

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2012年8月16日 (16.08.2012)



(10) 国际公布号
WO 2012/106923 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2011/077564
- (22) 国际申请日: 2011年7月25日 (25.07.2011)
- (25) 申请语言: 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)。
- (72) 发明人: 及
- (75) 发明人/申请人 (仅对美国): **卢庆明 (LU, Qingming)** [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: **北京同立钧成知识产权代理有限公司 (LEADER PATENT & TRADEMARK FIRM)**; 中国北京市海淀区西直门北大街32号枫蓝国际A座8F-6, Beijing 100082 (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, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, 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, 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))。
- 在修改权利要求的期限届满之前进行, 在收到该修改后将重新公布(细则48.2(h))。
- 根据申请人的请求, 在条约第21条(2)(a)所规定的期限届满之前进行。

(54) Title: OBJECT ADJUSTING METHOD, MIGRATION CONTROL DEVICE, NODE DEVICE AND STORAGE SYSTEM

(54) 发明名称: 对象调整方法、迁移控制设备、节点设备及存储系统

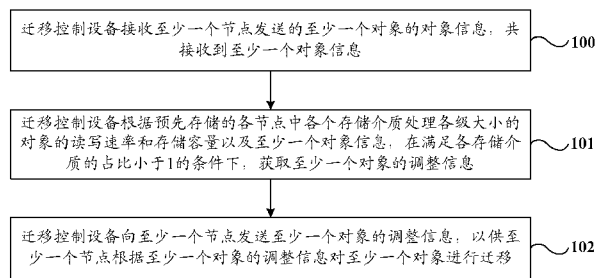


图1 / Fig.1

100 THE MIGRATION CONTROL DEVICE RECEIVES OBJECT INFORMATION ABOUT AT LEAST ONE OBJECT SENT BY AT LEAST ONE NODE, AND AT LEAST ONE PIECE OF OBJECT INFORMATION IS RECEIVED

101 THE MIGRATION CONTROL DEVICE ACQUIRES ADJUSTMENT INFORMATION ABOUT THE AT LEAST ONE OBJECT ACCORDING TO THE PRE-STORED READ/WRITE RATE AND THE STORAGE CAPACITY FOR EACH STORAGE MEDIUM IN EACH NODE TO PROCESS OBJECTS OF EACH LEVEL AND AT LEAST ONE PIECE OF OBJECT INFORMATION UNDER THE CONDITION THAT THE DUTY RATIO OF EACH STORAGE MEDIUM IS SMALLER THAN 1

102 THE MIGRATION CONTROL DEVICE SENDS THE ADJUSTMENT INFORMATION ABOUT THE AT LEAST ONE OBJECT TO THE AT LEAST ONE NODE, SO THAT THE AT LEAST ONE NODE MIGRATES THE AT LEAST ONE OBJECT ACCORDING TO THE ADJUSTMENT INFORMATION ABOUT THE AT LEAST ONE OBJECT

(57) Abstract: Provided are an object adjusting method, a migration control device, a node device and a storage system. The method includes: receiving object information about at least one object sent by at least one node; acquiring adjustment information about the at least one object according to the pre-stored read/write rate and the storage capacity for each storage medium in each node to process objects of each level and at least one piece of object information under the condition that the duty ratio of each storage medium is smaller than 1; and sending the adjustment information about the at least one object to the at least one node. The technical solution of the embodiments of the present invention can effectively ensure that the storage performance of the storage system of each node is optimal.

[见续页]



WO 2012/106923 A1



(57) 摘要:

本发明实施例提供一种对象调整方法、迁移控制设备、节点设备及存储系统。该方法包括:接收至少一个节点发送的至少一个对象的对象信息;根据预先存储的各节点中各个存储介质处理各级大小的对象的读写速率和存储容量以及至少一个对象信息,在满足各存储介质的占比小于1的条件下获取至少一个对象的调整信息;向至少一个节点发送至少一个对象的调整信息。本发明实施例的技术方案能够有效地保证包括各节点的存储系统的存储性能最优。

对象调整方法、迁移控制设备、节点设备及存储系统

技术领域

本发明实施例涉及数据存储技术领域，尤其涉及一种对象调整方法、
5 迁移控制设备、节点设备及存储系统。

背景技术

随着计算机网络技术的发展，用于存储计算机网络中数据的存储系统的使用越来越广泛。

10 在很多存储服务的场景（如视频点播、Web Cache 等）下，用户访问存储对象的频度是有很大差异的，根据统计，对象分发的热度分布符合齐普夫定律（Zipf's Law），热点对象的访问往往是冷门对象访问的数十倍，同时热点对象的规模远小于热度一般及冷门对象的规模，因此，提高热点对象的访问速度是提升存储系统整体性能的关键。为了有效地提高存储系
15 统的整体性能，现有技术中通常采用增加热点对象的副本数，例如从存储系统中的多个节点上同时向访问的用户提供热点对象的存储服务。该种方式在业界通常称为协作 Cache。其中对象可以为在存储系统中存储的数据单元，例如可以为一个数据块。

20 在实现本发明过程中，发明人发现现有技术中至少存在如下问题：现有的协作 cache 的方法，同时存在多个对象副本，在对多个对象进行更新时会导致多个对象副本不一致的问题，影响存储系统的性能。

发明内容

25 本发明实施例提供一种对象调整方法、迁移控制设备、节点设备及存储系统，以实现存储系统性能的最大化。

本发明实施例提供一种对象调整方法，包括：

接收至少一个节点发送的至少一个对象的对象信息，共接收到至少一个对象信息；所述对象信息中包括对象标识、对象大小、对象的第一位置和对象的访问热度；所述对象的访问热度指的是所述对象在上一周期内被访问的次数；

5 根据预先存储的各所述节点中各个存储介质处理各级大小的对象的读写速率和存储容量以及所述至少一个对象信息，在满足各所述存储介质的占比小于1的条件下，获取所述至少一个对象的调整信息；所述调整信息中包括所述对象标识和所述对象标识对应的对象被调整后的第二位置，所述至少一个对象的第二位置能够保证所述至少一个对象在对应的所述
10 第二位置上的读写速率与所述访问热度乘积之和最大；所述存储介质的占比指的是在上一周期内，所述存储介质上存储的各对象在所述存储介质上的访问热度和所述对象的大小的乘积除以所述存储介质的最大吞吐率之后的累加；

向所述至少一个节点发送所述至少一个对象的调整信息，以供所述至少一个节点根据所述至少一个对象的调整信息对所述至少一个对象进行
15 迁移。

本发明实施例还提供一种对象调整方法，包括：

向迁移控制设备发送第一对象的第一对象信息，以供所述迁移控制设备根据预先存储的至少一个节点中各所述节点中的各个存储介质处理各
20 级大小的对象的读写速率和存储容量、所述第一对象信息以及其他节点发送的第二对象信息，获取所述第一对象的调整信息；所述第一对象信息中包括第一对象标识、第一对象大小、第一对象的第一位置和第一对象的访问热度；所述对象的访问热度指的是所述对象在上一周期内被访问的次数；所述调整信息中包括所述第一对象标识和所述第一对象被调整后的第二位置，所述第一对象的所述第二位置和所述第二对象信息对应的对象被
25 所述迁移控制设备调整后的位置能够保证所述第一对象在所述第二位置

上和所述第二对象信息对应的对象在所述被调整后的位置上的读写速率与访问热度乘积之和最大；所述其他节点包括所述至少一个节点中当前节点之外的节点；

接收所述迁移控制设备发送的所述第一对象的调整信息；

5 根据所述第一对象的调整信息对所述第一对象进行迁移。

本发明实施例提供一种迁移控制设备，包括：

接收模块，用于接收至少一个节点发送的至少一个对象的对象信息，共接收到至少一个对象信息；所述对象信息中包括对象标识、对象大小、对象的第一位置和对象的访问热度；所述对象的访问热度指的是所述对象
10 在上一周期内被访问的次数；

获取模块，用于根据预先存储的各所述节点中各个存储介质处理各级大小的对象的读写速率和存储容量以及所述至少一个对象信息，在满足各所述存储介质的占比小于1的条件下，获取所述至少一个对象的调整信息；所述调整信息中包括所述对象标识和所述对象标识对应的对象被调整后的第二位置，所述至少一个对象的第二位置能够保证所述至少一个对象在
15 对应的所述第二位置上的读写速率与所述访问热度乘积之和最大；所述存储介质的占比指的是在上一周期内，所述存储介质上存储的各对象在所述存储介质上的访问热度和所述对象的大小的乘积除以所述存储介质的最大吞吐率之后的累加；

20 发送模块，用于向所述至少一个节点发送所述至少一个对象的调整信息，以供所述至少一个节点根据所述至少一个对象的调整信息对所述至少一个对象进行迁移。

本发明实施例提供一种节点设备，包括：

发送模块，用于向迁移控制设备发送第一对象的第一对象信息，以供
25 所述迁移控制设备根据预先存储的至少一个节点中各所述节点中的各个存储介质处理各级大小的对象的读写速率和存储容量、所述第一对象信息

以及其他节点发送的第二对象信息，获取所述第一对象的调整信息；所述第一对象信息中包括第一对象标识、第一对象大小、第一对象的第一位置和第一对象的访问热度；所述对象的访问热度指的是所述对象在上一周期内被访问的次数；所述调整信息中包括所述第一对象标识和所述第一对象

5 被调整后的第二位置，所述第一对象的所述第二位置和所述第二对象信息对应的对象被所述迁移控制设备调整后的位置能够保证所述第一对象在所述第二位置上和所述第二对象信息对应的对象在所述被调整后的位置上的读写速率与访问热度乘积之和最大；所述其他节点包括所述至少一个节点中当前节点之外的节点；

- 10 接收模块，用于接收所述迁移控制设备发送的所述第一对象的调整信息；
调整模块，用于根据所述第一对象的调整信息对所述第一对象进行迁移。

本发明实施例还提供一种存储系统，包括迁移控制设备和至少一个节点设备；

所述至少一个节点设备，用于分别向迁移控制设备发送对象信息；

- 15 所述迁移控制设备，用于分别接收所述至少一个节点设备发送的对象信息，共接收到至少一个对象信息；各所述对象信息中包括对象标识、对象大小、对象的第一位置和对象的访问热度；所述对象的访问热度指的是所述对象在上一周期内被访问的次数；根据预先存储的所述至少一个节点设备的各所述节点设备中各个存储介质处理各级大小的对象的读写速率
- 20 和存储容量以及所述至少一个对象信息，在满足各所述存储介质的占比小于1的条件下，获取所述至少一个对象的调整信息；所述调整信息中包括所述对象标识和所述对象标识对应的对象被调整后的第二位置，所述至少一个对象的第二位置能够保证所述至少一个对象在对应的所述第二位置上的读写速率与所述访问热度乘积之和最大；所述存储介质的占比指的是
- 25 在上一周期内，所述存储介质上存储的各对象在所述存储介质上的访问热度和所述对象的大小的乘积除以所述存储介质的最大吞吐率之后的累加；

向所述至少一个节点设备发送所述至少一个对象的调整信息；

所述至少一个节点设备，还用于接收所述迁移控制设备发送的所述对象的调整信息；根据所述对象的调整信息对所述对象进行迁移。

本发明实施例的对象调整方法、迁移控制设备、节点设备及存储系统，
5 通过采用本发明实施例的上述技术方案，能够实现各对象标识对应的对象在调整后的第二位置上使得各对象的读写速率与访问热度乘积之和最大，从而有效地保证了包括各节点的存储系统的存储性能达到最大。

附图说明

10 为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 为本发明一实施例提供的对象调整方法的流程图。

15 图 2 为本发明另一实施例提供的对象调整方法的流程图。

图 3 为本发明再一实施例提供的对象调整方法的流程图。

图 4 为本发明一实施例提供的迁移控制设备的结构示意图。

图 5 为本发明另一实施例提供的迁移控制设备的结构示意图。

图 6 为本发明一实施例提供的节点设备的结构示意图。

20 图 7 为本发明另一实施例提供的节点设备的结构示意图。

图 8 为本发明实施例提供的存储系统的结构示意图。

具体实施方式

为使本发明实施例的目的、技术方案和优点更加清楚，下面将结合本
25 发明实施例中的附图，对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。

基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

图 1 为本发明一实施例提供的对象调整方法的流程图。本实施例的对象调整方法的执行主体为一个迁移控制设备。如图 1 所示，本实施例的对象调整方法，具体可以包括如下：

100、迁移控制设备接收至少一个节点发送的至少一个对象的对象信息，共接收到至少一个对象信息；

其中每一个对象信息中包括对象标识、对象大小、对象的第一位置和对象的访问热度。

10 本实施例中的对象例如可以为用以存储数据的数据块等。本实施例中的至少一个节点均为用于存储对象的存储设备。该至少一个节点中共包括至少一个存储介质，具体地在每一个节点中包括几个存储介质不做限制，例如其中某一个节点中可以只有一个存储介质；也可以包括多个存储介质。而且在某一个节点中也可以包括两个以上存储介质类型相同的存储介
15 质。本实施例中的至少一个存储介质中可以包括固态硬盘(Solid state disk; 以下简称 SSD)、磁带库、物理硬盘等各种存储性能不相同的存储介质。

本实施例的应用场景为在某一个周期结束之后，迁移控制设备接收至少一个节点上报的该周期内该节点上存储的对象的对象信息，例如对象信息中包括对象标识、对象大小、对象的第一位置和对象的访问热度。其中
20 对象的访问热度即为该上一周期中该对象被访问的次数。对象的第一位置为该对象标识对应的对象在对应节点中存储的位置，可以根据该第一位置确定该对象在该节点中的哪一个存储介质上存储。

101、迁移控制设备根据预先存储的各节点中各个存储介质处理各级大小的对象的读写速率和存储容量以及至少一个对象信息，在满足各存储
25 介质的占比小于 1 的条件下，获取至少一个对象的调整信息；

其中调整信息中包括对象标识和对象标识对应的对象被调整后的第

二位置，至少一个对象的第二位置能够保证在各存储介质的容量以及占比
小于 1 的约束条件下，至少一个对象在对应的第二位置上使得至少一个对
象的读写速率与访问热度的乘积之和最大。这里的至少一个对象即指的是
至少一个节点上报的所有对象，可以理解为至少一个节点组成的存储系统
5 的所有对象。存储介质的占比指的是在上一周期内，该存储介质上存储的
各对象在该存储介质上的访问热度和该对象大小的乘积除以该存储介质
的最大吞吐率之后的累加；即可以采用如下公式表示：
$$P = \sum_i \frac{S_i \times L}{T_{\text{arg est}}}$$
。其中 P
表示存储介质的占比，i 表示上一周期该存储介质上被访问的对象， S_i 表
示对象 i 的大小，L 表示对象 i 在上一周期的访问热度， $T_{\text{arg est}}$ 表示对象 i
10 的最大吞吐率。换个角度，上述的占比也可以理解为一个周期内（例如可
以为单位周期内）存储介质满足其上所有对象存取所花费的时间与单位时
间的比值，用以体现存储介质的繁忙程度。

102、迁移控制设备向至少一个节点发送至少一个对象的调整信息，
以供至少一个节点根据至少一个对象的调整信息对至少一个对象进行迁
15 移。

例如，该迁移控制设备可以根据各对象的第一位置确定该对象存储在
哪个节点。然后向该节点发送该对象的调整信息，以供该节点根据该调整
信息对该对象进行迁移。

本实施例的对象调整方法，通过采用本实施例的上述技术方案，能够
20 实现各对象标识对应的对象在调整后的第二位置上使得各对象的读写速
率与访问热度乘积之和最大，从而有效地保证了包括各节点的存储系统的
存储性能达到最大。

可选地，上述实施例的技术方案中的 101 “迁移控制设备根据预先存
储的各节点中各个存储介质处理各级大小的对象的读写速率和存储容量
25 以及至少一个对象信息，在满足各存储介质的占比小于 1 的条件下，获取

至少一个对象的调整信息，具体可以包括如下：

(1) 迁移控制设备根据至少一个节点的各节点中各个存储介质处理各级大小的对象的读写速率和存储容量以及至少一个对象信息，在满足存储介质的占比小于 1 的条件下，获取至少一个对象被调整后的所述第二位置；

该步骤中存储介质的存储容量实质上可以理解为优化的约束条件。

(2) 迁移控制设备根据至少一个对象的第一位置和第二位置，获取至少一个对象的调整信息。

例如上述步骤 (1) 具体可以通过以下步骤实现：

10 (a) 迁移控制设备将至少一个对象信息按照对象大小和对象访问热度进行分类，得到至少一个对象组；

通过将至少一个对象分类得到对象组，能够进一步减少下述步骤 (b) 的计算复杂度。否则，若直接根据各节点中各个存储介质处理各级大小的对象的读写速率，在各存储介质的容量以及占比小于 1 的约束条件下，以至少一个对象的读写速率与访问热度乘积之和最大为目标，获取至少一个对象被调整后的第二位置，将会大大增加计算的复杂度。

20 (b) 迁移控制设备根据各节点中各个存储介质处理各级大小的对象的读写速率，在各存储介质的容量以及占比小于 1 的约束条件下，以至少一个对象的读写速率与访问热度乘积之和最大为目标，获取至少一个对象组在各存储介质上分配比例；

本实施例中，该步骤可以通过软件来实现。这里的至少一个对象的读写速率与访问热度乘积之和最大为目标，即指的是系统中所有对象的读写速率与访问热度乘积之和最大为目标。

25 (c) 迁移控制设备根据至少一个对象组在各存储介质上分配比例，获取至少一个对象被调整后的所述第二位置。

需要说明的是，当按照对象组在各存储介质上的分配比例，将该对象

组中的各个对象信息按照分配比例分配至对应的存储介质上时，可以有多种分配方式，此处不做限制，只要能够满足该对象组在该存储介质上的分配比例即可。

5 优选地，(c)中迁移控制设备可以这里还需要考虑各对象对应的第一位置，即该对象未被调整时的位置，即根据至少一个对象的第一位置以及至少一个对象组在各存储介质上分配比例，按照位置就近原则获取至少一个对象被调整后的所述第二位置，这样可以提高后续迁移效率。

10 之后，迁移控制设备根据至少一个对象的第一位置和第二个位置，可以获取到至少一个对象的调整信息，并将至少一个对象的调整信息发给对应的节点，各个节点可以将该对象从该节点中的第一位置迁移到该节点或者其他节点中的第二位置。

可选地，上述实施例中的步骤 100 之前，还包括：迁移控制设备收集至少一个节点中各节点中各个存储介质处理各级大小的对象的读写速率和存储容量。

15 图 2 为本发明另一实施例提供的对象调整方法的流程图。本实施例的对象调整方法的执行主体为一个节点，该节点可以位于存储系统中。如图 2 所示，本实施例的对象调整方法，具体可以包括如下：

20 200、节点向迁移控制设备发送第一对象的第一对象信息，以供迁移控制设备根据预先存储的至少一个节点各节点中各个存储介质处理各级大小的对象的读写速率和存储容量、第一对象信息以及其他节点发送的第二对象信息，在满足各存储介质的占比小于 1 的条件下，获取第一对象的调整信息；

25 其中第一对象信息中包括第一对象标识、第一对象大小、第一对象的第一位置和第一对象的访问热度。对象的访问热度指的是对象在上一周期内被访问的次数。该调整信息中包括第一对象标识和第一对象被调整后的第二位置，第一对象的第二位置和第一对象信息对应的对象被迁移控制设

备调整后的位置能够保证第一对象在第二位置上和第二对象信息对应的对象在被调整后的位置上的读写速率与访问热度乘积之和最大。存储介质的占比指的是在上一周期内，该存储介质上存储的各对象在该存储介质上的访问热度和该对象的大小的乘积除以该存储介质的最大吞吐率之后的累加。本实施例中执行主体的节点可以称为当前节点，其他节点包括至少一个节点中除当前节点之外的节点。

本实施例中该第一对象的第一位置和第二位置可以反映出该第一对象存储在哪个节点上以及在该节点的哪个存储介质上。

201、节点接收迁移控制设备发送的第一对象的调整信息；

10 202、节点根据第一对象的调整信息对第一对象进行迁移。

本实施例中是以一个节点上报一个对象信息为例来说明本发明实施例的技术方案。在该节点对应的存储系统中，可以预先设置一个周期，每一周期结束之后，节点都向迁移控制器发送该周期内各个对象的对象信息，以供迁移控制器根据接收到该节点上报的对象信息以及其他节点上报的对象信息，统筹考考虑，在满足所有对象在被调整后的位置上的读写速率与访问热度的乘积之和最大的原则下，为各个对象分配被调整后的第二位置，从而得到各个对象的调整信息，例如本实施例中迁移控制设备可以获取到第一对象的调整信息，并将第一对象的调整信息发送给该节点。该节点接收到该第一对象的调整信息之后，可以根据该第一对象的调整信息对第一对象进行迁移。

本实施例的对象调整方法，通过采用本实施例的上述技术方案，能够实现各对象标识对应的对象在调整后的第二位置上使得各对象的读写速率与对象的访问热度乘积之和最大，从而有效地保证了包括各节点的存储系统的存储性能达到最大。

25 可选地，在上述实施例的技术方案中的 200 之前，还包括统计第一对象的访问热度，生成第一对象信息。

可选地，上述实施例的技术方案中的 200 具体可以包括如下：

(1) 节点从第一对象的调整信息中获取第二位置;

(2) 节点判断第二位置为节点内的还是节点外的, 当第二位置为节点内, 执行(3); 否则当第二位置为节点外时, 执行(4);

(3) 节点将第一对象由第一位置迁移至第二位置;

5 (4) 节点向第二位置所在的节点发送迁移对象请求, 迁移对象请求中包括第二位置和第一对象, 以供第二位置所在的节点将第一对象存储在第二位置处。

可选地, 当第二位置为节点外时, 还包括:

10 (a) 节点接收第二位置所在节点发送的确认消息; 该确认消息中包括第一对象标识;

(b) 节点删除确认消息中的第一对象标识对应的第一对象, 以备存储新的对象。

需要说明的是, 当节点中存储有该节点中的对象的索引信息的时候, 节点还可以在删除确认消息中的第一对象标识对应的第一对象后, 更新第
15 一对象标识对应的索引信息。

图3为本发明再一实施例提供的对象调整方法的流程图。如图3所示, 本实施例的对象调整方法, 具体可以包括如下:

300、当周期结束, 各个节点统计该周期中所有对象的热度, 生成所有对象的对象信息, 执行301;

20 其中对象信息包括对象标识、对象大小、对象的第一位置和对象的访问热度。对象的第一位置即为对象被调整前的位置。

301、各个节点将所有对象的对象信息上报给迁移控制设备, 执行302;

302、迁移控制设备接收所有对象的对象信息, 将所有对象的对象信息按照对象大小和对象访问热度进行分类, 得到多个对象组, 执行303;

25 即将对象大小相当和访问热度相当的对象对象信息分为一类, 归为一个对象组。

303、迁移控制设备根据各节点中各个存储介质处理各级大小的对象的读写速率，在各存储介质的容量以及占比小于1的约束条件下，以至少一个对象的读写速率与访问热度乘积之和最大为目标，获取多个对象组在各存储介质上分配比例；执行304；

5 304、迁移控制设备根据所有对象的第一位置以及多个对象组在各存储介质上分配比例，按照位置就近的原则获取各个对象被调整后的第二位置，执行305；

305、迁移控制设备根据各个对象的第一位置和 second 位置，获取各个对象的调整信息，执行306；

10 306、迁移控制设备根据各个对象的第一位置向对应的节点发送该对象的调整信息，执行307；

各个对象的第一位置能够反映出该对象的节点信息。按照此步骤，迁移控制设备可以将每一个对象的调整信息发送至对应的节点。

307、各个节点接收迁移控制设备发送的对象的调整信息，执行308；

15 308、在每一个节点中，节点判断接收到的对象的调整信息是否为节点内的调整，如是执行309，否则执行310；

309、节点将对象从当前的第一位置迁移至第二位置；执行313；

310、节点向第二位置对应的节点发送迁移对象请求，该迁移对象请求中包括第二位置和要迁移的对象；执行311；

20 311、第二位置对应的节点接收迁移对象请求，并将要迁移的对象存储在第二位置处，并向第一位置所在的节点发送确认消息；执行312；

312、第一位置所在的节点接收确认消息，删除第一位置处存储的已经迁移成功的对象，以备存储新的对象；执行313；

313、各个节点在迁移对象成功之后，更新节点中的索引信息。

25 本实施例的对象调整方法，通过采用本实施例的上述技术方案，能够实现各对象标识对应的对象在调整后的位置上使得各对象的读写速率与访

问热度乘积之和最大，从而有效地保证了包括各节点的存储系统的存储性能达到最大。

本领域普通技术人员可以理解：实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成，前述的程序可以存储于一计算机可读存储介质中，该程序在执行时，执行包括上述方法实施例的步骤；而前述的存储介质包括：ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

图4为本发明一实施例提供的迁移控制设备的结构示意图。如图4所示，本实施例的迁移控制设备，具体可以包括：接收模块10、获取模块11和发送模块12。

其中接收模块10用于接收至少一个节点发送的至少一个对象的对象信息，共接收到至少一个对象信息；该对象信息中包括对象标识、对象大小、对象的第一位置和对象的访问热度；该对象的访问热度指的是该对象在上一周期内被访问的次数。获取模块11与接收模块10连接，获取模块11用于根据预先存储的各节点中各个存储介质处理各级大小的对象的读写速率和存储容量以及接收模块10接收的至少一个对象信息，在满足各存储介质的占比小于1的条件下，获取至少一个对象的调整信息；该调整信息中包括对象标识和对象标识对应的对象被调整后的第二位置，该至少一个对象的第二位置能够保证至少一个对象在对应的第二位置上的读写速率与访问热度乘积之和最大；该存储介质的占比指的是在上一周期内，该存储介质上存储的各对象在存储介质上的访问热度和对象的大小的乘积除以存储介质的最大吞吐率之后的累加。发送模块12与获取模块11连接，发送模块12用于向至少一个节点发送获取模块11获取的至少一个对象的调整信息，以供至少一个节点根据至少一个对象的调整信息对至少一个对象进行迁移。

本实施例的迁移控制设备，通过采用上述模块实现对象调整的实现机

制与上述相关方法实施例的实现机制相同，详细可以参考上述相关方法实施例的记载，在此不再赘述。

本实施例的迁移控制设备，通过采用本实施例的上述模块的技术方案，能够实现各对象标识对应的对象在调整后的第二位置上使得各对象的
5 读写速率与访问热度乘积之和最大，从而有效地保证了包括各节点的存储系统的存储性能达到最大。

图 5 为本发明另一实施例提供的迁移控制设备的结构示意图。如图 5 所示，本实施例的迁移控制设备在上述图 4 所示实施例的基础上，获取模块 11 具体用于根据各节点中各个存储介质处理各级大小的对象的读写速
10 率和存储容量以及接收模块 10 接收的至少一个对象信息，在满足存储介质的占比小于 1 的条件下，获取至少一个对象被调整后的第二位置；根据至少一个对象的第一位置和第二位置，获取至少一个对象的调整信息。

可选地，本实施例的迁移控制设备中的获取模块 11，包括：分类单元 111、第一获取单元 112 和第二获取单元 113。

15 其中分类单元 111 与接收模块 10 连接，分类单元 111 用于将接收模块 10 接收的至少一个对象信息按照对象大小和对象访问热度进行分类，得到至少一个对象组。第一获取单元 112 与分类单元 111 连接，第一获取单元 112 用于根据各节点中各存储介质处理各级大小的对象的读写速率，在各存储介质的容量以及占比小于 1 的约束条件下，以至少一个对象的读
20 写速率与访问热度乘积之和最大为目标，获取分类单元 111 分类得到的至少一个对象组在各存储介质上分配比例。第二获取单元 113 与第一获取单元 112 连接，第二获取单元 113 用于根据第一获取单元 112 处理后至少一个对象组在各存储介质上分配比例，获取至少一个对象被调整后的第二位置。优选地，第二获取单元 113 可以用于至少一个对象的第一位置以及第
25 一获取单元 112 处理后至少一个对象组在各存储介质上分配比例，按照位置就近的原则获取至少一个对象被调整后的第二位置。此时，发送模块 12

与第二获取单元 113 连接，发送模块 12 用于向至少一个节点发送第二获取单元 113 获取的至少一个对象的调整信息，以供至少一个节点根据至少一个对象的调整信息对至少一个对象进行迁移。

5 可选地，本实施例的迁移控制设备中还包括：收集模块 13。该收集模块 13 用于收集各节点中各个存储介质处理各级大小的对象的读写速率和存储容量。第一获取单元 112 与收集模块 13 连接，第一获取单元 112 用于根据收集模块 13 收集的各节点中各存储介质处理各级大小的对象的读写速率，在收集模块 13 收集的各存储介质的容量以及占比小于 1 的约束条件下，以至少一个对象的读写速率与访问热度乘积之和最大为目标，获取分类单元 111 分类得到的至少一个对象组在各存储介质上分配比例。

本实施例的迁移控制设备，通过采用上述模块实现对象调整的实现机制与上述相关方法实施例的实现机制相同，详细可以参考上述相关方法实施例的记载，在此不再赘述。

15 本实施例的迁移控制设备，通过采用上述模块能够实现各对象标识对应的对象在调整后的第二位置上使得各对象的读写速率与访问热度乘积之和最大，从而有效地保证了包括各节点的存储系统的存储性能达到最大。

图 6 为本发明一实施例提供的节点设备的结构示意图。如图 6 所示，本实施例的节点设备，包括：发送模块 20、接收模块 21 和调整模块 22。

20 其中发送模块 20 用于向迁移控制设备发送第一对象的第一对象信息，以供迁移控制设备根据预先存储的至少一个节点各节点中各个存储介质处理各级大小的对象的读写速率和存储容量、第一对象信息以及其他节点发送的第二对象信息，在满足各存储介质的占比小于 1 的条件下，获取第一对象的调整信息；该第一对象信息中包括第一对象标识、第一对象大小、第一对象的第一位置和第一对象的访问热度；对象的访问热度指的是
25 该对象在上一周期内被访问的次数；该调整信息中包括第一对象标识和第

一对象被调整后的第二位置，第一对象的第二位置和第二对象信息对应的对象被迁移控制设备调整后的位置能够保证第一对象在第二位置上和第二对象信息对应的对象在被调整后的位置上的读写速率与访问热度乘积之和最大；存储介质的占比指的是在上一周期内，该存储介质上存储的各对象在存储介质上的访问热度和对象的大小的乘积除以存储介质的最大吞吐率之后的累加。接收模块 21 用于接收迁移控制设备发送的第一对象的调整信息。调整模块 22 与接收模块 21 连接，调整模块 22 用于根据接收模块 21 接收的第一对象的调整信息对第一对象进行迁移。本实施例中节点设备可以称为当前节点，其他节点包括至少一个节点中当前节点之外的节点。

10 本实施例的节点设备，通过采用上述模块实现对象调整的实现机制与上述相关方法实施例的实现机制相同，详细可以参考上述相关方法实施例的记载，在此不再赘述。

本实施例的节点设备，通过采用本实施例的上述模块的技术方案，能够实现各对象标识对应的对象在调整后的第二位置上使得各对象的读写速率与访问热度乘积之和最大，从而有效地保证了包括各节点的存储系统的存储性能达到最大。

图 7 为本发明另一实施例提供的节点设备的结构示意图。如图 7 所示，本实施例的节点设备在上述图 6 所示实施例的基础上，还包括：处理模块 23。该处理模块 23 用于统计第一对象的访问热度；生成第一对象信息。此时发送模块 20 与处理模块 23 连接，发送模块 20 用于向迁移控制设备发送处理模块 23 生成的第一对象的第一对象信息。

可选地，本实施例的节点设备中的调整模块 22 还可以包括：获取单元 221 和处理单元 222。

其中获取单元 221 用于从接收模块 21 接收的第一对象的调整信息中获取第二位置。处理单元 222 与获取单元 221 连接，处理单元 222 用于判断获取单元 221 获取的第二位置为节点内的还是节点外的，当第二位置为节点内，将第一对象由第一位置迁移至第二位置；当第二位置为节点外时，

向第二位置所在的节点发送迁移对象请求，迁移对象请求中包括第二位置和第一对象，以供第二位置所在的节点将第一对象存储在第二位置处。

5 可选地，本实施例的节点设备中，接收模块 21 还用于当第二位置为节点外时，接收第二位置所在节点发送的确认消息；该确认消息中包括对象标识。处理模块 23 还与接收模块 21 连接，处理模块 23 用于根据接收模块 21 接收的确认消息，删除确认消息中的对象标识对应的对象。处理模块 23 还用于在删除确认消息中的对象标识对应的对象之后，更新对象标识对应的索引信息。

10 本实施例的节点设备，通过采用上述模块实现对象调整的实现机制与上述相关方法实施例的实现机制相同，详细可以参考上述相关方法实施例的记载，在此不再赘述。

本实施例的节点设备，通过采用上述模块能够实现各对象标识对应的对象在调整后的第二位置上使得各对象的读写速率与访问热度乘积之和最大，从而有效地保证了包括各节点的存储系统的存储性能达到最大。

15 图 8 为本发明实施例提供的存储系统的结构示意图。如图 8 所示，本实施例的存储系统包括：迁移控制设备 30 和至少一个节点设备 40。

其中至少一个节点设备 40 分别与迁移控制设备 30 连接，至少一个节点设备 40 用于分别向迁移控制设备 30 发送对象信息。

20 迁移控制设备 30 用于分别接收至少一个节点设备 40 发送的对象的对象信息，共接收到至少一个对象信息；该对象信息中包括对象标识、对象大小、对象的第一位置和对象的访问热度；对象的访问热度指的是对象在上一周期内被访问的次数。至少一个节点设备 40 中包括有至少一种存储介质。迁移控制设备 30 根据预先存储的至少一个节点设备 40 的各节点设备 40 中的各个存储介质处理各级大小的对象的读写速率和存储容量以及
25 接收的至少一个对象信息，在满足各个存储介质的占比小于 1 的条件下，获取至少一个对象的调整信息。该调整信息中包括对象标识和对象标识对应的对象被调整后的第二位置，至少一个对象的第二位置能够保证至少一

个对象在对应的第二位置上至少一个对象的读写速率与访问热度乘积之和最大；存储介质的占比指的是在上一周期内，该存储介质上存储的各对象在该存储介质上的访问热度和对象的大小的乘积除以该存储介质的最大吞吐率之后的累加；迁移控制设备 30 向至少一个节点设备 40 发送至少一个对象的调整信息。

至少一个节点设备 40 还用于接收迁移控制设备 30 发送的对象的调整信息；并根据对象的调整信息对对象进行迁移。

本实施例的存储系统中节点设备 40 可以采用上述图 6 或者 7 所示的节点设备。迁移控制设备 30 可以采用上述图 4 或者 5 所示实施例的迁移控制设备。本实施例的迁移控制设备实现对象调整的实现机制与上述相关实施例的实现机制相同，详细可以参考上述相关方法实施例的记载，在此不再赘述。

本实施例的存储系统，通过采用上述迁移控制设备和至少一个节点设备，能够实现各对象标识对应的对象在调整后的第二位置上使得各对象的读写速率与访问热度之和最大，从而有效地保证了包括各节点的存储系统的存储性能达到最大。

以上所描述的装置实施例仅仅是示意性的，其中作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到至少两个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性的劳动的情况下，即可以理解并实施。

最后应说明的是：以上实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

权 利 要 求

1、一种对象调整方法，其特征在于，包括：

接收至少一个节点发送的至少一个对象的对象信息，共接收到至少一个对象信息；所述对象信息中包括对象标识、对象大小、对象的第一位置
5 和对象的访问热度；所述对象的访问热度指的是所述对象在上一周期内被访问的次数；

根据预先存储的各所述节点中各个存储介质处理各级大小的对象的读写速率和存储容量以及所述至少一个对象信息，在满足各所述存储介质的占比小于 1 的条件下，获取所述至少一个对象的调整信息；所述调整信息
10 中包括所述对象标识和所述对象标识对应的对象被调整后的第二位置，所述至少一个对象的第二位置能够保证所述至少一个对象在对应的所述第二位置上的读写速率与所述访问热度乘积之和最大；所述存储介质的占比指的是在上一周期内，所述存储介质上存储的各对象在所述存储介质上的访问热度和所述对象的大小的乘积除以所述存储介质的最大吞吐率之
15 后的累加；

向所述至少一个节点发送所述至少一个对象的调整信息，以供所述至少一个节点根据所述至少一个对象的调整信息对所述至少一个对象进行迁移。

2、根据权利要求 1 所述的方法，其特征在于，根据预先存储的各所述节点中各个存储介质处理各级大小的对象的读写速率和存储容量以及
20 所述至少一个对象信息，在满足所述存储介质的占比小于 1 的条件下，获取所述至少一个对象的调整信息，包括：

根据各所述节点中各个存储介质处理各级大小的对象的读写速率和存储容量以及所述至少一个对象信息，在满足所述存储介质的占比小于 1
25 的条件下，获取所述至少一个对象被调整后的所述第二位置；

根据所述至少一个对象的所述第一位置和所述第二位置，获取所述至

少一个对象的调整信息。

3、根据权利要求 2 所述的方法，其特征在于，根据各所述节点中各个存储介质处理各级大小的对象的读写速率和存储容量以及所述至少一个对象信息，在满足所述存储介质的占比小于 1 的条件下，获取所述至少一个对象被调整后的所述第二位置，具体包括：

将所述至少一个对象信息按照对象大小和访问热度进行分类，得到至少一个对象组；

根据各所述节点中各所述存储介质处理各级大小的对象的读写速率，在各所述存储介质的容量以及所述占比小于 1 的约束条件下，以所述至少一个对象的读写速率与所述访问热度乘积之和最大为目标，获取所述至少一个对象组在各所述存储介质上分配比例；

根据所述至少一个对象组在各所述存储介质上分配比例，获取所述至少一个对象被调整后的所述第二位置。

4、根据权利要求 3 所述的方法，其特征在于，根据所述至少一个对象组在各存储介质上分配比例，获取所述至少一个对象被调整后的所述第二位置，具体包括：根据所述至少一个对象的第一位置以及所述至少一个对象组在各存储介质上分配比例，按照位置就近的原则获取所述至少一个对象被调整后的所述第二位置。

5、根据权利要求 1-4 任一所述的方法，其特征在于，还包括：

收集各所述节点中各个所述存储介质处理各级大小的对象的读写速率和存储容量。

6、一种对象调整方法，其特征在于，包括：

向迁移控制设备发送第一对象的第一对象信息，以供所述迁移控制设备根据预先存储的至少一个节点中各所述节点中的各个存储介质处理各级大小的对象的读写速率和存储容量、所述第一对象信息以及其他节点发送的第二对象信息，在满足各所述存储介质的占比小于 1 的条件下，获取

所述第一对象的调整信息；所述第一对象信息中包括第一对象标识、第一对象大小、第一对象的第一位置和第一对象的访问热度；所述对象的访问热度指的是所述对象在上一周期内被访问的次数；所述调整信息中包括所述第一对象标识和所述第一对象被调整后的第二位置，所述第一对象的所述第二位置和所述第二对象信息对应的对象被所述迁移控制设备调整后的位置能够保证所述第一对象在所述第二位置上和所述第二对象信息对应的对象在所述被调整后的位置上的读写速率与访问热度乘积之和最大；所述存储介质的占比指的是在上一周期内，所述存储介质上存储的各对象在所述存储介质上的访问热度和所述对象的大小的乘积除以所述存储介质的最大吞吐率之后的累加；所述其他节点包括所述至少一个节点中当前节点之外的节点；

接收所述迁移控制设备发送的所述第一对象的调整信息；

根据所述第一对象的调整信息对所述第一对象进行迁移。

7、根据权利要求6所述的方法，其特征在于，还包括：
15 统计所述第一对象的访问热度；生成所述第一对象信息。

8、根据权利要求6所述的方法，其特征在于，根据所述第一对象的调整信息对所述第一对象进行迁移，具体包括：

从所述第一对象的调整信息中获取所述第二位置；

20 判断所述第二位置为节点内的还是节点外的，当所述第二位置为节点内，将所述第一对象由所述第一位置迁移至所述第二位置；当所述第二位置为节点外时，向所述第二位置所在的节点发送迁移对象请求，所述迁移对象请求中包括所述第二位置和所述第一对象，以供所述第二位置所在的节点将所述第一对象存储在所述第二位置处。

9、根据权利要求8所述的方法，其特征在于，当所述第二位置为节点外时，还包括：

接收所述第二位置所在节点发送的确认消息；所述确认消息中包括所

述第一对象标识;

删除所述确认消息中的所述第一对象标识对应的第一对象。

10、根据权利要求 9 所述的方法,其特征在於,还包括:

删除所述确认消息中的所述第一对象标识对应的第一对象之后,更新
5 所述第一对象标识对应的索引信息。

11、一种迁移控制设备,其特征在於,包括:

接收模块,用于接收至少一个节点发送的至少一个对象的对象信息,
共接收到至少一个对象信息;所述对象信息中包括对象标识、对象大小、
对象的第一位置和对象的访问热度;所述对象的访问热度指的是所述对象
10 在上一周期内被访问的次数;

获取模块,用于根据预先存储的各所述节点中各个存储介质处理各级
大小的对象的读写速率和存储容量以及所述至少一个对象信息,在满足各
所述存储介质的占比小于 1 的条件下,获取所述至少一个对象的调整信息;
所述调整信息中包括所述对象标识和所述对象标识对应的对象被调整后
15 的第二位置,所述至少一个对象的第二位置能够保证所述至少一个对象在
对应的所述第二位置上的读写速率与所述访问热度乘积之和最大;所述存
储介质的占比指的是在上一周期内,所述存储介质上存储的各对象在所述
存储介质上的访问热度和所述对象的大小的乘积除以所述存储介质的最
大吞吐率之后的累加;

20 发送模块,用于向所述至少一个节点发送所述至少一个对象的调整信
息,以供所述至少一个节点根据所述至少一个对象的调整信息对所述至少
一个对象进行迁移。

12、根据权利要求 11 所述的设备,其特征在於:

所述获取模块,具体用于根据各所述节点中各个存储介质处理各级大
25 小的对象的读写速率和存储容量以及所述至少一个对象信息,在满足所述
存储介质的占比小于 1 的条件下,获取所述至少一个对象被调整后的所述

第二位置；根据所述至少一个对象的所述第一位置和所述第二位置，获取所述至少一个对象的调整信息。

13、根据权利要求 12 所述的设备，其特征在于，所述获取模块，包括：

5 分类单元，用于将所述至少一个对象信息按照对象大小和访问热度进行分类，得到至少一个对象组；

第一获取单元，用于根据各所述节点中各所述存储介质处理各级大小的对象的读写速率，在各所述存储介质的容量以及所述占比小于 1 的约束条件下，以所述至少一个对象的读写速率与所述访问热度乘积之和最大为
10 目标，获取所述至少一个对象组在各所述存储介质上分配比例；

第二获取单元，用于根据所述至少一个对象组在各所述存储介质上分配比例，获取所述至少一个对象被调整后的所述第二位置。

14、根据权利要求 13 所述的设备，其特征在于，还包括：所述第二获取单元，具体用于根据所述至少一个对象的第一位置以及所述至少一个
15 对象组在各存储介质上分配比例，按照位置就近的原则获取所述至少一个对象被调整后的所述第二位置。

15、根据权利要求 11-14 任一所述的设备，其特征在于，还包括：

收集模块，用于收集各所述节点中各个所述存储介质处理各级大小的对象的读写速率和存储容量。

20 16、一种节点设备，其特征在于，包括：

发送模块，用于向迁移控制设备发送第一对象的第一对象信息，以供所述迁移控制设备根据预先存储的至少一个节点中各所述节点中的各个存储介质处理各级大小的对象的读写速率和存储容量、所述第一对象信息以及其他节点发送的第二对象信息，在满足各所述存储介质的占比小于 1
25 的条件下，获取所述第一对象的调整信息；所述第一对象信息中包括第一对象标识、第一对象大小、第一对象的第一位置和第一对象的访问热度；

所述对象的访问热度指的是所述对象在上一周期内被访问的次数；所述调整信息中包括所述第一对象标识和所述第一对象被调整后的第二位置，所述第一对象的所述第二位置和所述第二对象信息对应的对象被所述迁移控制设备调整后的位置能够保证所述第一对象在所述第二位置上和所述

5 第二对象信息对应的对象在所述被调整后的位置上的读写速率与访问热度乘积之和最大；所述存储介质的占比指的是在上一周期内，所述存储介质上存储的各对象在所述存储介质上的访问热度和所述对象的大小的乘积除以所述存储介质的最大吞吐率之后的累加；所述其他节点包括所述至少一个节点中当前节点之外的节点；

10 接收模块，用于接收所述迁移控制设备发送的所述第一对象的调整信息；

调整模块，用于根据所述第一对象的调整信息对所述第一对象进行迁移。

17、根据权利要求 16 所述的设备，其特征在于，还包括：

处理模块，用于统计所述第一对象的访问热度；生成所述第一对象信息。

15 18、根据权利要求 17 所述的设备，其特征在于，所述调整模块，包括：

获取单元，用于从所述第一对象的调整信息中获取所述第二位置；

处理单元，用于判断所述第二位置为节点内的还是节点外的，当所述第二位置为节点内，将所述第一对象由所述第一位置迁移至所述第二位置；当所述第二位置为节点外时，向所述第二位置所在的节点发送迁移对象请求，所述迁移对象请求中包括所述第二位置和所述第一对象，以供所述

20 第二位置所在的节点将所述第一对象存储在所述第二位置处。

19、根据权利要求 18 所述的设备，其特征在于：

所述接收模块，用于当所述第二位置为节点外时，接收所述第二位置所在节点发送的确认消息；所述确认消息中包括所述对象标识；

25

所述处理模块，还用于删除所述确认消息中的所述对象标识对应的对

象。

20、根据权利要求 19 所述的设备，其特征在于：

所述处理模块，还用于删除所述确认消息中的所述第一对象标识对应的第一对象之后，更新所述第一对象标识对应的索引信息。

5 21、一种存储系统，其特征在于，包括迁移控制设备和至少一个节点设备；

所述至少一个节点设备，用于分别向迁移控制设备发送对象信息；

所述迁移控制设备，用于分别接收所述至少一个节点设备发送的对象信息，共接收到至少一个对象信息；各所述对象信息中包括对象标识、对象大小、对象的第一位置和对象的访问热度；所述对象的访问热度指的是所述对象在上一周期内被访问的次数；根据预先存储的所述至少一个节点设备的各所述节点设备中各个存储介质处理各级大小的对象的读写速率和存储容量以及所述至少一个对象信息，在满足各所述存储介质的占比小于 1 的条件下，获取所述至少一个对象的调整信息；所述调整信息中包括
10 所述对象标识和所述对象标识对应的对象被调整后的第二位置，所述至少一个对象的第二位置能够保证所述至少一个对象在对应的所述第二位置上的读写速率与所述访问热度乘积之和最大；所述存储介质的占比指的是在上一周期内，所述存储介质上存储的各对象在所述存储介质上的访问热度和所述对象的大小的乘积除以所述存储介质的最大吞吐率之后的累加；
15 向所述至少一个节点设备发送所述至少一个对象的调整信息；

所述至少一个节点设备，还用于接收所述迁移控制设备发送的所述对象的调整信息；根据所述对象的调整信息对所述对象进行迁移。

1/4

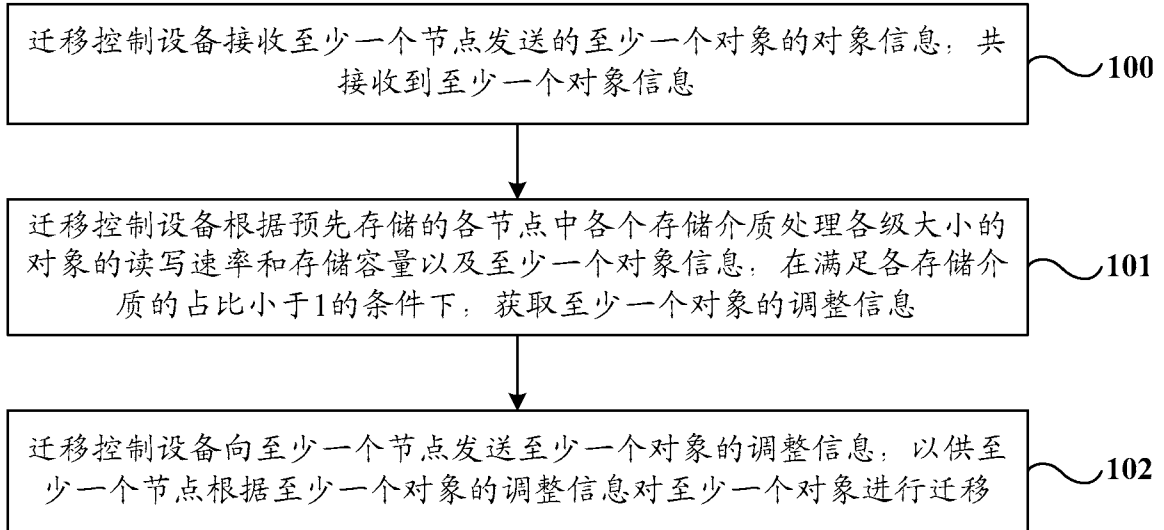


图 1

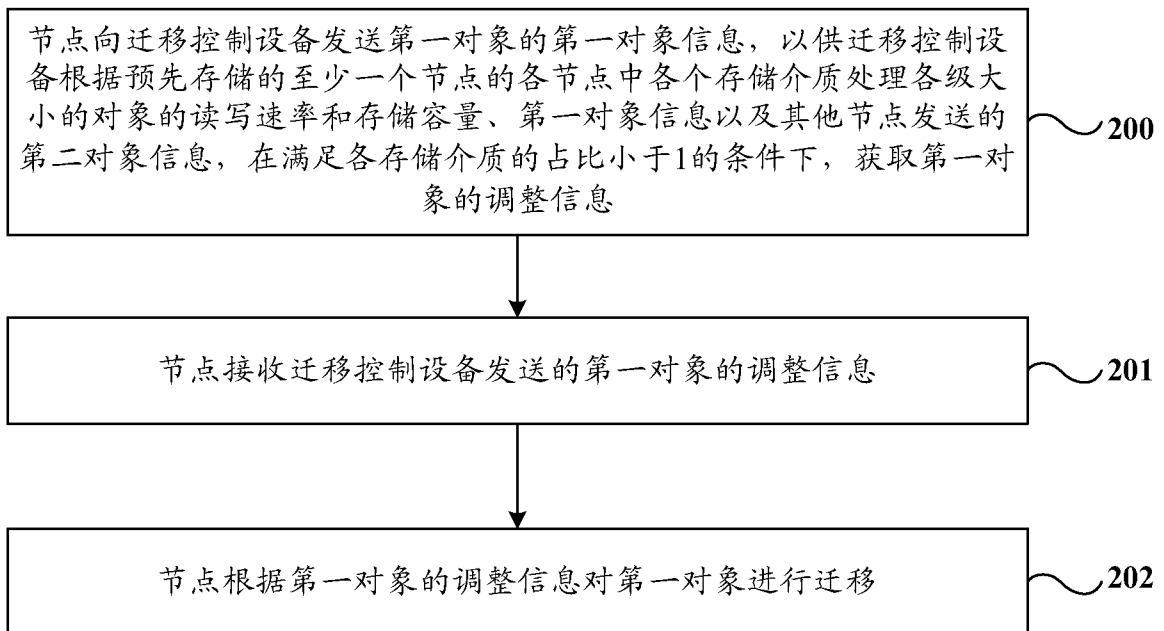


图 2

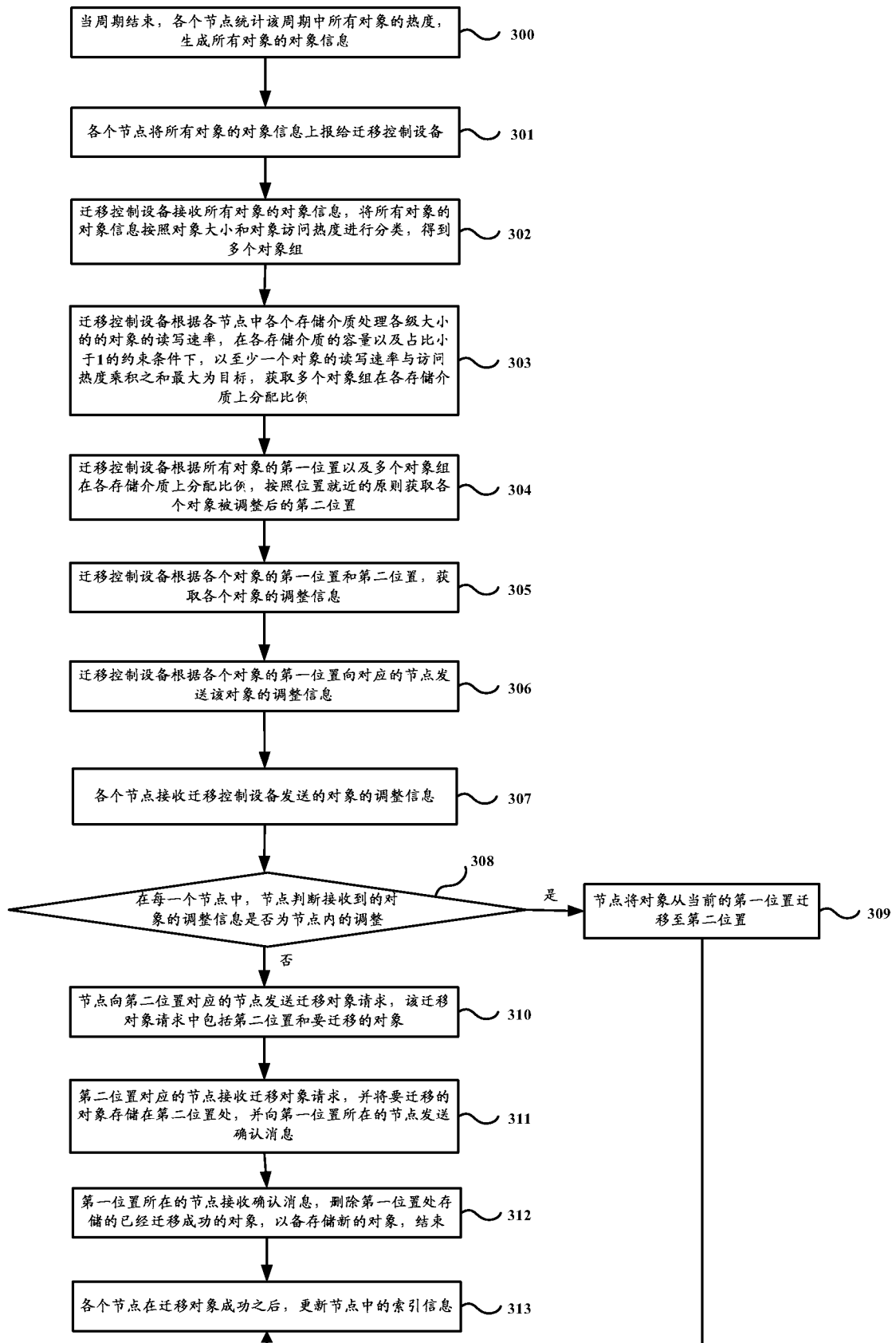


图 3

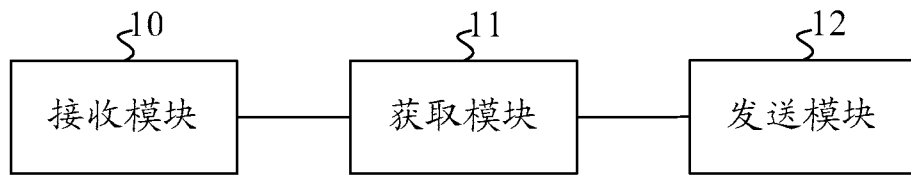


图 4

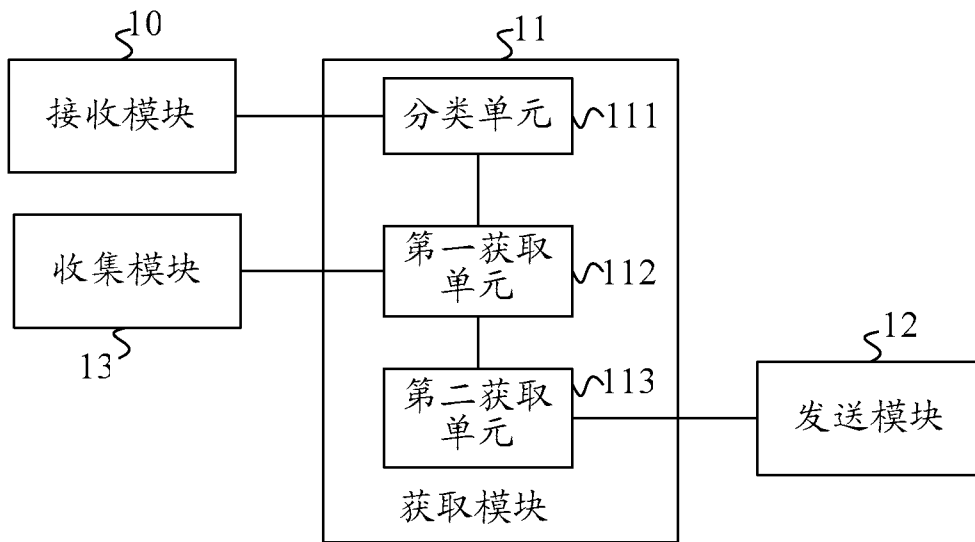


图 5

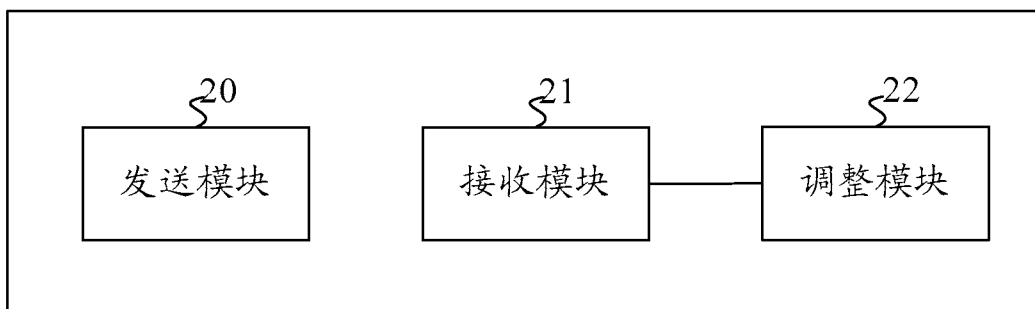


图 6

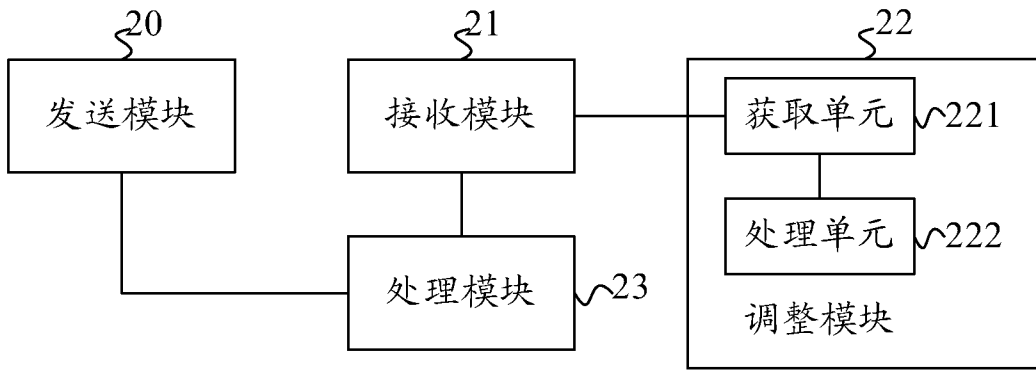


图 7

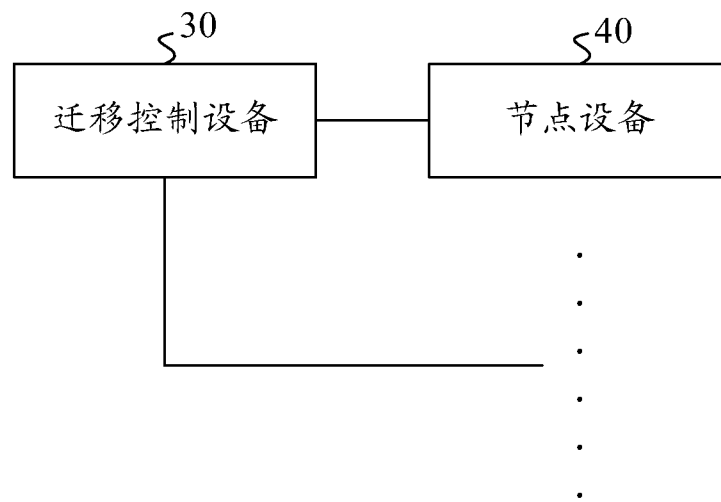


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2011/077564

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (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 17/-, H04L 29/-, H04L 12/-

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)

CPRSABS, CNKI, DWPI: node, heat, popular, network, note, media, object, video, popularity, hotspot, hot, cache, storage, memory, migrate, regulate

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 101184021 A (HUAWEI TECHNOLOGIES CO., LTD.), 21 May 2008 (21.05.2008), the whole document	1-21
A	CN 101552799 A (HUAWEI TECHNOLOGIES CO., LTD.), 07 October 2009 (07.10.2009), the whole document	1-21
A	CN 1758604 A (ZTE CORP.), 12 April 2006 (12.04.2006), the whole document	1-21
A	US 2004172501 A1 (HITACHI LTD.), 02 September 2004 (02.09.2004), the whole document	1-21
A	US 2004186861 A1 (PHATAK, S.H.), 23 September 2004 (23.09.2004), the whole document	1-21

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>
---	---

Date of the actual completion of the international search
27 March 2012 (27.03.2012)

Date of mailing of the international search report
19 April 2012 (19.04.2012)

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
QU, Fengli
Telephone No.: (86-10) 62411737

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2011/077564

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101184021 A	21.05.2008	CN 101184021 B	02.06.2010
CN 101552799 A	07.10.2009	None	
CN 1758604 A	12.04.2006	CN 100336343 C	05.09.2007
US 2004172501 A1	02.09.2004	JP 2004265110 A	24.09.2004
US 2004186861 A1	23.09.2004	WO 2004068469 A2	12.08.2004
		EP 1584036 A2	12.10.2005
		AU 2004207357 A1	12.08.2004
		JP 2006516341 A	29.06.2006
		CN 1754155 A	29.03.2006
		US 7103617 B2	05.09.2006
		US 2007198685 A1	23.08.2007

A. 主题的分类		
G06F17/30(2006.01)i		
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: G06F17/-,H04L29/-,H04L12/-		
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))		
CPRSABS,CNKI,DWPI: 网络, 节点, 媒体, 对象, 视频, 热度, 热点, 热门, 流行, 欢迎, 存储, 缓存, 调整, 迁移 network, note, media, object, video, popularity, hotspot, hot, cache, storage, memory, migrate, regulate		
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
A	CN101184021A(华为技术有限公司)21.5 月 2008(21.05.2008)全文	1-21
A	CN101552799A(华为技术有限公司)07.10 月 2009(07.10.2009)全文	1-21
A	CN1758604A(中兴通讯股份有限公司)12.4 月 2006(12.04.2006) 全文	1-21
A	US2004172501A1(HITACHI LTD)02.9 月 2004(02.09.2004) 全文	1-21
A	US2004186861A1(PHATAK S H) 23.9 月 2004(23.09.2004)全文	1-21
<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 “E” 在国际申请日的当天或之后公布的在先申请或专利 “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) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “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		“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 “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 “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 “&” 同族专利的文件
国际检索实际完成的日期 27.3 月 2012(27.03.2012)		国际检索报告邮寄日期 19.4 月 2012 (19.04.2012)
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		授权官员 曲凤丽 电话号码: (86-10) 62411737

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2011/077564

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN101184021A	21.05.2008	CN101184021B	02.06.2010
CN101552799A	07.10.2009	无	
CN1758604A	12.04.2006	CN100336343C	05.09.2007
US2004172501A1	02.09.2004	JP2004265110A	24.09.2004
US2004186861A1	23.09.2004	WO2004068469A2	12.08.2004
		EP1584036A2	12.10.2005
		AU2004207357A1	12.08.2004
		JP2006516341A	29.06.2006
		CN1754155A	29.03.2006
		US7103617B2	05.09.2006
		US2007198685A1	23.08.2007