Week1 Review Notes

ELEC0099: Introduce to Internet Protocol Networks 21/22

RUFENG DING

Contents

1 Introduce to the Internet

The Internet contains three main parts:

1. Connected computing devices:hosts/end-systems
   PCs,workstations,servers,smartphones,toasters, running network apps.
2. Comunication links
   fiber,copper,radio,satellite(**transmission rate = bandwidth**)
3. routers:forward packets
   chunk of data

1.1 History

1. 1961 *Leonard Kleinrock* create the Queuing Theory
2. 1965 *Bob Taylor and Larry Roberts* start tge ARPANET project
3. 1965 *Donald Davies (UK)* invents the packets switching
4. 1972 ARPAnet demonstrated publicly
   - NCP (Network Control Protocol) first host-host Protocol
   - first e-mail program
   - ARPAnet has 15 nodes
5. 1970 ALOHAnet: a wireless network to connect islands in Hawaii
6. 1973 Ethernet *Robert Metcalfe's PhD Thesis*

7. 1974 *Cerf and Kahn* define the principles for interconnecting networks (First network outside USA is in UCL)

   - minimalism,autonomy - no internal changes required to interconnect networks
   - best effort service model
   - stateless routers
   - decentralized control

8. 1983 TCP/IP and DNS
9. 1988 TCP Congestion Control
10. 1989 100,000 hosts a lot of new national networks
11. 1991 NSF allows for commmercial use
12. 1992 The World Wide Web developed by *Tim Berners Lee* in CERN
13. Late 90's New applications come out : p2p file sharing / Network security / Voice over IP
14. 2000 Web 2.0 Interactive app over-the-top applications (OTT)
15. 2007 iPhone use Cellular access (Cellular already dominant but not used in mobile devices)
16. Today's trends
    - video streaming
    - distribution network
    - clouds
    - software defined networks
    - privacy

Internet Access Statistics World Internet usage and population statistics

1.2 Internet design principles

End-to-end argument: Important functions(error control, encryption, delivery, acknowledgements, etc) should be implemented by the end systems.

Fate Sharing: if one put functionality in the end system, then that functionality only breaks if the end system breaks which would make the communication useless anyway

Russian Dolls: a lot of packets inside packests

Packet Switching: Statistical Multiplexing: Sequence of packets do not have fixed pattern

Note 1. Statistical multiplexing is a type of communication link sharing, very similar to dynamic bandwidth allocation (DBA). In statistical multiplexing, a communication channel is divided into an arbitrary number of variable bitrate digital channels or data streams. The link sharing is adapted to the instantaneous traffic demands of the data streams that are transferred over each channel. This is an alternative to creating a fixed sharing of a link, such as in general time division multiplexing (TDM) and frequency division multiplexing (FDM). When performed correctly, statistical multiplexing can provide a link utilization improvement, called the *statistical multiplexing gain*.

1.3 Future applications

Video is King: Video will be 82% of the traffic in 2021. 4K TV is being deployded: 25Mbits/s (5 times more than HD). 8K is next: 4 times more pixels and there will be more end-users.

Vitual Reality.

Augmented reality.

Holograms.

The Smart Home.

Transport.

Smart cities.

Internet of things (IoT) and Machine2machine (M2M).

5G.

1.4 Examples of Networks

JANET(The UK education network)

Topologies: Topology-zoo

2 Computer Networking

2.1 Types of Networks

2.1.1 LAN

Local Area Network. Connected computers that are physically close together.(<1 mile) eg. Data Center Networking
Characters:

- high speed
- multi-access

Technologies:

- Ethernet (10Mbps,100Mbps,1Gbps)
- Token Ring 16Mbps
- FDDI 100Mbps

2.1.2 WAN

Wide Area Network.Connects computers that are physically far apart."long-haul network".
Characters:

- Higher delay that LAN / Fast speeds
- Traditionally less reliable than a LAN
- point-to-point

Technologies:

- Telephone lines
- Satellite communications
- SONET/SDH - ATM
- Increasingly Ethernet

2.1.3 MAN

Metropolitan Area Network **Characters**:

- larger than LAN and smaller than WAN
- Campus-wide network (FDDI,DQDB,ATM)
  - FDDI: Fiber Dsitributed Data Interface
  - DQDB: Distributed-Queue Dual-Bus
  - ATM : Asynchronous Transfer Mode
- Interconnects LANs

Technologies:

- Coaxial cable
- Microware
- Optical

eg. A big MAN : LinX : The london internet eXchange point

2.1.4 Home Networking

New challenges and opportunities
Access:

- ADSL
- Optical fiber
- Wireless, Satellite

Access inside the house:

- Wifi, Bluetooth, power cables
- Wifi signal, wifi throughout, wifi leaking
- Internet of things(wifi zigbee)

2.1.5 Low Power Networking: LORA and Narrowband IoT

Characters:

- Long distance
- 9v battery
- up to 10 years
- Bytes per minutes

2.1.6 Personal Area Networks

A niche market but an important one.
Characters:

- Healthcare and Internet of things
- Bluetooth dominates (Wifi low energy may become relevant)
- energy is critical!(You don't want to charge these devices all the time)

2.1.7 Space Networking: Interplanetary Internet

To sustain life and travel to other planets and it has a very high delay (eg. Earth-Mars: 4 to 24 minutes)

2.2 Types of Addressing

There are 4 addressing modes:

- **unicast**: message sent to one destination
- **broadcast**: message sent to all hosts in a network
- **multicast**: message sent to all members of a group
- **anycast**: message sent to closest member of a group

2.3 Quality of Services metrics

Throughput
Measured in bits per second and it depends on:

- Transmission rate of the links in the path
- packets loss
- packets delay
- application requirements

Packet loss
caused by:

- congestion in the network
- active queuing policies
- link failures

Delay
some applications require low end-to-end delay:

- vocie: 150ms
- interactive services: 10ms
- financial services: each ms costs money

Five components:

- Propagation
  dominated by the speed of light on fiber (aprox 210,000 km/s) and increased by the fact that fibers do not follow straight lines.
- Queuing
  Each link has a queue. For lightly loaded links never more than 10 packets. (overloaded links can have 100ms delay)
- Processing(eg, middleboxes)
  Routers, wireless points, end hosts takes times to analyse the packets and determining what to do with them. Routers need to check for error, determine next link, etc. In normal conditions the processing delay is very small but if data processing is involved, it can be a major source.

- Packetzations(Increasing by lower layers)
  Time to require the data to be sent. (eg. Voice to be encoded. If it's coded at 32 Kbps you have to wait 20ms to get 80 bytes) Lower layers: ADSL interleaving, used to ”shuffle” fragments to increase reliability. And it add about 5ms of delay.

- Transmission(data rate)
  Determined by data rate links. (eg. On 1 Mbps link takes 10ms to send a 10000 bit packet/ On 10Gbps takes 1μs on 50Kbits/s, 200ms)

3 The OSI stack

3.1 Connection-Oriented vs connectionless

connection-oriented services uses *circuits* a single path is first established for each new connection(call setup/call release) and the *network* guarantees that the data are delievered in order, no loss or duplication.

If any thing goes wrong the connection is broken. It is possible to limit the number of connections. The network can guarantee bandwidth at connect time.(waste of bandwidth, if resources not used) And the network can refuse new connections.

eg. Telephony(PSTN), cellular network, ISDN, ATM

Connectionless Service uses *datagrams.* Each datagram is independently routed and includes the destination address.

No guarantee that the datagrams are delievered in order and are not lost or duplicated. It is direct transmission of data(no call set-up). And best-effort earnest attempt to deliver.

Resources are shared and little waste of bandwidth. If the network overly utilized, further traffic is still allowed.

eg. Postal Service, Ethernet, Internet Protocol(IP)

CO vs CL Services

3.2 Understand the need for Layering

Divide a task into separate functions and then define each piece independently.

Eastablishing a well defined interface between layers makes porting easier (Decoupling).

Major advantages:

Table 1: CO vs CL Services

	Type	Pros/Cons
PROS	CO	reliabilitym file transfer and terminal traffic main applications
		Faster forwarding, after call setup the path is constant
		Better to lock-out further cells, than to degrade service
		”Simple” Terninal Equipment, offload complexity to the core network
	CL	Fault tolerant, if link fails other paths available
		Applications like voice and video can tolerate datagram loss
		Better suited to bursty traffic, link reservation is a waste
		Fair, better allow user access, not only some lucky users
		Efficient for client-server applications with hundreds of clients
		”Simple” core network

- Code Reuse
- Extensibility
- Division of tasks. Each engineering implementing a layer does not need to know how the lower one is implemented.

3.3 OSI model and layers functions

OSI Protocol Reference Model

Parallel work to hte Internet development started in 1970s. It published its first version in 1984 by ISO(International Standard Organization). It consist two parts:

- reference model organizing networking functions in 7 layers
- a set of specific protocols for each layer(never really deployed)

It was thought to be the ”serious” standard.

Table 2: OSI Protocol Reference Model

No.	Layer	function	scope
7	application	application-specific protocols	end-to-end
6	presentation	data representation and encoding	
5	session	dialog and sychronisation	
4	transport	message transfer(error/flow/congestion control,CM)	
3	network	network routing/addressing(CM)	global
2	data link	data link control(framing,data transparency,EC)	local
1	physical	mechanical/electrical/optical interface	

CM: Connection management EC: Error control

3.3.1 Physical Layer

Responsibility:transmission of raw bits over a communication channel. (Encode format /cable types etc.)

Issues:

- mechanical and electrical interfaces
- time per
- distances

> **Note 2. Manchester Code** is a line code in which the encoding of each data bit is either low then high, or high then low, for equal time. It is a self-clocking signal with no DC component. Consequently, electrical connections using a Manchester code are easily galvanically isolated.

And there are lots of digital encoding format: NRZ-L/NRZI/Bipolar-AMI/Pseudoternary etc.

3.3.2 Data Link Layer:Ethernet Frame

Ethernet uses a method called **Carrier Sense Multiple Access with collision Detection** (CSMA/CD).

In Ethernet the medium is shared, computers **sense** the medium to check if anybody is tranmitting and if not, they start sending. While sending they check if the wave is changed. If it does than a **collision is detectd**. The computer waits a random amount of time and then retransmits again.

Another method is called the **Carrier-sense multiple access with collision avoidance** (CSMA/CA). It is uesd in Wifi/802.11. The Sender need to sned RTS to Revceiver to get a CTS and then send the Data packet. When finished get ACK from Receiver.(hand shake)

eg. EPC (Electronic Product Code) RFID reader and tag.

Error Detection
Networks are unrelable so the data link can help here:

Error correcting codes:

1. Hamming codes
2. Binary convolutional codes
3. Reed-Salomon codes
4. Low-Density Parity Check codes

Error detecting codes

1. Parity
2. Checksum
3. Cyclic Reducndancy Code (CRC)

3.3.3 Network Layer

Responsibilities:

- Path selection between end-systems(routing).
- Subnet flow control.
- Fragmentation and reassembly.
- Translation between different network types.

Issues:

- *packet* headers
- virtual circuits

Non-IP network layer: (historically) ATM, CLNP. There are other design choices include: admission control/ congestion control/ reliability.

3.3.4 Transport Layer

Responsibilities:

- provides virtual end-to-end links between peer processes.
- end-to-end flow control/ congestion control.
- reliable communication.

3.3.5 Session Layer

Responsibilities:

- Dialog Control(Keep track of whose turn is it to transmit)
- Token Management(preventing both parties from attempting the same critical operation simultaneously)
- Synchronization(checkpointing long transmission to allow time to pick up from where they left the event of a crash and subsequent recovery)

In TCP/IP it does not exist formally. Many people identify some functions provided by some protocols as being part of **Session Layer**. eg:

- RTP and RTCP for video conferencing
- Cookies for http
- SIP -Session Initiation Protocol
- ...

Session layer functionality is usually implemented by the application itself.

3.3.6 Presentation Layer

Responsibilities:

- data encryption
- data compression(lossless or lossy)
- data conversion

Many protocol suits do not include a Presentaion Layer.

Encryption Two kinds: Symmetric and Asymmetric. It allows for confidentiality and authentication and the management of public keys is the hardest part.

3.3.7 Application Layer

Responsibilities: anything not provided by any of the other layers.

Issues:

- application lebel protocols
- approprate selection of ”type of service”

Examples of application layer protocols:

- **HTTP** for the world wide web. This is the protocol used for web brosers to get web pages form web servers.
- **SMTP,POP,IMAP** to read and receive email.
- signaling protocols like **SIP** for voice over **IP**.
- **Routing protocols** are application level protocols that affect the Network Layer.

They are very easy to implement in any programming language.

3.4 How to connect networks

- Repeater: physical layer
- Bridge: datalink layer
- Router: network layer
- Gateway: network layer and above

Repeater copies bits form one part of the sme network to another and it does not look at any bits, just regenerates them. It allow the extension of a network beyond physical length limitations.

Hub is a star topology - multi-access device. It regenerate bits and provide muti-port device linking network segments. It copy all vilid data to all ports and allow the extension of a network beyond physical length limitations.

Bridge copies frames from one part of a network to another. And it can operate selectively - does not copy all frames(must look at data-link headers). Provides isolation and so improves performance. It can extends the network beyond physical length limitations.

Router transfer packets from one network to another. Makes decisions about what *route* a packet should take (looks at network headers).

Table 3: Criticisms about OSI and TCP/IP

OSI	TCP/IP
bad timing	not distinguish clearly between service/interface/protocols
bad technology/complex standard	not easily exatendatble/lacks generality
bad implementations	link layer not well defined
bad politics	many protocols have bad implemantations

Two elephants lead the world using TCP/IP (OSI using too many times define the standard).

Week2 Review Notes

ELEC0099: Introduce to Internet Protocol Networks 21/22

RUFENG DING

Contents

1 Ethernet

Why there is a need for a LAN?

- to share resources like files, printers, scanners, internet connections, WAN links.
- To share data and applications like common database help desk software.
- To increase productivity by making it easier to share data among users.

- To facilitate network management by making the networked compouters accessible to the administrator from a centralised site.

1.1 LAN

LAN topologies

- Bus

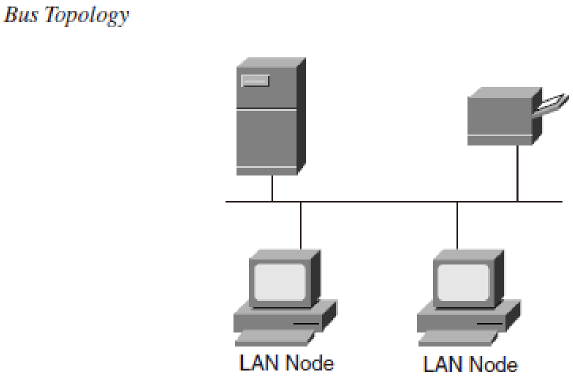


Figure 1: Star topology

- Tree

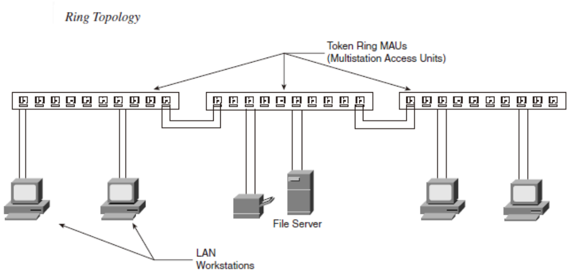


Figure 2: Star topology

- Ring
- Star

LAN transmission methods

- Unicast transmission a frame is sent from the source to the destination on a network

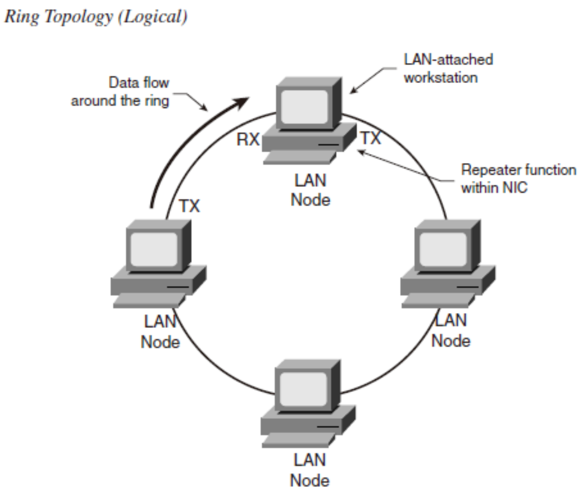


Figure 3: Star topology

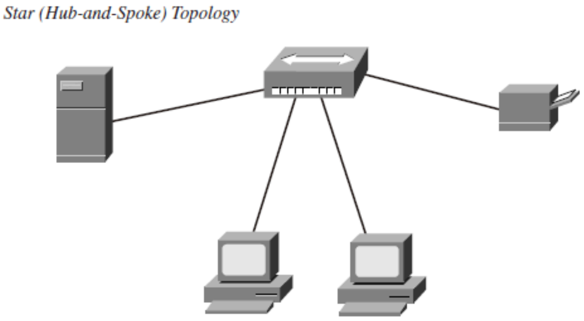


Figure 4: Star topology

- Multicast transmission a frame is snet from a source to a subset of nodes on the network
- Broadcast transmission a frame is sent to all nodes on the network

LAN protocols and OSI model

Table 1: LAN protocols and OSI model

Data Link layer	Logical Link Control (LLC)
	Medium Access Control (MAC)
Physical Layer	

LAN media access methods

- CSMA/CD where network devices contend for access to the physical network medium (eg Ethernet).
- Token passing where network devices access the physical network medium based on the possession of a token (Token ring and FDDI)

1.2 Ethernet details

Ethernet has survived as a LAN technology because:

- Flesxibility
- Relative simplicity
- Innovation
  - 100Mbps half duplex and full duplex
  - 100Mbps, 1Gbps, 10Gbps Ethernet
- costs
- Although critics claim that Ethernet cannot scale it continues to dominate the desktop market

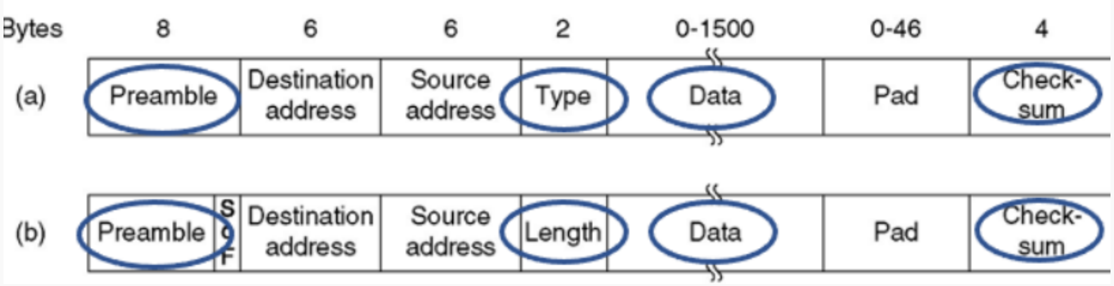
History Originally developed at the Xerox Palo Alto Research Centre in1973 patented in 1976.

In 1980 the first formal Ethernet standard was published when DEC, Intel and Xerox(DIX) joined together to publish a 10 Mbps Ethernet specification known as Ethernet Version 1.0. In 1982 the DIX alliance updated the standard to include additnal media types known as Ethernet Version 2.0.

In 1983 the IEEE 802 LAN/MAN Standards Committee published a specification for Ethernet ”IEEE 802.3 Carrier Sense Multiple Access whith Collision Detection (CSMA/CD) access method and Physical Layer Specifications”

Thus there are two types of ”Ethernet”: DIX Ethernet (original version of Ethernet) and IEEE 802.3 (standard Ethernet).

Note 1. IEEE 802.3 and DIX Ethernet Frame Formats



Ethernet type 2 uses a Type field after source address while 802.3 Ethernet use Length field (for the length of data).
Preamble - An alternating pattern of ones and zeros used to tell receiving stations athat an Ethernet fram is about to start.
Type - Specifies the upper layer protocol to receive the data after Ethernet processing is completed. (Only used in DIX Ethernet).
Length — Indicates the number of bytes if data that follow this field.
Data - Ethernet expects at least 46 bytes of data.
Frame Check Sequence (FCS) - This sequence contains a 4 bytes cyclic redundancy check value, used to checck for the presence of errors in the frame.

MAC address Destination and source addresses is called the MAC address. MAC addresses adentify network entities in Ethernet LANs.

Characters:

- Unique for each LAN interface.
- 48 bits in length
  - 22 bits identify the organisational unique identifier(OUl) and it is administered by the IEEE.
  - The last 24 bits are vendor assigned.
- The MAC address is burned in the ROM of a network interface card (NIC).
- The destination address maybe unicast.multicast or broadcast.

2 Medium Access Control

lt is the sublayer that controls the hardware responsible for interaction with the wired, optical or Wireless transmission medium.

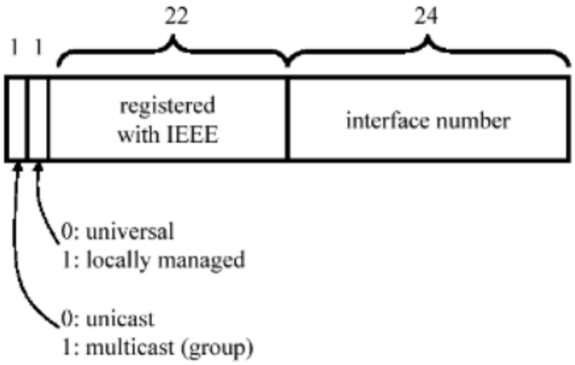


Figure 5: MAC address

2.1 Methods of Medium Access Control

Pure Aloha Translate when you want regardless of others.
Pure Aloha Collision is extremely inefficient,since the worst-case period of vulnerability is the time to transmit two frames.

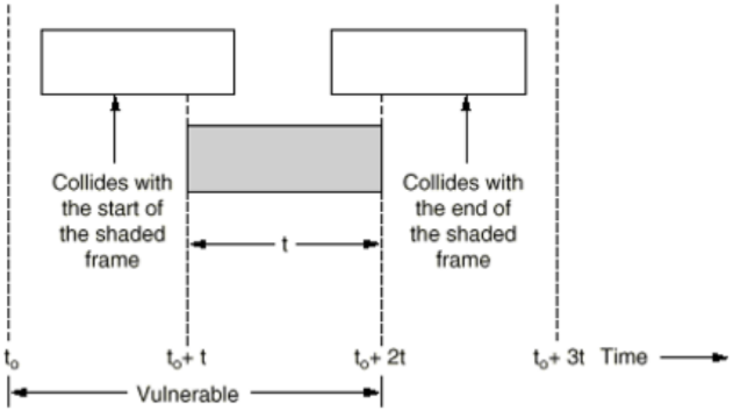

Collides with the start of the shaded frame
Collides with the end of the shaded frame
t
t_o
t_o+ t
t_o+ 2t
t_o+ 3t
Time
Vulnerable

Figure 6: Pure Aloha Collision Control

Slotted Aloha Translation only at the beginning of each Synchronized ”slot times”. And it is collision inefficient limited to one frame transmission time.
Comparation of Pure and Slotted Aloha Throughput efficiency increases dramatically for Slotted Aloha.(Figure 7)

2.2 CSMA/CD

CS - CarrierSenseIs (someone already talking?)
MA - MultipleAccess(I hear what you hear!)
CD - (CollisionDetectionHywe’re both talking!

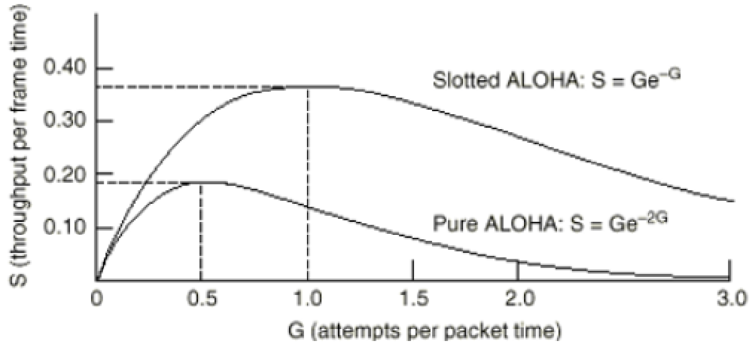


Figure 7: Pure Aloha vs Slotted Aloha

1. If the medium is idle. transmit any time.
2. lf the medium is busy, wait and transmit right after.
3. If a collision occured back off for a random period, then go back to 1.

CSMA/CD can be one of three state : contention, transmission or idle.

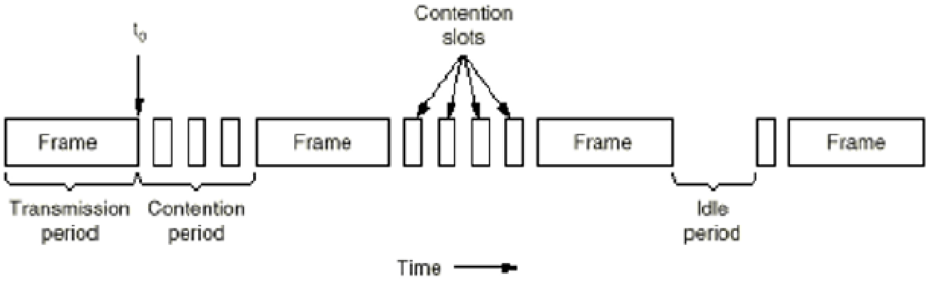


Figure 8: States in CSMA/CD

TIME is proportional to distance over the wire. (CSMA/CD on Ethernet Physical Layer). Host find the wire is clear, it began to transmit. When the packet has not arrived, another host also find the line is still clear so second hsot begins to transmit. As a result collision detected and the collision propagates and will be detected by both stations.

Performance Vertice is Throughput and horizon is Offered Load.

3 Ethernet Design

Factors Limiting the Length of Ethernet:

- Collision Detection - timing.

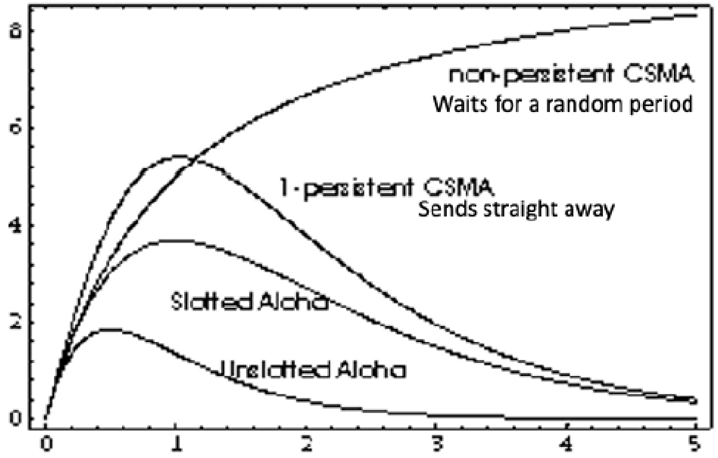


Figure 9: Performance

- Attenuation - the signal gets weaker as it propagates along the wire.
- Noise - longer wires pick up more noise Which masks the signal

Ethernet Performance An Ethernet with less than 20% utilization and less than 0.1% collisions is on **cruise control**. An Ethernet with more than 40% utilization and greater than 5% collisions **is in trouble**. If the same frame collides more than 16 times, the network interface card (NIC) will discard it.

Cable Types

- Coaxial cables
  - ThickNet
  - ThinNet
- Unshielded Twisted Pair(UTP)
  - CAT5
  - CAT6
  - CAT7

Fibre Types Firer contains three parts: Core, Cladding and Buffer. Multimode Step index / Multimode Graded Index / Singlemode.

3.1 The Collision Domain

A collision demain is defined as an area within which frames that have collided are propagated. Collision detection can take as long as 2τ, worse case. This "round-trip" delay defines the max Ethernet network diameter, or collision domain. Round-trip delay = 512 bits times for all Ethernet.

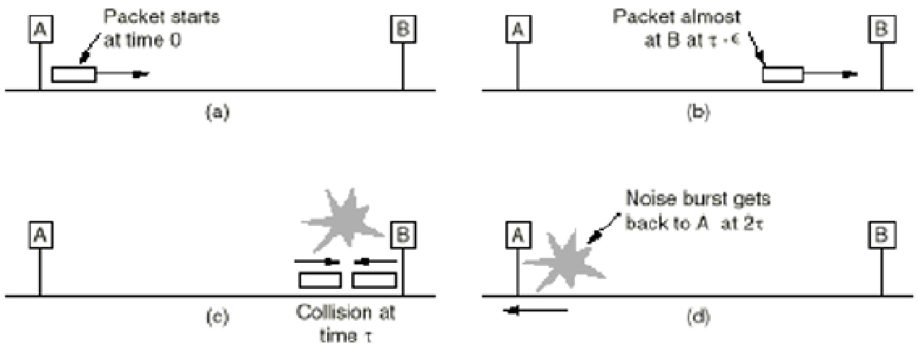


Figure 10: Collision Domain

3.2 Repeaters

Works at layer 1 (PHY layer) ONLY. *(repeaters don't understand frames they only understand BITS).* Repeat incoming signal from a port to all other ports with, restored timing, restored waveform shape, very little delay.

If 2 or more simultaneous receptions, transmit jam.

Class I repeaters may be used to repeat between media segments that use different signalling techniques (timing delays up tp 140 bits times). Class II repeaters can only connect segment using the same signalling technique (timing delays up to 92 bit times).

3.3 HUBs

HUBs are Multi-port Repeaters. All share the 10Mb ehternet bandwidth and Frames appear everywhere. It comprise a single COLLISION DOMAIN and everyone's frames collide with everyone else's/Every collision appears throughout the domain.

4 Ethernet Switch

4.1 Bridges

- Bridges separate collision domains
  - collision domains do not extend across the bridge
  - Timing rules ”restart” at a bridge port
  - Bridge is a store and forward device
- Bridges Properties

  - Frame Forwarding
    The bridge receives a frame on one port and transmit it on another port.

    The bridge stores (buffers) the frame:

    * The other port checks that the wire is clear
    * the other port transmits the frame
    * if a collision occurs, back off and retransmit

    Collision don't propagate across bridges.

    Bridges can connect dissimilar networks.
  - Learning
    * The bridge examines the layer 2 source addresses of every frame on the attached networks (promiscuous listening).
    * The bridge maintains a table , or cache, of which MAC addresses are attached to each of its ports.
  - Filtering
    * The bridge examines the destination MAC address of each frame on its attached networks.
    * If the destination is on the same port as the source, the frame is not forwarded.
    * The frame is forwarded ONLY to the port the destination is attached to.
    * Eliminates unnecessary traffic on the attached networks.
  - Spanning Tree

  More Forwarding and Filtering

  A broadcast is a frame destinated for every host on the network. Bridges forward broadcasts to every one of their ports - called ”Flooding”.

  If a bridge sees a destination address it has not yet learned, it also floods that frame.

  Bridges are called layer 2 devices because they examine layer 2 information and modify their behaviour accordingly.

4.2 Switch

A switch is a multiport bridge.

- Break up collision domain
  - Repeaters are inside the collision domain,shince they propagate collisions.
  - Bridges/Switches break up the domains, since they operate at layer 2 and buffer packets before sending them.
- Broadcast Domain
  A Network interconnectedd by bridges comprises a BROADCAST DOMAIN. Broadcasts form one host are seen by every other host aon the bridge network. If a NIC receives a frmae not addressed to it, the NIC ignores the frame. This decision is made without interrupting the CPU. BUT, broadcasts contain higher-layer information. Processor interrupt required.

- Managing Broadcast Domians:VLANs
  In a bridged network, broadcast and multicast traffic is sent everywhere. 100Mbps traffic could thus congest 10Mbps networks. It is therefore necessary to isolate broadcast domains. This may be done using multiple virtual local area networks VLANs within the switches or network.
- Switching implementation:crossbar

4.3 spanning tree protocol

In many scenarios ethernet switches are connected in network with redundancy. Broadcasts would be retransmitted forever. STP builds a spanning tree with a root. Broadcasts are never repeated.

5 Wi-Fi 802.11

Chanllenges of wireless networking:

- In wireless networks have a more limited range as the signal strength decreases more rapidly with distance (inverse square law), and is attenuated as it passes through different media.
- Wireless links have a higher Bit Error Rate due to noise, interference and multipath.

Relevant IEEE standard:

Table 2: IEEE Standards

Standard	Description
802.3	CSMA/CD(”Ethernet”)
802.5	Token Ring
802.11	Wireless LAN (”WiFi” family of standards)
802.15	Wireless personal area network (WPAN) - Bluetooth,Zigbee etc.
802.16	Fixed Broadband Wireless Access System (WiMax)

IEEE 802 standarisation framework

Table 3: IEEE 802

<table>
<tr><td colspan="6">802.2 Logical Link Control (LLC)</td></tr>
<tr><td rowspan="2">802.3 MAC PHY</td><td rowspan="2">802.5 MAC PHY</td><td colspan="4">802.11 Medium Access Control (MAC)(CSMA/CA)</td></tr>
<tr><td>802.11 PHY</td><td>802.11a PHY</td><td>802.11b PHY</td><td>802.11g PHY</td></tr>
</table>

IEEE 802.11 presented as the first true industry standard WLAN(released in 1997).

- The very first ”WiFi” standard.

- Provided data rates of 1Mbps or 2Mbps with range of 20 to 30m.

802.11 standard covers two aspects of the protocol stack:

- Physical transprot (PHY)
- Media access control(MAC)

Two network configurations are supported:

- ad-hoc - no structure and no fixed points (IBSS)
- infrastructure - fixed network access point, can bridge to fixed networks.(ESS)

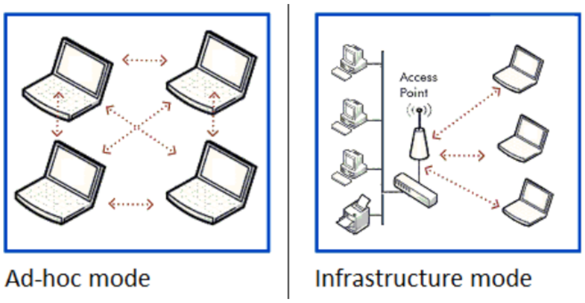


Figure 11: ad hoc model / infrastructure model

5.1 ad-hoc architecture 802.11

Direct communication within a limited range:

- Station(STA): terminal with access mechanisms to the wireless medium
- Independent Basic Service Set(IBSS): group of stations using the same radio frequency

5.2 infrastructure architecture 802.11

Direct communication within a limited range:

- Station(STA): terminal with access mechanisms to the wireless medium and radio contact to the access point
- Basic Service Set(BSS): group of stations using the same radio frequency
- Access Point(AP): station intergrated into the wireless LAN and the distribution system

- Portal: bridge to other (wired) networks
- Distribution System: interconnection network to form one logical network (ESS:Extended Service Set)based on several BSS

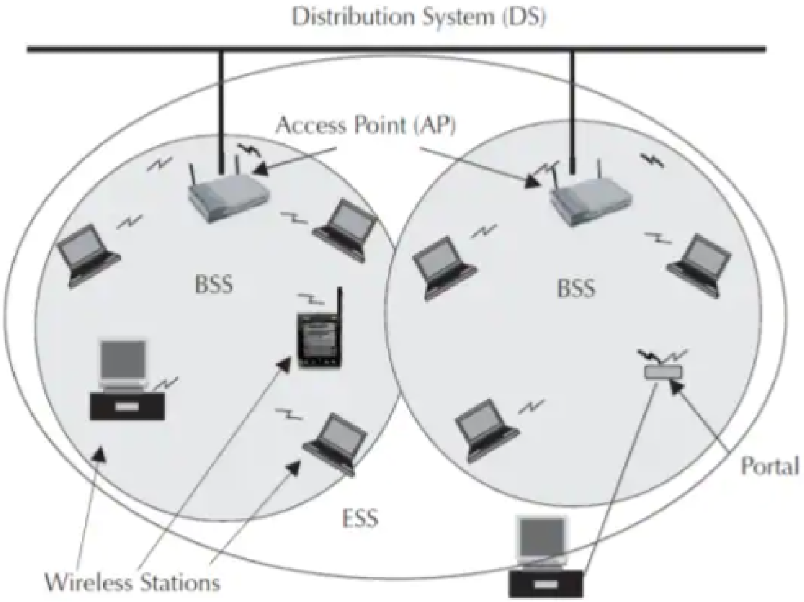


Figure 12: infrastructure architecture

WLAN frequency bands - ISM
WLAN systems make use of the ”industrial, Scientific, and Medical” (ISM) bands. These are unlicensed frequencies avaiable for free use in most countries, subject to power limitations.

5.3 802.11 frame format

- very similar to Ethernet.
- Duration of connections.
- serveral addresses, depending if we are using IBSS or BSS.

6 802.11 Protocol

6.1 802.11 Frame type

- management frames
  - beacon
  - probe Req/Res

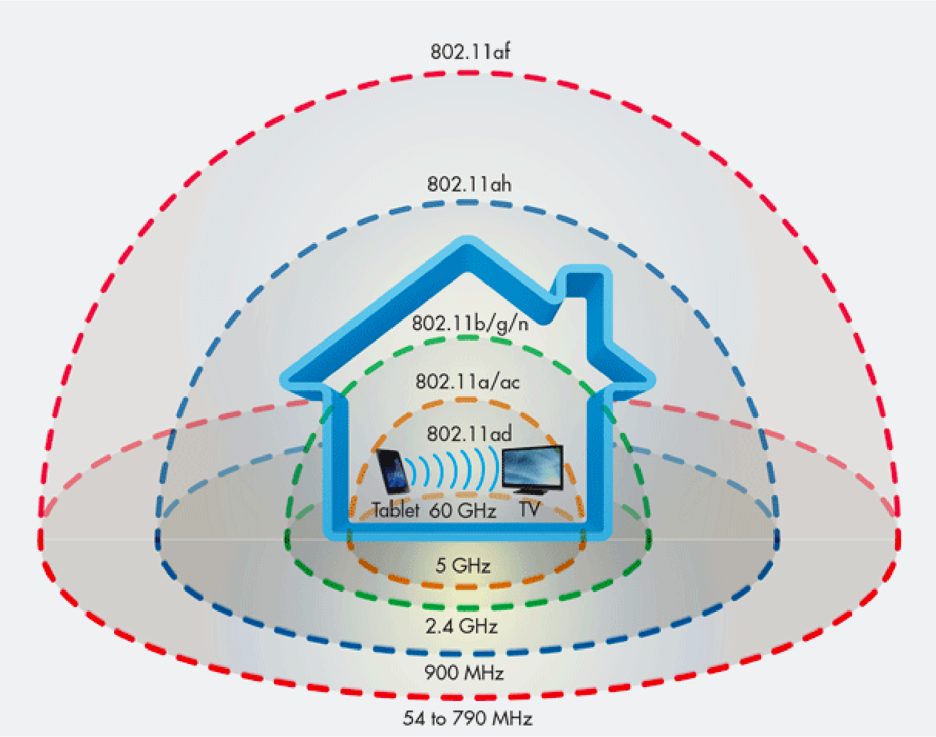
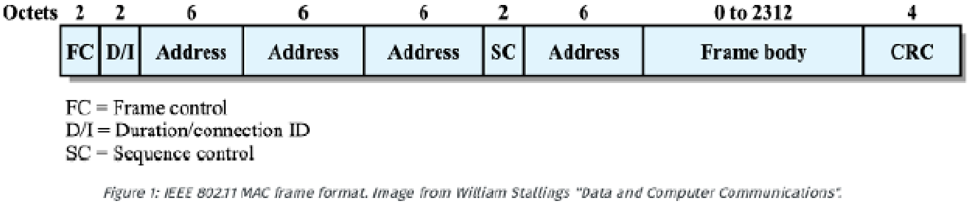


Figure 13: Range of 802.11 Standard



Octets	2	2	6	6	6	2	6	0 to 2312	4
	FC	D/I	Address	Address	Address	SC	Address	Frame body	CRC

FC = Frame control
D/I = Duration/connection ID
SC = Sequence control

Figure 1: IEEE 802.11 MAC frame format. Image from William Stallings "Data and Computer Communications".

Figure 14: 802.11 frame format

- association Req/Res
- reassociation Req/Res
- Authentication Frame
- Deauthentication and Disassociation
- Action Frames
- Channel Switch Announcement

- control frames
  - RTS(Request To Send)
  - CTS(Clear To Send)
  - ACK(Acknowledgement)
  - PS-Polling Some devices may want to go to sleep mode to save power. Node indicates to AP that it's going into power save mode. AP buffers the frames for the node. When node wakes up sends a PS-POLLING request to AP and gets all the frames.

Table 4: RTS

bytes	2	2	6	4
	Frame Control	Duration	Receiver Address	CRC

Table 5: CTS

bytes	2	2	6	6	4
	Frame Control	Duration	Receiver Address	Transmitter Address	CRC

- data frames

6.2 802.11 MAC

The CSMA/CA protocol also incudes an optional poihnt coordination function (PCF). Access point become a point coordinator (providing a contention free service). The point coordinator polls each client at a given time. No other station may transmit at that time. This provides a bounded delay service useful for voice, voice over IP (VolP) and other multimedia traffic. (However this option is very rarely implemented). The MAC layer also support authentication, network management and privacy.

6.3 PCF Point coordination function

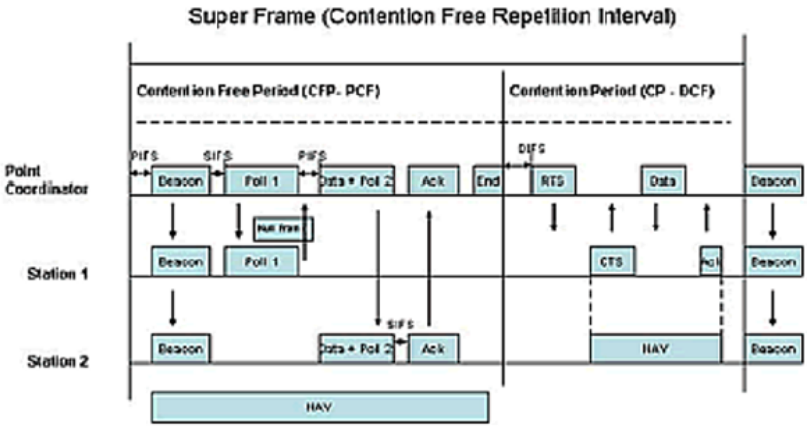


Figure 15: PCF

6.4 TWT Target wake up time (802.11ax)

New feature for IoT. AP tells devices to go to sleep and wake up at a specific time. This not only saces power in devices but reduces congestion.

Table 6: ACK

bytes	2	2	6	4
	Frame Control	Duration	Receiver Address	CRC

7 CSMA/CA

As in non-switched Ethernet a Multiple Access scheme is required to allow multiple users to all transmit within the allotted spectrum without intertering with each other.

Why not CSMA/CD?

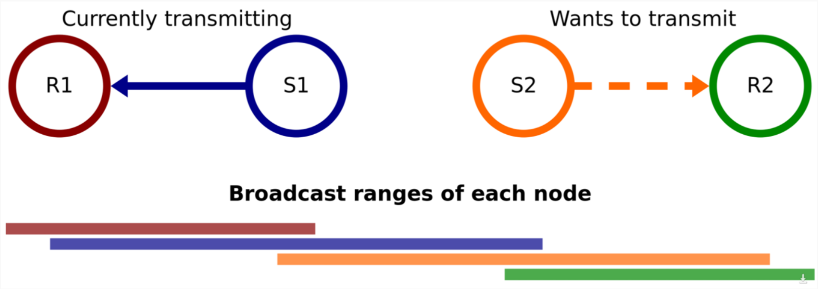
- lt is not possible to detect a collision: the power of a radio transmission decreases rapidly with distance,listening while transmitting only results in hearing yourself.
- On a wireless network it's not always possible for a station to hear all the other stations,so a sending station that is free to transmit has no way of knowing if the receiving station is free as well.This gives rise to **the Hidden Terminal Problem and the Exposed Terminal Problem**.

7.1 Hidden Terminal Problem

Note 2. In wireless networking, the hidden node problem or hidden terminal problem occurs when a node can communicate with a wireless access point (AP), but cannot directly communicate with other nodes that are communicating with that AP. This leads to difficulties in medium access control sublayer since multiple nodes can send data packets to the AP simultaneously, which creates interference at the AP resulting in no packet getting through.

7.2 Exposed Terminal Problem

Note 3. In wireless networks, the exposed node problem occurs when a node is prevented from sending packets to other nodes because of co-channel interference with a neighboring transmitter. Consider an example of four nodes labeled R1, S1, S2, and R2, where the two receivers (R1, R2) are out of range of each other, yet the two transmitters (S1, S2) in the middle are in range of each other. Here, if a transmission between S1 and R1 is taking place, node S2 is prevented from transmitting to R2 as it concludes after carrier sense that it will interfere with the transmission by its neighbor S1. However note that R2 could still receive the transmission of S2 without interference because it is out of range of S1.



7.3 CSMA/CA

7.3.1 Multiple Access

The first rule is only one people talk at a time, the other listen. Nobody interrupts or talks over someone else.
If you have something you wish to say, you first listen to ensure that nobody else is talking.If the channel is clear,then you can talk. This is the carrier sense part of CSMA.

7.3.2 Collision Avoidence(4 way handshake)

What happens if two dinner guests sense a lull the conversation and both start talking at the same time? This is a collision.

In 802.11 terminology,a collision occurs when two or more transmitters detect a quiet channel and both start transmitting at the same time. The collision will result in an undecipherable message to the intended receivers (listeners).

But in the wirless world one cannot detect these collisions. 802.11 handles collisions with a **4 way handshake**.

4 way handshake

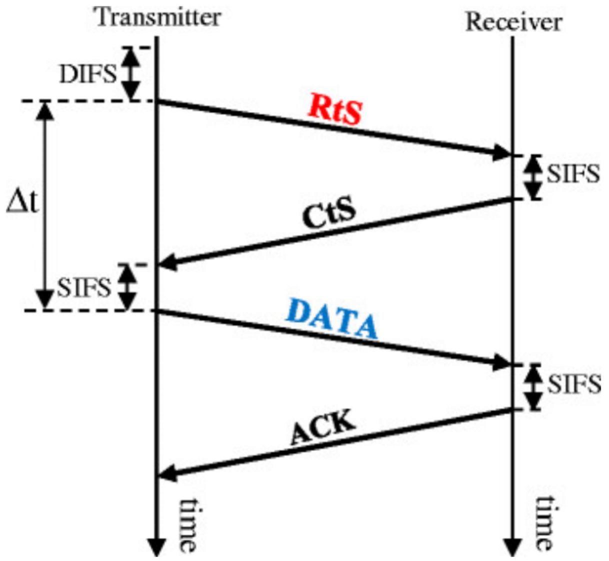


Figure 16: 4 way handshake flow

1. ”Listen before you talk”’: If the channel is busy, node backs-off for a random amount of time after waiting DIFS just as before.

2. But now,insead of packet sends a short message :Ready to Send(RTS) which lets the other nodes know that a message packet is coming.

3. RTS contains destination address and duration of message. The RTS tells everyone else to back-off for the duration.

4. If RTS reaches the destination successfully,the destination sends a Clear to Send (CTS) message after waiting a prescribed amount of time,called Short Inter Frame Space(SIFS).

5. After receiving the CTS,the original transmitter transmits the information packet.Other nodes in range of the receiver detect the CTS signal and refrain from transmitting for a time known as the Network Allocation Vector (NAV).

6. The receiver uses the CRC to determine if the packet has been received correctly. If so, the receiver sends out an ACK packet.

7. If the information packet is not ACKed, then the source start again and tried to retransmit the packet.

8 Other wireless technologies

8.1 Bluetooth

- Bluetooth was originally aimed at small form factor, low-cost,short-range radio links between mobile PCs, mobile phones and other portable devices.

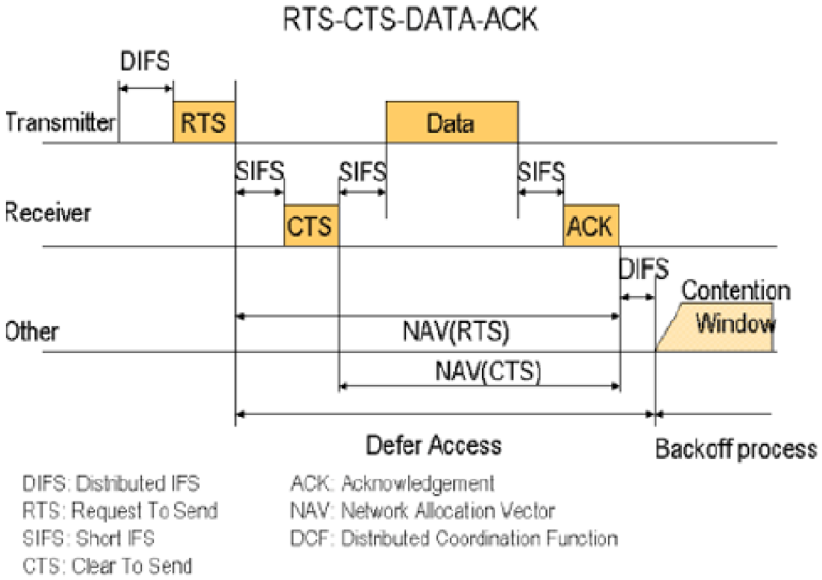


Figure 17: 4 way handshake protocol

- Often used for cordless computer pperipherals mouse,keyboard,trackpad etc.
- Very low cost hardware designed to be widely embedded in industrial and consumer equipment.

The basic idea:

- Universal radio interface for ad-hoc wireless connectivity (no infrastructure)
- Interconnecting computer and peripherals,handheld devices PDAs cell phones-replacement ofIrDA
- Embedded in other devices
- Short range (10 m),low power consumption, license-free 2.45 GHz ISM
- Voice and data transmission,approx.1 Mbit/s gross data rate

8.1.1 History

- 1994: Ericsson (Mattison/Haartsen), ”MC-link” project
- Renaming of the project Bluetooth according to Harald Blatand
- Gormsen [son of Gorm],King of Denmark in the 10^{th} century
- 1998: foundation of Bluetooth SIG, www.bluetooth.org
- 2001: first consumer products for mass market,spcversion 1.1 released

8.1.2 Link Types

- SCO(Synchronous connection Oriented)
  - FEC (forward error correction,no retransmission
  - point-to-point
  - 64 kbit/s duplex
  - circuit switched
  - Intended for voice transmission
- ACL(Asynchronous ConnectionLess)
  - Asynchronous fast acknowledge
  - point-to-multipoint
  - up to 433.9 kbit/s symmetric or 723.2/57.6 kbit/s asymmetric
  - packet switched
  - Intended for data transmission

8.2 Piconet

Collection of devices connected in an ad hoc fashion. One unit acts as master and the others as slaves for the life time of the piconet. **Master determines hopping pattern, slaves have to synchronize:This is the MAC layer** Each piconet has a unique hopping pattern. Each piconet hasand up to 7 simultaneous slaves (> 200 could be parked).

Forming a piconet
All devices in a piconet hop together. Master gives slaves its clock and device ID:

- Hopping pattern: determined by device ID(48 bit uique worldwide)
- Phase in hopping pattern determined by clock

Addressing

- Active Member Address (AMA,3 bit)
- Parked Member Address (PMA, 8 bit)

8.3 Scartternet

- Linking of multiple co-located piconets through the sharing of common master or slave devices - Devices can be slave in one piconet and master of another
- Communication between piconets - Devices jumping back and forth between the piconets

8.4 Bluetooth LE

Bluetooth Low-Energy(BLE) - or Bluetooth Smart,- is a significant protocol for IoT applications. It offers similar range to Bluetooth it has been designed to offer significantly reduced power consumption. Not backward-compatible with the previous"Classic" Bluetooth protocol, but the Bluetooth 4.0 specification permits devices to implement either or both of the LE and Classic systems.

Most used technology for wearable devices.

Standard: Bluetooth 4.2 core specification.

8.5 MANET

Adhoc networks: MANET Mobile Ad-hoc Networks.

- Vehicular ad hoc networks
- Military/Rescuenetworks
- UAV Ad hoc networks
- Wireless sensor networks

challenges:

- Decentralized routing algorithms
- Power

Week3 Review Notes

ELEC0099: Introduce to Internet Protocol Networks 21/22

RUFENG DING

Contents

1 IP backgrounds

A internetwork is :

- A collection of individual networks.
- Connected by intermediate networking devices.
- That functions as a single large network.

1.1 Internet Protocol suite

ISO/OSI and Internet Protocol suites

Table 1: Internet Protocol suites

7	Aplication	FTP/SMTP/Telnet/HTTP	Host-to-Host Protocols
6			
5			
4	Transport	TCP UDP	
3	Internet	IP (ICMP/ARP/RARP...)	
2	Data link	802.x/PPP/FR/ATM	Network specified ”outside” Internet specification
1	Physical	LANs/PSTN/LL/ADSL	

1.2 Two providers

A service describtion of Internet
The Internet allows Distributed Applications running on it end systems to exchange data with each other. It provides two service two its distributed applications:**a connection-oriented reliable service** and**a connectionless unreliable service**. It does not yet provide a service that makes promises about how long will it take to deliver the data from sender to receiver.

1.3 router

It is the main component of the Internet. It is responsibile to, step by step, forward packets to the destination. It can be simple and cheap but also very expensive.

Path determination:
route taken by packets from source to destination. Routing algorithms.

Forwarding:
move packets from router's input to appropriate router output.

Inside the router:
Two key router functions:

- run routing algorithms/protocol(RIP/OSPF/BGP).
- switching datagrams from incoming to outgoing link.

2 Addressing

2.1 IP internet Protocol RFC791

- Connectionless service
- Network addressing
- Best effort delivery
  - IP datagrams may arrive at destination host damaged duplicated of order, or not at all
  - no end-to-end delivery guarantees
- Handles data forwarding using routing tables prepared by other protocols such as:
  - Open shortest path first (OSPF)
  - Routing information protocol (RIP)
- Fragmentation and reassembly

2.2 IPv4 Datagram Structure

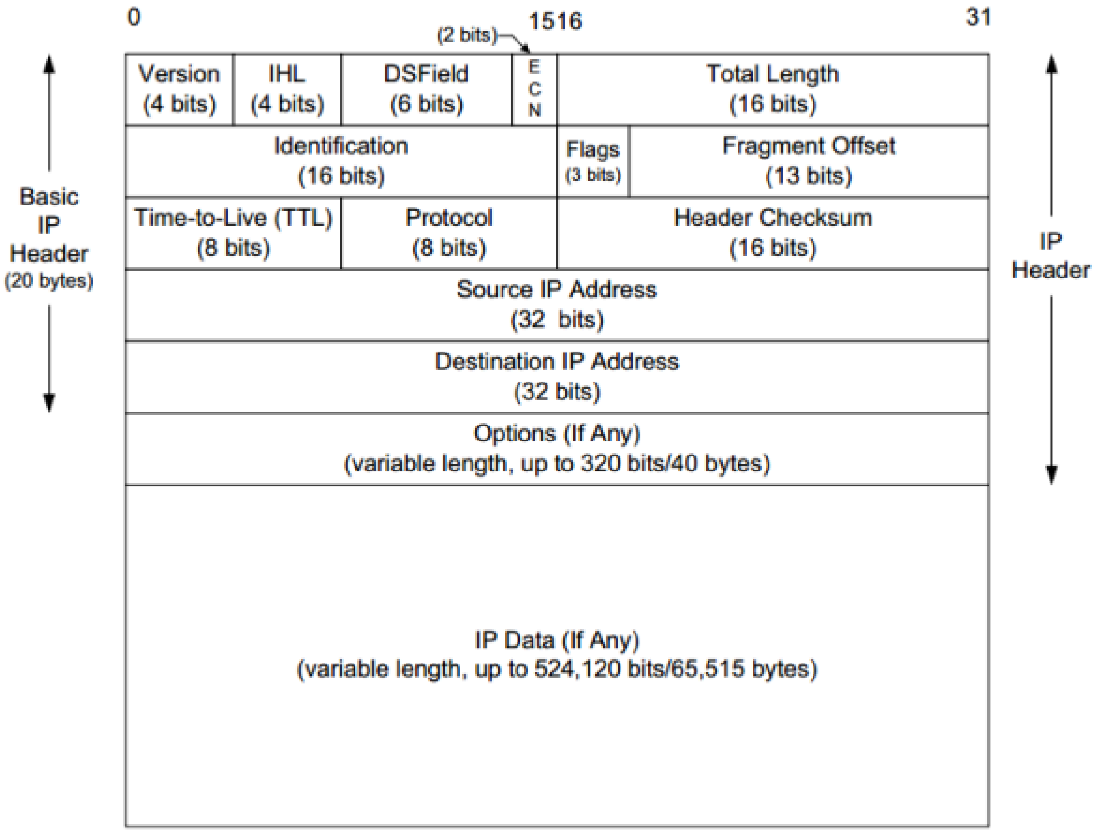

0
(2 bits)
1516
31
Version (4 bits)
IHL (4 bits)
DSField (6 bits)
E C N
Total Length (16 bits)
Identification (16 bits)
Flags (3 bits)
Fragment Offset (13 bits)
Time-to-Live (TTL) (8 bits)
Protocol (8 bits)
Header Checksum (16 bits)
Source IP Address (32 bits)
Destination IP Address (32 bits)
Options (If Any) (variable length, up to 320 bits/40 bytes)
IP Data (If Any) (variable length, up to 524,120 bits/65,515 bytes)
Basic IP Header (20 bytes)
IP Header

Figure 1: IPv4 Datagram Structure

2.3 IPv4 Addressing

Two main functions:

- uniquely identify a computer in a given internet.
- provide information so that routers deliver the packet to the correct destination.

They have 32 bits (4Bytes) and are representd by dot notation. eg.: *138.77.45.3*

10001010 01001101 00101101 00000011

addressdes will have two parts: The left part will identify a particular network/LAN. The right part will identify the host inside that network.

Important : The amount of bits for each part will vary.

2.4 Special IPv4 addresses

- All 0 host suffix - netwrok address 128.10.0.0 is a network with possible hosts like : 128.10.3.4, 128.10.0.3
- All 0s network - this network eg. 0.0.0.2 host 2 on this network
- all 1s host suffix - broadcast to all host on the same subnet
- public IP address are controlled by the internet
- private IP addresses RFC1918
  - any organization can use these inside their network. However thses addresses can't go on the internet
  - 10.0.0.0 - 10.255.255.255
  - 172.16.0.0 - 172.31.255.255
  - 192.168.0.0 - 192.168.255.255
- Loopback address : 127.0.0.1 All computers "have" this IP address

2.5 Classful addressing

host ID of all 0's indicate Network ID

host ID of all 1's indicate broadcast to Network ID

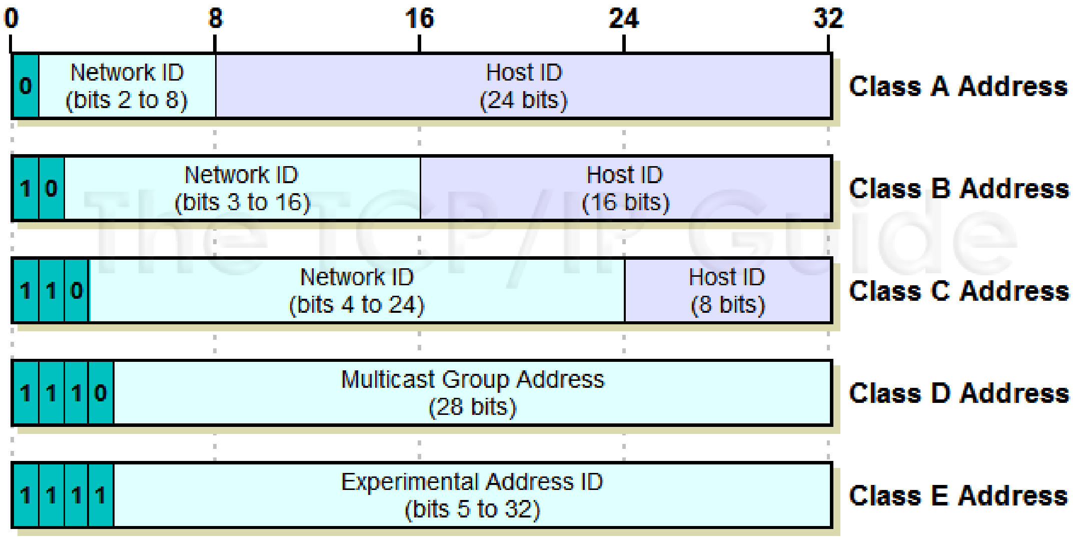


Figure 2: addresses classes

- 0.x.x.x -127.x.x.x : 128 big networks 16 million hosts.

- 128.x.x.x - 191.x.x.x : 16384 Medium Networks 65 thousand hosts.
- 192.x.x.x - 223.x.x.x : 2 million small networks 254 hosts.
- 224.x.x.x -239.x.x.x used for multicast group.
- 240.x.x.x - 255.x.x.x reserved for future use.

In this scheme each address is said to be self-identifying because the boundry between prefix and suffux can be computed from the address alone.

2.6 subnetting

sometiems organization may wish to partition their network, This can be done with subnetting.
Classless and subnet address extensions (CIDR)
class scheme is too rigid and many address can be wasted, subnettting permits to split class A,B,or C in smaller network. Host ID field is split in subnetwork field and host field. A subnet mask (Net ID + subnet field) identify all hosts that belong to a specific subnetwork address space.

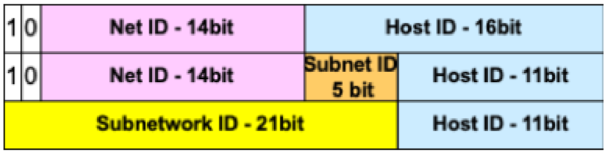


Figure 3: eg CIDR

In this example subnet mask is 255.255.248.0 or FF.FF.F8.00 Hex. This network permits to allocate up to 2046 Hosts.

IP address netmasks
Bit mask for the network part of the address: eg. 255.0.0.0/8. 255.255.240.0/20, etc.

eg: 128.16.20.1/16

/16 - 255.255.0.0

128.16.20.1 1000 0000 0001 0000 0001 0100 0000 0001

255.255.0.0 1111 1111 1111 1111 0000 0000 0000 0000

128.16.0.0 1000 0000 0001 0000 0000 0000 0000 0000

2.7 Problems with IPv4

- Sortage of IP addresses The 32bits address system in IPv4 can theoretically recongnize 4.3 billion hosts. This is not enough for widespread adoption of IP in multiple devices

- Insufficient security functions Scalability requires that robust security measures on the IP datagram are availble. In IPv4 security is typically a funciton of the upper layers
- Fragmentation introduces complexity
- No Quality of Service support
- Complex header

2.8 IPv6 Datagram Structure

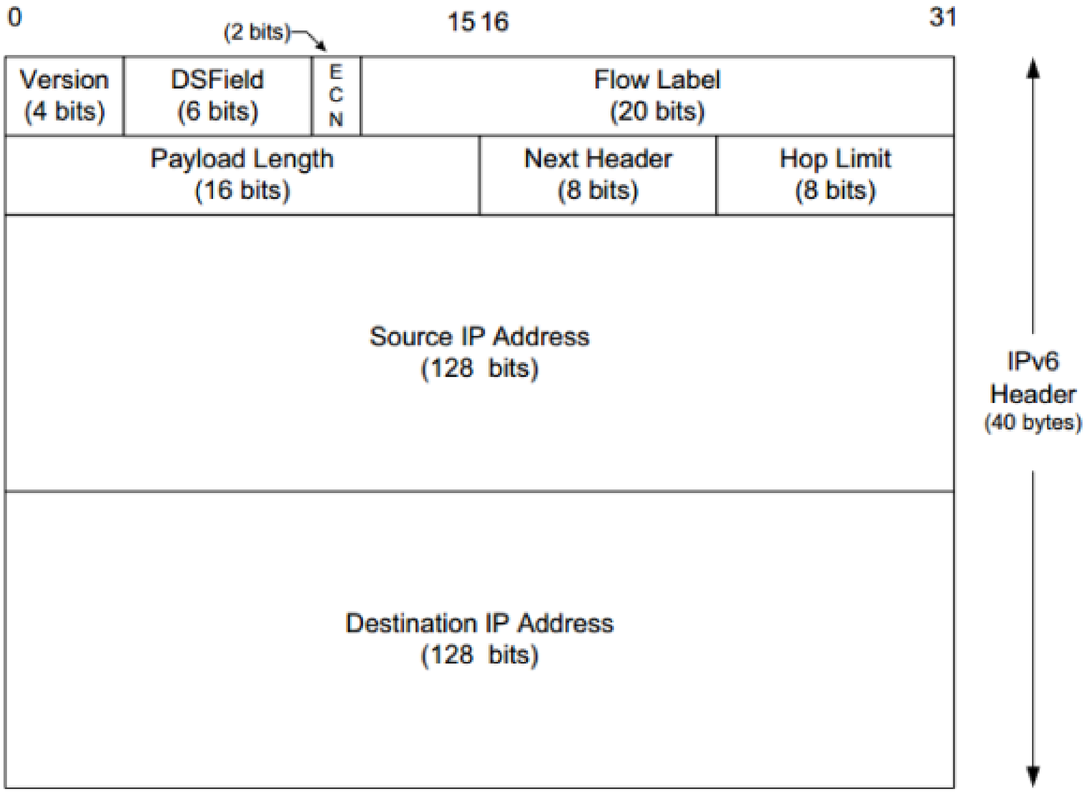


Figure 4: IPv4 Datagram Structure

2.9 IPv6 provides

- Expansion of IP address : An astronomical number is catered for.
- Hierarchical IP addresses
- IPsec function is installed as standard
- QOS control function
- No Fragmentation

- Automatical location of IP addresses
- Simplified header
- Allows Jumbograms (very big packets)

2.10 IPv6 address

2.10.1 notation

IPv6 uses 8 groups of 4 hexadecimal digits: 2001:0630:0013:0200:0000:0000:0000:ace0. This can be ”tided up” by removing leading 0’s and eliding runs of ”0”s: 2001:630:13:200::ace0 . The boundary between network and host part is indicated using /. eg. 2001:630:13:200::ace0/64 indicates a network address of 2001:630:13:200::and a host address of ::ace0.

2.10.2 address blocks

IPv6 addresses are allocated to interfaces. An interface can and will have many addresses. IPv6 address space has been split into blocks. RFC4291 describles IPv6 addressing.

There are :

special purpose blocks:

- 0000::/8 is reserved by IANA and includes
- The unspecified address (all 0’s) and the loopback address (::1) are assigned from this block
- IPv6 addresses mapped from IPv4 (E.g. ::128.40.42.82 is a valid IPv6 address,but cannot be globally routed)

multicast: Equivalent to addresses from the IPv4 block: 224.0.0.0/3 and IPv6 multicast range is FF00::/8

Note: There is no concept of broadcast in IPv6 : multicast must be used instead.

link-local unicast: Allocated from the block FE80::/10 and 64bit Interface ID appended to make an address. Interface ID constructed from interface MAC address and this address cannot be globally routed.

local IPv6: Designed to replace RFC1918 private IP addresses.

Block prefix: FC00::/7 8th bit indicates global or local management. Only a value of 1 (local management) has currently been standardised: Prefix becomes FDOO::/8 inpractice.

The remaining 56bits of the network address consist of a random 40bit ID and 16bit subnet number. The ID is randomly generated such that there is a good chance it is globally unique. Allows two organisations to merge Locally addressed networks without renumbering.

Global unicast: Equivalent to a standard IPv4 address such as 128.40.42.82 and allocated from the block: 2000::/3

Special case unicast address: *Anycast*: An anycast address may appear on several interfaces on different hosts,but the network layer only delivers packets to one of them.

A note on IPv6 addresses with embedded IPv4 addresses: Two forms of these were specified:

- IPv4 compatible lPv6 addresses: ::0000:128.40.42.82
  - Designed to allow automatic tunnelling of IPv6 over IPv4
  - This range has been deprecated and will no longer be used
- IPv4 mapped IPv6 addresses: ::FFFF:128.40.42.82
  - Designed to allow IPv6 hosts to exchange packets with IPv4 hosts

2.10.3 address allocation policy

The IPv6address assignment policies were designed to result in efficient routing tables. Assignments are hierarchical with the Regional Internet Registries getting large /12 allocations. The assignment policy recommended in RFC3177 is to allocate /48 prefixes to organisations and private individuals. Very large organisations may be assigned a /47 or a set of /48 prefixes. IPv6 was designed to a;low for 245 networks (/48 prefix, 35x1012). Compare IPv4 with 2.2x106 networks.

2.10.4 address in practice

recommended is to use 64/64 scheme:

- 64 bits for network part
- 64 bits for host part

Host part derived using EUI64. uses the 48 bit layer 2 MAC address (padded to 64bits). Means a subnet can have 224 hosts = 16M. Practivally, subnets should never run out of addresses.

3 Multicasting

3.1 Multicast in IPv4

Multicast is the transmission to a given set of members of a group.This group may contain members anywhere in the Internet. This process is quite complex.

Addresses of tthese groups must be in the range of 224.X.X.X to 239.X.X.X

Nodes need to subscribe to a multicast group (using multicast addresses). More on this later.

3.2 Multicast addresses in IPv6

Multicast is an integral part of IPv6 and is used for:

- Router discovery
- Address resolution
- Well-known service discovery

A multicast address is indicated by FF in first byte and the next byte is formed of 4 flag bits and 4 scope bits. The other 112 bits are the multicast group ID.

In practice,to make mapping to ethernet adresses simple,the group ID is usually 32bits.

3.3 Special Multicast addresses in IPv6

There are many permanently assigned multicast addresses:

- FF02::1 = All nodes on a link (LAN)
- FF02::2 = All routers on a link
- FF05::2 = All routers in a site
- FF05::3 = AII DHCP servers in a site

FF02::1:FFXX:XXXX is the solicited node multicast address. The last 24bits are copied from the last 24bits of the node's unicast address. This is likely to be site unique.

3.4 Mapping multicast addresses to Ethernet

The Multicast IP address has to be mapped to an Ethernet MAC address so that hosts can receive the datagrams.

In IPv4 the MAC prefix of 01:00:5e is used along with the last 24bits of the multicast IP address. (eg. 224.15.31.23 maps to 01:00:5e:0f:1f:17) ln IPv6 the MAC prefix of 33:33 is used along with the last 32bits of the IP address.

4 Resolution and Autoconfiguration

4.1 ARP address resolution protocol

Host broadcasts a request to ask for the MAC address of IP address. And the host whose IP is the one reply back the MAC address.

Sending a package:

Is the package for our network?

- Y (We need the Ethernet address) - Is the Ethernet address for this IP address in the cache?
  - Y Prepare an Ethernet Frame with the address - send the frame
  - N Send an ARP broadcast asking for the Ethernet address - Get the response update the Cache - go to Y
- N Looking in our routing table to check what is the router to send the packet - now the packet ”is” for our network. We are going to send the packet to the router.(Then go to Y)

4.2 ARP notes

ARP Cache timeout: typically 20 minutes.

Proxy ARP: sometimes router will reply ”instead of the host” to ”trick” the host into sending them the packets.

Gratuitous ARP: a ”reply” sent by a host without a reuest. Typically sent by a host waking up.

Remember: ARP is not just for Ehternet. It works for any layer 2 technology.

4.3 Replacement of ARP in IPv6

In IPv4 on a LAN, ARP is used to discover the link layer address so datagrams can be exchanged.

IPv6 uses a different method. **Remember there is no broadcast in IPv6**. A Neighbour Solicitation message is sent to the **neighbour solicitation multicast address**.

The node in the multicast group (remember, there should only be one) replies with a Neighbour Advertisement which contains the link-layer address.

The Address to link-layer mapping is stored in a cache in the host. Hosts can send periodic unsolicited neighbour advertisement message to the all nodes multicast address (FF02::1).

5 obtain own IP address: DHCP/SLAAC

5.1 DHCP

Sometimes we want an IP address allocated dynamically. For this we usually use DHCP.

- It allow dynamic configurations: automatically assigned address leasing.
- Uses LAN broadcast
- Require servers: central store of configuration information
- Useful for
  - mobile hosts
  - large numbers of hosts (can use static/manual address assignment)
- usually also returns:
  - default router
  - netmasks
  - DNS servers
- usually sent using UDP port 64 (destination 255.255.255.255 source 0.0.0.0)

5.2 SLAAC

IPv6 Host Autoconfiguration SLAAC = StateLess Address Autoconfiguration

How a host get its addresses:
as a host boots, it will bind eadch network interface to the following addresses:

- FF02::1 all nodes multicast address
- ::1 Loopback address

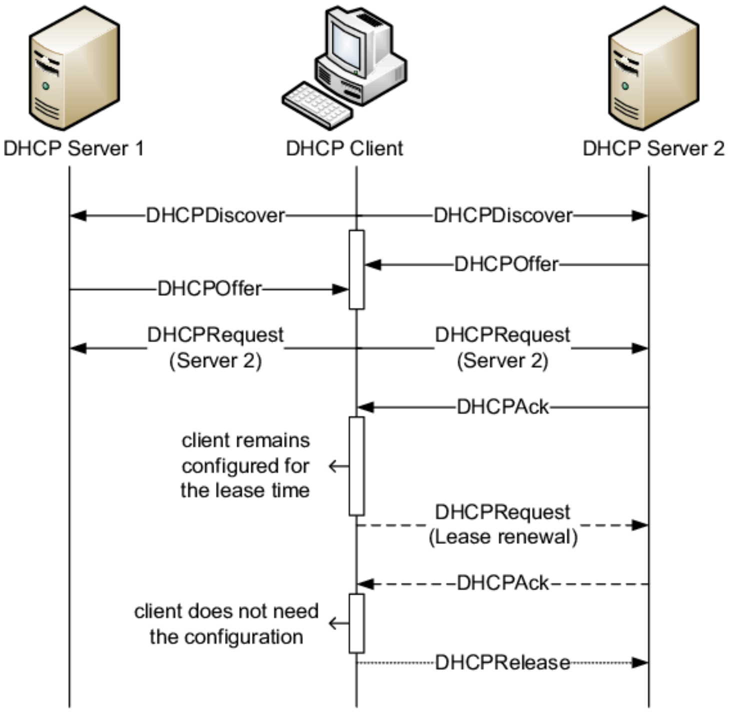


Figure 5: DHCP process

- FF02::1:ffxx:xxxx node solicitation multicast address(xx:xxxx = last 24 bits of MAC address(usually))
- FE90::interface ID (This is marked as a ”Tentative address” and cannot be used as the source address for any datagrams yet.)

5.3 Host autoconfiguration

At this point, a host has a usable IPv6 address. However, this address is link-local.(FE80 prefix).Which means it can not be used to talk to discover any router on the network.

The next stage is to discover any router on the network. A router solicitation message is snet to the all-routers multicast address FF02::2. All routers on the link reply with a router advertisement. Each advertisment that has the autoconfiguration flag set will cause the host to construct an address from the advertised router prefix and the host’s interface ID. If there is more than one router(and network prefix), an address for each network prefix will be assigned.

5.4 DAD

DAD = duplicate address detection: use a multicast mechanism

The host constructes a ICMP Neighbour Discovery packet which will be sent to the node solicitation multicast address corresponding to the last 24 bits of the interface ID.

This will discover if there is another node using the same address if there is, autoconfiguration will stop and the node will have to be assigned an address by another means. The other node will use a Neighbour Advertisement message to signal its presence.

Otherwise, the IPv6 address is unique and there usable.

5.5 Privacy Extension RFC4941

computer picks a series of bits randomly, and fills in the last 48 bits with the random bits. Reduces privacy concerns. A new IP address can be generated with varying frequency. In theory even one per connection. And it is widely available in operation systems.

6 fragmentation

6.1 Fragmentation in IPv4

in the beginning the Internet slogan was ”IP over everything”.

Because IP has to use several underlying technologies, IP packets may have to be fragmented. Fragmentation is done by routers. Defragmentation is done by end systems.

Related parts: Total length / Identificaiton / Flags /Fragment Offset.

6.2 Fragmentation in IPv6

IPv6 router do not do fragmentation. Instead senders implement a process called **Path MTU discovery**

Senders send a first packet. If a link in the middle cannot cope with that size, it drop the packet and sends back a ICMP message saying ”packet too big”. The process continues until the packet reached the destination.

7 ICMP

Internet Control Message Protocol (ICMP) ICMP is used by IP to send error and control messages.

It is sent by end hosts or routers.

Table 2: ICMP in IPv4

IPv4 Header	ICMP header	ICMP data

Table 3: ICMP in IPv6

IPv6 Header	IPv6 Extension Headers(if present)	ICMP header	ICMP data

7.1 ICMP messages

ICMP messages are carried in IP packets. But conceptually we see ICMP at the same level of IP.

7.2 ICMPv6

in IPv6 ICMP isresponsible not only for error and informational messages but also for IPv6 router and host configuration. Most messages achieve similar goals as IPv4.

Some messages:

0-127 errors.

128-255 informational.

7.3 ICMP tools

7.3.1 PING

used to see whether a specified IP address is reachable. Tool is available in Microsoft Windows operationg system and UNIX paltforms.

7.3.2 TraceRoute

Sned a packet with time to live = 1

the first router discards the packet and send a ICMP 'time to live exceeded message'

send a packet with time to live = 2

The second router discards the packet and send an ICMP 'time to live exceeded message'. This is repeated until a response is received from the destination.

Table 4: ICMP

Type 8 bits	Code 8 bits	Checksum 16 bits

Table 5: ICMP messages

type	ode	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest. host unreachable
3	2	dest. protocol unreachable
3	3	dest. port unreachable
3	6	dest. network unknown
3	7	dest. host unknown
4	0	source quench(congestion control not used)
8	0	echo request(ping)
9	0	route advertisement
10	0	routerr discovery
11	0	TTL expired
12	0	bad IP header

8 NAT

once a network is created, LANs can be point-to-point and routers have more than one LAN.

Network address tranlation NAT

sometimes the organization/ISP/home has many more computers than IP addresses. NAT is used to covert private IP address to public IP addresses and vice versa. Normally present as a router function. Used to connect an intranet using private IP addresses to the Internet. It may enhance security.

eg:

- inner host 10.0.0.1 sends datagram to 128.119.40.80
- NAT router changes datagram source address from 10.0.0.1 to 138.76.29.7,5001 updates table (NAT translation table)
- reply arrives destination address: 138.76.29.7,5001
- NAT router changes datagram destination address form 138.76.29.7 to 10.0.0.1,3345

8.1 NAT for UDP and ICMP

in UDP there is no connection termination. NAT routers need to use timers to check if they can reuse a port.

And it is similar for ICMP.

8.2 NAT traversal

originating connections TO a node behind a NAT is hard. There is no port to put in the TCP SYN packet.

There are many techniques: STUN, TURN, uPnP. Generally involve registering with a node with a public IP address which will then relay the connections.

8.3 Carrier Grade NAT

NAT can be done inside the network. Many networks behind the NAT box. Sometimes thousands of users.

Problems: port exhaustion. Some applicaitons use lots of TCP connections.

9 Mobility

9.1 Mobile IP

Nodes move to a different network. But they want to still receive the packets desined to their 'old' address. Sending applications do not need to be aware of the movement. Just one of the techniques for IP Mobility...

3 Step:

1. Discovery
2. Registration
3. tunnelling

9.2 Three Musketeers of Mobile IP

How does the Mobile Node find out where it is?

- Agent Discovery - ICMP Router Discovery

How does the Mobile Node Inform its current location?

- Registration - Authentication, location update and deregistration

How are packets delivered?

- Tunnelling - IP in IP or GRE

9.3 Mobile IP terminology

1. Agent Discovery To discovery a foreign agent a Mobile IP node uses ICMP Router Discovry (RFC 1256)

   Router periodically broadcasts or multicasts ICMP RD messages on all its links. Mobile IP advertisements contain defined extensions to ICMP RD.

2. Agent Advertisement routers propagate advertise packet in an ICMP packet.

3. registering

   - The mobile sends a registration request to the foreign agent
   - The foreign agent relays this request to the home agent
   - The foreign agent relays this reply to the mobile node

4. Tunnelling the home agent arrived at a foreign network and it discovered a router willing to act as a foreign agent and registered with its home agent.

9.4 IP over IP

senders sends the packet to the destination unaware of any tunnelling/mobility : original packets - encapsulated packet

9.5 The Triangle problem

Route not optimal. Server and Mobile node could communicate directly but...

Authenticaiton would have to be made with potentially every node in the internet.

9.6 Route Optimization

In IPv6 things are different when both nodes have IPv6 enabled.

The mobile node can send a IPv6 Destination options header message to the corresponding node. Packets can then be sent directly.

Week4 Review Notes

ELEC0099: Introduce to Internet Protocol Networks 21/22

RUFENG DING

Contents

1 What is routing?

The propagation of connectivity information in order build the routing tables. Routing tables are used for packet forwarding. Final hosts usually are configured to talk with one router.

Routing Tables

Routing:

- Routing protocols act before any data packets go on the network.
- They are not involved directly in data transmission.
- They are equivalent to the people who put traffic signs on roads.

Routing/Forwarding Plane

Routers are just computers with more than one network interface. They forwarding plane forwards every packet. It need to be fast. On the Routing Plane, one or more Routing daemons(normal programs) talks with daemons on other routers, to exchange routing information and updates the routing tables in the forwarding plane.

Building the Routing Tables

Routing tables can be built by hand but

- In some cases there are many entries
- They may have to change quickly when things change in the network

Today we need Routing Protocols to automatically build and update the Routing Tables.

1.1 The Big Pictures: Intra-domain vs inter-domain

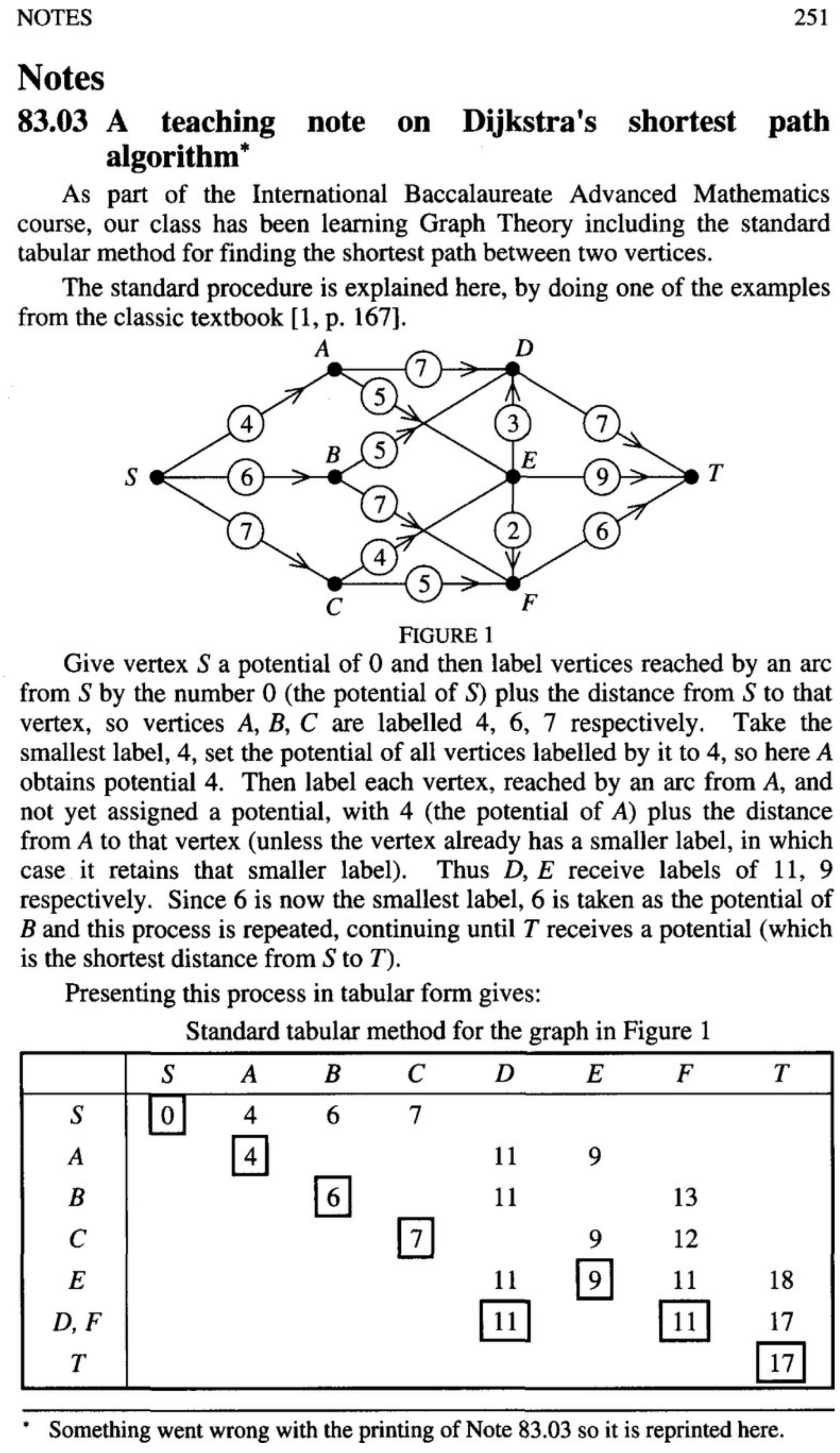
Inside each Autonomous System (AS) we run an **Intra-domain routing** protocol.

Between each **AS** we run a **Inter-domain Routing** protocol.

Graphs a useful mathematical abstraction several real-life problems can be approached by using graphs. These are used in lots of fields like road traffic optimization. Graph have: Nodes/Edges/Costs.

1.2 Dijkstra's Best Path algorithm

looks to the fig 1. The notes of Dijstra.



Notes

83.03 A teaching note on Dijkstra's shortest path algorithm*

As part of the International Baccalaureate Advanced Mathematics course, our class has been learning Graph Theory including the standard tabular method for finding the shortest path between two vertices.

The standard procedure is explained here, by doing one of the examples from the classic textbook [1, p. 167].

FIGURE 1

Give vertex *S* a potential of 0 and then label vertices reached by an arc from *S* by the number 0 (the potential of *S*) plus the distance from *S* to that vertex, so vertices *A*, *B*, *C* are labelled 4, 6, 7 respectively. Take the smallest label, 4, set the potential of all vertices labelled by it to 4, so here *A* obtains potential 4. Then label each vertex, reached by an arc from *A*, and not yet assigned a potential, with 4 (the potential of *A*) plus the distance from *A* to that vertex (unless the vertex already has a smaller label, in which case it retains that smaller label). Thus *D*, *E* receive labels of 11, 9 respectively. Since 6 is now the smallest label, 6 is taken as the potential of *B* and this process is repeated, continuing until *T* receives a potential (which is the shortest distance from *S* to *T*).

Presenting this process in tabular form gives:

Standard tabular method for the graph in Figure 1

	S	*A*	*B*	*C*	*D*	*E*	*F*	*T*
S	$\boxed{0}$	4	6	7				
A		$\boxed{4}$			11	9		
B			$\boxed{6}$		11		13	
C				$\boxed{7}$		9	12	
E					11	$\boxed{9}$	11	18
D, *F*					$\boxed{11}$		$\boxed{11}$	17
T								$\boxed{17}$

* Something went wrong with the printing of Note 83.03 so it is reprinted here.

Figure 1: Dijstra

2 RIP

Distance Vector vs Link-State

Bellman-Ford

Distribution of aggregate information

Distribution only to neighbours - Distance Vector Algorithms (eg. RIP Protocol)

Dijkstra

Distribution of information on local links

Distribution to every node - Link-state Algorithms (eg. OSPF Protocol)

RIP - Routing Information Protocol: Original routing protocol. Developted in 1969 Defined in RFC 2453. Gained popularity with inclusion in UNIX BSD in 1982. RIP itself is an Application level protocol running over UDP but it affects the network layer.

2.1 RIP Operation

a routing running RIP broadcast a routing message every 30 seconds to all it neighbours. Each update contains a pair <network,distance>. Wherer distance is usually the number of hops (in certain cases managers may increase the weight of a hop).

RIP participants updatae their routing tables based on these broadcasts.

RIP example
Importtant: routers never know the topology of the network. **Don't forget**: routers will also send message to same node which will also propagate it. Routers will choose the best metric to reach. **All routers** send equivalent messages.

Hysteresis - A RIP router does not change the route if the distance received is the same. If a router does not received an update after 180 seconds the router is deleted. The maximum distance for RIP is 16. If an administrator wnats to have bigger value it has to parrtition the network.

2.2 Split horizon

CURE: In RIP routers do not annouce routes on the links these were received - Split horizon.

2.3 RIP message

command: 1 = request 2 = reply.

- updates are replies whether asked for or not
- initializing node broadcasts request
- requests are replied to immediately

Version: 1 or 2.
Address family: 2 for IP.
IP address: non-zero network portion, zero host portion.

- identifies particular network

Metric:

- Path distance from this router to network
- Typically 1, so metric is hop count

3 OSPF

OSPF: open shortest path first. To encourage the use of link state routing protocols a working group of the IETF has designed OSPF. Defined in RFCs 1245, 1246, 1247, 2328(v2), 5340(v3 for IPv6).

3.1 OSPF operation

differences form RIP:

- lt is link-state based. Every router knows about every link in the network.
- incude type of service routing.
- authenticaiton of routing messages
- definition of areas
- allow load-balancing of traffic
- serveral minor improvements...

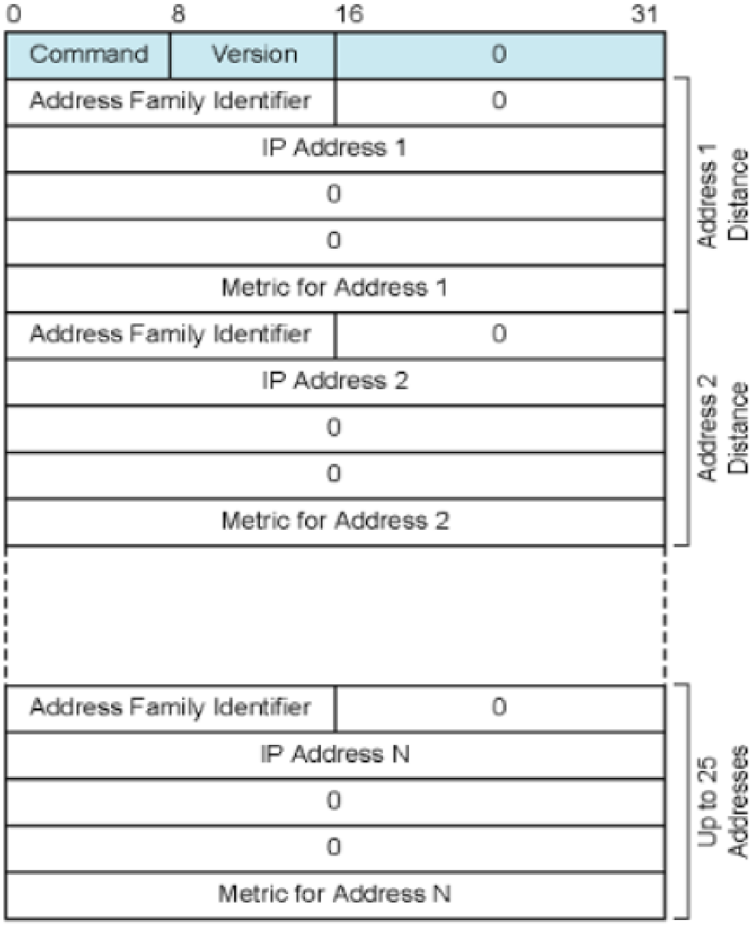


Figure 2: RIP message

3.2 OSPF Overview(OSPF = Flooding + Dijkstra)

Router maintains descriptions of state of local links as a directed graph:

1. 1 - Transmits updated state information to all routers it knows about using Flooding It sneds a message about each link to all the neighbours. These replicate that message to all the neighbours except the one that sent the message. If routers receive a message they already broadcasted they just drop it.

2. 2 - Router receiving update must acknowledge lots of traffic generated

3. 3 - after receeiving all the message form the network OSPF routers calculate what is the shortest path to reach all the destination.

4. 4 - with that information they calculate what is the link to the used for every network

5. 5 - Because all the routers have the same information and calculate the same paths using the same algorithm all the forwarding decision are consistent.

6. 6 - If for some reason information is not consistent packets may get looped. eg. link goes down and that information has not arrived to all the routers. Remember TTL!!!

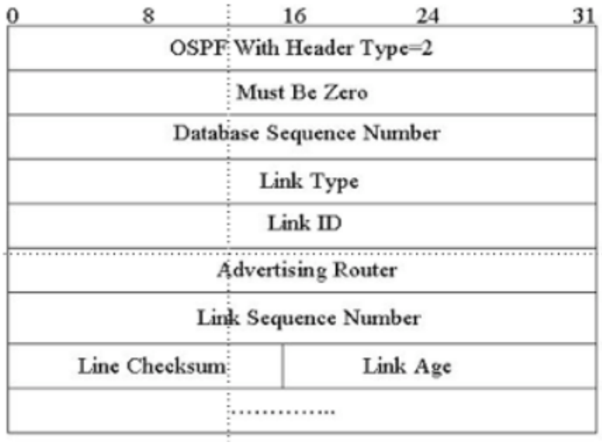


Figure 3: OSPF message

In reality OSPF announces links or networks. The type is indicated in the 'Link type' of the Database Description message

3.3 Link costs

cost of each hop in each direction is called routing metric. OSPF provides flexible metric scheme based on type of service (TOS)

- Normal (TOS) 0
- Minimize monetary cost (TOS 2)
- Maximize reliability (TOS 4)
- Maximize throughput (TOS 8)
- Minimize delay (TOS 16)

Each router generates 5 spanning tree (and 5 routing tables)

3.4 Area

mark large internets mroe managable. Configure as backbone and multiple areas.

- Area - colleciton of contiguous network and hosts plus routers connected to any included network.
- Backbone - contiguous collection of networks not contained in any area, their attached routers and routers belonging to multiple areas.

3.5 Operation of Areas

each area runs a separate copy of the link state algorithm.

- topological database and graph of just that area
- link state information broadcast to other routers in area
- reduced traffic
- intra-area routing relies solely on local link state information

3.6 OSPF Packet Header

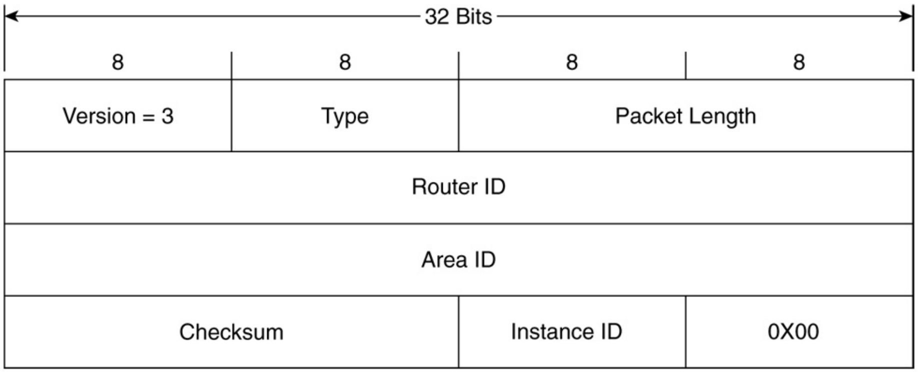


Figure 4: OSPF packet header

Packet format Notes:

- version number 2 is current
- type one of 5
- packet length: in octets including header
- router id: this packets source 32 bit
- area id: Area to which source router belongs
- Authentication type: null,simple password or encryption
- Authentication data: used by authenticaiton procedure

OSPF packet types

- hello: used in neighbour discovery

- databse description: defines set of link state information present in each router's database
- Link state request
- Link state update
- Link state acknowledgement
- Runs directly over IP

3.7 IS-IS

Has it roots in OSI. Similar to OSPF: also link state. also uses Dijstra. Runs directly over layer 2(paralled to IP).

Several minor differences:

- OSPF supports point to multipoint links
- OSPF supports virtual link
- an OSPF router can belong to multiple areas whereas an IS-IS router can belong to only one area.

3.8 OSPF weight calculation

There is no standard. $10^8/(interfacebandwidth)$.

More elaborate methods:

- Take into account traffic matrix
- spread the traffic over all links:minimizing the most used link

3.9 OSPF Configuration

OSPF works by itself...

But needs to be configured system administrators login into routers(these are just computers) and through Command line configure the parameters.

3.10 Summary

- Routing tables are built by applications running on the routers that communicate using routing protocols

- Graphs are mathematical abstraction used to calculate routing tables
- Each Autonomous System(AS)runs the same protocol across all its routers
- There are two types of routing protocols:link-state and distance-vector
- OSPF is the most used interior routing protocol
- OSPF uses a flooding algorithm to propagate link-state information and the Dijkstra algorithm to calculate the best path to every other node in the network
- Given the best path,calculating the routing table is straightforward

4 BGP

4.1 Routing continued

At this point you should know how IP packets traverse an Autonomous system to reach a destination inside taht AS.Here all the routers are under the same administration.

Now we will look at how are packets routed over several Autunomous systems?

- Many more networks
- Owned by different organizations

4.2 Autonomous System

The Internet is not controlled by any centeral authority. If an organization is big enough it can form an Autonomous System. These Autonomous Systems make bilateral agreements between themselves. There isn't any other form of organization!!!

An AS can be a big company, it can be a Research network and it can be an ISP. Each AS, runs one instance of an IGP (OSPF,RIP,etc).

4.3 BGP details

Border Gateway Protocol:

- connects Autonomous Systems
- Transmits reachability to networks (or prefixes)

allows routers (gateways) in different ASs to exchange routing information. Messages sent over TCP. Three funcitonal procedures:

- Neighbour acquisition
- Neighbour reachability
- Network reachability

4.4 BGP Messages

- Open Start neighbour relationship with another router.
- Update
  - Transmit information about single route.
  - List multiple routes to be withdrawn.
- Keepalive
  - Acknowledge open message.
  - Periodically confirm neighbour relationship.
- Notification Send when error condition detected.

4.5 BGP Message Formats

four types of mesages:

- OPEN - start a BGP peering
- KEEPALIVE - keeps alive the connection
- UPDATE - updates any information about network reachability using
  - withdraw routes
  - attributes
  - network layer reachability (network / prefixes)
- NOTIFICATION - notifies of any protocol errors. Must close TCP connection after it is sent.

Most used attributes

- ORIGIN - Tells how a route was learnt (1 - IGP/2 - iBGP/3 - Other)
- ASPATH - List of AS numbers that packets will traverse for this announcement.

- NEXT-HOP - IP address of the router that packets need to be sent.
- MUTL-EXIT DESCRIMINATOR(MED) - when a AS has several options to a given destination, it may include information for preferring one of them.
- LOCAL-PREF - allow a specific AS to signal internal preference for a route to a destnation (when there is nore than one)

4.6 BGP and CIDR

CIDR allows BGP to transmit much less information due to Aggregation of routes.

4.7 BGP = iBGP + eBGP

when we talk about BGP we usually mean eBGP.

iBGP it is the part of iBGP concerned with transmitting external connectivity information inside the Autonomous System. IT IS NOT an Interior Routing Protocol. Routers will not learn how to reach a router inside their network through iBGP.

4.8 BGP Route Reflection

some AS have a big number of external connections.

And If every router transmits every message revceived to every router in the AS there will be lots of traffic generated. iBGP uses sometimes **Route Reflectors** These reveice the messages and send summaries to all the routers. (For Ases with more than 100 nodes, it is recommended to use a router reflector instead of a full mesh)

4.9 Route selection

every router in an AS needs now to decide which route to take to all prefixes based on following rules(by order):

1. looks at LOCAL-PREF
2. the route with less ASes in the AS list
3. looks at the MED (Multi-Exit Discriminator)
4. IGP distance to the NEXT-HOP
5. if all routes were learned through iBGP, choose the neighbour with the lowest BGP identifier

4.10 Transit vs non-Transit

when a ISP is providing connectivity to its clients, it is also providing connectivity to clients of other ISPs!!!

Imaging that traffic form Customer 1 was going to customer 3 via ISP2. Neither C1 or C3 are paying ISP2 .. ISP2 would be carrying traffic that was not paid. Two options: Transit Service and Non-transit service.

4.11 BGP Polocies

In theory, BGP allows each domain to define its own routing policy. In practice, there are two common policies:

- customer-provider peering: customer C buys Internet connectivity from provider P
- shared-cost peering: Domains x and y agree to exchange packets by using a direct link or through an interconnection point

4.11.1 Customer-provider peering

Customer sends to its provider its internal routes and the routes learned form its own customers - Provider will advertise those routes to the entire Internet to allow anyone to reach the Customer.

Provider sends to its customers all know routes - Customer will be able to reach anyone on the Internet

4.11.2 Shared-cost peering

Peer X sneds to Peer Y its internal routes and the routes learned from its own customers.

- Peer Y will use shared link to reach Peer X and Peer Y's customer
- Peer X's providers are not reachable via the shared link

4.12 Traffic cost models

fixed cost : one fixed payment per month
linear cost: payment proportinal to traffic
95th percentile:

- Bandwidth is measured (or sampled) from the switch or router and recorded in a log file. In most casses, this is done every 5 minutes.
- At the end of the month, the sample are sorted from highest to lowest, and the top 5% (which equal to approximately 36 hours of a 30-day billing cycle) of data is thrown away.
- The next highest measurement becomes the billable use for the entire month.

4.13 Routing Policies

Routing policies implement business relationships between domains.

The routing policy of a domain is implemented via the route filtering mechanism on BGP routers:

- Inbound filtering: Upon reception of a route from a peer, a BGP router decide whether the route is acceptable, and if so whether to change some of its attribtes.
- Outbound filtering: Before sending its best route towards a destination, a BGP router decides which peers should receive this route and whether to change some of its attributes before sending it.

4.14 Route Instabilities

several types of Instabilities affect the behaviour of BGP:

- Instabilities of the IGP
- Hardware Failures
- Software problems
- Insufficient Horsepower
- Insufficient memory
- Network Upgrades
- Human Error
- Backup Link Overloads

4.15 Route Flap Dampening

imagine that the link AS1-AS2 goes up and down frequently. This generates disturbances in the entire Internet. This could be aggravated if this is done intentionally. In this case AS2 does Route Flap Dampening. That is it stops advertising routes to AS1 to the rest of the Internet.

4.16 Redundancy with Multi-homing

Multi-homing refers to a single network having more then one connection to the Internet. This can be done mainly for two reasons:

- Load-balancing
- Fault-tolerance

Two types:

- Single provider
- Multiple provider

Multi-homing with Multiple Providers more interesting and useful is when a customer has two providers. Multi-homing brings some problems:

- Destroys route symmetry
- May cause packet reordering
- Makes firewalls more difficult
- impact on the BGP tables: CIDR become difficult

4.17 Load Balacing

Inbound Load Balancing can be achieved by announcing some of our routes in one Link and some in the other(s).

An easy way of achieving Outbound Load Balancing is to use a Route Reflector. The Route Reflector can be configured to announce some routes to some of our routers and other routes to the others.

4.18 Ineraciton with the IGP

The interaction of the IGP with BGP needs to be careful thought by the network administrator. You can have the default that all you networks are annouced on all the BGP peering or you can decide which areas are announced to. You can even decide not to annouance some networks (eg. networks used on the administration of the AS) In certain circumstances networks obtained from BGP can be inserted into the IGP: Risky!!!

4.19 Anycast

anycast is another message paradigm: A packet is sent to the closest member of a group.

Today in the IPv4 internet this is achieved by announcing the same address/prefix from several locations in the Internet. BGP will redirect packets from different sources to the closed destinations. eg. DNS root servers.

4.20 confederations

another way of dealing with big ASes is to break them into confederations. A confederation is an AS broken into sub-ASes. These can run different IGPs uesd in company acquisitions.

4.21 Forwarding summary: longest Prefix Match

since with prefix/CIDR there is not a fixed distinction between ”network” and ”host” routes select the next hop using longest Prefix Match.

Imaging that the routing table has two entries:
138.39.0.0/16 - Link 1
138.39.16.0/24 - Link 2

if a packet arrives with destination address 138.39.16.32 the router uses link2 because 24 bits match against 16.

Week5 Review Notes

ELEC0099: Introduce to Internet Protocol Networks 21/22
RUFENG DING

Contents

1 Multicast

eg. a TV station want to send **n** packets to **n** receivers.

what is multicast? The transmission of information to more than one destinations.

1. computer asks its router to join a group
2. router joins the multicast distribution tree

1.1 IGMP Internet Group Management Protocol

Multicast STEP-1

1. a computer want to join a multicast group using IGMP.
2. it sends an IGMP join message to its router.
3. **membership is dynamic** routers will periodically **poll** host in the network to verify if they are still in the group.
4. IGMP we put it as a integral part of IP.

1.2 Tree formation

Multicast STEP-2

1. it is the difficult part.
2. using Multicast Routing Protocol.

3. after join in they will reach the router and hosts in the network.

4. routers in a group's multicast tree have the group address in their Routing Tables.

DVMRP-Distance Vector Multicast Routing Protocol

1. all routers receive traffic for all groups.

2. routers not in group send **Prune** message up to the tree.

3. routers no-member-neighbours remove multicast entry.

4. repeated periodically.

MOSPF Multicast Extensions to OSPF

1. DVMRP generates lots of useless traffic(Prune).

2. OSPF have the information How to get to each node.

3. MOSPF use the information to multicast.

4. **Big Disadvantage** each router has to know all the group memberships.

PIM-SM - Protocol Independent Multicast - Sparse Mode

1. router do not know in advance what to do to join a group.

2. each group has a Rendez-Vons Point in advance.

3. router sends a unicast to RP and RP replies and the routers on this path are automatically part of the distribution tree.

4. Hosts wishing to multicast just send to the RP.

1.3 Muticast Extra Notes

1. Multicast address allocation needs independently. eg. MADCAP(RFC 2730).

2. original model : anybody can send message to a group. Reality: only specified sources can send.(security and scale)

3. IP multicast maps directly into Ethernet multicast. *If there is more than one member in a LAN, potentially only one packet is sent.*

1.4 Problems

1. each groups requires **state** (routing table entries) in every router in the path.
2. routers in the core of the Internet will require informations for every groups.
3. *difficult to be used for time shifted content.*
4. difficult to be achieve reliability and congestion control.
5. ISPs selling bandwidth. Multicast saves bandwidth.
6. Multicast today is only available in limited intra-domain deployments.

2 SDN software Defined Networking

2.1 define

- hardware only do packet forwarding.
- all control is done in software in a centralized manner.
- maybe far away in a cloud.

2.2 OpenFlow Protocol

Secure Channel connected to a controller via SSL.

Flow Table connected to PCs.

1. supports layer2/3/4 protocols
2. rule can be flexibly defined by combination of different matching fields
3. no rule match the packet is dropped or escalated.

Packet Escalation packet don't match any rules will be sent back to the controller. And, the controller will then calculates rule and send it back to router. Rules can be reset. This work well with TCP harder with UDP.

2.3 SDN Interface

1. Northbound Interface To applications.
2. Southbound Interface OpenFlow.
3. Eastbound Interface To Controllers in other domains.

2.4 Trends

1. P4
2. Network Function Virtualization

3 Queuing

- arrival rate and service rate.
- Most common Arrival Distribution : Possion
- Most common Servicing Distribution : Exponential

3.1 Kendal's notation ASCKND

- A : Arrival Process
- S : Service Time Distribution
- C : Numbers of Servers
- K : Numbers of places on the queue
- N : Population size
- D : Queue Discipline (FIFO)

Because Possion and Exponential is both Markovian we use M in the notation.
Typically we only use the first 3 parts.We assume the queue size is infinite.The Queue Discipline is FIFO. eg. M/M/1
in practice in computer networking : we use MM1 or more often MD1.
D stand for Deterministic because it is often a fixed rate.

3.2 Queuing assumptions

1. FIFO
2. no bulking or reneging
3. arrivals are independent
4. service times are independent
5. arrival and service rates remain stable

3.3 Queuing performance for M/M/1

average number of packets in the queue

$$Lq = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

average time packet spends on the queue

$$Wq = \frac{\lambda}{\mu(\mu - \lambda)}$$

4 Active Queue Management

4.1 Multiple Queues

- each output port has many queues: 1perFlow 1perClass
- How to choose next ?
- several approaches: Fair Queuing / Weighted Fair Queuing / Priority Queuing

4.2 Fair Queuing (FQ)

simple scheme : serving the output queue for each flow in a round-robin way.(skip empty flow)

1. Greedy flows are implicitly penalised.
2. **Problem** flows with large packets are favoured in terms of throughput in comparison to flows with small packets

Bit Round Fair Queuing (BRFQ) using a bit-by-bit round-robin discipline not whole packets.

1. no preferential treatment
2. good for classless best-effort traffic only

4.3 Weighted Fair Queuing (WFQ)

add weight to a round-robin scheme

- each connection gets a queue

- each connection reserves some bandwidth
- every packet is assigned a virtual finishing time

$$F_i^k = F_i^{k-1} + \frac{L_i^k}{\phi_i}$$

F is Virtual finishing time
L is Time to transmit the packet
Φ is the Bandwidth allocated

4.4 Priority Queuing (PQ)

Queues are assigned priorities from high to low.
Good for important traffic but can lead low priority queues to **starvation**. And, they maybe dropped finally.

Priority Queuing with Rate Limits solve the starvation problem.
A priority is served until a cap bandwidth rate is reached its ”rate limit”

4.5 Proactive Packet Dscard

- as the queue occupancy grows, packets are dropped before the queue gets full
- TCP senders perceive packet loss as a congestion signal and back-off
- as such, TCP-based flows will back-off to helping alleviate congestion before transit router buffer get actually full
- the network do not oscillate between congestion and under-utilisation, with all TCP sources backing off almost simultaneously.

4.6 Random Early Detection (RED)

- RED used as a proactive packet discard mechanisms
- queue is split into 3 parts though two thresholds:0,TH_{min},TH_{max},1. TH is the abbreviation for threshold
- queue occupancy < part 1 added to the queue.
- queue occupancy in part 2 packet is dropped based on a calculated probability P_a which depends on various parameters and increases with queue utilisaion.
- queue occupancy > part3 the packet always dropped.

5 Quality of Service

1. Flow based Queuing : Integrated Services Architecture(RFC 1633) Uses RSVP
2. Class based Queuing : Differentiated Services Architecture RFC2475

5.1 IETF integrated Services Architecture

- providing QoS guarantees in IP networks for individual application sessions
- Resource reservation: routers maintain state info of allocated resources
- admit/deny new call setup requests

Call Admission arriving session must:

- Declare its QoS requirement:R-spec
- Characterize traffic it will send into network:T-spec

The protocol is RSVP : carry R-spec and T-spec to routers

RSVP

1. sender sends a PATH message along the way it checks if every router has resources
2. Receiver sends RESV message allocating resources in the router

attention:

1. every router needs to check every message
2. RSVP uses soft state: Messages need to be sent periodically otherwise reservation is removed

problems with IntServ/RSVP

- each connection need to store reservations in every routers on the PATH.
- very big cost in core routers.
- scalability has made intServ unsuccessful.

5.2 Diffserv Architecture

define two types of routers : edge router and core router.
edge router

- per-flow traffic management
- marks packets in classes
- as in-profile and out-profile

core router

- per-class traffic management
- buffering and scheduling
- based on marking at edge

Classification in IPv4

- packet is marked in the Type of Service (TOS) in IPv4 and Traffic Class in IPv6.
- 6 bits used for Differentiated Service Code Point (DSCP) and determine PHB that the packet will receive.
- 2 bits are currently unused.

PHB is **per-Hop Behaviors**

- Expedited Forwarding(EF) PHB
- Assured Forwarding(AF) PHB (more soft guarantees AF4 better AF3 better AF2 better AF1)
- Default(Best-Effort) with no QoS guarantees

Multidomain DiffServ To work across more than one domain ISPs have to agree on SLA(Service Level Agreements) for every type of traffic.
It is not clear how anybody can guarantee QoS to end-to-end applications.

Bandwidth Brokers would be a good solution to end-to-end

- each domain runs a Bandwidth Broker

- app contact BB in their domain
- requires standardization
- still an open research issue

5.3 The Over-Provisioning Alternative

- Make sure adequate resources are always available in the network given the expected traffic demand.
- it is possible in core network but harder in access network and in peering points between ISPs.

problems remains:

- no strict guarantees given the statistical nature of traffic.
- Moore's law : as capacity increases, demand will also increase to consume it.
- arguably not economical, especially for tier-2/tier-3 ISPs

In the other hand if we had admission control which accept flows in 99.999% of the cases. Do we need this?

Week6 Review Notes

ELEC0099: Introduce to Internet Protocol Networks 21/22
RUFENG DING

Contents

1 TCP

Two types of transport over IP:

1. TCP
   - Reliable Transport
   - Controlled rate of transmission
   - complex
2. UDP
   - Unreliable Transport
   - uncontrolled rate of transmission
   - very simple

TCP applications using TCP use usually the **client/server** paradigm
Clients connect to servers, ask for information, get the information and close the connection.

1.1 TCP functions

1. **Demultiplexing**. several flows get into the same destination and we have to deliver right packets to right process.
2. **Reliability**. some flows require no packet loss. TCP should manage retransmission.
3. **Congestion Control**. how to manage traffic go without congestion.

1.2 TCP complexity

1. in Linux TCP consists of 11443 lines of code while UDP is 1522 lines.
2. TCP is a protocol but people will also refer to the algorithm that congestion control.

1.3 TCP-3 way handshake

1. STEP1: client host send TCP SYN segment to server
   - specifies initial seq #
   - no data
2. STEP2: server host receives SYN, replies with SYNACK segment
   - server allocates buffers
   - specified server initial seq #
3. STEP3: client receives SYNACK, time replies with ACK segment, which may contain data

1.4 TCP packet

table:

Source Port								Destination Port	
Sequence Number									
Acknowledgment Number									
Data offset	Reserved	URG	ACK	PSH	RST	SYN	FIN	windows	windows
checksum								Urgent Pointer	
Options									Padding
Data									

1.5 TCP connection establishment

- during connection establishment the initial sequence number for both directions is chosen
- options like MSS(maximum segment size) may be negotiated
- REMEMBER: TCP connections are full-duplex with different counters and parameters in both directions
- The same packet can send data and acknowledge on the opposite direction

1.6 TCP multiplexing

1. each process ”listen” and sends to and from a port. These two ports are included in the TCP packet
2. destination ports are standardized, source ports are chosen dynamically by the operating system.

NAT: network address translation : Ports
NAT touter need to use layer form :TCP to know the ports.

1.7 TCP reliability

- each TCP packet has to be acknowledged: this is an ACK packet
- if data is bidirectional these ACK packets can contain the data in the reverse direction
- computer does not need to receive ACK of the n^{th} packet to send the $n + 1^{th}$ packet. Each time there are x number of packets unacknowledged. This is the **congestion window**

1.8 Nagle Algorithm

1. a computer only send data when user write more than MSS : avoid packet with one byte

2. this can be switched off
3. the PUSH flag is used for a sender to signal a receiver that the data should be given to the application straight away

1.9 Congestion Control

TCP sliding Window the congestion window ”contains” the packets in flight, those not acknowledged.

- every time an acknowledgment arrives the windows slides to the right.
- The size of the window will change with time.

roughly:

$$rate = \frac{Cwnd}{RTT} bytes/s \tag{1}$$

where $Cwnd$ is the size of the congestion Window and RTT is the Round Trip Time.

Slow Start .

Every TCP connection start with a congestion window of 1 and the sender then applied the **Slow Start** algorithm: every time an ACK arrives the Congestion window is increased by 1 until a value of **SSH thresh** is reached.

Congestion Avoidance

- in the second part of TCP algorithm, the window is increased by 1/cwnd each time a ACK is received
- every time a packet is lost, one reduces the cwnd by half. sshthresh is also reduced by half of the congestion windows at the time of loss

How do we detect loss?

- Nobody tells the end system that a packet was loss!!!
- The end system got two ways
  - ACK *Timesout* , that is a specific amount of time passes without havin received the correspondent ACK
  - We recive *3 dupacks* - 3 packets acknowledging packets sent after the one which loss we are detecting

Timeout Calculaion to estimate if a packet is lost we need to calculate RTO (Retransmission Timeout Value) First one has to estimate the RTT

$$R = aR + (1 - a)M \tag{2}$$

where M is the measured RTT, a is recommended to be 0.9

$$RTO = R \times b \tag{3}$$

where b is recommended to be 2
some implementations complicate this a bit by taking into account the mean deviation
because when we receive a ACK of a retransmitted packet we don't know for sure which packet is being acknowledged. These measurements are never taken into account for RTO calculation.(Karn's algorithm)

TCP with aprropriate byte counting RFC 3465 proposed that TCP counts bytes instead of packets : this prevents artificially increasing congestion windows with very small packets(only require changes in the server)

Acknowledgments in practice in reality TCP acknowledgement sends the next byte to be expected without any "holes"

2 Advanced TCP

2.1 Sharing congestion state

- normal operation each connection keeps its own information(state)
- if the server got several connections to the same client this information can be shared.
  - Temporal sharing : sharing with CLOSED connections
  - Ensemble sharing : sharing with ACTIVE connections

2.2 TCP options

TCP can be extended by defining options: define an option number, length and specific data.eg:

- MSS
- Window Scale Option
- Timestamp option
- PAWS-Protection against wrapped Segment numbers
- TCP extension for Transactions

2.3 TCP URGENT flag

app using TCP can signal to the other side that some data is urgent and should be delivered as soon as possible
This use two fields on the TCP packet

- URG bit is sent
- URGENT POINTER is set to the offset where the urgent data begins. This pointer should be added to the sequence number to determine where is the first byte of the urgent data

2.4 TCP connection Close

to close a connection 4 packets are needed. Both side should send a FIN and receive a ACK. This way data is delivered to the app on both side.

1. FIN(A)
2. ACK of FIN(B)
3. FIN(B)
4. ACK of FIN(A)

2.5 TCP Selective Acknowledgements Option

- TCP option allows receivers to transmit more details about packets not received.
- Receiver informs the data sender of non-contiguous blocks of data that have been received and queued.

2.6 delayed acknowledgements

often a receiver delays sending an ACK to save bandwidth:

- receives more data which is then acknowledged
- has do send some data back and piggybacks the ACK in the data packet

200ms in Windows by default

2.7 TCP Keepalive

- TCP connection can be alive for months without any data being exchanged
- TCP provides a *Keepalive* option but its use is controversial(keep telling other side: I am alve)

2.8 TCP Cubic

optimized for high bandwidth, high delay networks
two operating regions:

- first is a concave portion where the window quickly ramps up to the window size before the last congestion event.
- Next is the convex growth where CUBIC probes for more bandwidth, slowly at first then very rapidly

$Cwnd$ function is:

$$W(t) = C(t - K)^3 + Wmax \tag{4}$$

where C is a constant (default 0.4), t is elapsed seconds and K is the time period the function takes to increase W to $Wmax$.

2.9 TCP BBR

old idea in the Internet (see TCP Vegas) is to detect congestion before loss and to react less drastically

2.10 Buffer Bloat

Ironically, sometimes more memory in routers can cause worse performance. This happens particularly in home routers. 1-Application delay may increase. 2-More memory implies a bigger delay in congestion signal.

2.11 Explicit Congestion Notification(ECN)

Routers can tell the end-systems that the network is getting congested and therefore they should reduce their Congestion windows. This technique is very efficient but demands that the *routers play the game*. Uses two bits[DSCP][CU] of Type of service field in the IP packet.

3 UDP

User Datagram Protocol

- Provides multiplexing/demultiplexing through UDP ports.
- no reliability (no ACKs)
- no congestion control
- app just send data to IP layer at the rare they want/can:
  no connection establishment and teardown $->$ connectionless service.

3.1 UDP header

$| < - - - - - -32bits - - - - - - > |$

Source Port	Destination Port
Length	UDP checksum
data	

input to checksum:

- UDP header
- UDP data
- IP source and destination address
- in IPv4 is optional(all zeros)
- compulsory in IPv6

UDP reliability some applications do not require 100% reliability. If you want to use UDP you have to manage reliability yourself.

UDP Congestion control UDP is the best solution for now in : some app do not cope well with the saw-tooth behavior of TCP (not tolerant to changes in the transfer rate) but the widespread use of UDP could take the Internet to congestion collapse.

3.2 Who uses UDP?

mainly two types of applications:

1. applications with very small data transfers(eg. DNS) where establishing the connection would be a big overhead.

2. Real time applications where congestion control would be damaging:
   - voice over OP
   - video-conferencing
   - Telemetry

   Remember: Multicast has to use UDP.9

4 Applications

Operating System Kernel(From low level to high level)

- Ethernet, Wi-Fi, FDDI, SDH
- IP, ICMP, IGMP
- TCP, UDP

all upper : socket interface.
Processes or Threads running in user space:
Email, WWW, p2p, Voice, etc.

4.1 Elastic application

Elastic:

- Interactive eg. Telnet, X-windows
- Interactive bulk eg.FTP, HTTP
- Asynchronous eg.Email, Voice-mail

.
examples:

Email

- asynchronous
- message is not real-time
- delivery in several minutes is acceptable

File transfer

- interactive service
- require ”quick” transfer
- ”slow” transfer acceptable

Network file service

- interactive service
- similar to file transfer
- fast response required
- (usually over LAN)

WWW

- interactive
- file access mechanism(!)
- fast response required
- QoS sensitive content on WWW pages

4.2 Inelastic(real-time)applications

Inelastic(real-time)

- Tolerant
  - Adaptive
    - ∗ Delay adaptive
    - ∗ rate adaptive
    - ∗ ———————upper is newer real-time app——————-
    - ∗ ——————lower is traditional real-time app—————-
  - Non-adptive
- In-tolerant
  - Rate Adaptive
  - Non-adaptive

example:
.

Streaming voice and video

- not interactive
- end-to-end delay not important
- end-to-end jitter not important
- data rate and loss very important

Real-time voice and video

- person-to-person
- interactive(Virtual and Augmented Reality)
- Important to control:
  - end-to-end delay
  - end-to-end jitter
  - end-to-end loss
  - end-to-end data rate

4.3 Email

oldest internet application:example of an Elastic Asynchronous Application
different protocols to send and to receive due to user not always connected to server.

PC −− SMTP − > server −− SMTP − > server −− POP − > SMTP

- Originally designed as a simple ASCII text replacement for the office memo.
- now can transport any kind of file using Multipurpose Internet Mail Extensions (MIME)
  - allows files of different types to be sent as ’Attachments’
  - these files are encoded so that they are not corrupted in transit
- has led to an explosion in the bandwidth required by email
- has also led to an explosion in storage requirements.

4.3.1 SMTP

Simple Mail Transfer Protocol.

- only guarantee to transfer 7bits ASCII characters (hence the need to encode other file types)
- Protocol allows for a sender to identify themselves, specify a recipient and transfer the message
- SMTP ensures that the email is not deleted from the sending system until the receiver has saved the message to non-volatile storage.

Protocol Features

- consists of simple commands: MAIL,VRFY, EXPN, RSET, DATA...
- Every command received by the server side must be replied to with a result code.
  - 2XX action taken, Result OK
  - 3XX action pending
  - 4XX Non fatal error, transaction can be tried again
  - 5XX Fatal error, transaction should be aborted

4.3.2 Typical Mail setup

- unusual to allow desktop system to directly send email to any host on the Internet.
- Generally organizations will setup mail gateway which ensure:
  - headers are correct for the organization
  - virus screening and possible other security measures are taken
- same applies to receive mail: there are more than one servers to receive incoming mail and hold it for users.

4.3.3 Protocols for Accessing Mail

- client systems generally pull received mail form server
- use POP(Post Office Protocol) or IMAP
  - both will allow user to read the email received on the server
  - POP expects user to download email to Mail User Agent(MNA), so mail is stored and organized in client systems.
  - IMAP retains the email on the server, only download the head of email. Messages are downloaded as they are read.

  - POP system are vulnerable to loss of the client system and do not allow a coherent view of email from multiple clients.
  - IMAP requires an active net connection while reading messages.
  - some clients are now smart enough to synchronize email between client and server so that you get the best of both worlds.

4.4 WWW

30 years old and one of the most successful applications in the Internet.
We need to distnguish 3 **independent** entities:

- HTTP - the protocol to transfer data
- HTML - the language to describe the data
- URL's - The method of naming and identifying specific data(pages, images, etc.)

4.4.1 HTTP

the protocol allow HyperText Markup language pages to delivered over IP network.

- Client use browser to access the web.
- Services is provided by a Web server.
- Well know port number 80.
- Uses TCP: Potentially multiple connections per page.

Other characters:

- **Length encoding and Hearders.** These are in the beginning of a reply content.
- **Negotiation.** A client can negotiate which media it can accept
- **Conditional Requests.** A client can request a page **only** if it has been modified
- **Max forward.** a server can minimize the number of proxies on the path

Persistent HTTP .

Non-persistent HTTP issues:

- requires 2 RTTs per object

- OS must work and allocate host resources for each TCP connection
- but browsers often open parallel TCP connections to fetch referenced objects.

Persistent HTTP:

- server leaves connection open after sending response
- subsequent HTTP messages between same client/server are sent over connection

Persitent without pipelining:

- client issues new request only when previous response has been received.
- one RTT for each referenced object.

Persitent with pipelining:

- default in HTTP/1.1
- client sends requests as soon as it encounters a referenced object
- as little as one RTT for all the referenced objects
- requests contain more than one object
- Objects are sent one after the other

HTTP Other interactions

- POST method
- Chunk response

Maintaining State in HTTP for some services the web server must maintain information about the user's session and it is done with **Cookies**

1. server include in the reply with set-cookie
2. browser sees and stores the cookie for each site
3. when the browser accesses the same site again it just send the cookie
4. THIS FEATURE can be switched off
5. cookies are vital for : e-commerce

4.4.2 Current Trend:HTTP 2.0

- Reduce Delay in Web pages
- data compressions of HTTP headers
- Server Push technologies
- Fixing the head-of-line blocking problem in HTTP1.1
- Making persistent pipelining work in practice
- do not use text based request
- good adaption by several browsers

HTTP2 multiplexing:

- order of request and reply can be different
- allows for prioritization

HTTP2 Server Push:

- if server thinks client need objects, it will send them straight away
- reduce latency

HTTP2 Header compression:

headers in HTTP2 provide information about the request of response. Header take around 800 bytes of bandwidth and sometimes few KB if it carries cookies. Therefore compressing headers can reduce the bandwidth latency.
Header compression is not like request/respond body gzip rather it is like **not sending the same headers again.**

4.4.3 HTML

- A Markup Language provides hinds to the client on how to display the page.
- final decision about how the page looks like are taken by the browser
- Images can be embedded in text.
- Links can be inserted to allow the download of any sort of file type.
- Similar to MIME, in that the browser can invoke the required application to display the file.
- Cause of the bandwidth explosion on the Internet
- New version: HTML5

4.4.4 URLs

Uniform Resource Locator

- originally just web addresses, but is generalized to include any network service.
- eg. http://www.ee.ucl.ac.uk:8080/-tom/index.html
  - Protocol: HTTP (Could also be ftp, ldap)
  - //www.ee.uck.ac.uk : this is the host address
  - :8080 : when the port isn't the normal 80
  - /-tom/index.html : the filename to be retrieved

4.4.5 Web Proxing and Caching

sometimes administrators do not want browsers to access directly to the Internet

1. contact a proxy server
2. proxy server gets the documents from the server
3. file is returned to the proxy
4. the proxy returns the file to the browser

optionally, the server caches the page, so that other users access it faster.

4.4.6 Middleboxes

TCP packets form residential or mobile user often get modified by middleboxes inside ISPs. Several fields may get changed, including sequence numbers. Sometimes, proactive ACKing is used. The box acknowledges packets that did not arrive to speed up the congestion control mechanism

4.4.7 Peer2Peer File Sharing

- represents a paradigm shift on the Internet: the application is simultaneously **client and server**
- user's program serves files to other "clients" and gets files from other "servers". These "clients" and "servers" are bundled in the same applications.
- The p2p programs form an **Overlay network** where each node is always connected to a small set of 'neighbours'

some problems:

- Querying for the desired file. It could centralized in one server or by sending queries to other peer2peer nodes which will transmit it to the necessary one.
- Bootstrapping: how to find neighbours
- Finding good ”neighbours”
- choosing good neighbours to optimize QoS
- Creating incentives for cooperation

5 Multimedia

5.1 RTP

real-time protocol. It specifies a packet stucture for packets carrying audio and video data.

- RTP packets provide:
  - Payload type identification
  - packet sequence
  - numbering
  - timestamping
- RTP runs in the end system
- RTP packets are encapsulated in UDP segments
- Interoperability : if two runs RTP and they can work together.

RTP runs on the top of UDP RTP library provide a transport layer of interface that extend UDP:

- payload type identification
- packet sequence numbering
- time-stamping

RTP Header

- **Payload Type 7 bits** indicates type of encoding current used. If sender changes encoding in middle of conference, it will inform the receiver through this Payload Type field.
- **Sequence Number 16 bits** increments by one for each RTP packet sent, and may be used to detect packet loss and to restore packet sequence

RTCP Real-time control Protocol

- work in conjunction with RTP
- each participant in RTP session periodically send RTCP control packet to all other participants
- each RTCP packets contains sender and / or receiver reports: report statistics useful to application(number of packet sent/number of packets loss/inter-arrival jitter/etc.)
- feedback can be used to control performance : senders may modify its transmission based on the feedback.

5.2 RTSP

Real-Time Streaming Protocol
Defined in RFC 2326 and it is a n out-of-band protocol running on port 544. It can run over TCP or UDP.

- several app consist of several streams which need to be coordinated
- services like playback, fast-forwarding, pausing are very useful.

RTSP Operation .
Web browser $-----$ HTTP GET $->$ Web server
Web browser $<-$ presentations etc. $--$ Web server

media player(client)(show the process between client and server)
$SETUP \rightarrow$
$\leftarrow$
$PLAY \rightarrow$
$\leftarrow$

$\leftarrow mediastream$

$PAUSE \rightarrow$
$\leftarrow$

$TEARDOWN \rightarrow$
$\leftarrow$
media server(server)

TCP video streaming (youtube etc)

- use TCP
- they all do buffering
- 3 strategies:
  - short ON/OFF periods at the application layer
  - Long ON/OFF periods
  - No ON/OFF periods. app downloads everything in one go(need storage in the device and produce useless transmission)

5.3 DASH

dynamic adaptive streaming over HTTP

1. server send to client a manifest : containing all available encodings and speed in XML / URL for each video audio subtitles etc.
2. Client chooses depending on local conditions

5.4 CDNs

content distribution networks

- most content is replicated in CDNs
- Servers all around the world
- two big advantages:
  - save bandwidth
  - better user experience
- two main techniques : DNS / HTTP redirect

Week7 Review Notes

ELEC0099: Introduce to Internet Protocol Networks 21/22

RUFENG DING

Contents

1 DNS

Domain Name System

- solution to the directory problem for hosts.
- Characters:
  - Distributed
  - Local control and update
  - Replicable
  - Consistent
- translate host names to IP addresses
- support IPv6 addresses

The DNS specification defines a Name Space:

- name space is **hierarchical** (Tree like)
- extensible
- branches called **Domains** leaves are **Hosts**

Domains can be groups into Zones of control:

- administrator of zones has full control to everything
- administrator of upper branches can delegate control of lower ones and create a separate zone.

1.1 DNS implementation

Each zone has a Primary server

- it is the authority of the information about the zone
- it makes updates
- should have one or more secondary servers(for performance and availability)
- invalid the data after a time-to-live period if no update

1.2 DNS lookups

name resolution $- >$ query

- client in a library called a *resolver*
- configuration files contains address of local DNS server and domain of the client
- supplied by DHCP for desktop machines

The resolution process client $--$ ask $- >$ dns name server $--$ refers to $- >$ root server $--$ according to the name ask name server $- >$ [name server (continue asking according to the name hierachy)]

Caching DNS cache the address so no need to ask again
(Caching) Time to Live: time a record be cache
trade-off

- too small TTL the system will be queried a lot
- too big TTL the system cannot change the addresses efficiently
- one week TTLs are common

1.2.1 2 method of performing a query

- iterative mode
- recursive mode (the Root server do the ask part and let to target name server directly send the address to you)

1.2.2 answers

- from domain server (authoritative answer)
- from cache at local server (Non-authoritative answer)

1.3 Main types of DNS records

- A IPv4
- AAAA IPv6
- CNAME redirects a domain to another domains
- PTR maps a IP address to a domain name
- NS returns the name servers for a given domain

2 SNMP

simple network management protocol

2.1 History

ICMP - SGMP - SNMP (CMOT long term) - SNMPv2 - SNMPv3

2.2 information to manage

- static
- dynamic
- statistical

functional architecture:

- manager-agent model
- a model for summarisation

2.3 Monitoring applications

these components include the functions that are visible to the users/manager
main tasks:

- performance monitoring
- fault monitoring
- security monitoring
- accounting monitoring

2.3.1 Manager function

retrieving information from other elements of the configuration. Usually this consists of library code linked with the monitoring applications.

2.3.2 Agent function

gather and records management information for network elements and communicates info to them. Agent function usually runs in elements and listen/reply to info request.

2.3.3 Manage objects

the info represent resources and their activities.eg: queue sizes/numbers of packets with errors...

2.3.4 Monitoring Agent

- generates summaries and statistical analysis of management information
- usually separate from the manager and reports to it periodically
- very useful in a heterogeneous environment

2.4 Monitoring Methods

decide when to use polling or traps is often a crucial for network efficiency

2.4.1 Polling

management station periodically ask for the info

Polling frequency because not many TRAPs are defined most problems will be detected by sequential polling.

$$N \leq \frac{T}{\Delta} \tag{1}$$

N number of agents
T desired polling interval
Δ average time for a poll

Δ depends of many factors

- processing time for the request at the management station
- network delay
- processing time for the request at the agent
- numbers of request/responses

2.4.2 Traps

when somethings abnormal happened, a device send a trap to management station

2.5 MIB (and SNMP)

MIB: database each objects represent a resource

- structured in the form of Tree
- each system maintains one for its managed resources

2.5.1 Extensibility

- allows new equipment with new features to be added without significant changes to the infrastructures
- allows private MIBs

2.5.2 SMI

structure of management information

- define the general framework in which a MIB can be defined
- identified the data types that can be used in the MIB
- defined in RFC 1155

SMI defined types:

- networkaddress
- ipaddress
- counter
- gauge
- timeticks
- opaques

2.6 MIB-II

MIB-II is the part of MIB that deals with the management of internet based protocols. It contains:

- system
- interfaces
- at
- ip
- icmp
- tcp
- udp
- egp
- dot3
- snmp

2.7 Private MIBs

companies can register private MIBs for experimental testing and research.companies can register with IANA.

2.8 Limitations of MIB Objects

lack of granularity in the information limit the identification of problems.SNMP is designed to minimise the impact in the network.

2.9 SNMP Operations

- GET : *GetRequest* message to get a value of a managed object
- SET : *SetRequest* message to set a value in a object
- TRAP : agent send *Trap* message to inform the manager of events

GetNextRequest: (bacause the MIB info is structured in a tree, we can ask for the next one.) will allow us to traverse the MIB tree without knowing its structure

Generic-trap field

- coldStart(0)
- warmStart(1)
- linkDown(2)
- linkUp(3)
- authenticationFailure(4)
- egpNeighbourLoss(5)
- enterpriseSpecific(6)

2.10 SNMP interactins

- get values: manager $--$ get request PDU $->$ Agent $--$ get response PDU $->$ manager
- get next values: manager $--$ get next request PDU $->$ Agent $--$ get response PDU $->$ manager
- set values: manager $--$ set request PDU $->$ Agent $--$ get response PDU $->$ manager
- send trap:Agent $--$ trap PDU $->$ manager

2.11 SNMP over UDP

SNMP is designed to be transported by any protocols. In practice UDP is used.

- uses ports 161 for gets/sets and 162 for traps
- TRAPs should be used as soon a problems arises and before becoming critical
- SNMP traffic should be prioritized

2.12 More on TRAPS

- traps are not acknowledged by the monitoring station. May leed to problem
- Threshold values configured with a **Set** command
- a trap is usually packed with information in the form of MIB object and its values

2.13 SNMPv2

- No security mechanisms
- new structure of management information (SMI)
- manager-to-manager capabilities
- new protocol operations

2.14 SNMPv3

SNMPv2 + administration and security

security issues address

- modification of information
- masquerade
- message stream modification
- disclosure
- NOT denial of Service
- NOT traffic analysis

3 Network Security

- **confidentiality** man-in-the-middle attack: intercept the packets or messages between two. eg: using a packet sniffer
- **Authentication** fake email address/DNS poisoning/Phishing
- **Non Repudiation** sends message to someone and claims not send it/similar problem to authentication
- **Unauthorized Access** pretends to be one and access her/his resources: password capture/inadequate firewall
- **Denial-of Service** sends a huge amount of traffic disabling ones server. /Distributed:Botnets. /Source Spoofing. /SYN floods. /Email SPAM.

3.1 Policies

a security policy is first step.

- What to be protected?
- Who is it to be protected from?
- Is it affordable?

3.2 Physical Security

protect he quipment.(Theft for eg.)
Insurance.

3.3 Virus and Worms

- virus are attached to other files.
- Worms propagate by themselves.
- both are Exponential attacks as each program can infect many hosts and then proceed to search for and infect more.
- The mathematics and epidemiology is the same as that in biology.

3.4 Bot Nets

- commonly, virus/trojan infected machines to form a remotely controlled large networks
- these are commonly rented out
- used to launch distributed denial os services (DDoS) attacks (usually for Blackmail)
- used to accept and relay SPAM

4 Cryptography

4.1 Encryption

- standard encryption is symmetric. Same key to encrypted and decrypted message.
- AlgorithmsL: Data Encryption Standard (DES) and Advanced Encryption Standard (AES)
- same text encrypted with the same key will result in the same cipher text.
- use an random initialization vector and XOR the message

Main problem is key distribution.

4.2 Key

- Key exchange method:
  - exchange key through another method: telephone / letter etc.
  - Diffie-Hellman key exchange.
- Numbers of Key is proportional to N^2 for an N party network
  - use a Key Distribution Center
- Solution: Asymmetric Cryptography (Public Key Cryptography)

4.3 Public Key Cryptography

Encryption key : public key
Decryption key : private key
private key is never communicated and can not be eavesdropped.

one way function use large prime numbers and modular arithmetic

- a one way function is where it is very computationally expensive to reverse the calculation

The process:

1. choose 2 very large prime numbmers p and q
2. n = pq and $m = LCM\{p-1, q-1\}$(lowest common multiple)
3. choose r where r > 1 and coprime with m
4. find s such that $rs- > 1(mod\ m)$
5. n and r are the public key
6. p q s are the private key

4.4 Encrypting and Decrypting

- To encrypt a message M:
  $M_c = M_r(mod\ n)$ [$0 < M < n$ and M and n are coprime]
- To decrypt a message M:
  $(M_c)^s == M(mod\ n)$

4.5 Authentication

the same public key system can be used for authentication

4.6 SSL/TLS

- ACK
- using asymmetric to certificate
- transfer data normally

4.6.1 SSL/TLS characters

- Secure Sockets Layers/Transport Layer Security
- Widely used for web applications
- Overview of the protocols:

  - server must have pub/pri key pair
  - should have a certificate
  - connects to server use https
  - client and server should negotiate a cryptographic algorithm to use for session and the server assign a sessionID.
  - allows the same keys to be used for a period instead of using new ones for each connection.

process:

1. server sends the certificate to the client
2. client checks the certificate using the root certification authority's key(supplied with the browser)
3. checks the names of the server
4. self generated key warn (unless the pseudo root certificate is installed on the client browser)
5. the client generate the key material for encrypting the traffic
6. material signed using the server's public key and sent
7. server and client then exchange a message authentication code(MAC) to ensure that both agree with the key exchanges up to that point

Now all data is encrypted and has a MACs added.(If the client stop and comeback to the session, client can still use the same key as long as the session is still valid).
Caveats with SSL/TLS: all you know is that you are communicate with the server in an encrypted channel.

4.7 Certificate

- web server manager creates a pub/pri key
- web server signs the request (required by signing CA) with private key and send to CAs
- Responsible CAs then go through a great deal of efforts to check the request is valid
- CAs added its pertinent information to the certificate and signs the request using **CAs private key**
- then CAs send the certificate back to webserver and I install it on my webserver

4.8 Secure shell

- use public key cryptography
- do not use certificate
- main purpsoe is to encrypt telnet/rlogin and ftp sessions
- on connection the client downloads the server's public key and displays it to the user
- symmetric encryption key exchange is performed
- session secured using 3DES or IDEA
- SSH used to tunnel many other protocols:X11

4.9 IPSec and VPNs

IPSec 3 protocols:

- Encapsulating Security Protocol (ESP)
- Authentication Headers (AH)
- Internet Key Exchange (IKE)

4.10 DNS security

DNS Security Extensions (DNSSEC)
sign zone data with a private key
use TCP for transport(Instead of UDP)
use TSIG for authenticated updates

5 Firewalls

- provide second line of defence
- could be circumvented for other hosts you trust may become compromised

5.1 Define

- an entity on the network used to perform access control on the network traffic
- software running on a host : dependent on the host kernel being free of vulnerability
- a dedicate piece of hardware : more secure
- a dedicate virtual machine

5.2 Methods of Access control

- static filtering
- dynamic fitering
- content based access control (CISCO)

Static Fltering TCP and UDP packet will be identifiable by the 4-tuple of: source of IP address/source Port/Destination of IP address/Destination Port(allow combination)

Have to remember to allow traffic in the other direction as well.

Dynamic Filtering static + make a note of addresses and port numbers and dynamically installs a rule to allow the reverse traffic.

May have problem when the control channel and data channel negotiated dynamically whithin the application.
Also problematic with asymetric routing

Content Based Access Control look inside of the command stream and set up rules to enable the protocol to work

5.3 Next generation Firewall Features

- content inspection
- email attachment sandbox
- machine learning

6 Denial-of-Service

Dos:

- Internet is designed to lower transmission costs
- Default ON
- Increased by BotNets

6.1 Source Spoofing

- ISP/AS should check
- Other ASes can do Reverse Path Check
- (Big problem is asymetric)

6.2 SYN cookies

encode the client initial sequence number in the server sequence number

6.3 BotNets

- many computers are controlled by the Botnet Herder
- hard to detect
- traffic comes with the right source address
- attackers can return the right ACK to SYN-ACK

the solution:

- Cloud Computing helps
- Machine Learning helps

6.4 SPAM

sending identical messages to thousands of recipients.

Perpetrators often harvest addresses of prospective recipients from Usenet postings or from web pages, obtain them from databases or simply guess them by using common names and domains. By definition SPAM occurs without the permission of the recipients. Represents 95% of the mail sent in the Internet. Major problems for ISPs.

6.5 Fighting SPAM

- Bayesian filtering
- Governments have legal options...
- 'electronic stamps'(proposed)
- ISPs have several technologies

SPF sender policy framework.
SPF works by domains publishing ”reverse MX”records. only receive message from specified list.(domain/server)