

(12) 发明专利

(10) 授权公告号 CN 101258725 B

(45) 授权公告日 2011.04.13

(21) 申请号 200680032533.1

(74) 专利代理机构 中国国际贸易促进委员会专
利商标事务所 11038

(22) 申请日 2006.07.13

代理人 赵科

(30) 优先权数据

05108246.9 2005.09.08 EP

(51) Int. Cl.

H04L 29/06 (2006.01)

(85) PCT申请进入国家阶段日

2008.03.05

(56) 对比文件

CN 1358287 A, 2002.07.10,
 US 6775230 B1, 2004.08.10,
 CN 1543135 A, 2004.11.03,
 CN 1474275 A, 2004.02.11,
 US 2004/0078465 A1, 2004.04.22,

(73) 专利权人 国际商业机器公司

审查员 李晓利

地址 美国纽约

(72) 发明人 N · 豪斯汀 W · 穆勒

权利要求书 3 页 说明书 14 页 附图 3 页

U · 特彭斯

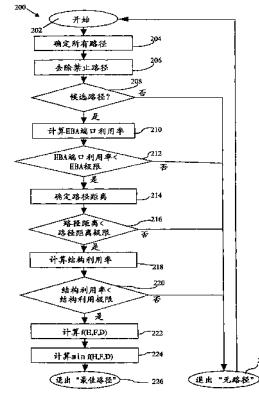
(54) 发明名称

存储区域网络中的负载分布

(57) 摘要

本发明提供了一种用于从一组 I/O 网络路径中识别输入 / 输出 (I/O) 网络路径的负载平衡方法。该组 I/O 网络路径将主机系统经网络连接至存储子系统。主机系统包括至少一个主机总线适配器 (HBA)，存储子系统包括至少一个 I/O 设备，网络包括至少一个网络设备。HBA、I/O 设备和网络设备中每一个包括至少一个 I/O 端口。对于每个 HBA 的每个 I/O 端口，确定 HBA 端口极限。此外，识别将每个 HBA 的 I/O 端口经网络设备的 I/O 端口连接至 I/O 设备的 I/O 端口的一组 I/O 网络路径。然后，为每个 I/O 网络路径确定结构利用率极限，并为该至少一个 HBA 的每个 I/O 端口确定 HBA 端口利用率。丢弃 HBA 端口利用率大于 HBA 端口极限的所有网络路径。为每个剩余网络路径确定网络路径距离。丢弃网络路径距离大于路径距离极限的所有 I/O 网络路径。然后为每个剩余网络路径确定结构利用率。丢弃结构利用率大于结构利用率极限的所有 I/O 网络路径。从剩余网络路径中确定 I/O 网络路径。

CN 101258725 B



1. 一种用于从一组 I/O 网络路径中识别输入 / 输出 I/O 网络路径 (170) 的负载平衡方法 (106)，其中所述一组 I/O 网络路径将主机系统 (102) 经网络 (112) 连接至存储子系统 (149)，所述主机系统 (102) 包括至少一个主机总线适配器 HBA (110, 111)，所述存储子系统 (149) 包括至少一个 I/O 设备 (120, 121, 125, 126)，所述网络 (112) 包括至少一个网络设备 (113, 114, 115)，每个所述 HBA (110, 111)、每个所述 I/O 设备 (120, 121, 125, 126) 和每个所述网络设备 (113, 114, 115) 包括至少一个 I/O 端口 (130, 131, 132, 133, 134, 140, 141, 142, 143, 144, 145, 150, 151, 152, 153, 160, 161, 162)，所述方法包括：

- a) 为所述至少一个 HBA (110, 111) 的所述至少一个 I/O 端口 (160, 161, 162) 中的每一个确定 HBA 端口极限 (180)；
- b) 确定将所述至少一个 HBA (110, 111) 中每一个的所述至少一个 I/O 端口 (160, 161, 162) 经所述至少一个网络设备 (113, 114, 115) 的 I/O 端口 (130, 131, 132, 133, 134, 140, 141, 142, 143, 144, 145) 连接至所述至少一个 I/O 设备 (120, 121, 125, 126) 中每一个的所述至少一个 I/O 端口 (150, 151, 152, 153) 的所述一组 I/O 网络路径；
- c) 为每个 I/O 网络路径确定结构利用率极限 (186)；
- d) 为所述至少一个 HBA (110, 111) 的所述至少一个 I/O 端口 (160, 161, 162) 中的每一个确定 HBA 端口利用率 (182)；
- e) 丢弃 HBA 端口利用率 (182) 大于 HBA 端口极限 (180) 的所有网络路径；
- f) 为每个剩余网络路径确定网络路径距离 (184)；
- g) 丢弃网络路径距离 (184) 大于路径距离极限 (185) 的所有网络路径；
- h) 为每个剩余网络路径确定结构利用率 (188)；
- i) 丢弃结构利用率 (188) 大于结构利用率极限 (186) 的所有网络路径；
- j) 从剩余网络路径中确定 I/O 网络路径 (170)。

2. 如权利要求 1 所述的方法，其中在步骤 a) 中，用所述至少一个 HBA (110, 111) 的所述至少一个 I/O 端口 (160, 161, 162) 中每一个可并行服务的 I/O 设备的数目，为所述至少一个 HBA (110, 111) 中每一个的所述至少一个 I/O 端口 (160, 161, 162) 中每一个确定 HBA 端口极限 (180)。

3. 如权利要求 1 所述的方法，其中在步骤 a) 中，用 HBA (110, 111) 的 I/O 端口 (160, 161, 162) 可处理的最大数目的每秒比特，为所述至少一个 HBA (110, 111) 中每一个的所述至少一个 I/O 端口 (160, 161, 162) 中每一个确定 HBA 端口极限 (180)。

4. 如权利要求 1-3 中任一项所述的方法，其中在步骤 d) 中，用分配给 HBA 的 I/O 端口的工作负荷来确定所述 HBA (110, 111) 的所述 I/O 端口 (160, 161, 162) 的 HBA 端口利用率 (182)。

5. 如权利要求 1-3 中任一项所述的方法，其中在步骤 d) 中，利用驱动器负担 (194) 确定 HBA 端口利用率 (182)，其中所述驱动器负担 (194) 被分配给 I/O 设备，并且所述驱动器负担 (194) 表示分配给 I/O 设备的工作负荷。

6. 如权利要求 1-3 中任一项所述的方法，其中在步骤 d) 中，利用驱动器负担 (194) 确定 HBA 端口利用率 (182)，其中所述驱动器负担由分配给 I/O 设备的工作负荷确定，

用压缩比 (192) 调整所述驱动器负担 (194)，且从所述 I/O 设备获取所述压缩比。

7. 如权利要求 1-3 中任一项所述的方法，其中在步骤 f) 中，通过加和包括在 I/O 网络路径中的所有互连的等待时间并加上包括在 I/O 网络路径中 I/O 端口的数目来确定每个剩余网络路径的网络路径距离 (184)。

8. 如权利要求 1-3 中任一项所述的方法，其中在步骤 h) 中，通过加和经过包括在网络路径中的所述至少一个网络设备的 I/O 端口的数据率来确定结构利用率 (188)。

9. 如权利要求 1-3 中任一项所述的方法，其中在步骤 h) 中，网络路径的结构利用率 (188) 用包括在所述网络路径中的活动 I/O 端口的数目与包括在所述网络路径中的 I/O 端口的总数目的比率来确定。

10. 如权利要求 1-3 中任一项所述的方法，其中在步骤 j) 中，I/O 网络路径 (170) 是考虑 HBA 端口利用率 (182)、网络路径距离 (184) 和结构利用率 (188) 的函数产生最低函数值的剩余网络路径。

11. 一种用于从一组将主机系统 (102) 经网络 (112) 连接到存储子系统 (149) 的 I/O 网络路径中识别输入 / 输出 I/O 网络路径 (170) 的负载平衡系统 (106)，所述主机系统 (102) 包括至少一个主机总线适配器 HBA (110, 111)，所述存储子系统 (149) 包括至少一个 I/O 设备 (120, 121, 125, 126)，所述网络 (112) 包括至少一个网络设备 (113, 114, 115)，每个所述 HBA (110, 111)、每个所述 I/O 设备 (120, 121, 125, 126)、和每个所述网络设备 (113, 114, 115) 包括至少一个 I/O 端口 (130, 131, 132, 133, 134, 140, 141, 142, 143, 144, 145, 150, 151, 152, 153, 160, 161, 162)，所述系统包括：

- a) 用于为所述至少一个 HBA (110, 111) 中每一个的所述至少一个 I/O 端口 (160, 161, 162) 中每一个确定 HBA 端口极限 (180) 的装置；
- b) 用于确定将所述至少一个 HBA (110, 111) 中每一个的所述至少一个 I/O 端口 (160, 161, 162) 经所述至少一个网络设备 (113, 114, 115) 的 I/O 端口 (130, 131, 132, 133, 134, 140, 141, 142, 143, 144, 145) 连接至所述至少一个 I/O 设备 (120, 121, 125, 126) 中每一个的所述至少一个 I/O 端口 (150, 151, 152, 153, 154) 的所述一组 I/O 网络路径的装置；
- c) 用于为每个 I/O 网络路径确定结构利用率极限 (186) 的装置；
- d) 用于为所述至少一个 HBA (110, 111) 的所述至少一个 I/O 端口 (160, 161, 162) 中每一个确定 HBA 端口利用率 (182) 的装置；
- e) 用于丢弃 HBA 端口利用率 (182) 大于 HBA 端口极限 (180) 的所有网络路径的装置；
- f) 用于为每个剩余网络路径确定网络路径距离 (184) 的装置；
- g) 用于丢弃网络路径距离 (184) 大于路径距离极限 (185) 的所有网络路径的装置；
- h) 用于为每个剩余网络路径确定结构利用率 (188) 的装置；
- i) 用于丢弃结构利用率 (188) 大于结构利用率极限 (186) 的所有网络路径的装置；
- j) 用于从剩余网络路径中确定 I/O 网络路径 (170) 的装置。

12. 如权利要求 11 所述的系统，其中所述用于为所述至少一个 HBA (110, 111) 中每一个的所述至少一个 I/O 端口 (160, 161, 162) 中每一个确定 HBA 端口极限 (180) 的装置被配置为：用所述至少一个 HBA (110, 111) 的所述至少一个 I/O 端口 (160, 161, 162)

中每一个可并行服务的 I/O 设备的数目，为所述至少一个 HBA(110, 111) 中每一个的所述至少一个 I/O 端口 (160, 161, 162) 中每一个确定 HBA 端口极限 (180)。

13. 如权利要求 11 所述的系统，其中所述用于为所述至少一个 HBA(110, 111) 中每一个的所述至少一个 I/O 端口 (160, 161, 162) 中每一个确定 HBA 端口极限 (180) 的装置被配置为：用 HBA(110, 111) 的 I/O 端口 (160, 161, 162) 可处理的最大数目的每秒比特，为所述至少一个 HBA(110, 111) 中每一个的所述至少一个 I/O 端口 (160, 161, 162) 的每一个确定 HBA 端口极限 (180)。

14. 如权利要求 11 到 13 中任一项所述的系统，其中所述用于为所述至少一个 HBA(110, 111) 的所述至少一个 I/O 端口 (160, 161, 162) 中每一个确定 HBA 端口利用率 (182) 的装置被配置为：用分配给 HBA 的 I/O 端口的工作负荷确定所述 HBA(110, 111) 的所述 I/O 端口 (160, 161, 162) 的 HBA 端口利用率 (182)。

15. 如权利要求 11 到 13 中任一项所述的系统，其中所述用于为所述至少一个 HBA(110, 111) 的所述至少一个 I/O 端口 (160, 161, 162) 中每一个确定 HBA 端口利用率 (182) 的装置被配置为：将所述驱动器负担 (194) 分配给 I/O 设备，并且所述驱动器负担 (194) 表示分配给所述 I/O 设备的工作负荷。

16. 如权利要求 11 到 13 中任一项所述的系统，其中所述用于为所述至少一个 HBA(110, 111) 的所述至少一个 I/O 端口 (160, 161, 162) 中每一个确定 HBA 端口利用率 (182) 的装置被配置为：利用驱动器负担 (194) 确定 HBA 端口利用率 (182)，其中用分配给 I/O 设备的工作负荷确定所述驱动器负担，用压缩比 (192) 调整所述驱动器负担 (194)，且从所述 I/O 设备获取所述压缩比 (192)。

17. 如权利要求 11 到 13 中任一项所述的系统，其中所述用于为每个剩余网络路径确定网络路径距离 (184) 的装置被配置为：通过加和包括在 I/O 网络路径中的所有互连的等待时间并加上包括在 I/O 网络路径中的 I/O 端口的数目来确定每个剩余网络路径的网络路径距离 (184)。

18. 如权利要求 11 到 13 中任一项所述的系统，其中所述用于为每个剩余网络路径确定结构利用率 (188) 的装置被配置为：通过加和经过包括在网络路径中的至少一个网络设备的 I/O 端口的数据率来确定结构利用率 (188)。

19. 如权利要求 11 到 13 中任一项所述的系统，其中所述用于为每个剩余网络路径确定结构利用率 (188) 的装置被配置为：用包括在所述网络路径中的活动 I/O 端口的数目与包括在所述网络路径中的 I/O 端口的总数目的比率确定网络路径的结构利用率 (188)。

存储区域网络中的负载分布

技术领域

[0001] 本发明一般涉及存储区域网络，尤其涉及存储区域网络内输入 / 输出工作负荷的动态分布。

背景技术

[0002] 在现代 IT 系统中，服务器和磁带驱动器经存储区域网络（例如光纤通道 SAN, iSCSI SAN）连接。此外，用于可移动介质管理的 IEEE1244 标准定义了跨异构服务器和应用程序边界有效地共享磁带驱动器的架构和协议。为了简化配置和存取控制，IEEE 1244 将驱动器组织为驱动器组。如果两个应用程序被授权访问公共驱动器组的驱动器，则这两个应用程序可共享 IEEE 1244 管理的驱动器。

[0003] 理论上，磁带管理系统、如 IEEE 1244 允许在几个应用程序之间非常灵活地共享磁带驱动器，但在实践上，该共享受到带宽因素的限制。由于主机 I/O 总线（例如 PCI 总线）和服务器 I/O 端口（例如光纤通道端口，iSCSI 端口）的带宽限制，I/O 端口仅可有效地利用一定数目的驱动器。如果 I/O 端口所利用的驱动器的数目高于某个阈值，则性能将受到影响。

[0004] 该数目经常被称为 I/O 端口的驱动器极限。而且，在异构环境中，不同 I/O 端口可以具有不同的驱动器极限。例如，在全驱动速度时，2Gbit 的光纤通道端口可以比 1Gbit 光纤通道端口或 1Gbit iSCSI 端口利用更多驱动器。术语“驱动器”不局限于磁带驱动器，其可以是任何其他 I/O 设备，诸如磁盘设备，光盘设备，随机存取存储设备，全息记录设备或纳米技术设备。

[0005] 现有技术使用以下方法和系统来确保在驱动器共享环境中，I/O 端口同时使用的驱动器的数目不超过其驱动器极限。

[0006] 1. 经由软件配置对资源分区。

[0007] 不超过端口的驱动器极限的典型方法是减少在端口处可见的驱动器的数目。假定有两个服务器，每个服务器装配有一个驱动器极限为 n 的 I/O 端口。如果存储器管理员知道这两个服务器在一天的不同时间需要 n 个驱动器，则通常将这 n 个相同的驱动器划分 (zone) 给这两个服务器。这使得驱动器能够在这两个服务器之间共享，同时确保不超过相应驱动器极限。

[0008] 该方法的缺点是缺少灵活性。例如，如果其中一个驱动器发生故障，则需要手动干预。管理员必须识别可以将其划分到这两个服务器的 I/O 端口的可用的备用驱动器。而且，他必须在应用程序可以使用新驱动器之前触发这两个服务器的操作系统以扫描新硬件。此外，第三服务器的存取特征必须在其可以被集成到该环境中而不与另两个服务器的资源要求冲突之前被公知。而且，如果其中一个服务器的存取特征变化，则还必须修改配置。如果这两个服务器随着时间的推移在一天中的相同时间需要驱动器或如果随着时间的推移可以以不超过一个驱动器来满足其中一个服务器的工作负荷，则要求额外的措施。该情形导致为了改善整个磁带基础设施 (infrastructure) 的集合驱动器利用

率所要求的 SAN 划分变化。 总之，这导致灵活性有限，因此该方法不适于大环境。

[0009] 2. 具有磁盘高速缓存的虚拟磁带服务器

[0010] 另一个方法是在服务器和磁带之间的 I/O 路径中设置磁盘高速缓存。 在这样的情形中，所谓的虚拟磁带服务器将虚拟磁带驱动器中的虚拟带盒 (cartridge) 提供给服务器，同时虚拟磁带服务器在内部将数据重新引导到磁盘高速缓存。 随后，虚拟磁带服务器经实际驱动器将数据从磁盘高速缓存转移 (destage) 到实际磁带盒。 该方法的缺点是需要额外的硬件和软件来构成虚拟磁带服务器。

[0011] 3. 无磁盘高速缓存的虚拟磁带服务器

[0012] 在另一情形中，虚拟磁带服务器输出虚拟磁带驱动器到服务器，其中服务器可使用的虚拟磁带驱动器的数量受到软件配置中的参数限制。 当服务器要安装带盒时，虚拟磁带服务器动态地将虚拟磁带驱动器连接到物理磁带驱动器。 以该方式，可实现物理磁带驱动器到服务器 I/O 端口的灵活映射。 该方法的缺点是需要额外的硬件和软件来形成虚拟磁带服务器。

[0013] 4. 具有多个端口的磁带驱动器

[0014] 某些磁带驱动器装配有多个端口。 额外的多路径软件可被用于在所有可用路径上平衡 I/O。 该方法的缺点是需要额外的硬件（磁带驱动器处的端口，连接第二驱动器端口的存储器网络的端口）和软件（诸如多路径软件）。 通常，一个端口在此时是活动的，而其他端口用作在经该活动端口执行 I/O 出现问题的情况下冗余目的。 因此，该方法不适于大环境。

[0015] 因此，需要一种方法和系统来促进在经多个主机总线适配器连接的可移动介质设备之间 I/O 工作负载分布的动态分布。

发明内容

[0016] 按照本发明，提供了一种负载平衡方法，用于从一组 I/O 网络路径中识别输入 / 输出 (I/O) 网络路径。 该组 I/O 网络路径中的每个路径将主机系统经网络连接到存储子系统。 主机系统包括至少一个主机总线适配器 (HBA)。 存储子系统包括至少一个 I/O 设备，且网络包括至少一个网络设备。 HBA、I/O 设备和网络设备中每一个都包括至少一个 I/O 端口。 对于该至少一个 HBA 的该至少一个 I/O 端口中的每一个，确定 HBA 端口极限。 然后确定经该至少一个网络设备的 I/O 端口将该至少一个 HBA 中每一个的该至少一个 I/O 端口连接到该至少一个 I/O 设备中每一个的该至少一个 I/O 端口的该组 I/O 网络路径。 然后，为每个 I/O 网络路径确定结构 (fabric) 利用极限，并为该至少一个 HBA 的该至少一个 I/O 端口中每一个确定 HBA 端口利用率。 HBA 端口利用率大于 HBA 端口极限的所有网络路径被丢弃。 对于剩余路径中的每一个，确定网络路径距离。 网络路径距离大于路径距离极限的所有网络路径被丢弃。 然后，对于每个剩余路径，结构利用率被确定。 结构利用率大于结构利用率极限的所有路径被丢弃，且从剩余网络路径中确定 I/O 网络路径。

[0017] 由于本发明通过考虑参数 HBA 端口利用率、路径距离和结构利用率而动态地确定从 HBA 到 I/O 设备的网络路径，所以本发明特别有用。 分别相对于参数 HBA 端口极限、路径距离极限和结构利用率极限检验这些参数。 以该方式，从所有合格的网络路径

中识别最佳网络路径。

[0018] 按照本发明的一个实施例，通过 HBA 的 I/O 端口可并行服务的 I/O 设备的数目，为该至少一个 HBA 中每一个的该至少一个 I/O 端口中每一个确定 HBA 端口极限。因为 HBA 的 I/O 端口可并行服务的 I/O 设备的数目通常被指定，所以这特别有利。按照本发明，规范例如可以由系统管理员提供或可以由负载平衡系统从相应的 I/O 端口读出。规范定义 HBA 的一个 I/O 端口可提供的带宽，并直接由可从所述 HBAI/O 端口接收 I/O 的 I/O 设备的数目表述。带宽主要定义一个 HBA I/O 端口能执行的每秒的 I/O 操作的数目，其通常以每秒兆字节 (MB/s) 测量。

[0019] 按照本发明的一个实施例，通过 HBA 的 I/O 端口可处理的每秒的最大比特数，为该至少一个 HBA 中每一个的该至少一个 I/O 端口中每一个确定 HBA 端口极限。特别有利的是使用 I/O 端口可处理的每秒的比特数作为 HBA 端口极限，因为该值可由负载平衡系统动态地确定。

[0020] 按照本发明的一个实施例，通过被分配给 HBA 的 I/O 端口的工作负荷（或带宽）确定 HBA 的 I/O 端口的 HBA 端口利用率。因为负载平衡系统可通过发送查询命令到 HBA 而动态地确定工作负荷，因此这特别有利。

[0021] 按照本发明的一个实施例，利用驱动器负担 (drive weight) 确定 HBA 端口利用率。驱动器负担被分配给 I/O 设备，并且驱动器负担表示 I/O 设备的工作负荷。

[0022] 按照本发明的一个实施例，利用驱动器负担确定 HBA 端口利用率，其中驱动器负担表示 I/O 设备的工作负荷并由压缩比 (compression ratio) 调整，而从 I/O 设备获取压缩比。

[0023] 本发明因为支持考虑该至少一个主机总线适配器的该至少一个 I/O 端口的 HBA 端口极限，并丢弃 HBA 端口利用率大于 HBA 端口极限的所有路径，所以特别有利。注意，每个 HBA 的每个 I/O 端口具有其自己特定的 HBA 端口极限。这确保 HBA 的 I/O 端口只利用最优数目的驱动器。这对整个 I/O 设备基础设施的集合带宽的整体改善有贡献。

[0024] 按照本发明的一个实施例，通过加和包括在 I/O 网络路径中的所有互连的等待时间并加上包括在 I/O 网络路径中的 I/O 端口的数目来确定每个剩余网络路径的网络路径距离。

[0025] 按照本发明的一个实施例，通过加和经过包括在网络路径中的该至少一个网络设备的 I/O 端口的数据率来确定结构利用率。通过网络设备的网络路径一般包括至少两个端口，一个输入端口和一个输出端口。

[0026] 按照本发明的一个实施例，通过包括在网络路径中的活动 I/O 端口的数目与包括在网络路径中的 I/O 端口的总数目的比率来确定网络路径的结构利用率。

[0027] 使用结构利用率来确定网络路径特别有利，因为考虑了包括在网络路径中的所有部件的工作负荷。选择具有最低结构利用率的路径提供了所有活动资源的最佳利用率。而且，本发明特别有利，因为其考虑了网络路径的结构利用极限，并检验结构利用率的大小是否低于结构利用极限。这确保了从不超过 I/O 设备基础设施的最大带宽。

[0028] 按照本发明的一个实施例，“最佳” I/O 网络路径是考虑 HBA 端口利用率、网络路径距离和结构利用率的函数产生最低函数值的剩余网络路径。这三个参数 HBA 利用

率、网络路径距离和结构利用率被分配给从中选择最佳路径的所有路径。例如，函数由系统管理员指定。使用函数来确定最佳路径具有这样的优点，即其使系统管理员能够对这三个参数同等地或不同地加权。

[0029] 本发明在另一方面中涉及一种计算机程序产品，其包括计算机可执行指令以便执行本发明的方法。

[0030] 本发明在再一方面中涉及一种负载平衡系统，用于从一组 I/O 网络路径中识别 I/O 网络路径，该组 I/O 网络路径将主机系统经网络连接到存储子系统。

附图说明

[0031] 下面将参考附图仅仅以示例的方式更详细地介绍本发明的优选实施例，其中：

[0032] 图 1 是实现本发明一个实施例的系统的方框图。

[0033] 图 2 的流程图示出在本发明一个优选实施例中如何确定 I/O 网络路径。

[0034] 图 3 的流程图示出在本发明另一优选实施例中如何确定 I/O 网络路径。

具体实施方式

[0035] 图 1 是实现本发明一个实施例的系统的方框图。系统包括可移动介质管理系统 118、主机系统 102、网络 112、存储子系统 149。可移动介质管理系统 118 是计算机系统，其包括负载平衡系统 106、微处理器 103、非易失性存储设备 105、和易失性存储设备 107。主机系统也是计算机系统，例如服务器系统，其包括应用程序 104、与可移动介质管理系统 118 通信的代理 109、设备驱动器 108、和分别包括 I/O 端口 160、161 及 162 的主机总线适配器 (HBA) 110、111。

[0036] 网络 112 包括网络元件 113、114、和 115，它们是典型的网络设备，诸如存储网络内的集线器、交换器和路由器（例如，光纤通道，iSCSI）。每个网络设备包括 I/O 端口 130、131、132、133、134、140、141、142、143、144、145、和 146。

[0037] 存储子系统 149 包括 I/O 设备 120、121、125、和 126。I/O 设备通常是可移动介质设备，诸如磁带类 I/O 设备、磁盘类 I/O 设备、全息类 I/O 设备、纳米技术类 I/O 设备或光学 I/O 设备。每个 I/O 设备 120、121、125 和 126 分别包括一个或多个 I/O 端口 150、151、152、153、和 154。

[0038] I/O 网络路径利用互连将 HBA 的 I/O 端口经一个或多个网络设备的 I/O 端口连接到 I/O 设备的 I/O 端口。例如，网络路径 190 利用互连将 HBA 111 的 I/O 端口 161 连接到网络设备 114 的 I/O 端口 131，并将网络设备 114 的另一 I/O 端口 145 连接至 I/O 设备 126 的 I/O 端口 153。应该指出，HBA 可包括一个以上 I/O 端口。通常多达 4 个 I/O 端口。而且，实践中，网络设备包括两个或更多 I/O 端口，且 I/O 设备也可包括一个以上的 I/O 端口。

[0039] 路径包括一个或多个互连。互连可以是将数据从一个 I/O 端口传输到另一 I/O 端口的任何设备。互连的例子是光纤通道缆线、以太网缆线、或 WAN 互连，如 CWDM、DWDM、ATM、Sonet/SDH、或 TCP/IP。互连也可由无线连接提供。例如，主机系统和存储子系统可通过无线局域网互连，从而数据通过光学信号从一个 I/O 端口发射到另一 I/O 端口，且光学信号具有例如红外频率范围内的频率。

[0040] 如果光学缆线被用作互连或如果使用无线互连，则互连自身可包括信号频率不同的几个通道。因此，互连可提供几个路径，而每个路径以通道为特征或相应地以用于传输数据的光学频率为特征。

[0041] 网络路径也可包括一个以上的网络元件。例如，在 I/O 网络路径 170 中，互连将 I/O 端口 160 连接到 I/O 端口 130，然后互连将 I/O 端口 140 连接到 I/O 端口 132，并且附加地，互连将 I/O 端口 142 连接到 I/O 设备 120 的 I/O 端口 150。

[0042] 网络路径也可不包括网络元件。I/O 设备的 I/O 端口可通过互连直接与 HBA 的 I/O 端口连接。网络路径由所述互连组成。

[0043] 应用程序 104 利用主机系统 102 访问 I/O 设备 120、121、125、和 126。主机系统 102，特别是应用程序 104 经代理 109 和适当的通信方法，诸如 TCPIP，联系包括在介质管理系统 118 中的负载平衡系统 106。负载平衡系统 106 从应用程序 104 可通过其读取或写入数据至 I/O 设备 120、121、125、126 的网络路径组中确定网络路径 170。作为数据存取请求，诸如安装、读取或写入命令，的结果，应用程序 104 可联系负载平衡系统 106。下面介绍用于从该组网络路径中识别 I/O 网络路径的方法的多种实施例。

[0044] 负载平衡系统 106 使用代理 109 来确定网络和存储子系统 149 的某些参数，这些参数被存储在易失性存储设备 107 上。可替换地，这些参数可被存储在非易失性存储设备 105 上。主机系统 102 采用一个或多个设备驱动器 108，以便访问 I/O 设备 120、121、125 和 126。

[0045] 然而，按照本发明的系统并不局限于该实施例。例如，包括负载平衡系统 106 的介质管理系统 118 可包括在主机系统 102 中或包括在存储子系统 149 中。介质管理系统 118 例如可以用内置到应用程序的库管理器的形式被实现到主机系统 102 中。介质管理系统 118 例如可以以用于磁带库的先进库管理器的形式被实现到存储子系统 149 中。然而，外部介质管理系统 118 的优点在于，其可同时被几个连接到不同网络和存储子系统或连接到同一网络和存储子系统的主机系统使用。

[0046] 这组网络路径包括经网络 112 将 HBA 110 或 111 的 I/O 端口 160、161 或 162 连接到存储子系统 149 的 I/O 设备 120、121、125、和 126 的所有路径。例如，I/O 网络路径 170 由从 HBA 110 的 I/O 端口 160 到网络设备 113 的 I/O 端口 130 的互连和从网络设备 113 的 I/O 端口 140 到网络设备 115 的 I/O 端口 132 的互连以及从网络设备 115 的 I/O 端口 142 到 I/O 设备 120 的 I/O 端口 150 的互连实现。

[0047] 按照本发明一个实施例，除了下面指定的其他参数之外，负载平衡系统 106 还确定被存储在易失性存储设备 107 或非易失性存储设备 105 上的以下参数：HBA 端口极限 180、HBA 端口利用率 182、网络路径距离 184、路径距离极限 185、结构利用率极限 186、结构利用率 188、压缩比 192、和驱动器负担 194。下面详细介绍如何确定这些参数。

[0048] 微处理器 103 响应于从应用程序 104 接收到数据存取请求而执行负载平衡方法。数据请求的例子包括安装带盒请求、建立从 SCSI 源到存储网络中 SCSI 目标的连接、和在存储网络中创建新区段。

[0049] 图 2 示出的流程图示出了在本发明的一个优选实施例中如何确定“最佳” I/O 网络路径。过程从 202 开始，前进到 204，其中确定所有路径的集合。每个 I/O 网络路径

描绘从系统的 HBA(主机系统可包括一个以上的 HBA, 其中每个 HBA 包括一个或多个 I/O 端口) 的 I/O 端口经网络到 I/O 设备的 I/O 端口(存储子系统也可包括一个以上的 I/O 设备, 其中 I/O 设备可包括一个以上的 I/O 端口)的路由。HBA 的 I/O 端口在下文中也被称为 HBA 端口。作为步骤 204 的结果, 集合 P_{204} 描述所有路径的集合。步骤 206 移除对于 I/O 不合格的所有路径。不合格的 I/O 网络路径的例子包括不是活动光纤通道区段的成员的物理 I/O 网络路径。不合格网络路径的另一例子是非空闲或不可用的设备。设备可能因为被另一应用程序或另一主机系统使用而不是空闲的。当设备有故障或离线时, 其可能是不可用的。注意, 对于新数据存取请求的处理, 步骤 204 仅考虑到当前未被使用的驱动器的路径。作为步骤 206 的结果, 集合 P_{206} 描述所有合格路径的集合。在一个可替换实施例中, 步骤 204 仅确定合格步骤, 因此步骤 206 可省略。

[0050] 然后, 过程进入判定检查点 208。如果此时没有遗留合格路径 ($P_{206} = \emptyset$), 则过程在步骤 299 结束。步骤 299 指示此时没有合格路径可用。步骤 299 中所执行的动作或步骤 299 的结果取决于环境。例如, 在可移动介质管理系统中, 各安装请求可能被阻挡, 这意味着安装请求被添加到安装请求队列中, 或被拒绝。在另一实施例中, 存储网络容量规划模块可拒绝各划分或存取请求。在再一实施例中, 步骤 299 可循环返回到该流程图的起点步骤 202, 以便重新尝试路径寻找。在又一实施例中, 可调用按需模块, 其动态地提供附加基础设施以便添加更多路径到路径结合 P_{204} 中。

[0051] 如果在检查点 208 存在满足存取请求的合格路径 ($P_{206} \neq \emptyset$), 则过程进入到步骤 210。步骤 210 建立主机总线适配器的所有 I/O 端口 h 的集合 H , 其中 h 是至少一个 I/O 网络路径 p 元素 P_{206} ($p \in P_{206}$) 的起点。而且, 在步骤 210 中, 在假设存取请求被调度的情况下, 为 H 的每个主机适配器 I/O 端口 h 元素确定 HBA 端口利用率 u_h 。计算 HBA 端口利用率的可替换实施例在后面详细说明。

[0052] 然后, 过程进入步骤 212。步骤 212 比较每个 HAB 端口 $h \in H$ 的 HBA 端口利用率 u_h 和 HBA 端口极限 l_h , 并丢弃 HBA 端口利用率 $u_h > HBA$ 端口极限 l_h 180 的所有路径 p 。在一个实施例中, 在应用程序已经触发过程 200 之后动态地确定 HBA 端口极限 l_h 。该确定基于媒体管理系统的处理器扫描和询问的网络中网络元件的参数和存储系统中的 I/O 设备特征。例如, 负载平衡系统询问与网络路径相关联的所有元件, 并确定路径中可能的最大数据率。该最大数据率于是相当于 HBA 端口极限。在另一实施例中, 用户为每个 HBA 端口配置 HBA 端口极限 l_h 。在后面进一步说明参数 HBA 端口极限 l_h 。步骤 212 的结果是网络路径集合 P_{212} , 其中 $P_{212} = \{p \in P_{206}, \text{其中 } "h \text{ 是 } p \text{ 的起点}" \text{ 且 } "新 HBA \text{ 端口利用率 } u_h \leq HBA \text{ 端口极限 } l_h"\}$ 。

[0053] 如果在步骤 212 中确定没有合格的候选者 ($P_{212} = \emptyset$), 则过程进入步骤 299, 表明没有发现合格路径。

[0054] 否则, 流程进入步骤 214, 其中确定所有剩余路径 $p \in P_{212}$ 的网络路径距离 N_p 。路径 $p \in P_{212}$ 的网络路径距离 N_p 是两个端口之间的物理距离或互连长度。路径 $p \in P_{212}$ 的网络路径距离 N_p 可由介质管理系统的微处理器例如通过在网络上发送测试分组并测量回传时间 (turnaround time) 而自动确定。回传时间与所述分组在网络上必须行进的距离成正比。在一个可替换实施例中, 网络路径距离 N_p 是用户可配置的, 这意味着用户为网

络上的任何给定路径配置距离。参数网络路径距离在下面进一步解释。

[0055] 下一步骤 216 丢弃 I/O 网络路径 p 的网络路径距离 N_p 大于路径距离极限 l_d 的所有路径 $p \in P_{212}$ 。在本发明的一个实施例中，路径距离极限 l_d 在应用程序已经触发过程 200 之后由按照本发明的负载平衡系统动态地确定。该确定基于连接主机系统和存储子系统的网络元件的参数，这些参数被扫描并被询问。在另一实施例中，用户为每个互连配置路径距离极限。后面进一步解释参数路径距离极限 l_d 。步骤 216 的结果是集合 P_{216} ，其中 $P_{216} = \{p \in P_{212} \text{, 其中 "I/O 网络路径 } p \text{ 的网络路径距离 } N_p \leq \text{I/O 网络路径 } p \text{ 的路径距离极限 } l_d\}$ 。如果没有合格候选者 ($P_{216} = \emptyset$)，则处理在步骤 299 结束。

[0056] 步骤 218 计算所有剩余路径 p 元素 P_{216} 的结构利用率。路径 $p \in P_{216}$ 的结构利用率 F_p 是由包括路径 p 的结构 F 必须管理的工作负荷来限定。路径 $p \in P_{216}$ 的结构利用率 F_p 可以由介质管理系统的微处理器自动确定，例如通过发送询问到结构 F 的网络元件，以询问每秒的 I/O 数目。然后，在结构 F 中与路径 p 相关的所有网络元件上累加该数目。参数结构利用率在后面进一步说明。

[0057] 下一步骤 220 丢弃其中路径 p 的结构利用率大于 I/O 网络路径 p 的结构利用率极限 l_f 的所有路径 p 元素 P_{216} 。在本发明的一个实施例中，结构利用率极限 l_f 由按照本发明的负载平衡系统在应用程序已经触发过程 200 之后动态地确定。该确定基于被扫描和询问的连接主机系统和存储子系统的网络中网络元件的参数。在另一实施例中，结构利用率极限由用户配置。参数结构利用率极限 l_f 在下面进一步说明。

[0058] 步骤 220 的结果是集合 P_{220} ，其中 $P_{220} = \{p \in P_{216} \text{, 其中 "路径 } p \text{ 的新结构利用率 } \leq \text{路径 } p \text{ 的结构利用率极限}\}$ 。如果 P_{220} 中没有合格候选者 ($P_{220} = \emptyset$)，则过程进入步骤 299，指示没有发现合格路径。否则，所有路径 $p \in P_{220}$ 不超过带宽，且因此是“最佳路径”的潜在候选者，并且过程进入步骤 222。因此，过程的剩余步骤不丢弃额外的路径，但确定最优路径。

[0059] 步骤 222 引入函数 f ，该函数加权这三个参数 HBA 端口利用率 u_h 、结构利用率 p_f 和距离 p_d 。在一个实施例中，函数 f 是加权和 $f(u_h, p_f, p_d) = a \times u_h + b \times p_f + c \times p_d$ ，其中 a 、 b 、和 c 是例如由系统管理员指定的常数值。

[0060] 最后，在步骤 224 中，选择所有 $p \in P_{220}$ 中上面给出的函数 f 产生最低函数值的 I/O 网络路径 $p_{min}: p_{min} = \min(f(u_h, p_f, p_d))$ 。处理在步骤 226 结束，其返回所确定的网络路径作为具有最佳性能特征的“最佳路径”。

[0061] 在本发明的一个实施例中，仅执行步骤 202 到 212，和步骤 226 或步骤 229。如上所述，步骤 212 的结果是网络路径集合 P_{212} ，其中 $P_{212} = \{p \in P_{206} \text{, 其中 "h 是 } p \text{ 的起点}\} \text{ 且 "新 HBA 端口利用率 } u_h \leq \text{HBA 端口极限 } l_h\}$ 。如果 $P_{212} \neq \emptyset$ ，则过程进入步骤 226，其中集合 P_{212} 中一个网络路径被选择，例如通过从列表 P_{212} 中选择第一个网络路径，或通过从集合 P_{212} 中随机选择网络路径。如果 $P_{212} = \emptyset$ ，则过程进入步骤 299。

[0062] 按照本发明，每个 I/O 网络路径 p 元素 P_{220} 的特征为三个参数 (a) HBA 端口利用率，(b) 网络路径距离 184 和 (c) 结构利用率，这三个参数被用于确定“最佳路径”。这些参数在下面说明。

[0063] HBA 端口利用率 u_h —该参数跟踪多少工作负荷当前被分配给 HBA 的 I/O 端口并

以每秒兆字节 (MB/s) 表示。如果当前在该 HBA 端口没有使用驱动器，则其值为零。如果一个或多个驱动器被调度用于在 HBA 端口处使用，则该值被设为各驱动器的驱动器负担之和。

[0064] 驱动器负担是表示 I/O 设备消耗多少 I/O 资源的额外参数。驱动器负担由以每秒兆字节 (MB/s) 为单位的 I/O 率表示。该参数可通过发送询问命令至驱动器以指示驱动器报告可实现的数据率而被自动获得。例如，询问命令是 SCSI INQUIRY 命令，其向驱动器询问特性，其中一个参数通常表示驱动器带宽。驱动器带宽例如表示 I/O 设备可处理的数据率，并通常以每秒兆字节表示。在一个可替换实施例中，参数驱动器负担可由用户配置。在又一实施例中，驱动器负担由按照本发明的介质管理系统基于 I/O 设备操作而动态地调整。I/O 设备通常保持关于先前操作中所处理的数据率的统计量。该统计量可由介质管理系统询问。询问命令的一个例子是 SCSI LOG SENSE 命令。基于该信息，介质管理系统可通过计算多个安装的平均 I/O 率而动态地更新驱动器负担。

[0065] 在一个实施例中，当驱动器负担涉及未压缩数据时，用压缩比调整驱动器负担参数。压缩通常在 I/O 设备级以一定比率进行，并影响 I/O 设备的 I/O 率。在该实施例中，驱动器负担参数在被加到 HBA 端口利用率并与 HBA 端口极限比较之前与压缩比率相乘。在一个实施例中，介质管理向 I/O 设备询问压缩比率参数。介质管理系统计算压缩比率平均值，并将其存储在易失性存储设备中。询问命令是 SCSI-3LOG SENSE 命令。在另一实施例中，压缩比率由用户配置。

[0066] 用于更新 HBA 端口利用率的方法相当简单。其从没有任何 I/O 设备经该端口被使用时开始，其中值为零。如果作为图 2 中过程的结果，I/O 设备被调度用于 HBA 端口（安装请求），则 HBA 端口利用率的值被增加驱动器负担参数，其中该参数可由 I/O 设备的压缩比调整。然后，HBA 端口利用率被存储在介质管理系统中。如果驱动器被从 HBA 端口取消调度 (de-scheduled) (不安装请求)，则 HBA 端口利用率的值被减少驱动器负担。

[0067] 网络路径距离 N_p – 网络路径距离主要由 I/O 网络路径的长度表示，其中 I/O 网络路径的长度是包括在 I/O 网络路径中的互连的物理长度。在优选实施例中包括 I/O 网络路径中包括的 I/O 端口的数目。其原因是，每个端口表示一定水平的开销，其影响数据率并因此应当被考虑。按照该优选实施例，网络路径距离以长度单位计量，如米 (m)。在一个可替换实施例中，该参数以数据分组在该路径上的回传时间来表述，其中回传时间以时间单位计量，诸如毫秒 (ms)。为了确定给定网络路径的长度，负载平衡系统过程在路径上发送分组并测量回传时间，类似于根据 TCPIP 协议的“ping”命令。回传时间作为网络路径距离参数被存储在介质管理系统的存储设备中。在一个可替换实施例中，路径距离由用户配置并以米表述。

[0068] 结构利用率 F_p – 描述路径 p 位于其中的结构利用率。结构由大量交换器或网络元件组成。网络元件或交换器由大量 I/O 端口组成。通过路径内 I/O 端口的 I/O 率之和来计算路径 p 的结构利用率。因此，不考虑整个结构，而仅考虑一个 HBA 和一个 I/O 设备之间的 I/O 网络路径中的元件。该参数用数据率表述，诸如每秒兆字节，以描述属于结构内路径 p 的端口的所包括的数据率。负载平衡系统通过适当命令确定属于路径的端口的数据率。例如，按照现有技术的 SAN 交换元件提供负载平衡系统可用其来询问 I/

O 端口的数据率的用于出站连接的应用编程接口 (API)。在一个可替换实施例中，路径 p 的结构利用率通过结构中活动端口数目与结构中总端口数目的比率来确定。端口的活动性由负载平衡系统 106 询问，例如通过 SAN 元件所提供的 API。作为图 2 所示过程中步骤 218 的一部分，确定该参数。因为结构利用率动态地改变，所以其可选地可在后台中被确定。

[0069] 通过利用该参数，按照本发明的负载平衡方法和系统考虑连接 HBA 端口和 I/O 设备的路径中所有元件的工作负荷。选择结构利用率最低的路径就提供所有资源的相等利用率。

[0070] 图 2 中所描述的过程还包括允许排除某些路径的下列这些参数的极限或阈值，诸如 (a) HBA 端口极限，(b) 网络路径距离极限和 (c) 结构利用率极限。将这些极限与上面在步骤 212、216 和 220 中所概述的三个参数之一进行比较，将在下面进一步说明这些极限。

[0071] HBA 端口极限 l_h -HBA 端口极限是 HBA 的每个端口的新元数据属性。在一个实施例中，HBA 端口极限表示该端口可并行服务的驱动器的数量。在另一实施例中，HBA 端口极限表示端口可处理的最大吞吐量 (例如，100Gbit)。在其中互连是包括几个通道的光学缆线或无线连接，因而利用不同光学频率在每个通道上传递数据的另一实施例中，HBA 端口极限也可以是 HBA 端口可处理的通道的数目。

[0072] 路径距离极限 l_d - 描述路径的最大允许距离，并与路径距离比较。路径距离极限是用户配置的常数值。

[0073] 结构利用率极限 l_f - 描述结构的最大允许的利用率，并与结构利用率比较。结构利用率极限由用户配置。

[0074] 图 3 教导了选择“最佳 I/O 网络路径”的另一实施例。与 200 所描述的过程的主要差别在于，过程 300 不移除禁止路径。具体地，过程 300 不考虑 HBA 端口利用率的极限、网络路径距离的极限、和结构利用率的极限。相反，该过程为上面所述的这三个参数 (a) HBA 端口利用率、(b) 网络路径距离和 (c) 结构利用率中的每一个选择最小值。该实施例与图 2 说明的实施例一起表示存在多个可替换方案来实现负载平衡系统。

[0075] 负载平衡系统 106 响应于来自应用程序 104 的数据存取请求执行过程 300。一个可替换实施例是，该过程在主机系统 102 的利用率较小时被执行。

[0076] 过程 300 从 302 开始，并进入 304，其中确定所有路径的集合。每个 I/O 网络路径描述从主机系统 102 的主机总线适配器 110 或 111 的 I/O 端口 160、161、或 162 经网络 112 到 I/O 设备 120、121、125 和 126 的该至少一个 I/O 端口的路由。例如，I/O 网络路径 170 包括 HBA110 的 I/O 端口 160、从 I/O 端口 160 到 I/O 端口 130 的互连、网络设备 113、从 I/O 端口 140 到 I/O 端口 132 的互连、网络设备 115、和从 I/O 端口 142 到 I/O 设备 120 的 I/O 端口 150 的互连。

[0077] 在步骤 305 中，检验是否存在任何网络路径。如果不存在网络路径，则过程进入步骤 399，并停止，因而不确定路径。如果存在至少一个网络路径，则过程进入步骤 306。在步骤 306 中，每个主机总线适配器 110 或 111 的每个 I/O 端口 160、161、或 162 的 HBA 端口利用率被确定。步骤 308 确定利用率最低的 I/O 端口 160、161 或 162。在步骤 310 中，HBA 端口 160、161 和 162 的 HBA 端口利用率被比较，且选择利用率最低

的 HBA 端口。如果仅存在一个 I/O 网络路径具有最小 HBA 端口利用率，则过程进入后面讨论的步骤 326。

[0078] 否则，如果存在一个以上的 I/O 网络路径具有最小利用率，则过程进入步骤 312，其中为具有相同的最小利用率的 I/O 网络路径确定主机总线适配器端口 160、161 和 162 与驱动器 120、121、125、126 的 I/O 端口 150、151、152、153、和 154 之间的距离。

[0079] 在步骤 314，确定步骤 312 中所考虑的 I/O 网络路径中的最小网络路径距离。在步骤 316，确定具有最小距离的路径数目是否大于 1。如果答案是否定的，即仅有一个路径，则过程进入后面讨论的步骤 326。如果有一个以上的 I/O 网络路径具有最小距离，则过程进入步骤 318，其中为所有具有最小网络路径距离的路径确定结构利用率。在步骤 320 中，确定 I/O 网络路径的最小结构利用率。

[0080] 然后，过程进入步骤 322，其中确定是否存在一个以上的 I/O 网络路径具有最小结构利用率。如果答案是否定的，即仅一个路径，则过程进入步骤 326。步骤 326 从步骤 310、步骤 316 或步骤 322 接收正好一个网络路径。在步骤 326 中，该 I/O 网络路径被标识为“最佳路径”。步骤 326 在 330 结束。

[0081] 如果在步骤 322 中的确定结果是存在一个以上的网络路径具有最小结构利用率，则过程进入步骤 324，其中所谓的第一网络路径被选择。在本发明的一个实施例中，步骤 322 中所确定的所有网络路径被存储在列表中。于是，所谓的第一网络路径是作为列表中第一元素的 I/O 网络路径。本发明的一个可替换实施例是选择列表中最后一个 I/O 网络路径。步骤 324 进入终点。

[0082] 在本发明的一个实施例中，仅步骤 302 到 310、以及步骤 326 或步骤 399 被执行。如果步骤 310 中所确定的路径的数目大于 1，则例如随机地选择一个路径，且方法在步骤 326 以“最佳路径”结束。

[0083] 附图标记列表

[0084]

100	方框图
102	主机系统
103	微处理器
104	应用程序
105	非易失性存储设备
106	负载平衡系统
107	易失性存储设备
108	设备驱动器

109	代理
110	主机总线适配器
111	主机总线适配器
112	网络
113	网络元件
114	网络元件
115	网络元件
118	可移动介质管理系统
120	I/O 设备
121	I/O 设备
125	I/O 设备
126	I/O 设备
130	I/O 端口
131	I/O 端口
132	I/O 端口
133	I/O 端口
134	I/O 端口
140	I/O 端口
141	I/O 端口
142	I/O 端口

[0085]

143	I/O 端口
144	I/O 端口
145	I/O 端口
146	I/O 端口
149	存储子系统
150	I/O 端口
151	I/O 端口
152	I/O 端口
153	I/O 端口
154	I/O 端口
160	I/O 端口
161	I/O 端口
162	I/O 端口
170	网络路径
180	HBA 端口极限
182	HBA 端口利用率
184	网络路径距离
185	路径距离极限
186	结构利用率极限
188	结构利用率
190	网络路径

192	压缩比
194	驱动器负担
199	客户计算机

[0086]

200	流程图
202	开始
204	确定网络路径集合
206	去除禁止路径

[0087]

208	选择候选 I/O 网络路径
210	确定 HBA 端口利用率
212	比较 HBA 端口利用率和 HBA 端口极限并选择
214	确定网络路径距离
216	选择
218	确定结构利用率
220	选择
222	确定 f
224	确定 f 的最低函数值并选择
226	退出 “最佳路径”
229	退出 “无路径”

[0088]

300	流程图
302	开始
304	确定网络路径集合

305	判定
306	确定 HBA 端口利用率
308	确定具有最小 HBA 端口利用率的 I/O 网络路径
310	判定
312	确定网络路径距离
314	确定具有最小网络路径距离的 I/O 网络路径
316	选择
318	确定结构利用率
320	确定最小结构利用率
322	选择
324	选择路径为“最佳路径”
326	确定“最佳路径”
330	结束
399	退出“无路径”

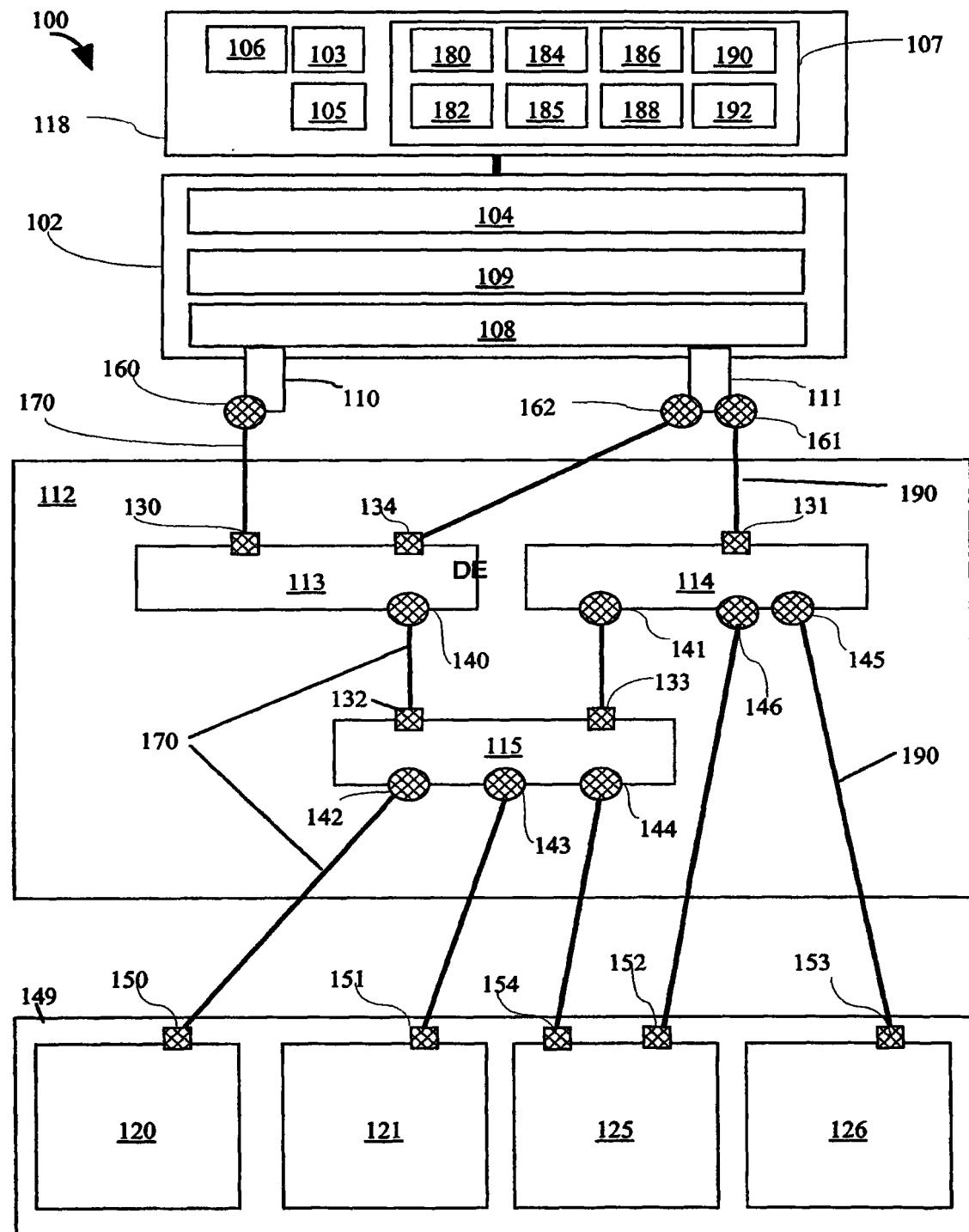


图 1

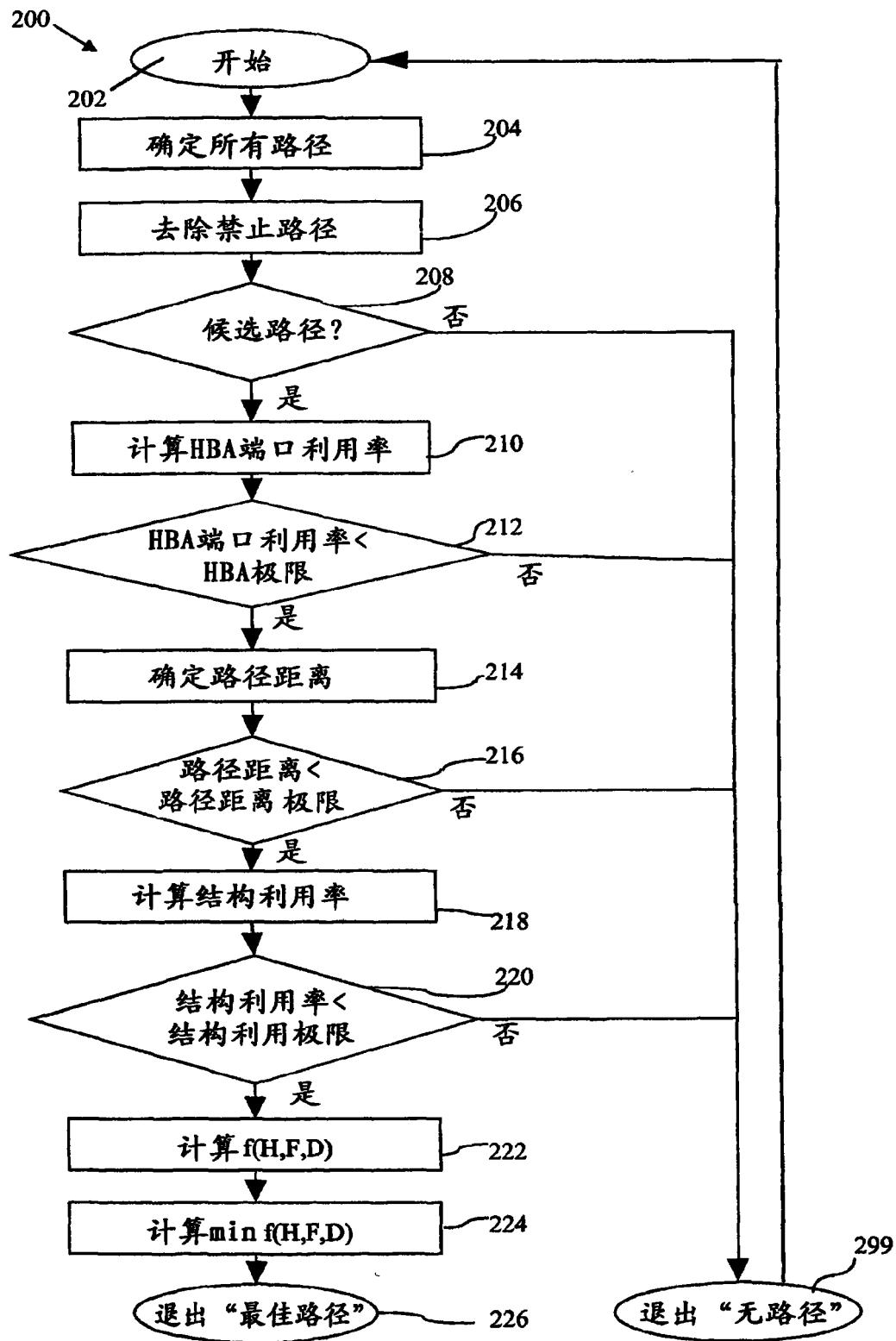


图 2

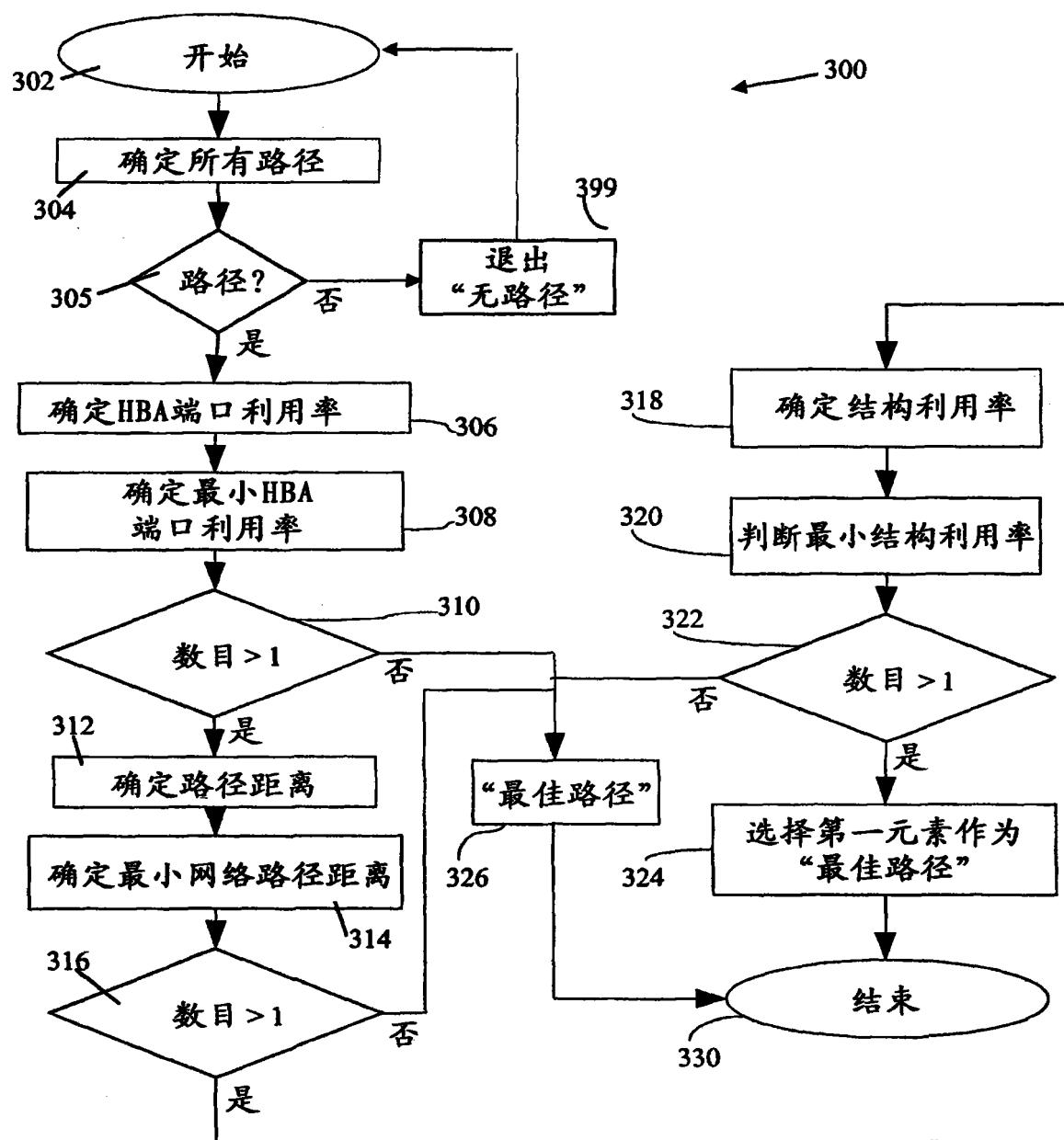


图 3