



Reti

(già "Reti di Calcolatori")

Livello Rete

Indizzamento IP (v4) e inoltro dei pacchetti

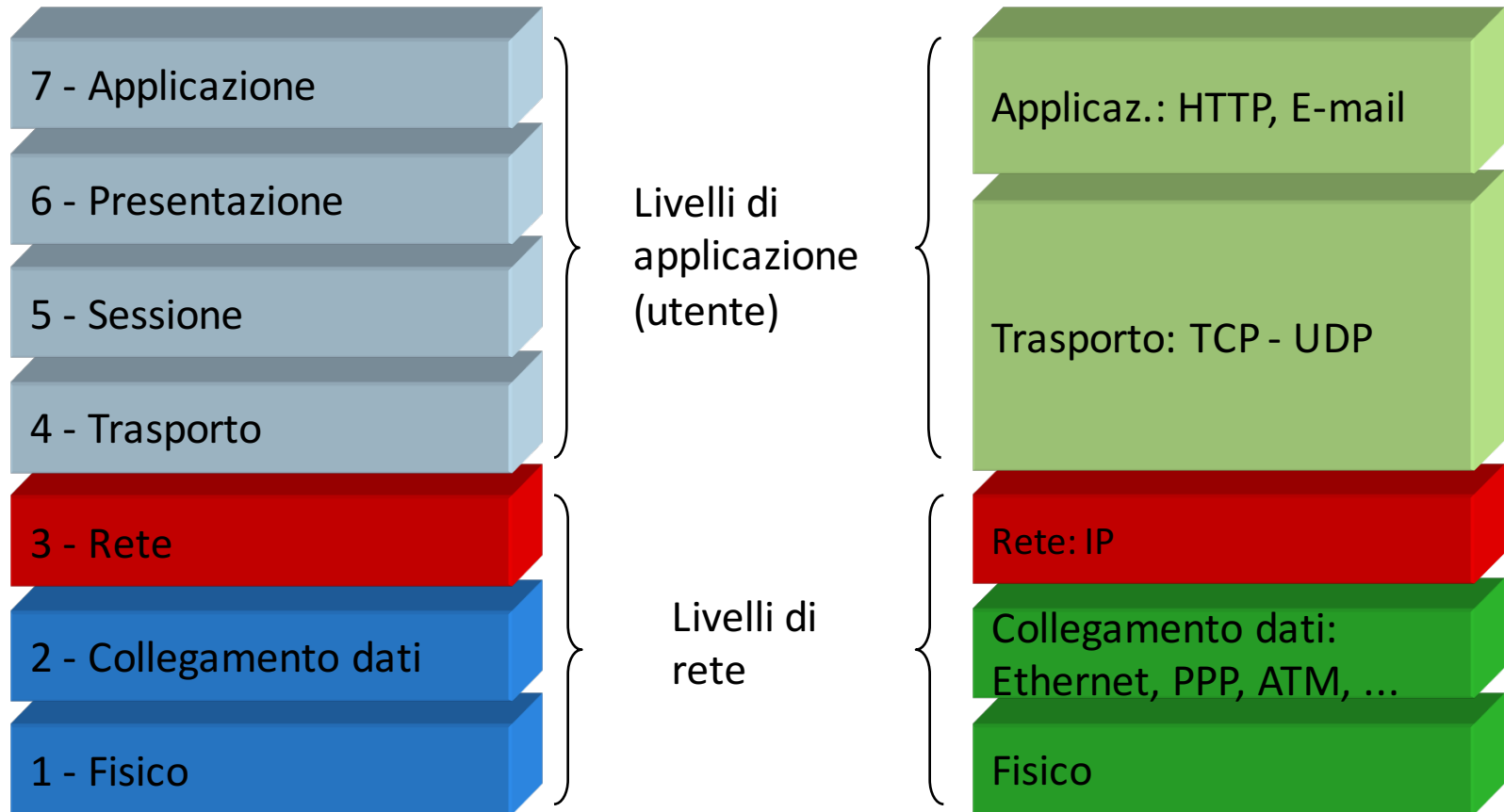
Renato Lo Cigno

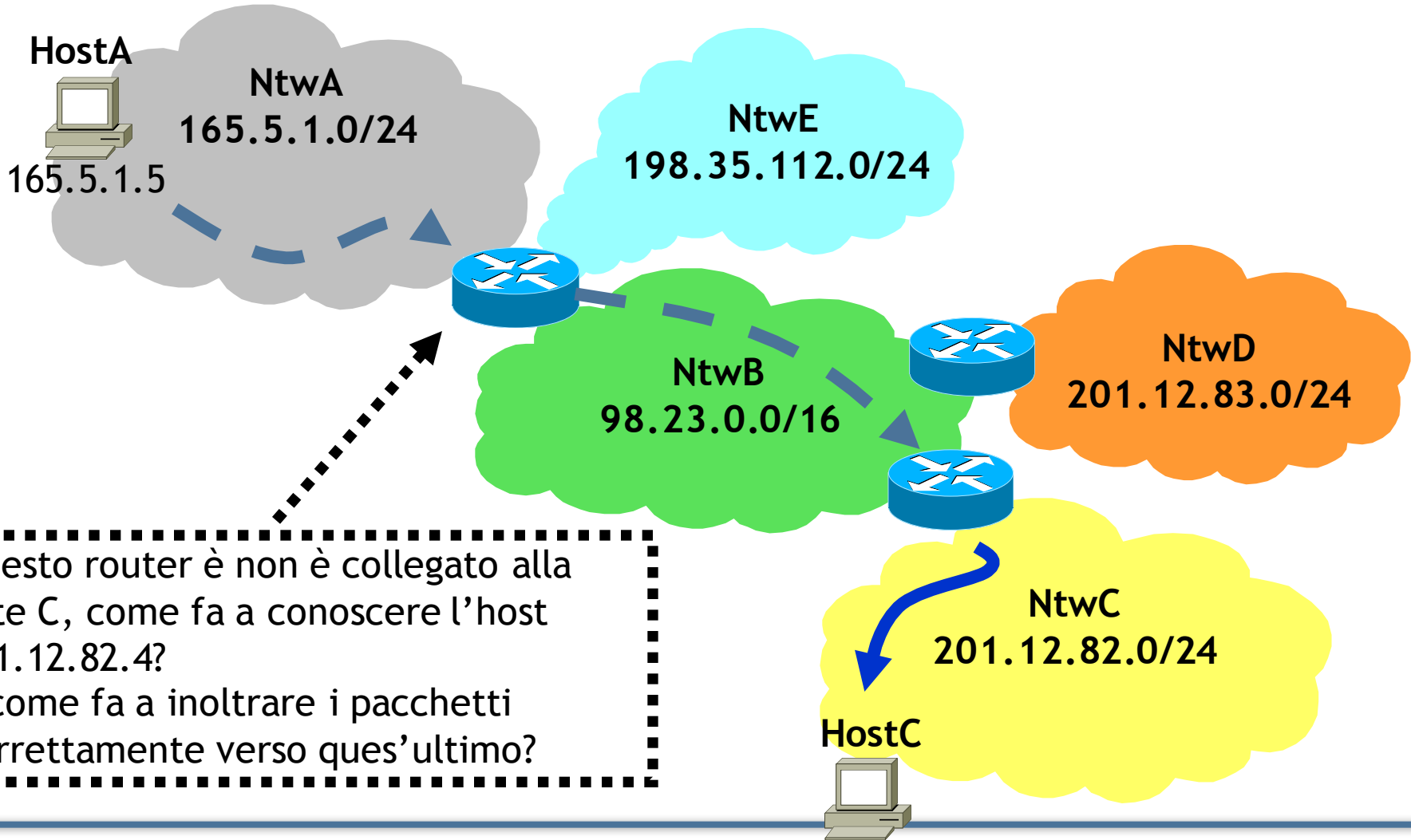
<http://disi.unitn.it/locigno/index.php/teaching-duties/computer-networks>

- *Credits*
 - *Part of the material is based on slides provided by the following authors*
 - *Jim Kurose, Keith Ross, “Computer Networking: A Top Down Approach,” 4th edition, Addison-Wesley, July 2007*
 - *Douglas Comer, “Computer Networks and Internets,” 5th edition, Prentice Hall*
 - *Behrouz A. Forouzan, Sophia Chung Fegan, “TCP/IP Protocol Suite,” McGraw-Hill, January 2005*
- La traduzione, se presente, è in generale opera (e responsabilità) del docente



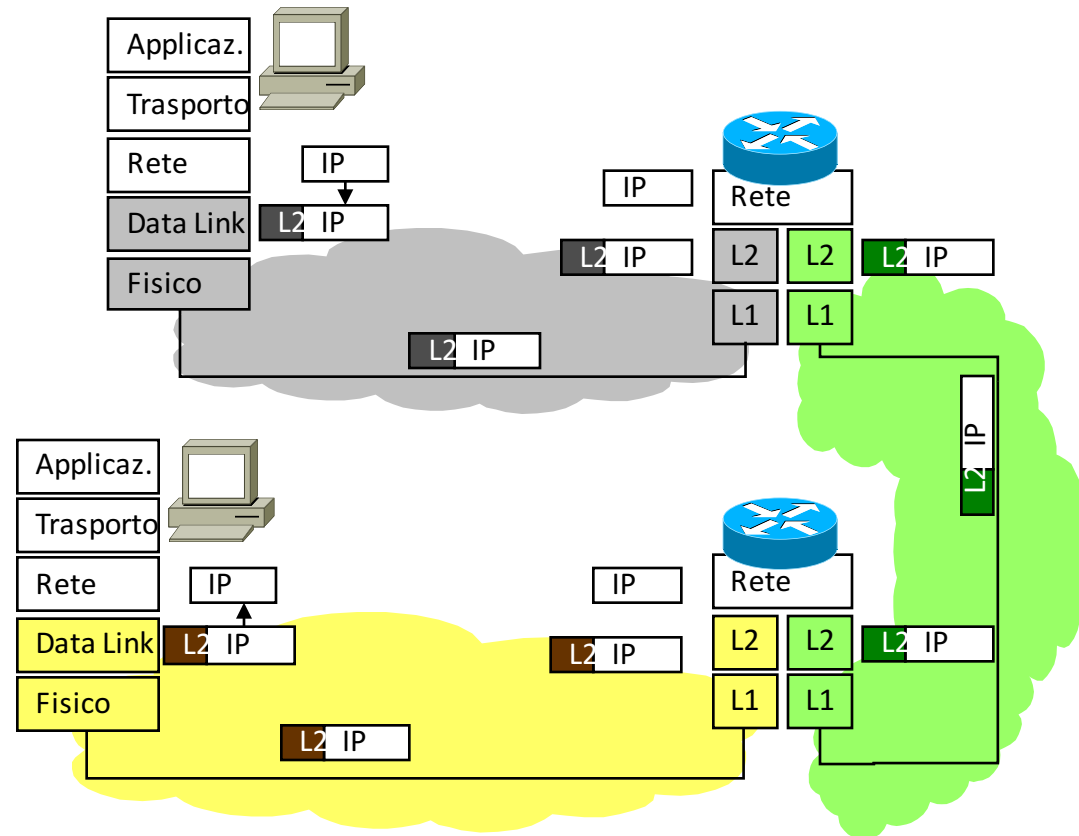
- **Spazio di indirizzamento**
- **Indirizzi IP e loro uso**
- **Consegna dei pacchetti**
- Configurazione dei PC e delle reti
- Instradamento e Routing

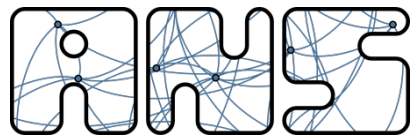




Questo router non è collegato alla rete C, come fa a conoscere l'host 201.12.82.4?
E come fa a inoltrare i pacchetti correttamente verso quest'ultimo?

- Trasporto dei pacchetti da sorgente a ricevitore. I pacchetti contengono un segmento di livello trasporto
- I pacchetti sono incapsulati in trame L2
- Al ricevitore i segmenti sono estratti dai pacchetti e consegnati al livello trasporto
- **I protocolli di rete sono in tutti gli host e router**
- Un router deve esaminare l'intestazione di tutti i pacchetti che lo attraversano





- **Instradamento (Routing)**

- Trovare il percorso dalla sorgente alla destinazione

- Algoritmi di Routing

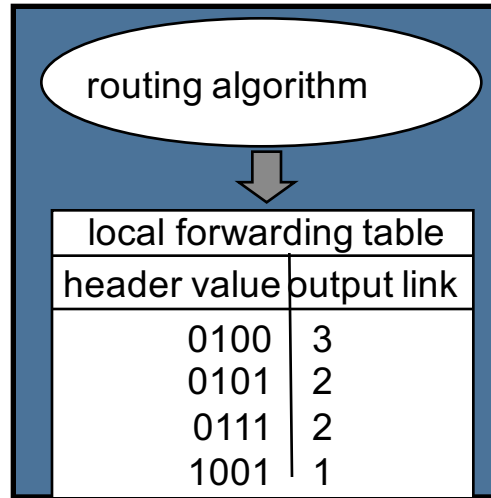
- Simile a pianificare un viaggio: devo determinare le strade da fare e gli incroci in cui cambiare la mia strada

- **Inoltro (Forwarding)**

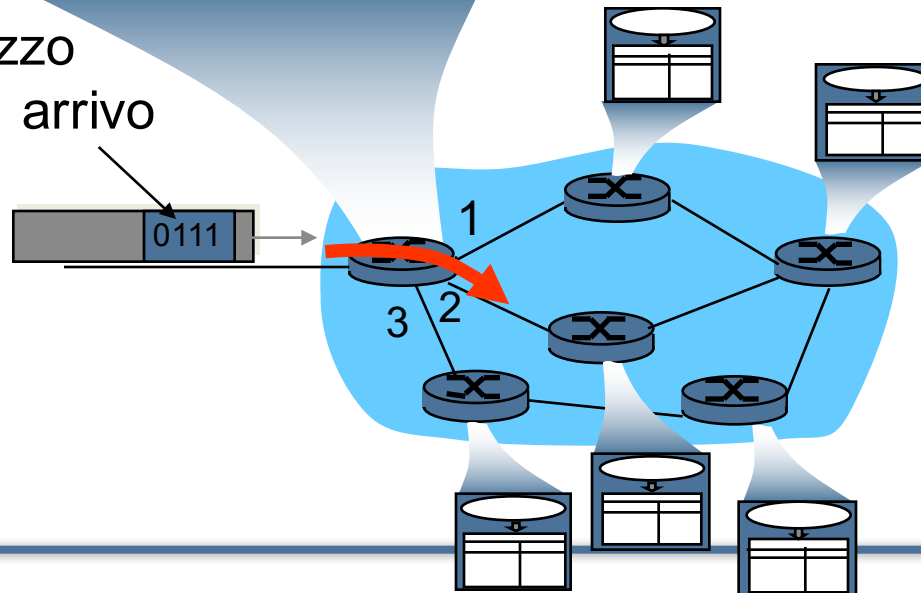
- Funzione che esegue il trasporto dei pacchetti dagli ingressi alle uscite dei router ... dato che il percorso è già noto

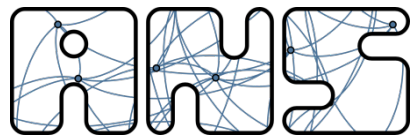
- Simile a prendere l'uscita giusta di una rotonda, sapendo che devo andare in una specifica direzione

- Entrambe richiedono uno spazio di indirizzamento appropriato ... e i relativi indirizzi

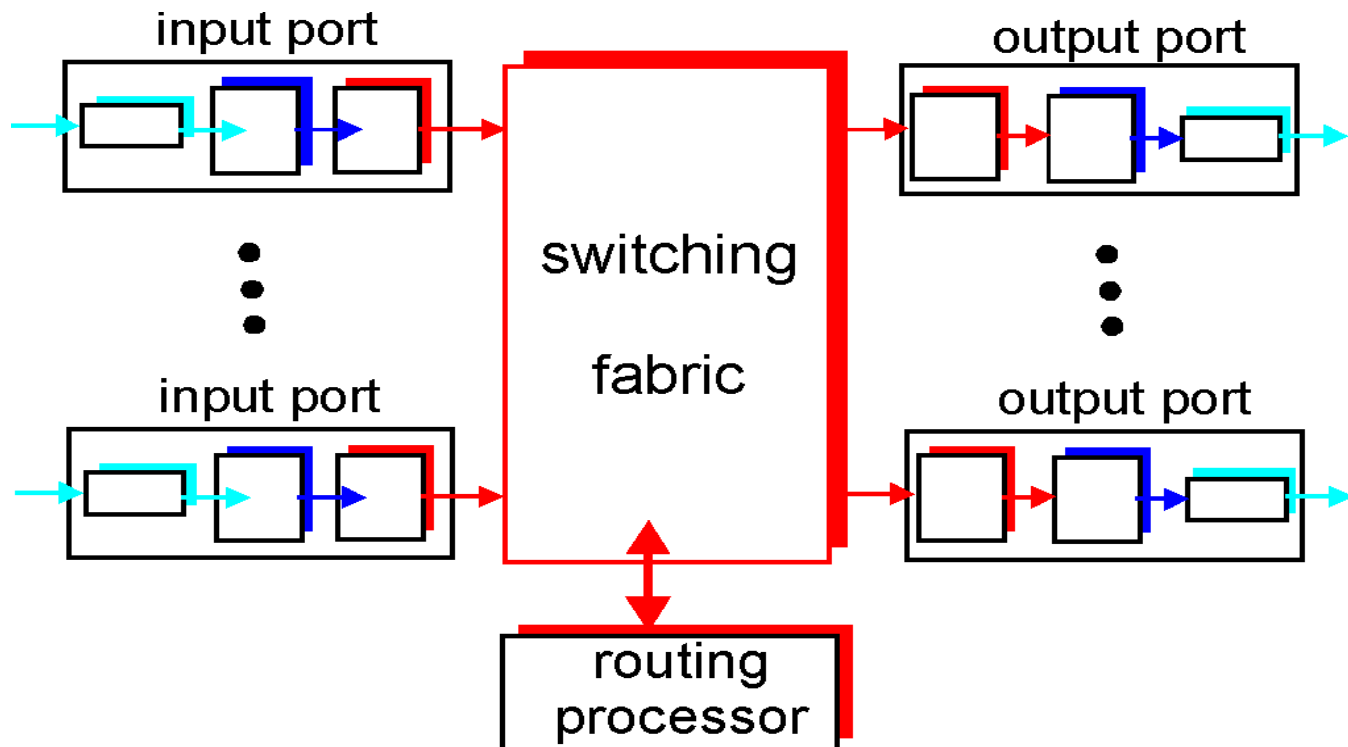


valore dell'indirizzo nel pacchetto in arrivo





- due funzioni fondamentali:
 - eseguire i protocolli e algoritmi di instradamento (RIP, OSPF, BGP)
 - inoltrare i datagrammi (pacchetti) dagli ingressi alle uscite





IL PROTOCOLLO IP

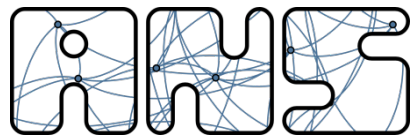
(VERSIONE 4)



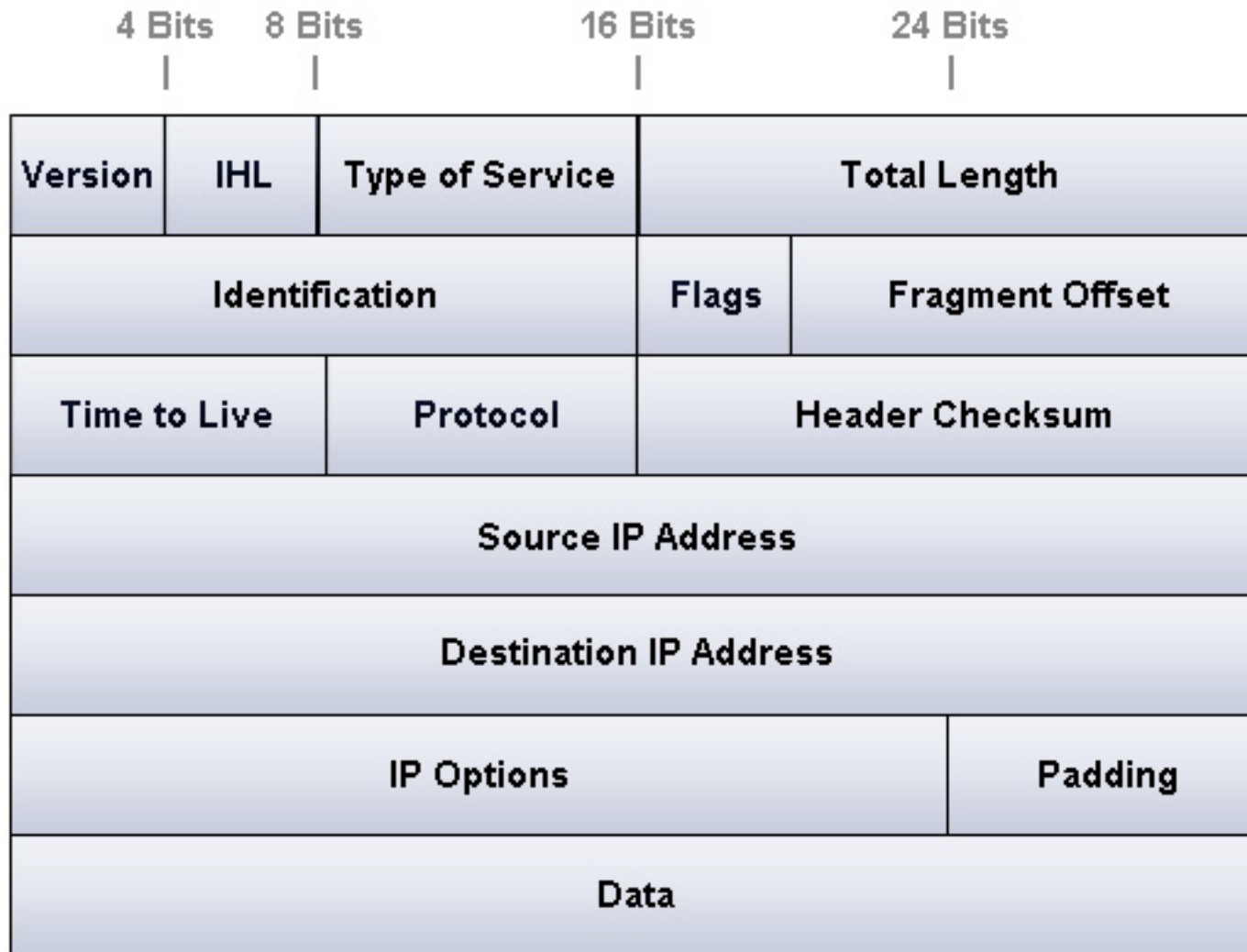
- TCP/IP usa il termine “IP datagram” per identificare un pacchetto di livello rete
- Ciascun datagramma è composto da una **intestazione (header)**
 - da 20 a 60 bytes che contengono le informazioni essenziali per l’instradamento e la consegna
- seguita dai **dati trasportati (payload)**
 - La dimensione dei payload non è fissa
 - La dimensione effettiva è determinata dall’applicazione e/o dal protocollo di trasporto
 - C’è una dimensione massima di 64kB (65536)

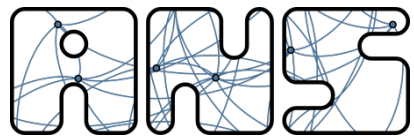


- L'header contiene informazioni utili per l'instradamento del datagramma:
 - L'indirizzo della sorgente (chi ha inviato per primo il datagramma)
 - L'indirizzo della destinazione (chi riceverà il datagramma se tutto va bene)
 - Il tipo di protocollo che ha generato i dati
 - ...
- Gli indirizzi sono di tipo IP (ovvio)
- Gli indirizzi MAC sono "esterni" al datagramma
- I campi sono di dimensione fissa per efficienza di manipolazione

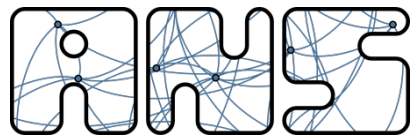


The IP Datagram Header Format





- **VERS:** Each datagram begins with a 4-bit protocol version number
- **H.LEN:** 4-bit header specifies the number of 32-bit quantities in the header
 - If no options are present, the value is 5
- Type of Service (**ToS**)
 - 8-bit field that carries a class of service for the datagram
 - potentially used for DiffServ and ECN (Explicit Congestion Notification)
 - seldom used in practice
- **TOTAL LENGTH:** 16-bit integer that specifies the total number of bytes including both the header and the data



- IDENTIFICATION
 - 16-bit number (usually sequential) assigned to the datagram
 - used to gather all fragments for reassembly to the datagram
- FLAGS
 - 3-bit field with individual bits specifying whether the datagram is a fragment
 - If so, then whether the fragment corresponds to the last piece of the original datagram
- FRAGMENT OFFSET
 - 13-bit field that specifies where in the original datagram the data in this fragment belongs
 - the value of the field is multiplied by 8 to obtain an offset



- **TIME TO LIVE (TTL)**
 - 8-bit integer initialized by the original sender
 - it is decremented **by each router** that processes the datagram
 - if the value **reaches zero (0) the datagram is discarded** and an error message is sent back to the source
- **PROTOCOL**
 - 8-bit field that specifies the type of the payload, i.e., the protocol above (e.g., 6 for TCP, 17 for UDP)
- **HEADER CHECKSUM**
 - 16-bit ones-complement checksum of header fields
- **SOURCE IP ADDRESS**
 - 32-bit Internet address of the original sender



- DESTINATION IP ADDRESS
 - The 32-bit Internet address of the ultimate destination
- IP OPTIONS
 - Optional header fields used to control routing and datagram processing
 - Most datagrams do not contain any options
- PADDING
 - If options do not end on a 32-bit boundary
 - zero bits of padding are added to make the header a multiple of 32 bits

FRAMMENTAZIONE DEI PACCHETTI IP

**L'USO DELLA FRAMMENTAZIONE È "DEPRECATO".
MOLTI ROUTER SEMPLICEMENTE NON LA IMPLEMENTANO
E SCARTANO IL PACCHETTO.**



- Each hardware technology specifies the maximum amount of data that a frame can carry
 - The limit is known as a **Maximum Transmission Unit (MTU)**
- Network hardware is not designed to accept or transfer frames that carry more data than the MTU allows
 - A datagram must be smaller or equal to the network MTU
 - or it cannot be encapsulated for transmission
- In an internet that contains heterogeneous networks, MTU restrictions create a problem
- A router can connect networks with different MTU values
 - a datagram that a router receives over one network can be too large to send over another network



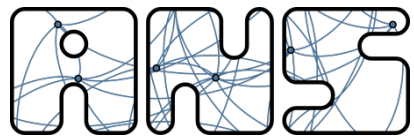
MTUs for some networks



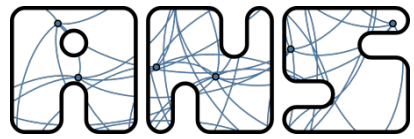
<i>Protocol</i>	<i>MTU</i>
Hyperchannel	65,535
Token Ring (16 Mbps)	17,914
Token Ring (4 Mbps)	4,464
FDDI	4,352
Ethernet	1,500
X.25	576
PPP	variabile



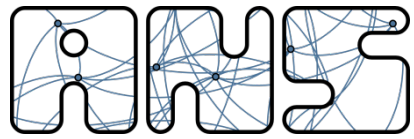
- Example: a router interconnects two networks with MTU values of 1500 and 1000
 - Host H_1 attaches to a network with an MTU of 1500
 - and can send a datagram that is up to 1500 octets
 - Host H_2 attaches to a network that has an MTU of 1000
 - which means that it cannot send/receive a datagram larger than 1000 octets
 - If host H_1 sends a 1500-octet datagram to host H_2
 - router R will not be able to encapsulate it for transmission across network 2



- When a datagram is larger than the MTU of the network over which it must be sent
 - the router divides the datagram into smaller pieces called fragments
 - and sends each fragment independently
- A fragment has the same format as other datagrams
 - a bit in the FLAGS field of the header indicates whether a datagram is a fragment or a complete datagram
- Other fields in the header are assigned information for the ultimate destination to reassemble fragments
 - to reproduce the original datagram
- The FRAGMENT OFFSET specifies where in the original datagram the fragment belongs



- A router uses the network MTU and the header size to calculate
 - the maximum amount of data that can be sent in each fragment
 - and the number of fragments that will be needed
- The router creates the fragments
 - It uses fields from the original header to create a fragment header
 - It copies the appropriate data from the original datagram into the fragment
 - Transmits the result



Flags field

D: Do not fragment
M: More fragments

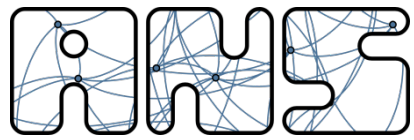




- Example: packets sent from H₁ to H₂
 - if host H₁ sends a 1500-octet datagram to host H₂, router R₁ will divide the datagram into two fragments, which it will forward to R₂
 - Router R₂ does not reassemble the fragments
 - Instead R uses the destination address in a fragment to forward the fragment as usual
 - The ultimate destination host, H₂, collects the fragments and reassembles them to produce the original datagram



- Requiring the ultimate destination to reassemble fragments has two advantages:
 - It reduces the amount of state information in routers
 - When forwarding a datagram, a router does not need to know whether the datagram is a fragment
 - It allows **routes** to change dynamically
 - If an intermediate router were to reassemble fragments, all fragments would need to reach the router
- By postponing reassembly until the ultimate destination
 - IP is free to pass some fragments from a datagram along different routes than other fragments



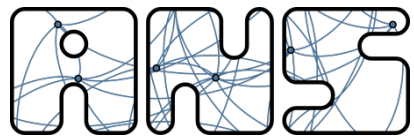
- A datagram cannot be reassembled until all fragments arrive
- The receiver must save (buffer) the fragments
 - In case missing fragments are only delayed
 - A receiver cannot hold fragments an arbitrarily long time
- IP specifies a maximum time to hold fragments
- When the first fragment arrives from a given datagram
 - the receiver starts a reassembly timer
- If all fragments of a datagram arrive before the timer expires
 - the receiver cancels the timer and reassembles the datagram
- Otherwise the receiver discards the fragments



GLI INDIRIZZI DI INTERNET: IPV4



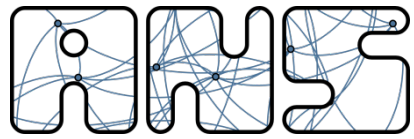
- Lo spazio di indirizzamento è un componente critico
- Tutti gli host e i router devono usare uno schema di indirizzamento **uniforme**
- Gli indirizzi Unicast, che identificano una specifica interfaccia devono essere **unici**
- Esistono due spazi di indirizzamento specificati per Internet
 - **IPv4**: quello attualmente in uso con indirizzi a 32 bit
 - IPv6: il sistema di indirizzamento che avrebbe dovuto sostituire IPv4, ma che continua a non farlo
 - indirizzi a 128 bit
 - funzioni “avanzate”
 - esistono molte “isole” IPv6 e ormai tutti i router dei maggiori vendor lo supportano



- IP addresses are supplied by protocol software
- Each network interface is assigned a unique 32-bit number
 - The interface **IP address** or **Internet address**
- When sending a packet across the Internet, sender's protocol software must specify
 - its own 32-bit IP address (the source address)
 - and the address of the intended recipient (the destination address)
- Routers only use the destination address for forwarding and routing

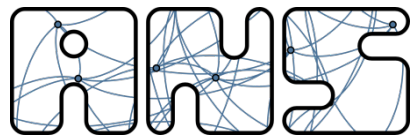


- Instead of writing 32 bits, a notation more convenient for humans to understand is used, known as **dotted decimal notation**
 - express each 8-bit section of a 32-bit number as a decimal value
 - use periods to separate the sections
- Dotted decimal treats each octet (byte) as an unsigned binary integer
 - the smallest value, 0
 - occurs when all bits of an octet are zero (0)
 - the largest value, 255
 - occurs when all bits of an octet are one (1)
 - dotted decimal addresses range
0.0.0.0 through 255.255.255.255

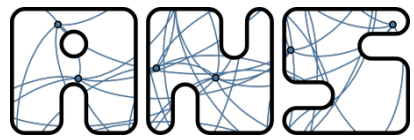


Dotted Decimal Notation: examples

32-bit Binary Number	Equivalent Dotted Decimal
1000001 00110100 00000110 00000000	129 . 52 . 6 . 0
11000000 00000101 00110000 00000011	192 . 5 . 48 . 3
00001010 00000010 00000000 00100101	10 . 2 . 0 . 37
10000000 00001010 00000010 00000011	128 . 10 . 2 . 3
10000000 10000000 11111111 00000000	128 . 128 . 255 . 0



- IP address is divided into two parts:
- A **prefix** → identifies the physical network to which the host is attached (also known as NetID)
 - Each network in the Internet is assigned a unique network number
- A **suffix** → identifies a specific interface on the network (also known as HostID)
 - Each NIC on a given network is assigned a unique suffix
- IP address scheme guarantees two properties:
 - Each computer is assigned a unique address
 - Network numbers (prefix) must be coordinated globally
 - Suffixes are assigned locally without global coordination



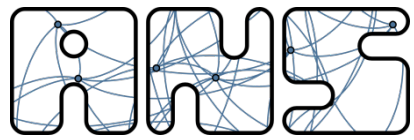
INDIRIZZAMENTO CON CLASSI (OBSOLETO)

Schema di organizzazione degli indirizzi usato fino alla metà degli anni '90 e basato su una divisione statica tra NetID e HostID

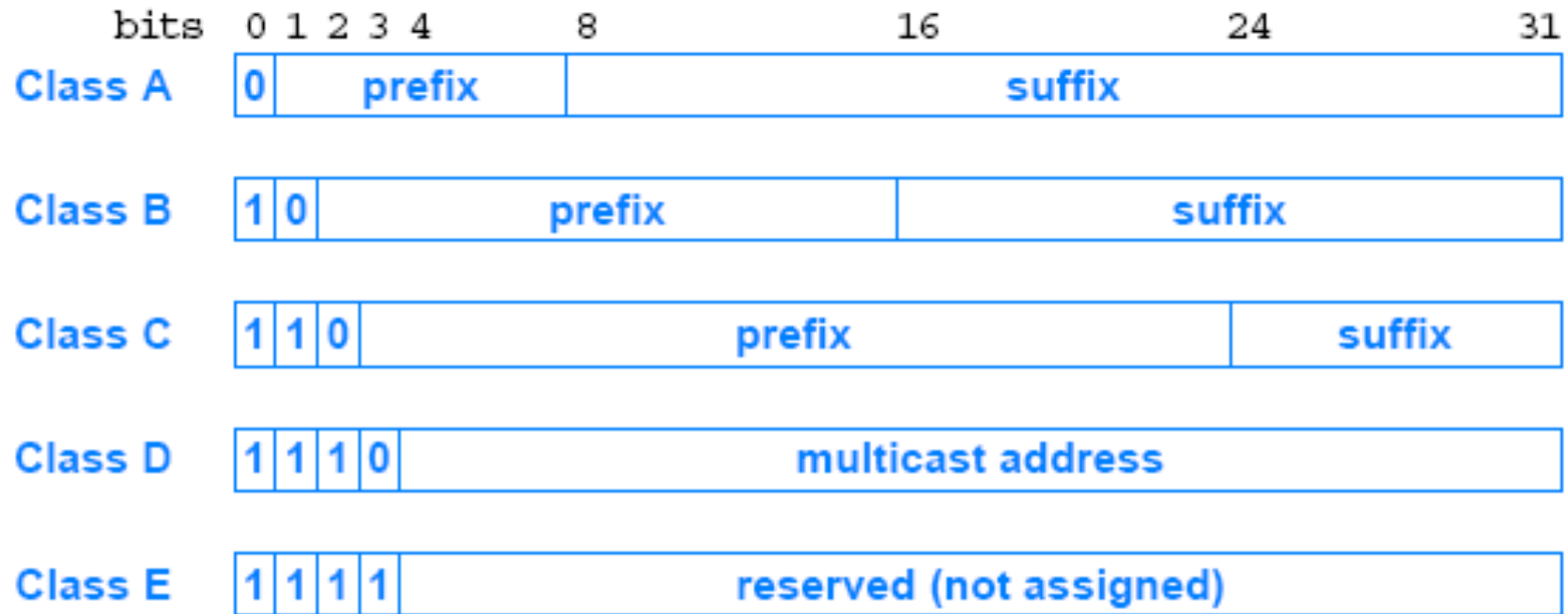
È uso ancora oggi riferire l'organizzazione degli indirizzi ad un concetto (e terminologia) di classe



- Internet contains a few large physical networks and many small networks
 - the designers chose an addressing scheme to accommodate a combination of large and small networks
- The original **classful** IP addressing divided the IP address space into 3 primary classes
 - each class has a different size prefix and suffix
- The first four bits of an IP address determined the class to which the address belonged
 - It specifies how the remainder of the address was divided into prefix and suffix



Original Classes of IP Addresses



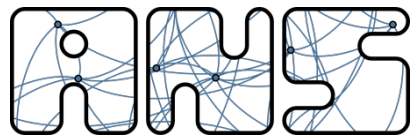


- The classful scheme divided the address space into unequal sizes
- The designers chose an unequal division to accommodate a variety of scenarios
 - For example, although it is limited to 128 networks, class A contains half of all addresses
 - The motivation was to allow major ISPs to each deploy a large network that connected millions of computers
 - But A classes were assigned to small networks all in the US ...
 - Similarly, the motivation for class C was to allow an organization to have a few computers connected on a LAN



Division of the Address Space

Address Class	Bits In Prefix	Maximum Number of Networks	Bits In Suffix	Maximum Number Of Hosts Per Network
A	7	128	24	16777216
B	14	16384	16	65536
C	21	2097152	8	256

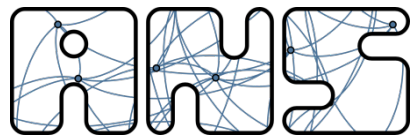


- Internet Corporation for Assigned Names and Numbers (**ICANN**) authority has been established
 - to handle address assignment and adjudicate disputes
- ICANN does not assign individual prefixes
 - Instead, ICANN authorizes a set of **registrars** to do so
- Registrars make blocks of addresses available to ISPs
 - ISPs provide addresses to subscribers
- To obtain a prefix
 - a corporation usually contacts an ISP

INDIRIZZAMENTO SENZA CLASSI E CIDR

Schema in uso attuale con divisione dinamica tra NetID e HostID

CIDR (Classless Inter-Domain Routing) consente l'instradamento globale senza usare la nozione di classe



- As the Internet grew the original classful addressing scheme became a limitation
- Everyone demanded a class A or class B address
 - So they would have enough addresses for future growth
 - but many addresses in class A and B were unused
- Two mechanisms, closely related, were designed to overcome the limitation
 - Subnet addressing
 - Classless addressing
- Instead of having three distinct address classes, allow the division between prefix/suffix on an arbitrary bit boundary



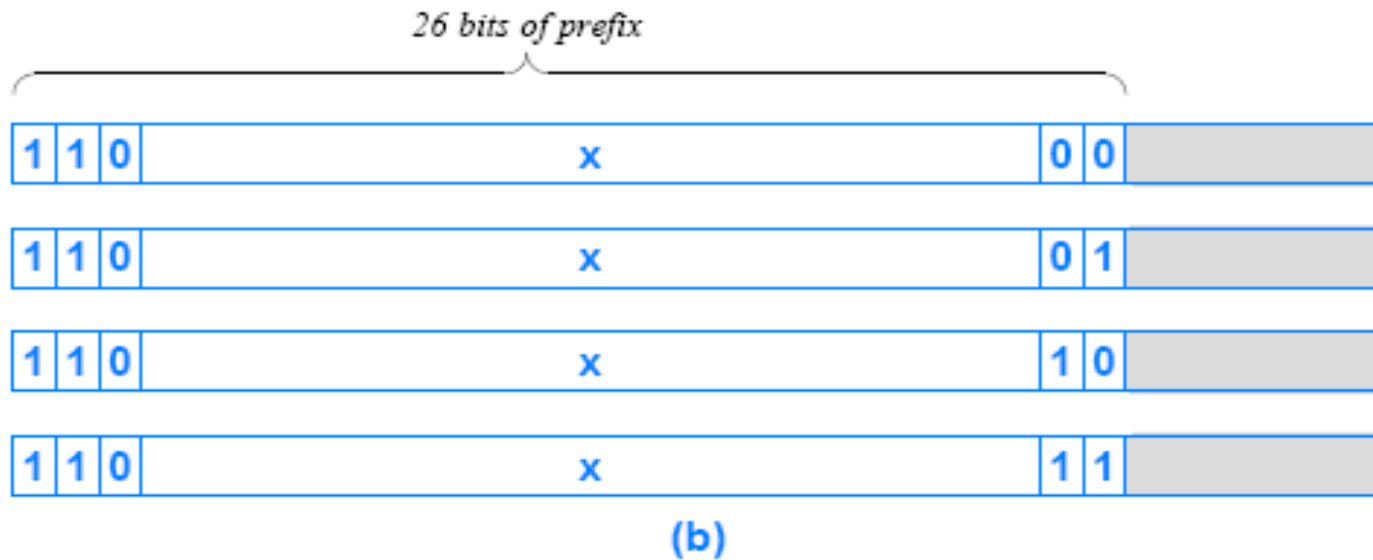
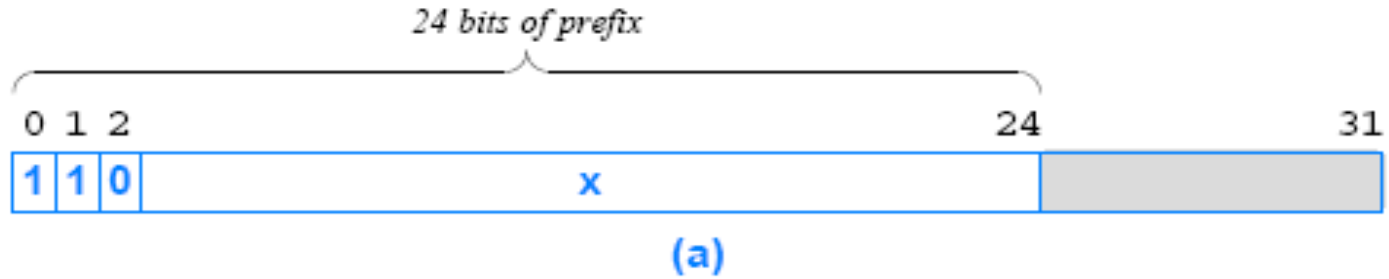
- Consider an ISP that hands out prefixes. Suppose a customer of the ISP requests a prefix for a network that contains 55 hosts
 - classful addressing requires a complete class C prefix
 - only 6 bits of suffix are needed to represent all possible host values
 - means 190 of the 254 possible suffixes would never be assigned
 - most of the class C address space is wasted
- For the above example
 - classless addressing allows the ISP to assign
 - a prefix that is 26 bits long
 - a suffix that is 6 bits long



- Assume an ISP owns a class C prefix
 - Classful addressing assigns the entire prefix to one organization
- With classless addressing
 - the ISP can divide the prefix into several longer prefixes
- For instance, the ISP can divide a class C prefix into 4 longer prefixes
 - each one can accommodate a network of up to 62 hosts
 - all 0s and all 1s are reserved
- The original class C address has 8 bits of suffix
 - and each of the classless addresses has 6 bits of suffix
- Thus, instead of wasting addresses
 - ISP can assign each of the 4 classless prefixes to a subscriber



Classless Addressing: Example

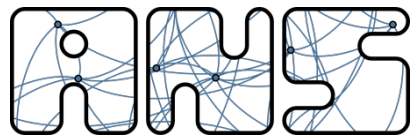




- How can an IP address be divided at an arbitrary boundary?
- The classless and subnet addressing schemes require hosts and routers to store an additional piece of information:
 - a value that specifies the exact boundary between prefix and suffix
- To mark the boundary, IP uses a 32-bit value
 - known as an **address mask**, also called a **subnet mask**
- Why store the boundary size as a bit mask?
 - A mask makes processing efficient
- Hosts and routers need to compare the network prefix portion of the address to a value in their forwarding tables
 - The bit-mask representation makes the comparison efficient



- Suppose a router is given
 - a destination address, D
 - a network prefix represented as a 32-bit value, N
 - a 32-bit address mask, M
- Assume the top bits of N contain a network prefix, and the remaining bits have been set to zero
- To test whether the destination lies on the specified network, the router tests the condition:
$$N == (D \& M)$$
- The router
 - uses the mask with a “logical and (&)” operation to set the host bits of address D to zero (0)
 - and then compares the result with the network prefix N



- Consider the following 32-bit network prefix:
10000000 00001010 00000000 00000000 → 128.10.0.0
- Consider a 32-bit mask:
11111111 11111111 00000000 00000000 → 255.255.0.0
- Consider a 32-bit destination address, which has a
10000000 00001010 00000010 00000011 → 128.10.2.3
- A logical & between the destination address and the address mask extracts the high-order 16-bits
10000000 00001010 00000000 00000000 → 128.10.0.0



- Classless Inter-Domain Routing (CIDR)
- Consider a mask defining a subnet with 2^6 nodes
 - It has 26 bits of 1s followed by 6 bits of 0s
 - In dotted decimal, the mask is: 255.255.255.192
- The general form of CIDR notation is: `ddd.ddd.ddd.ddd/m`
 - `ddd` is the decimal value for an octet of the address
 - `m` is the number of one bits in the mask
- Thus, one might write the following: 192.5.48.69/26
 - which specifies a mask of 26 bits



A CIDR Example

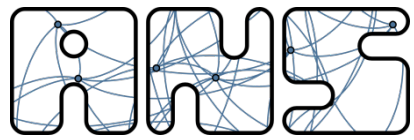
- Assume an ISP has the following block 128.211.0.0/16
- Suppose the ISP has 2 customers
 - one customer needs 12 IP addresses and the other needs 9
- The ISP can assign
 - customer1 CIDR: 128.211.0.16/28
 - customer2 CIDR: 128.211.0.32/28
 - both customers have the same mask size (28 bits), the prefixes differ



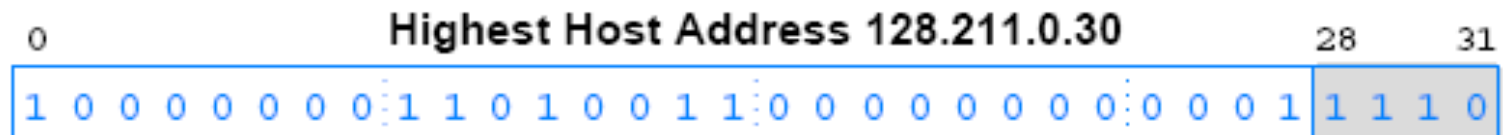
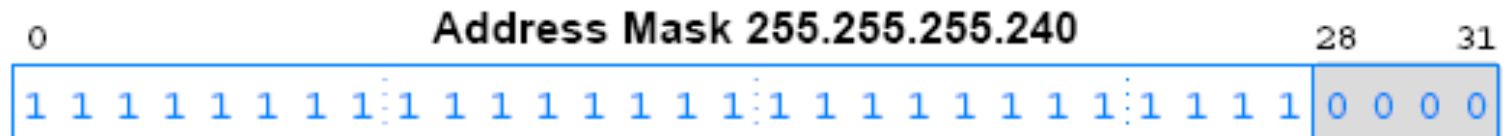
- The binary value assigned to customer1 is:
 - 10000000 11010011 00000000 00010000
- The binary value assigned to customer2 is:
 - 10000000 11010011 00000000 00100000
- There is no ambiguity
 - Each customer has a unique prefix
 - More important, the ISP retains most of the original address block
 - it can then allocate to other customers



- Once an ISP assigns a customer a CIDR prefix
 - the customer can assign host addresses for its network users
- Suppose an organization is assigned 128.211.0.16/28
 - the organization will have 4-bits to use as a host address field
- Disadvantage of classless addressing
 - Because the host suffix can start on an arbitrary boundary, values are not easy to read in dotted decimal



CIDR Host Addresses



INDIRIZZI PRIVATI, SPECIALI E INDIRIZZI DEI ROUTER

Non tutti gli indirizzi IP sono utilizzabili, alcuni indirizzi hanno significato solo interno al computer e altri consentono di fare il bootstrap delle macchine prima che abbiano un indirizzo IP con cui comunicare. I Router sono macchine con più indirizzi IP ... anche se non sempre con più interfacce fisiche di comunicazione.



- Non tutti gli indirizzi IP Unicast validi sono uguali
- Alcuni indirizzi sono stati definiti “privati” e non sono instradabili in Internet
 - Possono essere usati per costruire Intra-net private
- Un host con indirizzo IP privato ha bisogno di una apparato attivo che traduca opportunamente i suoi pacchetti per accedere a Internet
- NAT: Network Address Translator
 - Mappa la 5-tupla che identifica un flusso su un'altra 5-tupla con indirizzo pubblico, lavora a livello L3/L4
- Proxy
 - Gateway di L7, che interconnette a livello di singola applicazione



- 10.0.0.0 – 10.255.255.255
- 172.16.0.0 – 172.31.255.255
- 192.168.0.0 – 192.168.255.255

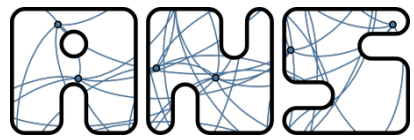
- Un indirizzo privato può essere riutilizzato in molti posti diversi
- All'interno di una stessa rete devono essere unici e possono essere “routati” fino a un “border router” che invece impedisce di andare verso Internet
- Normalmente sono assegnati tramite DHCP e non sono assegnati staticamente a un host
- Non c'è una reale differenza tra i tre gruppi di indirizzi, ma in genere i router “domestici” usano 192.168.x.y/24



- IP defines a set of special address forms that are reserved
 - That is, special addresses are **never assigned to hosts**
- Examples:
 - Network Address
 - Directed Broadcast Address
 - Limited Broadcast Address
 - This Computer Address
 - Loopback Address
 - Multicast addresses



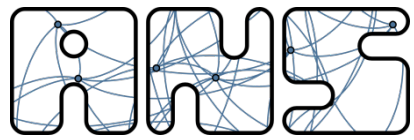
- It is convenient to have an address that can be used to denote the **prefix** assigned to a given network
- IP reserves host address zero
 - and uses it to denote a network
- Thus, the address 128.211.0.16/28 denotes a network
 - because the bits beyond the 28 are zero
 - 10000000 11010011 00000000 00010000
- A network address should never appear as the destination address in a packet



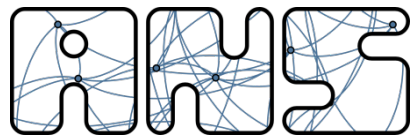
- To simplify broadcasting (send to all)
 - IP defines a directed broadcast address for each physical network
- When a packet is sent to a network's directed broadcast
 - a single copy of the packet travels across the Internet
 - until it reaches the specified network
 - the packet is then delivered to all hosts on the network
- The directed broadcast address for a network is formed by adding a suffix that consists of all 1 bits to the network prefix
 - 10000000 11010011 00000000 00011111



- Limited broadcast refers to a broadcast on a **directly-connected** network:
 - informally, we say that the broadcast is limited to a “single LAN” meaning that it will never be forwarded by a router, even if the “LAN” can be a huge Campus LAN with hundreds of computers
- Limited broadcast is used during system startup
 - by a computer that does not yet know the network number
- IP reserves the address consisting of 32-bits of 1s
 - 11111111 11111111 11111111 11111111
- Thus, IP will broadcast any packet sent to the all-1s address across the local network



- A computer needs to know its IP address
 - before it can send or receive Internet packets
- TCP/IP contains protocols a computer can use to obtain its IP address automatically when the computer boots
 - ... but the startup protocols also use an IP to communicate
- When using such startup protocols
 - a computer cannot supply a correct IP source address
 - To handle such cases IP reserves the address that consists of all 0s to mean this computer
 - 00000000 00000000 00000000 00000000



- Loopback address used to test network applications
 - e.g., for preliminary debugging after a network application has been created
- A programmer must have two application programs that are intended to communicate across a network
- Instead of executing each program on a separate computer
 - the programmer runs both programs on a single computer
 - and instructs them to use a loopback address when communicating
- When one application sends data to another
 - data travels down the protocol stack to the IP software
 - then forwards it back up through the protocol stack to the second program



- IP reserves the network prefix **127/8** for use with loopback
- The host address used with 127 is irrelevant
 - all host addresses are treated the same
 - programmers often use host number 1
 - so it makes **127.0.0.1** the most popular loopback address
- During loopback testing no packets ever leave a computer
 - the IP software forwards packets from one application to another
- The loopback address never appears in a packet traveling across a network



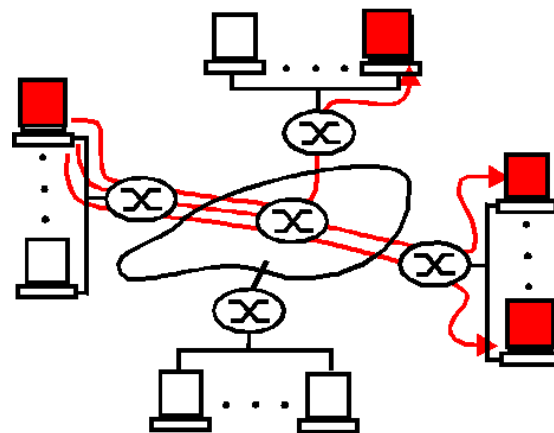
Summary of Special IP Addresses

Prefix	Suffix	Type Of Address	Purpose
all-0s	all-0s	this computer	used during bootstrap
network	all-0s	network	identifies a network
network	all-1s	directed broadcast	broadcast on specified net
all-1s	all-1s	limited broadcast	broadcast on local net
127/8	any	loopback	testing

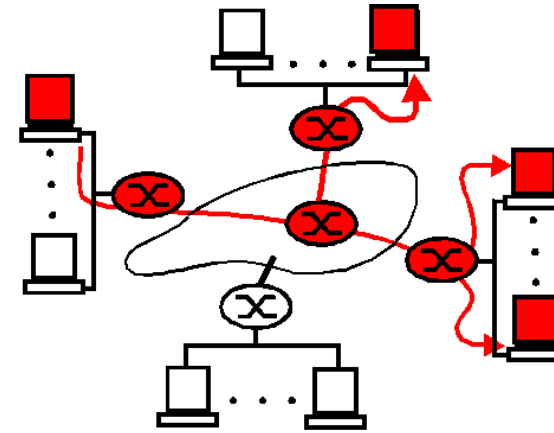


- Internet ammette l'invio di pacchetti a “molti”
- I router si preoccupano (complesso, non lo vediamo in questo corso) di capire il punto ottimo dove duplicare l'informazione
- Un pacchetto multicast è inviato a un indirizzo di “gruppo”
- In IPv4: Class D, iniziano per 1110
- 224.0.0.0 – 239.255.255.255
- Esistono gruppi multicast “well known”
- 224.0.0.1: All Hosts on this Subnet
- 224.0.0.2: All Routers on this Subnet
- Gli altri possono essere usati per applicazioni proprietarie o nuove
- Purtroppo non tutti gli ISP permettono traffico multicast se non per la gestione dei protocolli di routing stessi

- Multicast: delivery of same packet to a group of receivers with the minimum overhead
- Multiple unicast vs. multicast
 - Host based vs. network based



multicast via unicast

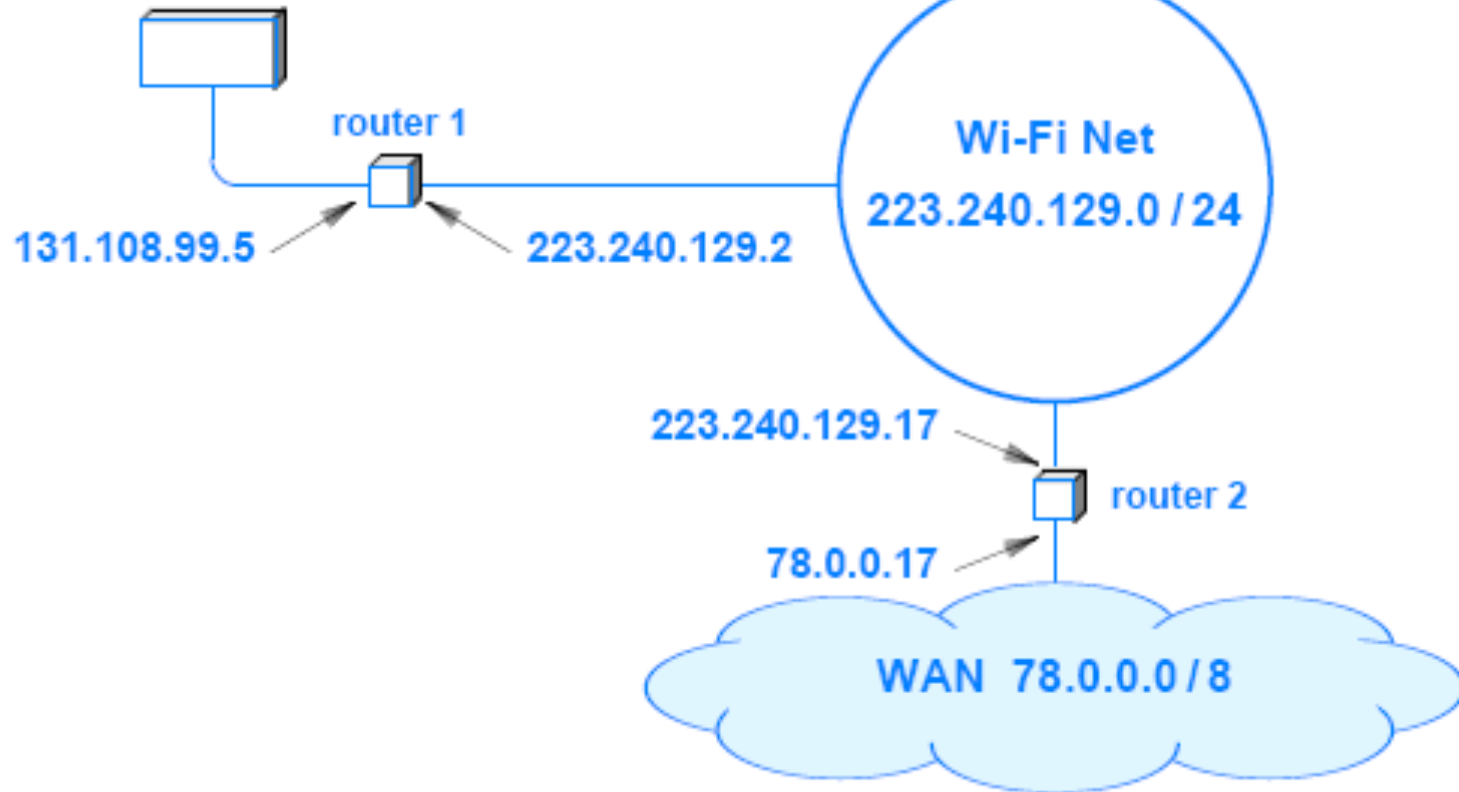


network multicast

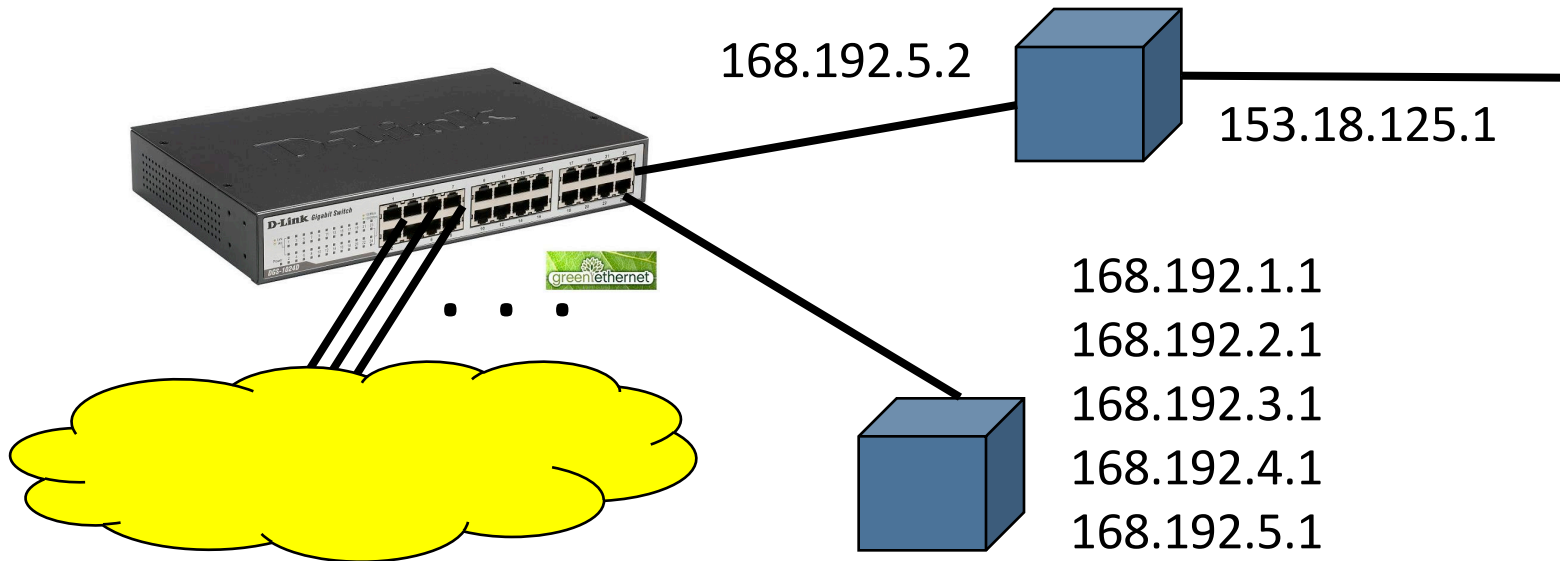


- Each router is assigned two or more IP addresses
 - one address for each logical network to which the router attaches
- To understand why, recall two facts:
 - A router connects multiple IP networks (by definition)
 - Each IP address contains a prefix that specifies a logical network
- A single IP address does not suffice for a router
 - because each router connects to multiple networks
 - and each network has a unique prefix
- The IP scheme can be explained by a principle:
 - An IP address does not identify a specific computer
 - each address identifies in interface, i.e., a logical connection between a computer and a network
 - A computer with multiple network connections (e.g., a router) must be assigned one IP address for each connection

Ethernet 131.108.0.0 / 16



Collegamento a livello ethernet



Router di campus che interconnette le 4 reti ed inoltra al Router NAT di collegamento verso Internet: 5 indirizzi IP tutti sulla stessa interfaccia fisica della campus LAN