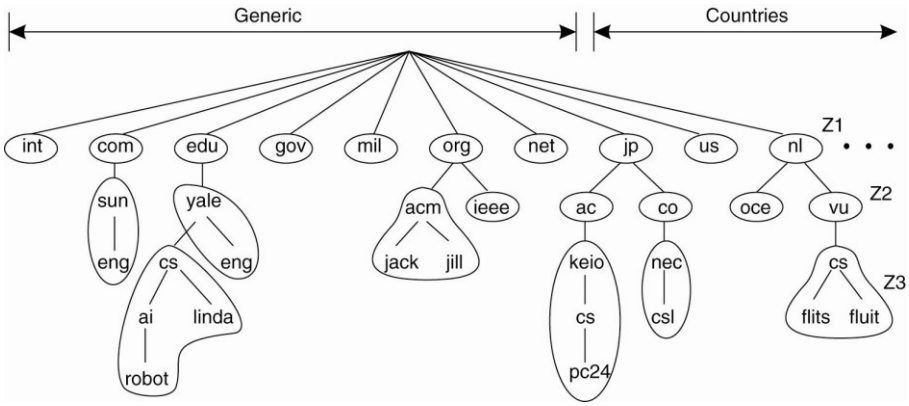
Challenges for Scalability

- The design of scalable distributed systems presents the following challenges:
  - Controlling the cost of physical resources: In general, the quantity of physical resources required should be at most O(n).
  - Controlling the performance loss: hierarchal structures scale better than linear ones.
  - Avoiding performance bottlenecks: decentralized algorithms are generally better than centralized.

Some Solutions to Scalability

- Techniques proven to be successful:
  - Replication and caching with consistency considerations.
  - Deployment of multiple servers to handle commonly performed tasks enabling concurrency.
  - Decentralized lookup tables (e.g. DNS name table).

Domain Name System



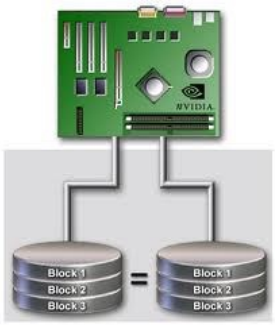
Challenges When Designing Distributed Systems

- Many issues arise when designing distributed systems:

1. Heterogeneity
2. Openness
3. Security
4. Scalability
5. Failure Handling
6. Concurrency
7. Transparency
8. Quality of Service

Challenge 5: Failure Handling

- Whenever multiple machines are used in cooperation with one another, the probability of failures rises.
- Failures in a distributed system are *partial.*
- Techniques for dealing with failures:
  - Detecting Failures
  - Masking Failures
  - Tolerating Failures
  - Recovery From Failures



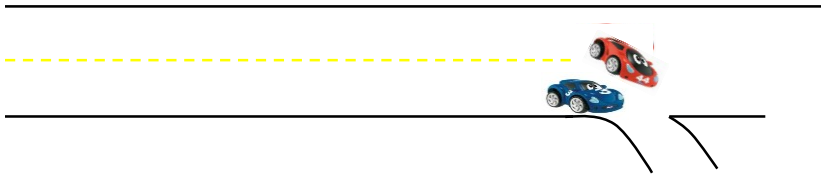
Challenges When Designing Distributed Systems

- Many issues arise when designing distributed systems:

1. Heterogeneity
2. Openness
3. Security
4. Scalability
5. Failure Handling
6. Concurrency
7. Transparency
8. Quality of Service

Challenge 6: Concurrency

- In distributed systems, several clients might attempt to access a shared resource concurrently.
- If objects do not operate correctly and *synchronize* with one another, inconsistencies might arise.



Challenges When Designing Distributed Systems

- Many issues arise when designing distributed systems:

1. Heterogeneity
2. Openness
3. Security
4. Scalability
5. Failure Handling
6. Concurrency
7. Transparency
8. Quality of Service

Challenge 7: Transparency

Transparency	Description
Access	Hide differences in data representation and how a resource is accessed
Location	Hide where a resource is located
Migration	Hide that a resource may move to another location
Relocation	Hide that a resource may be moved to another location while in use
Replication	Hide that a resource is replicated
Concurrency	Hide that a resource may be shared by several competitive users
Failure	Hide the failure and recovery of a resource

Challenges When Designing Distributed Systems

- Many issues arise when designing distributed systems:

1. Heterogeneity
2. Openness
3. Security
4. Scalability
5. Failure Handling
6. Concurrency
7. Transparency
8. Quality of Service

Challenge 8: Quality of Service

- Once users are provided with the functionality that they require of a service, we can go on to ask about the quality of the service (QoS) provided.
- The main nonfunctional properties of systems that affect QoS are:

  - Reliability
  - Security
  - Performance
  - Adaptability to meet changing system configurations
  - Availability of the necessary computing and network resources at the appropriate times

Next Lecture

Distributed Systems Architectures