



# **Felhő alapú hálózatok (VITMMA02)**

## **Data Center Bridging, Virtuális hálózati technológiák**

Dr. Maliosz Markosz

Budapesti Műszaki és Gazdaságtudományi Egyetem  
Villamosmérnöki és Informatikai Kar  
Távközlési és Médiainformatikai Tanszék

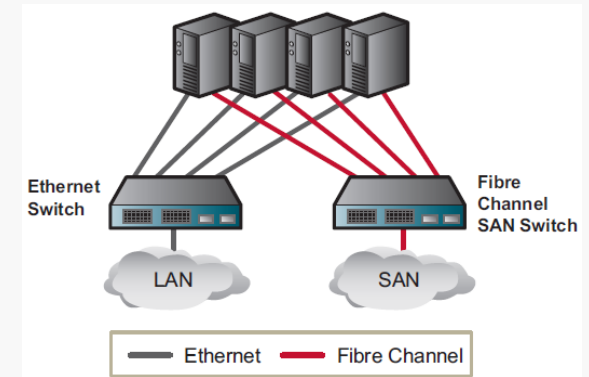
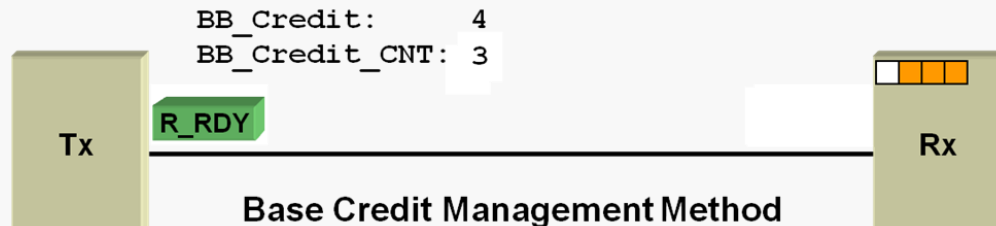
2017. tavasz



# DATA CENTER BRIDGING

# Háttértár forgalom az adatközpontban

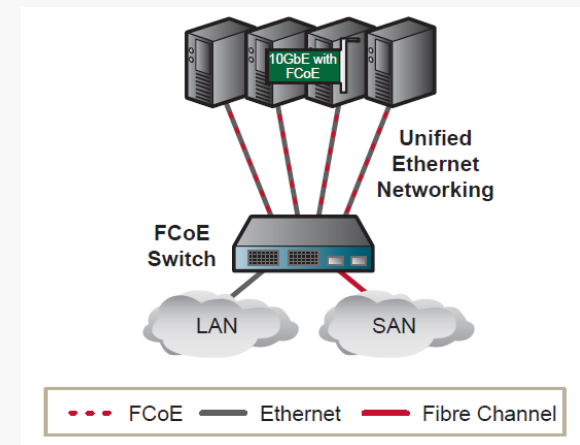
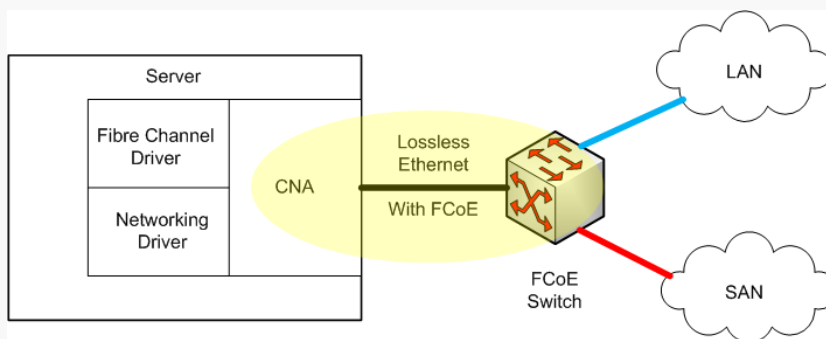
- » Adatközpontban korábban
  - » Ethernet adatforgalomra
  - » Fibre Channel a háttértár forgalomra (SAN – Storage Area Network)
    - » külön dedikált hálózat
    - » optikai vagy elektronikus interfész
    - » 2, 4, 8, 10, 16, 32, 128 Gbps
    - » nincs csomageldobás torlódás esetén
      - » puffer kredit alapú folyamvezérlés
      - » buffer to buffer credit



# Fibre Channel over Ethernet (FCoE)

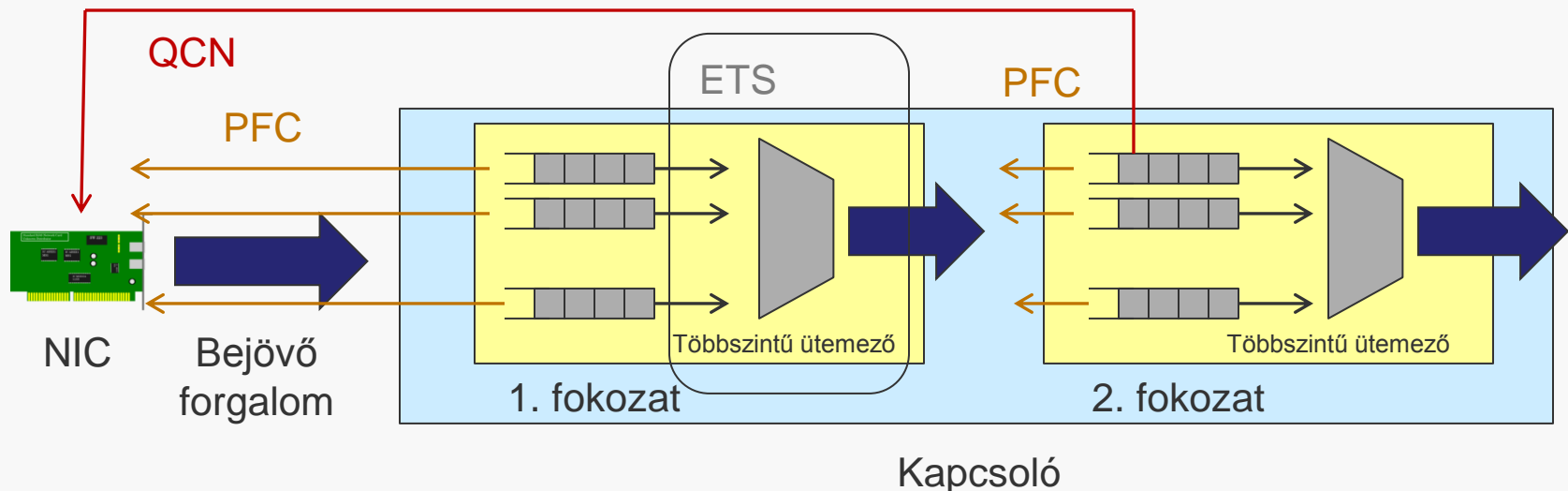
## » Ethernet

- » torlódás esetén csomageldobás
- » TCP garantálja a megbízható átvitelt (újraküld)
  - » késleltetés ingadozás
  - » video és háttértár forgalom számára ez nem ideális
- » kiegészítések szükségesek: DCB



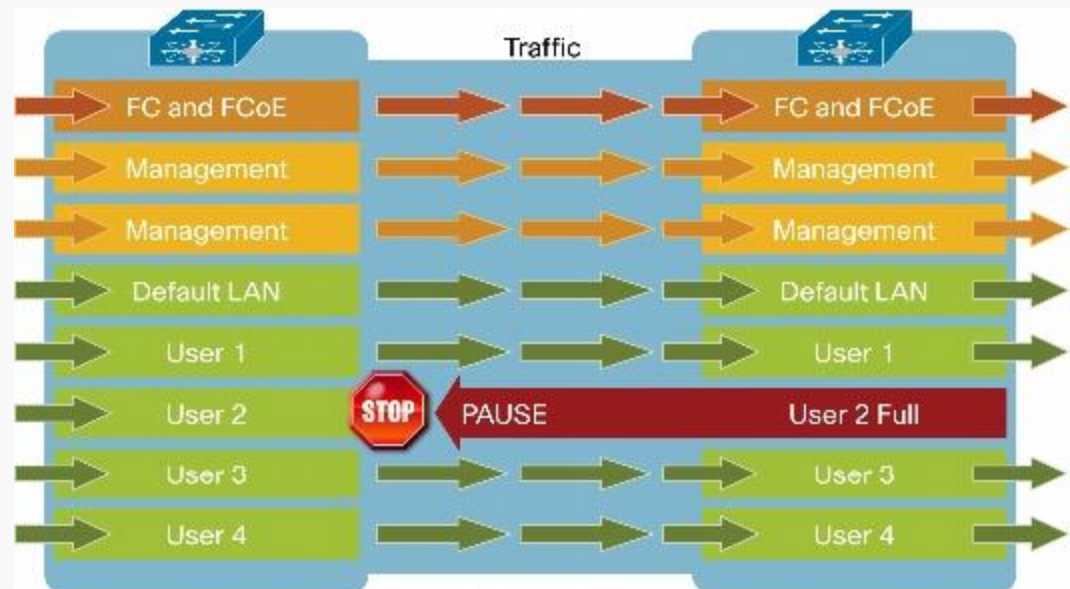
# Data Center Bridging

- » Ethernet kiegészítések: megbízható(bb) átvitel elérésére a TCP komplexitása nélkül
  - » Priority based Flow Control (PFC)
  - » Enhanced Transmission Selection (ETS)
  - » Quantized Congestion Notification (QCN)
  - » Data Center Bridging exchange (DCBx) protocol



# Priority based Flow Control

- » A torlódásból adódó csomageldobás kiküszöbölésére
- » IEEE 802.1Qbb
  - » link szintű
  - » kapcsolók, vagy kapcsolófokozatok között
- » 8 prioritás (802.1p): virtuális sávok
- » kapcsolóban: memória partíciók
  - » telítettségi szint
- » szünet üzenet: időtartamot tartalmaz



Forrás: Cisco



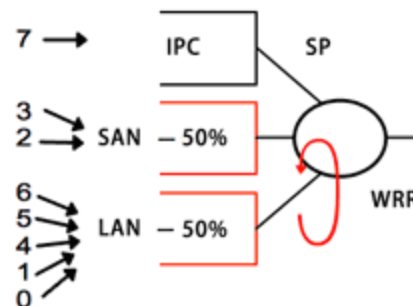
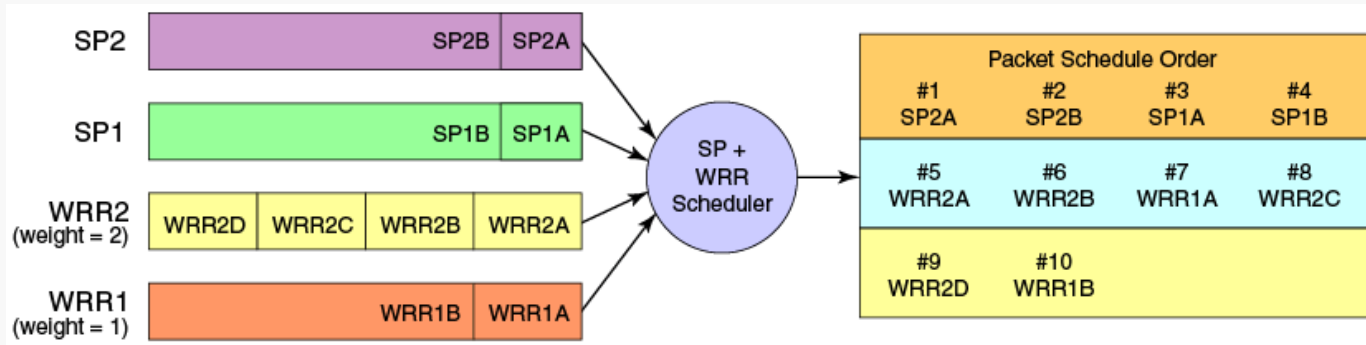
# Enhanced Transmission Selection

- » IEEE 802.1Qaz
- » Forgalmi osztályok
  - » osztályba sorolás
    - » szabályok alapján a fejléc mezők vizsgálatával: Access Control List (ACL)
    - » VLAN címke 3-bites prioritás mezője
  - » ütemező az osztályok csoportjait is kezelheti (Traffic Class Group – TCG)
    - » min. 3-at kell tudni az ETS-képes kapcsolónak



# Enhanced Transmission Selection

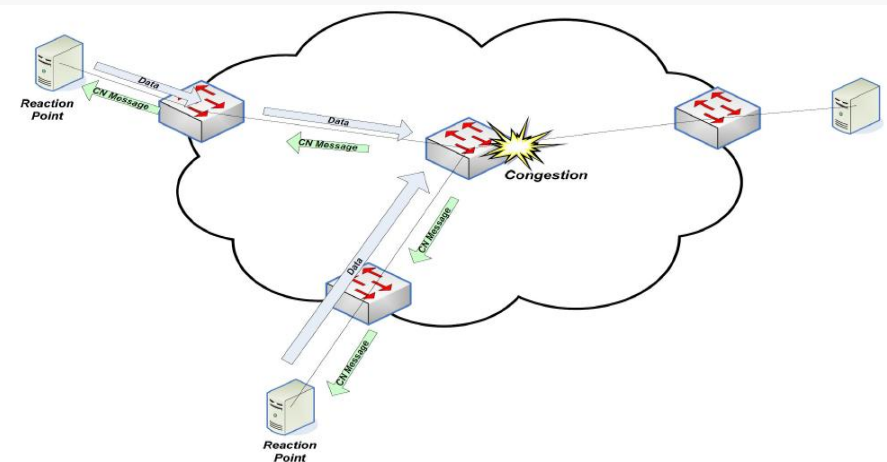
- » Sáv szélesség allokáció
  - » osztály csoportonként (max. 8) minimum garantált bitsebesség
    - » 1%-os egységekben,  $\pm 10\%$  pontosság
    - » a kihasználatlan sáv szélességet a többi osztály ki tudja használni
- » Megvalósítása: ütemezés (scheduling) és forgalom szabályozás (rate limiting, shaping)





# Quantized Congestion Notification

- » PFC + ETS
  - » veszteségmentes átvitel és sávszélesség garancia
  - » gyors reakcióidő
  - » azonban: többfokozatú kapcsolókban és adatközpontokban sok kapcsolón áthalad a forgalom
- » QCN (802.1Qau): tranziens torlódások kiküszöbölésére
  - » a forgalom forrását értesíti (ent-to-end)
  - » nagyobb időskála
  - » torlódási pont
    - » sorhosszak vizsgálata, mintavételezés (függ a telítettségtől)
    - » a telítettségtől függően visszajelzési érték kiszámítása (6 bitre kvantálás)
    - » forrás MAC címre küldés (beavatkozási pont)
      - » 1-10% közti vg.-gel
      - » mintavételezési idő utánállítás
  - » beavatkozási pont
    - » kimenő forgalom korlátozása a visszajelzett érték függvényében
    - » idővel újra növeli a sebességet





# Quantized Congestion Notification

- » Ritkán implementálják...
  - » a szabályozási kör sok tényezőtől függ
    - » torlódási pont reakcióideje, visszajelzés megérkezési ideje a forráshoz, a beavatkozási pont reakcióideje
  - » finom beállítást igényel
  - » hosszú ideig tartó forgalomra ideális
  - » a véletlen mintavételezés bizonytalansága
  - » a forrásnál minden torlódási ponthoz külön sort kellene fenntartani
  - » L2 alhálózatban működik
    - » egy útválasztón áthaladva másik QCN tartományba kerül
  - » hardveres implementáció lenne a megfelelő a sebességhez
    - » összes kapcsoló és hálózati kártya cseréje



# Data Center Bridging exchange (DCBx)

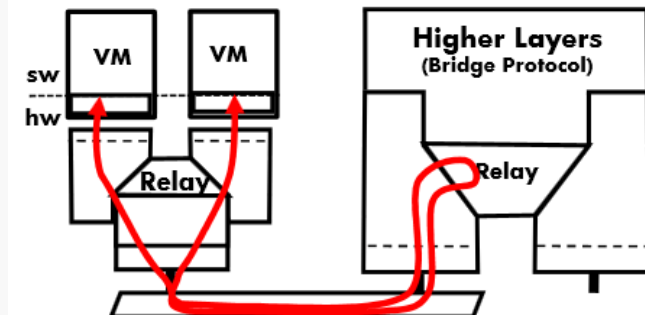
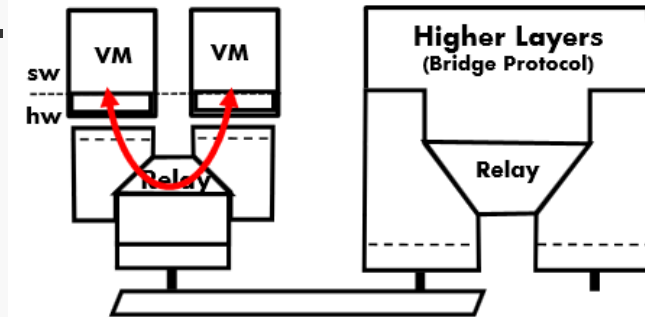
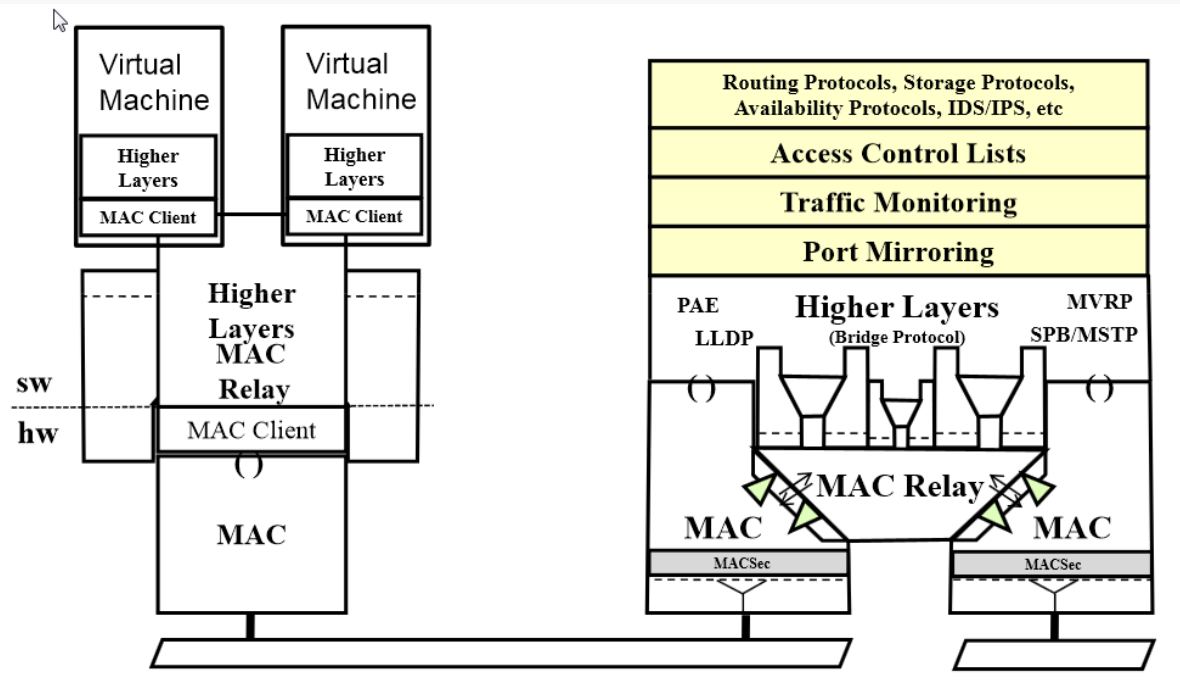
- » Koordináció a képességekről szomszédos kapcsolók között
  - » PFC
    - » prioritások, forgalmi sávok száma
  - » ETS
    - » sávszélesség egységek
- » Link Level Discovery Protocol (LLDP) üzenetekben
  - » Type-Length-Value struktúrában
- » Működés
  - » küldő oldal
    - » *javaslat* a másik oldali eszköznek a paraméterek beállítására
    - » periodikus küldés
  - » vevő oldal
    - » paraméterek beállítása a másik oldali eszköztől kapott konfiguráció figyelembe vételével
    - » adatbázis frissítés a vett adatok alapján
  - » nincsen nyugtázás
    - » nem foglalkozik azzal, hogy a másik oldal mit állít be



# VIRTUÁLIS HÁLÓZATI MEGOLDÁSOK

# Edge Virtual Bridging

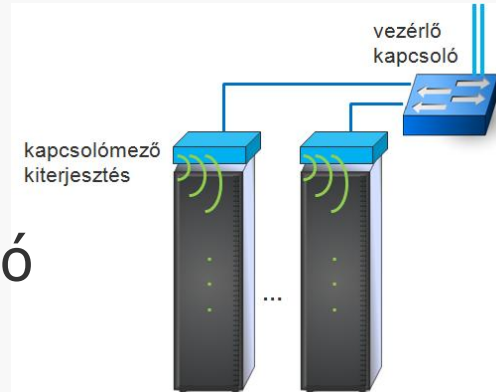
- » ToR fizikai kapcsoló  $\Leftrightarrow$  virtuális kapcsoló (Virtual Ethernet Bridge – VEB) képességek
  - » szűrés, biztonság, monitorozás, stb.



Forrás: Pat Thaler et al., IEEE 802 Tutorial: Edge Virtual Bridging, 2009.

# Edge Virtual Bridging

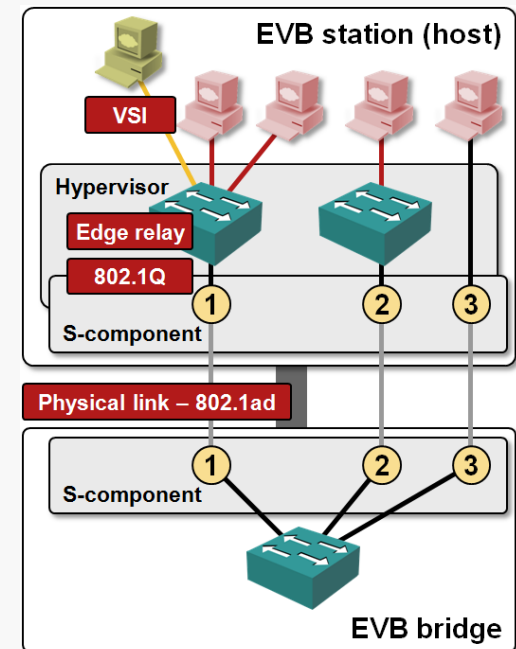
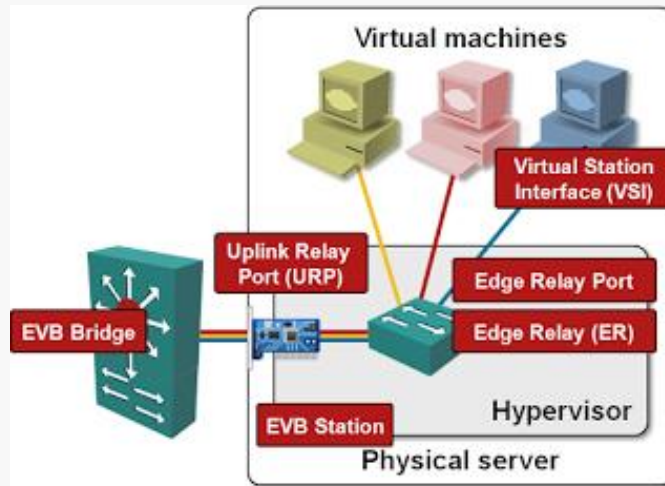
- » EVB: IEEE standard
  - » a virtuális és fizikai kapcsoló közötti interakció
    - » a fizikai kapcsoló képessége
    - » cél: minden forgalom egységes kezelése
  - » Virtual Ethernet Port Aggregation (VEPA) 802.1Qbg
    - » szerver oldali képesség
    - » minden forgalmat továbbít a szomszédos fizikai kapcsolónak
    - » multi-channel: S-Tag (Q-in-Q)
- » Virtuális interfészek azonosítása a fizikai porton
  - » Virtual Network Tag (VN-Tag), Bridge Port Extension 802.1Qbh, 802.1BR (E-Tag)
    - » vezérlő kapcsoló által konfigurált portok
      - » a kiterjesztett kapcsolómezőn (S-Tag)
      - » szerver fizikai hálózati kártyáján (VN-Tag)
        - » minden vNIC-hez külön VN-Tag
        - » extra fejléc, benne Virtual Interface (VIF)



# L2 konfiguráció automatizálása

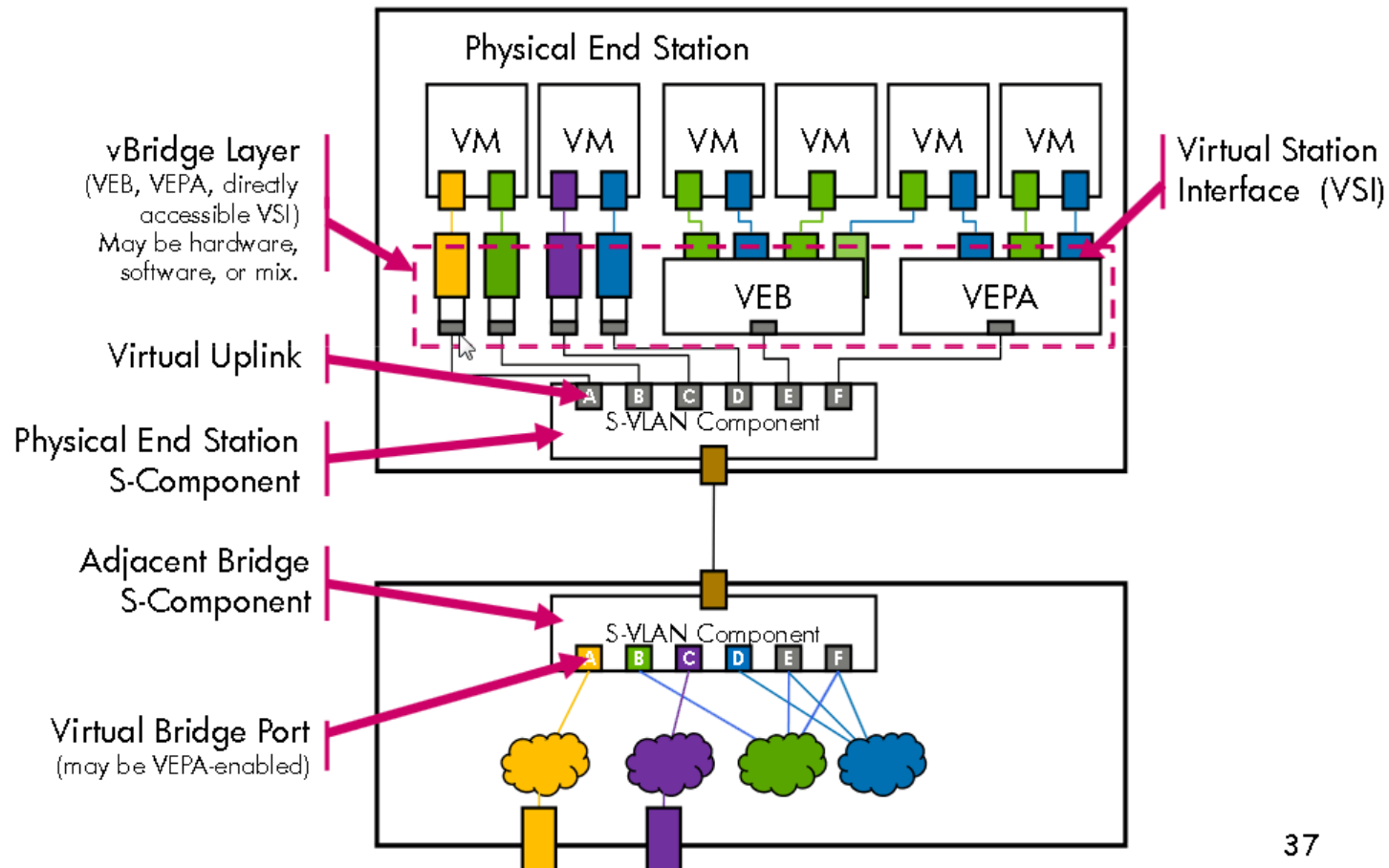
## » Edge Virtual Bridging

- » Virtual Station Interface (VSI): VM hálózati interfésze
- » VSI Discovery and Configuration Protocol (VDP)
  - » az EVB bridge információt kap a VM indítása előtt a hypervisortól
- » VN-Tag: plusz fejléc a virtuális interfészek azonosítására (Cisco)
  - » lokálisan a vezérlő és a kiterjesztett kapcsolómező között
- » S-component
  - » logikai 802.1Q összeköttetések multiplexléása egy fizikai szakaszon (Q-in-Q)



# Edge Virtual Bridging

» lehet kombinálni a megoldásokat



37

Forrás: Pat Thaler et al., IEEE 802 Tutorial: Edge Virtual Bridging, 2009.



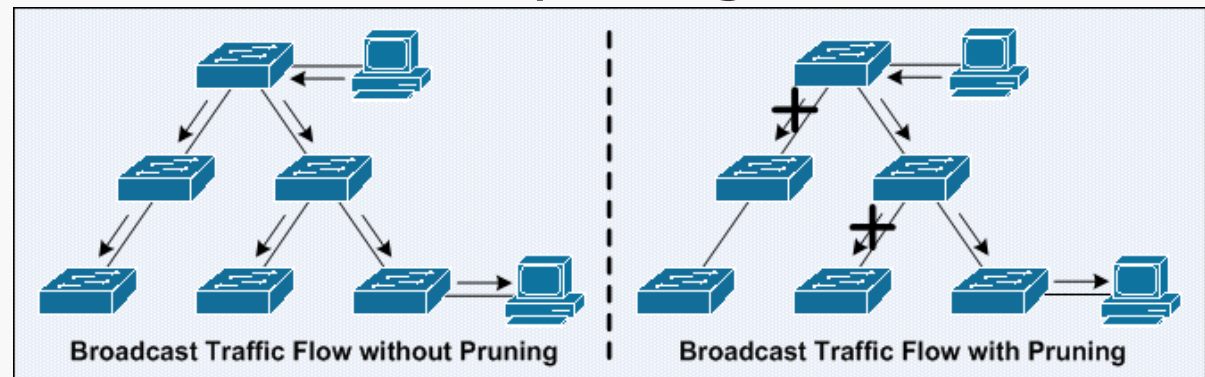


# Összehasonlítás

- » Virtuális kapcsoló (VEB)
  - » továbbítás MAC + VID alapján
  - » nem szükséges
    - » MAC cím tanulás, mert VM címek be lehetnek konfigurálva
    - » STP, mert a hálózat szélén helyezkedik el
  - » forgalom a szerveren belül marad
    - » kívülről nem látható, analizálható, szűrhető
    - » jobb teljesítmény azonos szerveren lévő VM-ek között
  - » nincs egységes menedzselés a fizikai kapcsolókkal
  - » CPU és memória használat
- » Alkalmazás szempontjából
  - » VEPA: hypervisor támogatás szükséges
  - » VN-Tag: speciális hálózati kártya szükséges
  - » irányok
    - » fizikai kapcsoló funkciók integrálása a virtuálisba
    - » más hálózat virtualizáció és alagút megoldások (VXLAN, NVGRE, stb.)
- » EVB
  - » minden forgalom fizikai kapcsolón keresztül (fejlettebb képességek)
  - » kevesebb hálózat konfigurációs feladat
  - » nagyobb forgalom és késleltetés
  - » VEPA
    - » továbbítás MAC + VID alapján
    - » virtuális kapcsoló funkció megmarad
    - » változatlan Ethernet keretek
    - » a fizikai kapcsolón a forgalom visszaküldése a beérkező porton (hairpin)
  - » VN-Tag
    - » továbbítás címke alapján
    - » új keretformátum

# Virtuális hálózati megoldások

- » STP problémák: útvonalválasztás (pl. IS-IS) MAC címekre
  - » Shortest Path Bridging MAC (SPBM)
- » VLAN-ok száma korlátos: még egy VLAN címke hozzáadása
  - » Q-in-Q, provider bridging, (IEEE 802.1ad)
- » MAC cím korlát: még egy MAC cím fejléc hozzáadása
  - » Provider Backbone Bridges (PBB), 802.1ah
  - » Transparent Interconnection of Lots of Links (TRILL)
    - » bridging + routing
- » Hypervisor elárasztás ellen: VM-eket figyelembe venni
  - » VLAN pruning (nyesés, metszés): felesleges forgalom eliminálása
- » Maghálózati elárasztás ellen: VLAN pruning a maghálózatban





# Hálózat virtualizáció

- » VN-Tag csak a VM-et azonosítja, az ügyfelet nem
- » Ügyfél szintű szeparáció támogatás (7. előadás)
  - » Virtual Extensible LAN (VXLAN) – RFC 7348
    - » Cisco, VMware
    - » virtuális L2 hálózati forgalom átvitele L3 fizikai hálózaton
  - » Network Virtualization using Generic Routing Encapsulation (NVGRE)
    - » Microsoft, Intel, HP, Dell
  - » Generic Network Virtualization Encapsulation (GENEVE)
    - » a fenti kettő fúziója
  - » Stateless Transport Tunneling (STT)
    - » Nicira ⇔ VMware



# Források

- » Pat Thaler et al., IEEE 802 Tutorial: Edge Virtual Bridging, 2009.
- » Overlay Virtual Networking Explained, Ivan Pepelnjak, NIL Data Communications, 2011.