

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06F 12/08 (2006.01)

G06F 9/46 (2006.01)

H04L 12/56 (2006.01)

H04L 12/24 (2006.01)



[12] 发明专利说明书

专利号 ZL 03807779.5

[45] 授权公告日 2007 年 5 月 23 日

[11] 授权公告号 CN 1317647C

[22] 申请日 2003.3.25 [21] 申请号 03807779.5

[30] 优先权

[32] 2002. 4. 5 [33] US [31] 10/118,386

[86] 国际申请 PCT/US2003/009332 2003. 3. 25

[87] 国际公布 WO2003/088050 英 2003. 10. 23

[85] 进入国家阶段日期 2004. 10. 8

[73] 专利权人 思科技术公司

地址 美国加利福尼亚州

[72] 发明人 克劳迪奥·德桑蒂

马尔科·迪贝内德托

[56] 参考文献

EP1011231A2 1999. 12. 15

CN1310925A 2001. 8. 29

WO0055750A1 2000. 9. 21

WO0195113A 2001. 12. 13

审查员 张 蕾

[74] 专利代理机构 北京东方亿思知识产权代理有限公司

代理人 王 怡

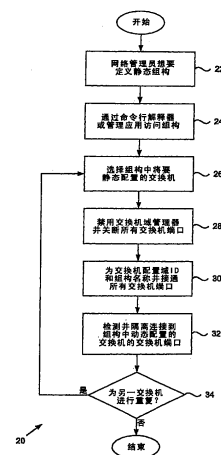
权利要求书 4 页 说明书 7 页 附图 3 页

[54] 发明名称

用于定义静态光纤通道组构的装置与方法

[57] 摘要

公开了一种存储区域网和用于定义不需要主要交换机的静态光纤通道组构的方法。该存储区域网包括一个或多个宿主机、一个或多个存储设备以及连接所述一个或多个宿主机和存储设备的静态组构。在该静态组构中，这些交换机静态设置它们的域 ID 和组构名称。所述方法包括：访问所述组构，选择组构中的一个交换机，以及为所选择的交换机静态配置域 ID 和组构名称。为静态组构中的每个交换机重复上述步骤。在一个实施例中，在静态配置之后，该交换机与组构中的任何动态设置的交换机相隔离。该交换机检测它有哪些端口连接到动态设置的交换机，之后隔离它们，同时保持连接到静态配置的交换机的端口的可操作性。



1. 一种定义存储区域网组构的方法，包括：
通过为所述组构中的每个所选择的交换机静态配置域 ID 和组构名称，静态定义所述组构。
2. 如权利要求 1 所述的方法，其中为所述组构中的每个交换机配置域 ID 和组构名称的步骤还包括：禁用每个交换机的域管理器。
3. 如权利要求 2 所述的方法，其中为所述组构中的每个交换机配置域 ID 和组构名称的步骤还包括：将所述域 ID 和组构名称分别写入每个交换机的存储器。
4. 如权利要求 3 所述的方法，还包括在将所述域 ID 和组构名称分别写入每个交换机的存储器之后，向耦合到静态配置的交换机的任何端设备分配它们的光纤通道地址。
5. 如权利要求 3 所述的方法，其中，在写入每个交换机的所述域 ID 和组构名称之后，将所述交换机与所述组构中任何动态配置的交换机在逻辑上相隔离。
6. 如权利要求 5 所述的方法，其中所述交换机的逻辑隔离还包括：
在所述交换机的端口处接收来自任何所述动态配置的交换机的第一消息，所述第一消息指示主要交换机的选择；
向发送所述第一消息的任何所述动态配置的交换机发送拒绝消息；以及
隔离所述交换机和发送所述第一消息的任何所述动态配置的交换机之间的端口。
7. 如权利要求 6 所述的方法，其中所述消息包括下述类型的消息之一：交换组构参数、构建组构、重新配置组构、域标识符已分配或请求域标识符。
8. 如权利要求 1 所述的方法，还包括通过管理应用或命令行解释器访问所述组构，来静态配置所述交换机。
9. 如权利要求 1 所述的方法，其中定义所述静态组构的步骤还包括：

- a. 访问所述组构;
- b. 选择所述组构中将要进行静态配置的第一交换机;
- c. 禁用与所述第一交换机相关联的域管理器;
- d. 配置所述第一交换机的所述域 ID 和组构名称;
- e. 为所述组构中的每个所述静态配置的交换机重复 b 至 d 步骤。

10. 如权利要求 1 所述的方法, 还包括通过使能所述组构中每个所选择的交换机的域管理器, 动态定义所述静态配置的组构。

11. 一种存储区域网, 包括:

一个或多个宿主机;

一个或多个存储设备; 以及

连接所述一个或多个宿主机和存储设备的静态组构, 所述静态组构具有一个或多个交换机, 所述一个或多个交换机静态配置它们的域 ID 和组构名称。

12. 如权利要求 11 所述的存储区域网, 其中静态配置其域 ID 和组构名称的所述一个或多个交换机在所述静态组构中是对等的。

13. 如权利要求 11 所述的存储区域网, 其中所述一个或多个交换机还包括隔离机构, 只要静态配置的交换机链接到动态配置的交换机时, 就调用所述隔离机构, 所述隔离机构在逻辑上隔离所述静态配置的交换机和所述动态配置的交换机。

14. 如权利要求 11 所述的存储区域网, 还包括耦合到所述组构用于访问所述组构的管理应用或命令行解释器。

15. 一种配置存储区域组构的方法, 包括:

- a. 访问所述组构;
- b. 选择所述组构中的交换机;
- c. 禁用与所选择的交换机相关联的域管理器;
- d. 静态配置所述所选择的交换机的域 ID 和组构名称; 以及
- e. 为所述组构中将要进行静态配置的每个交换机重复 b 至 d 步骤。

16. 如权利要求 15 所述的方法, 还包括在将所述域 ID 和组构名称分别写入每个所选择的交换机的存储器之后, 向耦合到所述静态配置的交换

机的任何端设备分配它们的光纤通道地址。

17. 如权利要求 15 所述的方法，其中，在写入每个所选择的交换机的所述域 ID 和组构名称之后，将所述交换机与所述组构中任何动态配置的交换机在逻辑上相隔离。

18. 如权利要求 17 所述的方法，其中每个所述所选择的交换机的逻辑隔离还包括：

在所述所选择的交换机处接收来自任何所述动态配置的交换机的第一消息，所述第一消息指示主要交换机的选择；

向发送所述第一消息的任何所述动态配置的交换机发送拒绝消息；以及

隔离所述所选择的交换机和发送所述第一消息的任何所述动态配置的交换机之间的端口。

19. 如权利要求 15 所述的方法，其中通过管理应用或命令行解释器访问所述组构。

20. 如权利要求 15 所述的方法，还包括在静态配置所述所选择的交换机的所述域 ID 和组构名称之前，禁用与所述所选择的交换机相关联的域管理器。

21. 一种方法，包括：

提供组构，所述组构具有多个交换机，每个所述交换机具有域 ID 和组构名称；

提供对所述组构的访问；以及

通过对所述组构的访问，使能所述组构的所述多个交换机的静态配置。

22. 如权利要求 21 所述的方法，还包括：

通过所述访问，在所述组构的所述多个交换机中，使能选择一个交换机；以及

在静态配置所选择的交换机的域 ID 和组构名称之前，禁用与所述所选择的交换机相关联的域管理器。

23. 如权利要求 22 所述的方法，其中，在写入所述所选择的交换机的

所述域 ID 和组构名称之后，将所述所选择的交换机与所述组构中任何动态配置的交换机在逻辑上相隔离。

24. 如权利要求 23 所述的方法，其中所述所选择的交换机的逻辑隔离还包括：

在所述所选择的交换机处接收来自任何所述动态配置的交换机的第一消息，所述第一消息指示主要交换机的选择；

向发送所述第一消息的任何所述动态配置的交换机发送拒绝消息；以及

隔离所述所选择的交换机和发送所述第一消息的任何所述动态配置的交换机之间的端口。

25. 如权利要求 21 所述的方法，其中提供对所述组构的访问的步骤还包括提供对所述组构之外的远程方的访问。

26. 如权利要求 22 所述的方法，还包括通过允许使能与所述所选择的交换机相关联的所述域管理器，动态重新配置所述组构。

用于定义静态光纤通道组构的装置与方法

技术领域

本发明一般地涉及存储区域网。更具体地说，本发明涉及用于定义不需要主要交换机的静态光纤通道组构（Fibre Channel Fabric）的方法与装置。

背景技术

随着因特网商务和网络中心计算的日益流行，企业和其它组织变得越来越依赖信息。为了操作全部数据，存储区域网或 SAN 已变得很流行。SAN 通常包括若干个存储设备、若干个宿主机和被布置在交换组构中的多个交换机。交换机有选择地连接 SAN 中的存储设备和宿主机。

大多数 SAN 依靠光纤通道协议在组构中通信。对于光纤通道协议和光纤通道交换组构的详细解释，参见 FC-FS（光纤通道成帧与发信号）和 FC-SW-2（光纤通道交换机组构-2），上述资料在这里为了全部目的作为参考而被引用。

在光纤通道中，由制造商指定的唯一八（8）字节宽节点名称标识每个设备（宿主机、存储设备和交换机）。当光纤通道设备进行互连以形成 SAN 时，节点名称（和其它参数一起）被用来标识每个设备。光纤通道帧被用于 SAN 中的设备之间的通信。但是，帧并不使用节点名称。而是通过三（3）字节光纤通道地址（或 FC_ID）来寻址每个端设备（宿主机和存储设备）的光纤通道端口，其中组构将三字节光纤通道地址（或 FC_ID）动态分配到端设备。每个端设备通过向交换组构执行组构登录过程来获取它的 FC_ID。在该过程中，端设备和组构交换它们的凭证和成功通过 SAN 进行通信所需的操作参数。最初，组构用唯一的组构名称标识它自身并用唯一的节点名称标识端设备。进而，组构将 FC_ID 分配给端设备的端口。

三字节宽光纤通道地址在结构上被分级为三个字段，每个字段一字节长：域 ID、区 ID 和端口 ID。组构中的每个交换机被分配一个域 ID。连接到特定交换机的端设备被分配该交换机的域 ID。交换机管理为每个端设备进行的区 ID 和端口 ID 字段分配，以保证在该域中所分配的地址的唯一性。例如，如果交换机被分为第五域，并且交换机将它的地址空间细分为两个区，每个区具有三个连接的端设备，那么可能的光纤通道地址分配为：5：1：1、5：1：2、5：1：3、5：2：1、5：2：2 和 5：2：3。

当交换组构初始化时，交换机之一被选为主要交换机。主要交换机向组构中的所有交换机分配域 ID，并且它的节点名称变成交换组构的组构名称。为了选择主要交换机，所有交换机相互交换被称作交换组构参数（EFP）的消息。EFP 除了含有其它参数外，还含有发送交换机的节点名称。具有最低节点名称的交换机被指定为主要交换机。所有其它交换机被称作非主要交换机。一旦选择了主要交换机，它向它的邻近交换机发送域标识符已分配（DIA）消息，其通知邻近交换机它已由主要交换机分配域 ID。作为应答，邻近交换机向主要交换机发送请求域标识符（RDI）消息。主要交换机分配域 ID 并以向每个交换机发送其域 ID 作出响应。此后，接收域 ID 的交换机向它的邻近交换机发送 DIA，接收作为应答的 RDI，并且向主要交换机转发 RDI，主要交换机向进行请求的交换机分配域 ID。该过程一直持续到所有交换机接收到域 ID。在已接收域 ID 之后，各个交换机为其域中的每个端设备分配区 ID 和端口 ID。当所有交换机已被分配域 ID 后，认为组构配置完成。结果端设备都被分配了它们的区 ID 和端口 ID。

光纤通道允许两个独立的交换组构合并成一个。这在每个属于不同组构的两个交换机之间建立了连接的时候发生。当发生这样的情况时，合并的组构的某些交换机的域 ID 和它们的端设备的 FC_ID 可能需要重新分配。例如，如果包括域 ID 一、二和三（1、2 和 3）的组构 A 将和包括域 ID 一和二（1 和 2）的第二组构 B 合并，则组构之一的重叠域 ID（1 和 2）必须进行重新分配。当两个组构连接时，通过连接两个组构的链路交换 EFP 消息，以确定在域 ID 之间是否有任何重叠。根据结果，会发生两

种情形之一。

如果在多个交换机之间存在域 ID 分配的重叠，则隔离连接两个原始组构的链路。链路在逻辑上被断开连接，并且不被设备识别，尽管物理链路仍旧保持。SAN 管理员随后可以请求联合组构的破坏性重新配置以解决域 ID 冲突。在此情况下，两个原始组构的所有交换机上大量充斥重新配置组构（RCF）消息。这将终止所有数据流量的输送，并且每个交换机撤销其域 ID。此后，选择主要交换机，将新的域 ID 分配给交换机，并且以和前述相同的方式将新的 FC_ID 分配给端设备。如此，两个组构被合并为一个。

如果在交换机之间没有域 ID 重叠，则自动执行非破坏性组构重新配置。在两个原始组构的所有交换机上大量充斥构建组构（BF）消息。数据帧输送没有停止，并且每个交换机保持它的域 ID。因为两个组构分别具有一个主要交换机，所以两者之一必须“辞去”其主要状态，只留下一个主要交换机用于合并的组构。结果，发生上面描述的主要交换机选择过程。每个非主要交换机随后向依然存在的主要交换机作出 RDI 请求，以请求与其在 BF 消息前具有的域 ID 相同的域 ID。这样，在没有改变任何交换机的域 ID 分配或被分配到端设备的任何 FC_ID 的情况下，合并了两个组构。

然而，对于两个原始组构之一，主要交换机改变，从而那些交换机的组构名称需要更新。假定组构名称是每个端设备保持的组构登录状态信息的一部分，则不稳定组构的交换机必须重新初始化它们的端设备以更新它们的状态。该过程引起失去主要交换机的组构中的数据流量的中断。

许多问题与在现有光纤通道标准下在组构的交换机之间分配域 ID 的方法相关联。首先，主要交换机是动态选择的。换句话说，主要交换机是在最初配置组构的时候或实现组构的改变的任何时候“在不工作状态下”选择的。此外，RCF 和 BF 过程对组构是破坏性的。当调用 RCF 过程时，在标识主要交换机和重新分配域 ID 的同时，合并的组构上的所有流量停止。BF 过程也是部分破坏性的。在原始主要交换机继续是合并组构的主要交换机的组构中，流量继续。然而对于另一个组构，必须更新交换机以

反映它们被包括在新的组构中。因此流量被中断，直到更新完成。此外，具有主要交换机的组构易于发生灾难性问题。例如，如果主要交换机发生故障，可能导致整个组构无法工作。并且，如果缆线被误插入错误的连接器，则会不小心将两个组构连接在一起，这可能引起整个组构重新配置。

因此，需要一种用于定义不需要主要交换机的静态光纤通道组构的方法与装置。

发明内容

为了实现前述方面，根据本发明的目的，公开了一种存储区域网和用于定义不需要主要交换机的静态光纤通道组构的方法。存储区域网包括一个或多个宿主机、一个或多个存储设备和连接所述一个或多个宿主机和存储设备的静态组构。在所述静态组构中，交换机手动设置它们的域 ID 和组构名称。所述方法包括访问所述组构，选择所述组构中的一个交换机，为所选择的交换机手动设置域 ID 和组构名称。为静态组构中的每个交换机重复上述步骤。在一个实施例中，在静态配置之后，该交换机与组构中的任何动态设置的交换机相隔离。该交换机检测它有哪些端口连接到动态设置的交换机，之后隔离它们，同时保持连接到静态配置的交换机的端口的可操作性。

附图说明

图 1 是存储区域网的交换组构；

图 2 是图示根据本发明定义静态光纤通道组构的步骤的流程图；

图 3 图示了根据本发明的与动态设置的光纤通道组构隔离的示例性静态光纤通道组构。

具体实施方式

在下面的描述中阐述了很多具体细节，以便于充分理解本发明。但是，没有全部或一些这些具体细节也可以实施本发明，对于本领域的技术人员来说是很明显的。在另外一些例子里，没有对公知的处理步骤进行详

细的描述，以避免不必要地混淆本发明。

参考图 1，示出了存储区域网的交换组构。存储区域网（SAN）10 包括交换组构 12，交换组构 12 包括多个光纤通道交换机 SW1 至 SW6。同时在 SAN 10 中包括多个宿主机 H1 至 H6 和多个存储设备 D1 至 D5。管理组构的工具例如命令行解释器或图形管理应用 14 通过宿主机 H4 被连接到组构 12。命令行解释器或管理应用 14 使得网络管理员 16 能够通过宿主机 H4 和交换机 SW4 管理组构 12。

根据本发明的多个实施例，宿主机 H1—H6 可以是任何类型的宿主机，包括但不限于运行在 Unix、Windows 或任何其它计算平台之上的个人计算机或服务器、或者它们的组合。类似地，存储设备 D1—D5 可以是任何类型的存储设备，包括但不限于磁带备份系统、仿真磁带备份系统、CD-ROM 存储阵列或诸如独立磁盘冗余阵列（RAID）的一个或多个盘、或者它们的组合。交换机 SW1—SW6 可以是任何类型的光纤通道交换机，例如那些可从本发明的受让人、Andiamo System、加州圣何塞市的 Brocade 商业获得的交换机。应当注意，图中所示的组构 12 只是对描述本发明有帮助的 SAN 的图示。不能将其的简单形式理解为对本发明的限制，本发明可以用于任何 SAN 配置中。

本发明提供了一种用于手动配置组构 12 的交换机 SW 的方法。该技术可以用在最初配置组构 12 的时候，或者用在在组构 12 工作以后网络管理员想要实现组构 12 上的改变的时候。

参考图 2，示出了流程图 20，流程图 20 图示了用于定义根据本发明的静态光纤通道组构的过程。当网络管理员 16 想要静态定义组构 12（块 22）时，管理员首先访问组构 12。在多个实施例中，该过程通过命令行解释器或管理应用 14 实现（块 24）。

一旦建立了对组构 12 的访问，管理员就选择将要进行静态配置的交换机 SW（块 26）。管理员随后禁用交换机 SW 的域管理器（交换机的组件，用于实现主要交换机选择和域 ID 分配状态机）（块 28）。禁用域管理器需要交换机的工作模式有明显改变。因此当禁用域管理器时，交换机关断它的所有端口，包括连接到其它交换机的那些端口和连接到端设备的

那些端口。一旦禁用了域管理器，管理员就在交换机的存储器中静态配置（即，手动写入）合适的域 ID 和组构名称。在静态配置了域 ID 和组构名称之后，交换机随后接通它的所有端口（块 30）。通过使用标准组构登录过程，使用静态配置的信息向连接到端口的端设备分配它们的 FC_ID。在另一个实施例中，管理员可以在禁用域管理器之前配置静态域 ID 和组构名称。在这种情形下，直到禁用了域管理器之后，交换机才使用静态配置的域 ID 和组构名称。

在进行了静态配置之后，该交换机与组构 12 中的动态设置的交换机相隔离。交换机检测它的哪些端口连接到动态设置的交换机，然后隔离它们，同时保持连接到静态配置的交换机的端口的可操作性。检测算法基于如 FC-SW-2 所定义的交换机端口行为。接收用于主要交换机选择的消息的任何端口指示发送该消息的交换机是动态配置的，其中所述消息例如是 EFP、BF、RCF、DIA、RDI。作出响应的静态配置的交换机向动态交换机发送拒绝（SW_RJT）消息和原因代码解释“E 端口被隔离”，并且随后隔离该端口。在接收到此消息后，动态配置的交换机需要隔离其连接到该静态配置的交换机的端口。如此，手动配置的交换机与组构 12 中进行动态配置的任何其它交换机在逻辑上相隔离。链接到任何其它静态配置的交换机的端口没有接收到主要交换机选择消息（EFP、BF、RCF、DIA、RDI），因此没有被隔离（块 32）。

在判定框 34 处，管理员决定是否要静态配置另一个交换机。如果没有，则过程结束。如果有，则管理员选择另一个交换机（块 26），并且重复上述步骤。为每个静态配置的交换机定义相同的组构名称和非冲突的域 ID 是管理员的责任。

分清管理员执行的动作和交换机执行的动作是有帮助的。管理员需要为每个静态配置的交换机禁用域管理器并写入静态配置信息。这些静态配置的交换机自动执行如上所述的端口动作，以在逻辑上将静态配置的交换机与动态设置的交换机相隔离。在管理员方不需要任何其它动作来静态定义组构 12。

因此，本发明提供了一种为组构 12 中的每个交换机静态配置域 ID 和

组名称的方法。此外，本发明为交换机提供了一种确定静态配置的交换机是连接到动态配置的交换机还是连接到另一个静态配置的交换机的方法。并且，这些静态设置的交换机被认为是对等的。换句话说，本发明可以用于定义不需要主要交换机的静态交换组。

参考图 3，示出了具有已进行静态配置的几个交换机 SW 的示例性组 40。在本例中，三个交换机 SW4、SW5 和 SW6 已经如上所述进行了静态配置。但是，余下的交换机 SW1、SW2 和 SW3 已经进行了动态配置。如上所述，只要交换机 SW 进行了静态配置，它就在逻辑上与任何动态配置的交换机 SW 相隔离。虚线 42 和 44 分别表示交换机 SW1 和 SW6 之间以及交换机 SW3 和 SW4 之间被隔离的链路。应当注意，隔离是逻辑上的而不必须是物理上的。换句话说，除了链路初始化消息（交换链路参数，ELP）外，在被隔离的链路上既没有数据流量也没有配置消息流动，其中链路初始化消息可以用于在邻近交换机变成静态配置的交换机的情况下退出隔离状态。因此，在本例中，组 40 已被有效地划分为彼此不进行通信的两个组。组 A 包括动态设置的交换机 SW1—SW3。组 B 包括静态设置的交换机 SW4—SW6。

或者，本发明还可以使管理员能够动态配置静态配置的组 12。该过程与前述过程基本相同，只是处理相反。首先需要管理员选择一个交换机并且使能该交换机的域管理器。根据一个实施例，当上述情形发生时，交换机关断它的所有端口，随后根据 FC-SW-2 标准再次接通它们。从而，端口根据 FC-SW-2 标准工作，并且接收何识别用于选择主要交换机的消息，例如 EFP、BF、RCF、DIA 和 RDI。因此，该交换机变成动态配置的，并且能够和组中的其它动态配置的交换机通信。为组中将要进行动态配置的每个交换机重复前述过程。

上面描述的本发明的实施例应当被认为是示例性的而非限制性的。例如，本发明不必只用于 SAN。而是可以由具有动态设置的光纤通道交换机的任何类型网络使用。本发明并不限于这里给出的细节，而是可以在所附权利要求的范围和等同方案内进行修改。

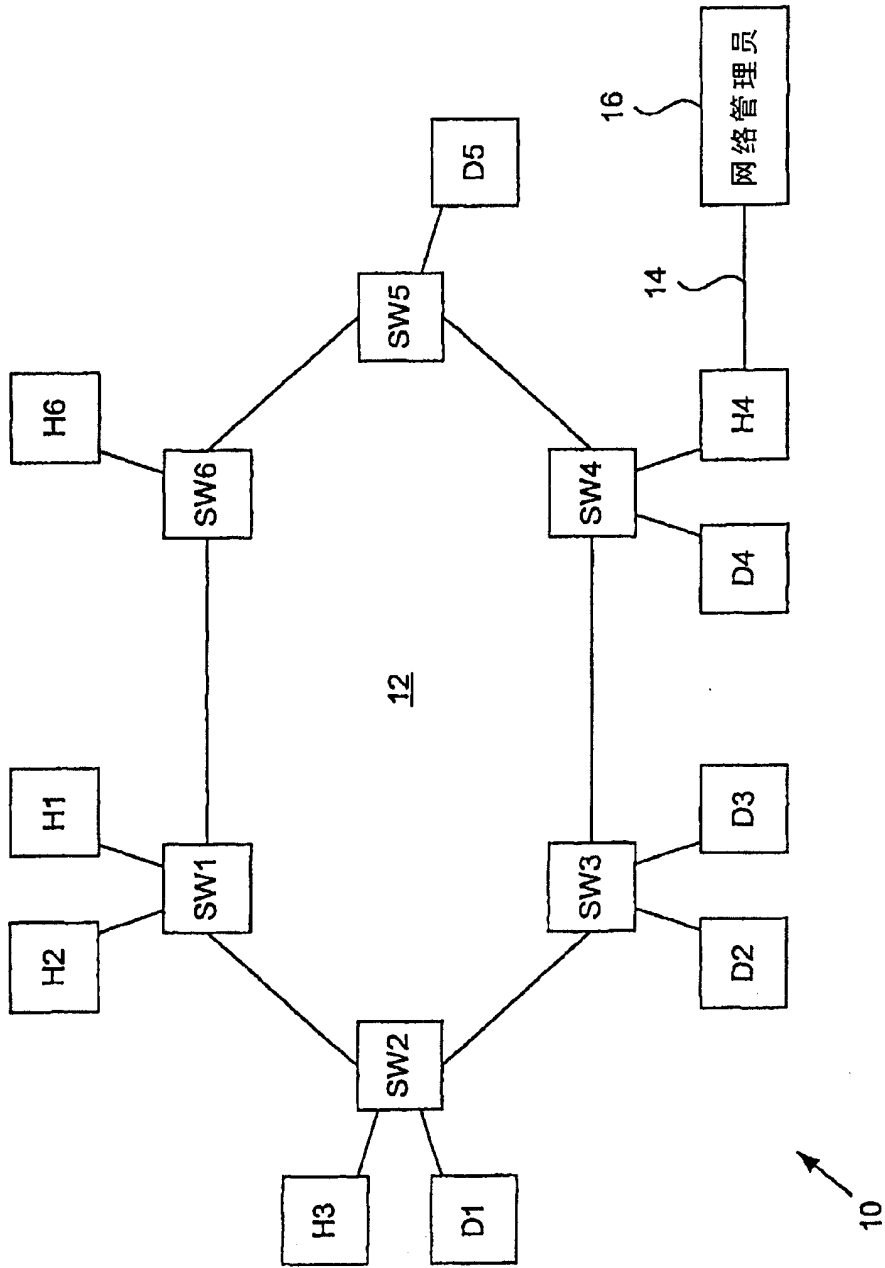


图1

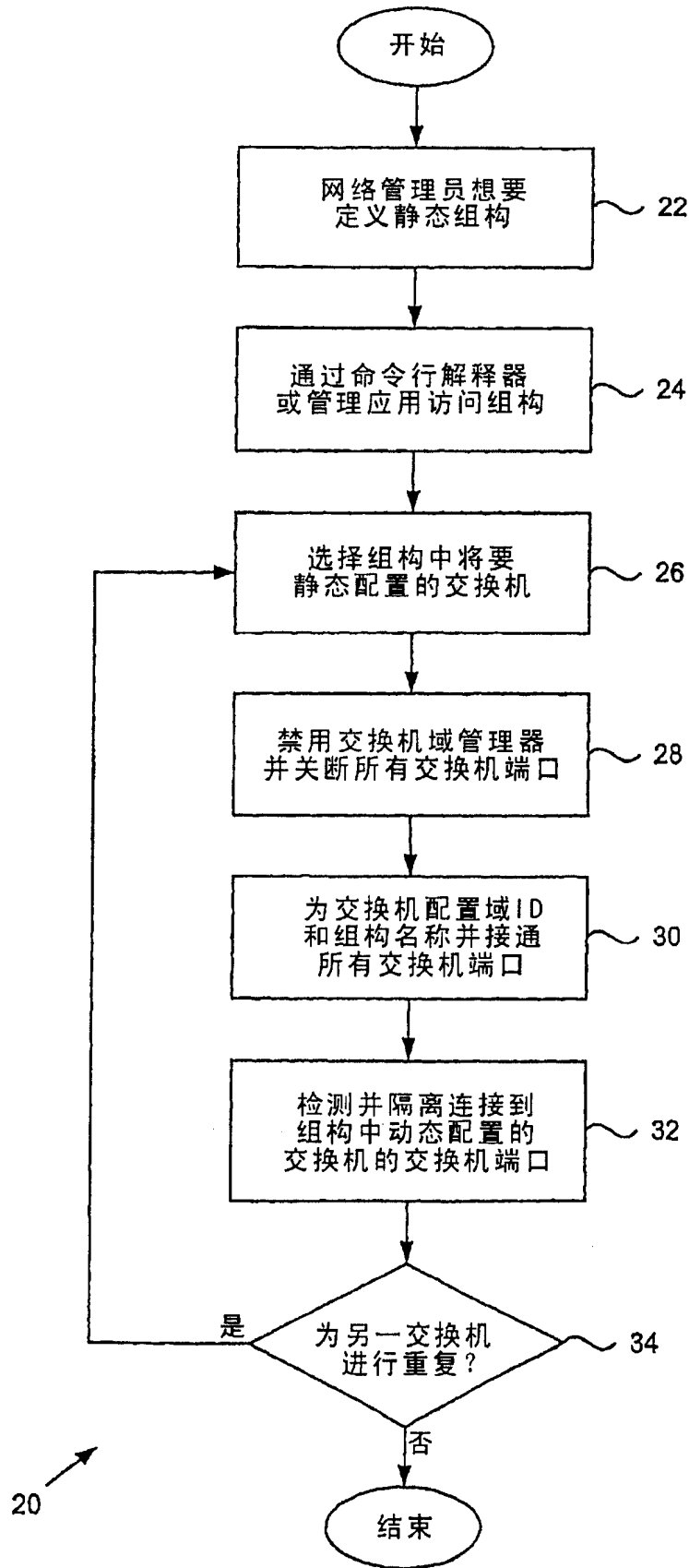


图2

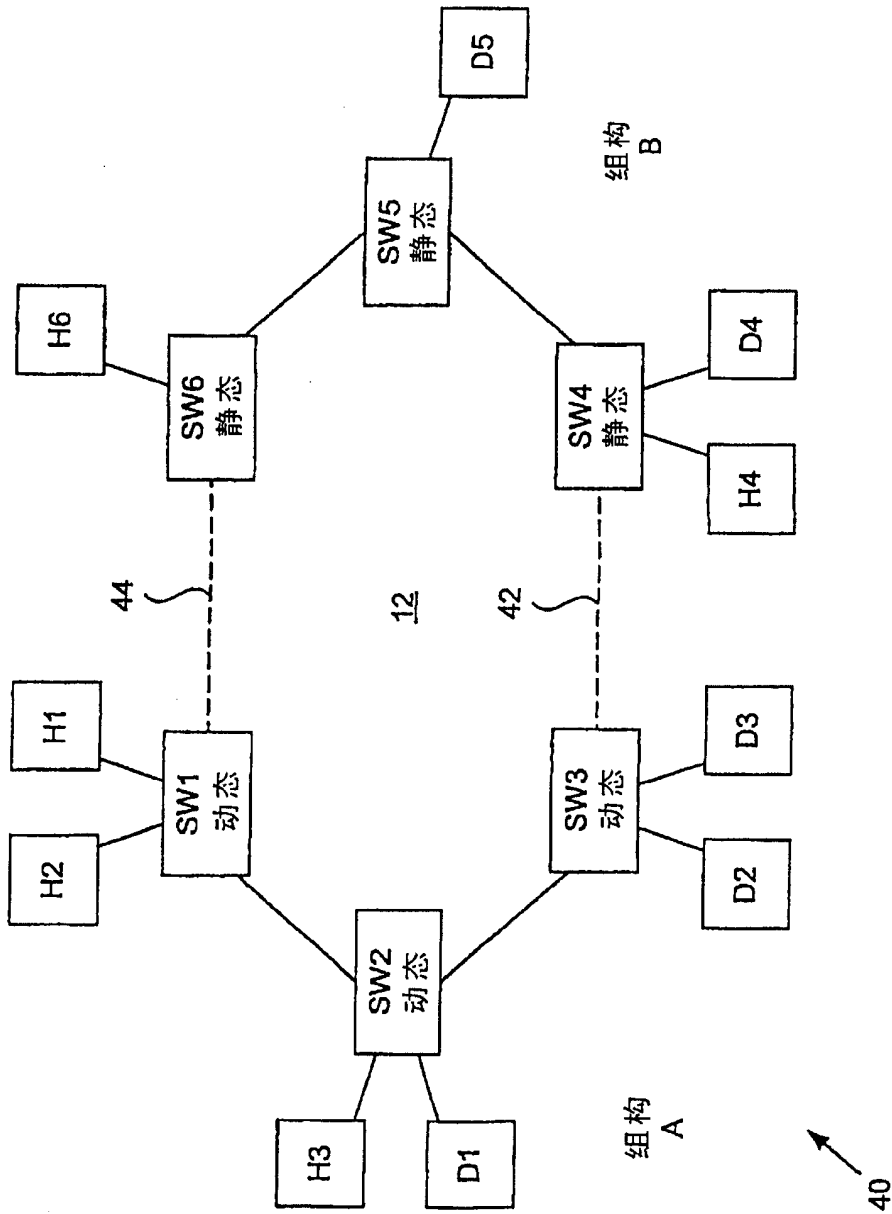


图3