# Modules

**A. IP**
- A1TCP/IP
- A2 Encapsulation
- A3 Domain Names & DNS
- A4 IP Addresses
- A5 ARP & ARP Cache
- A6 DHCP

**B. Routers**
- B1 ICMP
- B2 Default Route & Netmask
- B3 Routing & Traceroute

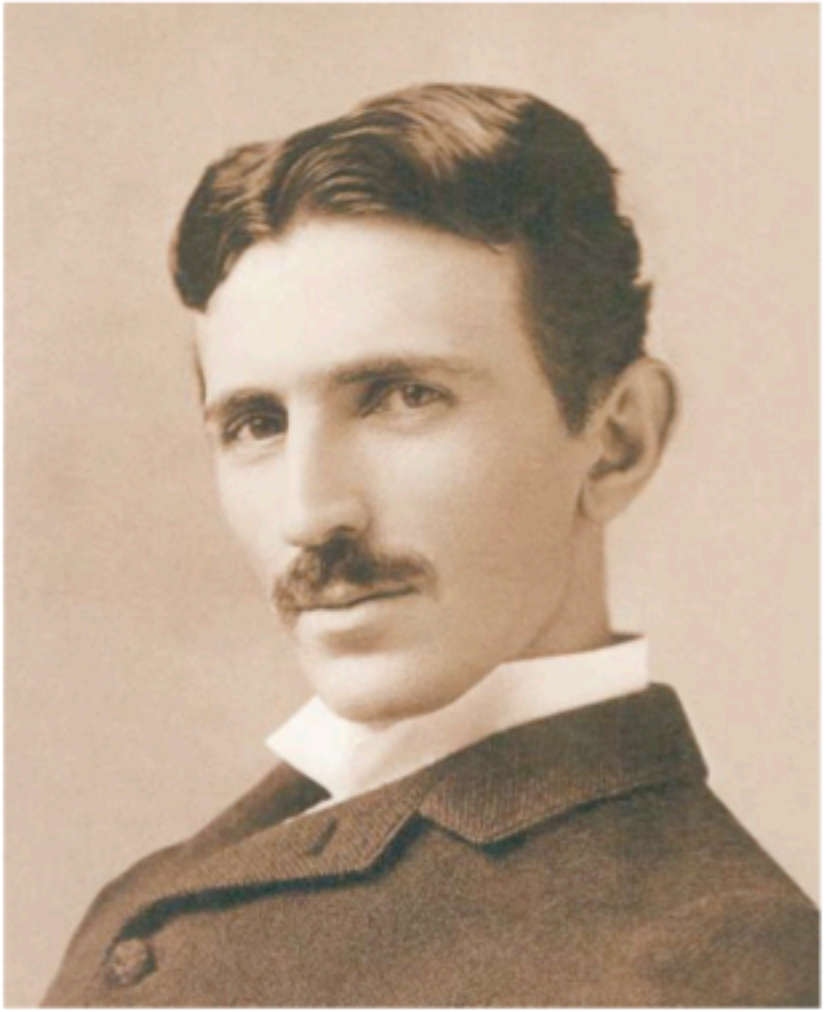
**C. UDP Transport**
- C1 UDP
- C2 Transport Ports
- C3 Decodes
- C4 IPv6
- C5 TCP Transport

# The TCP/IP Stack

G Fairhurst, http://www.erg.abdn.ac.uk

# Nikola Tesla, 1908

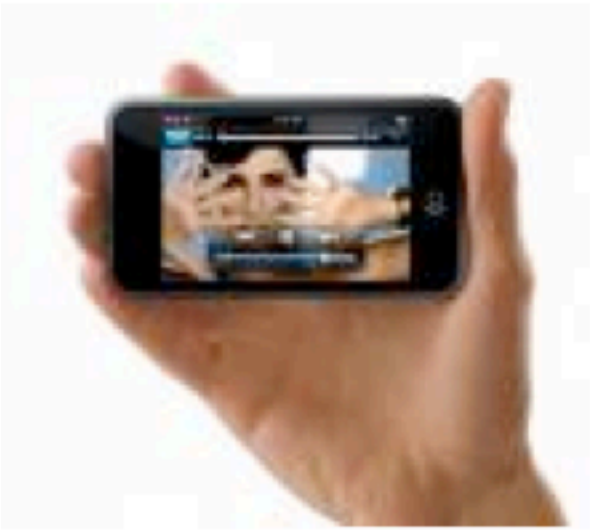


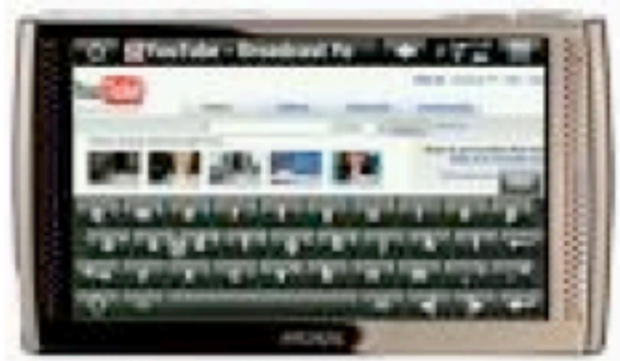


Envisioned a technology that would allow "... a business man in New York to dictate instructions and have them instantly appear in type at his office in London or elsewhere" and allow global access to "any picture, character, drawing, or print."

# People expect Internet connectivity

G Fairhurst, http://www.erg.abdn.ac.uk





*"By the year 2016, no one under the age of forty will remember a world without personal computer. The average twenty year old will find it hard to imagine a time when there wasn't any email to check or Web sites to visit."*

*– George Christian, 2006.*

# The Internet Message Processor



1968 IMP specification
Messages < 8KB divided into packet <1024B
Honeywell DDP-516 (12 kB memory)
50 kbps

1969 Contract won by BBN
First IMP delivered (Sept 1969)

1970 ALOHAnet

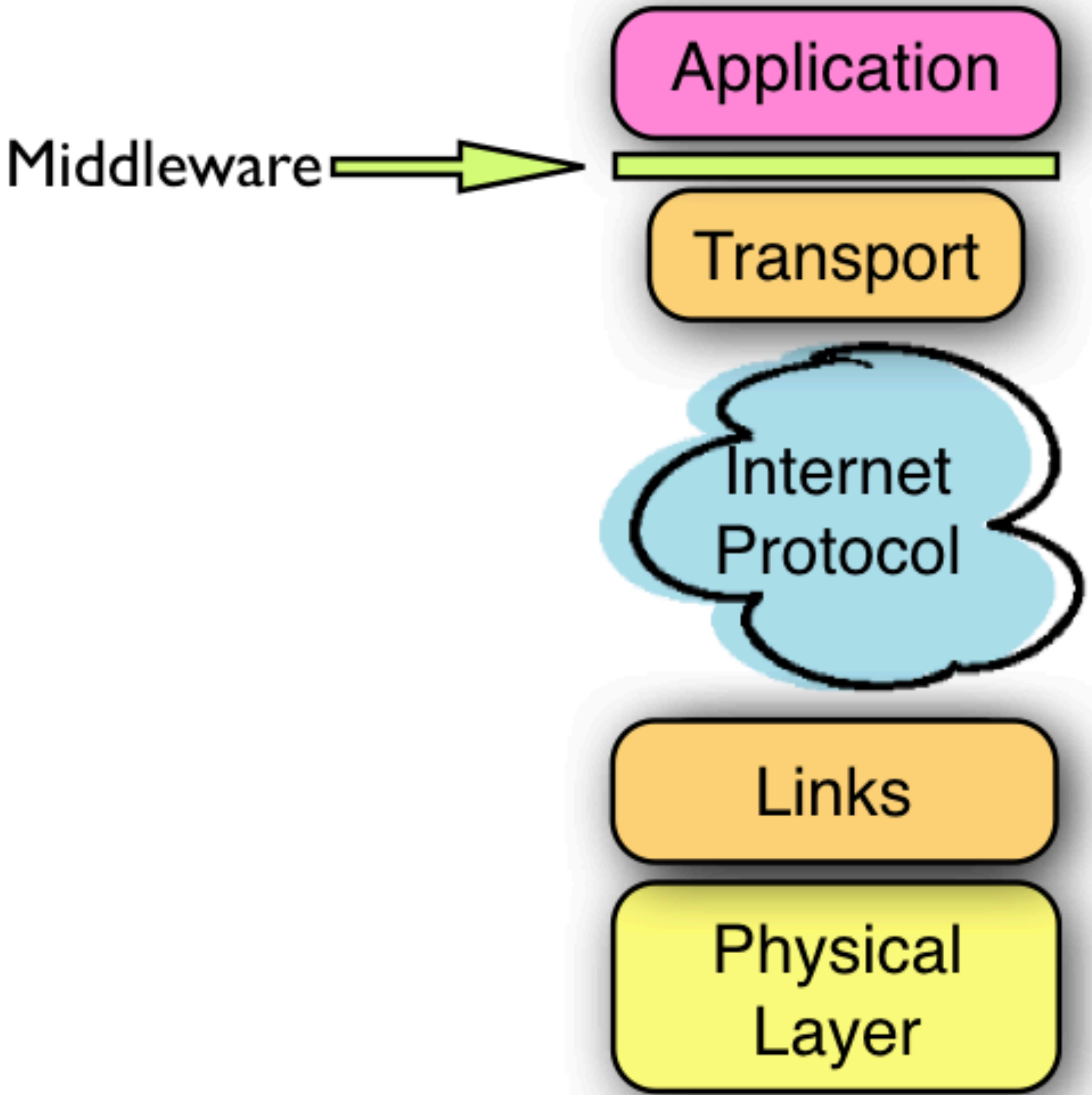
1971 Added email

1980 Layered design of TCP/IP (IPv4)

1998 IPv6

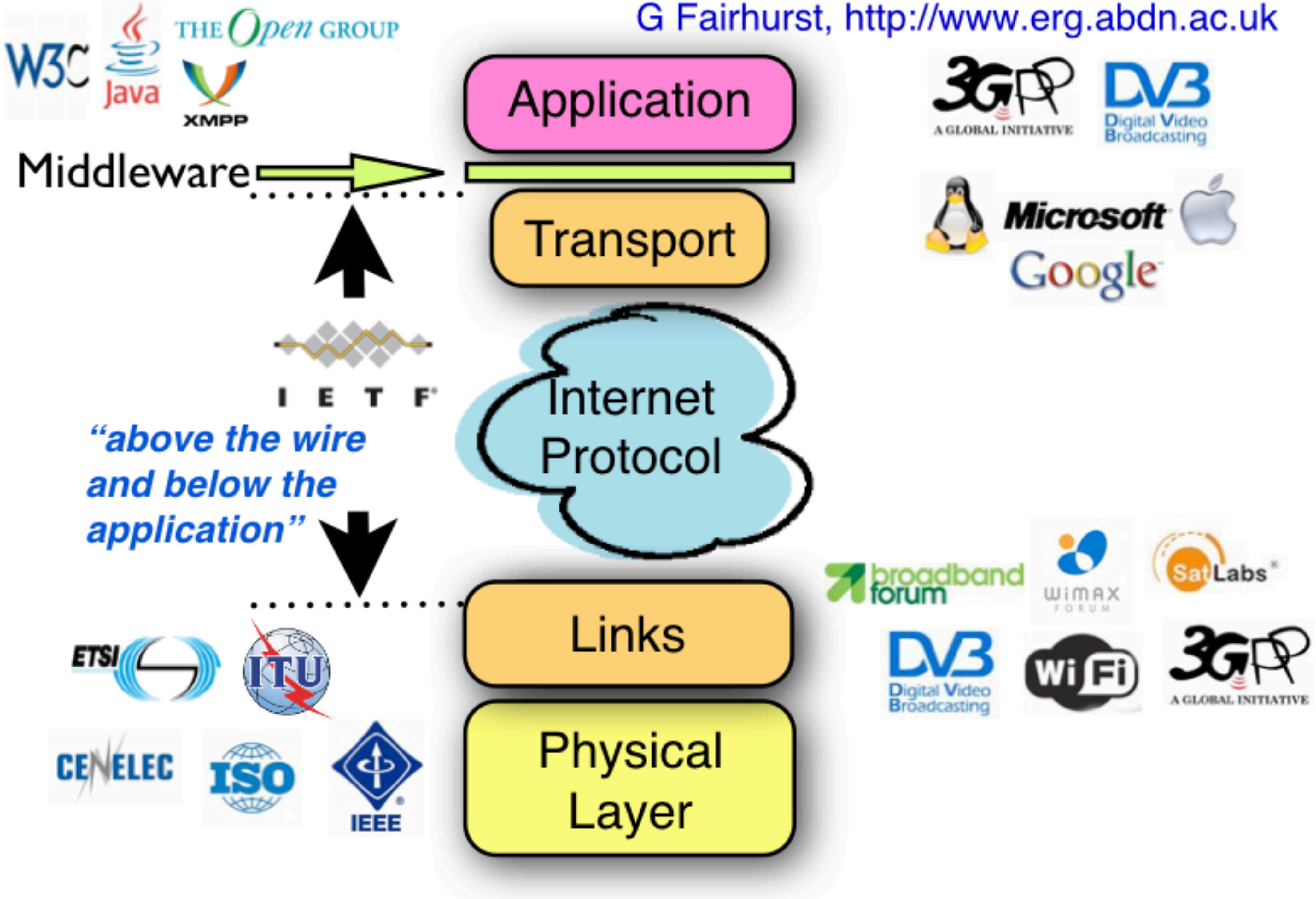
# Internet Protocol Stack

G Fairhurst, http://www.erg.abdn.ac.uk



# Some Internet Players

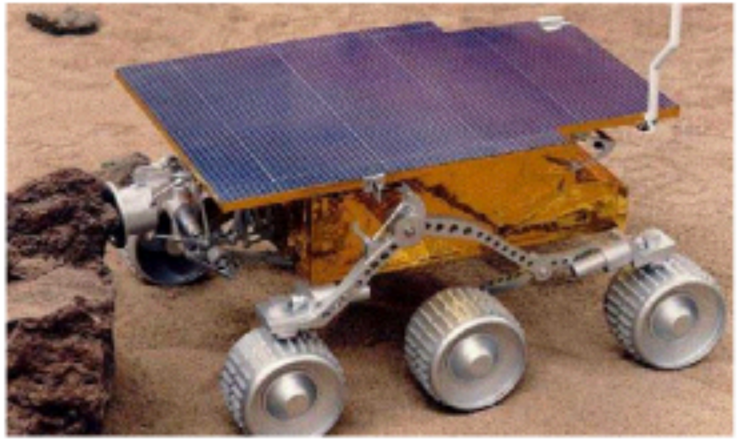
G Fairhurst, http://www.erg.abdn.ac.uk



# IP Appliances

G Fairhurst, http://www.erg.abdn.ac.uk










***"IP on everything"***
***"Everything on IP"***

G Fairhurst, http://www.erg.abdn.ac.uk

# Internet Protocol

- The Connection-Less Network Service
- The 20 byte IP Packet Header
- IP Network Layer Addresses
- Name Resolution (name to IP Address)

# IP Packets

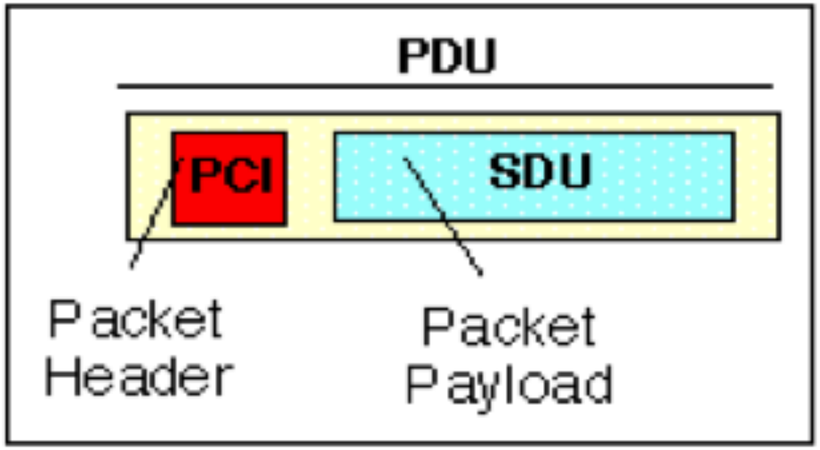
G Fairhurst, http://www.erg.abdn.ac.uk

Messages (large blocks of data)
are split into smaller pieces, called “Packets”

Each packet (also known as a PDU) has:

**A header (also known as the PCI)**
- Well-defined format
- Destination address , source address, type, ...

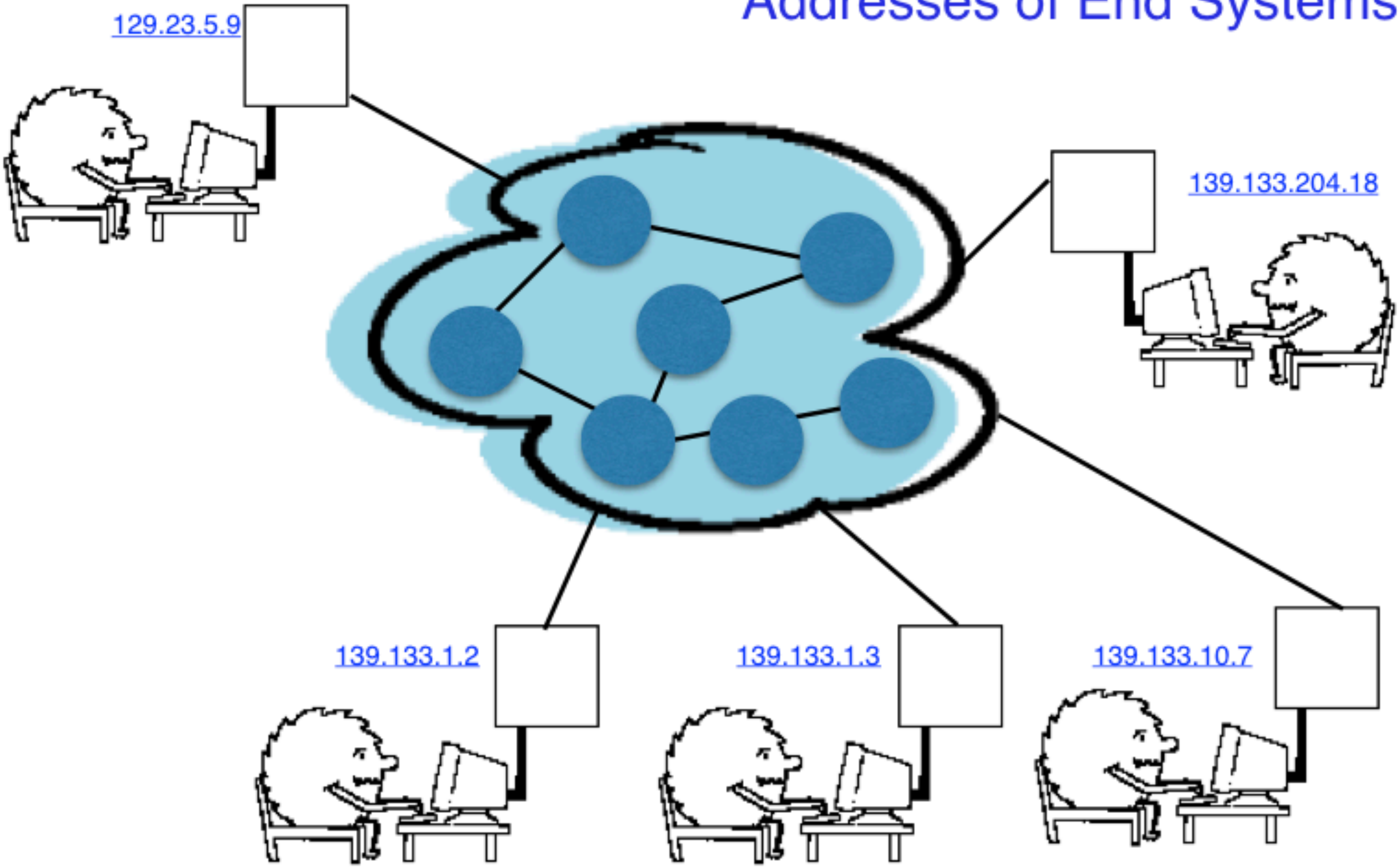
**A payload (also known as the SDU)**
- A piece of the data to be communicated



# Internet Addresses

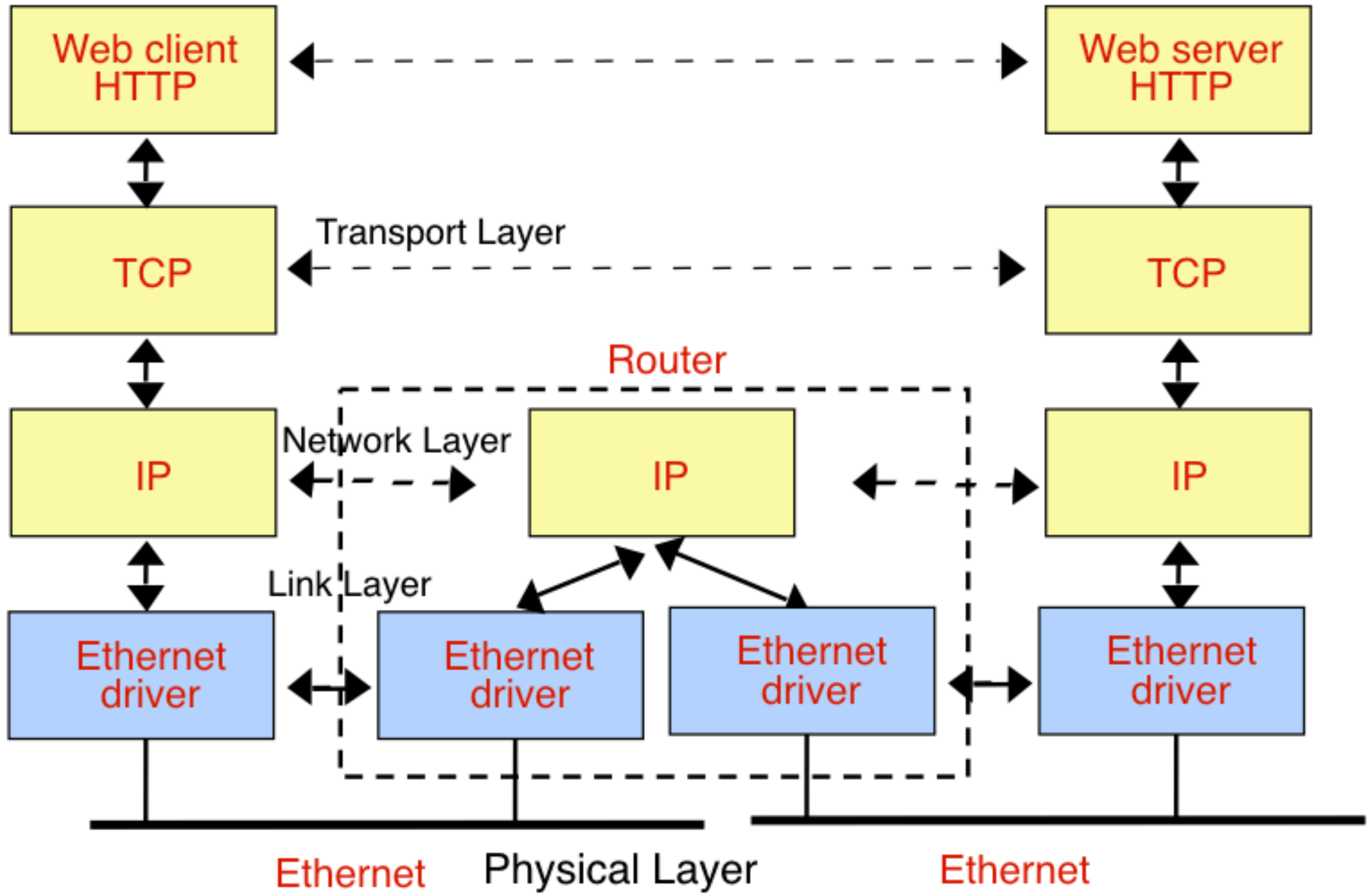
G Fairhurst, http://www.erg.abdn.ac.uk

## Addresses of End Systems



# The TCP/IP stack for web

G Fairhurst, http://www.erg.abdn.ac.uk

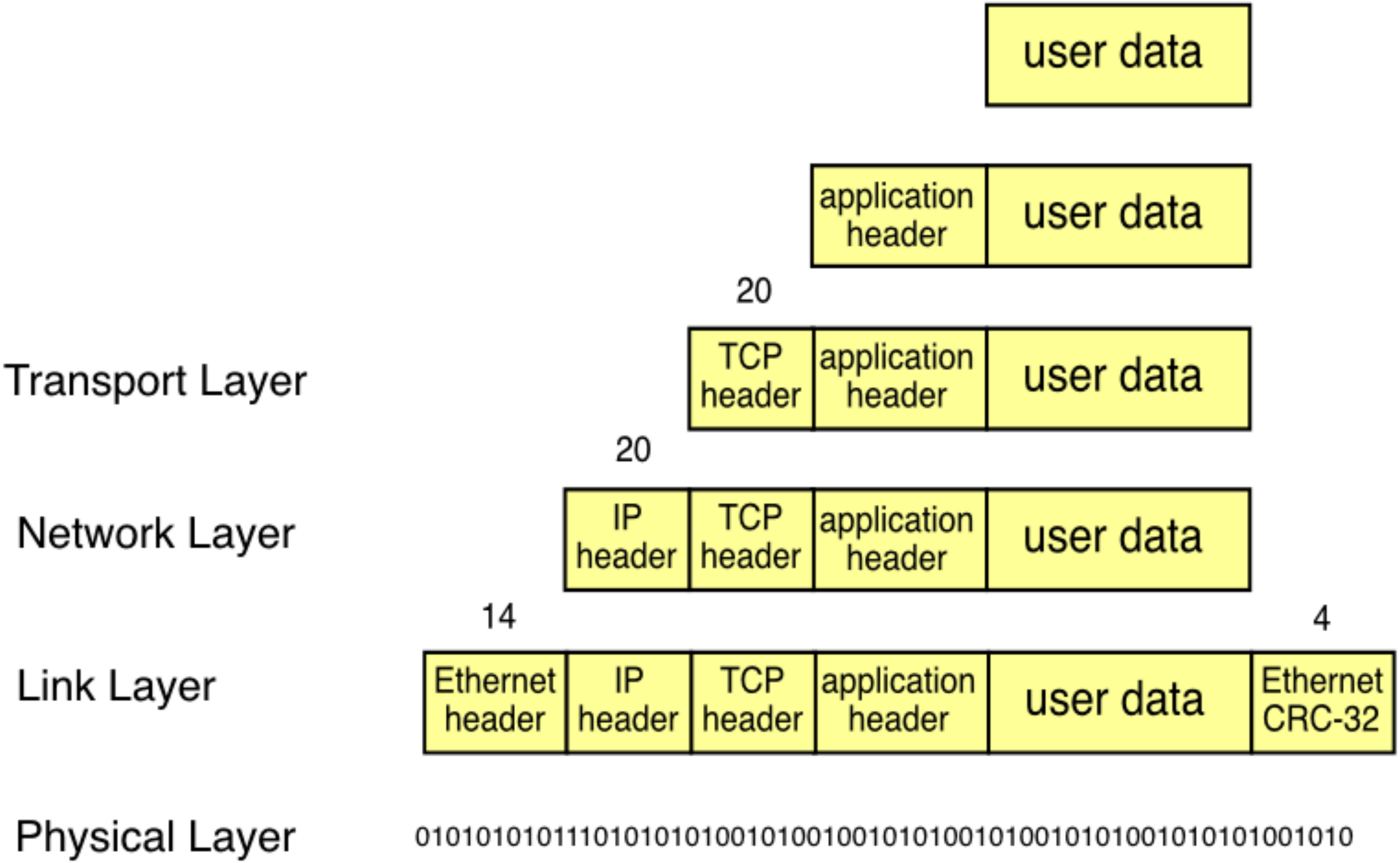


# Interface Layers (L1 & L2)

## Encapsulation for Ethernet

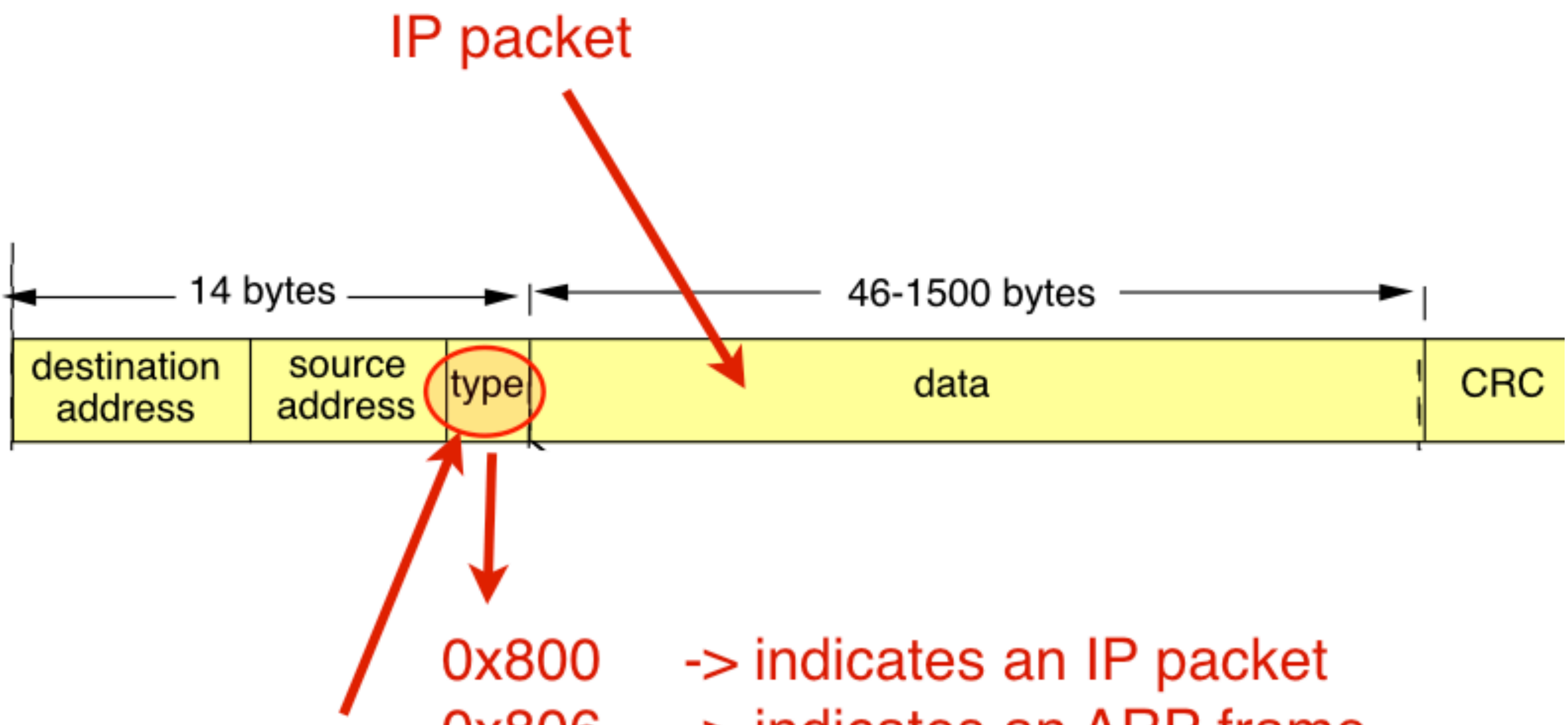
# Encapsulation

G Fairhurst, http://www.erg.abdn.ac.uk



# Ethernet Header

G Fairhurst, http://www.erg.abdn.ac.uk

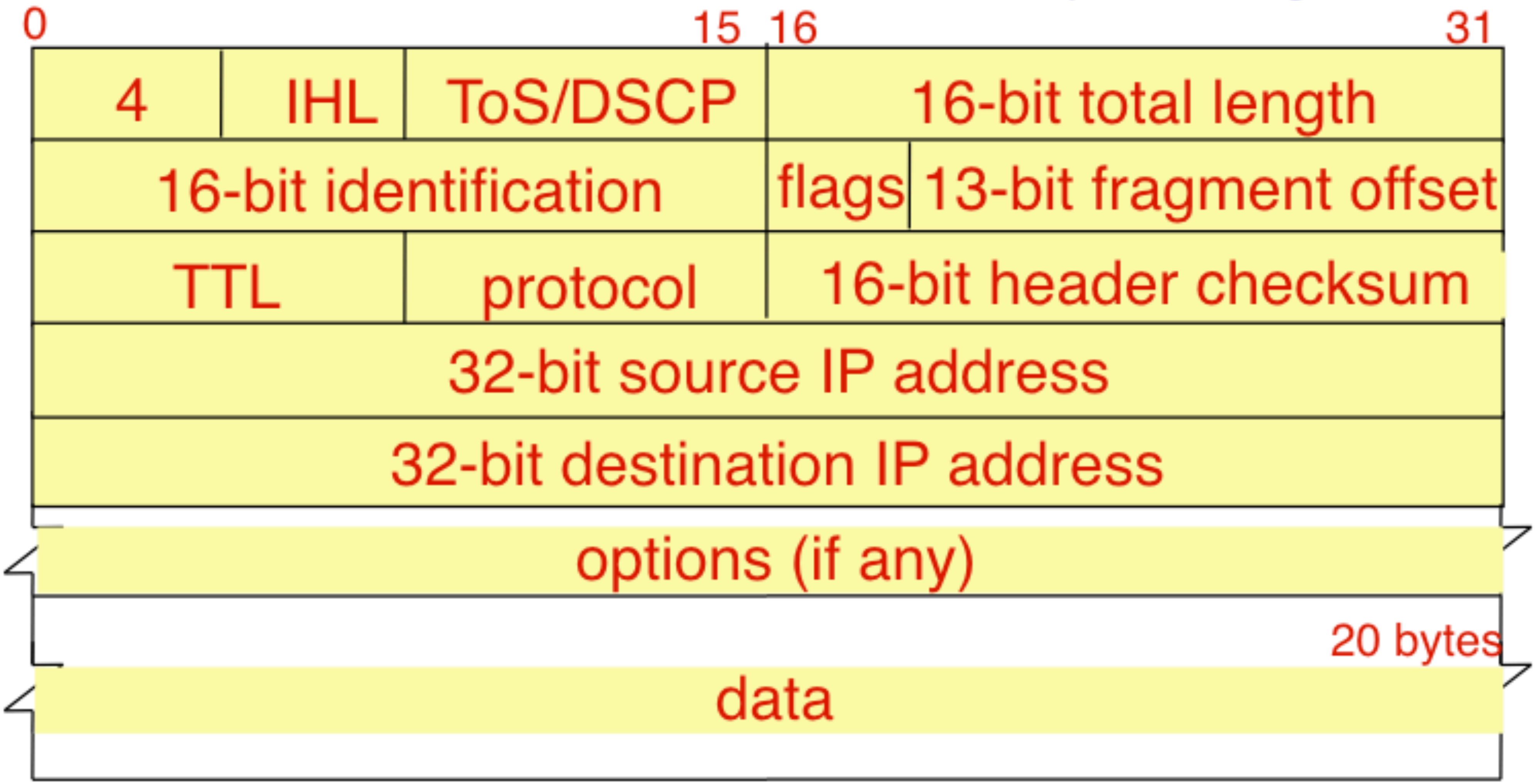



0x800 -> indicates an IP packet
0x806 -> indicates an ARP frame
0x86DD -> indicates an IPv6 packet

Layer 2 SAP

# IP Header

G Fairhurst, http://www.erg.abdn.ac.uk



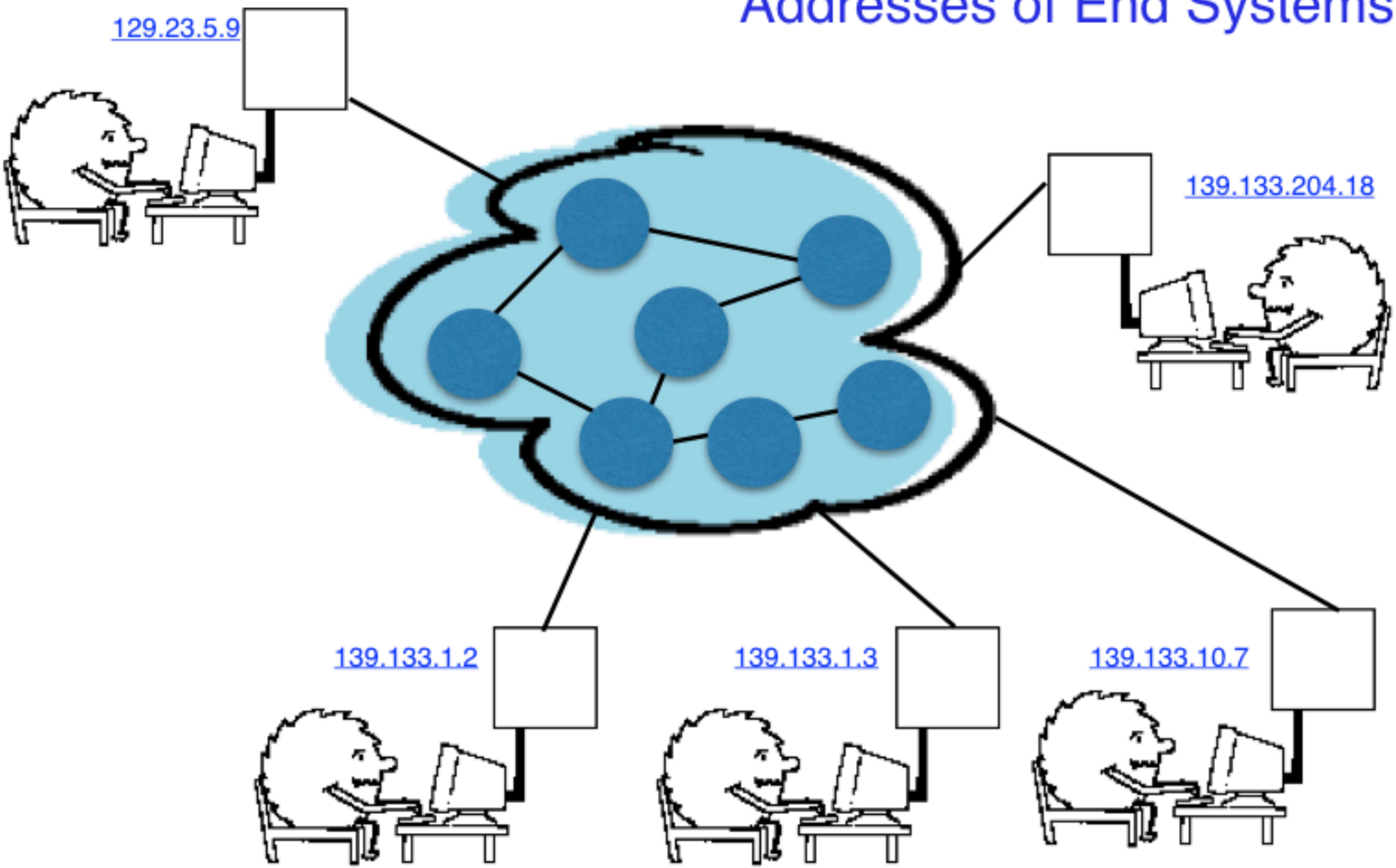

RFC 791

# Addresses

# Internet Addresses

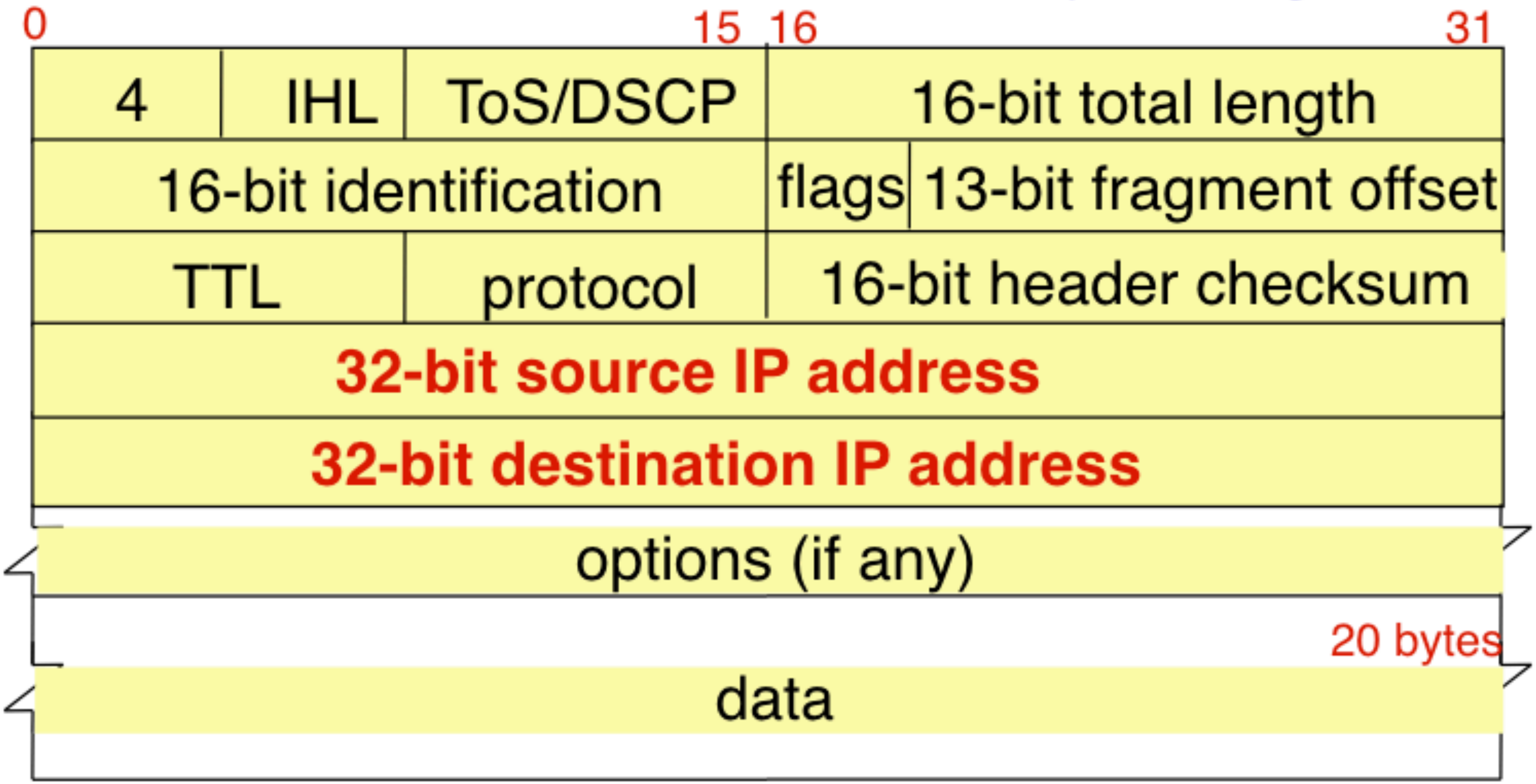
G Fairhurst, http://www.erg.abdn.ac.uk

## Addresses of End Systems



# IP Header

G Fairhurst, http://www.erg.abdn.ac.uk




RFC 791

G Fairhurst, http://www.erg.abdn.ac.uk

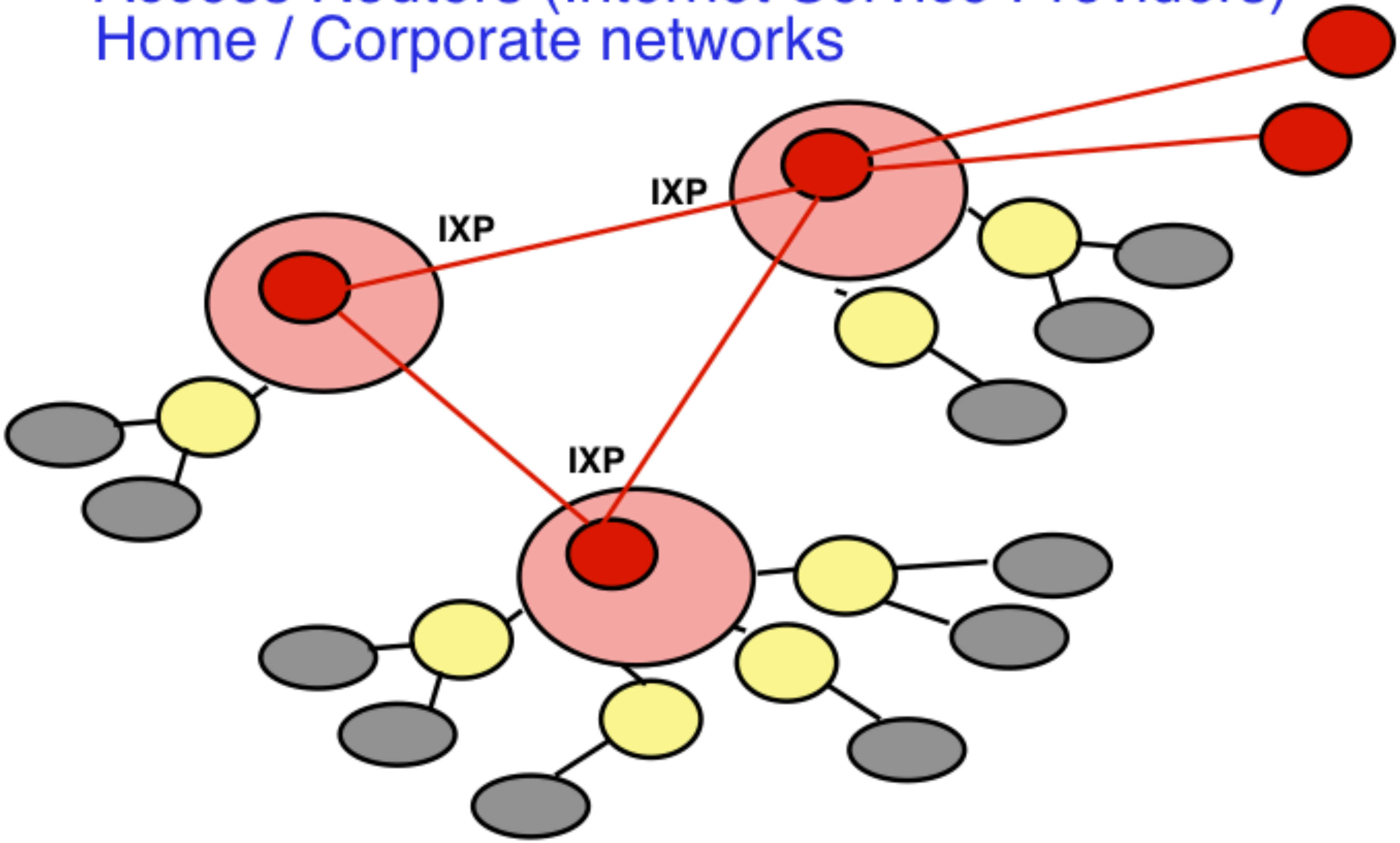

Blocks of addresses assigned to each service provider

Addresses then given (or sold) to customers/users
Sometimes addresses are "leased" for a period of time

G Fairhurst, http://www.erg.abdn.ac.uk

**Arranged in four levels:**

- Core Routers (No user networks connected)
- Distribution Routers (Regional networks)
- Access Routers (Internet Service Providers)
- Home / Corporate networks



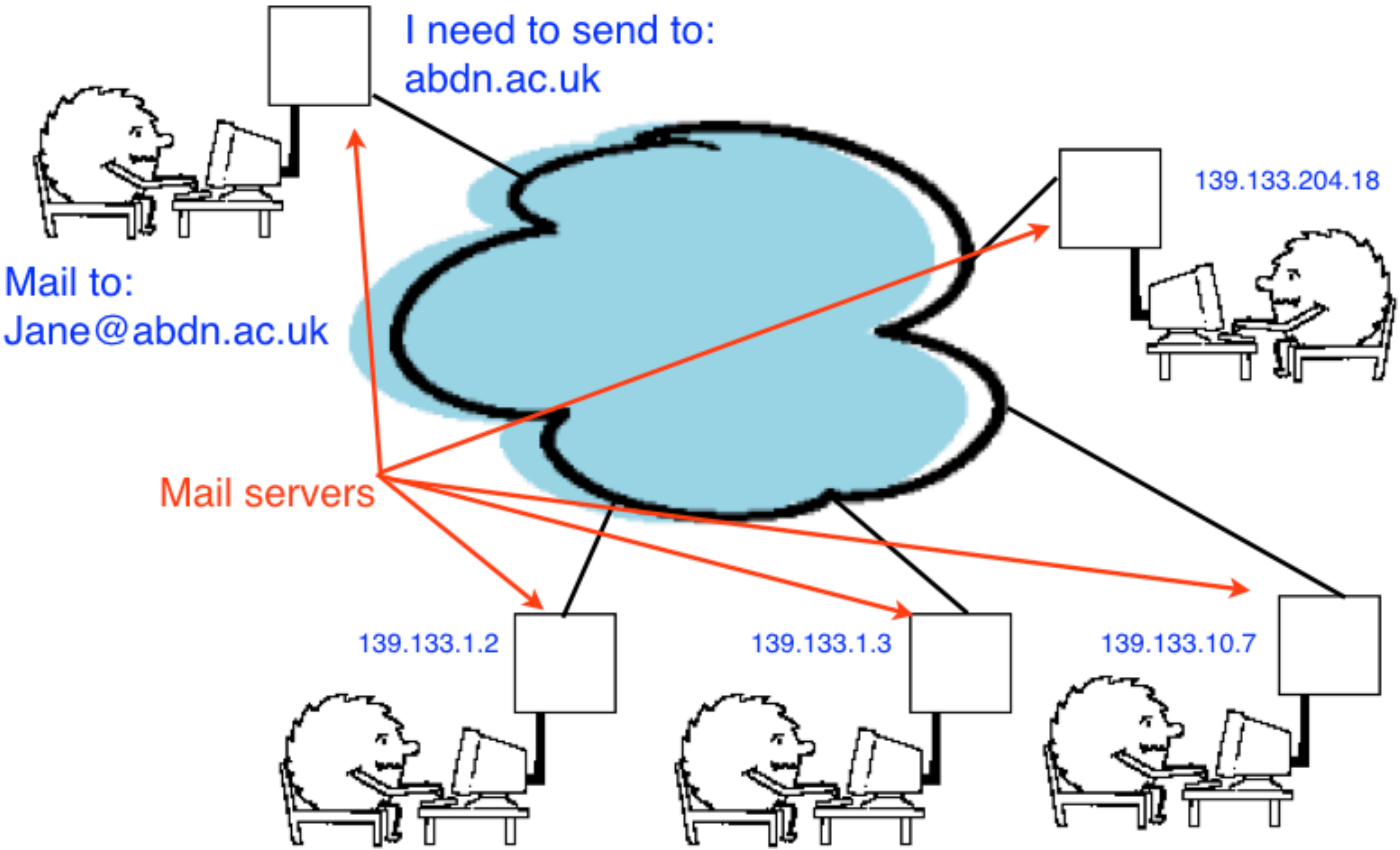
A4

## Name Resolution

- Name and Addresses
- Flat v. Hierarchical Structures
- The DNS

# Using Internet Addresses

G Fairhurst, http://www.erg.abdn.ac.uk



# *Organisation of names and addresses*

G Fairhurst

There are two ways of identifying a computer, using:

A name

A network address

Names and addresses may be organised using:

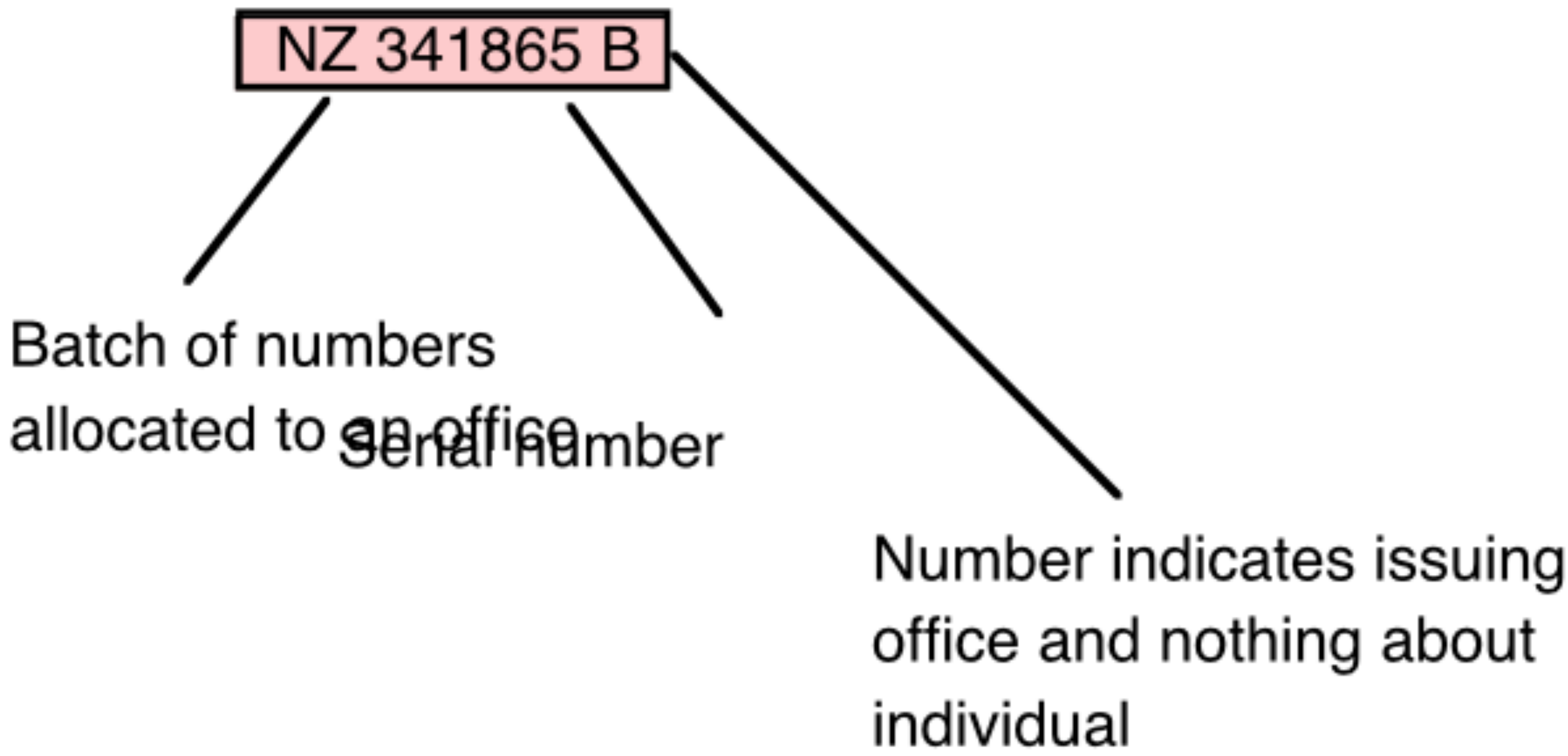
A flat structure

A hierarchical structure

# Flat Structure

G Fairhurst

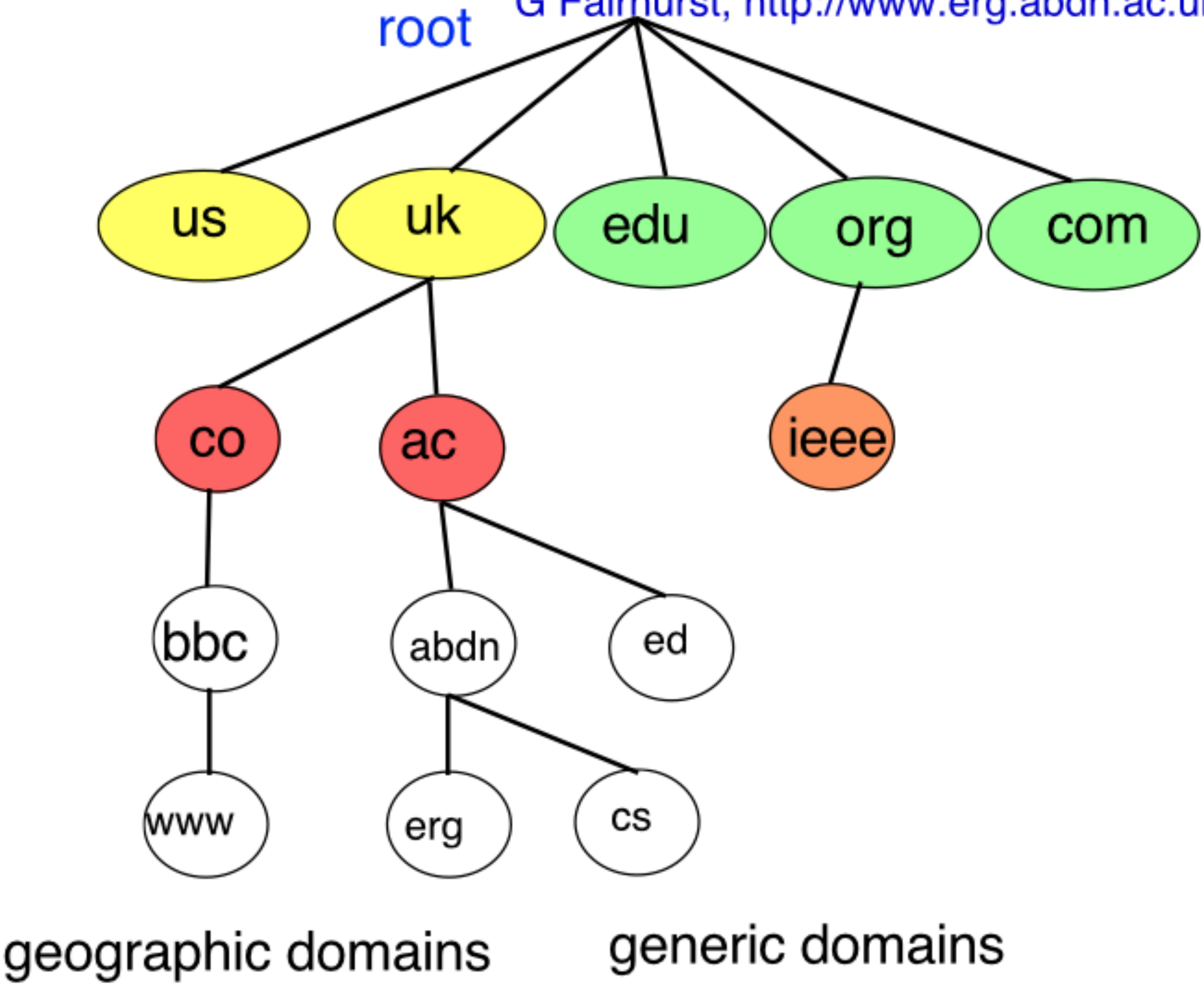
**National Insurance Number**



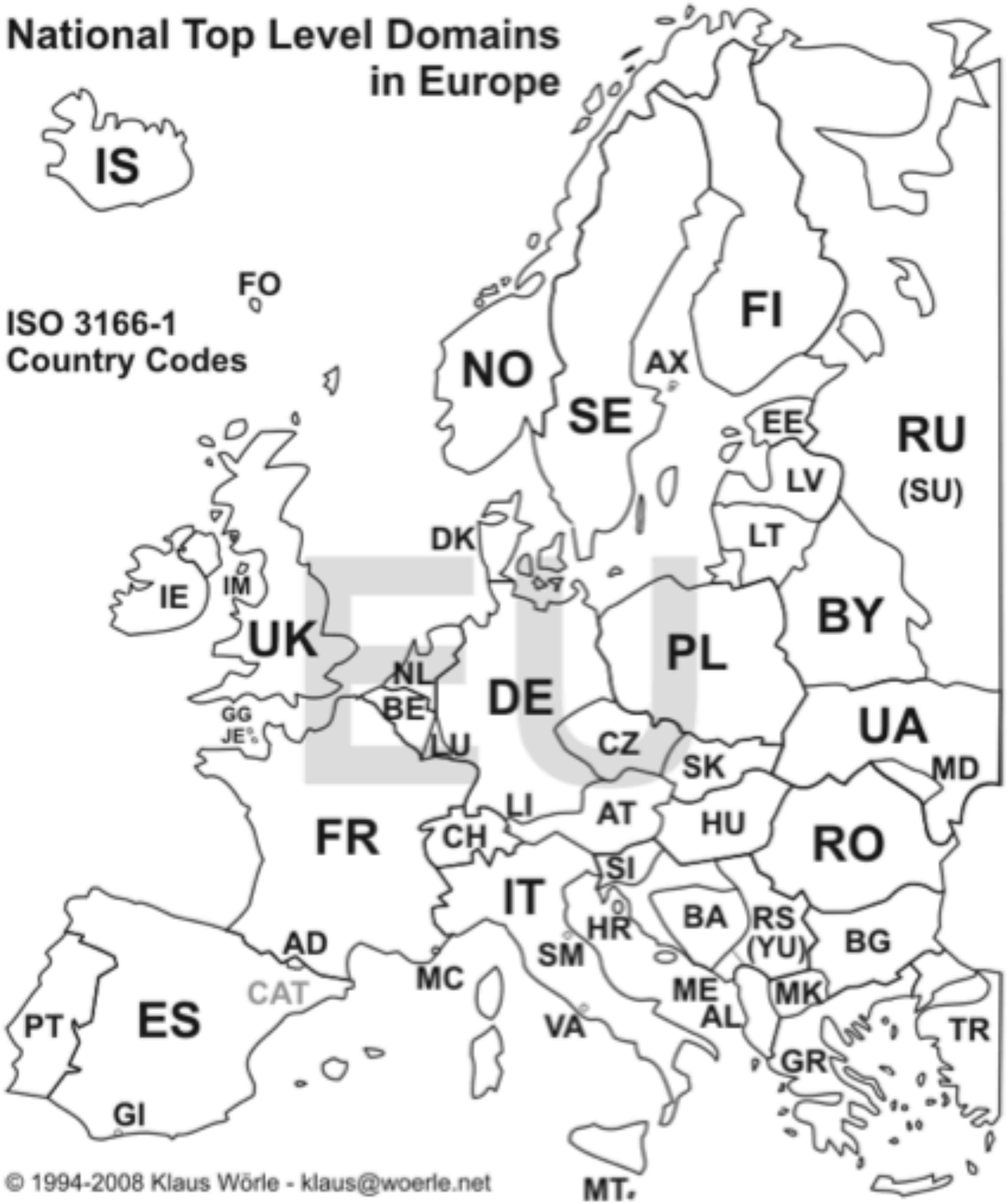

Flat address structures are used by IEEE for the MAC address assignments

# Domain Name Service Tree

G Fairhurst, http://www.erg.abdn.ac.uk



# Top-Level EU DNS Domains



# Flat v Hierarchical Structure

G Fairhurst

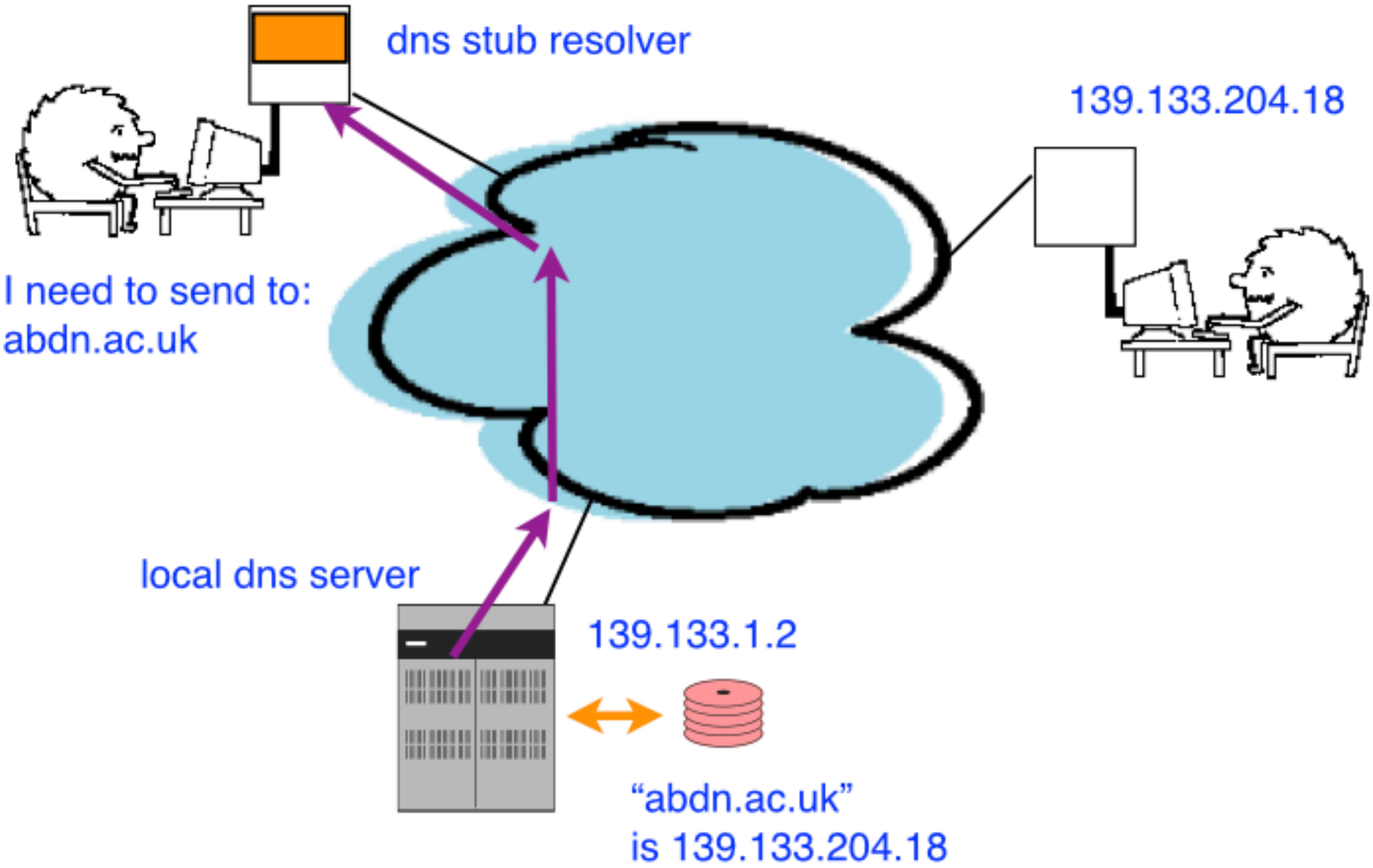
| Hierarchical | Flat |
|---|---|
| Easy to remember | Difficult to remember |
| Abbreviated name possible | No unique abbreviations |
| Easy to find location of name | Only uniquely identifies |
| Difficult to change location | Easy to change location |
| Locally administer names | Names allocated centrally |

e.g. telephone no.
Postcode
IP name (DNS)

e.g.
social security no.
IP address

# Internet Email: dns response

G Fairhurst, http://www.erg.abdn.ac.uk



# Evolution of the DNS

G Fairhurst, http://www.erg.abdn.ac.uk

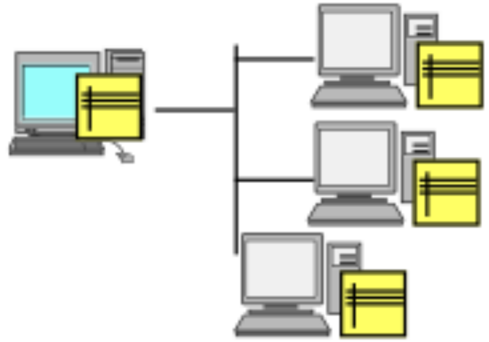
**A single file**

/etc/hosts (in unix)

entered by person setting-up computer

**A central file (at internic.arpa)**

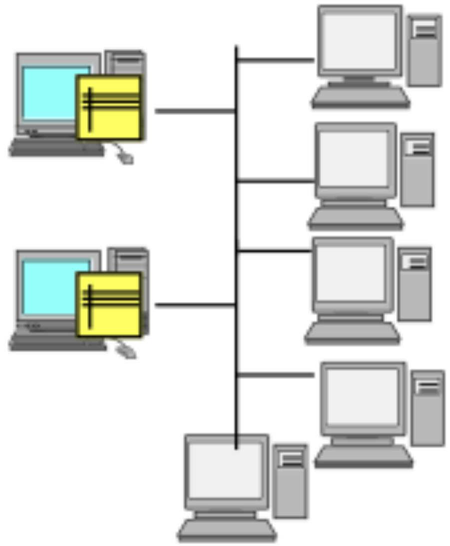
downloaded to /etc/hosts (using ftp)



**A distributed database**

clients send a request (query)

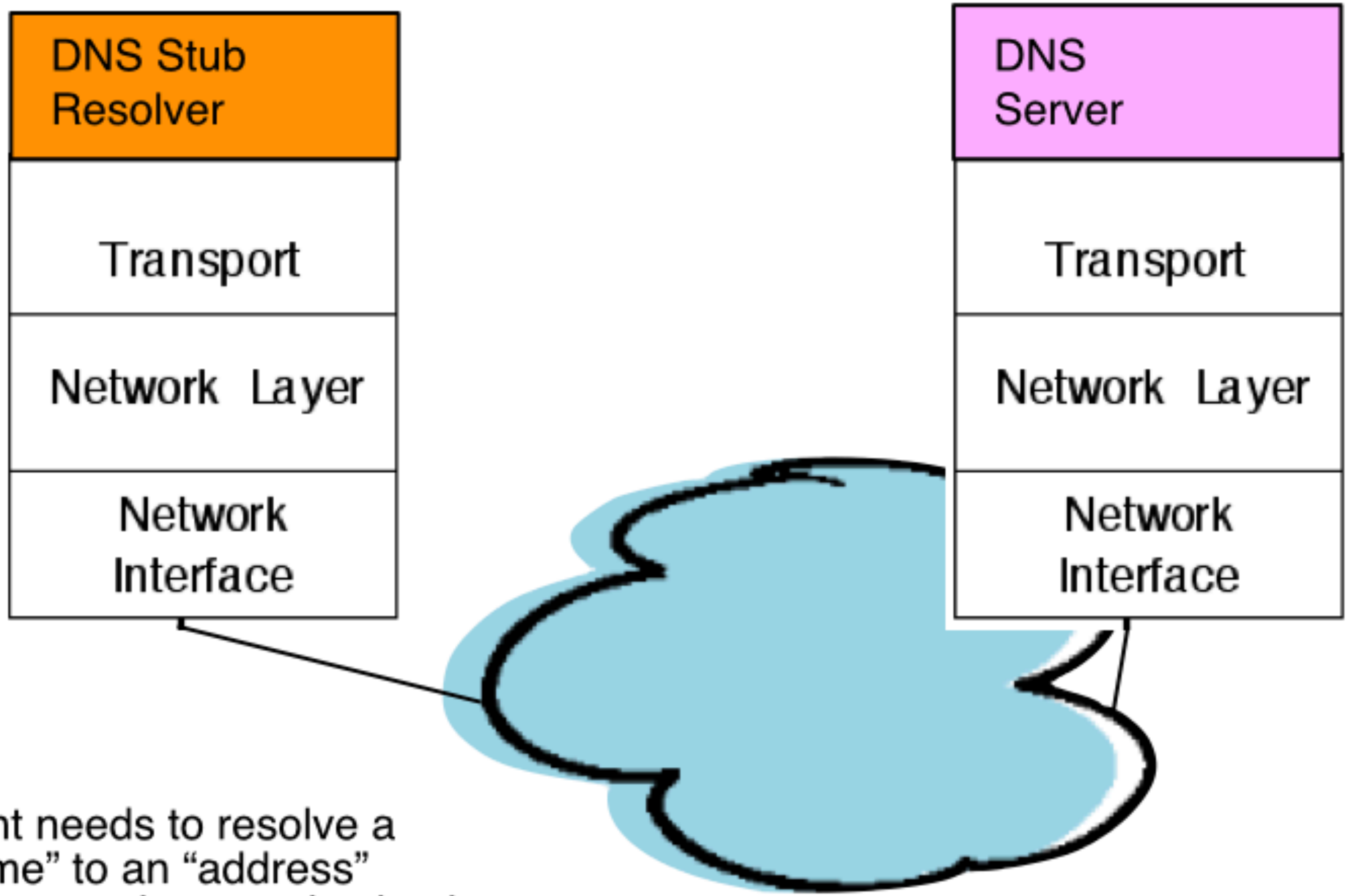
a dns sends a response (resolution)



Most systems still also have a "/etc/hosts" and some also use a LAN name server

# DNS Stack

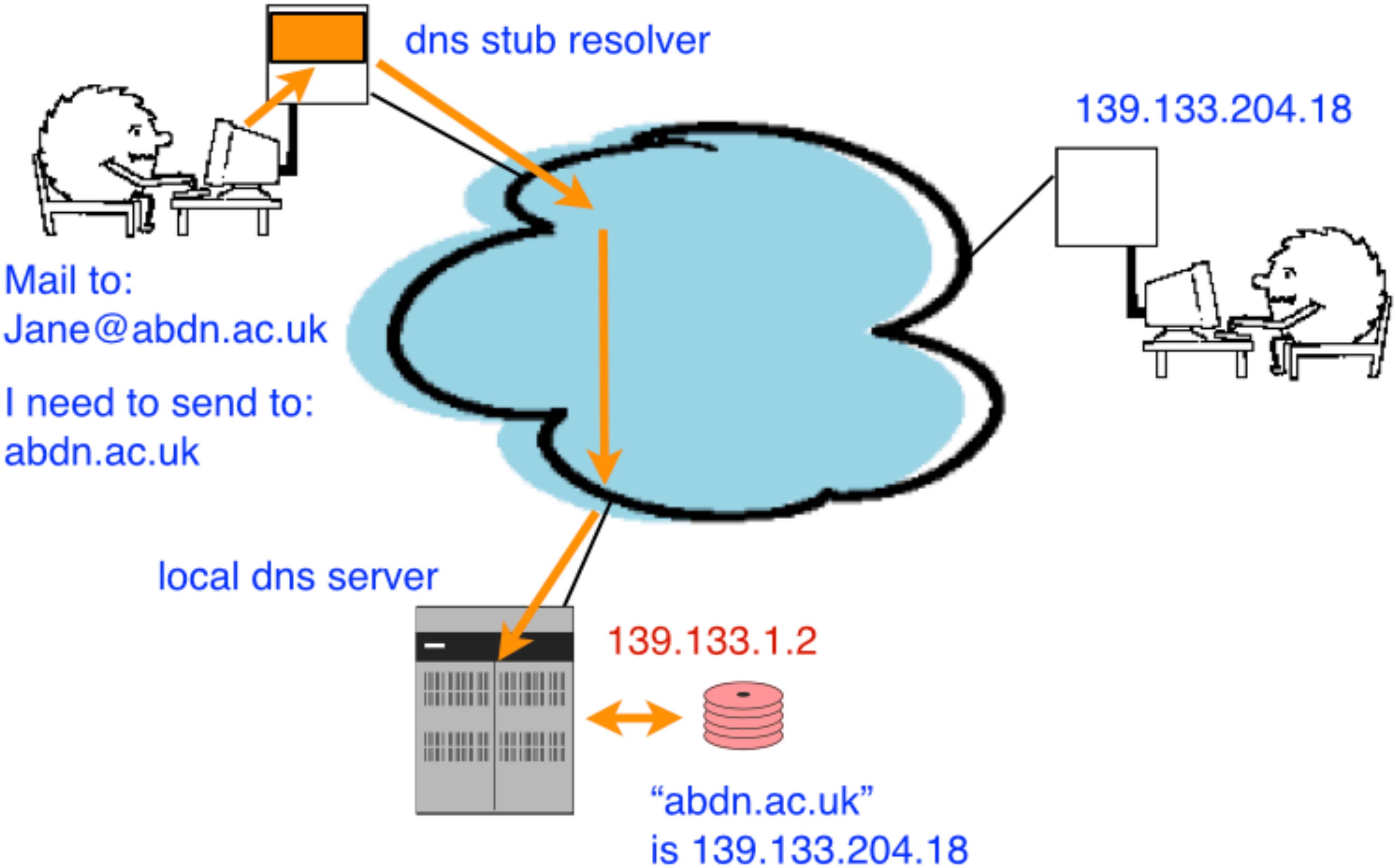
G Fairhurst, http://www.erg.abdn.ac.uk




client needs to resolve a "name" to an "address" to communicate to destination

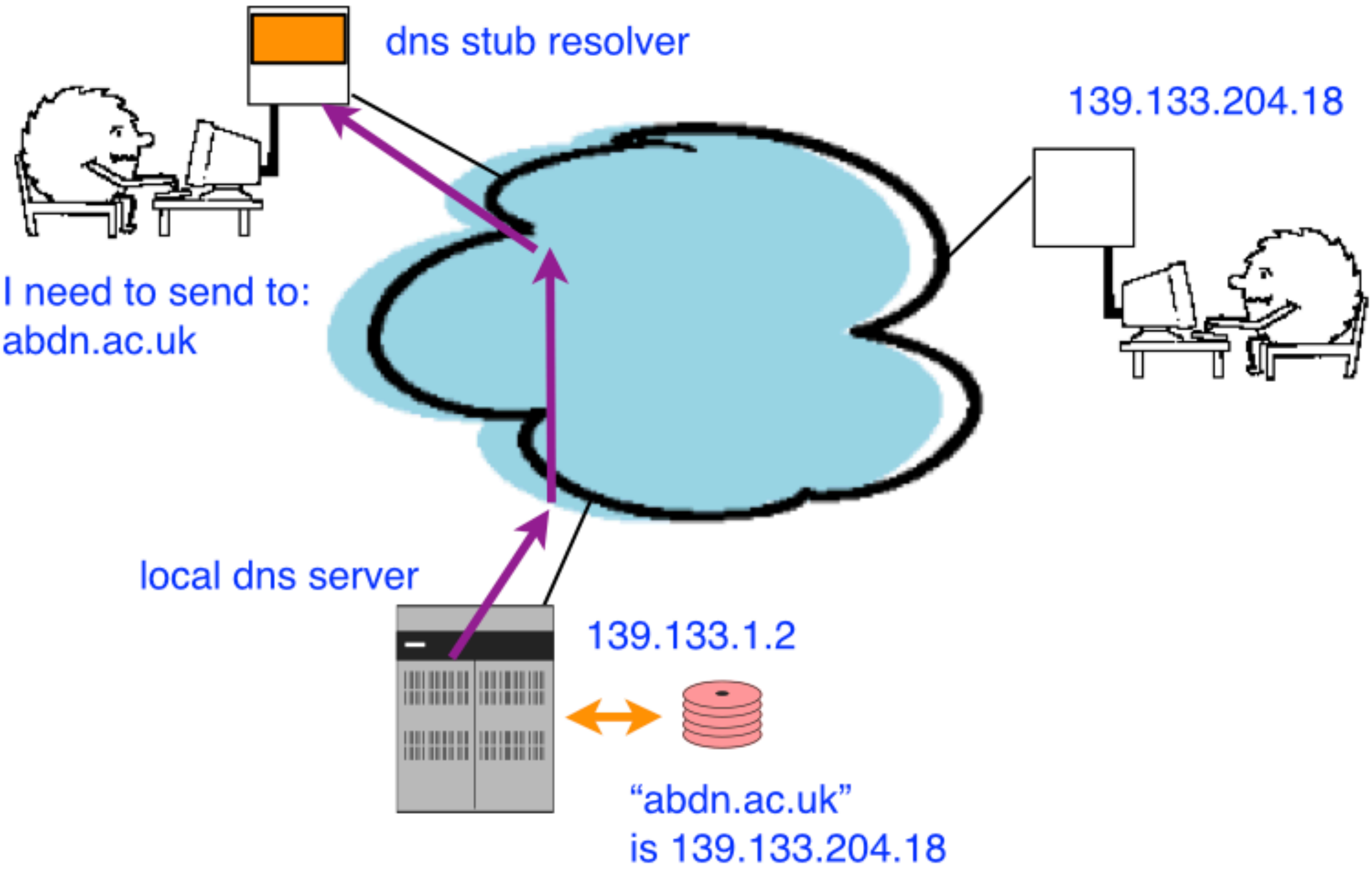
# Internet Email: dns query

G Fairhurst, http://www.erg.abdn.ac.uk



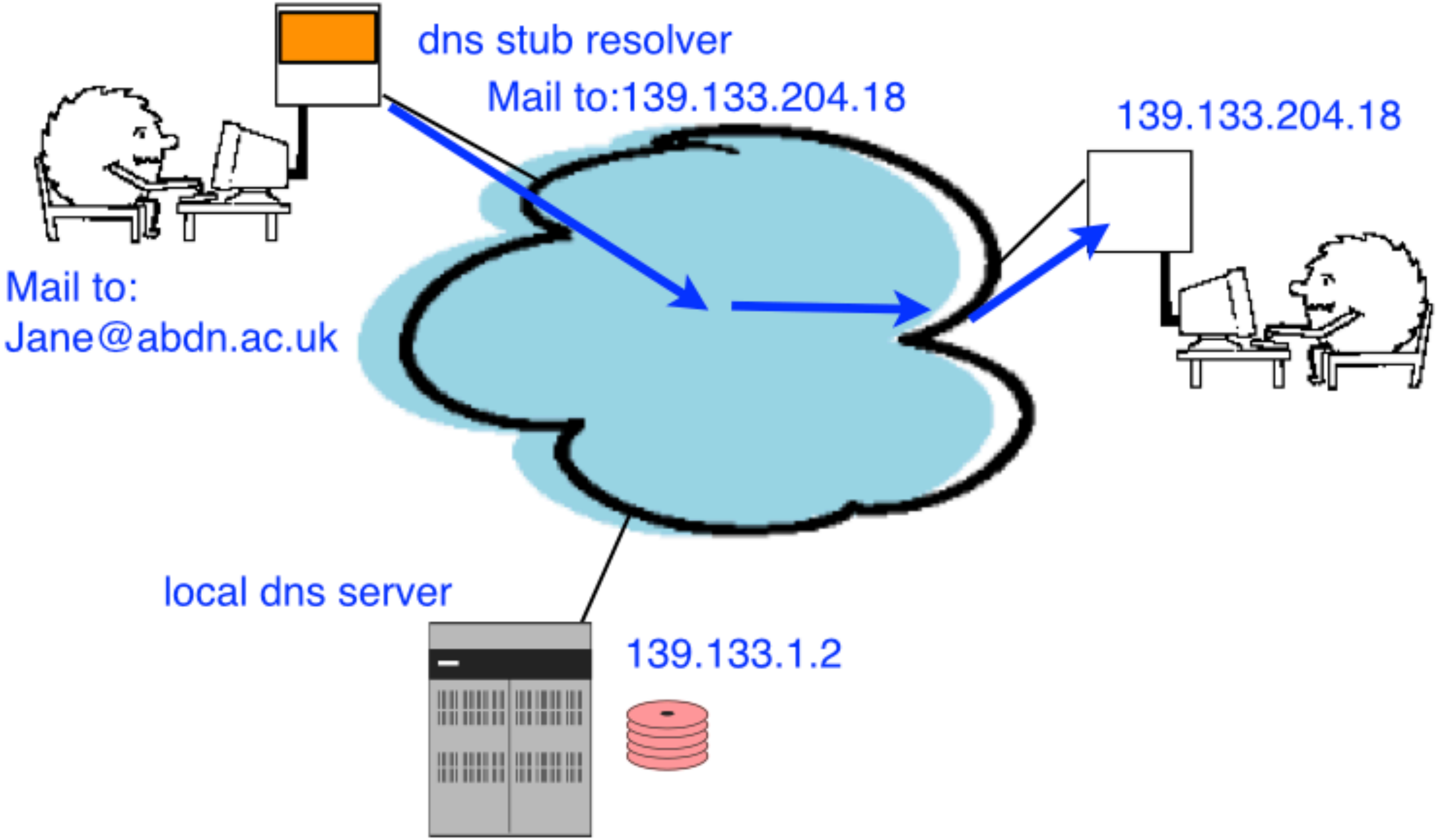
# Internet Email: dns response

G Fairhurst, http://www.erg.abdn.ac.uk



# Sending an Email

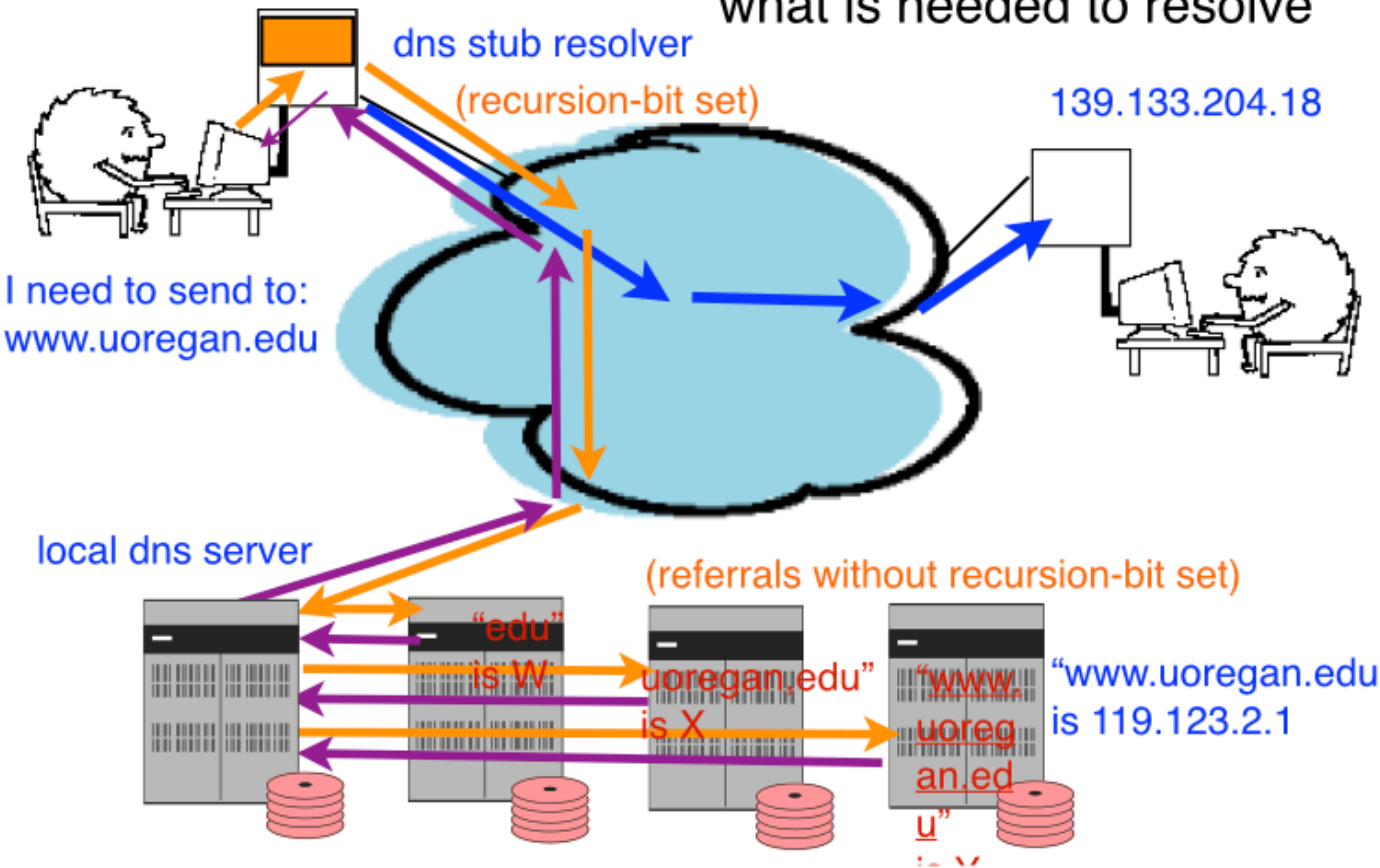
G Fairhurst, http://www.erg.abdn.ac.uk



# Recursive Lookup

G Fairhurst, http://www.erg.abdn.ac.uk

Recursion asks server to do what is needed to resolve



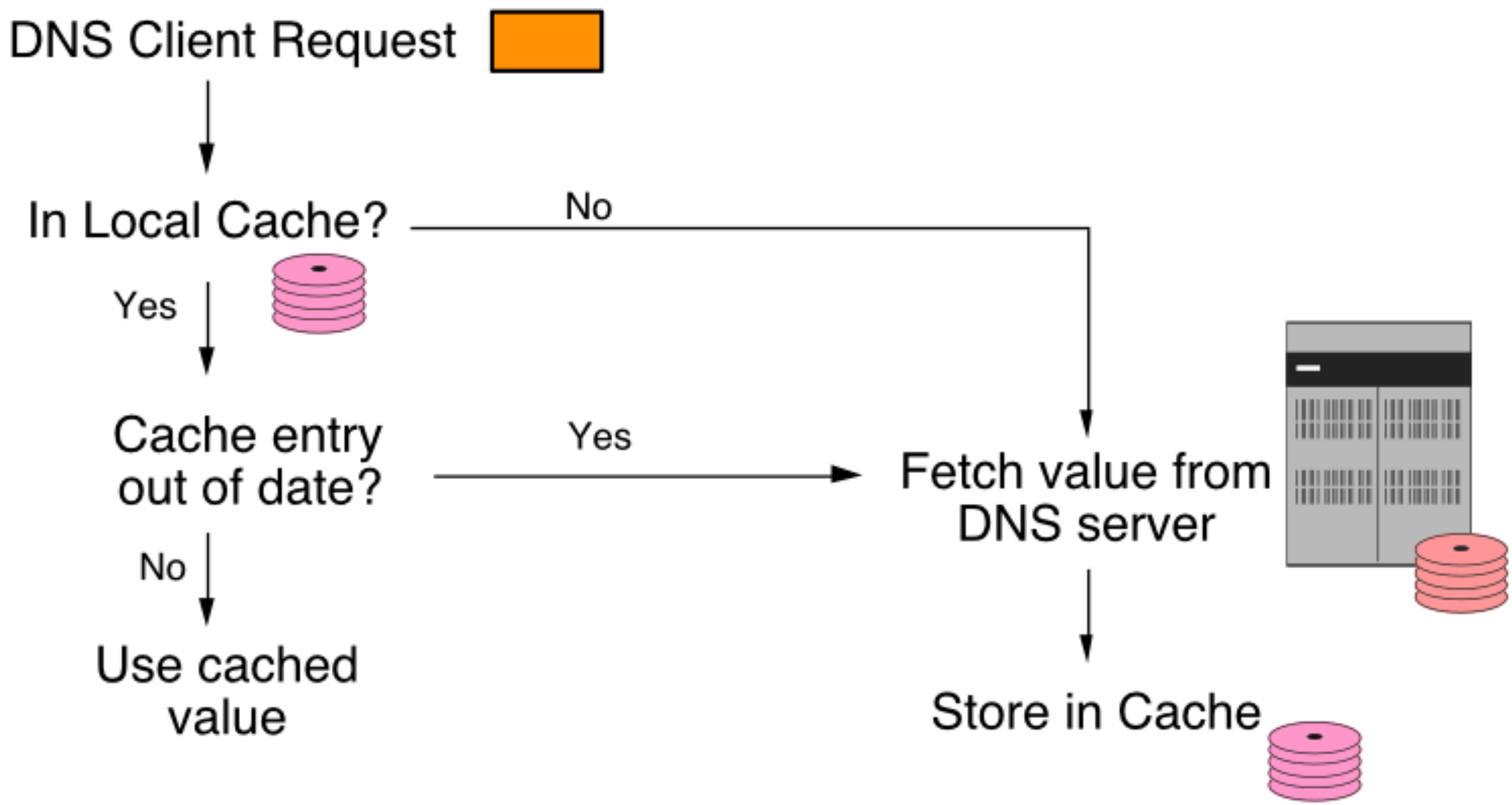
# Caching Lookups

G Fairhurst, http://www.erg.abdn.ac.uk


We can't do the lookup for every single packet!

- That would be incredibly wasteful
- But we could use a local cache
  - Can then keep track of recent lookup results and re-use them!

# DNS Client Cache

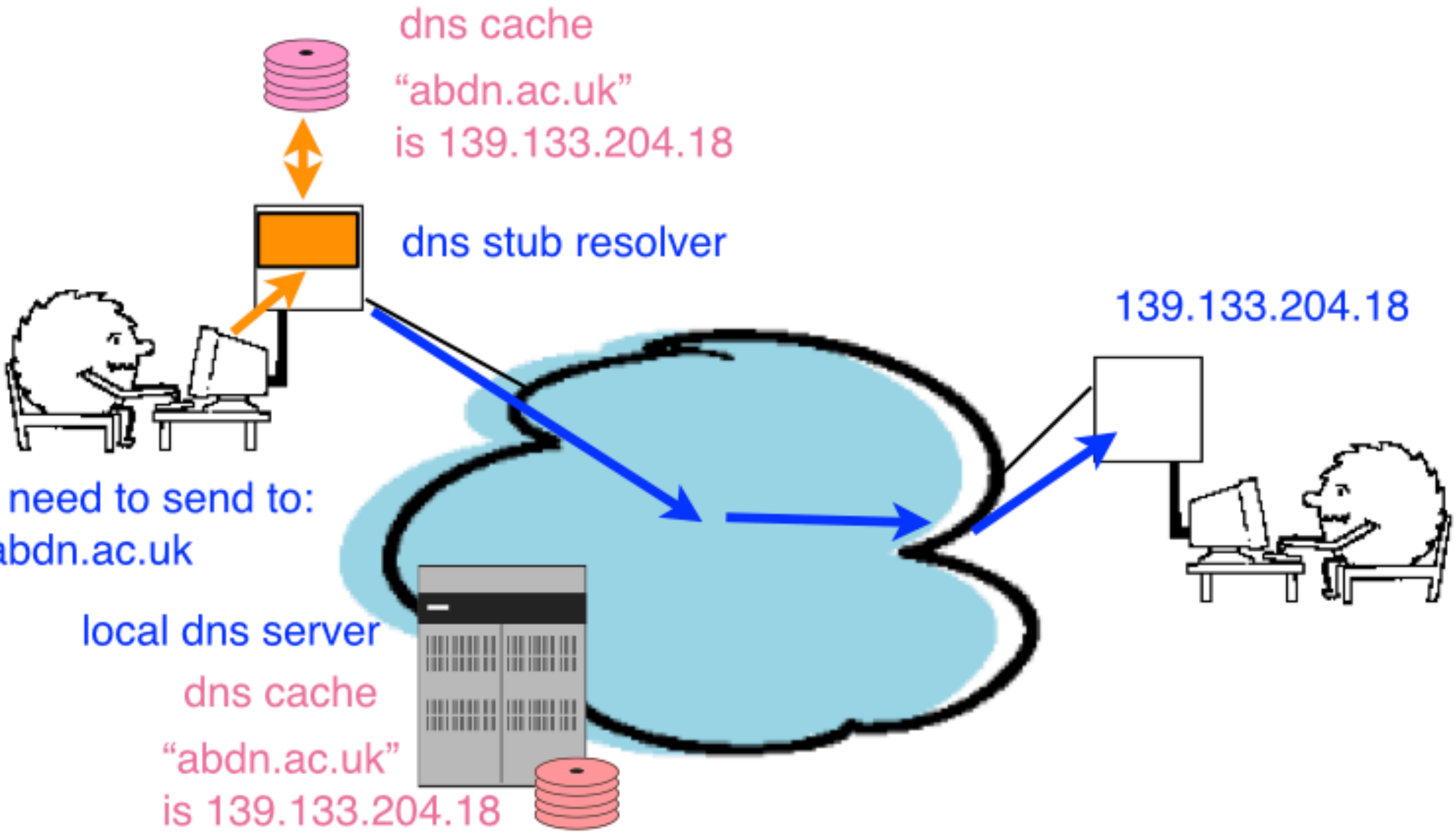


# DNS Cache

G Fairhurst, http://www.erg.abdn.ac.uk

stub cache: My client

Local dns server: Cached for all local users

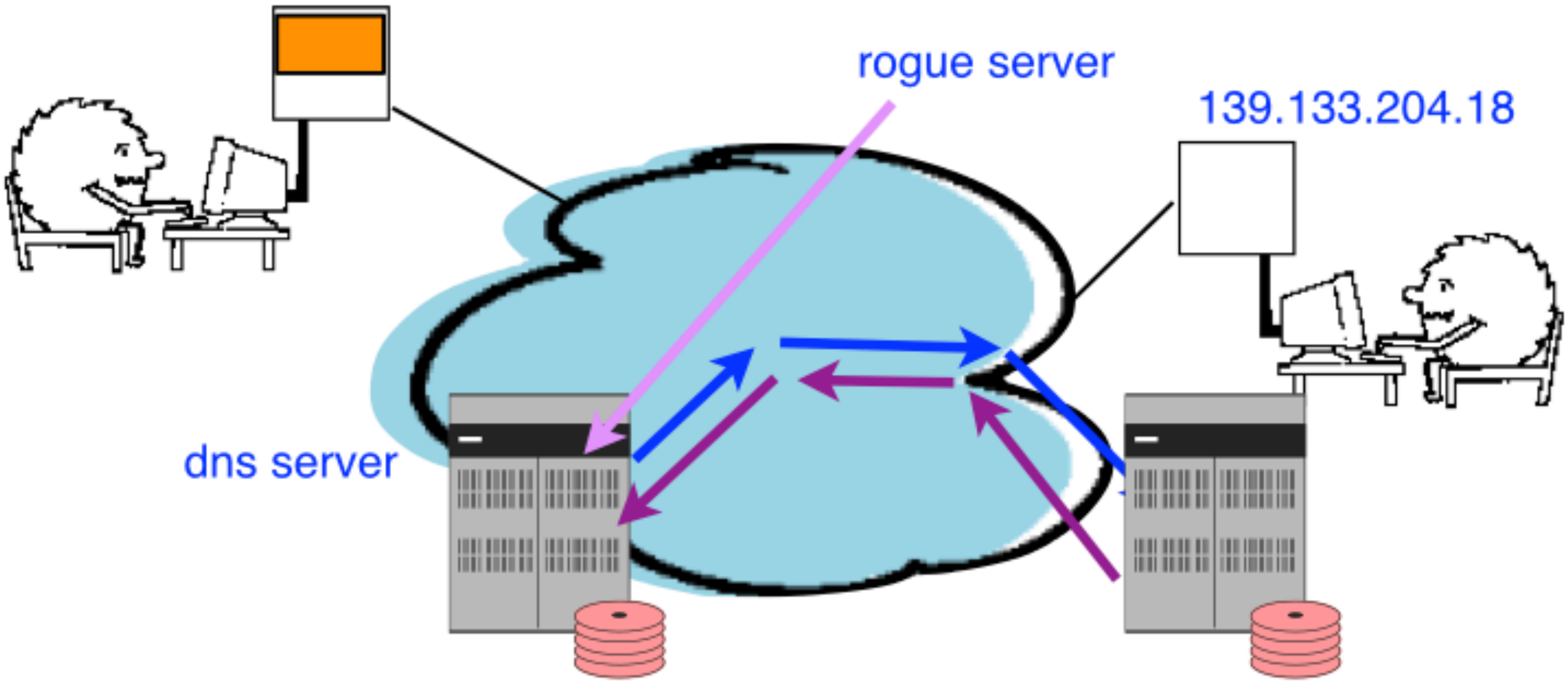


# DNS Sec

G Fairhurst, http://www.erg.abdn.ac.uk

DNS server may need to check referrals to stop *poisoning* the cache

- Check server is authorised to return its response
- Check the request ID and src port (randomised)
- Limits recursion to sub resolvers



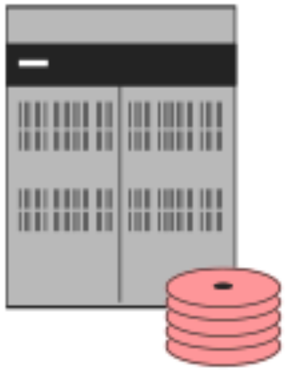
# DNS Resolution

G Fairhurst, http://www.erg.abdn.ac.uk

Browser/Application sends name to resolver (DNS client)

Resolver checks own cache (local files, etc)

If not resolved, contacts DNS Server
(resolver knows this IP address)



If not resolved, contacts root DNS server (.)
May redirect to other server(s)

Resolver given 1 or more addresses
(resolver caches the answer for some time)

Browser/Application given lowest numbered server

# Naming & Addressing: Summary

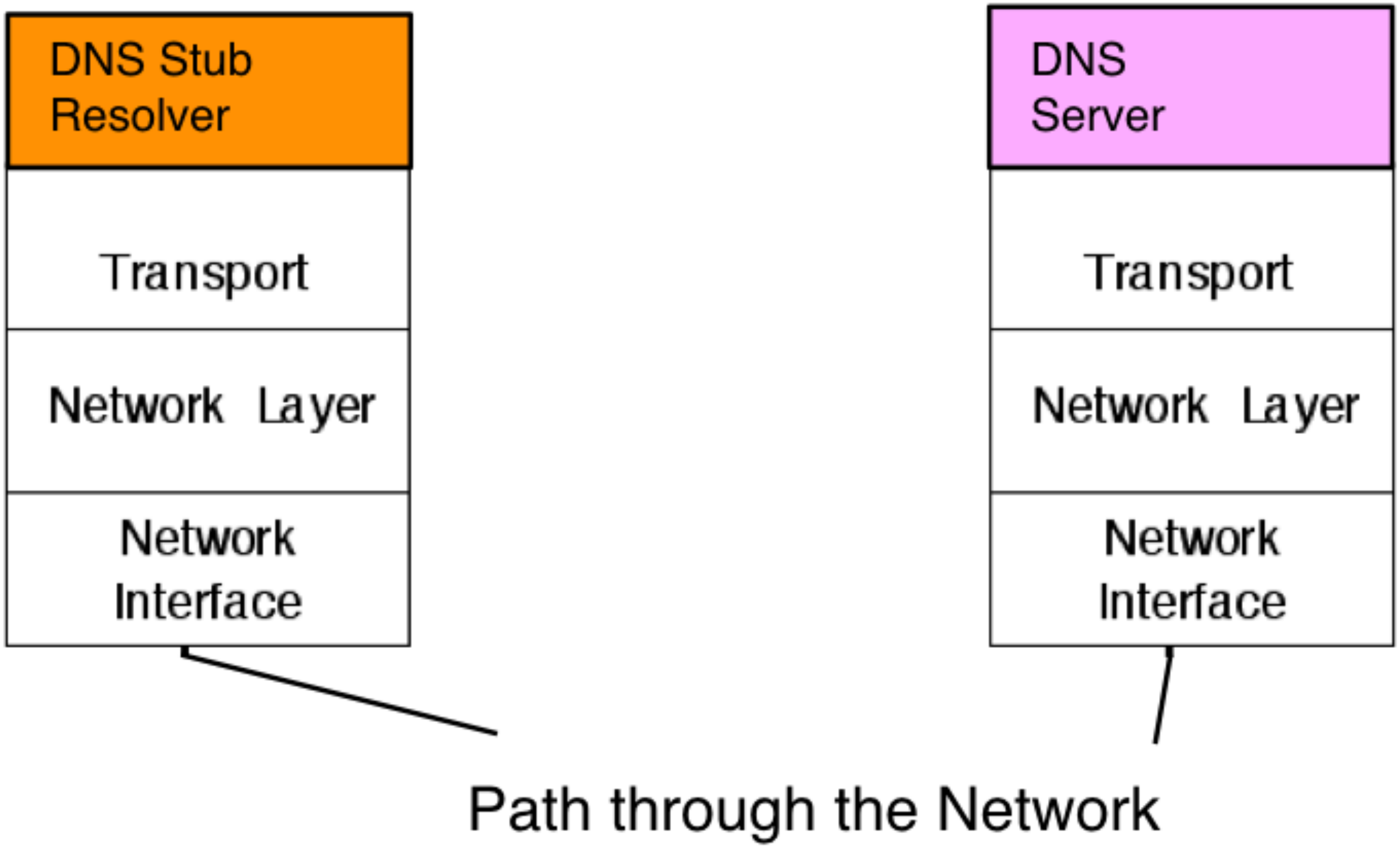
G Fairhurst

A *name* is a *symbol* - designed for human reading
An *address* is a *data structure* understood by a network
Organisation may be *hierarchical* or *flat*

A *name server* provides a service to change between network addresses and network names

To know who's who on the Internet a computer must know the *address* of a name server

# DNS Stack



DNS Stub Resolver
Transport
Network Layer
Network Interface
DNS Server
Transport
Network Layer
Network Interface
Path through the Network

# DNS Records

DNS Records have various types, can also point to another name:

MX records used for Mail Exchange

mail.abdn.ac.uk 3600 IN MX 500 backup.abdn.ac.uk
mail.abdn.ac.uk 3600 IN MX 5 mailserver.abdn.ac.uk
mail.abdn.ac.uk 3600 IN MX 10 mailserver1.abdn.ac.uk

Email uses the lowest numbered reachable mail server

Other formats also use the DNS:

http://www.abdn.ac.uk
ftp://ftp.abdn.ac.uk
sip://jane@swip.net

# DNS Resolution

Browser/Application sends name to resolver (DNS client)

Resolver checks own cache (local files, etc)

If not resolved, contacts DNS Server
  (resolver knows this IP address)

  If not resolved, contacts root DNS server (.)
    May redirect to other server(s)

Resolver given 1 or more addresses
  (resolver caches the answer for some time)

Browser/Application given lowest numbered server

# Naming & Addressing: Summary

G Fairhurst

A *name* is a *symbol* - designed for human reading
An *address* is a *data structure* understood by a network
Organisation may be *hierarchical* or *flat*

A *name server* provides a service to change between network addresses and network names

To know who's who on the Internet a computer must know the *address* of a name server

G Fairhurst, http://www.erg.abdn.ac.uk

# Address Resolution Protocol (arp)

# Address Resolution Protocol (arp)

G Fairhurst, http://www.erg.abdn.ac.uk

All systems connected to the Internet have
a unique IP address

Systems know (or find out from DHCP) their IP address

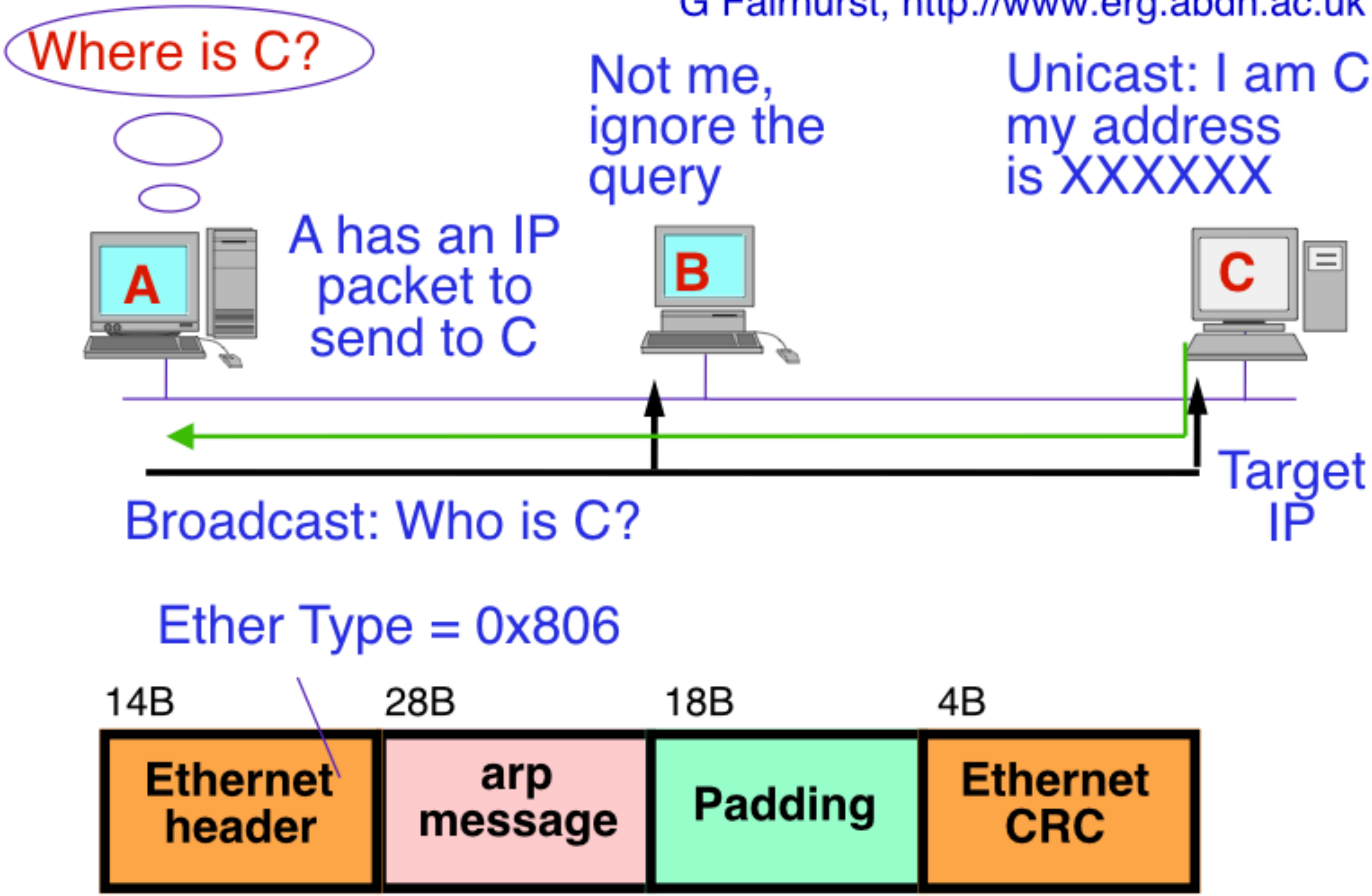
Systems know the IP address of the destination
(or find it out from the DNS)

Systems know their own MAC address
(or can look in the NIC ROM)

No obvious way of determining destination MAC address
- We will call the ***Next Hop IP*** address the ***Target-IP***

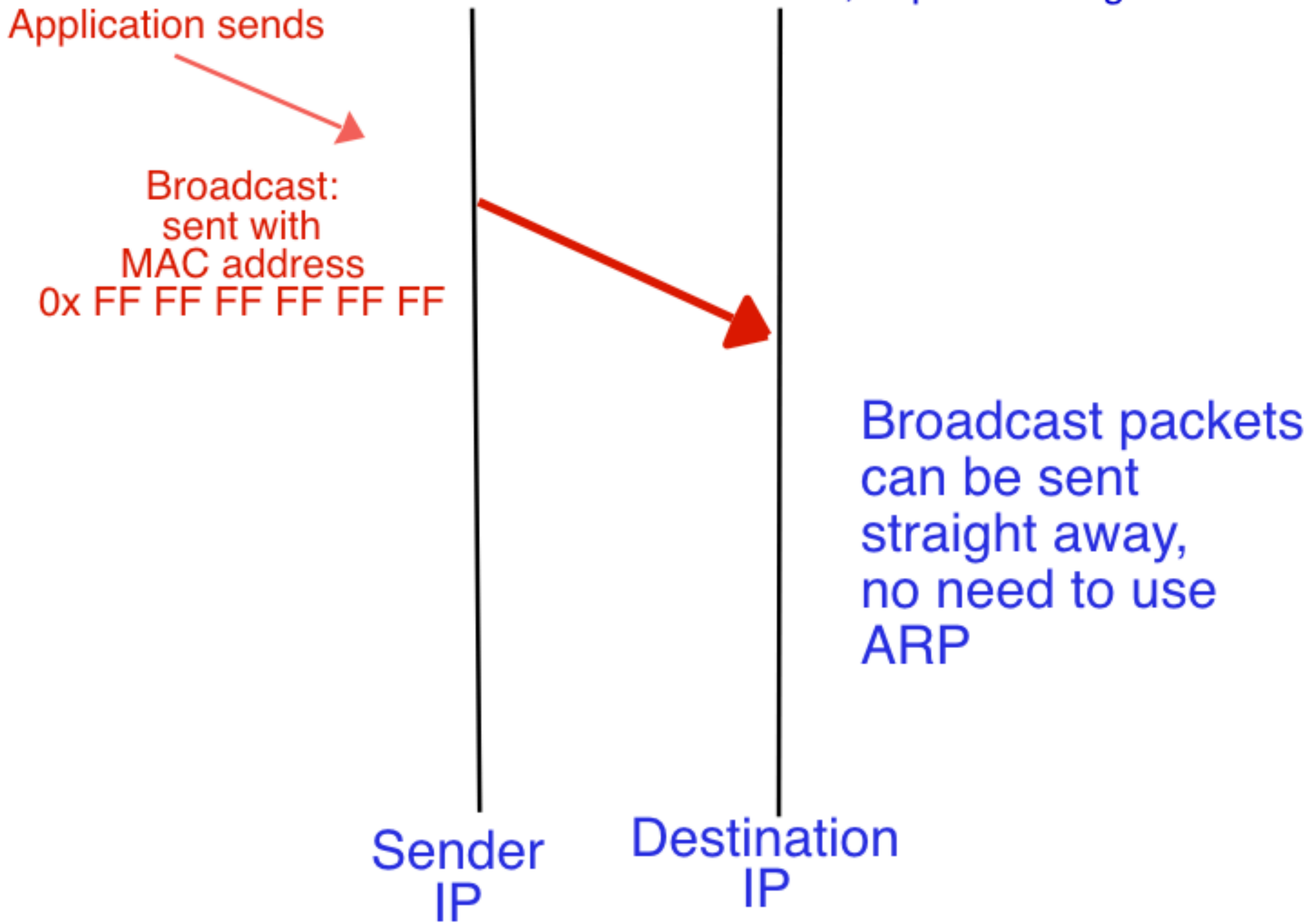
# ARP Request (send A -> C)

G Fairhurst, http://www.erg.abdn.ac.uk



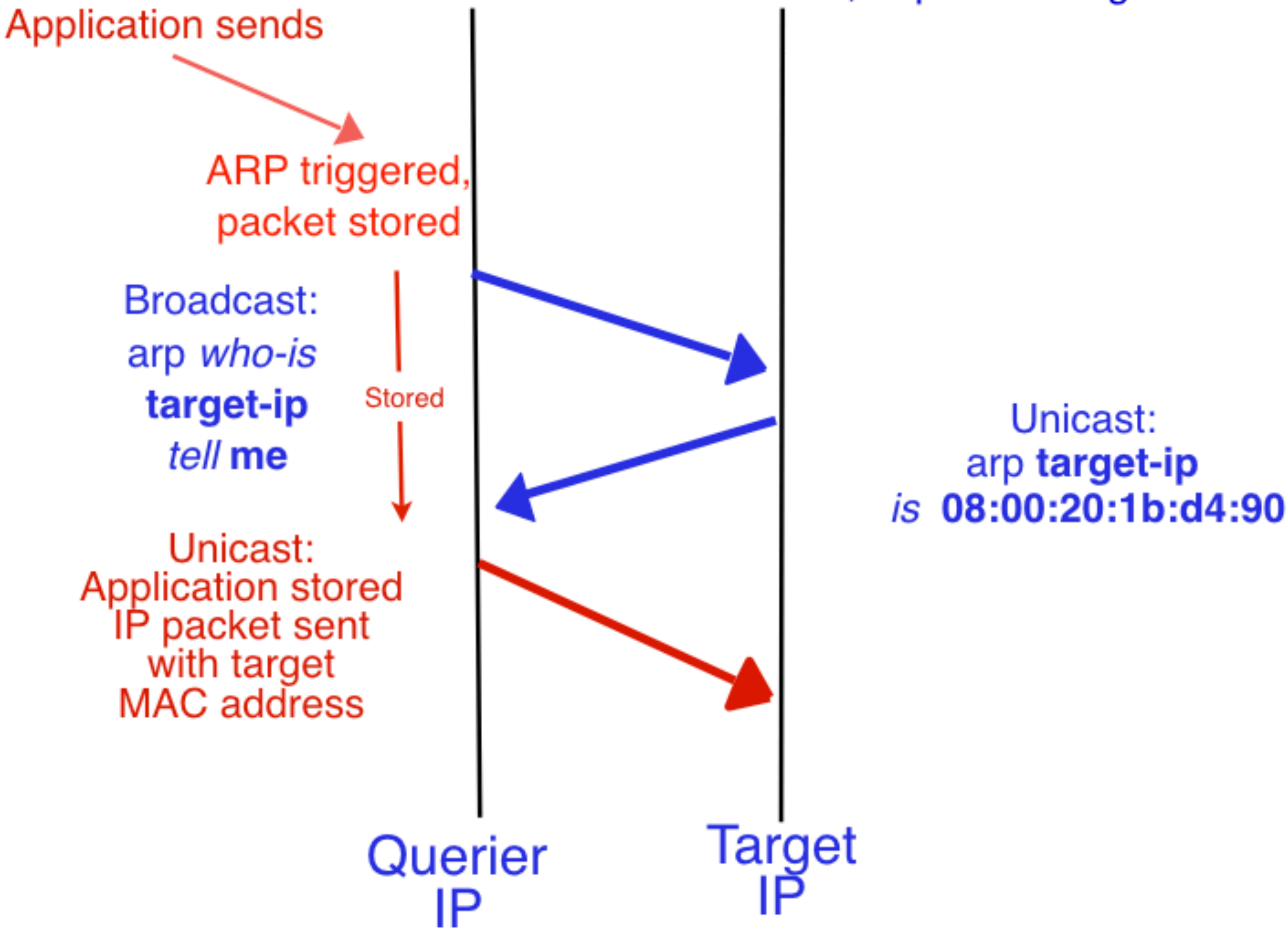
# *IP broadcast  Example*

G Fairhurst, http://www.erg.abdn.ac.uk



# ARP Example

G Fairhurst, http://www.erg.abdn.ac.uk



# ARP Frame

G Fairhurst, http://www.erg.abdn.ac.uk

| 0 | 8   15 | 16   31 |
|---|---|---|
| Hardware Type | | Protocol Type |
| HLEN | PLEN | Operation |
| Sender HA (octets 0-3) | | |
| Sender HA (octets 4-5) | | Sender IP (octets 0-1) |
| Sender !P (octets 2-3) | | Target HA (octets 0-1) |
| Target HA (octets 2-5) | | |
| Target !P (octets 0-3) | | |

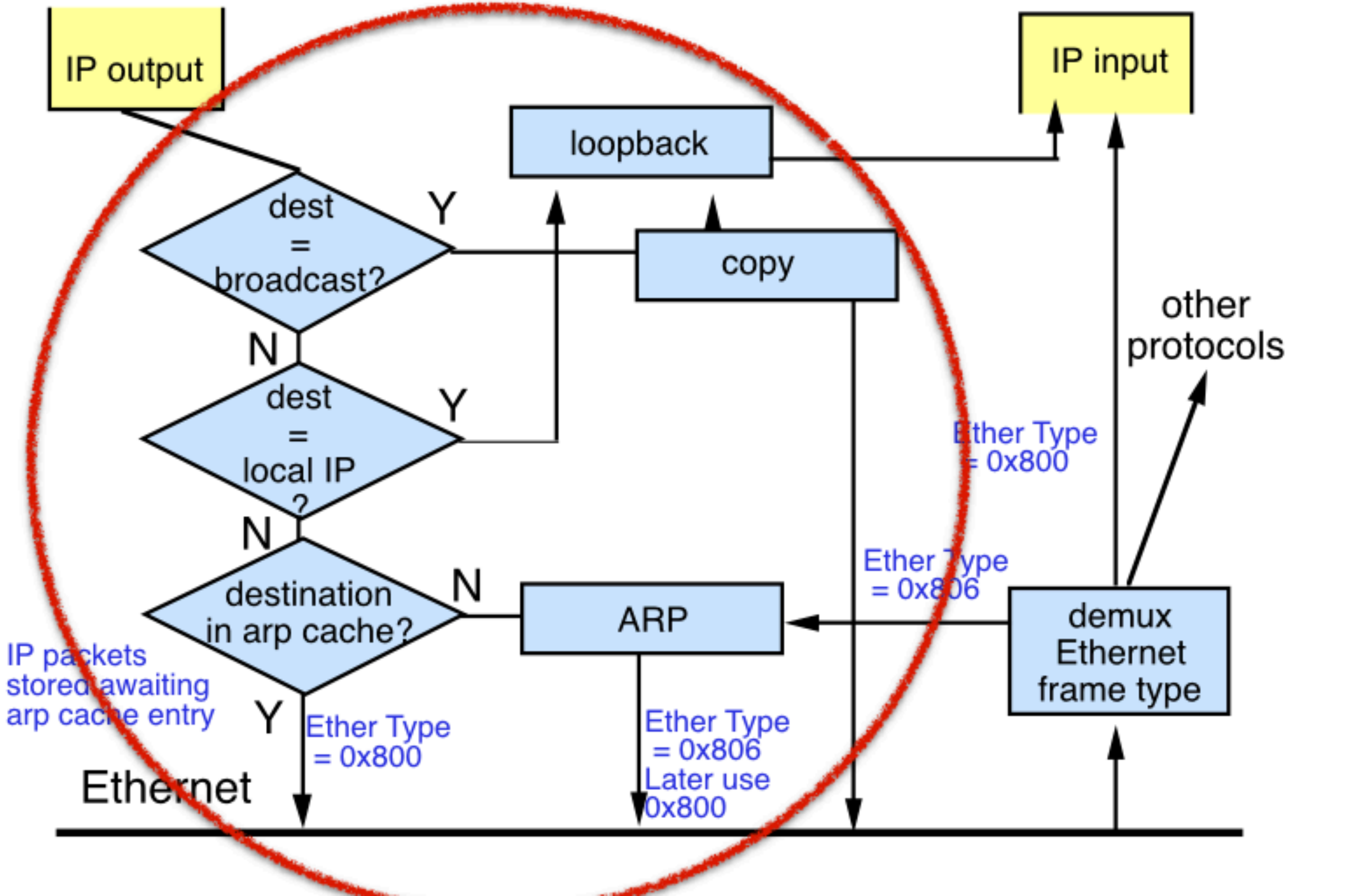
| operation | message |
|---|---|
| 1 | ARP request |
| 2 | ARP reply |

Ether Type = 0x806

RFC 826

# Ethernet Driver

G Fairhurst, http://www.erg.abdn.ac.uk



# ARP Frame

G Fairhurst, http://www.erg.abdn.ac.uk

```
001a 2f52 4841 000a 95cf ea5e 0806 0001
0800 0604 0002 000a 95cf ea5e 8b85 cf98
001a 2f52 4841 8b85 cf40
```

Must be a reply - it is an Ethernet unicast frame:

```
001a 2f52 4841 000a 95cf ea5e 0806 0001
0800 0604 0002 000a 95cf ea5e 8b85 cf98
001a 2f52 4841 8b85 cf40
```

Type code = 2, Target address was completed

```
001a 2f52 4841 000a 95cf ea5e 0806 0001
0800 0604 0002 000a 95cf ea5e 8b85 cf98
001a 2f52 4841 8b85 cf40
```

# ARP Frame

G Fairhurst, http://www.erg.abdn.ac.uk

```
001a 2f52 4841 000a 95cf ea5e 0806 0001
0800 0604 0002 000a 95cf ea5e 8b85 cf98
001a 2f52 4841 8b85 cf40
```

Source MAC address: 000a 95cf ea5e

Destination MAC address: 001a 2f52 4841

Ether Type: 0x0806 - arp message

# Example decode

```
001a 2f52 4841 000a 95cf ea5e 0806 0001
0800 0604 0002 000a 95cf ea5e 8b85 cf98
001a 2f52 4841 8b85 cf40
```

Decode of arp mesage:

Hardware Type = 0001 (Ethernet)

Protocol= 0800 (IP)

Hardware Length = 6 bytes

Protocol Length (IP address) = 4 bytes

Operation = 0002 (arp response)

SRC HA = 000a 95cf ea5e

SRC IP = 139.133.207.152 (0x8b85cf98)

DST HA = 001a 2f52 4841

DST IP = 139.133.207.64 (0x8b85cf40)

# Another ARP Frame

G Fairhurst, http://www.erg.abdn.ac.uk

What type is this?

```
ffff ffff ffff 0030 653a f222 0806 0001
0800 0604 0001 0030 653a f222 8b85 cf45
0000 0000 0000 8b85 cf40
```

# Another ARP Frame

G Fairhurst, http://www.erg.abdn.ac.uk

Three hints: Broadcast, Type, and Unknown address

```
ffff ffff ffff 0030 653a f222 0806 0001
0800 0604 0001 0030 653a f222 8b85 cf45
0000 0000 0000 8b85 cf40
```

Ethernet frame:
00:30:65:3a:f2:22 > Broadcast
Ethertype ARP (0x0806)

arp request:
arp who-has 139.133.207.64 tell 139.133.207.69

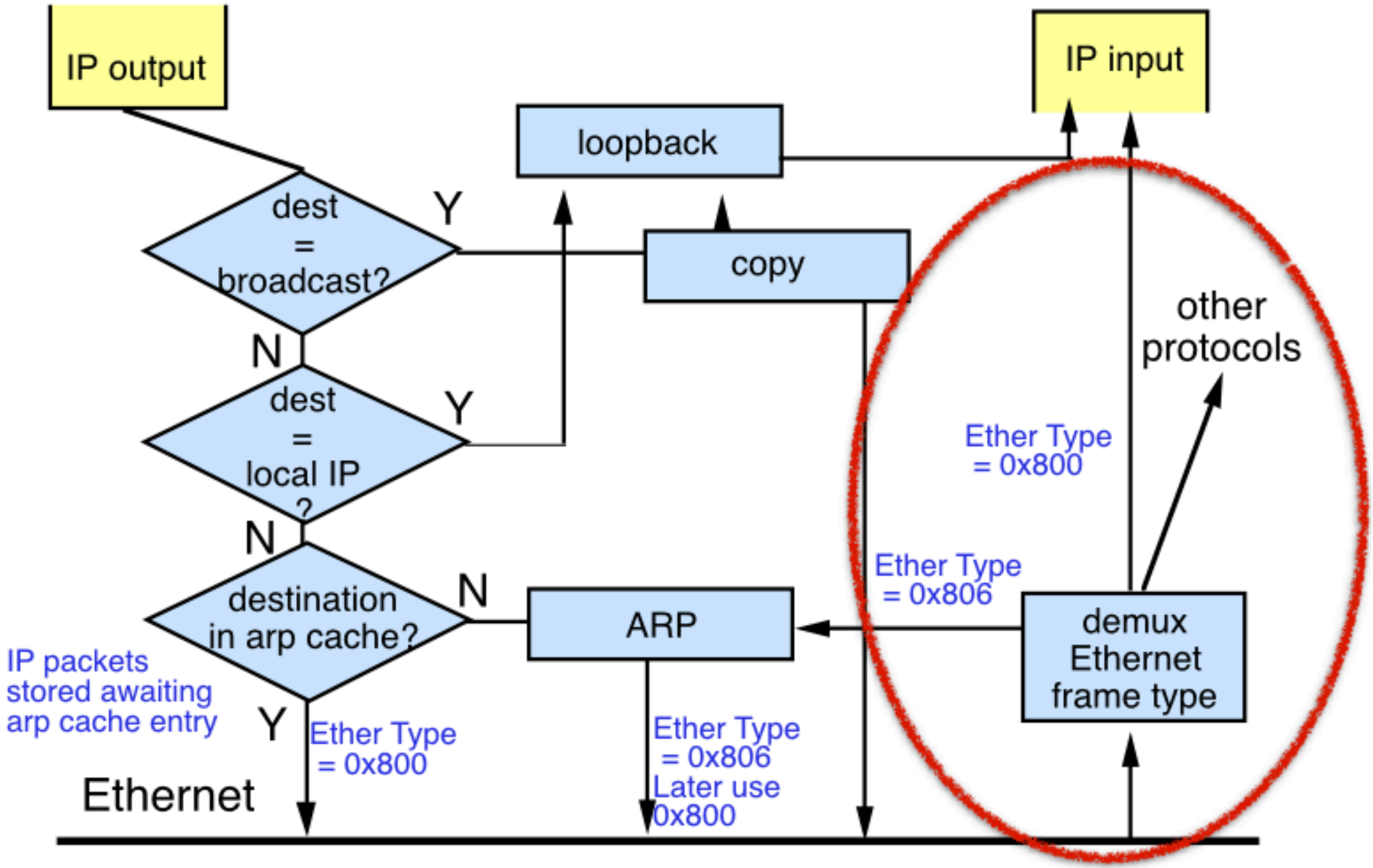
G Fairhurst, http://www.erg.abdn.ac.uk

# Address Resolution Protocol (arp)

ARP Cache

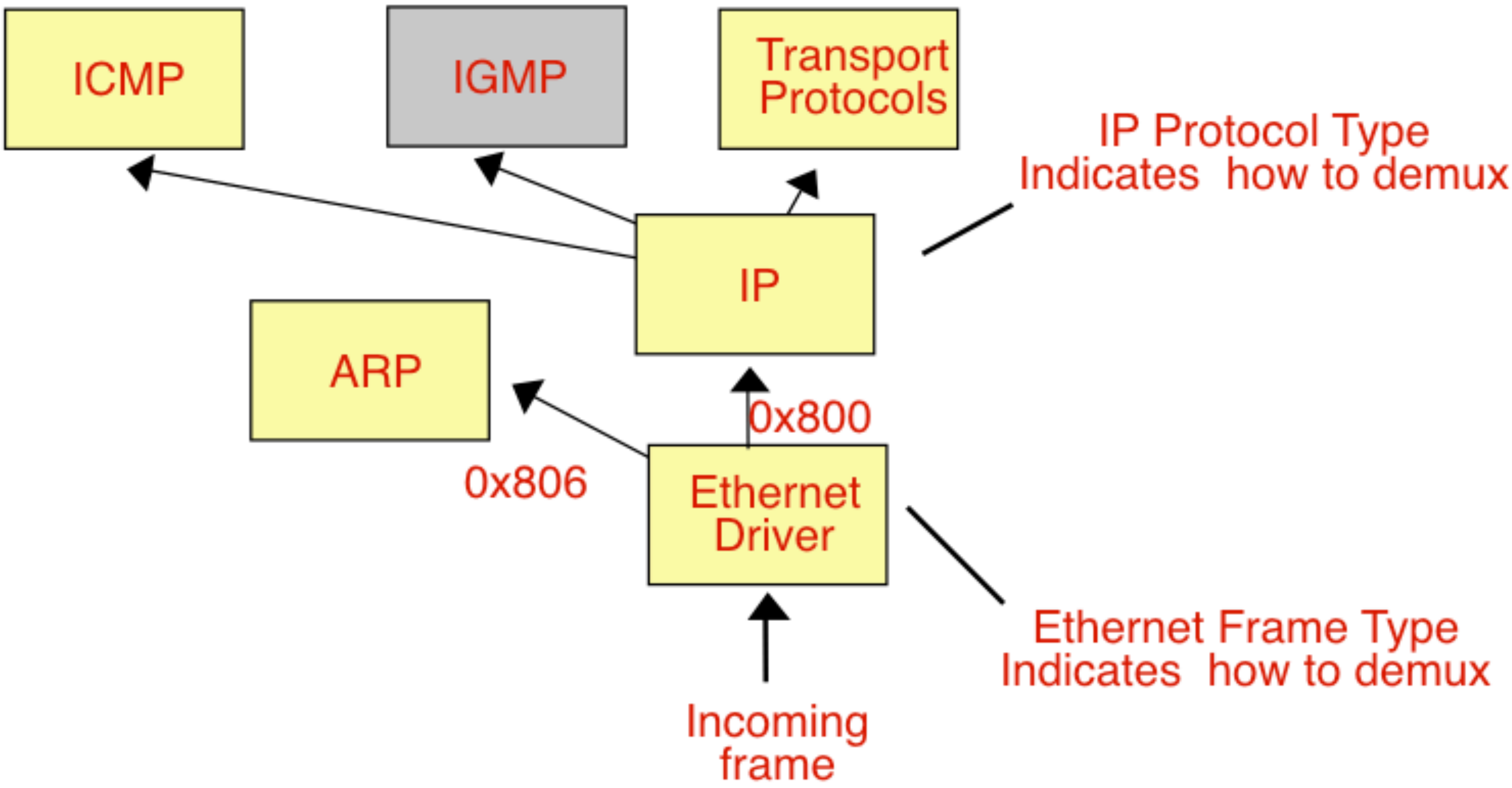
# Receiving a Packet

G Fairhurst, http://www.erg.abdn.ac.uk



# Protocol Demultiplexing

G Fairhurst, http://www.erg.abdn.ac.uk



# ARP Cache

G Fairhurst, http://www.erg.abdn.ac.uk

Senders use arp to find Target-IP's MAC addresses

An arp cache is needed to prevent overload!!

Stores IP & corresponding MAC addresses
Updated each time ***any*** query is seen
(Note: Other people's queries also fill the cache)

The arp cache entries expire after a fixed period

Significantly reduces the number of arp messages needed to run a large network!

# ARP Summary

G Fairhurst, http://www.erg.abdn.ac.uk

Senders know:

- IP source address (may use DHCP)
- IP destination address (may use DNS)
- .... and hence the Target-IP of the next-hop system
- MAC source address (may look in NIC ROM)

Senders use arp to find Target-IP's MAC addresses

An arp cache is needed to prevent overload!!

ARP is ***automatic*** when each IP packet is sent

# ARP Example

G Fairhurst, http://www.erg.abdn.ac.uk

## Use the "arp -a" command to examine ARP cache.

```
gresley:arp -a
milliways-mac.erg.abdn.ac.uk (139.133.207.64) at 0:d0:bb:f7:c6:c1 on en0
mavis-mac.erg.abdn.ac.uk (139.133.207.77) at 8:0:20:86:ec:df on en0
```

The cache consists of a table of address and bindings
There are currently two entries

# ARP Example

G Fairhurst, http://www.erg.abdn.ac.uk

## Use the “ping” command to send test packets

Each time a packet is made, arp is triggered as necessary to find the target-IP’s mac address.

1) IP packets sent to 139.133.207.111 were received an generate replies.

2) IP packets sent to 139.133.207.222 generate no replies, we can assume this address is not in use.

```
gresley:ping 139.133.207.111
PING 139.133.207.111 (139.133.207.111): 56 data bytes
64 bytes from 139.133.207.111: icmp_seq=0 ttl=60 time=1.732 ms
...

gresley:ping 139.133.207.222
PING 139.133.207.222 (139.133.207.222): 56 data bytes
ping: sendto: Host is down
ping: wrote 139.133.207.222 64 chars, ret=-1
...
```

# ARP Example

G Fairhurst, http://www.erg.abdn.ac.uk

```
gresley:arp -a
milliways-mac.erg.abdn.ac.uk (139.133.207.64) at 0:d0:bb:f7:c6:c1 on en0
mavis-mac.erg.abdn.ac.uk (139.133.207.77) at 8:0:20:86:ec:df on en0
```

```
gresley:ping 139.133.207.111
PING 139.133.207.111 (139.133.207.111): 56 data bytes
64 bytes from 139.133.207.111: icmp_seq=0 ttl=60 time=1.732 ms
...
```

```
gresley:ping 139.133.207.222
PING 139.133.207.222 (139.133.207.222): 56 data bytes
ping: sendto: Host is down
ping: wrote 139.133.207.222 64 chars, ret=-1
...
```

```
gresley:arp -a
milliways-mac.erg.abdn.ac.uk (139.133.207.64) at 0:d0:bb:f7:c6:c1 on en0
mavis-mac.erg.abdn.ac.uk (139.133.207.77) at 8:0:20:86:ec:df on en0
erg2-printer.erg.abdn.ac.uk (139.133.207.111) at 0:10:83:ba:c0:a5 on en0
? (139.133.207.222) at (incomplete) on en0 [ethernet]
```

The arp cache has two new entries:

139.133.207.111 has MAC: 0:10:83:ba:c0:a5

139.133.207.222 did not respond (no cache entry)

# ARP Question

G Fairhurst, http://www.erg.abdn.ac.uk


Two 10 Mbps Ethernet LANs are connected by a bridge.
When monitoring LAN A for 1 minute,
40 arp requests are observed and 30 arp responses.

(a) Calculate the Utilisation for the arp frames for LAN A.

(b) Give two reasons why there are fewer responses than queries.

# ARP Question

G Fairhurst, http://www.erg.abdn.ac.uk

Two 10 Mbps Ethernet LANs are connected by a bridge. When monitoring LAN A for 1 minute, 40 arp requests are observed and 30 arp responses.

Calculate the Utilisation for the arp packets for LAN A.

Size of ARP request/Response is
=8+14+28+4 (less than minimum Enet PDU) => 8+64 B

= (70/60) x 8x72/10^7x100 % = 0.007%

Give two reasons why there may be fewer responses than queries.

(1) Some arp requests fail to complete (IP addr not used)
(2) Some arp requests may have been sourced on LAN B and correspond to an IP address on LAN B. The response would not travel across the bridge.

# ARP/DHCP Packet

G Fairhurst, http://www.erg.abdn.ac.uk

48 bit
Ether hardware address

32 bit
IP target address

Where are my friends?

32 bit
IP source address

48 bit
Ether hardware address

Who am I?

RFC 2131

A6

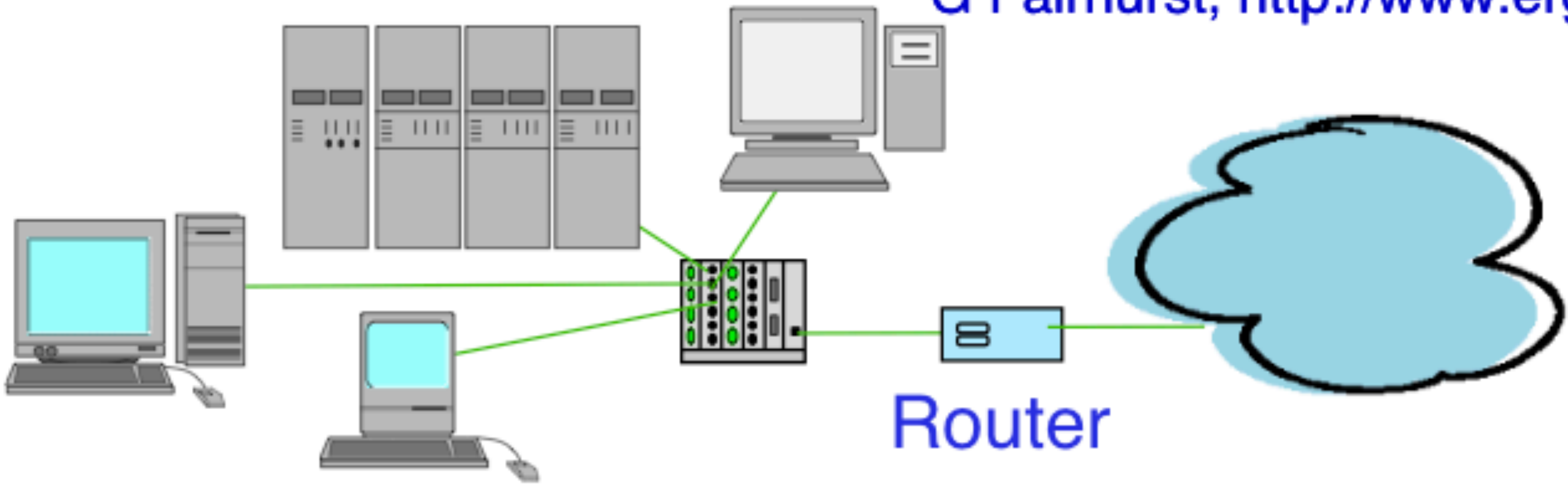
# Dynamic Host Configuration

IP Addresses

DHCP

# IP Address Allocation

G Fairhurst, http://www.erg.abdn.ac.uk




Addresses allocated to network as a ***flat*** address block

e.g. Aberdeen University allocated 139.133.x.x/16

i.e. addresses start with the same address prefix

e.g. 139.133.1.5, 139.133.208.1

Each system gets at least one IP address per interface

# Configuring Clients

G Fairhurst, http://www.erg.abdn.ac.uk

All network clients need to be configured:

MAC source address (may look in NIC ROM)
- Their own (source) IP address
- The Subnet mask
- The IP address of the default router
- List of IP address for local DNS servers

We could do this by hand!
... **but** in practice need a better way

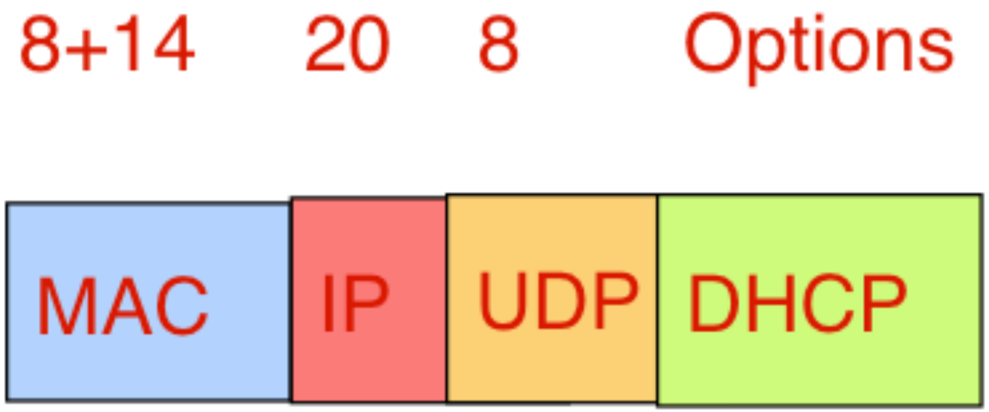
# Dynamic Host Configuration Protocol

G Fairhurst, http://www.erg.abdn.ac.uk

DHCP Server setup with a pool of addresses
This often executes on a router

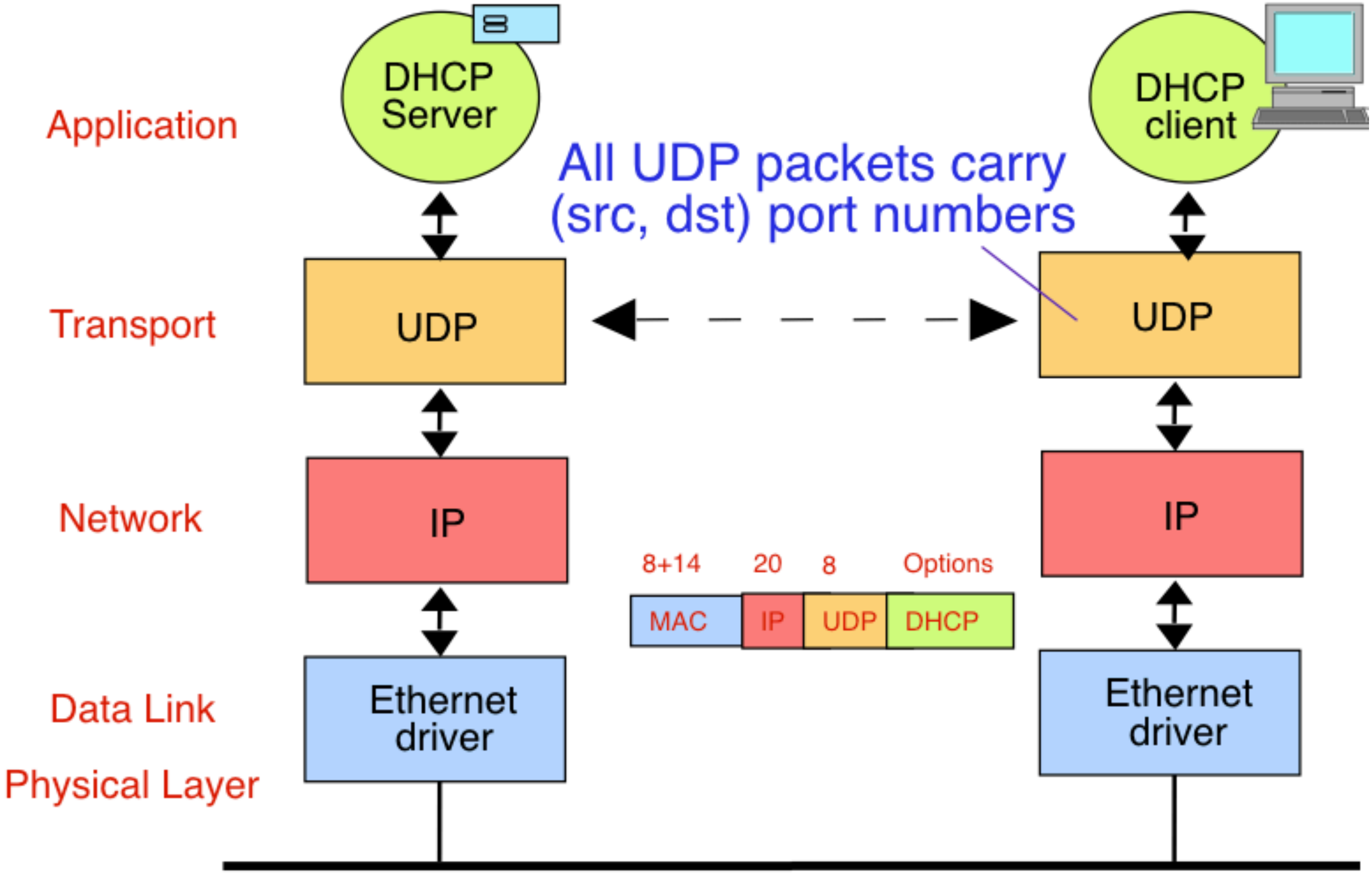
DHCP packets are sent using IP

Each DHCP packet uses UDP
Sent to the well-known DHCP port
Port 67 to server; Port 68 to client



# DHCP

G Fairhurst, http://www.erg.abdn.ac.uk



# *Dynamic Host Configuration Protocol*

G Fairhurst, http://www.erg.abdn.ac.uk

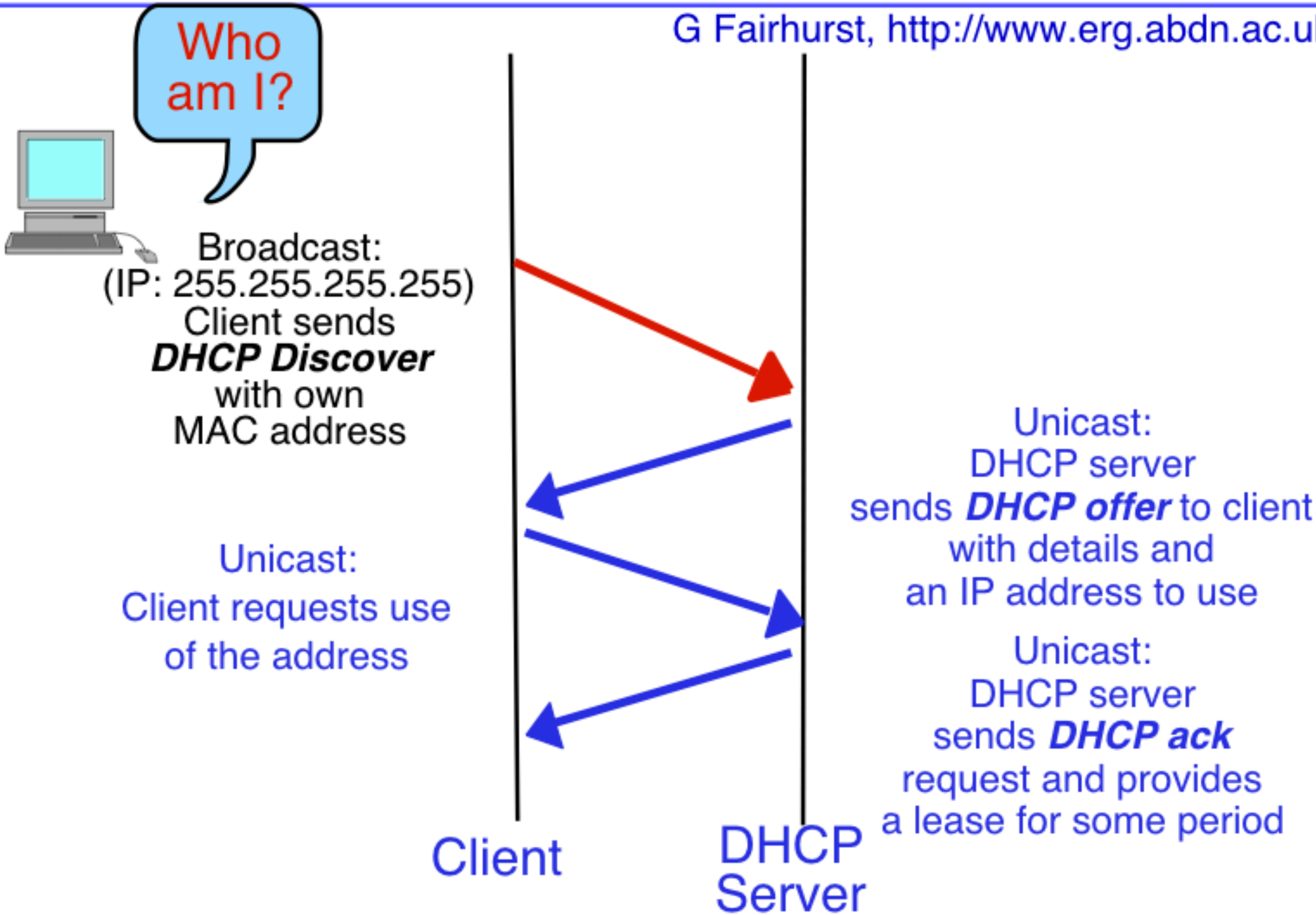
A client requests an IP address using ***DHCP***

***Sends*** its MAC address to the DHCP server (router)
- uses an IPv4 broadcast packet

Client offered an IPv4 address
Addresses may be loaned (for some time)
or static-assigned to a specific MAC address

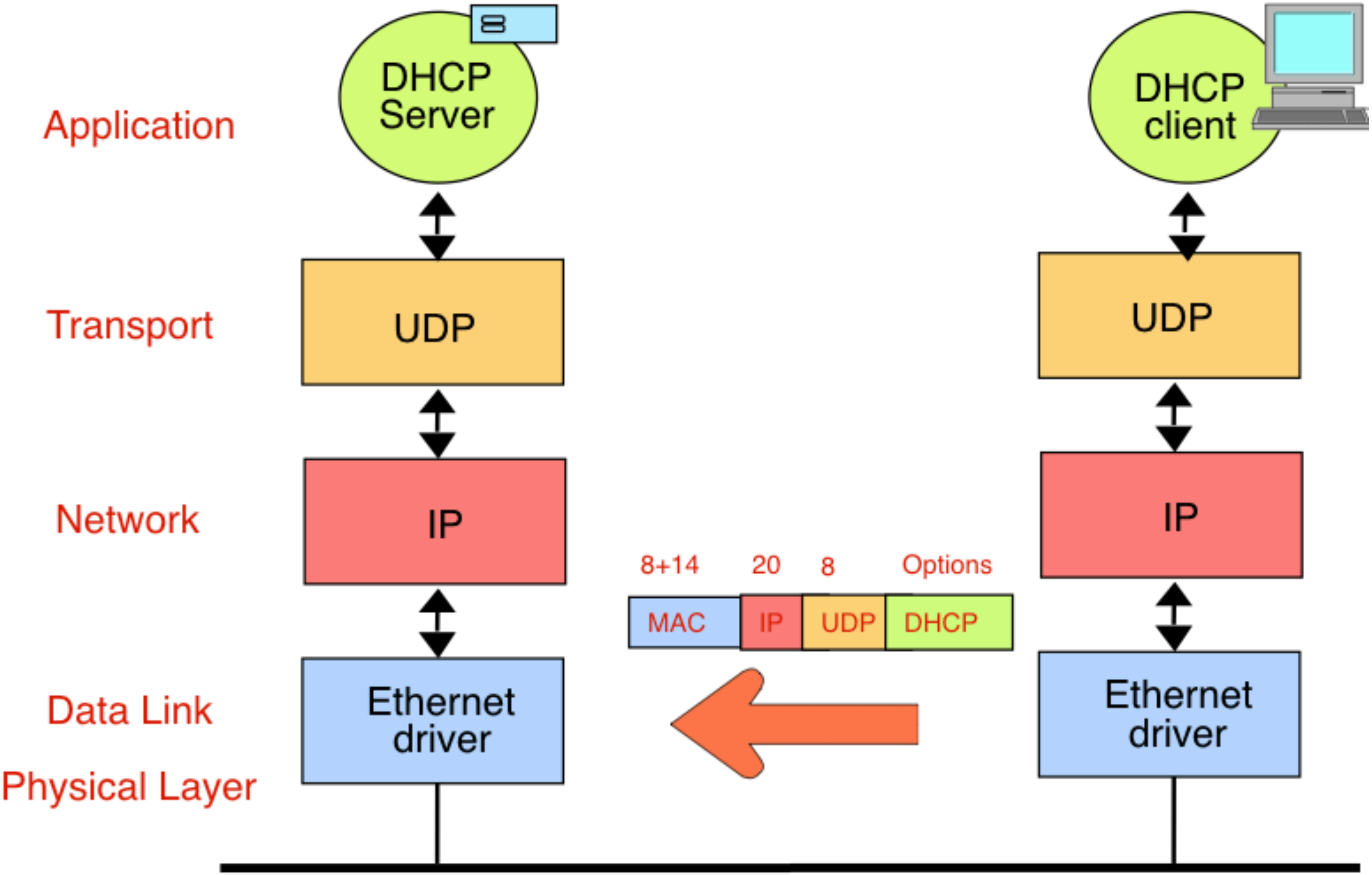
# DHCP Protocol for IPv4

G Fairhurst, http://www.erg.abdn.ac.uk



# DHCP Discover

G Fairhurst, http://www.erg.abdn.ac.uk



# *DHCP Discover message*

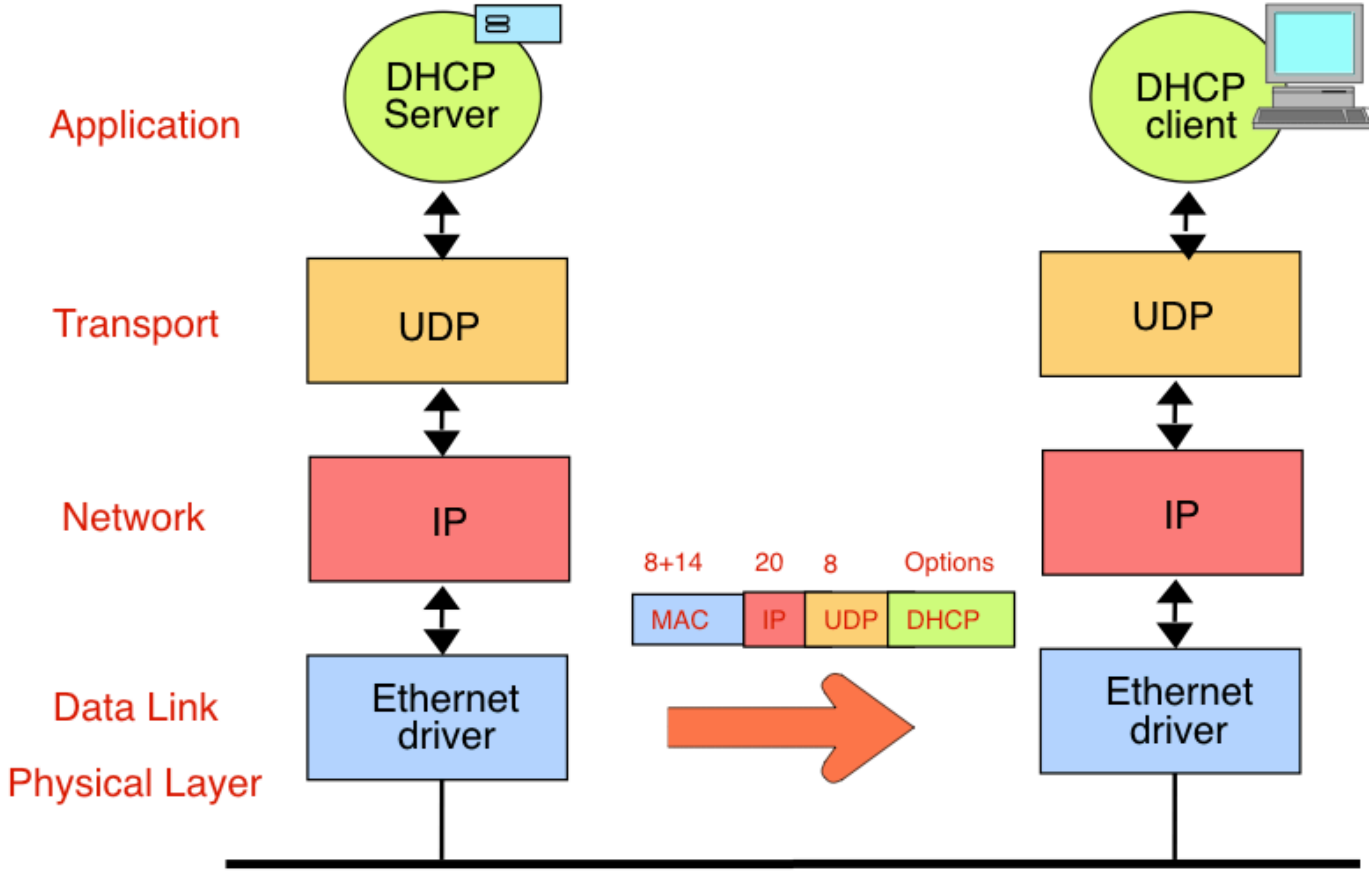
G Fairhurst, http://www.erg.abdn.ac.uk

Clients broadcast to LAN to *Discover* DHCP server
- includes own MAC address & "Magic Cookie"

Sent in a IP packet to "IP broadcast address"
- All clients and servers receive this!!
- This IP destination address is 255.255.255.255
- at the MAC level this dest is ff:ff:ff: ff:ff:ff

# DHCP Offer

G Fairhurst, http://www.erg.abdn.ac.uk



# *DHCP Offer message*



One or more DHCP Server responds with a *DHCP Offer*:

IP address that may be used;

IP Subnet mask (1);

IP address of default router (3);

IP address of DNS server (6);

IP lease time (51)

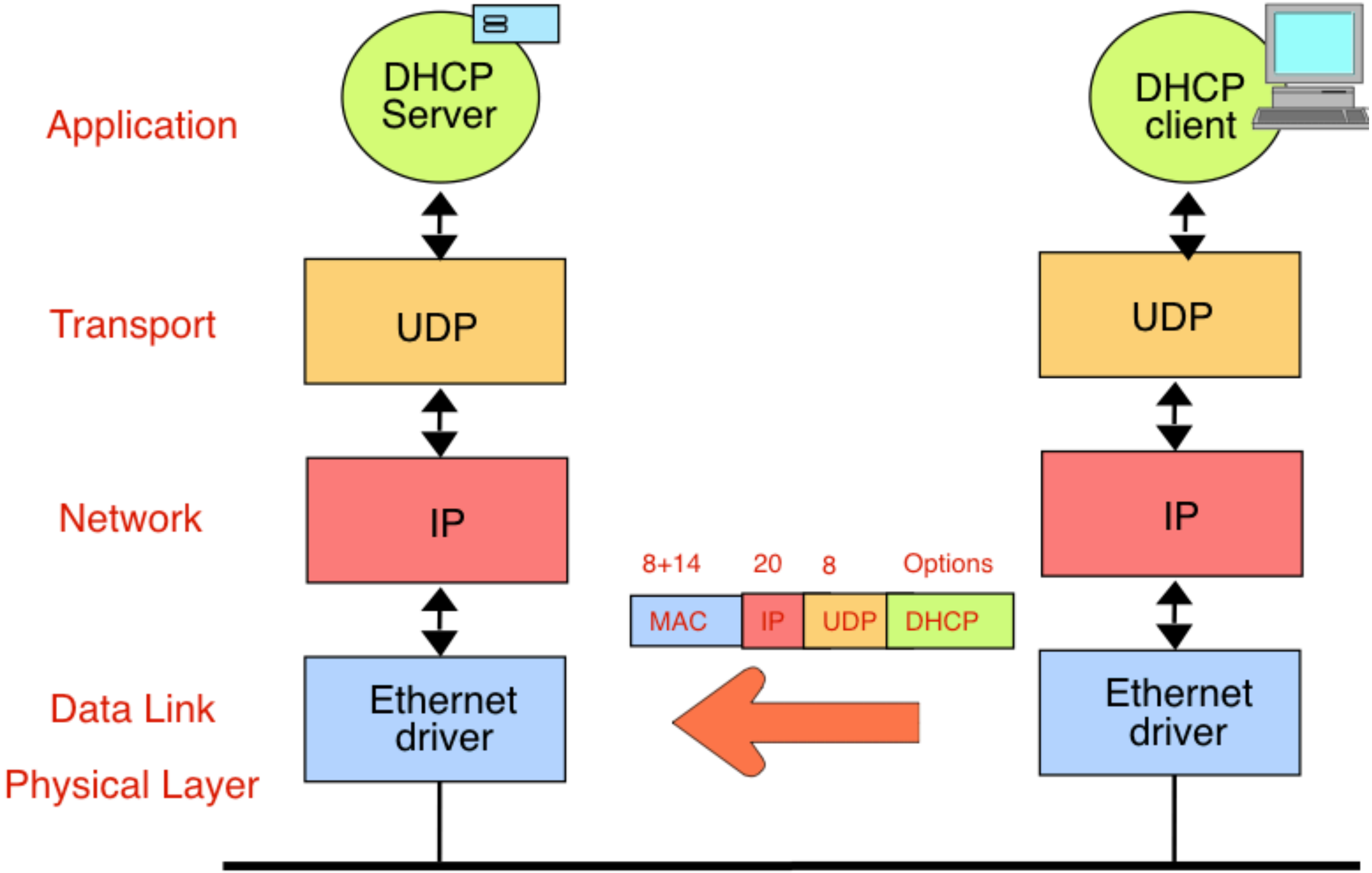
IP address of DHCP server (54);

... there may be other options also...

“Magic Cookie” - nonce to identify request at server

# DHCP Request

G Fairhurst, http://www.erg.abdn.ac.uk



# *DHCP Request/ACK*

G Fairhurst, http://www.erg.abdn.ac.uk

Several DHCP servers could offer - one is chosen:

Client responds to ONE server with a *DHCP Request*

- *Broadcast (because client still has no IP address yet)*

Server responds with a *DHCP Acknowledgment*

Value used only for a specified period (lease interval)

# DHCP Attacks



Several DHCP servers could offer....

**Bad** thing may happen with a rogue “server”

...could give a system an invalid address (blackhole)

...could give it some other system’s address (see later)

...could make itself the “gateway router”
it would then get a copy of all your packets...

Be aware, and don’t be stupid and do this by accident!

DHCP Server gives all information needed to use network!

- Your own IP Address
- The network subnet mask
- The address of your local router
- The address of the DNS server
- and more….

Information supplied valid for some time

- Could be days/weeks (in a company network)
- Could be minutes/hours in a WiFi hotspot

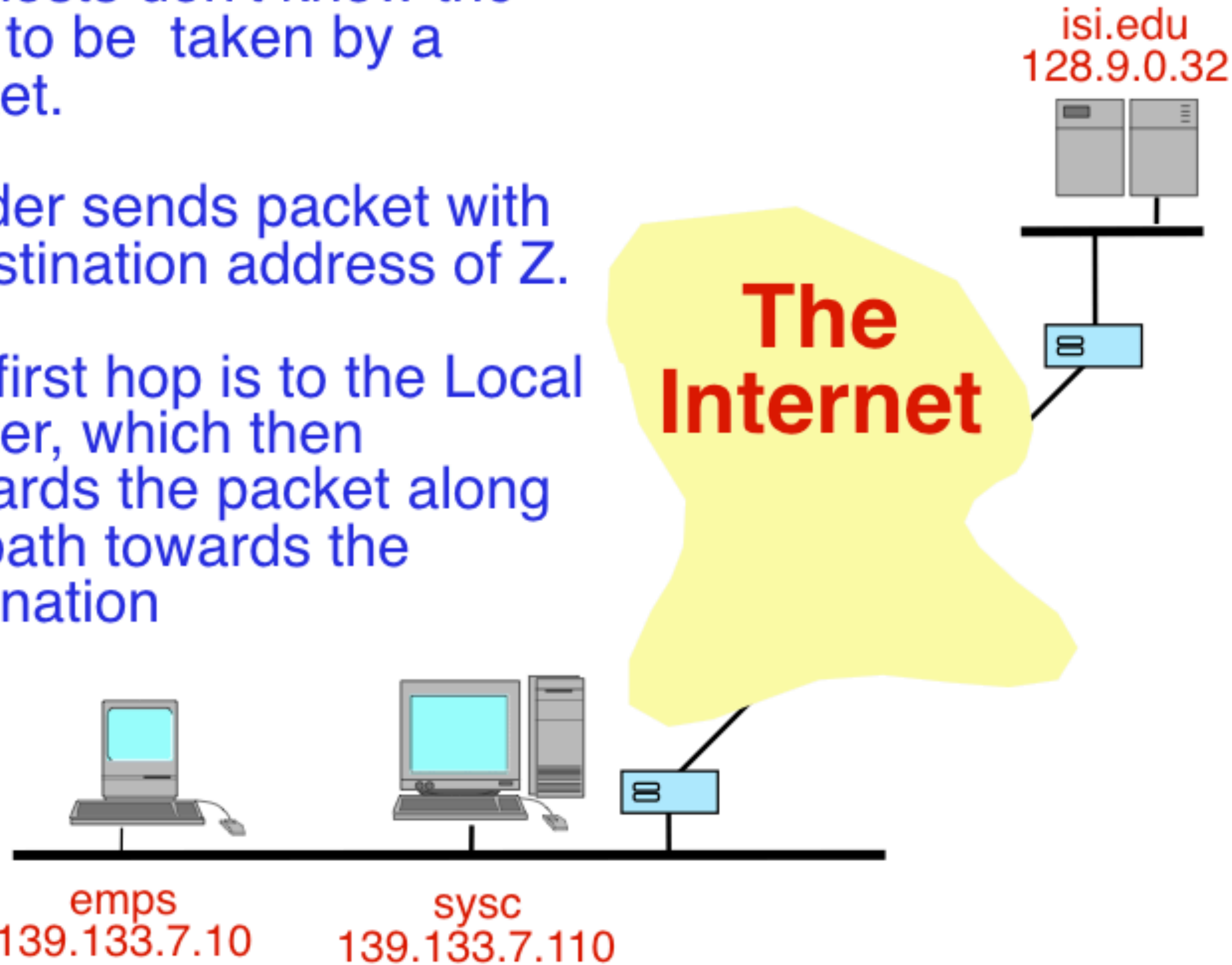
# Routers (L3)

G Fairhurst, http://www.erg.abdn.ac.uk

Endhosts don't know the path to be taken by a packet.

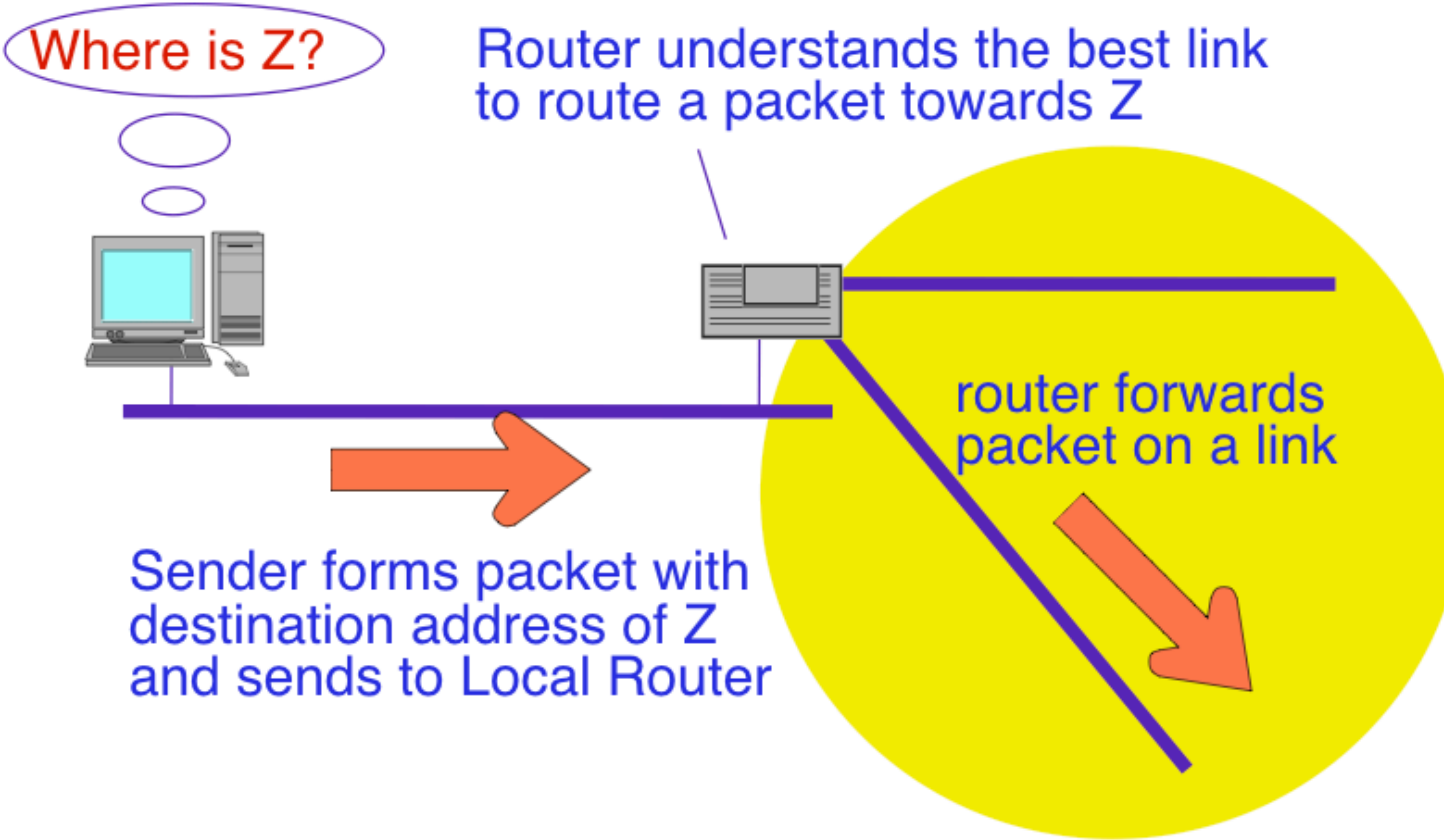
Sender sends packet with a destination address of Z.

The first hop is to the Local Router, which then forwards the packet along the path towards the destination



# Escaping from the LAN

G Fairhurst, http://www.erg.abdn.ac.uk



# Role of Routers

G Fairhurst, http://www.erg.abdn.ac.uk

Routers have multiple interfaces each connecting a LAN link

Connecting network segments
- Each interface configured with an IP address

Buffers Packets
- Configured with quality of Service

Security
- Configured with filters for addresses

RFC 1812

# Bridges v. Routers

G Fairhurst, http://www.erg.abdn.ac.uk

**Routers**

- ***Connect networks***
- ***Control traffic flow between networks***

More expensive

Work at Network Layer (e.g. IP)

Connect different IP networks

Need configuration

**Bridges/Switches**

- ***Separate work group traffic***
- ***Improve LAN performance***

Cheap

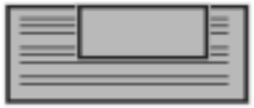
Work at MAC Layer (mostly self configuring)

Form one IP network (broadcast domain at L2)

# About An IP Network

G Fairhurst, http://www.erg.abdn.ac.uk

**End Systems send packet to an IP address**
know nothing about the network topology
"conversations" identified by their ***port number***



**Routers use IP address to forward packets**
know nothing about 'conversations'
recipients identified by their ***IP address***

G Fairhurst, http://www.erg.abdn.ac.uk

# ICMP

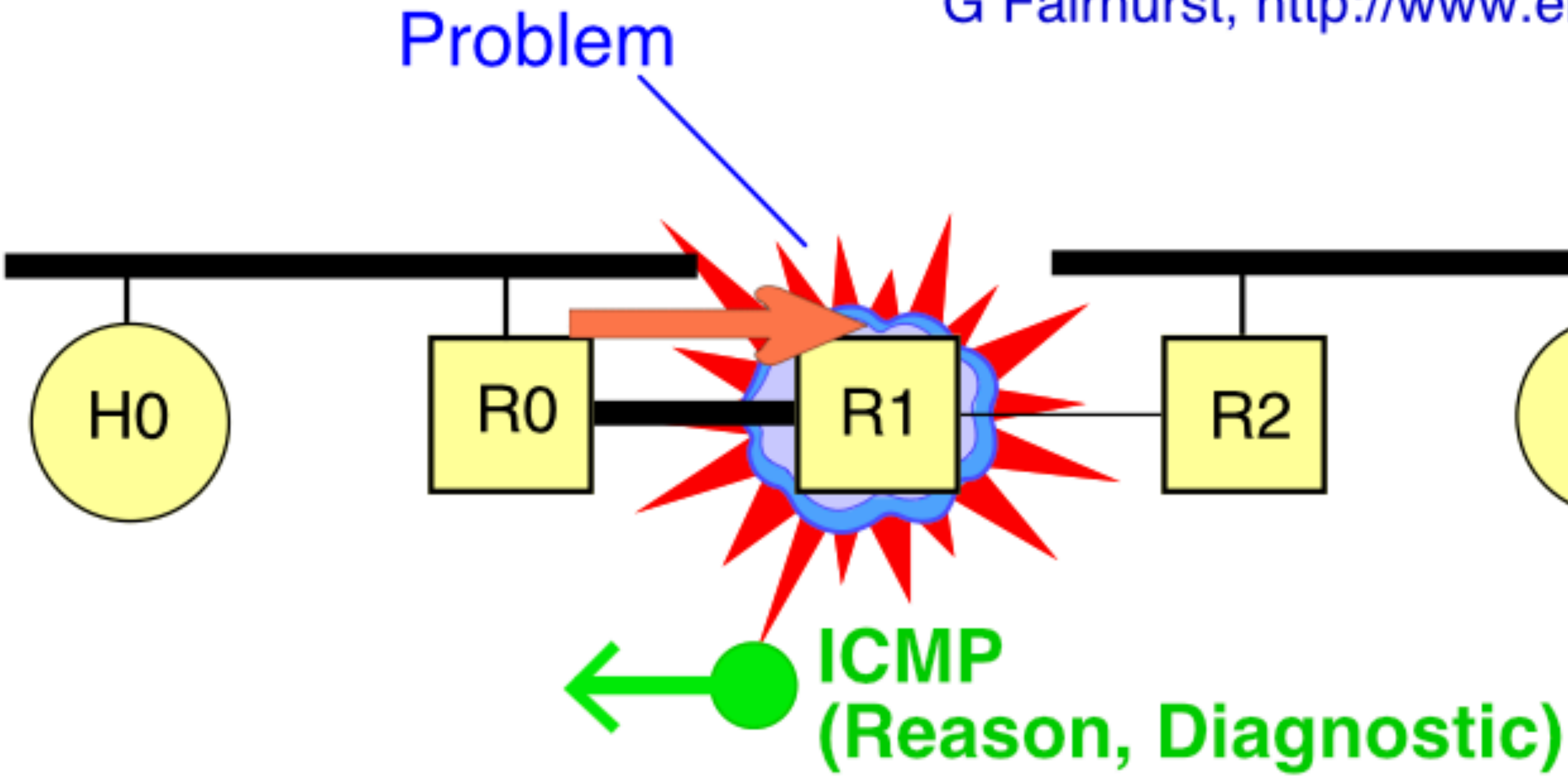
Internet Control Message Protocol

ICMP Encapsulation

Ping and ICMP Echo

# Internet Control Message Protocol

G Fairhurst, http://www.erg.abdn.ac.uk



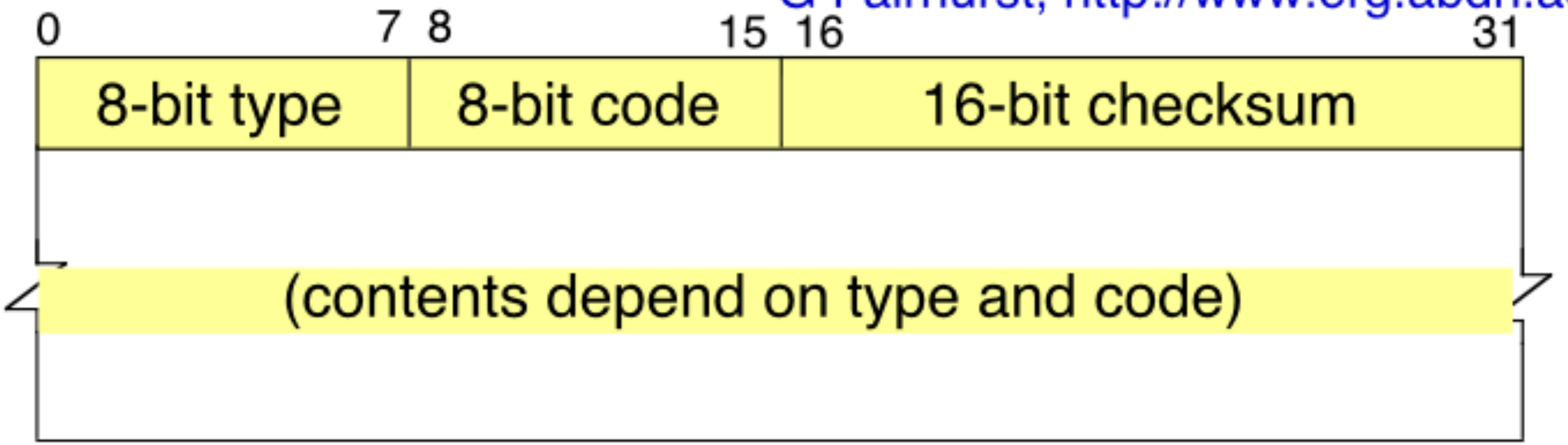

Routers / Computers send ICMP messages

Messages usually contain the header of the packet

Not usually sent when ICMP messages received
(An exception is an ICMP ECHO REQUEST)

# ICMP Message

G Fairhurst, http://www.erg.abdn.ac.uk

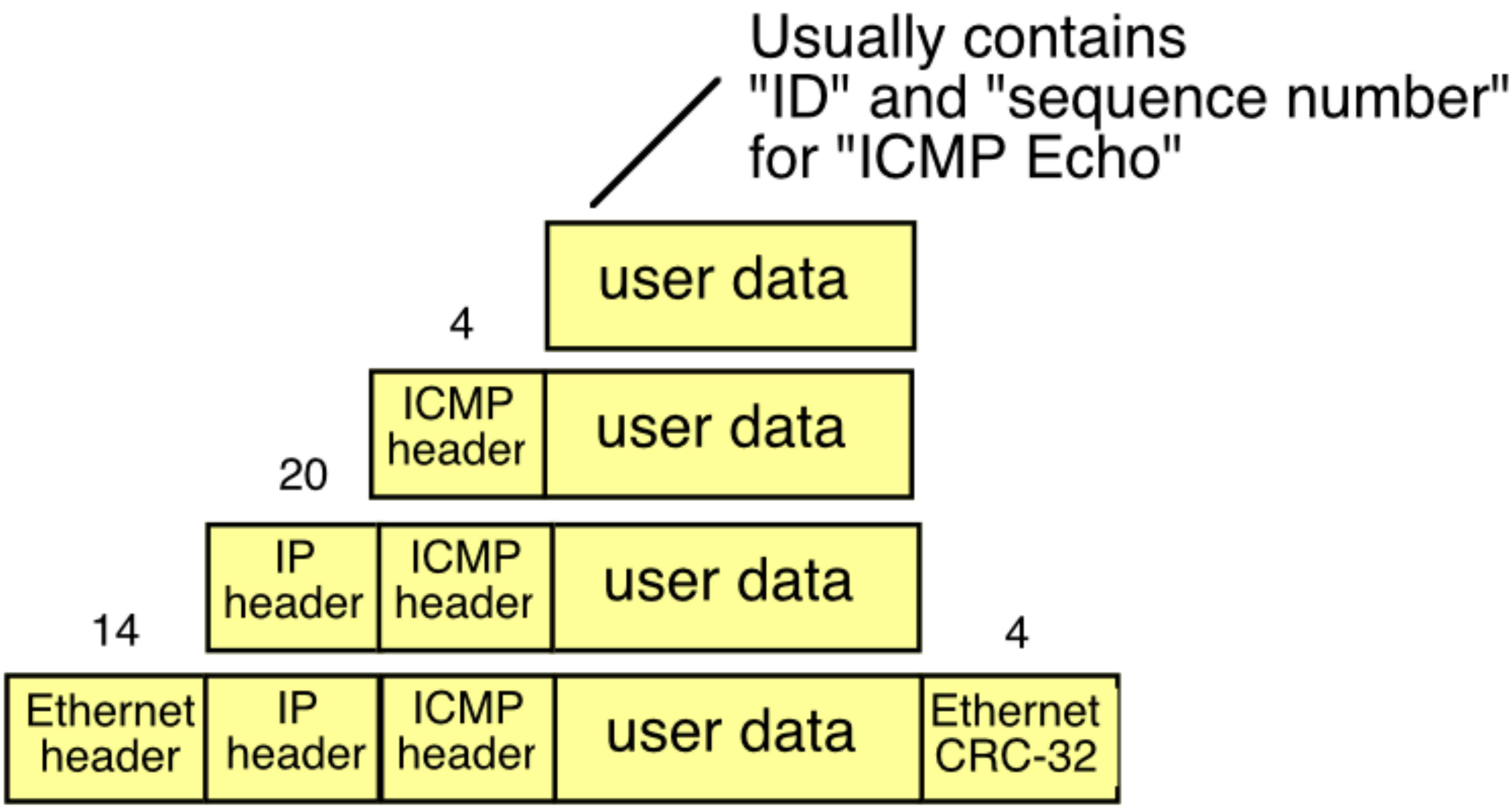



| Type | Message |
|---|---|
| 0 | Echo reply |
| 3 | Destination unreachable |
| 4 | Source quench |
| 5 | Redirect |
| 8 | Echo request |
| 11 | Time exceeded (i.e. TTL=0) |

Postel, J., "Internet Control Message Protocol", RFC792, STD 5, 1981.

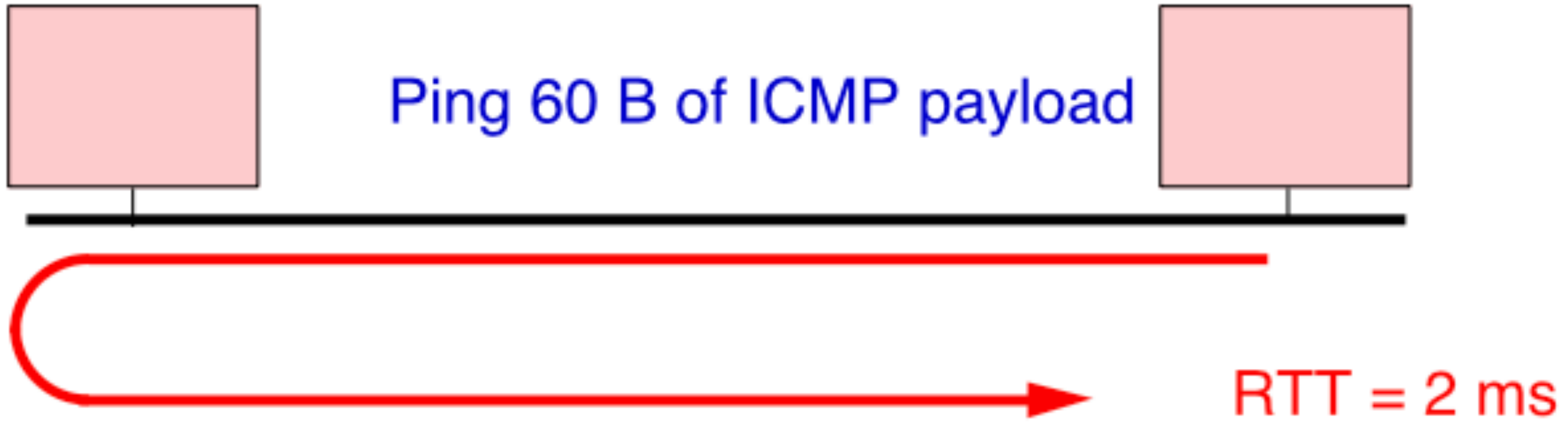
# ICMP Encapsulation

G Fairhurst, http://www.erg.abdn.ac.uk



# Ping of Local Host

G Fairhurst, http://www.erg.abdn.ac.uk




```
ping -s sysb
PING sysb: 56 data bytes
64 bytes from sysb (139.133.201.196): icmp_seq=0. time=3. ms
64 bytes from sysb (139.133.201.196): icmp_seq=1. time=3. ms
64 bytes from sysb (139.133.201.196): icmp_seq=2. time=3. ms
64 bytes from sysb (139.133.201.196): icmp_seq=3. time=3. ms
64 bytes from sysb (139.133.201.196): icmp_seq=4. time=3. ms
64 bytes from sysb (139.133.201.196): icmp_seq=5. time=3. ms
64 bytes from sysb (139.133.201.196): icmp_seq=6. time=3. ms
64 bytes from sysb (139.133.201.196): icmp_seq=7. time=3. ms
64 bytes from sysb (139.133.201.196): icmp_seq=8. time=3. ms
64 bytes from sysb (139.133.201.196): icmp_seq=9. time=4. ms
64 bytes from sysb (139.133.201.196): icmp_seq=10. time=5. ms
64 bytes from sysb (139.133.201.196): icmp_seq=11. time=3. ms
^C
----sysb PING Statistics----
12 packets transmitted, 12 packets received, 0% packet loss
round-trip (ms)  min/avg/max = 3/3/5
```

# Measuring Latency:Ping of Remote Host

G Fairhurst, http://www.erg.abdn.ac.uk

```
ping -s www.ksc.nasa.gov
PING zeno.ksc.nasa.gov: 56 data bytes
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=0. time=191. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=1. time=237. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=2. time=412. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=3. time=177. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=4. time=183. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=5. time=189. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=6. time=179. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=7. time=177. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=8. time=174. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=9. time=175. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=10. time=178. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=11. time=189. ms
64 bytes from zeno.ksc.nasa.gov (128.159.1.155): icmp_seq=12. time=322. ms
^C
----zeno.ksc.nasa.gov PING Statistics----
14 packets transmitted, 13 packets received, 7% packet loss
round-trip (ms)  min/avg/max = 174/214/412
```

*Not what it seems, no packet were really lost....*
*The sender sent 14 packets - each one second apart, and the user stopped the ping tool before the response to the last of these ICMP messages was received.*

G Fairhurst, http://www.erg.abdn.ac.uk

# Default Route

Subnetmasks

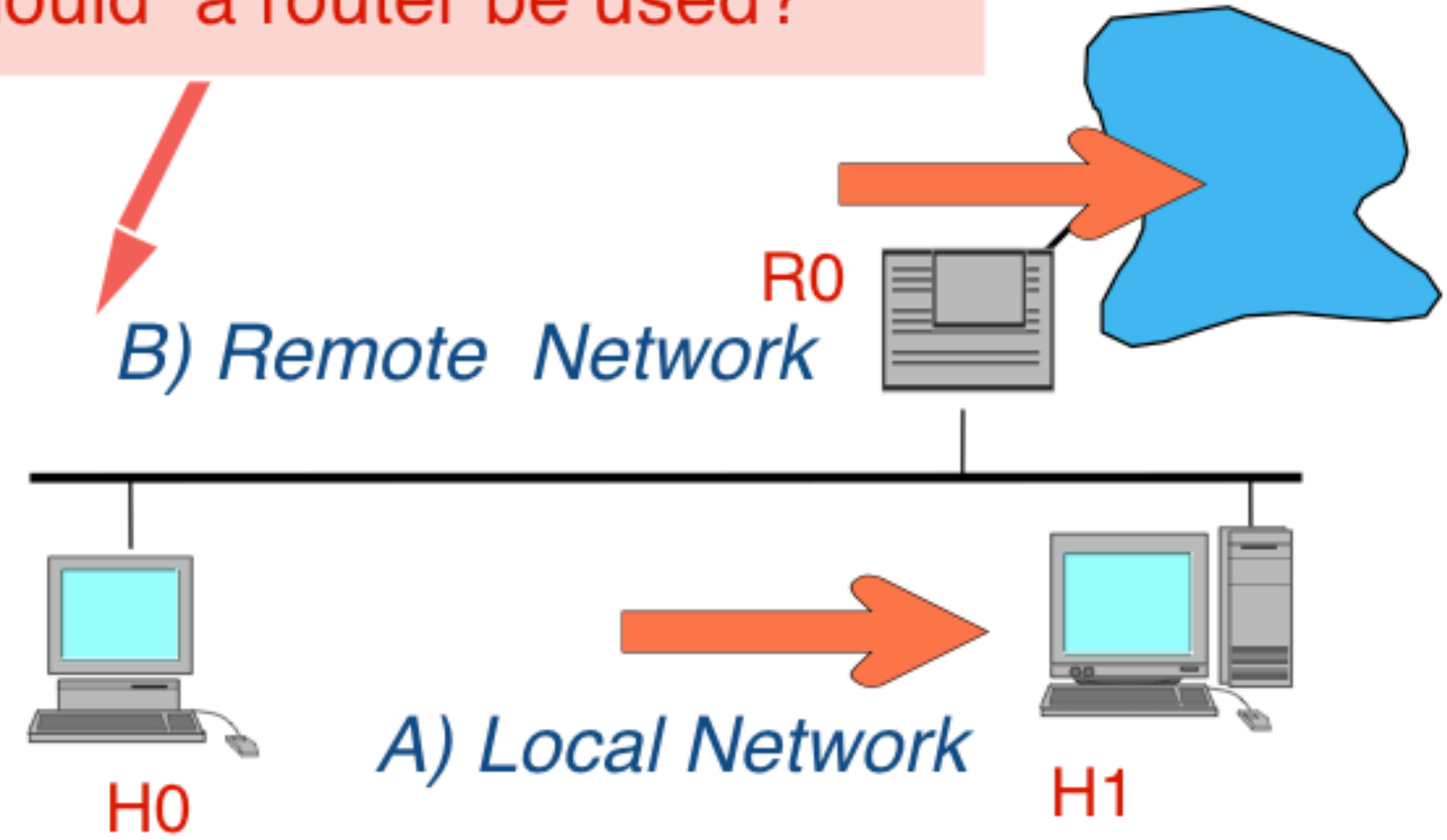
# Selecting a Route

G Fairhurst, http://www.erg.abdn.ac.uk

Should the Local Network be used?

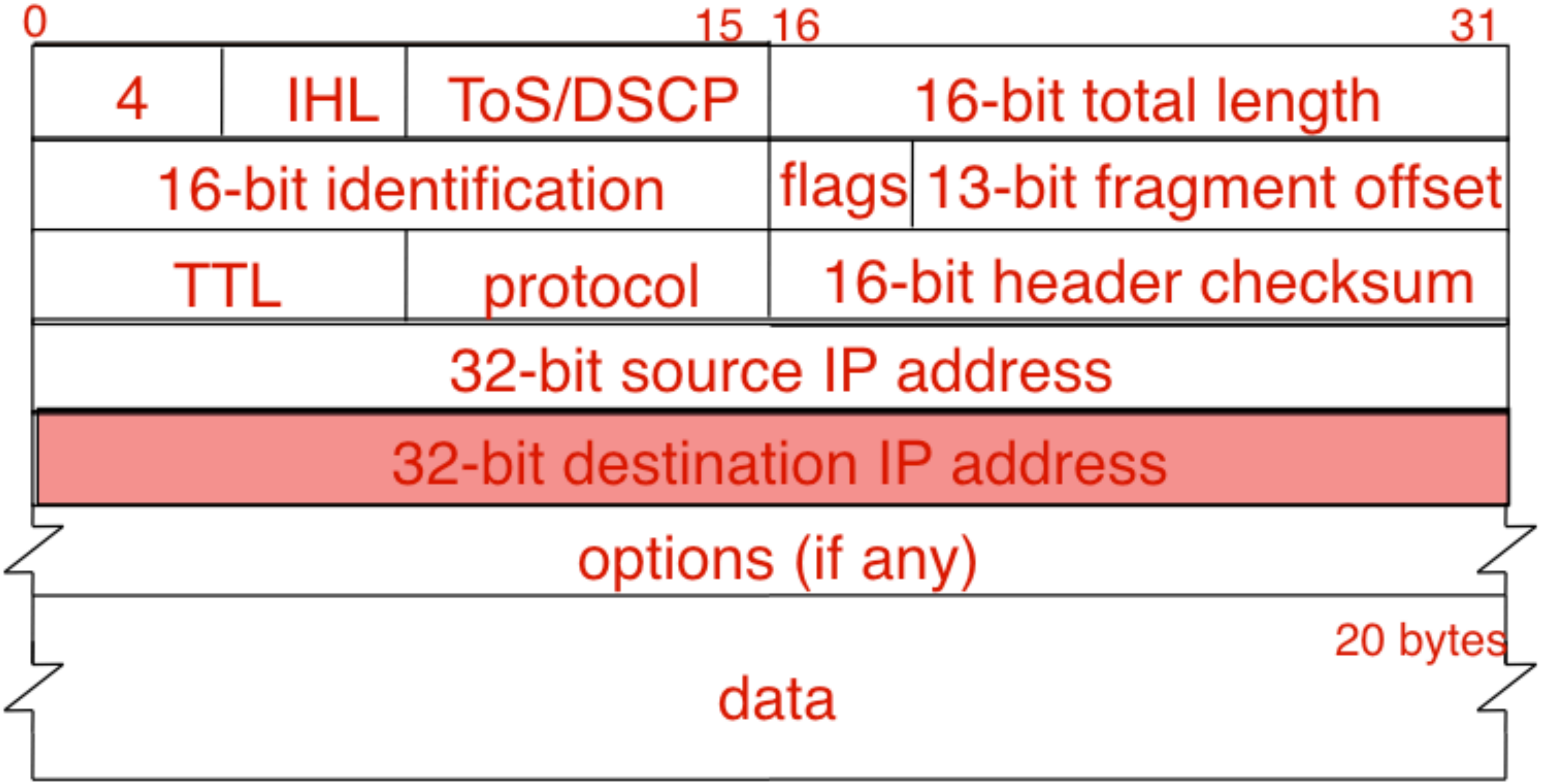
or

Should a router be used?



# IP Header

G Fairhurst, http://www.erg.abdn.ac.uk




ES and routers always examine the IP destination address

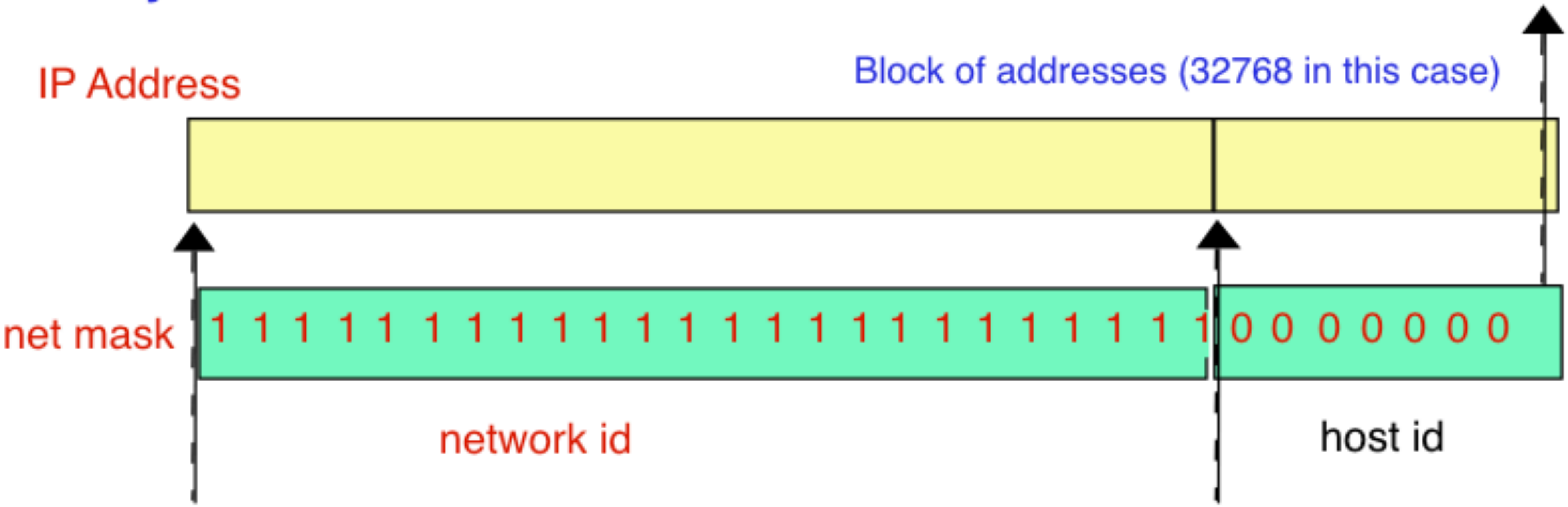
RFC 791

# IP Subnet Mask

G Fairhurst, http://www.erg.abdn.ac.uk

**A system need to know the *link netmask***

All systems in a subnet must share ***same*** subnet mask




```
IP address      139.133.7.110

netmask         0xffffff00 (255.255.255.0)

network ID      139.133.7.0/24
```

RFC 950

# IP Subnet Mask

G Fairhurst, http://www.erg.abdn.ac.uk

Subnet mask often written as a '/' followed by number of 1s in the mask, e.g. /8 = 8 '1s.

```
/8     11111111 00000000 00000000 00000000
/9     11111111 10000000 00000000 00000000
/10    11111111 11000000 00000000 00000000
/11    11111111 11100000 00000000 00000000
/12    11111111 11110000 00000000 00000000

/16    11111111 11111111 00000000 00000000
/20    11111111 11111111 11110000 00000000
/24    11111111 11111111 11111111 00000000
/28    11111111 11111111 11111111 11110000

/29    11111111 11111111 11111111 11111000
/30    11111111 11111111 11111111 11111100
/31    11111111 11111111 11111111 11111110
```

# Finding the Network ID



**Finding the network ID**

- Convert IP address to hexadecimal (or binary)
- Convert netmask to hexadecimal (or binary)
- Perform logical AND between the two 32b values

Example:

| | |
|---|---|
| IP address | 139.131. 63. 53 |
| & netmask | 255.255. 0 . 0 |
| Network id is: | 139.131. 0 . 0/16 |

# *Identifying the Destination Network (1)*

G Fairhurst, http://www.erg.abdn.ac.uk

*Local network calculation*

| | |
|---|---|
| local IP address | 139.133.7.110 |
| local network mask | 255.255.255.0 |
| local net +subnet id | 139.133.7.0 |

Compare

| | |
|---|---|
| dest IP address | 139.133.7.10 |
| local network mask | 255.255.255.0 |
| dest net +subnet id | 139.133.7.0 |

Match, therefore use local network

*Remote network calculation*

# Identifying the Destination Network (2)



*Local network calculation*

| | |
|---|---|
| local IP address | 139.133.7.110 |
| local network mask | 255.255.255.0 |
| local net +subnet id | 139.133.7.0 |

| | |
|---|---|
| dest IP address | 129.105.2.6 |
| local network mask | 255.255.255.0 |
| dest net +subnet id | 129.105.2.0 |

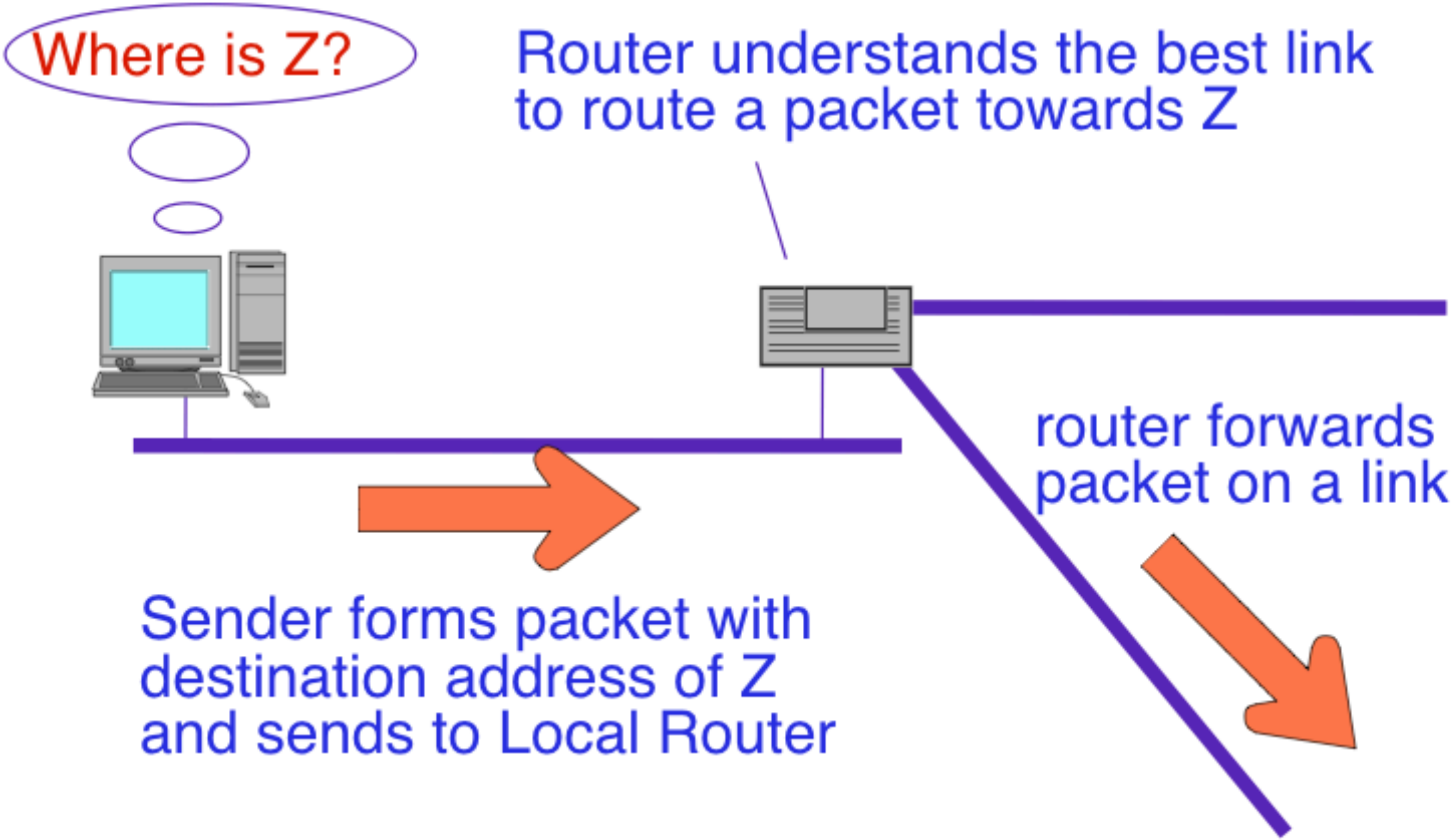
*Remote network calculation*

Compare

Differ, use a router

# Escaping from the LAN

G Fairhurst, http://www.erg.abdn.ac.uk



# *Finding the Broadcast Address*

G Fairhurst, http://www.erg.abdn.ac.uk

Broadcast address = the network ID + all 1's host ID

**Finding the broadcast address**

- Convert IP address to hex (or binary)
- Convert netmask to hex (or binary)
- Perform logical OR of the inverted netmask

Example:

netmask 255.255. 0 . 0 (/16)

| | |
|---|---|
| IP address | 139.131. 63. 53 |
| Logical OR | 0 . 0 .255.255 |
| broadcast is: | 139.131.255.255 |

*Same as host-id of all 1's*

# *IP Broadcast*

G Fairhurst, http://www.erg.abdn.ac.uk

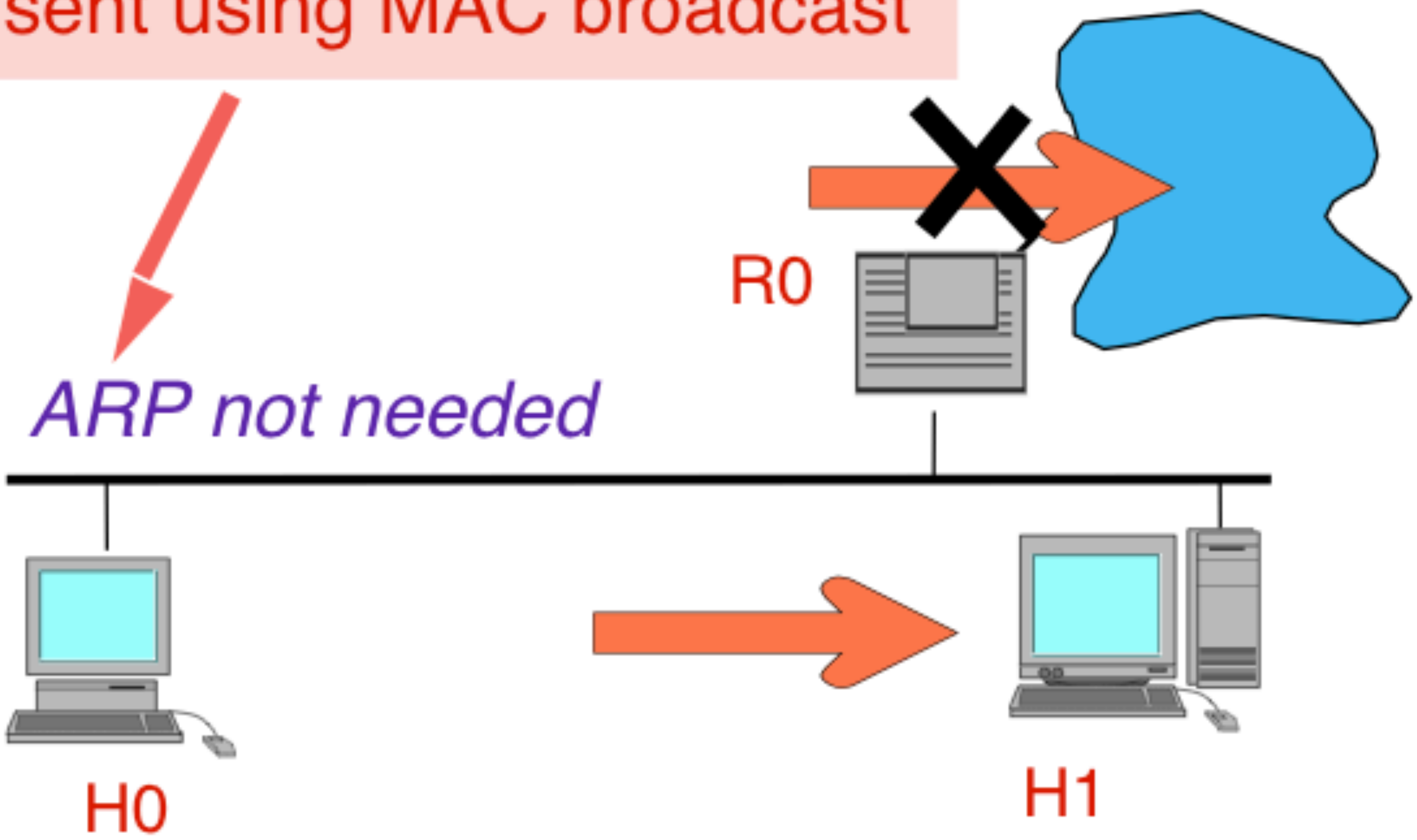
IP broadcast uses network address with a subnet value of all 1's

To all systems in an IP network

Always sent using MAC broadcast

*Routers* ***never*** *forward IP broadcast*

*ARP not needed*



G Fairhurst, http://www.erg.abdn.ac.uk

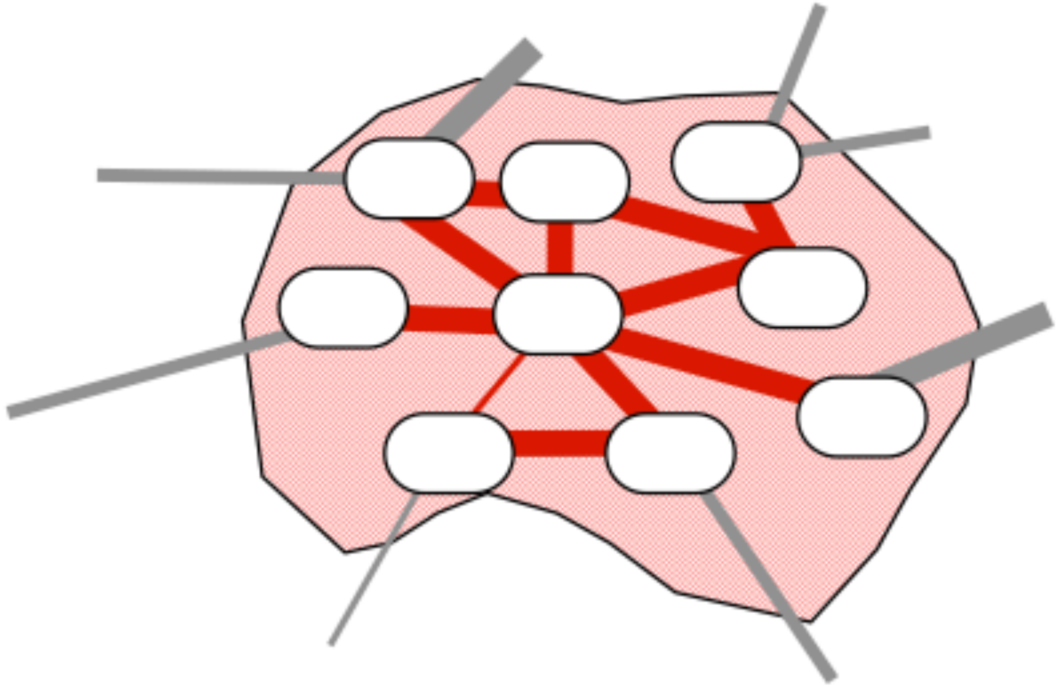
# Routing

Packet Forwarding

Routing & Traceroute

# Routing and Forwarding

G Fairhurst, http://www.erg.abdn.ac.uk



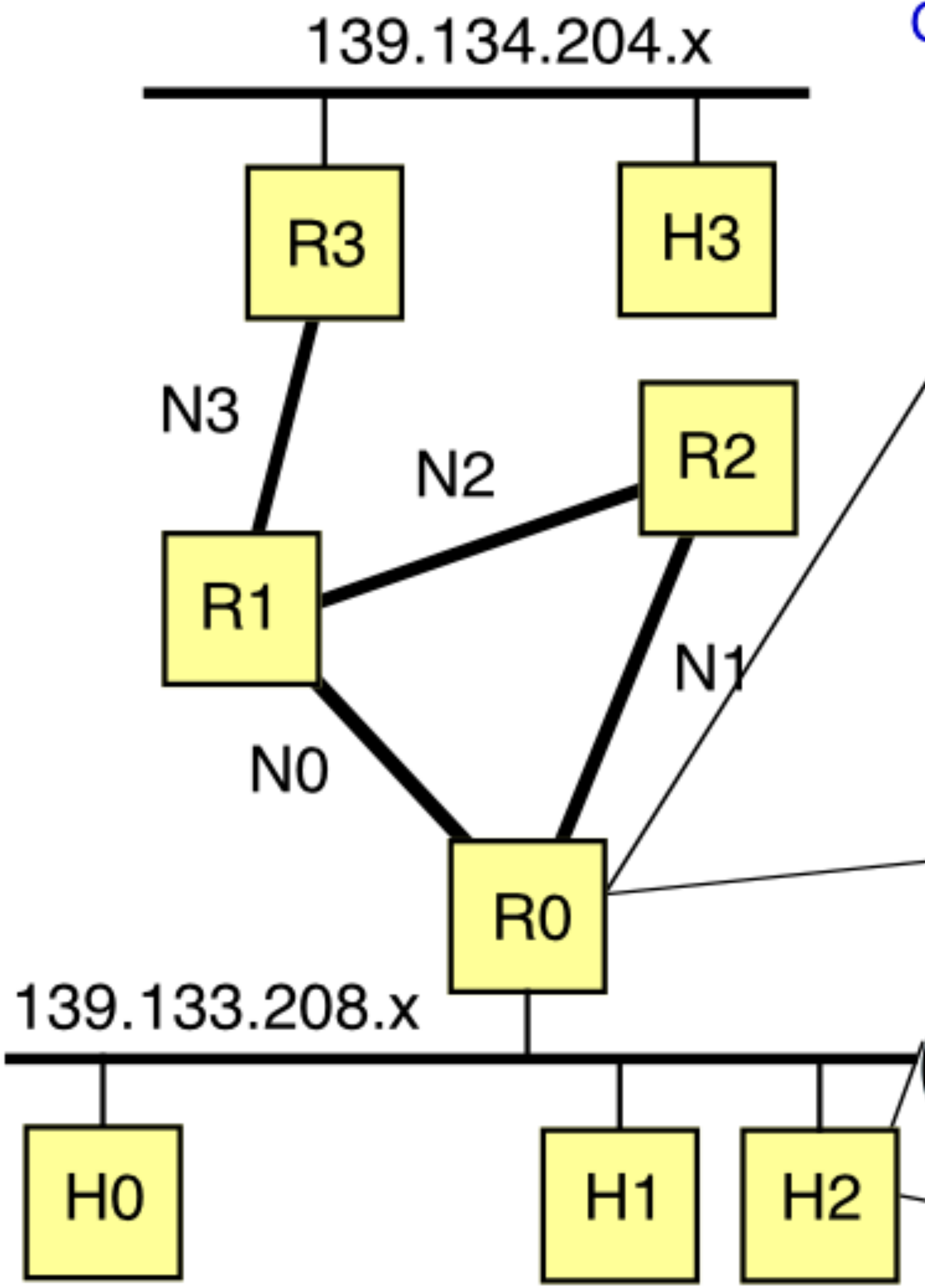
Used between routers in ***within*** a network

Each router informs others about connectivity

Automatically reconfigures around faults

# Routing Table

G Fairhurst, http://www.erg.abdn.ac.uk

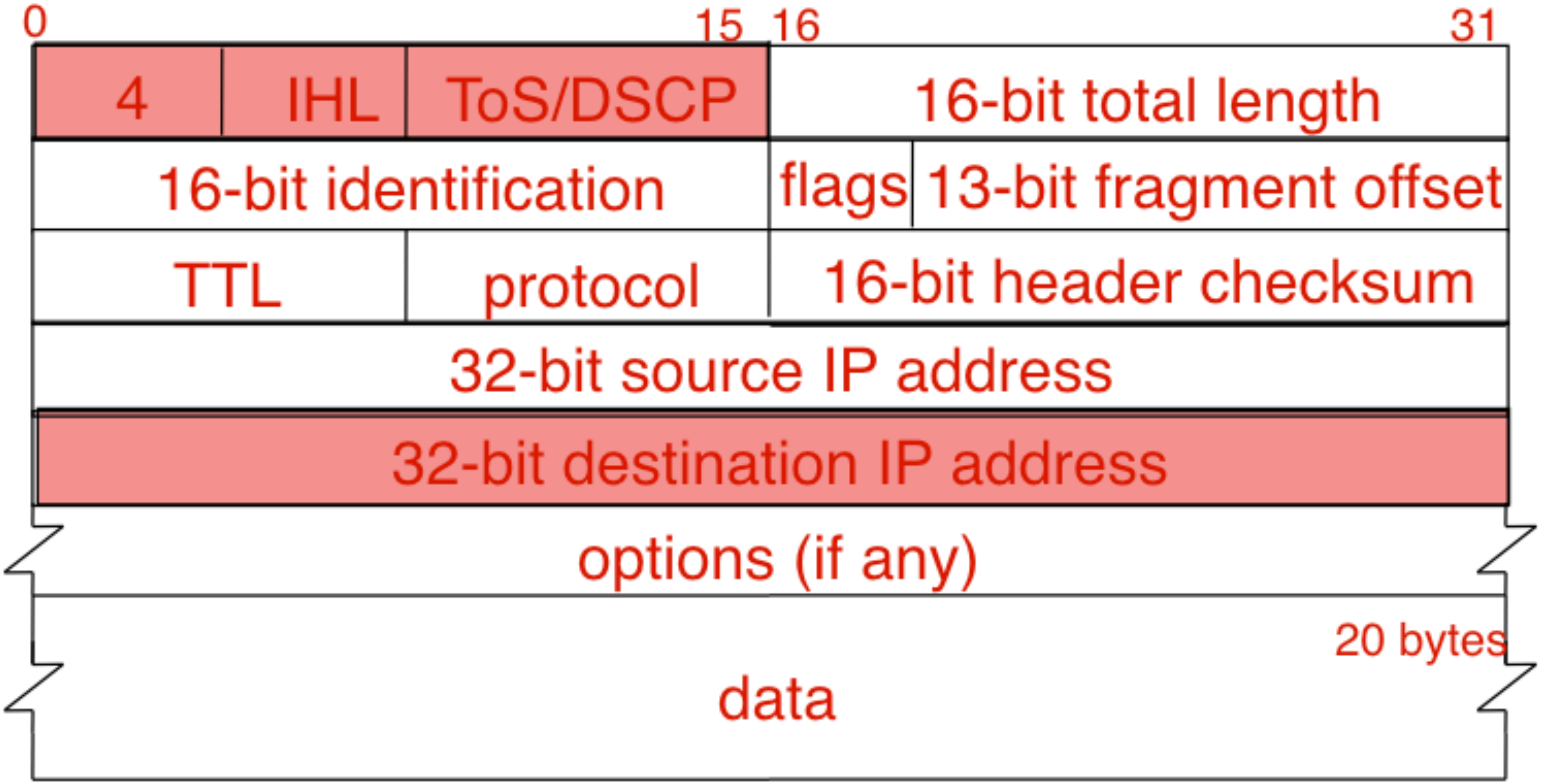



| destination | route |
|---|---|
| 139.133.208.x | local |
| N0 | local |
| N1 | local |
| N2 | R2 |
| N3 | R1 |
| 139.134.x.x | R1 |

| destination | route |
|---|---|
| 139.133.208.x | local |
| default | R0 |

# IP Fields used for Routing

G Fairhurst, http://www.erg.abdn.ac.uk




Router always examine the IP destination address
Routers may also utilise the ToS/DSCP value

RFC 791

# *Routing from a source*

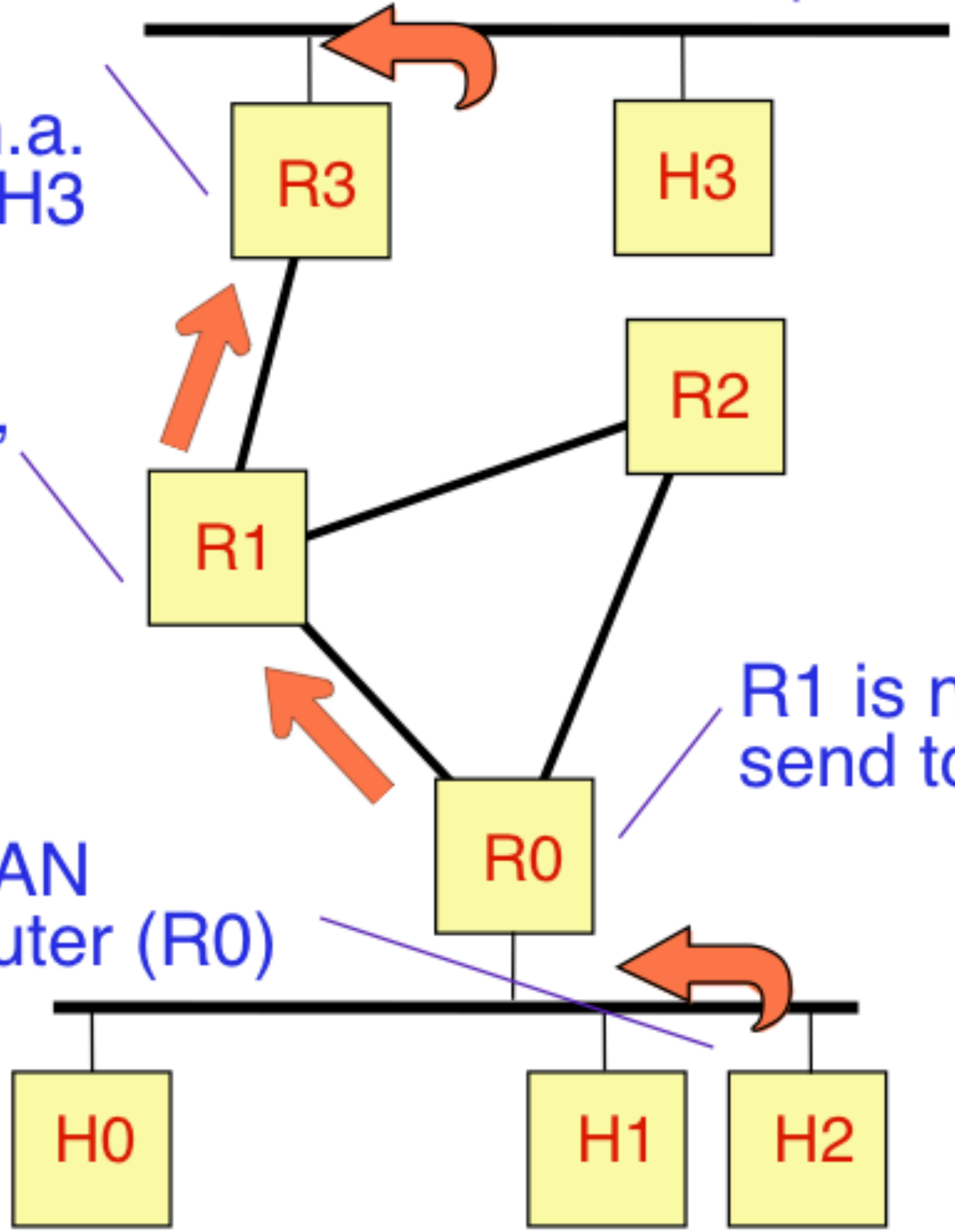
G Fairhurst, http://www.erg.abdn.ac.uk

It's on my LAN
use *arp* to find h.a.
send directly to H3

R3 is nearer H3,
send to R3

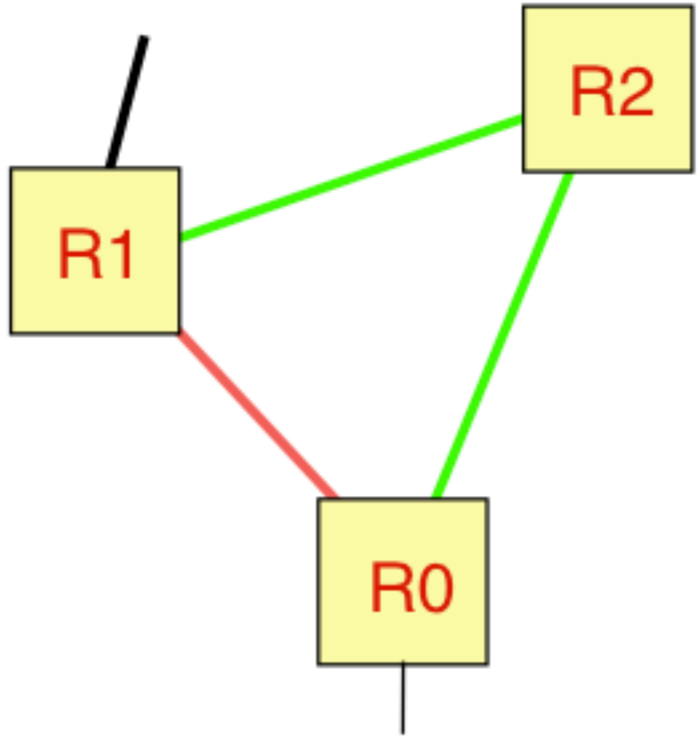
R1 is nearest H3,
send to R1

It's not for my LAN
send to local router (R0)



# Alternative Routes

G Fairhurst, http://www.erg.abdn.ac.uk




**"Best Effort" network service**

Packets are not always delivered

Some may be delivered twice!

Not all packets follow same route... can be reordered

Not all packets take the same time... can be buffered

# Routes from local networks

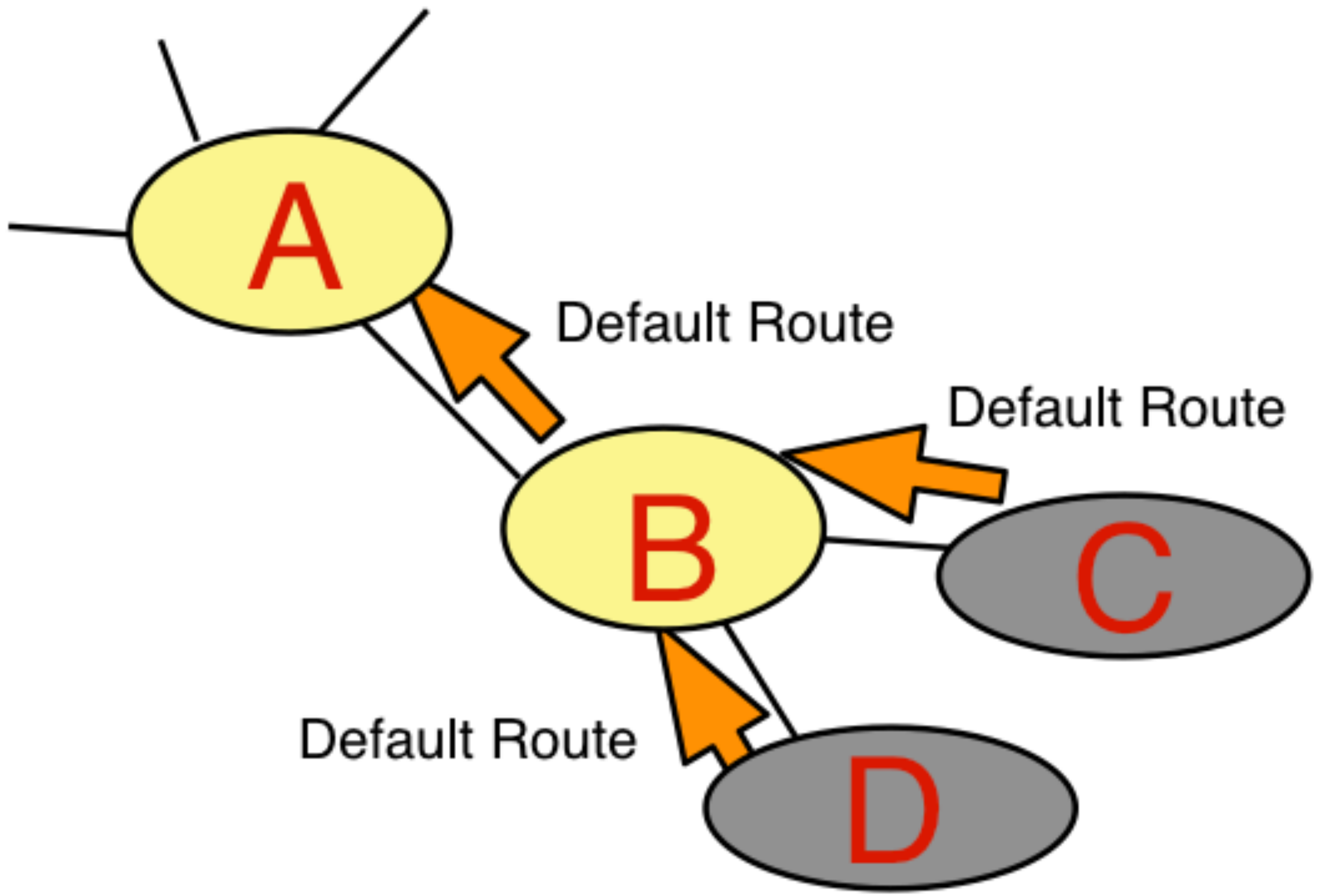
G Fairhurst, http://www.erg.abdn.ac.uk

Router B declares Router B as default route

Router C declares Router B as default route

Router D declares Router B as default route

All traffic to “unknown addresses flows towards “core”



# Routes to Remote Networks

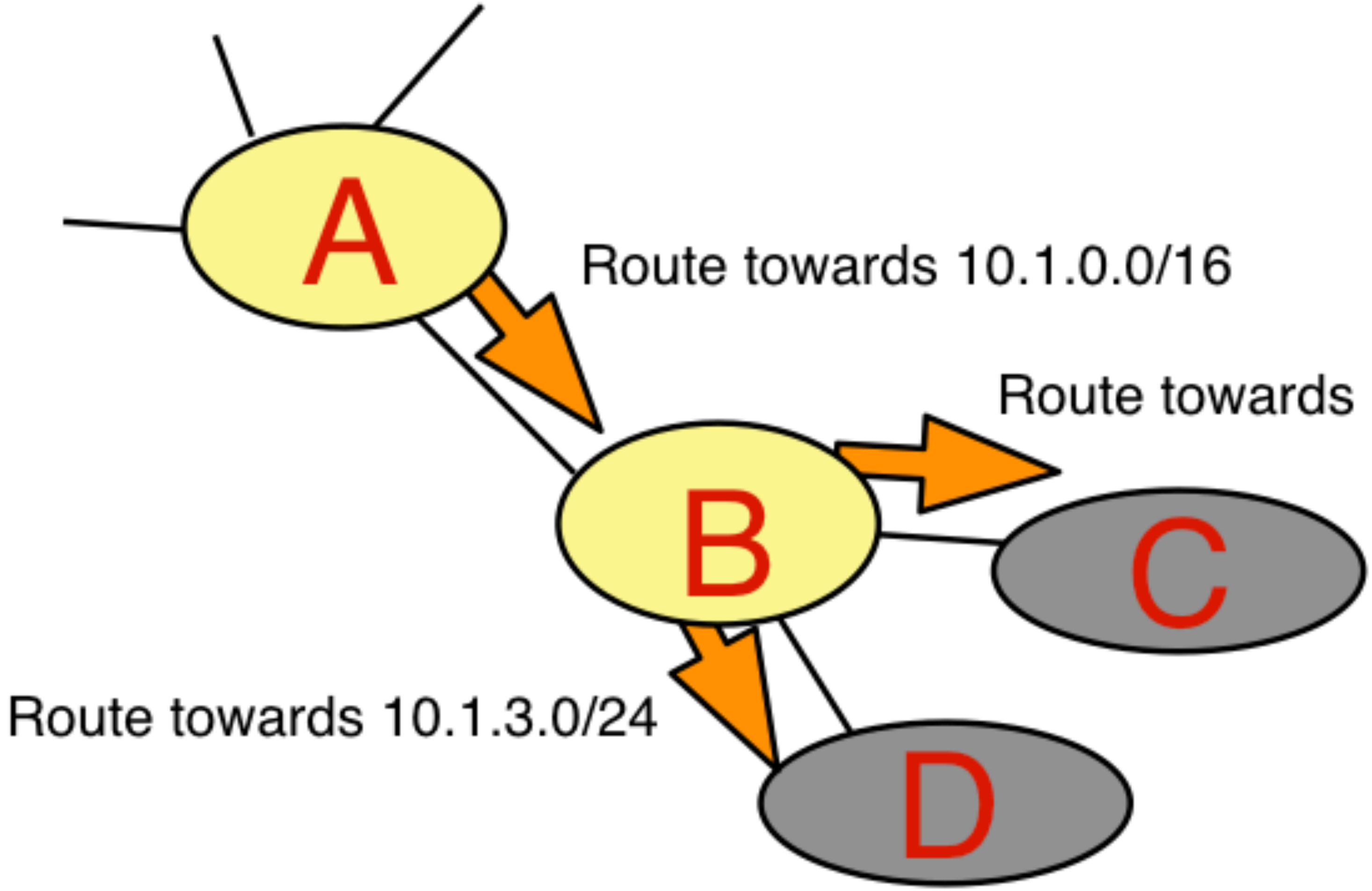
G Fairhurst, http://www.erg.abdn.ac.uk

Router A instructs Router B

Send B all traffic to addresses with network ID 10.1.0.0/16

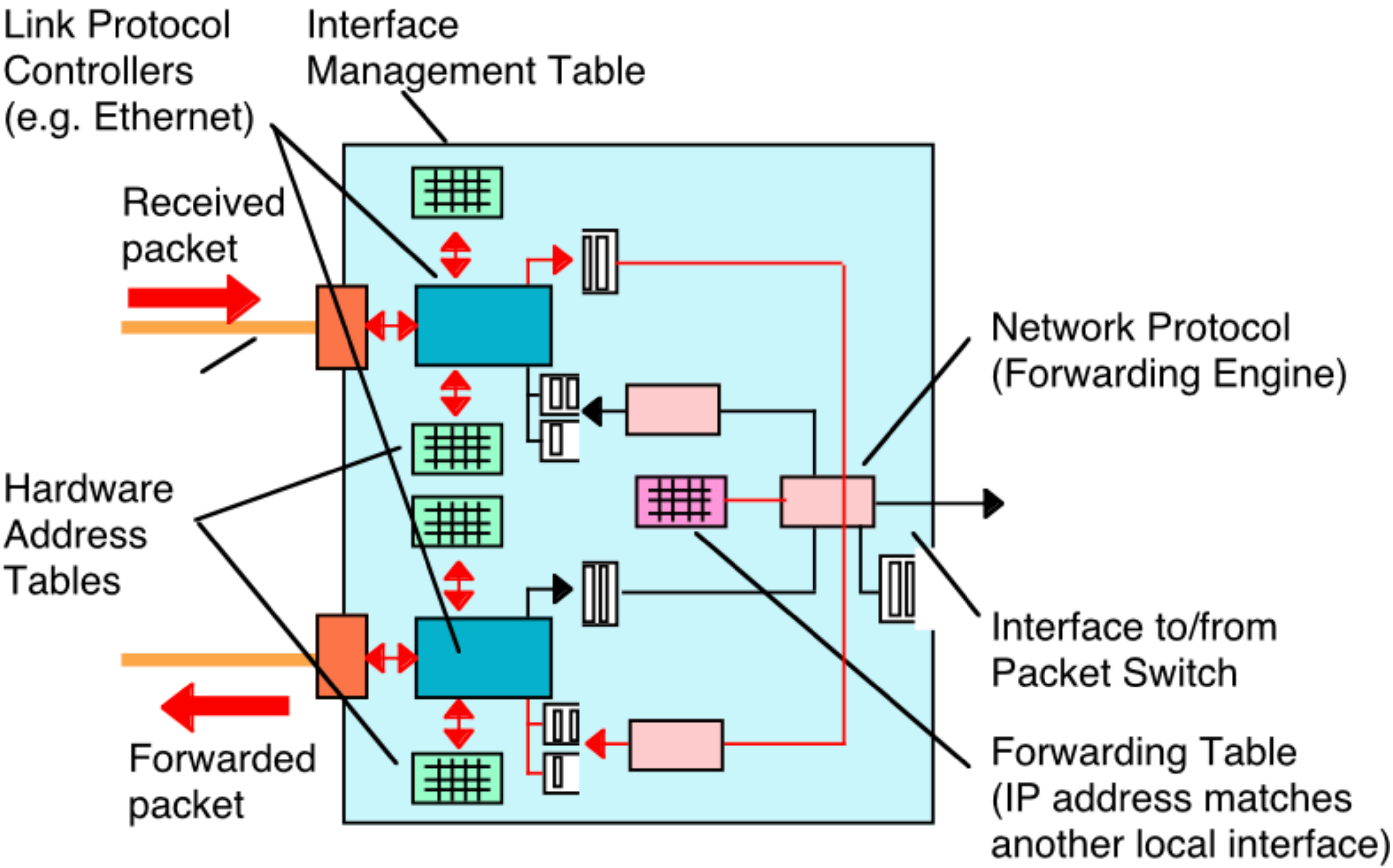
Router B sends traffic with network ID 10.1.1.0/24 to C

Router B sends traffic with network ID 10.1.3.0/24 to D



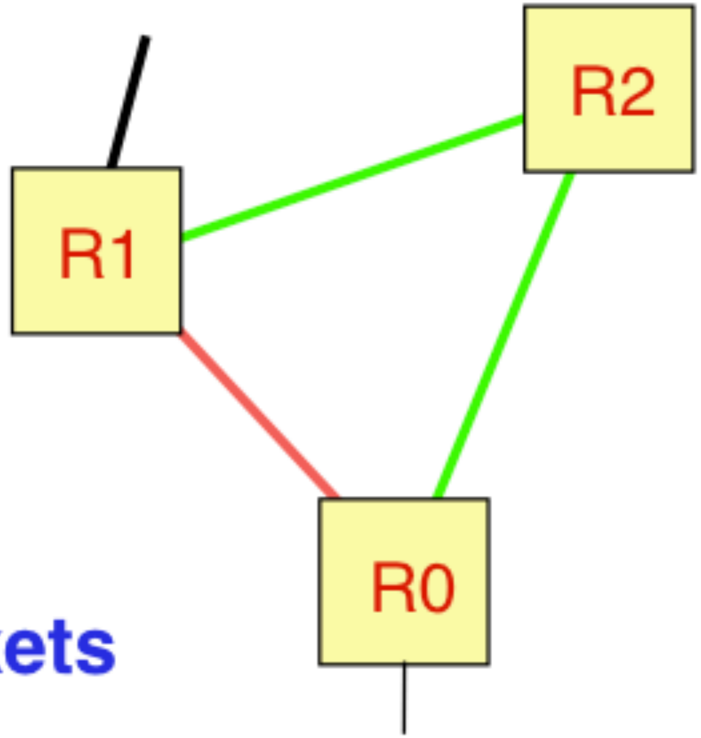
# Router Architecture

G Fairhurst, http://www.erg.abdn.ac.uk



# Routing Protocols

G Fairhurst, http://www.erg.abdn.ac.uk



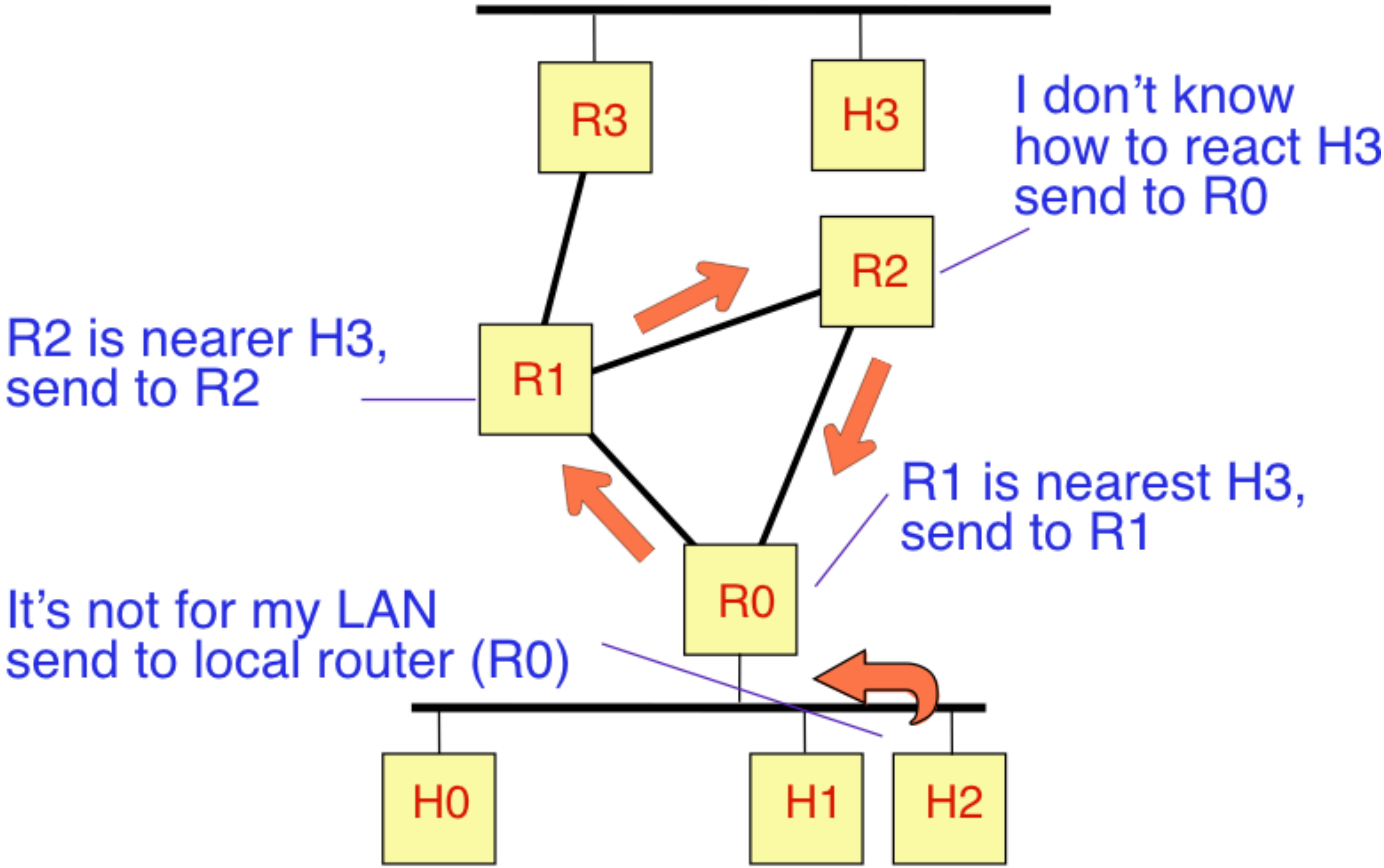

**Routers exchange control packets**

Routers send “control” messages
    If you don’t get them, you know link is “dead”

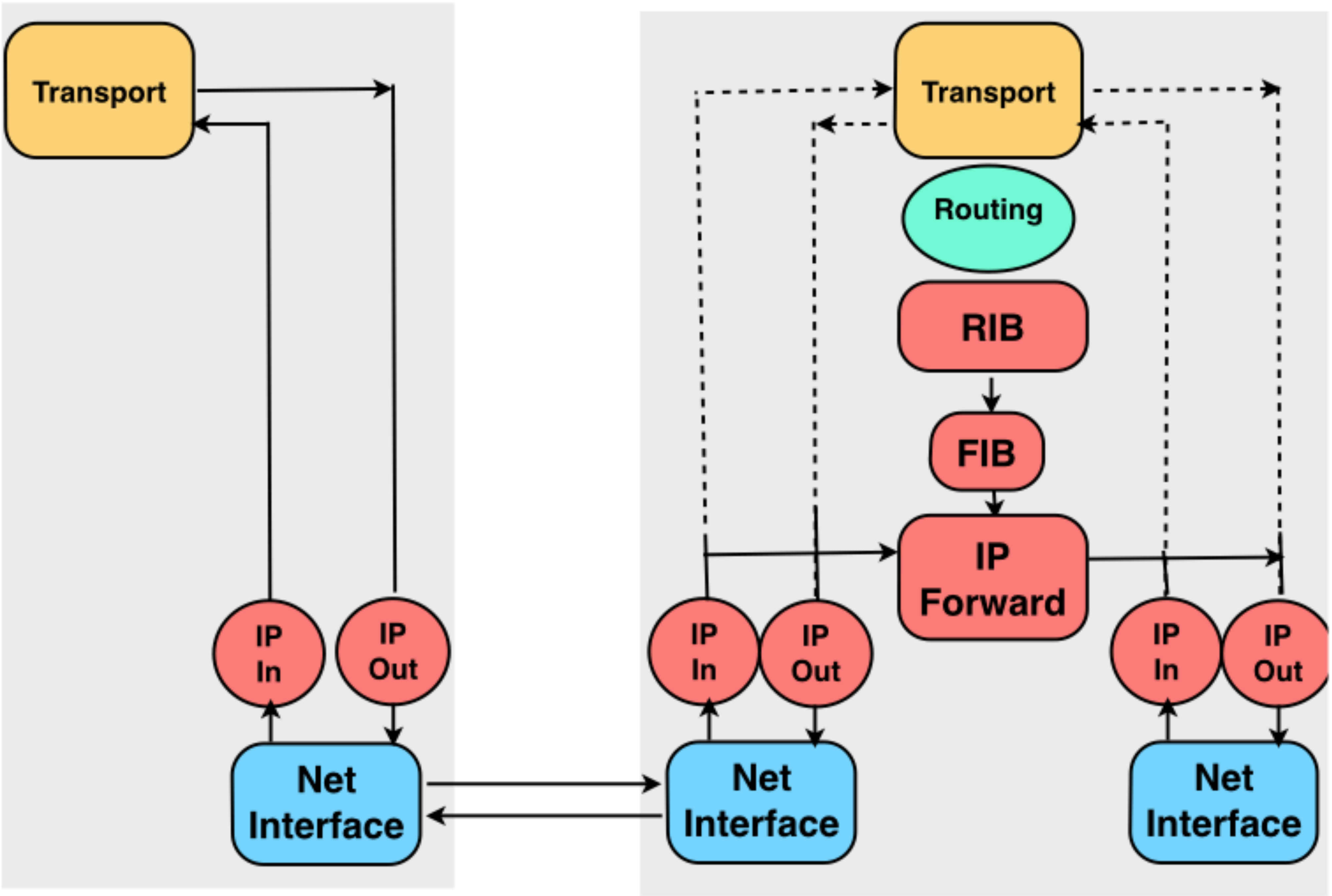
There’s no such thing as a reliable “I am dead indication”
This needs agreement between the two ends

# Routing Errors and Loops

G Fairhurst, http://www.erg.abdn.ac.uk

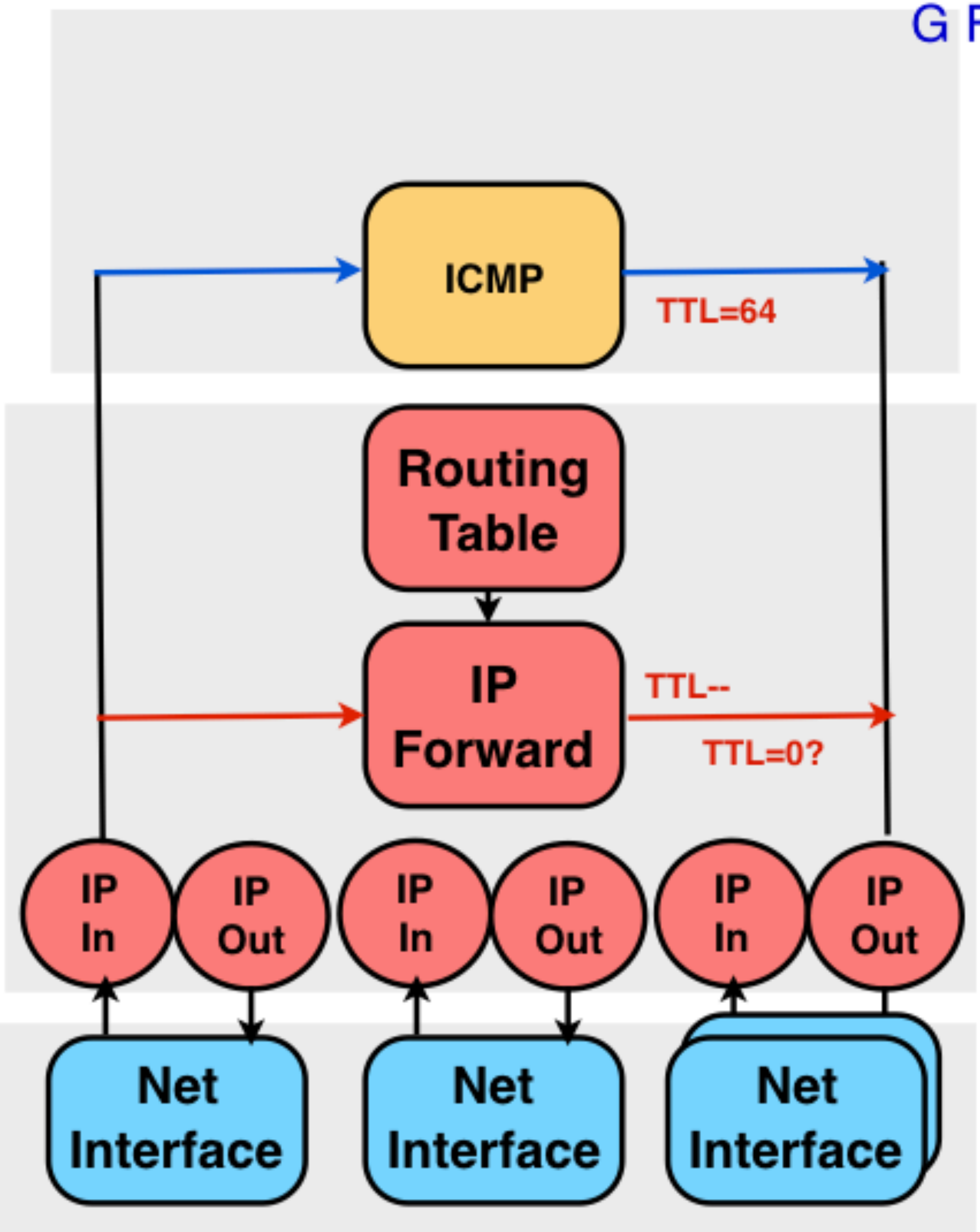


# Routing Process



# Forwarding and TTL

G Fairhurst, http://www.erg.abdn.ac.uk




**Control Plane**

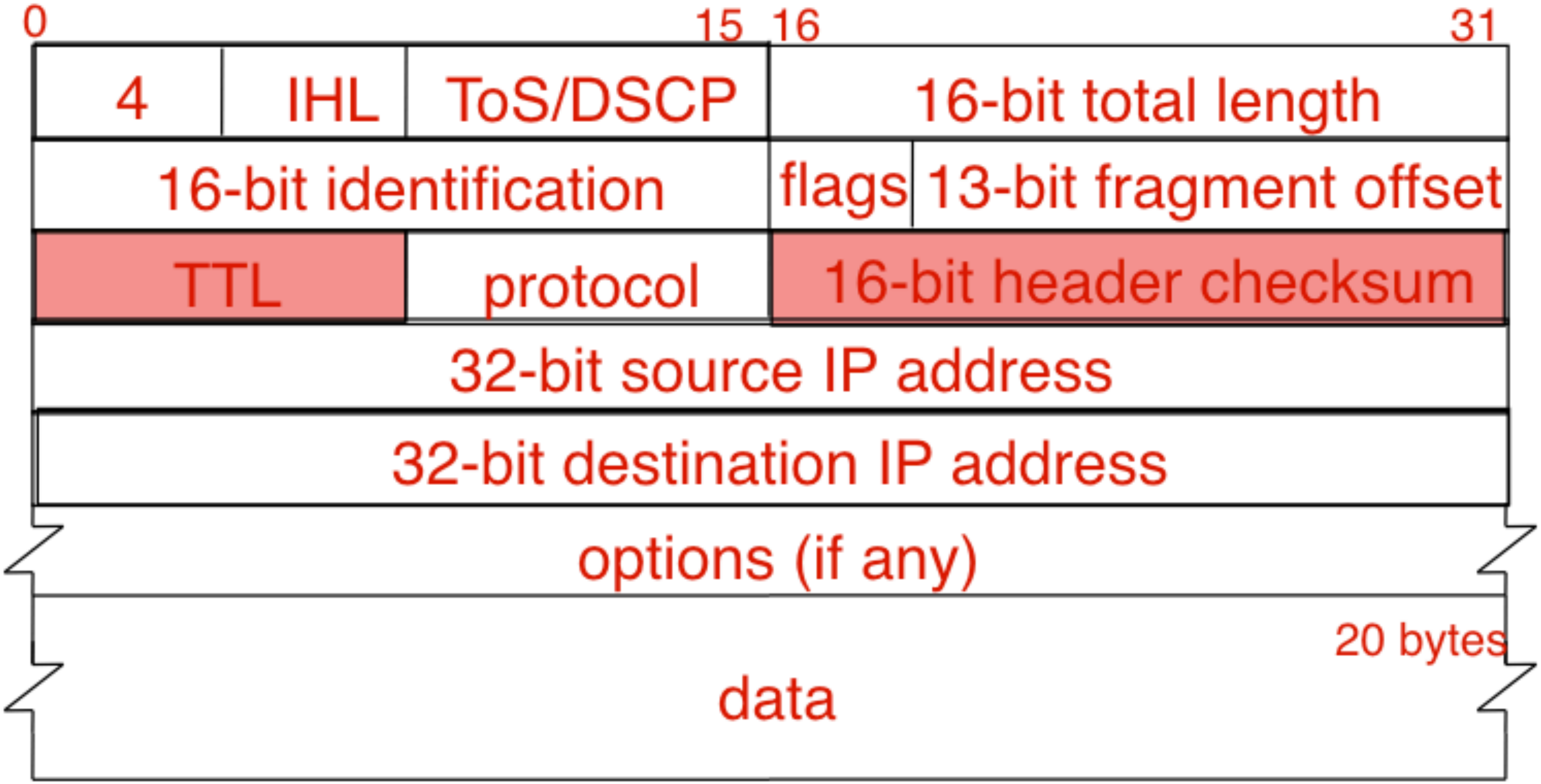
**Data Plane**

(switching/forwarding fabric)

**Interfaces**

# *Fields updated in the IP Header*

G Fairhurst, http://www.erg.abdn.ac.uk




The TTL is decremented and discarded if zero
The IP checksum is recalculated

RFC 791

# Ping v. Traceroute

G Fairhurst, http://www.erg.abdn.ac.uk

**Ping**

Checks connectivity to an endpoint host

- Reports DNS name of the host
- shows packet delivered all the way to the endpoint
- Shows round trip time (or loss) to the endpoint

**Traceroute**

Checks Internet path used to reach an endpoint host

- Reports intermediate routers on the path
- Reports IP address and DNS name of each router
- shows packet delivered up to the reported router
- Shows round trip time (pr loss) to each router on path
-

# Traceroute

G Fairhurst, http://www.erg.abdn.ac.uk

Route to zeno.ksc.nasa.gov (128.159.1.155),
30 hops max, 40 byte packets

```
 1  milliways-erg (10.0.0.64)                    2 ms    1 ms    1 ms
 2  139.133.210.1 (139.133.210.1)                7 ms    7 ms    7 ms
 3  abdn-gw.abdn.ac.uk (139.133.7.6)             3 ms    3 ms    3 ms
 4  smds-gw.ulcc.ja.net (193.63.203.33)         17 ms   16 ms   15 ms
 5  nsn-gw.ulcc.ac.uk (128.86.1.3)              16 ms   17 ms   16 ms
 6  128.161.165.1 (128.161.165.1)               96 ms  123 ms  147 ms
 7  GSFC6.NSN.NASA.GOV (192.100.13.6)           98 ms  115 ms  146 ms
 8  128.161.44.4 (128.161.44.4)                178 ms  189 ms  154 ms
 9  MSFC1.NSN.NASA.GOV (192.100.14.1)          170 ms  178 ms  175 ms
10  KSC.NSN.NASA.GOV (128.161.30.27)           192 ms  316 ms  168 ms
11  192.150.33.1 (192.150.33.1)                192 ms  205 ms  213 ms
12  128.159.215.239 (128.159.215.239)          172 ms  172 ms  193 ms
13  163.205.253.254 (163.205.253.254)          330 ms  222 ms  269 ms
14  zeno.ksc.nasa.gov (128.159.1.155)          220 ms  377 ms  177 ms
```

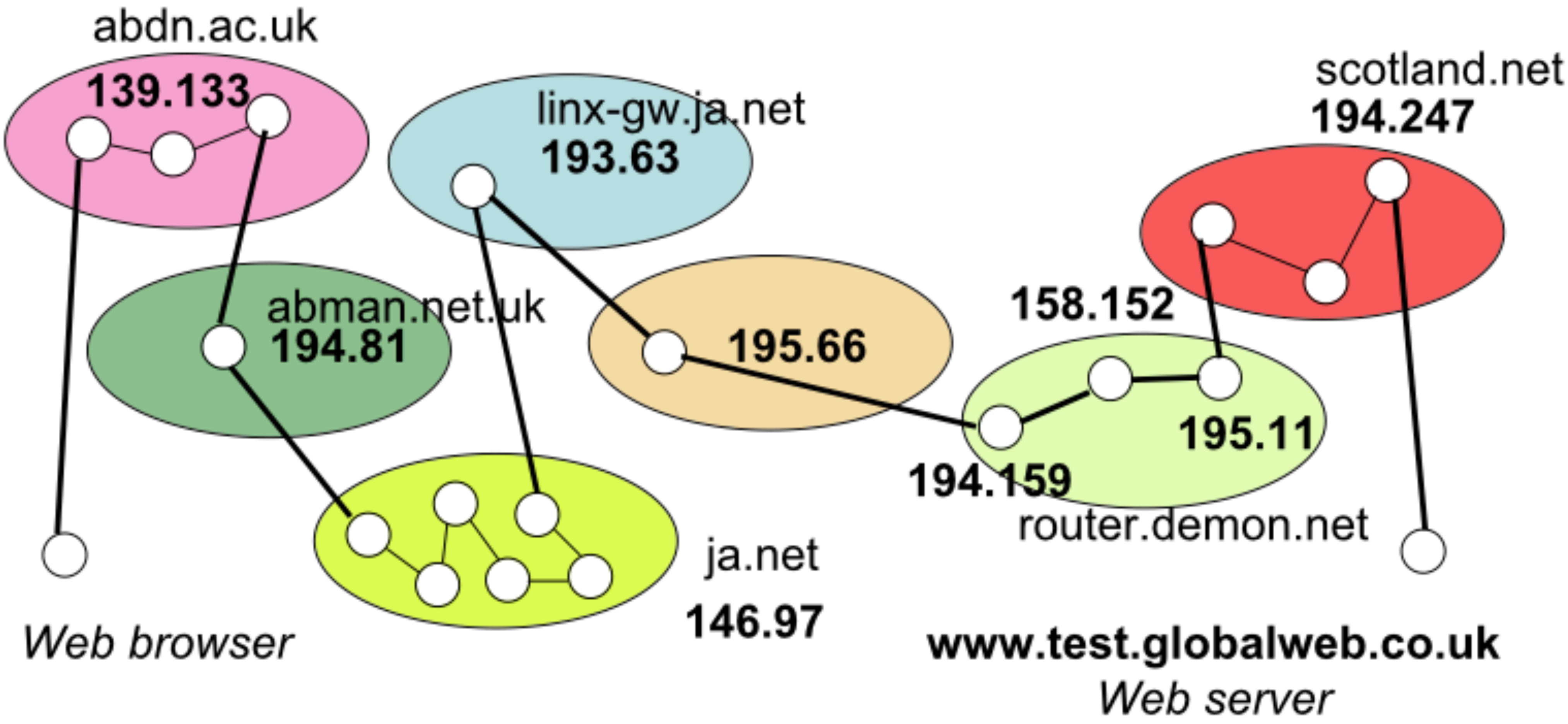
# Traceroute to globalweb.ac.uk

G Fairhurst, http://www.erg.abdn.ac.uk

```
 1  milliways (139.133.204.64)  2.831 ms  2.077 ms  2.167 ms
 2  gw34.abdn.ac.uk (139.133.34.1)  4.828 ms  4.955 ms  4.865 ms
 3  gwkccs.abdn.ac.uk (139.133.7.4)  16.989 ms  15.510 ms  5.331 ms
 4  aclarke-gw.abman.net.uk (194.81.60.94)  7.769 ms  5.545 ms  5.734 ms
 5  146.97.250.17 (146.97.250.17)  9.785 ms  12.061 ms  9.347 ms
 6  146.97.37.29 (146.97.37.29)  13.904 ms  16.689 ms  11.144 ms
 7  pos9-0.edin-scr.ja.net (146.97.35.61)  11.492 ms  16.527 ms  21.450 ms
 8  pos0-0.leed-scr.ja.net (146.97.33.26)  18.450 ms  27.231 ms  19.766 ms
 9  pos2-0.lond-scr.ja.net (146.97.33.30)  32.023 ms  35.862 ms  28.696 ms
10  146.97.35.6 (146.97.35.6)  26.864 ms  25.046 ms  24.458 ms
11  linx-gw.ja.net (193.63.94.249)  23.115 ms  32.644 ms  21.848 ms
12  linx-2.router.demon.net (195.66.224.13)  26.371 ms  26.082 ms  22.430 ms
13  tele-backbone-1-ge020.router.demon.net (194.159.252.54)
14  anchor-core-2-fxp1.router.demon.net (158.152.0.178)
15  demon-gw-2.sol.co.uk (195.11.50.130)  37.791 ms  33.314 ms  38.483 ms
16  atm1-0-0-1.core2.scotland.net (194.247.77.34)  50.325 ms  56.771 ms
17  fe12-0-0.core1.scotland.net (194.247.67.41)  44.368 ms  46.100 ms
18  ABZ-Sci-Park.LL.scotland.net (194.247.71.109)  50.041 ms  51.625 ms
```

# Route to globalweb.ac.uk

G Fairhurst, http://www.erg.abdn.ac.uk




**18 hops in total**
**over 9 domains (7 intermediate)**

# Traceroute might not report a router

G Fairhurst, http://www.erg.abdn.ac.uk

**Some reasons why Traceroute might not report a router:**

- Routers - might not send ICMP, due to rate limits
- There might be packet loss - IP is best effort
- Routers on path - might filter ICMP to hide network topology
  Routers might not have a return path route to the source

*Traceroute prints a star (*) when no ICMP message is received.*

# Required Information

G Fairhurst, http://www.erg.abdn.ac.uk

Own MAC hardware address (from NIC PROM)
- Used in MAC source address

Own IP address (given by network administrator)
- Used as IP source address

Own IP subnet mask (given by network administrator)
- Indicates remote destination addresses
- Indicates IP broadcast address (to all local systems)

IP default router (given by network administrator)
- IP address of router to send to for remote addresses
- (indicates MAC address for remote IP networks)

IP name server server (given by network administrator)
- IP address of a server to resolve names <-> address

# Transport,Middleware & Applications

- UDP (Header, Services)
- Demultiplexing (Protocol, Port)
- Checksum
- TCP (Header, Services)
- TCP Connections
- Applications