

# Ceph 学习计划

## 一、熟悉 **Linux** 常用命令工具

ps mpstat vmstat free top sar iostat fdisk lsblk df lsscsi mkfs fsck mount umount blkdiscard htop

## 二、文件系统

**Linux** 哲学：一切皆文件

文件系统体系结构：VFS/XFS/EXT4/Btrfs 等

文件的存储结构：superblock,inode,dentry

了解常见文件操作接口

了解 fuse,udev 等

## 三、IO 系统

iostat,dstat,atop 等工具的使用

了解 IO 系统整体结构：通用块层、io 调度层、块设备驱动层、SCSI 子系统

了解 Buffered IO、RAW IO

## 四、内存管理

物理内存和虚拟内存

**Linux** 中的伙伴系统

PageCache 和 BufferCache

掌握内存信息查看命令：free,vmstat

## 五、进程管理

ps,top,kill,pkill,mpstat 命令的使用

- 熟悉 ceph 的 nova driver
- 熟悉 nova 的常用命令
- 能够通过 log 定位简单的问题

### **cinder**

- 熟悉 cinder 的工作流程
- 熟悉 ceph 的 cinder driver
- 熟悉 cinder 的常用命令
- 能够通过 log 定位 cinder 的问题

### **glance**

- 熟悉 glance 的工作流程
- 熟悉 ceph 的 glance driver
- 熟悉 glance 的常用命令

### **keystone**

- 了解 keystone 的工作机制

### **neutron**

- 了解 neutron 的整体架构

### **manila**

- 了解 manila 的整体架构
- 了解 manila 的网络结构

## mysql

- 了解常见的数据库操作命令

了解 **nginx** 和 **rabbit MQ** 的工作机制

理解 **UOS** 系统整体部署架构及网络规划

## 十三、传统存储基础知识

磁盘/**SSD** 基础知识及工作原理

**NAS/SAN** 传统存储协议

**NFS/CIFS/9p** 协议

**hdparm** 磁盘故障检测、**smartctl** 的使用

容灾和备份

了解基本的 **raid** 知识

了解 **SCSI** 协议

## 十四、监控系统

能够通过 **zabbix** 了解 **ceph** 集群的状况

熟悉 **UOS** 的 **ceph** 监控系统 **Diamond+Carbon+Whisper+Grafana**

能够通过 **ceph** 监控定位 **ceph** 的问题

## 十五、虚拟化技术

熟悉 **QEMU/KVM**

熟悉 **virtIO**

熟悉 **libvirt**

- dd

## Linux 性能分析工具

- systemtap
- fstrace/strace
- lttng, perf

## ceph 性能分析

- ceph -w
- ceph tell osd.\* heap stat
- ceph --osd-daemon osd.\* dump\_historic\_ops
- ceph osd perf
- ceph perf dump

## 十七、Ceph

### ceph 核心理论

- ceph 架构
- Rados(mon,osd,librados,rados bench 性能压测)
- crush 算法
- RBD(librbd,krbd,snapshot,RBD 应用场景)
- cephfs(mds,libcephfs,cephfs 应用场景,cephfs 和 docker,cephfs 和 hadoop,cephfs 和 VDI)
- cephfuse 的使用
- RGW(civetWeb,RGW,S3 接口,swift 接口, 对象存储访问场景)
- ceph cache tier 和 EC
- ceph 社区

尝试向社区提交 commit

跟进每个 release 的主要 BP

积极参与问题的讨论

参与社区的研讨会

试图将想法转化为 BP

### **ceph 运维**

- ceph 集群部署初始化
- 常见故障处理
- ceph 集群扩容 (加磁盘和加节点)
- ceph 节点硬件更换
- ceph 运维知识的传递
- 集群数据重平衡

### **自动化部署**

#### **puppet**

- 了解 puppet 的机制
- 了解 puppet-ceph 模块
- 能够根据需求对 puppet-ceph 模块进行修改

#### **ansible**

- 熟练使用 ansible 安装 ceph 集群

## Ceph 知识总结

### 条带化设置

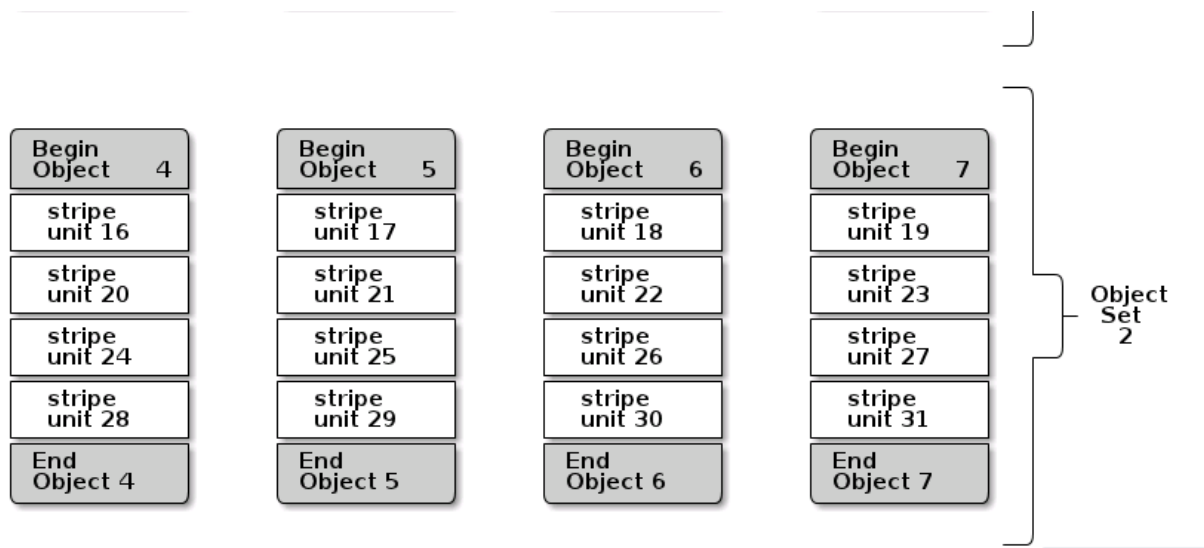


Figure 1: stripe\_unit\_count

如上图所示：

```
stripe_count = 4
```

(也可以表示为 1 个 object set 的大小) 即并发可顺序写的 thread，同时可向 4 个 object 写入，提高性能

```
stripe_unit = xxx
```

表示每次每块写入的最小单位

```
#102400M 大小 strip-unit 65535 byte strip-count 16
rbd create --image-format 2 disk1 --size 102400 --stripe-unit 65536 --stripe-count 16
```

[libradosstriper.cc](http://libradosstriper.cc)