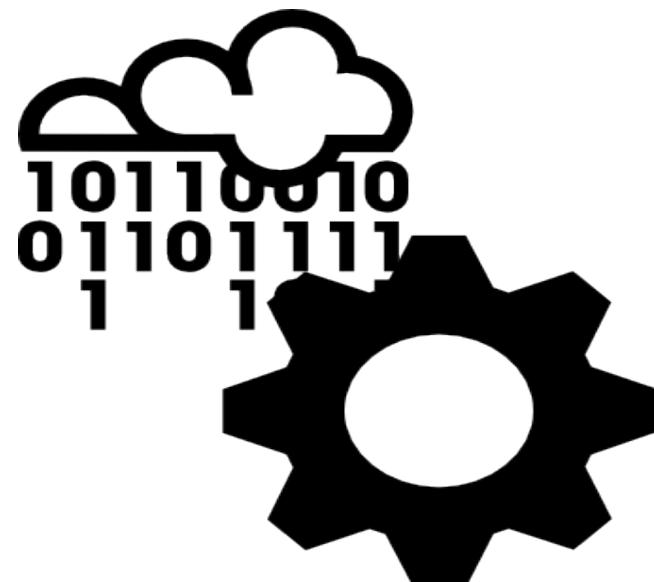
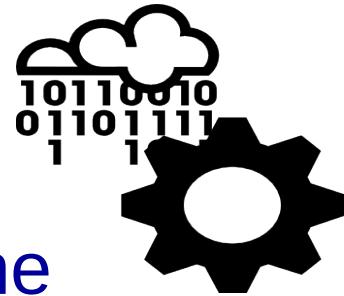


Libero Software per Libero Internet!



Cristiano Longo
opendatahacklab
Catania, 2017

COPYRIGHT: questa presentazione e tutti i materiali in essa contenuti sono rilasciate con licenza CC BY SA 4.0 (vedi <https://creativecommons.org/licenses/by-sa/4.0/>) ad eccezione dei contenuti per i quali è espressamente indicata un'altra licenza



A world where knowledge creates power for the many, not the few.

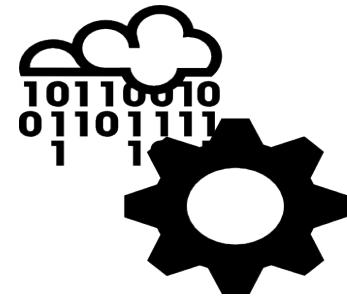
This is the world we choose.



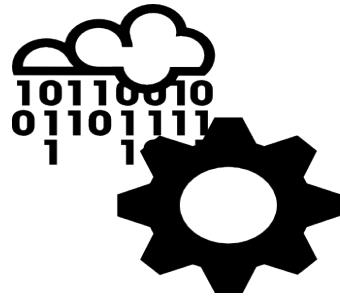
OPEN KNOWLEDGE

Open Knowledge Foundation (<https://okfn.org/>)
By Open Knowledge (OKFN) [CC BY 3.0], via Wikimedia Commons

power *not* the few



Detenere grosse moli di informazioni crea posizioni di rendita e di predominio.

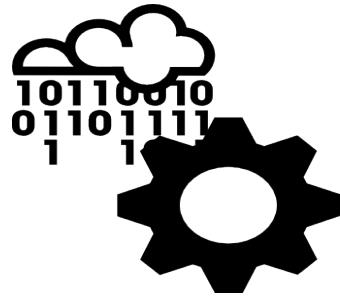


Detenere grosse moli di informazioni crea posizioni di rendita e di predominio.

Google Inc detiene grosse moli di informazione, ad esempio relativamente alle attività commerciali.

Per accedere a queste informazioni siamo costretti ad usare i servizi che mette a disposizione, venendo assoggettati a profilazione.

Le attività commerciali possono usare dei servizi Google per fare pubblicità mirata.



Detenere grosse moli di informazioni crea posizioni di rendita e di predominio.

Google Inc detiene grosse moli di informazione, ad esempio relativamente alle attività commerciali.

Per accedere a queste informazioni siamo costretti ad usare i servizi che mette a disposizione, venendo assoggettati a profilazione.

Le attività commerciali possono usare dei servizi Google per fare pubblicità mirata.

E' complesso avviare nuove imprese in presenza di competitor monopolistici.

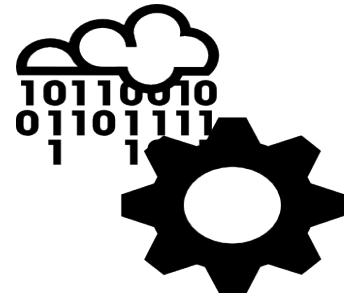
power to the people !



La *liberazione* dei dati rende impossibile la creazione di monopoli di informazione e il consolidamento di posizioni di rendita e di potere.

Abilità a produrre *valore* per l'intera comunità, ad esempio con la creazione di servizi e studi a partire dai dati.

Permette di esprimere il massimo potenziale dai dati riutilizzandoli e riaggregandoli.



HACKSPACE
CATANIA



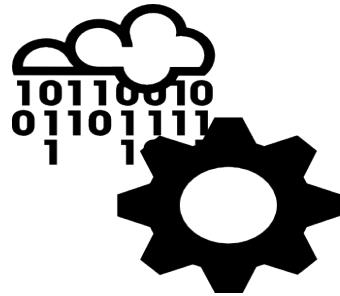
#opendatasicilia



UNIVERSITÀ
degli STUDI
di CATANIA

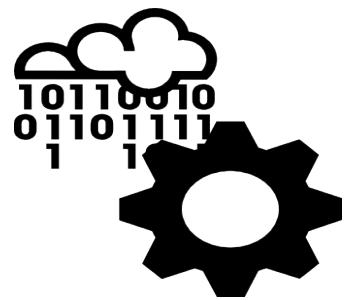


A servizio della comunità, produce studi e strumenti software per la generazione e l'utilizzo di open data.



I dati si dicono rilasciati come *open data* se:

- possono essere liberamente ottenuti tramite internet;
- sono riutilizzabili a qualunque fine con al più la limitazione di citare la fonte;
- vengono forniti con modalità che li renda elaborabili dalle macchine, permettendo quindi aggregazioni e mix con altri dati.



INTERNET

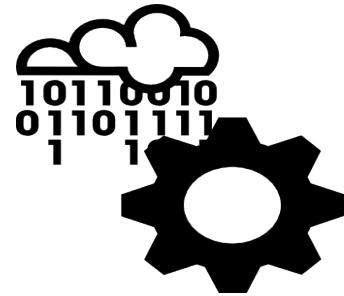


Galactic Network: “*a globally interconnected set of computers through which everyone could quickly access data and programs from any site.*”

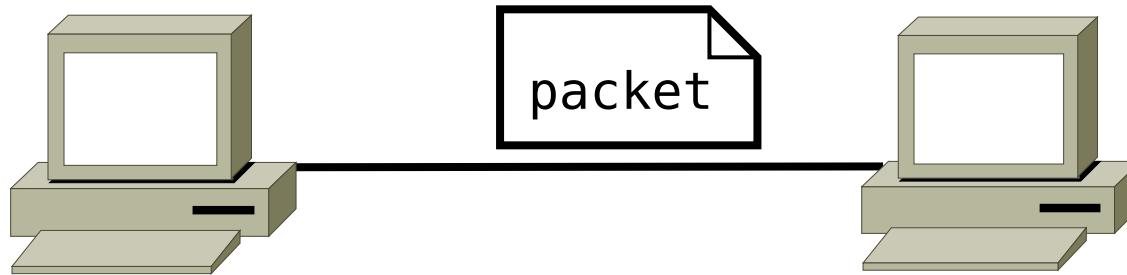
1962 - J.C.R. Licklider
Direttore del computer research program al
Defense Advanced Research Projects Agency (DARPA)

Foto da http://www.thocp.net/biographies/licklider_jcr.html

commutazione di pacchetti



Gli elaboratori si scambiano informazioni raggruppate in unità di dimensione fissata chiamate *pacchetti*.



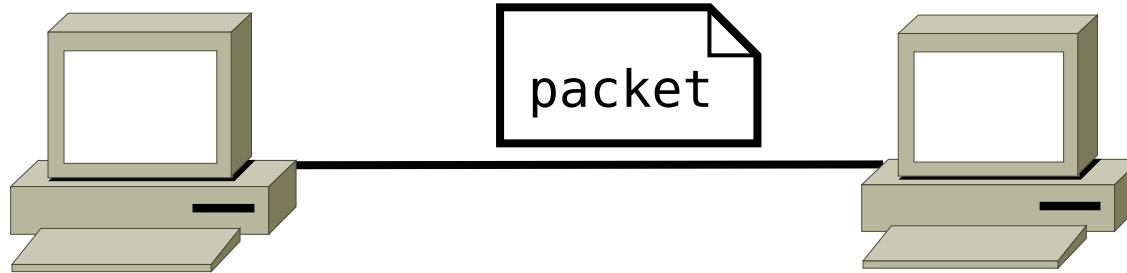
1961 - "Information Flow in Large Communication Nets", L. Kleinrock, MIT;

<http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet>

commutazione di pacchetti



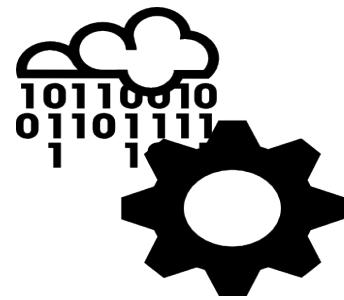
Gli elaboratori si scambiano informazioni raggruppate in unità di dimensione fissata chiamate *pacchetti*.



1961 - "Information Flow in Large Communication Nets", L. Kleinrock, MIT;

1969 - ARPANET: prima connessione e scambio di pacchetti tra due computer rispettivamente all'Università della California (UCLA) e allo Stanford Research Institute (SRI);

<http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet>



Una rete di computer è costituita da un insieme di elaboratori, detti nodi, collegati tra loro

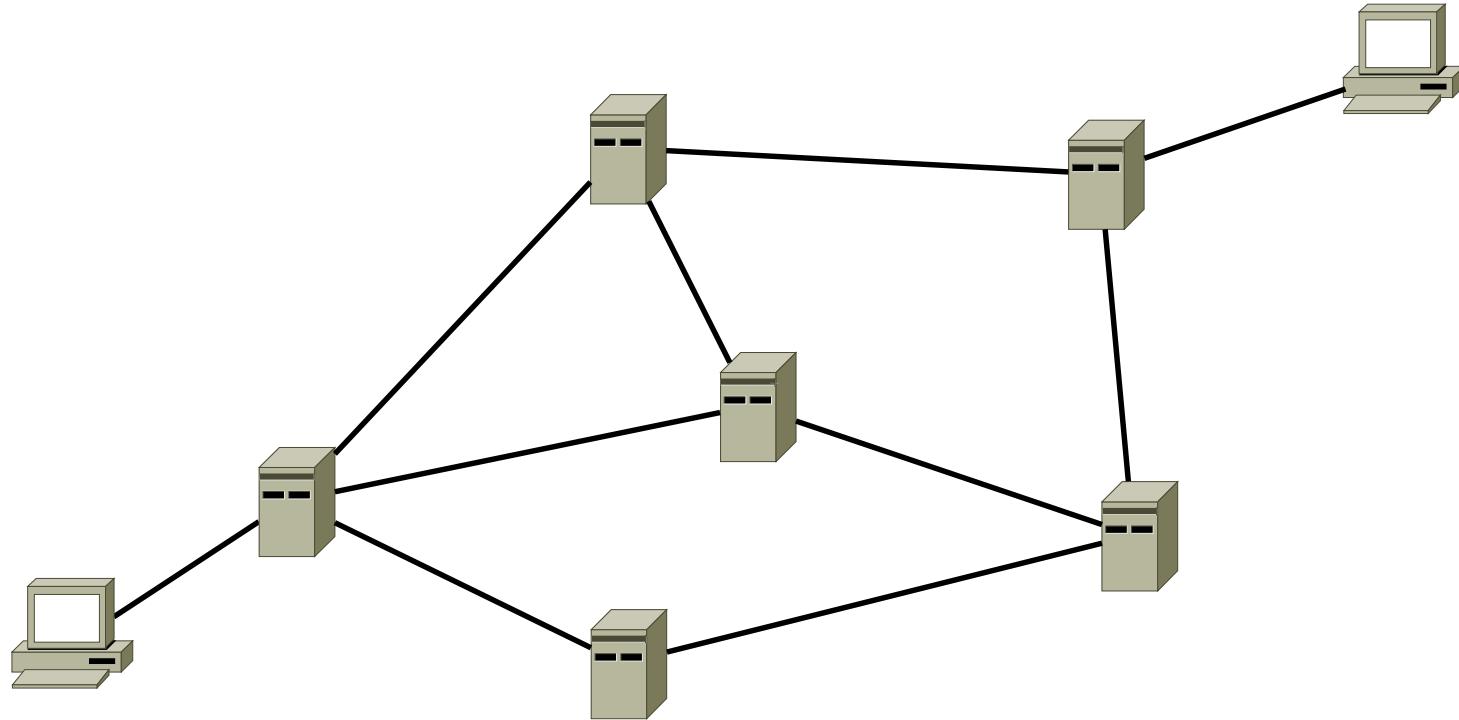
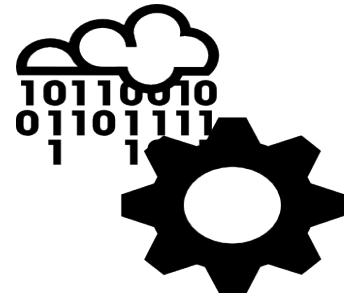


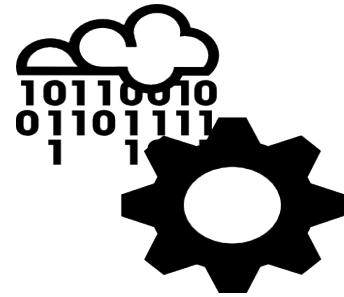
Diagramma realizzato con DIA

internet protocol (IP)



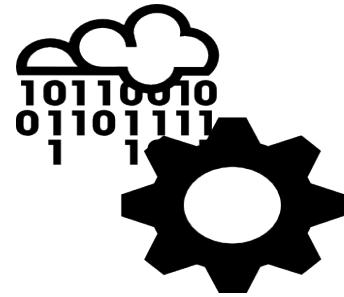
I nodi della rete sono identificati da indirizzi IP univoci. Generalmente, gli indirizzi sono di 4 byte, ossia di 4 elementi separati da . ognuno dei quali può variare tra 0 e 255. Un esempio è ad esempio 151.97.240.4 (UniCT).

internet protocol (IP)



I nodi della rete sono identificati da indirizzi IP univoci. Generalmente, gli indirizzi sono di 4 byte, ossia di 4 elementi separati da . ognuno dei quali può variare tra 0 e 255. Un esempio è ad esempio 151.97.240.4 (UniCT).

Per visualizzare gli IP assegnati alle schede di rete della propria macchina è disponibile (sistemi unix-like) il comando *ifconfig*.



I nodi della rete sono identificati da indirizzi IP univoci. Generalmente, gli indirizzi sono di 4 byte, ossia di 4 elementi separati da . ognuno dei quali può variare tra 0 e 255. Un esempio è ad esempio 151.97.240.4 (UniCT).

Per visualizzare gli IP assegnati alle schede di rete della propria macchina è disponibile (sistemi unix-like) il comando *ifconfig*.

Per visualizzare il proprio IP esposto verso l'esterno sono disponibili dei servizi online, ad esempio

<http://whatismyip.org>.

internet protocol (IP)



I nodi della rete sono identificati da indirizzi IP univoci. Generalmente, gli indirizzi sono di 4 byte, ossia di 4 elementi separati da . ognuno dei quali può variare tra 0 e 255. Un esempio è ad esempio 151.97.240.4 (UniCT).

Per visualizzare gli IP assegnati alle schede di rete della propria macchina è disponibile (sistemi unix-like) il comando *ifconfig*.

Per visualizzare il proprio IP esposto verso l'esterno sono disponibili dei servizi online, ad esempio

<http://whatismyip.org>

Dall'IP è possibile risalire alla posizione geografica del server.

<http://www.iplocationfinder.com>.



Un *pacchetto* inviato da A verso B attraversa un insieme di nodi.

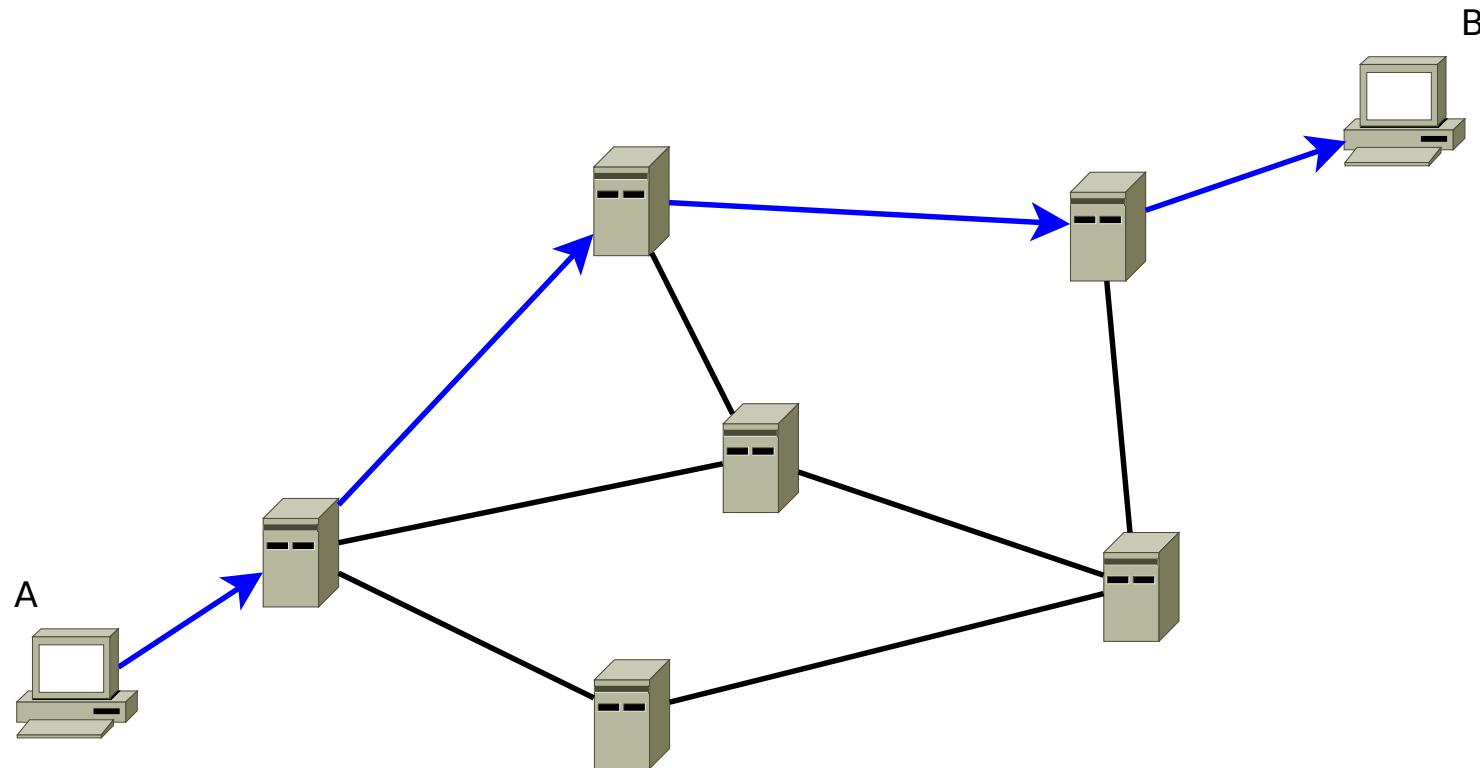
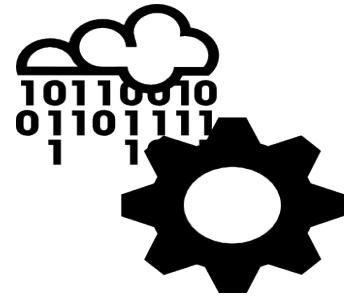


Diagramma realizzato con DIA



Il comando **tracepath** permette di individuare un percorso (in termini di Nodi) dal mio computer ad uno remoto.



Il comando **tracepath** permette di individuare un percorso (in termini di Nodi) dal mio computer ad uno remoto.

Visual Trace Root Tool

(<http://www.yougetsignal.com/tools/visual-tracert/>) permette di visualizzare graficamente il percorso da un host predefinito ad un qualsiasi nodo della rete.



Un *pacchetto* inviato da A verso B attraversa un insieme di nodi.
Il percorso non è fissato.

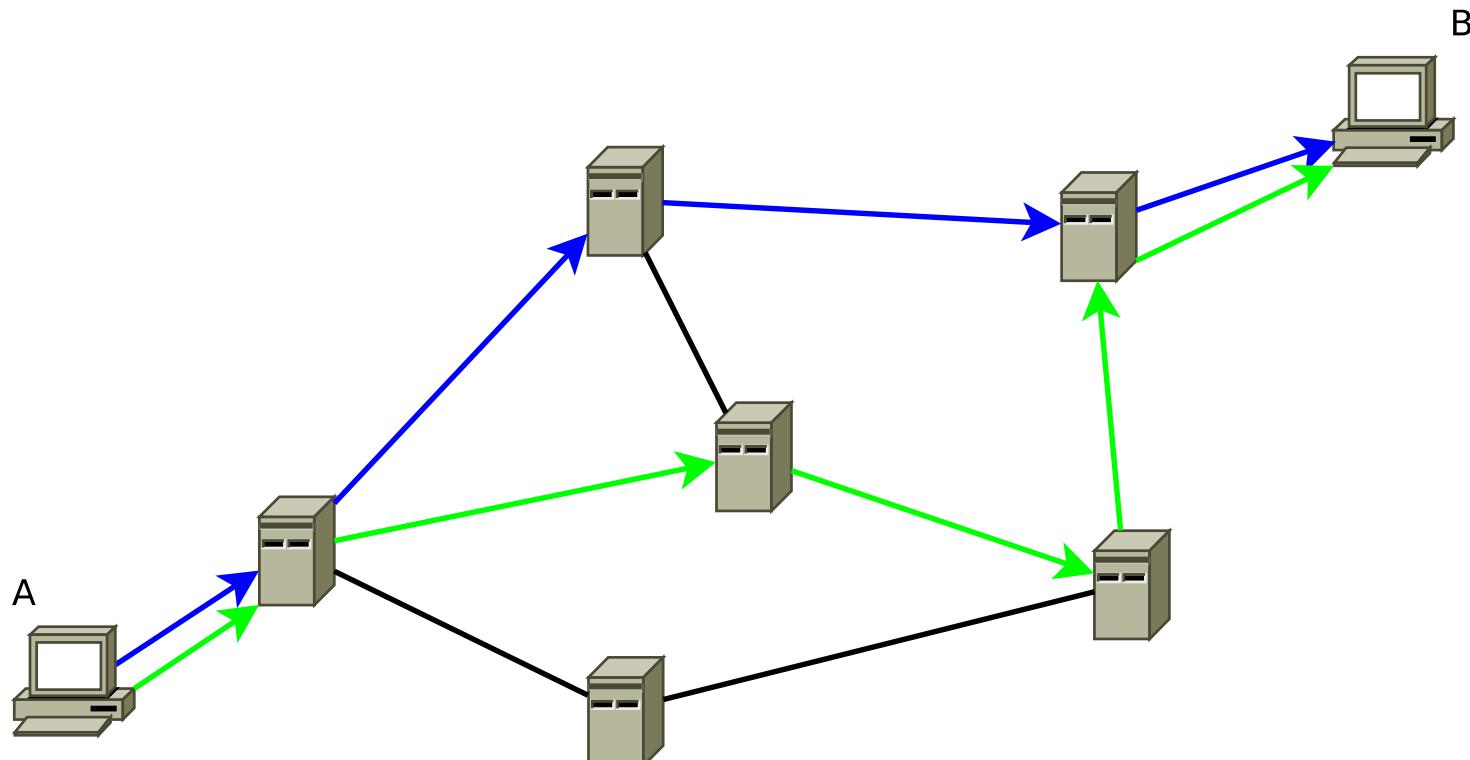


Diagramma realizzato con DIA



Se uno dei nodi sul percorso è sovraccarico o non funziona bene, viene utilizzato un percorso alternativo.

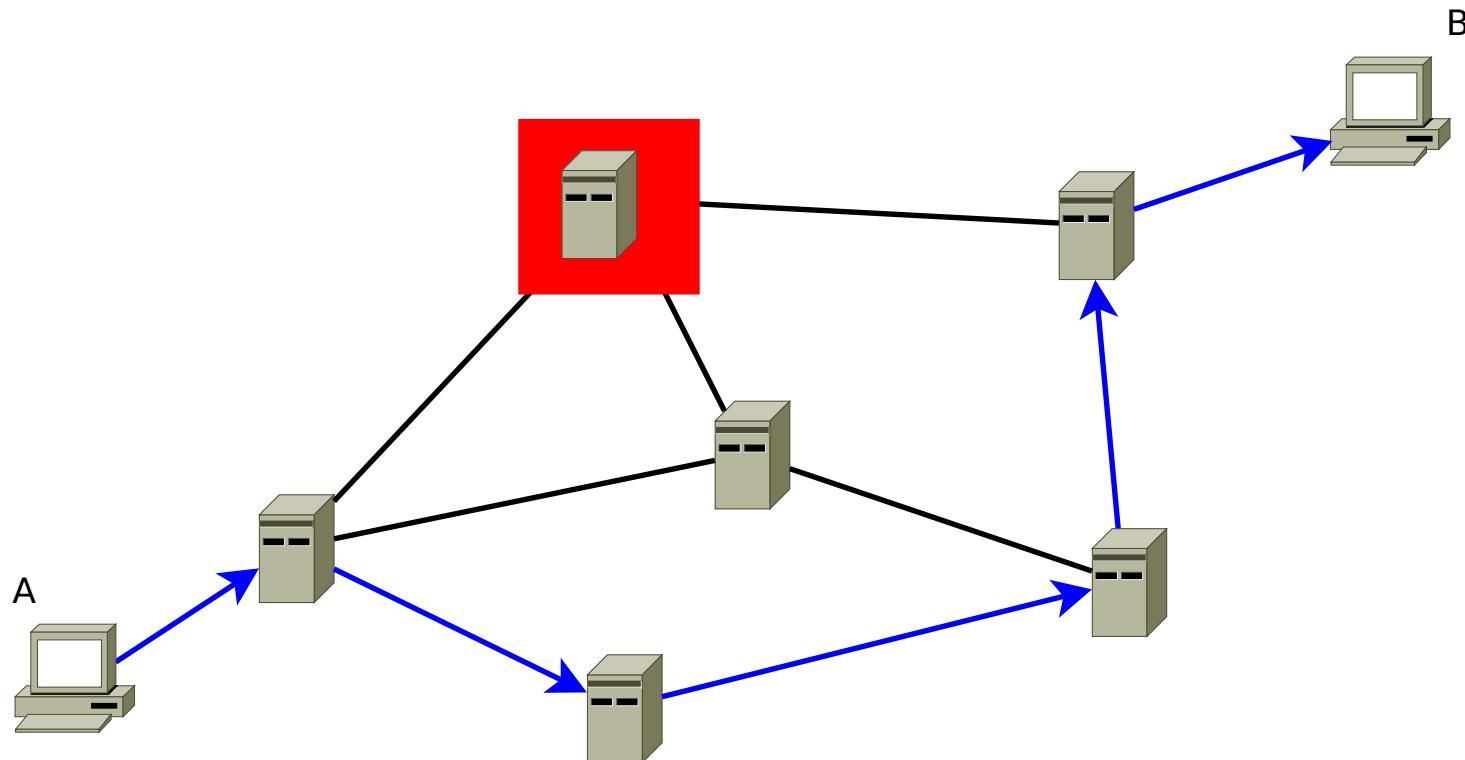
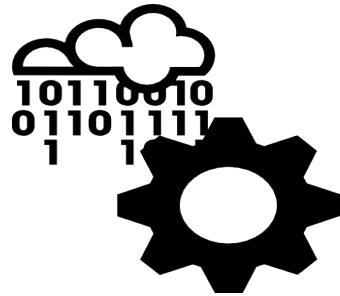
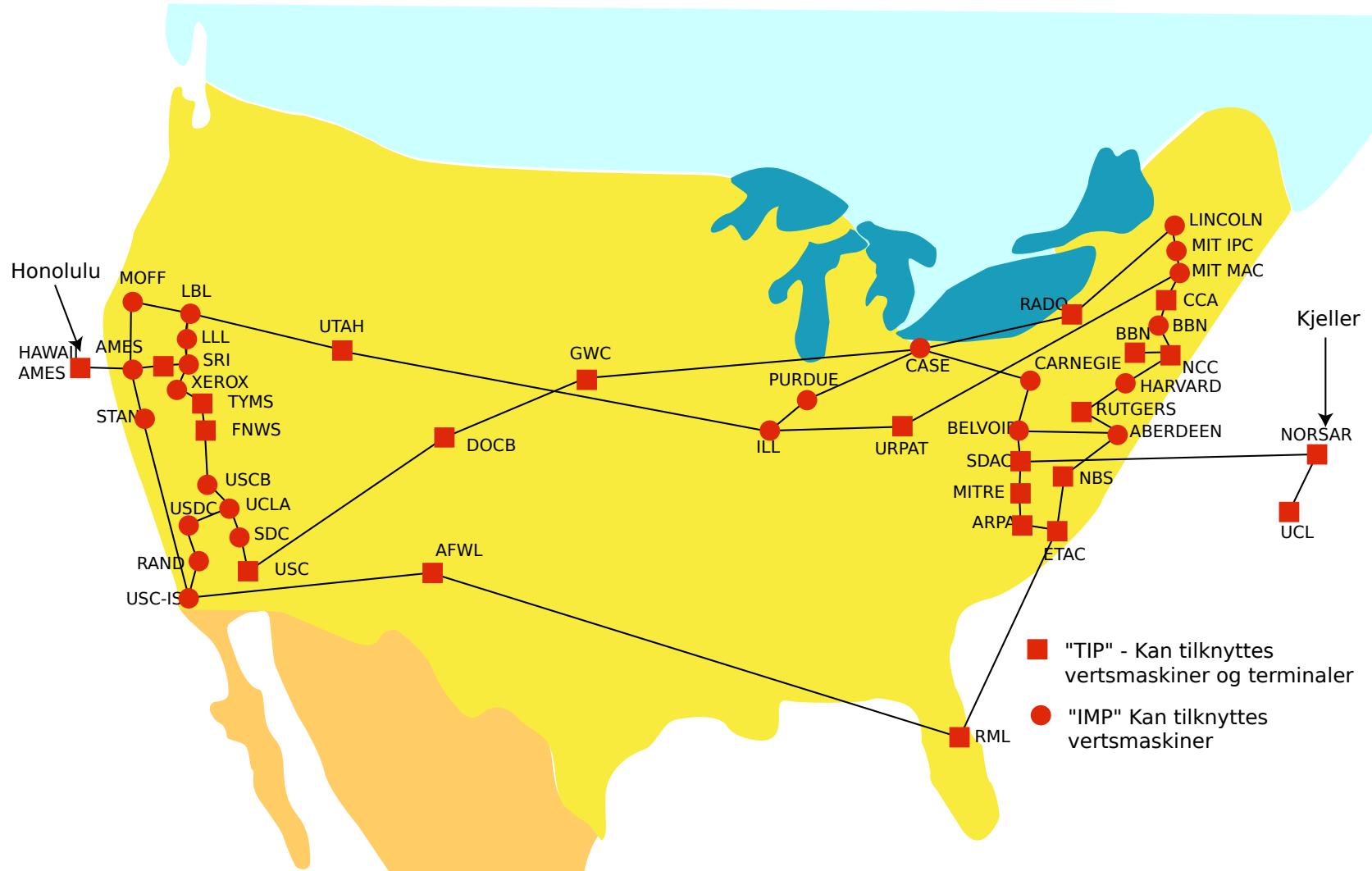


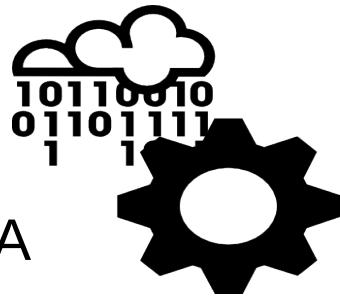
Diagramma realizzato con DIA



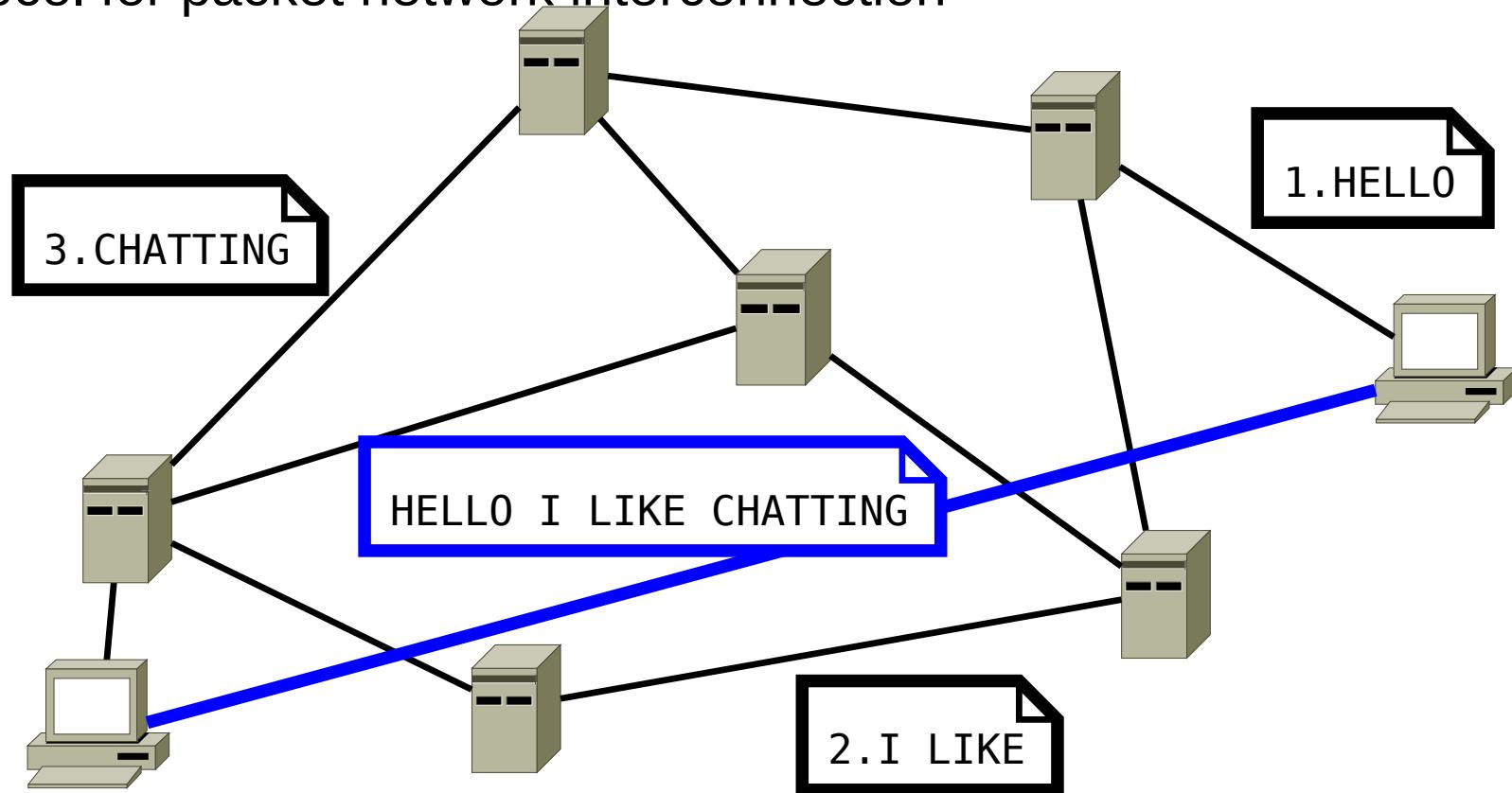
1972 - Viene ufficialmente presentata ARPANET, rete di computer di area-vasta.



Arpanet 1974 - By Yngvar [Public domain], via Wikimedia Commons

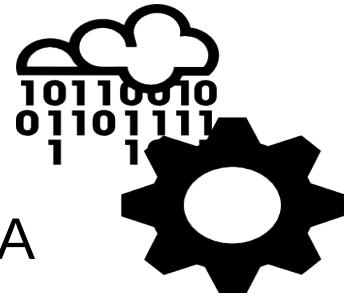


1974 - nasce il protocollo TCP/IP : V. G. Cerf and R. E. Kahn, "A protocol for packet network interconnection"



Permette di astrarsi dal concetto di pacchetto per usare quello di connessione vista come uno *stream* bidirezionale di dati.

Diagramma realizzato con DIA



1974 - nasce il protocollo TCP/IP : V. G. Cerf and R. E. Kahn, "A protocol for packet network interconnection"

Aggiunge il concetto di *porta*: ogni servizio su un server è in ascolto su una porta attraverso la quale contattarlo.

Alcune porte standard sono

20 - FTP

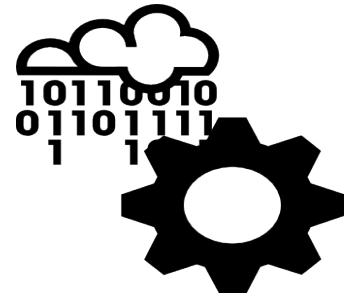
22 - SSH

25 - SMTP

80 - HTTP

190 - POP

443 - HTTPS



E' possibile stabilire una connessione TCP (justo per prova) con un servizio remoto in ascolto attraverso il comando

```
telnet <ip> <port>
```



I risultati tecnico-scientifici, le scoperte e le *invenzioni* che hanno permesso la realizzazione e lo sviluppo di internet sono nati dal contributo e la collaborazione di moltissimi studiosi, gruppi di ricerca e progetti industriali ...



National Physical Laboratory (logo non disponibile)

All website design, text, graphics, the selection and arrangement thereof and all other material on this site are ©NPL Management Limited.

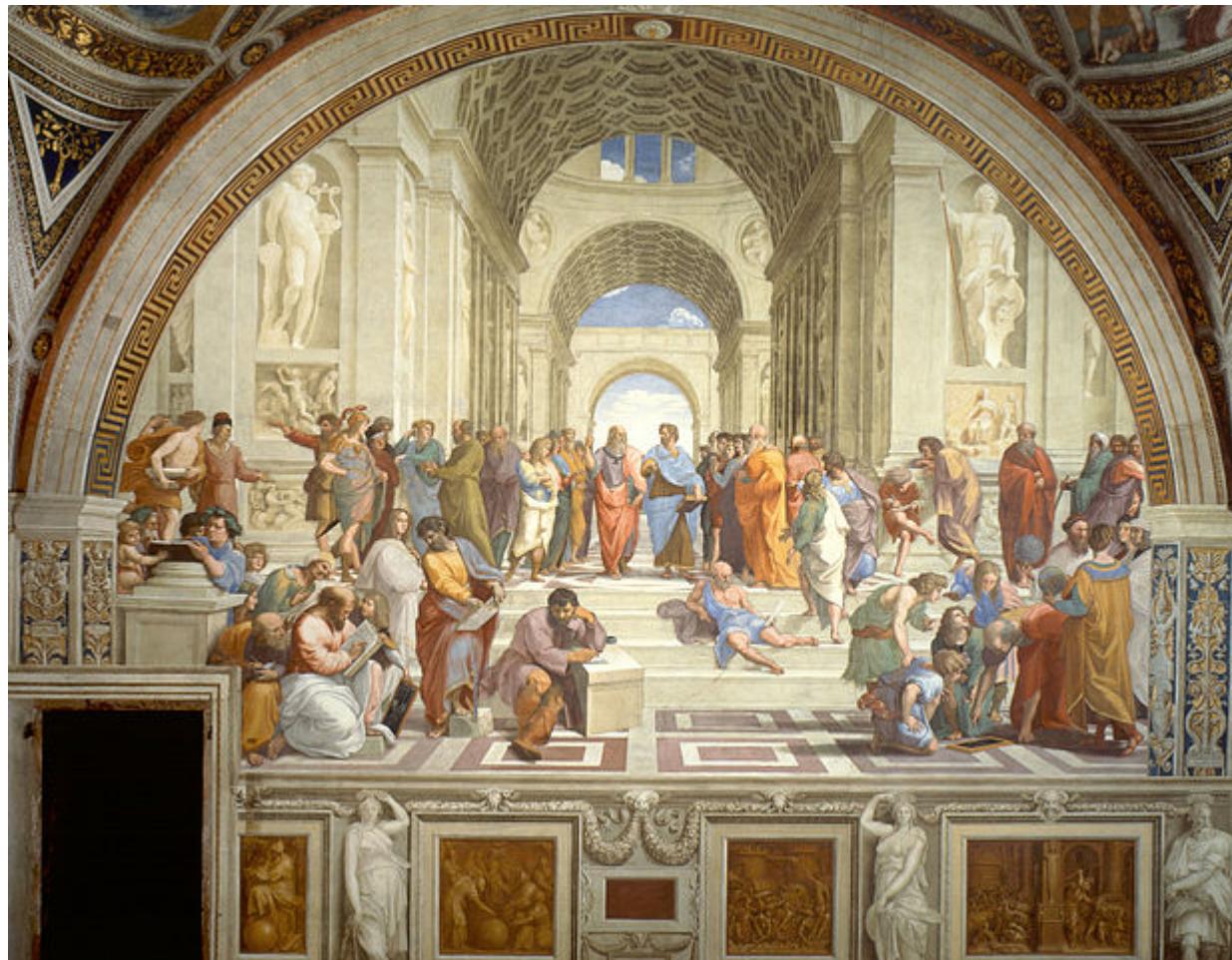
Non è chiaro se si può usare questo



By DARPA (DARPA.mil) [Public domain]; By Massachusetts Institute of Technology (<http://www.mit.edu/>, vectorized by User:Mysid) [Public domain], By Original University of California seal: probably Tiffany & Co.,; This SVG file: User:Casecrer (File:The_University_of_California_1868.svg) [Public domain], via Wikimedia Commons,



... che a loro volta hanno basato il loro lavoro su conoscenza sviluppata e condivisa nei secoli.



La scuola di Atene - Raffaello Sanzio, 1509-1511 circa, via Wikimedia Commons



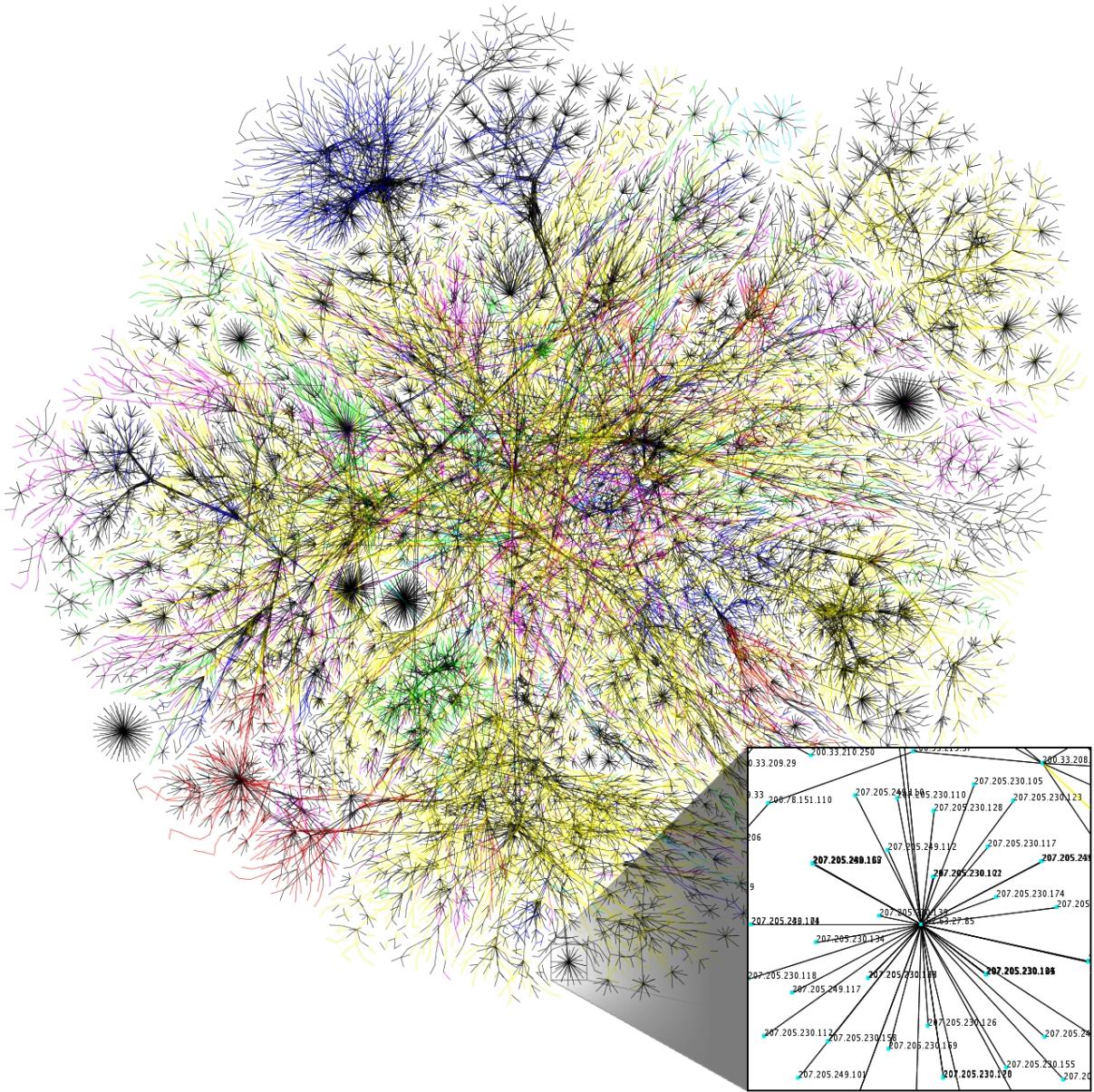
"Information is power. But like all power, there are those who want to keep it for themselves. The world's entire scientific and cultural heritage, published over centuries in books and journals, is increasingly being digitized and locked up by a handful of private corporations. Want to read the papers featuring the most famous results of the sciences? You'll need to send enormous amounts to publishers like Reed Elsevier."

Guerilla Open Access Manifesto - 2008

Aaron Swartz (1986 - 2013)



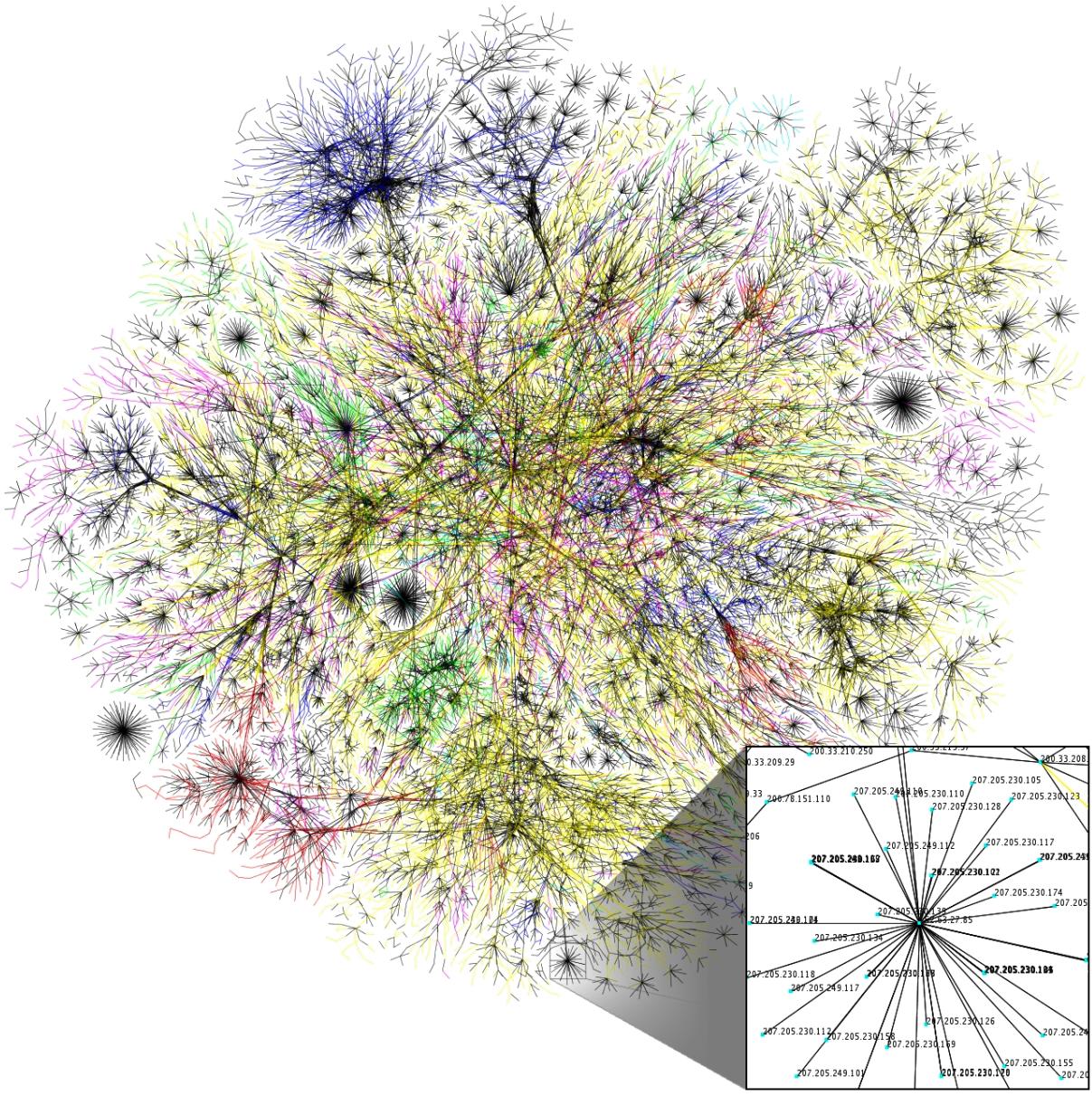
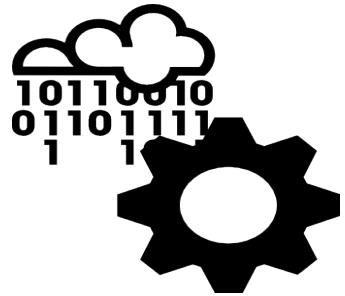
La scuola di Atene - Raffaello Sanzio, 1509-1511 circa, via Wikimedia Commons



“The Internet is the global system of interconnected computer networks that use the Internet protocol suite (TCP/IP) to link billions of devices worldwide. **It is a network of networks** that consists of millions of private, public, academic, business, and government networks of local to global scope, linked by a broad array of electronic, wireless, and optical networking technologies.”

Da Wikipedia

By The Opte Project [CC BY 2.5 (<http://creativecommons.org/licenses/by/2.5>) or CC BY 2.5 (<http://creativecommons.org/licenses/by/2.5>)], via Wikimedia Commons



“The Internet is the global system of interconnected computer networks that use the Internet protocol suite (TCP/IP) to link billions of devices worldwide. **It is a network of networks** that consists of millions of private, public, academic, business, and government networks of local to global scope, linked by a broad array of electronic, wireless, and optical networking technologies.”

Da Wikipedia

E’ (ancora) basata sullo stack di protocolli TCP/IP

By The Opte Project [CC BY 2.5 (<http://creativecommons.org/licenses/by/2.5>) or CC BY 2.5 (<http://creativecommons.org/licenses/by/2.5>)], via Wikimedia Commons



Internet si basa su tecnologie, protocolli e formati *aperti*.

Formato dei dati di tipo aperto: *un formato di dati reso pubblico, documentato esaustivamente e neutro rispetto agli strumenti tecnologici necessari per la fruizione dei dati stessi.* (AgID)

Open Format: *The work must be provided in an open format. An open format is one which places no restrictions, monetary or otherwise, upon its use and can be fully processed with at least one free/libre/open-source software tool.*

(The Open Definition, Open Knowledge Foundation)



“Free software means that the software's users have freedom. (The issue is not about price.) [...] Specifically, free software means users have the four essential freedoms: (0) to run the program, (1) to study and change the program in source code form, (2) to redistribute exact copies, and (3) to distribute modified versions.”

da Philosophy of the GNU Project

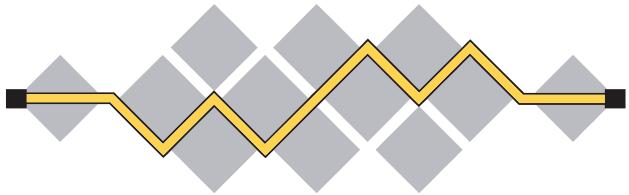


<https://www.gnu.org/philosophy/philosophy.html>
By Bill Ebbesen (Own work) [Public domain], via Wikimedia Commons



Internet si basa su tecnologie, protocolli e formati *aperti*. Vedi ad esempio <https://www.rfc-editor.org/rfc-index.html>

Gli standard vengono approvati e pubblicati da consorzi



I E T F®

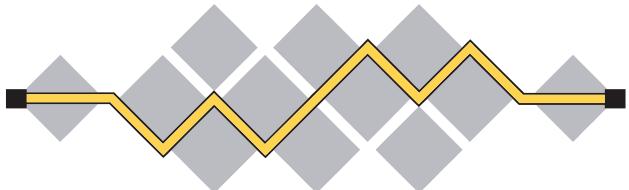
W3C®

Di the IETF Trust (<http://www.ietf.org/images/logos/>) [Public domain], attraverso Wikimedia Commons, via Wikimedia Commons



Internet si basa su tecnologie, protocolli e formati *aperti*. Vedi ad esempio <https://www.rfc-editor.org/rfc-index.html>

Gli standard vengono approvati e pubblicati da *consorzi* che coinvolgono tutti gli interessati.

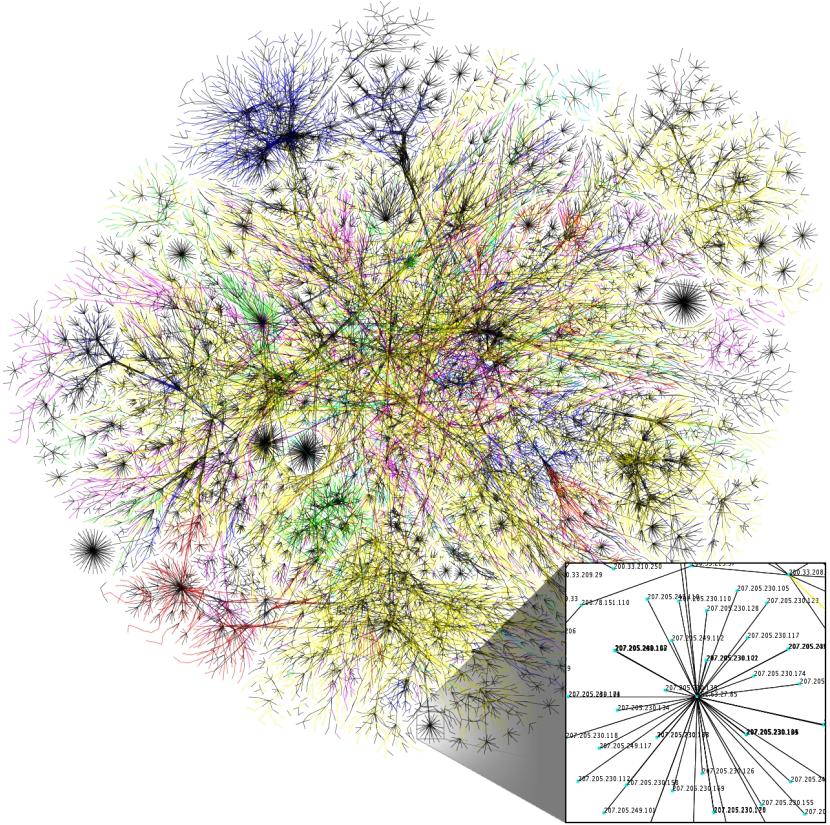


I E T F®

Ericsson, Oracle, Cisco, Isode,
ICANN, Fastly, Juniper Networks,
AT&T, ...

W3C®

Di the IETF Trust (<http://www.ietf.org/images/logos/>) [Public domain], attraverso Wikimedia Commons, via Wikimedia Commons



Internet è detta la *rete delle reti*.

“The Internet consists of lots of different bits and pieces, each of which has an owner.”

[http://computer.howstuffworks.com/internet/basics/wh-o-owns-internet1.htm](http://computer.howstuffworks.com/internet/basics/who-owns-internet1.htm)

“So who actually owns the Internet? There are two answers to this question:

- 1. Nobody*
- 2. Lots of people”*

internet è difficilmente controllabile



La struttura di internet rende intrinsecamente difficile impedire, controllare e alterare le comunicazioni.

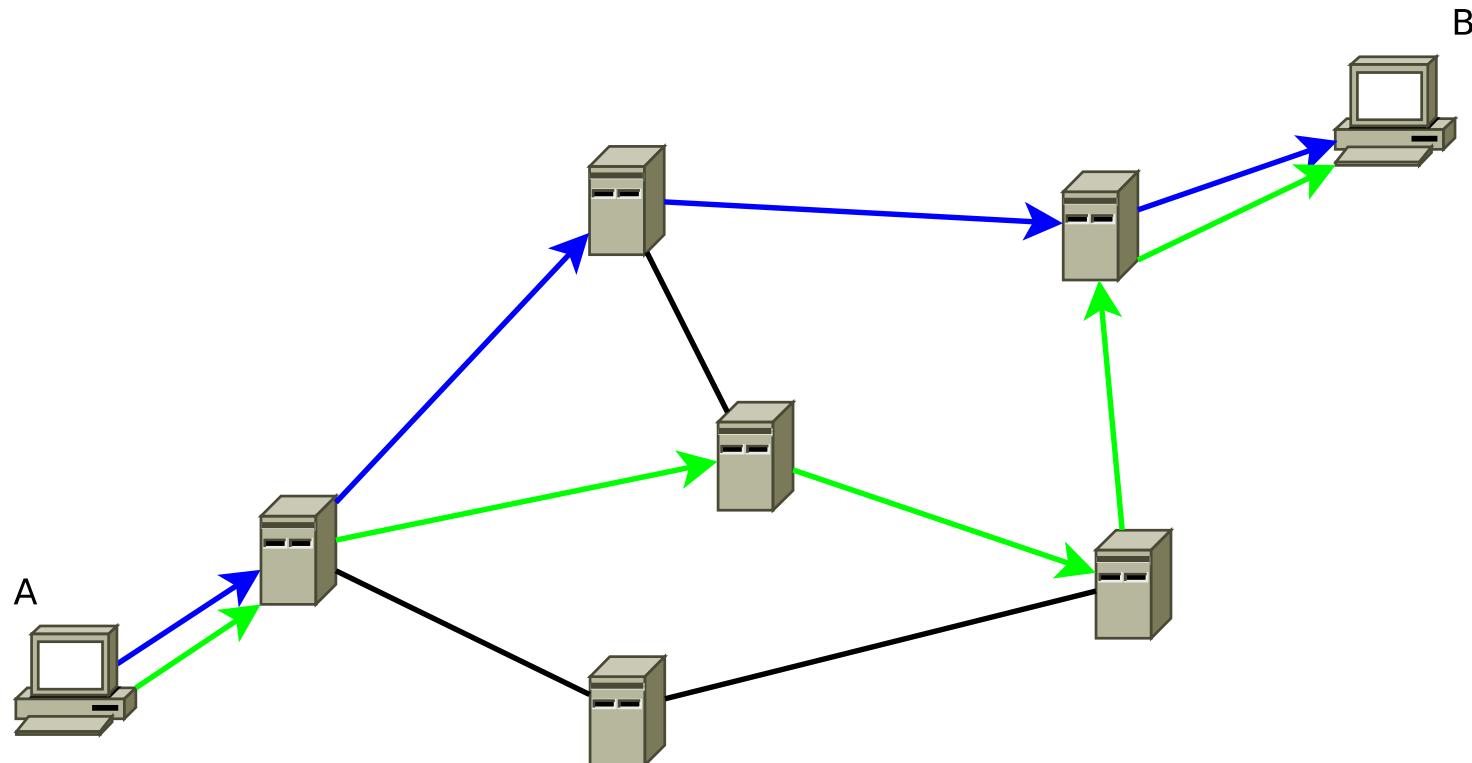


Diagramma realizzato con DIA

internet è difficilmente controllabile



La struttura di internet rende intrinsecamente difficile impedire, controllare e alterare le comunicazioni.

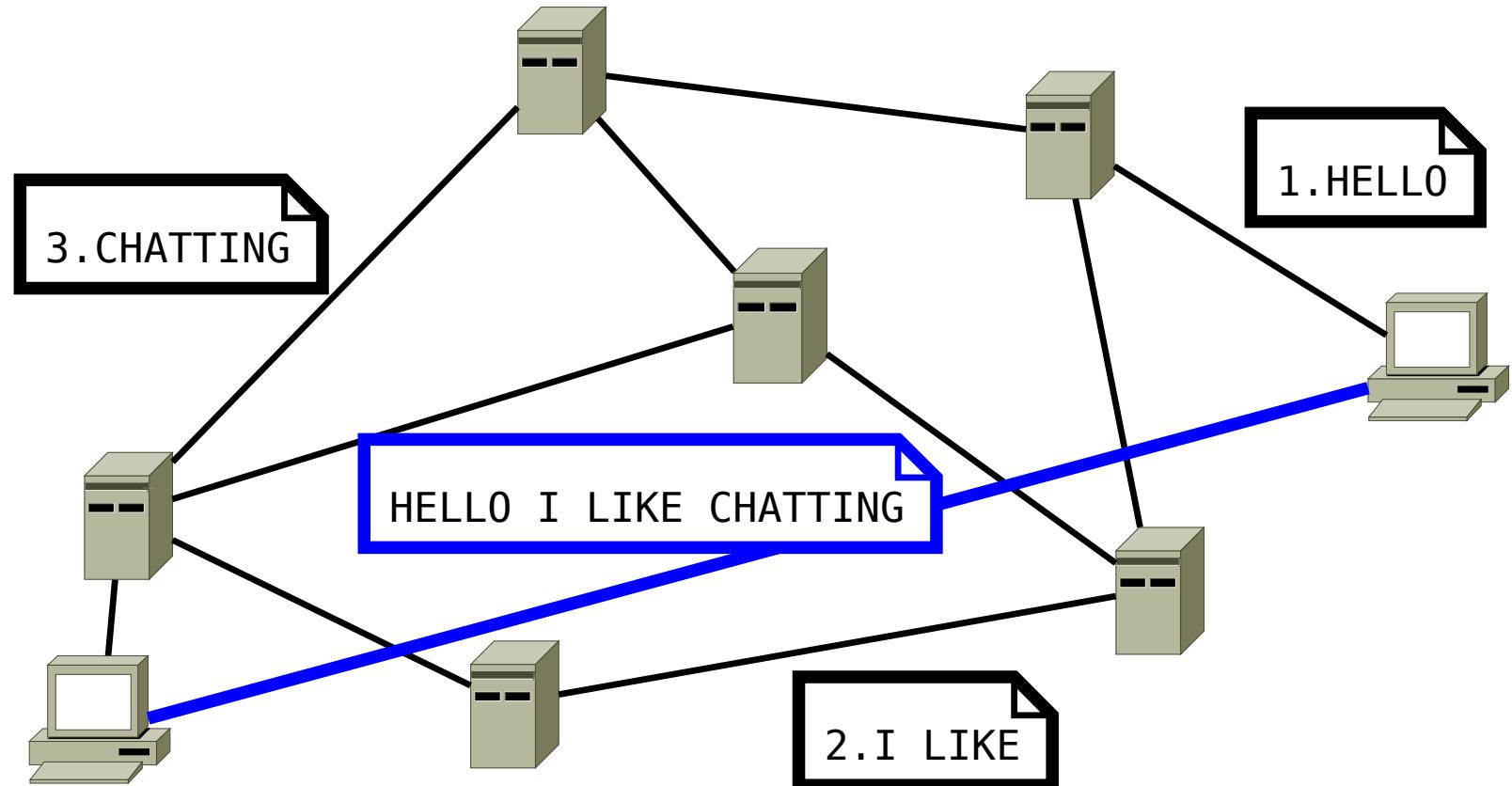
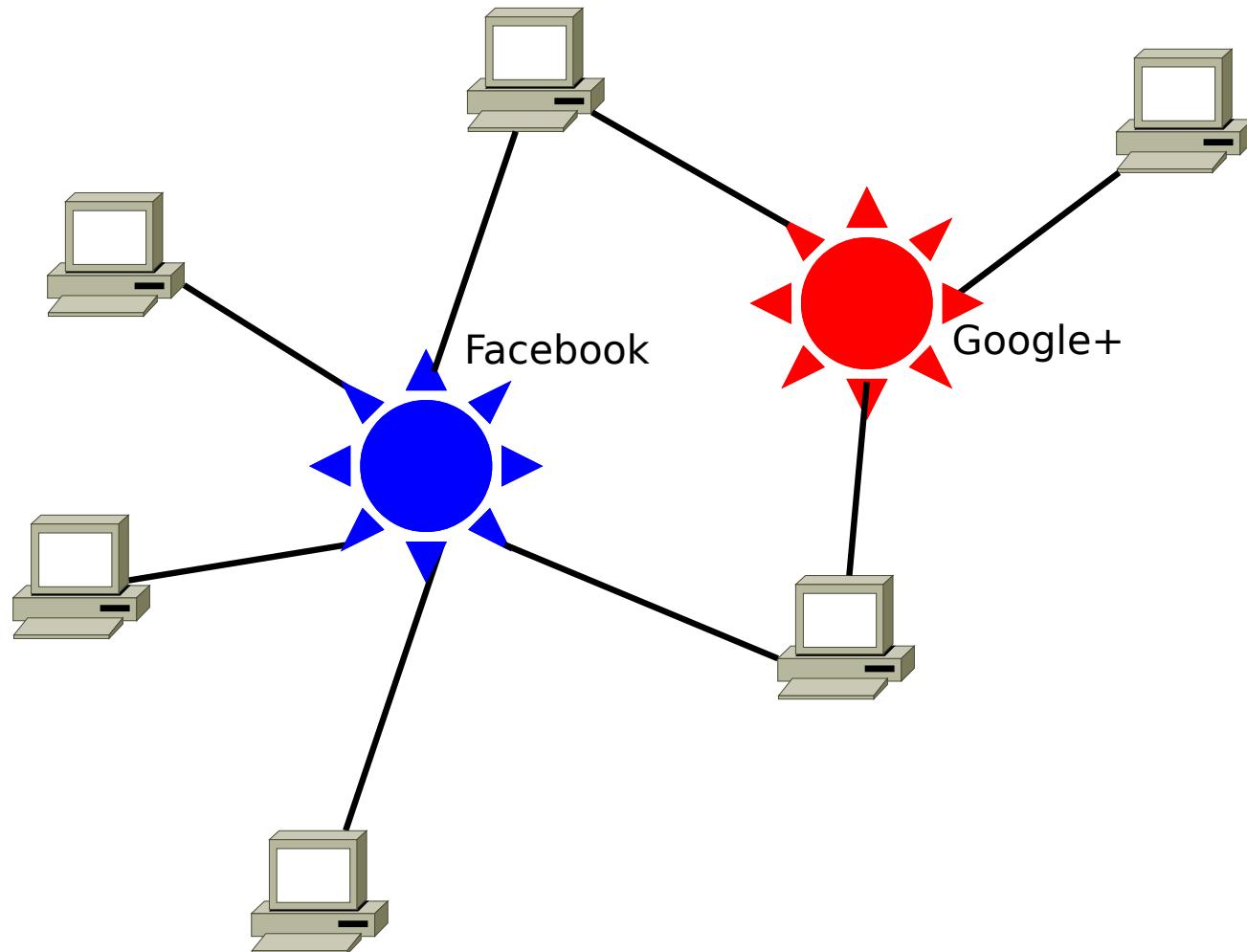


Diagramma realizzato con DIA

Social networking services



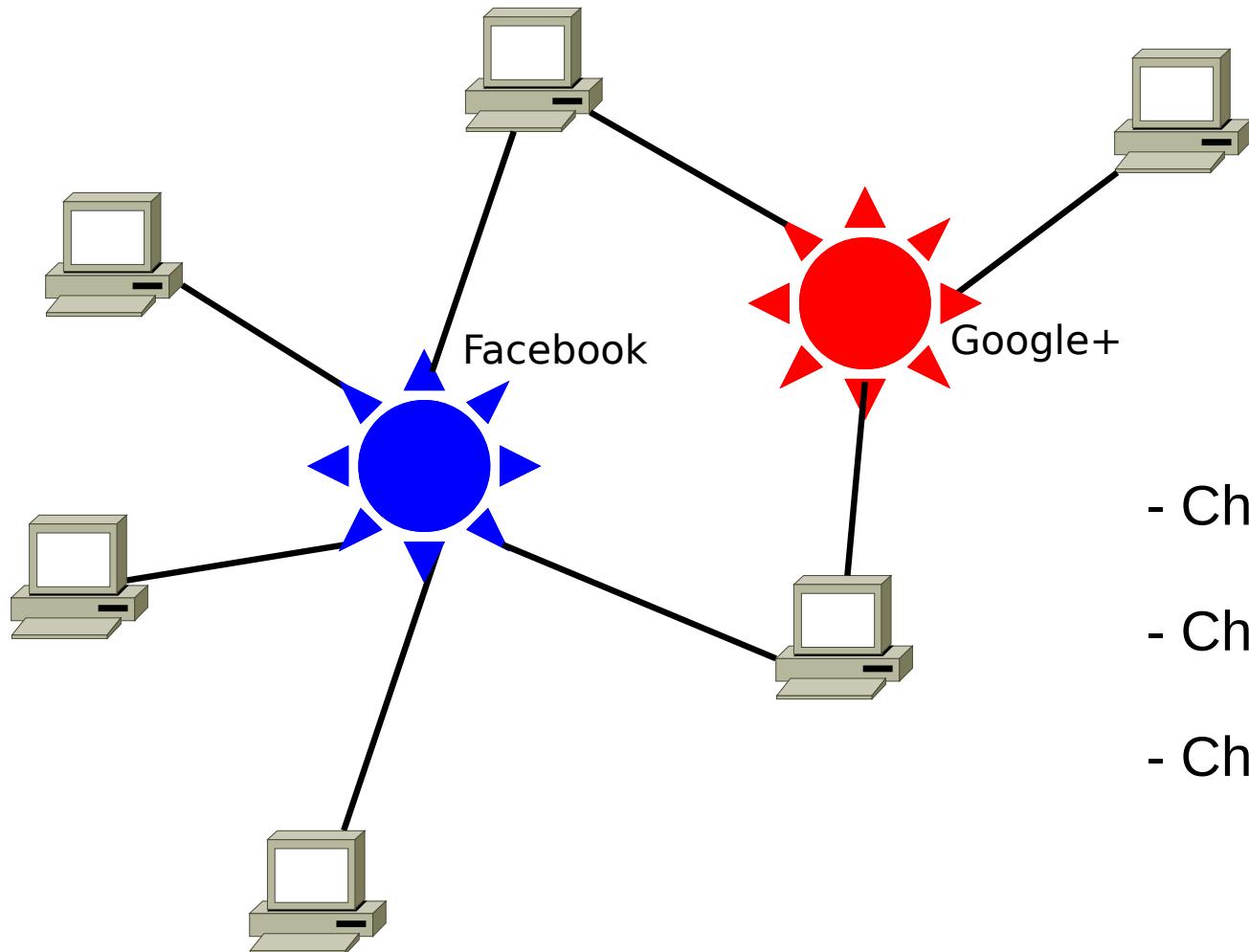
Nell'ultimo decennio la diffusione dei cosiddetti *social network* ha fatto sì che le comunicazioni su internet avvengano prevalentemente in maniera *intermediata*.



Social networking services

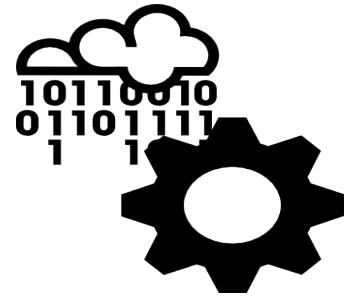


Nell'ultimo decennio la diffusione dei cosiddetti *social network* ha fatto sì che le comunicazioni su internet avvengano prevalentemente in maniera *intermediata*.

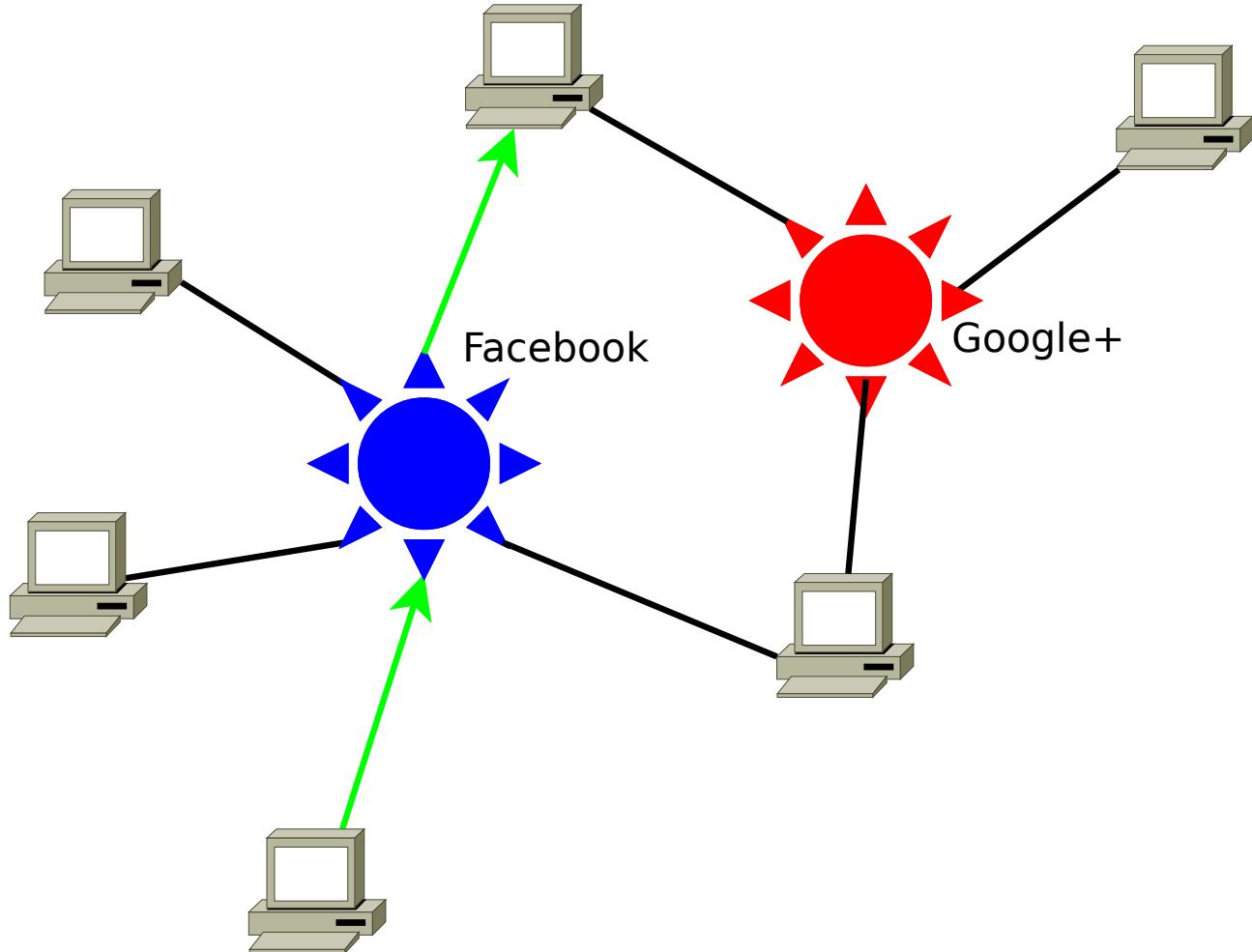


- Chi possiede Facebook?
- Chi possiede Google+?
- Chi possiede Telegram?

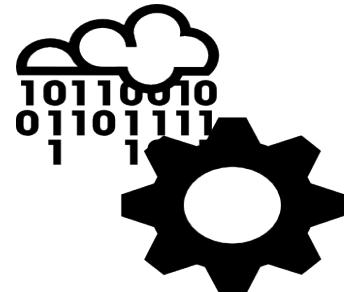
Social networking services



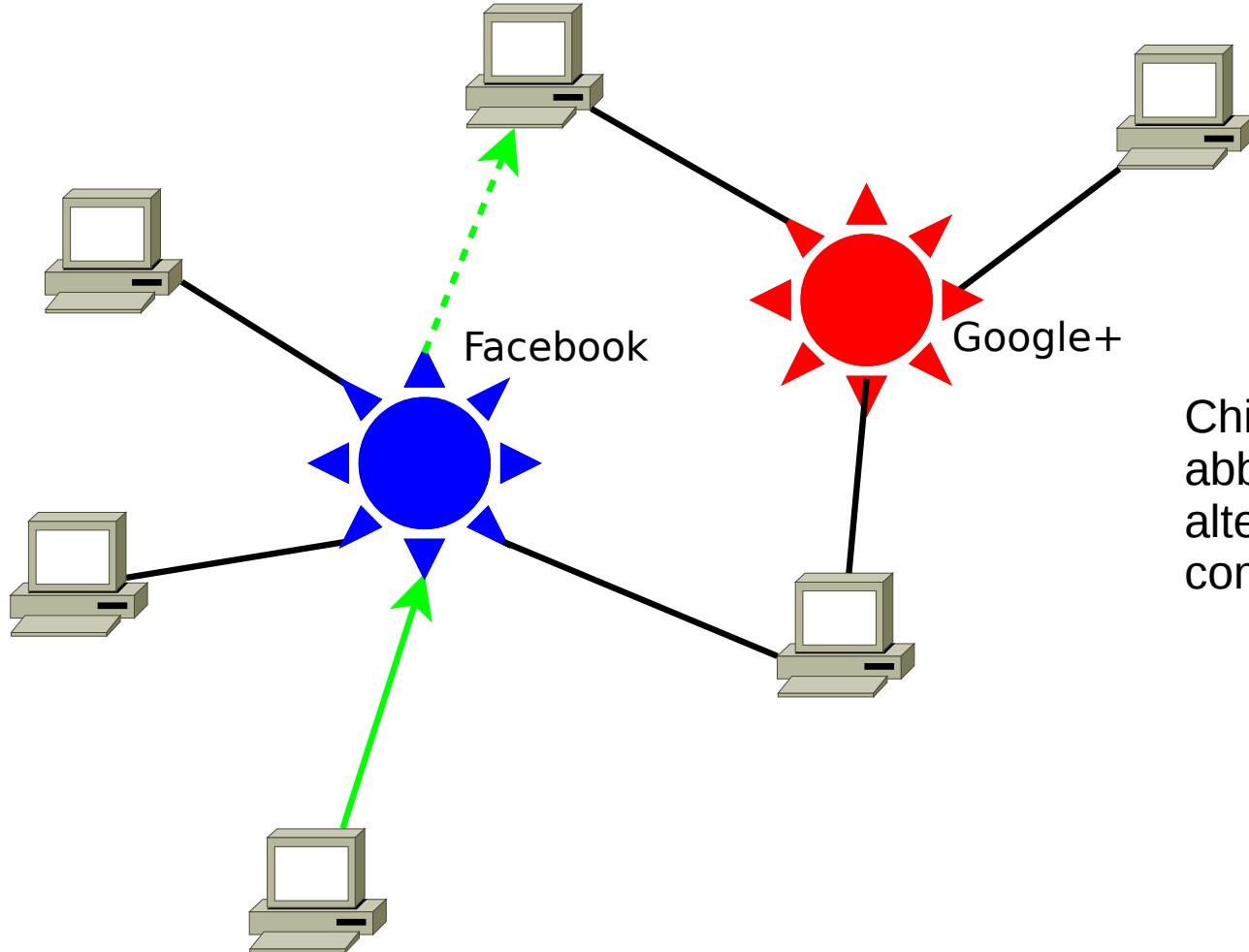
Nell'ultimo decennio la diffusione dei cosiddetti *social network* ha fatto sì che **le comunicazioni su internet avvengano prevalentemente in maniera *intermediata*.**



Social networking services



Nell'ultimo decennio la diffusione dei cosiddetti *social network* ha fatto sì che **le comunicazioni su internet avvengano prevalentemente in maniera *intermediata*.**



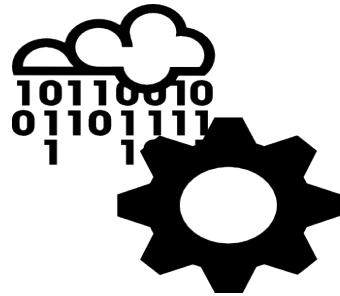
Chi gestisce i social network può abbastanza facilmente impedire, alterare e controllare le comunicazioni.



Molta gente si informa solo su internet.

L'accesso ad internet avviene principalmente attraverso due canali: Google e Facebook.

Distopia – Google e Facebook possono veicolare una visione alterata della realtà (vedi Orson Wells).



Molta gente si informa solo su internet.

L'accesso ad internet avviene principalmente attraverso due canali: Google e Facebook.

Distopia – Google e Facebook possono veicolare una visione alterata della realtà (vedi Orson Wells).

Realtà - l'ordine con il quale vengono presentate le risposte da google non è noto. Si pensa che Google sia in grado di influenzare il 20% dei votanti alle elezioni presidenziali degli US.



WORLD WIDE WEB



“Imagine, then, the references in this document all being associated with the network address of the thing to which they referred, so that while reading this document you could skip to them with a click of the mouse.”

1989 - Tim Berners Lee, Information Management: A Proposal

<https://www.w3.org/History/1989/proposal.html>

<http://home.cern/topics/birth-web>



“Imagine, then, the references in this document all being associated with the network address of the thing to which they referred, so that while reading this document you could skip to them with a click of the mouse.”

1989 - Tim Berners Lee, Information Management: A Proposal

1993 - il CERN rilascerà di pubblico dominio i primi software per il world wide web, tra cui il primo browser.

<https://www.w3.org/History/1989/proposal.html>
<http://home.cern/topics/birth-web>
<http://line-mode.cern.ch/>



“The World Wide Web (WWW, or simply Web) is an information space in which the items of interest, referred to as resources, are identified by global identifiers called Uniform Resource Identifiers (URI).”

da Architecture of the World Wide Web, Volume I

Le prime specifiche rilasciate furono:

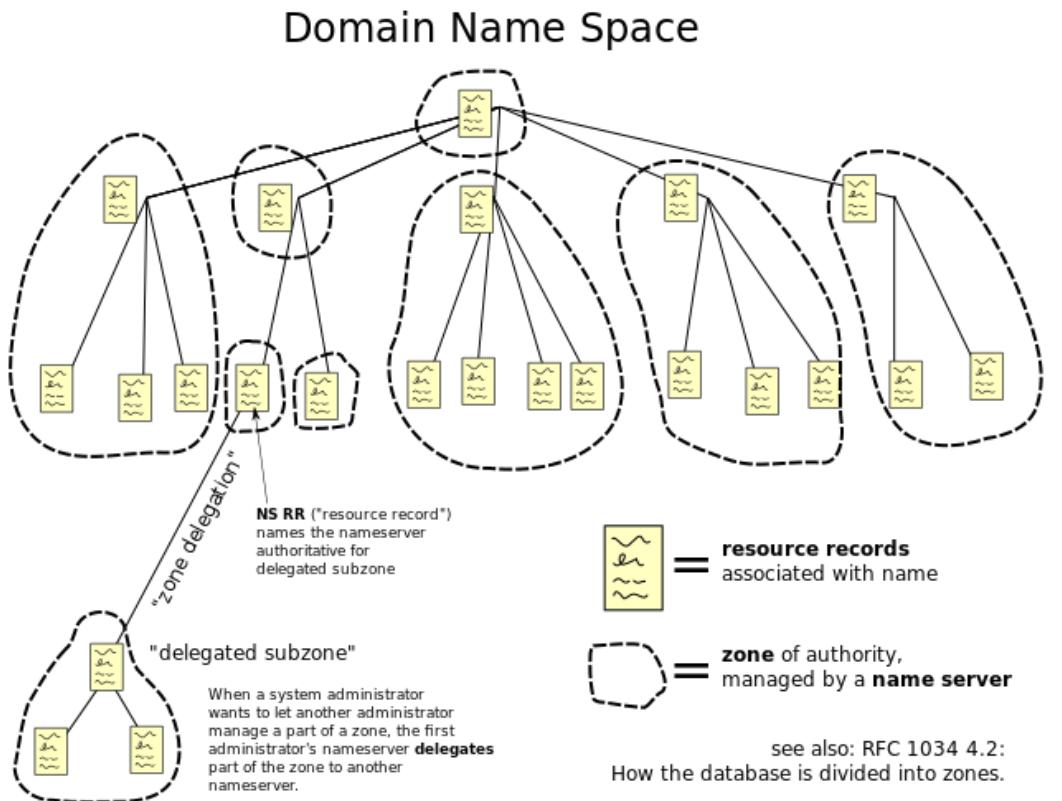
- Uniform Resource Locators (URLs),
- Hypertext Transfer Protocol (HTTP),
- Hypertext Markup Language (HTML).

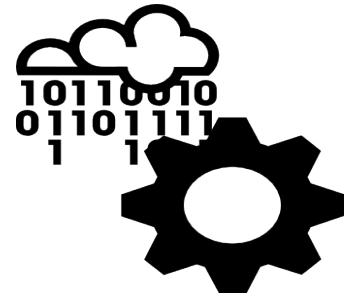
<https://www.w3.org/TR/webarch/>



Domain Name System (DNS, vedi RFC1034) è un sistema gerarchico decentralizzato che permette di associare a *nomi di dominio* delle risorse disponibili su Internet.

Nella sua versione più basilare associa ad un nome di dominio un indirizzo IP. Queste associazioni risiedono, in una organizzazione gerarchica, all'interno dei cosiddetti sever DNS che comunicano tra loro usando il protocollo DNS.





I nomi di dominio riflettono la struttura gerarchica del domain name system. I vari segmenti sono separati da “.” e vanno, da destra verso sinistra, dal livello più grande al più piccolo.

www.dmi.unict.it

Il comando nslookup <domainname> permette di ricavare l'ip associato al nome di dominio specificato.

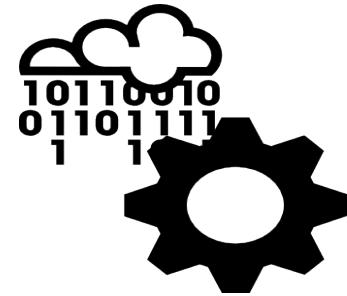
<https://www.w3.org/TR/webarch/>

Uniform Resource Locator



Uniform Resource Locator (URL, vedi RFC1738) è un formato di *nomi* per indicare risorse sul web.

Esempio: <https://www.ietf.org/rfc/rfc1738.txt>



<https://www.ietf.org/rfc/rfc1738.txt>

Common Internet Scheme Syntax:

<scheme>:// [<user> [:<password>]@] <host> [:<port>] [/<url-path>]



<https://www.ietf.org/rfc/rfc1738.txt>

Common Internet Scheme Syntax:

<scheme>:// [<user> [:<password>]@] <host> [:<port>] [/<url-path>]

schemi: http, https, mailto, ftp, ... (gli schemi solitamente indicano il tipo di risorsa ed il protocollo)



<https://www.ietf.org/rfc/rfc1738.txt>

Common Internet Scheme Syntax:

<scheme> : // [<user> [:<password>] @] <host> [:<port>] [/<url-path>]

schemi: http, https, mailto, ftp, ... (gli schemi solitamente indicano il tipo di risorsa ed il protocollo)

host: nome di dominio o indirizzo IP del server.



<https://www.ietf.org/rfc/rfc1738.txt>

Common Internet Scheme Syntax:

<scheme> : // [<user> [:<password>]@] <host> [:<port>] [/<url-path>]

schemi: http, https, mailto, ftp, ... (gli schemi solitamente indicano il tipo di risorsa ed il protocollo)

host: nome di dominio o indirizzo IP del server.

url-path: se presente, è una parte gerarchica nella quale tutti i livelli sono separati da /. Indica la risorsa all'interno del server.



`https://www.ietf.org/rfc/rfc1738.txt?id=1024&var=v`

Common Internet Scheme Syntax:

`<scheme> : // [<user> [:<password>]@] <host> [:<port>] [/<url-path>]`

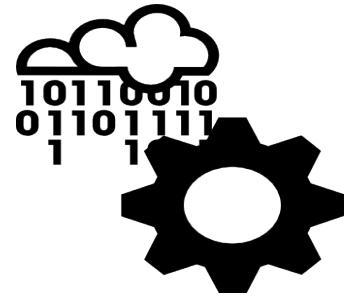
schemi: http, https, mailto, ftp, ... (gli schemi solitamente indicano il tipo di risorsa ed il protocollo)

host: nome di dominio o indirizzo IP del server.

url-path: se presente, è una parte gerarchica nella quale tutti i livelli sono separati da / . Indica la risorsa all'interno del server.

Dopo al termine del path alcuni URL Scheme prevedono la presenza di una *query-part*

HyperText Transfer Protocol



Hyper Text Transfer Protocol (HTTP, RFC2616), è un protocollo *richiesta-risposta* per comunicazioni client-server:

<https://www.w3.org/TR/webarch/>

HyperText Transfer Protocol



Hyper Text Transfer Protocol (HTTP, RFC2616), è un protocollo *richiesta-risposta* per comunicazioni client-server:

- il client apre una connessione TCP col server

<http://info.cern.ch>

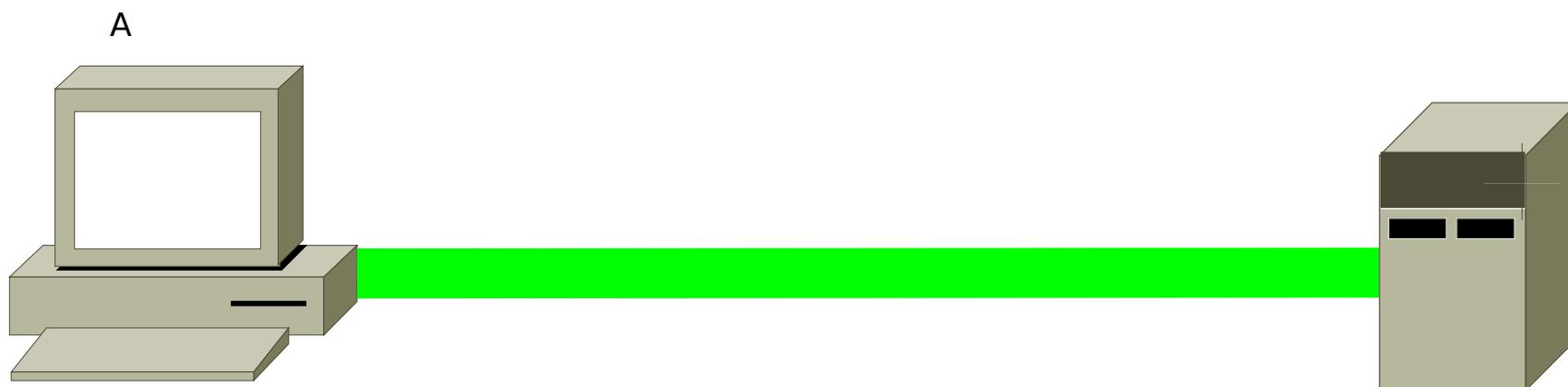
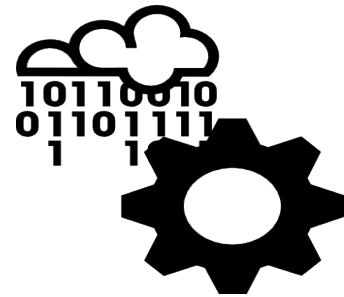


Diagramma realizzato con DIA

HyperText Transfer Protocol



Hyper Text Transfer Protocol (HTTP, RFC2616), è un protocollo *richiesta-risposta* per comunicazioni client-server:

- il client apre una connessione TCP col server
- il client invia una richiesta per una risorsa al server

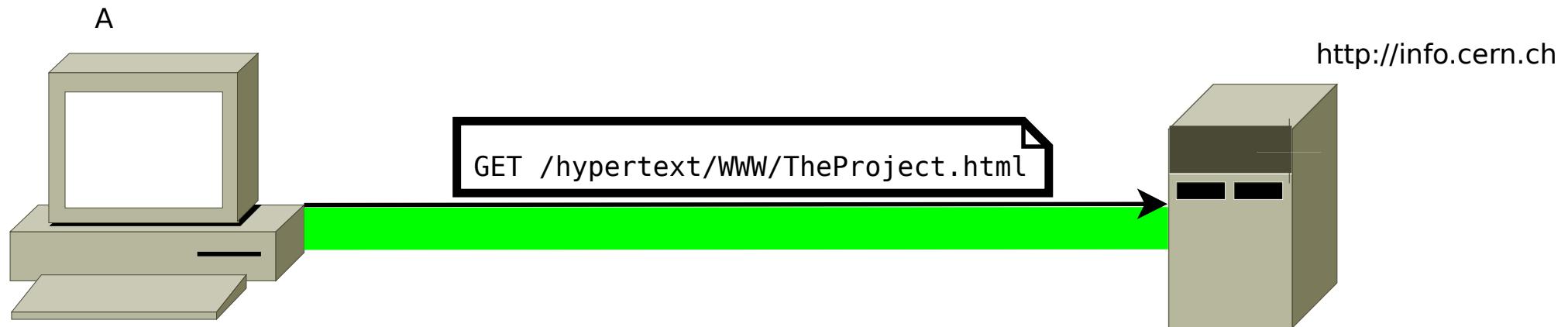


Diagramma realizzato con DIA

HyperText Transfer Protocol

Hyper Text Transfer Protocol (HTTP, RFC2616), è un protocollo *richiesta-risposta* per comunicazioni client-server:

- il client apre una connessione TCP col server
- il client invia una richiesta per una risorsa al server
- il server invia la risorsa al client

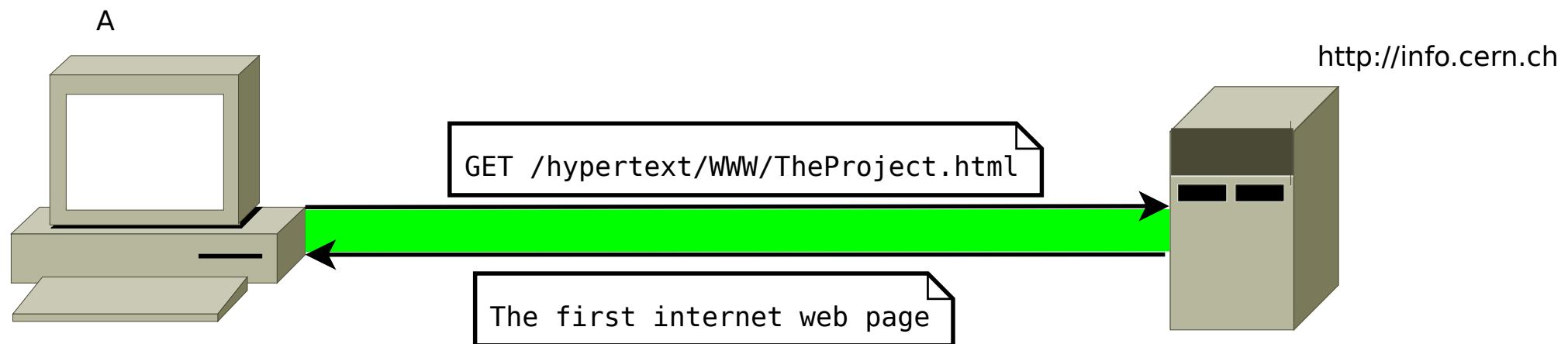


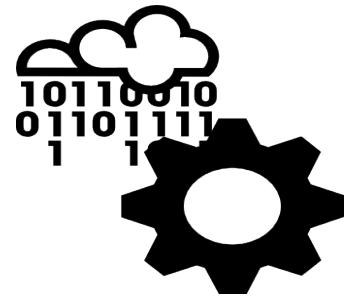
Diagramma realizzato con DIA



Le definizioni dei protocolli di comunicazione solitamente incorporano le specifiche dei *formati* da utilizzare per codificare i documenti trasmessi.

```
<HTML>
  <HEADER>
    <TITLE>The World Wide Web project</TITLE>
    <NEXTID N="55">
  </HEADER>
<BODY>
  <H1>World Wide Web</H1>
  The WorldWideWeb (W3) is a wide-area
  <A NAME=0 HREF="WhatIs.html">hypermedia</A>
  information retrieval initiative aiming to give universal
  . . .
```

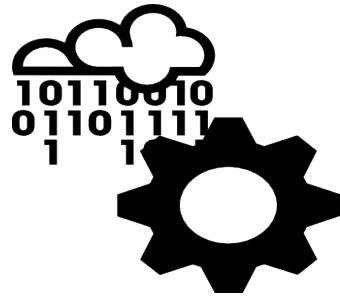
Frammento dai *sorgenti* della prima pagina Web della storia



RSS (*Really Simple Syndication* oppure *Rich Site Summary*) permette di pubblicare elenchi di contenuti presenti su internet. Viene usato generalmente per condividere news e aggiornamenti dei siti web.

La prima versione di questo protocollo RSS 0.9 risale al 1999. La versione attualmente in uso è RSS ATOM (RFC 4287 e RFC 5023).





Usando un *aggregatore* di feed RSS è possibile ricevere notifiche e visionare i contenuti dei feed sottoscritti. Il fatto che RSS sia un formato aperto ha permesso la creazione di svariati feed reader.



Usando un *aggregatore* di feed RSS è possibile ricevere notifiche e visionare i contenuti dei feed sottoscritti. Il fatto che RSS sia un formato aperto ha permesso la creazione di svariati feed reader.

Il formato RSS è molto diffuso per la condivisione di notizie e di post sui blog.

- **La Repubblica** <http://www.repubblica.it/rss/homepage/>
- **Gira Sicilia** <http://www.girasicilia.it/feed/>
- **Università di Catania** http://unict.it/in_evidenza.xml
- **Officina Rebelde** <http://www.officinarebelde.org/spip.php?page=backend>
- **Libera**
<http://www.libera.it/flex/cm/pages/ServeFeed.php/L/IT/fmt/rss20/feed/pages%3A1>

Formato dei feed



All'interno di un feed vengono enumerati alcuni contenuti disponibili su internet. Per ognuno di questi va indicata la URL, il titolo e la data di pubblicazione.

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">

    <title>Example Feed</title>
    <link href="http://example.org/" />
    <updated>2003-12-13T18:30:02Z</updated>
    <author>
        <name>John Doe</name>
    </author>
    <id>urn:uuid:60a76c80-d399-11d9-b93C-0003939e0af6</id>

    <entry>
        <title>Atom-Powered Robots Run Amok</title>
        <link href="http://example.org/2003/12/13/atom03" />
        <id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344efa6a</id>
        <updated>2003-12-13T18:30:02Z</updated>
        <summary>Some text.</summary>
    </entry>

</feed>
```



GRAZIE

Cristiano Longo - opendatahacklab

<mailto:cristianolongo@gmail.com>

Hackspace Catania - Via Grotte Bianche 112



GRAZIE

Cristiano Longo - opendatahacklab

<mailto:cristianolongo@gmail.com>

Hackspace Catania - Via Grotte Bianche 112

PS: il 4 Marzo è l'Open Data Day