



# **Felhő alapú hálózatok (VITMMA02)**

## **Hálózat virtualizálás: Overlay hálózatok**

### **OpenStack Neutron Networking**

Dr. Maliosz Markosz

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

2018. tavasz



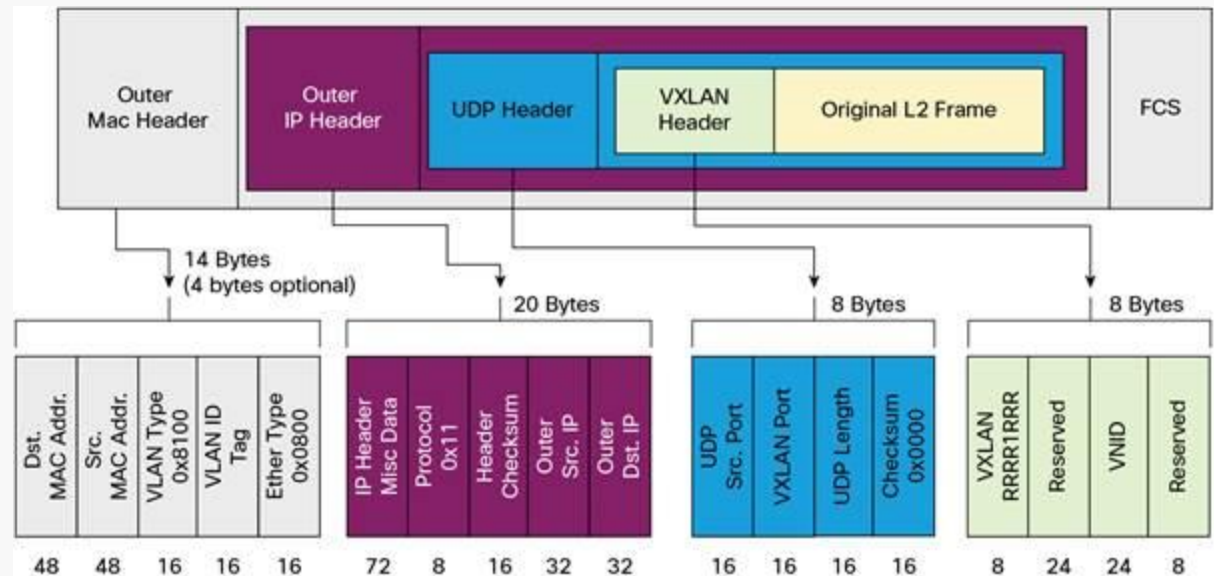
# OVERLAY HÁLÓZATOK

# Hálózat virtualizáció

- » Ügyfél szintű szeparáció támogatás
  - » 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

# VXLAN

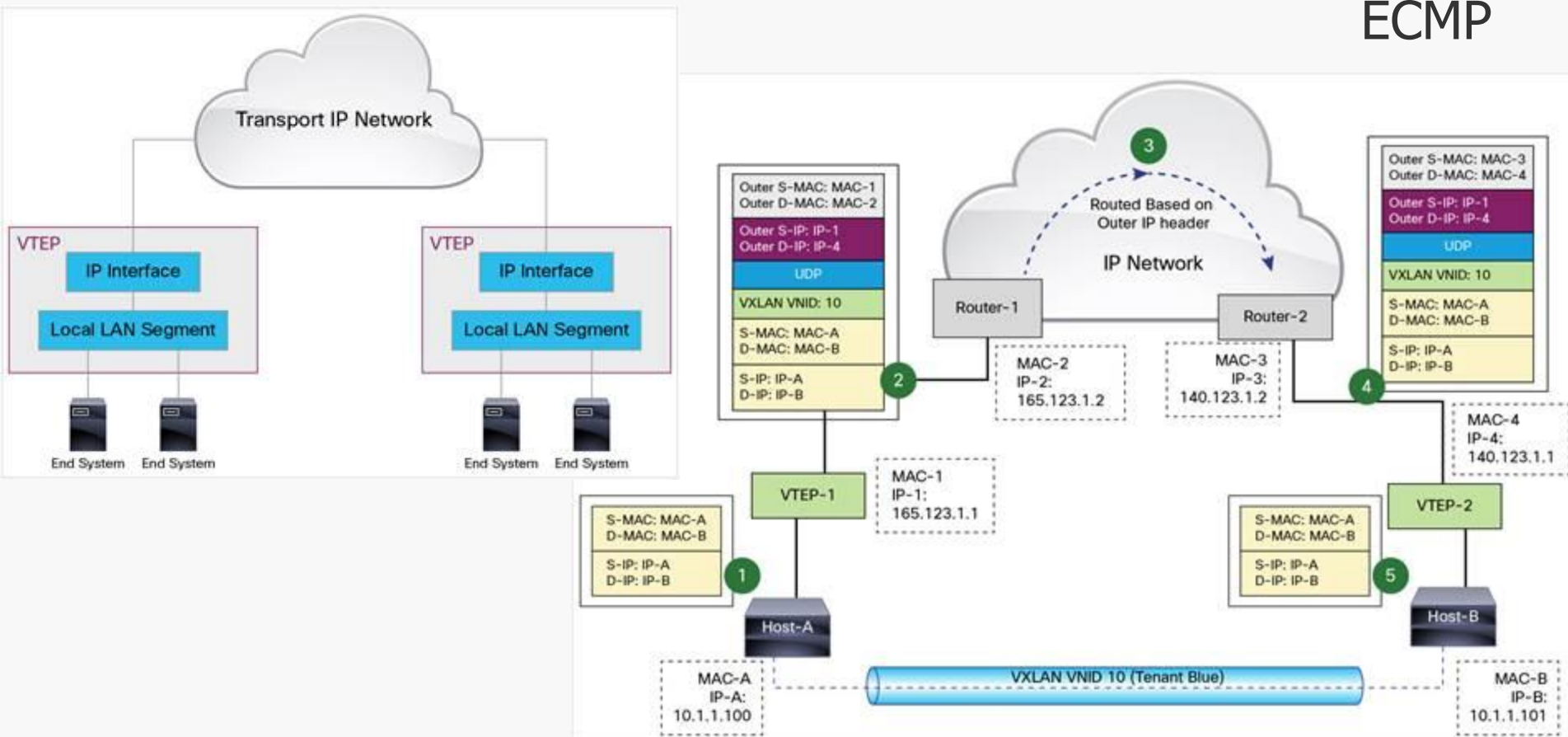
- » ügyfél eredeti L2 kerete
  - » eredeti MAC címmel és VLAN címkével
- » MAC-in-UDP
- » VXLAN és UDP fejléc
  - » VXLAN network ID (VNID) – ez azonosítja az ügyfelet
    - » 24 bit  $\Rightarrow$  16 millió ügyfél
- » fizikai hálózat: IP útvonalválasztás (Layer3)



# VXLAN

- » VXLAN Tunnel End Point (VTEP)
- » MAC-to-VTEP táblák tanulás útján (IP multicast)
  - » egy VNI összes VTEP-je egy multicast csoportban

ECMP



# NVGRE

- » hasonló a VXLAN-hoz
- » alapja: Generic Routing Encapsulation (GRE)
  - » általános fejléc
  - » sok különböző protokollra
  - » pont-pont kapcsolat
- » NVGRE
  - » GRE fejléc
    - »
 

```

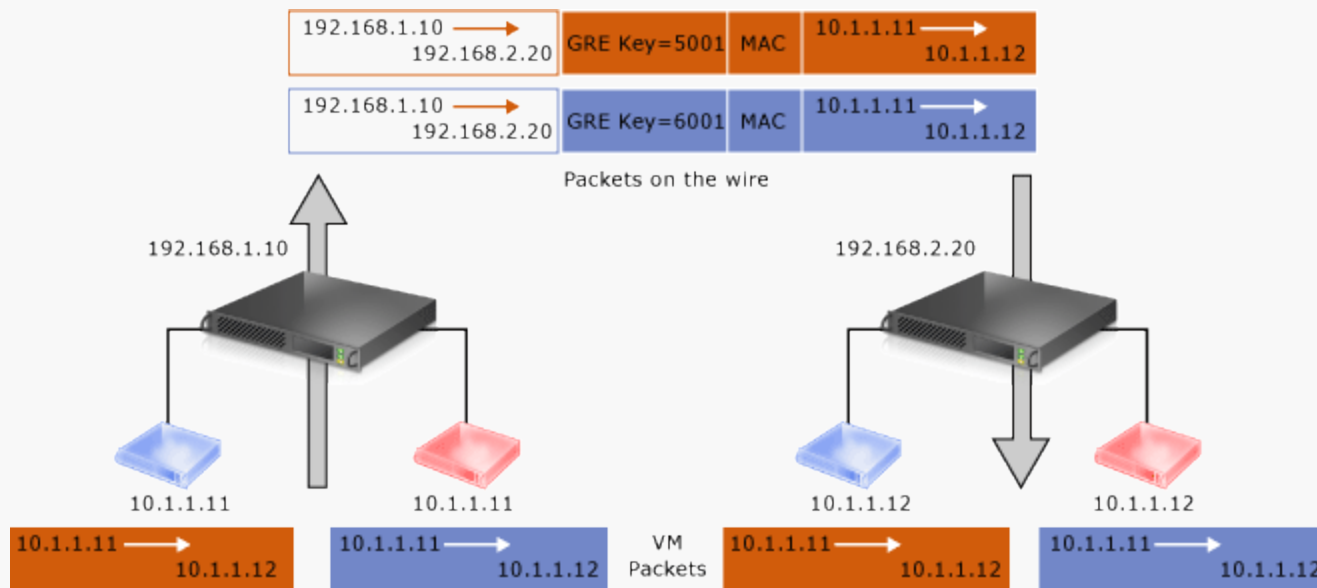
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|0| |1|0|   Reserved0   | Ver |   Protocol Type 0x6558   |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Virtual Subnet ID (VSID)           | FlowID |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
              
```

      - » Virtual Subnet Identifier (VSID) 24 bit  $\Rightarrow$  16 millió ügyfél
      - » FlowID: opcionális, egyedi folyamazonosító
        - » ECMP hash számításához
  - » belül nincs VLAN címke (vagy levételre kerül)
    - » VSID-be kódolják



# NVGRE

- » Network Virtual Endpoint (NVE)
  - » VSID és DMAC alapján a címzethez kapcsolódó NVE IP címére küldés
- » az Internet draft nem specifikálja
  - » a cím információk terjesztését
  - » VLAN információ helyreállítását





# Generic Network Virtualization Encapsulation

- » MAC-in-UDP over IPv4/IPv6
- » univerzális, kiterjeszthető megoldási javaslat
- » csak a beágyazási formátumot definiálja
- » opcionális mezők
  - » nem fix mezőhosszak, rugalmasság

## » Geneve fejléc:

```

+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|Ver|  Opt Len  |O|C|      Rsvd.  |              Protocol Type              |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|              Virtual Network Identifier (VNI)              |      Reserved      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|              Variable Length Options              |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```



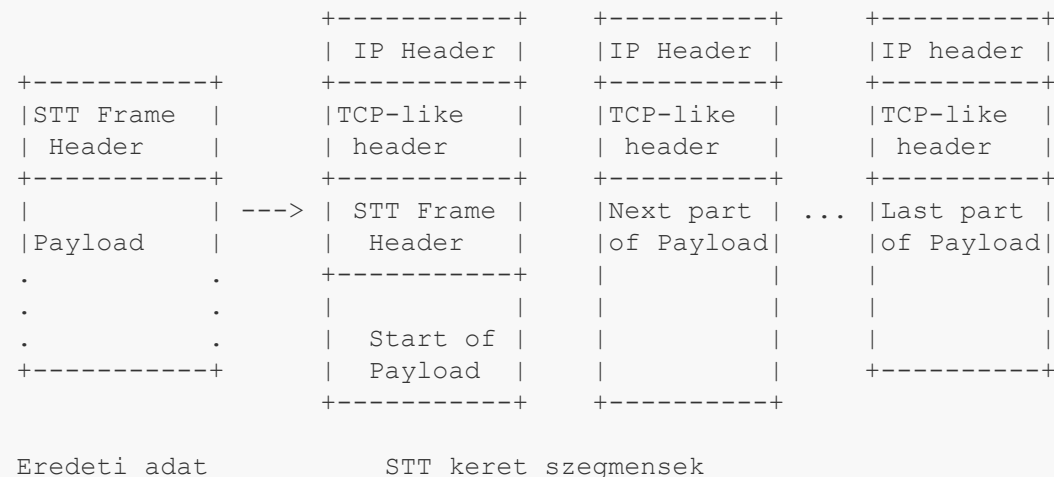
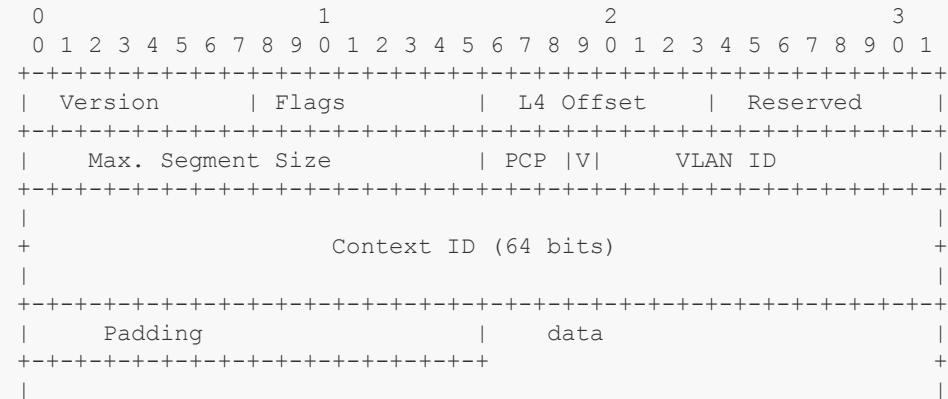
# Alagút végződés helye

- » hypervisor/vSwitch-ben
  - » ez az általánosan használt megoldás
  - » legközelebb a VM-ekhez, könnyű az ügyfél azonosítás
  - » CPU erőforrás
  - » TCP segmentation offload (TSO), checksum offload támogatás kérdéses
- » fizikai hálózati kártyán
  - » offload támogatás a tunnel protokoll fejlécre is
  - » kevés NIC támogatja
- » fizikai kapcsolón
  - » forrás VM nem ismert
  - » VNID/VSID meghatározásához kell a belső MAC cím



# Stateless Transport Tunneling (STT)

- » elsődlegesen vSwitch-ek közötti kommunikációra
- » komplexebb, mint az előzőek
- » max. 64 kbyte-os Ethernet keretet kezel
  - » maximum transmission unit (MTU)
  - » TCP segmentation offload kihasználása a hálózati kártyán
- » STT fejléc
  - » 64 bites Context ID mező
- » a feldarabolt adatok elé TCP-szerű/IP/Eth fejléc
  - » ez alapján állítja össze a nagyméretű keretet





# Összehasonlítás

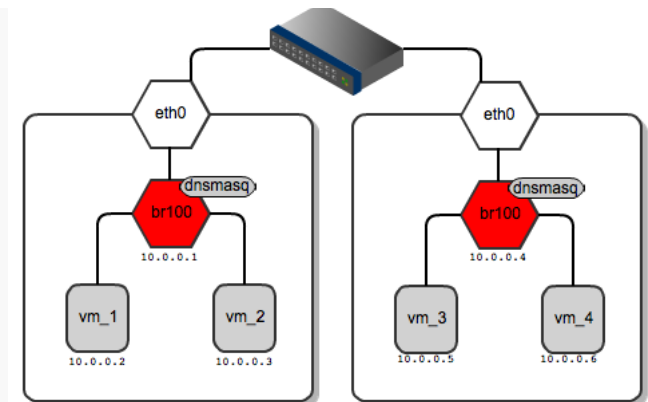
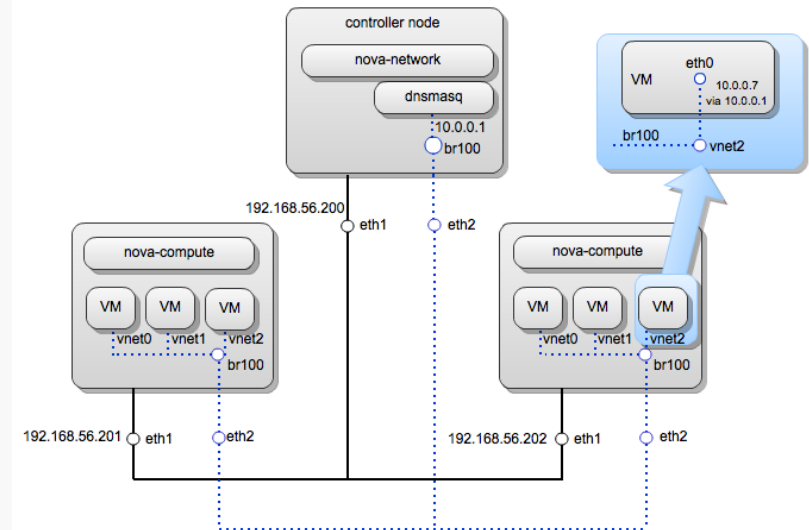
|                                                      | VXLAN           | NVGRE                   | STT                                              |
|------------------------------------------------------|-----------------|-------------------------|--------------------------------------------------|
| Plusz bájtok                                         | 50 (VLAN: +4)   | 42 (VLAN: +4)           | Első szegmens: 76<br>Továbbiak: 58<br>(VLAN: +4) |
| Protokoll                                            | UDP             | GRE                     | TCP                                              |
| Ügyfél<br>megkülönböztetés                           | 24 bit VNID     | 24 bit VSID             | 64 bit Context ID                                |
| ECMP-hez<br>megkülönböztetés<br>(belső⇒külső folyam) | Forrás UDP port | VSID +<br>FlowID (8bit) | Forrás TCP port                                  |



# OPENSTACK NEUTRON

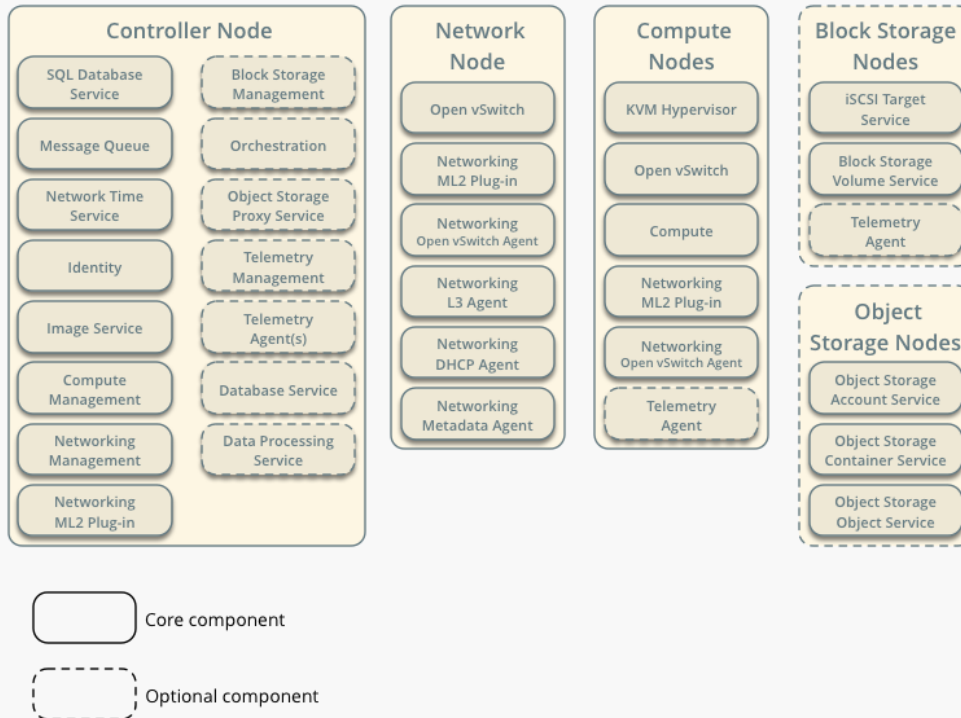
# OpenStack hálózati architektúra

- » „Networking in OpenStack is a complex, multifaceted challenge.” /OpenStack Operations Guide/
- » Network as a Service
- » feladatok
  - » IP címek kezelése
    - » statikus, DHCP
    - » floating IP
  - » virtuális hálózatok kezelése
    - » flat, VLAN
  - » önkiszolgáló módon
- » többféle megoldás
  - » Nova networking / Neutron
  - » single-host / multi-host
- » Neutron
  - » plug-in szemlélet
  - » SDN/OpenFlow

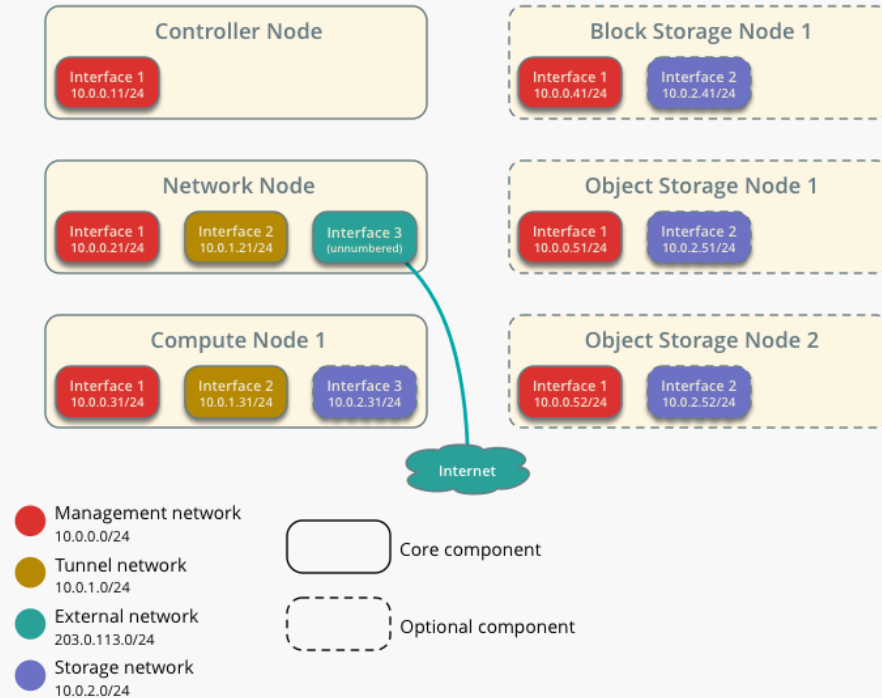


# Neutron network

Minimal Architecture Example - Service Layout  
OpenStack Networking (neutron)



Minimal Architecture Example - Network Layout  
OpenStack Networking (neutron)





# Nova és Neutron Network

## » Nova

- » alapfunkciók
  - » network address translation (NAT), DHCP, DNS
- » L2 hálózat
- » korlátozott skálázhatóság
  - » VLAN, DNS&DHCP (dnsmasq)

## » Neutron

- » hálózat absztrakció
- » L2/L3 hálózat, önkiszolgáló módon, szabályokkal
  - » pl. több szegmensből álló hálózat egy web alkalmazás számára
- » Load Balancing, Virtual IP, VPN, tűzfal
- » overlay VLAN tunneling
- » Distributed Virtual Router (Juno)



# Neutron hálózat absztrakció

- » Külső (external) /ez fizikai/ hálózathoz illesztés, pl. Internet
- » Belső hálózatok a VM-ek összekötésére
  - » virtuális: hálózat, alhálózat, útvonalválasztó
  - » VM-hez hozzárendelhető külső floating IP cím, hogy elérhető legyen
- » Security groups
  - » tűzfal szabályok
  - » VM-hez rendelt
- » Open vSwitch
  - » core plugin
  - » br-int (integration bridge)
    - » VM-ekhez kapcsolódik
  - » br-ex
    - » külső hálózathoz kapcsolódik

# Neutron komponensek

## » szerver + plugin + agent struktúra

### » neutron-server

- » controller node-on fut
- » API kérések kezelése
- » hálózati modell és portokhoz rendelt IP címek beállítása

### » plugin – kiterjesztés: neutron-\* -plugin

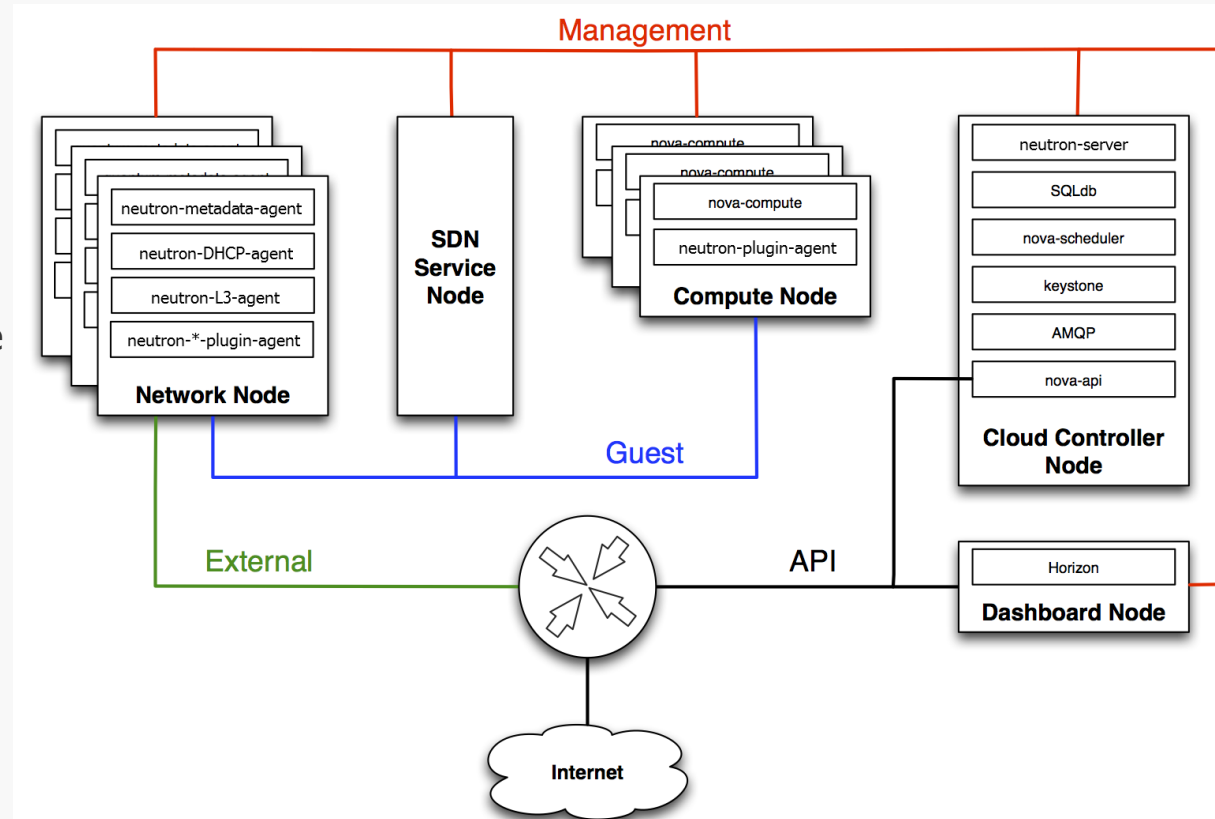
- » network node-on fut
- » agent-ek menedzselése

### » plugin-agent: neutron-\* -agent

- » compute node-on fut
- » menedzseli a lokális virtuális kapcsolót

### » általános agent-ek

- » DHCP: neutron-dhcp-agent
- » L3 agent: neutron-l3-agent
  - » L3/NAT funkció a külső hálózat felé
  - » megvalósítás: Linux IP stack és iptables





# Modular Layer 2 (ML2) plugin

- » Különböző L2 hálózati technológiákat kezel egységesen
- » Együttműködik az openvswitch, linuxbridge, és Hyper-V L2 agent-ekkel
- » Hálózat típusonkénti meghajtók (type drivers)
  - » Flat
  - » Local (DevStack single box)
  - » VLAN
  - » GRE
  - » VXLAN



# Hálózati névterek

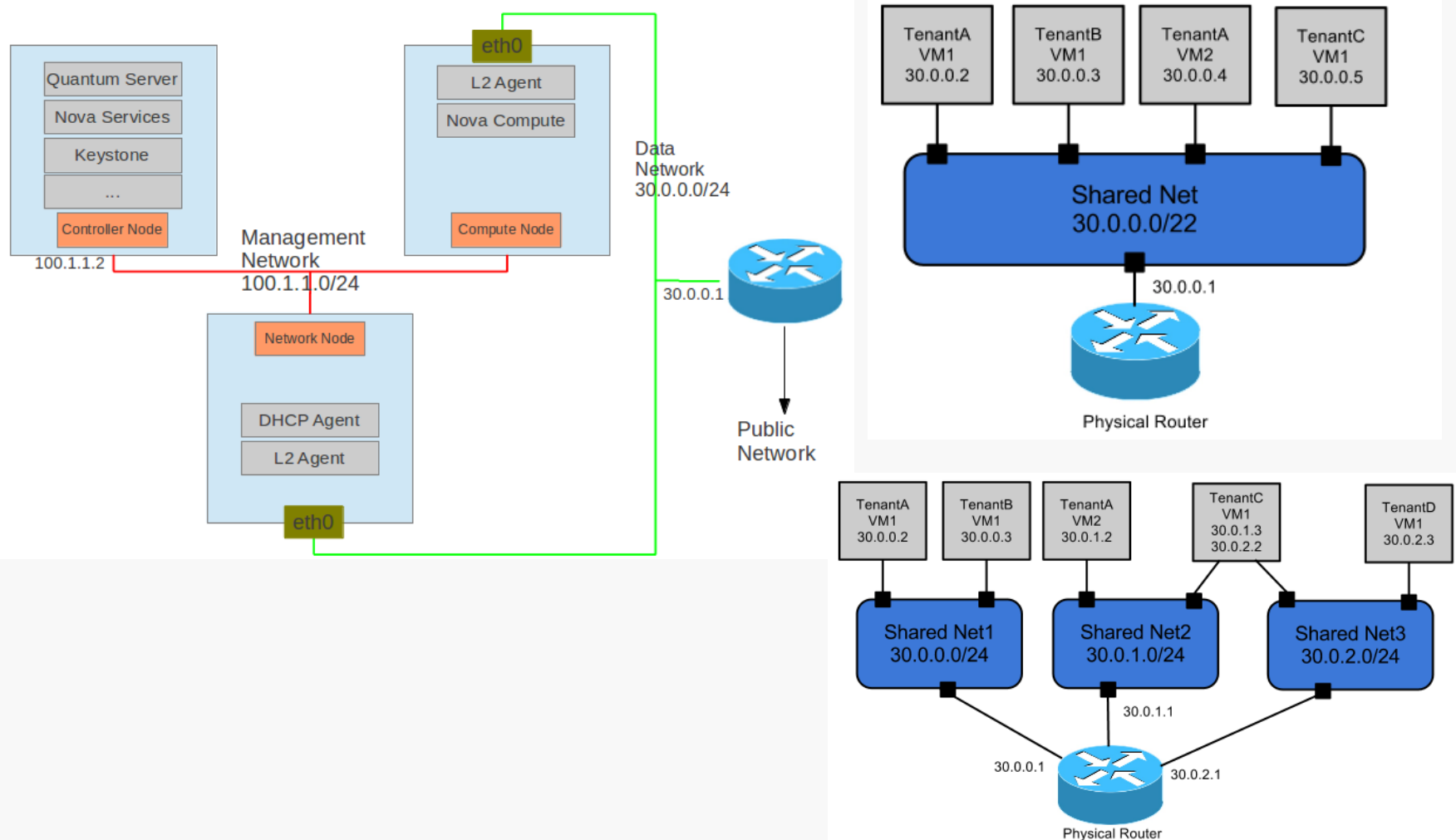
## » Network namespaces

- » kernel szintű megoldás, nem csak hálózatokra
  - » fájlrendszer, folyamat, felhasználó, stb.
- » izolált Layer2 hálózatok, átlapolódó IP címekkel
- » virtuális interfészek, útválasztók szeparálása
- » pl. dhcp-agent és l3-agent külön névtérben fut

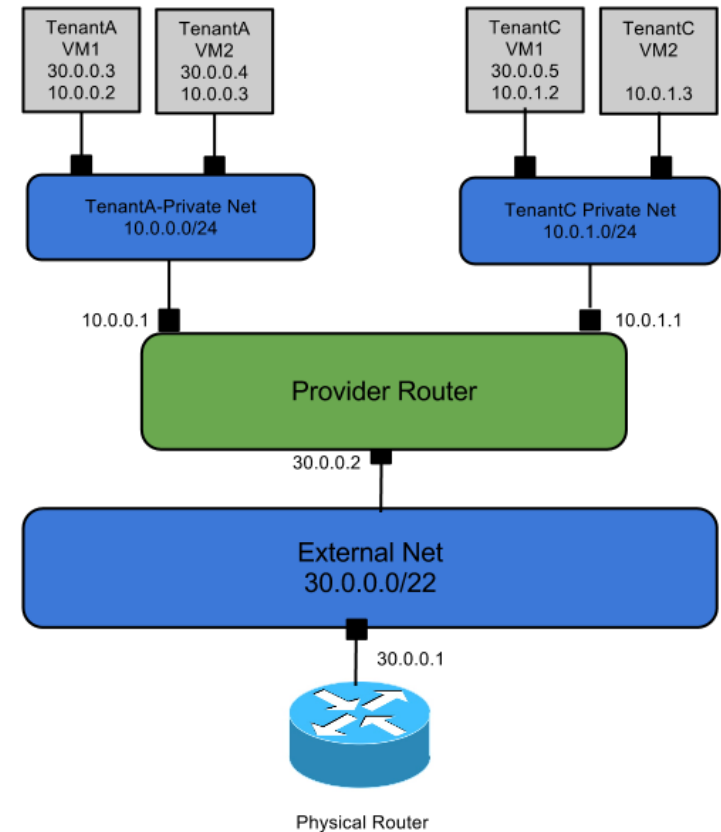
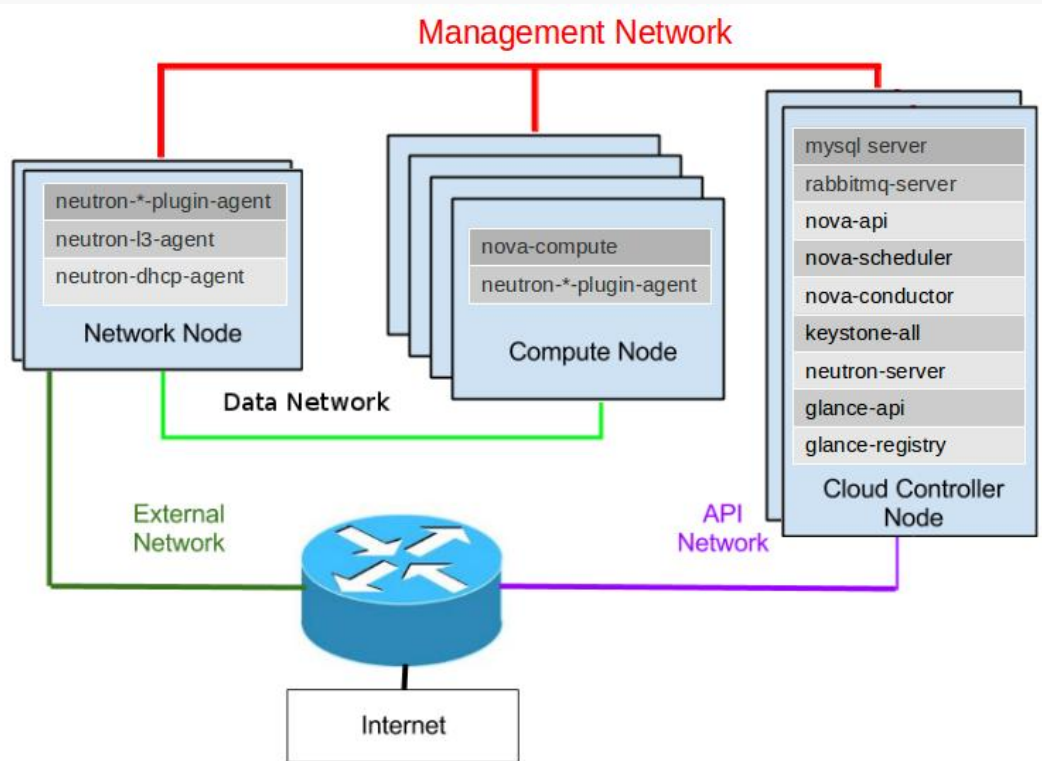
## » Gyakorlatban

- » `ip netns`
  - » kilistázza a névtereket
- » `ip netns exec <névtér> <névtérre vonatkozó parancs>`
  - » pl. `ip netns exec qdhcp-e521f9d0-a1bd-4ff4-bc81-78a60dd88fe5 ip a`

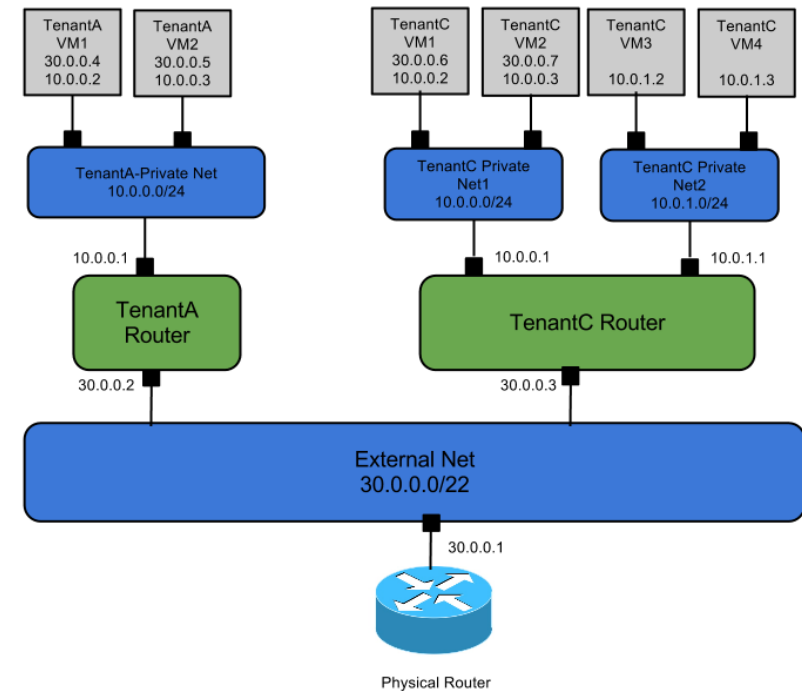
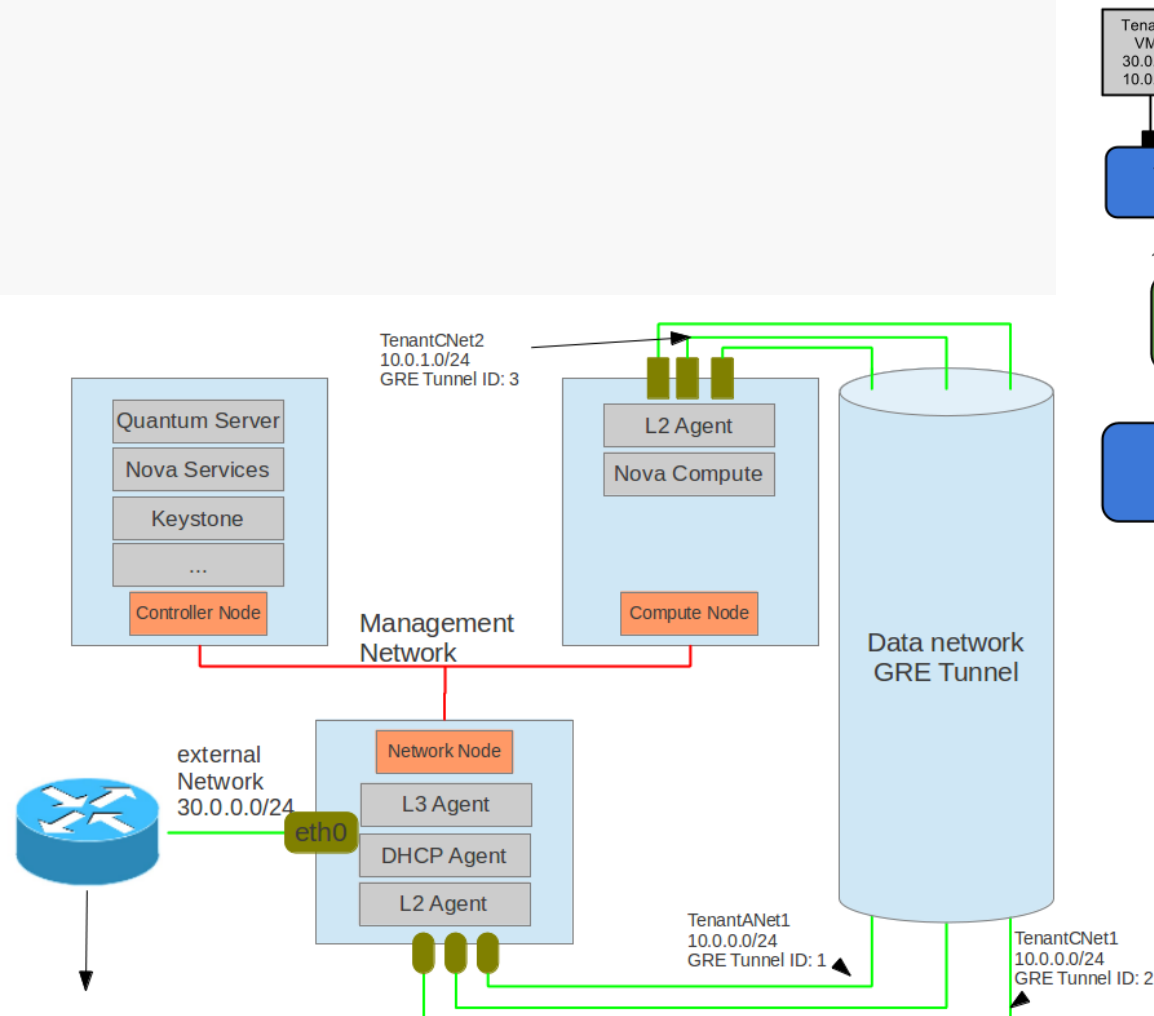
# Neutron: single/multiple flat hálózat



# Neutron: szolgáltatói útválasztóval



# Neutron: ügyfél útválasztókkal

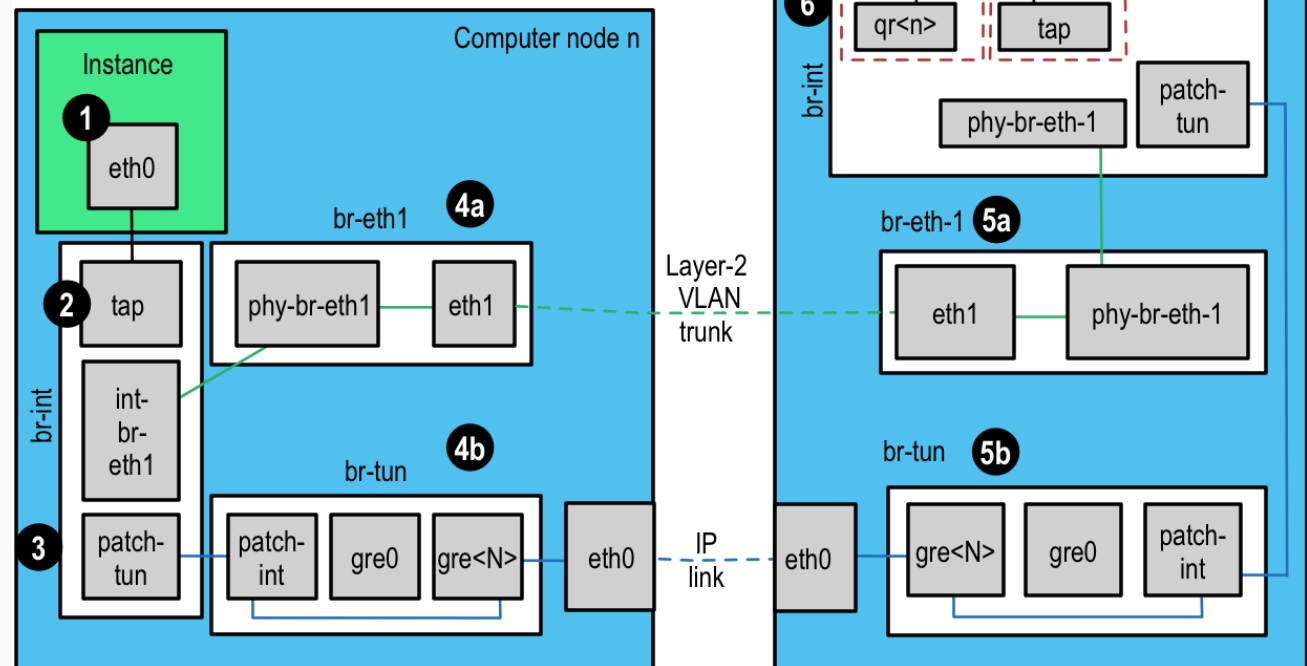


# A csomag útja

- » Test Access Point (TAP) device
- » int-br: integration bridge
- » br-eth1: VLAN internal/external címke fordítás
- » veth: int-br-eth1 és phy-br-eth1 között

Neutron network paths

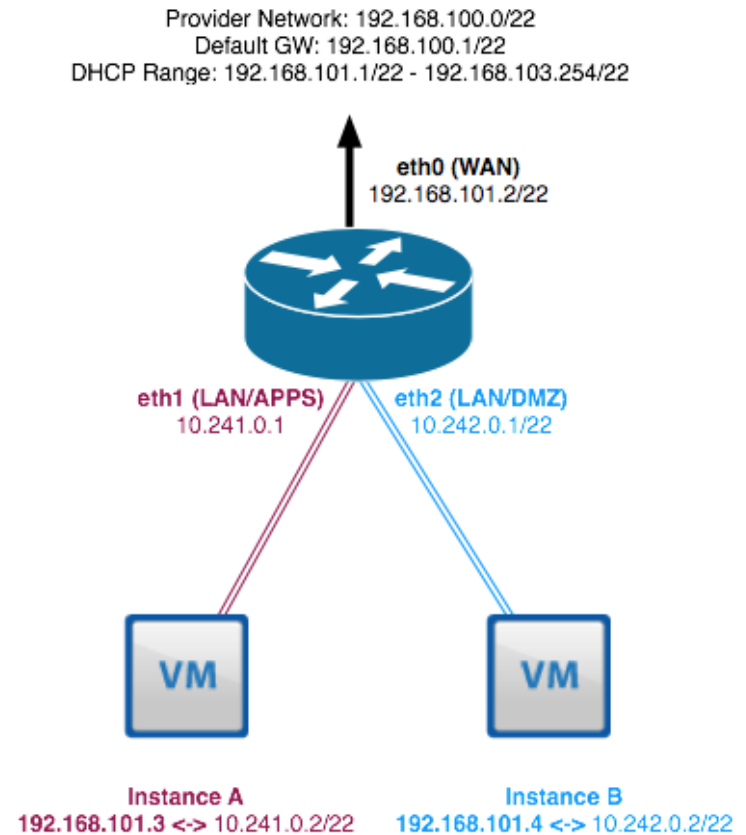
- VLAN networks
- GRE networks
- VLAN and GRE networks



# Floating IP

- » Neutron útválasztó
  - » gateway a VM-eknek
  - » iptables/NAT szabályok az útválasztó névterében
    - » nova network: a hypervisorban
  - » floating IP címek a fizikai útválasztó publikus címtartományából

**Diagram 1.1 - Logical Neutron Router**



- Diagram 1.1 -

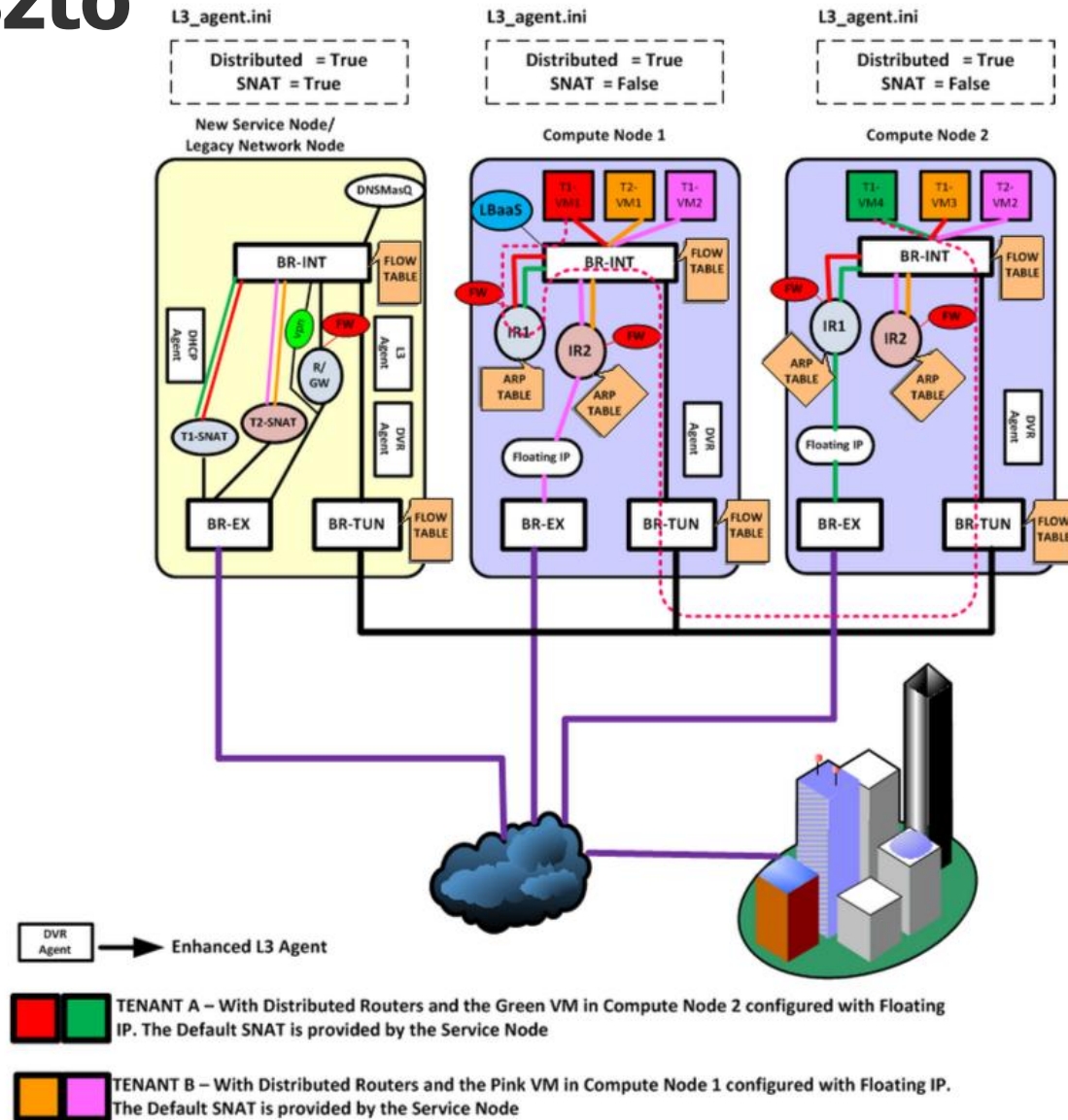
eth0 is connected to a PROVIDER network.  
 eth1 is connected to a TENANT network.  
 eth2 is connected to a TENANT network.

**Floating IPs** are assigned from the DHCP range of the PROVIDER network:

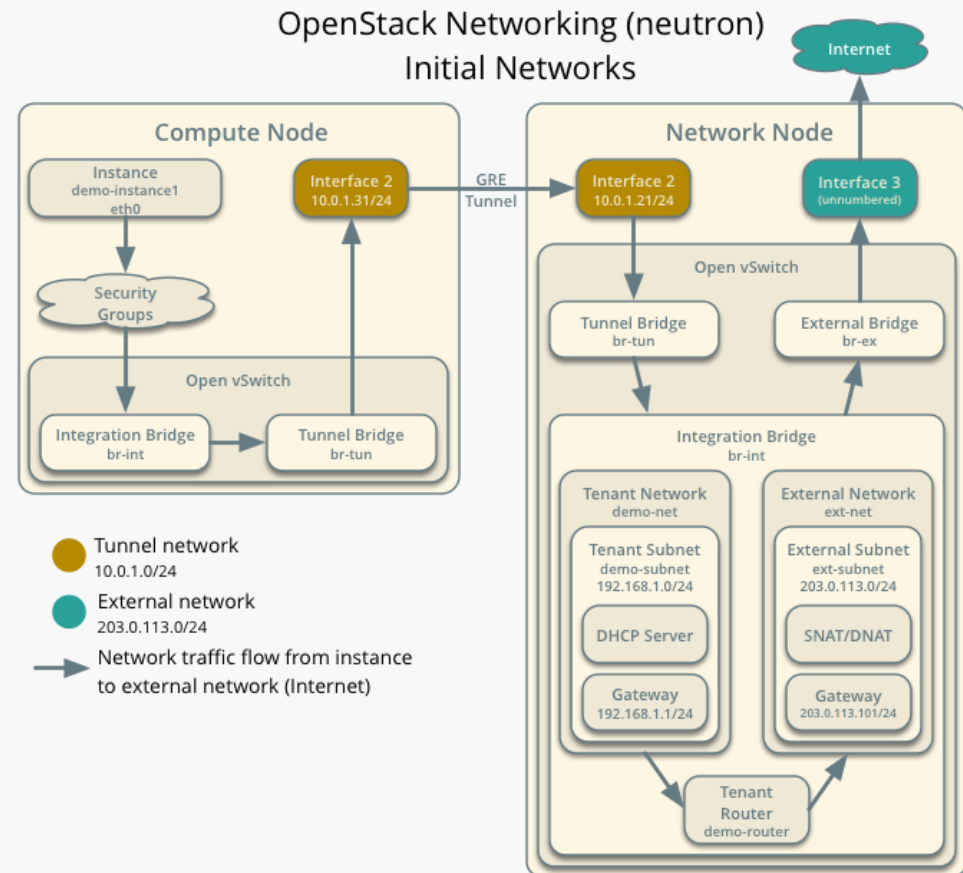
DHCP Range: 192.168.101.1/22 - 192.168.103.254/22

# Elosztott útválasztó

## » Distributed Virtual Router (DVR)



# Virtuális hálózatok kialakítása



## » Open vSwitch

- » szabályok megadása ovs-dpctl / OpenFlow segítségével
- » pl. leképezés a VM MAC címe és a hypervisor transport IP címe között



# Források

- » Overlay Virtual Networking Explained, Ivan Pepelnjak, NIL Data Communications, 2011.
- » <http://docs.openstack.org>
- » <https://developer.rackspace.com/blog/neutron-networking-l3-agent/>
- » <https://www.rdoproject.org/networking/networking-in-too-much-detail/>