



(12) 发明专利

(10) 授权公告号 CN 101878617 B

(45) 授权公告日 2013. 01. 16

(21) 申请号 200880118236. 8

代理人 周少杰

(22) 申请日 2008. 11. 14

(51) Int. Cl.

(30) 优先权数据

H04L 12/28 (2006. 01)

11/948, 833 2007. 11. 30 US

(56) 对比文件

(85) PCT申请进入国家阶段日

US 2003/0074449 A1, 2003. 04. 17, 全文.

2010. 05. 28

US 2004/0202155 A1, 2004. 10. 14, 全文.

US 2004/0085902 A1, 2004. 05. 06, 全文.

(86) PCT申请的申请数据

PCT/KR2008/006709 2008. 11. 14

审查员 白芳芳

(87) PCT申请的公布数据

W02009/069907 EN 2009. 06. 04

(73) 专利权人 国际商业机器公司

地址 美国纽约阿芒克

(72) 发明人 罗杰·G·哈索恩 布雷特·W·霍利

马修·J·卡洛斯

(74) 专利代理机构 北京市柳沈律师事务所

11105

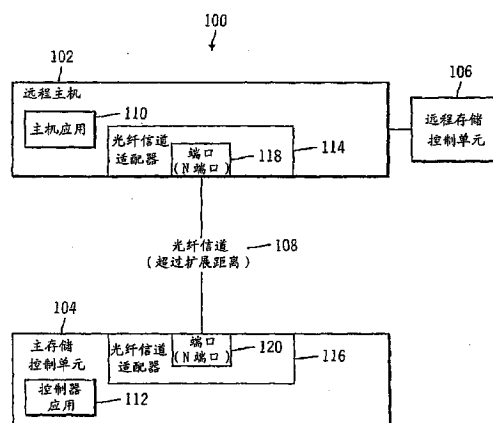
权利要求书 3 页 说明书 9 页 附图 7 页

(54) 发明名称

光纤信道通信中持续信息单元定步协议的使用

(57) 摘要

提供了一种方法、系统和制造产品,其中主存储控制单元通过光纤信道连接从远程主机接收信息单元,其中通过光纤信道连接实现持续信息单元定步。维持关于已经通过至少一个逻辑路径在主存储控制单元处接收多少大型写入的信息,所述至少一个逻辑路径在主存储控制单元和远程主机之间的光纤信道连接上建立,其中大型写入是处理的数据信息单元的数目超过信息单元定步信用的默认值的输入/输出(I/O)操作。主存储控制单元调整信息单元定步参数,所述信息单元定步参数包括在从主存储控制单元发送到远程主机的响应中,其中该调整至少基于维持的、关于已经通过所述至少一个逻辑路径在主存储控制单元处接收多少大型写入的信息。



1. 一种用于光纤信道通信中持续信息单元定步协议的使用的方法,包括:

在主存储控制单元处,通过光纤信道连接从远程主机接收信息单元,其中通过光纤信道连接实现持续信息单元定步;

维持关于已经通过至少一个逻辑路径在主存储控制单元处接收多少大型写入的信息,所述至少一个逻辑路径在主存储控制单元和远程主机之间的光纤信道连接上建立,其中大型写入是处理的数据信息单元的数目超过信息单元定步信用的默认值的输入/输出(I/O)操作;以及

通过主存储控制单元调整信息单元定步参数,所述信息单元定步参数包括在从主存储控制单元发送到远程主机的响应中,指示允许远程主机发送到主存储控制单元、而不用等待来自主存储控制单元的任何额外响应的信息单元的数目,其中该调整至少基于维持的、关于已经通过所述至少一个逻辑路径在主存储控制单元处接收多少大型写入的信息。

2. 如权利要求1所述的方法,其中如果不存在由大型写入导致的写入缓冲器的拥塞,则用于所述至少一个逻辑路径的信息单元定步参数设为大于16但是少于256的数。

3. 如权利要求1所述的方法,其中信息单元是光纤连接协议中的信道命令字,并且其中接收的信道命令字为信道命令字链的开始,该方法还包括:

确定信道命令字是否包括定义子系统操作命令,该定义子系统操作命令定义用于扩展远程复制操作的子系统操作;

确定与定义子系统操作命令相关联的读取记录集命令的数目,其中读取记录集命令对应于读取请求;以及

设置用于逻辑路径的信息单元定步参数值,使得在下一信道程序的开始能够发送增加数目的信息单元,其中如果使用了相同或更小数目的读取记录集命令,则能够立即发送整个信道命令字链。

4. 如权利要求1所述的方法,其中如果多个大型写入导致写入缓冲器的拥塞,则对于其中大型写入超过预定阈值的所有逻辑路径重置信息单元定步参数。

5. 如权利要求1所述的方法,其中远程主机执行从主存储控制单元到远程存储控制单元的数据的扩展远程复制,并且其中远程主机地理上与主存储控制单元分离超过一百千米的距离。

6. 一种用于光纤信道通信中持续信息单元定步协议的使用的系统,包括:

用于在主存储控制单元处,通过光纤信道连接从远程主机接收信息单元的装置,其中通过光纤信道连接实现持续信息单元定步;

用于维持关于已经通过至少一个逻辑路径在主存储控制单元处接收多少大型写入的信息的装置,所述至少一个逻辑路径在主存储控制单元和远程主机之间的光纤信道连接上建立,其中大型写入是处理的数据信息单元的数目超过信息单元定步信用的默认值的输入/输出(I/O)操作;以及

用于通过主存储控制单元调整信息单元定步参数的装置,所述信息单元定步参数包括在从主存储控制单元发送到远程主机的响应中,指示允许远程主机发送到主存储控制单元、而不用等待来自主存储控制单元的任何额外响应的信息单元的数目,其中该调整至少基于维持的、关于已经通过所述至少一个逻辑路径在主存储控制单元处接收多少大型写入的信息。

7. 如权利要求 6 所述的系统,其中还包括:用于如果不存在由大型写入导致的写入缓冲器的拥塞,则用于所述至少一个逻辑路径的信息单元定步参数设为大于 16 但是少于 256 的数的装置。

8. 如权利要求 6 所述的系统,其中信息单元是光纤连接协议中的信道命令字,并且其中接收的信道命令字为信道命令字链的开始,所述操作还包括:

用于确定信道命令字是否包括定义子系统操作命令的装置,该定义子系统操作命令定义用于扩展远程复制操作的子系统操作;

用于确定与定义子系统操作命令相关联的读取记录集命令的数目的装置,其中读取记录集命令对应于读取请求;以及

用于设置用于逻辑路径的信息单元定步参数值,使得在下一信道程序的开始能够发送增加数目的信息单元的装置,其中如果使用了相同或更小数目的读取记录集命令,则能够立即发送整个链。

9. 如权利要求 6 所述的系统,其中还包括:用于如果多个大型写入导致写入缓冲器的拥塞,则对于其中大型写入超过预定阈值的所有逻辑路径重置信息单元定步参数的装置。

10. 如权利要求 6 所述的系统,其中远程主机执行从主存储控制单元到远程存储控制单元的数据的扩展远程复制,并且其中远程主机地理上与主存储控制单元分离超过一百千米的距离。

11. 一种用于部署计算基础设施的方法,包括将计算机可读代码集成到计算系统中,其中结合计算系统的代码能够执行:

在主存储控制单元处,通过光纤信道连接从远程主机接收信息单元,其中通过光纤信道连接实现持续信息单元定步;

维持关于已经通过至少一个逻辑路径在主存储控制单元处接收多少大型写入的信息,所述至少一个逻辑路径在主存储控制单元和远程主机之间的光纤信道连接上建立,其中大型写入是处理的数据信息单元的数目超过信息单元定步信用的默认值的输入/输出(I/O)操作;以及

通过主存储控制单元调整信息单元定步参数,所述信息单元定步参数包括在从主存储控制单元发送到远程主机的响应中,指示允许远程主机发送到主存储控制单元、而不用等待来自主存储控制单元的任何额外响应的信息单元的数目,其中该调整至少基于维持的、关于已经通过所述至少一个逻辑路径在主存储控制单元处接收多少大型写入的信息。

12. 如权利要求 11 所述的用于部署计算基础设施的方法,其中如果不存在由大型写入导致的写入缓冲器的拥塞,则用于所述至少一个逻辑路径的信息单元定步参数设为大于 16 但是少于 256 的数。

13. 如权利要求 11 所述的用于部署计算基础设施的方法,其中信息单元是光纤连接协议中的信道命令字,并且其中接收的信道命令字为信道命令字链的开始,其中结合计算系统的代码能够执行:

确定信道命令字是否包括定义子系统操作命令,该定义子系统操作命令定义用于扩展远程复制操作的子系统操作;

确定与定义子系统操作命令相关联的读取记录集命令的数目,其中读取记录集命令对应于读取请求;以及

设置用于逻辑路径的信息单元定步参数值,使得在下一信道程序的开始能够发送增加数目的信息单元,其中如果使用了相同或更小数目的读取记录集命令,则能够立即发送整个链。

14. 如权利要求 11 所述的用于部署计算基础设施的方法,其中如果多个大型写入导致写入缓冲器的拥塞,则对于其中大型写入超过预定阈值的所有逻辑路径重置信息单元定步参数。

15. 如权利要求 11 所述的用于部署计算基础设施的方法,其中远程主机执行从主存储控制单元到远程存储控制单元的数据的扩展远程复制,并且其中远程主机地理上与主存储控制单元分离超过一百千米的距离。

16. 一种用于光纤信道通信中持续信息单元定步协议的使用的系统,包括:

主存储控制单元;

用于在主存储控制单元处、通过光纤信道连接从远程主机接收信息单元的部件,其中通过光纤信道连接实现持续信息单元定步;

用于维持关于已经通过至少一个逻辑路径在主存储控制单元处接收多少大型写入的信息的部件,所述至少一个逻辑路径在主存储控制单元和远程主机之间的光纤信道连接上建立,其中大型写入是处理的数据信息单元的数目超过信息单元定步信用的默认值的输入/输出(I/O)操作;以及

用于通过主存储控制单元调整信息单元定步参数的部件,所述信息单元定步参数包括在从主存储控制单元发送到远程主机的响应中,指示允许远程主机发送到主存储控制单元、而不用等待来自主存储控制单元的任何额外响应的信息单元的数目,其中该调整至少基于维持的、关于已经通过所述至少一个逻辑路径在主存储控制单元处接收多少大型写入的信息。

17. 如权利要求 16 所述的系统,其中如果不存在由大型写入导致的写入缓冲器的拥塞,则用于所述至少一个逻辑路径的信息单元定步参数设为大于 16 但是少于 256 的数。

18. 如权利要求 16 所述的系统,其中信息单元是光纤连接协议中的信道命令字,并且其中接收的信道命令字为信道命令字链的开始,该系统还包括:

用于确定信道命令字是否包括定义子系统操作命令的部件,其中该定义子系统操作命令定义用于扩展远程复制操作的子系统操作;

用于确定与定义子系统操作命令相关联的读取记录集命令的数目的部件,其中读取记录集命令对应于读取请求;以及

用于设置用于逻辑路径的信息单元定步参数值、使得在下一信道程序的开始能够发送增加数目的信息单元的部件,其中如果使用了相同或更小数目的读取记录集命令,则能够立即发送整个链。

19. 如权利要求 16 所述的系统,其中如果多个大型写入导致写入缓冲器的拥塞,则对于其中大型写入超过预定阈值的所有逻辑路径重置信息单元定步参数。

20. 如权利要求 16 所述的系统,其中远程主机执行从主存储控制单元到远程存储控制单元的数据的扩展远程复制,并且其中远程主机地理上与主存储控制单元分离超过一百千米的距离。

光纤信道通信中持续信息单元定步协议的使用

技术领域

[0001] 本公开涉及用于光纤信道通信中的持续信息单元定步 (pacing) 协议的使用的方法、系统和制造产品。

背景技术

[0002] 光纤信道 (Fibre Channel) 指由美国国家标准协会开发的一组综合的用于数据传送的架构标准。光纤连接 (FICON) 是光纤信道架构的协议, 并且还可以由正式名称 FC-SB-3 来指代。FC-SB-3 的进一步细节可以在 2003 年 3 月 26 日由 American National Standards for Information Technology (美国国家信息技术标准) 出版的出版物“FIBRE CHANNEL Single-Byte Command Code Sets-3 Mapping Protocol (FC-SB-3)” (Rev. 1.6)、以及在 2007 年 9 月 28 日由 American National Standards for Information Technology 出版的出版物“FibreChannel Single-Byte Command Code Sets-3” (AMENDMENT 1 (FC-SB-3/AM1), INCITS 374-2003/AM1-2007) 中找到。

[0003] 信道是各通信设备之间的直接的或切换的点到点连接。在光纤信道架构中, FICON 信道可以执行由 FC-SB-3 指定的功能, 以通过控制单元或仿真控制单元提供对输入 / 输出 (I/O) 设备的访问。FICON 信道可以依赖于用于在各通信设备之间传送数据的分组切换。

[0004] 信道命令字 (CCW) 是包括 I/O 请求的控制块, 并且可以指代特定系统架构的结构, 所述信道命令字与参数一起指定要执行的命令。信道程序 (program) 是控制信道操作的特定序列的、顺序执行的一个或多个信道命令字的序列。FICON 信道可以与用于任何写操作的相关联的数据一起、一次传输达 16 个信道命令字, 其中信道命令字可以称为“信息单元”。如果在信道程序中存在多于 16 个信息单元, 则在第一 16 个信息单元的传输后, 可以以 8 个为组传输剩余的信息单元, 直到信道程序完成。超过 16 的额外信息单元只能在响应于命令响应请求 (CRR) 已经从接收方接收命令响应后发送。要注意的是, 尽管以 8 个为组进行剩余信息单元的传输是常用方式, 但是剩余信息单元的传输不需要必须以 8 个为组。

[0005] 也称为全局镜像 Z (GMz 或 zGM) 的扩展远程复制 (XRC) 是可用于 z/OS* 和 OS/390* 操作系统的复制功能。XRC 在远程位置异步地维持数据的拷贝, 并且可以在扩展距离 (如超过一百千米的距离) 上实现。XRC 可以用于 IBM 企业存储服务器* (ESS)。XRC 的进一步细节可以在 2004 年 7 月由国际商业机器公司出版的出版物“IBM Total Storage Enterprise Storage Server: Implementing ESS Copy Services with IBM eServer zSeries”中找到。

[0006] 扩展距离 XRC 配置可以在远程主机和主站点控制单元之间使用光纤信道扩展技术。该扩展的光纤信道可以由远程主机使用, 以从主站点控制单元读取数据并将备份数据存储在远程站点处的存储上。远程主机可以发出命令链到主站点控制单元, 该命令链包括定义子系统操作 (DSO) 命令, 然后是一系列读取记录集 (read record set, RRS) 命令, 最后是执行子系统功能 (PSF) / 读取子系统数据 (RSSD) 命令集, 以便确定用于下一链的侧文件 (side-file) 入口的数目。可以在 DSO 命令参数中指示链中的 RRS 命令的计数。特定解决方案中的该计数可以是大约 150CCW。

[0007] FICON 架构中的称为“信息单元定步”的流程控制方法可以限制 FICON 信道在任何时间点正在处理 (in flight) 最多 16 个信息单元。信道程序可以通过设置第八个信息单元中的命令响应请求 (CRR) 位, 请求命令响应信息单元。当信道程序从控制单元接收命令响应信息单元时, 发送另外 8 个信息单元。这可能导致在扩展距离 (如超过一百千米的距离) 处的数据速率下降, 因为根据 FICON 定步协议, 控制单元在请求额外单元之前必须接收第一组 16 个信息单元, 因此增加了信道和控制单元端口之间的通信的额外往返。往返的数目依赖于信道命令字链的大小。如果在信道命令字链中存在 150 个信道命令字, 则可能存在达 17 个往返。在扩展距离 (如超过一百千米的距离) 处, 对于通信的额外往返可能增加显著开销。

[0008] 由于这个原因, 所谓的“欺骗”信道扩展器已经增加功能性来检查每个命令, 以确定命令是 DSO/RRS 命令链。如果是, 则远程信道扩展器可以生成 RRS 命令的 CCW 链以仿真预期由本地扩展器接收的命令。数据通过远程扩展器输送到本地扩展器, 然后在本地扩展器处, 对接收的数据执行 CCW 链, 就像它在控制单元中一样。这避免了由长距离要求的额外往返导致的数据速率下降。然而, 信道扩展器中的该功能性可能将客户限制到相对昂贵的信道扩展的解决方案。

发明内容

[0009] 提供了一种方法、系统和制造产品, 其中主存储控制单元通过光纤信道连接从远程主机接收信息单元, 其中通过光纤信道连接实现持续信息单元定步。维持关于通过至少一个逻辑路径、已经在主存储控制单元处接收多少大型写入的信息, 所述至少一个逻辑路径在主存储控制单元和远程主机之间的光纤信道连接上建立, 其中大型写入是处理的数据信息单元的数目超过信息单元定步信用 (credit) 的默认值的输入 / 输出 (I/O) 操作。主存储控制单元调整从主存储控制单元发送到远程主机的响应中包括的信息单元定步参数, 其中该调整至少基于维持的关于通过至少一个逻辑路径已经在主存储控制单元处已经接收多少大型写入的信息。

[0010] 在特定实施例, 信息单元定步参数指示允许远程主机发送到主存储控制单元、而不用等待来自主存储控制单元的任何额外响应的信息单元的数目, 并且其中如果不存在由大型写入导致的写入缓冲器的拥塞, 则用于至少一个逻辑路径的信息单元定步参数设为大于 16 但是少于 256 的数。

[0011] 在另外的实施例, 信息单元是光纤连接协议中的信道命令字, 并且其中接收的信道命令字开始信道命令字链。确定信道命令字是否包括定义子系统操作命令, 该定义子系统操作命令定义用于扩展远程复制操作的子系统操作。确定与定义子系统操作命令相关联的读取记录集命令的数目, 其中读取记录集命令对应于读取请求。设置用于逻辑路径的信息单元定步参数值, 使得在下一信道程序的开始可以发送增加数目的信息单元, 其中如果使用了相同或更小数目的读取记录集命令, 则可以立即发送整个链。

[0012] 在另外的实施例, 如果多个大型写入导致写入缓冲器的拥塞, 则对于其中大型写入超过预定阈值的所有逻辑路径重置信息单元定步参数。

[0013] 在特定实施例, 远程主机执行从主存储控制单元到远程存储控制单元的数据的扩展远程复制, 其中远程主机地理上与主存储控制单元分离超过一百千米的距离。

附图说明

- [0014] 现在参考附图,附图中贯穿各图相同的参考标号表示对应的部分;
- [0015] 图 1 图示根据特定实施例的计算环境的方块图;
- [0016] 图 2 图示根据特定实施例的、示出在远程主机和主存储控制器之间的通信的方块图;
- [0017] 图 3 图示根据特定实施例的、示出在信道和控制单元之间的通信的方块图;
- [0018] 图 4 图示根据特定实施例的、与光纤信道连接相关联的数据结构;
- [0019] 图 5 图示根据特定实施例的、与扩展远程复制相关联的数据结构;
- [0020] 图 6 图示根据特定实施例的、示出如何在操作环境中设置 IU 定步参数的方块图;
- [0021] 图 7 图示根据特定实施例的、示出用于设置 IU 定步参数的第一操作的流程图;
- [0022] 图 8 图示示出用于设置 IU 定步参数的第二操作的流程图;以及
- [0023] 图 9 图示计算系统的架构,其中在特定实施例中,可以根据计算系统的架构实现图 1 的计算环境的元件。

具体实施方式

[0024] 在下面的描述中,参考附图,该附图形成描述的一部分并图示若干实施例。要理解的是,可以利用其它实施例并且可以进行结构和操作的改变。

[0025] 特定示例实施例的概述

[0026] 持续 IU 定步协议允许信道一次发送多于 16 个信息单元到控制单元,因此减轻了扩展距离处的命令相关的数据速率下降。使用持续 IU 定步协议的特定实施例允许扩展距离远程复制 (XRC) 环境中的主存储控制单元修改 FICON 架构内的信息单元流,使得在某些时间点可以正在处理多于 16 个信息单元。在光纤信道网络中使用 FICON 的扩展距离远程复制实现中,远程主机可以执行扩展距离远程复制操作,以将数据从主存储控制单元复制到远程存储控制单元。通过允许在某些时间点正在处理多于 16 个信息单元,与其中限制信道在任何时间点不能正在处理多于 16 个信息单元的情况相比,特定实施例可以减少用于执行扩展远程复制的时间。

[0027] 特定实施例通过执行以下操作,在 FICON 架构上实现了可能干扰 XRC 实现中的读取的写入操作的拥塞控制:

[0028] (i) LW (大型写入) 计数跟踪每个逻辑路径的大型写入。大型写入定义为处理的数据 IU 的数目超过默认 IU 定步信用的 I/O 操作。

[0029] (ii) 每次在逻辑路径上接收具有 RRS 次等级 (suborder) 的 DSO 时,递减 LW 计数。如果此时 LW 计数达到 0,则将该逻辑路径的定步信用设为比 DSO 命令参数中指示的 RRS 命令的数目多 3,服从最大定步信用为 255。

[0030] (iii) 在链的末端,如果该链的写入 IU 计数超过默认定步信用 (或可调整的值),则递增 LW 计数。如果 LW 计数达到最大值,则将该逻辑路径放在大型写入阈值 (LWT) 列表上。为了处理非常大的 LW 链,可以将 LW 计数递增对于该链的 LW 计数除以默认定步计数的值。

[0031] (iv) 在 I/O 操作的结束,如果写入缓冲器拥塞已经出现,则对 LWT 列表上的所有逻辑

辑路径重置定步计数。

[0032] (v) LW 计数不允许递减到 0 以下或递增到选择的最大值以上。

[0033] (vi) 在链的开始, 在命令响应 (CMR) 中使用对于逻辑路径的当前存储的定步信用, 并且将对接收该 CMR 后通过信道开始的下一链起作用。

[0034] 示例性优选实施例

[0035] 图 1 图示利用耦合到主存储控制单元 104 和远程存储控制单元 106 的远程主机 102 的计算环境 100 的方块图。尽管图 1 只示出单个远程主机 102、单个主存储控制单元 104 和单个远程存储控制单元 106, 但是在特定替代实施例中, 多个远程主机可以耦合到多个主存储控制单元和远程存储控制单元。

[0036] 远程主机 102 可以通过数据接口信道 (如光纤信道 108 或本领域已知的任何其它数据接口机制) 连接到主存储控制单元 104。远程主机 102 可以是本领域目前已知的任何适合的计算设备, 如个人计算机、工作站、服务器、大型机、手持式计算机、电话设备、网络设备等。远程主机 102 可以包括本领域已知的任何操作系统, 如 IBM OS/390 或 z/OS 操作系统。

[0037] 主存储控制单元 104 和远程存储控制单元 106 可以包括多个逻辑容量 (volume)。主存储控制单元 104 和远程存储控制单元 106 可以控制多个物理存储设备, 每个物理存储设备可以包括一个或多个物理容量。

[0038] 远程主机 102 可以包括主机应用 110, 并且主存储控制单元 104 可以包括控制器应用 112。主机应用 110 与控制器应用 112 接口, 以从主存储控制单元 104 读取数据并将数据存储在远程存储控制单元 106 中。主机应用 110 和控制器应用 112 通过光纤信道 108 通信。在特定实施例中, 主机应用 110 通过光纤信道 108 使用扩展远程复制, 以将数据从主存储控制单元 104 复制到远程存储控制单元 106。

[0039] 通过光纤信道 108 在远程主机 102 和主存储控制单元 104 之间的通信, 可以通过远程主机 102 中包括的光纤信道适配器 114 和主存储控制单元 104 中包括的光纤信道适配器 116 来启用。远程主机 102 中包括的光纤信道适配器 114 包括端口 118, 并且主存储控制单元 104 中包括的光纤信道适配器 116 包括端口 120, 其中端口 118 和 120 在光纤信道术语中可以称为 N 端口。经由 FICON 协议的基于光纤信道的通信可以在远程主机 102 的端口 118 和主存储控制单元 104 的端口 120 之间执行。在两个光纤信道适配器 114 和 116 之间可以建立多个逻辑路径。

[0040] 因此, 图 1 图示了计算环境 100, 其中主机应用 110 通过使用扩展远程复制操作, 经由 FICON 协议将数据从主存储控制单元 104 复制到远程存储控制单元 106。在特定实施例中, 通过使用光纤信道扩展解决方案, 远程主机 102 和主存储控制单元 104 之间的距离可以超过光纤信道架构支持的最大距离。

[0041] 图 2 图示根据计算环境 100 中实现的特定实施例、示出在远程主机 102 和主存储控制器 104 之间的通信的方块图。

[0042] 远程主机 102 能够在远程主机 102 和主存储控制单元 104 之间根据 FICON 协议生成的信道上、顺序地发送一个或多个信道命令字 400。主存储控制单元 104 可以响应于特定信道命令字 200 发送命令响应信息单元 202。在特定实施例中, 通过调整命令响应信息单元 202 中包括的 IU 定步参数 (如 IU 定步参数 208), 主存储控制单元 104 可以修改可以在远

程主机 102 和主存储控制单元 104 之间正在处理的信道命令字的数目。

[0043] 图 3 图示根据特定实施例的、示出计算环境 100 中实现的信道 300 和控制单元 302 之间的通信的方块图。

[0044] 在特定实施例中,信道 300 是远程主机 102 的实体,并且包括端口 118。控制单元 302 是存储控制器 104 的实体,并且包括端口 120。可以在信道 300 和控制单元 302 之间建立逻辑路径。

[0045] 在特定实施例中,信道发送建立逻辑路径 (ELP) 请求 304 到控制单元 302,请求在信道 300 和控制单元 302 之间建立逻辑路径。响应于接收建立逻辑路径请求 304,控制单元 302 可以发送逻辑路径建立 (LPE) 响应 306 给信道 300,并且建立逻辑路径。

[0046] 在根据 FICON 架构实现持续 IU 定步的特定实施例中,当通过使用建立逻辑路径请求 304 和逻辑路径建立响应 306 建立逻辑路径时,如果控制单元 302 指示支持 IU 定步参数值的保持,则信道 300 可以保持随后的命令链的 IU 定步参数的值。

[0047] 图 4 图示根据特定实施例的、在计算环境 100 中通过光纤信道 108 实现的与光纤连接有关的数据结构。图 4 所示的数据结构称为 FICON 数据结构 400。

[0048] FICON 数据结构 400 包括一个或多个信道命令字 402、信息单元定步信用 404、具有信息单元定步参数 408 的命令响应信息单元 406,其中信息单元定步参数 408 也称为 IU 定步参数,并且信息单元定步信用 404 可称为 IU 定步信用。FICON 数据结构 400 还包括建立逻辑路径请求 304 和逻辑路径建立响应 306。

[0049] 信道命令字 402 是包括 I/O 请求的控制块。例如,在特定实施例中,信道命令字 402 可以包括从主机应用 110 到控制器应用 112 的读取请求,其中该读取请求是读取由存储控制器 104 存储的数据的请求。信道命令字 402 可以从信道 300 发送到控制单元 302。信道命令字 402 还可以称为信息单元。

[0050] 用于远程主机 102 和存储控制器 104 之间的光纤信道通信的每个信道 300 提供 IU 定步信用 404,其可以在每个信道程序的开始时或在重新连接以继续信道程序的执行期间初始化。IU 定步信用 404 是远程主机 102 从存储控制器 104 接收命令响应信息单元 406 之前、远程主机 102 可以发送给存储控制器 104 的信息单元的最大数目。

[0051] 命令响应信息单元 406 是响应于特定条件、从存储控制器 104 发送给远程主机 102 的信息单元。例如,命令响应信息单元 406 可以响应于特定信道命令字 402,从存储控制器 104 发送给远程主机 102。与命令响应信息单元 406 相关联的 IU 定步参数 408 可以从存储控制器 104 的端口 120 发送,以指示远程主机 102 可以通过信道发送的信息单元的最大数目。IU 定步参数 408 为零指示 IU 定步信用的值要重置为默认值。

[0052] 在信道程序的开始时或在每次重新连接时,信道 300 可以发送多个信息单元给控制单元 302。发送的信息单元的数目不可以超过信息单元定步信用 404 的值,其中信息单元定步信用 404 的值也称为信息单元定步信用值。

[0053] 建立逻辑路径请求 304 可以从信道 300 发送给控制单元 302,以指示信道 300 支持的可选特征,并且请求在信道 300 和控制单元 302 之间建立逻辑路径。来自控制单元 302 的逻辑路径建立响应 306 确认建立逻辑路径功能请求的成功完成和逻辑路径的建立,并且指示要在信道 300 和控制单元 302 之间发送的所有信息单元上使用的可选特征。

[0054] 建立逻辑路径请求 304 可以经由具有称为建立逻辑路径参数 308 的相关参数的建

立逻辑路径功能实现。建立逻辑路径参数 308 中包括的持续定步控制指示符位 310 可指示信道 300 是否提供对持续定步的支持,即,跨越多个命令链的信息单元定步参数 408 的值的保持,其中命令链是信道命令字的序列。持续定步也可以称为持续信息单元定步。

[0055] 逻辑路径建立响应 306 可以经由具有称为逻辑路径建立参数 312 的相关参数的功能实现。逻辑路径建立参数 312 中包括的持续定步控制响应指示符位 314 可以指示控制单元 302 是否提供对持续定步的支持。

[0056] 因此,图 4 图示特定实施例,其中称为持续定步控制指示符位 310 的第一指示符和称为持续定步控制响应指示符位 314 的第二指示符,指示信道 300 和控制单元 302 之间建立的逻辑路径是否启用持续定步。在特定实施例中,包括第一指示符 310 和第二指示符 314 作为对光纤信道协议的增强,其中对光纤信道协议的增强支持跨越多个命令链的持续信息单元定步。

[0057] 图 5 图示根据特定实施例的、与计算环境 100 中实现的扩展远程复制相关联的数据结构。图 5 所示的数据结构称为 XRC 数据结构 500。

[0058] XRC 数据结构 500 可以包括定义子系统操作 (DSO) 命令 502,其定义发信号传输执行当前命令链中的多个读取记录集 (RRS) 信道命令字的意图的子系统操作。在扩展远程复制操作期间,DSO 命令 502 可以在远程主机 102 和主存储控制单元 104 之间的通信期间定义子系统操作。RRS 命令 504 如果与 DSO 命令 502 相关联,则对主存储控制单元 104 指示远程主机 102 正在发送读取请求。

[0059] 扩展距离 XRC 配置可以使用远程主机 102 和主存储控制单元 104 之间的光纤信道扩展技术。扩展光纤信道可以由远程主机 102 用来从主存储控制单元 104 读取数据,并且将数据存储于远程存储控制单元 106 处的存储。在特定实施例中,从主存储控制单元 104 读取的数据作为备份数据存储于远程存储控制单元 106 中。

[0060] 在特定实施例中,远程主机 102 发出命令链到主存储控制单元 104,其中命令链包括 DSO 命令 502,然后是一系列读取记录集命令 504,最后是额外的命令 [例如,如 XRC 中定义的执行子系统功能 (PSF) 命令和读取子系统数据 (RSSD) 命令],以便确定用于下一链的侧文件入口的数目。链中的 RRS 命令 504 的计数在 DSO 命令参数中指示,并且在特定示例性实施例中,RRS 命令 504 的计数可以超过 150 信道命令字。

[0061] 因此,图 5 图示特定实施例,其中与 DSO 命令 502 相关联的读取记录集命令 504 对主存储控制单元 104 指示远程主机 102 请求了读取操作。在实现扩展远程复制的实施例中,这种 DSO 命令 502 可以从远程主机 102 发送到主存储控制单元 104。

[0062] 图 6 图示根据特定实施例的、示出如何在操作环境 600 中设置 IU 定步参数 408 的方块图。操作环境 600 可以出现在计算环境 100 中。

[0063] 在操作环境中,光纤信道上的 XRC 环境 (参考标号 602) 可在远程主机 102 和主存储控制单元 104 之间操作。启用持续信息单元定步 (参考标号 604)。然而,除了具有读取记录集命令 504 的 DSO 命令 502 外,干扰写入命令 606 也从远程主机 102 发送到主存储控制单元 104。干扰写入命令 606 可以是大型写入,其中大型单元定义为处理的数据信息单元的数目超过默认 IU 定步信用的 I/O 操作。

[0064] 在操作环境 600 中,作为 XRC 环境的 DSO 命令中的读取记录集命令 504 的功能,设置 IU 定步参数 408 (参考标号 608)。因此,当启用持续 IU 定步时,可以同时正在处理多于

16 个信息单元。然而,大型写入命令 606 可能导致干扰和问题。因此,在操作环境 600 中,基于可以处理的大型写入的最大数目跟踪大型写入的数目,并且对于存在太多大型写入的逻辑路径重置 IU 定步参数(参考标号 610)。其可以看作是一种形式的用于写入命令的拥塞控制,使得通过重置 IU 定步参数,在流控制期间限制了干扰写入命令的数目。

[0065] 图 7 图示根据特定实施例的、示出用于设置 IU 定步参数的第一操作的流程图。第一操作可以经由控制器应用 112 和主机应用 110,在计算环境 100 的远程主机 102 和主存储控制单元 104 中实现。

[0066] 控制在块 700 开始,其中通过光纤信道 108 在远程主机 102 和主存储控制单元 104 之间形成的信道处于空闲状态,并且信道命令字链开始(参考标号 702)。

[0067] 进行关于是否启用持续 IU 定步和第一个接收的命令是否具有读取记录集次等级的 DSO 命令的确定(在块 704)。如果是,则递减大型写入计数(在块 706)但不允许其下降至 0 以下。进行关于大型写入计数是否为 0 的确定(在块 708),并且如果是,则进行是否存在由大型写入导致的缓冲器的拥塞的确定(在块 710)。如果否,则在块 712,将对应逻辑路径的 IU 定步参数设为比 DSO 命令参数中指示的读取记录集命令的数目多 3。然而,IU 定步参数不能设为多于 255。

[0068] 如果在块 708 进行确定大型写入计数不为 0,则控制进行到块 718,其中进行关于对应逻辑路径是否在大型写入阈值列表上的确定。如果是,则从大型写入阈值列表移除该逻辑路径(在块 720)。

[0069] 如图 7 所示,控制可以从块 704、720、718、710 和 712 的任何进行到块 714,如图 7 提供的每个指示。在块 714,发送具有用于逻辑路径的当前 IU 定步参数的命令响应信息单元。控制进行到块 716,其中执行命令链,同时通过递增写入 IU 计数来跟踪写入信息单元,以便跟踪命令链中的写入信息单元的数目。要注意的是,在 I/O 操作的结束,如果已经出现写入缓冲器拥塞,则对大型写入阈值列表上的所有逻辑路径重置定步计数。

[0070] 图 8 图示示出用于设置 IU 定步参数的第二操作的流程图。第二操作可以经由控制器应用 112 和主机应用 110,在计算环境 100 的远程主机 102 和主存储控制单元 104 中实现。

[0071] 控制在块 800 开始,其中主存储控制单元 104 通过光纤信道连接 108 从远程主机 102 接收信息单元,其中通过光纤信道连接 108 实现持续信息单元定步。

[0072] 控制进行到块 802,其中维持关于通过至少一个逻辑路径、已经在主存储控制单元 104 处接收多少大型写入的信息,所述至少一个逻辑路径在主存储控制单元 104 和远程主机 102 之间的光纤信道连接 108 上建立,其中大型写入是处理的数据信息单元的数目超过信息单元定步信用的默认值的输入/输出(I/O)操作。

[0073] 在块 804,主存储控制单元 104 调整从主存储控制单元 104 发送到远程主机 102 的响应中包括的信息单元定步参数 408,其中该调整至少基于维持的关于通过至少一个逻辑路径已经在主存储控制单元 104 处接收多少大型写入的信息。

[0074] 控制可以从块 804 进行到块 806 和块 808。在块 806,如果不存在由大型写入导致的写入缓冲器的拥塞,则将用于至少一个逻辑路径的信息单元定步参数设为大于 16 但是小于 256 的数目。在块 808,如果多个大型写入导致写入缓冲器的拥塞,则对其中大型写入超过预定阈值的所有逻辑路径重置信息单元定步参数。

[0075] 特定实施例允许通过光纤信道的 XRC 实现,使得在启用持续 IU 定步的同时执行对于大型写入的拥塞控制。优选的是在与 XRC 读取操作相同的路径上不具有一直运行大型写入的应用,因为这种大型写入可能干扰持续定步的性能益处。特定实施例的大型写入检测设计为在需要时允许一些时段的大型写入 I/O。

[0076] 本发明的模式

[0077] 另外的实施例细节

[0078] 所述技术可以实现为方法、装置或涉及软件、固件、微代码、硬件和 / 或其任何组合的制造产品。如这里使用的术语“制造产品”指介质中实现的代码或逻辑,其中这种介质可以包括硬件逻辑 [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,如磁存储介质 (例如,硬盘驱动、软盘、带等)、光存储 (CD-ROM、光盘等)、易失性和非易失性存储器设备 [例如,电可擦除可编程只读存储器 (EEPROM)、只读存储器 (ROM)、可编程只读存储器 (PROM)、随机存取存储器 (RAM)、动态随机存取存储器 (DRAM)、静态随机存取存储器 (SRAM)、闪存、固件、可编程逻辑等]。计算机可读介质中的代码由处理器访问和执行。其中编码代码或逻辑的介质还可以包括通过空间或传输介质 (如光纤、铜线等) 传播的传输信号。其中编码代码或逻辑的传输信号还可以包括无线信号、卫星传输、无线电波、红外信号、蓝牙等。其中编码代码或逻辑的传输信号能够由发送站发送和由接收站接收,其中传输信号中编码的代码或逻辑可以在接收和发送站或设备被解码和存储在硬件或计算机可读介质中。此外,“制造产品”可以包括其中体现、处理和执行代码的硬件和软件组件的组合。当然,本领域技术人员将认识到,可以进行许多修改而不偏离实施例的范围,并且制造产品可以包括任何信息承载介质。例如,制造产品包括其中存储指令的存储介质,该指令在由机器执行时导致执行操作。

[0079] 特定实施例可以采取完全硬件实施例、完全软件实施例或包含硬件和软件元素的实施例的形式。在优选实施例中,本发明以软件实现,其包括但不限于固件、驻留软件、微代码等。

[0080] 此外,特定实施例可以采用从计算机可用或计算机可读介质可访问的计算机程序产品的形式,该计算机可用或计算机可读介质提供程序代码由计算机或任何指令执行系统使用或结合计算机或任何指令执行系统使用。为了该描述的目的,计算机可用或计算机可读介质可以是包含、存储、传递、传播或传送程序以用于由指令执行系统、装置或设备使用或结合指令执行系统、装置或设备使用的任何装置。该介质可以是电、磁、光、电磁、红外或半导体系统 (或装置或设备) 或传播介质。计算机可读介质的示例包括半导体或固态存储器、磁带、可移除计算机盘、随机存取存储器 (RAM)、只读存储器 (ROM)、刚性磁盘和光盘。光盘的当前示例包括致密盘只读存储器 (CD-ROM)、致密盘 - 读 / 写 (CD-R/W) 和 DVD。

[0081] 术语“特定实施例”、“一实施例”、“实施例”、“各实施例”、“该实施例”、“所述各实施例”、“一个或多个实施例”、“一些实施例”和“一个实施例”意味着一个或多个 (但不是所有) 实施例,除非明确另外指定。术语“包括”、“包含”、“具有”及其变化意味着“包括但不限于”,除非明确另外指定。列举的项目列表不暗示任何或所有项目相互排他,除非明确另外指定。术语“一”、“一个”和“该”意味着“一个或多个”,除非明确另外指定。

[0082] 相互通信的设备不需要相互持续通信,除非明确另外指定。此外,相互通信的设备可以直接或通过一个或多个中间物间接通信。此外,具有相互通信的若干组件的实施例的

描述不暗示要求所有这些组件。相反,描述了多个可选组件来图示种类广泛的可能的实施例。

[0083] 此外,尽管可以以序列顺序描述处理步骤、方法步骤、算法等,但是这种处理、方法和算法可以配置为以交替顺序工作。换句话说,可描述的步骤的任何序列或顺序不一定指示要求各步骤以该顺序执行。在此描述的处理步骤可以以任何实际顺序执行。此外,一些步骤可以同时、并行或并发执行。

[0084] 当在此描述单个设备或产品时,明显的是可以使用多于一个设备 / 产品 (无论它们是否合作) 来代替单个设备 / 产品。类似地,在这里描述多于一个设备或产品的情况下 (无论它们是否合作),将明显的是可以使用单个设备 / 产品来代替该多于一个设备或产品。设备的功能性和 / 或特征可以由没有明确描述为具有这种功能性和 / 或特征的一个或多个其它设备替代地体现。因此,其它实施例不需要包括设备自身。

[0085] 图 9 图示示例性计算机系统 900,其中在特定实施例中,图 1 的计算环境 100 的远程主机 102、主存储控制器 104 和远程存储控制器 106 可以根据计算机系统 900 的计算机架构而实现。计算机系统 900 还可以称为系统,并且可以包括电路 902,其在特定实施例中可以包括处理器 904。系统 900 还可以包括存储器 906 (例如,易失性存储器设备) 和存储 908。系统 900 的特定元件可以或可以不在图 1 的远程主机 102、主存储控制器 104 或远程存储控制器 106 中找到。存储 908 可以包括非易失性存储器设备 (例如,EEPROM、ROM、PROM、RAM、DRAM、SRAM、闪存、固件、可编程逻辑等)、磁盘驱动、光盘驱动、带驱动等。存储 908 可以包括内部存储设备、外接存储设备和 / 或网络可访问存储设备。系统 900 可以包括包含代码 912 的程序逻辑 910,该代码可以加载到存储器 906 并由处理器 904 或电路 902 执行。在特定实施例中,包括代码 912 的程序逻辑 910 可以存储在存储 908 中。在其他特定实施例中,程序逻辑 910 可以在电路 902 中实现。因此,尽管图 9 示出程序逻辑 910 与其它元件分离,但是程序逻辑 910 可以在存储器 906 和 / 或电路 902 中实现。

[0086] 特定实施例可能贯注于用于通过人或自动处理将计算机可读代码集成到计算系统中而部署计算指令的方法,其中启用与计算系统结合的代码以执行上述实施例的操作。

[0087] 图 9 所示的至少某些操作可以并行以及顺序执行。在替代实施例中,某些操作可以以不同顺序执行、被修改或删除。

[0088] 此外,为了说明目的已经以分开模块描述了许多软件和硬件组件。这种组件可以集成到更少数目的组件或划分为更多数目的组件。另外,描述为由特定组件执行的某些操作可以由其它组件执行。

[0089] 图 1-9 中示出或引用的数据结构和组件描述为具有特定类型的信息。在替代实施例中,各数据结构和组件可以不同地构造,并且具有比图中示出或引用的那些更少、更多或不同字段或不同功能。

[0090] 因此,为了说明和描述的目的已经呈现了本发明的前面描述。意图不在于穷尽性的或将实施例限于公开的精确形式。考虑上述教导,许多修改和变形是可能的。

[0091] *z/OS、OS/390 和企业存储服务器是国际商业机器公司的商标或注册商标。

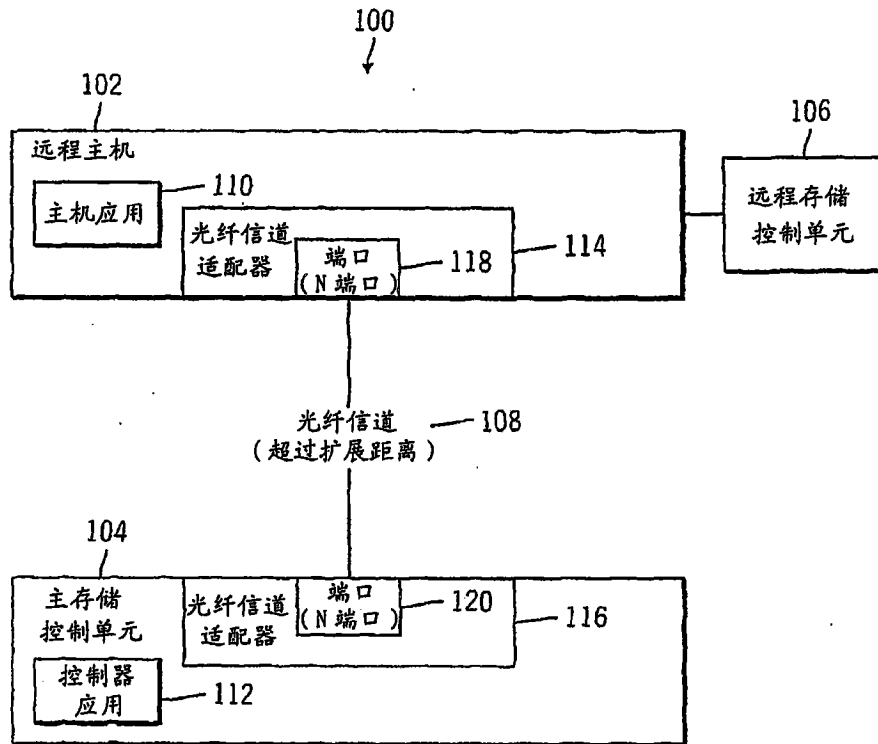


图 1

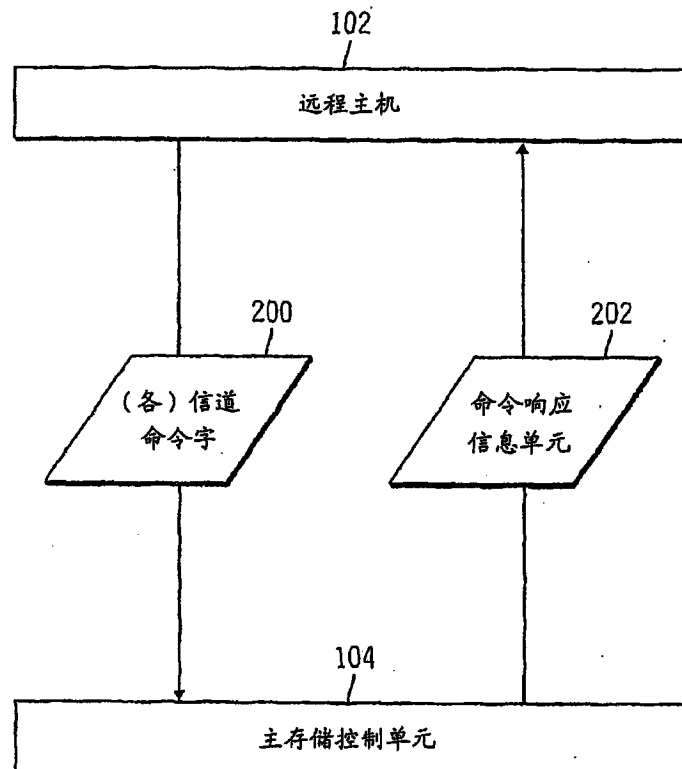


图 2

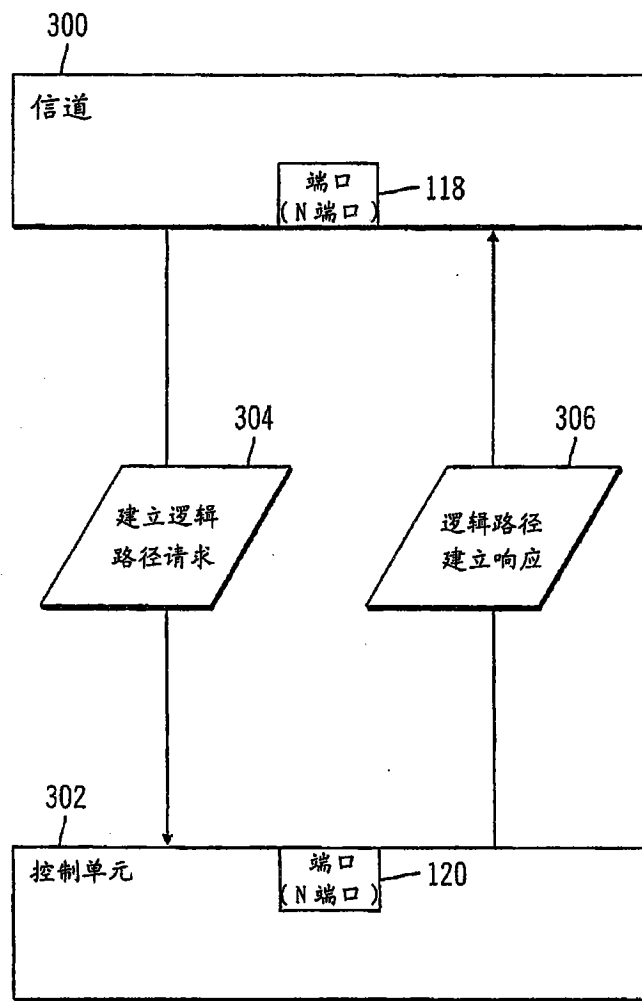


图 3

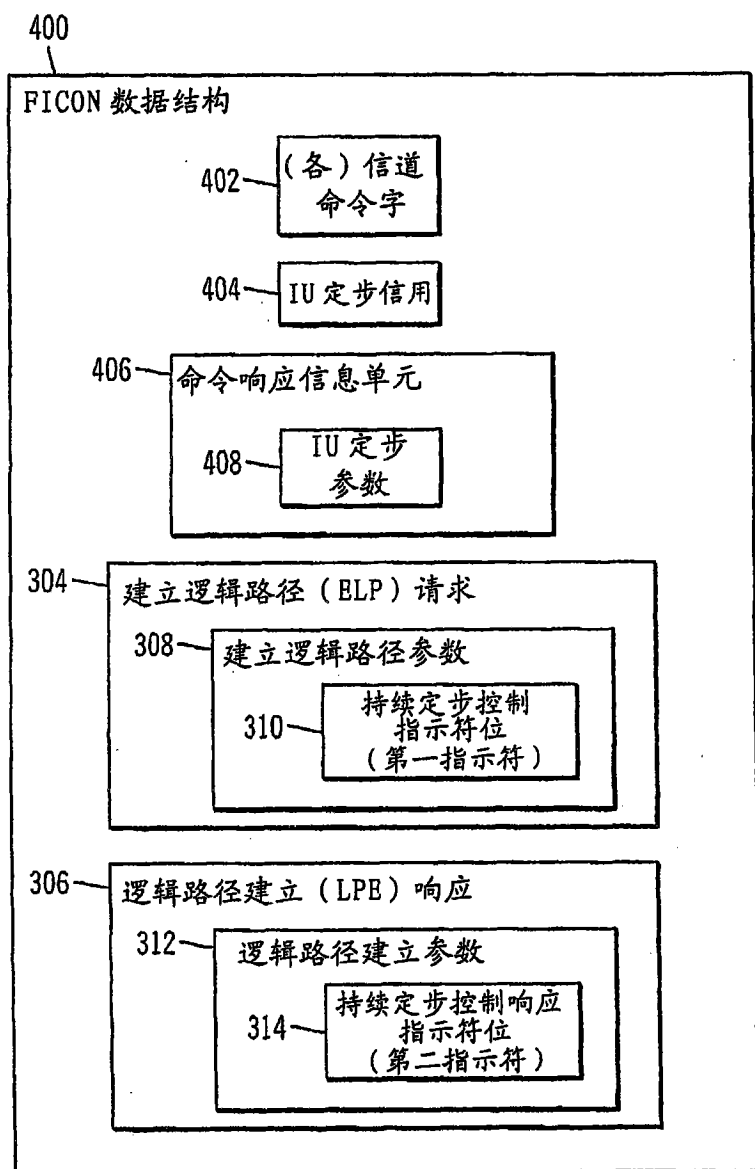


图 4

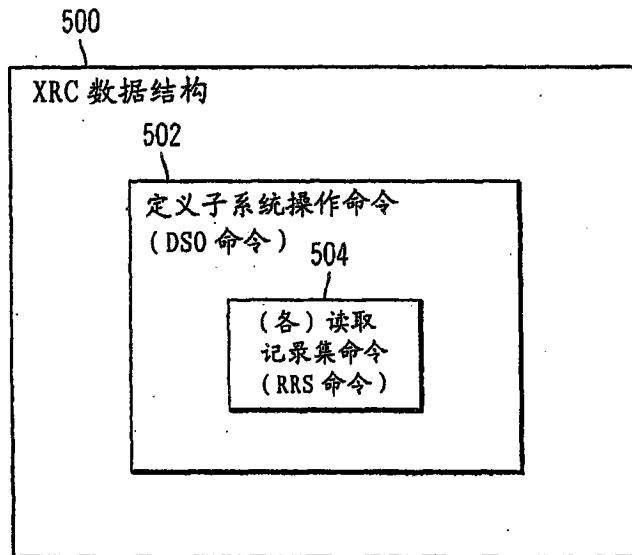


图 5

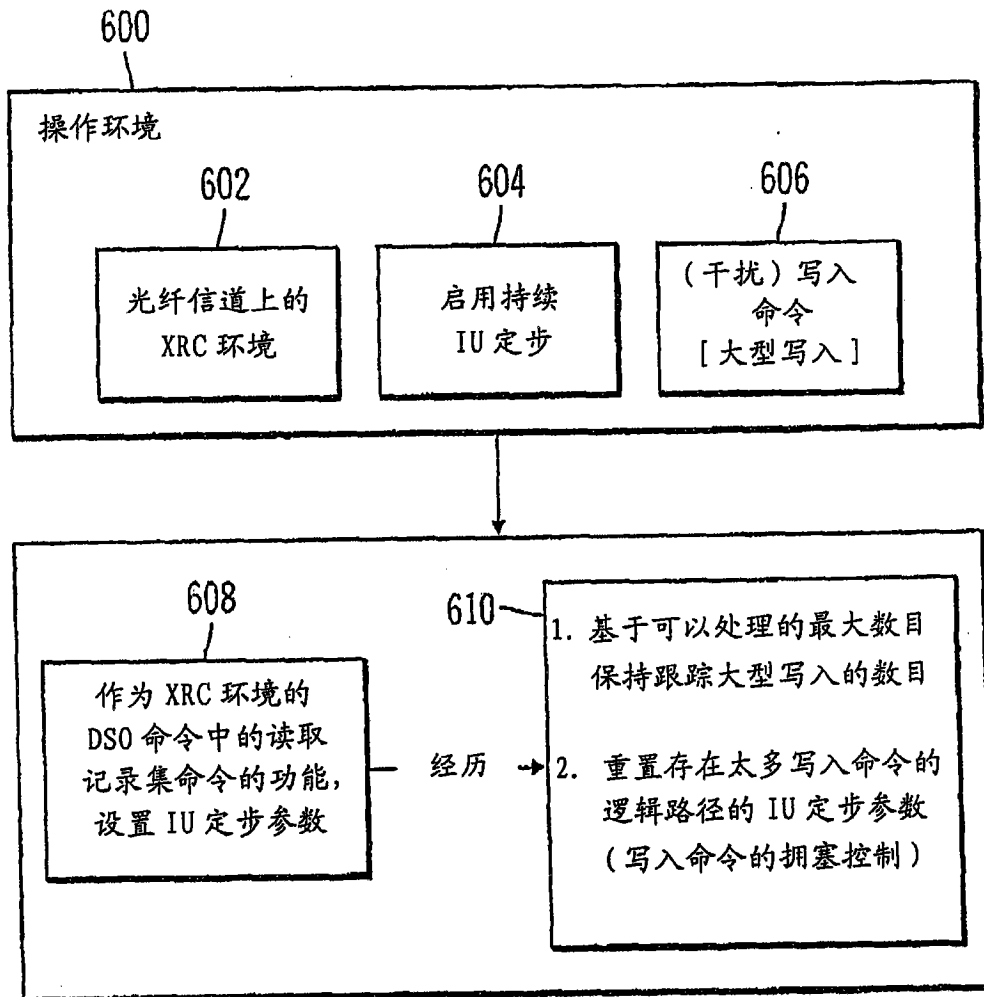


图 6

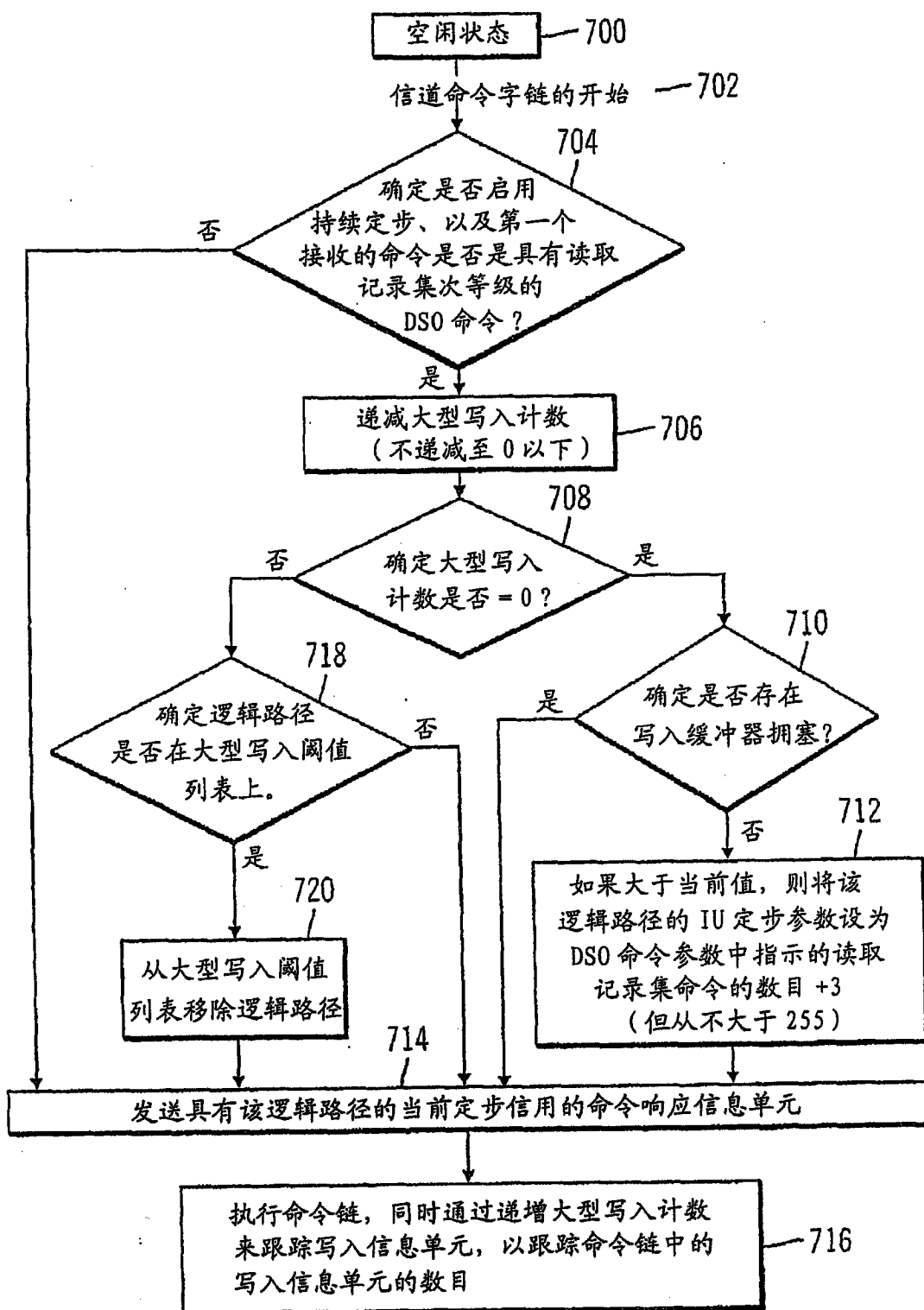


图 7

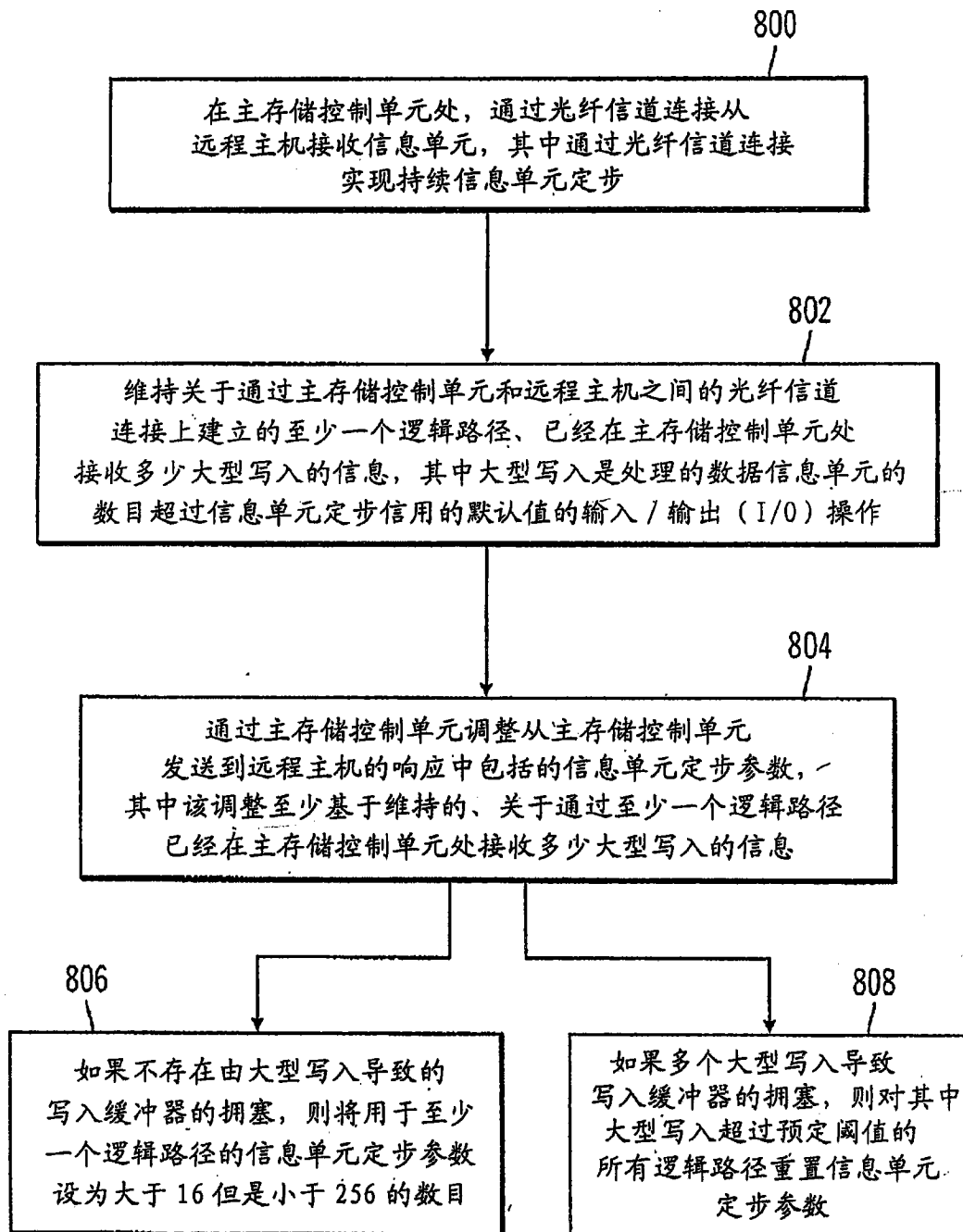


图 8

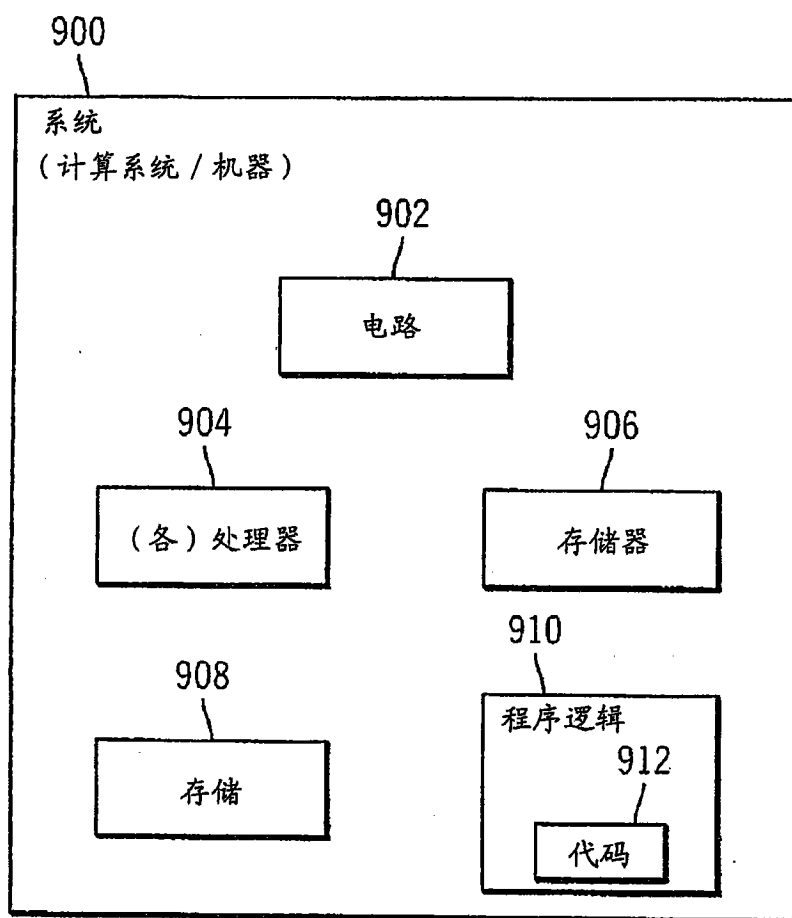


图 9