

# Reti di calcolatori

# Reti di calcolatori

- Rete di calcolatori: insieme di calcolatori interconnessi mediante una rete di comunicazione.
- Si possono distinguere
  - reti di calcolatori vere e proprie
    - ogni calcolatore mantiene una propria indipendenza, gli utenti interagiscono in modo esplicito con la rete
  - sistemi distribuiti
    - gli utenti non hanno visibilità sull'architettura del sistema
    - si presentano come sistemi omogenei progettati per eseguire un'applicazione particolare (Es: BANCOMAT)

# Finalità di una rete

- condivisione di risorse (memoria di massa, stampanti, CDROM, ...)
- comunicazione tra utenti (posta elettronica, file transfer,...)
- miglioramento dell'affidabilità
- risparmio attraverso la condivisione di risorse

# Classificazione

- tecnologia di comunicazione
  - reti punto a punto
  - reti multipunto (o broadcast)
- dimensione
  - Local Area Network (LAN): limitata estensione, calcolatori all'interno di un edificio
  - Metropolitan Area Network (MAN): area urbana
  - Wide Area Network (WAN): ampie dimensioni, collegamenti anche intercontinentali
  - reti di reti (internet working): collegamenti di più reti differenti mediante elementi di interfaccia

# Reti punto a punto

- Ogni coppia di calcolatori è connessa mediante una linea di comunicazione dedicata
- È una soluzione poco praticata perché
  - se i calcolatori da connettere sono fisicamente molto distanti, il costo della linea dedicata potrebbe essere molto elevato
  - se dei calcolatori devono essere collegati in modo non continuativo, l'uso di linee dedicate non ha molto senso ed è costoso
  - Nel caso si debbano connettere i calcolatori con modalità "tutti con tutti", il numero di connessioni dedicate da creare è molto elevato.

# Reti Broadcast

- Le reti broadcast sono caratterizzate da canali di comunicazione condivisi
- Ogni calcolatore sulla rete è caratterizzato da un identificatore univoco (indirizzo di rete)
- Ogni messaggio spedito da un calcolatore include l'indirizzo del destinatario.
- Ogni messaggio spedito raggiunge tutti i calcolatori connessi, solo il calcolatore il cui indirizzo di rete corrisponde a quello contenuto nel messaggio riceve il messaggio.

# WAN

- Nel caso più generale una rete geografica (WAN) può essere rappresentata su 3 livelli:
  - rete di trasmissione: si occupa della trasmissione dell'informazione fra i diversi nodi (IMP - Interface Message Processors), instradamento, verifica della correttezza dei messaggi
  - rete di calcolatori (hosts): i calcolatori sono collegati agli IMP. Lo scopo è quello di eseguire le applicazioni.
  - rete degli utenti: comunicazione fra utenti, esecuzione di applicazioni distribuite in maniera trasparente
- Le reti organizzate secondo questi livelli sono chiamate reti commutate

# WAN

- Le reti geografiche si classificano a seconda della strategia utilizzata per instradare i messaggi
  - commutazione di circuito
    - Viene inizialmente stabilita una connessione temporanea tra mittente e destinatario. Una volta che la connessione è stabilita, inizia il trasferimento dei dati.
    - Vantaggiosa nel caso in cui la comunicazione sia molto frequente e non si abbiano periodi di inattività della linea. (es. sistema telefonico).

# WAN

- commutazione di pacchetto
  - il messaggio è frammentato in pacchetti di piccole dimensioni. Ogni pacchetto contiene l'indirizzo del destinatario e un numero progressivo che indica la posizione del pacchetto nel messaggio.
  - Vantaggi: uso dei canali più efficiente, possibilità di invio anche in caso di traffico elevato, gestione di diversi livelli di priorità
- due metodi di instradamento
  - a datagrammi (spedizione dei pacchetti in maniera indipendente, anche su linee diverse)
  - a circuito virtuale (si stabilisce un cammino univoco tra mittente e destinatario per tutti i pacchetti).

# LAN

- Obiettivo: connettere dispositivi fisicamente vicini => tecnologie diverse da quelle impiegate nelle WAN (tipicamente usano un approccio broadcast).
- Le LAN sono generalmente
  - veloci (canali di comunicazione con ampia capacità di banda)
  - facilmente espandibili, modulari
  - economiche
  - affidabili

# LAN

- Le LAN possono essere caratterizzate dalla loro topologia di rete, ossia dalla struttura delle loro connessioni
  - topologia a bus (semplice, flessibile, economica, es: ethernet)
    - il guasto di un host non implica la disabilitazione dell'intera rete
    - svantaggio: unico mezzo di comunicazione, pessime prestazioni in caso di molto traffico
  - topologia a stella (connessione punto a punto)
    - controllo centralizzato delle risorse, ottime prestazioni
    - rischio di sovraccarico del server, dipendenza dall'affidabilità del server

# LAN

- Topologia ad anello: prevede una connessione circolare degli host
  - possono avere dimensioni molto elevate
  - limitata flessibilità, è difficile espanderla
  - l'affidabilità della rete dipende dall'affidabilità di tutti gli host (soluzione: reti a doppio anello)
- Topologia Wireless: approccio broadcast
  - trasmissione basata su onde radio

# Internetwork

- Una LAN non può crescere oltre una certa misura, se non al prezzo di un rilevante decadimento delle prestazioni.
- Soluzione: connettere più LAN insieme (LAN estesa) mediante
  - repeater (se le LAN sono identiche)
  - switch (permette un maggior numero di tx in contemporanea)
  - bridge (per filtrare i pacchetti e in caso di reti distinte ma con stesso schema di indirizzamento)
  - router (per reti con schemi di indirizzamento distinti, ma stesso protocollo di rete)
  - gateway (per reti completamente incompatibili)

# Architettura del SW di rete

- I calcolatori collegati in rete devono cooperare per stabilire una comunicazione e per poter trasmettere dei dati.
- Le regole che formalizzano questa cooperazione sono chiamate protocolli di comunicazione.
- Condividendo lo stesso protocollo due calcolatori possono comunicare.

# Protocolli di comunicazione

- mediante i protocolli di comunicazione si specificano tra l'altro
  - formati dei dati, struttura dei messaggi (pacchetti), velocità di trasmissione, come e quando comunicare...
- La comunicazione può essere piuttosto complessa, pertanto si specificano insiemi di protocolli, ognuno dei quali formalizza aspetti distinti della trasmissione. Ciò rende il progetto del SW di rete più semplice e flessibile

# Architettura di rete

- Un'architettura di rete dovrebbe essere organizzata gerarchicamente seguendo i principi qui elencati
  - ogni protocollo formalizza un diverso livello di astrazione della comunicazione
  - le funzioni associate ad ogni livello sono ben definite e omogenee tra loro
  - scopo di ogni livello è di fornire servizi al livello superiore, mascherando il modo in cui sono implementati. Ogni livello è caratterizzato da una interfaccia con il livello superiore.
  - Le interfacce sono progettate in modo da minimizzare gli scambi tra un livello e l'altro

# Architettura a livelli

- Consideriamo due processi (sorgente e destinazione), residenti su host distinti, che vogliono comunicare
- ognuno dei due host è dotato di un insieme di protocolli organizzati su livelli. In entrambi gli host, i livelli devono essere organizzati in maniera compatibile.
- La trasmissione avviene simulando la comunicazione tra livelli corrispondenti, seguendo le regole stabilite dai corrispondenti protocolli di livello.
- Ogni protocollo è caratterizzato da sintassi, semantica e temporizzazione
- Standard ISO/OSI

# Standard TCP/IP

- È l'architettura a livelli che permette l'interoperabilità tra reti fisiche diverse. È semplice ed efficiente
- È lo standard de facto di Internet
- È un'architettura basata su 5 livelli
  - livello applicazione
  - livello di trasporto (implementato dal protocollo TCP-Transmission Control Protocol)
  - livello internet (implementato dal protocollo IP-Internet Protocol)
  - livello di accesso alla rete
  - livello fisico

# Standard TCP/IP

- il livello fisico definisce l'interfaccia fisica tra stazione e mezzo di trasmissione
- il livello di accesso alla rete si occupa dello scambio di dati tra un sistema finale e la rete. I dati sono organizzati in frame.
- il livello internet si occupa di interfacciare sistemi su reti distinte (formato dei pacchetti e meccanismi per farli transitare attraverso uno o più router).
- il livello di trasporto si occupa di garantire l'affidabilità della trasmissione dei dati
- il livello applicazione specifica come le applicazioni possono utilizzare i servizi forniti dall'insieme di protocolli TCP/IP.

# Funzionamento di TCP/IP

- Ogni host individuato da un indirizzo univoco all'interno della rete.
- Ogni applicazione di un host è individuata da un indirizzo univoco all'interno del host (porta).
- Per poter comunicare con un processo in un host è necessario specificare l'indirizzo del host e la porta del processo.

# Funzionamento di TCP/IP

- Quando un processo in un host A vuole trasmettere un messaggio a un processo in un host B deve:
  - consegnare il messaggio al livello TCP indicando l'host destinatario (B) e la porta del processo destinatario
  - TCP aggiunge info di controllo al messaggio (lo può anche spezzettare) e lo invia al livello IP. Il livello IP aggiunge informazioni di controllo al messaggio e lo invia al livello di accesso alla rete
  - il livello di accesso alla rete aggiunge un'ulteriore intestazione al messaggio e finalmente lo invia al livello fisico che lo spedirà.

# Indirizzi Internet

- Il protocollo IP fornisce uno schema di indirizzamento universale.
  - indirizzi a 32 bit, di solito rappresentati come successione di 4 numeri decimali separati da punti: X.Y.Z.W, con X,Y,Z,W compresi tra 0 e 255.
  - Es.: 66.102.11.104
  - Esiste un'organizzazione che assegna univocamente gli indirizzi IP chiamata IANA (Internet Assigned Number Authority)
  - La versione attuale del protocollo IP non è più sufficiente: si è ormai vicini alla saturazione. Nuovo protocollo: IPv6, indirizzi a 128 bit, presenza di un supporto per audio e video,...

# Indirizzi Internet Simbolici

- Gli indirizzi numerici IP sono machine-oriented e difficilmente memorizzabili dall'uomo.
- soluzione: associare ad ogni indirizzo numerico un indirizzo simbolico mediante un sistema denominato DNS (Domain name system). Es. `www.google.com`
- I DNS eseguono la conversione degli indirizzi in maniera totalmente trasparente all'utente. Gli indirizzi simbolici sono suddivisi in domini e sottodomini.
- Es. Considerate `www.google.com`
  - `.com` indica il dominio di primo livello
  - `google.com` il dominio di secondo livello
  - `www.google.com` individua il calcolatore nel sottodominio `google.com`

# Affidabilità di TCP/IP

- Protocolli orientati alla connessione (connection-oriented)
  - TCP
- Protocolli non orientati alla connessione (connectionless)
  - IP, UDP (User Datagram Protocol)

# Il World Wide Web

- L'architettura client-server è alla base di quasi tutti gli applicativi di Internet.
- Il WWW (World Wide Web) è una tipica applicazione distribuita client-server composta da:
  - Client, ovvero i browser (navigatori) Internet: e.g. Safari, Netscape, Internet Explorer,...
  - i Web server (Apache, Internet Information Server, ...)
- Client e server comunicano mediante il protocollo HTTP (HyperText Transfer Protocol)

# HTTP

- il protocollo HTTP è utilizzato per condividere informazione sulla rete (pagine Web, documenti multimediali, etc.)
- Esempio d'uso:
- Il cliente (browser) richiede un'informazione indicando un indirizzo (Es.: `http://polygen.org/web/Ghezzi.645.0.html`) e genera un messaggio di tipo GET contenente indirizzo ed altre informazioni ausiliarie da inviare al server (e.g. versione del protocollo, tipo di browser, etc.).
- Il server riceve la richiesta da parte del browser, reperisce l'informazione nel suo disco locale e la spedisce al cliente insieme ad alcune informazioni di controllo (e.g. codice d'errore, tipo di documento, lunghezza del documento, tipo di web server,...)
- è un protocollo stateless: non mantiene la storia delle richieste ricevute.

# Altri servizi di rete

- altri servizi Internet sono implementati mediante applicazioni client server
  - TELNET, permette di impiegare il proprio calcolatore come terminale di un calcolatore (protocollo SSH: versione di telnet sicura)
  - posta elettronica (SMTP - Simple Mail Transfer Protocol, POP3, IMAP)
  - trasferimento file (FTP - File Transfer Protocol, SFTP - Secure File transfer Protocol,...)
  - Utilità di rete
    - ping
    - traceroute
    - ...

# Struttura del sistema email

- è gestito mediante un protocollo client-server.
  - Il programma client (e.g. Eudora, Outlook) si connette al mail server del mittente ed invia il messaggio mediante protocollo SMTP.
  - Sempre mediante SMTP il mail server del mittente invia il messaggio al mail server del destinatario.
  - L'utente destinatario scarica il messaggio dal server utilizzando uno dei seguenti protocolli
    - Post Office Protocol 3 (POP3)
    - Internet Message Access System (IMAP)

# Uniform Resource Locator

- Uniform Resource Locator (URL) permette di identificare in modo univoco la risorsa e il metodo di accesso alla risorsa sulla rete.
- In generale,

`protocollo://host:portaTCP/risorsa`

# Applicazioni su più livelli

- Un URL può individuare non solo dati ma anche programmi da dover eseguire in maniera remota.
- In questo caso il browser richiede l'esecuzione del programma al server web che a sua volta delega il compito al programma vero e proprio (back-end). Si tratta di un'architettura su 3 livelli (client, server, back-end). Il sottosistema LA è gestito dal back-end.
- È un sistema a deleghe successive: il server svolge anche un ruolo di client verso il back-end.
- La richiesta del client può contenere anche una lista di parametri sui quali eseguire il programma. Essi vengono specificati mediante lo standard Common Gateway Interface (CGI).

# Applicazioni su più livelli

- Nelle applicazioni distribuite molto complesse possiamo avere una struttura su 4 livelli. la quale prevede un'applicazione (DBMS - Data Base Management System) per l'implementazione del sottosistema di gestione dei dati connessa al Back-end.
- Browser <-> Web Server <-> Back-end <-> DBMS

# Sicurezza

- Tecniche per filtrare il traffico di rete che operano a livello IP (sui router)
  - packet filtering
- Tecniche di sicurezza a livello TCP
  - permettono di definire limitazioni nell'accesso ai servizi internet
  - Virtual Private Network (VPN)
- Esistono calcolatori destinati a proteggere reti private da eventuali intrusioni (FireWall)
  - Spamming, virus, Deny of service attacks, mail bombing,...

# Sicurezza

- Ipotizziamo che un'entità A si metta in comunicazione con un'entità B. I requisiti di sicurezza che si devono garantire sono
  - Privatezza dei dati trasmessi (evitare che un entità C intercetti il flusso di dati)
  - Integrità dei dati e servizi richiesti (C potrebbe frapporsi tra A e B e manipolare l'informazione in transito)
  - Identità dei partner (C potrebbe camuffarsi da A o B)
  - non ripudiabilità degli impegni (A potrebbe negare di aver inviato dati a B o B potrebbe negare di avere ricevuto dati da A).

# Sicurezza

- Per garantire i requisiti di sicurezza si utilizzano tecniche di crittografia:
  - A critta con una chiave il messaggio e lo invia a B, B riceve il messaggio e lo decritta con una chiave.
- Crittografia a chiave simmetrica
- Crittografia a chiave asimmetrica

# C. a Chiave simmetrica

- Le entità A e B devono condividere la stessa chiave per codificare e decodificare il messaggio e lo stesso programma di crittografia.
- Svantaggio: come trasferire la chiave in modo sicuro da A a B?
- La tecnica a chiave simmetrica non fornisce le capacità per garantire l'integrità dei dati e nemmeno l'identità del mittente.

# C. a chiave asimmetrica

- Non è necessario fornire un canale sicuro per il trasferimento delle chiavi. Transitano infatti sulla rete solo chiavi pubbliche.
- È necessario verificare a chi corrisponde una determinata chiave pubblica.
- Esistono Autorità di Certificazione che garantiscono l'identità del proprietario di una data chiave pubblica.
- Problema: chi controlla i controllori?

# C. a chiave asimmetrica

- Ogni entità dispone di una coppia di chiavi (SW: GnuPG, PGP):
  - la chiave pubblica
  - la chiave privata
- A critta il messaggio M con la sua chiave privata (firma digitale) ottenendo il messaggio firmato M1
- A critta M1 con la chiave pubblica di B e ottiene il messaggio crittato M2 da spedire a B.
- B usa la sua chiave privata per decodificare il messaggio e poi la chiave pubblica di A per verificarne l'autenticità (i.e. la firma).

# Motori di ricerca

- Sono delle risorse per il recupero dell'informazione tipicamente accessibili dalla rete
- Tipicamente i risultati ottenuti sono classificati secondo la loro rilevanza mediante qualche metrica.
- Permettono di ricercare informazione in differenti ambienti (Web, newsgroups, basi di dati...)

# Motori di ricerca

- Come funzionano?
  - Browsers automatizzati esplorano il web (Web Crawlers, Web Spider, Web Robots) alla ricerca di nuove pagine Web
  - Le nuove pagine Web vengono catalogate ed il loro contenuto è indicizzato rispetto a delle parole chiave (titolo, headings, metatags...)
  - L'utente interroga il motore di ricerca fornendo tipicamente delle parole chiave (query) ed altri criteri di ricerca (alcuni motori supportano operatori booleani AND,OR,NOT, etc.)

# Un esempio: Google



# Google: criteri di ricerca



**Trova risultati** che contengano **tutte** le seguenti parole  10 risultati ▾  
che contengano la **seguente frase**    
che contengano **una qualunque** delle seguenti parole   
**che non contengano** le seguenti parole

**Lingua** Visualizza solo le pagine scritte in  ▾

**Formato file**  ▾ i risultati in  ▾

**Data** Visualizza le pagine Web aggiornate riferite a  ▾

**Cerca in** Visualizza i risultati quando i termini selezionati appaiono  ▾

**Domini**  ▾ i risultati contenuti nel seguente sito o dominio   
Esempi: .org, google.com [Ulteriori Informazioni](#)

**Diritti di utilizzo** Trova risultati che  ▾

**SafeSearch**  Nessun filtro  Filtra con [SafeSearch](#)

## Cerca pagina specifica

**Pagine simili** Cerca pagine simili alla pagina    
Esempio: [www.google.com/help.html](http://www.google.com/help.html)

**Collegamenti** Cerca pagine che contengano collegamenti alla seguente pagina

# Query complesse

Ricerca frase: "frase da cercare"

Ricerca di una delle seguenti query: query<sub>1</sub> OR query<sub>2</sub> OR... OR query<sub>n</sub>

Ricerca negativa: -query

Ricerca formato: filetype:tipo\_file\_da\_cercare

Ricerca in un sito o dominio: site:sito\_o\_dominio\_in\_cui\_cercare

Ricerca pagine simili: related: pagina

Ricerca pagine che contengono link a una data pagina: link:pagina

Ricerca nell'url: allinurl:argomento

Ricerca nel titolo: allintitle:argomento

I criteri di ricerca possono essere composti per creare query più complesse.

# S.O. di rete

- Oltre ai tipici moduli per la gestione di processore, memoria, periferiche, file system, interprete comandi si aggiunge un modulo per la
  - gestione delle risorse di rete

# Modelli organizzativi

- I modelli organizzativi monolitico e a strati non sono più applicabili.
- Modelli Client-server
- Modelli ibridi

# M. Client-server

- Il nucleo del S.O. è di dimensioni minime, la gestione delle funzionalità affidate a processi serventi (printer server, memory server, file server,...)
- i processi clienti possono usufruire di un servizio, inviando una richiesta ai processi serventi (scambio messaggi).
- Il nucleo deve fornire le funzionalità necessarie per stabilire la comunicazione tra processi clienti e processi serventi
- I processi condividono informazione solo attraverso il nucleo mediante lo scambio di messaggi
- Sono S.O. di rete altamente portabili

# Vantaggi

- indipendenza fisica tra processi server e client.
- possibilità di specializzare l'interfaccia grafica dei client, lasciando invariati i processi server.
- riduzione della manutenzione (le parti che potrebbero essere comuni a molti processi sono contenute in un unico processo server)

# Modelli ibridi

- I S.O. Client-server permettono la comunicazione fra processi solo attraverso lo scambio messaggi.
- Nei S.O. ibridi, i processi (chiamati thread) sono suddivisi in gruppi
  - all'interno di uno stesso gruppo i thread possono scambiarsi informazione mediante l'uso di aree di memoria comuni.
  - gruppi di thread distinti comunicano mediante scambio messaggi
- Permettono un uso più efficiente della CPU

# Traduzione dei programmi in rete

- In una rete ci possono essere calcolatori con diverse architetture e diversi linguaggi macchina.
- Problema: come eseguire un programma su diversi calcolatori della rete?
- Soluzione:
- Dotare ogni calcolatore di un interprete per un linguaggio intermedio (bytecode)
- Tradurre il programma ad alto livello in un programma bytecode
- interpretare il programma bytecode sui differenti calcolatori
- Esempio: il linguaggio Java

# Gestione dei processi

- Una rete di calcolatori può essere vista come una generalizzazione di un sistema multiprocessore. Si ha almeno un processore per ogni nodo.
- Il gestore dei processi deve dunque
  - gestire l'evoluzione di un singolo processo su ogni nodo
  - definire come distribuire i vari processi da eseguire sui nodi a disposizione.
- Ogni nodo ha installato un kernel per gestire l'esecuzione dei processi che gli vengono assegnati
- Esiste un processo a livello globale (dispatcher) che distribuisce i processi da eseguire fra i vari nodi secondo un certo criterio. (Es.: bilanciare il carico)

# Gestione della memoria

- La memoria è distribuita in maniera più o meno uniforme fra i vari nodi.
- Per ragioni di efficienza si segue il principio di località: si tende a mantenere sullo stesso nodo il processo e la memoria che gli è stata assegnata.

# Gestione delle periferiche

- Si assegna ad ogni periferica un nome univoco.
- Si accede alle periferiche da ogni nodo mediante il loro nome.
- Ogni periferica è gestita da un processo responsabile.

# File system distribuiti

- Un S.O. che consente una gestione distribuita del file system deve:
  - integrare in modo organico i singoli file system dei calcolatori della rete
  - risolvere i problemi dell'univocità dei nomi
  - consentire un accesso efficiente anche ai file presenti su calcolatori remoti
- Generalmente si usa un modello client-server.
- Esempio il Network File System (NFS) di Unix

# Applicazioni come sistemi distribuiti

- Un'applicazione è definita mediante 3 sottosistemi: interfaccia utente (IU), sottosistema di logica applicata (LA), sottosistema di gestione dei dati (GD)
- In un'applicazione di rete i 3 sottosistemi possono risiedere su calcolatori distinti.
- Gestione tipica: client-server
  - sul nodo client (che potrebbe essere anche un palmare o un telefono cellulare) potrebbe essere implementata solo l'interfaccia utente
  - Applicazione server e applicazione client si scambiano i dati mediante un protocollo di comunicazione.

# Applicazioni peer to peer

- Utilizzate tipicamente per lo scambio di file o per la comunicazione sincrona
- Non esiste una gerarchia tra i calcolatori che partecipano alla rete
- Le applicazioni peer to peer integrano una componente SERVER con una componente cliENT, si parla di sistemi SERVENT.
  - sistema di condivisione file in Windows
  - Gnutella, Kazaa (file sharing P2P), ....