GIS e Geo WEB: piattaforme e architetture

Docente: Cristoforo Abbattista
eMail: abbattista@planetek.it

Il paradigma di riferimento del corso di laurea magistrale
“Sistemi informativi territoriali e Telerilevamento”

Le Immagini

I Sensori

Il Tempo reale

Presentazione del modulo didattico

Il paradigma

Le immagini

- Mappe prima di tutto
- Ma anche il design

I sensori

- Ogni cosa e ogni persona produce informazione
  - Troviamola (push or pull), acquisiamola (IT), aggiustiamola (GIS) e condividiamola (GeoWEB)

Il tempo reale

- Internet
- Mobile
- Accessibilità

Presentazione del modulo didattico

Obiettivi formativi

- Progettazione, messa in esercizio e gestione di un SIT
- Soluzioni GeoWEB
- Singoli elementi che compongono un SIT
- Caratteristiche dei servizi da erogare
- Tecnologie disponibili
- Servizi pubblici e privati
- Standard internazionali (ISO, OGC, INSPIRE…)
- Mode o buone pratiche?

Organizzazione

- 7 micro-moduli in 50 ore x 6 Crediti formativi
- Calendario
  - lezioni da 4, 6 o
  - 8 ore 

Presentazione del modulo didattico

7 micromoduli

- Componenti delle piattaforme
- Architetture dei SIT
- Linguaggi di sviluppo
- Web service
- Interoperabilità e standard
- Soluzioni commerciali e FOSS
- Gestione dei sistemi

Modalità d‘esame

- Presentazione Tesina
- Esame scritto
  - Con presentazioni aggiuntive

Presentazione del modulo didattico

I micro-moduli

Componenti delle piattaforme

- Panoramica sull'infrastruttura hardware e software necessaria per la messa a punto di un Sistema Informativo. Struttura e funzionamento di un calcolatore; gli archivi di massa; interazione con le periferiche. Processi ad alte performance. Reti LAN e WAN. Ethernet. Il protocollo TCP/IP
- La sicurezza dei sistemi. Sistemi di autenticazione e firma digitale. Sistemi di controllo di accesso.

Architetture dei SIT

- Componenti; architetture centralizzate e distribuite; caratteristiche di una architettura (affidabilità, disponibilità, prestazioni, scalabilità, sicurezza). Definizione di una Infrastruttura di Dati Territoriali (IDT). Le sfide di un SIT: la mole dell'informazione, la sua numerosità, la velocità di cambiamento. Il dimensionamento dei sistemi.

Presentazione del modulo didattico

I micro-moduli

Linguaggi di sviluppo

- Metodologie di sviluppo standard dei sistemi. Panoramica sullo standard di progettazione UML. I principali linguaggi di sviluppo lato server. I Content Management System. Il DHMTL e il linguaggio Javascript. XML. Mash-up. Cenni sul Semantic Web. Cenni sulle metodologie di sviluppo Agili.

Web service

- Il paradigma SOA (Sercvice Oriented Architecture). Cosa sono i Web service. Come funzionano. Vantaggi, Svantaggi e problematiche. Il protocollo SOAP. Il protocollo RESTful.

Presentazione del modulo didattico

I micro-moduli

Interoperabilità e standard

- Perchè essere Interoperabili. Gli standard e gli organismi: ISO, OGC, W3C. Approfondimento sugli standard OGC più importanti: WMS, WFS, WMC, CS-W, WCS, WPS, GML, SLD, KML, OLS. SensorWeb. La direttiva INSPIRE: le implementation rules e le technical guides.

Soluzioni commerciali e FOSS

- I top player presenti sul mercato: dati e servizi, analisi delle soluzioni, conformità agli standard, ambienti di sviluppo. Un top player particolare: "il mondo FOSS". Un nuovo attore: il cittadino (neo-geography).

Gestione dei sistemi

- La messa in esercizio. La struttura organizzativa. Il Service Level Agreement e l'utilizzo dei KPI (Key Performance Indicator). La manutenzione.

Studiate il programma e ...
fate le vostre richieste

Bibliografia

- GIS for Web Developers: Adding 'Where' to Your Web Applications
  Scott Davis
- Geospatial Services and Applications for the Internet
  John T. Sample, Kevin Shaw, Shengru Tu, Mahdi Abdelguerfi
- Geographical Information Systems: Principles, Techniques, Management and Applications
  Paul A. Longley, Michael F. Goodchild, David J. Maguire, David W. Rhind
- Spatial Database Systems: Design, Implementation and Project Management
  Albert K.W. Yeung, G. Brent Hall
- Spatial Data on the Web: Modeling and Management
  Alberto Belussi, Barbara Catania, Eliseo Clementini, Elena Ferrari
- Web mapping illustrated
  Tyler Mitchell
- The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society
  Arno Scharl, Klaus Tochtermann
- Spatial Portals: Gateways to Geographic Information
  Winnie Tang, Jan Selwood
- Research and Theory in Advancing Spatial Data Infrastructure Concepts
  Harlan Onsrud
- Next Generation Geospatial Information: From Digital Image Analysis to Spatiotemporal Databases
  Peggy Agouris, Arie Croitoru
- http://www.opengeospatial.org/
- http://inspire.jrc.ec.europa.eu/index.cfm

I lezione
Componenti delle piattaforme

GIS e Geo WEB: piattaforme e architetture

Il Sistema Informativo

Componenti

- Informazioni ed eventi
  - Input esterni
  - Accrescimento/generazione interna
  - Output per l'esterno
- Processi, regole ed obiettivi
  - Politiche/regolamenti/leggi
- Sistema informatico
  - Tecnologie a supporto dei processi di archiviazione, elaborazione e scambio delle informazioni

Il Sistema Informatico

Componenti

- Hardware centralizzato e/o distribuito
- Sistema di comunicazione di rete
- Software di ambiente e di sistema
- Basi di dati
- Applicazioni centralizzate e/o distribuite
- Interfaccia utente


Hardware
La ferraglia

Hardware

Tipologie

- **Server** - Computer con grandi capacità di calcolo e memoria.
- **Workstation** - Computer di medie capacità
- **Laptop -** Personal computer portatile
- **PDA –** Dispositivi mobili - smartphone

Componenti

- CPU
- Memoria
- Periferiche
- BUS

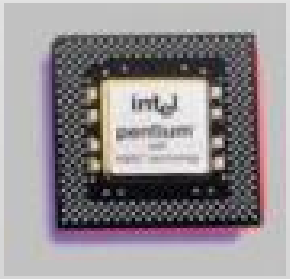
Hardware
CPU

Compiti

- Manipolazione dei dati.
- Coordinare e controllare l'esecuzione dei comandi
  - I comandi vengono interpretati secondo regole precise e differenti per ogni tipo di microprocessore.
- Trovare i dati e le istruzioni dalla memoria interna
- Collocarli nei registri, per poter elaborare
- Copiare i risultati nella memoria RAM

Potenza

- Velocità del clock in MHz (milioni di cicli per secondo)
- Milioni di operazioni che svolge in un secondo (MIPS)
- Potenza delle istruzioni in linguaggio macchina
- Capacità di indirizzamento (8, 16, 32 o 64 bit)



Hardware
La Memoria

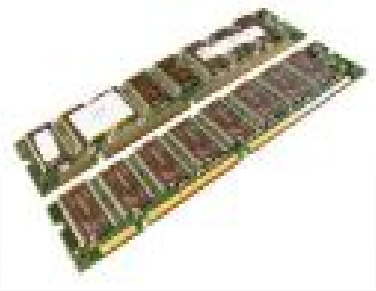
Come si misura

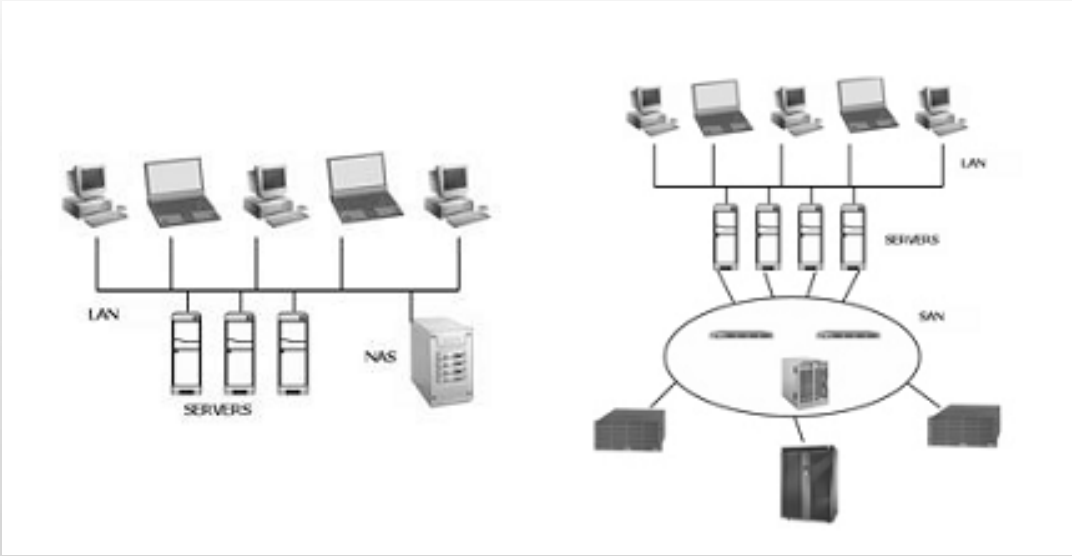
- 1 byte = 8 bit (binary digit)
- 1 Kb (chilo byte) = 1024 byte
- 1 Mb (mega byte) = 1024 Kb;
- 1 Gb (giga byte) = 1024 Mb;
- 1 Tb (tera byte) = 1024 Gb
- ……

Hardware
La Memoria

Tipologie

- Ram (Random Access Memory)
  - Cache memory
- Rom (Read Only Memory)
  - BIOS
- Memorie di Massa
  - Hard Disk
  - Network Attached Storage (NAS)
    - Ethernet
  - Storage Area Network (SAN)
    - Fibra
  - Floppy Disk
  - CD / DVD
  - Pen Drive
  - MMC – SD – XD
  - DOM
  - TAPE





Hardware
La Memoria

Caratteristiche

- Capacità
- Tempi d’accesso
- Costo

Compiti

- Contenere dati e istruzioni
- On-line – RAM e Cache
- Off-line – Storage

Hardware
Le periferiche

INPUT

Scanner

Mouse

Tastiera

Joystick

Cam

OUTPUT

Monitor

Stampante

Masterizzatore

Si collegano con dei cavi

ma oggi possono essere wi-fi

Si usano grazie ai driver software

INPUT-OUTPUT

Router

Modem

Floppy disk

Cd

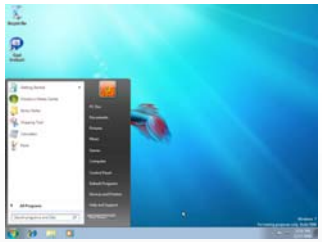
Pen Drive

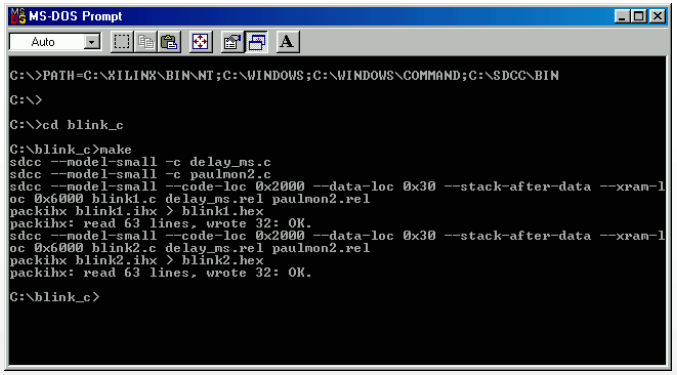
LAN adapter

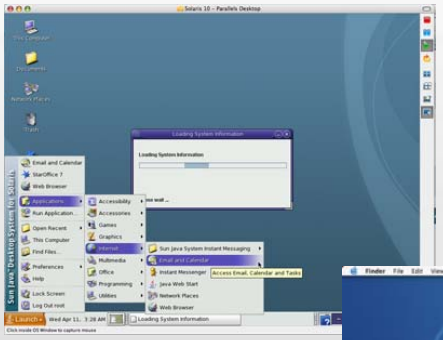

Software
Confusione controllata

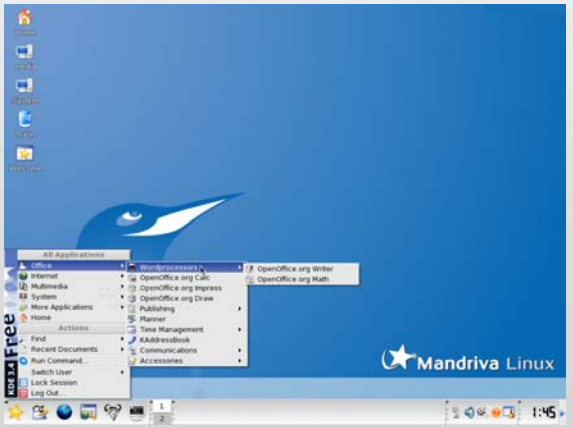
Software

Software di base



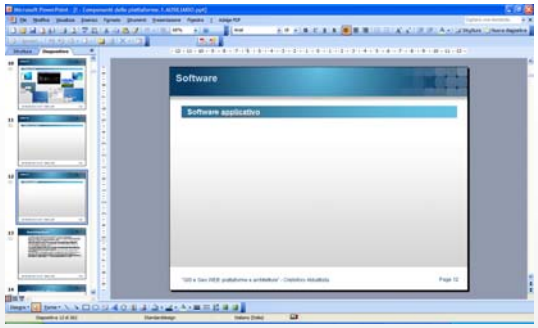


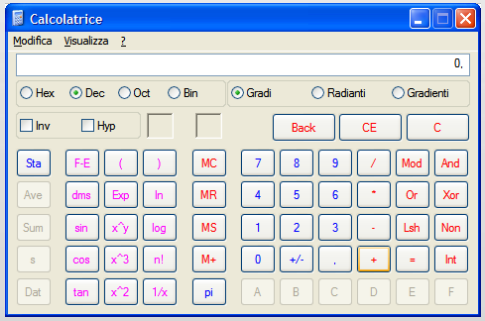


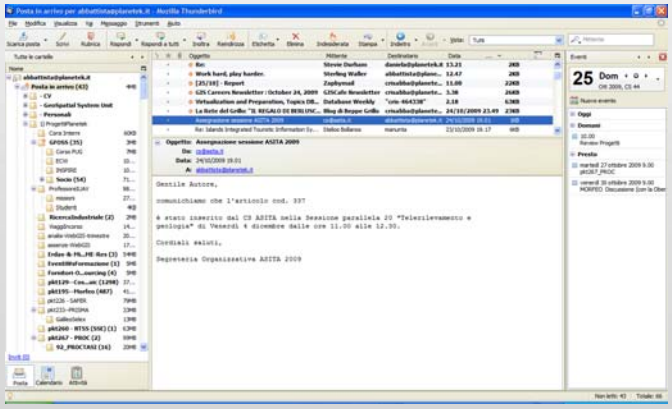


Software

Software applicativo

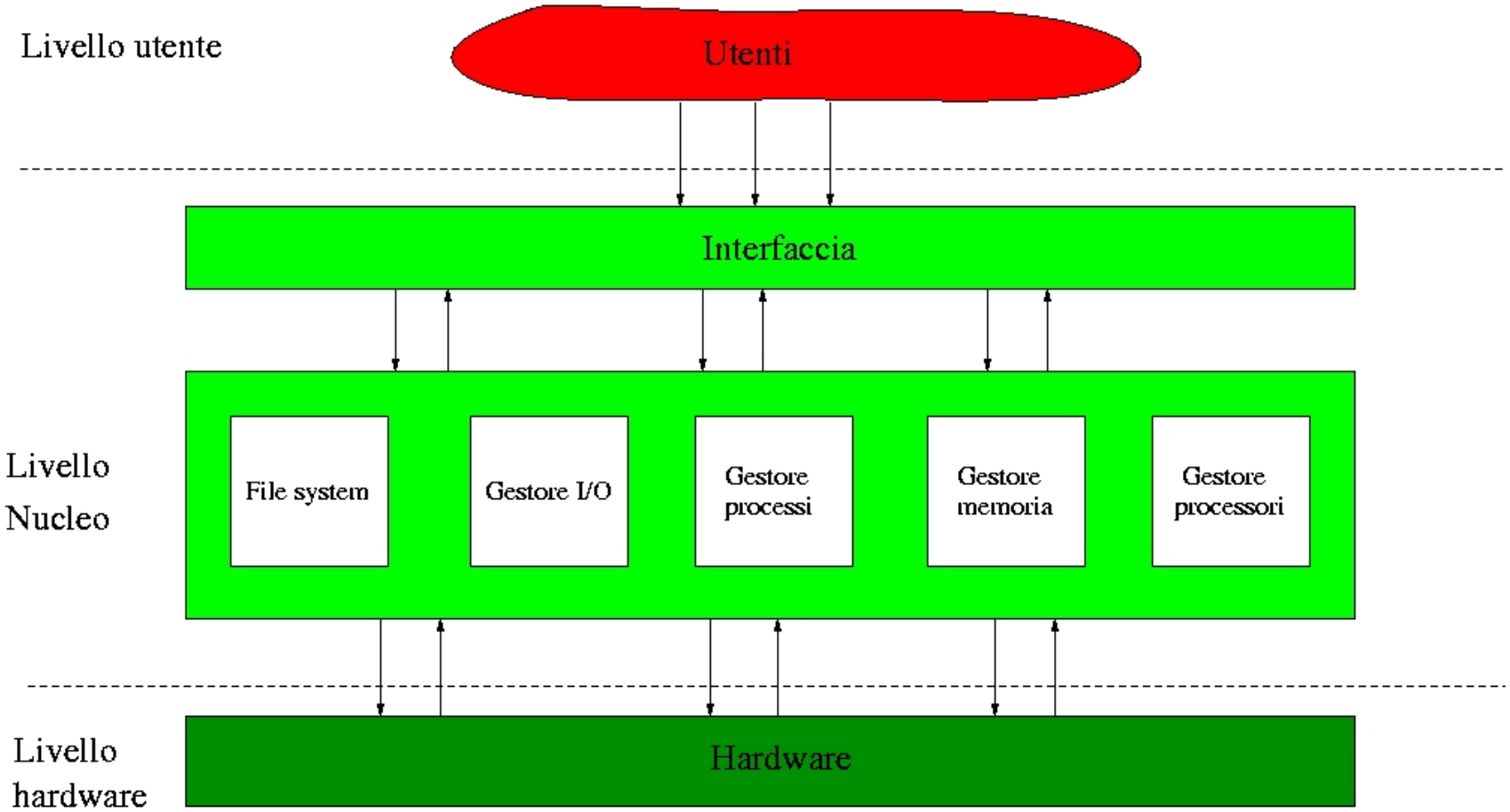






Il Sistema Operativo
Ci pensa lui!

Struttura logica



Quali sono le funzioni di un SO ?

Esegue applicazioni

- carica il programma binario prodotto della compilazione (e residente su disco) nella RAM,
- concede il processore all'applicazione da eseguire

Facilita l'accesso ai dispositivi di I/O

- interagisce con le periferiche facendosi carico di tutti i dettagli fisici (es. modem, reti, video…)
- mette a disposizione operazioni di lettura/scrittura

Quali sono le funzioni di un SO ?

Archivia dati e programmi

- mette a disposizione dell'utente una visione del file system basato sulle astrazioni di file e directory
- gestisce queste astrazioni sul supporto fisico nelle operazioni di lettura/scrittura dei settori fisici

Gestisce le risorse

- ripartisce le risorse disponibili (processore, RAM, periferiche) fra le varie applicazioni
- evita che ci siano malfunzionamenti dovuti all'uso contemporaneo di risorse
- ottimizza le prestazioni attraverso politiche idonee

Quali sono le funzioni di un SO ?

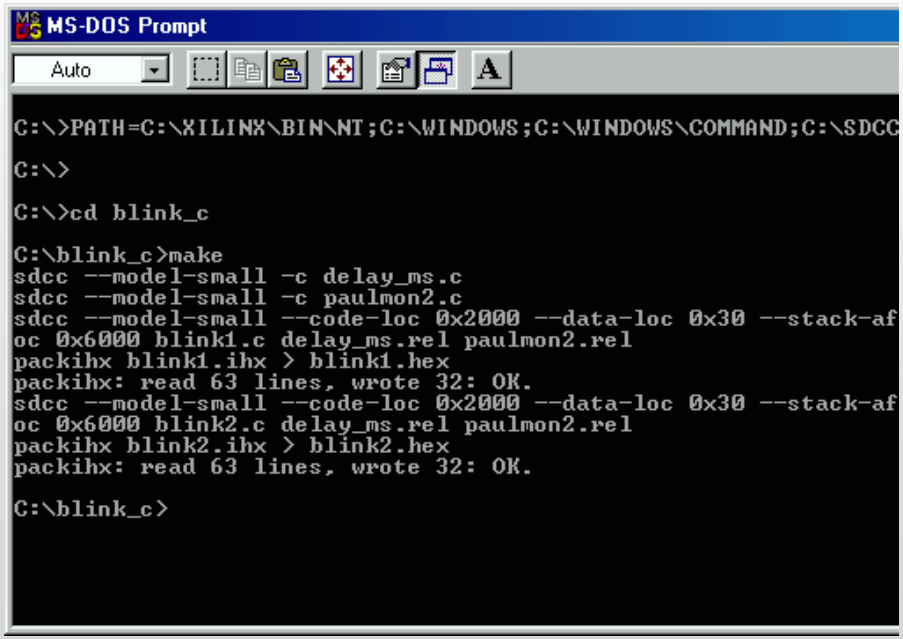
Gestisce malfunzionamenti del sistema

- rileva e gestisce situazioni anomale
  - es: se il disco ha un settore difettoso, il SO può evitare che venga usato e trasferire le informazioni residenti su quel settore da un'altra parte
  - es: se un'applicazione cerca di effettuare una operazione non permessa (come leggere i dati di un'altra applicazione) può bloccare l'applicazione segnalando all'utente la situazione erronea

L’interfaccia utente

Command line o Graphical User Interface

- È un programma (shell) che permette all'utente di interagire con il computer.
  - Interfaccia grafica
  - A linea di comando



Esecuzione di un programma

- Il SO ricopia lo spazio di indirizzamento di un programma dalla memoria di massa alla RAM

Una possibile organizzazione della RAM con più programmi attivi contemporaneamente

Area riservata, non accessibile in modalità utente

Esecuzione di un programma

- Il SO modifica il Program Counter (PC) del processore in modo che punti all'indirizzo della prima istruzione assembler da eseguire nel nostro programma 2
- Il processore esegue una dopo l'altra le istruzioni assembler che lo compongono

Interruzione di un programma

Terminazione

- Un processo termina :
  - Quando esegue un'istruzione assembler di terminazione
  - Quando effettua una operazione illecita
  - Quando c'è un errore
- In questi casi il processore ricomincia ad eseguire il sistema operativo

Interruzione

- Il sistema operativo può bloccare un processo attraverso le interruzioni
  - le periferiche possono 'richiedere attenzione' attraverso segnali di *interruzione* usando le linee di controllo del bus
  - dopo ogni istruzione il processore controlla la presenza di una interruzione
  - se è presente un'interruzione il controllo passa automaticamente al sistema operativo che la deve gestire

Il gestore del processore

Multitasking

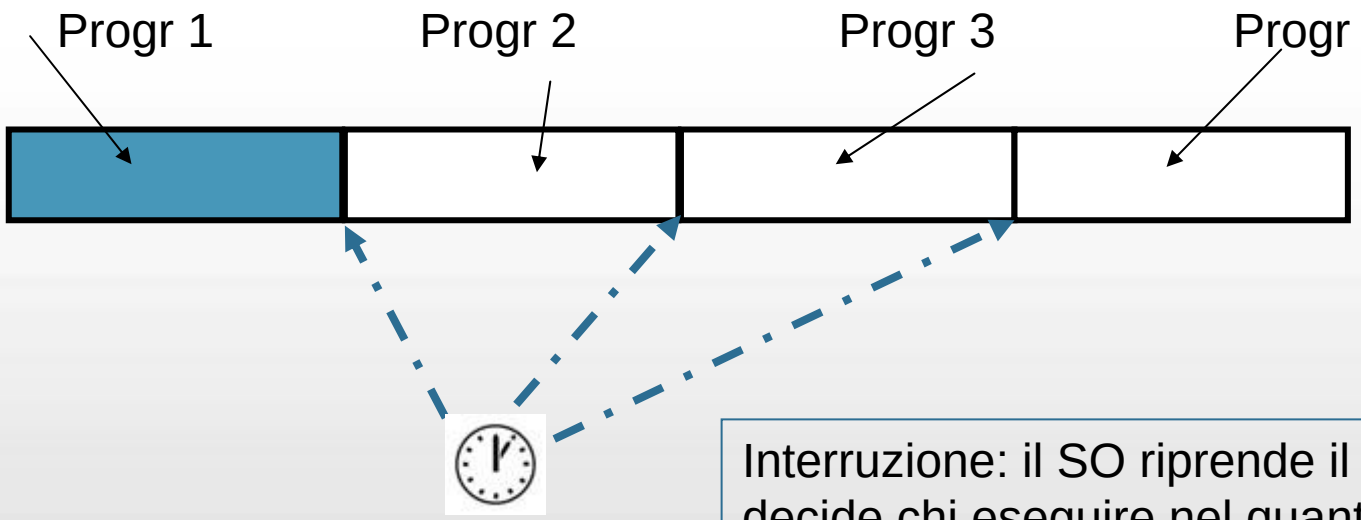
- Sia i programmi di SO che i programmi applicativi ‘avviati’ vengono eseguiti contemporaneamente
- Il gestore del processore si preoccupa di condividere il processore tra tutti i programmi attivi secondo politiche sensate
- Ogni programma pensa di avere un proprio processore

Multiutenza

- Più programmi di più utenti
- Senza che vi siano conflitti o operazioni non consentite

Il gestore del processore

Esecuzione ciclica



Interruzione: il SO riprende il controllo e decide chi eseguire nel quanto successivo

Periferica ‘clock interno’

Il gestore della memoria

Allocazione della memoria

- Sia i programmi di SO che i programmi applicativi ‘avviati’ usano contemporaneamente la RAM
- Il gestore della memoria si preoccupa di condividere la RAM tra i vari processi così che :
  - ogni processo abbia il suo spazio distinto ed inaccessibile agli altri
  - ogni processo abbia la memoria sufficiente per eseguire il proprio codice algoritmico e raccogliere i propri dati

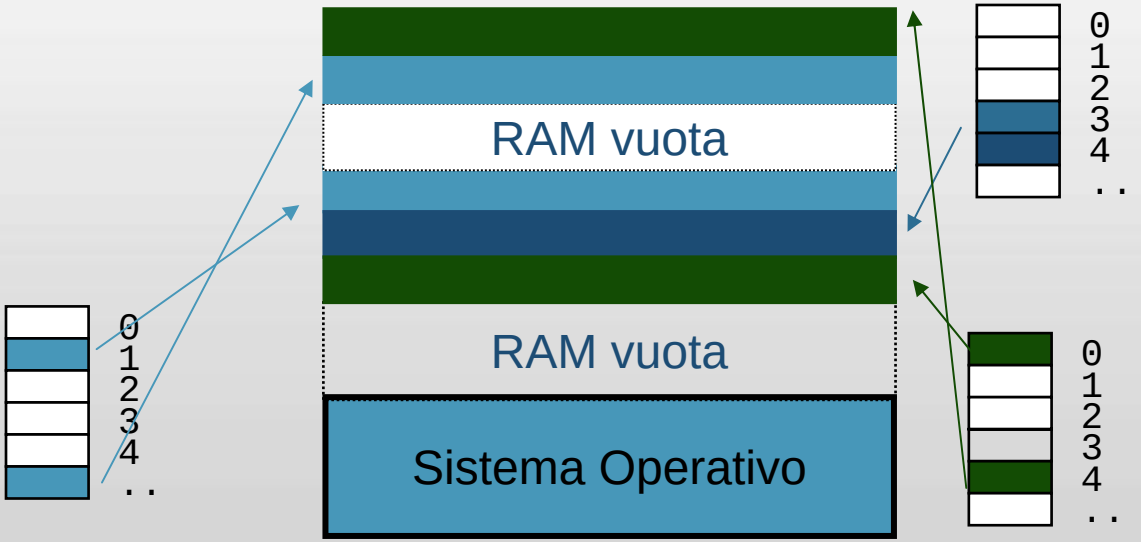
Il gestore della memoria

Allocazione Statica

- Ricopiare in memoria RAM tutto il codice archiviato in memoria di massa ed eseguirlo
- Problema: non si possono eseguire programmi con spazio di indirizzamento più grande della RAM!

Allocazione Dinamica – Memoria virtuale

- Ricopiare in memoria RAM solo i pezzi che servono per l'esecuzione corrente (pagine)
- Ogni programma è diviso in pagine della stessa dimensione



Il gestore delle periferiche

Controller

- Hardware che dialoga direttamente con la della periferica
  - Se non è generico si acquista con la periferica
  - Oppure può sfruttare le porte di comunicazione generiche con seriale, parallela, usb, ps2

Driver

- Software di sistema operativo che conosce le caratteristiche della periferica ed è capace di dialogare con il controller
  - Viene sviluppato da chi costruisce la periferica o il controller

Il gestore delle periferiche

Interrupt

- Interrupt di processo
  - Un processo che deve accedere ad una periferica manda un interrupt (interrupt software) e si mette in attesa
  - la CPU passa il controllo al kernel, che provvede all'I/O per conto del processo
  - la CPU fa ripartire il processo stesso o un altro se lo scheduling lo prevede
- Interrupt di periferica
  - Una periferica che ha terminato il suo lavoro ed ha bisogno della CPU (ad esempio l'utente ha battuto un tasto) manda un interrupt (interrupt hardware) alla CPU
  - la CPU interrompe il processo in corso e passa il controllo al sistema operativo, che tratta i dati letti dalla periferica e fa quanto necessario

Il gestore delle periferiche

Lo spooler di stampa

- Risolvere il problema delle stampe su carta.
  - Le stampanti sono lente
  - La CPU non può aspettare che la stampante finisca
- Non si lancia la stampa, ma un programma a priorità molto bassa che simula la stampante
- Lo spooler accumula i dati da stampare in memoria, e poi si fa carico di stamparli realmente

Il gestore del File System

File System

- È l'archivio di tutti i dati e i programmi in modo persistente
- Ogni SO ha un proprio tipo
  - FAT, FAT32, NTFS, NFS, Ext, Ext2, Ext3, Ext4, UDF, HFS, Joliet, ecc.

Il gestore del file system

- Attraverso le astrazioni di file e directory ed una gerarchia ad albero permette di:
  - creare file e directory con appositi comandi di SO
  - organizzare dati e programmi in modo da renderne semplice la localizzazione da parte dell'utente umano
  - Modificare la gerarchia per creare, aggiornare, eliminare file o directory
  - Associare all'estensione il programma che deve eseguire il dato (non è però uno standard)

Il gestore del File System

Sicurezza

- Ogni file appartiene ad un utente
- Il proprietario può specificare chi può fare cosa su ognuno dei propri file
  - Lettura, scrittura ed esecuzione
  - Il gestore del file system controlla la correttezza dell'operazione

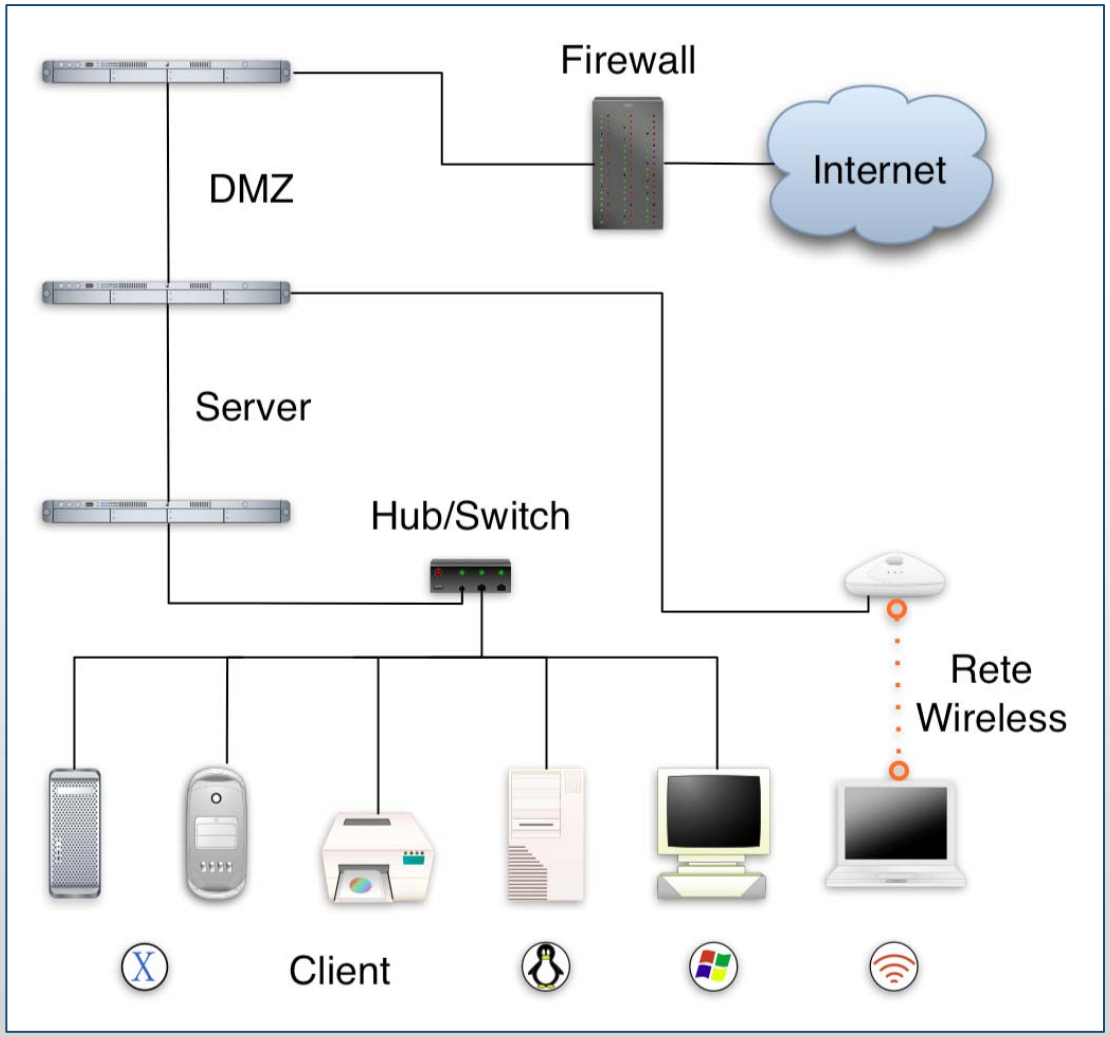
Diritti dell'amministratore

- Root o Administrator
  - Può accedere in maniera non ristretta dappertutto
  - Bisogna proteggerlo dagli attacchi


Le reti!
...finora abbiamo scherzato

Cos’è una rete

- Fili intersecati tra loro
  - cavi, telefono, satellite, wi-fi
- Ogni intersezione è un nodo della rete.
- In una rete di comunicazione, le informazioni vengono trasferite da un nodo all'altro.
- La trasmissione può essere
  - monodirezionale
  - interattiva
  - broadcast
  - punto-punto
  - Punto-multipunto



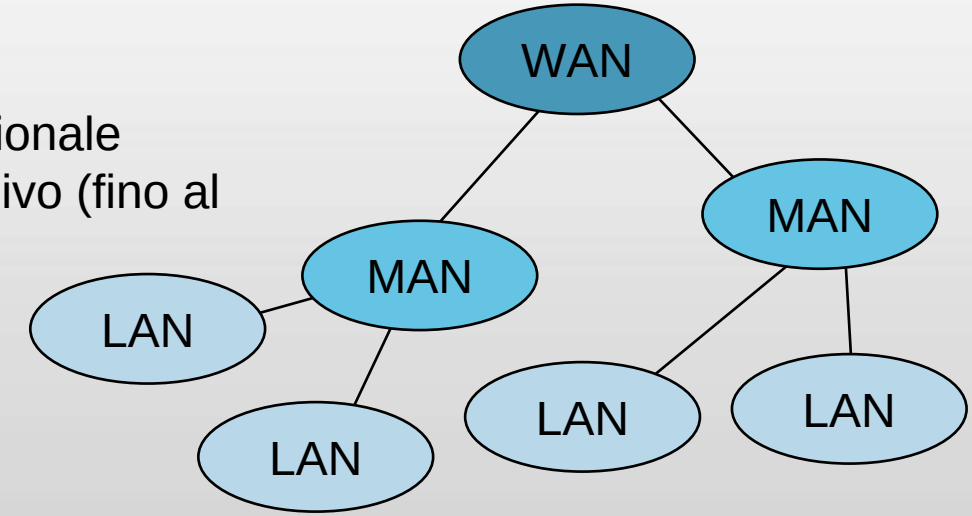
Classificazione delle reti

Massive Parallel	0.1 m	Piastra
Multi-Processor	1 m	Sistema
Cluster	10 m	Stanza
Reti Locali	100 m	Edificio
Reti Locali Estese	1 km	Comprensorio
Reti Metropolitane	10 km	Città
Reti Geografiche	100 km	Nazione
Interconnessione di reti geografiche	1000 km	Continente
	10.000 km	Pianeta

Classificazione delle reti

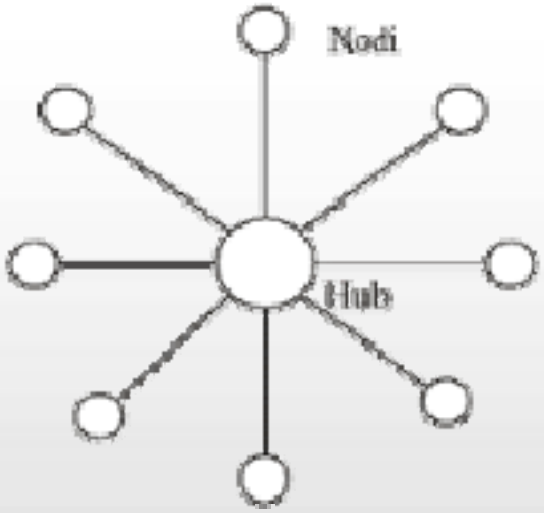
Nomenclatura

- LAN: Local Area Network
  - Rete senza attraversamento di suolo pubblico
- MAN: Metropolitan Area Network
  - Rete in ambito cittadino con disponibilità di canali trasmissivi veloci
- WAN: Wide Area Network
  - Rete in ambito nazionale o internazionale utilizzante qualsiasi mezzo trasmissivo (fino al satellite).



Topologie delle reti

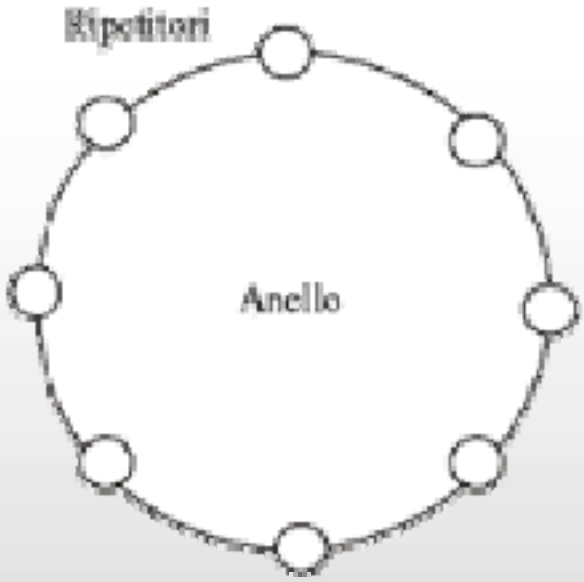
A stella



Tutti parlano con tutti

Topologie delle reti

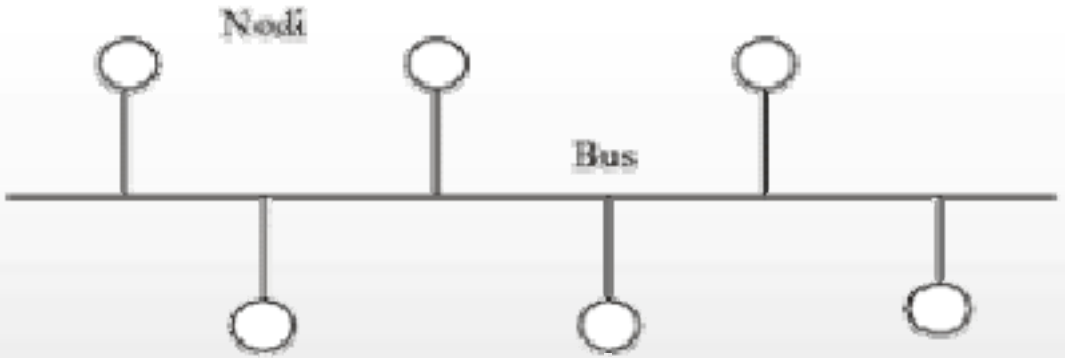
Ad anello



La stessa informazione a tutti
(ad esempio nella tv via cavo)

Topologie delle reti

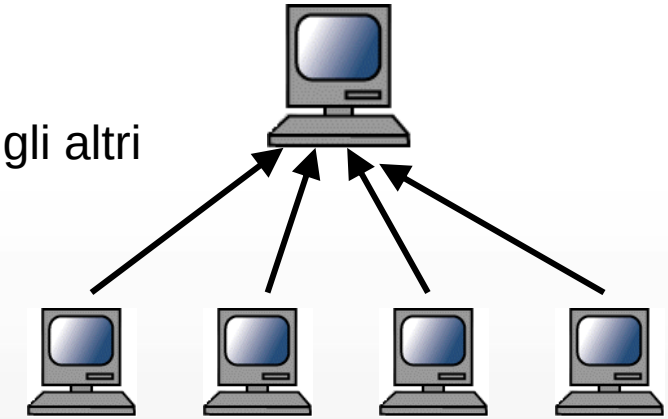
A BUS

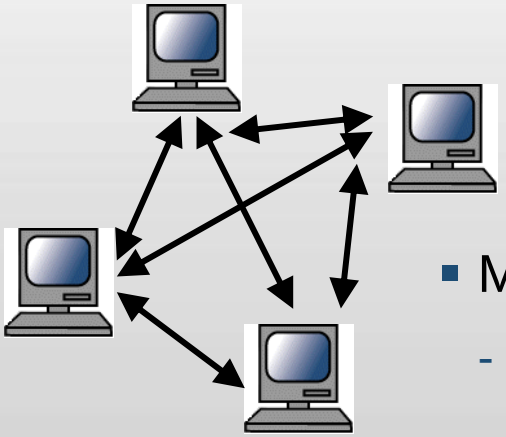


- Ogni nodo trasmette sul bus i propri dati che si propagano andando a toccare tutti i nodi rimanenti.
- Non ci sono nodi di controllo.
- Ogni nodo deve incorporare l'intelligenza necessaria per il controllo di flusso e il recupero in caso di errori.

Modelli di rete

- Modello client-server
  - un computer “possiede” le risorse, e gli altri vi accedono.





- Modello peer-to-peer
  - ogni computer può fungere da client e da server.

Scopo delle reti

Condividere le risorse

- risorse fisiche: stampanti, dischi, nastri, ecc.
  - Print server
  - File server
- risorse di calcolo: programmi residenti sul computer “remoto”
  - Application server

Comunicare informazioni

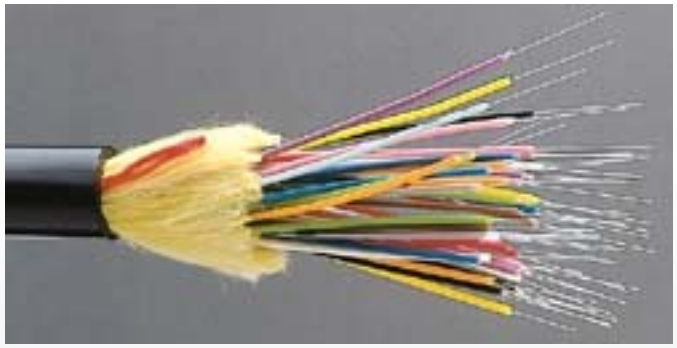
- posta elettronica
- chat (testuale, audio, video)
- messaggistica istantanea
- spazi di lavoro condivisi (groupware)

Elementi della comunicazione

- Mittente
- Destinatario
- I dati
- Il Mezzo trasmissivo
- Regole di linguaggio

Tipologie di mezzi trasmissivi

- Elettrici
  - Cavo coassiale
  - Cavi UTP, STP
- Ottici
  - Fibra ottica
  - Raggio laser
- Radio
  - Ponti radio
  - Satelliti
  - Reti cellulari



Fibra ottica

Thicknet Coassiale UTP

Protocolli e modi di comunicazione

Il protocollo

- Stessa lingua
- Stesse regole

Il modo

- Commutazione di circuito
  - Due fili vengono messi a contatto fisico
- Commutazione di pacchetto
  - L’informazione viene divisa in parti numerate (pacchetti) che contengono l'indirizzo del mittente e l'indirizzo del destinatario.
  - Quando i pacchetti arrivano, il computer-destinatario controlla che ci siano tutti, se ne manca qualcuno lo richiede nuovamente ed infine ricostruisce l'informazione

Il modello OSI

Open Systems Interconnection

- La prima rete di dati fu creata negli USA alla fine degli anni ‘60 dal DoD (Department of Defense)
- Ognuno usava SW e HW diverso
- Nel 1984 la ISO (International Organization for Standardization) rilasciò il modello OSI (Open Systems Interconnection).
- Suddivisione delle funzioni in 7 livelli (layers).

7. Applicazione (Application layer)
6. Presentazione (Presentation layer)
5. Sessione (Session layer)
4. Trasporto (Transport layer)
3. Rete (Network layer)
2. Collegamento dati (Data link layer)
1. Fisico (Physical layer)

I livelli del modello OSI

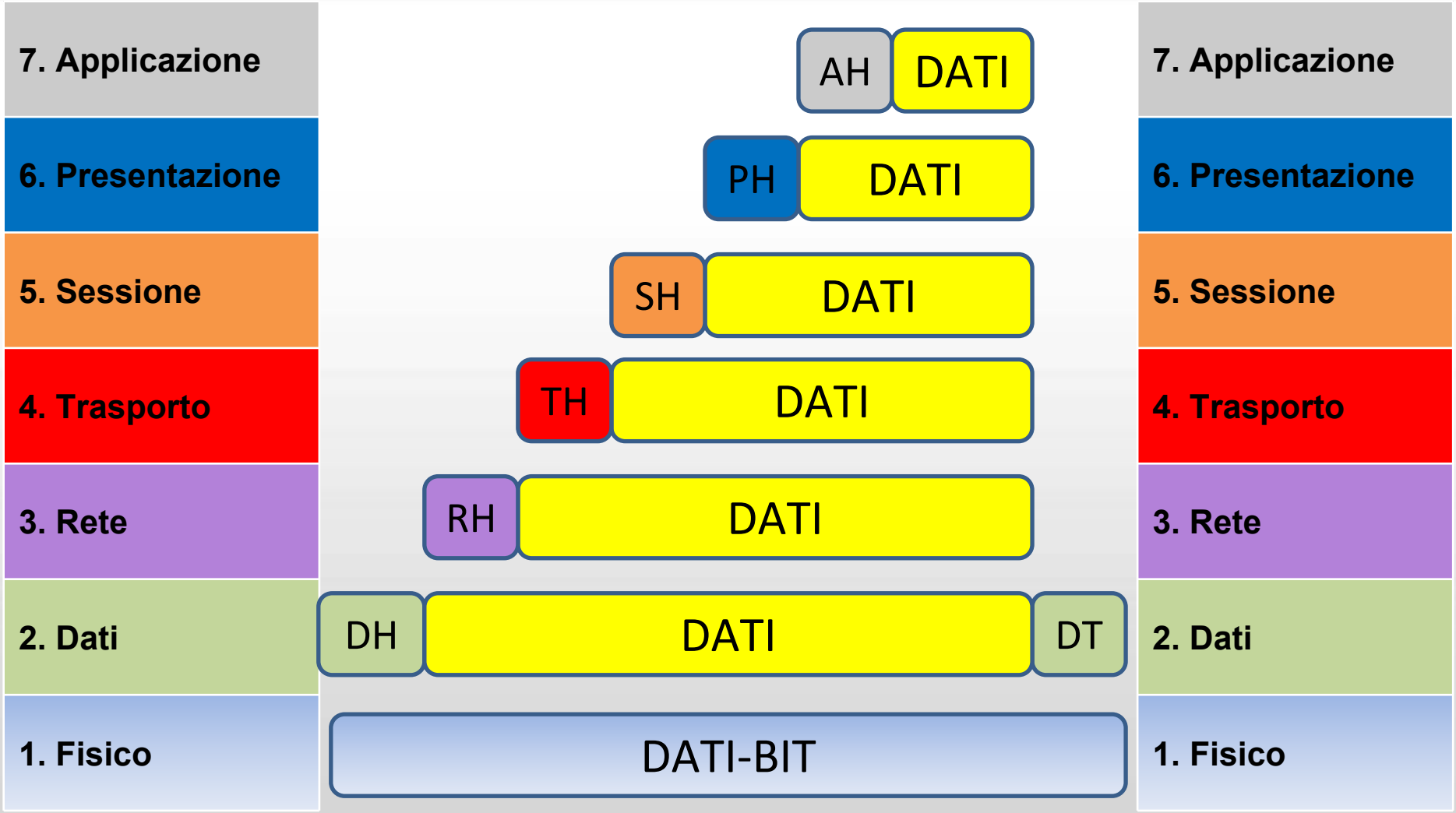
Lo scopo dei livelli

- Fornire servizi al livello superiore (in ricezione)
- Fornire dati al livello inferiore (in trasmissione)
- Due computer dialogano tra livelli con lo stesso numero
- I servizi messi a disposizione da un livello a un livello superiore, sono forniti, attraverso un'interfaccia software, in modo mascherato.
- In fase di trasmissione, viceversa, ogni livello passa dati e informazioni di controllo al livello sottostante, fino a quando si raggiunge il livello fisico.

L'incapsulamento

- **Encapsulation:** processo di apporre delle informazioni aggiuntive, tramite **headers** e **trailers**, ai dati da trasmettere.

Encapsulation



Esempio

Un messaggio e-mail spedito da un mittente

e ricevuto da un destinatario

Esempio

- **L'utente agisce a livello Applicazione**
- **Attraverso un sistema operativo e un client di posta elettronica genera dei dati.**

Esempio

- **A livello di Presentazione il messaggio viene codificato secondo un preciso standard.**

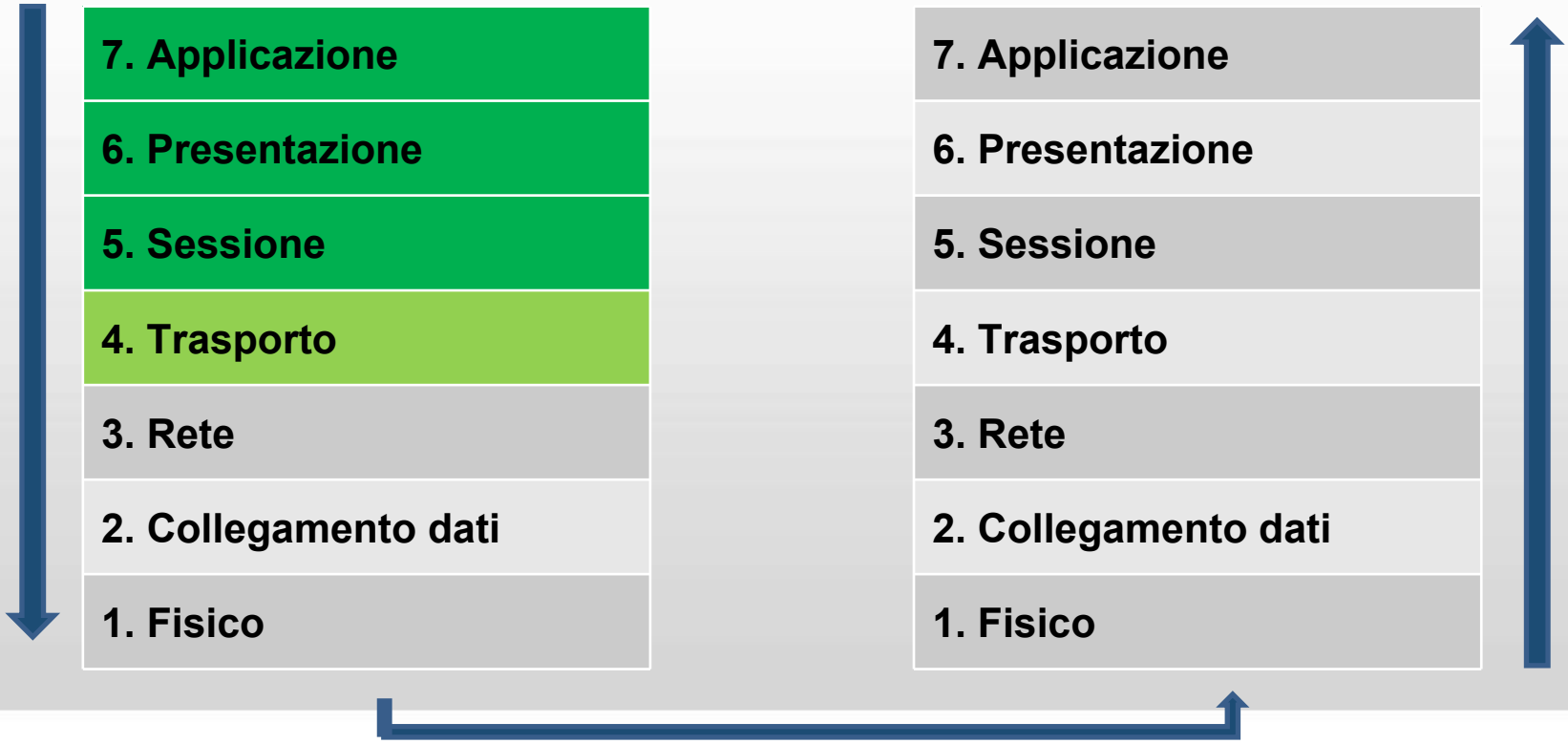
7. Applicazione	7. Applicazione
6. Presentazione	6. Presentazione
5. Sessione	5. Sessione
4. Trasporto	4. Trasporto
3. Rete	3. Rete
2. Collegamento dati	2. Collegamento dati
1. Fisico	1. Fisico

Esempio

- **Il livello Sessione apre sessioni di comunicazione virtuali**
- **Si occupa, inoltre, di gestire la sincronizzazione di tutta la comunicazione.**

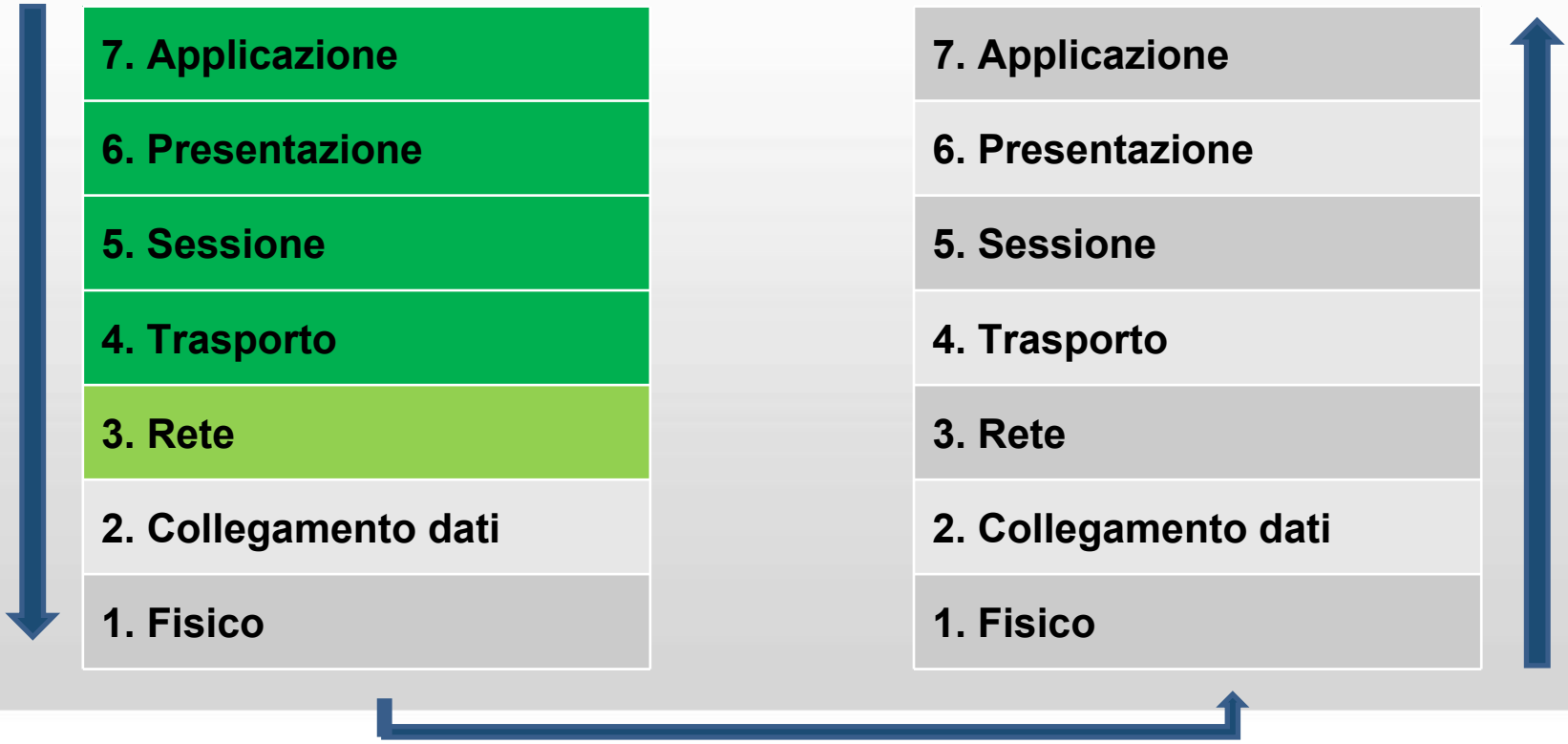
Esempio

- **Il livello Trasporto ha il compito di ridurre gli effetti negativi dei servizi offerti dallo strato di rete sottostante.**
- **Se la rete ha problemi il livello di trasporto si occupa di rimandarlo senza generare duplicati.**



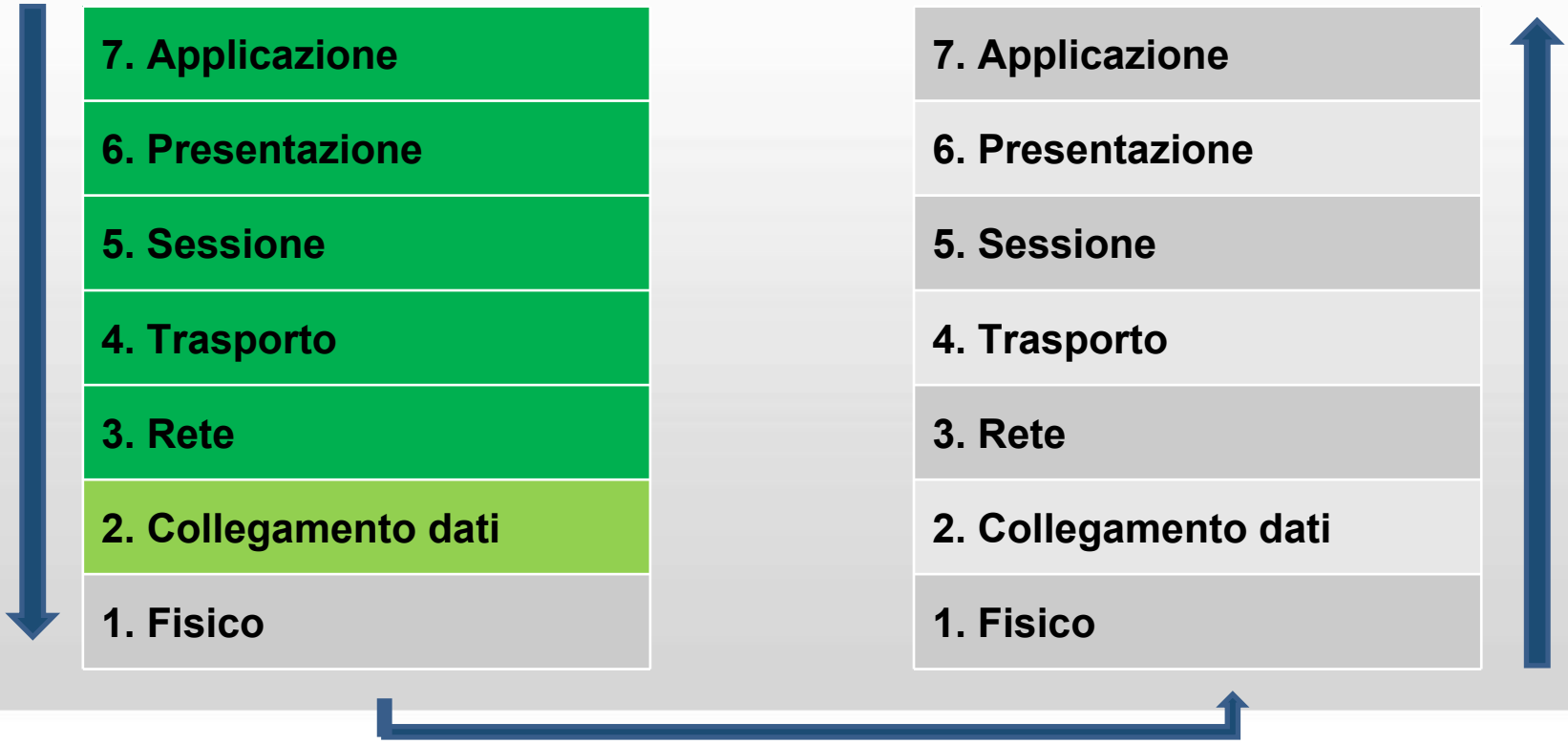
Esempio

- **Il livello Rete determina il modo in cui i messaggi sono instradati dal nodo di provenienza a quello di destinazione**
- **I percorsi possono essere basati su tabelle statiche o essere impostati dinamicamente ad ogni trasmissione.**



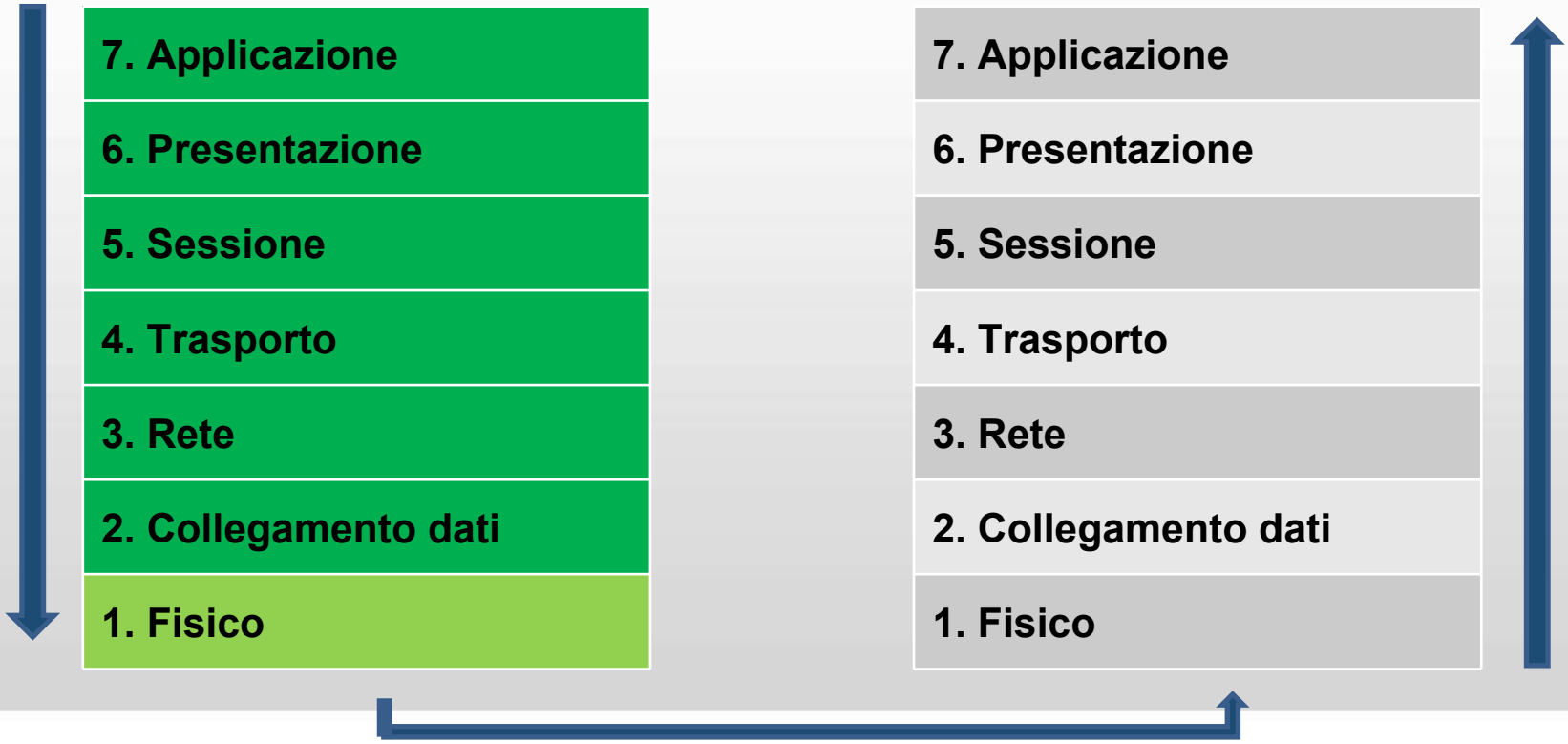
Esempio

- **Il livello Collegamento dati fornisce al livello Rete una linea esente da errori di trasmissione;**
- **Per ottenerla gestisce tutta una serie di parametri riguardanti il controllo degli errori e dei flussi sulla linea.**



Esempio

- **Il Livello Fisico riceve il messaggio sotto forma di BIT (0 ed 1) e lo invia attraverso il canale di trasmissione a cui è connesso.**



Esempio

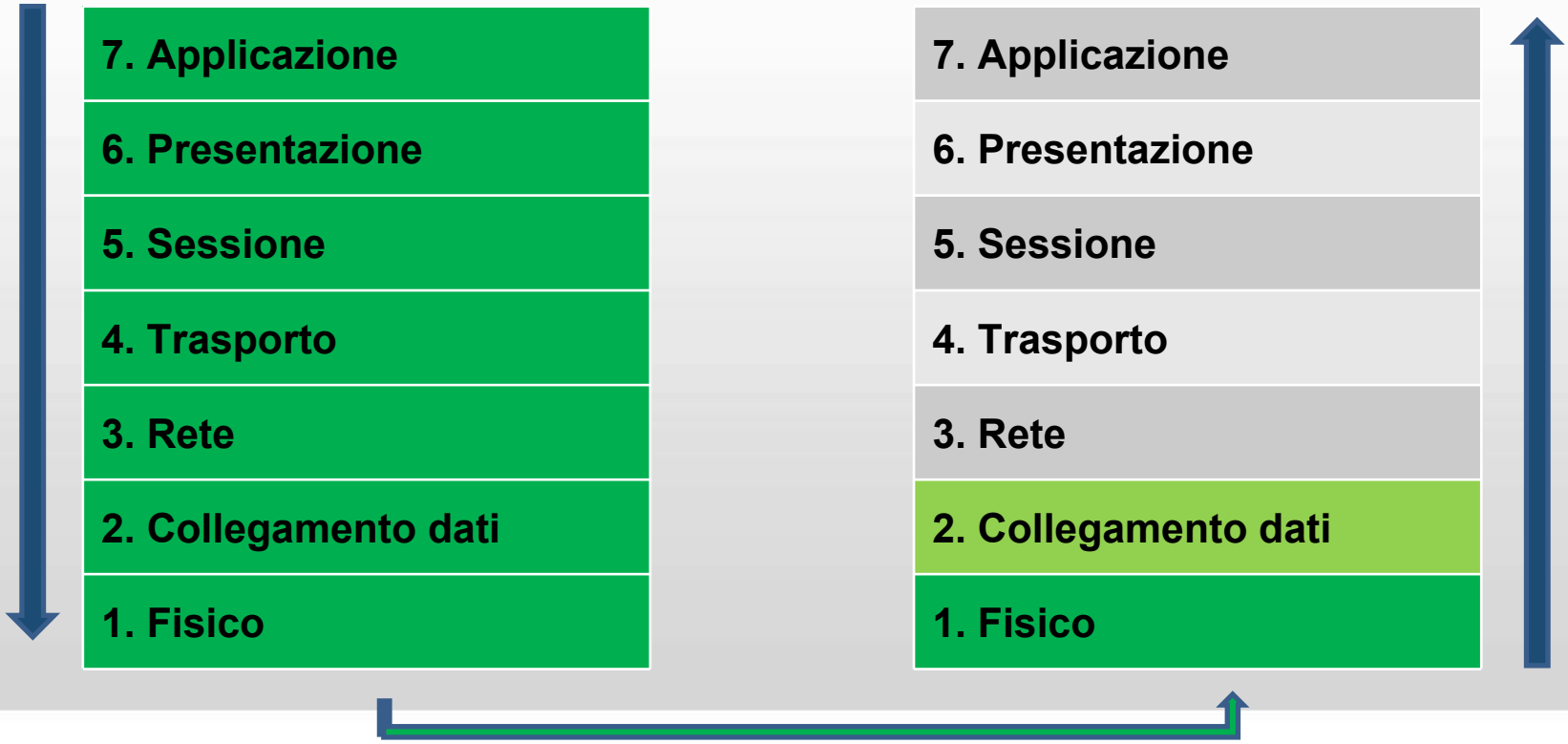
- **Fase di trasferimento dei dati attraverso il mezzo trasmissivo.**

Esempio

- **Il Livello Fisico ha ricevuto la sequenza di BIT e li trasferisce al livello Collegamento dati soprastante gli 0 e 1 ricevuti.**

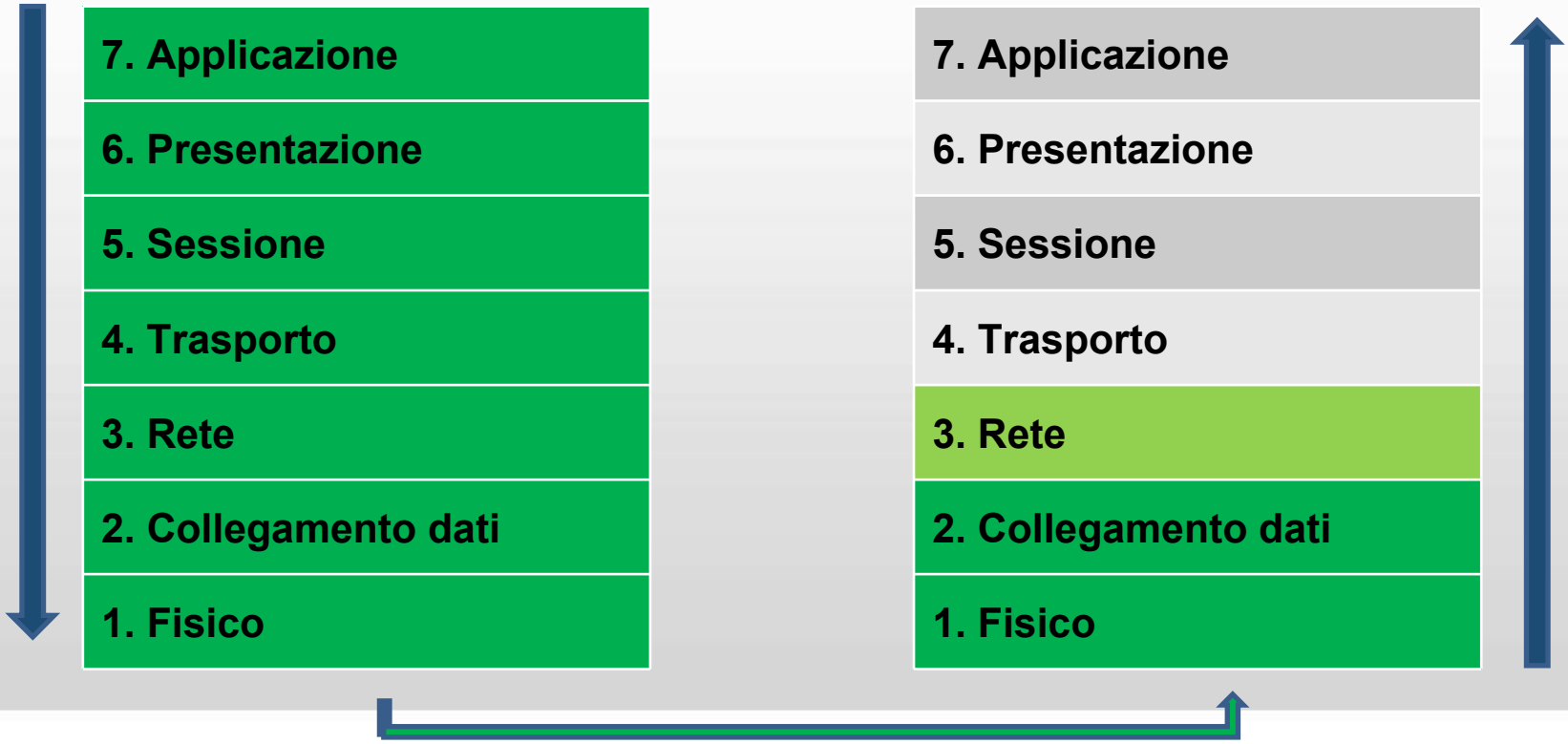
Esempio

- **Il livello Collegamento dati controlla gli eventuali errori e modifica, a seconda delle esigenze, le velocità di ricezione per rendere la trasmissione ottimale**



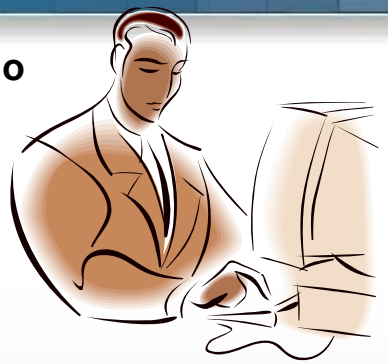
Esempio

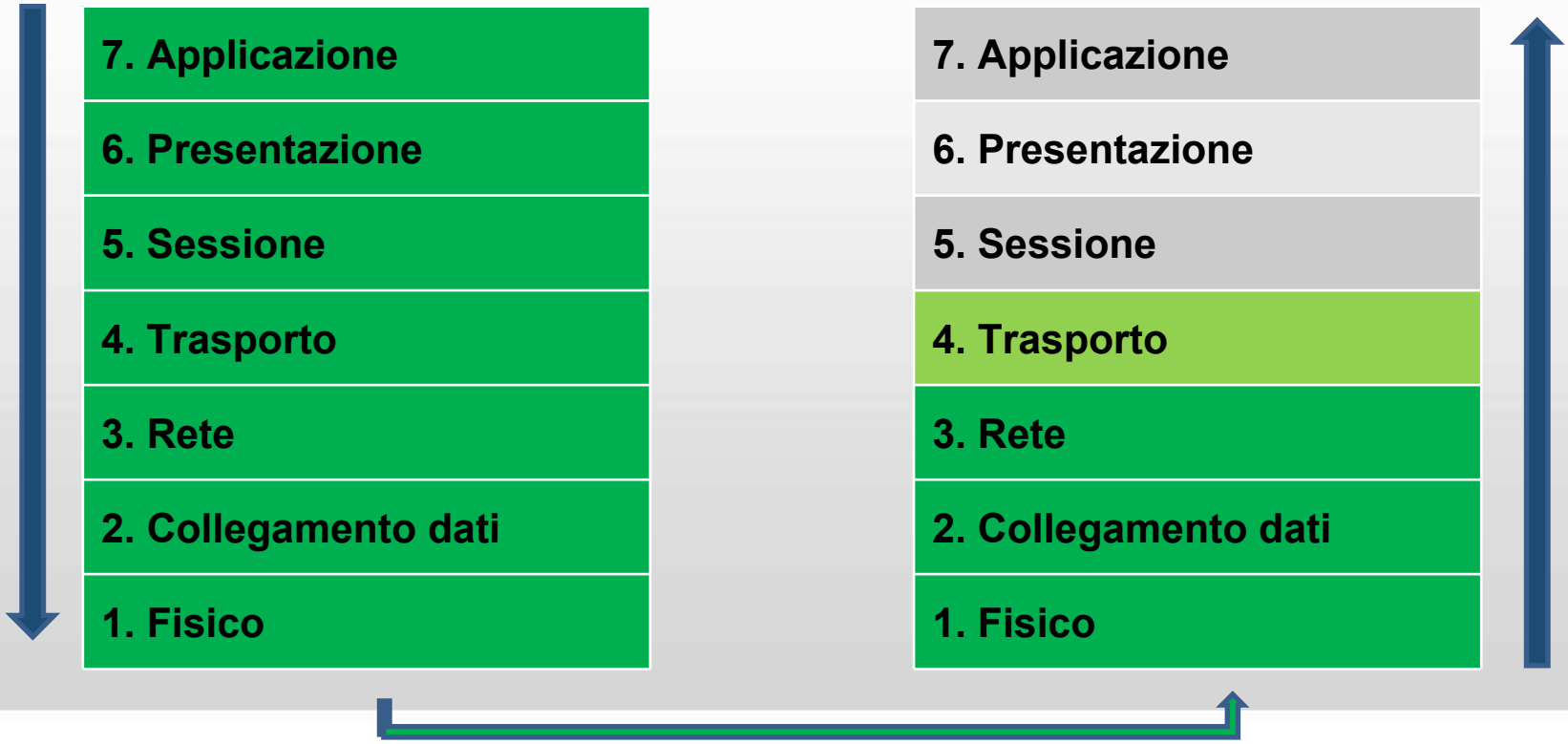
- **Il livello Rete comunica con il livello Rete dell'altro elaboratore e, in accordo con esso, è responsabile dei percorsi virtuali**



Esempio

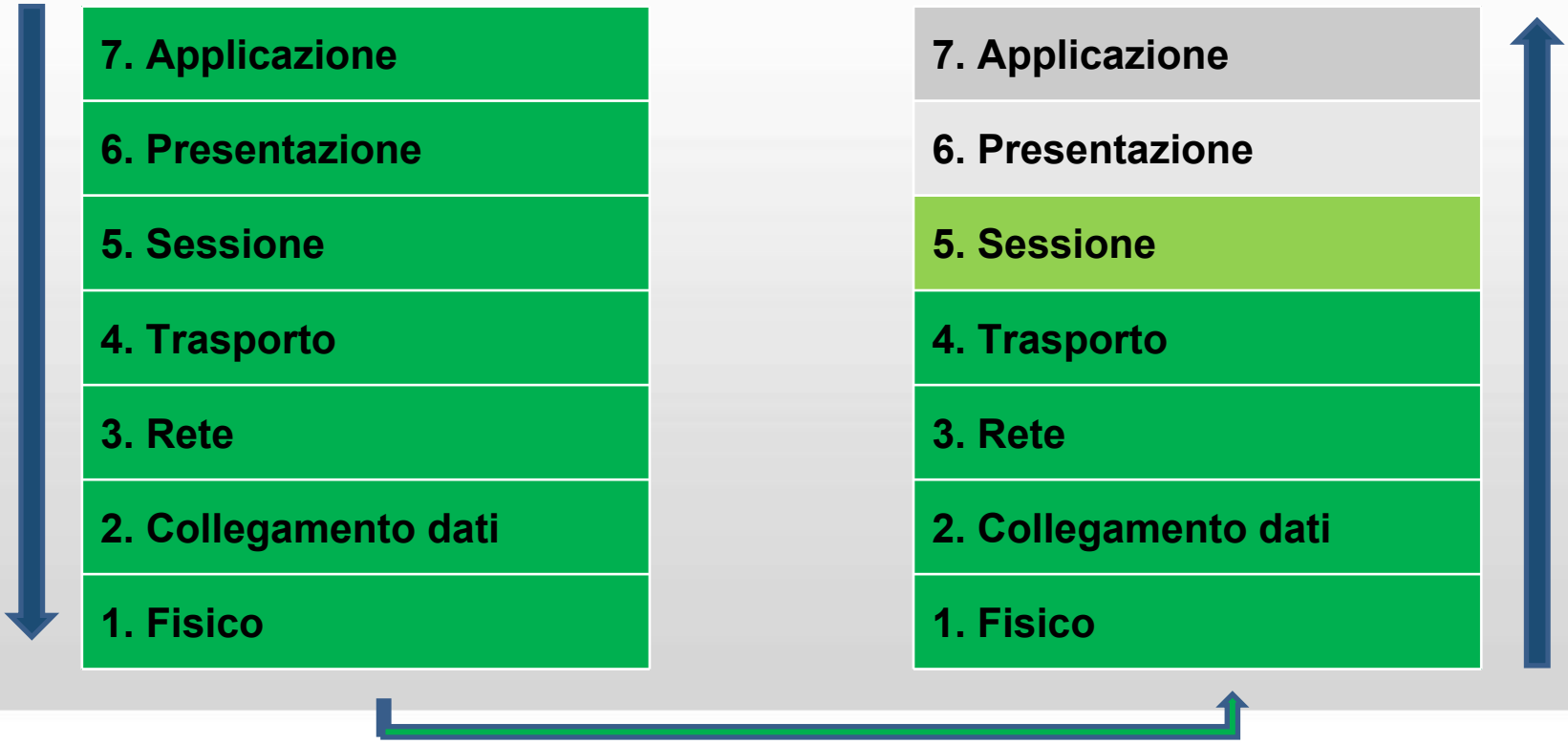
- **Il livello Trasporto segnala al livello Trasporto adiacente l'eventuale perdita o duplicazione di informazioni**





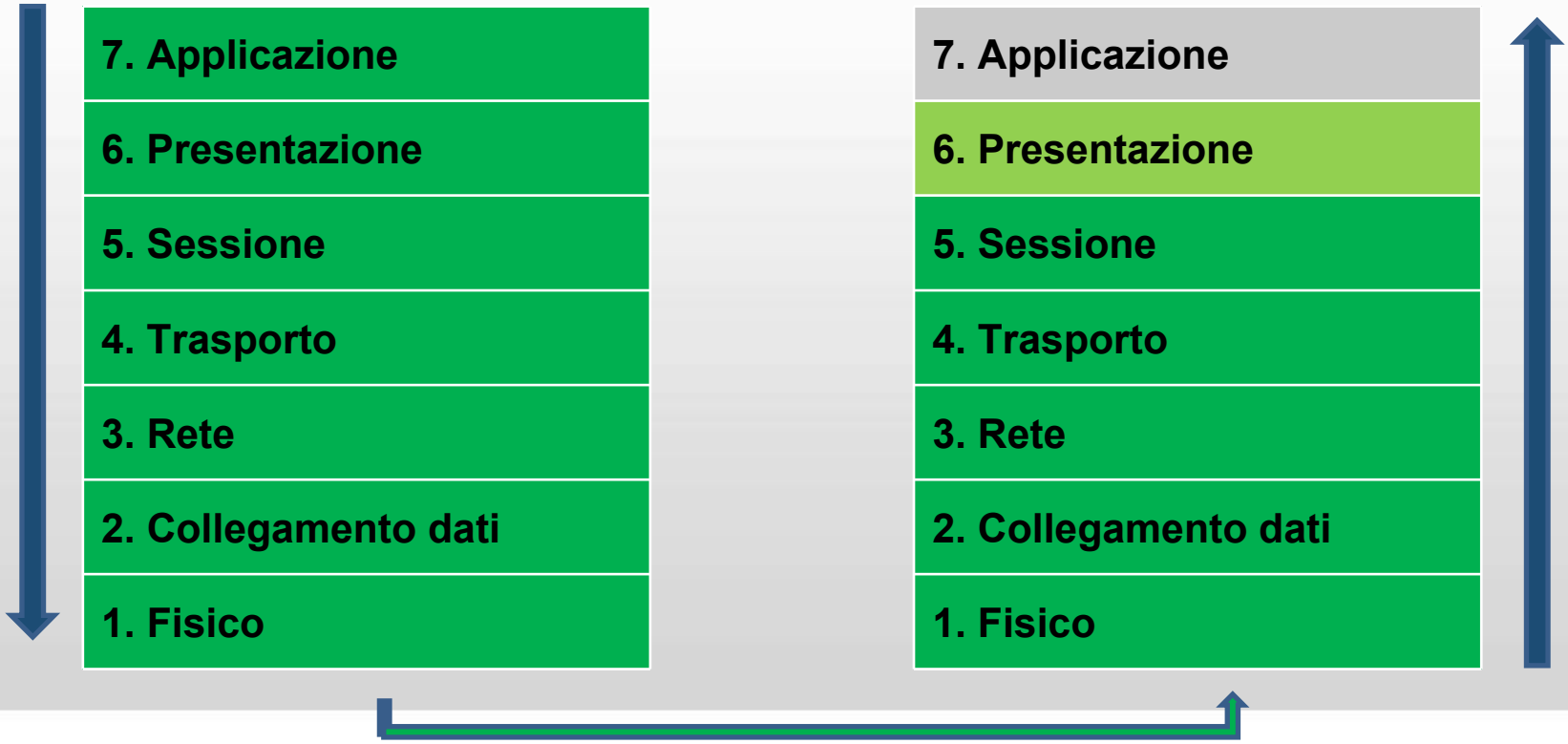
Esempio

- **Il livello Sessione adesso può chiudere la sessione virtuale nel caso sia finito lo scambio di messaggi o lasciarla aperta nel caso aspetti ancora altre informazioni.**



Esempio

- **Il livello Presentazione, in questa fase, decodifica i dati ottenuti rendendoli comprensibili alle varie applicazioni**



Esempio

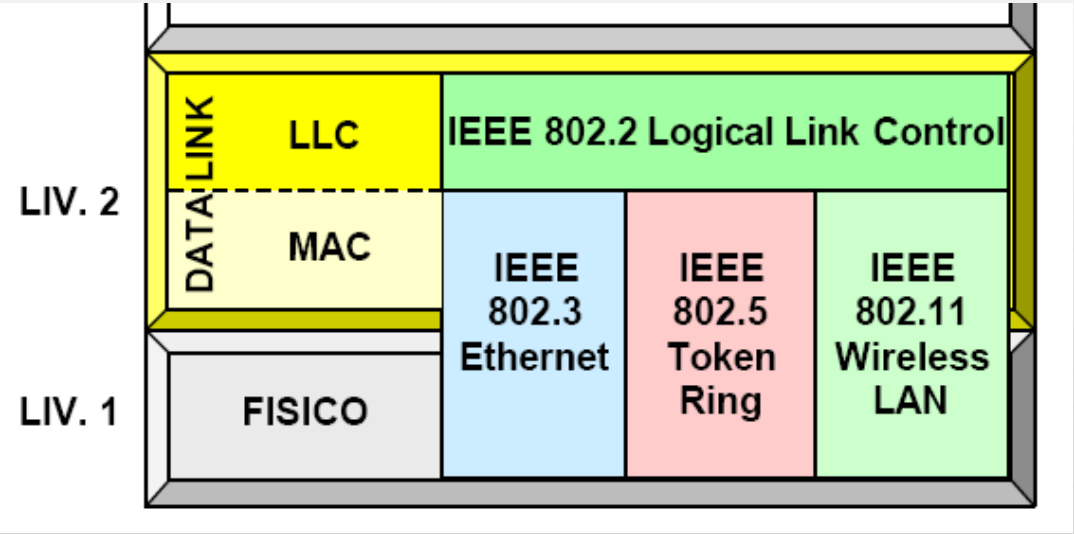
- **Il messaggio è giunto al livello Applicazione.**
- **L'utente può usufruirne**

I protocolli di Internet

Remo Romagnuolo - *"Internet"* - 2007

IEEE 802

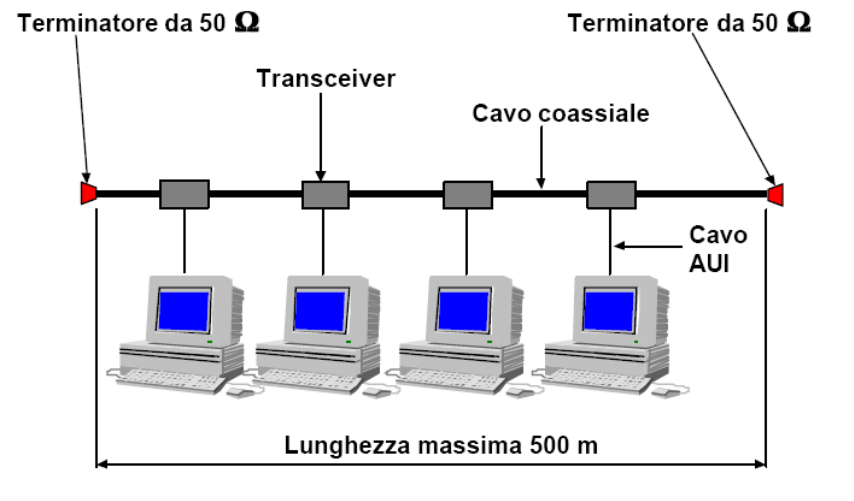
- Definizione di LAN
  - *La LAN è un sistema di comunicazione che permette ad apparecchiature indipendenti di comunicare tra di loro entro un'area delimitata utilizzando un canale fisico a velocità elevata e con basso tasso di errore*.
- Il progetto 802 lavora a livello fisico e dati della pila ISO



Ethernet 802.3

Il protocollo

- CSMA-CD
  - *Carrier Sense Multiple Access Collision Detection*
- quando una persona vuole parlare si mette in ascolto per verificare che non ci sia già qualcuno che sta parlando;
- se c'è silenzio inizia a parlare con uno o più destinatari
- tutti però ricevono il messaggio;
- la persona a cui è riferito il messaggio ne analizza il contenuto;
- le persone a cui non è riferito il messaggio non prendono in considerazione il contenuto.



Wifi 802.11

Elementi del successo

- Reti indoor
- Reti Outdoor
  - Hot spot
- Eliminazione del cablaggio
  - Riduzione dei costi associati alle infrastrutture di rete
  - Il mezzo trasmissivo non si guasta: si riducono i costi associati alla manutenzione
  - Possibilità di collegare ambienti con una logistica complessa non adatti al cablaggio (edifici storici etc...)
  - Viene facilitata l'implementazione di reti temporanee: fiere, eventi, convegni
- Elemento di equità sociale

Essere su Internet

Come si fa

- Serve un'**interfaccia** di rete
- Ogni interfaccia ha un indirizzo IP del tipo 157.138.204.250 (quattro byte separati da punti)
- Ogni interfaccia possiede 65535 "porte" di ingresso
- Ad un indirizzo IP si può associare un nome
  - www.iuav.it <-> 157.138.204.250

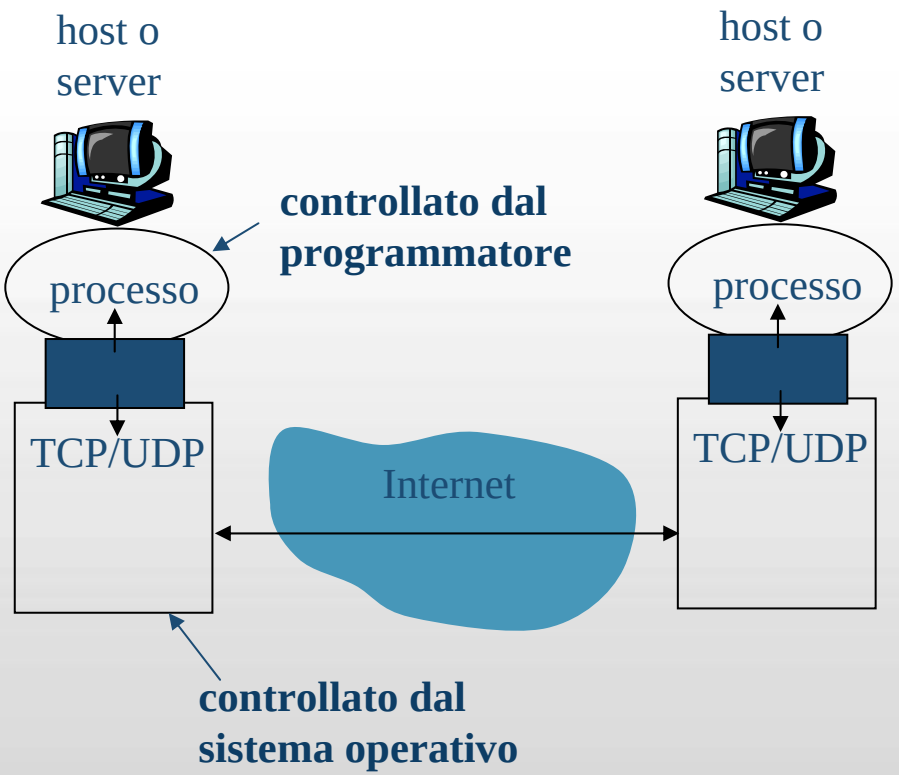
Cosa mettiamo su Internet

- eMail
- Web
- Instant messaging
- P2P file sharing
- Streaming stored video clips
- Internet telephone
- Real-time video conference
- Massive parallel computing

Essere su Internet

Cosa succede

- Processi su host diversi devono lanciarsi messaggi conformi al protocollo applicativo
- Indipendentemente dall'architettura
  - Il processo server attende che qualcuno lo chiami (ascolta su una porta del'linterfaccia. Ad es: 157.138.204.250:80)
  - Il processo client chiama il server: apre un socket sulla porta in ascolto del server
  - Il processo server risponde sul socket instanziato



Esigenze del protocollo applicativo

Applicazione	Affidabilità	Banda	Latenza
file transfer	no loss	elastic	no
e-mail	no loss	elastic	no
Web documents	no loss	elastic	no
real-time audio/video	loss-tolerant	audio: 5kbps-1Mbps video:10kbps-5Mbps	yes, 100’s msec
stored audio/video	loss-tolerant	same as above	yes, few secs
interactive games	loss-tolerant	few kbps up	yes, 100’s msec
instant messaging	no loss	elastic	yes and no

TCP e UDP

Transmission Control Protocol

- Orientato alla connessione
- Affidabile
- Controllo di flusso e congestione: i pacchetti arrivano in ordine e anche se produciamo tanti dati basta inviare
- Nessuna garanzia su latenza e banda

User Datagram Protocol

- Non affidabile
- Non garantisce: connessione, affidabilità, controllo di flusso e congestione, garanzie di latenza o banda, sequenza di arrivo

Alcuni protocolli standard

Application	Application layer protocol	Underlying transport protocol
e-mail	SMTP [RFC 2821]	TCP
remote terminal access	Telnet [RFC 854], SSH	TCP
Web	HTTP [RFC 2616]	TCP
file transfer	FTP [RFC 959]	TCP
streaming multimedia	RTP [RFC 1889] proprietary	TCP or UDP
Internet telephony	(es., Vonage,Dialpad Skype)	typically UDP

Come funziona la posta elettronica

I componenti

- Applicazione client: thunderbird, outlook, riga di comando
- Mail Server
  - contengono le caselle di posta
  - Gestisce la coda dei messaggi in partenza
  - Utilizza il protocollo SMTP
    - client: mail server che invia
    - server: mail server che riceve
- Protocollo d'invio: SMTP
- Protocollo di richiesta: POP3 - IMAP

Simple Mail Transport Protocol

- Usa TCP sulla porta **25**
- Trasferimento diretto da mittente a destinatario
- Tre fasi nel trasferimento
  - handshaking (saluti)
  - trasferimento dei messaggi
  - saluti
- comandi: testo ASCII leggibile!
- risposta: un codice di ritorno e una frase

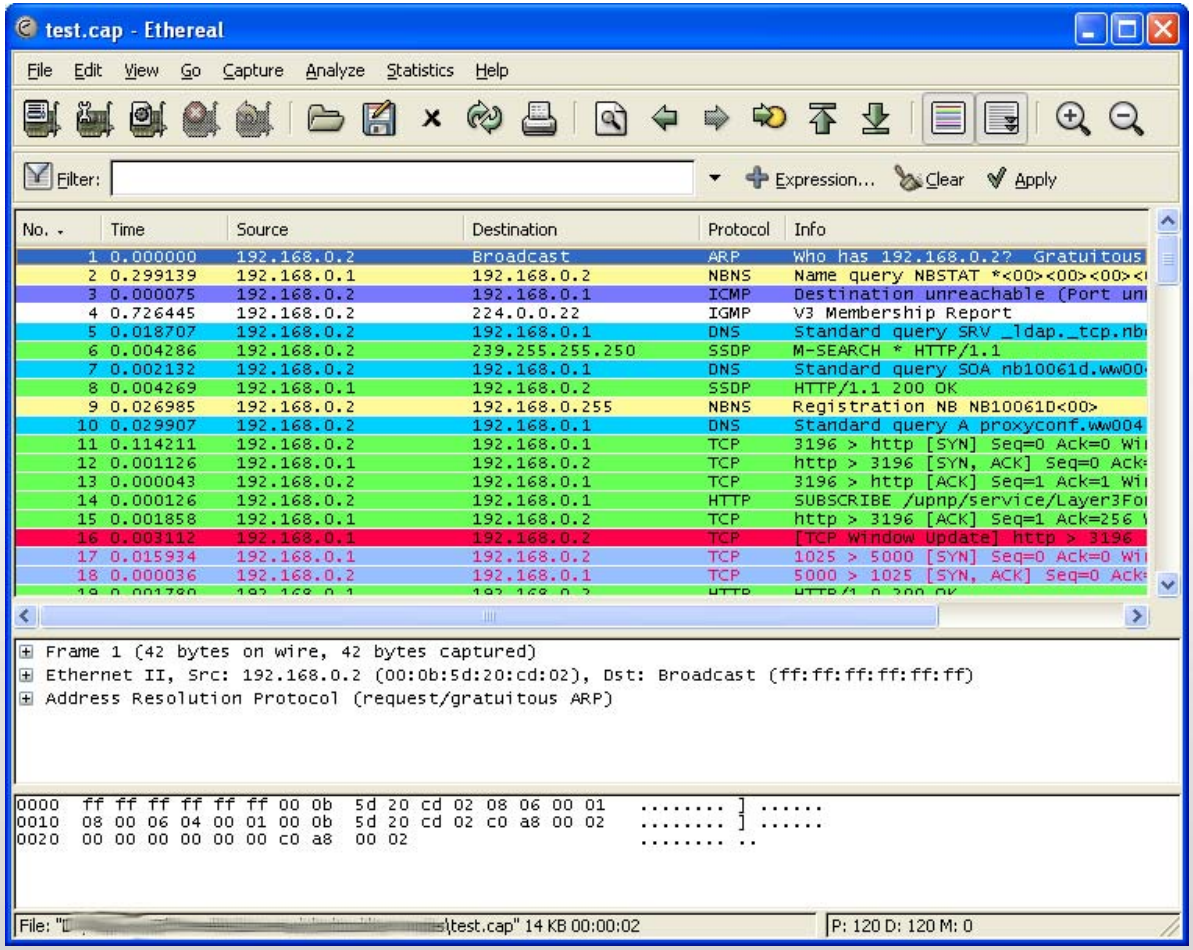
```
telnet server 25

S: 220 hamburger.edu
C: HELO crepes.fr
S: 250  Hello crepes.fr, pleased to meet you
C: MAIL FROM: <alice@crepes.fr>
S: 250 alice@crepes.fr... Sender ok
C: RCPT TO: <bob@hamburger.edu>
S: 250 bob@hamburger.edu ... Recipient ok
C: DATA
S: 354 Enter mail, end with "." on a line by itself
C: Do you like ketchup?
C: How about pickles?
C: .
S: 250 Message accepted for delivery
C: QUIT
S: 221 hamburger.edu closing connection
```

Sicurezza, Autenticazione, Credibilità

Ethereal

http://www.ethereal.com/



POP / IMAP

Post Office Protocol

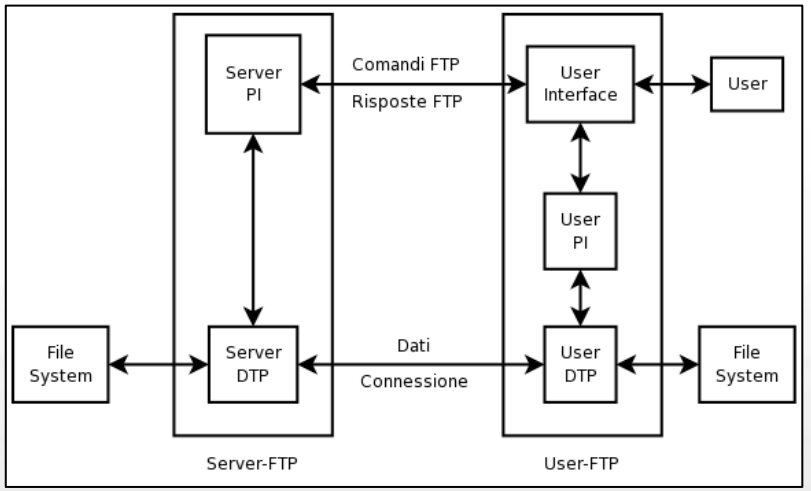
- Usa TCP sulla porta **110**
- Fasi della richiesta
  - Autorizzazione (**username, password**)
  - Recupero (**list, retr**)
  - Eliminazione (**dele**)
  - Saluti (**quit**)

Internet Message Access Protocol

- Usa TCP sulla porta **143**
- Alcuni vantaggi rispetto al POP3
  - I messaggi restano sul server
  - Si possono creare cartelle sul server
  - Accesso alla posta sia online che off-line
  - Più utenti possono utilizzare la stessa casella di posta
  - Possibilità di fare ricerche sul server

File Transfer Protocol

- Trasferisce file da e per un host remoto
- Usa TCP sulla porta **20** e **21**
- La negoziazione avviene sulla 20
  - Si navigano le directory sulla connessione 20
- Una connessione dati separata (21) viene aperta per trasferire i file
- Dopo aver trasferito i file il server chiude la connessione dati
- Es: ftp pds-geosciences.wustl.edu



Sicurezza

- FTP non prevede cifratura per i dati scambiati tra client e server
- Username, password, comandi, codici di risposta e file trasferiti possono essere "sniffati" o visionati da malintenzionati
- Il problema è comune anche a HTTP, TELNET e SMTP
- FTPS che aggiunge al protocollo FTP originale un layer di cifratura SSL/TLS

HTTP

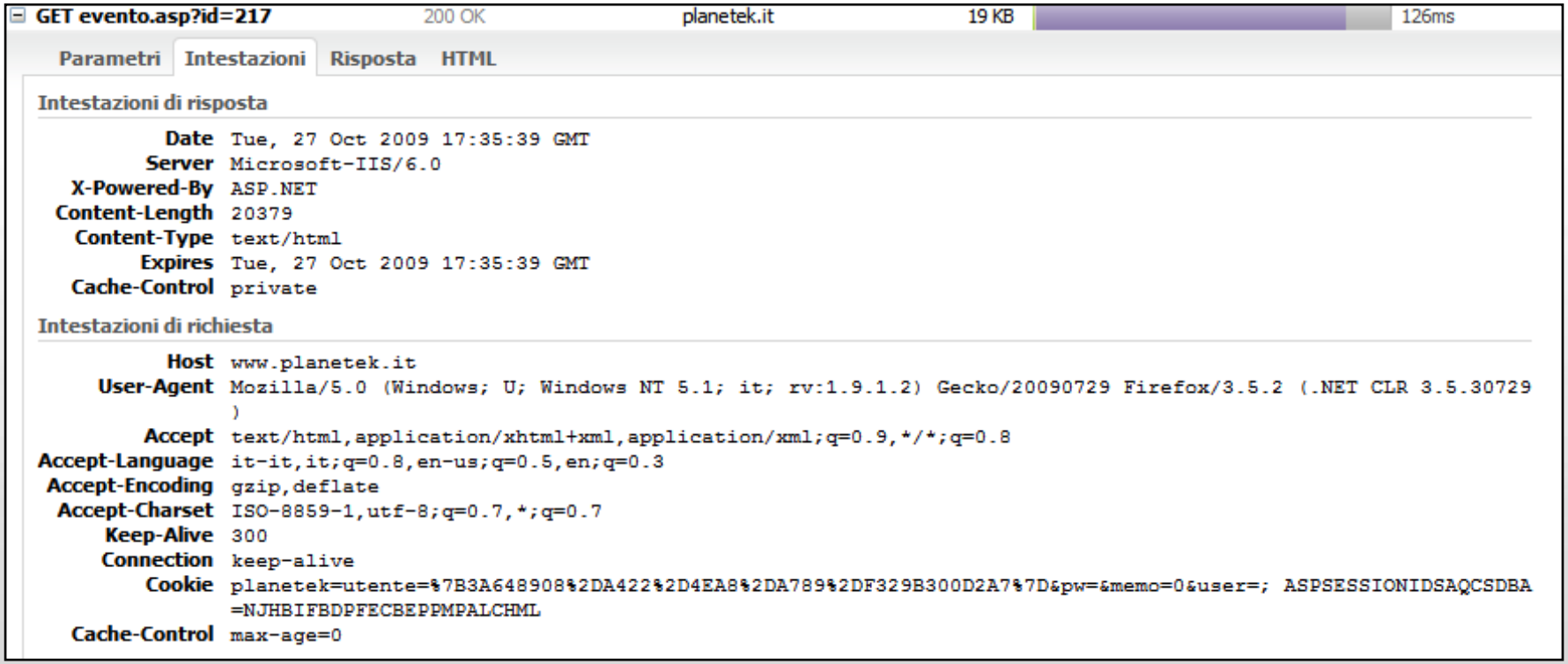
Hypertext Transfer Protocol

- Trasferisce oggetti web da un host remoto
- Usa TCP sulla porta **80**
- Comunicazione client-server
  - Il client chiama il server e chiede un oggetto web (HTTP/1.0 non persistente)
  - Il server lo invia
  - La connessione viene chiusa
- HTTP è stateless!
  - Necessità dei cookie
- HTTP/1.1
  - Più siti www sullo stesso server
  - Riuso delle connessioni disponibili
  - Ulteriori meccanismi di sicurezza

I messaggi HTTP

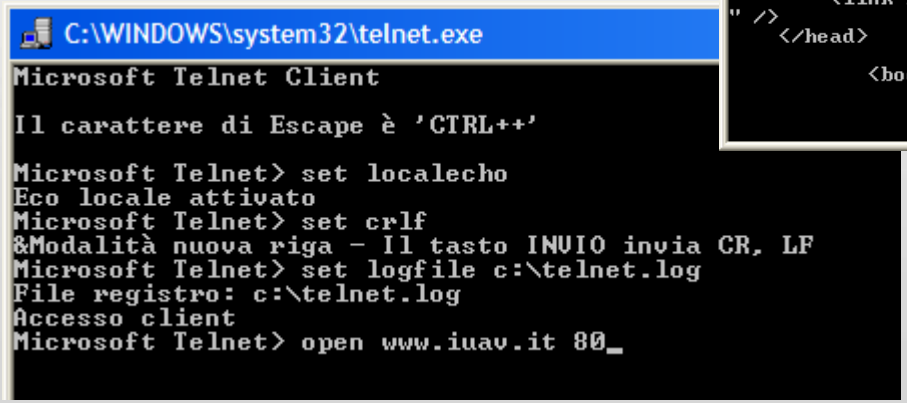
- ASCII leggibile
- Request o Response
- Ogni comando è fatto da
  - Metodo + URI + versione del protocollo
  - GET /homepage/index.htm HTTP/1.1
- Comandi
  - GET, POST, HEAD, PUT, DELETE, TRACE, OPTIONS
- Codici di ritorno
  - 1xx: Informational
  - 2xx: Success
  - 3xx: Redirection
  - 4xx: Client error
  - 5xx: Server error

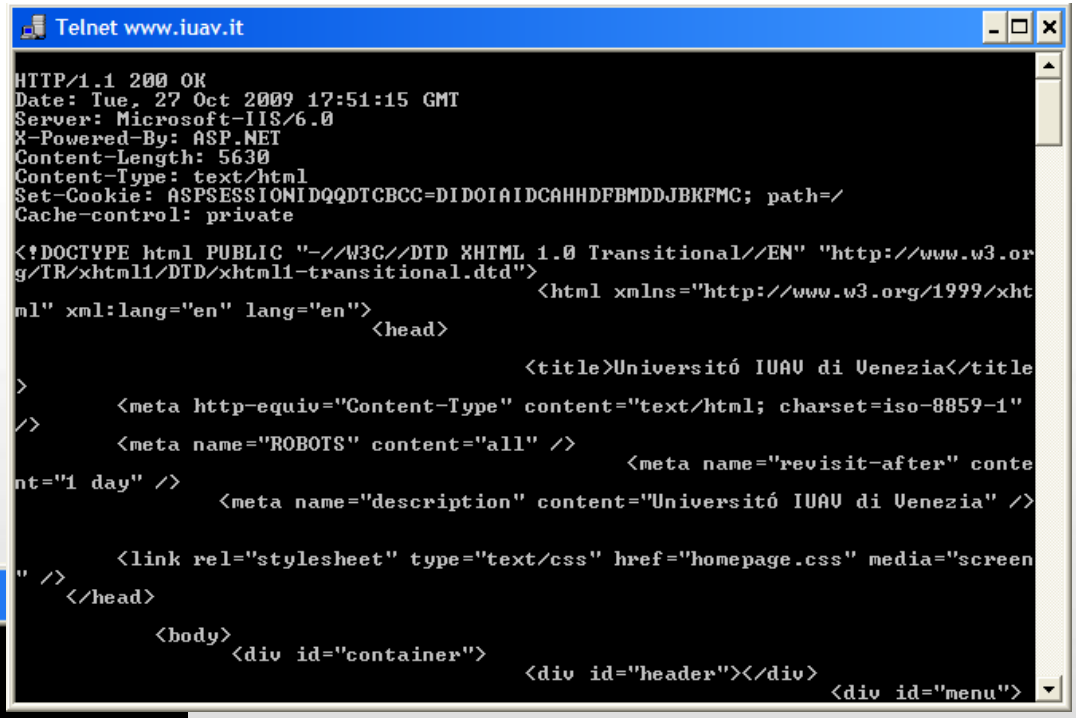
Debug di Firebug



Cose mai viste!

- Esegui…. telnet
  - set localecho
  - set crlf
  - set logfile c:\telnet.log
  - open www.iuav.it 80
- Chiediamo l'homepage
  - GET /homepage/
- E voilà!





GET condizionale e PROXY

La rete è una risorsa pregiata

- Scopo: non mandare l'oggetto in giro se non necessario

GET condizionale

- Il client quando fa la richiesta indica la data della propria copia
  - `If-modified-since: <date>`
- Il server risponde con pochi byte se la copia del client è aggiornata:
  - `HTTP/1.0 304 Not Modified`

Proxy (cache server)

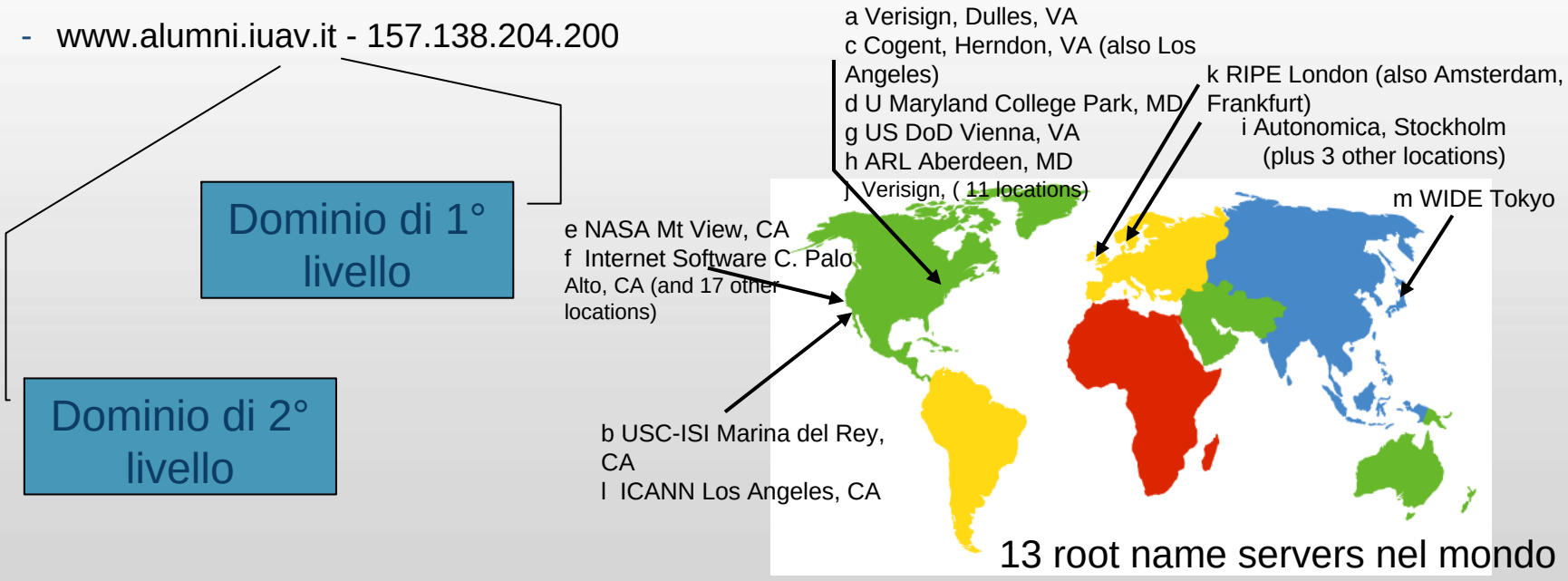
- Il client fa tutte le richieste al proxy
  - Se l'oggetto è in cache viene ritornato
  - Altrimenti il proxy si occupa di richiedere l'oggetto e di ritornarlo al client

Domain Name Service

- Trasformare nomi mnemonici in indirizzi IP
  - www.iuav.it - 157.138.250.253

Database distribuito

- Naming gerarchico
  - www.alumni.iuav.it - 157.138.204.200



Caratteristiche

- Richieste
  - Iterative
  - Ricorsive
- Permanenza in cache delle associazioni
- DNS dinamico
  - DHCP e query di UPDATE
- Manipolazione delle risposte
  - **protezione da abusi**: il server DNS può filtrare le query relative a siti pericolosi per gli utenti
  - **Censura**
  - **man-in-the-middle**: il traffico viene reindirizzato verso un server che agisce da proxy trasparente, intercettando il traffico degli utenti.
  - **redirezione degli errori**: alle query per nomi inesistenti viene risposto con l'indirizzo IP di un server, che tipicamente ospita un motore di ricerca e tenta di aiutare gli utenti a trovare il sito cercato.
- Strumenti utili
  - NSLOOKUP

Uno per tutti, tutti per uno

Processi ad alte prestazioni

Sempre più potenti

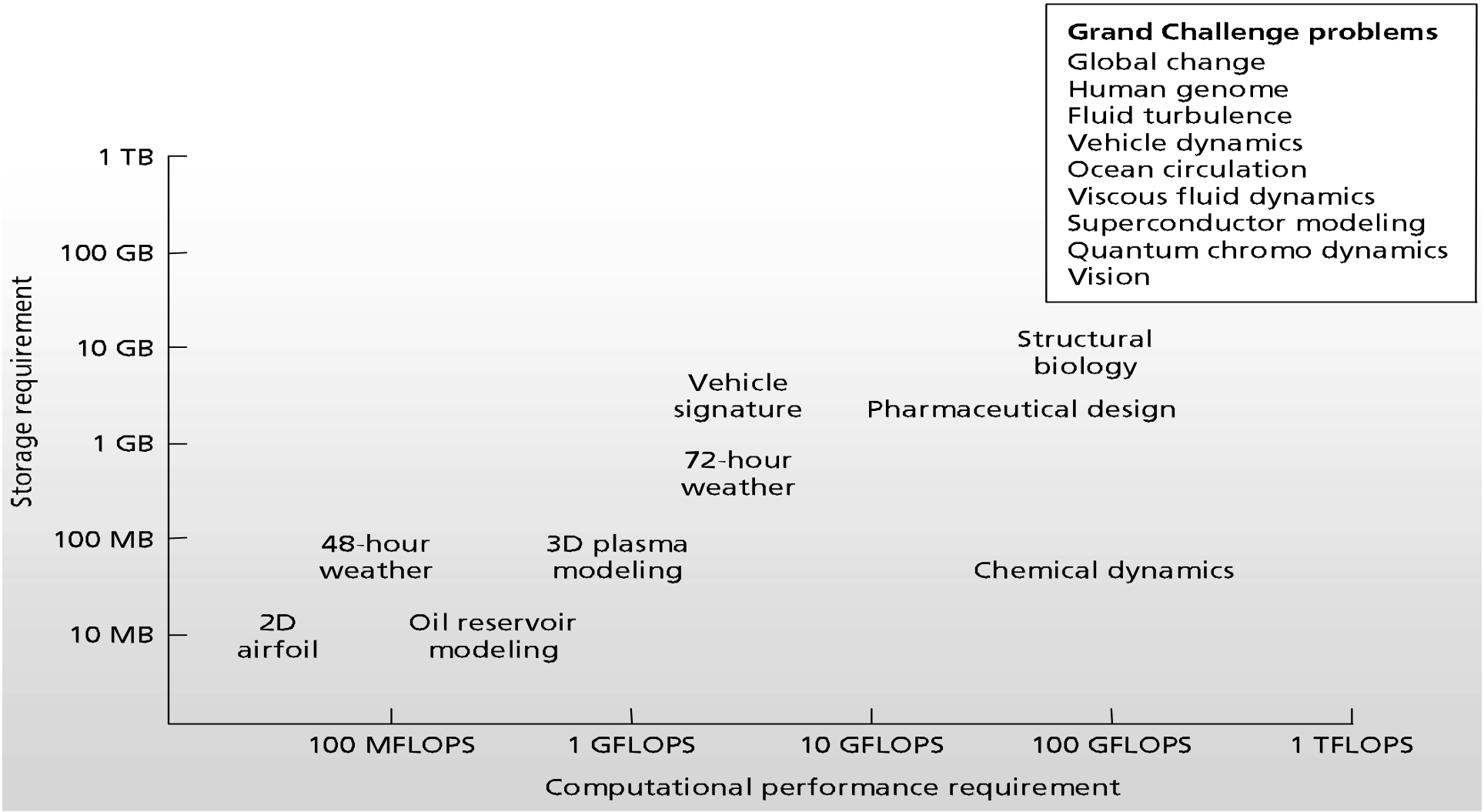
Nuove sfide

- Più piccoli, più potenti, più interconnessi e meno costosi.
- Capacità di risolvere problemi più complicati

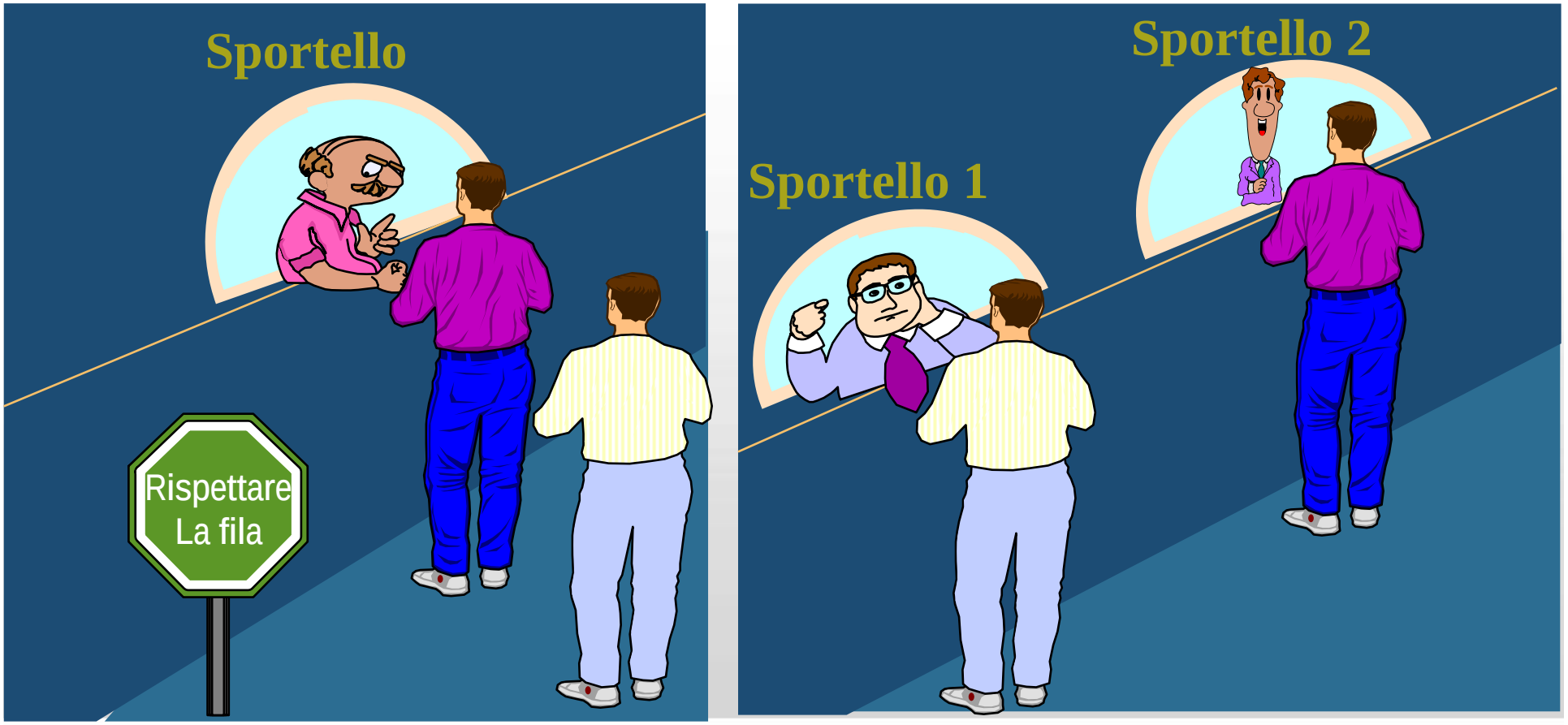
L'evoluzione

- 20 anni fa
  - 10^6 Floating Point Ops/sec (Mflop/s) - processori scalari
- 10 anni fa
  - 10^9 Floating Point Ops/sec (Gflop/s) - processori vettoriali, memoria condivisa
- Oggi
  - 10^{12} Floating Point Ops/sec (Tflop/s) - parallelismo massivo, calcolo distribuito
- Domani
  - 10^{15} Floating Point Ops/sec (Pflop/s)

Di cosa abbiamo bisogno



Seriale vs Parallelo



La tassonomia di Flynn

Lo schema di Flynn

- Flynn propone di raggruppare le architetture a seconda del numero di istruzioni e di dati che si possono maneggiare simultaneamente
  - SISD (Single Instruction and Single Data)

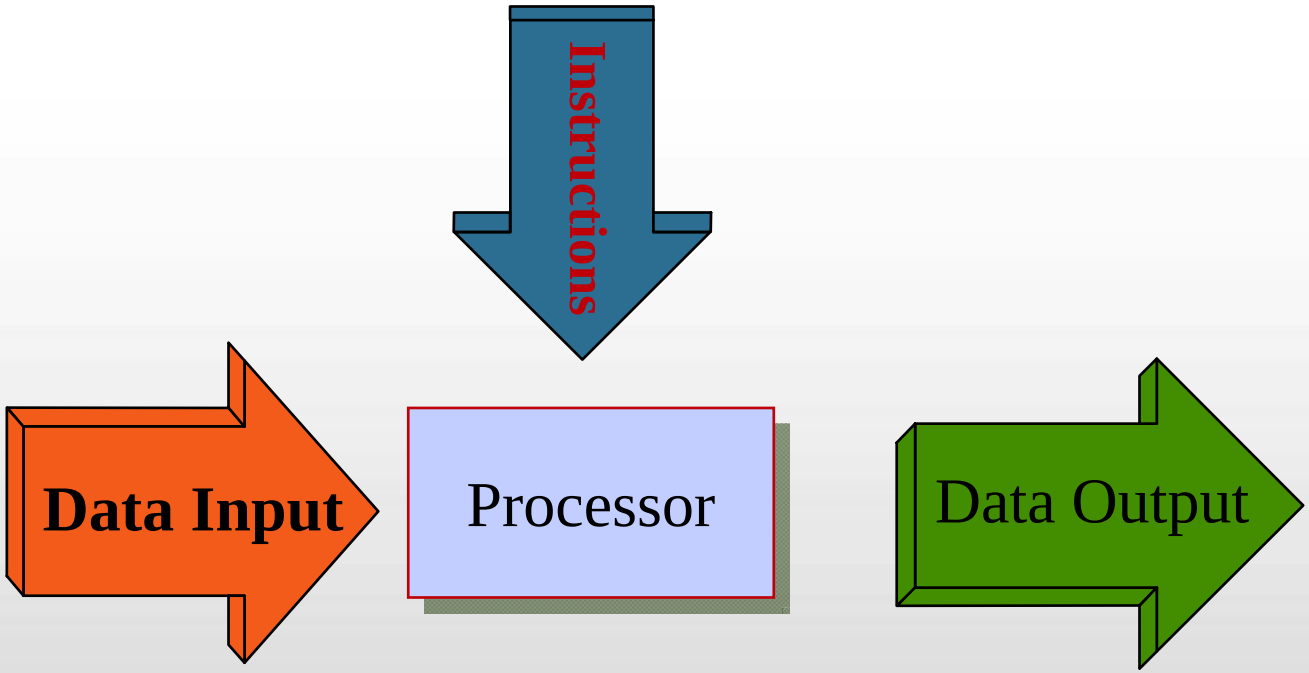
    Conventional computers
  - SIMD (Single Instruction and Multiple Data)

    Data parallel, vector computing machines
  - MISD (Multiple Instruction and Single Data)

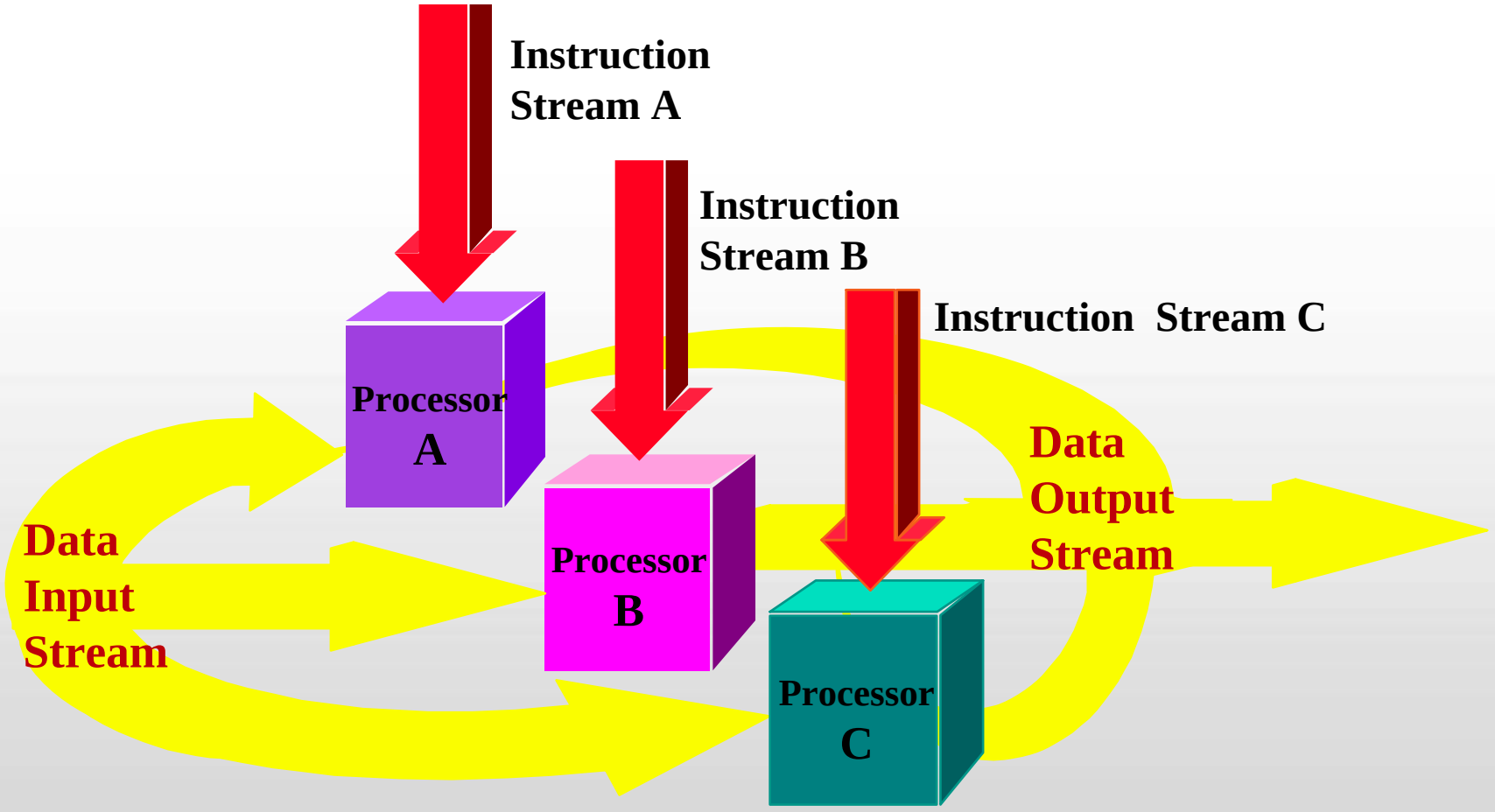
    Systolic arrays
  - MIMD (Multiple Instruction and Multiple Data)

    General purpose machine

Architettura SISD

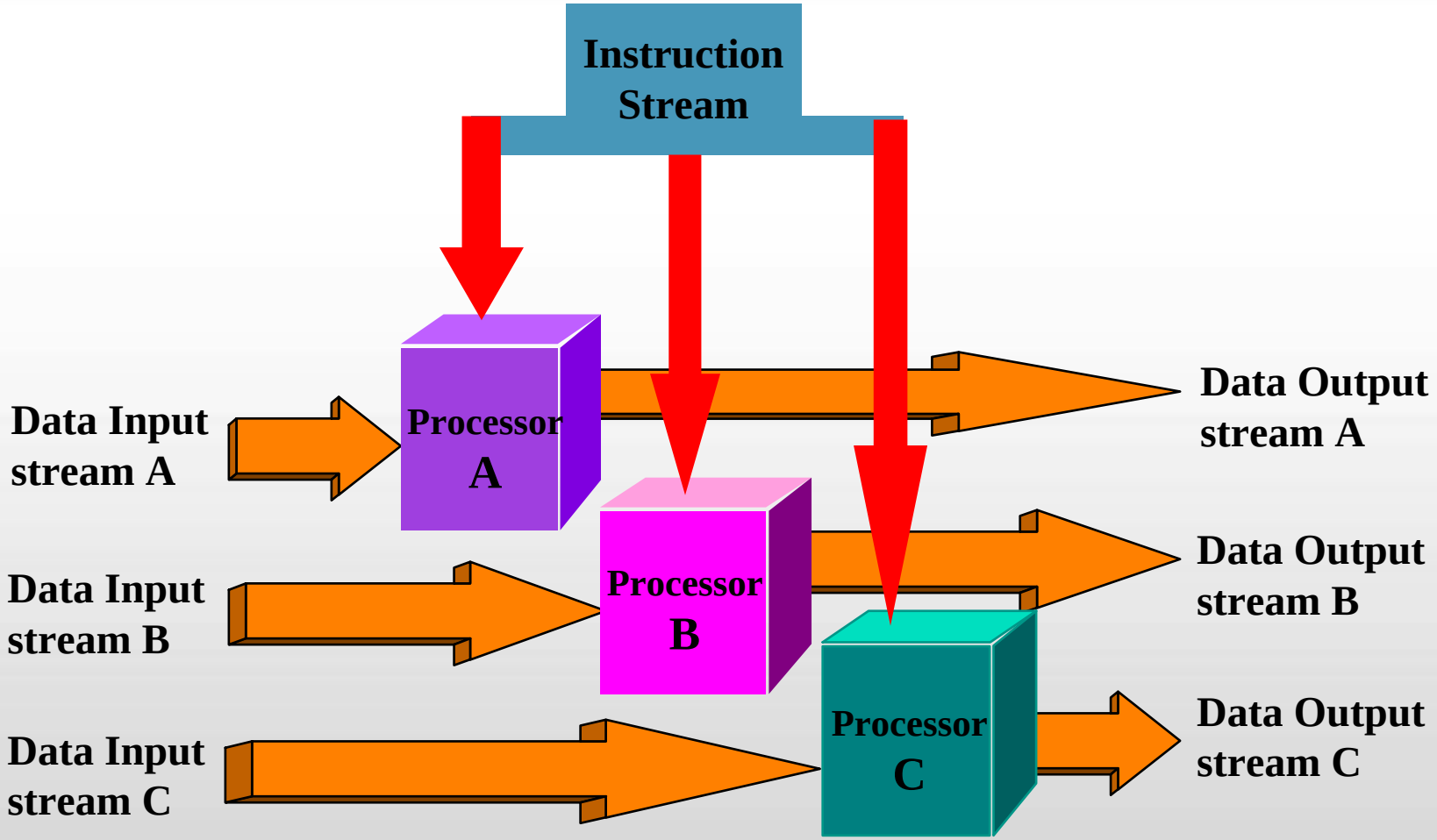


Architettura MISD



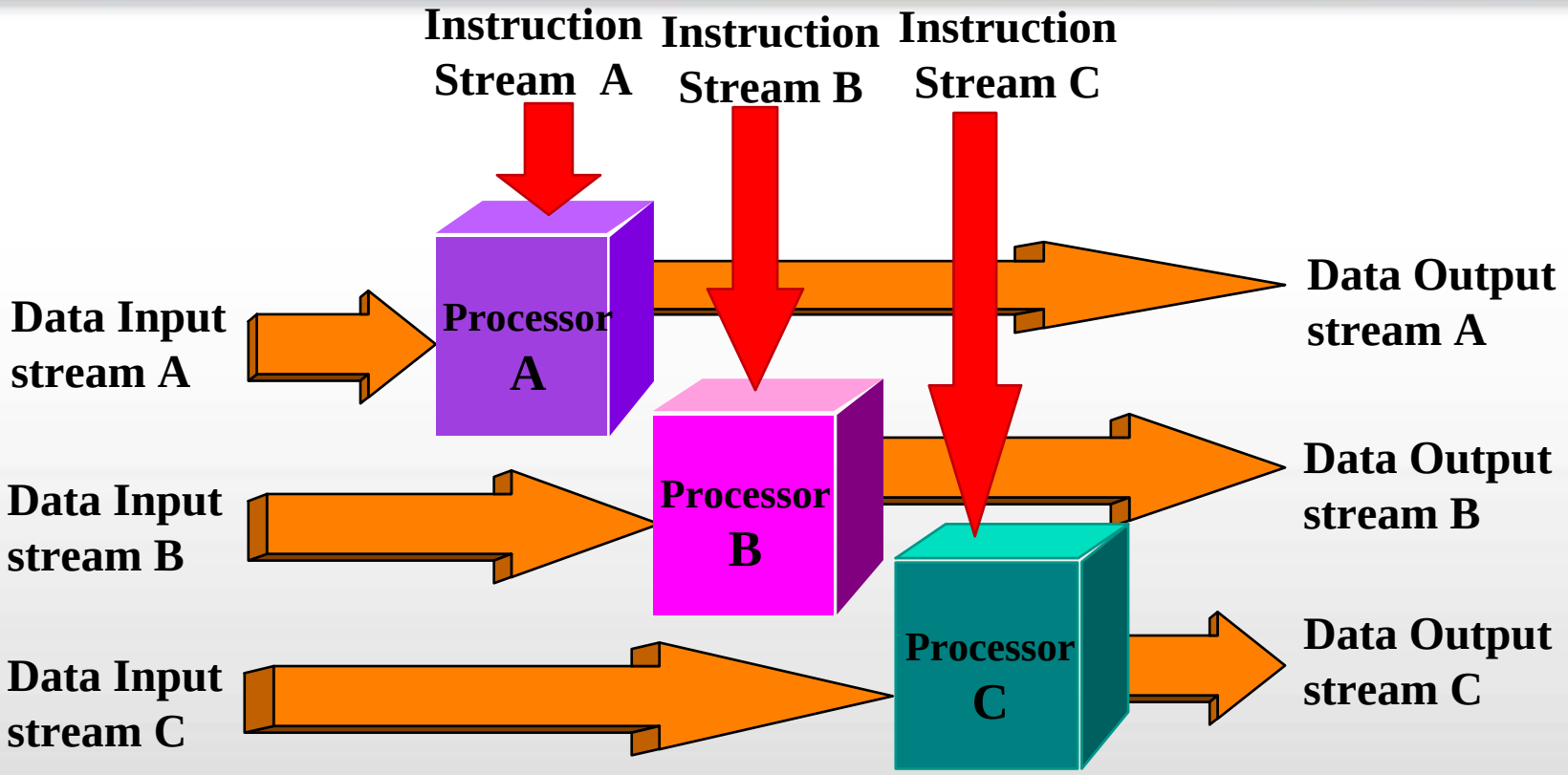
Non esistono!

Architettura SIMD



Intel MMX (multimedia)

MIMD Architecture



Computer lavorano in maniera asincrona.

- Memoria condivisa (fortemente accoppiati)
- Memoria distribuita (accoppiamento lasco)

Memoria condivisa

Vantaggi

- Al programmatore è quasi trasparente
- I sistemi operativi possono essere portati facilmente

Svantaggi

- Poco scalabile
  - più processori ci sono più traffico e richieste della memoria ci sono
- Poco affidabile
  - Se c'è un problema sulla memoria o su un processore tutto il sistema ne risente

Memoria distribuita

Vantaggi

- Facilità ed immediatezza d'estensione
- Alta affidabilità: qualsiasi problema non si ripercuote sul sistema
- La memoria aumenta col numero dei processori
- Ogni processore
  - parla con la propria memoria
  - non necessita di interfacce
  - Non deve mantenere coerenti memorie intermedie
- Basso costo: si può usare HW standard

Svantaggi

- Va gestita tutta la comunicazione
- Non è banale portare le strutture dei dati verso un'organizzazione così diversa di memoria

Possibili architetture

Tre modi diversi

- **Calcolo Parallelo**: un singolo sistema con molti processori che lavorano sullo stesso problema
- **Calcolo Distribuito**: molti sistemi gestiti da uno scheduler per lavorare su problemi tra loro connessi
- **Grid Computing**: molti sistemi strettamente accoppiati, anche distribuiti geograficamente, per lavorare insieme su un singolo problema o su problemi tra loro connessi

Utilizzo della memoria

- **Shared memory:** unico spazio di indirizzamento a cui hanno accesso tutti i processori
- **Distributed memory:** ogni processore ha il suo spazio di memoria locale Si deve effettuare il “message passing”

Più veloce della luce…

1 + 1 ≠ 2

- In media la potenza di calcolo è proporzionale alla radice quadrata del prezzo
- L'accelerazione aumenta col $\log_2$ del numero dei processori

Costo	Potenza
10	3,2
20	4,5
30	5,5
40	6,3
50	7,1
60	7,7
70	8,4
80	8,9
90	9,5
100	10,0

Accelerazione(pr)				
2	3	4	5	6
3,2	5,0	6,3	7,3	8,2
4,5	7,1	8,9	10,4	11,6
5,5	8,7	11,0	12,7	14,2
6,3	10,0	12,6	14,7	16,3
7,1	11,2	14,1	16,4	18,3
7,7	12,3	15,5	18,0	20,0
8,4	13,3	16,7	19,4	21,6
8,9	14,2	17,9	20,8	23,1
9,5	15,0	19,0	22,0	24,5
10,0	15,8	20,0	23,2	25,8

Costo(pr)				
2	3	4	5	6
20	30	40	50	60
40	60	80	100	120
60	90	120	150	180
80	120	160	200	240
100	150	200	250	300
120	180	240	300	360
140	210	280	350	420
160	240	320	400	480
180	270	360	450	540
200	300	400	500	600

Legge di Amdahl

1 + 1 ≠ 2

- L’accelerazione è funzione della parte di codice parallelizzabile
  - Speedup = 1 / (1 – P)
- Se N sono i processori
  - Speedup = 1/(1-P + P/N).

Nuove piattaforme

- **GPU** (Graphics Processing Unit): è il microprocessore per le schede video per il rendering di immagini 3D.
- **GpGPU** (General Purpose computation using GPU): utilizzo dei processori grafici per accelerare applicazioni fortemente parallelizzabili.
- GPU Tesla: processore Nvidia con un core a 128 (240) processori.

Il futuro quantistico

Informatica quantistica

- D.Deutch (1984), un fisico britannico, riformula l'informatica sulla base della fisica quantistica.
- In un computer quantistico le varianti si risolvono tutte contemporaneamente, in una "sovrapposizione" di stati (o universi) paralleli.

Come funziona

- Ad ogni bit corrisponde una particella subatomica, dotata di spin
- La particella potrebbe ruotare in un verso o nell'altro
- Due diverse rotazioni = 0 e 1

Est	Est	Ovest	Ovest	Est	Ovest	Ovest
1	1	0	0	1	0	0

Il futuro quantistico

Come funziona

- A livello subatomico le particelle obbediscono alle leggi della meccanica quantistica e possono essere in una sovrapposizione di stati.
  - Ogni particella può essere simultaneamente in universi alternativi: in uno ruota verso ovest e nell'altro verso est
  - Eccitando le particelle esse possono trovarsi in 2^7 (128) stati diversi

I Qbit

- I Qbit possono eseguire 128 calcoli nel tempo necessario ad uno solo
  - un'entità singola che esegue simultaneamente i calcoli con 128 numeri diversi;
  - 128 entità distinte che eseguono, in universi paralleli, un solo calcolo, diverso da tutti gli altri
- Nel 2001 IBM realizza il primo computer quantistico a 7-qbit.
  - Viene fattorizzato il numero 15 utilizzando 1018 molecole identiche ognuna contenente 7 atomi.


Da qui non si passa
Sicurezza e Crittografia

La sicurezza

Definizione

- Un sistema informatico è sicuro se le informazioni in esso contenute sono garantite da misure di sicurezza.

I servizi di sicurezza

- **Disponibilità**: reperibilità dei dati e continuità dei processi
- **Riservatezza**: l'informazione è cifrata
- **Integrità**: Il destinatario può verificare che il contenuto della transazione non sia stato alterato
- **Autenticazione**: Il destinatario è certo dell'identità del mittente
- **Non Ripudio**: Il mittente non può negare di aver eseguito la transazione

La sicurezza

Metodi di intrusione

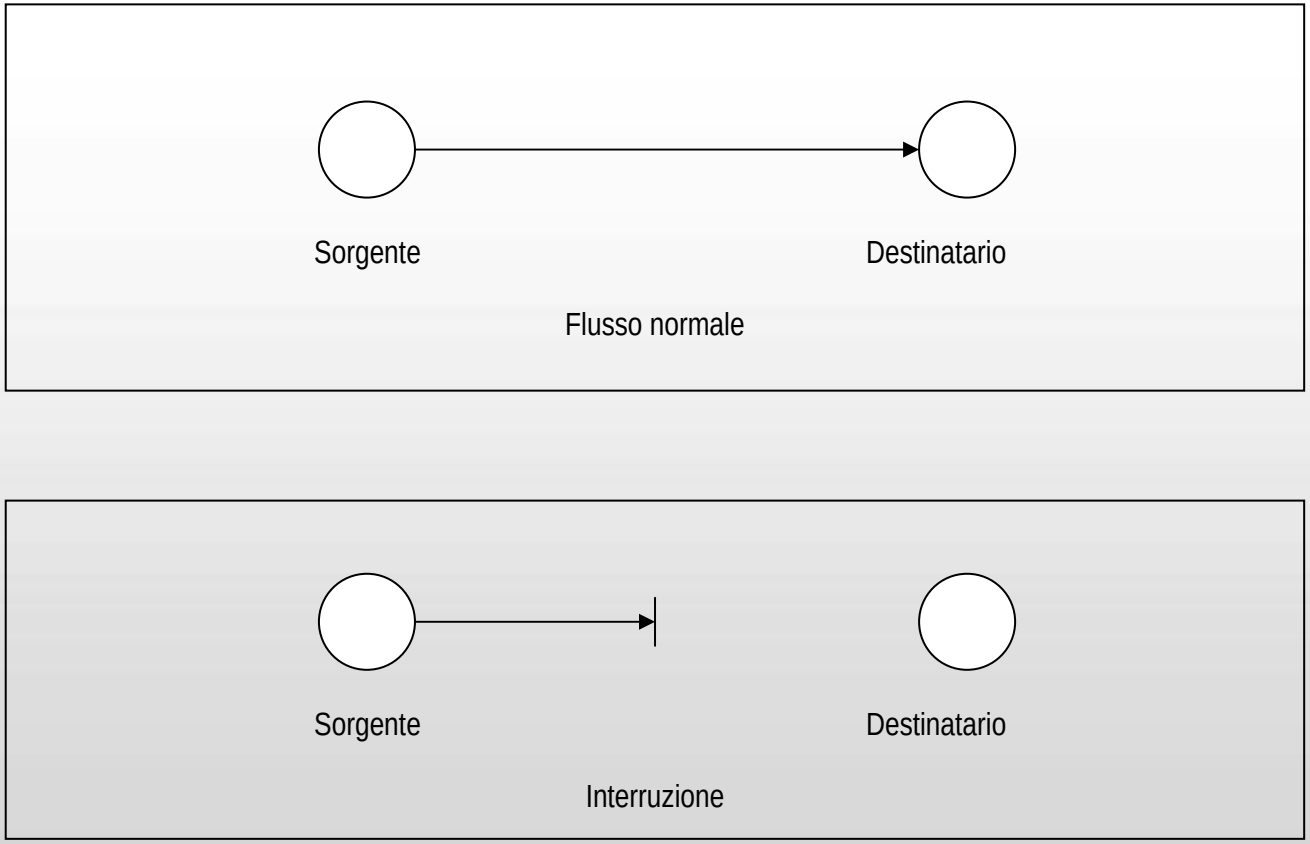
- Intrusione via macro
- Intrusione via e-mail
- Introduzione dall'interno
- Intercettazione dei dati trasmessi

Strategie di difesa

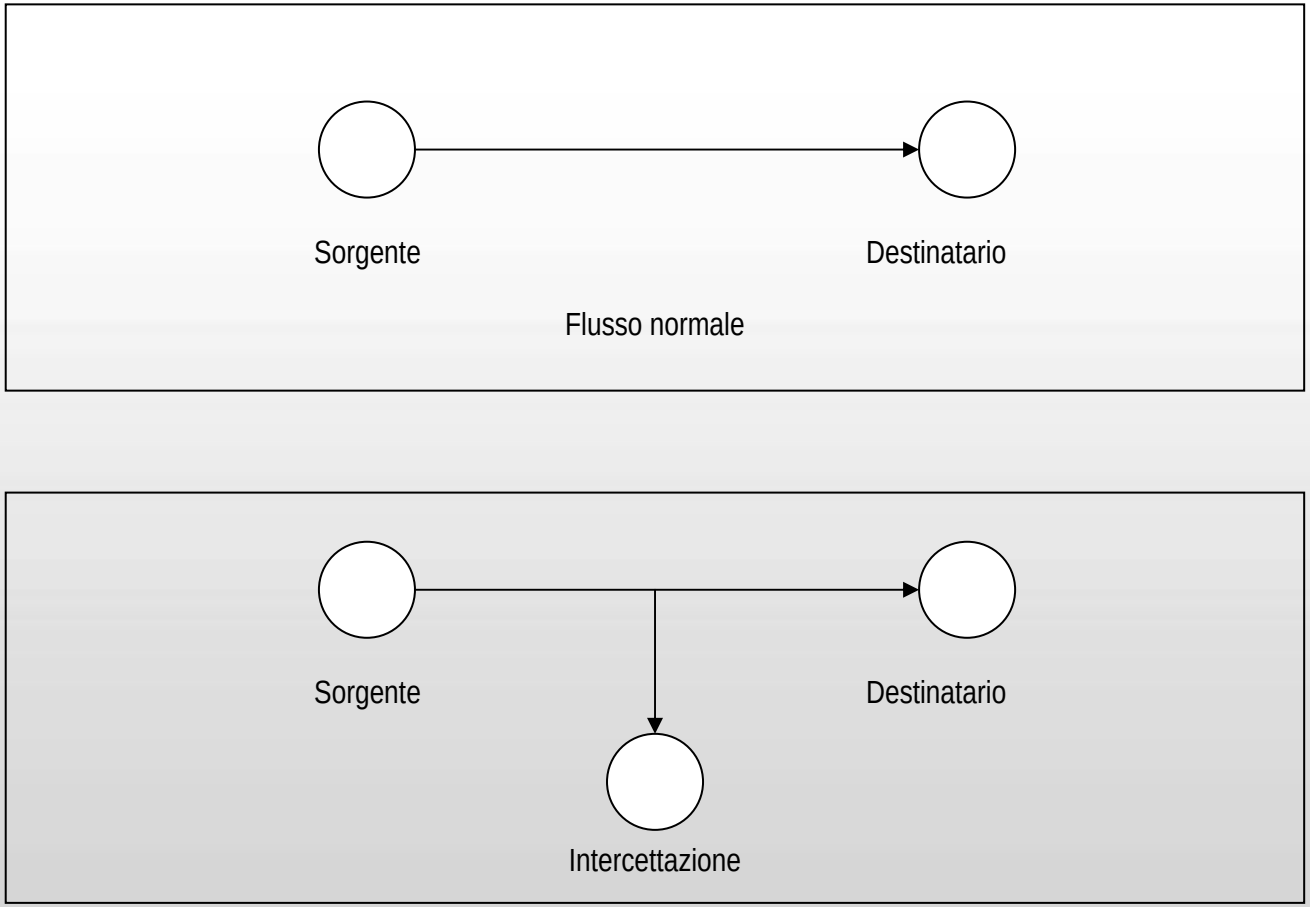
- Protezione degli accessi al sistema
  - solo gli utenti autorizzati possano usufruire delle risorse offerte dal sistema
    - Fisiche, biometriche, meccaniche, Logiche, Cifratura, firma digitale
- Protezione con software specifici
  - Firewall, Antivirus
- Segmentazione della LAN

L’attacco

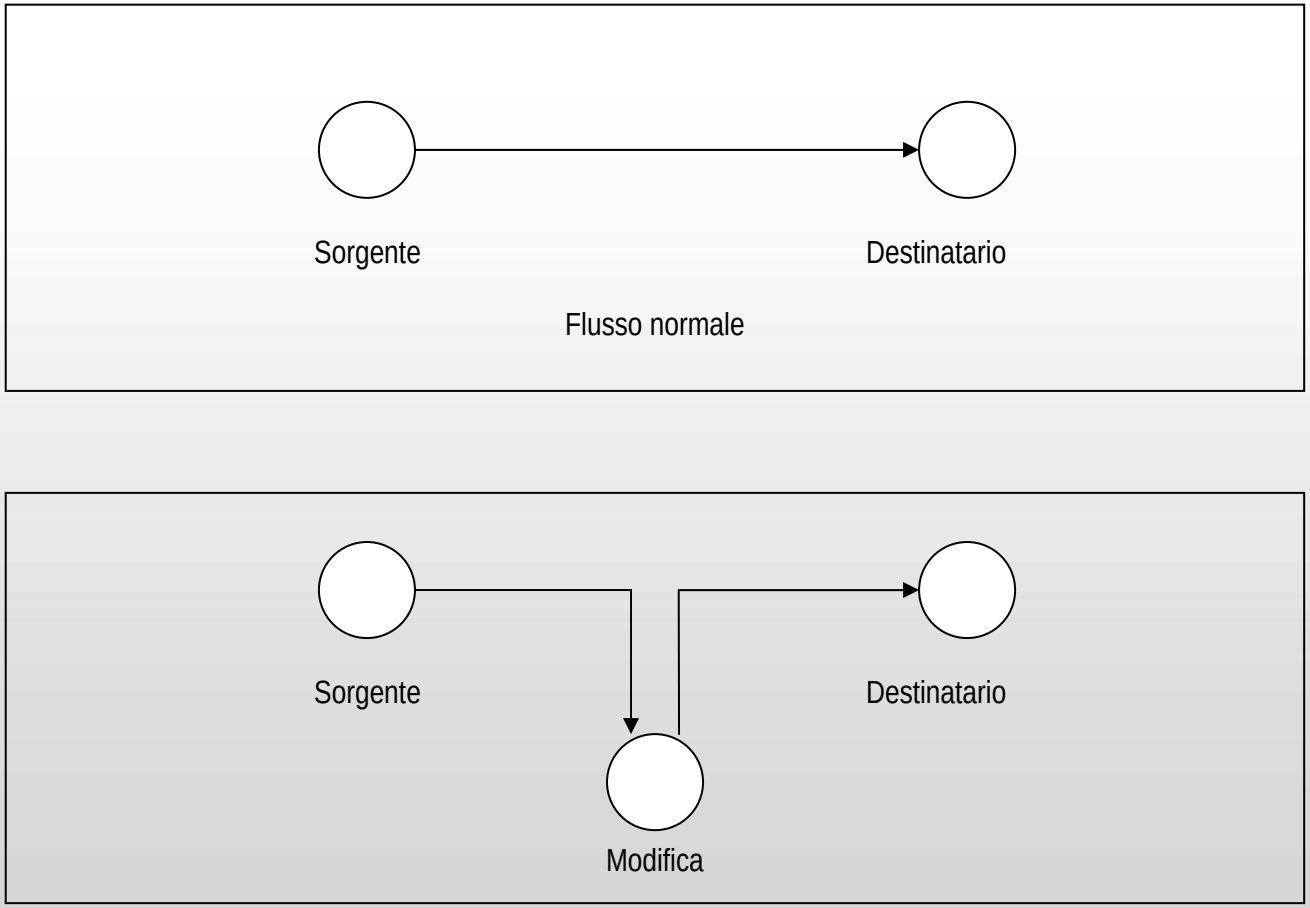
Interruzione



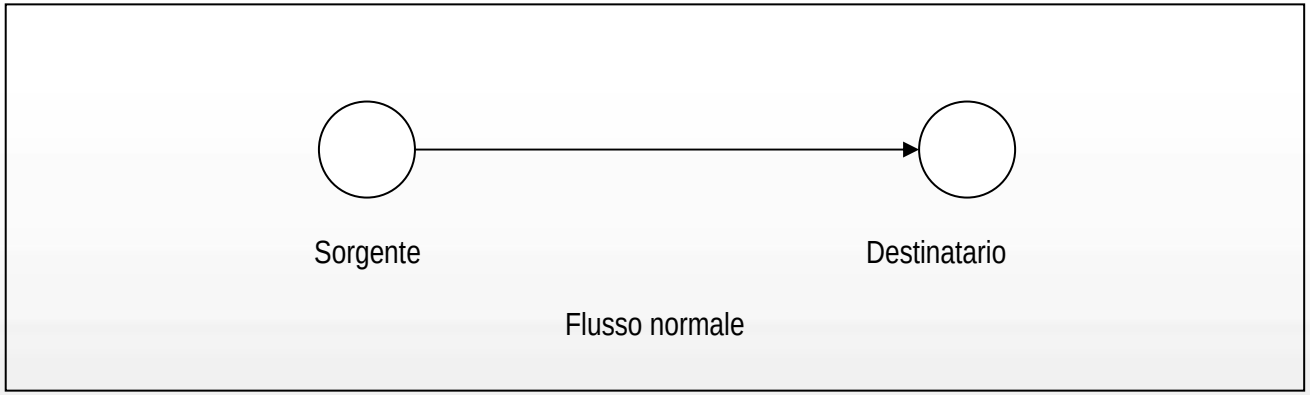
Intercettazione

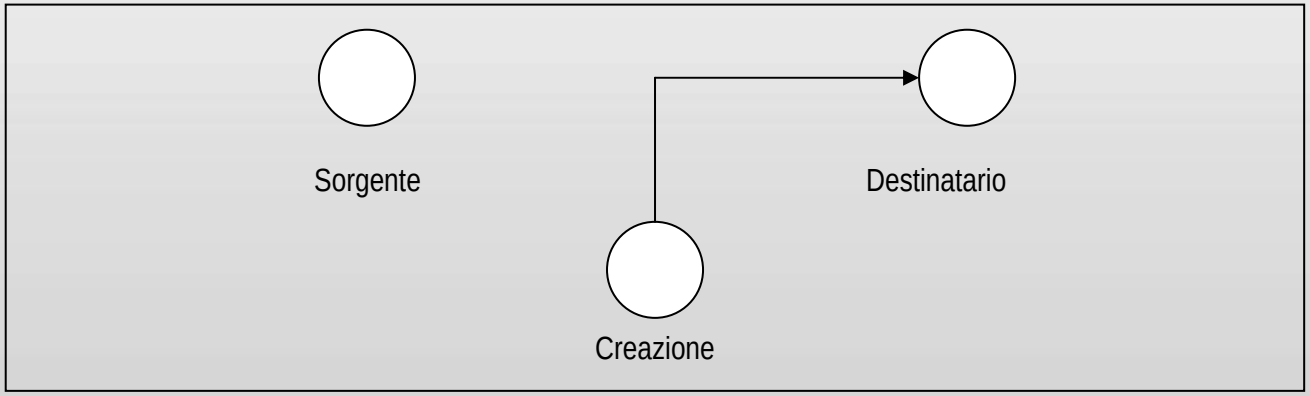


Modifica


Sorgente
Destinatario
Flusso normale
Sorgente
Destinatario
Modifica

Creazione


Sorgente
Destinatario
Flusso normale


Sorgente
Destinatario
Creazione

DoS

Denial of Service

- Tentativi di rendere una risorsa di un computer non disponibile ai suoi legittimi utilizzatori.
  - Tipici bersagli sono web server di alto profilo, come banche, gateways per i pagamenti via carta di credito o server DNS
- Saturare la macchina della vittima di richieste false
  - non può più rispondere alle richieste legittime
  - lo fa lentamente

Il Backup dei dati

Definizioni

- Backup
  - Effettuare copie di dati o programmi sensibili
- Disaster Recovery
  - Misure atte a ripristinare il sistema dopo un evento distruttivo

Strategie

- Backup totali
- Backup Incrementali
- Schedulazione dei backup
- Protezione delle copie

Il Backup dei dati

Tecnologie e processi

- Streamer - Tape
- Dischi ottici
- Dischi rimovibili
- Disk Array
- Mirroring

Attività

- Ripristino dei dati dal backup
- Ripristino dei sistemi hardware
- Ripristino dei sistemi software

Crittografia e Steganografia

Kryptos (segreto) e Grafein (scrittura)

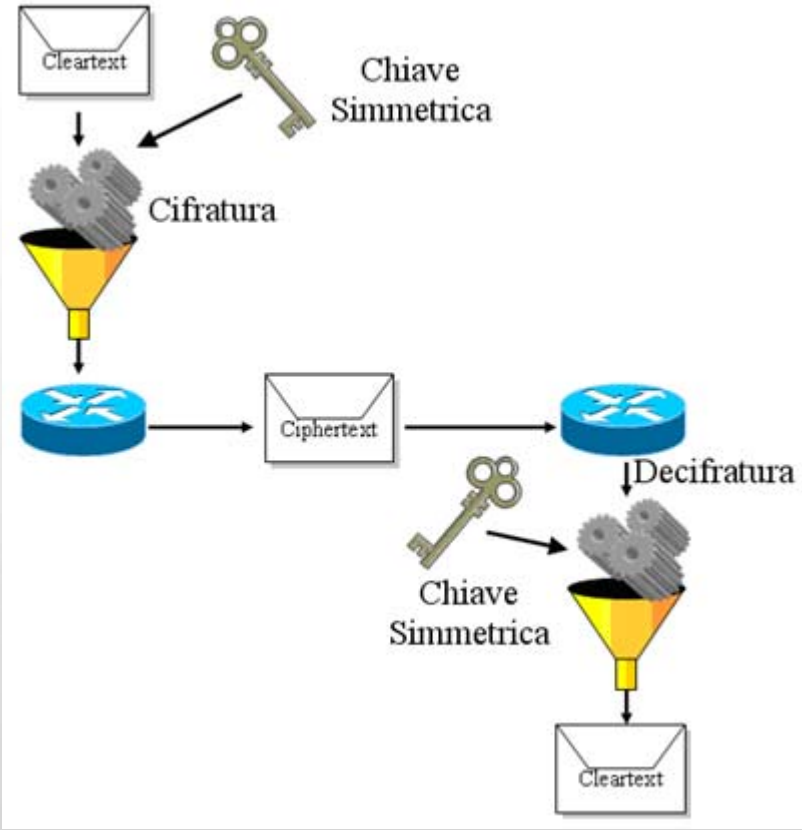
- Nascondere un messaggio da trasmettere tra due entità così che una terza non possa **capirne il significato**.
  - Trasposizione monoalfabetica lineare (traslazione)
  - Trasposizione plurialfabetica

Steganos (coperto) e Grafein (scrittura)

- Nascondere un messaggio da trasmettere tra due entità così che una terza non lo posso **trovare**
  - Antica Grecia: Scrittura su tavolette ricoperte di cera
  - Tatuaggi sulla testa rasate degli schiavi
  - Antica Cina: mess. scritti sulla seta avvolta e ricoperta di cera fatta ingoiare
  - Messaggi di testo in file MP3 o JPG

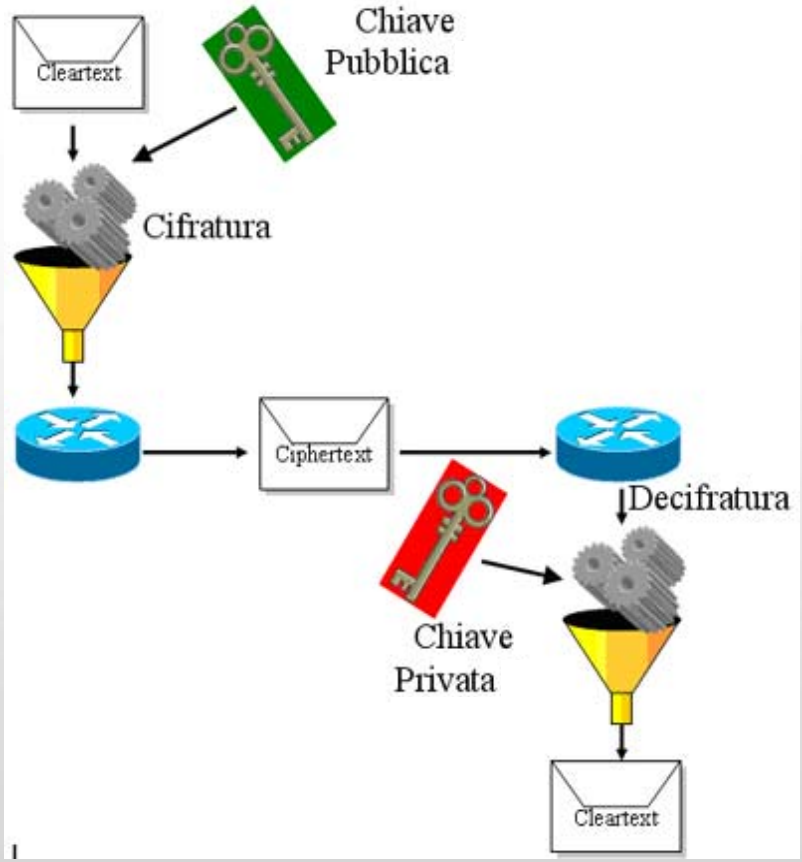
Simmetrica (a chiave privata)

- Protocollo:
  - A e B si accordano su un sistema di cifratura
  - A e B si accordano su una chiave di cifratura
  - A cifra il suo messaggio con la chiave decisa
  - A invia il messaggio a B
  - B decifra il messaggio con la stessa chiave



Asimmetrica (a chiave pubblica)

- Idea di base:
  - 2 chiavi diverse per cifrare e decifrare
- Vantaggio principale:
  - Non serve scambiarsi la chiave di cifratura (punto debole della cifr. simmetrica)
- Funzionamento:
  - A genera una propria coppia di chiavi (pubblica e privata)
  - B cifra messaggi con la chiave pubblica di A
  - Solo il VERO destinatario può decifrare il messaggio con la propria chiave privata



Crittografia

Requisiti crittografia asimmetrica

- Deve essere computazionalmente facile generare una coppia di chiavi
- Deve essere computazionalmente facile cifrare un messaggio con la chiave pubblica
- Deve essere computazionalmente facile decifrare un messaggio con la chiave privata
- Deve essere computazionalmente impossibile calcolare la chiave privata dalla chiave pubblica
- Deve essere computazionalmente impossibile risalire al messaggio conoscendo la chiave pubblica e il messaggio cifrato
- Ognuna della 2 chiavi deve decifrare un messaggio cifrato con l'altra chiave (importante per la firma)

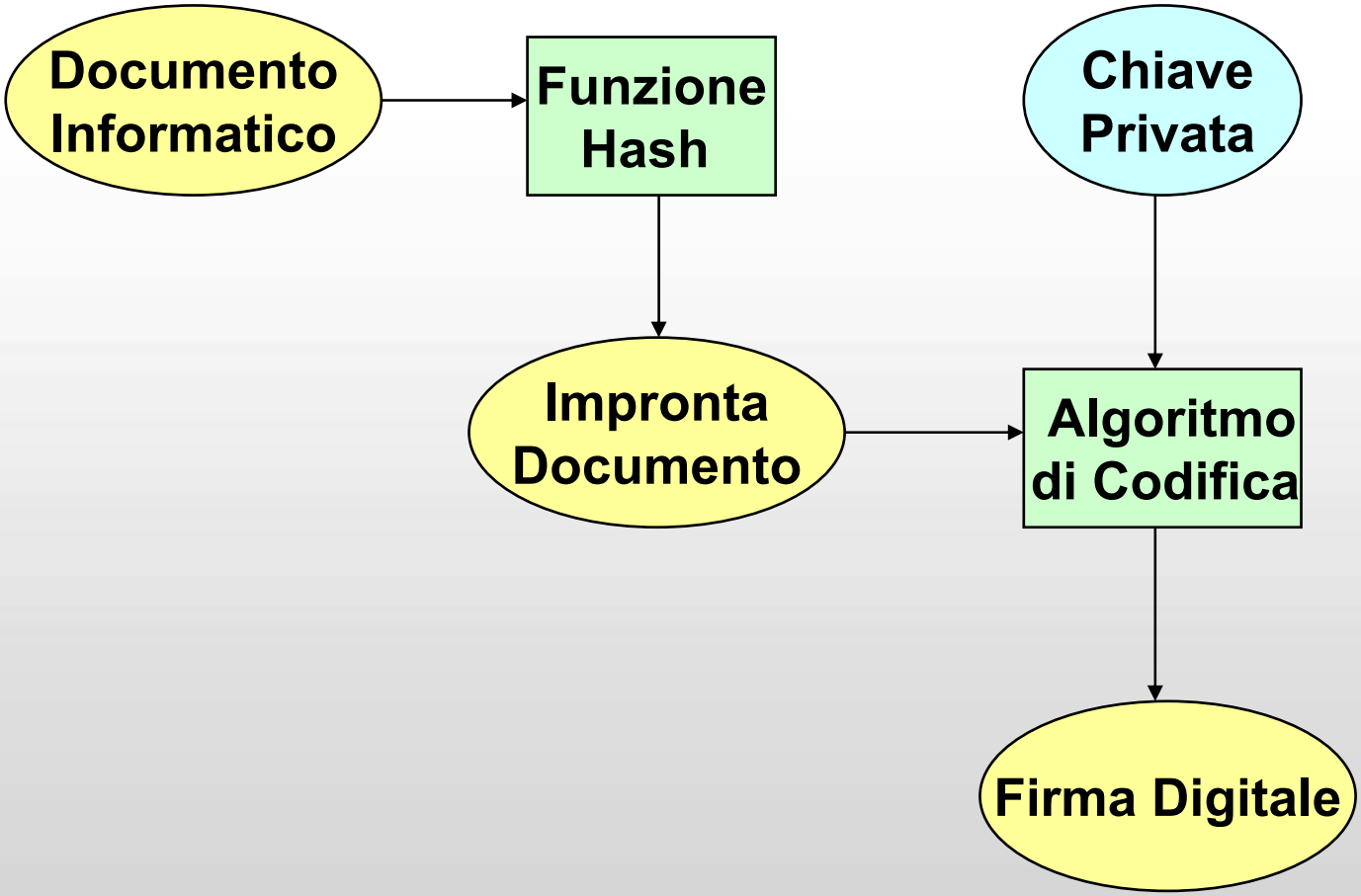
Alternativa mista

- In realtà il sistema non è così semplice
  - Per trasmettere grandi quantità di dati occorre tanto tempo
- Quindi A e B si scambieranno con questo sistema una chiave segreta (che non occupa molto spazio)
  - La useranno per comunicare tra loro usando un sistema a crittografia simmetrica, più semplice e veloce.

Firma digitale

- Se il mittente cifra il messaggio con la propria chiave privata, il destinatario puo' verificare (tramite chiave pubblica del mittente) che il messaggio provenga proprio da lui!
  - In questo caso il messaggio e' pero' leggibile da tutti!
  - Ma se il messaggio viene cifrato sia con la chiave privata del mittente, sia con la chiave pubblica del destinatario, solo il destinatario puo' leggerlo ed essere anche sicuro della provenienza!

Firma digitale



SSL e TSL

Secure Sockets Layer - Transport Layer Security

- Negoziare i parametri iniziali per la connessione, per poi instaurare un dialogo cifrato con algoritmo simmetrico
  - Negoziazione dell'algoritmo da utilizzare
  - Scambio di chiavi segrete tramite cifratura a chiave pubblica e identificazione tramite l'utilizzo di certificati
  - Cifratura del traffico tra le parti a chiave (segreta) simmetrica

Garanzie

- Autenticazione tramite crittografia asimmetrica
  - Client (opzionale) e server
- Confidenzialità tramite crittografia simmetrica
- Affidabilità tramite funzioni hash sicure
  - Prevenire alterazioni casuali

VPN

Virtual Private Network

- Rete privata instaurata tra soggetti che utilizzano un sistema di trasmissione pubblico e condiviso come Internet
- Ottenere le stesse possibilità delle linee private in affitto ad un costo inferiore

Garanzie

- Tutto il traffico su una VPN sicura deve essere criptato ed autenticato
  - Segretezza dei messaggi
  - Autenticazione dell'utente
  - Integrità dei messaggi
- Protocolli impiegati
  - IPsec (IP security)
  - PPTP (point-to-point tunneling protocol)
  - SSL/TLS, utilizzate per il "Tunneling" dell'intera rete

HTTPS

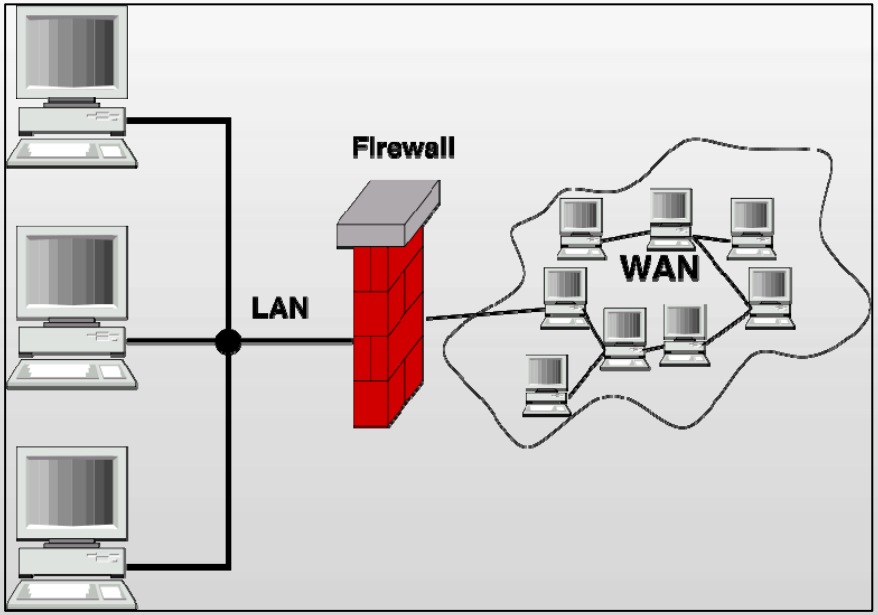
Hypertext Transfer Protocol over Secure Socket Layer

- Aggiungere sicurezza alle pagine web
  - applicazioni di commercio elettronico
  - accesso ai conti bancari
- Creazione di un canale di comunicazione criptato tra il client ed il server, attraverso i meccanismi SSL.
- All'interno di tale canale si usa HTTP per la comunicazione.

Il firewall

Cos‘è questo Muro Tagliafuoco

- Dispositivo Hardware o Software
  - Protegge/filtra il traffico dei dati da e verso una LAN, controllandone i pacchetti
  - Evita accessi non autorizzati
  - Ciò che non è **espressamente** permesso è vietato
- Limiti
  - Non garantisce l’integrità dei dati
  - Non protegge da eventi dannosi
  - Non autentica le fonti



Tipologie

- Firewall a livello di rete
- Firewall a livello di applicazione
  - Firewall a livello di circuito
- Personal o Perimetral firewall

Tipi di firewall

Caratteristiche

- Packet filter
  - analizza gli header di ciascun pacchetto (stateless)
- Stateful
  - tiene traccia delle relazioni tra i pacchetti per riconoscere pacchetti TCP malevoli che non fanno parte di alcuna connessione
- Application filter: controlla fino al livello 7 della pila ISO/OSI
  - Valuta anche il contenuto applicativo dei pacchetti
  - Riconosce e blocca dati appartenenti a virus o worm noti in una sessione HTTP o SMTP
- I proxy rientrano tra gli application firewall
  - permette le connessioni in modo selettivo, solo per i protocolli supportati.
- Personal Firewall
  - si installa sul sistema da proteggere
  - effettua controlli su tutti i programmi che tentano di accedere ad Internet
  - ZoneAlarm – Comodo

DPS

Documento programmatico annuo sulla sicurezza

- Aggiornato entro il 31 marzo di ogni anno
- Contiene
  - l'analisi dei rischi che incombono sui dati
  - le misure per garantire l'integrità e la disponibilità dei dati
  - la previsione di idonei interventi formativi degli incaricati del trattamento per renderli edotti dei rischi che incombono sui dati
  - la descrizione dei criteri da seguire per garantire l'adozione delle misure minime di sicurezza in caso di outsourcing dei trattamenti.
- Per i dati personali (stato di salute e la vita sessuale) trattati da organismi sanitari
  - i criteri per la cifratura o per la separazione dei dati dai dati personali
  - disgiunzione dei dati anagrafici da quelli riferiti alla salute.

Fine I lezione
Componenti delle piattaforme

GIS e Geo WEB: piattaforme e architetture