

Cenni di Programmazione di Rete

A cura di:
Giuseppe Pozzi

Dipartimento di Elettronica e Informazione
Politecnico di Milano

giuseppe.pozzi@polimi.it
8 maggio 2001

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

1

Indice (i)

- La programmazione di rete (i):
 - Architettura delle reti:
 - reti fisiche;
 - gli indirizzi IP (Internet Protocol), istradamento e gateway;
 - stratificazione dei protocolli;
 - port di comunicazione e UDP;
 - internet e TCP/IP.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

2

Indice (ii)

- La programmazione di rete (ii):
 - Alcuni concetti di base:
 - le applicazioni distribuite ed il protocollo applicativo;
 - le interfacce ed i protocolli;
 - l'interfaccia di socket;
 - il modello client/server (C/S);
 - l'indirizzamento;
 - il modello di comunicazione e la connessione.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

3

La Programmazione di Rete e di Sistema

- Bibliografia:
 - Pelagatti G., Fondamenti di Informatica II - Programmazione in ambiente di rete. Progetto Leonardo, Bologna, 1999.
 - Comer D. E., Internetworking with TCP/IP: Principles, Protocols and Architecture, Prentice Hall, 4th ed. 2000.
Edizione italiana: Comer D. E., Internetworking con TCP/IP: principi, protocolli, architettura. Gruppo Editoriale Jackson, Milano, 1992.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

4

Architettura delle reti

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

5

Tipologia di reti

- La comunicazione tra due calcolatori avviene attraverso un mezzo trasmissivo, come ad es.:
 - doppino di rame;
 - cavo coassiale;
 - fibra ottica;
 - onde elettromagnetiche;
 - porte a infrarossi;
 - o una loro combinazione.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

6

La tecnologia di rete

- E' definita da un insieme di tipi di mezzi trasmissivi e di regole di connessione dei calcolatori e degli altri apparati di rete, tenuti insieme da un protocollo di trasmissione comune.
- I parametri della tecnologia di rete sono:
 - la distanza;
 - la velocità di trasmissione (bit al secondo);
 - il costo.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

7

Connessione a commutazione di circuito

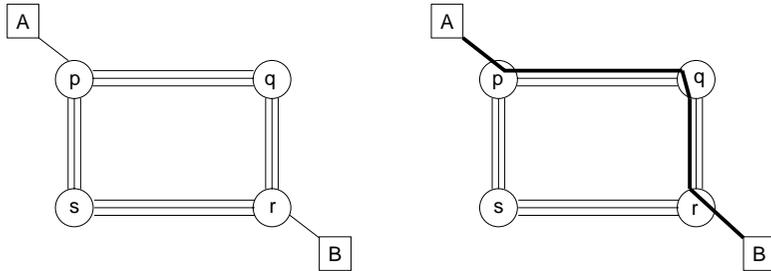
- è una connessione diretta, punto a punto;
- garantisce una banda (ad es. 64 kbps);
- tipicamente basata su linee telefoniche;
- lo sfruttamento della banda non è genericamente continuo:
(ad es. occupo la linea senza sfruttarne la banda quando leggo una pagina Web che ho appena scaricato in locale);
- elevati costi.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

8

Commutazione di circuito



A seguito della richiesta di connessione, viene creato un circuito fisico da A a B.

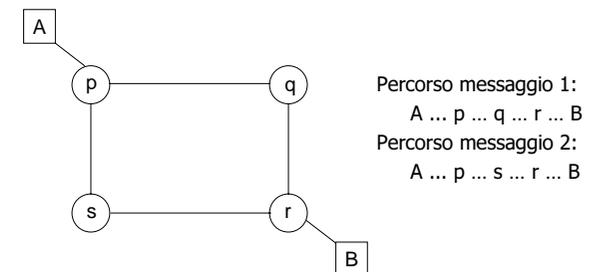
Connessione a commutazione di pacchetto

- il canale è unico, ma il traffico è diviso in piccoli messaggi (pacchetti) di poche centinaia di byte;
 - quando un calcolatore collegato in rete (host) non utilizza la banda, questa può essere utilizzata da altri calcolatori, consentendo più comunicazioni simultanee;
 - non è garantito il percorso effettuato;
 - presenta lo svantaggio di comunicazione **frammentata**.

Connessione a commutazione di pacchetto

- Non garantendo il percorso seguito dai messaggi, si realizza un servizio che può essere:
 - a circuito virtuale:
 - mantiene l'ordinamento dei messaggi durante la trasmissione;
 - oppure
 - a datagramma:
 - ogni messaggio è trattato in modo indipendente.

Commutazione di pacchetto



A seguito della richiesta di connessione, la rete non crea nessun collegamento.

Definizioni

- Tipo di rete: tecnologia di rete utilizzata.
- Istanza di rete: una specifica rete di un certo tipo.
- Categoria di reti:
 - LAN (local area network): velocità 4 Mbps - 2 Gbps;
 - MAN (metropolitan area network): velocità 56 kbps - 100 Mbps;
 - WAN (wide area network): velocità 9.6 kbps - 45 Mbps.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

13

Reti standard

- La maggior parte delle reti è stata standardizzata:
 - IEEE 802.3 (Ethernet): di tipo LAN;
 - Fast Ethernet: di tipo LAN;
 - Token-Ring (IBM): di tipo LAN.
- La comunicazione tra calcolatori anche se su reti differenti e' difficile ma irrinunciabile. L'utente vuole vedere tutto come una **unica** rete.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

14

internet

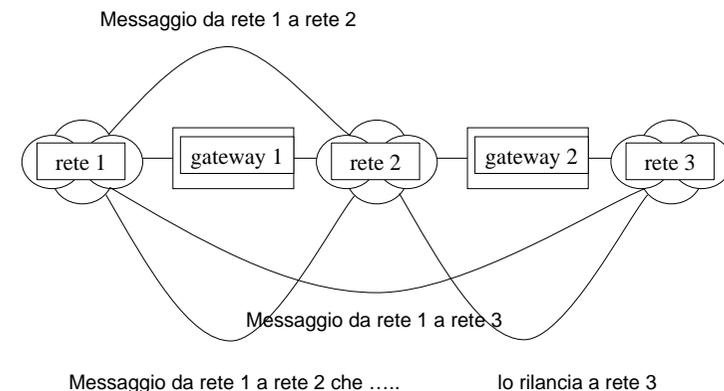
- Una rete di reti è detta internet (con la *i* minuscola).
- Una internet è definita fornendo i protocolli per trasferire le informazioni tra le varie reti. Il TCP/IP serve proprio a questo.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

15

Interconnessione tra reti



8-May-01

Fondamenti di Informatica II - Programmazione di Rete

16

Indirizzo IP

- Ogni host collegato ad una internet ha un suo proprio indirizzo (detto indirizzo IP):
 - univoco: non esistono cioè due macchine di una stessa internet che abbiano indirizzo IP uguale;
 - composto da netid e hostid, per un totale di 32 bit;
 - tutte le macchine di una rete hanno lo stesso netid.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

17

Classi di reti internet

- Classe A: reti con più di 65536 host;
- Classe B: reti con host tra 256 e 65536;
- Classe C: reti con meno di 256 host;
- Classe D: indirizzo multicast;
- Classe E: per usi futuri.

Tramite il broadcast, i messaggi sono inviati a tutti gli host della rete.

Tramite il multicast, i messaggi vengono inviati ad un gruppo di host anche se appartenenti a LAN fisiche separate.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

18

Classi di reti internet

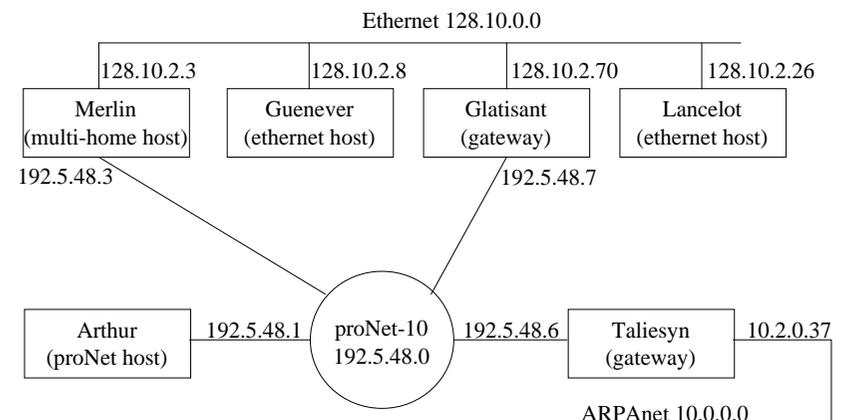
	0	1	2	3	4	5	6	7	8	16	24	31	
A	0	netid							hostid				
B	1	0	netid						hostid				
C	1	1	0	netid					hostid				
D	1	1	1	0	indirizzo multicast								
E	1	1	1	1	0	riservato per usi futuri							

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

19

Esempio di internet (Purdue University, 1980 circa)



8-May-01

Fondamenti di Informatica II - Programmazione di Rete

20

Il protocollo IP

- Il servizio realizzato da IP è la consegna del datagramma.
- Il datagramma è un pacchetto di bit contenente:
 - i dati;
 - le informazioni ausiliare quali ad es:
 - indirizzo del mittente;
 - indirizzo del destinatario.

Il protocollo IP

- Il servizio realizzato da IP è la consegna del datagramma:
 - senza connessione. Ogni pacchetto è trattato indipendentemente dagli altri;
 - inaffidabile. La consegna del messaggio al destinatario non è garantita, possono esservi pacchetti persi, duplicati, ritardati o consegnati fuori sequenza. Il mittente non ha riscontro dell'avvenuta consegna;
 - best effort. Il software compie ogni tentativo per consegnare i pacchetti correttamente.

Il protocollo IP

- Fornisce:
 - formato esatto di tutti i dati;
 - funzioni di instradamento (routing), realizzato proprio in base all'indirizzo IP;
 - ogni gateway dispone di opportune tabelle (di routing) per l'instradamento;
 - insieme di regole che inglobano l'idea di consegna non affidabile.

Stratificazione del protocollo

- Viene utilizzata una famiglia (o suite) di protocolli.
- La suddivisione in protocolli consente di dividere un problema complesso in più sottoproblemi ed una migliore analisi di ogni singolo protocollo.
- I vari protocolli gestiscono differenti problemi.

Problemi gestiti dai vari protocolli

- malfunzionamenti hardware (host o gateway);
- congestione della rete: una macchina congestionata può sopprimere ulteriore traffico;
- ritardo o perdita di pacchetti: il software deve riconoscere ed adattarsi a ritardi lunghi e di durata variabile;
- alterazione dei dati dovuta a possibili errori di trasmissione;
- duplicazione dei dati o errore nella sequenza di trasmissione, ad es. dovuti a reti che offrono più percorsi alternativi.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

25

I protocolli stratificati

Il modello ISO/OSI prevede la definizione di 7 livelli di protocollo. Partendo dal più alto:

(ISO/OSI)

- 7) di applicazione;
- 6) di presentazione;
- 5) di sessione;
- 4) di trasporto
- 3) di rete;
- 2) di collegamento dei dati;
- 1) fisico.

Internet Protocol Suite

- http, smtp, ftp, telnet ...
- (compressione pacchetti)
- (DNS)
- TCP e UDP
- IP e routing
- non specificati
- non specificati

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

26

I protocolli stratificati

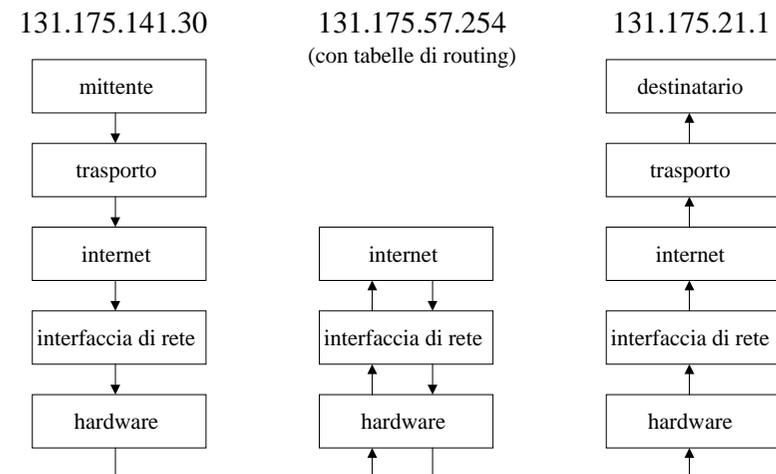


8-May-01

Fondamenti di Informatica II - Programmazione di Rete

27

Comunicazione tra due host



8-May-01

Fondamenti di Informatica II - Programmazione di Rete

28

Svantaggi della stratificazione

- Se la trasmissione è nella stessa sottorete, conviene utilizzare pacchetti di grandi dimensioni: se la trasmissione è tra sottoreti differenti, conviene utilizzare pacchetti di dimensioni inferiori.
- La stratificazione rende difficile l'ottimizzazione delle comunicazioni.
- Ma ... è possibile utilizzare dei buffer.

Terminologia

- internet: una rete di reti;
- Internet (INTERconnected NETwork): la più diffusa internet del mondo.
- TCP/IP: il più diffuso protocollo per creare internet, ed usato da Internet.
- intranet: una rete privata basata sulle stesse tecnologie di Internet.
- extranet: diverse intranet collegate tra loro.

User Datagram Protocol

- IP non consente di distinguere più destinazioni di datagrammi all'interno della stessa macchina (vede **solo** l'indirizzo IP).
- E' una limitazione molto pesante: basti pensare che un semplice *telnet* su un host, lo isolerebbe da altre connessioni.
- Per superare tale limite è stato definito il protocollo User Datagram Protocol.

UDP

- Per distinguere più connessioni verso uno stesso calcolatore, si introduce il concetto di *port*, inteso come punto di destinazione di una connessione.
- Ogni port è identificato univocamente da un *unsigned int*.
- Un processo accede ad un port in modo sincrono e bufferizzato.
- La connessione UDP richiede quindi un indirizzo IP ed un numero di port.

UDP

- A parte la specifica di un port, UDP ha gli stessi vincoli di IP:
 - senza connessione;
 - inaffidabile;
 - best-effort.

Il datagramma UDP

0	15	16	31
port provenienza UDP		port destinazione UDP	
lungh. messaggio UDP		checksum UDP	
dati			
dati			

Alcuni port riservati

num.	nome	descrizione	
11	USERS	systat	utenti attivi
13	DAYTIME	daytime	ora del giorno
15		netstat	stato della rete
17	QUOTE	quotd	citazione del giorno
20	FTP-DATA	ftp-data	dati FTP
21	FTP	ftp	FTP
23	TELNET	telnet	connessione di terminale
25	SMTP	smtp	simple mail transport protocol
37	TIME	time	tempo
42	NAMESERVER	name	server di nomi dell'host
53	DOMAIN	nameserver	server di nomi del dominio
79	FINGER	finger	
80	WEBSERVER	web server	server web
93	DCP	device control protocol	

Servizio affidabile di trasporto di stream

- TCP (Transmission Control Protocol)/IP:
 - orientamento allo stream: trasmette sequenze di byte;
 - connessione di circuito virtuale: chiamante e chiamato stabiliscono una connessione full-duplex utilizzando un circuito virtuale. In caso di fallimento, entrambi se ne accorgono;
 - trasferimento bufferizzato;
 - stream non strutturato: ad es. un record può essere frammentato su più pacchetti.

L'affidabilità di TCP/IP

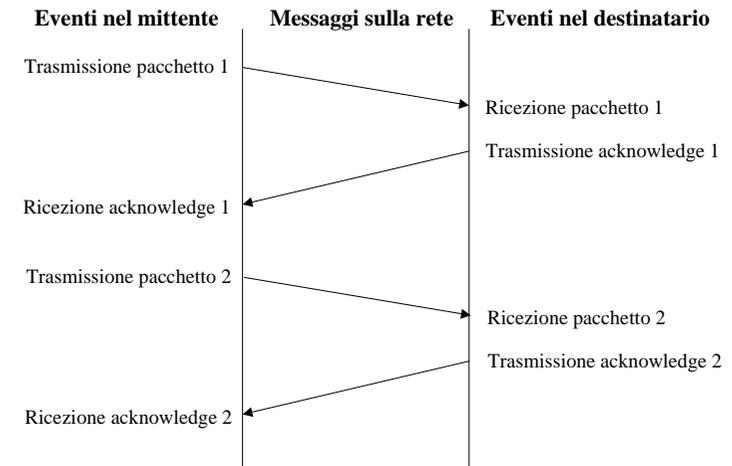
- E' una caratteristica imprescindibile.
- E' basata sul riscontro positivo di ricezione (PAR - Positive Acknowledge with Retransmission).
- Il destinatario informa il mittente della ricezione del messaggio.
- Il mittente se non ottiene riscontro dal destinatario entro un certo tempo (time-out), arguisce la perdita del pacchetto.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

37

Protocollo con riscontro positivo di ricezione

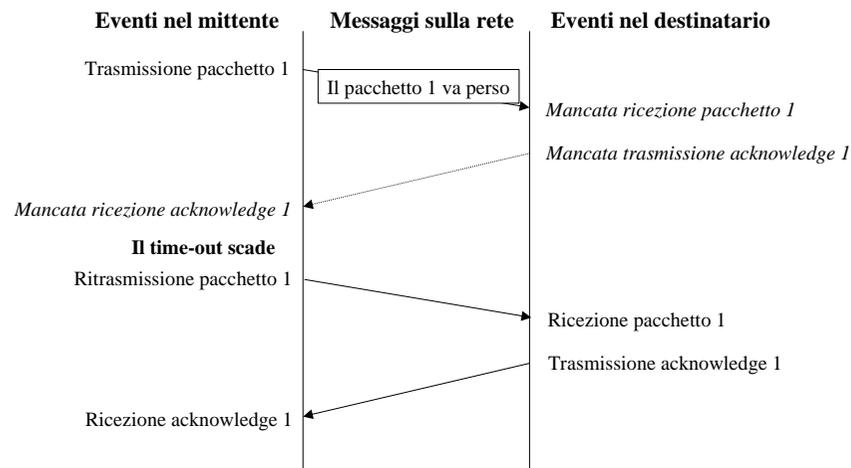


8-May-01

Fondamenti di Informatica II - Programmazione di Rete

38

Trasmissione con perdita di pacchetto

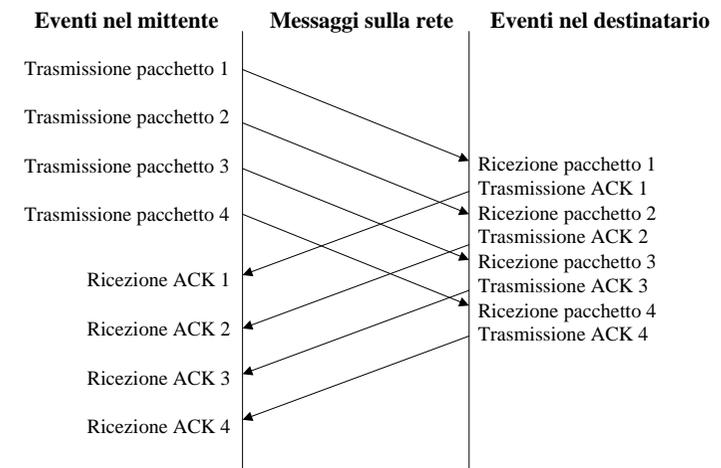


8-May-01

Fondamenti di Informatica II - Programmazione di Rete

39

Trasmissione contemporanea di più pacchetti



8-May-01

Fondamenti di Informatica II - Programmazione di Rete

40

Controllo di flusso

- Ogni riscontro dal ricevente indica anche il numero di byte che il ricevente è in grado di accettare.
- In tal modo il ricevente può indicare al trasmittente se si ha congestione sulla linea e addirittura annullare la trasmissione di altri pacchetti indicando di essere disposto a ricevere 0 byte.

Il datagramma TCP

- Il datagramma TCP descrive il formato dei pacchetti.
- Esso prevede:
 - una parte intestazione, di lunghezza fissa pari a 6 blocchi da 32 bit ciascuno (quindi 192 bit);
 - una parte dati, di lunghezza variabile.

Il datagramma TCP

- port di provenienza: port del mittente (16 bit);
- port di destinazione: port del destinatario (16 bit);
- numero sequenziale: posizione dei dati nello stream del mittente (32 bit);
- numero riscontro: numero del byte che il destinatario si aspetta di ricevere successivamente;
- hlen: lunghezza del segmento in multipli di 32 bit (4 bit). E' necessario perché il campo opzioni (vedi sotto) ha lunghezza variabile;

Il datagramma TCP

- riservato: per usi futuri (6 bit);
- codice: indica lo scopo ed il contenuto del segmento (6 bit):
 - URG: il campo puntatore urgente è valido;
 - ACK: il campo riscontro è valido;
 - PSH: il campo richiede un push;
 - RST: effettua il reset della connessione
 - SYN: sincronizza i numeri di sequenza;
 - FIN: il mittente ha raggiunto la fine del suo stream.

Il datagramma TCP

- finestra: dimensione del buffer dei dati (16 bit);
- checksum: controllo (16 bit);
- puntatore urgente: il pacchetto deve essere processato con urgenza - ad es. contiene ctrl-c ed i pacchetti residui non devono essere più inviati - (16 bit). Se il codice URG = 1, specifica la posizione della fine dei dati urgenti nella finestra;

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

45

Il datagramma TCP

- opzioni: per informazioni aggiuntive, come ad es. stabilire la dimensione massima del segmento - maximum segment size - (24 bit);
- riempimento: completa le opzioni fino a 32 bit (8 bit);

ed infine

- dati: i dati effettivamente trasmessi dal mittente al destinatario.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

46

Il datagramma TCP

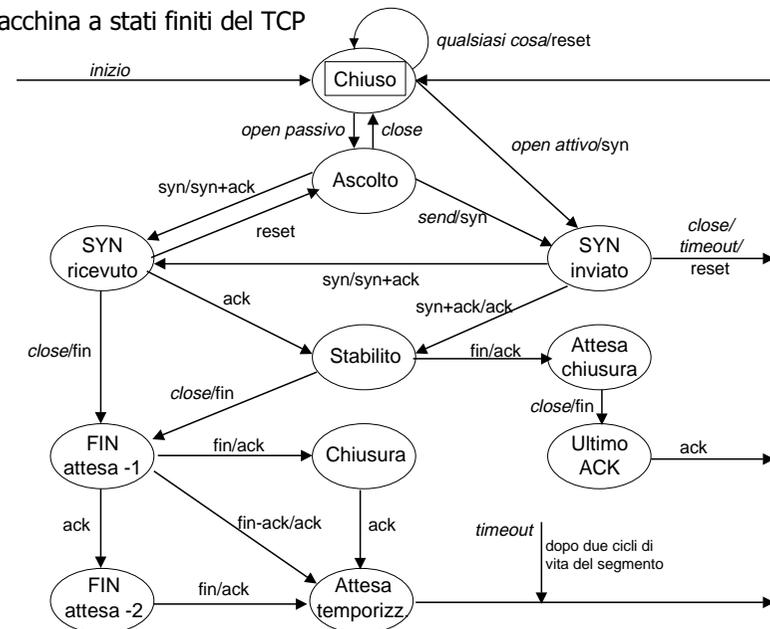
0		4		10		15		16		24		31	
port provenienza						port destinazione							
numero sequenziale													
numero di riscontro													
hlen		riservato		bit codice		finestra							
cheksum						puntatore urgente							
opzioni (eventuali)										riempimento			
lung. messaggio UDP						checksum UDP							
dati													

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

47

La macchina a stati finiti del TCP



8-May-01

Fondamenti di Informatica II - Programmazione di Rete

48

Introduzione alla programmazione di rete

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

49

Applicazioni distribuite

- Applicazione: un insieme di programmi coordinati per svolgere una data funzione.
- Un'applicazione è distribuita se prevede più programmi eseguiti (o processi) su differenti calcolatori connessi tramite una rete.
Es: Web Browser (Netscape, IE, Opera ...) e Web Server (server http)

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

50

Protocollo applicativo

- Le regole per la comunicazione in una applicazione distribuita sono dette *protocollo applicativo*.
Es. il protocollo applicativo della navigazione Web è detto HyperText Transfer Protocol - http.
- Il protocollo applicativo deve essere definito opportunamente e comune a tutti i programmi dell'applicazione.
Es. ogni messaggio scambiato è terminato dalla stringa "\0 \0 \0".

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

51

Interfacce e protocolli

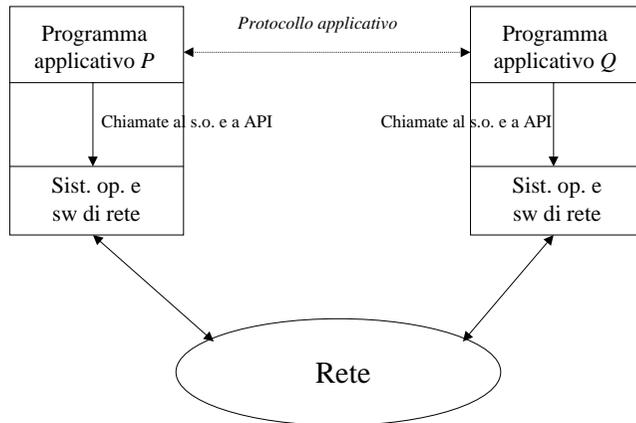
- I programmi applicativi utilizzano opportune interfacce (API - application program interface), fornite dal sistema operativo e dal software di rete, per comunicare:
 - le API forniscono il canale (o supporto) di comunicazione;
 - il protocollo applicativo rappresenta le regole di comunicazione, e considera il contenuto della comunicazione.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

52

Interfacce e protocolli



8-May-01

Fondamenti di Informatica II - Programmazione di Rete

53

L'interfaccia di socket

- L'API standard per TCP/IP (il protocollo di internet) si chiama "interfaccia di socket" ed è stata definita a Berkley agli albori di internet (circa 1980).
- Socket funziona anche per altri protocolli differenti da TCP/IP.
- L'interfaccia di socket è in linguaggio C ed è per calcolatori Unix: per calcolatori Windows si utilizza WinSocket.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

54

L'interfaccia di socket

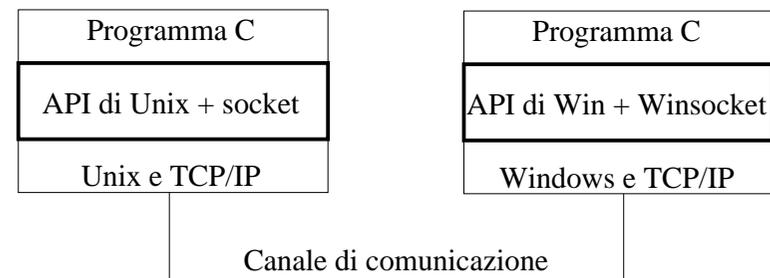
- Il socket è il descrittore di una connessione.
- TCP/IP garantisce l'interoperabilità tra calcolatori anche se equipaggiati con sistemi operativi differenti.
- Calcolatori Unix e calcolatori Windows possono comunicare utilizzando, rispettivamente, le interfacce Socket e WinSocket.

8-May-01

Fondamenti di Informatica II - Programmazione di Rete

55

Connessione tra calcolatori Unix e Windows



8-May-01

Fondamenti di Informatica II - Programmazione di Rete

56

Il modello Client/Server

- Il software di rete TCP/IP consente la comunicazione tra calcolatori in modalità peer-to-peer (pari-a-pari) purché condividano un protocollo comune.
- Spesso i processi che cooperano sono due:
 - client: effettua la richiesta di un servizio ad un calcolatore che svolge la funzione di server;
 - server: soddisfa la richiesta del client. Il server viene pertanto attivato **prima** del client.

Il modello C/S

- Generalmente un processo conserva lo stesso ruolo (è o sempre Client o sempre Server).
- Esempio:
- il Web client (Netscape, IE, Opera) richiede una data pagina ad un Web server;
 - il Web server rintraccia la pagina specificata e la fornisce a chi ne abbia fatto richiesta (client);
 - client e server utilizzano un protocollo comune (nel caso, http).

Indirizzamento

- Siano P e Q i due processi che devono comunicare. Q (client) deve richiedere un servizio a P (server).
- E' necessario che Q sappia come *raggiungere* P (per inviargli la richiesta) e che poi P sappia come raggiungere Q (per restituire quanto richiesto o segnalare il completamento dell'operazione richiesta dallo stesso Q).

Indirizzamento

- Q indirizza P specificando il calcolatore sul quale P è in esecuzione ed il numero di *port*, che identifica il processo P sul suo calcolatore.
- Il calcolatore è specificato tramite l'indirizzo IP (Internet Protocol):
 - l'indirizzo IP è composto da 4 byte (quindi 32 bit);
 - un esempio di indirizzo IP è 131.175.21.8

Indirizzamento

- Il port:
 - identifica un processo nell'ambito del calcolatore ai fini della comunicazione;
 - generalmente non coincide con il PID del processo (che è assegnato dal sistema operativo ad ogni processo: Q non avrebbe modo di conoscerlo e quindi non potrebbe indirizzare P);
 - alcuni port (compresi tra 0 e 1023) sono assegnati a servizi standard e non possono essere utilizzati per sviluppare propri server.

Indirizzamento

- L'indirizzamento completo è quindi specificato fornendo l'indirizzo IP della macchina ed il port.
 - Es. <131.175.21.8, 80> significa:
 - port 80 (di solito è il port del Web server)
 - della macchina 131.175.21.8

Il modello di comunicazione e la connessione

- La comunicazione TCP è connection-oriented (orientata alla connessione): la prima cosa da fare è stabilire una connessione, seguiranno poi lo scambio dei dati.
- E' necessario che P (server) rimanga in attesa di una richiesta di connessione da Q (client). Viene poi stabilita la connessione.

Il modello di comunicazione e la connessione

- Sequenza per la connessione:
 - P e Q si identificano come *punti terminali* della connessione;
 - ogni punto terminale è identificato dalla coppia <indirizzo IP, numero di port>;
 - ogni connessione è quindi identificata da 4 numeri: <indirizzo IP di P, numero di port di P>, <indirizzo IP di Q, numero di port di Q>;
 - dopo la connessione, il canale è bidirezionale, affidabile e orientato allo stream.

