

Rapid spanning tree protocol

Mario Baldi

Politecnico di Torino

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

Pietro Nicoletti

Studio Reti

<http://www.studioreti.it>

Basato sul capitolo 4 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 dell'Istruzione, dell'Università e della Ricerca, per scopi istituzionali, non a fine di lucro. In tal caso non è richiesta alcuna autorizzazione.

Ogni altro utilizzo 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 dell'edizione. 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.

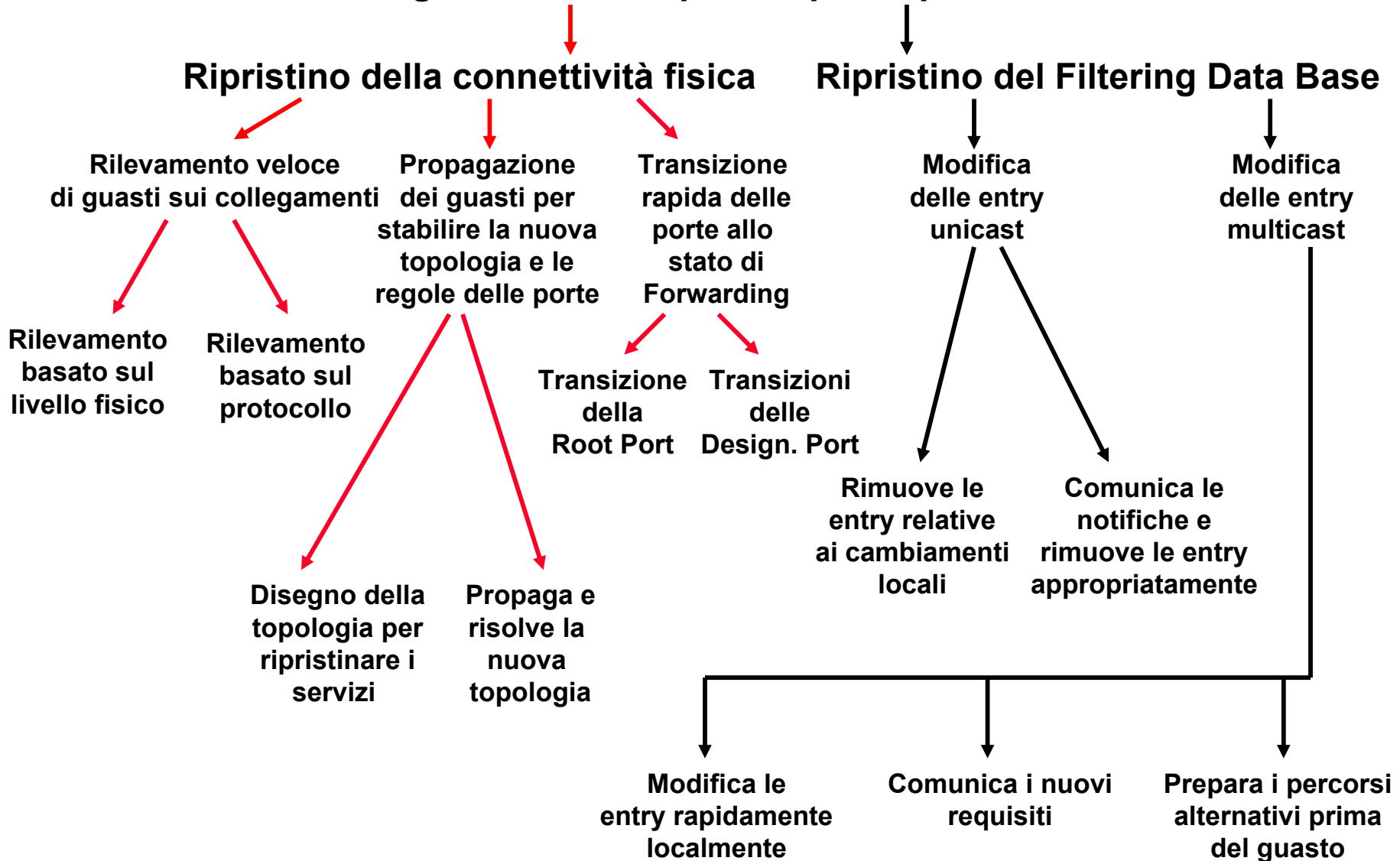
In ogni caso questa nota di copyright non deve mai essere rimossa e deve essere riportata anche in utilizzi parziali.

Rapid Spanning Tree Protocol (RSTP)

- Spanning tree con convergenza veloce
- Standard IEEE 802.1w approvato a fine anno 2001
- Interoperabile con lo spanning tree IEEE 802.1D
 - Senza convergenza veloce
- Soluzione di livello 2 ideale per reti mission critical
- Opera solamente in presenza di link punto-punto
 - Collegamenti diretti (no hub)
- Sostituisce soluzioni proprietarie per resiliency con convergenza veloce adottate da alcuni costruttori

Migliorie rispetto STP

Migliorie adottate per il rapido ripristino dei servizi



Ruolo delle porte

- Alternate
 - offre un percorso alternativo in direzione del Root Bridge
 - porta alternativa alla Root Port da utilizzare in caso di guasto
- Backup
 - entra in funzione a seguito del guasto della Designated port
 - presente solo se due o più porte dello stesso bridge connettono una LAN
- Root
- Designated
- Disabled

Alternate port e Backup port

- IEEE 802.1w non prevede lo stato di Blocking
- Le porte connesse a percorsi alternativi verso la radice vengono selezionate come Alternate
- Più porte dello stesso bridge collegate alla stessa LAN:
 - una Designated
 - le altre Backup
 - connessione punto-punto tra due porte di un bridge
 - due o più porte di un bridge collegate ad un repeater
 - Le due porte sono collegate alla stessa LAN

Stato delle porte: relazioni tra STP e RSTP

STP - stato delle porte	Stato amministrativo delle porte	Operatività del MAC	RSTP - stato delle porte	Topologia attiva RSTP (ruolo delle porte)
DISABLED	Disabled	FALSE	Discarding	Excluded (Disabled)
DISABLED	Enabled	FALSE	Discarding	Excluded (Disabled)
BLOCKING	Enabled	TRUE	Discarding	Excluded (Alternate, Backup)
LISTENING	Enabled	TRUE	Discarding	Included (Root, Designated)
LEARNING	Enabled	TRUE	Learning	Included (Root, Designated)
FORWARDING	Enabled	TRUE	Forwarding	Included (Root, Designated)

Simboli grafici convenzionali

Legenda simboli grafici	
Simbolo grafico	Significato
<p>Bridge Identifier</p> <p>Port Identifier e Port Cost</p> <p>Root Identifier</p> <p>Root Path Cost</p>	BRIDGE
<p>A</p>	LAN

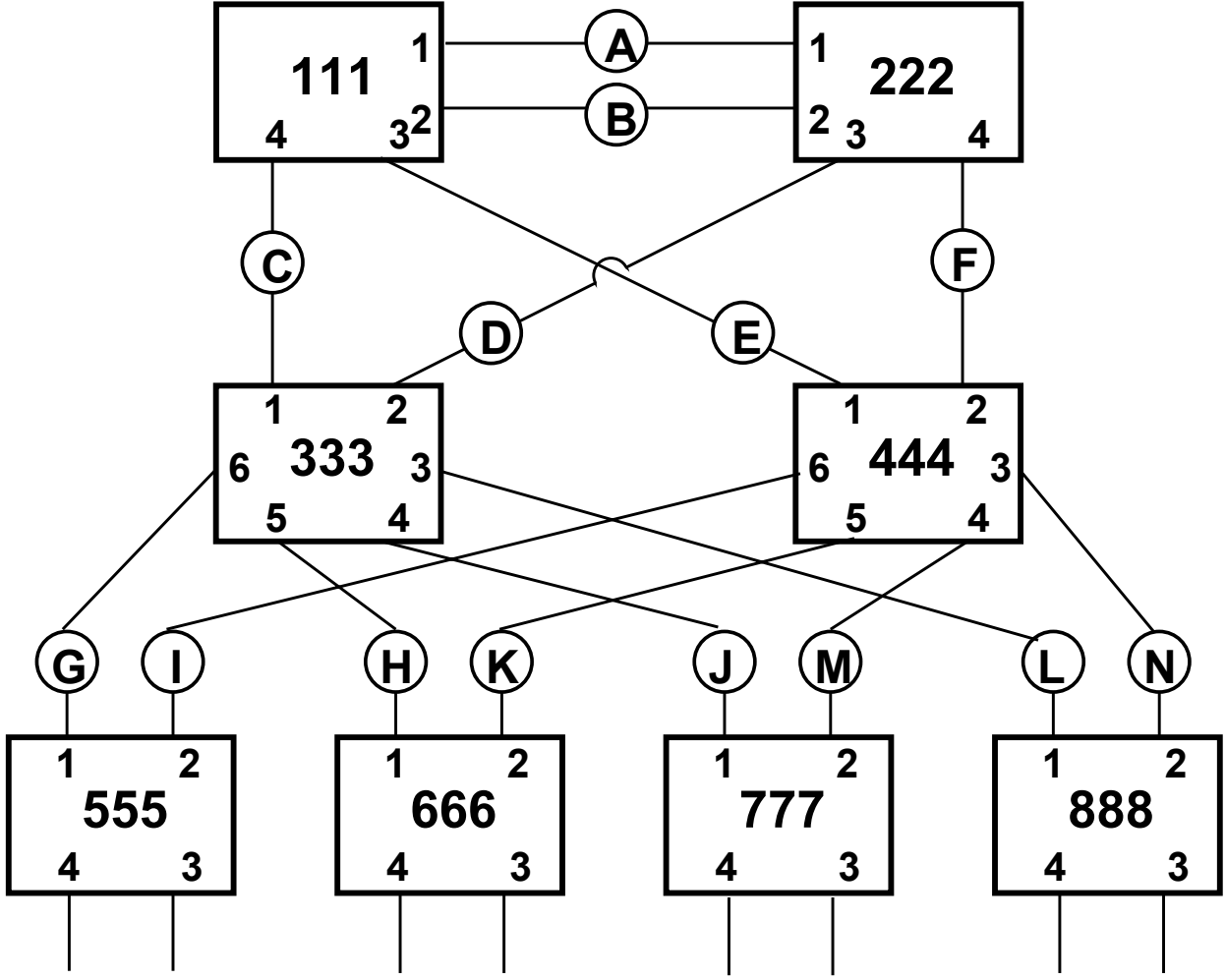
Simboli grafici convenzionali

Ruolo della porta	Stato della porta	Legenda
Designated su porta ridondata ----- su porta terminale	Discarding Learning Forwarding	●+———
		●++———
	Forwarding	●◇———
Root Port	Discarding Learning Forwarding	○+——— ○++——— ○———
Alternate	Discarding	—++———
Backup	Discarding	—+++———
Disabled	-----	—/———
BPDU trasmesse		
Designated		————→
Designated Proposal		————→→
Root		————▷
Root Proposal		————▷▷

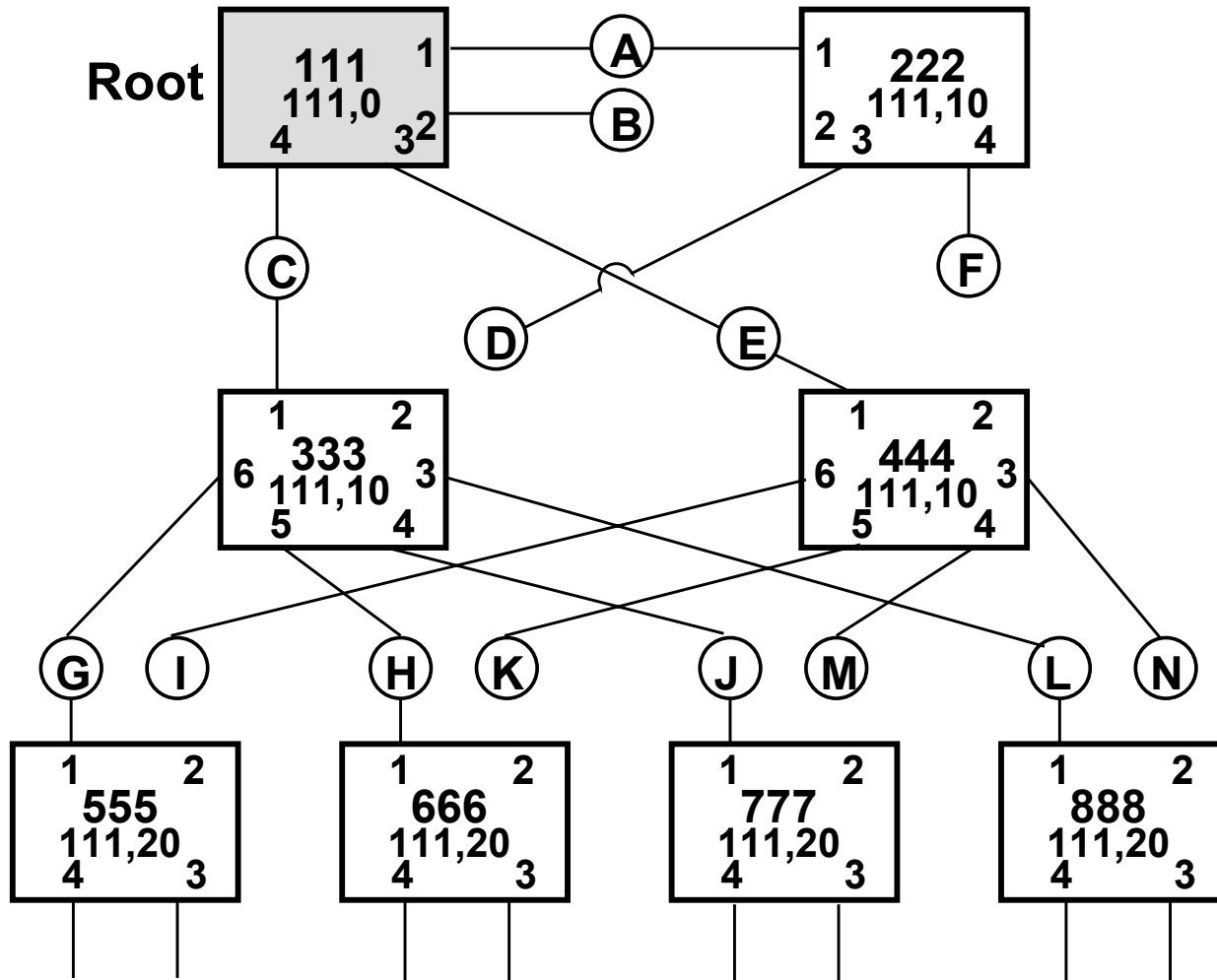
Principi di funzionamento

- L'albero viene creato come nello STP tradizionale
 - Elezione del root bridge
 - Definizione della root port
 - Definizione delle designated port
- Le porte rimanenti (né root né designated)
 - Alternate se collegate ad una porta di un altro bridge
 - Backup se collegate ad una porta dello stesso bridge

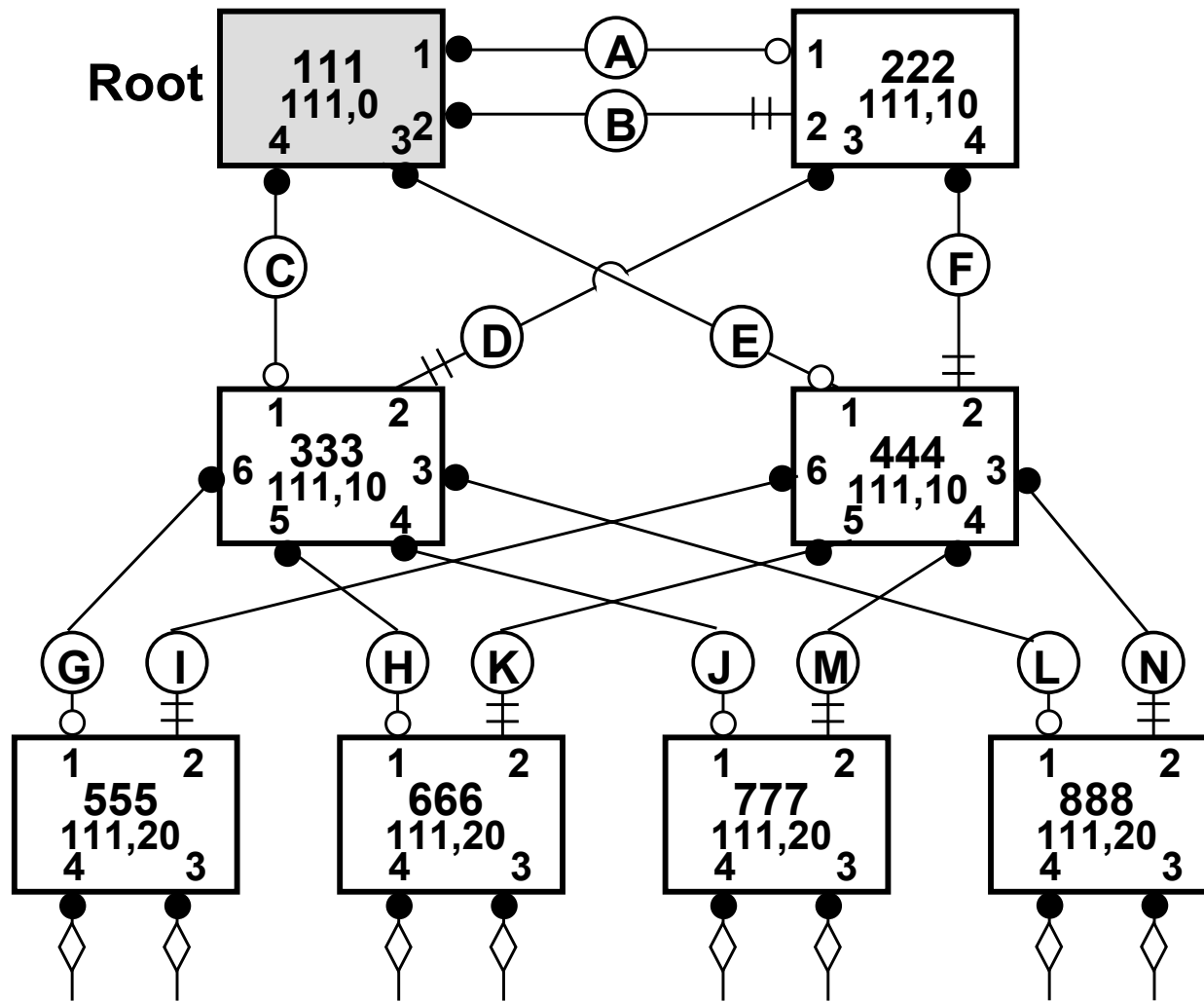
Esempio: topologia fisica



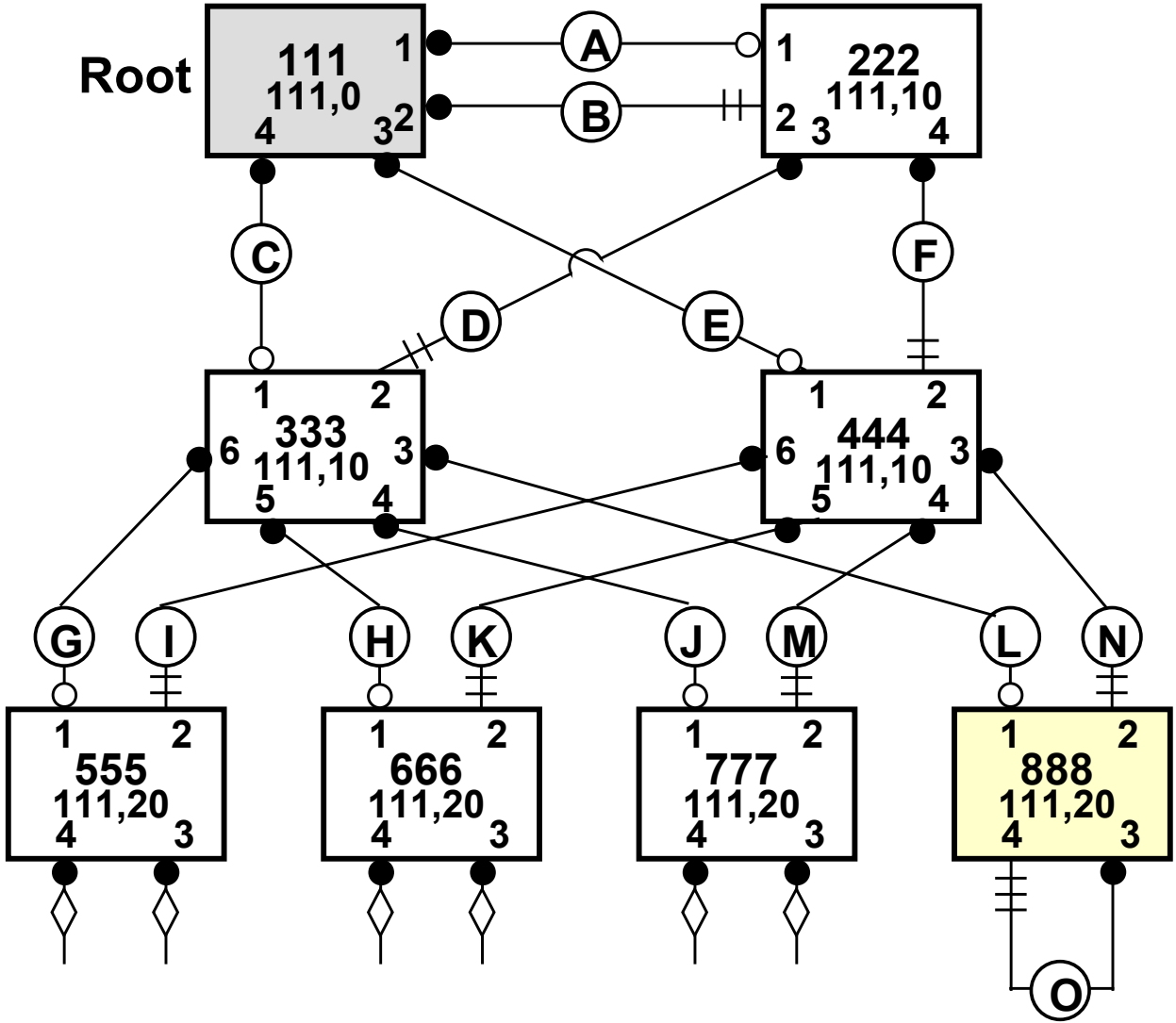
Topologia attiva RSTP



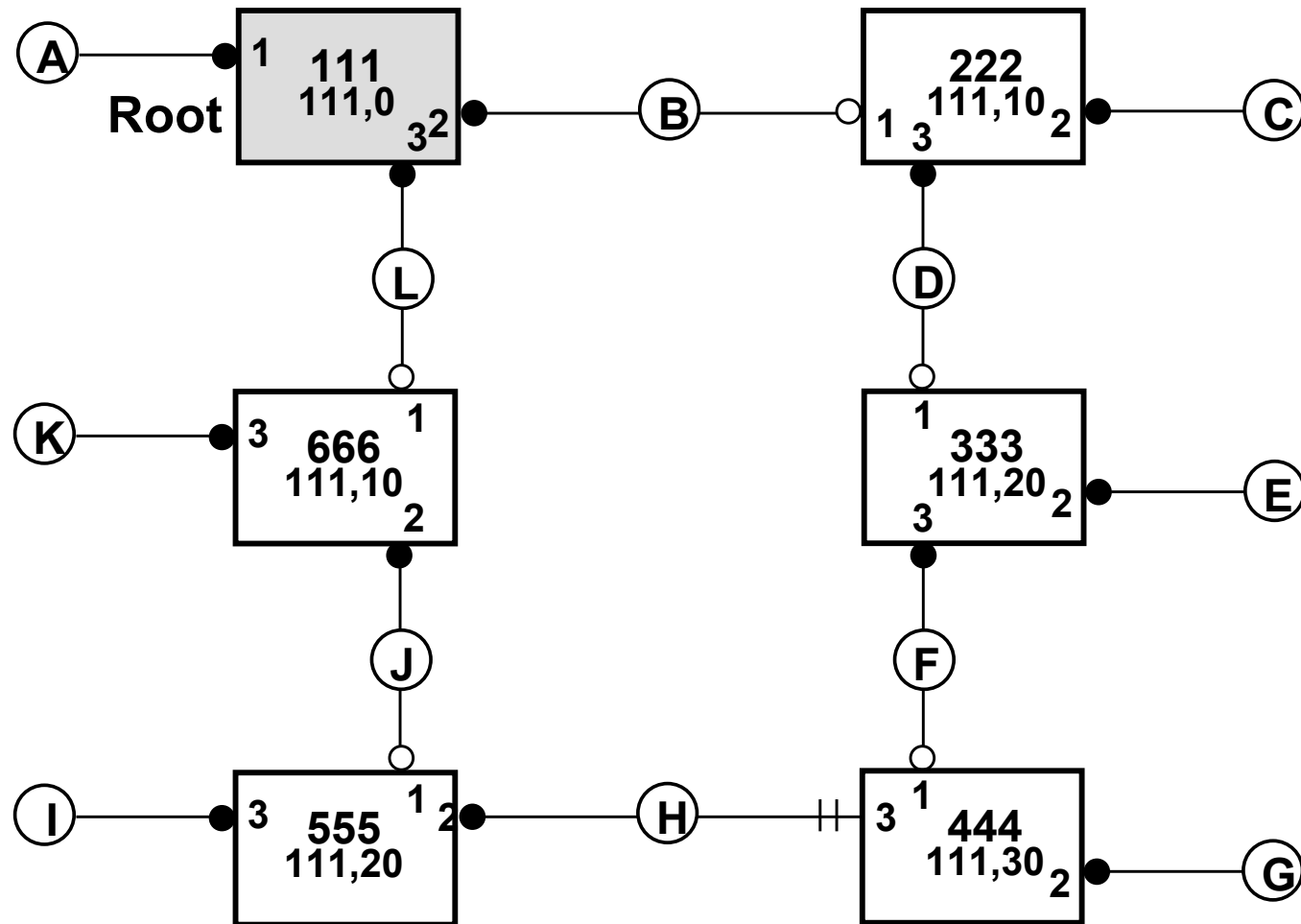
Stato delle porte



Porta backup



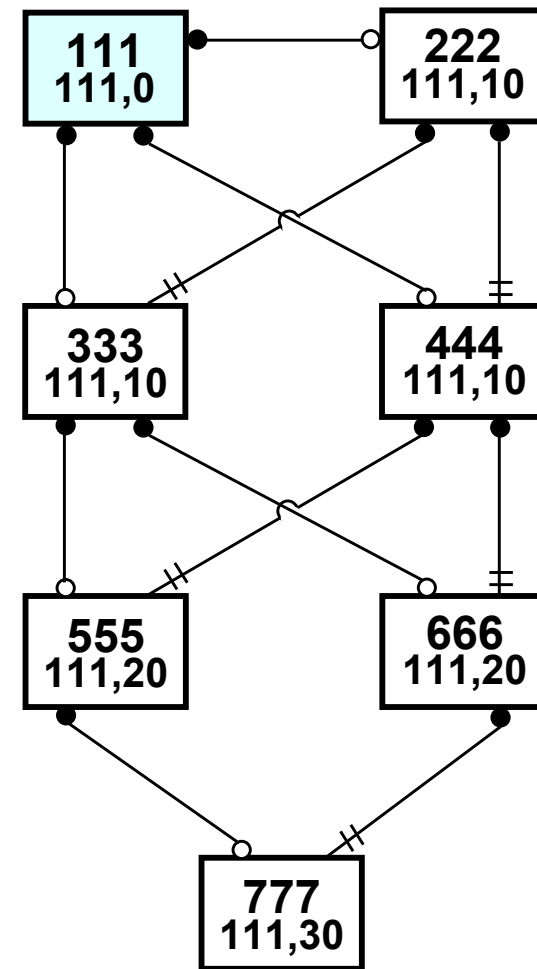
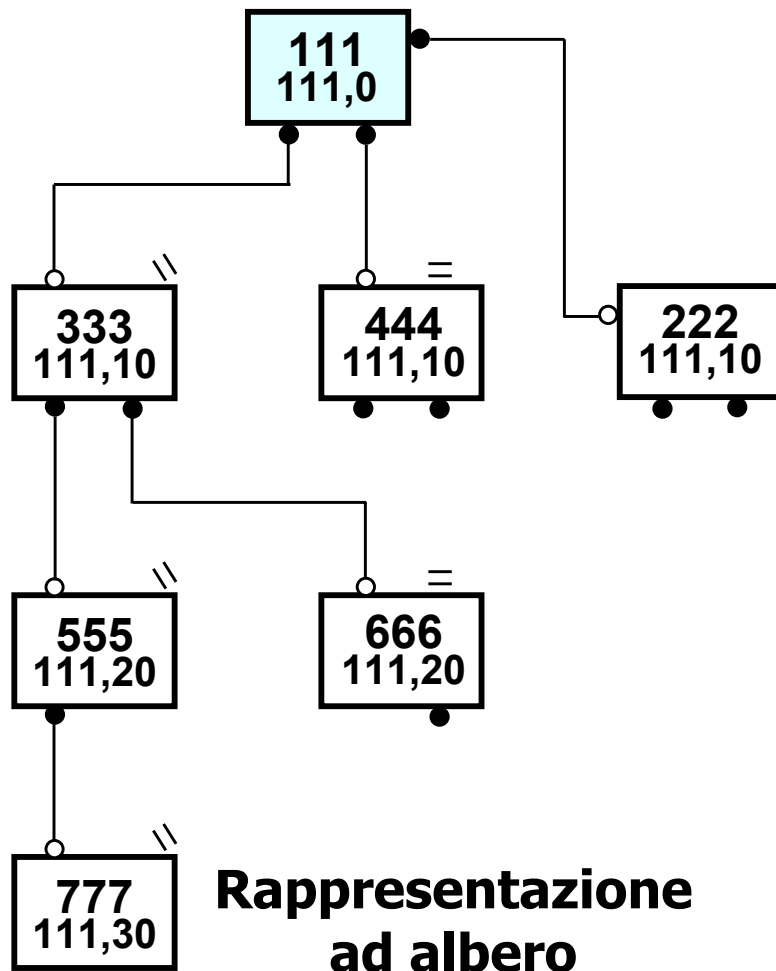
Dorsale ad anello



Reazione ai guasti

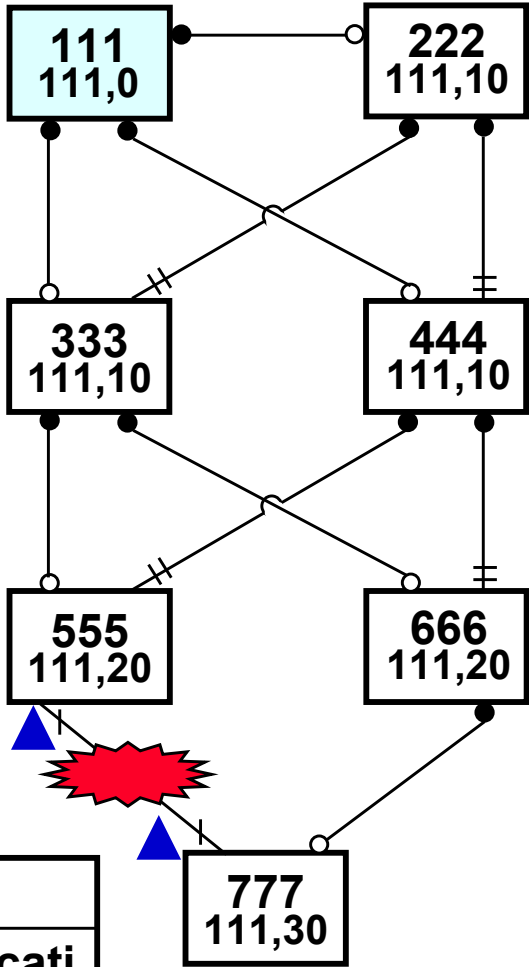
- Eliminazione entry relative alla porta guasta
 - I due bridge agli estremi di un link guasto
 - I bridge direttamente collegati ad un bridge guasto
- Se si guasta la root port una porta alternate diviene root
- Invio di un messaggio ***Topology Change Notification*** (TCN)
 - Su tutte le porte in stato forwarding
- Alla ricezione di un messaggio TCN
 - Rimozione delle entry di tutte le porte, eccetto quella da cui il messaggio TCN è ricevuto
 - Invio di un messaggio TCN sulle porte in stato forwarding
 - Eccetto quella da cui il messaggio TCN è ricevuto

Flush entry: configurazione iniziale



Flush entry: fase 1

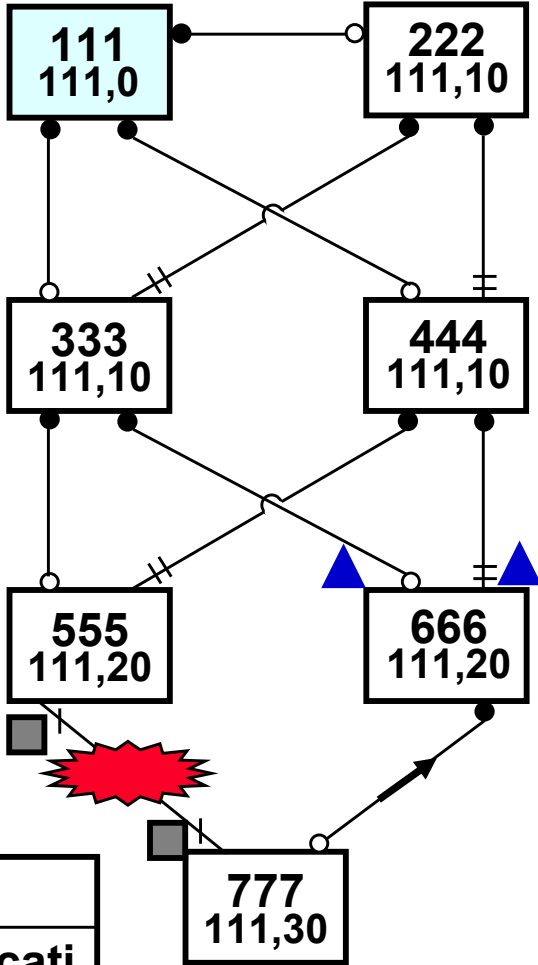
- Guasto sul link tra 555 e 777
- I bridge 555 e 777 rilevano il guasto
- Rimuovono (flush) le entry relative alle porte coinvolte



Legenda simboli relativi alle fasi di flush	
▲	Indirizzi appresi sulle porte che devono essere scaricati
■	Indirizzi appresi sulle porte che sono stati scaricati
→	Pacchetti TCN trasmessi in direzione della freccia

Flush entry: fase 2

- I bridge 555 e 777 hanno rimosso le entry relative alle porte coinvolte nel guasto
- Il bridge 777 fa diventare root la porta alternate
- Il bridge 777 invia un messaggio TCN sulla nuova root port
 - Il messaggio TCN ha lo stesso effetto di una Configuration-BPDU con il flag TC impostato al valore 1

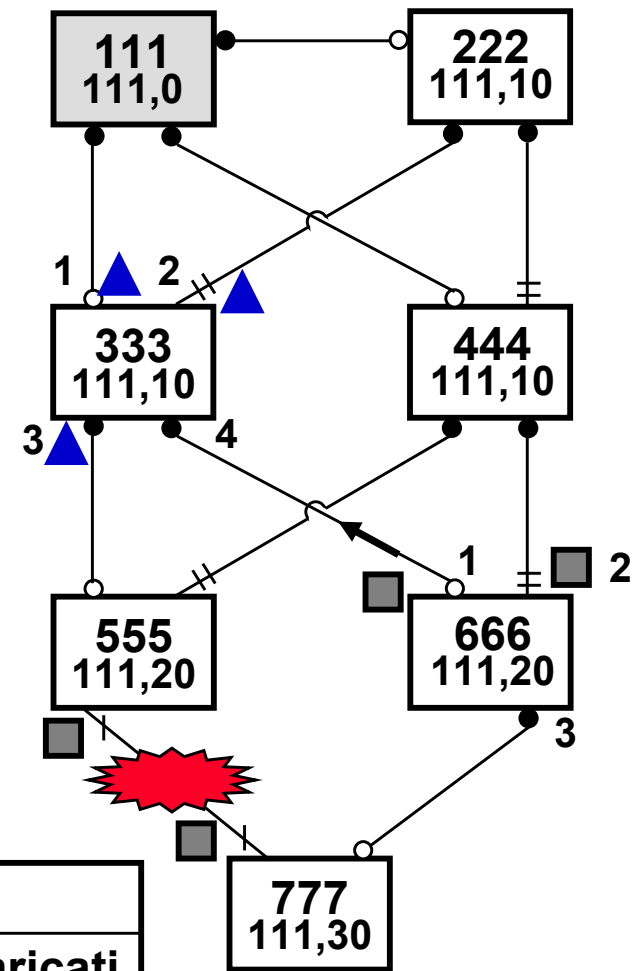


Legenda simboli relativi alle fasi di flush	
▲	Indirizzi appresi sulle porte che devono essere scaricati
■	Indirizzi appresi sulle porte che sono stati scaricati
→	Pacchetti TCN trasmessi in direzione della freccia

Flush entry: fase 3

Il bridge 666

- Riceve il messaggio TCN su una porta e
- Rimuove le entry relative a tutte le altre porte
 - Porta 1 e porta 2
- Invia il messaggio TCN su ogni altra porta in stato forwarding
 - La porta 1 verso il bridge 333



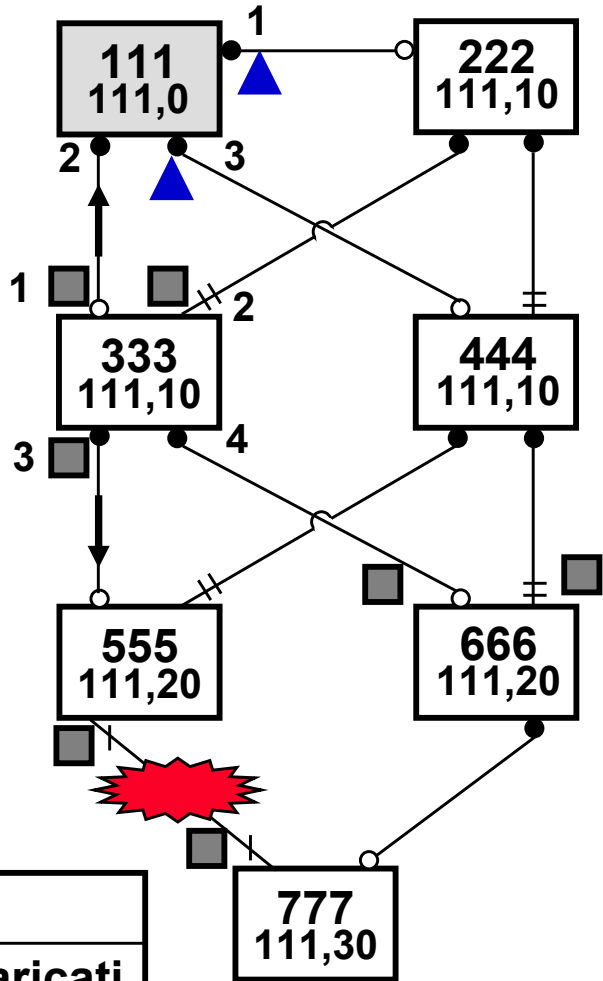
Legenda simboli relativi alle fasi di flush

- ▲ Indirizzi appresi sulle porte che devono essere scaricati
- Indirizzi appresi sulle porte che sono stati scaricati
- Pacchetti TCN trasmessi in direzione della freccia

Flush entry: fase 4

Il bridge 333

- Riceve il messaggio TCN sulla porta 4
- Rimuove le entry relative a tutte le altre porte
 - Porta 1, porta 2 e porta 3
- Invia il messaggio TCN su tutte le sue porte in stato forwarding
 - Porta 1 verso il bridge 111
 - Porta 3 verso il bridge 555

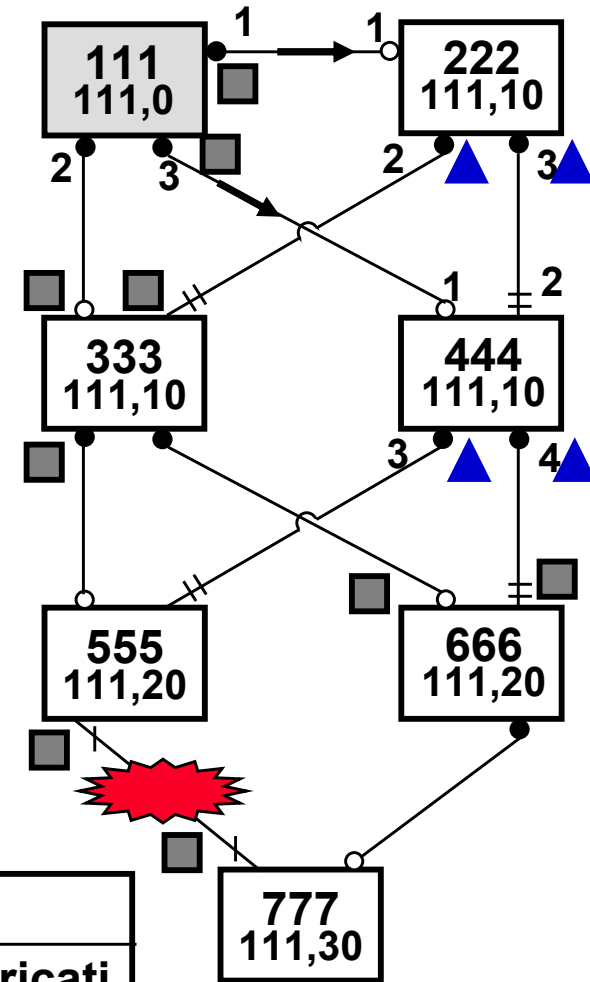


Legenda simboli relativi alle fasi di flush	
▲	Indirizzi appresi sulle porte che devono essere scaricati
■	Indirizzi appresi sulle porte che sono stati scaricati
→	Pacchetti TCN trasmessi in direzione della freccia

Flush entry: fase 5

Il bridge 111

- Riceve il messaggio TCN sulla porta 2
- Elimina le entry relative a tutte le altre porte
 - Porta 1 e porta 3
- Trasmette il messaggio TCN sulle altre porte in stato forwarding
 - Porta 1 verso il bridge 222
 - Porta 3 verso il bridge 444

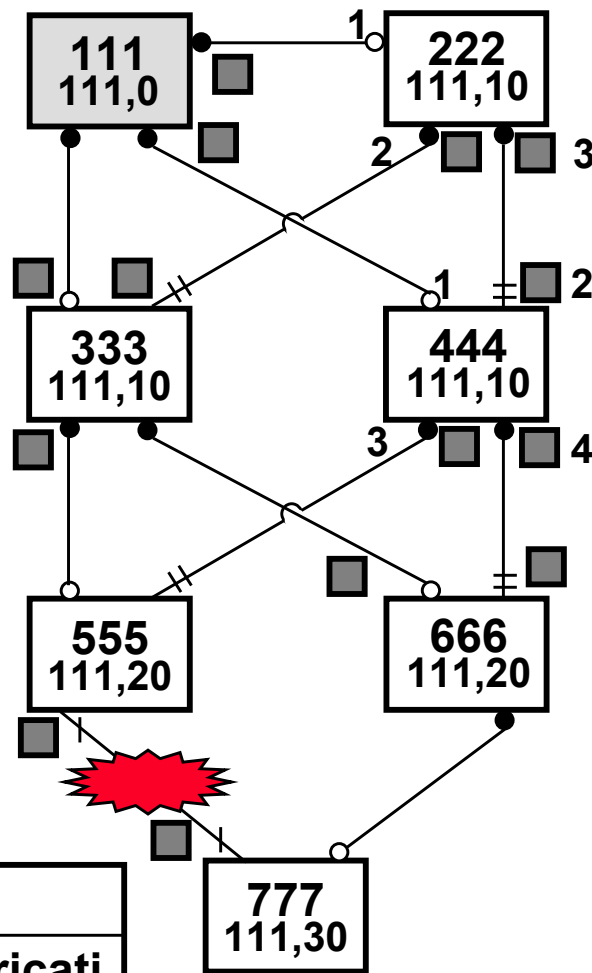


Legenda simboli relativi alle fasi di flush

- ▲ Indirizzi appresi sulle porte che devono essere scaricati
- Indirizzi appresi sulle porte che sono stati scaricati
- Pacchetti TCN trasmessi in direzione della freccia

Flush entry: fase 6

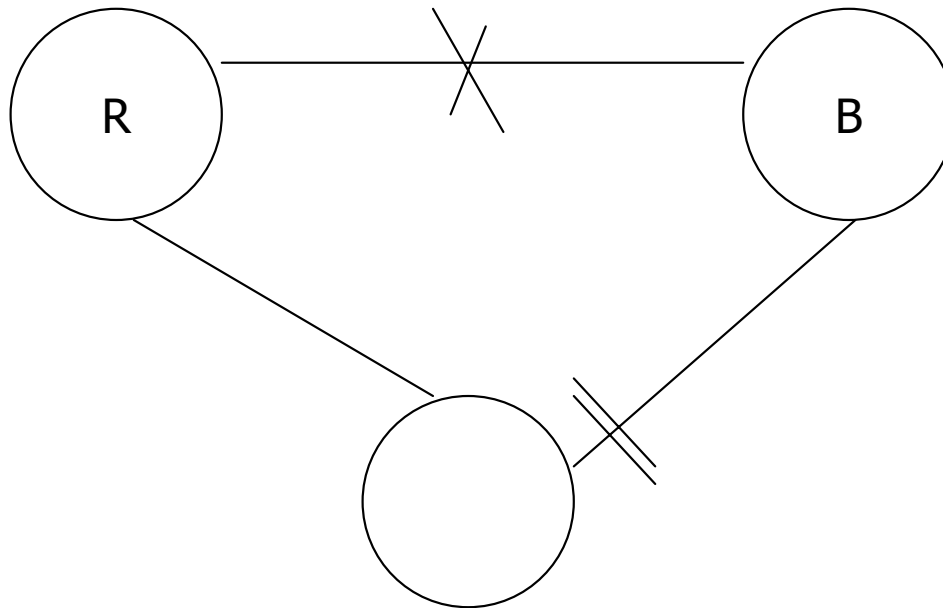
- I bridge 222 e 444
- Ricevono il messaggio TCN su una delle loro porte
- Eliminano le entry relative a tutte le altre porte (si veda figura)
- I messaggi TCN da loro trasmessi raggiungono solo porte in stato discarding ???



Legenda simboli relativi alle fasi di flush	
▲	Indirizzi appresi sulle porte che devono essere scaricati
■	Indirizzi appresi sulle porte che sono stati scaricati
→	Pacchetti TCN trasmessi in direzione della freccia

Meccanismo proposal/agreement

- Un bridge vicino al root, al cadere del link con il root si propone come root.
- Gli altri bidge gli fanno sapere che c'è un altro root e lui agrees.



Path Cost raccomandato da IEEE 802.1w



Velocità porta	Valore raccomandato	Intervallo raccomandato	Intervallo accettabile
<= 100 Kb/s	200.000.000	20.000.000 - 200.000.000	1 - 200.000.000
1 Mb/s	20.000.000	2.000.000 - 200.000.000	1 - 200.000.000
10 Mb/s	2.000.000	200.000 - 20.000.000	1 - 200.000.000
100 Mb/s	200.000	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	2000	200 - 20.000	1 - 200.000.000
100 Gb/s	200	20 - 2000	1 - 200.000.000
1 Tb/s	20	2 - 200	1 - 200.000.000
10 Tb/s	2	1 -20	1 - 200.000.000

DSAP	SSAP	Length	DSAP	SSAP	Control	
Multicast 01-80-C2-00-00-00	Singlecast Bridge Address	XY	042H	042H	XID	Configuration message

00000000 = STP
 00000010 = RSTP

Byte 1÷2

3

4

5

6÷13

14÷17

18÷25

26÷27

28÷29

30÷31

32÷33

34÷35

Protocol Identifier		00-00		
Protocol Ver. Identifier		00		
BPDU Type		00		
TC	Flags			TCA
	P	PR	PR	A
Root Identifier				
primi 2 byte = Bridge Priority				
successivi 6 byte = indirizzo MAC del Root Bridge				
Root Path Cost				
Bridge Identifier				
primi 2 byte = Bridge Priority				
successivi 6 byte = indirizzo MAC del Bridge trasmittente la BPDU				
Port Identifier				
Primo byte = Port Priority				
Secondo Byte = numero di porta				
Message Age				
Max Age				
Hello Time				
Forward Delay				

Configuration BPDU STP e RSTP

RSTP e compatibilità con STP

- I bridge RSTP possono essere configurati ad operare in modalità STP
 - Fondamentale in presenza di ripetitore tra due bridge
 - Funzionamento in modalità RSTP potrebbe portare a instabilità e loop temporanei
- I bridge RSTP (IEEE 802.1w) operano automaticamente in modalità STP (IEEE 802.1D) in presenza di 1 o più bridge operanti in modo STP
 - Cambiamento di modalità alla ricezione di BPDU con valore 0 nel campo protocol version identifier
 - L'operatività contemporanea in modalità STP e RSTP può portare a instabilità nella definizione dello spanning tree
 - Possibile duplicazione di pacchetti
 - Possibile recapito di pacchetti fuori sequenza

Rapidità di convergenza di RSTP

- Un percorso alternativo potrebbe essere attivato in 10 - 20 ms
- La rapidità di convergenza dipende da
 - Capacità di rilevare un guasto in modo affidabile
 - Capacità di rilevare un guasto in tempi brevi
 - Criteri di livello fisico
- A questo fine concorrono i sistemi di trasmissione utilizzati
 - Elevata stabilità dovuta alla componentistica elettronica molto affidabile
 - Guasti intermittenti creerebbero problemi di stabilità
 - Transceiver con funzionalità di rilevamento guasto locale e remoto su collegamenti punto-punto

Problemi dovuti alla rapidità di convergenza

- Aumento della probabilità di duplicazione di pacchetti
- Aumento delle probabilità di consegna di pacchetti fuori sequenza
- Dovuto alla probabilità non nulla di avere
 - Percorsi chiusi (loop) temporanei
 - Cambiamenti intermittenti dello spanning tree

Attenzione con protocolli che assumono l'uso di Ethernet!

- LAT, NETBEUI e LLC2
 - Non sono in grado di gestire pacchetti duplicati e fuori ordine
- Consigliabile disabilitare la modalità RSTP (IEEE 802.1w)

Compatibilità ed interlavoro prodotti

- Prodotti di costruttori diversi conformi a IEEE 802.1w devono interoperare senza problemi
- Attenzione ai prodotti che annunciavano la disponibilità di RSTP nel 2001 o prima!!
 - Lo standard è stato ratificato alla fine 2001
 - Tali prodotti devono per forza avere una versione pre-standard
 - Potrebbe portare a problemi di interlavoro con prodotti effettivamente conformi allo standard