# Reti di Elaboratori

Livello di Collegamento: Introduzione



Alessandro Checco@uniroma1.it



Capitolo 6

# Esercizi sullo strato di rete

Ognuno dei seguenti indirizzi appartiene a un blocco. Trovare il primo e l'ultimo indirizzo di ogni blocco

- 1. 14.12.72.8/24
- 2. 200.107.16.17/18

Spiegare la differenza tra routing e forwarding

I messaggi OSPF e quelli ICMP vengono incapsulati direttamente nei datagrammi IP. Se intercettiamo un datagramma IP, come possiamo capire se il payload è relativo all'OSPF o all'ICMP?

Dato il grafo  $\{(A,B),(B,C),(B,D),(C,D)\}$  con costi  $c_{AB}=3$ ,  $c_{BC}=2$ ,  $c_{BD}=c_{CD}=1$  disegnare:

- 1. il grafo
- 2. Il distance vector al primo step per ogni nodo
- 3. il link state database alla fine del flooding

Qual è la differenza tra i pacchetti Distance Vector e i pacchetti Link State?

- Che tipo di informazione viene propagata?
- Come si propaga l'informazione sulla rete (come cambia all'attraversamento dei nodi)?

Supponiamo che la distanza minima tra I nodi a,b,c,d rispetto al nodo y ed i costi dal nodo x ai nodi a,b,c,d siano:

$$D_{ay}=5$$
,  $D_{by}=6$ ,  $D_{cy}=4$ ,  $D_{dy}=3$   
 $C_{xa}=2$ ,  $C_{xb}=1$ ,  $C_{xc}=3$ ,  $C_{xd}=1$ 

Qual è la distanza minima  $D_{xy}$  tra il nodo x e il nodo y, usando l'equazione di Bellman-Ford?

Quali campi dell'intestazione IP cambiano all'attraversamento di un router?

# Link layer e LAN: obiettivi

- comprendere i principi alla base dei servizi a livello di collegamento:
  - rilevamento e correzione degli errori
  - condivisione di un canale di trasmissione: accesso multiplo
  - indirizzamento al livello di collegamento
  - reti locali: Ethernet, VLAN
- reti nei data center
- istanziazione e implementazione di varie tecnologie a livello di collegamento

## Livello di collegamento e LAN: sommario

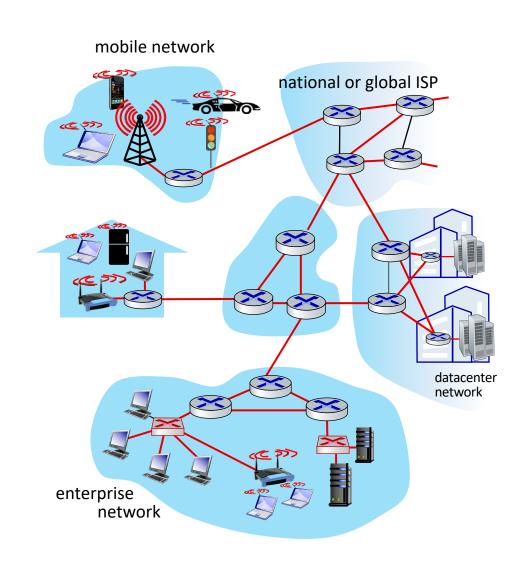
- introduzione
- rilevamento e correzione degli errori
- protocolli di accesso mutiplo
- LAN
  - indirizzamento, ARP
  - Ethernet
  - switch
  - VLAN
- virtualizzazione dei collegamenti: MPLS
- data center
- un giorno nella vita di una richiesta web

# Link layer: introduzione

### terminologia:

- host e router: nodi
- canali di comunicazione che collegano nodi adiacenti lungo il percorso di comunicazione: link
  - cablato
  - senza fili
  - LAN
- pacchetto di livello 2: frame, incapsula il datagramma

il livello di collegamento ha la responsabilità di trasferire datagrammi tra due nodi fisicamente adiacenti lungo un link



## Livello di collegamento: contesto

- datagramma trasferito da protocolli di collegamento diversi su collegamenti diversi:
  - ad esempio, WiFi sul primo collegamento, Ethernet sul collegamento successivo
- ogni protocollo di collegamento fornisce servizi diversi
  - ad esempio, può fornire o meno un trasferimento di dati affidabile lungo il link

#### analogia del trasporto:

- viaggio da Princeton a Losanna
  - limousine: da Princeton a JFK
  - aereo: JFK a Ginevra
  - treno: da Ginevra a Losanna
- turista = datagramma
- segmento di trasporto = link
- modalità di trasporto = protocollo a livello di link
- agente di viaggio = algoritmo di routing

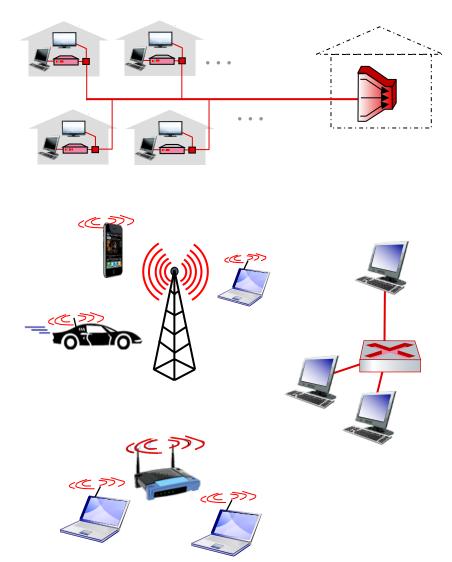
## Livello di collegamento: servizi

### framing, link access:

- incapsula il datagramma nel frame, aggiungendo header, trailer
- accesso al canale se mezzo condiviso
- Gli indirizzi "MAC" nelle intestazioni dei frame identificano origine, destinazione (diversi dall'indirizzo IP!)

### consegna affidabile tra nodi adiacenti

- sappiamo già come farlo!
- usato raramente su collegamenti a basso tasso di bit error
- collegamenti wireless: alti tassi di bit error
  - D: perché avere affidabilità sia a livello di collegamento che end-to-end?



# Livello di collegamento: servizi (2)

#### controllo del flusso (flow control):

 evitare sovraccarico dei buffer del nodo di destinazione

#### rilevamento errori:

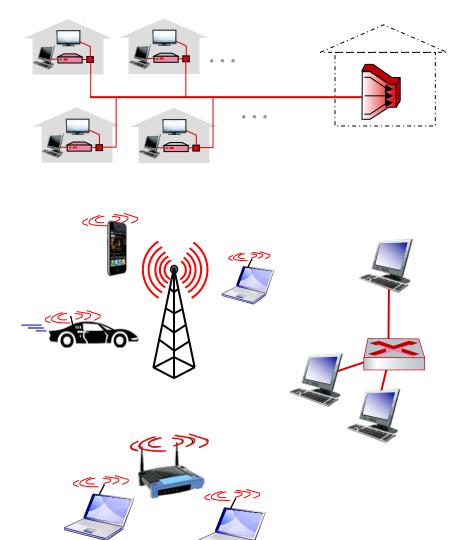
- errori causati da attenuazione del segnale, rumore
- il ricevitore rileva errori, chieve la ritrasmissione o scarta il frame

#### Correzione dell'errore:

 ricevente identifica e corregge errori di bit senza ritrasmissione

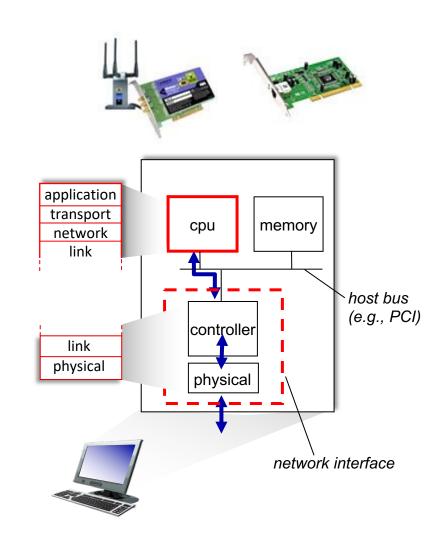
### half duplex e full duplex:

• con half duplex, i nodi ad entrambe le estremità del collegamento possono trasmettere, ma non contemporaneamente

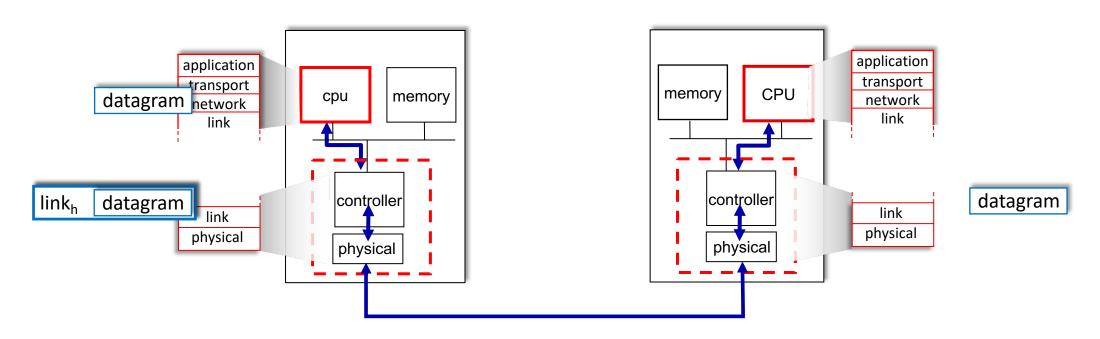


## Dove è implementato il livello di collegamento?

- in ogni singolo host
- livello di collegamento implementato nella scheda di interfaccia di rete (NIC) o su un chip
  - Ethernet, scheda WiFi o chip
  - implementa i livelli di collegamento e quello fisico
- si collega ai bus di sistema dell'host
- combinazione di hardware, software, firmware



### Comunicazione tra interfacce



#### lato mittente:

- incapsula il datagramma nel frame
- aggiunge bit di controllo degli errori, trasferimento dati affidabile, controllo del flusso, ecc.

#### lato ricevente:

- cerca errori, trasferimento dati affidabile, controllo del flusso, ecc.
- estrae il datagramma, passa al livello superiore sul lato ricevente

## Due sottolivelli del livello di collegamento

#### **Data Link Control (DLC)**

si occupa di tutte le questioni **comuni** sia ai collegamenti puntopunto che a quelli broadcast

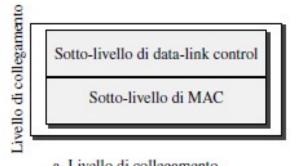
- Framing
- Controllo del flusso e degli errori
- Rilevamento e correzione degli errori

si occupa delle procedure per la comunicazione tra due nodi adiacenti (comunicazione nodo-a-nodo), indipendentemente dal fatto che il collegamento sia dedicato o broadcast

#### **Media Access Control (MAC)**

si occupa solo degli aspetti specifici dei canali broadcast

 Controllo dell'accesso al mezzo condiviso



 a. Livello di collegamento di un collegamento broadcast



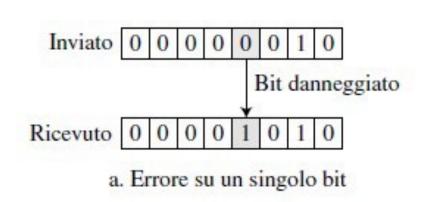
 b. Livello di collegamento di un collegamento punto-punto

## Livello di collegamento e LAN: sommario

- introduzione
- rilevamento e correzione degli errori
- protocolli di accesso mutiplo
- LAN
  - indirizzamento, ARP
  - Ethernet
  - switch
  - VLAN
- virtualizzazione dei collegamenti: MPLS
- data center
- un giorno nella vita di una richiesta web

### Errori su singolo bit o a burst

Gli errori sono dovuti a interferenze che possono cambiare la forma del segnale

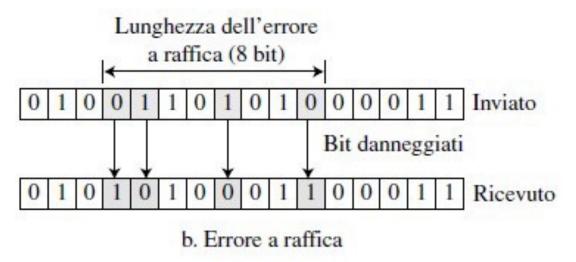


di tipo burst (a raffica) è più elevata rispetto a quella di un errore sul singolo bit, in quanto la durata dell'interferenza (detta anche rumore) normalmente è più lunga rispetto a quella di un solo bit

La probabilità che avvenga un errore

Il numero di bit coinvolti dipende dalla velocità di trasferimento dati e dalla durata del rumore.

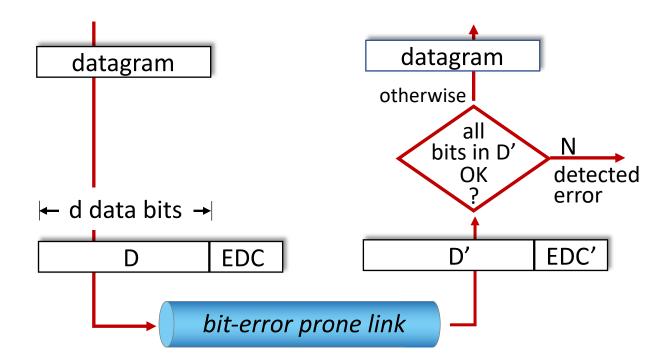
ES. 1 Kbps con un rumore di 1/100 sec può influire su 10 bit



## Rilevamento degli errori

EDC: error detection and correction bits (ad es. ridondanza)

D: dati protetti dal controllo degli errori, possono includere campi di intestazione



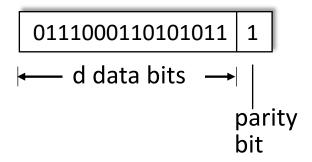
Rilevamento errori non affidabile al 100%!

- il protocollo può non rilevare alcuni errori, ma raramente
- un campo EDC più ampio produce una maggiore capacità rilevazione e correzione

# Parity checking

#### parità a bit singolo:

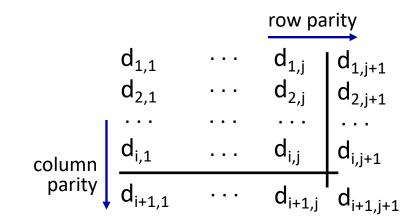
rilevare errori singoli

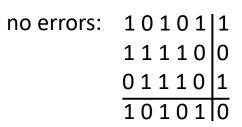


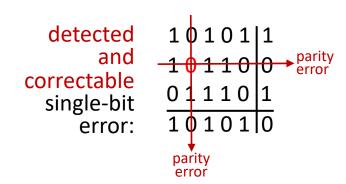
Even parity: il bit di parità viene impostato in modo che la somma di tutti i bit (incluso il bit di parità) sia pari

#### parità di bit bidimensionale:

■ rilevare e **correggere** errori singoli







## Internet checksum (ripasso)

Obiettivo: rilevare gli errori (cioè i bit capovolti) nel segmento trasmesso

#### mittente:

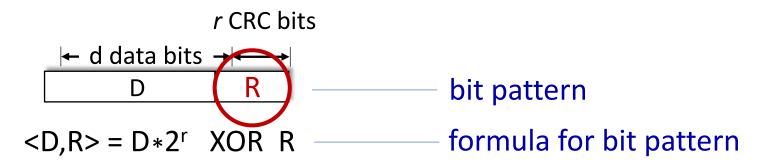
- tratta i contenuti del segmento UDP (compresi i campi di intestazione UDP e gli indirizzi IP) come una sequenza di numeri interi a 16 bit
- checksum: addizione (somma in complemento a uno) del contenuto del segmento
- valore di checksum inserito nel campo checksum UDP

#### ricevitore:

- calcola il checksum del segmento ricevuto
- controlla se il checksum calcolato è uguale al valore del campo checksum:
  - non uguale errore rilevato
  - uguale nessun errore rilevato. Ma potenzialmente possono comunque esserci errori

## Cyclic Redundancy Check (CRC)

- codifica di rilevamento degli errori più potente
- D: bit di dati (dati, trattati come un numero binario)
- G: pattern (generator) di *r+1* bits (predeterminato)



*obiettivo*: scegliere *r* bit CRC, R tali che <D,R> esattamente divisibile per G (mod 2)

- il ricevitore conosce G, divide <D,R> per G. Se resto diverso da zero: errore rilevato!
- può rilevare tutti gli errori burst inferiori a r+1 bit
- ampiamente utilizzato in pratica (Ethernet, WiFi 802.11)

# Cyclic Redundancy Check (CRC): esempio

#### Vogliamo trovare R per cui:

 $D \cdot 2^r XOR R = nG$ 

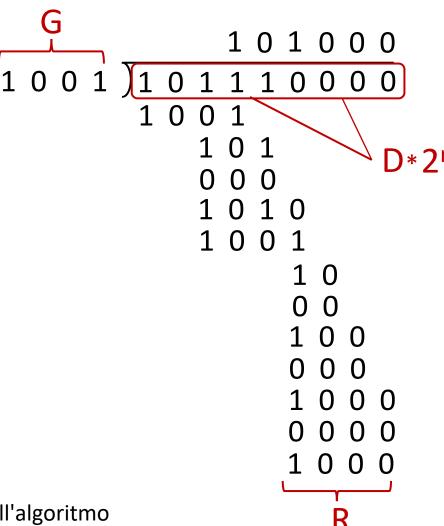
equivalente a (XOR R both sides):

 $D \cdot 2^r = nG XOR R$ 

#### equivalente a:

se dividiamo D 2<sup>r</sup> per G (mod 2), il resto R deve soddisfare:

$$R = remainder \left[ \frac{D \cdot 2^r}{G} \right]$$



tutte le operazioni sono modulo 2, quindi la sottrazione nell'algoritmo di divisione è una XOR

## Livello di collegamento e LAN: sommario

- introduzione
- rilevamento e correzione degli errori
- protocolli di accesso mutiplo
- LAN
  - indirizzamento, ARP
  - Ethernet
  - switch
  - VLAN
- virtualizzazione dei collegamenti: MPLS
- data center
- un giorno nella vita di una richiesta web

# Collegamenti e protocolli di accesso multiplo

### due tipi di "link":

- punto-punto
  - collegamento punto-punto tra switch Ethernet, host
  - PPP per l'accesso dial-up
- trasmissione broadcast (cavo o mezzo condiviso)
  - Ethernet vecchio stile
  - HFC a monte nella rete di accesso via cavo
  - LAN senza fili 802.11, 4G/4G. satellitare



cavo condiviso (ad es. Ethernet cablata)



radio condivisa: 4G/5G



radio condivisa: Wi -Fi



radio condivisa: satellitare



umani a un party (aria condivisa, acustico)

## Protocolli di accesso multiplo

- singolo canale di trasmissione condiviso
- due o più trasmissioni simultanee da parte dei nodi: interferenza
  - collisione se il nodo riceve due o più segnali contemporaneamente

### protocollo di accesso multiplo

- algoritmo distribuito che determina come i nodi condividono il canale, cioè determina quando il nodo può trasmettere
- la comunicazione sulla condivisione del canale deve utilizzare il canale stesso!
  - nessun canale fuori banda per il coordinamento

## Un protocollo di accesso multiplo ideale

dato: canale di accesso multiplo (MAC) di velocità R bps obiettivi:

- 1. quando un nodo vuole trasmettere, può inviare alla velocità R.
- 2. quando M nodi vogliono trasmettere, ognuno può trasmettere a velocità media R/M
- 3. completamente decentralizzato:
  - nessun nodo speciale per coordinare le trasmissioni
  - nessuna sincronizzazione di orologi o slot temporali
- 4. semplice

### Protocolli MAC: tassonomia

### tre grandi classi:

- partizionamento dei canali
  - dividere il canale in "pezzi" più piccoli (fasce orarie, frequenza, codice)
  - allocare quel pezzo al nodo per uso esclusivo

#### accesso casuale

- canale non suddiviso, consentire collisioni
- "recuperare" dalle collisioni

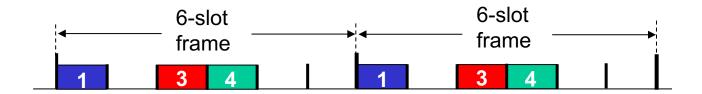
#### a rotazione

• i nodi si alternano, ma i nodi con più informazioni da inviare possono richiedere turni più lunghi

## Protocolli MAC di partizionamento del canale: TDMA

### TDMA: time division multiple access

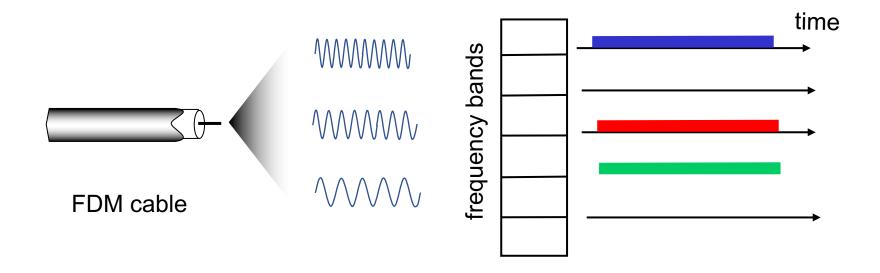
- accesso al canale in "round"
- ogni stazione ottiene slot di lunghezza fissa (lunghezza = tempo di trasmissione del pacchetto) in ogni round
- gli slot inutilizzati diventano inattivi
- esempio: LAN a 6 stazioni, 1,3,4 hanno pacchetti da inviare, slot
   2,5,6 inattivi



## Protocolli MAC di partizionamento del canale: FDMA

### FDMA: frequency division multiple access

- spettro del canale suddiviso in bande di frequenza
- a ciascuna stazione è assegnata una banda di frequenza fissa
- il tempo di trasmissione inutilizzato nelle bande di frequenza diventa inattivo
- esempio: LAN a 6 stazioni, 1,3,4 hanno pacchetti da inviare, bande di frequenza 2,5,6 inattive



### Protocolli di accesso casuale

- quando il nodo ha un pacchetto da inviare
  - trasmette alla massima velocità dati del canale R
  - nessun coordinamento a priori tra i nodi
- due o più nodi trasmittenti: "collisione"
- Il protocollo MAC ad accesso casuale specifica:
  - come rilevare le collisioni
  - come recuperare dalle collisioni (ad esempio, tramite ritrasmissioni ritardate)
- esempi di protocolli MAC ad accesso casuale:
  - ALOHA, slotted ALOHA
  - CSMA, CSMA/CD, CSMA/CA

### Slotted ALOHA

#### assunzioni:

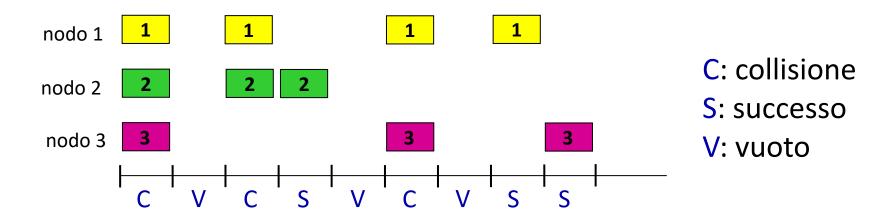
- tutti i frame della stessa dimensione
- tempo diviso in slot di uguali dimensioni (tempo per trasmettere 1 frame)
- i nodi iniziano a trasmettere solo all'inizio dello slot
- i nodi sono sincronizzati
- se 2 o più nodi trasmettono nello slot, tutti i nodi rilevano la collisione

#### funzionamento:

- quando il nodo ottiene un nuovo frame, lo trasmette nello slot successivo
  - se non c'è collisione: il nodo può inviare un nuovo frame nello slot successivo
  - in caso di collisione: il nodo ritrasmette il frame a ogni slot successivo con probabilità p fino al successo

randomizzazione : perché ?

### Slotted ALOHA



#### Pro:

- se un singolo nodo è attivo, può trasmettere continuamente alla massima velocità del canale
- altamente decentralizzato: solo gli slot nei nodi devono essere sincronizzati
- semplice

#### Contro:

- collisioni, spreco di slot
- slot inattivi (dopo collisione)
- i nodi potrebbero essere in grado di rilevare la collisione in un tempo minore di uno slot
- sincronizzazione dell'orologio

### Slotted ALOHA: efficienza

efficienza: frazione a lungo termine di slot tramessi con successo (assumendo molti nodi, tutti con molti frame da inviare)

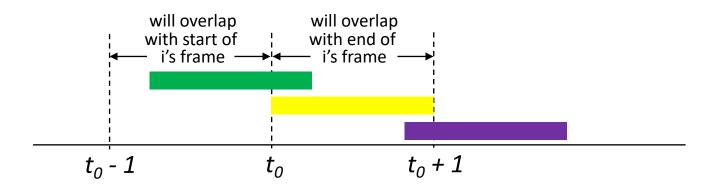
- assumiamo: N nodi con molti frame da inviare, ognuno trasmette in slot con probabilità p
  - prob quel dato nodo ha successo in uno slot =  $p(1-p)^{N-1}$
  - prob che qualsiasi nodo ha successo =  $Np(1-p)^{N-1}$
  - massima efficienza: trova  $p^*$  che massimizza  $Np(1-p)^{N-1}$
  - per molti nodi, il limite di  $Np^*(1-p^*)^{N-1}$  per  $N \to \infty$ :

efficienza massima = 1/e ~= 0.37

• al massimo: canale utilizzato per trasmissioni utili il 37% delle volte!

## **Pure ALOHA**

- Aloha senza slot: più semplice, nessuna sincronizzazione
  - quando arriva un frame viene trasmesso senza aspettare l'inizio di uno slot
- la probabilità di collisione aumenta senza sincronizzazione:
  - il frame inviato a t<sub>0</sub> andrà in collisione con frame inviati nell'intervallo temporale[t<sub>0</sub>-1,t<sub>0</sub>+1] (due volte il tempo di trasmissione di un frame, questo intervallo è anche noto come **tempo di vulnerabilità**)



Efficienza di pure Aloha: 18%!

# CSMA (carrier sense multiple access)

#### CSMA semplice - ascolta prima di trasmettere:

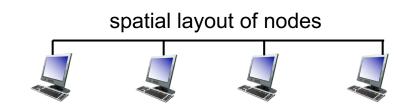
- se il canale viene rilevato inattivo: trasmette l'intero frame
- se il canale è occupato: differire la trasmissione
- analogia umana: non interrompere gli altri!

#### CSMA/CD: CSMA con rilevamento delle collisioni

- collisioni *rilevate* in un tempo breve
- le trasmissioni in collisione vengono immediatamente interrotte, riducendo lo spreco di canale
- rilevamento delle collisioni facile nel cablato, difficile con wireless
- analogia umana: il conversatore educato

## CSMA: collisioni

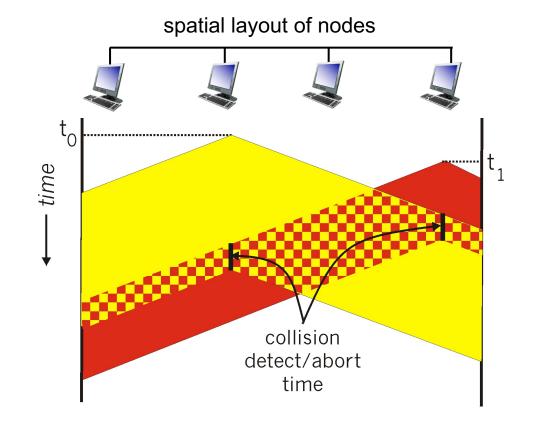
- collisioni possono ancora verificarsi con il carrier sense:
  - ritardo di propagazione significa che due nodi potrebbero ognuno non sentire la trasmissione appena iniziata dall'altro
- collisione: tempo di trasmissione dell'intero pacchetto sprecato
  - la distanza e il ritardo di propagazione svolgono un ruolo nel determinare la probabilità di collisione
- Tempo di vulnerabilità: T<sub>p</sub>





# CSMA/CD:

- CSMA/CD riduce la quantità di tempo sprecato nelle collisioni
  - trasmissione interrotta al rilevamento della collisione
- Rilevazione della collisione
  - facile nelle LAN cablate (a singolo canale)
  - difficile nelle LAN wireless (motivi fisici, differenza tra potenza di emissione e di ricezione) quindi cercheranno invece di evitare problemi successivi (collision avoidance)
- Non cambia il tempo di vulnerabilità



## Dimensione minima del frame

- □ Cosa succederebbe se un mittente finisse di trasmettere un frame prima di ricevere il primo bit di un'altra stazione (che ha già iniziato a trasmettere)?
- □ Una stazione una volta inviato un frame non tiene una copia del frame, né controlla il mezzo trasmissivo per rilevare collisioni
- Perché il Collision Detection funzioni, il mittente deve poter rilevare la trasmissione mentre sta trasmettendo ovvero prima di inviare l'ultimo bit del frame!
- $\square$  Il tempo di trasmissione di un frame deve essere almeno due volte il tempo di propagazione  $T_p$
- ☐ Quindi la prima stazione deve essere ancora in trasmissione dopo 2Tp

## Esempio

Una rete che utilizza il CSMA/CD ha un rate 10 Mbps. Se il tempo di propagazione massimo è 25,6 μs, qual è la dimensione minima del frame?

#### **Soluzione**

Il tempo di trasmissione minimo del frame è:

$$T_{fr} = 2 \times T_p = 51.2 \ \mu s.$$

Ciò significa, nel peggiore dei casi, che una stazione deve trasmettere per un periodo di 51,2 µs per poter rilevare la collisione.

La dimensione minima del frame è quindi 10 Mbps  $\times$  51,2  $\mu$ s = 512 bit o 64 byte.

Questa è proprio la dimensione minima del frame nell'Ethernet Standard

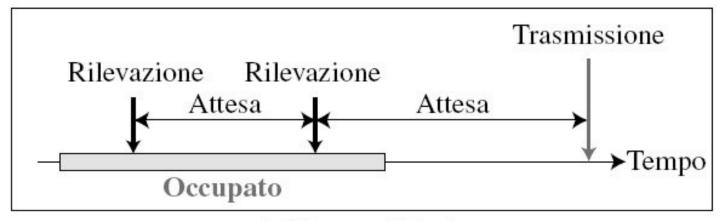
## Metodi di persistenza

Non persistente, 1-persistente, p-persistente (slottizzato)

- Cosa fa un nodo se trova il canale libero?
  - o Trasmette subito
    - Non persistente
    - 1-persistente
  - Trasmette con probabilità p
    - p-persistente
- Cosa fa un nodo se trova il canale occupato?
  - Desiste: riascolta dopo un tempo random
    - Non persistente
  - o Persiste: rimane in ascolto finché il canale non si è liberato
    - 1-persistente: trasmette con prob 1 quando si è liberato
    - p-persistente: trasmette con prob p quando si è liberato

## Non persistente

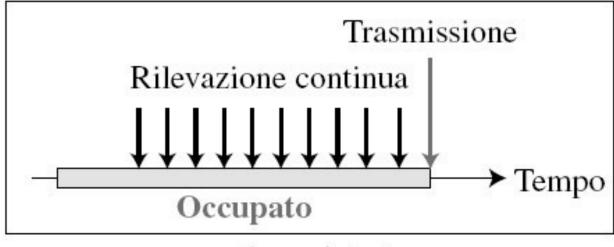
- O Se il canale è libero trasmette immediatamente
- Se il canale è occupato attende un tempo random e poi riascolta il canale (carrier sense a intervalli)
- ☐ Se collisione back-off



b. Non persistente

## 1 persistente

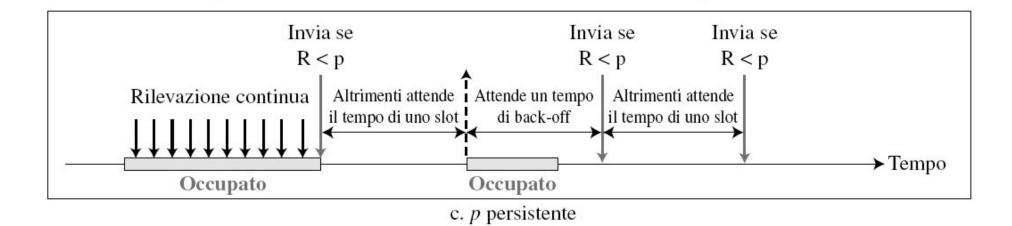
- □ Se il canale è libero trasmette immediatamente (p=1)
- Se il canale è occupato continua ad ascoltare (carrier sense continuo)
- Se collisione backoff



a. 1 persistente

## p persistente (slottizzato)

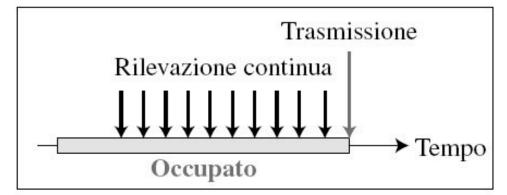
- Se il canale è libero
  - Trasmette con probabilità p
  - Aspetta l'inizio del prossimo slot con probabilità (1-p)
- Se il canale è occupato usa la procedura di back-off (attesa di un tempo random e nuovo ascolto del canale)
- Se collisione back-off



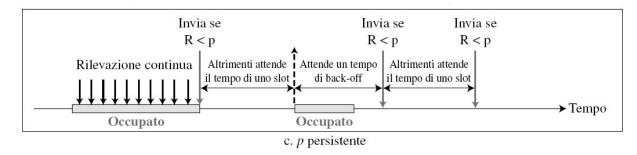
## Persistenza

# Rilevazione Rilevazione Attesa Attesa Occupato Trasmissione Attesa Tempo

b. Non persistente



a. 1 persistente



#### A cosa si riferisce?

all'ascolto del canale

## Algoritmo Ethernet CSMA/CD

- 1. NIC riceve il datagramma dal livello di rete, crea il frame
- 2. il NIC ascolta (sense) il canale:

se inattivo: avvia la trasmissione del frame

se occupato: attende che il canale sia libero, quindi trasmette

È un algoritmo 1-persistente!

- 3. Se il NIC trasmette l'intero frame senza collisioni, ok.
- 4. Se il NIC rileva un'altra trasmissione durante l'invio: interrompe e invia segnale di jam (48bit, per avvissare tutti gli altri NIC)
- 5. Dopo l'interruzione, il NIC entra nel backoff binario (esponenziale) :
  - dopo la m-esima collisione, NIC sceglie K a caso tra  $\{0,1,2,...,2^m-1\}$  . Il NIC attende K slot, torna al passaggio 2
  - più collisioni: intervallo di backoff più lungo (prima collisione {0,1}, seconda [0,3], ... decima [0,1023]...

(slot è il tempo per tramettere un frame di 512 bit)

## Efficienza di CSMA/CD

- t<sub>prop</sub> = massimo ritardo di propagazione tra 2 nodi nella LAN
- t<sub>trans</sub> = tempo per trasmettere il frame di dimensioni massime

$$efficiency = \frac{1}{1 + 5t_{prop}/t_{trans}}$$

- l'efficienza va a 1
  - quando  $t_{prop}$  va a 0
  - quando  $t_{trans}$  tende all'infinito (pacchetti grandi)
- prestazioni migliori di ALOHA (throughput massimo in condizioni ragionevoli ~50%): è semplice, economico, decentralizzato!

## Protocolli MAC a rotazione - motivazione

#### protocolli MAC di partizionamento dei canali:

- condividere il canale in *modo efficiente* ed *equo* a carico elevato
- inefficiente a basso carico: ritardo nell'accesso al canale, larghezza di banda 1/N allocata anche se solo 1 nodo attivo!

#### protocolli MAC ad accesso casuale

- efficiente a basso carico: il singolo nodo può utilizzare completamente il canale
- carico elevato: riduzione di banda per collisioni

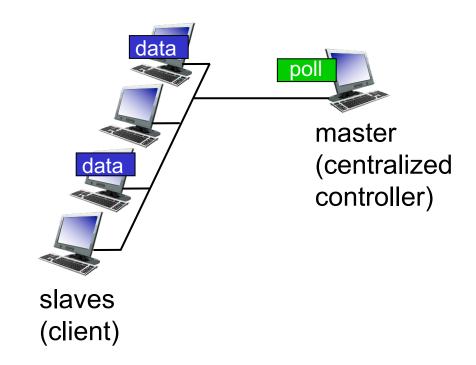
#### Protocolli a rotazione

prende il meglio da entrambi i mondi!

## Protocolli MAC a rotazione

#### polling:

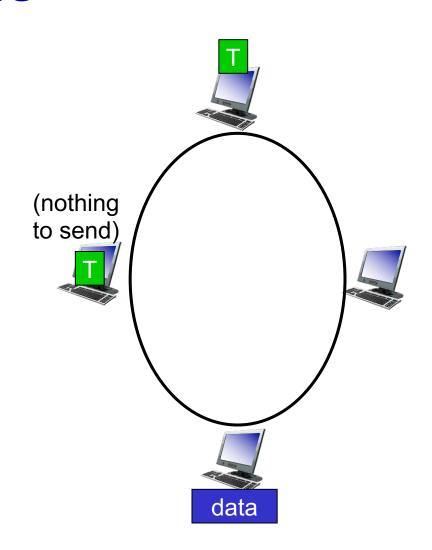
- il nodo master "invita" altri nodi a trasmettere a turno
  - se un nodo non ha niente da trasmettere si passa al prossimo
  - serve un protocollo per entrare/uscire
- tipicamente utilizzato con dispositivi "dumb"
- problemi:
  - overhead dovuto al polling
  - latenza di accesso (attesa turno)
  - single point of failure (master)



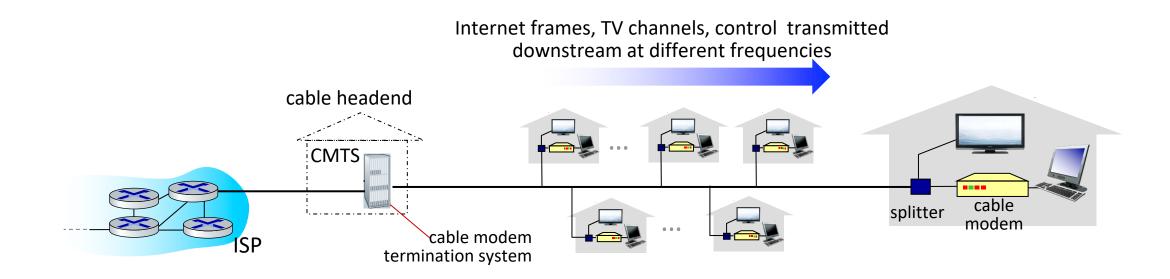
## Protocolli MAC a rotazione

#### token passing:

- token di controllo passato da un nodo al successivo in sequenza
  - di trasmette quando si possiede il token
- problemi:
  - token overhead (basso)
  - latenza
  - single point of failure (token)
    - se non viene passato il canale può fallire

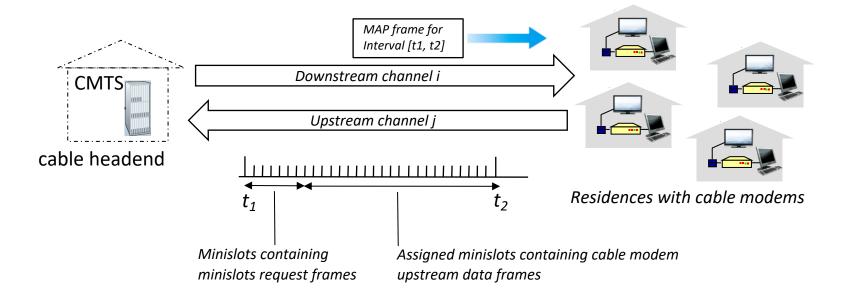


## Rete di accesso via cavo: FDM, TDM e accesso casuale!



- Canali multipli downstream (broadcast) FDM: fino a 1.6 Gbps/canale
  - un unico CMTS trasmette in questi canali
- Canali multipli upstream FDM (up to 1 Gbps/channel)
  - accesso multiplo: una parte degli slot assegnata via TDM, in più tutti gli utenti si contendono (random access) dei time slot

## Rete di accesso via cavo



#### DOCSIS: data over cable service interface specification

- FDM su canali upstream e downstream
- TDM upstream: alcuni assegnati, alcuni contesi
  - downstream MAP frame: assegna gli slot upstream
  - richieste di slot (e dati) possono essere trasmesse ad accesso casuale (binary backoff) in un sottoinsieme di slot

## Sintesi dei protocolli

- partizione dei canali, per tempo, frequenza (o codice)
  - Divisione di tempo, divisione di frequenza
- accesso casuale (dinamico),
  - ALOHA, S-ALOHA, CSMA, CSMA/CD
  - carrier sensing: facile in alcune tecnologie (filo), difficile in altre (wireless)
  - CSMA/CD utilizzato in Ethernet
  - CSMA/CA utilizzato in 802.11
- accesso a rotazione
  - polling dal sito centrale, passaggio di token
  - Bluetooth, FDDI, token ring