

Storia, principi e architettura di **Internet** e del **Web**

Slide originali di Simone Basso riviste e integrate da Juan Carlos De Martin

Risorse

- [History of the Internet](#) (Wikipedia)
- [Come funziona Internet](#) (opuscolo informativo tradotto dal Centro Nexa su Internet e Società)



Storia di Internet

Nei '50 si usavano reti telefoniche

- Un filo di rame collega un apparecchio telefonico a una centrale
- Un filo di rame collega tra loro due centrali
- **Fare il numero** significa ordinare alla rete di creare un **circuito elettrico** che mette insieme tutti i fili necessari per far parlare A con B
- si chiama: **commutazione di circuito**

evoluzione naturale:
trasmettere dati sempre con
commutazione di circuito
(o sue evoluzioni)

ideazione di
un modo diverso di
trasmettere dati

Anni '50: reti a pacchetto

- Tema di ricerca in USA, UK, Francia
- [Joseph Licklider](#) definisce il concetto di **commutazione di pacchetto**
 - un messaggio viene suddiviso in pacchetti
 - ogni pacchetto ha un indirizzo di destinazione
 - la rete fa da “sistema postale” elettronico
 - ogni pacchetto viaggia indipendentemente dagli altri

perché?

- Idea che potesse essere una tecnica trasmissiva **più efficiente**
- Idea che la rete potesse essere **più robusta**
- Idea che, a parità di prestazioni, potesse essere **più economica**

Reti telefoniche vs reti a pacchetti

- Reti telefoniche
 - **qualità garantita** (filo dedicato)
 - **non robuste** (se unico filo tagliato = non si parla)
- Reti a pacchetto
 - **migliore utilizzo risorse** (più pacchetti possono usare lo stesso filo perché l'indirizzo di destinazione permette di distinguere tra questi)
 - **qualità non garantita** (il filo è condiviso)
 - **robuste** (sono possibili più percorsi)

Scontro di visioni: TLC contro Informatici

Recepimento reti a pacchetti nei '50

- Il mondo **TELECOMUNICAZIONI** (che aveva grandi laboratori di ricerca e poteva orientare la ricerca):
 - aveva **seri dubbi che si potesse garantire qualità sufficiente usando le reti a pacchetti** (avevano in mente in primis le telefonate)
 - pensava che sarebbe stato **meglio investire in meccanismi per rendere più robuste le reti telefoniche** permettendo circuiti multipli
 - **ATT rifiuta di lavorare su commutazione di pacchetto**
(altro esempio della “legge” della soppressione del potenziale radicale?)

Anni '60: Paul Baran

- studia la **robustezza delle reti**
- conclude che una **rete robusta a guasti** (dalle cessoie di un pazzo, all'escavatore distratto all'attacco nucleare) **deve essere a pacchetti**
- **intendendo con questo:** la rete a pacchetti secondo Baran era la **tecnologia migliore per implementare una rete robusta ai guasti**

Rete a pacchetti meglio perché

- la rete è **robusta by design**
 - basta prevedere percorsi multipli e poi,
 - siccome ogni pacchetto viene smistato in modo indipendente dagli altri,
 - se c'è un guasto la rete cambia percorso
- questo **costa meno** di rendere robusta una rete telefonica (dove un numero di telefono indica un percorso fisico)

Nello stesso periodo in UK

- In UK anche Paul Davies, che aveva lavorato con Alan Turing, sta definendo il concetto di “commutazione di pacchetto” (v. <http://www.wired.com/2012/09/donald-davies/>)

Anni '60: Leonard Kleinrock (UCLA)

- **dimostra** che le **reti a pacchetti** sono **tecnicamente superiori** alla rete telefonica (o “rete a commutazione di circuito”)
- “[p]acket-switching provides **better bandwidth utilization and response times** than the traditional circuit-switching technology used for telephony, particularly on resource-limited interconnection links.” ([Wikipedia](#))

Visto che il settore privato (ATT)
rifiuta di studiare la
commutazione di pacchetto,
allora interviene il Governo



Sputnik-1

4 ottobre 1957

DARPA
Defense Advanced
Research Projects Agency

7 febbraio 1958

29 Ottobre 1969: nasce ARPANET

- ARPA = Advanced Research Projects Agency
è una **agenzia del governo USA che
finanzia progetti di ricerca avanzati** con
l'obiettivo **espandere la frontiera
tecnologica** andando oltre le esigenze
militari immediate
- ARPANET = prima rete a pacchetto
funzionante

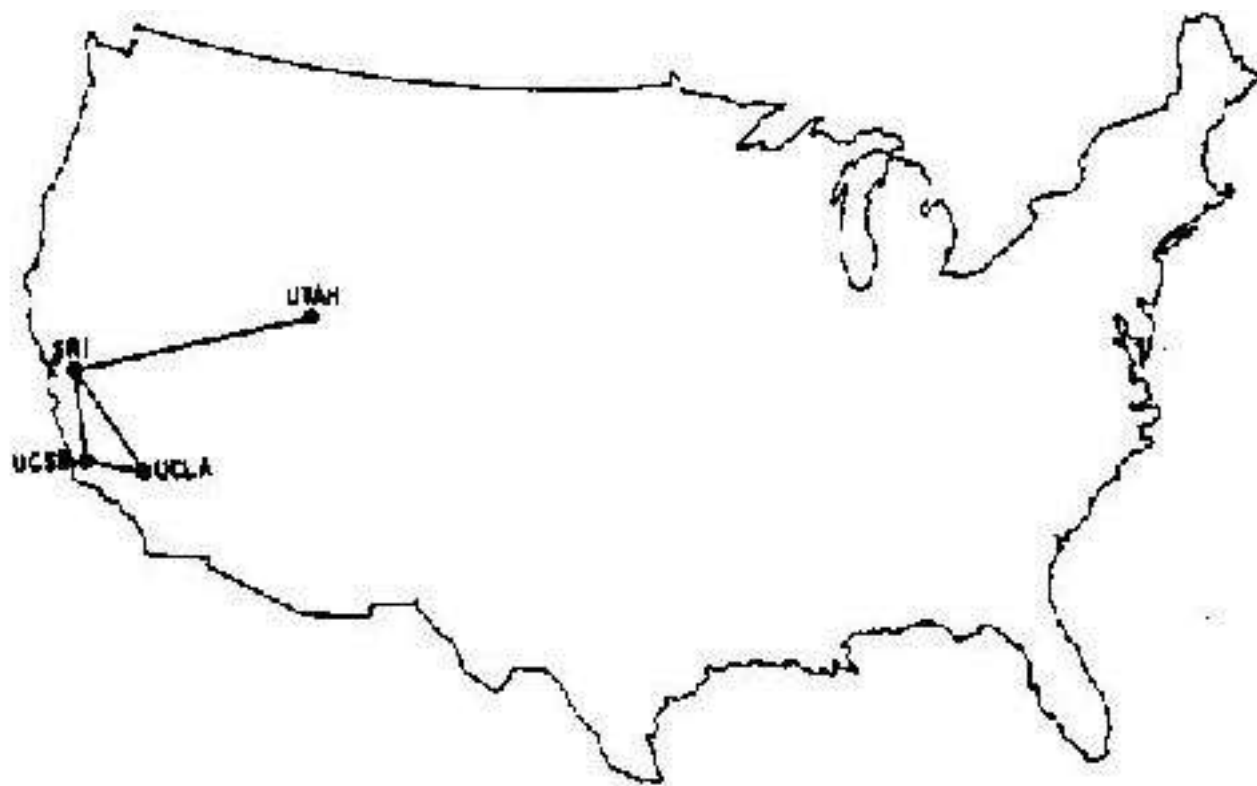
ARPANET

- un progetto gestito in maniera **collaborativa**, non rigidamente top-down
- **utenti** hanno voce in capitolo nelle scelte e propongono idee e applicazioni (come l'email)
- è qui che nasce e si radica l'idea di “**un popolo della rete**”

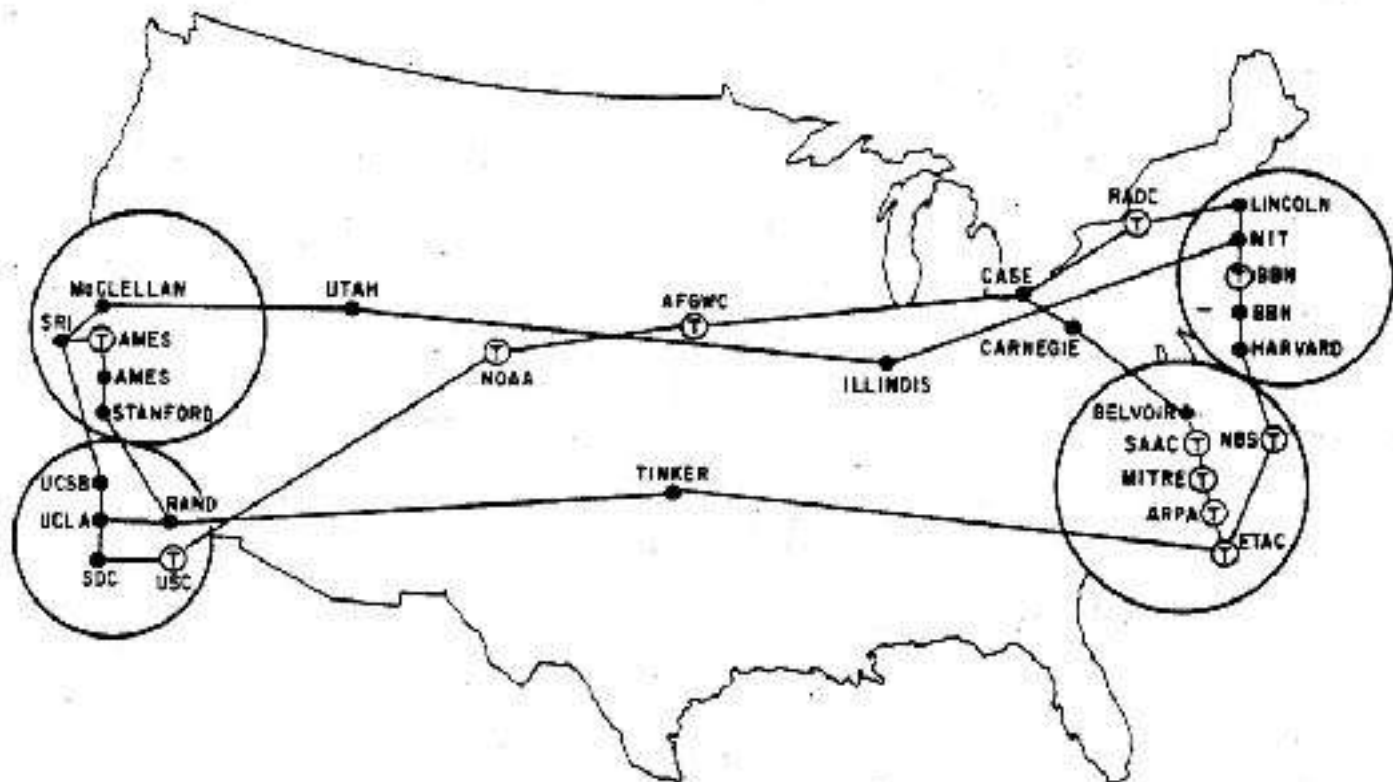
Storia di ARPANET

- Primi esperimenti tra UCLA (Los Angeles) e SRI (Stanford), problemi di stabilità
- I problemi vengono poi risolti e i nodi si moltiplicano molto rapidamente

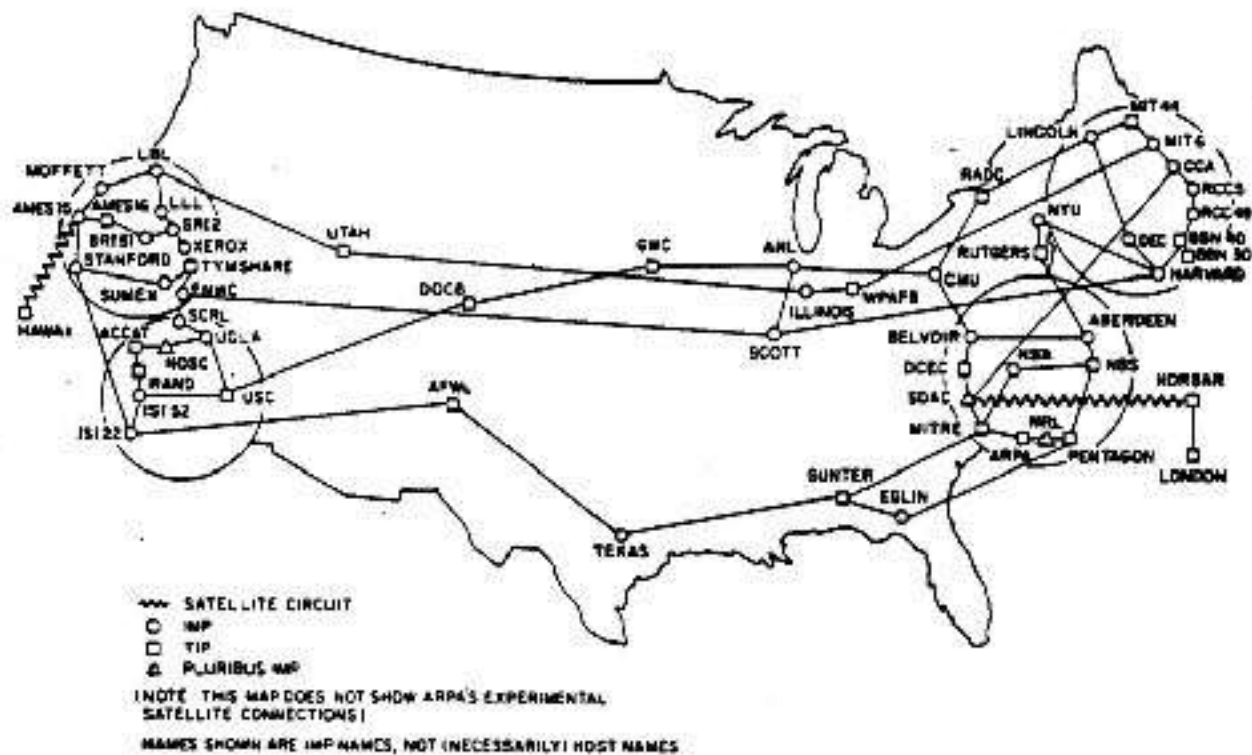
ARPANET, Dicembre 1969



ARPANET, Agosto 1972



ARPANET, Luglio 1977



ARPANET

- a proposito... certe applicazioni avanzate, come **comunicazioni vocali interattive in tempo reale**, le avevano in mente fin dall'inizio! (*video*)

Reti alternative a ARPANET

- X.25
 - rete a pacchetti con emulazione di circuito
 - standard ITU
 - inizialmente in UK poi tutto il mondo
- UUCP/USENET
 - “gruppi di discussione” usati ancora oggi
- Altre
 - Cominciava a esserci il **bisogno di una tecnologia per connettere queste reti tra loro**

Sempre più reti di computer:

- **satellitari, wireless, locali, ecc.**

1973: “Inter networking” -> Internet

- Bob Kahn, Vint Cerf, altri
- protocollo per **permettere a reti a pacchetti diverse di scambiarsi informazioni**
- idea chiave: **nascondere le differenze** tra le reti creando un protocollo comune a tutte

Internet Protocol (IP):
precisa le informazioni di base per rendere
possibile la commutazione di pacchetto

nessuna garanzia di qualità
“best effort”
perdite di pacchetti

**alcune applicazioni possono
tollerare qualche pacchetto perso
(VoIP, video streaming, ecc.)**

**altre applicazioni invece hanno
bisogno di tutti i dati (email,
trasferimento file, ecc.)**

**come fare a recuperare i
pacchetti persi quando
l'applicazione lo richiede?**

Transfer Control Protocol TCP

apre una connessione tra due computer

**ritrasmette gli eventuali pacchetti persi fino a
quando non arrivano tutti**

se i pacchetti arrivano fuori ordine, li mette in ordine

TCP usa (presuppone) IP

RFC = Request For Comments

- Documenti testuali che descrivono gli **standard usati su Internet**
- Emessi dalla **Internet Engineering Task Force (IETF)**, che rivedremo più avanti
- Uno dei più importanti standard sarà la **suite di protocolli TCP/IP** utilizzata dalla rete Internet a partire da inizio anni '80

Settembre 1981: TCP/IP

- Vengono specificati i protocolli che utilizziamo ancora adesso su Internet (**TCP/IP v4**)
- IP (**Internet Protocol**) si occupa di **instradare i pacchetti** verso la destinazione, è il **protocollo eseguito dagli apparati di rete** (detti “router”)
- I **computer connessi a Internet** (detti invece “end host”) utilizzano **TCP e UDP**
- **TCP/IP come lingua franca che permette a reti diverse tra loro di interconnettersi**

Punto su IP e TCP:

- **La rete usa il protocollo IP**
- La rete e i computer connessi in rete riescono a **parlarsi** perché **TCP** (o altri protocolli, come UDP) **internamente usano IP**
- **IP è la “lingua franca” parlata da tutti**

Note su TCP/IP: “La rete è stupida”

- Siccome IP è più semplice di TCP e di UDP e siccome la rete parla solo IP, si usa dire nel mondo Internet che
 - la rete è stupida
 - l'intelligenza sta negli end host (ossia ai bordi)
- Questa caratteristica del design di Internet si chiama **principio end-to-end**

Note su TCP/IP: “I router costano poco”

- Siccome la rete è stupida **non è troppo difficile e costoso costruire router efficienti** che smistino velocemente il traffico
- Chiave del successo di Internet:
 - i router **costano poco** (rispetto alle centrali della rete telefonica)
 - i router smistano **molto velocemente** il traffico
- Questo **spiega** (in parte) **come mai** alla fine **si sono affermate le reti a pacchetto**

1 Gennaio 1983: si passa a TCP/IP

- Data a partire dalla quale ufficialmente si passa dai protocolli ARPANET ai protocolli di Internet (ossia al TCP/IP)
- Il **Governo USA** praticamente dice che da ora in poi **acquisterà soltanto apparati che sono conformi alle specifiche TCP/IP (il grande potere del 'procurement' governativo)**

1980 - 1985: NSF finanzia l'upgrade

- Sistemata la questione dei protocolli universali per parlarsi in rete, si comincia a usare Internet
- Si rende necessario **aggiornare l'infrastruttura fisica di ARPANET per avere più capacità**
- Aggiornamento finanziato dalla **National Science Foundation (NSF)**
- Da 56 kb/s in dorsale nel 1986 a 45 Mb/s in dorsale nel 1991 (dove “in dorsale” significa nei collegamenti tra quelli che sono i router più importanti)

Fine anni '80-Inizio anni '90

- Molte tecnologie di rete in competizione tra loro: IP è solo una delle proposte, da molti considerata non “seria”...
- IBM, Microsoft, Novell, Digital, ecc. hanno le loro tecnologie di rete
- Gli operatori telefonici investono moltissime risorse in tecnologia **ATM** - Asynchronous Transfer Mode, uno standard ANSI e ITU per comunicazioni integrate, che poi verrà soppiantato da TCP-IP
- Ancora a inizio anni '90 ci sono importanti imprenditori italiani che ritengono che “Internet non abbia futuro”

COMPUSERVE - PRODIGY

- Reti PROPRIETARIE in concorrenza con INTERNET

1995+: apertura commerciale

- Nel 1995 il **governo USA** rende **possibile usare Internet per scopi commerciali**
- **Prima del 1995 non era possibile** perché era stata **finanziata con soldi pubblici**
- Da questo momento in poi diventano sempre più importanti gli ISP (**Internet Service Providers**) ossia operatori commerciali che vendono il servizio di accesso a Internet

Principi di Internet

Quattro principi base di Internet:

1. Inter-networking
2. Principio end-to-end
3. Best effort (= no garanzie di qualità)
4. Net neutrality

Inter-networking

- Internet nasce per **unire tra loro reti diverse** come ARPANET e X.25
- Basata su una **lingua franca**
- Ai giorni nostri Internet viene usata per unire tra loro reti diversissime come:
 - **4G** ossia le reti cellulari
 - **WiFi**
 - reti in **fibra**, etc

Principio end to end

- La rete implementa protocolli più semplici così si può andare più veloce e costa meno implementare una rete robusta
- Le applicazioni sono infatti molto complesse se confrontate con il protocollo IP
- Questo significa che **la complessità sta ai bordi della rete** (“end to end”)

E2E e innovazione

- esplosione di innovazione
- abilitata dal fatto che “basta parlare IP”, senza dover chiedere permesso a nessuno (sicuramente non a chi possiede la Rete).

Best effort = no garanzie di qualità

- Siccome la rete è semplice **non cerca di fare cose strane con i pacchetti**
- Semplicemente si limita a **smistarli nell'ordine in cui gli arrivano**
guardando solo l'indirizzo di destinazione
- Se c'è **troppo traffico** i pacchetti possono subire **ritardi o venir scartati**

Net neutrality

- Il fatto che Internet sia stata progettata e implementata seguendo i **principi end-to-end e best effort**
- Ha generato l'**aspettativa che questi principi continuino a valere** come “legge naturale” di Internet
- Ossia ha generato l'**aspettativa che la rete sia neutrale** e faccia solo il suo lavoro di smistamento dei pacchetti
- Anche se oggi **gli ISP hanno la tecnologia per bloccare o rallentare il traffico** di BitTorrent, Skype, etc.
- Questo ha generato numerosi dibattiti e controversie come ad esempio [BitTorrent vs Comcast](#)

Net neutrality

- Situazione USA
- Situazione EU
- Ora si parla molto di “**zero rating**”

Protocolli applicativi notevoli

- **SMTP** e **IMAP**: basati su TCP implementano l'invio e la ricezione della **posta elettronica**
- **HTTP**: basato su TCP implementa il **web**
- **uTorrent**: basato su UDP, usato per il **file sharing**
- **Skype**: basato su UDP, si usa per fare telefonare
- **DNS**: basato su UDP, si usa per mappare **indirizzi Internet** su **nomi a dominio** (ne parliamo tra poco)

Applicazioni notevoli

- **Mozilla Thunderbird**: client di posta che implementa i protocolli **SMTP** e **IMAP**
- **Mozilla Firefox**: web browser che implementa il protocollo **HTTP**
- **BitTorrent**: programma di file sharing che implementa il protocollo **uTorrent**
- ...

Indirizzi Internet

Indirizzi Internet v4 e v6

- La versione più utilizzata di **TCP/IP** è la **versione 4** (v4), che ha indirizzi fatti così:
 - a.b.c.d
 - dove a, b, c, d sono numeri compresi tra 0 e 255
- Per esempio: **130.192.91.211**
- Per completezza: **esiste (e si sta lentamente affermando) anche la versione 6** (v6) di TCP/IP
- Gli indirizzi IPv6 (molto più numerosi di quelli IPv4) li vedrete nei corsi di reti

DNS

Domain Name System

DNS: motivazione

- Gli indirizzi IP erano difficili da ricordare in quanto numeri (130.192.91.211)
- Per questo motivo sono stati inventati i **nomi di dominio**, ossia per aiutare gli umani a ricordarsi (esempio: diderot.polito.it)

DNS: funzionamento

- Il protocollo DNS richiede dei **server DNS** che di solito sono **forniti dal vostro ISP**
- Viene inviata una **richiesta DNS** che contiene un nome a dominio (es., x.org)
- Ricevete una **risposta DNS** dal server contenente l'indirizzo corrispondente (es., 18.7.25.161) al nome che avete chiesto

Tra Internet e Web

Architettura “client-server” (cliente-servitore)

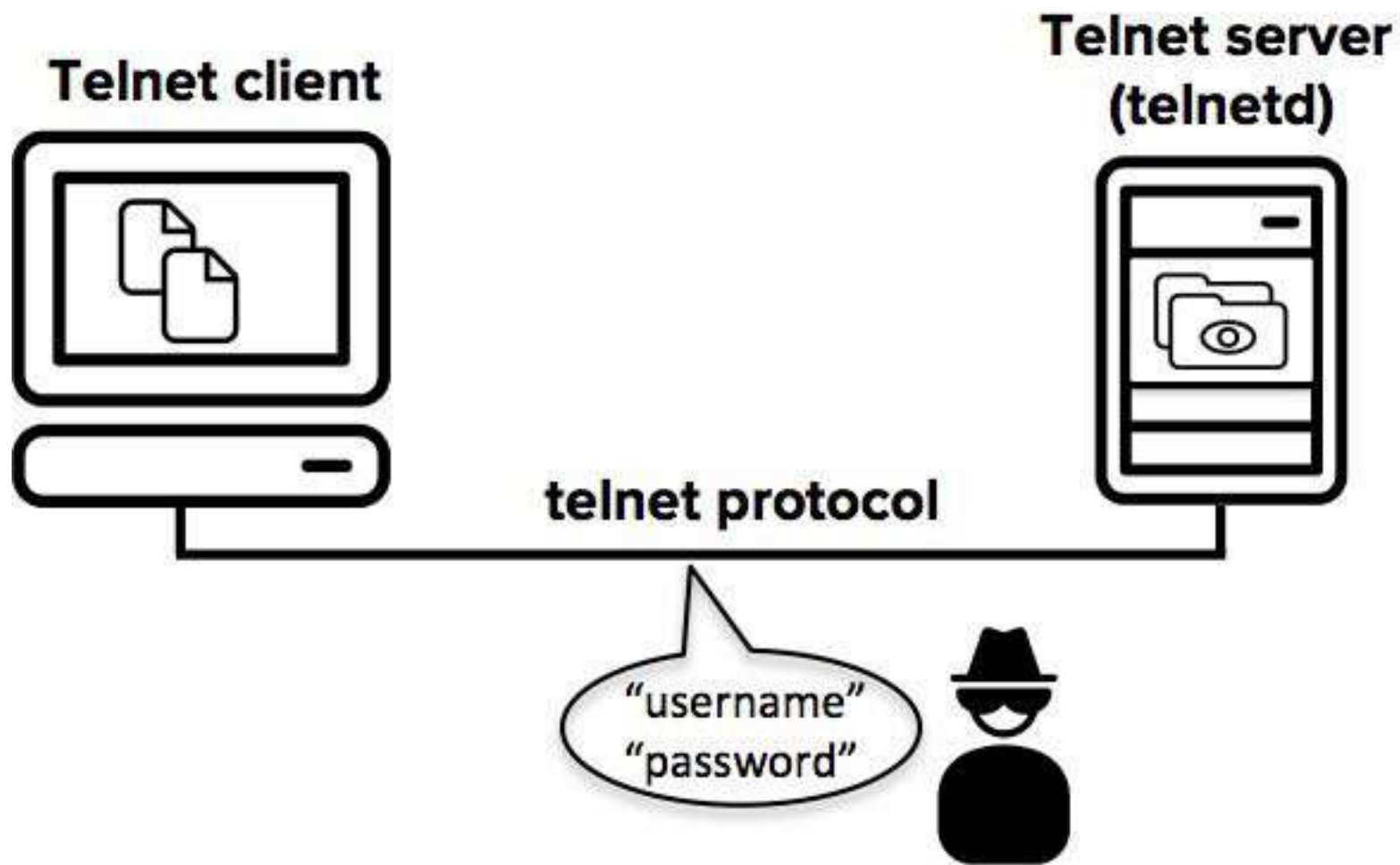
TELNET, FTP, ecc.

- interfaccia **testuale**
- occorre sapere **dove andare a cercare** le cose
- si scaricavano i file sulla fiducia (in base al nome del file)
- occorre avere il **programma adatto** per aprire il file (a seconda del formato)

Telnet: protocollo di rete per connessione remota a un computer (RFC 854 e RFC 855)

“Terminale remoto” (**teletype network**)

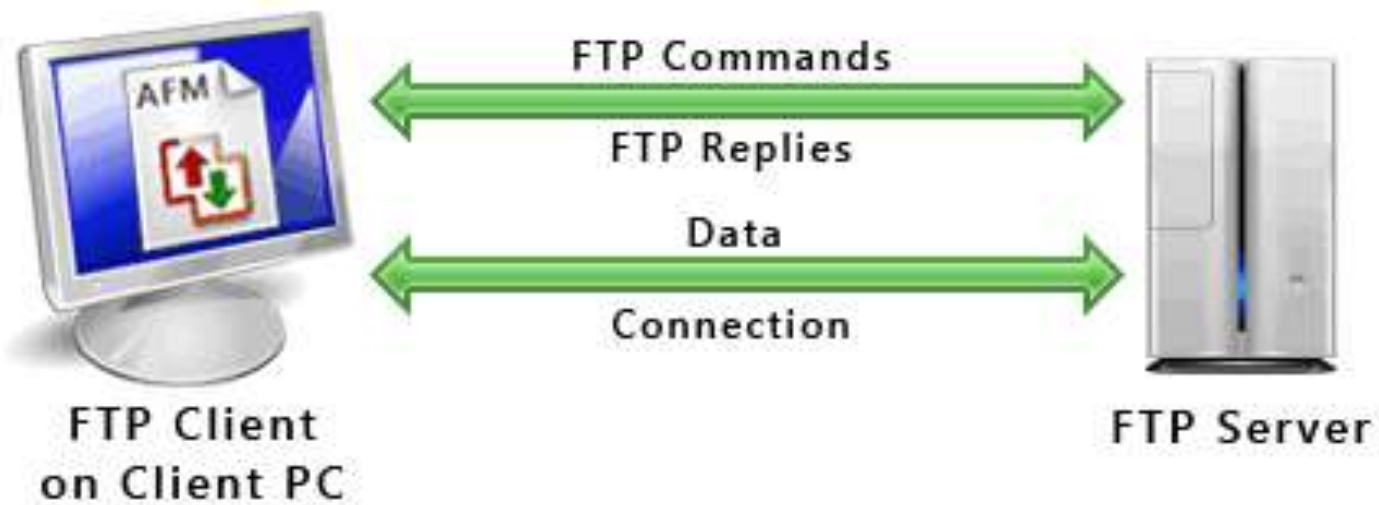
Client-server, usa TCP



FTP: protocollo di rete per trasmissione di file (risale al 1971, ultima specifica è RFC 959)

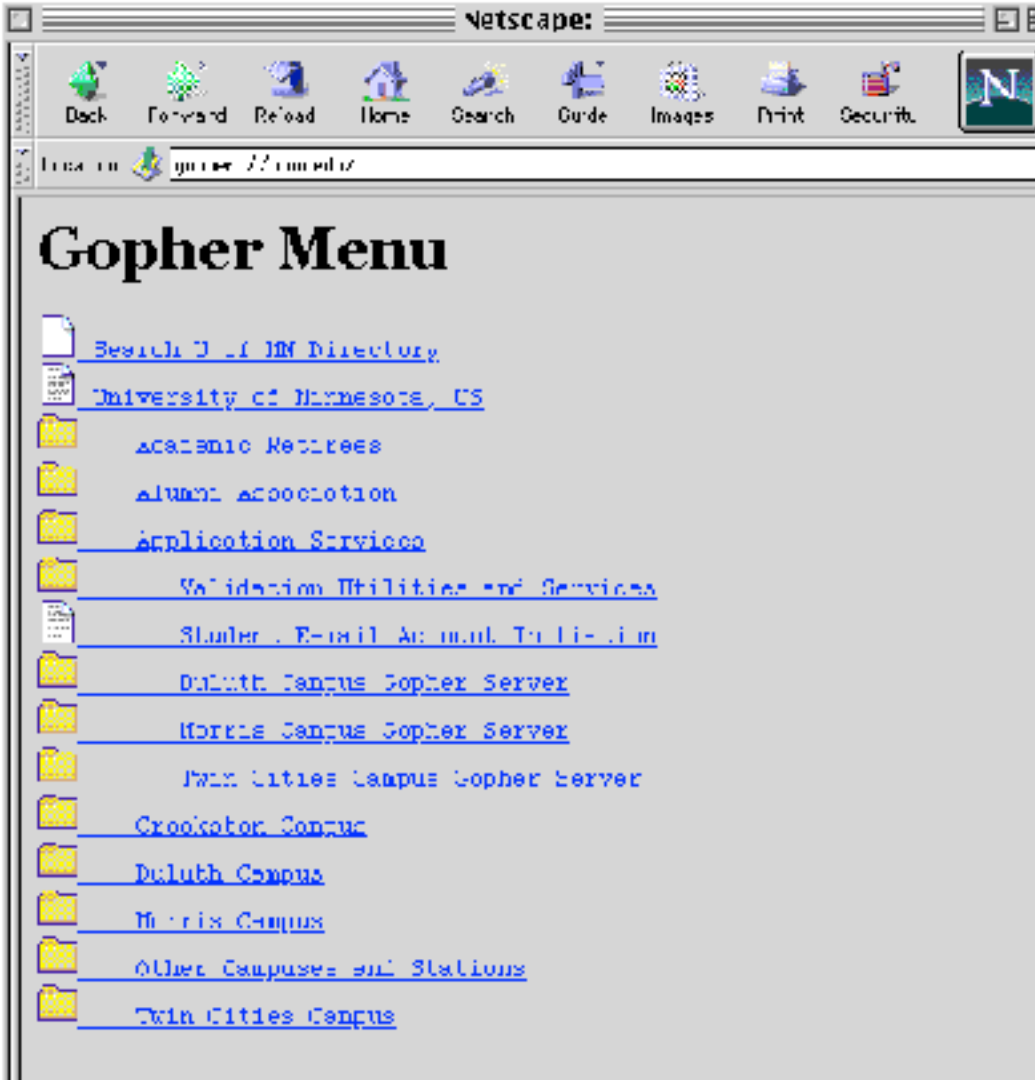
“File Transfer Protocol”

Client-server, usa TCP



GOPHER

- <https://www.minnpost.com/business/2016/08/rise-and-fall-gopher-protocol>



II World Wide Web (WWW)

Storia

- **Tim Berners-Lee** (e altri) al CERN
- **Problema:** come condividere documenti tra gruppi di ricerca che usavano computer, programmi e formati diversi?

Storia

- Ci volevano due cose:
 1. un **linguaggio comune per descrivere i documenti** (tra l'altro rendendo possibili link tra loro)
 2. un modo standard per **“pubblicare”** tali documenti sulla rete

PREMESSA

- Non era la prima volta che ci si poneva il problema. Anzi, erano già state sviluppate soluzioni che però non avevano attecchito, come per esempio **GOPHER**

HTML

HyperText Mark-up Language

Testo con all'interno comandi per corsivo, sottolineato, a capo, link, ecc.
Al momento alla **versione 5** (2014).
Prima versione ufficiale nel 1993.

HTTP

HyperText Transport Protocol

Protocollo che specifica come richiedere un documento da un computer che lo possiede (il server) da parte di un computer che lo vuole (il client). Versione 1.0 rilasciata nel 1991 (e poi standardizzata da RFC 1945 del 1996).

HTTP

HyperText Transport Protocol

E' un'**architettura 'client-server'**: un computer (il 'servitore') ha dei documenti. Quando questi interessano ad un altro computer (il 'cliente'), il cliente invia una richiesta al servitore, che glieli manda. Questo scambio avviene tramite **messaggi HTTP**, che a loro volta usano TCP per garantire affidabilità, ecc. della comunicazione.

Il servitore è detto **web server**. Si tratta di un computer connesso a Internet che a bordo ha sia i **contenuti** che vuole rendere pubblici sia **un software apposito** per renderli disponibili tramite HTTP (come [apache.org](https://en.wikipedia.org/wiki/Comparison_of_web_server_software) o IIS di Microsoft, v. https://en.wikipedia.org/wiki/Comparison_of_web_server_software).

Il cliente è qualsiasi computer (inclusi smartphone) connesso a Internet che ha a bordo **un apposito software** che 1) implementa un cliente HTTP e 2) è in grado di visualizzare contenuti espressi in HTML, ovvero un 'browser' o 'navigatore Web' (come Firefox, Chrome, Safari, ecc. v. https://en.wikipedia.org/wiki/Comparison_of_web_browsers)

Tim Berners-Lee sviluppa **sia HTML** (basandosi su iniziative precedenti), **sia HTTP** e poi pubblica entrambi con una prima implementazione software (usando un computer NeXT)

**Successivamente (nel 1994)
viene creato il **WWW
Consortium**, ovvero **W3C**, che
dal standardizza i protocolli e
linguaggi del Web (non di
Internet). V. <http://w3c.org>.**

TBL donò
sia HTML, sia HTTP
al **pubblico dominio**

era già successo
con ARPANET/INTERNET

TBL avrebbe potuto diventare molto ricco,
ma probabilmente il Web non sarebbe
mai diventato l'incredibile successo che,
grazie alla sua generosità, è diventato...

Universal Resource Locator (URL)

- Una URL è un indirizzo sul web e indica una risorsa (una pagina web, un video, una foto, un file di testo) su un certo server
 - <http://x.org/antani> significa:
 - la risorsa identificata da /antani
 - sul server x.org
 - usando il protocollo HTTP

Il browser

- Dato <http://x.org/antani> il browser
 - usa il **DNS** per risolvere x.org e ottenere il corrispondente indirizzo 18.7.25.161
 - usa **TCP** per connettersi al server 18.7.25.161 sulla porta 80 (quella usata dal web)
 - usa **HTTP** (che a sua volta usa TCP) per mandare al server una richiesta per la risorsa /antani
 - chiude la connessione TCP
 - vi fa vedere la risorsa /antani (che potrebbe essere una foto, un video, una pagina web)

“A Read-Write Web”

Relazione tra Internet e web

- Il web è implementato da un protocollo che si chiama HTTP e che si basa su TCP
- In altre parole questo significa che possiamo vedere **il web** come una **applicazione distribuita che gira su Internet**
- Ci sono dei server web con dei contenuti e dei client che usano HTTP per scaricarli

Esempio

Corso per corrispondenza

il corso per corrispondenza (dispense, esami, ecc.) è l'**applicazione** (p.es., il WWW)

il sistema postale che trasporta le buste che contengono il corso è **Internet**

Reazioni?

COMPUTER 'WEB' TO CHANGE BILLIONS OF LIVES (YEAH, RIGHT)

A BRITISH computer geek's brain-wave could be one of the greatest inventions ever, it was claimed last night.

Tim Berners-Lee, 35, has enabled computer users to see documents and pictures made available by others in "cyberspace".

He uses the "Internet" system, which so far only

By DOT COMME

links academics but could eventually include anyone.

Berners-Lee, who works at a nuclear research base near Geneva, calls his idea the "World Wide Web".

One scientist said: "This could be huge. The idea of strangers worldwide sharing ideas instantly is mind-boggling." But another sneered: "They said Sinclair's C5 would change the world. Now you'd struggle to give one away."

Riddle of 'E' mail - Page 5



Web feat... Berners-Lee

Conclusioni

personal computer
+
Internet
+
WWW

**Un'architettura complessiva che concede
molta autonomia e controllo agli utenti**

**Un'architettura che però negli ultimi 15
anni ha subito un progressivo
scivolamento verso nuove forme di
centralizzazione e intermediazione**