

Viaggiare informatici

Autostrade della ricerca

di Massimo Carboni

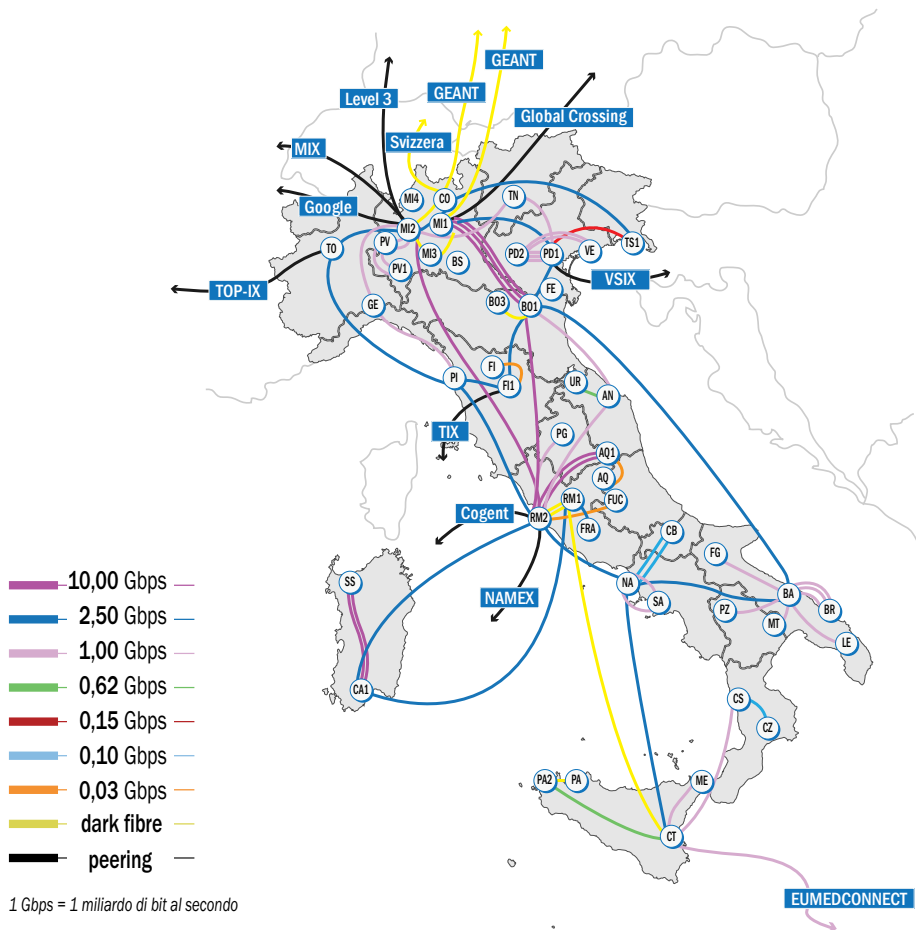
a.
Come una rete da pesca, internet consiste di un insieme di nodi e collegamenti ed è resistente a possibili danneggiamenti esterni, perché è stata studiata in modo che ogni nodo possa continuare a elaborare (e trasmettere) dati anche nel caso i nodi vicini siano inutilizzabili.



La parola “rete” per molti di noi è legata al ricordo di quando eravamo bambini, intenti a pescare sulla riva di un torrente durante una calda estate. Ripensando alla nostra infanzia, una rete può essere rappresentata in modo semplice come un insieme di nodi e fili. Oltre alla leggerezza, la caratteristica più importante di una rete è la resistenza, che resta invariata anche nel caso di danneggiamento parziale. Queste semplici proprietà sono anche alla base della più famosa delle reti: internet. Di fatto internet è l'insieme di reti singole, che collegate tra di loro costituiscono la più grande rete di comunicazione esistente al mondo. Può essere rappresentata come un insieme di nodi e collegamenti che, unitamente a regole d'instradamento, permettono a oltre due miliardi di persone di comunicare, scambiandosi informazioni in tempo

reale. I nodi sono gli strumenti con cui accediamo alla rete come pc, smartphone o server, i collegamenti sono le “strade” attraverso cui viaggiano le informazioni e possono essere, ad esempio, circuiti di rame o fibre ottiche. I vari nodi hanno funzioni diverse: alcuni sono nodi terminali, da cui si accede alla rete, come il computer, altri hanno la funzione di veicolare il traffico all'interno della rete e si chiamano *router* (dall'inglese “instradatore”). La loro funzione assomiglia a quella svolta dai controllori di volo che gestiscono le diverse rotte degli aeroplani in modo da evitare quelle più saturate, il maltempo o le collisioni. Così i router si scambiano le informazioni sullo stato di funzionamento dell'intero sistema modificando i percorsi secondo regole prestabilite. La rete internet è costituita da un insieme di reti indipendenti tra loro,

ognuna delle quali viene identificata mediante un codice numerico. Alla base del suo funzionamento c'è un protocollo di comunicazione che disciplina l'invio e l'organizzazione dei bit di informazione tra computer diversi e che si chiama *Tcp/Ip*, acronimo di *Transmission Control Protocol/Internet Protocol*. In realtà si tratta dell'unione di due protocolli. *Tcp/Ip* è indipendente dalla rete di comunicazione su cui si poggia, sia essa radio (Wi-fi, Gsm, Umts), cavo (fibra o rame) o satellitare, e si integra facilmente con diverse tecnologie hardware in un'unica struttura logica di comunicazione. Le caratteristiche e l'unione di questi due protocolli hanno determinato il successo e la diffusione di internet. Abbiamo detto che a ciascuna rete è associato un codice numerico, ma questo sistema, pur funzionando ottimamente per lo scambio



b. Schema della rete Garr. Sono evidenziate a colori le velocità di trasmissione dei dati tra un nodo e un altro. Una *dark fibre* è una fibra ottica che non trasporta servizi forniti da un operatore di telecomunicazioni, ma che viene "illuminata" mediante apparati acquistati e gestiti dal Garr. Con il termine *peering* si considera il rapporto che due organizzazioni stabiliscono al fine di scambiare traffico dati, sia a titolo oneroso che gratuito. Ad esempio, con Google (fornitore di contenuti su internet) è a titolo gratuito, mentre con Level3/Global Crossing/Cogent è oneroso.

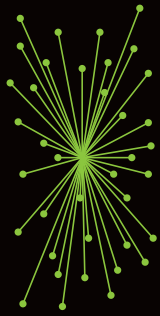
d'informazione, non è intuitivo e dunque funzionale per gli utenti che normalmente usano la rete. Al fine di semplificare le cose sono state introdotte delle agende di indirizzi Ip associate a nomi di più facile memorizzazione: il meccanismo in questione si chiama *Domain Name System (Dns)* ed è interrogabile da ogni nodo della rete.

La comunità accademica, per sua natura, è abituata a lavorare in gruppi di ricerca distribuiti a livello nazionale e internazionale che, da sempre, hanno avuto la necessità di scambiarsi informazioni in modo rapido e semplice. Per rispondere a questa esigenza i ricercatori hanno sviluppato una rete dati che collegasse i diversi gruppi di ricerca a partire dal progetto di rete realizzato per il Dipartimento della difesa Usa (Arpanet) (vd. approfondimento, ndr). In Italia le iniziative accademiche dei primi anni '80 (vd. p. 11, ndr) si sono consolidate rapidamente con la costituzione del Gruppo di Armonizzazione delle Reti della Ricerca (Garr). Il Garr aveva il compito di fare sinergia tra i diversi enti pubblici di ricerca italiani come il Cnr, l'Enea, l'Infn e le università, rendendo disponibile una rete di telecomunicazioni comune a tutti. La capacità di rete del Garr era decisamente lontana dagli attuali valori, ma le basi tecniche, così come quelle funzionali, erano del tutto simili a quelle in uso oggi. Dai primi passi della rete ad oggi sono passati 30 anni; attualmente la rete Garr, gestita da un decennio dal Consortium Garr, collega tutti gli enti pubblici di ricerca, le università, gli Istituti di Alta Formazione (Afam), gli Istituti di Ricovero e Cura a Carattere

Scientifico (Irccs), le Biblioteche Nazionali e molte altre strutture di ricerca e istruzione italiane. Complessivamente ne fanno parte oltre 500 sedi distinte, ognuna delle quali è a sua volta una rete composta da diversi nodi e collegamenti e fornisce l'accesso ai servizi di rete dati a oltre due milioni tra ricercatori, professori, dipendenti e studenti.

Anche le sedi Infn sono connesse alla rete Garr e hanno una capacità di accesso compresa tra i 0,1 Gbps e i 3 Gbps, con una sola sede connessa a 30 Gbps: il Cnaf di Bologna (vd. fig. b). Con l'evoluzione attuale di rete nota come Progetto Garr-X anche le sedi Infn potranno disporre di accessi alla rete Garr a 10 Gbps, mentre il Cnaf sarà la prima sede italiana connessa a 100 Gbps alla rete accademica di ricerca. Mediante queste straordinarie possibilità i fisici hanno costruito negli anni un'infrastruttura di calcolo distribuita, nota come Wlcg (Worldwide Lhc Computing Grid), in grado di elaborare i dati provenienti dagli esperimenti di fisica presenti sull'acceleratore di particelle Lhc (vd. p. 21, ndr). Il Cnaf è uno degli undici centri di elaborazione al mondo ed è in grado di immagazzinare su disco oltre 10 petabyte, ovvero dieci milioni di gigabyte, e su nastro una quantità tre volte superiore (il byte, pari a 8 bit, è l'unità di misura della capacità di memoria di un computer, ndr). Il numero di nodi di calcolo presenti al Cnaf è dell'ordine di 100.000. Il sistema è utilizzabile da differenti comunità di ricercatori, sia fisici che di altre discipline. Ancora una volta ci troviamo di fronte alla nostra rete di nodi, ognuno con la

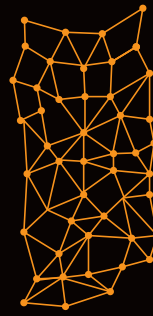
In principio fu Arpanet



centralizzata



decentralizzata



distribuita

La prima versione di internet si chiama Arpanet e nasce alla fine degli anni '60 negli Stati Uniti da una collaborazione fra il Dipartimento della Difesa, che si occupava di sviluppare nuove tecnologie militari attraverso la Defense Advanced Research Projects Agency (Darpa), e le principali università americane. In piena guerra fredda, l'esigenza era quella di poter disporre di uno strumento di comunicazione digitale in grado di operare anche in caso di distruzione di una sua parte. Fino a quel momento le tipologie di reti erano due: le reti *centralizzate*, con tutti i collegamenti che conducono a un unico nodo

centrale, e le reti *decentralizzate*, dove i collegamenti si dipanano da un certo numero di nodi principali collegati tra loro. Gli ingegneri che lavoravano ad Arpanet trovarono un terzo approccio: la rete *distribuita* (vd. fig. 1), composta da moltissimi nodi collegati gli uni agli altri in modo ridondante. Il cosiddetto *livello di ridondanza* indica il grado di connettività tra i nodi di una rete e un buon livello di sicurezza sarebbe stato garantito da una rete con livello di ridondanza tre o quattro (ogni nodo può essere raggiunto attraverso tre o quattro link diversi). La seconda idea sviluppata dagli ingegneri di Arpanet

1.

Le tre tipologie di rete: centralizzata, decentralizzata e distribuita. La prima versione di internet è stata concepita negli Stati Uniti come una rete distribuita dal nome Arpanet.

fu ancora più rivoluzionaria: riguardava la tecnica di trasmissione dei dati, frammentati in piccole porzioni, che successivamente fu denominata *packet switching*. I dati, prima di essere inviati, vengono scissi e spediti autonomamente a destinazione, dove vengono ricomposti.

Nel 1969, l'anno passato alla storia per la prima volta dell'uomo sulla Luna, Arpanet collegava quattro nodi: l'Università della California di Los Angeles, l'Università della California di Santa Barbara, lo Stanford Research Institute di Stanford e l'Università dello Utah. Nel 1972 fu inventata la posta elettronica. Con l'introduzione dei protocolli che disciplinano lo scambio di pacchetti di informazione, la rete Arpa varcò l'oceano e vi si collegarono prima la Norvegia, poi il Regno Unito e, nel 1986, l'Italia (vd. p. 13, ndr). All'aumentare progressivo del numero di utenti e della quantità di informazione scambiata tra i diversi nodi corrisponde una diminuzione della sicurezza e della capacità di controllo sulla rete. Così nel 1983 il Dipartimento della Difesa separò da Arpanet una parte della rete, che riservò esclusivamente a scopi militari (Milnet) e la restante parte, ormai chiamata da tutti internet, continuò a servire prevalentemente il mondo accademico della ricerca fino all'invenzione del World Wide Web. [E.C.]

propria specializzazione, che può essere l'elaborazione, l'immagazzinamento e la movimentazione dei dati o la presentazione dei risultati. Questo modello è di fatto la nuova frontiera che la comunità scientifica sta sviluppando da oltre dieci anni e che oggi comincia a entrare nella vita di tutti i giorni con il nome evocativo di "Cloud".

c.

Il Tier 1 di Bologna, il nodo italiano della rete Grid di Lhc al Cnaf. I ricercatori indossano delle cuffie a causa dell'eccessivo rumore prodotto dai computer.



Biografia

Massimo Carboni lavora presso i Laboratori Nazionali di Frascati. Esperto di sistemi informatici, dal 1999 lavora per conto dell'Infn all'interno del Consortium Garr, dove ricopre l'incarico di responsabile dell'infrastruttura di rete e del progetto di rete Garr-X.

Link sul web

<http://www.garr.it/>

<http://lcg.web.cern.ch/>