



# PANASAS 预自我管理功能介绍

技术白皮书

## 内容

介绍	3
监控和修复	3
ActiveScan	3
ActiveGuard	3
监控和报警	4
风扇，电源和电池	4
温度	4
PanActive 管理	4
“热点”计算节点监控	5
SNMP	5
负载均衡	5
被动对象负载均衡	5
主动对象负载均衡	5
读结合	5
数据重组负载均衡	6
硬件冗余	6
风扇	6
供电系统	6
电池备份系统	6
ActiveScale 镜像	7
ActiveUpgrade 升级	7
摘要	7

---

一种崭新水平的可靠性和可用性----  
通过连续地监控；智能地预报和预空闲  
机制，使集群存储系统能够保持连续地  
性能。

---

## 介绍

Panasas ActiveScale™ 操作环境提供了一个基于对象的存储架构，它能够通过将数据动态地分布在多个智能存储刀片上，实现数据的并行存取。Panasas 存储刀片存储数据对象。Panasas 指挥刀片控制和监控计算节点和存储刀片之间的数据存取，这种独特的能力使得Panasas 存储集群系统具备强大的高性能并行I/O能力，是高性能Linux计算机集群运行并行计算达到空前性能的理想集群存储系统。

Panasas 预自我管理包含了一系列关键技术，他们协同工作，能够监控设备自身环境，并且自愈，这样使得Panasas 存储集群连续工作并能够提供最高的性能给运行集群应用的用户。这些技术主要体现在4个领域：监控和修复；监控和报警；负载均衡和冗余。

所有这些技术内部相关联并且支持ActiveScale操作环境，Panasas 集群存储系统提供最高的连续性能和集群存储从未有的可管理性能。Panasas 已经扩展了产品线，支持包括交互式的应用，给所有类型的计算集群提供一个同一的产品平台。这使得像预自我管理这样的功能变得更加重要。统一的存储系统可能达到Petabyte 级别的扩展，这个时候，可靠性变得尤为重要。这有别于传统的使用在HPC环境中的基于SAN 的包含三个独立的管理环节的存储系统。Panasas ActiveStor 提供给用户是一个包含预自我管理的集成系统，从而显著地降低了投资和总体拥有成本。

## 监控和修复

ActiveScan 和 ActiveGuard 是Panasas ActiveScale操作环境版本3.0的新功能。

### ActiveScan

ActiveScan 这项 Panasas 的技术能够连续监控存储刀片上的内容，并且在检测到问题发生时及时采取行动来保证存储集群系统连续工作。ActiveScan 持续监控数据对象，RAID校验，磁盘介质和磁盘驱动器属性。如果检测到磁盘的块存在潜在的问题，之后数据就会被移动到同一磁盘的空闲区域。通过使用统计学的分析方法来分析磁盘驱动器的SMART属性，潜在的磁盘故障就会被预先发现，因此在磁盘真正故障前，系统就会预先采取行动以保护数据。如多磁盘故障被预先发现，然后数据目标就会抢先移动到其它存储刀片上，这样就避免了数据重构。替换的存储刀片插入后，指挥刀片将数据对象并行的移到新的刀片上。这时，新插入的刀片的空间可以和其它刀片一样使用。当存储刀片出现故障时，只对损坏的部分进行数据重构。这样就大大降低了数据重构的时间，保证了存储集群系统的性能。

### ActiveGuard

Panasas ActiveGuard包含两个主要的高可用性技术。指挥刀片切换和存储网络自动切换。当发现指挥刀片失效时，所有的卷的服务会自动切换到备份的智能刀片上。

所有的智能刀片配置成活动/活动状态，也就是说，不需要“Spares”。这样，所有的指挥刀片都在充当着集群存储系统管理者的角色。不需要中止服务，系统可以不被中断的连续工作。所有使用 DirectFLOW 数据路径的计算节点都能够做连续的数据处理而一无所知。

Panasas ActiveStor 系统基于千兆以太网和iSCSI。每个机箱中的所有的刀片均含有两条完全的数据路径。当ActiveGuard 的可选服务启动后，网络可以无缝地切换到另外一条数据路径上，而此时，所有前端计算节点的应用可以连续地进行数据存取。

这两个主要的功能并结合标准的基于对象的RAID包含在 ActiveScale 操作环境当中，增强了 Panasas存储集群系统的高冗余性能，最高的操作可用性，能够满足客户运行他们的关键应用。

## 监控和报警

在ActiveScale操作环境中包含了许多功能，它几乎可以监控Panasas 存储集群的所有部件，并且适时地给系统管理员报警。

### 风扇，供电系统和电池

机箱里的所有电子元器件都会被监控，并且会和正常的状态比较。一旦有属性改变或者超过了限定值的情况发生，系统就会给管理员发送警报，这样系统就能够及早采取动作。避免真正的故障产生。例如：风扇不工作，电源故障和电池问题等。

### 温度

保证存储机箱和机架系统的温度，是系统高可靠性的一个最重要的因素。这就是为什么 Panasas 在保证机箱正确温度方面做了非常细心的设计。系统监控机箱温度和温度变化率，在真正错误出现前采取措施。在动作产生前先给管理系统发送信息。如果温度持续上升到了不安全的水平，而这时还没有采取措施的话，这时系统会执行完全关机，以保护数据。

## PanActive 管理

The Panasas ActiveStor存储集群系统内置了Web 图形界面和命令行 (CLI) 管理界面。系统使用一个单一的管理界面管理所有的刀片上的数据，无论存储系统是一个单一的机箱或者是 Petabytes级别的存储系统。因此，最小化了您的总体拥有成本。PanActive Manager™ 同时提供安全管理。使得系统管理员可以提供一致的控制存取和权限控制，这包含所有的协议，DirectFLOW™, NFS and CIFS甚至还可以控制对根文件系统的IP地址级别的控制。所有的监控和控制功能都可以通过Web图形用户接口(GUI) 或者 CLI来实现不同的功能。通过GUI 来实现存储资源管理(SRM) 生成报告，包括历史性能报告，资产报告和历史的以及预监控分析报告。

### “热”计算节点监控

PanActive Manger 在存储集群上连续监控计算节点和应用进程。它管理并且连续更新一个表，这个表里记录着使用DirectFLOW 或NFS协议最活跃的计算节点。这样就允许系统管理员来及时发现失控的作业或者重新分布提交作业，将存储需求分割给更多的计算节点。

### SNMP

警报信息可以通过SNMPv3发送给系统管理员。SNMP 被存储集群系统用作发出警报 (traps) 和允许查询系统状态 (gets)，但是，出于系统安全方面的考虑，不允许改变配置(sets)。这对数据中心环境的单点性能和故障的监控非常有价值，

### 动态负载均衡

具有自动负载均衡，给数据存取提供最高的性能的能力是Panasas ActiveScale 操作环境最主要的功能之一。以下是部分主要功能介绍：

#### 被动对象负载均衡

只有对所有磁盘驱动器进行数据存取的并行存储架构才可能提供最高的性能。在Panasas 存储集群系统的ActiveScale 操作环境下，数据存放在所有的存储刀片上。一个叫做性能管理器 (Performance Manager) 的服务分布在所有的存储刀片上用来维护系统表，来维护系统的性能和容量。当计算节点请求写数据时，指挥刀片使用这个表决定使用最佳的一组存储刀片来写数据。因此，系统使用的越多，就越是变得均衡，从而提高了性能并且避免了刀片上的“热点”。

#### 主动对象负载均衡

当新的磁盘容量增减或者替换时，使用被动对象负载均衡或许不能够使系统取得完全均衡。这时，ActiveScale 操作环境将使用性能管理器建立起来的对象表，由性能管理器决定将那些对象在存储集群中移动到哪些存储刀片上以得到最佳的性能。然后，指挥刀片就在性能管理器的调度下来移动这些对象。

#### 数据读组合

在存储系统中有一个功能，存储刀片将收集读请求，以便自己能够执行最有效的操作顺序，为存储刀片缓存提供最优化的性能。当并行应用启动时，大量的计算节点会对同一数据进行请求，这是，系统会给数据存取提供最优化的性能。

## 数据重构负载均衡

ActiveScan 减少了像传统RAID的数据重构的可能性。当不可预知的磁盘故障时，数据是安全的，因为这时数据由RAID保护，但此时，丢失的对象需要重构。数据重构负载均衡能够使用在存储集群系统中的所有控制刀片参与，这样，使得数据重构的时间远远少于传统的RAID数据重构时间。在同等情况下，传统的RAID控制器对500GB磁盘进行数据重构，可能需要花费一整天的时间，但是，ActiveStor 存储集群，数据重构可以在90分钟内轻松完成。从而大大降低了数据在没有保护状态下的时间，增强了系统性能。

## 冗余

Panasas ActiveStor存储集群系统在诸多层次上提供冗余，最大可能性来保护用户数据。

### RAID

Panasas 面向对象的ActiveRAID™ 根据您在虚拟卷中存取文件的大小，内部会自动选择RAID 1或RAID 5。系统用户和系统维护人员不再需要关心RAID这一层，这与传统RAID的系统管理员需要给不同的RAID建立单独的卷相比，增加了效率并降低了维护成本，能够腾出更加充裕的时间来管理和维护用户的应用。使用ActiveRAID，小文件被镜像以提供更高的性能；大文件使用RAID5。数据被分布在多个刀片上以使得到最高的性能。PanFS 并行文件系统具有独一无二的智能数据重构能力，替换磁盘空间横跨在所有的存储刀片上，并且所有这些刀片并行工作加速任何需要的数据重构。这与传统的RAID相比，戏剧性的减少了数据重构时间。

### 风扇

在ActiveStor 存储集群的一个机箱中，有三个可变速风扇。在风扇失效时，工作的风扇就会自动加速，给机箱提供最佳的温度环境。同时系统会给系统管理员发出提醒，以便及时替换。无论是一个刀片，电源或者电池从机箱中移走，机箱会自动关闭空位，保证空气气流的正确导向。

### 电源单元

双冗余，可热插拔的电源，给存储系统机箱提供连续的供电能力。如果一个电源发现问题，另外一个就会接管，并且保证机箱的正常供电。这时，带着机箱ID的警报会自动发送给系统管理员，以方便尽快替换。

### 电磁备份

电池备份给整个机箱存储的数据提供安全保障，即使意外断电，也不会对数据造成威胁  
这个内置的UPS 允许存储集群系统优先使用缓存中的数据，以提高性能。

这样，保证了数据总是被写入可用的刀片中。机箱中的电池提供充足的电源，保证在缓存中的所有数据写回磁盘。之后，系统才会执行正常的关机。ActiveStor 系统对电池进行周期检查以保证充足的充电，当需要替换时，系统会及时通知系统维护人员。

#### ActiveScale 镜像

ActiveScale 操作环境无论是在指挥刀片还是在存储刀片上，都是以镜像的形式存在，因此，即使一个系统出现问题，另外的镜像会即时提供文件访问，而不需要进行数据重构。这样 ActiveScale 操作环境给每一个刀片提供了高可用性。使得 ActiveStor 系统具备非常高的可靠性。

#### ActiveUpgrade

所有的存储系统都需要升级自己的系统以保证增加新的功能和修复问题。Panasas ActiveUpgrade 可以在升级的同时不中断应用和用户的访问。同时 Panasas 的客户可以通过他们自己的 MyPanasas 入口来进行。升级时，可以在后台以并行的方式同时对系统中的所有刀片进行升级，根本不需要中断客户应用。

#### 摘要

Panasas ActiveStor 存储集群系统给客户提供了一个软件，硬件一体化的集成存储解决方案。提供了业界领先的带宽和随机 I/O 性能。系统提供的 PanFS 并行文件系统是一个基于可扩展的以对象为基础架构的文件系统，它具有超级的性能，可扩展和管理能力。

Panasas 预自我管理包含了一系列关键技术，他们协同工作，能够监控设备自身环境，并且自愈，这样使得 Panasas 存储集群连续工作和提供最高的性能给运行集群应用的用户。避免了传统的存储架构下的系统管理的工作，将系统管理员从预自我管理的工作中解放出来，这样降低了系统管理的花销，一个系统管理员可以对数 terabytes 级别的系统进行管理。

Panasas 为所有的 HPC 环境用户提供完整的存储解决方案，系统具有最高的带宽和 I/O 性能，可扩展的性能和容量；以及简单的统一管理。Panasas 软件硬件一体化的解决方案已经为全球数以千计的 Linux 集群客户提供了他们的对象集群存储系统。如果希望了解更多的信息，请和我们联系 [www.panasas.com](http://www.panasas.com) 或 1-888-PANASAS.



---

#### Accelerating Time to Results™ with Clustered Storage

6520 Kaiser Drive Fremont, California 94555 Phone: 1-888-PANASAS Fax: 510-608-4798 [www.panasas.com](http://www.panasas.com)

© 2006 Panasas Incorporated. Panasas Incorporated. All rights reserved. Panasas, the Panasas logo, Accelerating Time to Results, [My.Panasas.com](http://My.Panasas.com), PanActive Link, PanActive Support Program, Panasas ActiveScale File System, Panasas ActiveScale Storage Cluster, Panasas ActiveRAID, Panasas ActiveStor, Panasas DirectorBlade, Panasas DirectFLOW, and Panasas StorageBlade are trademarks of Panasas in the United States and other countries. All other trademark names are property of their respective companies