

1 STP: IL PROTOCOLLO DI COMUNICAZIONE TRA GLI SWITCH

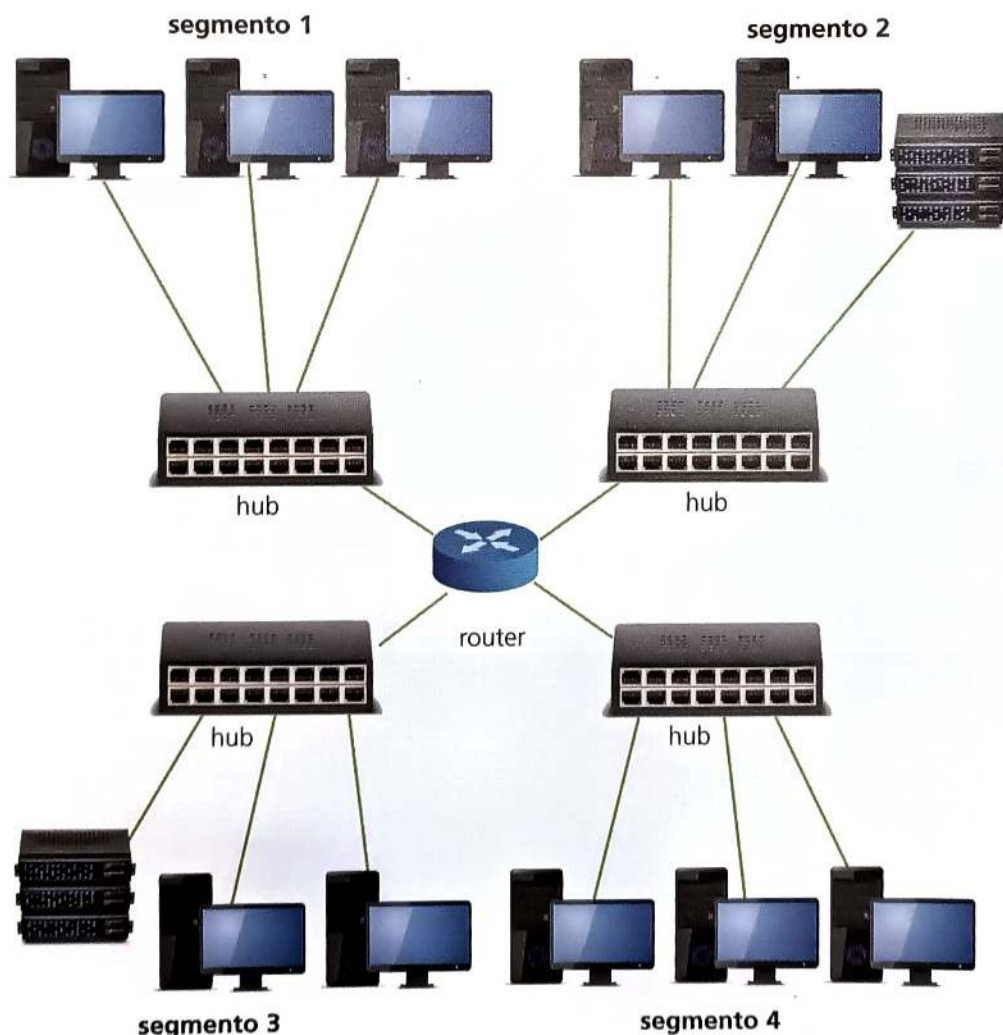
1.1 Reti locali "segmentate"

Le moderne reti locali sono "segmentate", cioè suddivise in parti più piccole, dette **segmenti**, tramite **switch** (o bridge, ma nel seguito parleremo solo di switch) al fine di isolare il traffico tra i segmenti e raggiungere una maggiore ampiezza di banda per ogni computer grazie alla creazione di domini di collisione più piccoli.

Il **dominio di collisione** di una rete è un'area in cui può verificarsi una collisione. Per esempio, se abbiamo 5 computer connessi allo stesso mezzo condiviso, i dati inviati da uno di essi possono collidere con i dati inviati da un altro. In questo caso abbiamo un dominio di collisione che contiene 5 host.

Oltre alle LAN segmentate con switch, è possibile creare segmenti di LAN utilizzando i **router** (FIGURA 1). Questa soluzione rende più lenta la trasmissione rispetto alla soluzione con gli switch (a meno di usare router che implementino funzionalità di switching).

FIGURA 1 LAN segmentata con un router in 4 segmenti



Infatti il router è un apparato che opera a livello Network, basa le sue decisioni sugli indirizzi di rete (IP address) e non su quelli fisici (MAC address) e implementa algoritmi per trovare il percorso migliore che richiedono più tempo di elaborazione. Spesso le reti con switch sono progettate con **percorsi fisici ridondanti** al fine di evitare che il guasto di un cavo o di una porta di un apparato possa portare al blocco delle trasmissioni sulla rete.

Se da un lato la duplicazione dei percorsi offre maggiori garanzie in termini di affidabilità e fault tolerance, dall'altro può dar luogo a effetti indesiderati (side effects), come la creazione di loop che portano al fenomeno detto "broadcast storm" che può in breve tempo bloccare la rete.

Un **broadcast storm** (letteralmente "tempesta di broadcast") avviene quando ci sono così tanti frame broadcast in un loop da impegnare tutta la banda disponibile. Un broadcast storm è inevitabile su una rete con percorsi ridondanti (FIGURA 2).

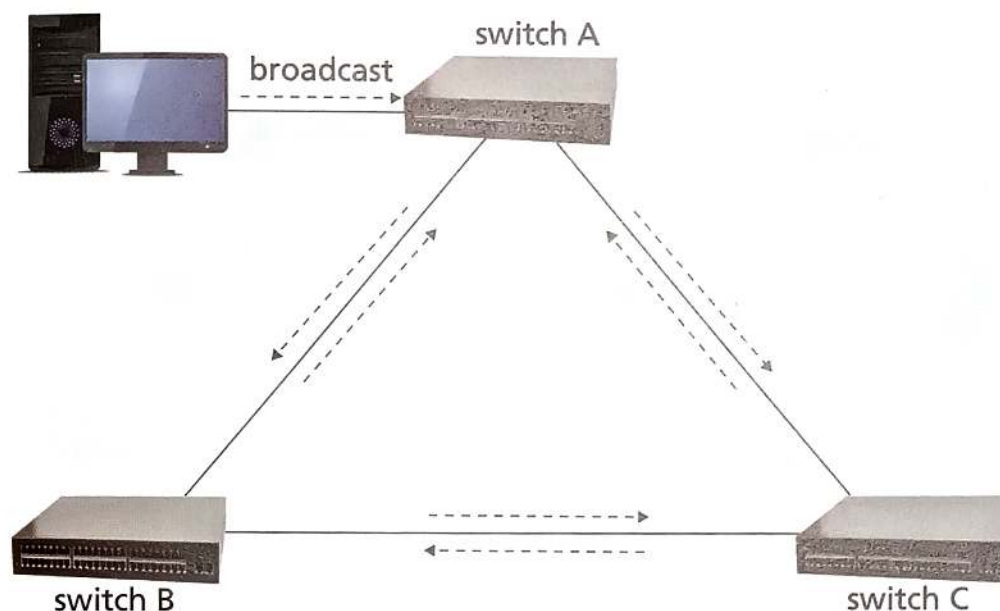


FIGURA 2 Loop con broadcast storm

Per evitare questi loop gli switch usano un protocollo per la gestione dei collegamenti, denominato **#STP** (Spanning Tree Protocol).

1.2 Spanning Tree Protocol

STP è un protocollo definito nello standard IEEE 802.1 per realizzare reti LAN complesse senza loop. Con STP si crea un albero gerarchico che mantiene ancora disponibili i percorsi alternativi da usare in caso di necessità, quindi si lasciano loop fisici ma si eliminano a livello di topologia logica.

Infatti, una volta rilevato che esistono più percorsi tra i nodi della rete (loop), il protocollo STP crea una struttura ad albero relegando i percorsi ridondanti a uno stato di standby (*blocked*) della relativa porta dello switch.

STP permette che venga stabilito un solo percorso attivo alla volta tra due dispositivi della rete per evitare i loop, tuttavia stabilisce collegamenti ridondanti come alternative nel caso in cui il collegamento primario dovesse non essere più disponibile. Se in un certo istante un segmento della rete dovesse diventare irraggiungibile, l'algoritmo di Spanning Tree riconfigurerà la topologia logica, ristabilendo il

#techwords

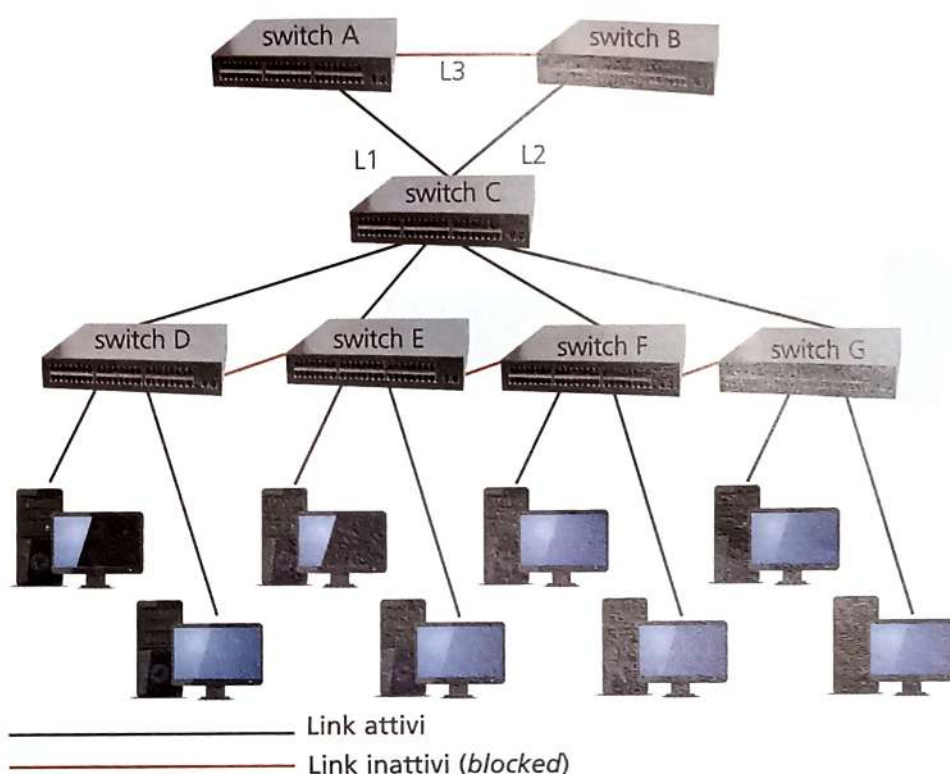
STP (Spanning Tree Protocol) è così denominato in quanto il risultato dell'eliminazione dei loop è quello di creare un albero logico gerarchico.

collegamento attraverso l'attivazione del percorso in standby (cioè attivando la porta prima inattiva).

Se non ci fosse la struttura di Spanning Tree, entrambi i collegamenti ridondanti potrebbero essere considerati il percorso primario, producendo così un loop infinito di traffico sulla LAN.

Nella FIGURA 3 si mostra un esempio di rete locale con 3 switch interconnessi A, B e C.

FIGURA 3 Esempio di LAN con STP



In assenza di STP, tra questi switch si avrebbe un loop, ma se STP è attivo, il link L3 (rosso) è messo in stato *blocked*, così non può essere usato per la trasmissione dati. Se si verificasse un guasto sui link L1 oppure su L2, il link L3 verrebbe attivato automaticamente. Questa soluzione fornisce quindi una condizione di ridondanza per la rete, senza incorrere nel problema dei loop.

Anche gli switch D, E, F e G forniscono una ridondanza sui link (per via dei collegamenti segnati in rosso) e se STP è attivo, vale quanto indicato prima per gli switch A, B e C.

Annotiamo che, per tradizione, il termine "bridge" continua a essere usato anche quando STP si applica a una rete con switch, poiché STP fu sviluppato per essere utilizzato con i bridge. Quindi quando si legge "bridge" (per esempio, root bridge) si deve pensare "switch" (root switch).

Ogni switch della LAN invia dei messaggi detti **BPDU** (Bridge Protocol Data Unit), trasmessi da tutte le porte per conoscere l'esistenza di altri switch e per eleggere un **root bridge** nella rete (cioè la radice dell'albero logico che si verrà a creare).

I BPDU contengono informazioni per:

- selezionare un solo switch come root dello Spanning Tree;
- calcolare il percorso più breve da ogni switch alla root;

- eleggere il **designated switch**, che per ogni LAN è lo switch più vicino alla root, attraverso cui passano tutte le comunicazioni della LAN;
- scegliere per ogni switch la **root port**, cioè l'interfaccia che dà il miglior percorso verso la root.

Le porte che fanno parte dello Spanning Tree sono le **designated port**, le altre sono bloccate.

Quando si attiva l'algoritmo per la creazione dello Spanning Tree, trascorre un certo tempo, da 30 a 50 secondi, prima che la topologia logica della rete *converga*, ossia che tutte le porte degli switch siano nello stato *blocked* o *forwarding*. Quando la topologia cambia, gli switch ricalcolano lo Spanning Tree.

Quando la LAN ha ottenuto la convergenza e si è creato l'albero gerarchico (FIGURA 4), si hanno i seguenti elementi:

- un root bridge per LAN;
- una root port per i non root bridge;
- una designated port per segmento;
- porte non usate.

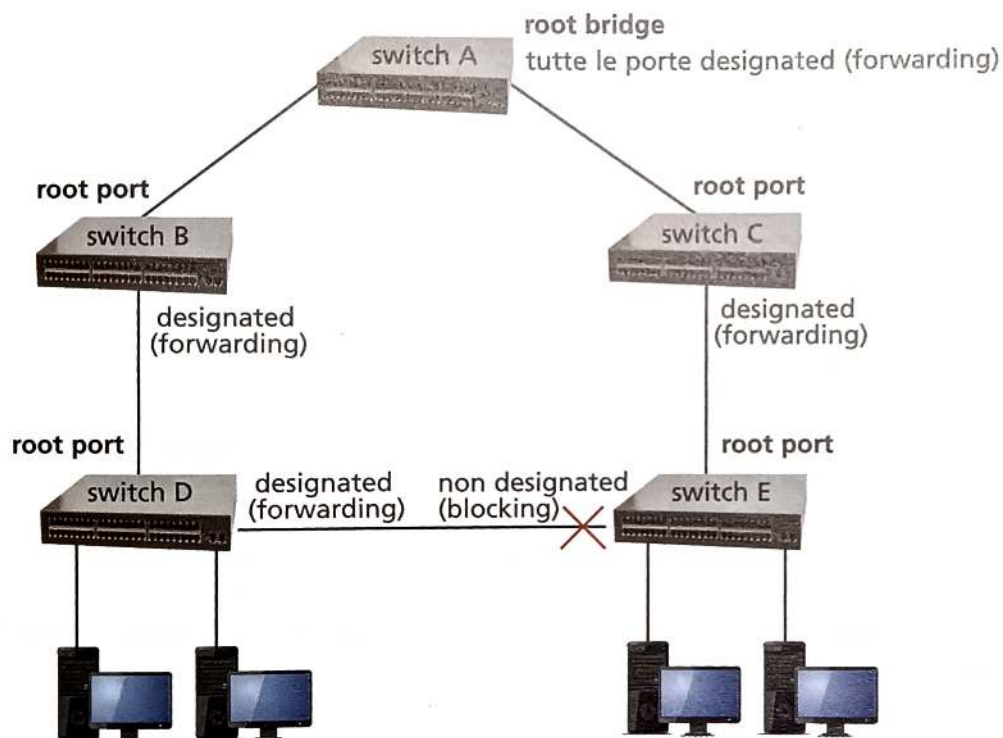


FIGURA 4 Esempio di gerarchia creata con STP

Ogni porta dello switch in STP si può quindi trovare in uno dei seguenti stati:

- **blocking**: può solo ricevere le BPDU, scarta i frame e non è in grado di apprendere nessun indirizzo fisico;
- **listening**: sta costruendo la topologia "attiva" della rete, ossia lo switch determina se ci sono altri percorsi verso il root bridge;
- **learning**: sta costruendo la tabella di bridging, quindi è in grado di apprendere gli indirizzi fisici, ma non invia né riceve dati;
- **forwarding**: può inviare e ricevere dati e in generale è in grado di svolgere tutte le funzioni possibili;
- **disabled**: è stata disabilitata (cioè resa *down*) dall'amministratore.

#techwords

In generale, il **tempo di convergenza** della rete è il tempo che la rete stessa impiega ad aggiornare i percorsi in seguito a una modifica (per esempio, il guasto o l'aggiunta di un nodo, l'attivazione dell'algoritmo Spanning Tree).

Nelle reti attuali un **#tempo di convergenza** di 30-50 secondi risulta inadeguato alle velocità elevate delle LAN, di conseguenza molti produttori hanno sviluppato delle modifiche, proprietarie, al protocollo STP standard al fine di ottenere tempi di convergenza inferiori.

Un'altra modifica è quella di permettere all'amministratore di configurare manualmente le porte alle quali è connesso un computer (e non un altro switch), così da evitare che la porta transiti attraverso tutti gli stati previsti dal protocollo, ma passi direttamente da *blocked/disabled* a *forwarding*.

Tutte queste migliorie hanno consentito di abbassare il tempo di convergenza, ma, nonostante ciò, sui link a elevata velocità che necessitano di ridondanza, gli switch di livello 2 sono sostituiti da apparati di livello superiore, chiamati **MultiLayer Switch**.

1.3 Evoluzione del protocollo Spanning Tree: RSTP

Negli anni STP ha subito varie evoluzioni: sul sito di IEEE (www.ieee.org) è possibile avere informazioni sugli sviluppi in corso che riguardano STP.

Nel 2001 IEEE definì il nuovo protocollo **Rapid Spanning Tree Protocol (RSTP)**, identificato dalla sigla IEEE 802.1w, allo scopo di rendere più veloce la convergenza dell'algoritmo Spanning Tree a fronte di cambiamenti della topologia.

Infatti, RSTP è in grado di intervenire entro: $3 \cdot \text{Hello time}$, dove *Hello time* è un intervallo di tempo usato da RSTP in vari ambiti; il suo valore di default è 2 secondi.

Nel 2004 è stata emessa una nuova specifica dello standard, denominata IEEE Std 802.1D™-2004, che include il nuovo protocollo RSTP e rende obsoleto STP.

Pur mantenendo la struttura di STP, con un root bridge da cui parte lo Spanning Tree e un path costituito dagli switch, RSTP calcola un percorso alternativo definendo le porte che lo costituiscono *alternate*; eventuali porte da cui partono altri percorsi con costo maggiore rimangono in stato *blocking*. Quando una porta in *forwarding* smette di ricevere i frame di *Hello* e i timer scadono, viene subito attivato un altro percorso aprendo le porte *alternate*, velocizzando notevolmente il processo di convergenza.

Quando uno switch ha due porte in un stesso segmento, quella a costo più alto viene definita in un nuovo modo: *backup*. Quando la porta in *forwarding* non funziona più, quella in *backup* prende il suo posto. Anche in questo caso la convergenza è molto veloce.

Da sottolineare che RSTP non migliora la convergenza nei link condivisi, infatti lo standard è stato sviluppato pensando alle reti moderne che non usano più hub ma switch e che lavorano in full-duplex.

FISSA LE CONOSCENZE

- Che cosa significa "segmentare" una rete locale?
- Che cosa si intende con dominio di collisione?
- Che differenza c'è tra una LAN segmentata usando uno switch e una segmentata usando un router?
- Quali problemi può creare l'operazione di rendere ridondante la rete?
- Spiega il protocollo Spanning Tree Protocol.
- Quali sono gli stati in cui si può trovare una porta di uno switch in cui è attivo STP?