



Fondamenti di Internet e Reti

Antonio Capone, Matteo Cesana,
Ilario Filippini, Guido Maier



5 – Livello di Linea e Reti Locali

Antonio Capone, Matteo Cesana,
Ilario Filippini, Guido Maier

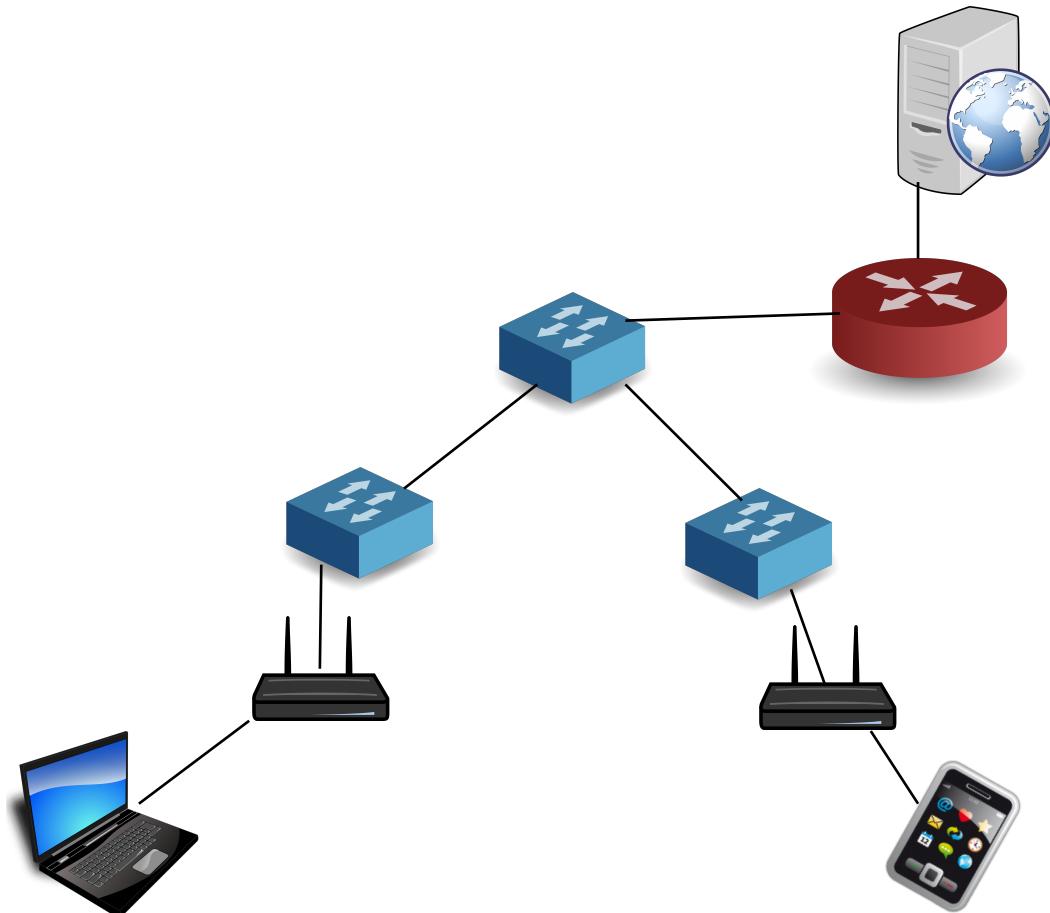
Outline

- **Introduzione**
- **Collegamenti punto-punto**
 - Framing, HDLC, PPP
- **Collegamenti broadcast**
 - Aloha, CSMA, CSMA-CD, CSMA-CA
- **Tecnologie di reti locali**
 - Ethernet, WiFi
- **Collegamenti commutati**
 - Switching, Spanning Tree, VLAN
- **Esempi di protocolli di linea**
 - HDLC, PPP
- **Spanning tree protocol**



Dal livello di rete al livello di linea

- Abbiamo visto che i router possono scambiare pacchetti IP attraversando collegamenti e reti locali di tipo eterogeneo
- Ma come viaggiano i pacchetti tra un router ed il successivo?
- Esistono indirizzi a livello di rete locale?
- Che differenza c'è tra i collegamenti diretti (P2P) e quelli condivisi (ad es. WiFi)?



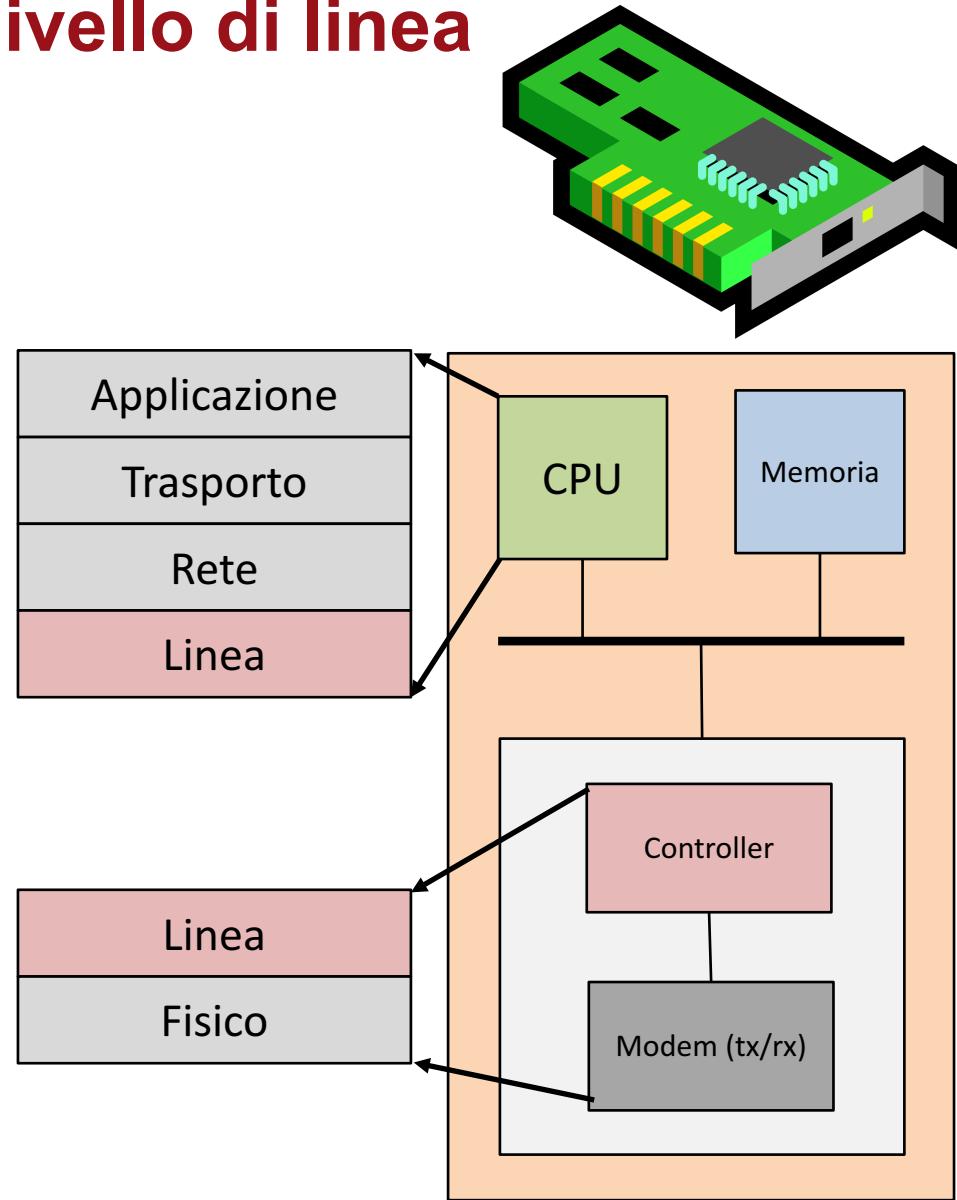
Il livello di linea

- E' il primo livello logico presente nella modalità a pacchetto
- **Funzionalità**
 - identificare logicamente i bit o gruppi di bit scambiati col livello fisico (framing)
 - segnalare o correggere gli errori (opzionale)
 - Multiplazione (opzionale)
 - Accesso multiplo (opzionale)



Dov'è implementato il livello di linea

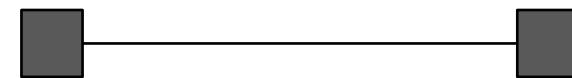
- Il livello di linea è normalmente parte della scheda di rete (**Network Interface Card - NIC**)
- Insieme al livello fisico è di solito implementato su chipset dedicato (controller)
- Alcune delle funzionalità (gestione degli indirizzi, preparazione della trama) sono svolte in software dall'host



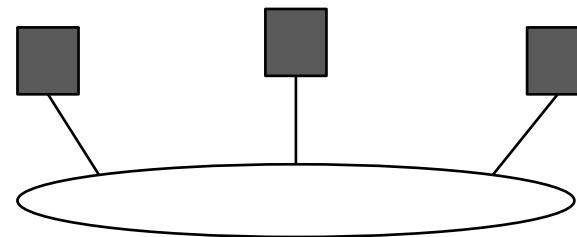
Tipi di livelli di linea

- **Esistono fondamentalmente tre tipologie di livelli di linea:**

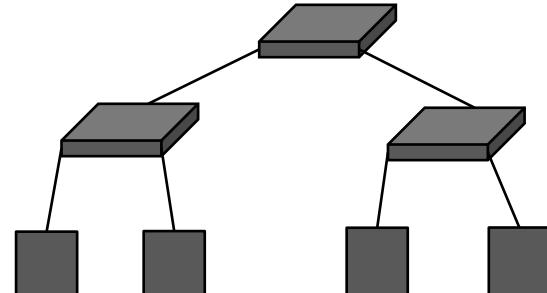
- Collegamenti punto-punto (P2P)



- Collegamenti broadcast



- Collegamenti commutati
(variante del P2P ma con altri elementi di rete locale)



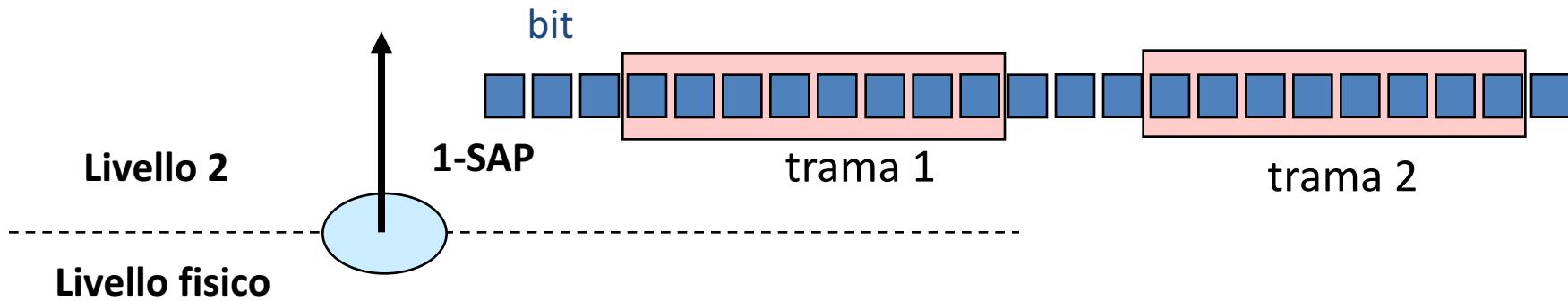


5a – Collegamenti punto-punto

Framing, HDLC, PPP

La costruzione della trama

- La prima funzione del livello logico è di individuare il significato dei bit scambiati con il livello fisico

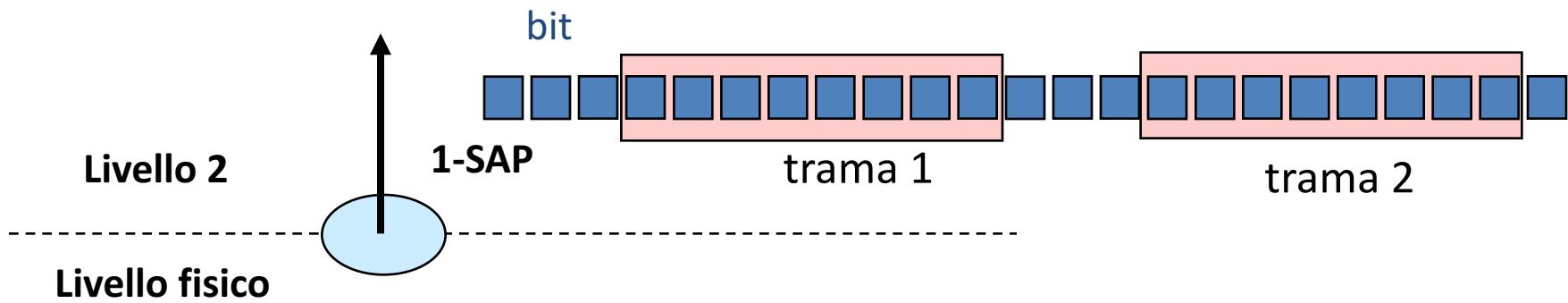


- Allo scopo i bit sono raggruppati in una struttura dati definita dal protocollo e chiamata “trama” (PDU-2)



La costruzione della trama

- Occorre un mezzo per identificare la posizione delle varie trame all'interno del flusso di bit



- Per questo si usano
 - i delimitatori di trama (particolare stringa di bit)
 - o segnalazioni passate dal livello fisico



Protocolli orientati al bit

- Si utilizzano dei “flag” (particolari sequenze di bit) per trovare l'allineamento di trama
- Esempio: HDLC
 - Sequenza di flag all'inizio e alla fine di una trama

0 1 1 1 1 1 1 0

- Problema: come impedire una casuale presenza della sequenza di flag nei dati ?

Soluzione:

Bit stuffing: PRIMA DI TRASMETTERE, si inserisce uno 0 dopo aver osservato cinque 1 consecutivi (n.b. *indipendentemente dal valore del bit successivo ai cinque 1 consecutivi*)



Bit Stuffing

informazione
da trasmettere

1111000111111000010010101111101

inserimento bit di stuffing dopo 5 “1”

trama 0111110 111100011111010000100101011111001 0111110
flag flag

ricezione 0111110 111100011111010000100101011111001 0111110
↑
riconoscimento
flag d'inizio ↑
eliminazione di un bit
dopo 5 uno consecutivi →
riconoscimento
flag di fine



Controllo d'errore

- Abbiamo già trattato il controllo d'errore e la ritrasmissione (ARQ) per il livello di trasporto
- A differenza del livello di trasporto dove l'obiettivo è il recupero dei segmenti persi, nel livello di linea l'obiettivo è il recupero degli errori di livello fisico
- Esiste la ritrasmissione anche nel caso di collegamento broadcast e può servire a recuperare anche le contese (collisioni sul canale)



Multiplazione

- Nei collegamenti punto-punto i protocolli di linea possono essere istanziati su più canali fisici
- In alcuni casi un canale viene diviso in più sotto-canali a livello fisico
- Quest' operazione viene definita multiplazione fisica



Multiplazione fisica

- La multiplazione a livello fisico consiste nel suddividere la capacità di un canale a velocità costante in sottocanali di velocità costante (e inferiore)



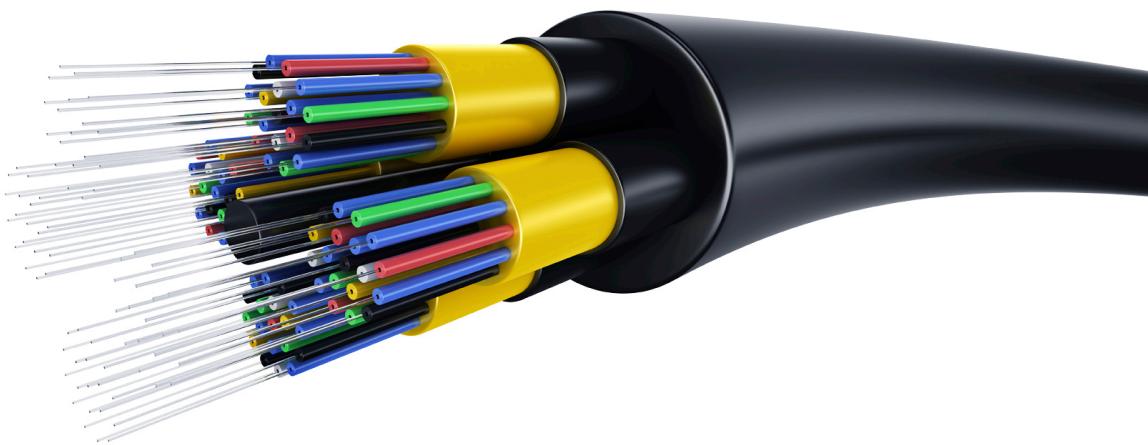
Multiplazione fisica

- Si distingue per la caratteristica fisica attraverso la quale i diversi segnali vengono separati
 - Divisione di spazio
 - Divisione di frequenza (FDM Frequency Division Multiplexing)
 - Divisione di tempo (TDM Time Division Multiplexing)
 - Divisione di codice (CDM Code Division Multiplexing)
 - Divisione di lunghezza d'onda (WDM Wavelength Division Multiplexing)



Multiplazione a divisione di spazio

- L'esempio tipico si ha in un cavo a coppie, usato per concentrare i doppini d'utente in telefonia
- Oppure in cavi che portano diverse fibre ottiche
- ...

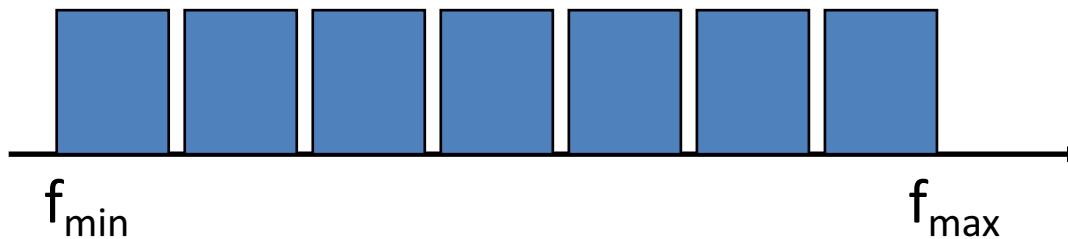


Multiplazione FDM

- Il mezzo trasmittivo è caratterizzato da una banda di frequenze utilizzabili



- la banda complessiva può essere divisa in sotto-bande cui associare un canale

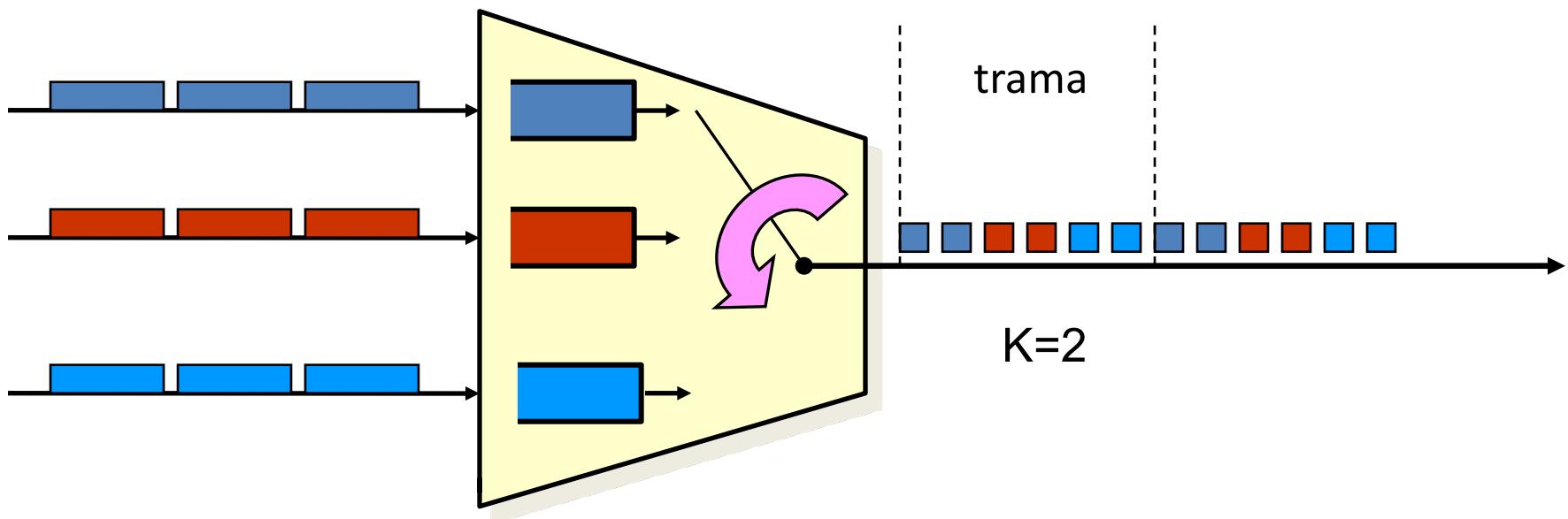


- Esempi: Digital TV, ADSL, ...



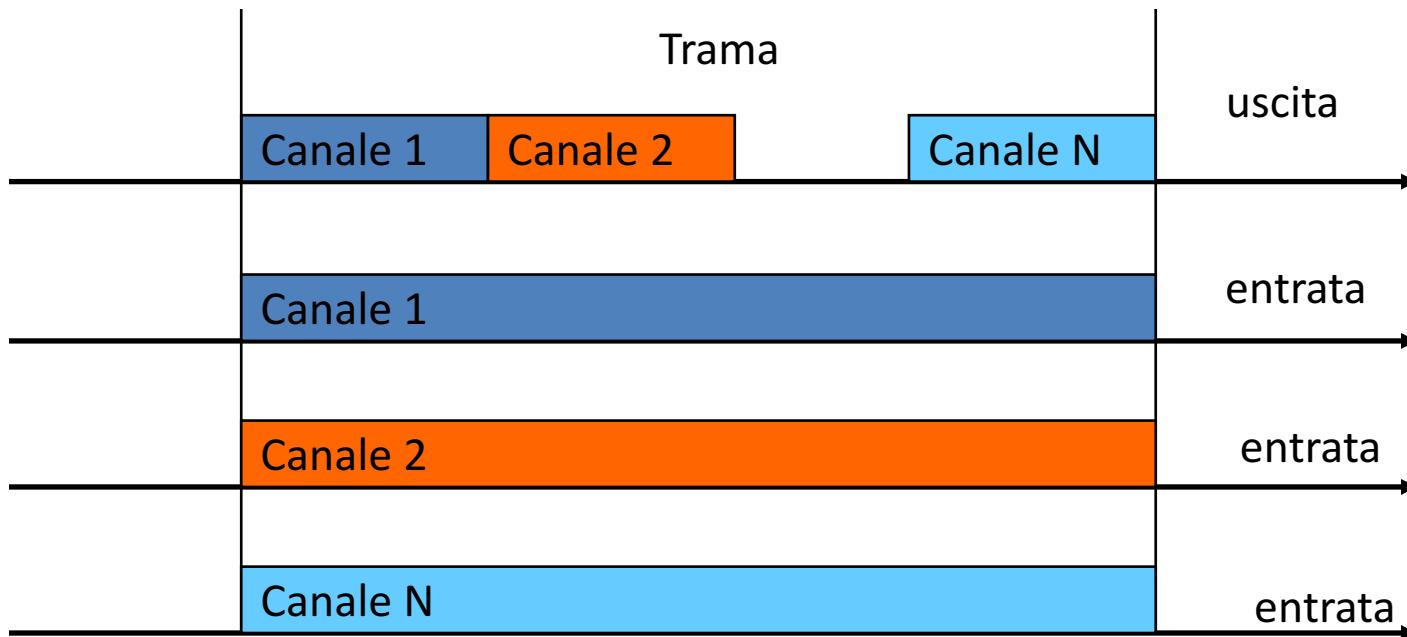
Multiplazione TDM

- I bit di N flussi vengono raccolti in code e trasmessi sul flusso di uscita a gruppi di K (interlacciamento di K bit)



Multiplazione TDM

- La durata della trama deve uguagliare l'intervallo di tempo in cui sul singolo canale in entrata arrivano i bit in numero pari a quelli trasmessi nella trama



TDM: relazioni fra velocità

V: velocità del flusso tributario (entrata)

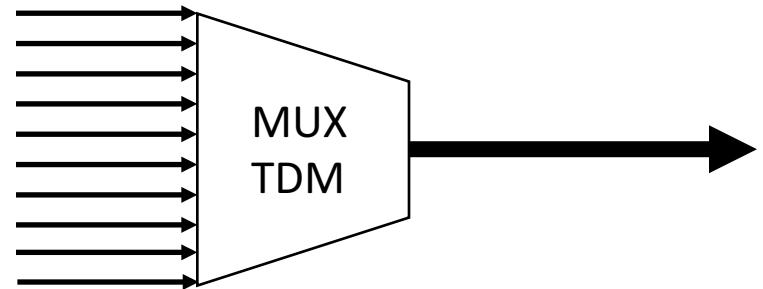
W: velocità del multiplex

N: n. di tributari

k: grado di interlacciamento (bit nello slot)

T_T : durata della trama

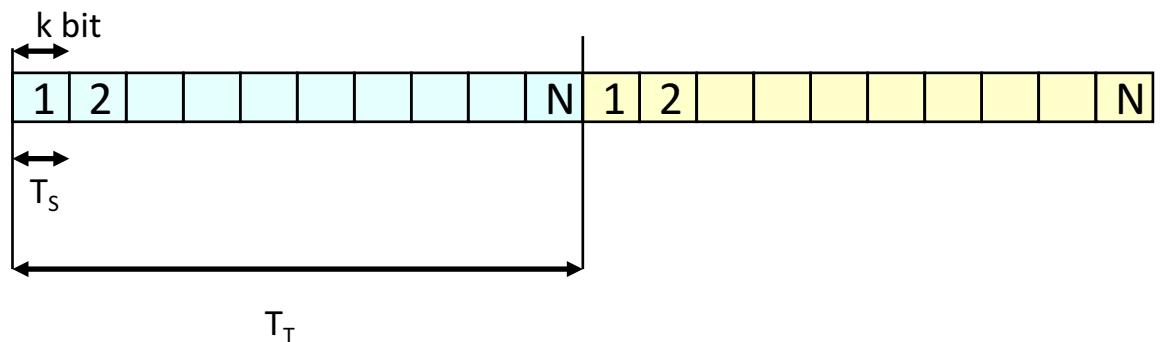
T_S : durata dello slot



$$T_S = \frac{T_T}{N}$$

$$k = W \cdot T_S \Rightarrow T_T = N \cdot T_S = N \frac{k}{W} = \frac{k}{V}$$

$$V = \frac{k}{T_T} = \frac{W}{N}$$





5b – Collegamenti broadcast

Aloha, CSMA, CSMA-CD, CSMA-CA

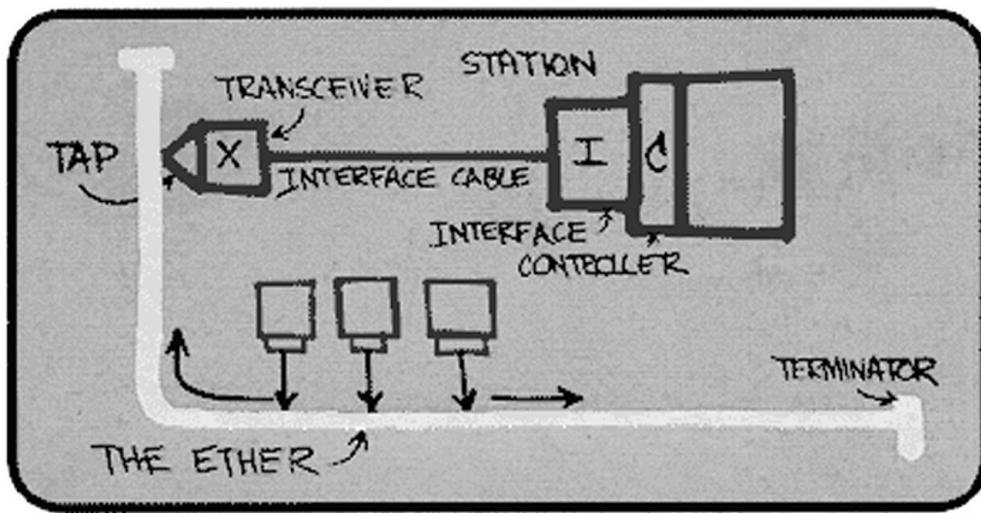
Gli esordi delle reti locali (anni '70 secolo scorso)

- **La funzione di rete è un'operazione che richiede capacità computazionale** (analisi dell'header, lookup alla tabella di instradamento, inoltro alla coda d'uscita, gestione della coda, ecc.)
- All'inizio della storia di Internet quando i nodi della rete ARPANET gestivano collegamenti dell'ordine di poche decine di kbps, esistevano già reti locali con velocità dell'ordine dei Mbps
- **Il segreto era il canale broadcast** senza funzione di rete (commutazione/switching)
- Tutti ricevono le trame, solo il destinatario preleva la trama e la inoltra ai livelli superiori.



Gli esordi delle reti locali (anni '70 secolo scorso)

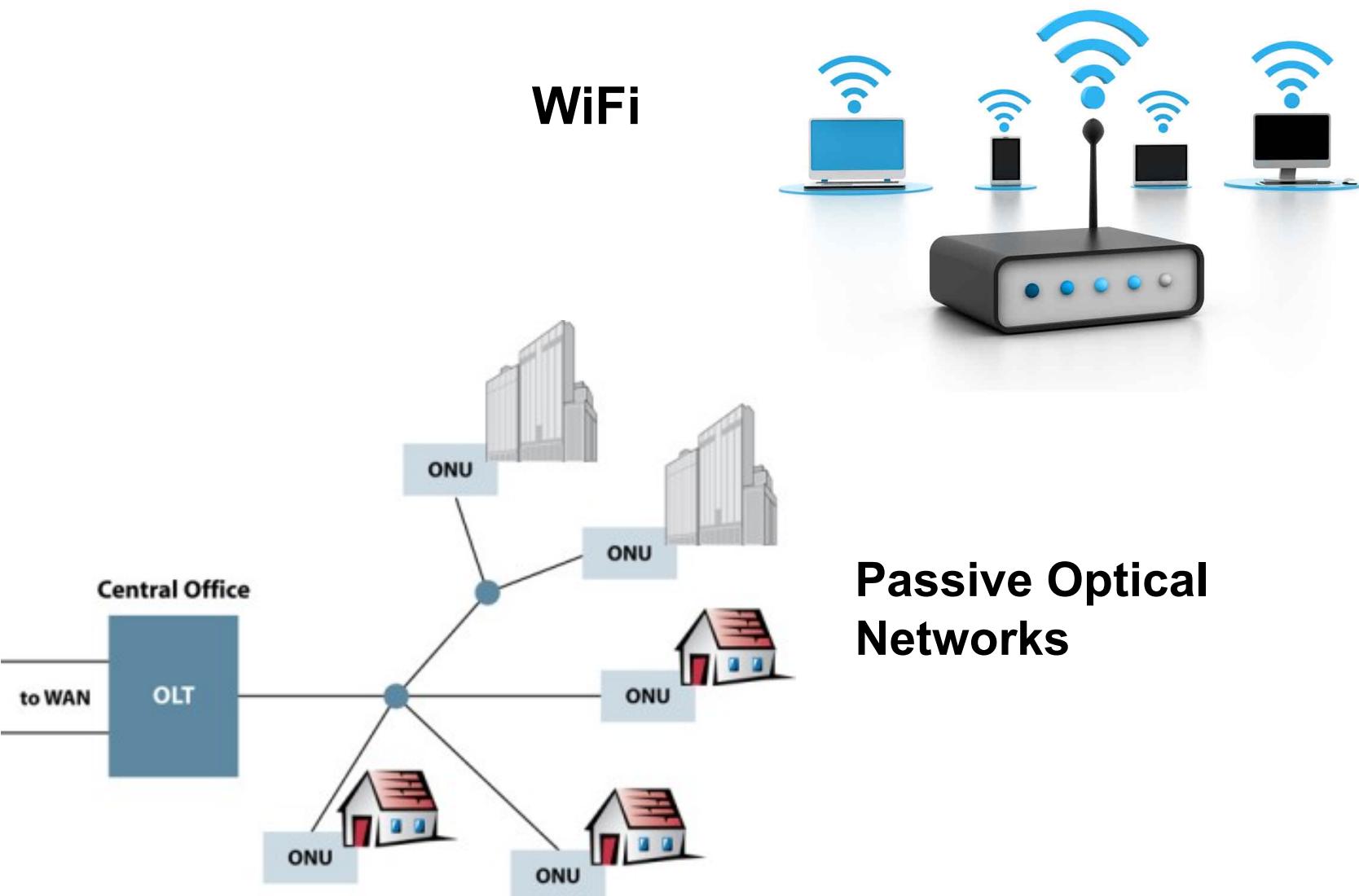
Ethernet 1976



AlohaNET 1971



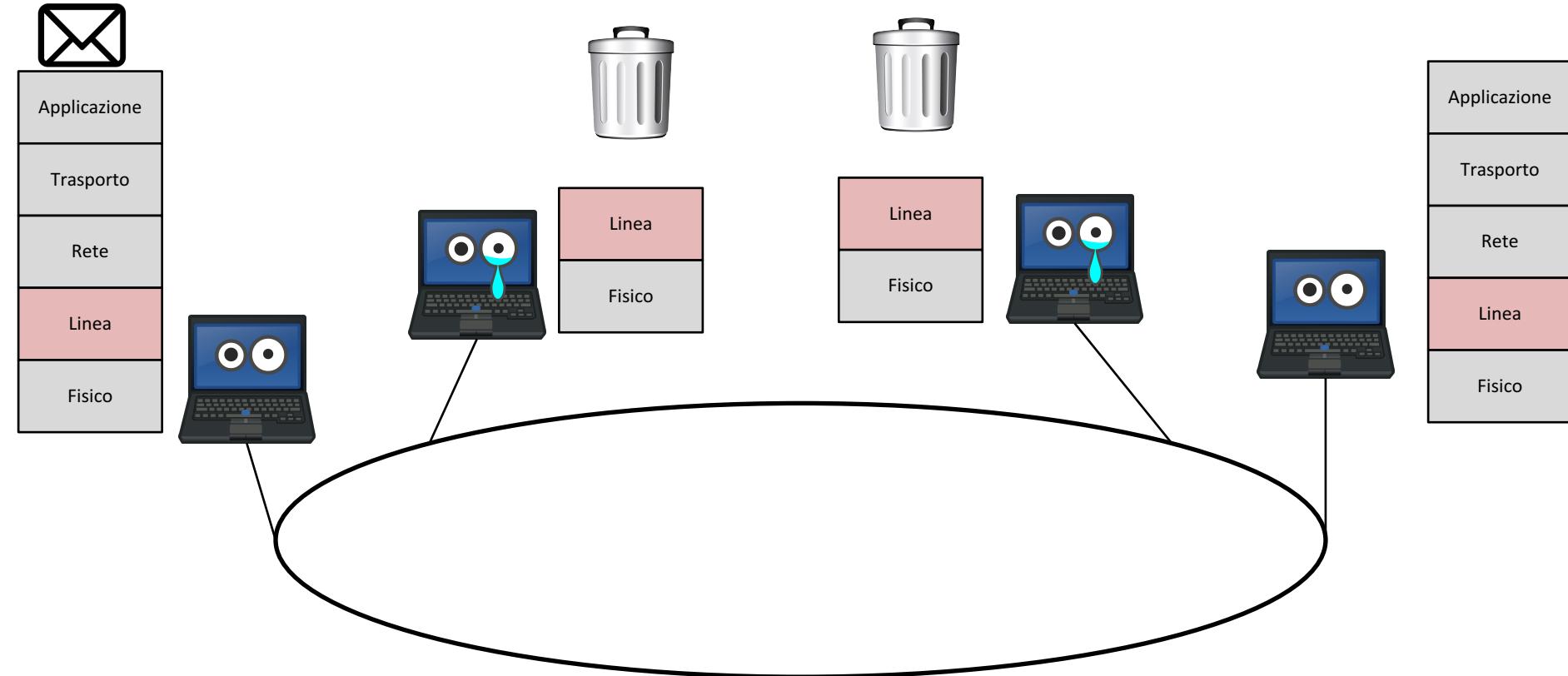
Le reti broadcast oggi



Passive Optical Networks



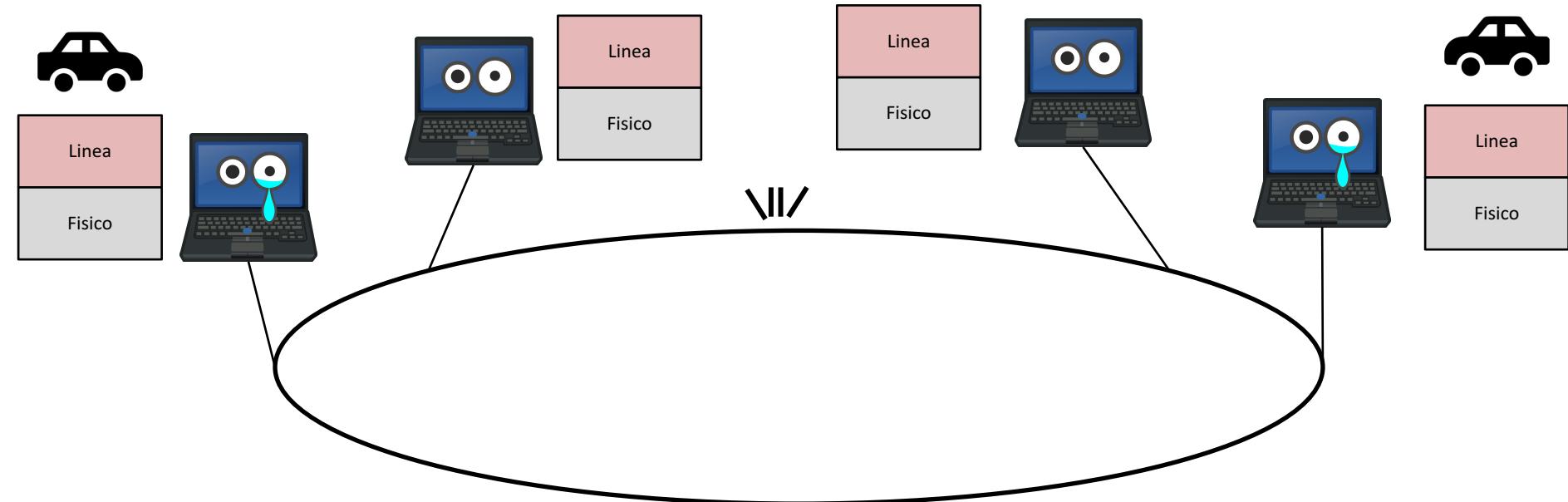
Canale broadcast



- **La scelta se inoltrare ai livelli superiori o scartare dipende dall'indirizzo di destinazione**
- **Ah, un altro indirizzo! ... su questo punto dobbiamo tornare**



Canale broadcast



- Ma sul canale broadcast le trasmissioni contemporanee (o quasi) provocano “collisioni”



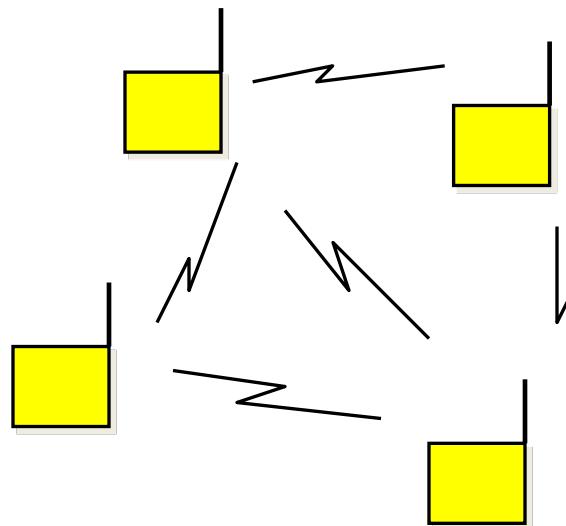
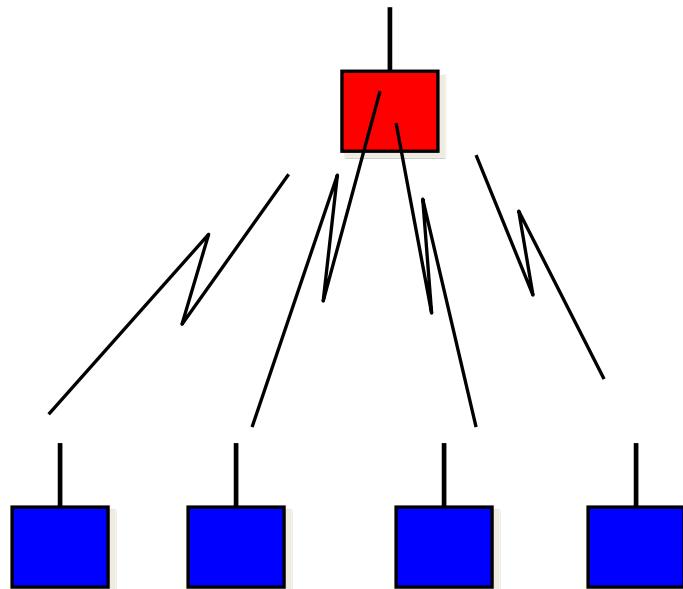
L'accesso multiplo

- L'accesso multiplo è la funzione che consente di regolare l'accesso al canale ed evitare le collisioni
- La funzione di accesso multiplo può essere implementata
 - a livello fisico, dividendo staticamente le risorse tra le stazioni (host)
 - a livello del protocollo di linea, gestendo l'accesso pacchetto per pacchetto



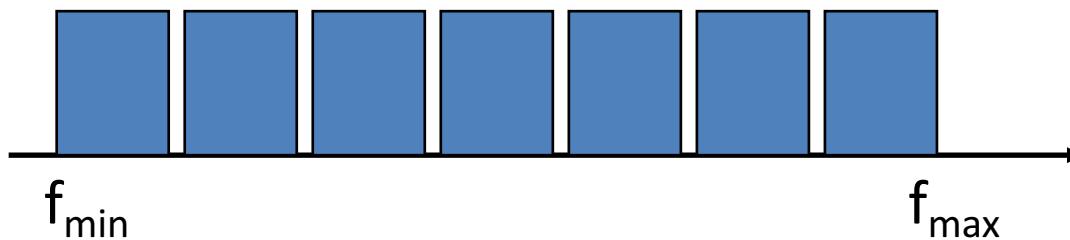
L'accesso multiplo fisico

- L'accesso multiplo fisico è equivalente alla moltiplicazione, ma è relativo al caso in cui i diversi sottocanali sono gestiti da trasmettitori diversi
- Esempio: trasmissione radio con mezzo condiviso



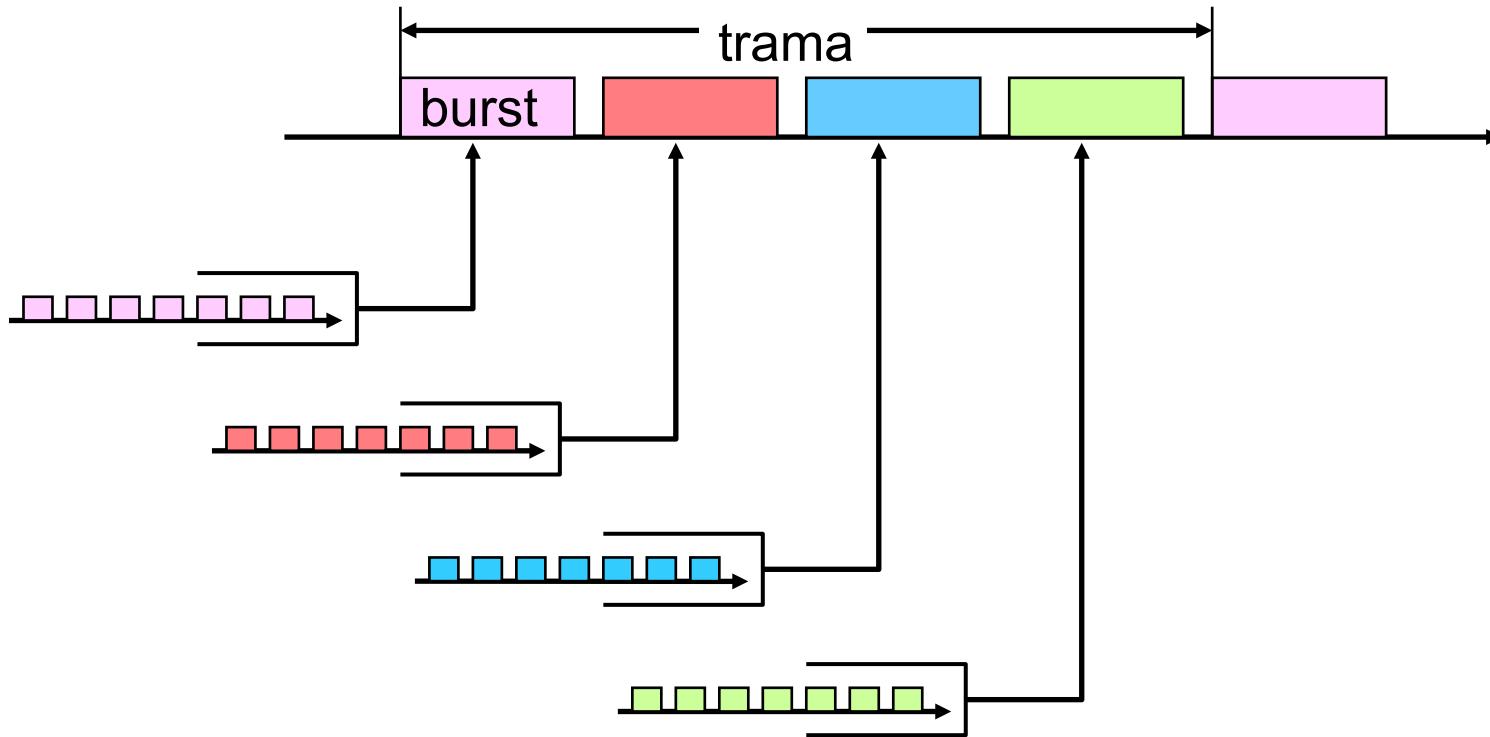
FDMA (Frequency Division Multiple Access)

- E' completamente equivalente al FDM
- Esempi: Canali WiFi, canali cellulari, ...



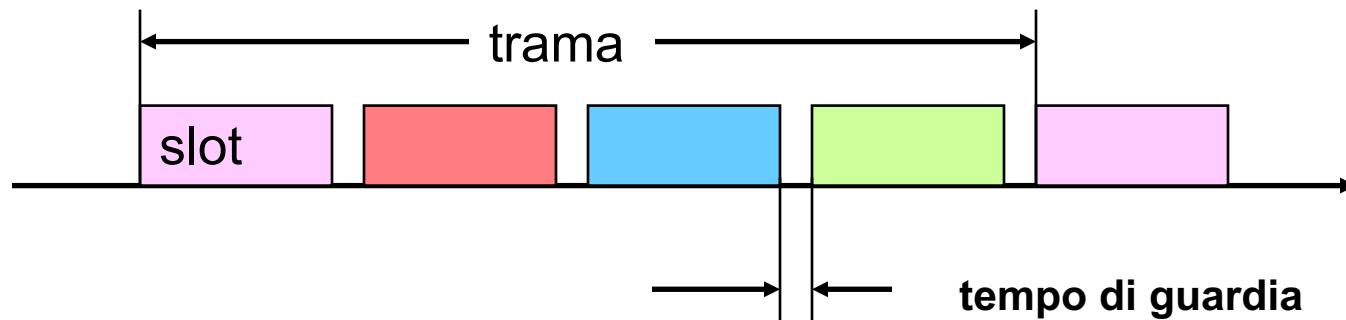
TDMA

- E' l'analogo del TDM
- Vengono definiti degli "slot" temporali dedicati alla trasmissione delle diverse stazioni



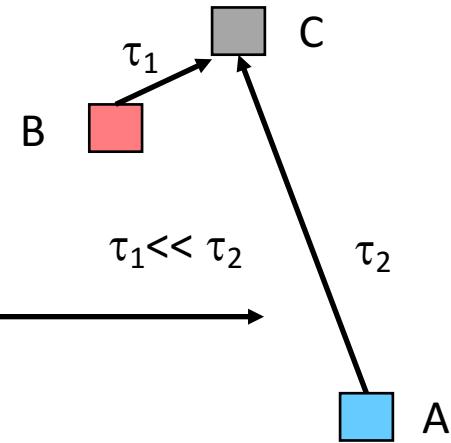
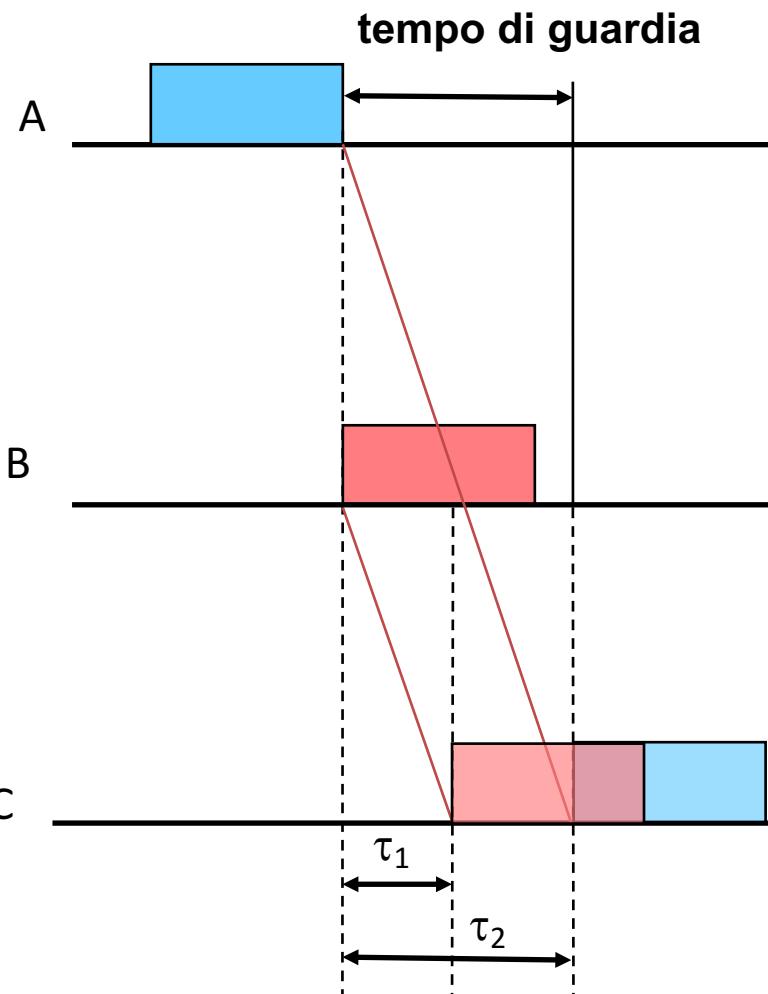
TDMA

- il flusso di bit fra i vari utenti generalmente non è sincrono (con alcune importanti eccezioni)
- il ricevitore deve sincronizzarsi su un particolare flusso
- vanno adottati “tempi di guardia” fra gli slot



TDMA

- Perché i tempi di guardia?



Duplexing

- È la modalità con la quale si ricavano i due sensi di trasmissione da un unico mezzo trasmisivo

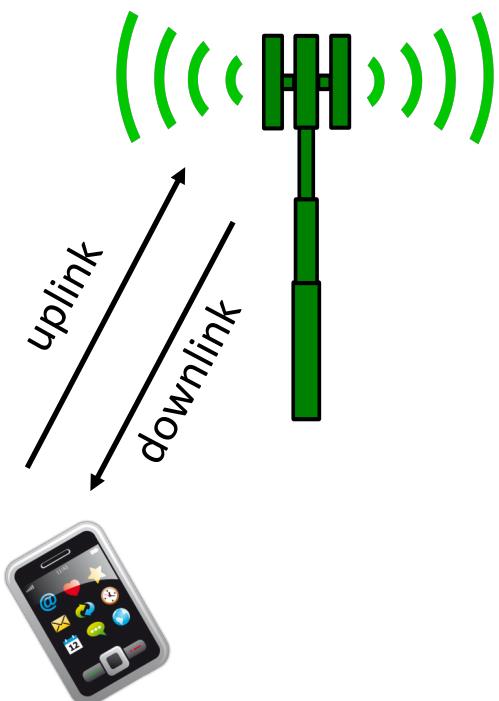


- Può essere visto come un caso particolare di accesso multiplo
- In alcuni casi particolari si può ottenere che si possa trasmettere e ricevere contemporaneamente (**FULL DUPLEX** a livello fisico)



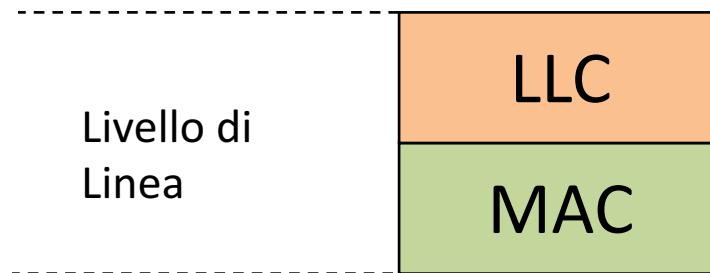
Duplexing

- Nel caso in cui il canale non sia full duplex a livello fisico occorre ricorrere a tecniche di suddivisione della capacità trasmissiva:
 - a divisione di spazio
 - a divisione di frequenza (FDD Frequency Division Duplexing)
 - a divisione di tempo (TDD Time Division Duplexing)
- Esempio: downlink e uplink cellulare



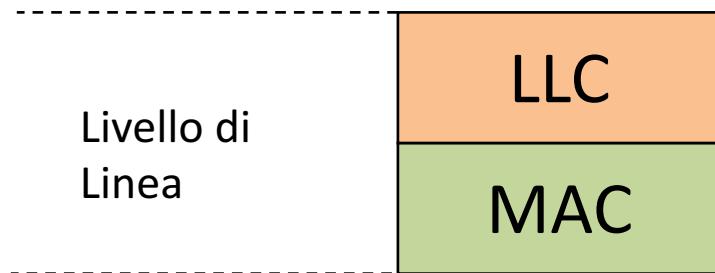
Accesso multiplo livello di linea

- A livello di pacchetto (protocollo di linea) l'accesso multiplo è gestito con dei meccanismi che regolano l'istante di trasmissione dei singoli pacchetti
- Il coordinamento può essere gestito da una entità centrale, ma molto più spesso (e nei casi di nostro interesse) è gestito in modo distribuito dalle singole stazioni
- In questi casi il livello di linea è diviso in due sotto-livelli: **MAC (Medium Access Control), LLC (Logical Link Control)**



Accesso multiplo livello di linea

- A livello di pacchetto (protocollo di linea) l'accesso multiplo è gestito con dei meccanismi che regolano l'istante di trasmissione dei singoli pacchetti
- Il coordinamento può essere gestito da una entità centrale, ma molto più spesso (e nei casi di nostro interesse) è gestito in modo distribuito dalle singole stazioni
- In questi casi il livello di linea è diviso in due sotto-livelli: **MAC (Medium Access Control), LLC (Logical Link Control)**

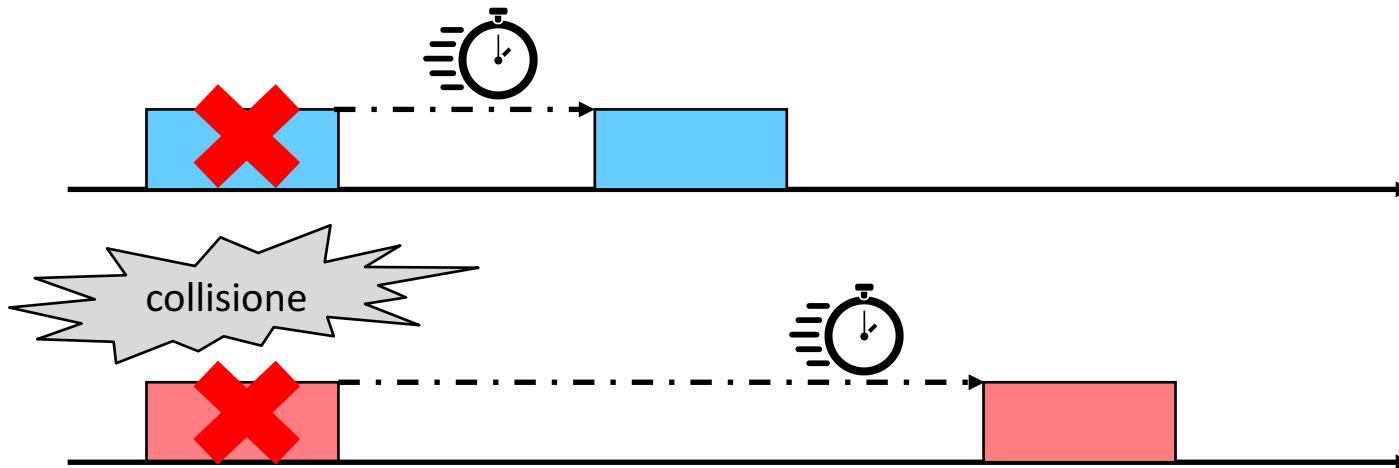


Il livello MAC si occupa dell'accesso multiplo mentre il livello LLC delle altre funzioni tipiche del livello di linea

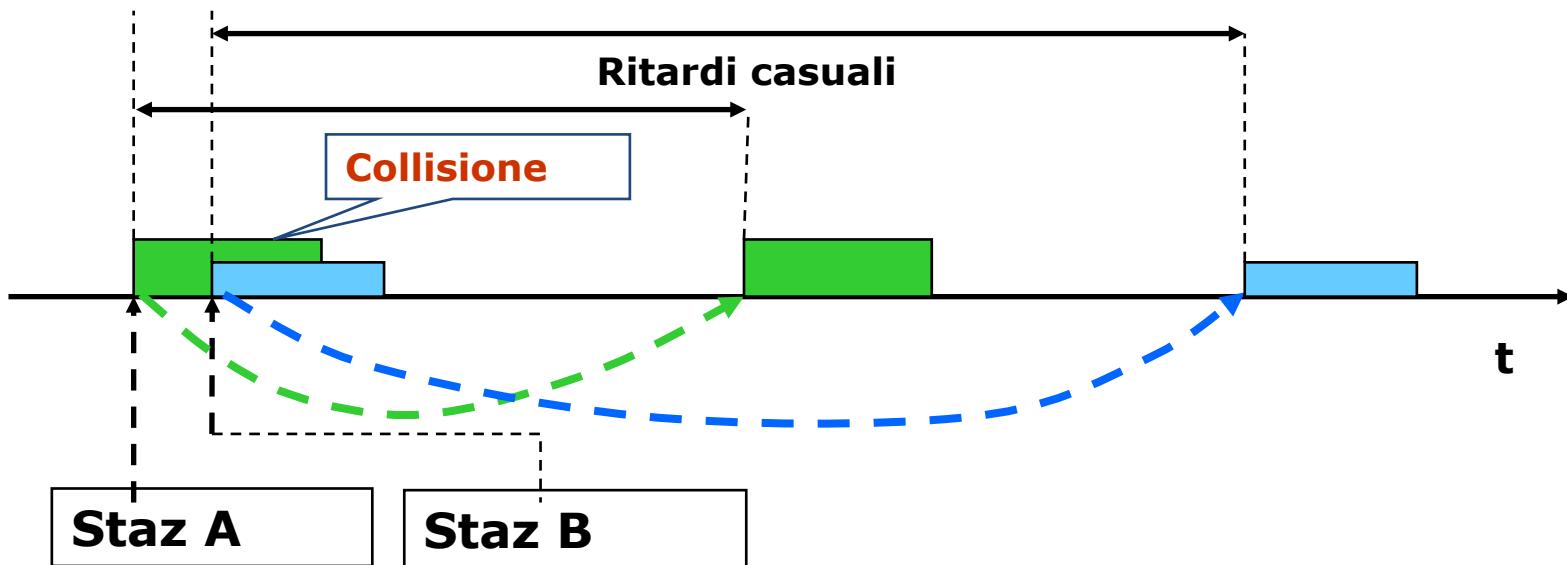


Accesso multiplo casuale

- Una categoria molto utilizzata di meccanismi di accesso multiplo a livello di linea è denominata **accesso casuale**
- Una stazione che ha un pacchetto decide autonomamente quando trasmettere osservando il canale
- Se la trasmissione collide, le stazioni coinvolte ritrasmettono dopo un tempo casuale
- La casualità consente con buona probabilità che la collisione non si ripresenti

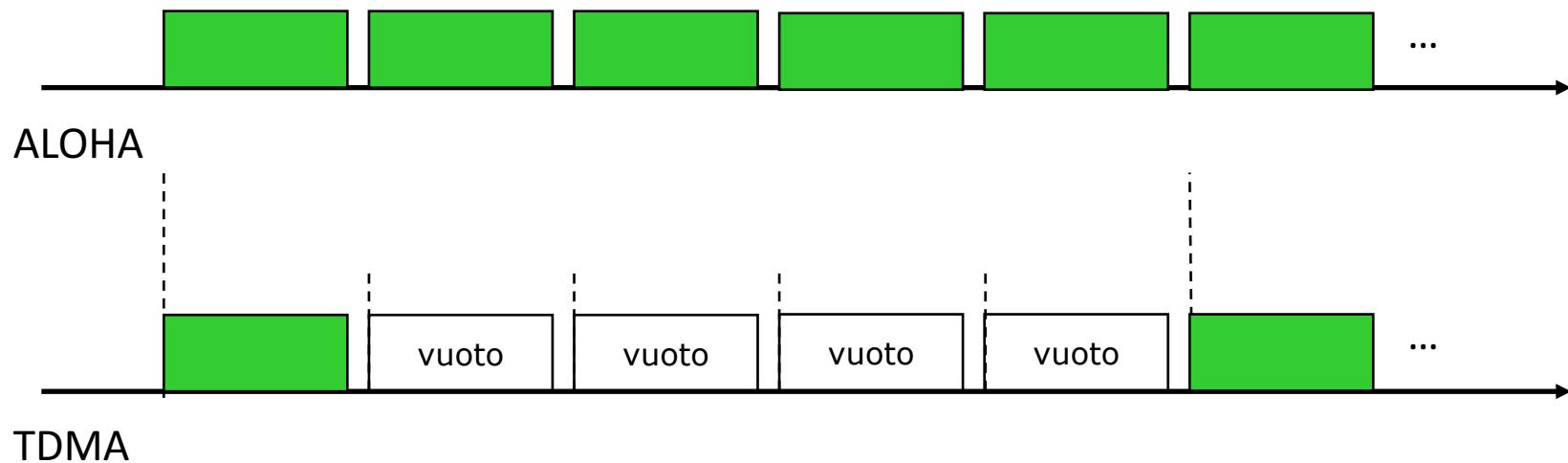


- E' il protocollo della prima rete locale utilizzata (AlohaNET)
- Estremamente semplice:
 - Una stazione che ha un pacchetto da trasmettere lo trasmette subito senza osservare il canale
 - Se c'è una collisione lo ritrasmette dopo un tempo casuale



ALOHA

- Se c'è solo una stazione che trasmette il vantaggio dell'ALOHA è evidente:

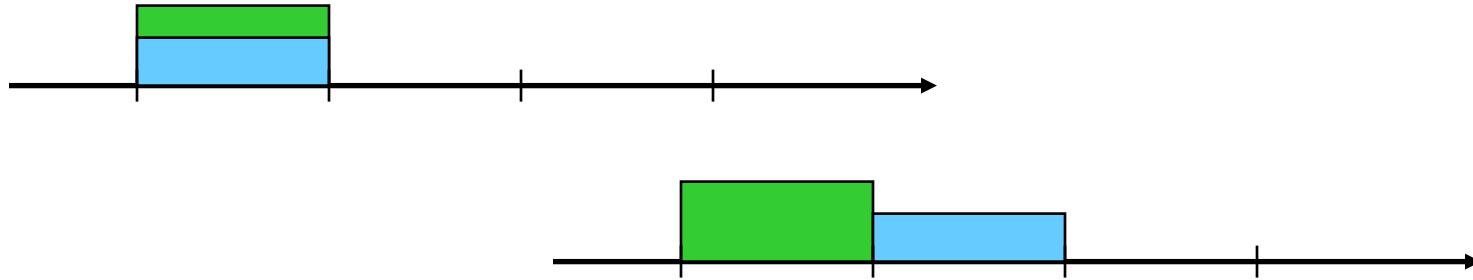


- Ma cosa succede con N stazioni?



Slotted ALOHA

- Facciamo una piccola analisi partendo da una variante dell'ALOHA che assume trasmissioni sincronizzate in slot: **Slotted ALOHA**



- In questo caso le collisioni ci possono essere solo se le trasmissioni sono nello stesso slot



Analisi Slotted ALOHA

- Consideriamo uno scenario con
 - N stazioni
 - Ogni stazione trasmette in uno slot con probabilità p
- Se una stazione trasmette, la probabilità di successo è data dalla probabilità che le altre $N-1$ non trasmettano:

$$P_S = (1 - p)^{N-1}$$



Analisi Slotted ALOHA

- La probabilità che in uno slot arbitrario una particolare stazione trasmetta e abbia successo è dunque

$$p(1 - p)^{N-1}$$

- E dunque la probabilità che una qualunque stazione trasmetta e abbia successo è

$$S = Np(1 - p)^{N-1}$$

- Questo è anche il numero medio di trasmissioni con successo in uno slot, che chiamiamo *throughput* (S)



Analisi Slotted ALOHA

- Il numero medio di trasmissioni sul canale, che chiamiamo traffico (G) è dato da:

$$G = Np$$

- Sostituendo nella formula del throughput $p = G/N$ si ha:

$$S = Np(1 - p)^{N-1} = G \left(1 - \frac{G}{N}\right)^{N-1}$$

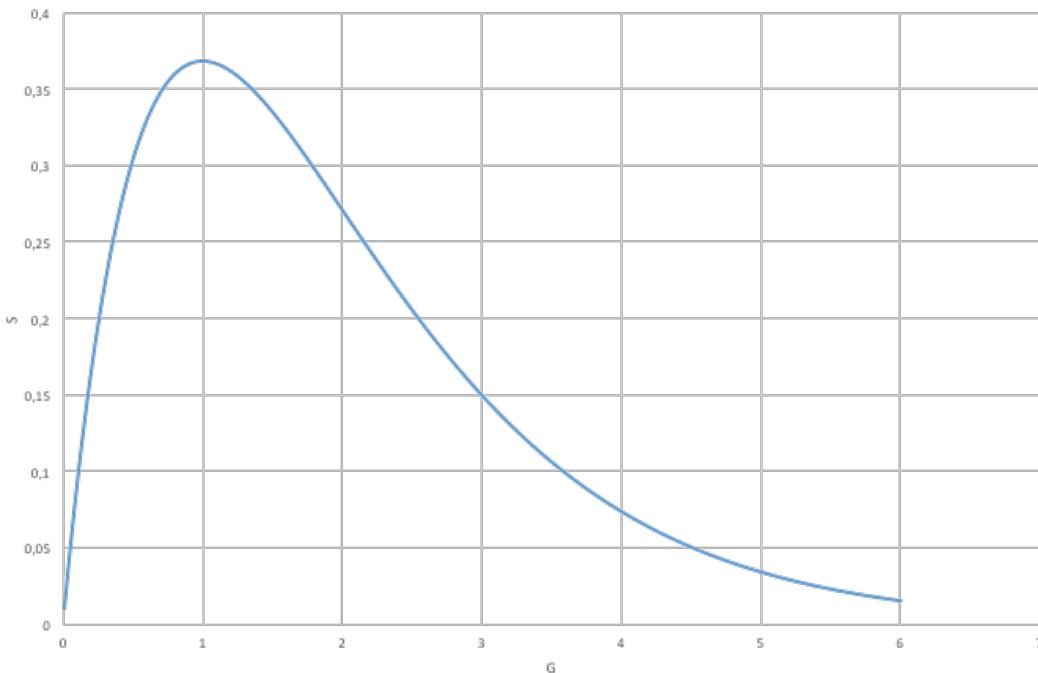
- Questa formula ci dà il numero medio di successi in funzione del numero medio di trasmissioni
- E' dunque la frazione di slot utilizzati proficuamente



Analisi Slotted ALOHA

- Il limite per N che tende ad infinito del throughput è noto (vedi corso Analisi 1) ed è

$$S = Ge^{-G}$$



Massimo in
 $G = 1$
 $S = 1/e \cong 0.37$



Analisi ALOHA

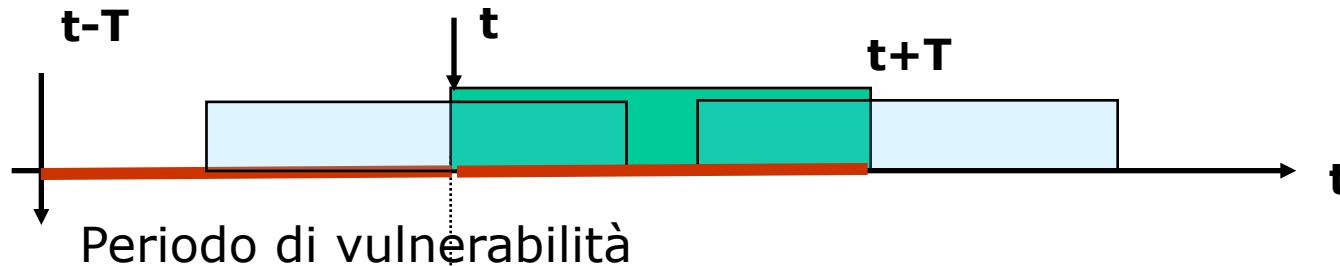
- Nel caso dell'ALOHA (niente slot), l'analisi è molto simile
- Basta osservare che, essendo possibili collisioni con sovrapposizione parziali, per aver successo non deve trasmettere nessun altro sia nell'intervallo di tempo prima che in quello dopo l'inizio della trasmissione considerata

$$P_S = (1 - p)^{2(N-1)}$$

- E quindi:

$$S = Np(1 - p)^{2(N-1)}$$

$$S = G \left(1 - \frac{G}{N}\right)^{2(N-1)}$$

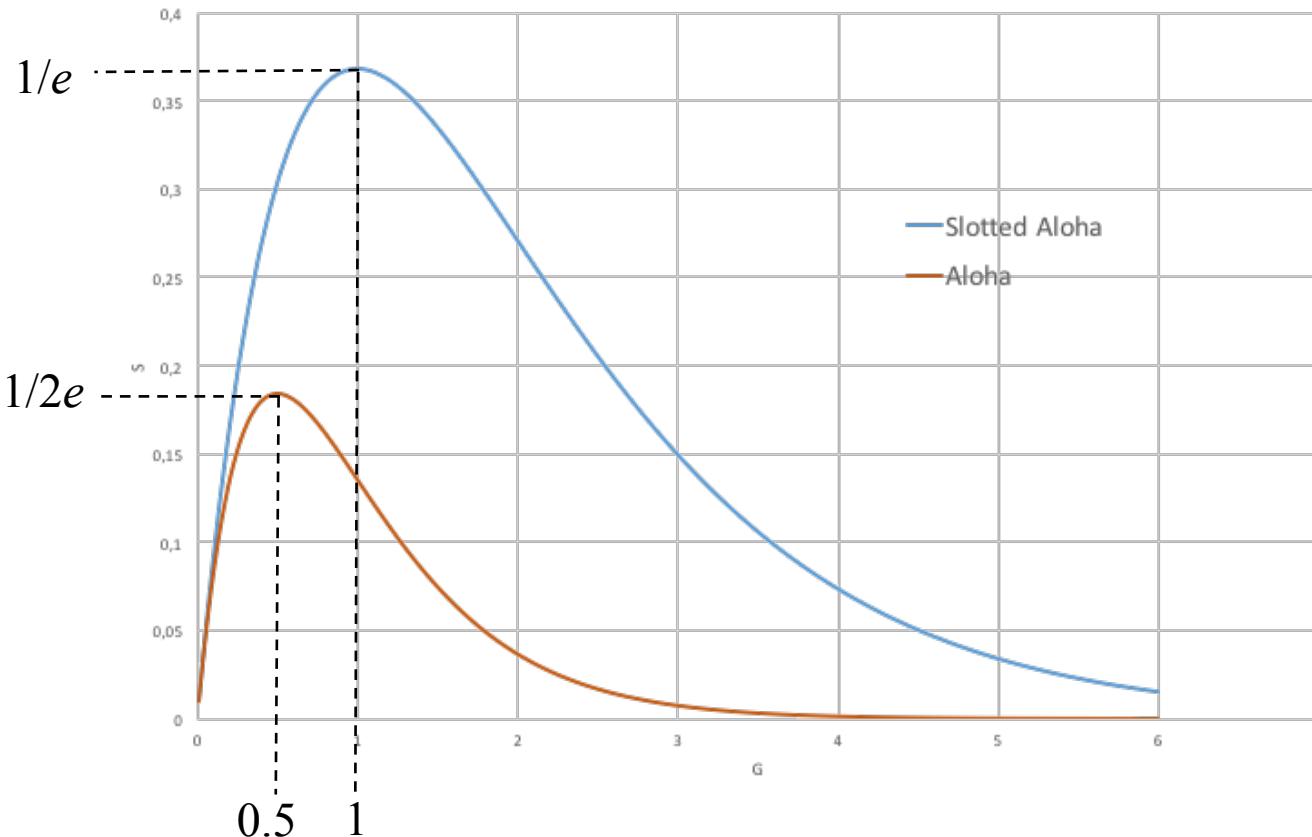


Analisi ALOHA

- Facendo il limite per N che tende ad infinito

$$S = Ge^{-2G}$$

Massimo in
 $G = 0.5$
 $S = 1/2e \approx 0.18$



Ascoltare prima di parlare

- Aloha è un protocollo “maleducato” che non verifica che qualcuno stia usando il canale prima di trasmettere (in realtà adatto a lunghi ritardi di propagazione)
- La buona educazione ci dice che non bisogna sovrapporsi a chi sta già parlando (“ascoltare prima di parlare”)
- Anche se a volte è difficile dire qualcosa se l’interlocutore parla troppo ☹



Ascoltare prima di parlare: CSMA

- **Carrier Sense Multiple Access (CSMA)** è un protocollo “educato”
 - Una stazione che ha una trama da trasmettere, prima verifica che il canale sia libero rilevando la presenza di segnale a livello fisico (carrier o portante)
 - (\rightarrow il motivo del nome si chiarirà parlando del livello fisico)
 - Se il canale è libero, la stazione trasmette la trama
 - Se il canale è occupato
 - La trasmissione viene rimandata ripetendo il sensing dopo un tempo casuale
 - In altre implementazioni, la stazione resta in ascolto e trasmette appena il canale si è liberato

Ma allora niente più collisioni???



Ascoltare prima di parlare

- In realtà sappiamo che anche in questo caso le collisioni di chi parla contemporaneamente sono sempre possibili



Perché si possono verificare lo stesso le collisioni?

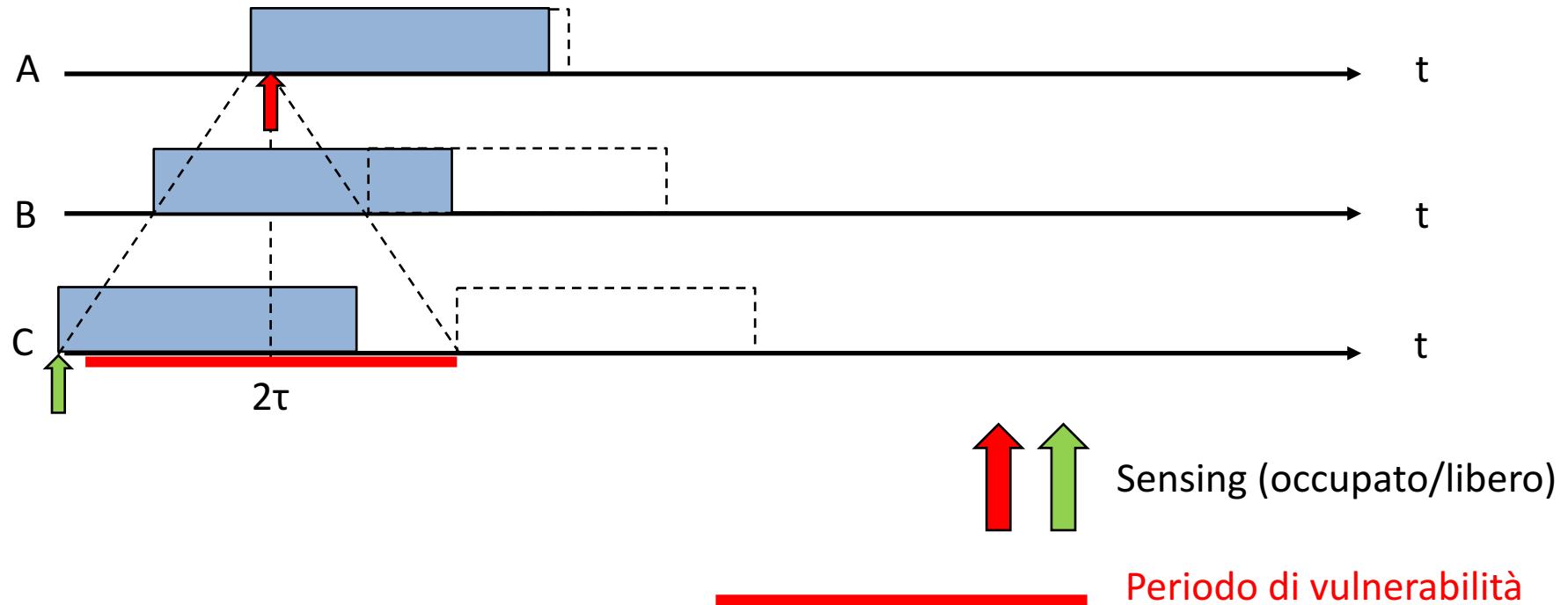
Il motivo è legato al ritardo di propagazione (...)



Carrier Sense Multiple Access (CSMA), $T > \tau$

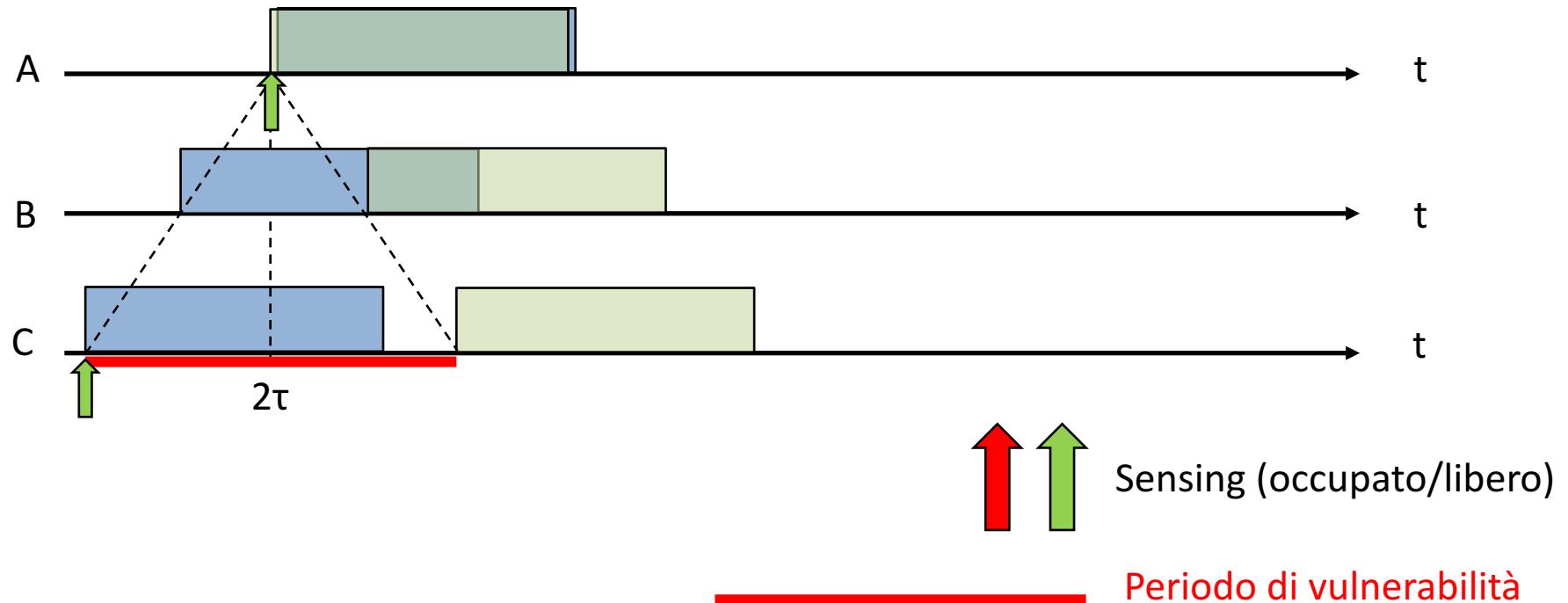
[Kleinrock 1975]

- C trova libero \rightarrow trasmette
- A trova il canale occupato \rightarrow non trasmette \rightarrow nessuna collisione



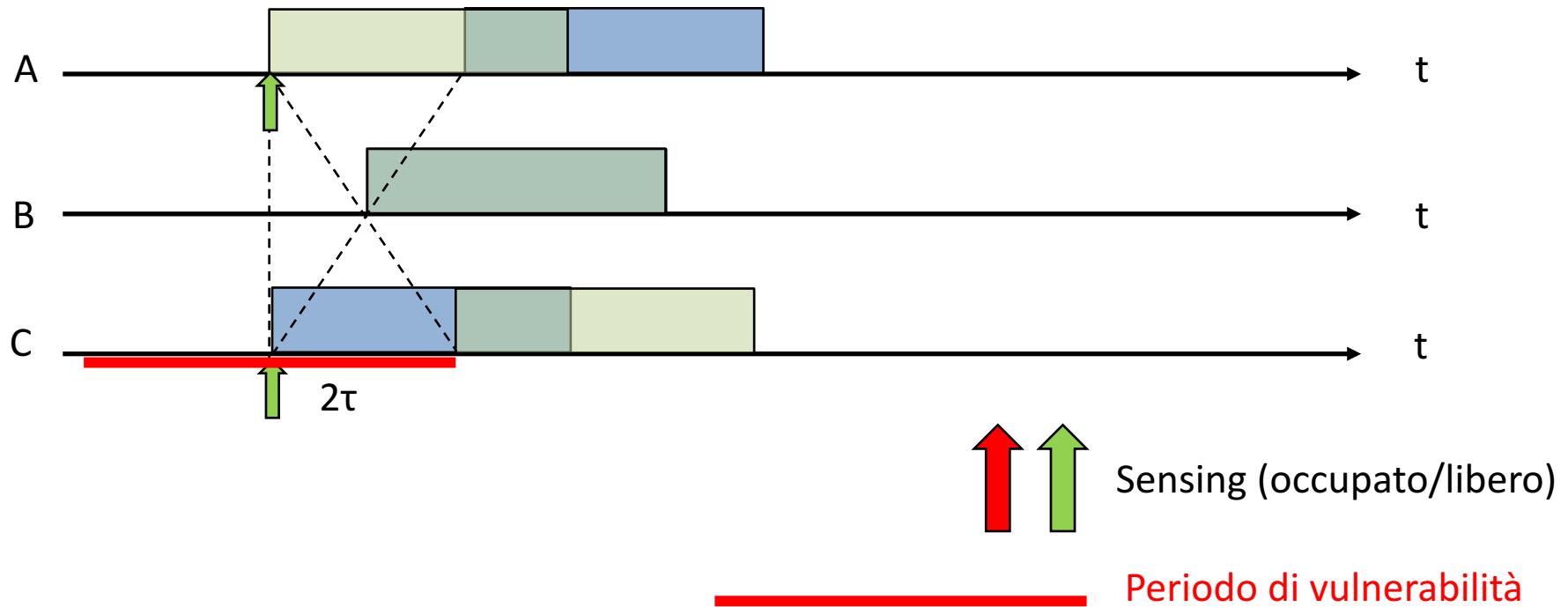
Carrier Sense Multiple Access (CSMA), $T > \tau$

- C trova libero \rightarrow trasmette
- A trova il canale libero \rightarrow trasmette \rightarrow collisione in B e A



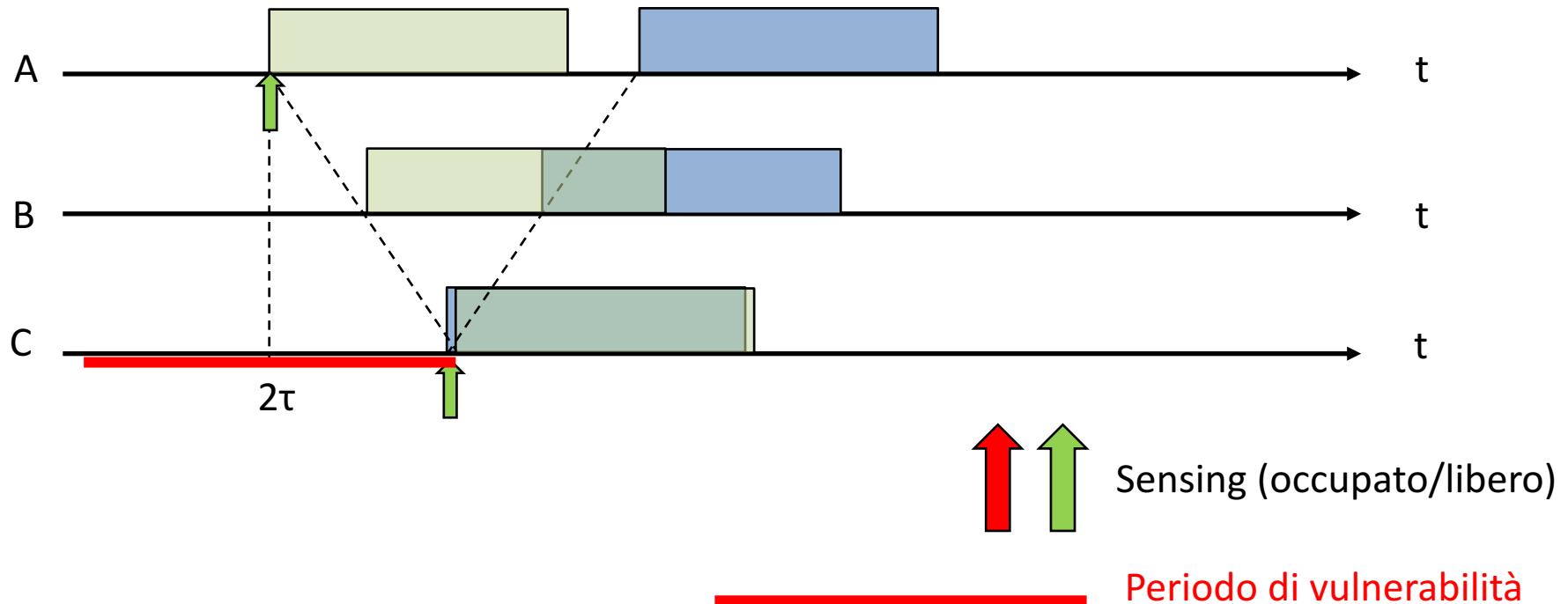
Carrier Sense Multiple Access (CSMA), $T > \tau$

- C trova libero \rightarrow trasmette
- A trova il canale libero \rightarrow trasmette \rightarrow collisione in C, A e B



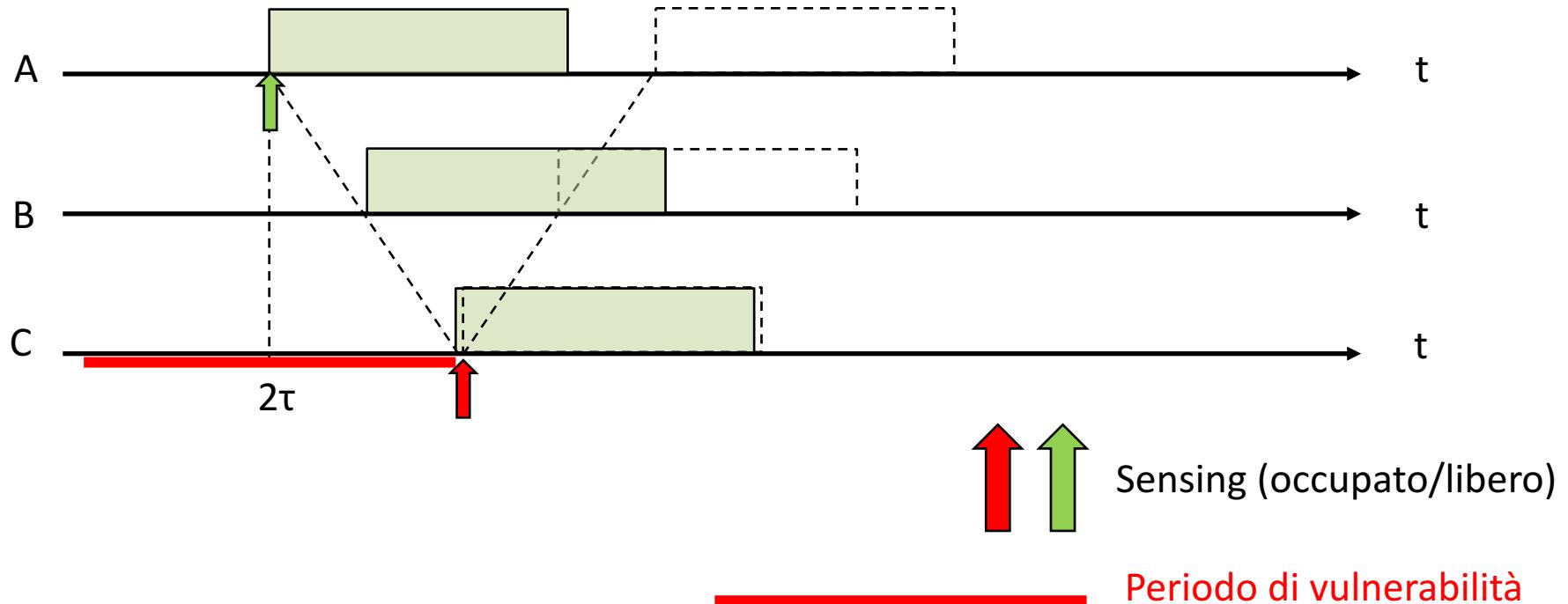
Carrier Sense Multiple Access (CSMA), $T > \tau$

- A trova libero \rightarrow trasmette
- C trova il canale libero \rightarrow trasmette \rightarrow collisione in C e B



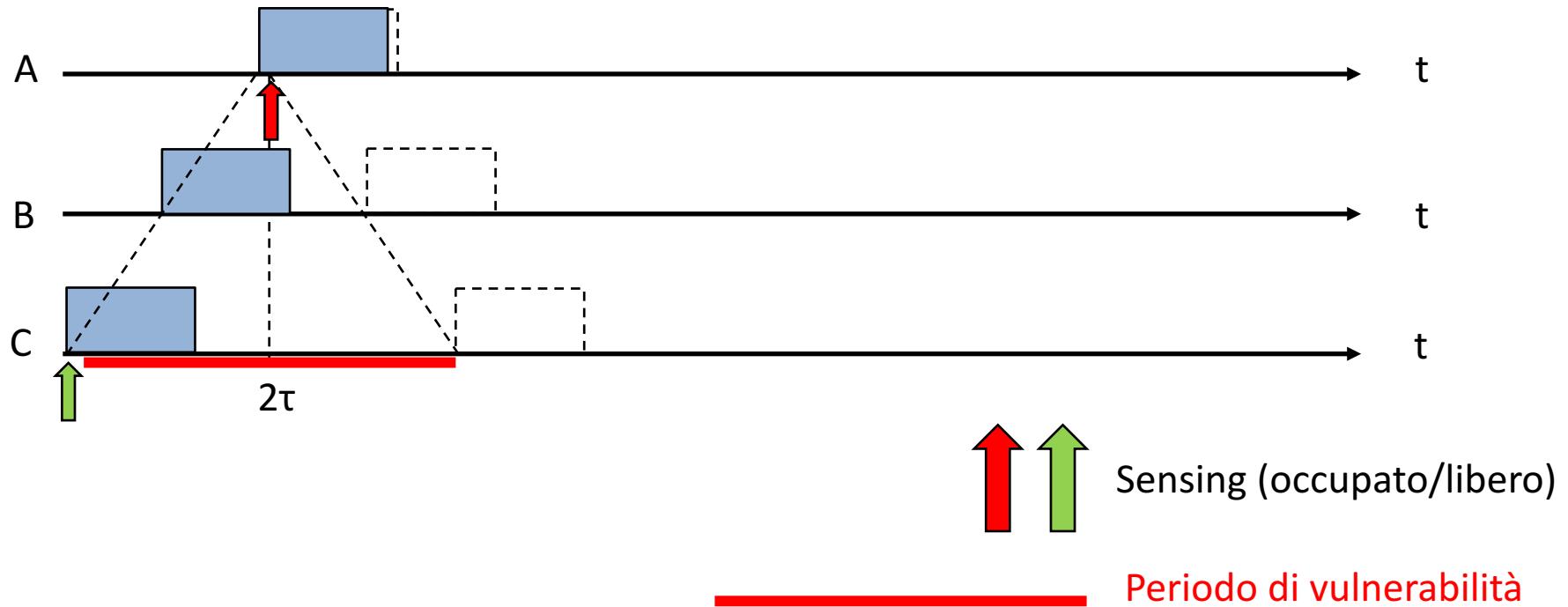
Carrier Sense Multiple Access (CSMA), $T > \tau$

- A trova libero \rightarrow trasmette
- C trova il canale occupato \rightarrow non trasmette \rightarrow nessuna collisione



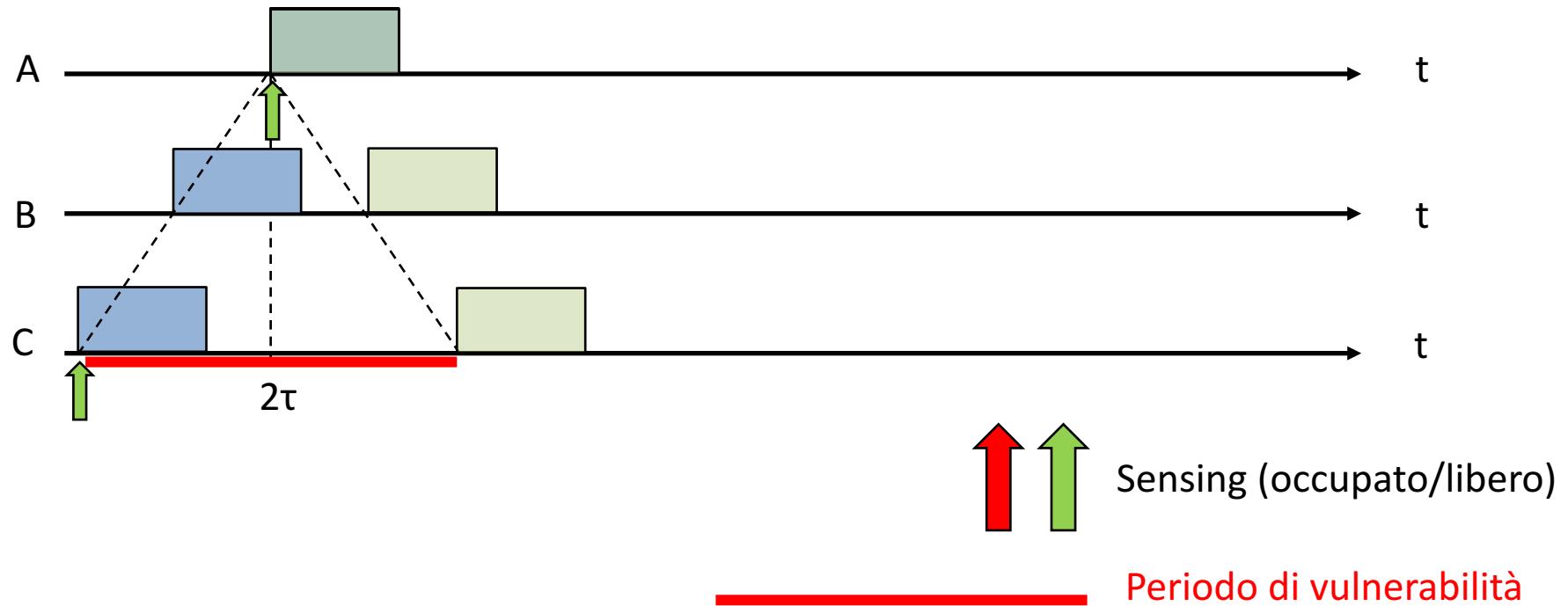
Carrier Sense Multiple Access (CSMA), $T < \tau$

- Collisione evitata



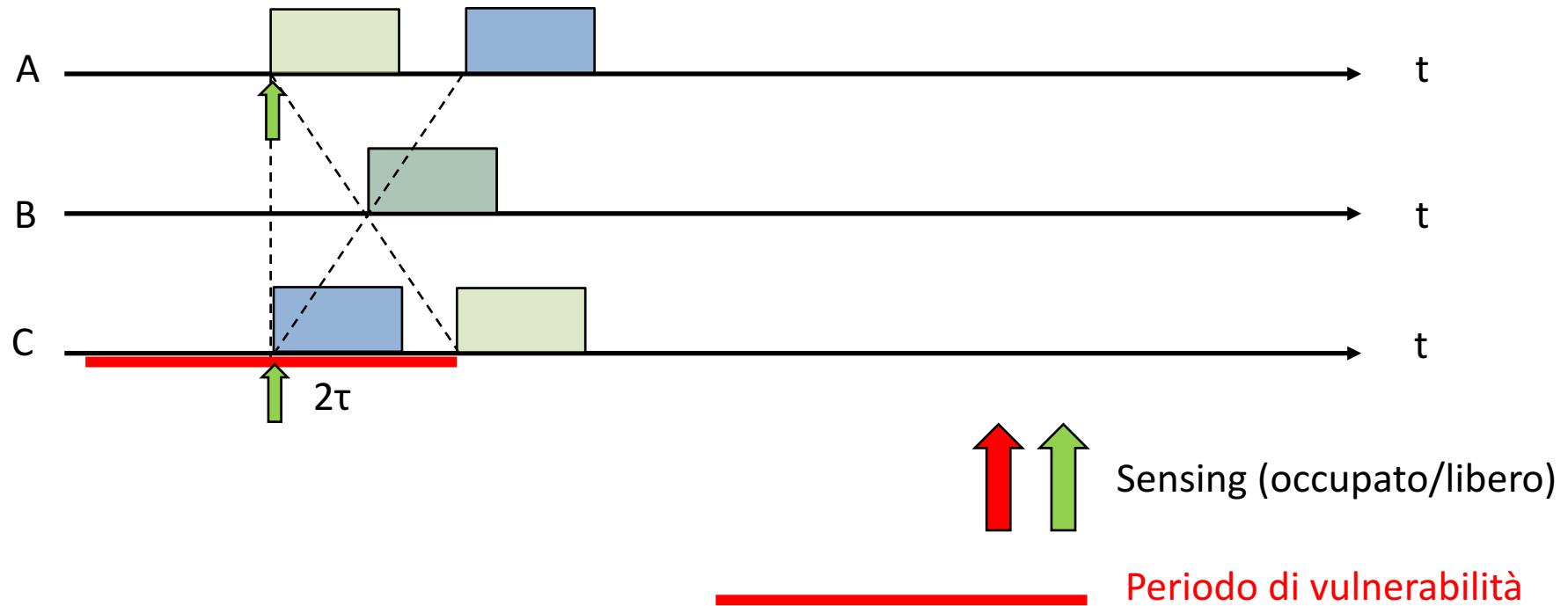
Carrier Sense Multiple Access (CSMA), $T < \tau$

- Collisione in A



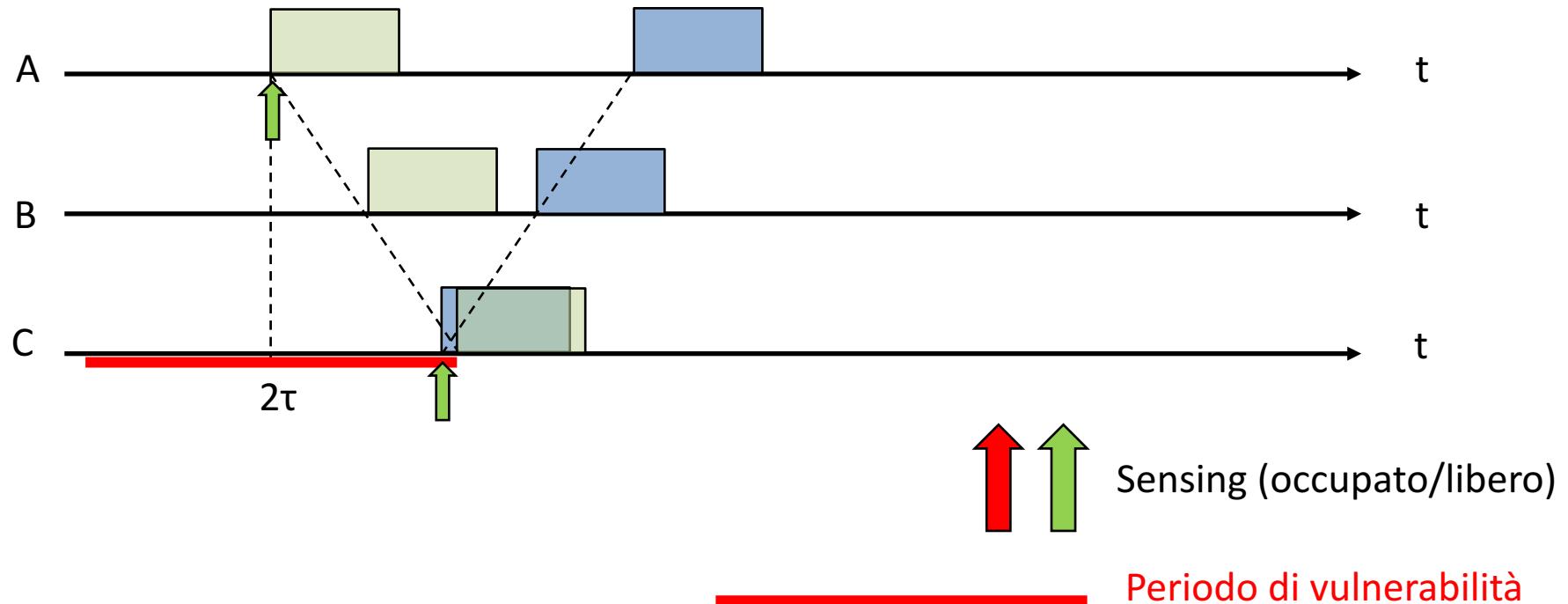
Carrier Sense Multiple Access (CSMA), $T < \tau$

- Collisione in B



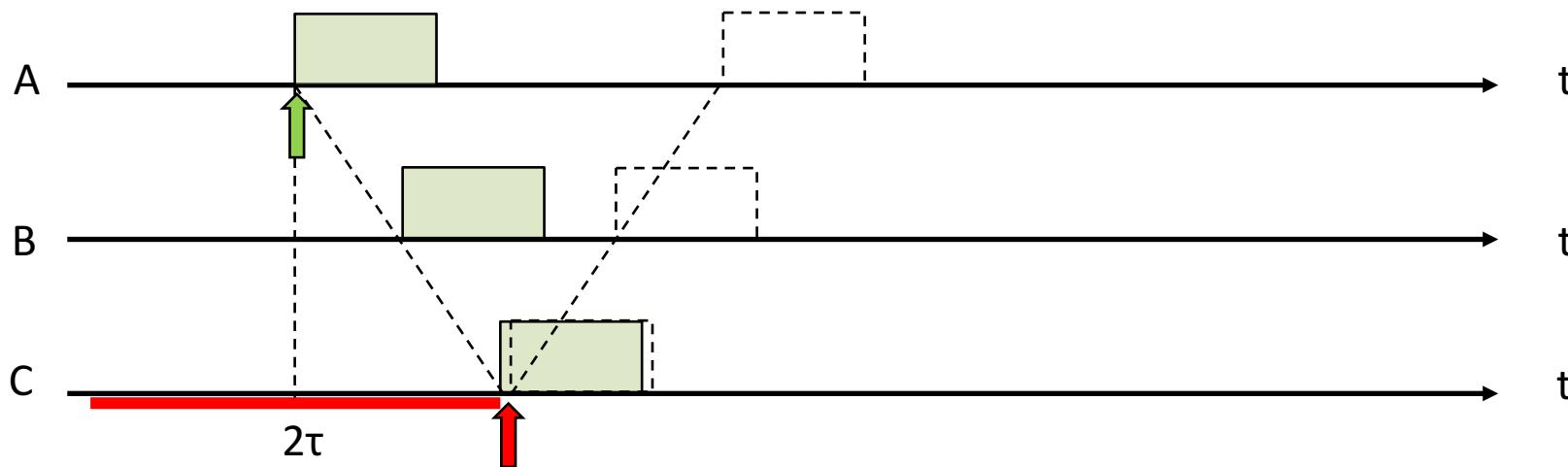
Carrier Sense Multiple Access (CSMA), $T < \tau$

- Collisione in C



Carrier Sense Multiple Access (CSMA), $T < \tau$

- Collisione evitata
- E' come prima, ma questa volta il periodo di vulnerabilità è molto più grande rispetto al tempo di trasmissione → la probabilità di evitare collisioni col CSMA è molto più piccola



Nota: il periodo di vulnerabilità in ALOHA dipende dal tempo di trasmissione, mentre in CSMA al ritardo di propagazione



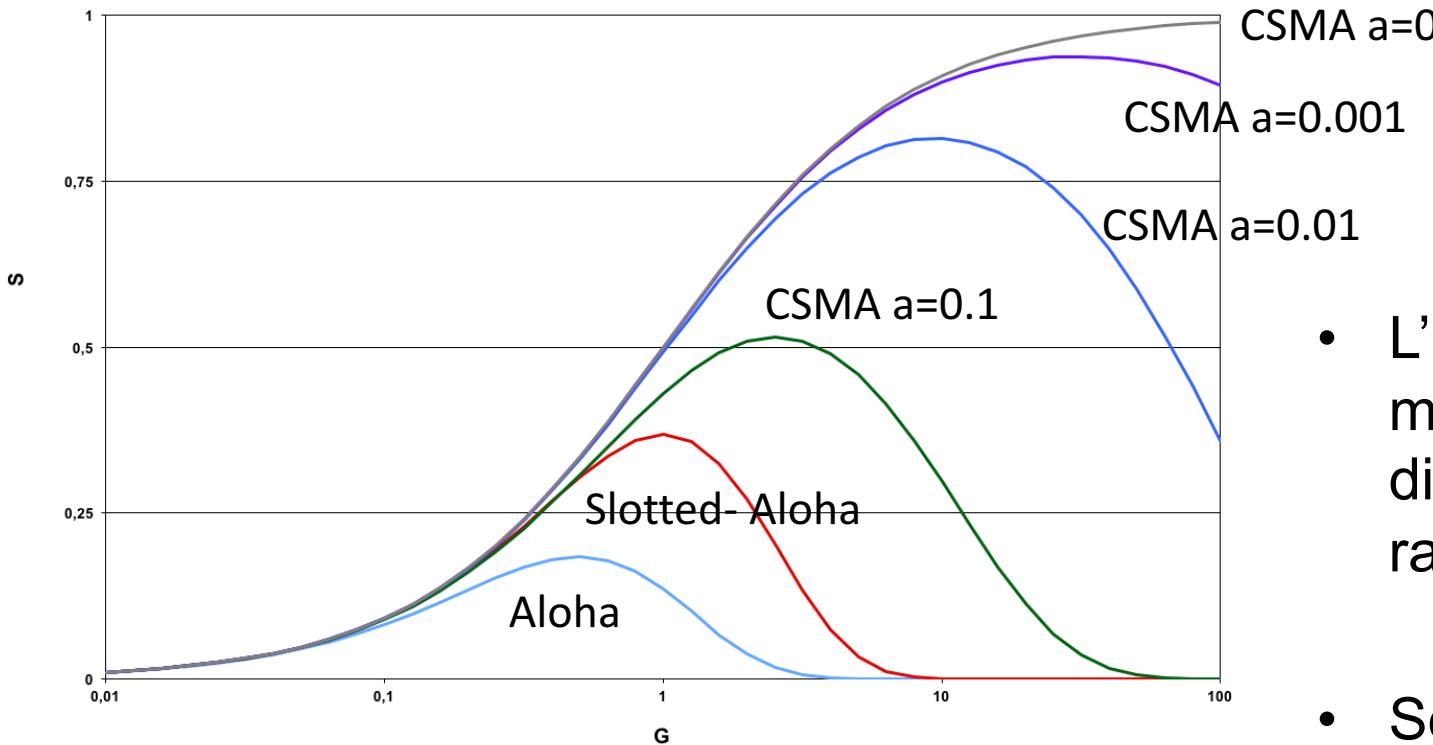
Sensing (occupato/libero)



Periodo di vulnerabilità



Carrier Sense Multiple Access (CSMA)



Formula throughput (no dimostrazione)

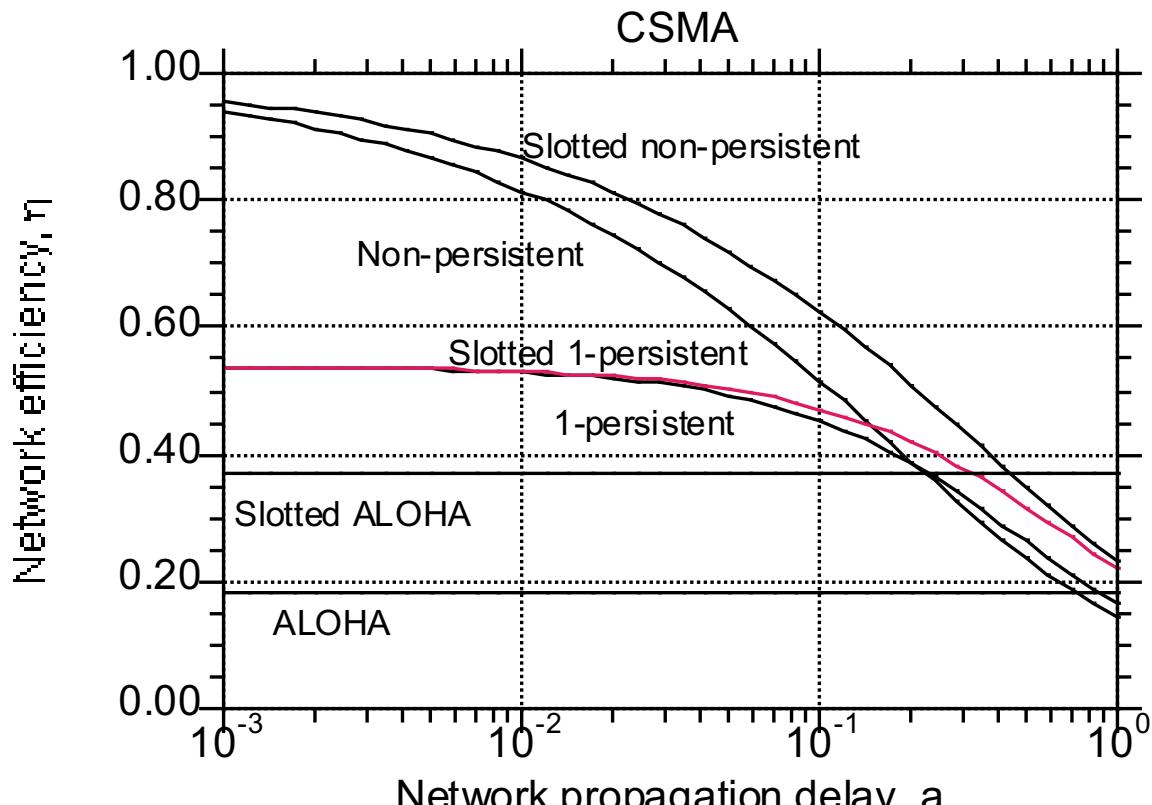
$$S = \frac{Ge^{-aG}}{G(1 + 2a) + e^{-aG}}$$

- L'efficienza del meccanismo dipende dal rapporto $a = \tau/T$
- Se $a \ll 1$, allora l'efficienza del CSMA può essere elevata



Carrier Sense Multiple Access (CSMA)

- Efficienza di rete (= valore del massimo throughput)



Nota: per gli ALOHA η è costante e pari ai valori 0.18 e 0.37 calcolati prima

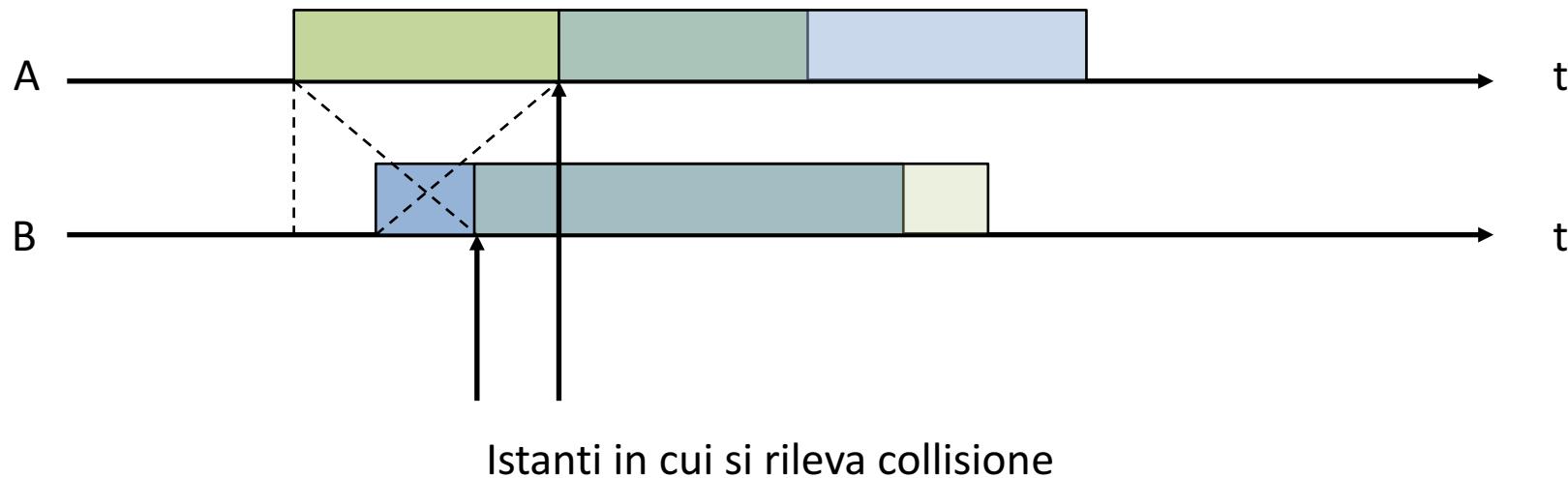
$$a = \tau/T$$



CSMA – Collision Detect (CSMA-CD)

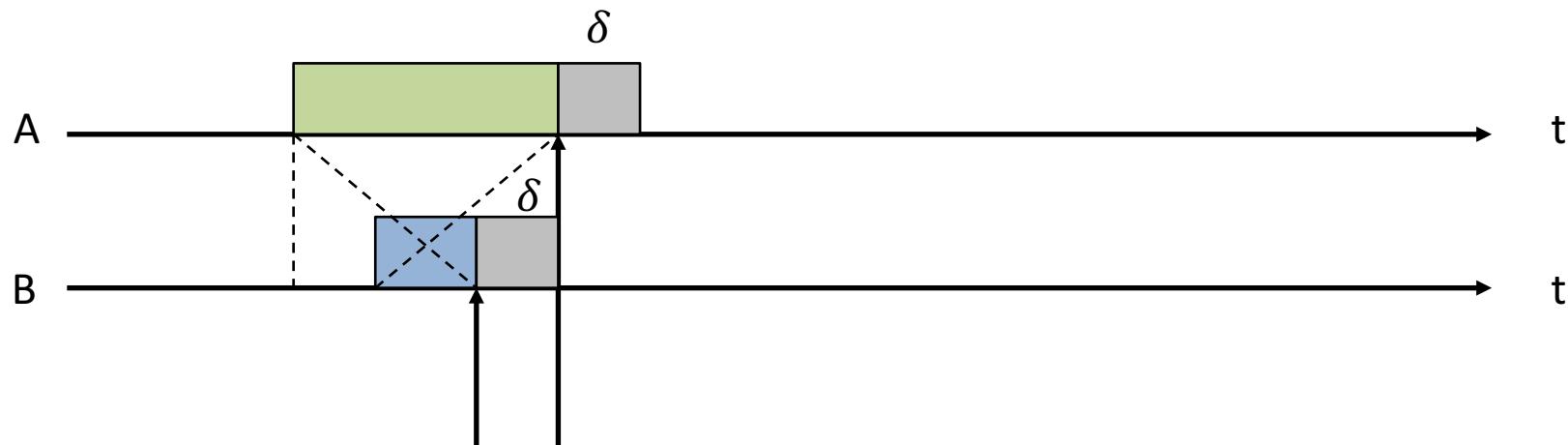
[Metcalfe 1976]

- Anche le stazioni che trasmettono possono accorgersi della collisione (“ascoltare mentre si parla”, per niente facile!)
- Quando se ne accorgono possono interrompere la trasmissione per risparmiare tempo



CSMA – Collision Detect (CSMA-CD)

- Dopo la rivelazione si può attendere un piccolo intervallo di tempo δ e poi interrompere

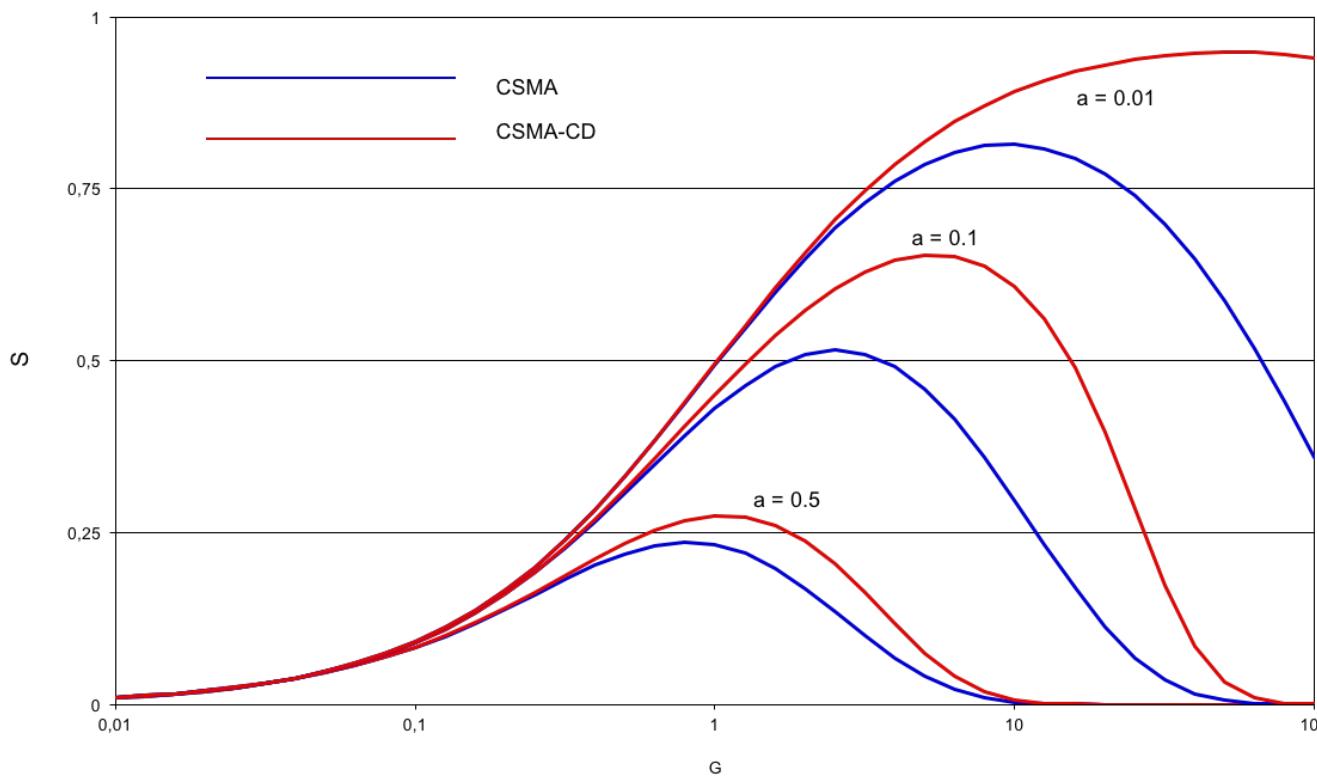


Istanti in cui si rileva collisione

Sistema usato da
Ethernet



CSMA – Collision Detect (CSMA-CD)



Formula throughput (no dimostrazione)

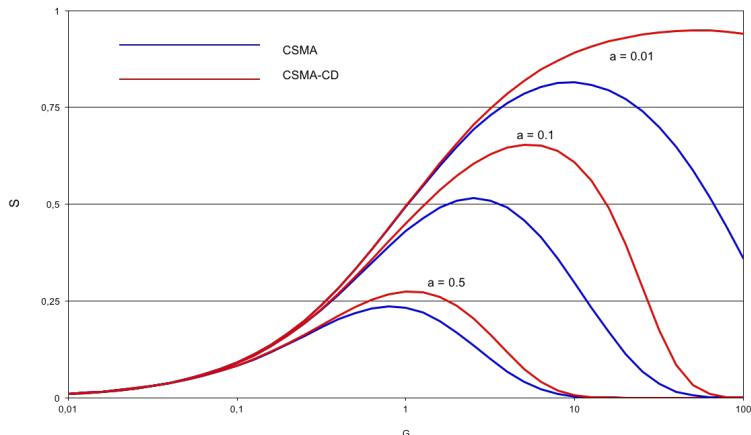
$$S = \frac{Ge^{-aG}}{G(1+2a) + e^{-aG} - G(1-\delta)(1-e^{-aG})}$$

Formula massimo approssimata
(no dimostrazione)

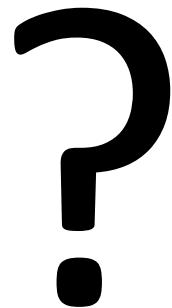
$$S_{\max} = \frac{1}{1+5a}$$



CSMA – Collision Detect (CSMA-CD)



Ma a che serve il
Collision Detect se
migliora le prestazioni
solo di poco



D: Come fa una stazione a rilevare una collisione con Aloha e CSMA?

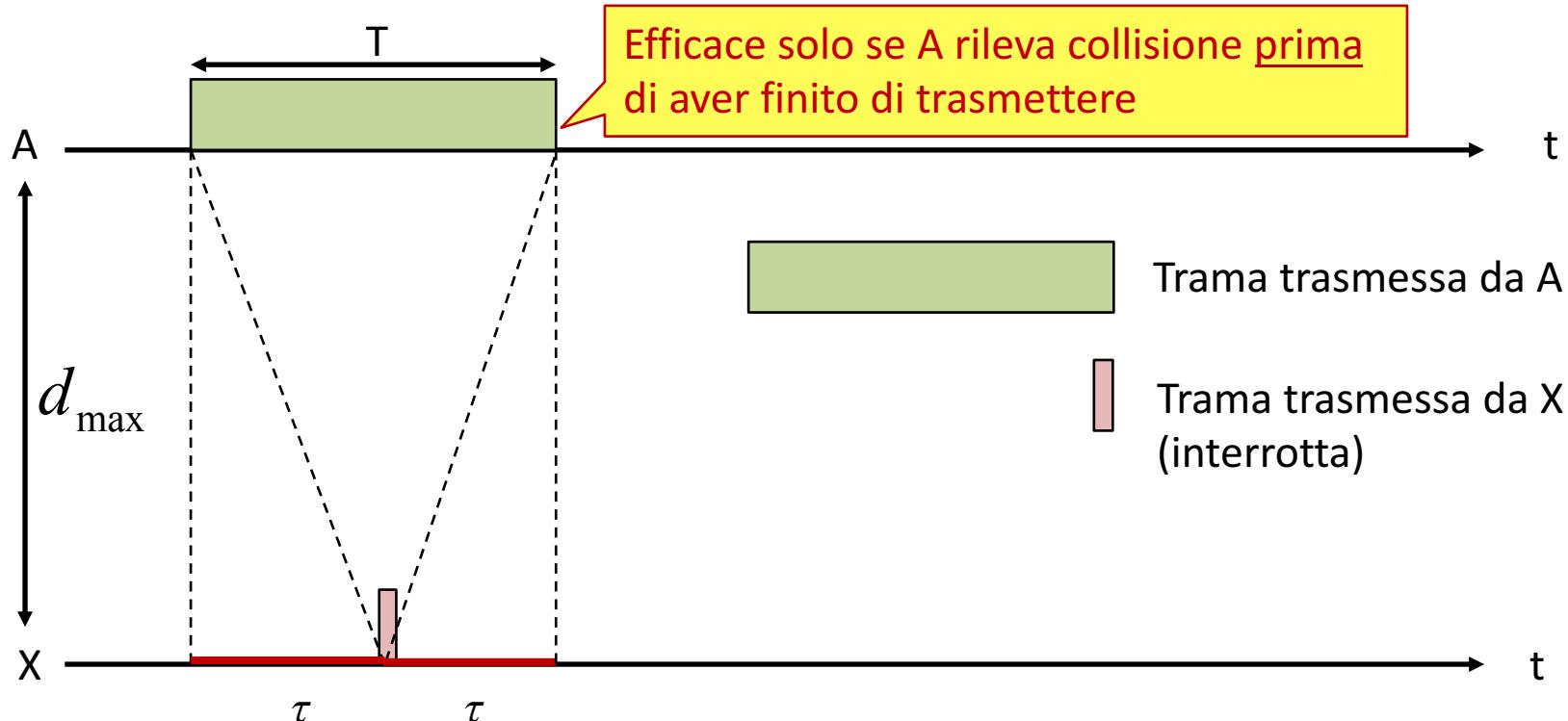
R: Il solito meccanismo di ACK è necessario: se c'è collisione, il ricevitore non manda ACK e il timeout scade

Con il Collision Detect il trasmettitore può rivelare lui stesso la collisione → si elimina la necessità di un riscontro (ACK) dal ricevitore



CSMA – Collision Detect (CSMA-CD)

- Relazione tra parametri che consente la funzione “CD”



$$T \geq 2\tau$$

$$\frac{L_{\min}}{C} \geq 2 \frac{d_{\max}}{v} \Rightarrow L_{\min} \geq 2 \frac{d_{\max} C}{v}$$

L_{\min} Lungh. minima trama [bit]

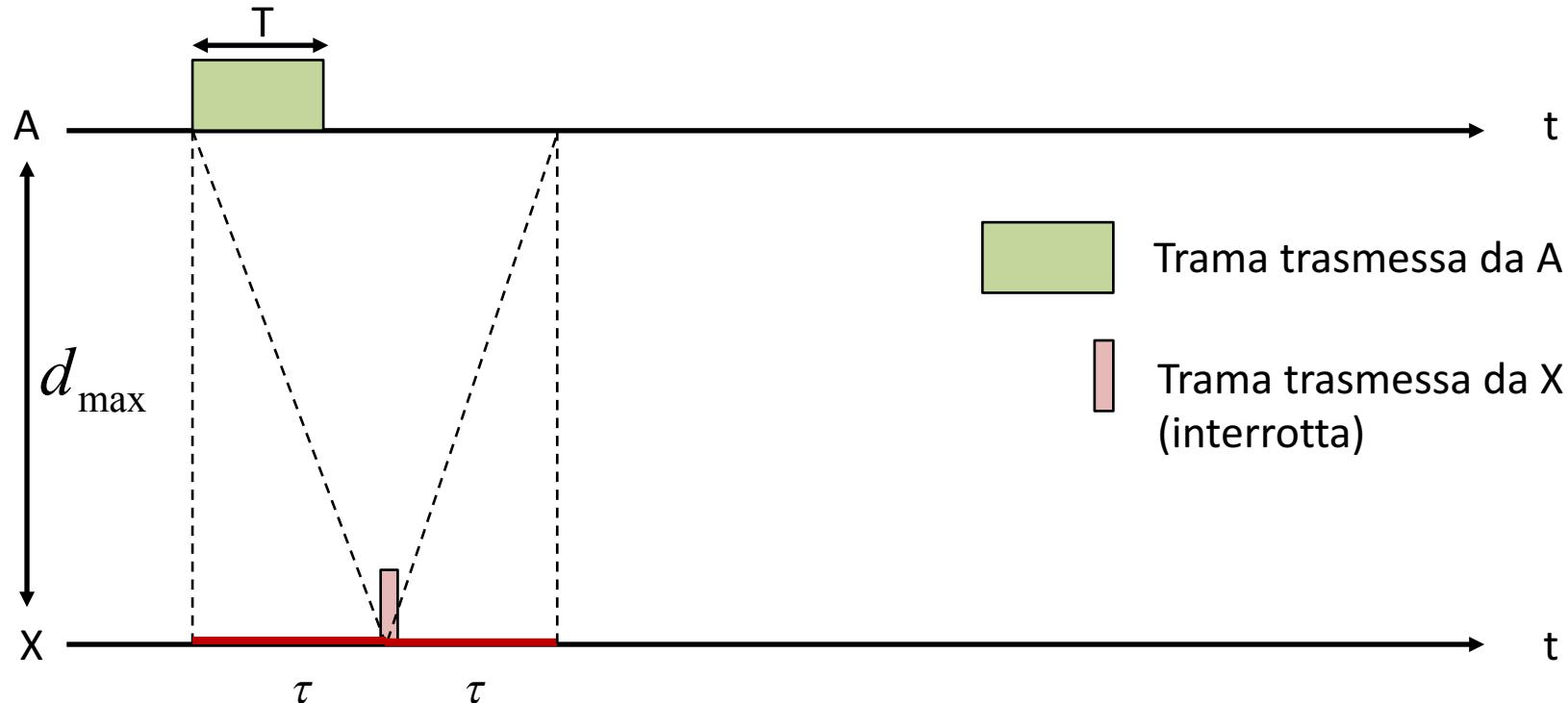
C Capacità del mezzo [bit/s]

v Vel. Prop. nel mezzo [m/s]



CSMA – Collision Detect (CSMA-CD)

- Dimensionamento trama non corretto



$$L_{\min} < 2 \frac{d_{\max} C}{v}$$

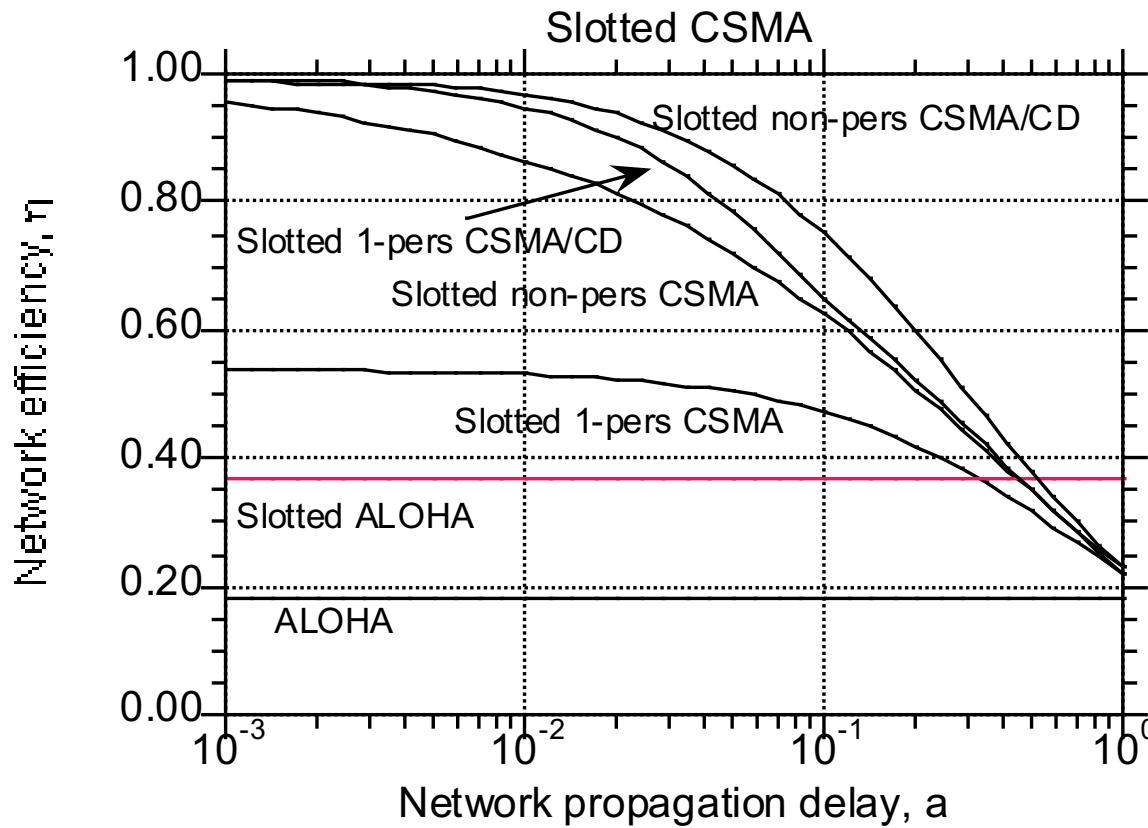


“A” non rileverà alcuna
collisione sulla propria trama



CSMA – Collision Detect (CSMA-CD)

- Efficienza di rete (= valore del massimo throughput)

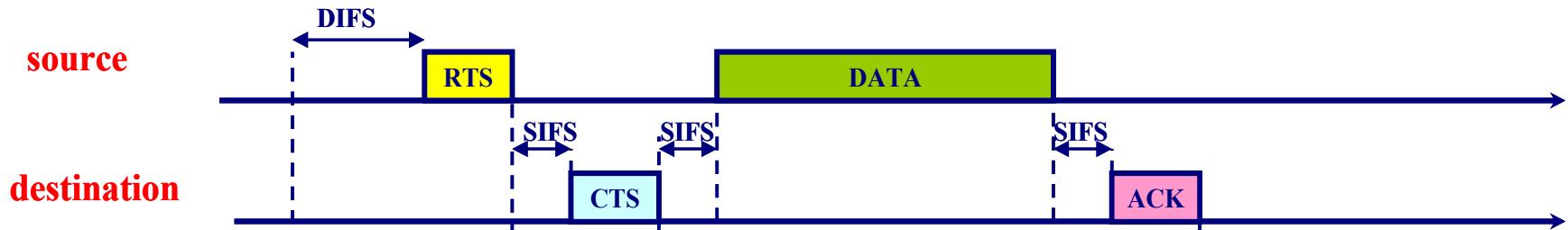


Collision Detect

- Il **Collision Detect** si basa sul fatto che nelle reti locali cablate come Ethernet l'attenuazione del segnale è piccola ed il livello di segnale ricevuto dalle altre stazioni è simile al proprio: ci si può dunque accorgere se c'è più di una trasmissione (collisione)
- Nelle reti radio (wireless) il mezzo attenua molto e quindi non è più possibile usare il Collision Detect
- E' come cercare di ascoltare i grilli quando qualcuno ti urla nelle orecchie



CSMA Collision Avoidance (CSMA-CA)



- **RTS:** Request to Send
- **CTS:** Clear to Send
- La collisione può avvenire solo su RTS
- Se si riceve il CTS si prosegue con la trama dati
- C'è anche l'ACK. Come mai?



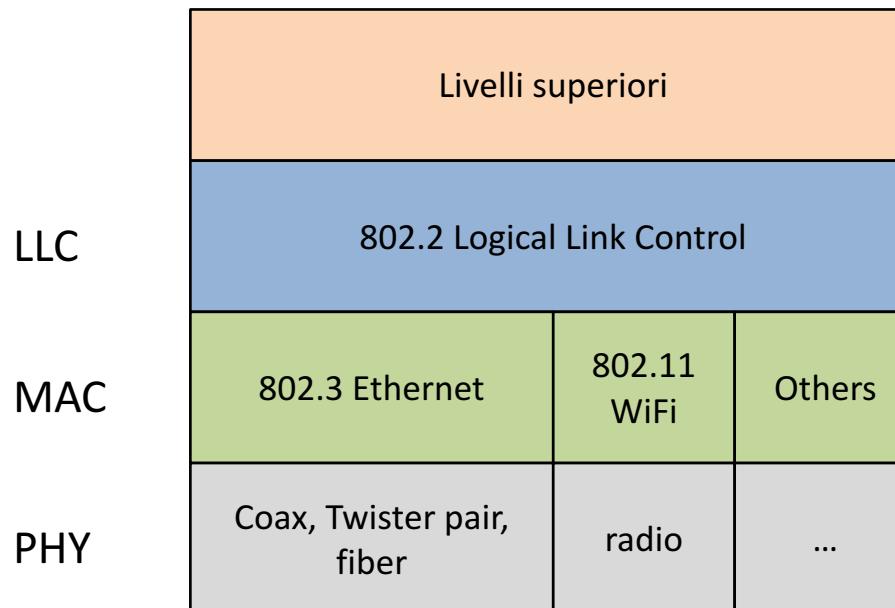


5c – Tecnologie di reti locali

Ethernet, WiFi

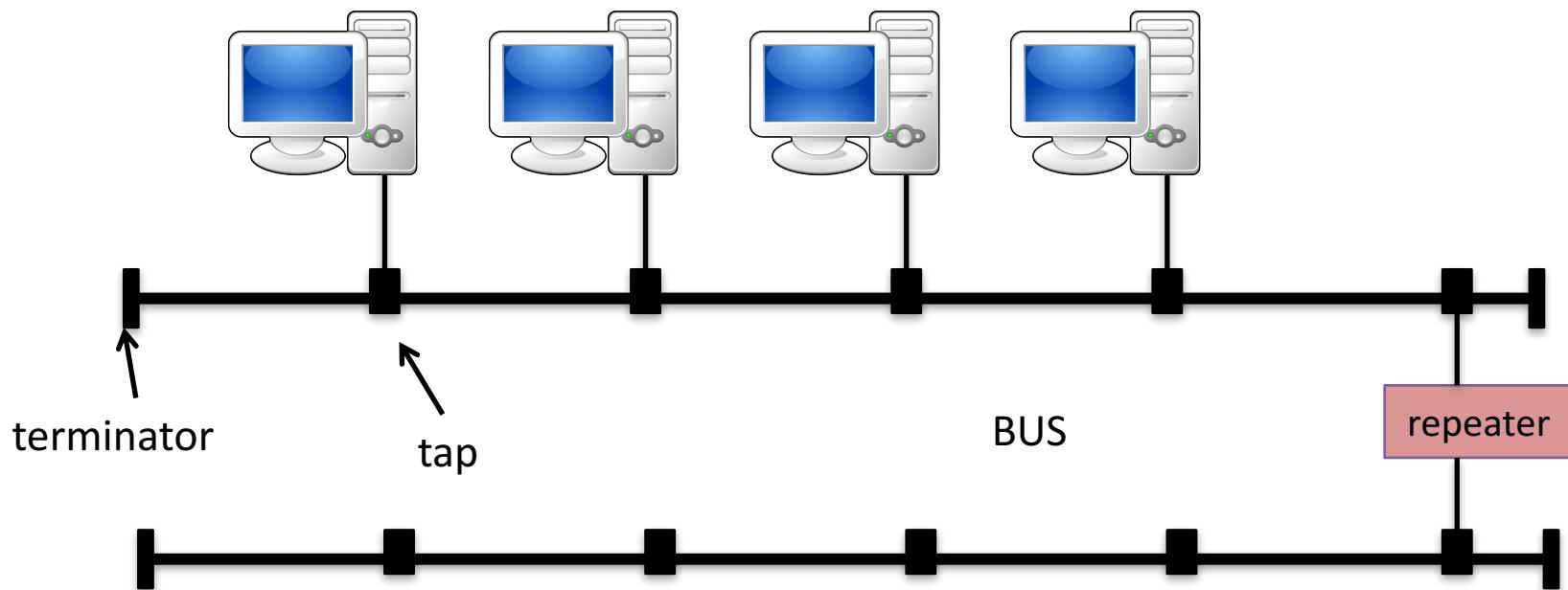
Tecnologie di reti locali (LAN)

- IEEE (Institute for Electrical and Electronics Engineers) è il principale organismo di standardizzazione delle tecnologie per reti locali con il suo **802 Working Group** project
- Differenti tecnologie sono standardizzate da IEEE 802: i livelli LLC e superiori sono in comune, MAC e Livello Fisico sono diversi



Ethernet

- **Ethernet** è stato progettato da Xerox (1976) e poi standardizzato da **IEEE 802.3 WG**
- Il mezzo fisico inizialmente adottato era un cavo coassiale passivo (BUS) a cui si connettevano le stazioni mediante un transceiver



Ethernet – cavi cossiali

- **Ethernet 10Base5**

- Cable RG-213
 - (Thick Ethernet)



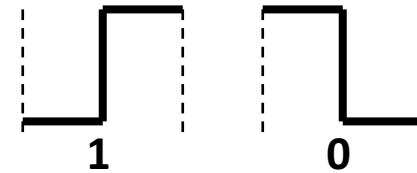
- **Ethernet 10Base2**

- Cable RG-58
 - (Thin Ethernet)



- **XBaseY**

- X: bit rate in Mb/s
 - Base: trasmissione in banda base (codifica Manchester)
 - Y: massima lunghezza (in centinaia di metri)



tap

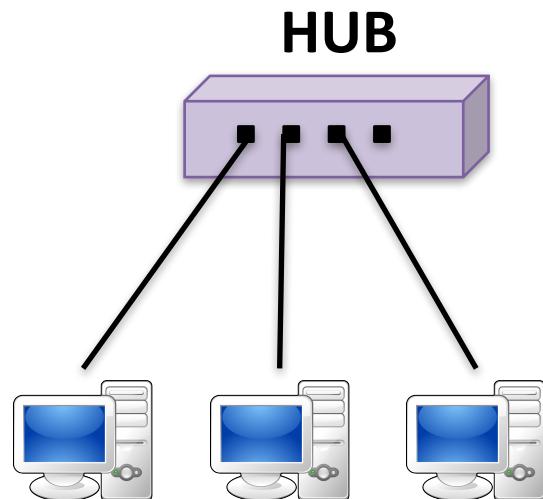


terminator



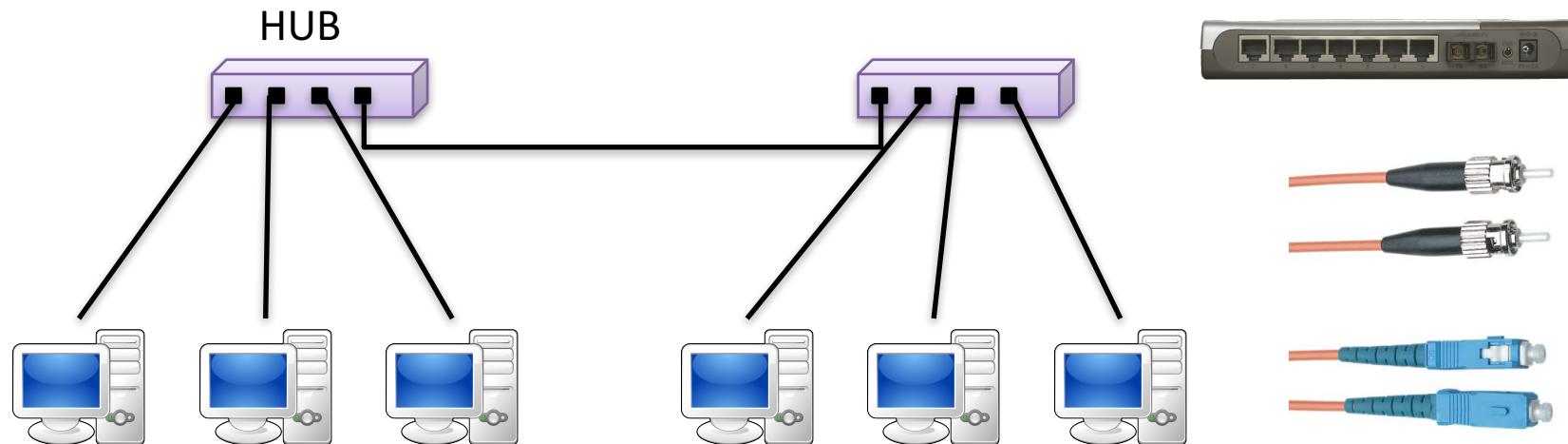
Ethernet – topologie a stella

- A partire da metà anni '90, i BUS con cavo coassiale sono stati sostituiti da topologie a stella
- Le topologie a stella sono erano basate su ripetitore di segnale a livello fisico multi-porta, denominati **HUB**
- Il mezzo trasmissivo è rimpiazzato da **doppini in rame (twisted pairs)** (Ethernet 10BaseT)



Fast Ethernet

- A fine anni 90', il rate di trasmissione viene aumentato da 10 Mb/s a 100 Mb/s con **Fast Ethernet**
- In aggiunta ai twisted pair (100BaseTX), si iniziano a usare le **fibre ottiche** (100BaseFX)



Con Fast Ethernet anche l'apparato di interconnessione cambia radicalmente da hub a SWITCH (→ vedi dopo)



Gigabit Ethernet

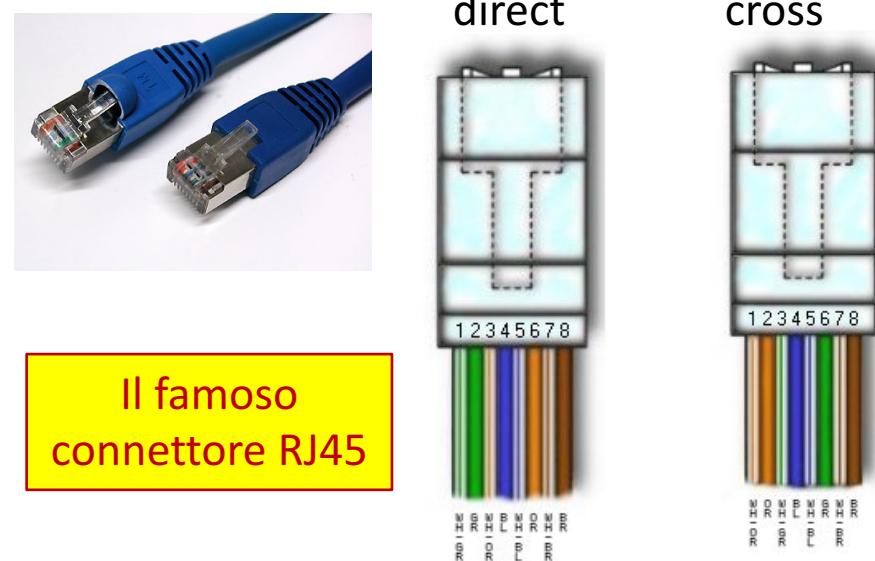
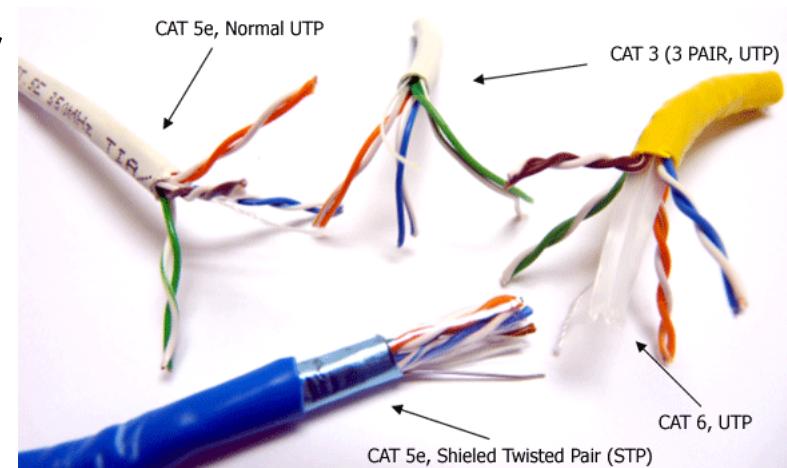
- Nei primi anni 00', il rate di trasmissione viene aumentato da 100 Mb/s a 1 Gb/s con **Gigabit Ethernet** e poi a 10 Gb/s a metà anni 00' con **10 Gigabit Ethernet**
- Molti mezzi trasmissivi sono utilizzati: ancora doppini di diversi tipi, fibre ottiche multimodali e monomodali
- 40 e 100 Gbit/s sono già standard
 - Principalmente in fibra ottica, o in rame solo per cortissimo raggio (es. patch cord nei datacenter)
- Oggi si va verso la standardizzazione del Tbit/s (400 G già commerciale; 1 T quasi)



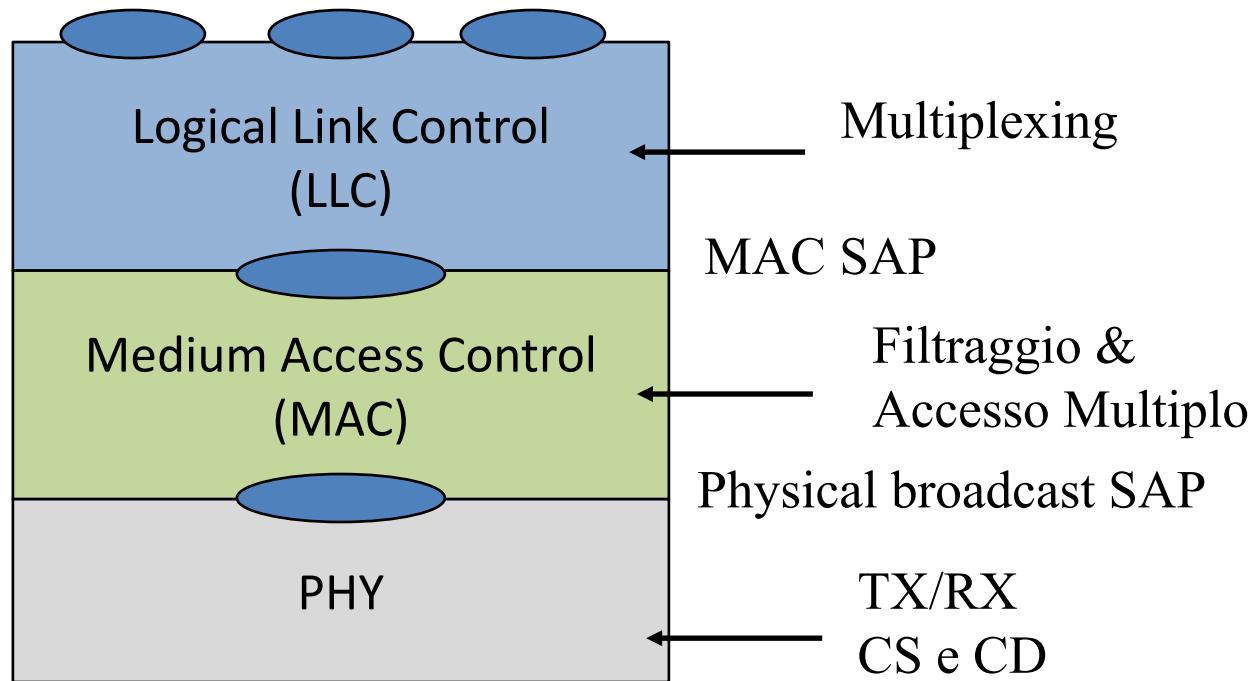
Tipi di doppini in rame

UPT – Unshielded Twisted Pair

- **Cat. 1:** analog telephony line
- **Cat. 2:** digital telephony line
- **Cat. 3:** $B=16$ MHz
 - Ethernet 10BaseT
- **Cat. 4:** $B=20$ MHz
 - Token Ring 16 Mb/s
- **Cat. 5:** $B=100$ MHz
 - Ethernet 100Base-TX
- **Cat. 5E:** $B=100$ MHz
 - Gigabit Ethernet 1000Base-TX
- **Cat. 6:** $B=250$ MHz
- **Cat. 7:** $B=600$ MHz

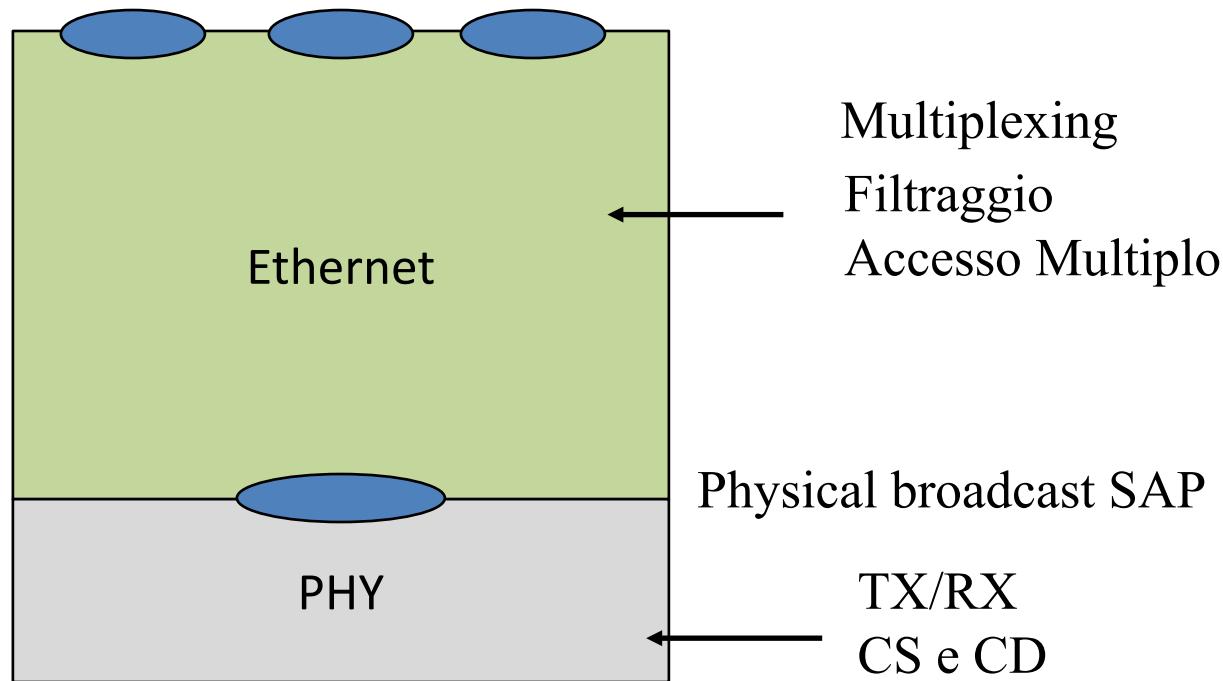


Protocolli di rete locale (IEEE 802) protocol stack



Protocolli di rete locale (IEEE 802) protocol stack

- Esiste una versione dello stack per Ethernet che mette insieme LLC e MAC in un unico “livello ethernet”



Trame Ethernet

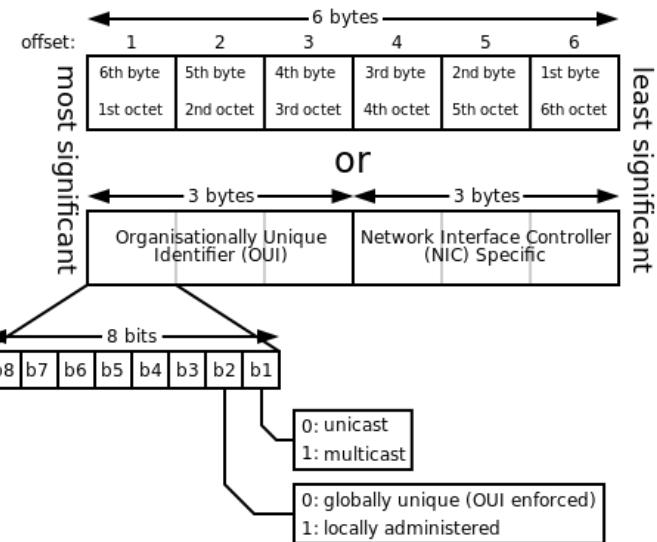
7	6	6	2	46-1500	4	bytes
Sync	Destinaz.	Sorgente	Type	Dati	FCS	

- **Sync:** Preamble di sincronizzazione di livello fisico
- **Indirizzi:** (Destinazione e sorgente): indirizzi di 48 bit definiti dal costruttore nella scheda di rete (NIC)
- **Type:** serve per la multiplazione di più livelli superiori (ad esempio IP ha il suo Type=0800)
- **Dati:** Campo dati per PDU proveniente dal livello superiore
- **FCS:** Frame Check Sequence per il controllo d'errore
- Le trame hanno una LUNGHEZZA MINIMA (oltre che massima):
 - $(46 + 18) \text{ byte} = 64 \text{ byte}$ [il preamble Sync non conta]
 - Questo garantisce il funzionamento del CD nel CSMA/CD fino al Fast Ethernet (500 m @ 100 Mbit/s)
 - Gli Ethernet a bit-rate maggiori non vengono mai utilizzati con mezzi broadcast (c'è la possibilità di farlo con GigaEth., ma ormai è in disuso)
 - Notare che la lunghezza max. corrisponde al MTU = 1500 byte



Indirizzi

- Gli indirizzi di rete locale sono detti **indirizzi MAC** o **indirizzi fisici**
- Servono per la funzione di filtraggio
- Primi 3 byte identificano il costruttore
- Ultimi 3 byte identificano la scheda
- 48 bit di solito indicati con notazione esadecimale (HEX)
- L'indirizzo con 48 bit a “1” è l'indirizzo broadcast (tutte le stazioni ricevono e processano la trama)



48-bit MAC address					
00	0C	42	28	79	45
00000000	00001100	01000010	00101000	01111001	01000101

broadcast

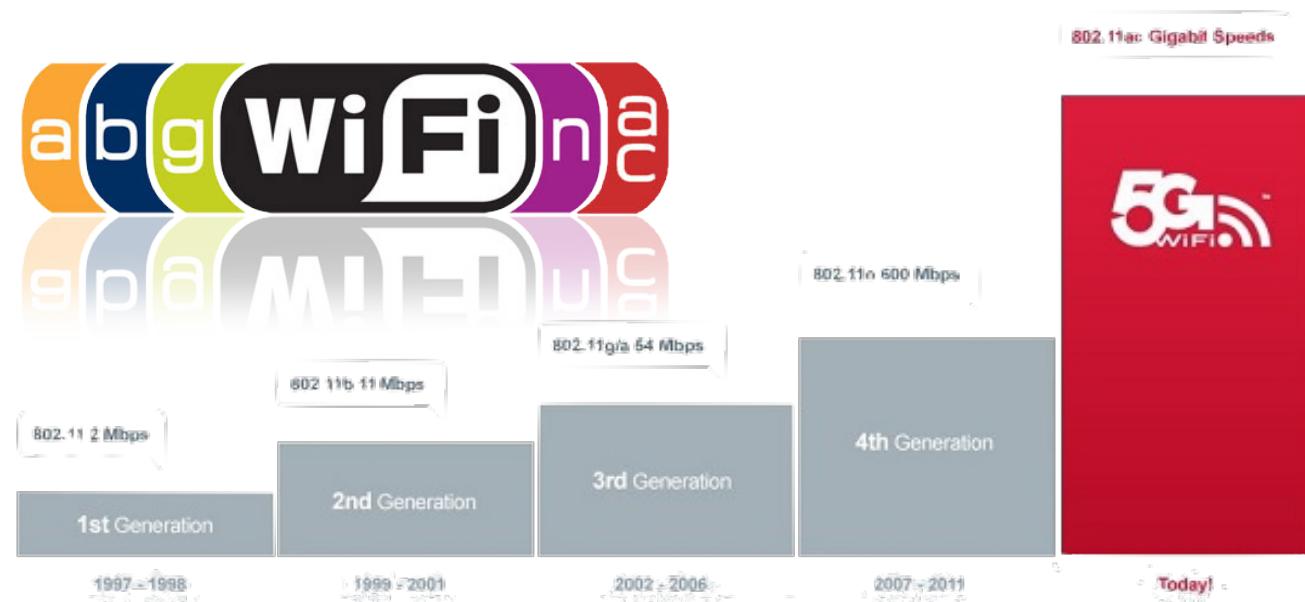
FF:FF:FF:FF:FF:FF

In origine il MAC addr. era "cablato" nella NIC, oggi si può riconfigurare



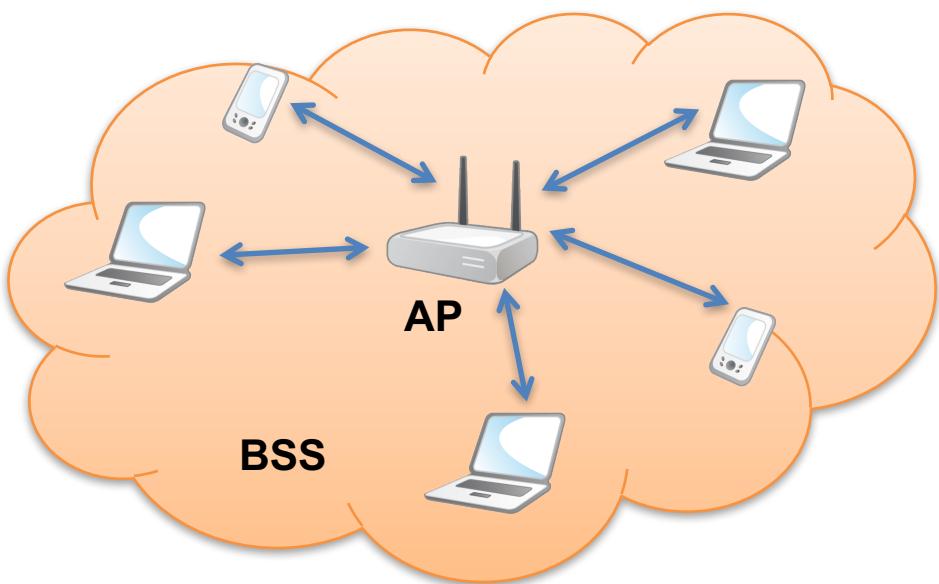
WiFi

- La tecnologia WiFi è standardizzata dal gruppo di lavoro IEEE 802.11
- Rappresenta la versione wireless di Ethernet ed è largamente usata
- Esistono molte versioni di livello fisico che operano a velocità e bande di frequenze diverse

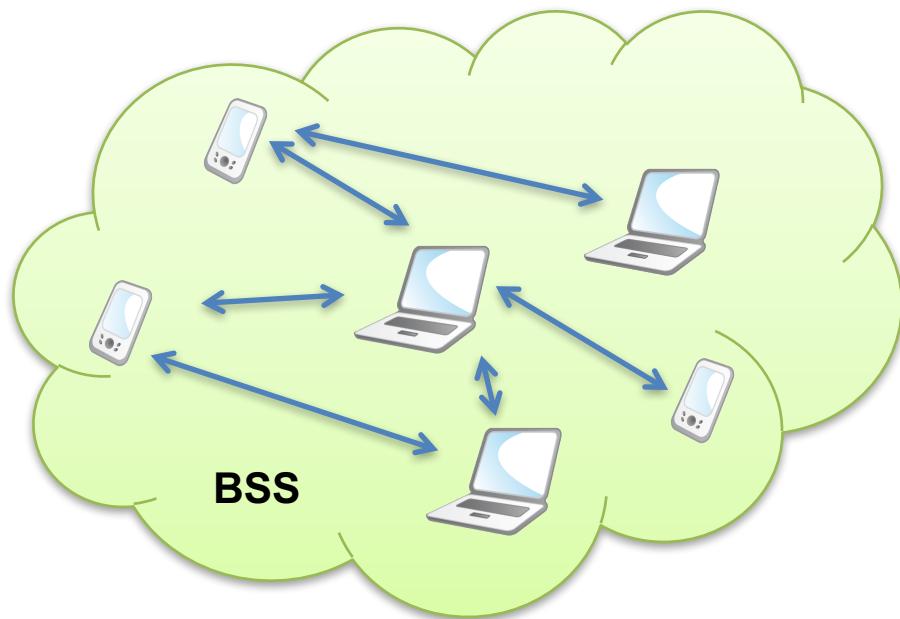


WiFi: modalità AP e Ad hoc

- **BSS:** Basic Service Set
- **AP:** Access Point



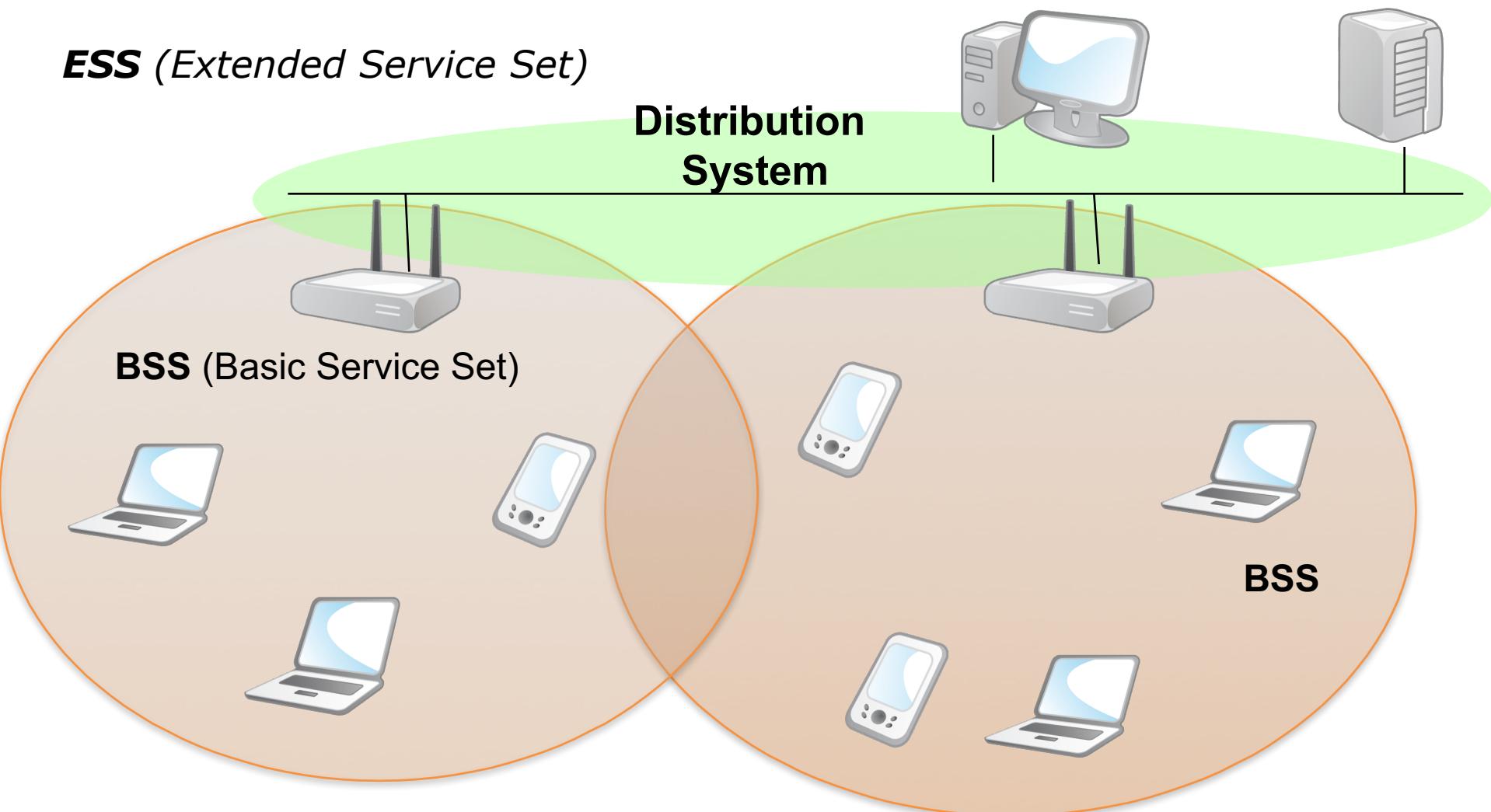
Centralized interconnection



Distributed ad hoc mode



WiFi: architettura di rete



Frame format: *Frame Control Field*

Bytes:

2	2	6	6	6	2	6	0-2312	4
Frame Control	Duration ID	Addr 1	Addr 2	Addr 3	Sequence Control	Addr 4	Frame Body	CRC
802.11 MAC Header								

- Stesso formato degli indirizzi, ma fino a 4 indirizzi per trama
 - Varie modalità di trasmissione, da wireless a wired, da wired a wireless, da wireless a wireless (ad-hoc), wireless distribution system
 - Necessità di capire quale AP ha trasmesso / è destinata la trasmissione radio
- Numero di sequenza necessario perché ogni trama viene riscontrata (non c'è CSMA/CD)
- E' indicata la durata temporale (μs) delle trasmissione della trama (usato per aiutare meccanismo anti-collisioni)
- Frame Control: versione protocollo, tipo di trama, gestione energetica dei dispositivi, frammentazione, etc.



Wireshark

- **Guardiamo insieme le trame Ethernet e WiFi con Wireshark**
 - Analisi indirizzi
 - Differenze tra 802.3 e “livello Ethernet”
- (...)



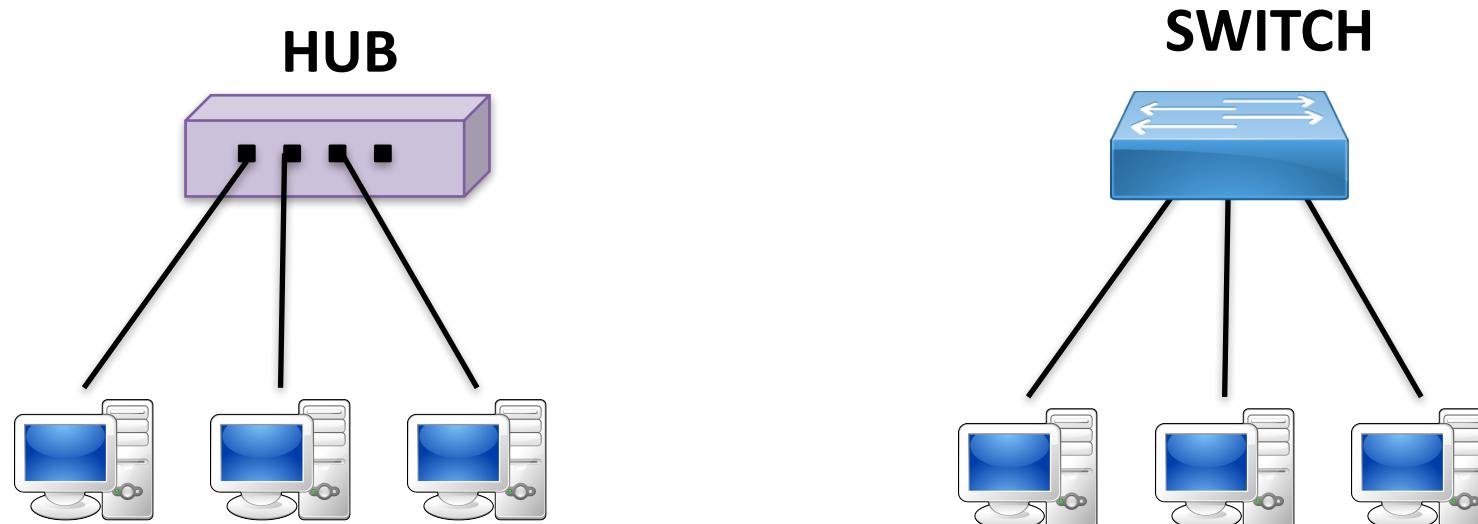


5d – Collegamenti commutati

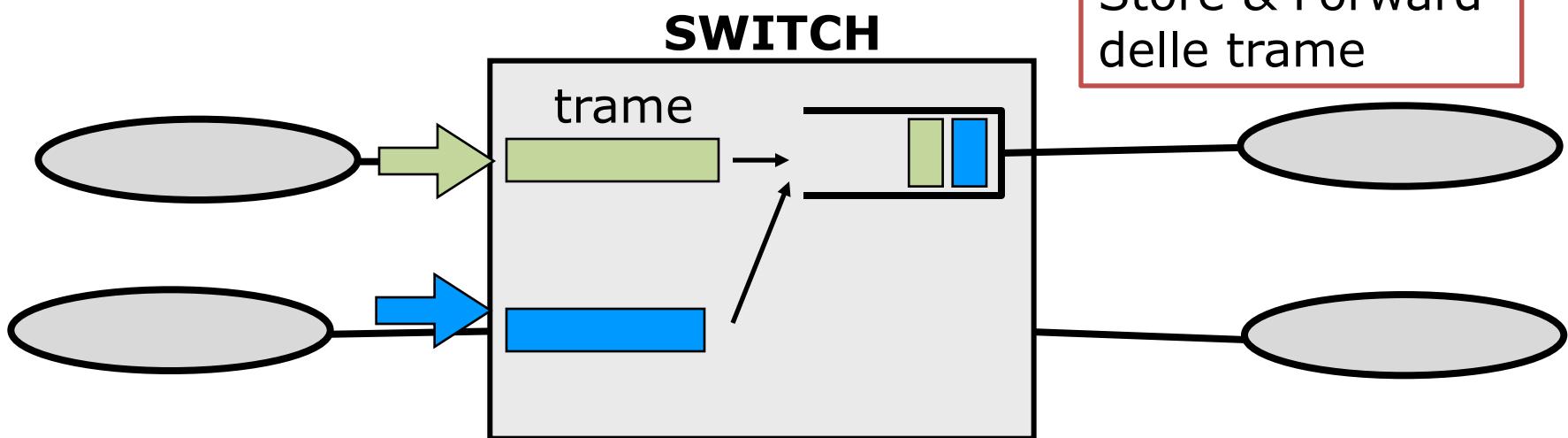
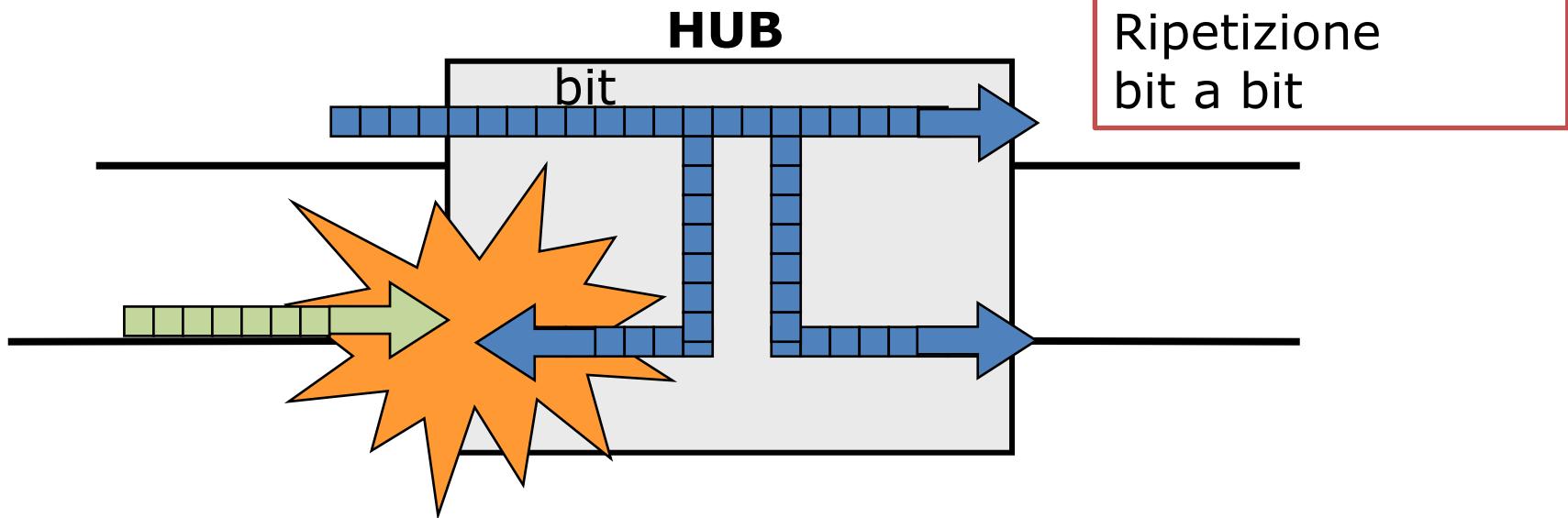
Switching, Spanning Tree, VLAN

LAN commutate (switched)

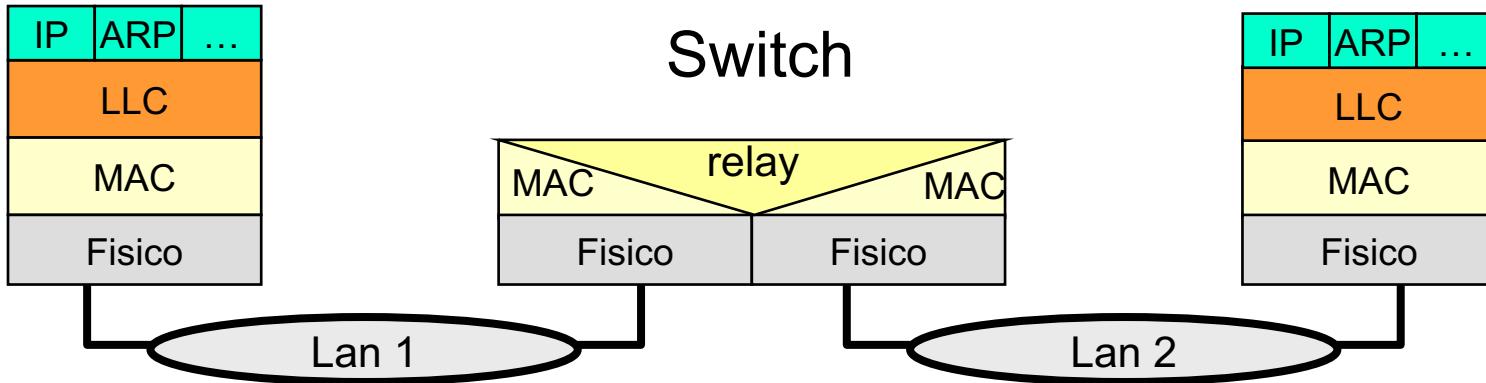
- Finora abbiamo visto LAN con livello di linea broadcast
- Nel caso di Ethernet tuttavia abbiamo una possibilità in più, costituita dalla LAN commutate o switched, oggi largamente usate
- Conosciamo già la commutazione di pacchetto, per implementarla in Ethernet occorre sostituire l'HUB con un dispositivo detto SWITCH (o bridge)



Confronto HUB e SWITCH



Switch

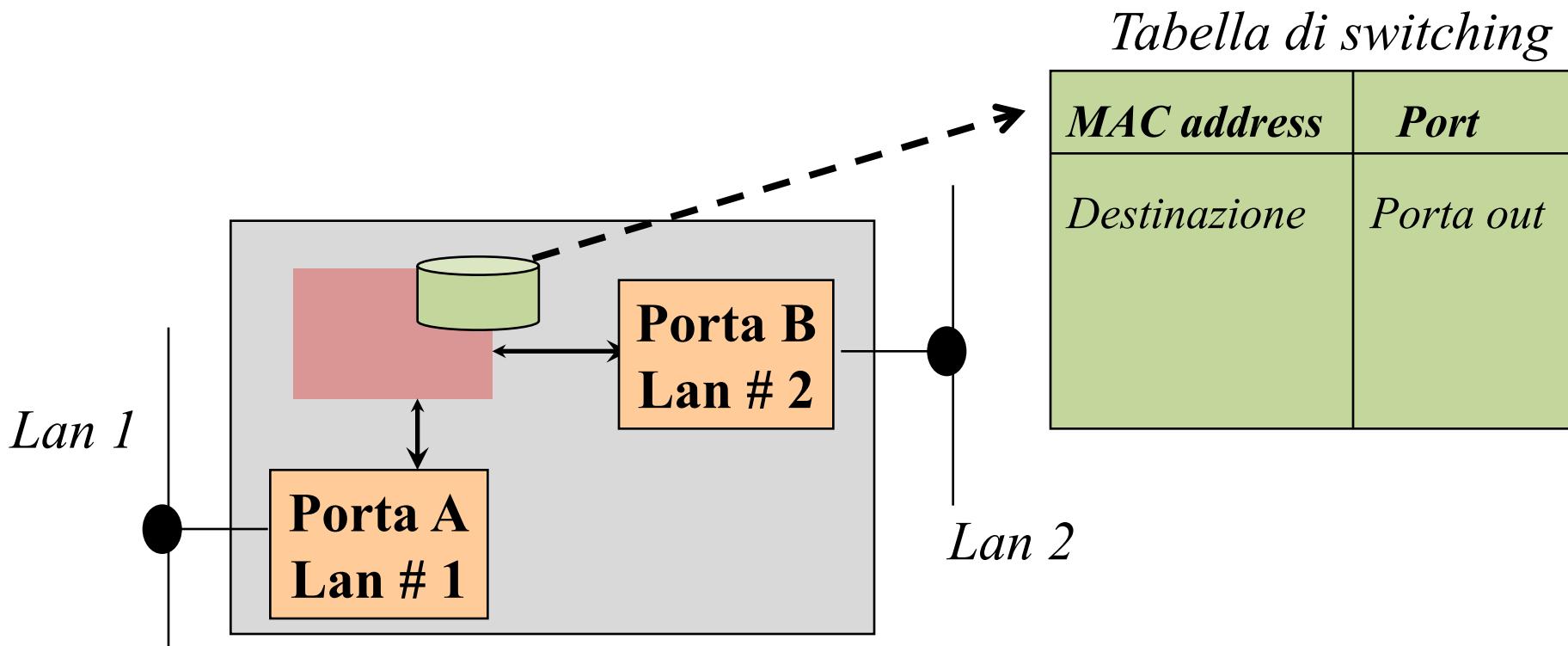


- **Funzioni dello Switch**
 - Filtering: se una trama ricevuta da Lan 1 è indirizzata ad una stazione di Lan 1, viene scartata
 - Relay: se una trama ricevuta da Lan 1 è indirizzata ad una stazione di Lan 2, viene trasmessa su Lan 2



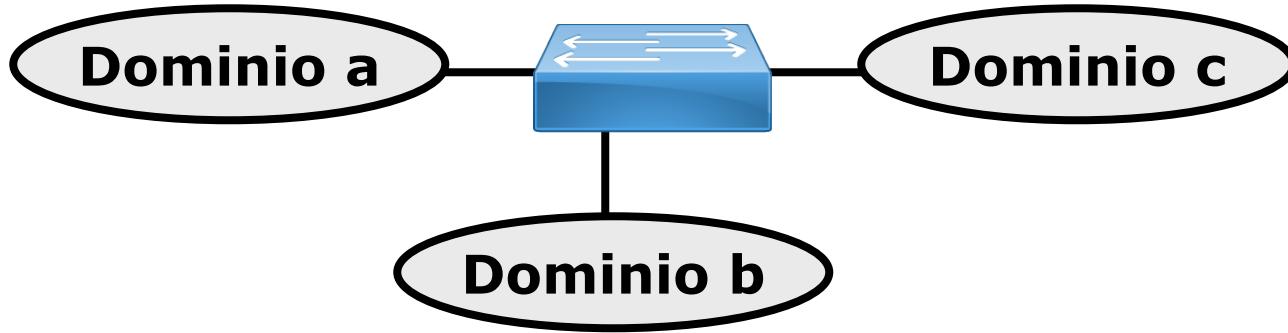
Switch

- Per stabilire se filtrare/instradare una trama si consulta una tabella di instradamento locale chiamata forwarding data base (o FDB)



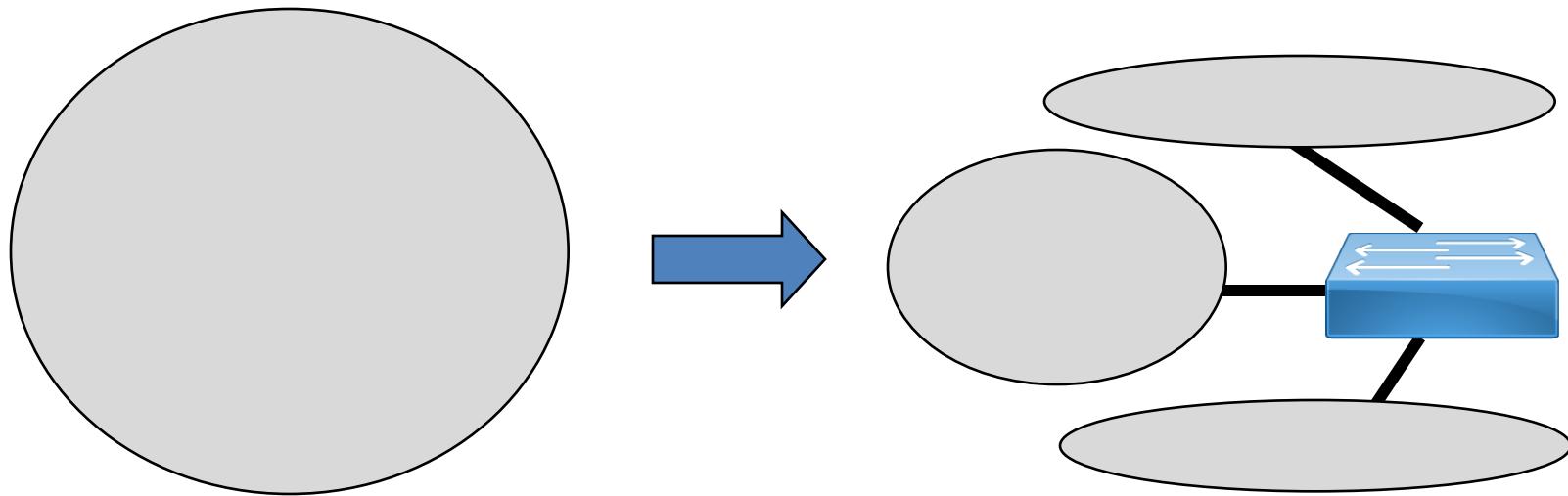
Switch

- A ciascuna porta di uno switch è collegato un **dominio** (rete) che può contenere una rete broadcast tradizionale o anche solo una stazione (LAN completamente commutata)
- **Inoltro:** lo switch assicura che le trame provenienti da ciascun dominio siano inoltrate al dominio di destinazione
- **Funzione broadcast:** le trame con destinazione indirizzo broadcast sono inoltrate su tutti i domini tranne quello di origine



Efficienza della segmentazione in domini

- Lo switch segmenta domini di accesso multiplo (detti anche **domini di collisione**) mantenendo la rete come un unico dominio broadcast (grazie alla funzione broadcast)
- La segmentazione consente di aumentare l'efficienza



Efficienza della segmentazione in domini

- Assumendo che il traffico massimo (efficienza) S_M smaltibile in un dominio di accesso multiplo sia indipendente dalla dimensione e dal numero di utenti (es. $S_M=0.98$)
- α è la frazione di traffico che esce da ogni dominio ed è uniformemente diretta verso gli altri domini,
- X è il massimo traffico (efficienza) smaltibile dalla rete complessiva
- Allora

$$S_M = \underline{X(1 - \alpha)} + X\alpha + X\alpha = X(1 + \alpha)$$

traffico interno traffico uscente traffico entrante



Efficienza della segmentazione in domini

- Il nuovo traffico smaltibile da un dominio diventa dunque

$$X = \frac{S_M}{1 + \alpha}$$

- E, se la partizione è in N domini, si ha

$$X_N = \frac{NS_M}{1 + \alpha}$$



NS_M per $\alpha = 0$

$NS_M/2$ per $\alpha = 1$



Transparent Bridging

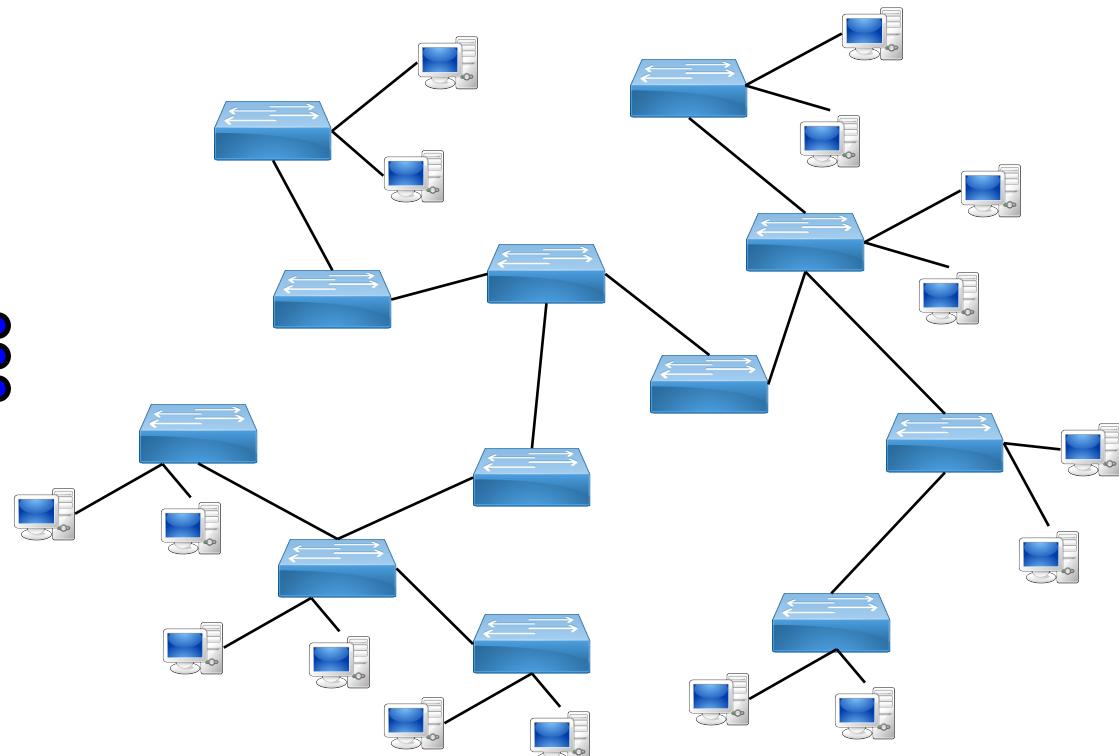
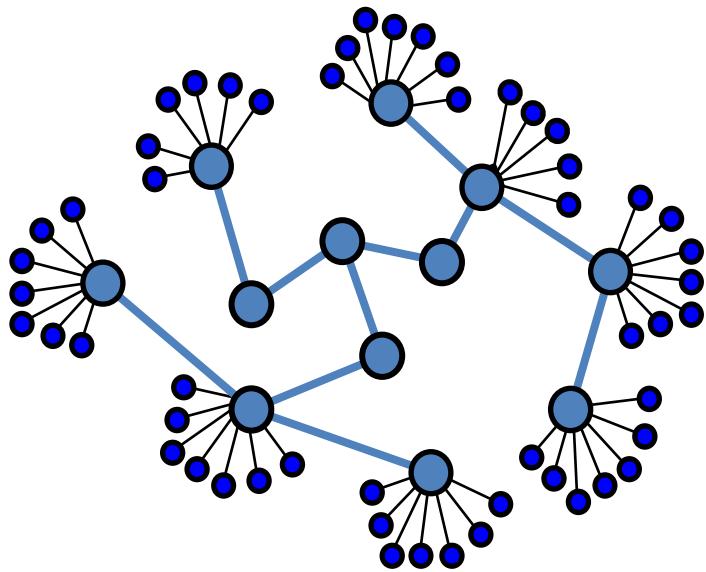
- Ma come si compila la tabella di switching?
- La presenza dello switch nella rete è completamente **trasparente** alle stazioni che continuano ad implementare le funzionalità del livello di linea come prima
- Lo switch non ha indirizzo MAC (o meglio non compare nelle trame)
- Le tabella di switching sono **compilate automaticamente**

Tabella di switching	
MAC address 1	Porta A
MAC address 2	Porta B
MAC address 3	Porta C



Transparent Bridging

- Come può uno switch **imparare automaticamente** da quale porta si raggiunge una stazione?
- **Facilmente se la rete è ad albero!**
- C'è un solo percorso tra ogni coppia di nodi



Transparent Bridging : autoapprendimento

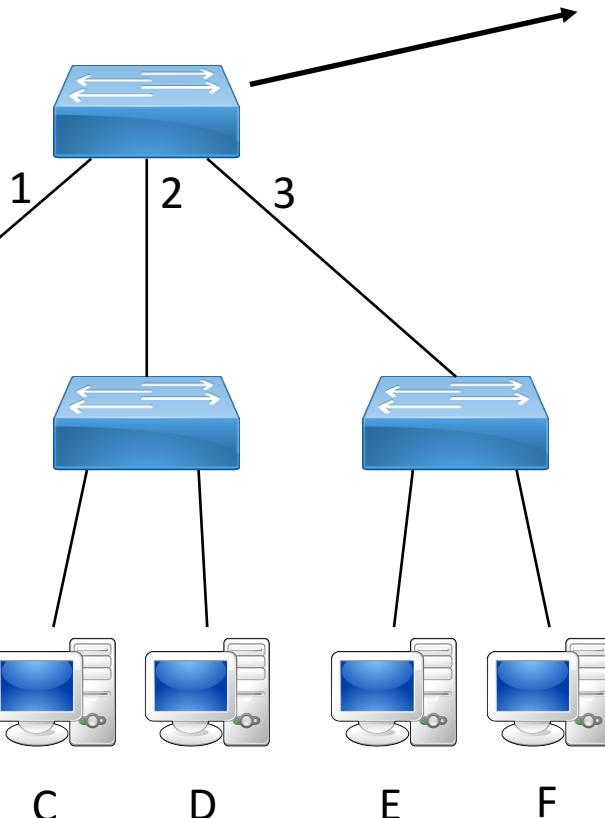
- **Inizializzazione:**
 - Tabella di Switching vuota
- **All'arrivo di ogni pacchetto:**
 - Indirizzo sorgente inserito in tabella
- **Inoltro:**
 - Se indirizzo in tabella: inoltro porta corrispondente
 - Se indirizzo non in tabella: inoltro broadcast



Transparent Bridging : autoapprendimento

Pacchetto 1

From: A – To: C



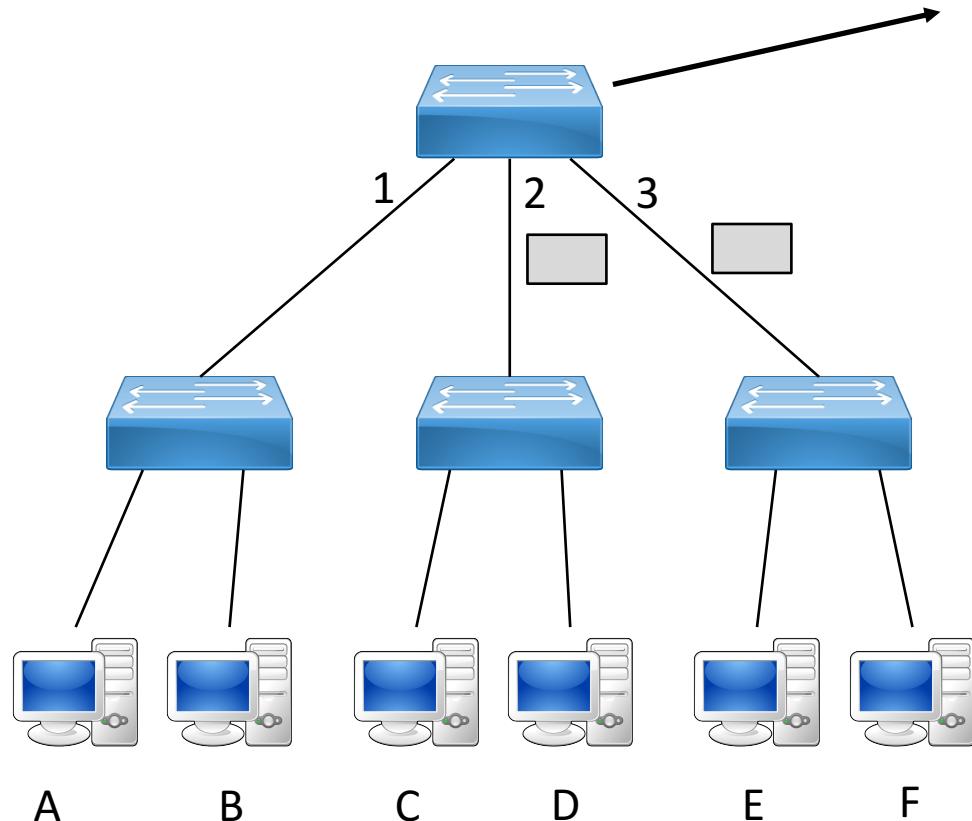
Destinazione	Porta	time



Transparent Bridging : autoapprendimento

Pacchetto 1

From: A – To: C



Destinazione	Porta	time
MAC_A	1	8:01

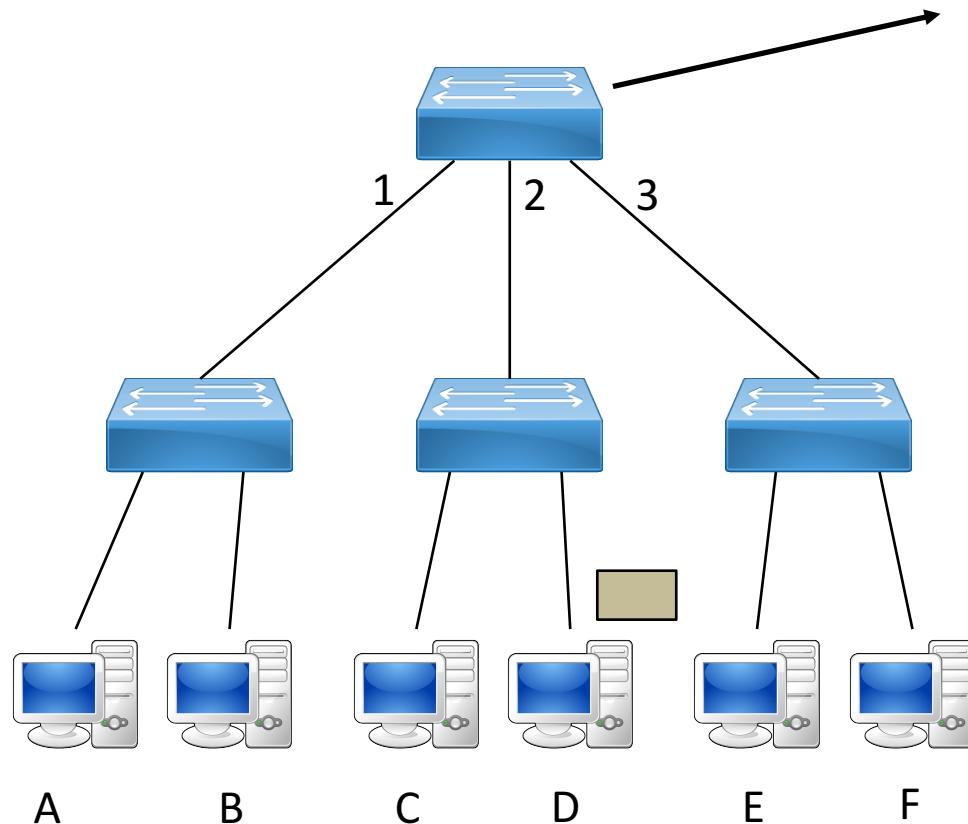
Inoltro:
Porta 2 e Porta 3



Transparent Bridging : autoapprendimento

Pacchetto 2

From: D – To: E



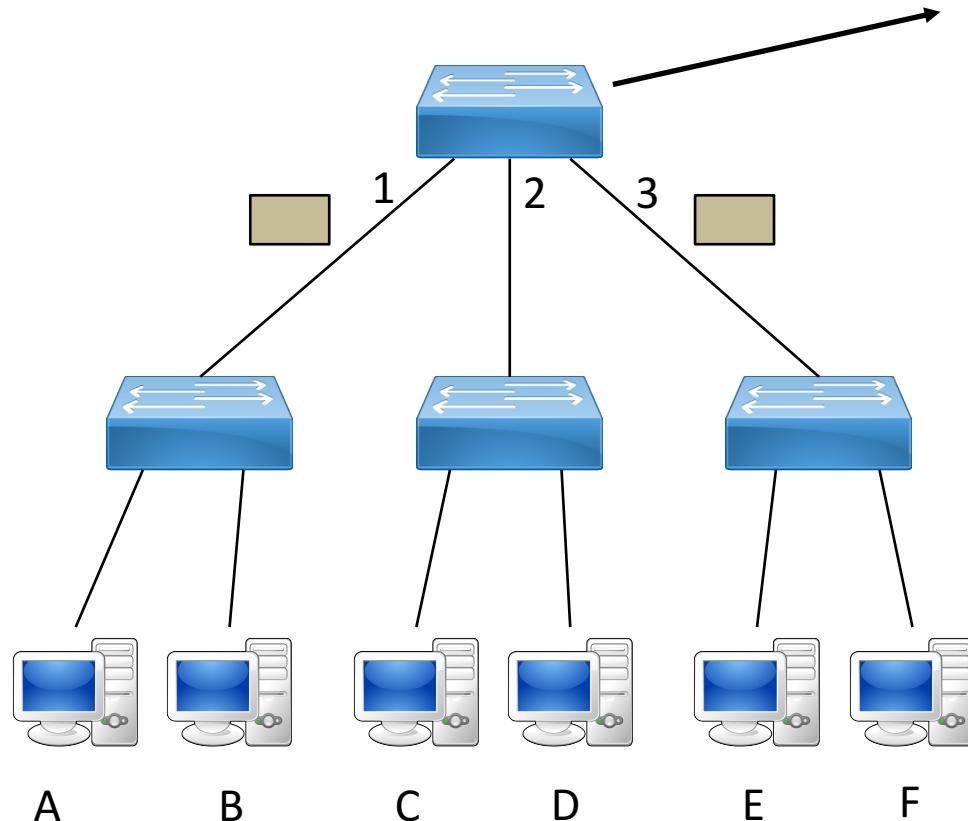
Destinazione	Porta	time
MAC_A	1	8:01



Transparent Bridging : autoapprendimento

Pacchetto 2

From: D – To: E



Destinazione	Porta	time
MAC_A	1	8:01
MAC_D	2	8:04

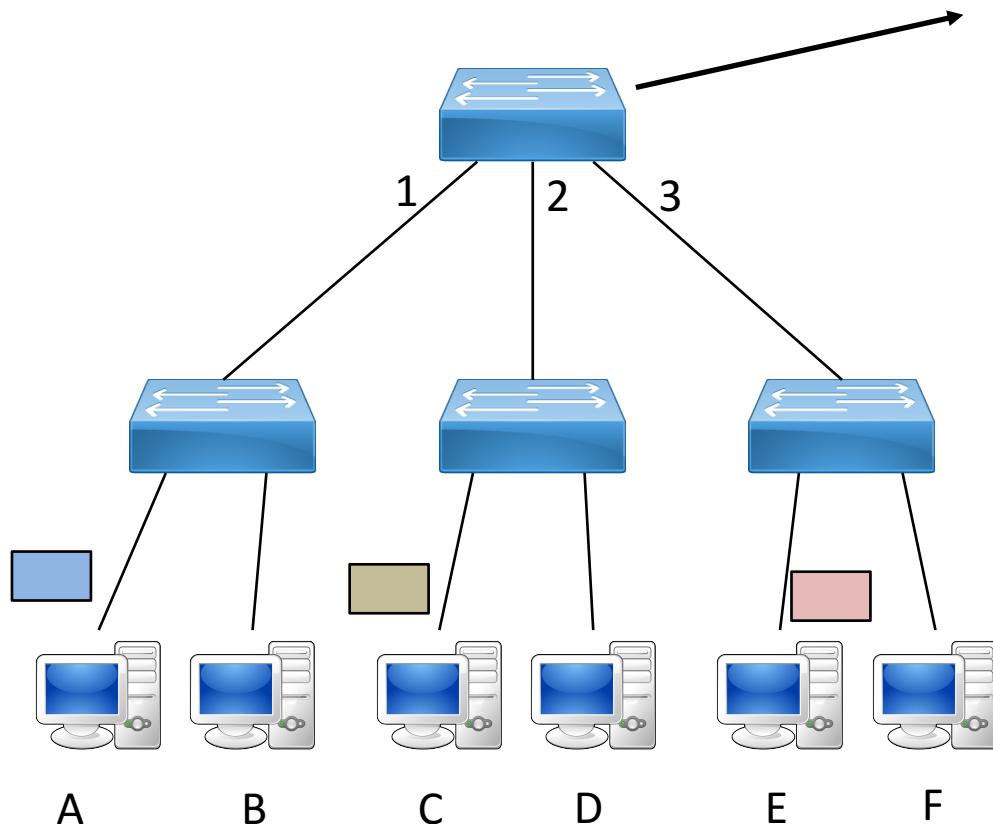
Inoltro:
Porta 1 e Porta 3

A B C D E F



Transparent Bridging : autoapprendimento

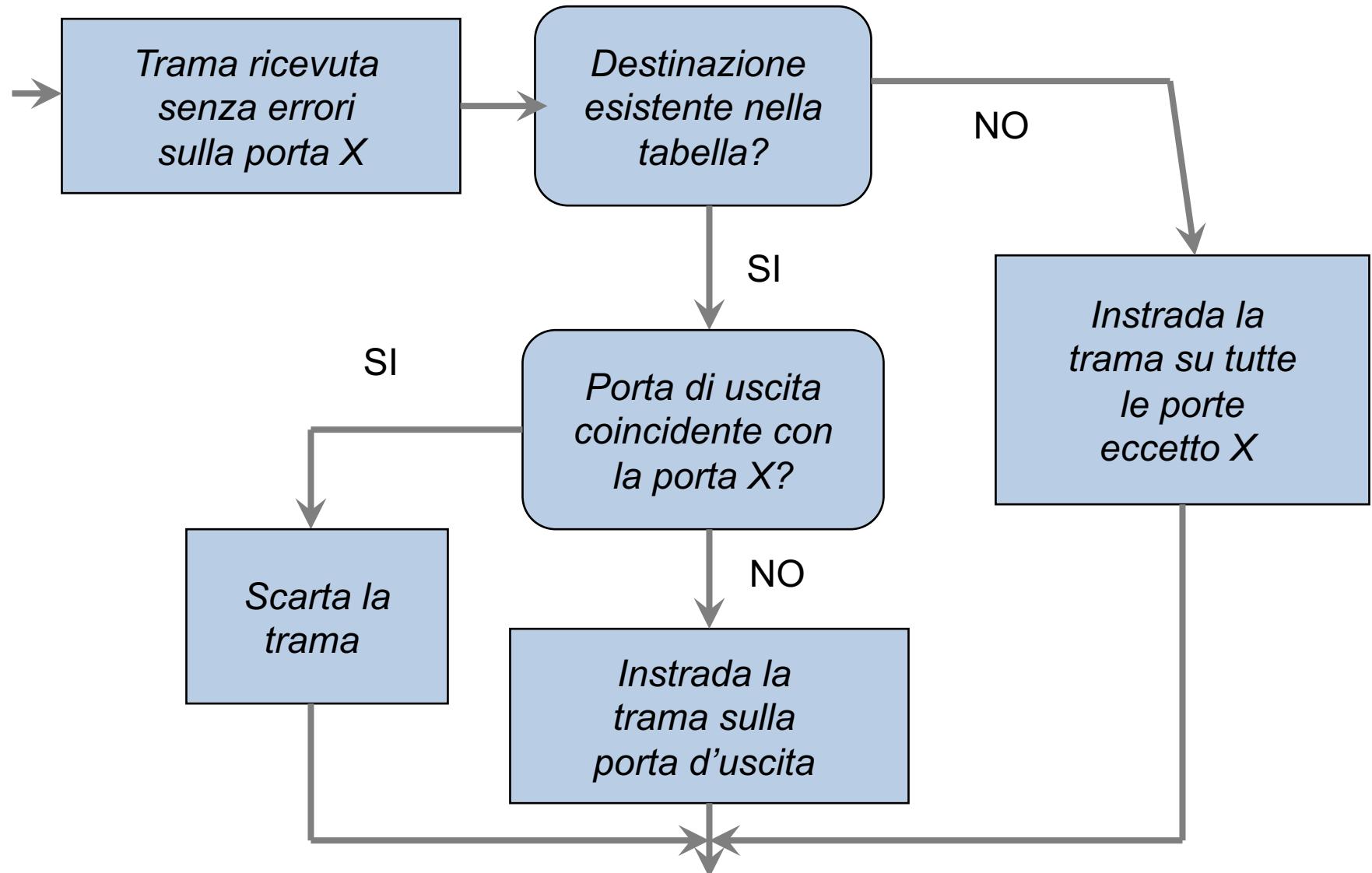
Dopo che tutte le stazioni hanno trasmesso almeno un pacchetto



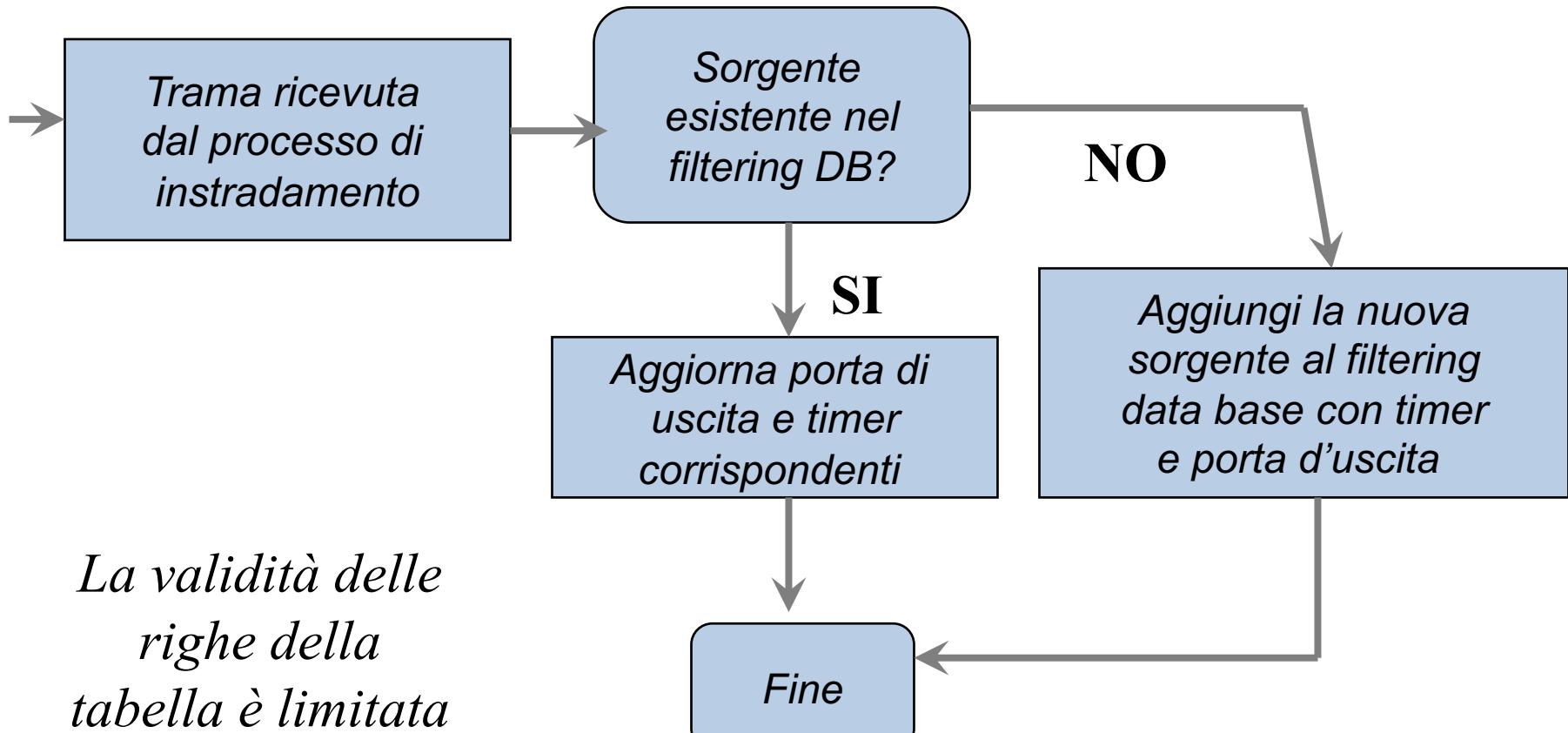
Destinazione	Porta	time
MAC_A	1	8:01
MAC_D	2	8:04
MAC_C	2	8:06
MAC_E	3	8:09
MAC_B	1	8:11
MAC_F	3	8:13



Forwarding

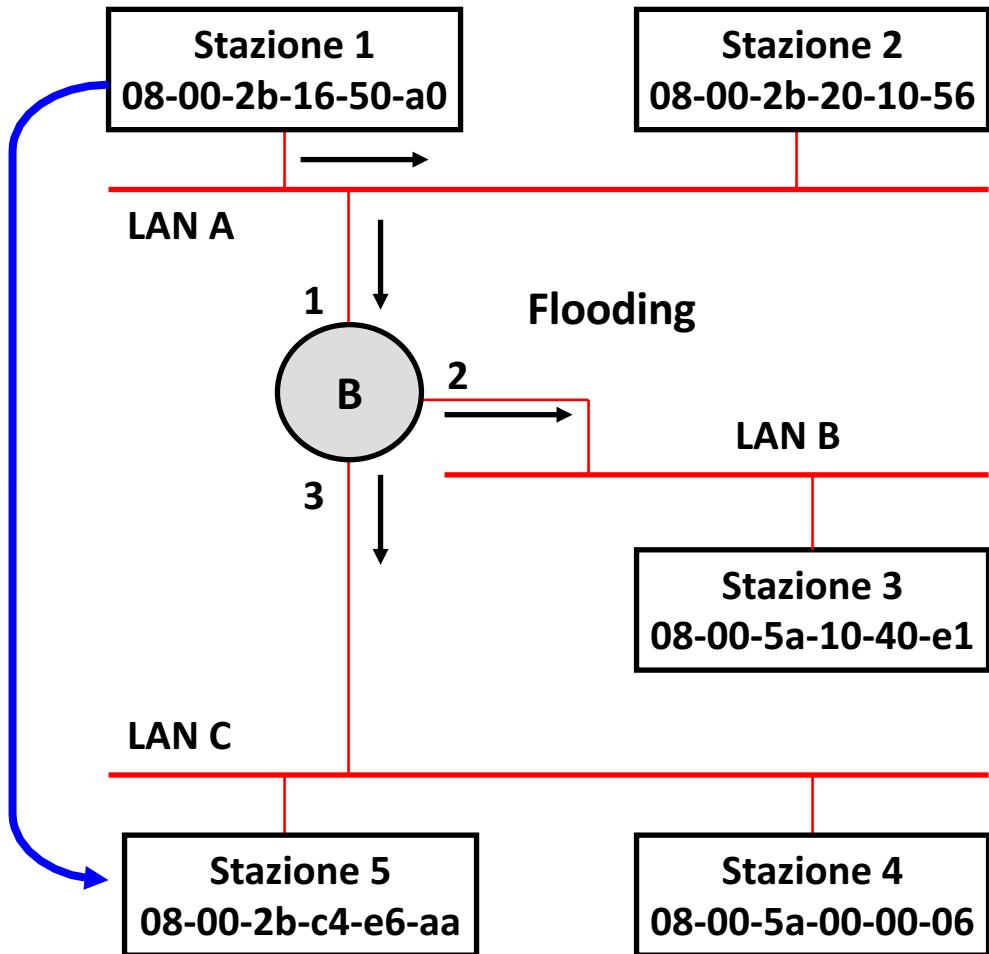


Learning



Esempio - 1

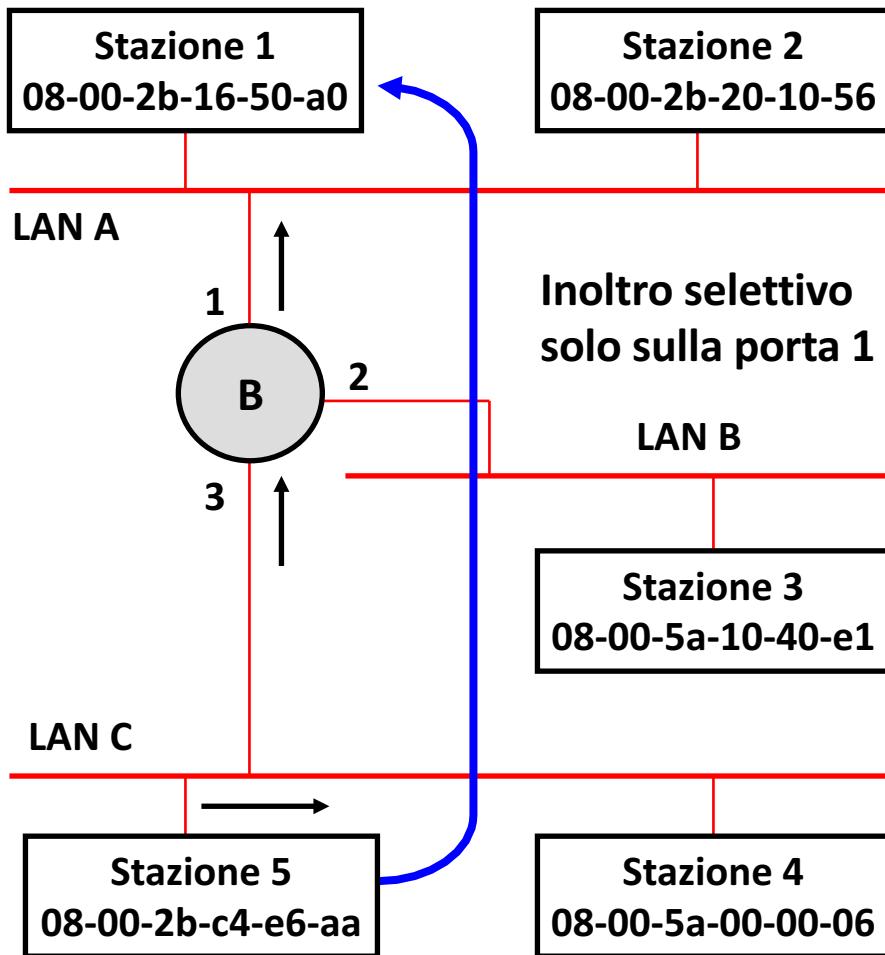
Inoltro di trama



Port	MAC address	Ageing time
1	08-00-2b-16-50-a0	0

Esempio - 2

Inoltro di trama

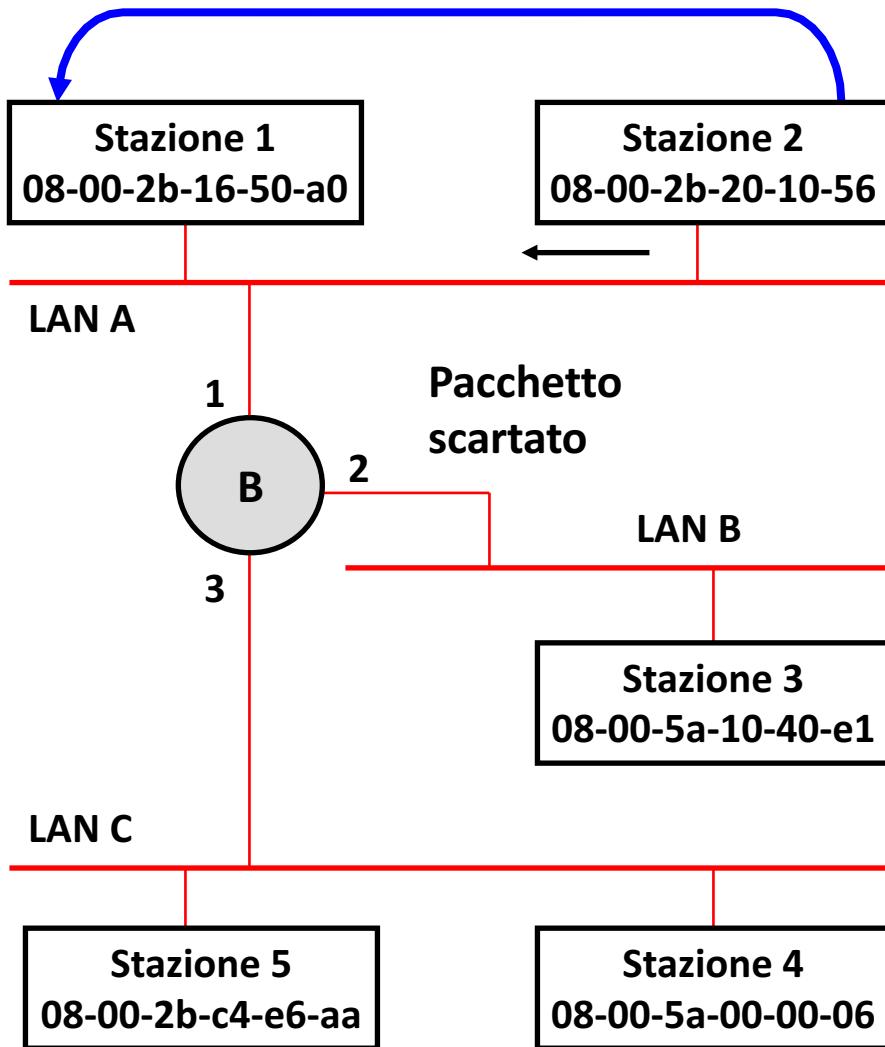


Inoltro selettivo
solo sulla porta 1

Port	MAC address	Ageing time
1	08-00-2b-16-50-a0	5
3	08-00-2b-c4-e6-aa	0

Esempio -3

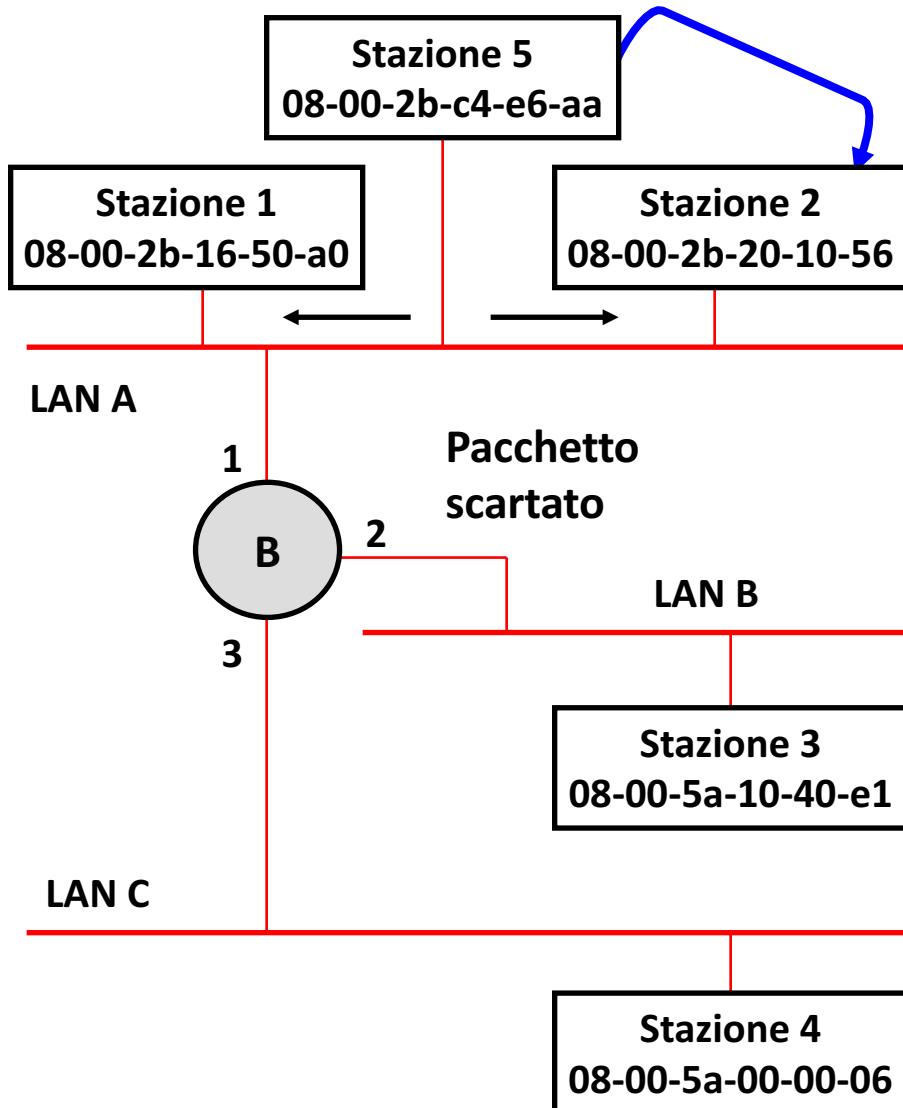
Limitazione del traffico



Port	MAC address	Ageing time
1	08-00-2b-16-50-a0	18
3	08-00-2b-c4-e6-aa	13
1	08-00-2b-20-10-56	0

Esempio - 4

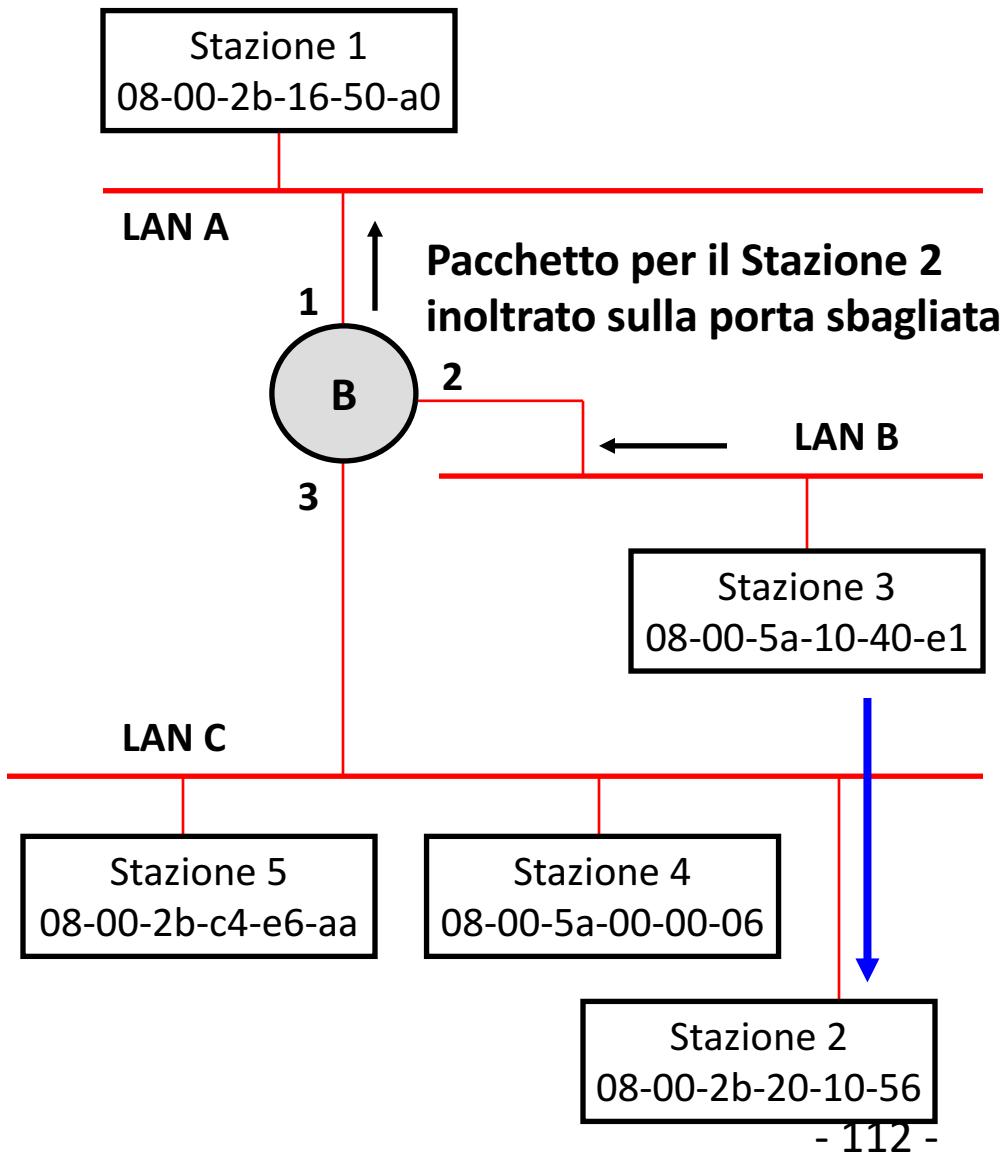
Spostamento Stazione 5 e aggiornamento



Port	MAC address	Ageing time
1	08-00-2b-16-50-a0	40
1	08-00-2b-c4-e6-aa	0
1	08-00-2b-20-10-56	20

Esempio - 5

Spostamento Stazione 2 e inoltro errato

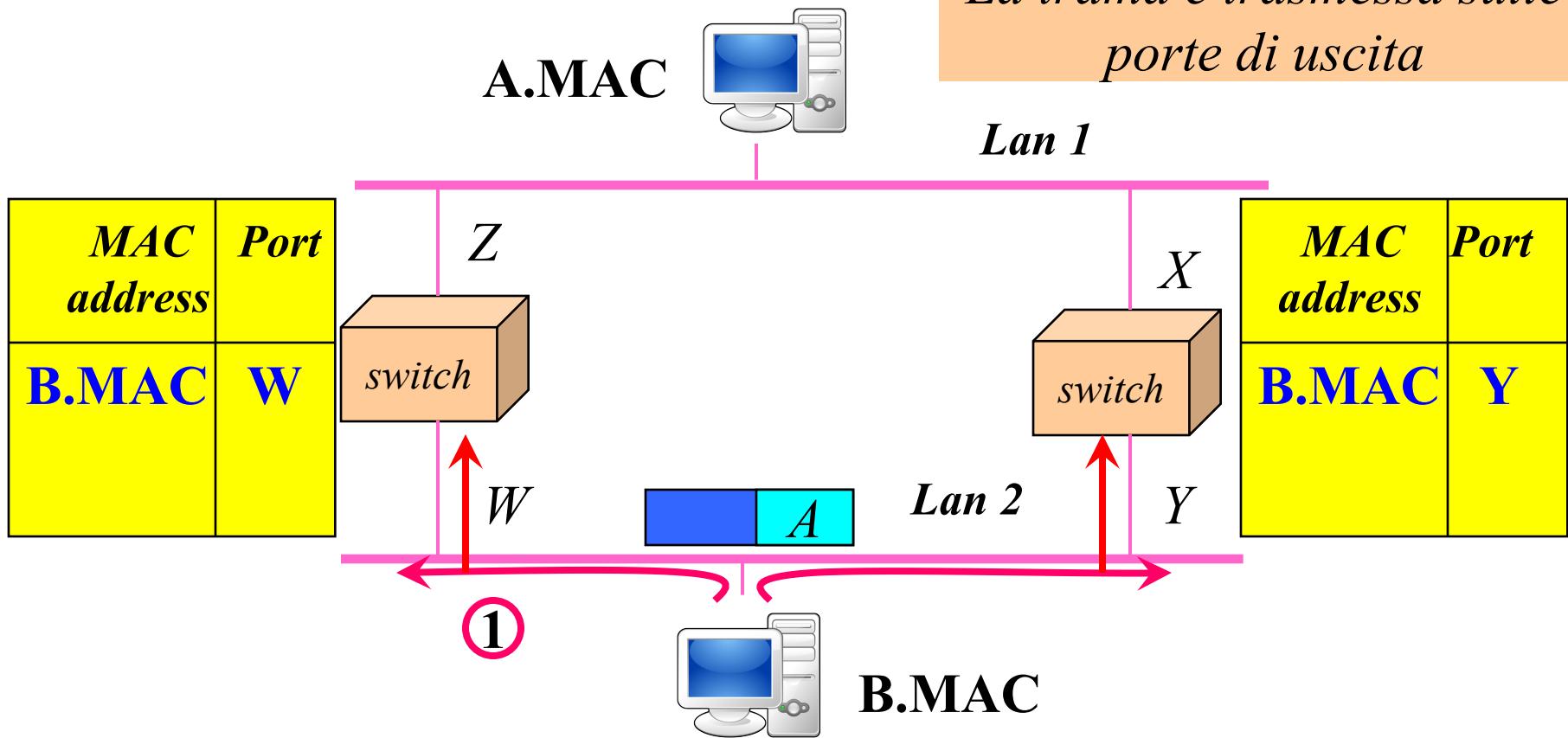


Port	MAC address	Ageing time
1	08-00-2b-16-50-a0	50
3	08-00-2b-c4-e6-aa	51
1	08-00-2b-20-10-56	40
2	08-00-5a-10-40-e1	0

Importanza dell'ageing-time

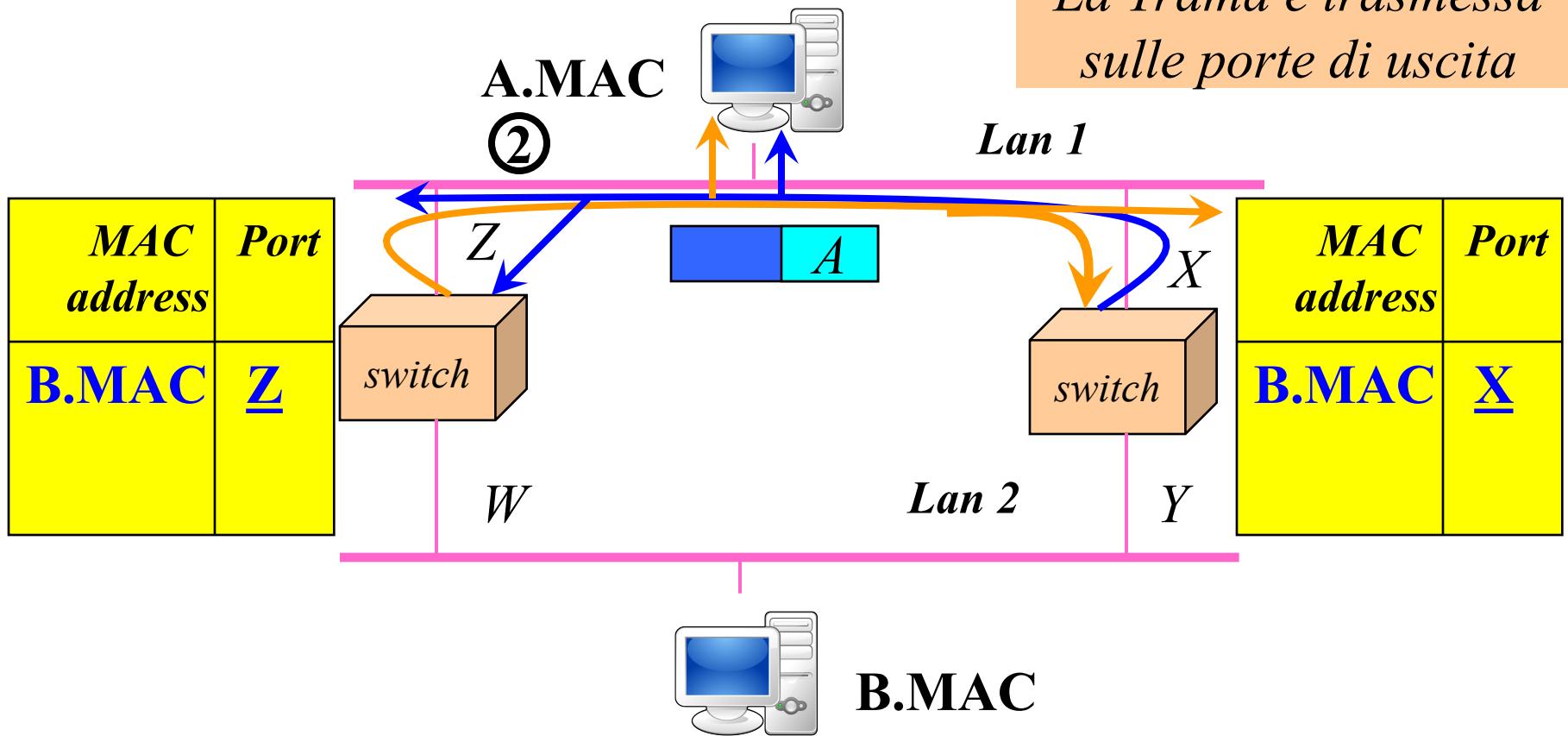
Se la rete non è un albero: Broadcast Storm

*La sorgente B è inserita nelle tabelle degli switch.
La trama è trasmessa sulle porte di uscita*



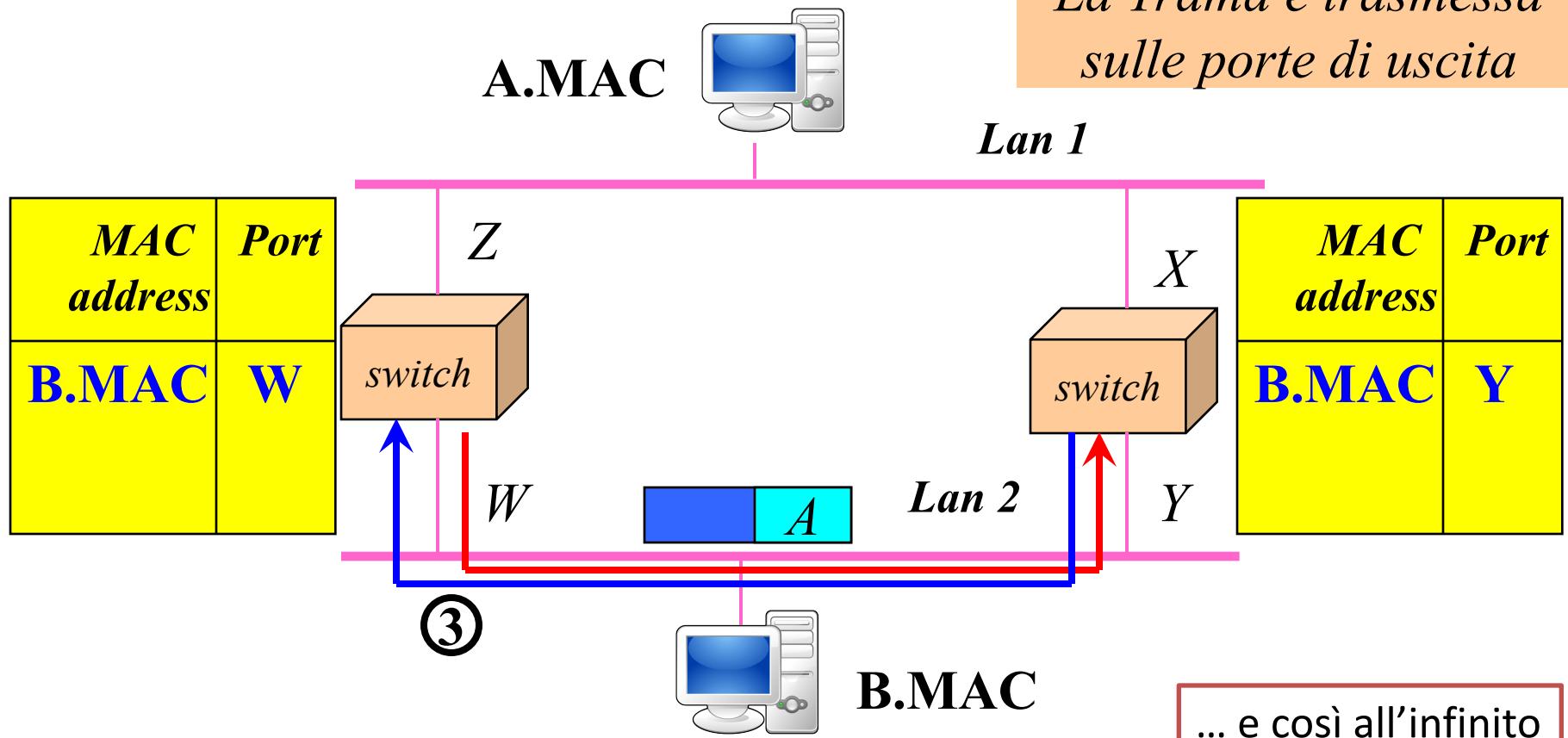
Se la rete non è un albero: Broadcast Storm

*Le righe della sorgente
B sono modificate.
La Trama è trasmessa
sulle porte di uscita*



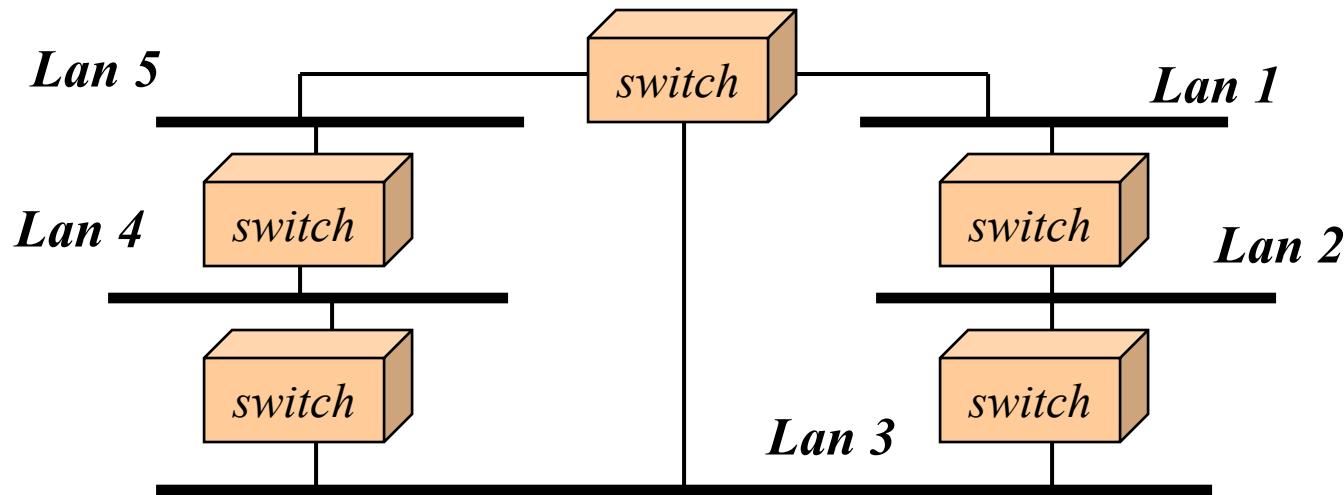
Se la rete non è un albero: Broadcast Storm

*Le righe della sorgente B sono modificate.
La Trama è trasmessa sulle porte di uscita*



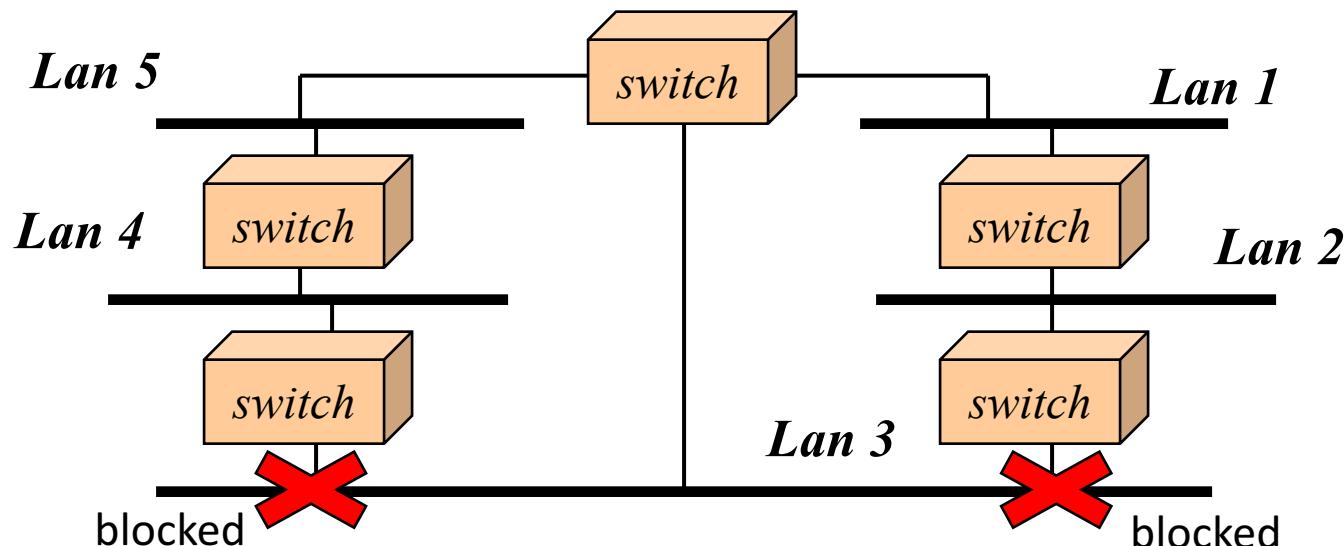
Spanning Tree

- Nella pratica la topologia è di solito una topologia magliata per garantire tolleranza ai guasti
- Per evitare il problema del broadcast storm: i bridge rendono inattive alcune porte in modo da ridurre la rete ad albero (spanning tree) nel funzionamento normale
- Un protocollo distribuito calcola lo spanning tree e lo modifica in caso di guasti (spanning tree protocol 802.1D)



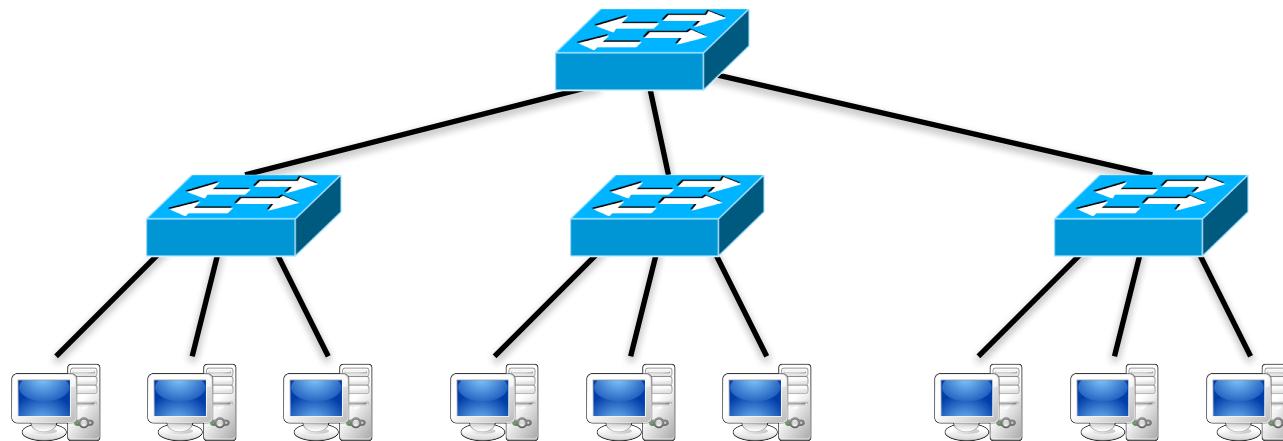
Algoritmo di Spanning Tree

- Permette di ricavare, a partire da una topologia fisica magliata, una topologia logica ad albero.
- La topologia logica ad albero e' realizzata ponendo in stato di 'blocco' delle porte.
- Una porta bloccata lascia passare i messaggi del protocollo di spanning tree ma non le trame dati.



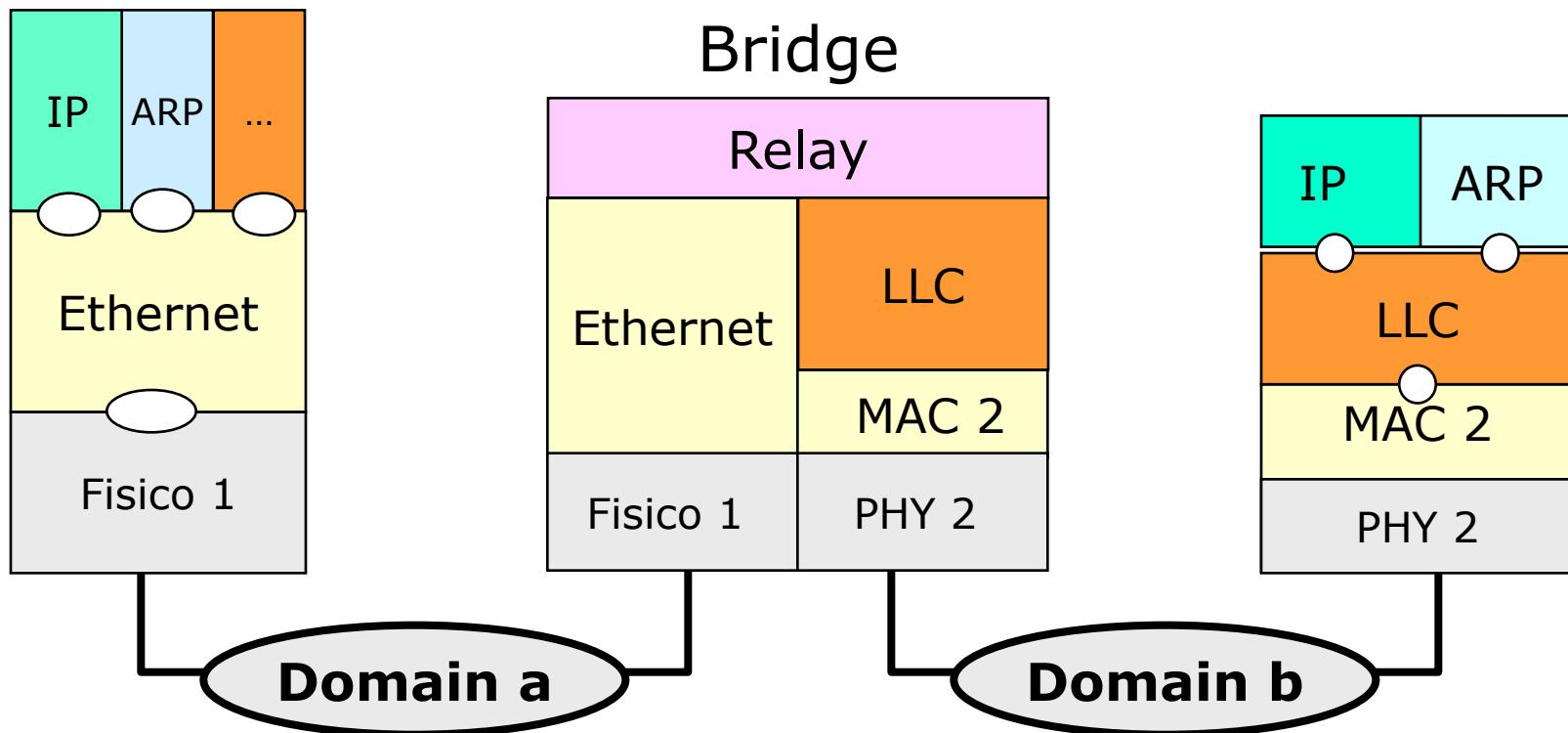
LAN completamente commutate

- La modalità **Ethernet full duplex** consente la trasmissione su doppino (twisted pairs) o fibra senza meccanismo di accesso multiplo (no CSMA-CD)
- Ciò consente la creazione di **LAN completamente commutate** che costituiscono la soluzione usata universalmente oggigiorno



Bridge con tecnologie eterogenee

- I Bridge/switch possono anche operare con tecnologie LAN differenti
- Come ad esempio i WiFi Access Point che sono collegati anche a un Distribution System basato su Ethernet



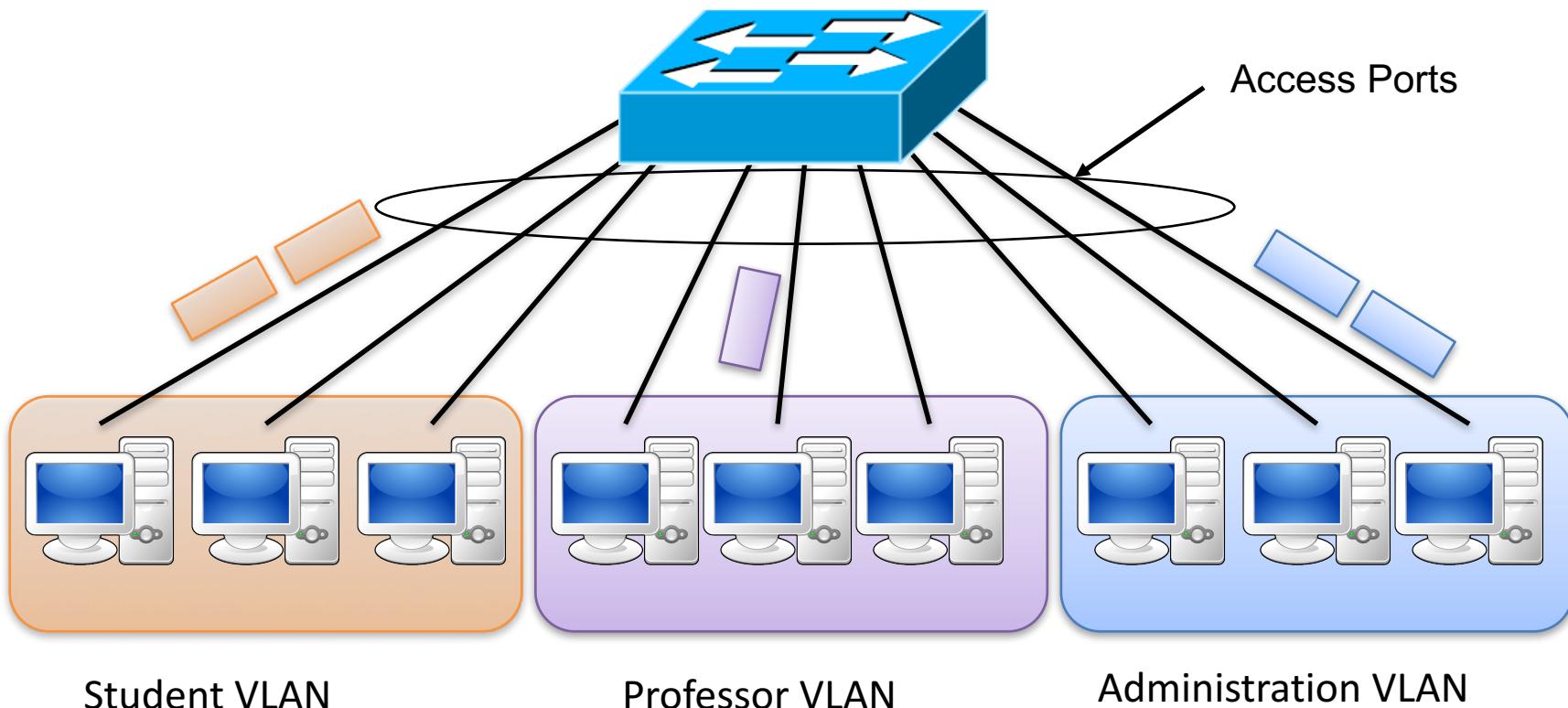
Virtual LAN (VLAN)

- Le **Virtual LANs** (VLAN) consentono di creare LAN logicamente separate su un'unica LAN fisica (commutata)
- Anche se le VLAN sono sulla stessa rete fisica, le stazioni su VLAN differenti non possono comunicare direttamente a livello 2 (esse sono su differenti domini broadcast)
- La comunicazione tra le diverse VLAN può però avvenire a livello 3 attraverso un router
- La separazione in VLAN è spesso dovuta a motivi di sicurezza e separazione di traffico



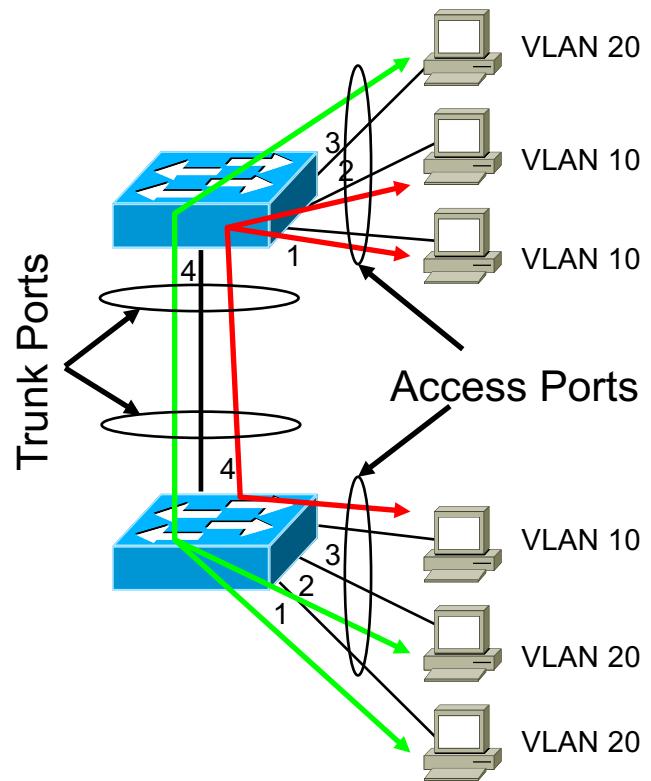
VLAN port based

- Le VLAN possono essere separate staticamente su base porta (solo reti ad un livello)
- Dominio di broadcast disgiunto per ogni VLAN, definito a priori assegnando le porte degli switch (access port) a VLAN diverse
- Indipendente dalla configurazione fisica delle VLAN



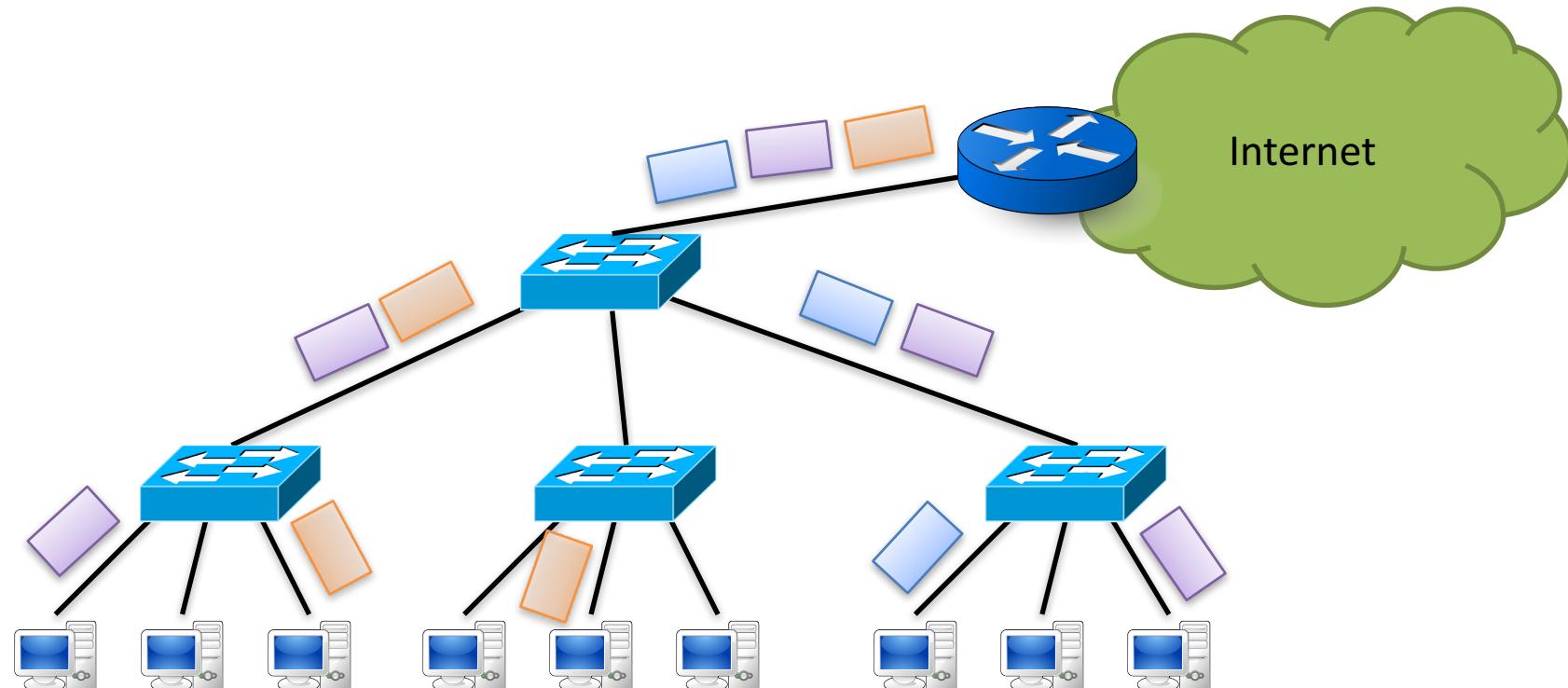
Interconnessione di VLAN

- I link di collegamento tra gli switch (**trunk**) trasportano il traffico di tutte le VLAN (sono condivisi)
 - Una **trunk port** non può essere assegnata ad un'unica VLAN
 - E' necessario un metodo per distinguere le trame di VLAN diverse sui trunk



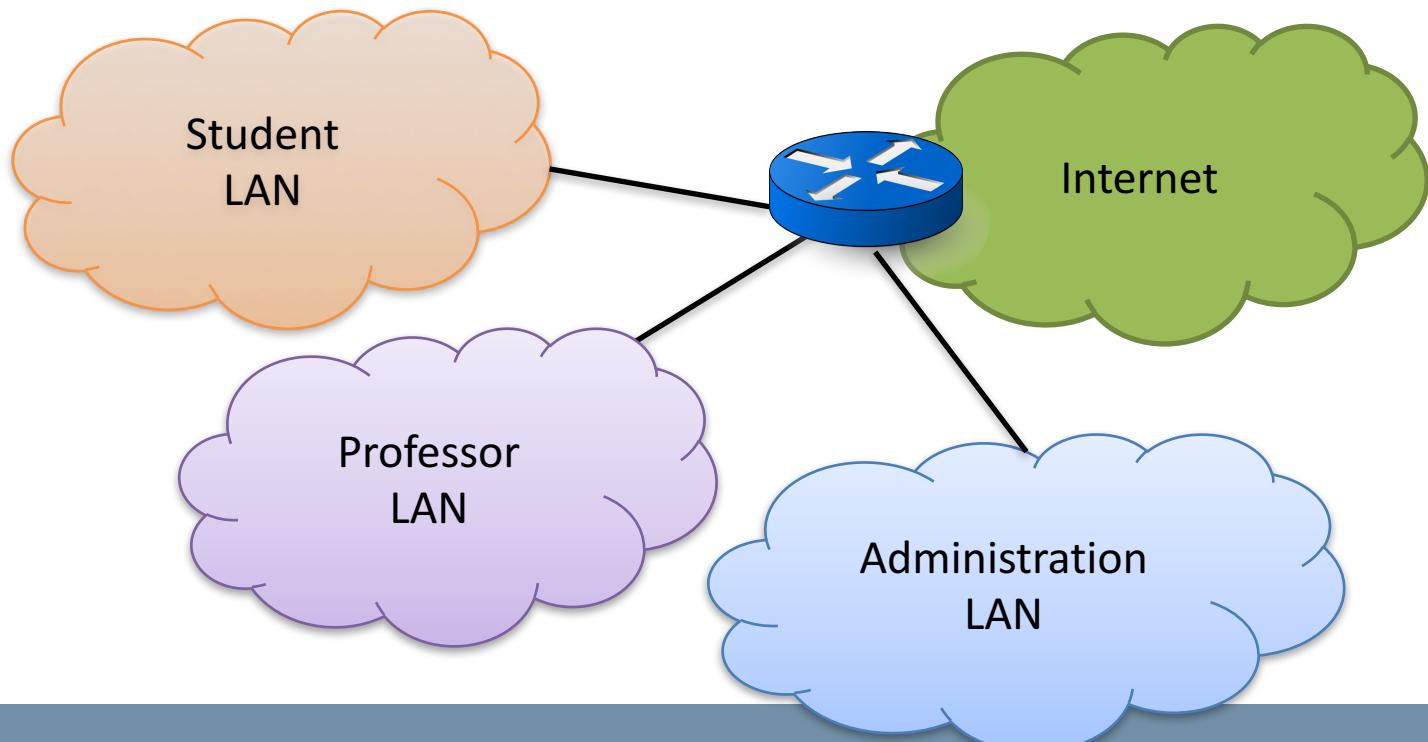
VLAN tagging

- Sui link che interconnettono switch e router (o switch e switch), le trame di differenti VLAN devono essere differenziate da etichette (tag)
- A questo scopo è usato il protocollo 802.1q (LAN tagging)

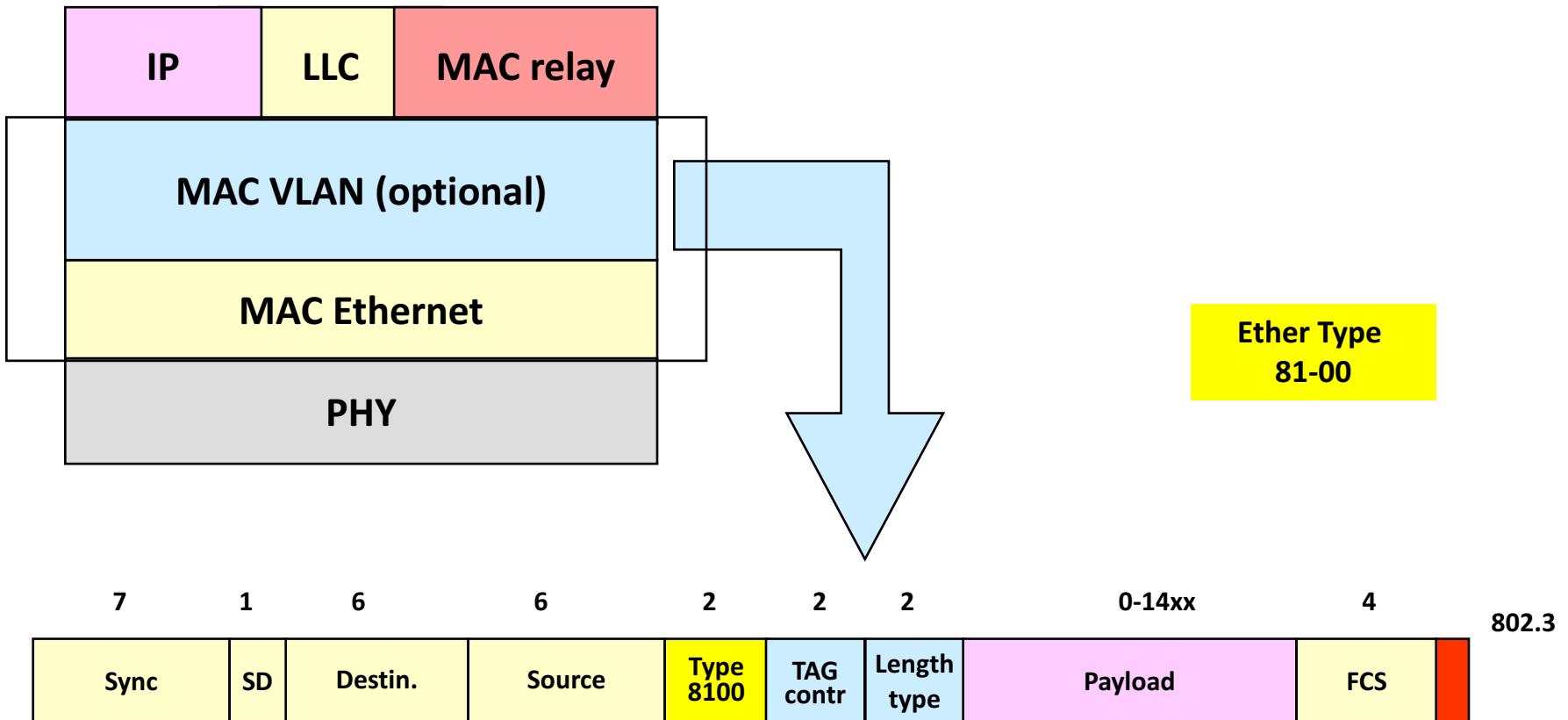


VLAN tagging

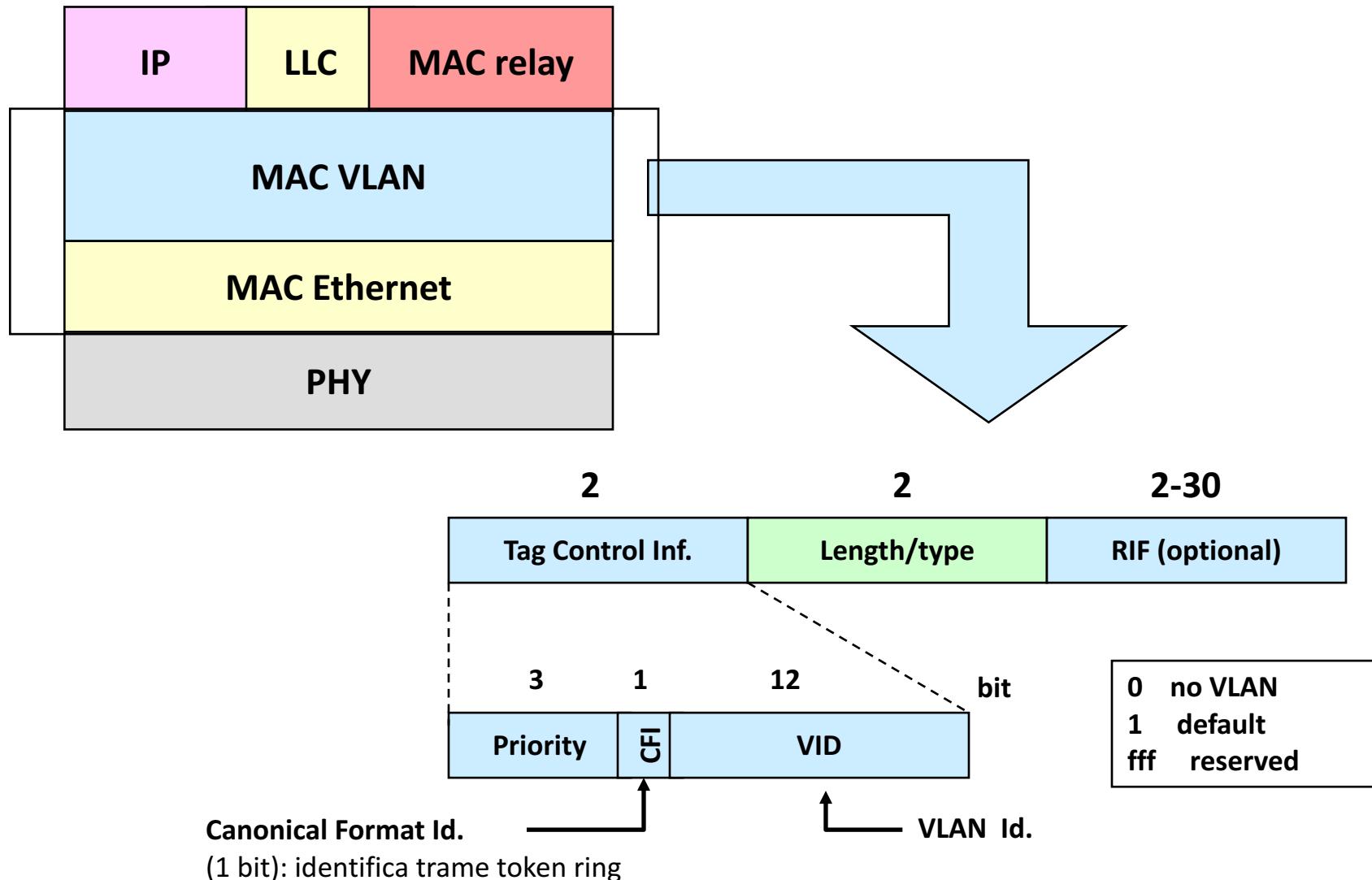
- Dal punto di vista dei dispositivi di livello 3 (router), le VLAN equivalgono a LAN fisiche separate
- Tuttavia, le interfacce dei router connessi agli switch che gestiscono le VLAN devono supportare il protocollo di VLAN tagging



VLAN tagging: 802.11q



VLAN tagging: 802.11q





5 appendice – Esempi di protocolli di linea HDLC, PPP

High-level Data Link Control

- Vediamo un esempio di protocollo di linea (il più diffuso)
- E' istruttivo vedere come anche in un caso semplice come il protocollo di linea punto-punto ci siano molte opzioni e parametri nel protocollo
- standard ISO degli anni 60
- Caratteristiche:
 - orientato al bit
 - può operare in molti modi differenti e con diversi meccanismi di controllo d'errore e di flusso
 - half-duplex o full-duplex
 - master-slave o peer-to-peer



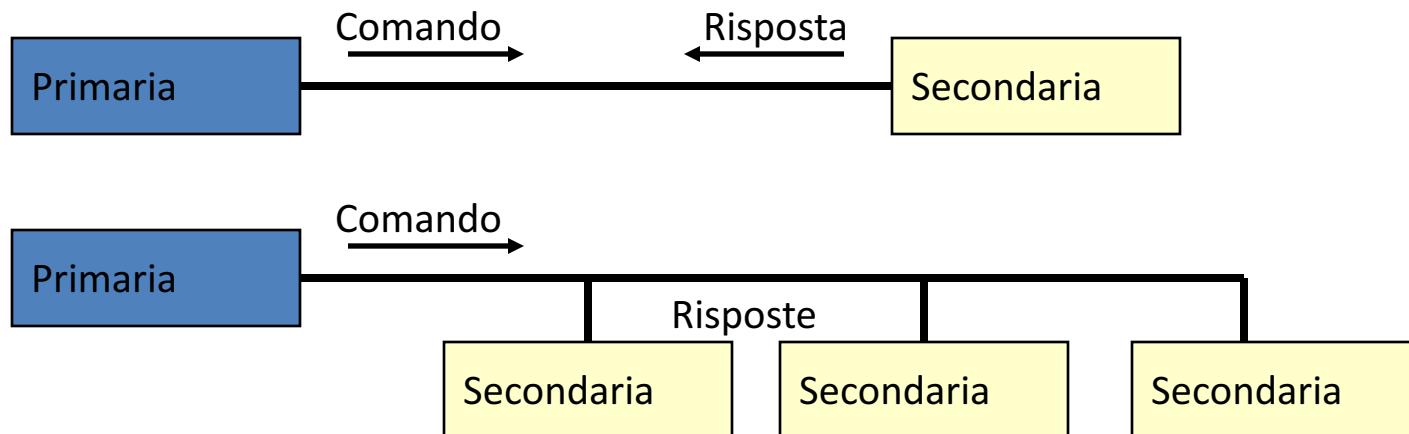
HDLC: configurazioni

- **Tipo di stazione**
 - *Primaria*: responsabile del collegamento, emette comandi
 - *Secondaria*: asservita alla primaria, emette risposte
 - *Combinata*: emette sia comandi, sia risposte
- **Configurazione del collegamento**
 - *Sbilanciata*: 1 primaria, ≥ 1 secondarie
 - *Bilanciata*: 2 stazioni combinate
- **Modi di trasferimento**
 - *Asynchronous Balanced Mode (ABM)*: configurazione bilanciata
 - 2 stazioni combinate
 - Trasmissione di tipo full-duplex
 - *Normal Response Mode (NRM)*: configurazione sbilanciata
 - 1 stazione primaria e almeno 1 stazione secondaria
 - Trasmissione di tipo half-duplex
 - *Asynchronous Response Mode (ARM)*: configurazione sbilanciata
 - Come NRM ma il secondario può iniziare la trasmissione senza permesso



HDLC: Modalità di funzionamento

- **Normal Response Mode (NRM)**
 - Una stazione primaria è collegata a una o più stazioni secondarie in modalità *half-duplex*.
 - Solo la stazione primaria può inviare i comandi e le stazioni secondarie trasmettono solo a seguito di un permesso (*polling*) esplicito inviato dalla stazione primaria: *half-duplex*



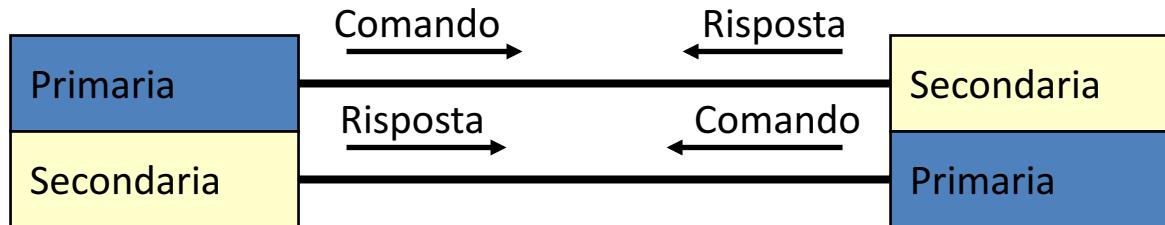
HDLC: Modalità di funzionamento

- ***Asynchronous Response Mode (ARM)***
 - Anche in questo caso come nel NRM il colloquio è di tipo sbilanciato, ma la stazione secondaria ha la possibilità di iniziare una trasmissione senza il permesso esplicito della stazione primaria iniziando così un colloquio *full-duplex*. (poco usata)



HDLC: Modalità di funzionamento

- **Asynchronous Balanced Mode (ABM)**
 - Fornisce una modalità di funzionamento bilanciato su configurazioni punto-punto tra stazioni combinate che possono, in modalità *full-duplex*, inviare informazioni in modo indipendente ed asincrono.



Trama HDLC: Flag

1 byte

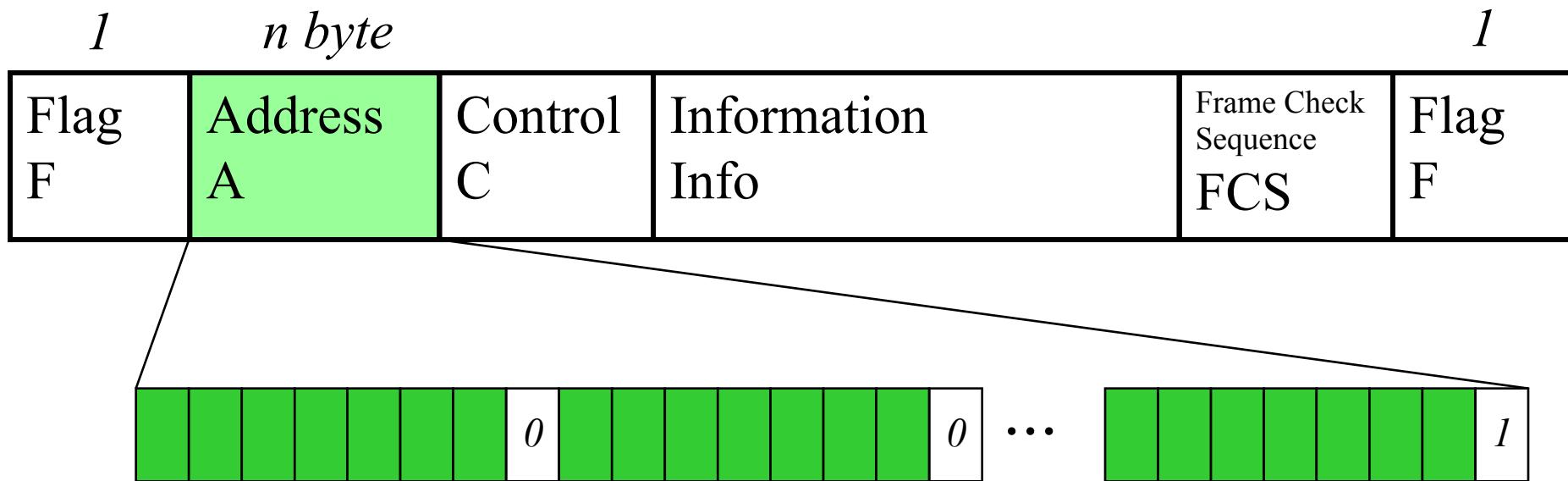
1 byte

Flag F	Address A	Control C	Information Info	Frame Check Sequence FCS	Flag F
-----------	--------------	--------------	---------------------	--------------------------------	-----------

- uso del bit stuffing
- di solito in caso di mancanza di informazione si esegue l'invio continuo dei flag



Trama HDLC: Indirizzo



- normalmente di 8 bit, ma può essere esteso a n byte (modalità EXTENDED)
- l'ultimo bit di ogni byte è usato per indicare se segue un ulteriore byte del campo A



Trama HDLC: indirizzo

$1 \quad 1 \div n$

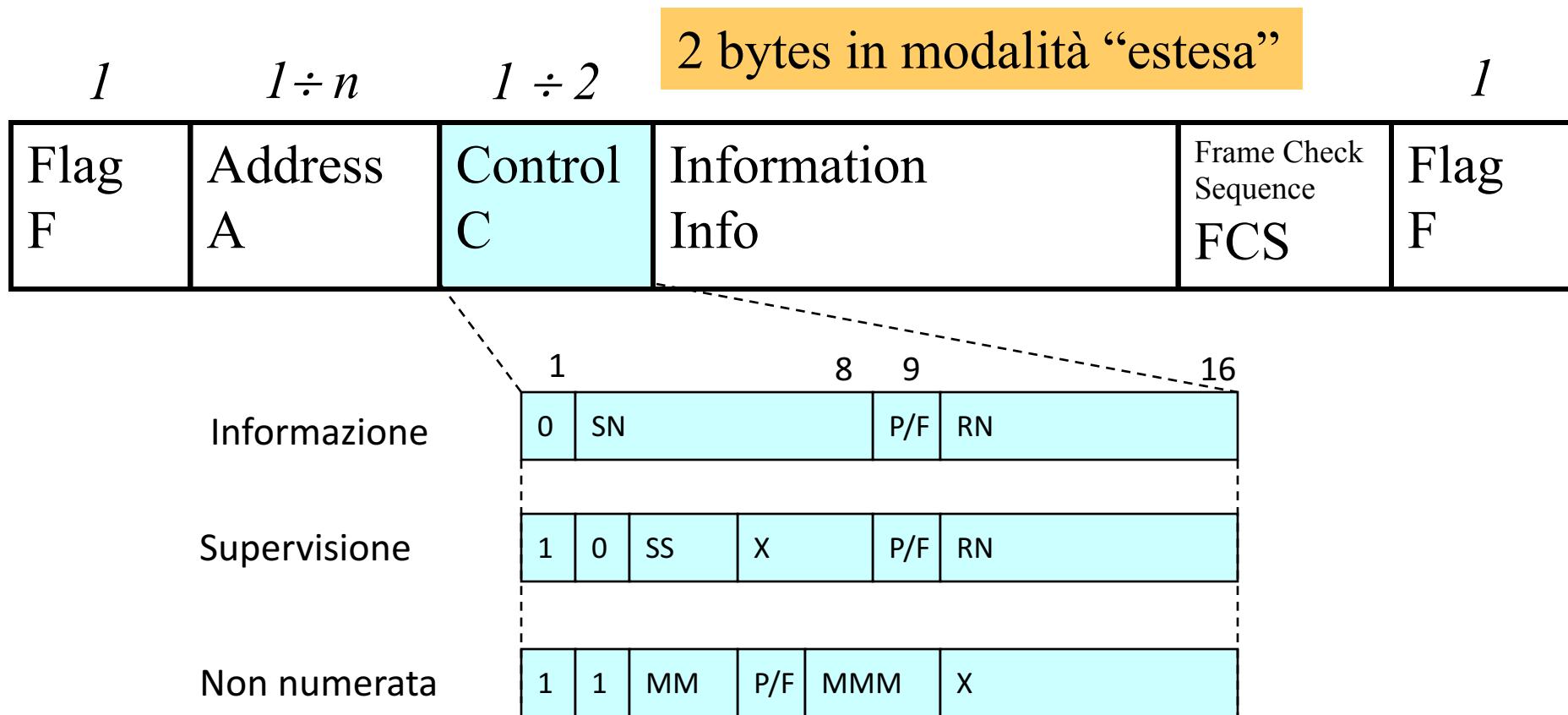
1

Flag F	Address A	Control C	Information Info	Frame Check Sequence FCS	Flag F
-----------	--------------	--------------	---------------------	--------------------------------	-----------

- **L'indirizzo contenuto può essere quello della stazione destinataria o quello della stazione sorgente**
 - nelle modalità sbilanciate (NRM, ARM) è sempre quello della stazione secondaria
 - nella modalità ABM è quello della stazione destinataria



Trama HDLC: Campo di controllo



SN - Send Number

RN - Request Number

P/F - Polling bit

SS - indicatore trame di supervisione

M - Modificatore di funzione



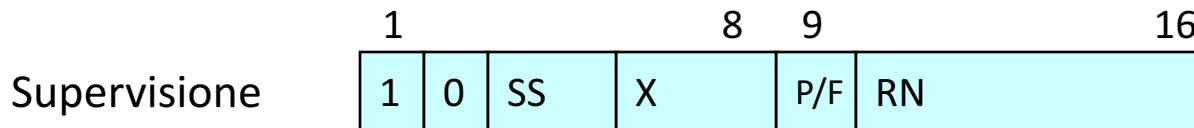
Trama HDLC: Trame di informazione (I)



- Sono trame numerate per la trasmissione di informazione d'utente contenuta nel campo I
- Consentono il riscontro delle trame ricevute in modalità piggybacking
- Consentono il polling (bit P alzato)
- e la chiusura (bit F (Final) alzato) della controparte



Trama HDLC: Trame di supervisione (S)



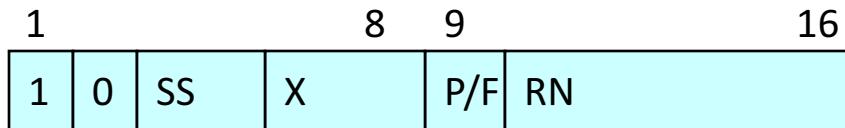
Sono trame numerate per il controllo dell'invio del flusso di informazione

ACK, NAK e controllo di flusso.

Comandi	SS	Risposte
RR Receiver Ready	00	RR Receiver Ready
RNR Receiver Not Ready	10	RNR Receiver Not Ready
REJ Reject	01	REJ Reject
SRJ Selective Reject	11	SRJ Selective Reject



Trama HDLC: Trame di supervisione (S)



- **RR (Receiver Ready)**, è normalmente usato come ACK e il campo RN contiene la prossima trama attesa (riscontro delle trame fino a RN-1)
- **RNR (Receiver Not Ready)**, serve a bloccare l'invio di trame da parte dell'altra stazione (controllo di flusso) e, contemporaneamente a riscontare le trame fino a RN-1
- **REJ (Reject)**, serve a richiedere la ritrasmissione delle trame da RN in avanti e, contemporaneamente, a riscontrare le trame fino a RN-1 (NAK)
- **SREJ (Selective Reject)**, è usato per richiedere la ritrasmissione della sola trama con numero RN (NAK)



Trama HDLC: Trame non numerate (U)



- Sono usate per l'invio di informazione di controllo (ad esempio per l'instaurazione delle connessioni) o per l'invio di informazione in modalità senza connessione.
- MM (modifier) indica il tipo di trama



Trama HDLC: Trame non numerate (U)

Utilizzate per l'instaurazione e il controllo della connessione

Comandi	Risposte
SNRM Set Normal Response Mode SARM Set Asynchronous Response Mode SABM Set Asynchronous balanced Mode SNRME SNRM estesa SARME SARM estesa SABME SABM estesa SIM Set Initialization Mode DISC Disconnect	UA Unnumbered Ack DM Disconnect RIM Request Inizialization Mode
RSET Reset	FRMR Frame
XID Exchange Identification	XID Exchange Identification RD RejectRequest Disconnect

Utilizzate per scambio di informazione

UI	Unnumbered Information	UI	Unnumbered Information
UP	Unnumbered Poll		



Trama HDLC: Campo informazione

1	$1 \div n$	$1 \div 2$	≥ 0	1
Flag F	Address A	Control C	Information Info	Frame Check Sequence FCS

Contiene l'informazione d'utente (dei livelli superiori)
può non essere presente

è presente solo nella trame I e nella trame UI usate per
trasferimento di informazione in modalità connectionless
lunghezza variabile



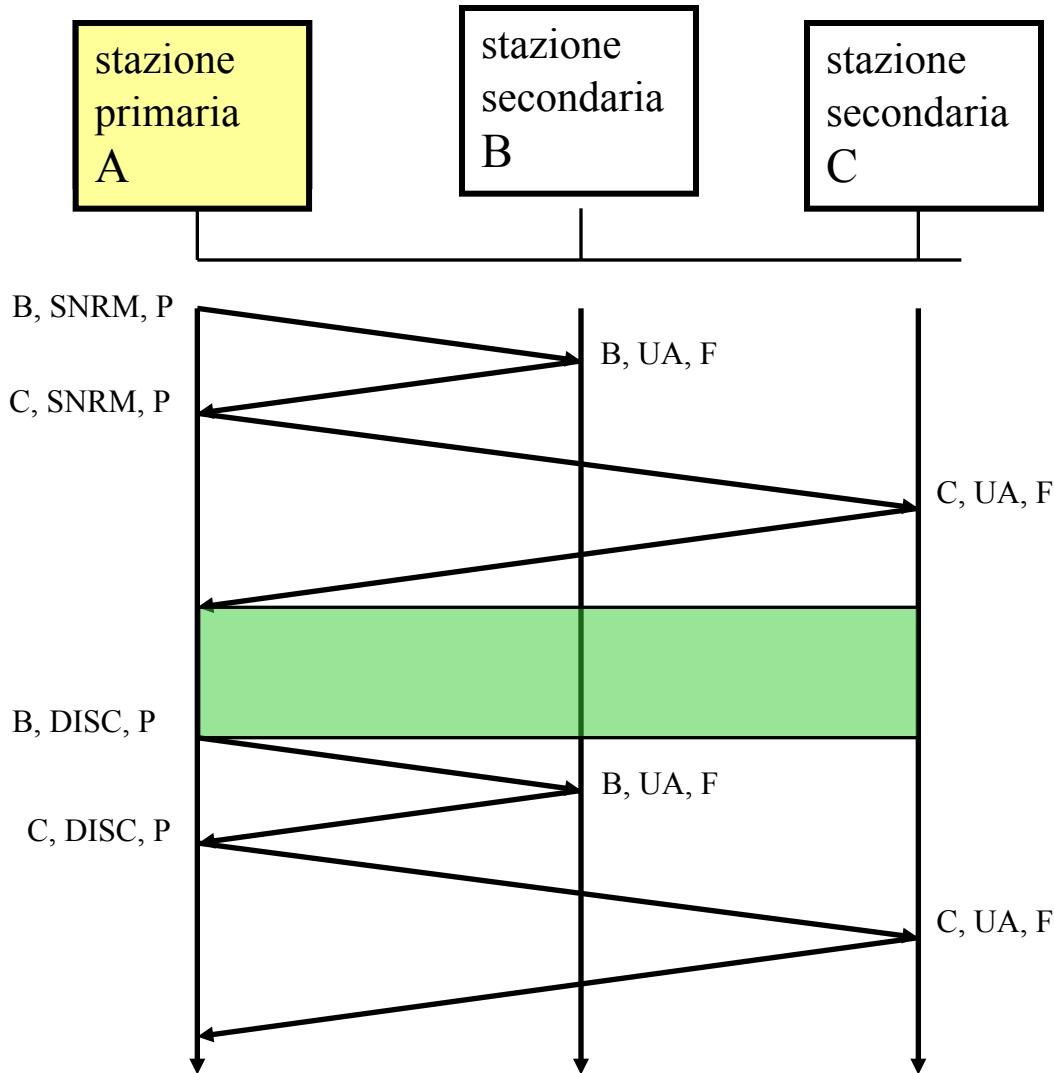
Trama HDLC: Campo di parità

1	$1 \div n$	$1 \div 2$	≥ 0	2	1
Flag F	Address A	Control C	Information Info	Frame Check Sequence FCS	Flag F

Contiene il codice rivelatore d'errore usato per riconoscere le trame errate



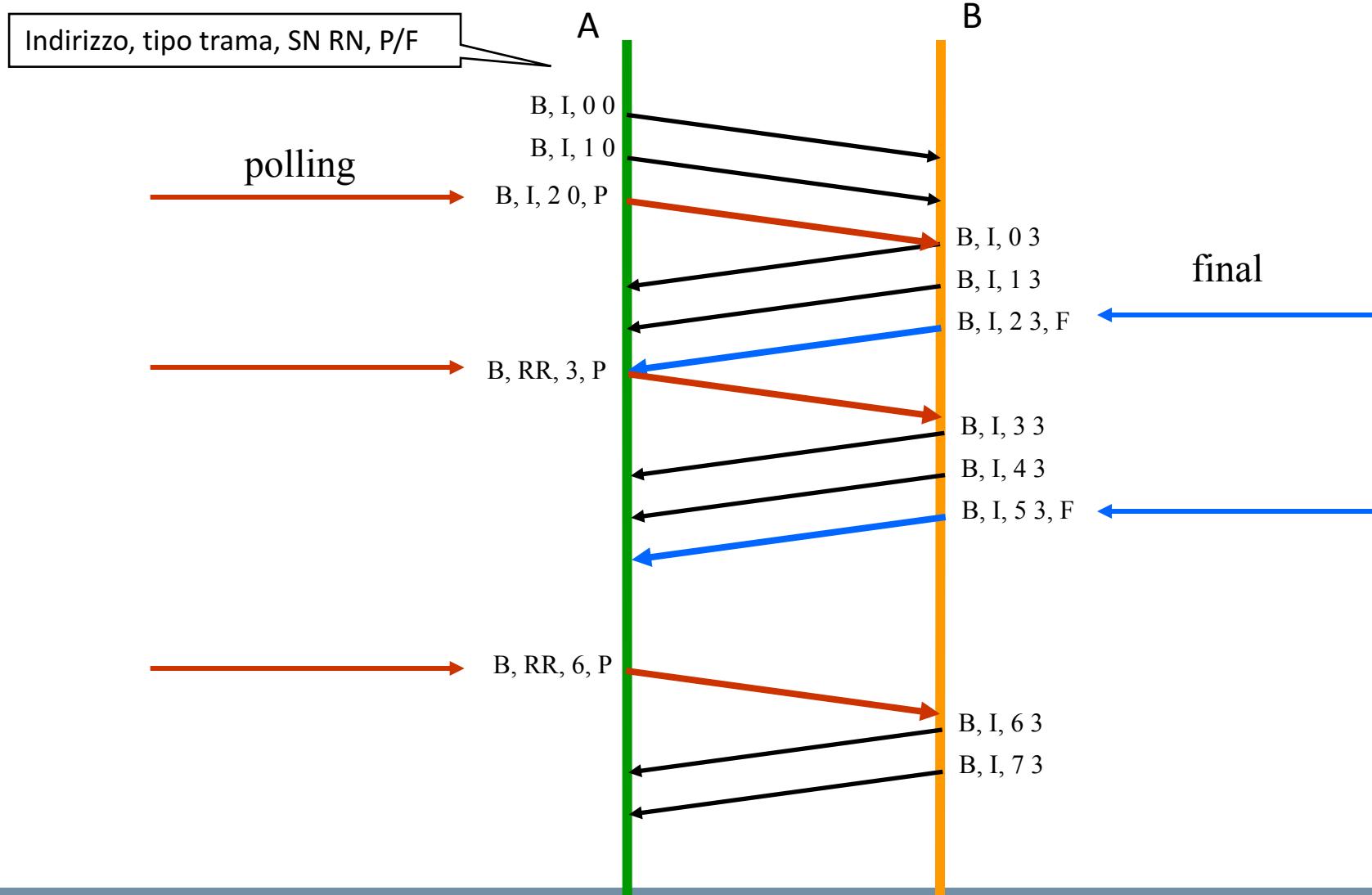
HDLC Instaurazione della connessione - modalità NRM



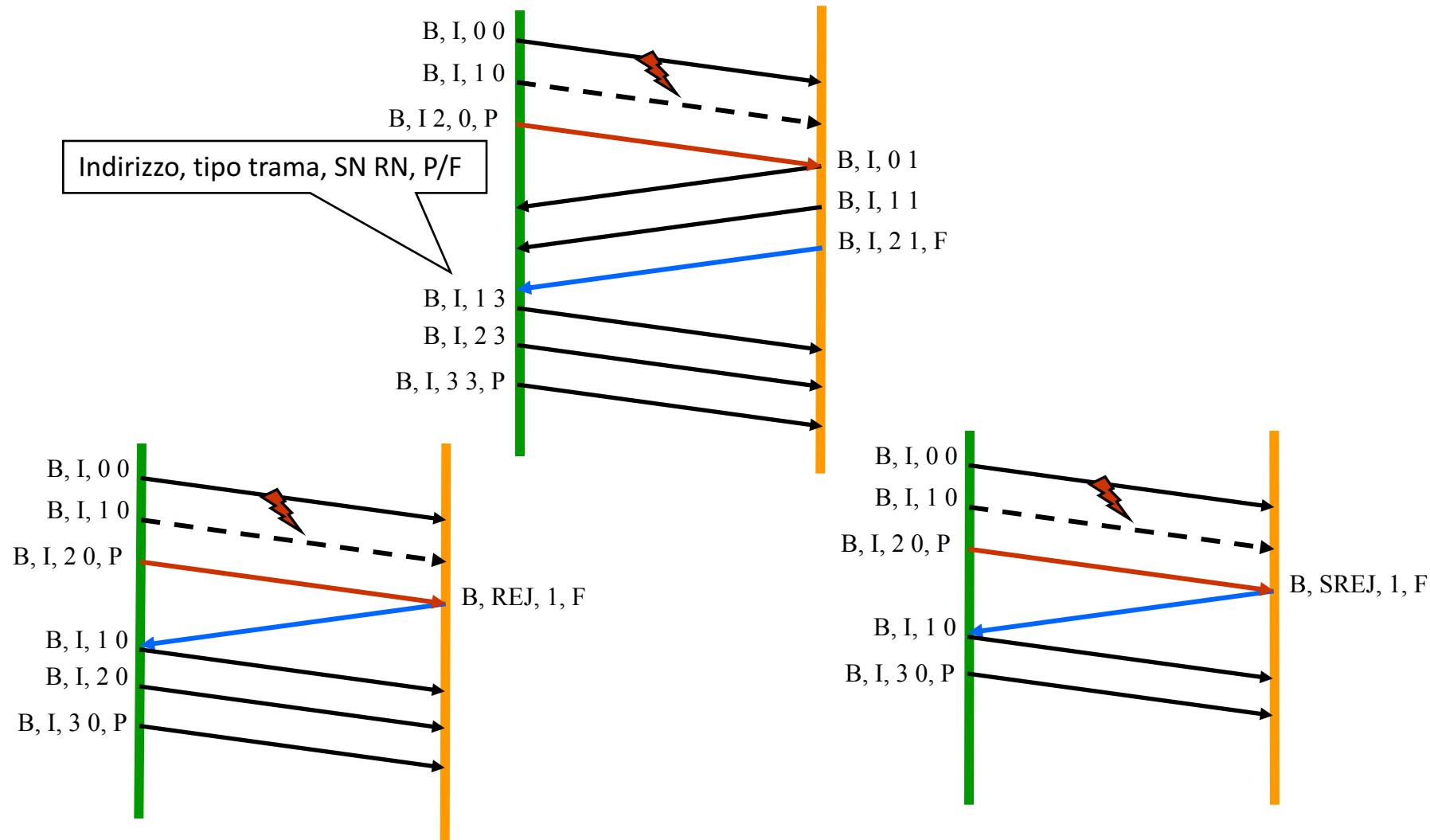
- **il bit P/F serve per passare il controllo dalla stazione primaria a quella secondaria e viceversa (Polling/Final)**



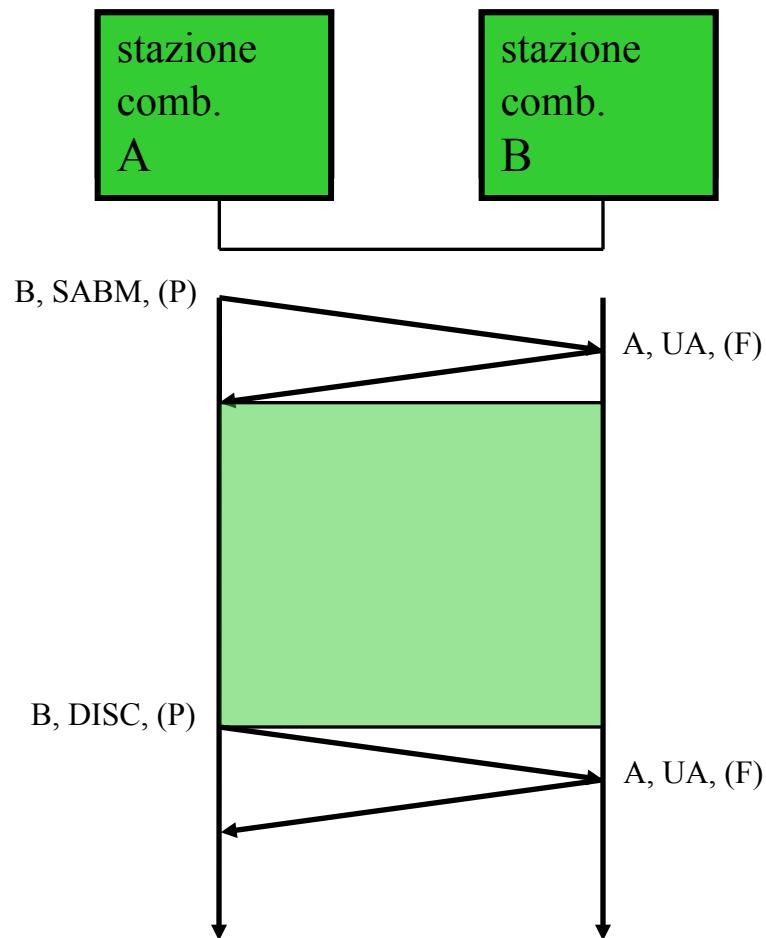
HDLC: Esempi di trasferimento dell'informazione - modalità NRM



HDLC: Esempi di trasferimento dell'informazione - modalità NRM



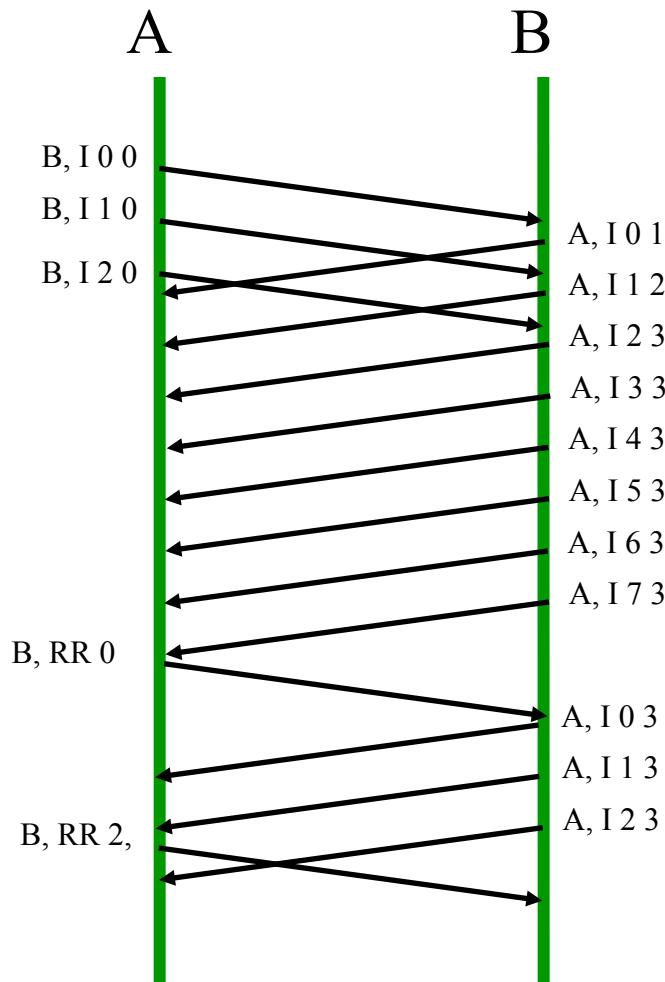
HDLC: Instaurazione della connessione - modalità ABM



HDLC: Esempi di trasferimento dell'informazione

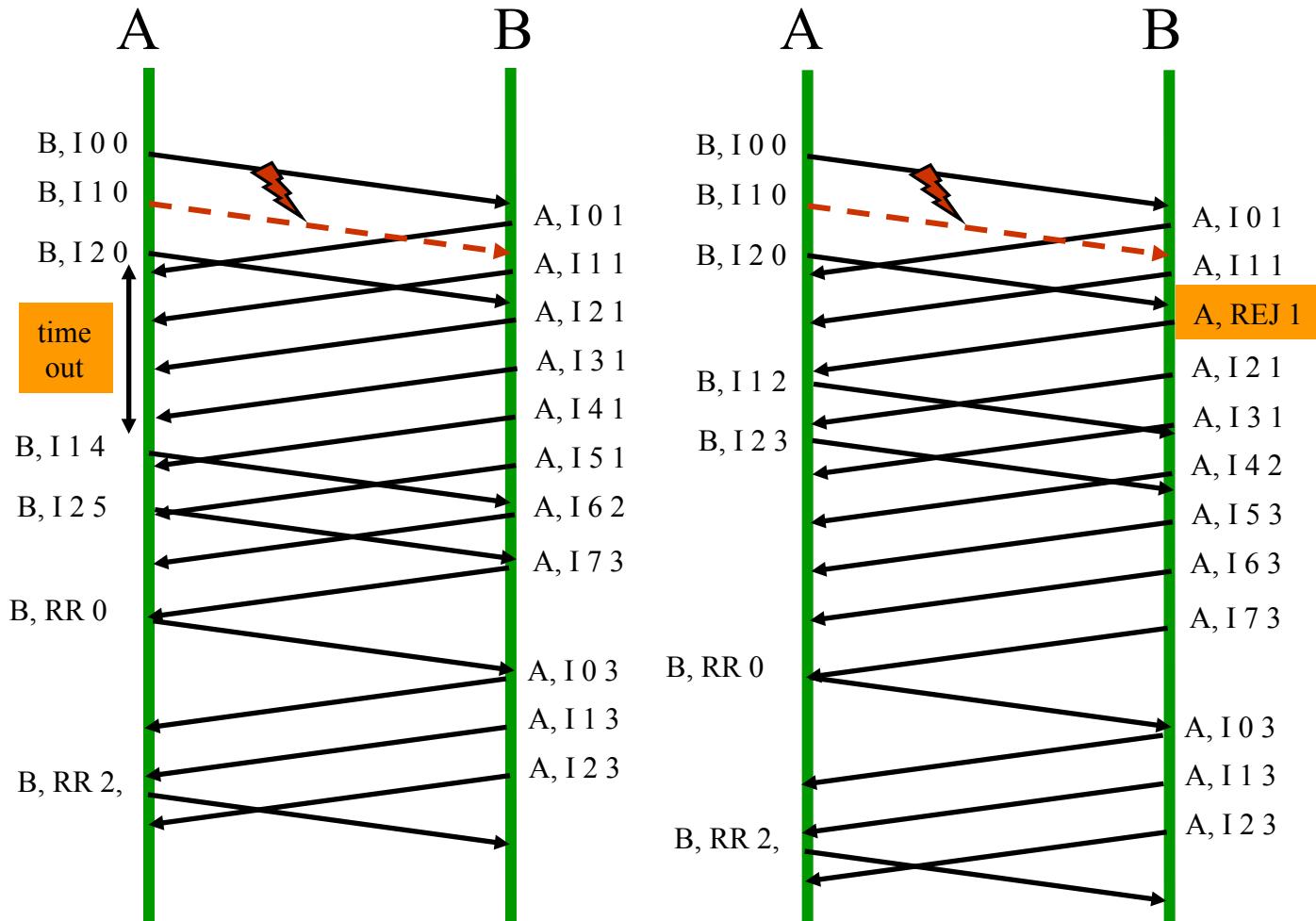
- modalità ABM

- **modalità full-duplex**



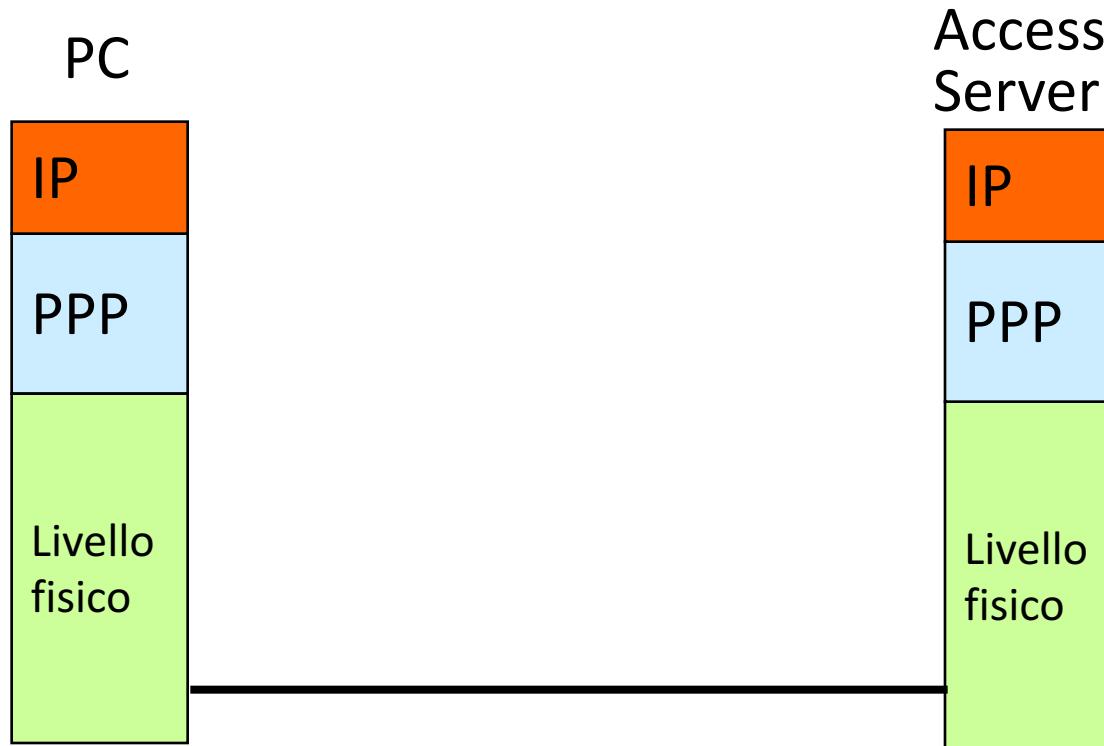
HDLC: Esempi di trasferimento dell'informazione

- modalità ABM



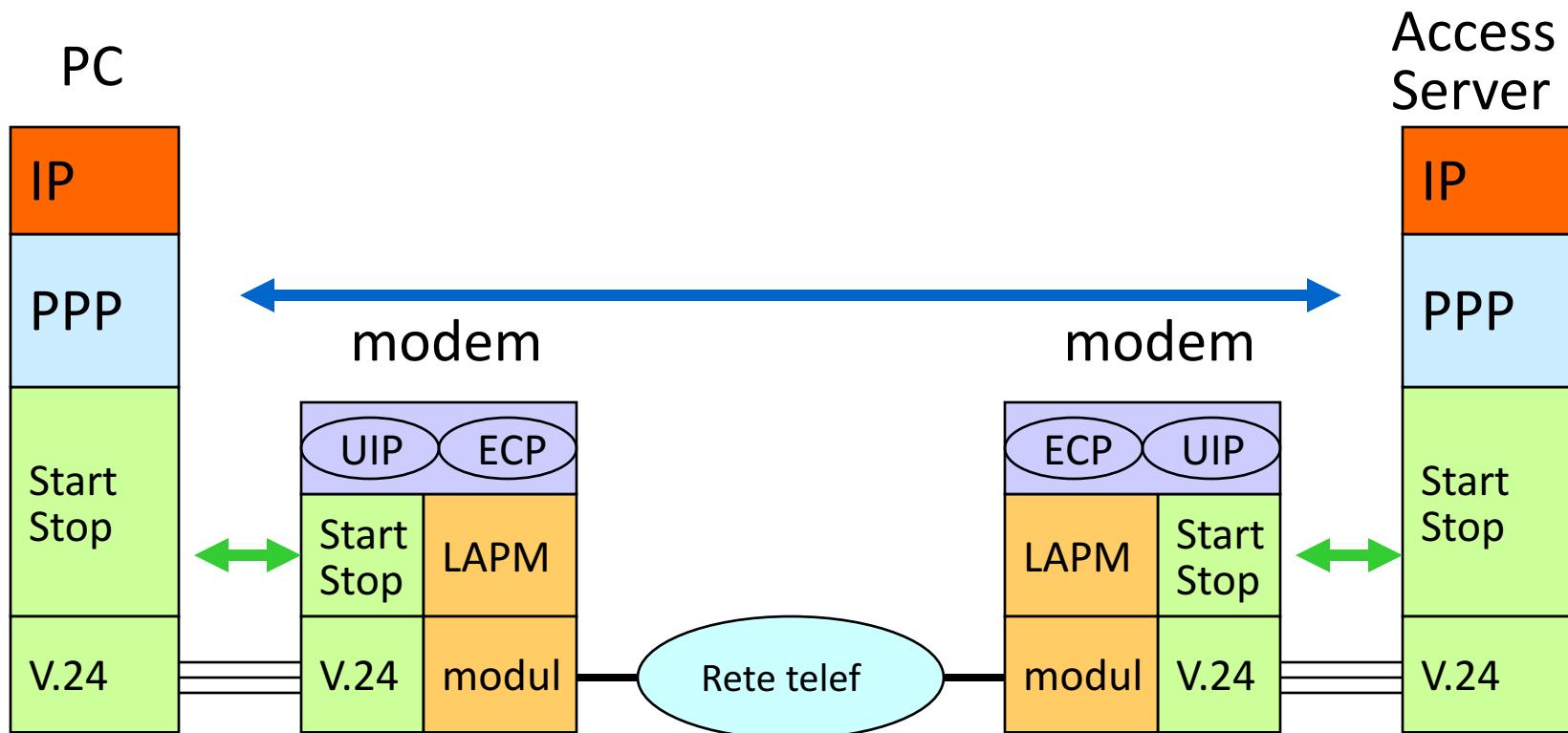
Il protocollo PPP (Point to Point Protocol)

- E' nato in ambito IETF per connessioni punto-punto su collegamenti senza errori e per consentire procedure di accesso a Internet



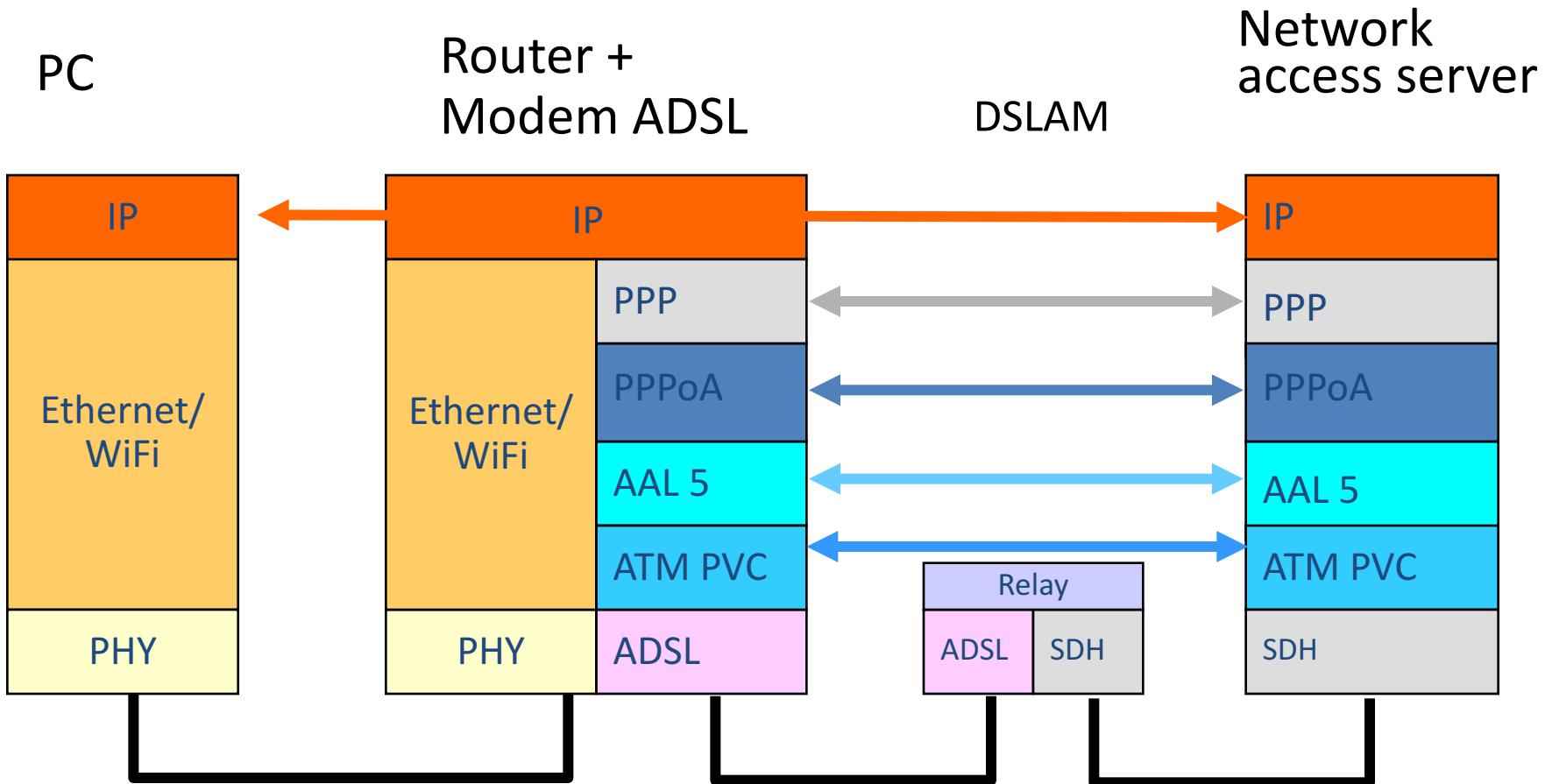
II PPP su modem

- Viene anche usato con connessioni fisiche ottenute attraverso i modem e la rete telefonica (vecchio accesso a Internet dial-up)



II PPP su ADSL

- Viene anche usato per accesso ADSL con router in modalità MPOA

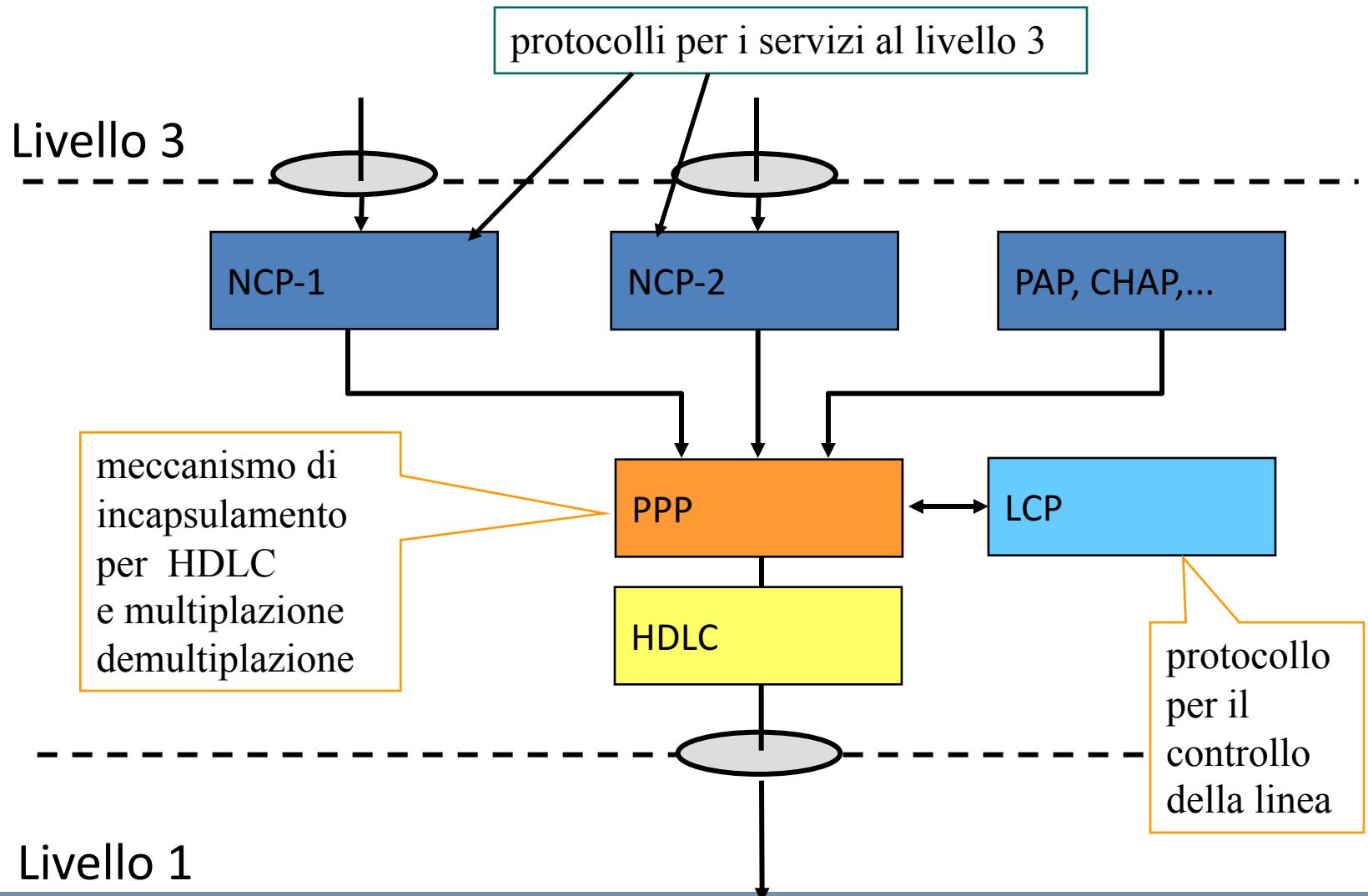


Il protocollo PPP

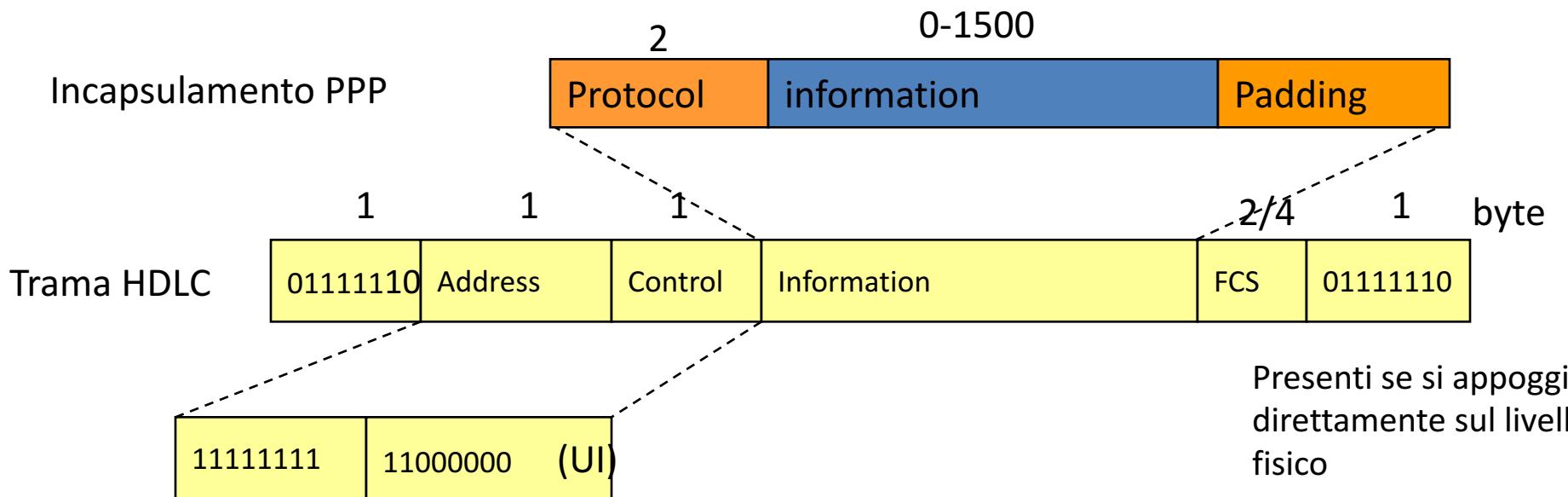
- E' in realtà un insieme di protocolli diversi che offrono supporto ai protocolli di livello 3 per effettuare la negoziazione degli indirizzi IP e l'autenticazione
- La trama e i meccanismi base sono basati direttamente su HDLC, con aggiunta di un livello (header) per multiplare vari flussi di livello superiore



PPP: architettura protocollare



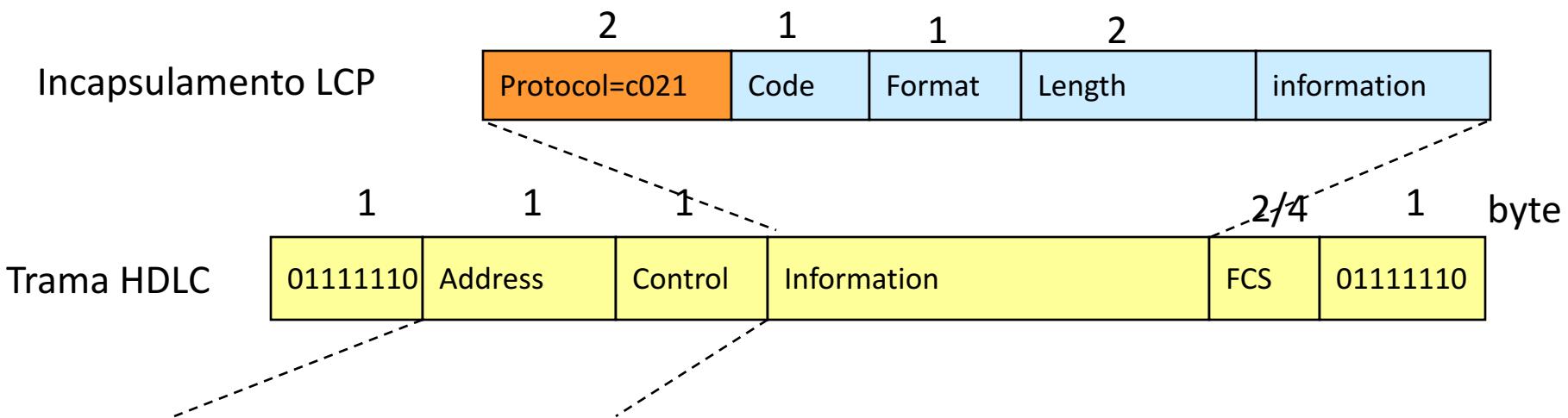
PPP-Trama e Incapsulamento



- Il campo address contiene sempre 11111111
- Il campo Control contiene sempre 11000000 ovvero l'indicazione di trama UI (datagram e niente recupero d'errore)
- Information: campo protocol + PDU degli altri protocolli



PPP - Link Control Protocol (LCP)

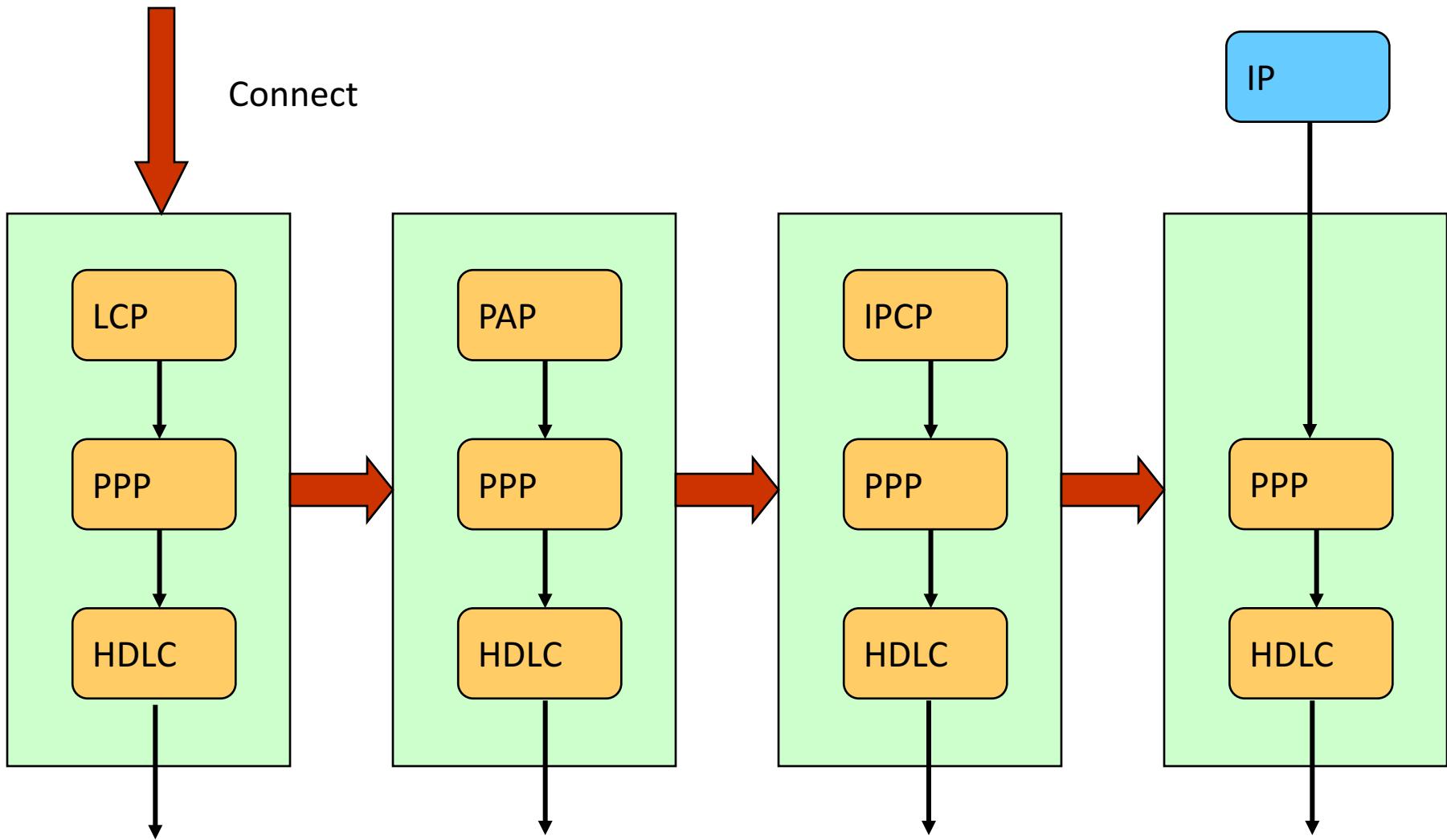


Usa i seguenti messaggi di controllo

- 1. Configure-Request
- 2. Configure-ACK
- 3. Configure-NAK
- 4. Configure-Reject
- 5. Terminate-Request
- 6. Terminate-ACK
- 7. Code-Reject
- 8. Protocol-Reject
- 9. Echo-Request
- 10. Echo-Replay
- 11. Discard-Request



PPP: uso dei diversi protocolli



PPP: protocolli ausiliari

- Dopo la connessione si hanno le seguenti possibilità
 - c023 Password Authentication Protocol
 - c025 Link Quality Report
 - c223 Challenge Handshake Authentication Protocol
- poi si passa ai NCP per i protocolli di livello 3



NCP: IP Control Protocol (IPCP)

- E' il protocollo NCP che gestisce il trasporto e il controllo di IP
- Il protocollo di controllo (protocol 8021) stabilisce una connessione in cui si decide
 - l'assegnazione dinamica dell'indirizzo IP
 - Il tipo di compressione
- Il protocollo di trasporto incapsulato:
 - 0021 IP non compresso
 - 002d TCP/IP compresso
 - 002f TCP non compresso





6 appendice – Spanning Tree Protocol

STP

Algoritmo di Spanning Tree

- Viene eletto il root bridge (la radice dello spanning tree)
- Ciascun bridge individua la root port (la porta a distanza ‘minore’ dal root bridge)
- Per ciascuna LAN si sceglie il ‘designated bridge’ di interconnessione con il root bridge. La porta di connessione del designated bridge con la LAN e’ detta ‘designated port’.
- Le root port e le designated port sono lasciate attive, mentre tutte le altre porte sono messe in uno stato di blocking la topologia logica risultante e’ un albero ricoprente.



Spanning Tree Protocol

Sistema degli identificatori e dei costi

Bridge Identifier, Port Identifier

- Ogni switch e ogni sua porta viene configurato assegnando un identificatore
- L'identificatore contiene un codice di priorità: più è basso, più è elevata la priorità
- Il bridge con ID a minima priorità diventerà il Root Bridge

Costi di attraversamento delle reti consigliati (802.1D 2004)

Table 17-3—Port Path Cost values

Link Speed	Recommended value	Recommended range	Range
<=100 Kb/s	200 000 000 ^a	20 000 000–200 000 000	1–200 000 000
1 Mb/s	20 000 000 ^a	2 000 000–200 000 000	1–200 000 000
10 Mb/s	2 000 000 ^a	200 000–20 000 000	1–200 000 000
100 Mb/s	200 000 ^a	20 000–2 000 000	1–200 000 000
1 Gb/s	20 000	2 000–200 000	1–200 000 000
10 Gb/s	2 000	200–20 000	1–200 000 000
100 Gb/s	200	20–2 000	1–200 000 000
1 Tb/s	20	2–200	1–200 000 000
10 Tb/s	2	1–20	1–200 000 000

^aBridges conformant to IEEE Std 802.1D, 1998 Edition, i.e., that support only 16-bit values for Path Cost, should use 65 535 as the Path Cost for these link speeds when used in conjunction with Bridges that support 32-bit Path Cost values.

Spanning Tree Protocol

Regole

Root identifier	Root path cost	Bridge identifier	Port identifier
Byte	8	4	8

BPDU scambiate tra bridge

- Root identifier (RI): bridge che si reputa essere il root
- Root Path Cost (RPC): costo totale delle reti attraversate nel percorso fino al root
- Bridge (Port) Identifier (RI, PI): identificatore del bridge (porta) che ha trasmesso la BPDU

Algoritmo basato su confronti tra BPDU ricevute e trasmesse

- Confronto campo-a-campo in sequenza: RI, RPC, BI, PI

- If $RI(a) \neq RI(b)$ then
 - if $RI(a) < RI(b) \rightarrow$ return [BPDU(a) > BPDU(b)] else return [BPDU(a) < BPDU(b)]
- elseif $RPC(a) \neq RPC(b)$ then
 - if $RPC(a) < RPC(b) \rightarrow$ return [BPDU(a) > BPDU(b)] else return [BPDU(a) < BPDU(b)]
- elseif $BI(a) \neq BI(b)$ then
 - if $BI(a) < BI(b) \rightarrow$ return [BPDU(a) > BPDU(b)] else return [BPDU(a) < BPDU(b)]
- else
 - if $PI(a) < PI(b) \rightarrow$ return [BPDU(a) > BPDU(b)] else return [BPDU(a) < BPDU(b)]
- end

Spanning Tree Protocol

Regole

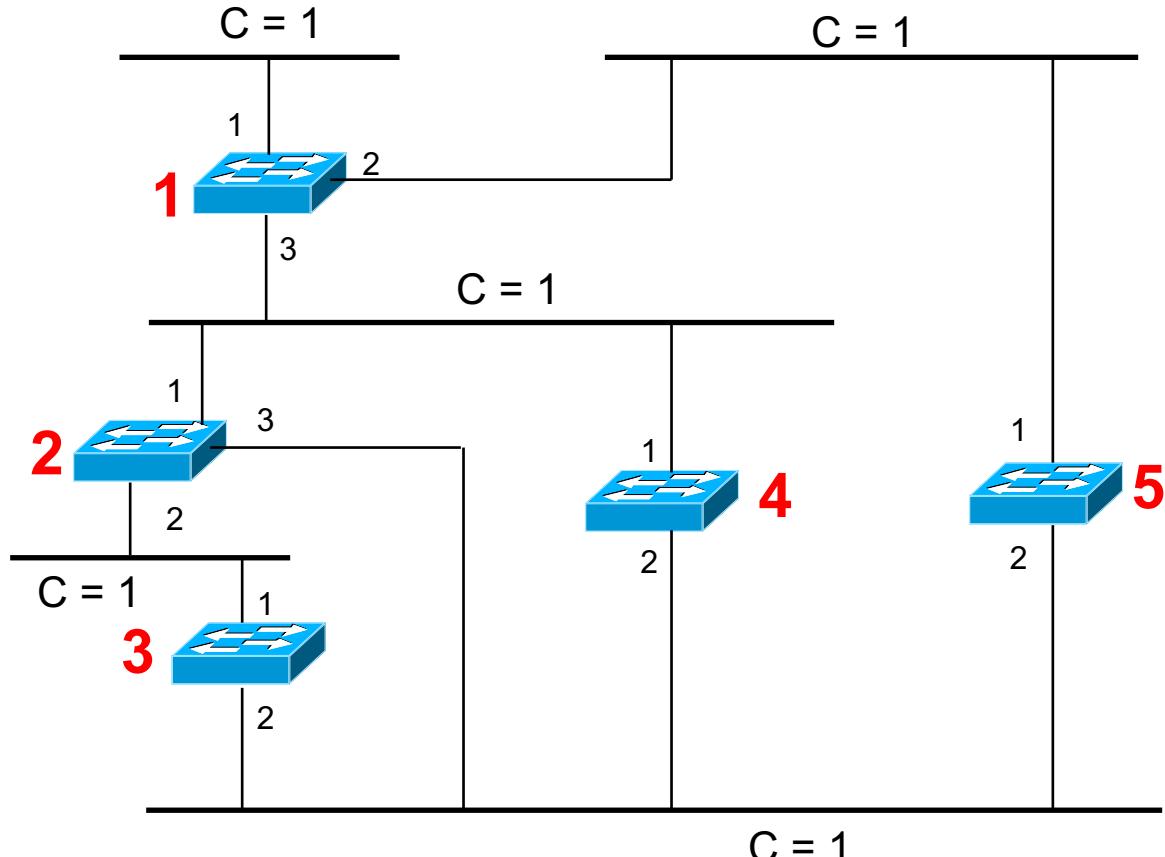
Stato di una porta p di un bridge b

- Stato Designated (D): trasmette BPDU con BI = b e PI = p
- Stato Root (R): riceve BPDU
- Stato Blocked (B): riceve BPDU

Stato di un bridge b

- Root bridge:
 - tutte le porte in stato D
 - trasmette da ciascuna porta p in stato D una BPDU con: RI = b; RPC = 0; BI = b; PI = p
- Non-root bridge:
 - una porta in stato R e le altre in stato D o B
 - tra le BPDU ricevute dalla porta in stato R, ne seleziona una con RI = x e RPC = y
 - trasmette da ciascuna porta p in stato D una BPDU con: RI = x; RPC = y + c, BI = b; PI = p
dove c è il costo della rete connessa alla porta in stato R

SPT - Esercizio 16.1



Ipotesi

- Gli ID degli switch costituiscono l'ordine di priorità: lo switch 1 sarà quello con priorità maggiore (stesso discorso vale per le porte)
- Costi unitari delle reti

Spanning Tree Protocol

Fasi dell'algoritmo

L'algoritmo opera nei seguenti 3 passi

- Root Bridge election
- Root Port selection (one per bridge)
- Designated/Blocking Port selection (one Designated per LAN)

STP – step 0: costruzione tabella

PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (1)												
Tx- BPDU												
Rx- BPDU												
Rx- BPDU cost- update												

Costruire una tabella con

- Una colonna per ogni porta (riportando porte dello stesso switch in colonne adiacenti e BI,PI in ogni colonna)
- Quattro righe in cui scrivere i campi delle BPDU trasmesse e ricevute

STP – step 0: costruzione tabella

-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2	
PORTE	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (1)												
Tx-BPDU												
Rx-BPDU												
Rx-BPDU cost-update												
-	1	1	1	1	1	1	1	1	1	1	1	1

Conviene riprodurre a margine della tabella la topologia della rete, riportando per ogni porta

- Tutte le porte connesse
- Il costo della rete connessa

STP – step 1: Root Bridge Election

-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2	
PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (1)												
Tx-BPDU												
Rx-BPDU												
Rx-BPDU cost-update												
	-	1	1	1	1	1	1	1	1	1	1	1

Ciascun bridge inizialmente si pone in stato root

Vengono inviate le trame BPDU

Si riportano tutte le BPDU ricevute da ciascuna porta

STP – step 1: Root Bridge Election

-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2	
PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (1)	D	D	D	D	D	D	D	D	D	D	D	D
Tx-BPDU	1,0,1,1	1,0,1,2	1,0,1,3	2,0,2,1	2,0,2,2	2,0,2,3	3,0,3,1	3,0,3,2	4,0,4,1	4,0,4,2	5,0,5,1	5,0,5,2
Rx-BPDU	-	5,0,5,1	2,0,2,1 4,0,4,1	4,0,4,1 1,0,1,3	3,0,3,1	3,0,3,2	2,0,2,2	2,0,2,3	1,0,1,3	2,0,2,3 3,0,3,2 5,0,5,2	1,0,1,2	2,0,2,3 3,0,3,2 4,0,4,2
Rx-BPDU cost-update												
	-	1	1	1	1	1	1	1	1	1	1	1

L'aggiornamento del costo si ottiene aggiungendo al Root Path Cost della BPDU ricevuta da una determinata porta il corso della rete connessa alla porta stessa

Il valore di RPC aggiornato si utilizza per compilare l'ultima riga

STP – step 1: Root Bridge Election

-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2	
PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (1)	D	D	D	D	D	D	D	D	D	D	D	D
Tx-BPDU	1,0,1,1	1,0,1,2	1,0,1,3	2,0,2,1	2,0,2,2	2,0,2,3	3,0,3,1	3,0,3,2	4,0,4,1	4,0,4,2	5,0,5,1	5,0,5,2
Rx-BPDU	-	5,0,5,1	2,0,2,1 4,0,4,1	4,0,4,1 1,0,1,3	3,0,3,1	3,0,3,2 4,0,4,2 5,0,5,2	2,0,2,2 4,0,4,2 5,0,5,2	2,0,2,3 2,0,2,1 5,0,5,2	1,0,1,3 2,0,2,1 3,0,3,2 5,0,5,2	2,0,2,3 3,0,3,2 5,0,5,2	1,0,1,2 3,0,3,2 4,0,4,2	2,0,2,3 3,0,3,2 4,0,4,2
Rx-BPDU cost-update	-	5,1,5,1	2,1,2,1 4,1,4,1	4,1,4,1 1,1,1,3	3,1,3,1	3,1,3,2 4,1,4,2 5,1,5,2	2,1,2,2 4,1,4,2 5,1,5,2	2,1,2,3 4,1,4,2 5,1,5,2	1,1,1,3 2,1,2,1 3,1,3,2 5,1,5,2	2,1,2,3 3,1,3,2 5,1,5,2	1,1,1,2 3,1,3,2 4,1,4,2	2,1,2,3 3,1,3,2 4,1,4,2
	-	1	1	1	1	1	1	1	1	1	1	1

Per l'elezione del Root Bridge si confronta in ogni bridge il BI con i RI delle BPDU ricevute con RPC aggiornato.
 Ogni bridge

- Continua a ritenersi root se $BI < RI \forall$ BPDU ricevuta (es. bridge 1)
- Passa in stato non-root se \exists una BPDU ricevuta | $BI \geq RI$

NOTA: temporaneamente ci possono essere anche più bridge che rimangono root, anche se a convergenza ce ne sarà uno solo

STP – step 2: Root Port Selection

-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2	
PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (1)	D	D	D	D	D	D	D	D	D	D	D	D
Tx-BPDU	1,0,1,1	1,0,1,2	1,0,1,3	2,0,2,1	2,0,2,2	2,0,2,3	3,0,3,1	3,0,3,2	4,0,4,1	4,0,4,2	5,0,5,1	5,0,5,2
Rx-BPDU	-	5,0,5,1	2,0,2,1 4,0,4,1	4,0,4,1 1,0,1,3	3,0,3,1	3,0,3,2 4,0,4,2 5,0,5,2	2,0,2,2	2,0,2,3 4,0,4,2 5,0,5,2	1,0,1,3 2,0,2,1 3,0,3,2 5,0,5,2	2,0,2,3 3,0,3,2 5,0,5,2	1,0,1,2	2,0,2,3 3,0,3,2 4,0,4,2
Rx-BPDU cost-update	-	5,1,5,1	2,1,2,1 4,1,4,1	4,1,4,1 1,1,1,3	3,1,3,1	3,1,3,2 4,1,4,2 5,1,5,2	2,1,2,2	2,1,2,3 4,1,4,2 5,1,5,2	1,1,1,3 2,1,2,1 3,1,3,2 5,1,5,2	2,1,2,3 3,1,3,2 5,1,5,2	1,1,1,2	2,1,2,3 3,1,3,2 4,1,4,2
	-	1	1	1	1	1	1	1	1	1	1	1

Ogni bridge non-root deve eleggere la propria Root port

- Root port → porta con Rx BPDU (aggiornata) a priorità maggiore tra tutte le BPDU (aggiornate) ricevute

STP – step 2: Root Port Election

-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2	
PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (2)	D	D	D	R	D	D	R	D	R	D	R	D
Tx-BPDU												
Rx-BPDU												
Rx-BPDU cost-update												
	-	1	1	1	1	1	1	1	1	1	1	1

Aggiornamento dello stato delle porte

Nei non-root bridge:

- Le porte R non trasmettono più
- Le porte D ritrasmettono le BPDU ricevute dalle porte R allo step precedente

I root bridge continuano a comportarsi come prima

STP – step 2: Root Port Election

-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2	
PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (2)	D	D	D	R	D	D	R	D	R	D	R	D
Tx-BPDU	1,0,1,1	1,0,1,2	1,0,1,3	-	1,1,2,2	1,1,2,3	-	2,1,3,2	-	1,1,4,2	-	1,1,5,2
Rx-BPDU	-	-	-	1,0,1,3	-	2,1,3,2 1,1,4,2 1,1,5,2	1,1,2,2 1,1,4,2 1,1,5,2	1,1,2,3 1,1,4,2 1,1,5,2	1,0,1,3	1,1,2,3 2,1,3,2 1,1,5,2	1,0,1,2	1,1,2,3 2,1,3,2 1,1,4,2
Rx-BPDU cost-update	-	-	-	1.1.1.3	-	2,2,3,2 1,2,4,2 1,2,5,2	1.2.2.2	1,2,2,3 1,2,4,2 1,2,5,2	1.1.1.3	1,2,2,3 2,2,3,2 1,2,5,2	1.1.1.2	1,2,2,3 2,2,3,2 1,2,4,2
	-	1	1	1	1	1	1	1	1	1	1	1

Nota: «nessuna BPDU» ha sempre priorità minore di «qualsiasi BPDU»

Confronto nei bridge delle Rx BPDU con il costo aggiornato tra di loro e con il BI per confermare lo stato delle porte

Lo step 2 deve durare un po' di iterazioni in modo che gli stati delle porte si stabilizzino

STP – step 3: Blocked/Designated Ports

	-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2
PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (3)	D	D	D	R	D	D	R	D	R	D	R	D
Tx-BPDU	1,0,1,1	1,0,1,2	1,0,1,3	-	1,1,2,2	1,1,2,3	-	1,2,3,2	-	1,1,4,2	-	1,1,5,2
Rx-BPDU	-	-	-	1,0,1,3	-	1,2,3,2 1,1,4,2 1,1,5,2	1,1,2,2 1,1,4,2 1,1,5,2	1,1,2,3	1,0,1,3 1,2,3,2 1,1,5,2	1,1,2,3	1,0,1,2 1,2,3,2	1,1,2,3 1,1,4,2
Rx-BPDU cost-update	-	-	-	1,1,1,3	-	1,3,3,2 1,2,4,2 1,2,5,2	1,2,2,2 1,2,4,2 1,2,5,2	1,2,2,3 1,2,4,2 1,2,5,2	1,1,1,3	1,2,2,3 1,3,3,2 1,2,5,2	1,1,1,2	1,2,2,3 1,3,3,2 1,2,4,2
	-	1	1	1	1	1	1	1	1	1	1	1

Selezione porte D/B: confronto di Tx BPDU con Rx BPDU con il costo NON aggiornato della porta attraversata

- Se Tx ha priorità più alta → la porta si pone in stato D
- Se Rx aggiornato ha priorità più alta → la porta si pone in stato B

STP – step 3: Blocked/Designated Ports

	-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2
PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (3)	D	D	D	R	D	D	R	D	R	D	R	D
Tx-BPDU	1,0,1,1	1,0,1,2	1,0,1,3	-	1,1,2,2	1,1,2,3	-	1,2,3,2	-	1,1,4,2	-	1,1,5,2
Rx-BPDU	-	-	-	1,0,1,3	-	1,2,3,2 1,1,4,2 1,1,5,2	1,1,2,2 1,1,4,2 1,1,5,2	1,1,2,3	1,0,1,3 1,2,3,2 1,1,5,2	1,1,2,3	1,0,1,2 1,2,3,2	1,1,2,3 1,1,4,2
Rx-BPDU cost-update	-	-	-	1,1,1,3	-	1,3,3,2 1,2,4,2 1,2,5,2	1,2,2,2	1,2,2,3 1,2,4,2 1,2,5,2	1,1,1,3	1,2,2,3 1,3,3,2 1,2,5,2	1,1,1,2	1,2,2,3 1,3,3,2 1,2,4,2
	-	1	1	1	1	1	1	1	1	1	1	1

La selezione porte D/B si può fare solo in uno step del protocollo in cui le porte R non sono variate rispetto lo step precedente

- In questo caso si poteva fare anche allo step 2

STP – step 3: Blocked/Designated Ports

-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2	
PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (4)	D	D	D	R	D	D	R	B	R	B	R	B
Tx-BPDU	1,0,1,1	1,0,1,2	1,0,1,3	-	1,1,2,2	1,1,2,3	-	-	-	-	-	-
Rx-BPDU							!	!				
Rx-BPDU cost-update							!	!				
-	1	1	1	1	1	1	1	1	1	1	1	1

- Le porte in stato B smettono di trasmettere
- Le porte in stato D continuano a trasmettere

STP – step 3: Blocked/Designated Ports

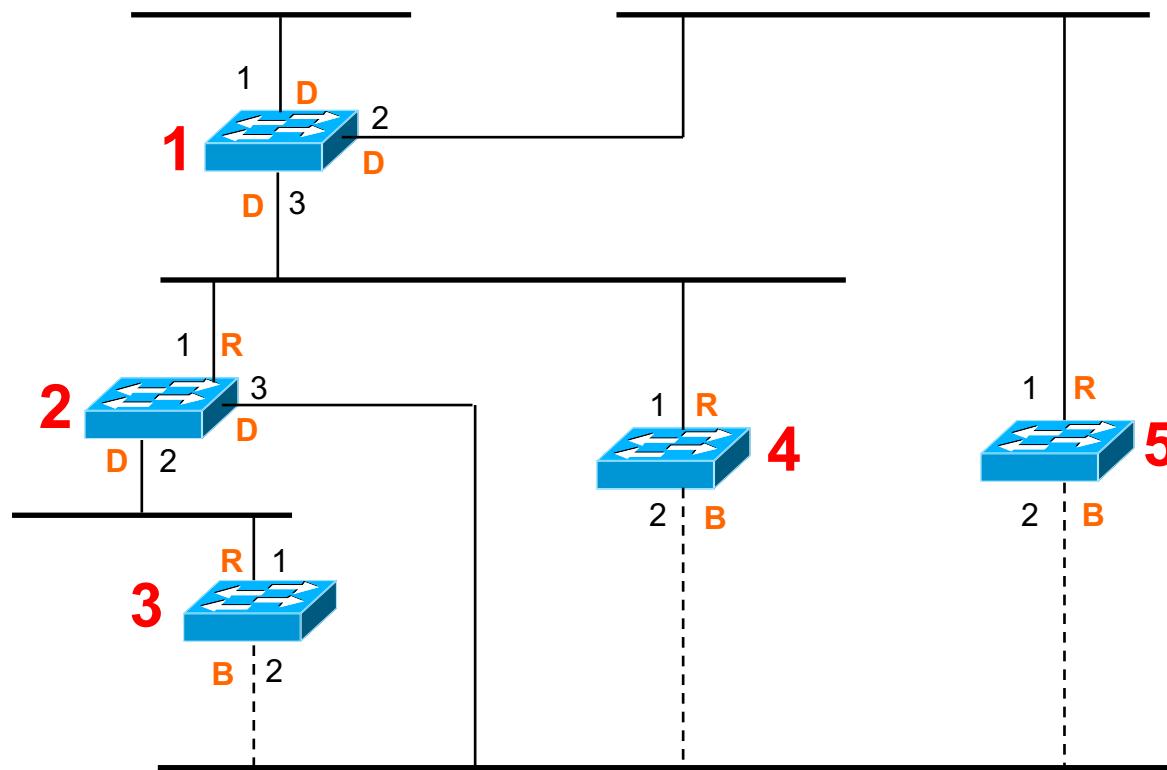
-	5,1	2,1 4,1	1,3 4,1	3,1	3,2 4,2 5,2	2,2	2,3 4,2 5,2	1,3 2,1	2,3 3,2 5,2	1,2	2,3 3,2 4,2	
PORTA	1,1	1,2	1,3	2,1	2,2	2,3	3,1	3,2	4,1	4,2	5,1	5,2
STATO (4)	D	D	D	R	D	D	R	B	R	B	R	B
Tx-BPDU	1,0,1,1	1,0,1,2	1,0,1,3	-	1,1,2,2	1,1,2,3	-	-	-	-	-	-
Rx-BPDU	-	-	-	1,0,1,3	-	-	1,1,2,2	1,1,2,3	1,0,1,3	1,1,2,3	1,0,1,2	1,1,2,3
Rx-BPDU cost-update	-	-	-	1,1,1,3	-	-	1,2,2,2	1,2,2,3	1,1,1,3	1,2,2,3	1,1,1,2	1,2,2,3
	-	1	1	1	1	1	1	1	1	1	1	1

- Una volta arrivato a convergenza, gli step si ripetono identici fino a quando la topologia della rete non varia
- Attenzione! Attraverso le porte in stato B non passano trame dati, ma le porte B continuano a ricevere le BPDU
 - Altrimenti la disconnessione di un bridge non sarebbe rilevata e porterebbe all'isolamento di altri bridge

Se una porta B non riceve nulla, passa in stato D

Esercizio

Soluzione



Selezione della Root Port

- Una volta completata l'elezione del Root Bridge, ciascun Bridge identifica la sua porta ‘piu’ vicina’ al Root Bridge come Root Port.
- La distanza e’ espressa in termini di costo tramite il parametro Root Path Cost, e, a parita’ di costo dei diversi link (situazione comune nelle reti locali) corrisponde al numero di hop attraversati



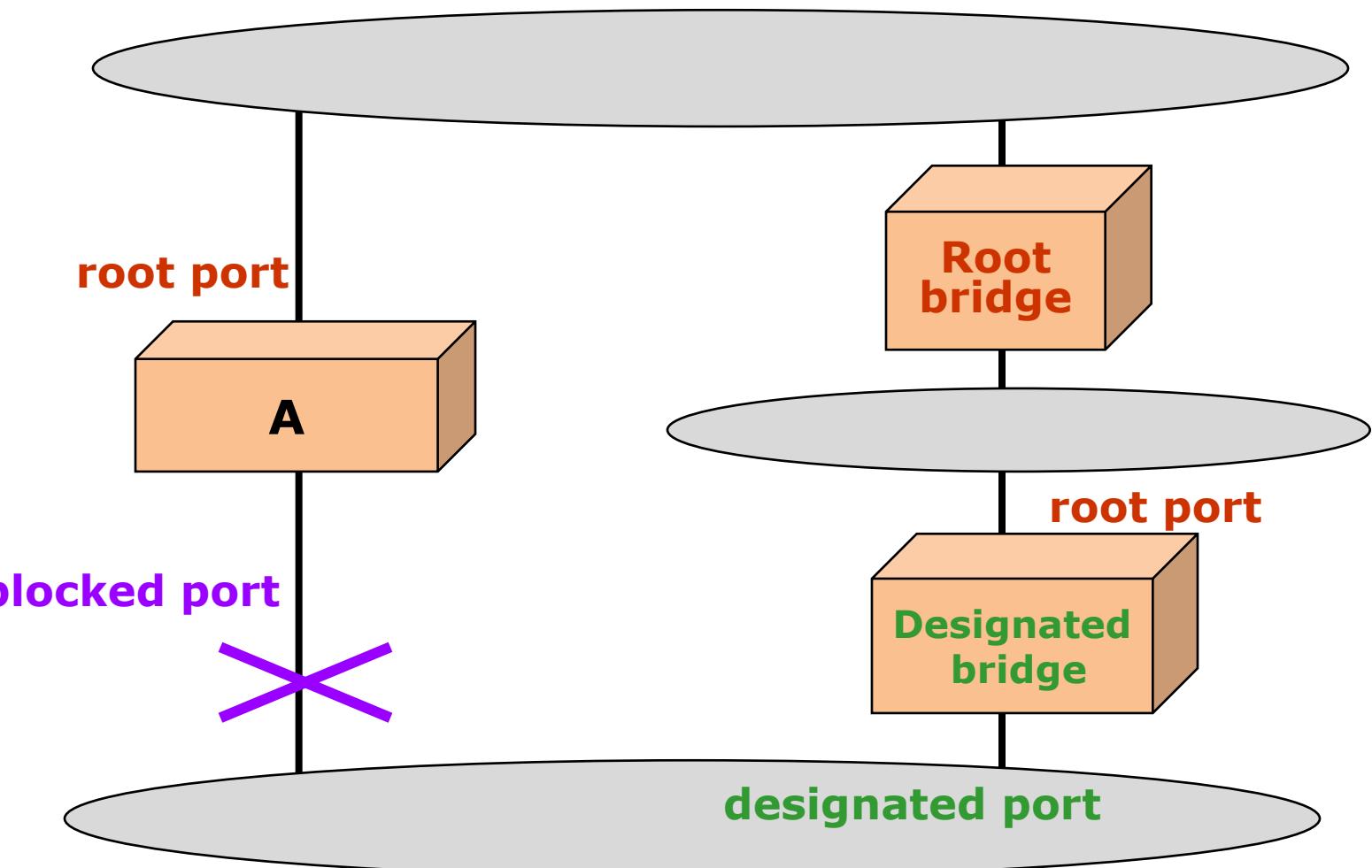
Selezione della Designated Bridge Port

- Su ciascuno dei segmenti di LAN a cui siano connessi più di un Bridge viene eletto un Designated Bridge incaricato di inoltrare le trame nella direzione del root Bridge
- La porta tramite cui il Designated Bridge è connesso alla LAN prende il nome di Designated Bridge Port.
- Viene scelto come Designated Bridge il Bridge a distanza minima dal Root Bridge e, a parità di distanza, il Bridge con minor Bridge ID.

Le porte del Root Bridge sono Designated Bridge Ports !

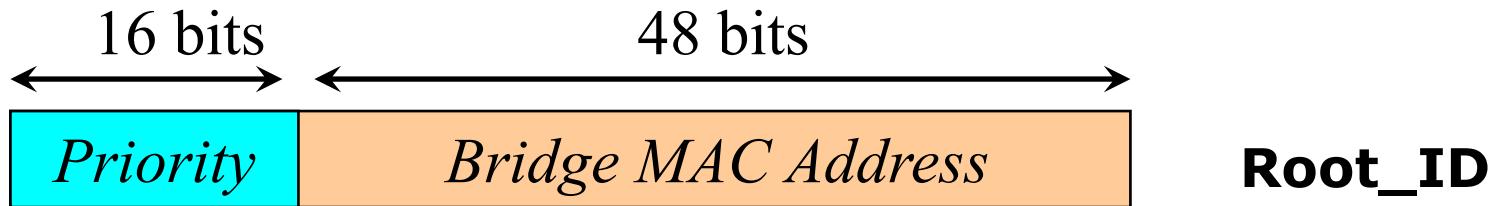


Meccanismo di definizione dell'albero



Selezione del Root Bridge

- Il processo inizia con ciascun switch (bridge) che considera se stesso come root
- Trame di segnalazione del protocollo STP con il `root_ID` sono inviate su tutte le porte
- Quando una trama è ricevuta se il `root_ID` contenuto è più piccolo di quello memorizzato, questo viene aggiornato ed una nuova trama viene trasmessa
- Alla fine del processo è eletto root lo switch con ID più piccolo

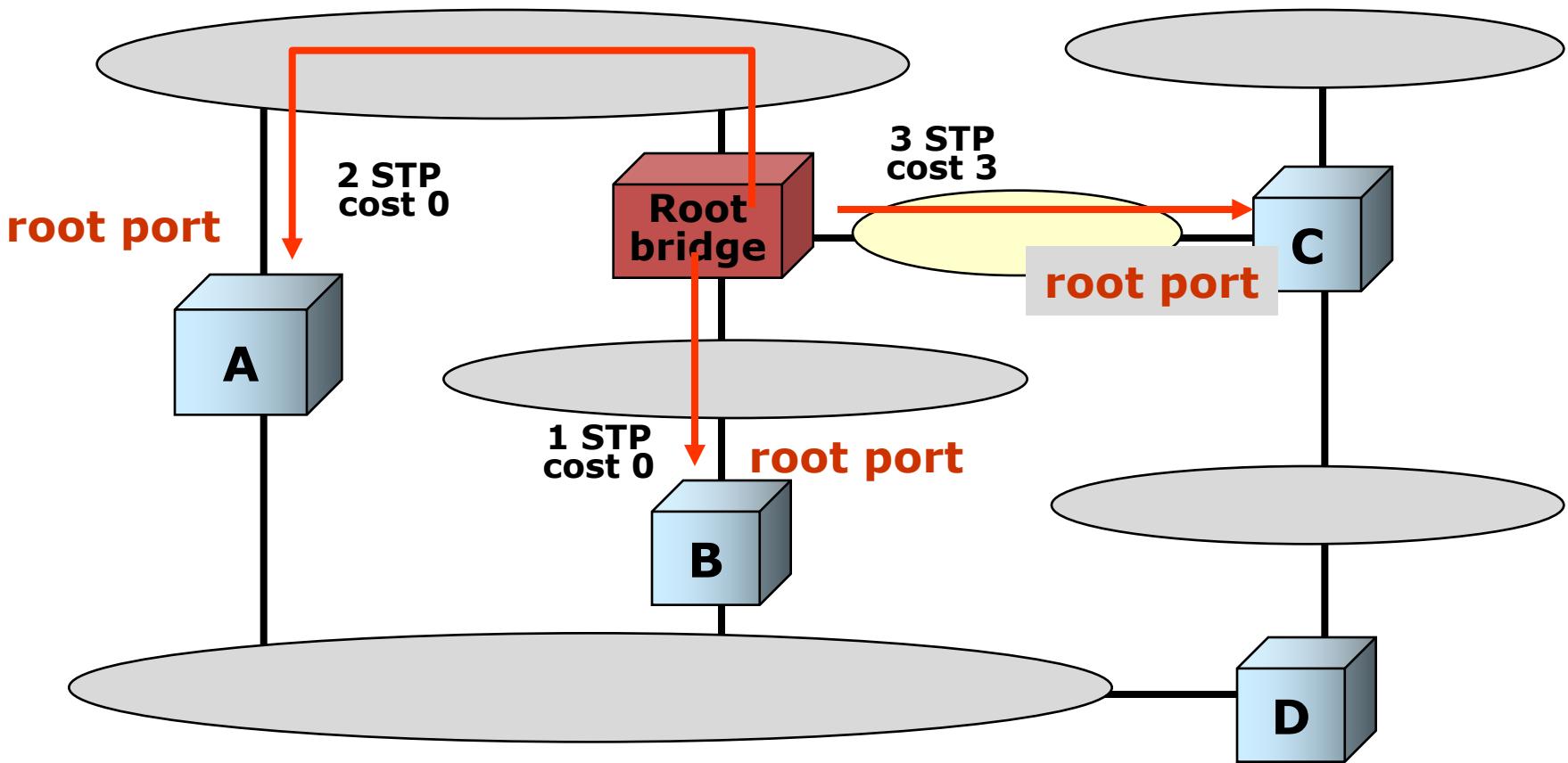


Selezione della Root Port

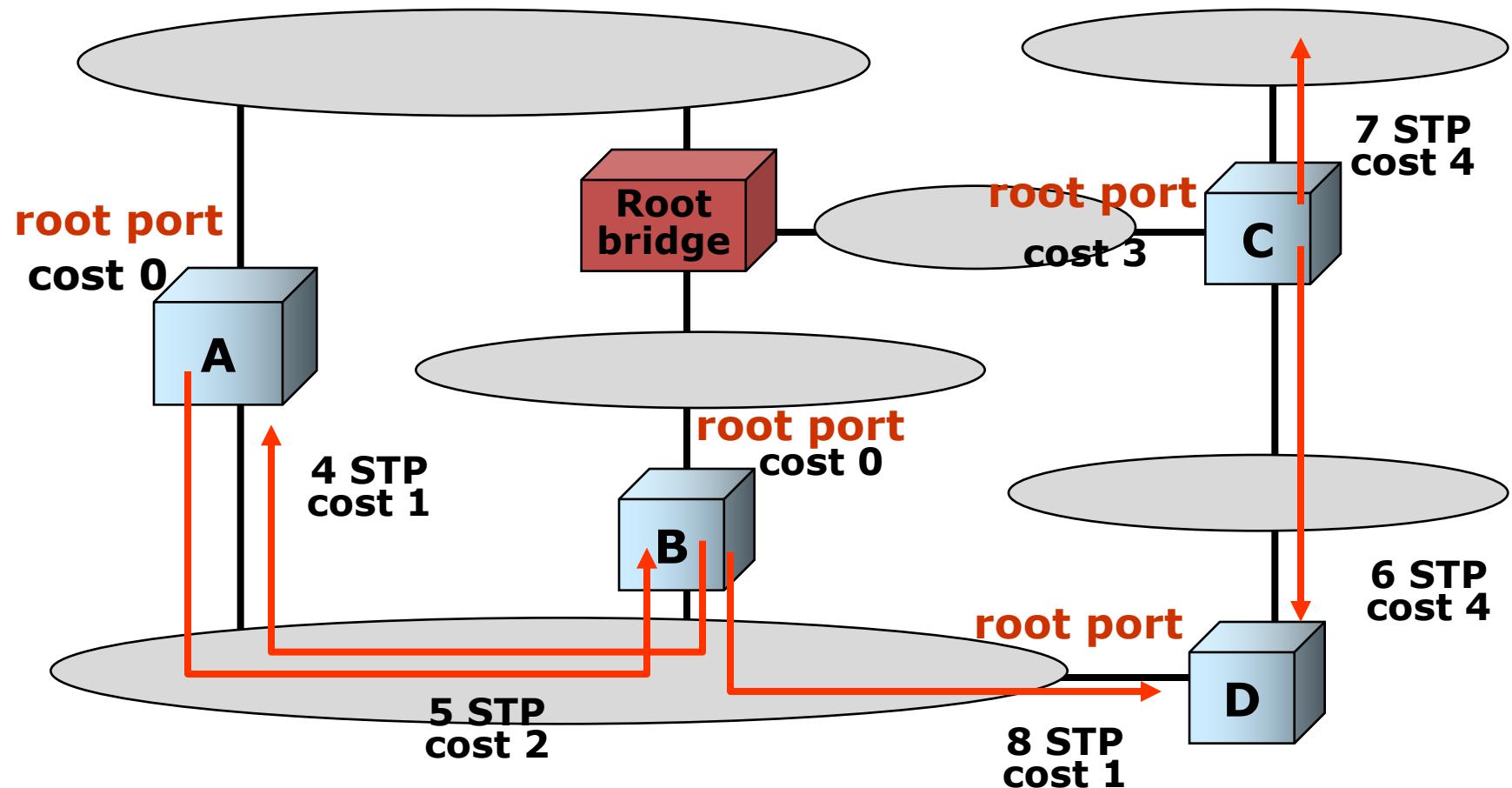
- Ogni volta che una trama STP è ricevuta si calcola il costo corrispondente sommando il costo indicato nella trama con quello associato alla porta di ricezione
- Il costo (Root Path Cost) più basso calcolato è memorizzato e incluso nelle trame STP trasmesse
- La porta corrispondente al costo più basso è selezionata come Root Port
- Sia il Root Path Cost che la Root Port sono aggiornata quando altre trame STP sono ricevute
- Il processo di elezione della root e calcolo di costo e root port sono eseguiti in parallelo



Esempio: root port



Esempio: root port

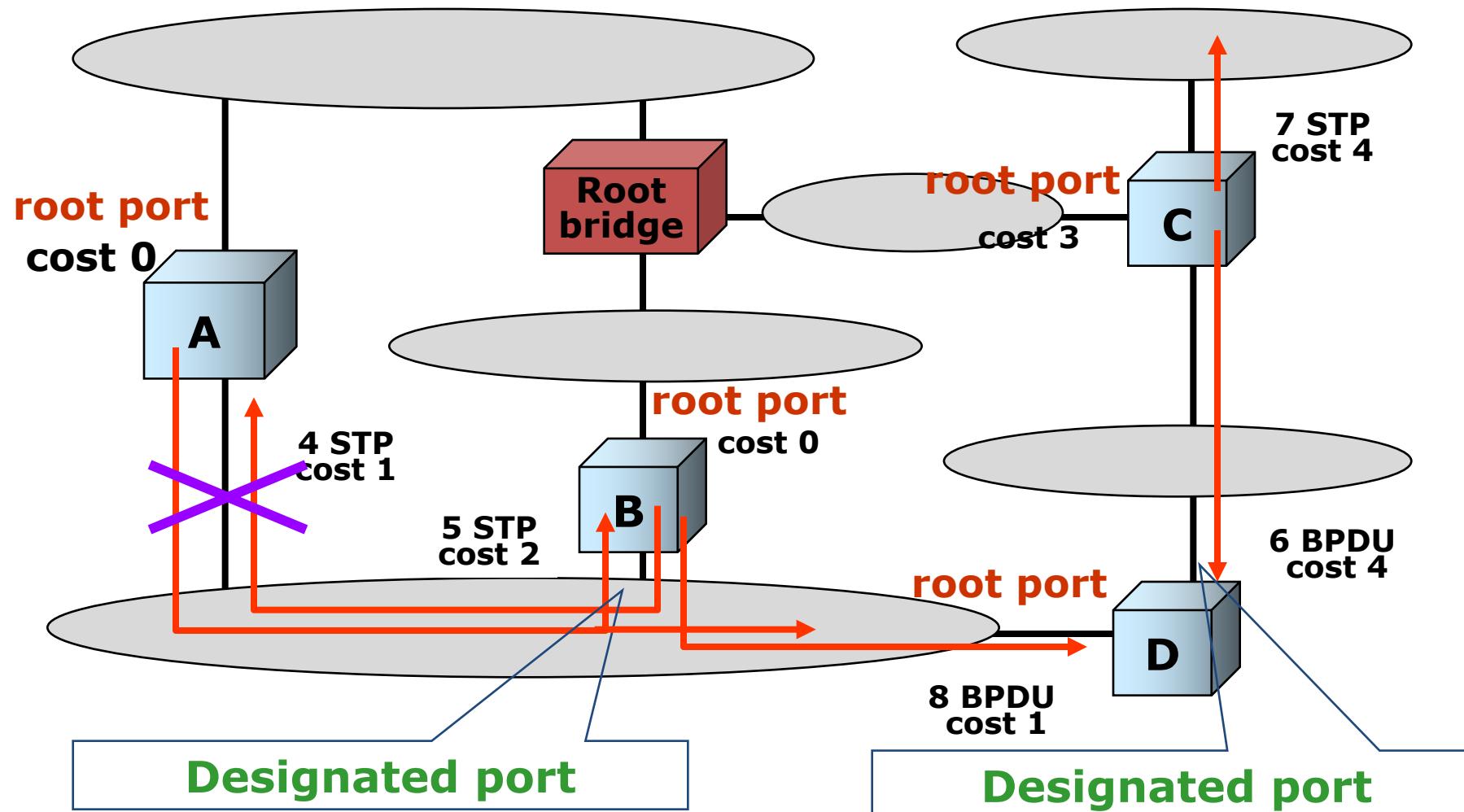


Selezione della Designated Port

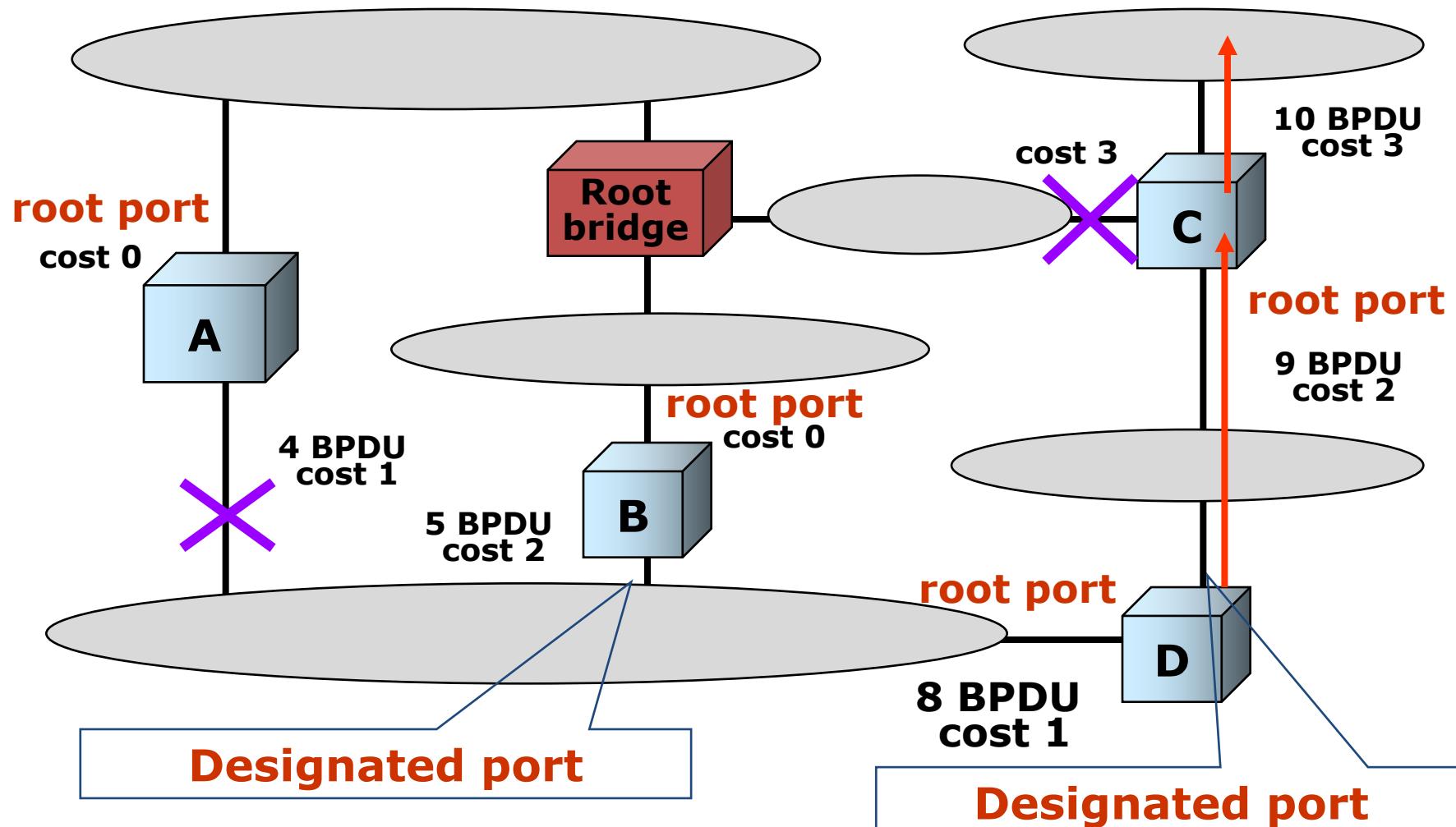
- Durante lo scambio di trame STP, ciascuno switch può apprendere la distanza dalla root degli switch vicini sullo stesso segmento di LAN (dominio di collisione)
- Se la sua distanza è la minore la corrispondente non-root port è selezionata come Designated Port
- In caso di uguale distanza si considera ID più piccolo
- Tutte le root port sono anche designated
- Tutte le altre sono blocked



Esempio: designed port



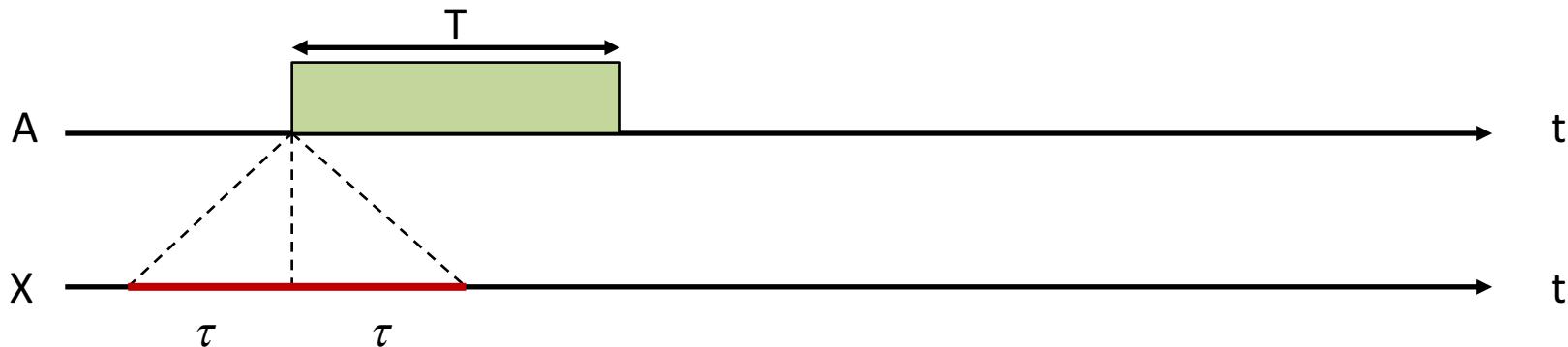
Esempio: designated port





Carrier Sense Multiple Access (CSMA)

- Risultato: esiste un periodo di vulnerabilità pari a due volte il massimo ritardo di propagazione τ



- L'efficienza del meccanismo dipende dal rapporto
$$a = \tau/T$$
- Se $a > 1$, il primo bit arriva quando la trasmissione è già finita e ascoltare il canale non serve a niente (meglio Aloha)
- Se $a \ll 1$ allora l'efficienza del CSMA può essere elevata

