

基于 GlusterFS 的数据分布策略研究

陈辉

Outline

Introduction

GlusterFS 简介

概述

本研究致力于在分布式（或高性能计算）环境下，针对不同工作负载进行冷热数据动态负载均衡。主要思想是：

- ▶ 对用户程序实时监控，抽取与 IO 操作相关性较大的行为特征；
- ▶ 根据用户 IO 读写历史对文件系统中的数据进行聚类；
- ▶ 将用户、用户程序特征与聚类后的数据进行关联；并根据访问时序对聚类后的数据进行粗略排序。
- ▶ 针对正在运行或即将运行的用户程序特征，将关联的数据（主要针对小文件）按时序加载至 SSD。

GlusterFS 简介

GlusterFS 架构简述

Some examples of commonly used commands and features are included, to help you get started.

GlusterFS tiering 模块

日志分析

日志分析能够比较直接地比对分析用户程序运行与系统层 IO 活动的逻辑关联性.