Panasas

概述

Panasas公司于1999年由卡内基梅隆大学的Garth Gibson等人创建。其提供的并行文件系统PanFS是通用的并行文件系统，目前其主要的应用领域和luster类似，都是用于高性能计算（HPC）领域，这就要求其具有高可用，可扩展行，高性能，通过分布式锁提供的强一致性。

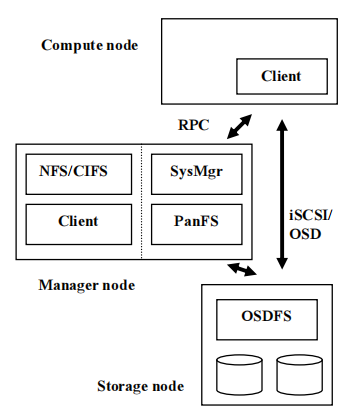
解决的问题

Panasas存储系统是被设计为一个具有极强的可伸缩性的存储集群。同时，它也是一个生产系统，为世界上一些最大的计算集群、科学实验室、地震数据处理、数字动画工作室、计算流体力学等提供文件服务。

在这些环境中，数百或数千个文件系统客户端共享数据，并在文件系统上生成非常高的聚合I/O负载。而 Panasas就为了支持数千个客户设计的、存储容量超过1兆字节的分布式存储系统。

设计架构

Panasas系统组件



Client:客户机（Client）是运行在Linux内核中的可安装内核模块。 内核模块实现了标准的VFS接口，使客户端主机能够挂载文件系统并使用存储系统的接口。

OSDFS:OSDFS是存储节点使用实现对象存储原语的专用本地文件系统。 它们实现iSCSI目标和OSD命令集。 OSDFS对象存储和iSCSI target/OSD命令处理器是内核模块。 OSDFS关注传统的块级文件系统问题，如有效的磁盘臂、媒体管理（即错误处理）、高吞吐量以及OSD接口。

SysMerg：集群管理器(SysMerg)维护全局配置，并控制存储集群中的其他服务和节点。 有一个关联的管理应用程序，它同时提供命令行接口(CLI)和HTML接口(GUI)。 这些都是运行在管理器节点子集上的用户级应用程序。 集群管理器涉及存储集群中的成员资格、故障检测、配置管理和软件升级和系统重新启动等操作的总体控制。

Pan FS：元数据管理器(Pan FS)实现文件系统语义，并管理跨对象存储设备的数据条带。 这是一个运行在每个管理器节点上的用户级应用程序。 元数据管理器涉及分布式文件系统问题，如安全的多用户访问、保持一致的文件和对象级元数据、客户端缓存一致性以及从客户端、存储节点和元数据服务器崩溃中恢复。 容错基于本地事务日志，该日志被复制到不同管理器节点上的备份。

NFS/CIFS：NFS和CIFS服务为不能使用Linux可安装文件系统客户端的主机提供对文件系统的访问。 NFS服务是运行在内核中的标准FreeBSD的 NFS服务器的调优版本。 CIFS服务基于Samba并在用户级别运行。 反过来，这些服务使用本地实例文件系统客户端，它们运行在FreeBSD系统的内核中。 这些网关服务运行在每个管理器节点上，提供集群NFS和CIFS服务。

关键技术

对象存储、容错、缓存和缓存一致性以及

简化的管理模型：Panasas存储系统将自己呈现为具有POSIX接口的文件系统，并隐藏存储管理的大部分复杂性。 客户端对整个系统有一个安装点。 /etc/fstab文件引用集群管理器，客户端从中学习元数据服务实例的位置。 管理员可以在系统上线时添加存储，自动发现新资源。 为了管理可用的存储，我们引入了两个基本的存储概念：称为刀片集（BladeSet)的物理存储池和称为卷的逻辑配额树。

总结