

(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
2014 年 11 月 13 日 (13.11.2014)

WIPO | PCT

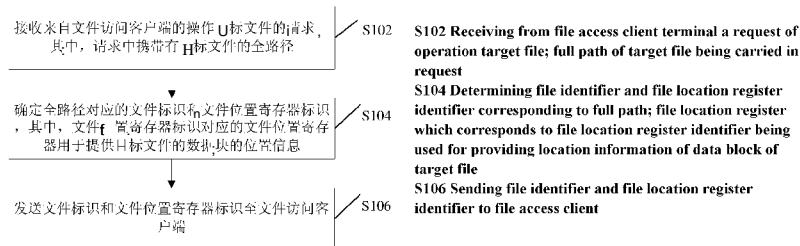
(10) 国际公布号
WO 2014/180232 A 1

- (51) 国际分类号 : G06F 17/30 (2006 .01) H04L 29/06 (2006.01)
- (21) 国际申请号 : PCT/CN2014/075522
- (22) 国际申请日 : 2014 年 4 月 16 日 (16.04.2014)
- (25) 申报语言 : 中文
- (26) 公布语言 : 中文
- (30) 优先权 : 2013 10601094.3 2013 年 11 月 25 日 (25.11.2013) CN
- (71) 申请人 :中兴通讯股份有限公司 (ZTE CORPORATION) [CN/CN]; 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦 ,Guangdong 5 18057 (CN)。
- (72) 发明人 :洪建峰 (HONG, Jianfeng); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦 ,Guangdong 5 18057 (CN)。雷爱民 (LEI, Aimin); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦 ,Guangdong 5 18057 (CN)。王俊 (WANG, Jun); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦 ,Guangdong 5 18057 (CN)。
- (74) 代理人 :北京康信知识产权代理有限公司 (KANGXIN PARTNERS, P.C.); 中国北京市海淀区知春路甲 48 号盈都大厦 A 座 16 层 ,Beijing 100098 (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, BN, 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, IR, IS, JP, KE, KG, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, 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, KM, 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))。

[见续页]

- (54) Title: METHOD AND DEVICE FOR RESPONDING TO A REQUEST, AND DISTRIBUTED FILE SYSTEM
- (54) 发明名称 :请求响应方法、装置及分布式文件系统



(57) Abstract: The present invention discloses a method and device for responding to a request, and a distributed file system, the method comprising: receiving from a file access client terminal a request of an operation target file; the full path of the target file being carried in the request; determining the file identifier and file location register identifier corresponding to the full path; the file location register which corresponds to the file location register identifier being used for providing location information of a data block of the target file; sending the file identifier and the file location register identifier to the file access client. The present invention solves the problem of a distributed file system making high requirements on file location register hardware, reduces the requirements on file location register hardware, and improves system stability.

(57) 摘要: 本发明公开了一种请求响应方法、装置及分布式文件系统, 其中, 该方法包括: 接收来自文件访问客户端的操作目标文件的请求, 其中, 该请求中携带有目标文件的全路径; 确定全路径对应的文件标识和文件位置寄存器标识, 其中, 文件位置寄存器标识对应的文件位置寄存器用于提供目标文件的数据块的位置信息; 发送文件标识和文件位置寄存器标识至文件访问客户端。通过本发明, 解决了分布式文件系统对文件位置寄存器硬件要求高的问题, 降低了对文件位置寄存器硬件的要求, 提升了系统的稳定性。



2 14/180232 A1



-
- 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。

请求响应方法、装置及分布式文件系统

技术领域

本发明涉及通信领域，具体而言，涉及一种请求响应方法、装置及分布式文件系统。

5 背景技术

在相关技术中的分布式文件系统（Distributed File System，简称为DFS）架构中，只有文件位置寄存器、文件访问客户端和文件访问服务器，其中：

文件位置寄存器：负责文件的数据块分布和管理，其功能包括：块分布、垃圾收集和清理、负载均衡、数据库复制和修复。

10 文件访问客户端：负责应用程序访问文件系统的代理，提供给应用程序通用的文件操作接口，如：打开（Open）、关闭（Close）、读（Read）、写（Write）、查找（Seek）等；

文件访问服务器：是文件系统中实际存储用户数据的地方。

15 在上述的系统架构中，文件访问客户端和文件访问服务器是可以扩充的，都可以有多个。然而，用于存储整个分布式文件系统的数据块信息的文件位置寄存器往往只有一个，所有的文件访问客户端都通过这一个文件位置寄存器请求所要操作的文件对应的数据块信息。

20 因此，相关技术中的DFS系统的容量和性能受到文件位置寄存器性能的影响，对文件位置寄存器服务器硬件的要求也非常高。然而，在硬件很难满足需求的情况下必须对文件位置存储器做技术上的创新。

针对相关技术中分布式文件系统对文件位置寄存器硬件要求高的问题，目前尚未提出有效的解决方案。

发明内容

25 本发明提供了一种请求响应方法、装置及分布式文件系统，以至少解决相关技术中分布式文件系统对文件位置寄存器硬件要求高的问题。

根据本发明的一个方面，提供了一种请求响应方法，包括：接收来自文件访问客户端的操作目标文件的请求，其中，所述请求中携带有所述目标文件的全路径；确定所述全路径对应的文件标识和文件位置寄存器标识，其中，所述文件位置寄存器标识对应的文件位置寄存器用于提供所述目标文件的数据块的位置信息；发送所述文件标识和所述文件位置寄存器标识至所述文件访问客户端。

优选地，在所述请求为写入所述目标文件的请求的情况下，确定所述全路径对应的所述文件标识和所述文件位置寄存器标识包括：根据所述全路径，生成所述目标文件的文件标识；根据预定策略，确定与所述文件标识对应的文件位置寄存器标识。

10 优选地，在确定所述全路径对应的所述文件标识和所述文件位置寄存器标识之后，所述方法还包括：保存所述全路径与所述文件标识、所述文件位置寄存器标识的关联关系。

15 优选地，在发送所述文件标识和所述文件位置寄存器标识至所述文件访问客户端之后，所述方法还包括：所述文件访问客户端获取所述位置信息，其中，所述位置信息是所述文件位置寄存器根据所述文件标识生成并发送给所述文件访问客户端的；所述文件访问客户端将所述目标文件写入所述位置信息所对应的磁盘位置。

优选地，在所述请求为读取或者删除所述目标文件的请求的情况下，确定所述全路径对应的所述文件标识和所述文件位置寄存器标识包括：查找所述全路径对应的所述文件标识和所述文件位置寄存器标识。

20 优选地，在发送所述文件标识和所述文件位置寄存器标识至所述文件访问客户端之后，所述方法还包括：所述文件访问客户端获取所述位置信息，其中，所述位置信息是所述文件位置寄存器根据所述文件标识从保存的文件记录中查找并发送给所述文件访问客户端的；所述文件访问客户端读取或者删除所述位置信息所对应的磁盘位置中保存的所述目标文件。

25 根据本发明的另一个方面，还提供了一种请求响应装置，包括：接收模块，设置为接收来自文件访问客户端的操作目标文件的请求，其中，所述请求中携带有所述目标文件的全路径；确定模块，设置为确定所述全路径对应的文件标识和文件位置寄存器标识，其中，所述文件位置寄存器标识对应的文件位置寄存器用于提供所述目标文件的数据块的位置信息；发送模块，设置为发送所述文件标识和所述文件位置寄存器标识至所述文件访问客户端。

优选地，所述确定模块包括：生成单元，设置为根据所述全路径，生成所述目标文件的文件标识；确定单元，设置为根据预定策略，确定与所述文件标识对应的文件位置寄存器标识。

5 优选地，所述装置还包括：保存模块，设置为保存所述全路径与所述文件标识、所述文件位置寄存器标识的关联关系。

优选地，所述确定模块包括：查找单元，设置为查找所述全路径对应的所述文件标识和所述文件位置寄存器标识。

10 根据本发明的另一个方面，还提供了一种分布式文件系统，包括文件访问客户端、文件位置寄存器和文件访问服务器，其中，所述文件位置寄存器的数目为多个，所述系统还包括：目录树元数据存储服务器，设置为根据来自所述文件访问客户端的操作目标文件的请求中携带的目标文件的全路径，确定所述全路径对应的文件标识和文件位置寄存器标识，并发送所述文件标识和所述文件位置寄存器标识至所述文件访问客户端，其中，多个文件位置寄存器中所述文件位置寄存器标识对应的文件位置寄存器用于提供所述目标文件的数据块的位置信息。

15 通过本发明，采用接收来自文件访问客户端的操作目标文件的请求，其中，该请求中携带有目标文件的全路径；确定全路径对应的文件标识和文件位置寄存器标识，其中，文件位置寄存器标识对应的文件位置寄存器用于提供目标文件的数据块的位置信息；发送文件标识和文件位置寄存器标识至文件访问客户端的方式，解决了分布式文件系统对文件位置寄存器硬件要求高的问题，降低了对文件位置寄存器硬件的要求，
20 提升了系统的稳定性。

附图说明

此处所说明的附图用来提供对本发明的进一步理解，构成本申请的一部分，本发明的示意性实施例及其说明用于解释本发明，并不构成对本发明的不当限定。在附图中-

25 图 1 是根据本发明实施例的请求响应方法的流程示意图；

图 2 是根据本发明实施例的请求响应装置的结构示意图；

图 3 是根据本发明实施例的请求响应装置的优选结构示意图一；

图 4 是根据本发明实施例的请求响应装置的优选结构示意图二；

图 5 是根据本发明实施例的分布式文件系统的结构示意图；

图 6 是根据本发明优选实施例的分布式文件系统的系统构架示意图；

图 7a~图 7c 是根据本发明优选实施例的用户操作文件的流程示意图；

图 8 是根据本发明优选实施例的写文件过程的流程示意图；

5 图 9 是根据本发明优选实施例的读文件过程的流程示意图；

图 10 是根据本发明优选实施例的删除文件过程的流程示意图。

具体实施方式

需要说明的是，在不冲突的情况下，本申请中的实施例及实施例中的特征可以相互组合。下面将参考附图并结合实施例来详细说明本发明。

10 在附图的流程图示出的步骤可以在诸如一组计算机可执行指令的计算机系统中执行，并且，虽然在流程图中示出了逻辑顺序，但是在某些情况下，可以以不同于此处的顺序执行所示出或描述的步骤。

本实施例提供了一种请求响应方法，图 1 是根据本发明实施例的请求响应方法的流程示意图，如图 1 所示，该流程包括：

15 步骤 S102，接收来自文件访问客户端的操作目标文件的请求，其中，请求中携带有目标文件的全路径；

步骤 S104，确定全路径对应的文件标识和文件位置寄存器标识，其中，文件位置寄存器标识对应的文件位置寄存器用于提供目标文件的数据块的位置信息；

步骤 S106，发送文件标识和文件位置寄存器标识至文件访问客户端。

20 通过上述步骤所提供的请求响应方法，在文件访问客户端请求操作目标文件的情况下，通过请求中携带的全路径确定目标文件对应的文件标识和文件位置寄存器标识，并响应给文件访问客户端。由此可见，在上述方法所应用的系统中，可以包括多个文件位置寄存器，这些文件位置寄存器通过文件位置寄存器标识进行区别，而文件访问客户端通过上述响应过程获得文件标识和文件位置寄存器标识后，就可以通过对应的
25 文件位置寄存器对文件进行处理，即通过上述的请求响应方法，对文件访问客户端的请求根据全路径进行负载均衡处理，使得可以支持多个文件位置寄存器来共同实现相

关技术中仅有的一个文件位置寄存器的功能，从而降低了对文件位置寄存器的硬件要求，解决了相关技术中分布式文件系统对文件位置寄存器硬件要求高的问题，提高了系统的稳定性。

需要说明的是，上述的目标文件的全路径、文件标识是用于在海量的文件中唯一标识目标文件的信息，文件位置寄存器标识是用于在多个文件位置寄存器中唯一标识文件位置寄存器的信息。其中，全路径也可以作为唯一标识目标文件的文件标识。优选地，还可以通过相关技术中的特定算法将全路径唯一的映射为文件标识。

优选地，上述的操作目标文件的请求包括：对目标文件的写入、读取、查找、删除等操作的请求。

10 优选地，在请求为写入目标文件的请求的情况下，在确定全路径对应的文件标识和文件位置寄存器标识时，可以根据目标文件的全路径，采用预定的算法生成目标文件的文件标识，或者直接为目标文件分配文件标识；根据预定策略，确定与文件标识对应的文件位置寄存器标识，例如，通过对系统中各个文件位置寄存器负载的统计情况，确定在统计区间中负载最小的文件位置寄存器的标识为与文件标识对应的文件位置寄存器标识。

15 优选地，在确定全路径对应的文件标识和文件位置寄存器标识之后，还包括：保存全路径与文件标识、文件位置寄存器标识的关联关系。其中，保存的关联关系用于在请求读取或者删除目标文件的情况下，根据文件访问客户端发送的请求中携带的全路径查找对应的文件标识和文件位置寄存器标识。

20 优选地，在文件访问客户端接收到文件标识和文件位置寄存器标识之后，文件访问客户端可以根据携带有文件标识的请求，向对应的文件位置寄存器请求并获取目标文件的数据块的位置信息，其中，该位置信息是文件位置寄存器根据文件标识生成并发送给文件访问客户端的；文件访问客户端接收到目标文件的位置信息之后，将目标文件写入位置信息所对应的磁盘位置，其中，该磁盘位置位于相应的文件访问服务器上。

25 优选地，在请求为读取或者删除目标文件的请求的情况下，通过在保存的关联关系中查找全路径对应的文件标识和文件位置寄存器标识来确定该目标文件对应的文件标识和文件位置寄存器标识。

30 优选地，在文件访问客户端接收到文件标识和文件位置寄存器标识之后，文件访问客户端可以根据携带有文件标识的请求，向对应的文件位置寄存器请求并获取目标

文件的数据块的位置信息，其中，该位置信息是文件位置寄存器根据文件标识从保存的文件记录中查找并发送给文件访问客户端的；文件访问客户端接收到目标文件的位置信息后，读取或者删除位置信息所对应的磁盘位置中保存的目标文件，其中，该磁盘位置位于相应的文件访问服务器上。

5 本实施例还提供了一种请求响应装置，该装置用于实现上述请求响应方法，该装置的功能实现已经在上述方法实施例中进行了说明，在此不再赘述。

图 2 是根据本发明实施例的请求响应装置的结构示意图，如图 2 所示，该装置包括：接收模块 22、确定模块 24 和发送模块 26，其中，接收模块 22，设置为接收来自文件访问客户端的操作目标文件的请求，其中，请求中携带有目标文件的全路径；确定模块 24 耦合至接收模块 22，设置为确定全路径对应的文件标识和文件位置寄存器标识，其中，文件位置寄存器标识对应的文件位置寄存器用于提供目标文件的数据块的位置信息；发送模块 26 耦合至确定模块 24，设置为发送文件标识和文件位置寄存器标识至文件访问客户端。

15 本发明的实施例中所涉及到的模块、单元可以通过软件的方式实现，也可以通过硬件的方式来实现。本实施例中的所描述的模块、单元也可以设置在处理器中，例如，可以描述为：一种处理器包括接收模块 22、确定模块 24 和发送模块 26。其中，这些模块的名称在某种情况下并不构成对该模块本身的限定，例如，接收模块还可以被描述为“用于接收来自文件访问客户端的操作目标文件的请求的模块”。

图 3 是根据本发明实施例的请求响应装置的优选结构示意图一，如图 3 所示，优选地，确定模块 24 包括：生成单元 242 耦合至接收模块 22，设置为根据全路径，生成目标文件的文件标识；确定单元 244 耦合至生成单元 242，设置为根据预定策略，确定与文件标识对应的文件位置寄存器标识。

图 4 是根据本发明实施例的请求响应装置的优选结构示意图二，如图 4 所示，优选地，该装置还包括：保存模块 42，耦合至确定模块 24，设置为保存全路径与文件标识、文件位置寄存器标识的关联关系。

25 优选地，确定模块 24 还包括：查找单元 246，设置为查找全路径对应的文件标识和文件位置寄存器标识。

本实施例还提供了一种分布式文件系统，图 5 是根据本发明实施例的分布式文件系统的结构示意图，如图 5 所示，该系统包括：文件访问客户端 52、文件位置寄存器 30 54 和文件访问服务器 56，其中，文件位置寄存器 54 的数目为多个，该系统还包括：

目录树元数据存储服务器 58，设置为根据来自文件访问客户端 52 的操作目标文件的请求中携带的目标文件的全路径，确定全路径对应的文件标识和文件位置寄存器标识，并发送文件标识和文件位置寄存器标识至文件访问客户端 52，其中，多个文件位置寄存器 54 中文件位置寄存器标识对应的文件位置寄存器 54 用于提供目标文件的数据块的位置信息。

其中，文件的元数据分为两部分进行存储，文件命名空间信息存储在目录树元数据存储服务器上，文件块信息（数据块信息）存储在文件位置寄存器上；文件位置寄存器个数可以在线扩展，从而可以最大限度的提高系统的容量；写文件时通过目录树元数据存储服务器分配文件位置寄存器，并将文件的路径和生成的 FILEID 关联，存储在目录树服务器（目录树元数据存储服务器）上，后面操作文件时可以通过目录树服务器定位到文件位置寄存器，在文件位置寄存器上对文件进行操作；文件块操作在文件位置寄存器上进行，通过多个文件位置寄存器对归属自己的数据库进行操作，提高系统的性能。

需要说明的是：上述的分布式文件系统目录树元数据存储服务器 58 的功能相当于上述请求响应装置，可以结合上述请求相应装置和请求响应方法进行结合说明和描述，在此将不再赘述。

下面结合优选实施例进行描述和说明。

本优选实施例提供了一种基于 DFS 通过多个文件位置寄存器实现海量文件存储的实现方法，涉及存储领域，尤其涉及利用分布式文件系统海量数据存储。

在本优选实施例中提供了一种当单个服务器无法满足存储规模时，通过多个文件位置寄存器来满足海量存储的方案。

为了实现上述目的，本优选实施例提供的方案包括：

新增目录树元数据存储服务器，目录树元数据存储服务器的功能包括两部分，一部分是负责数据配置和管理；另一个部分是负责管理文件的命名空间。

数据配置和管理部分用于配置文件位置寄存器、文件访问服务器和文件访问客户端的个数，服务器的地址等，并对这些服务器进行维护。

管理文件的命名空间负责根据文件路径将文件对应到某个文件位置寄存器，到对应的文件位置寄存器上去查找数据库的存储位置。

图 6 是根据本发明优选实施例的分布式文件系统的系统构架示意图，在图 6 中示出了包括多个文件位置寄存器的系统架构。

如图 6 所示，当用户需要读写文件的时候，需要把文件全路径发送到目录树服务器（相当于上述目录树元数据存储服务器），如果是写文件，目录树服务器需要存储文件命名空间，然后分配可用的文件位置寄存器，然后通过文件访问客户端将目录树服务器上生成的文件标识（ID）发送到目录树服务器分配的文件位置寄存器上，文件位置寄存器存储文件 ID（简称为 FLRID），并分配写文件的数据块位置信息，返回到文件访问客户端，将文件信息写入磁盘。

如果是读文件，目录树服务器通过文件全路径到命名空间中定位，找到写文件时分配的文件 ID 和文件位置寄存器 ID，返回给文件访问客户端，文件访问客户端通过文件 ID 到对应的文件位置寄存器上去查找数据库所在的磁盘信息，文件位置寄存器将相应信息返回给文件访问客户端，文件访问客户端到磁盘中读取文件。

下面通过附图举例说明多文件位置寄存器文件系统中文件的写流程。

图 7a~图 7c 是根据本发明优选实施例的用户操作文件的流程示意图，该流程包括如下步骤：

如图 7a 所示，用户将操作文件请求发送到文件访问客户端，文件访问客户端将文件名全路径发送到目录树服务器，目录树服务器通过查找命名空间，取得数据表中存储的文件信息，将文件信息中的文件标识（FILEID）和文件位置寄存器 ID 返回给文件访问客户端；

如图 7b 所示，文件访问客户端通过目录树服务器返回的 FILEID 和文件位置寄存器 ID，将操作文件的请求发送到文件位置寄存器 ID 对应的文件位置寄存器上，文件位置寄存器通过 FILEID 找到 FILEID 对应的记录，并通过 FILEID 对应的记录关联到文件存储的数据块信息，将数据块信息返回给文件访问客户端；

如图 7c 所示，文件访问客户端通过文件位置寄存器返回的数据块信息，将数据库信息中数据存储的磁盘位置，数据块大小等信息发送到文件访问服务器，通过文件访问服务器操作数据块信息，文件访问服务器操作完成后，给文件访问客户端响应；以及，文件访问客户端收到文件访问服务器的响应，本次操作文件流程完成，给用户回复操作文件的应答。

可见,在本优选实施例中提出了一种新的分布式文件系统中元数据管理实现方案。通过本优选实施例,可以通过多个文件位置寄存器来管理文件的数据块分布和数据块管理,从而达到大容量和高性能的需求。而由于在现有的分布式文件系统架构中,难以实现大容量和高性能的需求,因此,本发明优选实施例提出了一种新的适用于分布式文件系统的提高性能和容量的实现方法。

下面将分别通过写文件、读文件和删除文件这三个通用的文件处理操作,对本发明优选实施例提供的系统以及系统的工作流程进行说明。

优选实施例一

写文件过程

10 图 8 是根据本发明优选实施例的写文件过程的流程示意图,如图 8 所示,该流程包括如下步骤:

步骤 S801, 用户通过文件访问客户端写文件,写文件请求发送到文件访问客户端线程,文件访问客户端将要写的文件对象(称为目标文件)的全路径发送给目录树服务器,请求写文件;

15 步骤 S802, 目录树服务器先在命名空间中查找,查看待写文件是否已经存在。如果已存在则给文件访问客户端回复文件已存在的失败码;

步骤 S803, 如果不存在则生成写的文件 ID, 分配空闲的文件位置寄存器 ID 等信息,生成字典表记录,存储文件名。生成文件的 FILEID 对应记录,存储 FILEID、文件位置寄存器 ID 等信息,并给文件访问客户端返回创建文件成功;

20 步骤 S804, 文件访问客户端收到消息之后给对应的文件位置寄存器发送创建文件消息,文件位置寄存器通过 FILEID 定位,如果定位成功,给文件访问客户端回复文件已存在。如果定位失败则通过创建 FILE 记录,存储 FILEID, 创建时间等信息,给文件访问客户端回复创建文件成功应答;

25 步骤 S805, 文件访问客户端收到创建文件应答,通过 FILEID 给文件位置寄存器发送创建数据块请求,文件位置寄存器根据存储规则选择写数据块的目的磁盘,并生成数据块对应记录,存储数据块所在磁盘信息,给文件访问客户端回复创建数据块所在磁盘信息;

步骤 S806，文件访问客户端收到信息，根据文件位置寄存器返回的磁盘信息到文件访问服务器创建数据块，并写入文件内容；

步骤 S807，文件访问服务器写入之后，给文件访问客户端回复写入结果，数据块大小信息，文件访问客户端上报到文件位置寄存器，要求文件位置寄存器将数据块大小记录到数据块记录中，文件位置寄存器记录完成之后给文件访问客户端回复；

步骤 S808，文件访问客户端收到回复之后，写文件完成，给用户写文件完成应答。

优选实施例二

读文件过程

图 9 是根据本发明优选实施例的读文件过程的流程示意图，如图 9 所示，该流程包括如下步骤：

步骤 S901，用户通过文件访问客户端读文件，读文件请求发送到文件访问客户端线程，文件访问客户端将要读的文件对象（即目标文件）全路径发送给目录树服务器，请求读文件；

步骤 S902，目录树服务器先在命名空间中查找，查看待读文件是否已经存在。如果不存在则给文件访问客户端回复文件不存在的失败码；

步骤 S903，如果存在找到全路径对应的文件记录，将文件的 FILEID 和文件位置寄存器 ID 返回给文件访问客户端；

步骤 S904，文件访问客户端收到消息之后给对应的文件位置寄存器发送读文件消息，文件位置寄存器通过 FILEID 定位，如果定位失败，给文件访问客户端回复文件不存在。如果定位成功则将 FILEID 对应的文件记录信息发送给文件访问客户端；

步骤 S905，文件访问客户端收到读文件应答，通过 FILEID 给文件位置寄存器发送读文件的数据块的请求，文件位置寄存器根据 FILEID 找到数据块记录，将数据块所在磁盘信息发送给文件访问客户端；

步骤 S906，文件访问客户端收到信息，根据文件位置寄存器返回的磁盘信息到文件访问服务器读取数据块信息，并将数据块内容读出；

步骤 S907，读出文件内容之后，给用户返回文件内容信息。

优选实施例三

删除文件过程

图 10 是根据本发明优选实施例的删除文件过程的流程示意图，如图 10 所示，该流程包括如下步骤：

5 步骤 S1001，用户通过文件访问客户端删除文件，删除文件请求发送到文件访问客户端线程，文件访问客户端将要删除的文件对象（即目标文件）全路径发送给目录树服务器，请求删除文件；

 步骤 S1002，目录树服务器先在命名空间中查找，查看待写文件是否已经存在。如果不存在则给文件访问客户端回复文件不存在的失败码；

10 步骤 S1003，如果存在则根据文件名全路径找到文件对应的记录，将命名空间中文件信息删除，删除文件记录和字典表中文件名记录，将文件的 FILEID 和文件位置寄存器 ID 信息返回给文件访问客户端；

 步骤 S1004，文件访问客户端收到消息之后给对应的文件位置寄存器发送删除文件消息，文件位置寄存器通过 FILEID 定位，如果定位失败，给文件访问客户端回复文件不存在。如果定位成功则通过 FILEID 找到对应的文件记录记录，并通过 FILEID 找到对应的数据块记录，通知文件访问客户端删除数据块信息，并将 FILEID 对应的数据块记录和文件记录删除；

 步骤 S1005，文件访问客户端删除数据块请求，给文件访问服务器发送删除数据块消息，文件访问服务器根据消息中数据块的位置删除对应数据块；

20 步骤 S1006，文件访问客户端收到文件访问服务器删除数据块响应之后给用户回复删除文件成功。

 通过本发明的上述实施例、优选实施例和优选实施方式的技术方案，增加了目录树服务器，并将元数据分为两部分存储，其中，目录树服务器只负责对命名空间进行管理和存储，文件位置寄存器对文件具体的位置和数据块信息进行管理和存储。而在
25 相关技术中，只有单个分布式文件系统服务器（相当于文件位置寄存器），并且所有的数据存储在同一个服务器上，如果系统容量要求比较高的话，对服务器内存的要求就很高；存储对象操作并发数也受到服务器性能的影响。

综上所述，通过本发明实施例的方案设计中目录树服务器和文件位置寄存器的分工合作，目录树服务器上完成存储对象命名空间的管理和存储，文件位置寄存器完成对文件和数据块的操作，降低了对服务器的内存要求，提高了系统的性能，并且文件位置寄存器可以在线扩展，提高系统扩容能力，对系统的容量有很大的提高。

- 5 显然，本领域的技术人员应该明白，上述的本发明的各模块或各步骤可以用通用的计算装置来实现，它们可以集中在单个的计算装置上，或者分布在多个计算装置所组成的网络上，可选地，它们可以用计算装置可执行的程序代码来实现，从而，可以将它们存储在存储装置中由计算装置来执行，或者将它们分别制作成各个集成电路模块，或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样，本发明
- 10 不限制于任何特定的硬件和软件结合。

以上所述仅为本发明的优选实施例而已，并不用于限制本发明，对于本领域的技术人员来说，本发明可以有各种更改和变化。凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

工业实用性

- 15 如上所述，本发明实施例提供的一种请求响应方法、装置及分布式文件系统具有以下有益效果：降低了对服务器的内存要求，提高了系统的性能，并且文件位置寄存器可以在线扩展，提高系统扩容能力，对系统的容量有很大的提高。

权 利 要 求 书

1. 一种请求响应方法，包括：

接收来自文件访问客户端的操作目标文件的请求，其中，所述请求中携带有所述目标文件的全路径；

确定所述全路径对应的文件标识和文件位置寄存器标识，其中，所述文件位置寄存器标识对应的文件位置寄存器用于提供所述目标文件的数据块的位置信息；

发送所述文件标识和所述文件位置寄存器标识至所述文件访问客户端。

2. 根据权利要求 1 所述的方法，其中，在所述请求为写入所述目标文件的请求的情况下，确定所述全路径对应的所述文件标识和所述文件位置寄存器标识包括：

根据所述全路径，生成所述目标文件的文件标识；

根据预定策略，确定与所述文件标识对应的文件位置寄存器标识。

3. 根据权利要求 2 所述的方法，其中，在确定所述全路径对应的所述文件标识和所述文件位置寄存器标识之后，所述方法还包括：

保存所述全路径与所述文件标识、所述文件位置寄存器标识的关联关系。

4. 根据权利要求 2 所述的方法，其中，在发送所述文件标识和所述文件位置寄存器标识至所述文件访问客户端之后，所述方法还包括：

所述文件访问客户端获取所述位置信息，其中，所述位置信息是所述文件位置寄存器根据所述文件标识生成并发送给所述文件访问客户端的；

所述文件访问客户端将所述目标文件写入所述位置信息所对应的磁盘位置。

5. 根据权利要求 1 至 4 中任一项所述的方法，其中，在所述请求为读取或者删除所述目标文件的请求的情况下，确定所述全路径对应的所述文件标识和所述文件位置寄存器标识包括：

查找所述全路径对应的所述文件标识和所述文件位置寄存器标识。

6. 根据权利要求 5 所述的方法，其中，在发送所述文件标识和所述文件位置寄存器标识至所述文件访问客户端之后，所述方法还包括：

所述文件访问客户端获取所述位置信息，其中，所述位置信息是所述文件位置寄存器根据所述文件标识从保存的文件记录中查找并发送给所述文件访问客户端的；

所述文件访问客户端读取或者删除所述位置信息所对应的磁盘位置中保存的所述目标文件。

7. 一种请求响应装置，包括：

接收模块，设置为接收来自文件访问客户端的操作目标文件的请求，其中，所述请求中携带有所述目标文件的全路径；

确定模块，设置为确定所述全路径对应的文件标识和文件位置寄存器标识，其中，所述文件位置寄存器标识对应的文件位置寄存器用于提供所述目标文件的数据块的位置信息；

发送模块，设置为发送所述文件标识和所述文件位置寄存器标识至所述文件访问客户端。

8. 根据权利要求7所述的装置，其中，所述确定模块包括：

生成单元，设置为根据所述全路径，生成所述目标文件的文件标识；

确定单元，设置为根据预定策略，确定与所述文件标识对应的文件位置寄存器标识。

9. 根据权利要求8所述的装置，其中，所述装置还包括：

保存模块，设置为保存所述全路径与所述文件标识、所述文件位置寄存器标识的关联关系。

10. 根据权利要求7至9中任一项所述的装置，其中，所述确定模块包括：

查找单元，设置为查找所述全路径对应的所述文件标识和所述文件位置寄存器标识。

11. 一种分布式文件系统，包括文件访问客户端、文件位置寄存器和文件访问服务器，所述文件位置寄存器的数目为多个，所述系统还包括：

目录树元数据存储服务器，设置为根据来自所述文件访问客户端的操作目标文件的请求中携带的目标文件的全路径，确定所述全路径对应的文件标识和文件位置寄存器标识，并发送所述文件标识和所述文件位置寄存器标识至所述

文件访问客户端，其中，多个文件位置寄存器中所述文件位置寄存器标识对应的文件位置寄存器用于提供所述目标文件的数据块的位置信息。

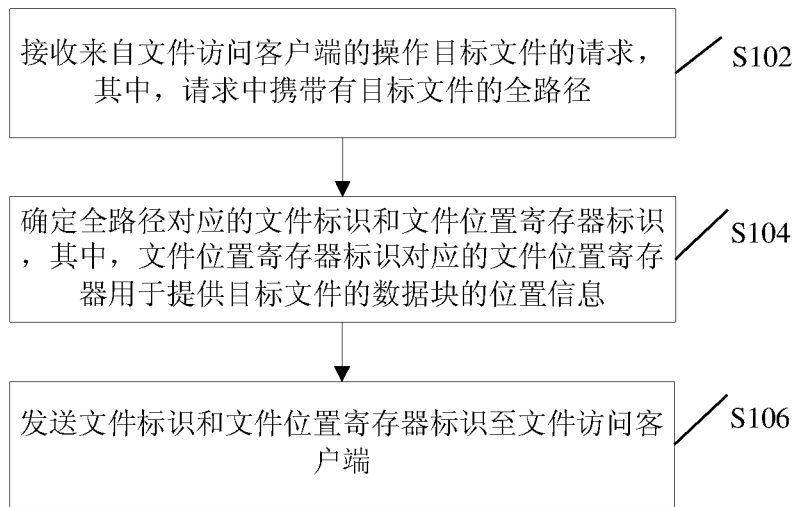


图 1



图 2

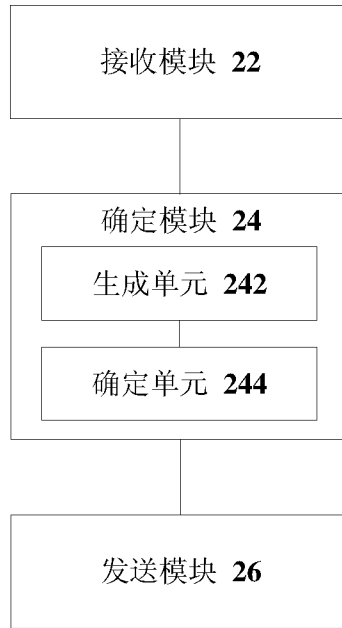


图 3

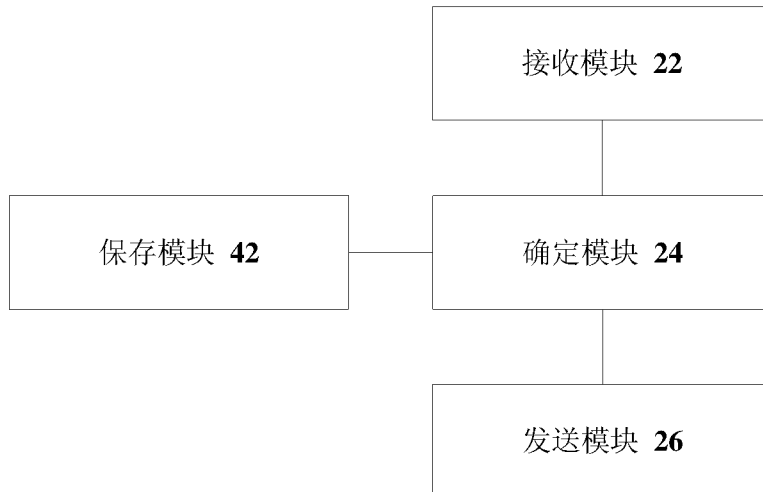


图 4

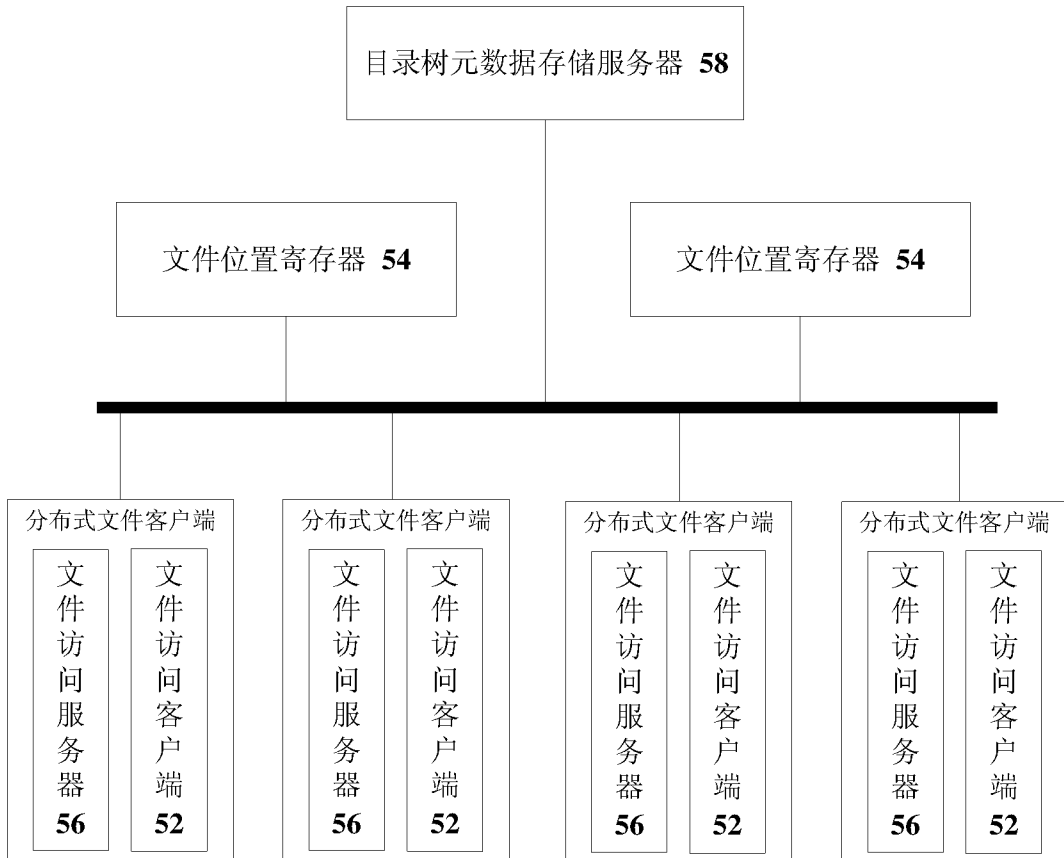


图 5

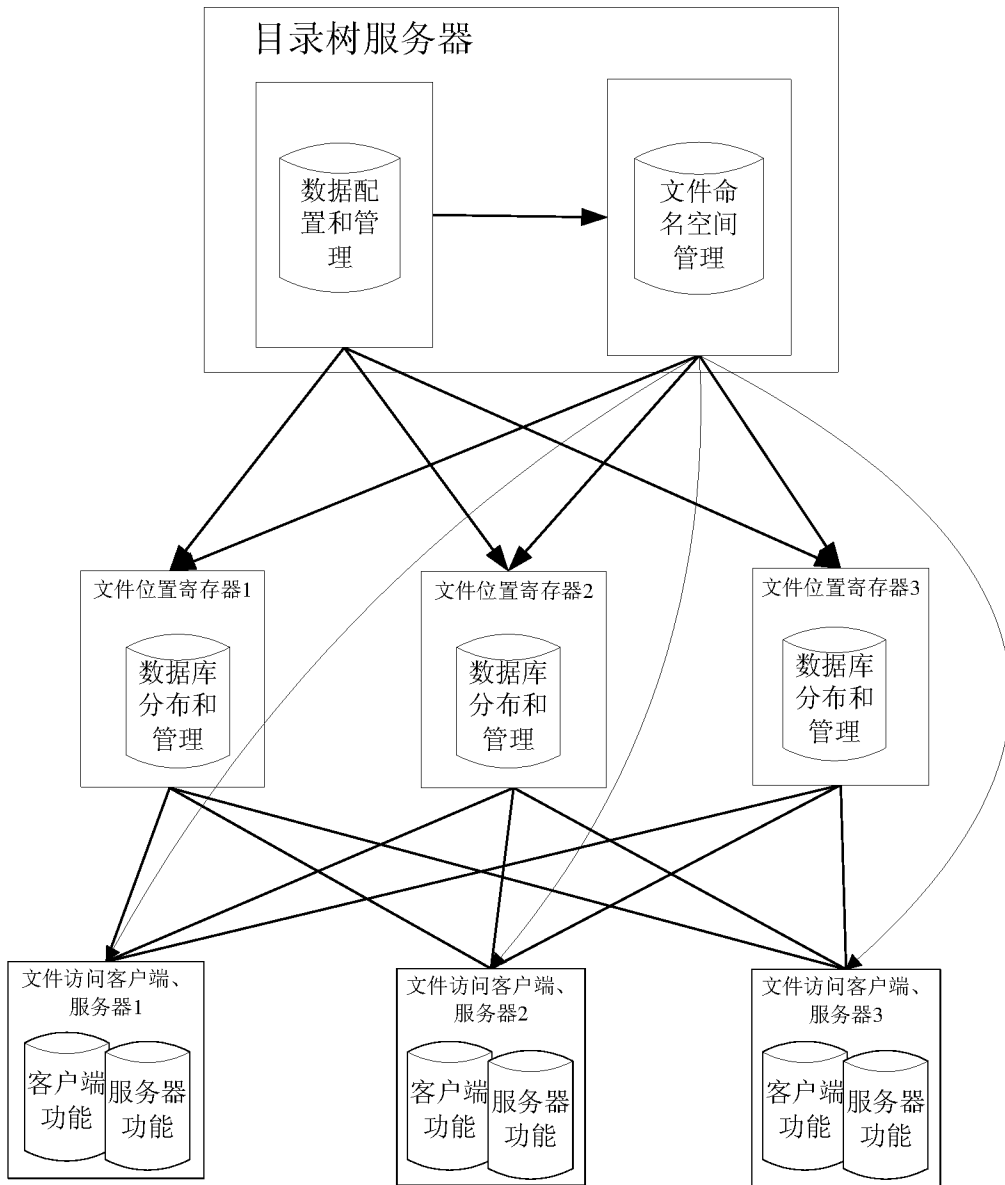


图 6

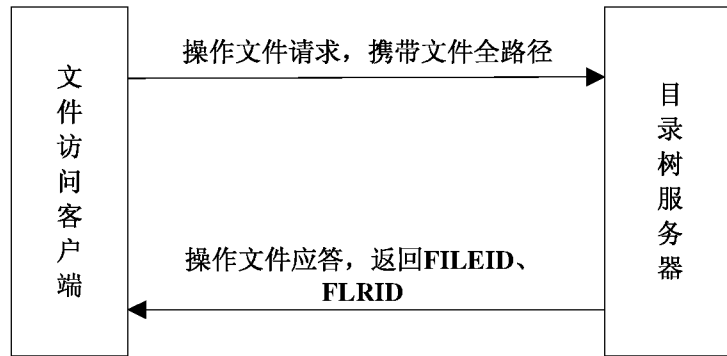


图 7a

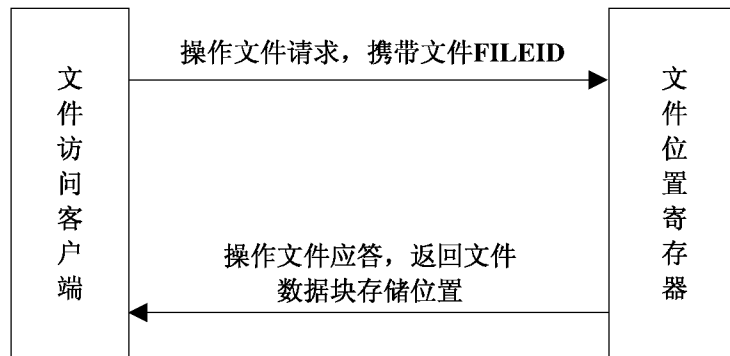


图 7b

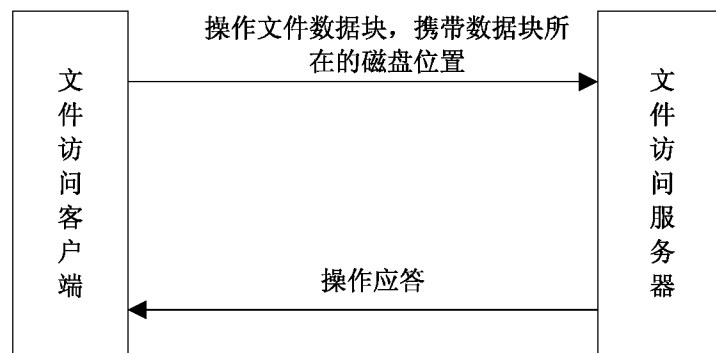


图 7c

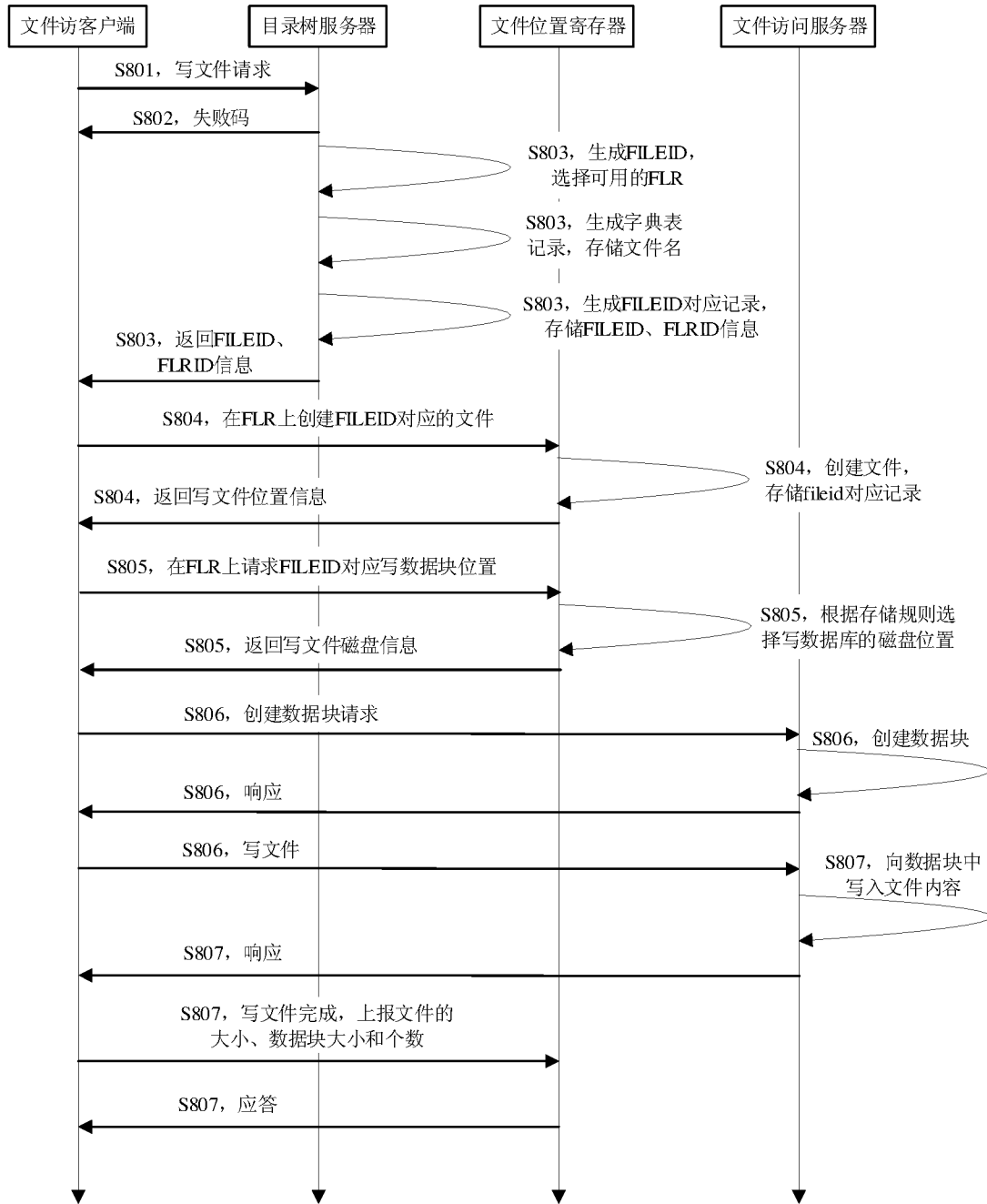


图 8

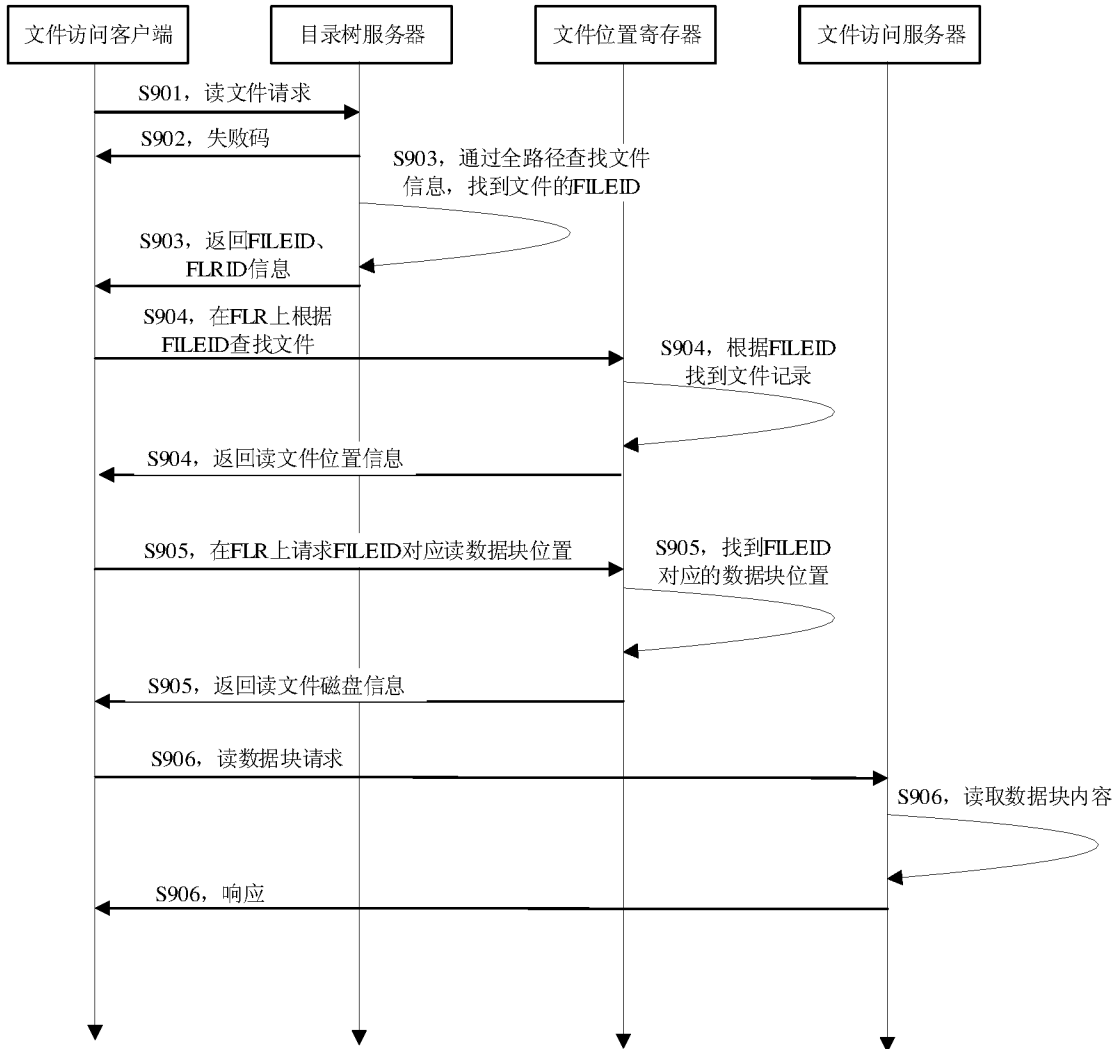


图 9

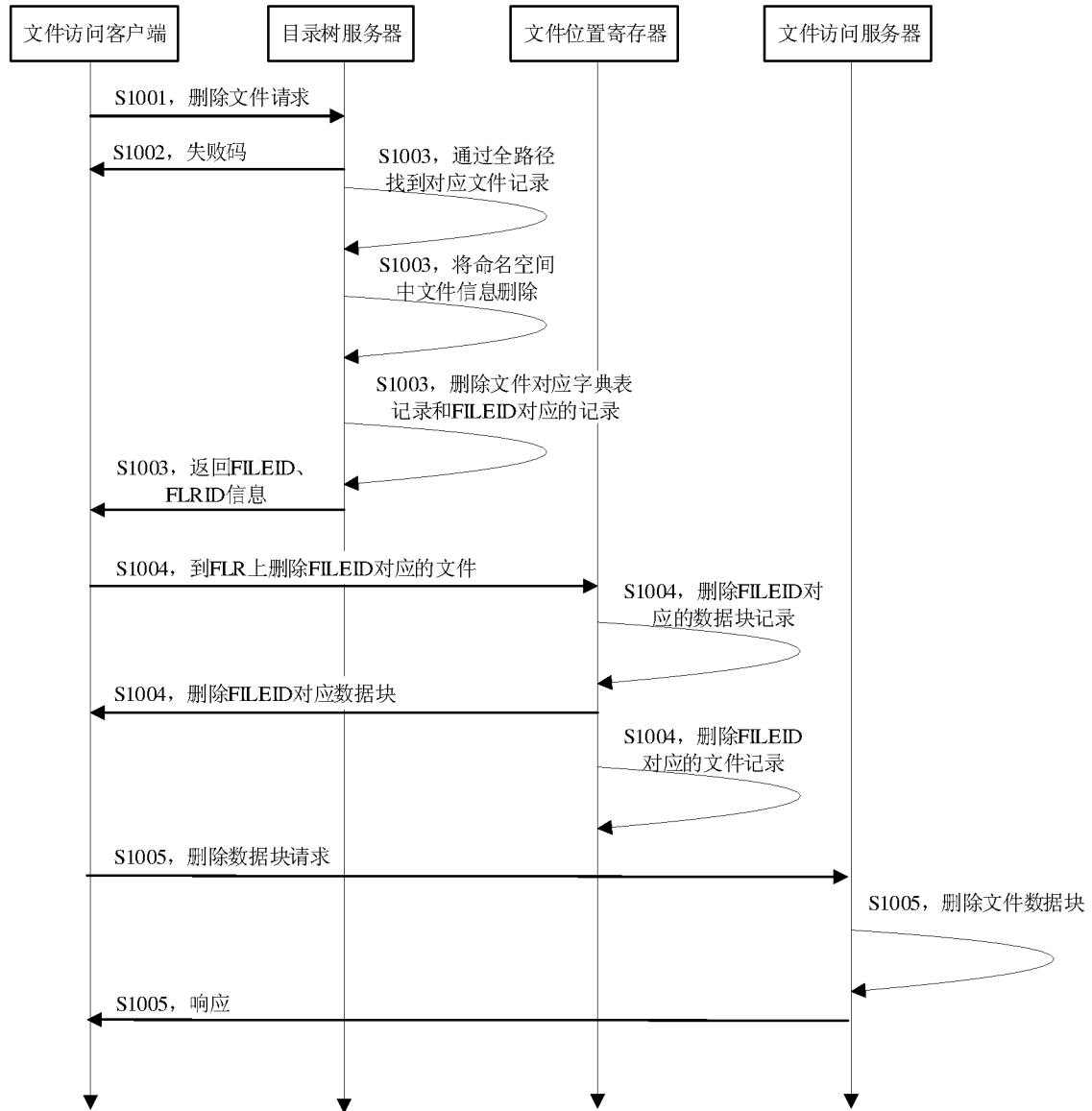


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2014/075522

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (2006.01) i; H04L 29/06 (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)

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)

CPRSABS; CNTXT; CNKI; VEN: FLR, DFS, ID, location register, metadata server, distributed file system, identifier, file, request

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 101997823 A (LENOVO BEIJING CO., LTD.) 30 March 2011 (30.03.2011) description, [paragraphs [0050]-[0058], [0062]-[0064], and [0082] -[0095], and figures 2 and 7	1-11
A	CN 102890716 A (NANJING ZHONGXING NEW SOFTWARE CO., LTD.) 23 January 2013 (23.01.2013) the whole document	1-11
A	CN 101520805 A (ZTE COMMUNICATION CO., LTD.) 02 September 2009 (02.09.2009) the whole document	1-11

II 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 15 August 2014 (15.08.2014)	Date of mailing of the international search report 26 August 2014 (26.08.2014)
--	---

Name and mailing address of the ISA 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 YU, Ruifu Telephone No. (86-10) 62089387
--	---

INTERNATIONAL SEARCH REPORT
 Information on patent family members

International application No.
 PCT/CN2014/075522

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101997823 A	30 March 2011	CN 101997823 B	02 October 2013
CN 102890716 A	23 January 2013	W O 2014048396 A I	03 April 2014
CN 101520805 A	02 September 2009	U S 2011307534 A I	15 December 2011
		CN 101520805 B	11 May 2011
		EP 24 13251 A I	01 February 2012
		W O 2010108368 A I	30 September 2010

<p>A. 主题的分类</p> <p>G06F 17/30 (2006. 01) i ; H04L 29/06 (2006. 01) i</p> <p>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</p>														
<p>B. 检索领域</p> <p>检索的最低限度文献 (标明分类系统和分类号)</p> <p>G06F ; H04L</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CPRSABS ;CNTXT ;CNKI ;VEN:位置寄存器, FLR, 元数据服务器, 分布式文件系统, DFS, 标识, ID, 识别符, 文件, 请求 ; location register, metadata server, distributed file system, identifier, file, request</p>														
<p>C. 相关文件</p> <table border="1" style="width:100%; border-collapse: collapse;"> <thead> <tr> <th style="width:10%;">类 型 *</th> <th style="width:70%;">引用文件, 必要时, 指明相关段落</th> <th style="width:20%;">相关的权利要求</th> </tr> </thead> <tbody> <tr> <td style="text-align:center;">X</td> <td>CN 101997823 A (联想北京有限公司) 2011 年 3 月 30 日 (2011 - 03 - 30) 说明书第 [0050] - [0058]、[0062] - [0064] 和 [0082] - [0095] 段、图2和7</td> <td style="text-align:center;">1-1 1</td> </tr> <tr> <td style="text-align:center;">A</td> <td>CN 102890716 A (南京中兴新软件有限责任公司) 2013 年 1 月 23 日 (2013 - 01 - 23) 全文</td> <td style="text-align:center;">1-1 1</td> </tr> <tr> <td style="text-align:center;">A</td> <td>CN 101520805 A (中兴通讯股份有限公司) 2009 年 9 月 02 日 (2009 - 09 - 02) 全文</td> <td style="text-align:center;">1-1 1</td> </tr> </tbody> </table>			类 型 *	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 101997823 A (联想北京有限公司) 2011 年 3 月 30 日 (2011 - 03 - 30) 说明书第 [0050] - [0058]、[0062] - [0064] 和 [0082] - [0095] 段、图2和7	1-1 1	A	CN 102890716 A (南京中兴新软件有限责任公司) 2013 年 1 月 23 日 (2013 - 01 - 23) 全文	1-1 1	A	CN 101520805 A (中兴通讯股份有限公司) 2009 年 9 月 02 日 (2009 - 09 - 02) 全文	1-1 1
类 型 *	引用文件, 必要时, 指明相关段落	相关的权利要求												
X	CN 101997823 A (联想北京有限公司) 2011 年 3 月 30 日 (2011 - 03 - 30) 说明书第 [0050] - [0058]、[0062] - [0064] 和 [0082] - [0095] 段、图2和7	1-1 1												
A	CN 102890716 A (南京中兴新软件有限责任公司) 2013 年 1 月 23 日 (2013 - 01 - 23) 全文	1-1 1												
A	CN 101520805 A (中兴通讯股份有限公司) 2009 年 9 月 02 日 (2009 - 09 - 02) 全文	1-1 1												
<p><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。</p>														
<p>* 引用文件的具体类型：</p> <table style="width:100%;"> <tr> <td style="width:50%; vertical-align: top;"> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“V” 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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <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</p> </td> <td style="width:50%; vertical-align: top;"> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p> </td> </tr> </table>			<p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“V” 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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <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</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>										
<p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“V” 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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <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</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>													
<p>国际检索实际完成的日期</p> <p style="text-align:center;">2014 年 8 月 15 日</p>	<p>国际检索报告邮寄日期</p> <p style="text-align:center;">2014 年 8 月 26 日</p>													
<p>ISA/CN 的名称和邮寄地址</p> <p>中华人民共和国国家知识产权局 (ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国</p> <p>传真号 (86-10) 62019451</p>	<p>授权官员</p> <p style="text-align:center;">于瑞甫</p> <p>电话号码 (86-10) 62089387</p>													

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2014/075522

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	101997823	A	2011年 3月 30日	CN	101997823	B	2013年 10月 02日
CN	102890716	A	2013年 1月 23日	wo	2014048396	AI	2014年 4月 03日
CN	101520805	A	2009年 9月 02日	us	2011307534	AI	2011年 12月 15日
				CN	101520805	B	2011年 5月 11日
				EP	2413251	AI	2012年 2月 01日
				Wo	2010108368	AI	2010年 9月 30日