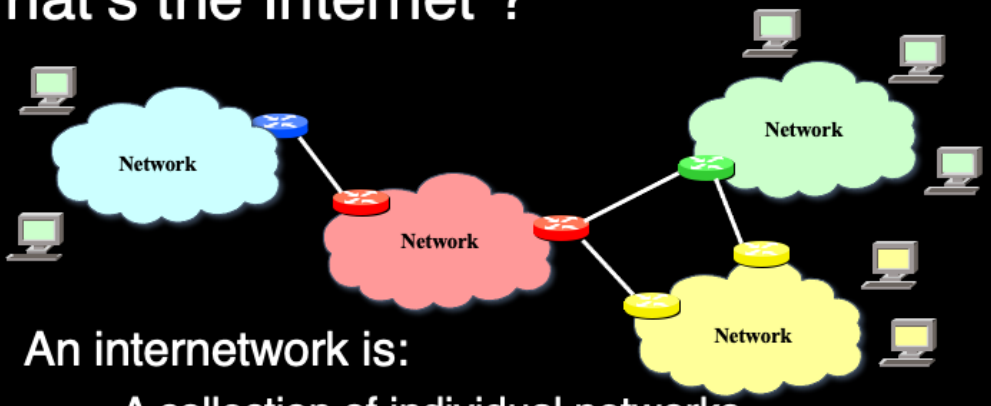

TCP/IP
Miguel Rio
UCL

From the LAN to the WAN

- Until now we learned how computers in the same local network communicate.
- It is very easy for a computer to know the address of every other computer in its network.
- It is also easy to know how to reach every other computer in the network
- In the Internet/WAN everything gets much more complicated

What's the Internet ?

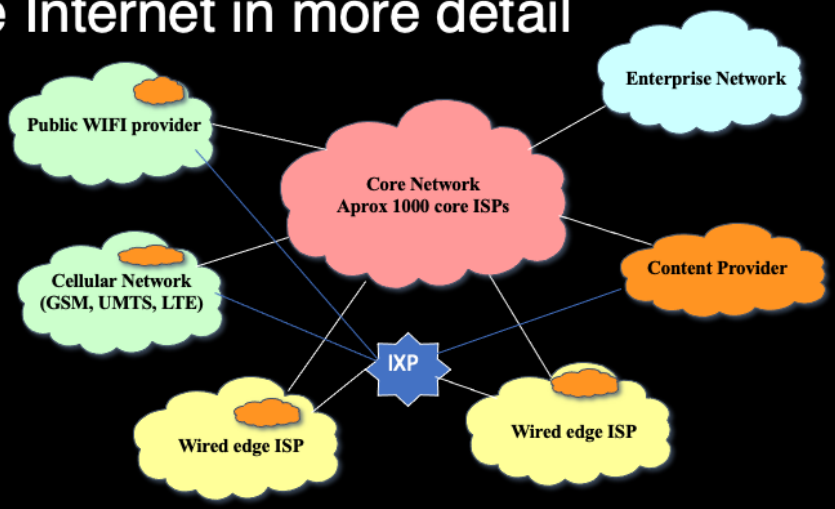


An internetwork is:

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

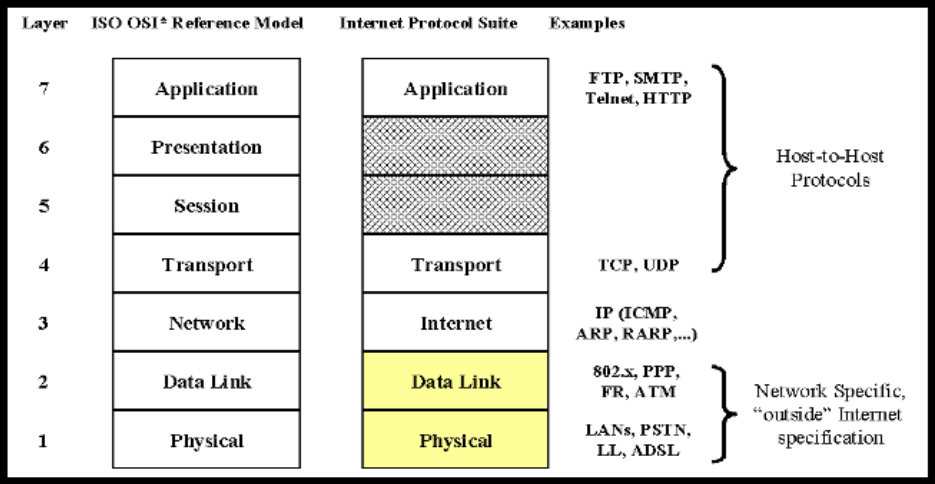
UCL

The Internet in more detail



UCL

ISO/OSI & Internet Protocol suites



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

UCL

A Service description of the Internet

- The Internet allows Distributed Applications running on it end systems to exchange data with each other
- It provides two services to 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

UCL

The Router

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

UCL

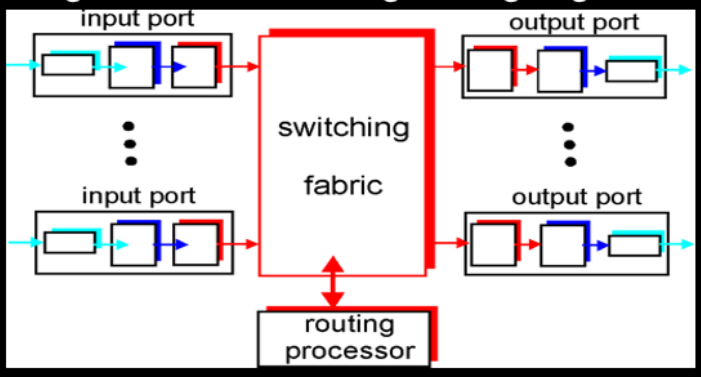
Forwarding and Routing

- Path determination: route taken by packets from source to destination. Routing algorithms
- Forwarding: move packets from router's input to appropriate router output

UCL

Inside the Router

- Two key router functions:
- run routing algorithms/protocol (RIP, OSPF, BGP)
- switching datagrams from incoming to outgoing link



UCL

Summary

- The Internet
- Types of providers
- The TCP/IP stack
- The router

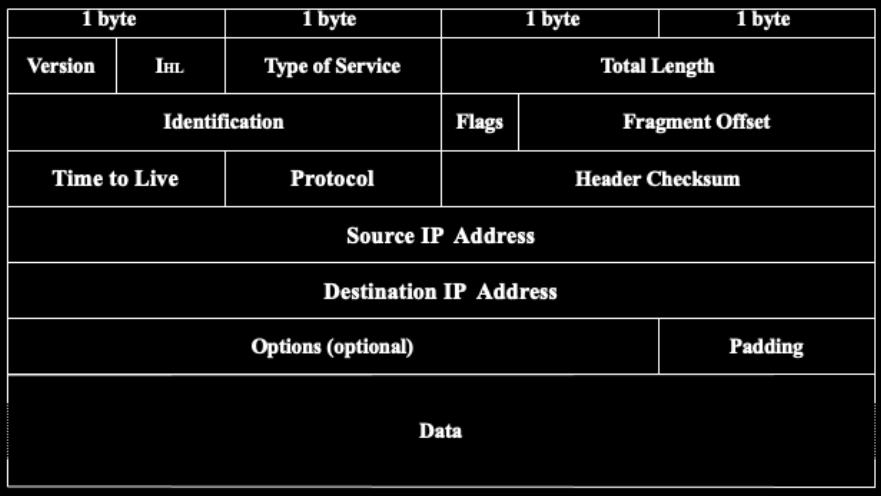

Addressing
Miguel Rio
UCL

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

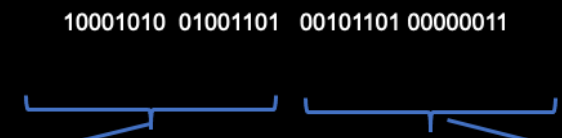
- This have two main functions
  - Uniquely identify a computer in a given internet
  - Provide information so that routers deliver the packet to the correct destination
- They have 32 bits (4 bytes) and are represented by dot notation. E.g:
- 138.77.45.3

10001010 01001101 00101101 00000011

UCL

Network and host part

- Addresses will have two parts:



The left part will identify a particular network/LAN

The right part will identify the host inside that network

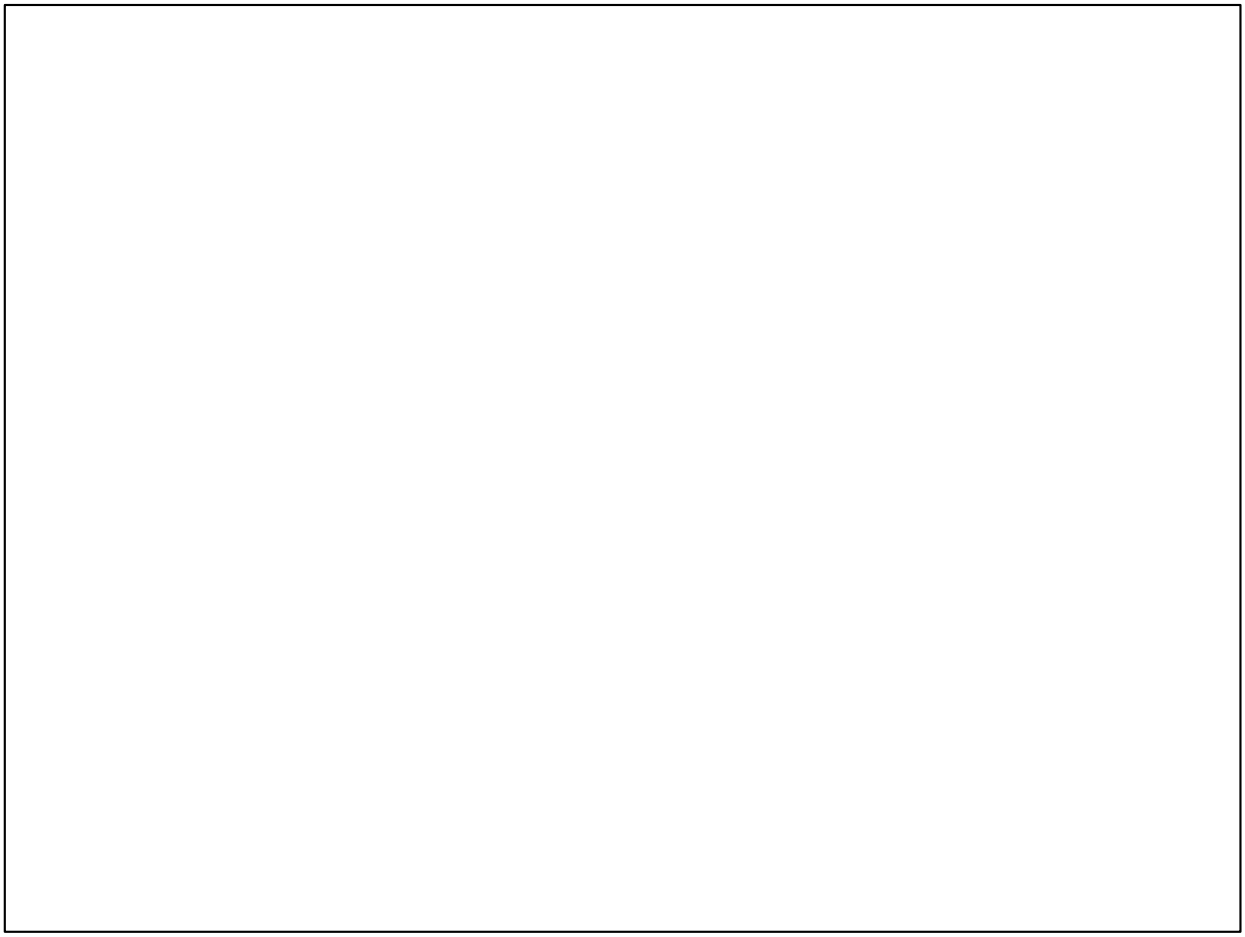
Important: The amount of bits for each part will vary

UCL

Special IPv4 Addresses

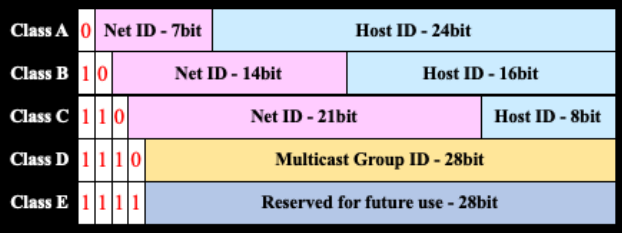
- All 0 host suffix => Network Address
  - 128.10.0.0 is a network with possible hosts like: 128.10.3.4, 128.10.0.3
- All 0s network => This network
  - (e.g. 0.0.0.2, host 2 on this network)
- All 1s host suffix => broadcast to all host on the same subnet
- Public IP address are controlled by the Internic
- 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



Classfull addressing

Classically address space is divided in 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

Host ID of all 0's indicate Network ID

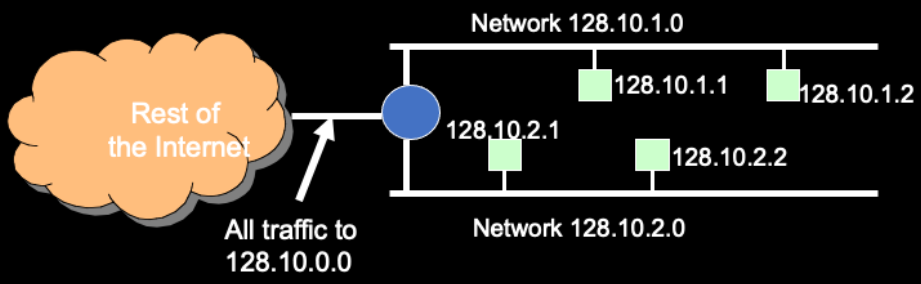
Host ID of all 1's indicate broadcast to Network ID

In this scheme each address is said to be self-identifying, because the boundary between prefix and suffix can be computed from the address alone

UCL

Subnetting

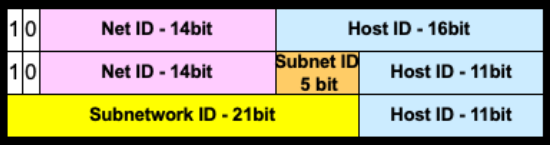
- Sometimes organisations may wish to partition their network. This can be done with subnetting



UCL

Classless and Subnet Address Extensions (CIDR)

- Class scheme is too rigid and many address can be wasted, subnetting permits to split class A, B or C in smaller networks
- 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



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

UCL

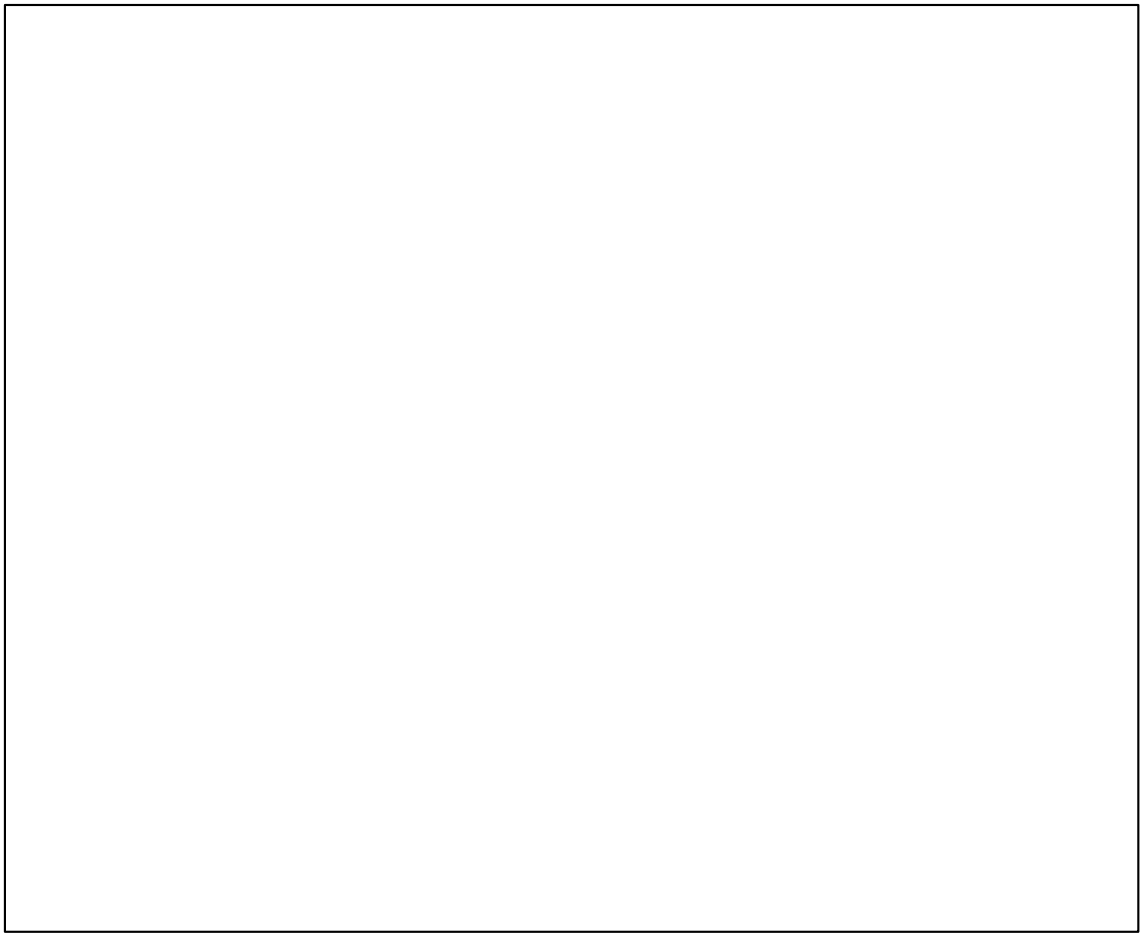
IP address netmasks

- Bit mask for the network part of the address:
  - e.g. 255.0.0.0/8, 255.255.240.0/20, etc.
- For example: 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

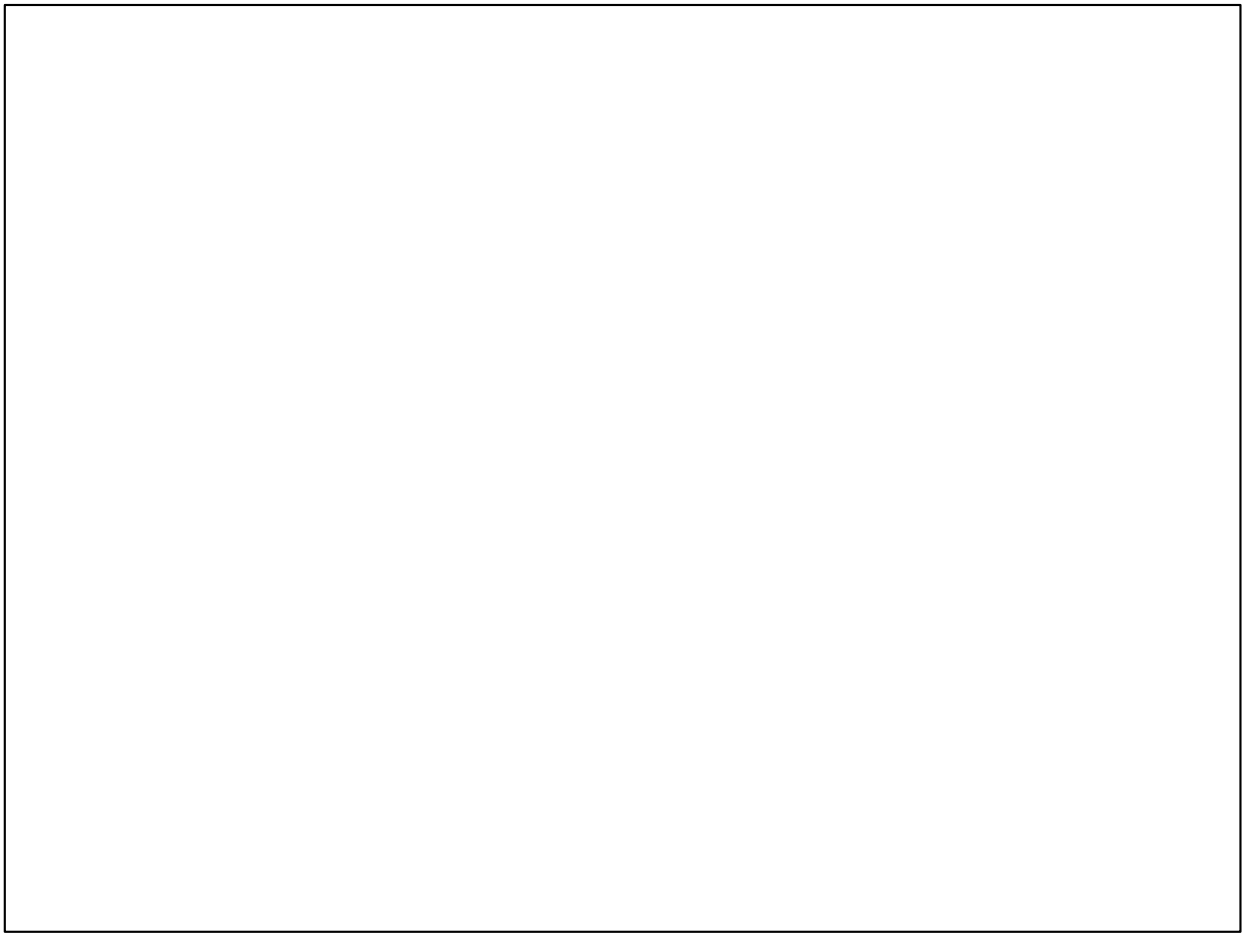
UCL



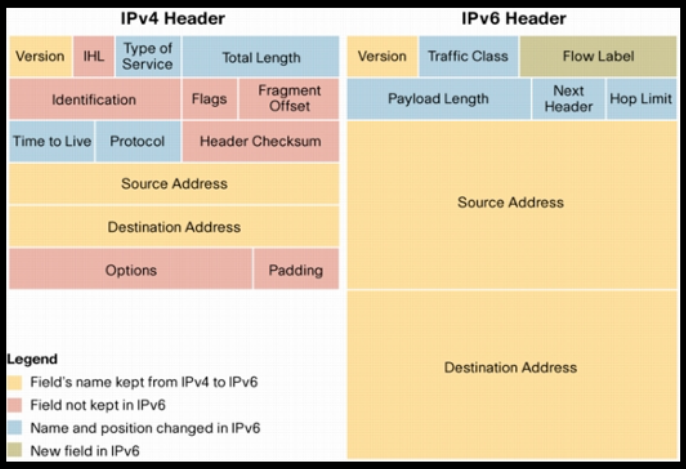
Problems with IPv4

- 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



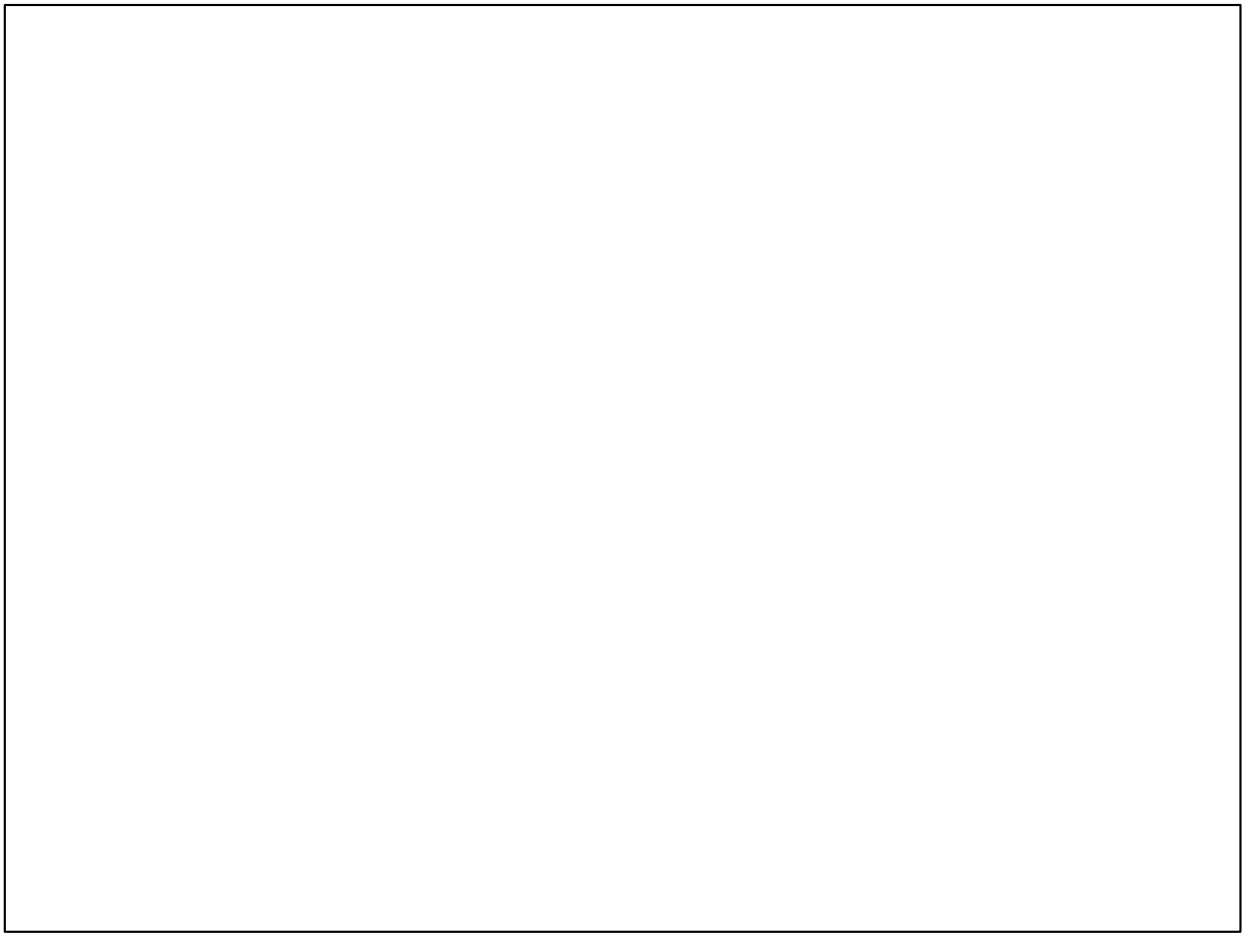
IPv6


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

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

- IPv4 addresses are written as 4 groups of 3 decimal digits separated by '.' 128.40.42.82
- 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

IPv6 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 describes 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)

UCL

IPv6 Address Blocks

- Multicast
  - Equivalent to addresses from the IPv4 block: 224.0.0.0/3
  - 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
  - 64bit Interface ID appended to make an address
  - Interface ID constructed from interface MAC address
  - This address cannot be globally routed

IPv6 Address Blocks

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 FD00::/8 in practice
- 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.

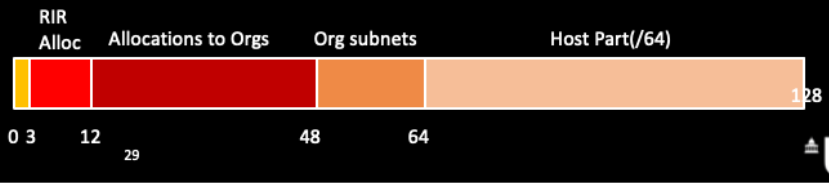
IPv6 Address Blocks

- Global unicast
  - Equivalent to a standard IPv4 address such as 128.40.42.82
  - 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 IPv6 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

UCL

IPv6 Address allocation Policy

- The IPv6 address 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 allow for 245 networks (/48 prefix, ~35x1012).
- Compare IPv4 with 2.2x106 networks.



UCL

IPv6 Addressing in practice

- Recommendation is to use 64/64 scheme:
  - 64 bits for network part
  - 64 bits for host part
- Host part derived using EUI64
  - Uses the 48bit layer 2 MAC address (padded to 64bits)
- Means a subnet can have 224 hosts = ~16M
- Practically, subnets should never run out of addresses.

Summary

- The IP protocol
- IPv4 header
- IPv4 addresses
- Subnetting
- IPv6 header
- IPv6 addresses
- IPv6 address allocation


Multicasting
Miguel Rio
UCL

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 these 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.

UCL

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
- 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

Special Multicast addresses

- 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 = All 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.

UCL

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

UCL

Summary

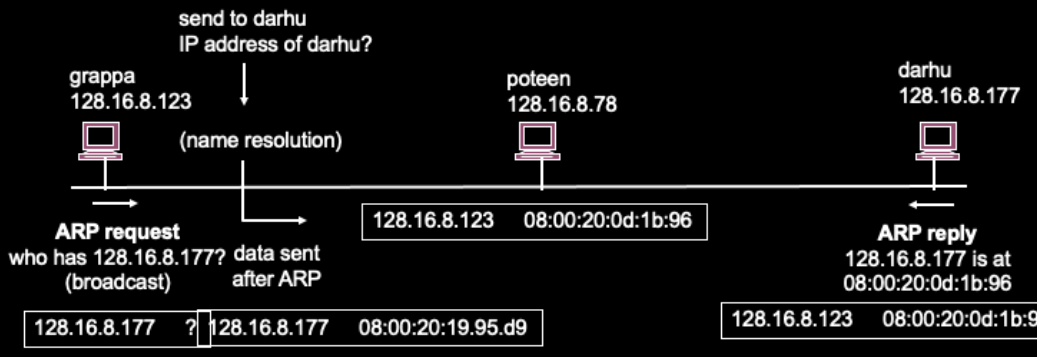
- Multicast in IPv4
- Special Multicast addresses in IPv6
- Mapping multicast addresses to Ethernet

Resolution and Auto-configuration

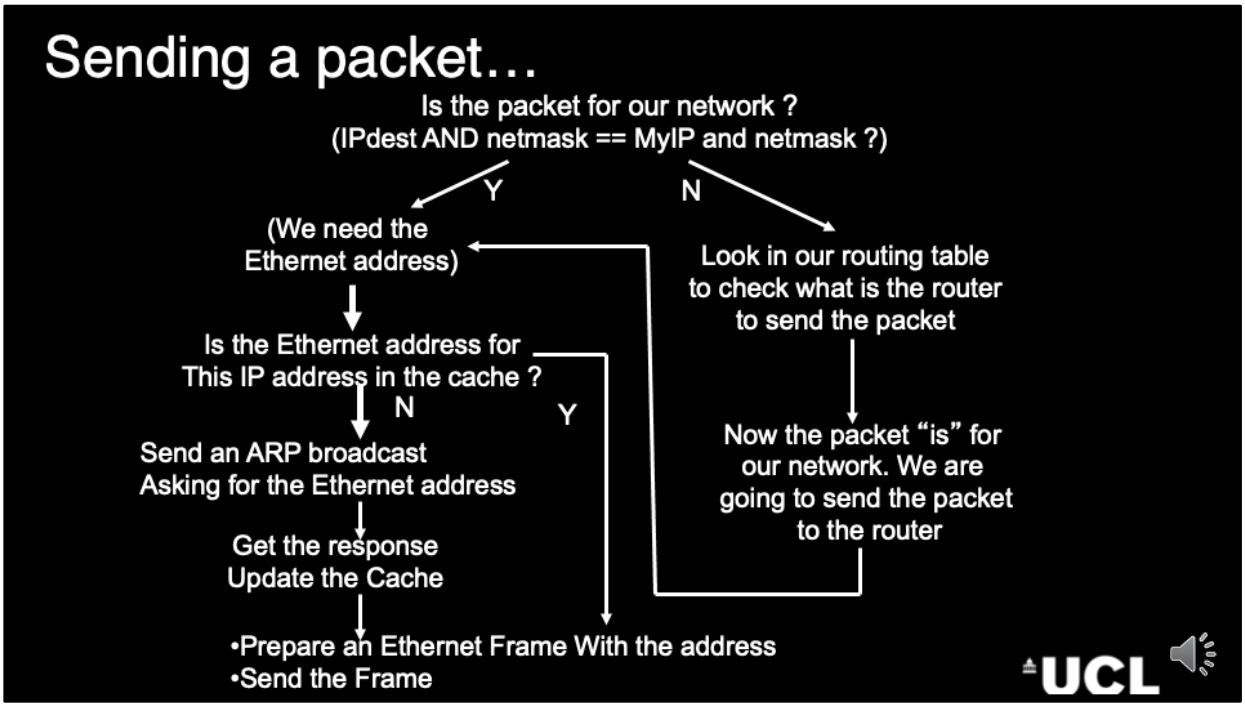
Miguel Rio

Address Resolution Protocol

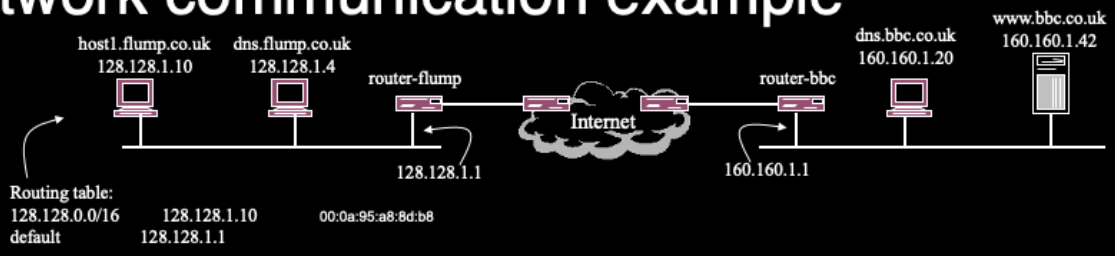
- Given an IP address what is the MAC address of machine? The solution:
  - The host broadcasts a request, What is the MAC address of 10.10.12.02?
  - The host whose IP address is 10.10.12.02 replies back, “The MAC address is ……”



Sending a packet…



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

Some Extra notes on ARP

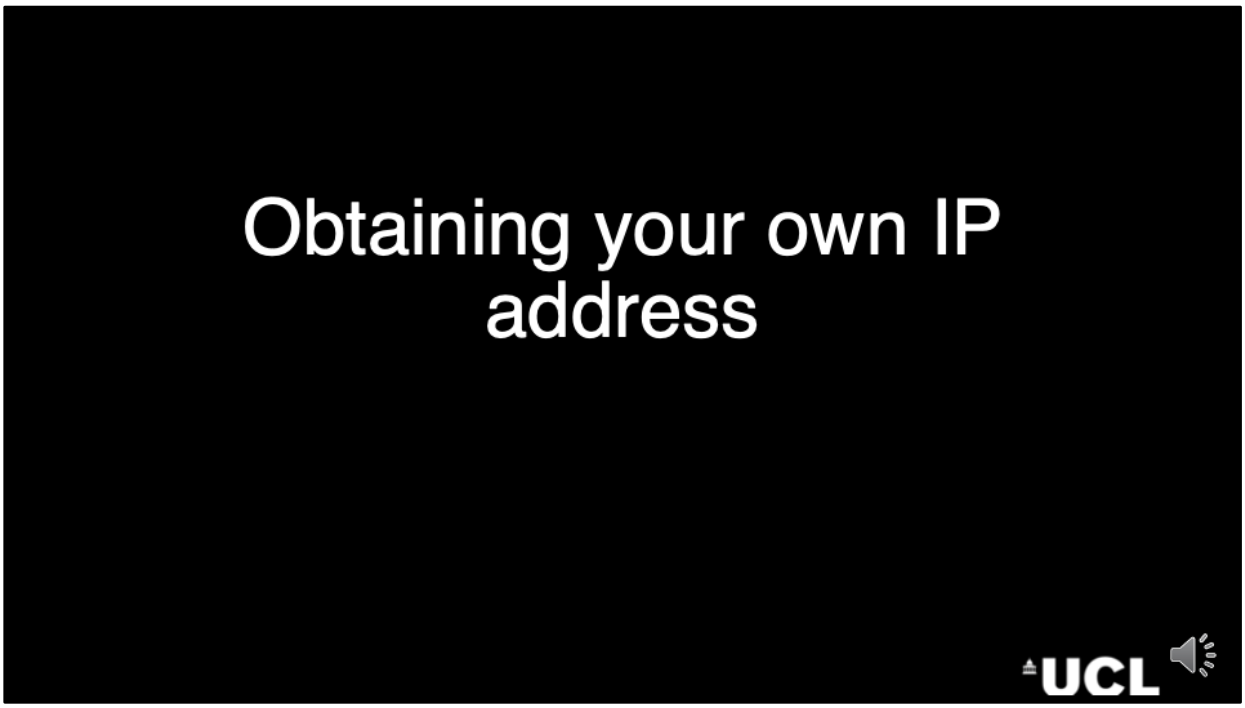
- ARP Cache timeout: typically 20 minutes
- Proxy ARP: Sometimes routers will reply “instead of the host” to “trick” the host into sending them the packets
- Gratuitous ARP: A “reply” sent by a host without a request. Typically sent by a host waking up.
- Remember: ARP is not just for Ethernet. It works for any layer 2 technology

UCL

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)

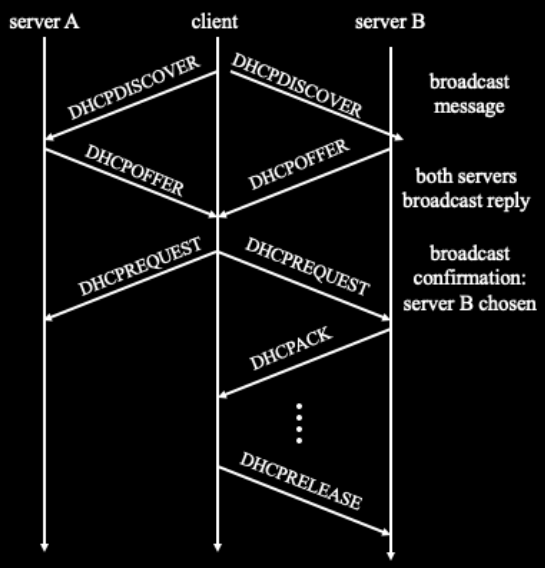
UCL


Obtaining your own IP address
UCL

Dynamic Host Configuration Protocol (DHCP)

- Sometimes we want an IP address allocated dynamically. For this we usually use DHCP
- Allow dynamic configuration:
  - automatically assigned address leasing
- Uses LAN broadcast
- Requires server(s):
  - central store of configuration information
- Useful for:
  - mobile hosts
  - large numbers of hosts (can use static/manual address assignment)
- Usually also returns:
  - default router
  - netmask
  - DNS server
- Usually sent using UDP port 64

(destination 255.255.255.255, source 0.0.0.0)



UCL

IPv6 Host Autoconfiguration (SLAAC)

- SLAAC = StateLess Address AutoConfiguration
- How a host gets its addresses:
- As a host boots, it will bind each network interface to the following addresses:
  - FF02::1 all nodes multicast address
  - ::1 Loopback address
  - FF02::1:ffxx:xxxx node solicitation multicast address
    - xx:xxxx = last 24 bits of MAC address (usually)
  - FE80::Interface ID
    - This is marked as a “Tentative address” and cannot be used as the source address for any datagrams yet.

UCL

Host Autoconfiguration

- At this point, a host has a usable IPv6 address.
- However, this address is link-local (FE80 prefix)
  - It cannot be used to talk to anything not on the same LAN.
- The next stage is to discover any routers on the network
- A router solicitation message is sent to the all-routers multicast address FF02::2
- All routers on the link reply with a router advertisement
  - Each advertisement 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.

Host Autoconfiguration: DAD

- DAD = Duplicate Address Detection
  - Uses a multicast mechanism
- The host constructs a ICMP Neighbour Discovery packet which will be sent to the Node solicitation multicast address corresponding to the last 24bits 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 therefore usable.

UCL

Privacy Extensions – RFC 4941

- 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.
- Widely available in operating systems.

UCL

Summary

- Converting IP addresses to MAC addresses with ARP
- Obtaining your own IP address:
  - DHCP
  - IPv6 auto-configuration

UCL

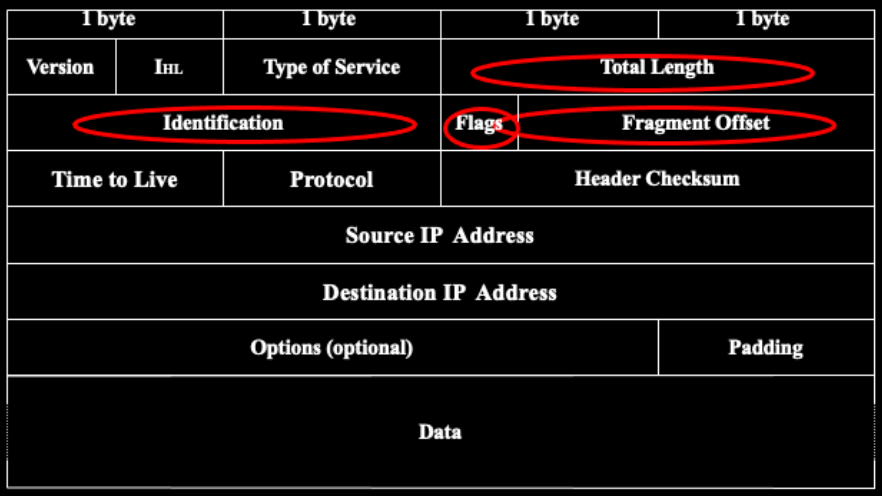

Fragmentation
Miguel Rio
UCL

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.

UCL

IP Version 4 Fragmentation



UCL

Fragmentation in IPv6

- As we saw IPv6 routers 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 drops the packet and sends back a ICMP message saying “packet too big”
- The process continues until the packet reached the destination

UCL

Summary

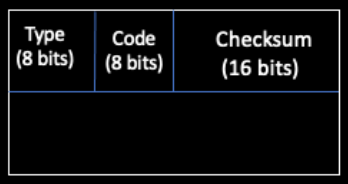
- Fragmentation in IPv4
- Fragmentation in IPv6

UCL


ICMP
Miguel Rio
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	source 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

ICMPv6

- In IPv6 ICMP is responsible 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:

UCL

ICMP based tools

- Several tools developed using ICMP messages
  - PING
    - Used to see whether a specified IP address is reachable. Tool is available in Microsoft Windows operating system and UNIX platforms
  - TRACEROUTE
    - Send a packet with time to live = 1
    - The first router discards the packet and send an 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

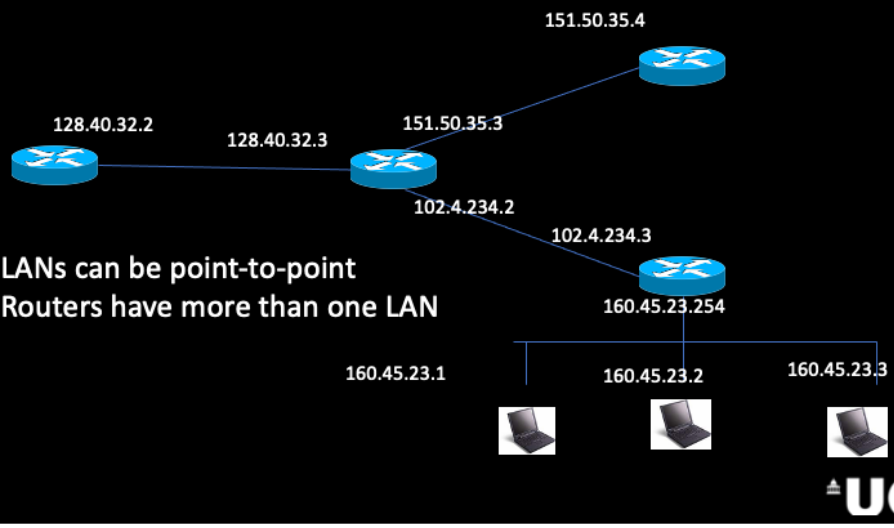
UCL

Summary

- ICMP for IPv4 and IPv6
- Types of messages
- ICMP based tools


NAT
Miguel Rio
UCL

Once the Network is created: example



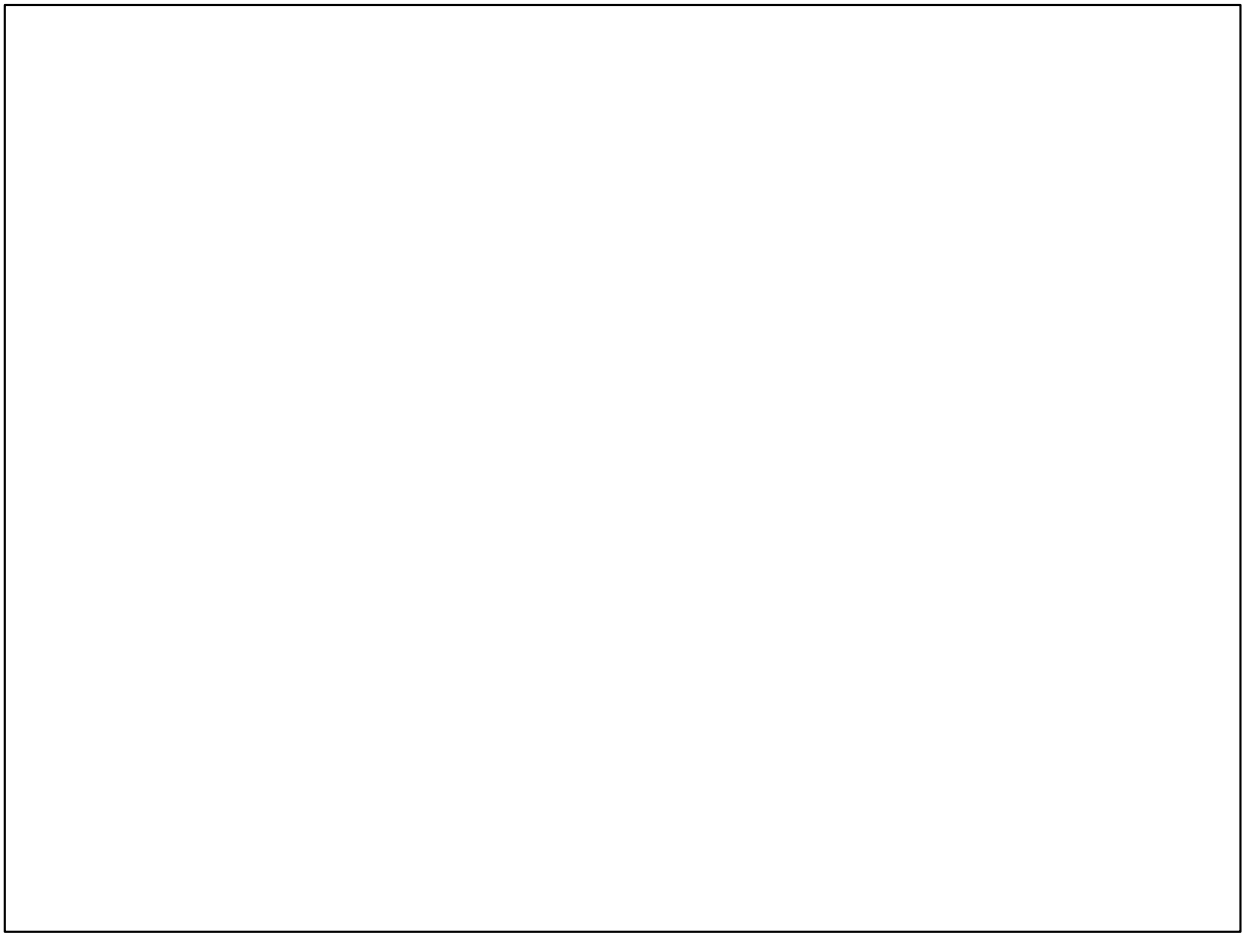
- LANs can be point-to-point
- Routers have more than one LAN

UCL

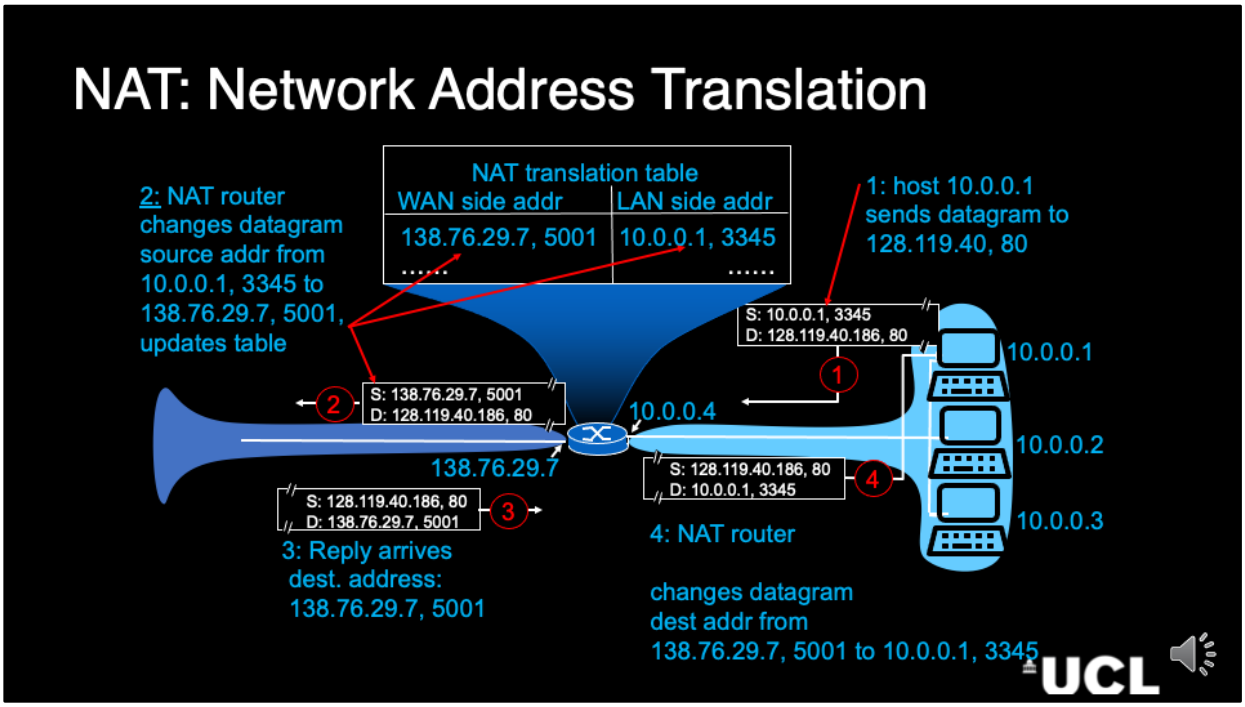
Network address translation (NAT)

- Sometimes the organisation/ISP/home has many more computers than IP addresses
- NAT is Used to convert 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

UCL



NAT: Network Address Translation



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
- Similarly for ICMP.

UCL

NAT traversal

- Originating connections TO a node behind a NAT is hard. There is no port to put in the TCP SYN packet
- Many techniques: STUN, TURN, uPnP
- Generally involve registering with a node with a public IP address which will then relay the connections.

UCL

Carrier Grade NAT

- NAT can be done inside the network. Many networks behind the NAT box
- Sometimes thousands of users
- Problem: Port exhaustion. Some applications use lots of TCP connections

UCL

Summary

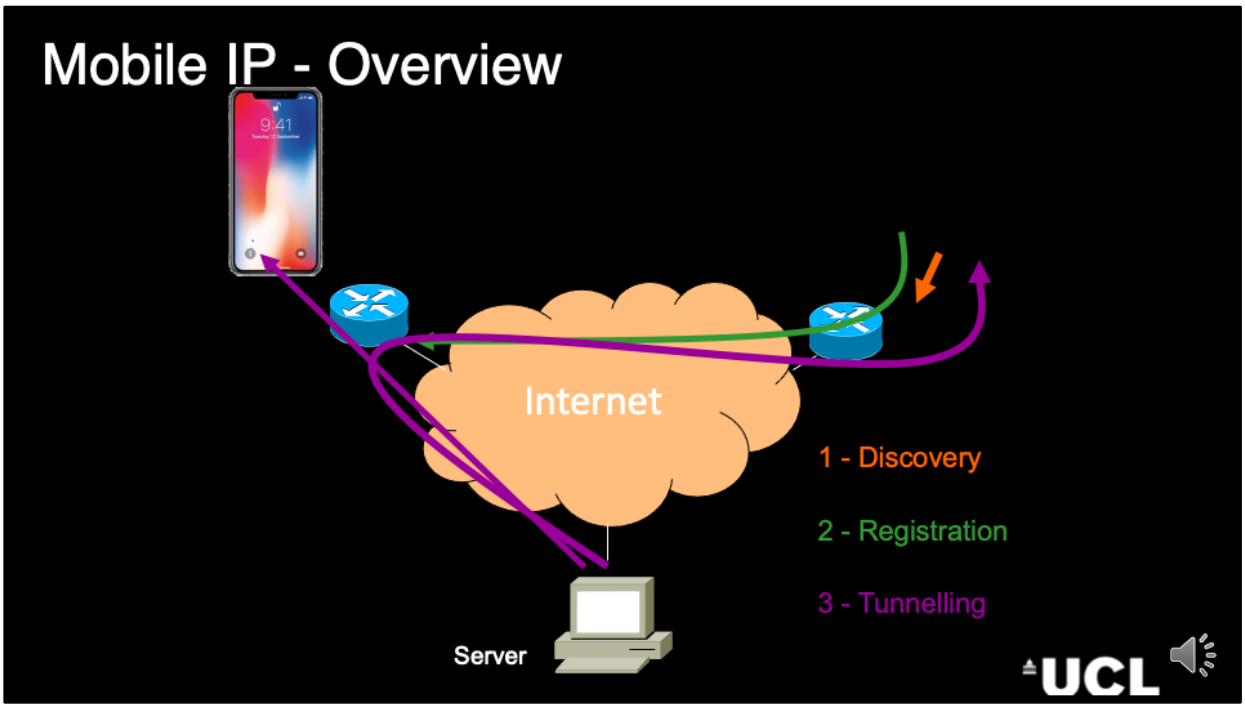
- NAT allows for 1 IP address to be shared among several computers


Mobility
Miguel Rio
UCL

Mobile IP

- Nodes move to a different network
- But they want to still receive the packets destined to their “old” address
- Sending applications do not need to be aware of the movement
- Just one of the techniques for IP Mobility…`

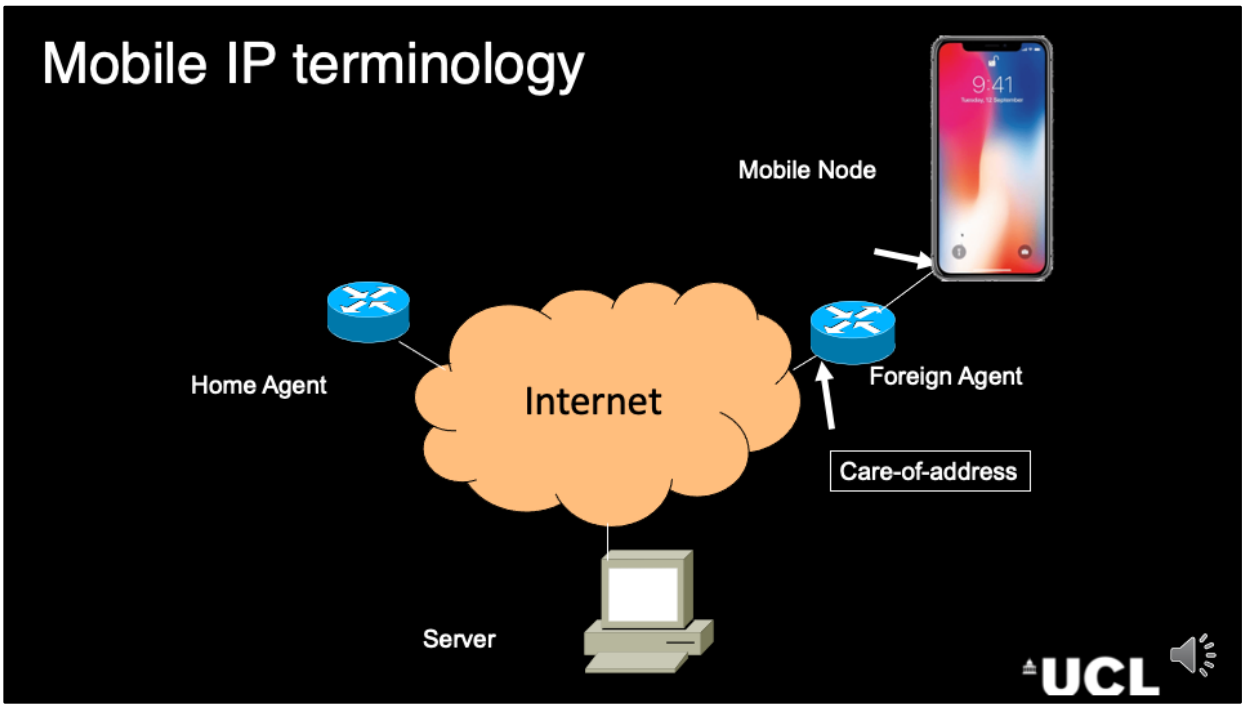
UCL


Mobile IP - Overview
Internet
1 - Discovery
2 - Registration
3 - Tunnelling
Server
UCL

Three Musketeers of Mobile IP

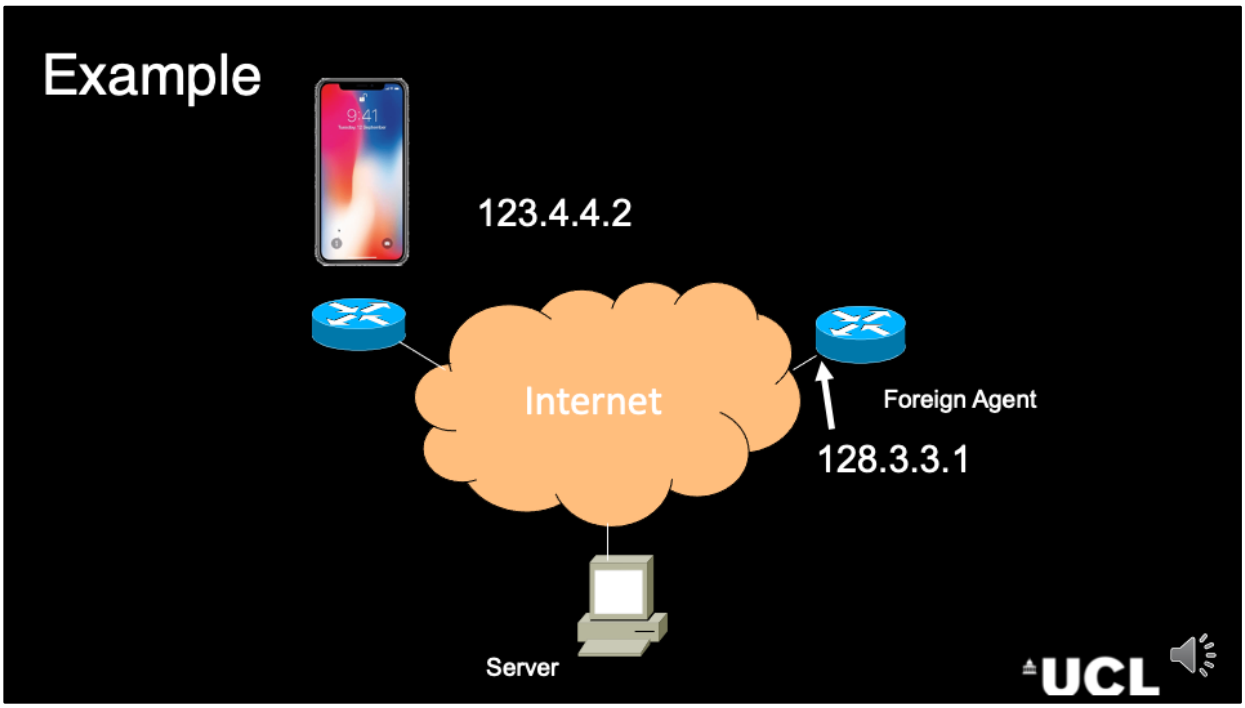
- How does the Mobile Node find out where it is?
  - Agent Discovery — ICMP Router Discovery
- How does the Mobile Node Inform its current location?
  - Registration—Authentication, location update and deregistration
- How are packets delivered?
  - Tunneling— IP in IP or GRE

UCL



Introduce terminology

Mention Collocated care-of-address



Introduce terminology

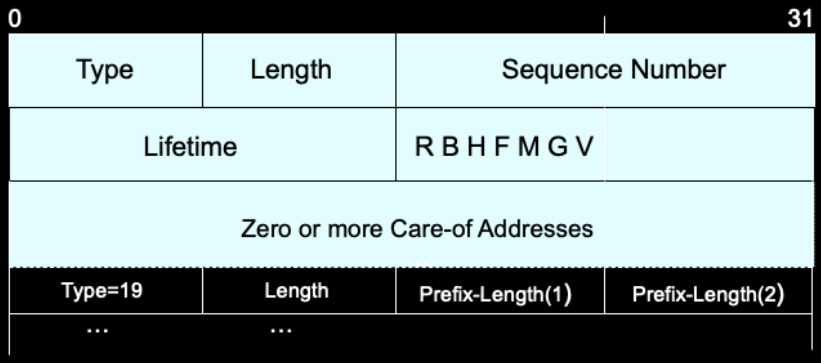
Mention Collocated care-of-address

ICMP router discovery….it is used not only for mobile ip but when a host does not know which router to use.

Routers pedriodically broadcast messages on their local networks showing they are available to route packets.

Agent Advertisement

- Routers propagate advertise packet in an ICMP packet



UCL

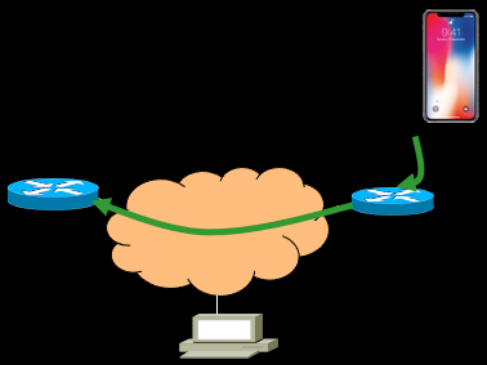
This is our first mobile IP packet. It is sent as an extension of an ICMP Router discovery packet.

It is 32 bits long. We see several fields like Type length and Sequence number and in the middle there a set of 7 flags.

The area below the dashed line is optional. As we will see this will be used to detect movement.

We'll now look at some of these fields in detail…

2 - Registering



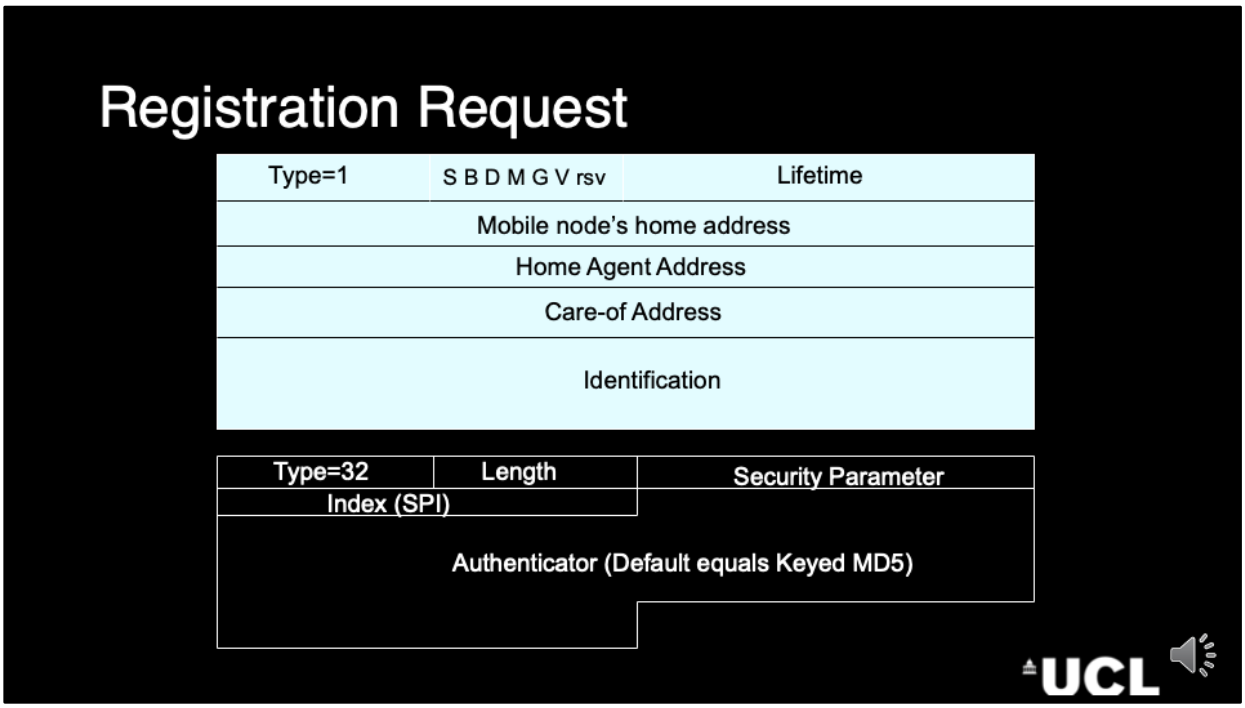
1 – The mobile sends a registration request to the foreign agent

2 – The foreign agent relays this request to the home agent

3 – The home agent accepts or denies the request

4 – The foreign agent relays this reply to the mobile node

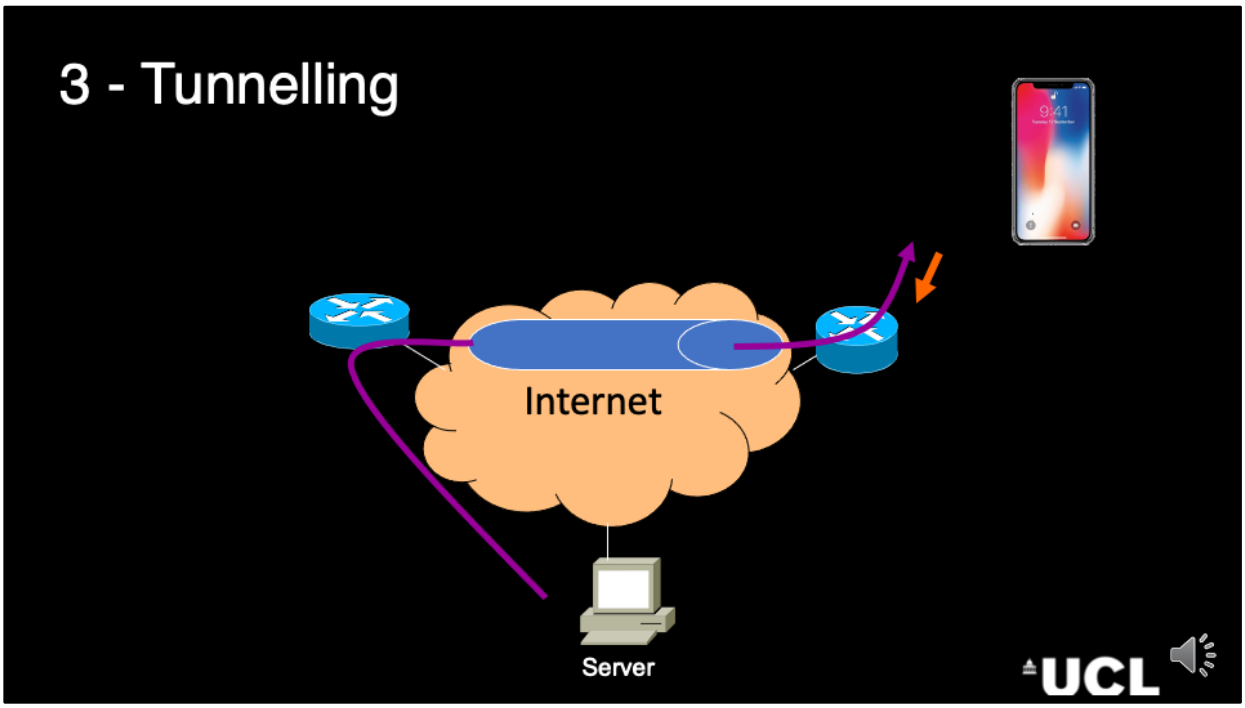
UCL



This is the message sent by the mobile node in a foreign network to its home agent. It tells what is its home address so that the home agent knows who is the node wishing to received tunnelled data.

The identification is an unique number to each request so that when a reply is received the mobile node knows which request was successful

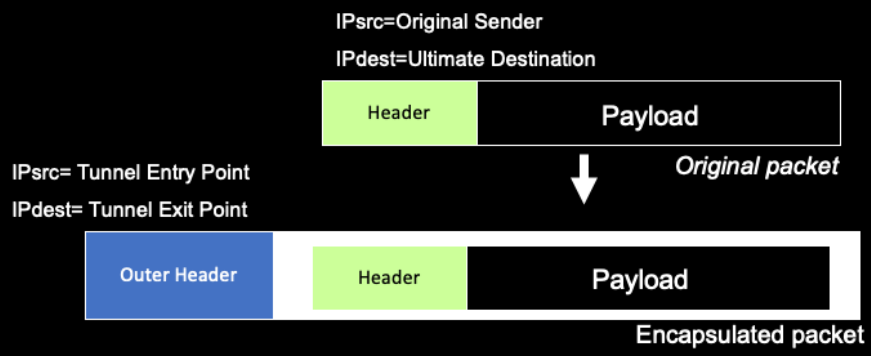
The part in white deals with security and authentication which we'll go into more depth later



Introduce Tunneling…

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

IP over IP

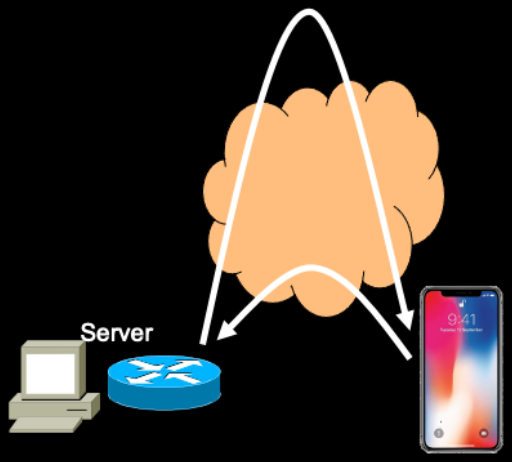


- Sender sends the packet to the destination unaware of any tunnelling/mobility

UCL

Explain this slide swithcing with previous one.

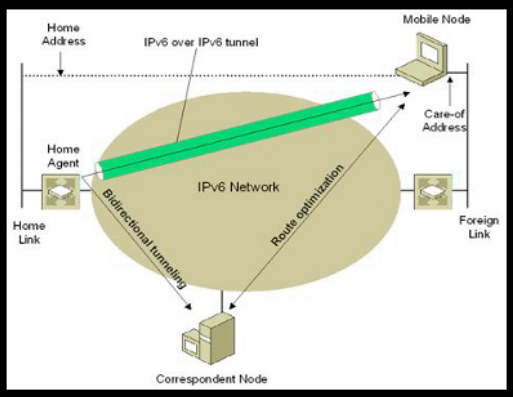
The Triangle problem



- Route not optimal
- Server and Mobile node could communicate directly but…
- Authentication would have to be made with potentially every node in the internet

UCL

Route Optimization



In IPv6 things are different when both nodes have IPv6 enabled

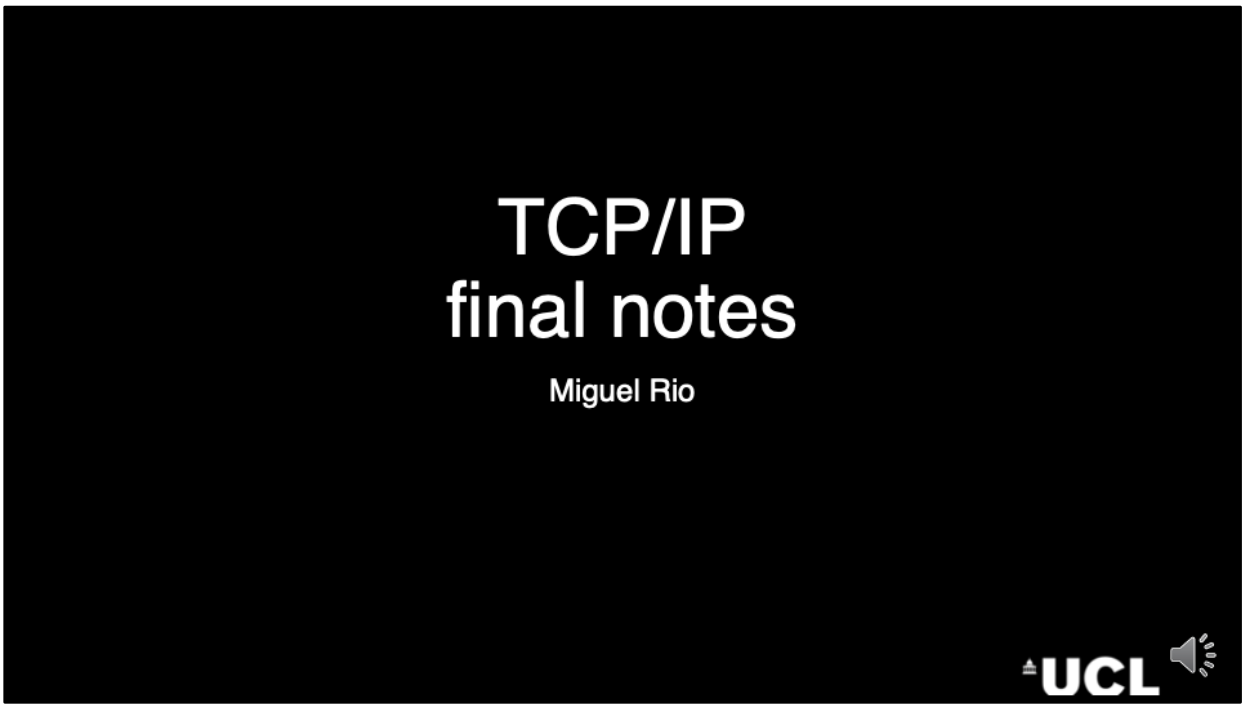
- The mobile node can send a IPv6 Destination options header message to the corresponding node
- Packets can then be sent directly

UCL

Summary

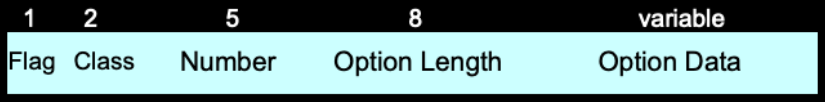
- Mobile IP
- The Three Musketeers
  - Discovery
  - Registration
  - Tunneling
- Mobile IPv6

UCL


TCP/IP
final notes
Miguel Rio
UCL

IPv4 Options

- The IP packet header contains a field (of variable size) that allows for extra options



- Defined options include:

[0,7] - Record Route

[2,4] - Time Stamp

[0,3] - Loose Source Route

[0,10] - Strict Source Route

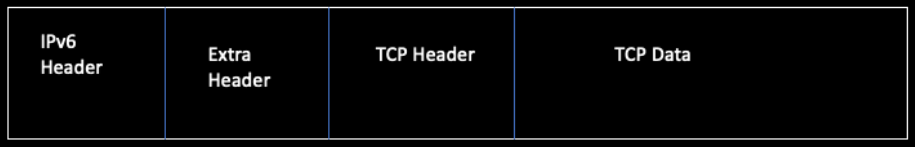
[0,20] - Router Alert

- Complete list at http://www.iana.org/assignments/ip-parameters

UCL

IPv6 Extension headers

- Similar to IPv4 options but there are more of them
- Includes a Routing header for source routing
- Fragment header indicates that the packet got fragmented by the source



UCL

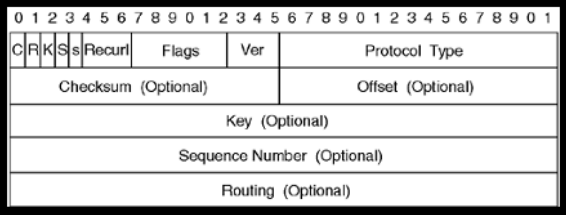
Tunneling

Tunneling refers to carrying lower level traffic in higher (or equal) layer. Examples:

- Ethernet over UDP
- IPv4 in an IPv6 packet

Several protocols:

- GRE [RFC 2784]
- L2TP[RFC 3931]
- Used for VPNs (Virtual Private Networks)



GRE Header

UCL

Transition From IPv4 To IPv6

- Not all routers can be upgraded simultaneous
  - no “flag days”
  - How will the network operate with mixed IPv4 and IPv6 routers?
- Two proposed approaches:
  - *Dual Stack*: some routers with dual stack (v6, v4) can “translate” between formats
  - *Tunneling:* IPv6 carried as payload in IPv4 datagram among IPv4 routers

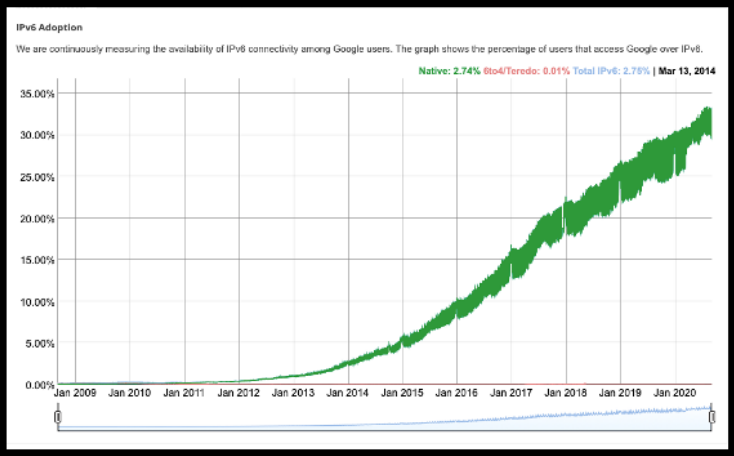
UCL

Discussion: IPv6 ? Where is it ?

- IPv4 addresses are over. They cost money now
- NATs and Carrier Grade NATs
- Internet of things and Smart Cities driving IPv6

UCL

IPv6 current status



- http://www.google.com/intl/en/ipv6/statistics.html
- Also check: http://6lab.cisco.com/index.php

UCL

Summary

- IP options
- IP tunneling
- State of IPv6