



Controllo di flusso negli switch IEEE 802.3x

Mario Baldi

Politecnico di Torino

<http://staff.polito.it/mario.baldi>

Pietro Nicoletti

Studio Reti

<http://www.studioreti.it>

Basato sul capitolo 8 di:

M. Baldi, P. Nicoletti, "Switched LAN", McGraw-Hill, 2002, ISBN 88-386-3426-2

Nota di Copyright

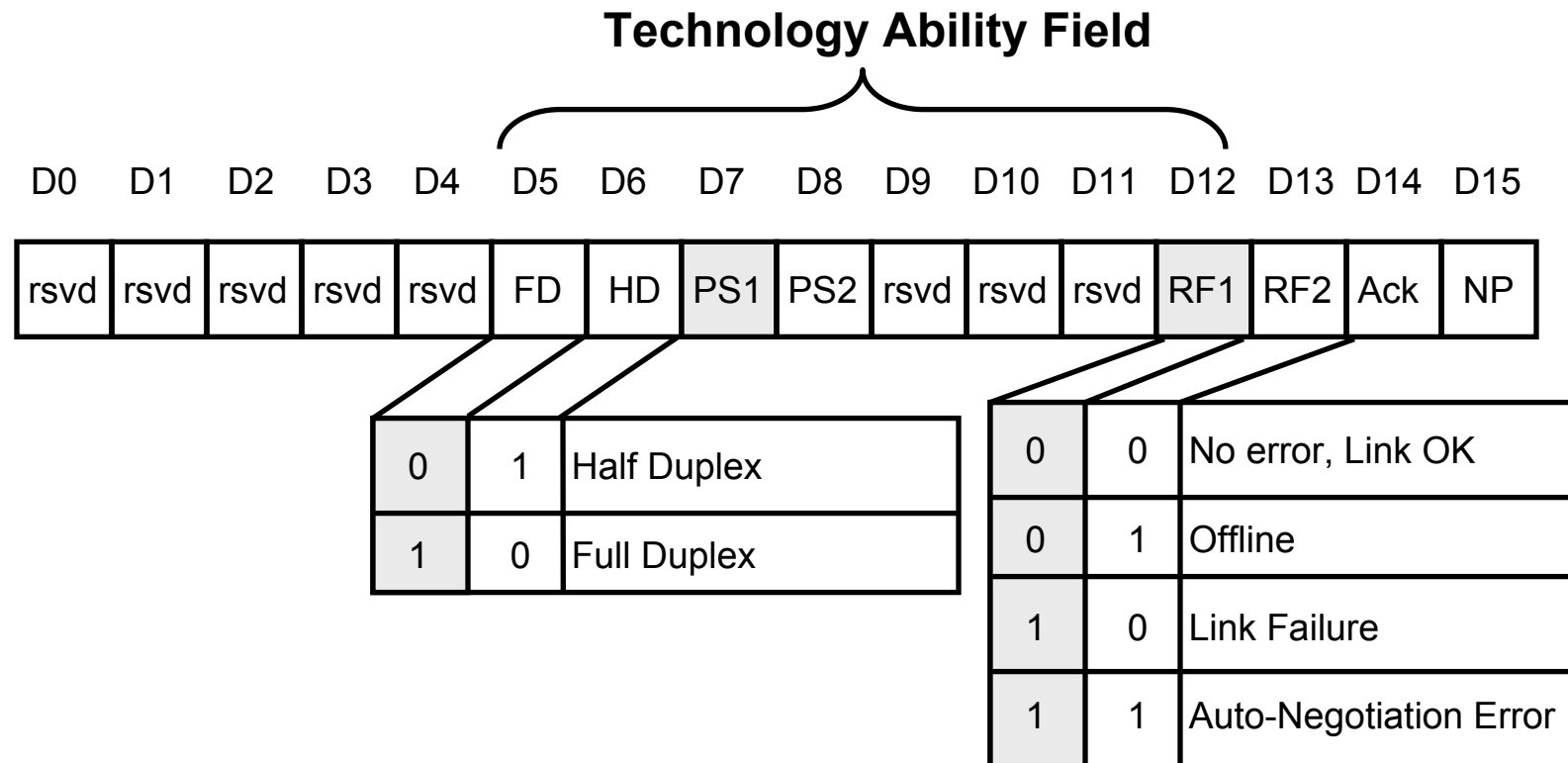
- Questo insieme di trasparenze (detto nel seguito slide) è protetto dalle leggi sul copyright e dalle disposizioni dei trattati internazionali. Il titolo ed i copyright relativi alle slide (ivi inclusi, ma non limitatamente, ogni immagine, fotografia, animazione, video, audio, musica e testo) sono di proprietà degli autori indicati a pag. 1.
- Le slide possono essere riprodotte ed utilizzate liberamente dagli istituti di ricerca, scolastici ed universitari afferenti al Ministero della Pubblica Istruzione e al Ministero dell'Università e Ricerca Scientifica e Tecnologica, per scopi istituzionali, non a fine di lucro. In tal caso non è richiesta alcuna autorizzazione.
- Ogni altra utilizzazione o riproduzione (ivi incluse, ma non limitatamente, le riproduzioni su supporti magnetici, su reti di calcolatori e stampate) in toto o in parte è vietata, se non esplicitamente autorizzata per iscritto, a priori, da parte degli autori.
- L'informazione contenuta in queste slide è ritenuta essere accurata alla data della pubblicazione. Essa è fornita per scopi meramente didattici e non per essere utilizzata in progetti di impianti, prodotti, reti, ecc. In ogni caso essa è soggetta a cambiamenti senza preavviso. Gli autori non assumono alcuna responsabilità per il contenuto di queste slide (ivi incluse, ma non limitatamente, la correttezza, completezza, applicabilità, aggiornamento dell'informazione).
- In ogni caso non può essere dichiarata conformità all'informazione contenuta in queste slide.
- In ogni caso questa nota di copyright non deve mai essere rimossa e deve essere riportata anche in utilizzi parziali.

Half o full duplex?

- Le LAN sono strutture intrinsecamente half-duplex:
 - trasmette una sola stazione per volta
- Lo switching ridimensiona molto il ruolo del mezzo fisico condiviso:
 - spesso il mezzo trasmissivo diventa punto-punto: sono collegati unicamente la stazione e lo switch
- I mezzi trasmissivi punto-punto possono essere full-duplex:
 - entrambe le stazioni possono trasmettere contemporaneamente
 - le trasmissioni avvengono su canali fisici diversi

Full duplex e standard 802.3x

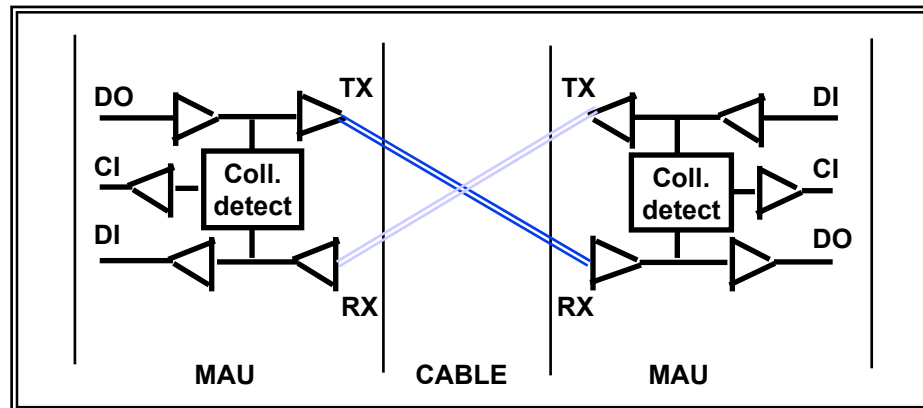
- Lo standard 802.3x definisce la modalità di funzionamento full duplex
- Negozia la modalità tra gli apparati operativa e la memorizza el registro 4



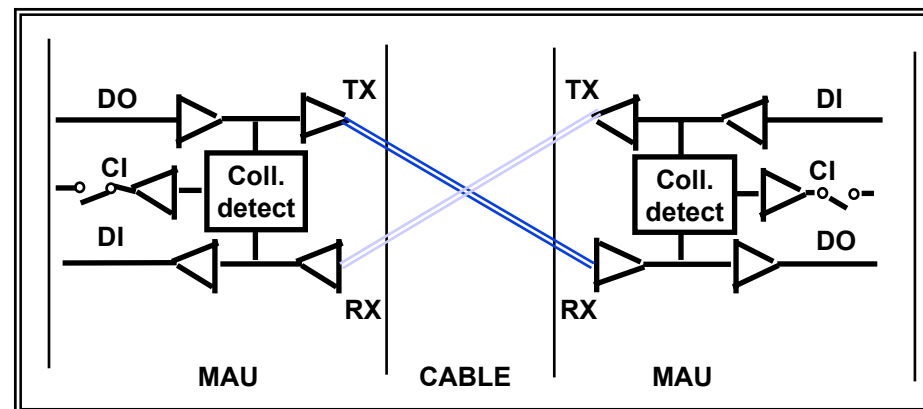
Modalità di funzionamento full duplex

- Non viene più utilizzato il MAC CSMA-CD
 - la stazione Ethernet trasmette immediatamente il pacchetto senza mettersi in ascolto sul canale trasmissivo
- Adottato sicuramente nelle connessioni tra switch, meno frequentemente tra switch e stazione
- Necessita di particolari transceiver in cui non viene rilevata la collisione:
 - i transceiver normali inviano un segnale di collisione all'interfaccia quando si ha la presenza di attività contemporanea su TX e RX
- La distanza tra due stazioni Ethernet full-duplex
 - dipende solo dalle caratteristiche del canale trasmissivo
 - è indipendente dal diametro del dominio di collisione

I transceiver half e full duplex



Connessioni tra Transceiver tradizionali (più vecchi)



Connessioni tra Transceiver per link Full-Duplex

DO = Data Output DI = Data Input CI = Collision Input

Limiti di distanza

- In Ethernet full-duplex:
 - nel caso del doppino telefonico la distanza massima è di 100 m
 - nel caso di fibra ottica multimodale 62.5/125 μm la distanza massima è di 2 Km
 - nel caso di fibra ottica monomodale e transceiver dotati di laser di categoria II (caso FDDI e Fast-Ethernet) la distanza massima può raggiungere i 50 Km
 - in Gigabit Ethernet con laser ad alta potenza:
 - fibra monomodale 75 Km
 - fibra monomodale dispersion-shift 100 Km

Controllo di flusso: IEEE 802.3x

- Standard che definisce:
 - le modifiche necessarie al MAC IEEE 802.3 per supportare la modalità operativa Full-Duplex
 - un meccanismo di controllo di flusso per link Full-Duplex
 - vale per tutte le diverse tipologie di reti di Ethernet (10/100/1000 Mb/s)
- Obbligatorio su Gigabit Ethernet, opzionale su Ethernet e Fast Ethernet

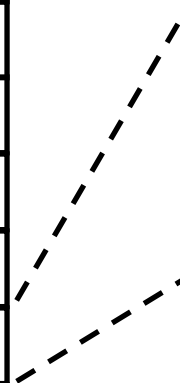
Modello OSI e IEEE 802.3x

- IEEE 802.3x introduce un sottolivello (**MAC Control**) tra il MAC 802.3 e il sottolivello superiore (Bridge Relay Entity, LLC, ...)

Modello ISO/OSI



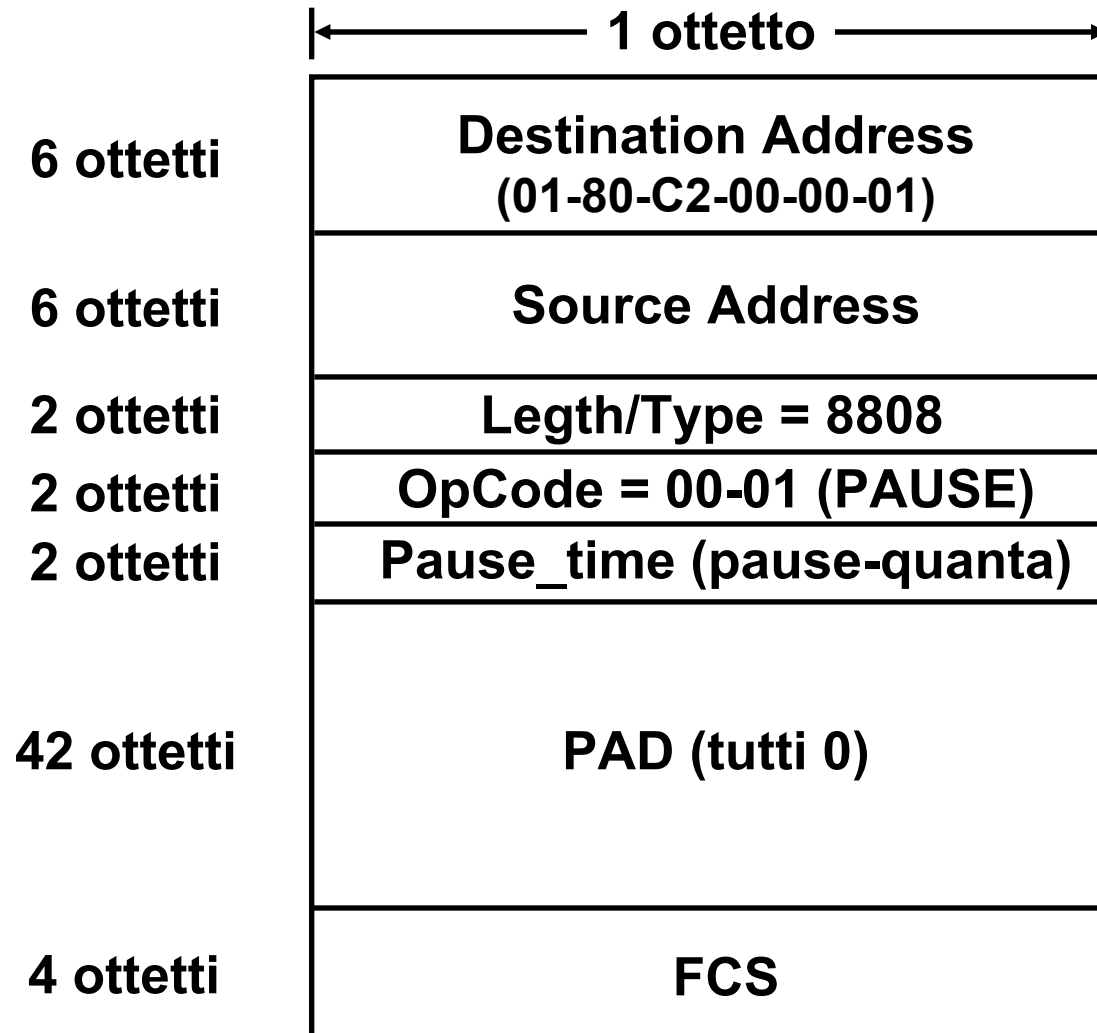
Sottolivelli per LAN 802



Controllo di flusso

- Meccanismo di controllo di flusso definito dall'IEEE 802.3x:
 - il dispositivo che vuole fermare la trasmissione invia un pacchetto **PAUSE** in multicast a tutti i partner coinvolti nella trasmissione
 - il pacchetto contiene l'indicazione del tempo (in slot time) per il quale ogni partner deve fermare la trasmissione
 - il tempo può essere esteso o annullato con l'invio di un ulteriore pacchetto PAUSE

Pacchetto PAUSE



Tempo di pausa

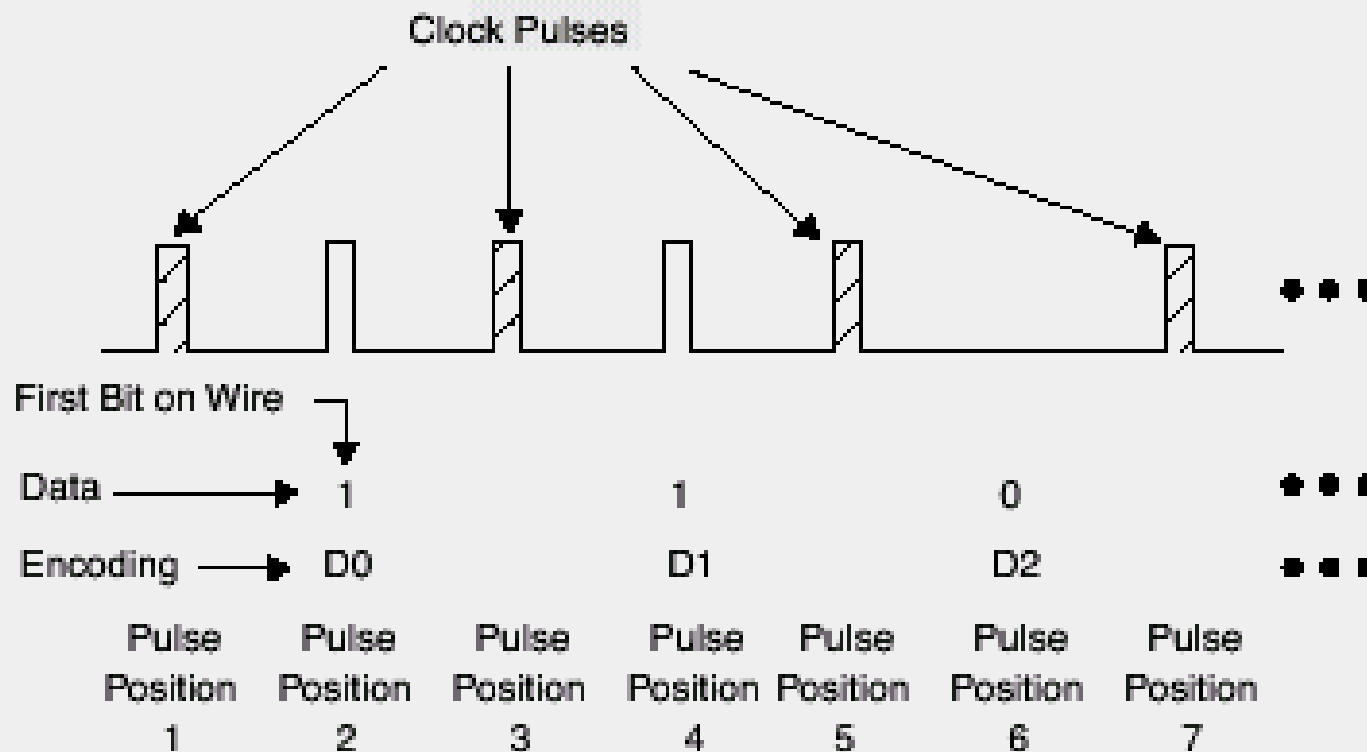
- Campo `Pause_time`: contiene il numero di pause-quanta (da 0 a 65535) che indicano il tempo di pausa
 - pause-quanta = 512 bit time
 - velocità uguali o inferiori a 100 Mb/s
 - T-Pause in bit time = pause-quanta * 512
 - velocità superiori a 100 Mb/s
 - T-Pause in bit time = pause-quanta * 512 * 2

IEEE 802.3x: modalità flow control

- Esistono due meccanismi di controllo di flusso:
 - modalità **asimmetrica**
 - solo uno dei due apparati può inviare il pacchetto pause, l'altro si limita a riceverlo e fermare l'invio di dati
 - modalità **simmetrica**
 - entrambi gli apparati all'estremità di un link possono trasmettere e ricevere il pacchetto pause

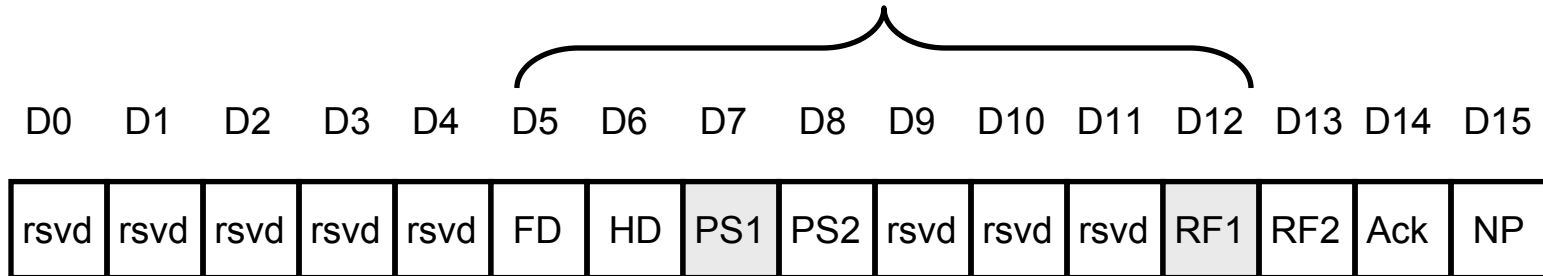
Flow Control negotiation

- Negoziazione flow control tramite codifica su Burst FLP (Fast Link Pulse)
 - inviato inizialmente dai due partner in fase di negoziazione dei parametri del link



Flow control: codifica su FLP Reg.4

Technology Ability Field



0	0	No Pause
0	1	Asymmetric Pause verso il link partner
1	0	Symmetric Pause
1	1	Symmetric Pause e Asymmetric Pause (tipo "Both") verso il local device

0	0	No error, Link OK
0	1	Offline
1	0	Link Failure
1	1	Auto-Negotiation Error

Flow control: codifica su FLP Reg.4

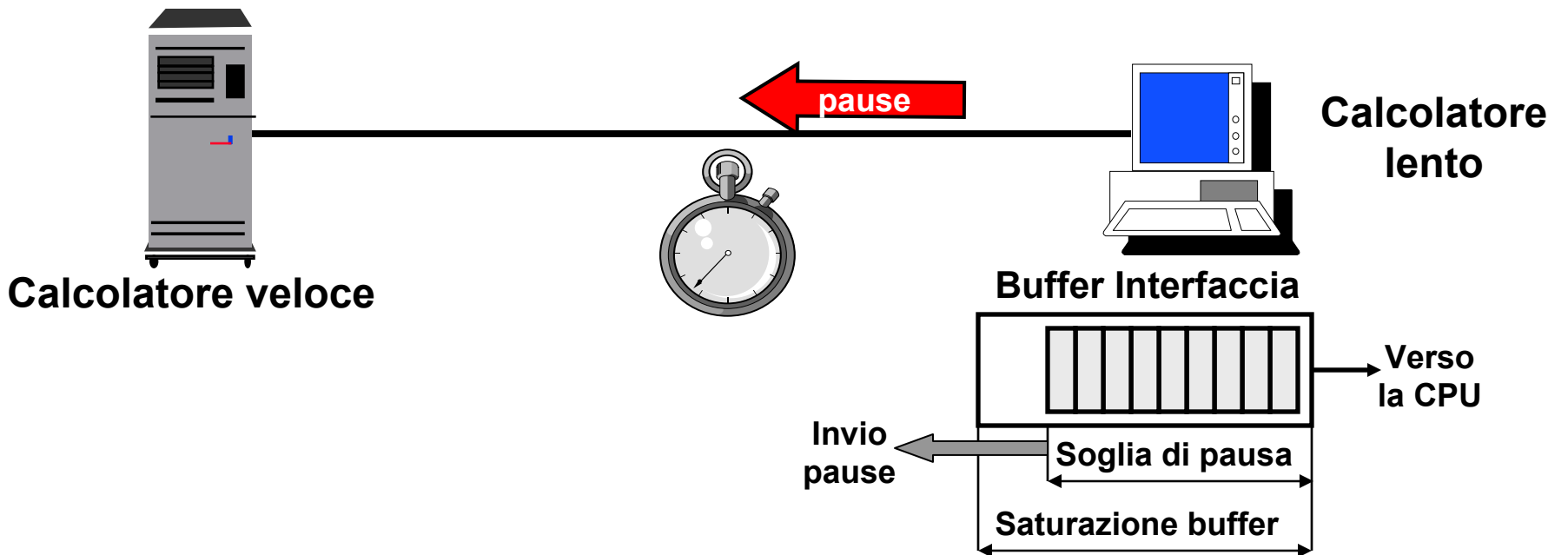
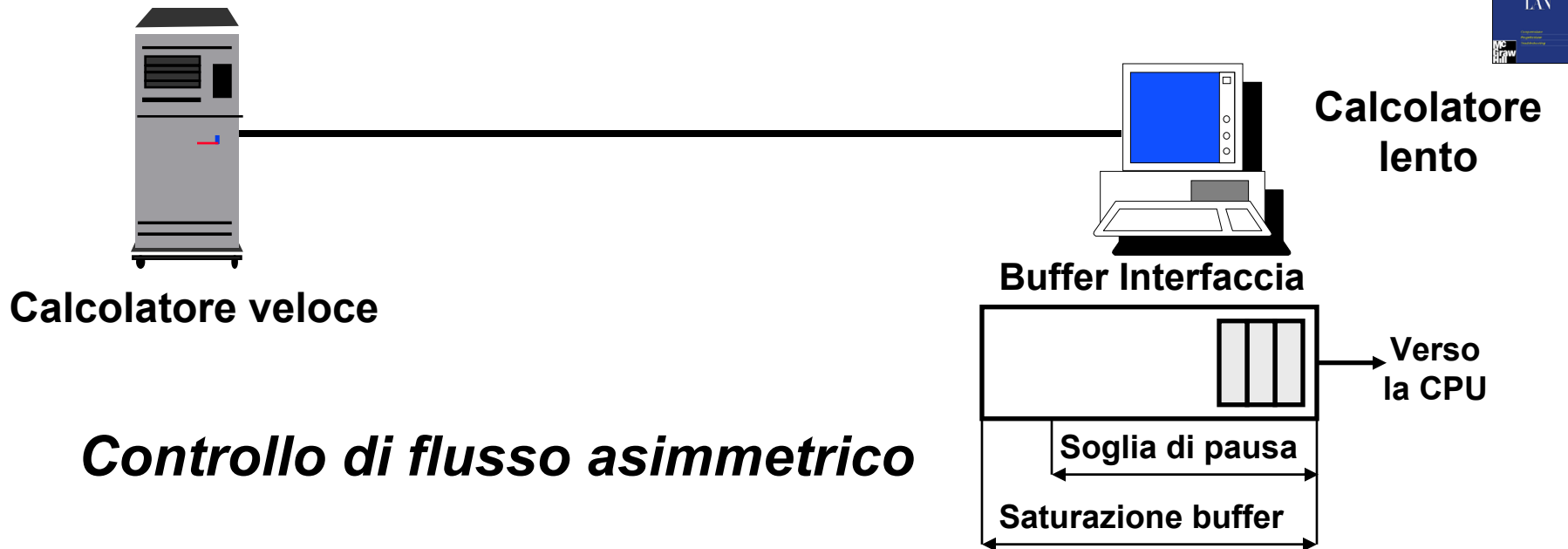
Register 4: Auto-Negotiation Advertisement Register

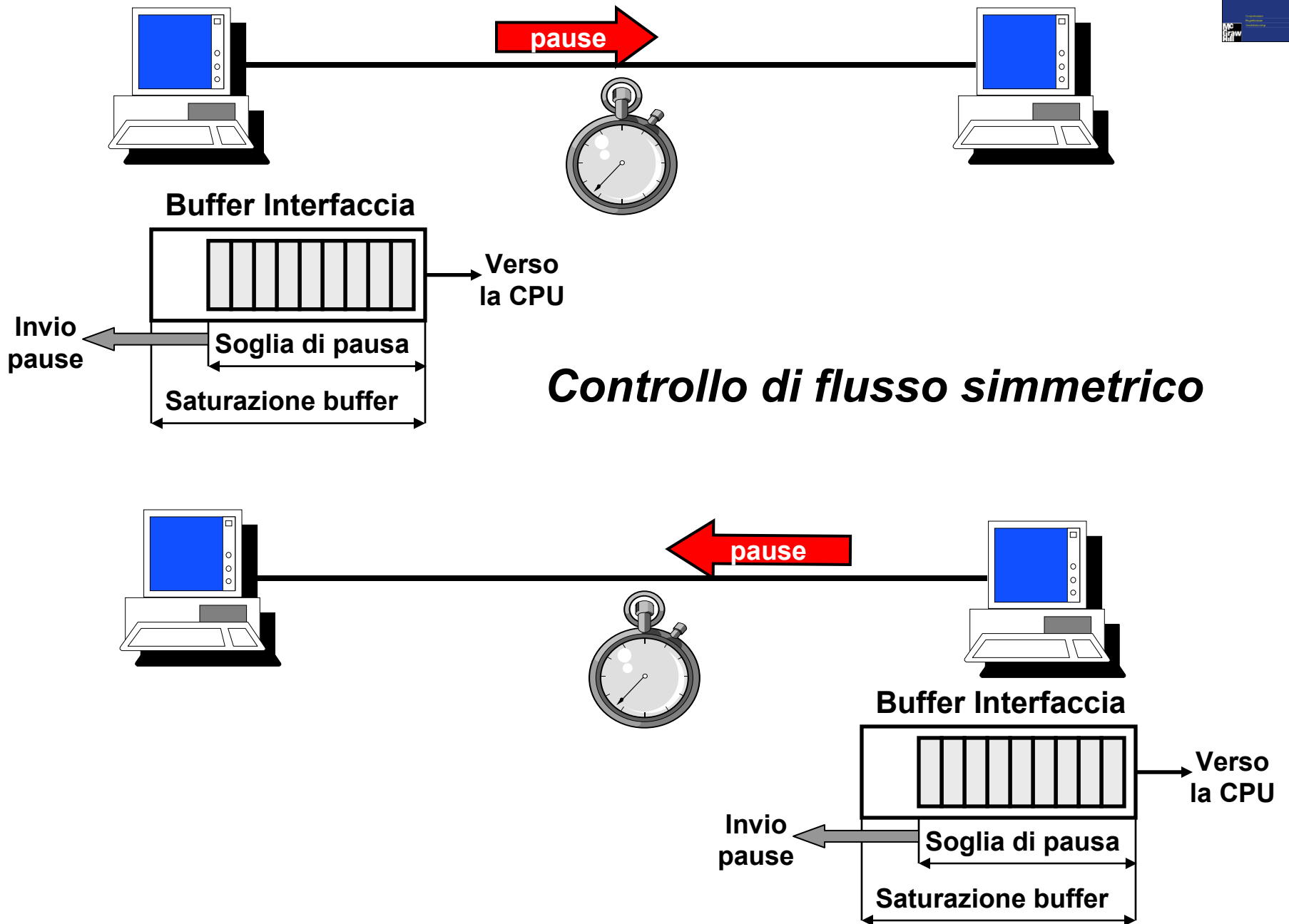
Bit(s)	Nome	Descrizione	Default	R/W
15	Next Page	Constant 0 = trasmissione della pagina con le capacità primarie	0	RO
14	Reserved	Riservato. Deve essere impostato al valore 0.	0	RO
13	Remote Fault	1 = malfunzionamento al lato opposto del collegamento 0 = nessun malfunzionamento	0	RW
12:5	Technology Ability Field	Campo di 8 bit contenente informazioni sulle funzionalità specifiche delle tecnologie identificate dal valore del campo selector field cui si offre supporto.	00101111	RW
4:0	Selector Field	Campo di 5 bit che identifica il tipo di messaggio inviato per la negoziazione. Nel circuito Intel 82559 questo campo non è scrivibile (read only) e contiene il valore 00001b che indica lo standard IEEE 802.3.	00001	RO

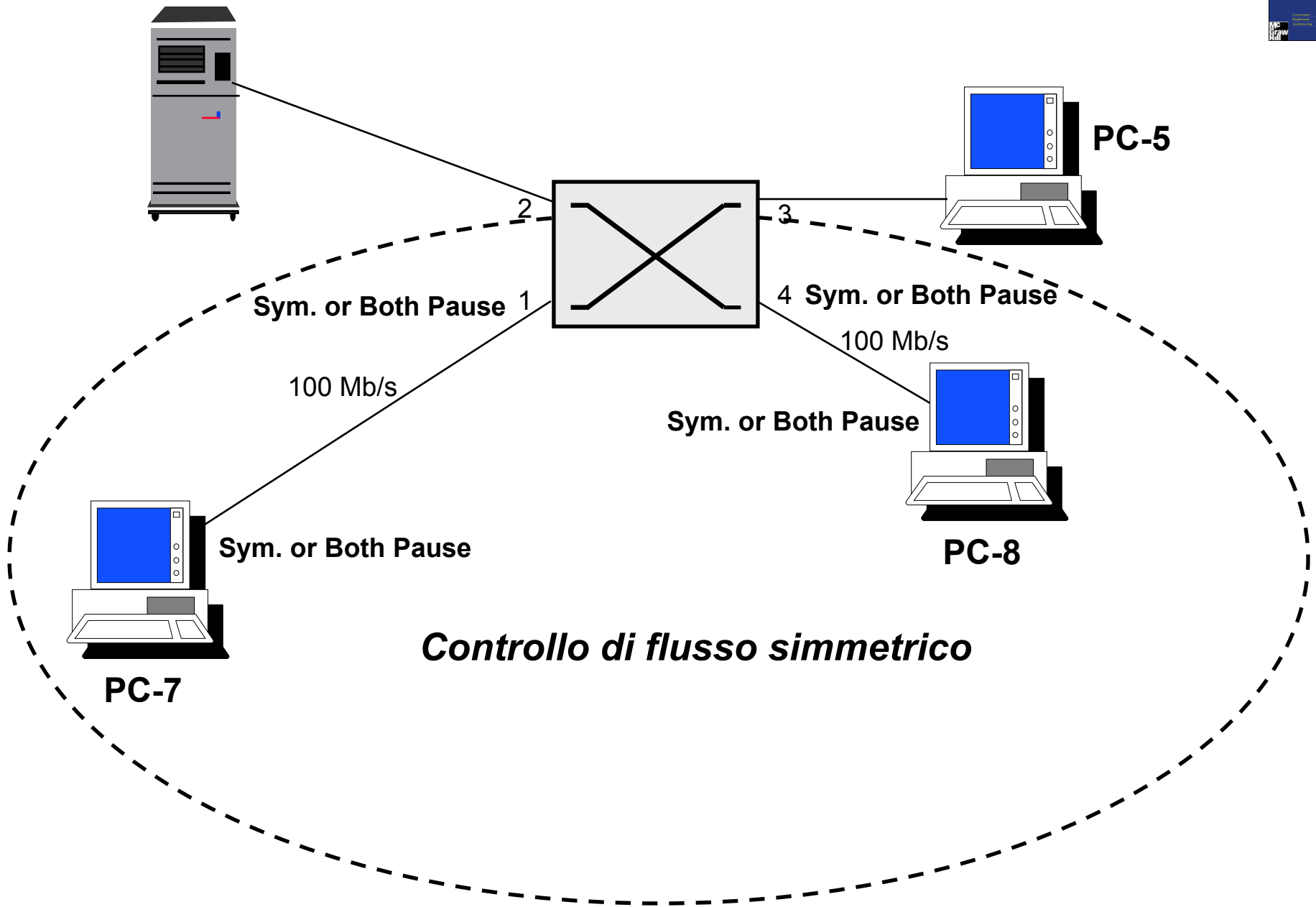
Local Device		Link Partner		Local Device resolution	Link Partner resolution
Bit - PS1	Bit - PS2	Bit - PS1	Bit - PS2		
0	0	Don't care	Don't care	Disable PAUSE TX & RX	Disable PAUSE TX & RX
0	1	0	Don't care	Disable PAUSE TX & RX	Disable PAUSE TX & RX
0	1	1	0	Disable PAUSE TX & RX	Disable PAUSE TX & RX
0	1	1	1	Enable PAUSE TX Disable PAUSE RX	Enable PAUSE RX Disable PAUSE TX
Asymmetric PAUSE		Both Sym & Asym Pause			
1	0	0	Don't care	Disable PAUSE TX & RX	Disable PAUSE TX & RX
1	Don't care	1	Don't care	Enable PAUSE TX & RX	Enable PAUSE TX & RX
Symmetric or Both		Symmetric or Both			
1	1	0	0	Disable PAUSE TX & RX	Disable PAUSE TX & RX
1	1	0	1	Enable PAUSE RX Disable PAUSE TX	Enable PAUSE TX Disable PAUSE RX
Both Sym & Asym Pause		Asymmetric PAUSE			

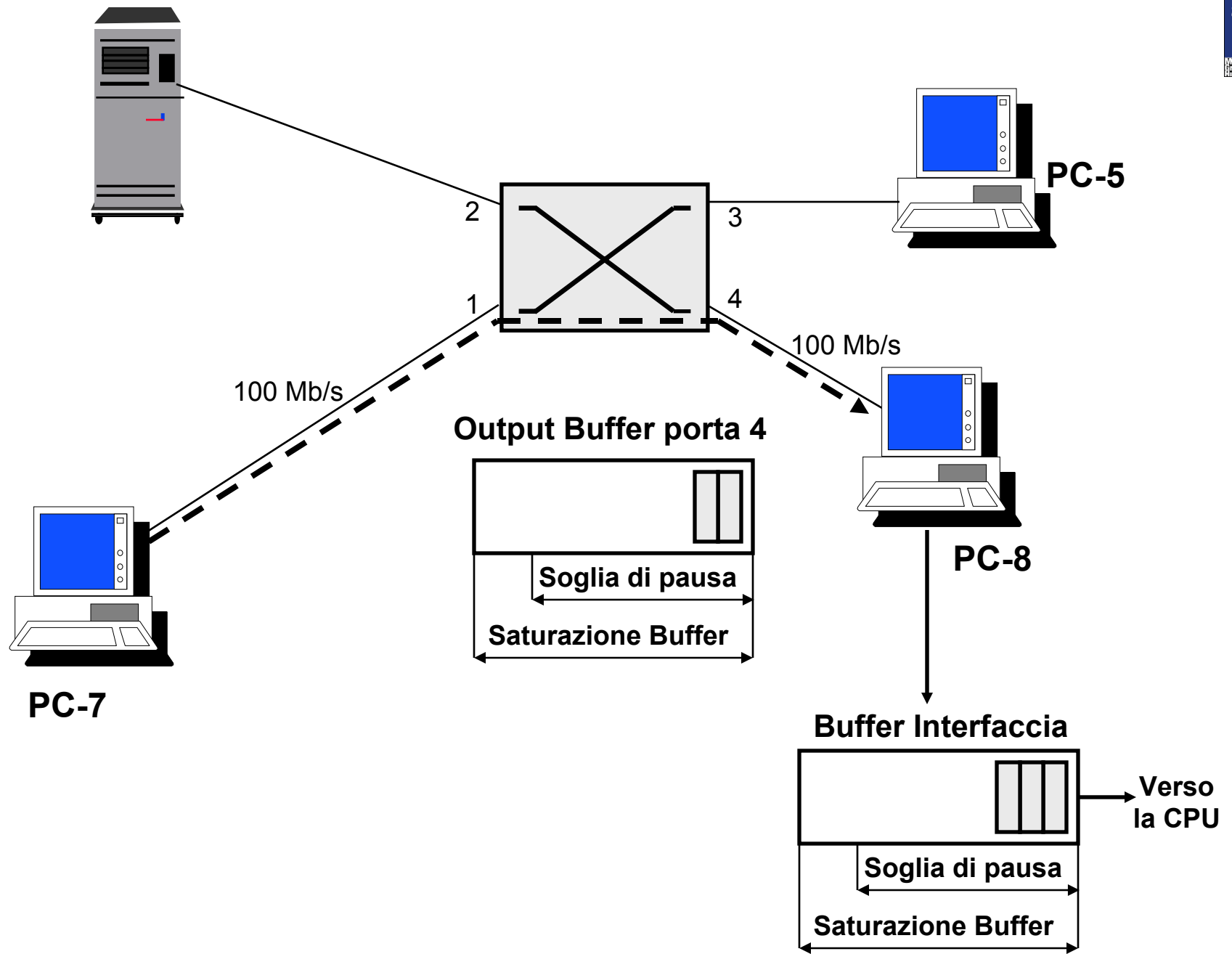
Flow control su switch: con buffer in uscita

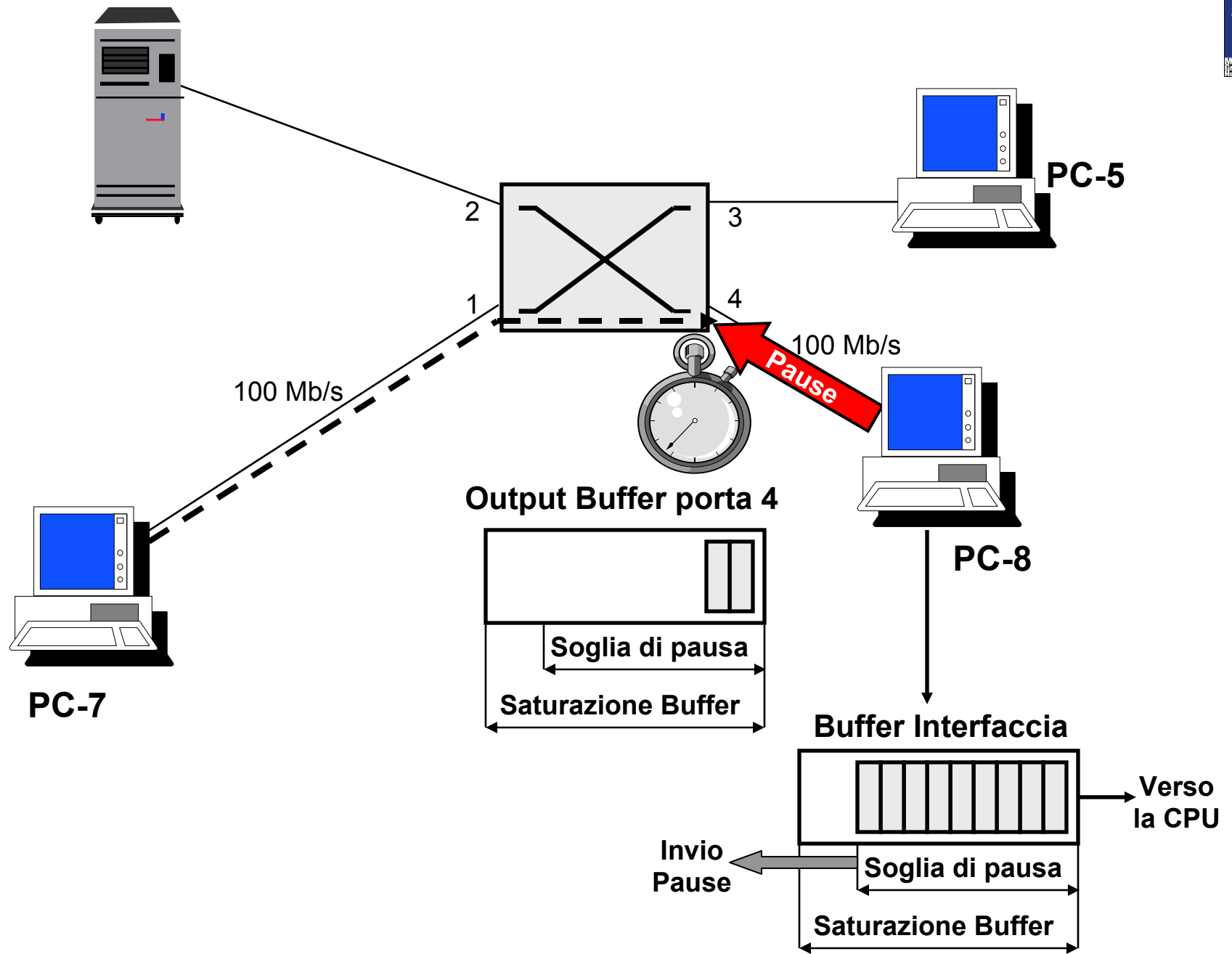
- Switch con concatenamento tra output buffer e input port
 - la saturazione di un buffer di uscita provoca l'invio di Pause sulla porta d'ingresso ad esso concatenata col flusso di traffico
 - approccio non adottato dai produttori perché presenta più svantaggi che vantaggi
 - il pacchetto di Pause su un link tra due switch penalizza anche i flussi di traffico che non congestionavano i buffer di uscita
- Soluzione adottata dai produttori:
 - lo switch blocca la trasmissione su una porta se riceve il pacchetto di Pause, ma non può scatenare l'invio del pacchetto di pause

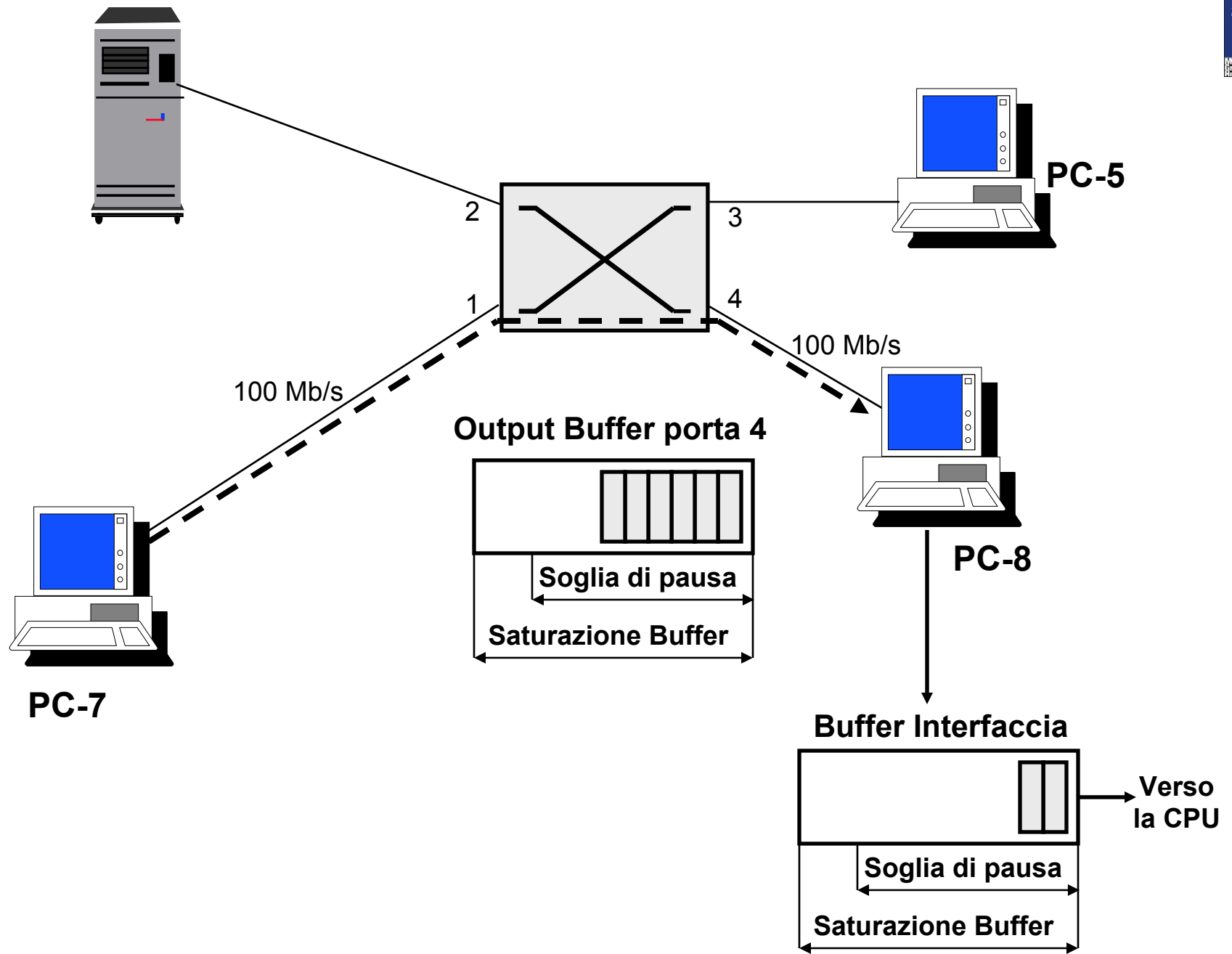


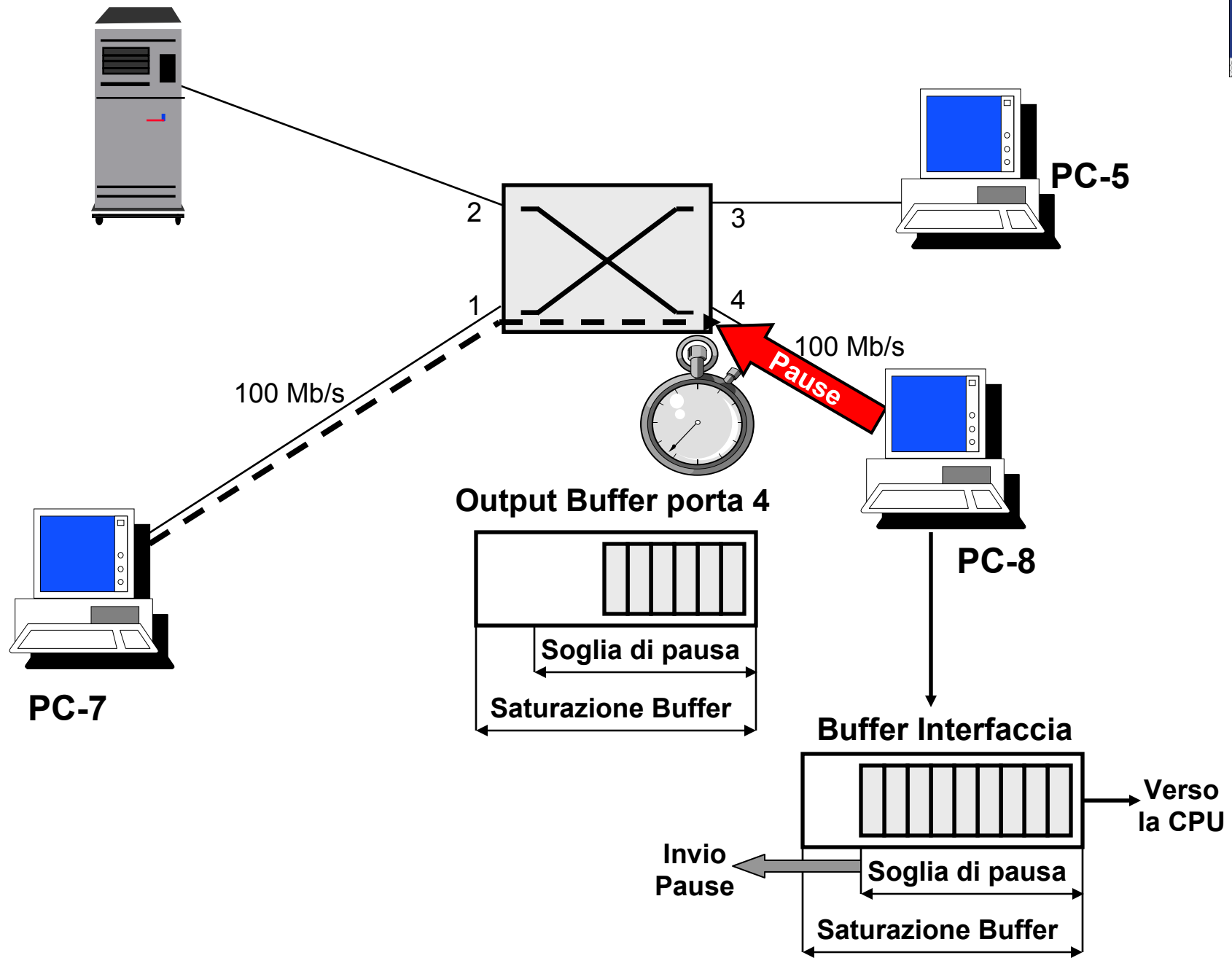


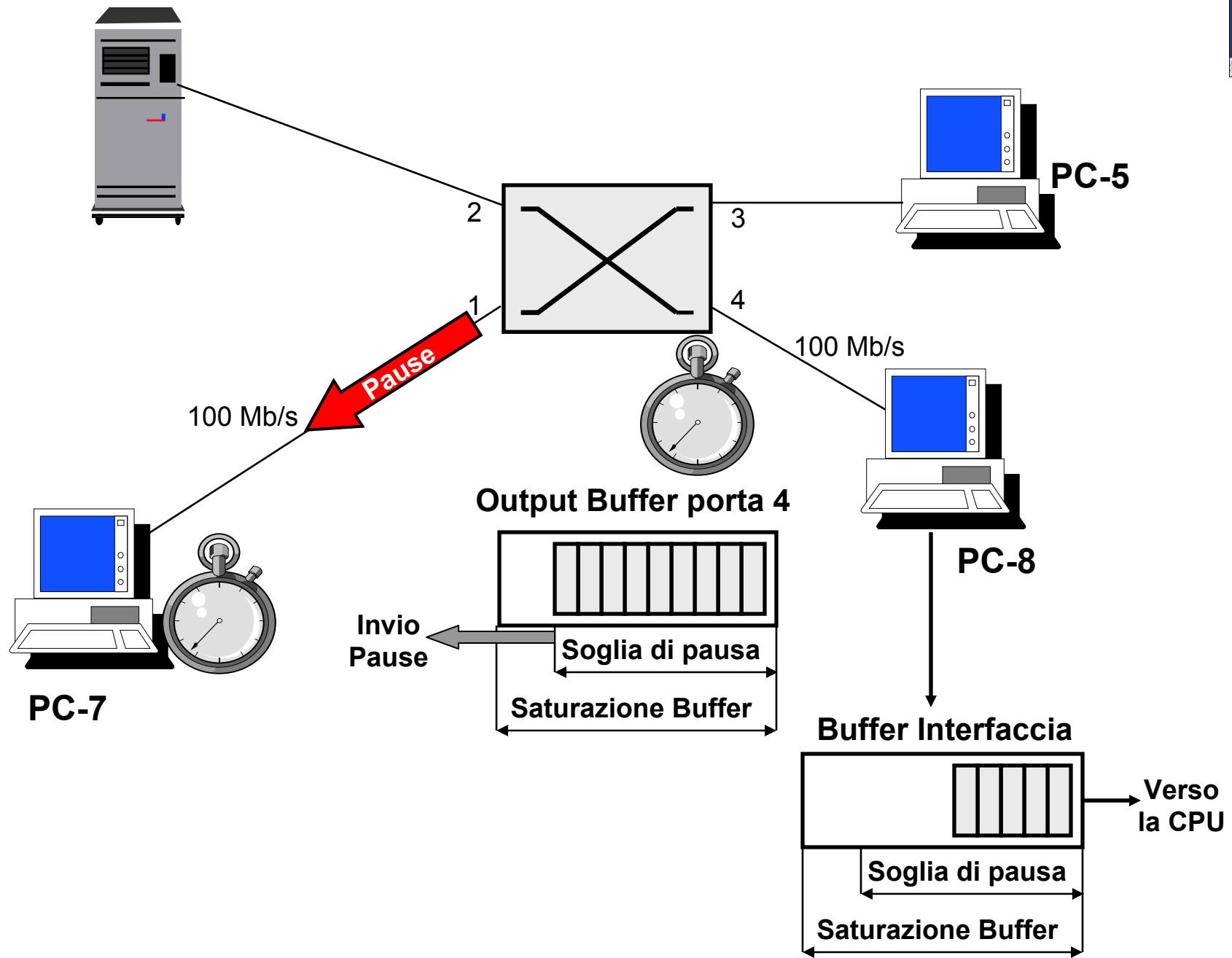


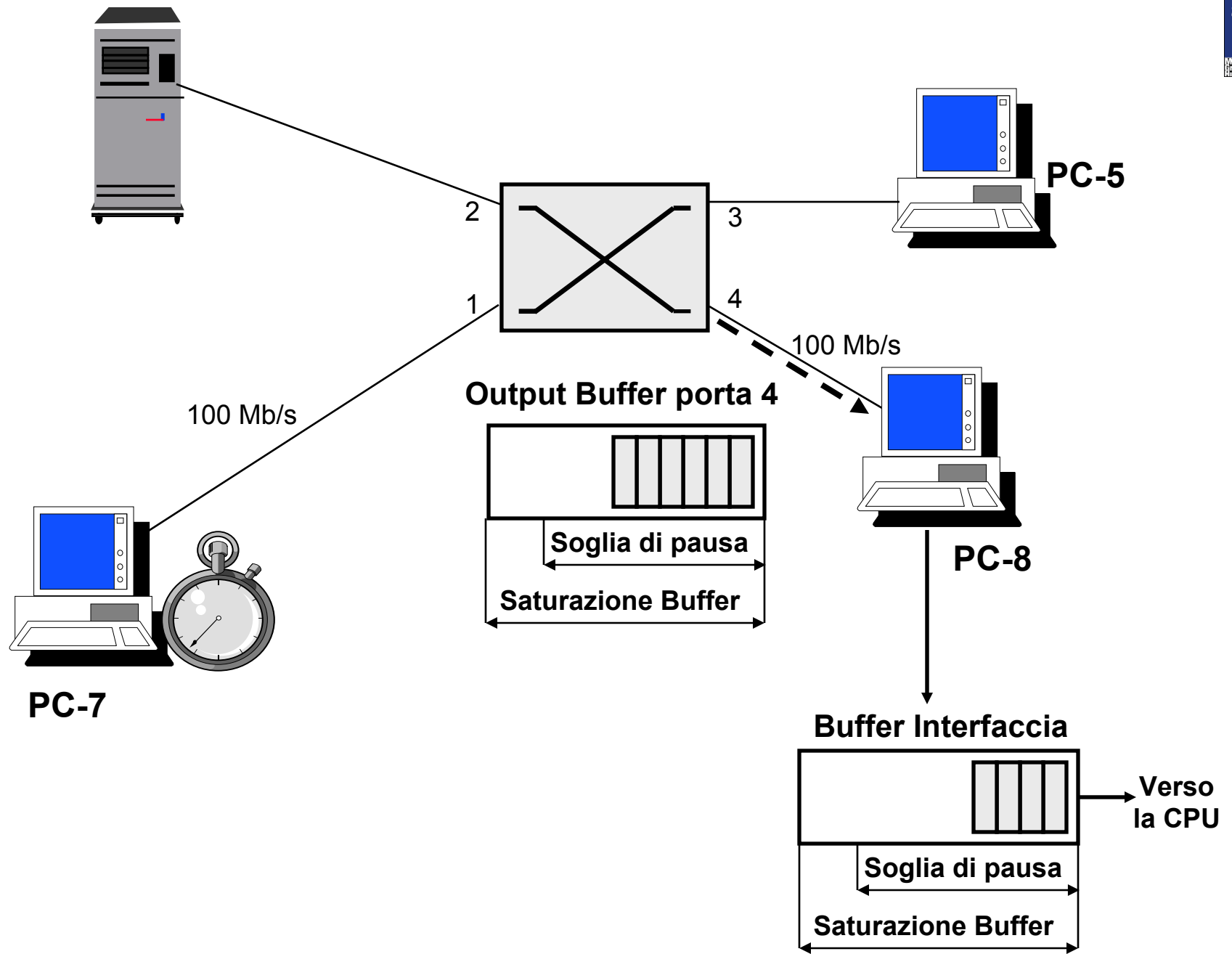


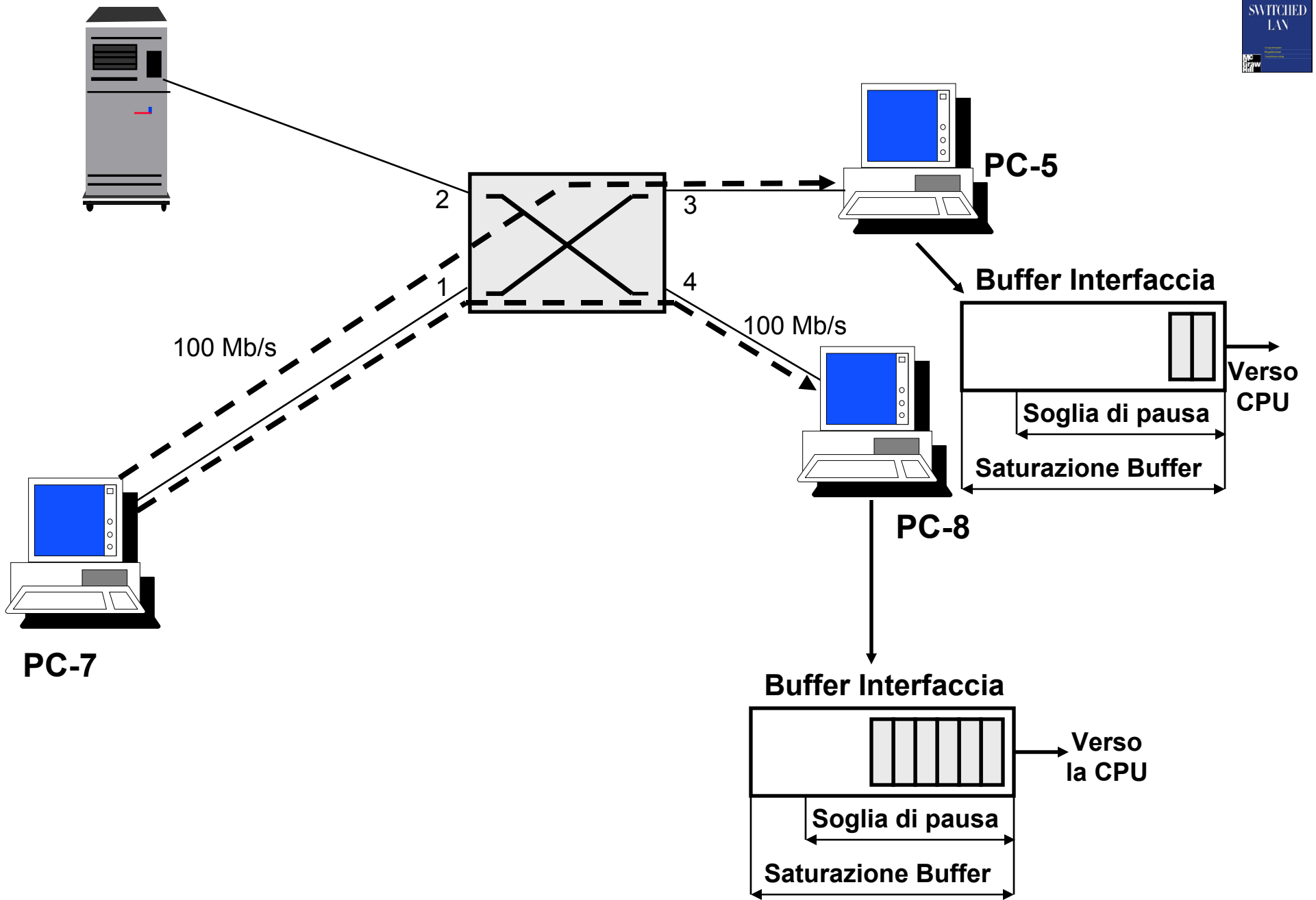


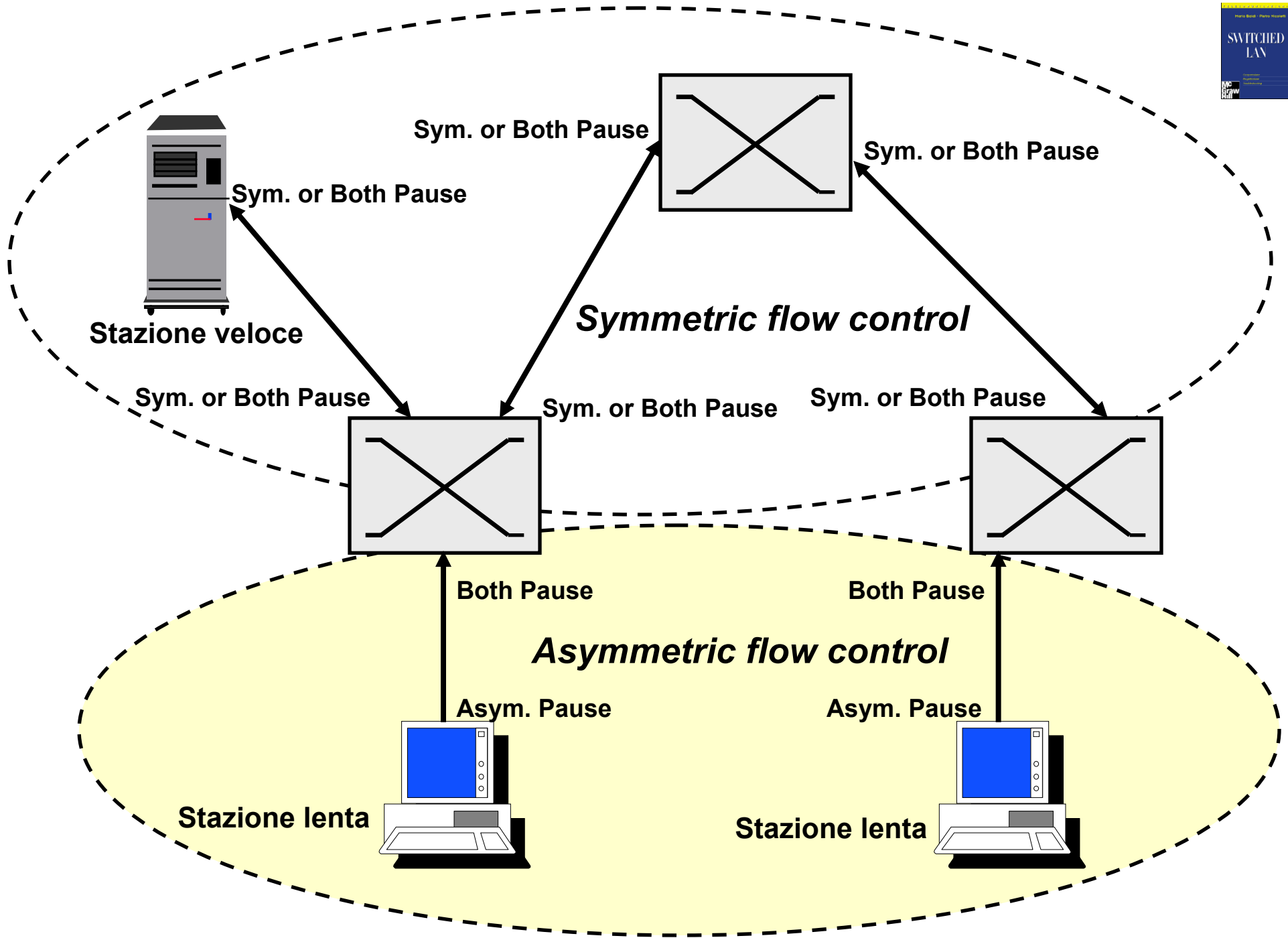


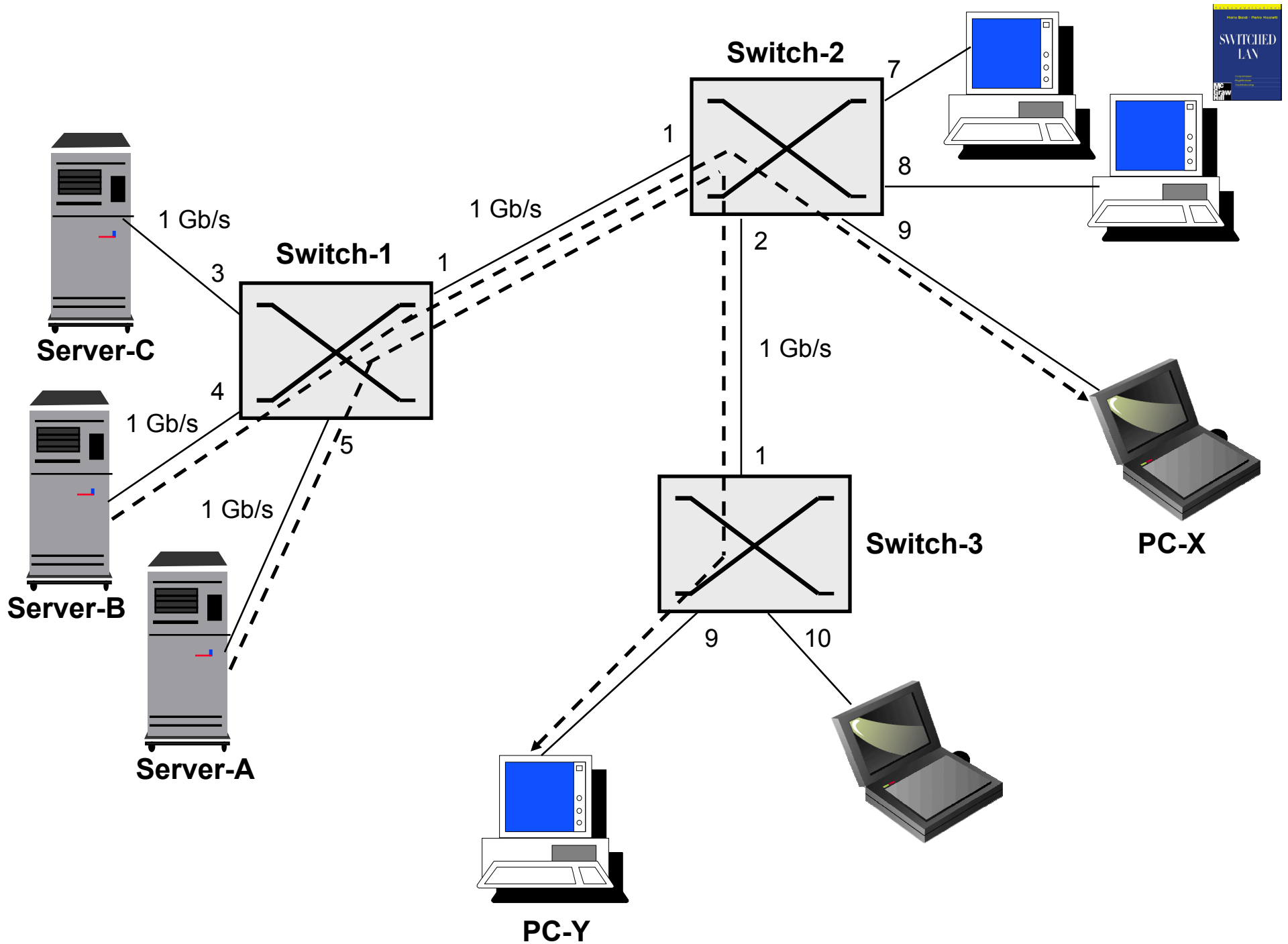


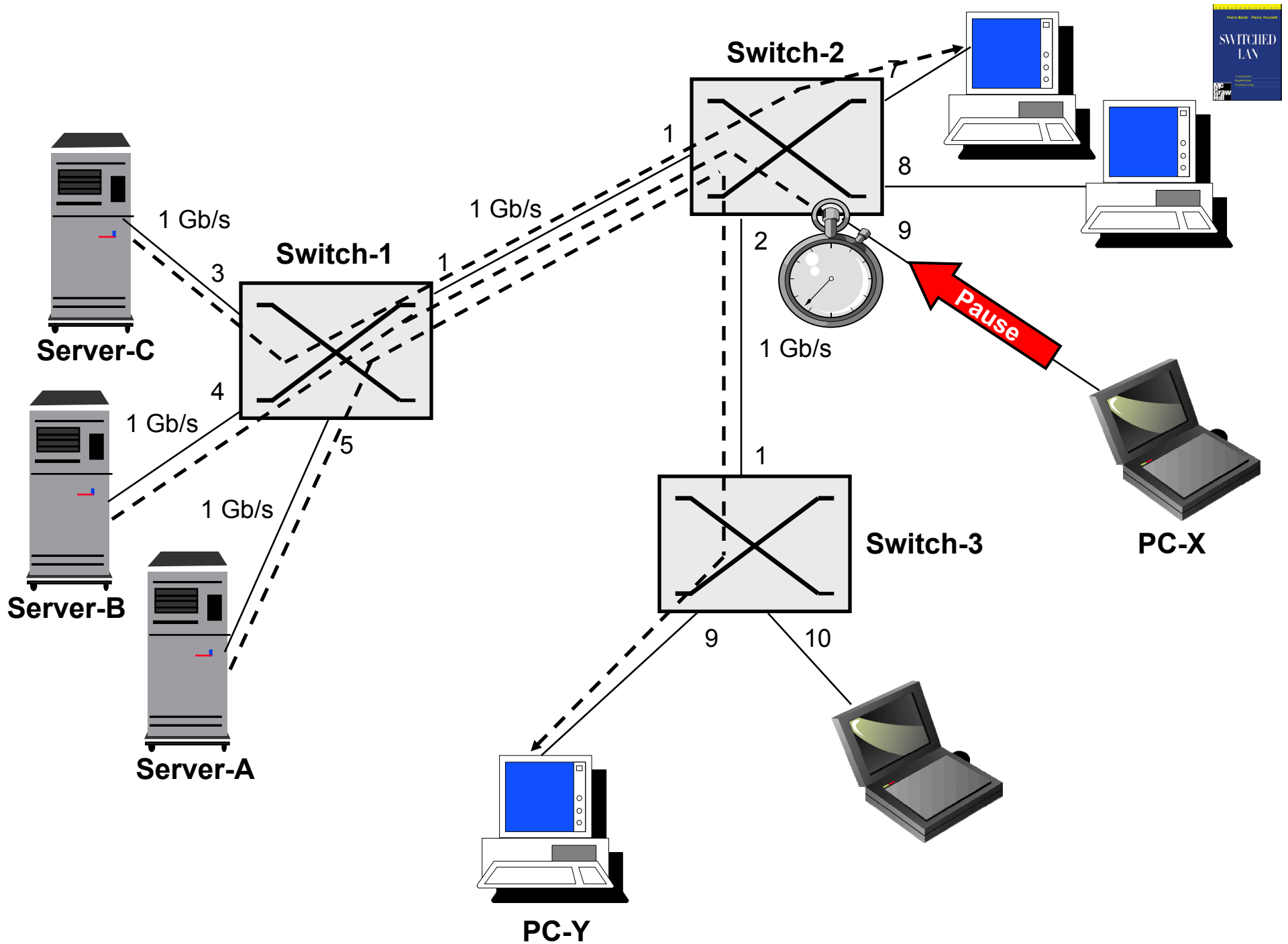


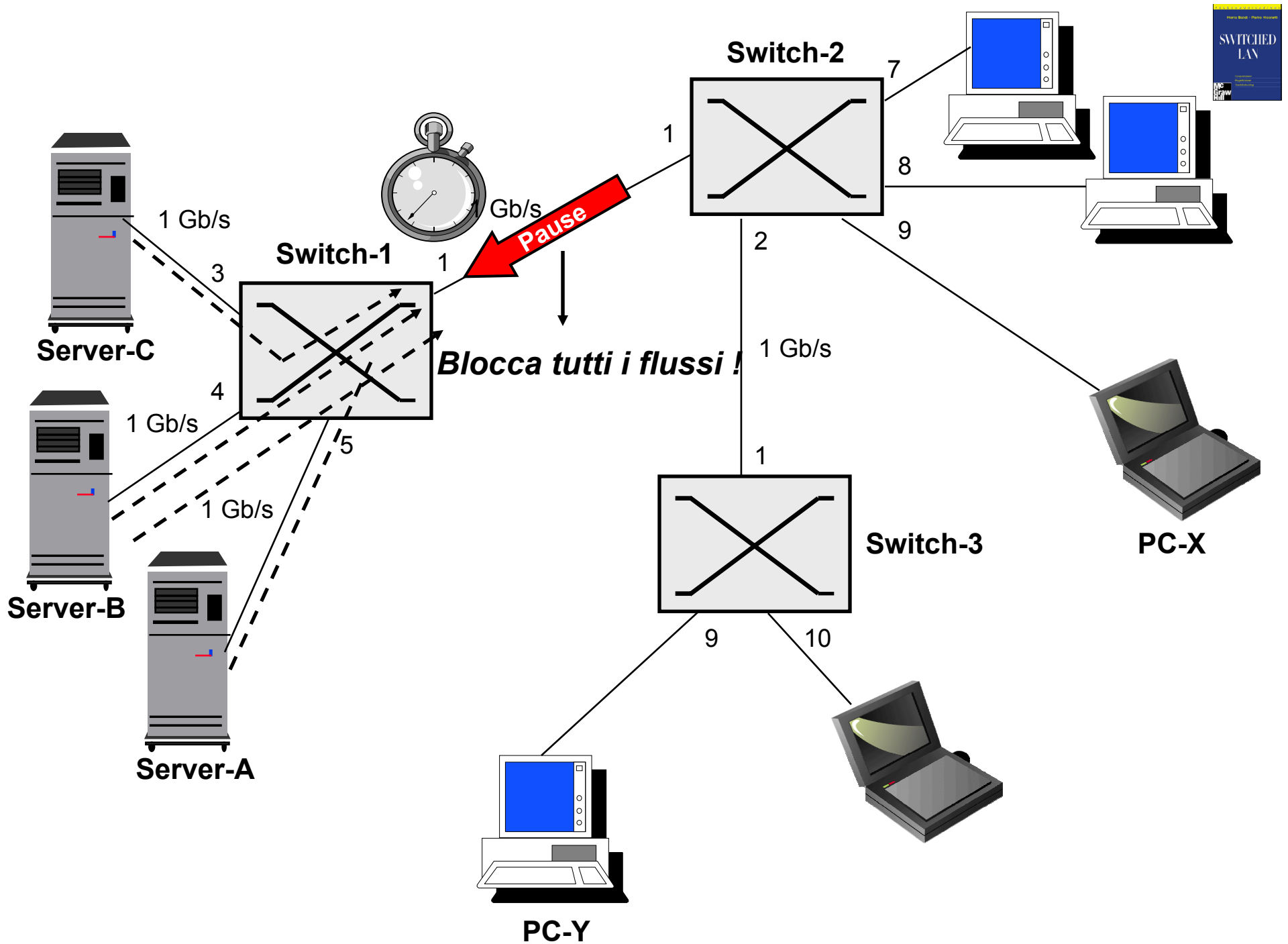






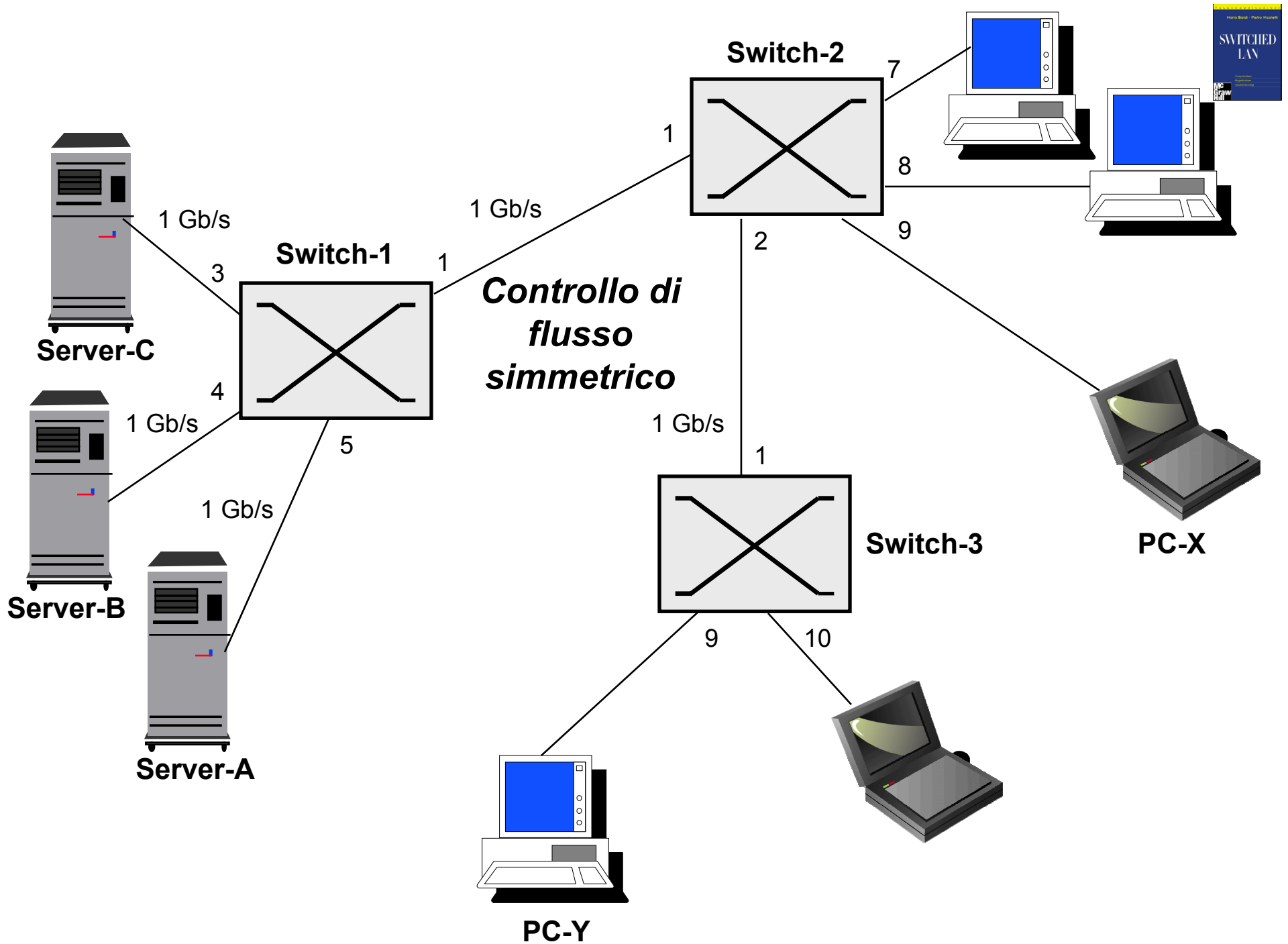


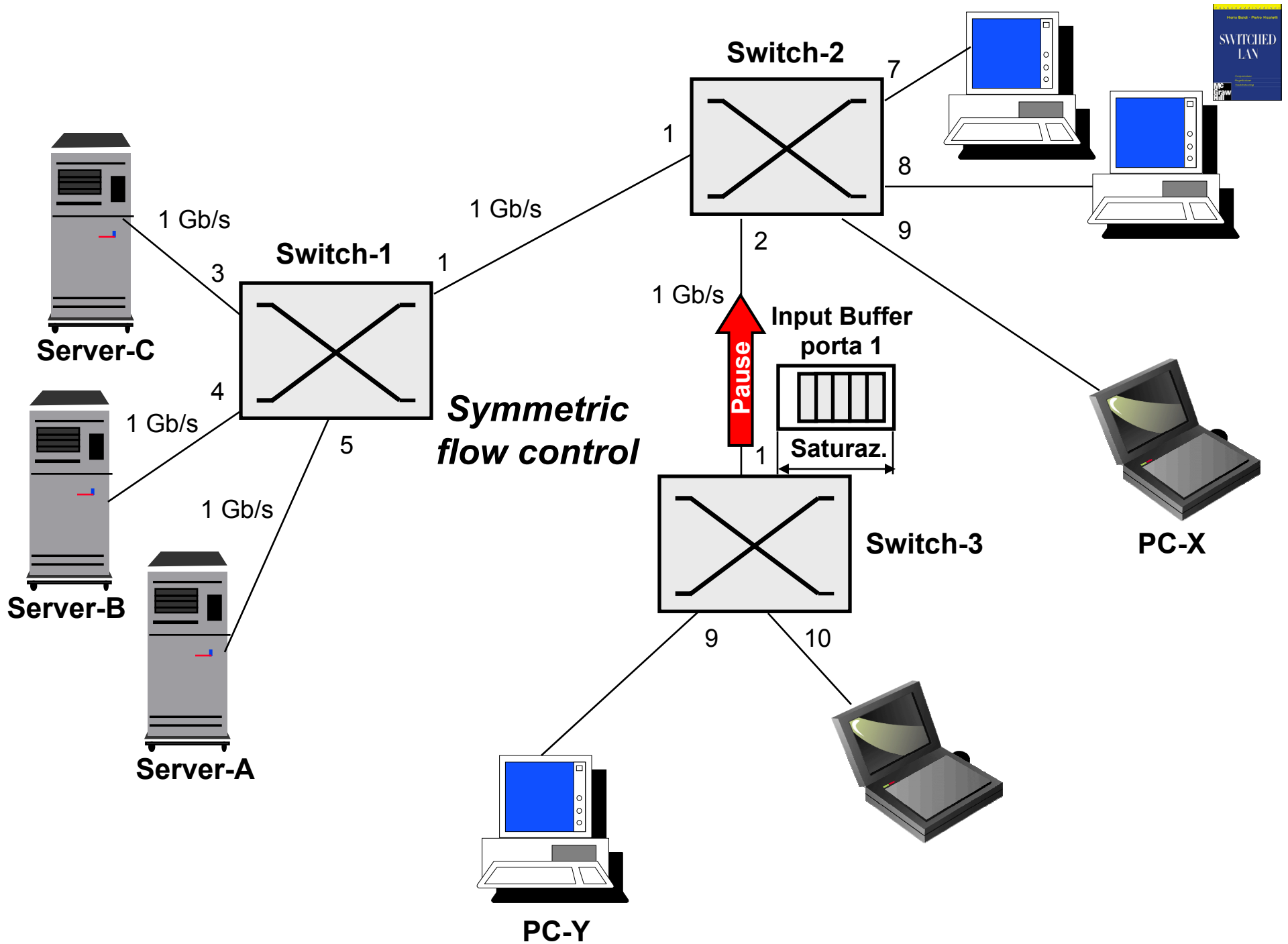


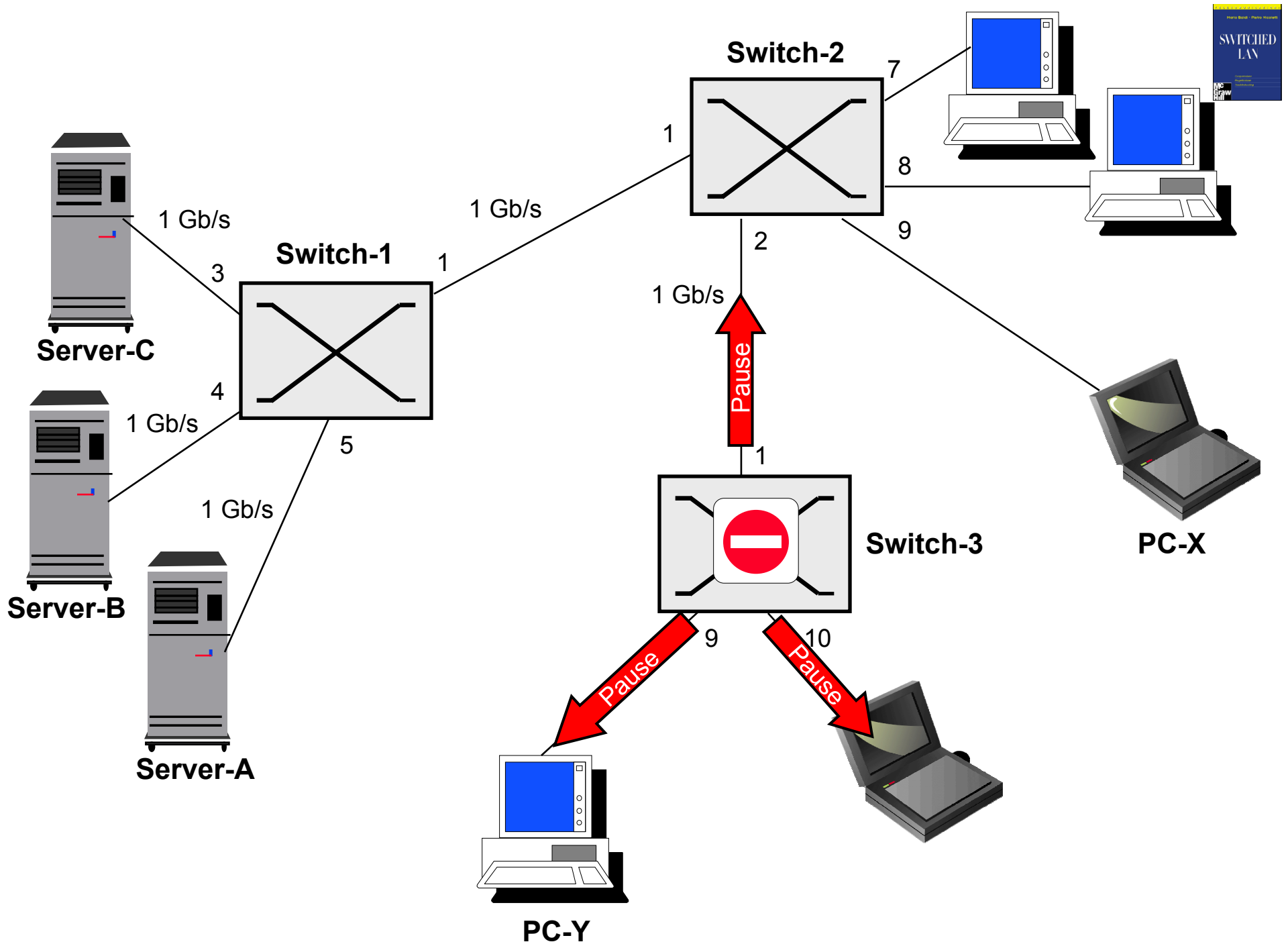


Flow control su switch con buffer in ingresso

- Se un buffer di ingresso raggiunge la saturazione, a causa della congestione sulla porta di ingresso, invia un pacchetto Pause sulla porta interessata
 - Se la matrice di commutazione è bloccante
 - Se c'è contesa per le porte di uscita
 - No buffer in uscita
 - Buffer in uscita saturi
- Se la matrice di commutazione dello switch è congestionata scatena l'invio un pacchetto Pause su tutte le porte
 - L'invio del pacchetto di Pause avviene solo negli switch con matrice di tipo bloccante







Approccio realistico su switch con buffer in uscita

- Si abita il flow control di tipo asimmetrico solo nelle connessioni tra switch e stazione
 - compensa congestioni temporanee dei buffer d'ingresso delle interfacce di rete delle stazioni
 - soluzione non ideale, ma può essere un buon compromesso in certe condizioni di traffico

