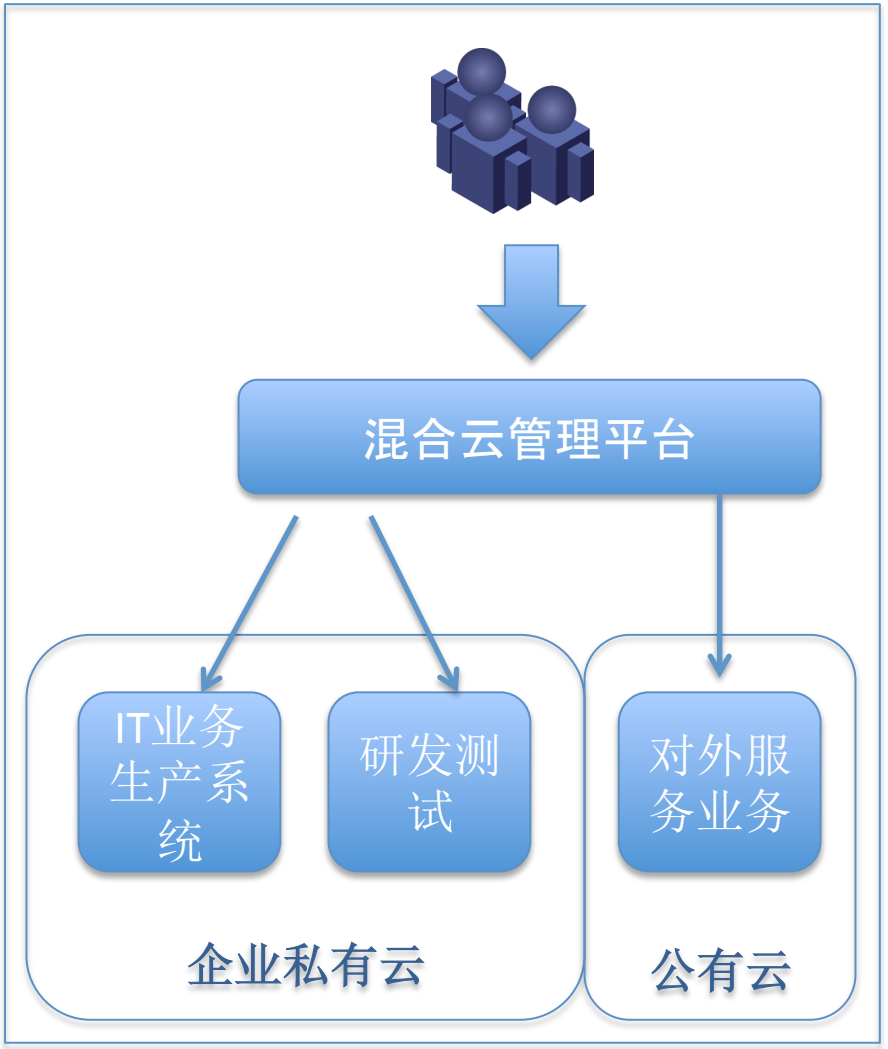
CMRI 混合云建设部署实践

目录

- 混合云项目简介
- OpenStack简介
- 实施情况
  - 物理机配置
  - 私有云架构
  - 数据网络方案
  - 存储方案选择
  - 安装部署
  - 监控管理
  - 实际问题：流量控制、安全策略等.

混合云项目目标

- 构建统一的云平台，整合分散的测试、办公、对外服务的IT资源。
- 构建混合云，连通内部私有云平台和外部公有云，充分发挥云平台的优势，支持业务的快速部署上线，弹性扩展。
- 依托统一的运营运维平台，实践云计算运营运维的管理要求，提高资源使用和管理效率。



OpenStack简介

- 开放源码的云平台管理项目
  - 组件化，功能组件可根据实际部署环境选择部署安装
  - 社区活跃，成员支持度高，软件本身发展稳定

关键组件:

Nova

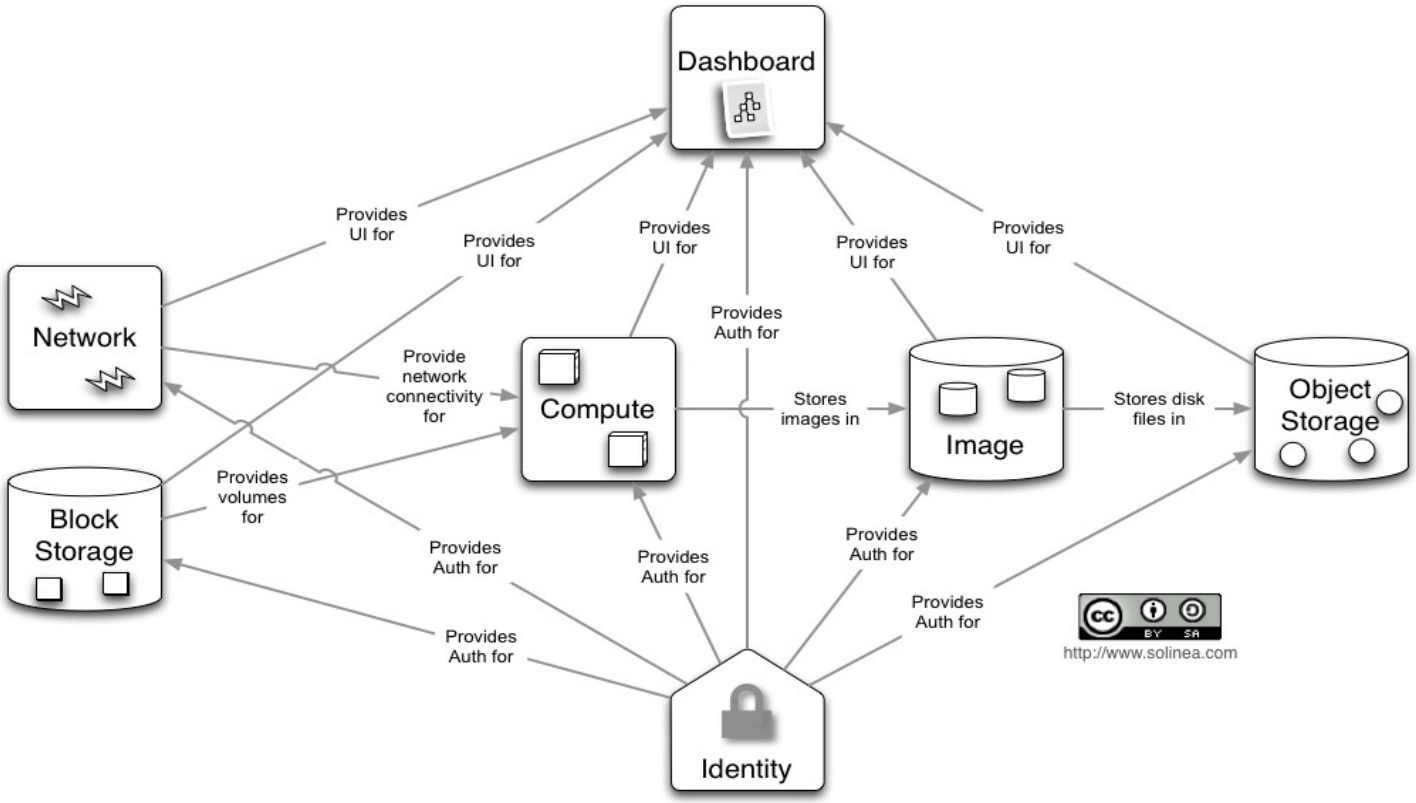
Neutron

Horizon

Glance

Keystone

Swift



目录

- 混合云项目简介
- OpenStack简介
- 实施情况
  - 物理机配置
  - 私有云架构
  - 数据网络方案
  - 存储方案选择
  - 安装部署
  - 监控管理
  - 实际问题：流量控制、安全策略等.

物理环境

计算/控制节点 ：

- 两路 E5-2650 v2 @ 2.60GHz
- 网络 ：4* 1Gb Ethernet ; 2* 10Gb Ethernet

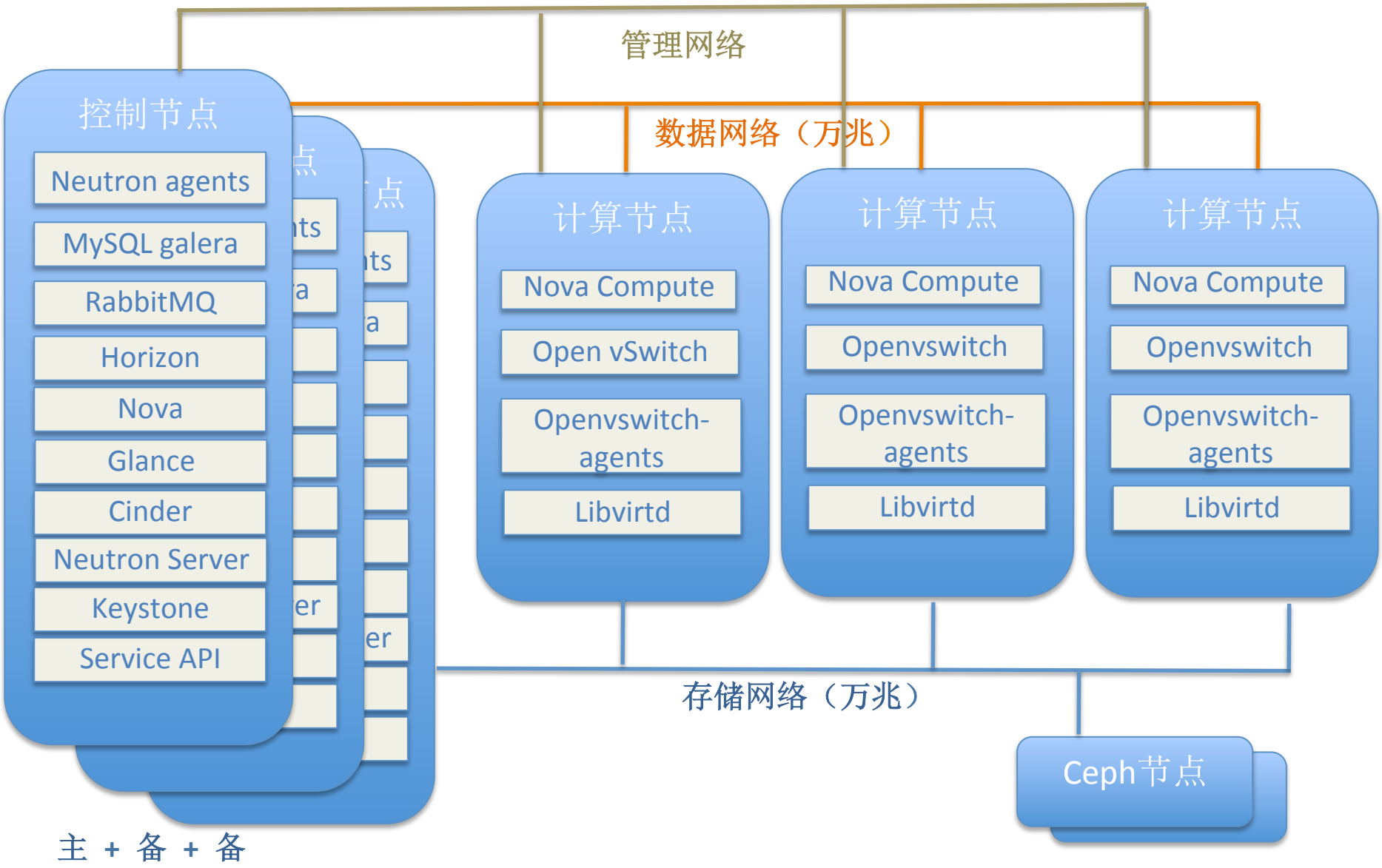
存储节点：

- 两路 E5-2640 v2 @ 2.0GHz
- 硬盘 ：8*3T SATA
- 网络 : 3 * 1Gb Ethernet ; 2* 10Gb Ethernet

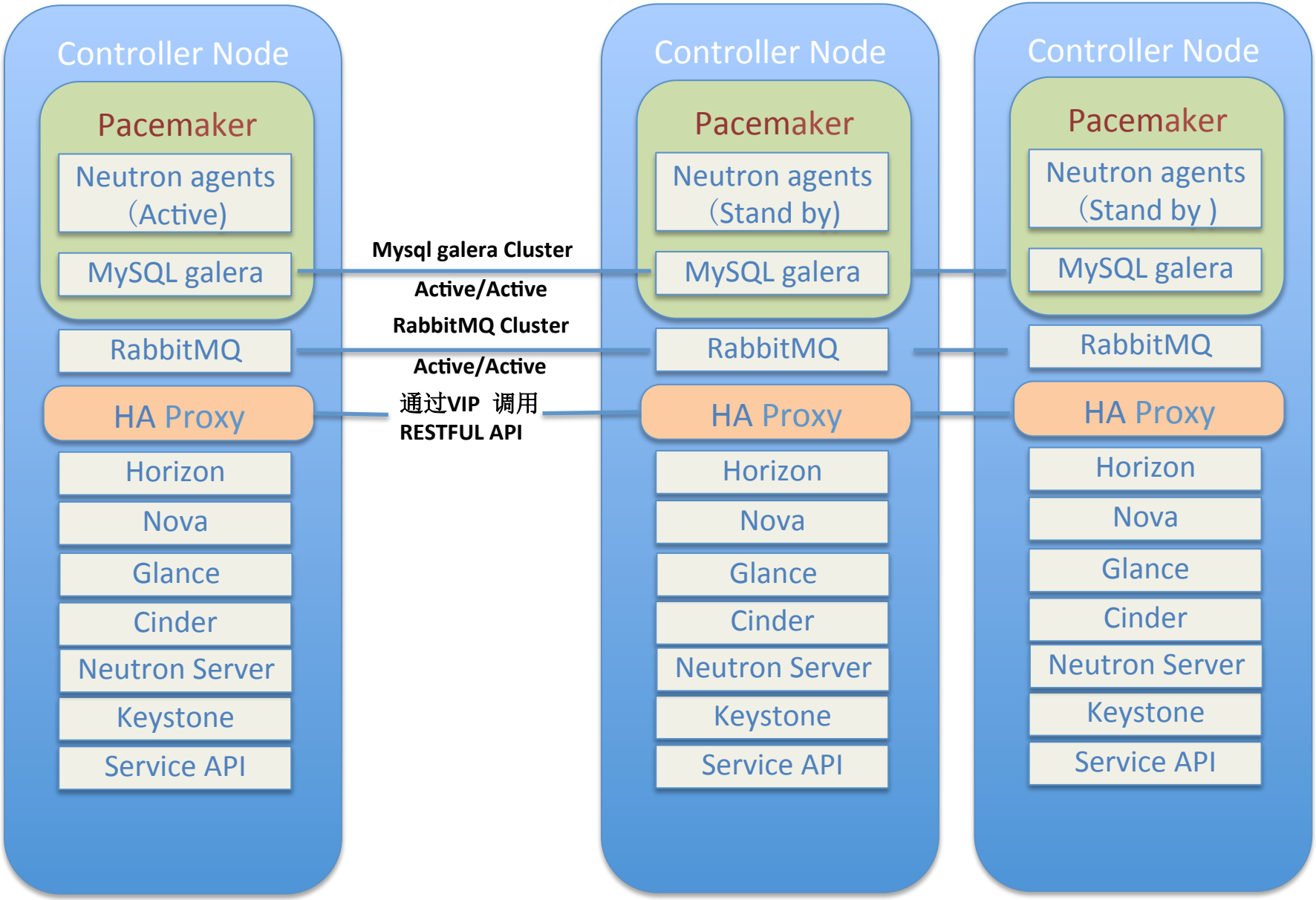
交换网络：

节点间交换使用万兆网络，网络出口接入到已有环境千兆交换出口

基础设施架构


管理网络
数据网络（万兆）
控制节点
Neutron agents
MySQL galera
RabbitMQ
Horizon
Nova
Glance
Cinder
Neutron Server
Keystone
Service API
计算节点
Nova Compute
Open vSwitch
Openvswitch-agents
Libvirtd
计算节点
Nova Compute
Openvswitch
Openvswitch-agents
Libvirtd
计算节点
Nova Compute
Openvswitch
Openvswitch-agents
Libvirtd
存储网络（万兆）
Ceph节点
主 + 备 + 备

控制节点 高可用


Controller Node
Pacemaker
Neutron agents (Active)
MySQL galera
RabbitMQ
HA Proxy
Horizon
Nova
Glance
Cinder
Neutron Server
Keystone
Service API
Mysql galera Cluster
Active/Active
RabbitMQ Cluster
Active/Active
通过VIP 调用
RESTFUL API
Controller Node
Pacemaker
Neutron agents (Stand by)
MySQL galera
RabbitMQ
HA Proxy
Horizon
Nova
Glance
Cinder
Neutron Server
Keystone
Service API
Controller Node
Pacemaker
Neutron agents (Stand by)
MySQL galera
RabbitMQ
HA Proxy
Horizon
Nova
Glance
Cinder
Neutron Server
Keystone
Service API

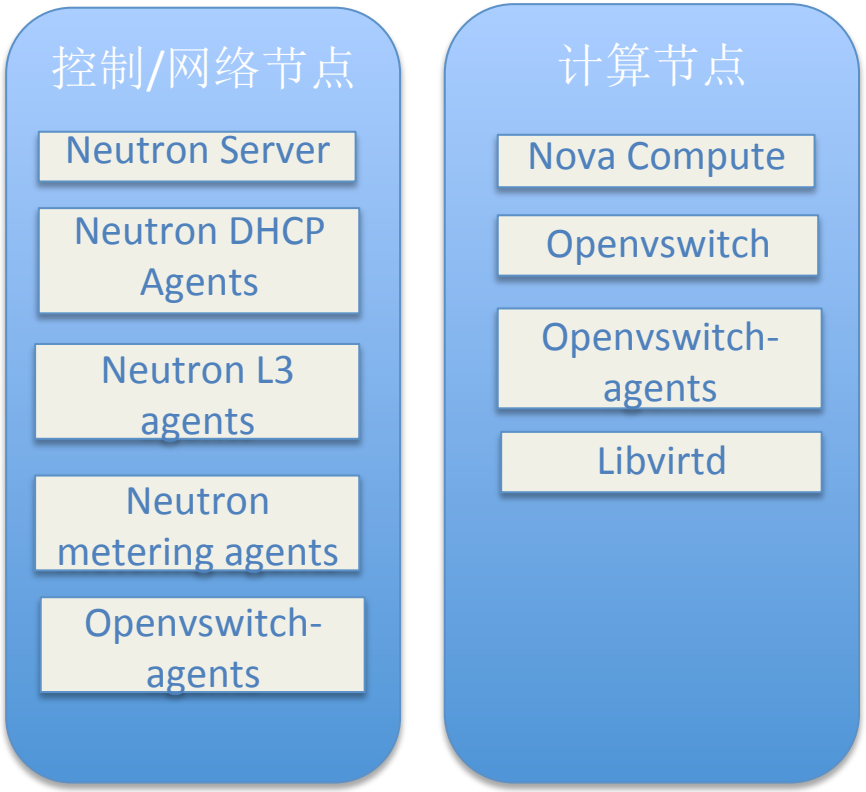
网络部署

Neutron Server

- Neutron Server: 守护进程，用来提供外部调用的API和与其它组件交互的接口。

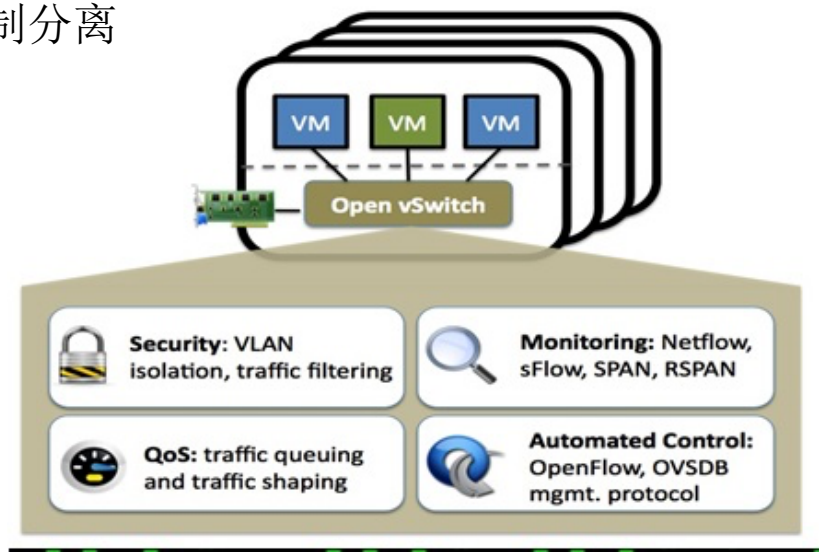
Neutron Agents

- plug-in agent(neutron-*-agent)插件代理，部署在每一个运行hypervisor的主机上，它提供本地的vSwitch配置，例如OpenVswitch 的openvswitch-agents
- Neutron-DHCP-Agent：为每一个子网配置DHCP
- Neturon-L3-Agent：设置iptables/routing/NAT表
- Neutron-metering-agent： 计量代理，为租户网络提供三层网络流量数据计量服务

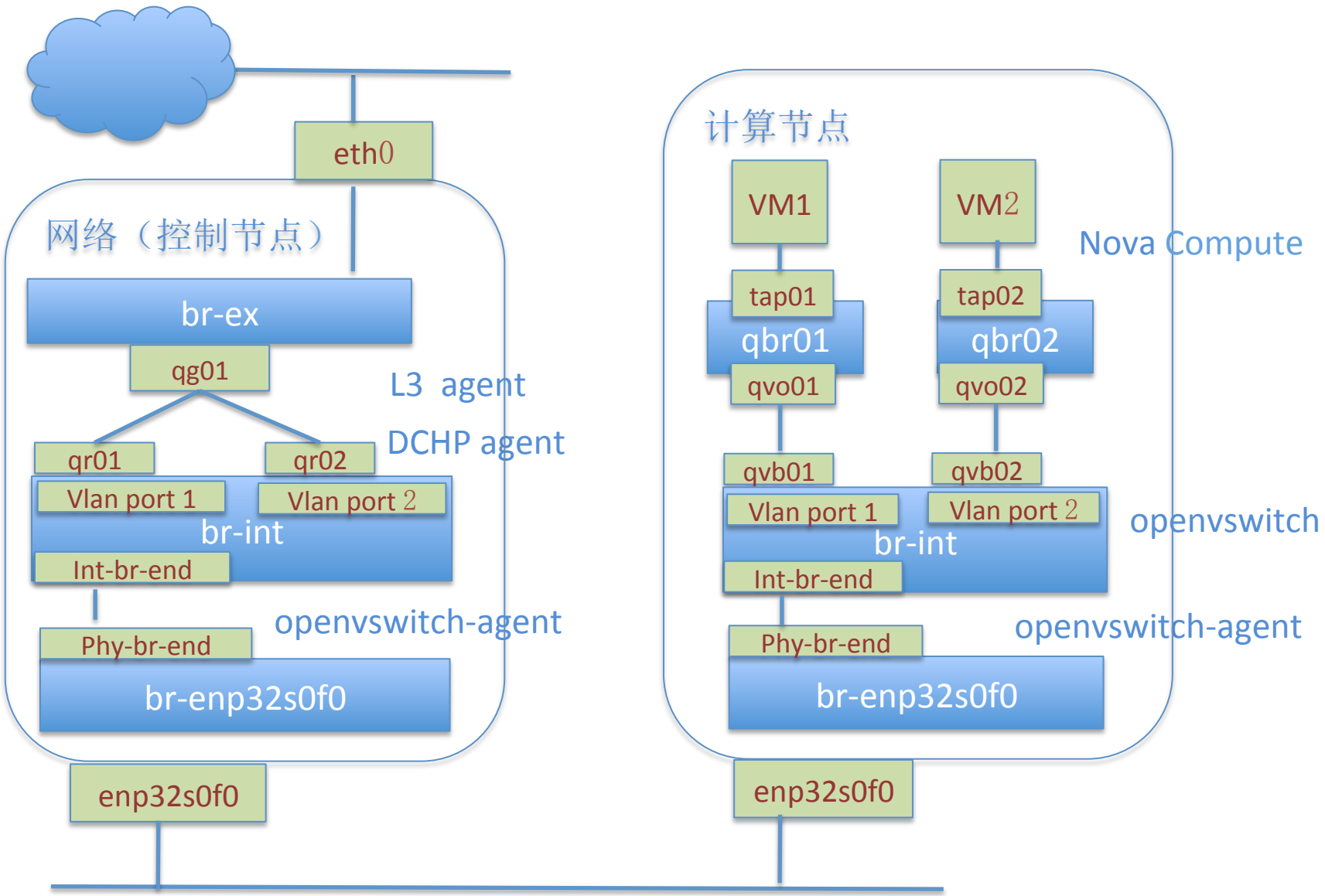


OVS vs GRE

- 基本构架相同，GRE会在通道网桥上实现点对点到的Tunnel协议封装，OVS不需要
- GRE
  - 可以隔离广播风暴.
  - 没有vlan id个数限制，3层隧道技术可以实现跨机房部署。
  - 在IP头中增加Tunnel ID，势必减少vm的mtu值，同样大小的数据，需要更多的ip包来传，传输效率有影响。
- OVS
  - 针对每个vm做流量限制、流量监控、数据包分析
  - 同时可以引入OpenFlow，实现流量和控制分离



实施网络 OVS（+Linux Bridge）+ VLAN



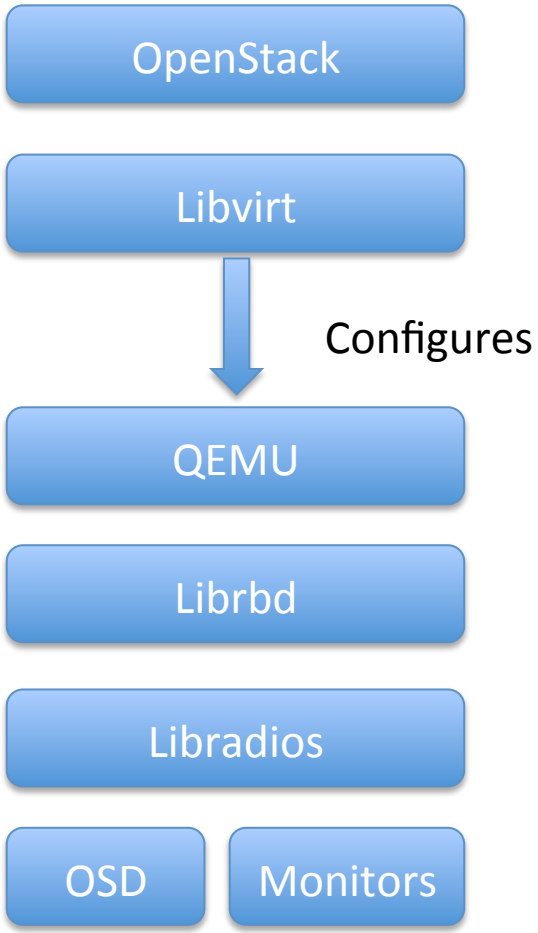
存储部署

Ceph

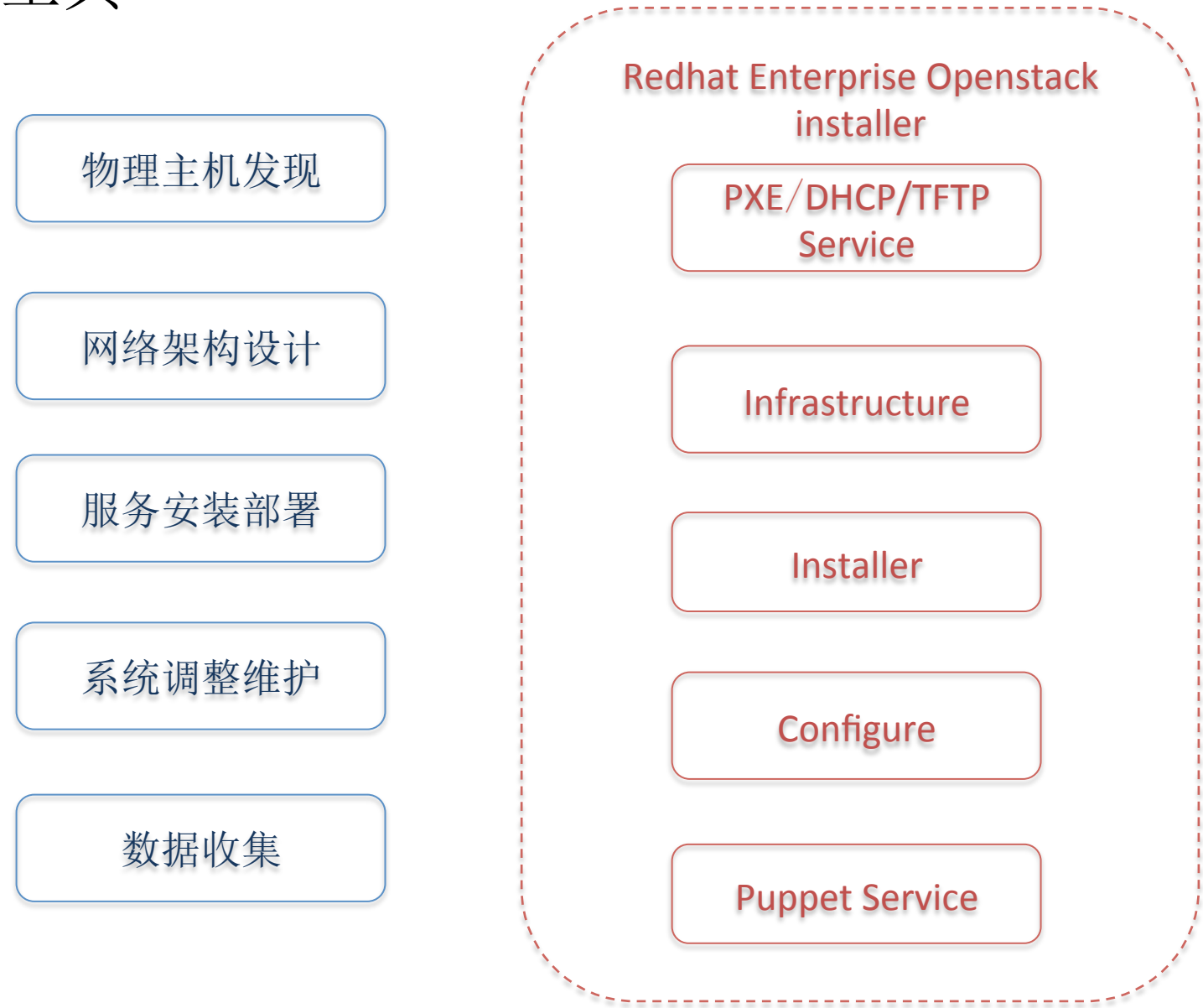
- 它主要提供块服务，也能提供对象存储服务
- 能很好的支持顺序IO和随机IO。除了用作cinder后端给虚拟机提供卷服务，同时也可以作为glance的后端，libvirt配置了librbd的QEMU接口
- 为了同时保证写性能和一致性，节点间通讯的延迟要尽量小

Swift

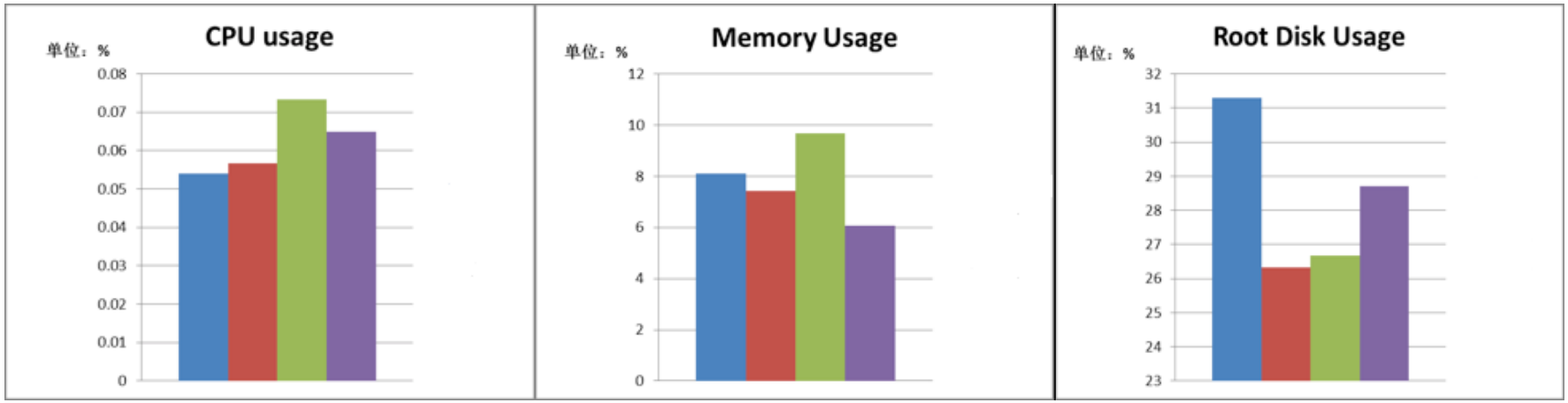
- 支持不同的存储后端，不依赖后端的任何如多副本，RAID特性来提供可靠性
- 最终一致性模型
- 支持跨机房跨地域部署。

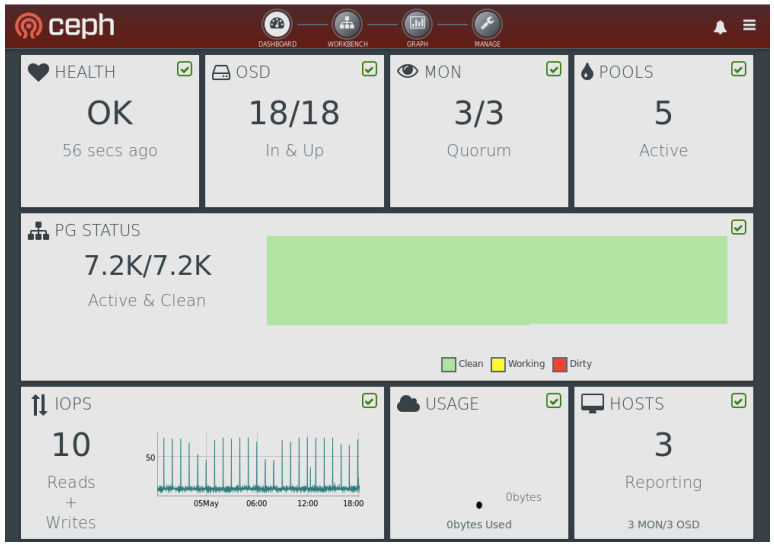


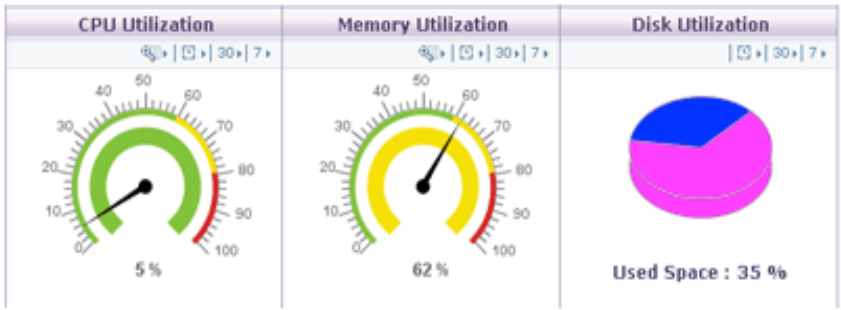
部署工具


物理主机发现
网络架构设计
服务安装部署
系统调整维护
数据收集
Redhat Enterprise Openstack installer
PXE/DHCP/TFTP Service
Infrastructure
Installer
Configure
Puppet Service

监控管理


CPU usage
单位：%
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
Memory Usage
单位：%
12
10
8
6
4
2
0
Root Disk Usage
单位：%
32
31
30
29
28
27
26
25
24
23


ceph
DASHBOARD
WORKBENCH
GRAPH
MANAGE
HEALTH
OK
56 secs ago
OSD
18/18
In & Up
MON
3/3
Quorum
POOLS
5
Active
PG STATUS
7.2K/7.2K
Active & Clean
Clean
Working
Dirty
IOPS
10
Reads
+
Writes
50
05May
06:00
12:00
18:00
USAGE
0bytes
0bytes Used
HOSTS
3
Reporting
3 MON/3 OSD


CPU Utilization
Memory Utilization
Disk Utilization
5 %
62 %
Used Space : 35 %

实际问题

- 虚拟机的可用性
  - 提供live-migration/ migration /flavor resize的配置
  - 解决的问题：多个计算节点之间ssh 互信配置
- 流量控制
  - 现有OpenStack 构架中，Qos设置相关功能较弱
  - 解决办法：提供修改网桥接口流量的工具;定制flavor。

实际问题

- 安全策略
  - 现有openstack 构架中对安全策略支持较弱
  - 解决办法：在外部网络出口的物理交换机旁路架设安全设备，做安全扫描。
- 出口带宽
  - 现有构架中对出口带宽受构架限制
  - 解决办法：修改网络节点Neutron agent状态策略。

谢 谢！