

综上所述, BWMMMS 采用的基于集中共享虚拟存储的文件系统元数据存储模型, 基于文件动态访问的元数据请求分布管理机制, 基于元数据动态迁移的文件系统一致性协议, 基于请求优先级、投机执行请求的请求并发控制机制, 能够很好地支持系统的扩展需求。同时, 系统在实现上还存在不足, 需要进行大量的优化工作。

第八章 结束语

存储技术、计算技术和网络技术的进步共同催生网络化存储技术。集中化网络存储有利于提高系统的存储资源利用率、存储资源管理的扩展能力和系统的可管理性，为用户提供方便的存储资源共享。

在网络存储系统基础上，分布式文件为用户提供便捷的数据共享。文件系统由元数据和数据构成，元数据访问效率是文件数据访问性能的关键因素。应用规模的不断增大和新应用的不断涌现，对分布式文件系统的元数据服务提出越来越高的要求。在新的存储系统架构下，针对分布式文件系统元数据服务的扩展能力进行深入研究，对推动网络存储技术的发展具有至关重要的意义。

8.1 本文工作总结

本文着重于分布式文件系统元数据服务的可扩展性研究，期望以多个服务器组成的服务器集群方式提供具有较强扩展能力的分布式文件系统元数据服务，以促进网络存储系统更好地支持多种应用环境，提高其适用性。

通过分析已有的研究成果，结合用户的元数据访问负载的特点，本研究获得了高扩展能力的存储资源组织、管理和使用，动态高效的元数据请求分布管理，简单有效的文件系统一致性协议，以及高效的元数据请求并发控制算法等研究成果。具体包括：

1. 基于集中虚拟存储模型的存储资源组织、分布式层次化的存储资源管理、完全共享的存储资源使用模式的文件系统元数据存储服务。虚拟存储支持存储资源的独立扩展，层次化的管理机制支持存储资源的有效利用，完全共享的存储资源使用模式解决非常困难的数据放置问题，支持动态灵活的元数据请求服务。在此基础上，元数据的存储和访问彻底分离，元数据的存储不再限制其请求的分布，元数据的存储问题得到有效解决。
2. 动态灵活的元数据请求分布管理机制和策略，支持元数据服务器规模和用户元数据访问负载的动态变化。通过仅管理活跃元数据的请求分布，大幅度缩小请求分布管理的对象集，支持请求分布管理的扩展。根据元数据活跃性管理元数据请求分布，为用户元数据访问负载的动态变化提供有效支持。采用层次化的分布管理机制，支持元数据服务器规模的动态扩展。元数据请求分布策略结合请求的动态性和元数据间客观存在的访问相关性，结合元数据请求的处理效率，尽可能地平衡服务器的负载。
3. 简单高效的更改文件系统的跨服务器请求的原子性保证协议。对称的元数据服

务器系统结构、用户元数据访问的统计特征、动态灵活的元数据请求分布管理，为简单高效的元数据请求一致性协议提供足够的支持。通过动态改变活跃元数据在服务器间的分布，完成元数据在服务器间的迁移，将跨服务器请求集中化，利用相关技术保证文件系统一致性，减小协议对系统正常请求处理的影响。

4. 基于系统的请求并发情形分析，提出简单的请求同步控制算法。利用元数据请求分布信息，结合元数据请求语义，充分分析可能的请求并发，根据请求的优先级进行判定，尽可能地并发元数据请求。
5. 根据上述结果，设计和实现了 1 个应用到 BWFS 的原型系统（蓝鲸多元数据服务器系统，BWMMS），并通过实验验证其扩展能力的有效性。

8.2 未来工作展望

通过本研究，分布式文件系统可扩展元数据服务研究的系统平台得以实现，为未来的研究提供必需的支持。初步的评估结果表明，系统采用的集中共享虚拟存储架构和基于数据共享的对称服务器结构，灵活的元数据请求分布管理，轻量级的文件系统一致性协议和有效的请求并发控制机制等，能够为系统可扩展的分布式文件系统元数据服务提供必要的支持。但是，由于本研究主要关注分布式文件系统元数据服务的扩展能力。在后续的研究中，还需要在研究内容的深化、系统适用范围的推广和新研究领域的拓展等方面进行大量的工作。

8.2.1 研究内容的深化

本研究着重于分布式文件系统元数据服务的扩展性的研究，主要集中在元数据存储服务、集群的服务器结构、元数据请求分布管理和元数据请求并发控制等关键问题。在很大程度上，系统的元数据请求处理性能和扩展能力依赖于元数据请求分布决策算法，如何提供有效的元数据请求分布决策是首先需要进一步深化的问题。除此之外，还需要在系统的实现上进行优化。

1. 元数据负载预测模型的扩展和评估。动态的元数据请求分布管理的核心是元数据服务器负载的预测。在保证元数据请求的处理效率的前提下，尽可能地将用户的负载分布到多个服务器。元数据服务器负载的自学习模型、元数据负载参数的丰富、元数据分布策略的优化，都是需要进一步深入的研究内容。
2. 集中决策服务器结构的改进。从系统的实验评估结果看，单一的集中决策机制和元数据分布信息缓存，能够很好支持元数据活跃性变化不频繁的应用的扩展需求。在未来工作中，为了支持更大规模的系统，还需要进一步考虑集中决策服务器结构的扩展。通过服务器集群或层叠方式组织多个决策服务器，扩展集中决策服务器结构，进一步提高系统的扩展能力。

3. 针对应用特性的元数据请求分布管理策略。BWMMS 根据的应用特征统计结果中，文件创建非常稀少，其处理过程时延较大。但是，可能存在如频繁的文件创建/删除、临时文件等应用，其元数据活跃性变化频繁，元数据分布信息缓存的作用将非常有限，导致集中决策服务器的压力非常大。为很好地支持更多类型的应用，元数据分布策略将扩展其考虑的因素范围。不仅需要考虑元数据服务器的负载、单个请求涉及的元数据之间的相关性，还需要抽取更多应用的访问特征，从单个用户角度考虑元数据请求的分布。
4. 大规模系统的验证及优化。本研究在相对较小系统规模下，对系统的扩展能力进行了初步的验证。在未来工作中，还需要从理论验证和实际环境使用两个方面，验证在更大规模系统中的扩展能力，并根据结果对系统进行优化。
5. 元数据读写的优化。由于元数据表现出频繁的小读小写特征，为降低请求服务器与存储服务器间的通信开销，需要结合应用的特征和系统的管理策略，优化请求服务器的元数据读写性能。

8.2.2 系统适应范围的扩展

本研究依赖于已有研究结果的结论：文件系统的活跃数据比例不大，活跃文件的数目比例在 5%-20%左右。需要更改文件系统的分布式元数据请求所占比例非常小，在请求数量比例上不超过 5%。基于这些结论，BWMMS 以动态灵活的机制决定元数据请求的分布，通过元数据动态迁移、集中化处理跨服务器请求等方式保证文件系统的一致性。

但是，已有研究针对的应用范围毕竟有限，如何拓展 BWMMS 的适用范围将成为下一步研究的重点。目前学术界已有的研究结果主要针对软件开发、web 服务、E-Mail 和其他商业应用环境，提取用户的数据访问特征，且研究成果的发布时间已经比较早。对于新兴的应用，如遥感信息处理、生物计算等，还没有显著的成果发表。为更好地支持网络存储系统在多种环境的应用，还需要通过与相关研究的合作，提取应用的 I/O 特征，改进和完善文件系统的相关问题。通过对更多应用特征的综合，推广系统的适用范围，推动网络存储技术的进一步发展。

8.2.3 研究领域的扩展

本研究主要关注的是分布式文件系统元数据服务的性能和扩展能力，没有过多的关注元数据服务的高可用问题。

分布式系统的高可用性是系统广泛使用的必备条件。必须得到充分的研究，系统才能很好地满足更多应用的需求。分布式系统的高可用性要求通过系统服务器间以某种方式进行互补，以性能的降低或者部分功能的不能使用为代价，提供服务的可用性。系统的故障只会带来性能的降低、或者系统功能的部分缺陷，整个系统依然能够继续为系统用户提供服务。

本研究仅仅通过各个服务器维护的元数据分布信息实现相互冗余，部分地支持元数据分布管理服务的高可用功能，保证同一个元数据不会同时具有多个宿主服务器。目前仅仅能够做到：1) 通过 MS 缓存的分布信息合并，能够恢复出现故障的 BS 的元数据缓存信息；或者：2) 各个 MS 出现故障但 BS 正常工作时，各个 MS 恢复后，可以通过请求驱动的方式重新缓存元数据分布信息。系统还不能支持 MS 和 BS 任意故障的情况，BWMMS 的高可用问题还需要深入的研究。

总之，本研究通过可扩展分布式文件系统元数据服务关键问题的解决，完成分布式文件系统可扩展元数据服务的平台构建工作，还需要通过未来的工作提升系统的可用性，并加强系统在高可用性等领域的研究，推动网络存储系统在更大应用范围的适用性。