



Link Aggregation - IEEE 802.3ad

Mario Baldi

Politecnico di Torino

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

Pietro Nicoletti

Studio Reti

<http://www.studioreti.it>

Basato sul capitolo 8 di:

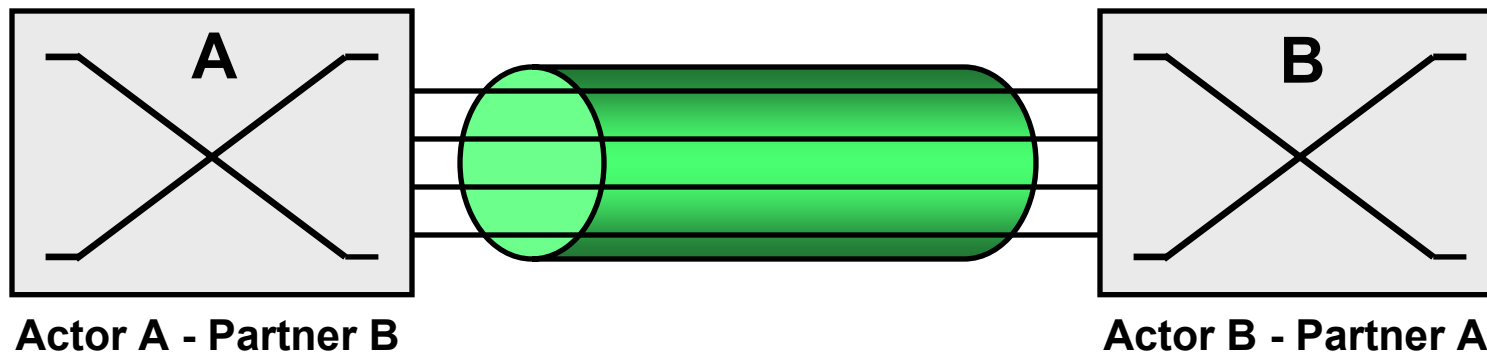
M. Baldi, P. Nicoletti, "Switched LAN", McGraw-Hill, 2002, ISBN 88-386-3426-2

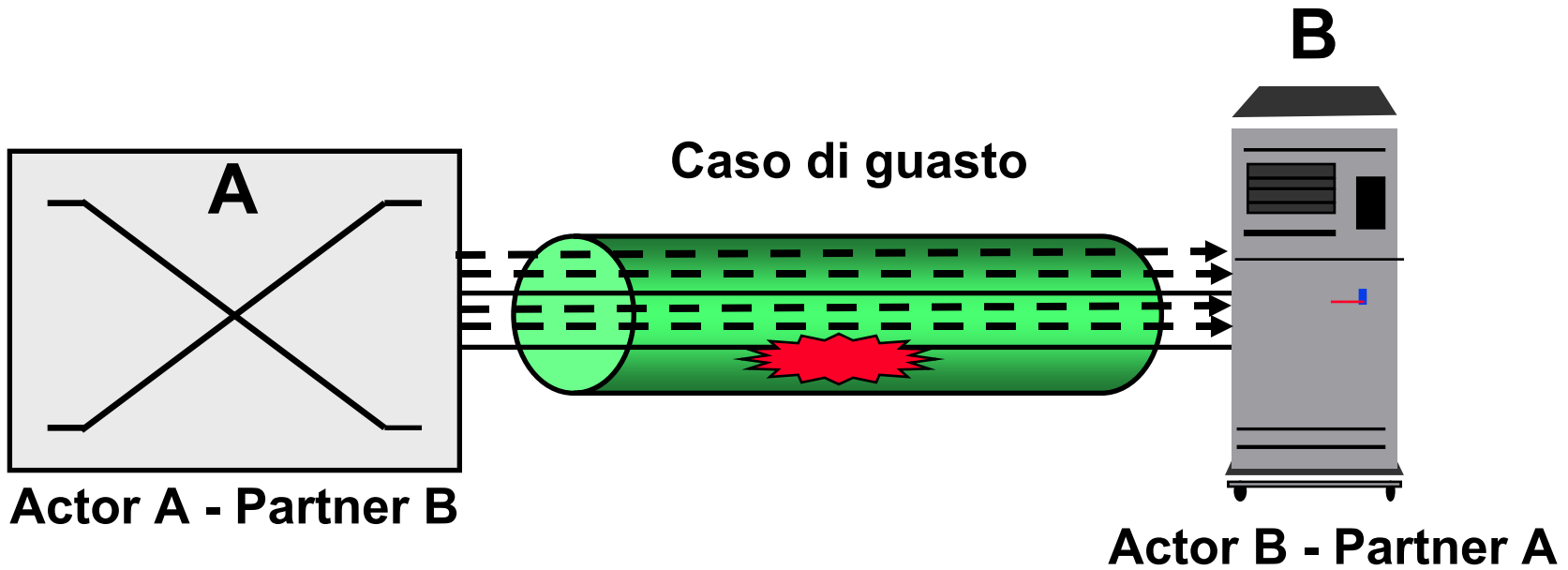
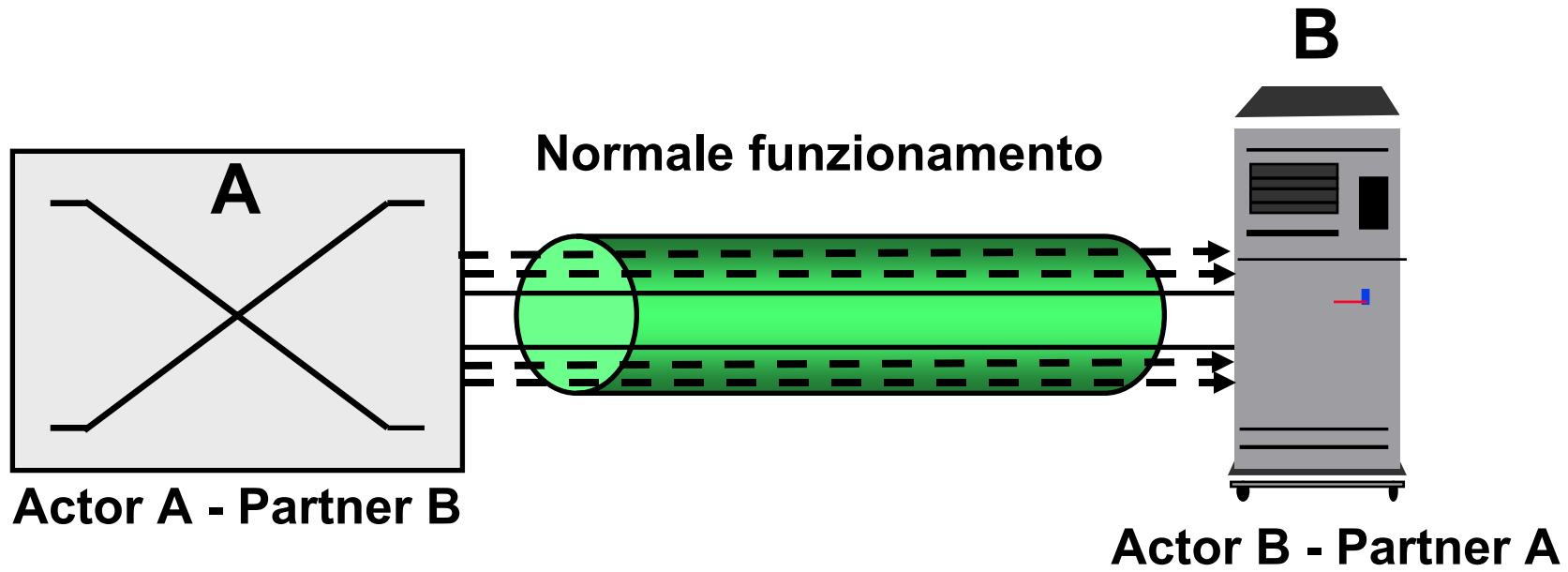
Nota di Copyright

- 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.

IEEE 802.3ad

- Standard che sostituisce soluzioni proprietarie per raggruppare/aggiungere più porte
 - serve per aumentare la banda trasmissiva
 - è in grado di utilizzare in modo efficiente link ridondanti tra due apparati
 - soluzione normalmente applicata su link tra switch, meno frequentemente tra switch e computer





Generalità 802.3ad

- Link aggregation possibile solo su connessioni full duplex
- Aumento della banda
 - link multipli vengono raggruppati in un link logico
 - la crescita della banda è incrementale
- Load sharing
 - traffico di un client che può essere distribuito su link multipli
- Aumento di affidabilità
 - il guasto di un link di un set di aggregazione non pregiudica la comunicazione tra i partner presenti alle due estremità

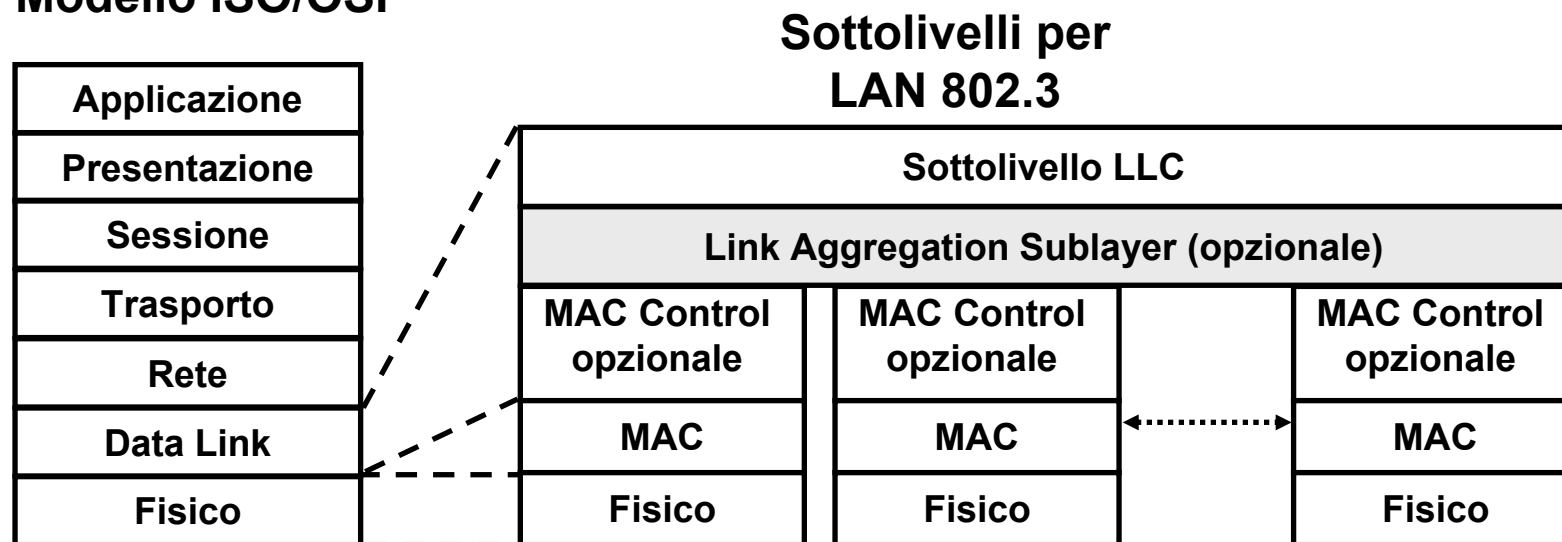
Generalità 802.3ad

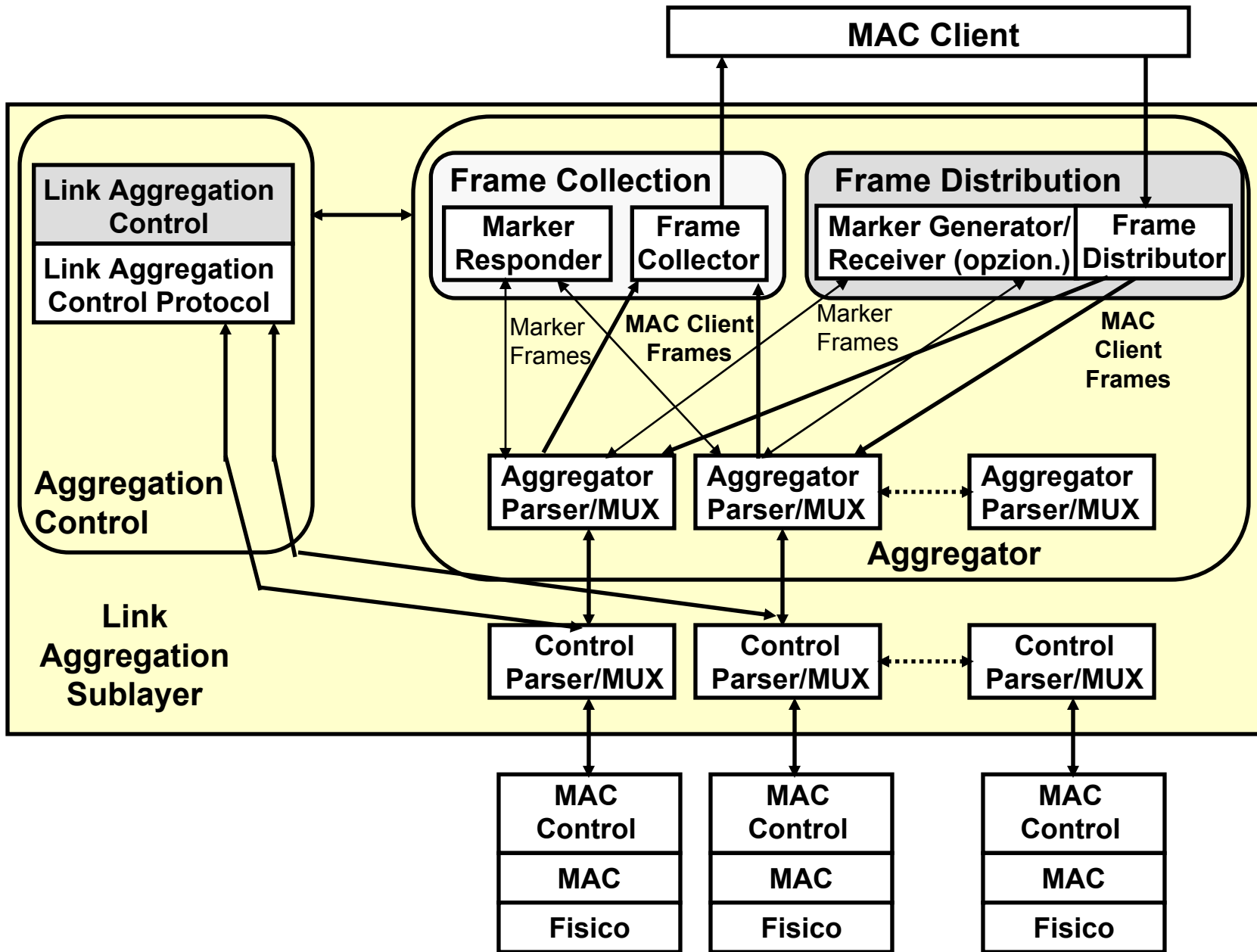
- Convergenza rapida
 - un set di aggregazione può convergere in una nuova configurazione in meno di 1 secondo
- Tutti i link di un gruppo di aggregazione devono avere la stessa velocità
- Configurazione automatica di aggregazione attraverso il protocollo LACP (Link Aggregation Control Protocol)
 - impiego di LACPDU trasmessi in multicast

802.3ad e layer OSI

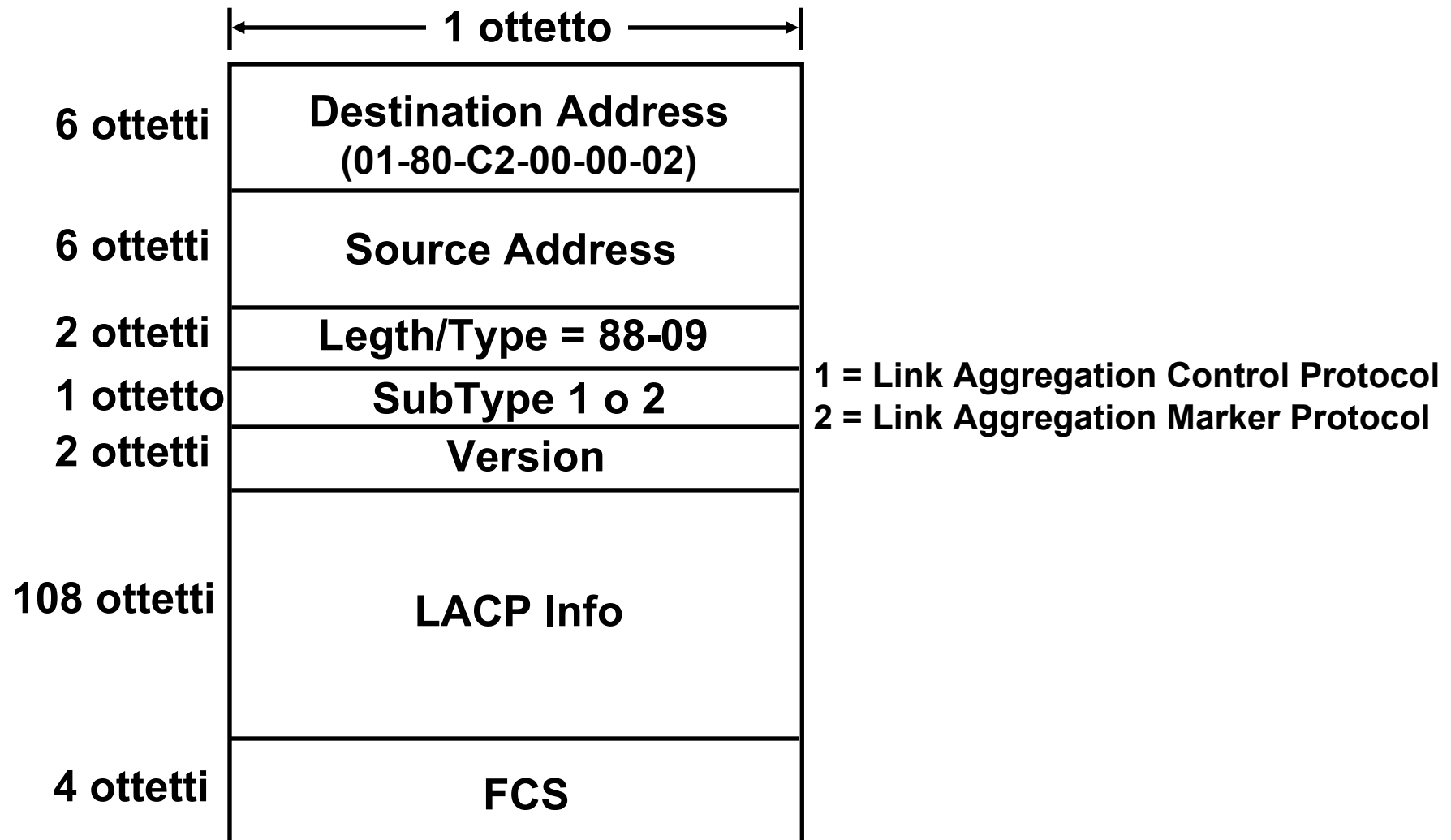
- Link aggregation sublayer si inserisce tra il livello LLC e quelli MAC delle singole porte da aggregare

Modello ISO/OSI





Formato LACPDU



Link Aggregation Group Identifier (LAG ID)

- Serve per stabilire automaticamente l'appartenenza allo stesso gruppo di aggregazione di due porte di un link
 - il LAG ID completo è costituito dal LAG ID locale e da quello remoto
 - Actor LAG ID è quello locale
 - Partner LAG ID è quello remoto

Link Aggregation Group Identifier (LAG ID)

- Il LAG ID è costituito dai seguenti parametri:
 - System Identifier (System Priority + MAC Address)
 - Operational Key assegnata alle porte nel LAG
 - Port Identifier (Port Priority + numero di porta)
 - parametro non necessario in alcune circostanze che viene impostato, in tal caso, al valore zero

Link Aggregation Group Identifier (LAG ID)

- Per stabilire l'appartenenza allo stesso gruppo di aggregazione, tra due porte all'estremità del link, nel pacchetto LACP viene codificato il LAG ID locale (Actor) e quello remoto (Partner)
 - variabili S e T per il System ID locale e remoto
 - variabili K e L per il valore Operational Key locale e remoto
 - variabili P e Q per il Port ID locale e remoto

Esempio di Partner parameter per costruire il LAG ID completo del link

| | Partner SKP | Partner TLQ |
|--------------------------|--|--|
| System Parameters (S, T) | System Priority = 0x8000 (see 43.4.2.2) System Identifier = AC-DE-48-03-67-80 | System Priority = 0x8000 (see 43.4.2.2) System Identifier = AC-DE-48-03-FF-FF |
| Key Parameter (K, L) | Key = 0x0001 | Key = 0x00AA |
| Port Parameters (P, Q) | Port Priority = 0x80 (see 43.4.2.2) Port Number = 0x0002 | Port Priority = 0x80 (see 43.4.2.2) Port Number = 0x0002 |

The complete LAG ID derived from this information is represented as follows, for an Individual link:

[(SKP), (TLQ)] = [(8000,AC-DE-48-03-67-80,0001,80,0002), (8000,AC-DE-48-03-FF-FF,00AA,80,0002)]

Distribuzione dei pacchetti sulle porte aggregate

- Lo standard non stabilisce l'algoritmo da adottare per distribuire i pacchetti sulle varie porte
 - non ammette la segmentazione ed il riassettaggio dei pacchetti
 - su una porta ci possono essere più conversazioni
 - una conversazione può essere spostata tra le porte di un aggregato per ragioni di load balancing o *in caso di guasto su un link*
 - si può creare un'aggregazione costituita da un singolo link o più link
 - lo standard elenca i possibili criteri di distribuzione dei pacchetti sulle varie porte

Distribuzione dei pacchetti sulle porte aggregate

- Lo standard non stabilisce l'algoritmo da adottare per distribuire i pacchetti sulle varie porte
 - due switch di produttori differenti possono usare algoritmi diversi di distribuzione dei pacchetti
 - questo fatto non ha alcun impatto sulla funzionalità
 - si potrebbe avere una distribuzione non ottimale
 - è preferibile che gli switch posti all'estremità di un'aggregazione di link siano dello stesso produttore

Criteri di distribuzione dei pacchetti sulle porte aggregate

- Assegnazione delle conversazioni in base:
 - Source MAC Address
 - Destination MAC Address
 - porta di ricezione
 - tipo di destinazione (unicast, multicast, broadcast)
 - valore contenuto nel campo Length/Type
 - Higher Layer Protocol (es. L4 porte)
 - combinazione dei precedenti criteri

Criterio di distribuzione di Cisco

| Coppie di indirizzi MAC Source e Destination | Ultimi 2 bit | Risultato X-OR | Link Scelto |
|---|-----------------|-------------------|----------------|
| Source MAC Address 00-00-00-00-00-01 Destination MAC Addr. 00-00-00-00-00-04 | 01 00 | 01 | Link 2 |
| Source MAC Address 00-00-00-00-00-02 Destination MAC Addr. 00-00-00-00-00-05 | 10 01 | 11 | Link 4 |
| Source MAC Address 00-00-00-00-00-03 Destination MAC Addr. 00-00-00-00-00-07 | 11 11 | 00 | Link 1 |
| Source MAC Address 00-00-00-00-00-06 Destination MAC Addr. 00-00-00-00-00-08 | 10 00 | 10 | Link 3 |

Standby link

- Possibilità di abilitare dei link per operare come backup di altri
- Link appartenenti a due gruppi di aggregazione
 - quello con priorità maggiore diventa attivo
 - quello con priorità minore diventa standby
- Quando il numero di link aggregabili tra due apparati è *inferiore* a quelli presenti ed ai quali è stata assegnata la stessa chiave amministrativa:
 - vengono selezionati come attivi i link più prioritari e gli altri vengono considerati standby

Standby link e assegnazione dinamica delle chiavi

- Assegnazione della chiave amministrativa ai link che sono attivi
 - ai link standby viene assegnata una chiave differente da quelli attivi
 - un link standby utilizza una nuova chiave ma tiene traccia della chiave del gruppo di link attivo di cui è backup
 - in caso di guasto di un link attivo si deve formare un nuovo gruppo di aggregazione che escluda il link guasto ed includa quello che era standby
 - riassegnazione della chiave al nuovo gruppo formato

Assegnazione dinamica delle chiavi

- I due apparati usano la stessa chiave sull'aggregato di porte
- La chiave viene assegnata dall'apparato che ha la priorità più alta
 - il criterio di priorità è identico a quello adottato dallo spanning tree
 - system priority (bridge priority nel caso di switch) + indirizzo MAC
 - è prioritario il valore risultante più basso
 - esempio 8000-08-00-2B-50-20-00 è prioritario rispetto a 8000-08-00-2B-C4-E6-AA

Standby e equal level cost load sharing

- Si assegna una chiave amministrativa unica a tutte le porte, in pratica si riuniscono insieme tutte le porte, sebbene possa essere aggregato un numero inferiore
 - una porta può essere attiva per un gruppo e standby per un altro gruppo
 - l'aggregato prioritario diventa immediatamente attivo utilizzando la chiave amministrativa
 - priorità = System-ID - Port ID
 - dopo pochi secondi l'aggregato meno prioritario diventa attivo utilizzando un'altra chiave

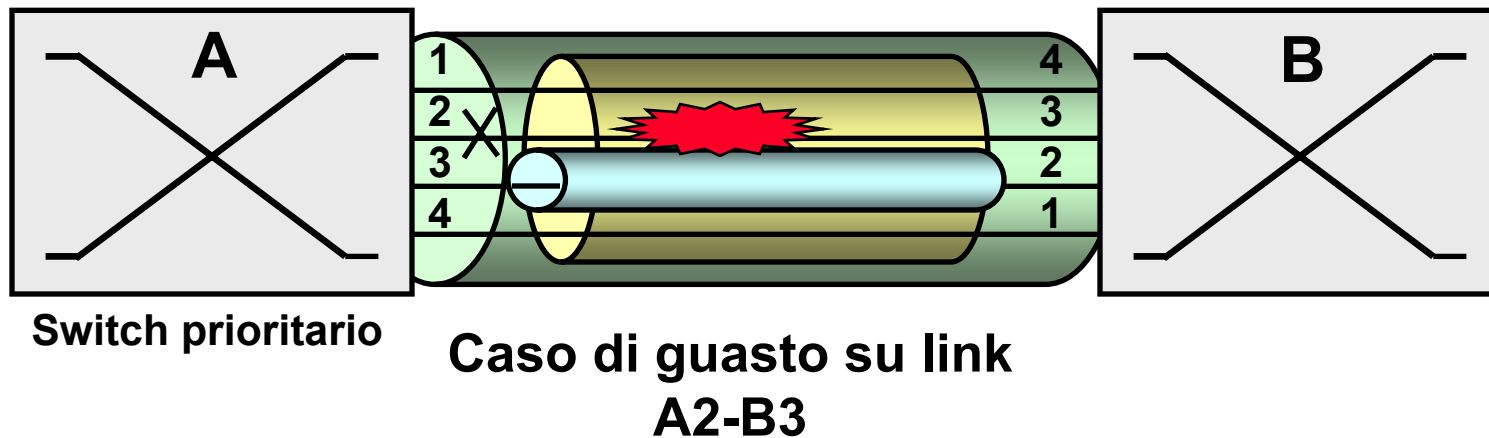
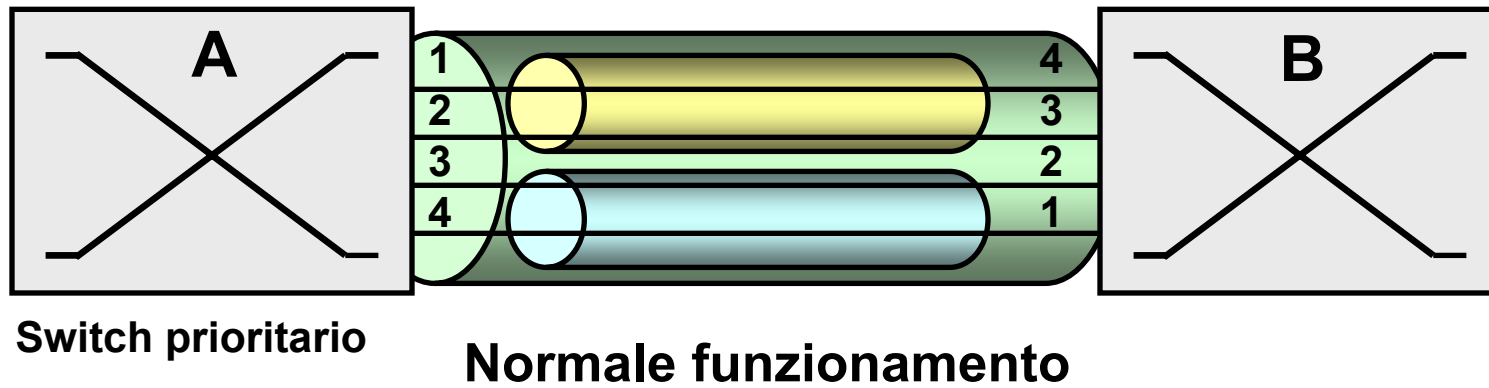
Esempio 1: standby e equal level cost load sharing

- Esempio 1: 4 link paralleli di cui solo 2 possono essere aggregati
 - chiave di aggregazione = 1
- Switch A prioritario
- Regole di aggregazione e backup su Switch A:
 - porta 1 con porta 2 o 4
 - porta 2 con porta 3 o 1
 - porta 3 con porta 4 o 2
 - porta 4 con porta 1 o 3

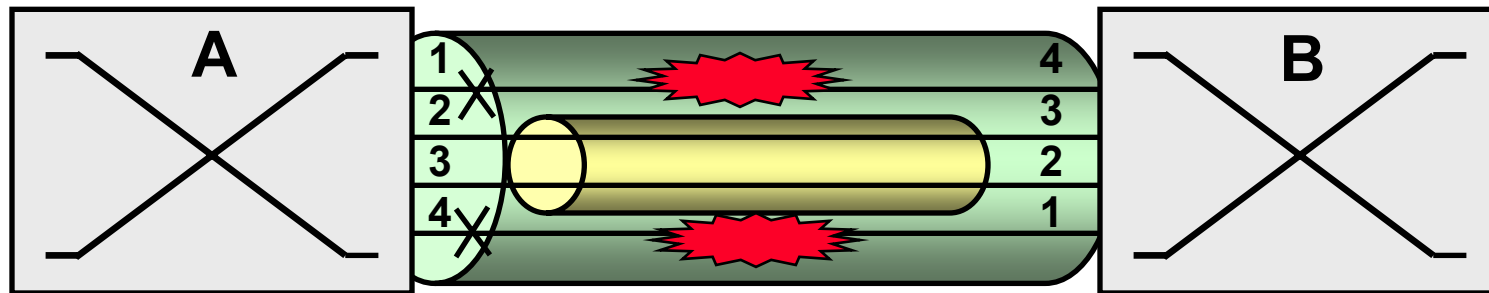
Esempio 1: standby e equal level cost load sharing

- Come opera l'algoritmo:
 - seleziona i due link più prioritari come attivi
 - assegna la chiave = 1 ai link A1-B4 e A2-B3
 - considera dapprima i restanti link come standby
 - mantiene traccia della chiave 1 sul link A4-B1 ed effettua il cambio chiave sui link A3-B2 e A4-B1
 - dopo pochi secondi attiva i link A3-B2 e A4-B1 come secondo aggregato di porte
 - load sharing sui due aggregati

Esempio 1: condizioni di lavoro e di guasto

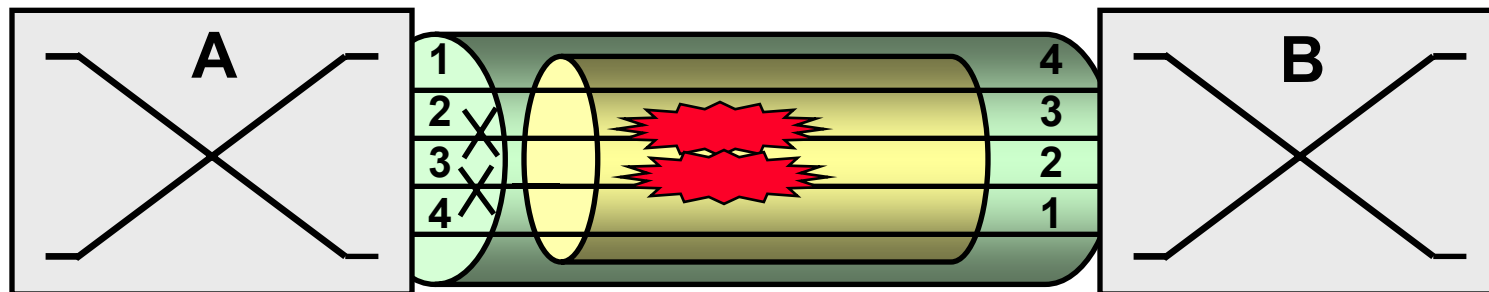


Esempio 1: condizioni di guasto



Switch prioritario

Guasti su link
A1-B4 e A4-B1



Switch prioritario

Guasti su link
A2-B3 e A3-B2

Link Aggregation e STP/RSTP

- Non pregiudica l'impiego di STP o RSTP
- Impostare i valori di Path cost alle porte che riflettano il valore di banda aggregata
 - disabilitare l'impostazione automatica di path cost sulle porte
 - preferibile usare RSTP perché ha un range più ampio di valori di Path Cost associabili alle porte

Link Aggregation e STP/RSTP

