

解析Linux新技术对象存储文件系统（3）PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/144/2021_2022__E8_A7_A3_E6_9E_90Linu_c103_144604.htm 2.3 Lustre对象存储文件系统

Lustre对象存储文件系统就是由客户端（client）、存储服务器（OST，Object Storage Target）和元数据服务器（MDS）三个主要部分组成。Lustre的客户端运行Lustre文件系统，它和OST进行文件数据I/O的交互，和MDS进行命名空间操作的交互。为了提高Lustre文件系统的性能，通常Client、OST和MDS是分离，当然这些子系统也可以运行在同一个系统中。其三个主要部分如图1所示。图1 Lustre文件系统的组成

Lustre是一个透明的全局文件系统，客户端可以透明地访问集群文件系统中的数据，而无需知道这些数据的实际存储位置。客户端通过网络读取服务器上的数据，存储服务器负责实际文件系统的读写操作以及存储设备的连接，元数据服务器负责文件系统目录结构、文件权限和文件的扩展属性以及维护整个文件系统的数据库一致性和响应客户端的请求。Lustre把文件当作由元数据服务器定位的对象，元数据服务器指导实际的文件I/O请求到存储服务器，存储服务器管理在基于对象的磁盘组上的物理存储。由于采用元数据和存储数据相分离的技术，可以充分分离计算和存储资源，使得客户端计算机可以专注于用户和应用程序的请求；存储服务器和元数据服务器专注于读、传输和写数据。存储服务器端的数据备份和存储配置以及存储服务器扩充等操作不会影响到客户端，存储服务器和元数据服务器均不会成为性能瓶颈。Lustre的全局命名空间为文件系统的所有客户端提供了一个有效的全局

唯一的目录树，并将数据条块化，再把数据分配到各个存储服务器上，提供了比传统SAN的"块共享"更为灵活的共享访问方式。全局目录树消除了在客户端的配置信息，并且在配置信息更新时仍然保持有效。

三、测试和结论

1、Lustre iotest

测试针对对象存储文件系统，我们对Lustre文件系统作了初步测试，具体配置如下：3台双至强系统：CPU：1.7GHz，内存：1GB，千兆位以太网
Lustre文件系统：lustre-1.0.2
Linux版本：RedHat 8
测试程序：iotest
测试结果如下：

写(MB/s/thread)	单线程	两个线程	Lustre 1个OST	2个OST	1个OST 2个OST	NFS
21.7	50	12.8	24.8	12	5.8	

从以上的测试表明，单一OST的写带宽比NFS好，2个OST的扩展性很好，显示strip的效果，两个线程的聚合带宽基本等于饱和带宽，但lustre客户方的CPU利用率非常高（90%以上），测试系统的规模（三个节点）受限，所以没有向上扩展OST和client数量。另外，lustre的cache对文件写的性能提升比NFS好。通过bonnie初步测试了lustre的元数据处理能力，和NFS比，文件创建速度相对快一些，readdir速度慢。

100Test 下载频道开通，各类考试题目直接下载。详细请访问 www.100test.com