Computer Networks II
IPv4 addressing

Luca Beccheti
Luca.Becchetti@dis.uniroma1.it

A.A. 2009/2010

IP addressing (v4)

- Basic idea: every network interface assigned unique 32 bit address
  - Host/router may possess multiple interfaces
  - Some addresses (belonging to specific blocks) may be assigned multiple times --> VPN, NAT (see further)
- Interface --> network card
- Address format
  - Classful --> original proposal
  - Classless --> allows better use of address space

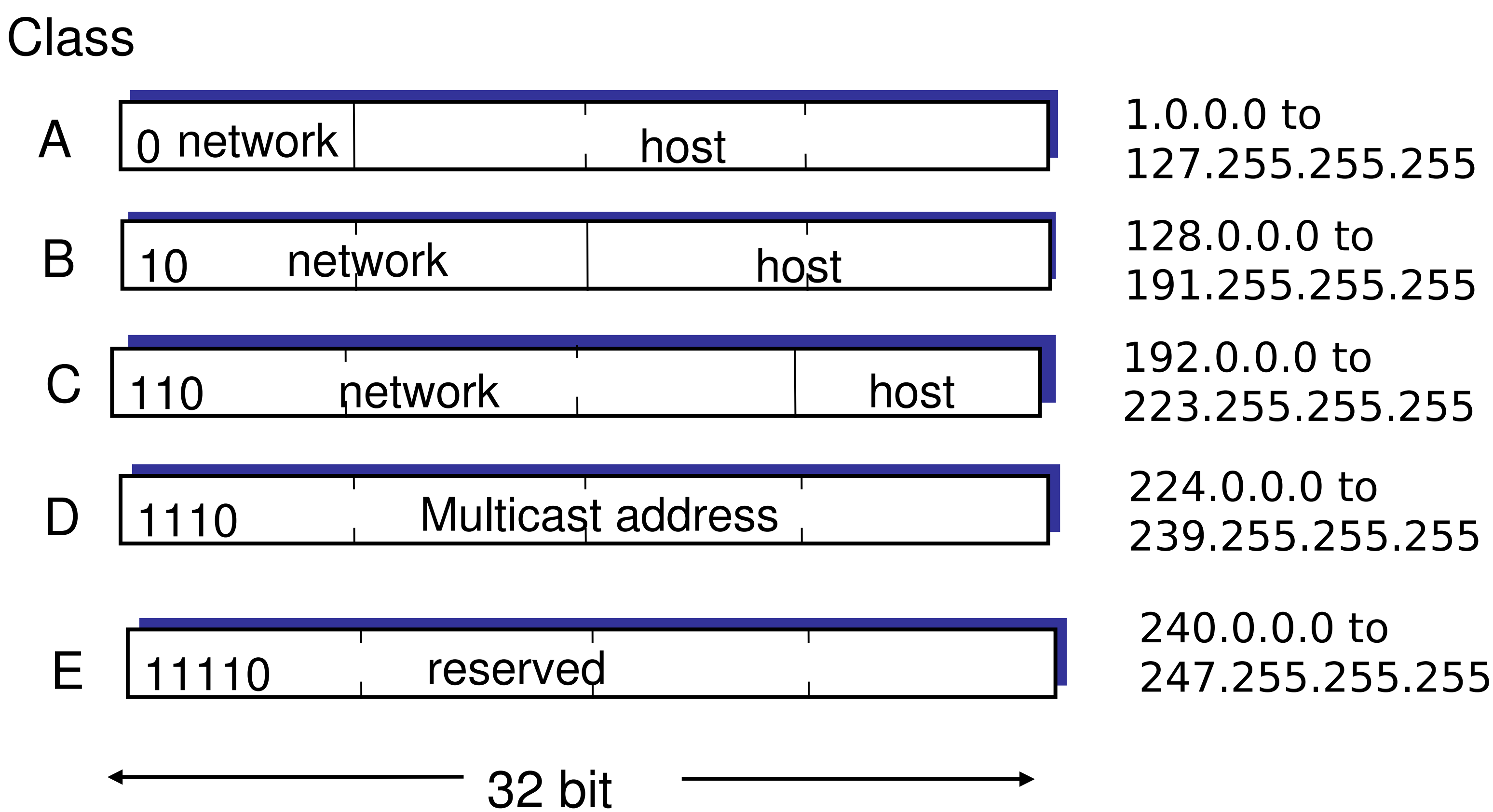
IP addressing (v4)/cont.

- IP address:
  - originally (1981, RFC 1166) two components
    - Net_Id: subnet identifier (prefix)
    - Host_Id: host identifier within subnet (suffix)
- IP_Address = Net_Id . Host_Id
- Boundary between Net_Id and Host_Id not fixed

IP addresses - classful

Classe	Bit iniziali	Net_Id	Host_Id	"Reti" disponibili	"Host" disponibili
A	0	7 bit	24 bit	128	16.777.216
B	10	14 bit	16 bit	16384	65.536
C	110	21 bit	8 bit	2.097.152	256
D	1110	Indirizzo multicast: 28 bit Indirizzi possibili: 268.435.456			
E	11110	Riservata per usi futuri: 27 bit Indirizzi possibili: 134.217.728			

IP addresses - classful/cont.



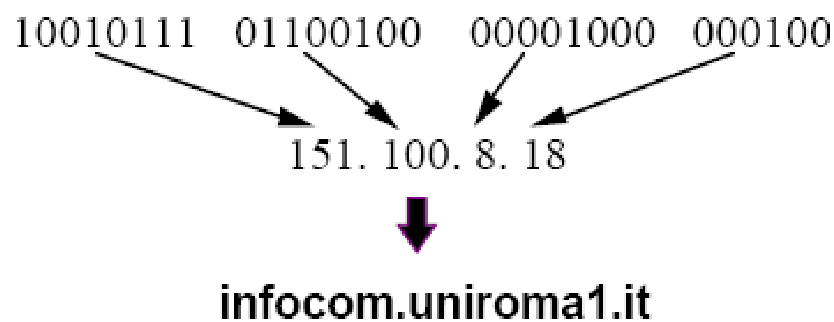
Addressing scheme

- Numeric, “dotted” and “mnemonic” notations

Numeric notation

Dotted notation

Mnemonic notation



- A specific protocol (DNS) provides numeric to mnemonic translation and viceversa

Addressing scheme /cont.

- If host moves to a different network its IP address must generally be changed
  - Mobility: Mobile IP protocol
- Conventions

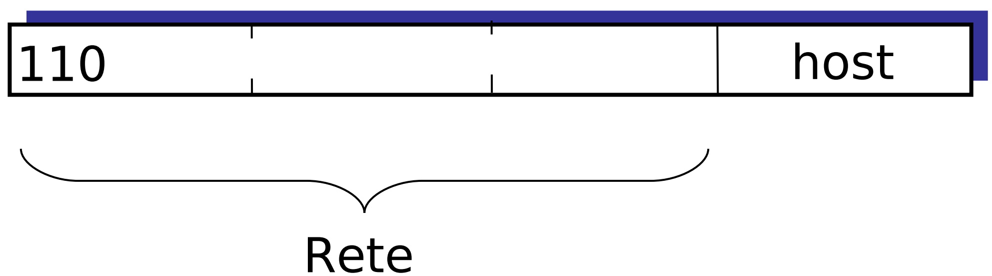
Rete locale	Tutti "0"	
Host nella rete locale	Tutti "0"	Host_Id
Broadcast sulla rete locale	Tutti "1"	
Broadcast sulla rete Net_Id	Net_Id	Tutti "1"

Broadcast - conventions

- Broadcast to local networks
  - All bits set to 1 (Net_Id + Host_Id)
- Broadcast towards a different network
  - Valid network prefix
  - Host suffix bits all set to 1
  - Non standard: Host suffix bits all set to 0
- Broadcast IP: Hw broadcast on destination network (if possible)

Classful addressing - pros

- Boundary between network and host adrresses immediately identified
  - E.g.: highest 3 bits are 110 --> class C address



- No further info needed

Classful addressing - cons

- Problem: potential address waste
- Example:
  - 2000 hosts network
  - Class B address block required
  - -> roughly 63000 unused addresses
- Solution (RFC 1519): Classless InterDomain Routing
  - Further in this lecture

IP address assignment

- All interfaces on the same network share same network prefix
  - Prefixes assigned by central authority
  - Required from ISP or other instution
- Every host on the same network has distinct suffix
  - Locally assigned
  - Local network administrator responsible for uniqueness

Some open issues

- Multicast
  - (1 to many) transmissions
    - Obvious trivial and costly solution (which?)
  - Many proposals but no unique standard
- Multi-homed hosts
  - E.g.: mobile users
  - Discussed further in this course

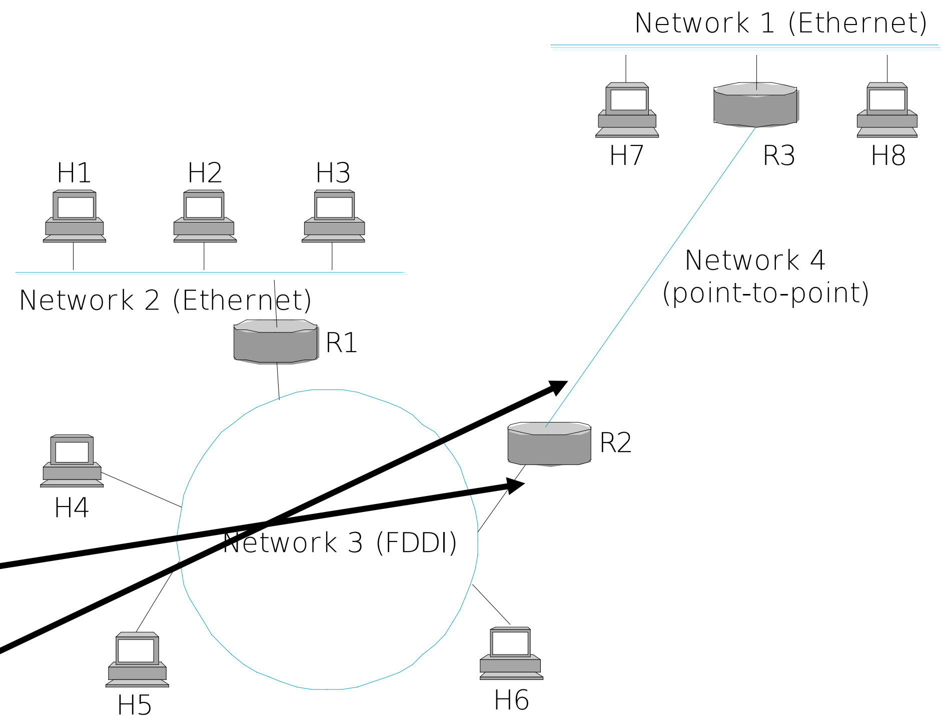
Datagram forwarding

- **Goal:** choose correct exit port at router based on i) IP destination address and ii) local Routing Table
- Routing is different: used to *build* Routing Tables
- Basic structure of RT line:
  - (DestinationNetwork, Cost, NextHop)
  - NextHop is address of next interface on (ideally shortest) path towards destination network
- Default NextHop (Gateway) always present

Datagram forwarding/cont.

- RT at router R2

DestNet	NextHop
1	R3
2	R1
3	Interface1
4	Interface0



Cost = hop number (in this case)

Forwarding algorithm

if (DestNet == ThisNetNum on some interface) /* let it be x */

<forward packet on interface x>

else if

(DestNet is in RT)

<forward packet to NextHop>

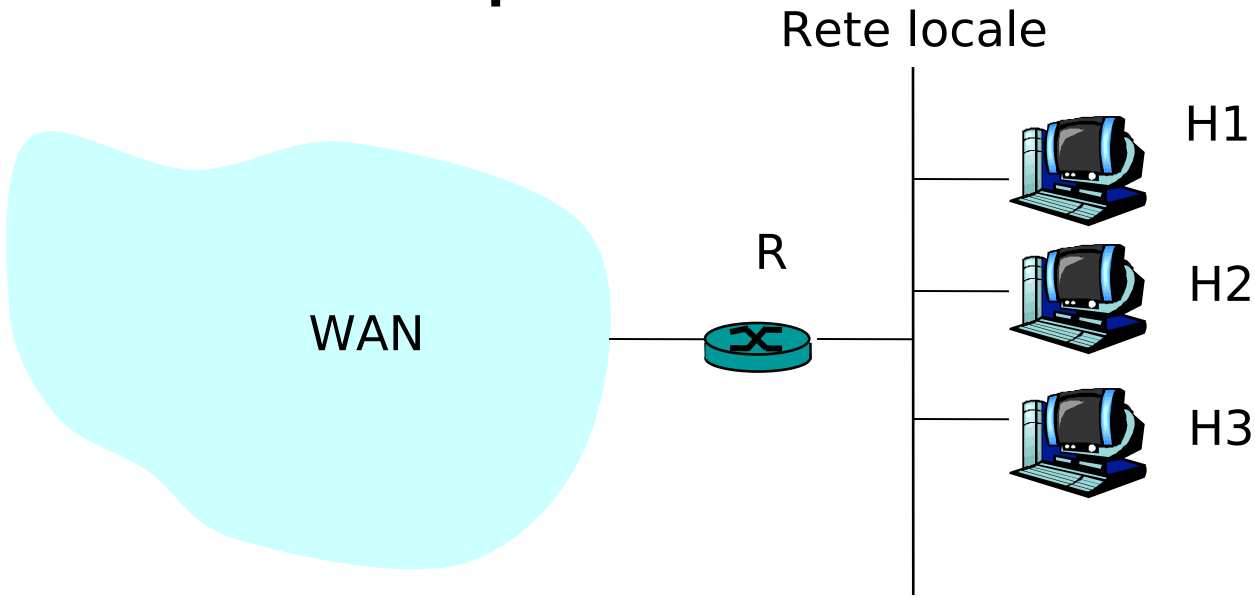
else

<send packet to default router>

Extensions

- <u>Main shortcoming</u>: great address waste
  - Especially for class B blocks
- Solutions
  - Transparent routers and proxy ARP (obsolete)
  - Subnet addressing
  - <u>CIDR - Classless Inter-Domain Routing</u>

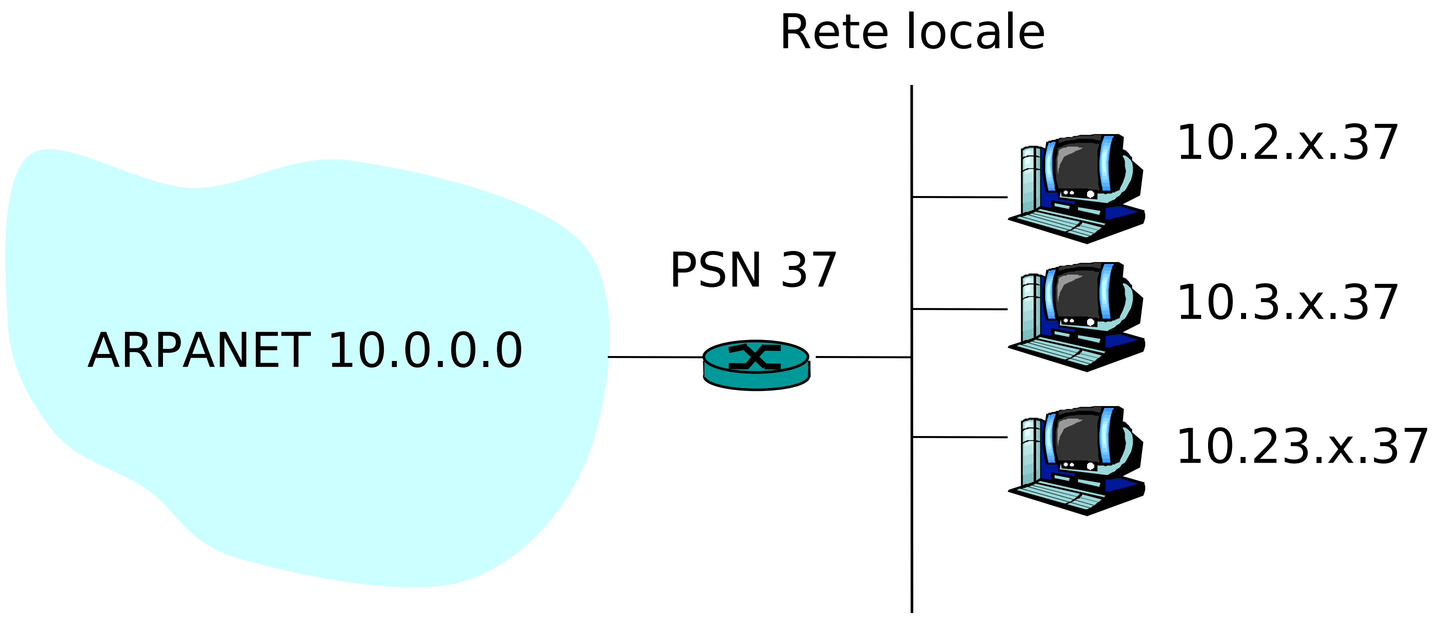
Router trasparenti



- H1, H2 E H3 “credono” di essere connessi direttamente alla WAN
- La rete locale non ha un proprio prefisso IP

- R demultipla i datagrammi da e per la rete locale
- R suddivide gli indirizzi IP in porzioni che interpreta separatamente

Esempio: ARPANET



- Indirizzo di rete in forma 10.h.x.p
- h-->host, p--> PSN, x non interpretato
- <u>Nota</u>: 10.2.5.37 e 10.2.10.37 --> stesso host

Vantaggi/svantaggi

- Vantaggi
  - Necessari meno indirizzi di rete
  - Possibile bilanciamento del carico
- Svantaggi
  - Non funziona con reti di classe C
  - I router trasparenti possono non offrire tutti i servizi standard (ICMP, SNMP….)

Indirizzi LAN e IP (cont.)

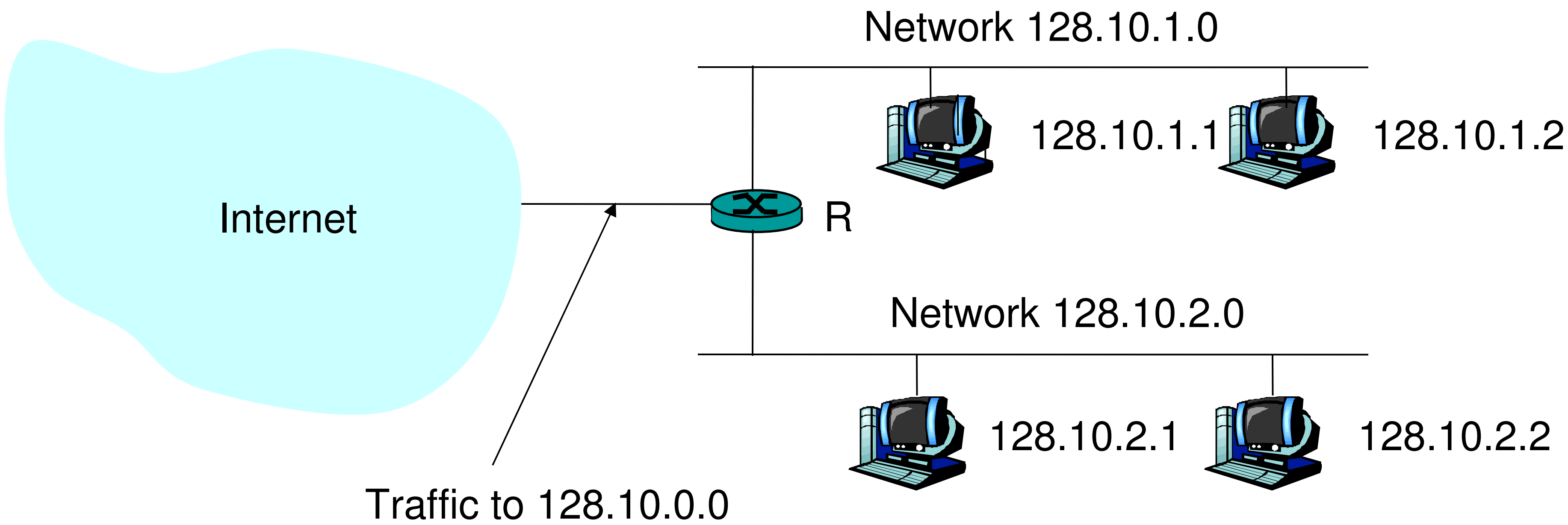
- Gli indirizzi MAC sono amministrati dalla IEEE (Institute of Electrical and Electronics Engineers)
- Ogni costruttore acquista una porzione dello spazio di indirizzamento (per assicurare unicità)
- Analogia:
  - (a) indirizzo MAC -> codice fiscale
  - (b) indirizzo IP -> indirizzo postale
- Indirizzamento MAC piatto => portabilità
- Indirizzo IP gerarchico non portabile
- Indirizzo Broadcast LAN: 1111………….1111
- D.: perché non si usano soltanto gli indirizzi IP?

Subnet addressing

- Departs from 2-level hierarchy of first Internet
  - In 1984 third hierarchical level added: Subnet
  - Standard technique
  - Often used in class B networks
  - Idea
    - Router (R) responsible for to/from exterior
    - R knows internal network organization into subnets
    - Rest of the world "sees" unique network, accessible through R
    - Some bits of Host_Id suffix used to encode Subnet_Id (subnet identifier)

Network_Id	Subnet_Id	Host_Id

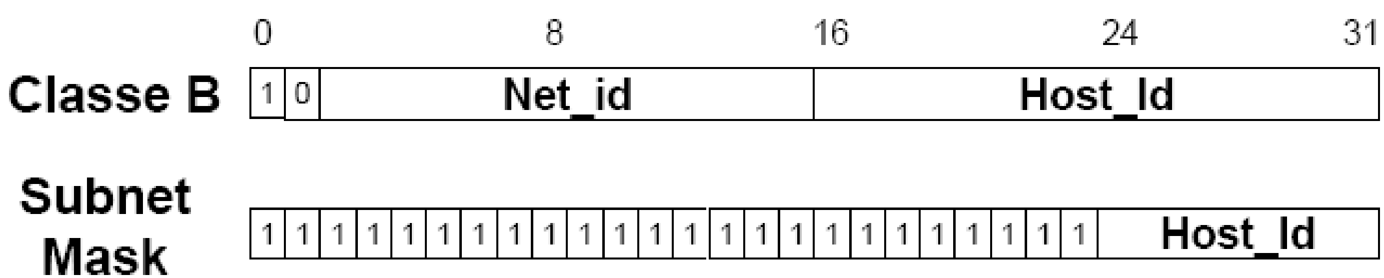
Example



- Networks 128.10.1.0 and 128.10.2.0 *not visible outside*
- R interprets destination address of datagrams coming from the Internet

Subnetting

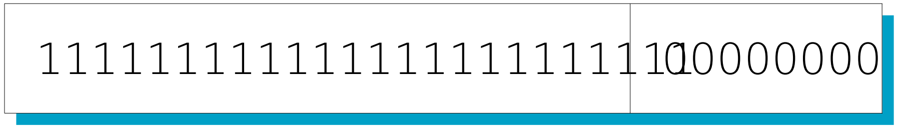
- Subnet.Id field identified by “Subnet Mask”
- Subnet Mask: 32 bit word in which
  - Bits set to “1” identify positions of Net_Id and Subnet_Id bits
  - Bits set to “0” identify positions of Host_Id bits
- Subnet_Id only meaningful at router connected to subnets (e.g., R in former slide)



Subnetting

- Adding another level to address/routing hierarchy: *subnet*
- *Subnet masks* define variable partition of host part
- Subnets visible only within site

Class B address

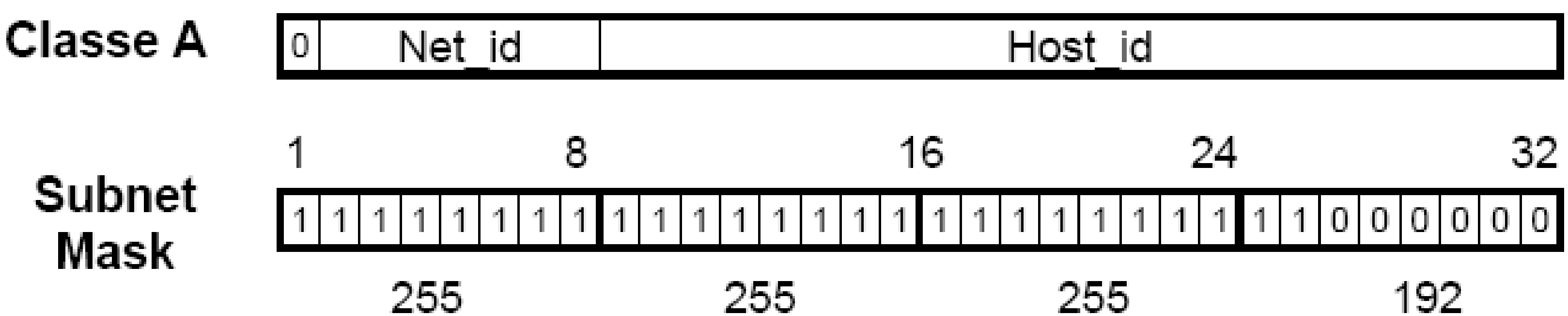


Subnet mask (255.255.255.0)

Subnetted address

Static subnetting

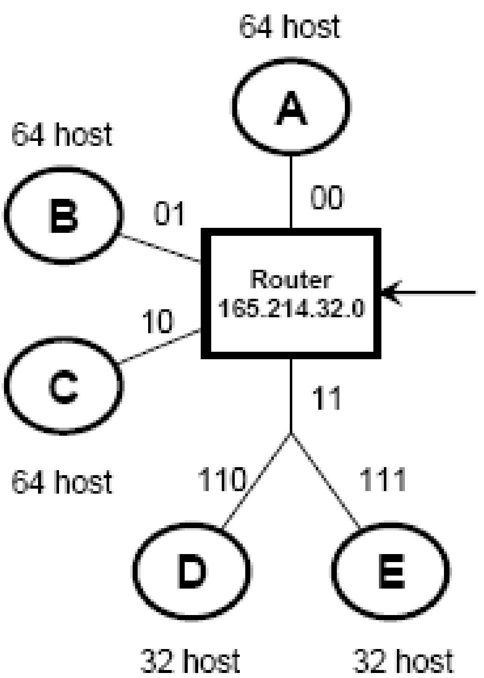
- All subnets share same subnet mask
- example:



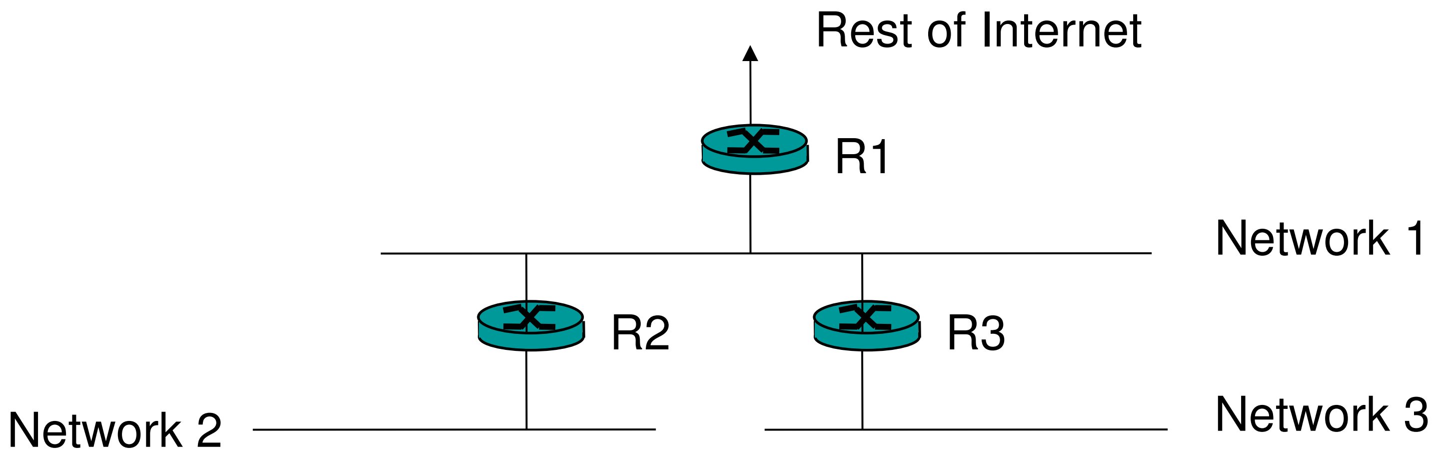
- numero massimo di sottoreti possibili = $2^{18} - 2 = 226.143$
- numero massimo di host per sottorete = $2^6 - 2 = 62$

Variable length subnetting

- Subnets are assigned different subnet masks
  - Allows definition of variable size subnets
- Example:
  - Router with class C address [165.214.32.0]
  - 5 subnets:
    - Subnet A, Subnet B, Subnet C: 50 hosts
    - Subnet D, Subnet E: 30 hosts
  - Subnetting
    - 4 subnets with 64 addresses each (Host_id: 6 bits) (subnet mask 255.255.255.192)
    - 1 subnet split into 2 further subnets with 32 addresses each (Host_id: 5 bits) (subnet mask 255.255.255.224)



Example



- R2 must know the number of bits identifying network 3's subnet prefix
- Rest of Internet ignores existence of R2 and R3

Routing tables with subnetting

- Routing tables' structure changes (costs omitted in the sequel for simplicity)
- Generic RT's line:

  (M, D, NH)

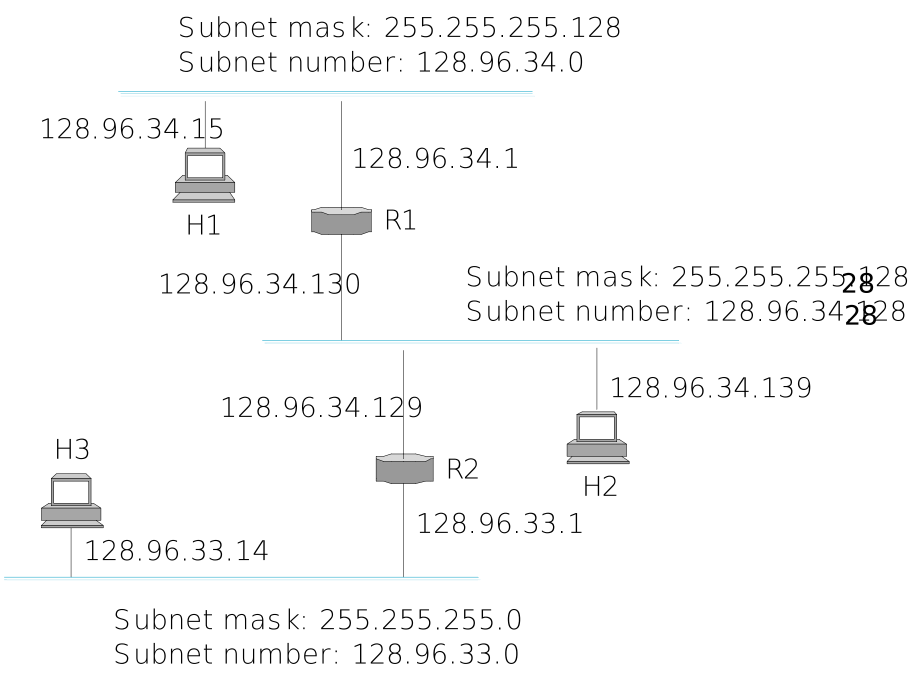
- Fields
  - D --> destination network address
  - NH --> address of next hop on the path towards destination netowork
  - M --> mask: 32 bits field that identifies network prefix

Routing tables with subnetting/cont.

	……….		
I	255.255.0.0	151.100.0.0	220.190.17.22
I+1	255.255.255.0	220.190.16.0	220.190.15.1
	…………		

- Example: resolve address 220.190.16.3
  - Try entry I: bit to bit AND with 255.255.0.0 returns 220.190.0.0 ≠ 151.100.0.0
  - Try entry I+1: bit to bit AND with 255.255.255.0 returns 220.190.16.0 --> next hop is 200.190.15.1

Subnet Example



Forwarding table at router R1

Subnet Number	Subnet Mask	Next Hop
128.96.34.0	255.255.255.128	interface 0
128.96.34.128	255.255.255.128	interface 1
128.96.33.0	255.255.255.0	R2

Forwarding Algorithm

```
D = destination IP address
for each entry (SubnetMask, SubnetNum, NextHop)
   D1 = SubnetMask & D
   if D1 = SubnetNum
      if NextHop is an interface
         deliver datagram directly to D
      else
         deliver datagram to NextHop
```

- Use a default router if nothing matches
- Not necessary for all 1s in subnet mask to be contiguous
- Can put multiple subnets on one physical network
- Subnets not visible from the rest of the Internet

Maschere di sottorete

- Sono consentite maschere arbitrarie
- Soluzione solitamente scelta (non e' una scelta obbligata):
  - Numero costante di bit della parte locale per individuare la sottorete
  - Allocazione contigua
  - Esempio: data la rete 151.100.0.0 di classe B, gli 8 bit piu' significativi della parte locale per la sottorete e gli 8 bit meno significativi per l'host
  - In questo caso la maschera di sottorete sarebbe 255.255.255.0

Classless Inter Domain Routing [CIDR]

- Situation in 1996
  - 100 % of class A addresses assigned
  - 61.95 % of class B addresses assigned (approaching exhaustion)
  - 36.44 % of class C addresses assigned
- CIDR conceived to
  - Slow down IP address space exhaustion (number of hosts doubling every year)
  - Reduce size of routing tables
  - Speed up routing
- CIDR obsoletes classful addressing but is compatible with it
- RTs’ structure different with respect to classful addressing

Classless Inter Domain Routing [CIDR]/cont.

CIDR based on Supernetting

- Upper half of A class (64 to 127) reserved for future use
- Class B address block assigned only if
  - Network contains at least 32 subnets
  - Networks has at least 4096 hosts overall
- Lower half of class C address space (192.0.0 to 207.255.255) partitioned into 8 blocks, each assigned to an authority responsible for a specific geographic area
  - Upper half of class C address space (208.0.0 to 223.255.255) not assigned
- If a network does not satisfy requirements for assignment of a class B block -> A sufficient number of *contiguous* class C blocks assigned
  - Network has unique prefix (defined by most significant bits)
  - Network identified only by prefix in routing tables

Classless Inter Domain Routing [CIDR]/cont.

- Geographic allocation of class C address blocks

Multiregional	192.0.0	193.255.255
Europe	194.0.0	195.255.255
Others	196.0.0	197.255.255
North America	198.0.0	199.255.255
Central/South America	200.0.0	201.255.255
Pacific Rim	202.0.0	203.255.255
Others	204.0.0	205.255.255
Others	206.0.0	207.255.255

- All networks belonging to same area have same 7 bits prefix
  - Example: Europe
    - 11000010 0 [194]
    - 11000011 1 [195]

Classless Inter Domain Routing [CIDR]/cont.

- Example 1
  - Address assignment in North America
    - CIDR mask for North America = 198.0.0.0/7
  - A major Internet Service Provider (ISP) is assigned 2048 class C blocks
    - From 198.24.0.0 (11000110.00011000.00000000.0)
    - To 198.31.255.0 (11000110.00011111.11111111.0)
    - CIDR mask for this ISP = 198.24.0.0/13
  - A local ISP requests 16 class C address blocks from the major ISP. It is assigned blocks
    - From 198.24.16.0 (11000110.00011000.00010000.0)
    - To 198.24.31.0 (11000110.00011000.00011111.0)
    - CIDR mask for local ISP = 198.24.16.0/20

Classless Inter Domain Routing [CIDR]/cont.

- Example 2
  - Address assignment in Europe
    - CIDR mask for Europe = 194.0.0.0/7
  - A european organization is assigned 2048 class C addresses
    - From 194.32.136.0 (11000010.00100000. 10001000.0)
    - To 194.32.143.0 (11000010. 00100000.10001111.0)
    - CIDR mask for this organization = 194.32.136.0/21

Classless Inter Domain Routing [CIDR]/cont.

- Network prefix computed with bit to bit AND between 32 bits IP address and (32 bits) network mask
- A whole address block represented by one line in Routing Table, corresponding to networks prefix (Supernetting)
- Correct entry chosen according to Longest Prefix Matching rule

```
11000000 00100000 10001000 00000000 = 192.32.136.0 (class C address)
11111111 11111111 11111000 00000000   255.255.248.0 (network mask)
  ===================================== logical_AND
11000000 00100000 10001             = 192.32.136 (IP prefix)
11000000 00100000 10001111 00000000 = 192.32.143.0 (class C address)
11111111 11111111 11111000 00000000   255.255.248.0 (network mask)
  ===================================== logical_AND
11000000 00100000 10001             = 192.32.136 (same IP prefix)
```

Longest Prefix Matching

Instradamento

- indirizzo 198.15.7.3
- indirizzo 198.15.7.4

198.15.7.3

- porta 1: matching prefisso 16
- porta 7: matching prefisso 24
- porta 4: matching prefisso 32

198.15.7.4

- porta 1: matching prefisso 16
- porta 7: matching prefisso 24
- porta 4: no matching

Tabella di instradamento

Prefix	Porta d'uscita
198.15.0.0/16	1
198.15.7.0/24	7
198.15.7.3/32	4

198.15.7.3 $\Rightarrow$ porta 4

198.15.7.4 $\Rightarrow$ porta 7

Indirizzamento senza classi

- Problema: una rete di classe C corrisponde a 256 indirizzi IP
- Molte organizzazioni hanno bisogno di piu' indirizzi
- Soluzione: assegnare ad una stessa organizzazione blocchi contigui di indirizzi in classe C
- Esempio: l'organizzazione X riceve i tre blocchi contigui 220.123.16.x, 220.123.17.x e 220.123.18.x
  - 768 indirizzi disponibili

Indirizzamento senza classi

- Come rappresentare il blocco di indirizzi assegnati?
- Informazioni necessarie: indirizzo + basso del blocco e No. blocchi
- In pratica:
  - CIDR usa le seguenti informazioni:
    - Valore a 32 bit dell'indirizzo piu' basso del blocco
    - Maschera a 32 che funziona come una maschera di sottorete standard
- CIDR compatibile con versioni meno recenti dei protocolli di routing

Differences with subnetting

- CIDR mask uses contiguous prefix
  - Not required in subnetting
- CIDR requires every address block to be a power of 2
- example:
  - 11111111 11111111 00011000 01000000 is a valid subnet mask but not a correct CIDR mask

CIDR rules/recap.

- Agregate block corresponds to number of entries in RT without aggregation that is a power of 2
  - 2, 4, 8, 16, … entries can be aggregated
  - For example 5 entries cannot be aggregated into 1 block
- Blocks correspond to contiguous address sequences - no “holes” allowed
- For every Routing Table entry: interface to which packets are forwarded needs to be the same for all addresses in aggregated block
- If non aggregated addresses differ in n-th byte, the value of that byte for the lowest address in the block must be multiple of block size

CIDR notation

- Network prefix + Number of prefix bits
- Example: address block 220.16.128.0 to 220.16.255.255:

220.16.128.0/17

11111111.11111111.10000000.00000000

Example

- An ISP receives CIDR address block 210.20.128.0/17
- ISP defines (for example) 128 networks with 256 IP addresses each:
  - 210.20.128.0/24
  - 210.20.129.0/24
  - …………
- Network mask corresponding to main block is 255.255.128.0
- Each of 128 sub-block has network mask 255.255.255.0

Routing with CIDR

- Indirizzamento senza classi: occorre separare il prefisso di rete dall'indirizzo dell'host
- Nelle tabelle si memorizzano soltanto i prefissi di rete
- Esempio: arriva Dgram con IP destinazione 150.122.19.30
  - Qual e' il prefisso di rete?
  - Non e' possibile stabilirlo a priori se si usa la convenzione CIDR

Forwarding with CIDR

- Generic Routing Table entry has format (Mask, Dest. network, Next hop)
- Try entries according to decreasing length of network mask
  - Longest prefix first
- Example: address 150.122.19.30
  - RT contains 2 entries, corresponding to destination networks 150.122.19./24 and 150.122.16./20
  - Correct entry is the first one, corresponding to longest prefix matching

Implementation

- Binary tree search structures (trie) or evolutions thereof used
- Every root-leaf path corresponds to a possible prefix
- Search of longest matching prefix fast -> occurs in router's main memory (RAM)

Example

- Routing table (only destinations listed):

32 bit address unique prefix

01101010000000000000000000000000 01101

01000110000000000000000000000000 0100

01010110000000000000000000000000 0101

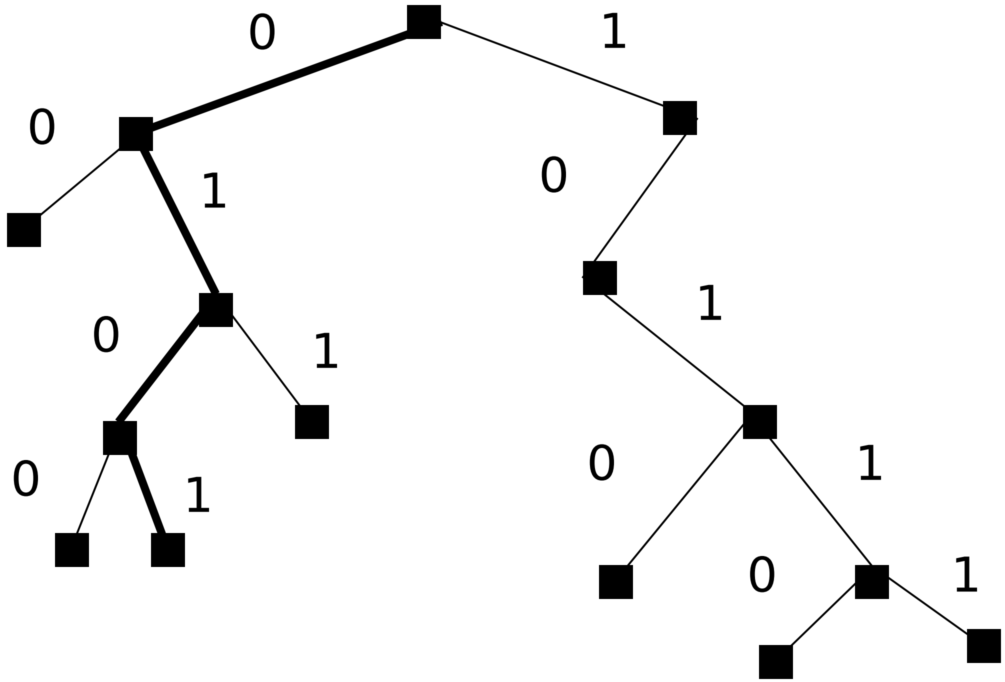
01100001000000000000000000000000 01100

10101010111100000000000000000000 1010

10110000000000100000000000000000 10110

10111011000010100000000000000000 10111

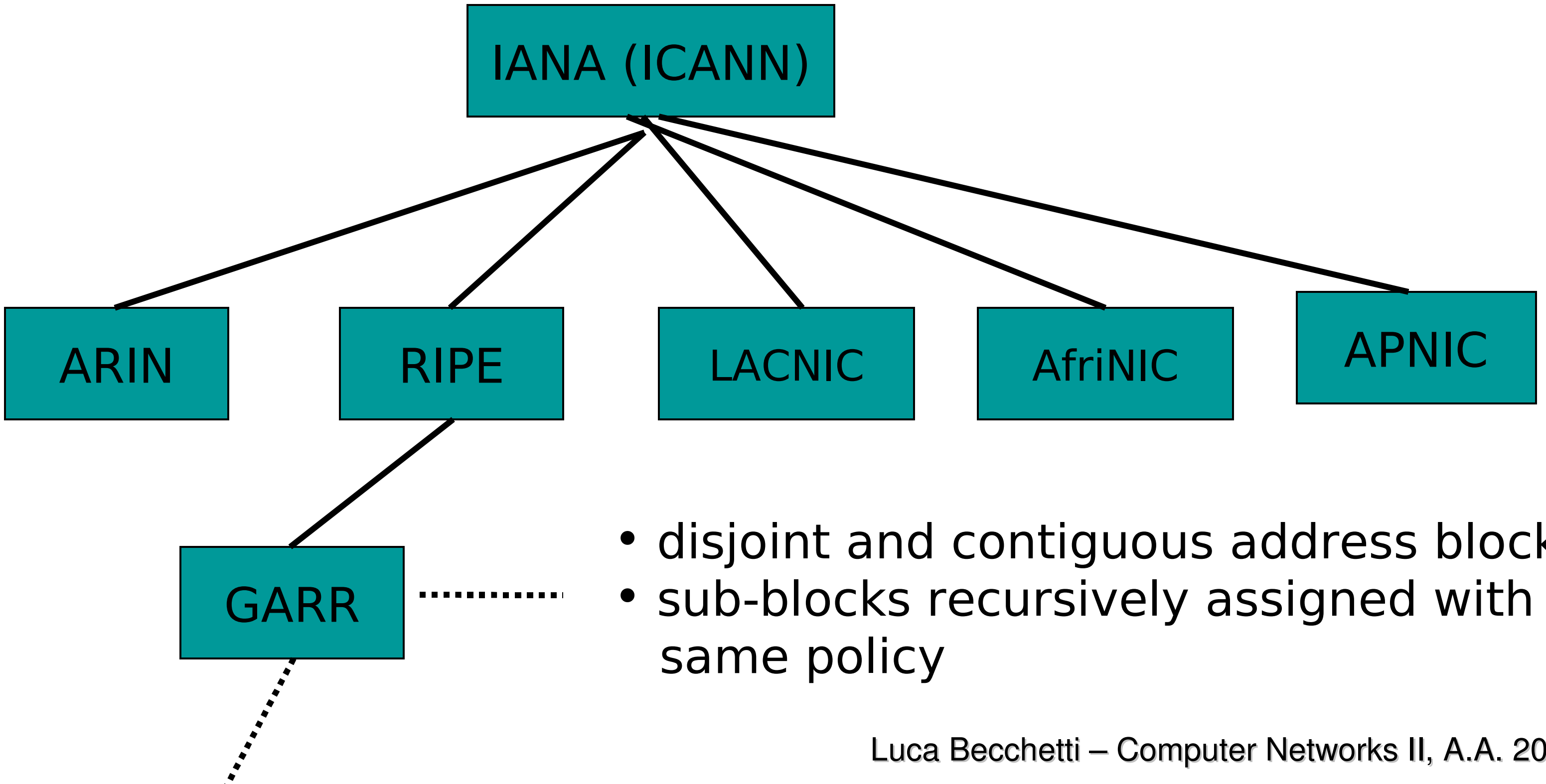
Example/cont.



- Search for address 01010010.x.y.z
- Every leaf corresponds to a possible destination (entry) in the Routing Table

Why it works

- Hierarchical assignment of IP addresses
- Check for acronyms on Wikipedia



- disjoint and contiguous address blocks
- sub-blocks recursively assigned with same policy

Private IP addresses

- The following address blocks are reserved for private networks, not requiring (direct) Internet connectivity

16 blocchi di Indirizzi di classe B	172.16.0	172.31.0
256 blocchi di Indirizzi di classe C	192.168.0	192.168.255

- Every network can use any of this addresses internally
- Border routers of a network using private addresses do not externally propagate routing information concerning internal hosts
- A host using a private address cannot directly connect to the Internet. It needs to access indirectly through an "Application Level Gateway (Proxy)"

References

- TCP/IP guide: http://www.tcpipguide.com/free/t_IPAddressing.htm
- Kurose – Ross, 4.4
- Description of a Routing Table on Windows server 2003: http://technet2.microsoft.com/WindowsServer/en/library/63158f32-9fcd-42ea-ba2f-8008bb7bb5241033.mspx?mfr=true
- ARIN: http://www.arin.net/index.shtml
- RIPE: http://www.ripe.net/ -> contains teaching material