

腾讯云TStack IaaS 升级实践

徐凯 李瑜洁
腾讯TStack架构师



目录

01

TStack简介

02

OpenStack升级实践 (From Kilo To Pike)

03

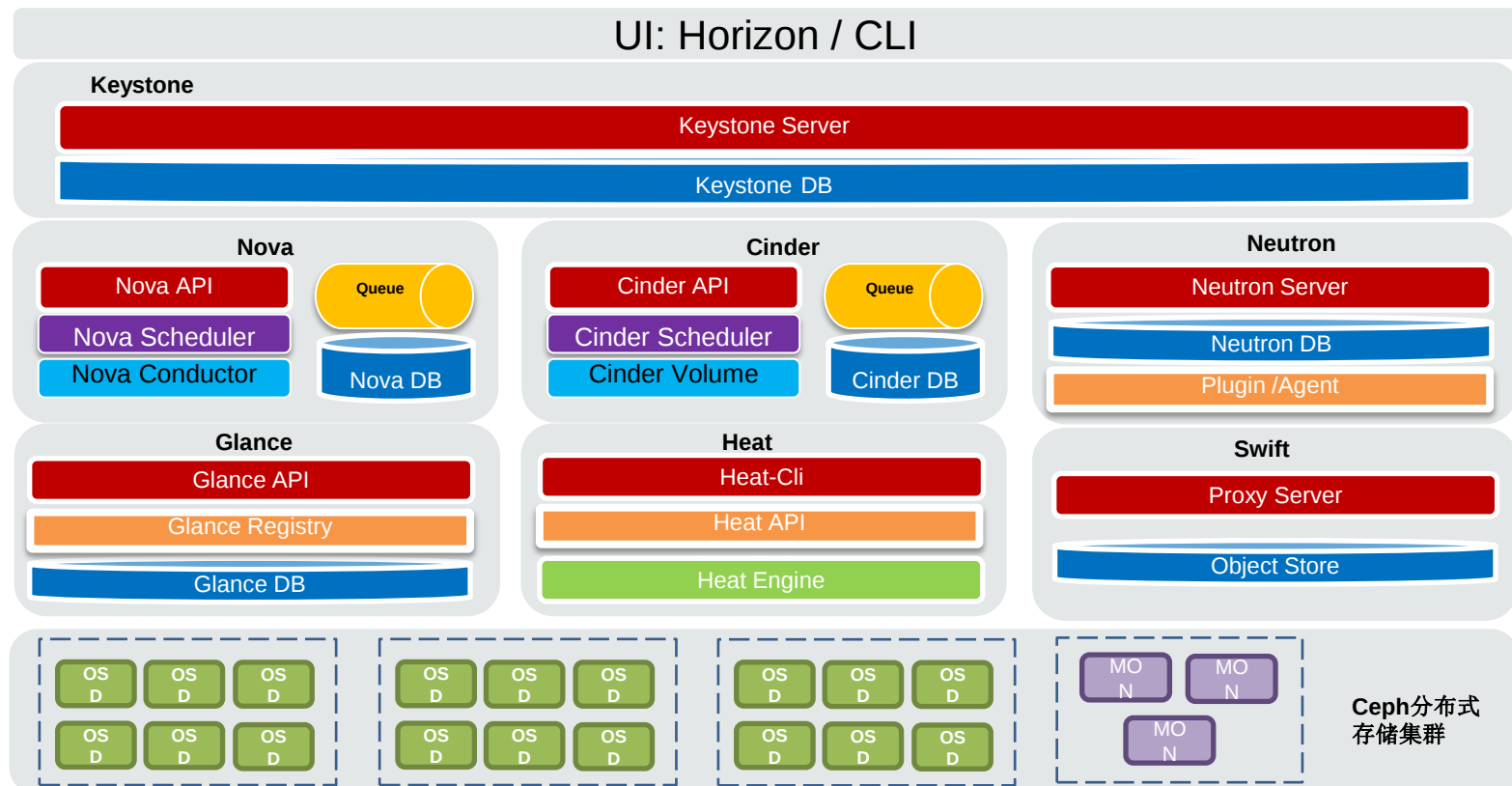
Ceph升级实践 (From Hammer to Jewel)

TStack简介

TStack是腾讯云基于自身强大技术能力和海量运营经验推出的私有云平台，提供集IaaS、PaaS和SaaS为一体的综合云服务解决方案；大量私有化部署经验，具备混合云管理能力，具有高稳定性、统一管理、可视化运营等特点，助力政府、企业构建稳定安全的云环境和健康的云生态。



TStack简介



目录

01

TStack简介

02

OpenStack升级实践 (From Kilo To Pike)

03

Ceph升级实践 (From Hammer to Jewel)

Openstack 半年发布一次，新版本功能带来惊喜，升级不可避免。



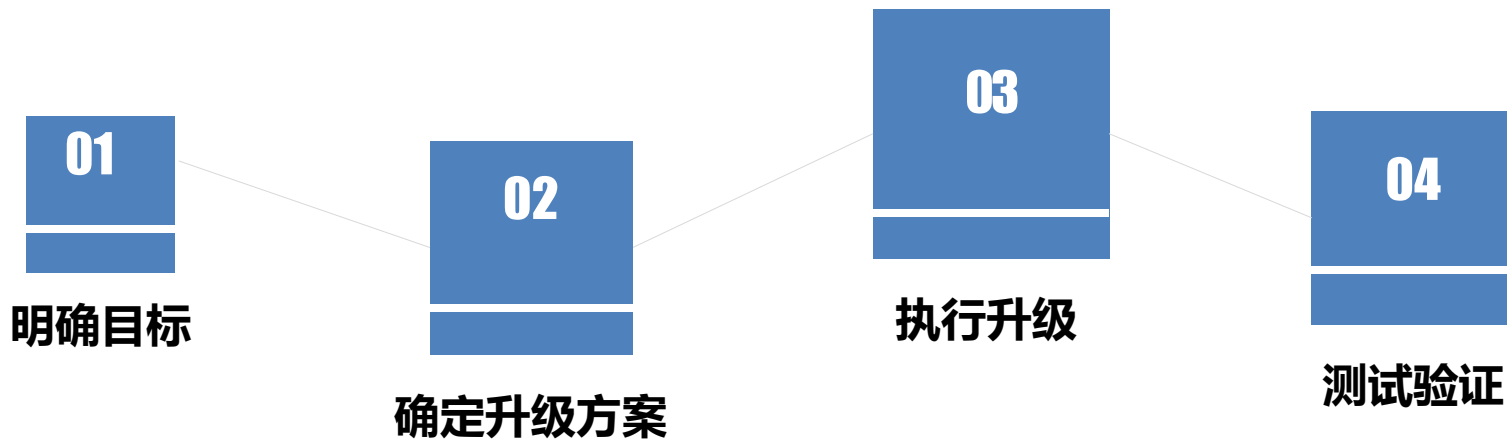
- 不断增强对容器的支持
- Cells V2实现了Nova计算环境的横向扩展
- Lbaas、虚拟机调度等不断完善
- 虚拟Gpu、volume多挂载
- ...

TStack IaaS升级实践

挠头...



TStack IaaS升级实践



TStack IaaS升级实践

升级目标

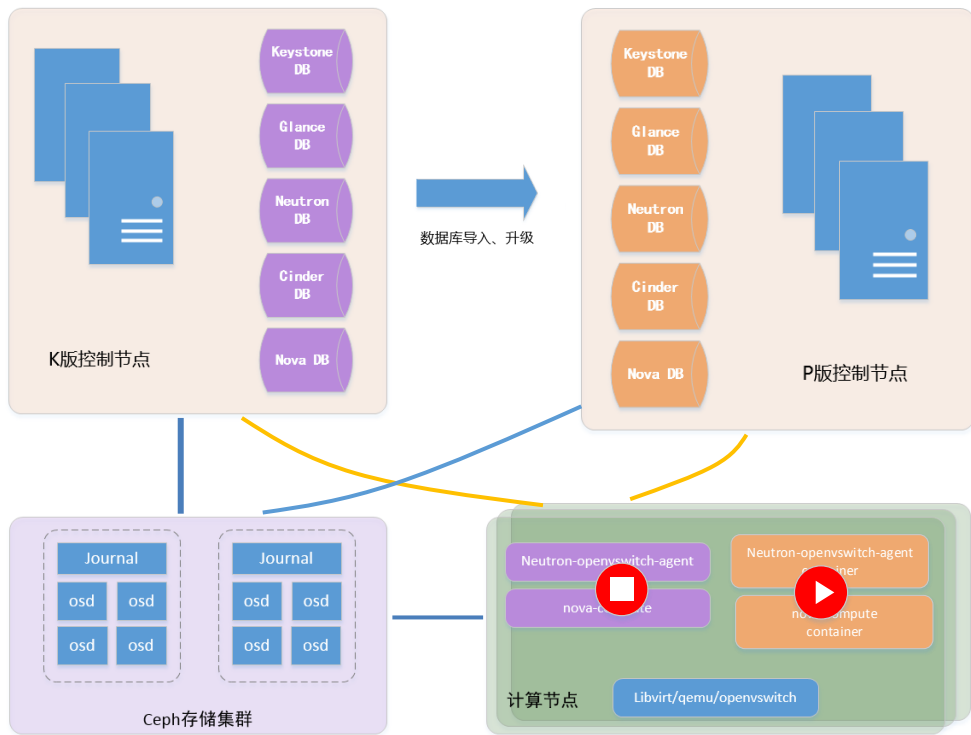
平滑升级
业务无影响

最短服务暂停

升级可回退

TStack IaaS升级实践

结合部署架构确定升级方案



+ 01.控制节点 替换升级

- 需额外三台物理服务器
- 将K版数据库导入P版，并进行升级。

+ 02.计算节点 容器升级

- 停止K版openstack服务
- 启动P版openstack容器

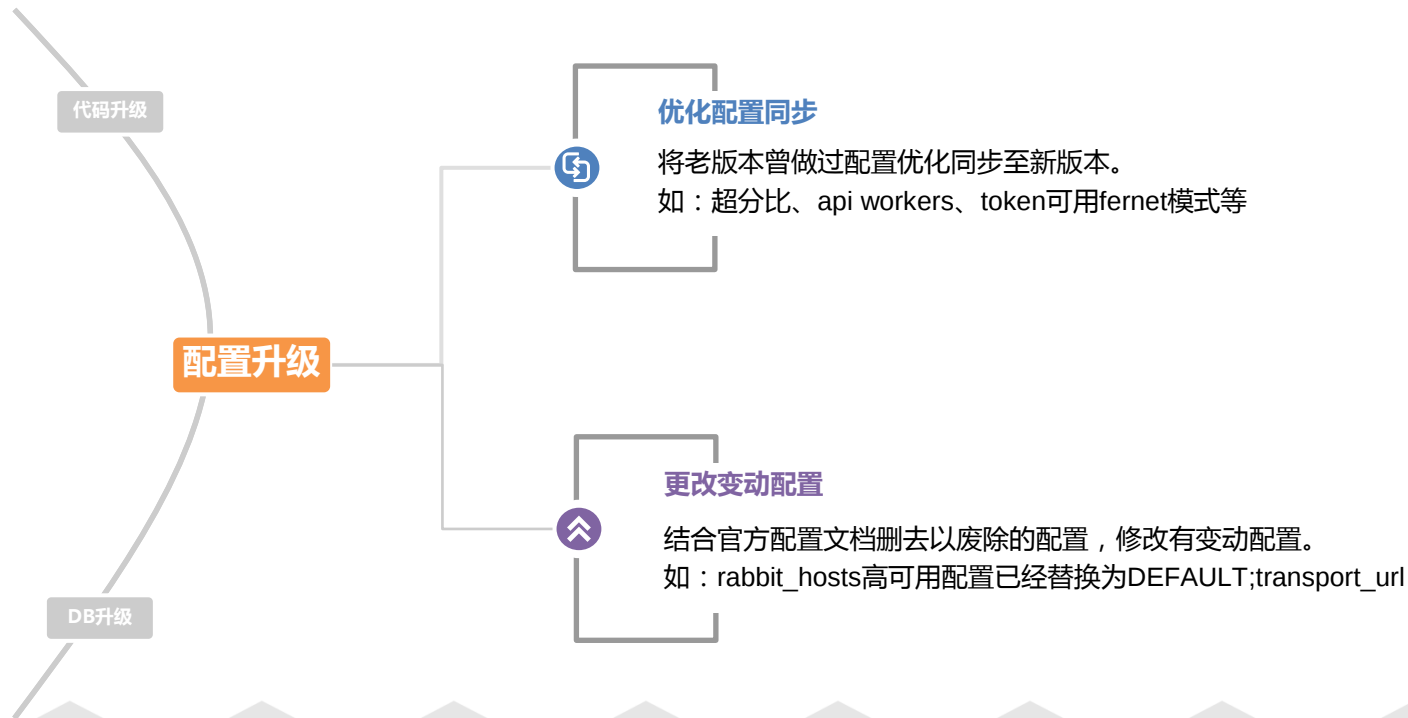
TStack IaaS升级实践

OpenStack From Kilo to Pike



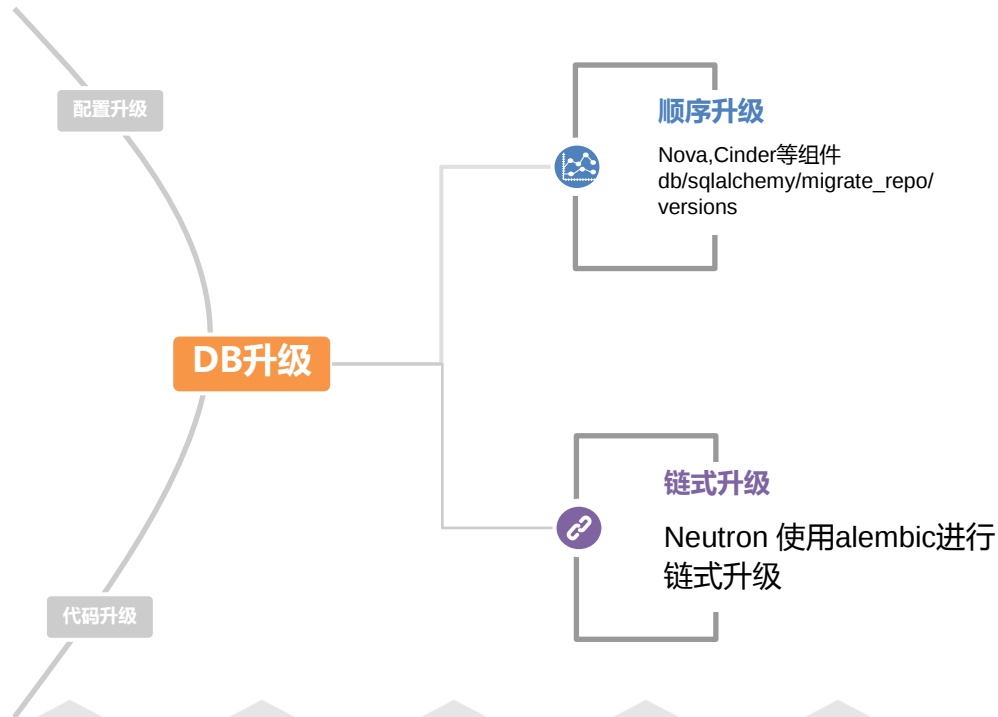
TStack IaaS升级实践

OpenStack From Kilo to Pike



TStack IaaS升级实践

OpenStack From Kilo to Pike



```
291_enforce_flavors_migrated.py
292_drop_nova_volumes_tables.py
293_add_migration_type.py
294_add_service_heartbeat.py
295_add_virtual_interfaces_uuid_index.py
296_add_missing_db2_fkeys.py
297_add_forced_down_for_services.py
298_mysql_extra_specs_binary_collation.py
299_service_version_number.py
300_migration_context.py
301_add_cpu_and_ram_ratios_for_compute_nodes.py
302_pgsql_add_instance_system_metadata_index.py
```

```
030a959ceafa_uniq_routerports0port_id.py
0f5bef0f87d4_add_qos_minimum_bandwidth_rules.py
30107ab6a3ee_provisioning_blocks.py
3d0e74aa7d37_add_flavor_id_to_routers.py
45f8dd33480b_qos_dscp_db_addition.py
5abc0278ca73_add_support_for_vlan_trunking.py
5cd92597d11d_add_ip_allocation_to_port.py
67daae611b6e_add_standard_attr_to_qos_policies.py
6b461a21bcfc_uniq_floatingips0floating_network_.py
a5648cfceadf_add_subnet_service_types.py
a963b38d82f4_add_dns_name_to_portdnses.py
c415aab1c048_add_revisions_column.py
d3435b514502_add_device_id_index_to_port.py
```

```
# revision identifiers, used by Alembic.
revision = '030a959ceafa'
down_revision = '3d0e74aa7d37'
```

控制节点DB升级

- 01 Keystone
 - 备份K版数据库，导入到P版控制节点
 - `su -s /bin/sh -c "keystone-manage db_sync" keystone`
 - 补充新增服务、用户、endpoint等 (placement,cinderV3)
- 02 Nova
 - 备份K版数据库，导入到P版控制节点。
 - `su -s /bin/sh -c "nova-manage db sync" nova`
 - `su -s /bin/sh -c "nova-manage api_db sync" nova`
 - `su -s /bin/sh -c "nova-manage db online_data_migrations" nova`
- 03 Neutron
 - 备份K版数据库，导入到P版控制节点。
 - `neutron-db-manage upgrade --expand`
 - `neutron-db-manage upgrade --contract`

跨越5版本，多个组件无法自动升级。



控制节点DB升级

- `ValidationError: Migration cannot continue until all these have been migrated to the api database. Please run `nova-manage db online_data_migrations` on Newton code before continuing.`
 - There are still 1 unmigrated records in the `compute_nodes` table. Migration cannot continue until all records have been migrated.
- 330_enforce_mitaka_online_migrations.py**

```
[root@controller ~]# glance-manage db sync
/usr/lib/python2.7/site-packages/oslo_db/sqlalchemy/enginefacade.py:1328: OsloDBDeprecationWarning: EngineFacade is deprecated; please use oslo_db.sqlalchemy.enginefacade
  expire_on_commit=expire_on_commit, _conf=conf)
Database is currently not under Alembic's migration control.
('Legacy head: ', 41)
The current database version is not supported any more. Please upgrade to Liberty release first.
[root@controller ~]#
```

TStack IaaS升级实践

计算节点升级难点

01

- 保证业务无影响

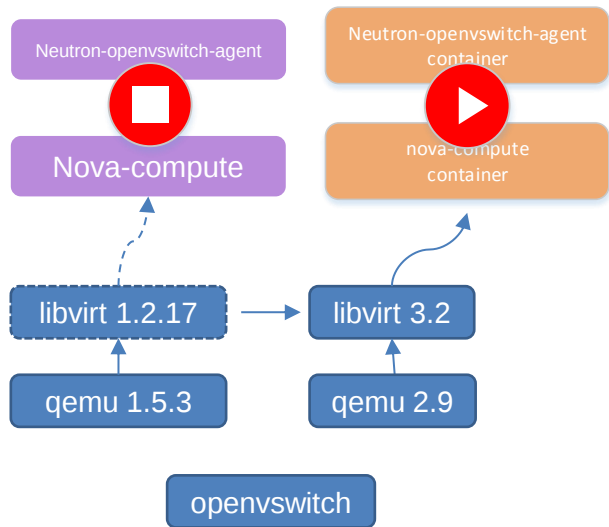
02

- 平滑升级Libvirt和qemu
(Nova requires QEMU version 2.1.0 or greater)

03

- 保证升级失败可回退

计算节点升级



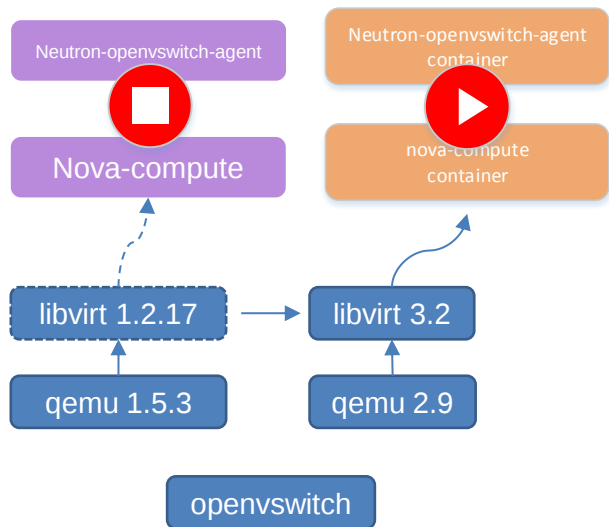
+ 01.Qemu

- 重新编译Qemu,改变新版本安装目录。
- 保留旧版本动态加载库,保证旧版本qemu进程正常运行。

+ 02.Libvirt

- 重新编译libvirt,识别新版本Qemu。
- 开启tcp远程访问Libvirtd服务。

计算节点升级



- Nova-compute 和 Neutron-openvswitch-agent采用kolla 容器升级。

- 通过docker-compose启动容器

neutron-openvswitch-agent:

container_name: "neutron_openvswitch_agent"

image: "neutron-openvswitch-agent:pike"

environment:

- neutron_plugin_agent=openvswitch

- KOLLA_CONFIG_STRATEGY=COPY_ALWAYS

privileged: True

net: "host"

volumes:

- "/root/upgrade/docker/volume/neutron-openvswitchagent/:var/lib/kolla/config_files/:ro"

- "/lib/modules:/lib/modules:ro"

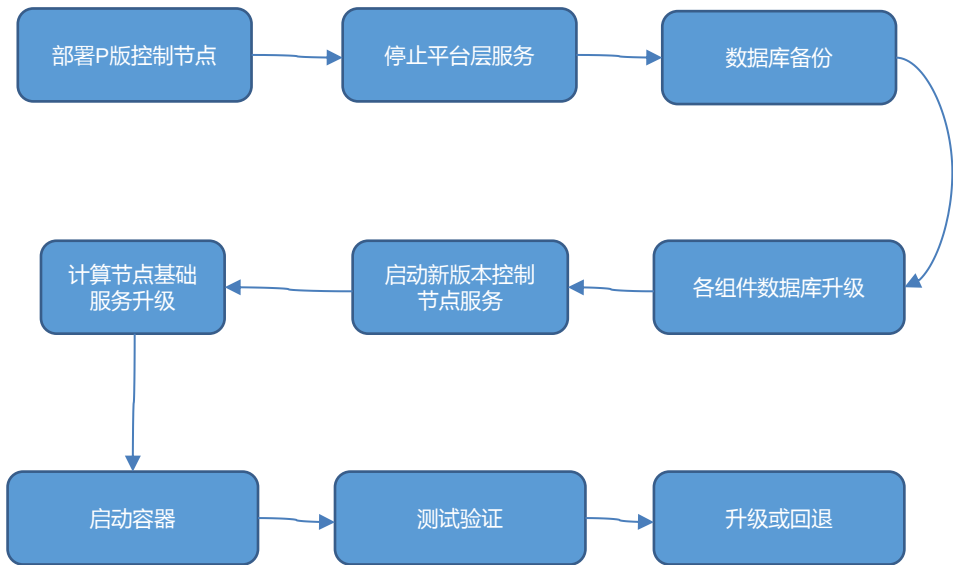
- "/run:/run:shared"

- "/var/log/kolla:/var/log/kolla:rw"

- 容器化部署方便后续向更高版本升级

TStack IaaS升级实践

OpenStack From Kilo to Pike



升级步骤多，若手动操作导致平台服务停止时间过长，且人为操作容易出错。自动化升级是必然选择。

目录

01

TStack简介

02

OpenStack升级实践 (From Kilo To Pike)

03

Ceph升级实践 (From Hammer to Jewel)

Ceph 升级实践

Ceph From Hammer to Jewel

佛系运维



能不升级吗？

一线运维同事来报：

每天都会有虚拟机进程自杀

BUG：librbd 会 double free同一内存。

某个周一早上九点，OSD随机自杀。

原因：scrub超时，触发osd suicide timeout。

Ceph 升级实践

Ceph From Hammer to Jewel

Ceph Server 端：

MON 组件升级：ceph-mon

OSD 组件升级：ceph-osd

Ceph Client 端：

librbd 库升级：ceph-common

升级第一准则

客户无感知！

Ceph 升级实践

Ceph From Hammer to Jewel

Server 端升级主要步骤:

1. 配置 yum 源，固定 Ceph 版本。
2. 依次升级MON，修改systemd脚本，重启MON。 *注1
3. 依次升级OSD，修改systemd脚本，重启OSD，等待PG a+c。 *注2
4. 配置ceph.conf，解除告警。*注3

划重点！

- *注1: H版和J版的服务启动停止指令不一样，需要enable服务。J版要求在 ceph:ceph 权限下运行进程。
- *注2: 全盘改权限耗时过久，使用 root 用户运行OSD，每个OSD节点升级时间约 5 分钟内，改权限需要 3-4 小时。
- *注3: 使用ceph tell mon.* injectargs 方式来更新配置。

Ceph 升级实践

Ceph From Hammer to Jewel

Client 端升级 主要步骤:

1. 配置 yum 源，固定 Ceph 版本。
2. 预留一台母机装好 Jewel 版客户端。
3. 将第一台宿主机上的虚机热迁移到空母机上。
4. 升级第一台宿主机客户端到 Jewel。
5. 将第二台宿主机上的虚机热迁移到第一台宿主机上。
6. 依次循环迁移完所有虚拟机。

Thank You

