SACC 2014中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA 2014
发现架构之美

OpenStack网络介绍

龚永生 @UnitedStack
微博：云计算推手
微信：robingong2004
www.ustack.com

Contents

- 关于OpenStack
- 部署openstack
- Neutron的架构
- Neutron的核心模型
- Neutron网络技术

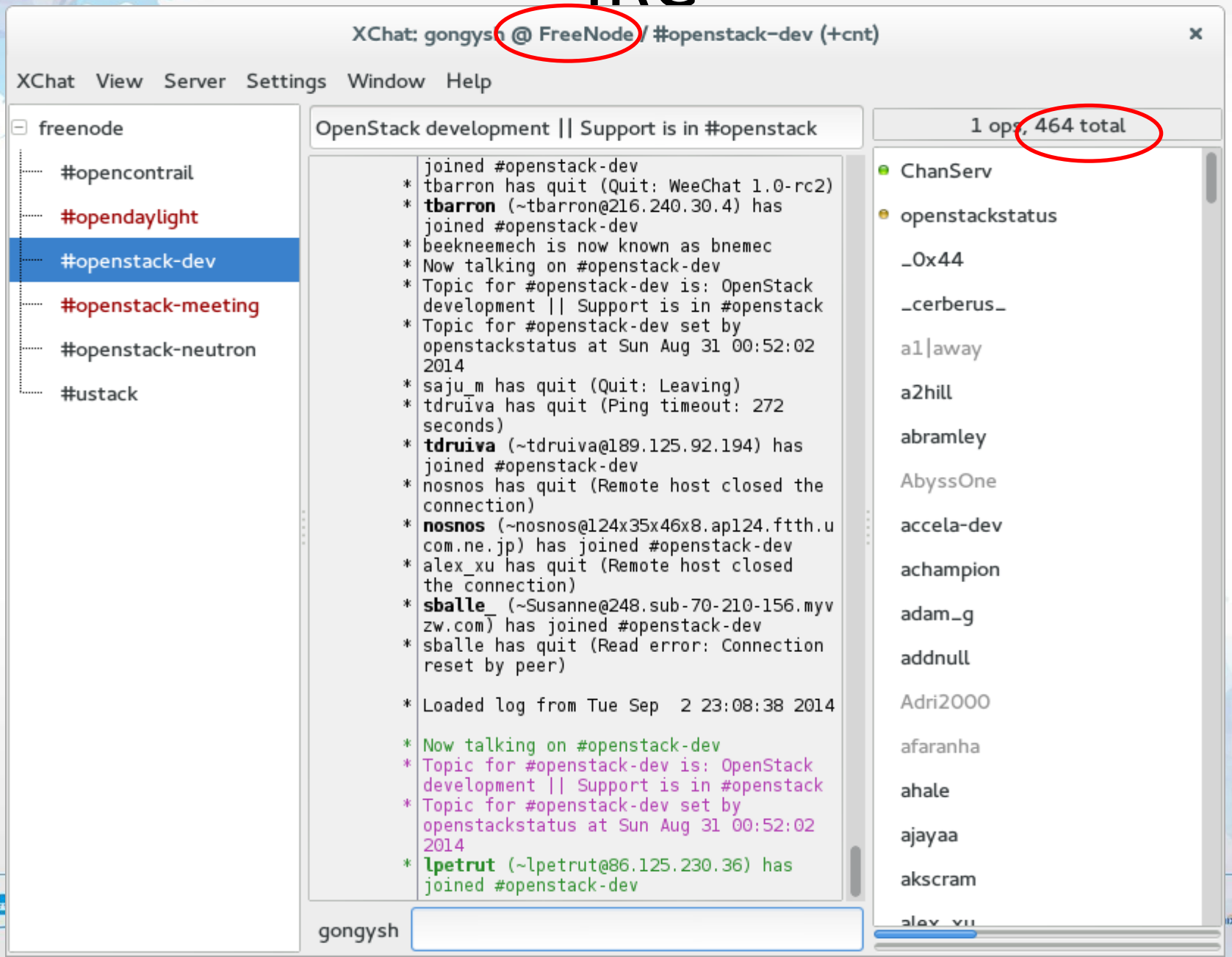
关于OpenStack

- 社区
- 项目管理方法
- CODE
- 文档
- 培训视频
- IaaS, PaaS, SaaS

社区

- community with more than 17,114 individual members and 390+ supporting organizations across 139 countries
- The Icehouse release had **1,202** contributors, a 32 percent increase from the Havana release six months ago. Approximately **350** new features and **2,902** bug fixes were added in the Icehouse release cycle, with **a focus on testing, maturity and stability**
- communications
  - IRC, mailist, various meet-up across the world

IRC


XChat: gongysh @ FreeNode / #openstack-dev (+cnt)
XChat View Server Settings Window Help
freenode
#opencontrail
#opendaylight
#openstack-dev
#openstack-meeting
#openstack-neutron
#ustack
OpenStack development || Support is in #openstack
1 ops, 464 total
joined #openstack-dev
* tbarron has quit (Quit: WeeChat 1.0-rc2)
* tbarron (~tbarron@216.240.30.4) has joined #openstack-dev
* beekneemech is now known as bnemec
* Now talking on #openstack-dev
* Topic for #openstack-dev is: OpenStack development || Support is in #openstack
* Topic for #openstack-dev set by openstackstatus at Sun Aug 31 00:52:02 2014
* saju_m has quit (Quit: Leaving)
* tdruiva has quit (Ping timeout: 272 seconds)
* tdruiva (~tdruiva@189.125.92.194) has joined #openstack-dev
* nosnos has quit (Remote host closed the connection)
* nosnos (~nosnos@124x35x46x8.ap124.ftth.ucom.ne.jp) has joined #openstack-dev
* alex_xu has quit (Remote host closed the connection)
* sballe_ (~Susanne@248.sub-70-210-156.myvzw.com) has joined #openstack-dev
* sballe has quit (Read error: Connection reset by peer)
* Loaded log from Tue Sep 2 23:08:38 2014
* Now talking on #openstack-dev
* Topic for #openstack-dev is: OpenStack development || Support is in #openstack
* Topic for #openstack-dev set by openstackstatus at Sun Aug 31 00:52:02 2014
* lpetrut (~lpetrut@86.125.230.36) has joined #openstack-dev
ChanServ
openstackstatus
_0x44
cerberus
a1|away
a2hill
abramley
AbyssOne
accela-dev
achampion
adam_g
addnull
Adri2000
afaranha
ahale
ajayaa
akscram
gongysh

Meetup

[OpenStack Meetup] 8月份meet up资料已上传，9月份meet up预告并征集演讲嘉宾

发件人：Daisy Guo <daisy.ycguo@gmail.com> 自动归档

时　间：2014年9月11日(星期四) 下午4:29

收件人：undisclosed-recipients:

大家好。

8月份的meet up资料已经上传，您可以在http://www.meetup.com/China-OpenStack-User-Group/files/下载，需要注册成为组成员才有下载权限。

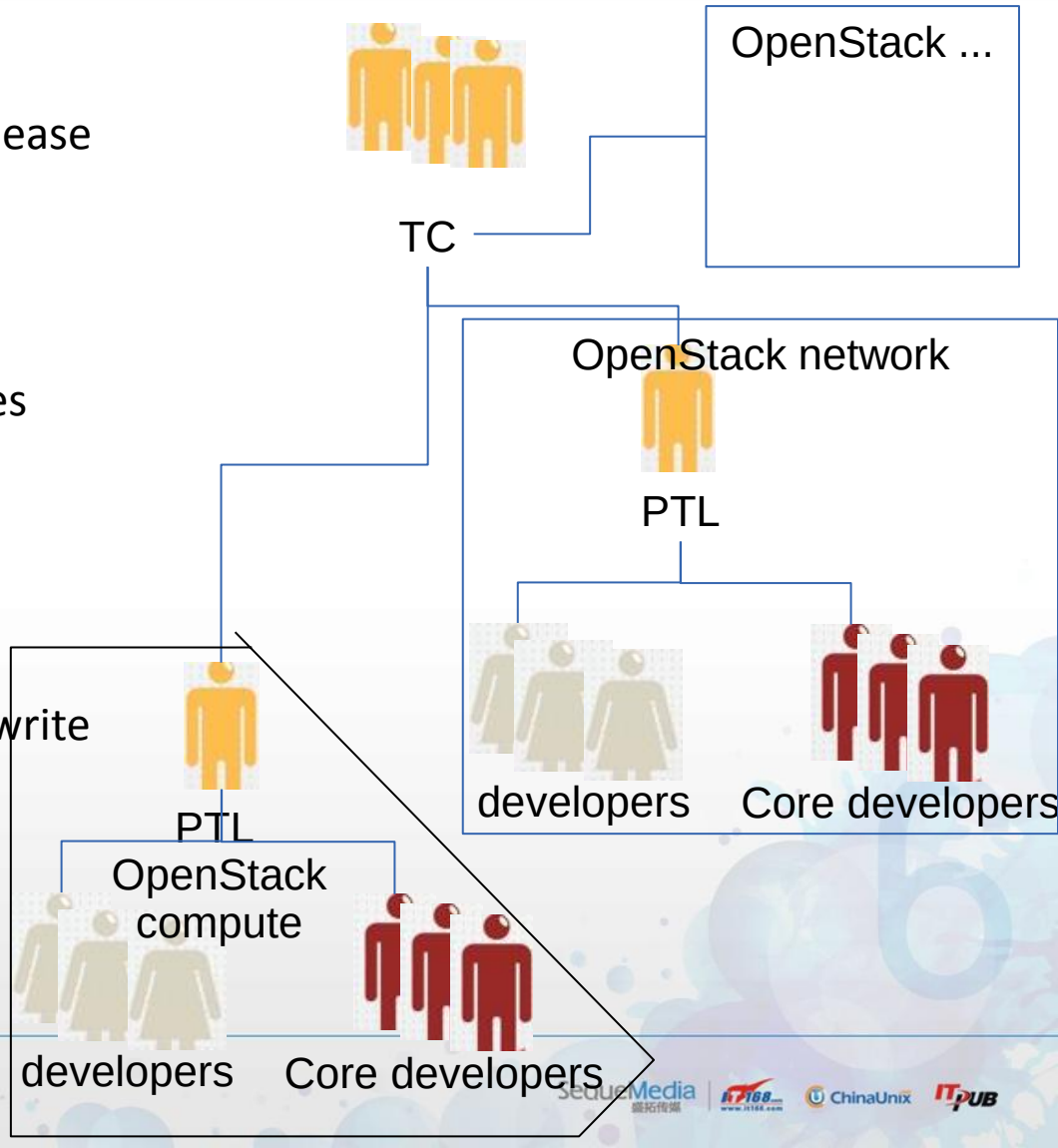
9月份的meet up的主题是目前的大热话题——Docker，已经邀请到Docker社区合办，Docker社区会出一位专家来为我们深度解读Docker，同时在OpenStack社区这边，希望邀请到一位嘉宾，来为大家讲解：OpenStack中的Docker使用现状及趋势，目前还没有合适的人选。如果有对这个话题有研究、或者愿意花时间对这个话题做调研的朋友，请跟我联系。同时，也征集赞助场地费用的公司，有兴趣的公司代表请和我联系。谢谢！

另外，为了沟通方便，我建立一个OpenStack Meet up的微信群，大家可以在里面踊跃发言，提出需求，推荐嘉宾，等等。请扫码：

Daisy

项目管理组织

- technique committee - scrum of scrum
  - elected by ATC for every release
- Project Technique lead - PTL
  - elected by developers for every release
- core developers
  - to triage bugs, reviewing and approving patches
  - design and implement new features
  - nominated by PTL from active developers
- developers
  - any one can submit the codes to implement new features, fix bugs write documents.
  - encouraged review patches.

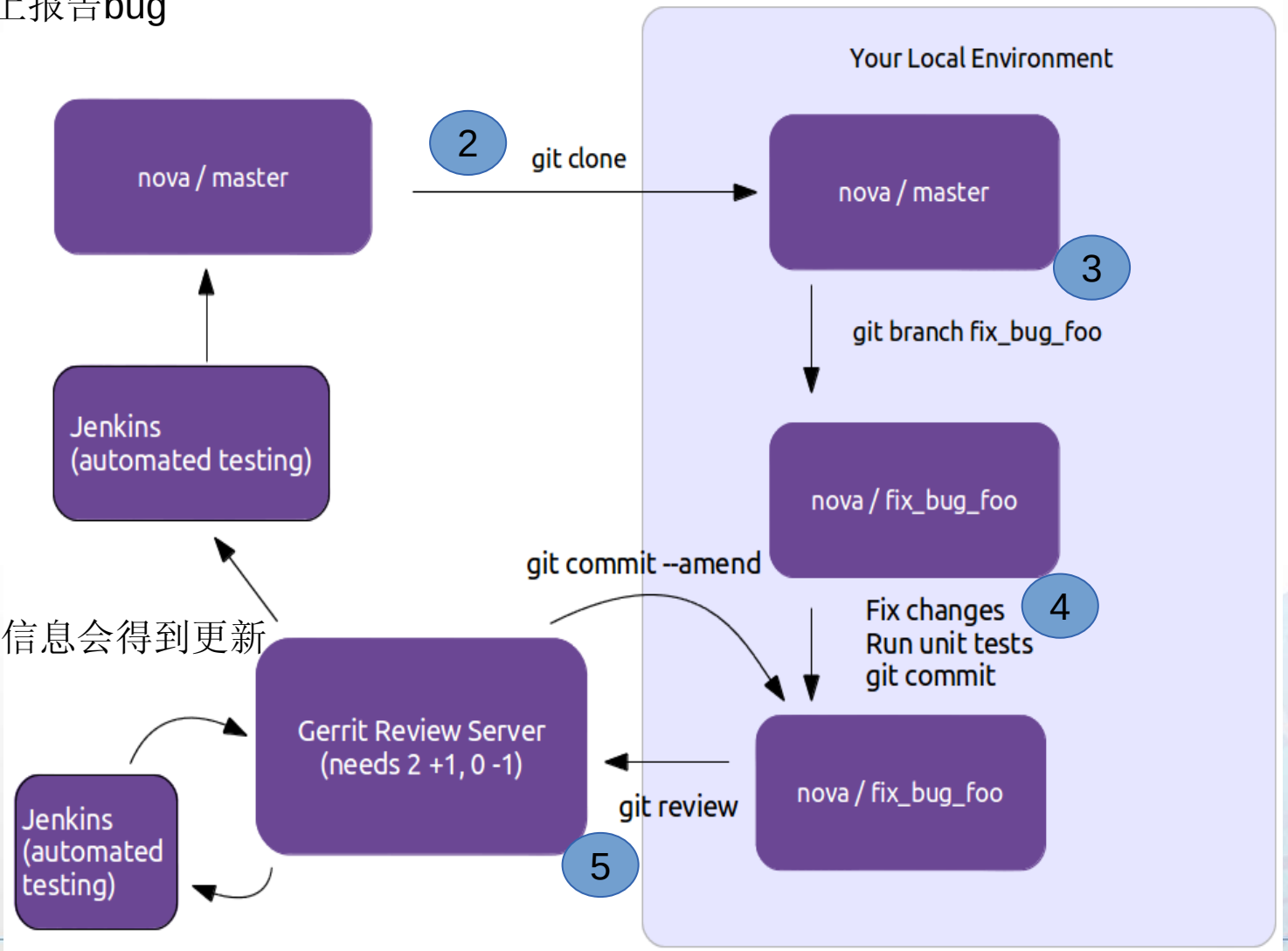


项目管理及其工具

- Launchpad
  - 基于web界面的功能和bug管理平台
- Gerrit
  - 基于web界面的功能，文档和代码评审系统（平民策略，大家都参与）
- CI
  - Jenkins based build and gate system
- devstack
  - 简单和快速的部署系统

修复 bug流程

1 在launchpad上报告bug

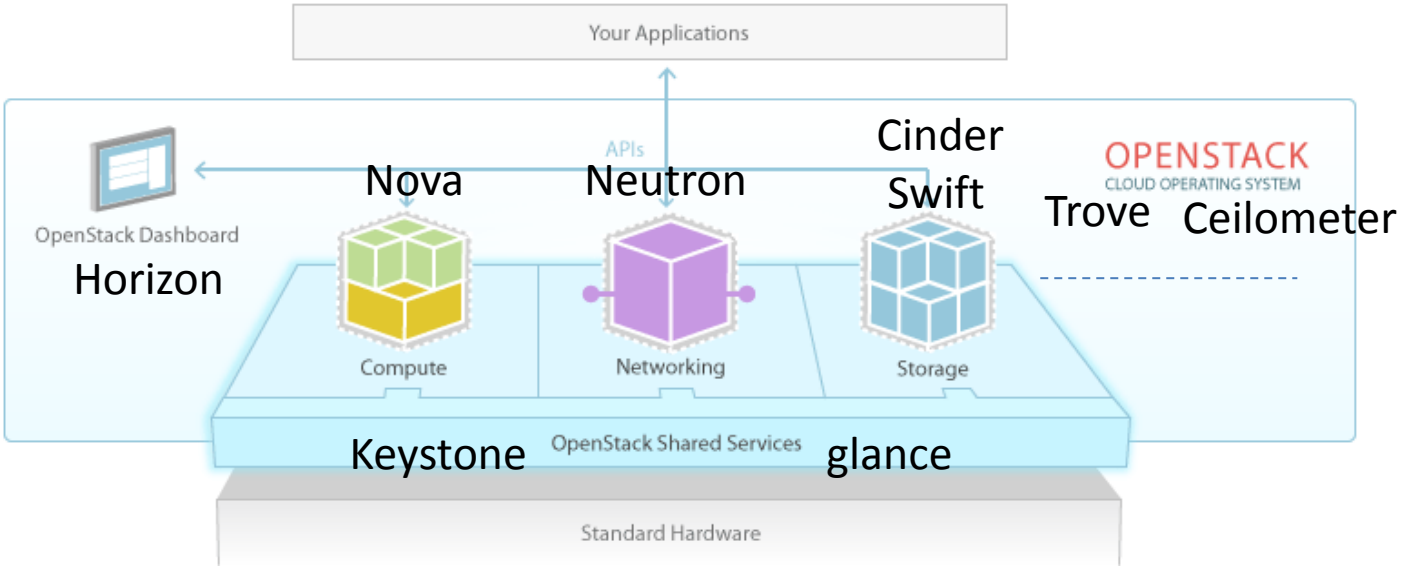


当补丁被合并，bug的信息会得到更新

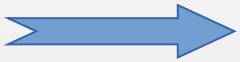
OpenStack项目

虚拟机	Nova
权限认证	Keystone
镜像管理	Glance
控制面板	Horizon
对象存储	Swift
虚拟网络	Neutron
虚拟块存储	Cinder
应用编排	Heat
监控	ceilometer
数据库服务	Trove
Hadoop服务	Savana
…	…

OpenStack核心之一Neutron



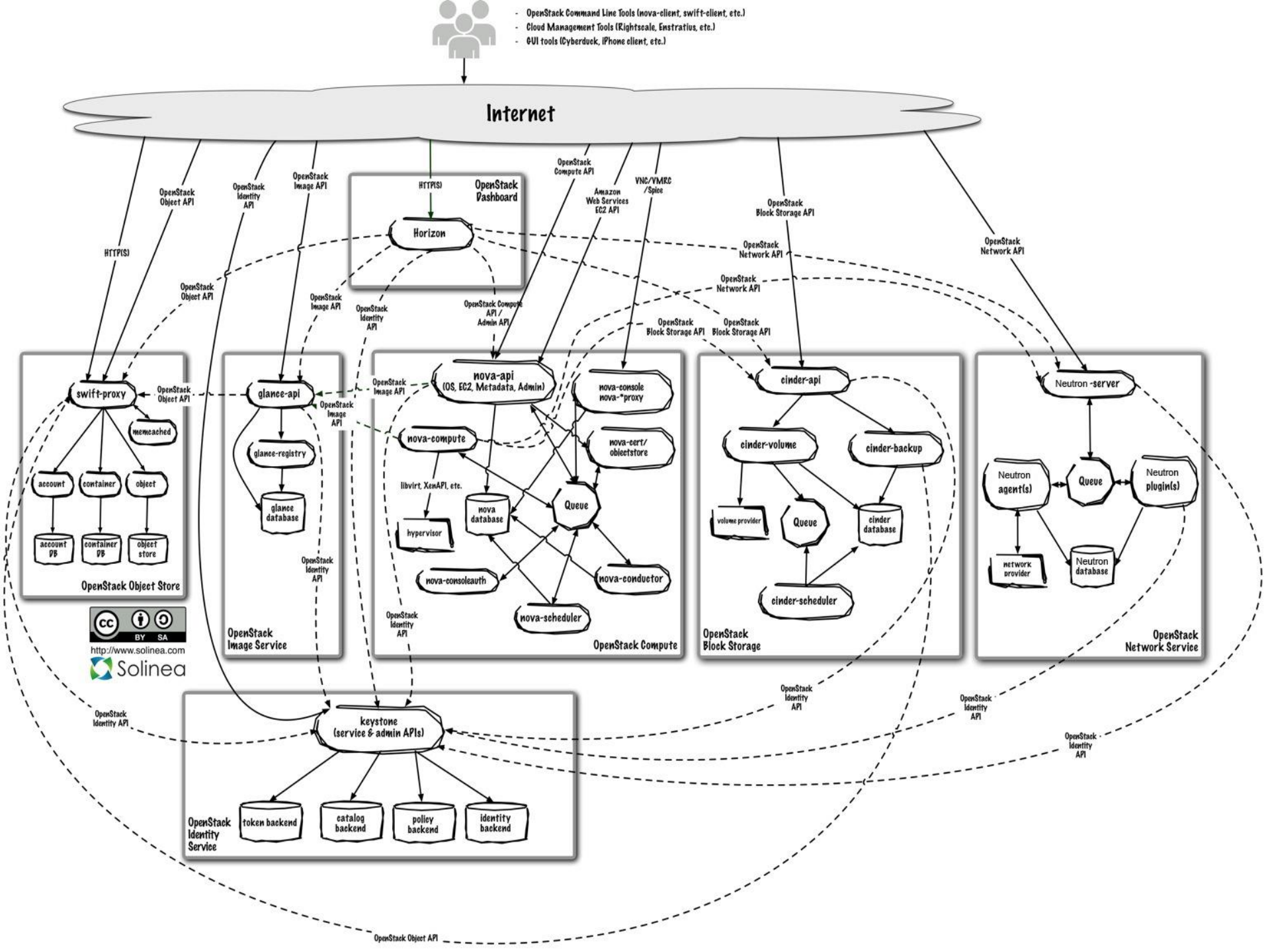
nova sub module
nova-network



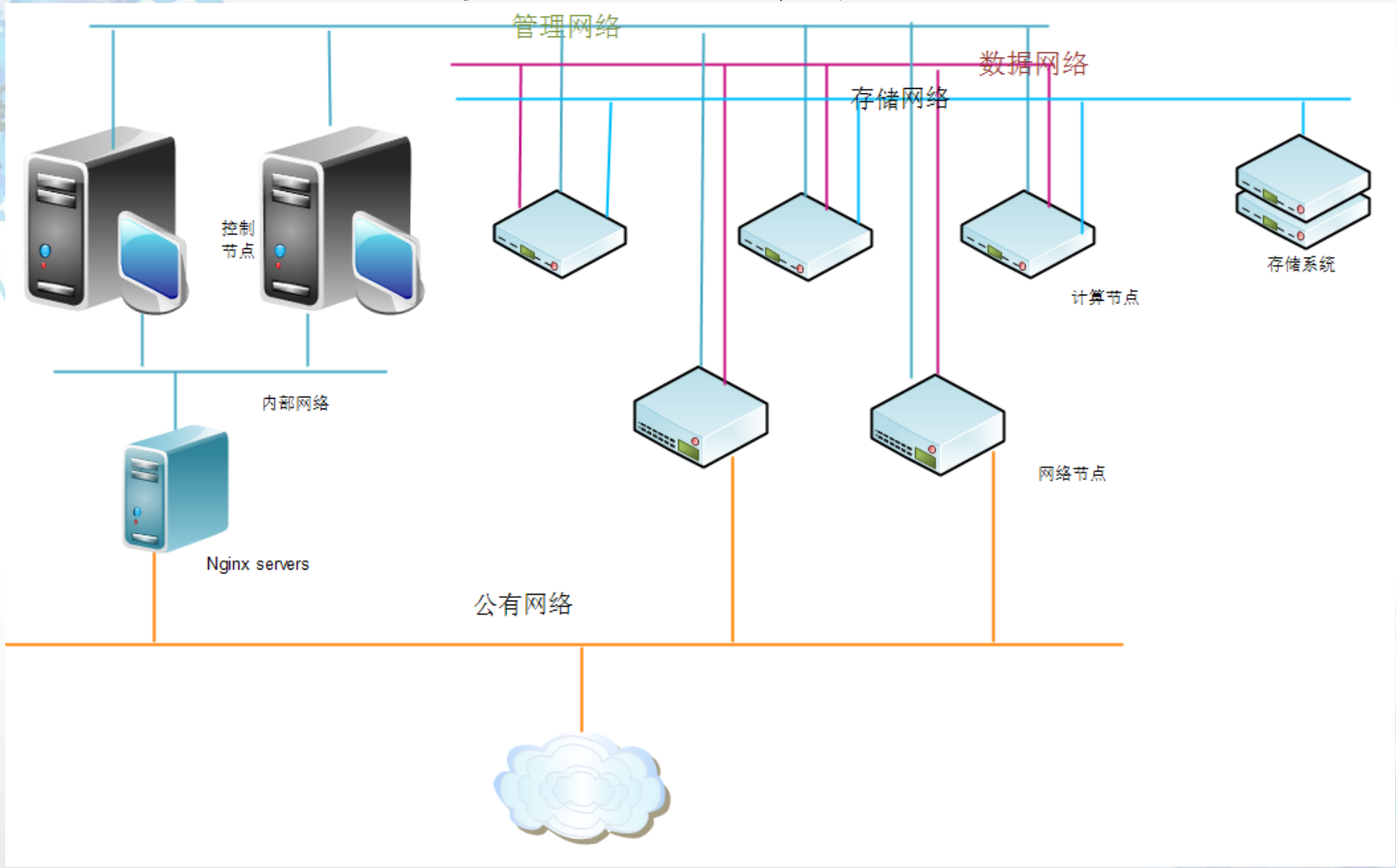
neutron

内容

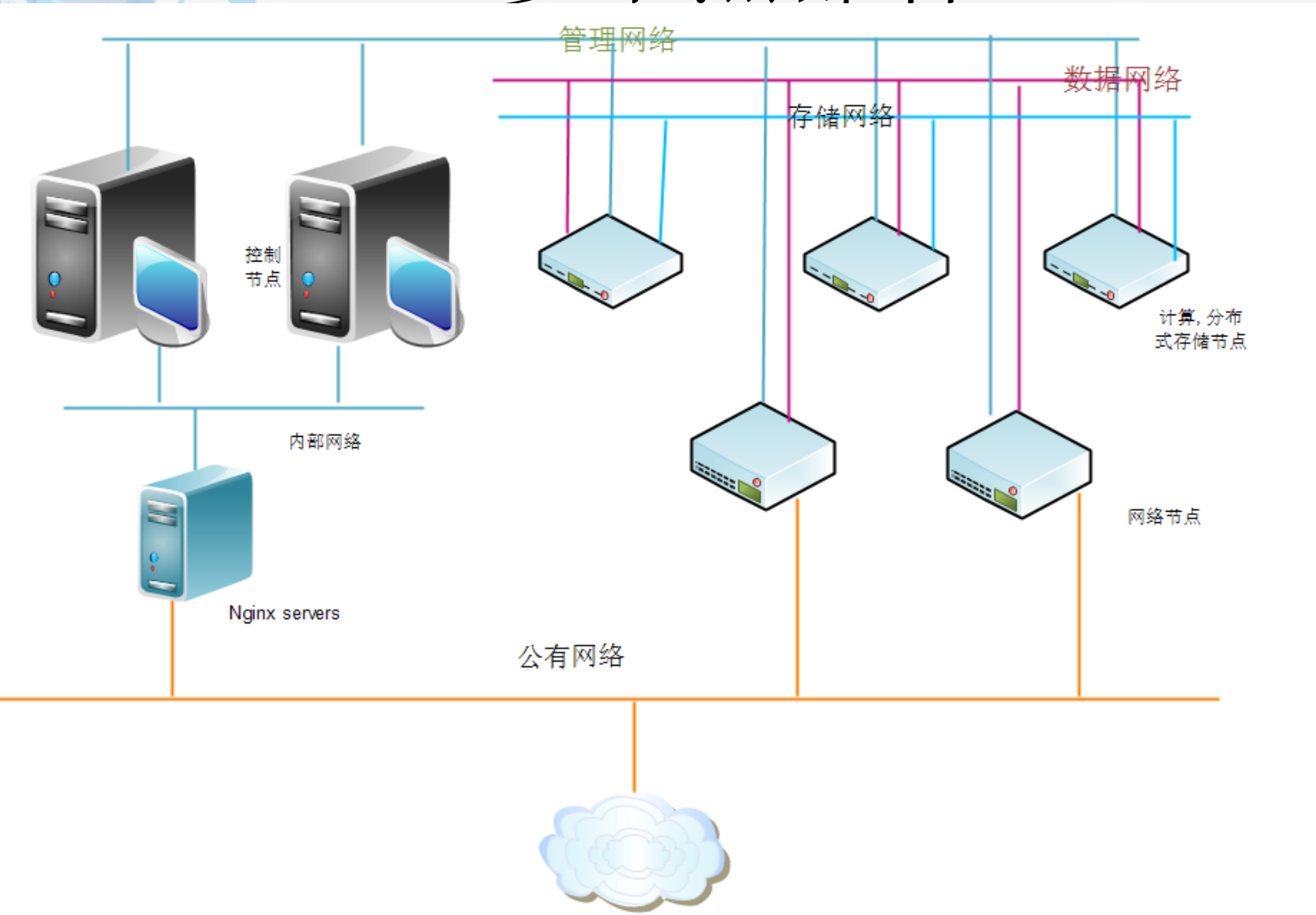
- 关于OpenStack
- 部署openstack
- Neutron的架构
- Neutron的核心模型
- Neutron网络技术


- OpenStack Command Line Tools (nova-client, swift-client, etc.)
- Cloud Management Tools (Rightscale, Enstratius, etc.)
- GUI tools (Cyberduck, iPhone client, etc.)
Internet
HTTP(S)
OpenStack Dashboard
Horizon
OpenStack Object API
OpenStack Identity API
OpenStack Image API
OpenStack Compute API
Amazon Web Services EC2 API
VNC/VMRC /Spice
OpenStack Block Storage API
OpenStack Network API
OpenStack Compute API / Admin API
swift-proxy
memcached
account
container
object
account DB
container DB
object store
OpenStack Object Store
glance-api
glance-registry
glance database
OpenStack Image Service
nova-api (OS, EC2, Metadata, Admin)
nova-console nova-*proxy
nova-compute
nova-cert/ objectstore
libvirt, XenAPI, etc.
hypervisor
nova database
Queue
nova-consoleauth
nova-conductor
nova-scheduler
OpenStack Compute
cinder-api
cinder-volume
cinder-backup
volume provider
Queue
cinder database
cinder-scheduler
OpenStack Block Storage
Neutron -server
Neutron agent(s)
Queue
Neutron plugin(s)
network provider
Neutron database
OpenStack Network Service
keystone (service & admin APIs)
token backend
catalog backend
policy backend
identity backend
OpenStack Identity Service
http://www.solinea.com
Solinea

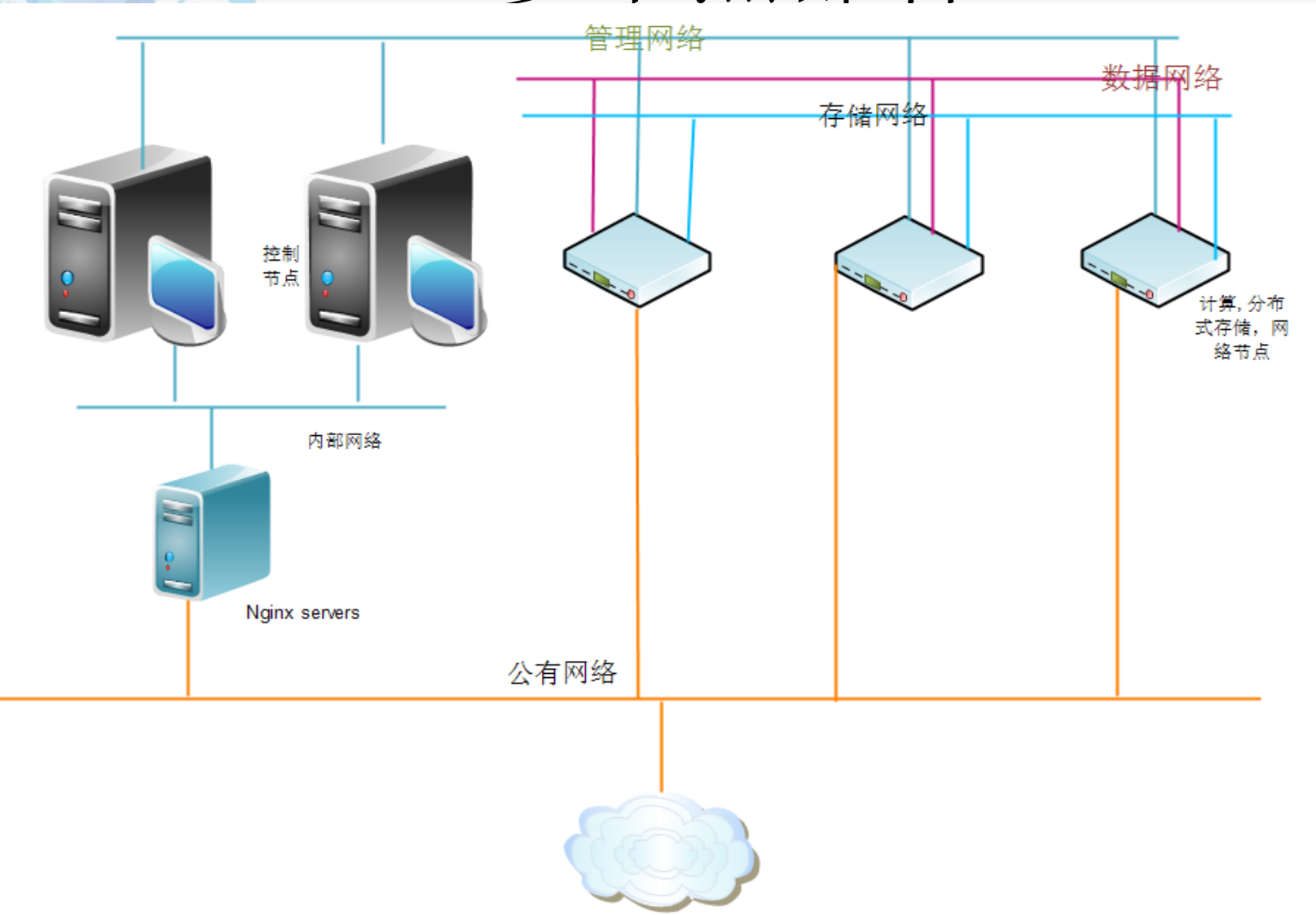
多节点部署一



多节点部署二



多节点部署三



Devstack sample control

```
[[local|localrc]]
disable_all_services
ENABLED_SERVICES=g-api,g-reg,key,n-api,n-cond,n-sch,n-novnc,n-cauth,
horizon,neutron,q-svc,rabbit,mysql
HOST_IP=192.168.13.57
Q_PLUGIN=ml2
DATABASE_PASSWORD=root
RABBIT_PASSWORD=password
SERVICE_TOKEN=token
SERVICE_PASSWORD=password
ADMIN_PASSWORD=admin
…
[[post-config|/etc/neutron/neutron.conf]]
[DEFAULT]
router_distributed = True
[[post-config|/etc/neutron/plugins/ml2/ml2_conf.ini]]
[ml2]
type_drivers = local,vxlan
tenant_network_types = vxlan
mechanism_drivers = openvswitch,l2population
[ml2_type_vxlan]
vni_ranges = 100:300
```

Devstack 计算，分布式网络节点

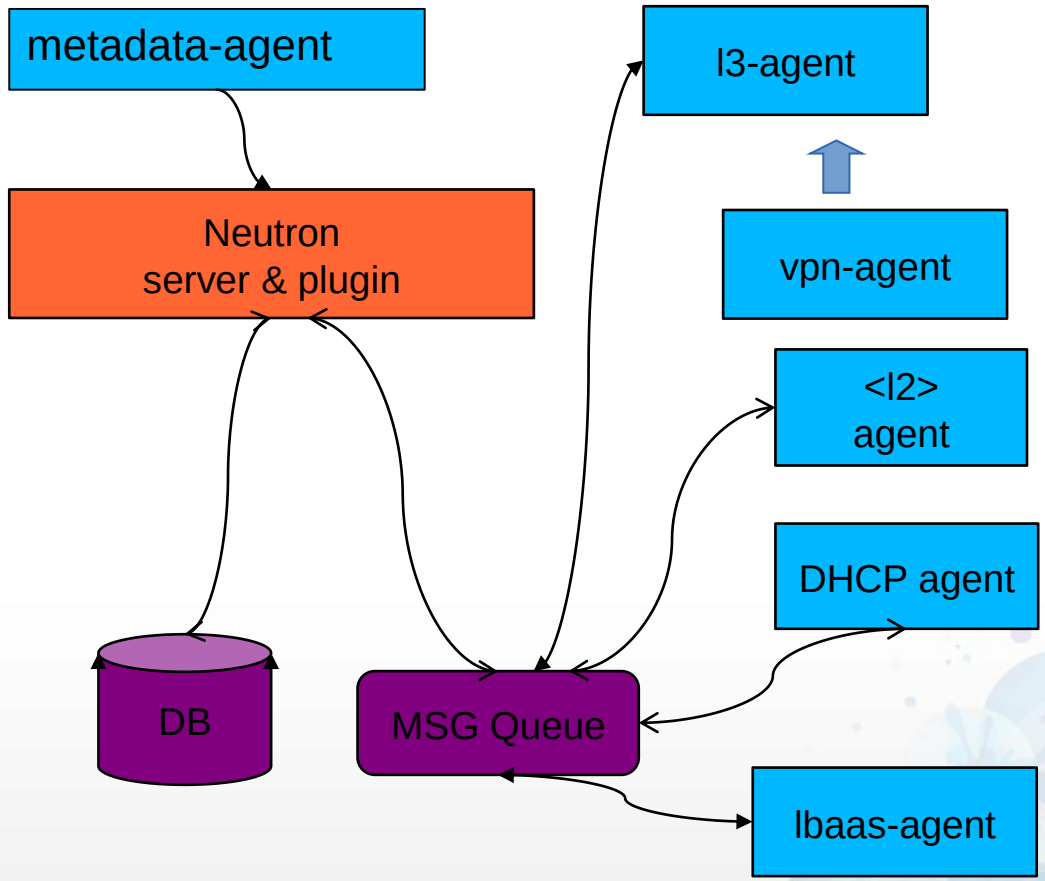
```
[[local|localrc]]
disable_all_services
ENABLED_SERVICES=n-cpu,rabbit,neutron,q-agt,q-l3
SERVICE_HOST=192.168.13.47
HOST_IP=192.168.13.49
MYSQL_HOST=$SERVICE_HOST
RABBIT_HOST=$SERVICE_HOST
…
[[post-config|/etc/neutron/neutron.conf]]
[[post-config|/etc/neutron/l3_agent.ini]]
[DEFAULT]
agent_mode = dvr
[[post-config|/etc/neutron/plugins/ml2/ml2_conf.ini]]
[ovs]
local_ip = 192.168.13.49
[agent]
tunnel_types = vxlan,gre
enable_distributed_routing=True
arp_responder = True
l2_population = True
```

内容

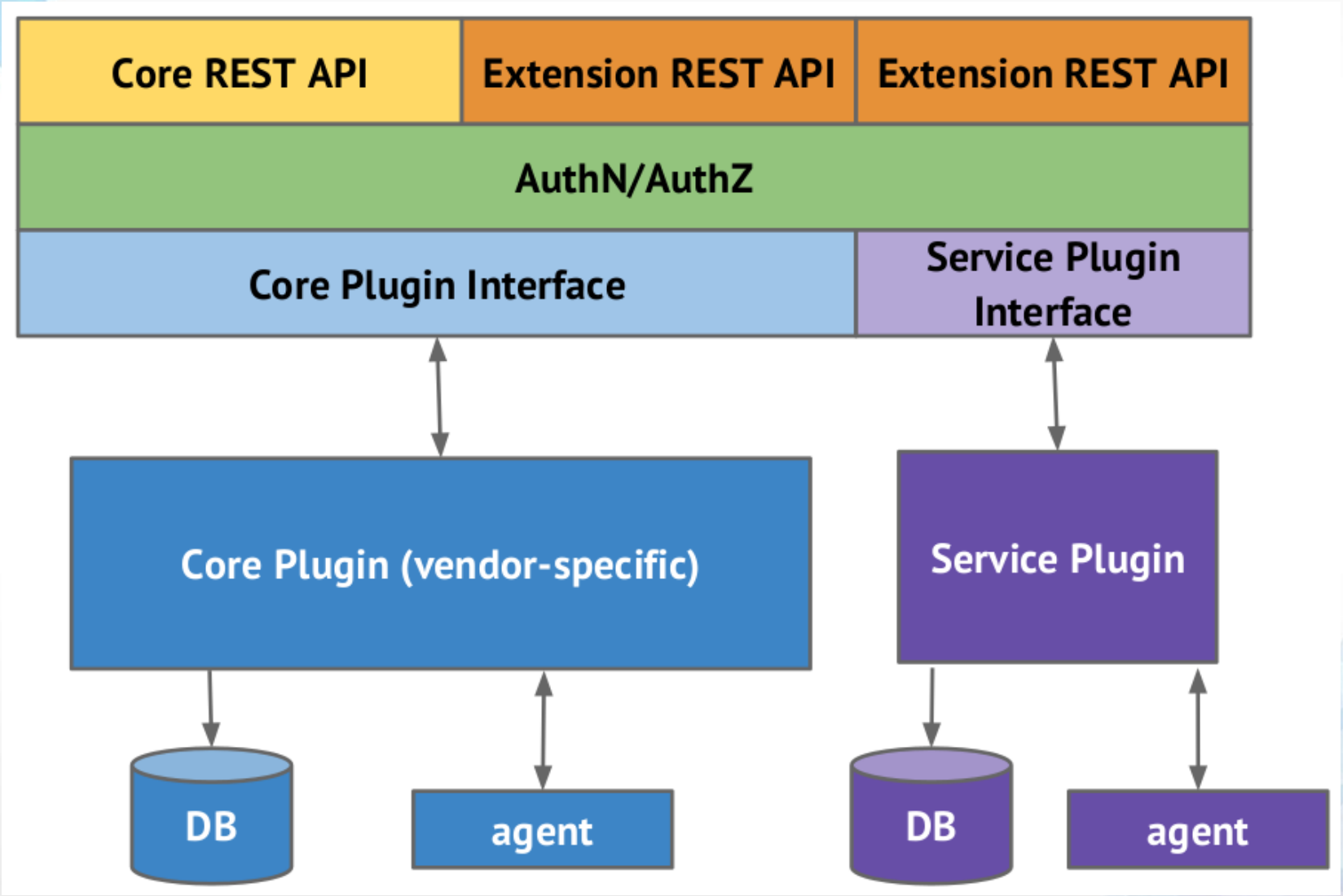
- 关于OpenStack
- 部署openstack
- Neutron的架构
- Neutron的核心模型
- Neutron网络技术

OpenStack neutron

- virtual network management
- L3 router service
- VPN service
- Firewall service
- Load balance service



neutron server的层次架构



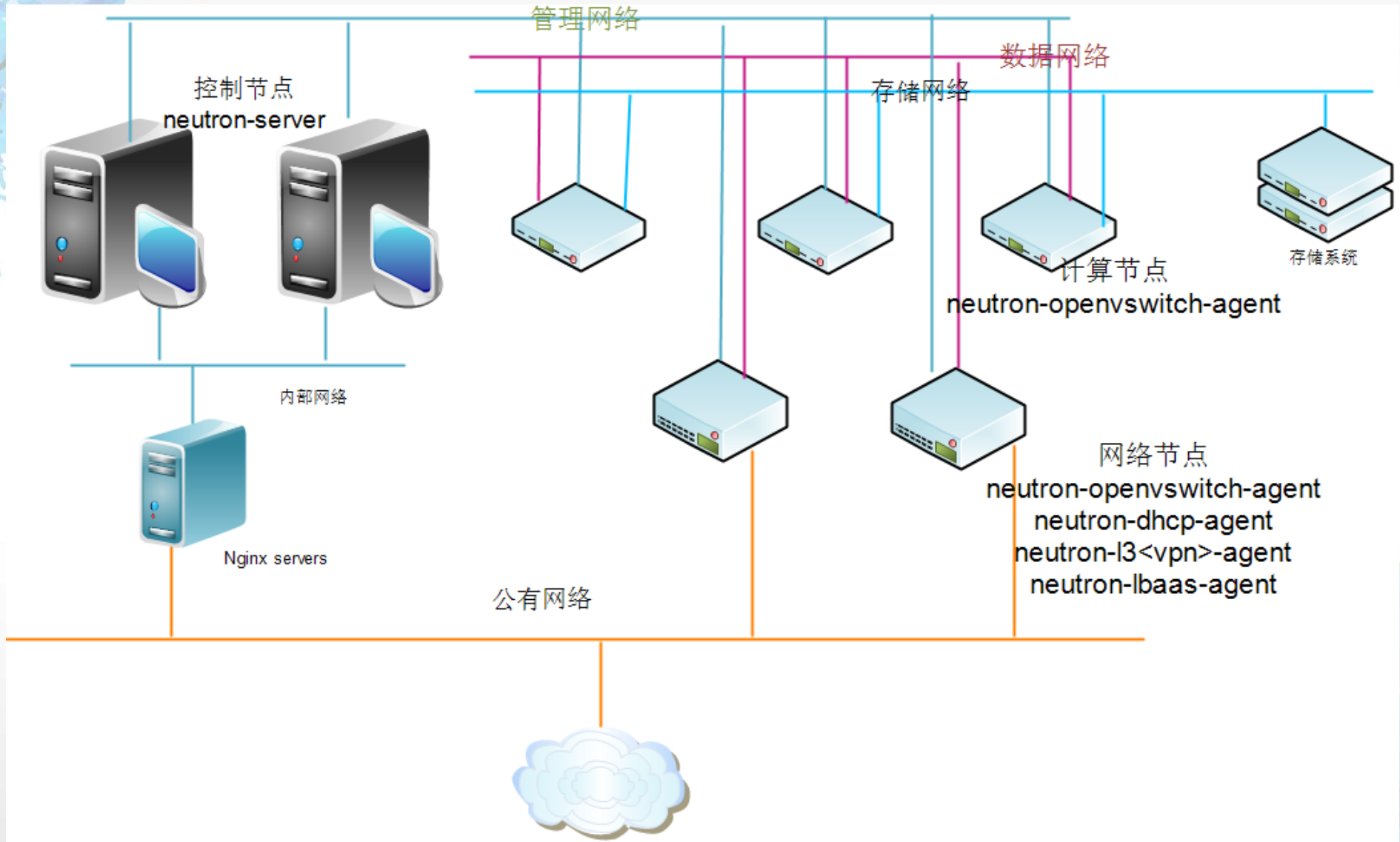
http://www.slideshare.net/gongys2004/inside-neutron-2

Neutron plugins

	Core plugins by neutron community
1	Open vSwitch Plugin (deprecated)
2	Linux Bridge Plugin(deprecated)
3	Modular Layer 2 Plugin (linuxbridge, ovs bridge, openddaylight ...
4	Ryu OpenFlow Controller Plugin (NTT)
5	Big Switch Controller Plugin
6	Cisco UCS/Nexus Plugin
7	NEC OpenFlow Plugin
8	Nicira Network Virtualization Platform (NVP) Plugin
9	Cloudbase Hyper-V Plugin
10	MidoNet Plugin
11	Brocade Neutron Plugin
12	…

Service plugins
Load Balancer Service Plugin
L3 router Service Plugin
Firewall Service Plugin
VPN Service Plugin

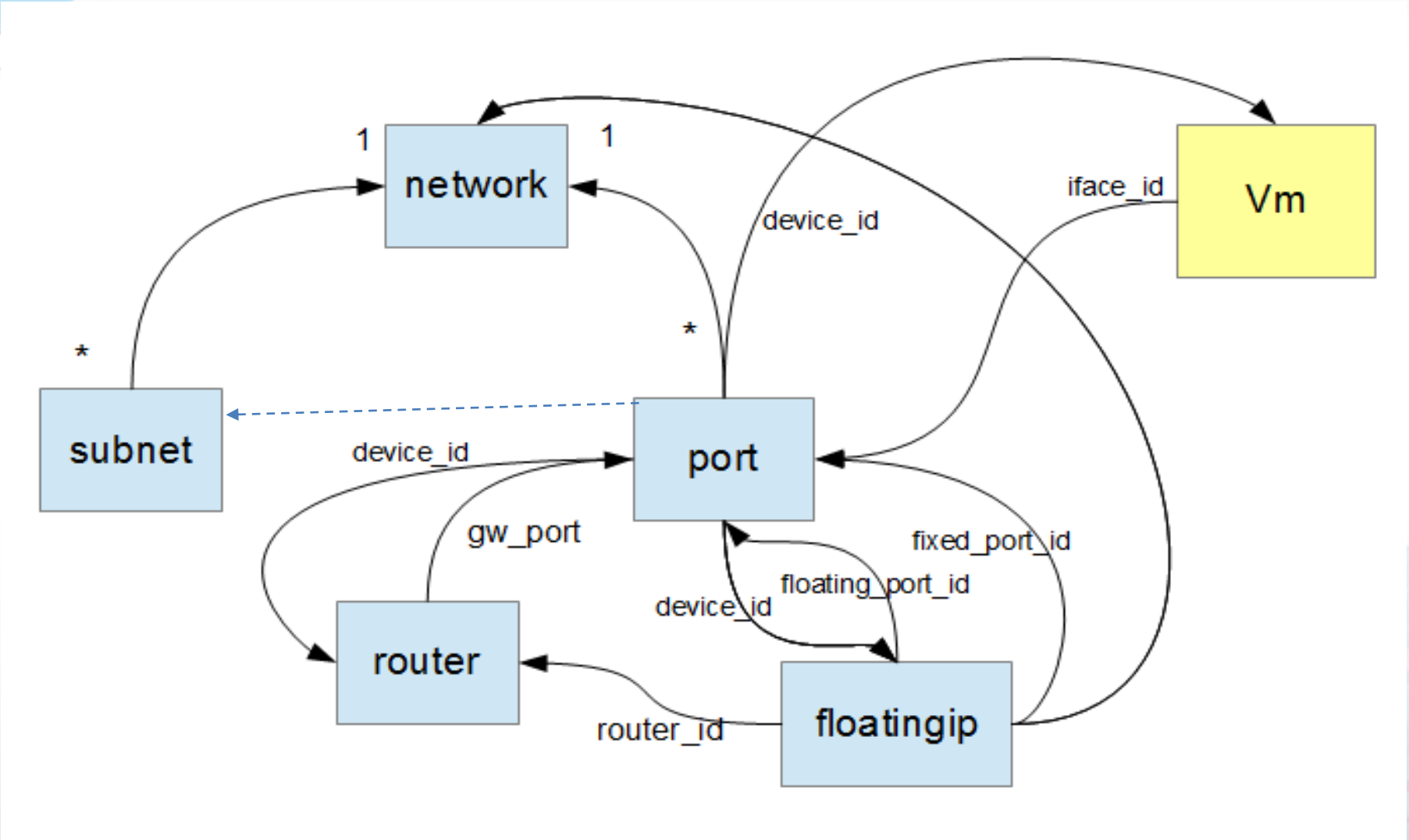
组件部署


管理网络
数据网络
存储网络
控制节点
neutron-server
内部网络
Nginx servers
计算节点
neutron-openvswitch-agent
存储系统
网络节点
neutron-openvswitch-agent
neutron-dhcp-agent
neutron-l3<vpn>-agent
neutron-lbaas-agent
公有网络

内容

- 关于OpenStack
- 部署openstack
- Neutron的架构
- Neutron的核心模型
- Neutron网络技术

Neutron Core models



show neutron features on ustack

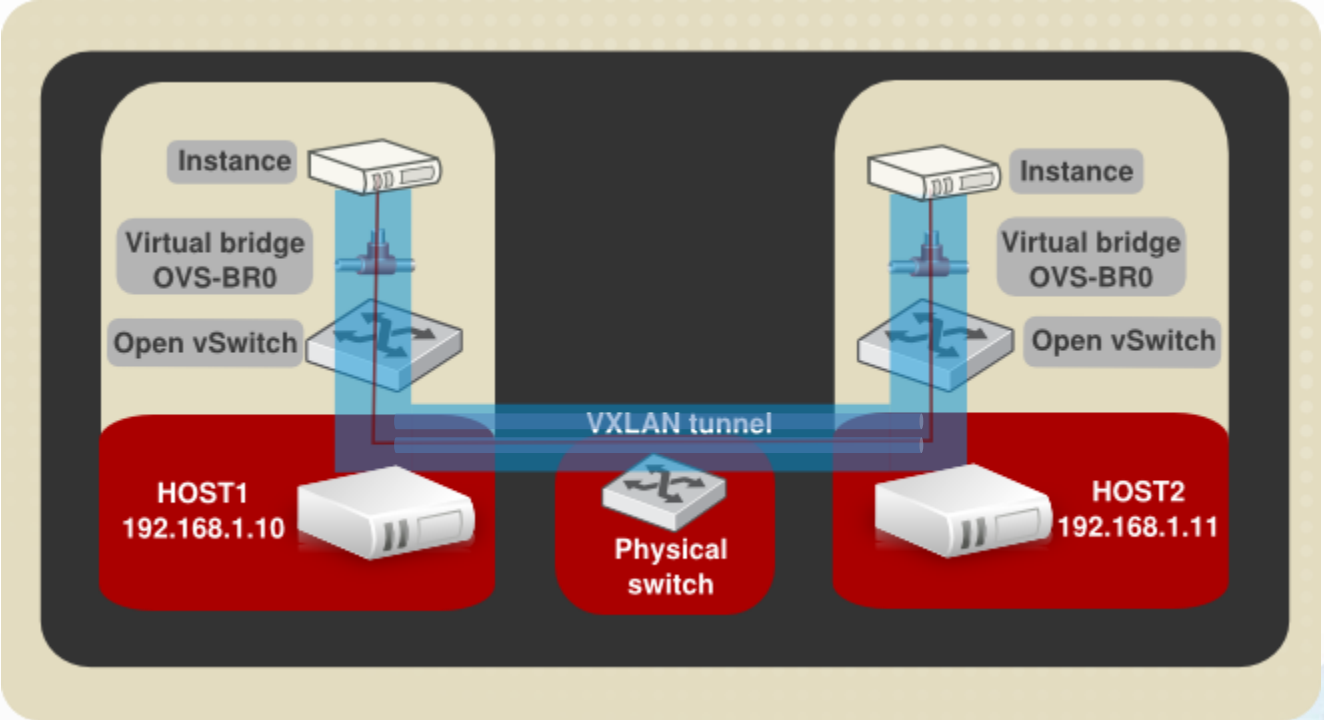
- www.ustack.com
  - floating IP on router
  - HA router
  - port forwarding
  - PPTP VPN
  - multiple subnets on floatingip
  - hundreds of bugs

内容

- 关于OpenStack
- 部署openstack
- Neutron的架构
- Neutron的核心模型
- Neutron网络技术

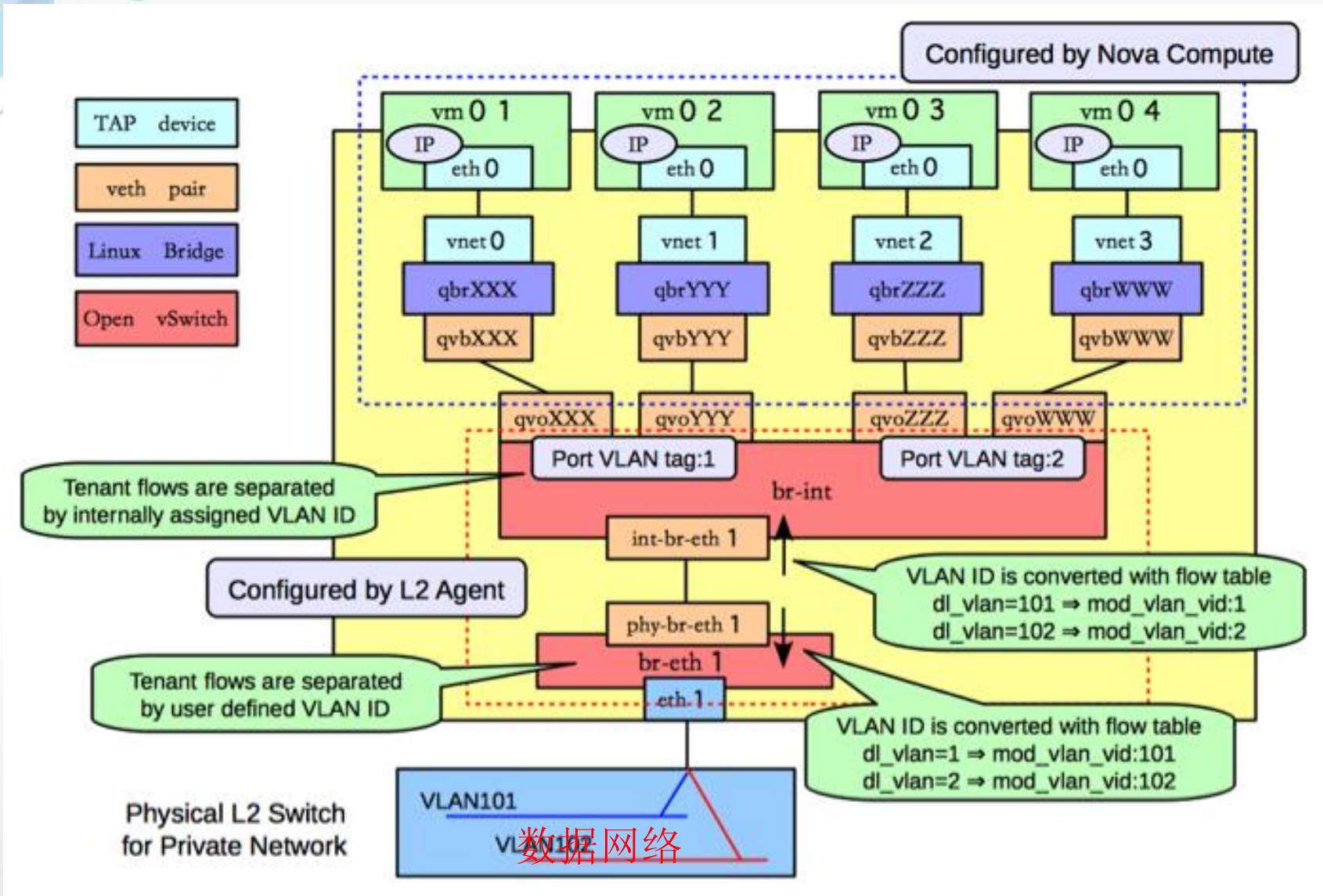
网络隔离

1. Ovs bridge is used to connect VMs to physical eth nic
2. VXLAN tunnel is used to isolate Networks. Vlan or GRE can also be used



src：http://docs.openstack.org/admin-guide-cloud/content/figures/14/a/a/common/figures/OVStunneling.png

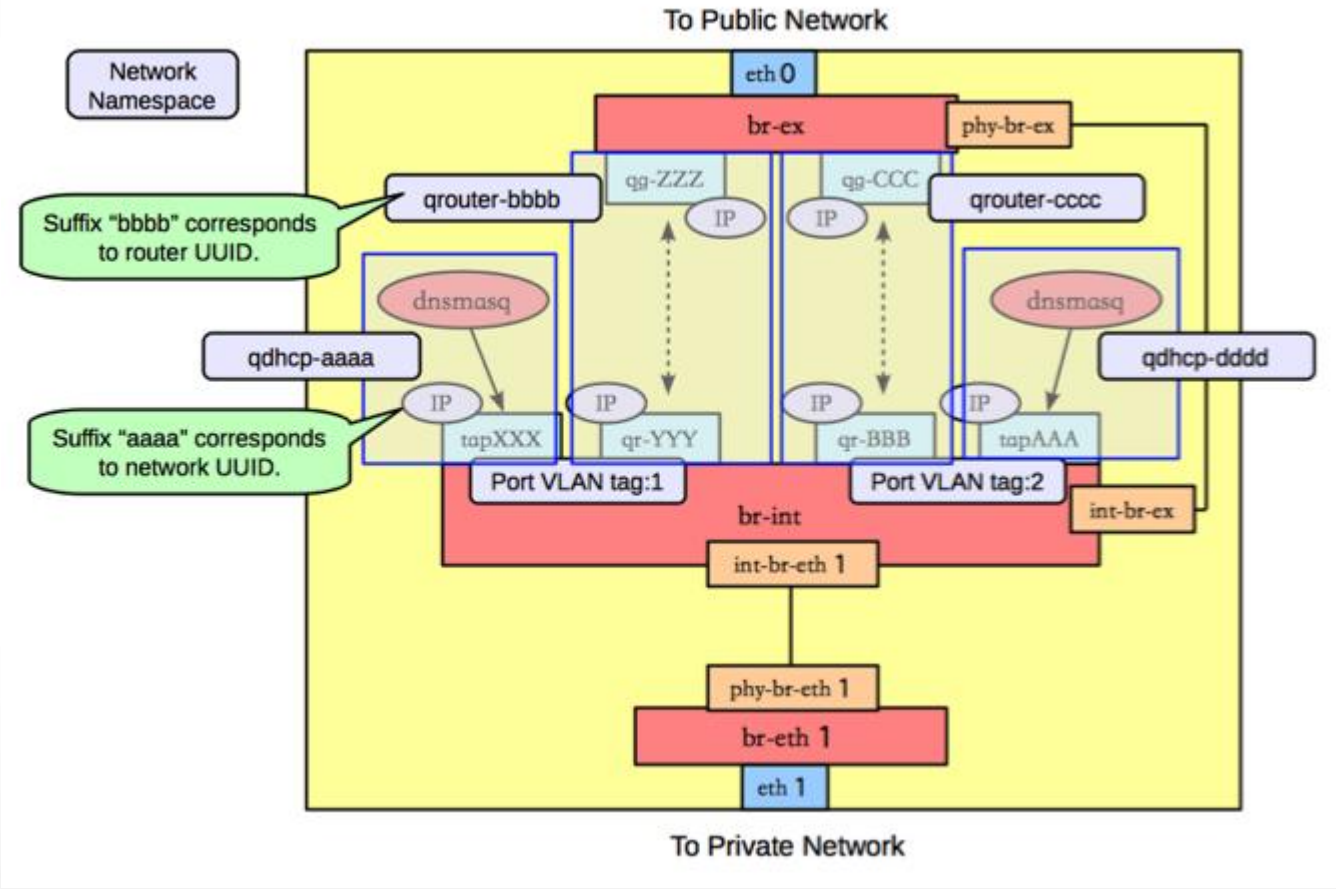
计算节点



http://docs.openstack.org/admin-guide-cloud/content/under_the_hood_openvswitch.html#under_the_hood_openvswitch_scenario1

网络节点

公有网络



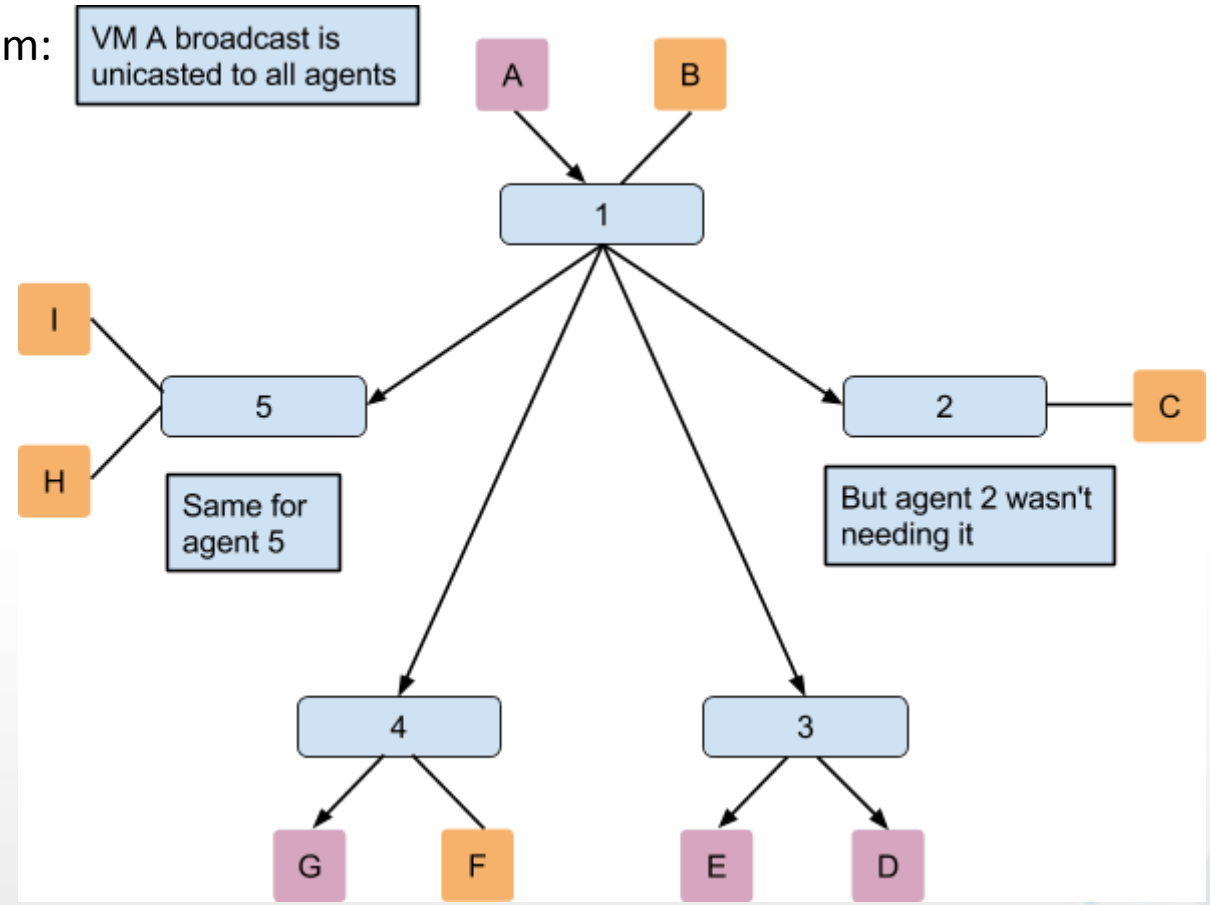
数据网络

http://docs.openstack.org/admin-guide-cloud/content/under_the_hood_openvswitch.html#under_the_hood_openvswitch_scenario1

L2pop和arpresponder

arp-request:who-has 10.1.23.7?

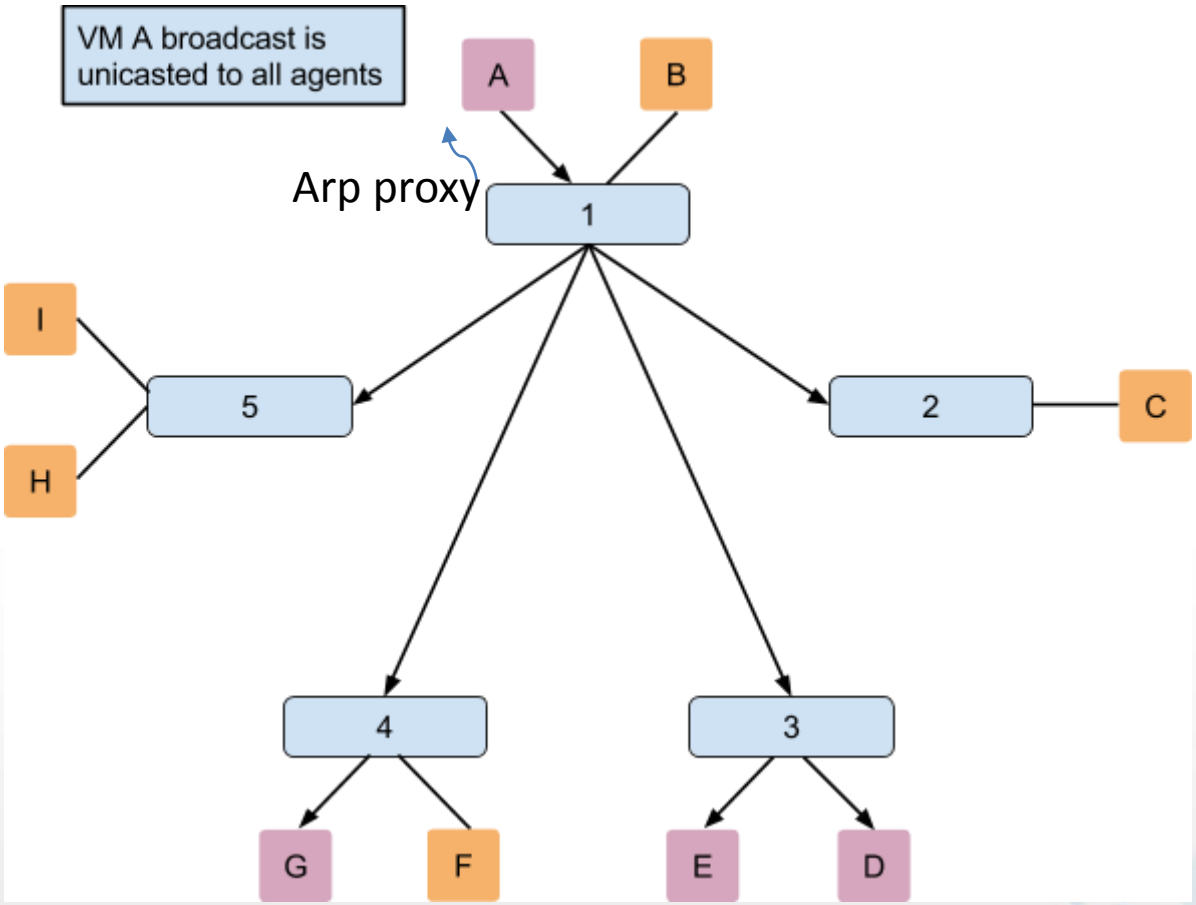
Problem:



fa:16:3e:73:47:84

10.1.23.7

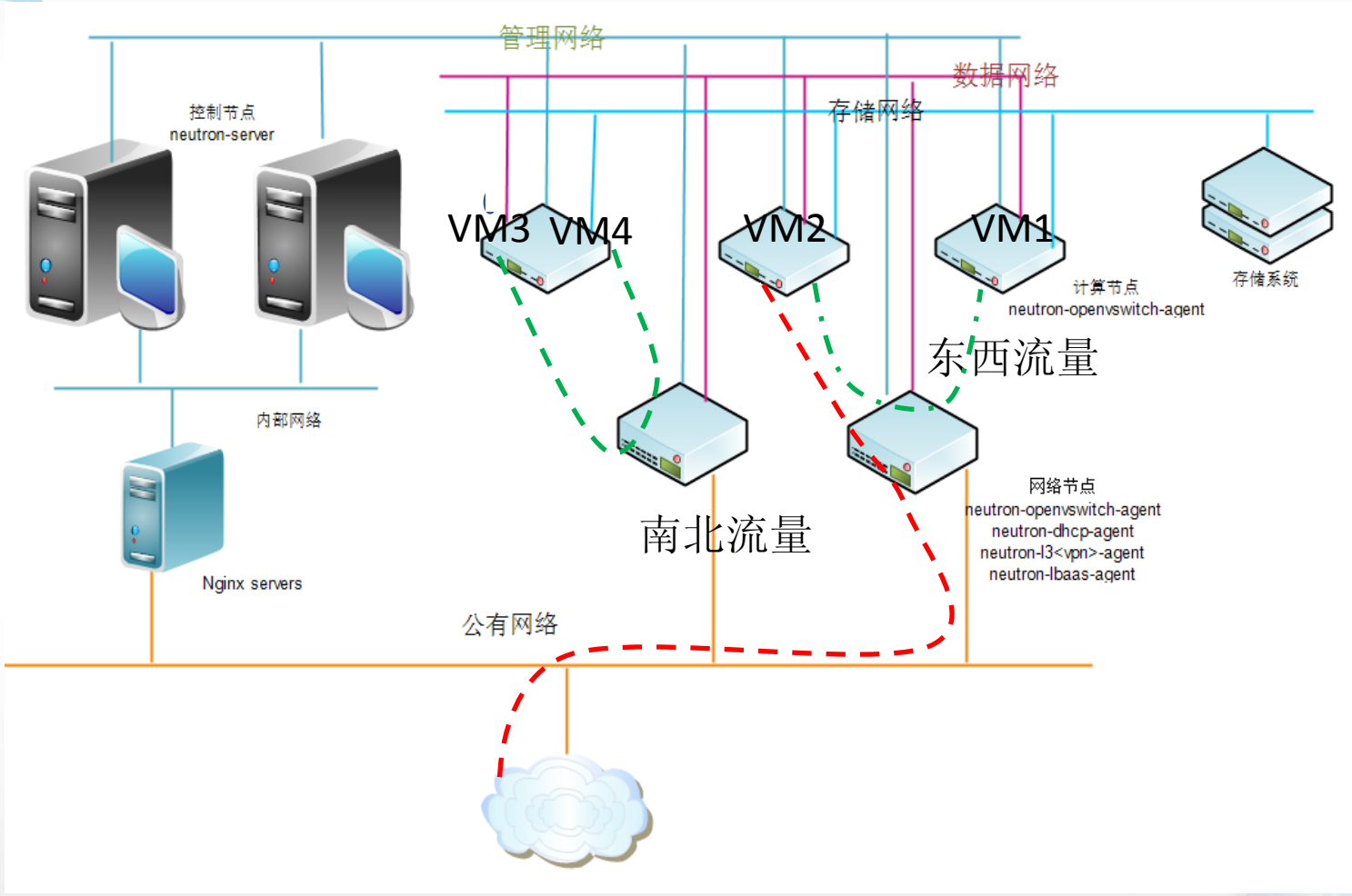
ARP responder avoids the broadcaset



FDB entry in the form of ovs flow

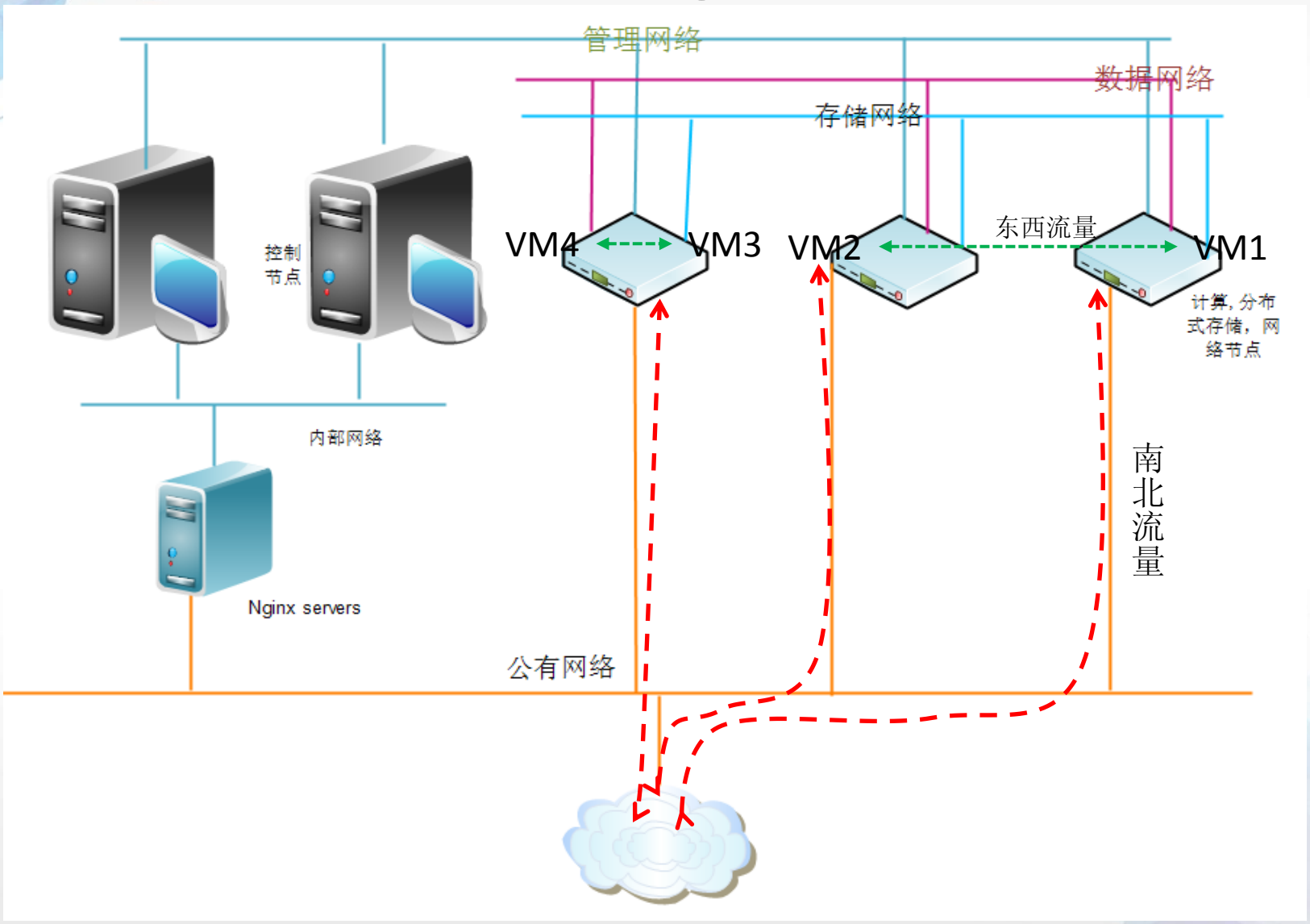
- cookie=0x0, duration=13.576s, table=20, n_packets=0, n_bytes=0, idle_age=20, priority=2,dl_vlan=1,dl_dst=fa:16:3e:73:47:84 actions=strip_vlan,set_tunnel:0x66,output:2
- cookie=0x0, duration=13.763s, table=21, n_packets=0, n_bytes=0, idle_age=20, priority=1,arp,dl_vlan=1,arp_tpa=10.1.23.7 actions=move:NXM_OF_ETH_SRC[]->NXM_OF_ETH_DST[],mod_dl_src:fa:16:3e:73:47:84,load:0x2->NXM_OF_ARP_OP[],move:NXM_NX_ARP_SHA[]->NXM_NX_ARP_THA[],move:NXM_OF_ARP_SPA[]->NXM_OF_ARP_TPA[],load:0xfa163e734784->NXM_NX_ARP_SHA[],load:0xa011707->NXM_OF_ARP_SPA[],IN_PORT

DVR

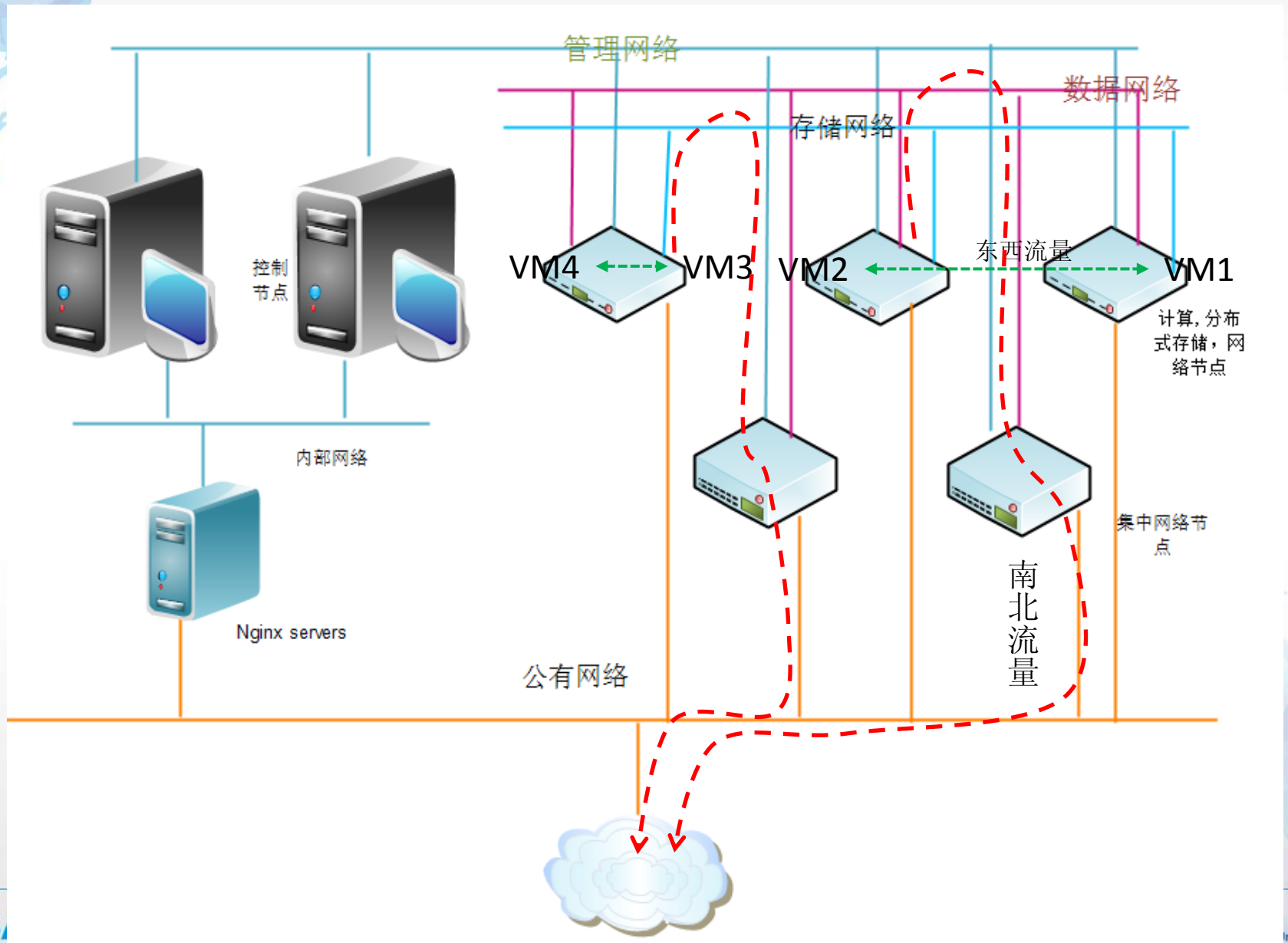


1. 跨子网的东西流量需要经过集中的网络节点，甚至在一台节点上
2. 虚拟机的南北流量需要经过集中的网络节点

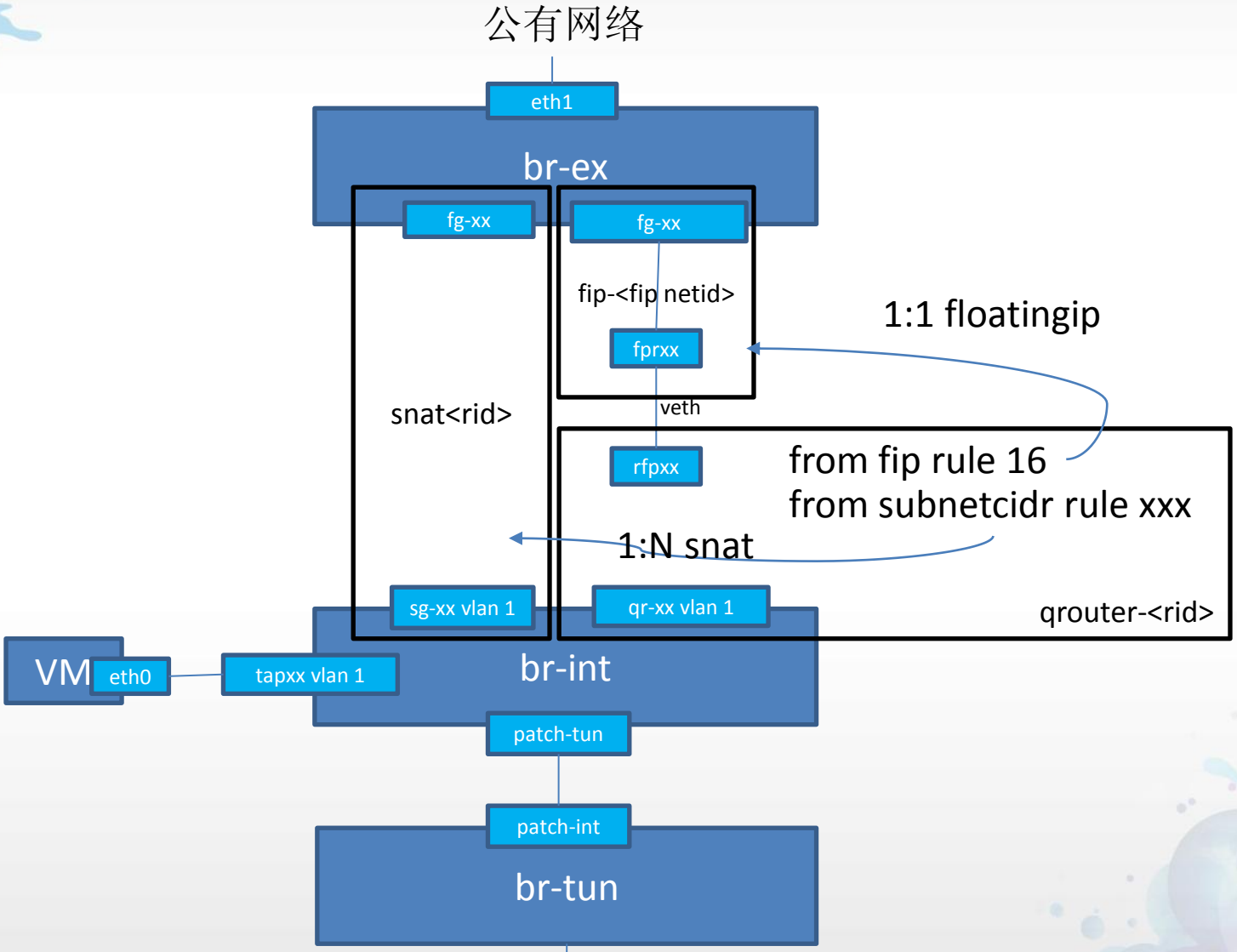
DVR Floating IP 1:1 NAT



DVR 1:N SNAT


管理网络
数据网络
存储网络
控制节点
VM4
VM3
VM2
东西流量
VM1
计算，分布式存储，网络节点
内部网络
集中网络节点
南北流量
Nginx servers
公有网络

DVR floatingip traffic flows



NFV

- 优化网络转发能力数据平面上：

1. Packets per second target -> either SR-IOV or an accelerated DPDK-like data plane:
2. "SR-IOV Networking Support" (https://blueprints.launchpad.net/nova/+spec/pci-passthrough-sriov)
3. "userspace vhost in ovs vif bindings" (https://blueprints.launchpad.net/nova/+spec/libvirt-ovs-use-usvhost)
4. "Snabb NFV driver" (https://blueprints.launchpad.net/neutron/+spec/snabb-nfv-mech-driver)
5. "VIF_VHOSTUSER" (https://blueprints.launchpad.net/nova/+spec/vif-vhostuser)

https://wiki.openstack.org/wiki/Meetings/NFV

参考资料

- OpenStack管理文档 http://docs.openstack.org/admin-guide-cloud/content/
- Neutron wiki https://wiki.openstack.org/wiki/Neutron
- inside the Neutron http://sdrv.ms/YdnQYS
- inside neutron 2 http://www.slideshare.net/gongys2004/inside-neutron-2
- UnitedStack blog http://www.ustack.com/blog/

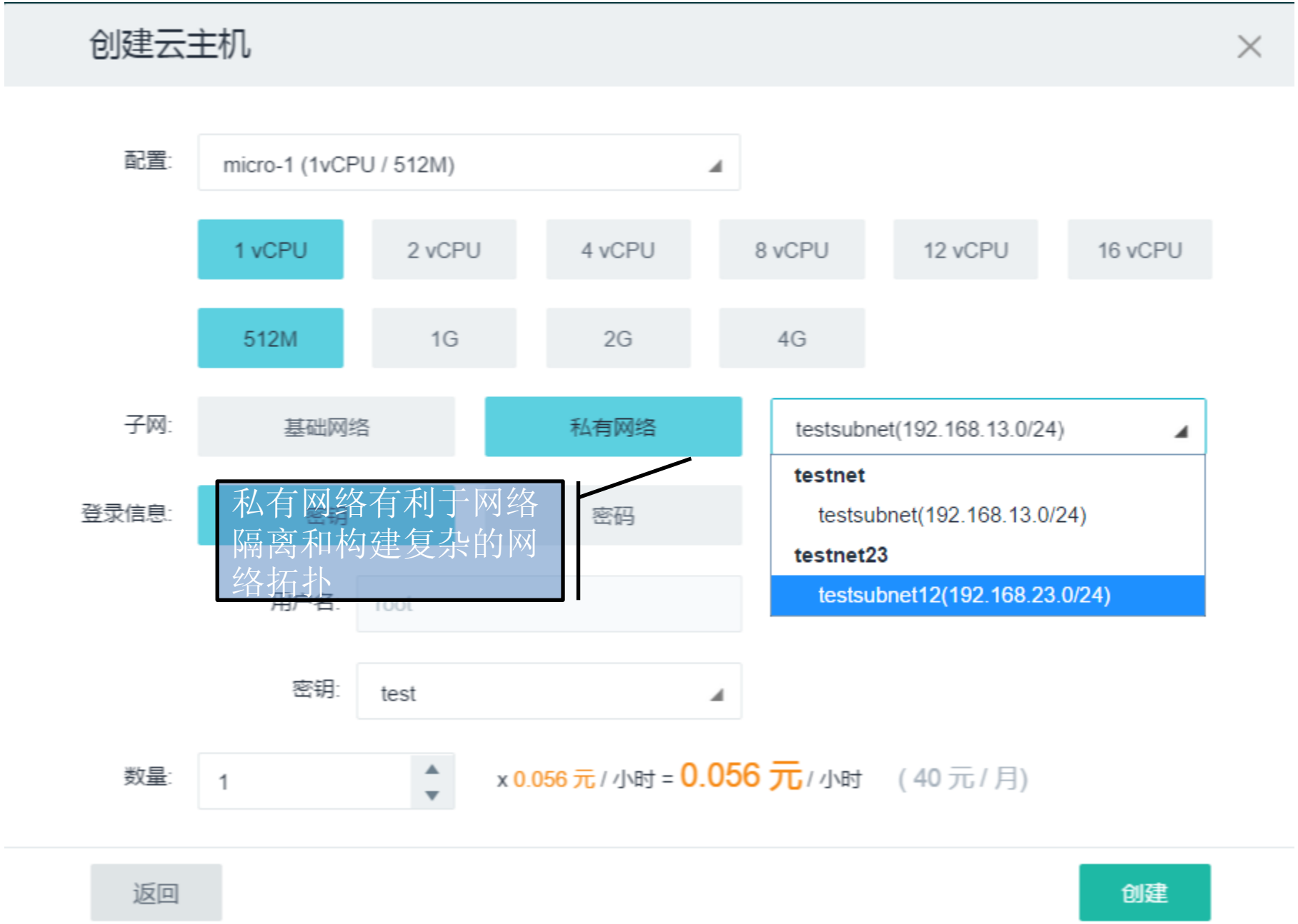
Q&A

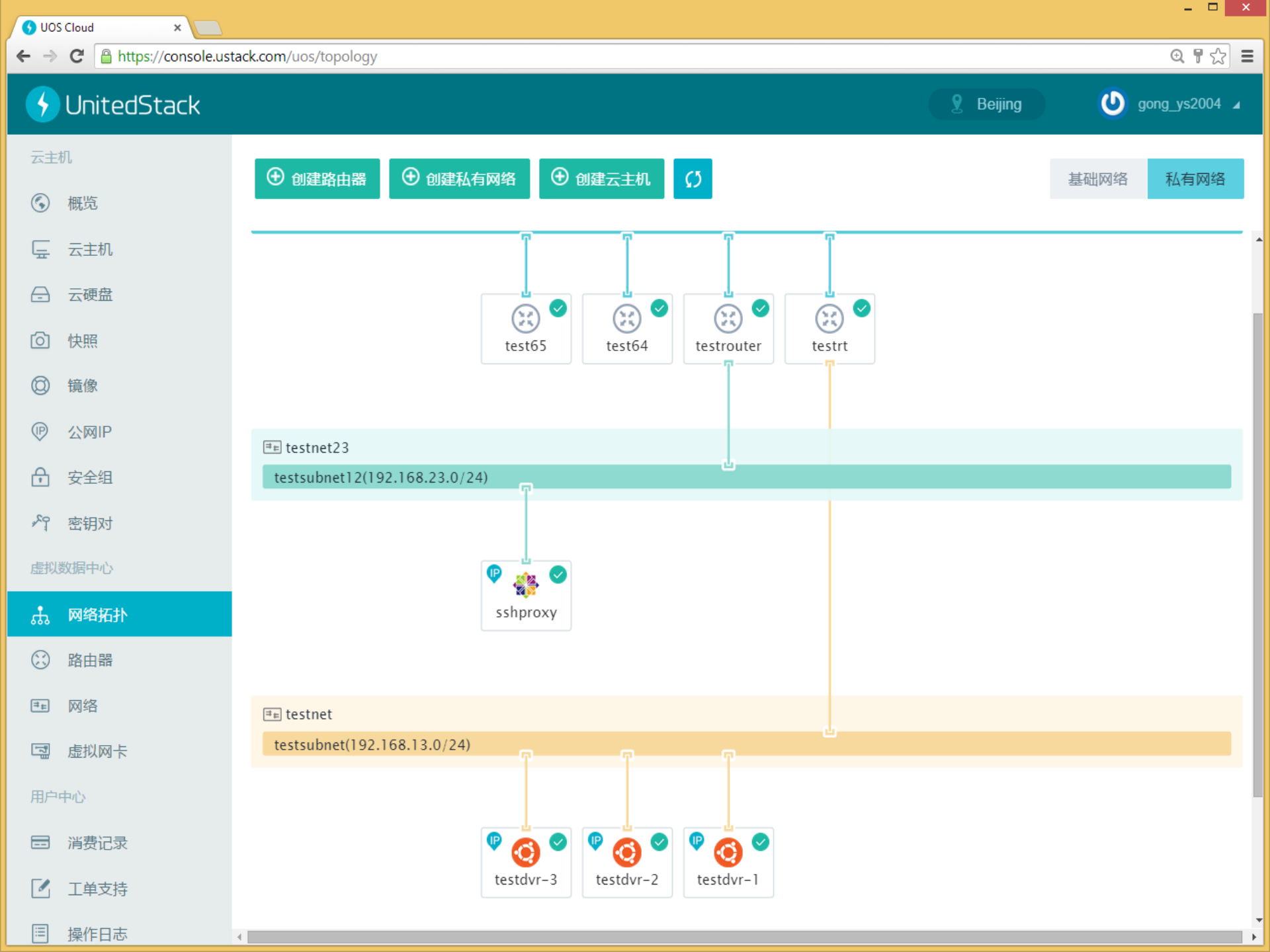
THANKS

SequeMedia 盛拓传媒 | IT168 www.it168.com | ChinaUnix | ITPUB

backup


创建云主机
配置:
micro-1 (1vCPU / 512M)
1 vCPU
2 vCPU
4 vCPU
8 vCPU
12 vCPU
16 vCPU
512M
1G
2G
4G
子网:
基础网络
私有网络
登录信息:
密钥
密码
用户名:
root
密钥:
test
数量:
1
x 0.056 元 / 小时 = 0.056 元 / 小时
(40 元 / 月)
返回
创建
基础网络实现物理交换机交换，但是在同一个二层下。主要为了入门用户的简单需求


创建云主机
配置:
micro-1 (1vCPU / 512M)
1 vCPU
2 vCPU
4 vCPU
8 vCPU
12 vCPU
16 vCPU
512M
1G
2G
4G
子网:
基础网络
私有网络
testsubnet(192.168.13.0/24)
testnet
testsubnet(192.168.13.0/24)
testnet23
testsubnet12(192.168.23.0/24)
登录信息:
密码
私有网络有利于网络隔离和构建复杂的网络拓扑
密钥:
test
数量:
1
x 0.056 元 / 小时 = 0.056 元 / 小时
(40 元 / 月)
返回
创建


UOS Cloud
https://console.ustack.com/uos/topology
UnitedStack
Beijing
gong_ys2004
云主机
概览
云主机
云硬盘
快照
镜像
公网IP
安全组
密钥对
虚拟数据中心
网络拓扑
路由器
网络
虚拟网卡
用户中心
消费记录
工单支持
操作日志
创建路由器
创建私有网络
创建云主机
基础网络
私有网络
test65
test64
testrouter
testrt
testnet23
testsubnet12(192.168.23.0/24)
sshproxy
testnet
testsubnet(192.168.13.0/24)
testdvr-3
testdvr-2
testdvr-1


云主机
概览
云主机
云硬盘
快照
镜像
公网IP
安全组
密钥对
虚拟数据中心
网络拓扑
路由器
网络
虚拟网卡
私有网络
子网
创建私有网络
名称
test
testnet23
testnet
创建私有网络
私有网络是一个隔离的二层网络广播域，在这个二层网络中，可以创建多个三层网络，即子网。在UOS VDC中，私有网络主要作用作为子网容器。
* 私有网络名称:
创建子网
* 子网名称:
网络地址:
192 . 168 . 0 . 0 / 24
更多选项
创建