



# **Felhő alapú hálózatok (VITMMA02)**

## **Szerver virtualizáció, Adatközpont hálózatok**

Dr. Maliosz Markosz

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

2017. tavasz



# **SZERVER VIRTUALIZÁCIÓ**



# Szerver virtualizáció

- » Alacsony szerver kihasználtság  $\Rightarrow$  virtualizálás
  - » PC-k, szerverek: 10%
  - » háttértár: 50%
- » Szerver CPU és hálózati sávszélesség kihasználtság nő
  - » Jelenleg: 2-4(-10) VM / fizikai szerver (Virtual/Physical Machine)
  - » 2018-ra előrejelzés: 16 VM/PM
  - » kis teljesítményigényű folyamatok esetén akár 100 VM/PM
- » Hypervisor
  - » virtual machine monitor/manager (VMM)
  - » lehetővé teszi VM-ek futtatását
    - » memória kezelés, CPU ütemezés
  - » hardver: hoszt, VM: vendég (guest)

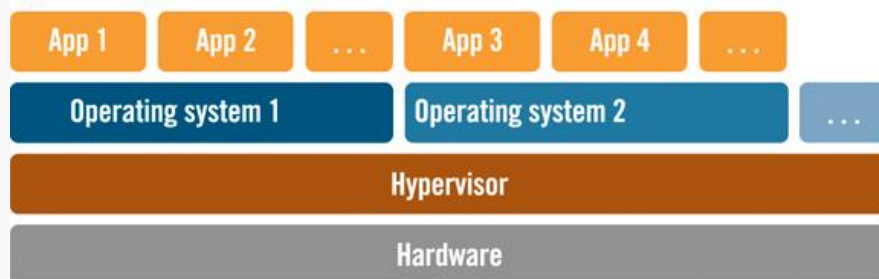


# Szerver virtualizáció

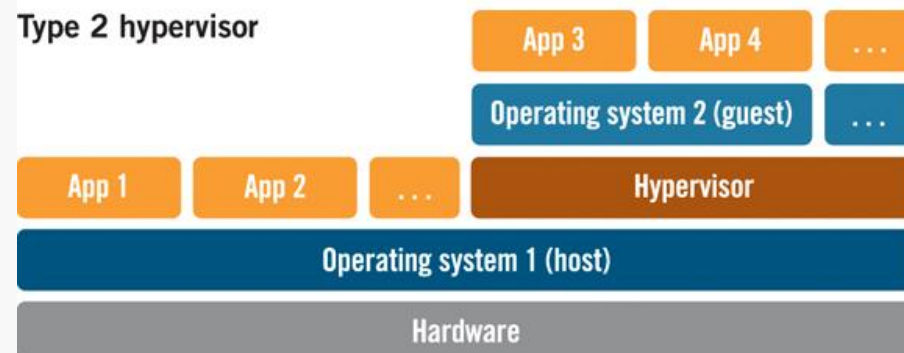
## » Hypervisor

- » típusok, besorolás: Gerald J. Popek és Robert P. Goldberg, „Formal Requirements for Virtualizable Third Generation Architectures”, 1974
  - » 1: natív (bare metal)
    - » a hypervisor közvetlenül a hardveren fut
    - » pl. Citrix XenServer, VMware ESX/ESXi, Microsoft Hyper-V
  - » 2: hosztolt
    - » a hypervisor operációs rendszeren fut (host), futtatott VM: guest
    - » pl. VMware Workstation/Player, VirtualBox
  - » egyéb: Linux Kernel-based VM (KVM)
    - » kernel modulként fut, a hoszt OS-t 1-es típusra konvertálja
    - » inkább a 2-es típusba sorolják

Type 1 hypervisor

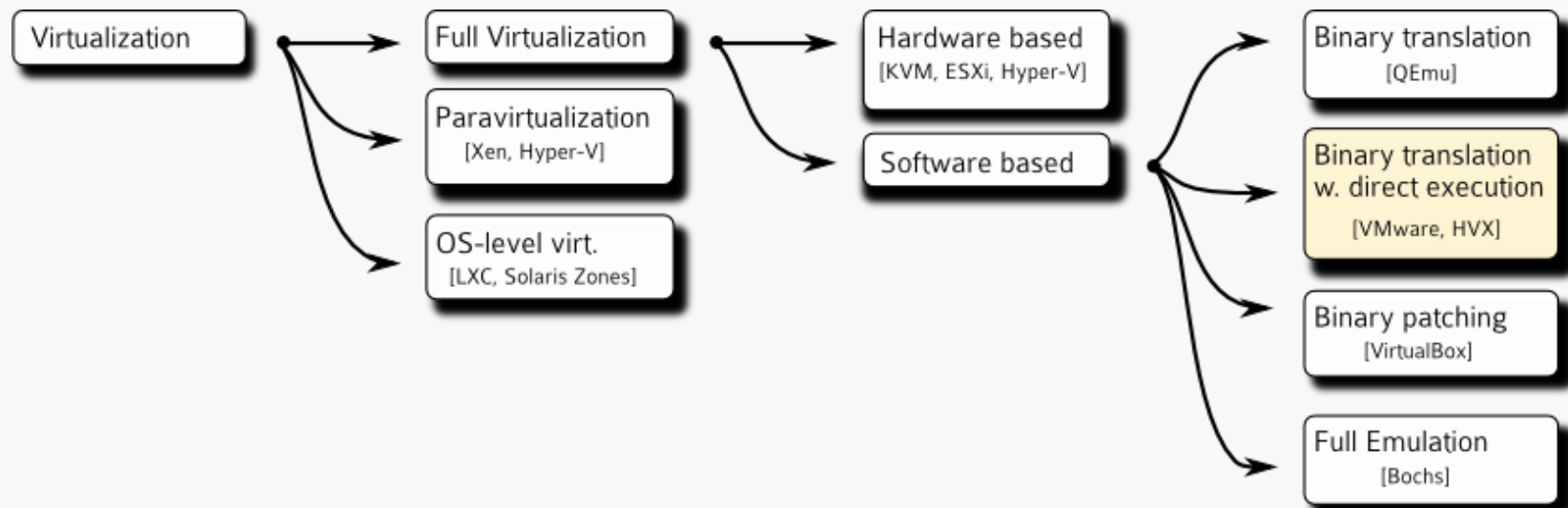


Type 2 hypervisor



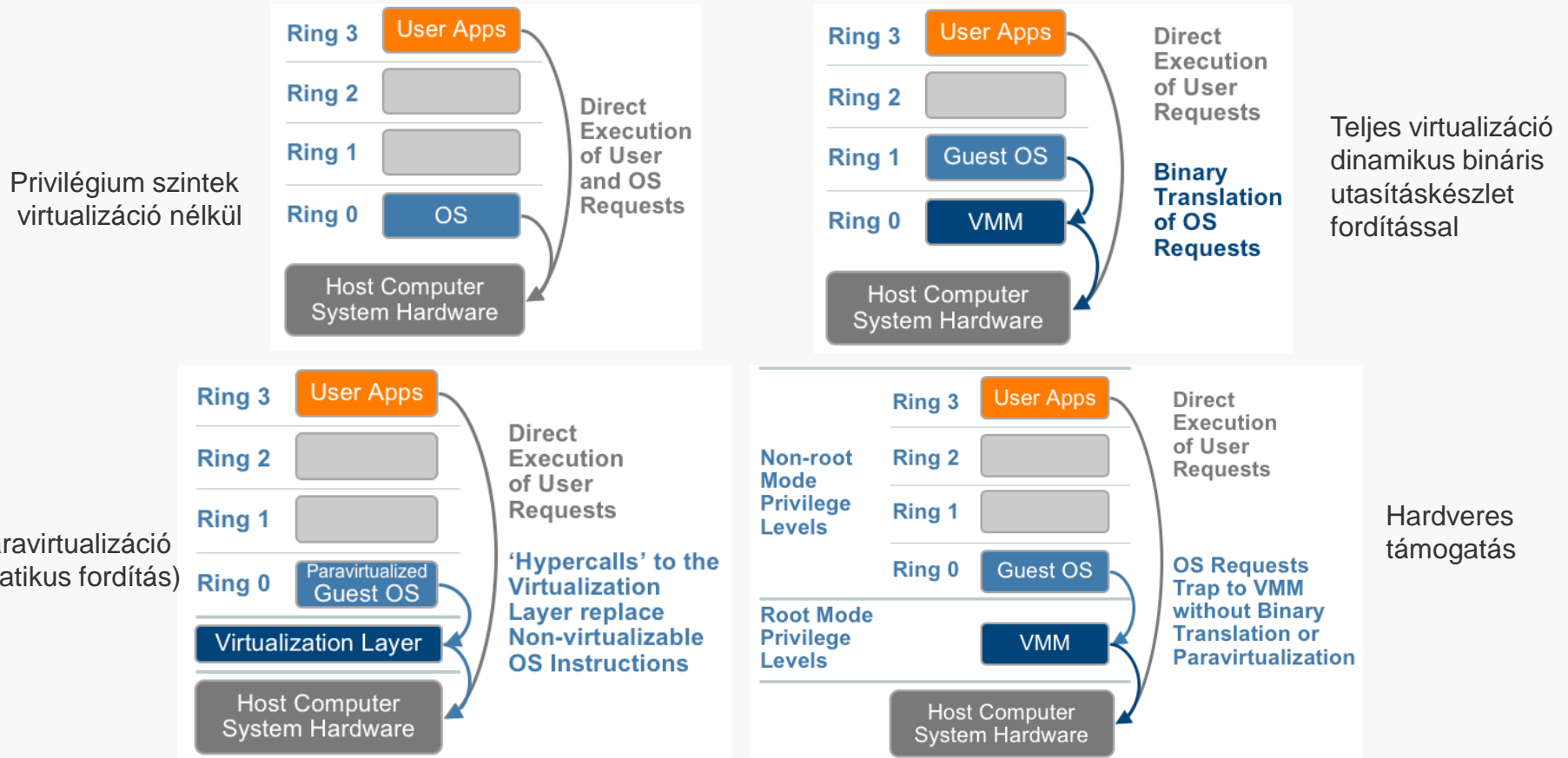


# Virtualizáció típusai



# CPU Virtualizáció

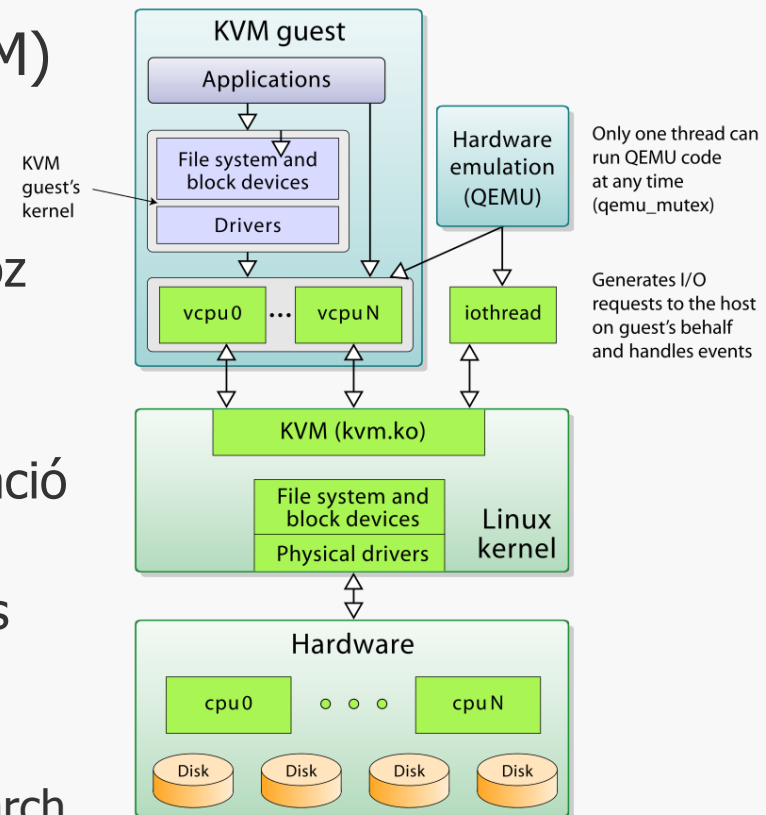
- » 2005-2006. hardver támogatás a virtualizációhoz: Intel VT-x és AMD-V
- » Virtualizációs szoftverek elterjedése
- » x86 CPU virtualizáció



Ábrák forrása: VMware, Understanding Full Virtualization, Paravirtualization, and Hardware Assist, White Paper, 2007

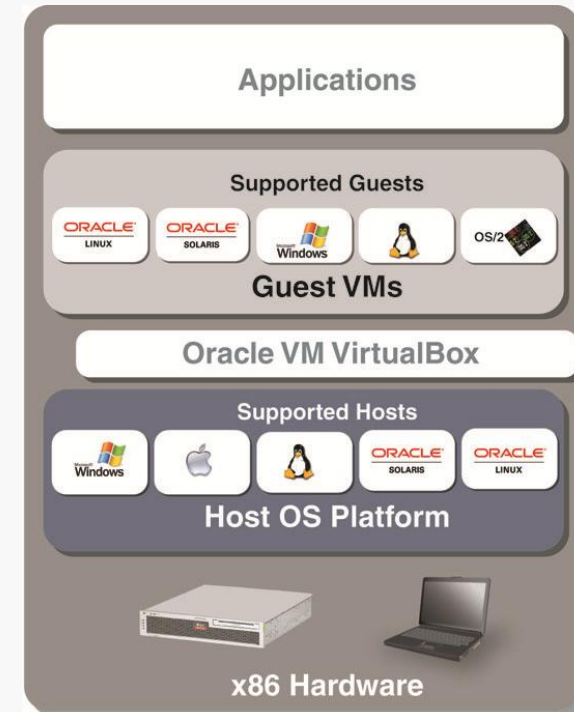
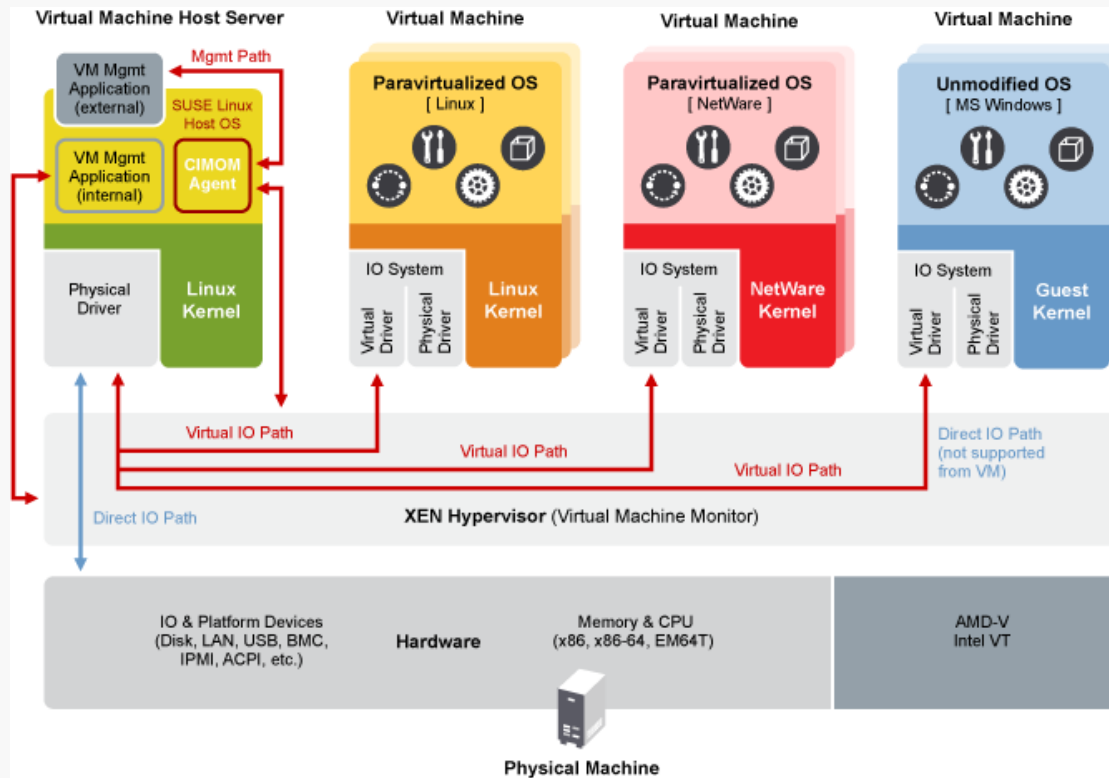
# Platform virtualizáló szoftverek

- » Nyílt forráskódú / szabad szoftver
  - » Kernel-based Virtual Machine (KVM)
    - » 2-es típus
    - » Linux kernel része
    - » a hardver támogatás virtualizációhoz követelmény
  - » QEMU
    - » általános célú emuláció és virtualizáció
    - » emuláció: más architektúrára írt OS vagy program futtatható, dinamikus fordítással
    - » virtualizáció: Xen vagy KVM
      - » ha a hoszt és vendég ugyanaz az arch.
      - » különben szoftver virtualizáció



# Platform virtualizáló szoftverek

- » Nyílt forráskódú / szabad szoftver
  - » Oracle VirtualBox
    - » 2-es típus
    - » szoftveres vagy hardveres támogatottságú virtualizáció
  - » Xen
    - » 1-es típus
    - » paravirtualizáció vagy hardveres támogatás virtualizációhoz





# Platform virtualizáló szoftverek

- » Nem nyílt / kereskedelmi termék
  - » VMware ESXi
    - » 1-es típus
    - » paravirtualizáció vagy hardveres támogatottságú virtualizáció
    - » kis méret: kb. 200 MB
    - » monolitikus VMkernel
      - » a hypervisor minden eszközmeghajtót tartalmaz és kezel
  - » Microsoft Hyper-V
    - » 1-es típus
    - » partíciók
      - » kell egy szülő partíció: x86-64 Windows Server
      - » gyerek partíciók: VM-ek
    - » paravirtualizáció vagy hardveres támogatottságú virtualizáció
    - » nagyobb méret: kb. 5GB core, vagy 10GB teljes
    - » mikorkernel
      - » eszközmeghajtók VM szinten





# **ADATKÖZPONTOK HÁLÓZATÁNAK KIALAKÍTÁSA: HÁLÓZATI ESZKÖZÖK**

# Adatközpont hálózat

- » Szerverek rack szekrénybe rendezve
  - » néhány 10 vagy 100 ezer szerver
- » Befektetési és működtetési költségek csökkentése
- » Hálózati eszközök
  - » hálózati kártyák (Network Interface Card – NIC)
  - » kapcsolók (switch, bridge)
  - » útválasztók (router)
  - » vezetékezés (réz vagy optika)
  - » új trend
    - » hálózati eszközök legyártatása egyedi specifikáció alapján és saját szoftver írása (pl. Google) ⇨ Software Defined Networking (SDN)



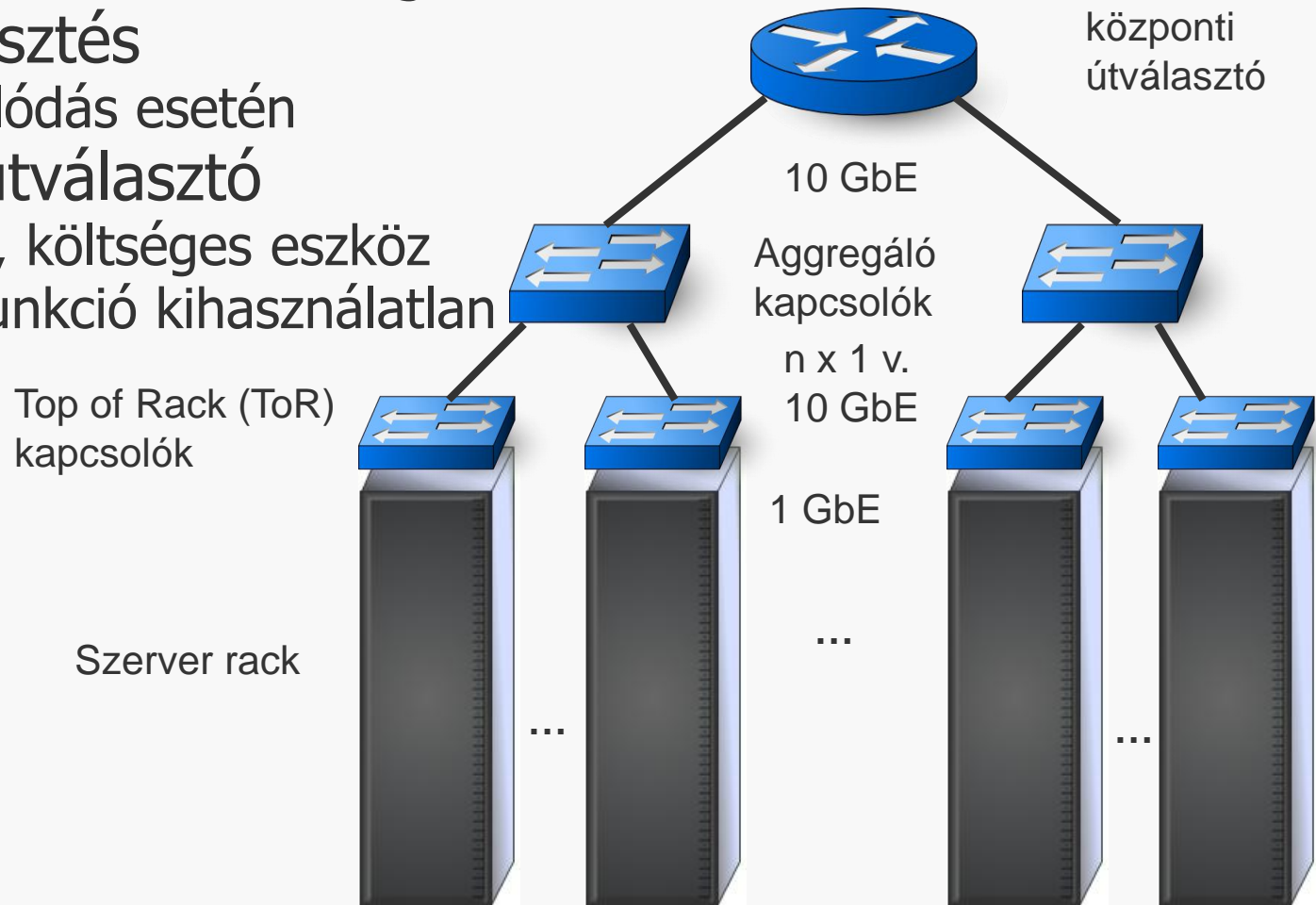


# Adatközpont hálózat

- » Miért Ethernet?
  - » általában: max. teljesítmény min. költséggel
  - » kis költségű Layer2 hálózat
  - » nagy bitsebesség: 1, 10, 40 (, 100) GbE
  - » 10 GbE
    - » 2006-tól jelentek meg eszközök
    - » 2012-től széleskörű elterjedés
  - » háttértár hálózati forgalom
    - » Fibre Channel over Ethernet (FCoE)
- » Kihívások
  - » más követelmények, mint a LAN-ban: skálázhatóság, hibatűrés, keresztmetszeti sáv szélesség, automatizált cím allokáció
- » Konvergencia
  - » nem csak az adatkommunikáció, hanem a tároló egységek (storage) forgalma is egyazon hálózaton fut
    - » adatvesztés nem tolerálható
    - » minimum bitsebesség garancia

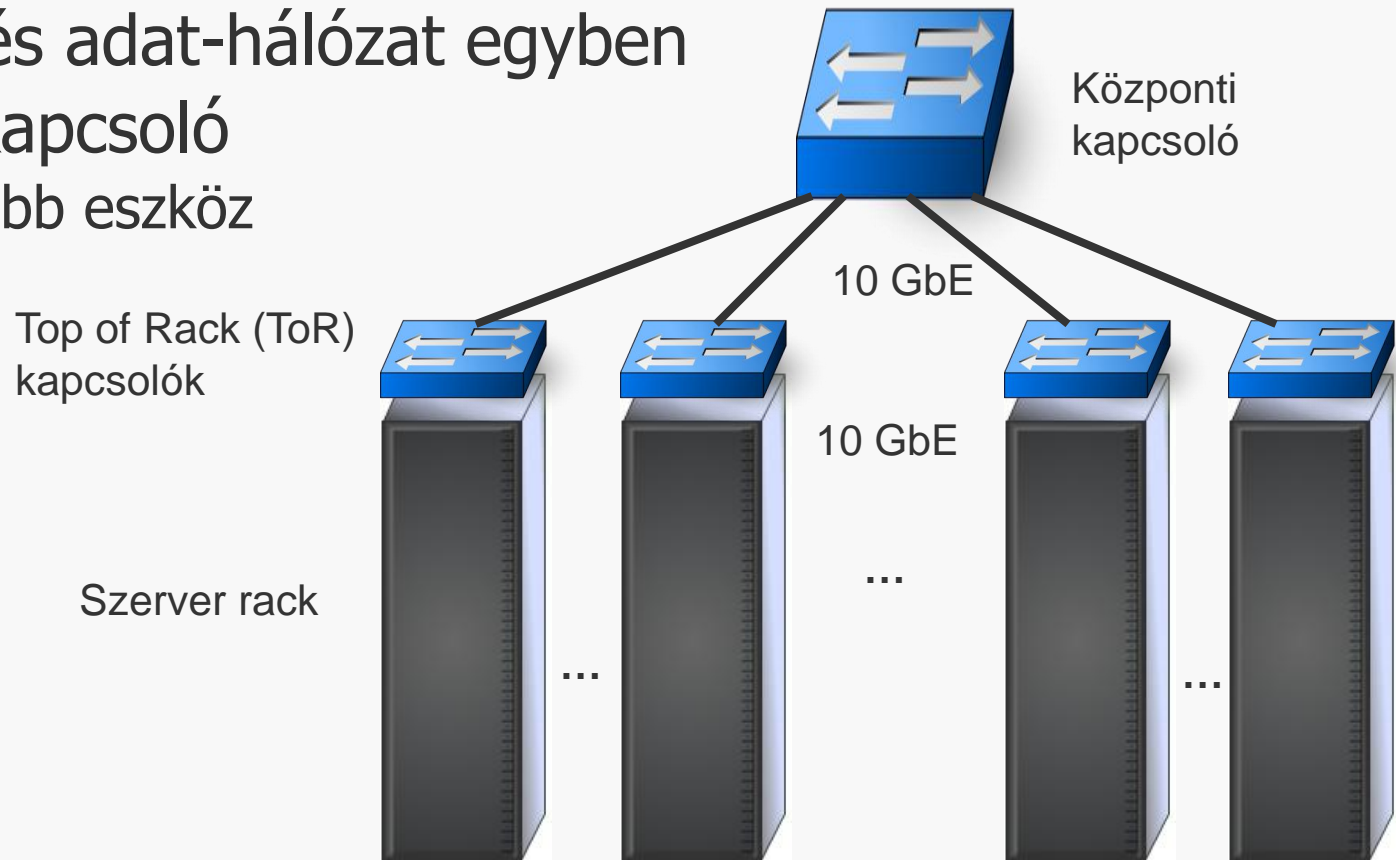
# Vállalati adatközpont

- » szerverek közötti forgalom több eszközön keresztül
  - » késleltetés, késleltetés ingadozás
- » csomagvesztés
  - » nagy torlódás esetén
- » központi útválasztó
  - » komplex, költséges eszköz
  - » sok L3 funkció kihasználatlan



# Felhő adatközpont

- » szerverek közötti forgalom kevesebb köztes eszközön halad
  - » lapos(abb) /flat/ hálózati topológia
- » háttértár és adat-hálózat egyben
- » központi kapcsoló
  - » egyszerűbb eszköz



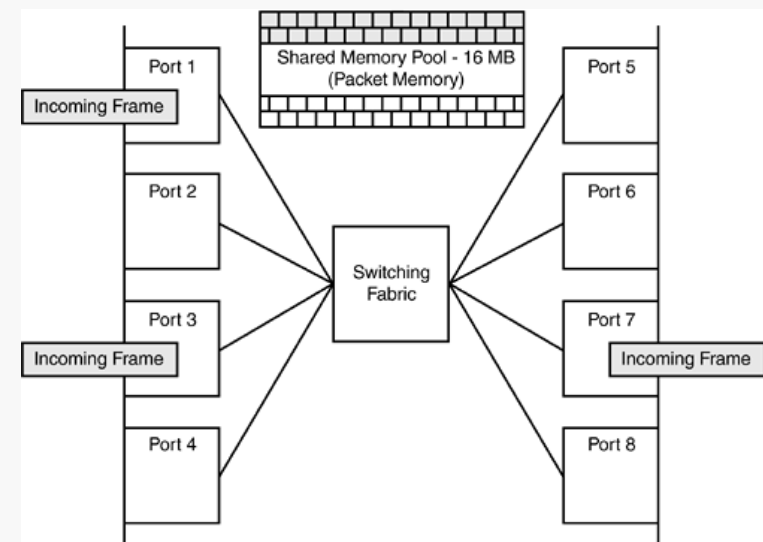
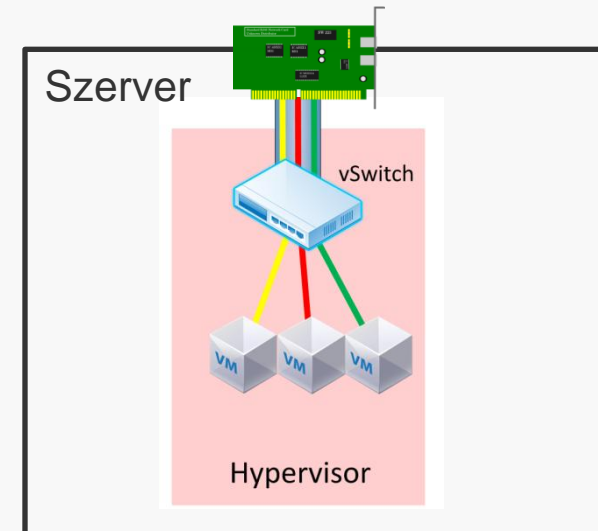


# Adatközpont kapcsoló típusok

- » Virtuális kapcsoló (virtual switch, vSwitch)
  - » VM-ek között ugyanazon fizikai szerveren
- » Top of Rack (ToR) kapcsoló
- » End of Row (EoR) kapcsoló
- » Aggregáló kapcsoló
- » Központi kapcsoló/útválasztó

# Virtuális kapcsoló

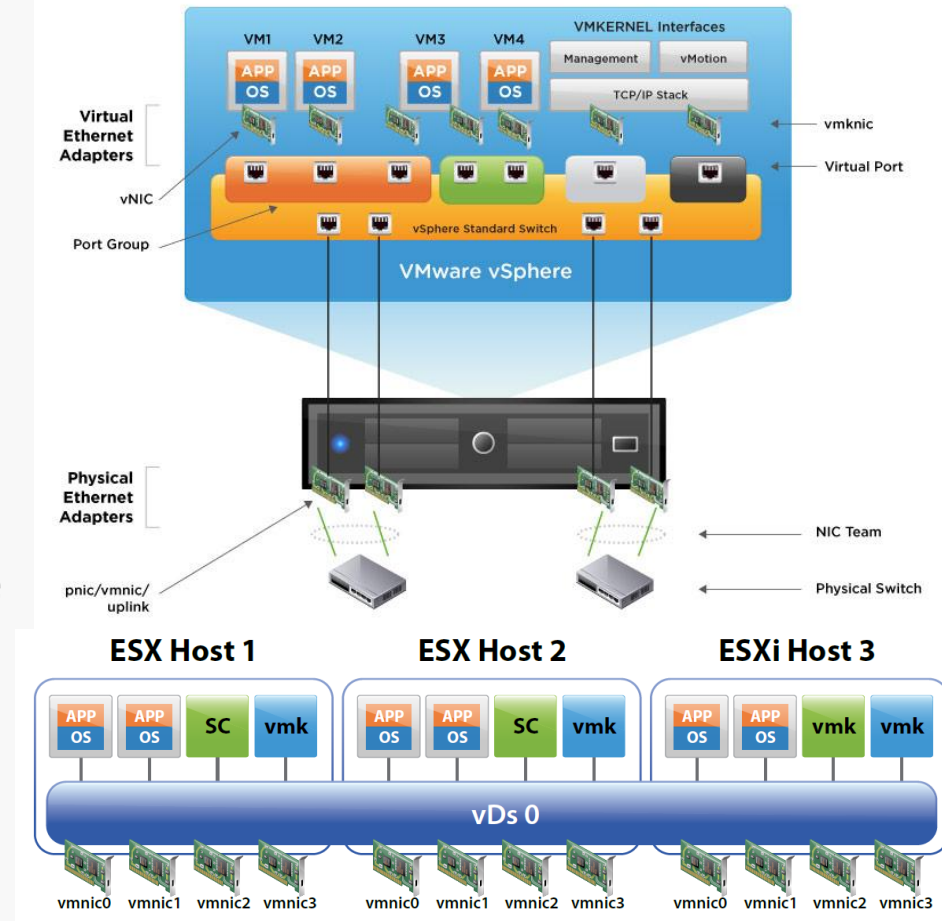
- » Hypervisor
  - » VM-ek és virtuális kapcsolók konfigurálása
- » vSwitch
  - » a fizikai hálózati kártyához kapcsolódik, jellemzően egyet használ az összes VM
    - » a serveren lévő VM-ek ki- és bemenő forgalmának bitsebességét korlátozza
  - » a server CPU végzi a műveleteket
  - » gyakorlatilag osztott memóriával megvalósított kapcsoló
    - » a server memóriájában vannak az adatok (csomagok, keretek)
    - » VM-ek a mutatókat cserélik egymás között
    - » nagy bitsebesség!
  - » ugyanúgy a hálózat része
    - » egységes menedzselés, konfigurálás lenne az ideális





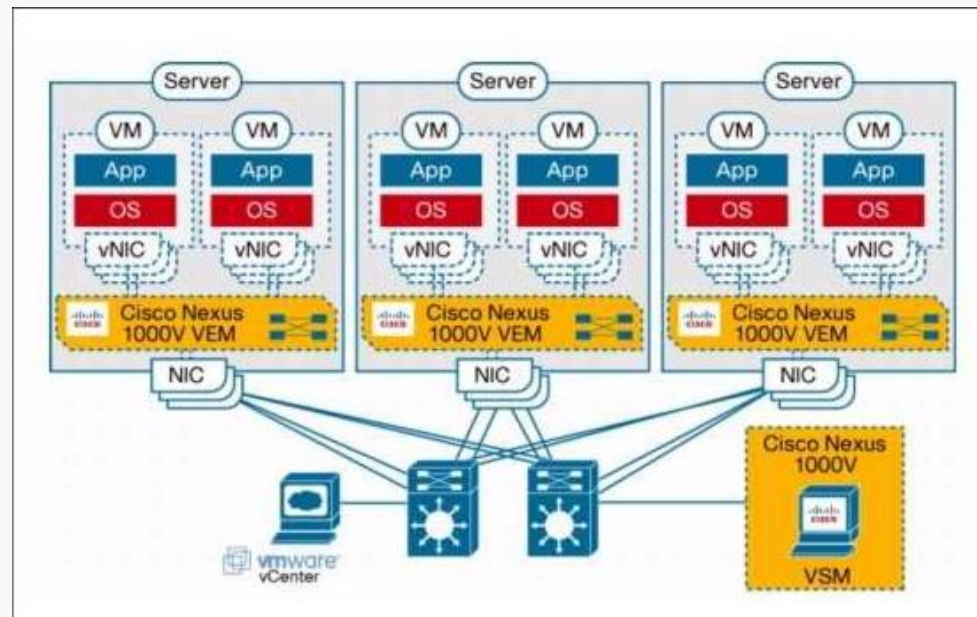
# Virtuális kapcsolás: VMware vSwitch

- » több VM – egy (vagy több) hálózati kártya
- » szoftveres vSwitch – hypervisor
  - » VMware vSphere (ESXi)
    - » vNetwork Standard Switch (VSS)
    - » absztrakt, elosztott kapcsoló: vSphere Distributed Switch (VDS)
      - » több fizikai szervert fog össze
    - » Új képességek (jelentős CPU használat)
      - » forgalom monitorozás
      - » VLAN izoláció
      - » forgalom formázás (max. sávszélesség)
      - » vMotion támogatás



# Virtuális kapcsolás: VMware vSwitch

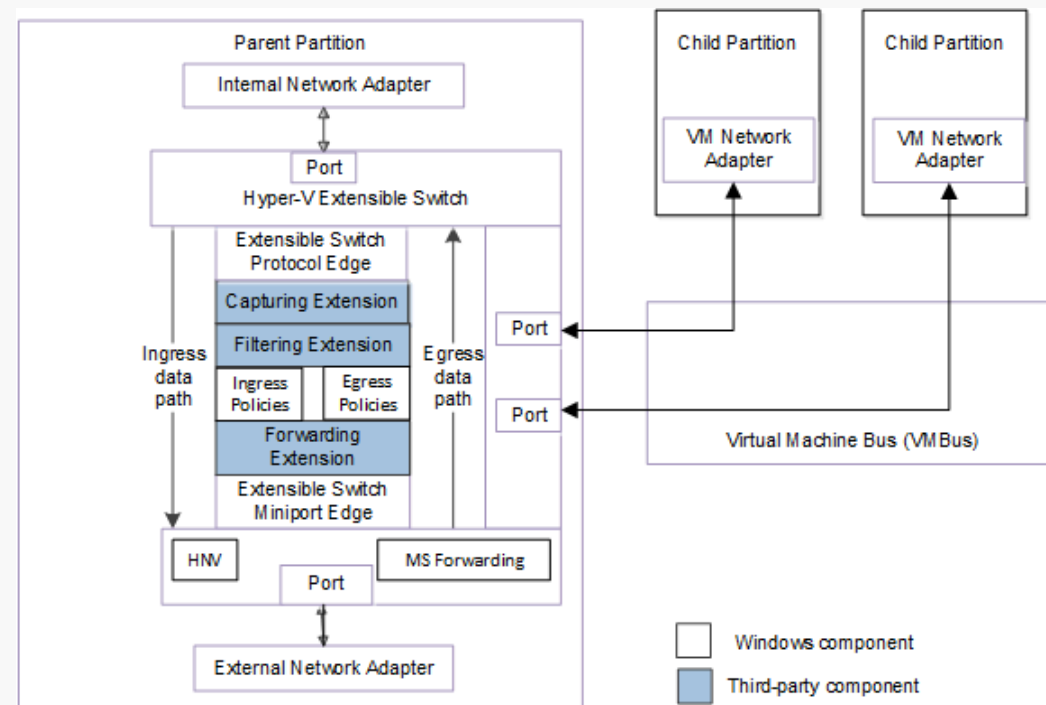
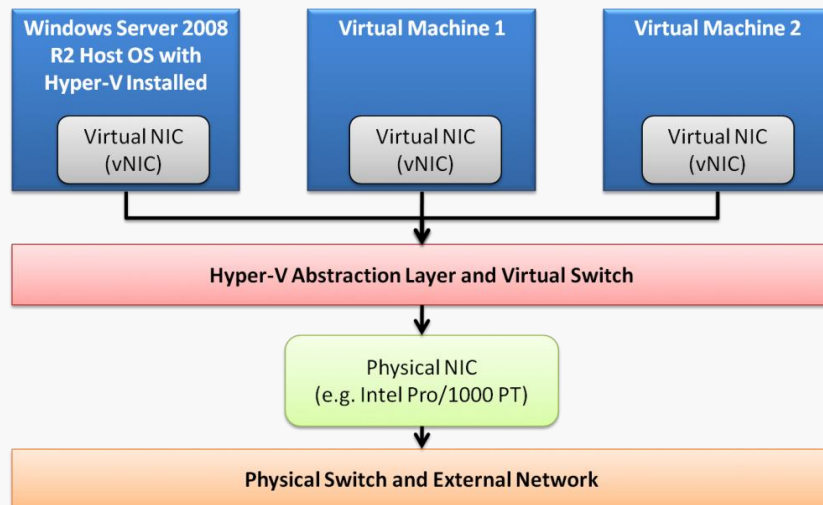
- » VMware vSphere (ESXi)
  - » Cisco Nexus 1000V
    - » Cisco/VMware kollaboráció
    - » közelít egy fizikai kapcsoló képességeihez, de korlátozásokkal
    - » Komponensek
      - » Virtual Ethernet Module
      - » Virtual Supervisor Module
    - » integrálva Cisco Command Line Interface (CLI) és VDS API
    - » VXLAN támogatás



# Virtuális kapcsolás: Microsoft Hyper-V

- » Private/Internal/Public mód
- » Hyper-V 3.0: Windows Server 2012
  - » Hyper-V Extensible Switch
  - » forgalom osztályozás és szűrés és monitorozás
  - » minimum sávszélesség garancia és maximum limit
  - » torlódás vezérlés
  - » VM-enként várakozási sorok
  - » live migration
  - » egyedi kiterjesztések fejleszthetők
  - » Cisco Nexus 1000V integrálható

## Hyper-V Networking Basic Diagram

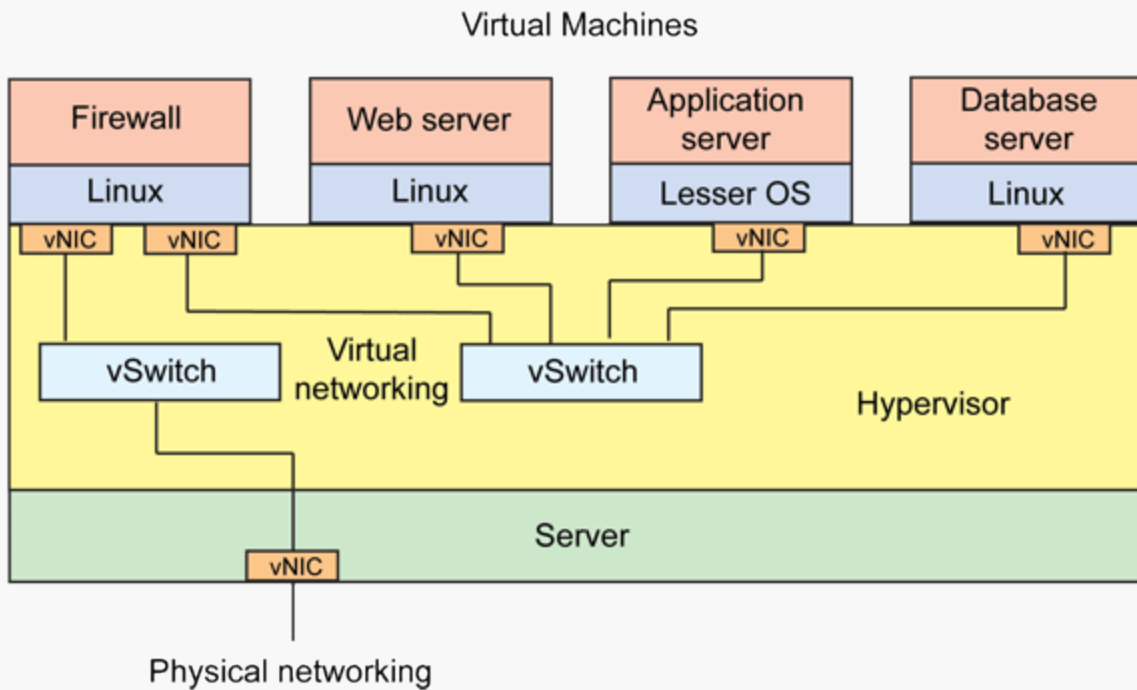




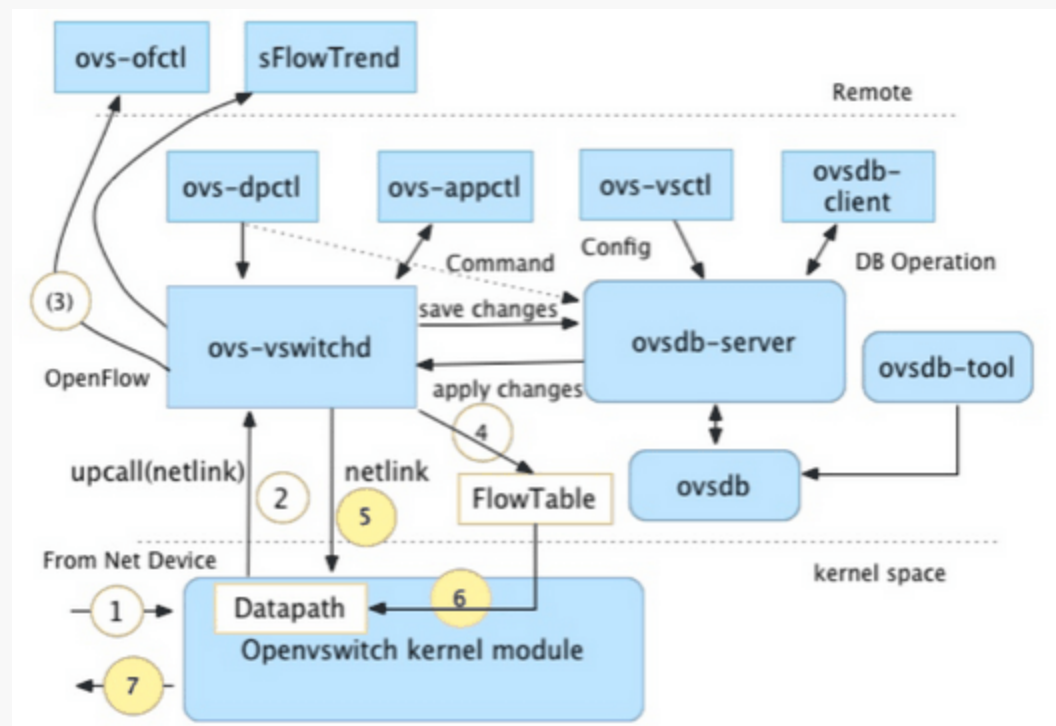
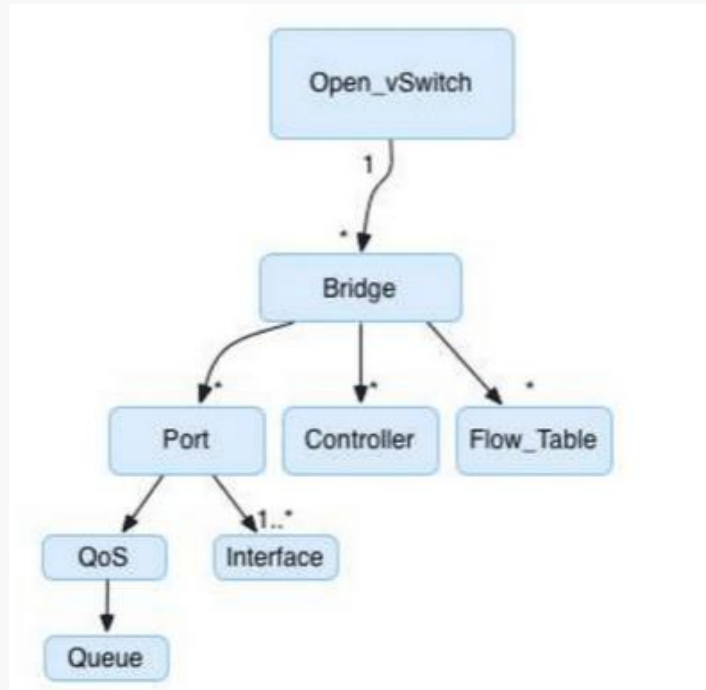
# Virtuális kapcsolás: Open vSwitch

- » nyílt forráskódú implementáció
- » szabványos menedzsment protokollok
- » Képességei
  - » VM-ek közötti forgalom monitorozása, pl. NetFlow segítségével
  - » 802.1Q VLAN
  - » STP (IEEE 802.1D-1998)
  - » QoS kezelés
  - » VM interfészenkénti forgalom szabályozás
  - » hálózati kártyák összefogása
  - » OpenFlow protokoll támogatás
  - » IPv6
  - » alagút prtokollok: GRE, VXLAN, IPsec
  - » Kernel vagy user-space csomagtovábbítás
  - » user-space vezérlés
- » Jellemzői
  - » hálózati állapot mobilitás, pl. bejegyzések a továbbítási táblákba, hozzáférési és QoS szabályok
  - » hálózati dinamika kezelése: VM-ek indítása, leállítása, mozgatása
  - » címkék kezelése: alagutazáshoz, VM azonosításhoz
  - » hardver integráció lehetősége: vezérlő rétege lehet fizikai kapcsolónak
  - » több fizikai szerveren elosztott működés: OpenFlow segítségével
- » Platformok
  - » XenServer, Xen, KVM, VirtualBox, OpenStack, OpenNebula, Linux (kernel), FreeBSD

# Virtuális kapcsolás: Open vSwitch



# Open vSwitch adatszerkezet és felépítés



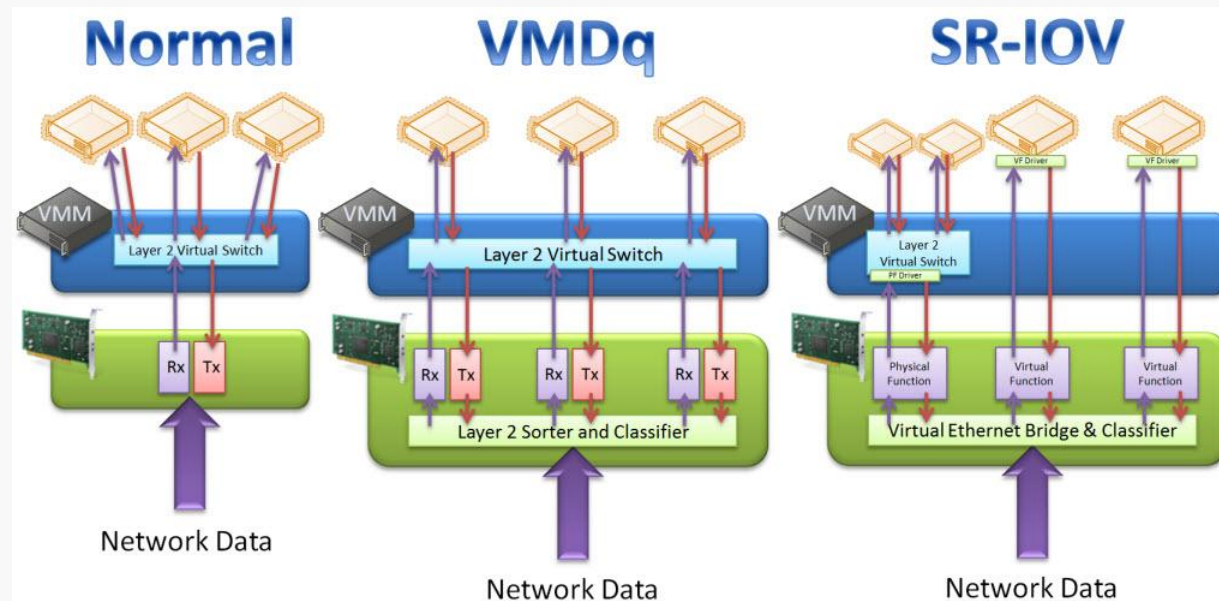


# Linux Bridge

- » egyszerűbb az OVS-nél
  - » kevesebb funkció
  - » nem folyam alapú modell
  - » kisebb kódbázis: kevesebb hibalehetőség
  - » könnyebb a hibakeresés
  - » beállítás: brctl, ip route
- » tunneling
  - » korábban csak GRE volt támogatott
  - » VXLAN támogatás: Linux kernel 3.7, 2012.

# VM hálózati teljesítmény növelése

- » Virtual machine device queues (VMDq) – Intel
  - » vSwitch intenzív CPU használata csökkenti a VM-ek teljesítményét
  - » VMDq a hálózati kártyán implementálva
  - » VM-enkénti küldési és fogadási sorok
    - » MAC és VLAN címke alapján
  - » előnyök
    - » párhuzamosítás
    - » hardver kezeli a csomagok leválogatását
- » PCIe single-root IO virtualization (SRIOV)
  - » külön virtuális funkció (VF)
  - » kikerüli a vSwitch-et, ezért csak speciális esetekben alkalmazzák
  - » DMA VF és VM között
  - » támogatás
    - » VM hálózati meghajtó
    - » hypervisor
  - » kb. 10-15% CPU terhelés csökkenés





# ToR kapcsoló

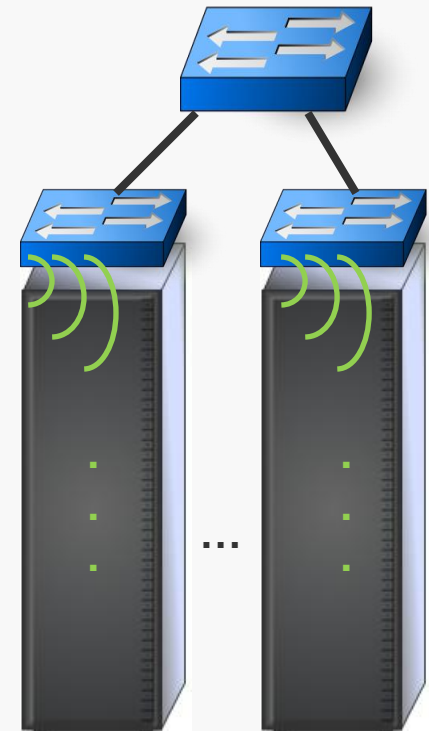
- » Jellemző konfiguráció
  - » szerverek csillag topológiában
  - » 48 x 10 GbE port a szerverek felé
  - » 4 x 40 GbE az aggregáló kapcsoló(k) felé
  - » 480 Gbps  $\leftrightarrow$  160 Gbps
    - » 3:1 oversubscription: túljegyzés, túlfoglalás
      - » tipikus értékek: 2,5:1 – 8:1
  - » 1 rack – 1 kapcsoló
    - » rack szintű redundancia
- » ToR kapcsoló
  - » kis késleltetés
  - » nagyméretű címtér kezelés
  - » háttértár forgalmat is kezelje
- » Lehetséges extra funkciók
  - » alagutazás
  - » szűrés
  - » mérés
  - » terhelés kiegyenlítés

## Előnyök:

- rövidebb és egyszerűbb vezetékvezés
- rack szintű menedzsment/redundancia

## Hátrányok:

- több kapcsolót kell menedzselni
- skálázhatósági korlátok (STP, portszám)
- kapcsolónkénti vezérlősík



# EoR kapcsoló

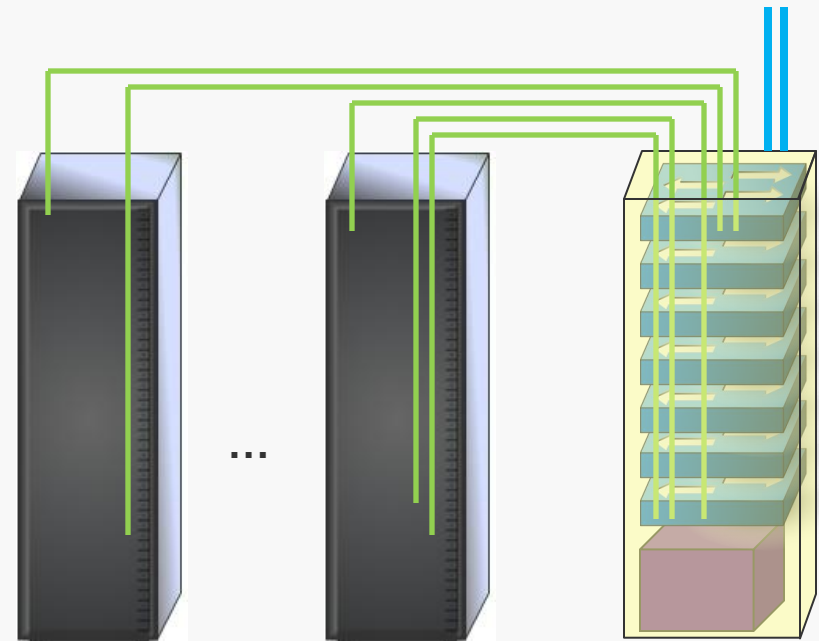
- » **Költségcsökkentés**
  - » több kapcsoló komponens, kapcsoló kártyák
    - » szerverek felé
    - » központi kapcsoló felé
  - » közös áramforrás, hűtés, menedzsment
  - » kártyák közös buszon
- » **Vezetékezés**
  - » max. 100 m a szerverekig
  - » ellenben ez költséges lehet

## Előnyök:

- kevesebb kapcsolót kell menedzselni
- kevesebb aggregációs portot igényel
- kevesebb STP példány
- egy vezérlősík

## Hátrányok:

- költséges, rugalmatlan vezetékezés
- hosszú, nehezen menedzselhető vezetékek
- soronkénti menedzsment/redundancia



# Kapcsolómező kiterjesztés

- » Fabric extender (FEX)
- » ToR / EoR hibrid megoldás
  - » logikailag elosztott EoR kapcsolóként viselkedik
  - » minden rack tetején kiterjesztett kapcsolómező
    - » fizikai kapcsoló korlátozott képességekkel
  - » egy vezérlő sík az összes porthoz
  - » vezetékezés, mint ToR kapcsolóknál

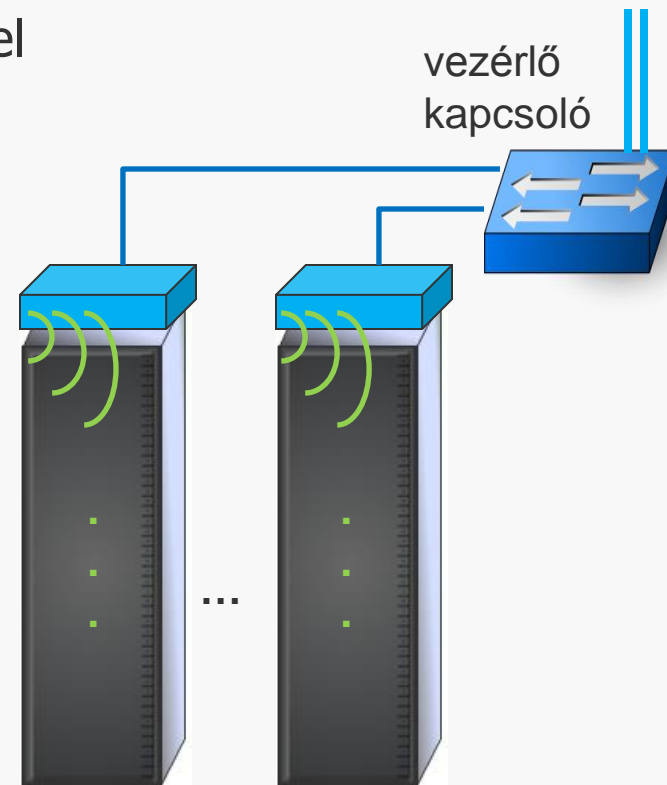
## Előnyök:

- kevesebb kapcsolót kell menedzselni
- kevesebb aggregációs portot igényel
- kevesebb STP példány
- egy vezérlősík
- költséghatékony vezetékezés
- rack szintű menedzsment/redundancia

## Hátrányok:

- kevésbé elterjedt

kapcsolómező  
kiterjesztés





# Aggregáló és központi kapcsoló

- » aggregáló: ToR és központi között
- » központi
  - » kapcsolódás a külső hálózathoz
  - » nagyszámú, nagysebességű port
  - » moduláris felépítés
    - » vonali kártyák (első fokozat)
    - » kapcsoló kártyák (második fokozat)
    - » feldolgozó kártyák (CPU, memória) – tűzfal, terheléselosztó
    - » menedzsment kártyák
  - » cél: egyszerű csomagtovábbítási szabályok címkék segítségével