
RETI DI CALCOLATORI

Corso A

Autore

Giuseppe Acocella

2025/26

<https://github.com/Peenguino>

Ultima Compilazione - December 3, 2025

Contents

1 Introduzione alle Reti	4
1.1 Tipi di Rete per Dimensione	5
1.2 Gestione di Internet come Rete di Reti	5
1.3 Reti e Mezzi d'Accesso	5
1.3.1 Schema dell'Accesso DSL	5
1.3.2 Tipi d'Accesso - Cavo, Aziendale, Domestico, Mobile e Satellitare	6
1.4 Nucleo della Rete - Commutazione Circuito vs Pacchetto	7
1.5 Metriche - Prestazioni della Rete (Latenza, Throughput, Prod. Rate/Ritardo)	8
1.6 Modello Stratificato	9
1.6.1 Stack Protocolare TCP/IP	9
1.6.2 Stack Protocolare ISO/OSI - Cenno Storico	10
1.7 Breve Introduzione alla Sicurezza di Reti	10
2 Livello Applicazione	11
2.1 Introduzione a Paradigmi, Protocolli e Tipi d'Utilizzo	11
2.2 TCP vs UDP nel Contesto del Livello Applicazione	12
2.2.1 API e Socket	12
2.3 Protocollo HTTP - HyperText Transfer Protocol	13
2.3.1 Versioni HTTP, Persistenza e Pipelining	14
2.3.2 Messaggi HTTP	15
2.3.3 Content Negotiation	16
2.3.4 Metodi dell'HTTP	16
2.3.5 Web Caching	17
2.3.6 Cookies	17
2.4 Protocollo Telnet	18
2.5 Protocollo SMTP - Simple Mail Transport Protocol	19
2.6 Protocolli POP3/IMAP	20
2.7 DNS - Domain Name System	20
2.7.1 Gerarchia dei Name Servers	21
2.7.2 Protocollo DNS	23
3 Livello Trasporto	25
3.1 Introduzione al Livello Trasporto	25
3.2 TCP vs UDP e Multiplexing/Demultiplexing	26
3.3 Protocollo UDP	27
3.4 Principi di Trasferimento Affidabile	28
3.4.1 RDT 1.0 - Trasferimento su Canale Affidabile	29
3.4.2 RDT 2.0 - Trasferimento su Canale con Errori su Bit	29
3.4.3 RDT 2.1 - Checksum per Controllo Corruzione ACK	30
3.4.4 RDT 2.2 - Protocollo NAK-free	31
3.4.5 RDT 3.0 - Timer in Attesa di ACK	31
3.4.6 Prestazioni e Reale Utilizzo del Canale	31
3.4.7 Approcci di Pipelining - Go-Back-N (GBN) e Selective Repeat (SR)	32
3.4.8 Riassunto Meccanismi Affidabili	34

3.5	Protocollo TCP	34
3.5.1	Formato Segmento TCP	35
3.5.2	Gestione della Connessione TCP - Handshaking e Chiusura	36
3.5.3	Trasferimento Affidabile del TCP	38
3.5.4	TCP RTT, RTO, Stime Timeout e Finestre di Ricezione	39
3.5.5	Controllo di Flusso del TCP	40
3.5.6	Controllo di Congestione del TCP	40
3.5.7	Tipologie di TCP che Istanzano Algoritmo di Controllo Congestione - (Reno, Tahoe, Cubic e Vegas)	41
3.5.8	Throughput, Equità e Cenni di Attacchi TCP	43
3.6	Protocollo QUIC e Approfondimento HTTP/2 HTTP/3	44
3.6.1	HTTP/1.1 e Head of Line Blocking	44
3.6.2	HTTP/2	44
3.6.3	HTTP/3 e QUIC	45
4	Livello Rete	46
4.1	Servizi e Modelli di Servizio - Forwarding/Routing	46
4.2	Protocollo IP	47
4.2.1	Funzioni IP	47
4.2.2	Meccanismi IP	48
4.2.3	Formato Datagramma IP	48
4.2.4	Frammentazione di Datagramma IP	49
4.2.5	Indirizzi IP	50
4.2.6	Classful/Classless Addressing e Subnet Mask	51
4.2.7	Assegnazione a Blocchi e Indirizzi Speciali	51
4.2.8	IP Forwarding - Diretto vs Indiretto	52
4.3	Protocollo DHCP - Dynamic Host Configuration Protocol	53
4.4	NAT - Network Address Translation	54
4.5	ICMP - Internet Control Message Protocol	54
4.6	Architettura di Router	56
4.6.1	Descrizione Fisica dell'Architettura di un Router	56
4.7	Protocolli di Routing	58
4.7.1	Algoritmo Distance Vector	58
4.7.2	Algoritmo Link-State	59
4.7.3	Confronto Distance Vector e Link State	60
4.8	Sistemi Autonomi AS ed Instradamento Gerarchico	61
4.8.1	Intra/Inter-AS	61
4.9	Protocolli Reali per Intra-AS	61
4.9.1	RIP - Routing Information Protocol	61
4.9.2	OSPF - Open Shortest Path First	62
4.10	Protocolli Reali per Inter-AS	63
4.10.1	BGP - Border Gateway Protocol	63
4.11	Protocollo IPv6	64

1 Introduzione alle Reti

Come facciamo a costruire una rete di computer che sia scalabile e che supporti **svariati tipi di applicazioni** come streaming, messaggistica e videochiamate? Internet ci permette di farlo, e tecnicamente è un insieme di miliardi di host (servers, laptops, smartphones) connessi tra loro.

Internet - Vista dei Componenti Questa connessione è permessa da specifici componenti:

1. **Link di Comunicazione:** Fibra ottica, rame, onde radio, microonde.
2. **Dispositivi di Interconnessione:** I più comuni sono:
 - (a) **Switch:** Dispositivo che collega tra loro **più host**.
 - (b) **Router:** Dispositivo che collega tra loro **più reti**.
3. **Reti:** Insieme di host, dispositivi di interconnessione e link gestiti da una stessa organizzazione.

Internet - Vista dei Servizi Da un punto di vista di servizi offerti invece Internet:

1. E' un **infrastruttura** che offre servizi alle **applicazioni distribuite**.
2. Offre un **interfaccia di programmazione** (socket) che permette alle applicazioni, distribuite su **host diversi**, di scambiarsi informazioni.

Internet - Vista delle Entità Software

1. **Applicazioni:** Elaborano e si scambiano informazioni.
2. **Protocolli:** Regolano la trasmissione di queste informazioni, definendo quindi il formato e l'ordine dei messaggi scambiati tra due o più entità in comunicazione. Due esempi noti di protocolli internet sono:
 - (a) **Transmission Control Protocol (TCP)**
 - (b) **Internet Protocol (IP)**

Standardizzazione di Protocolli I protocolli di Internet vengono standardizzati da diversi enti, come:

1. Internet Engineering Task Force (IETF). Si basano su una procedura detta RFC (Request for Comments) che seguendo passi rigorosi definisce gradualmente nuovi standard per Internet.
2. IEEE 802 LAN Standards.
3. World Wide Web Consortium (W3C): Comunità internazionale che sviluppa standard aperti per favorire lo sviluppo del Web.

1.1 Tipi di Rete per Dimensione

Descrizione dei tipi di rete per estensione fisica:

1. **Personal Area Network (PAN)**: Connessione personale, ad esempio una Bluetooth.
2. **Local Area Network (LAN)**: Reti circoscritte ad un area limitata, utilizzano cavi o mezzi wireless per connettersi e sono solitamente proprietarie.
3. **Metropolitan Area Network (MAN)**: Rete di calcolatori che non supera mai l'estensione fisica indicativa di una "città".
4. **Wide Area Network (WAN)**: Reti che cercano di interconnettere host separati da distanze geografiche tramite cavi in fibra ottica o ponti radio.

1.2 Gestione di Internet come Rete di Reti

Ogni rete locale si espone alle altre reti grazie agli **Internet Service Provider (ISP)**. Come possiamo immaginare però non connettiamo ogni punto terminale ad ogni altro punto terminale perché provocherebbe un alta complessità da un punto di vista di collegamenti. Di conseguenza saranno presenti dei **Global ISP** che si occupano della gestione del route della comunicazione. Non esiste un solo **Global ISP** ma **molteplici** e questi si connettono tra loro con dei **Peering Point** su punti d'incontro tra le reti detti **Internet Exchange Point (IXP)**.

Content Provider Network Idealmente si segue lo schema descritto sopra, ma a volte grandi aziende (Amazon, Google) si posizionano trasversalmente a questa organizzazione descritta, in modo tale da gestire la propria rete per avvicinare i propri servizi all'utente finale.

1.3 Reti e Mezzi d'Accesso

Descriviamo questi elementi caratterizzanti dell'accesso ad una rete:

1. **Rete d'Accesso**: Rete che connette fisicamente un sistema terminale al **edge router**, ossia il primo router da sorgente a destinazione.
2. **Mezzi d'Accesso**: Come passiamo fisicamente dal calcolatore al edge router:
 - (a) **Cablati (Vincolati)**: Doppino, coassiale, fibra.
 - (b) **Wireless (Non Vincolati)**: Onde Radio, microonde, infrarossi.

1.3.1 Schema dell'Accesso DSL

Uno dei primi schemi d'accesso è stato quello Digital Subscriber Line (DSL), dove la compagnia telefonica rappresentava l'ISP. Risultava quindi necessario splittare la linea telefonica già esistendo, suddividendo il traffico in bande di frequenze diverse:

1. **Canale Telefonico**: $0 - 4kHz$
2. **Upstream**: $4 - 50kHz$
3. **Downstream**: $50kHz - 1MHz$

In questo modo sarà necessario un multiplexer di accesso alla linea digitale (DSLAM) che permetta di separare i due tipi di comunicazione. Esistono **versioni diverse di questo schema**, riferiremo infatti allo schema **xDSL**, proprio perchè è possibile migliorare le caratteristiche della connessione ad esempio basandosi su una connessione in **fibra** tra DSLAM ed ISP, ad esempio:

1. **Fibre To The Cabinet (FTTC)**: Standard VDSL e VDSL2.
2. **Fibre To The Derivation Point**.
3. **Fibre To The House (FTTH)**: Caratterizzato dai seguenti elementi:
 - (a) Fibra ottica fino all'interno delle abitazioni.
 - (b) Fibra uscente dalla centrale locale con terminatore ottico di linea **OLT** viene condivisa da diverse abitazioni.
 - (c) Il terminatore ottico di rete **ONT** invece è connesso ad uno splitter tramite fibra dedicata, questo permette la gestione fino al centinaio di abitazioni.
 - (d) L'**OLT** si connette al router dell'ISP e tramite questo ad Internet.

1.3.2 Tipi d'Accesso - Cavo, Aziendale, Domestico, Mobile e Satellitare

Descriviamo queste tipologie di accesso alla rete:

1. **Accesso Via Cavo**: Il segnale ottico viene convertito in segnale elettromagnetico e inviato sulle linee di cavi coassiali per la distribuzione del segnale alle varie abitazioni. La velocità potrebbe degradare a causa della distanza o della congestione del canale condiviso tra più abitazioni.
2. **Accesso Aziendale**: Nelle aziende i sistemi sono collegati al router di frontiera con una LAN.
3. **Accesso Domestico**: Utilizzate tecnologie Ethernet o WiFi per creare LAN domestiche.
4. **Accesso Mobile**: Infrastruttura radio mobile definita dagli ISP, può subire degrado di prestazioni a causa di distanza oppure ostacoli.
5. **Accesso Satellitare**: Accesso permesso da trasmettitori terrestri che si connettono a satelliti geostazionari (GEO) oppure satelliti a bassa quota (LEO). Questo sistema risulta essere il più duttile e non richiede complesse installazioni, ma potrebbe risultare inferiore da alcuni punti di vista come la latenza più alta.

1.4 Nucleo della Rete - Comutazione Circuito vs Pacchetto

Come determiniamo il percorso effettuato da un messaggio? Come trasferiamo dalla porta di uscita a quella di ingresso di due calcolatori? A queste domande risponde il tipo di commutazione.

Commutazione di Circuito Si instaura un cammino tra gli host in comunicazione e tutte le risorse del canale sono dedicate a loro. Questo causa un utilizzo esclusivo del canale e conseguente spreco di risorse. Per poter permettere molteplici comunicazioni utilizzando questo tipo di commutazione esistono due politiche:

1. **Frequency Division Multiplexing (FDM)**: Divido il canale in bande di frequenze e le dedico alle diverse comunicazioni.
2. **Time Division Multiplexing (TDM)**: Dedico tutto il canale alla coppia di host ma per una finestra di tempo limitata.

Commutazione di Pacchetto Il flusso dati di una comunicazione viene suddiviso in **pacchetti**, e gli viene assegnato un **header** che mantenga dei metadati riguardanti il flusso dati originale. Questo tipo di instradamento segue queste fasi:

1. Suddivisione in pacchetti ed assegnamento header.
2. Instradamento del singolo **pacchetto indipendentemente** dagli altri. Pacchetti di flussi dati diversi possono condividere i canali di comunicazione.
3. Fase di **Store and Forward**, ossia prima che il commutatore (router) consegni i pacchetti dovrà prima aspettare che siano del tutto arrivati. Il tipo di Multiplexing quindi è statistico e non prefissato. Seguendo questa politica potremmo incorrere anche in problemi di **perdita di pacchetti** in base al tipo di gestione del buffer di ciascun commutatore. Non abbiamo quindi alcuna garanzia sulle prestazioni utilizzando questa tecnica.

Confronto Caratteristiche Comutazione Circuito/Pacchetto Elenchiamo pro e contro di ciascuna tecnica:

1. **Commutazione Circuito - Pro:**
 - (a) Prestazioni garantite.
 - (b) Tecnologie di switching efficienti.
 - (c) Tariffazioni per ISP semplici da definire.
2. **Commutazione Circuito - Contro:**
 - (a) Segnalazione e configurazione delle tabelle di switching.
 - (b) Poca ottimizzazione del uso di risorse.

3. Commutazione Pacchetto - Pro:

- (a) Utilizzo ottimale delle risorse.
- (b) Non richiede fase di setup e tabelle di switching predefinite.

4. Commutazione Pacchetto - Contro:

- (a) Tecnologie di inoltro non efficienti.
- (b) Ritardi variabili.
- (c) Rischio di congestione.

1.5 Metriche - Prestazioni della Rete (Latenza, Throughput, Prod. Rate/Ritardo)

Descriviamo due metriche fondamentali delle prestazioni della rete:

1. **Latenza:** Tempo richiesto dal primo bit partito dalla sorgente fino all'arrivo a destinazione. Questa definizione è ideale, infatti dipende anche dal tipo di commutazione, quella di pacchetto con Store & Bound introduce altre difficoltà. Nello specifico le cause specifiche di ritardo nella commutazione di pacchetto sono:

(a) **Ritardo di Elaborazione del Nodo**

- i. Controllo errori sui bit.
- ii. Determinazione porta di uscita.

(b) **Ritardo di Accodamento**

Componente di ritardo più complessa da stabilire, si basa su una **condizione di stabilità** ossia un rapporto $\rho = \frac{L_a}{R}$ dove a è velocità media di arrivo, R velocità di trasmissione ed L lunghezza media dei pacchetti in bit. La condizione $\rho < 1$ è detta **condizione di stabilità**.

- i. Attesa di trasmissione.
- ii. Dipende da intensità e tipo di traffico.

(c) **Ritardo di Trasmissione L/R**

Tempo impiegato dal router per trasmettere un pacchetto sul link, con $R(\text{bit/sec})$ indichiamo il rate, ossia la velocità di trasmissione del collegamento mentre con $L(\text{bit})$ la lunghezza del pacchetto. Il ritardo sarà quindi $d_{trasm} = \frac{L}{R}$. Quindi rappresenta il ritardo di passaggio nel commutatore.

(d) **Ritardo di Propagazione**

Quanto un bit ci mette per spostarsi fisicamente, rappresenta il ritardo di passaggio nel link.

Il ritardo complessivo è calcolato come somma di tutti i ritardi calcolati sopra.

2. **Throughput** Rappresenta la capacità effettiva di un link da host A verso host B , difficile da definire formalmente, ma diamo questi due riferimenti:

- (a) **Throughput Istantaneo:** Velocità in bit/sec a cui host B riceve in ogni istante.
- (b) **Throughput Medio:** Rapporto tra quantità L di dati trasferiti e il tempo T impiegato per il loro trasferimento, ovvero $\frac{L}{T}$.

Un potenziale bottleneck influenzerebbe pesantemente questa metrica, in caso quindi di più di un collegamento q_i il calcolo sarà $\min(q_1, \dots, q_n)$. Il Throughput quindi è differente dalla velocità di trasmissione del collegamento che si pone come upper bound alla metrica del Throughput.

3. **Prodotto Rate per Ritardo** Il prodotto rate per ritardo rappresenta il "volume" del collegamento, ossia il massimo numero di bit che possono riempire il collegamento ad un certo istante.

1.6 Modello Stratificato

Lo schema base di gestione delle reti è di tipo **stratificato**, ossia ciascun livello ha le proprie caratteristiche e responsabilità seguendo i due criteri di:

1. **Separation of Concern:** Divisione di interessi e responsabilità tra i livelli.
2. **Information Hiding:** Si espone solo l'interfaccia di ciascun livello ma non la sua implementazione.

Quindi ogni livello fornisce **servizi** al livello superiore e scambia informazioni con i livelli adiacenti tramite **interfacce**. In questo caso anche in caso di rifattorizzazione di un livello gli altri livelli sarebbero indipendenti dalle modifiche apportate a meno di modifica dell'interfaccia. Bisogna quindi definire dei protocolli che guidino la gestione di questi livelli.

Cosa è un Protocollo In protocollo si specificano **sintassi** di un messaggio, la **semantica** di un messaggio e **azioni** da intraprendere in caso di ricezione di messaggio.

1.6.1 Stack Protocollare TCP/IP

Si compone di 5 livelli (ciascun livello sarà guidato a sua volta da protocolli):

1. **Livello Applicazione:** Rappresenta delle applicazioni di rete, oppure collegamenti logici e scambi di messaggi end-to-end tra due processi su host diversi.
 - (a) **Protocolli:** **ftp**, **smtp**, **http**.
 - (b) **Entità Informazione del Livello:** L'informazione a questo livello è detta **messaggio**.

2. **Livello Trasporto:** Trasferimento dati end-to-end tra processi su host remoti.
 - (a) **Protocolli:** **tcp**, **udp**.
 - (b) **Entità Informazione del Livello:** L'informazione a questo livello è detta:
 - i. **Segmento** se in **Protocollo TCP**.
 - ii. **Datagramma Utente** se in **Protocollo UDP**.
3. **Livello Rete:** Trasferimento e instradamento di **datagrammi** dalla sorgente alla destinazione.
 - (a) **Protocolli:** **IP** o **ICMP**.
 - (b) **Entità Informazione del Livello:** L'informazione a questo livello è detta **data-gramma**.
4. **Livello Link:** Trasferimento dati in **frame** attraverso il collegamento tra elementi di rete vicini.
 - (a) **Protocolli:** **PPP** o **Ethernet**.
 - (b) **Entità Informazione del Livello:** L'informazione a questo livello è detta **frame**.
5. **Livello Fisico:** Trasferimento dei bit di un frame sul mezzo trasmissivo.

1.6.2 Stack Protocollare ISO/OSI - Cenno Storico

A differenza dello Stack TCP/IP, che era stato definito dall'IETF, l'ISO/OSI voleva essere uno standard effettivo e formale, definito invece dall'ISO. Si differenziava dall'TCP/IP perchè aggiungeva due livelli ossia Presentazione e Sessione (compresi nel livello Applicazione nel TCP/IP). Non dimostrò nessun pratico vantaggio al TCP/IP quindi viene solo usato come riferimento formale.

1.7 Breve Introduzione alla Sicurezza di Reti

Un software malevolo può penetrare in un host e può rendere questo host parte di una botnet che distribuisce richieste spam. Esistono varie tecniche:

1. **DDoS** (Distributed Denial of Service):
 - (a) **Bandwidth Flooding:** Inondazione di pacchetti
 - (b) **Connection Flooding:** Inondazione di richieste TCP
2. **Sniffing:** Man in the middle host C collegato ad un mezzo condiviso può ottenere copia di ogni pacchetto trasmesso tra due host A e B.
3. **Spoofing:** Man in the middle host C che immette pacchetti spacciandosi per un altro utente tra due host A e B.

Parametri e Criteri di Sicurezza Esistono dei metodi per difendersi da potenziali attacchi:

1. **Autenticazione:** Dimostrazione della propria identità.
2. **Crittografia:** Garantisce confidenzialità delle informazioni.
3. **Integrità ed Autenticità:** Assicurare che i dati non siano stati alterati e provengano da fonti affidabili.
4. **Controllo Accessi:** Limitare l'uso di risorse in base ai ruoli.
5. **Utilizzo di Firewall:** Filtrare il traffico tramite dispositivi hardware o software dedicato.

2 Livello Applicazione

2.1 Introduzione a Paradigmi, Protocolli e Tipi d'Utilizzo

Livello che definisce programmi applicativi che comunicano tra i vari host. Si seguono in genere due paradigmi di comunicazione tra host in questo livello:

1. **Client/Server:** Host server sempre attivo che attende richieste dei vari host client.
2. **Peer to Peer:** Tutti i client alla pari, spesso utilizzato nella condivisione di file.

Protocolli La comunicazione si basa su protocolli aperti (HTTP, SMTP) oppure proprietari. Una web app ad esempio si basa sul protocollo HTTP, invece un client di posta elettronica può utilizzare il protocollo SMTP.

Socket e API Ogni processo, essendo che comunicherà con **processi di altre macchine**, dovrà essere **identificato dall'IP della macchina** (da 32 bit) e dalla **porta dedicata al processo** (da 16 bit). Questo avviene grazie all'astrazione dei socket, che utilizzando IP e porta permettono ai processi di scrivere/leggere da ciascun lato.

Caratterizzazione di un Applicazione Ciascuna applicazione dovrà scegliere un protocollo da utilizzare in base alla sua caratterizzazione ed in base a specifici criteri:

1. **Throughput:** Tasso effettivo di trasferimento dati, ad esempio un client di mail non avrà bisogno di un alto Throughput.
2. **Integrità dei Dati:** Richiesta d'integrità del flusso dei dati, ad esempio una stream potrebbe non richiedere un integrità assoluta.
3. **Sensibilità ai Ritardi:** Richiesta di una bassa latenza, ad esempio un videogioco multiplayer potrebbe richiedere una bassa latenza.

L'analisi di questi criteri in relazione all'applicazione permette la scelta del protocollo di trasporto (livello inferiore), quindi nel nostro caso se TCP (orientato allo stream) o UDP (orientato al messaggio).

2.2 TCP vs UDP nel Contesto del Livello Applicazione

Si sceglie un protocollo del livello trasporto in base alle caratteristiche dell'applicazione in questione. Descriviamone le caratteristiche:

1. TCP

- (a) Orientato allo stream.
- (b) Necessario un primo setup tra client e server, questo ha un **costo** in termini di **latenza**.
- (c) Utilizzo di un buffer limitato d'appoggio per lo stream.
- (d) Nessuna garanzia su proprietà di **Timing** e **Throughput**.
- (e) Attualmente utilizzato in diversi contesti.

2. UDP

- (a) Orientato al messaggio.
- (b) Non necessario alcun setup.
- (c) Nessuna garanzia su proprietà di **Timing**, **Throughput**, **Trasferimento Dati**, **Controllo Flusso**, **Aampiezza Minima di Banda**.
- (d) Si accettano potenziali perdite di pacchetti.

2.2.1 API e Socket

API - Application Programming Interface Un API è un insieme di regole che un programmatore deve rispettare per utilizzare delle risorse. Se un processo a livello applicazione vuole comunicare un messaggio ad un altro host allo stesso livello allora dovrà interagire con il sistema operativo che implementa i sottostanti livelli dello stack TCP/IP attraverso un interfaccia.

Interfaccia Socket API che viene utilizzata da interfaccia tra gli strati di applicazione e trasporto, è l'API di Internet per eccellenza. Dato che un host abbiamo detto essere individuato grazie ad un IP da 32 bit, identificheremo un processo e la relativa socket con la porta da 16 bit.

Per fare un esempio di API basato su uno dei protocolli del livello trasporto citati prima, mostriremo una bozza di setup tra client e server in TCP e qualche pseudocomando:

1. Client:

```
1 //Apertura Connessione  
2 connection TCPopen(IPaddress, int)  
3
```

2. Server:

```
1 //Binding
2 int tcpbind(int port)
3 //Attesa Richieste
4 connection TCPaccept(int)
5 //Spedire Dati
6 void TCPsend(connection, data)
7 //Ricevere Dati
8 data TCPreceive(connection)
9 //Chiusura Connessione:
10 void TCPclose(connection)
11
```

Meccanismi di Sicurezza in TCP/UDP Nativamente questi livelli non implementano alcun meccanismo di sicurezza, dunque è stato aggiunto un layer con un socket TLS (Transport Layer Security) che garantisce proprietà come cifratura o integrità dei dati.

2.3 Protocollo HTTP - HyperText Transfer Protocol

Protocollo del livello Applicazione che risulta essere **generico, stateless e basato sulla tipizzazione** che non deve per forza essere HTML, si negozia infatti il tipo di contenuto.

Inizialmente nato per le pagine HTML, attualmente utilizzato anche per la comunicazione tra applicazione tramite API REST che permettono esposizione di servizi.

Dati un client ed un server, il client tramite un **browser** ottiene una **pagina HTML** e tante **risorse dereferenziabili** con **URL**.

URI - Uniform Resource Identifier Permette l'identificazione di una risorsa:

1. **URN - Uniform Resource Name:** Sottoinsieme dell'URI che deve identificare univocamente una risorsa anche quando questa cessa di esistere.
2. **URL - Uniform Resource Locator:** Sottoinsieme dell'URI che identifica una risorsa attraverso il suo meccanismo d'accesso.

Sintassi di un URL Un URL segue questo schema per essere definito:

```
< scheme > // < user > : < password > @ < host > : < port > / < path >
```

1. `< user > : < password >` è attualmente deprecata.
2. `< port >` è spesso omessa, dato che se non definita viene utilizzata la porta standard 80.
3. `< path >` segue lo stesso schema di un file system, dato che storicamente i server web erano statici e l'accesso ad una risorsa avveniva tramite il suo path.

Istanziando al protocollo HTTP come `< scheme >` otteniamo:

```
http // < host > : < port > / < path > ? < query >
```

Client/Server HTTP Protocollo basato su richiesta/risposta tra client e server. La **risposta** è **stateless**, ossia non dipende dalla richiesta effettuata. Il protocollo su cui si basa, del livello trasporto, è il TCP, dato che risulta necessario prima **setuppare** una connessione **tra server e client**.

2.3.1 Versioni HTTP, Persistenza e Pipelining

HTTP/1.0 - Connessione Non Persistente Nella versione HTTP 1.0 la connessione non era persistente, e per ciascuna risorsa richiesta veniva aperta una nuova connessione TCP tra client e server. Questo comportava un overhead significativo, poiché ogni richiesta (ad esempio per immagini, fogli di stile, script) necessitava di una nuova connessione, aumentando la latenza e il carico sul server.

HTTP/1.1 - Connessione Persistente e Pipelining Con HTTP/1.1 è stata introdotta la **connessione persistente**: una singola connessione TCP può essere riutilizzata per più richieste e risposte tra client e server, riducendo l'overhead di apertura e chiusura delle connessioni. Inoltre, HTTP/1.1 supporta il **pipelining**, cioè la possibilità di inviare più richieste bufferizzando tramite una pipe di richieste. Tuttavia, il pipelining non è stato ampiamente adottato a causa di problemi di gestione dell'ordine delle risposte e di compatibilità.

2.3.2 Messaggi HTTP

In questo protocollo il messaggio può essere o una **richiesta** o una **risposta**. E' facilmente riconoscibile il tipo di messaggio già dalla sua **startline**, ossia la prima linea del messaggio.

1. **HTTP Request:** Composta da due parti:

- (a) **Request Line:**

METODO URI VERSIONE_HTTP

ad esempio:

GET /dump/page.html HTTP/1.1

- (b) **Header:** Insieme di coppie (nomi-valori), che mantengono informazioni riguardanti formati preferiti, charset o encodings, ad esempio:

Accept: text/plain; text/html, image/png, application/json

Accept-Charset iso-8859-5, unicode-1-1

Durante questa fase è possibile che si instauri una negoziazione tra client e server sulla versione di HTTP da utilizzare. Dato che la comunicazione partirà sempre da una richiesta lato client, in base al confronto tra le versioni dell'HTTP in utilizzo di client e server si stabilisce quale usare.

Un messaggio di HTTP Request deve avere necessariamente una request line.

Un messaggio HTTP Response deve necessariamente avere una status line.

2. **HTTP Response:**

- (a) **Status Line:**

VERSIONE_HTTP STATUS_CODE FRASE

ad esempio:

HTTP/1.1 200 OK

dove lo status code segue questo criterio:

1xx	Informational
2xx	Success
3xx	Redirection
4xx	Client Error
5xx	Server Error

- (b) **Response Headers:** Informazioni relative alla risposta che non possono essere inserite nella Status Line, come ad esempio la **location della risorsa** oppure il **software** utilizzato dal **server** per handlare la richiesta.

(c) **Entity Headers:** Mantiene informazioni come:

Content-Base	URI Assoluta per dereferenziare il body
Content-Encoding	Codifica del body
Content-Language	Lingua del body
Content-Type	Tipo del body
Expires	Validità Temporale, usato nel Caching
Last Modified	Ultima modifica della risorsa, usato nel Caching

2.3.3 Content Negotiation

Il client conosce l'URL più generico della risorsa a cui vuole riferirsi. In base ai parametri forniti nell'header della richiesta sarà il server a dereferenziare l'URL più specifico, magari in base alle informazioni sulla lingua, il tipo di compressione o il formato.

2.3.4 Metodi dell'HTTP

Descriviamo le caratteristiche di ciascun metodo messo a disposizione da questo protocollo:

1. **OPTION:** Permette di listare quali siano i metodi chiamabili su una risorsa.

(a) **Sintassi:** `OPTION [RISORSA]`

2. **GET:** Permette la richiesta di una risorsa, anche in maniera **parziale** o **condizionale**.

(a) **Sintassi:** `GET [RISORSA]`

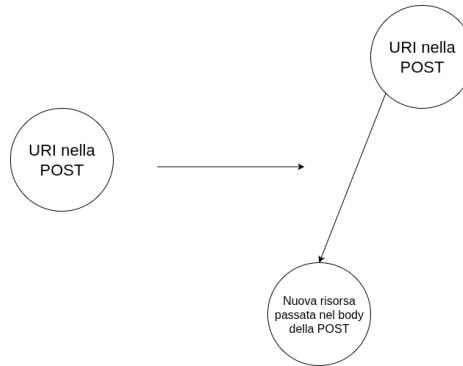
Un comune utilizzo di una GET condizionale è quella basata sull'*If-Modified-Since*, che permette di ottenere la risorsa solo se questa è stata modificata.

3. **HEAD:** Permette la richiesta solo della *response line* e dell'*header* di una risorsa, senza il suo *body*.

(a) **Sintassi:** `HEAD [RISORSA]`

4. **POST:** Permette che il server accetti l'*entità* passata nel *body* della richiesta come **nuova subordinata** della risorsa identificata dall'*URI*.

(a) **Sintassi:** `POST [RISORSA]`



Sarà quindi responsabilità del server trovare un nuovo URI per la nuova risorsa.

5. **PUT/DELETE**: Permette di inserire (PUT) o cancellare (DELETE) una risorsa all'URI passato al metodo.

(a) **Sintassi:** PUT/DELETE [RISORSA]

In questo caso invece il client dovrà conoscere a priori l'URI della nuova risorsa che sta inserendo.

Metodi Safe ed Idempotenti Definiamo queste due proprietà e listiamo i metodi che le rispettano:

1. **Metodi Safe**: Non hanno effetti collaterali, dunque non modificano la risorsa, quindi i metodi: GET, HEAD, OPTION e TRACE.
2. **Metodi Idempotenti**: Hanno lo stesso effetto se eseguiti una o n volte sulla stessa risorsa, quindi i metodi: GET, HEAD, PUT, DELETE, OPTION, TRACE.

2.3.5 Web Caching

Immaginando di avere una rete locale, assumendo di avere $\{host_1, \dots, host_k\}$ nella rete, cosa succede se $host_1$ richiede una risorsa su un host remoto $host_x$ e lo stesso viene fatto da $host_2$ subito dopo? Questo procedimento attualmente non viene per nulla ottimizzato, di conseguenza sono state formulate politiche di caching che permettono, tramite l'utilizzo di host intermedi, di mantenere le informazioni riguardanti le richieste per non ripeterle.

1. **Server Proxy**: Un approccio di **Web Caching** infatti è quello del **Server Proxy**, che si pone tra gli host della rete locale ipotizzata ed il server remoto. Se il server proxy cachasse infatti le richieste allora il secondo host potrebbe acquisire quelle informazioni senza effettuare richieste al server remoto ma solo richiedendole al server proxy stesso.
2. **User Agent Cache**: Un metodo alternativo è quello che si basa sul mantenimento in locale da parte del browser delle risorse visitate da un utente.

2.3.6 Cookies

Il protocollo HTTP è per definizione stateless, di conseguenza risulta necessario un meccanismo che permetta di mantenere uno stato utente riguardante la sua navigazione su una applicazione web. Vengono quindi utilizzati per autenticazione, ricordare il profilo utente e creare sessioni su un protocollo stateless.

Si vuole quindi avere la possibilità di numerare e far riconoscere gli utenti sul sito presentando ogni volta un cookie. Il meccanismo nello specifico si basa su questi passaggi:

1. **Riga di intestazione nel messaggio di risposta HTTP**.
2. **Riga di intestazione nel messaggio di richiesta HTTP**.
3. **Cookie file** mantenuto dal **browser** dell'utente.
4. **Database** di gestione **cookie** del **server web**.

Cookie First/Third Party Esistono due tipologie diverse di cookie:

1. **First Party Cookie:** Creati direttamente dal sito web che l'utente sta visitando per mantenere informazioni riguardanti lo stato.
2. **Third Party Cookie:** Memorizzano informazioni sensibili riguardanti gli utenti, sono creati e memorizzati da un dominio diverso rispetto a quello del sito web visitato.

2.4 Protocollo Telnet

Protocollo di terminale il cui scopo è quello di permettere l'uso interattivo di macchine remote. Chiaramente i due calcolatori devono poter comunicare a prescindere dai sistemi operativi che utilizzano, di conseguenza si utilizza un **interfaccia minima a caratteri**.

Protocollo TELNET [RFC 854] Questo modello standardizzato si basa su un programma client ed uno server, nello specifico si seguono questi passi:

1. Il client stabilisce una **connessione TCP** con il server.
2. Il **client** acquisisce l'**input dei tasti** e lo invia al server, successivamente **accetta i caratteri** che il **server manda indietro** e li visualizza sul terminale utente.
3. Il server accetta la connessione TCP e trasmette i dati al sistema operativo locale.

Il client si connette alla porta 23 del server e la connessione TCP persiste per la durata della sessione di login.

NVT - Network Virtual Terminal L'obiettivo di TELNET è quello di poter operare con quanti più calcolatori diversi e di conseguenza sistemi operativi eterogenei. Di conseguenza si definisce un **terminale virtuale** che permette la conversione dei comandi dei vari terminali locali e viceversa, creando una vera e propria interfaccia comune di riferimento. Si definiscono quindi dei set di caratteri universali per trasformare quelli delle macchine locali in comunicazione tra loro.

SSH - Secure Shell Lo standard attuale per comunicazioni tra macchine remote via terminale è SSH, ossia un architettura di protocolli nata per sostituire TELNET ed i suoi problemi di sicurezza. Uno dei servizi principali di SSH è quello di tunneling, infatti grazie al port forwarding possiamo mettere in comunicazione due processi su due macchine diverse mappando le loro porte a quella di SSH, lasciando quindi viaggiare la comunicazione tra i due processi tramite il tunnel SSH.

2.5 Protocollo SMTP - Simple Mail Transport Protocol

Uno dei protocolli applicativi più antichi, si basa su due "protagonisti":

1. **User Agent:** Client dell'user che permette l'editing e l'invio di messaggi di posta.
2. **Mail Server:** Server dei provider che devono archiviare i messaggi ricevuti, dato che il tipo di comunicazione tra i vari User Agent è asincrona, di conseguenza dovranno immagazzinare in una coda i messaggi in uscita.

Indirizzo Mail - Sintassi Solitamente un indirizzo è indicato con:

local-part @ domain-name

dove con *local-part* si indica la specifica "cassetta" dell'user nel mail server, mentre *domain-name* indica l'effettivo mail server.

Spooling dei Mail Server I Mail Server seguono una tecnica di spooling, ossia tentano di inviare i messaggi di posta settando una connessione TCP con la macchina destinazione. Se la connessione viene aperta e il trasferimento va a buon fine allora il Mail Server cancella la copia della mail in locale. Altrimenti, se qualcosa dovesse andare storto ritenterebbe la connessione TCP in maniera periodica fino ad un certo intervallo di tempo, se viene superato allora si scarta il messaggio, notificando il mittente della condizione.

Gestione Alias dei Mail Server Gestione a puntatori di indirizzi di casella postale, ne esistono di due tipi:

1. **Uno a Molti:** Un singolo indirizzo alias riferisce a più indirizzi mail reali.
2. **Molti a Uno:** Un singolo indirizzo mail reale ha molti indirizzi alias che lo riferiscono.

Protocollo SMTP [RFC5321] Si stabilisce una connessione TCP sulla porta 25, si segue una politica PUSH da parte di ogni client fino a quello precedente al destinatario. Quindi dal primo user fino all'ultimo Mail Server tutti agiranno da client nei confronti del server destinatario successivo. Tre fasi:

1. Handshaking a livello applicativo.
2. Trasferimento messaggi in coda.
3. Chiusura.

Comandi SMTP Questo standard definisce specifici comandi:

1. HELO < client identifier >
2. MAIL FROM: < reverse-path > < CRLF >
3. RCPT TO: < forward-path > < CRLF >
4. DATA
5. QUIT

Formato Messaggi SMTP I comandi sono formato ASCII, mentre le risposte sono ASCII a 7 bit. Nello specifico, il messaggio è così composto:

1. **Header:** Contiene le linee d'intestazione, come ad esempio *To*, *From*, *Subject*, che non corrispondono ai parametri passati ai comandi SMTP citati prima.
2. **Body:** Contiene il messaggio effettivo, composto da caratteri ASCII a 7 bit.

MIME - Multipurpose Internet Mail Extension Fino ad ora i body erano composti solo da caratteri ASCII a 7 bit. Si definisce quindi una codifica per trasferimento media e testo non-ASCII. Si aggiungono quindi dei parametri nelle intestazioni per dichiarare l'utilizzo del MIME nel body.

2.6 Protocolli POP3/IMAP

Fino ad ora abbiamo descritto il protocollo SMTP che pusha l'informazione fino all'ultimo mail server. Essendo però un tipo di comunicazione asincrona, sarà il client del destinatario a pullare il messaggio. Si definiscono quindi i protocolli POP e IMAP.

1. **POP3:** Il client apre una connessione TCP sulla porta 110, si autentica con utente e password e richiede l'elenco dei messaggi e li preleva uno alla volta.
2. **IMAP:** Più feature di POP3, permette la gestione logica in folder dei messaggi memorizzati sul server e permette di estrarre solo alcuni componenti dei messaggi.
3. **HTTP:** In alternativa, anche HTTP può effettuare richieste ai Mail Server tramite browser.

2.7 DNS - Domain Name System

Vogliamo avere la possibilità di dereferenziare nomi e non indirizzi numerici, dato che questi ultimi risultano essere più difficili da ricordare. La domanda quindi è come **associare indirizzi IP a nomi**. Inizialmente la gestione di queste associazioni era statica, ma attualmente risulta impossibile utilizzare questa politica, date le dimensioni di Internet.

Architettura Centralizzata vs Distribuita Potremmo immaginare un unico server con relativo unico database a cui riferirsi per risolvere queste associazioni. Questo però ci vincolerebbe ad una **pesante dipendenza** da quest'ultimo. Questo dovrebbe gestire miliardi di query al giorno, ed in caso di fallimento del server si bloccherebbe tutto Internet. Si preferisce quindi un architettura distribuita.

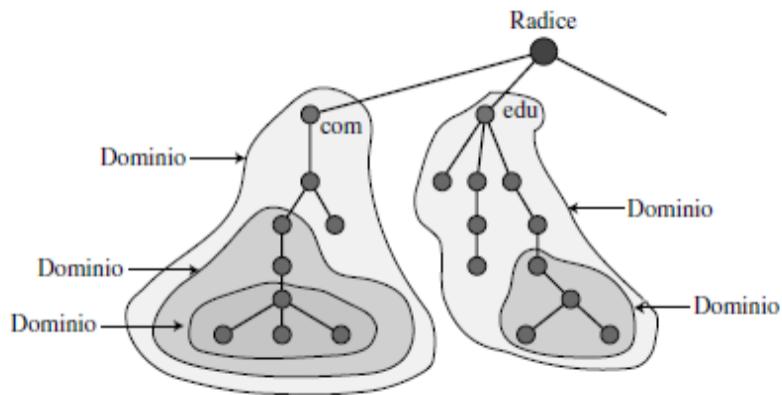
Composizione Essenziale del DNS Si compone di tre elementi fondamentali:

1. Uno **schema di assegnazione dei nomi** gerarchico basato sui domini.
2. Un **database distribuito** contenente nomi e corrispondenti IP.
3. Un **protocollo** per la **distribuzione delle informazioni** tra **name servers**:
 - (a) **Name Servers:** Comunicano per risolvere nomi.
 - (b) **Porta:** Utilizzano porta 53 in UDP (sono a volte TCP).

Servizi Offerti dal DNS Elenchiamo i servizi offerti dal DNS:

1. Risoluzione nomi (hostnames) in indirizzi IP.
2. **Host Aliasing:** Più alias possono dereferenziare lo stesso hostname, e di conseguenza, lo stesso indirizzo IP.
3. **Distribuzione Carico:** In caso di potenziale carico elevato assegnato ad un server, si fa corrispondere un singolo hostname a più IP per evitare problemi di sovraccarico su una macchina specifica.

Struttura Gerarchica di Gestione Domini Lo spazio dei nomi è definito da una **struttura gerarchica ad albero** dove ogni nodo è identificato da un **etichetta**, ossia un **nome di dominio**.

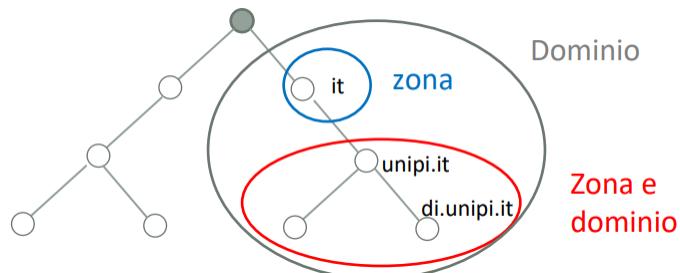


I domini top-level (.com, .edu, .gov, .net, .org) vengono gestiti dallo IANA.

2.7.1 Gerarchia dei Name Servers

Avendo definito il **DNS** come un **database distribuito** in una **gerarchia** di più **name servers**, allora bisogna definire cosa sia un **name server**, ossia un programma che gestisce la conversione da nome di dominio ad indirizzo IP.

Zone e Domini di Name Servers Tutte le informazioni per le risoluzioni di nomi sono ripartite su più name server, dunque ciascuno di essi immagazzina le info relative alla propria **zona** ed i riferimenti ai **name servers** dei **domini di livello inferiore**.



Ruoli dei Server nella Gerarchia Elenchiamo tutti i ruoli e le loro responsabilità:

1. **Root Name Server**: Responsabile dei record della zona radice, riconosce tutti i domini TLD (Top Level Domain) e conosce il server TLD che li risolve, quindi restituisce le info sui name server TLD. A questo ruolo corrispondono circa 13 indirizzi IP.
2. **Top Level Domain Server**: Mantiene le info dei nomi di dominio che appartengono ad un certo TLD, e restituisce informazioni sui name server di competenza dei sottodomini.
3. **Authoritative Name Server**: Responsabile di una certa **zona**, memorizza nome ed indirizzo IP di un insieme di host di cui può effettuare la traduzione. Questo ruolo può essere suddiviso in altri due sottoruoli:
 - (a) **Primary Authoritative Name Server**: Mantengono il file di zona.
 - (b) **Secondary Authoritative Name Server**: Ricevono il file di zona e offrono il servizio di risoluzione.

A questa gerarchia si aggiunge un **ulteriore ruolo**, che **non appartiene** propriamente alla **gerarchia**, ma che si occupa da prima interfaccia nel caso in cui un client (es. Browser) tenti di riferire ad un nome:

1. **Local Name Server**: Un server appartenente ad un ISP (Internet Service Provider) che assume il ruolo di proxy, per fare prima un controllo nelle tabelle locali per cercare la risoluzione del nome richiesto dal client, e solo successivamente, in caso di fault, richiede la query effettiva al DNS, nello specifico ad un Root Name Server.

Query Ricorsiva/Iterativa Esistono due tipi diversi di risoluzione dei nomi:

1. **Query Ricorsiva**: Il Local Name Server richiede una conversione completa di nome.
2. **Query Iterativa**: Viene restituito al client o al Local Name Server il riferimento al server successivo da contattare per completare la risoluzione di nome.

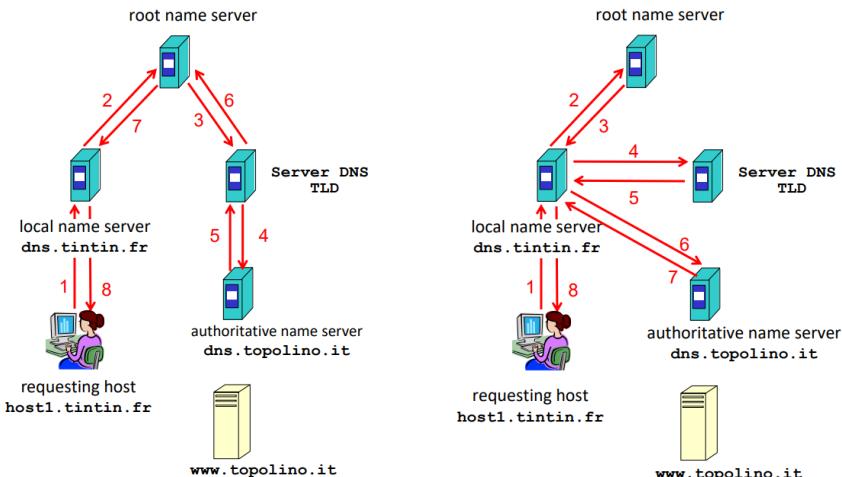


Figure 1: Query Ricorsiva

Figure 2: Query Iterativa

Caching DNS Nel DNS si utilizza molto il caching per ridurre richieste che rimbalzano continuamente tra server. Solitamente quindi si mantiene in cache di server DNS il dato in cache in modo non autorevole, e i server invalideranno queste informazioni dopo un periodo fissato. Un altro approccio potrebbe essere quello di mantenere in cache direttamente degli indirizzi IP di server DNS TLD (Top Level Domain), evitando di riferirsi più volte ai server radice.

Record DNS Un DNS è un database distribuito di *resource records* (RR) che seguono la seguente sintassi:

RR format : (name, value, type, ttl)

dove gli elementi vengono così interpretati:

1. **TTL**: Tempo residuo del record in cache.
2. **Name e Value**: Dipendono da type.
3. **Type**: Esistono vari tipi:

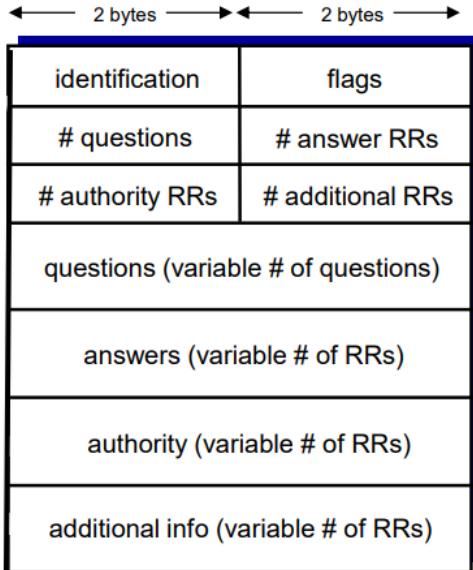
Type	Name	Value
A	hostname	indirizzo IP
CNAME	hostname (sinonimo)	nome canonico dell'host
NS	nome di dominio	hostname dell'Authoritative Name Server per quel dominio
MX	nome di dominio	nome canonico del server di posta associato a name

2.7.2 Protocollo DNS

L'effettivo protocollo del livello applicazione che si occupa di questa risoluzione di nomi utilizza prevalentemente **UDP** sulla **porta 53**. L'applicazione è detta **resolver**, effettua polling al server DNS e dopo n richieste o richiede ad un altro server oppure informa l'applicazione con il fallimento.

Perchè UDP e Quando TCP Viene utilizzato il protocollo UDP per una questione di riduzione di ritardi, riduzione di overhead causato dal setup di TCP e perchè solitamente le query DNS richiedono pochi dati, quindi rispettano le dimensioni di un pacchetto UDP. A volte viene utilizzato il protocollo TCP, ma questo accade solo se è necessaria l'affidabilità del trasferimento oppure se si sforano i 512 byte dei pacchetti UDP.

Messaggi del Protocollo DNS I messaggi di query e reply seguono lo stesso formato:



Dove ogni campo indica

1. **Identification:** Campo di 16 bit per l'identificazione di una query, la risposta ad una query usa lo stesso identification id.
2. **Flags:** Indicano se query o reply, se desiderata o disponibile la ricorsione e se la risposta è autorevole.
3. **Questions:** Nome richiesto e tipo di query.
4. **Answers:** RR (*resource records*) in risposta alla query.
5. **Authority:** Record per i server di competenza.
6. **Additional Info:** Informazioni utili aggiuntive.

DNS e Sicurezza Dato il funzionamento molto decentralizzato del DNS, esistono vari attacchi che sfruttano questa sua caratteristica:

1. **DNS Cache Poisoning:** Un attaccante invia risposte DNS falsificate ad un DNS Server o ad un client, avvelenandone la cache.
2. **Man-in-the-Middle Attack:** Vengono intercettate le comunicazioni tra un client ed un server DNS, ed in questo modo viene indirizzato un obiettivo verso host malevoli.
3. **DNS Hijacking:** Si restituiscono risposte non corrette alle query DNS, reindirizzando il client verso siti malevoli alterando le impostazioni DNS. Questo può essere fatto a diversi livelli, infatti può essere un Local, Router oppure Rogue Hijacking.

3 Livello Trasporto

L'obiettivo di questo livello è quello di creare una connessione logica tra processi applicativi su host diversi, assumono quindi di essere collegati a livello fisico, astraendo dalla reale connessione fisica di cui si occuperanno i livelli sottostanti. I protocolli di trasporto vengono eseguiti nei sistemi terminali, mantenendo la politica per cui si mantengono dumb i dispositivi intermedi e smart quelli terminali.

3.1 Introduzione al Livello Trasporto

In questo sottocapitolo definiamo concetti base del livello trasporto.

Servizi del Livello Trasporto Elenchiamo i servizi offerti da questo livello:

1. Offre servizi al livello applicazione, permettendo alle applicazioni di scegliere il modo in cui inviare/ricevere dati secondo uno **stile** a **messaggi** oppure a **stream**.
2. Comunica con il livello Rete sottostante, che si occupa a sua volta della consegna del datagramma all'host destinatario.

Servizio Connection/Non Connection Based Seguendo gli stili definiti sopra possiamo definire anche servizi orientati o meno alla connessione. Nello stile orientato alla connessione si instaura una logica di server/client dopo una fase di setup, mentre in uno stile non orientato alla connessione i dati vengono visti come messaggi indipendenti.

Azioni del Livello Trasporto Elenchiamo le azioni di questo livello in base a chi le esegue:

1. **Azioni Mittente:**

- (a) Riceve un messaggio dal livello applicazione.
- (b) Determina i valori dell'header del segmento.
- (c) Crea il segmento.
- (d) Passa il segmento al livello Rete.

2. **Azioni Destinatario:**

- (a) Riceve il segmento dal livello Rete.
- (b) Controlla i valori dell'header.
- (c) Estrae il messaggio del livello applicazione.
- (d) Smista il messaggio all'applicazione attraverso la socket.

3.2 TCP vs UDP e Multiplexing/Demultiplexing

I due protocolli di trasporto principali hanno caratteristiche differenti:

1. **Protocollo TCP:**

- (a) Orientato alla Connessione.
- (b) Orientato allo Stream.
- (c) Gestione della connessione.
- (d) Consegna affidabile dei dati.
- (e) Controlli di congestione e flow.

2. **Protocollo UDP:**

- (a) Senza connessione, non affidabile.
- (b) Orientato al messaggio.
- (c) Estensione del servizio *host to host* del livello rete.

Entrambi i protocolli però offrono controlli sugli errori e servizi di multiplexing e demultiplexing.

Multi/Demultiplexing Il livello di trasporto ha la responsabilità di smistare i messaggi ottenuti dal livello di rete tra i vari processi tramite la risoluzione di socket address. Questo risulta necessario se più processi utilizzano lo stesso protocollo di trasporto per la comunicazione su rete, cosa che accade molto spesso. Quindi abbiamo due attori duali:

1. **Mittente in Multiplexing:** Gestisce i dati da più socket e si occupa dell'aggiunta del header di trasporto.
2. **Destinatario in Demultiplexing:** Usa le informazioni nell'header per recapitare i segmenti alla corretta socket.

Nello specifico, nel demultiplexing vengono utilizzati dall'host indirizzi IP e numeri di porta per smistare il segmento verso la corretta socket dato che ogni datagramma si porta dietro un segmento del livello trasporto ed ogni segmento nell'header mantiene un numero di porta sorgente ed uno di destinazione.

Definizione di Porta Ogni comunicazione a livello trasporto è univocamente identificata tramite la combinazione *IP/porta*, una **porta** è quindi definita come un intero a 16 bit che viene assegnato ad un processo, più nello specifico è un **punto di demultiplexing** dei protocolli TCP o UDP. Esistono dei range di porte predefinite, come le *System Ports*, *User Ports* o *Dynamic Ports*. In generale è il sistema operativo ad occuparsi dell'assegnamento di porte.

Multiplexing TCP/UDP Esistono differenze tra i tipi di multiplexing dei due protocolli, infatti il protocollo TCP esegue il multiplexing basandosi su una quadrupla con tutte le informazioni su sorgente e destinatario ossia ($IP_{sorg}, porta_{sorg}, IP_{dest}, porta_{dest}$), mentre il protocollo UDP utilizza la coppia ($IP_{dest}, porta_{dest}$).

3.3 Protocollo UDP

Servizio di trasporto best effort, i datagrammi UDP possono essere persi o consegnati fuori sequenza. Questo protocollo è orientato al messaggio, quindi ogni datagramma è indipendente e di dimensioni limitate. Questo garantisce però una semplicità non indifferente, risultando vantaggioso in contesti di applicazioni time-sensitive.

Impieghi e Caratteristiche del Trasporto UDP E' facile e leggero da gestire, non esiste alcun controllo sulla congestione e le intestazioni non superano gli 8 byte. Viene impiegato in contesti di applicazioni multimediali, dove sono tollerate le piccole perdite, oppure in altre applicazioni come il DNS, HTTP/3 o SNMP (Simple Network Management Protocol).

Azioni Mittente/Destinatario Elenchiamo le azioni effettuate in base al tipo di attore:

1. Azioni Mittente:

- (a) Riceve un messaggio del livello Applicazione.
- (b) Determina i valori dei campi dell'header UDP.
- (c) Crea lo user datagram.
- (d) Passa lo user datagram al livello Rete.

2. Azioni Destinatario:

- (a) Riceve lo user datagram dal livello Rete.
- (b) Controlla la checksum
- (c) Estrae il messaggio di livello Applicazione.
- (d) Smista il messaggio attraverso la socket.

Composizione di un Datagramma UDP Descriviamo l'entità base del protocollo UDP:

numero porta sorgente	numero porta destinazione
lunghezza	checksum
payload	

Il messaggio si compone di 8 byte di intestazione, i numeri di porta per la fase di demultiplexing e la lunghezza del messaggio massima del datagramma UDP è di 65535 byte.

3.4 Principi di Trasferimento Affidabile

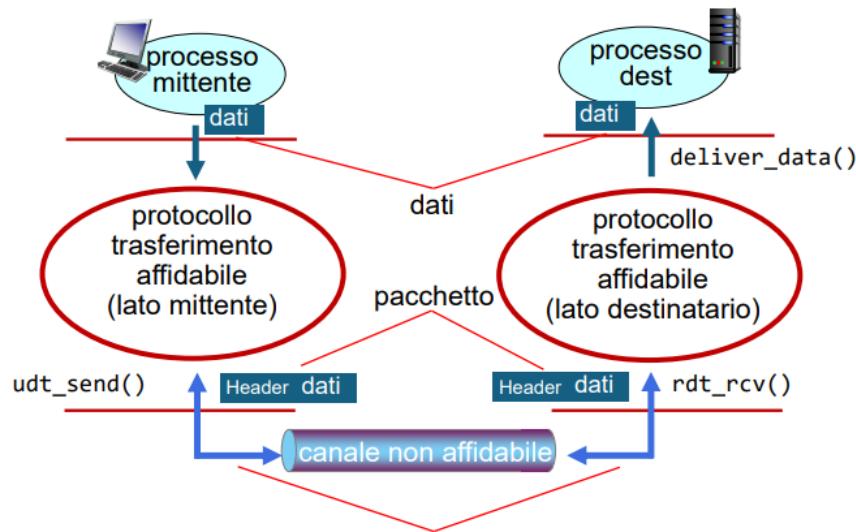
Definiamo in maniera incrementale un **protocollo generico** per il trasferimento di dati affidabile, quindi **senza errori su bit o perdite**. La complessità del protocollo dipende anche dalle caratteristiche del canale non affidabile su cui si sta appoggiando. In questo caso il **livello di rete** è considerato **non affidabile**, dato che può causare **perdite, corruzioni o consegne fuori ordine**.

Interfacce per Trasferimento Affidabile Definiamo le interfacce

1. **RDT**: Reliable Data Transfer
2. **UDT**: Unreliable Data Transfer

che comunicano grazie ai seguenti metodi:

1. *rdt_send()*: Chiamata dall'applicazione per chiedere di consegnare i dati al processo destinatario.
2. *udt_send()*: Chiamata da RDT per inviare il pacchetto al destinatario tramite il canale inaffidabile.
3. *rdt_receive()*: Chiamata dal protocollo inaffidabile UDT quando il pacchetto arriva sul lato destinatario del canale.
4. *deliver_data()*: Chiamata da RDT per consegnare i dati al livello applicazione.



3.4.1 RDT 1.0 - Trasferimento su Canale Affidabile

Questo schema si basa sull'**assunzione** che il **canale di rete** sottostante sia **perfettamente affidabile**. Quindi tutto il funzionamento si riduce a queste due entità che eseguono le seguenti operazioni:

1. Il **mittente invia** al livello di rete i dati che riceve dal livello applicazione, quindi viene effettuato l'**incapsulamento del messaggio**, compenendo un **pacchetto**¹.
2. Il **destinatario inoltra** al livello applicazione i dati che riceve dal livello di rete quindi viene effettuato il **decapsulamento del pacchetto** estraendo un **messaggio**.

3.4.2 RDT 2.0 - Trasferimento su Canale con Errori su Bit

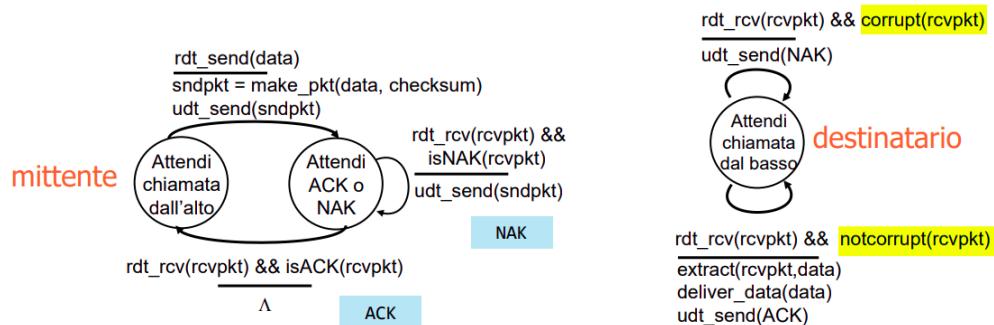
Indebolendo l'ipotesi dello schema precedente, assumiamo di poter ottenere **errori sui bit** a causa del **transito sul canale inaffidabile**.

Ritrasmissione con ACK/NAK Si utilizzano quindi riscontri positivi **ACK** per confermare la corretta ricezione e negativi **NAK** per richiedere ritrasmissione. I protocolli basati su ritrasmissione sono anche detti **ARQ** (Automatic Repeat ReQuest).

Meccanismi Caratteristici dei ARQ I protocolli ARQ si basano su questi meccanismi:

1. **Checksum per rilevamento errori sui bit.**
2. Meccanismo di **riscontro errore esplicito** con bit settato per **ACK o NAK**.
3. **Ritrasmissione** dei **pacchetti** ricevuti con errori.

Il tipo di protocollo ARQ che trattiamo è quello **stop and wait**, ossia il mittente invia un messaggio e poi aspetta riscontro positivo da parte del destinatario prima di inviare il nuovo pacchetto.



In questo schema però non si tiene conto di possibili corruzioni di ACK e NAK. Questo problema viene risolto nel RDT2.1 grazie ad una checksum per il controllo dell'ACK e alla ritrasmissione in caso di corruzione. Sarà necessario anche quindi un meccanismo di numerazione per capire se il pacchetto sia un duplicato o meno.

¹Nel protocollo TCP si parla di segmenti, non di pacchetti, ma nel contesto di trasferimento affidabile generico parleremo di pacchetti.

3.4.3 RDT 2.1 - Checksum per Controllo Corruzione ACK

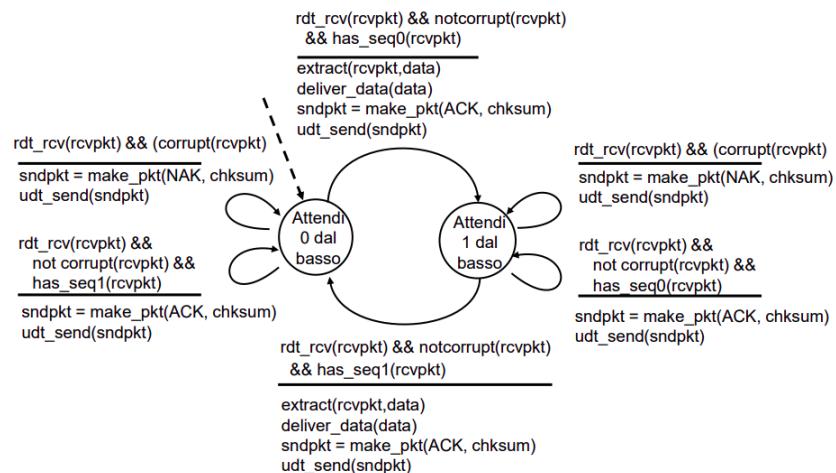
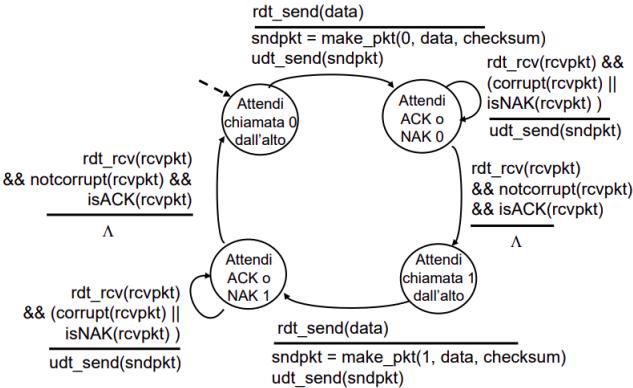
Viene aggiunto un numero di sequenza ai pacchetti lato mittente, che nel caso dello stop and wait di prima sarà quindi $\{0, 1\}$. Quindi i passi sono:

1. Mittente:

- (a) Aggiunta numero di sequenza al pacchetto.
- (b) Verifica corruzione ACK/NAK tramite checksum.
- (c) Quantità stati della macchina raddoppiati: Lo stato tiene conto della numerazione del pacchetto atteso.

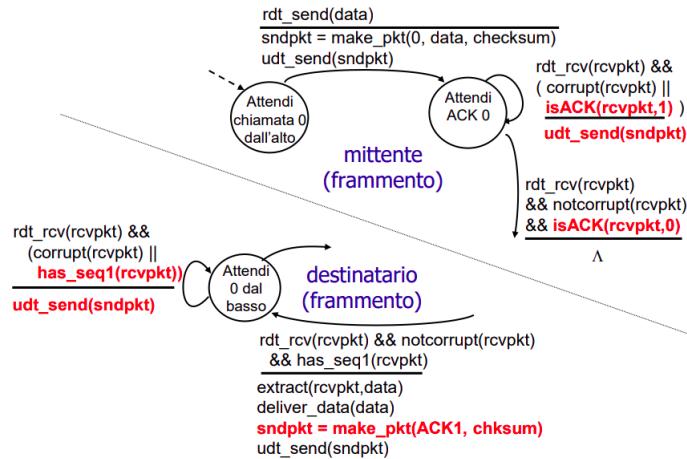
2. Destinatario:

- (a) Verifica se pacchetto ricevuto sia duplicato, grazie alla numerazione.



3.4.4 RDT 2.2 - Protocollo NAK-free

Il destinatario non invia NAK esplicativi, ma invia gli ACK dell'ultimo pacchetto ricevuto correttamente. Quindi risulta importante includere il numero di sequenza del pacchetto da riscontrare. Quando il mittente riceve ACK duplicati, ritrasmette il pacchetto corrente.

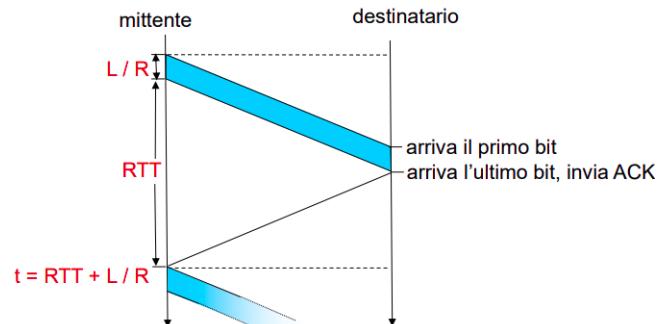


3.4.5 RDT 3.0 - Timer in Attesa di ACK

Se la perdita pacchetti esiste nel canale sottostante, allora il mittente potrebbe non ricevere alcuna risposta dal destinatario. Il mittente dunque aspetta ogni ACK per un massimo di tempo detto **timeout**. Se il pacchetto non è perso ma solo in ritardo, allora la ristrasmissione invia un duplicato, che sarà individuabile grazie al numero di sequenza da riscontrare.

3.4.6 Prestazioni e Reale Utilizzo del Canale

RDT3.0 risulta essere molto valido da un punto di vista funzionale, ma da un profilo di prestazioni notiamo che l'utilizzo del canale è molto ridotto a causa dell'attesa del mittente. Volendo quantificare l'utilizzo del canale possiamo definire questo schema con relativa formula:



$$U_{send} = \frac{L/R}{RTT + L/R} = \text{tempo utilizzato dal mittente per inviare pacchetti}$$

Questo tipo di problematiche si risolvono con approcci di **pipelining**.

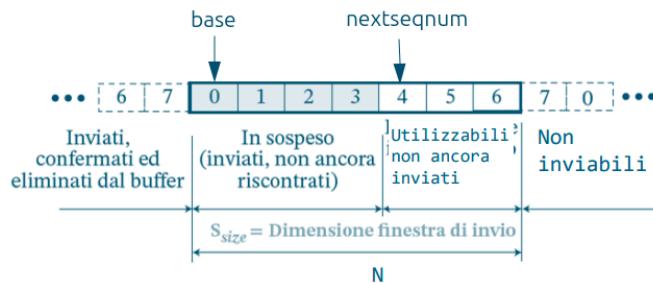
3.4.7 Approcci di Pipelining - Go-Back-N (GBN) e Selective Repeat (SR)

Esistono due approcci di pipelining: Go-Back-N (GBN) e Selective Repeat (SR).

Go-Back-N (GBN) Definiamo ogni responsabilità dei ruoli di mittente e destinatario.

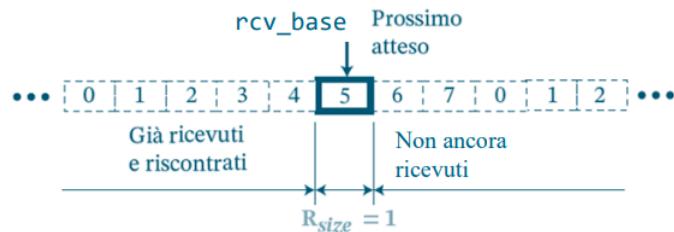
1. Mittente

- (a) Il mittente può inviare fino ad n pacchetti, ossia la dimensione della sua finestra, senza ricevere ACK.
- (b) Si mantengono due puntatori sulla finestra:
 - i. *base*: primo pacchetto, il più vecchio non riscontrato.
 - ii. *nextseqnum*: prossimo pacchetto da inviare.
- (c) Possono essere riscontrati pacchetti in maniera cumulativa, scorrendo la finestra di n posizioni su cui si chiama la $ACK(n)$.
- (d) Il mittente ritrasmette il pacchetto e tutti quelli con un numero di sequenza maggiore di quello non riscontrato nel caso in cui scade il timer di quello puntato da *base*.



2. Destinatario

- (a) Invia un ACK ogni volta che riceve un pacchetto.
- (b) L'ACK è sempre riferito all'ultimo pacchetto ricevuto in ordine.
- (c) Mantiene un solo puntatore *rcv_base*.
- (d) I pacchetti ricevuti fuori ordine vengono scartati.



Questo tipo di approccio ha problemi di prestazioni quando il livello di rete è soggetto a numerose perdite di pacchetti.

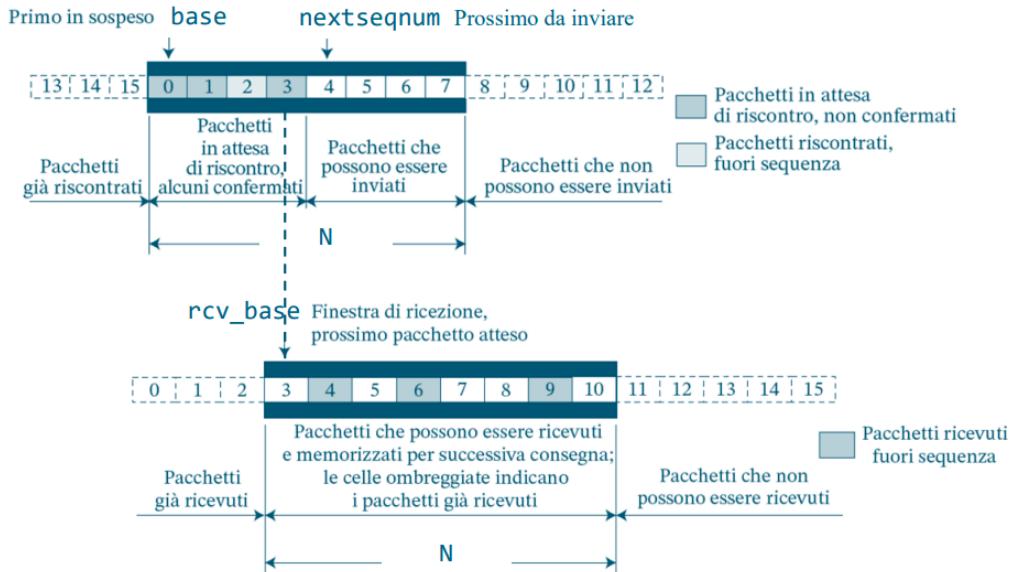
Selective Repeat (SR) Questo approccio utilizza ACK individuali per ciascun pacchetto ricevuto correttamente, si appoggia quindi ad un buffer da cui verranno estratti i pacchetti e consegnati in ordine all'applicazione. Nello specifico, vengono utilizzate **due finestre** di pacchetti, una **lato mittente** ed una **lato destinatario**.

1. Mittente

- (a) Invocazione dall'alto, se c'è un numero di sequenza disponibile allora invia il pacchetto.
- (b) In caso di $timeout(n)$, il pacchetto n viene ritrasmesso e risettato il timer.
- (c) La gestione di $ACK(n)$ in $[base, base + N + 1]$ è definita come:
 - i. Segna n come ricevuto.
 - ii. Se $n = base$ allora si fa scorrere la finestra fino al prossimo pacchetto non riscontrato.

2. Destinatario

- (a) Se $pktOk \in [rcv_base, rcv_base + N + 1]$ allora:
 - i. Invia $ACK(n)$, quindi un nuovo pacchetto è stato correttamente ricevuto.
 - ii. Se il pacchetto in questione è fuori ordine allora lo si salva nel buffer.
 - iii. Se il pacchetto è in ordine allora si effettua una consegna di gruppo di pacchetti consecutivi in ordine nel buffer e si lascia scorrere la finestra fino al prossimo pacchetto non ricevuto.
- (b) Se $pktOk \in [rcv_base - N, rcv_base - 1]$ allora:
 - i. Invia $ACK(n)$, per indicare che il pacchetto già è stato riscontrato.
- (c) Altrimenti scarta il pacchetto.



3.4.8 Riassunto Meccanismi Affidabili

Listiamo una serie di meccanismi riguardanti il trasferimento affidabile con le relative caratteristiche:

- **ACK:** Indica al mittente che un pacchetto è stato ricevuto correttamente dal destinatario.
- **NAK:** Indica al mittente che un pacchetto è stato **non** ricevuto correttamente dal destinatario.
- **Ritrasmissione:** Permette al ricevitore di ricevere un pacchetto che è stato corrotto o perso in una trasmissione precedente.
- **Numeri di Sequenza:** Consente il rilevamento di duplicati lato ricevitore.
- **Checksum:** Utilizzato dal ricevitore per rilevare i bit invertiti durante la trasmissione di un pacchetto.

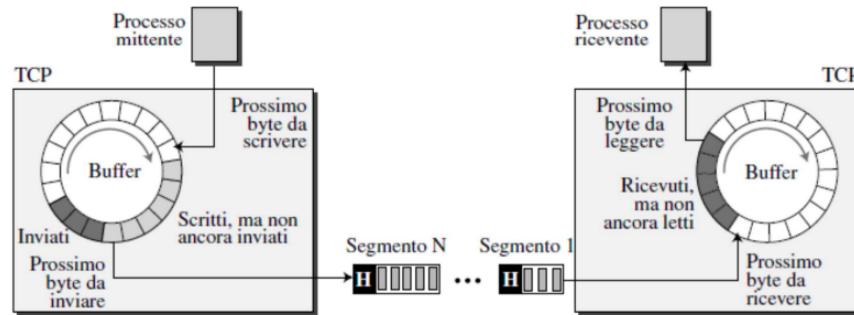
3.5 Protocollo TCP

Iniziamo la descrizione in verticale sul protocollo TCP.

Proprietà del Protocollo TCP Il protocollo TCP garantisce queste proprietà:

- **Orientato alla Connessione:** Si effettuano handshake prima dell'inizio della connessione, e lo stato della connessione risiede sui punti terminali della connessione e non su quelli intermedi.
- **Connessione Full Duplex:** La connessione punto a punto è per definizione bidezionale.
- **Multiplexed:** Consente l'assegnamento di una connessione ad un determinato processo.
- **Controllo della Connessione:** Fornisce meccanismi di inizio e fine trasmissione.
- **Trasferimento di Dati Ordinato e Affidabile:** Permette il controllo di tutti i tipi d'errore quindi dati corrotti, segmenti persi, segmenti duplicati e segmenti fuori sequenza.
- **Controllo di Flusso:** Evita di spedire più dati di quanti non ne possa handlare il destinatario.
- **Controllo di Congestione:** Permette di recuperare da situazioni di sovraccarico della rete.

Bufferizzazione del Trasferimento Il flusso di dati in byte fornito dall'applicativo viene suddiviso in segmenti TCP, e per poter permettere questo si utilizzano **due buffer** uno lato **mittente** ed uno lato **destinatario**. Quindi il flusso di byte viene partizionato in segmenti, ciascuno di essi avrà un header e viene consegnato al livello IP.



3.5.1 Formato Segmento TCP

Dato un flusso di **byte numerati** TCP suddivide questo stream in segmenti, mantenendo quindi un enumerazione sui segmenti in base a specifici byte, quindi un **segmento è identificato dal suo primo byte di dati nel segmento**. In questo modo il riscontro è rappresentato dall'ultimo byte ricevuto correttamente +1.



Descriviamo a pagina successiva questa struttura passo passo.

- **Porte sorgente e Destinazione**
- **Numero di Sequenza:** Si riferisce al primo byte nel flusso che caratterizza il segmento.
- **Numero di ACK:** Numero di sequenza del prossimo byte atteso, se FLAG = 1 allora è un ACK.
- **FLAGS di Bit Codice:** Sei FLAGS che hanno specifiche funzionalità:
 - **URG:** Campo puntatore urgente, da utilizzare se ci sono dati da trasferire in via prioritaria.
 - **ACK:** Se settata ad 1 indica che il numero di riscontro contiene dati significativi.
 - **PSH:** Se settata ad 1 causa il trasferimento dei dati immediatamente all'applicativo.
 - **RST:** Richiede il reset della connessione.
 - **SYN:** Sincronizza il Numero di Sequenza.
 - **FIN:** Se settata ad 1 indica la richiesta di chiusura della connessione.
- **Finestra di Ricezione:** Numero di byte che il destinatario² può ricevere.
- **Checksum.**
- **Opzioni TCP:** Permette la negoziazione di parametri, come ad esempio la dimensione massima di un segmento (MSS).

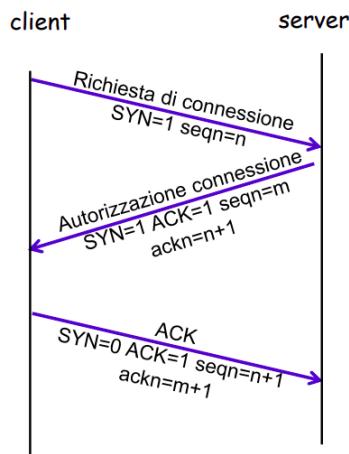
3.5.2 Gestione della Connessione TCP - Handshaking e Chiusura

Analizziamo le fasi di setup, di effettivo trasferimento di dati e chiusura di una connessione TCP.

Three Way Handshake Prima di trasmettere effettivi dati mittente e destinatario si concordano sull'apertura di una connessione, stabilendo anche i parametri della connessione. I primi segmenti quindi non sono di dati effettivi, permettono anche l'inizializzazione dei due buffer, necessari per il controllo del flusso e della congestione. Al momento di arrivo del terzo segmento la connessione è instaurata.

Mostriamo a pagina successiva passo passo i dettagli dello scambio di questi 3 segmenti iniziali.

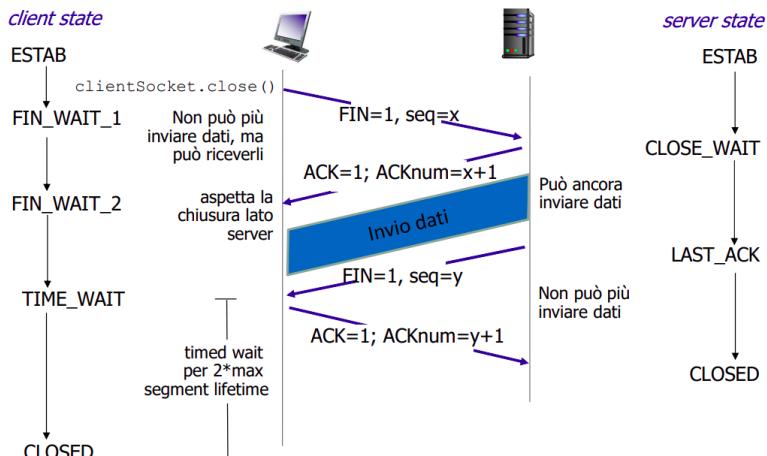
²Il destinatario in questo caso è chi sta componendo questo segmento, in veste di futuro destinatario, comunica all'altra parte quanto può ricevere nel suo buffer. Questo appiattimento dei ruoli di mittente e destinatario è causato dal fatto che dopo l'handshake i due attori sono posti allo stesso livello.



- Il **client** invia una **richiesta di connessione** ad un **server TCP**, settando come attiva la flag **SYN** e trasmettendo anche un **numero di sequenza iniziale**.
- Il server alloca i buffer e risponde con un segmento di connessione garantita detto **SYNACK**, nello specifico:
 - Con **SYN** attivo è significativo il valore di sequenza, che è il valore iniziale.
 - Anche **ACK** è attivo.
- Il client alloca il buffer dal suo fato e invia un riscontro positivo del messaggio del server.
 - Si setta **SYN = 0** per fare in modo che dopo questo terzo segmento inizi l'effettivo scambio di dati.

Quindi da questo momento in poi server e client sono sullo stesso livello.

Chiusura di Connessione TCP Client e server chiudono ciascuno il loro lato della connessione, inviando segmenti dove **FIN = 1**. Ciascuno risponderà ad un segmento del genere con un **ACK**.



Si noti che sulla precedente rappresentazione erano riportati degli stati di una connessione TCP. Ne esistono diversi e in particolare lo stato **TIME_WAIT** è lo stato finale in cui chi sta richiedendo la chiusura resta prima di passare alla reale chiusura. Questo stato è caratterizzato dalla presenza di un timer che lo fa terminare, e la durata di questo timer è correlata alla MSL, ossia **Maximum Segment Lifetime**. L'esistenza di questo stato garantisce l'implementazione in maniera affidabile della terminazione di una connessione in entrambe le direzioni e l'eliminazione di segmenti duplicati in rete.

Chiusura Parziale di Connessione TCP Se uno dei due lati lancia un segmento con **FIN = 1**, allora non potrà inviare dati ma potrà ancora riceverli mettendosi quindi in stato di **FIN_WAIT_2**.

3.5.3 Trasferimento Affidabile del TCP

Un segmento può essere smarrito o corrotto, quindi il TCP crea un servizio di trasferimento affidabile basandosi su questi meccanismi:

- **Checksum:** Permette di scartare segmenti corrotti.
- **Numero di Sequenza:** Enumerazione di segmenti basata sui primi byte di ciascun segmento.
- **Numero di Riscontro:** Numero di sequenza del byte che l'host sta attendendo dall'altro.
- **Riscontro Cumulativo:** Si riscontra fino al primo byte mancante nel flusso.
- **Timer.**

Eventi Lato Mittente Descriviamo il comportamento del mittente:

- Riceve i dati dall'applicativo.
- Incapsula di dati in segmenti e li numera con la logica a byte definita prima.
- Avvia il timer di ritrasmissione RTO solo se non è già assegnato a qualche altro precedente segmento.
- Viene effettuata una **ritrasmissione** se:
 - Scade un timeout: Ritrasmette il segmento che non è stato riscontrato, riavviando il timer.
 - Riceve tre ACK duplicati: Questo è indice di perdita del segmento successivo all'ultimo riscontrato, quindi viene effettuata una **fast transmission** prima della scadenza del timer.

I segmenti ottenuti fuori sequenza possono essere temporaneamente memorizzati dal destinatario, ma la loro gestione dipende dall'implementazione del TCP a cui stiamo facendo riferimento.

Eventi Lato Destinatario Descriviamo il comportamento del destinatario:

- Tutti i segmenti inviati contengono ACK.
- Se il destinatario **non ha dati da inviare e riceve un segmento in ordine** allora ritarda l'invio di ACK di 500ms
- Se il destinatario **riceve il segmento atteso e il precedente non è stato riscontrato** allora invia immediatamente un ACK.
- Se il destinatario riceve un **segmento fuori sequenza, mancante in una sequenza** oppure **duplicato** allora invia immediatamente un ACK.

3.5.4 TCP RTT, RTO, Stime Timeout e Finestre di Ricezione

Il tempo di timeout RTO è fondamentale per il funzionamento del TCP e non può essere né troppo breve né troppo lungo, va quindi accuratamente stimato. Ci basiamo quindi su due misure:

- **SampleRTT**: Misura del tempo trascorso da quando si invia un segmento a quando si riceve il suo riscontro.
- **EstimatedRTT**: Stima del valore precedente calcolata iterativamente come:

$$\text{EstimatedRTT}_{i+1} = (1 - \alpha) \text{EstimatedRTT}_i + \alpha \text{SampleRTT}$$

- **RTT_DEV**: Stima della variabilità di RTT determinata con:

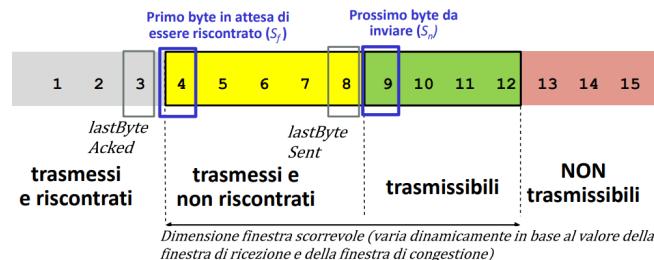
$$\text{RTT}_{DEV_{k+1}} = (1 - \beta) \text{RTT}_{DEV_k} + \beta |\text{RTT}_{SAMPLE} - \text{RTT}_{ESTIMATED}|$$

- **RTO**: Una volta ottenuti questi valori il timeout viene calcolato come:

$$RTO = \text{RTT}_{ESTIMATED} + 4 \text{RTT}_{DEV}$$

Le quantità α e β sono determinate arbitrariamente, spesso nello svolgimento di esercizi si settano come $\alpha = \frac{1}{8}$ e $\beta = \frac{1}{4}$.

Finestra di Ricezione TCP I dati inviati dall'applicativo sono mantenuti nel buffer di invio, e la trasmissione si basa su una sliding window che viene fatta scorrere a tempo di ricezione di un ACK.



3.5.5 Controllo di Flusso del TCP

Assumendo che sia mittente sia destinatario abbiano buffer, bisogna valutare cosa succeda nel caso in cui il **livello di rete** fornisca i dati a **velocità maggiore** rispetto a quella dell'**applicativo a consumarli** dal buffer.

L'obiettivo del **controllo di flusso** è quindi quello di mettere in **relazione la frequenza di invio del mittente** con la **frequenza di lettura** dell'applicazione **ricevente**.

Questo meccanismo è implementato tramite una variabile caratteristica **rwnd** mantenuta dal mittente calcolata come:

$$rwnd = RcvBuffer - (LastByteReceived - LastByteRead)$$

Il mittente fornirà questa variabile al destinatario nell'header del segmento e quest'ultimo verificherà se:

$$LastByteSent - LastByteAcked \leq rwnd$$

Esistono anche dei segmenti sonda di dimensione *1byte* che il mittente invia per ottenere la dimensione della **rwnd**.

3.5.6 Controllo di Congestione del TCP

Il fenomeno della congestione è causato da più sorgenti che tentano l'invio di dati tramite lo stesso percorso. Questo può provocare lunghi ritardi e perdita di pacchetti.

Approcci per Controllo Congestione Esistono due tipologie di controllo di congestione di rete:

1. **End To End**: I mezzi intermedi non forniscono alcun feedback, sono quindi gli host terminali tramite sintomi a rilevare la congestione della rete.
2. **Assistito dalla Rete**: I router segnalano la congestione³. Definito anche come **ECN** (Explicit Congestion Notification), dove si utilizzano 2bit nell'**header IP** per indicare **congestione**. Nello specifico abbiamo un bit **ECE** settato dalla **destinazione** nel segmento di riscontro e un bit **CWR lato mittente** per indicare di aver ricevuto la notifica di congestione.

In TCP quindi il controllo di congestione è implementato imponendo a ciascun mittente di **limitare la frequenza di invio di pacchetti** sulla connessione in funzione della **congestione percepita**.

La dimensione della finestra di invio sarà quindi $\min\{rwnd, cwnd\}$ con **rwnd** receiver window e **cwnd** congestion window. Quindi il rate di invio sarà:

$$rate_{invio} \leq \frac{\min\{rwnd, cwnd\}}{RTT}$$

³Questo infrange gli schemi impostati fino ad ora di totale isolamento tra i vari livelli del TCP/IP.

Algoritmo Controllo Congestione di TCP Si compone di tre fasi:

- Slow Start
- AIMD (Additive Increase Multiplicative Decrease/Congestion Avoidance)
- Reazione agli Eventi Perdita:
 - Fast Recovery
 - Reazione al Timeout

Prima della descrizione effettiva delle fasi è utile definire la **cwnd** (Congestion Window) misurata solitamente in **MSS** (Maximum Segment Size), ossia la massima quantità di dati trasportabile da un segmento.

Slow Start Fase in cui ad ogni ACK la cwnd viene incrementata di 1 MSS, quindi si radoppia ad ogni RTT. Segue quindi un andamento esponenziale.

AIMD (Congestion Avoidance) Fase in cui TCP mittente aumenta proporzionalmente la propria finestra di congestione ad ogni ACK ricevuto:

- Ad ogni ACK la finestra di congestione viene incrementata, nello specifico aumento di 1 MSS per ogni RTT, quindi $cwnd+ = \frac{1}{cwnd}$.

3.5.7 Tipologie di TCP che Istanzano Algoritmo di Controllo Congestione - (Reno, Tahoe, Cubic e Vegas)

TCP Reno Istanza delle fasi che abbiamo appena definito, aggiungendo anche gli eventi di perdita:

- **Gestione cwnd:** Viene definita una variabile soglia ad un valore alto.
 - Se $cwnd < \text{soglia}$ allora si incrementa cwnd aumenta esponenzialmente in **Slow Start**.
 - Se $cwnd > \text{soglia}$ allora si incrementa cwnd aumenta linearmente in **Additive Increase (Congestion Avoidance)**.
- **Eventi di Perdita:** Esistono varie condizioni:
 - **Timeout:** Si setta $\text{soglia} = 1/2\text{cwnd}$ e $\text{cwnd} = 1 \text{ MSS}$ in **Slow Start**.
 - **3 ACK Duplicati:** $\text{soglia} = 1/2\text{cwnd}$ e $\text{cwnd} = \text{soglia} + 3\text{MSS}$ in **Fast Recovery**. In particolare in questa fase possono anche accadere altri eventi:
 - * Se avviene un **timeout** si va in **Slow Start**.
 - * Fino a che arrivano **ACK duplicati** allora $\text{cwnd} = \text{cwnd} + 1\text{MSS}$.
 - * Se arriva un **ACK non duplicato** allora $\text{cwnd} = \text{soglia}$ e si va in **Congestion Avoidance**.

A pagina successiva si mostra un esempio di controllo congestione in TCP Reno.

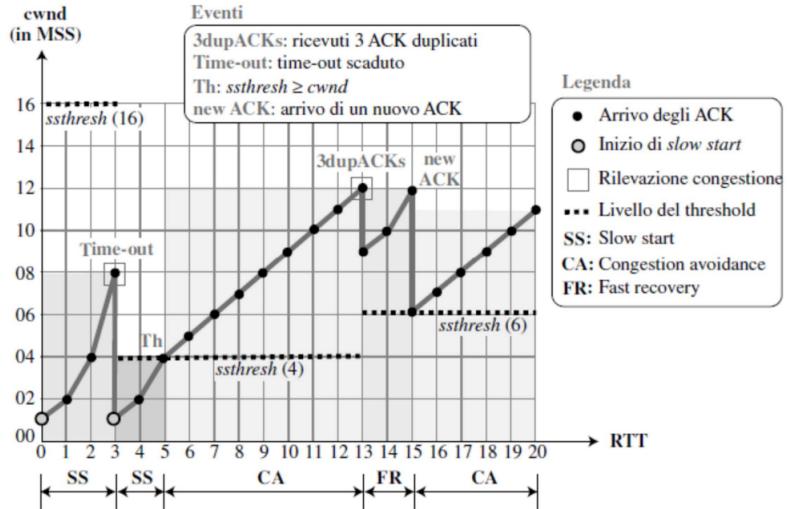


Figure 3: Gestione di congestione in TCP Reno

Si notano quindi i tre possibili stati ossia Slow Start SS, Congestion Avoidance CA e Fast Recovery FR.

TCP Tahoe Versione meno recente che comprende solo gli stati di Slow Start e Congestion Avoidance ed è così definito:

- **Timeout e 3 ACK Duplicati** sono gestiti allo stesso modo, ossia in **Slow Start** settando **soglia = cwnd/2** e **cwnd = 1MSS**.

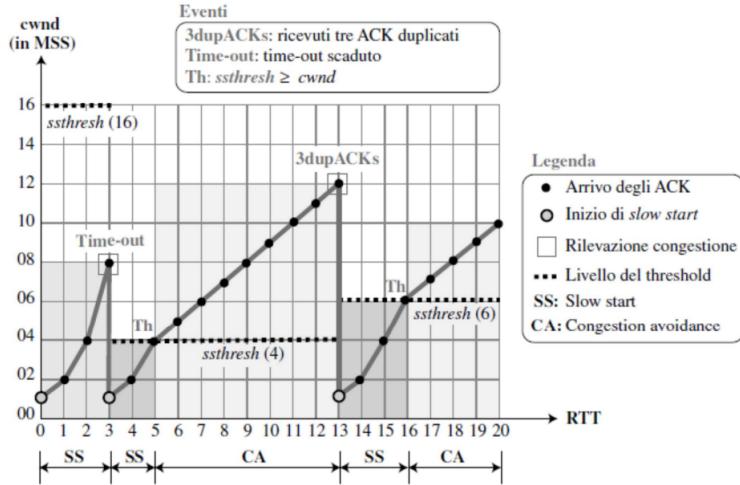
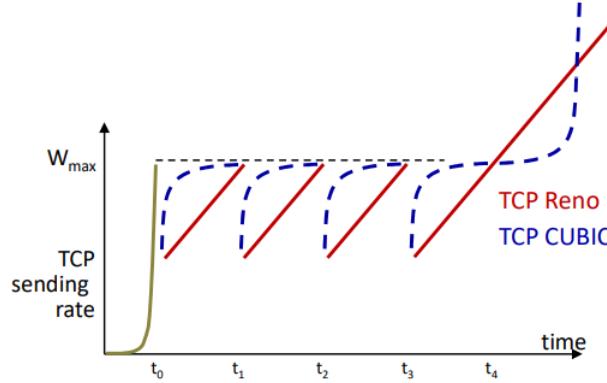


Figure 4: Gestione di congestione in TCP Tahoe

TCP Cubic Cerca di ottimizzare l'andamento "a dente di sega" del TCP Reno settando degli intervalli di ampiezza k e gli incrementi vengono così gestiti:

- Incrementi maggiori quando siamo lontani da k .
- Incrementi minori quando siamo vicini a k .



3.5.8 Throughput, Equità e Cenni di Attacchi TCP

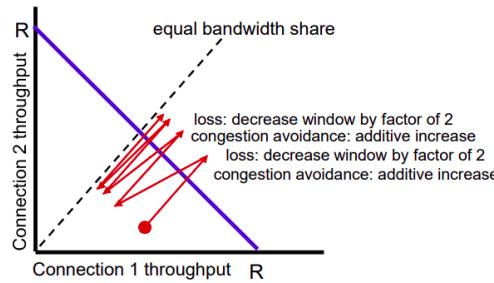
Avendo mostrato i tre stati di TCP Reno, possiamo dire che macroscopicamente il TCP ha un andamento "a dente di sega". Da questo possiamo analizzarne il throughput, assumendo che il rate di trasmissione sia circa costante. Quindi:

- W : Finestra di taglia massima.
- w : Finestra di taglia minima, ossia $w = \frac{W}{2}$.
- **Throughput**: Se la finestra è W allora è $\frac{W}{RTT}$, invece dopo un evento di perdita quando la finestra è w il throughput diventa $\frac{w}{2RTT}$.

Calcolando quindi una media tra le due possibili casistiche otteniamo:

$$\text{throughput} = \frac{1}{2} \left(\frac{W}{RTT} + \frac{w}{RTT} \right) = \frac{3}{4} \frac{W}{RTT}$$

Equità Il TCP è equo perché ipotizzando un bottleneck settato ad una quantità R e un numero di host k , allora ciascuna delle connessioni TCP trasmetterà $\frac{R}{k} \text{bit/sec}$, seguendo quindi un andamento simile a quello rappresentato in figura.



Throughput e Connessioni TCP Parallelle Nel paragrafo precedente ipotizzavamo che ciascun applicativo aprisse esattamente una connessione TCP. Questo in un contesto reale potrebbe risultare una semplificazione, ed un applicativo potrebbe richiedere l'apertura di più connessioni TCP. Quindi ciascuna di queste connessioni sarà alla pari delle altre, ed è a questo livello che si considera l'equità tra le connessioni stesse.

Attacco SYN Flood e Potenziale Mitigazione Attacco DDoS basato sul funzionamento del handshake TCP:

- L'attaccante finge di essere una moltitudine di client, i quali inviano un segmento di SYN al server.
- Il server, per handlare queste richieste, alloca un buffer per ogni SYN e risponde con dei SYN-ACK.
- L'attaccante non riscontra mai i SYN-ACK del server, lasciando le risorse del server allocate. Questo può quindi causare un esaurimento di risorse del server.

Mitigazione dell'Attacco Invece di allocare direttamente il buffer, il server genera un numero di sequenza iniziale detto **cookie**. Successivamente il server risponde con un SYN-ACK senza memorizzare informazioni sullo stato della connessione. Un client normale risponderà quindi con un $ACK = \text{cookie} + 1$. A questo punto il server calcola il risultato con il valore contenuto nell'ACK del client, e solo se c'è corrispondenza crea la connessione ed alloca le risorse.

3.6 Protocollo QUIC e Approfondimento HTTP/2 HTTP/3

In questo sottocapitolo definiamo come migliorare il tempo di caricamento di una pagina agendo a livello di protocollo. Quindi definiremo un "wrapper" di UDP, ossia QUIC, e i suoi utilizzi nel protocollo applicativo HTTP/3.

3.6.1 HTTP/1.1 e Head of Line Blocking

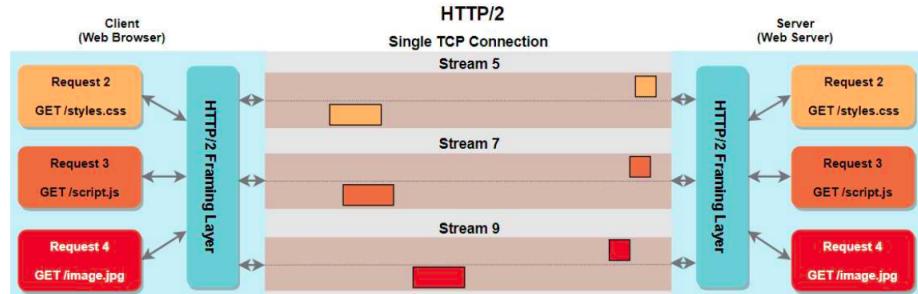
Utilizzando HTTP/1.1 in pipelining, la richiesta di grandi oggetti potrebbe ritardare l'ottenimento di oggetti piccoli richiesti dopo. Questo fenomeno è detto HOL Blocking (Head Of Line). Una possibile alternativa sarebbe quella di aprire **più connessioni TCP per browser**, ma questo porterebbe ad un sistema **poco equo**.

3.6.2 HTTP/2

Questo protocollo applicativo permette maggiore flessibilità lato server, permettendo:

- **Ordine di trasmissione** basata sulla **priorità** degli oggetti richiesti.
- **Push** di oggetti **dal server al client**, che non corrispondono a richieste esplicite dal client.
- **Suddivisione** dei **messaggi HTTP** in **frame HTTP**, e schedulazione dei frame per mitigare la HOL Blocking.

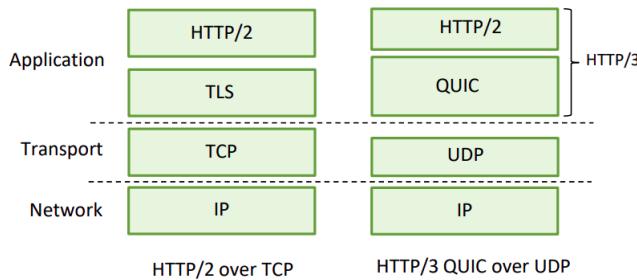
Stream e Multiplexing Vengono definiti dei flussi bidirezionali tra client e server all'interno di una singola connessione TCP. Questo meccanismo permette quindi il multiplexing di oggetti sui vari stream della connessione.



Limiti dell'HTTP/2 questo approccio ottimizza l'utilizzo di risorse del protocollo TCP, ma non risolve del tutto il problema del HOL Blocking, dato che la gestione a Frame HTTP non è nota al protocollo TCP stesso. Per il TCP stanno passando solo una quantità generica di byte, senza dare a questi ultimi un'interpretazione. È infatti responsabilità degli applicativi dare una semantica alla suddivisione dei messaggi. Quindi un problema comune potrebbe essere quello dei tre timeout, che provoca uno stall per tutti gli stream della connessione.

3.6.3 HTTP/3 e QUIC

Evoluzione del protocollo HTTP/2 basato sul protocollo di trasporto UDP "wrappato" dal protocollo QUIC.

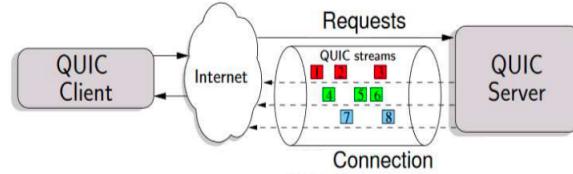


QUIC (Quick UDP Internet Connections) Protocollo di trasporto che sfrutta UDP ed è connection-oriented. Aggiunge al protocollo UDP elementi come:

- Trasmissione affidabile, controllo di errori e congestione.
- Apertura delle connessioni, affidabilità, autenticazione e cifratura.
- Lo stato della connessione è stabilito in un solo RTT, a differenza di TCP che ne richiede due⁴.

⁴Abbiamo sempre definito il costo in RTT dell'Handshake TCP come 1 RTT, ma in realtà vanno considerati anche il protocollo aggiuntivo per cifratura ed autenticazione TLS, che porta il costo complessivo a 2 RTT.

Stream QUIC Più stream a livello applicazione vengono multiplexati su un'unica connessione QUIC. Questo garantisce trasferimento dati affidabile, controllo di flusso e controllo comune della congestione. Il concetto di stream è un'astrazione al livello di trasporto, quindi ciascuno di essi è indipendente dall'altro e noto al protocollo di trasporto utilizzato.



Confronto Prestazioni QUIC vs TCP In generale QUIC performance meglio di TCP, ma le performance di QUIC diminuiscono in contesti mobile e reti cellulari.

4 Livello Rete

L'obiettivo di questo livello di astrazione è quello di trasferire pacchetti da un host ad un altro, **offrendo servizi al livello di trasporto** ed **utilizzando** servizi del **livello collegamento**.

Quindi il livello rete è presente negli host di mittente e destinatario e nei router intermedi. Questo livello interconnette reti eterogenee, implementando poche funzioni.

4.1 Servizi e Modelli di Servizio - Forwarding/Routing

Questo livello offre due principali servizi:

- **Forwarding:** Spostamento di pacchetti dal collegamento di ingresso a quello corretto tra gli n in uscita.
- **Routing:** Determinazione del percorso seguito dai pacchetti da sorgente a destinazione.

Ipotetici Servizi di Rete e Offerta Protocollo IP I servizi di rete potrebbero offrire servizi come:

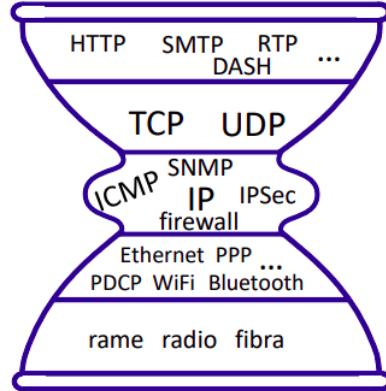
- Garanzia di consegna.
- Consegnare con garanzia temporale.
- Consegnare in ordine.
- Disponibilità di banda garantita.
- Servizi di sicurezza.

Il protocollo IP è di tipo **best-effort** da questo punto di vista, quindi **non garantisce alcuna di queste proprietà**⁵.

⁵ Esistono degli standard basati su IP, come INTSERV e DIFFSERV che garantiscono o possono garantire proprietà come tempo, banda, ordine e perdite.

4.2 Protocollo IP

Protocollo di tipo connection-less, non esiste alcun circuito fisico o virtuale tra host sorgente e destinazione. In origine il livello di rete si riferiva esclusivamente a questo protocollo, attualmente invece attualmente si compone di più protocolli.

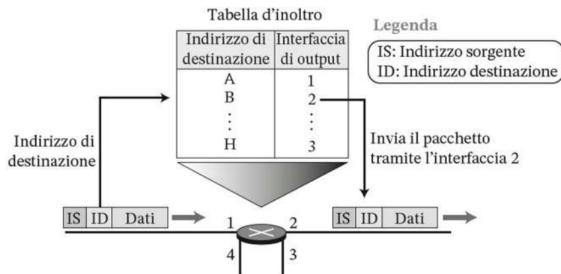


Quindi come definito prima non si ha nessun tipo di garanzia sulla qualità del servizio.

4.2.1 Funzioni IP

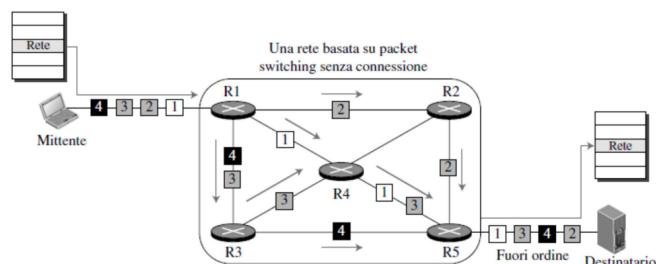
Come già definito prima, il protocollo IP fornisce due principali tipi di funzioni:

- **Forwarding (Inoltro):** Trasferimento del pacchetto sull'appropriato collegamento di uscita. Questa tabella d'inoltro va però prima popolata, questa si determina infatti



tramite la funzione di instradamento.

- **Intradamento:** Processo decisionale di scelta del percorso verso una destinazione, tramite algoritmi di routing.



4.2.2 Meccanismi IP

Il protocollo IP definisce due meccanismi:

- **Indirizzamento:** Strumento per identificare gli host.
- **Modello Datagram:** Consegna dei dati in segmenti basata su specifiche fasi:
 - Frammentazione e riunificazione dei pacchetti.
 - Controllo degli errori.
 - Verifica TTL⁶.
- **Multiplexing/Demultiplexing:** Anche al livello di rete si effettua multiplexing, infatti grazie ad un campo presente nel datagramma IP si riesce a capire a quale protocollo del livello superiore debba essere inviata l'informazione decapsulata.

4.2.3 Formato Datagramma IP

Descriviamo i campi significativi di un datagramma IP:

4 bits	4 bits	8 bits	16 bits		
Version	Header Length	Type of Service	Total Length		
Identification (16 bits)					
Flags (3 bits)	Fragment Offset (13 bits)				
Time To Live (8 bits)	Protocol (8 bits)	Header Checksum (16 bits)			
Source IP Address (32 bits)					
Destination IP Address (32 bits)					
Options (variable, if any)					
Data (variable length, typically TCP/UDP segment)					

Table 1: Formato del Datagramma IP (IPv4 Header)

- **Versione:** Indica la versione utilizzata del protocollo IP.
- **Header Length:** Lunghezza dell'intestazione espressa in parole da 32 bit.
- **Service Type:** Permettono la caratterizzazione di Datagrammi IP:
 - **Differentiated Services:** I pacchetti vengono caratterizzati per la loro classe di servizio o tipi di accodamento ai router (6 bit).
 - **Explicit Congestion Notification (ECN):** Supporto a livello rete della gestione della congestione a livello di trasporto (2 bit).
- **Lunghezza del datagramma (16 bit):** è la lunghezza di tutto il datagramma in byte, header incluso.
 - Dim. Max 65535 byte (nella pratica 1500 byte)

⁶Con TTL si intende Time To Live, ossia quanti salti il pacchetto può fare al massimo.

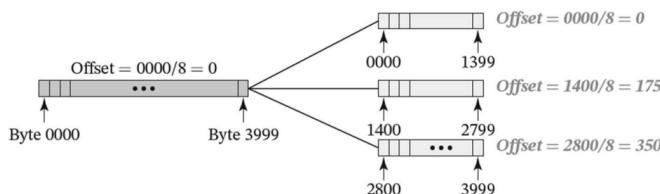
- **Identificatore, flag, offset:** sono campi per la frammentazione (vedi più avanti).
- **Tempo di vita (8 bit):** ad ogni passaggio da un router viene decrementato, quando raggiunge lo zero viene scartato. Assicura che eventuali percorsi ad anello non provochino traffico perpetuo nella rete.
 - Valore iniziale dipende dal S.O.: esempi 30, 64, 128, 255
- **Protocollo (8 bit):** in ricezione all'host destinatario indica quale protocollo dello strato superiore deve ricevere i dati. Inizialmente la tabella era nella RFC1700 (STD2), attualmente è un database online che si trova su iana.org
 - Demux, es. TCP 6, UDP 17
- **Checksum dell'intestazione (16 bit):** viene calcolato il checksum della sola intestazione (ponendo questo campo pari a zero) ad ogni router (il TTL cambia ad ogni hop!). Se si ottiene un errore si scarta il datagramma.
- **Indirizzi sorgente e destinazione (32 bit):** indirizzi IP del mittente e del destinatario.
- **Opzioni (lunghezza variabile, multiplo di 32 bit):** uso sporadico, da 0 a 40 byte
- **Dati:** dati trasportati dal datagramma IP.

4.2.4 Frammentazione di Datagramma IP

Dato che il livello rete dipenderà dal livello collegamento dovrà scendere a compromessi con la dimensione Maximum Transfer Unit (MTU), ossia la dimensione massima di invio del collegamento. Quindi l'idea è quella di suddividere un **Datagramma IP** in tanti **segmenti diversi**, che dovranno poi essere riassemblati nel sistema di destinazione.

Ovviamente questa gestione si porta dietro vari **punti critici**:

- Se uno solo dei frammenti si perde, il sistema terminale è costretto a scartare tutto il resto del datagramma ricevuto.
- I frammenti possono arrivare fuori ordine e di conseguenza vanno riordinati tramite **identificazione, offset eflag**.
 - **Identificazione:** Mantiene un riferimento al **datagramma IP originale**, conservando anche IP sorgente e IP destinazione.
 - **Offset:** Indica la posizione relativa al frammento nel datagramma in multipli di 8 byte.

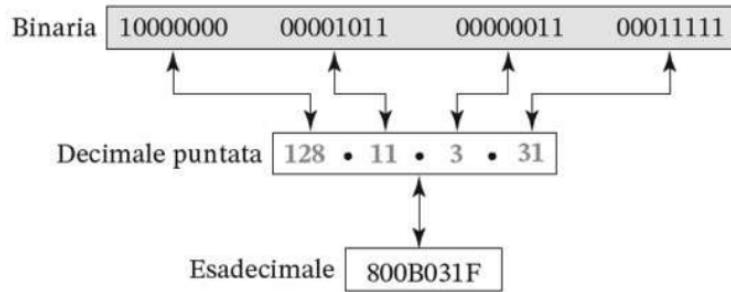


- **Flag:** Tre bit dedicati a dettagli specifici:
 - * **Primo Bit:** Bit riservato.
 - * **Secondo Bit:** Not fragment, se settato ad 1 non si frammenta il datagramma.
 - * **Terzo Bit:** Se settato ad 1 indica che quello è l'ultimo frammento del datagramma.

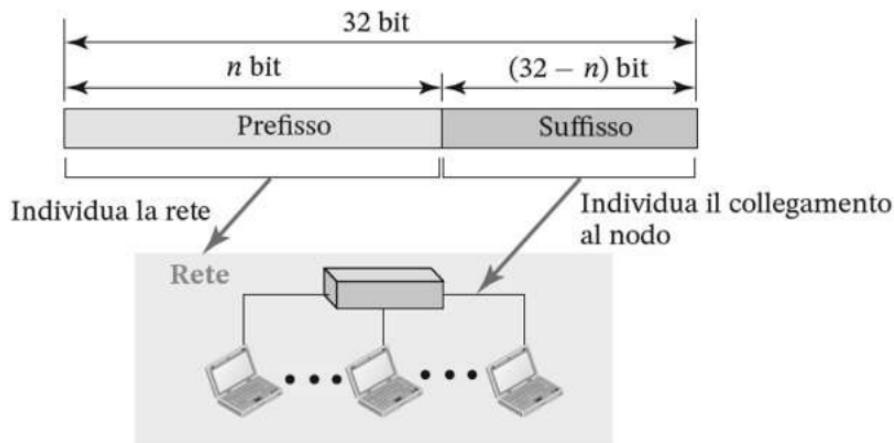
Dati tutti questi elementi di complessità, il **processo di frammentazione è critico**, quindi in tante occasioni si preferisce non frammentare.

4.2.5 Indirizzi IP

Ogni host è connesso ad Internet tramite un **interfaccia di rete**, ossia il **confine tra host e collegamento** su cui vengono trasmessi i datagrammi. Ogni indirizzo IP, secondo il protocollo IPv4, è definito da 4 byte, quindi 32 bit che solitamente vengono riportati nella seguente notazione decimale puntata:



Qgnuno di questi indirizzi si compone di due parti, ossia un **Network ID** ed un **Host ID**, che identificano rispettivamente una rete su Internet e l'host su quella rete IP.



L'assegnazione di specifici bit alla rete ed i restanti agli host determina un tipo di indirizzamento detto **Classful Addressing**.

4.2.6 Classful/Classless Addressing e Subnet Mask

Definiamo delle **classi di indirizzi IPv4** in base ai byte dedicati all'identificazione di reti:

- **Indirizzamento Classful:**

- **Classe A:** Si dedica il primo byte all'identificazione di una rete e i restanti tre per gli host.
- **Classe B:** Si dedicano due byte alla rete e gli altri due per gli host.
- **Classe C:** Si dedicano tre byte alla rete ed il restante per gli host.
- **Classe D:** Riservata al multicasting.
- **Classe E:** Riservata ad usi futuri.

- **Indirizzamento Classless:** Definiamo un modo custom per indicare il numero di bit dedicati alla rete, ossia **x.x.x.x/num_bit_rete**.

Subnet Mask Basandoci sulla notazione classless, possiamo definire una maschera di sottorete, che se messa in AND con un indirizzo IPv4 permette di ricavare l'indirizzo della rete:

Indirizzo IP	150.217.8.42	10010110 11011001 00001000 00101010
Subnet Mask	255.255.255.0	11111111 11111111 11111111 00000000
AND	150.217.8.0	10010110 11011001 00001000 00000000

La rete ha indirizzo
150.217.8.0/24

4.2.7 Assegnazione a Blocchi e Indirizzi Speciali

Assegnazione Blocchi di Indirizzi Se un ISP deve suddividere un blocco in 8 blocchi di indirizzi.

Blocco dell'ISP	<u>11001000 00010111 00010000 00000000</u>	200.23.16.0/20
Organizzazione 0	<u>11001000 00010111 00010<u>000</u> 00000000</u>	200.23.16.0/23
Organizzazione 1	<u>11001000 00010111 00010<u>010</u> 00000000</u>	200.23.18.0/23
Organizzazione 2	<u>11001000 00010111 00010<u>100</u> 00000000</u>	200.23.20.0/23
...
Organizzazione 7	<u>11001000 00010111 00011<u>110</u> 00000000</u>	200.23.30.0/23

- Il numero di indirizzi N in ogni sottoreti deve essere una potenza di 2.
- La lunghezza del prefisso di ogni sottorete va calcolata con la seguente formula $n = 32 - \log_2 N$ dove N è il numero di indirizzi della sottorete.
- Si assegnano blocchi di indirizzi contigui.
- Se i blocchi sono di dimensioni diverse allora si parte dai blocchi più grandi.

Indirizzi Speciali Esiste una serie di indirizzi riservati:

- **This-Host 0.0.0.0:** Usato come placeholder quando un host ancora non conosce il proprio indirizzo IP.
- **Limited Broadcast 255.255.255.255:** Usato quando un router o un host deve inviare un pacchetto a tutti gli host in una rete.
 - **Directed Broadcast:** Serve per indirizzare tutti i nodi di una specifica sottorete esterna.
- **Loopback 127.0.0.1** Il datagramma con questo indirizzo di destinazione non lascia l'host locale.

4.2.8 IP Forwarding - Diretto vs Indiretto

Ogni diagramma IP è soggetto a forwarding, ossia inoltro verso l'uscita. Esistono due tipi diversi di inoltro:

- **Inoltro Diretto:** Il pacchetto IP ha come destinazione un host nella propria rete, quindi **non viene interpellato alcun router** di quella rete.
 - L'Host A deve inviare un pacchetto ad un Host B, conoscendone l'indirizzo IP.
 - L'Host A confronta il proprio indirizzo con quello del destinatario, capisce che si tratta di un host della stessa sottorete e quindi sceglie un inoltro diretto.
 - L'Host A consulta quindi una tabella di corrispondenza **IP_ADD → MAC_ADD**.
 - L'Host A passa il pacchetto al livello di trasporto, che gestirà il frame con header contenente gli indirizzi MAC sorgente e destinatario.
- **Inoltro Indiretto:** Il pacchetto IP ha come destinazione un host di un'altra rete IP. Quindi il pacchetto viene inoltrato **almeno ad un router**.
 - L'Host A deve inviare un pacchetto ad un Host B, conoscendone l'indirizzo IP.
 - L'Host A confronta il proprio indirizzo con quello del destinatario, capisce che si tratta di un host di un'altra sottorete e quindi sceglie un inoltro indiretto.
 - L'Host A quindi recupera l'indirizzo MAC del router dalla tabella di corrispondenza e passa il pacchetto a livello inferiore in sua direzione.
 - Il pacchetto viene costruito e spedito dall'interfaccia del router.

Longest Prefix Matching Quando si cerca una voce della tabella di inostro per un dato indirizzo di destinazione, in caso di più match nella tabella, si seleziona la entry con il prefisso di indirizzo più lungo tra tutte quelle che matchano. Questo perché un prefisso più lungo per la rete garantisce che sia più specifico rispetto agli altri match.

Aggregazione di Indirizzi e Routing Gerarchico Assumiamo che un ISP abbia assegnato 4 piccoli blocchi di indirizzi ad organizzazioni. L'ISP può aggregare questi blocchi in uno singolo e annuncia alla rete il singolo blocco per questioni di efficienza.

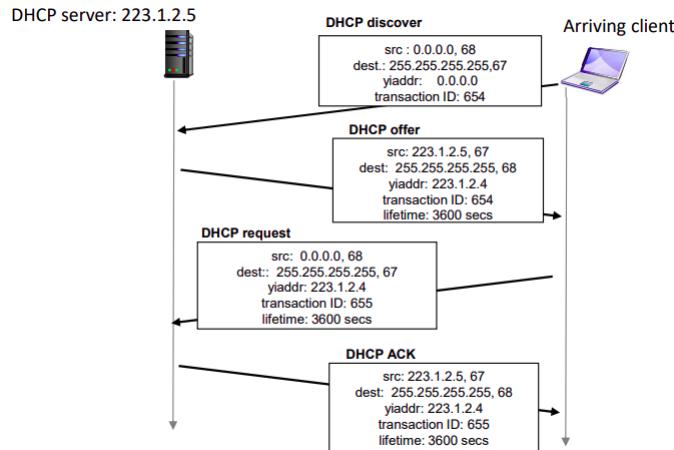
4.3 Protocollo DHCP - Dynamic Host Configuration Protocol

Un indirizzo IP può essere **configurato manualmente** tramite assegnazione di indirizzo IP, indirizzo gateway/router, netmask e indirizzo IP di un server DNS.

In alternativa alla configurazione manuale esiste un assegnazione automatica tramite DHCP, ossia quando un host si aggiunge ad una rete ottiene dinamicamente un indirizzo IP da un programma server nella rete.

Protocollo DHCP secondo Paradigma Client/Server L'assegnazione dinamica si basa su queste fasi:

- L'host invia in broadcast un messaggio DHCP **discover**.
- Il server DHCP risponde con un messaggio DHCP **offer**.
- L'host richiede un indirizzo IP tramite un messaggio DHCP **request**.
- Il server DHCP invia un messaggio DHCP **ack** se la richiesta va a buon fine.



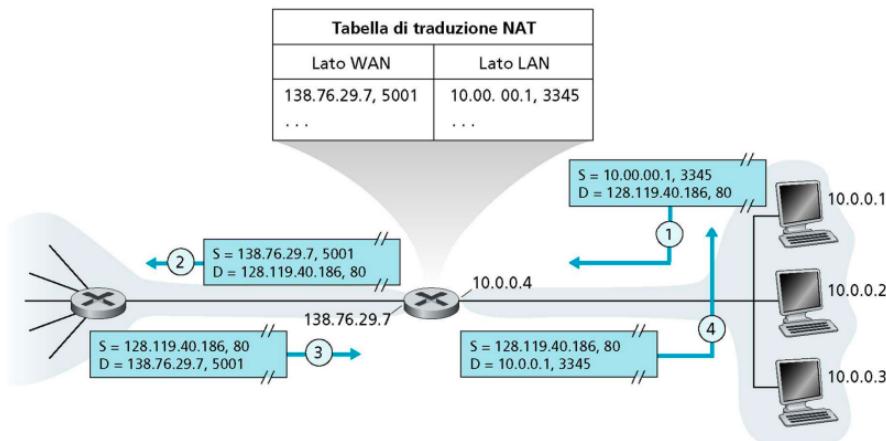
Il DHCP si basa sul protocollo UDP e non si occupa solo di assegnazioni di IP ma anche di informazioni riguardanti indirizzo gateway/router, netmask ed indirizzo IP di almeno un server DNS.

4.4 NAT - Network Address Translation

Questo meccanismo permette di trasmettere su Internet il traffico proveniente da sistemi di sottoreti private, dove l'accesso ad Internet è **abilitato** da un **router** che ha abbia **almeno** un **IP pubblico**.

Quindi tutto il traffico proveniente da ciascuno degli host della sottorete apparirà come in uscita dall'IP pubblico, e tutto il traffico in entrata avrà come destinazione l'IP del router. Ciascun host sarà quindi differenziato da una porta, ma si esporrà all'esterno della sottorete con lo stesso IP di tutti gli altri host della sottorete.

Tabella di Traduzione NAT Il router di accesso alla sottorete ha quindi in memoria una tabella di associazioni per riferire ad un host specifico della sua sottorete.



Critiche al NAT Il NAT è stato visto in modo controverso da chi definiva formalmente le specifiche di Internet per questioni quali ad esempio l'utilizzo di porte per riferire host e non processi, violazione dell'argomento end-to-end, carenza di indirizzi già risolta da IPv6 e attraversamento del NAT per connettersi ad un server nattato.

4.5 ICMP - Internet Control Message Protocol

Il **meccanismo IP** di base **non ha** alcun meccanismo integrato per **notifica di errori**. Oltre a questo è anche sprovvisto di meccanismi per effettuare richieste sullo stato di un sistema remoto.

Alcune Caratteristiche Messaggi ICMP I pacchetti ICMP sono incapsulati all'interno di datagrammi IP. Per pacchetti IP frammentati, i messaggi ICMP sono relativi solo al frammento con offset 0. I messaggi ICMP non sono mai inviati in risposta a pacchetti IP con destinazione IP broadcast o multicast, e allo stesso tempo non sono mai inviati in risposta ad altri messaggi di errore ICMP ma possono essere utilizzati in risposta a messaggi ICMP di interrogazione.

Tipi di Messaggio ICMP I messaggi si distinguono in:

- **Messaggi di Segnalazione Errore**
- **Messaggi di Richiesta/Risposta**

Oltre alla tipologia, ogni messaggio ICMP è caratterizzato anche da due attributi **tipo 8bit** e **codice 8bit**, che indicano il significato dei messaggi.



Messaggio di segnalazione errori



Messaggio di richiesta

Valori del tipo e dei codici

Messaggi di segnalazione errori

- 03:** destinazione non raggiungibile (codici da 0 a 15)
- 04:** source quench (solo codice 0)
- 05:** reindirizzamento (codici da 0 a 3)
- 11:** tempo scaduto (codici 0 e 1)
- 12:** problema a un parametro (codici 0 e 1)

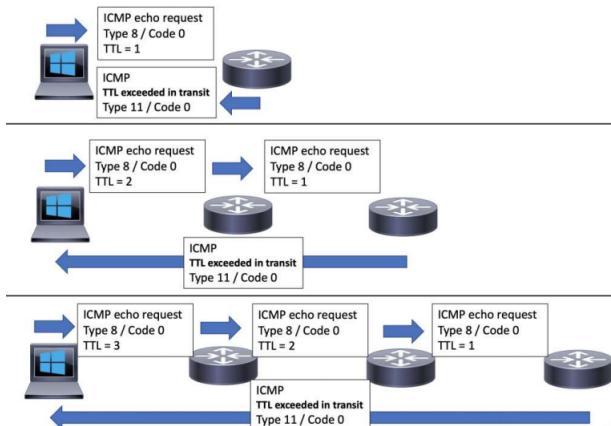
Messaggi di richiesta

- 08 and 00:** richiesta e risposta eco (solo codice 0)
- 13 and 14:** richiesta e risposta timestamp (solo codice 0)

Esempi di Utilizzo ICMP - Ping e Traceroute Descriviamo l'utilizzo di questi due meccanismi che si basano su messaggi ICMP:

- **Ping:** Si basa su uno scambio di messaggi ICMP richiesta/risposta di tipo ECHO, rispettivamente **tipo 8, codice 0** dal mittente e **tipo 0, codice 0** dal destinatario se raggiungibile.

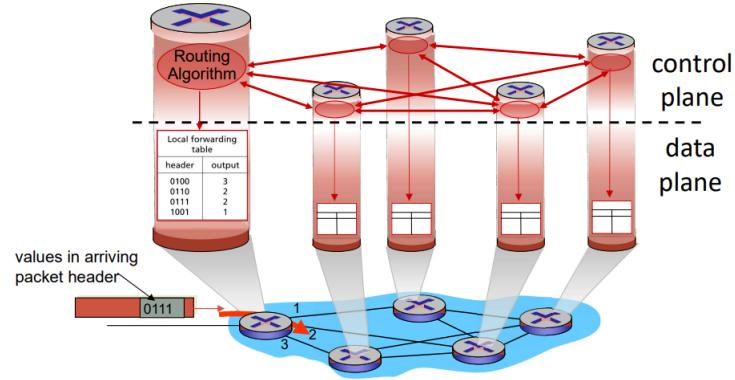
1. **Versione UNIX:** Traceroute si basa sulla configurazione che permette di scatenare l'invio di messaggi di errore correlati a tempo scaduto o porte di destinazione non raggiungibili.
2. **Versione Windows:** Tracert utilizza invece messaggi echo ICMP di tipo request/reply.



4.6 Architettura di Router

Come già accennato, un router deve occuparsi di due tipologie di operazione:

- **Forwarding:** Trasferire i pacchetti sull'appropriato collegamento in uscita, **utilizzando una tabella di inoltro**, l'operazione è detta anche di tipo **data plane**.
- **Routing:** Processo **decisionale** di scelta del percorso verso una destinazione, **popola** quindi la **tabella di inoltro**, azione definita da algoritmi di routing, definita anche come di tipo **control plane**.

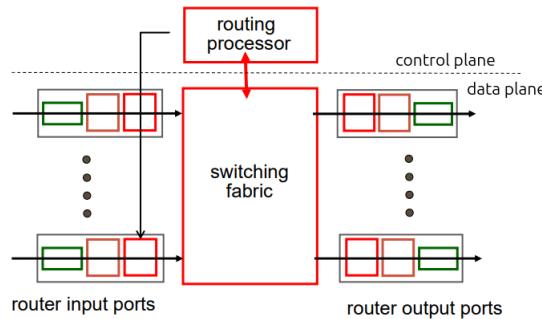


Possiamo quindi definire due tipi di routing:

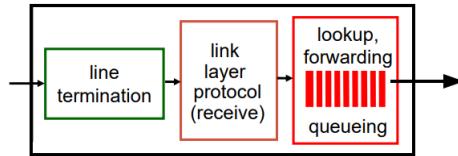
- **Routing Decentralizzato:** L'algoritmo di routing è eseguito su ciascuno dei router, e questi ultimi si scambiano messaggi.
- **Routing Logicamente Centralizzato:** Un controller remoto con Control Agents CA locali si occupa delle informazioni sui collegamenti e sul traffico, ed invia ai singoli CA i valori da inserire nella tabella di inoltro. Questo tipo di gestione è definita anche come **Software-defined Networking SDN**.

4.6.1 Descrizione Fisica dell'Architettura di un Router

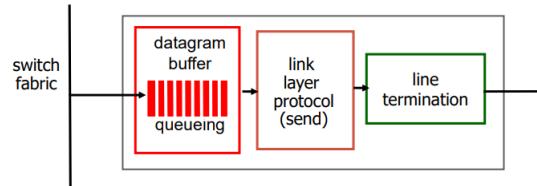
Si basa su delle componenti principali quali porte di input, porte di output, processore di routing e struttura di commutazione detta anche switching fabric.



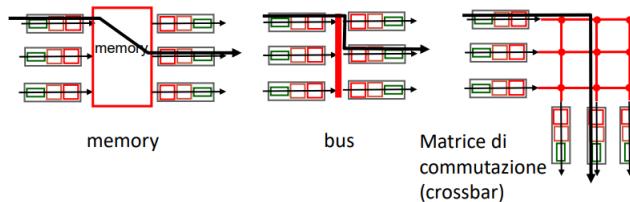
- **Porte di Input:** Si compone di tre parti, una ricezione dei bit, un livello di collegamento ed un modulo per la gestione di lookup, forwarding ed accodamento.



- **Porte di Output:** Si occupa del buffering dei datagrammi quando arrivano dalla struttura di commutazione ad una velocità maggiore rispetto a quella di trasmissione del collegamento in uscita. Oltre a questo devono anche definire delle drop policy per scegliere quali datagrammi scartare in caso di mancanza di spazio e politiche di scheduling per definire l'ordine di trasmissione dei datagrammi in coda.



- **Struttura di Commutazione:** Si occupa del trasferimento del pacchetto dal buffer di input al buffer di output appropriato, caratterizzata da una velocità con cui i pacchetti possono essere trasferiti da un buffer ad un altro.



Ritardi e Perdite Si possono formare delle code di pacchetti sia presso porte di ingresso che di uscita, la lunghezza delle code dipende da vari fattori come traffico di rete, velocità della struttura di comunicazione e velocità della linea.

4.7 Protocolli di Routing

L'obiettivo del routing è quello di determinare buoni percorsi da host mittenti ad host destinatari attraverso host intermedi, definendo quindi percorsi come sequenza di host intermedi tra sorgente e destinazione. Con buono intendiamo che abbia un costo minore.

Astrazione del Grafo Immaginiamo che la rete sia un grafo, ciascun arco ha un peso, quindi ci chiediamo quale sia il **cammino di costo minimo** tra sorgente e destinazione. Un **algoritmo di routing calcola** quindi il **cammino a costo minimo** tra i due nodi.

Routing Statico/Dinamico Esistono due tipi di routing:

- **Statico:** Un admin definisce il path ottimale, popolando manualmente le entry delle tabelle di routing. Questo approccio viene utilizzato per reti di piccole dimensioni.
- **Dinamico:** Definito da protocolli specifici che popolano automaticamente le tabelle di routing con i possibili percorsi. Utilizzato in reti di media/grande dimensione.

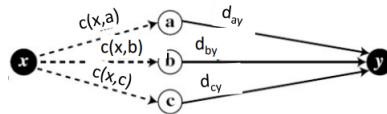
Algoritmi di Routing Globali/Decentralizzati Classifichiamo gli algoritmi di in-stradamento in:

- **Globali:** Basati sulla conoscenza dell'intera topologia di rete, il calcolo può essere fatto in un unico sito, utilizzando informazioni su tutti i nodi e gli archi. Ne è un esempio l'**algoritmo Link State**.
- **Decentralizzati:** Nessun nodo conosce l'intera topologia di rete, di conseguenza si distribuisce il calcolo e si reitera. Ne è un esempio l'**algoritmo Distance Vector**.

4.7.1 Algoritmo Distance Vector

L'algoritmo Distance Vector è distribuito, iterativo ed asincrono.

Equazione di Bellman-Ford Formula per trovare il percorso a costo minimo tra un nodo sorgente x ed una destinazione y attraversando nodi intermedi, detti vicini v .



- **Costo:** Definito un costo $c(x, v)$ tra nodo sorgente x e vicino v .
- **Distanza Minima d_{vy} :** Sia nota la distanza minima tra vicino v e destinazione y .

Da questo definiamo quindi:

$$d_{xy} = \min_v \{c(x, v) + d_{vy}\}$$

dove $c(x, v)$ è il costo da x al vicino v e d_{vy} è la distanza dal vicino v alla destinazione y .

Funzionamento Algoritmo Distance Vector L'algoritmo del Distance Vector si basa su queste fasi:

- Ogni nodo calcola il proprio vettore distanza e ne invia una copia a ciascuno dei suoi vicini.
- Quando un nodo riceve un vettore distanza, lo salva ed applica l'equazione di Bellman-Ford per aggiornare il proprio vettore distanza, quindi:

$$D_{xy} = \min_v \{c(x, v) + D_{vy}\} \text{ per ogni } y \in N$$

- Finché tutti i nodi continuano a cambiare i propri Distance Vector in maniera asincrona, ciascuna stima dei costi D_{xy} converge all'effettivo costo del percorso a costo minimo d_{xy} .

Per questo ad inizio sottocapitolo si definiva questo algoritmo distribuito, ciascun nodo infatti:

1. Attende un **cambiamento di costo** di un **arco locale** oppure un **messaggio da un vicino**.
2. **Ricalcola la stima** dei propri **Distance Vector** usando il Distance Vector ottenuto dal vicino.
3. Se il **proprio Distance Vector varia** allora lo **notifica ai vicini**.

Notifica Cambiamento Costo Questo algoritmo ha due diversi comportamenti in caso di variazione di costo degli archi:

- **Decremento - "Good News Travels Fast"**: Se il cambiamento del costo di un arco è un decremento, allora questa informazione è facilmente propagata ai nodi vicini.
- **Incremento - "Bad News Travels Slow"**: Se il cambiamento del costo è un incremento si ricade nella problematica del **count to infinity**, sono necessarie molte iterazioni prima che l'algoritmo si stabilizzi. Questa condizione si risolve con due approcci, che possono anche essere combinati:
 - **Split Horizon**: Tutte le rotte ricevute tramite un interfaccia **non devono** essere ritrasmesse a quell'interfaccia.
 - **Poisoned Reverse**: Se un nodo x trasmette pacchetti a v che sono diretti ad y allora li etichetta come irraggiungibili da x , passando ad esempio un etichetta ∞ .

4.7.2 Algoritmo Link-State

Tipo di **algoritmo globale**, la **topologia di rete** ed i suoi collegamenti sono **noti a tutti** i nodi. Si calcola quindi il cammino a costo minimo da un nodo origine a tutti gli altri nodi tramite l'utilizzo dell'**Algoritmo di Dijkstra**.

Notazione Algoritmo di Dijkstra Utilizzeremo questa notazione:

- $c(x, y)$: Costo dei collegamenti dal nodo x al nodo y , settato ad ∞ se non sono collegati.
- $D(v)$: Costo del cammino dal nodo origine alla destinazione v per l'iterazione corrente.
- $p(v)$: immediato predecessore di v lungo il cammino.
- N' : Sottoinsieme di nodi per cui il cammino a costo minimo dall'origine è definitivamente noto.

Complessità Algoritmo di Dijkstra Dati n nodi, per ogni iterazione è necessario controllare w non presenti in N . La complessità è quindi $O(n^2)$ ma esistono versione che scendono ad $O(n \log n)$. Da un punto di vista dei messaggi, ogni router deve trasmettere le informazioni sullo stato del proprio collegamento ad altri n router.

4.7.3 Confronto Distance Vector e Link State

Analizziamo alcune caratteristiche mettendo a confronto i due algoritmi:

- **Complessità Messaggi:**
 - **Distance Vector:** Richiede scambi tra nodi adiacenti.
 - **Link State:** Con n nodi richiede l'invio di $O(n^2)$ messaggi.
- **Velocità di Convergenza:**
 - **Distance Vector:** Può convergere lentamente, può presentare cicli nell'instradamento e problemi di conteggio all'infinito.
 - **Link State:** L'algoritmo $O(n^2)$ richiede $O(n^2)$ messaggi.
- **Robustezza:** Immaginiamo una condizione in cui un router funzioni male:
 - **Distance Vector:**
 1. Un nodo può comunicare cammini a costo minimo errati a tutte le destinazioni.
 2. La tabella di ciascun nodo può essere usata dagli altri.
 3. Un calcolo errato si può diffondere per l'intera rete.
 - **Link State:**
 - * Un router potrebbe inviare in broadcast un costo sbagliato per uno dei suoi collegamenti.
 - * In nodi si occupano di calcolare soltanto le proprie tabelle.

4.8 Sistemi Autonomi AS ed Instradamento Gerarchico

Sarebbe irrealistico immaginare un routing basato sulla visualizzazione di tutti i miliardi di hosts che compongono Internet. Tutta la rete è composta da sistemi autonomi **AS**, ossia un **gruppo di router** sotto lo **stesso controllo amministrativo**.

Tipi di AS Elenchiamo i vari tipi di sistemi autonomi:

- **AS Stub:** Collegato solo ad un altro stub, trasporta solo traffico di cui è origine o destinazione.
- **AS Multihomed:** Collegato a più di un altro AS.
- **AS di Transito.**

4.8.1 Intra/Inter-AS

Si differenzia il tipo di routing in base a se sorgente e destinazione facciano o meno parte dello stesso AS:

- **Intra Domain Routing:** Routing all'interno dello stesso AS, tutti i router al suo interno utilizzano lo stesso protocollo intra-domain. Un **gateway router** si trova al confine del suo AS ed ha i link con i router di altri AS.
- **Inter Domain Routing:** I gateway router si occupano dell'instradamento inter-domain, quindi tra AS diversi.

4.9 Protocolli Reali per Intra-AS

Elenchiamo alcuni protocolli reali di routing interno ad un sistema autonomo.

4.9.1 RIP - Routing Information Protocol

Protocollo di routing intra-dominio con le seguenti caratteristiche:

- Basato su Distance Vector con Poisoned Reverse.
- Gestione del **costo** basata sul **numero di hop** effettuati tra sottoreti per raggiungere la destinazione.
- Basato su richieste/risposte.
- Utilizza un processo demone che comunica tramite protocollo UDP con porta standard 520.
- Non viene inviata solo la distance vector ma tutta la tabella di routing.

Algoritmo del RIP Un nodo invia la tabella di inoltro ai propri vicini in maniera periodica, solitamente 30 secondi. Si definisce quindi il costo di hop tra R ed Y con:

$$D_R[Y] = 1 + D_V[Y]$$

dove

- $D_R[Y]$: Corrisponde alla nuova distanza tra R sorgente ed Y destinazione.
- $D_V[Y]$: Corrisponde alla distanza che il nodo V afferma di avere per raggiungere la destinazione Y .
- $+1$: Costo causato dal collegamento tra i due nodi vicini, ossia R e V .

Quindi il nuovo percorso da R ad Y viene inserito in tabella se è un percorso non presente, ha un costo inferiore rispetto a quello del vecchio percorso o se V è nextHop del vecchi e nuovo percorso, ma con costo aggiornato.

Prestazioni RIP Così come per l'algoritmo DV abbiamo una potenziale convergenza lenta delle tabelle di inoltro, specialmente in domini ampi. Allo stesso tempo però si garantisce robustezza in caso di guasto di un router a causa della sua logica distribuita.

Protocollo RIP e Utilizzo UDP Nella precedente pagina abbiamo definito come questo protocollo a livello di rete si basi sull'utilizzo dell'UDP, che è un protocollo di trasporto. Non stiamo però nell'ambito del data plane ma del control plane, quindi non siamo limitati solo all'utilizzo del livello di rete, ma possiamo anche arrivare fino al livello applicazione. Quindi non si è limitati solo al livello di rete nei router per il Control Plane, proprio perchè il compito è la comunicazione per la creazione delle regole, e non per l'inoltro effettivo.

4.9.2 OSPF - Open Shortest Path First

Basato sul Link State, in maniera periodica viene effettuata una operazione di **flooding** che permette la propagazione dello stato globale della rete.

Caratteristiche OSPF Elenchiamo alcune caratteristiche di questo protocollo:

- Il payload OSPF viene incapsulato in datagrammi IP.
- Questo protocollo è utilizzato soprattutto in reti grandi, per evitare quindi di inondare tutta la rete di messaggi si effettua l'invio in maniera globale ma relativamente ad un area/dorsale della rete totale.

Prestazioni OSPF Definiamo le prestazioni di questo protocollo:

- Il formato dei payload e il metodo di trasmissione a flooding potrebbero generare molto traffico e sprecare banda.
- La convergenza delle tabelle di inoltro è molto più rapida del RIP.
- La robustezza è garantita dall'indipendenza ci ciascun router.

4.10 Protocolli Reali per Inter-AS

Elenchiamo alcuni protocolli reali di routing tra sistemi autonomi diversi.

4.10.1 BGP - Border Gateway Protocol

Permette ai Gateway di diversi AS di interfacciarsi secondo un protocollo comune, dando la possibilità a ciascuna AS di pubblicizzare la propria esistenza.

In generale le operazioni permesse da BGP sono ad esempio comunicare la propria esistenza, ottenere informazioni riguardanti le reti raggiungibili di AS vicini, determinare percorsi verso altre reti basandosi su politiche di raggiungibilità, propagare alle reti vicine informazioni di raggiungibilità.

Oltre a questo gestisce in maniera aggregata gli indirizzi e distribuisce **informazioni di raggiungibilità** in due modi:

- **eBGP - external BGP:** Gestione dello scambio di informazioni di routing tra AS diversi, tramite comunicazione dei router Gateway con altri router Gateway di AS esterni.
- **iBGP - internal BGP:** Gestione della distribuzione interna delle informazioni sulle rotte apprese dall'esterno dal Gateway del AS locale.

Messaggi BGP Due router BGP, detti peer, si scambiano messaggi BGP su una connessione TCP semi-permanente, ciascun nodo mantiene quindi informazioni sul percorso completo verso una destinazione, non solo sulla distanza o sulla prossima tappa. Esistono vari tipi di messaggi BGP, ad esempio OPEN, UPDATE, KEEP ALIVE, NOTIFICATION.

Annunci ed Attributi BGP Definiamo elementi principali di questo meccanismo:

- **Annuncio BGP:** Messaggio centrale che un router BGP invia ai vicini per comunicare una rottura verso una o più destinazioni.
- **Prefisso CIDR⁷:** Identifica la destinazione raggiungibile come blocco di indirizzi IP rappresentato in notazione CIDR.
- **Attributo AS-PATH:** Lista logica ordinata degli AS attraversati per raggiungere il corrente router.
- **Attributo NEXT-HOP:** Indirizzo IP del prossimo router fisico che il pacchetto deve raggiungere.

Selezione Rotta con BGP Un router potrebbe conoscere più di un percorso verso l'**AS destinazione**, quindi questo viene selezionato in base a vari elementi come una **preferenza locale** assegnata alle rotte, l'**AS-PATH più breve**, il **NEXT-HOP più vicino** oppure **altri criteri aggiuntivi**.

⁷Classless Inter Domain Routing

Policy Based Routing Il gateway che riceve l'annuncio di una rotta usa una import policy per accettare/rifiutare un percorso. Le policy vengono utilizzate anche per pubblicizzare o meno un percorso ad altri AS vicini.

Riassunto Caratteristiche Protocolli Inter/Intra AS Elenchiamo le caratteristiche mostrate negli ultimi due capitoli:

- **Politiche:**

- **Inter-AS:** Si vuole che l'amministratore abbia controllo su come viene instradato il suo traffico.
- **Intra-AS:** Singolo amministratore, la gestione della politica è meno problematica.

- **Scalabilità:** L'instradamento gerarchico consente di ridurre le dimensioni delle tabelle e del traffico di aggiornamento.

- **Performance:**

- **Inter-AS:** Si tiene conto delle performance.
- **Intra-AS:** Le politiche hanno maggiore importanza delle performance.

4.11 Protocollo IPv6

Nato per soddisfare la potenziale mancanza di indirizzi generati da IPv4, a causa di un forte utilizzo del Natting non è stato utilizzato in larga scala come ci si aspettava.

Caratteristiche Questo protocollo si basa su queste caratteristiche:

- Indirizzi a 128 bit.
- Lunghezza dell'header fissa a 40 byte.
- Eliminazione della checksum.
- Nessuna frammentazione.

Tunneling Dato che non è possibile switchare da IPv4 ad IPv6 si definisce un meccanismo di "convivenza" tra i due protocolli, grazie al quale si wrappa tutto il datagramma IPv6 nel payload del datagramma IPv4 in modo tale che questo possa essere trasmesso in una rete IPv4.

