

openstack™
CLOUD SOFTWARE

OpenStack从入门到精通实践 第十二周

DATAGURU专业数据分析社区

【声明】 本视频和幻灯片为炼数成金网络课程的教学资料，所有资料只能在课程内使用，不得在课程以外范围散播，违者将可能被追究法律和经济责任。

课程详情访问炼数成金培训网站

<http://edu.dataguru.cn>

炼数成金逆向收费式网络课程



- Dataguru (炼数成金) 是专业数据分析网站，提供教育，媒体，内容，社区，出版，数据分析业务等服务。我们的课程采用新兴的互联网教育形式，独创地发展了逆向收费式网络培训课程模式。既继承传统教育重学习氛围，重竞争压力的特点，同时又发挥互联网的威力打破时空限制，把天南地北志同道合的朋友组织在一起交流学习，使到原先孤立的学习个体组合成有组织的探索力量。并且把原先动辄成千上万的学习成本，直线下降至百元范围，造福大众。我们的目标是：低成本传播高价值知识，构架中国第一的网上知识流转阵地。
- 关于逆向收费式网络的详情，请看我们的培训网站 <http://edu.dataguru.cn>

■ 高可用 (High Availability , 简称 HA)

- 高可用性是指提供在本地系统单个组件故障情况下，能继续访问应用的能力。
- HA 的计算公式是 $[1 - (\text{宕机时间}) / (\text{宕机时间} + \text{运行时间})]$
 - 2 个9 : $99\% = 1\% * 365 = 3.65 * 24 \text{ 小时/年} = 87.6 \text{ 小时/年的宕机时间}$
 - 4 个9: $99.99\% = 0.01\% * 365 * 24 * 60 = 52.56 \text{ 分钟/年}$
 - 5 个9 : $99.999\% = 0.001\% * 365 = 5.265 \text{ 分钟/年的宕机时间}$ ，也就意味着每次停机时间在一到两分钟。
 - 11 个 9 : 几乎就是几年才宕机几分钟。

■ HA 按服务状态两类：

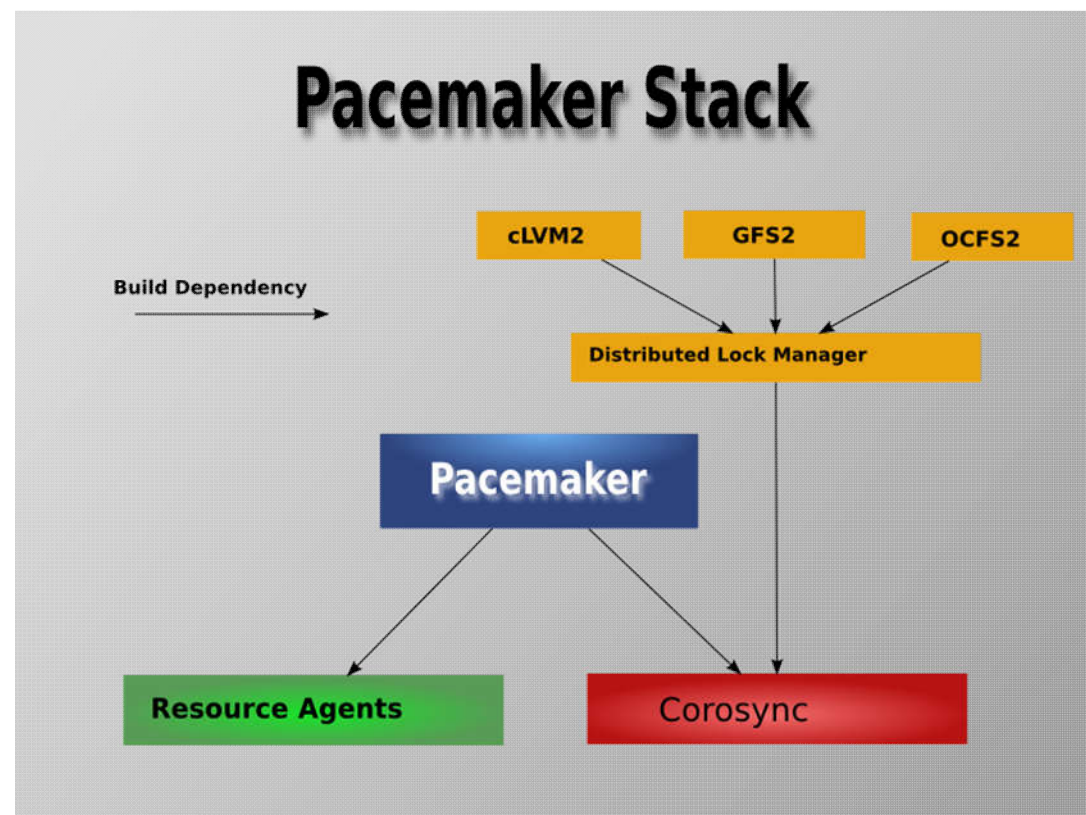
- 有状态服务：后续对服务的请求依赖于之前对服务的请求。
- 无状态服务：对服务的请求之间没有依赖关系，是完全独立的。

■ HA按冗余分类：

- Active/Passive HA：集群只包括两个节点简称主备。在这种配置下，系统采用主和备用机器来提供服务，系统只在主设备上提供服务。在主设备故障时，备设备上的服务被启动来替代主设备提供的服务。典型地，可以采用 CRM 软件比如 Pacemaker 来控制主备设备之间的切换，并提供一个虚拟IP 来提供服务。
- Active/Active HA：集群只包括两个节点时简称双活，包括多节点时成为多主（Multi-master）。在这种配置下，系统在集群内所有服务器上运行同样的负载。以数据库为例，对一个实例的更新，会被同步到所有实例上。这种配置下往往采用负载均衡软件比如 HAProxy 来提供服务的虚拟 IP。

Pacemaker

- Pacemaker 承担集群资源管理者 (CRM - Cluster Resource Manager) 的角色，它是一款开源的高可用资源管理软件，适合各种大小集群：
 - Pacemaker - 资源管理器(CRM)，负责启动和停止服务
 - Corosync - 消息层组件 (Messaging Layer)，管理成员关系、消息和仲裁
 - Resource Agents - 资源代理，实现在节点上接收 CRM 的调度对某一个资源进行管理的工具



OpenStack HA (一)



- 从功能服务维度划分
 - 基础服务
 - 计算
 - 网络
 - 存储
- OpenStack 部署环境中，各节点可以分为几类：
 - Controller Node (控制节点)
 - Neutron Node (网络控制节点)
 - Storage Node (存储控制节点)
 - Compute node (计算节点)

OpenStack HA (二)



- OpenStack HA原则：
 - 能 A/A 尽量 A/A，不能的话则 A/P
 - 有原生（内在实现的）HA方案尽量选用原生方案，没有的话则使用额外的HA 软件比如 Pacemaker 等
 - 需要考虑负载均衡
 - 方案尽可能简单，不要太复杂
 - 结合具体的使用场景

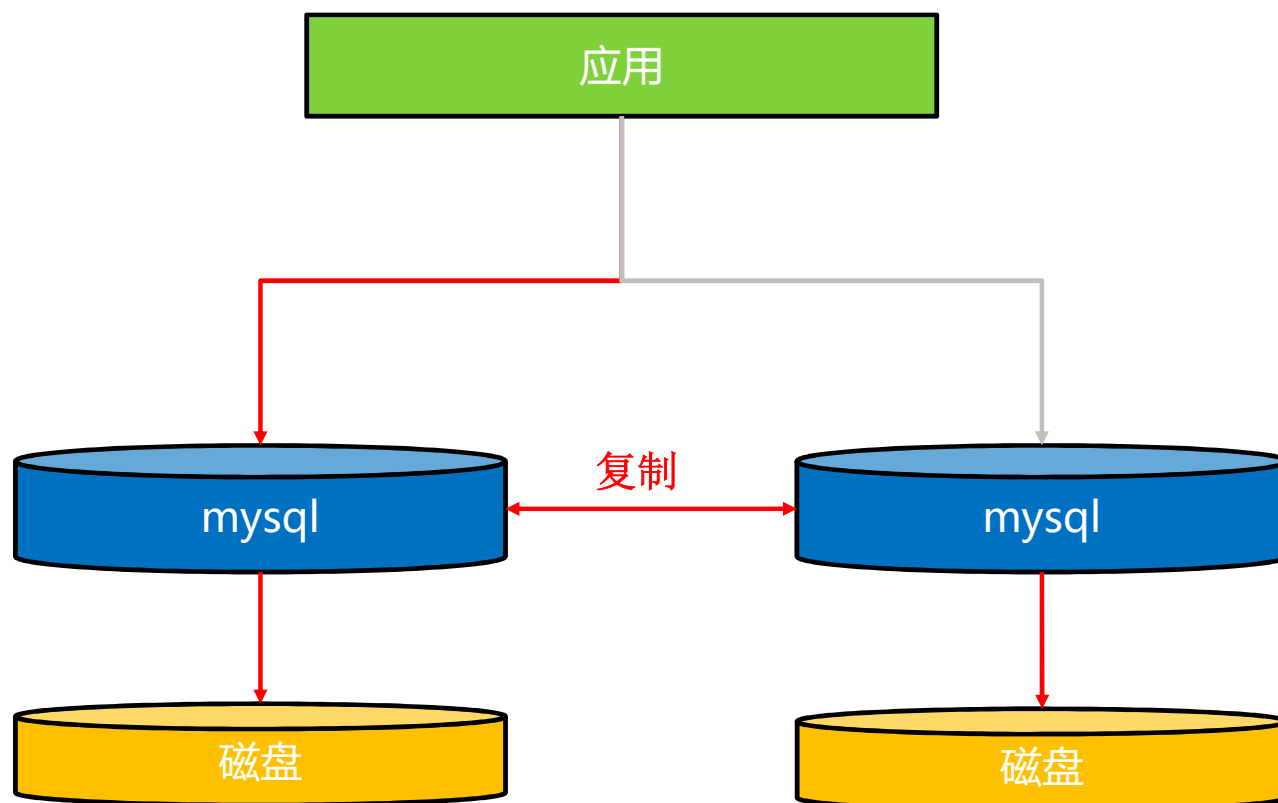
基础服务的高可用(mysql)

■ 实现原则

- 数据库可用
- 数据可靠

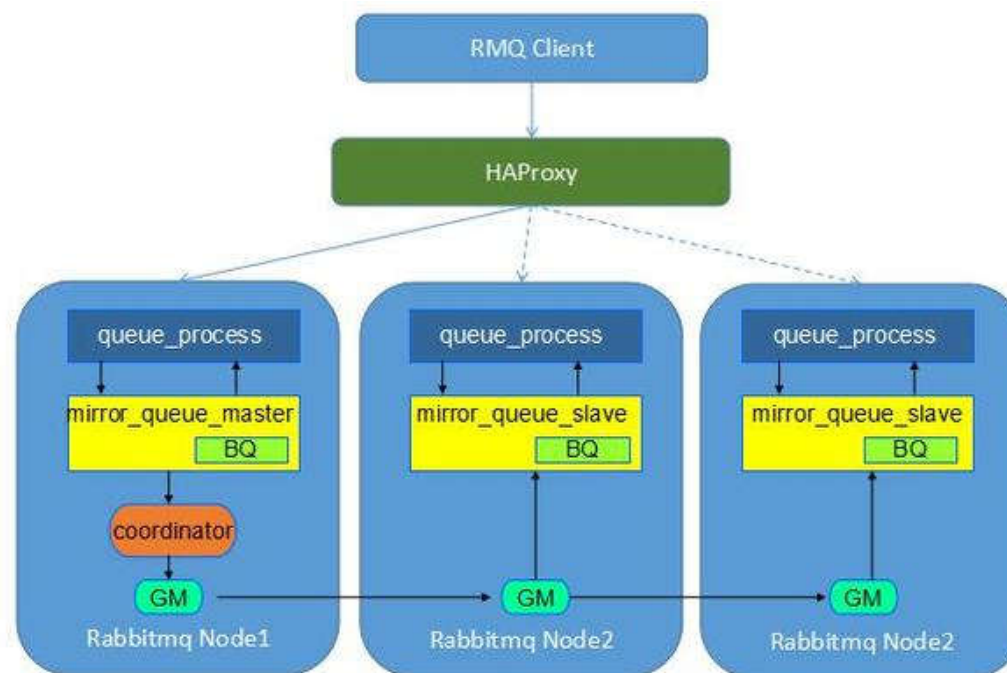
■ mysql高可用常见实现方式：

- M-M+HA
- M-S+MHA
- Haproxy+Galera

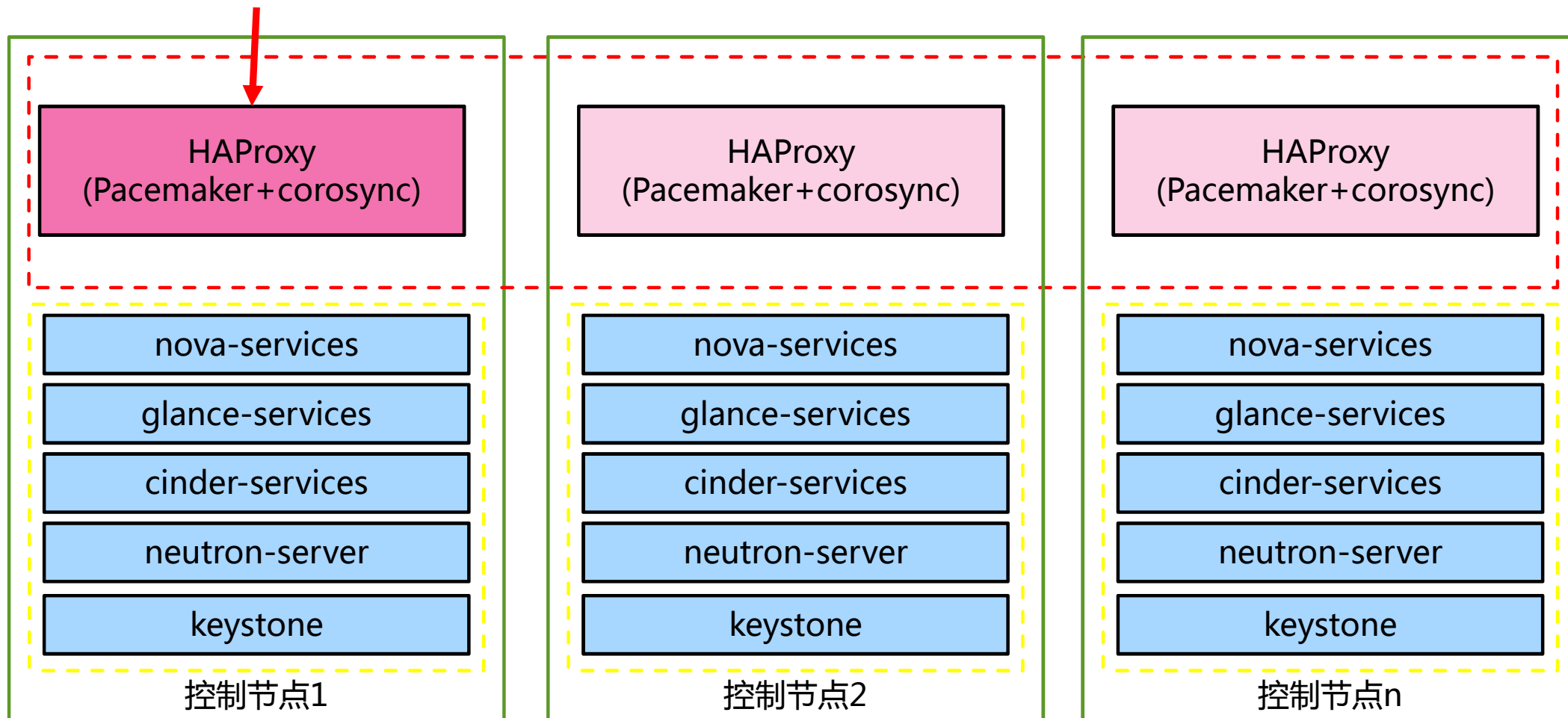


基础服务的高可用性(消息队列)

- openstack除了支持rabbitMQ，还支持Qpid和ZeroMQ
- oslo.messaging:封装了处理消息队列相关逻辑的一个公共库(rabbit_hosts)



控制节点高可用



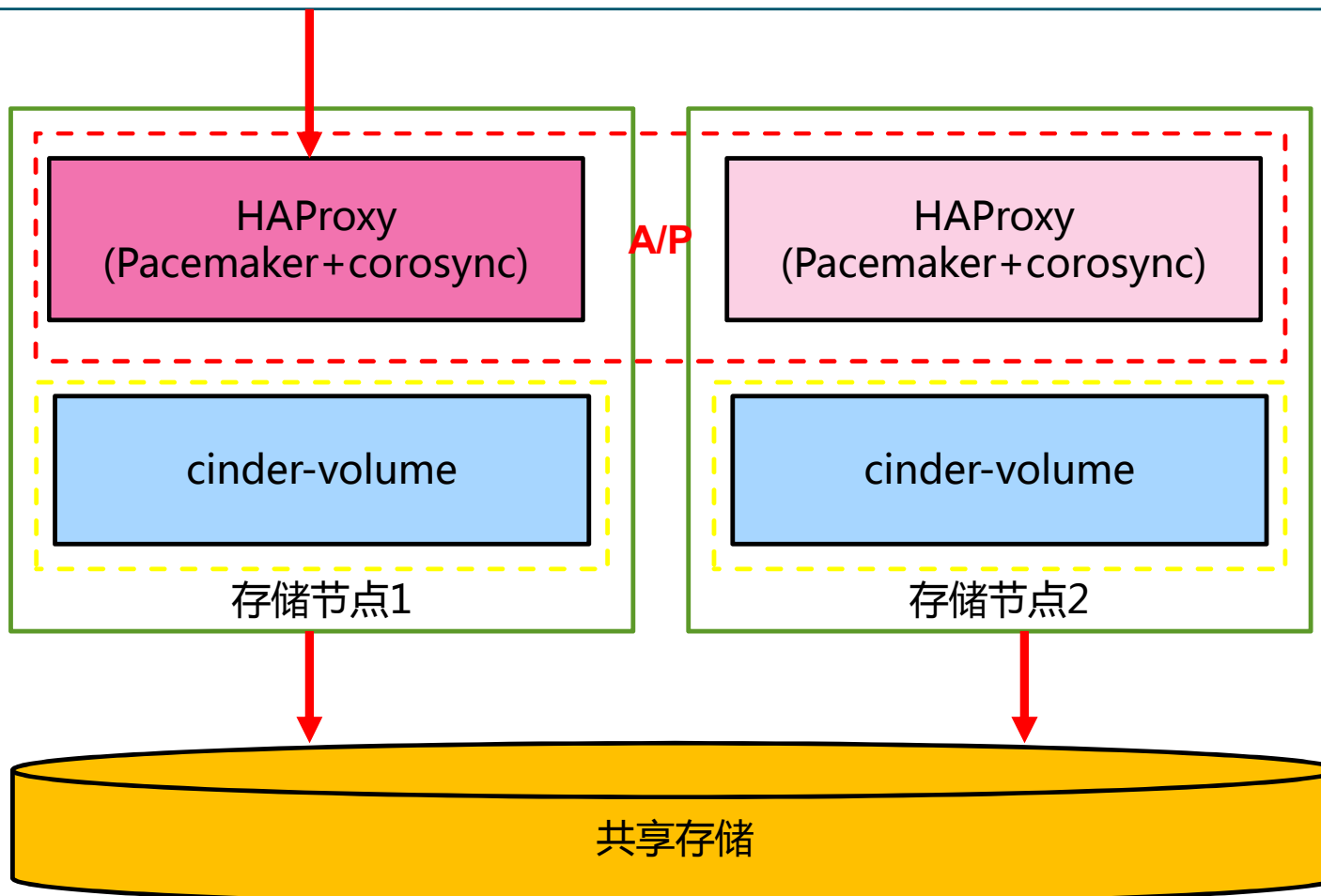
DATAGURU专业数据分析社区

keystone服务高可用(token)

- keystone支持将token存储到mysql、memache、redis等
 - driver = keystone.token.persistence.backends.memcache.Token
 - driver = keystone.token.persistence.backends.sql.Token
 - driver = keystone.token.persistence.backends.kvs.Token

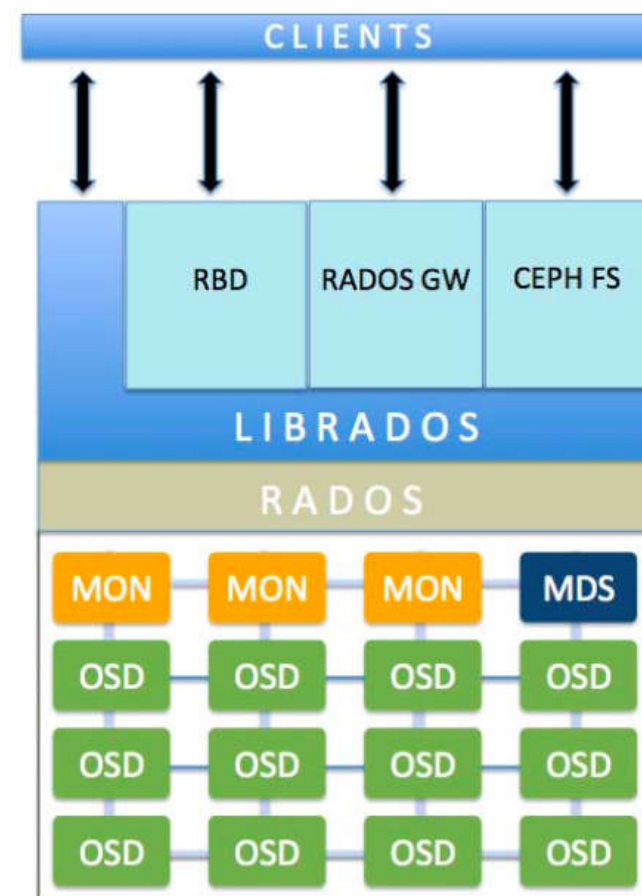


cinder服务高可用（一）

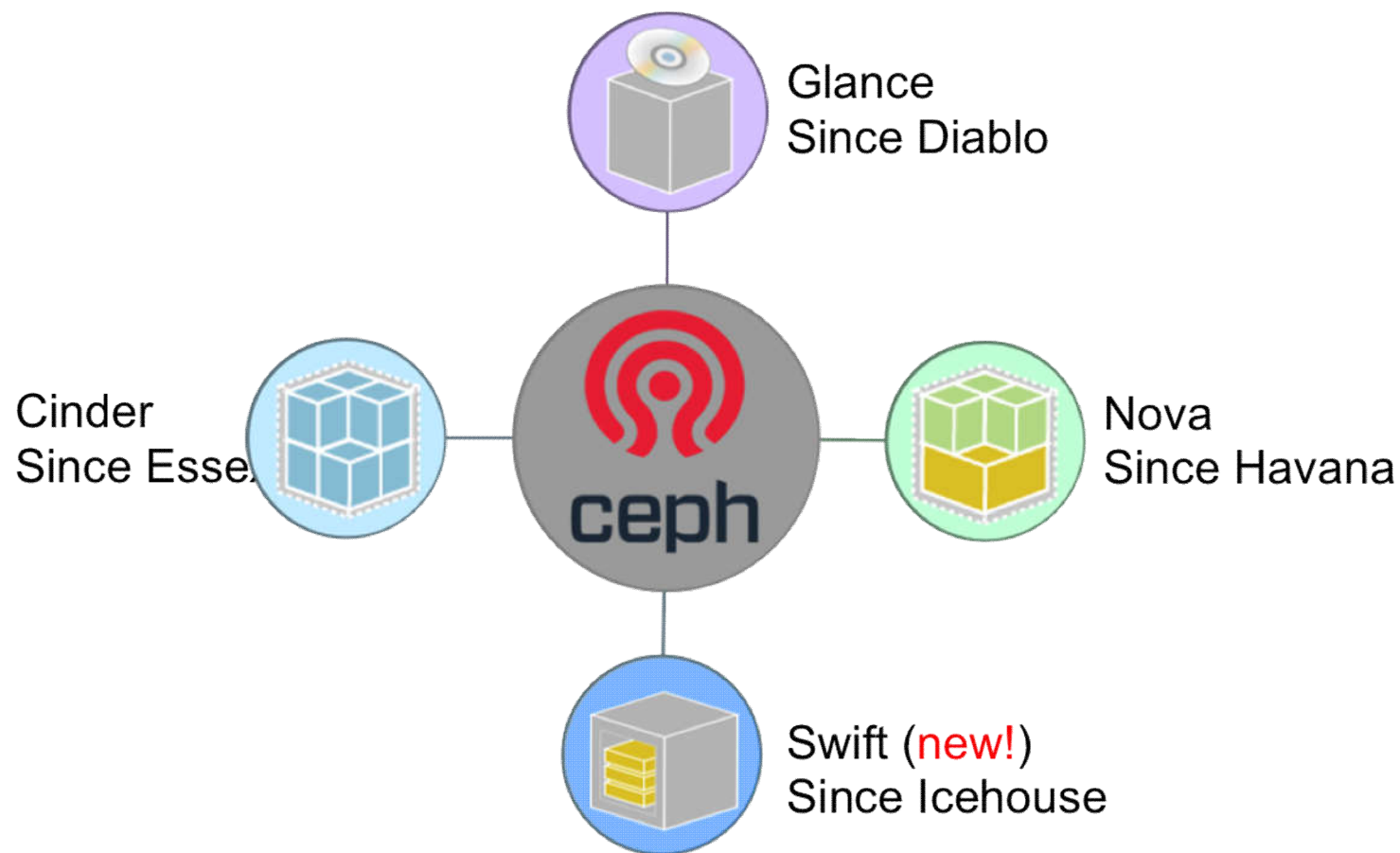


cinder服务高可用（二）

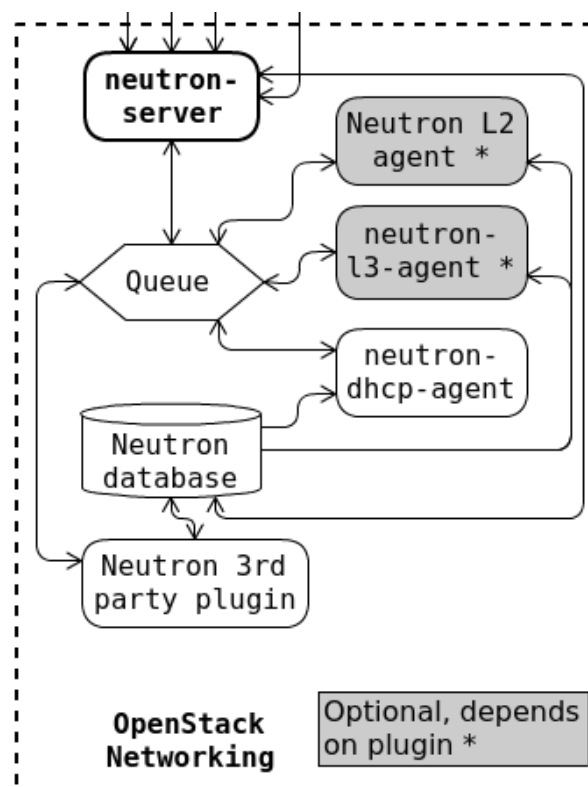
- ceph
 - 提供分布式的对象存储、块设备存储和文件系统服务



cinder服务高可用（三）



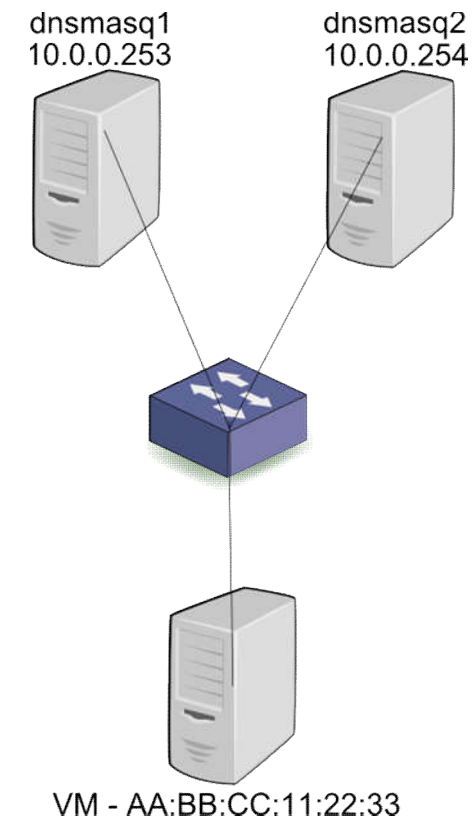
neutron服务高可用（一）



neutron服务高可用（二）

■ DHCP Agent

- DHCP 协议自身就支持多个 DHCP 服务器
- neutron.conf :
 - dhcp_agents_per_network = X

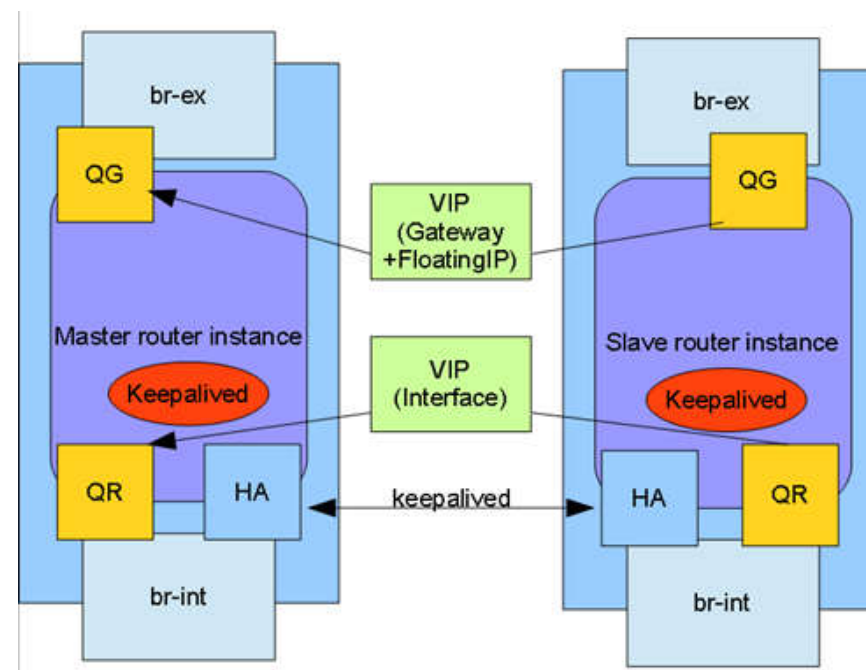


- VRRP (Virtual Router Redundancy Protocol)
 - 虚拟路由冗余协议。它能够自动选择可用的路由路径，从而提高了路由路径的可用性和稳定性。
- DVR(Distributed Virtual Router)
 - 分布式虚拟路由。顾名思义就是 Neutron 的 router 将不单单部署在网络节点上，所有启动了 Neutron L3 Agent 的节点，都会在必要时在节点上创建 Neutron router 对应的 namespace，并更新与 DVR router 相关的 Openflow 规则，从而完成 DVR router 在该节点上的部署。
 - 同一租户，不同子网之间通信（直接由计算节点处理）
 - 虚拟机(有浮动IP)访问外部网络（直接由计算节点处理）
 - 虚拟机（无浮动IP)访问外部网络（由网络节点处理）

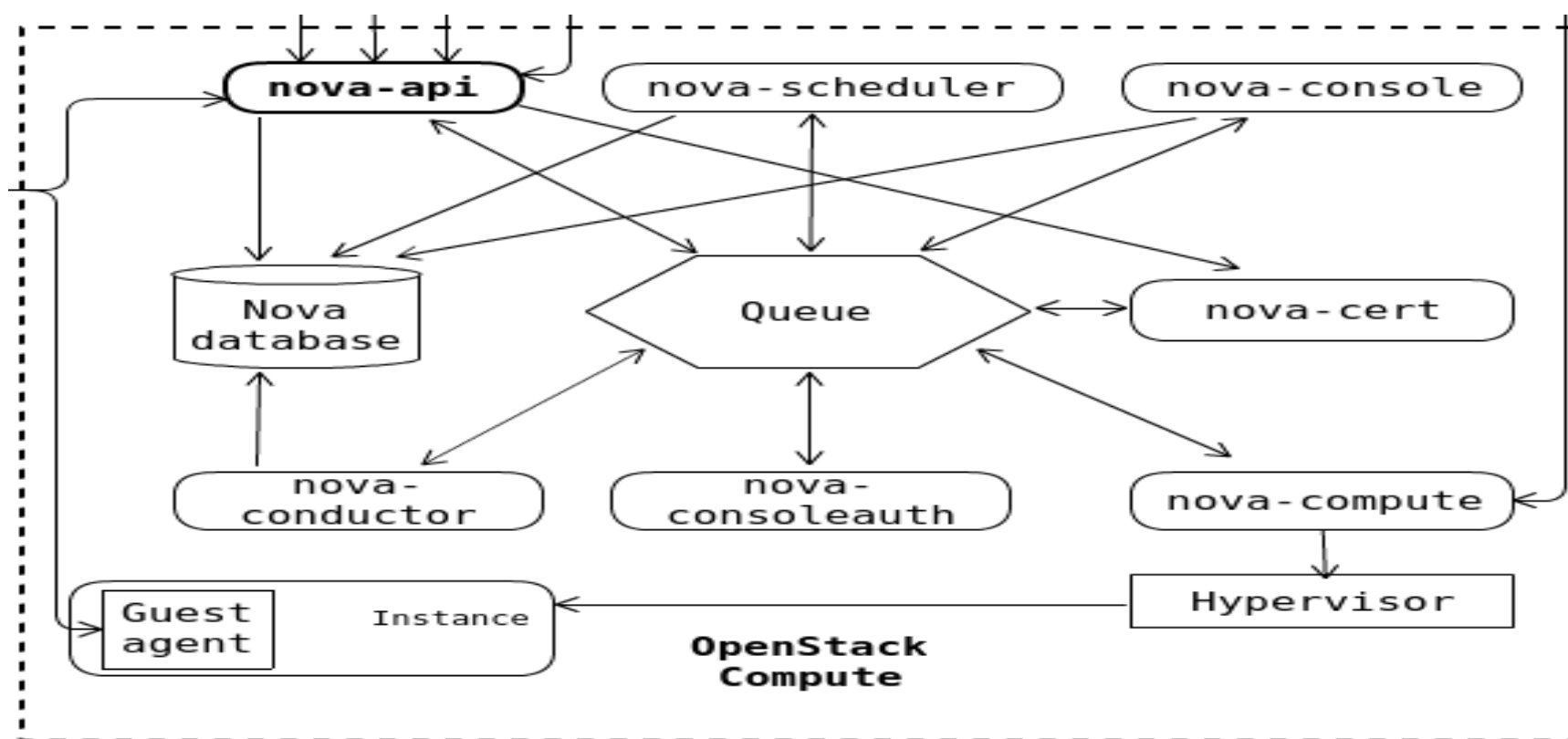
neutron服务高可用（四）

■ L3 Agent HA(基于VRRP)

- 修改 Neutron配置文件，支持创建多个 router 实例部署到不同物理服务器上的 L3 Agent 中
- 借助Keepalived，实现保证多个 L3 Agent 的 HA，即其中一个是 Master，其他是 Backup



nova服务高可用（一）



nova服务高可用（二）

■ 方案前提

- 虚拟机部署在共

■ 方案一

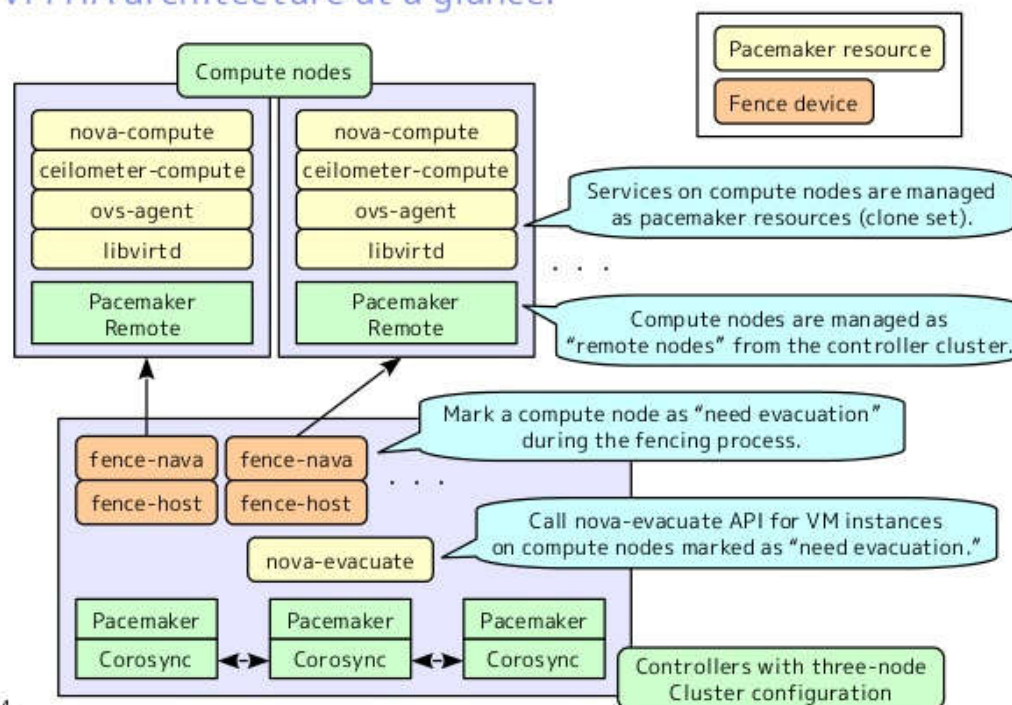
- 在控制节点执行
- 通过nova eva

■ 方案二

- Pacemaker+C
- Pacemaker-re
- Pacemaker+C
移

Red Hat Enterprise Linux OpenStack Platform 7 VM Instance HA Architecture

VM HA architecture at a glance.

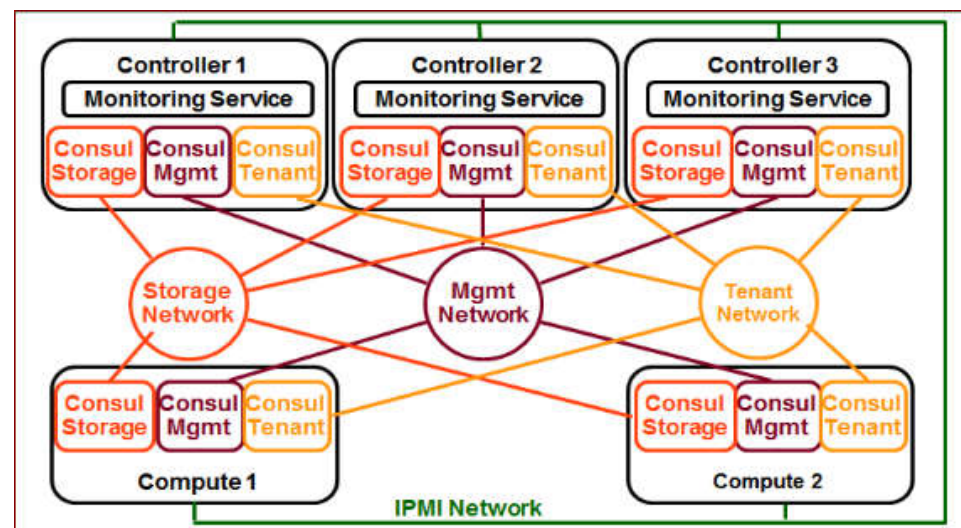
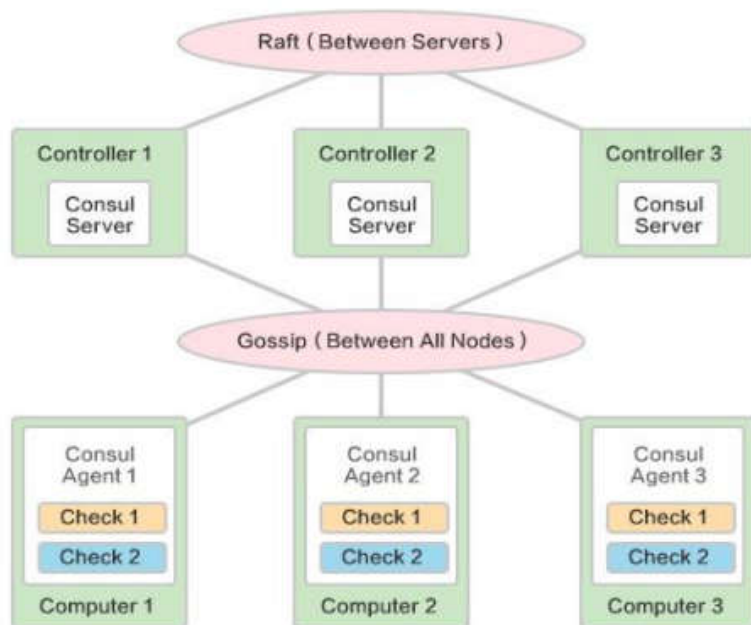


节点进行隔离和虚拟机迁

nova服务高可用（三）

■ 方案三

- 基于Consul的分布式健康检查(<https://www.consul.io/>)



Thanks

FAQ时间