

国家重点基础研究发展规划（973）项目
下一代互联网信息存储的组织模式和核心技术研究

项目编号：2004CB318200

简 报

[2007] 01 号 总第 7 期

项目办公室 编

2007 年 1 月 27 日

项目顺利通过中期总结与评估

2006 年 12 月 4-5 日国家重点基础研究发展计划项目（973 计划项目）“下一代互联网信息存储的组织模式和核心技术研究”（编号为 2004CB318200）的中期总结与评估会议在华中科技大学光电国家实验室举行。

会议由项目首席科学家、华中科技大学冯丹教授主持，项目组咨询专家张以谟教授、金怡濂院士和陈俊亮院士，湖北省科技厅王红斌处长，项目专家组专家沈绪榜院士、叶朝辉院士、何炎祥教授、韩承德教授、谢长生教授和代亚非教授，以及中国地质大学（武汉）康立山教授，中船重工 709 研究所袁由光研究员，武汉理工大学钟珞教授组成评估小组，项目组下属六个课题组主要学术骨干和项目依托单位华中科技大学科研主管人员等参加了会议。

会议上，项目各课题负责人汇报了两年来研究进展情况和取得的创新性成果，项目评估组专家在听取汇报总结的基础上进行了评估。

第一课题组

在理论、方法和系统等多个层面开展工作，取得以下创新性成果：

1、研究网络存储数据行为规律，给出一种分析 I/O 负载自相似性的判定方法，发现网络存储 I/O 行为具有突发性和自相似性，且读自相似性占主导地位；采用多分形模型反映 I/O 流量长相关性和短时间范围内突发特性，研制了一种分布式 I/O trace 系统。

2、提出可扩展主动存储对象概念，引入方法（Method）对象及策略（Policy）对象；采用对象分级方法研究了对象生命周期现象；通过柔性空闲空间管理和渐近空间分配策略实现对象文件系统性能的持久性；针对流媒体和事务处理两类应用，分别以阻塞概率和响应时间作为代价函数研究对象的放置策略，使性能最佳。

3、提出存储协议 vSCSI 并实现了 VISA 存储系统，根据存储芯片级软硬件结合缩短 I/O 路径方法的需要研制了系列具有自主知识产权的 IP core，完成了基于 I/O 处理器和交换式内部互联结构的对象存储控制器硬件实现。

4、构建了基于可扩展对象的海量存储系统原型，提出三种元数据管理与对象布局的优化方法：一种根据文件大小决定采用哈希映射或分片映射策略进行文件到对象映射的数据柔性分布算法，一种基于热点数据规律预测数据副本迁移的均衡策略，一种采用非对称共享元数据结构的主从备三重链式元数据服务。改进并实现了对象存储系统三方安全协议；提出一种基于热度的多线程重建调度优化算法提高存储系统可靠性，相关论文被 FAST07 录用；初步实现移动 Cache 体系结构和运行机制的研究。

5、形成对象存储原型系统，作为主要部分与课题三和课题六合作参加 Super Computing 2006 的存储挑战项目，成为四个入选项目之一。

6、将自组织概念和对象存储系统结合，初步形成支持主动存储服务的自组织对象存储系统理论框架。

该课题已发表论文 67 篇，SCI 收录 19 篇，EI 收录 15 篇，申请专利 15 项，批准专利 2 项，获得一项软件著作权，获得一项省技术发明一等奖。

课题经过两年多的科研工作，和课题三一起建立了一个用于主动对象存储实验的平台，已可以支持 GIS 示范应用。该平台容量达到 100TB，可以扩展到 1PB 的容量，核心传输速度达到 10Gbps。研发的对象存储原型系统聚合带宽已达到 11.1GB/s。目前该平台包含 100 个各种类型的存储节点，已接入 CERNET 和 CERNET2，同时支持 IPv4 和 IPv6。由于构成系统的设备具有多样性，其软件多样性和复杂性将为项目进一步深入研究提供试验平台。

第二课题组

按照项目预定计划，全面完成了课题预定任务，实现了预期目标。

研究了面向下一代互联网的网络存储系统中多层次数据通道的构建方法与优化机制，提出了“VINE”通道体系结构，系统地描述了功能完备的集成网络存储通道架构、透明可靠的网络存储通道接口、轻量级网络存储通道协议以及弹性灵活的网络存储通道服务机制等研究内容，并研制完成包括多链路可配置网关系统在内的多个软硬件系统。

与其它课题组合作，通过通道协议优化提高了基于对象存储系统的服务性能，利用网关阵列结构扩展了对存储系统的访问聚合带宽，同时，通过虚拟化、多层次的高速通道机制，实现与开放、动态存储网络的连接，并对提高数据通道的可靠性与可用性提出了新的途径和方法。

研究工作在如下几个方面有创新性：

- 1、提出了一种集成网络存储通道架构，设计并实现了可适用于多网络通道环境下的多链路可配置网关系统，扩展了对基于对象存储系统的访问聚合带宽。

- 2、提出了具有透明的网络存储通道接口，能够支持异构、分布、异质的数据存储资源。

- 3、提出并实现了一套轻量级网络存储通道协议，深入研究了其在基于对象存储、系统负载平衡等技术中的应用。

- 4、提出了具有弹性灵活特征的网络存储通道服务机制，研究了在服务计算环境中应用此机制的重要意义。

第三课题组

从系统层面研究了海量数据网络存储系统的时空演化规律和进化结构,建立了包括统一存储网和基于 IPv6 的对象集群存储的大规模实验平台,并在该平台上进行了研究工作。课题组在两年的时间内,做了大量而有效的研究工作,完成了预定的研究目标,并取得了部分超过原计划任务书的研究成果,为后三年的工作打下了良好基础。

研究工作取得了以下主要进展和创新性的成果:

1、设计并实现了基于 IPv6 的大规模对象存储集群系统,不仅成为 100TB 对象网络存储研究平台的重要组成部分,还为用于下一代互联网的集群存储模式提供了核心技术。

2、提出了统一存储网机制,将多种异构网络存储结构进行融合,并根据数据属性进行分布优化。

3、提出存储系统的进化原理、机制和实现方法,使海量存储系统的性能随时间和系统规模的扩展而增强。

4、从理论上探讨海量存储系统自发演化规律,为优化海量存储系统的时空演化过程提供了依据。

5、与课题 1、6 合作解决了 GIS 系统的存储挑战性问题,参加了国际 Super Computing 存储挑战方案竞赛,进入最后 4 名决赛,获得 Finalist 奖。

第四课题组

在前两年的工作中在理论、方法和系统方面开展了大量的工作,完成了任务书中规定的科研任务,已经取得了部分研究成果。

课题所做的研究工作及创新点包括:

1、提出开放的存储体系结构的思想,对 P2P 模式下的存储系统的基本属性和功能进行抽象,定义了 P2P 存储功能层次,给出了各个层次之间的接口和服务模式。

2、研究 P2P 系统的持久性问题,发现了国际上通行测量可持续性方法的不足,提出了新的测量和计算方法。

3、研究 P2P 系统的信誉机制，在用户行为分析和检测方面提出独创性的新方法。在纠删码研究上，取得了比以往研究更加深入的成果。

4、完成了一个 P2P 增量备份存储系统，将研究结果在系统中进行验证。

5、直接面向互联网，实现在线备份和数字财富仓库两个应用，同时支持 IPv6 和 IPv4 的功能，支持跨平台，符合异构性的要求。

课题前两年研究中共申请专利 2 项，发表文章 18 篇，其中国际会议 12 篇，国内重要期刊 5 篇，存储年会 1 篇，其中 SCI 索引 3 篇，EI 索引 7 篇。

第五课题组

完成了前两年预定目标，所做的研究工作及创新性成果包括：

1、以存储与数据分离的理念为基础，建立了存储对象模型。通过引入 Storage Fault 概念，使存储对象能动态扩展和数据按需部署。将存储对象模型扩展到广域存储领域，研究了广域存储范围内的数据缓存和部署机制。

2、以计算与存储分离的理念为基础，在数据对象模型的基础上，初步建立了服务对象模型，完成了一个通用性支持服务动态绑定嵌入式装置研究。

3、研究并提出了快速系统数据克隆、配置和管理方法——迭代式写拷贝，支持大量网络服务的动态生成。研究并提出了缓存资源和网络资源的共享机制，较大地提高了资源利用率和系统效率。

4、提出了多维度的存储资源预定与分配算法、SCSI 命令队列的动态调度方法，可适应多种应用模式对存储服务的需求。

5、提出海量存储环境中虚拟空间的组织与动态调整方法，可适应存储系统动态变化，测试表明，Cellio92 和 TPC-C Trace 下负载调整使 I/O 操作平均响应时间减少 20%以上。

6、提出大规模异构存储环境的数据块副本技术，提高了存储系统中的数据可用性。测试表明，多副本选择调度使得读操作的带宽能提高 17%，在磁盘失效的情况下仍能使读操作吞吐率提高 10%。

7、提出了一种循环条带卷快速扩容的方法，它比传统的基于映射表的扩容方法快 40%，该方法基于滑动窗口的映射管理，支持数据重分布又不失扩展性，并用元数据懒惰更新技术大幅削减元数据写操作次数。

目前，课题五发表论文 31 篇，其中 SCI 检索 10 篇，EI 检索 16 篇，其中有多篇论文发表(录用)在 IEEE Trans. on Computers(影响因子 2.419)、ACM Trans. on Storage 和 IEEE Conference on Mass Storage Systems and Technologies 等国际著名期刊和会议上，并申请 11 项发明专利。该课题的相关的研究成果已经应用到石油、审计、公安等单位。

第六课题组

实现了一种可支持国家对地观测需求的地球空间信息系统原型 GlobeSIGht。该课题已发表学术论文 64 篇，专著 1 部。其中，SCI 已收录 1 篇，EI 已收录 14 篇，国际会议报告论文 28 篇，特邀报告 1 篇。该课题组本阶段的研究工作实现了预期目标，完成示范系统的第一阶段建设，部分研究工作进展超过了预期目标。

本阶段研究工作的创新性主要有：

1、提出了一种“基于对象存储的多媒体网络 GIS 体系结构”，实现海量空间数据的自主分布式存储和透明服务，为高性能大型 GIS 系统的研究开发提供了新的参考模型。

2、提出并实现了一种多媒体电子地图插件和扩展引擎，支持用户自定义多媒体插件，将多媒体技术集成到数字虚拟地球的应用中，非常方便地扩展了数字地球的应用范围，实现了真正意义上的大型分布式多媒体网络 GIS 系统框架。

3、提出了一种“多层次可扩展的地球空间信息海量数据组织模型”。从逻辑上规划全球的空间信息，建立空间数据组织模型，采用“垂直分层、水平剖分”的总体思想，建立多层次可扩展的全球统一的立体空间信息虚拟视图。

4、提出了一种“虚拟地球空间信息的无缝融合与可视化调度方法”，实现了一种中高分辨率全球空间数据的观测和无缝漫游机制，自动融合多

源多分辨率空间数据，支持多达 20 级海量影像数据的分布式获取与无缝融合，最高分辨率可达 0.1 米。

5、提出了一种具有空间语义的“基于对象存储的 GIS 空间数据存储模型”，为构建空间对象数据库的完全一体化组织管理提供了新的途径，发展了新一代海量空间信息存储的理论与方法。

目前，课题六已培养硕士 20 名，博士 5 名。整个课题组以 4 名教授、4 名副教授为核心组建，研究队伍梯队建设情况好，培养出一些优秀中青年研究人才成为课题组的骨干力量。

项目评估组听取了各课题两年来的研究进展报告，参观了各课题的成果演示，并对项目前两年的研究情况进行了充分的讨论和评估，认为：各课题的研究工作按照计划进行，均完成甚至超额完成了前两年的研究任务，并积累了一定的创新性成果；研究队伍状态良好，水平较高。专家组认为各课题下一阶段的研究计划切实可行，从目前的研究情况来看，各课题均有能力出色地完成后续阶段的研究任务。另外专家组也指出存在的问题：课题之间需要进一步加深相互结合，在科学理论的层面加强研究的力度，以期取得突破性成果。