

解析Linux新技术对象存储文件系统（1）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\\_144611.htm](https://www.100test.com/kao_ti2020/144/2021_2022__E8_A7_A3_E6_9E_90Linu_c103_144611.htm) 随着高性能计算由传统的主机方式向网络化集群演变，传统的基于主机的存储架构已逐渐向网络化存储发展，计算和存储分离的趋势越来越明显。针对SAN和NAS的不足，国际上已开展针对Linux集群的新型文件系统——对象存储文件系统的研究，本文重点论述了存储对象文件系统的架构、技术特点，并针对Lustre对象存储文件系统进行了初步测试，结果表明对象存储在可扩展性、性能、易用性等方面都有显著提高，随着网络化存储技术的不断成熟，对象存储文件系统将成为重要的发展方向。

一、引言 高性能计算已由传统的主机方式逐渐向集群方式演变，如TOP500中，1998年只有2台系统是集群方式，而到2003年已有208台为集群系统。随着高性能计算体系结构的发展变化，传统的基于主机的存储架构已成为新的瓶颈，不能满足集群系统的需求。集群的存储系统必须有效解决两个主要问题：（1）提供共享访问数据，便于集群应用程序的编写和存储的负载均衡；（2）提供高性能的存储，在I/O级和数据吞吐率方面能满足成百上千台规模的Linux集群服务器聚合访问的需求。目前，网络化存储已成为解决集群系统高性能存储的有效技术途径。国际上主要有两类网络化存储架构，它们是通过命令集来区分的。第一类是SAN（Storage Area Network）结构，它采用SCSI块I/O的命令集，通过在磁盘或FC（Fiber Channel）级的数据访问提供高性能的随机I/O和数据吞吐率，它具有高带宽、低延迟的优势，在高性能计算

中占有一席之地，如SGI的CXFS文件系统就是基于SAN实现高性能文件存储的，但是由于SAN系统的价格较高，且可扩展性较差，已不能满足成千上万个CPU规模的系统。第二类是NAS（Network Attached Storage）结构，它采用NFS或CIFS命令集访问数据，以文件为传输协议，通过TCP/IP实现网络化存储，可扩展性好、价格便宜、用户易管理，如目前在集群计算中应用较多的NFS文件系统，但由于NAS的协议开销高、带宽低、延迟大，不利于在高性能集群中应用。针对Linux集群对存储系统高性能和数据共享的需求，国外已开始研究全新的存储架构和新型文件系统，希望能有效结合SAN和NAS系统的优点，支持直接访问磁盘以提高性能，通过共享的文件和元数据以简化管理，目前对象存储文件系统已成为Linux集群系统高性能文件系统的研究热点，如Cluster File Systems公司的Lustre、Panasas公司的ActiveScale文件系统等。Lustre文件系统采用基于对象存储技术，它来源于卡耐基梅隆大学的Coda项目研究工作，2003年12月发布了Lustre 1.0版，预计在2005年将发布2.0版。Lustre在美国能源部(U.S.Department of Energy：DOE)、Lawrence Livermore 国家实验室，Los Alamos国家实验室，Sandia 国家实验室，Pacific Northwest国家实验室的高性能计算系统中已得到了初步的应用，IBM正在研制的Blue Gene系统也将采用Lustre文件系统实现其高性能存储。ActiveScale文件系统技术来源于卡耐基梅隆大学的Dr. Garth Gibson，最早是由DARPA支持的NASD（Network Attached Secure Disks）项目，目前已是业界比较有影响力的对象存储文件系统，荣获了ComputerWorld 2004年创新技术奖。

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](http://www.100test.com)