University of St Andrews

CS2003: Internet and the Web
IPv4 Addressing and Address Management

IPv4 addresses – classless

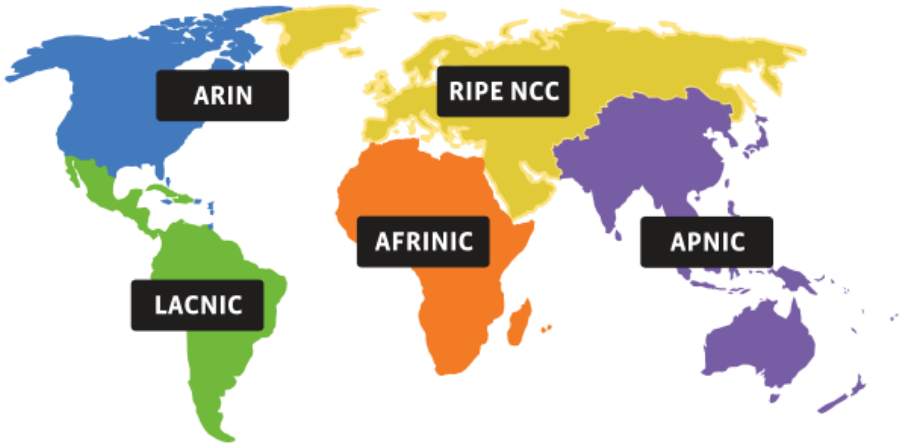
- Special addresses:
  - (also many for IPv6, not shown here)

Address example	Src	Dst	Description
255.255.255.255	No	Yes	Limited broadcast. If the local network supports broadcast then broadcast the datagram. This is never passed on by routers.
138.251.255.255 210.50.160.255	No	Yes	Net directed broadcast. If the target network supports broadcast then broadcast the datagram on it.
127.x.x.x	Yes	Yes	Loopback. Send to yourself.
0.0.42.6	Yes	No	Used by a host which does not know its network prefix.
0.0.0.0	Yes	No	Used by a host that does not know its IP address.

- Classless addressing:
  - remove use of Class A, B and C
  - **address mask** → network prefix
  - IPv4 address plus mask: e.g. 138.251.195.61/24
- More flexibility in address allocation.
- Routing information aggregation:
  - (sub-netting & super-netting)
  - **CIDR: Classless InterDomain**

Global address management (1)

- Regional Internet Registry (RIR)
- National Internet Registry (NIR)
- Local Internet Registry (LIR)
- **Delegated allocation of address prefixes.**



https://www.iana.org/numbers

REGISTRY	AREA COVERED
AFRINIC	Africa Region
APNIC	Asia/Pacific Region
ARIN	Canada, USA, and some Caribbean Islands
LACNIC	Latin America and some Caribbean Islands
RIPE NCC	Europe, the Middle East, and Central Asia

Global address management (2)

University of St Andrews

- Hierarchy of address allocation:
  - Global registries delegate prefixes to regional registries.
  - Regional registries delegate prefixes to local registries.
  - Local registries delegate prefixes to ISPs.
- ISPs allocate addresses to users:
  - Typically a single address, with NAT for IPv4.
  - (IPv6 can be single address with NAT, but can also be a prefix, e.g. /48, /56, /64)

Private IPv4 address space

- Private IPv4 networks:
  - We have run out of IPv4 addresses.
  - Internet access via Network Address Translation (NAT).
- Using IPv4, so need some IPv4 addresses!
  - Could possibly use any address.
  - Hosts **must not** be directly connected to the Internet.
- IPv4 prefixes for private network use (RFC1918):
  - 10/8 prefix.
  - 172.16/12 prefix.
  - 192.168/16 prefix.
  - Not forwarded by routers.
  - NAT maps these to one (or more) public IPv4 addresses.

Local address management (1)

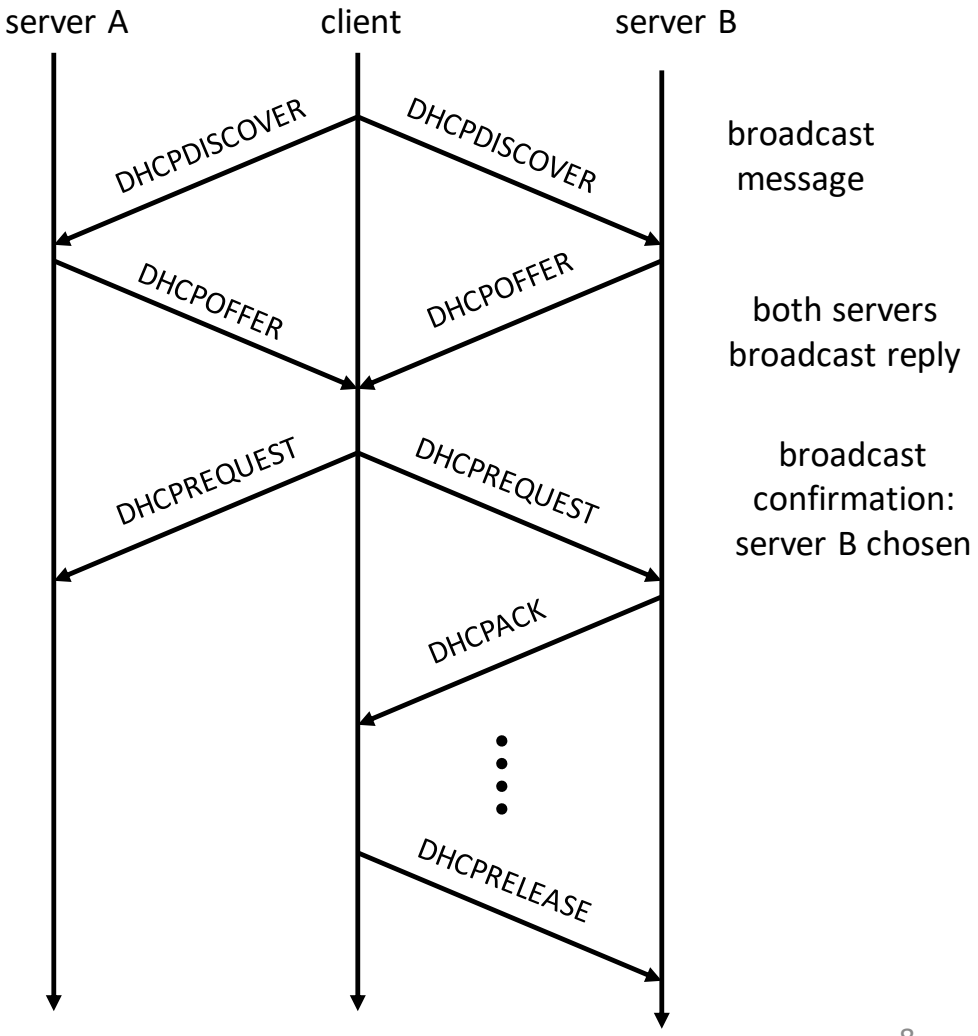
- Allocation of a prefix or address range (or a single address!) from your upstream provider.
- Manual configuration per host:
  - not practical / scalable / maintainable.
- PPP can assign addresses:
  - "dial-up" / but also used on broadband services.
- **Dynamic Host Configuration Protocol (DHCP):**
  - allows automated address assignment.
  - also: default router, DNS resolver(s), mail server …
  - (IPv6 – DHCPv6 – but IPv6 does not "need" it like IPv4)

Local address management (2)

- Client systems – for use by human users:
  - Usually assigned addressed dynamically (e.g. via DHCP).
  - Clients often have “private” addresses (used with NAT).
  - Clients do not normally provide services:
    - Do not usually have to wait for incoming connections.
    - Do not need DNS entries.
- Servers – provide services accessible by clients:
  - Often have globally unique addresses.
  - Can be semi-permanently assigned, but can also use DHCP.
  - Clients lookup server address, so DNS entry required.

DHCP

- Allow dynamic configuration:
  - automatically assigned address "leasing" from a **pool** of addresses
  - global or private address pool
  - can configure network parameters: IP address, subnet mask, default gateway, local DNS resolver, …
- Uses LAN broadcast
- Requires server(s):
  - central store of configuration information, central administration
- Useful for:
  - "nomadic" devices
  - large numbers of hosts (can use static/manual address assignment)



University of St Andrews

Basic routing and forwarding at a edge / site network

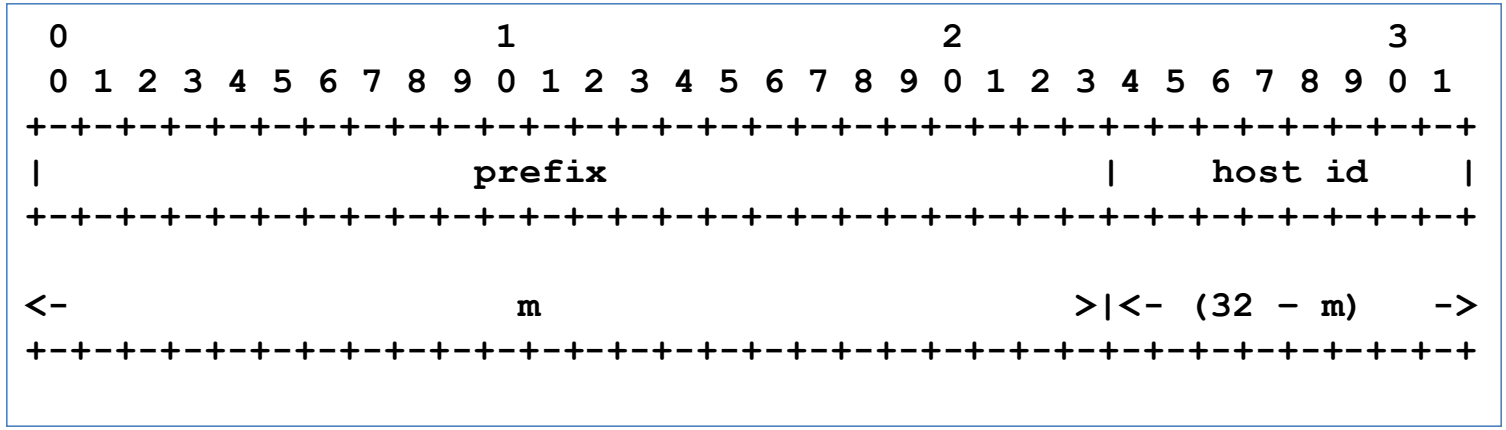
IP address allocation (1)

- IP prefix “assigned” to a network by administrator, e.g.:
  - 138.251.195.0/24
  - 192.168.1.0/24
- Prefix of **m** bits, for IPv4, so number of hosts:
  - **Number of hosts: $2^{(32-m)} - 2$**, e.g. 121.16.20.0/24:
  - “zero” host address is “the network”: 121.16.20.0
  - “all 1s” host address is IPv4 broadcast (remember previous discussion): 121.16.20.255
- **Prefix size constrains number of global addresses.**

IP address allocation (2)

University of St Andrews

IPv4 prefix 121.16.20.1/24



binary

```
255: 1111 1111
121: 0111 1001
 16: 0001 0000
 20: 0001 0100
  1: 0000 0001
```

masks

```
/16: 255.255.0.0
/20: 255.255.240.0
/24: 255.255.255.0
```

m = 24 **number of hosts** $= 2^{(32-m)}-2$

$= 2^{(32-24)}-2$

$= 253$

address

121.16.20.2 **0111 1001 0001 0000 0001 0100** 0000 0010

121.16.20.4 **0111 1001 0001 0000 0001 0100** 0000 0100

121.16.21.6 **0111 1001 0111 1001 0001 0101** 0000 0101 **wrong prefix**

IPv4 routing and forwarding (1)

University of St Andrews

- **Forwarding:**
  - Receiving a packet, looking up destination address in a table, and sending the packet in the direction indicated by the table
- **Routing:**
  - The process by which forwarding tables are built
  - Discover paths through network, gather routing information, using a routing protocol

IPv4 routing and forwarding (2)

- IP routers perform routing and forwarding
  - Discover paths, assign values, *and* transmit packets using discovered routing information
- IP hosts perform forwarding
  - Discover *locally* available routes
  - Make *local* forwarding decisions

IPv4 routing and forwarding (3)

- Routing protocol finds paths to **destinations**:
  - A **destination** is another network (with a prefix).
  - Discovery messages, discover hops along path.
  - Router typical has **multiple interfaces**.
  - A **destination** is reachable on a particular **interface**.
- Router constructs a **forwarding table**:
  - For a given destination, based on the routing protocol information, select which is the best interface on which to transmit the packet.
  - Similar process for forwarding on a host.

IPv4 routing and forwarding (4)

- Choose route for **forwarding** using **longest prefix match**:
  - apply network mask of routing entry to destination address in packet (address AND mask)
  - the longest mask that produces a match between the local forwarding entry and destination address is used
- If no prefix matches, use **default route**:
  - IPv4 default route entry for any destination is 0.0.0.0.

Forwarding table on Linux

University of St Andrews

```
File  Edit  View  Search  Terminal  Help
tristan@pc3-005-l:~ $ ip link show
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN mode DEFAULT group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
2: enp3s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP mode DEFAULT group default qlen 1000
    link/ether 40:16:7e:a8:fc:74 brd ff:ff:ff:ff:ff:ff
3: virbr0: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc noqueue state DOWN mode DEFAULT group default qlen 1000
    link/ether 52:54:00:0b:d8:88 brd ff:ff:ff:ff:ff:ff
4: virbr0-nic: <BROADCAST,MULTICAST> mtu 1500 qdisc fq_codel master virbr0 state DOWN mode DEFAULT group default qlen 1000
    link/ether 52:54:00:0b:d8:88 brd ff:ff:ff:ff:ff:ff
tristan@pc3-005-l:~ $ ip address show dev enp3s0
2: enp3s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
    link/ether 40:16:7e:a8:fc:74 brd ff:ff:ff:ff:ff:ff
    inet 138.251.29.167/23 brd 138.251.29.255 scope global dynamic noprefixroute enp3s0
       valid_lft 40996sec preferred_lft 40996sec
    inet6 fe80::4216:7eff:fea8:fc74/64 scope link noprefixroute
       valid_lft forever preferred_lft forever
tristan@pc3-005-l:~ $
tristan@pc3-005-l:~ $ ip route show
default via 138.251.29.254 dev enp3s0 proto dhcp metric 100
138.251.28.0/23 dev enp3s0 proto kernel scope link src 138.251.29.167 metric 100
192.168.122.0/24 dev virbr0 proto kernel scope link src 192.168.122.1 linkdown
tristan@pc3-005-l:~ $
tristan@pc3-005-l:~ $ netstat -4rn
Kernel IP routing table
Destination     Gateway         Genmask         Flags   MSS Window  irtt Iface
0.0.0.0         138.251.29.254  0.0.0.0         UG        0 0          0 enp3s0
138.251.28.0    0.0.0.0         255.255.254.0   U         0 0          0 enp3s0
192.168.122.0   0.0.0.0         255.255.255.0   U         0 0          0 virbr0
tristan@pc3-005-l:~ $
```

IPv4 forwarding example

	destination	mask	next hop	metric	interface
A	121.16.0.0	16	121.16.16.14	1	eth0
B	121.16.20.0	24	121.16.20.1	2	eth1
C	121.16.16.0	20	121.16.16.14	3	eth2
D	0.0.0.0	0	121.16.16.14	1	eth0

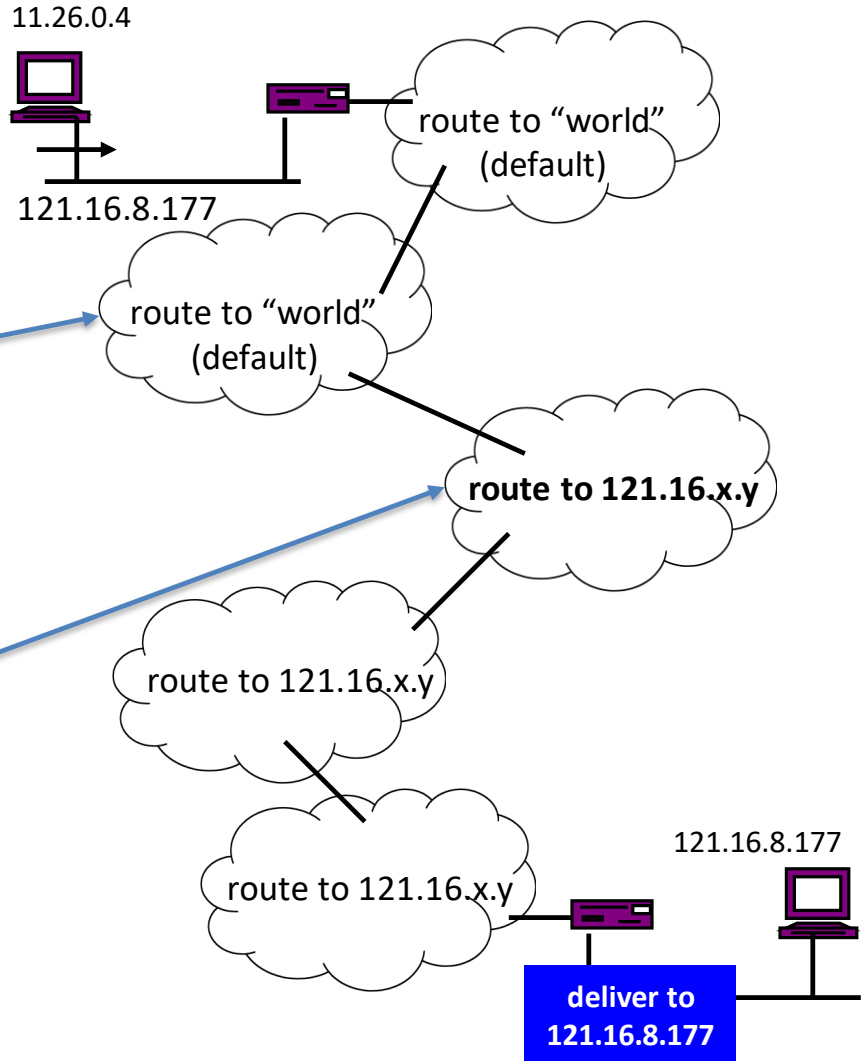
dst addr	
121.16.20.2	→ **B**
121.16.23.4	→ **C**
121.16.6.8	→ **A**
64.28.67.150	→ **D**

masks	entry
/16:	255.255.0.0
/20:	255.255.240.0
/24:	255.255.255.0

binary	
255:	1111 1111
240:	1111 1111
23:	0001 0111
20:	0001 0100
16:	0001 0000
8:	0000 1000
4:	0000 0100
2:	0000 0010
1:	0000 0001

Internet routing: addresses and hierarchy

- **Network prefix**:
  - far fewer networks than hosts
  - reduces routing information
- **Non-backbone networks:**
  - **default routes**
  - any network that is not directly connected
  - "the rest of the world"
- **Backbone networks**:
  - **no default routes (default free zone – DFZ)**
  - **large** routing tables



IPv4 sub-netting (1)

- Separate LANs:
  - need router to interconnect
  - need separate IP networks
- In IPv4 use part of hostID as **sub-network ID**:
  - e.g. /8: ~16.8M hosts
  - e.g. /16: ~65K hosts
  - e.g. /24: 253 hosts
- "Extend" network part of IP address as required:
  - need to specify **network mask**
- (Sub-netting works differently for IPv6 addresses.)

Example: 2 networks, 100 hosts each

192.192.192/24: split into two

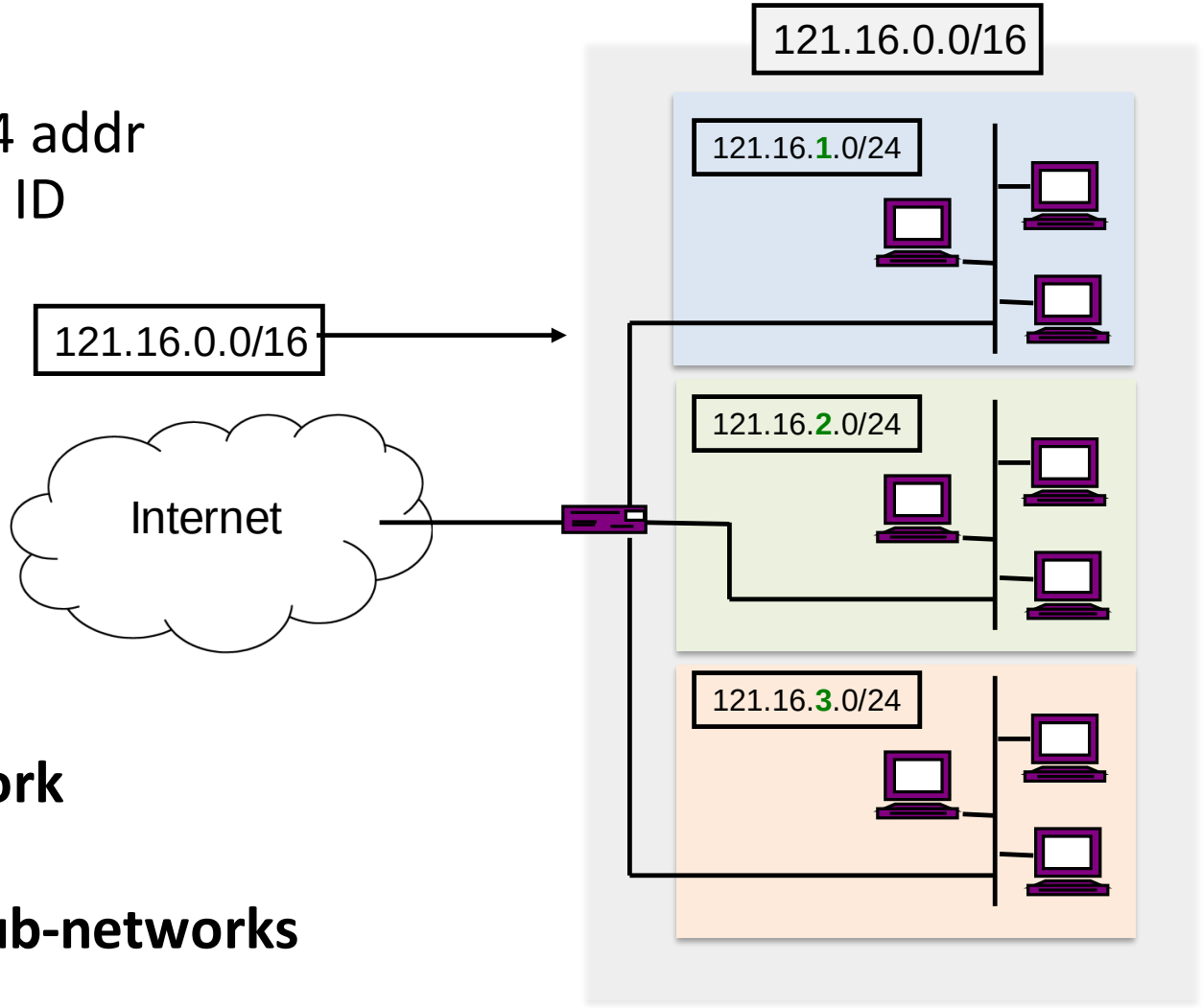
192.192.192/25
2 sub-nets, 126 hosts per sub-net

192.192.192.0/25: networkID
192.192.192.1 – 192.192.192.126: hosts
192.192.192.127: broadcast

192.192.192.128/25: networkID
192.192.192.129 – 192.192.192.254: hosts
192.192.192.255: broadcast

IPv4 sub-netting (2)

Use 3rd byte of /24 addr address as subnet ID



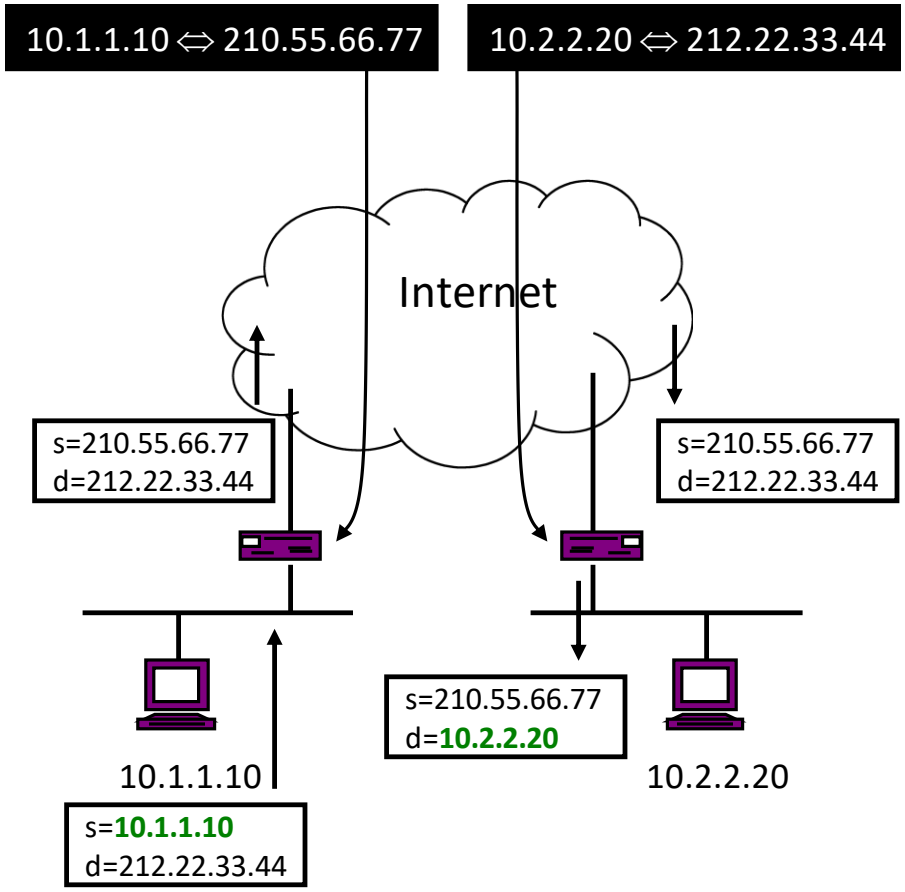
- External view:
  one network
- Internal view:
  many **IP sub-networks**

Private IP address space

- Run out of IPv4 address space!
- Private IP networks:
  - a private IP-based network.
- Use IP, so need IP addresses:
  - can use any address, in theory as long as …
  - … hosts **not** directly connected to the Internet
- Instead, use special network prefixes:
  - IPv4 RFC1918 (BCP)
  - 10/8
  - 172.16/12
  - 192.168/16
  - **packets with private source addresses are not forwarded to public network by default**

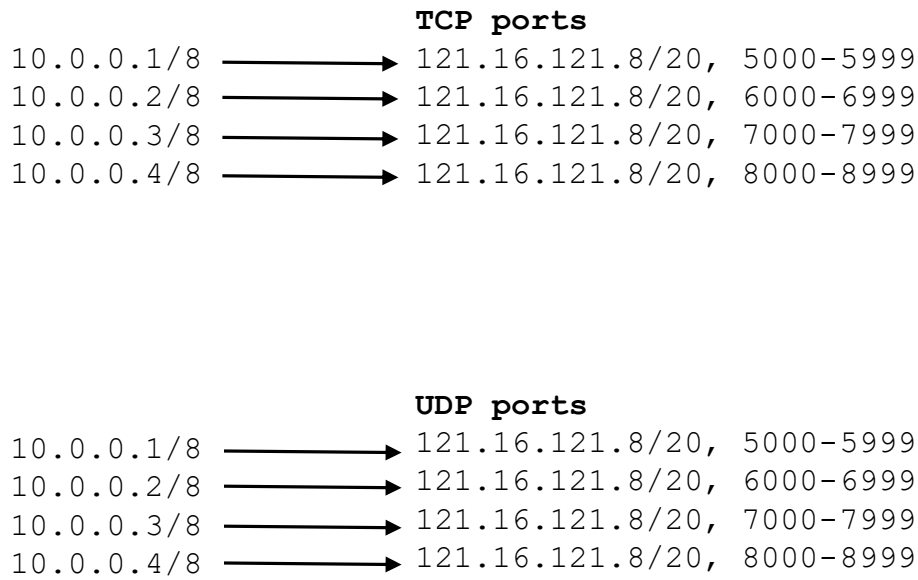
Network Address Translation (NAT) (1)

- Private address to global address mapping (RFC3022)
- Should use RFC1918 addresses
- IP, TCP and UDP checksums
- Caution: applications that use IP address in messages
- **NAT re-writes addresses and port values in packet (so it has to update checksum)**
- **Can map multiple private addresses to a single public address**
- **(NAT not recommended for IPv6, but still used.)**



Network Address Translation (NAT) (2)

- Network address and port translation (NAPT)
- Use layer-4 protocol port numbers:
  - further use of the same IP address
- Typically, private IP address **set**, {P1, P2, P3, P4} maps to **single** global address, G1, using port number range for sharing G1:
  - restriction on number of sessions per private IP address
- This is what a broadband router does for a home network:
  - typically, 192.168/16 is used
  - single global IPv4 address from ISP



Problems with NAT

- RFC2993 (Architectural Implications of NAT) says we should not use NAT!
- Lose end-to-end model:
  - IP addresses different at end-point of connections and flows.
  - some applications may need special handling if they rely on addresses for configuration, e.g. may require proxies or application-level gateways (ALGs)
- NAT makes end-to-end security harder:
  - site-to-site still possible.
- Other problems, e.g. single point of failure.

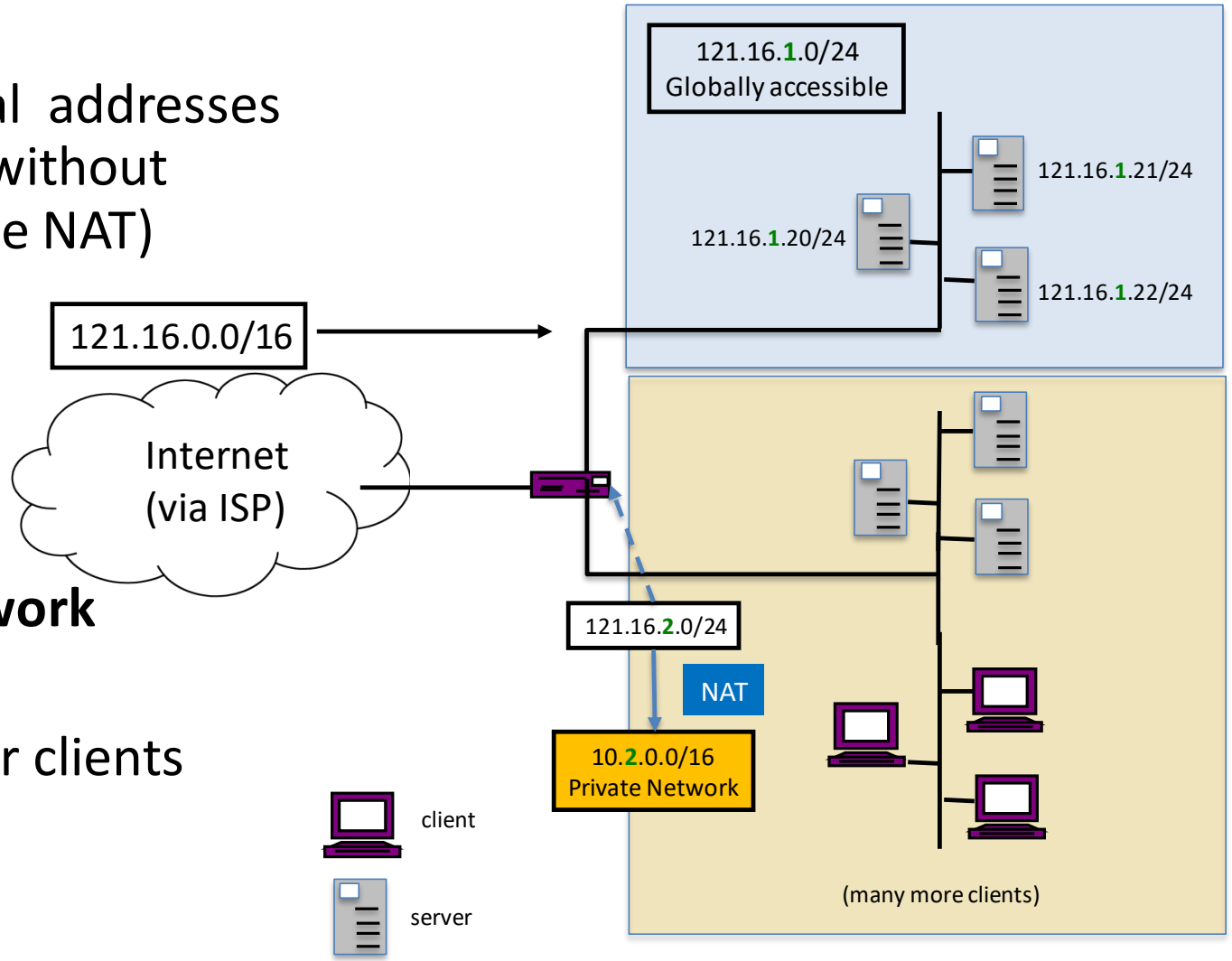
RFC 2993 - conclusion

"The original IP architecture is powerful because it provides a general mechanism on which other things (yet unimagined) may be built. While it is possible to build a house of cards, time and experience have lead to building standards with more structural integrity. IPv6 is the long-term solution that retains end-to-end transparency as a principle. NAT is a technological diversion to sustain the lifetime of IPv4."

Example: site network, NAT + global (1)

Hybrid scheme:

- global servers, global addresses
- clients can operate without global addresses (use NAT)



External view: **one network**

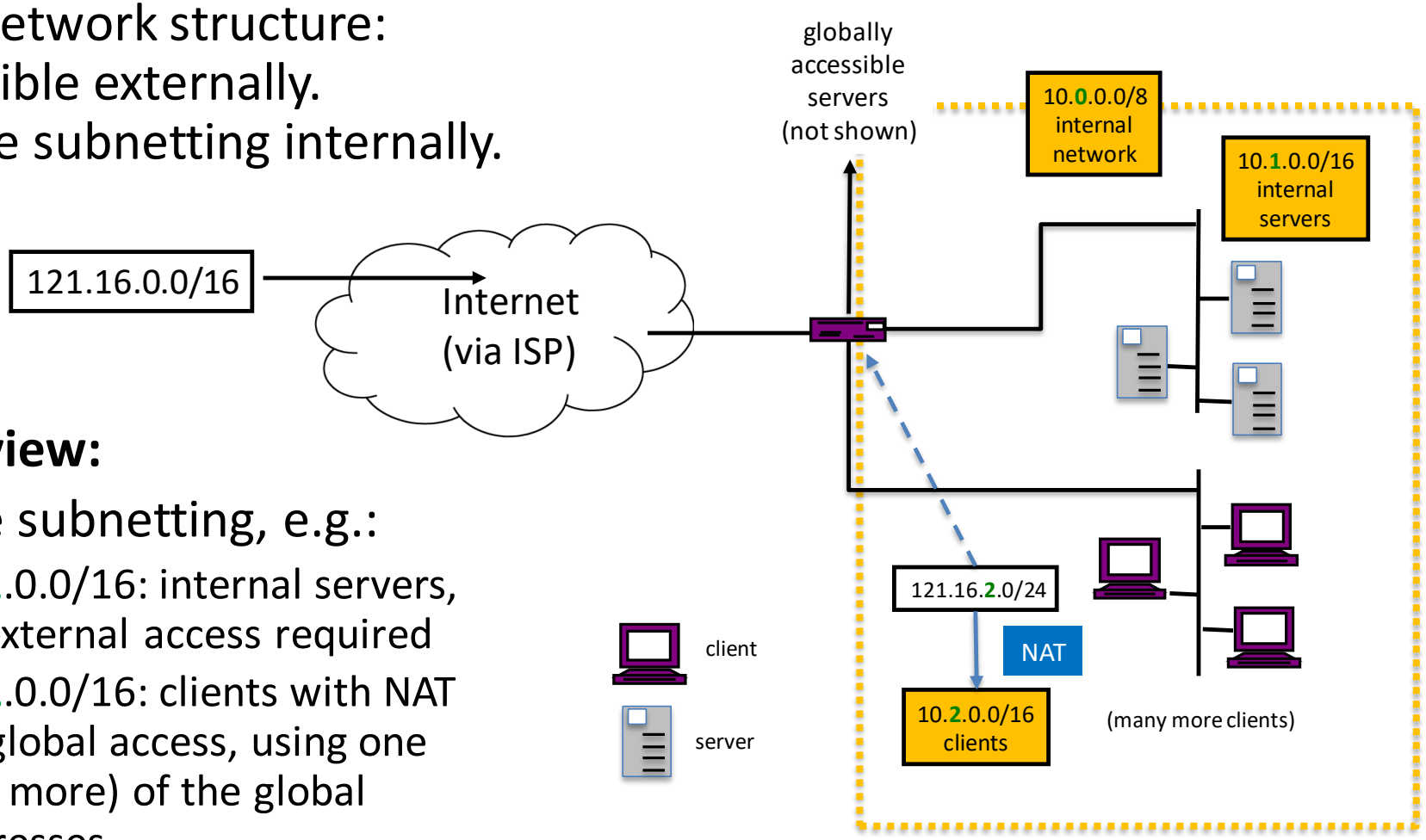
Internal view:

- private addressing for clients and internal servers.
- **NAT for clients.**

Example: site network, NAT + global (2)

Internal network structure:

- Not visible externally.
- Can use subnetting internally.



Internal view:

- private subnetting, e.g.:
  - 10.1.0.0/16: internal servers, no external access required
  - 10.2.0.0/16: clients with NAT for global access, using one (ore more) of the global addresses.

University of St Andrews

Summary

- IP address allocation:
  - Hierarchy and registries.
- Routing and forwarding:
  - Distinction between routing and forwarding.
  - Use of longest prefix match.
- Private IP address space and NAT.
- Further reading: Peterson & Davie Ch 3.3-3.4; Kurose & Ross Ch 4.1-4.3; RFC 2993