Q：有一个场景：每天有近百GB数据增加，数据内容有WORD文档和图像等多种类型。用什么存储或文件系统比较合适？

A: HDFS、HBase、Hive不太适合存文档、图片大小的文件，HDFS适用于存大文件，后两者适用于数据库场景，每天近百GB的文档、图像，那只有用SWIFT了。

Q：SWIFT与Ceph、GlusterFS、MongoDB/GridFS之间的比较呢？

A：SWIFT是专为这种场景设计的，性能可以调节。

Q：如果图像比较大的情况，一张图接近GB时，哪个合适？

A：我仍然推荐用SWIFT，SWIFT官方认为对几个G的支持都没有问题，但是我认为处理几个G的大文件性能上可能会比HDFS差，因为没有条带化。但是考虑到文档、图片类型存储，你会遇到很多几兆、几十兆的，这些文件的存储，HDFS就不如SWIFT了。从适用范围的角度来看，咱们这个场景里还是首选SWIFT。

Q：GlusterFS它的优点在哪里，在文件数量太大时Gluster会出现性能问题正常么？

A：GlusterFS，它的优点在于如果你需要使用POSIX接口，那GlusterFS是一个很好的选择，但是Gluster在文件数量、目录数量太大的情况下会出现性能衰减，这是文件系统本身的特点决定的，不论是哪种POSIX兼容的文件系统，都存在类似的问题。

Q：SWIFT、Ceph、GlusterFS、MongoDB那个更好？

A：关于SWIFT、Ceph、GlusterFS、MongoDB之间的比较，没有孰优孰劣，只有是不是适合。MongoDB是NoSQL数据库，和HBase是同一类，但是和其他几个分布式存储显然不是同一类的，不适合用于此类场景。

Q：Ceph，SWIFT分别适合什么样的情况？

A：如果你只需要用块存储，那当然是Ceph，如果只需要用对象存储，那当然是SWIFT。各自有各自擅长的地方，不过现在因为Ceph也支持对象存储，SWIFT和Ceph又是OpenStack社区非常频繁提到的两种存储形式，所以，有很多人会把SWIFT和Ceph在一起比较。

如果只要用对象存储，就选择SWIFT；如果只要用块存储，那就Ceph；即要用对象存储又要用块存储的场合，是用SWIFT还是Ceph呢？我们一般是这样推荐的：

1.如果节点数量很大，推荐用Ceph单独做块，用SWIFT做对象存储，因为在节点数量较大时，Ceph的维护成本比SWIFT要高得多，大多数场景实际应用的时候会发现，大部分数据都可以放到对象存储上（这一点如果有疑问，欢迎随时讨论）；

2.如果节点数量少，那就用Ceph统一搞定，因为一般认为生产环境中最小的分布式存储应当有五个节点，所以，如果节点数量少于十个或者刚到十来个，那构建两个分布式存储显然是不理想的（考虑到空间划分问题）；

3.如果团队里有牛人能轻松解决Ceph大规模部署问题，那就果断用Ceph；

4.如果希望对象存储能够和OpenStack其他项目无缝结合，如果希望实现多租户，果断用SWIFT来实现对象存储。

Q：对象存储的本质是什么？

A：对象存储本身有两个概念，一种是九几年提出的，针对NAS性能问题提出的一种新的存储架构，其实Gluster也是那种对象存储。另一种是06年亚马逊推出S3以后逐渐被人接受的对象存储，特点是RESTful接口和扁平的数据组织形式。我们说的CEPH和SWIFT提供对象存储，指的是后者，但是CEPH的底层RADOS，又是前一类对象存储。

Q：如果现在有一堆文件要存，如何选择是用对象还是用文件系统呢？

A：所谓文件系统的本质是POSIX接口，“对象”这个名词是做对象存储的人为了把自己做的东西和文件系统区分开而用的术语，把存在对象存储里的文件叫做“对象”，所以选择文件系统还是对象存储，跟你把这堆数据称作对象还是文件并没有直接关系，而是要看你是需要POSIX还是RESTful HTTP接口，是需要目录结构还是适用扁平数据管理结构。