



## (12) 发明专利

(10) 授权公告号 CN 102947815 B

(45) 授权公告日 2016. 01. 20

(21) 申请号 201180029757. 8

H04L 9/00(2006. 01)

(22) 申请日 2011. 06. 06

(56) 对比文件

(30) 优先权数据

12/818, 515 2010. 06. 18 US

US 2009/0070356 A1, 2009. 05. 12,

US 2009/0190760 A1, 2009. 07. 30, 说明书第  
1 页第 4 段至第 6 页 52 段。

(85) PCT 国际申请进入国家阶段日

2012. 12. 17

CN 101582076 A, 2009. 11. 18,

(86) PCT 国际申请的申请数据

PCT/US2011/039318 2011. 06. 06

审查员 陈英

(87) PCT 国际申请的公布数据

W02011/159517 EN 2011. 12. 22

(73) 专利权人 微软技术许可有限责任公司

地址 美国华盛顿州

(72) 发明人 E·C·布朗 T·E·乔利

J-T·普芬宁

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 陈斌

(51) Int. Cl.

G06F 15/16(2006. 01)

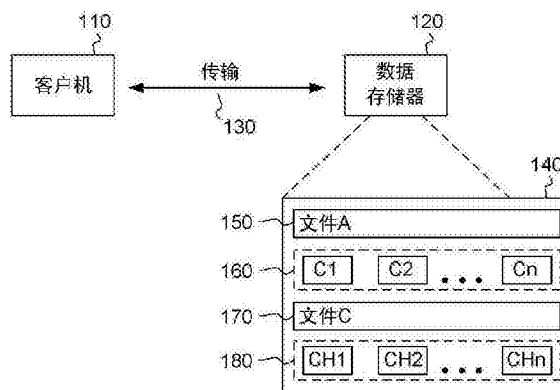
权利要求书2页 说明书11页 附图4页

(54) 发明名称

数据的存储与发送的优化

(57) 摘要

本发明涉及数据存储和数据的传输的端对端优化的方法、系统、和计算机程序产品。数据如何被存储在数据存储器上的细节被展示给客户机和应用。能使客户机和应用向数据存储器做出请求来获得在数据存储器中实际所存储的数据,从而消除对于所请求的数据的冗余处理。在数据存储器中利用对数据的压缩和去重来增加效率且减少在 LAN 或 WAN 上传输的数据的延时。



1. 一种在包括客户机和数据存储服务器的计算环境中的方法,所述方法用于将所述数据存储服务器中的存储优化的细节展示给所述客户机,所述方法包括:

由所述客户机访问描述在所述数据存储服务器上的文件数据的存储的元数据,其中访问所述元数据包括向数据存储服务器发送对于所述元数据的请求,并且其中所述文件数据被以与所述文件数据的原本形式不同的形式存储在所述数据存储服务器上,且其中所述元数据展示存储在所述数据存储服务器上的所述文件数据的所述不同的形式;

在所述客户机处从所述数据存储服务器接收信息,所述信息描述在数据存储服务器上的文件数据的存储的元数据,其中描述在数据存储服务器上的文件数据的存储的元数据包括描述针对所述文件的压缩信息的数据,所述压缩信息能够被用于对所述文件数据进行解压缩,其中所述数据存储服务器远离所述客户机;

由所述客户机发送对于文件数据的随后的请求,其中所述对于文件数据的随后的请求至少部分基于描述在数据存储服务器上的文件数据的存储的元数据,并且其中所述随后的请求针对存储在所述数据存储服务器上的所述文件数据的所述不同的形式;

在所述客户机处,接收所述文件数据,其中所述文件数据是以存储在所述数据存储服务器上的所述文件数据的所述不同的形式被接收;以及

由所述客户机基于所述元数据将所述文件数据的所述不同的形式转换成所述文件数据的原本形式。

2. 如权利要求 1 所述的方法,其特征在于,描述在所述数据存储服务器上的文件数据的存储的所述元数据还包括:描述对所述数据存储服务器上的所述文件数据进行去重复所得的文件数据的存储的数据。

3. 如权利要求 1 所述的方法,其特征在于,描述在所述数据存储服务器上的文件数据的存储的所述元数据还包括所述文件数据的子集的加密散列。

4. 如权利要求 1 所述的方法,其特征在于,描述在所述数据存储服务器上的文件数据的存储的所述元数据还包括所述文件数据的多个子集的每一个的加密散列。

5. 如权利要求 1 所述的方法,其特征在于,描述在所述数据存储服务器上的文件数据的存储的所述元数据还包括描述所述文件数据的经压缩的子集的数据。

6. 一种在包括客户机和数据存储服务器的计算环境中的方法,所述方法用于将所述数据存储服务器中的存储优化的细节展示给所述客户机,所述方法包括:

在所述数据存储服务器处接收访问描述所述数据文件的存储的元数据的请求;

响应于接收到所述请求,从所述数据存储服务器发送描述在所述数据存储服务器上的文件数据的存储的元数据,其中所述文件数据被以与所述文件数据的原本形式不同的形式存储在所述数据存储服务器上,且其中所述元数据展示存储在所述数据存储服务器上的所述文件数据的所述不同的形式;

在所述数据存储服务器处接收来自所述客户机的对于以存储在所述数据存储服务器上的所述文件数据的所述不同的形式的文件数据的请求,其中所述对于文件数据的请求至少部分基于描述在数据存储服务器上的文件数据的存储的元数据,并且其中所述客户机远离所述数据存储服务器;和

从所述数据存储服务器发送信息,所述信息包括文件数据和描述在所述数据存储服务器上的文件数据的存储的附加元数据,其中所述附加元数据包括描述针对所述文件数据的

压缩信息的数据,所述压缩信息能够被用于对所述文件数据进行解压缩,并且其中所述附加元数据能够被用于将所述文件数据的所述不同的形式转换成所述文件数据的所述原本形式。

7. 如权利要求 6 所述的方法,其特征在于,描述在所述数据存储服务器上的文件数据的存储的所述元数据包括:描述对所述数据存储服务器上的所述文件数据进行去重复所得的文件数据的存储的数据。

8. 如权利要求 6 所述的方法,其特征在于,描述在所述数据存储服务器上的文件数据的存储的所述元数据包括所述文件数据的子集的加密散列。

9. 如权利要求 6 所述的方法,其特征在于,描述在所述数据存储服务器上的文件数据的存储的所述元数据包括所述文件数据的多个子集的每一个的加密散列。

10. 如权利要求 6 所述的方法,其特征在于,描述在所述数据存储服务器上的文件数据的存储的所述元数据包括描述所述文件数据的经压缩的子集的数据。

11. 一种用于将数据存储服务器中的存储优化的细节展示给客户机的方法,所述方法包括:

从所述客户机向数据存储服务器发送访问描述文件的存储的元数据的请求;

在所述客户机处,从所述数据存储服务器接收信息,所述信息包括描述在所述数据存储服务器上的文件数据的存储的信息,其中,所述文件数据在所述数据存储服务器上是被压缩的,并且其中描述文件数据的存储的信息包括描述针对所述文件的压缩信息的数据,所述压缩信息能够被用于对所述文件数据进行解压缩;以及

由所述客户机发送对于经压缩的文件数据的随后的请求,其中所述对于文件的随后的请求至少部分基于描述在所述数据存储服务器上的文件数据的存储的元数据,并且其中所述对于文件的随后的请求是针对存储在所述数据存储服务器上的所述经压缩的文件数据;

在所述客户机处接收所述经压缩的文件数据;以及

由所述客户机基于所述元数据对所述经压缩的文件数据解压。

12. 如权利要求 11 所述的方法,其特征在于,包括描述在所述数据存储服务器上的所述文件数据的存储的信息的所述信息包括:描述源自对于所述数据存储服务器上的所述文件数据去重复的文件数据的存储的数据。

13. 如权利要求 11 所述的方法,其特征在于,包括描述在所述数据存储服务器上的文件数据的存储的信息的所述信息包括:所述文件数据的子集的加密散列。

14. 如权利要求 11 所述的方法,其特征在于,包括描述在所述数据存储服务器上的文件数据的存储的信息的所述信息包括:所述文件数据的多个子集的每一个的加密散列。

15. 如权利要求 11 所述的方法,其特征在于,包括描述在所述数据存储服务器上的文件数据的存储的信息的所述信息包括:描述所述文件数据的经压缩的子集的数据。

## 数据的存储与发送的优化

[0001] 背景

[0002] 存储优化功能变得越来越重要从而在文件服务器和数据存储市场中有竞争性。网络话务优化在计算机和网络环境中以及结合到现有网络基础结构中的设备中也是重要的，且执行网络话务的实时优化可提供有用的益处。

[0003] 在计算机上被生成、发送、和存储的数据量快速持续增长。消费者和竞争者推动着向着使用数据优化技术的增长的趋势，从而减少静止数据的存储要求。例如，数据可被压缩且所存储的数据中的冗余可被减少从而减少存储数据所要求的空间。类似技术还可被应用于减少在网络上传输的数据量，因此减少 LAN 和 WAN 带宽成本并降低应用延时。然而，数据存储和数据传输的当前解决方法大为分开且不同的，没有统一的解决方案被已知。由于存储和传输技术是分开的，当一起来看数据存储和数据传输时，存在冗余、不兼容、和不必要的开销。

[0004] 作为示例，当被存储在数据存储服务器上时，被存储在服务器（即，数据存储器）上的文件可被压缩并存储在分开的段（如，组块）中。当客户机请求文件从服务器传输至客户机时，在将文件传输至客户机之前，服务器必须重新组装这些组块并解压缩文件来重构文件。

[0005] 类似地，为了减少（如，在网络上的）传输带宽、延时、或传输成本，则网络代理可在传输前获得文件或再次压缩、将被压缩的文件传输至另一个端点、且然后在传输路径的另一端处解压缩。

[0006] 可能有用的是，统一的数据优化工具和涵盖存储器、传输协议、文件系统 API、数据存储、服务器、客户机、应用、和云的各种技术。通过传递可由数据传输和传输协议所利用的静止数据的优化存储，这样的工具和技术可扩展并增强现有的零零碎碎和分开的数据存储器和数据传输解决方法。

### 发明内容

[0007] 本发明涉及数据的存储和传输的端对端优化的方法、系统、设备、和计算机程序产品。例如，此处公开的实施例提供了利用和增加数据存储和数据传输二者的效率和优化。

[0008] 一个示例性实施例提供了一方法，用于将数据存储器服务器中的存储优化的细节展示给客户机。该方法包括访问描述数据存储服务器上的文件数据的存储的元数据，其中文件数据被以与文件数据的原本形式不同的形式存储在数据存储服务器上。该元数据将文件数据的存储形式展示为存储在数据存储服务器上。

[0009] 客户机可发送对于文件数据的请求至存储服务器，且客户机可从数据存储服务器接收信息，该信息包含文件数据、描述在数据存储服务器上的文件数据的存储的附加元数据、和 / 或代表该文件数据的至少一部分的数据。

[0010] 另一个示例性实施例提供了用于将数据存储器服务器中的存储优化的细节展示给客户机。这个方法包括发送描述在数据存储服务器上的文件数据的存储的元数据。该文件数据被以与该文件数据的原本形式不同的形式存储在数据存储服务器上，且元数据将该

文件数据的存储形式展示为存储在数据存储服务器上。

[0011] 数据存储服务器接收来自计算系统的对于文件数据的请求且数据存储服务器发送信息,该信息包含文件数据、描述在数据存储服务器上的文件数据的存储的附加元数据、和 / 或代表该文件数据的至少一部分的数据。

[0012] 另一个示例性实施例提供了一计算机程序产品,用于将数据存储服务器中的存储优化的细节展示给客户机。该计算机程序产品包括计算机可执行指令,特别地,用于从计算系统发送对于文件数据的请求至数据存储服务器并从数据存储服务器接收信息,该信息包括描述在该数据存储服务器上的文件数据的存储的信息。

[0013] 本发明的附加特征和优点将在以下描述中叙述,并且其一部分根据本描述将是显而易见的,或者可通过对本发明的实践来获知。本发明的特征和优点可通过在所附权利要求书中特别指出的工具和组合来实现和获取。本发明的这些以及其他特征将通过以下描述和所附权利要求书变得更加显而易见,或者可通过对下文中所阐述的本发明的实践来获知。

[0014] 注意到提供本概述以便使用简化的形式介绍将在以下详细描述中进一步描述的一些概念。本发明内容既不在旨在标识要求保护的的主题的关键特征或必要特征,也不旨在用于帮助确定要求保护的的主题的范围。

[0015] 附图简述

[0016] 为了描述可获得本发明的上述和其他优点特征的方式,将通过引用附图中示出的本发明的特定实施例来呈现以上简要描述的本发明的更具体描述。应当理解,这些附图只描述了本发明的典型实施例,并且因此不被认为是对其范围的限制,本发明将通过使用附图用附加特征和细节来描述和说明,在附图中:

[0017] 图 1 示出数据的存储与传输的端对端优化的示例。

[0018] 图 2 示出数据的存储与传输的端对端优化的示例性体系结构。

[0019] 图 3 示出从客户机的角度来看,用于将数据存储服务器中的存储优化的细节展示给客户机的示例性方法。

[0020] 图 4 示出从服务器的角度来看,用于将数据存储服务器中的存储优化的细节展示给客户机的示例性方法。

## 具体实施方式

[0021] 本发明涉及数据的存储和传输的端对端优化的方法、系统、设备、和计算机程序产品。例如,此处公开的实施例提供了利用数据存储和传输的效率和优化。本发明扩展为用于将数据存储服务器中的存储优化的细节展示给客户机的方法、系统、和计算机程序产品。本发明的实施例可以包括含有各种计算机硬件或模块的专用或通用计算机,这将在全文做出进一步讨论。

[0022] 一个示例性实施例提供了一方法,用于将数据存储服务器中的存储优化的细节展示于客户机。该方法包括访问描述数据存储服务器上的文件数据的存储的元数据,其中文件数据被以与文件数据的原本形式不同的形式存储在数据存储服务器上。该元数据将文件数据的存储形式展示为存储在数据存储服务器上的形式。

[0023] 客户机可发送对于文件数据的请求至存储服务器,且客户机可从数据存储服务器

接收信息,该信息包含文件数据、描述在数据存储服务器上的文件数据的存储的附加元数据、和 / 或代表该文件数据的至少一部分的数据。

[0024] 另一个示例性实施例提供了用于将数据存储服务器中的存储优化的细节展示给客户机。这个方法包括发送描述在数据存储服务器上的文件数据的存储的元数据。该文件数据被以与该文件数据的原本形式不同的形式存储在数据存储服务器上,且元数据展示了如被存储在数据存储服务器上一样的文件数据的存储形式。

[0025] 数据存储服务器接收来自计算系统的对于文件数据的请求且数据存储服务器发送信息,该信息包含文件数据、描述在数据存储服务器上的文件数据的存储的附加元数据、和 / 或代表该文件数据的至少一部分的数据。

[0026] 另一个示例性实施例提供了一计算机程序产品,用于将数据存储服务器中的存储优化的细节展示于客户机。该计算机程序产品包括计算机可执行指令,特别地,用于从计算系统发送对于文件数据的请求至数据存储服务器并从数据存储服务器接收信息,该信息包括描述在该数据存储服务器上的文件数据的存储的信息。

[0027] 本发明的各实施例可包括或利用专用或通用计算机,该专用或通用计算机包括诸如一个或多个处理器和系统存储器等计算机硬件,如以下更详细讨论的。本发明范围内的各实施例还包括用于承载或存储计算机可执行指令和 / 或数据结构的物理和其他计算机可读介质。这样的计算机可读介质可以是可由通用或专用计算机系统访问的任何可用介质。存储计算机可执行指令的计算机可读介质可以是物理存储介质。承载计算机可执行指令的计算机可读介质可以是传输介质。由此,作为示例而非限制,本发明的各实施例可包括至少两种显著不同的计算机可读介质:计算机存储介质和传输介质。

[0028] 计算机存储介质包括 RAM、ROM、EEPROM、CD-ROM 或其他光盘存储、磁盘存储或其他磁存储设备、或可用于存储计算机可执行指令或数据结构形式的所需程序代码装置且可由通用或专用计算机访问的任何其他介质。

[0029] 计算机程序产品可包括一个或多个计算机可读介质,其上具有计算机可执行指令,当在一个或多个计算机处理器上执行这些指令时,执行此处描述的方法、步骤、和动作。

[0030] “网络”被定义为允许在计算机系统和 / 或模块和 / 或其他电子设备之间传输电子数据的一个或多个数据链路。当信息通过网络或另一个通信连接(硬连线、无线、或者硬连线或无线的组合)传输或提供给计算机时,该计算机将该连接适当地视为传输介质。传输介质可包括可用于携带计算机可执行指令或数据结构形式的所需程序代码装置且可由通用或专用计算机访问的网络和 / 或数据链路。上述的组合也应被包括在计算机可读介质的范围内。

[0031] 此外,在到达各种计算机系统组件之后,计算机可执行指令或数据结构形式的程序代码装置可从传输介质自动传输到计算机存储介质(或反之亦然)。例如,通过网络或数据链路接收到的计算机可执行指令或数据结构可被缓存在网络接口模块(例如,“NIC”)内的 RAM 中,然后最终被传输到计算机系统 RAM 和 / 或计算机系统处的较不易失性的计算机存储介质。因而,应当理解,计算机存储介质可被包括在还利用(或甚至主要利用)传输介质的计算机系统组件中。

[0032] 计算机可执行指令例如包括,当在处理器处执行时使通用计算机、专用计算机、或专用处理设备执行某一功能或某组功能的指令和数据。计算机可执行指令可以是例如二进

制代码、诸如汇编语言之类的中间格式指令、或甚至源代码。尽管用结构特征和 / 或方法动作专用的语言描述了本主题,但可以理解,所附权利要求书中定义的主题不必限于上述特征或动作。相反,上述特征和动作是作为实现权利要求的示例形式而公开的。

[0033] 本领域的技术人员将理解,本发明可以在具有许多类型的计算机系统配置的网络计算环境中实践,这些计算机系统配置包括个人计算机、台式计算机、膝上型计算机、消息处理器、手持式设备、多处理器系统、基于微处理器的或可编程消费电子设备、网络 PC、小型计算机、大型计算机、移动电话、PDA、寻呼机、路由器、交换机等等。本发明也可在其中通过网络链接(或者通过硬连线数据链路、无线数据链路,或者通过硬连线和无线数据链路的组合)的本地和远程计算机系统两者都执行任务的分布式系统环境中实施。在分布式系统环境中,程序模块可以位于本地和远程存储器存储设备二者中。

[0034] 如此处所使用的,术语“模块”或“组件”指的是在计算系统上执行的软件对象或例程。此处所描述的不同组件、模块、引擎,以及服务可以实现为在计算系统上执行的对象或进程(例如,作为分开的线程)。尽管此处描述的系统和方法可用软件较佳地实现,但用硬件或软件和硬件的组合的实现也是可能的且已被想到。在此描述中,“计算实体”可以是如上文定义的任何计算系统,或者是在计算系统上运行的任何模块或模块的组合。

[0035] 图 1 示出其中可运行本发明的示例性环境。图 1 示出客户机 110、数据存储器 120、和客户机 110 与数据存储器 120 之间的数据传输 130。数据可以很多不同形式存储在数据存储器 120 上。

[0036] 此处呈现的实施例描述了方法、系统、和计算机程序产品,来结合和优化在诸如图 1 所示的环境中的数据的存储 140 和传输 130。

[0037] 文件可被以其原本形式存储在数据存储器中,作为连续文件。例如,文件 A 150 以包括所有比特、字节、和文件的数据(如可由应用呈现或期待的那样)的未改变的原始或原本的格式存储在数据存储器 120 中。数据还可被以各种可选格式存储。例如,数据可被以压缩格式存储从而减少必须的存储空间,且可使用技术来存储数据从而减少冗余并对于存储在数据存储器上的数据去重复。

[0038] 数据可被以组块或块(其中文件被分解为分开的且不同的数据子集)存储在数据存储器上。例如,文件可在数据存储器中被存储为组块 160C1 到 Cn。来自文件的组块、数据子集,有时也被称为块,且这两个术语,组块和块,此处被互换地使用。(要注意的是,如此处使用的术语,文件,描述了任何逻辑上相关的组或数据量。)

[0039] 数据存储器可具有将文件分开为组块的算法,从而优化数据的存储。例如,文件可被分开为组块 160C1 到 Cn,从而以更有效或紧凑的方式将文件存储在数据存储器中。通过减少文件中的冗余,分为组块的文件还可更有效地存储。例如,组块 C1 可出现在文件中多于一次。通过将文件分开为组块,组块 C1 仅被写入数据存储器一次,且在文件中的组块 C1 的每一个重复出现可通过对于组块 C1 的引用或指针所替代。

[0040] 如可理解的,组块或块并不必须是任何固定长度且可以是任何长度、任何数据量、或文件的任何部分,包括整个文件。文件的组块或块可以是文件中的任意长度和 / 或偏移量。将文件分为组块或块可遵循任何算法或技术,且组块的大小可由数据存储的特定考量所影响或规定,该特定考量有关于何数据要被持久存储,或有关于数据要通过何传输路径来传输。

[0041] 数据还可被以压缩格式存储在数据存储器中。例如,文件 C170 被以压缩格式存储,其中使用压缩算法压缩原始文件来创建文件,文件 C 170,其相比原始、未压缩的文件数据,在数据存储器中占据较少的存储空间。可通过业界中已知的技术,诸如 Lempel-Ziv (LZ), Lempel-Ziv-Welch (LZW), 和 MPEG 压缩,来执行文件和数据压缩。

[0042] 也可在数据存储器上采用压缩和组块化(或块化)的组合。例如,文件可被分为组块,然后组块可被压缩并存储为经压缩的组块 180CH1 到 CHn。

[0043] 通过对存储在数据存储器中文件和的数据去重复,来获得另一个优化。去重复标识了存储在数据存储器中的相同文件或在相同文件部分(可能在不同文件中出现),并将重复的文件或数据部分中的除了一个之外的所有用对于该文件或数据部分的引用副本的引用替代。通过对文件进行去重复,特定文件或数据部分的仅一个副本将被存储在数据存储器中,藉此节省了被多个、重复的文件或数据部分所占据的存储空间。

[0044] 还可在文件组块级别上执行去重复。例如,如果两个或更多个文件被组块化为数据组块,则在数据存储器中,重复的组块可用对于冗余组块的副本的引用所替代。例如,文件可被存储在数据存储器 120 上,存储为组块 C1 和对于与以组块格式存储在数据存储器 120 中的其他文件相关联的已经存储的其他组块的引用。例如,文件 X (fileX)可被存储为对于组块 C1 到 Cn 的引用;文件 Y (fileY)可被存储为对于组块 CH1、C1、和 C2 的引用;且文件 Z (fileZ)可被存储为对于组块 C1 和压缩组块 CH2 到 CHn 的引用的列表。

[0045] 对于文件数据的去重复、组块化、和压缩也可被组合地执行。例如,文件可被存储在数据存储器上作为一个或多个组块,其中每一个组块已经被压缩。文件数据还可被以任何组合存储,其中一些文件被存储为未压缩、一些文件被存储为压缩的、一些文件被存储为组块化的格式、且一些文件被存储为组块藉此一些组块是被压缩的且一些组块是未压缩的。

[0046] 一般而言,当客户机从数据存储器请求数据时,客户机将要求整个文件或文件的某些逻辑部分的数据。例如,客户机可通过文件系统请求 `get(fileX)`,或可通过文件系统请求 `getFileBytes(fileX;bytes=100-1000)`。当文件或文件的一部分被从数据存储器 120 传输 130 至客户机 110 时,负担落在数据存储器上来对于压缩的数据解压缩并对于数据的组块重新组装,从而以客户机或应用所期待的格式重新组装并传输所请求的数据至客户机。

[0047] 此处描述的实施例允许客户机请求或访问数据存储器上与文件数据的存储相关的信息,从而通过向客户机提供与存储在数据存储器上的数据的存储细节相关的信息来获得效率和优化。例如,客户机 110 可请求数据存储器 120 告知客户机如何将文件 X 存储在数据存储器上。数据存储器可提醒客户机,文件 X 被存储为压缩组块 CH1 到 CH3。由于以压缩形式将经压缩的组块传输至客户机更为有效,然后客户机可请求数据存储器将组块 CH1 到 CH3 传输至客户机,而不是请求 `get(fileX)`,该命令将要求数据存储器在将文件传输至客户机前,必须解压缩组块 CH1 到 CH3 并重新组装该文件。

[0048] 实施例还允许客户机访问与数据存储器上文件数据的存储相关的信息,从而通过向客户机提供与存储在数据存储器上的数据的存储细节相关的信息来获得效率和优化。例如,客户机 110 可访问局部缓存或存储的信息,标识文件 X 如何被存储在数据存储器上。这个信息可由之前的请求获取,或可在客户机与数据存储器之间的之前的事务过程中被缓存。

[0049] 如果客户机已经具有被局部存储的或从相比数据存储器 120 具有更低的延时或传输成本的存储器位置获得的组块 CH1 的副本,可获得附加效率。在这样的情况下,然后客户机可仅从数据存储器请求 getChunk(CH3)。

[0050] 此处描述的实施例减少了客户机与数据存储器 and / 或集中式服务器之间的冗余的 LAN 和 / 或 WAN 通信。此处的实施例对于各网络文件系统协议,能进行存储和传输优化。例如,通过所描述的设备和技术, SMB 和 HTTP 协议可被扩展增强。

[0051] 标准文件系统协议(如, SMB 和 HTTP) 可被扩展来提供 API,其能使客户机从数据存储器请求数据,当该数据由数据存储器提供时,该数据展示了文件或数据部分如何被存储在数据存储器上的细节。例如,客户机 110 可从数据存储器 120 请求关于文件 X 如何被存储在数据存储器 120 上的数据。例如,客户机 110 可调用文件系统扩展,诸如 getStorageDetails(fileX),且数据存储器可用 {fileX:=chunks CH1,CH3} 来响应。现在具有对于文件 X 如何被存储在数据存储器上的细节的了解,然后客户机可决定如何从数据存储器请求与文件 X 相关联的数据。以标准方式,客户机可请求其原始或原本格式的整个文件。反之,此处的实施例能使客户机请求数据存储器将压缩的组块 CH3 传输至客户机。

[0052] 在一个实施例中,如图 3 中所示,客户机可访问 310 描述在数据存储服务器上的文件数据的存储的元数据,其中文件数据被以与文件数据的原本形式不同的形式存储在数据存储服务器上,且其中元数据展示了存储在数据存储服务器上的文件数据的存储形式。描述在数据存储服务器上的文件数据的存储的元数据可以是描述文件数据在数据存储器上如何被组块化、文件数据如何在数据存储器上被压缩、或者文件数据在数据存储器上如何被组块化且被压缩的信息。

[0053] 文件如何被组块化的细节可包括文件的哪些部分对应于服务器上所存储的每一个组块。组块化的细节还可包括构成文件的每一个组块的加密散列。组块的加密散列能使客户机、应用、和数据存储器唯一地标识每一个组块。使用这个信息,客户机、应用、或其他数据存储器可标识它是否已经具有可用的相同组块,如其加密散列所标识的那样。

[0054] 文件或数据部分(如,组块)如何被压缩的细节可包括原始未压缩的数据的加密散列,从而唯一地标识该数据。还可包括压缩数据的加密散列,从而唯一地标识经压缩的数据。细节还可包括被用于执行压缩的压缩类型(其是必须的,用于从数据存储器传输经压缩数据至另一个端点后,对于经压缩的数据解压缩)。压缩类型可包括,例如, LZ、LZW、MPEG 等。

[0055] 通过访问元数据,客户机可知道存储在数据存储器上的数据的存储细节。当客户机知道在数据存储器上的数据的存储细节时,客户机可发送 320 对于文件数据的请求至存储服务器。通过采用此处所述的实施例,客户机不需要请求整个文件,客户机可仅请求它可能需要的文件的这些组块,或可请求文件的压缩版本或文件的组块的压缩版本。在已经发送 320 了对于文件数据的请求后,客户机可从存储服务器接收 330 信息,包括所请求的文件数据、描述在存储服务器上的文件数据的存储的附加元数据、和 / 或代表了文件数据的至少一部分的数据。

[0056] 接收 330 文件数据信息可包括以下各项中的至少一项:文件数据、描述在存储服务器上的文件数据的存储的附加元数据、和 / 或代表了文件数据的至少一部分的数据。如在客户机处的传统应用可能期待的,信息可包括标准格式的文件数据。该信息可包括描述在存储服务器上的文件数据的存储的信息。该信息可包括代表了文件数据的至少一部分的

数据。

[0057] 访问 310 描述文件数据的存储的元数据可包括将对于描述文件数据的存储的信息的请求发送至服务器。这样的请求可以形式为文件系统扩展,其能使客户机做出多于文件系统(或网络文件系统)的调用来请求文件、文件数据、或数据部分如何被存储在数据存储器上的细节。

[0058] 访问 310 描述文件数据的存储的元数据可,可选地,包括为描述文件数据的存储的信息访问本地存储器。本地存储器中的信息可以是,之前响应于之前的请求从文件服务器中已经被接收的,或可被本地缓存为正在进行的一系列文件系统进程的一部分。访问 310 描述文件数据的存储的元数据可包括文件系统调用(通过常规文件系统 API 的扩展所引入),其返回细节,将在数据存储器服务器上的文件数据的存储形式或如何本地地存储本地缓存的副本,展示给客户机。

[0059] 例如,描述在数据存储器服务器上的文件数据的存储的元数据可包括描述源自对于数据存储器服务器上的文件数据去重复的文件数据的存储的数据。该元数据可包括构成文件的组块的组块列表,且可包括构成文件的每一个组块的加密散列的散列列表。然后客户机可使用所返回的组块列表或散列列表来建立对于要被传输的一个或多个组块的请求,或可使用散列列表来与已经接收到或本地地缓存的组块列表相比较,从而确定是否需要从数据存储器请求任何组块。

[0060] 例如,当下载文件时,客户机可从文件服务器请求散列列表,且可为期望的数据来询查对等客户机和/或询查对等文件服务器。客户机可接收 330 包括散列列表的信息作为对于询查的响应。散列列表可代表在数据存储器上所存储的数据,且客户机能仅请求其所需要的数据部分(如,组块)。当对等机具有期望数据且对等机与客户机之间的数据传输的传输成本或延时低于客户机与数据存储器之间的传输成本或延时时,还可从对等机读取数据。

[0061] 描述在数据存储器服务器上的文件数据的存储的元数据还可包括描述文件数据的经压缩的子集的数据或描述文件数据的经压缩的版本的版本的数据。使用这个信息,客户机可建立对于经压缩的文件数据子集的请求或者建立对于文件数据的经压缩的版本的请求。这可提供效率在于,响应于对于文件数据的请求,在传输数据前,数据存储器不需要解压缩文件数据或文件数据的子集。

[0062] 在一个实施例中,客户机可发送 320 对于文件数据的请求,该请求可包括对于整个文件的请求或对于文件的一部分的请求。例如,对于文件的请求, `get(fileX)`,或对于文件的一部分的请求, `getFileBytes(fileX;bytes=100 - 1000)`,可通过文件系统被发送至数据存储器服务器。作为响应,数据存储器服务器可通过不发送文件或文件部分,而是包含所请求的文件或文件部分的可能不同形式的数据,以进行响应。

[0063] 例如,数据存储器服务器可返回包含完全覆盖了所请求的文件或所请求的文件的一部分的经压缩的组块范围的文件数据。此外,数据存储器服务器可与组块一起返回文件存储元数据,其标识了包含所请求的数据(和可能比请求的更多的数据)的返回的组块。

[0064] 此外,如果所返回的组块是被压缩的,数据存储器服务器可返回文件存储元数据,其标示所返回的数据(或数据组块)是被压缩的且可标识使用何种压缩技术或算法来压缩数据或需要使用何种压缩技术或算法来压缩数据。如可理解的,在没有也返回标识特定压缩

或解压缩结束的元数据的而返回经压缩的数据和 / 或经压缩的阻抗的情况下,可假设存在默认压缩或解压缩技术。

[0065] 然后客户机可从数据存储服务器接收 330 这个数据和 / 或元数据,且在客户机侧执行合适的解压缩和 / 或块组装来重建所请求的数据。如可理解的,此举,相比在传输至客户机和 / 或由客户机接收前,使数据存储器解压缩和 / 或组装实际由客户机所请求的特定数据,由于数据传输成本或传输延时的原因,是更为有效率的。

[0066] 文件存储元数据可包括块或经压缩的块的加密散列列表、和与那些块包含文件数据的哪些部分相关的标识。通过使用块或经压缩的块的加密散列列表、和哪些块包含文件数据的哪些部分相关的标识,客户机可合适地解压缩经压缩的数据和 / 或重新组装块,其包含了客户机所期望或所请求的数据范围的全部或更多。

[0067] 图 2 示出了文件存储与传输的结合的方法的示例性体系架构。客户机与服务器 210 可包括知晓优化应用和 / 或服务。客户机与服务器可与文件系统接口 250 通信,接口可包括文件系统应用编程接口(API)且还可包括优化 API。文件系统 API 可包括常规文件系统和 / 或网络文件系统的所有常规调用和函数。优化 API 包括扩展的 API 元素(如,函数调用和接口),其展示了存储在数据存储器上的数据 260、270、和 280 的存储细节。

[0068] 文件系统接口 250 能使客户机请求描述在数据存储服务器上的文件数据的存储的元数据。文件系统接口 250 还能使客户机以各种格式请求来自数据存储服务器的数据。客户机可使用常规文件系统 API (如,标准或传统文件系统 API)请求数据,从而以其原始或原本格式获得完整的文件。客户机还可使用优化 API 请求数据,从而仅请求文件的特定块、如被存储在服务器上的文件的压缩形式,且可请求如被存储在服务器上的文件的经压缩的块。

[0069] 通过对于文件系统 API (保持传统文件系统 API 的所有功能)做出调用,不知晓增强的和 / 或扩展的文件系统接口 250 的客户机、应用、和服务 220 仍可常规地、不变地、且不受妨碍地操作。

[0070] 知晓优化的客户机、应用、和服务 230 可对于优化 API 做出调用来调用此处描述的实施例的所有功能。知晓优化的客户机、应用、和服务可从数据存储器或服务器请求散列列表、块列表、经压缩的数据等。例如,文件 foo.vhd 260 可被存储在数据存储器上作为指向块存储 / 索引 270 的块列表。块存储 / 索引可包括块(如,块 160C1-Cn),可包括经压缩的块(如,块 180CH1-CHn),且可包括对于能对于文件和数据存储去重复并优化的所存储的块的引用、指针、和索引。

[0071] 客户机可通过优化 API 请求描述 foo.vhd 的存储的元数据,且从数据存储器接收描述如何存储 foo.vhd 的元数据。一旦客户机访问了元数据,其可通过优化 API 向存储服务发出对于文件数据的请求。请求可以是对于其原本格式的整个文件的请求,或,请求可以是仅对于如被存储在块存储 / 索引 270 中的文件的一个或多个块或经压缩的块的请求。

[0072] 然后,客户机可从数据存储服务器接收信息,信息包括以下各项中的一个或多个:文件数据、描述在数据存储服务器上的文件数据的存储的附加元数据、和代表文件数据的至少一部分的数据。客户机可以其原本格式接收整个文件。客户机可接收被压缩在数据存储器中的整个文件。客户机可接收文件的块。客户机可接收文件的经压缩的块。客户

机可接收描述在数据存储服务器上的文件数据的存储的附加元数据、且可接收包括文件数据的一部分的数据。客户机所接收到的响应可对应于通过扩展的优化 API（能使知晓数据存储服务器中的数据存储的细节的客户机和应用做出请求）所做出的请求。

[0073] 在另一个示例中，文件 bar.doc 已经由优化服务 240 压缩、组块化、和去重复，且被存储为组块存储 / 索引 270 中的指针。在此处的实施例中，客户机可请求描述在数据存储服务器上的 bar.doc 的存储的元数据，且在接收描述在数据存储服务器上的 bar.doc 的存储的元数据后，对于存储在组块存储 / 索引 270 中的一个或多个经压缩的 bar.doc 的组块发送请求。由于客户机请求经压缩的组块，数据存储服务器不需要对于 bar.doc 的组块解压缩，数据存储服务器也不需要重新组装 bar.doc 的组块来响应于来自客户机的对于 bar.doc 的请求。

[0074] 在另一个实施例中，提供了用于将数据存储服务器中的存储优化的细节展示给客户机的方法。这个方法包括，发送描述在数据存储服务器上的文件数据的存储的元数据，其中文件数据被以与文件数据的原本形式不同的形式存储在数据存储服务器上，且其中元数据展示了存储在数据存储服务器上的文件数据的存储形式。该方法还包括在数据存储服务器处，接收来自计算系统的，对于文件数据的请求。该方法还包括从数据存储服务器发送信息，信息包括以下各项中的至少一个：文件数据、描述在数据存储服务器上的文件数据的存储的附加元数据、和代表文件数据的至少一部分的数据。

[0075] 如图 4 中所示，服务器或数据存储服务器可发送 410 描述在数据存储服务器或数据存储服务器上的文件数据的存储的元数据。文件数据被以与文件数据的原本形式不同的形式存储在数据存储服务器上。例如，文件数据可以组块化格式、压缩格式、或压缩和组块化格式的组合，被存储在存储服务器上。

[0076] 被发送的元数据提供信息，该信息展示了在数据存储服务器上所存储的文件数据的存储格式。例如，元数据可包括展示文件数据被存储为组块化格式、压缩格式、或压缩和组块化格式的组合的信息。元数据可包括信息，该信息包括构成被存储在数据存储服务器上的文件数据的组块的散列列表。存储在数据存储服务器上的组块可以是，源自被存储在存储服务器上的文件数据（以及其他文件数据）的去重复的组块。

[0077] 元数据可包括信息，该信息包括文件数据的子集的加密散列。数据子集的加密散列可由客户机、传输设备、或另一个数据存储服务器使用来标识组块是否与另一个组块相同。通过使用文件数据子集的加密散列，使得客户机、传输设备、和其他数据存储服务器能确定特定数据子集是否本地地可用或可以较低的延时或传输成本从源处获得。通过标识相同的数据子集，可确定特定数据子集是否需要被请求或传输。

[0078] 文件数据的子集可以是整个文件或文件数据。数据子集还可以是已经被数据存储服务器组块化作为存储优化或去重复方案的一部分的文件数据的一个或多个组块。

[0079] 描述在数据存储服务器或数据存储服务器上的文件数据的存储的元数据还可包括描述文件数据的一些或全部被压缩在数据存储服务器或数据存储服务器上的数据。元数据可包括文件数据的被组块化的格式的一个或多个组块已经被压缩的信息。通过使用表示文件数据的某个部分被压缩的信息，客户机可请求文件或文件的一个或多个组块，在对于客户机的响应中，以被存储于数据存储服务器中的被组块化或被压缩的格式，被返回。通过请求文件的特定组块或被压缩的组块，由于在将文件或文件的组块传输至发出请求的客户机前，数据存储服务器不需要解压缩文件或文件的组块吗，从而减少了开销。

[0080] 图 4 还示出了从计算系统接收 410 对于文件数据的请求。可从客户机、从另一个存储服务器、从在远程计算系统上执行的应用、等,接收请求。通过使用对应于扩展和 / 或增强标准网络文件系统 API 的优化 API 的协议来格式化请求。

[0081] 对于文件数据的请求可包括标识被请求的文件的特定组块的信息。该请求还可包括标识所请求的文件数据应该被以经压缩或未压缩格式发送的信息。该请求可包括信息,仅文件的组块的子集应该被发送而其他组块已经本地可获得。

[0082] 图 4 还示出发送 430 文件数据信息,其包括以下各项中的至少一项:文件数据、描述在存储服务器上的文件数据的存储的附加元数据、和 / 或代表了文件数据的至少一部分的数据。文件数据信息的发送 430 可响应于对于所接收 420 到的对文件数据的请求。如上所述,对于文件数据的请求可以是对于被存储在数据存储器上作为组块、经压缩的格式、或任意组合,的文件数据的请求。

[0083] 发送 430 文件数据信息可包括以下各项中的至少一项:文件数据、描述在存储服务器上的文件数据的存储的附加元数据、和 / 或代表了文件数据的至少一部分的数据。由于在客户机处的传统应用可期待的,信息可包括标准格式的文件数据。该信息可包括描述在存储服务器上的文件数据的存储的信息。该信息可包括代表了文件的至少一部分的数据。

[0084] 所接收到的请求可具有客户机所期望的数据的被标识的特定组块。响应于这个请求,数据存储器可将所请求的数据的组块发送至发出请求的客户机。所接收到的请求可具有客户机所期望的数据的被标识的经压缩的子集。响应于这个请求,数据存储器可将所请求的数据的经压缩的子集发送至发出请求的客户机。所接收到的请求可具有标识客户机所期望的数据的组块的被标识的特定加密散列。响应于这个请求,数据存储器可将由加密散列所标识的数据的特定组块发送至发出请求的客户机。

[0085] 在一个实施例中,数据存储器可接收 420 对于文件或文件部分的请求。例如,数据存储器可接收对于文件的请求 `get(fileX)` 或可接收对于文件部分的请求 `getFileBytes(fileX;bytes=100-1000)`。数据存储器可构建对应请求的响应并发送文件数据信息,该信息包括被存储在数据存储器上的文件数据且包括标识所存储的文件数据的存储细节的元数据。例如,数据存储器可返回一组组块和标志哪些组块包括所请求的数据的哪些部分的元数据。此外,数据存储器可返回元数据,包括恰当的压缩和 / 或解压缩信息,从而对于以经压缩格式返回的数据进行解压缩。

[0086] 在一些实施例中,在不执行发送元数据 410 的之前的步骤的情况下,请求可被接收 420 且文件数据信息可被发送 430。例如,知晓优化的客户机可仅请求文件数据,数据存储器可接收请求 420,且数据存储器可组成响应并将该响应发送至客户机,假设客户机在需要时可合适地处理所返回的文件数据和 / 或元数据且合适地重新组装组块和 / 或解压缩数据。

[0087] 实施例还提供了对于数据的存储和传输的写路径优化的支持。例如,具有对于文件本地修改的客户机可生成所修改的文件的散列列表表示。然后该散列列表可被传输至数据存储服务器。然后该数据存储服务器可将所接收到的代表经修改的文件的散列列表与维持在数据存储服务器上的标识了存储在数据存储服务器上的文件组块的全面的散列列表比较。

[0088] 基于这个比较,数据存储服务器然后可将已经存储在数据存储服务器上的组块列表返回至客户机。数据存储服务器还可将没有存储在数据存储服务器上的组块列表返回至客户机。基于所返回的在数据存储服务器上存储的组块列表(或未存储的组块列表),然后客户机可将没有被存储在数据存储服务器上的这些组块传输至数据存储服务器。

[0089] 接收到代表经修改的文件的散列列表、且接收到没有被存储在数据存储服务器上的经修改的文件的组块,数据存储服务器现在存储完整的经修改的文件(包括已经被存储在服务器上的一些组块、服务器新接收到的一些组块、和代表完整的经修改的文件的散列列表(或组块列表))。通过传输代表完整文件的散列列表(或组块列表)以及仅传输没有被存储在数据存储服务器上的那些组块,可实现从客户机到数据存储器的数据传输的优化。

[0090] 例如,数据存储服务器可从客户机处接收散列列表且将代表文件的被传输的散列列表与存储在组块存储/索引 270 中的散列列表(包括存储在数据存储服务器上的组块和存储在数据存储服务器上的组块的加密散列的索引)进行比较。然后数据存储器将代表了没有被存储在组块存储和索引 270 中的组块的散列列表返回至客户机。然后,客户机可将没有被存储在组块存储中的组块传输至数据存储器。然后,数据存储器将所接收到的组块与代表整个被修改的文件的散列列表仪器存储在组块存储 270 中。以此方式,数据存储服务器现在可存储经修改的文件的整个代表(形式为代表文件的组块列表和对应的组块),而不需要客户机传输构成文件的所有组块。

[0091] 在另一个示例中,包括五个组块,组块 C1-C5,的文件,可由客户机仅在组块 C4 中修改(导致了经修改的组块 Cm4)。客户机可将代表组块 C1-C3, Cm4, 和 C5 的散列列表发送至数据存储服务器。该散列列表现在代表了整个经修改的文件。然后,数据存储服务器可响应于客户机,表示已经将组块 C1-C3、和 C5 存储在服务器上,但是没有组块 Cm4。然后,客户机可将组块 Cm4 发送至数据存储服务器。然后数据存储服务器可将 Cm4 存储在数据存储服务器上,还有所接收到的代表了组块 C1 - C3, Cm4, 和 C5 的散列列表,且已经存储了组块 C1 - C3, 和 C5,现在将完整的经修改的文件存储在数据存储器上。

[0092] 如可理解的,对于新创建的文件以及对于经修改的文件,以类似方式,能允许这个写路径实施例。客户机可创建任何文件(无论是经修改的文件或新创建的文件)的组块列表,并将组块列表发送至数据存储服务器,从而数据存储服务器可将所接收到的组块列表与已经被存储在服务器上的组块列表进行比较。此外,组块列表可以是唯一地标识构成文件的每一个组块的加密散列列表。组块,其本身,如此处所讨论地,可以是经压缩的组块、原始数据格式的组块,或甚至是已经以某种方式(加密或其他)被改变的组块。

[0093] 当被传输时,组块可以原始数据格式、经压缩的格式、或其他格式被传输。如可理解的,当文件数据部分以压缩的格式被传输时,可引起这样的优化:传输体系结构不需要压缩数据来获得传输中的效率且数据存储器不需要压缩数据来优化数据存储服务器上的存储。通过仅传输未被存储在或呈现在传输的接收端上的那些经压缩的组块,在文件数据的传输和存储方面均实现优化。

[0094] 本发明可具体化为其它具体形式而不背离其精神或本质特征。所描述的实施例在所有方面都应被认为仅是说明性而非限制性的。因此,本发明的范围由所附权利要求书而非前述描述指示。落入权利要求书的等效方案的含义和范围内的所有改变被权利要求书的范围所涵盖。

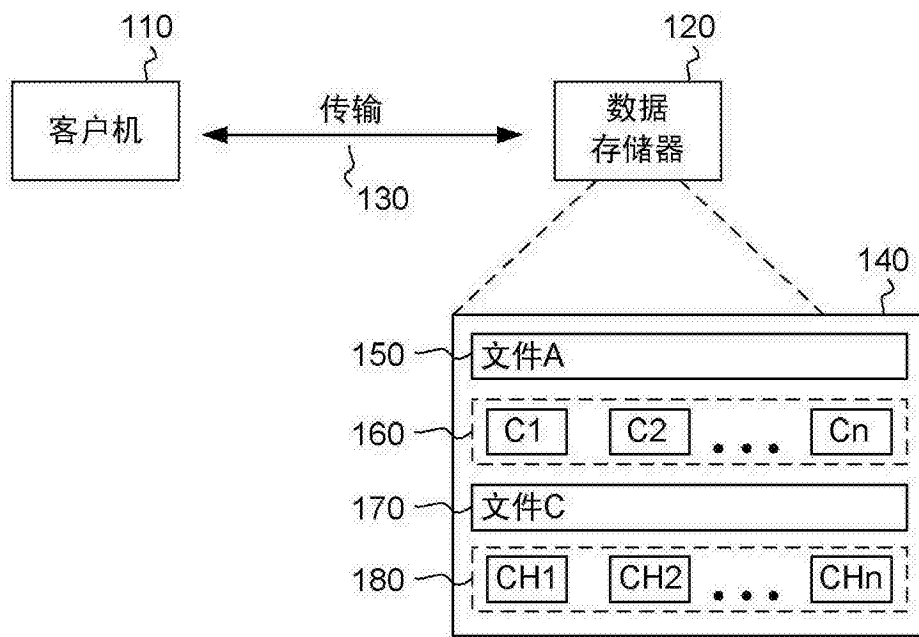


图 1

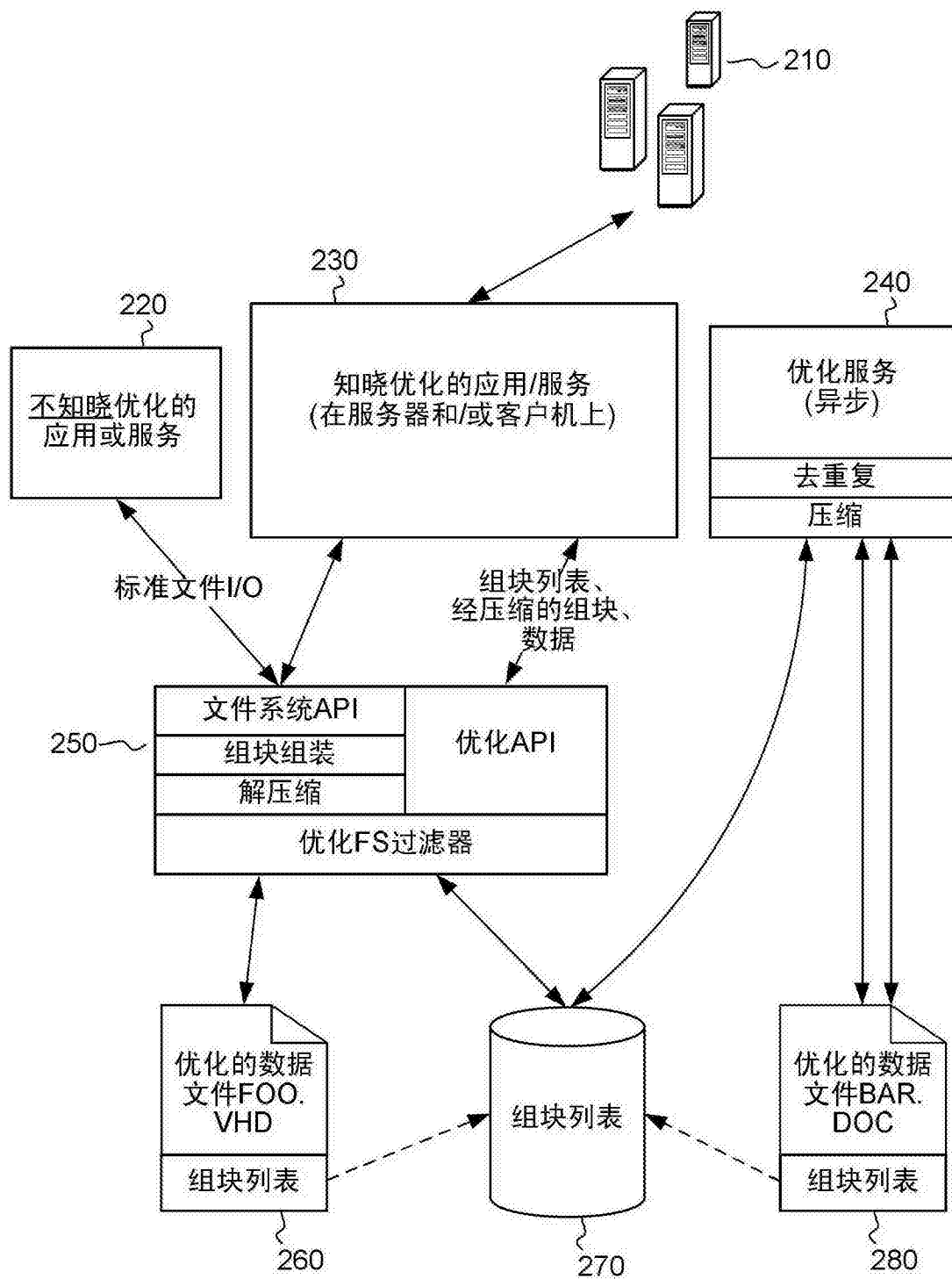


图 2

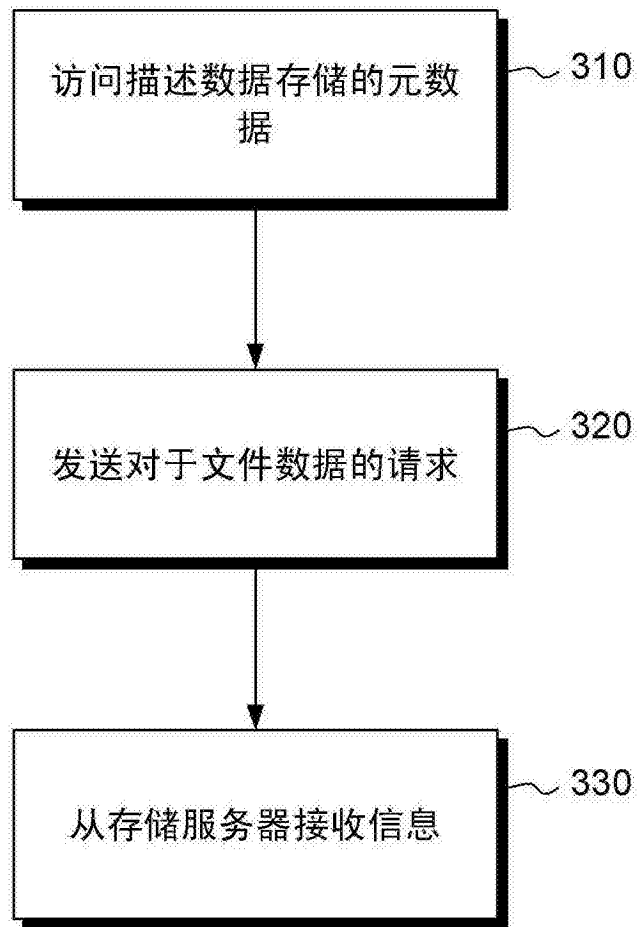


图 3

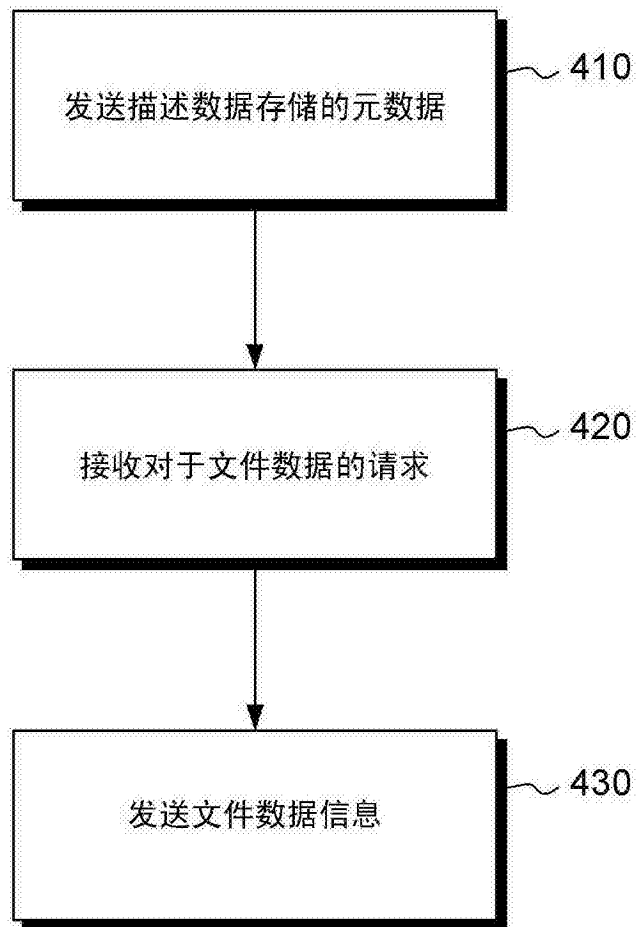


图 4