基于Azure托管磁盘配置高可用共享文件系统

* 背景介绍

在当下，共享这个概念融入到了人们的生活中，共享单车，共享宝马，共享床铺等等。其实在IT界，共享这个概念很早就出现了，通过SMB协议的windows共享目录，NFS协议的网络文件系统等等。在Azure平台，也提供了Azure共享文件服务供客户访问。但是不管是CIFS,NFS，还是Azure File Share都有其不足之处。比如CIFS,NFS无法实现高可用，一旦CIFS server或者NFS server不可用，整个共享目录就变得不可用。那么如何实现共享文件系统的高可用呢？

今天就给大家介绍一下：

基于Azure托管磁盘+Gluster分布式文件系统来构建高可用的共享文件系统。

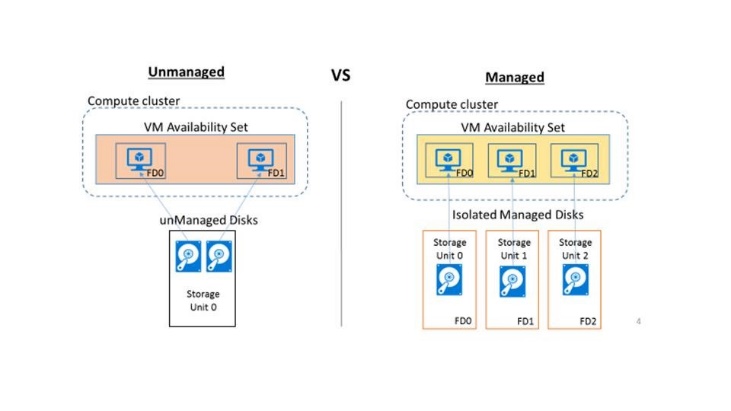
* 名词解释

**Azure托管磁盘**

在托管磁盘发布之前，Azure的磁盘分为普通存储和高端存储，在创建存储账号之后，来存储虚拟机的磁盘（VHD页文件）。而且对于磁盘的IOPS都有所限制，比如不能超过20,000 IOPS/账户。而且存储单元是没有高可用的功能。一旦存储单元出现故障，那么该存储单元下的磁盘都会受到影响。在这种架构下，基于这些磁盘部署的虚拟机即使部署在同一个高可用集合内，依然会变得不可用。同理Azure File Share也会受到上述的限制。

最近Azure发布了托管磁盘（中国区还没正式上线），就完美的解决了上述的问题。其优点参考如下：

* + **简单且可缩放的虚拟机部署，**在使用托管磁盘时，磁盘的IOPS不再受到存储账号的IOPS上线限制。在进行扩展时，无需额外创建存储账号等。
  + **可用性集更加可靠，**托管磁盘彼此之间完全隔离以避免单点故障，通过高可用集，自动将磁盘防止在不同的存储单元。如果该存储单元不可用，那么只能影响到使用该磁盘的虚拟机实例。

[](https://msdnshared.blob.core.windows.net/media/2017/03/92.jpg)

* + 其他优点及详情，参考链接如下：

<https://docs.microsoft.com/zh-cn/azure/storage/storage-managed-disks-overview>

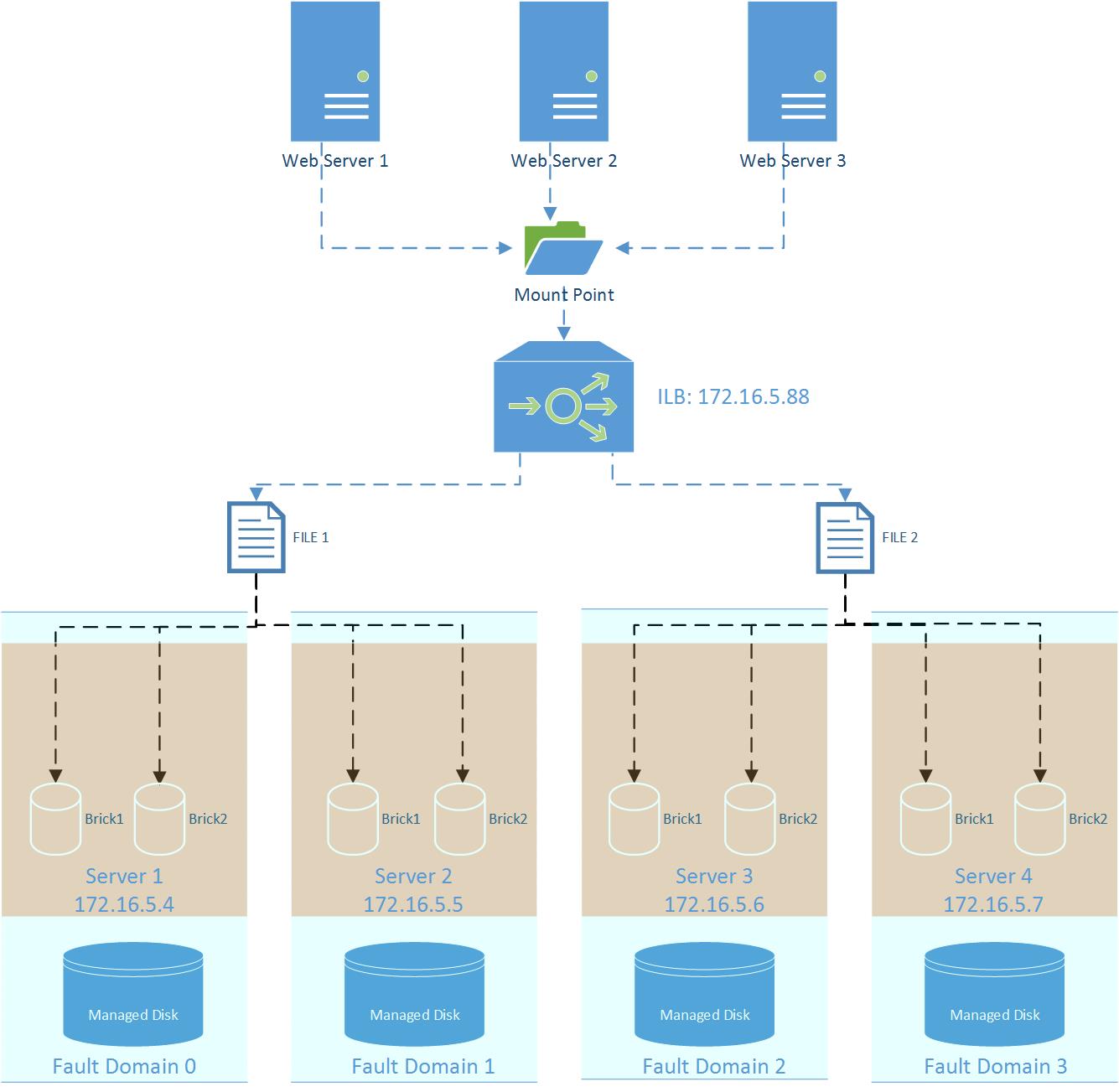
**GlusterFS**

**GlusterFS**是一个开源的可扩展分布式文件系统，适用于云计算，流媒体等互联网应用。与**Lustre、MooseFS、CEPH**并列成为四大开源分布式文件系统。由于GlusterFS新颖的KISS（KeepIt as Stupid and Simple）的系统[架构](http://lib.csdn.net/base/architecture)，使其在扩展性、可靠性、性能、维护性等方面具有独特的优势，目前开源社区风头有压倒之势，国内外有大量用户在研究、[测试](http://lib.csdn.net/base/softwaretest)和部署应用。有优势，也势必有其短板之处，比如处理元数据和海量小文件的效率和性能表现一般，所以在选择时，一定要评估自己的应用场景是否适用。

* 架构简介

基于上述对托管磁盘的分析，我们可以看到，从磁盘层面，我们已经做了高可用，保证同一可用集合内，不会所有的磁盘（虚拟机）同时不可用。之后，通过Gluster分布式文件系统的备份卷，即使有磁盘因为计划内的维护升级变得不可用，从文件系统层面，依然可以提供数据访问的可用性。

再加上Azure平台面向内部负载均衡器（ILB）,从前端访问不用担心节点故障时，需要手动切换挂载点。



* 实施方案

**部署基于托管磁盘的虚拟机**

1. 点击链接：<https://portal.azure.cn登录Azure>门户管理网站。
2. 根据向导，部署4台基于ARM的CentOS 7.3虚拟机服务器。

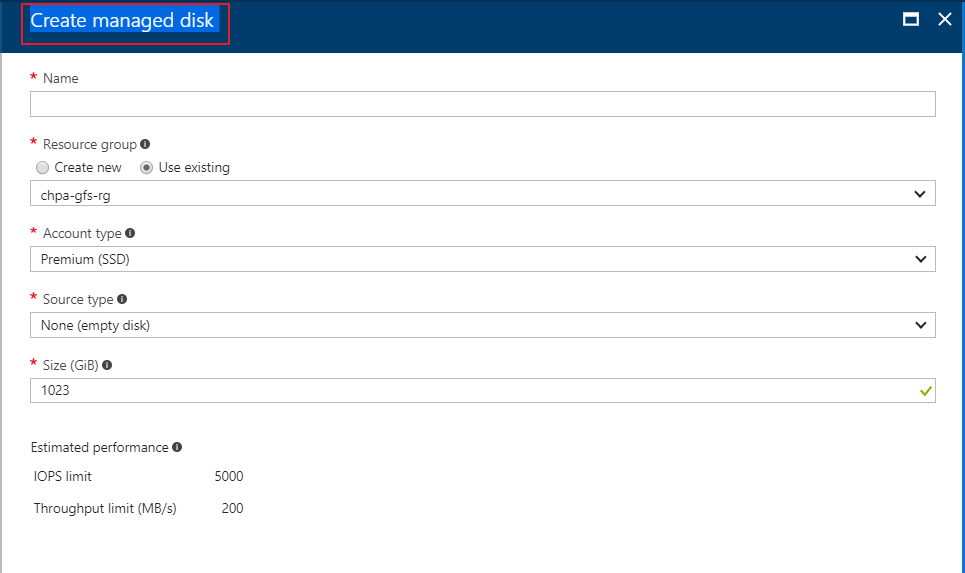
gfs-node1 ( 172.16.5.4 ）

gfs-node2 ( 172.16.5.5 ）

gfs-node3 ( 172.16.5.6 ）

gfs-node4 ( 172.16.5.7 ）

1. 在创建时，一定要把虚拟机放在同一高可用集内。
2. 虚拟机创建完毕之后，给每一台虚拟机附加一块1TB的托管磁盘。



1. 登录虚拟机，通过fdisk -l检查托管磁盘是否附加上。
2. 当托管磁盘附加上之后，执行命令进行初始化：
3. # mkfs.xfs /dev/sdc
4. # mkdir /sdc
5. # mount /dev/sdc /sdc
6. 在之后，/sdc就作为Gluster文件系统里的Brick来使用。

**安装配置Gluster**

1. **登录每台虚拟机，执行下列操作，安装配置Gluster服务**
2. # yum install centos-release-gluster37
3. # yum install glusterfs gluster-cli glusterfs-libs glusterfs-server
4. 启动Gluster服务：# service glusterd start
5. 创建存储池：# gluster peer probe <hostname>
6. 查看存储池状态：# gluster peer status

**创建管理Gluster卷**

1. 为了兼顾可用性和性能，创建分布式条带化备份卷。命令如下：
2. # gluster volume create dsr\_vol stripe 2 replica 2 172.16.5.{4,5,6,7}:/sdc force
3. 关于更多卷的类型和说明，[**点击这里**](https://access.redhat.com/documentation/en-US/Red_Hat_Storage/2.1/html/Administration_Guide/sect-User_Guide-Setting_Volumes-Distributed_Striped_Replicated.html)
4. 执行命令启动卷：# gluster volume start dsr\_vol
5. 执行命令查看卷状态：# gluster volume status

**配置ILB实现负载均衡**

1. 登录Azure门户管理网站
2. 根据向导创建负载均衡gfs-ilb
3. 指定前端IP地址为：172.16.5.88
4. 指定后端IP池：172.16.5.4, 172.16.5.5, 172.16.5.6, 172.16.5.7
5. 指定健康探测端口：22
6. 指定负载均衡规则：
7. Tcp/udp-24007/24008
8. 上述端口为Gluster服务端口
9. Tcp 49152为数据端口，一个卷对应一个端口，其实为49152，依次加1.

**客户端挂载Gluster卷**

1. 执行命令挂载Gluster卷：

# mount.glusterfs 172.16.5.88:/dsr\_vol /dsr\_data

至此，你就拥有了几乎不会offline的共享目录！