

(12) 按照专利合作条约所公布的国际申请

7<sub>a</sub>

(19) 世界知识产权组织  
国际局



(43) 国际公布日  
2013年7月4日 (04.07.2013)

W I P O | P C T

(10) 国际公布号  
W O 2013/097119 A 1

- (51) 国际分类号 : G06F 12/08 (2006.01)
  - (21) 国际申请号 : PCT/CN201 1/084830
  - (22) 国际申请日 : 2011年12月28日 (2011.12.28)
  - (25) 申报语言 : 中文
  - (26) 公布语言 : 中文
  - (71) 申请人 (对除美国外的所有指定国): 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
  - ( ) 发明人 : 及
  - ( ) 发明人/申请人 (仅对美国): 陆海翔 (LU, Haixiang) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
  - (74) 代理人 : 北京凯特来知识产权代理有限公司 (BEIJING CATALY IP ATTORNEY AT LAW); 中国北京市西城区阜成门外大街甲9号国宾酒店B座505室郑立明 Beijing 100037 (CN)。
  - (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, ML, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。
  - (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。
- 本国际公布 :  
- 包括国际检索报告 (条约第21条(3))。

(54) Title: METHOD AND DEVICE FOR REALIZING MULTILEVEL STORAGE IN FILE SYSTEM

(54) 发明名称 : 在文件系统中实现多级存储的方法和装置

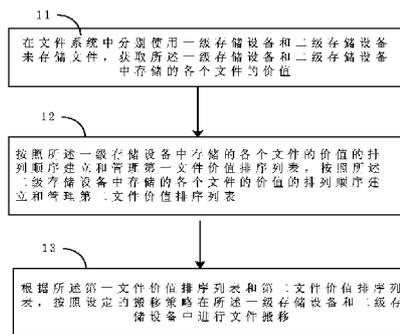


图 / Fig. 1

11 IN A FILE SYSTEM, A PRIMARY STORAGE DEVICE AND A SECONDARY STORAGE DEVICE ARE USED TO STORE FILES RESPECTIVELY TO ACQUIRE THE VALUE OF EACH FILE STORED IN THE PRIMARY STORAGE DEVICE AND THE SECONDARY STORAGE DEVICE

12 A FIRST FILE VALUE RANK LIST IS ESTABLISHED AND MANAGED ACCORDING TO THE ORDER OF THE VALUE OF EACH FILE STORED IN THE PRIMARY STORAGE DEVICE, AND A SECOND FILE VALUE RANK LIST IS ESTABLISHED AND MANAGED ACCORDING TO THE ORDER OF THE VALUE OF EACH FILE STORED IN THE SECONDARY STORAGE DEVICE

13 FILE TRANSFER IS PERFORMED IN THE PRIMARY STORAGE DEVICE AND THE SECONDARY STORAGE DEVICE ACCORDING TO A PRESET TRANSFER POLICY IN ACCORDANCE WITH THE FIRST FILE VALUE RANK LIST AND THE SECOND FILE VALUE RANK LIST

(57) Abstract: Provided is a method for realizing multilevel storage in a file system. In a file system, a primary storage device and a secondary storage device are used to store files respectively to acquire the value of each file stored in the primary storage device and the secondary storage device; a first file value rank list is established and managed according to the order of the value of each file stored in the primary storage device, and a second file value rank list is established and managed according to the order of the value of each file stored in the secondary storage device; and file transfer is performed in the primary storage device and the secondary storage device according to a preset transfer policy in accordance with the first file value rank list and the second file value rank list. The embodiments of the present invention can store files with high value in a storage medium with higher performance and store files with low value in a slow storage medium, thus making the best of the overall performance of the system.

(57) 摘要 :

[见续页]



A1

2013 09 11

V

---

本发明实施例提供了一种在文件系统中实现多级存储的方法。在文件系统中分别使用一级存储设备和二级存储设备来存储文件，获取所述一级存储设备和二级存储设备中存储的各个文件的价值；按照所述一级存储设备中存储的各个文件的价值的排列顺序建立和管理第一文件价值排序列表，按照所述二级存储设备中存储的各个文件的价值的排列顺序建立和管理第二文件价值排序列表；根据所述第一文件价值排序列表和第二文件价值排序列表，按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移。本发明实施例可以实现将价值高的文件存放到性能高的存储介质上，将价值低的文件存放在慢速存储介质上，使系统达到整体性能最佳。

## 说明书

### 发明名称：在文件系统中实现多级存储的方法和装置

[1] 技术领域

[2] 本发明涉及电子技术领域，尤其涉及一种在文件系统中实现多级存储的方法和装置。

[3] 发明背景

[4] 伴随着存储技术的发展，在磁带、机械磁盘等传统存储介质之外不断涌现出新的存储介质以及存储接口，这些新的存储介质和存储接口通常意味着更高的传输速率，其中，SSD (Solid State Disk, 固态硬盘) 就是一种典型技术。

[5] SSD也称作电子硬盘或者固态电子盘，它由控制单元和固态存储单元组成，其接口的功能、形式和使用方法与磁介质硬盘相同。目前SSD大部分应用在笔记本电脑上，为了和现有的笔记本相匹配，大多数产品外形尺寸与传统笔记本硬盘一致，也有其他规格尺寸的产品。由于SSD内部无机械结构，读取也不用马达旋转，所以抗震性和噪音方面都很有优势，另外SSD的芯片可以在很宽的温度范围(-40~85℃)内工作。因此，SSD目前广泛应用于军事、电力、医疗、航空、车载、工控、视频监控、网络监控、网络终端、导航设备等各个领域。

[6] 随着SSD技术的不断成熟，用户开始使用SSD构建存储系统，出于性能、成本的综合考虑以及保护现有的投资，用户存在着使用新的存储介质和原有存储介质搭建混合存储系统的需求。将昂贵高速的SSD作为一级存储系统，将大容量、廉价的低速机械磁盘作为后端二级存储系统，最终达到高性能，大容量，低成本的整体最佳配置。

[7] 在混合存储系统中，一般把经常访问的数据存放在快速存储设备上，不经常访问的数据存放在慢速存储设备上，因此使用较少的高速存储设备就可以使系统整体性能达到最优。

[8] 现有技术中的一种多级存储系统的构建方案为：将SSD加入到传统机械HDD (Hard Disk Drive, 硬盘驱动器) 构成的存储系统中，形成一个混合存储池，将SSD作为系

统中的第二级缓存，将磁盘上被频繁访问的数据自动缓存到SSD上，当SSD被写满之后，这些缓存数据根据ARC算法进行淘汰。

[9] 上述现有技术中的缺点为：将SSD作为系统的二级缓存，整体上提升了系统的性能，但是SSD仅仅作为缓存的作用存在，其存储空间被浪费了。上述多级存储系统只适用于热点分布非常集中的系统，如1/99的热度模型系统。但是对于热点分布集中程度较低的系统，如30/70模型（30%的数据提供70%的访问量），需要将30%的最热数据缓存在SSD中，那么系统需要额外配备的SSD存储空间就占到总存储容量的30%，这一部分存储空间对于应用来说是不感知的，不能体现在系统总存储容量中，无疑是一个巨大的浪费。

[10] 发明内容

[11] 本发明的实施例提供了一种在文件系统中实现多级存储的方法和装置，以在文件系统中实现基于文件的价值的多级存储。

[12] 本发明的目的是通过以下技术方案实现的：

[13] 一种在文件系统中实现多级存储的方法，包括：

[14] 在文件系统中分别使用一级存储设备和二级存储设备来存储文件，获取所述一级存储设备和二级存储设备中存储的各个文件的价值；

[15] 按照所述一级存储设备中存储的各个文件的价值的排列顺序建立和管理第一文件价值排序列表，按照所述二级存储设备中存储的各个文件的价值的排列顺序建立和管理第二文件价值排序列表；

[16] 根据所述第一文件价值排序列表和第二文件价值排序列表，按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移。

[17] 一种在文件系统中实现多级存储的装置，包括：

[18] 文件存储模块，用于在文件系统中分别使用一级存储设备和二级存储设备来存储文件；

[19] 文件价值处理模块，用于获取所述一级存储设备和二级存储设备中存储的各个文件的价值，按照所述一级存储设备中存储的各个文件的价值的排列顺序建立和管理第一文件价值排序列表，按照所述二级存储设备中存储的各个文件的价值的排列顺序建立和管理第二文件价值排序列表；

[20] 搬移处理模块，用于根据所述第一文件价值排序列表和第二文件价值排序列表，按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移。

[21] 由上述本发明的实施例提供的技术方案可以看出，本发明实施例通过在文件系统中分别使用一级存储设备和二级存储设备来存储文件，并基于各个文件的价值和按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移，可以实现将价值高的文件存放到性能高的存储介质（即一级存储设备）上，将价值低的文件存放在慢速存储介质（即二级存储设备）上，使系统达到整体性能最佳。

[22] 附图简要说明

[23] 图1为本发明实施例一提供的一种在文件系统中实现多级存储的方法的具体处理流程图；

[24] 图2为本发明实施例提供的一种文件的元数据信息中的字段示意图；

[25] 图3为本发明实施例二提供的一种当只根据文件的最近访问时间来确定文件的价值时，在文件系统中实现多级存储的方法的具体处理流程图；

[26] 图4为本发明实施例三提供的一种在文件系统中实现多级存储的装置的具体构图。

[27] 实施本发明的方式

[28] 下面将参考附图详细说明本发明实施例。

[29] 实施例一

[30] 该实施例提供了一种在文件系统中实现多级存储的方法，该方法的具体处理流程如图1所示，包括如下的处理步骤：

[31] 步骤11、在文件系统中分别使用一级存储设备和二级存储设备来存储文件，获取所述一级存储设备和二级存储设备中存储的各个文件的价值。

[32] 比如，可以将SSD作为一级存储设备，将机械磁盘作为二级存储设备，上述一级存储设备和二级存储设备都用于存储文件，构成了文件系统的多级整体存储空间。

[33] 本发明实施例需要计算上述一级存储设备和二级存储设备中存储的各个文件的价值，具体计算方法如下：

[34] 按照设定的文件价值更新时间间隔（比如30分钟），根据文件的最近访问时间、设定时间段内的访问次数、文件的内容、系统配置信息中的至少一项，按照设定的算法计算所述一级存储设备和二级存储设备中存储的各个文件的价值。

[35] 比如，当只根据文件的最近访问时间来确定文件的价值时，则访问时间最近的文件的价值最大。

[36] 又比如，当同时根据文件的最近访问时间、设定时间段内的访问次数来确定文件的价值时，可以将文件的最近访问时间乘以对应的权重之后，再加上设定时间段内的访问次数乘以对应的权重来计算各个文件的价值。

[37] 又比如，系统管理员可以根据实际需要来直接设定某个或者多个文件的价值。

[38] 该实施例提供的一种文件的元数据信息中的字段示意图如图2所示，在实际应用中，可以利用每个文件的元数据信息中的现有部分字段以及其他可选的扩展字段中记录的信息来计算文件的价值，并且将文件的价值信息、排序信息和文件的存储位置信息存储在元数据信息中的可选的扩展字段中。比如，在文件被访问后，系统对上述文件的元数据信息进行更新后，利用图2中时间戳字段中记录的文件的最近访问时间来计算文件的价值；或者同时利用图2中时间戳字段中记录的文件的最近和历史访问时间、文件长度字段、文件块数字段中记录的信息中的至少一项来计算文件的价值。然后，将图2中文件价值字段中存储的文件的原有价值删除，将刚计算出的文件的价值存储在文件价值字段中，并对排序信息字段进行相应的更新。

[39] 上述图2中存储位置字段中存储文件是存储在所述一级存储设备或二级存储设备中的信息，上述图2中排序信息字段中存储文件在下面将描述的第一文件价值排序列表或者第二文件价值排序列表中的排序情况。上述图2中排序信息字段、存储位置字段和文件价值字段是元数据信息中的扩展字段。

[40] 由于上述一级存储设备是快速硬盘，而上述二级存储设备是慢速硬盘，因此，在上述一级存储设备中存储价值比较高的文件，在上述机械硬盘中存储价值比较低的文件。在理想情况下，上述一级存储设备中存储的任何一个文件的价值都大于上述机械硬盘中存储的所有文件的价值。

[41] 步骤12、按照所述一级存储设备中存储的各个文件的价值的排列顺序建立和管

理第一文件价值排序列表，按照所述二级存储设备中存储的各个文件的价值的排列顺序建立和管理第二文件价值排序列表。

[42] 将所述一级存储设备中存储的各个文件按照价值的大小从大到小排序建立第一文件价值排序列表。当按照设定的文件价值更新时间间隔，对所述一级存储设备中存储的各个文件的价值进行更新后，则需要对所述第一文件价值排序列表进行相应的更新。

[43] 将所述二级存储设备中存储的各个文件按照价值的大小从大到小排序建立第二文件价值排序列表。当按照设定的文件价值更新时间间隔，对所述二级存储设备中存储的各个文件的价值进行更新后，则需要对所述第二文件价值排序列表进行相应的更新。

[44] 上述第一文件价值排序列表和第二文件价值排序列表中排列的是各个文件的标识。

[45] 步骤 13、根据所述第一文件价值排序列表和第二文件价值排序列表，按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移。

[46] 预先设定在所述一级存储设备和二级存储设备中进行文件搬移的时间间隔，并开始所述时间间隔的计时，当所述时间间隔的计时结束之后，则在所述一级存储设备和二级存储设备中进行文件搬移处理，具体文件搬移处理的过程可以为：

[47] 如果所述第二文件价值排序列表中的一个或者多个文件的价值超过了所述第一文件价值排序列表中至少一个文件的价值，将所述第二文件价值排序列表中的一个或者多个文件从二级存储设备中搬移到所述一级存储设备中，然后，对所述第一文件价值排序列表和所述第二文件价值排序列表进行更新；或者

[48] 如果所述一级存储设备的空间利用率超过了设定的阈值，将所述第一文件价值排序列表中排列靠后的一个或者多个文件从所述一级存储设备中搬移到所述二级存储设备中，直到所述一级存储设备的空间利用率没有超过设定的阈值，然后，对所述第一文件价值排序列表和所述第二文件价值排序列表进行更新。

[49] 需要进一步说明的是，如果所述二级存储设备的空间利用率超过了设定的阈值，则将所述第二文件价值排序列表中排列靠后的一个或者多个文件从所述二级

存储设备中删除，直到所述二级存储设备的空间利用率没有超过设定的阈值，然后，对所述第二文件价值排序列表进行更新。

[50] 实施例二

[51] 当只根据文件的最近访问时间来确定文件的价值时，该实施例提供了一种在文件系统中实现多级存储的方法的具体处理流程如图3所示，包括如下的处理步骤：

[52] 步骤31、文件系统中的某个文件被访问，上述文件可以为一级存储设备或者二级存储设备中存储的文件。

[53] 步骤32、由于只根据文件的最近访问时间来确定文件的价值，当前被访问的上述文件的价值最高，将上述文件的价值设置为最大。

[54] 步骤33、判断上述文件是否存储在一级存储设备中，如果是，则执行步骤34；否则，执行步骤35。

[55] 步骤34、将上述文件的标识排列在一级存储设备中的所述第一文件价值排序列表中的最前列。

[56] 步骤35、将上述文件加入到所述二级存储设备的搬移文件列表中，当预先设定的文件搬移的时间间隔的计时结束之后，将所述二级存储设备的搬移列表中的文件从二级存储设备中搬移到一级存储设备中。

[57] 实施例三

[58] 本发明实施例还提供了一种在文件系统中实现多级存储的装置，其具体结构如图4所示，包括如下的模块：

[59] 文件存储模块41，用于在文件系统中分别使用一级存储设备和二级存储设备来存储文件，比如，可以将SSD作为一级存储设备，将机械磁盘作为二级存储设备；

[60] 文件价值处理模块42，用于获取所述一级存储设备和二级存储设备中存储的各个文件的价值，按照所述一级存储设备中存储的各个文件的价值的排列顺序建立和管理第一文件价值排序列表，按照所述二级存储设备中存储的各个文件的价值的排列顺序建立和管理第二文件价值排序列表；

[61] 搬移处理模块43，用于根据所述第一文件价值排序列表和第二文件价值排序列表

表，按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移。

[62] 具体的，所述的文件价值处理模块42可以包括：

[63] 文件价值计算模块421，用于按照设定的文件价值更新时间间隔，根据文件的最近访问时间、设定时间段内的访问次数、文件的内容、系统配置信息中的至少一项确定所述一级存储设备和二级存储设备中存储的各个文件的价值。比如，在文件被访问后，系统对所述文件的元数据信息进行更新后，利用所述文件的元数据信息中的时间戳字段中记录的文件的最近访问时间来计算所述文件的价值；或者，在文件被访问后，系统对所述文件的元数据信息进行更新后，利用所述文件的元数据信息中的时间戳字段中记录的文件的最近和历史访问时间，以及文件长度字段、文件块数字段中记录的信息中的至少一项来计算文件的价值。

[64] 文件排序列表建立模块422，用于将所述一级存储设备中存储的各个文件按照价值的大小从大到小排序建立第一文件价值排序列表，根据所述一级存储设备中存储的各个文件的价值的更新结果对所述第一文件价值排序列表进行更新；

[65] 将所述二级存储设备中存储的各个文件按照价值的大小从大到小排序建立第二文件价值排序列表，根据所述二级存储设备中存储的各个文件的价值的更新结果对所述第二文件价值排序列表进行更新。

[66] 具体的，所述的搬移处理模块43：还用于预先设定在所述一级存储设备和二级存储设备中进行文件搬移的时间间隔，并开始所述时间间隔的计时，当所述时间间隔的计时结束之后，

[67] 如果所述第二文件价值排序列表中的一个或者多个文件的价值超过了所述第一文件价值排序列表中至少一个的文件的价值，将所述第二文件价值排序列表中的一个或者多个文件从二级存储设备中搬移到所述一级存储设备中；

[68] 对所述第一文件价值排序列表和所述第二文件价值排序列表进行更新。

[69] 具体的，所述的搬移处理模块43：还用于预先设定在所述一级存储设备和二级存储设备中进行文件搬移的时间间隔，并开始所述时间间隔的计时，当所述时间间隔的计时结束之后，如果所述一级存储设备的空间利用率超过了设定的阈

值，将所述第一文件价值排序列表中排列靠后的一个或者多个文件从所述一级存储设备中搬移到所述二级存储设备中，直到所述一级存储设备的空间利用率没有超过设定的阈值；

[70] 对所述第一文件价值排序列表和所述第二文件价值排序列表进行更新。

[71] 应用本发明实施例的装置在文件系统中实现多级存储的具体过程与前述方法实施例类似，此处不再赘述。

[72] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，所述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory, ROM）或随机存储记忆体（Random Access Memory, RAM）等。

[73] 综上所述，本发明实施例通过在文件系统中分别使用一级存储设备和二级存储设备来存储文件，并基于各个文件的价值和按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移，可以实现将价值高的文件存放到性能高的存储介质（即一级存储设备）上，将价值低的文件存放在慢速存储介质（即二级存储设备）上，使系统达到整体性能最佳。

[74] 本发明实施例支持对异构存储介质的统一管理，对用户呈现的整体存储空间为各级存储的空间之和；用户使用时不必区分各级存储介质，由文件系统对用户提供统一的管理。从而避免了使用额外的分级存储管理软件，减少了复杂度。

[75] 本发明实施例提供灵活的策略配置，可以灵活控制文件迁移的频度周期和文件迁移所遵循的算法。

[76] 以上所述，仅为本发明较佳的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到的变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应该以权利要求的保护范围为准。

## 权利要求书

- [权利要求 1] 一种在文件系统中实现多级存储的方法，其特征在于，包括：  
在文件系统中分别使用一级存储设备和二级存储设备来存储文件，获取所述一级存储设备和二级存储设备中存储的各个文件的价值；  
按照所述一级存储设备中存储的各个文件的价值的排列顺序建立和管理第一文件价值排序列表，按照所述二级存储设备中存储的各个文件的价值的排列顺序建立和管理第二文件价值排序列表；  
根据所述第一文件价值排序列表和第二文件价值排序列表，按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移。
- [权利要求 2] 根据权利要求 1所述的在文件系统中实现多级存储的方法，其特征在于，所述的获取所述一级存储设备和二级存储设备中存储的各个文件的价值，包括：  
在文件系统中按照设定的文件价值更新时间间隔，根据文件的最近访问时间、设定时间段内的访问次数、文件的内容、系统配置信息中的至少一项确定所述一级存储设备和二级存储设备中存储的各个文件的价值。
- [权利要求 3] 根据权利要求 2所述的在文件系统中实现多级存储的方法，其特征在于，所述的获取所述一级存储设备和二级存储设备中存储的各个文件的价值，包括：  
在文件被访问后，系统对所述文件的元数据信息进行更新后，利用所述文件的元数据信息中的时间戳字段中记录的文件的最近访问时间来计算所述文件的价值；  
或者，  
在文件被访问后，系统对所述文件的元数据信息进行更新后，利用所述文件的元数据信息中的时间戳字段中记录的文件的最近和历史访问时间，以及文件长度字段、文件块数字段中记录的信息

中的至少一项来计算文件的价值。

[权利要求 4]

根据权利要求3所述的在文件系统中实现多级存储的方法，其特征在于，所述的按照所述一级存储设备中存储的各个文件的价值的排列顺序建立和管理第一文件价值排序列表，按照所述二级存储设备中存储的各个文件的价值的排列顺序建立和管理第二文件价值排序列表，包括：

将所述一级存储设备中存储的各个文件按照价值的大小从大到小排序建立第一文件价值排序列表，根据所述一级存储设备中存储的各个文件的价值的更新结果对所述第一文件价值排序列表进行更新；

将所述二级存储设备中存储的各个文件按照价值的大小从大到小排序建立第二文件价值排序列表，根据所述二级存储设备中存储的各个文件的价值的更新结果对所述第二文件价值排序列表进行更新。

[权利要求 5]

根据权利要求1所述的在文件系统中实现多级存储的方法，其特征在于，所述的一级存储设备包括固态硬盘；所述的二级存储设备包括机械磁盘。

[权利要求 6]

根据权利要求1-5任一项所述的在文件系统中实现多级存储的方法，其特征在于，所述的根据所述第一文件价值排序列表和第二文件价值排序列表，按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移，包括：

预先设定在所述一级存储设备和二级存储设备中进行文件搬移的时间间隔，并开始所述时间间隔的计时，当所述时间间隔的计时结束之后，

如果所述第二文件价值排序列表中的一个或者多个文件的价值超过了所述第一文件价值排序列表中至少一个的文件的价值，将所述第二文件价值排序列表中的一个或者多个文件从二级存储设备中搬移到所述一级存储设备中；

对所述第一文件价值排序列表和所述第二文件价值排序列表进行更新。

[权利要求 7]

根据权利要求 1-6 任一项所述的在文件系统中实现多级存储的方法，其特征在于，所述的根据所述第一文件价值排序列表和所述第二文件价值排序列表，按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移，包括：

预先设定在所述一级存储设备和二级存储设备中进行文件搬移的时间间隔，并开始所述时间间隔的计时，当所述时间间隔的计时结束之后，

如果所述一级存储设备的空间利用率超过了设定的阈值，将所述第一文件价值排序列表中排列靠后的一个或者多个文件从所述一级存储设备中搬移到所述二级存储设备中，直到所述一级存储设备的空间利用率没有超过设定的阈值；

对所述第一文件价值排序列表和所述第二文件价值排序列表进行更新。

[权利要求 8]

一种在文件系统中实现多级存储的装置，其特征在于，包括：

文件存储模块，用于在文件系统中分别使用一级存储设备和二级存储设备来存储文件；

文件价值处理模块，用于获取所述一级存储设备和二级存储设备中存储的各个文件的价值，按照所述一级存储设备中存储的各个文件的价值的排列顺序建立和管理第一文件价值排序列表，按照所述二级存储设备中存储的各个文件的价值的排列顺序建立和管理第二文件价值排序列表；

搬移处理模块，用于根据所述第一文件价值排序列表和所述第二文件价值排序列表，按照设定的搬移策略在所述一级存储设备和二级存储设备中进行文件搬移。

[权利要求 9]

根据权利要求 8 所述的在文件系统中实现多级存储的装置，其特征在于，所述的文件价值处理模块包括：

文件价值计算模块，用于按照设定的文件价值更新时间间隔，根据文件的最近访问时间、设定时间段内的访问次数、文件的内容、系统配置信息中的至少一项确定所述一级存储设备和二级存储设备中存储的各个文件的价值；

文件排序列表建立模块，用于将所述一级存储设备中存储的各个文件按照价值的大小从大到小排序建立第一文件价值排序列表，根据所述一级存储设备中存储的各个文件的价值的更新结果对所述第一文件价值排序列表进行更新；

将所述二级存储设备中存储的各个文件按照价值的大小从大到小排序建立第二文件价值排序列表，根据所述二级存储设备中存储的各个文件的价值的更新结果对所述第二文件价值排序列表进行更新。

[权利要求 10]

根据权利要求9所述的在文件系统中实现多级存储的装置，其特征在于：

所述的文件价值计算模块，还用于在文件被访问后，系统对所述文件的元数据信息进行更新后，利用所述文件的元数据信息中的时间戳字段中记录的文件的最近访问时间来计算所述文件的价值；

或者，

在文件被访问后，系统对所述文件的元数据信息进行更新后，利用所述文件的元数据信息中的时间戳字段中记录的文件的最近和历史访问时间、文件长度字段、文件块数字段中记录的信息中的至少一项来计算文件的价值。

[权利要求 11]

根据权利要求8或9或10所述的在文件系统中实现多级存储的装置，其特征在于：

所述的搬移处理模块，还用于预先设定在所述一级存储设备和二级存储设备中进行文件搬移的时间间隔，并开始所述时间间隔的计时，当所述时间间隔的计时结束之后，

如果所述第二文件价值排序列表中的一个或者多个文件的价值超过了所述第一文件价值排序列表中至少一个的文件的价值，将所述第二文件价值排序列表中的一个或者多个文件从二级存储设备中搬移到所述一级存储设备中；

对所述第一文件价值排序列表和所述第二文件价值排序列表进行更新。

[权利要求 12]

根据权利要求8或9或10所述的在文件系统中实现多级存储的装置，其特征在于：

所述的搬移处理模块，还用于预先设定在所述一级存储设备和二级存储设备中进行文件搬移的时间间隔，并开始所述时间间隔的计时，当所述时间间隔的计时结束之后，如果所述一级存储设备的空间利用率超过了设定的阈值，将所述第一文件价值排序列表中排列靠后的一个或者多个文件从所述一级存储设备中搬移到所述二级存储设备中，直到所述一级存储设备的空间利用率没有超过设定的阈值；

对所述第一文件价值排序列表和所述第二文件价值排序列表进行更新。

1/2

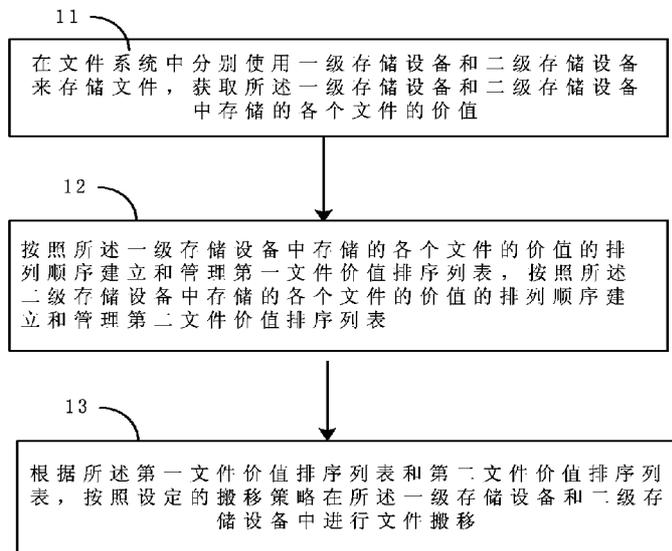


图1

保护模式
用户和组
文件块数
文件长度
时间戳
链接数
直接指针
一级间接指针
二级间接指针
排序信息
文件价值
存储位置

元数据信息

图2

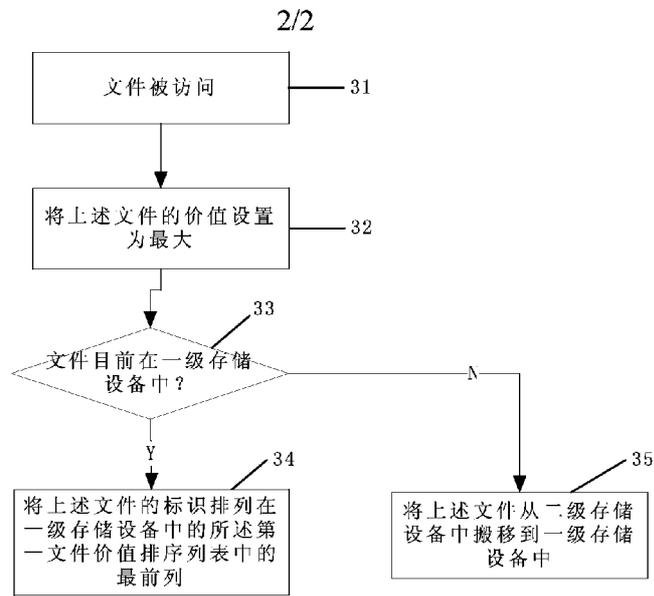


图3

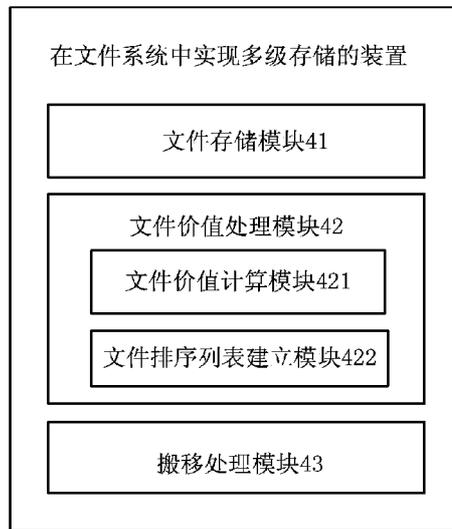


图4

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2011/084830

## A. CLASSIFICATION OF SUBJECT MATTER

G06F 12/08 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06F, H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, WPI, EPODOC, CNKI, GOOGLE: HSM, migration, movement, ILM, data w transfer, data w migration, life w cycle, value or valuation, hierarchical w storage, file, list

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LV, Shuai et al., Research on the Data Migration Strategy of Hierarchical Mass Storage System, COMPUTER ENGINEERING & SCIENCE, vol. 31, no. A1, 2009, pages 163-167., ISSN 1007-130X, sections 1-4, and figure 2	1-5, 8-10
Y		6-7, 11-12
Y	CN 10 110 1563 A (TSINGHUA UNIVERSITY), 09 January 2008 (09.01 .2008), claim 1	6-7, 11-12
A	CN 102291450 A (LANGCHAO ELECTRONIC INFORMATION INDUSTRY CO., LTD.), 21 December 2011 (21.12.2011), the whole document	1-12
A	US 2006015529 A I (HITACHI, LTD.), 19 January 2006 (19.01 .2006), the whole document	1-12
A	US 2008154988 A I (FUJITSU LIMITED), 26 June 2008 (26.06.2008), the whole document	1-12

Further documents are listed in the continuation of Box C.  See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search  
06 September 2012 (06.09.2012)

Date of mailing of the international search report  
18 October 2012 (18.10.2012)

Name and mailing address of the ISA/CN:  
State Intellectual Property Office of the P. R. China  
No. 6, Xitucheng Road, Jimenqiao  
Haidian District, Beijing 100088, China  
Facsimile No.: (86-10) 62019451

Authorized officer  
FAN, Yuxia  
Telephone No.: (86-10) 62413706

INTERNATIONAL SEARCH REPORT  
 Information on patent family members

International application No.  
 PCT/CN2011/084830

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101 101563 A	09.01 .2008	None	
CN 102291450 A	21.12.201 1	None	
US 2006015529 A I	19.01 .2006	JP 2006031668 A	02.02.2006
US 2008154988 A I	26.06.2008	W O 2006131978 A I	14.12.2006
		JPWO 20061 31978 A I	08.01 .2009

A. 主题的分类		
G06F 12/08(2006.01)1		
按照国际专利分类(IPC) 或者同时按照国家分类和 IPC 两种分类		
B. 检索领域		
检索的最低限度文献(标明分类系统和分类号)		
IPC: G06F, H04L		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))		
CNP AT, WPI,EPODOC,CNKI, GOOGLE: 分级存储, HSM, 迁移, 搬移, 生命周期, ILM, 文件, 价值, 列表, data w transfer, data w migration, life w cycle, value or valuation, hierarchical w storage, file, list		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	吕帅等, 海量信息分级存储数据迁移策略研究, 计算机工程与科学, 第 31 卷第 A 1 期, 2009 年, 第 163-167 页, ISSN 1007-130X, 第 1-4 节, 图 2	1-5,8-10
Y		6-7, 11-12
Y	CN101 101563 A (清华大学) 09. 1 月 2008(09.01.2008) 权利要求 1	6-7, 11-12
A	CN102291450A (浪潮电子信息产业股份有限公司) 21. 12 月 201 1 (21. 12.201 1) 全文	1-12
A	US2006015529A1 (HITACHI, LTD. ) 19. 1 月 2006(19.01.2006) 全文	1-12
A	US2008154988A1 (FUJITSU LIMITED ) 26. 6 月 2008(26.06.2008) 全文	1-12
<input type="checkbox"/> 其余文件在 C 栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。		
* 引用文件的具体类型:		
"A" 认为不特别相关的表示了现有技术一般状态的文件	"T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件	
"E" 在国际申请日的 3/4 之前公布的在先申请或专利	"X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性	
"L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)	"Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性	
"O" 涉及口头公开、使用、展览或其他方式公开的文件	"&" 同族专利的文件	
"P" 公布日先于国际申请日但迟于所要求的优先权日的文件		
国际检索实际完成的日期 06. 9 月 20 12 (06.09.2012)	国际检索报告邮寄日期 18.10 月 2012 (18.10.2012)	
ISA/CN 的名称和邮寄地址: 中华人民共和国国家知识产权局 中国北京市海淀区蓟门桥西土城路 6 号 100088 传真号: (86-10)62019451	受权官员  范玉霞  电话号码: (86-10) 62413706	

国际检索报告

关于同族专利的信息

国际申请号

PCT/CN201 1/084830

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN101 101563A	09.01.2008	无	
CN102291450A	21. 12.201 1	无	
US2006015529A1	19.01.2006	JP200603 1668A	02.02.2006
US2008154988A1	26.06.2008	WO200613 1978A1	14. 12.2006
		JPWO200613 1978A1	08.01.2009