软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2017,28(8):1927−1928 [doi: 10.13328/j.cnki.jos.005205]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

分布式云存储:理论、技术、系统专题前言*

黄 宇[1], 吴维刚[2], 赵军平[3]

[1](计算机软件新技术国家重点实验室(南京大学),江苏 南京 210023)

[2](中山大学 数据科学与计算机学院,广东 广州 510006)

[3](EMC 中国研究院,北京 100084)

通讯作者: 黄宇, E-mail: yuhuang@nju.edu.cn

中文引用格式: 黄宇,吴维刚,赵军平.分布式云存储:理论、技术、系统专题前言.软件学报, 2017,28(8):1927−1928. http://www.jos.org.cn/1000-9825/5205.htm

在 Internet 开放环境下,以云服务和移动终端为计算平台,以大数据为内容资源的新一代应用,对云平台的开发者和云服务提供者提出了诸多挑战.随着大规模分布式应用关注的焦点逐渐从“计算”向“数据”迁移,分布式云存储技术作为云计算的关键性支撑技术得到越来越多的关注.一方面,持续增加的应用规模对云存储数据访问的低延迟、高可用、高容错、可扩展等特性提出了更高的要求;另一方面,计算平台的水平扩展和存储介质的更新换代给云储存系统性能的进一步提升带来了机遇和挑战.源于上述挑战和机遇,近年来,分布式云存储相关理论、技术和系统方面的研究得到了学术界和工业界的广泛关注,研究活动较为活跃.为了总结我国在分布式云储存方面的最新成果,促进该领域的学术交流,推动进一步研究,我们组织了本专题.

本专题经过两轮征稿,共收到 18 篇有效投稿,审稿过程历经 5 个多月,有约 20 名相关领域的专家和学者参与了审稿工作.审稿过程共分两个阶段.在第 1 阶段,每篇投稿由两位领域专家和一位特约编辑分别审阅,根据评审结果从 18 篇投稿中遴选出 6 篇论文作为条件录取.在第 2 阶段,通过第 1 阶段审稿的论文作者获邀在全国软件与应用学术会议(NASAC 2016,昆明)云存储专题 Session 中报告交流.特约编辑现场汇总听众意见并复审,进而提出论文的进一步修改意见.复审及修改结果经由本刊编委会终审认定.经过以上严格程序,最终确定收录 6 篇论文,现概述如下.

在分布数据管理技术方面,《一种基于跳跃 hash 的对象分布算法》基于跳跃哈希的方法将存储节点映射到二维矩阵,进而数据对象的分布、定位只需从矩阵的行内/行间计算目标节点的行号和列号,有效提升了数据分布的公平性、紧凑性和自适应性,并降低了数据查找的开销和数据对象迁移的开销.《基于简单再生码的带宽感知的分布式存储节点修复优化》基于网络编码实现大规模分布数据的容错存储,结合实际物理网络拓扑结构,将链路带宽引入到简单再生码的修复过程中,建立了带宽感知的节点修复时延模型,提出了基于最优瓶颈路径和最优修复树的并行修复树构建算法,实现了低延迟、高修复率的容错存储.《一种自适应文件系统元数据服务负载均衡策略》针对分布式文件系统元数据管理问题,提出一种自适应元数据服务负载均衡策略.该策略基于实时的元数据服务器的性能评价模型,自适应地监控和更新元数据服务器的负载情况,并根据服务器负载指标,实时地完成元数据的迁移,实现负载均衡.

在面向分布式云存储的底层支撑技术方面,《一种基于小数据同步写的回写 I/O 调度器》着重关注普遍存在的小数据同步带来的 I/O 性能瓶颈的问题,提出一种新的 I/O 调度器.该调度器可以识别小数据写,并通过对其他数据块中的数据进行压缩,将小数据嵌入到压缩出来的空间中,从而将小数据和该数据块一起写入到磁盘上,以异步回写的方式完成小数据的同步写,不仅有效缓解了磁盘的写放大问题,也极大地提高了小数据同步写的效率.《虚拟化环境下面向多目标优化的自适应 SSD 缓存系统》关注新型存储介质 SSD 在虚拟化环境下的应

* 收稿时间: 2017-01-11; jos 在线出版时间: 2017-1-12
CNKI 网络优先出版: 2017-01-12 10:36:02, http://www.cnki.net/kcms/detail/11.2560.TP.20170112.1036.007.html

用,基于对虚拟机和应用状态的监控,动态检测局部 SSD 缓存抢占状态;基于聚类方法生成虚拟机的优化放置方案,依据全局 SSD 缓存供给能力确定虚拟机迁移顺序和时机,有效缓解 SSD 缓存资源的争用,同时满足应用对虚拟机放置的需求,提升应用的性能并兼顾应用的可靠性.

在基于分布式云储存的领域应用方面,《面向海量高清视频数据的高性能分布式存储系统》通过对视频监控数据的特点和传统存储方案进行分析,提出一种高性能分布式存储系统解决方案.不同于传统的基于文件存储的方式,该方案设计了一种逻辑卷结构,将非结构化的视频流数据以此结构进行组织并直接写入 RAW 磁盘设备,解决了传统存储方案中随机磁盘读写和磁盘碎片导致存储性能下降的问题.该方案将元数据组织为两级索引结构,分别由状态管理器和存储服务器管理,极大地减少了状态管理器需要管理元数据的数量,消除了性能瓶颈,并提供精确到秒级的检索精度.此外,该方案灵活的存储服务器分组策略和组内互备关系使得存储系统具备容错能力和线性扩展能力.

本专题主要面向分布式系统、云计算和大数据等相关领域的研究人员,反映了我国学者在上述领域的最新进展.在此,我们感谢所有向本专题踊跃投稿的各位作者,感谢发起本专题的中国计算机学会系统软件专委会和软件工程专委会,感谢对投稿论文进行认真审阅并提供宝贵意见的各位审稿专家,感谢细致和辛勤工作的《软件学报》编辑部的各位编辑.

黄宇(1982－),男,江苏淮安人,博士,副教授,博士生导师,CCF 专业会员,现任中国计算机学会系统软件专委会委员,主要研究领域为分布式算法,分布式系统,网构软件.

吴维刚(1976－),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为网络与分布式计算,云计算资源管理,大数据管理,车联网.

赵军平(1982－),男,硕士,EMC 中国研究院首席系统架构师,CCF 专业会员,ACM 会员,主要研究领域为分布式系统,NVRAM,微服务架构和实时大数据分析系统,已申请美国/中国技术发明专利 50 余项(其中已授权美国专利 4 项).