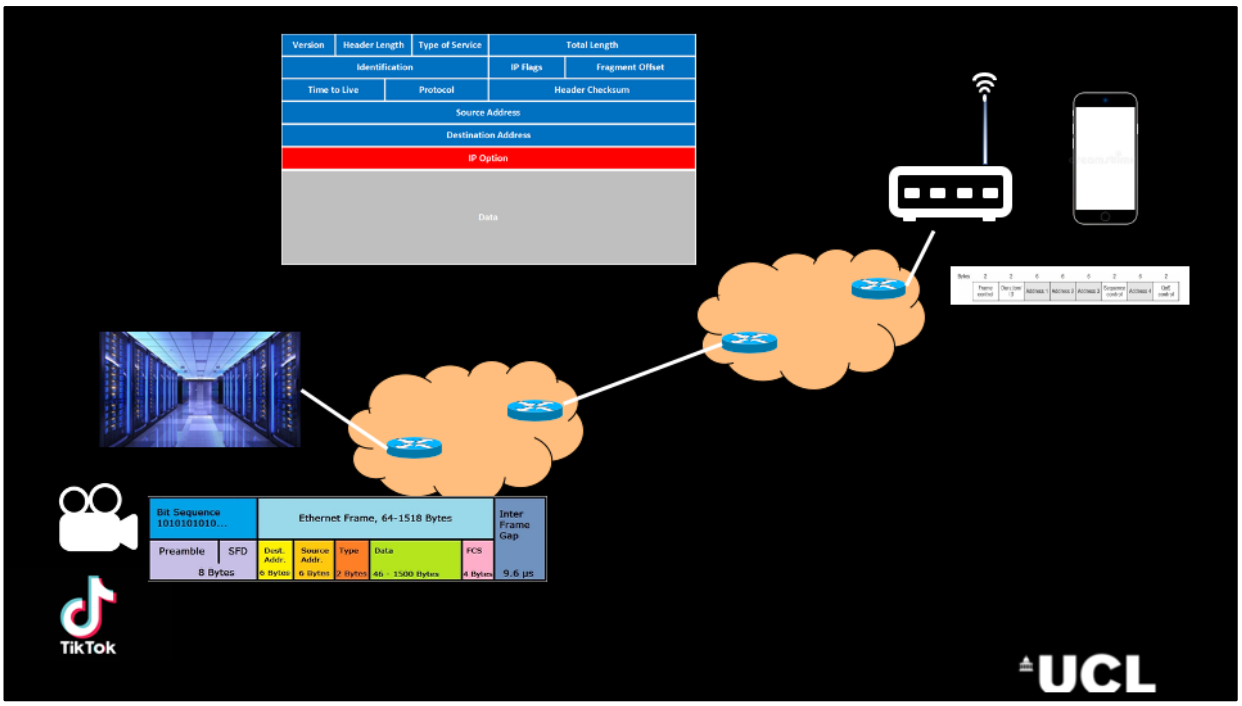
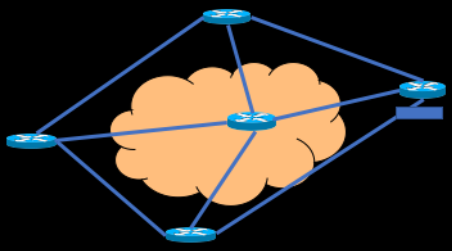
Introduction to TCP/IP networks

Miguel Rio

UCL


Version
Header Length
Type of Service
Total Length
Identification
IP Flags
Fragment Offset
Time to Live
Protocol
Header Checksum
Source Address
Destination Address
IP Option
Data
Bit Sequence
1010101010...
Ethernet Frame, 64-1518 Bytes
Inter Frame Gap
Preamble
SFD
Dest. Addr.
Source Addr.
Type
Data
FCS
8 Bytes
6 Bytes
6 Bytes
2 Bytes
46 - 1500 Bytes
4 Bytes
9.6 µs
TikTok
UCL

Intra-domain routing



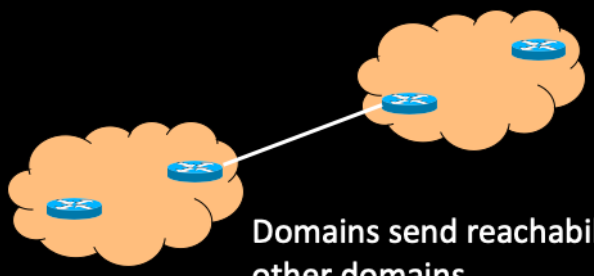
Exchange of routing messages

Calculation of shortest paths

Creation of routing tables

UCL

Inter-domain Routing

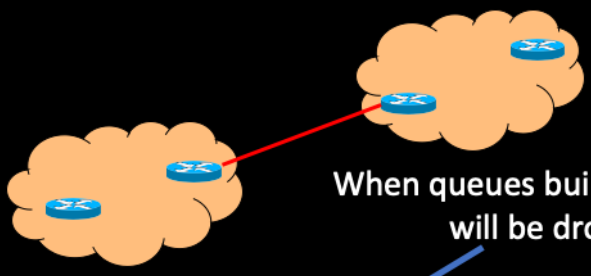


Domains send reachability messages to other domains

Domains decide to which domain to send

UCL

Transport



When queues build-up packets will be dropped

Retransmissions

Slow Down Transmi

UCL

TCP/IP outline

1. Data Link Layer (Ethernet and Wi-Fi)
2. The IP protocol (versions 4 and 6)
3. Intra—domain routing
4. Inter-domain Routing
5. Transport
6. State of the Internet

UCL

The Data Link Layer

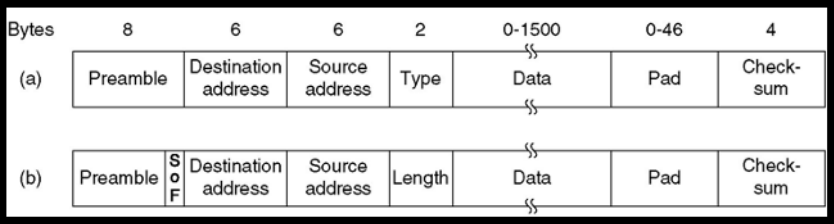
UCL

What is Ethernet ?

- Main technology to connect Local Area Networks
  - Also used in wide area networks
- Developed in the 1980s
- Information is put in frames
- Two ways of operating
  - CSMA/CD
  - Switching

UCL

IEEE 802.3 & DIX Ethernet Frame Formats

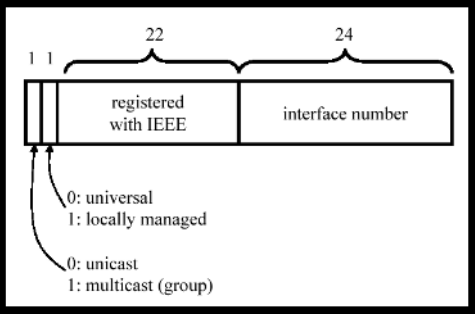


- Ethernet type II Frame Format (DIX): (a)
  - Uses a Type field after source address
  - 8 bytes Preamble, (no SFD)
- 802.3 Ethernet Frame Format: (b)
  - Length field for length of data
  - 7 bytes preamble, 1 byte SFD

UCL

MAC address

- Destination and Source addresses also known as MAC addresses
- MAC addresses identify network entities in Ethernet LANs
- Unique for each LAN interface
- 48 bits in length
- 22 bits identify the organisational unique identifier (OUI) 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 may be unicast, multicast or broadcast



UCL

CSMA/CD Defined

CS - Carrier Sense (Is someone already talking?)

MA - Multiple Access (I hear what you hear!)

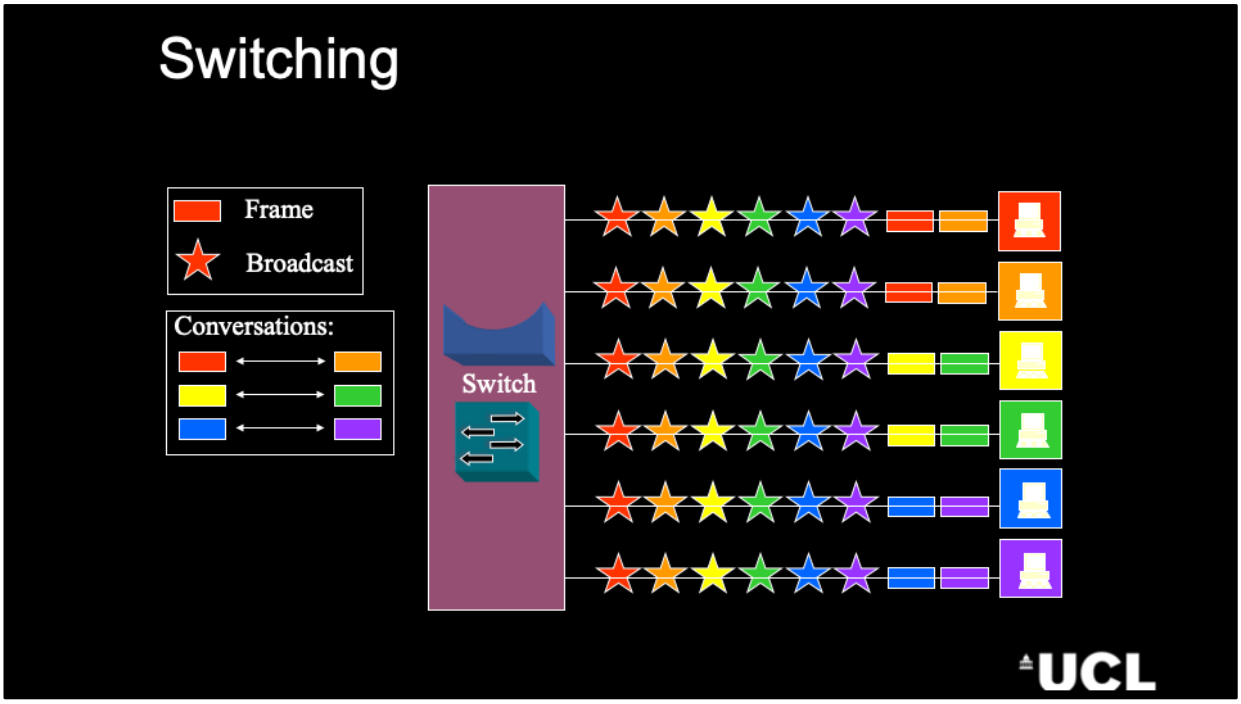
CD - Collision Detection (Hey, we're both talking!)

1. If the medium is idle, transmit anytime.
2. If the medium is busy, wait and transmit right after.
3. If a collision occurs, backoff for a random period, then go back to 1.

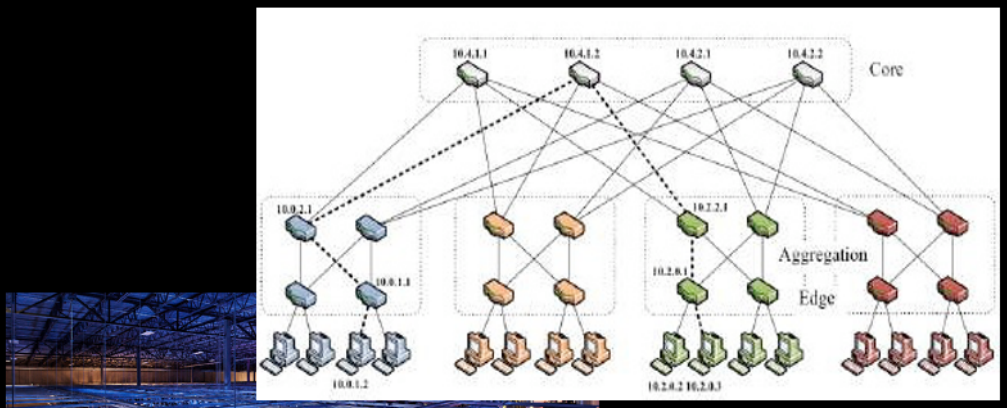
We use CSMA/CD in normal group conversation.

UCL

Switching



Inside the cloud

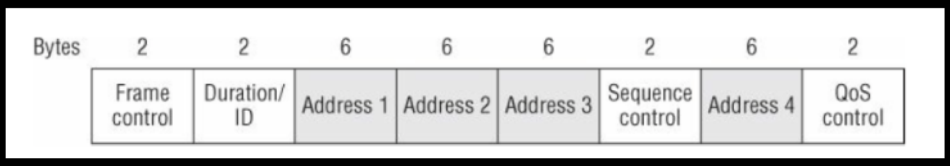


UCL


802.11
UCL

802.11: Wi-Fi

- Similar Frame to Ethernet but more complex
- Addresses are the same



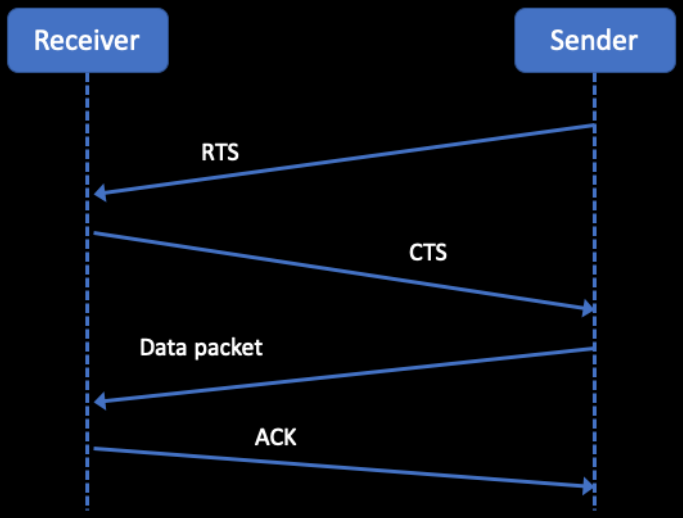
UCL

Collision Management : 4 Way Handshake

1. “Listen before you talk” : If the channel is busy, node backs-off for a random amount of time
2. But now, instead 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)
5. After receiving the CTS, the original transmitter transmits the information packet.
6. Other nodes in range of the receiver detect the CTS signal and refrain from transmitting
7. If the packet has been received correctly the receiver sends out an ACK packet.
8. If the information packet is not ACKed, then the source starts again and tries to retransmit the packet.

UCL

4 Way Handshake flow



UCL

Summary

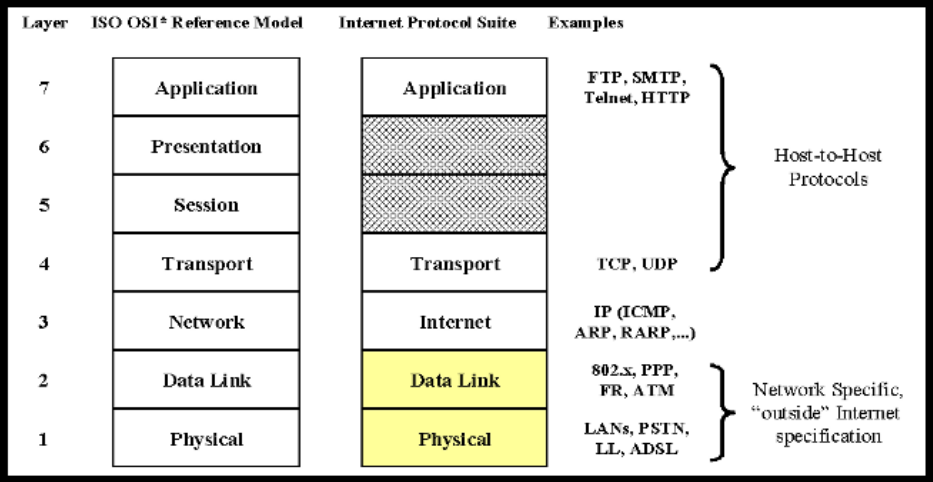
- Ethernet Frame for point-to-point communication
- MAC addresses
- Two ways of communication:
  - CSMA/CD
  - Switching
- Wi-Fi: 802.11
  - 4-way handshake

UCL

IP – The internet Protocol

UCL

ISO/OSI & Internet Protocol suites



* OSI: Open Systems Interconnection – Basic Reference Model, ISO 7489

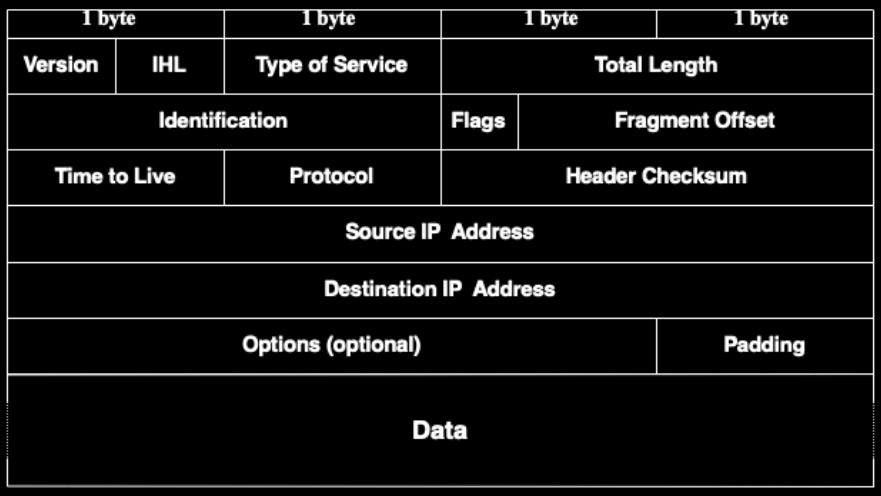
UCL

IP - Internet Protocol (RFC 791)

- Connectionless service
- Network addressing
- Best effort delivery
  - IP datagrams may arrive at destination host damaged, duplicated, out 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

UCL

IP Version 4 Datagram Structure



UCL

IPv4 Addresses

- This have 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 (4 bytes) and are represented by dot notation. E.g:
- 138.77.45.3

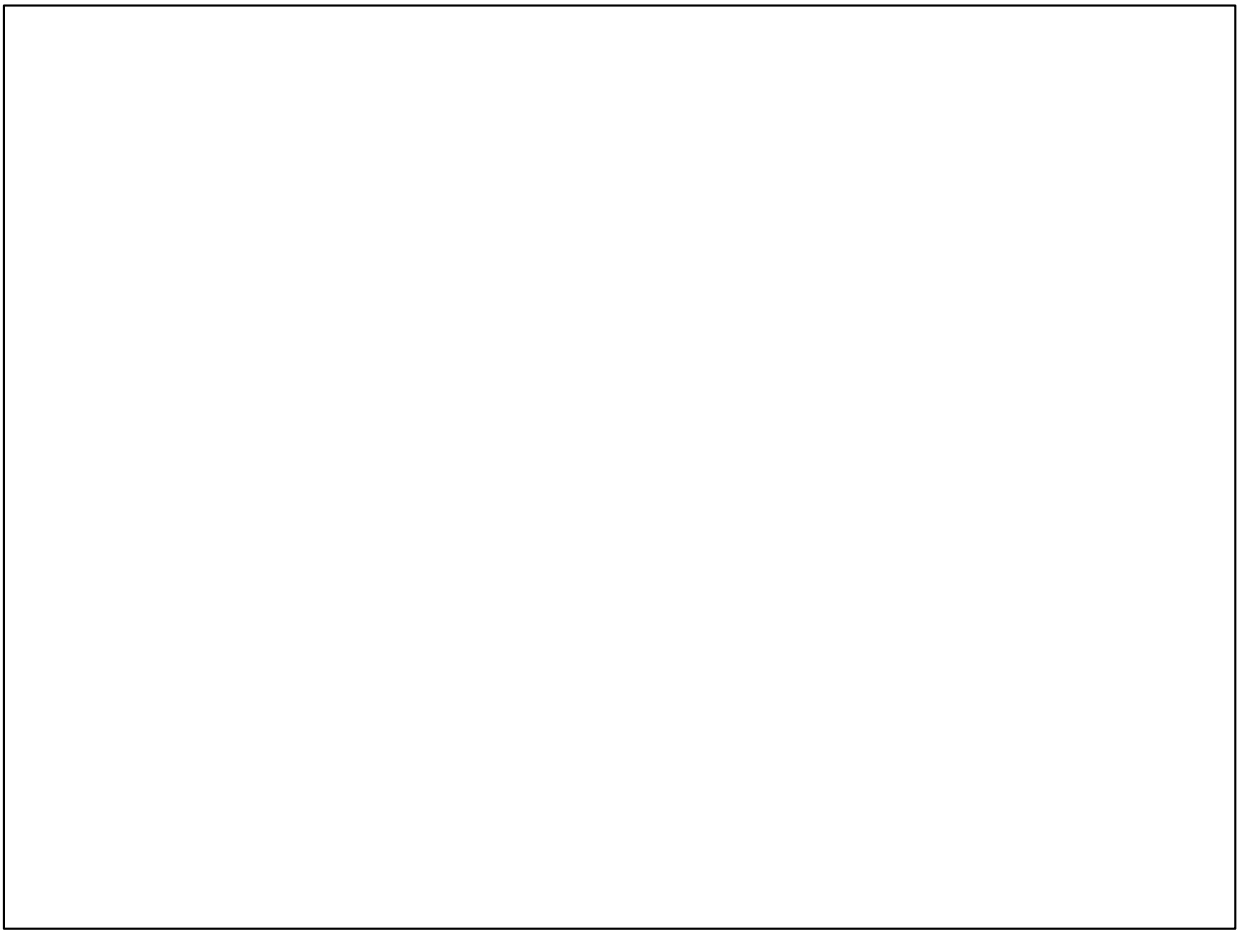
10001010 01001101 00101101 00000011

UCL

Special IPv4 Addresses

- All 0 host suffix => Network Address
- All 0s network => This network
  - (e.g. 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 Internic
- Private IP addresses (RFC 1918)
  - Any organisation can use these inside their network. However these 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

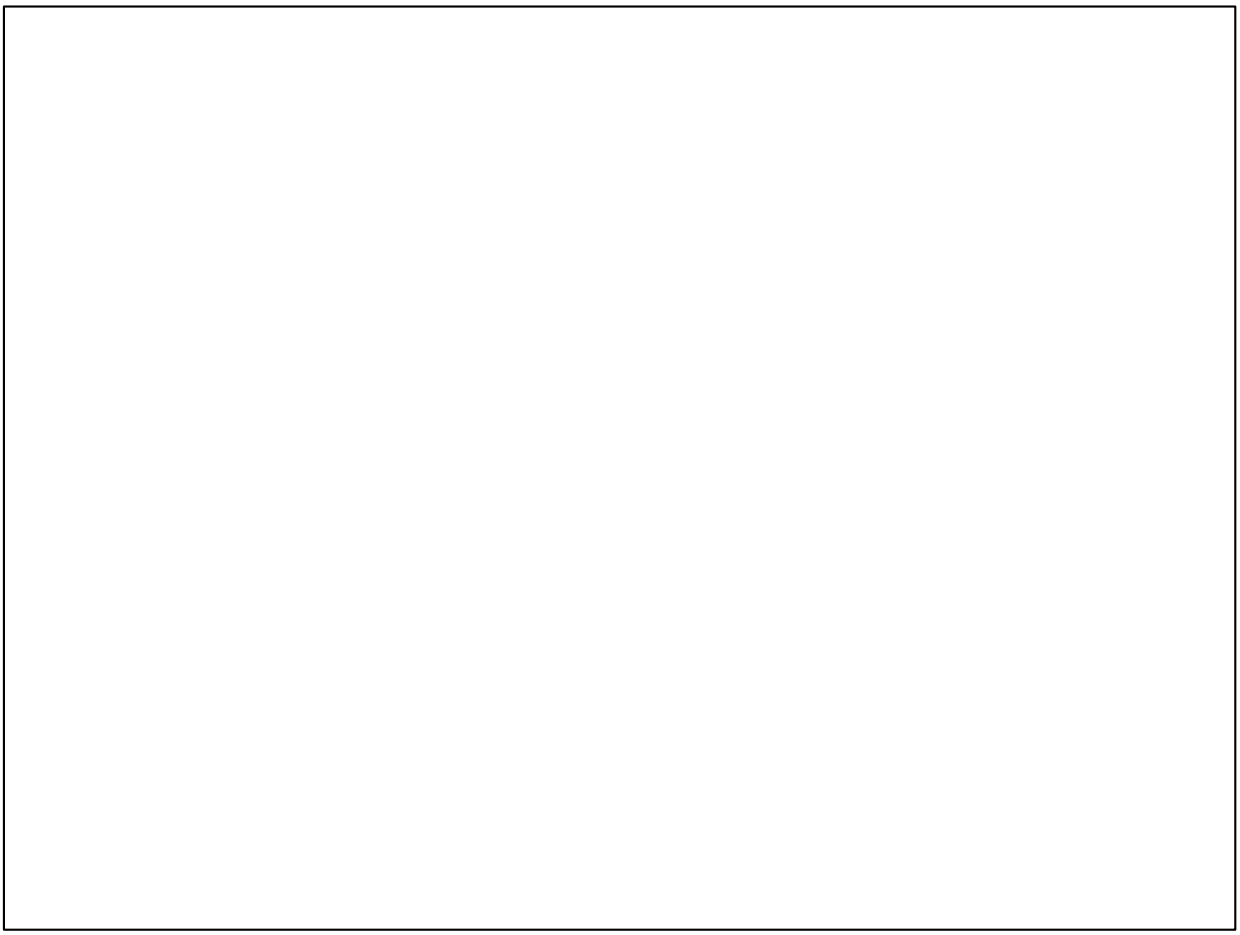
UCL



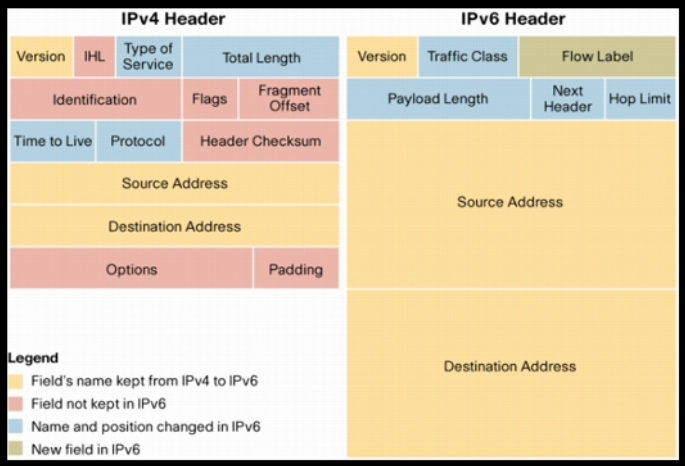
Problems with IPv4

- Shortage of IP addresses
  - The 32bit address system in IPv4 can theoretically recognise 4.3 billion hosts. This is not enough for widespread adoption of IP in multiple devices
- Insufficient security functions
  - In IPv4 security is typically a function of the upper layers. Scalability requires that robust security measures on the IP datagram are available
- Fragmentation introduces complexity
- No Quality of Service support
- Complex header

UCL



IPv6

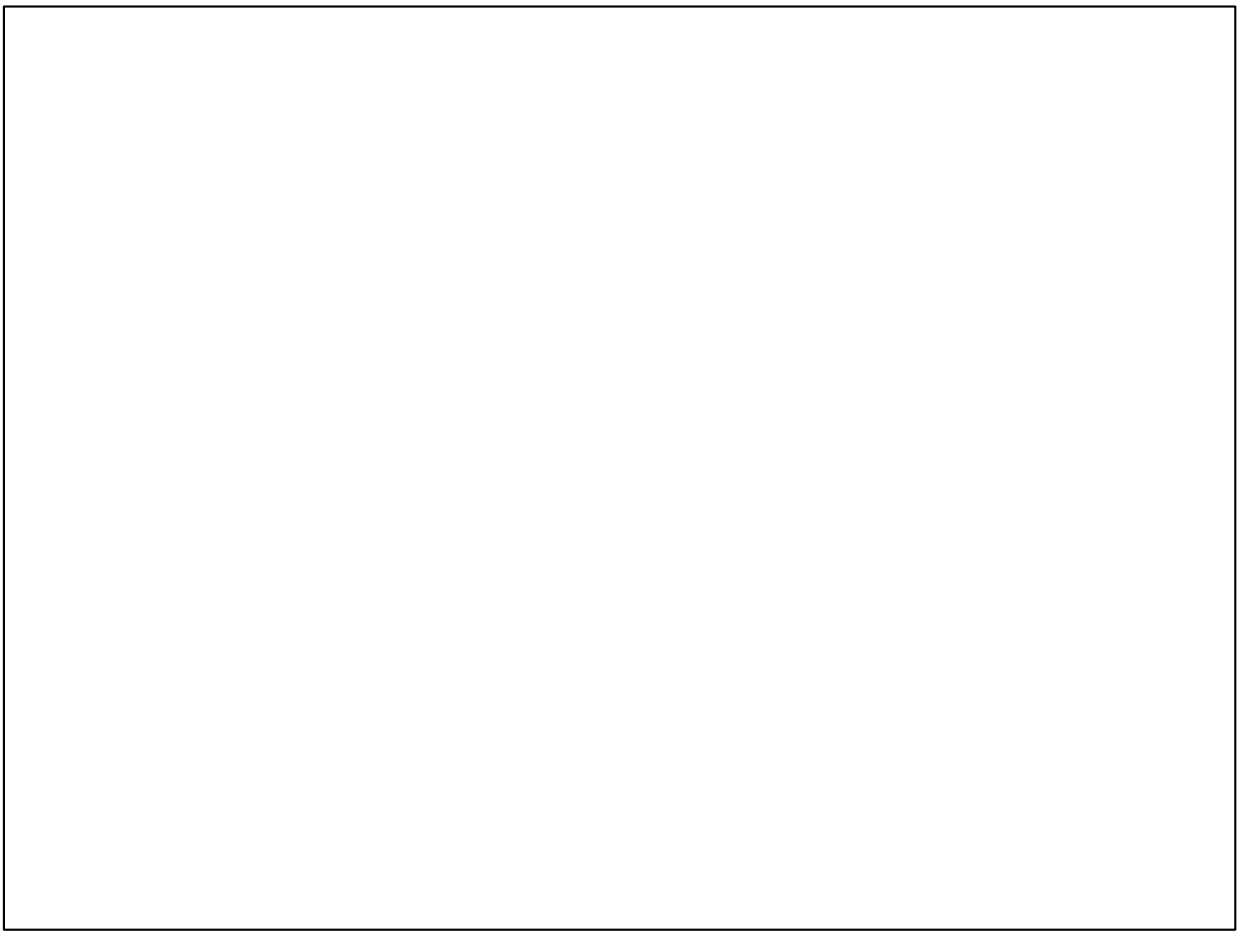

IPv4 Header
Version
IHL
Type of Service
Total Length
Identification
Flags
Fragment Offset
Time to Live
Protocol
Header Checksum
Source Address
Destination Address
Options
Padding
IPv6 Header
Version
Traffic Class
Flow Label
Payload Length
Next Header
Hop Limit
Source Address
Destination Address
Legend
Field's name kept from IPv4 to IPv6
Field not kept in IPv6
Name and position changed in IPv6
New field in IPv6

UCL

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
- Automatic allocation of IP addresses
- Simplified header
- Allows Jumbograms (very big packets)

UCL



IPv6 Address notation

- IPv6 uses 8 groups of 4 hexadecimal digits: 2001:0630:0013:0200:0000:0000:0000:ace0
- This can be ‘tidied 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 /:
- E.g. 2001:630:13:200::ace0/64 indicates a network address of 2001:630:13:200:: and a host address of ::ace0

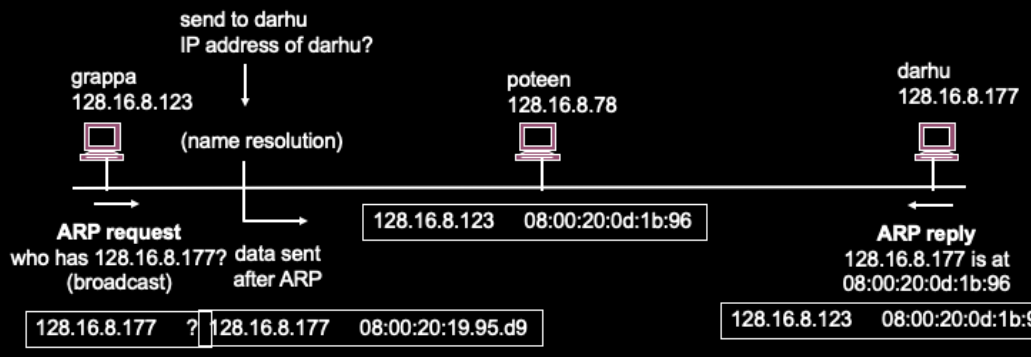
UCL

Address Resolution

UCL

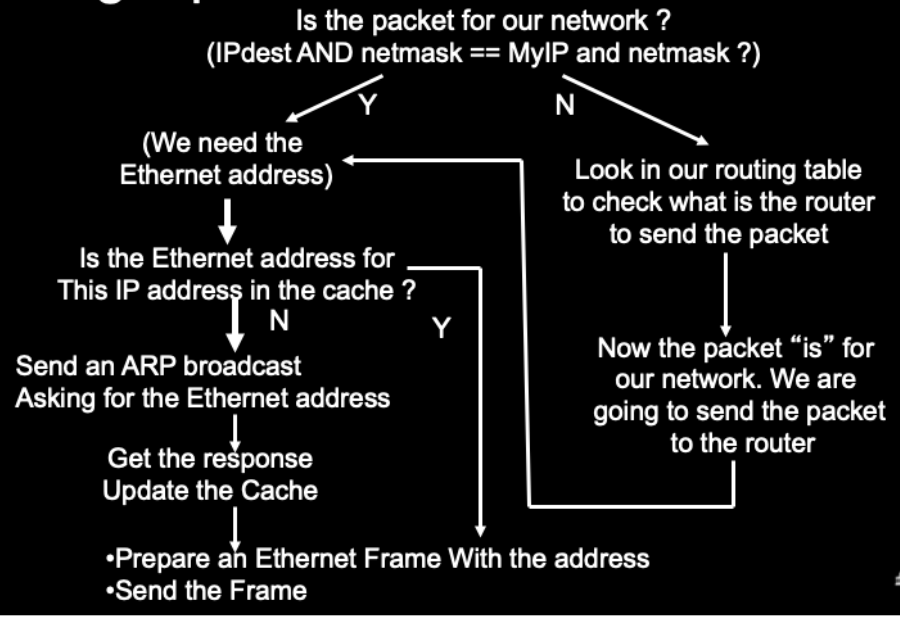
Address Resolution Protocol

- Given an IP address what is the MAC address of machine? The solution:
  - The host broadcasts a request, What is the MAC address of 10.10.12.02?
  - The host whose IP address is 10.10.12.02 replies back, "The MAC address is ……"



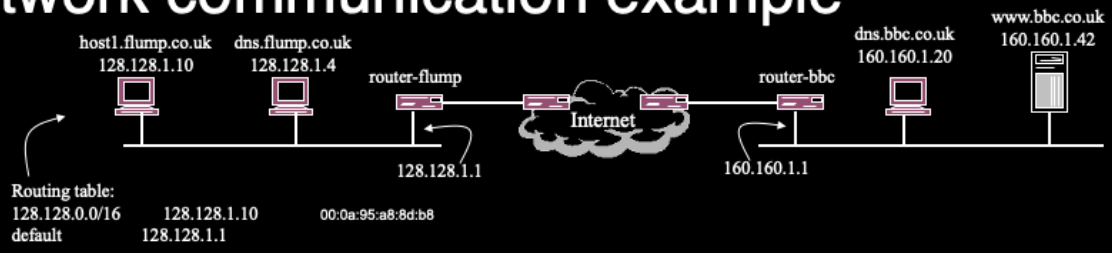
UCL

Sending a packet…



UCL

Network communication example



1. host1 wants to communicate with www.bbc.co.uk (160.160.1.42)
2. host1 verifies that 160.160.42 is not in its network. It therefore needs to send it to a router. It checks in its routing table and verifies that the router it should use is 128.128.1.1
3. 128.128.1.1 is in its network (it had to !!!) so the packet can be sent directly. host1 issues an ARP request to know what is the Ethernet address of 128.128.1.1. This is broadcasted to all hosts in the network
4. 128.128.1.1 replies to host1 with its Ethernet address 00:0a:95:a8:8d:b8
5. host1 now puts the IP packet (the Destination address of the IP packet is 160.160.1.42) in an Ethernet Frame with Destination address 00:0a:95:a8:8d:b8. It sends this frame to the

UCL

ICMP

UCL

Internet Control Message Protocol (ICMP)

- ICMP is used by IP to send error and control messages
- Sent by end hosts or by routers

UCL

ICMP Messages

Type	Code	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	soçurce quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

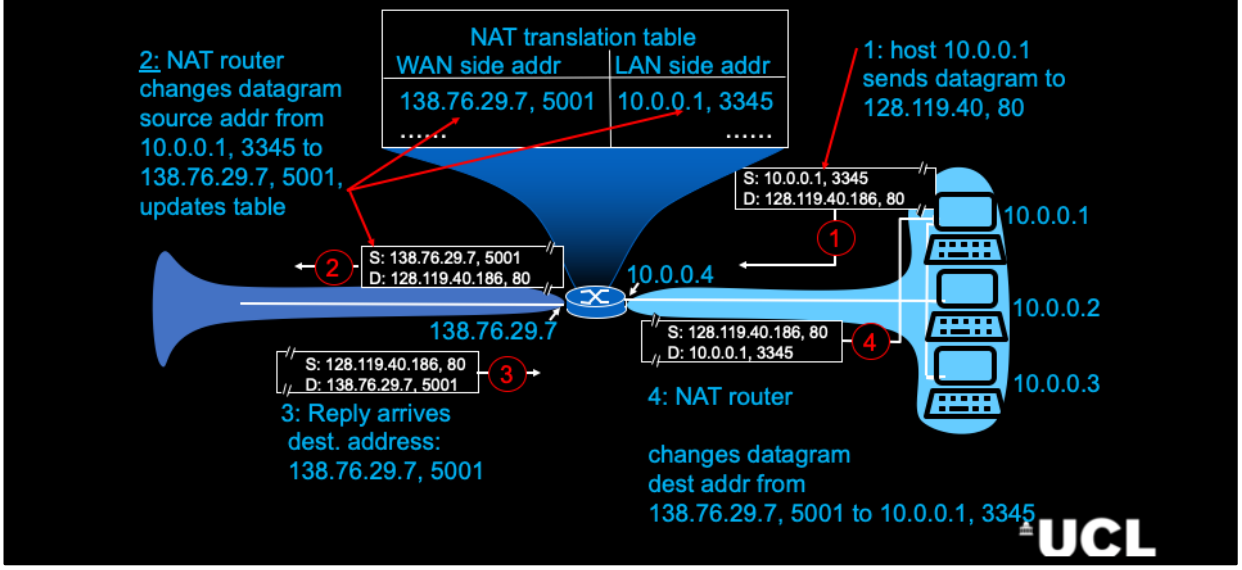
- ICMP messages are carried in IP packets
- …but conceptually we see ICMP at the same level of IP

UCL

Network Address Translation

UCL

NAT: Network Address Translation



Summary

- The IP protocol
  - Best-effort
  - Connectionless
  - Addressing
  - Forward packets
- Packet headers (IPv4 and IPv6)
- Address Resolution Protocol
- ICMP
- NAT

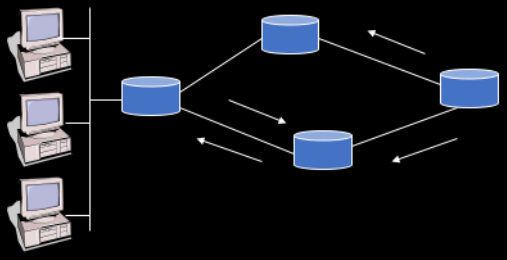
UCL

Intra-domain Routing

UCL

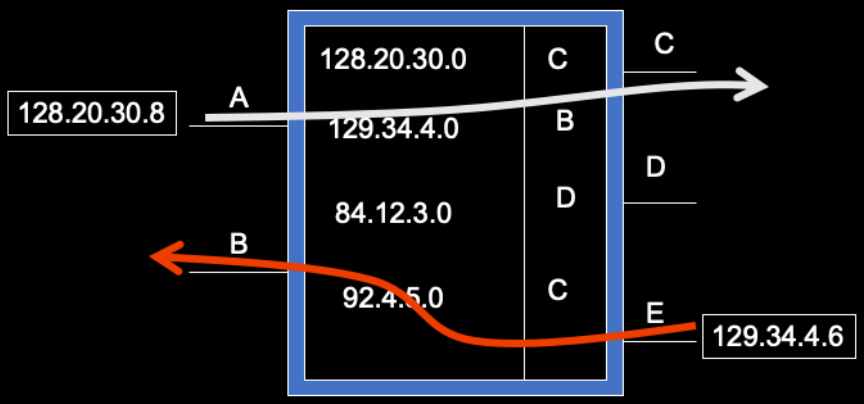
What is routing ?

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



UCL

Routing Tables



UCL

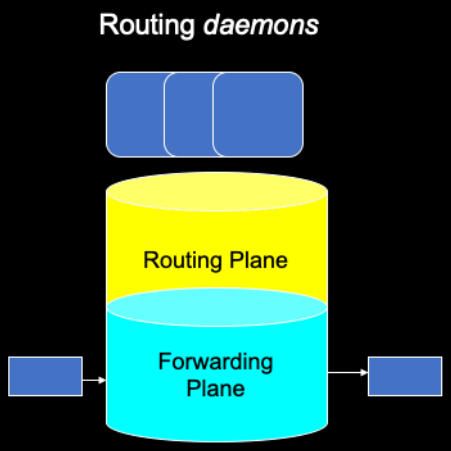
Routing

- 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

UCL

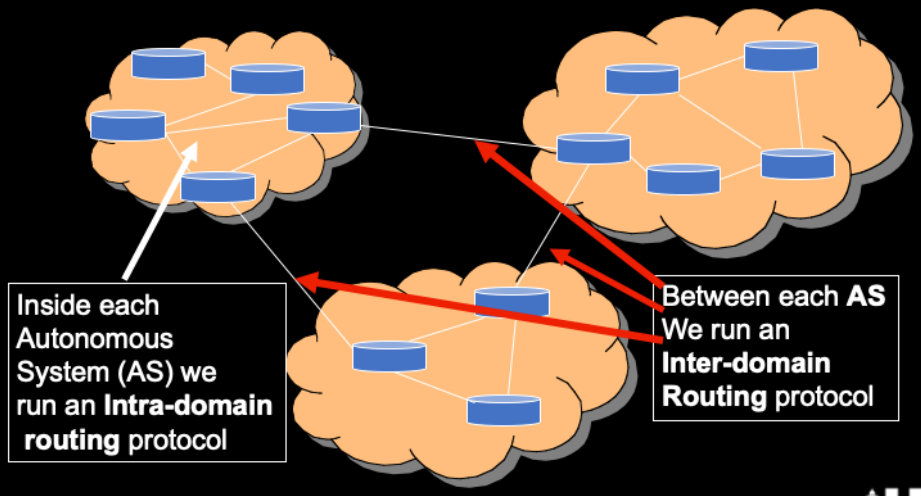
Routing/Forwarding Plane

- Routers are just computers with more than one network interface
- The forwarding plane forwards every packet. It needs 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



UCL

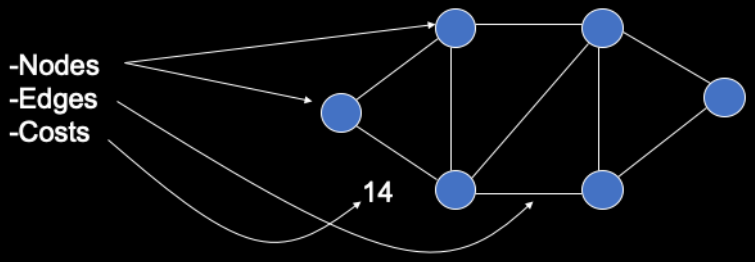
The Big Picture: Intra-domain vs inter-domain



UCL

Graphs - A useful mathematical abstraction

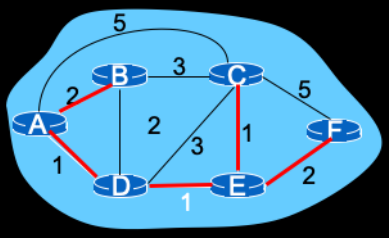
- Several real-life problems can be approached by using graphs
- These are used in lots of fields like road traffic optimization
- Graphs have:



UCL

Abstraction

- Mathematically we model networks with graphs G=(V,E) with costs f(A,B)=Cab
- Nodes are routers, edges are links
- It becomes easier to think about algorithms
- The information is stored in a matrix
- It just represents a snapshot in time



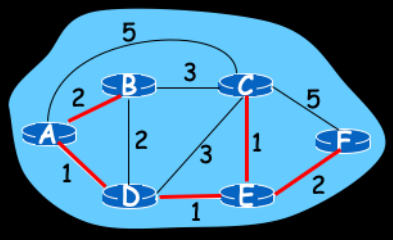
$$\begin{bmatrix} 0 & 2 & 5 & \infty & 1 & \infty \\ 2 & 0 & 3 & 2 & \infty & \infty \\ 5 & 3 & 0 & 3 & 1 & 5 \\ 1 & 2 & 3 & 0 & 1 & \infty \\ \infty & \infty & 1 & 1 & 0 & 2 \\ \infty & \infty & 5 & \infty & 2 & 0 \end{bmatrix}$$

UCL

Dijkstra's Best Path algorithm: example

Step	start N	D(B),p(B)	D(C),p(C)	D(D),p(D)	D(E),p(E)	D(F),p(F)
0	A	2,A	5,A	1,A	infinity	infinity
1	AD	2,A	4,D		2,D	infinity
2	ADE	2,A	3,E			4,E
3	ADEB		3,E			4,E
4	ADEBC					4,E
5	ADEBCF					

```
1  Initialization:
2  N = {A}
3  for all nodes v
4    if v adjacent to A
5      then D(v) = c(A,v)
6      else D(v) = infinity
7
8  Loop
9    find w not in N such that D(w) is a minimum
10   add w to N
11   update D(v) for all v adjacent to w and not in N:
12     D(v) = min( D(v), D(w) + c(w,v) )
13   /* new cost to v is either old cost to v or known
14    shortest path cost to w plus cost from w to v */
15 until all nodes in N
```



UCL

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)

UCL

OSPF Overview (OSPF=Flooding + Dijkstra)

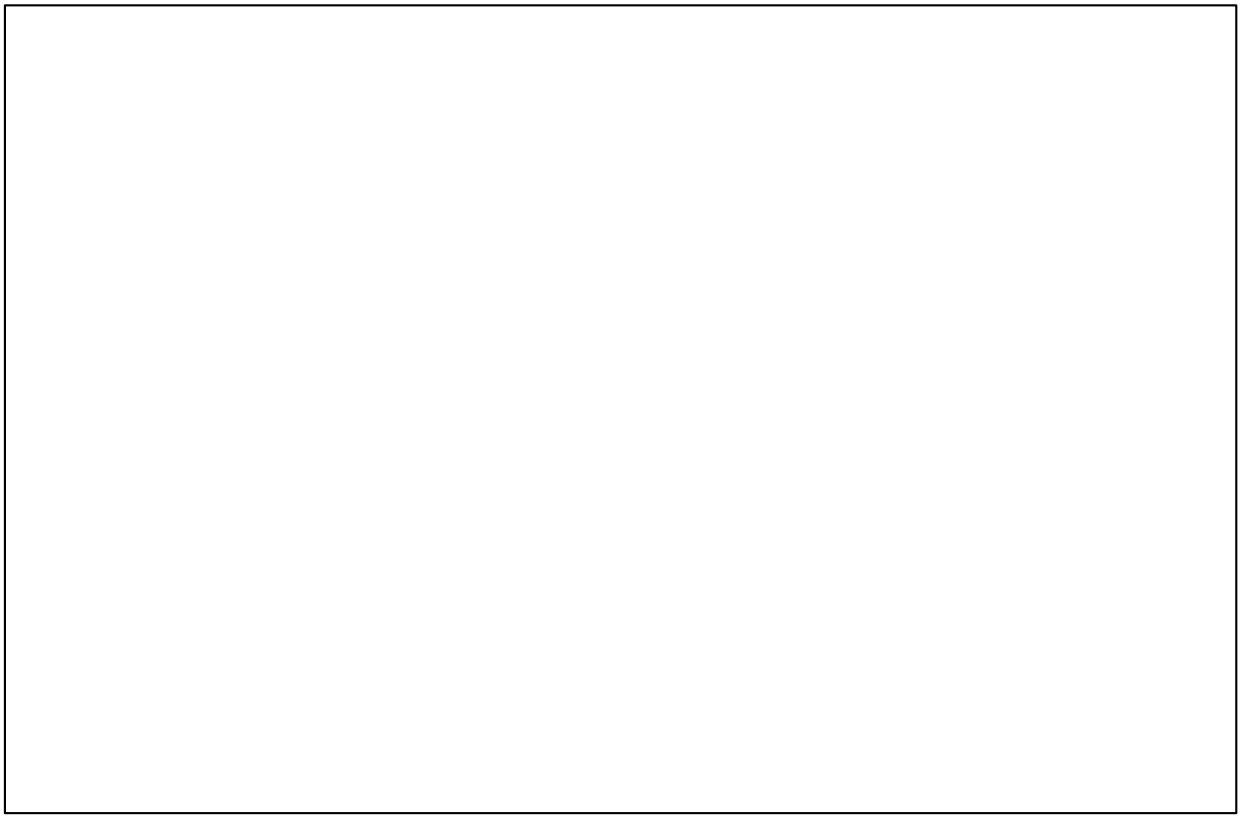
Router maintains descriptions of state of local links as a directed graph

1 - Transmits updated state information to all routers it knows about using flooding

- It sends 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 - Router receiving update must acknowledge

UCL

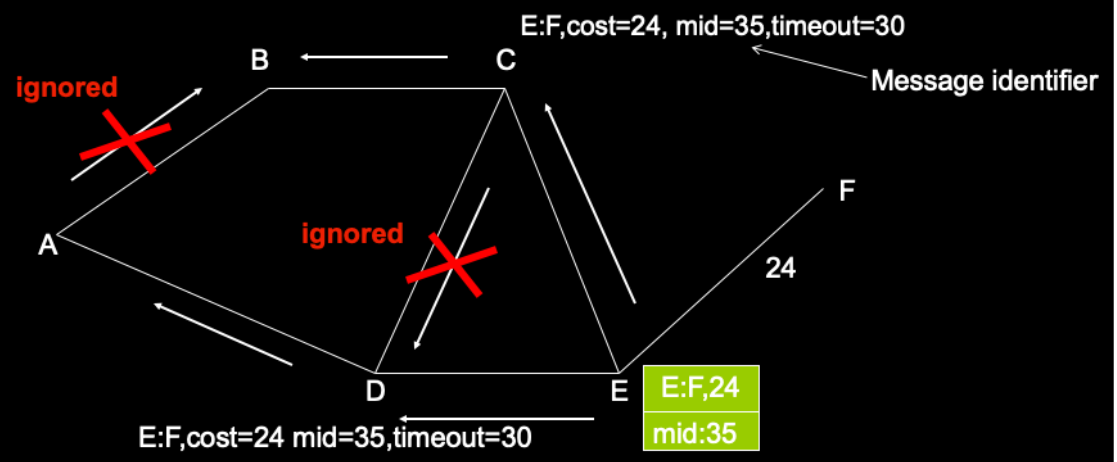


More OSPF…

3. After receiving all the message from the network OSPF routers calculate what is the shortest path to reach all the destinations
4. With that information they calculate what is the link to be used for every network
5. Because all the routers have the same information and calculate the same paths using the same algorithm all the forwarding decisions are consistent
6. If for some reason information is not consistent packets may get looped. Example: link goes down and that information has not arrived to all the routers. Remember TTL !!!

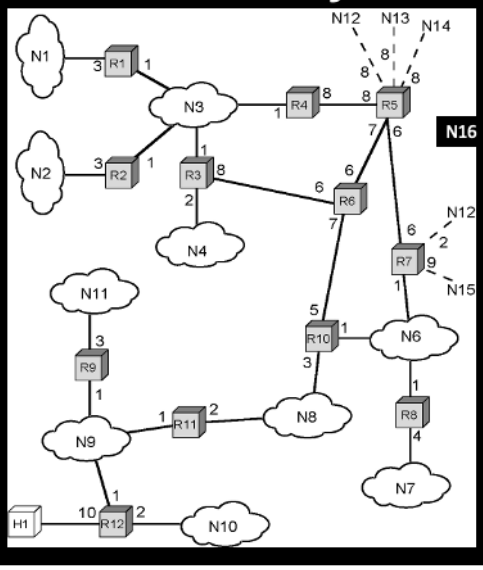
UCL

OSPF flooding in practice - example



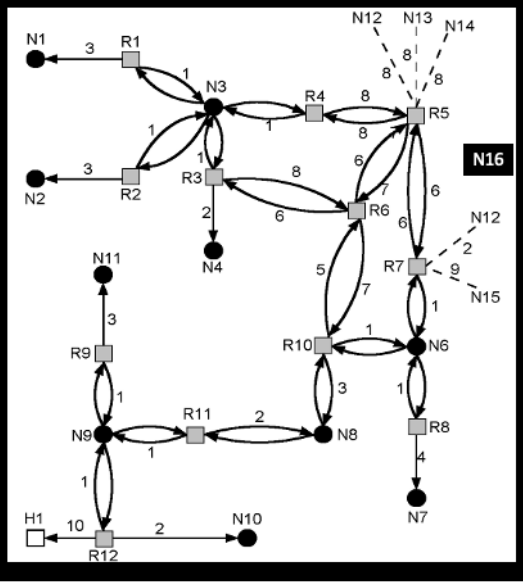
UCL

Sample Autonomous System



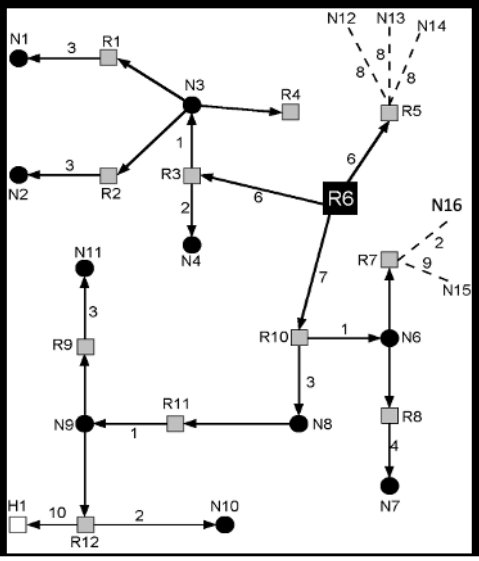
UCL

Directed Graph of Autonomous System



UCL

The Shortest Path Tree for Router R6



Routing Table:

N1: R3
N2:R3
N3:R3
N4:R3
N6:R10
N7:R10
N8:R10
N9:R10
N10:R10
N11:R10
N12:R5
N13:R5
N14:R5
N15:R10
N16:R10

UCL

Summary

- What is Routing
- Intra vs Inter domain Routing
- Dijkstra shortest path algorithm
- OSPF
  - Flooding
  - Shortest path
  - Routing tables

UCL

Inter-domain Routing

UCL

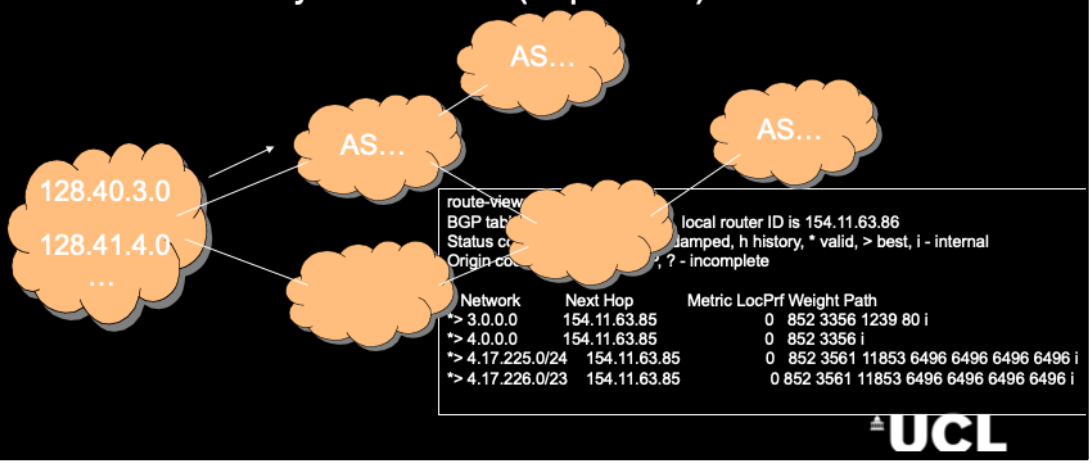
Routing Continued

- At this point you should know how IP packets traverse an Autonomous system to reach a destination inside that AS. Here all the routers are under the same administration
- Now we will look at how are packets routed over several Autonomous systems ?
  - Many more networks
  - Owned by different organizations

UCL

BGP - Border Gateway Protocol

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



UCL

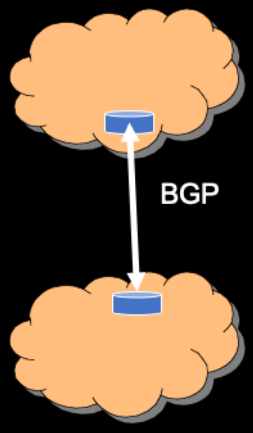
What is an Autonomous System ?

- The Internet is not controlled by any central 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 (e.g IBM), it can be a Research network (e.g. JANET) and it can be an ISP
- Each AS, runs one instance of an IGP (OSPF, RIP, etc)

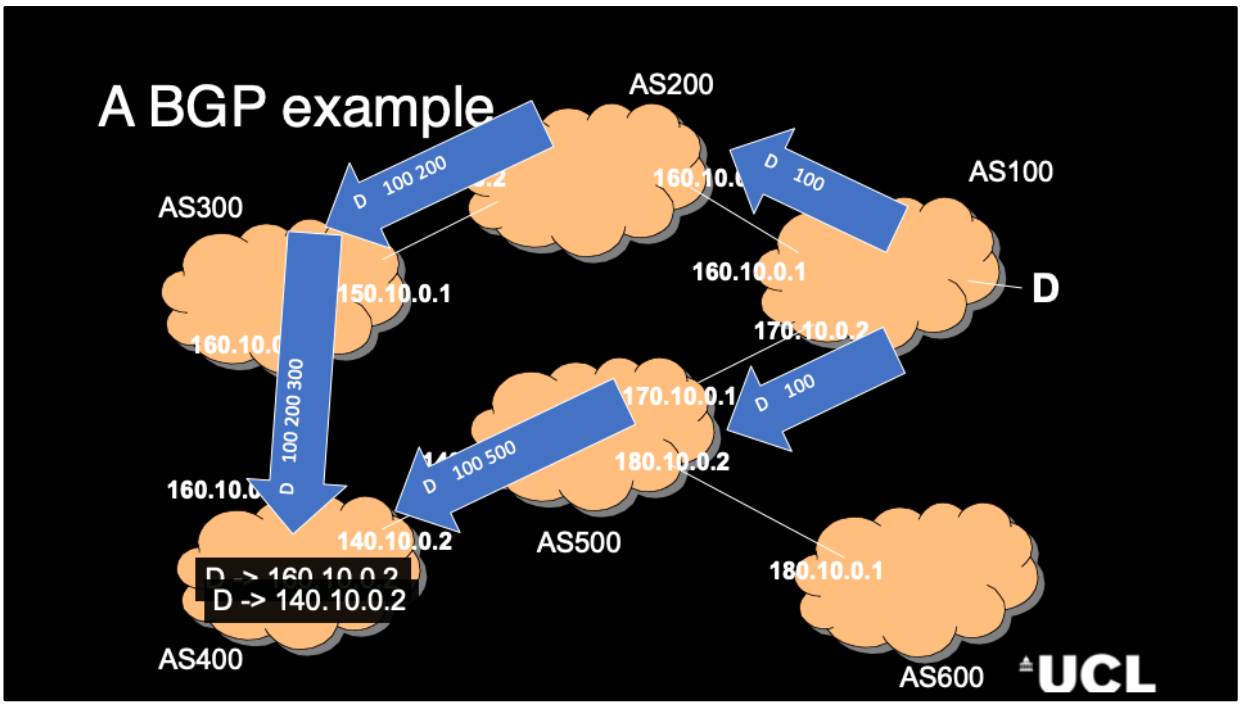
UCL

Border Gateway Protocol (BGP)

- Allows routers (gateways) in different ASs to exchange routing information
- Messages sent over TCP



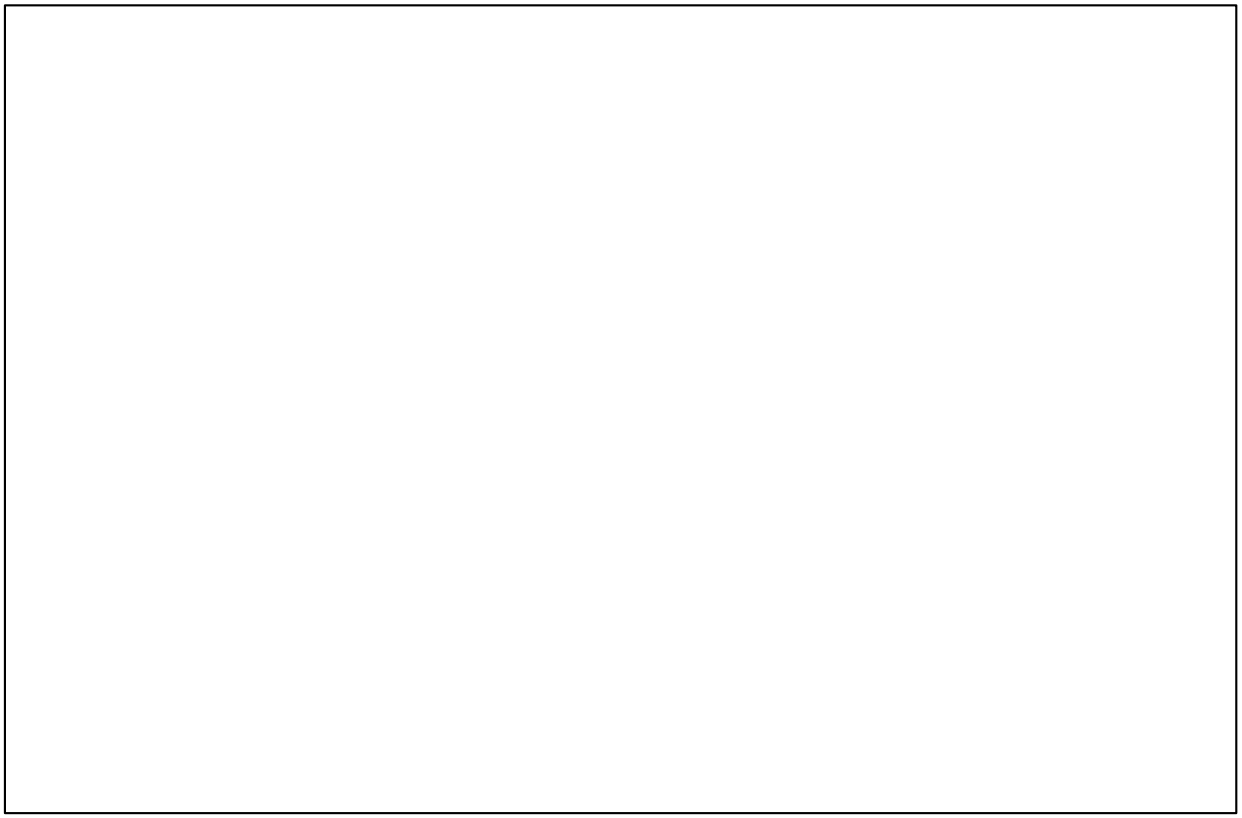
UCL

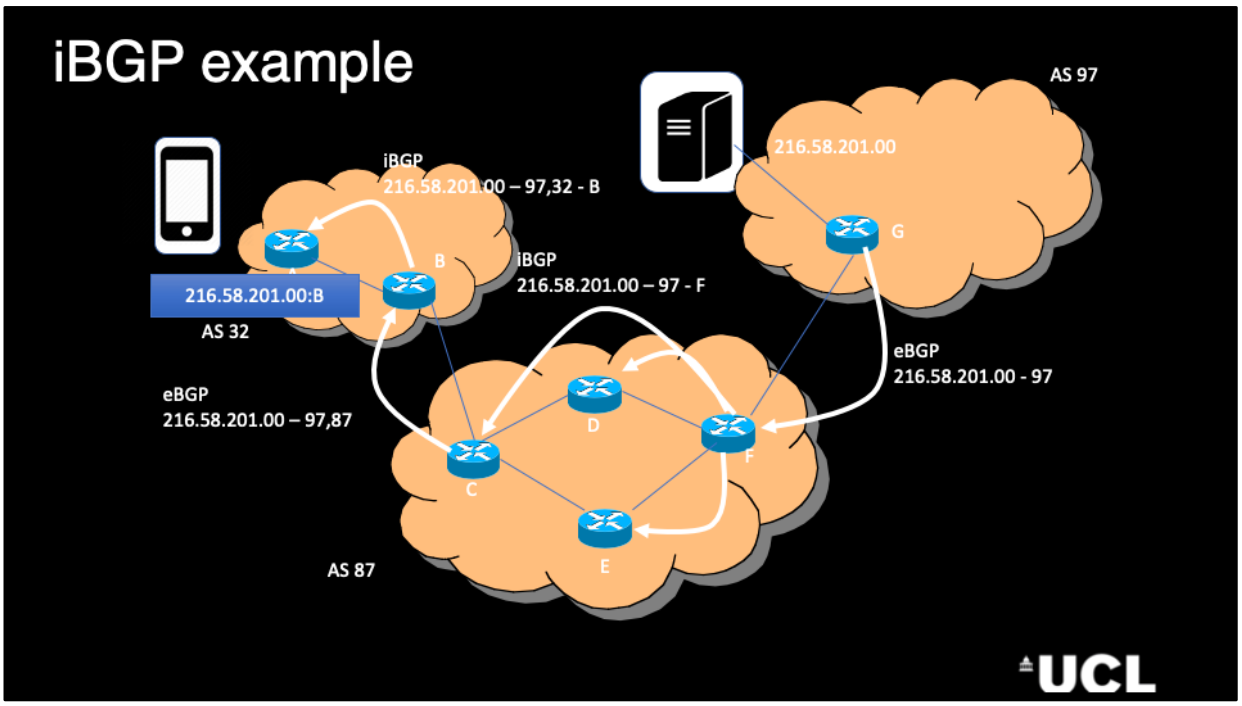

A BGP example
AS200
AS100
AS300
AS400
AS500
AS600
D
D 100 200
D 100
D 100 200 300
D 100 500
D 100
160.10.0.1
150.10.0.1
170.10.0.2
170.10.0.1
180.10.0.2
140.10.0.2
180.10.0.1
D -> 140.10.0.2
UCL

BGP= eBGP+ iBGP

- When we talk about BGP we usually mean eBGP
- iBGP it is the part of BGP 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

UCL




iBGP example
AS 97
216.58.201.00
iBGP
216.58.201.00 – 97,32 - B
G
B
iBGP
216.58.201.00 – 97 - F
216.58.201.00:B
AS 32
eBGP
216.58.201.00 - 97
eBGP
216.58.201.00 – 97,87
D
F
C
E
AS 87
UCL

Routing Summary

1. Destinations in the same LAN are sent directly through a switch
2. Destinations in the same Autonomous system use Intra-domain routing
    1. Flooding + shortest path
3. Destinations outside are learned through BGP
4. And are disseminated inside an AS through iBGP

UCL

Transport

UCL

…Until now…

- … we saw how information, in the form of IP packets, reaches any computer in the Internet…but:
- Several issues remain:
  - Information has to reach applications running on devices. Several applications run on each device
  - Packets may get lost
  - How fast should we send the data ?

UCL

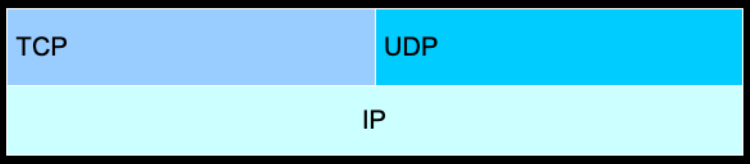
Two types of transport over IP

TCP

- Reliable Transport
- Controlled rate of transmission
- Complex

UDP

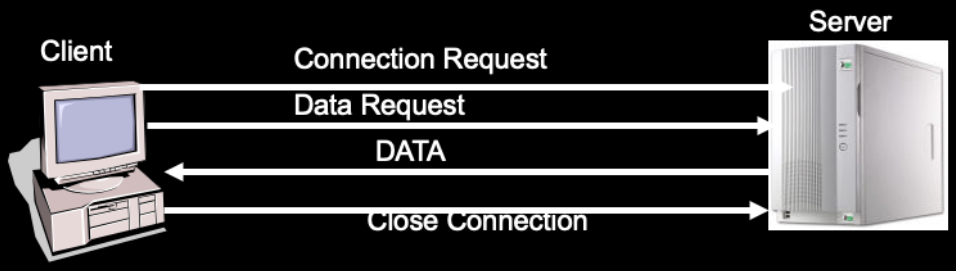
- Unreliable Transport
- Uncontrolled rate of transmission
- Very simple



UCL

The Client/Server paradigm

- Applications using TCP use usually the **client/server** paradigm
- Clients (web browser, email reader) connect to servers (web server, mail server), ask for information, get the information and close the connection



UCL

TCP functions

- **Demultiplexing** - Several flows arrive at the same destination host. We need a way of delivering the right packets to the right processes
- **Reliability** - Some flows require no packet loss. IP does not guarantee this. Somebody needs to manage retransmissions
- **Congestion Control** - How much traffic can we send without creating Congestion

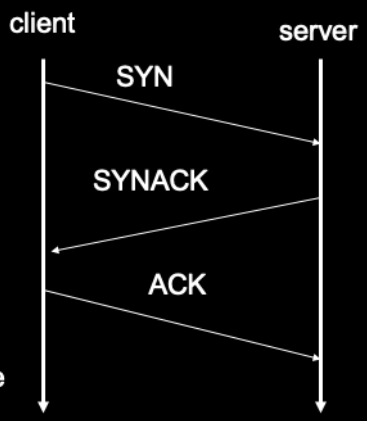
UCL

TCP -3 way handshake

Step 1: client host sends TCP SYN segment to server
- specifies initial seq #
- no data

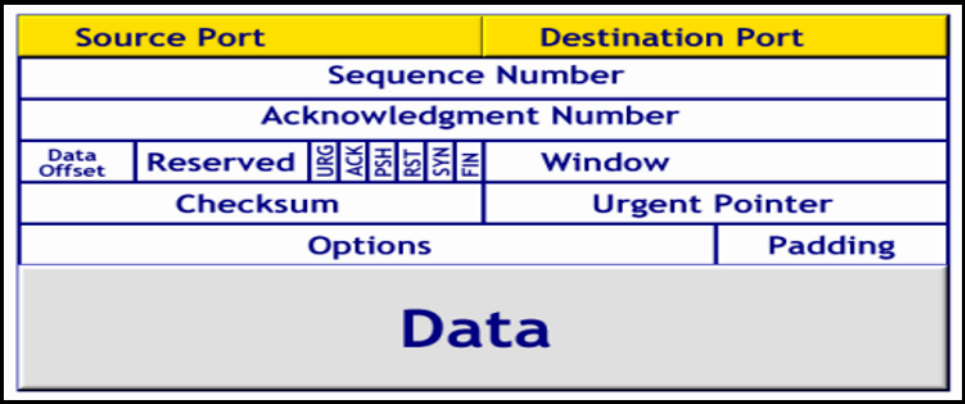
Step 2: server host receives SYN, replies with SYNACK segment
- server allocates buffers
- specifies server initial seq. #

Step 3: client receives SYNACK, time replies with ACK segment, which may contain data



UCL

The TCP packet



UCL

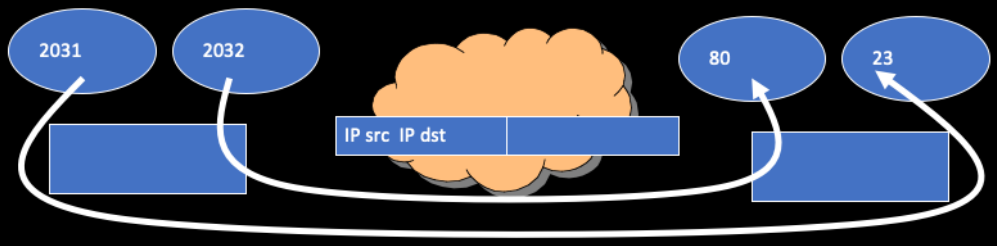
TCP Multiplexing

1 - IP Packets contain information to reach the end system

2 - But processes need to send data to processes

3 - To do this each process “listens” and sends to and from a port. These two ports are included in the TCP packet

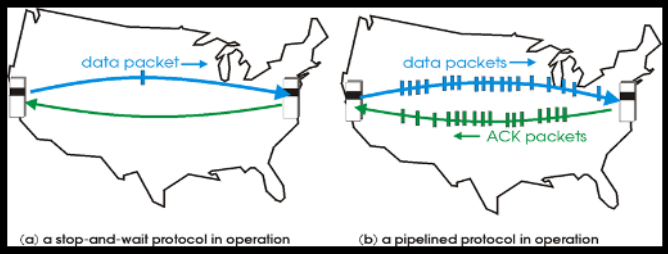
4 - Destination ports are standardized, source ports are chosen dynamically by the operating system



UCL

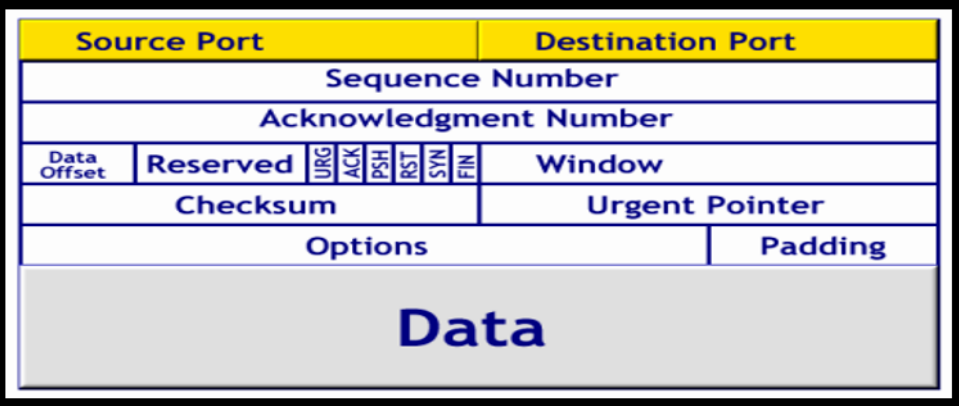
TCP reliability

- 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
- A computer does not need to receive an ACK of the nth packet to send the (n+1)th packet. At each moment in time there are x number of packet unacknowledged. This is called the **congestion window**. The Bigger the congestion window the faster the sending rate



UCL

The TCP packet



UCL

Congestion Control

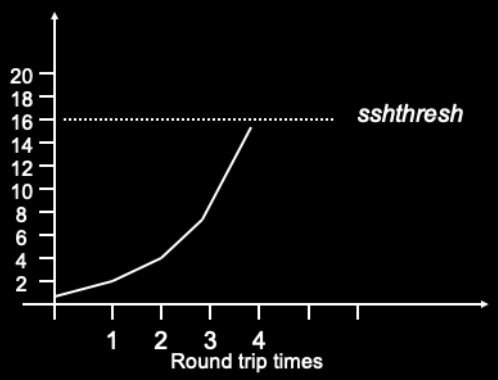
The Problem: A sender has to send packets to a remote destination but it does not know the status of all the links in the path

How fast should it send? That is: how big should the congestion window be?

UCL

Slow Start

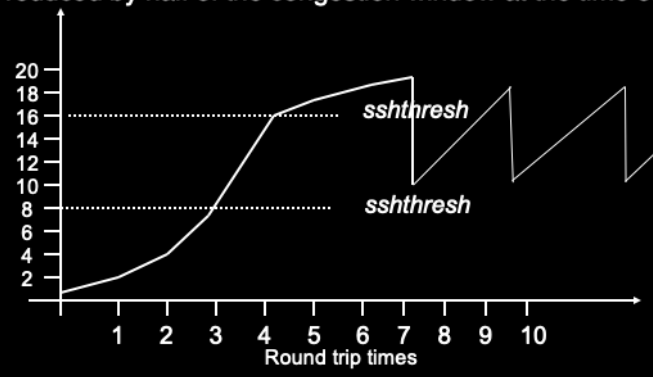
- Every TCP connection starts with a congestion window of 1
- The sender then applies the **Slow Start** algorithm: Every time an ACK arrives the Congestion window is increased by 1
- This is done until a value **sshthresh** is reached



UCL

Congestion Avoidance

- In the second part of TCP algorithm, the window is increased by 1/cwnd each time a ACK is received (this is equivalent to an increase in 1 every RTT)
- Every time a packet is lost, one reduces the cwnd by half. sshthresh is also reduced by half of the congestion window at the time of loss



TCP Sawtooth
Additive
Increase
Multiplicative
Decrease

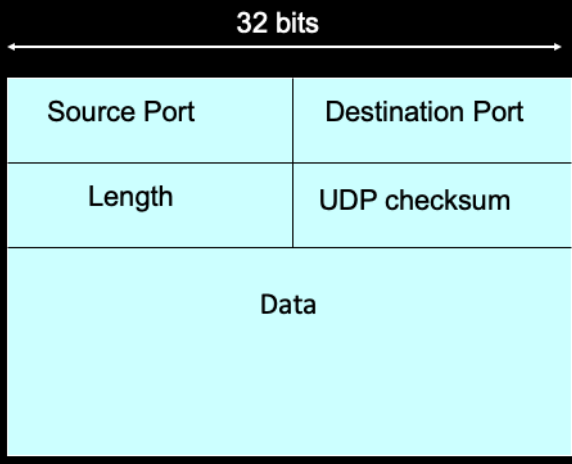
UCL

UDP - User Datagram Protocol

- Extremely simple transport protocol. Provides multiplexing/demultiplexing through UDP ports
- Does not provide reliability - no ACKs
- Does not provide congestion control
- Applications just send the data to the IP layer at the rare they want/can
  - There is no connection establishment and teardown -> connectionless service

UCL

UDP header

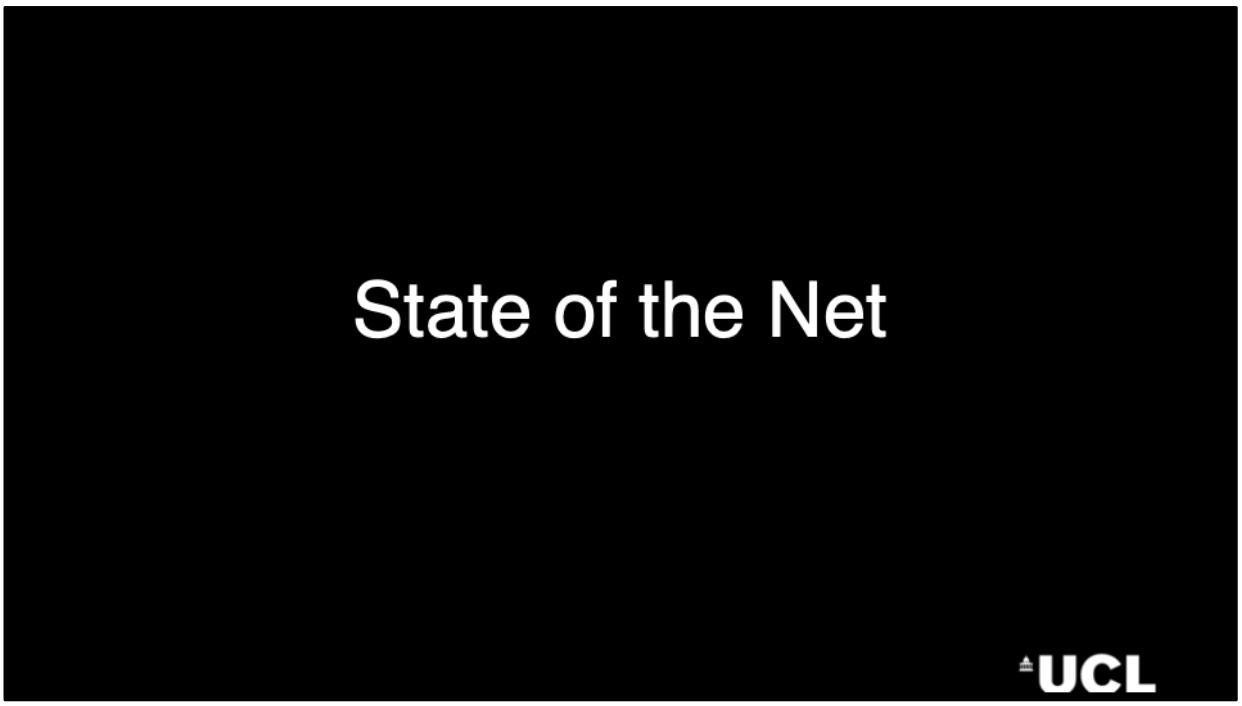


UCL

Summary: Transport

1. Multiplexing (TCP and UDP): Ports allow for end devices to know which applications to give data to
2. Reliability (TCP only) When packets get lost, they have to be retransmitted. This uses TCP sequence numbers
3. Congestion Control (TCP only): When packets get delivered, increase rate. Then they are dropped reduce rate

UCL


State of the Net
UCL

Internet Access Statistics

WORLD INTERNET USAGE AND POPULATION STATISTICS 2020 Year-Q2 Estimates						
World Regions	Population (2020 Est.)	Population % of World	Internet Users 30 June 2020	Penetration Rate (% Pop.)	Growth 2000-2020	Internet World %
Africa	1,340,598,447	17.2 %	566,138,772	42.2 %	12,441 %	11.7 %
Asia	4,294,516,659	55.1 %	2,525,033,874	58.8 %	2,109 %	52.2 %
Europe	834,995,197	10.7 %	727,848,547	87.2 %	592 %	15.1 %
Latin America / Caribbean	654,287,232	8.4 %	467,817,332	71.5 %	2,489 %	9.7 %
Middle East	260,991,690	3.3 %	184,856,813	70.8 %	5,527 %	3.8 %
North America	368,869,647	4.7 %	332,908,868	90.3 %	208 %	6.9 %
Oceania / Australia	42,690,838	0.5 %	28,917,600	67.7 %	279 %	0.6 %
WORLD TOTAL	7,796,949,710	100.0 %	4,833,521,806	62.0 %	1,239 %	100.0 %

http://www.internetworldstats.com/stats.htm

UCL

IPv6 adoption

UCL

Video Traffic

- Already 60% of Internet traffic is video.
  - Youtube: 15% (of total)
  - Netflix: 12%
- Expected to grow to 80% by 2022
- Social Networking 10%
- WWW: 8%
- File sharing 5%

UCL

By 2023…

- More than two-thirds of the global population will have Internet access
- The number of devices connected to IP networks will be more than three times the global population
- M2M (machine-to-machine) connections will be half of the global connected devices and connections by 2023
- Over 70 percent of the global population will have mobile connectivity

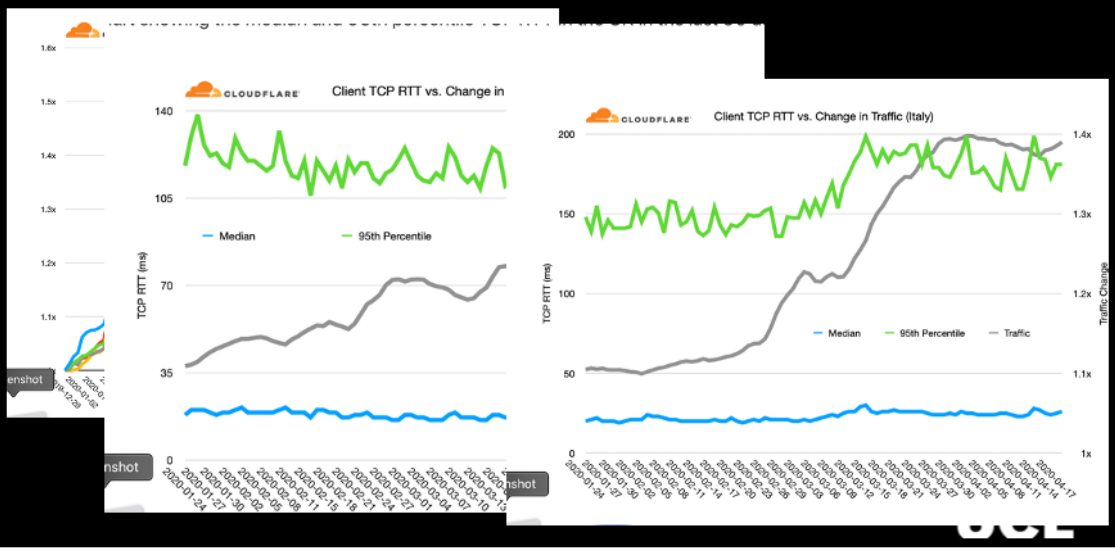
UCL

2023:Wi-Fi hotpots

Source: Maravedis, Cisco Annual Internet Report, 2018–2023

UCL

Covid 19 and the Internet



New Applications

UCL

Virtual Reality

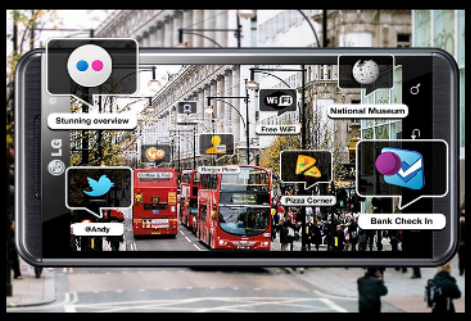
- Revolution in Healthcare, Education and Entertainment
- Interactivity will demand very short latency ≈10ms

89

UCL

Augmented reality

- Image processing for precise superimposition
- Delay is even more crucial



UCL

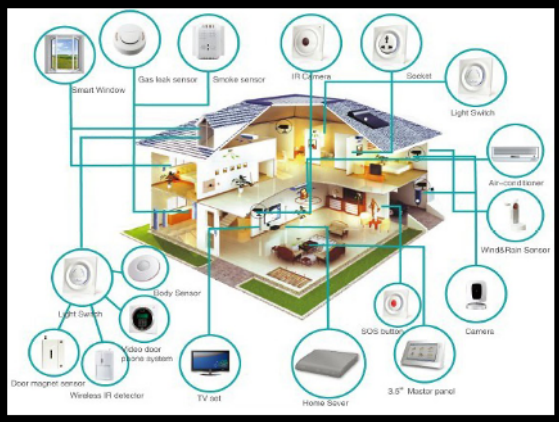
Holograms

- 3 orders of magnitude more data
- Interactivity
- Use cases in Education, Health, Entertainment
- Telepresence

UCL

The Smart Home

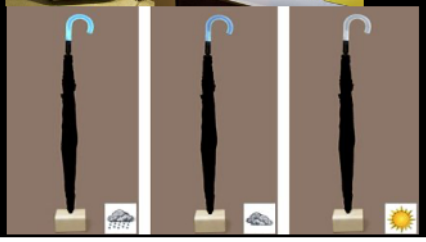
- Hundreds of data sensors, microphones, actuators
- Holograms/Virtual Reality for telepresence
- Privacy concern



UCL

http://smarthomeenergy.co.uk/what-smart-home

Internet of Things and Machine2machine



UCL

Summary

- Number of users and devices set to grow
- Video to become the dominant type of traffic
- Lot of exciting applications that are going to change society.

UCL

The End

- https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html
- https://www.sandvine.com/hubfs/Sandvine_Redesign_2019/Downloads/2020/Phenomena/COVID%20Internet%20Phenomena%20Report%2020200507.pdf
- Cloudflare blog

UCL