Operazione Rif. PA 2022-17295/RER approvata con DGR 1379/2022 del 01/08/2022 finanziata con risorse del Programma Fondo sociale europeo Plus 2021-2027 della Regione Emilia –Romagna.

Progetto n. **1** - Edizione n. **1**

MODULO: N. 6

Titolo: SICUREZZA DEI SISTEMI INFORMATICI

DOCENTE: MARCO PRANDINI

Parte 2 – Architettura di Internet e sicurezza delle reti

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Introduzione alle reti di calcolatori

-

Marco Prandini
marco.prandini@unibo.it

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

A cosa serve una rete

Dove viene utilizzata l'informatica esistono:

- **Una molteplicità di risorse**
  - **di calcolo**
  - **di memorizzazione**
  - **per l'ingresso e l'uscita dei dati**

  associate a stazioni di lavoro diverse
- **Una continua necessità**
  - **di accedere alle risorse**
  - **di fornire o recuperare informazioni**

  indipendentemente dalla stazione su cui si lavora

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Connessione di calcolatori

Idea di fondo: far "parlare" tra loro due calcolatori collegandoli elettricamente.

- **Vantaggi rispetto allo spostamento fisico dei dati:**
  - **Capacità (sempre?)**
  - **Trasparenza**
  - **Robustezza**
  - **Interattività**
- **Problemi da risolvere:**
  - **costruire l'infrastruttura**
  - **concordare un linguaggio comune**
  - **proteggere il sistema dalle intrusioni**

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

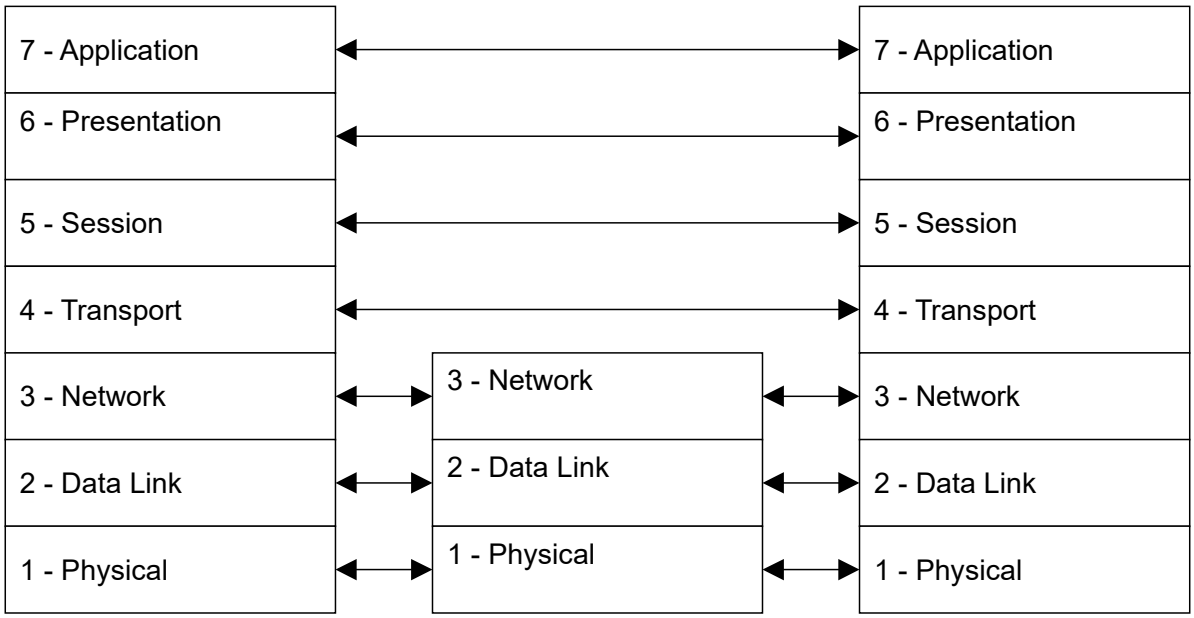
Stabilire le regole del dialogo

Perchè i calcolatori connessi si capiscano, è necessario stabilire uno standard sul formato fisico dei segnali, la temporizzazione dei messaggi, ed il significato dei simboli trasmessi dall'uno all'altro.

Se affrontato in modo monolitico, il problema risulta complesso e le soluzioni non generalizzabili.

Per questo motivo, l'ISO ha formalizzato il modello OSI, che rappresenta la comunicazione come una pila (stack) di livelli, ognuno deputato a svolgere un ruolo preciso senza che i dettagli implementativi siano noti agli altri livelli.

Modello ISO/OSI



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Strati 1-2-3

- **Lo strato 1 (Physical) definisce le caratteristiche elettriche e meccaniche dei mezzi di trasporto delle informazioni.**
- **Lo strato 2 (Data Link) gestisce le comunicazioni tra sistemi connessi alla stessa rete fisica. Esegue il controllo d'errore sui dati trasportati dal livello fisico, definisce i metodi per identificare un sistema all'interno una rete locale e le politiche di accesso dei sistemi al livello fisico.**
- **Lo strato 3 (Network) si occupa dell'*instradamento (routing)*, ovvero della ricerca del percorso per connettere due sistemi che vogliono comunicare tra loro anche se sono su reti fisiche distinte. I metodi per identificare un sistema su scala globale (come gli indirizzi IP usati in Internet) sono definiti in questo strato.**

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Strati 4-5-6-7

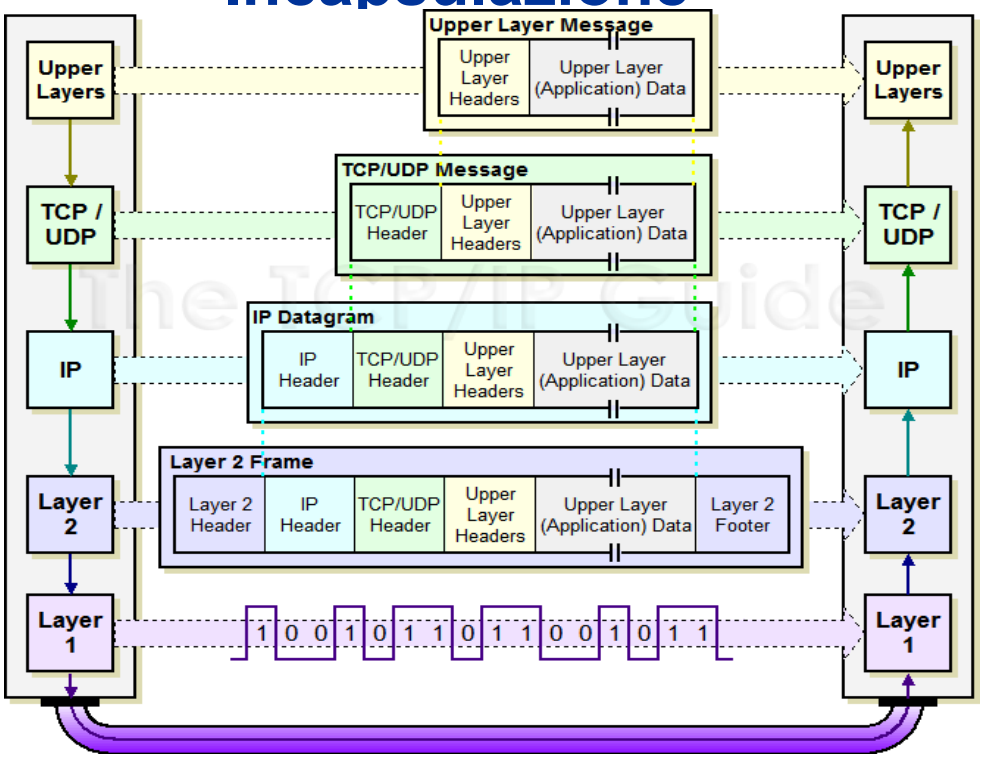
- Lo strato 4 (Transport) si occupa del trasporto dei dati *end-to-end*. Fornisce allo strato superiore canali affidabili tra i due estremi della comunicazione, occupandosi del controllo di flusso, della correzione degli errori, della divisione e ricomposizione dei dati. Nasconde la frammentazione dei percorsi "reali" che gli sono forniti dallo strato 3.
- Gli strati 5 e 6 (Session-Presentation) sono raramente implementati, e modellano gli strumenti quali le procedure di negoziazione della localizzazione, i controlli di integrità e validità dei documenti, e la sincronizzazione tra applicazioni interattive.
- Lo strato 7 (Application) è quello in cui si collocano le applicazioni che originano e ricevono i dati oggetto della comunicazione

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Percorso logico e fisico dei dati

- Fisicamente un messaggio originato da un'applicazione (strato 7) percorre il sistema di origine scendendo di strato in strato, attraverso le *interfacce*.
  - Ogni strato sottopone il messaggio alle elaborazioni necessarie per lo svolgimento del proprio compito, lo *incapsula* in un pacchetto arricchito di metadati in modo che il proprio omologo nel sistema di destinazione possa effettuare le operazioni inverse, e lo passa allo strato inferiore.
- Logicamente, durante l'interazione di due sistemi avviene uno scambio di messaggi tra ciascuno degli strati omologhi, secondo le regole stabilite da un *protocollo*.
  - Il vantaggio della struttura a strati risiede nella possibilità di scegliere qualsiasi protocollo per uno strato senza influenzare quelli circostanti. (Es. passaggio di Microsoft a TCP/IP)

Incapsulazione



Reti Locali

- Una rete locale può essere definita come un insieme di calcolatori connessi da un'infrastruttura tale per cui ogni partecipante ha visibilità fisica diretta di tutti gli altri.
  - Infrastruttura basata sui soli strati 1 e 2
  - Caratterizzate da
    - *Topologia (bus, anello, stella)*
    - *Mezzo trasmissivo (coassiale, doppino, fibra, etere, ...)*
    - *Politica di accesso (token-based, collisione, ...)*
- L'identificazione di ogni nodo avviene per mezzo di un indirizzo che può avere caratteristche arbitrarie, perchè per definizione deve essere raggiungibile senza istruire intermediari --> problemi di interconnessione

Indirizzo globale e indirizzo locale

- Indirizzo globale
  - È valido per tutta la rete
  - Deve essere **univoco** (non devono esistere indirizzi replicati) per evitare ambiguità
  - **Va "assegnato" seguendo una procedura di gestione "globale" che assicura la non replicazione**

- Indirizzo locale
  - È valido limitatamente ad un certa sottoporzione della rete
    - Internamente ad un terminale
    - In un dominio di rete specifico
  - Può non essere globalmente univoco
  - Può essere assegnato con una procedura puramente "locale"

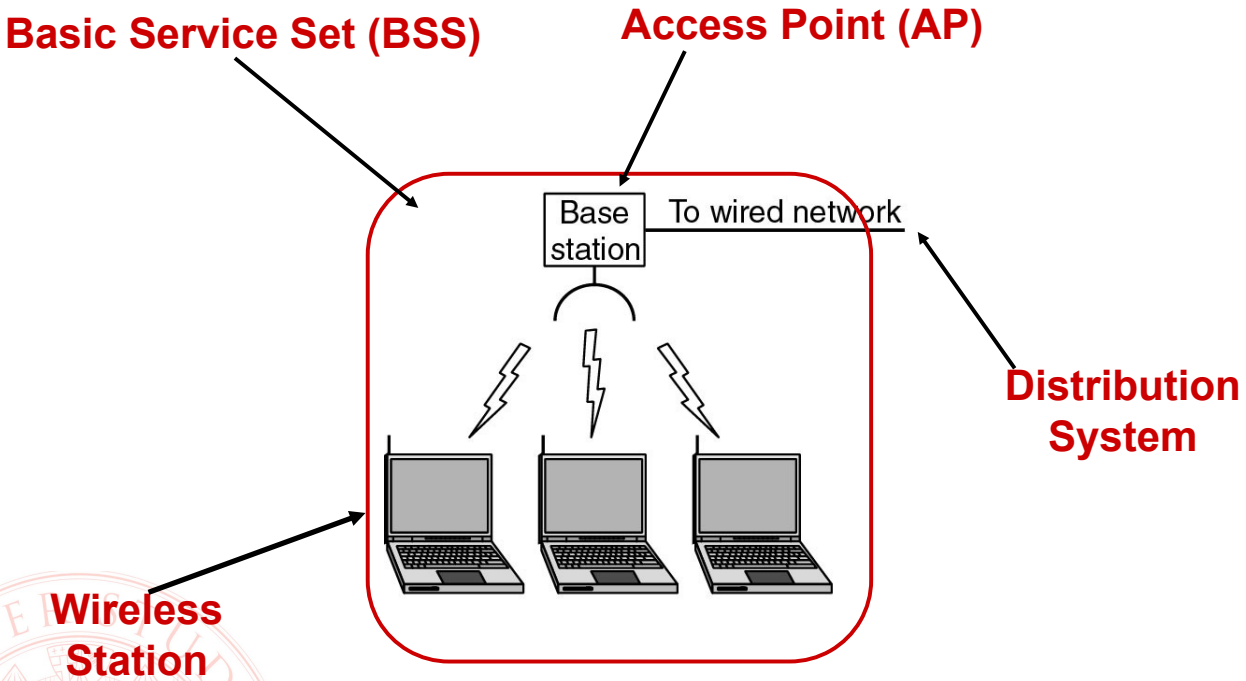
Rete logica e rete fisica

- Nella terminologia di Internet si definisce
  - **Rete logica**: la network IP (o **subnet**) a cui un Host appartiene logicamente
  - **Rete fisica**: la rete (tipicamente **LAN**) a cui un Host è effettivamente connesso
- La rete fisica normalmente ha capacità di instradamento e può avere indirizzi locali (es. indirizzi MAC)
- L'architettura a strati nasconde gli indirizzi fisici e consente alle applicazioni di lavorare solo con indirizzi IP

La tecnologia

- Ogni network IP può essere implementata con una tecnologia specifica
- Esempio
  - **Wi-Fi** : Network realizzata con tecnologia wireless in area locale
  - **ADSL e xDSL**: Network realizzata con tecnologia a media distanza via cavo tramite infrastruttura di uno specifico fornitore di servizio pubblico
  - **Ethernet**: Network realizzata con tecnologia a breve distanza via cavo privata in area locale
  - **GPRS/EDGE/LTE**: Network realizzata con tecnologia radio a media distanza tramite infrastruttura di uno specifico fornitore di servizio pubblico

Tipica rete wireless: Architettura di rete 802.11



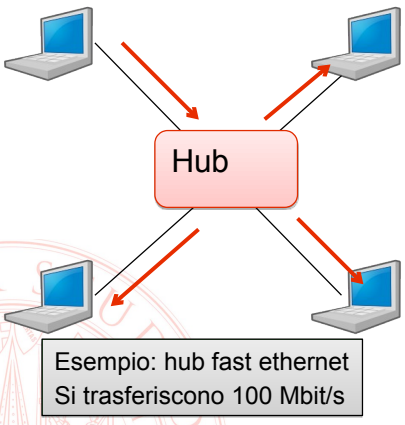
Da A.S. Tanenbaum, "Reti di Calcolatori"

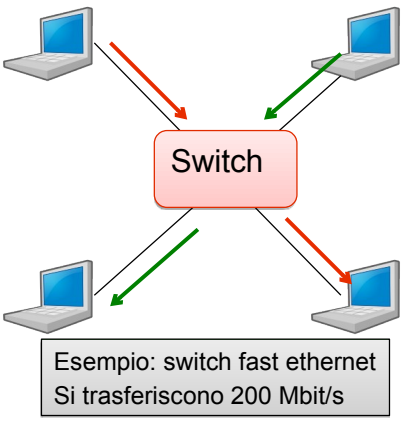
Tipica rete cablata: Ethernet basata su SWITCH

- Un bridge è un ponte tra due diverse LAN
- Un bridge tra più di due LAN (ma tipicamente della stessa tecnologia) è denominato HUB
  - Tipicamente ad ogni porta è connessa una sola stazione
- Uno switch Ethernet svolge una funzione simile all'hub ma garantendo maggiori prestazioni
  - È in grado di trasferire contemporaneamente trame da più porte di ingresso a più porte di uscita
  - Opera una funzione di commutazione a livello 2 basata sull'indirizzo MAC

Differenza fra hub e switch

- Hub
  - bus collassato = mezzo condiviso, trasmissione broadcast delle trame
  - Capacità aggregata = capacità della singola porta
- Switch
  - Sistema di commutazione = ri-trasmissione selettiva delle trame
  - Capacità aggregata superiore a quella della singola porta





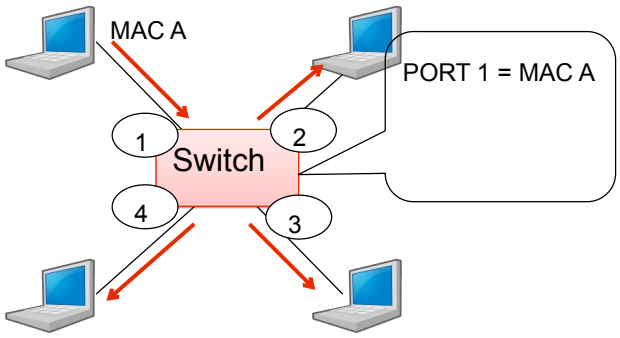
Switch come learning bridge

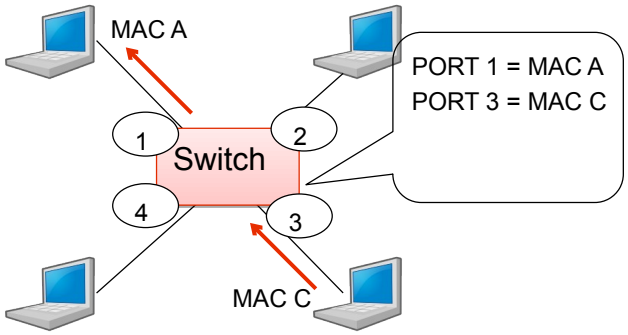
- Lo Switch costruisce una “Tabella di inoltro”
  - Associa il Mac Address delle interfacce alla porta dello switch su cui si trova collegata l'interfaccia stessa
  - Se alla porta è connesso un altro switch, da essa si raggiungono molteplici MAC di destinazione
- Esempio switch a 8 porte → implementazione CAM (content addressable mem)

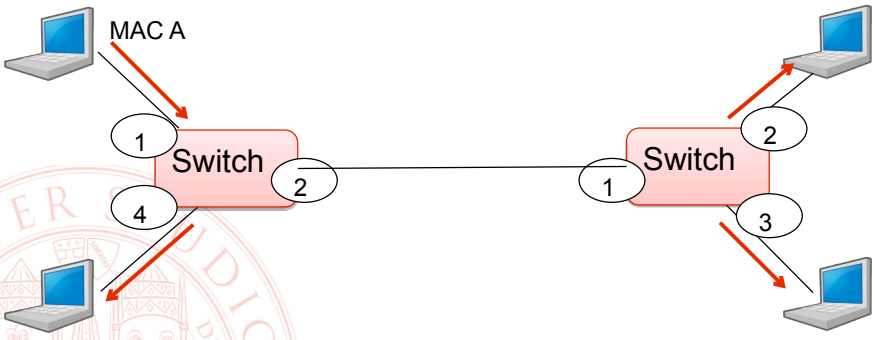
Porta	Lista MAC raggiungibili
1	eb:a6:99:de:1c:b0 2c:65:1e:b1:9f:44
2	
3	0c:2e:22:b0:8e:16
4	
5	
6	5b:06:72:1b:3c:03 e4:b0:56:d5:2d:0f 92:ff:9e:6c:b0:8e
7	
8	

Porta	MAC raggiungibile
1	eb:a6:99:de:1c:b0
1	2c:65:1e:b1:9f:44
3	0c:2e:22:b0:8e:16
6	5b:06:72:1b:3c:03
6	e4:b0:56:d5:2d:0f
6	92:ff:9e:6c:b0:8e

Learning Switch







Internetworking

- Usando il solo livello 2 ...
  - Come garantire l'univocità degli indirizzi a livello globale?
  - Come limitare la propagazione del traffico?
- Usando il livello 3 si sovrappone uno schema di indirizzi configurabili all'indirizzamento di livello 2
  - Gli indirizzi sono concessi, sia numericamente che organizzativamente, secondo una gerarchia
  - La contiguità numerica degli indirizzi rispecchia quella fisica dei corrispondenti nodi della rete
- Da questo momento non tratteremo più concetti generali ma faremo riferimento allo standard mondiale di fatto: TCP/IP

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Internet

EnAIP Tecnico Informatico

Origini di TCP/IP

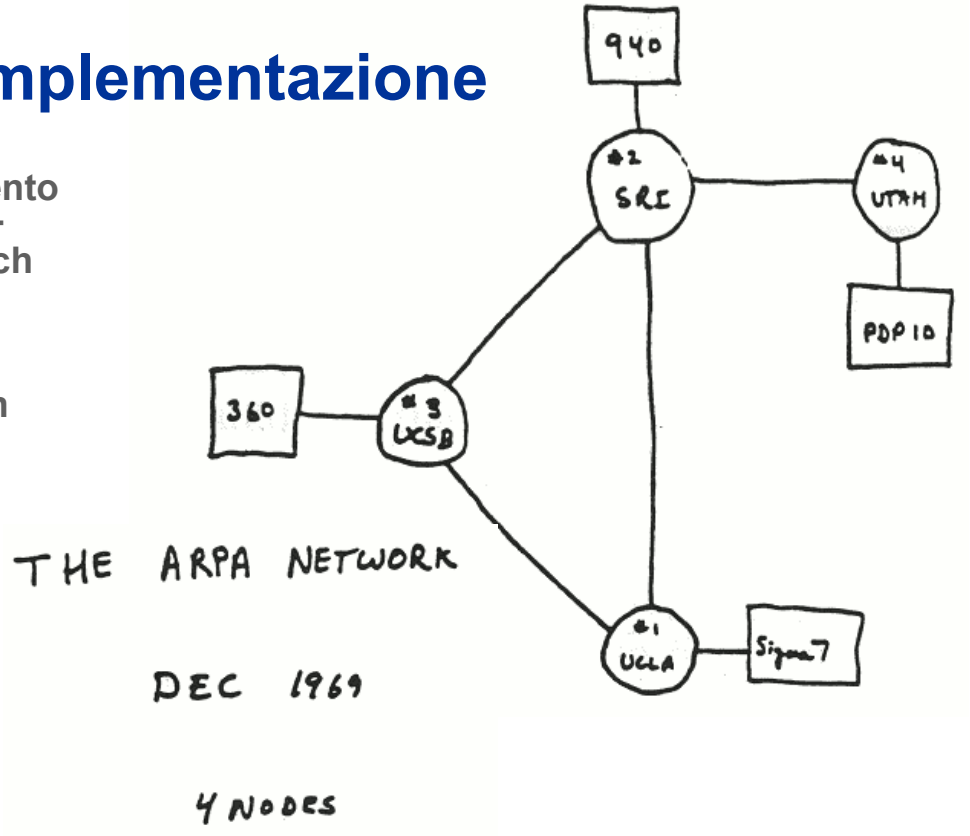
Le radici di Internet affondano nella Guerra Fredda.

- **Problema:**
  - **Un sistema di comunicazione a prova di attacco nucleare.**
- **Soluzione (RAND Corporation, 1964):**
  - **Una *rete*, anzichè un albero di interconnessioni, modello tradizionalmente utilizzato nelle telecomunicazioni**
  - **Moltitudine di nodi nessuno dei quali abbia un "ruolo di comando"**
  - **Moltitudine di connessioni tra i nodi (ridondanza dei percorsi)**
  - **Traffico suddiviso in pacchetti**
  - **La funzione di ogni nodo è il *packet forwarding***

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

La prima implementazione

- **DARPA: Dipartimento della Difesa USA + Advanced Research Projects Agency**
- **BBN**
- **RAND Corporation**



EnAIP Tecnico Informatico

Alcune tappe fondamentali

- 1972 Invenzione della posta elettronica.
- 1973 Primi nodi internazionali (UK, Norvegia).
  Vinton Cerf pubblica TCP.
- 1983 Smilitarizzazione.
  Primo uso del termine Internet.
  Introduzione dei nomi.
- 1986 NSF collega i propri centri e le università al di fuori di ARPAnet (che ora conta 1000 nodi) usando TCP/IP.
- 1990 ARPAnet viene dismessa.
  I nodi TCP/IP superano le 100.000 unità.
- 1992 Nascita del World Wide Web.
  Più di 1 milione di nodi.
- 1993 Primo browser grafico (NCSA Mosaic)
  Il tasso di crescita di Internet sfiora il 350% annuo.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Organismi, standard, documentazione

L'impulso all'evoluzione di Internet è stato dato da una forma di collaborazione estremamente libera. I documenti fondamentali (Request For Comments – RFC) prodotti dal Network Working Group seguivano queste "regole" (RFC 3):

«Il contenuto di una nota del NWG può essere qualsiasi riflessione, suggerimento, o altro soggetto relativo al software per host o ad aspetti diversi della rete. È incoraggiata la sottomissione tempestiva delle note piuttosto che la cura della forma. Questioni teoriche prive di esempi o applicazioni immediate, suggerimenti specifici, tecniche di implementazione prive di spiegazioni introduttive e di contesto, domande esplicite prive di qualunque tentativo di risposta sono tutti argomenti accettabili. La lunghezza minima per una nota del NWG è di una frase.»

Organismi, standard, documentazione

- La RFC 1 “Host Software” è del 7 aprile 1969. Ora sono oltre 4400.
- Attualmente, esistono diversi organismi che regolano i vari aspetti di Internet, dalla ricerca avanzata ai dettagli applicativi:
  - IAB (Internet Architecture Board) svolge il coordinamento e decide le strategie a lungo termine
  - IRTF (Internet Research Task Force) si occupa della ricerca di base, con progetti a lungo termine
  - IETF (Internet Engineering Task Force) si occupa degli aspetti tecnici, e produce gli Internet Drafts che precorrono le RFC.
  - IANA (Internet Assigned Numbers Authority) si occupa della registrazione degli IP, degli host, e dei servizi TCP e UDP.
- Il coordinamento, le attività di promozione, finanziamento e formazione sono svolte dall’Internet Society.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

TCP/IP - generalità

La sigla TCP/IP indica l’insieme dei protocolli usati in ambito Internet.

Alcune caratteristiche:

- È uno standard aperto, slegato da qualsiasi produttore
- È definito ufficialmente dall’IETF nelle Request For Comments
- È indipendente dall’hardware
- Svolge principalmente le funzioni proprie degli strati 3 e 4:
  - Prevede una modalità di indirizzamento globale
  - Supporta diverse modalità di routing
  - È basato sulla commutazione di pacchetto
- Fornisce agli strati superiori connessioni multiple e di diversi tipi

TCP/IP – schema a livelli

Lo stack TCP/IP è più semplice di quello OSI, e prevede solo 4 livelli

Application Layer	**(~ 5,6,7)**	**Protocolli di posta, trasferimento file, web, condivisione dischi, ...**
Transport Layer	**(~ 4)**	**Servizi di multiplazione e di garanzia della completezza dei dati**
Internet Layer	**(~ 3)**	**Indirizzamento dei nodi ed instradamento dei pacchetti**
Network Access Layer	**(~ 1,2)**	**Mappatura degli indirizzi Internet sulla LAN ed accesso al mezzo**

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Internet Protocol

IP (*Internet Protocol*) è il protocollo che svolge le funzioni classificate come strato 3 dal modello OSI; fornisce un servizio di trasporto dati attraverso un'internet, mediante il meccanismo della commutazione di pacchetto:

- **tutti i dati consegnati ad IP dallo strato superiore vengono frammentati in unità di dimensione massima prefissata (*pacchetti*)**
- **per ogni singolo pacchetto viene individuato un percorso che porti dal sistema sorgente al sistema di destinazione, eventualmente attraverso sistemi intermedi**

IP è un protocollo di tipo *best effort*, ovvero non dà garanzie:

- **sull'effettivo arrivo di un pacchetto**
- **sull'ordine di arrivo dei pacchetti**
- **sulla velocità di trasmissione o sui tempi di percorrenza**

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Indirizzi IP

- **Associazione univoca indirizzo –> sistema (*host*)**
  - **(non è imposto il contrario –> *multi-homed host*)**
- **Indirizzi IPv4: 32 bit divisi in 4 byte, normalmente rappresentati con 4 numeri in base 10 separati da punti (*dotted decimal notation*)**
  - **Esempio: 137.204.59.1**
- **Ogni indirizzo fa parte di una rete *(subnet)* che inizia da un indirizzo di *network***
- **In origine l' estensione era implicita, ora è specificata da una *netmask***
  - **Una subnet logica coincide con una LAN fisica**
  - **Per instradare un pacchetto verso la destinazione non serve considerare il suo indirizzo, basta la subnet cui appartiene**

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Reti class-based

Originariamente sono state definite *classi* di indirizzi, ovvero raggruppamenti da usare per i sistemi di una rete locale, in cui i byte erano rigidamente divisi tra network id e host id:

- **128 reti di Classe A (contenenti fino a 16 milioni di host circa):**
  - **indirizzi da 0.*.*.* a 127.*.*.***
- **16.384 reti di Classe B (contenenti fino a 65000 host circa):**
  - **indirizzi da 128.0.*.* a 191.255.*.***
- **2.097.152 reti di Classe C (contenenti fino a 254 host):**
  - **indirizzi da 192.0.0.* a 223.255.255.***
- **Gli indirizzi da 224.*.*.* a 239.*.*.* sono riservati al *multicast***
- **Gli indirizzi da 240.*.*.* a 255.*.*.* sono riservati a usi futuri**

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Il vantaggio sulle LAN

- Host su una stessa rete locale:
  - collocati in una stessa area, servita da un determinato apparato di rete
  - numericamente identificati da indirizzi di una stessa subnet

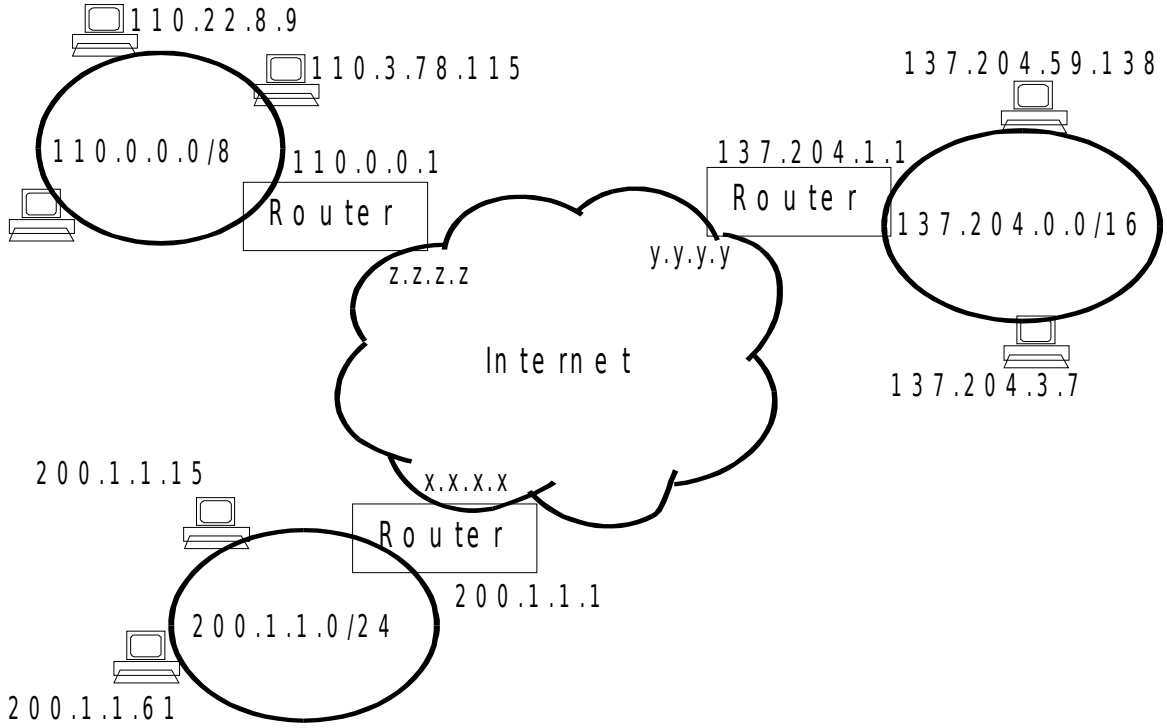
- Per raggiungerli non ho bisogno di sapere dove si trova ognuno, mi basta sapere come raggiungere la subnet

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Instradamento del traffico

- Ogni host di una rete locale IP “sa” quali altri host può raggiungere direttamente, grazie a network address e netmask.
- Ogni host può essere configurato per raggiungere destinazioni al di fuori della rete locale usando degli intermediari: i *router*
  - Un router per instradare traffico tra LAN diverse deve avere un'interfaccia connessa ad ognuna, a cui deve essere associato un indirizzo della corrispondente subnet.
  - Più in generale, un router sa come raggiungere *ogni* host/subnet, direttamente o attraverso altri router
  - L'elenco delle destinazioni non è conservato da ogni router in modo estensivo, ma tipicamente contiene l'indicazione di un router a cui delegare l'instradamento verso tutte le subnet non esplicitamente note: il *default gateway*.

Internetworking con TCP/IP



CIDR

- Poche classi di dimensioni fissate = spreco di indirizzi
- Soluzione: CIDR (*classless inter-domain routing*)
- Con classless-IP gli indirizzi sono visti come una stringa di 32 bit divisa in net-id e host-id in un punto arbitrario, anzichè per byte.
  - Unico vincolo (ovvio): la dimensione di una rete è potenza di 2 (in questo esempio 2^6 = 64 indirizzi)

144 . 156 . 166 . 151

10010000 10011100 10100110 10010111

26 bit net-id 6 bit host-id

- Una rete locale è identificata per mezzo di un *network address* e di una *netmask*, noti a tutti gli host che ne fanno parte.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Netmask, network, broadcast

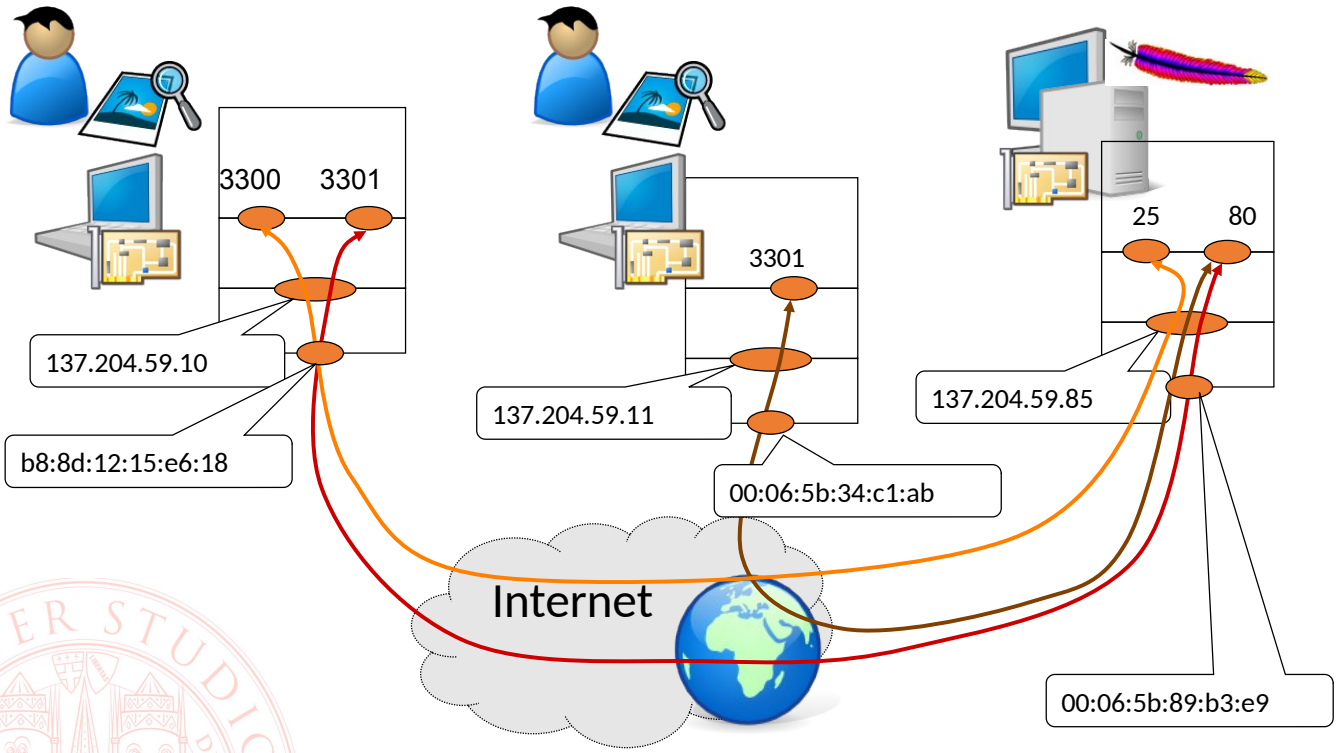
- Serve un modo per specificare dove cade la divisione: la *netmask* è un valore di 32 bit composto da tanti “1” quanti sono i bit che identificano la subnet, e tanti “0” quanti sono i bit che specificano l'host al suo interno
  - Nell'esempio precedente:
    11111111.11111111.11111111.11000000 = 255.255.255.192
- Due valori in ogni subnet hanno un significato speciale e non possono essere usati per indirizzare un host:
  - Network address (ottenuto mettendo a 0 tutti i bit dell’host-id)
    - 10010000.10011100.10100110.10000000 = 144.156.166.128
  - Broadcast (ottenuto mettendo a 1 tutti i bit dell’host-id)
    - 10010000.10011100.10100110.10111111 = 144.156.166.191

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Netmask, network, broadcast... perchè?

- Il sistema è stato pensato per velocizzare il lavoro degli apparati di instradamento
  - ogni pacchetto contiene un indirizzo di destinazione
  - deve essere inviato verso la subnet che contiene tale indirizzo
  - i router conoscono il modo di raggiungere le subnet per mezzo di una tabella che elenca (network, netmask, interfaccia da usare)
  - l'operazione di AND bit-a-bit tra un indirizzo e una netmask restituisce il valore da confrontare con il nework
- Facile costruzione di gerarchie di subnet, per diminuire ulteriormente il numero di regole di instradamento memorizzata su ogni router

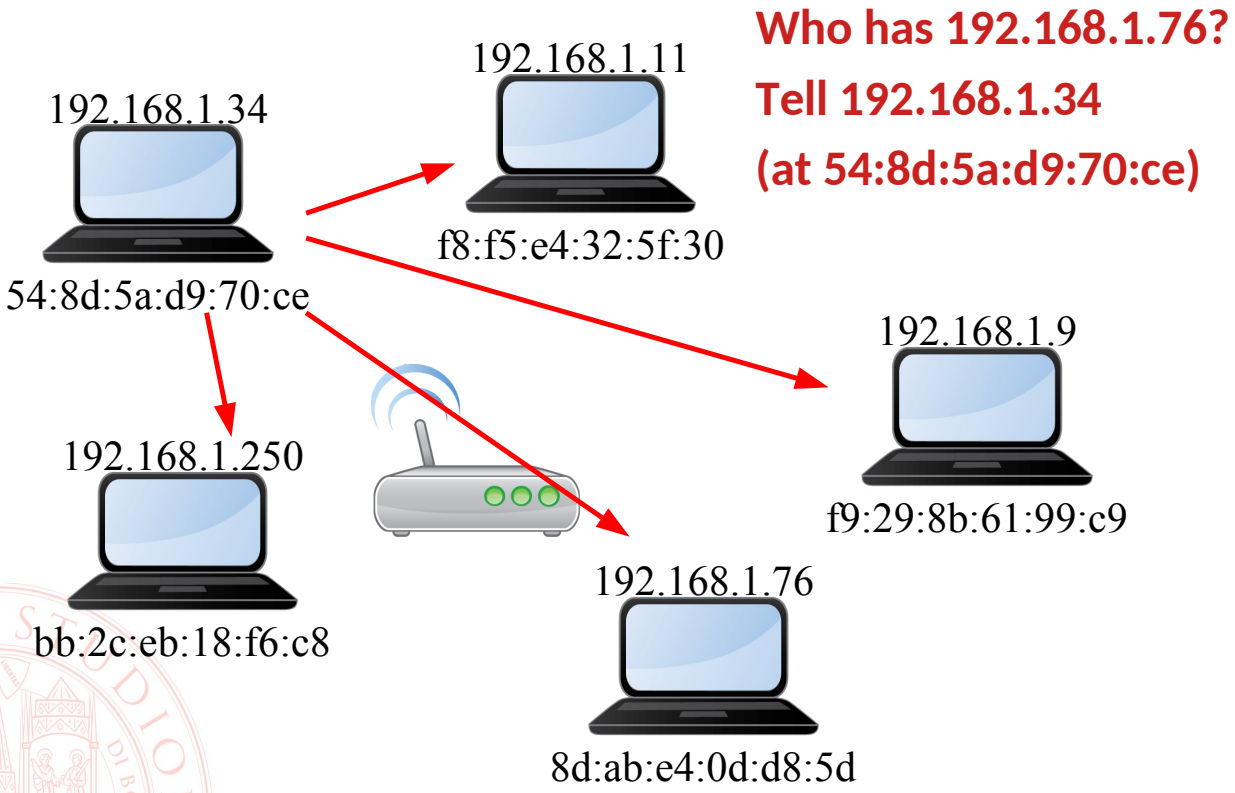
Flussi di comunicazione



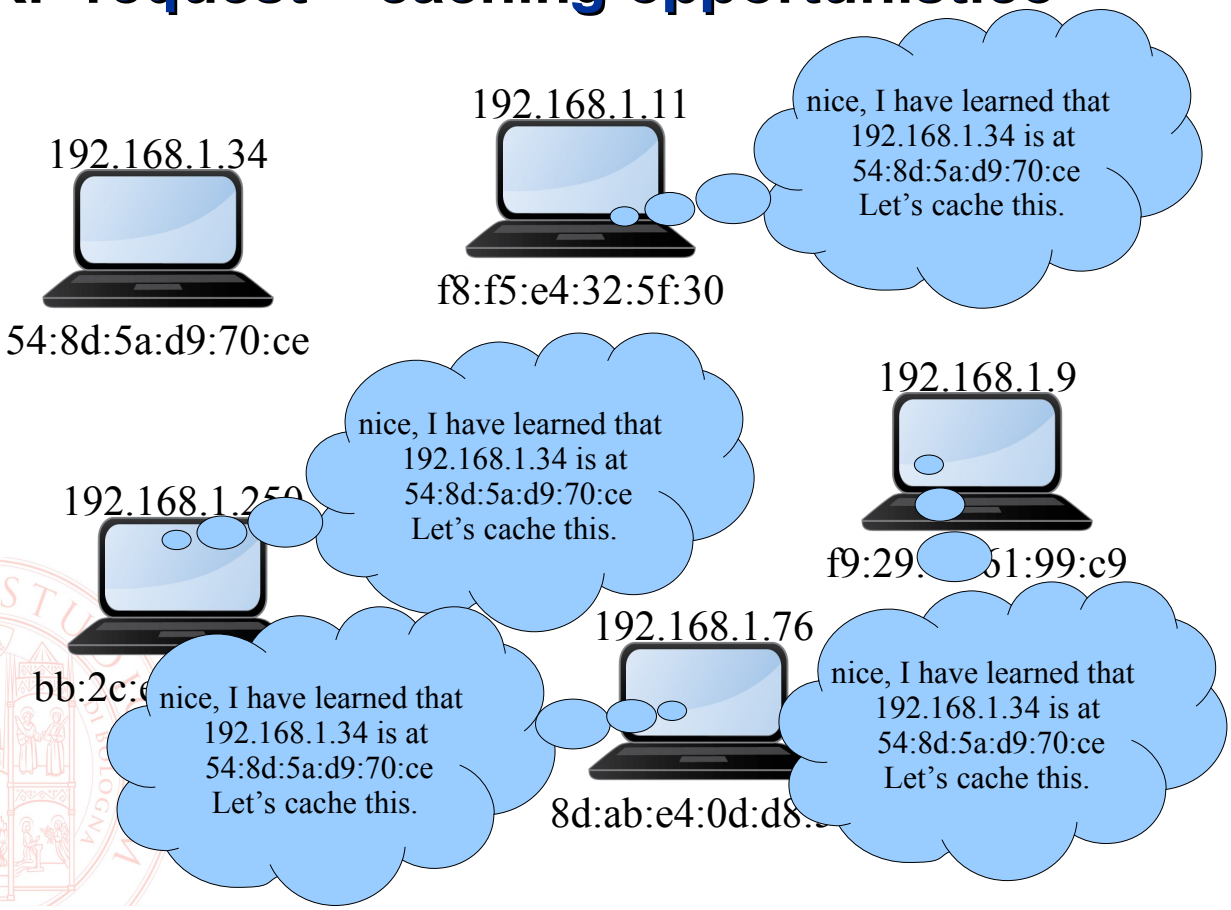
Doppio indirizzamento nella LAN/subnet!

- Ogni dispositivo della LAN ha un MAC address
  - L'inoltro fisico del traffico avviene tra le schede di rete
- Ma è anche un dispositivo della rete IP con un indirizzo
  - Le applicazioni si "conoscono" come endpoint IP
- Come tradurre un indirizzo nell'altro?
- Address Resolution Protocol (ARP – RFC 826)

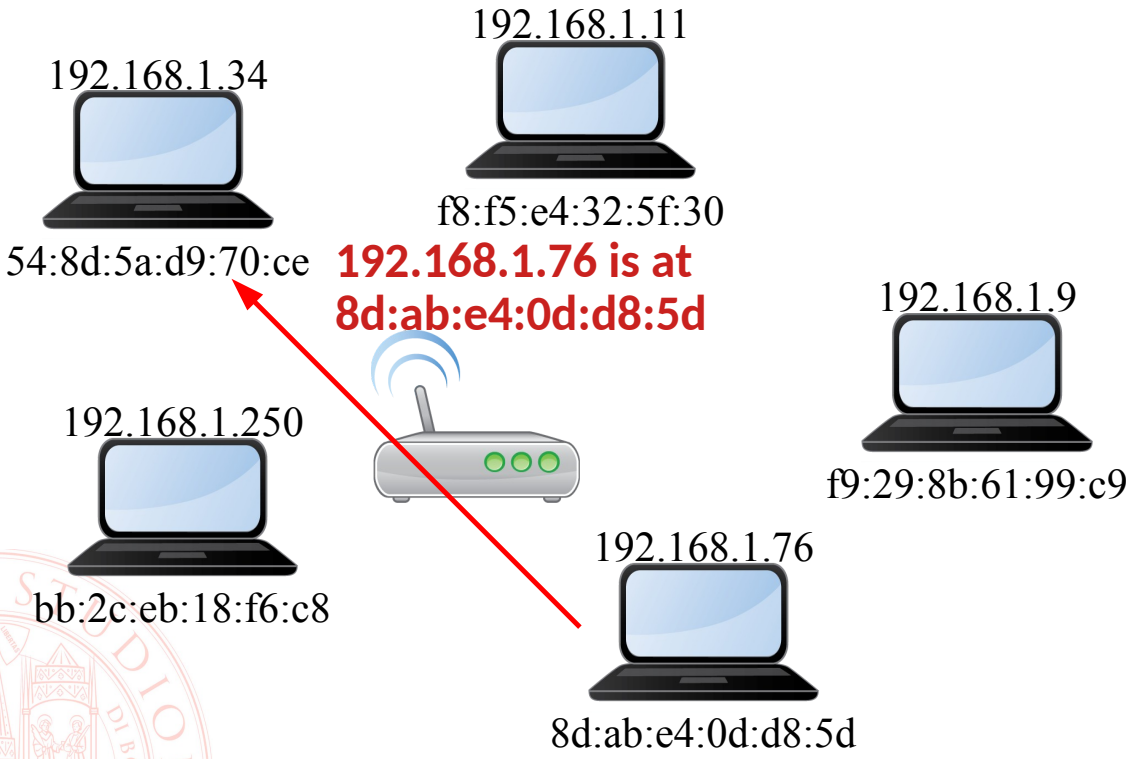
ARP request - broadcast



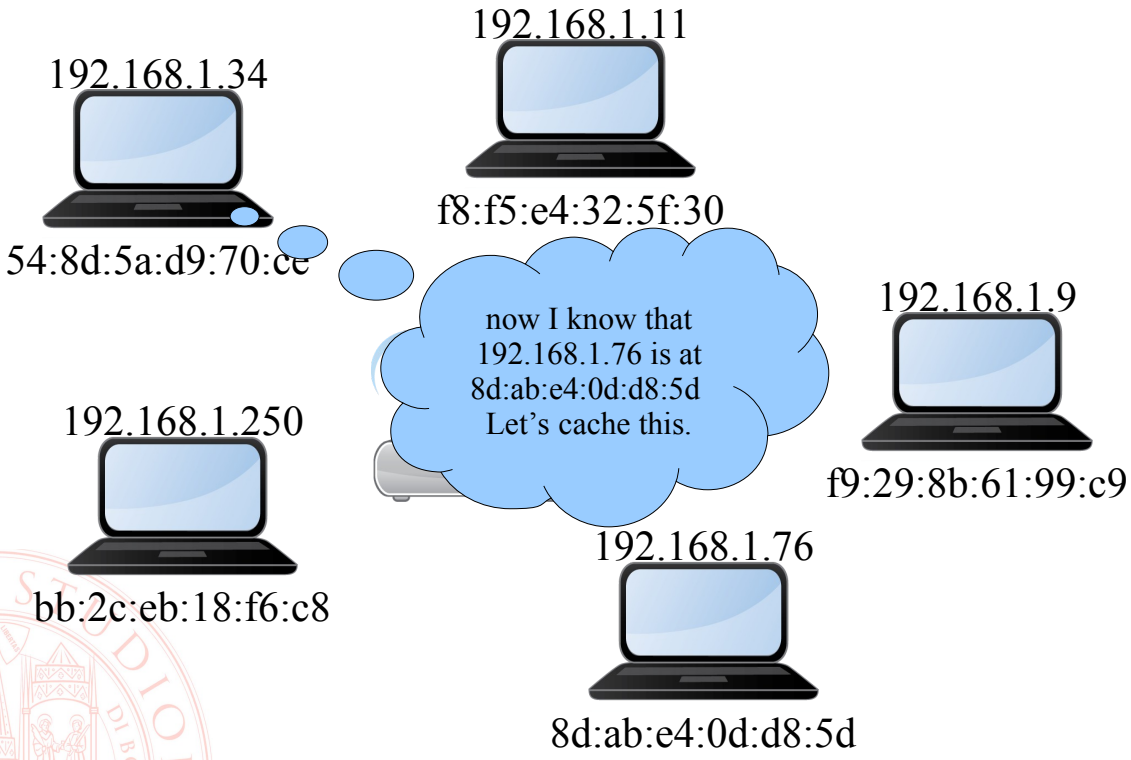
ARP request – caching opportunistico



ARP reply - unicast



ARP reply - unicast



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Il livello di trasporto

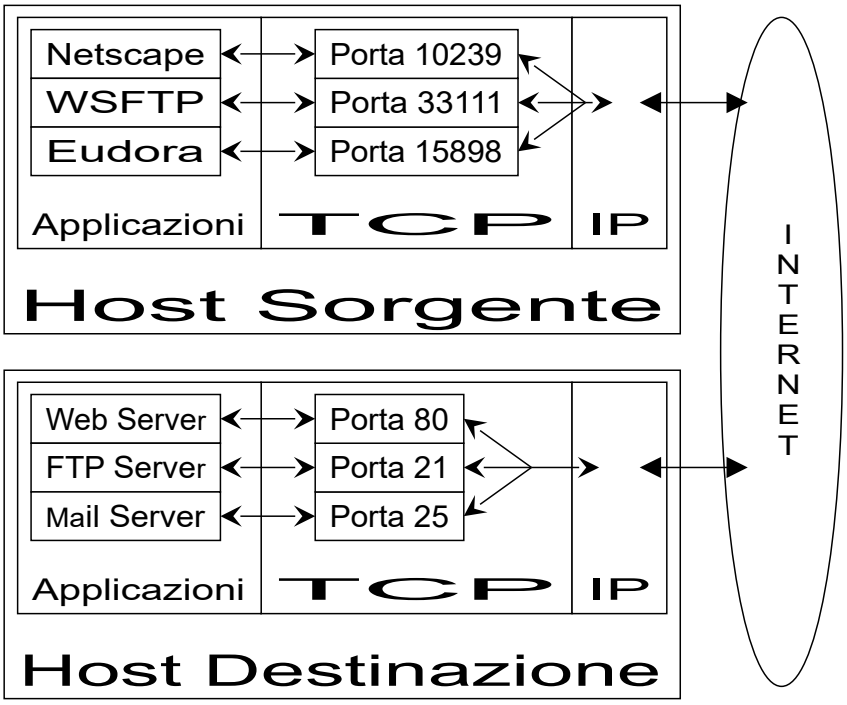
- TCP = Transmission Control Protocol
- UDP = User Datagram Protocol
- Le funzioni svolte da questi protocolli sono classificabili nell'ambito dello strato 4 del modello OSI. Conseguentemente, operano su di un canale di comunicazione dei dati end-to-end, fornito da IP che nasconde i dettagli della gestione dei pacchetti.
  - IP non permette di distinguere diverse applicazioni sorgente o destinazione dei pacchetti su di uno stesso host
  - IP non fornisce garanzie sulla consegna dei pacchetti

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Multiplexing

- Il multiplexing è la funzione svolta sia da TCP che da UDP è per consentire l'indirizzamento di applicazioni specifiche all'interno di un host
  - IP fornisce un canale trasmissivo tra due host, TCP ed UDP permettono a più applicazioni di usare questo canale contemporaneamente-
  - Questo servizio viene offerto associando ogni applicazione ad una *porta*.
    - Una porta è indicata da un numero compreso tra 0 e 65535.
    - In ogni istante, su di un host, ciascuna porta non può essere utilizzata da più di un'applicazione.
- Per raggiungere un'applicazione su di un host, bisogna conoscere la porta associata. Per questo i servizi di ampio interesse sono collocati su porte standard, dette *well-known ports*

Multiplexing



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

TCP

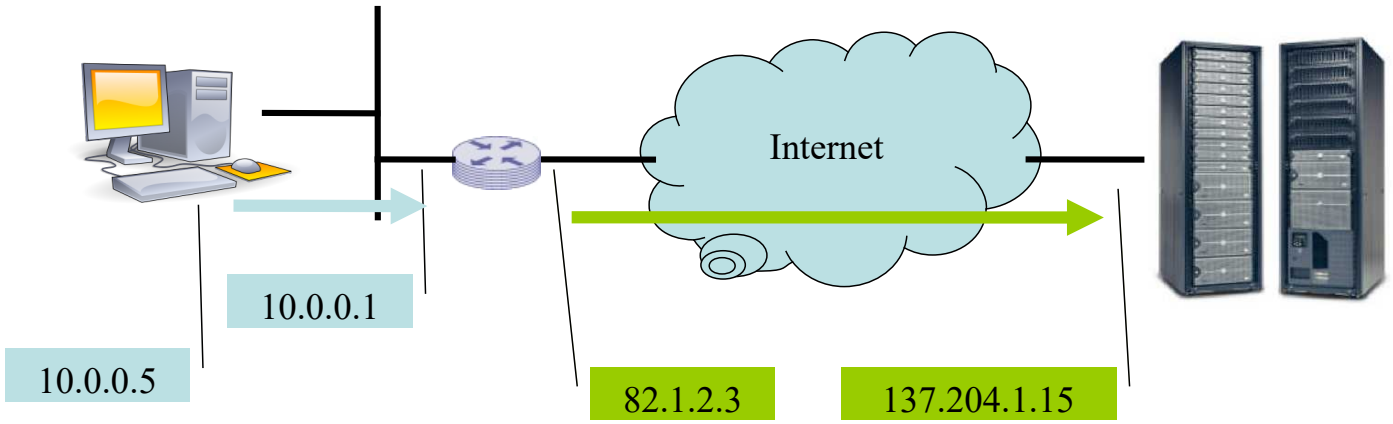
- UDP è un protocollo connection-less, ovvero i messaggi vengono inviati senza accordi preliminari, di conseguenza è molto efficiente per l'invio di piccole quantità di dati la cui perdita non è critica (es.: NTP, domain)
- TCP è un protocollo connection-oriented, ovvero:
  - ogni scambio di messaggi è preceduto da una fase preliminare in cui le due parti concordano di voler comunicare tra loro, che permette di marcare i pacchetti come appartenenti ad una precisa connessione e di numerarli in sequenza
  - di conseguenza:
    - richiede un tempo iniziale aggiuntivo per stabilire la connessione
    - è affidabile perché rileva la perdita di pacchetti
    - esegue il riordino dei pacchetti che arrivano fuori sequenza

Network Address Translation (NAT)

- La prospettiva di un esaurimento degli IP disponibili ha fatto esplodere l'utilizzo di una tecnica che consente di utilizzare un solo indirizzo pubblico, senza rinunciare alla possibilità di realizzare in tecnologia TCP/IP una rete anche di grandi dimensioni e di connetterla ad Internet.
- L'efficacia della tecnica si basa sull'osservazione che la gran parte degli host è client e non server
  - non necessitano di essere raggiunti da richieste
  - originano richieste e devono poter essere raggiunti dalle risposte

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Network Address Translation (NAT)



Richiesta		
Source	10.0.0.5	: 34567
Destination	137.204.1.15	: 80

Richiesta traslata		
Source	82.1.2.3	: 34567
Destination	137.204.1.15	: 80

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Network Address Translation (NAT)

- La quintupla
  (*protocollo, ip_sorgente, porta_sorgente, ip_destinazione, porta_destinazione*)
  identifica univocamente una connessione
- Nel NAT molti IP sorgente vengono sostituiti dall'unico IP pubblico del router
  - possibilità di modificare la porta sorgente per disambiguare le connessioni originate con tutti i parametri identici a parte l'IP sorgente
  - memorizzazione delle traslazioni per poter riconoscere il destinatario delle risposte

Source IP	Source port	Router IP	Router port	Dest. IP	Dest. port
10.0.0.5	11111	82.1.2.3	11111	137.204.1.15	80
10.0.0.9	11111	82.1.2.3	11111	137.204.1.15	80

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

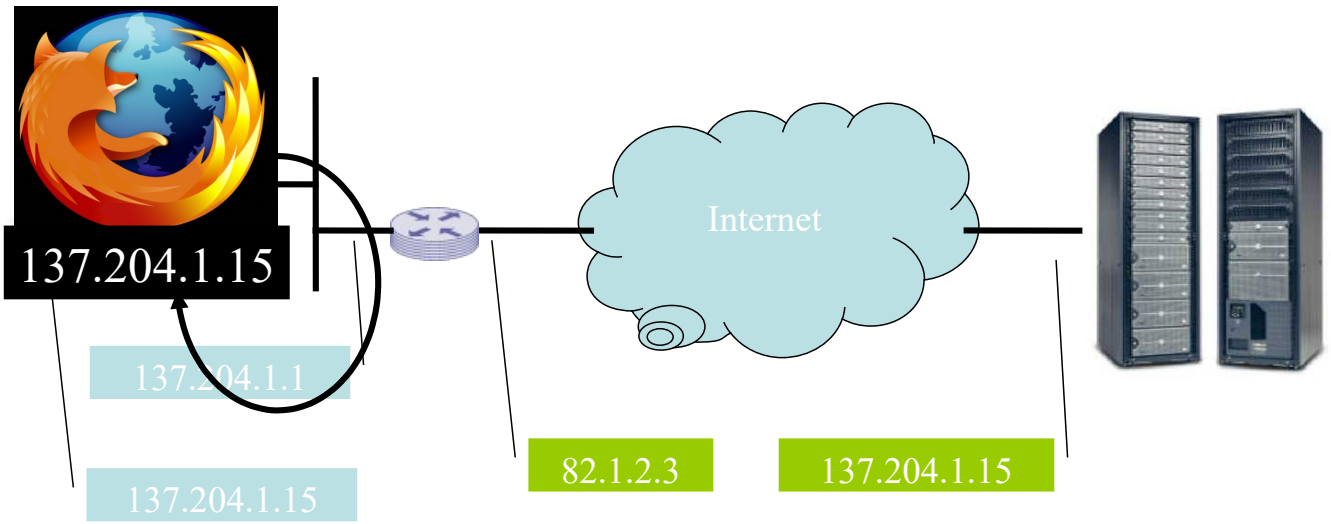
Network Address Translation (NAT)

- La quintupla
  (*protocollo, ip_sorgente, porta_sorgente, ip_destinazione, porta_destinazione*)
  identifica univocamente una connessione
- Nel NAT molti IP sorgente vengono sostituiti dall'unico IP pubblico del router
  - possibilità di modificare la porta sorgente per disambiguare le connessioni originate con tutti i parametri identici a parte l'IP sorgente
  - memorizzazione delle traslazioni per poter riconoscere il destinatario delle risposte

Source IP	Source port	Router IP	Router port	Dest. IP	Dest. port
10.0.0.5	11111	82.1.2.3	11111	137.204.1.15	80
10.0.0.9	11111	82.1.2.3	22222	137.204.1.15	80

Network Address Translation (NAT)

- Gli IP della rete interna risultano del tutto nascosti
- Cosa capiterebbe scegliendoli arbitrariamente?



IP privati (RFC 1918)

- Per evitare il problema del possibile "oscuramento" di IP validi, uno standard definisce alcuni intervalli di indirizzi che non possono essere utilizzati su Internet
  - 10.x.y.z
  - 172.16.x.y ... 172.31.x.y
  - 192.168.x.y

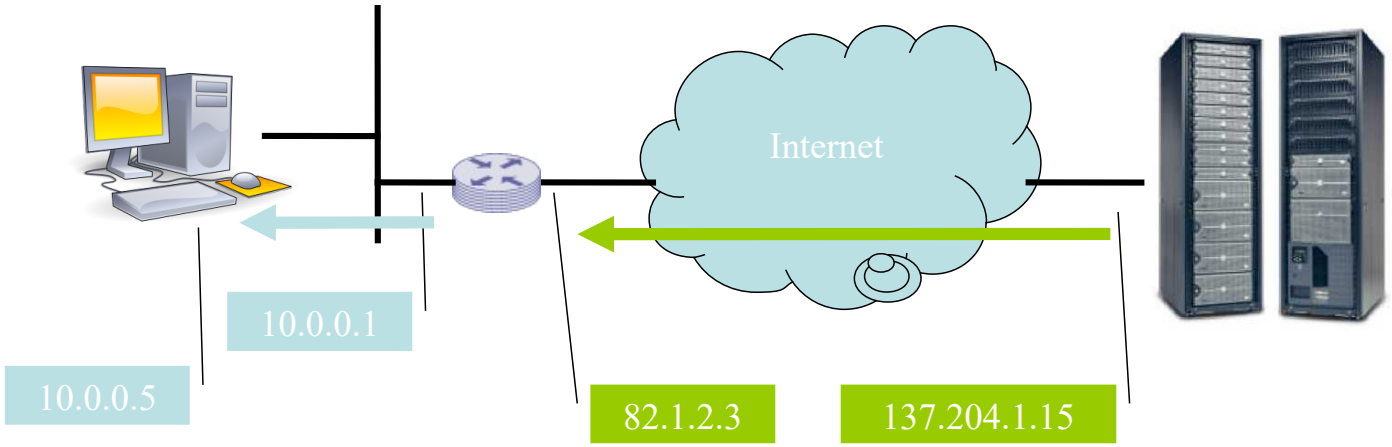
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

SNAT / DNAT

- Per poter utilizzare una rete di client con un solo IP pubblico si modifica l'IP sorgente → Source NAT (SNAT)
  - comportamento trasparente ed automatico
- Lo stesso dispositivo consente anche di rendere raggiungibili alcuni host della rete privata, modificando l'indirizzo di destinazione quando riceve richieste dall'esterno → Destination NAT (DNAT)
  - la mappatura tra porta (servizio) di destinazione ed host interno a cui inoltrare la richiesta va esplicitamente configurata

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DNAT



Richiesta traslata	
Source	137.204.1.15 : 34567
Destination	10.0.0.5 : 80

Richiesta	
Source	137.204.1.15 : 34567
Destination	82.1.2.3 : 80

Aspetti gestionali

- La concessione degli IP è compito dello IANA.
  - http://www.iana.org/ipaddress/ip-addresses.htm
- All'atto pratico vengono affidati blocchi di IP a grandi gestori (Regional Internet Registrar - RIR), che a loro volta ne assegnano sottoinsiemi ai gestori locali (Local Internet Registrar - LIR), come ad esempio i maggiori ISP (Internet Service Provider). I RIR sono:
  - AfriNIC serving Africa
  - APNIC serving the Asia Pacific region
  - ARIN serving North America
  - LACNIC serving South America and the Caribbean
  - RIPE NCC serving Europe, Central Asia and the Middle East.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Configurazione di Linux

- Si può verificare la configurazione delle interfacce di rete con il comando *ifconfig*

```
eth0      Link encap:Ethernet  HWaddr 00:0F:B0:93:FC:D5
          inet addr:137.204.57.118  Bcast:137.204.57.255  Mask:255.255.255.0
          UP BROADCAST MULTICAST  MTU:1500  Metric:1
```

- Si può verificare la configurazione delle tabelle di routing con il comando *route*:

```
Kernel IP routing table
Destination     Gateway         Genmask         Flags Metric Ref    Use Iface
137.204.57.0    0.0.0.0         255.255.255.0   U     0      0        0 eth0
0.0.0.0         137.204.57.254  0.0.0.0         UG    0      0        0 eth0
```

Diagnostica della connettività IP

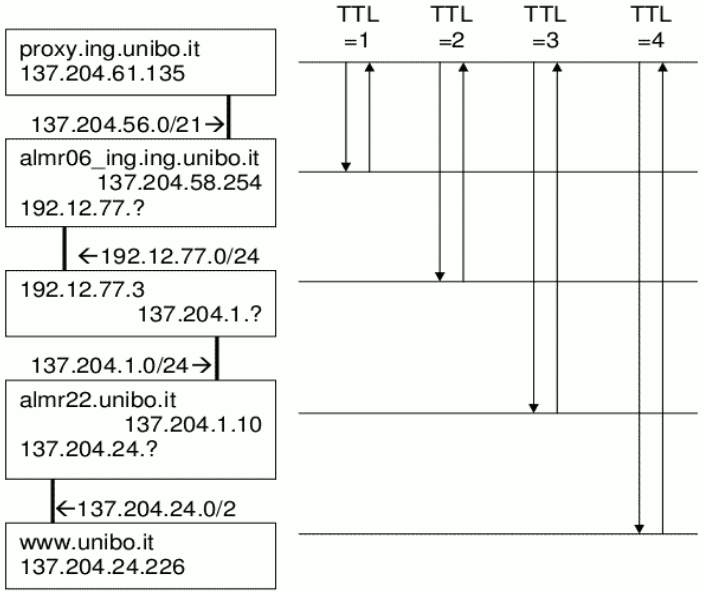
- ping

```
sansone.ing.unibo.it%~% ping -s www.netscape.com
PING www-de1.netscape.com: 56 data bytes
64 bytes from 194.25.242.193: icmp_seq=2. time=172. ms
64 bytes from 194.25.242.193: icmp_seq=6. time=165. ms
64 bytes from 194.25.242.193: icmp_seq=7. time=285. ms
64 bytes from 194.25.242.193: icmp_seq=8. time=371. ms
^C
----www-de1.netscape.com PING Statistics----
10 packets transmitted, 4 packets received, 60% packet loss
round-trip (ms)  min/avg/max = 165/248/371
sansone.ing.unibo.it%~%
```


ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Diagnostica della connettività IP

- traceroute




```
[proxy@proxy proxy]$
     /usr/sbin/traceroute www.unibo.it

traceroute      to      cesias.alma.unibo.it
(137.204.24.226),  30  hops  max,  40  byte
packets

 1  almr06_ing (137.204.58.254)  1.578 ms
 2  192.12.77.3 (192.12.77.3)  4.246 ms
 3  almr22.unibo.it (137.204.1.10)  3.504 ms
 4  cesias.alma.unibo.it (137.204.24.226)
                                  3.108 ms
[proxy@proxy proxy]$
```

Servizi

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Il modello client-server

- Con il termine *server* indichiamo un'applicazione che rende disponibile, mediante un'interfaccia standard, un servizio.
- Con il termine *client* indichiamo un'applicazione in grado di utilizzare i servizi messi a disposizione da un server.

Attenzione: a volte gli stessi termini vengono usati impropriamente per indicare l'host che ospita l'applicazione

- Il modello CS centralizza dati e metodi di elaborazione
  - evita duplicazioni, incoerenze di versione, proliferazione di interfacce
  - introduce problemi di disponibilità e prestazioni

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Domain Name Service (DNS)

- Gli host sono individuati da indirizzi IP, con cui gli apparati lavorano bene perché sono numeri binari.
- Gli esseri umani faticano a ricordare dei numeri, preferiscono usare un nome significativo della funzione dell'host.
- É utile disporre di un servizio che associ un nome all'indirizzo
- Ogni host viene identificato per mezzo del proprio nome (*hostname*) e del nome del proprio dominio (*domain name*).
- I dominî sono definiti in maniera gerarchica.
- I componenti della gerarchia sono separati con un punto
  es.: www.deis.unibo.it

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Top Level Domain (TLD)

- Il primo componente a destra (la parte più elevata della gerarchia) è detto TLD (Top Level Domain) e deve essere scelto essenzialmente:
  - fra i TLD nazionali:
    - sigle di due lettere concesse solo alle nazioni come definite e riconosciute dall'ONU
    - es.: .it, .uk, .fr, .de, .tv, …
  - fra le sigle identificative di determinate categorie (gTLD):
    - dal 1985 esistono: .com, .org, .net, .edu, .gov, .mil, .int
    - dal 2001 anche: .aero, .biz, .coop, .info, .name, .museum, .pro,
    - dal 2005 anche: .jobs, .travel, .cat
    - dal 2006 anche: .mobi
    - Poi .asia, .tel, .xxx, ma altri lasciati in attesa di decisione per anni
  - La critica dell'arbitrarietà delle procedure per ottenere l'attivazione dii un gTLD ha infine portato alla possibilità di richiedere gTLD qualsiasi (decisione 2008, avvio 2012, attivazione 2013)
    - 185.000$ + 25.000$/anno

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

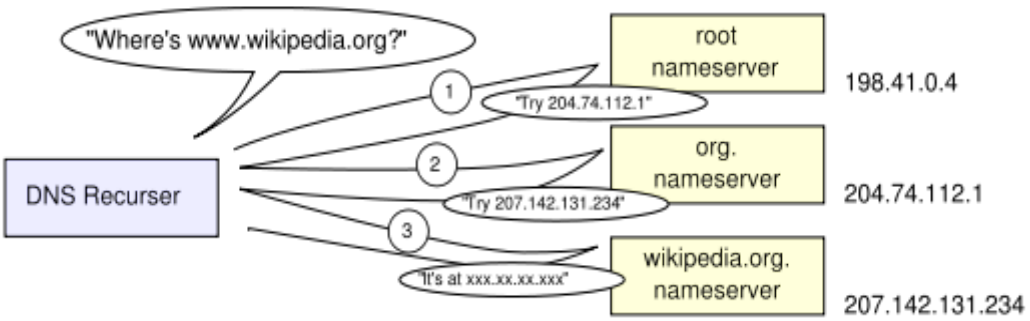
Aspetti gestionali

- La gestione dei TLD è in carico allo IANA.
  - http://www.iana.org/domain-names.htm
- All'atto pratico la gestione di un dominio viene esercitata per mezzo di deleghe ai gestori dei sottodomini:
  - sponsor o operatori dei TLD generici
    - .com --> Verisign, .aero --> SITA, ...
  - enti nazionali responsabili dei TLD geografici
    - .it --> IIT-CNR, ...
  - qualsiasi ente o persona responsabile dei domini al di sotto dei TLD (domini di 2° livello)

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Risoluzione dei nomi in indirizzi

- Il DNS è un sistema client-server.
  - Il client prende il nome di *resolver*, e rivolge le proprie interrogazioni (*query*) tramite il protocollo UDP ad un server in ascolto sulla porta 53.
  - Il server può farsi carico del reperimento dell'indirizzo richiesto, o solo di fornire un puntatore da cui proseguire le ricerche



Risoluzione inversa

- Per la mappatura inversa (da IP a nome) si usa lo stesso meccanismo con un “trucco”: si sfrutta la gerarchia degli IP e si definisce un spazio dei nomi con origine in-addr.arpa. Si noti che
  - gli IP sono scritti dal byte che rappresenta la rete più ampia a quello che rappresenta la parte più specifica dell'indirizzo
    - Es: 137.204.57.1 == RIPE --> UNIBO --> DEIS --> PROMET1
  - i nomi sono scritti a partire dal nome specifico dell'host verso il TLD che rappresenta il dominio più ampio nella gerarchia
    - www.deis.unibo.it
- quindi per usare la stessa infrastruttura la ricerca di 137.204.57.1 viene effettuata con una query del nome 1.57.204.137.in-addr.arpa

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Identificazione dei domini e degli IP

- I registri degli indirizzi IP e dei nomi di dominio sono pubblicamente consultabili per poter risalire ai titolari
  - utilizzo delle query DNS per associare nome <--> indirizzo
  - utilizzo dei servizi *whois* per identificare i responsabili tecnici e legali del domino o del blocco di indirizzi
    - i registri non sono integrati a livello globale
    - utilizzo dei servizi whois del RIR appropriato per l'indirizzo
    - utilizzo dei servizi whois del registrar appropriato per il TLD

Posta elettronica

La posta elettronica (e-mail) è stato uno dei primi servizi introdotti su ARPAnet, ed è tuttora uno dei più utilizzati su Internet.

I problemi affrontati per la ricezione dei messaggi sono:

- È richiesta l'identificazione univoca del destinatario
  - ∏ indirizzo costruito in modo gerarchico: utente@host.dominio
- È necessario poter depositare i messaggi fino alla effettiva lettura
  - casella postale, ovvero spazio su di un server di posta
- È utile un sistema di trasporto che tolleri la temporanea irraggiungibilità del server di destinazione
  - ridondanza, per ogni destinazione sono indicati punti di appoggio alternativi (inizialmente una vera "rete sulla rete")

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Posta elettronica

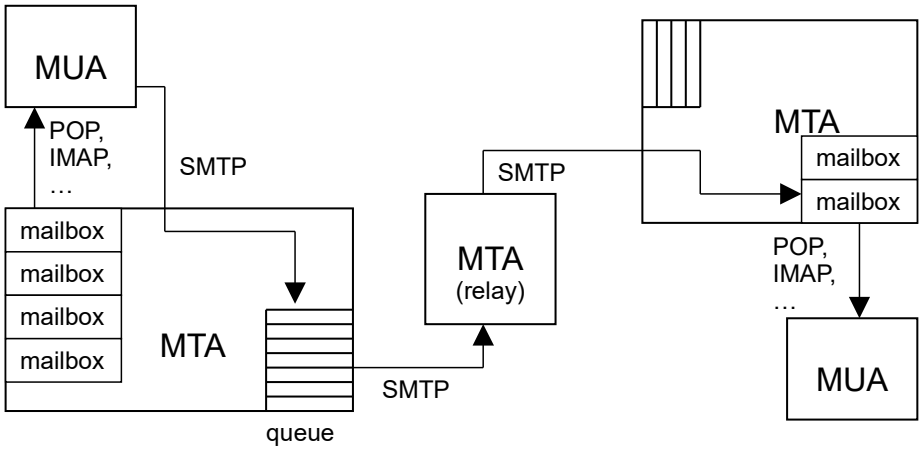
Il reperimento delle informazioni per individuare i server di posta di un dominio è affidato al DNS - stretta integrazione tra i sistemi

I problemi affrontati per la spedizione dei messaggi sono:

- È desiderabile che l'utente veda un'interazione veloce
  - queuing, il server accetta i messaggi e li accoda per spedirli
- L'utente non deve essere obbligato a mantenere la connessione al server durante la lettura e la composizione dei messaggi
  - client di posta elettronica e protocolli di colloquio col server

Posta elettronica

- Applicazioni:
  - MUA (Mail User Agent): Client (es. Eudora)
  - MTA (Mail Transport Agent): Server di transito e smistamento
- Protocolli:
  - SMTP (Simple Mail Transfer Protocol), via TCP alla porta 25
  - POP (Post Office Protocol),), via TCP alla porta 110
  - IMAP (Internet Mail Access Protocol), via TCP alla porta 143



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Posta elettronica

- Originariamente pensata per trasferire solo testi
  - *header*, contenente le informazioni per la consegna del messaggio (destinatario, mittente, soggetto, data, …)
  - *body*, contenente il testo del messaggio.
- Presto riconosciuta come utile per trasferire dati di ogni genere
  - estendere le possibilità di codifica del body con lo standard MIME (Multipurpose Internet Mail Extension)
  - Il corpo della mail rimane in formato ascii, ma in esso sono codificate le informazioni binarie originali (attach).
    - Possibilità di spedire più dati distinti in un solo messaggio
    - Possibilità di spedire una indicazione del tipo di dato platform-independent
    - Consente di definire nuovi tipi di dato
  - del tutto trasparente per i server

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Formato di un messaggio

```
From luca.ghedini@mail.ing.unibo.it Wed Nov 12 23:06:49 1997
Received: (from ghedo@localhost)
          by mail.ing.unibo.it (8.8.8/8.8.5) id XAA04169
          for luca.ghedini; Wed, 12 Nov 1997 23:06:49 +0100 (MET)
Date: Wed, 12 Nov 1997 23:06:49 +0100 (MET)
From: Luca Ghedini <luca.ghedini@mail.ing.unibo.it>
To: luca.ghedini
Subject: Saluti
Mime-Version: 1.0
Content-Type: multipart/mixed;
boundary=6e57_3bae-6e62_7e55-1cf_41
Content-Length: 590

--6e57_3bae-6e62_7e55-1cf_41
Content-Type: text/plain; charset=us-ascii
Content-Transfer-Encoding: 7bit
Content-MD5: DjPyOvmsowwWAIx7HqbxRg==
X-Sun-Data-Type: text

Ciao, ti mando le cose che mi avevi chiesto
--6e57_3bae-6e62_7e55-1cf_41
Content-Type: image/gif; name=2torri.gif
Content-Transfer-Encoding: base64
Content-MD5: LcOMV2MEgHKz/Ufm58vvlQ==
Content-Description: 2torri.gif
Content-Disposition: attachment; filename=2torri.gif
X-Sun-Data-Type: gif-file

R0lGODlh2AA0AfcAACAgQCAgAg/yBAQCBAwCBA/yBgQCBggCBgwCBg/yCAQCCA
cmtzIEluYy4NDVVzZSBubyBob29rcwA7
--6e57_3bae-6e62_7e55-1cf_41--
```

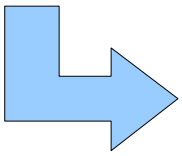

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

World Wide Web

Il problema: recuperare informazioni in un contesto altamente eterogeneo.

Cosa esisteva prima del web ?

- **programmi specializzati per accedere a informazioni codificate in differenti maniere.**
- **difficoltà di integrazione delle informazioni provenienti da fonti non omogenee**
- **interfacce di accesso non consistenti**



- **Costi elevati**
- **Difficoltà di apprendimento**
- **Rapida obsolescenza dei programmi**
- **Necessità di “aggiustamenti” manuali**

EnAIP Tecnico Informatico

Scopi del WWW (CERN 1989)

- differenti rappresentazioni delle informazioni (es. per le immagini : GIF, JPEG, BMP,...)
- differenti modalità di recupero delle informazioni (database query, ftp site, ...)
- differenti modalità di autenticazione (password, chiavi pubbliche, DES,...)
- trasparenza accesso e allocazione
- presentazione multimediale
- unicità di interfaccia utente

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Ipertesti

L'elemento base del www è l'ipertesto.

- Un ipertesto e' un testo arricchito di immagini, suoni, e riferimenti ad altre informazioni (link)
- Ciascun link punta ad un'altra informazione che può essere ovunque (World Wide)
- L'insieme dei link forma una rete (Web) in cui si incrociano i percorsi logici secondo cui l'informazione può essere esplorata.
- Il risultato è un ipertesto distribuito su una rete di calcolatori

Componenti del WWW

- Browser
  - accetta le richieste dell'utente e gli presenta i risultati della esplorazione
- Server:
  - Immagazzina gli ipertesti, elabora le richieste ricevute dai browser ed invia loro i risultati della elaborazione
- Helper
  - moduli per aggiungere ai browser particolari funzionalità
- Common Gateway Interface/Server Side Scripting
  - permettono agli utenti di eseguire programmi sui server (esecuzione remota).

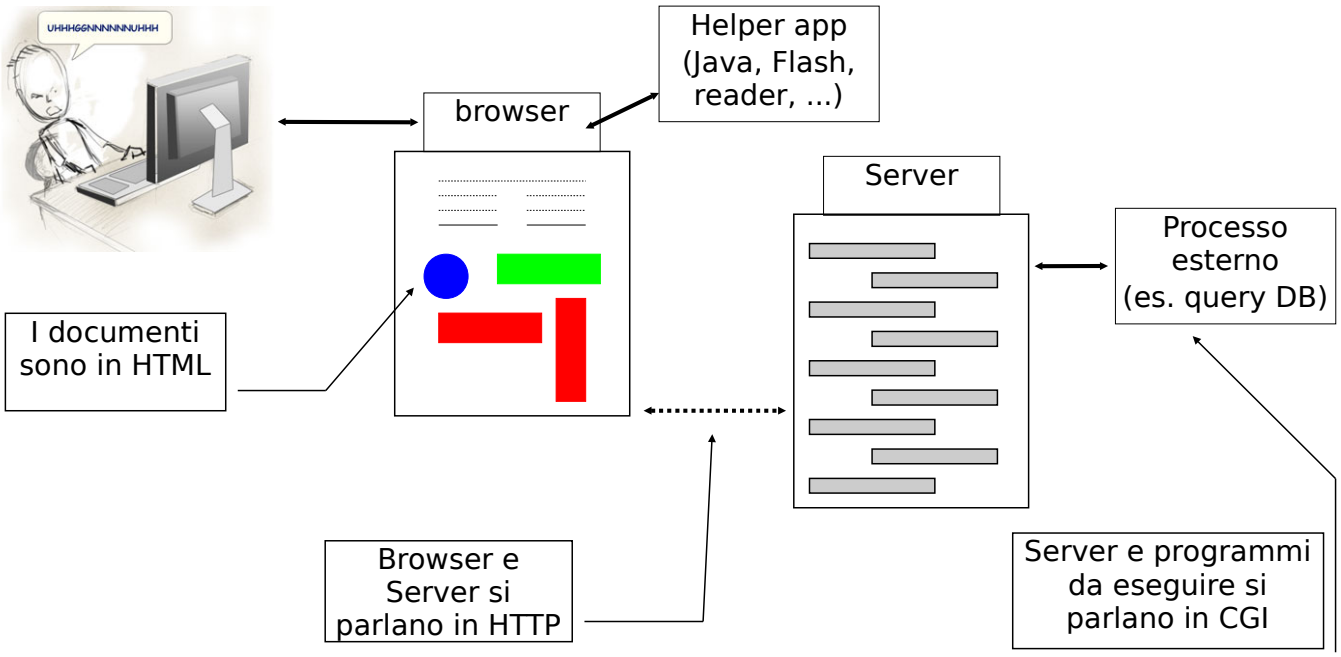
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Standard relativi al WWW

Occorrono degli standard:

- per identificare le risorse in rete
  - Uniform Resource Identifier/Locator (URI/URL)
- per far comunicare server e browser
  - HyperText Transfer Protocol (HTTP)
- per permettere al browser di presentare i documenti all'utente
  - HyperText Markup Language (HTML)
  - Cascading Style Sheets (CSS)
- per gestire l'esecuzione lato server
  - Common Gateway Interface (CGI)

Standard relativi al WWW



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Transazione HTTP

- **Il protocollo è il più semplice possibile**
  - **V. 1.1 (RFC 2068)**
  - **REQUEST/RESPONSE**
    - **nessuna memoria del passato influenza le risposte successive**
  - **Capacità di negoziazione**
    - **il browser indica che tipi di dato, lingue, codifiche è in grado di gestire**
    - **è possibile elencare alternative in ordine di preferenza**
  - **Riutilizzo degli standard**
    - **messaggi codificati secondo MIME**

GET /pippo.gif HTTP/1.0

HTTP/1.0 200 OK
Date: Sunday, 20-Apr-1995 20:23:57 GMT
Server: apache/1.2
MIME-version: 1.0
Content-type:image/gif
last-modified: Sunday, 20-Apr-1995 20:23:57 GMT
Content-lenght: 2245

.
. (dati della immagine)
.

EnAIP Tecnico Informatico

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

HTTP Request

- 2 parti separate da una riga vuota:
  - header <metodo> <URL> <versione>
    <opzioni>
  - body (opzionale) dati relativi alla richiesta
- Metodo: Es. GET, HEAD, POST, PUT
- URL solo la parte relativa (tolti protocollo, host, porta)
- Versione HTTP/1.0 o HTTP/1.1
- Opzioni Es. Accept:…, If-modified-since:…, Referer:…, User-Agent:…,

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

HTTP Response

- 2 parti separate da una riga vuota:
  - header <status code> <status description>
    <opzioni>
  - body <risposta>
- Status:

1xx	Information
2xx	Success
3xx	Redirection
4xx	Client Error
5xx	Server Error

- Opzioni Es. Content-type:…, Expires:…, Content-encoding:…,

DNS – NTP

Breve descrizione e configurazione dei servizi su Linux

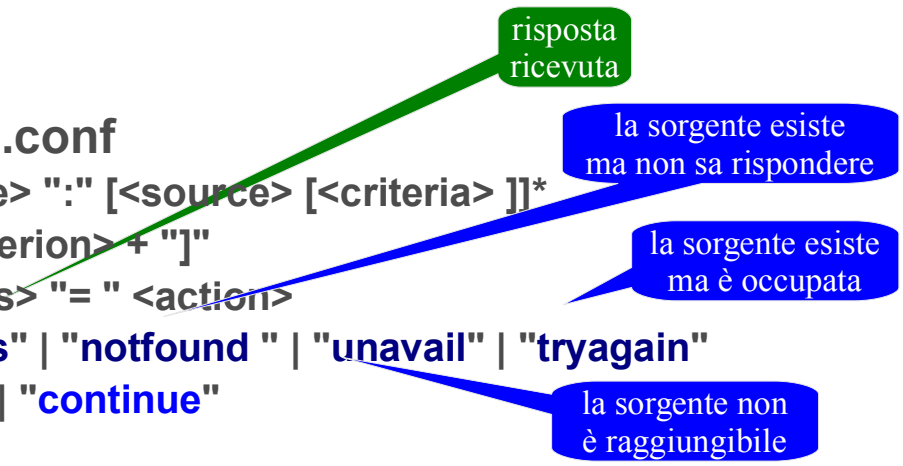
Marco Prandini

Risoluzione dei nomi - generalità

- La mappatura da nomi di host a indirizzi IP e viceversa è uno dei tanti casi in cui il sistema ha bisogno di un dizionario di nomi
- Il primo accorgimento adottato da GNU/Linux riguarda la *scelta della sorgente di informazioni*
  - *Name Service Switch*
  - svolta dalla libreria C di sistema
  - supporta un set fisso di possibili database
  - configurata tramite `/etc/nsswitch.conf`
    - vedi man page omonima

NSS

- Sintassi di nsswitch.conf
  - <entry> ::= <database> ":" [<source> [<criteria>]]*
  - <criteria> ::= "[" <criterion> + "]"
  - <criterion> ::= <status> "= " <action>
  - <status> ::= "**success**" | "**notfound** " | "**unavail**" | "**tryagain**"
  - <action> ::= "**return**" | "**continue**"



(i colori indicano l'azione di default)

- Es.

```
passwd: files nis ldap
group: files ldap
hosts: ldap [NOTFOUND=return] dns files
```

Risoluzione dei nomi – host e IP

```
hosts: ldap [NOTFOUND=return] dns files
```

- **`files`** → la sorgente di informazioni è il file **`/etc/hosts`**
  - formato: `<IP>  <FQDN>  [<ALIAS> …]`
  - esempio: `8.8.8.8  dns.google.com  gdns`
- **`dns`** → la sorgente di informazioni è il sistema DNS
  - l'interrogazione di server DNS è un'ulteriore set di funzioni della libreria C di sistema, il *resolver*
  - si configura attraverso **`/etc/resolv.conf`**
  - esempio

```
nameserver 137.204.58.1
domain disi.unibo.it
search ing.unibo.it
```

DNS caching

- Spesso si trova un server DNS locale
  - Miglioramento prestazioni
  - Maggiore flessibilità per contesti dinamici

```
~$ cat /etc/resolv.conf
# Dynamic resolv.conf(5) file for glibc resolver(3) generated by resolvconf(8)
#     DO NOT EDIT THIS FILE BY HAND -- YOUR CHANGES WILL BE OVERWRITTEN
nameserver 127.0.1.1
```

  - Tutti gli IP che iniziano per 127 puntano a localhost

```
sudo ss -naup | grep 127.0.1.1:53
...
UNCONN     0      0      127.0.1.1:53     *:*      users:(("dnsmasq",pid=2154,fd=4))
```

Risoluzione di nomi via NSS

- Il comando getent permette di interrogare i database del name service switch

```
getent <db name> <keyword>
```

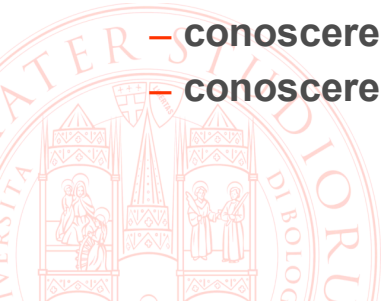
Esempi:

```
$ getent passwd las
las:x:1000:1000:Lab Amministrazione Sistemi,,,:/home/las:/bin/bash

$ getent hosts www.unibo.it
137.204.24.35   atrproxy.unibo.it www.unibo.it
```

Risoluzione nomi DNS diretta

- Per interrogare direttamente il DNS e avere più controllo sulle query si usano tipicamente host e dig
  - non considerano nsswitch
  - usano i nameserver di resolv.conf di default
  - possono interrogare un server specifico
- *host* (tipicamente per conversioni IP ←→ nome)
  - query di un nome: `host www.unibo.it`
  - query a un server specifico: `host www.unibo.it 8.8.8.8`
- *dig* (tipicamente per ottenere informazioni legate a un dominio diverse da nomi host)
  - conoscere i Mail eXchanger: `dig mx example.com`
  - conoscere i Name Server: `dig ns example.com`

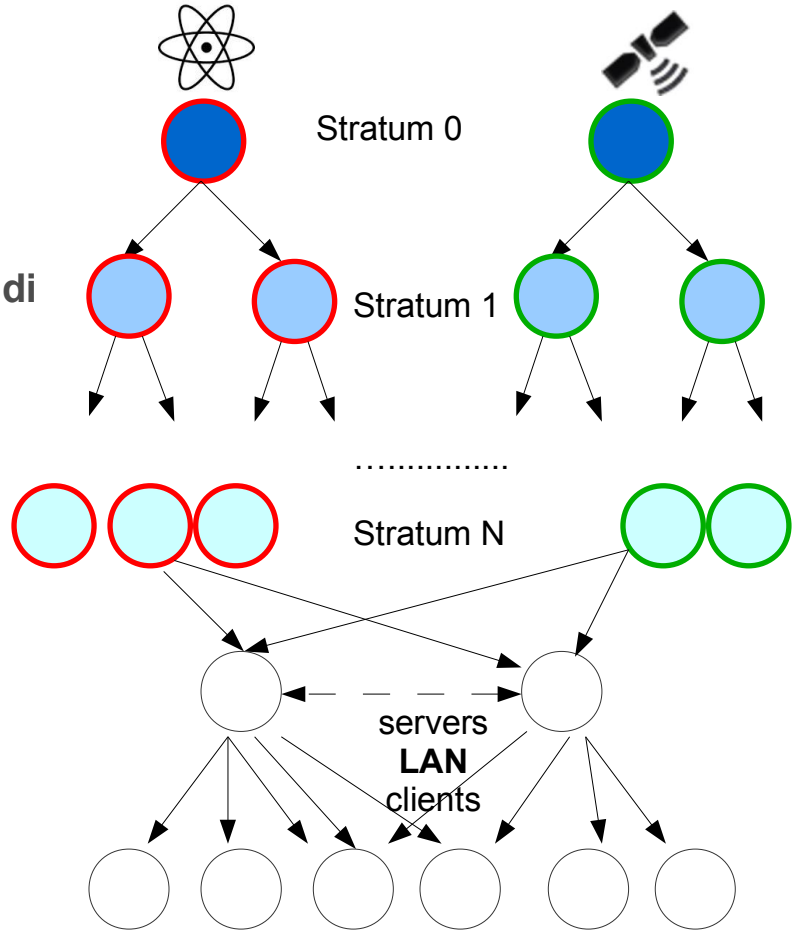


Sincronizzazione

- L'allineamento dell'ora di un sistema ad un orologio di riferimento è cruciale
  - per la diagnostica dei problemi (timestamp su log)
  - per i protocolli di autenticazione e autorizzazione (i messaggi hanno una vita limitata)
  - per la sincronizzazione di azioni distribuite
  - per il valore legale di azioni compiute attraverso i computer
- È possibile usare un protocollo che compensa i ritardi di rete per ottenere informazioni precise via Internet: *Network Time Protocol* (NTP)
  - sito ufficiale: http://www.ntp.org/
  - grande quantità di informazioni su: http://www.eecis.udel.edu/~mills/ntp.html

NTP in breve

- Preciso
  - poche decine di millisecondi di scarto su WAN
  - <1 millisecondo su LAN
  - supporto di sorgenti HW (oscillatori, GPS, ...)
- Standard
  - portato su ogni architettura nota
- Scalabile e affidabile
  - *multi-server*
  - *strata*
  - *peering*
  - *auto-keying*



NTP su Linux

- Il demone *ntpd* è client e/o server in funzione della configurazione
- /etc/ntp.conf – esempio

```
server 0.ubuntu.pool.ntp.org
server 1.ubuntu.pool.ntp.org
peer fellow.server.lan
# By default, exchange time with everybody, but don't allow
configuration.
restrict -4 default kod notrap nomodify nopeer noquery
restrict -6 default kod notrap nomodify nopeer noquery
# Local users may interrogate the ntp server more closely.
restrict 127.0.0.1
restrict ::1
```

NTP – inizializzazione e uso sporadico

- Il tool `ntpdate` permette di sincronizzare l'orologio locale a un server NTP
  - senza parametri usa i server in `ntp.conf`
    - `ntpd` non deve essere attivo
  - accetta come parametro un server specifico

- L'ora viene modificata in due modi
  - se la differenza è più di 0.5 secondi: step
  - se la differenza è meno di 0.5 secondi: slew con `adjtime()`

- Non rimpiazza ntpd, che usa algoritmi sofisticati
  - per compensare errori e ritardi dei pacchetti dai server
  - per profilare il comportamento dell'orologio locale