

Università di Roma Tor Vergata
Corso di Laurea triennale in Informatica

Sistemi operativi e reti

A.A. 2017-18

Pietro Frasca

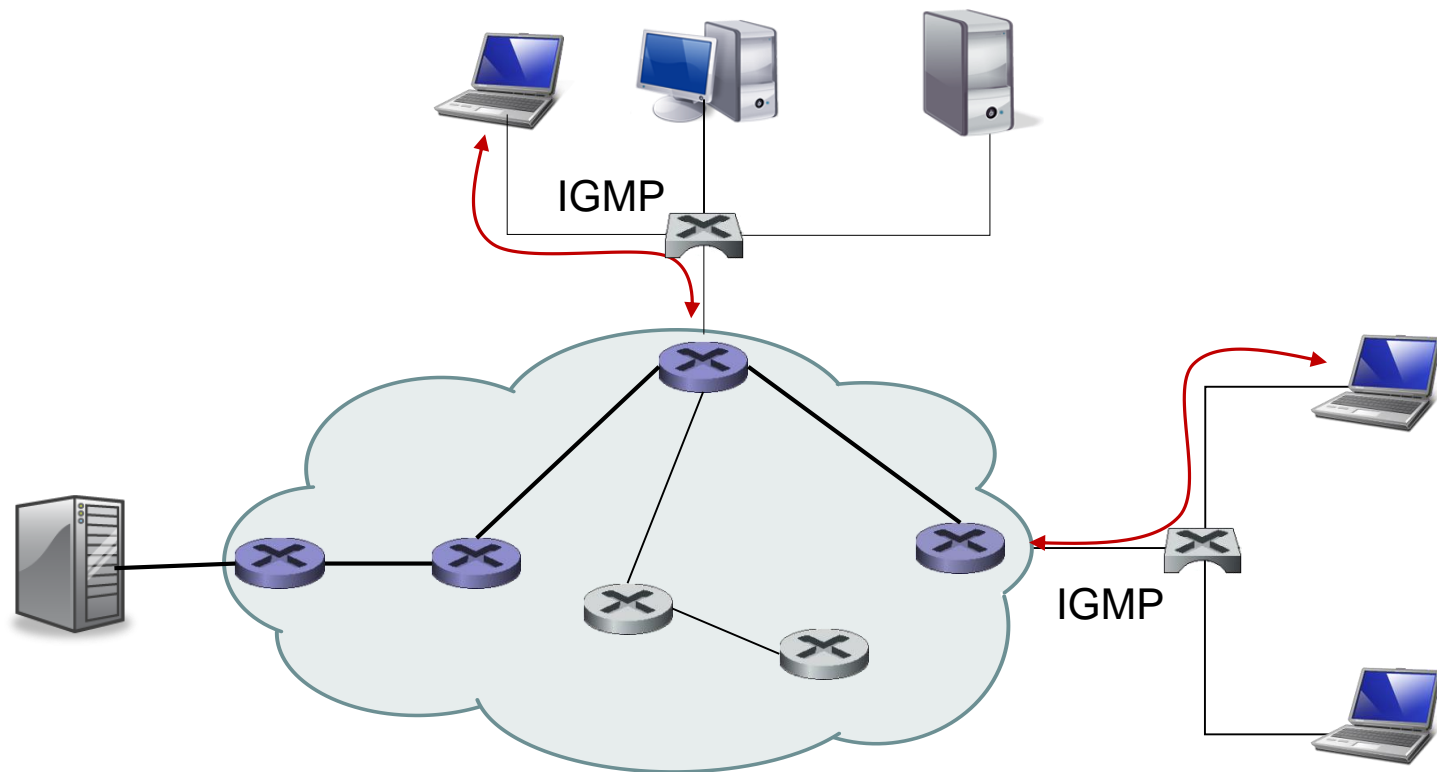
Parte II: Reti di calcolatori

Lezione 20 (44)

Giovedì 24-05-2018

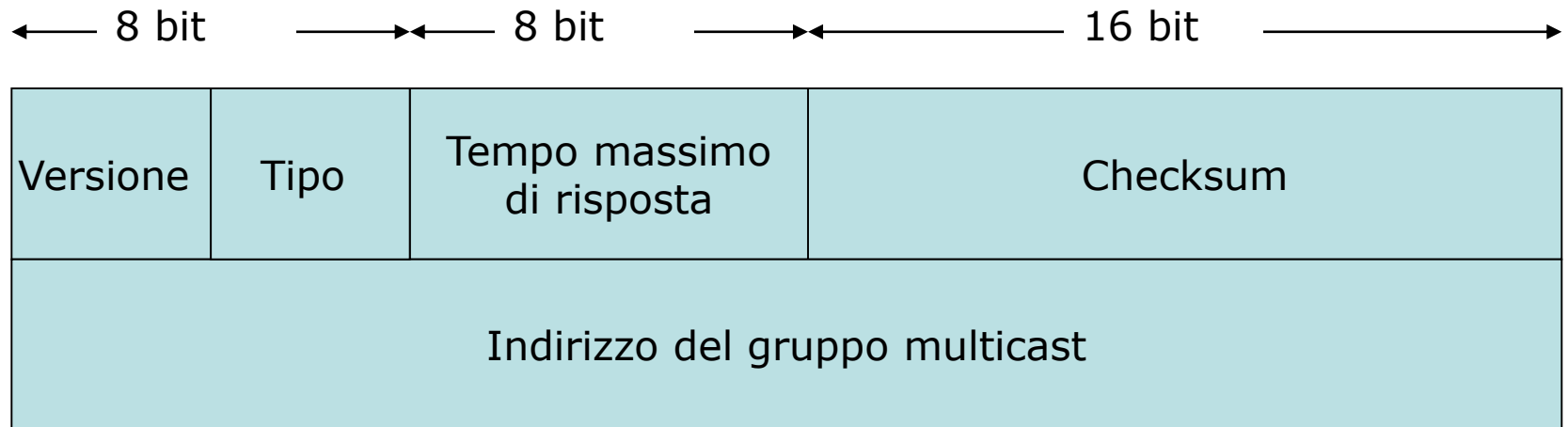
Il protocollo IGMP

- Il funzionamento del multicast dello strato di rete si basa su due componenti: **l'IGMP** e i **protocolli di instradamento multicast** (PIM, DVMRP e MOSFP).
- Il protocollo **IGMP (*Internet Group Management Protocol*, Protocollo Internet Gestione Gruppi)** consente a un processo di rete, in esecuzione su un host, di comunicare al suo router di default che vuole associarsi a un gruppo multicast.
- Poiché IGMP funziona soltanto tra un host e il suo router di default, per creare un percorso multicast tra i router, compresi tra l'host mittente e l'host destinatario, è necessario un altro protocollo per instradare i datagrammi multicast. Questa funzionalità è svolta dagli **algoritmi di instradamento multicast** dello strato di rete (*network layer multicast routing algorithm*) come **PIM, DVMRP e MOSFP**.



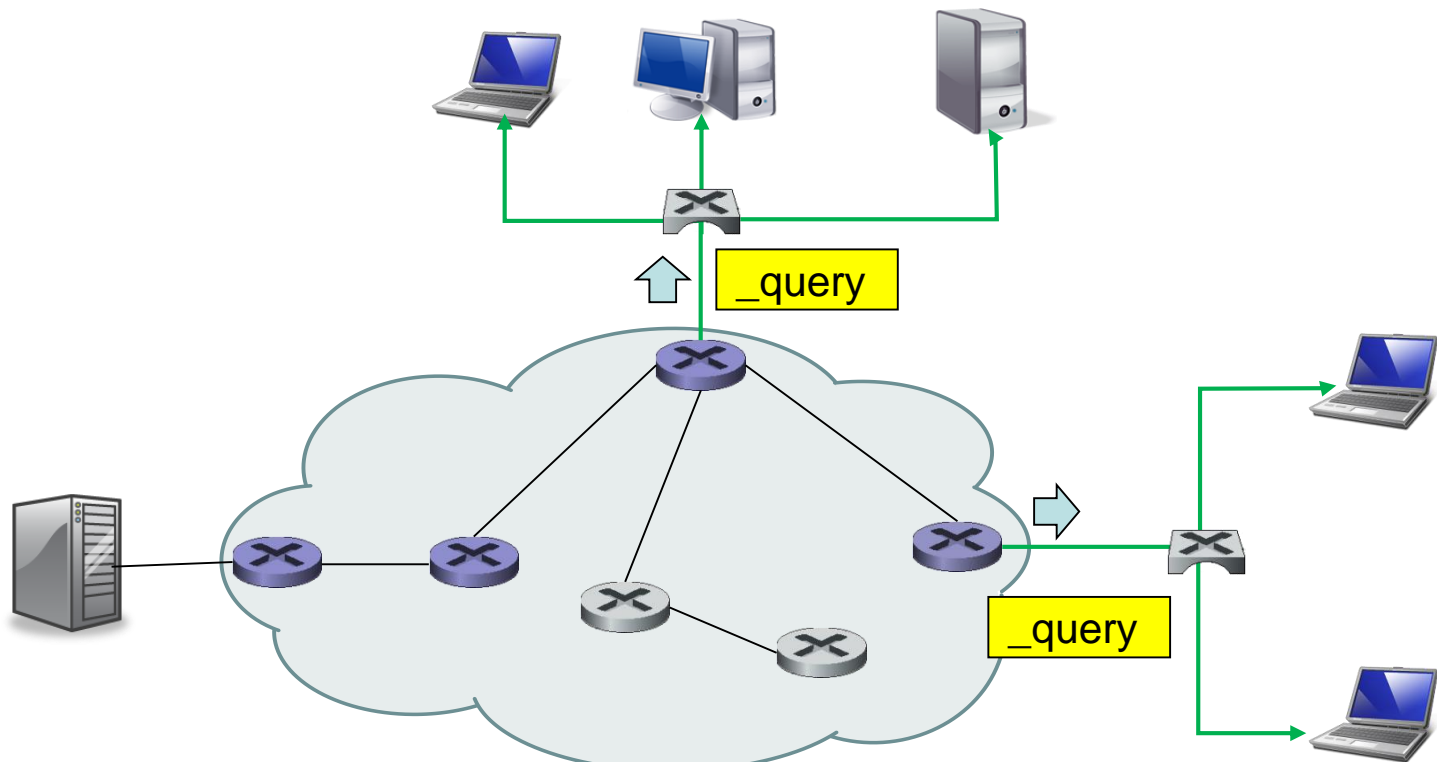
Protocolli di instradamento multicast

- I messaggi IGMP sono incapsulati all'interno di un datagram IP, con il valore del campo **numero di protocollo** pari a **2**.
- L'IGMP ha tre tipi di messaggio: **membership_query**, **membership_report** e **leave_group**.

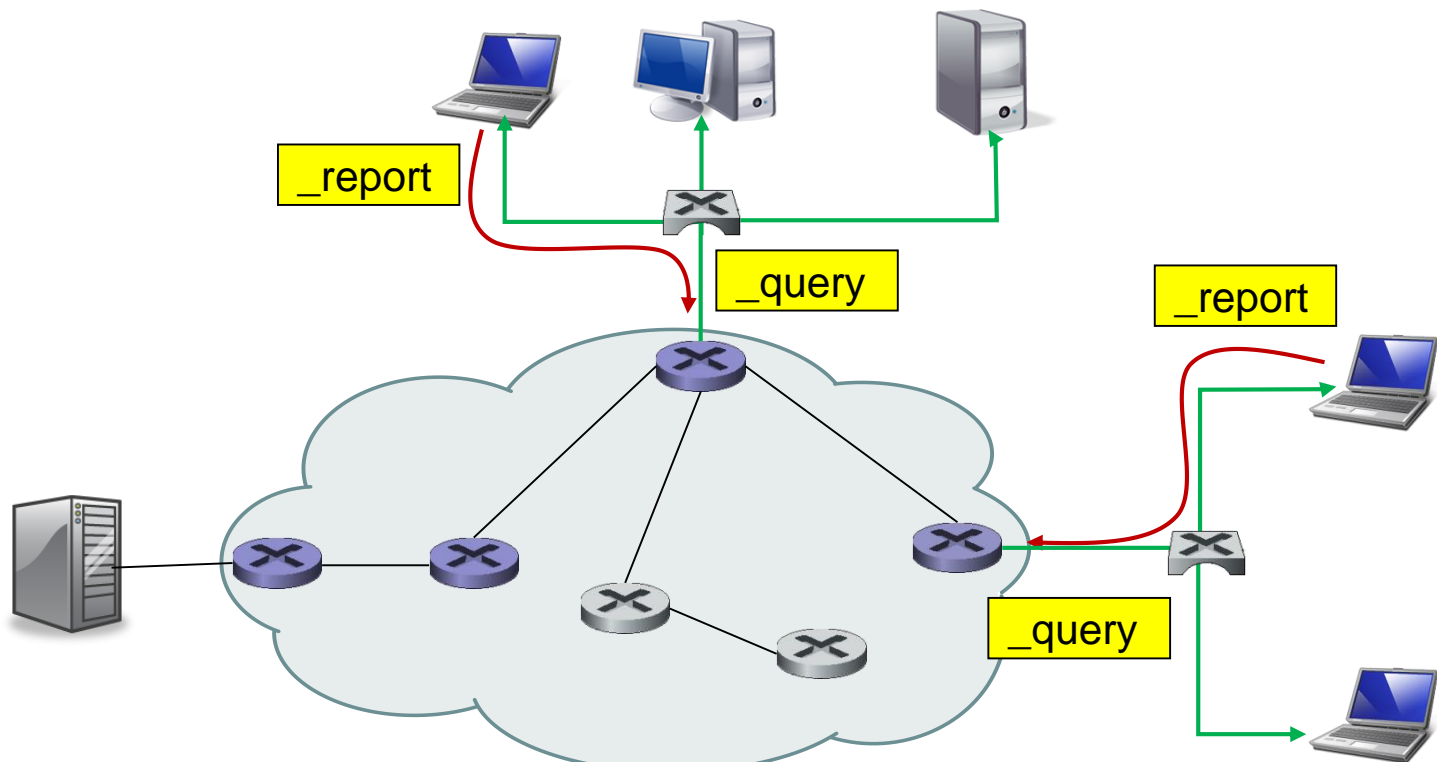


Formato del messaggio IGMP

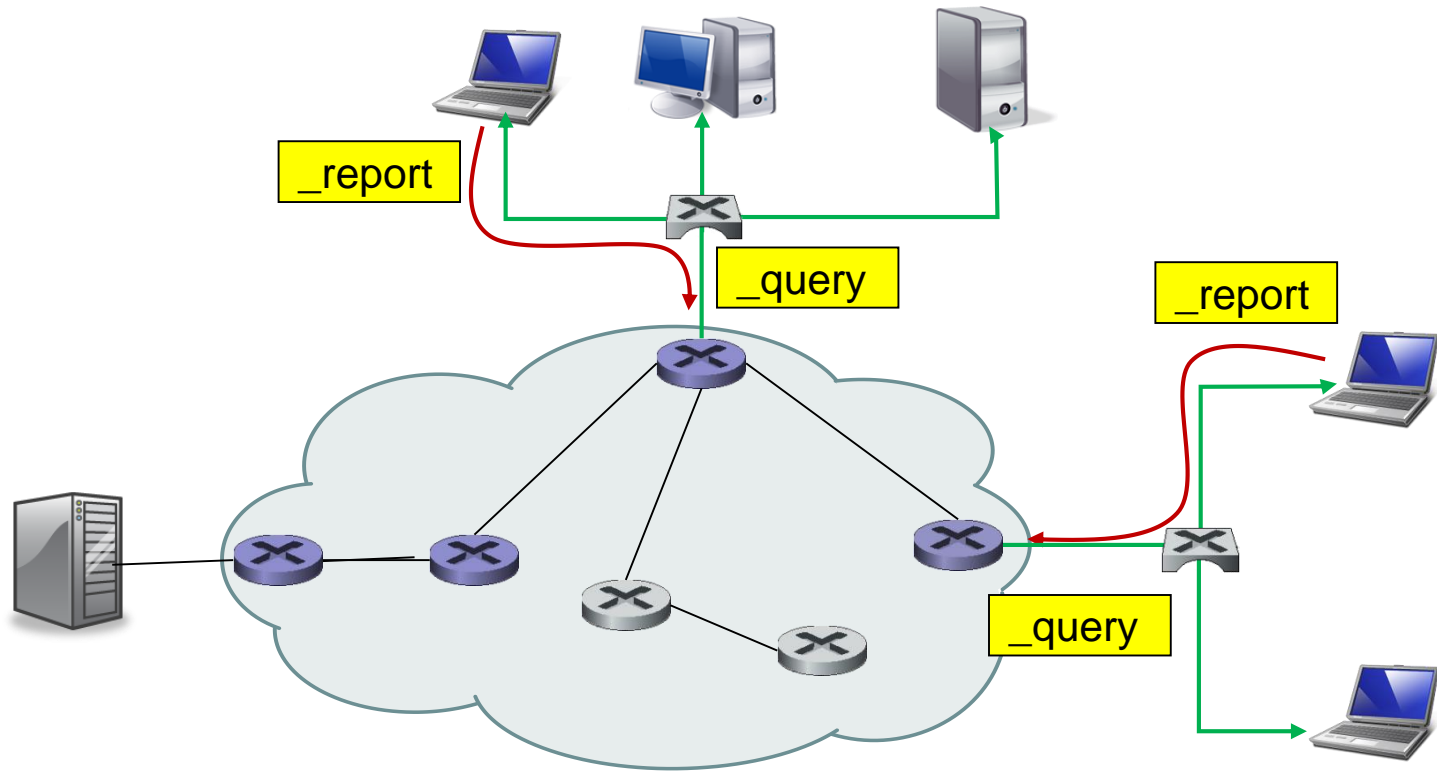
- Il messaggio **membership_query** è inviato da un router attraverso una sua interfaccia, collegata per esempio ad una LAN, per determinare se uno o più host su quella LAN si sono uniti a **gruppi multicast**.
- Questo messaggio può anche specificare un particolare gruppo multicast. In tal caso, l'indirizzo del gruppo multicast è inserito nel campo **indirizzo del gruppo multicast**



- Gli host rispondono a un messaggio **membership_query** con un messaggio **membership_report** (rapporto di appartenenza), come illustrato nella figura.
- I messaggi membership_report possono anche essere inviati da un host quando **un'applicazione si connette per la prima volta a un gruppo multicast**, senza attendere un messaggio membership_query dal router.



- Il messaggio **membership_report** contiene **l'indirizzo multicast del gruppo** cui l'host che risponde o che invia la prima volta è membro.
- I messaggi membership_report sono ricevuti sia dal router, che da tutti gli host collegati alla stessa LAN.
- A un router non interessa conoscere *quali e quanti* host fanno parte di uno specifico gruppo multicast, al router basta la conoscenza di **uno solo degli host per ogni gruppo**.
- L'IGMP è dotato quindi di un meccanismo che ha lo scopo di ridurre il numero dei messaggi membership_report generati quando più host appartengono allo stesso gruppo multicast.
- In particolare, il formato del messaggio IGMP possiede il campo "**tempo massimo di risposta**" (*maximum response time*).



Messaggi membership_query e membership_report in IGMP

- Quando un host riceve un messaggio **membership_query**, prima di inviare un messaggio **membership_report** per un dato gruppo multicast, esso attende un intervallo di tempo casuale compreso tra zero e il valore del "*tempo massimo di risposta*". Se l'host riceve un messaggio **membership_report**, per il suo stesso gruppo, che è stato inviato da qualche *altro* host, esso non invia il suo messaggio **membership_report**, dato che ormai il router è stato informato della presenza di membri di quel gruppo multicast nella LAN.
- In questo modo si evita la trasmissione di messaggi **membership_report** non necessari.
- L'ultimo tipo di messaggio IGMP è **leave group**. Questo messaggio è opzionale. Infatti, il router deduce che nessun host è associato a uno specifico gruppo multicast quando nessun host risponde a un messaggio **membership_query** con l'indirizzo di quel dato gruppo.

- Da quanto descritto fino ad ora con IGMP, qualsiasi host inviando un messaggio IGMP **membership_report** può unirsi a un gruppo multicast a livello di rete.
- L'appartenenza a un gruppo multicast è quindi decisa dal destinatario (un client).
- Un mittente (tipicamente un server) non verifica l'identità dei client che si uniscono ad un gruppo multicast, e non può nemmeno controllare chi si unisce al gruppo e quindi chi riceve i datagram spediti a quel gruppo.

Tipo messaggio	Inviato da	scopo
Membership query: generale	Router	Richiede i gruppi multicast
Membership query: specifico	Router	Richiede uno specifico gruppo multicast
Membership report	Host	Registra che un host si vuole unire o che appartiene ad un dato gruppo multicast
Leave group	Host	Registra l'abbandono da un gruppo multicast

Messaggi IGMP

Algoritmi di instradamento multicast

- Il **PIM** (Protocol Independent Multicast) è implementato in numerose piattaforme di router ed è stato sviluppato nell'ambito di trasferimento di stream multimediali.
- **MOSPF** (Multicast Open Shortest Path First) funziona nei sistemi autonomi (AS) che usano il protocollo OSPF per l'instradamento unicast. Il MOSPF estende l'OSPF.
- Il **DVMRP** è il protocollo "de facto" di instradamento multicast inter-AS. Comunque non funziona molto bene.
- **Lo sviluppo di un protocollo multicast inter-AS è un'attiva area di ricerca e sviluppo.**

Esempio di applicazione multicast in java

- L'applicazione seguente mostra l'uso della comunicazione multicast in java.
- Il server spedisce in multicast, usando l'indirizzo 230.10.10.1 e la porta 2014, la data e l'ora.

```
import java.io.*;
import java.net.*;
import java.util.*;
import java.lang.*;
public class MulticastServer{
    public static void main (String args[]) throws Exception {
        MulticastSocket socket = new MulticastSocket();
        byte[] buf = new byte[64];
        InetAddress gruppoIP = InetAddress.getByName("230.10.10.1");
        String sData;
        DatagramPacket packet=null;
        while (true) {
            sData = new Date().toString();
            System.out.println(sData);
            buf = sData.getBytes();
            packet = new DatagramPacket(buf, buf.length, gruppoIP, 2014);
            socket.send(packet);
            Thread.sleep(5000);
        }
        //socket.close();
    }
}
```

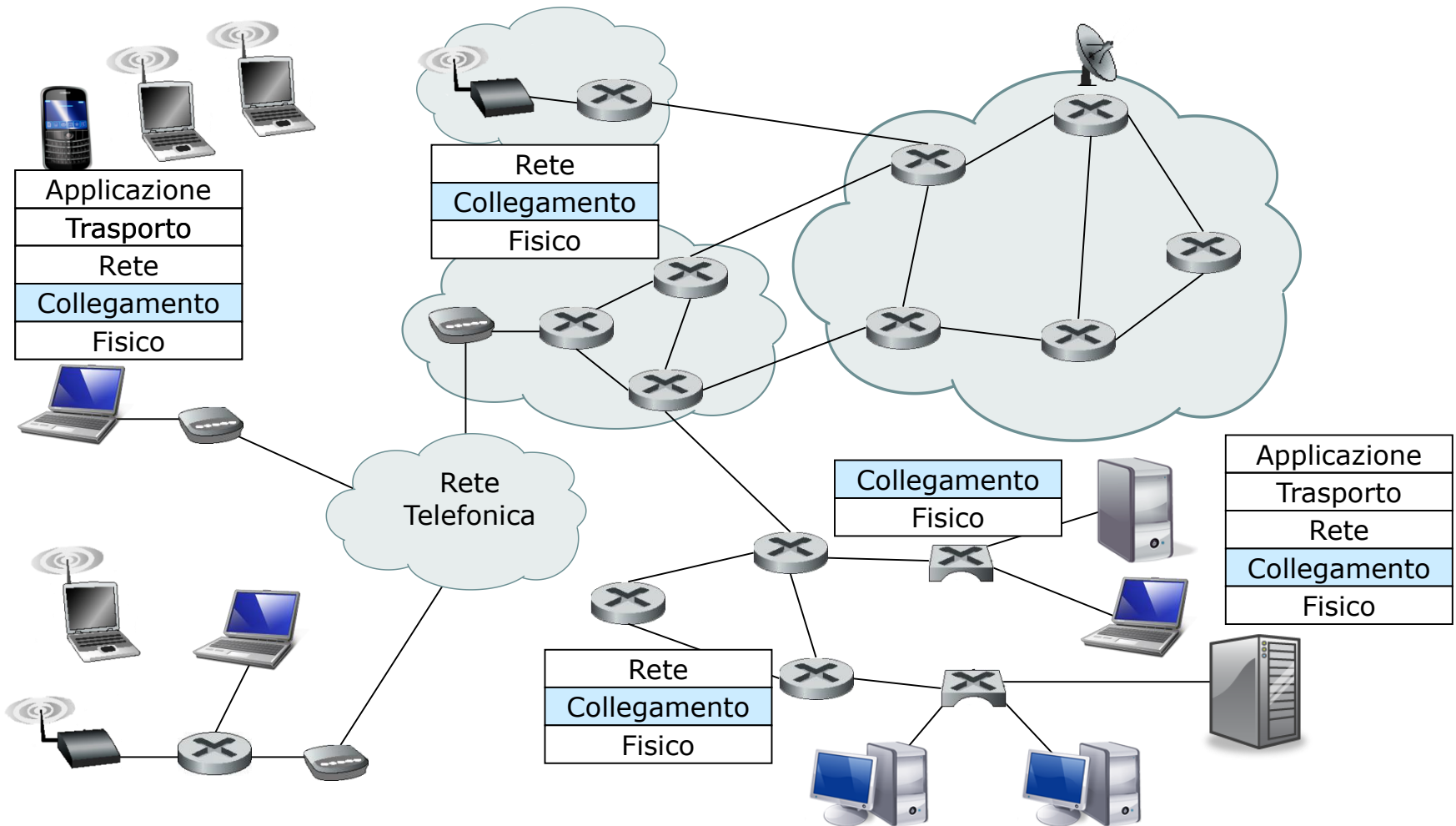
```
import java.io.*;
import java.net.*;
import java.util.*;

public class MulticastClient {
    public static void main(String[] args) throws IOException {
        MulticastSocket socket = new MulticastSocket(2014);
        InetAddress gruppoIP = InetAddress.getByName("230.10.10.1");
        socket.joinGroup(gruppoIP);
        DatagramPacket packet;
        byte[] buf = new byte[64];
        String risposta;
        while (true) {
            packet = new DatagramPacket(buf, buf.length);
            socket.receive(packet);
            risposta = new String(packet.getData());
            System.out.println("Sono le ore: " + risposta);
        }
        // socket.leaveGroup(gruppoIP);
        // socket.close();
    }
}
```

- Gli indirizzi multicast assegnati sono registrati nel sito <http://www.iana.org/assignments/multicast-addresses>.

Lo strato di collegamento

- In questo strato indicheremo con il termine **nodi** sia gli host che i router.
- Le unità di dati (PDU-2) inviate da un protocollo dello strato di collegamento sono chiamate **frame (trame)**.
- Mentre lo strato di rete si occupa di trasferire segmenti dello strato di trasporto **da host a host**, un protocollo dello strato di collegamento esegue il trasferimento di datagram dello strato di rete **da nodo a nodo** per ogni **link** del percorso.
- Un nodo mittente incapsula il datagram in un **frame**, e lo trasmette su un link; un nodo destinatario riceve il frame ed estrae il datagram.
- Alcuni protocolli dello strato di collegamento molto diffusi sono: **Ethernet, LAN wireless 802.11, token ring e PPP.**



- Una caratteristica importante dello strato di collegamento è che un datagram può essere incapsulato **in diversi protocolli dello strato di collegamento** nei differenti link del percorso.
- Per esempio, un datagram può essere trasportato da PPP sul primo link, da wireless 802.11 sull'ultimo link, e da Ethernet su tutti i link intermedi.
- È da notare che i vari protocolli dello strato di collegamento possono offrire servizi diversi allo strato soprastante. Per esempio, un protocollo dello strato di collegamento può essere affidabile nel trasferimento dati mentre un altro può non esserlo.

I servizi forniti dallo strato di collegamento

- I principali servizi offerti da un protocollo dello strato di collegamento sono: il **trasferimento affidabile dei dati**, il **controllo di flusso**, la **verifica di errori**, e l'**accesso casuale al canale condiviso**.

Trasferimento affidabile

- In modo analogo al TCP, l'affidabilità è ottenuta per mezzo di riscontri e ritrasmissioni.
- Un servizio per la trasmissione affidabile dei dati è generalmente usato per collegamenti che sono soggetti ad **alti tassi di errore**, come può essere un **link wireless**, con l'obiettivo di correggere localmente, a livello di protocollo di collegamento, un errore piuttosto che propagarlo ai protocolli degli strati superiori e affidare quindi a un protocollo dello strato di trasporto o di applicazione il problema della rilevazione ed eventuale correzione.

- Generalmente, il servizio di trasmissione affidabile dello strato di collegamento non si utilizza nelle tecnologie per reti cablate (ad esempio Ethernet) nelle quali i collegamenti sono realizzati con mezzi trasmissivi guidati come ad esempio fibre ottiche e cavi UTP formati da doppiini in rame. Il servizio non è ritenuto necessario in quanto le reti cablate hanno bassi valori del tasso di errore.

Accesso a link condivisi

- Un **protocollo di controllo di accesso al mezzo (MAC, *Media Access Control Protocol*)** specifica le regole di trasmissione di un frame sul link. Per i link punto-punto che hanno un singolo trasmettitore da un lato e un singolo ricevitore dall'altro lato del link, il protocollo di accesso al link è semplice poichè il trasmettitore può inviare un frame ogni volta che il link non è occupato.
- Il caso più complesso si verifica quando più nodi condividono un singolo link broadcast: il cosiddetto **problema dell'accesso multiplo**.

In questo caso il protocollo MAC serve a coordinare la trasmissione di frame da diversi nodi.

Framing

- Un frame è costituito da un **campo dati**, in cui è inserito il datagram dello strato di rete, e da **campi di intestazione**. Il protocollo di collegamento specifica la struttura del frame. Le intestazioni del frame contengono anche campi per l'**indirizzo fisico** del nodo, che è *diverso* dall'indirizzo del nodo dello strato di rete (per esempio, IP).

Controllo di flusso

- I nodi in ciascun lato di un link hanno buffer di limitata capacità per contenere i frame. Questo può essere un problema, poiché un nodo ricevente, in certi momenti, potrebbe ricevere frame a una velocità che supera la sua capacità di elaborazione. Senza controllo di flusso, il buffer di ricezione potrebbe saturarsi e i frame essere scartati come visto per lo strato di trasporto.

Rilevazione di errori

- Un nodo ricevente può ricevere frame con errori nei bit. Questi errori sono generati dall'attenuazione del segnale e da disturbi elettromagnetici. Per non consegnare un datagramma che contiene errori, molti protocolli dello strato di collegamento implementano funzioni per rilevare la presenza di errori. Ciò si realizza con l'inserimento, nel nodo mittente, di un gruppo di bit per la rilevazione di errori, e con il nodo ricevente che compie un controllo degli errori. Ricordiamo che lo strato di trasporto (protocolli UDP e TCP) sia quello di rete (protocollo IPv4 ma non IPv6) forniscono, pur se in forma limitata, una rilevazione dell'errore (a livello software). **La rilevazione nello strato di collegamento è più sofisticata ed è implementata in hardware.**

Correzione degli errori

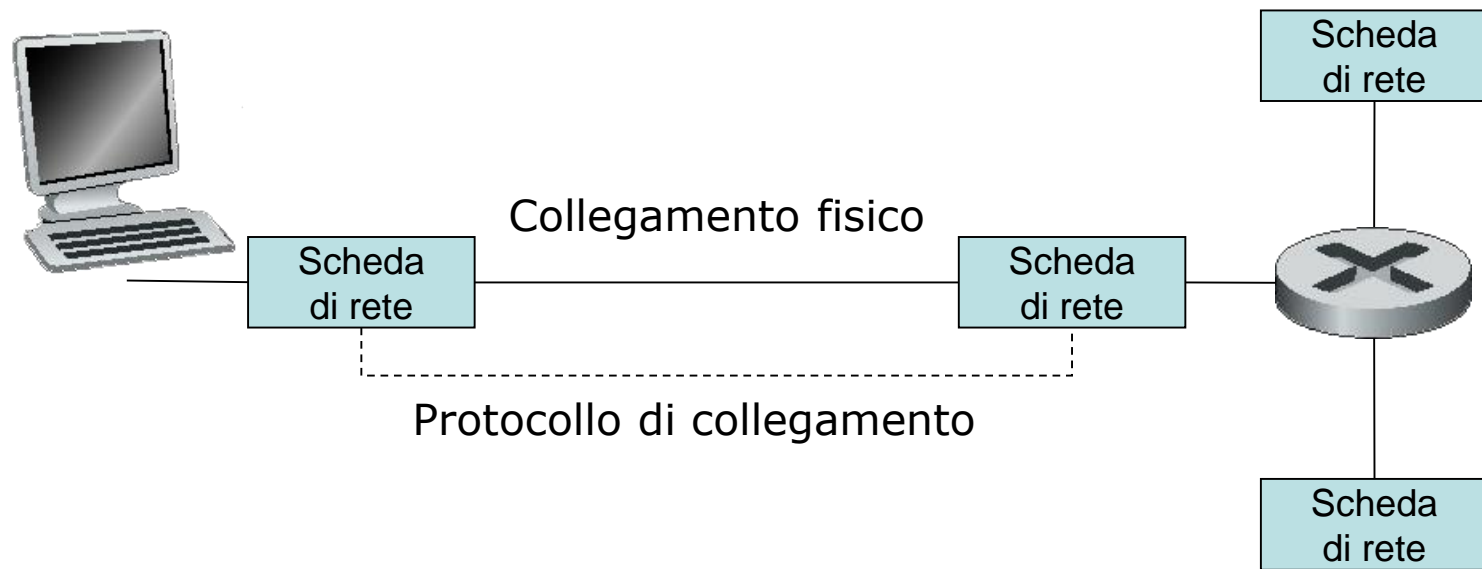
- un nodo ricevente può determinare la presenza di errori nel frame, ma anche correggerli.

Half-duplex e full-duplex

- Con la trasmissione full-duplex i nodi di entrambe le estremità di un link possono trasmettere frame contemporaneamente. Con la trasmissione half-duplex un nodo non può trasmettere e ricevere nello stesso tempo.

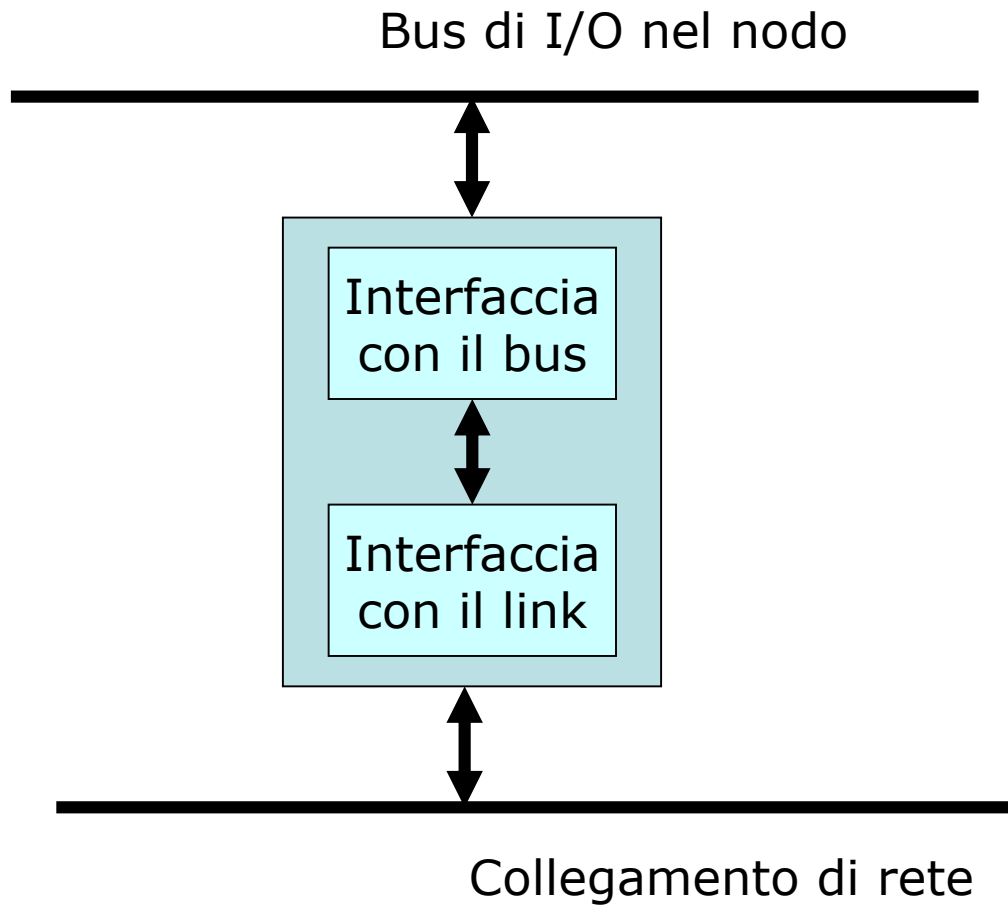
Adattatori (schede di rete)

- Il protocollo dello strato di collegamento è, per la gran parte, implementato in un **adattatore** (***scheda di rete***).
- Un adattatore è un dispositivo elettronico composto da chip DSP (Digital Signal Processor), RAM, un'interfaccia per il bus dell'host (esempio PCI o USB) e un'interfaccia per il link.
- Come mostra la figura seguente, lo strato di rete nel nodo mittente passa un datagram all'adattatore il quale incapsula il datagram in un frame e lo trasmette nel link. Nel destinatario, l'adattatore riceve l'intero frame, estrae il datagram e lo passa allo strato di rete.
- Se il protocollo dello strato di collegamento implementa il controllo dell'errore, allora è l'adattatore mittente che esegue le operazioni per la rilevazione dell'errore ed è l'adattatore ricevente che esegue il controllo dell'errore.
- Se il protocollo dello strato di collegamento fornisce il trasferimento affidabile, allora il meccanismo (per esempio, numeri di sequenza, timer e riscontri) è implementato completamente dagli adattatori.



Il protocollo di collegamento è implementato nelle schede di rete

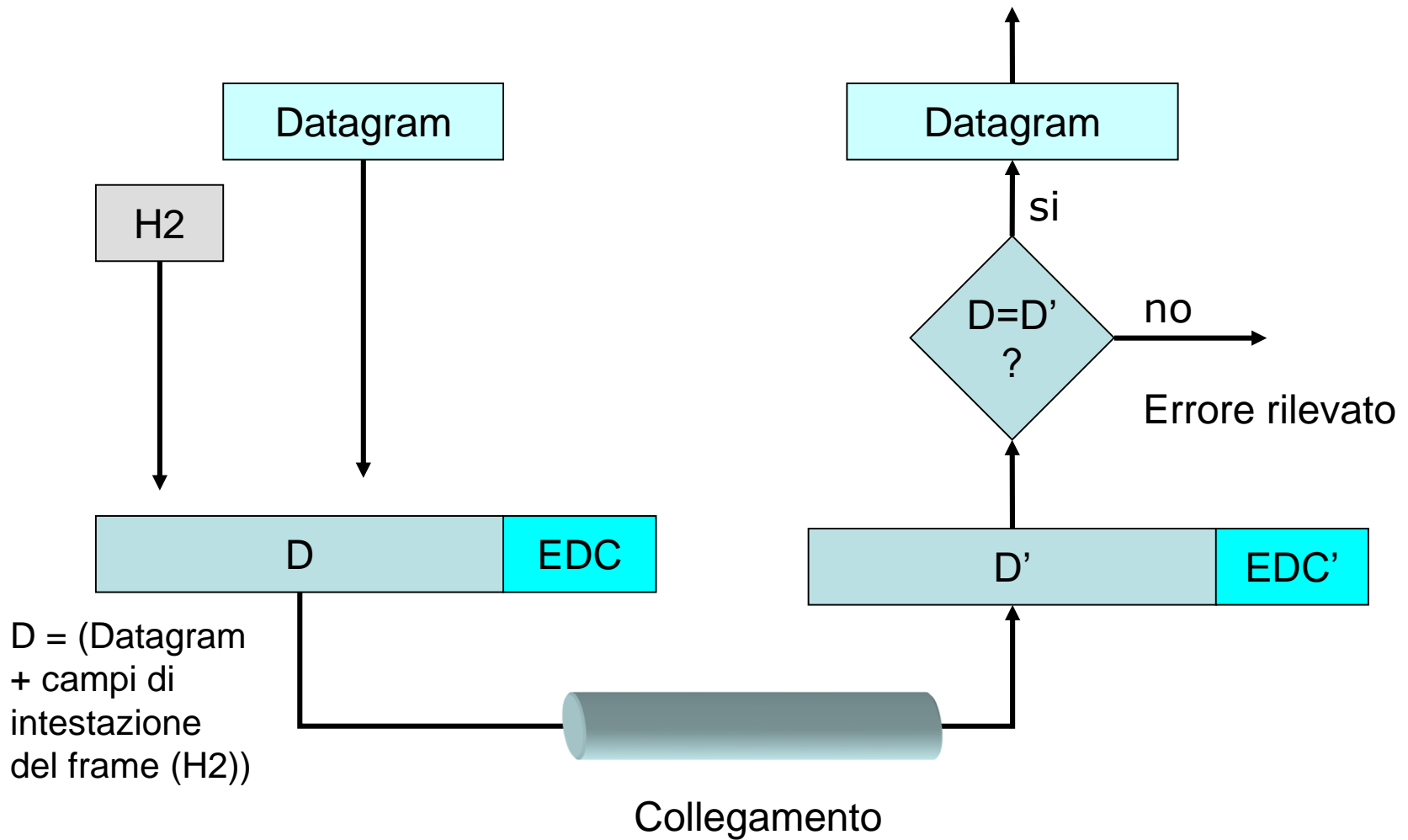
- Se il protocollo dello strato di collegamento fornisce l'accesso casuale il protocollo di accesso casuale è allora implementato completamente negli adattatori.
- Un adattatore è una **unità semi-autonoma**. Per esempio, un adattatore può ricevere un frame, determinare se ha un errore e scartarlo senza passarlo allo strato soprastante.
- In modo analogo, quando un nodo passa un datagram a un adattatore, l'adattatore svolge il compito di trasmettere il datagram attraverso quel link.
- Tuttavia, un adattatore non è un'unità completamente autonoma. A livello hardware, esso è cablato sulla stessa main board del nodo cui appartiene (o è connesso con esso tramite bus) ed è sotto il controllo del SO del nodo con il quale comunica, attraverso il suo driver.



- I componenti principali di un adattatore sono le interfacce con il bus e con il link. L'interfaccia con il bus è responsabile della comunicazione con il nodo collegato all'adattatore.
- L'interfaccia con il link è responsabile dell'implementazione del protocollo dello strato di collegamento. Oltre a framing e de-framing dei datagram, essa può fornire la rilevazione dell'errore, l'accesso casuale e altre funzioni dello strato di collegamento.
- La scheda **Ethernet** è la più diffusa tecnologia dello strato di collegamento.

Tecniche di rilevazione e correzione degli errori

- La figura seguente rappresenta uno schema basilare per lo studio delle tecniche di rilevazione e correzione degli errori.
- Nel nodo mittente, ai dati ***D*** da inviare, che devono essere protetti da possibili errori, sono aggiunti un certo numero di bit per la rilevazione e correzione dell'errore, ***EDC (Error Detection and Correction)***.
- Generalmente, i dati da proteggere sono i bit dell'intero frame, costituito dal datagram dello strato di rete e dall'intestazione del protocollo dello strato di collegamento.



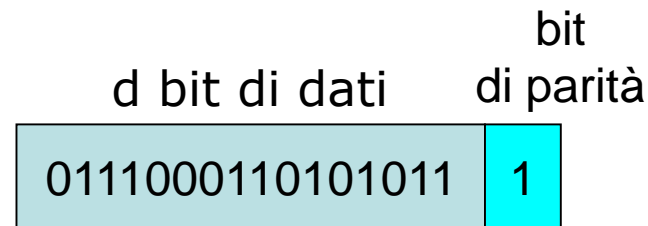
Rilevazione e correzione degli errori

- **D** e **EDC** formano un frame che il nodo mittente invia al nodo ricevente.
- Il nodo ricevente, riceve una sequenza di bit **D'** ed **EDC'** che potrebbero essere diversi dagli originali D e EDC , a causa di qualche **errore di trasmissione**.
- Il compito del ricevente è di determinare se D' è uguale all'originale D , considerando il fatto che esso ha ricevuto D' e EDC' .
- Le tecniche di **rilevazione e di correzione degli errori** permettono al nodo ricevente di individuare se si sono verificati errori nei bit.
- Tuttavia, anche con l'uso dei bit EDC c'è la probabilità che ci siano errori non rilevati. Pertanto, il protocollo di collegamento nel lato ricevente potrebbe consegnare un datagram errato allo strato di rete.

- In genere, per limitare la probabilità che non siano individuati errori nei bit, le tecniche di rilevazione e correzione degli errori più potenti e sofisticate utilizzano molti bit EDC.
- Esaminiamo ora tre tecniche per rilevare gli errori nei dati trasmessi:
 - **controllo di parità** (*parity check*) per spiegare i concetti di base della ricerca e della correzione dell'errore;
 - **metodi di checksum**, (metodi della somma di controllo) di solito usati nello strato di trasporto;
 - **controlli a ridondanza ciclica** (**CRC**, *cyclic redundancy check*) generalmente impiegati nello strato di collegamento negli adattatori di rete.

Controllo di parità

- Una semplice tecnica di rilevamento degli errori usa un solo bit, detto **bit di parità**.
- Supponiamo che l'informazione da spedire ***D*** contenga ***d*** bit. In uno schema di **parità pari**, il nodo trasmittente aggiunge **un bit** di controllo e sceglie il suo valore in modo che il numero totale di numeri **1** nei ***d* + 1** bit (l'informazione originale più un bit di parità) **sia pari**.



- Analogamente, per lo schema di **parità dispari**, il valore del bit di parità è scelto in modo che ci siano un numero **dispari di 1**.
- Con un solo bit di parità l'operazione che esegue il ricevitore è semplice in quanto esso deve solo contare il numero di 1 nei bit $d + 1$ ricevuti. Nel caso di schema di parità pari, il ricevitore se trova un numero dispari di bit con valore 1 può stabilire che si è verificato almeno un errore in un bit. Più esattamente, può determinare che si è verificato **un numero dispari di errori**.
- Se si verifica un **numero pari di errori** il ricevitore non è in grado di rilevare l'errore.
- Se il canale fisico è molto affidabile e quindi la probabilità di errore è piccola e si può assumere che l'errore possa verificarsi in un bit indipendentemente dal successivo, la probabilità di errori multipli in un pacchetto sarà molto bassa. In questo caso, un solo bit di parità potrebbe essere sufficiente.

- Però, le statistiche sulle misure hanno dimostrato che spesso gli errori si presentano a "**raffica**", invece che verificarsi indipendentemente tra loro.
- Se gli errori si verificano a raffica, la probabilità di non rilevare gli errori in un frame protetto da un solo bit di parità è del **50%**, in quanto è in ugual misura la probabilità che il numero di bit alterati sia pari o dispari.
- Per tali situazioni, è necessario ricorrere ad una tecnica più potente e sofisticata per la rilevazione degli errori.
- Consideriamo ora una semplice estensione dello schema di parità a un bit per mostrare una semplice soluzione della tecnica di **correzione dell'errore**.
- La figura seguente mostra come, i ***d*** bit in ***D*** sono suddivisi in ***i* righe** e ***j* colonne**. Un valore di parità è calcolato per ciascuna riga e per ciascuna colonna. I risultanti ***i* + *j* + 1** bit di parità contengono i bit per la rilevazione dell'errore del frame del link dati.

d1,1	d1,2	d1,j	d1,j+1
d2,1	d2,2	d2,j	d2,j+1
.....				
.....				
di,1	di,2	di,j	di,j+1
di+1,1	di+1,2		di+1,j	di+1,j+1

D = 1 0 1 0 1 1 0 1 1 0 0 1 0 1 0 1

Senza errori

1	0	1	0	0
1	1	0	1	1
1	0	0	1	0
0	1	0	1	0
1	0	1	1	1

Errore correggibile (1 bit)

1	0	1	0	0
1	1	0	1	1
1	1	0	1	0
0	1	0	1	0
1	0	1	1	1

Parità pari a due dimensioni.

- Supponiamo ora che si verifichi un solo errore nei bit degli originali ***d*** bit dei dati. Con questo schema di parità a due dimensioni, le parità sia della colonna sia della riga che contengono il bit errato presenteranno l'errore.
- Il ricevitore può quindi non solo ***rilevare*** che si è verificato un errore in un singolo bit, ma può usare gli indici della colonna e della riga che presentano l'errore di parità per identificare il bit che è stato alterato e ***correggerlo***.
- La figura precedente mostra un esempio in cui il bit di valore **0** in posizione **(3,2)** ha cambiato valore passando da 0 a 1.
- L'errore può essere rilevato e corretto anche se si verifica nei bit di parità.
- La parità a due dimensioni può anche rilevare, ma non correggere, qualsiasi combinazione di due errori in un pacchetto.

- La capacità del ricevitore di **rilevare e correggere** gli errori è detta **FEC** (**FEC**, *Forward Error Correction*, correzione degli errori in avanti).
- Queste tecniche sono comunemente usate in dispositivi audio di registrazione e riproduzione, come nei CD audio.
- Le tecniche FEC sono fondamentali perché possono diminuire il numero di ritrasmissioni richieste al nodo trasmittente. Infatti con la correzione automatica dell'errore il ricevitore non deve inviare indietro al mittente un pacchetto **NAK** (negative acknowledgment) e quindi attendere che il mittente ritrasmetta il pacchetto che conteneva errori, ritrasmissione che invece è necessaria nel caso di sola rilevazione dell'errore.

Metodi di checksum

- Le tecniche basate sulla checksum suddividono i pacchetti di ***d*** bit da proteggere in parole di ***k*** bit.
- Come già descritto per UDP e TCP, la *checksum di Internet* si basa su questa tecnica: il nodo mittente calcola la checksum facendo il complemento a 1 della somma delle parole di 16 bit del segmento e di alcuni campi dell'intestazione IP.
- Il nodo ricevente controlla la checksum effettuando il complemento a 1 della somma dei dati ricevuti (compresa la checksum) e verifica che il risultato sia formato da tutti bit pari a 1. Se qualche bit è pari a 0, viene indicato un errore.
- I metodi checksum forniscono una protezione debole contro gli errori in confronto alle tecniche CRC, che generalmente sono usate nello strato di collegamento.

- Il motivo per cui il metodo checksum è implementato nello strato di trasporto è perché questo è implementato a livello software e fa parte del sistema operativo e quindi è necessario utilizzare una tecnica di rilevazione di errore semplice e veloce.
- La rilevazione dell'errore nello strato di collegamento è invece **implementato in hardware nelle schede di rete**, per cui si possono effettuare più rapidamente le più complesse operazioni di controllo di errori.
- Inoltre è meglio correggere l'errore immediatamente nell'adattatore piuttosto che propagarlo al sistema operativo.