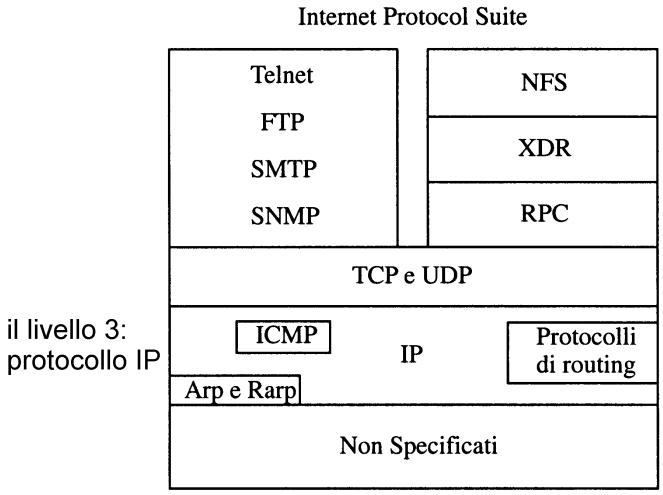
Il livello Network del TCP/IP.

Il protocollo IP (versione 4)

L'architettura TCP/IP (il cui nome più preciso è **Internet Protocol Suite**) è formata da diversi componenti, che si posizionano nello stack dei protocolli a partire dal livello 3 (network).
I protocolli appartenenti a questa architettura sono specificati tramite standard denominati RFC (Request For Comments) disponibili in rete. Ad es. l'RFC 791 specifica il protocollo IP.



Il protocollo IP (**Internet Protocol**) è il protocollo principale del livello 3 dell'architettura TCP/IP. Si tratta di un protocollo semplice, di tipo datagram ovvero **senza connessione** e **non affidabile**;

- **non affidabile**: il pacchetto inviato può essere perso, duplicato ritardato o consegnato fuori sequenza, ma il protocollo IP non informerà nè il trasmettitore nè il ricevitore.
- **senza connessione**: ogni pacchetto viene trattato in maniera indipendente dagli altri, pacchetti diversi aventi stesso mittente e stesso destinatario possono seguire percorsi diversi, alcuni possono essere consegnati ed altri no. Se le risorse della rete lo consentono il pacchetto viene portato a destinazione, in caso contrario verrà scartato.

Il protocollo IP versione 6

Attualmente lo standard per il livello network dello stack TCP/IP è rappresentato dal protocollo IP versione 4, ma già dal 1995 è stata proposta una nuova versione del protocollo IP, nota col nome di IP versione 6 (RFC 1883: Internet Protocol, Version 6 (IPv6) Specification, December 1995 R. Hinden, S. Deering).
IP versione 4 versione soffre di almeno 3 problemi principali, che IPv6 vuole correggere:

• **numero degli indirizzi IP disponibili ormai insufficiente**: gli indirizzi IP sono composti da 4 byte (32 bit) e a causa del grande incremento del numero degli hosts nel mondo la disponibilità degli indirizzi è in forte calo. Il protocollo IPv6 prevede invece indirizzi formati da 16 byte (128 bit) e quindi rende disponibili un numero enorme di indirizzi.

• **traffico gestito esclusivamente in modo best-effort**: in IPv4 tutti i pacchetti sono trattati allo stesso modo dai router, anche se esiste nell'header dei pacchetti IPv4 un campo priorità, che però non viene utilizzato. Con IPv6 si vuole definire delle classi di servizio a cui assegnare priorità diverse. Si vuole anche gestire la comunicazione con un meccanismo simile ad un protocollo con connessione, cioè implementando un flusso di dati.

• **sicurezza**: in IPv6 saranno rese standard e disponibili alcune primitive per l'autenticazione e la cifratura dei dati.

Nonostante queste prospettive, il protocollo IP versione 6 è ancora poco diffuso, e rimane ancora a livello di sperimentazione, forse perchè l'adozione del nuovo protocollo costringerebbe a modificare fortemente gli apparati di rete esistenti, con un grande dispendio di denaro. Esistono a tuttoggi, in un oceano IPv4, solo delle **isole** in cui si parla IPv6.

Nel seguito parleremo di IPv4, che rappresenta ancora lo standard più diffuso.

Funzioni del protocollo IPv4

Il protocollo IP svolge le seguenti funzioni:

• **distingue** ogni hosts, o meglio **ogni scheda di rete mediante un identificatore**, detto **indirizzo IP**. Un indirizzo IP di tipo single o **unicast** identifica un unico host, ma uno stesso host può avere più indirizzi IP unicast, tanti quante sono le schede di rete che possiede. Si parla allora di MultiHomed Systems. Ad es. i router hanno più indirizzi, perchè dovendo fungere da centri di smistamento dispongono di più schede di rete. Un host può comunque disporre di più schede di rete anche senza essere un router, cioè anche se non effettua un servizio di instradamento per pacchetti destinati ad altri hosts, ma dovrà prevedere una politica che definisca quale scheda di rete utilizzare per inviare i dati.

• **riceve i dati** (una sequenza di PDU) **dal livello trasporto** (4).

• **incapsula ciascuna PDU ricevuta in pacchetti** di dimensione massima 64 Kbyte (normalmente circa 1500 byte), **aggiungendovi un proprio header** (o intestazione).

header del Datagram IP	area dati del Datagram IP (PDU del livello trasporto)

• eventualmente **frammenta i pacchetti** all'inizio o durante il trasporto, per inserirli nei frame di livello 2.

• **instrada i pacchetti** sulla rete,

• **effettua la rilevazione**, non la correzione, **degli errori**,

• **alla destinazione**, se necessario, **riassembla i frammenti** ricostruendo i pacchetti originali,

• **estrae dai pacchetti i dati** (PDU) **del livello trasporto**,

• **consegna al livello trasporto i dati nell'ordine in cui sono arrivati a destinazione**, che può essere diverso dall'ordine in cui sono partiti.

Il pacchetto IPv4: lo HEADER

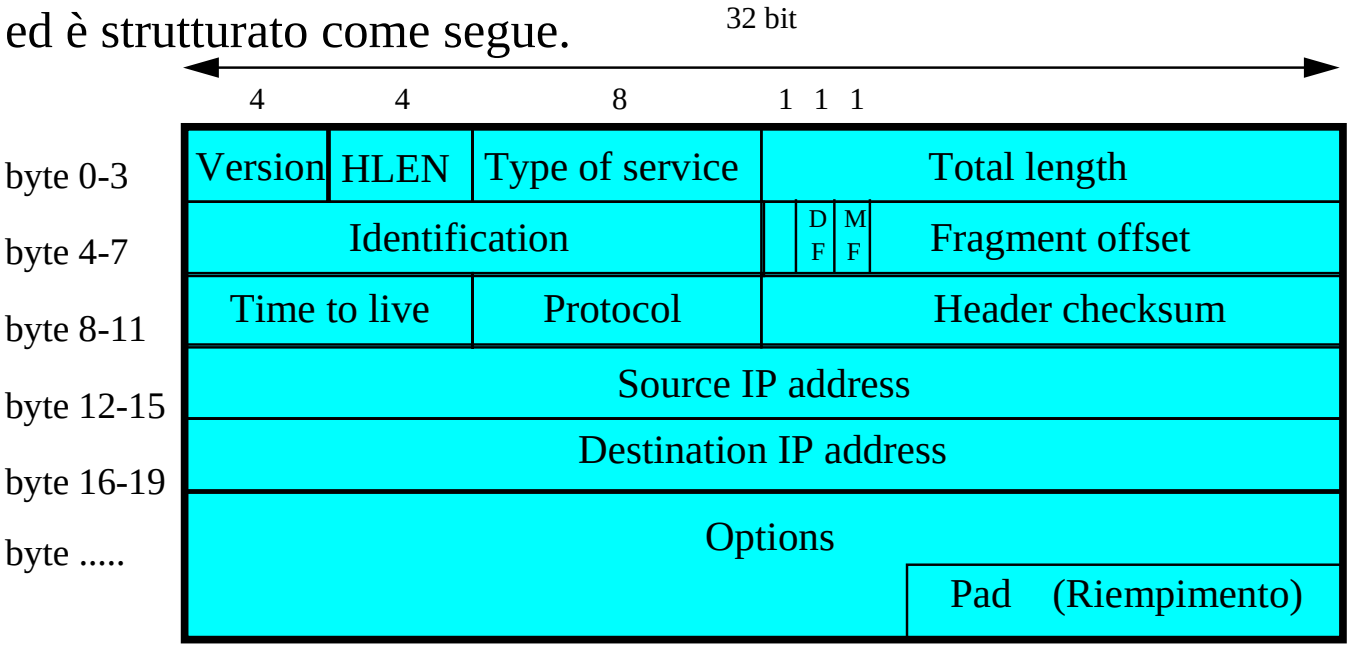
Un pacchetto IP è costituito da un header e da una parte dati, che rappresenta la porzione di dati del livello trasporto da trasferire.

header del Datagram IP	area dati del Datagram IP (PDU del livello trasporto)

L'header IP ha:

- una prima parte fissa di 20 bye,
- una seconda parte, opzionale, di lunghezza variabile, ma sempre multiplo di 4 byte

ed è strutturato come segue.



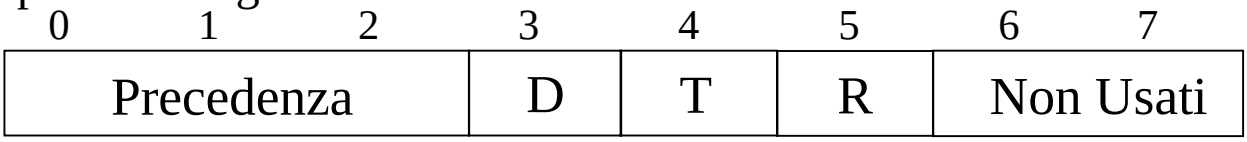
• Il campo **Version** (di 4 bit) indica la versione del protocollo IP che ha generato il pacchetto. Serve al ricevente per capire il formato del pacchetto che stà ricevendo. Se lo standard cambia, ad es. il passaggio da IPv4 (Version=4) ad IPv6 (version=6) il ricevente capisce da questo campo come deve trattare il pacchetto.

• Il campo **HLEN** (anch'esso di 4 bit) indica la lunghezza dell'Header IP misurata in **word di 4 byte**. Se HLEN vale 7 l'header è lungo 4*7=28 byte. Tutti i campi dell'header hanno lunghezza fissa, tranne il campo Options, di lunghezza variabile, ed il corrispondente campo PAD (Riempimento) che serve a rendere l'header lungo un multiplo di 4 byte.

HEADER IPv4 (2)

• Il campo **Total Length** (di 16 bit) indica la lunghezza totale del pacchetto IP, **espressa in byte**, e comprende sia l'header che il campo dati. Quindi se Total Length = 50 il pacchetto IP è lungo 50 byte. Poichè il campo è lungo 16 bit la massima lunghezza possibile per un pacchetto IP è di 2^{16}-1=65535 byte.

• Il campo **Type of Service** (di 8 bit) rappresenta un'**indicazione ai router sulla qualità del trasporto che possibilmente il pacchetto IP dovrebbe sperimentare**. Il router dovrebbe basarsi anche su queste indicazioni per decidere la precedenza dei pacchetti nelle sue code, e l'instradamento. Ad es. se il pacchetto richiede un certo ritardo massimo, il router potrebbe decidere di instradarlo su un percorso attraverso una rete ATM a cui chiedere garanzie sul ritardo massimo piuttosto che attraverso una rete con servizio di tipo Best Effort che non può offrire garanzie.



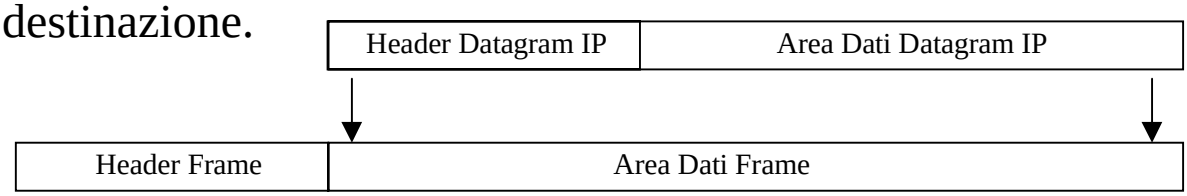
- I primi 3 bit definiscono un campo **Precedenza** con valori da 0 a 7, al crescere del quale cresce l'importanza del pacchetto IP. Il valore 0 indica precedenza normale, il valore 7 un pacchetto di controllo della rete, e quindi maggiore importanza.
- I bit 3 (D=Delay), 4 (T=Throughput) e 5 (R=Reliability) indicano il tipo di trasporto preferito: D settato (D=1) indica richiesta di basso ritardo, T settato chiede un elevato throughtput (ampia banda di trasmissione), R settato indica richiesta di massima affidabilità.
- I bit 6 e 7 non sono usati.

Di solito però i router non tengono conto delle preferenze espresse mediante il campo Type of Service. Attualmente si studiano delle politiche (note come **Differentiated Services**) che cercano di realizzare routing e buffering basandosi sulla classificazione delle applicazioni proprio in base a indicazioni di questo tipo.

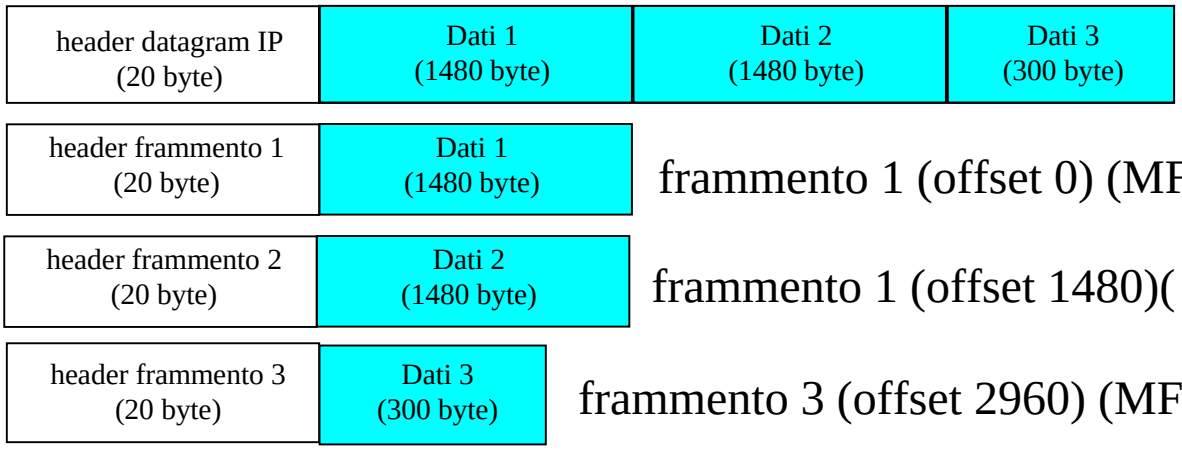
La Frammentazione dei pacchetti IP

L'hardware di ogni tipo di rete impone un limite superiore alla dimensione del frame di livello 2, e quindi anche alla quantità di dati di livello 3 che possono essere trasportati in un unico frame a livello 2. La dimensione massima di dati di livello 3 che possono essere trasportati in un frame del Data Link viene chiamata **Massima Unità di Trasferimento** (**MTU**: **Maximum Transfer Unit**), ed è caratteristico di ogni tipologia di rete. Per Ethernet MTU=1500 bytes.

• Se la porzione dati di un datagram IP (più la dimensione dell'header IP) è più piccola della MTU della rete sottostante, il datagram IP potrà essere inserito completamente in un frame di livello 2 e inviato a destinazione.



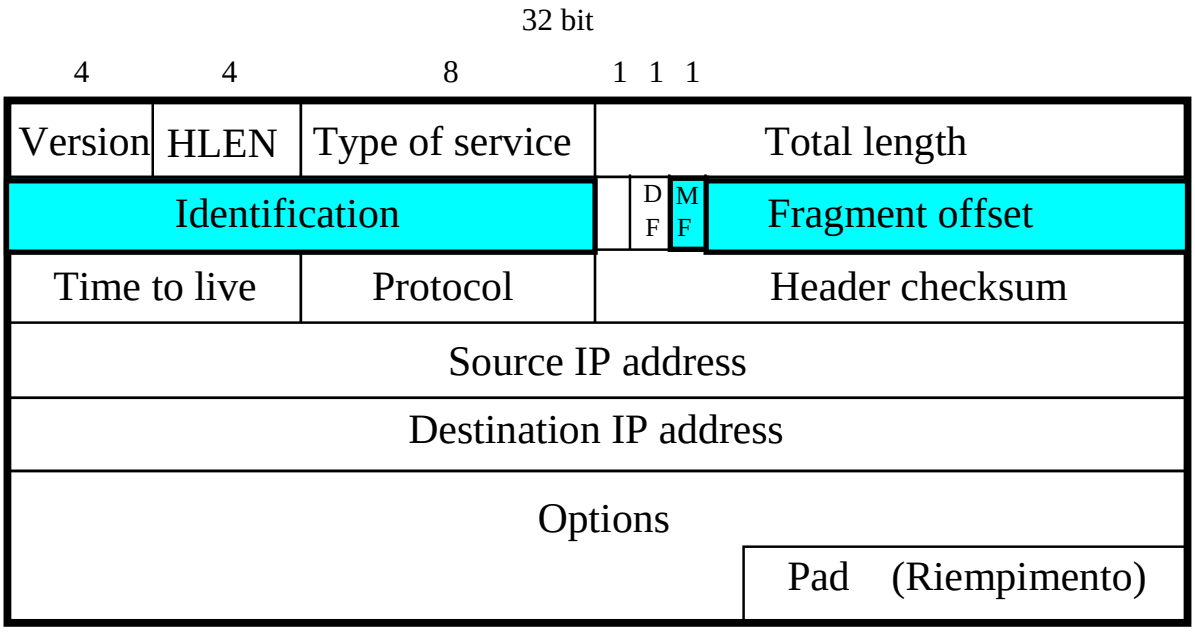
• Se invece la porzione dati del datagram IP (più la dimensione dell'header IP) è più grande della MTU della rete sottostante, **la porzione dati** dovrà essere spezzata in più pezzi che verranno incapsulati in datagram IP (detti frammenti) più piccoli della MTU, e ciascun frammento dovrà essere inserito in un frame diverso e verrà spedito separatamente dagli altri verso la destinazione finale, dove il protocollo IP provvederà a rimettere assieme i diversi frammenti e a ricostituire il datagram originale. Nell'esempio un pacchetto IP con 3260 byte di dati frammentato per una MTU di 1500 byte.



La Frammentazione dei pacchetti IP

il problema della frammentazione si propone ogni volta che nel percorso seguito dai pacchetti IP, si deve attraversare una porzione di rete avente una MTU minore della porzione di rete precedentemente attraversata, sempre se la dimensione dei pacchetti IP è maggiore della MTU più piccola.
Il router preleva allora la porzione dati del datagram IP e lo spezza in più porzioni, in modo che ciascuna (aggiungendovi l'header) stia in un frame, e in modo che ogni frammento dei dati, tranne l'ultimo, abbia dimensione multipla di 8 byte, perchè così è definito il campo offset dell'header IP.
L'ultimo pezzo in genere sarà il più corto e verrà identificato come ultimo settandovi a zero il flag **MoreFragment**, ad indicare che è l'ultimo frammento. Negli altri frammenti **MF** è settato a 1.

Il protocollo IP usa tre campi dell'header per controllare il meccanismo della frammentazione e permettere il riassemblaggio dei datagram frammentati. Questi campi sono **Identificatio**n (16 bit), **Fragment Offset** (15 bit) e il flag **More Fragment** (**MF**).



La Frammentazione dei pacchetti IP

L'header del datagram originale verrà copiato integralmente nei frammenti (con qualche modifica per il campo Options) e in più verrà cambiato il campo **Fragment Offset che indica il punto dell'area dati del datagram originale in cui comincia la porzione di dati trasportata nel frammento**. Tale offset è pensato come un multiplo di 8 byte. Se l'offset indica ad es. 185, il frammento porta la porzione di dati che inizia nella posizione 185*8=1480 byte.

• **Tutti i frammenti** sono caratterizzati dall'avere lo stesso identificatore (il campo **Identification**) del datagram originale. Tale numero è assegnato univocamente dal trasmettitore (che mantiene un contatore dei datagram IP trasmessi), e la coppia (IP Provenienza, Identification) rende univocamente identificabile un certo datagram IP, e tutti i suoi frammenti.

Il Riassemblaggio dei Frammenti

Dopo la frammentazione, ogni frammento viaggia separatamente dagli altri fino alla destinazione finale. Solo alla fine del viaggio avrà luogo il riassemblaggio dei frammenti, nel tentativo di ricostruire il datagram Originale.

Il ricevitore riconosce di avere ricevuto un frammento e non un datagram intero in due modi:
• il pacchetto IP ricevuto ha un **offset uguale a zero**, ma ha il flag **More Fragment settato ad uno** (è il primo frammento).
• il pacchetto IP ricevuto ha un **offset diverso da zero** (è un frammento successivo). Se il More Fragment è zero è l'ultimo frammento.

Il protocollo IP del ricevente identifica univocamente i frammenti di uno stesso datagram mediante la coppia (**IP trasmettitore, Identification del datagram**).

Il ricevente non conosce la dimensione del datagram originale perchè ogni frammento mantiene nel campo Total Length la lunghezza del frammento stesso, e non quella del datagram originale. Solo quando riceverà il frammento con il flag **More Fragment settato a zero**(che indica l'ultimo frammento del datagram originale), si potrà capire la dimensione totale del datagram originale sommando all'offset dell'ultimo frammento la lunghezza dei dati trasportati nell'ultimo frammento.

Se un solo frammento viene perso, è impossibile ricostruire il datagram IP originale.

Per evitare di aspettare inutilmente un frammento perso, il ricevitore nel momento in cui riceve un primo frammento inizializza un timer. Se il timer scade prima che tutti i frammenti siano giunti a destinazione il ricevitore butta via tutti i frammenti.

HEADER IPv4 (continua 3)

• Il campo **Time To Live** (TTL, tempo di vita, di 8 bit) indica in modo approssimato il tempo, in secondi, che un pacchetto IP può rimanere all'interno di una rete prima di essere scartato.
-Tale valore viene inizializzato a 255 dall'host che spedisce il pacchetto IP. Ogni volta che un router processa quel pacchetto IP, decrementa di una unità questo contatore. Quando il contatore raggiunge il valore zero, il router scarta il pacchetto.
-Per ovviare al problema della congestione della rete che causa lunghe attee in coda, quando il router riceve un pacchetto e lo mette in coda in attesa per spedirlo, salva il valore corrente dell'orologio.
-Quando il pacchetto deve essere spedito si calcola il tempo (in secondi) trascorso in coda e si decrementa di questo tempo il contatore TTL.
-Grazie al contatore TTL, i pacchetti IP non possono viaggiare in eterno nella rete anche se per un errore i router li instradano in un percorso ciclico, e si evitano così congestioni nella rete.
-Quando un pacchetto IP viene frammentato durante il percorso, tutti i suoi frammenti vengono incapsulati con il TTL residuo del pacchetto.

• Il campo **Protocol** (di 8 bit) indica di quale tipo è il dato trasportato nell'area dati del pacchetto IP, ovvero indica qual'è il protocollo di livello 4 (o 3) che ha generato i dati trasportati dal pacchetto IP. In tal modo il livello Network sa a quale protocollo dovrà consegnare i dati trasportati. Tra i protocolli che possono essere trasportati nell'area dati ricordiamo:

•0	ICMP	Internet Control Message Protocol
•4	IP	IP in IP (incapsulamento, tunneling)
•6	TCP	Transmission Control Protocol
•17	UDP	User Datagram Protocol
•29	ISO-TP4	ISO Transport Protocol Class 4
•93	AX.25	AX.25 frames

HEADER IPv4 (4)

• Il campo **Header Checksum** (di 16 bit) serve a verificare che **lo header IP sia arrivato integro a destinazione**. Viene codificato utilizzando i byte del solo header. Se durante il trasporto l'header subisce delle modifiche, la checksum risulta diversa e il protocollo IP capisce che c'è stato un errore.
Notare che la checksum verifica l'integrità del solo header, e non dei dati trasportati.
Il **vantaggio di avere una checksum separata per header** e dati è che:

- i protocolli di livello superiore possono scegliere una loro codifica per il controllo degli errori,
- i router diminuiscono il tempo necessario a calcolare la checksum, perchè devono processare solo l'header, che di solito è più piccolo dei dati trasportati.

Di contro, lo svantaggio di avere a livello IP una checksum solo per l'header IP impedisce al livello 3 di accorgersi di eventuali errori sui dati di livello 4, fino a che tali dati non sono giunti a destinazione finale, e solo allora il protocollo di livello 4 effettuerà il controllo sui dati con delle proprie checksum. Però il livello 2 potrebbe avere già riconosciuto un errore e scartato il frame.

• I campi **Source IP address** e **Destination IP address** contengono gli indirizzi IPv4 a 32 bit della provenienza originale e della destinazione finale del datagramma IP. Tali indirizzi quindi non cambiano mai durante tutto il percorso, comunque venga instradato il pacchetto, e comunque venga frammentato.

• Il campo **Options**, è di lunghezza variabile, e poichè l'header IP deve avere lunghezza pari ad un multiplo di 4 byte, viene introdotto ove necessario un ultimo campo **Padding** di riempimento, per arrivare alla giusta lunghezza.

HEADER IPv4 (5)

Il campo **Options** non è necessario in tutti i datagram IP. Le opzioni sono utilizzate allo scopo di testare la funzionalità della rete sottostante. Evitiamo di addentrarci su come sono organizzati i sottocampi del campo Opzioni IP, e analizziamone soltanto le funzionalità previste. Essenzialmente le opzioni sono classificabili in 3 categorie:

1) Opzione di registrazione del percorso.

• Quando il trasmettitore setta l'opzione di registrazione del percorso, indica il numero massimo di hop che vuole memorizzare e crea spazio sufficiente nel campo opzioni per memorizzare tali hop, 32 bit per ogni indirizzo IP da memorizzare.

• Quando il pacchetto IP viaggia per la rete, ogni router toccato dal datagram IP aggiunge il proprio indirizzo IP alla lista di registrazione del percorso, almeno fino a che tale lista non è piena, nel qual caso il router si limita ad inoltrare il messaggio.

• Quando il pacchetto IP giunge alla destinazione finale, il protocollo IP, se vuole, può estrarre la lista dei router toccati dal pacchetto.

•Questa opzione viene utilizzata ad es. per implementare l'applicazione detta "traceroute" che visualizza i router toccati da un pacchetto.

2) Opzioni di Instradamento di Provenienza.

Le opzioni di Instradamento di Provenienza, consentono al trasmettitore di imporre ad un pacchetto IP un certo percorso attraverso la rete, anche se i router normalmente sceglierebbero un percorso diverso. Ciò può essere utile per effettuare dei test sulla rete. Naturalmente per imporre l'instradamento è necessario conoscere la topologia della rete.

Esistono due modalità per l'instradamento di provenienza. la prima, detta **instradamento di provenienza severo** specifica una sequenza di salti consecutivi, e causa errore se due router non sono consecutivi nella rete, cioè non stanno sulla stessa rete fisica o se il router non può seguire quel percorso. La seconda detta **instradamento di provenienza permissivo** specifica una sequenza di indirizzi IP, ma consente di attraversare più reti tra due indirizzi consecutivi.

HEADER IPv4 (6)

3) Opzione di contrassegno temporale.
E' simile all'opzione di registrazione del percorso, ma aggiunge all'indirizzo IP di ogni router attraversato anche la data e l'ora in cui il router gestisce il datagram IP, espresso secondo l'ora del meridiano di Greenwich.

• **Elaborazione delle Opzioni durante la frammentazione**.
Ciascuna Opzione IP, viene identificata mediante un byte nel campo Opzioni. Il primo bit (detto bit COPIA) di questo byte stabilisce, quando posto ad 1, che l'opzione deve essere copiata in tutti gli eventuali frammenti del pacchetto IP. In caso contrario l'opzione verrà copiata solo in uno dei frammenti.
Questo diverso comportamento viene configurato in modo diverso per le diverse Opzioni IP.

• Per l'opzione di registrazione del percorso, si vuole che l'opzione sia copiata in uno solo dei frammenti, perchè essendo ogni frammento gestito separatamente, potrebbe seguire percorsi diversi verso la destinazione. Si avrebbero così più liste di registrazione del percorso potenzialmente diverse. Il flag COPIA viene perciò posto a zero, e l'opzione copiata in uno solo dei frammenti.

• Al contrario, per l'opzione di instradamento di provenienza si vuole che tutti i frammenti seguano lo stesso percorso, e quindi il flag COPIA viene posto ad uno.

INDIRIZZAMENTO IP

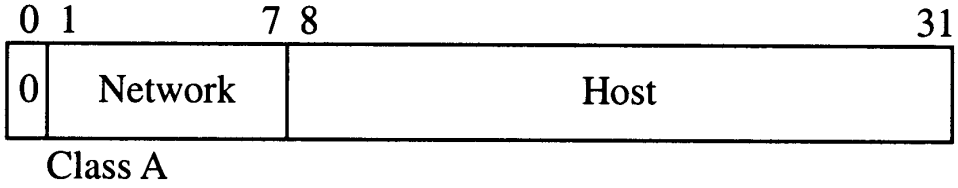
Gli indirizzi IP, che devono essere univoci sulla rete, sono lunghi 32 bit (quattro byte) e sono espressi scrivendo i valori decimali di ciascun byte separati dal carattere punto.

• Gli indirizzi IP comprendono due o tre parti. La prima parte indica l'indirizzo della rete (**network**), la seconda (se presente) quello della sottorete (**subnet**) e la terza quello dell'**host**.

• Si noti che non sono gli hosts ad avere un indirizzo IP, bensì le interfacce di rete. Quindi se un nodo ha tre interfacce, esso ha tre indirizzi IP. Poichè la maggior parte degli hosts ha una sola interfaccia di rete, parlando di indirizzo IP di un host si fa riferimento all'indirizzo della sola interfaccia di rete presente.

Gli indirizzi IP sono suddivisi in cinque classi denominate A, B e C, che differiscono per il numero di host che ciascuna rete può indirizzare, e altre due classi D ed E assai differenti:

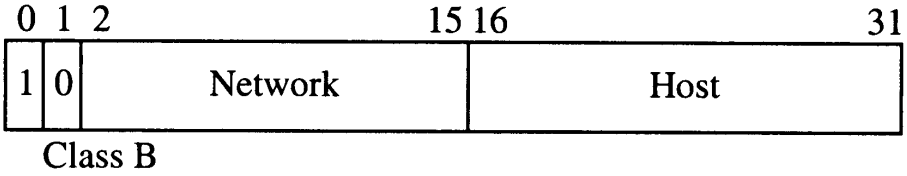
- **Classe A**. Sono concepiti per poche reti di dimensioni molto grandi. Gli indirizzi di classe A sono riconoscibili in quanto il il bit più significativo del primo byte è posto a zero, e quindi il primo campo dell'indirizzo è compreso tra 0 e 127. I bit che indicano la rete sono 7 e quelli che indicano l'host 24. Quindi si possono avere al massimo 128 reti di classe A, ciascuna con una dimensione massima di circa 16 milioni di indirizzi.



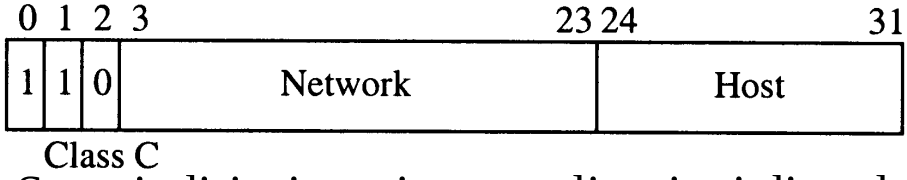
Class A

INDIRIZZAMENTO IP (2)

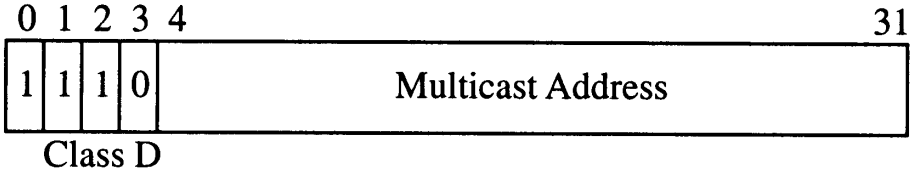
- **Classe B**. Sono pensati per un numero medio reti di dimensioni medio-grandi. Gli indirizzi di classe B si riconosconno perchè i 2 bit più significativi del primo byte sono posti a 10, quindi il primo campo dell'indirizzo è compreso tra 128 e 191. I bit che indicano la rete sono 14 e quelli che indicano l'host 16. Si possono avere circa 16000 reti di classe B, ciascuna con circa 64000 indirizzi.



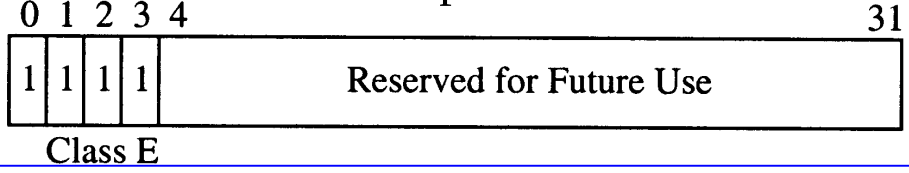
- **Classe C**. Sono concepiti per avere reti di dimensioni piccole. Gli indirizzi di classe C sono riconoscibili perchè i primi 3 bit sono settati a **110** e quindi il primo campo dell'indirizzo è compreso tra **192 e 223**. I bit che indicano la rete sono 21 e quelli che indicano l'host 8. Quindi si possono avere al massimo **2 milioni di reti** di classe C, ciascuna con una dimensione massima di **256** indirizzi.



- **Classe D**. Sono indirizzi usati per applicazioni di multicast . Gli indirizzi di classe D si riconoscono perchè i primi 4 bit del primo byte sono settati a 1110, e quindi il primo campo dell'indirizzo è compreso tra 224 e 239.

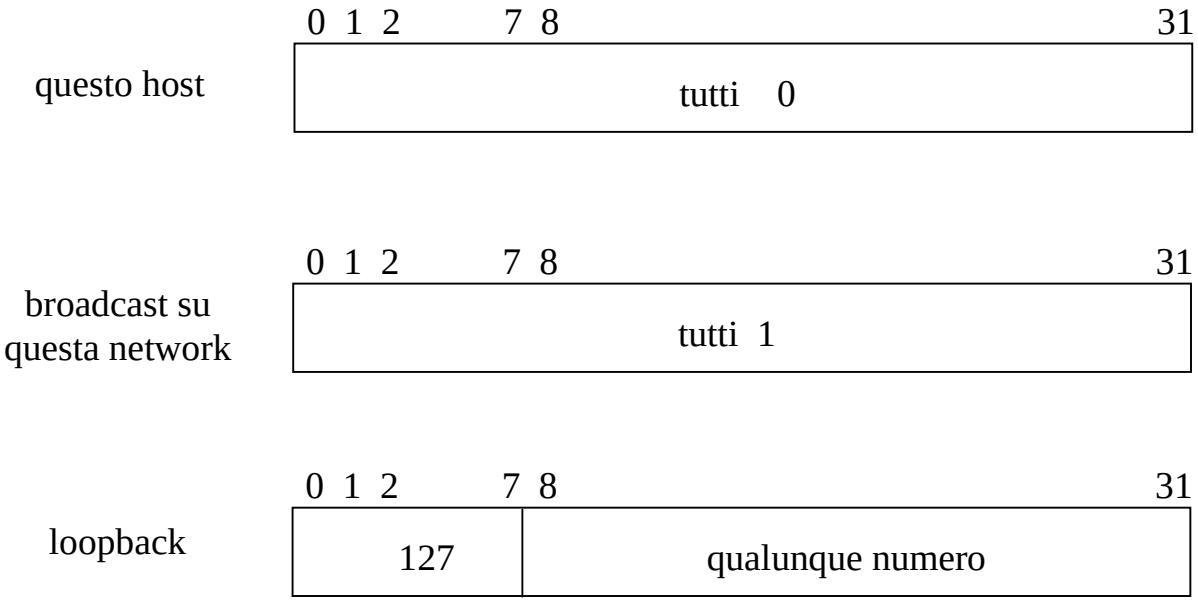


- **Classe E**. Questi indirizzi sono riservati per usi futuri. Gli indirizzi di classe E hanno i 4 bit più significativi settati a 1111, e quindi il primo campo dell'indirizzo è compreso tra 240 e 255.



INDIRIZZAMENTO IP (3)

Esistono inoltre indirizzi IP con significato particolare, ad esempio per gli indirizzi di broadcast e per il loopback.



Quando si utilizza il loopback, il pacchetto non viene inviato sulla rete ma viene elaborato come se fosse in arrivo: ciò serve per effettuare localmente dei test su un software di rete in fase di sviluppo.