CSCI1680
Network Layer: IP & Forwarding

Nick DeMarinis

Based partly on lecture notes by Rodrigo Fonseca, David Mazières, Phil Levis, John Jannotti

Administivia

- Snowcast: due Wednesday (9/27)
- IP project: out Thursday, fill out group preference form by Thursday 11:59pm

  ↳ GEARUP THURS 5-7PM +ZOOM + REC.

- **HW1: out later today, due next Thurs**
  - **Some practice for IP!**

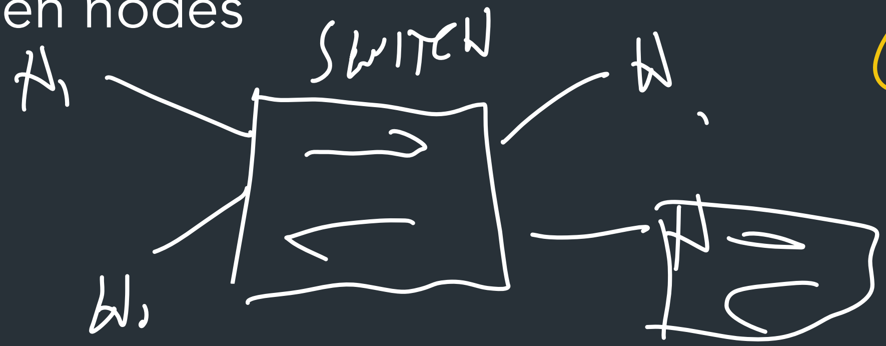
Today

Start of network layer

- Network layer: Internet Protocol (IP) (v4)
- Mechanics of IP forwarding
- Intro to IP project

Last time: link-layer

- Hosts connect to network via <u>interfaces</u>
- Every interface has a link-layer address
  Ethernet/Wifi: MAC address (`0c:45:22:c1:be:03`)
- Switches, Wifi APs: in-network devices that forward packets between nodes



(OFFICE, HOUSE)
≈ 100 ISH HOSTS

SWITCH MAPS
MAC ADDRS TO PACKETS.

Last time: link-layer

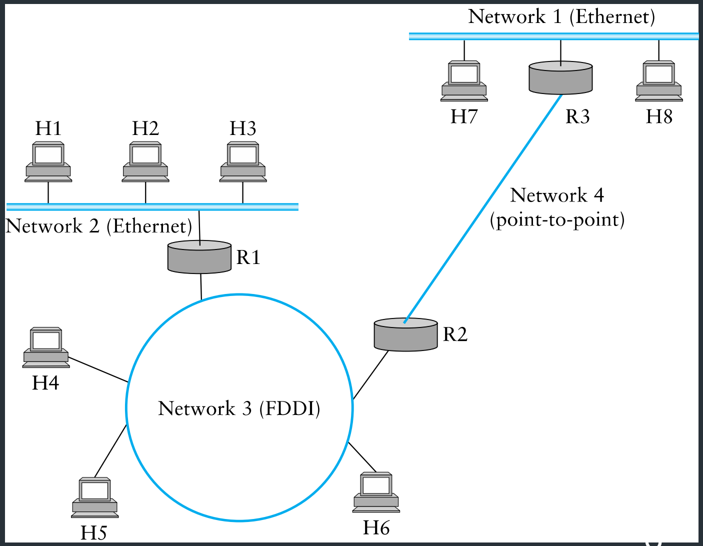
- Hosts connect to network via interfaces
- Every interface has a link-layer address
  Ethernet/Wifi: MAC address (`0c:45:22:c1:be:03`)
- Switches, Wifi APs: in-network devices that forward packets between nodes

Mental model for the link layer
=> How to connect hosts on a "small" network
=> Given link-layer address, know how to reach host on your network

Layers, Services, Protocols

Application	Service: user-facing application. Application-defined messages
Transport	Service: multiplexing applications Reliable byte stream to other node (TCP), Unreliable datagram (UDP)
L3 Network	Service: move packets to any other node in the network Internet Protocol (IP)
L2 Link	Service: move frames to other node across link. May add reliability, medium access control
Physical	Service: move bits to other node across link

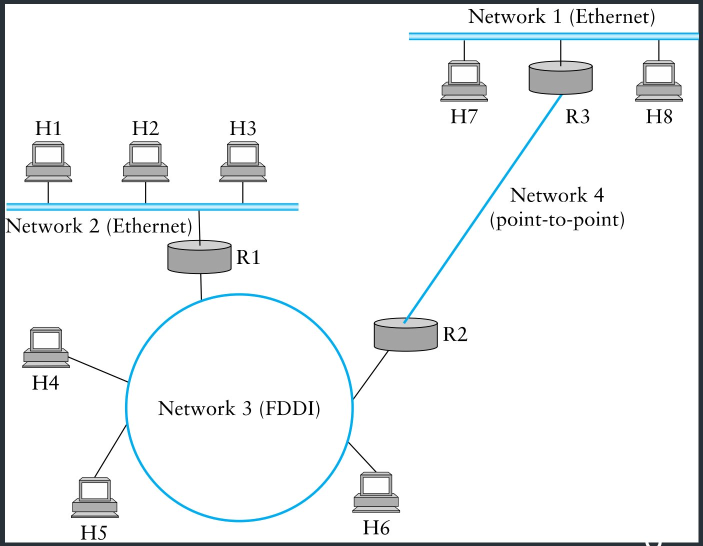
Internet Protocol (IP) Goals



Internet Protocol (IP) Goals

How to connect everyone?

- Glue lower-level networks together
- A network of networks!
- Router: device that forwards packets between networks

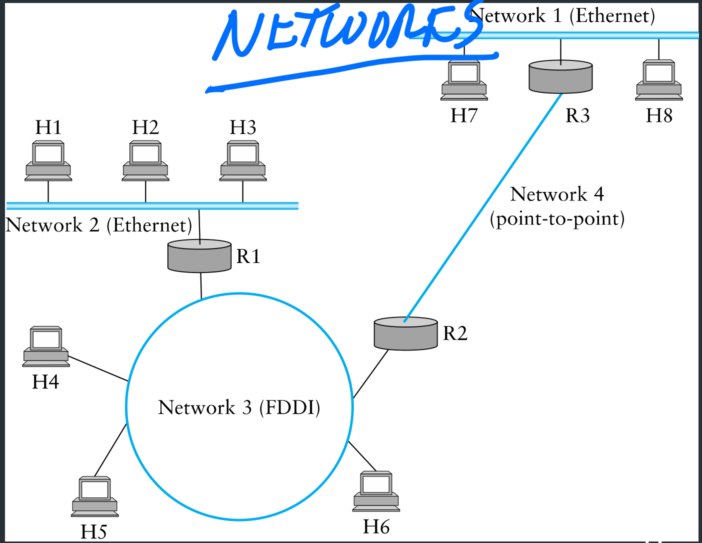


Internet Protocol (IP) Goals

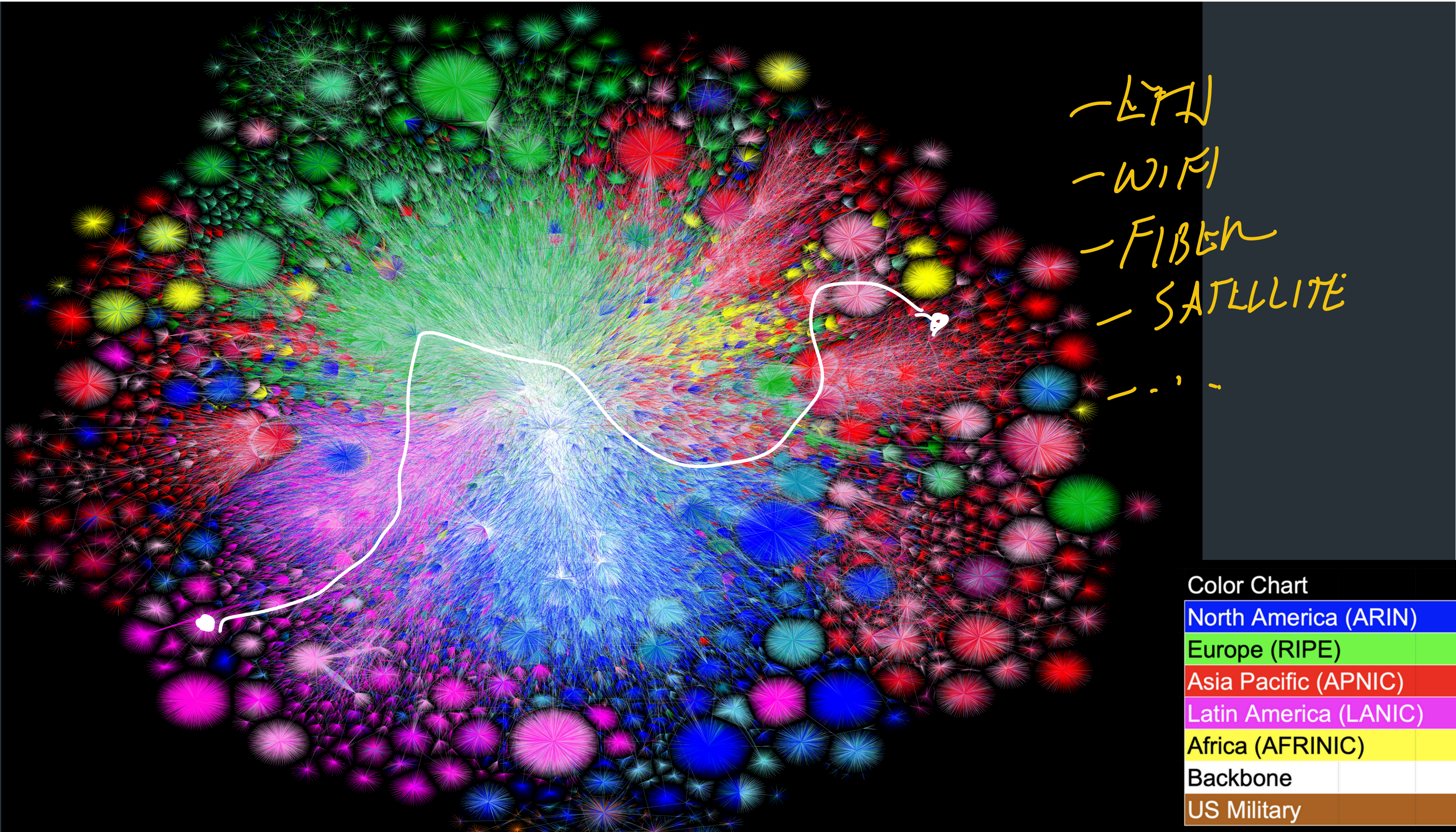
How to connect everyone?

- Glue lower-level networks together
- A network of networks!
- Router: device that forwards packets between networks

=> Doesn't that sound like switching?



-ETH
-WIFI
-FIBER
- SATELLITE
- ...



Map of the Internet, 2021 (via BGP)
OPTE project

New Challenges

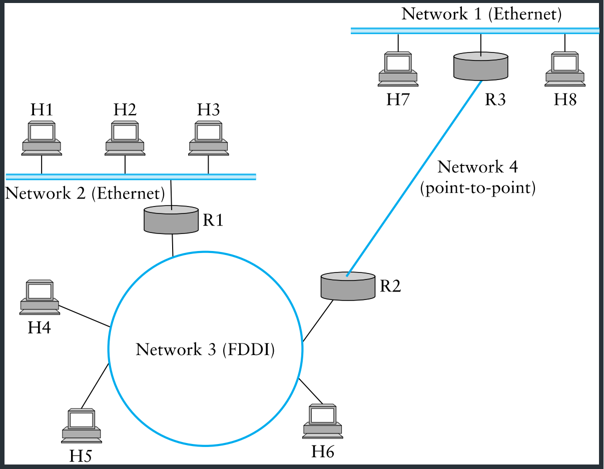
- Networks are heterogeneous (eg. Wifi vs. Ethernet)
  - Different frame formats
  - Some are more reliable than others
  - Different packet sizes/bandwidths
- Scaling: link-layer strategies don't work!

CAN'T MAKE TOO MANY ASSUMPTIONS ABOUT L2

SWITCH: TABLE

MAC ADDR => PORT

NEED DIFF STRATEGIES!



What came before the Internet?

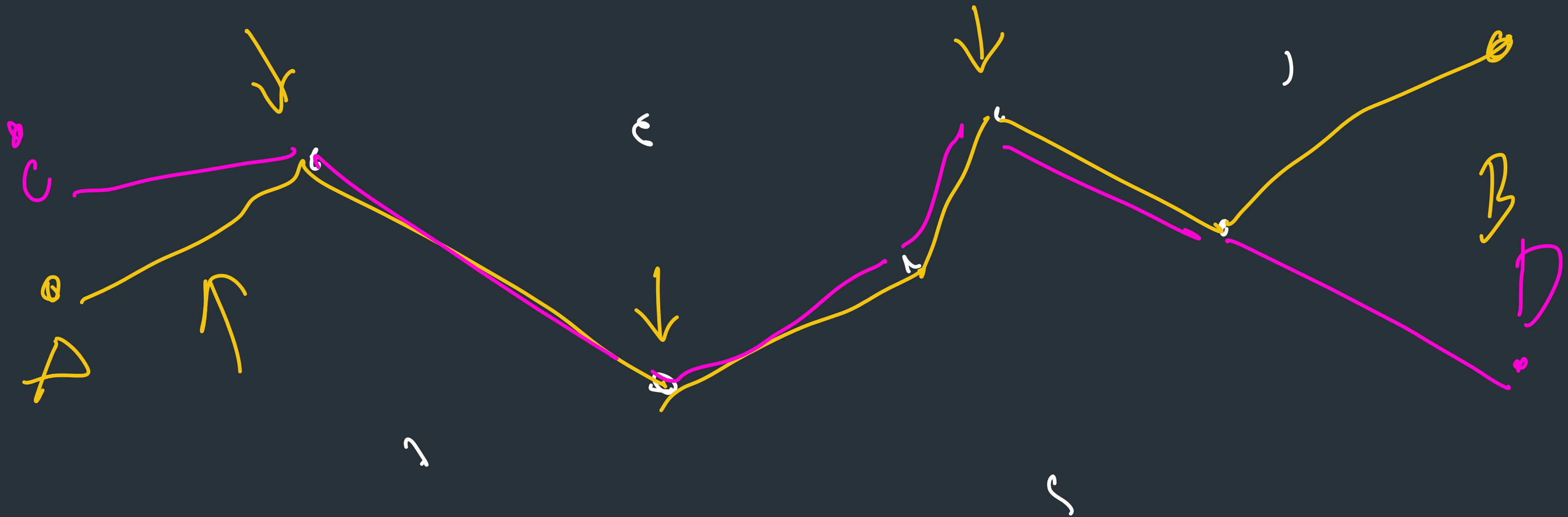
The (landline) telephone network

(Plain Old Telephone Service (POTS))

At the time the internet was forming (~1960s), landline phones were a big network that connected lots of end-users (160M worldwide)

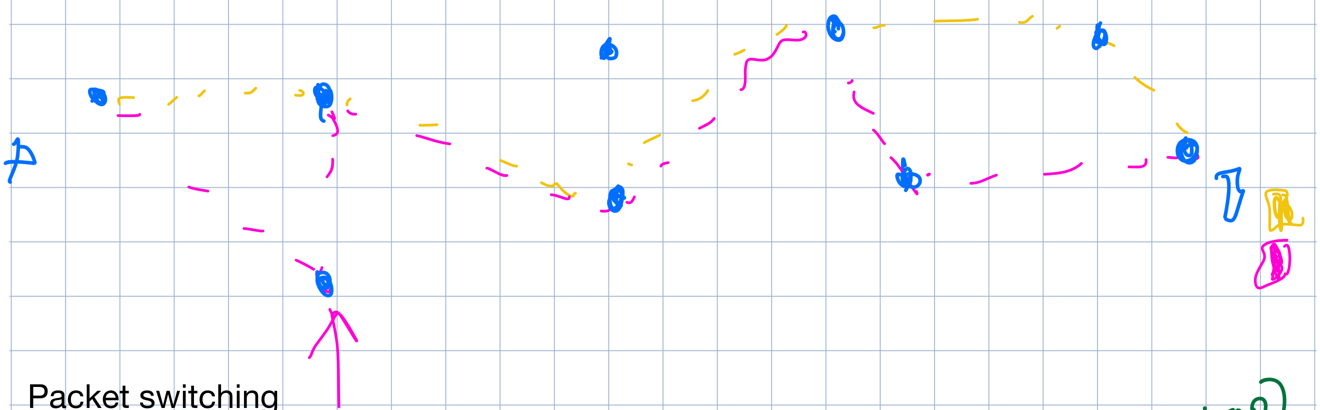
Why should we care?
⇒ Provides a useful comparison
⇒ Understand some of the design goals of the time

EARLY PHONE NETWORKS: CIRCUIT SWITCHING



- SETUP CONNECTION BEFORE YOU TALK
- NODES NEED TO REMEMBER CONNECTIONS

PACKET SWITCHING: ROUTERS MOVE SMALL CHUNKS OF DATA ("PACKETS")



Packet switching

- Routers just need to pick destination for next packet
- No initial setup
- Messages must get divided up into small units (and reassembled at other end (TCP)
- Packets may take any path (can be different for each packet)

Early telephone networks

Used circuit switching: set up whole path for call beforehand

- Does it scale?

WORKED FOR TELEPHONES UNTIL 1980s,
BUT INTERNET HAS VERY DIFFERENT DESIGN GOALS.

A Bit of History

Early Packet switched networks: Arpanet's IMPs

- Late 1960's => RFC 1, 1969!
- Reliable network with many features we know today

- Build apps with Network Control Program (NCP)
  - Built on reliable IMPs
  - Used by programs like telnet, mail, file transfer

A Bit of History

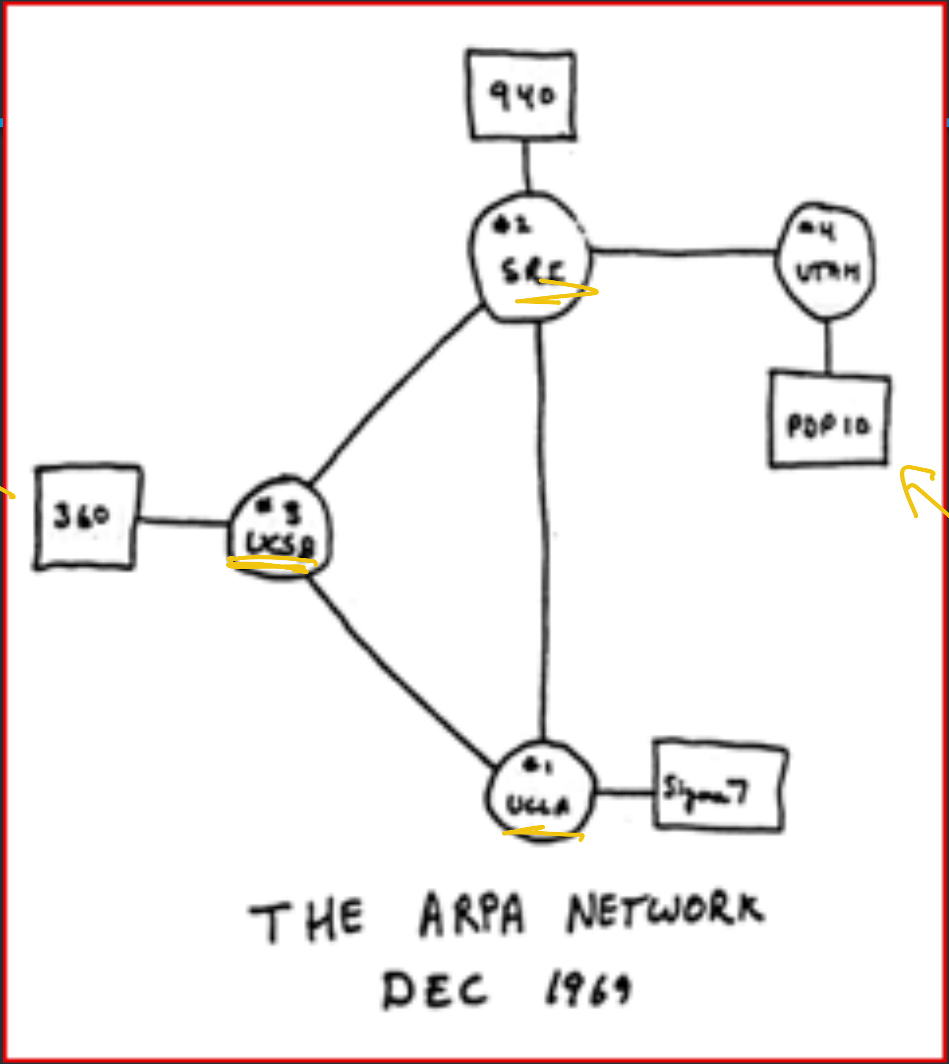
Early Packet switched networks: Arpanet's IMPs

- Late 1960's => RFC 1, 1969!
- Reliable network with many features we know today

- Build apps with Network Control Program (NCP)
  - Built on reliable IMPs
  - Used by programs like telnet, mail, file transfer

What about when network isn't reliable?

S/360



PDP 10
(CONNECTING SINGLE COMPUTERS!!)

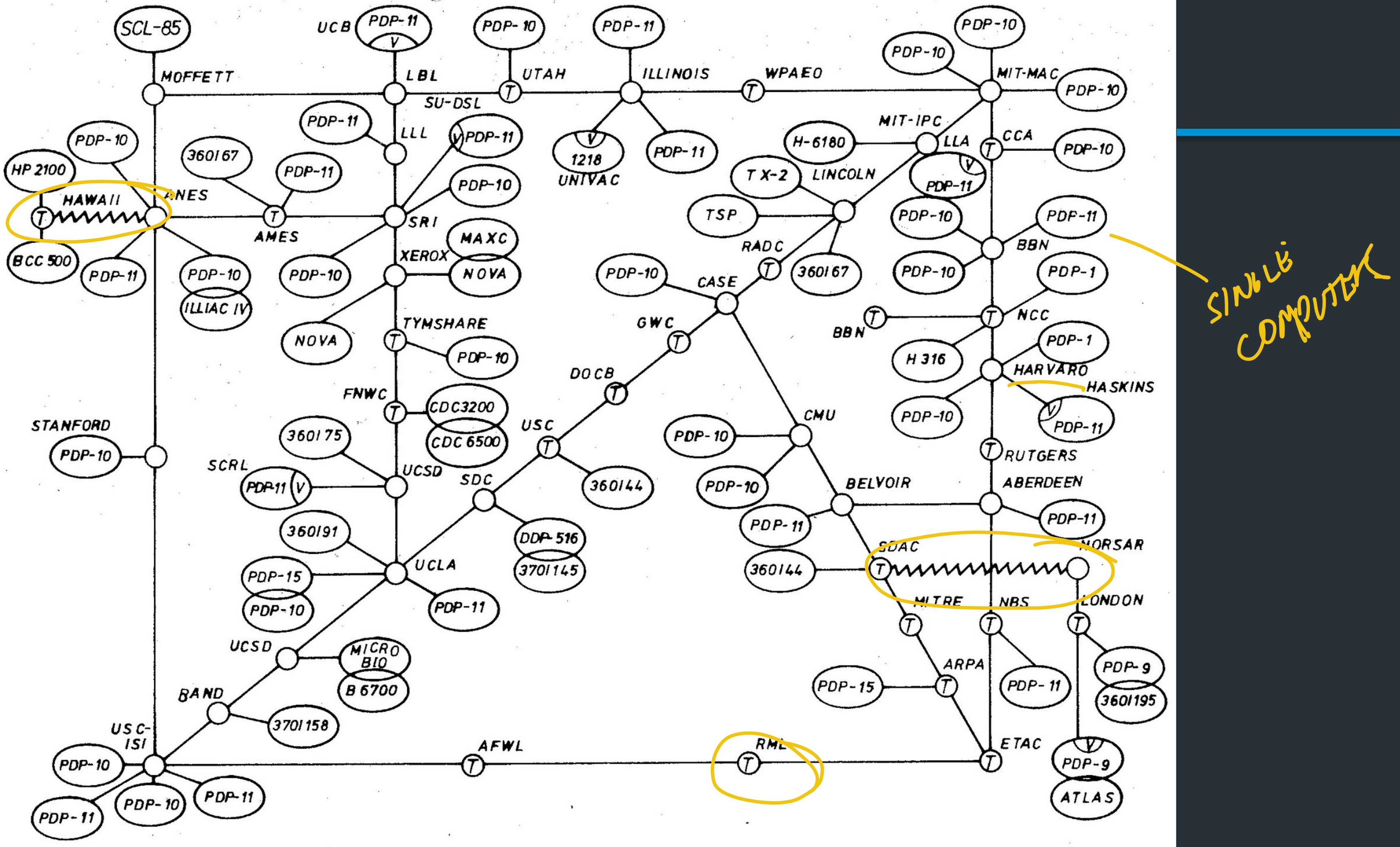


Abb. 4 ARPA NETwork, topologische Karte. Stand Juni 1974.

How to make such a protocol?

- How to deal with heterogeneous networks?
- How to find hosts?
- Should messages be reliable or unreliable?
- What to do when a device joins/leaves?
- …

NEED TO WORK W/ "LEAST COMMON DENOMINATOR" OF FUNCTIONALITY"

Big concerns

⇒ Not every application needs all features

⇒ Can't assume much functionality from (heterogeneous) link layers

1974: TCP/IP Introduced

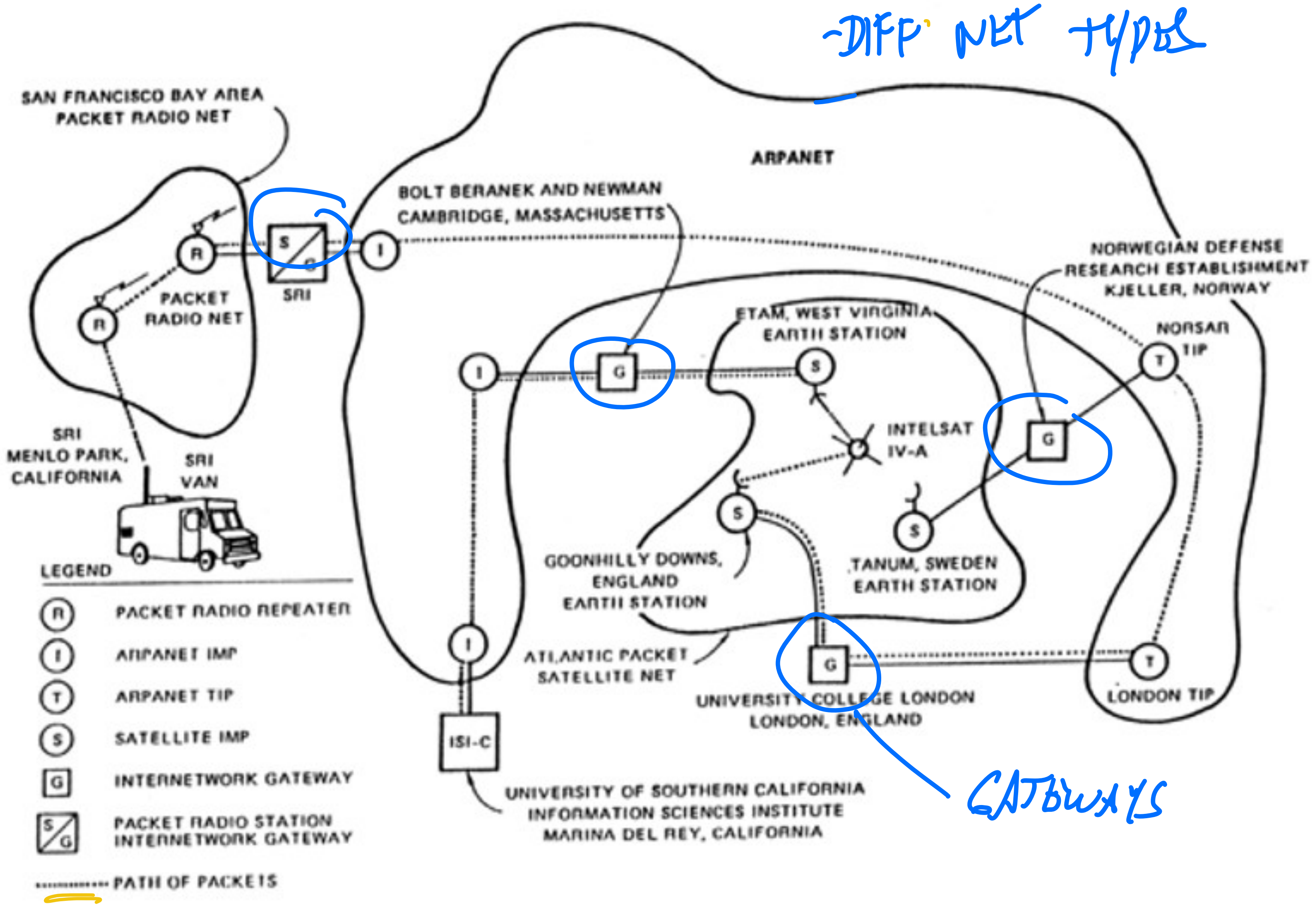
- Vint Cerf, Robert Kahn build protocol to replace NCP
- Initial design: single protocol providing a reliable pipe

1974: TCP/IP Introduced

- Vint Cerf, Robert Kahn build protocol to replace NCP
- Initial design: single protocol providing a reliable pipe

- Eventually, separated into different protocols we know today
  - IP: basic datagram service among hosts
  - TCP: reliable transport
  - UDP: unreliable multiplexed datagram service

"PACKET" OR "MESSAGE"



-DIFF NET TYPES
SAN FRANCISCO BAY AREA
PACKET RADIO NET
ARPANET
BOLT BERANEK AND NEWMAN
CAMBRIDGE, MASSACHUSETTS
PACKET
RADIO NET
SRI
NORWEGIAN DEFENSE
RESEARCH ESTABLISHMENT
KJELLER, NORWAY
ETAM, WEST VIRGINIA
EARTH STATION
NORSAR
TIP
INTELSAT
IV-A
SRI
MENLO PARK,
CALIFORNIA
SRI
VAN
GOONHILLY DOWNS,
ENGLAND
EARTH STATION
TANUM, SWEDEN
EARTH STATION
ATLANTIC PACKET
SATELLITE NET
UNIVERSITY COLLEGE LONDON
LONDON, ENGLAND
LONDON TIP
ISI-C
UNIVERSITY OF SOUTHERN CALIFORNIA
INFORMATION SCIENCES INSTITUTE
MARINA DEL REY, CALIFORNIA
GATEWAYS
LEGEND
PACKET RADIO REPEATER
ARPANET IMP
ARPANET TIP
SATELLITE IMP
INTERNETWORK GATEWAY
PACKET RADIO STATION
INTERNETWORK GATEWAY
PATH OF PACKETS

IP's Decisions

- Connectionless, packet-switched network

  = NO CONNECTION SETUP.

- "Best-effort" service: other layers add reliability if you need it

  NO GUARANTEES ON RELIABILITY W/ IP ALONE.

How to reach hosts?

- Common message format: IP header
- Every host identified by an IP address

IP's Decisions

- Connectionless, packet-switched network

  => Routers are "simple" => no connection state

- "Best-effort" service: other layers add reliability if you need it

  => Packets might be dropped, reordered, delayed, …

How to reach hosts?

- Common message format: IP header
- Every host identified by an IP address

An excellent read

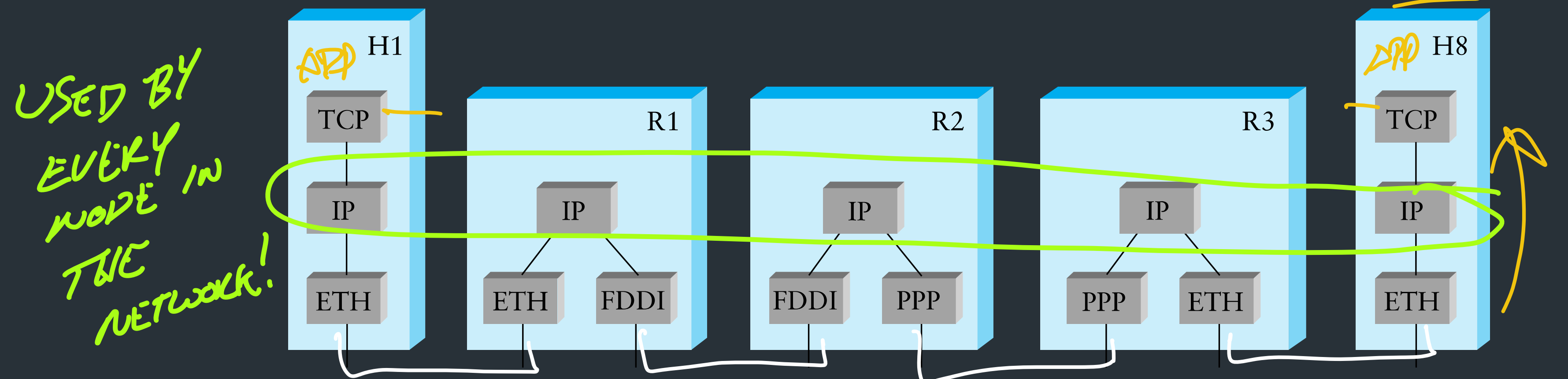
David D. Clark, "The design Philosophy of the DARPA Internet Protocols", 1988

- Primary goal: multiplexed utilization of existing interconnected networks
- Other goals:
  - Communication continues despite loss of networks or gateways
  - Support a variety of communication services
  - Accommodate a variety of networks
  - Permit distributed management of its resources
  - Be cost effective
  - Low effort for host attachment
  - Resources must be accountable

The Internet Protocol

IP runs on all hosts and routers

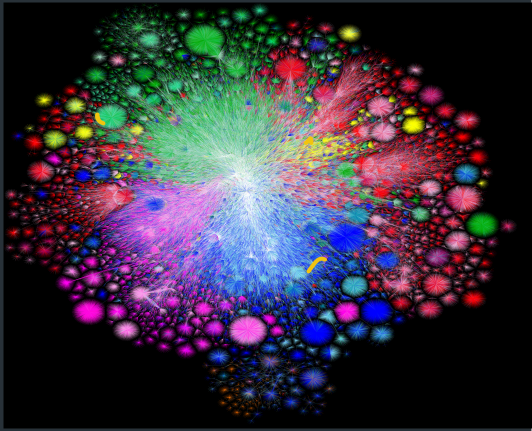
- Addressing: how we name nodes in an IP network
- Provides forwarding: how routers move packets based on the destination address
- (later) Routing: how routers build forwarding rules



IP Addressing

Overview

- Unique number to identify “all” hosts on the Internet
- A number with structure => the number tells the network where the host is



Analogy: back to phones

Telephone numbers have a structure to them

+ 1 401 863 1000

COUNTRY CODE

AREA CODE (RI)

BROWN UNIVERSITY

Analogy: back to phones

Telephone numbers have a structure to them

+ 1 401 863 1000

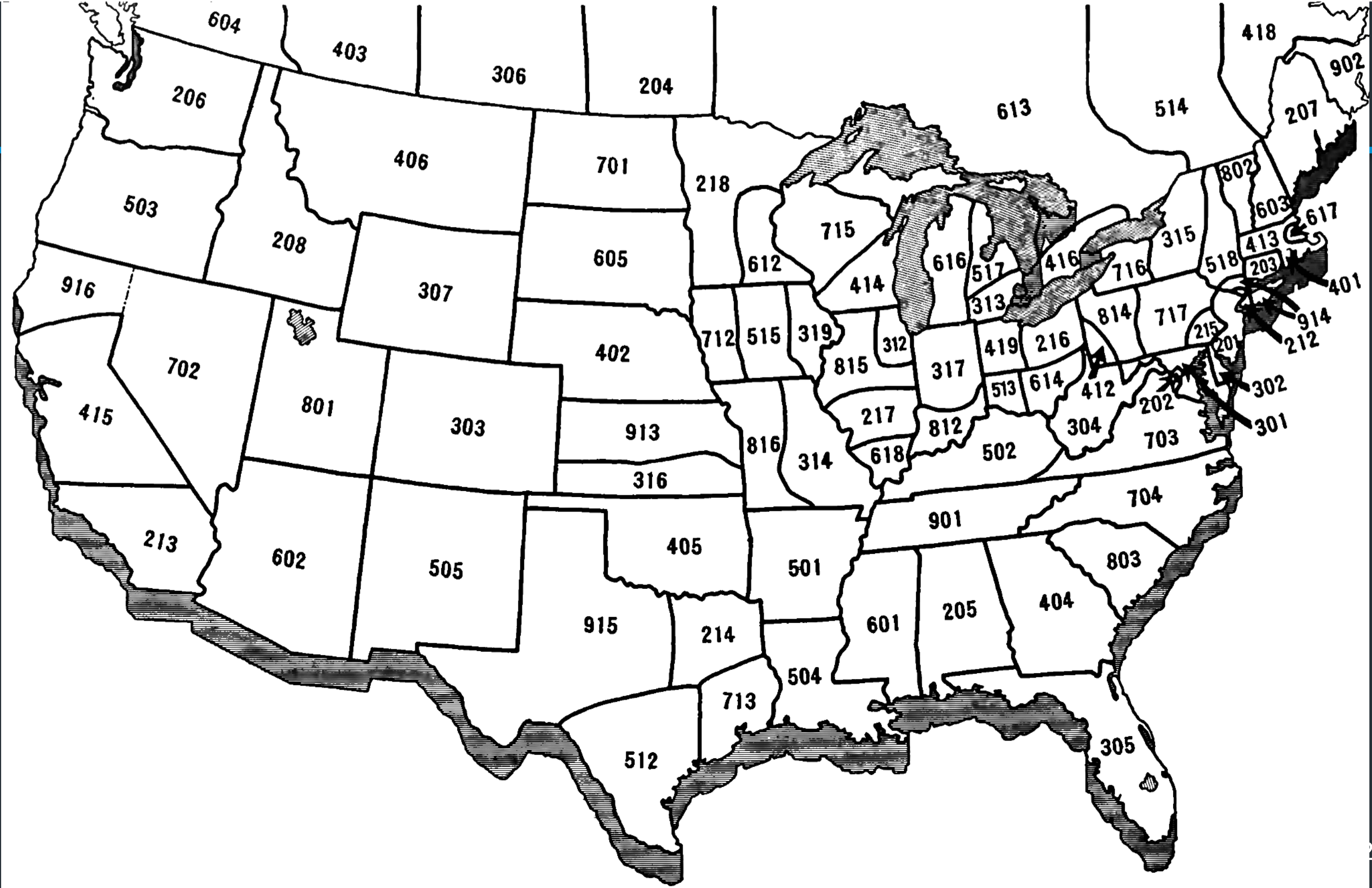
RI

+1 212 555 4253

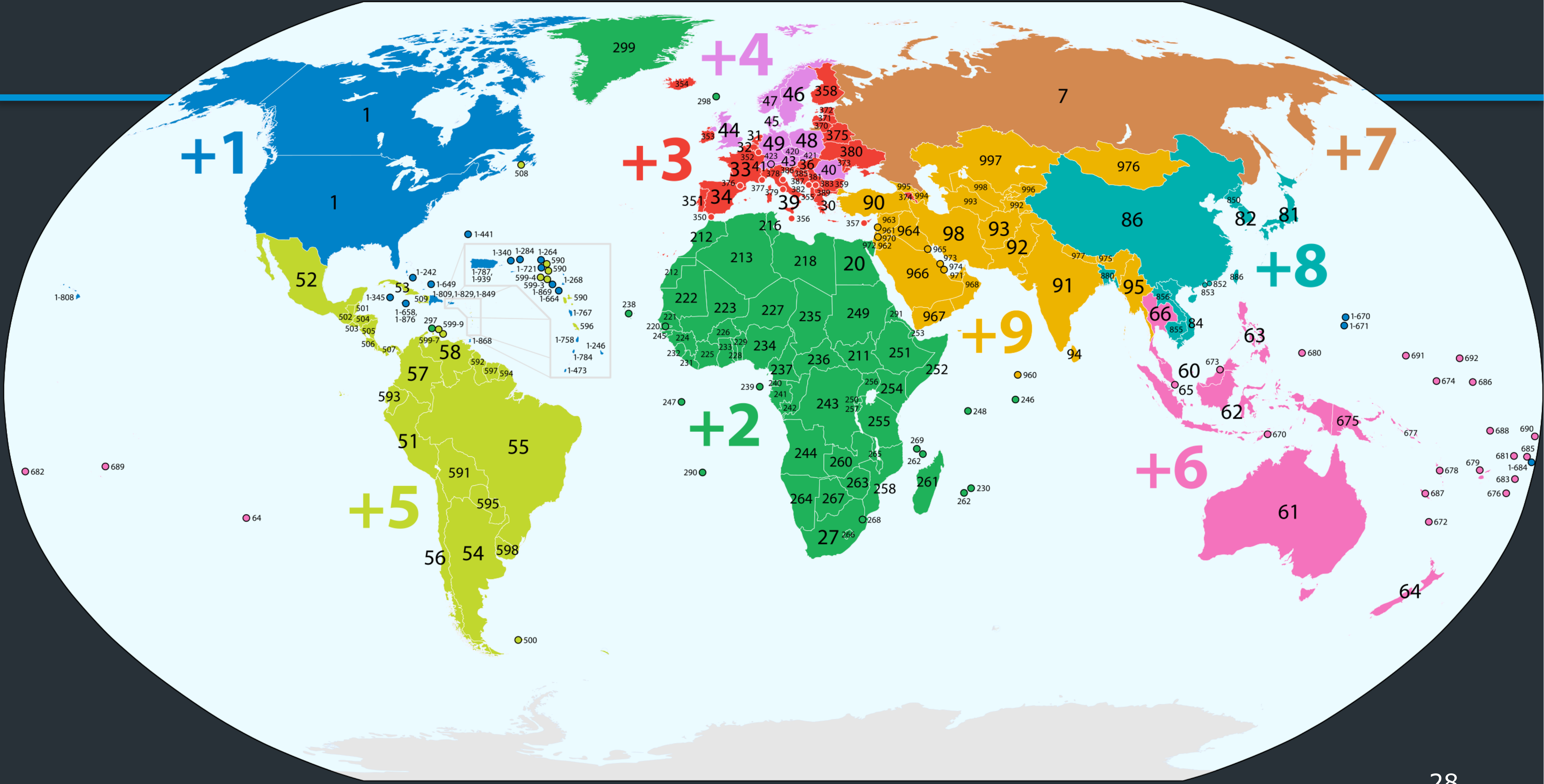
MANHATTAN

⇒ AT ONE TIME, USED TO ROUTE CALLS TO A GEOGRAPHIC LOCATION.

Part of the number tells where you are!
(or at least it did before cell phones)



604
403
306
204
418
902
206
613
514
207
406
701
218
503
802
603
617
715
315
208
605
612
616
517
416
413
716
518
203
401
916
307
414
313
814
717
914
712
515
319
312
402
815
419
216
215
201
212
702
317
513
614
412
302
202
801
217
415
303
913
816
314
812
304
703
301
618
502
316
704
213
901
602
505
405
501
803
404
205
601
915
214
504
713
305
512


+1
+2
+3
+4
+5
+6
+7
+8
+9
1
1
299
354
298
47
46
358
44
45
353
31
32
49
48
375
380
352
423
420
421
43
36
40
373
33
41
386
385
381
378
387
383
359
376
377
379
382
389
355
351
34
39
30
356
350
357
90
995
374
994
963
961
964
970
972
962
965
973
974
971
968
966
967
98
93
92
997
998
996
993
992
976
7
86
850
82
81
886
852
853
977
975
880
91
95
856
66
855
84
94
960
63
673
60
65
62
670
675
61
64
680
691
692
674
686
677
688
690
685
681
1-684
679
678
683
687
676
672
1-670
1-671
212
213
218
20
216
222
223
227
235
249
291
253
221
220
245
224
226
229
234
232
231
225
233
228
237
236
211
251
252
240
241
242
239
243
250
257
256
254
255
247
238
244
260
265
269
262
263
258
261
230
264
267
268
27
266
290
248
246
52
53
501
502
504
503
505
506
507
1-345
1-658,
1-876
509
1-242
1-649
1-809,1-829,1-849
1-441
1-340
1-284
1-264
1-787,
1-939
1-721
590
599-4
599-3
1-268
1-869
1-664
590
1-767
596
1-758
1-246
1-784
1-473
1-868
297
599-9
599-7
58
592
597
594
57
593
51
55
591
595
56
54
598
500
508
1-808
682
689
64

IP Addressing

IP Version 4: Each address is a 32-bit number: 128.148.16.7

IPV6 = 128 BIT ADDRESSES.

128.148.16.7

OCTET

10000000 10010100 00010000 00000111

32 BITS = 2^{32} POSSIBLE ADDRESSES

≈ 4 BILLION

Notation

- Write each byte (“octet”) as a decimal number
- This is called ”dotted decimal” or ”dotted quad” notation

IP Addressing

IP Version 4: Each address is a 32-bit number:

128.148.16.7

10000000 10010100 00010000 00000111

32 bits => 2^{32} possible addresses… problem?

<u>Notation</u>

- Write each byte ("octet") as a decimal number
- This is called "dotted decimal" or "dotted quad" notation

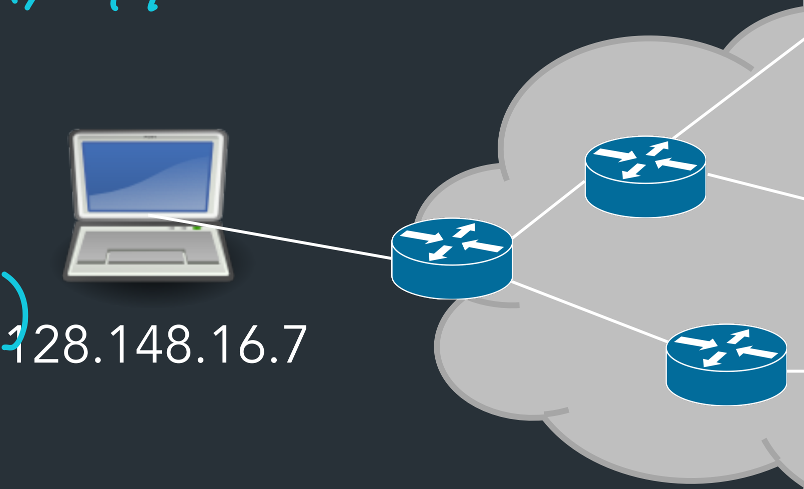
IP Addressing

An IP address identifies…

HOST #1247

- Who a host is: A unique number
- Where it is on the Internet

  ↳ IDENTIFIES THE NETWORK
  (EG. Brown, Amazon, Google,…)



- Networks are allocated ranges of IPs by global authority (ICANN)
  - Further subdivided by regions, ISPs, …
- Some IPs have special uses (eg. 127.0.0.1)

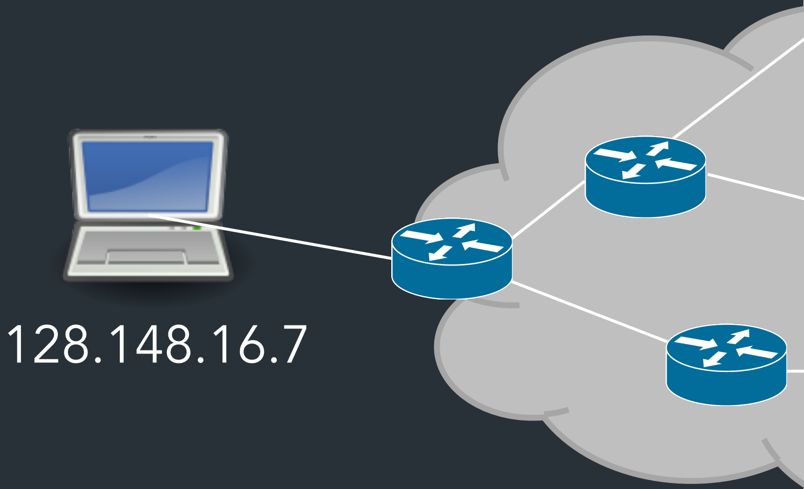
  ↳ LOCALHOST

*ICANN (Internet Corporation for Assigned Names and Numbers)

IP Addressing

An IP address identifies…

- Who a host is: A unique number
- Where it is on the Internet



- Networks are allocated ranges of IPs by global authority (ICANN)
  - Further subdivided by regions, ISPs, …
- Some IPs have special uses (eg. 127.0.0.1)

eg. Brown owns 128.148.xxx.xxx, 138.16.xxx.xxx

*ICANN (Internet Corporation for Assigned Names and Numbers)

IP Addressing

BROWN OWNS 128.148.0.0/16

Brown owns the range:

128.148.xxx.xxx

16 BITS

$2^{16} \approx 65K$ HOSTS

10000000 10010100 xxxxxxxx xxxxxxxx

16 BITS

INDIVIDUAL HOSTS

<u>Network part</u>

Identifies Brown (to the Internet)

<u>Host part</u>

Denotes individual hosts within the Brown Network

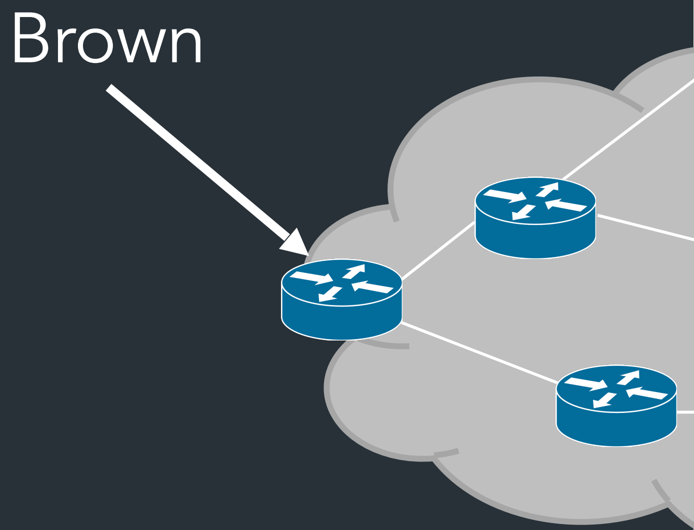
⇒ USED BY INTERNET TO SEND PACKETS TO THIS NETWORK.

Formal way to write this:
128.148.0.0/16

IP Addressing

A network can designate IP addresses for its own hosts within its address range

For 128.148.xxx.xxx:



IP Addressing

A network can designate IP addresses for its own hosts within its address range

For 128.148.xxx.xxx:

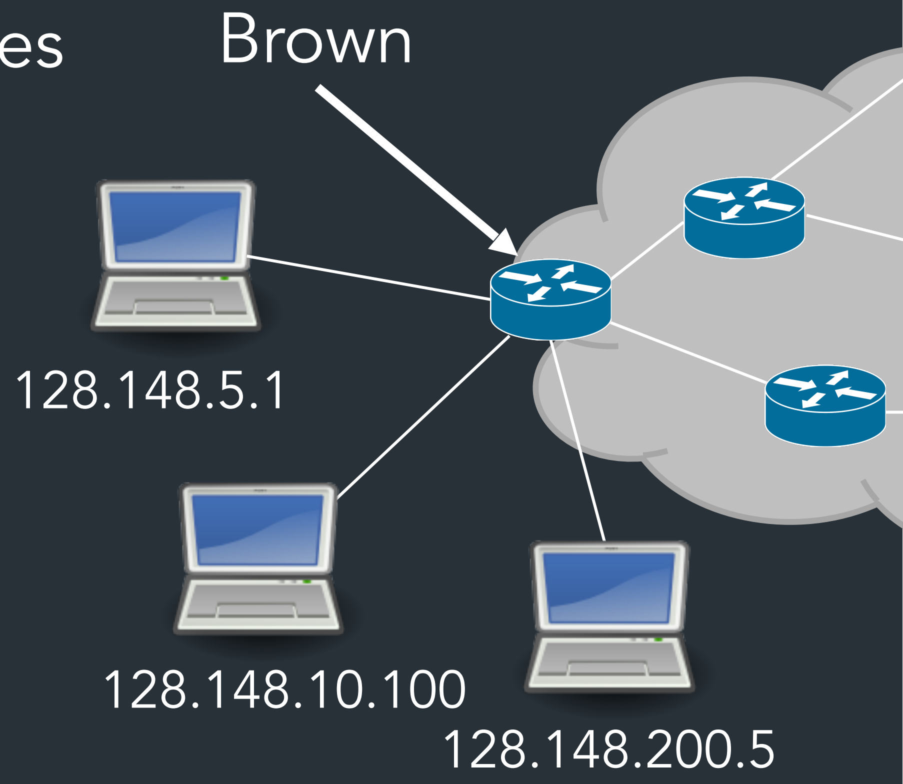
10000000 10010100 xxxxxxxx xxxxxxxx

Brown uses the the prefix 128.148.0.0/16

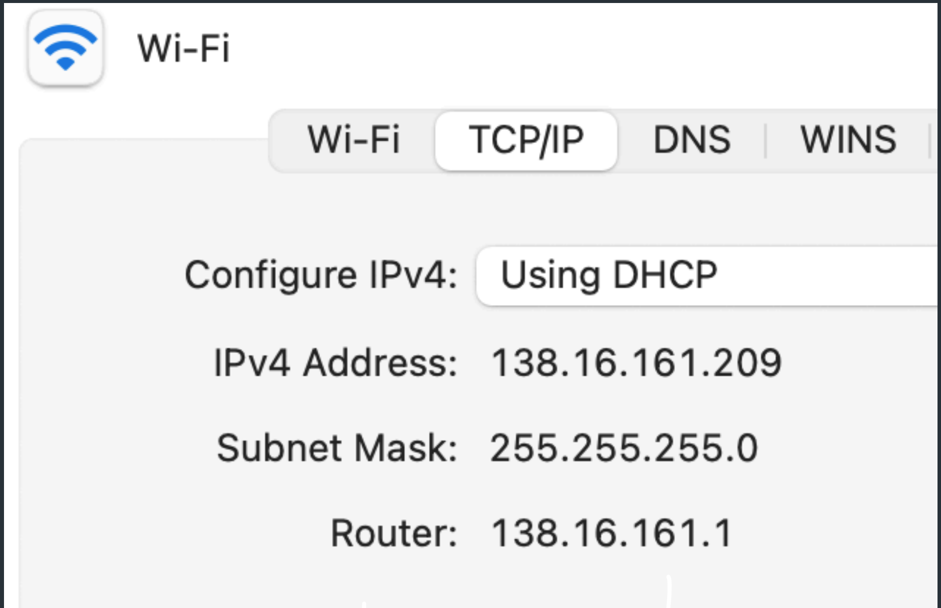
Some other ways to write this:

128.148/16

128.148.0.0 + subnet mask 255.255.0.0



A typical configuration



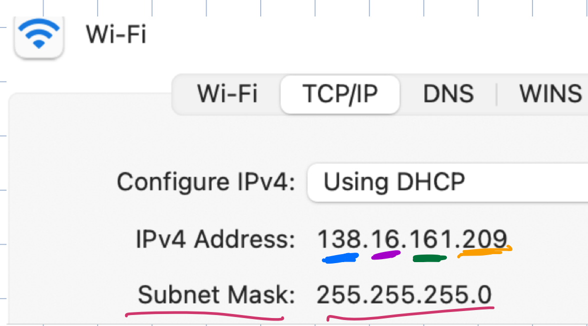
IP ADDRESS

138	16	161	209
1000 1010	0001 0000	1010 0001	1101 1001

SUBNET MASK

138.16.161.0/24

UNPACKING IP SETTINGS: EXAMPLE



IP ADDR	138 1000 1010	16 0001 0000	161 1010 0001	209 1101 1001
SUBNET MASK	255 1111 1111	255 1111 1111	255 1111 1111	0 0000 0000

24 BITS

BITS WITH ONES IN SUBNET MASK IDENTIFY THE NETWORK PART.
SO THIS IS NETWORK IS

138.16.161.0/24

IE, THE SUBNET MASK IS 24 BITS LONG.

Common prefixes

1.2.0.0/16	00000001 00000010 xxxxxxxx xxxxxxxx
8.0.0.0/8	00001000 xxxxxxxx xxxxxxxx xxxxxxxx
123.10.1.0/24	01111011 00001010 00000001 xxxxxxxx
201.112.10.200/30	11001001 01110000 00001010 110010xx

HOST PART: 16 BITS

$2^{16} \approx 65K$ HOSTS

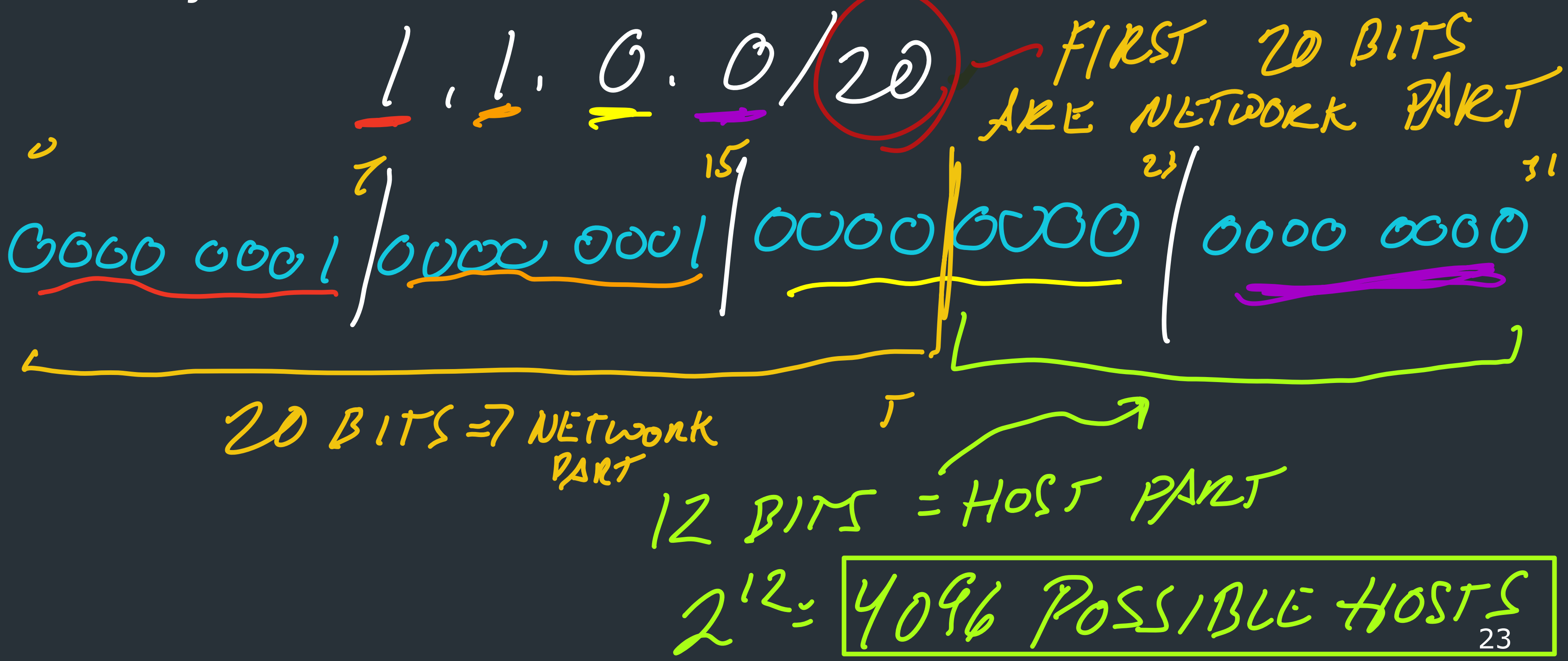
HOST PART $= 2^{32-8} = 2^{24} \approx 16M$ HOSTS

$2^8 = 256$ HOSTS

$2^2 = 4$ HOSTS

Example

How many addresses are in the network 1.1.0.0/20?



IP PREFIX:

A RANGE OF ADDRESSES FOR A NETWORK.

How do we move packets between networks?

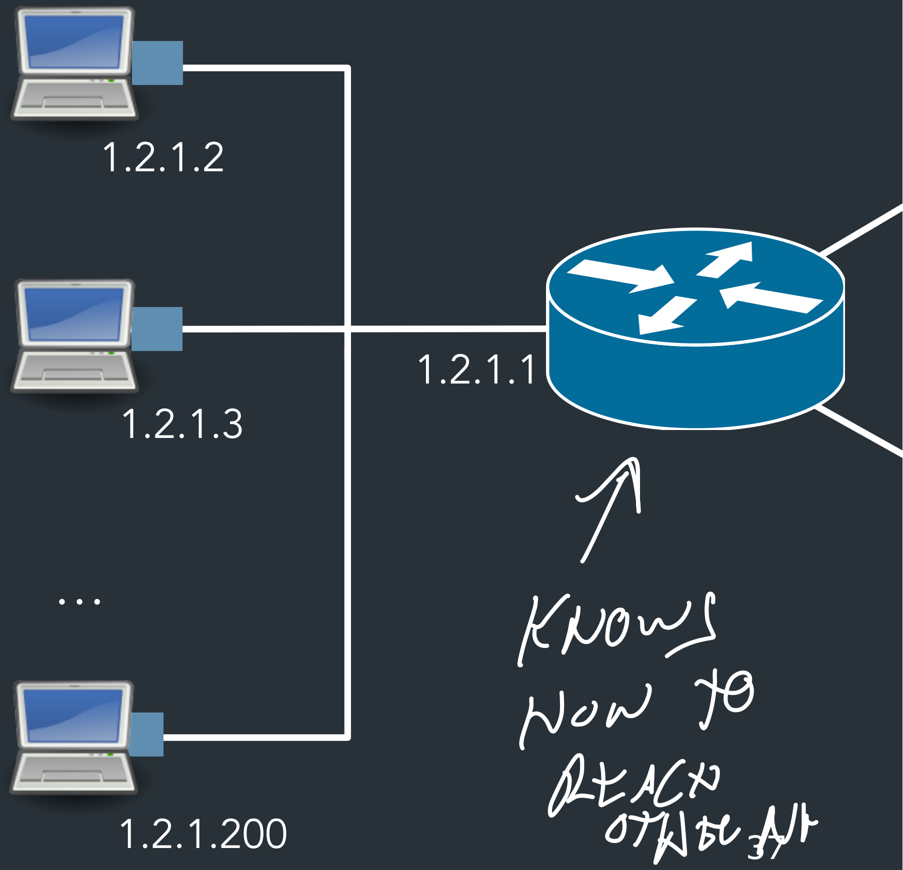
How IP forwarding works

Assume:

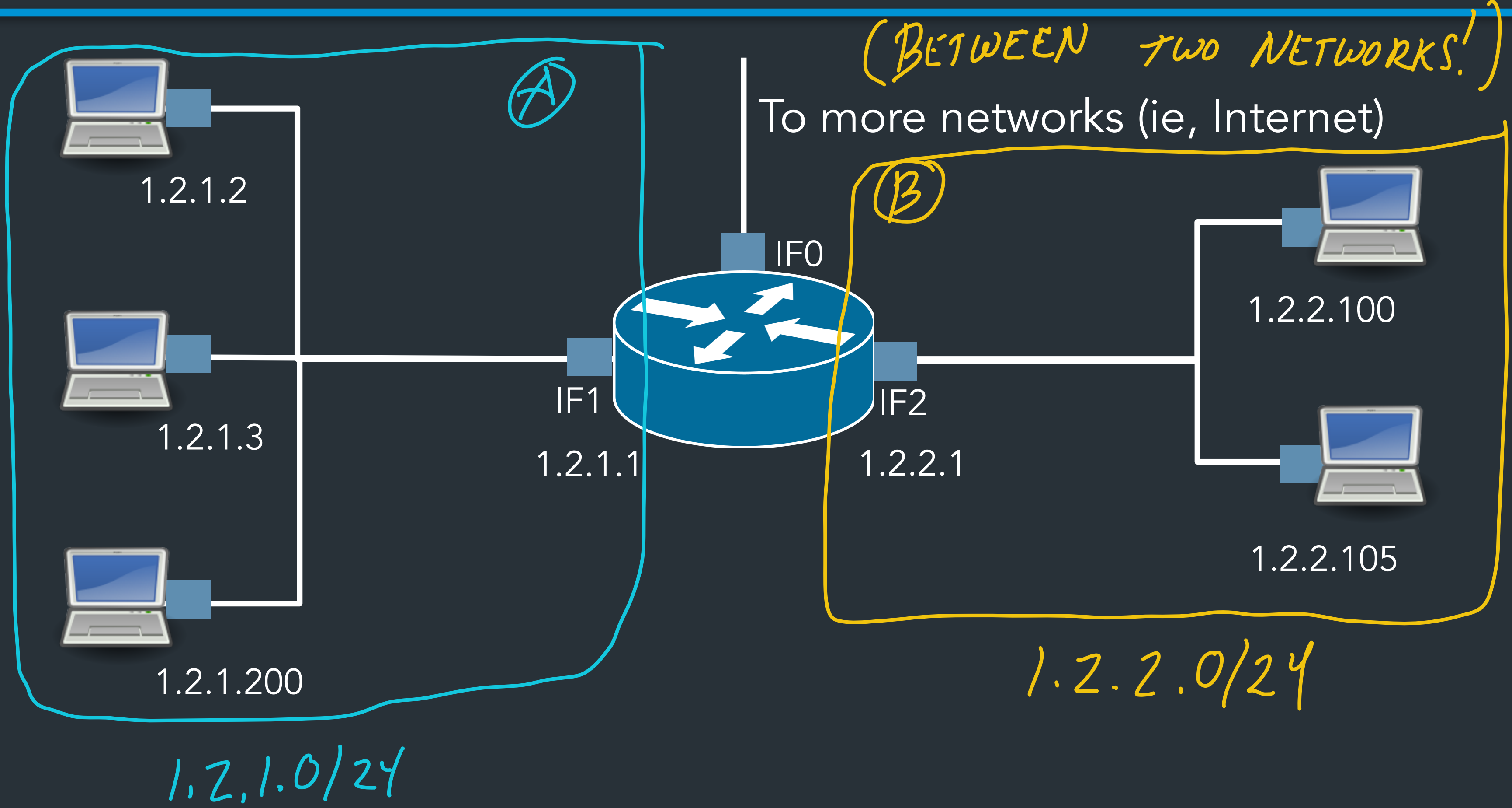
- Communicating on same network is easy—this is the link-layer's job!
- Can map IP addresses to MAC addresses (more on this later)

How to reach an address outside this network?

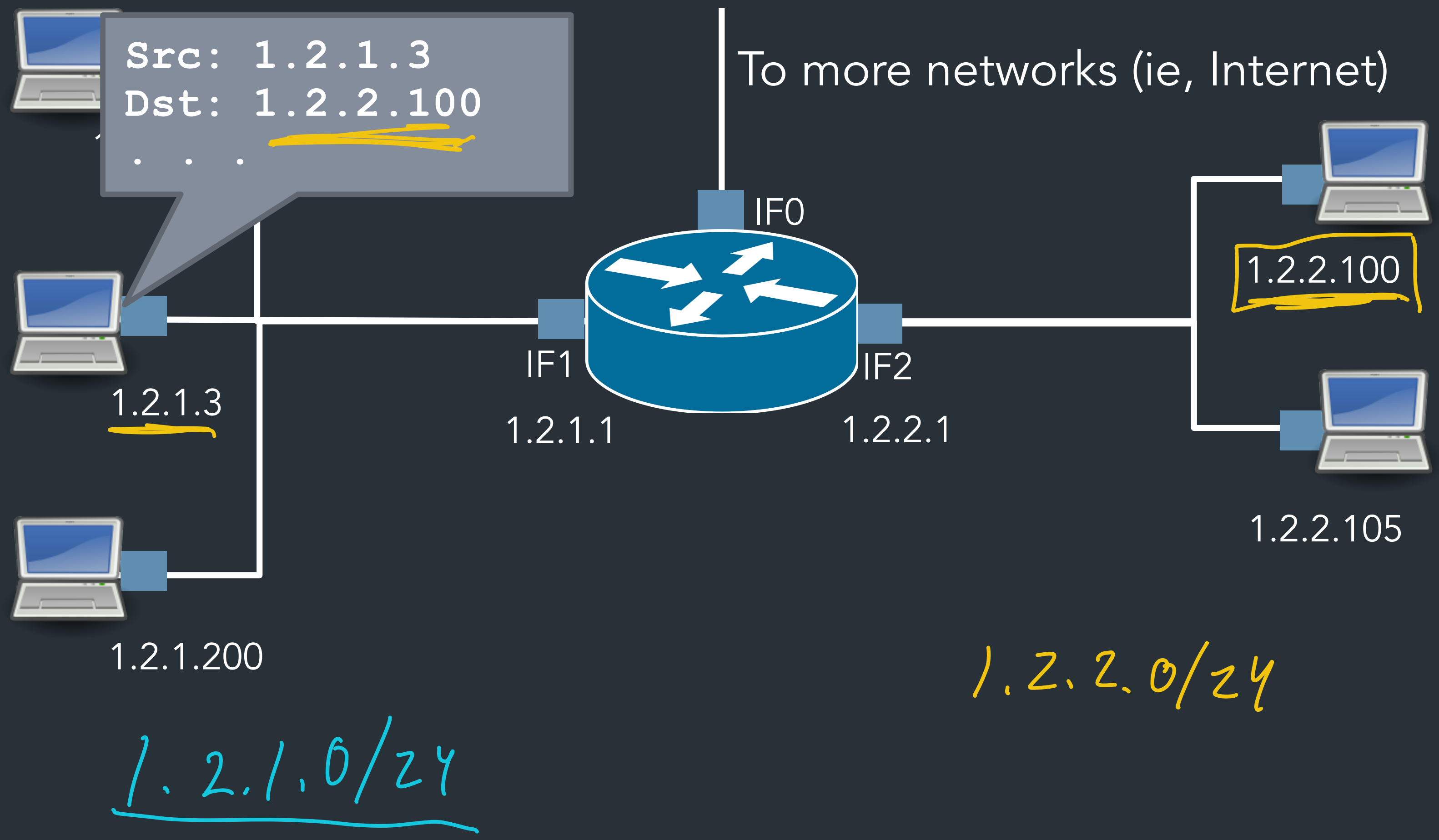
Send packets to a <u>router</u>, which <u>forwards</u> IP packets to other networks



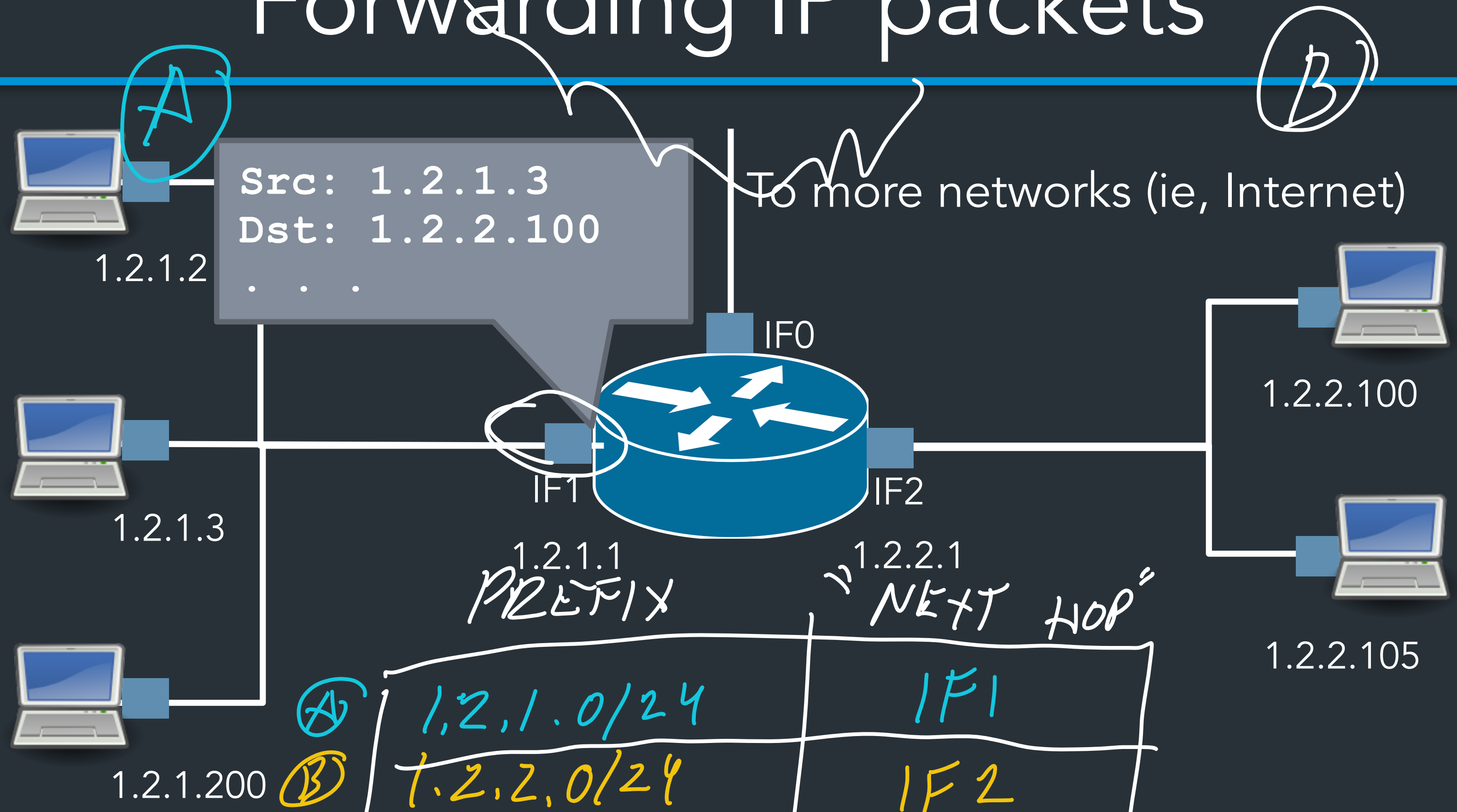
Forwarding IP packets



Forwarding IP packets

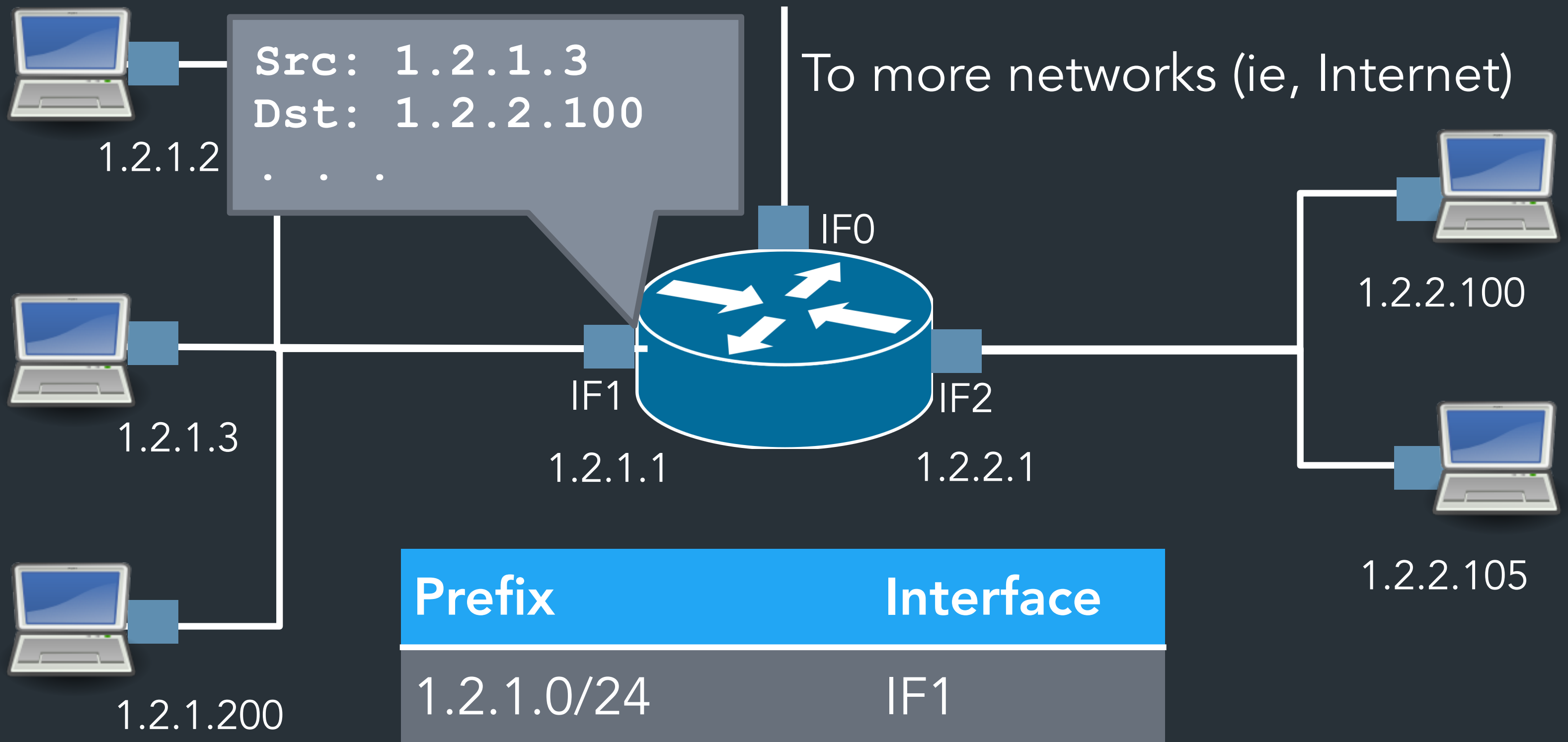


Forwarding IP packets



	PREFIX	"NEXT HOP"
A	1.2.1.0/24	IF1
B	1.2.2.0/24	IF2
	*	IF0

Forwarding IP packets



Prefix	**Interface**
1.2.1.0/24	IF1
1.2.2.0/24	IF2
<everything else>	(IF0)

A routing table

```
R6#sh ip ro
Gateway of last resort is 108.34.215.1 to network 0.0.0.0

S*     0.0.0.0/0 [1/0] via 108.34.215.1
       10.0.0.0/8 is variably subnetted, 7 subnets, 3 masks
C         10.1.0.0/24 is directly connected, wlan-ap0
L         10.1.0.2/32 is directly connected, wlan-ap0
O IA      10.1.44.1/32 [110/1001] via 10.20.30.33, 3w4d, Tunnel0
C         10.1.48.0/24 is directly connected, Loopback0
L         10.1.48.1/32 is directly connected, Loopback0
C         10.20.30.32/31 is directly connected, Tunnel0
L         10.20.30.32/32 is directly connected, Tunnel0
       108.0.0.0/8 is variably subnetted, 2 subnets, 2 masks
C         108.34.215.0/24 is directly connected, GigabitEthernet0/0
L         108.34.215.208/32 is directly connected, GigabitEthernet0/0
       172.16.0.0/16 is variably subnetted, 2 subnets, 2 masks
C         172.16.98.0/24 is directly connected, Vlan98
L         172.16.98.1/32 is directly connected, Vlan98
       172.17.0.0/16 is variably subnetted, 6 subnets, 3 masks
O IA      172.17.44.0/24 [110/1001] via 10.20.30.33, 3w4d, Tunnel0
C         172.17.48.0/24 is directly connected, Vlan20
```

A routing table

```
R6#sh ip ro
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2
       i - IS-IS, su - IS-IS summary, L1 - IS-IS level-1, L2 - IS-IS level-2
       ia - IS-IS inter area, * - candidate default, U - per-user static route
       o - ODR, P - periodic downloaded static route, H - NHRP, l - LISP
       + - replicated route, % - next hop override


Gateway of last resort is 108.34.215.1 to network 0.0.0.0


S*    0.0.0.0/0 [1/0] via 108.34.215.1
      10.0.0.0/8 is variably subnetted, 7 subnets, 3 masks
C        10.1.0.0/24 is directly connected, wlan-ap0
L        10.1.0.2/32 is directly connected, wlan-ap0
O IA     10.1.44.1/32 [110/1001] via 10.20.30.33, 3w4d, Tunnel0
C        10.1.48.0/24 is directly connected, Loopback0
L        10.1.48.1/32 is directly connected, Loopback0
```

A large table

```
rviews@route-server.ip.att.net>show route table inet.0 active-path

inet.0: 866991 destinations, 13870153 routes (866991 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both


0.0.0.0/0          *[Static/5] 5w0d 19:43:09
                    > to 12.0.1.1 via em0.0
1.0.0.0/24         *[BGP/170] 1d 10:24:47, localpref 100, from 12.122.83.238
                      AS path: 7018 3356 13335 I, validation-state: valid
                    > to 12.0.1.1 via em0.0
1.0.4.0/22         *[BGP/170] 1d 10:24:47, localpref 100, from 12.122.83.238
                      AS path: 7018 3356 4826 38803 I, validation-state: valid
                    > to 12.0.1.1 via em0.0
1.0.4.0/24         *[BGP/170] 1d 10:24:47, localpref 100, from 12.122.83.238
                      AS path: 7018 3356 4826 38803 I, validation-state: valid
                    > to 12.0.1.1 via em0.0
1.0.5.0/24         *[BGP/170] 1d 10:24:47, localpref 100, from 12.122.83.238
                      AS path: 7018 3356 4826 38803 I, validation-state: valid
                    > to 12.0.1.1 via em0.0
1.0.6.0/24         *[BGP/170] 1d 10:24:47, localpref 100, from 12.122.83.238
```