

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 9/54 (2006.01)



[12] 发明专利说明书

专利号 ZL 200580042164. X

[45] 授权公告日 2010 年 1 月 20 日

[11] 授权公告号 CN 100583048C

[22] 申请日 2005.11.17

[21] 申请号 200580042164. X

[30] 优先权

[32] 2004.12.10 [33] US [31] 11/009,937

[86] 国际申请 PCT/EP2005/056040 2005.11.17

[87] 国际公布 WO2006/061316 英 2006.6.15

[85] 进入国家阶段日期 2007.6.8

[73] 专利权人 国际商业机器公司

地址 美国纽约

[72] 发明人 M·J·安格林 A·H·霍赫贝格

J·维克斯涅

[56] 参考文献

EP0294499A1 1988.12.14

CN1531684A 2004.9.22

THE ALLOC STREAM FACILITY A RE-
DESIGN OF APPLICATION - LEVEL STREAM I/
O. Orran Krieger, Michael Stumm. COMPUTER
IEEE SERVICE CENTER, Vol. 27 No. 3. 1994

审查员 马雅凡

[74] 专利代理机构 北京市中咨律师事务所

代理人 于 静 张亚非

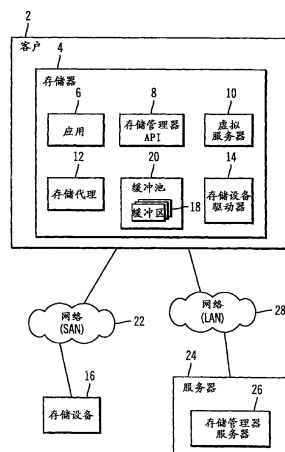
权利要求书 6 页 说明书 13 页 附图 5 页

[54] 发明名称

传送数据的方法和系统

[57] 摘要

提供了一种用于在共享缓冲区中在系统和存储器之间传送数据的方法、系统和程序。应用从组件请求缓冲区。该组件分配一缓冲区并将缓冲区中用于应用数据的第一偏移和缓冲区中用于第一首部的第二偏移返回该应用。该应用在缓冲区中的第一偏移处写入应用数据；并在缓冲区中的第二偏移处写入第一首部。第二首部被写入缓冲区的第三偏移。



1. 一种传送数据的方法，包括：

由应用从组件请求缓冲区；

由该组件分配缓冲区；

由该组件将在该缓冲区中用于应用数据的第一偏移和在该缓冲区中用于第一首部的第二偏移返回给该应用，所述第一偏移和第二偏移被计算为在缓冲区中提供足够量的空间；

由该应用在该缓冲区中的该第一偏移处写入应用数据；

由该应用在该缓冲区中的该第二偏移处写入第一首部；以及将第二首部写入该缓冲区中的第三偏移。

2. 根据权利要求 1 所述的方法，其中所述第一偏移比所述第二偏移大，并且所述第二偏移比所述第三偏移大。

3. 根据权利要求 1 所述的方法，其中所述组件将所述第二首部写入所述缓冲区，还包括：

将第三首部写入在该缓冲区中的第四偏移，所述第一偏移比所述第二偏移大，所述第二偏移比所述第三偏移大，而所述第三偏移比所述第四偏移大。

4. 根据权利要求 1 所述的方法，其中所述组件包含存储管理器 API，所述方法进一步包括：

由虚拟服务器替换所述第二首部为用于将缓冲区内容传送到存储设备的信息。

5. 根据权利要求 1 所述的方法，其中所述组件包括存储管理器 API，并且在所述第三偏移处的所述第二首部指示包含所述应用数据和首部的缓冲区内容将被传送到第一设备或第二设备，还包括：

由所述存储管理器 API 将指向所述第二首部的指针传递给至虚拟服务器；

由所述虚拟服务器处理所述第二首部以判定是否将缓冲区内内容发送到第一设备或第二设备；以及

响应于判定所述第二首部指示将缓冲区内内容发送到所述第一设备，将用于将缓冲区内内容传送到所述第一设备的设备信息写入所述第二首部，其中所述设备信息替换由所述存储管理器 API 写入的所述第二首部信息。

6. 根据权利要求 5 所述的方法，其中所述第一设备包括存储设备，进一步包括：

响应于判定缓冲区内内容将被发送到所述存储设备，由用于所述存储设备的存储设备驱动器使用所述第二首部中的所述设备信息，将缓冲区内内容从所述缓冲区传送到所述存储设备。

7. 根据权利要求 6 所述的方法，其中由所述存储管理器 API 传递的指针包括第一指针，还包括：

响应于判定缓冲区内内容将被发送到所述存储设备，由所述虚拟服务器发送指向缓冲区的第二指针至存储代理，其中所述存储代理将所述设备信息写入所述第二首部；以及

由所述存储代理将指向缓冲区的第三指针发送至所述存储设备驱动器，其中所述存储设备驱动器使用所述第三指针将缓冲区的內容写入所述存储设备中的存储介质。

8. 根据权利要求 5 所述的方法，其中所述第二设备包括存储服务器，还包括：

响应于判定所述第二首部指示将缓冲区内内容发送到所述存储服务器，由所述虚拟服务器将缓冲区内内容发送到所述存储服务器。

9. 根据权利要求 1 所述的方法，进一步包括：

由所述组件计算所述第一偏移，以在所述第一偏移之前的缓冲区内提供足够量的空间，以便在从所述第一偏移写入所述应用数据之后写入所述第一首部和第二首部；以及

由所述组件计算所述第二偏移，以在所述第二偏移之前的缓

冲区中提供足够量的空间，以便在写入所述应用数据和所述第一首部之后写入所述第二首部。

10. 根据权利要求 9 所述的方法，进一步包括：

由所述组件计算所述第三偏移以在所述第三偏移之前的缓冲区中提供足够量的空间，以便在写入所述应用数据和第一首部之后写入所述第二首部。

11. 根据权利要求 1 所述的方法，进一步包括：

判定所述应用数据是否是固定数的倍数，其中响应于判定所述应用数据不是该固定数的倍数，所述第一和第二偏移被选择在所述第一偏移处写入的所述应用数据之前维持空白空间，使得所述空白空间和应用数据占据该固定数的倍数，其中所述第一首部从所述第二偏移延伸到所述空白空间的起始处。

12. 根据权利要求 11 所述的方法，进一步包括：

判定所述第一首部是否是固定数的倍数，其中响应于判定所述第一首部不是固定数的倍数，所述第三偏移被选择在所述第一首部之前维持空白空间，使得所述第一首部和空白空间占据该固定数的倍数，其中所述第二首部从所述第三偏移延伸到在所述第一首部之前的所述空白空间的起始处。

13. 根据权利要求 11 所述的方法，其中所述固定数包括 4 千字节，且其中如果在所述第一偏移处写入缓冲区的所述应用数据没有对准在缓冲区中的 4 千字节边界上，则所述应用数据不是所述固定数的倍数。

14. 根据权利要求 1 所述的方法，进一步包括：

由所述应用请求多个缓冲区；

由所述组件将所述多个缓冲区返回到所述应用；

由所述应用将应用数据和第一首部并行地写入所述多个缓冲区。

15. 一种传送数据的系统，包括：

由应用从组件请求缓冲区的装置；

由该组件分配缓冲区的装置；

由该组件将在该缓冲区中用于应用数据的第一偏移和在该缓冲区中用于第一首部的第二偏移返回给该应用的装置，所述第一偏移和第二偏移被计算为在缓冲区中提供足够量的空间；

由该应用在该缓冲区中的该第一偏移处写入应用数据的装置；

由该应用在该缓冲区中的该第二偏移处写入第一首部的装置；以及

将第二首部写入该缓冲区中的第三偏移的装置。

16. 根据权利要求 15 所述的系统，其中所述第一偏移比所述第二偏移大，并且所述第二偏移比所述第三偏移大。

17. 根据权利要求 15 所述的系统，其中所述组件将所述第二首部写入所述缓冲区，所述系统还包括：

将第三首部写入在该缓冲区中的第四偏移的装置，所述第一偏移比所述第二偏移大，所述第二偏移比所述第三偏移大，而所述第三偏移比所述第四偏移大。

18. 根据权利要求 15 所述的系统，其中所述组件包含存储管理器 API，所述系统进一步包括：

由虚拟服务器替换所述第二首部为用于将缓冲区内容传送到存储设备的信息的装置。

19. 根据权利要求 15 所述的系统，其中所述组件包括存储管理器 API，并且在所述第三偏移处的所述第二首部指示包含所述应用数据和首部的缓冲区内容将被传送到第一设备或第二设备，所述系统还包括：

由所述存储管理器 API 将指向所述第二首部的指针传递给至虚拟服务器的装置；

由所述虚拟服务器处理所述第二首部以判定是否将缓冲区内

容发送到第一设备或第二设备的装置；以及

响应于判定所述第二首部指示将缓冲区内容发送到所述第一设备，将用于将缓冲区内容传送到所述第一设备的设备信息写入所述第二首部的装置，其中所述设备信息替换由所述存储管理器 API 写入的所述第二首部信息。

20. 根据权利要求 19 所述的系统，其中所述第一设备包括存储设备，所述系统进一步包括：

响应于判定缓冲区内容将被发送到所述存储设备，由用于所述存储设备的存储设备驱动器使用所述第二首部中的所述设备信息，将缓冲区内容从所述缓冲区传送到所述存储设备的装置。

21. 根据权利要求 20 所述的系统，其中由所述存储管理器 API 传递的指针包括第一指针，所述系统还包括：

响应于判定缓冲区内容将被发送到所述存储设备，由所述虚拟服务器发送指向缓冲区的第二指针至存储代理的装置，其中所述存储代理将所述设备信息写入所述第二首部；以及

由所述存储代理将指向缓冲区的第三指针发送至所述存储设备驱动器的装置，其中所述存储设备驱动器使用所述第三指针将缓冲区的内容写入所述存储设备中的存储介质。

22. 根据权利要求 19 所述的系统，其中所述第二设备包括存储服务器，所述系统还包括：

响应于判定所述第二首部指示将缓冲区内容发送到所述存储服务器的装置，由所述虚拟服务器将缓冲区内容发送到所述存储服务器。

23. 根据权利要求 15 所述的系统，进一步包括：

由所述组件计算所述第一偏移，以在所述第一偏移之前的缓冲区中提供足够量的空间，以便在从所述第一偏移写入所述应用数据之后写入所述第一首部和第二首部的装置；以及

由所述组件计算所述第二偏移，以在所述第二偏移之前的缓

冲区中提供足够量的空间，以便在写入所述应用数据和所述第一首部之后写入所述第二首部的装置。

24. 根据权利要求 23 所述的系统，进一步包括：

由所述组件计算所述第三偏移以在所述第三偏移之前的缓冲区中提供足够量的空间，以便在写入所述应用数据和第一首部之后写入所述第二首部的装置。

25. 根据权利要求 15 所述的系统，进一步包括：

判定所述应用数据是否是固定数的倍数的装置，其中响应于判定所述应用数据不是该固定数的倍数，所述第一和第二偏移被选择以在所述第一偏移处写入的所述应用数据之前维持空白空间，使得所述空白空间和应用数据占据该固定数的倍数，其中所述第一首部从所述第二偏移延伸到所述空白空间的起始处。

26. 根据权利要求 25 所述的系统，进一步包括：

判定所述第一首部是否是固定数的倍数的装置，其中响应于判定所述第一首部不是固定数的倍数，所述第三偏移被选择以在所述第一首部之前维持空白空间，使得所述第一首部和空白空间占据该固定数的倍数，其中所述第二首部从所述第三偏移延伸到在所述第一首部之前的所述空白空间的起始处。

27. 根据权利要求 25 所述的系统，其中所述固定数包括 4 千字节，且其中如果在所述第一偏移处写入缓冲区的所述应用数据没有对准在缓冲区中的 4 千字节边界上，则所述应用数据不是所述固定数的倍数。

28. 根据权利要求 15 所述的系统，进一步包括：

由所述应用请求多个缓冲区的装置；

由所述组件将所述多个缓冲区返回到所述应用的装置；

由所述应用将应用数据和第一首部并行地写入所述多个缓冲区的装置。

传送数据的方法和系统

技术领域

本发明涉及一种用于在共享缓冲区中在系统和存储器之间传送数据的方法、系统和程序。

背景技术

在网络备份环境中，客户系统可以将数据备份在存储设备中并和备份服务器协调该备份。例如，国际商业机器公司（“IBM”®）Tivoli®存储管理器产品为客户和服务系统提供软件以备份客户数据（IBM 和 Tivoli 是 IBM 的注册商标）。客户机器可以包括多个程序组件，所述程序组件处理在客户的应用和存储设备之间传送的数据。每个组件从前一组件使用的缓冲区中读取数据并接着执行存储器复制操作以将数据复制到分配给该组件的缓冲区。作为当数据在组件分配的缓冲区中传递时处理数据的一部分，程序组件可以将首部信息加入该数据。程序组件可从前一组件的缓冲区读取数据，处理该数据，将新的首部加入该数据，接着将这个数据和新的首部写入分配到该组件的缓冲区。在处理数据的组件链中的下一程序组件可进行同样的处理。

发明内容

提供了一种在共享缓冲区中在系统和存储器之间传送数据的方法、系统和程序。在本发明的第一方面，应用从组件请求缓冲区。该组件分配一缓冲区并将在缓冲区中用于应用数据的第一偏移和在缓冲区中用于第一首部的第二偏移返回该应用。应用在缓冲区中的第一偏移处写入应用数据；并在缓冲区中的第二偏移处写入第一首部。第二首部被写入缓冲区的第三

偏移处。

根据本发明的第二方面，第一组件从第二组件请求缓冲区。第二组件分配一缓冲区并将在缓冲区中用于应用数据的第一偏移和在缓冲区中用于第一首部的第二偏移返回该应用。第二组件在缓冲区中的第一偏移处写入应用数据并在缓冲区中的第二偏移处写入第一首部。第二首部被写入缓冲区的第三偏移处。

本发明提供了一种传送数据的方法，包括：由应用从组件请求缓冲区；由该组件分配缓冲区；由该组件将在该缓冲区中用于应用数据的第一偏移和在该缓冲区中用于第一首部的第二偏移返回给该应用，所述第一偏移和第二偏移被计算为在缓冲区中提供足够量的空间；由该应用在该缓冲区中的该第一偏移处写入应用数据；由该应用在该缓冲区中的该第二偏移处写入第一首部；以及将第二首部写入该缓冲区中的第三偏移。

按照本发明的一个方面，上述方法还包括：由应用从组件请求缓冲区；由该组件分配缓冲区；由该组件将在该缓冲区中用于应用数据的第一偏移和在该缓冲区中用于第一首部的第二偏移返回给该应用，所述第一偏移和第二偏移被计算为在缓冲区中提供足够量的空间；由该应用在该缓冲区中的该第一偏移处写入应用数据；由该应用在该缓冲区中的该第二偏移处写入第一首部；以及将第二首部写入该缓冲区中的第三偏移。

按照本发明的另一个方面，上述方法中所述第一偏移比所述第二偏移大，并且所述第二偏移比所述第三偏移大。

按照本发明的另一个方面，上述方法中所述组件将所述第二首部写入所述缓冲区，还包括：将第三首部写入在该缓冲区中的第四偏移，所述第一偏移比所述第二偏移大，所述第二偏移比所述第三偏移大，而所述第三偏移比所述第四偏移大。

按照本发明的另一个方面，上述方法中所述组件包含存储管理器 API，所述方法进一步包括：由虚拟服务器替换所述第二首

部为用于将缓冲区内容传送到存储设备的信息。

按照本发明的另一个方面，上述方法中所述组件包括存储管理器 API，并且在所述第三偏移处的所述第二首部指示包含所述应用数据和首部的缓冲区内容将被传送到第一设备或第二设备，还包括：由所述存储管理器 API 将指向所述第二首部的指针传递给至虚拟服务器；由所述虚拟服务器处理所述第二首部以判定是否将缓冲区内容发送到第一设备或第二设备；以及响应于判定所述第二首部指示将缓冲区内容发送到所述第一设备，将用于将缓冲区内容传送到所述第一设备的设备信息写入所述第二首部，其中所述设备信息替换由所述存储管理器 API 写入的所述第二首部信息。

按照本发明的另一个方面，上述方法中所述第一设备包括存储设备，进一步包括：响应于判定缓冲区内容将被发送到所述存储设备，由用于所述存储设备的存储设备驱动器使用所述第二首部中的所述设备信息，将缓冲区内容从所述缓冲区传送到所述存储设备。

按照本发明的另一个方面，上述方法中由所述存储管理器 API 传递的指针包括第一指针，还包括：响应于判定缓冲区内容将被发送到所述存储设备，由所述虚拟服务器发送指向缓冲区的第二指针至存储代理，其中所述存储代理将所述设备信息写入所述第二首部；以及由所述存储代理将指向缓冲区的第三指针发送至所述存储设备驱动器，其中所述存储设备驱动器使用所述第三指针将缓冲区的内容写入所述存储设备中的存储介质。

按照本发明的另一个方面，上述方法中所述第二设备包括存储服务器，还包括：响应于判定所述第二首部指示将缓冲区内容发送到所述存储服务器，由所述虚拟服务器将缓冲区内容发送到所述存储服务器。

按照本发明的另一个方面，上述方法进一步包括：由所述组

件计算所述第一偏移，以在所述第一偏移之前的缓冲区中提供足够量的空间，以便在从所述第一偏移写入所述应用数据之后写入所述第一首部和第二首部；以及由所述组件计算所述第二偏移，以在所述第二偏移之前的缓冲区中提供足够量的空间，以便在写入所述应用数据和所述第一首部之后写入所述第二首部。

按照本发明的另一个方面，上述方法进一步包括：由所述组件计算所述第三偏移以在所述第三偏移之前的缓冲区中提供足够量的空间，以便在写入所述应用数据和第一首部之后写入所述第二首部。

按照本发明的另一个方面，上述方法进一步包括：判定所述应用数据是否是固定数的倍数，其中响应于判定所述应用数据不是该固定数的倍数，所述第一和第二偏移被选择以在所述第一偏移处写入的所述应用数据之前维持空白空间，使得所述空白空间和应用数据占据该固定数的倍数，其中所述第一首部从所述第二偏移延伸到所述空白空间的起始处。

按照本发明的另一个方面，上述方法进一步包括：判定所述第一首部是否是固定数的倍数，其中响应于判定所述第一首部不是固定数的倍数，所述第三偏移被选择以在所述第一首部之前维持空白空间，使得所述第一首部和空白空间占据该固定数的倍数，其中所述第二首部从所述第三偏移延伸到在所述第一首部之前的所述空白空间的起始处。

按照本发明的另一个方面，上述方法中所述固定数包括 4 千字节，且其中如果在所述第一偏移处写入缓冲区的所述应用数据没有对准在缓冲区中的 4 千字节边界上，则所述应用数据不是所述固定数的倍数。

按照本发明的另一个方面，上述方法进一步包括：由所述应用请求多个缓冲区；由所述组件将所述多个缓冲区返回到所述应用；由所述应用将应用数据和第一首部并行地写入所述多个缓冲

区。本发明还提供了一种传送数据的系统，包括：由应用从组件请求缓冲区的装置；由该组件分配缓冲区的装置；由该组件将在该缓冲区中用于应用数据的第一偏移和在该缓冲区中用于第一首部的第二偏移返回给该应用的装置，所述第一偏移和第二偏移被计算为在缓冲区中提供足够量的空间；由该应用在该缓冲区中的该第一偏移处写入应用数据的装置；由该应用在该缓冲区中的该第二偏移处写入第一首部的装置；以及将第二首部写入该缓冲区中的第三偏移的装置。

按照本发明的一个方面，上述系统还包括：中所述第一偏移比所述第二偏移大，并且所述第二偏移比所述第三偏移大。

按照本发明的一个方面，上述系统中所述组件将所述第二首部写入所述缓冲区，所述系统还包括：将第三首部写入在该缓冲区中的第四偏移的装置，所述第一偏移比所述第二偏移大，所述第二偏移比所述第三偏移大，而所述第三偏移比所述第四偏移大。

按照本发明的一个方面，上述系统中所述组件包含存储管理器 API，所述系统进一步包括：由虚拟服务器替换所述第二首部为用于将缓冲区内容传送到存储设备的信息的装置。

按照本发明的一个方面，上述系统中所述组件包括存储管理器 API，并且在所述第三偏移处的所述第二首部指示包含所述应用数据和首部的缓冲区内容将被传送到第一设备或第二设备，所述系统还包括：由所述存储管理器 API 将指向所述第二首部的指针传递给至虚拟服务器的装置；由所述虚拟服务器处理所述第二首部以判定是否将缓冲区内容发送到第一设备或第二设备的装置；以及响应于判定所述第二首部指示将缓冲区内容发送到所述第一设备，将用于将缓冲区内容传送到所述第一设备的设备信息写入所述第二首部的装置，其中所述设备信息替换由所述存储管理器 API 写入的所述第二首部信息。

按照本发明的一个方面，上述系统中所述第一设备包括存储

设备, 所述系统进一步包括: 响应于判定缓冲区内容将被发送到所述存储设备, 由用于所述存储设备的存储设备驱动器使用所述第二首部中的所述设备信息, 将缓冲区内容从所述缓冲区传送到所述存储设备的装置。

按照本发明的一个方面, 上述系统中由所述存储管理器 API 传递的指针包括第一指针, 所述系统还包括: 响应于判定缓冲区内容将被发送到所述存储设备, 由所述虚拟服务器发送指向缓冲区的第二指针至存储代理的装置, 其中所述存储代理将所述设备信息写入所述第二首部; 以及由所述存储代理将指向缓冲区的第三指针发送至所述存储设备驱动器的装置, 其中所述存储设备驱动器使用所述第三指针将缓冲区的内容写入所述存储设备中的存储介质。

按照本发明的一个方面, 上述系统中所述第二设备包括存储服务器, 所述系统还包括: 响应于判定所述第二首部指示将缓冲区内容发送到所述存储服务器的装置, 由所述虚拟服务器将缓冲区内容发送到所述存储服务器。

按照本发明的一个方面, 上述系统进一步包括: 由所述组件计算所述第一偏移, 以在所述第一偏移之前的缓冲区中提供足够量的空间, 以便在从所述第一偏移写入所述应用数据之后写入所述第一首部和第二首部的装置; 以及由所述组件计算所述第二偏移, 以在所述第二偏移之前的缓冲区中提供足够量的空间, 以便在写入所述应用数据和所述第一首部之后写入所述第二首部的装置。

按照本发明的一个方面, 上述系统进一步包括: 由所述组件计算所述第三偏移以在所述第三偏移之前的缓冲区中提供足够量的空间, 以便在写入所述应用数据和第一首部之后写入所述第二首部的装置。

按照本发明的一个方面, 上述系统进一步包括: 判定所述应

用数据是否是固定数的倍数的装置，其中响应于判定所述应用数据不是该固定数的倍数，所述第一和第二偏移被选择以在所述第一偏移处写入的所述应用数据之前维持空白空间，使得所述空白空间和应用数据占据该固定数的倍数，其中所述第一首部从所述第二偏移延伸到所述空白空间的起始处。

按照本发明的一个方面，上述系统进一步包括：判定所述第一首部是否是固定数的倍数的装置，其中响应于判定所述第一首部不是固定数的倍数，所述第三偏移被选择以在所述第一首部之前维持空白空间，使得所述第一首部和空白空间占据该固定数的倍数，其中所述第二首部从所述第三偏移延伸到在所述第一首部之前的所述空白空间的起始处。

按照本发明的一个方面，上述系统中所述固定数包括 4 千字节，且其中如果在所述第一偏移处写入缓冲区的所述应用数据没有对准在缓冲区中的 4 千字节边界上，则所述应用数据不是所述固定数的倍数。

按照本发明的一个方面，上述系统进一步包括：由所述应用请求多个缓冲区的装置；由所述组件将所述多个缓冲区返回到所述应用的装置；由所述应用将应用数据和第一首部并行地写入所述多个缓冲区的装置。

附图说明

现在将仅作为示例并参考下面的附图来描述本发明的实施例，其中：

图 1 例示了网络计算环境的一实施例。

图 2 例示了缓冲区格式的一实施例。

图 3、4 和 5 例示了将数据从应用传送到存储设备的操作的一实施例。

具体实施方式

在下列的描述中将参考附图，所述附图形成本说明书的一部分并例示

了本发明的一些实施例。可以理解的是，可以利用其他的实施例并可以对结构和操作进行变化而不背离本发明的范围。

图 1 例示了具有包括存储器 4 的存储客户 2 的网络计算环境。存储客户 2 包括作为程序组件的应用 6 例如数据库应用等、存储管理器应用程序接口 (API) 8、虚拟服务器 10、存储代理 12 和提供写入存储设备 16 的接口的存储设备驱动器 14。存储管理器 API 8 管理在缓冲池 20 中的缓冲区 18。

在一个实施例中，存储设备驱动器 14 通过网络 22 例如存储区域网 (SAN) 传送数据以将其写入存储设备 16。在另一个可选实施例中，客户 2 可以通过总线接口，例如外设部件互连 (PCI) 总线、小型计算机系统接口 (SCSI) 总线等，连接到存储设备 16。

在一个实施例中，虚拟服务器 10 可以通过第二网络 28 例如局域网 (LAN) 将关于写入存储设备 16 的数据的信息传送至具有存储管理器服

务器 26 的服务器 24。这样，通过第一网络例如 SAN 传输去往和来自存储器的数据，而通过另一网络例如局域网发送关于被传送的数据的控制信息。在另一可选实施例中，数据以及用于数据传送操作的信息和命令可以在同一网络上传输，而不是如图 1 所示的在分开的网络上传输。服务器 26 在管理备份数据集的备份数据库顶端中维护关于存储在存储设备 16 中的数据的信息。

存储管理器 API 8 将缓冲池 20 中的缓冲区 18 分配给客户 2 中的组件以用于存储在应用 6 和存储设备 16 之间传送的数据。图 2 例示了组件 6、8、10、12 和 14 添加到缓冲区 18 的数据，包括应用数据 40 和由应用 6 写入的应用首部 44、由存储管理器 API 8 写入的 API 首部 46、指示如何处理存储管理器 API 8 写入的缓冲区的动词 (verb) 信息 48。在一个实施例中，存储代理 12 可以用设备特定信息 4 盖写动词信息 48，以使存储设备驱动器 14 能够将数据传送到存储设备 16。在描述的实施例中，例如组件 6、8、10、12 和 14 的组件将指向缓冲区内容包括该组件写入缓冲区的任何首部或者其他信息的指针传送到例如组件 8、10、12 和 14 的下一组件，以处理该缓冲区内容。这样，缓冲区 18 由组件共享，因此处理缓冲区中的数据并添加首部信息的每个组件不必分配在其中复制内容的另外的缓冲区。相反，每个组件可以处理数据并添加首部和其他信息给相同的共享缓冲区，因而不需要执行去往和来自分开的缓冲区的数据的存储器复制操作。

图 3、4 和 5 例示了例如组件 6、8、10、12 和 14 的组件执行的操作，以在组件之间传送数据以便传送到存储设备 16。参照图 3，应用 6 通过从存储管理器 API 8 请求（在块 102）的一个或多个缓冲区 18，来启动（在块 100）将应用数据传送到存储设备 16 的操作。该请求可以包括应用 6 打算写入的首部的最大长度，如果应用 6 将不写入首部，则该长度可以为零。作为响应，存储管理器 API 8 分配（在块 104）在缓冲池 20 中被指示为自由缓冲区 18 的一个或多个缓冲区 18 给应用。第一组件或存储管理器 API 8 计算（在块 106）用于应用数据 40 的第一偏移，以在第一偏移之前的缓冲区 18 中提供足够量的空间，以便在从第一偏移写入应用数据 40 之后写入

第一首部和第二首部，例如应用首部 42 和 API 首部 44。第一偏移可以在缓冲区 18 中提供用于由请求中指示的首部的长度的足够的空间。因此第一偏移寻址应用数据 40 在缓冲区 18 中的何处开始。第一组件 8 还计算（在块 108）第二偏移以在第二偏移之前的缓冲区中提供足够量的空间，以便在写入应用数据 40 和应用首部 42 之后，写入第二首部，例如 API 首部 44。这样，偏移被计算为允许一个组件在由前一组件写入的数据之前写入数据，由此允许组件在先前写入的应用数据 40 和首部之前添加首部。另外，在缓冲区 18 的顶端可以有数据结构，该数据结构具有关于所有的首部，例如首部 48、46、44 和数据 40 的信息，包括在缓冲区 18 中的偏移和每个首部的长度，以使得任何组件 6、8、10、12 和 14 能够确定首部和数据在缓冲区 18 中的位置。

在描述的实施例中，存储管理器 API 8 可能想要确保数据保持在固定的边界区域例如 4K 边界上。当应用 6 输入/输出（I/O）到存储设备 16 时，被缓冲区 18 传递到内核以进行输入/输出。为了实现直接存储器寻址（DMA），内核可将缓冲区 18 “钉”（pin）到物理存储器 4（而不是虚拟存储器），从而输入/输出设备（PCI 总线、SCSI 总线、SAN 等）能访问数据并将其发送到适当的设备。这种“钉”住缓冲区 18 防止了当输入/输出传送在进行中时（在总线上），其他的应用使用该数据。

如果（在块 110）指针应用数据 40 不是一固定数（例如，4 千字节（Kb））的倍数，那么第一组件（存储管理器 API 8）计算（在块 112）第一和第二偏移以在第一偏移处写入的应用数据之前维持空白空间（empty space），使得空白空间和应用数据占据固定数或 4Kb 的倍数。例如，第二偏移可以包含从第一偏移的偏移，其等于需要在 4Kb 边界上排列应用数据 40 的空白空间 42 的量和将在第二偏移开始写入的首部例如应用首部 44 的长度。第一组件 8 将第一和第二偏移以及用于所分配的缓冲区 18 的句柄（handle）返回到应用 6（在块 114）。

响应于所述偏移，应用 6 在缓冲区 18 中的第一偏移处写入（在块 116）应用数据 40，并在缓冲区 18 中的第二偏移处写入（在块 118）第一首部，

例如应用首部 44。应用首部 44 的写入可以是可选的。在写入之后，应用 6 将缓冲区句柄传递（在块 120）给第一组件（存储管理器 API 8）。

参照图 4，当接收（在块 130）来自应用 6 的缓冲区句柄时，第一组件（存储管理器 API 8）计算（在块 132）第三偏移，以在第三偏移之前的缓冲区中提供足够量的空间，以便在写入应用数据 40 和第一首部（应用首部 44）之后写入第二首部（API 首部 46）。在一个实施例中，如果（在块 134）第一首部（应用首部 44）是固定数例如 4Kb 的倍数，那么存储管理器 API 8 选择（在块 136）第三偏移以在第一首部之前维持空白空间，使得第一首部和空白空间占据固定数（4Kb）的倍数。第二首部（API 首部 46）被写入（在块 138）到缓冲区 18 中的第三偏移处。存储管理器 API 8 还将第三首部（动词首部 48）写入（在块 140）缓冲区中的第四偏移处，指示缓冲区内容是否将被传送到存储设备 16 或服务器 24。存储管理器 API 8 指示将打算用于存储设备 16 的数据传送到存储设备 16，并将备份命令和相关信息传送到管理客户 2 备份操作的服务器 24。存储管理器 API 8 接着将指向第三首部的指针传递（在块 142）至第二组件例如虚拟服务器 10。

响应于接收到指向第三首部的指针，虚拟服务器 10 判定（在块 144）第三首部例如动词首部 48 是否指示将缓冲区内容发送到存储设备 16 或服务器 24。如果（在块 144）该内容是打算用于服务器 24，那么虚拟服务器 10 通过网络 28 将缓冲区内容传输（在块 146）到服务器 24。否则，如果（在块 144）该内容是打算用于存储设备 16，那么虚拟服务器 10 将指向缓冲区 18 的第二指针发送（在块 148）至第三组件例如存储代理 12。

参照图 5，响应于接收到（在块 150）第二指针，第三组件（存储代理 12）将用于将缓冲区内容传送到第一设备的设备信息写入（在块 152）到第三首部。该设备特定信息替换第一组件（存储管理器 API 8）写入的第二首部信息。存储代理 12 接着将指向缓冲区 18 的第三指针发送（在块 154）至存储设备驱动器 14。作为响应，存储设备驱动器 14 使用第三指针访问缓冲区 18 内容和将其传送到存储设备 16 中的存储介质，例如磁存储介质（例如，硬盘驱动器、磁带等）、光存储介质或本领域已知的其他适当的

非易失性存储介质。

在其他实施例中，某些首部可以不被写入。例如，如果应用 6 或存储管理器 API 8 没有写入一个首部 44、46，那么第二首部会包含动词数据 48。或者，组件可以写入需要进一步的偏移计算的另外的首部。此外，任一被写入的首部之上的偏移可考虑在先前写入的首部的之上添加空白空间，以确保空白空间和写入的首部落在 4Kb 边界上。这样，在空白空间 42 之后的偏移处写入的首部 44 从该偏移延伸到空白空间 42 的起始处。

在一个实施例中，第一偏移比第二偏移大，第二偏移比第三偏移大，而第三偏移比第四偏移大。这样，处理数据的每个随后的组件可以向一偏移处的数据写入首部，以将该附加的首部放置在写入缓冲区的先前写入的应用数据和首部之上。

在另一实施例中，应用 6 可以从存储设备 16 请求数据。作为响应，存储设备驱动器 14 访问数据并将该数据返给存储代理 12。存储代理 12（或某其他组件）接着可以从存储管理器 API 8 请求缓冲区 18 以添加取回的数据。接着，组件可以执行偏移计算以提供偏移，以便在缓冲区 18 中提供空间以允许组件将对取回的应用数据的首部添加至缓冲区 18，从而将指向缓冲区的指针最终传递给应用 6 以访问数据。

在描述的实施例中，不同的组件共享缓冲区并从该相同缓冲区访问数据和将首部信息写入这个相同的缓冲区，以及将指向缓冲区的指针传递至另外的组件，使得另外的组件能访问该缓冲区。

在另外的实施例中，应用可以从存储管理器 API 8 请求多个缓冲区，接着将应用数据 40 和应用首部 44 数据并行地写入多个缓冲区，接着将对缓冲区的句柄传递给存储管理器 API 8 来进一步处理。

另外的实施例细节

描述的操作可以实现为使用标准编程和/或工程技术来产生软件、固件、硬件或其任何组合的方法、装置或者制造物品。这里使用的术语“制造物品”是指在硬件逻辑（例如，集成电路芯片、可编程门阵列（PGA）、

专用集成电路 (ASIC) 等) 或计算机可读介质中实现的代码或逻辑, 所述计算机可读介质例如为磁存储介质 (例如, 硬盘驱动器、软盘、磁带等)、光存储器 (CD-ROM、光盘等)、易失性和非易失性存储设备 (例如, EEPROM、ROM、PROM、RAM、DRAM、SRAM、固件、可编程逻辑等)。在计算机可读介质中的代码可以由处理器访问和执行。优选实施例在其中实现的编码还可以经过传输介质或通过网络从文件服务器访问。在这种情况下, 该代码在其中实现的制造物品可以包含传输介质, 例如网络传输线, 无线传输介质, 通过空间、无线电波、红外信号传播的信号等。因此, “制造物品” 可以包含其中包含该代码的介质。另外地, “制造物品” 可以包含其中包含、处理和执行该代码的硬件和软件组件的组合。当然, 本领域技术人员会认识到, 可以对该配置进行多种修改而不会背离本发明的范围, 而且 “制造物品” 可以包含本领域已知的任何信息承载介质。

图 3-5 例示的操作示出以某种次序发生的某些事件。在其他可选实施例中, 某些操作可以不同的次序执行、被修正或删除。而且, 步骤可以被添加到上述的逻辑而仍符合描述的实施例。此外, 在此描述的操作可以顺序发生, 或者某些操作可以并行处理。还有, 操作可以由单个处理单元或分布的处理单元执行。

为了例示和描述的目的, 给出了对本发明的各种实施例的以上描述。这种描述并非旨在是穷举的或将本发明限制于所公开的精确形式。根据上述示教可以进行多种修改和变化。本发明的范围旨在不是受到这种具体的描述的限制, 而是由所附的权利要求限制。上述的说明书、示例和数据提供了对本发明的组成的制造和使用的完整描述。因为可以实现本发明的许多实施例而不会背离本发明的精神和范围, 所以本发明在于随后所附的权利要求。

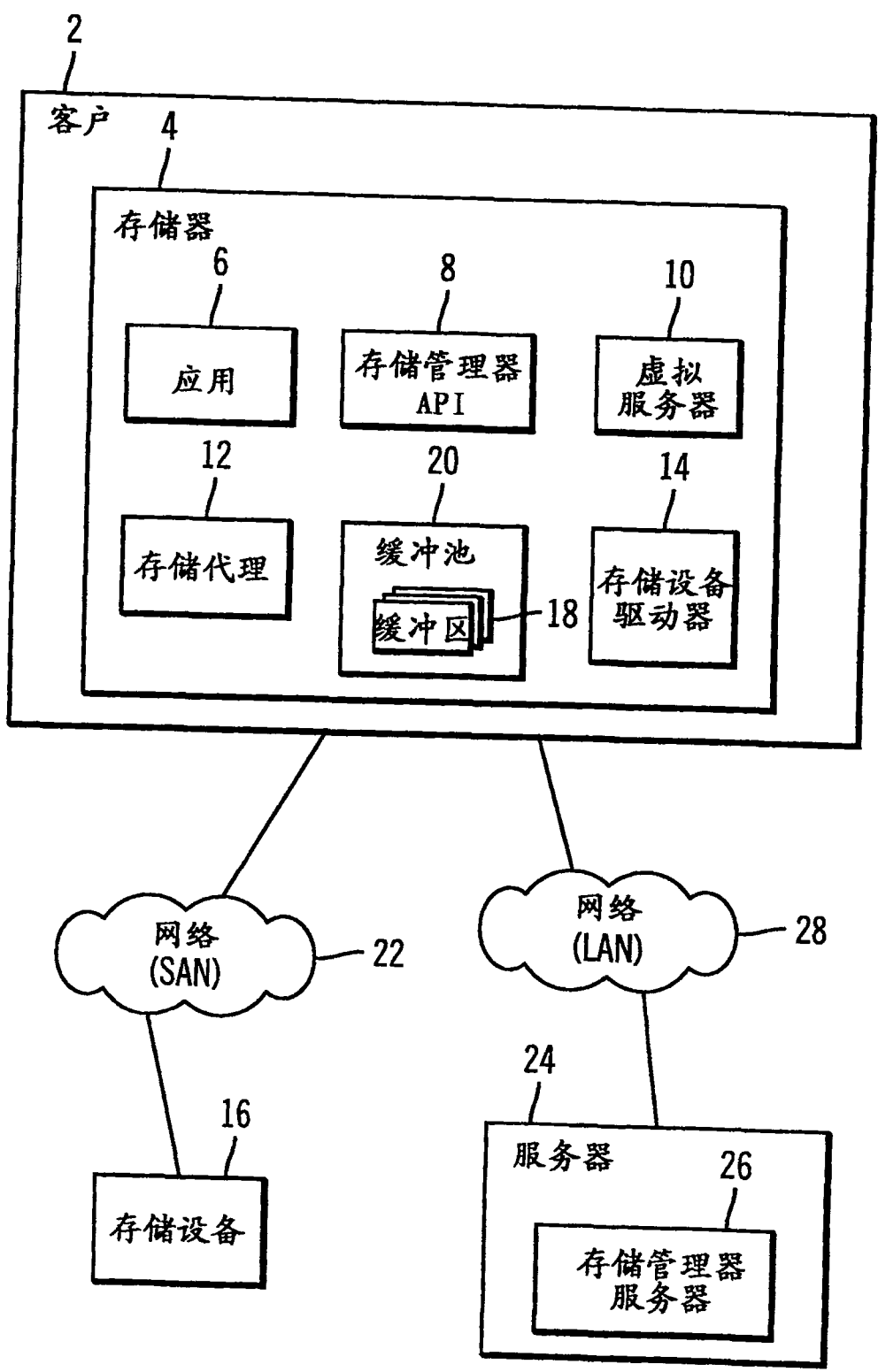


图 1

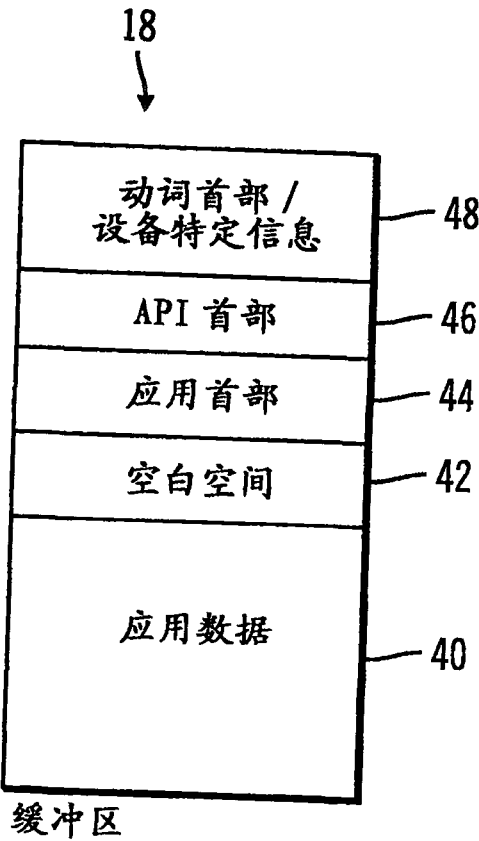


图 2

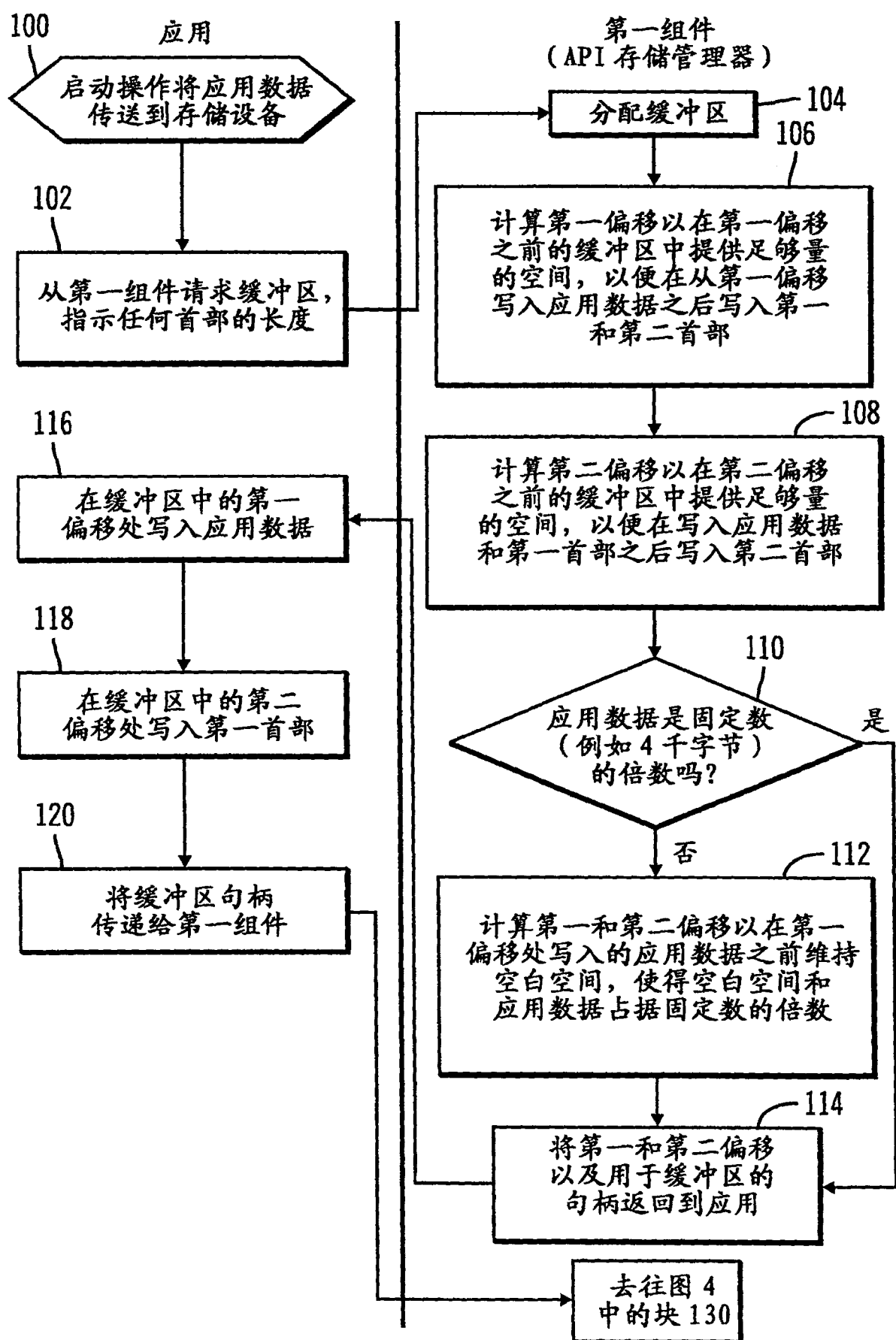


图 3

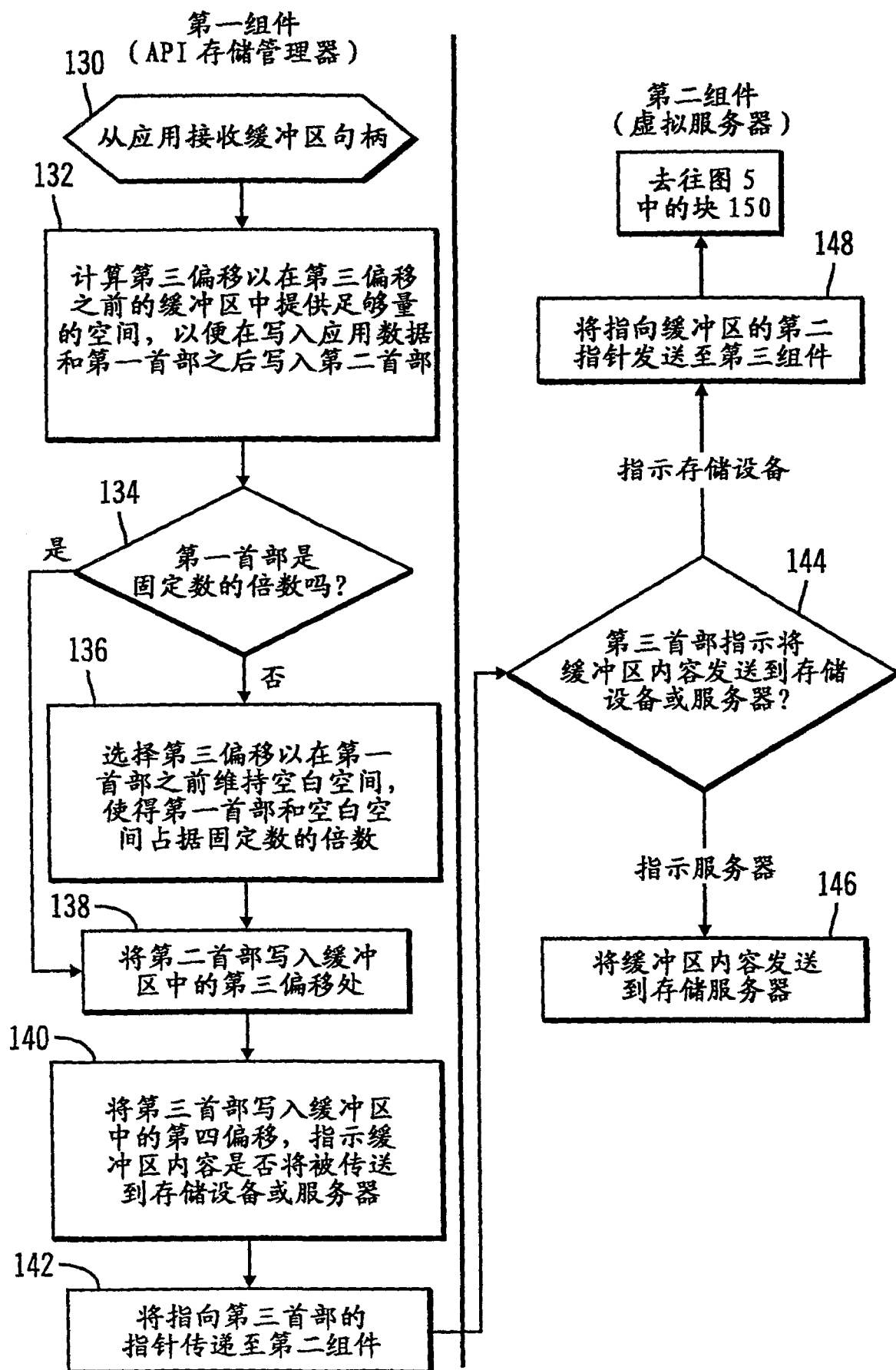


图 4

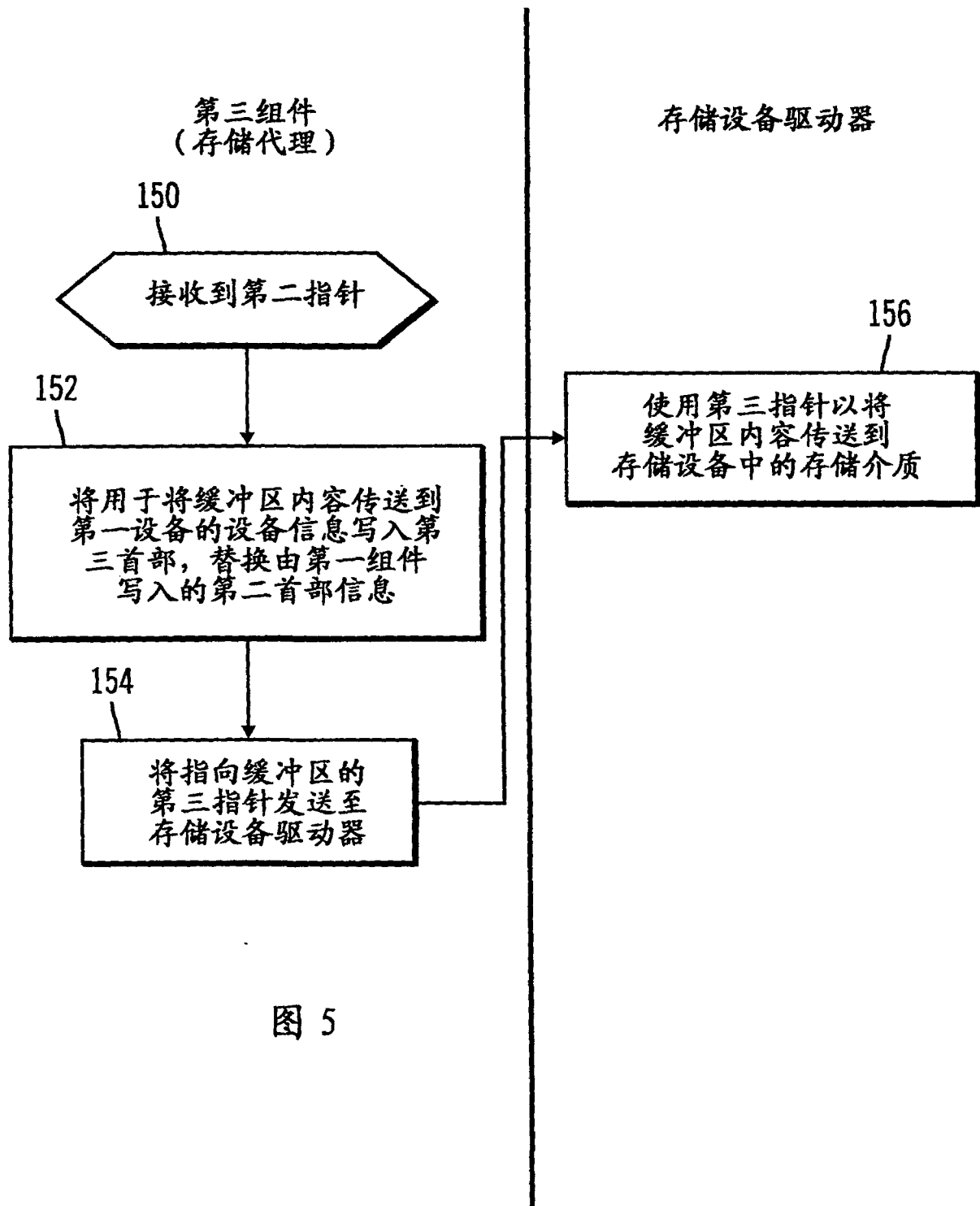


图 5