

灵活、随用的数据存储时代

客座编辑导言 • Sundara Nagarajan • 2013 年 3 月



数据中心里唯一一个在规模和数量上都保持持续增长的组件是存储，它也是计算学科中颇具吸引力的一个研究领域。企业 IT 负责人在组织和经营自己的数据中心时，一直都在追求最高效率。由于 IT 已经成为现代企业的业务性技术，数据不可用——即使持续时间很短——都变得不可忍受。同时，IT 员工每天都要面对日益复杂的业务问题，其背后往往是大量的数据。

在日复一日的工作中，“热”数据并不多，比如那些需要马上访问的数据（例如刚刚收到的电子邮件），而大量数据是“冷”的（并不需要立即访问，比如两年前的邮件）。也就是说，冷数据的一小部分会因为需要而突然变热，在数据和用户规模都很大的情况下，这是一个颇为棘手的难题。

这给存储研究和工程人员提出了不少挑战：在不中断访问的情况下（甚至是在升级和维护时）如何确存储基础设施可扩展、高效、可靠？如何在尽可能快地访问热数据的同时用尽可能低的成本存储冷数据？如何确保所管理的数据在其整个生命周期中都保持连续统一？存储行业的领导者和初创企业需要继续在人才和资源方面进行投入，才可能解决这些现实问题。

数据管理不断变换的风景线

目前，用户可以直接投资建立自己的存储基础设施，也可以购买按需付费的存储服务。这两种情况背后的管理问题都是多租户能力问题，即能够把用户组完全隔离开来，并为这些租户提供相应的安全访问和服务确保级别。CIO 和 IT 经理可以借助先进的经济模型和经纪服务来理解投资和运行成本的动态本质，以建立最合意的总拥有成本（TCO）。

存储组件技术的进展，包括闪存/固态驱动器、存储级内存以及不断提升的磁盘驱动器，为存储系统供应商提供了新的选择。用户需要的是隐藏了元件选择、数据部署和数据保护等复杂细节的解决方案。IT 研究公司 IDC 预测，在企业级存储方面，可扩展架构在 2013 年增长最快[出处]。系统工程师在寻求可以长期使用并扩大规模的集群架构方面的努力开始于 70 年代，这一架构曾成功应用于高端大型机和昂贵

的容错系统。基于普通组件和集群互联结构的存储系统架构创新正在创造物美价廉的产品，这些系统提供了高度灵活的、可扩展的虚拟化存储基础架构，可以实现不中断运行。

现代存储系统吸收了几十年来的大量算法和体系结构创新。构建产品所必需的成套基础技术，虽然还不足以用来区分存储系统，但是可以用来表征它们。这些基础技术包括：磁盘布局层次的强数据保护算法，可恢复文件系统，数据去重和压缩等高效存储技术，文件/文件系统/卷快照以及克隆技术。集群架构大幅度提高了最小化中断的能力，无论造成中断的原因是故障还是正常升级或日常维护。已有证据表明，集群已成为主流存储体系结构，它与普通组件的良好组合可以实现最佳客户价值。

本月主题

我为“今日计算”本期主题收集的这组文章可以说是对现代企业级数据存储架构和集群存储解决方案的一次检阅。

在《[IaaS 云架构：从虚拟数据中心到联邦式基础设施](#)》一文中，Rafael Moreno-Vozmediano, Rubén Montero 和 Ignacio Llorente 对云基础设施进行了很好的剖析，还解释了云 OS 的概念以及它在现代企业数据中心运营中的角色。

Lee Garber 的《[聚合基础设施：直面效率挑战](#)》是一则《Computer》新闻故事，中心是数据中心组件（计算服务器、存储和网络）的集成问题。这种架构集成的组件可以来自单一供应商，也可以来自不同组件的最佳供应商。本文介绍了实现灵活的、集成货架产品的基础设施系统（即数据中心）面临的挑战和机遇。

Boris Grot 及其同事在《[用可扩展的处理器优化数据中心的总体拥有成本](#)》一文中全面分析了基于低端处理器的可扩展方案如何达成最优 TCO（总体拥有成本）。这篇文章把 TCO 定义为“一种综合考虑房屋、供电和制冷设施以及硬件采购成本和运营费用的优化指标”。这一优秀的研究将对存储系统架构产生深远影响。

Yaoguang Wang 和他的同事在随后的《[HAaaS：走向高可用分布式系统](#)》中讨论了对互补存储架构的重要需求：高可用性。本文介绍了用共享存储解决状态管理有关问题的方法。

最后一篇是《[一种面向云服务的新型分布式文件存储方案](#)》，这是 Yu Zhang、Weidong Liu 和 Jiaying Song 的一篇会议论文，讨论了云服务的属性，提供了有关共享存储服务的重要信息。

为了增强本期主题，我联系了 Tim Russell，他是 NetApp 数据生命周期生态系统方案的副总裁，询问他对面向企业用户的集群存储架构的发展和挑战的看法。他

领导的团队负责开发支持数据从创建到长期使用的方案，之前他领导过一个负责识别和评估长期市场趋势的产品策略组，可以说他一直在和存储系统的大用户打交道。在这份独家视频回应中，他分享了他对企业存储系统的看法。

【Tim Russell 采访视频】

我还请教了计算机科学与软件工程教授 Rajkumar Buyya，他是澳大利亚研究理事会的未来院士，澳大利亚墨尔本大学云计算和分布式系统实验室主任。他在应答视频中，谈了他在集群系统领域的工作和研究视角。

【Rajkumar Buyya 采访视频】

企业级解决方案

企业 IT 用户，包括员工、客户、供应商和合作伙伴，在他们快节奏的商业和个人生活中，需要无处不在的 IT 体验。对于 CIO 和企业 IT 技术人员，业务服务中断不在选项之列。快速变化的技术和交付模式，促使他们重新思考向企业用户交付业务技术的方法。数据管理体系结构是解决方案的关键，集群存储系统正在成为企业现代业务技术的基础。



Sundara Nagarajan 是 NetApp 技术总监和印度班加罗尔国际信息职业技术学院访问教授。他是 IEEE 今日计算地区代表，负责 IEEE 计算学会在印度的联络。可通过 [s.nagarajan at computer dot org](mailto:s.nagarajan@computer.org) 联系他。

([黄铁军](#) 译)