Understanding OpenStack networking – Neutron

Университет ИТМО
Олег Лазо
oll@niif.spb.su

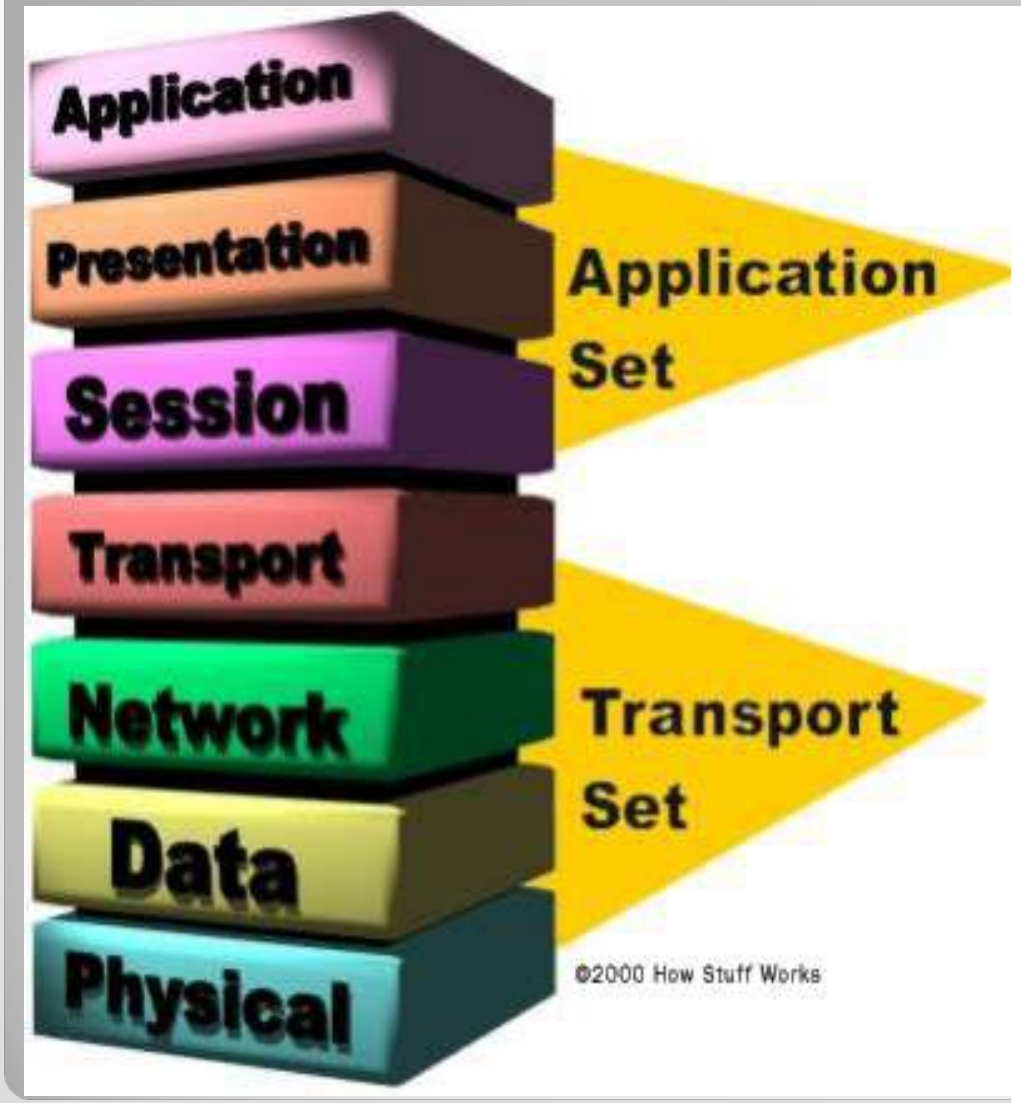
Программа

- OSI сетевая модель
- Транспортный набор, примеры
- Пример сетевой архитектуры
- Структура OpenStack Neutron
- Neutron Open vSwitch plug-in

Программа

- OSI сетевая модель
- Транспортный набор, примеры
- Пример сетевой архитектуры
- Структура OpenStack Neutron
- Neutron Open vSwitch plug-in

OSI network model



L7.Приложений – обеспечивает взаимодействие пользовательских приложений с сетью

L6.Представления – обеспечивает преобразование протоколов и шифрование/расшифровку данны

L5.Сеансовый – обеспечивает поддержание сеанса связи, позволяя приложениям взаимодействовать между собой длительное время

L4.Транспортный – предназначен для обеспечения надёжной передачи данных от отправителя к получателю

L3.Сетевой – предназначен для определения пути передачи данных в сети

L2.Канальный – предназначен для обеспечения взаимодействия с сетью по физическому уровню и контролем над ошибками

L1.Физический – определяет метод передачи данных, представленных в двоичном виде, от одного устройства к другому (физ.соединение)

Программа

- OSI сетевая модель
- Транспортный набор, примеры
- Пример сетевой архитектуры
- Структура OpenStack Neutron
- Neutron Open vSwitch plug-in

Transport set exampls

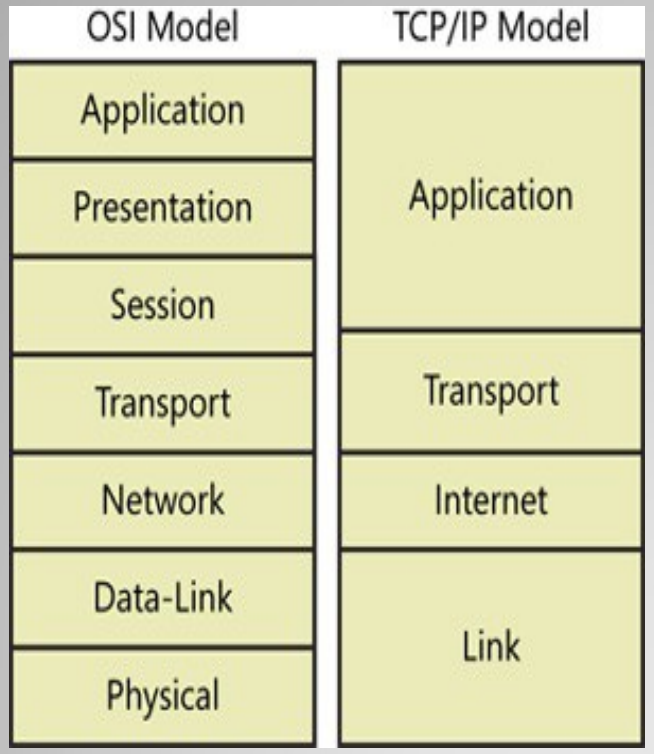
Level 1: IEEE 802.15 (Bluetooth), IRDA, EIA RS-232, EIA-422, EIA-423, RS-449, RS-485, DSL, ISDN, SONET/SDH, 802.11 Wi-Fi, GSM Um radio interface, ITU, ITU-T, TransferJet, ARINC 818 и G.hn/G.9960.

Level 2: ARCnet, ATM, Econet, Ethernet, FDDI (Fiber Distributed Data Interface), Frame Relay, HDLC (High-Level Data Link Control), IEEE 802.2 (provides LLC functions to IEEE 802 MAC layers), LAPD (Link Access Procedures, D channel), IEEE 802.11 wireless LAN, LocalTalk, MPLS (Multiprotocol Label Switching), PPP (Point-to-Point Protocol), PPPoE (Point-to-Point Protocol over Ethernet), SLIP (Serial Line Internet Protocol, obsolete), StarLan, Token ring, UDLD (Unidirectional Link Detection), x.25.

Level 3: IP/IPv4/IPv6 (Internet Protocol), IPX (Internetwork Packet Exchange), x.25 (part of this protocol is implemented at the level 2), CLNP (Connection Less Network Protocol), IPsec (Internet Protocol Security).

Level 4: ATP (AppleTalk Transaction Protocol), CUDP (Cyclic UDP), DCCP (Datagram Congestion Control Protocol), FCP (Fiber Channel Protocol), IL (IL Protocol), NBF (NetBIOS Frames protocol), NCP (NetWare Core Protocol), RTP (Real-time Transport Protocol), SCTP (Stream Control Transmission Protocol), SPX (Sequenced Packet Exchange), SST (Structured Stream Transport), TCP (Transmission Control Protocol), UDP (User Datagram Protocol).

Note



Пока комитеты ISO спорили о своих стандартах, за их спиной менялась вся концепция организации сетей и по всему миру внедрялся протокол TCP/IP.

…

И вот, когда протоколы ISO были наконец реализованы, выявился целый ряд проблем:

- эти протоколы основывались на концепциях, не имеющих в современных сетях никакого смысла;
- их спецификации были в некоторых случаях неполными;
- по своим функциональным возможностям они уступали другим протоколам;
- наличие многочисленных уровней сделало эти протоколы медлительными и трудными для реализации.

…

Сейчас даже самые ярые сторонники этих протоколов признают, что OSI постепенно движется к тому, чтобы стать маленькой сноской на страницах истории компьютеров.

— Эви Немет

TCP/IP protocols

Семейство TCP/IP имеет три транспортных протокола:

- TCP, полностью соответствующий OSI, обеспечивающий проверку получения данных;
- UDP, отвечающий транспортному уровню только наличием порта, обеспечивающий обмен датаграммами между приложениями, не гарантирующий получения данных;
- SCTP, разработанный для устранения некоторых недостатков TCP, в который добавлены некоторые новшества.

В семействе TCP/IP есть ещё около двухсот протоколов, самым известным из которых является служебный протокол ICMP, используемый для внутренних нужд обеспечения работы; остальные также не являются транспортными протоколами.

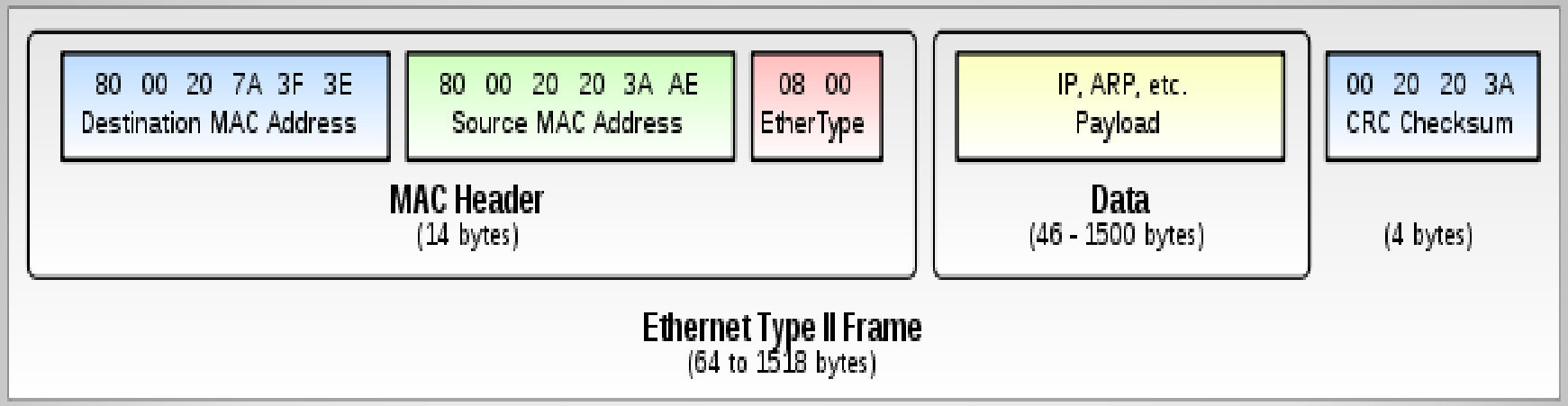
Популярные протоколы маршрутизации LAN:

- RIP (Routing Information Protocol);
- OSPF (Open Shortest Path First)
- ну и STATIC :)

Ethernet technology

Ethernet — семейство технологий пакетной передачи данных для компьютерных сетей. В стандарте первых версий (Ethernet v1.0 и Ethernet v2.0) указано, что в качестве передающей среды использовался коаксиальный кабель, в дальнейшем появилась возможность использовать витую пару и оптический кабель.

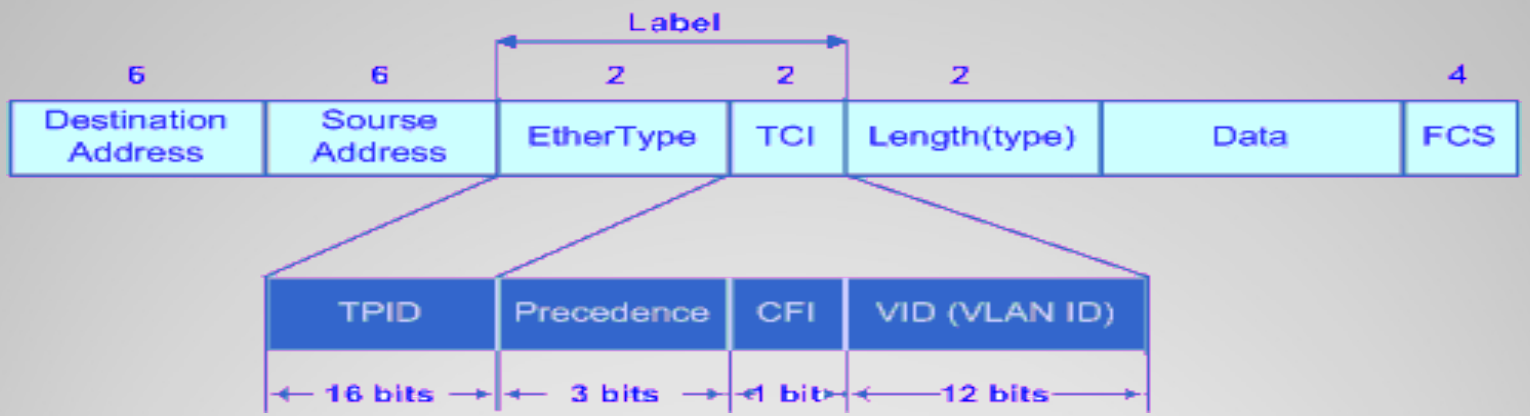
Наиболее распространенный формат кадра Ethernet II:



Разные типы кадра имеют различный формат и значение MTU.

Ethernet technology

В качестве дополнения Ethernet-кадр может содержать тег IEEE 802.1Q для идентификации VLAN, к которому он адресован, и IEEE 802.1p для указания приоритетности.

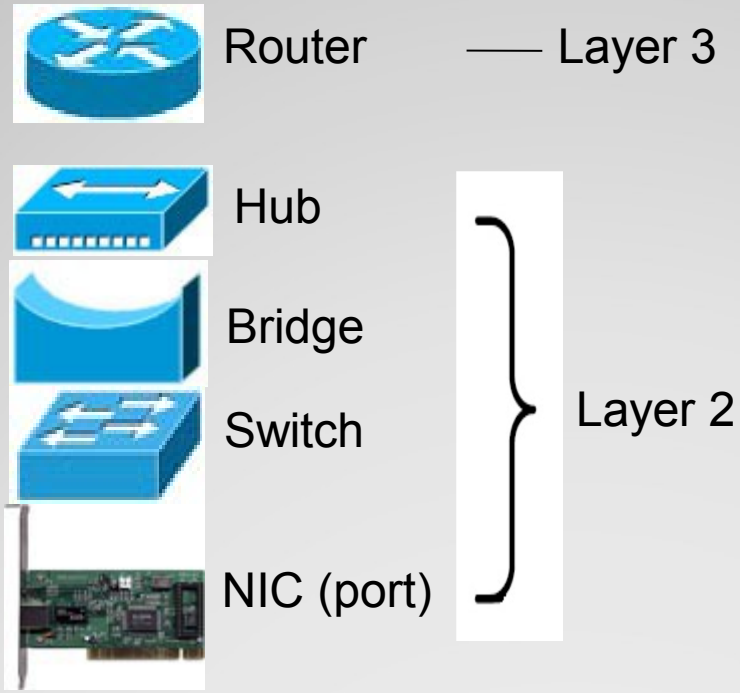


EtherType используется как TPID (Tagged Protocol Identifier). Для 802.1Q используется значение 0x8100.
TCI (Tagged Control Information) содержит CoS (Class of Service), Используется стандартом IEEE 802.1p для задания приоритета передаваемого трафика.
CFI (Canonical Format Identifier) and 12-bit field VID (VLAN ID), значения от 0 до 4095.

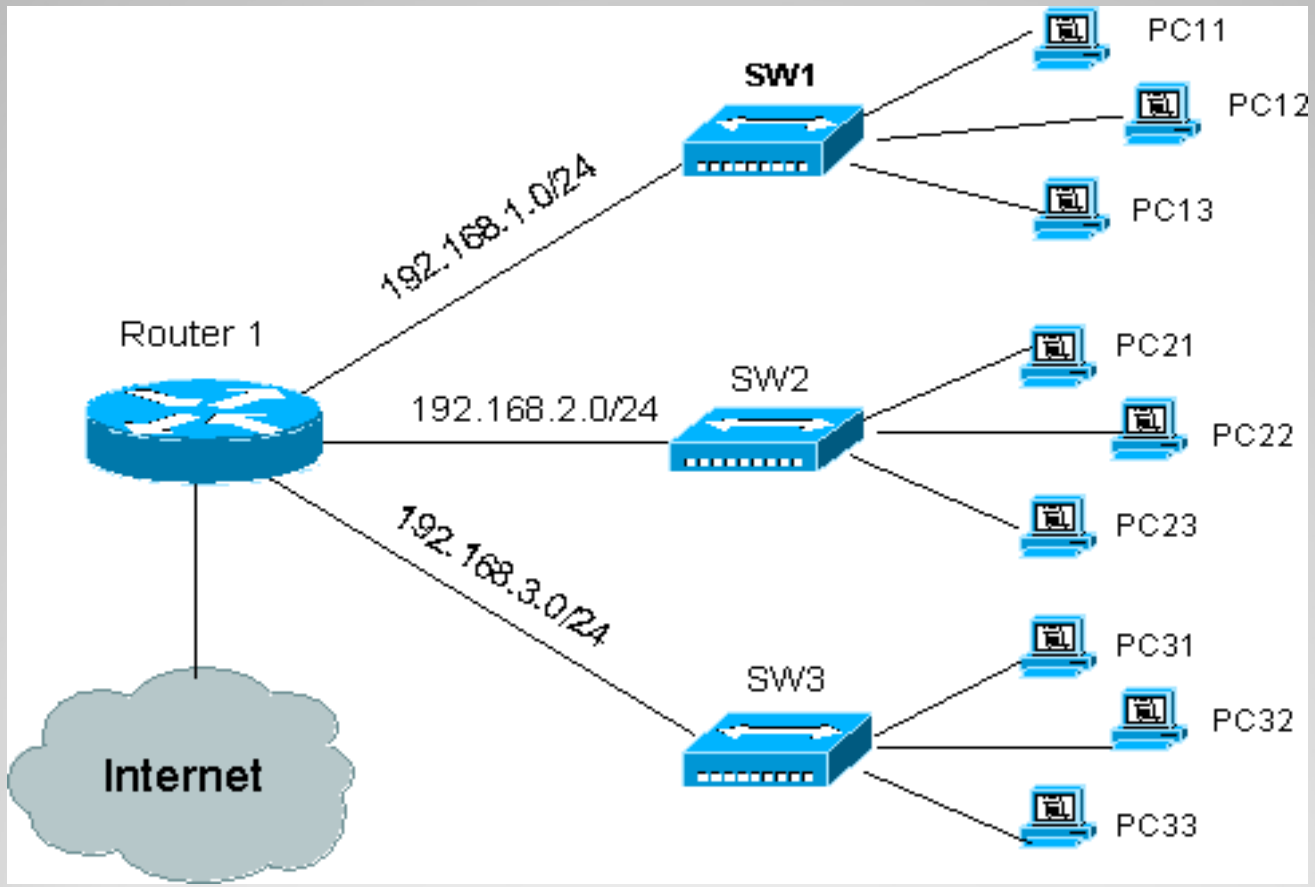
Программа

- OSI сетевая модель
- Транспортный набор, примеры
- Пример сетевой архитектуры
- Структура OpenStack Neutron
- Neutron Open vSwitch plug-in

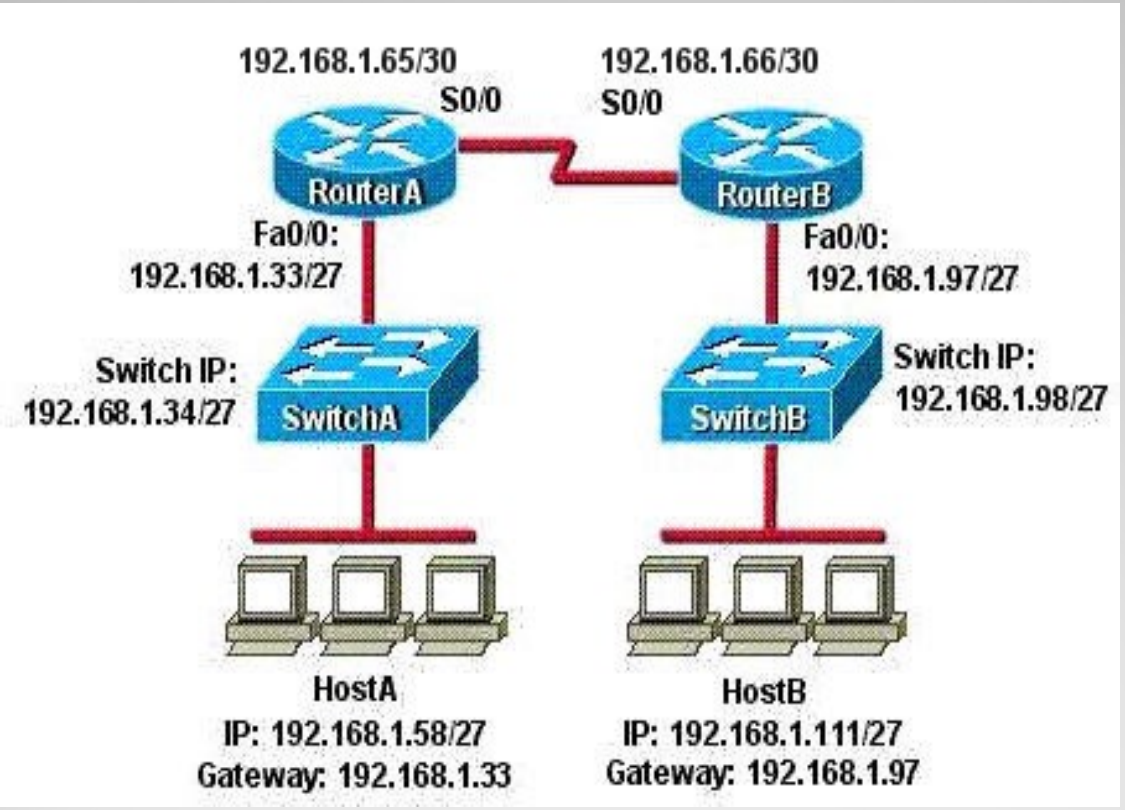
Compliance with network equipment OSI layer



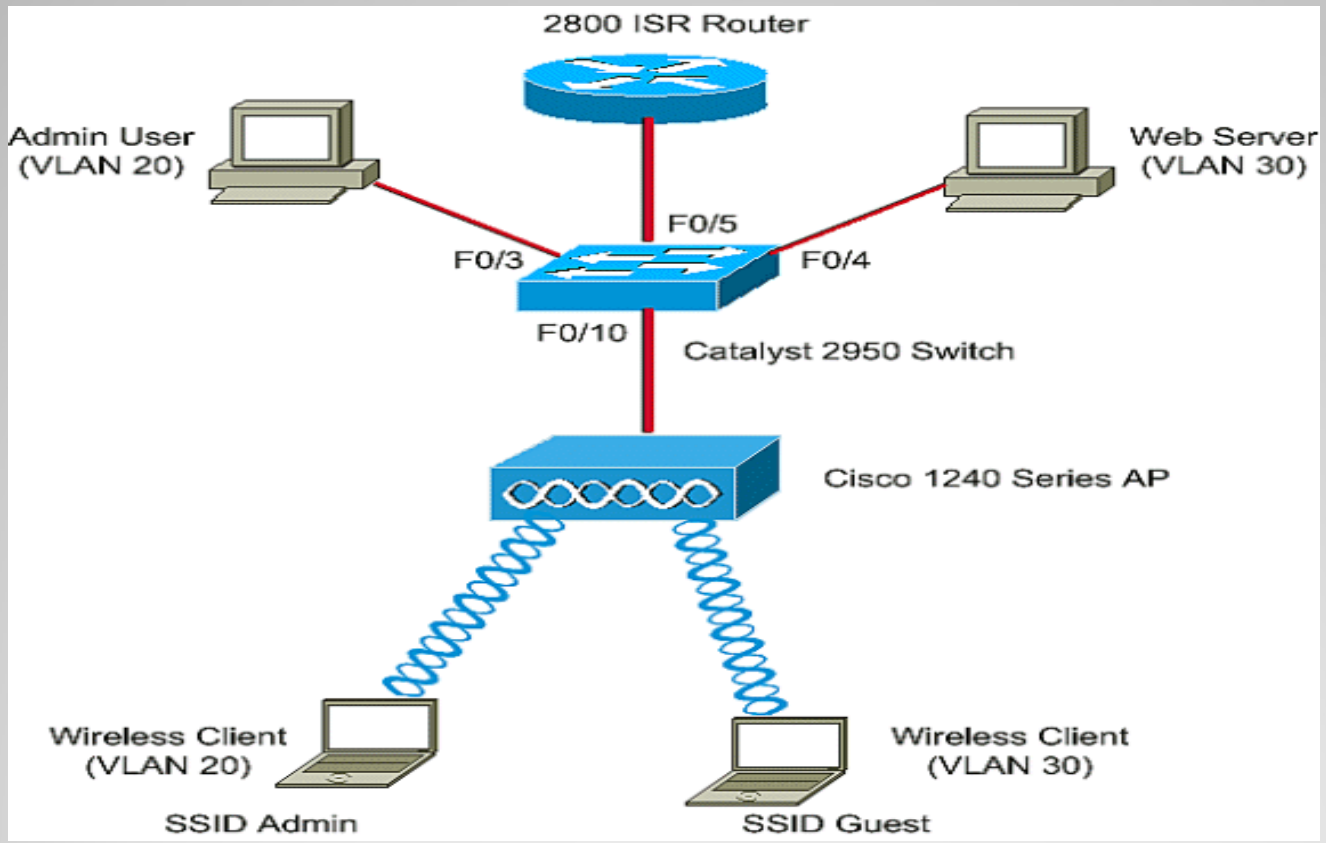
Typical LAN


PC11
SW1
PC12
PC13
192.168.1.0/24
Router 1
PC21
SW2
192.168.2.0/24
PC22
PC23
192.168.3.0/24
PC31
SW3
PC32
Internet
PC33

Another example of a 2 LANs



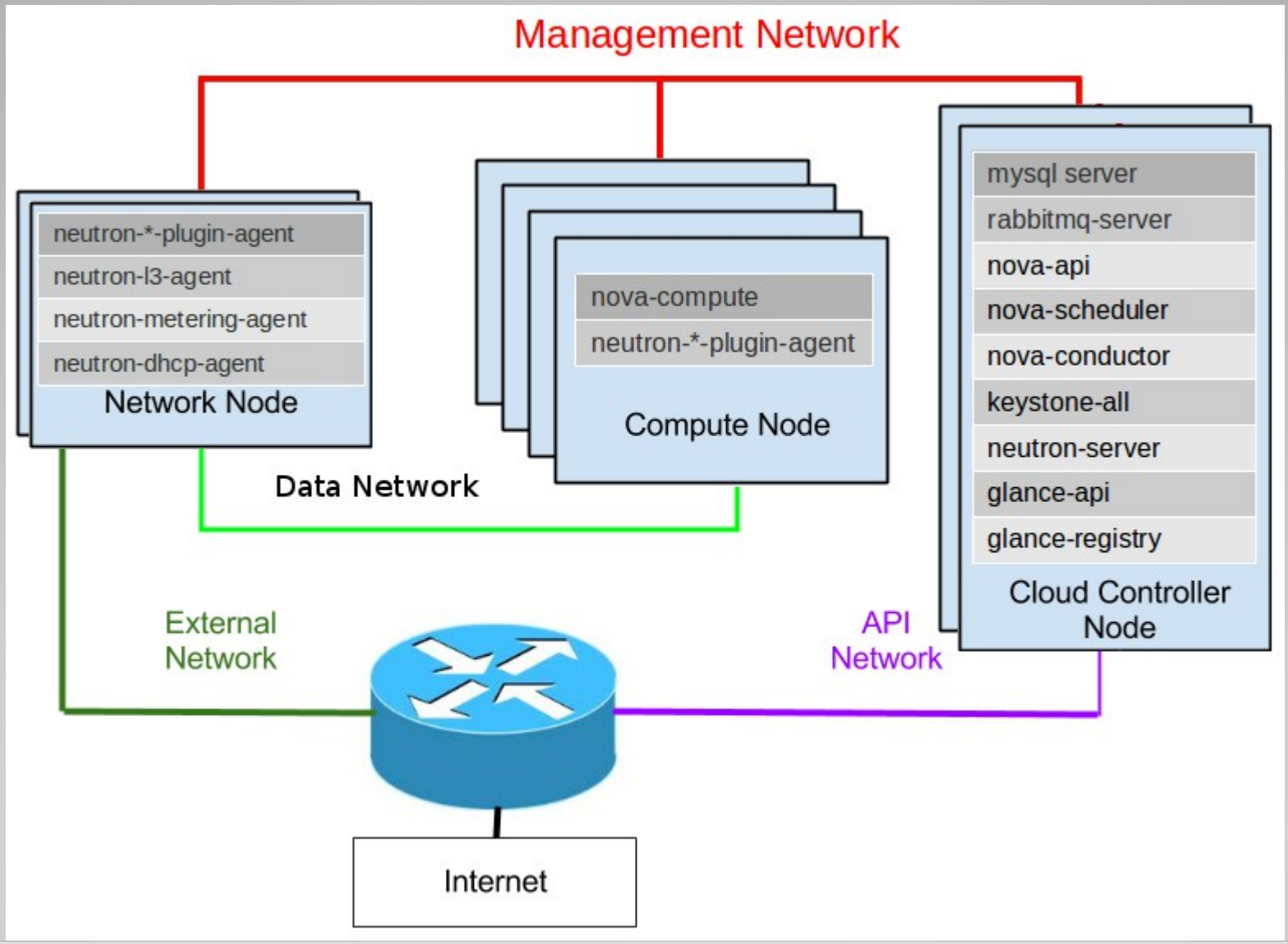
Example: LAN with VLANs and different connections



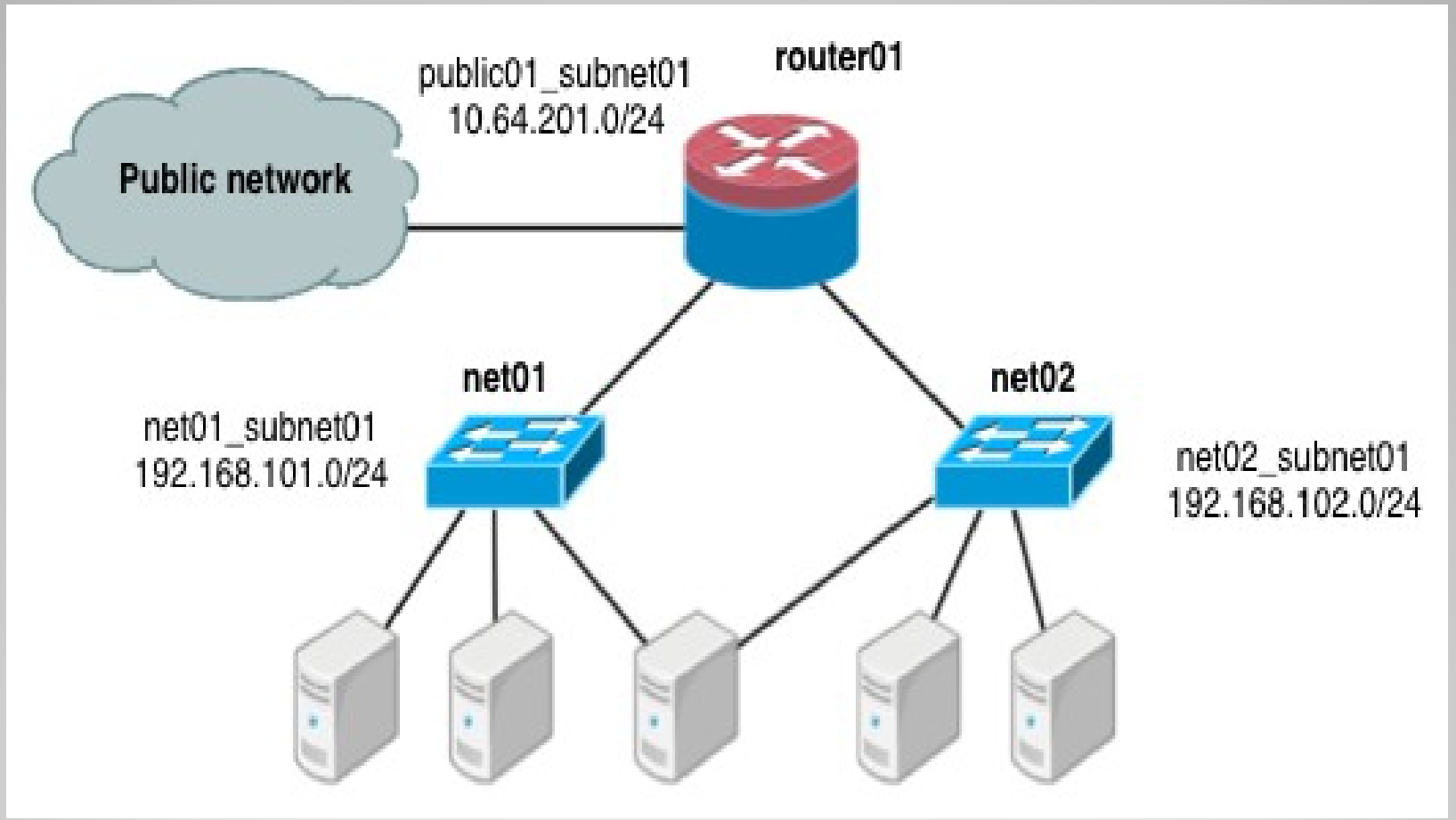
Программа

- OSI сетевая модель
- Транспортный набор, примеры
- Пример сетевой архитектуры
- Структура OpenStack Neutron
- Neutron Open vSwitch plug-in

Neutron physnet diagram



Example: virtual network Neutron

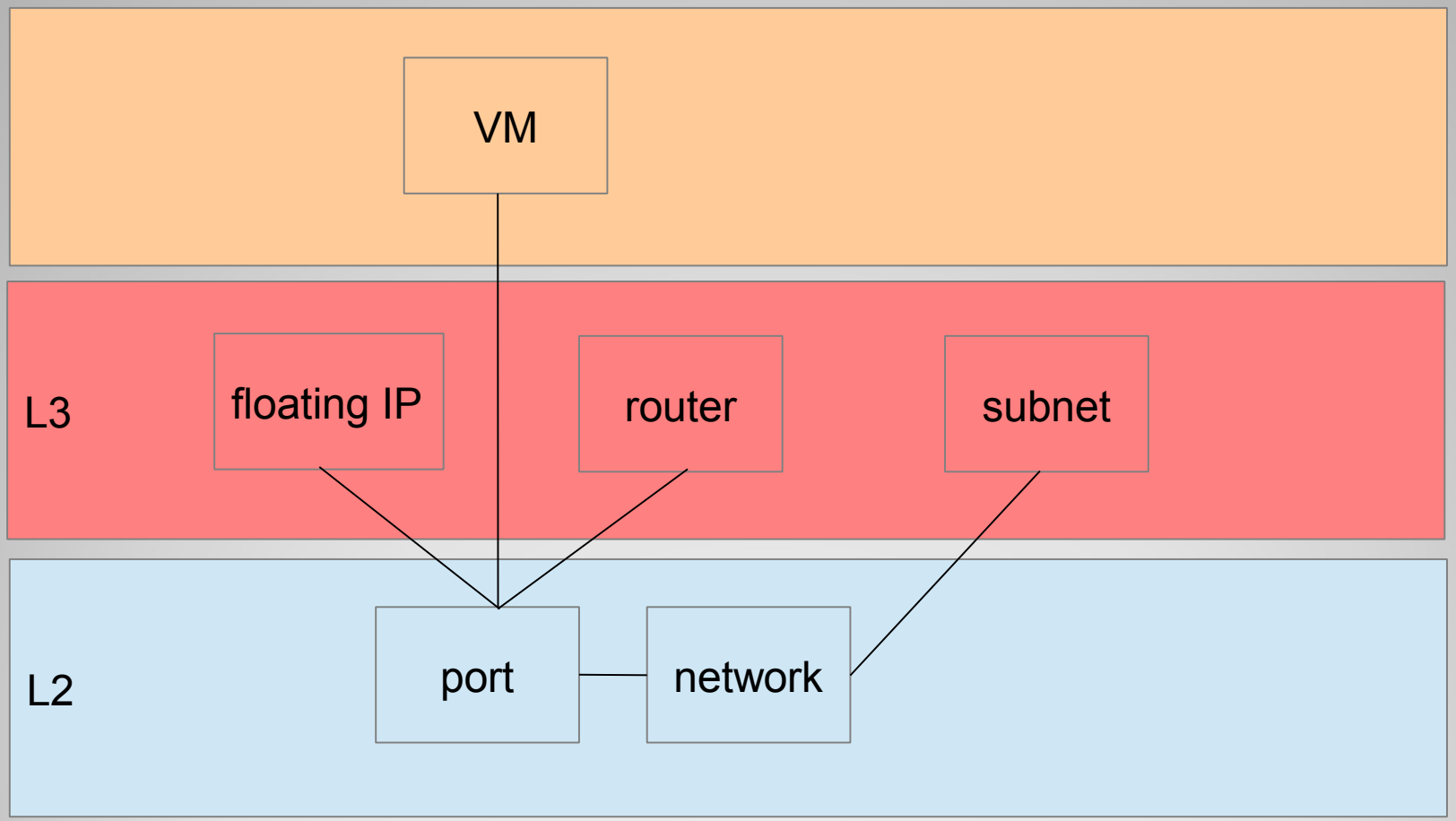


Neutron Objects

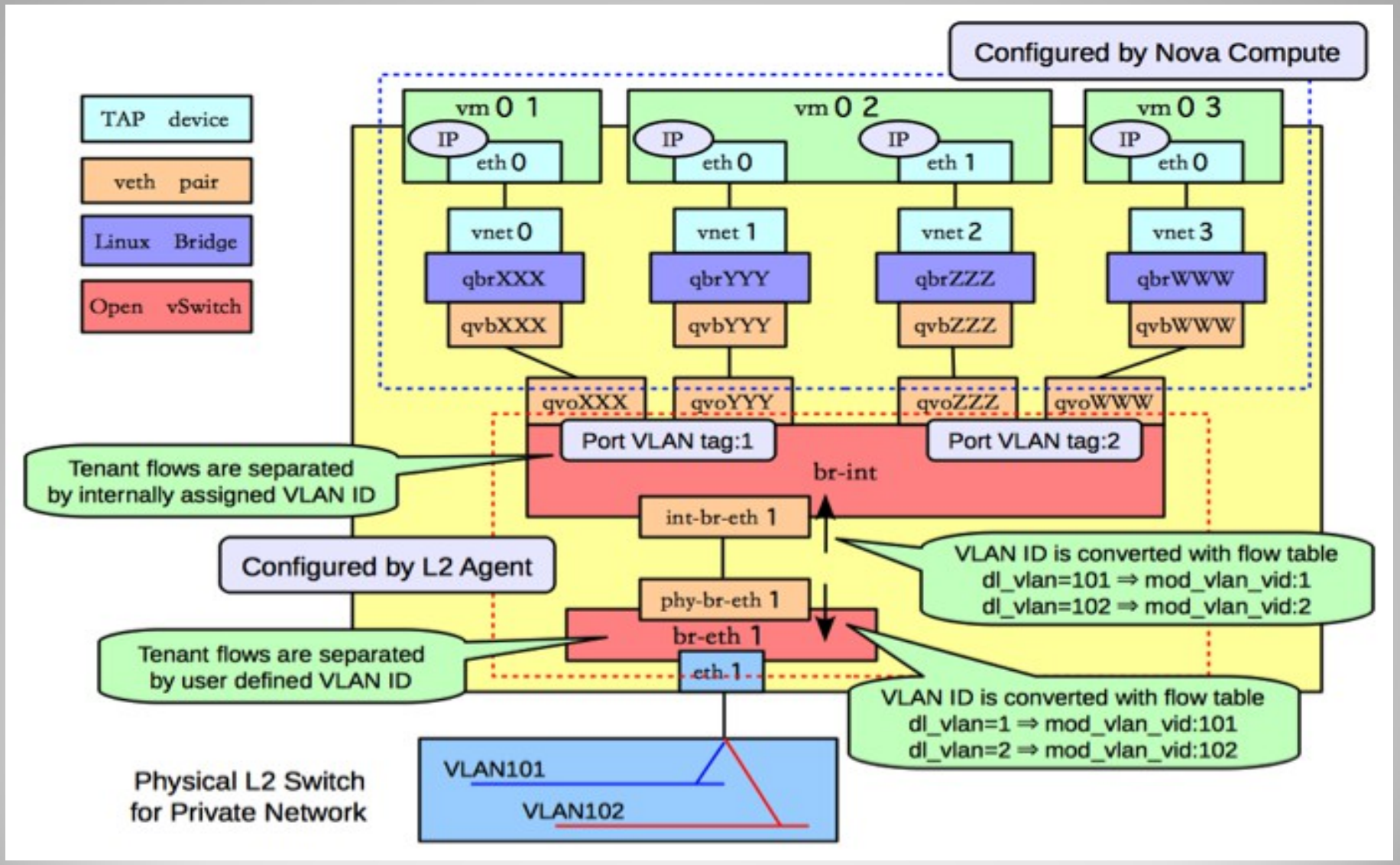
Core API objects

- Port – точка подключения к сети;
- Network – изолированный L2 сегмент сети;
- Subnet – блок IP адресов, связывающий L3 объекты сети;
- Router – роутер между сетями :)
- Floating IP – статическое мапирование из public IP внешней сети в private IP внутренней локальной сети;

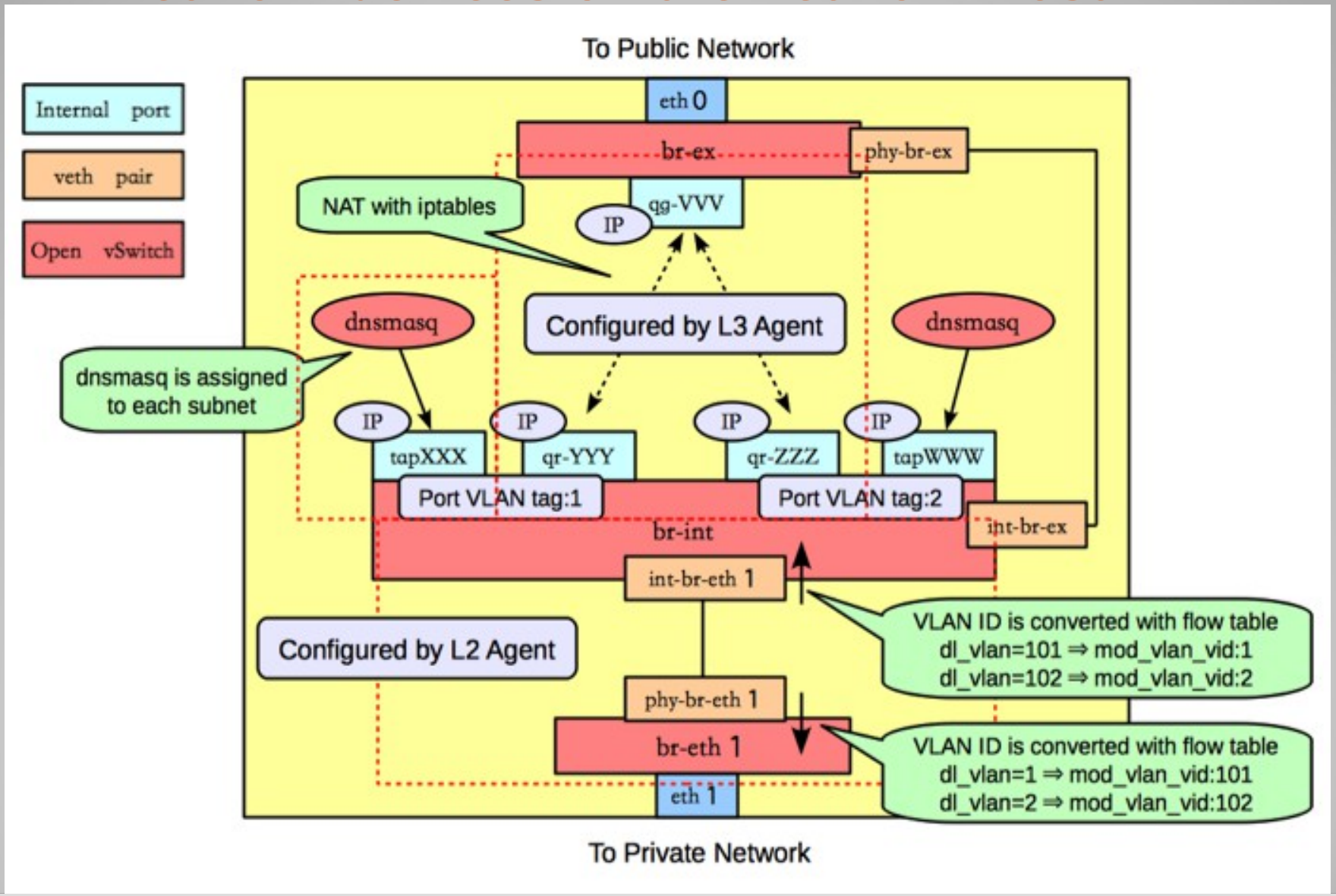
Neutron Object OSI Relations



Linux network devices on the comuter host



Network devices on the network host



Программа

- OSI сетевая модель
- Транспортный набор, примеры
- Пример сетевой архитектуры
- Структура OpenStack Neutron
- Neutron Open vSwitch plug-in

Open vSwitch plug-in

- Open source programmable virtual switch
- One of the most popular core plug-ins
- Supports OpenFlow, 802.1Q VLANs, GRE, LACP, STP
- Supports KVM and Xen
- OVS is the basis for different SDN/network virtualization platforms
- Flexible controller in user-space
- Fast datapath in kernel

Open vSwitch concepts

- Ports represent connections to other things, such as physical interfaces and patch cables;
- Packets from any given port on a bridge are shared with all other ports on that bridge;
- Bridges can be connected through OVS virtual patch cables or through Linux virtual Ethernet cables (veth);
- IEEE 802.1Q support attaching VLAN tags to interfaces;
- Support Generic Routing Encapsulation (GRE);
- Use network namespaces that enable Linux to group adapters into unique namespaces that are not visible to other namespaces;
- Fine-grained ACLs and QoS (L2-L4 matching, actions) by Hybrid IPTables/Open vSwitch plug-in;
- Centralized control via OpenFlow to transfer packets between VLANs;

Open vSwitch tools

ovs-vswitchd - daemon that implements a switch with help of kernel module

ovsdb-server - database server

ovs-vsctl - utility for working with the configuration

ovs-appctl - tool for controlling Open vSwitch daemon

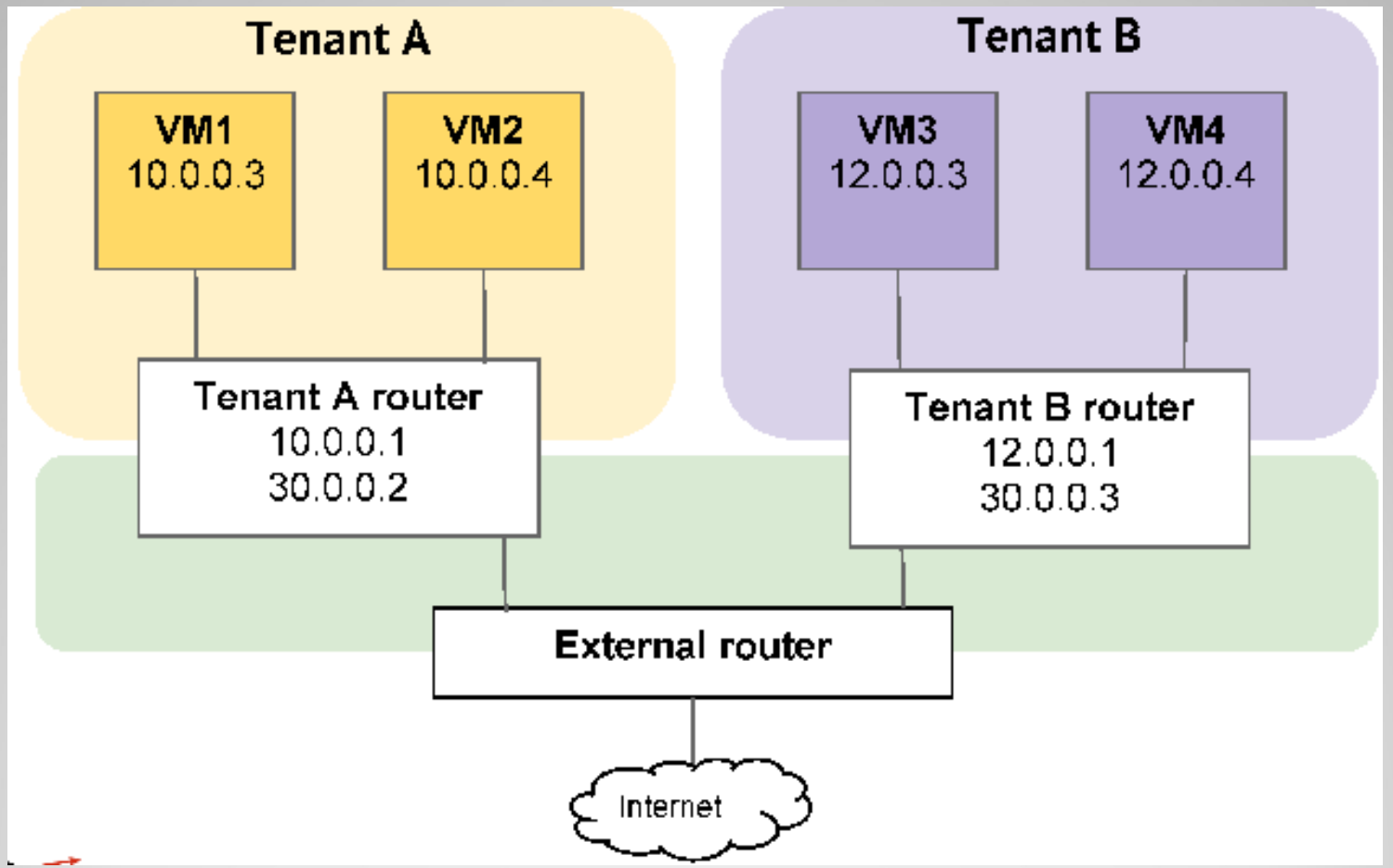
ovs-dpctl - datapath management utility

ovs-controller - simple OpenFlow controller

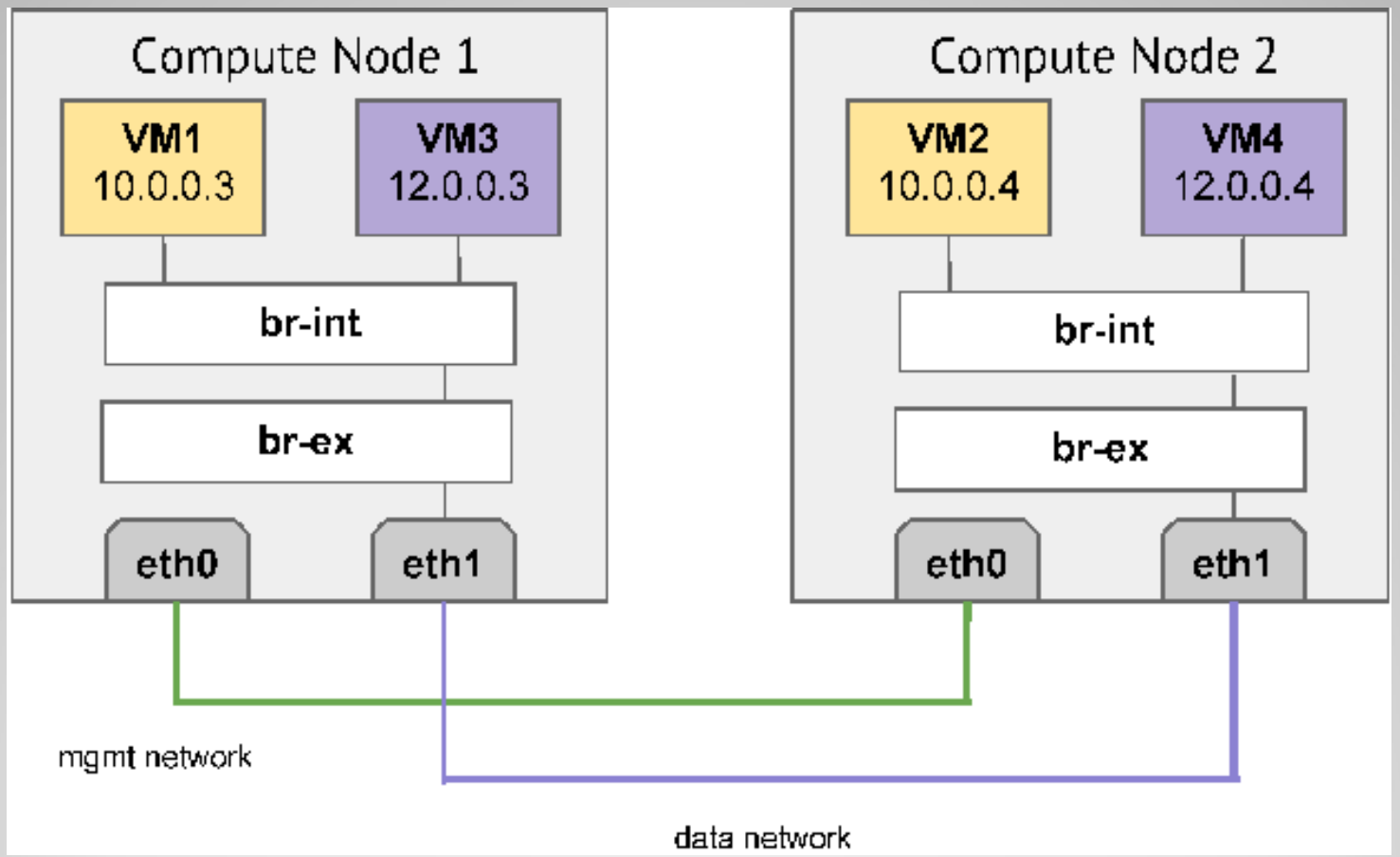
ovs-ofctl - OpenFlow switch management utility

ovs-pki - utility for managing public-key infrastructure

Example: logical view



Example: physical view



Let's Start (with example)!

We have:

- tenant A and network (10.0.0.0/24)
- router that wires private network with external
- DCHP enabled (neutron port is create)

Commands we need:

- **brctl show** - to show all bridges
- **ovs-vsctl show** - to show all interfaces
- **ip netns exec** - to show contents of namespace
- **neutron port-list**, **neutron net-list**

ovs-vsctl show

```
Bridge br-int
    Port "qr-9b80a882-55"
        tag: 1
        Interface "qr-9b80a882-55"
            type: internal
    Port "tap66a249f1-bf"
        tag: 1
        Interface "tap66a249f1-bf"
            type: internal
    Port br-int
        Interface br-int
            type: internal
Bridge br-ex
    Port "qg-e41c368d-a8"
        Interface "qg-e41c368d-a8"
            type: internal
    Port br-ex
        Interface br-ex
            type: internal
ovs_version: "1.11.0"
```

internal interface of router

port of DHCP server

external interface of router

brctl show

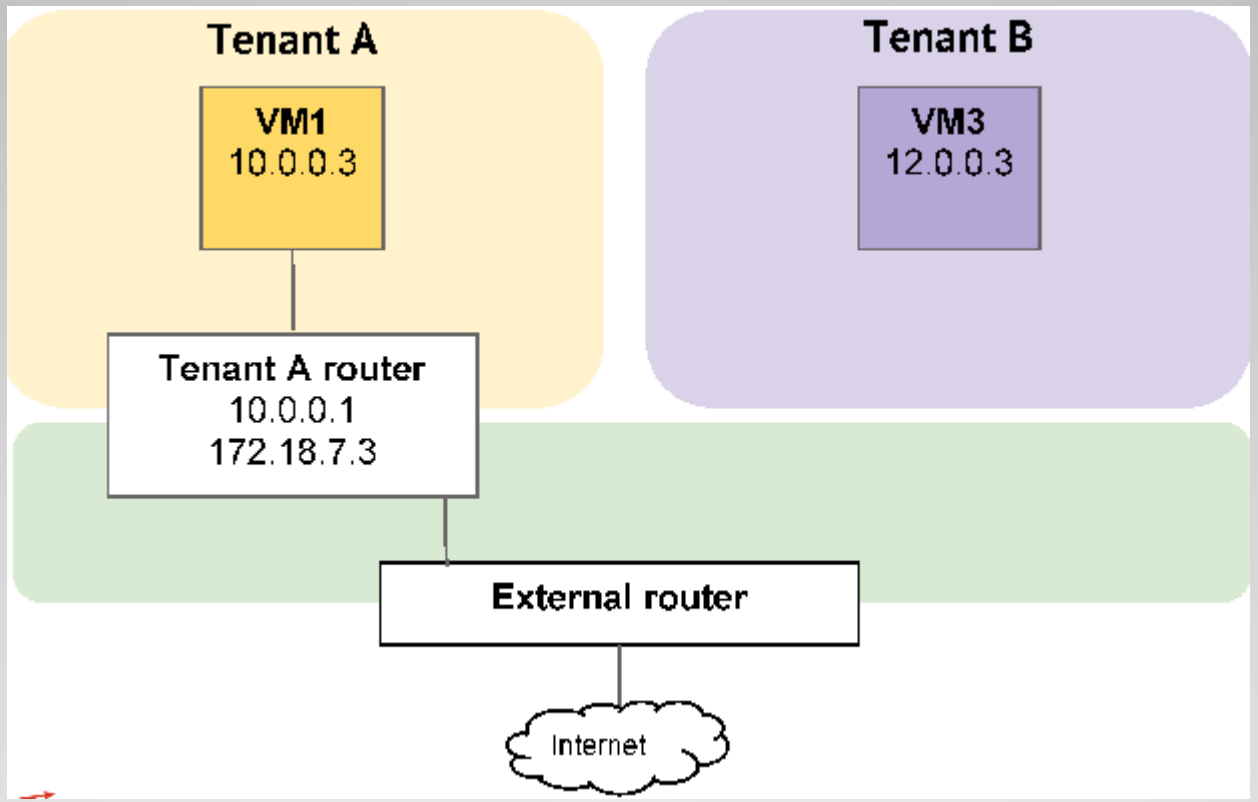
bridge name	bridge id	STP enabled	interfaces
br-ex	0000.6eed69b21a4b	no	qg-e41c368d-a8
br-int	0000.f68d58076046	no	qr-9b80a882-55
			tap66a249f1-bf

Let's begin

Expand the configuration:

- Launch VM in tenant A (10.0.0.3)
- Create network for tenant B (12.0.0.0/24) with DHCP enabled (12.0.0.2) but without router
- Launch VM in tenant B (12.0.0.3)

logical view



ovs-vsctl show

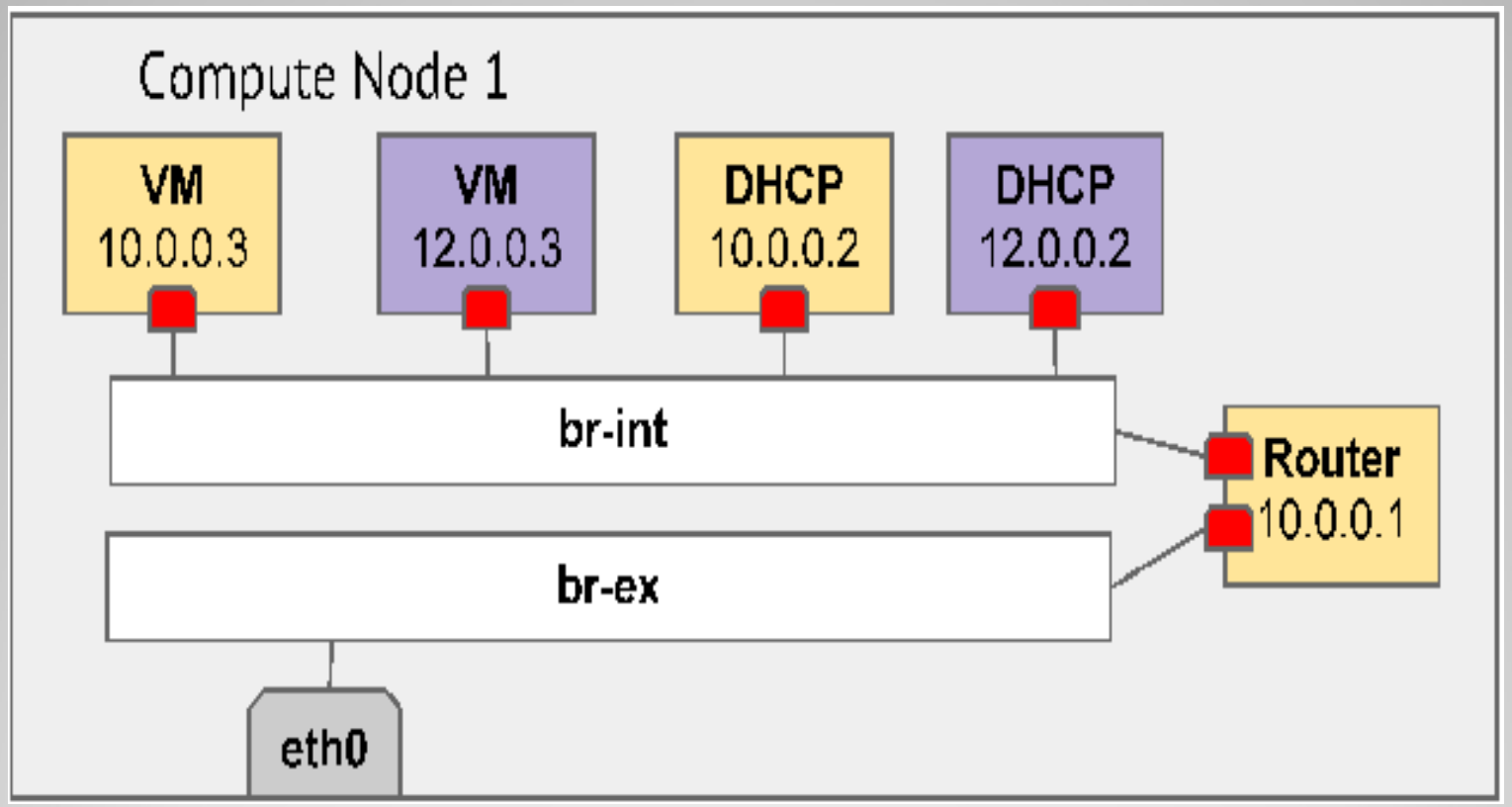
```
Bridge br-int
    Port "qvo8b0b577a-2c"
        tag: 1
        Interface "qvo8b0b577a-2c"
    Port "qr-9b80a882-55"
        tag: 1
        Interface "qr-9b80a882-55"
            type: internal
    Port "tap66a249f1-bf"
        tag: 1
        Interface "tap66a249f1-bf"
            type: internal
    Port "qvo4a744a65-92"
        tag: 2
        Interface "qvo4a744a65-92"
    Port "tap3aa4a560-d2"
        tag: 2
        Interface "tap3aa4a560-d2"
            type: internal
    Port br-int
        Interface br-int
            type: internal
```

interface for VM in tenant A

interface for VM in tenant B

interface of DHCP server in tenant B

physical view



brctl show

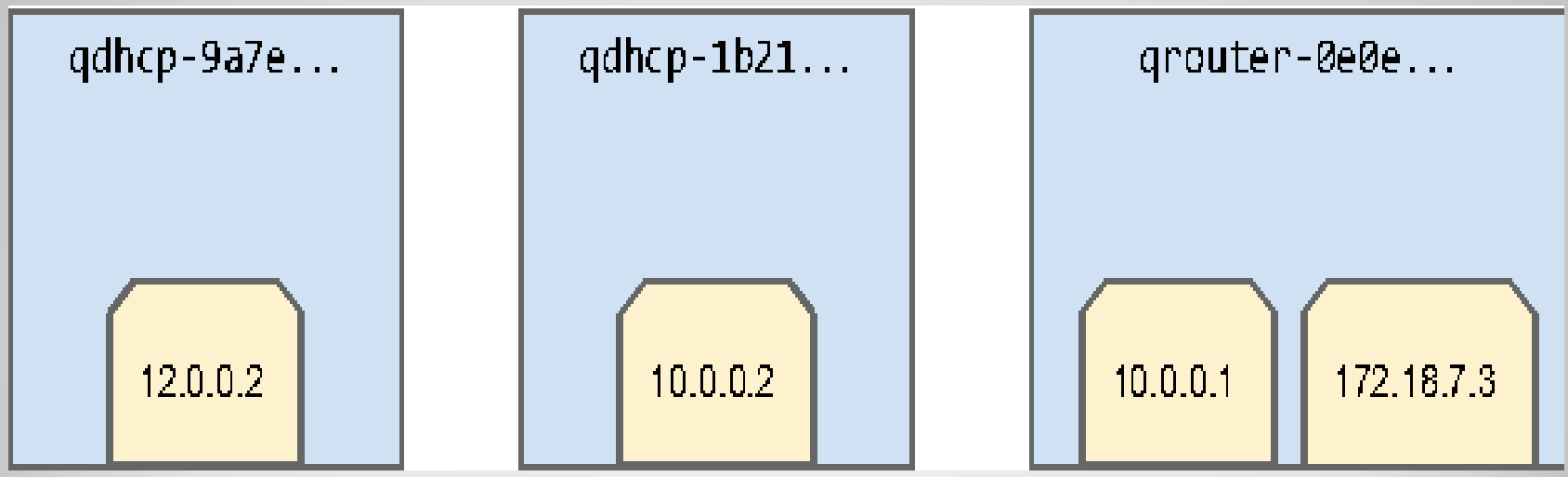
bridge name	bridge id	STP enabled	interfaces
br-ex	0000.6eed69b21a4b	no	qg-e41c368d-a8
br-int	0000.f68d58076046	no	qr-9b80a882-55
			qvo4a744a65-92
			qvo8b0b577a-2c
			tap3aa4a560-d2
			tap66a249f1-bf
qbr4a744a65-92	8000.7a95a8a2b9bd	no	qvb4a744a65-92
			tap4a744a65-92
qbr8b0b577a-2c	8000.de84d986f61e	no	qvb8b0b577a-2c
			tap8b0b577a-2c

one bridge for VM
(created by VIF driver in Nova)

ip netns list

List of namespaces. There's one namespace per DHCP port and per router port

qdhcp-9a7e9331-2508-4615-889b-b99a6f260eef
qdhcp-1b2101e0-cefa-4347-a581-e1f1f02215a1
qrouter-0e0e2e6e-a60b-4808-b914-8f45cae02b2e



ip netns exec qrouter-0e0e... ifconfig

Show interfaces for namespace associated with router

```
qg-e41c368d-a8 Link encap:Ethernet HWaddr fa:16:3e:27:a1:85
          inet addr:172.18.7.3 Bcast:172.18.7.7 Mask:255.255.255.248
          inet6 addr: fe80::f816:3eff:fe27:a185/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
          RX packets:60 errors:0 dropped:0 overruns:0 frame:0
          TX packets:69 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:6608 (6.6 KB)  TX bytes:5298 (5.2 KB)

qr-9b80a882-55 Link encap:Ethernet HWaddr fa:16:3e:9e:ed:50
          inet addr:10.0.0.1 Bcast:10.0.0.255 Mask:255.255.255.0
          inet6 addr: fe80::f816:3eff:fe9e:ed50/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
          RX packets:18878 errors:0 dropped:0 overruns:0 frame:0
          TX packets:3958 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:3269696 (3.2 MB)  TX bytes:349880 (349.8 KB)
```

FIN.