使用map-reduce读写data

Placement group的原因，如何分布同一文件的不同object

Ceph消除了同时读写的bottleneck

对同一文件的读写，分配权限时加以限制，只可写某一object

当某个device坏掉时，怎么知道里面有些什么。 - monitor？

怎么知道什么时候该移动，比如有可能一会儿又好了？

可以把某个object标记为down然后看其log

SSD快但是不如hdd可靠，怎么才能知道primary怎么分配load到各个replica的

不考虑地域，还可以将replica放到离user比较近的地方

高rate的进入ssd，低rate的退出并重新走crush进行分配可保证uniform

假设有控制变量，应该长什么样？应该是个类似weight的东西？

用不同的rule即可保证数据的放置？同一个pool可以使用不同的rule？

提问，常访问的object占用多大比例

Crush commit policy – ssd commit – 1

1. Location
2. Server load

Metrics: data movement

Bandwidth

Osd performance

energy