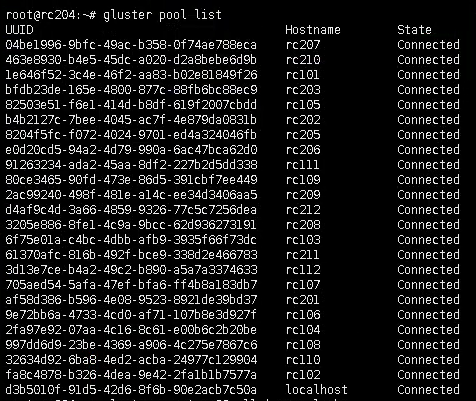
柳州arm集群检查分析

1. 现象

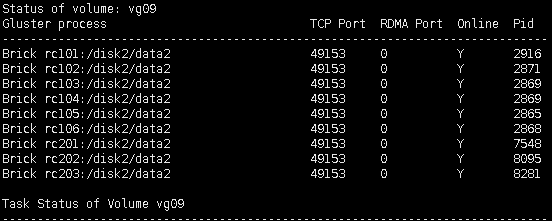
用户29号左右使用集群（vg09）发现读写较慢，性能较差

1. 集群状态

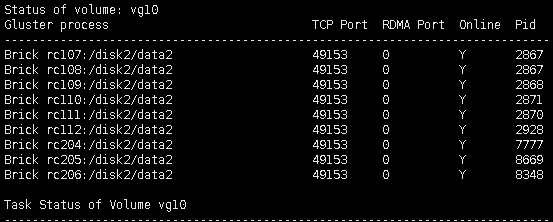


1. 卷状态

Vg09

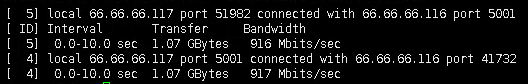


Vg10

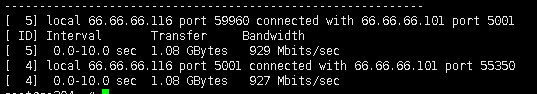


1. arm节点间通信

同一服务器两节点

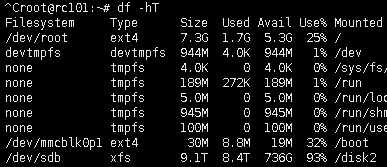


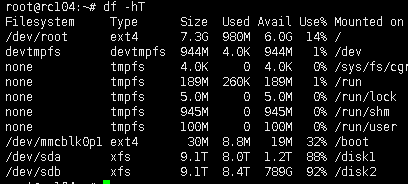
两台服务器两节点

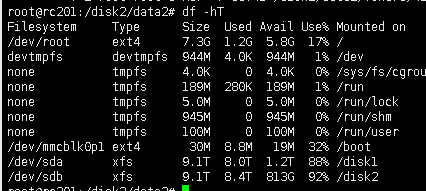


1. 磁盘容量（vg09）

3子卷的磁盘容量均超过了90%





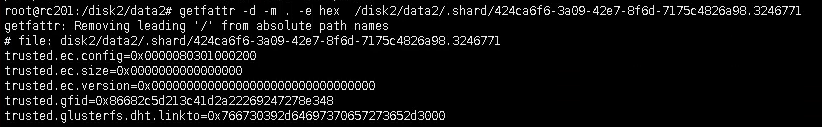


1. 分析

Glusterfs本身设计有在处理分布式节点不均衡时的链接技术，假设某个文件hash到某个brick，而brick容量不足会创建一个T文件（链接文件），然后在其他节点真实的创建文件写入数据。由于磁盘在容量超过90%时性能会出现断崖式下跌，因此gluster设置了一个阈值90%，当节点容量超过此值时开始使用T链接方式存储文件，客户端根据返回的扩展属性获取真实文件所在的brick节点，最终客户端在真实brick节点上读写。如图：

【以 .3246771为例】





覆盖写入该.3246771数据时会多次查找本身gfid，失败后再跳转到真实地址，导致延时较大，性能表现较差。

日志：





1. 建议

我们的环境磁盘容量均衡分布，并且开启了shard功能，因此建议用户使用在90%以下（建议值80%-85%），防止此现象影响读写性能。

Gluster本身也有对其值的修改设置：

gluster v set volname cluster.min-free-disk 10% (默认) ，可修改为5%