Chapter 03

Networking and Internetworking

- Introduction
- Types of network
- Network principles
- Internet protocols
- Network case studies
  - Ethernet, wireless LAN
- Summary
Introduction

- Why should we study network?
- Networking issues for distributed systems

Concepts

- Transmission media
  - wire, cable, fiber and wireless channels
- Hardware devices
  - router, switch, bridge, hub, repeater and network interface
- Software components
  - protocol stacks, communication handler and driver
- Communication subsystem
  - Hardware devices + software components
Concepts ... continued

- **Host**
  - computers and devices that use the network for communication purpose

- **Node**
  - computer or switching device attached to a network

- **Subnet**
  - a unit of routing (delivering data from one part of Internet to another); it’s a collection of nodes that can be reached on the same physical network.

---

Performance

- **latency**
  - the delay that occurs after a send operation is executed before data starts to arrive at the destination computer.
  - Software overheads, routing delays, delay of accessing to channels

- **data transfer rate**
  - the speed at which data can be transferred between two computers in the network once transmission has begun, bits/s

- **Message transmission time**
  - latency + length / data transfer rate
Performance ... continued

- **total system bandwidth**
  - the total volume of traffic that can be transferred across the network in a given time
    - Ethernet: system bandwidth is as same as data transfer rate
    - WAN: multiple channels, deteriorates when there are too many messages

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>local network</td>
<td>a null message transmission time is under a millisecond</td>
</tr>
<tr>
<td>local memory</td>
<td>1000 or more times faster than local network</td>
</tr>
<tr>
<td>local hard disk</td>
<td>500 times slower than fast local network</td>
</tr>
<tr>
<td>Internet</td>
<td>round-trip latencies are in 300-600ms</td>
</tr>
</tbody>
</table>

Scalability

- **future Internet**: several billion nodes and hundreds of millions of active hosts, new addressing and routing mechanisms

<table>
<thead>
<tr>
<th>上网的计算机数</th>
<th>3630/4560万</th>
</tr>
</thead>
<tbody>
<tr>
<td>上网的用户数</td>
<td>8700/10300万</td>
</tr>
<tr>
<td>cn下注册的域名</td>
<td>382,216/622,534</td>
</tr>
<tr>
<td>WWW站点数</td>
<td>626,200/677,500</td>
</tr>
</tbody>
</table>
Reliability

- Networks are highly reliable, whereas client and server computers and their software often aren’t, so error detection and recovery is best performed end-to-end at the highest feasible level.
- Errors: e.g.
  - failures in software of sender or receiver
  - buffer overflow

Security

- Firewall: to protect the resources in all of the computers inside the organization from access by external users or processes and to control the use of resources outside the firewall by users inside the organization, always runs on a gateway
- Secure network environment, e.g. VPN
Mobility

- Although the current mechanisms have been adapted and extended to support them, the expected future growth in the use of mobile devices will require further extension.

Addressing and routing schemes of the Internet and other networks

QoS

- require guaranteed bandwidth and bounded latencies.

Streams of real-time multimedia data transmission
Multicasting

- Need for one-to-many communication

Networking and Internetworking

- Introduction
- Types of network
- Network principles
- Internet protocols
- Network case studies: Ethernet, wireless LAN
- Summary
Local area networks (LANs)

- High speed, connected to a single communication medium, no routing of messages, may have switches or hubs
- Ethernet, Token rings and ... (1970s),
- 10M -> 100M -> 1000M (G)

Metropolitan area networks (MANs)

- Ethernet, ATM, DSL (digital subscriber line)
Wide area networks (WAN)

- lower speeds, links between different cities, countries or continents
- communication medium is a set of circuits linking a set of dedicated computers called routers

Wireless networks

- IEEE 802.11 (WLAN): 2-11 mbps over 150 meters
Wireless networks ... continued

- WPANs (wireless personal area networks): infra-red links, BlueTooth (1-2 mbps over 10 meters)

Wireless networks ... continued

- digital mobile phone network: European GSM/USA CDPD (up to 2 mbps)
- WAP (Wireless Application Protocol)
Internetworks

- several networks are linked together to provide common data communication facilities that conceal the technology differences.
- They are interconnected by routers and gateways
- e.g. Internet

Network comparisons

- Different failure model
  - TCP vs. UDP
- Packets may be lost
  - Delays
  - Overflow at the destination
Networking and Internetworking

- Introduction
- Types of network
- Network principles
- Internet protocols
- Network case studies: Ethernet, wireless LAN
- Summary

Packet, message, data streaming

- **Messages**
  - arbitrary length

- **Packet**
  - restricted length
  - Sufficient buffer, avoid undue delays

- **Data streaming**
  - Video stream: bandwidth requirement, continues flow e.g. frame N arrives no more than N/24 seconds after the first frame arrives
  - Bandwidth, latency and reliability must be guaranteed: predefined route, e.g. RSVP/IPv6
Network Switching technology

- **Switching schemes**
  - Broadcast: no switching, Ethernet
  - Circuit switching: telephone system
  - Packet switching: postal system

- **Packet switching**
  - Store and forward, share communication link, asynchronous

- **Integrated telecommunication system?**
  - Next generation network

Network Protocols

- **Protocol**
  - A specification of the sequence of messages that must be exchanged
  - A specification of the format of the data in the messages

- **Protocol suits (protocol stack)**
  - Simplify and generalize the communication interface, but bring significant cost, e.g., N copies
  - Seven-layer Reference Model for OSI adopted by ISO
  - TCP/IP
Protocol Layers

- Each layer presents an interface to the layers above it that extends the properties of the underlying communication system.

Encapsulation of a packet

- **Layer encapsulations**

  - AL message
  - PL Header
  - AL message
  - SL Header
  - PL
  - TL Header
  - SL
  - NL Header
  - TL
  - DL Header
  - NL

- **Packet assembly**
  - Network-layer protocol packets: MTU (Maximum transfer unit), 1500Bytes in Ethernet, 64K in IP
**Ports**

- **Echo**
  - Protocol: 7/tcp
  - Description: Echo
- **Echo**
  - Protocol: 7/udp
  - Description: Echo
- **File Transfer [Default Data]**
  - Protocol: 20/tcp
- **File Transfer [Default Data]**
  - Protocol: 20/udp
- **File Transfer [Control]**
  - Protocol: 21/tcp
- **File Transfer [Control]**
  - Protocol: 21/udp
- **Telnet**
  - Protocol: 23/tcp
- **Telnet**
  - Protocol: 23/udp
- **Simple Mail Transfer**
  - Protocol: 25/tcp
- **Simple Mail Transfer**
  - Protocol: 25/udp
- **Domain Name Server**
  - Protocol: 53/tcp
- **Domain Name Server**
  - Protocol: 53/udp
- **World Wide Web [HTTP]**
  - Protocol: 80/tcp
- **World Wide Web [HTTP]**
  - Protocol: 80/udp

**Packet delivery**

- **Datagram packet delivery, e.g. IP, and most wired and wireless LAN**

[Diagram of packet delivery]
Packet delivery ... continued

- Virtual circuit packet delivery

Routing algorithm

- Determine the route taken by each packet.
  - Predetermined for circuit switching (e.g. X.25) and frame relay switching (e.g. ATM);
  - determine on the fly for packet switching;
- Dynamically update its knowledge of the network
**Distance vector**

- Bellman-Ford shortest path algorithm [1957]
- RIP (router information protocol)
  - Periodically, and whenever the local routing table changes, send the router table to all accessible neighbours
  - When a table is received from a neighbouring router, make necessary update

- Convergence problem
- Improvement: link cost include bandwidth information, speed convergence, etc

---

**Distance-Vector Routing table**

### Routings from A

<table>
<thead>
<tr>
<th>To</th>
<th>Link</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>local</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

### Routings from B

<table>
<thead>
<tr>
<th>To</th>
<th>Link</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>local</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>E</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>

### Routings from C

<table>
<thead>
<tr>
<th>To</th>
<th>Link</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>local</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>E</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

### Routings from E

<table>
<thead>
<tr>
<th>To</th>
<th>Link</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>local</td>
<td>0</td>
</tr>
</tbody>
</table>

### Routings from D

<table>
<thead>
<tr>
<th>To</th>
<th>Link</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>local</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>

---

*Copyright ©2005 YAN Wei. All rights reserved.*
**Pseudo-code for RIP routing algorithm**

*Send:* Each \( t \) seconds or when \( T_l \) changes, send \( T_l \) on each non-faulty outgoing link.

*Receive:* Whenever a routing table \( T_r \) is received on link \( n \):

for all rows \( R_r \) in \( T_r \) {
  if \((R_r.link <> n)\) {
    \( R_r.cost = R_r.cost + 1; \)
    \( R_r.link = n; \)
    if \((R_r.destination \) is not in \( T_l \)) add \( R_r \) to \( T_l \);
    // add new destination to \( T_l \)
  } else for all rows \( R_l \) in \( T_l \) {
    if \((R_r.destination = R_l.destination \) and \((R_r.cost < R_l.cost \) or \( R_l.link = n))\) \( R_l = R_r; \)
    // \( R_r.cost < R_l.cost \) : remote node has better route
    // \( R_l.link = n \) : remote node is more authoritative
  }
}

**Routing ... continued**

- **Link state algorithm**
  - Each node maintains all knowledge of the network
  - Each node can compute appropriate routes based on the knowledge
  - Avoid slow convergence and undesirable intermediate states
  - E.g. OSPF
Connection and connectionless

- Different to the concepts of Connection oriented (TCP)/Connectionless (UDP) in transport layer protocol

![Diagram of connection and connectionless]

Congestion control

- A rule of thumb
  - when the load on a network exceeds 80% of its capacity, the total throughput tends to drop as a result of packet losses

- Congestion control
  - instead of allowing packets to travel through the network until they reach over-congested nodes, hold them at earlier nodes
Control approaches

- Informing nodes along the congested route to reduce packet transmission rate, i.e. buffering for long time at intermediate nodes or queue packets at source host.
- In the Internet, congestion control rely on the end-to-end traffic control, e.g. choke packets requesting a reduction in transmission rate in TCP.

Internetworking

- **Internetwork**
  - Integrate many subnets that use different network technologies.
- **Requirements**
  - Unified internetwork addressing scheme that enables packets to be addressed to any host connected to any subnet.
  - A protocol defining the format of internetwork packets and giving rules according to which they are handled.
  - Interconnecting components that route packets to their destinations in terms of internetwork addresses, transmitting the packets using subnets with a variety of network technologies.
QMW Computer Science network

Internetworking components

- **Router**
  - Conduct routing, additionally link networks of different types

- **Bridge**
  - Link networks of different types, but not conduct routing

- **Hub**
  - Connect hosts and extend segments of Ethernet and other broadcast local network

- **Switch**
  - Perform similar function to router, but for LANs only

- **Tunnel**
  - A software layer that transmits packets through an alien network environment
Tunnelling

- IPv6 encapsulated in IPv4 packets

- IP encapsulated in PPP packets
Networking and Internetworking

- Introduction
- Types of network
- Network principles
- Internet protocols
- Network case studies: Ethernet, wireless LAN
- Summary

Internet protocols

- Protocol layers
  - TCP(UDP)/IP, web [HTTP], Email [SMTP,POP], news [NNTP], FTP, SSL, etc
- Exceptions to the universal adoption of TCP/IP
  - The use of WAP for wireless applications on portable devices
  - Special protocols to support multimedia streaming applications
Heterogeneous underlying networks support
- The success of TCP/IP: independence of the underlying transmission technology
- E.g., IP over ATM, IP over Ethernet, IP over PPP, etc
The programmer’s conceptual view of a TCP/IP Internet

- SOCKET
- TCP
- UDP
- IP
- Underlying network

Schemes for naming and addressing hosts and for routing IP packets to their destination is challenging.

- Requirement
  - It must be universal
  - It must be efficient
  - The addressing scheme must lend itself to the development of routing scheme

- The scheme
  - A 32-bit numeric identifier containing a network identifier and a host identifier
  - There are four allocated classed of Internet address-A, B, C, D
### Internet address structure

1.0.0.0 ~ 127.255.255.255
128.0.0.0 ~ 191.255.255.255
192.0.0.0 ~ 223.255.255.255
224.0.0.0 ~ 239.255.255.255
240.0.0.0 ~ 247.255.255.255

<table>
<thead>
<tr>
<th>0</th>
<th>Network</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Network</td>
<td>Host</td>
</tr>
<tr>
<td>110</td>
<td>Network</td>
<td>Host</td>
</tr>
<tr>
<td>1110</td>
<td>Multicast address</td>
<td></td>
</tr>
<tr>
<td>11110</td>
<td>Reversed for future use</td>
<td></td>
</tr>
</tbody>
</table>

Two steps were taken: IPv6, Classless interdomain routing (CIDR)

### IP protocol

- Transmits datagrams from one host to another, if necessary via intermediate routers
  - Unreliable (best-effort) delivery semantics
    - Packets can be lost, duplicated, delayed or delivered out of order
  - Address resolution: Address Resolution Module (ARP)
    - IP address \(\rightarrow\) Ethernet address mapping, (IP address, Ethernet address) pairs cache on each host
IP routing

- Routes packets from source to destination
  - Internet topology
    - Autonomous System, Areas
  - Routing algorithms
    - RIP-1
    - RIP-2
    - Open Short Path First (OSPF)
  - Default routes: trade routing efficiency for table size
  - Classless interdomain routing (CIDR): create subnet by means of subdividing address or aggregating addresses by mask field, e.g. 162.105.203.0/24

Internet topology

- Backbone topology
- Graph of Internet topology
- Internet multicast topology
### Future of IP

- **IPv6**
  - $2^{128}$ ($3 \times 10^{38}$) addresses, 1000 IP addresses per square meter of the Earth’s surface
  - Routing speed: no checksum, no fragmentation
  - Real time: priority and flow label which is used to reserve resources
  - Extension header (information of router, authentication, etc).
  - Multicast and anycast
  - Security through extension header type

- **Migration from IPv4**
  - IPv6 router island
  - Depend on economics

### IPv6 header layout

<table>
<thead>
<tr>
<th>Base Header</th>
<th>Extension Header1</th>
<th>......</th>
<th>Extension HeaderN</th>
<th>Data</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>4</th>
<th>4</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Pri</td>
<td>Flow label</td>
</tr>
</tbody>
</table>

- Payload length
- Next header
- Hop limit

- Source address(16B)
- Destination address(16B)
The MobileIP routing mechanism

- **Foreign agent**: 解移动节点的位置
- **Home agent**: 跟踪记录移动节点的位置

1. 寻址移动节点家乡地址的分组
2. 家乡代理截获该分组并转发给外地代理
3. 外地代理接收此分组并转发给移动节点
4. 移动节点直接响应对等节点

<table>
<thead>
<tr>
<th>IP Address</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>162.105.203.0/24</td>
<td>Permanent address: 162.105.203.16</td>
</tr>
<tr>
<td>166.111.4.0/24</td>
<td>Care of address: 166.111.4.1</td>
</tr>
</tbody>
</table>

TCP and UDP

- **Use of ports**
  - Provide process-to-process communication
- **UDP features**
- **TCP features**

<table>
<thead>
<tr>
<th>Port</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Port-QQ</td>
<td>162.105.111.118</td>
</tr>
<tr>
<td>Port-IE</td>
<td>162.105.111.118</td>
</tr>
<tr>
<td>Port-MSN</td>
<td>162.105.111.118</td>
</tr>
</tbody>
</table>
UDP features

- **Connectionless**
- **Datagram delivery**
  - A UDP datagram is encapsulated inside an IP packet, up to 64kb in size

- **Con**
  - unreliable delivery due to unreliable IP

- **Pro**
  - minimal additional cost and transmission delays

TCP features

- **Connection oriented**
  - two side must shake hands to establish a bi-directional communication channel

- **Message delivery**
  - Deliver arbitrary long sequences of bytes via stream-based programming abstraction
  - Sequencing: divide stream into data segments, sequence number on each segment
  - Checksum: cover the header and the data in the segment

- **Flow control**
  - Receiver send the highest number of received segment and window size to sender by acknowledge message
  - Buffering: receiver buffer and sender buffer used for flow control
  - In interactive application, receiver inform sender when timeout or the buffer reaches the MTU limit
  - Retransmission: retransmit the segment when no acknowledgement within a specified timeout
Domain names

- Symbolic names for hosts and networks
  - net.pku.edu.cn
  - yanwei.grids.cn
- The DNS would not workable without the extensive use of caching.

Firewall

- The purpose of a firewall is to monitor and control all communication into and out of an intranet
  - service control
  - behavior control
  - user control
- Filter approaches
  - IP packet filtering, e.g. router/filter
  - TCP gateway, e.g. bastion
  - Application level gateway, e.g. telnet proxy process
- Virtual private networks (VPN)
  - Secure connections located at different sites using public Internet links
  - By the use of cryptographically protected secure channels at the IP level
Networking and Internetworking

- Introduction
- Types of network
- Network principles
- Internet protocols
- Network case studies: Ethernet, wireless LAN and ATM
- Summary

Ethernet

- IEEE 802.3[Xerox 1973]
  - Carrier sensing, multiple access with collision detection
  - Frame broadcasting
  - Bandwidth: 3m -> 10m -> 100m -> 1000m

- Packet collisions
  - Carrier sensing
    - wait until no signal is present then transmit
  - Collision detection
    - When transmit through output port, also listen on the input port, and compare the two signals, if differ, send jamming signal
  - Back-off
    - wait a time $n\tau$ before retransmitting, $n$: a random integer
Ethernet ... continued

Ethernet packet layout

- $2^{48}$ different addresses

<table>
<thead>
<tr>
<th>Preamble</th>
<th>SFD</th>
<th>DA</th>
<th>SA</th>
<th>Length</th>
<th>data</th>
<th>PAD</th>
<th>FCS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

“10101010”

“10101011”

Collision detection / stop transmission

t0

t1

© 2005 YAN Wei. All rights reserved.
Ethernet ... continued

- Ethernet efficiency
  - Efficiency = number of packets transmitted successfully / theoretical maximum number without collision
  - Affected by
    - A finite time for a signal inserted at a point in the media to reach all other points
    - number of stations on the network
    - stations’ level of activity

Wireless LAN

- Wireless LAN types
  - Infrastructure network, e.g. IEEE 802.11
  - Ad hoc network: network built on the fly

- Collision detection failures in 802.11
  - Hidden stations: carrier sensing fail to detect that another station on the network is transmitting, lead to collision at base station
  - Fading: the strength of radio signals diminishes rapidly with the distance from the transmitter, so that defeating both carrier sensing and collision detection
  - Collision masking
### Hidden stations

- A
- B
- C

There is a conflict between devices A and B due to hidden stations.

### Wireless LAN configuration

- **AP Scheme**
- **Ad hoc Scheme**

The network configuration includes devices connected via a backbone and wireless connections.
Slot reservation added to the MAC protocol in 802.11

- Firstly, sense the medium, if no carrier signal, then
- Sender send a RTS (Request To Send) frame to receiver; Receiver reply a CTS (Clear To Send) frame to sender. The effect of the exchange is:
  - the station within range of sender will pick up the RTS frame and take note of the duration
  - the station within range of receiver will pick up the CTS frame and take note of the duration
- Begin to transmit