

Cenni sulle reti di calcolatori

- Nel corso degli anni si è passati da sistemi **centralizzati** (un mainframe con tanti terminali collegati) a sistemi **distribuiti** (tanti piccoli calcolatori collegati in rete).
- Questo fenomeno è noto come **downsizing**.
- Nel caso delle reti di calcolatori un fattore critico è la **qualità dei collegamenti**:
 - a livello fisico (cavi e segnali elettrici);
 - a livello della gestione del trasferimento di informazione (protocolli di comunicazione).

Tipi di reti

- Essenzialmente vi sono tre tipi di reti:
 - **Reti locali** (LAN: Local Area Network)
 - **Reti metropolitane** (MAN: Metropolitan Area Network)
 - **Reti geografiche** (WAN: Wide Area Network)

Reti locali

- Sono costituite da un insieme di calcolatori collegati da un mezzo trasmissivo **veloce** ed a **basso tasso d'errore** (ad esempio Ethernet) in **ambito locale**.
- La velocità di trasmissione di una rete locale tradizionale varia da 1 a 20 Mbit/s.
- Le reti locali ad alta velocità hanno velocità di trasmissione da 100 Mbit/s a 10 Gbit/s.

Reti metropolitane

- Sono un'estensione delle reti locali in ambito urbano.
- Sono caratterizzate dall'utilizzo di fibre ottiche per i collegamenti fisici (dorsali).
- Sono soggette ad una legislazione particolare in quanto fanno transitare dei dati su “suolo pubblico”.

Reti geografiche

- Sono gestite dai fornitori nazionali di telecomunicazioni (per trasmettere diverse tipologie di dati).
- Hanno solitamente velocità trasmissive modeste (ma possono arrivare a 10/100 Mbit/s).
- Due nodi (stazioni) sulla rete non condividono generalmente un unico mezzo trasmissivo (i dati passano per una serie di nodi intermedi detti **nodi di commutazione**).

L'informazione nel World Wide Web

- L'ormai capillare diffusione di **Internet** (ed in particolare del **World Wide Web**) sta rivoluzionando molte attività.
- L'innovazione chiave della “rete delle reti” è che consente virtualmente a chiunque di **pubblicare delle informazioni** rendendole immediatamente e facilmente accessibili ad una vasta comunità di utenti.
- La ricchezza delle informazioni reperibili in rete è data anche dal fatto che non sono limitate a contenuti testuali, ma possono essere **multimediali** (integrando testo, immagini, suoni, animazioni/filmati).

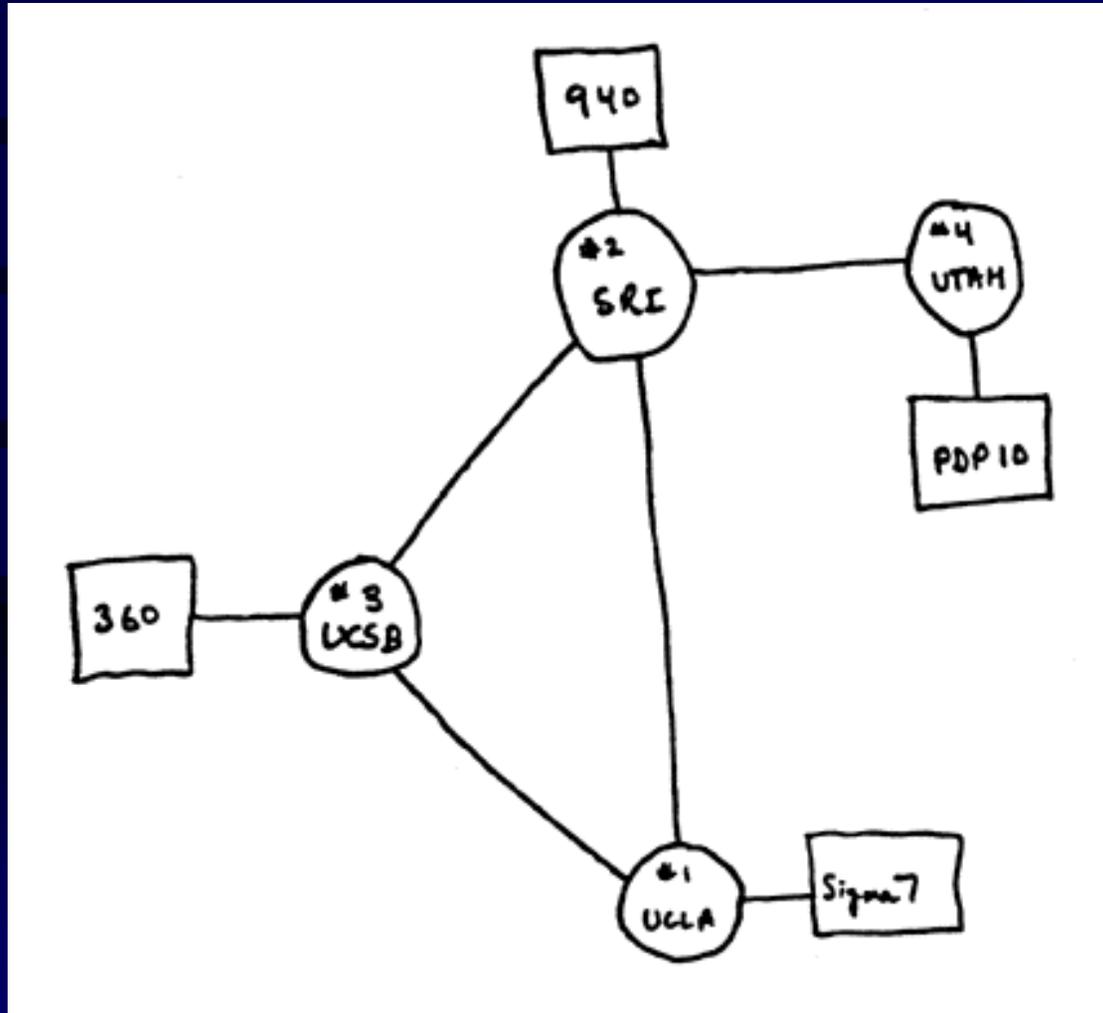
L'informazione nel World Wide Web

- La libertà di comunicazione offerta comporta anche degli svantaggi:
 - difficoltà nel reperimento delle informazioni cercate (data la mole di dati presente in Internet),
 - manca di una struttura uniforme,
 - attendibilità delle informazioni.
- L'espressione “autostrada dell'informazione” (information highway) viene spesso usata come sinonimo di Internet e dei servizi offerti dalla rete.
- Negli Stati Uniti è in corso un dibattito per stabilire se sia il caso di imporre un controllo sul flusso di informazioni da parte di enti governativi (come avviene per la gestione del traffico su una comune autostrada).

Le origini di Internet

- **1957**: gli Stati Uniti istituiscono la **Advanced Research Projects Agency** (ARPA - Agenzia per i Progetti di Ricerca Avanzata). Si tratta di un dipartimento del Ministero della Difesa con l'incarico di preservare il primato americano nella scienza e nella tecnologia militare.
- **1969**: l'ARPA sviluppa **ARPANET** (il precursore di Internet): si tratta di una rete che collega i quattro computer di quattro università del sud-ovest degli Stati Uniti (UCLA, SRI, UCSB e l'Università dello Utah).
- **1970-1971**: alla rete si aggiungono il MIT, Harvard, Carnegie-Mellon e molti altri centri di ricerca ed università. La rete inizia ad espandersi velocemente.

ARPANET (1969)



Internet negli anni '70 – '80

- Grazie all'introduzione del protocollo (UUCP: Unix to Unix Copy Protocol) inizia a diffondersi l'uso dei **gruppi di discussione** (newsgroup), mettendo in contatto sistemi Unix presenti in tutto il mondo.
- Con lo sviluppo della suite di protocolli **TCP/IP** (Transmission Control Protocol/Internet Protocol) Internet si evolve; vengono introdotti i servizi di **e-mail**, **FTP** e **telnet** che permettono l'utilizzo della rete anche ai non addetti ai lavori.
- Vista la crescita continua di Internet, si sente la necessità di creare degli **strumenti per tenere traccia delle risorse disponibili**. Nascono **Archie** (Peter Deutsch, McGill University in Montreal, Canada) e **WAIS** (Wide Area Information Service).

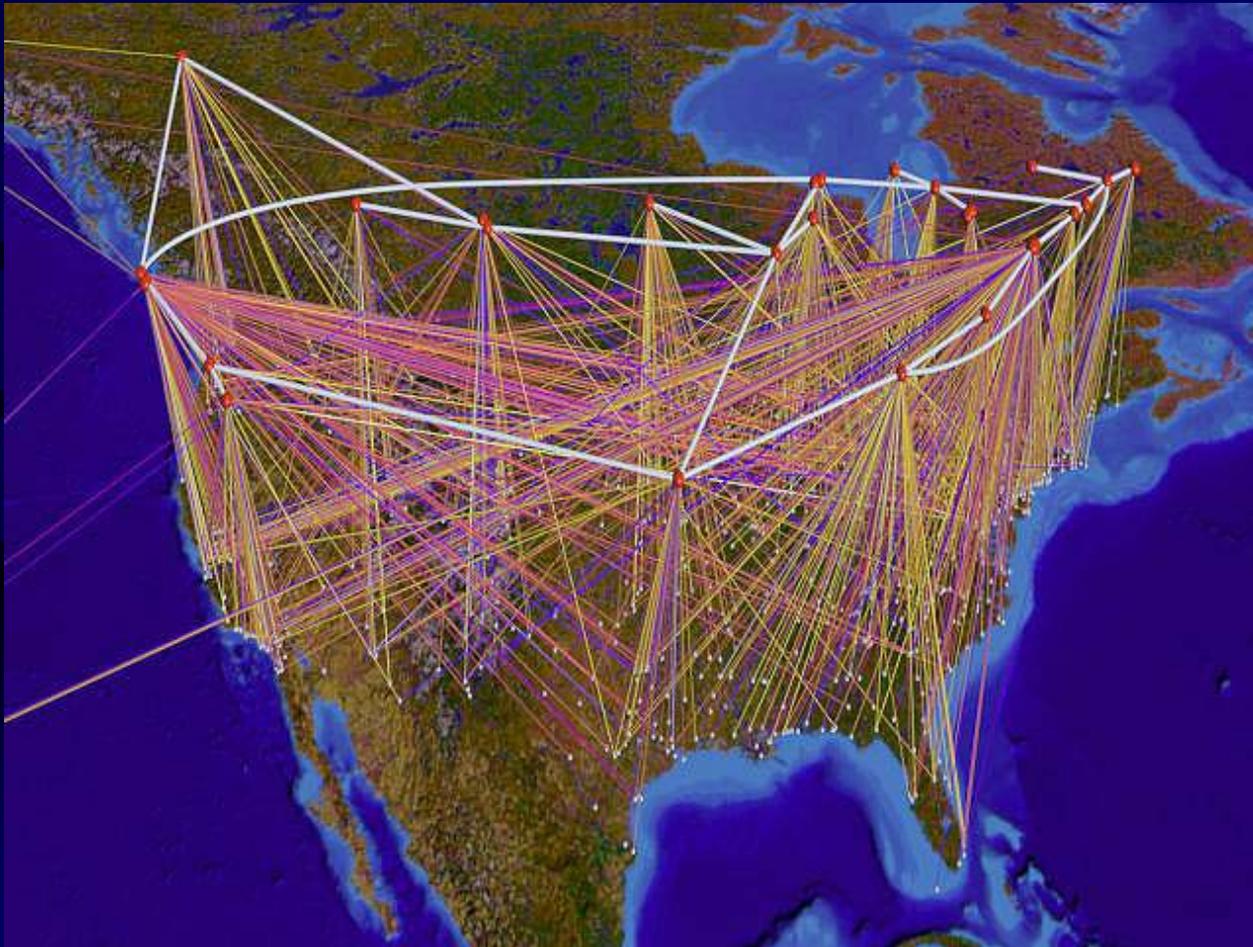
Internet negli anni '90

- **1991**: presso l'Università del Minnesota viene creata la prima interfaccia “semplice” per Internet che trova una larga diffusione: **Gopher** (basato su un'architettura client-server).
- La popolarità di Gopher cresce a dismisura quando l'Università del Nevada (Reno) sviluppa un indice consultabile di siti gopher chiamato **VERONICA**: Very Easy Rodent-Oriented Netwide Index to Computerised Archives (indice di rete per roditori di facile uso degli archivi informatici).

Internet negli anni '90

- **1989**: **Tim Berners-Lee** ed altri ricercatori del CERN propongono un nuovo protocollo (HTTP: HTML Transfer Protocol) per lo scambio di documenti: **nasce il World Wide Web**.
- **1993**: Marc Andreessen (NCSA) sviluppa con la sua équipe Mosaic (il primo browser grafico).
- **1995**: spariscono tutte le limitazioni all'uso commerciale di Internet.
- **Nel 1993 i siti Web erano solo 130**. Al giorno d'oggi ce ne sono milioni.

Internet oggi negli USA



Indirizzi IP

- Per comunicare fra loro i computer (detti anche **host**) utilizzano dei **protocolli** (i.e., insiemi di regole). Tra questi il protocollo IP (Internet Protocol) stabilisce il formato degli indirizzi IP.
- Un **indirizzo IP** è formato da quattro numeri (ognuno può assumere un valore da 0 a 255) separati da punti.
- La funzione degli indirizzi IP è quella di **identificare in modo univoco una macchina** (host) all'interno di una rete (network).

Indirizzi IP

- Quindi ogni host ha un indirizzo IP unico all'interno della rete in cui si trova, in modo da permettere di creare un canale virtuale di comunicazione fra due host.
- Esistono varie classi (tipologie) di indirizzi IP, a seconda del tipo di rete a cui l'host è connesso.

Classi di indirizzi IP

- Nel seguito ‘n’ indica un network bit, ‘h’ un host bit:
- **Classi di tipo A** -
0nnnnnnn.hhhhhhhh.hhhhhhhh.hhhhhhhh (primo bit: 0; 7 network bit; 24 host bit; primo byte: 0 – 127).
Esistono quindi **128 classi di tipo A**. Ci possono essere **16.777.216 host in ogni classe A**.
- **Classi di tipo B** -
10nnnnnnn.nnnnnnnn.hhhhhhhh.hhhhhhhh (primi 2 bit: 10; 14 network bit; 16 host bit; primo byte: 128 – 191). Esistono quindi **16.384 classi di tipo B** ed ognuna di esse può contenere **65.536 host**.

Classi di indirizzi IP

- Nel seguito ‘n’ indica un network bit, ‘h’ un host bit, ‘m’ un multicast bit e ‘r’ un bit riservato:
- **Classi di tipo C** - 110nnnnn.nnnnnnnn.nnnnnnnn.hhhhhhhh (primi 3 bit: 110; 21 network bit; 8 host bit; primo byte: 192 – 223). Esistono quindi **2.097.152 classi di tipo C** ed ognuna di esse può contenere **256 host**.
- **Classi di tipo D** - 1110mmmm.mmmmmmmm.mmmmmmmm.mmmmmmmm (primi 4 bit: 1110; 28 bit per il multicast address, i.e., trasmissione di dati a più host contemporaneamente; primo byte: 224 – 247).
- **Classi di tipo E** - 1111rrrr.rrrrrrrr.rrrrrrrr.rrrrrrrr (primi 4 bit: 1111; 28 bit per indirizzi riservati; primo byte: 248 – 255).

Classi di indirizzi IP - Esempi

- L'indirizzo IP del web server dell'Università di Udine è 158.110.1.44.
- Analizzando il valore del primo numero (158) si evince che l'host appartiene ad una rete con 65.536 host (indirizzo di classe B) in quanto $128 < 158 < 191$.
- Ogni host può “comunicare con se stesso” utilizzando l'indirizzo IP 127.0.0.1 (indirizzo di classe A). Infatti quest'ultimo rappresenta il local host (macchina locale).
- Tutti gli host che appartengono alla stessa rete (network) hanno un indirizzo IP tale che la parte di quest'ultimo che individua la rete stessa è invariante.
- Ad esempio, tutti i computer dell'università di Udine hanno indirizzi IP della forma 158.110.x.y, ovvero, variano soltanto gli ultimi due numeri (infatti i primi due individuano la rete di tipo B).

Nomi simbolici

- Nonostante gli indirizzi IP siano un modo efficace per individuare in modo univoco un host sulla rete, le persone li trovano scomodi in quanto sono dei dati numerici, difficili da ricordare.
- Quindi fin dagli inizi di Internet ad ogni indirizzo IP si associava un nome simbolico (il nome dell'host).

Nomi simbolici

- In tal modo le persone potevano utilizzare il nome simbolico (facile da ricordare) per riferirsi ad un host; il software di rete poi si occupava di tradurre tale nome nell'indirizzo IP corrispondente in maniera da stabilire la connessione.
- Inizialmente le associazioni tra nomi simbolici ed indirizzi IP erano mantenute in un file di testo (una per riga). Quest'ultimo era memorizzato e mantenuto aggiornato presso il **NIC (Network Information Center)**; ogni host doveva periodicamente scaricarlo per essere al passo con i cambiamenti di Internet (e.g., comparsa di nuovi host).

Nomi simbolici - DNS

- La soluzione **centralizzata** del NIC, con il rapido evolversi di Internet, cominciava a mostrare dei limiti:
 - Gli aggiornamenti si rendevano via via più frequenti a causa della forte crescita della rete (aggiunta di nuovi host).
 - La necessità di scaricare frequentemente il file aggiornato dal server INTERNIC creava dei problemi di traffico di rete.
 - Diventavano frequenti i casi in cui due host distinti (quindi con indirizzi IP diversi) si registravano presso l'INTERNIC con lo stesso nome simbolico.
- Per superare tali problemi si passò ad un database **distribuito** noto come **DNS (Domain Name System)**. Tale soluzione permise un **controllo locale** sulle singole parti della mappatura nomi-indirizzi, mantenendo accessibile nel contempo l'informazione a tutta la rete.



DNS

- L'architettura su cui si basa il DNS è di tipo **client-server**:
 - I **name server** contengono parte delle informazioni sulle associazioni nomi-indirizzi.
 - I **client** (o resolver) formulano delle **query** (interrogazioni) ai name server per ottenere l'indirizzo IP corrispondente ad un dato nome.

DNS

- La struttura del database del DNS è quella di un albero rovesciato (con la radice in alto e le foglie in basso):



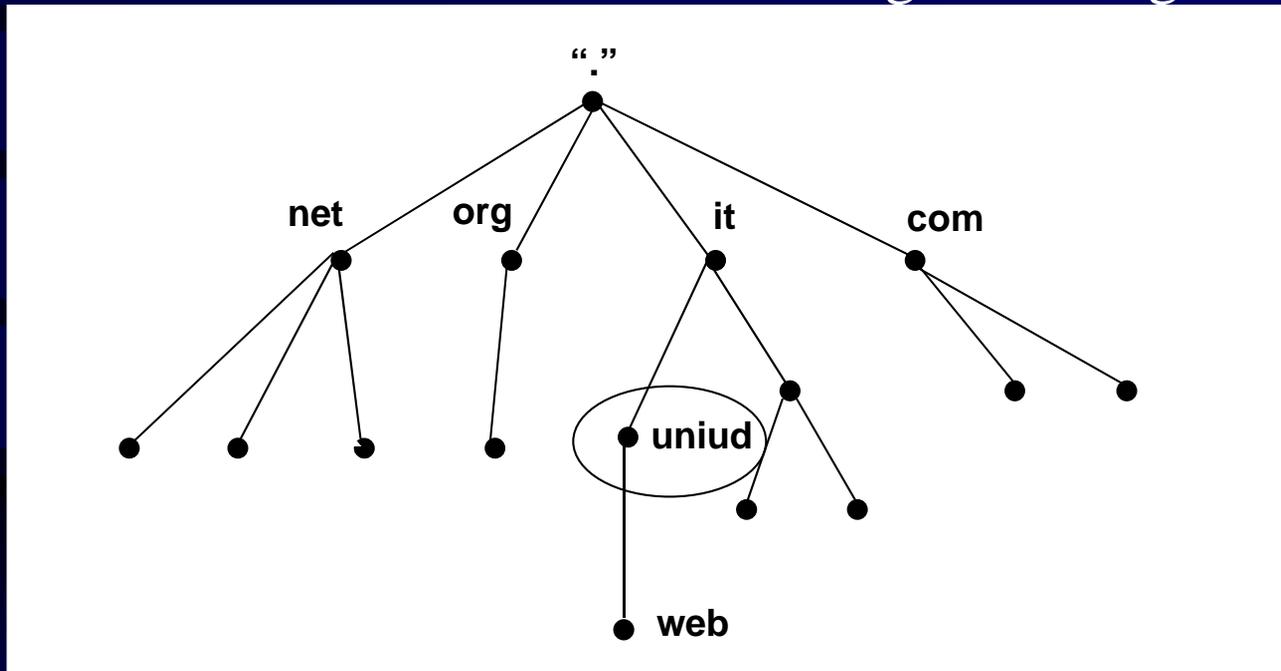


DNS

- Ogni nodo dell'albero può essere visto come la radice del sottoalbero ottenuto considerando soltanto i rami ed i nodi “sottostanti” al nodo considerato.
- Ogni sottoalbero corrisponde ad una porzione del database ed è chiamato **dominio**.
- Ogni dominio ha un nome unico, ottenuto concatenando tramite dei punti (“.”) le etichette dei nodi attraversati a partire dal nodo radice del dominio al nodo radice di tutto l'albero (tale nodo, detto anche “world”, ha come etichetta il punto, che solitamente non si scrive nel nome risultante).

DNS - Esempio

- Consideriamo il nodo cerchiato nel seguente diagramma:



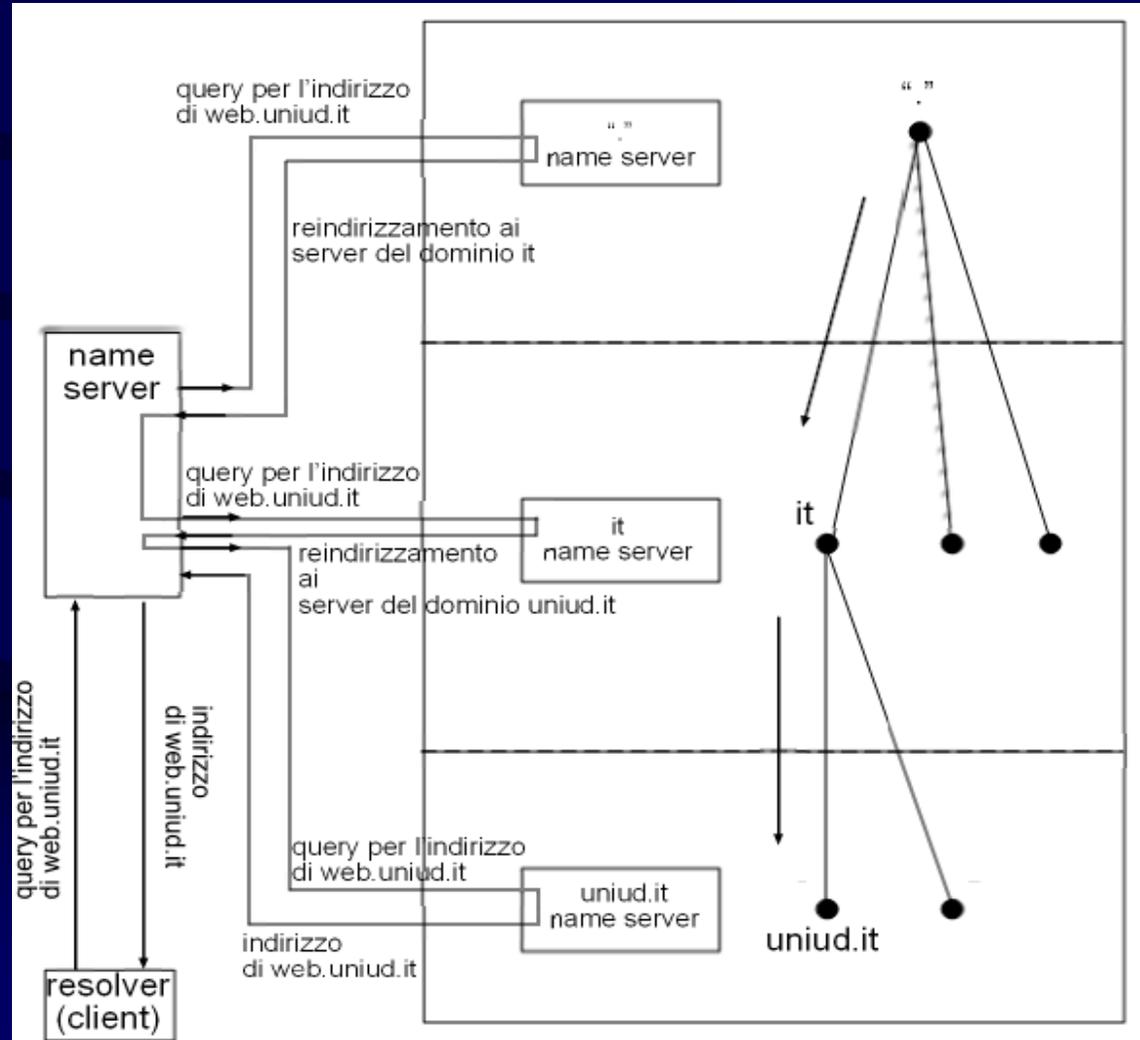
- Il nome del dominio corrispondente è: uniud.it (il punto finale, ovvero l'etichetta della radice dell'albero, non si scrive).

DNS

- I nodi figli della radice (“.”) sono i cosiddetti **domini top level**: **com** (e.g., IBM: ibm.com), **edu** (e.g., Università di Berkeley: berkeley.edu), **gov** (e.g., NASA: nasa.gov), **mil** (e.g., Esercito USA: army.mil), **net** (e.g., NSFNET: nfs.net), **org** (e.g., Electronic Frontier Foundation: eff.org), **int** (e.g., NATO: nato.int), **it** (e.g., Università degli Studi di Udine: uniud.it), ...
- Vantaggi del DNS:
 - Essendo il database distribuito, il carico di lavoro dovuto all’interrogazione, al mantenimento ed all’aggiornamento delle varie parti non va a gravare su un unico server.
 - Grazie alla struttura ad albero non vi sono rischi di duplicazione di nomi; infatti se anche due nodi appartenenti a domini distinti avessero la stessa etichetta, i rispettivi nomi completi (FQDN, ovvero, fully qualified domain name) sarebbero comunque diversi.

Risoluzione dei nomi

- Il diagramma a lato mostra i passi necessari per ottenere l'indirizzo (IP) del web server dell'Università di Udine (web.uniud.it):



Risoluzione dei nomi

- La catena di eventi descritta nel precedente diagramma si verifica quando, ad esempio, vogliamo collegarci al sito web dell'Università di Udine (web.uniud.it):
 - Il resolver (e.g., il browser) richiede al name server della propria rete locale (e.g., al name server del proprio provider) l'indirizzo IP corrispondente al nome simbolico web.uniud.it
 - Se il name server non conosce la risposta a tale query (perché già eseguita recentemente), inoltra quest'ultima ad uno dei name server del mondo (".") di cui conosce l'indirizzo IP.
 - (continua ...)

Risoluzione dei nomi

- (...continua)
 - Il name server del mondo risponde fornendo l'indirizzo IP di uno dei name server del dominio it, a cui il name server della rete locale inoltra la query.
 - Il name server del dominio it risponde fornendo l'indirizzo IP di uno dei name server del dominio uniud.it, a cui il name server della rete locale inoltra la query.
 - Finalmente il name server del dominio uniud.it fornisce l'indirizzo IP corrispondente al nome simbolico web.uniud.it.
 - Tale indirizzo viene poi inviato al resolver.