Seminario
architetture Layer 2 switching

Massimiliano Sbaraglia

Ethernet 802.1q

- Ethernet protocol
- Ethernet 802.1q tagging vlan
- Ethernet Q in Q tagging
- Ethernet Q in Q application
- Ethernet untagged vs tagged
- Ethernet LACP 802.1ad
- Ethernet MC-LAG 802.1aq
- PBB 802.1ah Provider Bridge
- Ethernet ARP details
- VTP Vlan Trunking Protocol

ETHERNET Protocol

Ethernet è una frame di livello 2 (data-link) del modello ISO/OSI per il trasporto di informazioni di dati; essa è composta da:

- Preambolo: ha il compito di sincronizzare il mittente con il destinatario a livello fisico con un valore binario = 10101010
- SFD: indica l'inizio di una frame con valore = 10101011
- DA: indica il MAC address dell'unità di destinazione
- SA: indica il MAC address dell'unità sorgente
- Type: indica il tipo di protocollo utilizzato (802.3 ethernet) oppure la lunghezza del campo data
- Data: contiene il payload (carico) dei dati e/o informazione
- CRC: è un controllo ciclico che permette la rivelazione di eventuali errori di trasmissione

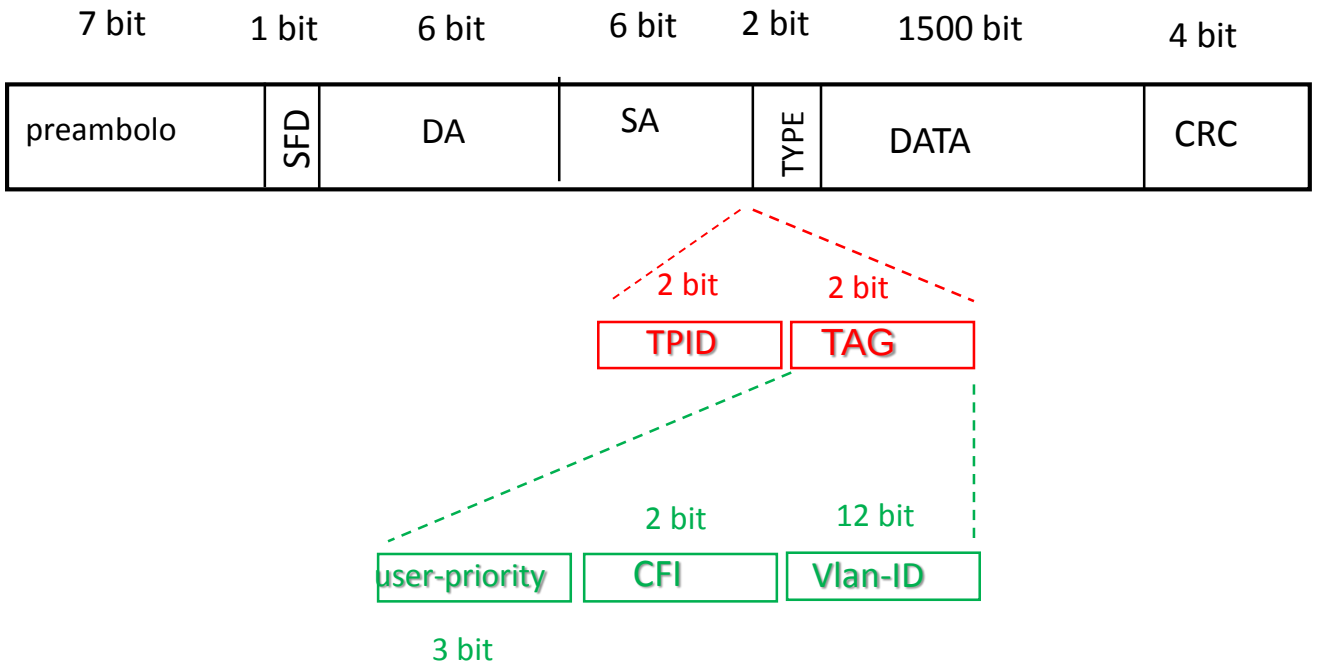
7 bit	1 bit	6 bit	6 bit	2 bit	1500 bit	4 bit
preambolo	SFD	DA	SA	TYPE	DATA	CRC

Original Frame Ethernet

ETHERNET Protocol with 802.1q tagging

802.1q (protocollo standard): introduce il concetto di vlan (virtual LAN) permettendo a questi segmenti logici di rete di condividere lo stesso media fisico

- TPID: indica il tipo di ethertype che assume il valore 0x8100 indicando il nuovo frame 802.1q tagged
- TAG: contiene tre sotto-infomrazioni:
- user-priority: indica un livello di priorità della frame; l'utilizzo di questo campo è definito in 802.1p (definisce classi di servizio cos)
- CFI: indica se i MAC address della frame sono in forma canonica
- VLAN-ID: assume un valore numerico in un range sino a 4096 possibili segmenti logici di rete.



ETHERNET Protocol Q in Q tunneling

E' una tecnica utilizzata per tunnellizzare un tag vlan in una secondo tag vlan all'interno di una frame Ethernet; questo permette di separare L2-VPN traffico utente all'interno di una rete services provider (ad esempio una Metro Ethernet).
E' molto utilizzato anche in ambito datacenters multi-tenants, dove è prevista una quantità di vlans molto elevata.

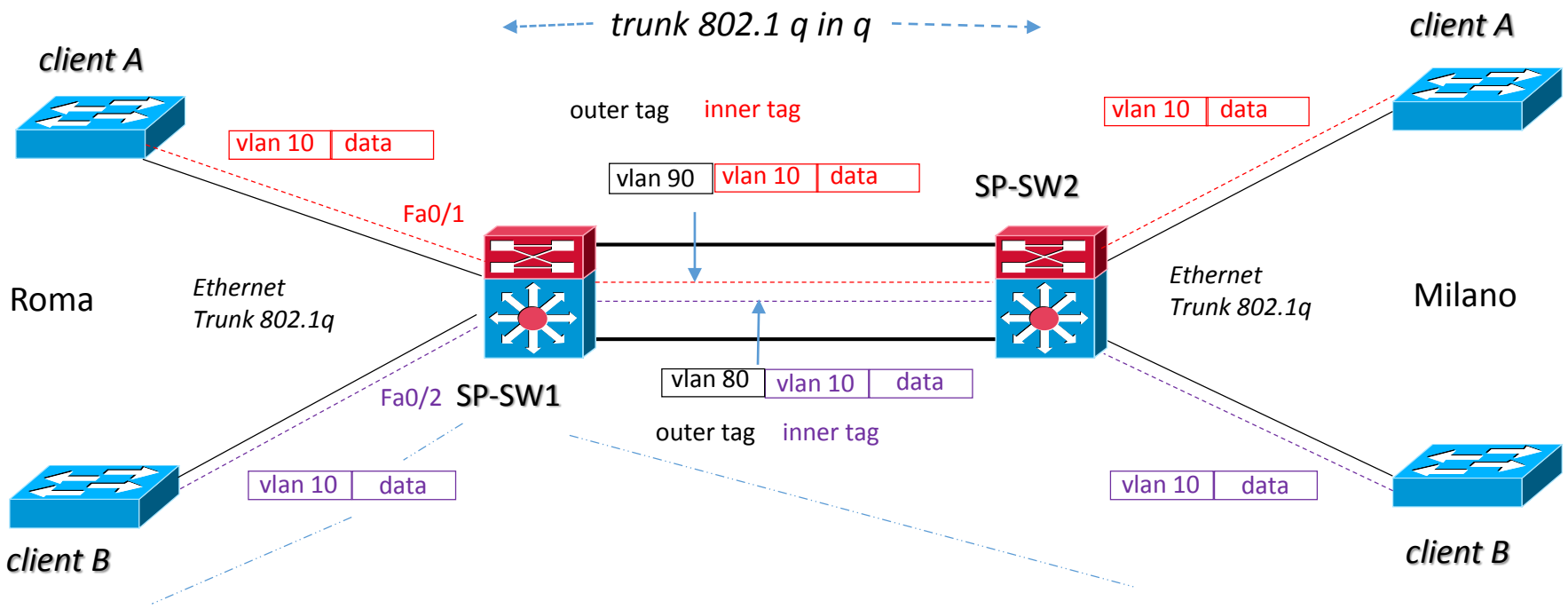
Original Frame Ethernet

Tag Vlan Customers 802.1q

Tag Vlan Providers Q in Q tunneling

ETHERNET Protocol Q in Q tunneling application example

Trova applicazione in backbone services provider tipo Metro Ethernet e per estendere domini layer 2 tra custmers geograficamente distanti




```
interface FastEthernet0/1
description «to Client-A Roma»
switchport access vlan 90
switchport mode dot1q-tunnel
```

```
interface FastEthernet0/2
description «to Client-B Roma»
switchport access vlan 80
switchport mode dot1q-tunnel
```

ETHERNET Protocol Q in Q tunneling application example

Qualsiasi frame ethernet proveniente dallo switch cliente è taggato con un nuovo VID dal services provider; questo processo è realizzato a livello Switch Services Provider sulla porta che è connessa al cliente.

Considerazioni:

- L'interfaccia dello SP - Switch che "vede" il cliente deve settare BPDU filter e e Root-Guard per prevenire lo switch cliente di diventare root-switch STP (gli switch L2 client vedono le BPDU trasmesse dai corrispettivi switch all'altro capo del dominio STP)

- Di default il VTP protocol è disabilitato tra i due switch client attraverso la rete del Services Provider (è possibile abilitarlo in caso di necessità)

- UDLD, PAgP e CDP (disabilitati per default) possono essere attivi;

- Il numero massimo di vlans che può essere configurato è 4096 (il valore del tag vlan-ID 802.1qè di 12 bit)

ETHERNET Protocol untagged vs tagged

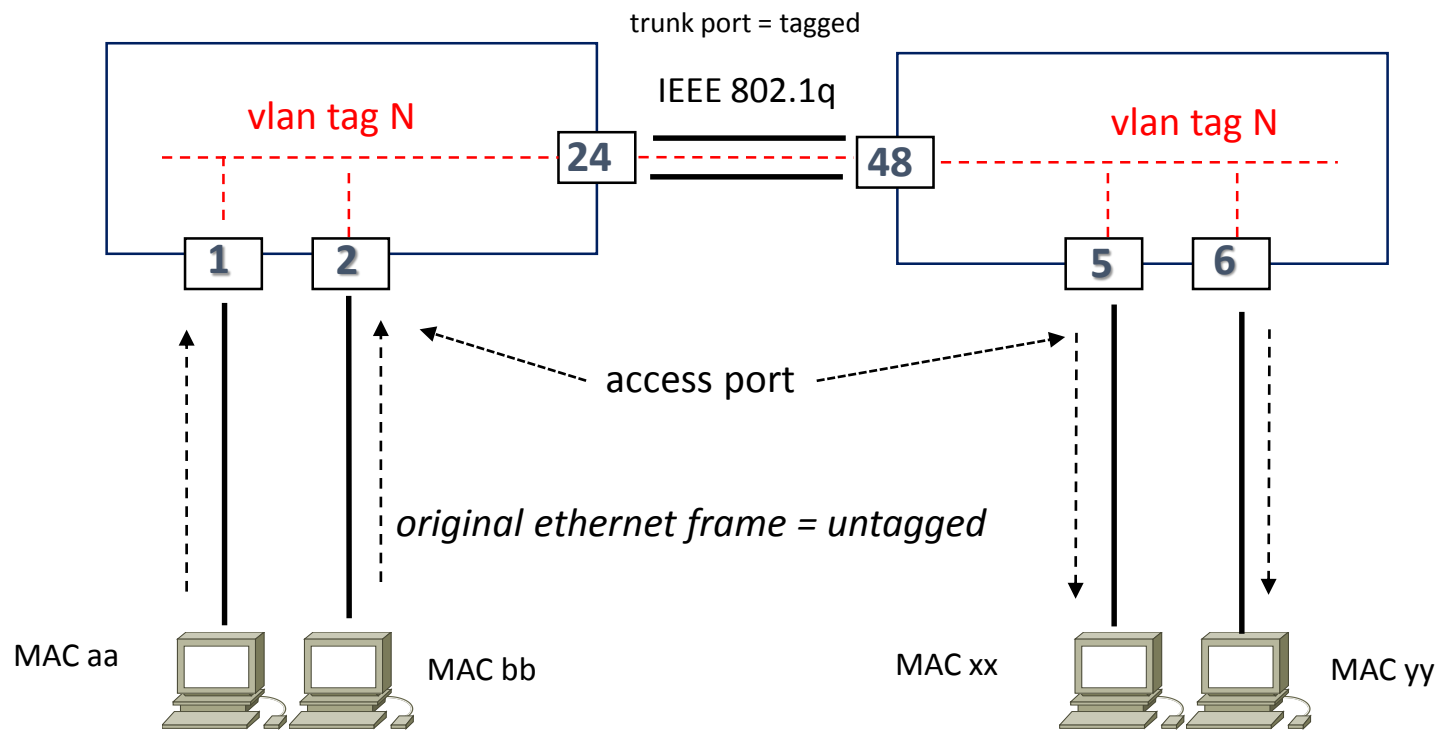
Le porte in access mode sono di tipo untagged; Le porte in trunk mode sono di tipo tagged

Gli switch trasmettono e ricevono frame ethernet sulla base della loro MAC table address Frame di tipo broadcast (FFFF.FFFF.FFFF) e unknown (sconosciuto il MAC di destinazione) vengono trasmesse su tutte le porte ad eccezione di quella dove hanno ricevuto la frame.

Cut-Through è una tecnica di forwardind della frame non appena lo switch riceve (e legge) il MAC di destinazione; bassa latenza, modalità sincrona (stesso bit rate tra porta sorgente e destinazione)

Store and Forward è una tecnica che permette la trasmissione di una frame interamente ricevuta e letta dallo switch; alta latenza ma con un maggior controllo via FCS.

ETHERNET Protocol untagged vs tagged



switch A

MAC	PORTS
aa	1
bb	2
xx	24
yy	24

switch B

MAC	PORTS
xx	5
yy	6
aa	48
bb	48

ETHERNET LACP 802.1ad Link Aggregation Protocol

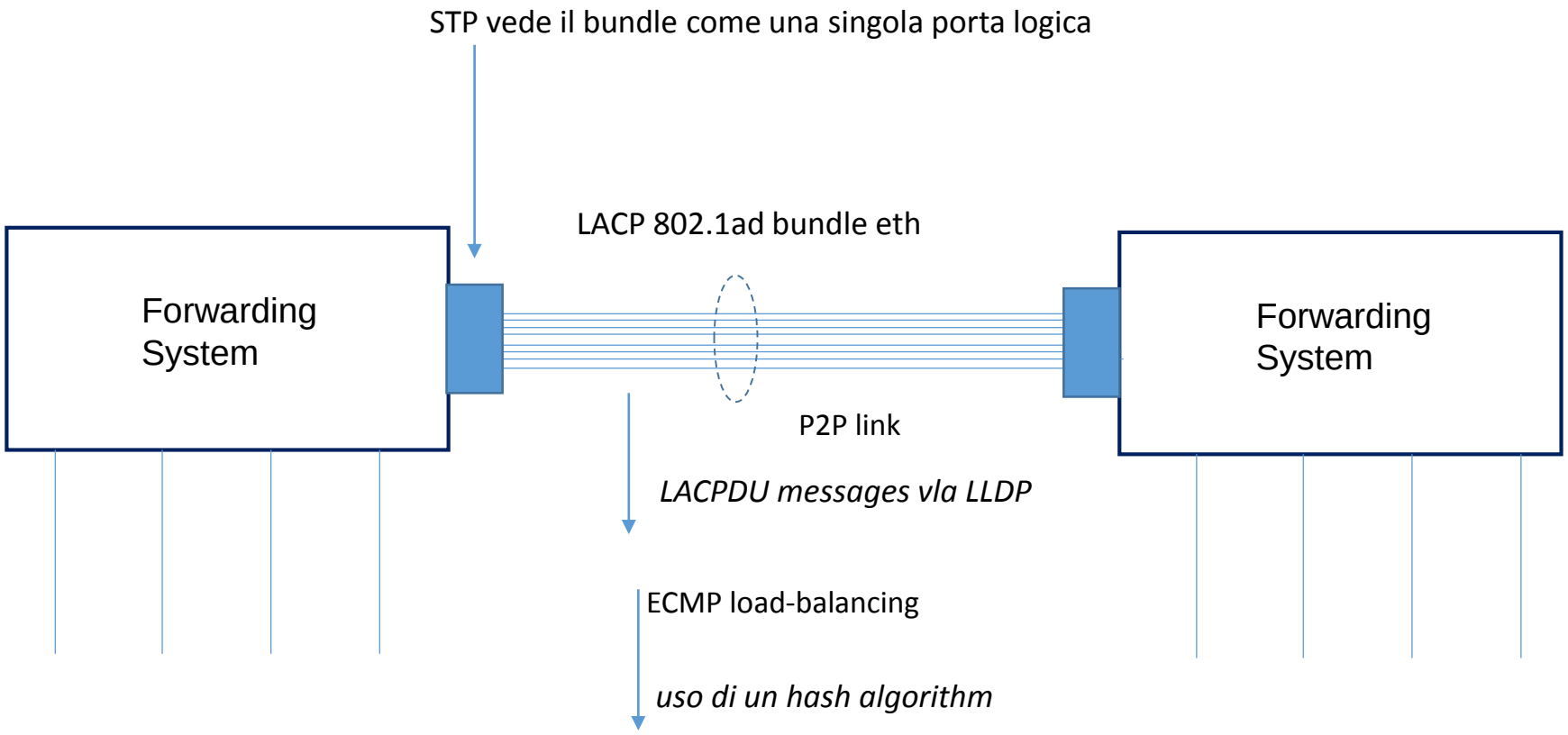
LACP è un protocollo standard 802.1ad; i pacchetti LACP sono trasmessi ad un gruppo multicast MAC con indirizzo 01-80-c2-00-00-02; durante la negoziazione i pacchetti sono trasmessi ogni secondo.
Le combinazione agli estremi del channel possono essere:

- on - on
- auto – desiderable
- desiderable – desiderable (only for PAgP; proprietario Cisco)
- active – active (only for LACP)
- passive – active (only for LACP)

Il significato di queste opzioni sono:

- auto = negoziazione passiva del link aggregation
- on = nessun protocollo è usato; si assume che su entrambi i capi del channel è abilitato il link aggregation
- desiderable = negoziazione attiva del link aggregation solo per PAgP
- active = negoziazione attiva del link aggregation solo per LACP

ETHERNET LACP 802.1ad Link Aggregation Protocol



ETHERNET LACP 802.1aq Multi-Chassis LAG and Shortest Path Bridge

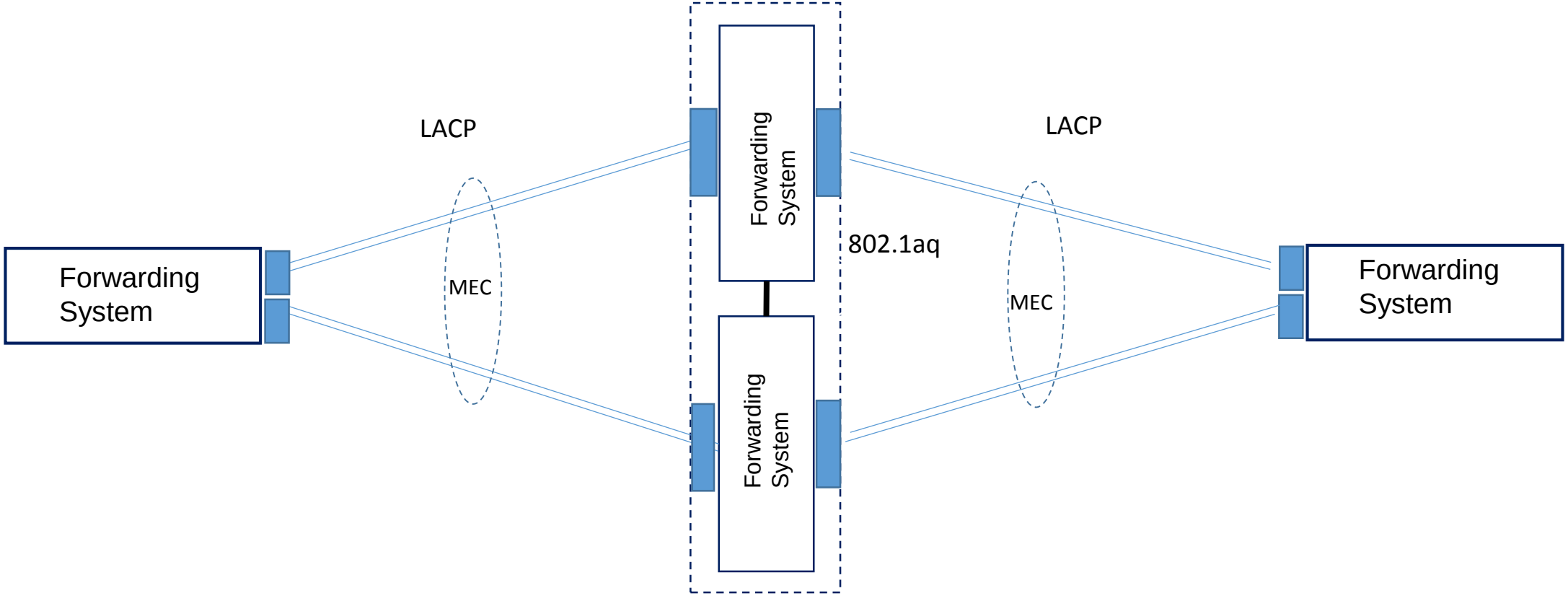
MC-LAG consiste di una connessione di links aggregati un in bundle terminati ad un «sistema» comune di Bridge, ma a differenza del LACP semplice p2p, i due switches/bridges hanno un indipendente piano di forwarding formando tra loro un multi-chassis (802.1aq)

Shortest Path Bridging (SPB) è un protocollo che applica i principi del routing al forwarding di pacchetti layer 2, sostituendo la funzionalità propria dello spanning-tree di disabilitare tutti i path ridondanti meno uno, ed introducendo invece la funzionalità di instradamento multipath a costo eguale (ECMP) ed accelera i tempi di convergenza con l'utilizzo di maggiore banda tra un nodo sorgente ed in nodo destinazione.

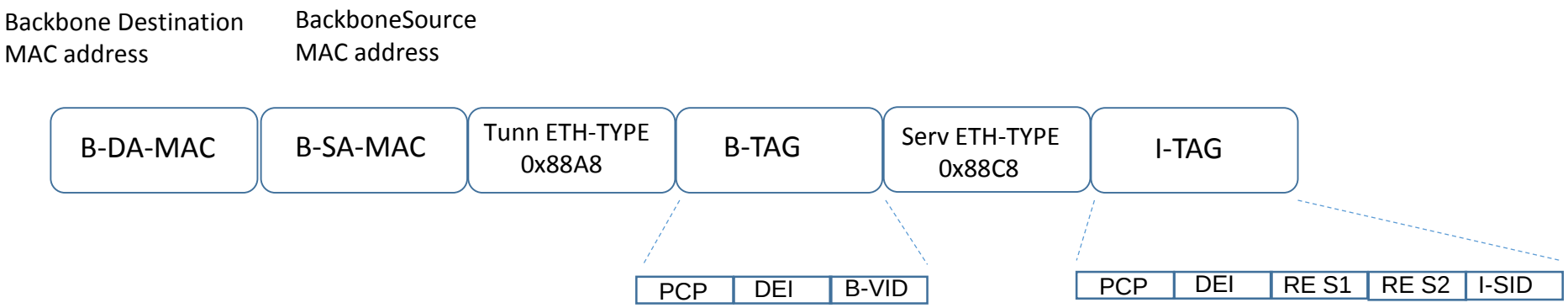
SPB utilizza un Layer 2 ISIS protocol (ISIS è un protocollo che a differenza di OSPF, lavora direttamente a livello data-link; ISIS non ha bisogno di IP sulle interfacce per scambiare informazioni di adiancenza ma invece di indirizzi chiamati NET network entity titles: ISIS è un protocollo CNLS no IP)con specifiche estensioni come control-plane per determinare multipli percorsi tra sistemi switches attraverso due modalita:

- SPBM: Shortest Path Bridging MAC = grande rete composta da 2 a 1000 bridges
- SPBV: Shortest Path Bridging VID (vlan-id) = piccola rete conposta da 1 to 100 bridges

ETHERNET LACP 802.1aq Multi-Chassis LAG and Shortest Path Bridge



ETHERNET PBB 802.1ah Provider Backbone Bridged



- B-DA = Backbone Destination Address = Tunnel destinazione address (può essere solo un indirizzo unicast; NO multicast address)
- B-SA = Backbone Source Address = Tunnel sorgente address (utilizzato per identificare questo nodo all'interno della rete Provider)
- B-VID = Tunnel Vlan ID 802.1q compliant
- I-TAG = contiene il campo I-SID il quale identifica come una frame ethernet dovrebbe essere mappata all'interno di un backbone di un SP e pertanto offre una instanza VID per il dominio di backbone (I-SID può assumere un valore tra 1 e 16 milioni = campo di 24 byte)
- PCP = identifica un priorità di servizio (range 0 – 7)
- DEI = un indicatore di drop (1 = drop eligible; 0 = not drop eligble)
- RES = campi riservati

ETHERNET PBB 802.1ah Provider Backbone Bridged encapsulation

Original Frame Ethernet

MAC DA	MAC SA	DATA

Frame Ethernet 802.1q (0x8100)

MAC DA	MAC SA	TAG VID	DATA

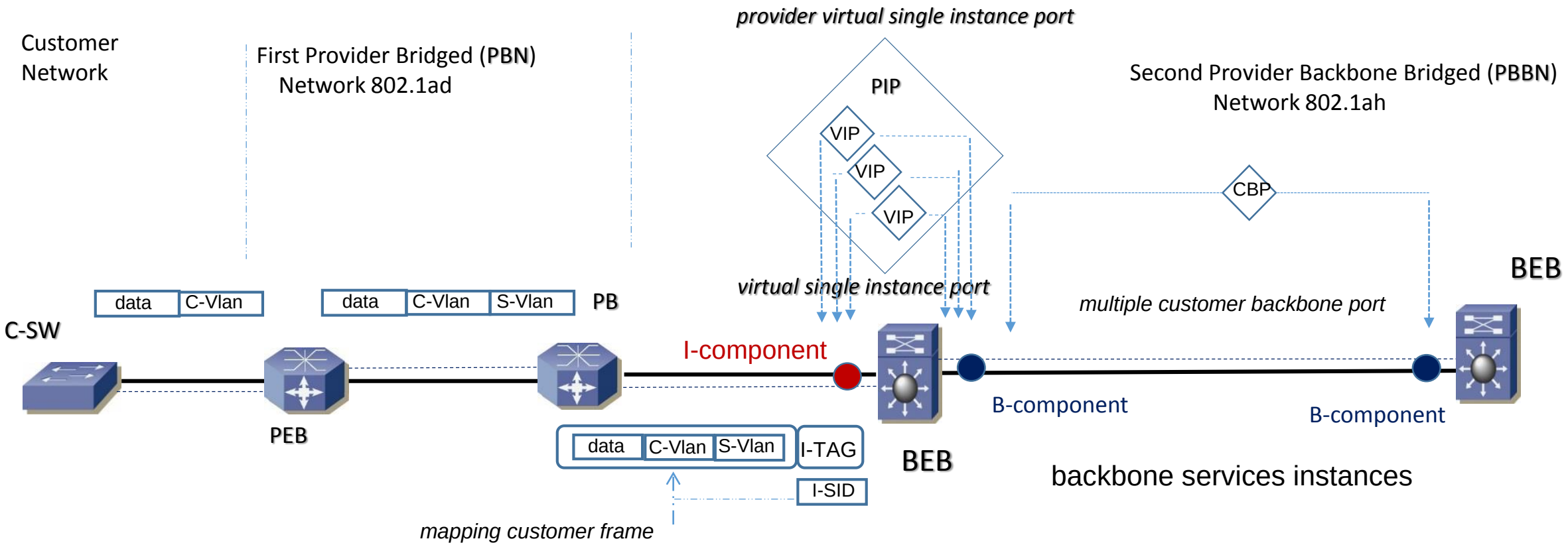
Frame Ethernet 802.1ad (0x88a8)

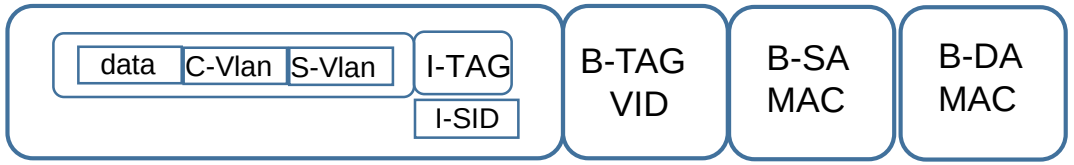
MAC DA	MAC SA	S-TAG VID	C-TAG VID	DATA

Frame Ethernet PBB 802.1ah

B-MAC DA	B-MAC SA	B-VID	0x88e7	I-TAG I-SID	C-MAC DA	C-MAC SA	S-TAG VID	C-TAG VID	DATA

ETHERNET PBB 802.1ah Provider Backbone Bridged flow-chart





I BEB mappano le frame dentro il B-VLAN con il suo tag ID contenuto nel campo B-TAG collegando tutti i BEB e BCB della PBBN;
Il B-SA di una frame encapsulata identifica una PIP che performa l'encapsulation (PIP sorgente);
Il B-DA di una frame encapsulata identifica PIPs per le quali le frame dovrebbero essere deliverate (PIP destinazione)

ETHERNET PBB 802.1ah Provider Backbone Bridged flow-chart

- I-Component: è responsabile di incapsulare frame ricevute dal customer ed assegnare ad ogni frame una backbone service instance; Il backbone service instance consiste in un set di BEB che supportano una customer S-Vlan identificata con un valore I-SID all'interno del PBBN.

  - Comprende una S-Vlan single
  - Riconosce ed usa il campo I-TAG in una o più customer network port or provider instance port (PIP)
  - Supporta 1:1 mapping tra S-Vlan e I-SID value
  - Termina il dominio STP con l'inibizione di trasmissione di PBN BPDU at PIP

- S-Component: è responsabile di ritrasmettere le customer frame encapsulate verso e da I-Component (ingress and egress direction) con il valore di I-SID utilizzato per assegnare un indirizzo VID (Vlan-ID) di backbone all'interno del PBBN con la ritrasmissione di queste frame attraverso le provider network port che provvedono al collegamento con altri Bridges facenti parte del backbone

  - Comprende una S-Vlan single
  - Riconosce ed usa il campo I-TAG in una o più customer backbone port
  - Termina il dominio STP con l'inibizione di trasmissione di PBBN at customer backbone port

ETHERNET ARP Address Resolution Protocol

ARP è un protocollo per il mapping tra un indirizzo MAC (indirizzo fisico di un nodo host Ethernet) ed il suo indirizzo IP associato, all'interno di un dominio di broadcast (vlan).

Quando una pacchetto entrante è destinato ad una certa destinazione, questo arriva al suo default gateway, il quale genera una richiesta ARP per trovare il corrispondente indirizzo fisico del nodo di destinazione che mappa l'indirizzo IP di destinazione indicato dal pacchetto

Un lookup ARP table è quindi performato dal nodo IP gateway e se trova la corrispondenza IP/MAC (entry), il pacchetto viene inoltrato nella giusta maniera; se non trova la corrispondenza IP/MAC (NO entry), il gateway genera una richiesta di ARP broadcast in modo tale che tutti i nodi presenti nella LAN possano verificare se l'indirizzo IP destinazione è presente in uno dei nodi è creare la corretta associazione tra sorgente e destinazione (il nodo che riconosce come proprio l'indirizzo IP destinazione genera una risposta di conferma)

Il nodo gateway aggiorna la sua cache ARP per la " nuova associazione " per future referenze di comunicazione.

VTP Vlan Trunking Protocol

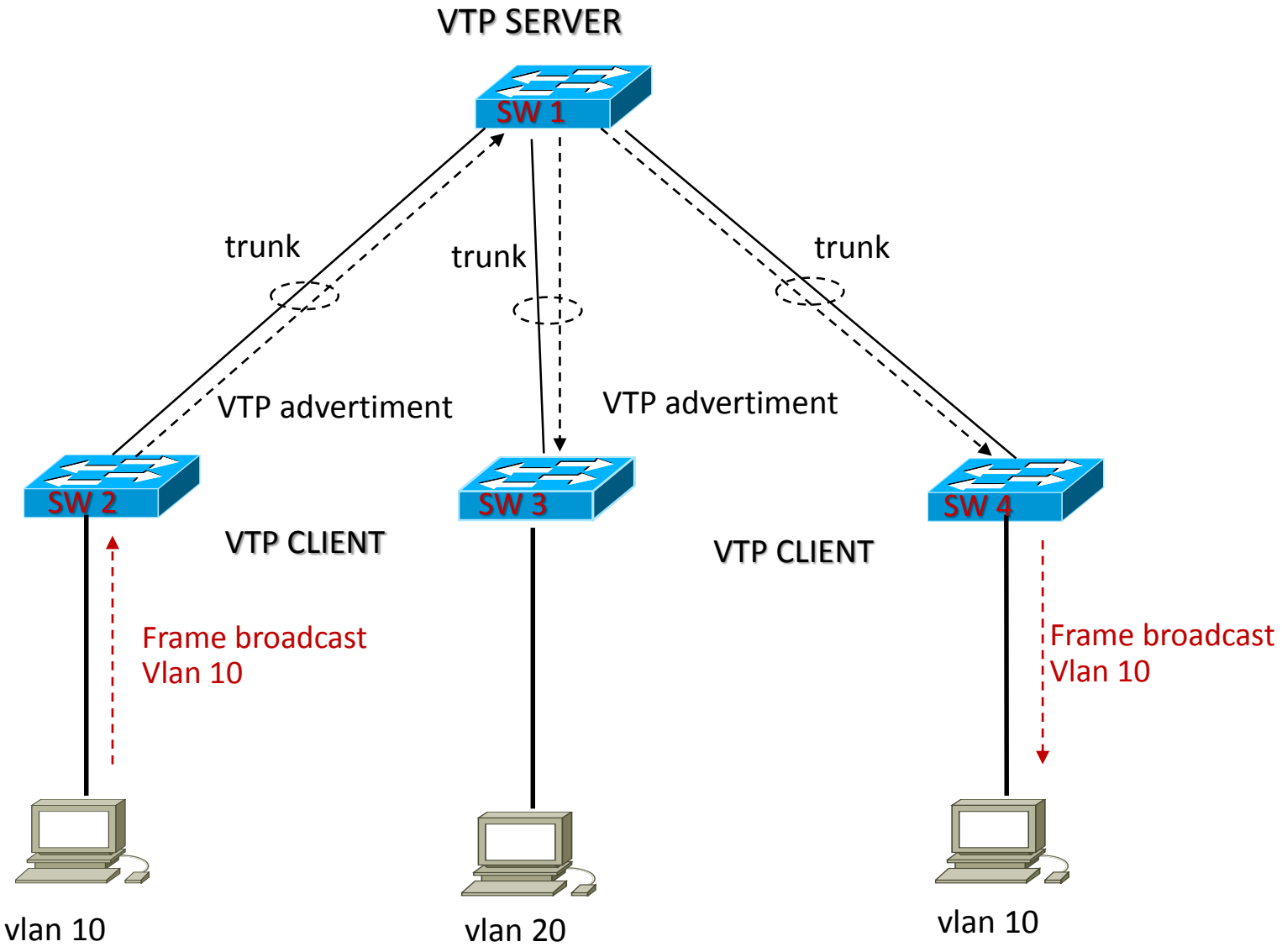
VTP è un protocollo che assicura uniformità e consistenza d vlans all'interno di un dominio LAN

Server Mode: può creare, modificare, eliminare vlans e decidere nuovi parametri di configurazione come ad esempio version e pruning per tutto il dominio VTP; VTP server annuncia e sincronizza tutto il vlans database a tutti gli switches client dello stesso dominio VTP attraverso gli advertisement ricevuti via trunk links (VTP server è configurato di default su switches cisco)

Client Mode: rappresentano altri switches settati in questo modo, appartenenti allo stesso dominio VTP del VTP server switch; questi switches non possono creare, modificare o eliminare nessun tipo di configurazione.

Trasparent Mode: rappresentano switches che non partecipano a nessun protocollo VTP (VTP off mode) e pertanto non annuncia e non sincronizza nessuna vlans del suo database; gli updates VTP sono ignorati e ritrasmessi in egress mode solo su trunks link

VTP Vlan Trunking Protocol



STP Spanning Tree Protocol

- STP 802.1d details
- STP 802.1d port role and status
- RSTP 802.1w port role and status
- RSTP 802.1w link type
- RSTP 802.1w path cost
- PVST+ path cost short and long
- MSTP 802.1s and PVST details
- BPDU details
- Considerazioni architetturali STP
- UDLD details
- Loop Guard details
- Root Guard details
- BPDU Guard details
- Storm Control Traffic
- Broadcast Suppression

STP 802.1d Spanning Tree Protocol

STP è un protocollo che assicura assenza di loop all'interno di un dominio di switching layer 2

Il protocollo utilizza un valore dello switch chiamato BRIDGE-ID per eleggere il root-switch; è un valore trasportato all'interno delle BPDU di 8 bytes:

- i primi due bytes rappresentano il Bridge Priority assumendo un range di valori decimali da 0 a 65.353 (di default = 32.768);

- gli ultimi sei bytes rappresentano l'indirizzo MAC proprio dello switch

In STP il valore più basso di Bridge-ID è preferito: in caso di parità della priority (valore per il quale uno switch assume il ruolo di root), il secondo item è quello di confrontare il più basso valore di indirizzo MAC tra i due switch oggetto della contesa.

STP 802.1d Port Role and Status

Port Role in STP 802.1d:

listening: la porta sta trasmettendo e ricevendo BPDU;

learning: si occupa di costruire la sua tabella bridging STP free loop;

blocking: non permette nessuna trasmissione di dati utente in egress alla porta;

fowarding: permette la trasmissione e ricezione di dati utente;

disable: non fa parte di nessuna azione STP; porta non attiva.

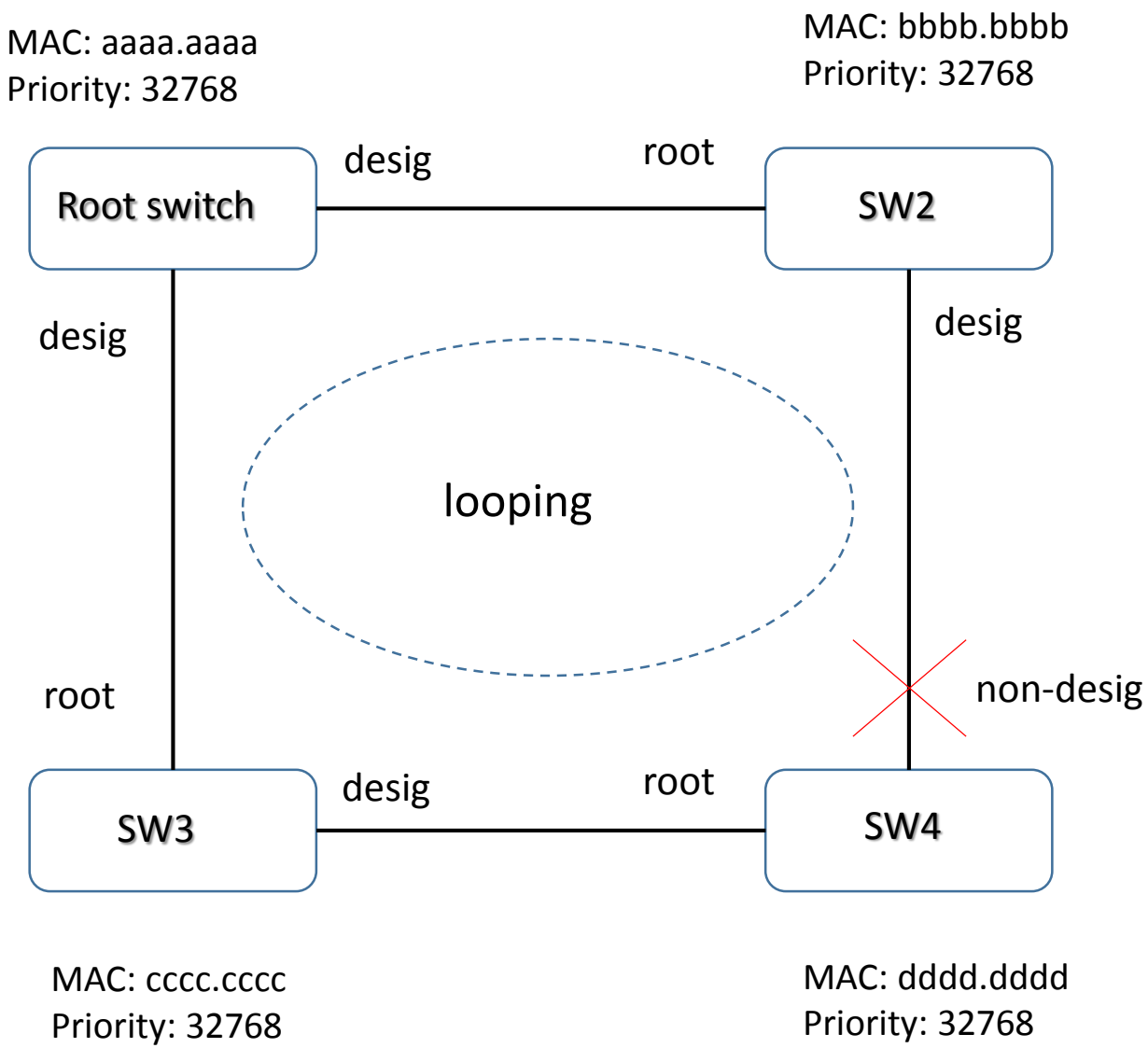
Port Status in STP 802.1d:

root: è la best port in forwarding status con direzione verso il root bridge;

designated: in forwarding status port per ogni segmento di rete LAN;

non-designated: in blocking status

STP 802.1d Port Role and Status



RSTP 802.1w Port Role and Status

Port Role in RSTP 802.1w (usato per fast convergence):

root: è la best port in forwarding status con direzione verso il root bridge;

designated: in forwarding status port per ogni segmento di rete LAN;

alternate: in blocking status rappresenta il path alternativo verso il root bridge; questo path è differente da quello usato dalla best port ;

backup: in blocking status rappresenta il backup/ridondanza di un path verso un segmento LAN dove un altro switch è invece già connesso;

disable: non necessariamente facente parte dello STP; è possibile manualmente disabilitare una porta.

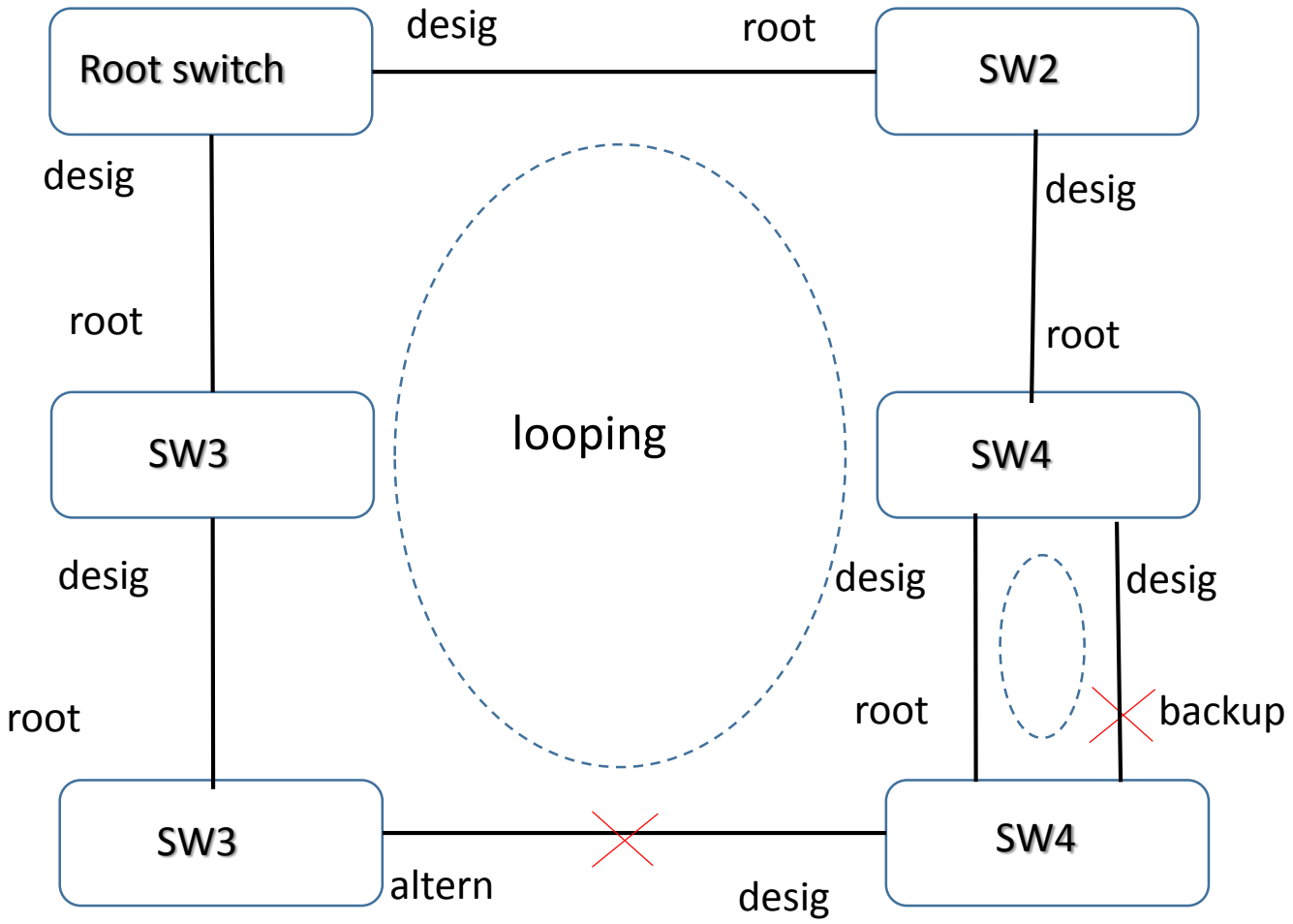
Port Status in RSTP 802.1w:

discarding: nessun pacchetto è trasmesso dalla porta;

learning: popolazione della tabella Mac-Address free loop;

forwarding: operativa

RSTP 802.1w Port Role and Status



RSTP 802.1w link type to forwarding

RSTP 802.1w Links Type to Forwarding

RSTP ha facoltà di categorizzare alcune porte in stato di forwarding senza attendere i timers del protocollo STP attraverso due variabili:

LINK TYPE: eseguito attraverso la configurazione della porta in half-duplex oppure full-duplex

full-duplex = link type point-to-point = connessione diretta tra due switches

half-duplex = link shared = connessione attraverso un terzo media dove multipli switches possono esistere

EDGE PORT: sono porte direttamente collegate tra uno switch ed un End-Point quali Host, Servers. (ricordarsi che queste porte se ricevono BPDU perdono il loro ruolo/status di edge porte e diventano normali porte STP e, quindi, generano all'arrivo di BPDU un TCN

RSTP 802.1w path cost

Il best path cost è calcolato su base del valore Bandwidth Interface del link associato ad un valore di costo della porta

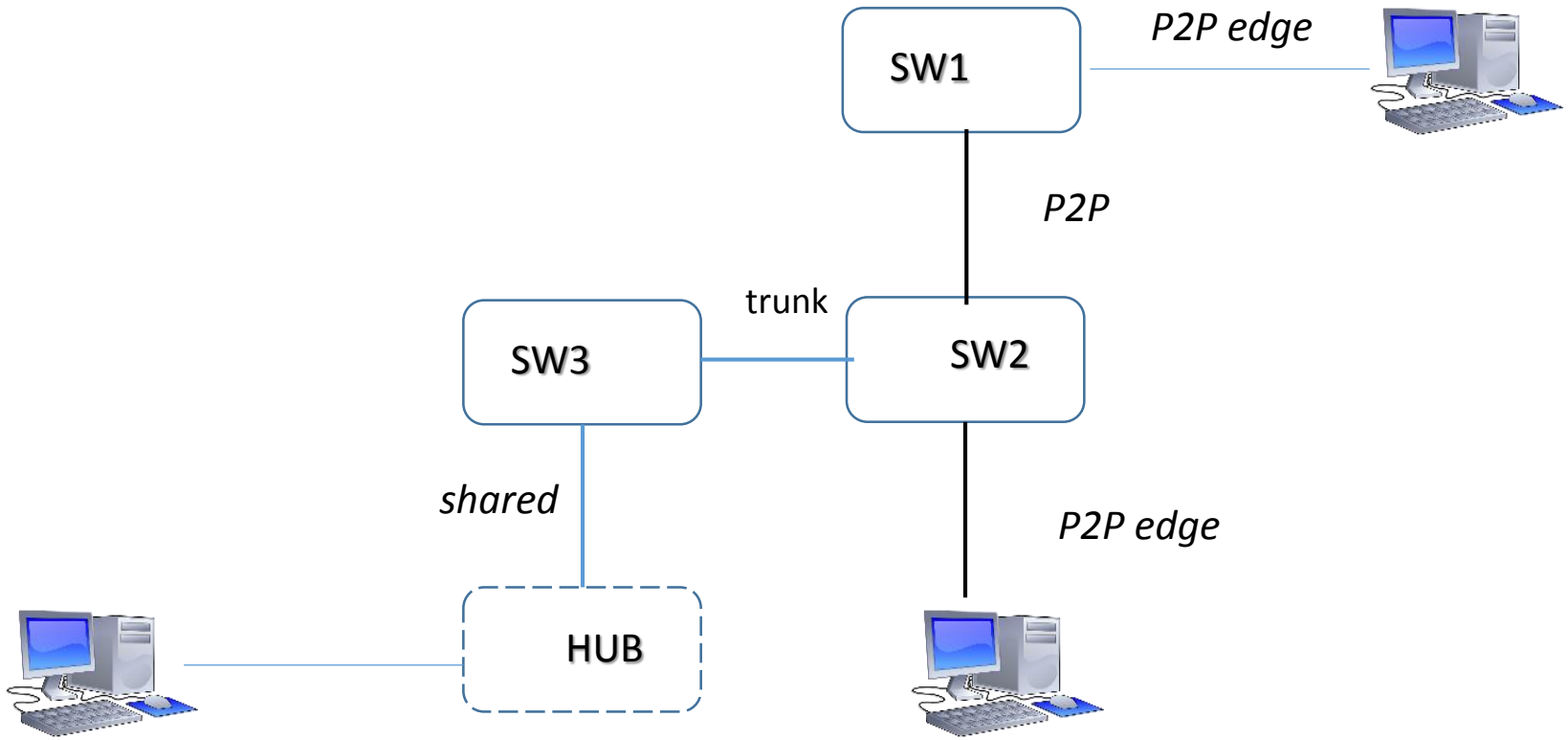
Data Rate	STP cost (802.1d)	RSTP cost (802.1w)
4 Mbit/s	250	5,000,000
10 Mbit/s	100	2,000,000
16 Mbit/s	62	1,250,000
100 Mbit/s	19	200,000
1 Gbit/s	4	20,000
2 Gbit/s	3	10,000
10 Gbit/s	2	2,000
100 Gbit/s	N/A	200
1 Tbit/s	N/A	20

PVST+ path cost

Il best path cost è calcolato su base del valore Bandwidth Interface del link associato ad un valore di costo della porta

Data Rate	PVST+ short path cost	PVST+ long path cost
10 Mbps	100	2,000,000
100 Mbps	19	200,000
1 Gigabit Ethernet	4	20,000
10 Gigabit Ethernet	2	2,000

RSTP 802.1w link type to forwarding



MSTP 802.1s and PVST+

802.1s = MSTP = Multiple Spanning Tree = una istanza STP per un set o region di vlans.

Risulta molto utile nel caso di molte vlans in uso, dove invece di avere una istanza per singola vlans, è possibile ottenere differenti regioni associate a diversi gruppi di vlans (ad esempio una regione per un range di vlans 1-600 ed un'altra regione per un range di vlans 601 – 1000).

Supporta un numero ridotto di instanze di STP e risparmia CPU di uno switch rispetto a RSTP.

Example configuration:

spanning-tree mst configuration
instance-id < id >vlan < vlan-range >
instance-id <id > vlan < vlan-range >
name < region-x >
revision 1 → *(0 to 65535)* → *traccia eventuali cambiamenti di configurazione MST (deve essere lo stesso per tutti gli switch appartenenti alla stessa region)*

Proprietario Cisco = PVST+ = Per Vlan Spanning Tree plus = standard compliant 802.1d = una istanza STP per singola vlan

BPDU Bridge Protocol Data Unit

BPDU (Bridge Protocol Data Unit): uno switch trasmette un pacchetto BPDU usando un unico MAC-Address inerente la porta fisica come indirizzo sorgente e come indirizzo di destinazione un STP Multicast address 01:80:C2:00:00:00

RSTP usa una estensione BPDU composta da:

CBPDU = Configuration BPDU utilizzato per la costruzione dell'albero, la scelta root-bridge e lo stato delle porte
TCN = Topology Change Notification BPDU

Il processo di selezione basato su BPDU (Bridge Protocol Data Unit) è trasmesso tra switches per ogni porta fisica; gli switches usano quattro step di processo per stabilire e salvare una copia del best BPDU visto da ciascuna porta (quando una porta riceve un miglior BPDU, ferma la trasmissione delle proprie BPDU).

Dopo un intervallo di tempo (20 sec di default), se le BPDU smettono di arrivare, la porta fisica inizia a ritrasmettere le proprie BPDU. I quatto step di processo Best BPDU sono:

- Il più basso Bridge ID (BID);
- Il più basso Path Cost Root Bridge: dipende dalla throughput di bandwidth del link;
- Il più basso Sender BID;
- Il più basso port ID (esempio una porta fa0/1 è più bassa di una porta fa0/2)

BPDU Bridge Protocol Data Unit

Le BPDU sono inviate ogni 2 secondi (hello time; tempo modificabile tra 1 e 10 sec)

Se entro 20 sec uno switch non riceve BPDU si considera perso lo switch mittente (max-age; modificabile tra 6 e 40 sec)

Il passaggio di una porta dallo stato di blocking a forwarding permane un tempo di 15 sec in cui la porta transita negli stadi di listening e learning (forwarding-delay; modificabile tra 4 e 30 sec)

per RSTP il valore è 2

<table>
<tr><td>Protocol-ID (0)</td><td>Protocol-version (0)</td><td>BPDU Type 0x80)</td></tr>
<tr><td>FLAG</td><td colspan="2">ROOT-ID specifica il root-bridge</td></tr>
<tr><td colspan="3">ROOT Path Cost contiene il costo del path verso il root-bridge</td></tr>
<tr><td colspan="3">BRIDGE –ID contiene priority e MAC del bridge</td></tr>
<tr><td colspan="3">INTERFACE-ID specifica la porta su cui il bridge sta inviando la config BPDU</td></tr>
<tr><td colspan="3">Message Age = val. incrementale al passaggio di uno switch</td></tr>
<tr><td>max age</td><td>hello time</td><td>forward delay</td></tr>
</table>

Considerazioni architetturali STP

Scalabilità: partendo da un root-bridge l'algoritmo sceglie sempre un solo path a costo minore di collegamento tra tutti i bridge del dominio L2; questo causa un 50% di link inutilizzati ma comunque necessari ai fini di fault e ridondanza.

Posizione root-bridge: è utile considerare la posizione che copre un root-bridge in una architettura STP; il traffico transita sempre da un root-bridge e pertanto si consiglia sempre una topologia a stella dove ciascun bridge è collegato direttamente al root-bridge, abbiamo una maggiore tolleranza ai guasti. Si consiglia di posizionare e collegare devices quali Servers oppure ambianti datacenters sempre il più vicino possibile allo switch root-bridge per ridurre possibili tempi di latenza nella trasmissione dei dati.

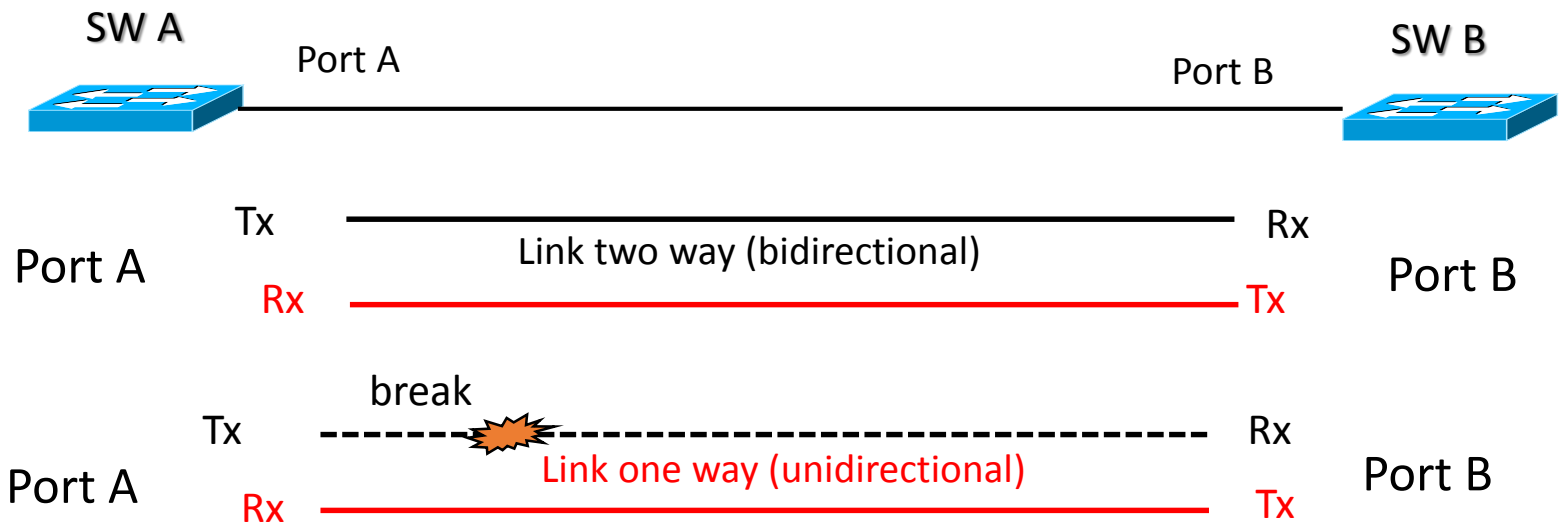
Posizione backup-bridge: entra in gioco qualora il primario root-bridge è in fault come pure un link primario; è importante quindi mantenere una topologia a stella dove abbiamo a partire da un bridge unl ink primario collegato al root-bridge ed un link secondario collegato al suo backup-bridge ed inoltre si consiglia di personalizzare i valori di priority dei due switches avendo quello di backup con un valore leggermente più alto rispetto a quello di primario.

Dimensionamento STP: si ricorda che il configuration BPDU come pacchetto di informazione tra due end-system hosts non può attraversare più di sette switches in cascata (questo per permettere all'informazione BPDU di attraversare la rete due volte entro il forwarding time = 15 sec di default; (relazione tra il message-age ed il forwarding time)

UDLD (Unidirectional Link Detection)

UDLD è un protocollo proprietario Cisco che pemette di prevenire problemi di loops tra una connessione in fibra ottica oppure in rame attraverso il riconoscimento di un link in one-way (unidirectional), disabilitando la porta in errore ed inviando un alert

- UDLD Normal Mode: opera su una connessione di sola fibra ottica
- UDLD Aggressive Mode: opera su connessioni sia in fibra ottica che rame
- UDLD deve essere abilitato su base interfaccia o globalmente



loop-guard e dove deve essere configurato

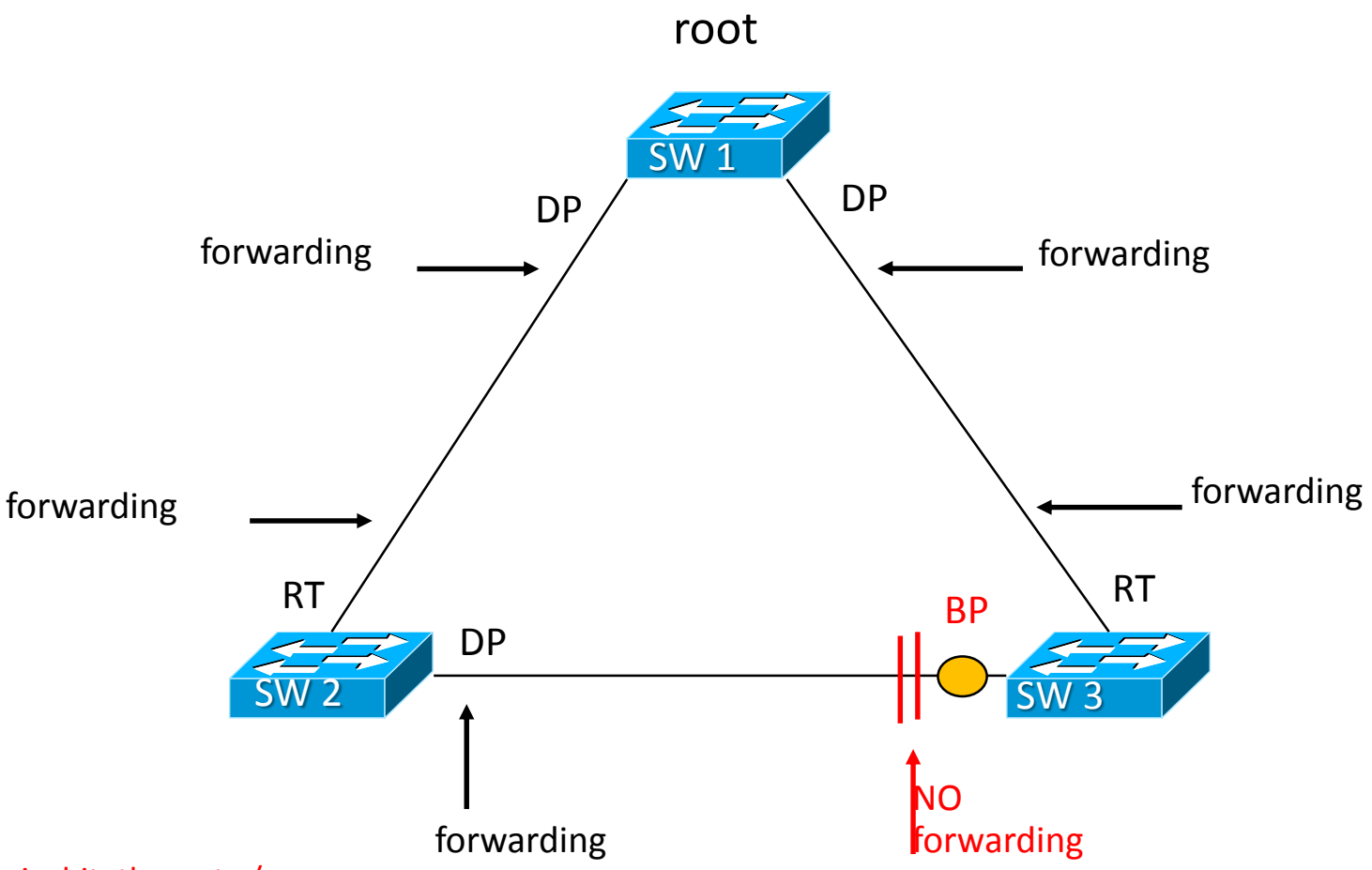
Loop Guard è una feature che protegge un link di collegamento tra switches a causare un loop; quindi previene una transazione da uno stato di blocking ad uno di forwarding, quando per cause di errore dovute a UDLD presente in uno specifico link.

Vedi diagramma slide successiva:

Lo swith 3 ha una porta in blocked state e, se ad esempio, non ricevesse più BPDU hello dallo switch 2 a causa di un ipotetico fault (ricordiamo che una porta blocking resta in ascolto delle BPDU), lo switch 3, pensando ad un nuovo processo STP transita in stato di forwarding la sua porta (per trasmettere le sue BPDU) ma essendo tutte le altre porte in stato di forwarding, si crea un loop

Se invece fosse presente loop-guard, la porta in stato blocking dello switch 3, in conseguenza di mancata ricezione di BPDU su quella porta, anziche transitare in uno stato di forwarding, transita in uno stato di loop-inconsistence che ha lo stesso valore e comportamento di una porta in blocking.
Di default loop-guard è disabilitato

loop-guard e dove deve essere configurato



interface gigabitethernet x/y
spanning-tree loopguard

OR
spanning-tree loopguard default

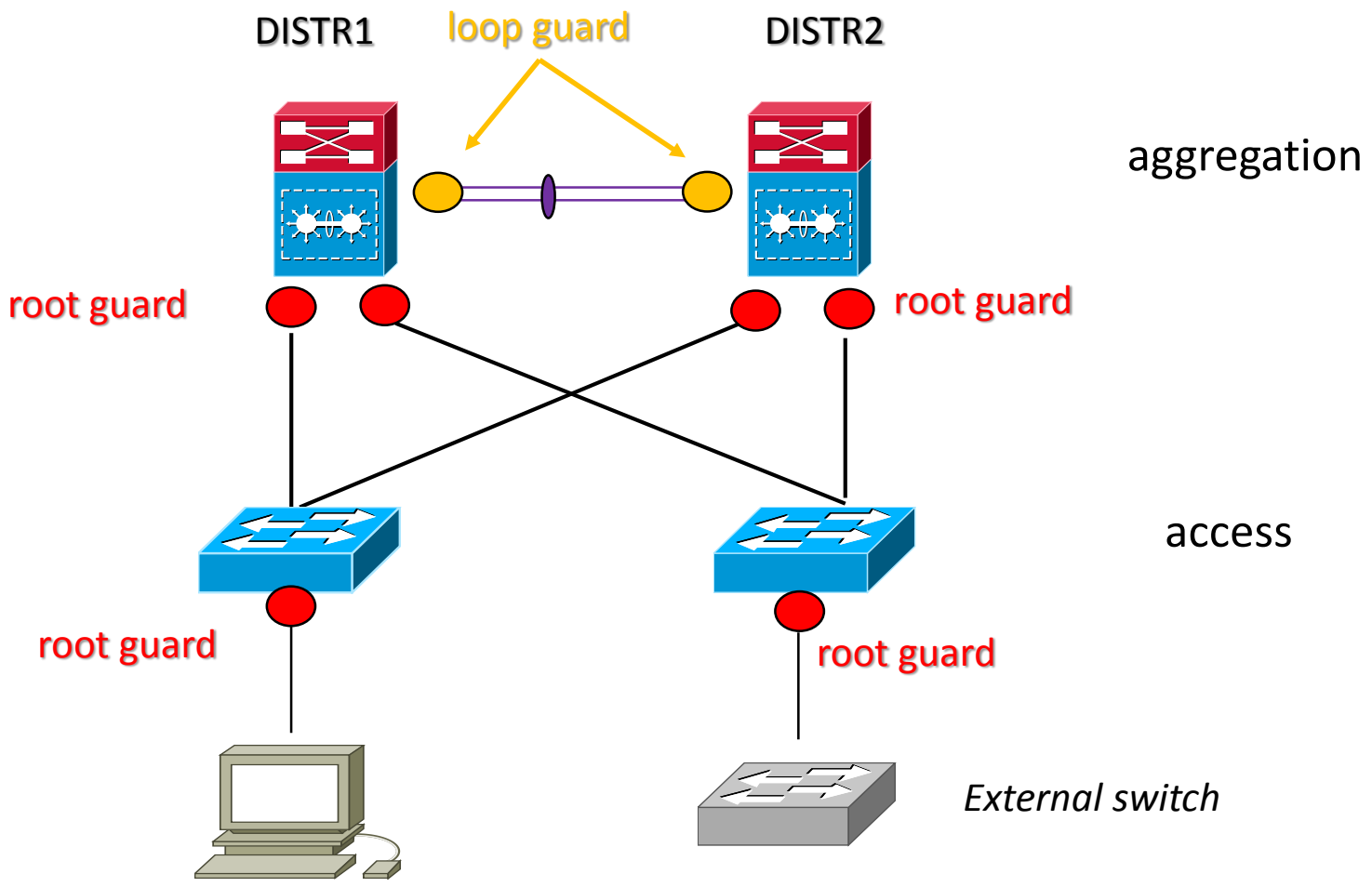
root-guard e dove deve essere configurato

Root Guard è una feature che previene una porta (differente dalla legittima root port) di uno switch (non il root-switch) a diventare root-port

Questa feature è raccomandabile a livello di aggregazione di un campus che vede la parte di accesso, evitando cosi che un eventuale errore di configurazione possa far ricalcolare un processo STP con una nuova elezione del root-bridge per una specifica vlan o istanza.

Anche a livello di accesso è buona norma configurare il root-guard per quelle porte di accesso alle quali sono collegati untrusted host che potrebbero introdurre malevoli traffici oppure, ad esempio, collegare external switch con un valore più basso di bridge ID rispetto a quello legittimo della rete, scatenando un ricalcolo dello STP con conseguenze distruttive per l'intera architettura.

root-guard e dove deve essere configurato



interface gigabitethernet x/y
spanning-tree guard root

bpdu-guard e dove deve essere configurato

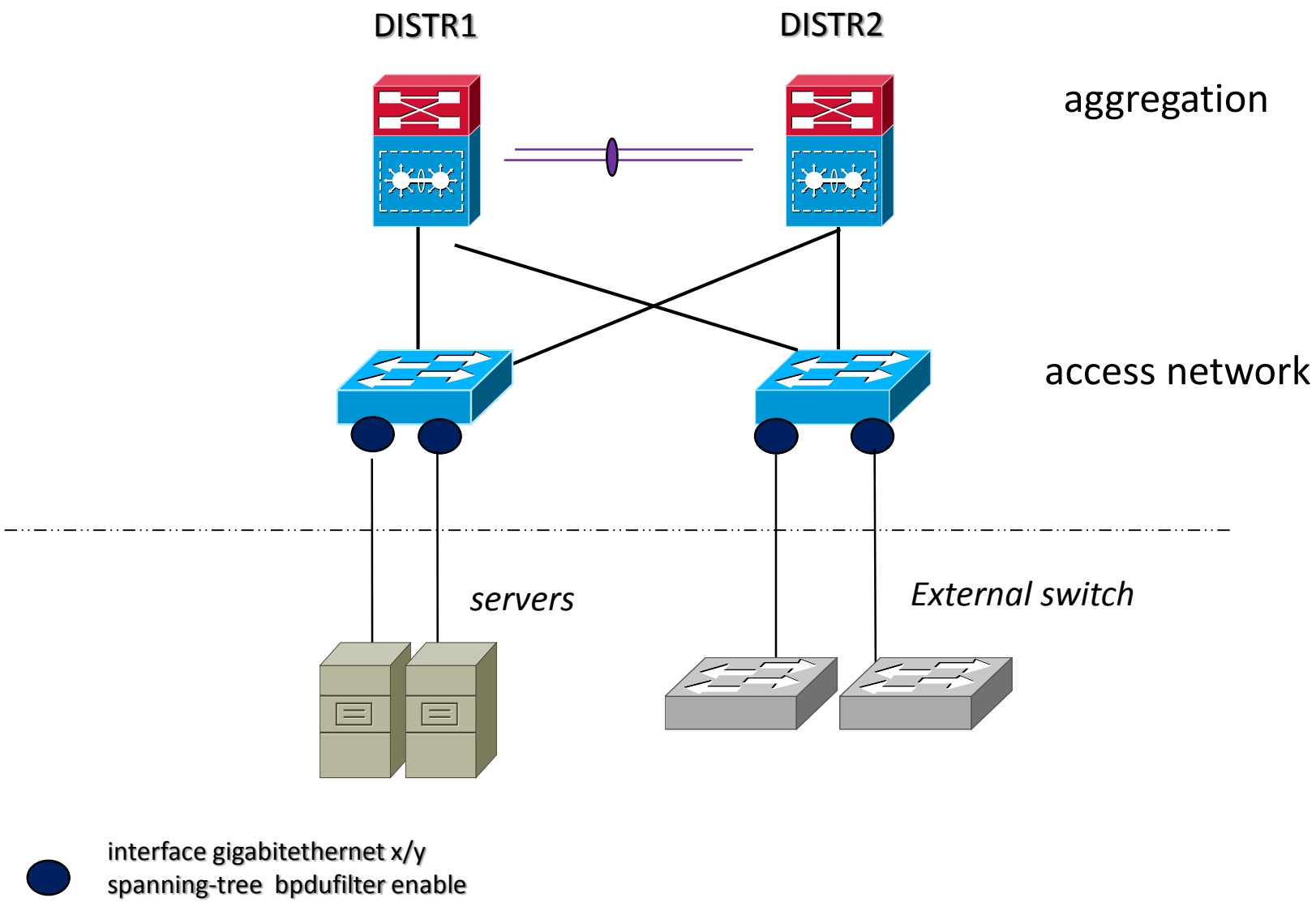
BPDU Guard è una feature che disabilita (shutdown) una porta di uno switch, non appena questa riceve una BPDU proveniente da un external switch o devices STP, ponendo la porta sulla quale era stato abilitato il bpdu-guard lo stato di err-disable.

In pratica si usa abbinare la configurazione bpdu guard con il port-fast per quelle connessioni verso servers per i quali è richiesta una rapida convergenza a trasmettere dati senza aspettare tempi (circa 50 sec) per cui una porta stabilisce il suo stato di forwarding (bypassando di fatto gli stati intermedi di listening e learning);

In ogni caso considera che STP è sempre attivo e pertanto potrebbe portare in blocking queste porte; il BPDU guard previene da questa situazione.

BPDU guard + Port-fast garantiscono un dominio STP mantenendo la configurazione legittima e non permettendo a devices esterni al dominio di influenzarne il processo

bpdu-guard e dove deve essere configurato



storm-control traffic

Traffico di tipo Storm si verifica quando vi è un eccesso di pacchetti (flooding) per una data porta fisica LAN, degradando di conseguenza le prestazioni.

Il controllo per questo tipo di traffico, chiamato STORM CONTROL, previene le porte LAN di uno switch dall' essere "disturbate" da traffico storm sia broadcast, multicast, unicast; in altre parole, questo controllo monitorizza il livello di traffico in ingresso in un dato intervallo di tempo (1 sec) , comparando la quantità di traffico in uso su quella determinata interfaccia fisica e il livello di storm control che manualmente viene configurato. Il livello di controllo di traffico storm è una percentuale del valore di bandwitdh della porta; ogni porta ha un singolo livello di controllo di traffico di tipo storm per qualsiasi tipo di trasmissione sia broadcast, multicast, unicast.

All'interno dell'intervallo di tempo (1 sec), quando il traffico in ingresso per il quale il controllo di storm traffic è abilitato, raggiunge il livello settato sotto la porta fisica, il traffico viene scartato (drop).

```
interface x/y
storm-control [ broadcast | multicast | unicast ] level [ percentage or pps ]
```

broadcast-suppression traffic

Uno storm broadcast si verifica quando pacchetti di tipo broadcast o multicast inondano (flooding) un dominio di rete LAN creando malfunzionamenti e degrado delle prestazioni.

Broadcast suppression usa dei filtri di misurazione dell'attività broadcast all'interno della LAN per un periodo di circa 1 sec e varia a seconda del tipo di line-card e velocità settata sulla porta fisica dello switch, comparando questa misurazione con una soglia (threshold) predefinita.

Se viene raggiunta questa soglia, traffico broadcast viene soppresso per una durata di tempo specifico.

Broadcast soppression è di default disabilitato

```
interface x/y broadcast suppression [ threshold ]

Example :
Threshold = 0,00 for suppress all broadcast traffic
Threshold =0,01 for 0,01% (1/100 percent)
Threshold = 0,50 for 0,50% (one-half percent)
```

Architetture Layer2

- Architettura di Accesso
- Architettetura di Distribuzione
- Achiettetura di Core
- Looped Switching Cisco Models
- Looped Triangle Cisco design
- Looped Square Cisco design
- Looped Free Cisco Models
- Loop-free type U Cisco design
- Loop-free inverted U Cisco design
- Flex-link Cisco Models
- Flex-link Cisco design

Architetture di Accesso

- Access Level: è il livello edge di accesso per il collegamento di host quali PC, stampanti, IP-Phone, WIFI access-point, camera, etc.

- Discovery: CDP and LLDP

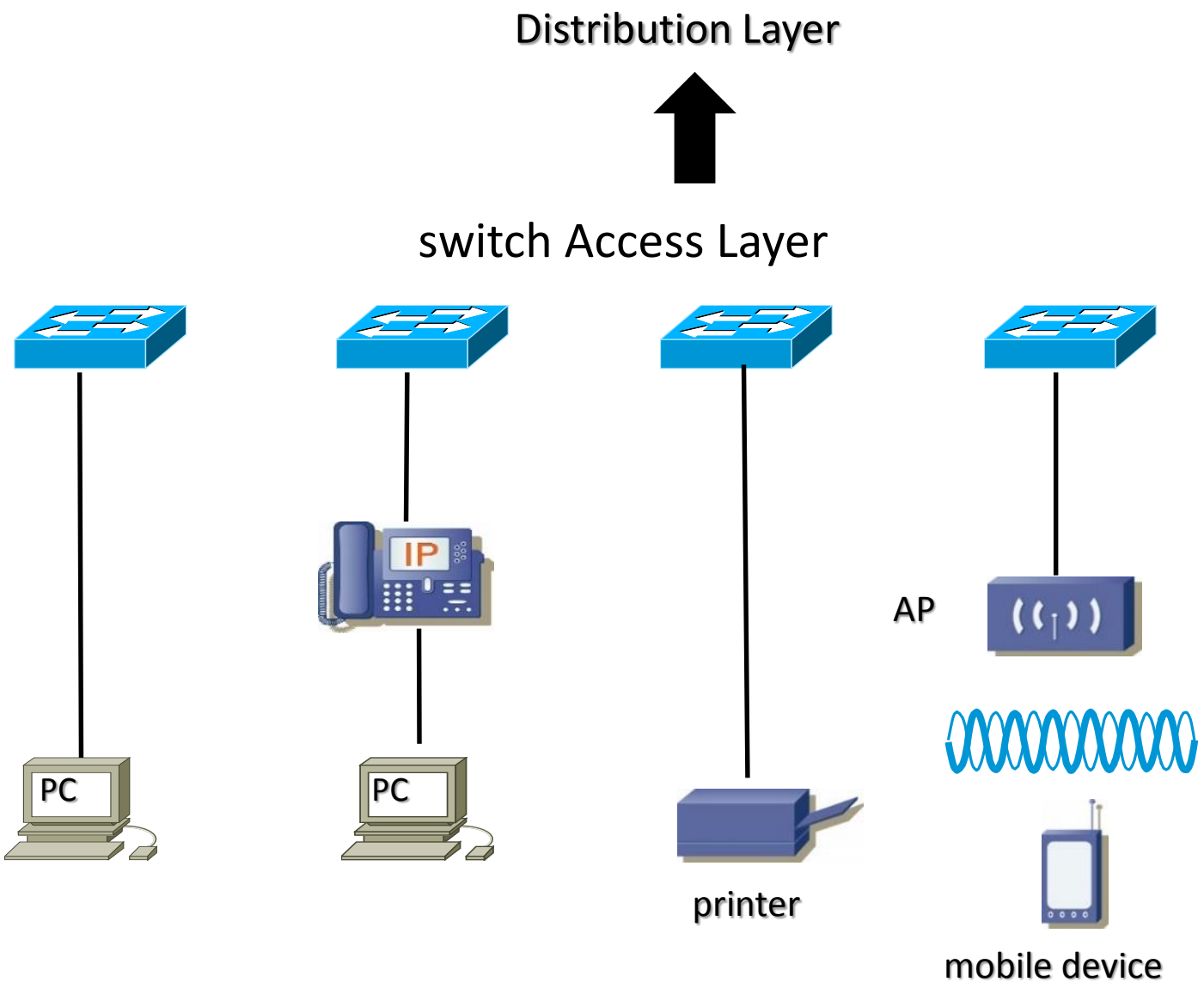
- Security and Network Identity: 802.1x, port-security, DHCP snooping, IBNS (Identity Base Network Services) , IPSG (IP source-guard), Web-Auth, DAI (Dynamic ARP Inspection)

- Application Recognition Services: QoS marking, policing, queueing, deep packet inspection NBAR

- Network Control Service: STP, RPST, PVST+, VTP, LACP, PAgP, UDLD, port-fast, uplink-fast, backbone-fast, loop-guard, BPDU-guard, rot-guard, port-security, EIGRP, OSPF (access routing layer)

- Physical Infrastructure Services: PoE (Power of Ethernet)

Architetture di Accesso



Architetture di Distribuzione

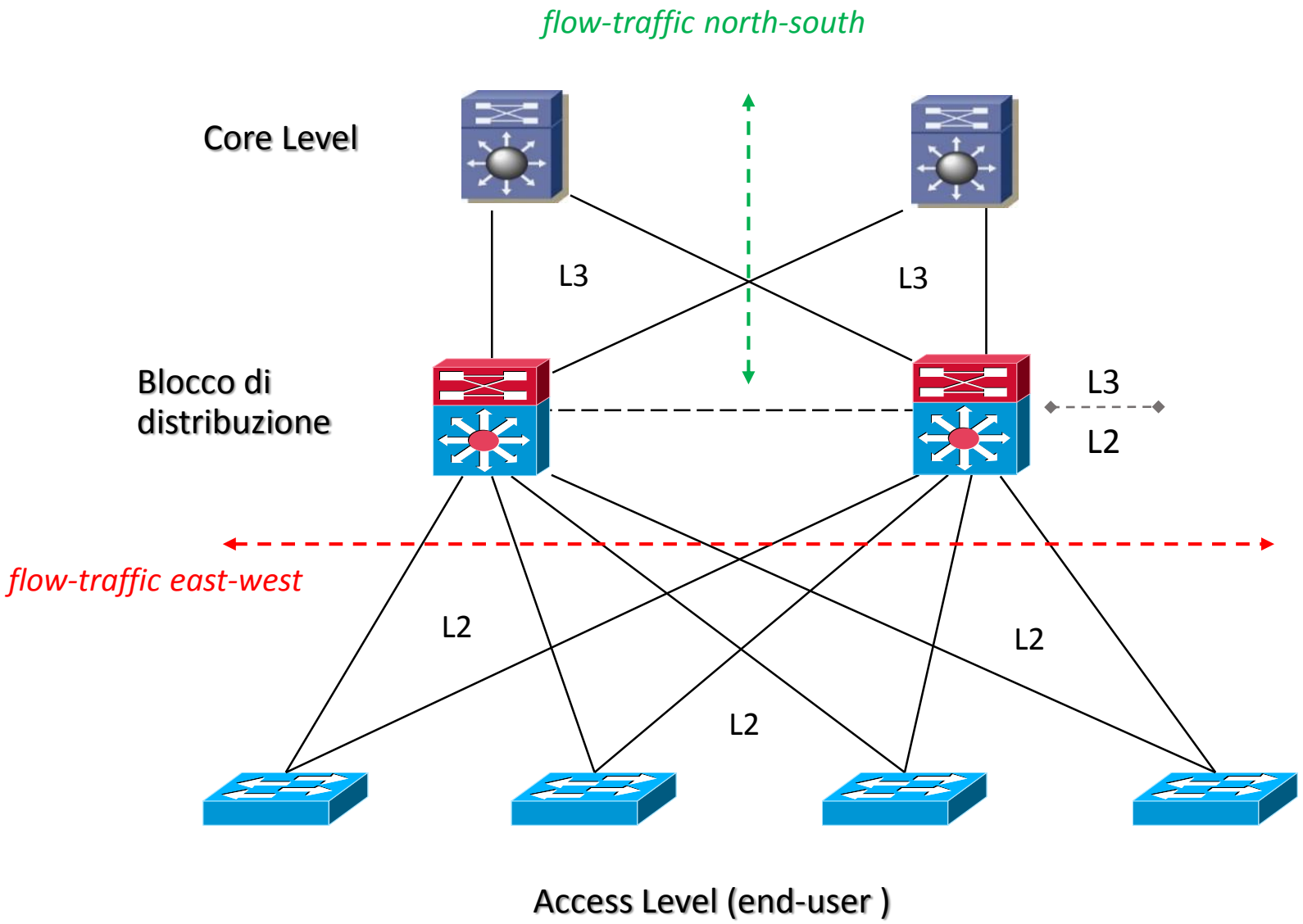
Distribution Level: è il livello che consente un punto di aggregazione tra il livello di accesso (end-user) ed il livello di core (external domain)

Provvede alla connettività e policy di servizio, all'interno di un singolo blocco di distribuzione, per flussi di traffico transitanti tra end-user node (east-west flow traffic)

Provvede alla connettività, policy di controllo ed un punto di demarcazione tra il blocco di distribuzione (campus di rete) ed il resto della rete per flussi di traffico diretti verso external network domain.

Partecipa a configurazioni di routing per scalabilità e performance di convergenza tra il livello di distribuzione ed il livello di core

Architetture di Distribuzione



Architetture di Core

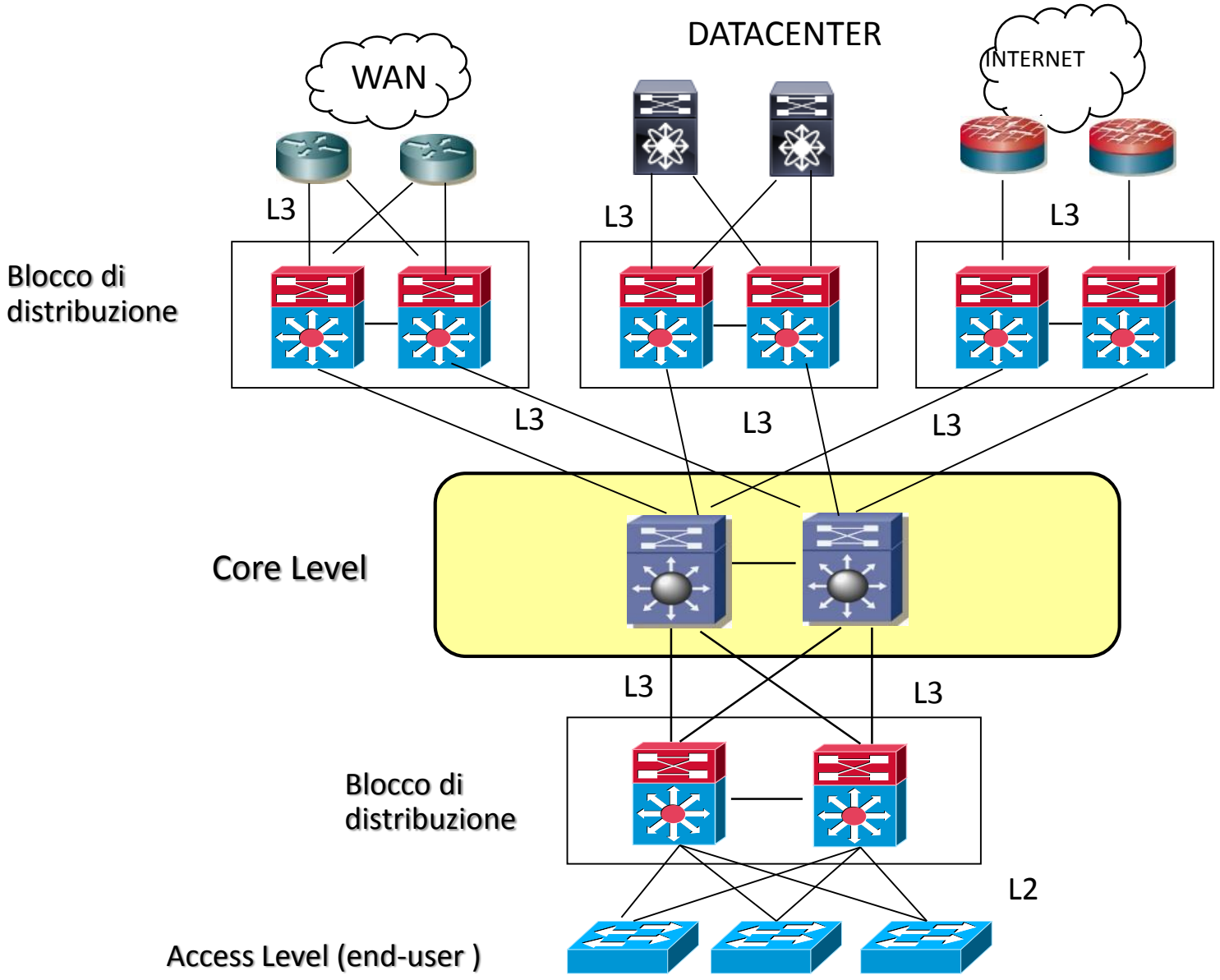
Core Level: è il livello che deve garantire alta affidabilità, ridondanza, resilienza, security, non-stop service capability

Consente un livello di aggregazione tra differenti blocchi di distribuzione

Consente l'interoperabilità tra differenti ed external network domain (Internet, WAN, DataCenters) con il livello di accesso (End-User)

Consente immediato data-flow recovery in caso di fault di qualsiasi componente della rete

Architetture di Core



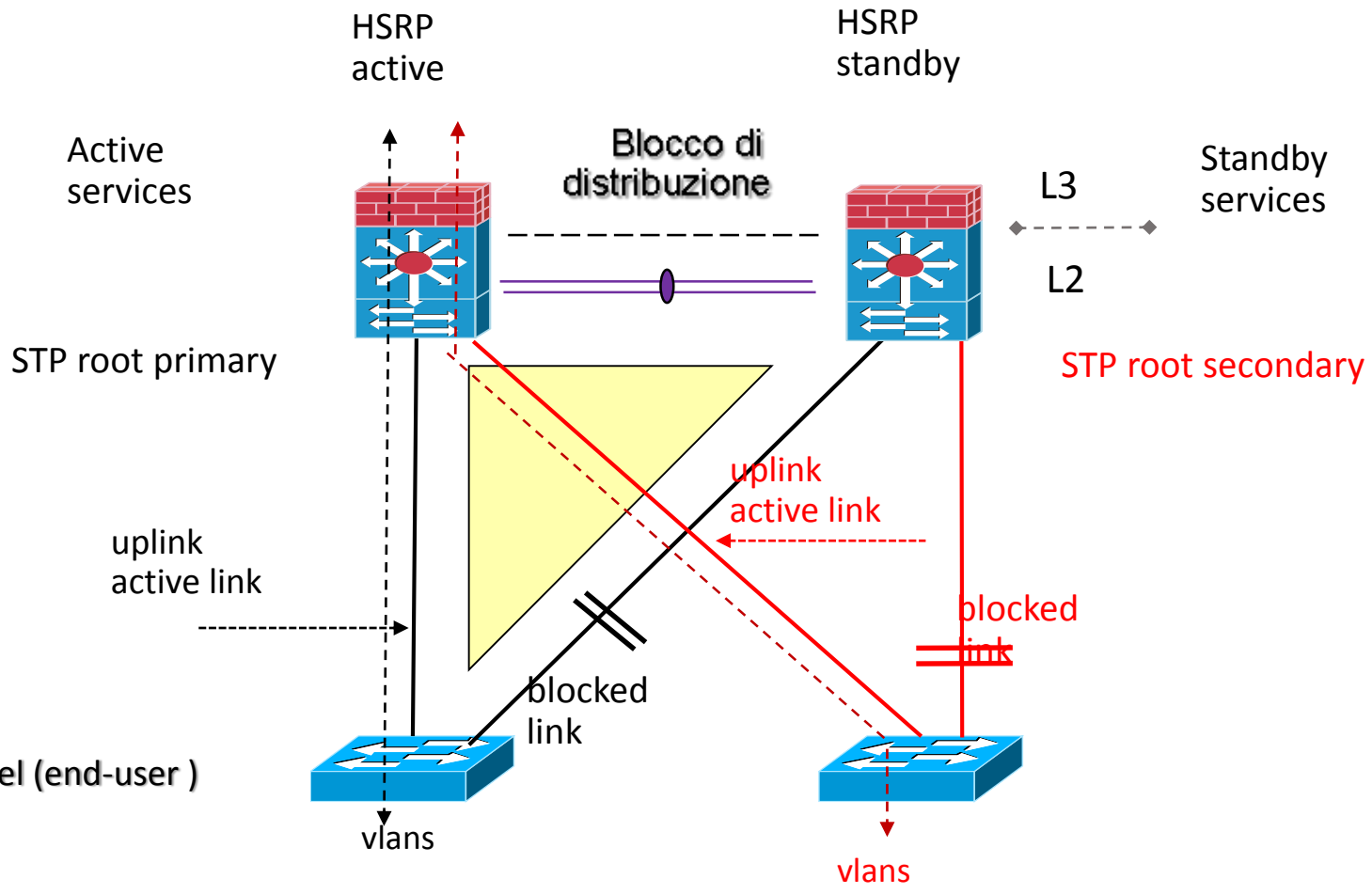
Looped Switching Architectures Design Cisco Models

Layer 2 Looped : tutte le vlans sono estese a livello di aggregazione ed offre particolari benefici per servizi di tipo statefull (significa la capacità di memorizzare dati per essere utilizzati a scopi specifici come quelli di FWSM o SLB); il livello 3 (routing) è performato dal livello di aggregazione in alto

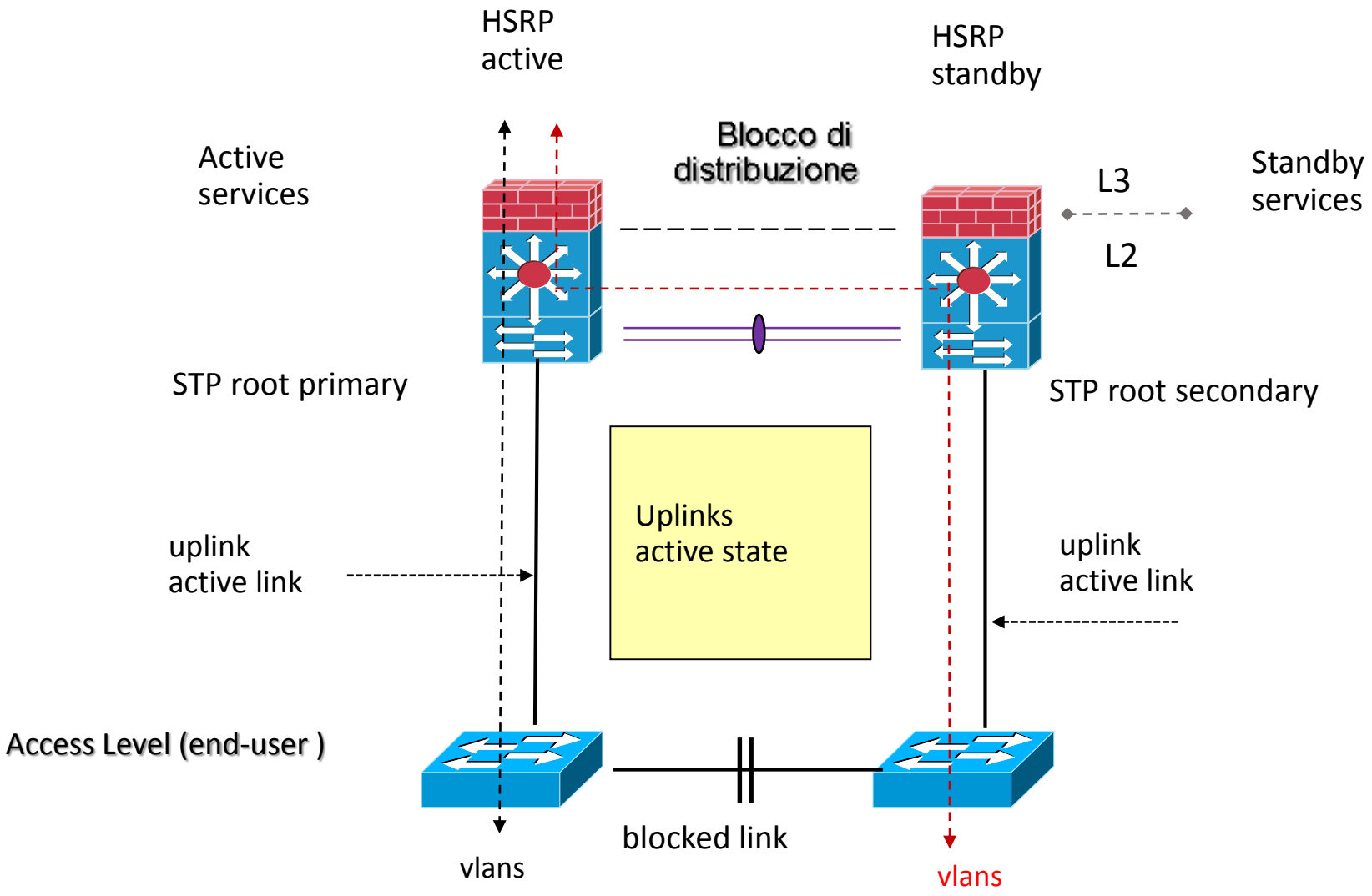
Looped Triangle : provvede ad una ridondanza di tipo active standby tra due peers di aggregazione (Distribution Switch), garantendo HA tramite protocolli HSRP o VRRP e, per effetto dello spanning tree, abbiamo un link in stato active e l'altro in stato blocked (o standby) tra il livello di accesso ed il livello di aggregazione allineando di fatto tutti i componenti e protocolli di rete performati a livello aggregation in stato active solo su uno dei due peers di aggregazione. Il throughput di banda a disposizione è pari al 50% del valore effettivo utilizzabile; non è presente nessun collegamento tra gli switch di accesso (access inter-switch link)

Looped Square : rispetto al modello di cui sopra, la differenza consiste nel permettere collegamento tra gli switch di accesso (access inter-switch link) in stato blocked; permette un solo collegamento tra lo switch di accesso e lo switch di aggregazione.

Looped Triangle Switching Architectures Design Cisco Models



Looped Square Switching Architectures Design Cisco Models



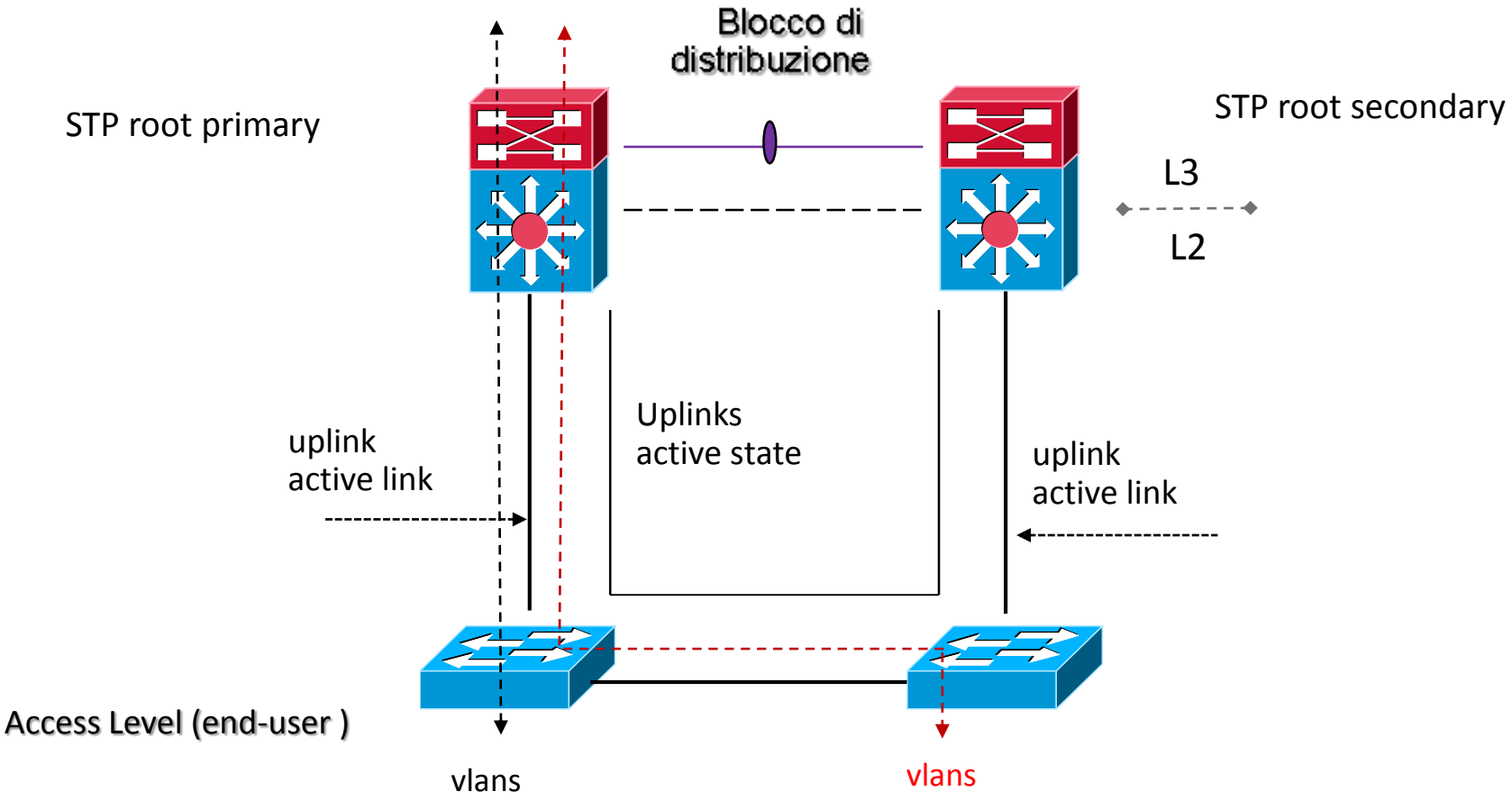
Looped-FREE Switching Architectures Design Cisco Models

Layer 2 Looped Free : le vlans non sono estese a livello di aggregazione; il livello 3 (routing) è performato dal livello di aggregazione in su; lo spanning-tree protocol (STP) in questo scenario è in uno stato di background (non è assente) in caso di fault a livello fisico (cabling); Il throughput di banda utilizzato è del 100%.

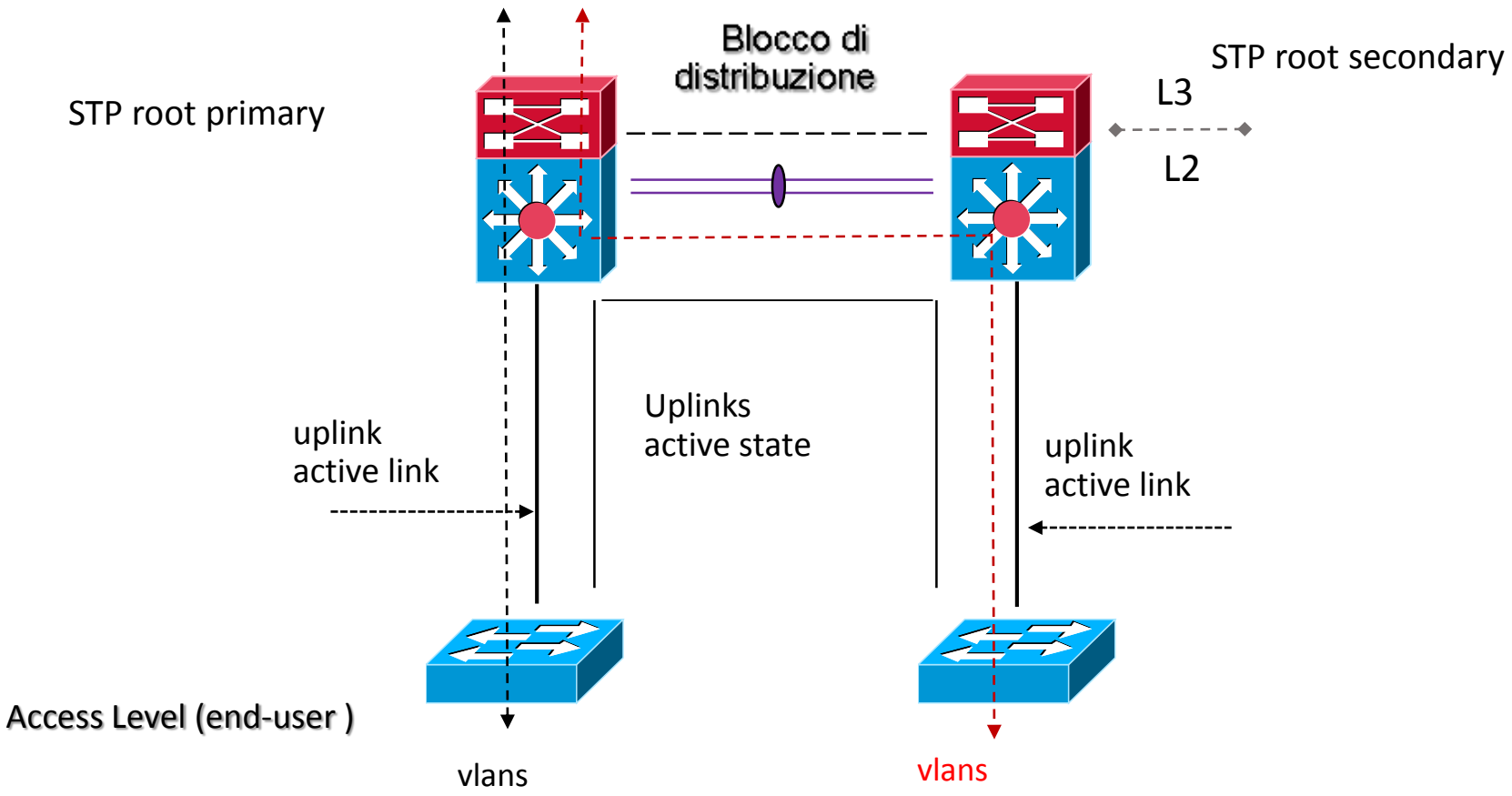
Loop-Free type U : in questo modello le vlans transitano tra gli switch di accesso avendo un collegamento tra loro (inter-switch link) e terminano per L3 SVI gateway tra i due peers di distribuzione (non tutti i modular services sono supportati da questo modello): lo svantaggio prevede un black-holing traffic a causa di un eventuale fault di un singolo link perchè le vlans non sono estese a livello di switch di aggregazione.

Loop-Free Inverted U : rispetto al modello di cui sopra, la differenza consiste nel permettere la distribuzione delle vlans a livello di aggregazione e non attraverso gli switch di accesso; tutti i modular services supportano questo modello

Loop-Free type U Switching Architectures Design Cisco Models



Loop-Free type inverted U Switching Architectures Design Cisco Models



Flex-Link Switching Architectures Design Cisco Models

Flexlink : ogni switch di accesso ha due link verso gli switch di aggregazione con una configurazione di tipo Flex-Link

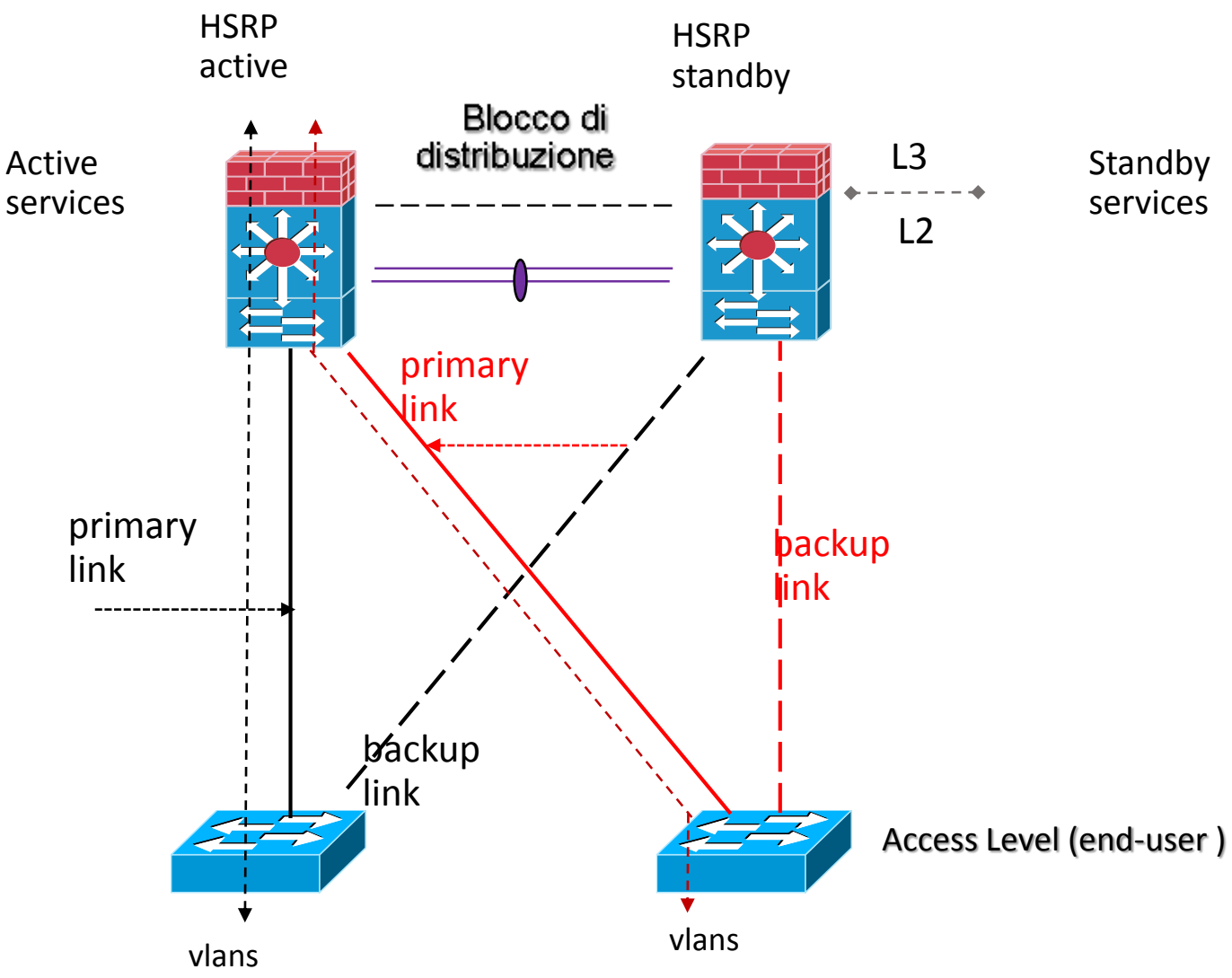
Flex-link disabilità il protocollo STP (No BPDU propagation) rendendo non conforme il modello di ricalcolo new-path in caso di fault di un link a livello fisico; in ogni caso il failover tra i due links è all'interno di 1 o 2 secondi

Gli switch di distribuzione non sono consapevoli della configurazione Flex-Link

La configurazione prevede che a livello di link primario venga inserito questo comando:

switchport backup *interface interface-id*

Flex-Link Switching Architectures Design Cisco Models



Ottimizzazioni architetture layer 2

- STAR switching Cisco Models
- VSS Virtual Switching System
- MEC Multi-Chassis Etherchannel
- Stackwise Virtual System
- vPC Nexus archiecture Cisco

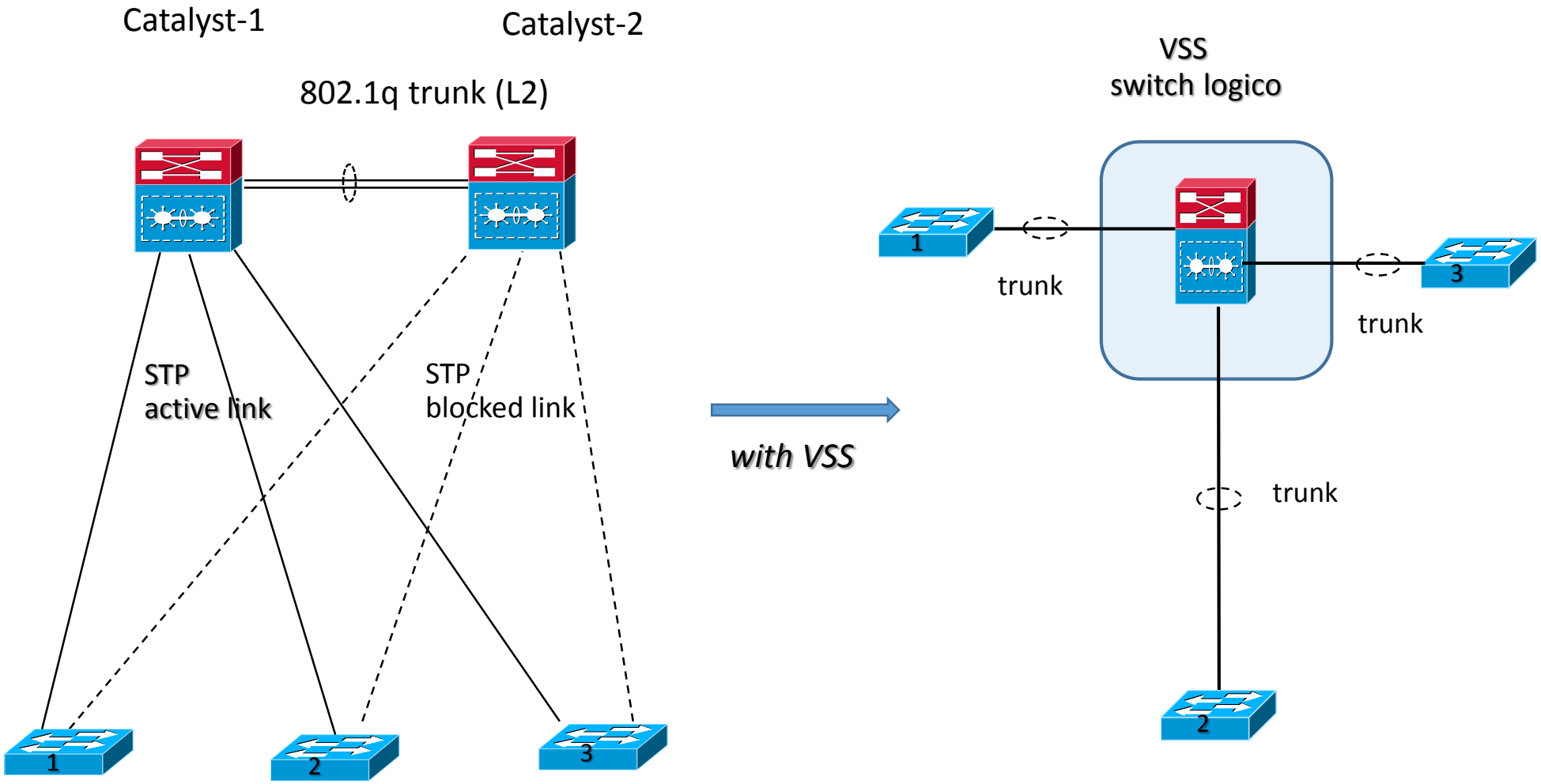
STAR Switching Architectures Design Cisco Models

Star : utlizzata quando un sistema VSS è impiegato insieme ad una configurazione MEC (MultiChassis Etherchannel)

Questa configurazione logicamente significa un solo switch logico di aggregazione avente un solo link MEC in trunk verso lo switch di accesso

Tre switch di accesso sono quindi rappresentati come tre raggi ciascuno collegati in modo indipendente al proprio hub di centro-stella rappresentato dalla coppia di switch VSS.

STAR Switching Architectures Design Cisco Models



50% di utilizzo della rete layer 2

100 % di utilizzo della rete layer 2

VSS Virtual Switching Cisco Models

Un sistema VSS opera via SSO (Stateful Switch Over) tra peers attraverso le Supervisor Engine active e standby ospitate nei rispettivi chassis

VSS Supervisor Engine active controlla le funzionalità layer 2 (switching) and layer 3 (routing) per entrambi gli chassis (single control-plane, single management point, dual active forwarding planes)

Il piano di forwarding del traffico è performato da entrambi i peers VSS

In caso di fault della Supervisor Engine active, quella in stato standby assume il suo ruolo(switchover)

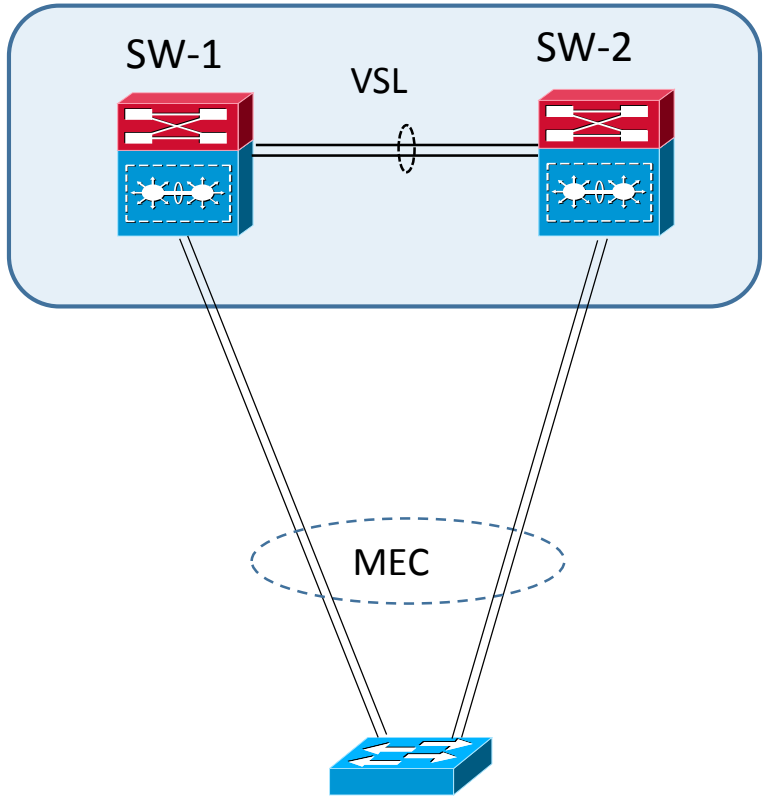
VSL (Virtual Switch Link) è un collegamento tra i peers VSS per lo scambio di messaggi di controllo processati dalla Supervisor Engine active ma trasmessi e ricevuti su interface presenti nel peer VSS standby

VSS opera in un contesto di Spanning Tree Protocol; il VSS standby redirige le BPDU STP via VSL verso il peer active

Il STP bridge ID è un valore comune ed è calcolato sul MAC address chassis; non cambia a seguito di uno switchover peers

VSS Virtual Switching Cisco Models

```
SW1#
redundancy
 mode sso
!
router
 nsf
!
switch virtual domain 10
 switch 1
!
interface port-channel 10
 switch virtual link 1
 no shut
!
interface range gi0/1-2
 channel-group 10 mode on
 no shut
!
switch convert mode virtual
```




```
SW2#
redundancy
 mode sso
!
router
 nsf
!
switch virtual domain 10
 switch 2
!
interface port-channel 20
 switch virtual link 2
 no shut
!
interface range gi1/5-6
 channel-group 20 mode on
 no shut
!
switch convert mode virtual
```

MEC Multi-Chassis Etherchannel Cisco Models

Un MEC è un collegamento Multihassis Etherchannel tra un nodo di accesso verso entrambi i peers VSS active e standby

Un VSS MEC può collegare qualsiasi elemento di rete che supporti etherchannel (quale host, server, switch, router)

Un VSS MEC supporta protocolli quali LACP (Link Aggregation Protocol) oppure PAgP (Port Aggregation Protocol)

Il MSFC (Multilayer Switch Feature Card) presente nella Supervisor Engine active lavora a livello 3 (routing) ed entrambi i peers VSS performano il forwarding del traffico sulle rispettive interface sia in ingresso che in uscita

In genere un traffico in ingresso è trasmesso (forwarding) da una interfaccia di uscita appartenente allo stesso chassis per ridurre cosi la quantità di dati transitante via VSL.

Stackwise Virtual System Cisco Models

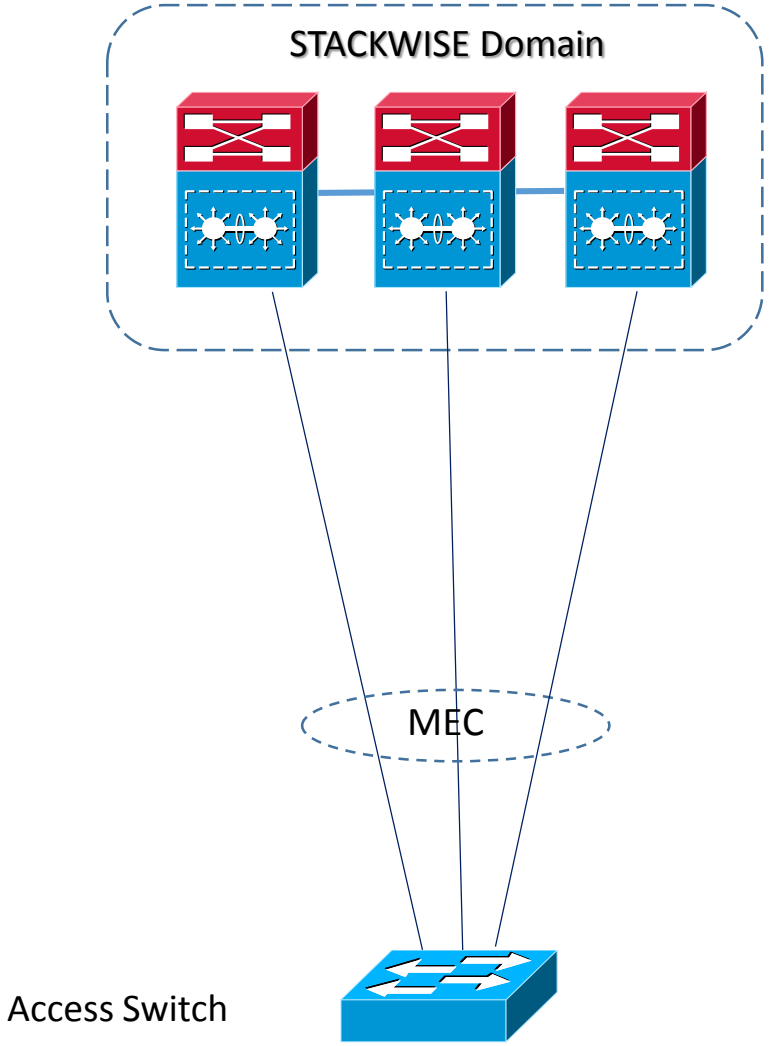
Un sistema “stackwise virtual” cisco opera in modo che due (o più) switch catalyst sono unificati all’interno di un’unico sistema di switching:

- con un solo piano di controllo (control-plane) ed un piano di management (gestione);
- un piano di forwarding distribuito (significa che ogni switch facente parte dello stackwise è capace di trasmettere traffico attraverso le sue porte locali senza coinvolgere quelle di altri switch membri;
- un pacchetto entrante in una determinate porta locale, comunque, è soggetto ad una trasmissione su differente porta distribuità su tutti gli switch membri sino a considerare lo switch egress per quella specifica destinazione

Attraverso un meccanismo di elezione un membro dello stack viene eletto master e gli altri standby (lo switch master è responsabile del piano di controllo, di forwarding e di management.

Gli switch membri dello stack comunicano tra loro attraverso un virtual software module chiamato VCM (Virtual Communication Manager) via links.

Stackwise Virtual System Cisco Models



vPC Architectures Nexus Design Cisco Models

Con vPC, ogni Nexus mantiene il proprio control-plane ed il proprio management:

- vPC IEEE 802.3ad: è un port-channel tra un devices in downstream e due Nexus Devices con stessa release software;

- vPC peer: è uno dei due devices (or VDC) che formano la coppia di Nexus in aggregazione;

- vPC member port: è una interfaccia che appartiene ad uno specific vPC port-channel di uno dei due vPC peers;

- vPC domain: un unico identificativo per coppia di Nexus (ogni Nexus switch or VDC supporta un solo dominio);

- vPC peer-link: usato per la sincronizzazione dei rispettivi status switches e per il forward del traffic o tra i due peers;

- vPC peer-keepalive: usato per la verifica heartbeat tra I due switches vPC peers (in modo esplicito per verificare failure dei peer-link e peer-keepalive);

- CSF (Cisco Fabric Services): è automaticamente abilitato in un vPC e rappresenta la capacità di sincrinizzare lo stato e la configurazione tra i due vPC peers

vPC Architectures Nexus Design Cisco Models

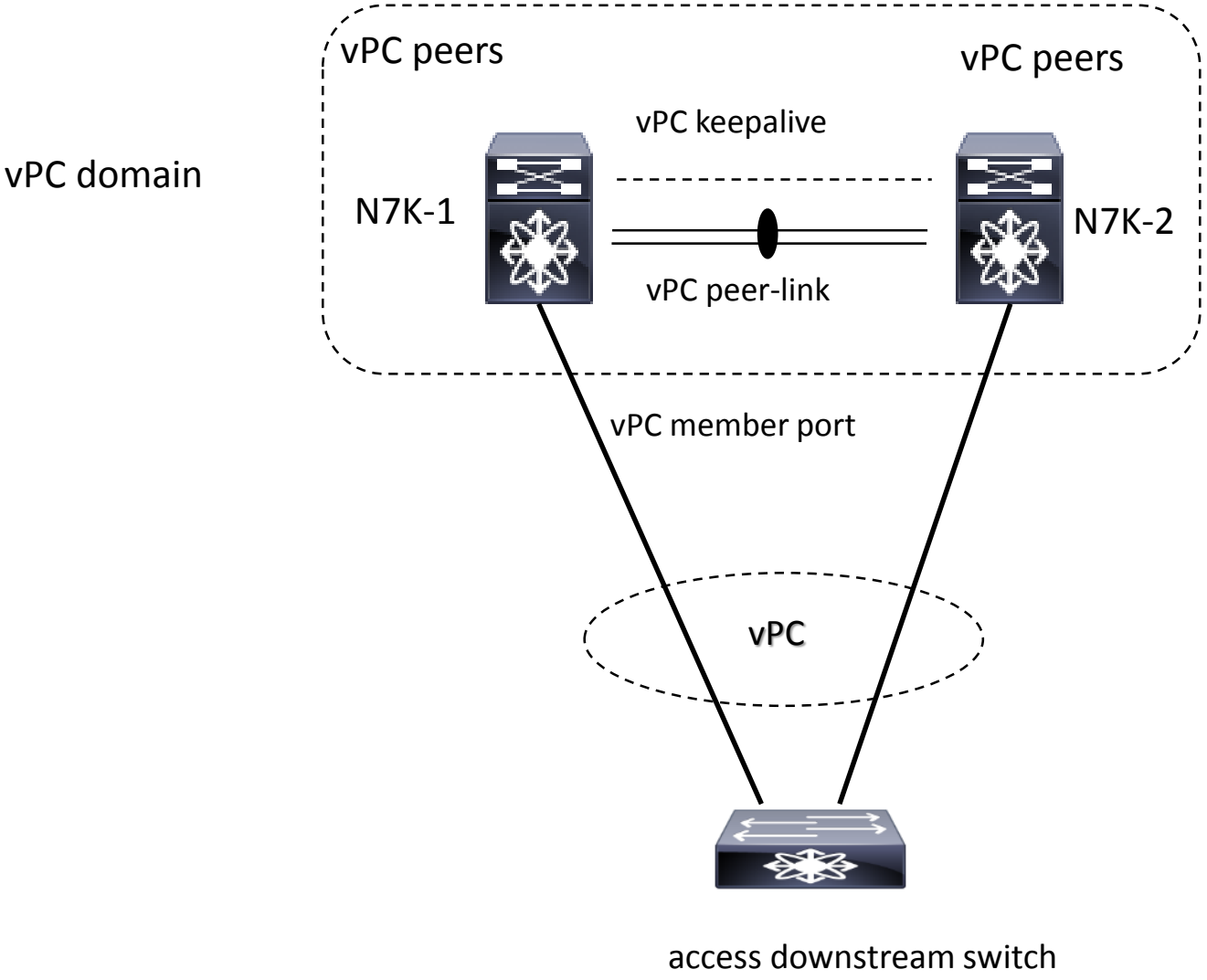
NOTA CISCO: quando si configura un vPC domain tra i due vPC peers, questi generano un condiviso MAC address che viene usato come un logico switch bridge ID in SpanningTree Protocol.

All’interno di un vPC domain ad ogni peer è assegnato un ruolo: primario e secondario (di default lo switch con il più basso MAC address diventa il primario; è un valore comunque che l’operatore può cambiare)

La comunicazione per il control-plane avviene attraverso il peer-link tra i due peers.

Esiste un meccanismo di loop-avoidance split-horizon loop via port-channeling dove il traffico entrante in un port-channel non può uscire dallo stesso port-channel)

vPC Architectures Nexus Design Cisco Models



Layer 2 security

- DHCP Snooping
- IP source guard
- Unicast Reverse Path Forwarding
- MAC address limitng (port security)

DHCP SNOOPING

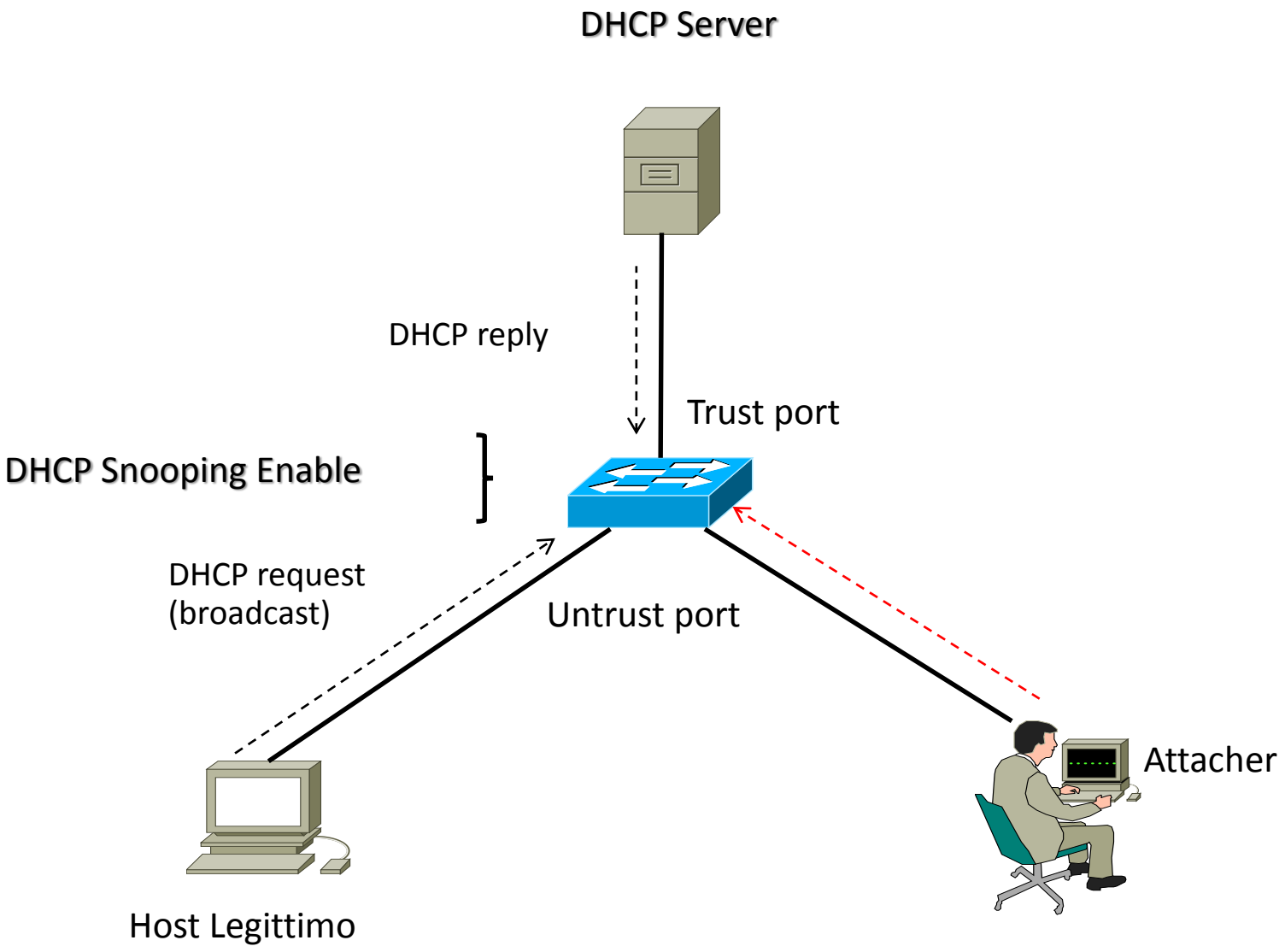
DHCP Snooping è una tecnica di sicurezza attraverso il filtering di DHCP messages ed attraverso un database che permette il controllo di questi messaggi ed accessi.

DHCP Snooping agisce come un firewall tra untrusted host ed il DHCP Server, permettendo solo messaggi di tipo trusted.

Quando DHCP Snooping è abilitato a livello switch, le porte sono classificate come trustred oppure untrusted ; le porte trusted hanno il permesso di trasmettere tutti i tipi di messaggi DHCP, viceversa le porte untrusted hanno autorizzazione a richiedere solo richieste DHCP (se lo switch vede un DHCP reply attraverso una porta untrusted questa viene disabilitata (shutdown)

E’ solito utilizzare DHCP Snooping con IP source guard dove quest’ultimo controlla sia il MAC sorgente associato all’indirizzo IP verificando il match con il DHCP database (se il match è negativo la frame è filtrata)

DHCP SNOOPING



IP Source Guard e Unicast Reverse Path Forwarding

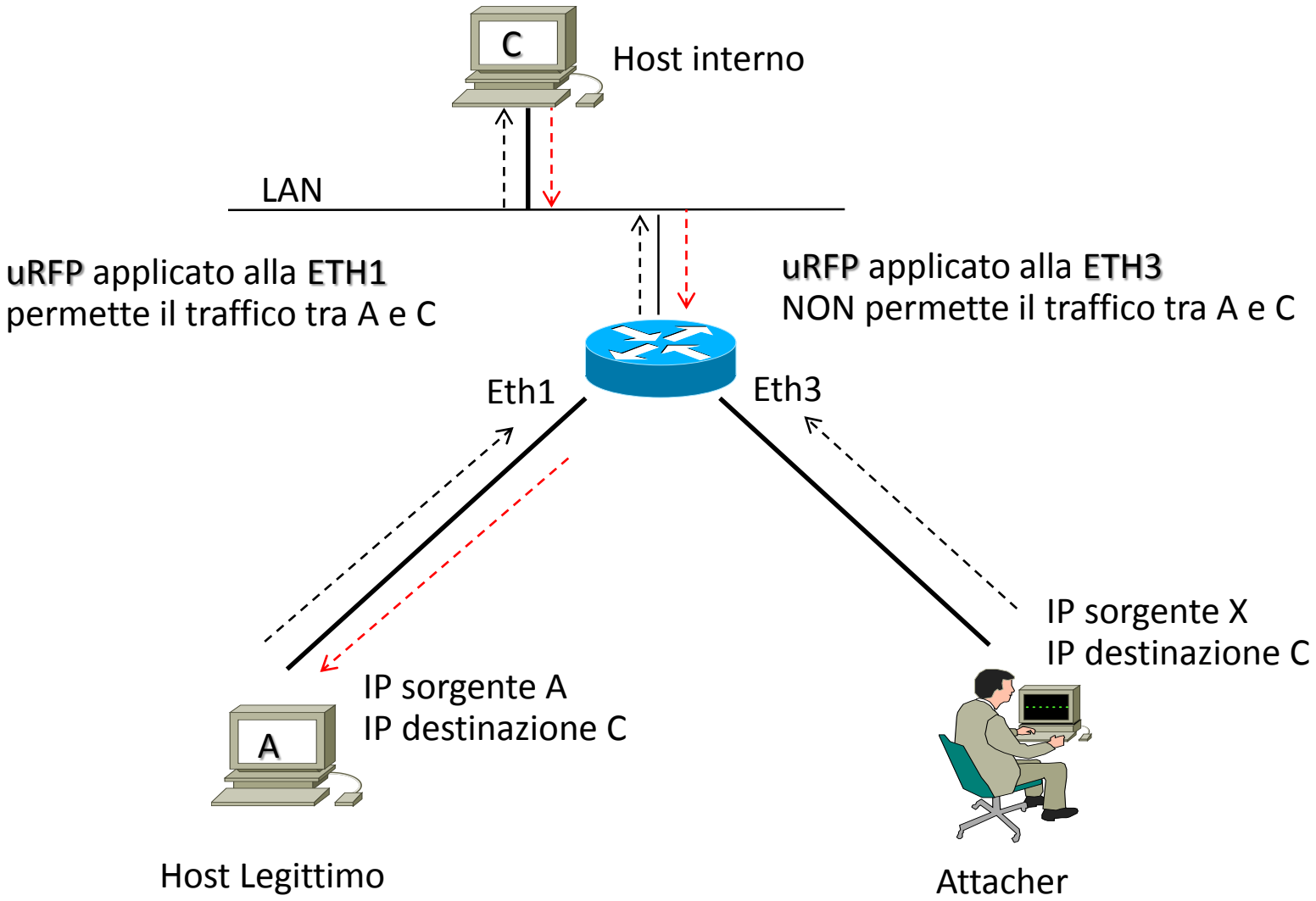
IP Source Guard è una tecnica di sicurezza contro IP spoofing dove un attacher utilizza un indirizzo IP legittimo della rete per cercare di guadagnare un accesso ad essa.

IP Source Guard, quindi, controlla e verifica l'indirizzo IP di un host associato ad una determinata porta di uno switch e previene traffic o dati se sorgente da un differente indirizzo IP da quello legittimo.

IP Source Guard può verificare anche il binomio IP address e MAC address di un host collegato allo switch

Unicast Reverse Path Forwarding (uRPF) garantisce sicurezza attraverso la verifica dell'indirizzo IP sorgente per traffico transitante attraverso un router, con il drop dei pacchetti se l'indirizzo IP sorgente non è compatibile e verificato con quello legittimo

IP Source Guard e Unicast Reverse Path Forwarding



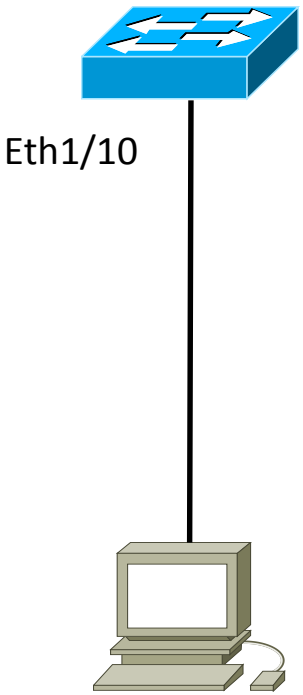
MAC Address Limiting (port security)

E’ una funzionalità di protezione layer 2 switching, applicata a livello di porta di accesso, contro attacchi che usano MAC addresses quali MAC flooding e MAC spoofing (DoS attack).

MAC limit: permette di specificare il numero massimo di MAC addresses che possono essere appresi attraverso una singola porta di accesso; una volta che lo switch raggiunge il numero limite di MAC, tutto il traffico sorgente da nuovi MAC address sono droppati, sulla base della azioni previste in configurazione

MAC allowed: permette di definire MAC addresses per una specifica porta di accesso; qualsiasi MAC addresses che non è specificato nella lista per quella determinata porta non sarà preso in considerazione e pertanto negato.

MAC Address Limiting (port security)




```
interface gigabitethernet 1/10
switchport mode access
switchport port-security mac-address sticky
---
switchport port-security maximum  <value>
---
switchport port-security violation  [ restriction | shutdown ]
```