Distributed Information Processing

1st Lecture

Eom, Hyeonsang (엄현상)
Department of Computer Science & Engineering
Seoul National University

©Copyrights 2014 Eom, Hyeonsang All Rights Reserved
Outline

- Introduction to Distributed Systems
- Clock
- Q&A
Distributed Systems

- Systems Where Components of Networked Computers Communicate and Coordinate Their Actions via Message Passing
  - Concurrency
  - No Global Clock
  - Independent Failures

- Systems Consisting of Collections of Spatially Separated Processes Communicating by Exchanging Messages
  - Sharing States
  - Providing Services
  - Having Global Properties
  - Heterogeneity
  - Unreliable, Insecure, Costly Comm.
  - Scalability
  - Autonomy
Different-Scale Systems

- Systems of Increasing Scale & Decreasing Integration

<table>
<thead>
<tr>
<th>System Type</th>
<th>Heterogeneity, Geographic Distribution</th>
<th>Lack of Centralized Control</th>
<th>Exemplar Computational Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>End (Single) System</td>
<td></td>
<td></td>
<td>Multithreading</td>
</tr>
<tr>
<td>Cluster</td>
<td></td>
<td></td>
<td>Distributed Shared Memory</td>
</tr>
<tr>
<td>Intranet</td>
<td>○</td>
<td></td>
<td>Manager/Worker</td>
</tr>
<tr>
<td>Internet</td>
<td>○</td>
<td>○</td>
<td>Collaborative Systems</td>
</tr>
</tbody>
</table>
Example of Distributed Systems

A typical portion of the Internet

desktop computer: 
satellite link:
server: 
network link:
Distributed System Goals

- Accessibility
  - As Connectivity and Sharing Increases, Security and Privacy Matter

- Distribution Transparency
  - Access: Regarding Data Representation & Access Method
  - Location: Regarding Resource Location
  - Migration: Regarding Resource Movement
  - Relocation: Regarding Movement in Use
  - Replication: Regarding Resource Replication
Distributed System Goals (Cont’d)

- Distribution Transparency (Cont’d)
  - Concurrency: Regarding Competitive Sharing
  - Failure: Regarding Resource Failure and Recovery

- Performance & Comprehensibility Issues

- Openness:
  - Following Standard Rules That Describe the Syntax and Semantics of Services
  - Interoperability
Distributed System Goals (Cont’d)

- **Openness (Cont’d)**
  - Portability
  - Extensibility
  - Separating Policy from Mechanism

- **Scalability**
  - Limitations
    - Centralized services
    - Centralized data
    - Centralized algorithms
    - Synchronous communication
Distributed System Goals (Cont’d)

- Scalability (Cont’d)
  - Decentralized Characteristics
    - No local maintenance of global system state
    - Decision making based only on local information
    - Localized Failure
    - No global clock
  - Scaling Techniques
    - Asynchronous communication
    - Client and Server load balancing
    - Distribution: e.g., DNS & WWW
    - Replication & caching
Distributed System Goals (Cont’d)

- Pitfalls
  - Network Reliability
  - Network Security
  - Network Homogeneity
  - Static Topology
  - Zero Latency
  - Infinite Bandwidth
  - Zero Transport Cost
  - One Administrator
Technical Goals

- Heterogeneity
  - H/W, S/W, and Data Components
- Varying Component Size and Extent
- Network Connection
- Uniform Set of Services
- Certain Global Properties
Distributed Computing Approaches

- Grid
  - Addressing Infrastructure

- Peer-to-Peer
  - Addressing Failure
    - Self-organizing into network topologies
  - w/o a Global Server or Authority
Distributed System Types

- Distributed Computing Systems
  - Cluster Computing Systems
    - Collection of computers connected in a high-speed network
  - Grid Computing Systems
    - Federation of computer systems possibly in different administrative domains

- Distributed Information Systems
  - Transaction Processing Systems
    - Atomic, Consistent, Isolated, and Durable Transaction
Distributed System Types (Cont’d)

- Distributed Information Systems (Cont’d)
  - Enterprise Application Integration
    - Communicating Independent Components
      - Remote procedure call
      - Remote method invocation
      - Message-oriented middleware w/ logical contact points
        - Message queuing model
        - Message brokers as application-level gateway w/ subscription & publication
  - Distributed Pervasive Systems
    - Home Systems Possibly w/ UPnP
    - Electronic Health Care Systems in a BAN
    - Sensor Networks
Issues

- Problem
  - No Global Clock
- Issues
  - How to Determine an Order of Events
  - How to Determine Global States
    - Consistency
Clock

- **Skew**
  - Difference between the Readings of Any Two Clocks

**Skew between computer clocks in a distributed system**

- **Drift**
  - Divergence of Clocks due to Counting Times at Different Rates
UTC (Coordinated Universal Time)

- **International Time Standard**
  - Formerly, Greenwich Mean Time or GMT
    - Zero hours UTC: midnight Greenwich (0 degrees longitude)
  - Based on Atomic Time (Drift Rate: $1/10^{13}$ Seconds/Second)
    - Signals synchronized and broadcast regularly
      - From land-based radio stations and satellites
Synchronous vs Asynchronous

- **Synchronous Systems**
  - Known Bounds
    - Drift rate of clocks
    - Max message transmission delay
    - Time to execute each step of a process

- **Asynchronous Systems**
  - No Bounds