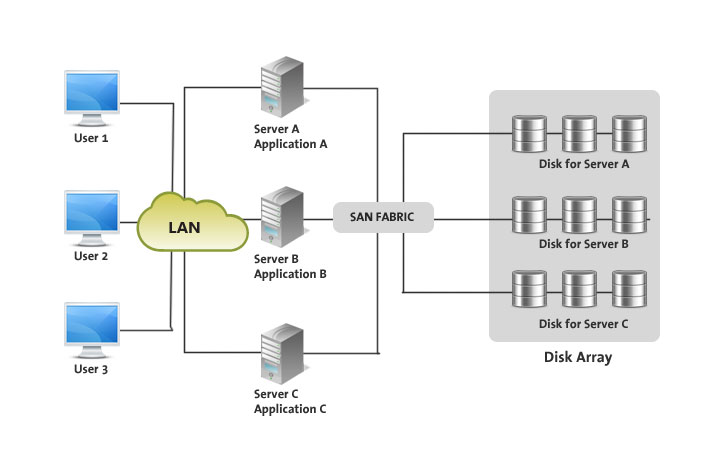
# 云计算的那些事之存储虚拟化

- 柳清风的专栏 - 博客频道 - CSDN.NET

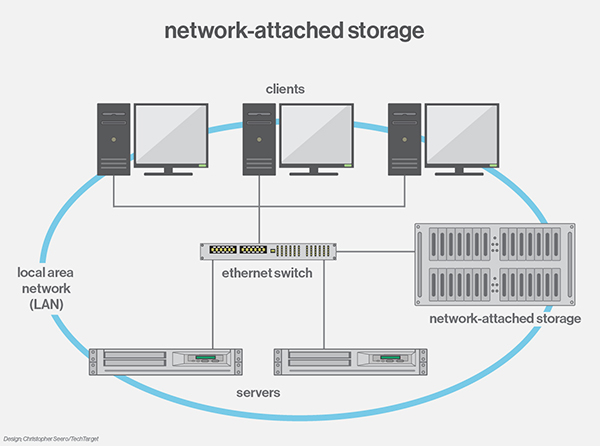
http://blog.csdn.net/u010278923/article/details/70990666

数据在整个计算机系统是最重要的一部分，也是最珍贵的。数据的存储一直是一个热议的话题，在没有[**云计算**](http://lib.csdn.net/base/hadoop)之前数据存储已经发展了很久，可以说和云计算没有特别关系，我这边将的云计算存储通常指的是亚马逊的S3存储或者EBS存储等，将统一的资源池划分给多个用户。

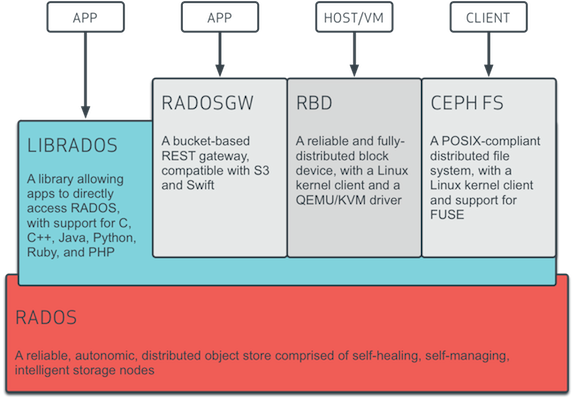
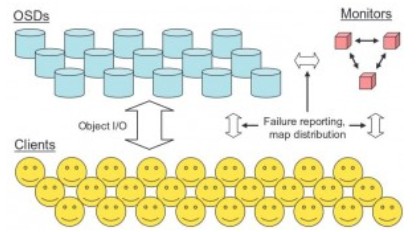
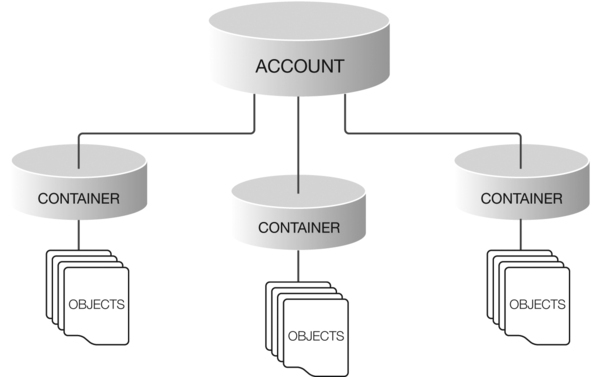
### SAN

那就先从存储说起，现在高端存储应该EMC、IBM和HDS的天下，这些年外置存储跟随着廉价磁盘不断提升容量和性能，推动了SAN网络、主机FC接口不断成熟，在数据中心变得很普遍，尤其在金融领域。   
  
SAN提供的是块存储，譬如磁盘阵列里面有10块I T的数据盘，然后可以通过做RAID或者逻辑卷（LVM）的方式划分出10个的数据盘，但这个10个数据盘已经和之前的物理盘不一样了，一个逻辑盘可能有第一个物理盘提供100G，第二个物理盘提供300G。对于[**操作系统**](http://lib.csdn.net/base/operatingsystem)来说，完全无法感知是物理盘还是逻辑盘，这是存储资源池的理念。通过RAID或者LVM不仅可以提供数据保护还能够重新划分盘的大小，提高读写速率。但SAN也不是毫无缺点，它价格也是比较昂贵的，光纤口，光纤交换机价格高，所以才有了IPSAN存储，通过IP协议承载存储协议；无法提供数据共享，一个盘只能挂给一个主机，所以这就有了NAS存储。

### NAS

NAS是文件存储，文件存储相比块存储最大的优势是能共享数据，它基于标准的网络协议，SAN是有自己一套存储协议的。常见的NAS包括NFS、FTP和HTTP文件服务器等，由于这种设备通常都有一个IP，所以一般客户机充当数据网关服务器可以直接对其访问。NAS建立在传统网络之上，所以可以更远距离的传输，并且NAS具有安装容易易于维护的特点，但其速度通常要比SAN慢很多。   


### 分布式存储

伴随着x86性能提升，以x86芯片构建的小型存储系统，在中端存储领域开始崭露头角。通过将X86本地的磁盘利用起来构建一个大存储集群。这里不得介绍一个和[**OpenStack**](http://lib.csdn.net/base/openstack)结合紧密的分布式存储ceph，下图是ceph官网的一个整体模块图，它提供了CEPH FS文件存储系统和POSIX接口、对象存储以及最常用的快存储。   
   
它的基石是下面的RADOS，下面就是系统组件。   
   
CEPH OSDs：CEPH的OSD（Object Storage Device）守护进程。主要功能包括：存储数据、副本数据处理、数据恢复、数据回补，平衡数据分布。并将数据相关的一些监控信息提供给CEPH Moniter，以便CEPH Moniter来检查其他OSD的心跳状态。一个CEPH存储集群，要求至少两个CEPH OSDs，才能有效的保存两份数据。注意，这里的两个CEPH OSD是指运行在两台物理服务器上的，并不是在一台物理服务器上开两个CEPH OSD的守护进程。   
Moniters：CEPH的Moniter守护进程，主要功能是维护集群状态的表组，这个表组中包含了多张表，其中有Moniter map、OSD map 、PG(Placement Group) map、CRUSH map。   
MDSs：CEPH的MDS (Metadata Server)守护进程，主要保存的是CEPH Filesystem的元数据。注意，对于CEPH的块设备和CEPH对象存储都不需要CEPH MDS守护进程。CEPH MDS 为基于POSIX文件系统的用户提供了一些基础命令的执行，比如ls、find等等，这样可以很大层度上降低CPEH 存储集群的压力。   
还有一个开源的对象存储就是openstack的[**Swift**](http://lib.csdn.net/base/swift)，swift的初衷就是用廉价的成本来存储容量特别大的数据，swift使用容器来管理对象，允许用户存储、检索和删除对象以及对象的元数据，而这些操作都是通过用户友好的RESTful风格的接口完成。swift有三层数据模型，Account、[**Container**](http://lib.csdn.net/base/docker)、Object，Account划分了不同账号命名空间，同一个Account内Container不能重名，当然不同的Account下可以重名。   
   
数据分布式通过一致性hash来维护的，这个以后再深入讲解