PART IX

INTERNET PROTOCOL: CLASSLESS AND SUBNET ADDRESS EXTENSIONS (CIDR)

Recall

In the original IP addressing scheme, each physical network is assigned a unique network address; each host on a network has the network address as a prefix of the host's individual address.

- Routers only examine prefix (small routing tables)

An Observation

- Division into prefix and suffix means: site can assign and use IP addresses in unusual ways provided
  - All hosts and routers at the site honor the site's scheme
  - Other sites on the Internet can treat addresses as a network prefix and a host suffix

Classful Addressing

- Three possible classes for networks
- Class C network limited to 254 hosts (cannot use all-1s or all-0s)
- Personal computers result in networks with many hosts
- Class B network allows many hosts, but insufficient class B prefixes

Question

- How can we minimize the number of assigned network prefixes (especially class B) without abandoning the 32-bit addressing scheme?

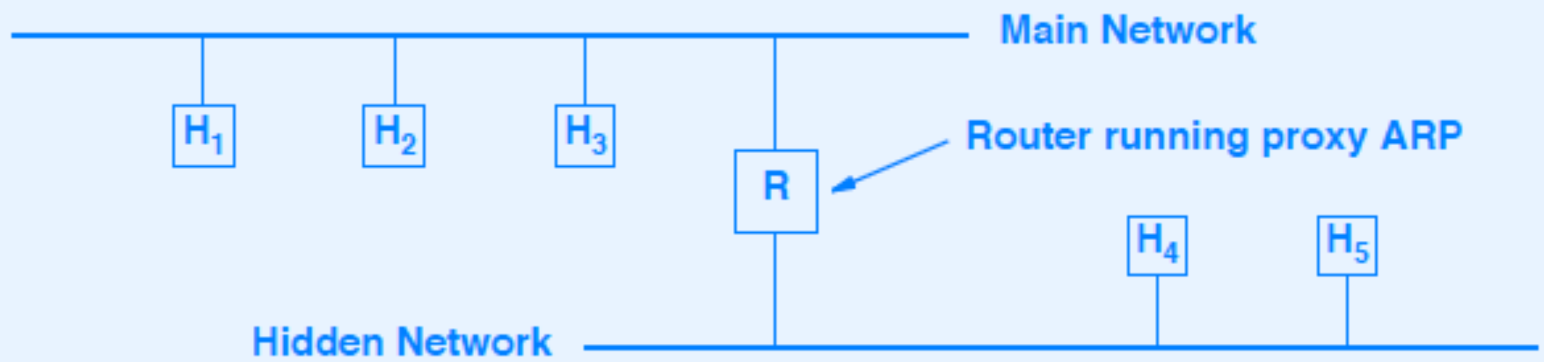
Two Answers To The Minimization Question

- Proxy ARP
- Subnet addressing

Proxy ARP

- Layer 2 solution
- Allow two physical networks to share a single IP prefix
- Arrange special system to answer ARP requests and forward datagrams between networks

Illustration Of Proxy ARP



- Hosts think they are on same network
- Known informally as *the ARP hack*

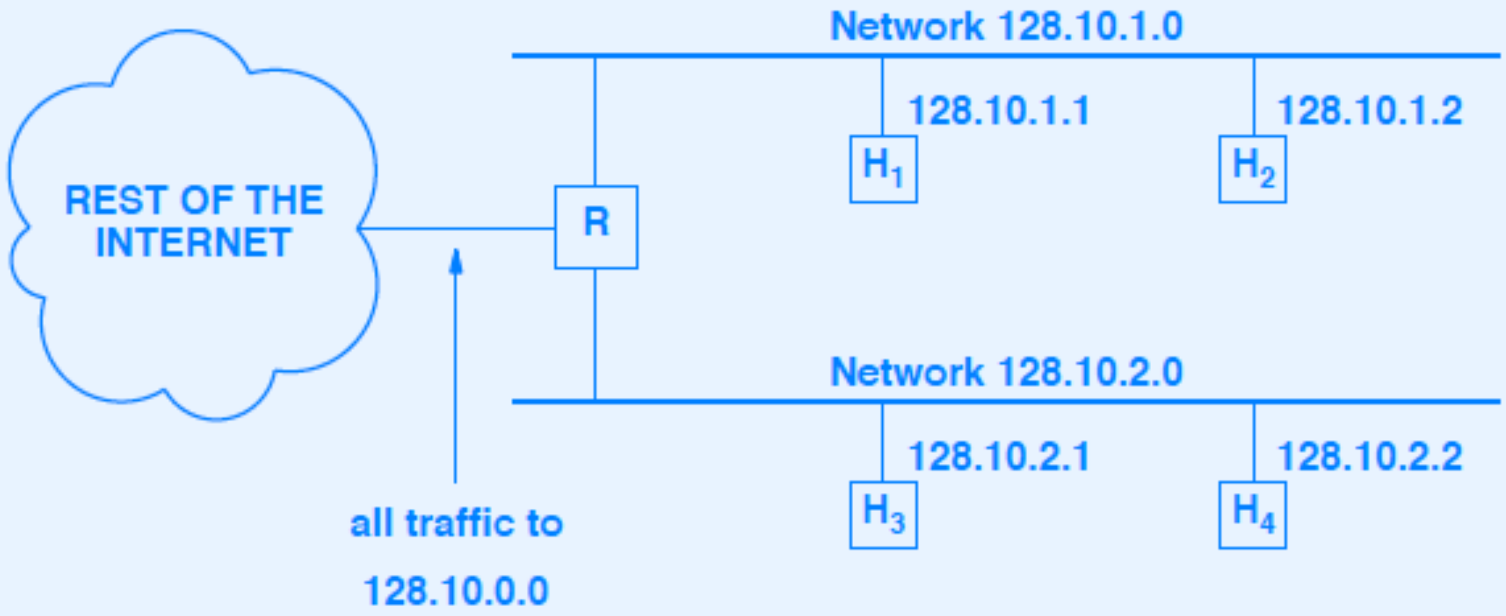
Assessment Of Proxy ARP

- Chief advantages
  - Transparent to hosts
  - No change in IP routing tables
- Chief disadvantages
  - Does not generalize to complex topology
  - Only works on networks that use ARP
  - Most proxy ARP systems require manual configuration

Subnet Addressing

- Not part of original TCP/IP address scheme
- Allows an organization to use a single network prefix for multiple physical networks
- Subdivides the host suffix into a pair of fields for physical network and host
- Interpreted only by routers and hosts at the site; treated like normal address elsewhere

Example Of Subnet Addressing

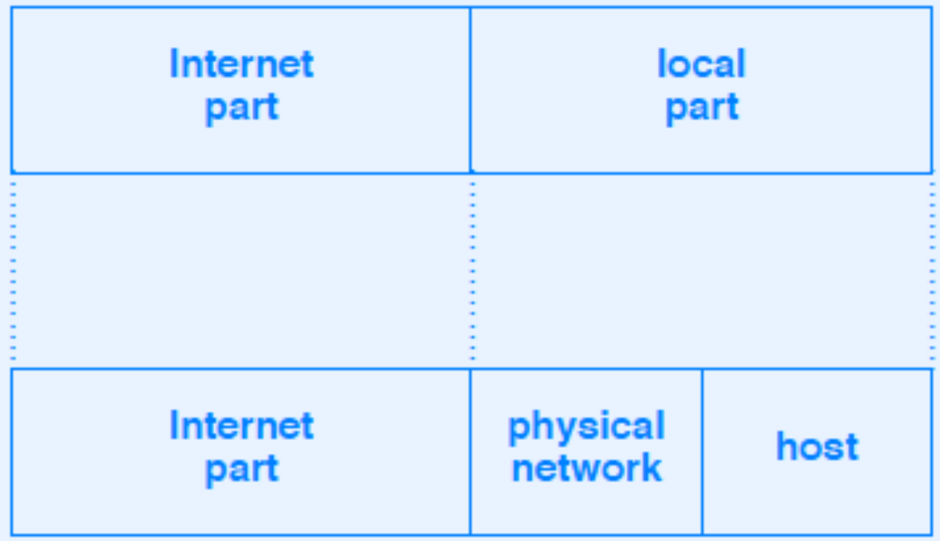


- Both physical networks share prefix 128.10
- Router R uses third octet of address to choose physical net

Interpretation Of Addresses

- Classful interpretation is two-level hierarchy
  - Physical network identified by prefix
  - Host on the net identified by suffix
- Subnetted interpretation is three-level hierarchy
  - Site identified by network prefix
  - Physical net at site identified by part of suffix
  - Host on the net identified by remainder of suffix

Example Of Address Interpretation (Subnetted Class B Address)



Note: in this case, 16-bit host portion is divided into two 8-bit fields

Choice Of Subnet Size

- How should host portion of address be divided?
- Answer depends on topology at site and number of hosts per network

Example Of Site With Hierarchical Topology

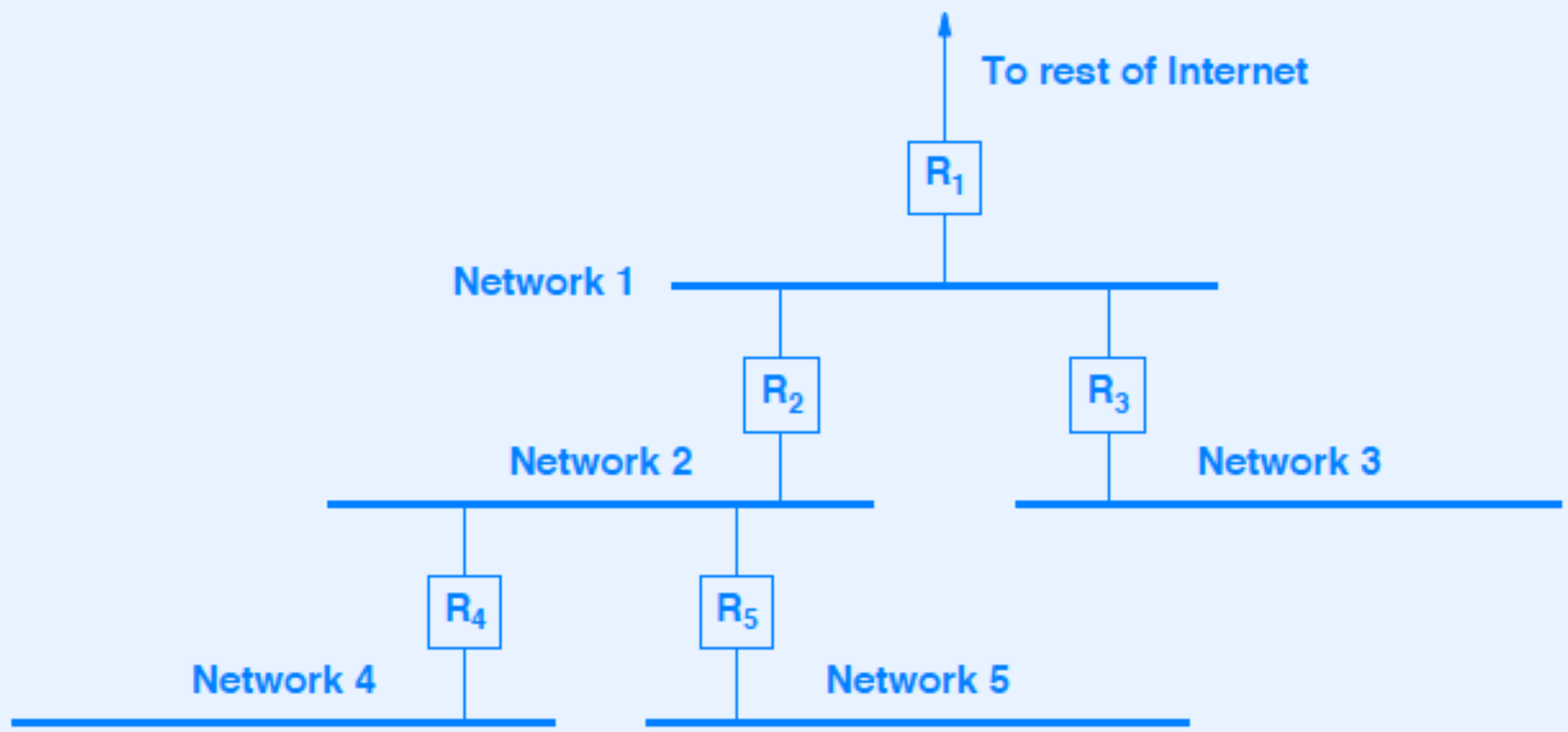
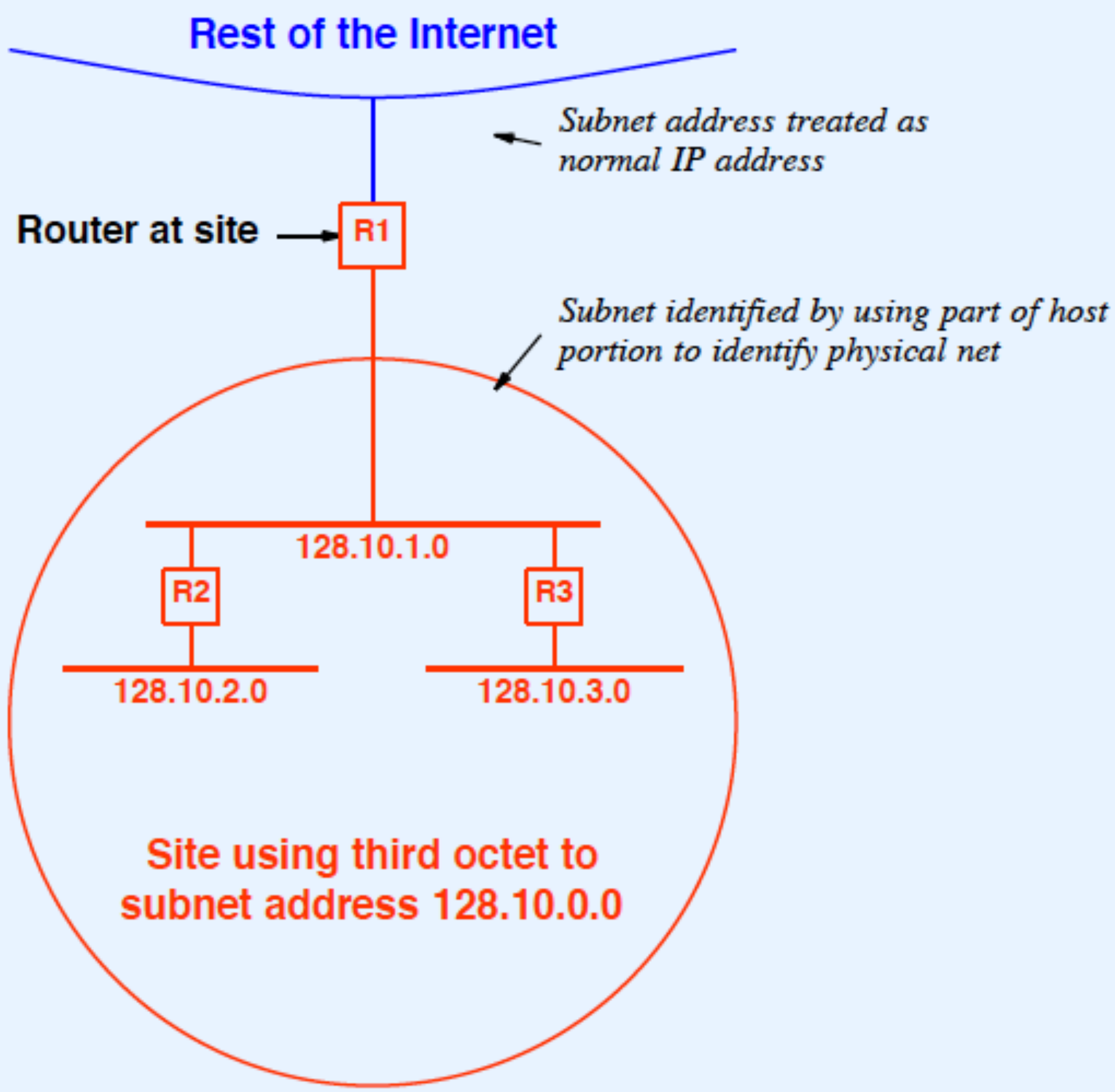


Illustration Of Subnet Addressing



Address Mask

- Each physical network is assigned 32-bit *address mask* (also called *subnet mask*)
- One bits in mask cover network prefix plus zero or more bits of suffix portion
- Logical *and* between mask and destination IP address extracts the prefix and subnet portions

Two Possible Mask Assignments

- Fixed-length subnet masks
- Variable-length subnet masks

Fixed-length Subnet Masks

- Organization uses same mask on all networks
- Advantages
  - Uniformity
  - Ease of debugging / maintenance
- Disadvantages
  - Number of nets fixed for entire organization
  - Size of physical nets fixed for entire organization

Possible Fixed-Length Subnets For Sixteen Bit Host Address

Bits in mask	# subnets	# hosts/subnet
16	1	65534
18	2	16382
19	6	8190
20	14	4094
21	30	2046
22	62	1022
23	126	510
24	254	254
25	510	126
26	1022	62
27	2046	30
28	4094	14
29	8190	6
30	16382	2

- All-*0*s and all-*1*s values must be omitted
- Organization chooses one line in table

Variable-Length Subnet Masks (VLSM)

- Administrator chooses size for each physical network
- Mask assigned on per-network basis
- Advantages
  - Flexibility to mix large and small nets
  - More complete use of address space
- Disadvantages
  - Difficult to assign / administer
  - Potential address ambiguity
  - More routes

Use Of Address Space
(Start With 16 Bits Of Host Suffix)

- One possible VLSM assignment (92.9% of addresses used)
  - 11 networks of 2046 hosts each
  - 24 networks of 254 hosts each
  - 256 networks of 126 hosts each
- Another possible VLSM assignment (93.1% of addresses used)
  - 9 networks of 2046 hosts each
  - 2 networks of 1022 hosts each
  - 40 networks of 510 hosts each
  - 160 networks of 126 hosts each

Subnet Details

- Two interesting facts
  - *Can* assign all-0's or all-1's subnet
  - *Can* assign noncontiguous subnet mask bits
- In practice
  - Should avoid both
- Discussion question: why does the subnet standard allow the all-1's and all-0's subnet numbers?

VLSM Example

- Use low-order sixteen bits of 128.10.0.0
- Create seven subnets
- Subnet 1
  - Up to 254 hosts
  - Subnet mask is 24 bits
- Subnets 2 through 7
  - Up to 62 hosts each
  - Subnet mask is 26 bits

Example VLSM Prefixes

- Subnet 1 (up to 254 hosts)

  ```
  mask:   11111111 11111111 11111111 00000000
  prefix: 10000000 00001010 00000001 00000000
  ```

- Subnet 2 (up to 62 hosts)

  ```
  mask:   11111111 11111111 11111111 11000000
  prefix: 10000000 00001010 00000000 10000000
  ```

- Subnet 3 (up to 62 hosts)

  ```
  mask:   11111111 11111111 11111111 11000000
  prefix: 10000000 00001010 00000000 11000000
  ```

Example VLSM Prefixes
(continued)

- Subnet 4 (up to 62 hosts)

```
mask:   11111111 11111111 11111111 11000000
prefix: 10000000 00001010 00000001 00000000
```

- Subnet 5 (up to 62 hosts)

```
mask:   11111111 11111111 11111111 11000000
prefix: 10000000 00001010 00000001 01000000
```

- Subnet 6 (up to 62 hosts)

```
mask:   11111111 11111111 11111111 11000000
prefix: 10000000 00001010 00000001 10000000
```

Example VLSM Prefixes
(continued)

- Subnet 7 (up to 62 hosts)

```
mask:   11111111 11111111 11111111 11000000

prefix: 10000000 00001010 00000001 11000000
```

Address Ambiguity

- Address of host 63 on subnet 1 is

```
mask:   11111111 11111111 11111111 00000000
prefix: 10000000 00001010 00000001 00000000
host:   10000000 00001010 00000001 00111111
```

- Directed broadcast address on subnet 4 is

```
mask:   11111111 11111111 11111111 11000000
prefix: 10000000 00001010 00000001 00000000
bcast:  10000000 00001010 00000001 00111111
```

Address Ambiguity

- Address of host 63 on subnet 1 is

```
mask:   11111111 11111111 11111111 00000000
prefix: 10000000 00001010 00000001 00000000
host:   10000000 00001010 00000001 00111111
```

- Directed broadcast address on subnet 4 is

```
mask:   11111111 11111111 11111111 11000000
prefix: 10000000 00001010 00000001 00000000
bcast:  10000000 00001010 00000001 00111111
```

- Same value!

More Address Ambiguity

- Directed broadcast address on subnet 1 is

```
mask:      11111111 11111111 11111111 00000000
prefix:    10000000 00001010 00000001 00000000
broadcast: 10000000 00001010 00000001 11111111
```

- Directed broadcast address on subnet 7 is

```
mask:     11111111 11111111 11111111 11000000
prefix:   10000000 00001010 00000001 11000000
broadcast:10000000 00001010 00000001 11111111
```

More Address Ambiguity

- Directed broadcast address on subnet 1 is

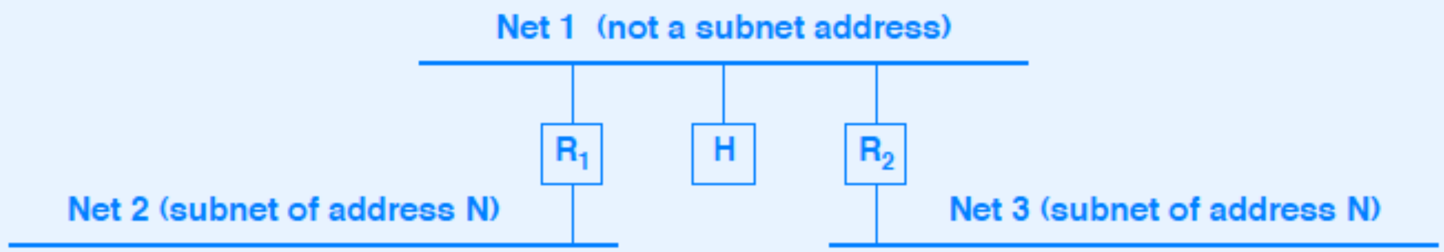
```
mask:      11111111 11111111 11111111 00000000
prefix:    10000000 00001010 00000001 00000000
broadcast: 10000000 00001010 00000001 11111111
```

- Directed broadcast address on subnet 7 is

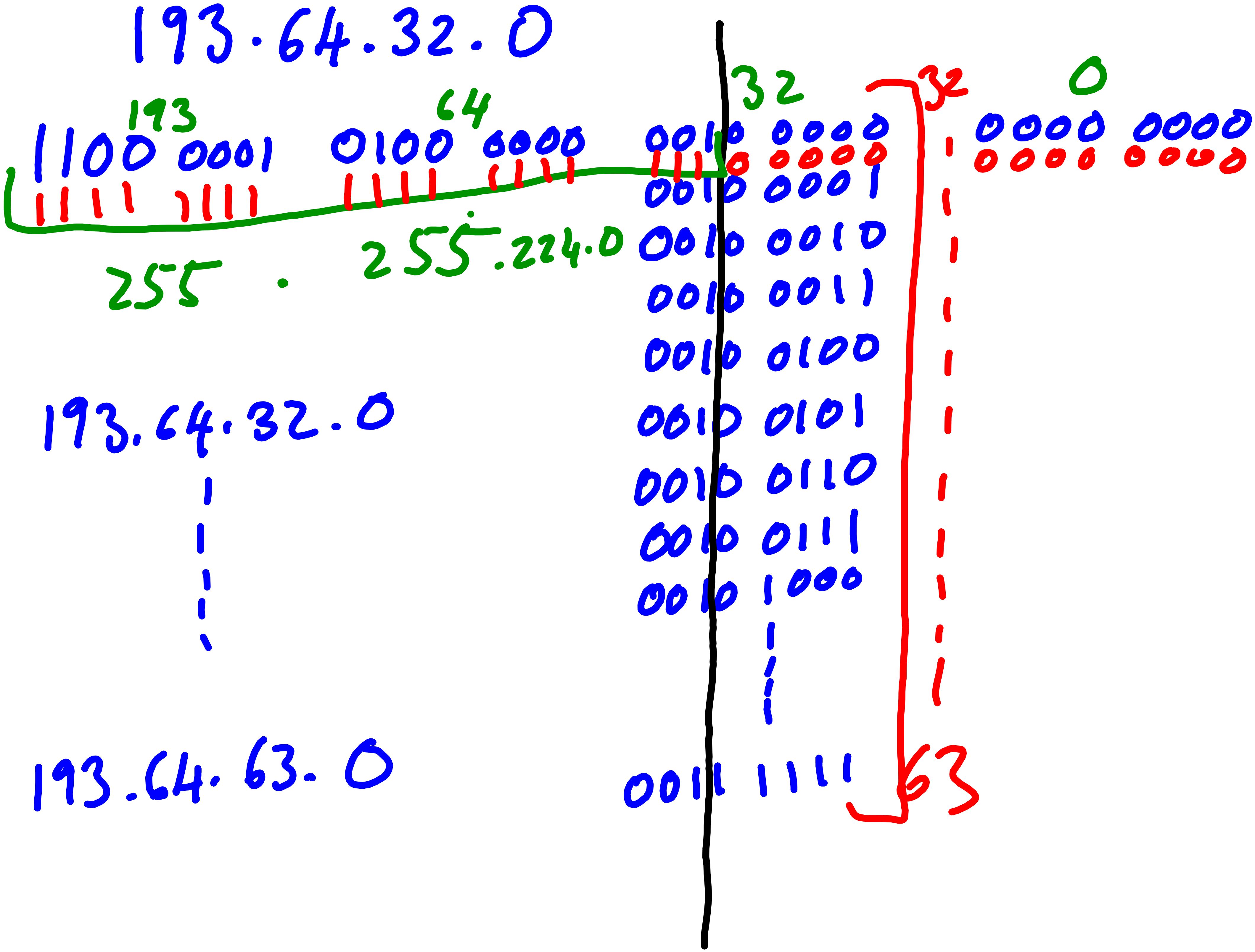
```
mask:     11111111 11111111 11111111 11000000
prefix:   10000000 00001010 00000001 11000000
broadcast:10000000 00001010 00000001 11111111
```

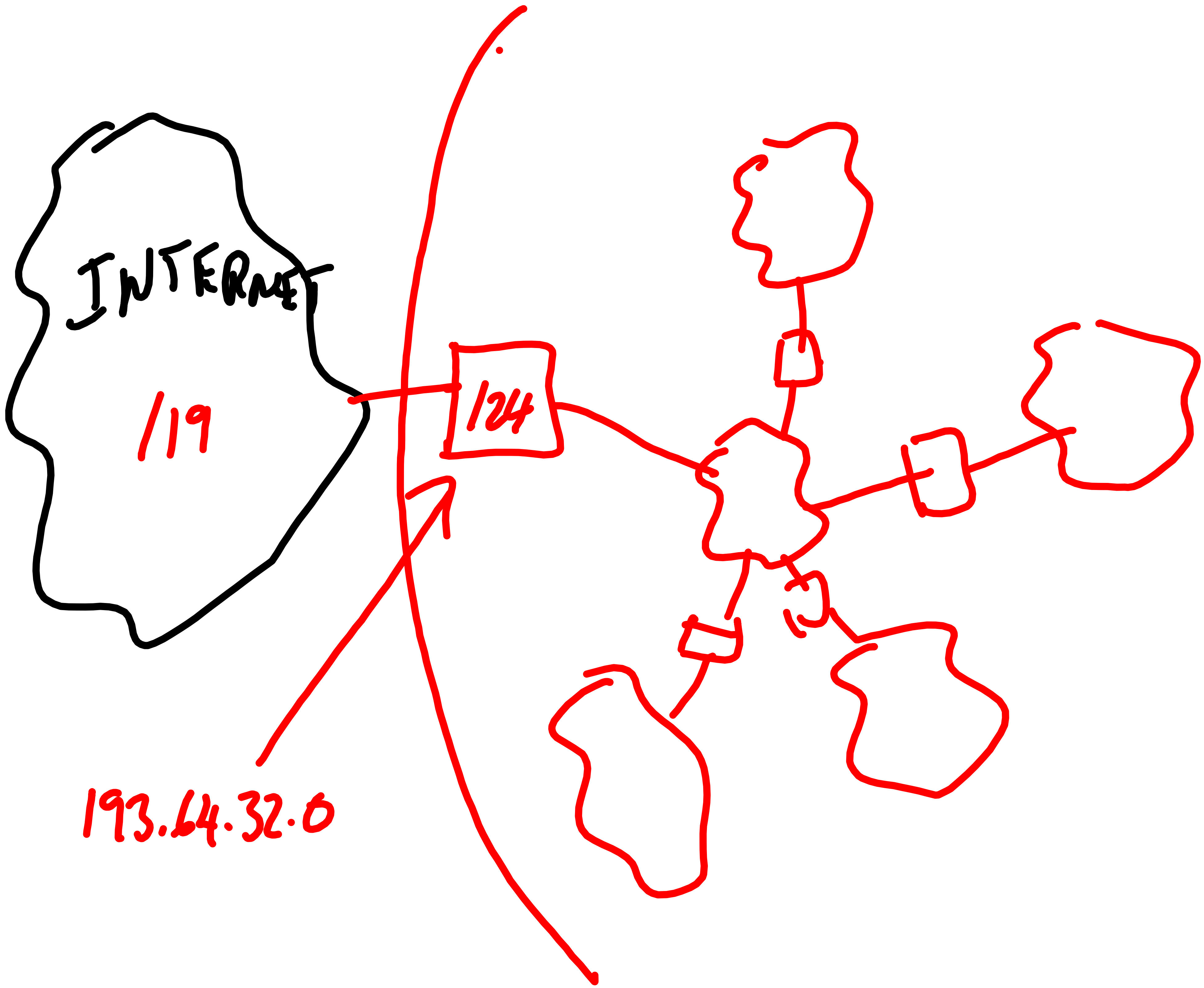
- Same value!

Example Of Illegal Subnet Assignment



- Host cannot route among subnets
- Rule: subnets *must* be contiguous!


193.64.32.0
193
64
1100 0001
0100 0000
1111 1111
1111 1111
255 . 255.224.0
32
0010 0000
111 0 0000
32
0
0000 0000
0000 0000
0010 0001
0010 0010
0010 0011
0010 0100
0010 0101
0010 0110
0010 0111
0010 1000
193.64.32.0
193.64.63.0
0011 1111
63


INTERNET
/19
/24
193.64.32.0

Variety Of Routes

- Forwarding must accommodate
  - Network-specific routes
  - Subnet-specific routes
  - Host-specific routes
  - Default route
  - Limited broadcast
  - Directed broadcast to network
  - Directed broadcast to specific subnet
- Single algorithm with address masks can accommodate all the above

Use Of Address Masks

- Each entry in routing table also has address mask
- All-1s mask used for host-specific routes
- Network mask used for network-specific routes
- Subnet mask used for subnet-specific routes
- All-0s mask used for default route

Unified Forwarding Algorithm

Algorithm:

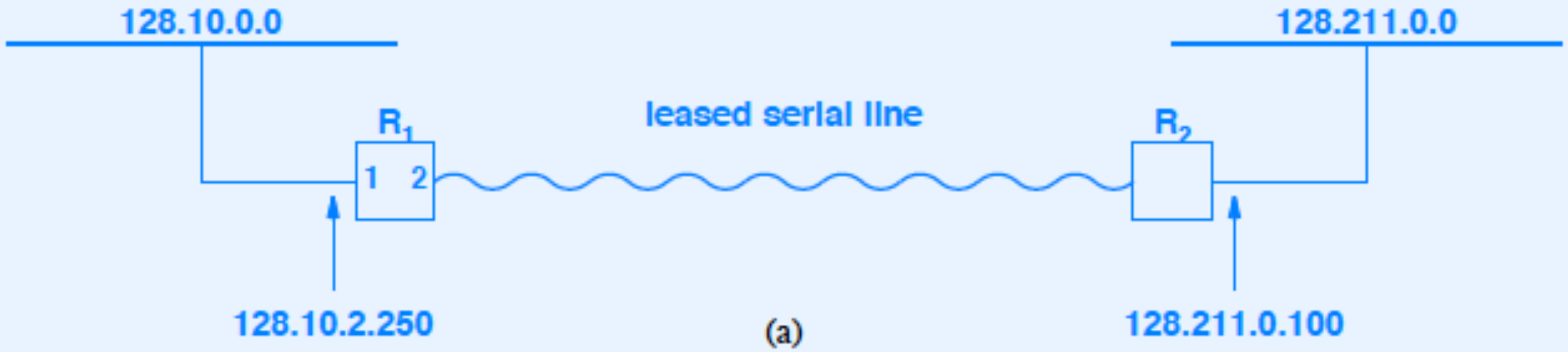
```
    Forward_IP_Datagram (datagram, routing_table)

Extract destination IP address, ID, from datagram;
If prefix of ID matches address of any directly connected
   network send datagram to destination over that network
   (This involves resolving ID to a physical address,
   encapsulating the datagram, and sending the frame.)
else
   for each entry in routing table do
      Let N be the bitwise-and of ID and the subnet mask
      If N equals the network address field of the entry then
         forward the datagram to the specified next hop address
   endforloop
If no matches were found, declare a forwarding error;
```

Special Case: Unnumbered Serial Network

- Only two endpoints
- Not necessary to assign (waste) network prefix
- Trick: use remote IP address as next hop

Example Unnumbered Serial Network



(a)

TO REACH HOSTS ON NETWORK	ROUTE TO THIS ADDRESS	USING THIS INTERFACE
128.10.0.0	DELIVER DIRECT	1
default	128.211.0.100	2

Classless Inter-Domain Routing (CIDR)

- Problem
  - Continued exponential Internet growth
  - Subnetting insufficient
  - Limited IP addresses (esp. Class B)
- Dire prediction made in 1993:

  We will exhaust the address space ‘‘in a few years’’.

Note: address space is *not* near exhaustion

CIDR Addressing

- Solution to problem
  - Temporary fix until next generation of IP
  - Backward compatible with classful addressing
  - Extend variable-length subnet technology to prefixes
- CIDR was predicted to work ‘‘for a few years’’
  - Extremely successful!
  - Will work for at least 25 years!

One Motivation For CIDR: Class C

- Fewer than seventeen thousand Class B numbers (total)
- More than two million Class C network numbers
- No one wants Class C (too small)
- CIDR allows
  - Merging 256 Class C numbers into a single prefix that is equivalent to Class B
  - Splitting a Class B along power of two boundaries

CIDR Notation

- Addresses written *NUMBER* / *m*
  - *NUMBER* is IP prefix
  - *m* is ‘‘address mask’’ length
- Example

214.5.48.0/20

– Prefix occupies 20 bits

– Suffix occupies 12 bits

- Mask values must be converted to dotted decimal when configuring a router (and binary internally)

Route Proliferation

- If classful forwarding used, CIDR addresses result in more routes
- Example:
  - Single CIDR prefix spans 256 Class C network numbers (*supernetting*)
  - Classful routing table requires 256 separate entries

Route Condensation

- Solution: change forwarding as well as addressing
- Store address mask with each route
- Send pair of (address, mask) whenever exchanging routing information
- Known as a *CIDR block*

Example Of A CIDR Block

	Dotted Decimal	32-bit Binary Equivalent
lowest	128.211.168.0	10000000 11010011 10101000 00000000
highest	128.211.175.255	10000000 11010011 10101111 11111111

Dotted Decimal Equivalents

CIDR Notation	Dotted Decimal	CIDR Notation	Dotted Decimal
/1	128.0.0.0	/17	255.255.128.0
/2	192.0.0.0	/18	255.255.192.0
/3	224.0.0.0	/19	255.255.224.0
/4	240.0.0.0	/20	255.255.240.0
/5	248.0.0.0	/21	255.255.248.0
/6	252.0.0.0	/22	255.255.252.0
/7	254.0.0.0	/23	255.255.254.0
/8	255.0.0.0	/24	255.255.255.0
/9	255.128.0.0	/25	255.255.255.128
/10	255.192.0.0	/26	255.255.255.192
/11	255.224.0.0	/27	255.255.255.224
/12	255.240.0.0	/28	255.255.255.240
/13	255.248.0.0	/29	255.255.255.248
/14	255.252.0.0	/30	255.255.255.252
/15	255.254.0.0	/31	255.255.255.254
/16	255.255.0.0	/32	255.255.255.255

Example Of /30 CIDR Block

	Dotted Decimal	32-bit Binary Equivalent
lowest	128.211.176.212	10000000 11010011 10110000 11010100
highest	128.211.176.215	10000000 11010011 10110000 11010111

- Useful when customer of ISP has very small network

Implementation Of CIDR Route Lookup

- Each entry in routing table has address plus mask
- Search is organized from most-specific to least-specific (i.e., entry with longest mask is tested first)
- Known as *longest-prefix lookup* or *longest-prefix search*

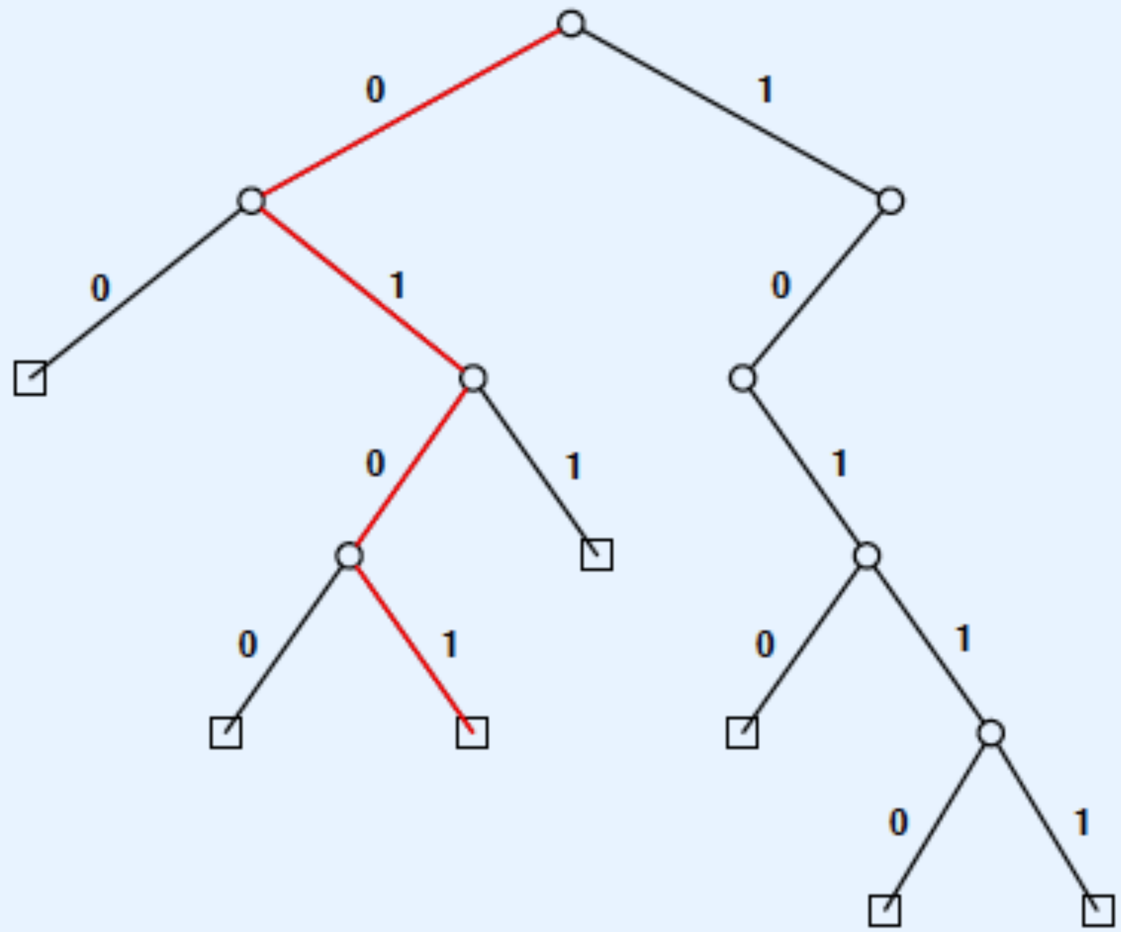
Implementing Longest-Prefix Matching

- Cannot easily use hashing
- Data structure of choice is *binary trie*
- Identifies unique prefix needed to match route

Example Of Unique Prefixes

32-Bit Address	Unique Prefix
00110101 00000000 00000000 00000000	00
01000110 00000000 00000000 00000000	0100
01010110 00000000 00000000 00000000	0101
01100001 00000000 00000000 00000000	011
10101010 11110000 00000000 00000000	1010
10110000 00000010 00000000 00000000	10110
10111011 00001010 00000000 00000000	10111

Example Binary Trie For The Seven Prefixes



- Path for 0101 is shown in red

Modifications And Extensions

- Several variations of trie data structures exist
  - PATRICIA trees
  - Level-Compressed tries (LC-tries)
- Motivation
  - Handle longest-prefix match
  - Skip levels that do not distinguish among routes

Nonroutable Addresses

- CIDR blocks reserved for use within a site
- Must never appear on the Internet
- ISPs do not maintain routes
- Also called *private addresses*

Prefix	Lowest Address	Highest Address
10 / 8	10.0.0.0	10.255.255.255
172.16 / 12	172.16.0.0	172.31.255.255
192.168 / 16	192.168.0.0	192.168.255.255
169.254 / 16	169.254.0.0	169.254.255.255

Summary

- Original IP addressing scheme was classful
- Two extensions added
  - Subnet addressing
  - CIDR addressing
- Subnetting used only within a site
- CIDR used throughout the Internet
- Both use 32-bit address mask
  - CIDR mask identifies division between network prefix and host suffix
  - Subnet mask identifies boundary between subnet and individual host

Summary
(continued)

- Single unified forwarding algorithm handles routes that are
  - Network-specific
  - Subnet-specific
  - Host-specific
  - Limited broadcast
  - Directed broadcast to network
  - Directed broadcast to subnet
  - Default
- Longest-prefix match required
  - Typical implementation: binary trie

128.10.211.78 — IP

255.255.240.0 — SNM

NET #				SUBNET	HOST #		
1000	0000	0000	1010	1101	0011	0100	1110
1111	1111	1111	1111	1111	0000	0000	0000

32768	16384	8192	4096	2048	1024	512	256	128	64	32	16	8	4	2	1
1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0

8	4	2	1
1	1	0	1

SUBNET # 13

32768
10

NET # 32,778

2048	1024	512	256	128	64	32	16	8	4	2	1
0	0	1	1	0	1	0	0	1	1	1	0

1 1
512
256
64
14

HOST # 846