（54）发明名称
通信系统和收集端口信息的方法

（57）摘要
公开了包括多个分组传送装置和控制每个分组传送装置的控制装置的通信系统。该通信系统能够将各分组传送装置在分组传送时采用的通信网络和控制装置用于控制每个分组传送装置的通信网络组织成单种通信网络。控制装置（80）识别目的装置上的尚未查找的端口，并且向其上存在尚未探索的端口的目的地装置发送响应请求。分组传送装置（90）判定是否要将其上接收到的响应请求传送到另外的分组传送装置。在判定要将响应请求传送到另外的分组传送装置时，分组传送装置（90）将响应请求传送到另外的分组传送装置。在判定不将响应请求传送到另外的分组传送装置时，分组传送装置（90）沿着去到控制装置的路径返回包括该分组传送装置的端口信息的响应。
1. 一种通信系统，包括：
多个分组传送单元；以及
控制每个分组传送单元的控制单元，
其中：控制单元包括

拓扑信息存储装置，该拓扑信息存储装置存储包括关于控制单元或每个分组传送单元的端口的互连信息的拓扑信息，

未检测连接目的地端口确定装置，该未检测连接目的地端口确定装置根据拓扑信息在控制单元或每个分组传送单元的所有端口中仅确定未检测连接目的地端口，未检测连接目的地端口是连接目的地未被检测的端口，

答复请求发送装置，该答复请求发送装置经由未检测连接目的地端口向未检测连接目的地端口的连接目的地发送答复请求，

答复接收装置，该答复接收装置从成为未检测连接目的地端口的连接目的地的分组传送单元接收包括关于该分组传送单元的端口的信息的答复，以及

拓扑信息添加装置，该拓扑信息添加装置向拓扑信息添加与成为未检测连接目的地端口的连接目的地的分组传送单元的端口有关，并被包括在从该分组传送单元接收的答复中的信息，并且如果该分组传送单元的端口之一是连接目的地未被检测的端口并且不是该分组传送单元用来接收回答复请求的端口，则拓扑信息添加装置还向拓扑信息添加指示出该端口未被检测的信息；并且

分组传送单元包括

传送判定装置，该传送判定装置基于接收到的答复请求的内容来判定该答复请求是否要被传送到另外的分组传送单元，

传送装置，在判定答复请求要被传送到任何其他分组传送单元时，该传送装置基于答复请求的内容将答复请求传送到任何其他分组传送单元，以及

答复发送装置，在判定答复请求不被传送到任何其他分组传送单元时，该答复发送装置返回包括关于该分组传送单元的端口的信息的答复，该答复被沿着通向控制单元的路径发送。

2. 根据权利要求1所述的通信系统，

其中：如果未检测连接目的地端口确定装置判定不存在未检测连接目的地端口，则答复请求发送装置停止答复请求的发送。

3. 根据权利要求1所述的通信系统，

其中：答复请求发送装置在答复请求中记述一列表作为答复请求传送路径，该列表是关于从控制单元的端口延伸到未检测连接目的地端口的路径上的每个分组传送单元要用来输出答复请求的端口的信息的顺序阵列；

传送装置沿着答复请求中记述的传送路径传送答复请求，并且

答复发送装置基于答复请求中记述的传送路径来辨识去到控制单元的答复发送路径。

4. 根据权利要求1所述的通信系统，

其中：控制单元包括：控制信道确定装置，该控制信道确定装置使用拓扑信息来设定作为控制单元和每个分组传送单元之间的路径的控制信道；以及控制信道通知发送装置，该控制信道通知发送装置发送控制信道通知以将控制信道通知给每个分组传送单元；并且
分组传送单元包括控制信道存储装置，该分组传送单元在接收到控制信道通知时，使用控制信道通知来存储互连控制单元和该分组传送单元的控制信道。

5. 一种用于控制多个分组传送单元的控制单元，该控制单元包括：

拓扑信息存储装置，该拓扑信息存储装置存储包括关于控制单元或每个分组传送单元的端口的互连信息的拓扑信息，

未检测连接目的端口确定装置，该未检测连接目的端口确定装置根据拓扑信息在控制单元或每个分组传送单元的所有端口之中仅确定未检测连接目的端口，未检测连接目的端口是连接目的地未被检测的端口，

答复请求发送装置，该答复请求发送装置经由未检测连接目的端口向未检测连接目的端口的连接目的地发送答复请求，

答复接收装置，该答复接收装置从成为未检测连接目的端口的连接目的地的分组传送单元接收包括关于该分组传送单元的端口的信息的答复，以及

拓扑信息添加装置，该拓扑信息添加装置向拓扑信息添加与成为未检测连接目的端口的连接目的地的分组传送单元的端口有关并被包括在从该分组传送单元接收的答复中的信息，并且如果该分组传送单元的端口之一是连接目的地未被检测的端口并且不是该分组传送单元用来接收到答复请求的端口，则拓扑信息添加装置还向拓扑信息添加指示出该端口未被检测的信息。

6. 一种受控制单元控制的分组传送单元，该分组传送单元包括：

传送判定装置，该传送判定装置在从控制单元接收到被寻址到连接目的地未被检测的未检测连接目的端口的连接目的地的答复请求时，基于答复请求的内容来判定该答复请求是否要被传送到另外的分组传送单元；

传送装置，在判定答复请求要被传送到任何其他分组传送单元时，该传送装置基于答复请求的内容将答复请求传送到任何其他分组传送单元，以及

答复发送装置，在判定答复请求不被传送到任何其他分组传送单元时，该答复发送装置返回包括关于该分组传送单元的端口的信息的答复，该答复被沿着通向控制单元的路径发送。

7. 一种端口信息收集方法，其中控制多个分组传送单元的控制单元执行：

根据拓扑信息在控制单元或每个分组传送单元的所有端口之中仅确定未检测连接目的端口，未检测连接目的端口是连接目的地未被检测的端口；

经由未检测连接目的端口向未检测连接目的端口的连接目的地发送答复请求；

从成为未检测连接目的端口的连接目的地的分组传送单元接收包括关于该分组传送单元的端口的信息的答复；以及

向拓扑信息添加与成为未检测连接目的端口的连接目的地的分组传送单元的端口有关并被包括在从该分组传送单元接收的答复中的信息，并且如果该分组传送单元的端口之一是连接目的地未被检测的端口并且不是该分组传送单元用来接收到答复请求的端口，则还向拓扑信息添加指示出该端口未被检测的信息。
说明 书

通信系统和收集端口信息的方法

技术领域
[0001] 本发明涉及包括多个分组传送装置和用于控制分组传送装置的控制单元的通信系统。本发明还涉及应用于该通信系统的分组传送单元、控制单元、端口信息收集方法和控制单元用程序。

背景技术
[0002] 有一种包括各自传送分组的多个交换机和控制每个交换机的控制单元的通信系统，在该通信系统中，控制单元设定每个交换机中的流条目信息并且交换机根据流条目信息来传送接收到的分组。流条目信息依分组的头部规定接收到的分组要经历的处理的类型（例如传送、丢弃、更新或其他）。这种通信系统需要控制信道，控制单元使用该控制信道来控制交换机。。例如，如果交换机接收到在流条目信息中未限定的分组，则交换机将此事件通知给控制单元，然后控制单元根据该具体分组向交换机发送适当的流条目信息。控制信道被用于诸如从交换机到控制单元的指令和流条目信息从控制单元到交换机的发送之类的用途。控制单元用于控制交换机的协议被称为 OpenFlow 开放体系。交换机可被认为是传送分组的分组传送单元。控制信道相当于 OpenFlow 协议中的“安全信道”。另外，控制单元在 OpenFlow 中被称为控制器。

[0003] OpenFlow 的规范在非专利文献 1 中记载。在 OpenFlow 的规范中规定控制单元和交换机应当使用 SSL（安全套接字层）来以端口号 6633 的 TCP（传输控制协议）通信。从而，应用了 OpenFlow 协议的通信系统通常包括两个独立的通信网络。一个是用于按顺序传送分组多个交换机的通信网络，另一个是充当控制单元与每个交换机之间的控制信道的通信网络。图 27 是示出应用了 OpenFlow 协议的一般通信系统的示例的说明图。图 27 中所示的每个交换机 92 使用分组传送通信网络 94 来根据流条目信息向其他交换机传送接收到的分组。此外，独立于分组传送通信网络 94 提供了用作控制信道的控制通信网络 93，并且控制单元 91 经由控制通信网络 93 控制每个交换机 92。控制通信网络 93 使用诸如 STP（生成树协议）、TCP/IP（传输控制协议 / 因特网协议）、RIP（路由信息协议）或 OSPF（开放最短路径优先）之类的路由协议。

[0004] 另外，在图 27 中作为示例示出的通信系统中，当交换机的拓扑被改变时，每个交换机在通信系统中交换控制帧并且认识到新拓扑。用于认识到拓扑的控制帧的示例例如包括 STP 中的 BPDU（网桥协议数据单元）以及其他。

[0005] 引文列表
[0006] 非专利文献
说明 书

[0008] 技术问题

[0009] 图27中的基 于 OpenFlow 的一般通信系统使用控制通信网络93和分组传送通信网络94。然而，在包括多个分组传送装置和控制每个分组传送单元的控制单元的通信系统中，最好每个分组传送单元用于传送分组的通信网络和控制单元用于控制每个分组传送单元的通信网络应被集成为一种。

[0010] 从而，本发明希望提供：包括多个分组传送装置和控制每个分组传送单元的控制单元的通信系统，该通信系统被配置成使得每个分组传送单元用于传送分组的通信网络和控制单元用于控制每个分组传送单元的通信网络可被集成成为一种；以及应用于该通信系统的分组传送单元、控制单元、端口信息收集方法和控制单元用程序。

[0011] 用于解决问题的手段

[0012] 根据本发明的一种通信系统包括多个分组传送单元和控制每个分组传送单元的控制单元，其中：控制单元包括：拓扑信息存储装置，该拓扑信息存储装置存储包括关于控制单元和每个分组传送单元的端口的互连信息的拓扑信息；未检测连接地端口确定装置，该未检测连接地端口确定装置根据拓扑信息在控制单元和每个分组传送单元的所有端口之间仅限定未检测连接地端口，未检测连接地端口是连接地端口检测的端口；

[0013] 复制请求发送装置，该复制请求发送装置包括由未检测连接地端口的端口前向未检测连接地端口的连接地端口发送的复制请求；以及复枚接收到装置，该复枚接收到装置从未检测连接地端口的连接地端口的端口前向未检测连接地端口的端口发送的复制请求包括关于该分组传送单元的端口的信息的复制，而且分组传送单元包括：传送判定装置，该传送判定装置基于接收到的复枚请求的内容来判断该复制请求是否要被发送到另外的分组传送单元；传送装置，在判断复制请求被发送到任何其他分组传送单元时，该传送装置基于复制请求的内容将复制请求发送到任何其他分组传送单元；以及复枚发送装置，在判断复制请求不被发送到任何其他分组传送单元时，该复枚发送装置返回包括关于该分组传送单元的端口的信息的复制，该复枚被沿著通向控制单元的路径发送。

[0014] 根据本发明的分组传送单元包括控制单元，该分组传送单元包括：传送判定装置，该传送判定装置在接收到被寻址到连接地端口的未检测连接目的地端口的连接请求时，基于复制请求的内容来判断该复制请求是否要被发送到另外的分组传送单元；传送装置，在判断复制请求被发送到任何其他分组传送单元时，该传送装置基于复制请求的内容将复制请求发送到任何其他分组传送单元；以及复枚发送装置，在判断复制请求不被发送到任何其他分组传送单元时，该复枚发送装置返回包括关于该分组传送单元的端口的信息的复制，该复枚被沿着通向控制单元的路径发送。
[0015] 根据本发明的一种端口信息收集方法，其中控制多个分组传送单元的控制单元执行，根据拓扑信息在控制单元或每个分组传送单元的所有端口之中已确定未检测连接目的
地端口，未检测连接目的地端口是连接目的地未被检测的端口；经由未检测连接目的地端口
前未检测连接目的地端口的连接目的地发送答复请求，以及从成为未检测连接目的地端口
的连接目的地的分组传送单元接收包括关于该分组传送单元的端口的信息的答复。
[0016] 根据本发明的一种控制单元用程序欲安装在控制多个分组传送单元的计算机上，
该程序使得该计算机执行，根据包括关于该计算机或每个分组传送单元的端口的连通信
息的拓扑信息在该计算机中每个分组传送单元的所有端口之中已确定未检测连接目的地端
口，未检测连接目的地端口是连接目的地未被检测的端口；经由未检测连接目的地端口向
未检测连接目的地端口的连接目的地发送答复请求，以及从成为未检测连接目的地端口的
连接目的地的分组传送单元接收包括关于该分组传送单元的端口的信息的答复。
[0017] 本发明的有利效果
[0018] 根据本发明，在包括多个分组传送装置和控制每个分组传送单元的控制单元的通
信系统中，每个分组传送单元用于传送分组的通信网络和控制单元用于控制每个分组传
送单元的通信网络可被集成为一种。

附图说明
[0019] 图1描绘了示意性示出从控制单元未存储拓扑信息的状态直到控制单元开始基
于隧道传输的通信为止的处理步骤的说明图。
[0020] 图2描绘了示意性示出本发明的通信系统中配备的控制单元和交换机的配置的示例
的框图。
[0021] 图3描绘了示意性示出源路由的示例的说明图。
[0022] 图4描绘了示意性示出特定分组的格式的说明图。
[0023] 图5描绘了示意性示出请求分组中包括的信息的示例的说明图。
[0024] 图6描绘了示意性示出答复分组中包括的信息的示例的说明图。
[0025] 图7描绘了示意性示出分组中包括的信息的示例的说明图。
[0026] 图8描绘了示意性示出隧道分组中包括的信息的示例的说明图。
[0027] 图9描绘了示意性示出控制单元生成拓扑信息并形成控制信道的处理步骤的示例图。
[0028] 图10描绘了示意性示出控制单元生成拓扑信息并形成控制信道的处理步骤的另一示例
图。
[0029] 图11描绘了示意性示出控制单元生成拓扑信息并形成控制信道的处理步骤的另一示例
图。
[0030] 图12描绘了示意性示出控制单元生成拓扑信息并形成控制信道的处理步骤的另一示例
图。
[0031] 图13描绘了示意性示出控制单元生成拓扑信息并形成控制信道的处理步骤的另一示例
图。
[0032] 图14描绘了示意性示出控制单元和每个交换机交换请求分组和答复分组的处理步骤的
示例的说明图。
[0033] 图15描绘了示意性示出生成拓扑信息的处理步骤的示例的说明图。
具体实施方式

[0046] 以下，将参考附图来描述本发明的示例性实施例。

[0047] 本发明的通信系统包括各自发送分组的多个分组传送装置，以及控制多个交换机（分组传送装置）的控制单元。控制单元向每个分组传送单元发送流表指令，并且分组传送单元在接收到分组时根据流表信息来处理分组。在此序列完成后，连接到一个分组传送单元的终端单元可经由给定的路径向连接到另一分组传送单元的任何其他终端单元发送分组，或者经由各路径向连接到其他分组传送装置的其他终端装置广播分组。

[0048] 本发明中的分组传送单元例如是 OpenFlow 中的交换机，并且本发明中的控制单元例如是 OpenFlow 中的控制器。以下描述分组传送单元和控制单元分别是 OpenFlow 中的交换机和控制器的示例。然而，本发明也可应用到除 OpenFlow 以外的协议；本发明可应用到被构造成为使得控制单元承担存在与通信网络上的每个分组传送单元的集中管理的任何通信系统。

[0049] 本发明的通信系统不包括控制单元（控制器）与交换机之间的控制专用通信网络。控制单元只需要连接到至少一个交换机。控制单元利用交换机用来传送分组的通信网络来在每个交换机与控制单元自身之间提供控制信道。更具体而言，控制单元使用此通信网络的链路来形成控制单元自身与从控制单元看来连接到控制单元的每个交换机之间的控制信道。然后，控制单元经由控制信道通过隧道传输与每个交换机通信并且控制交换机。换言之，在本发明中，可以控制控制单元利用交换机用来传送分组的通信网络形成了每个交换机与控制单元自身之间的星形结构的覆盖网络。

[0050] 在本发明的通信系统中，控制单元还存储表示控制单元自身和每个交换机的端口之间的互连关系的拓扑信息。各交换机不需要存储拓扑信息。控制单元在其初始状态下不存储拓扑信息，向交换机查询其端口的连接状态，并且创建拓扑信息。在识别到了交换机的拓扑，控制单元向每个交换机确定且当该交换机与控制单元自身之间的控制信道的路径，并且发送分组来将通信控制单元的路径通知给该交换机。这使得控制单元能够识别到通信交换机的控制信道，并且交换机也同样能够识别到通信控制单元的控制信道。然后，交换机和控制单元通过隧道传输经由控制信道彼此通信。
图 1 是示意图，示出从控制单元未存储拓扑信息的状态直到控制单元和每个交换机开始基于隧道传输的通信为止的处理步骤的说明图。图 1 (a)示意图，示出了本发明的通信系统的初始状态。在此状态下，控制单元 10 虽然连接到交换机群组的至少一个交换机，但没有认识到控制单元连接到哪个交换机的哪个端口。

根据拓扑信息，控制单元 10 确定控制单元自身或交换机群组拥有的端口或在状态的端口，其中处于连接起来的状态的端口是知道该实体端口连接到哪个交换机的哪个端口的端口。以下将此端口称为未检测连接目的地端口。这里，用于检测连接的目的地端口输出用于指示处于连接起来的状态的端口向交换机进行查询的请求分组，控制单元 10 向交换机发送请求分组，如图 1 (b) 中所示。由于控制单元 10 在初始状态下未存储拓扑信息，所以控制单元 10 指定控制单元自身的连接目的地交换机未知的端口作为未检测连接目的地端口，并且从该未检测连接目的地端口发送请求分组。处于连接起来的状态的端口是连接到任何其他交换机的端口。

接收到请求分组的每个交换机 20 从该交换机 20 的用于接收到该请求分组的端口向控制单元 10 返回分组（以下其被称为答复分组）。此状态在图 1 (c) 中示出。答复分组包括指示处于连接起来的状态的端口的信息。指示出交换机的用于接收到该请求分组的端口的信息以及标识交换机 20 自身的信息。

在本发明中，控制单元 10 根据答复分组生成并存储拓扑信息。控制单元 10 还根据拓扑信息确定未检测连接目的地端口，并且发送请求分组，使得此请求将被从未检测连接目的地端口输出。如果未检测连接目的地端口是交换机 20 的端口，则控制单元 10 通过源路由每个交换机传送请求分组，使得具有所关注的未检测连接目的地端口的交换机 20 将未检测连接目的地端口输出请求分组。接收到请求分组的交换机 20 也将从该交换机 20 向该交换机 20 的用于接收到该请求分组的端口发送答复分组，并且控制单元 10 将把答复分组的内容包含到拓扑信息中。

这样，控制单元 10 重复确定未检测连接目的地端口、发送请求分组、接收答复分组和生成拓扑信息。当控制单元 10 得到不能根据拓扑信息确定未检测连接目的地端口时，控制单元判定拓扑信息已完成。

然后，控制单元 10 根据拓扑信息为每个交换机 20 确定充当控制信道的路径。从而，控制单元 10 认识到每个交换机的控制信道。控制单元 10 随后发送设立分组来将路径或者控制信道通知给每个交换机（参见图 1 (d)）。接收到设立分组的每个交换机 20 根据设立分组存储从交换机 20 自身延伸到控制单元 10 的路径或者控制信道。这使得交换机 20 能够认识到通过控制单元 10 的各控制信道。

随后，控制单元 10 用于控制交换机 20 的通信是经由专用于每个交换机 20 的控制信道通过隧道传输进行的，如图 1 (e) 中所示。例如，如果交换机 20 收到在流条目信息中未限定的分组，则交换机 20 将此通知给控制单元 10。此时，交换机 20 经由控制信道通过隧道传输进行该通知。另外，接收到该通知的控制单元 10 生成新的流条目信息并将该信息发送到交换机，此时控制单元 10 经由控制信道通过隧道传输发送流条目信息。经由控制信道通过隧道传输在控制单元 10 和交换机 20 之间交换的分组在下文中被称为隧道分组。
来，控制单元 10 再一次开始请求分组生成，并且在基于收集的答复分组生成新的拓扑信息之后，向每个交换机发送设立分组。控制单元 10 重复该序列。即使通过拓扑信息的周期性再生由修改了交换机间连接关系，也可将其广泛包含在拓扑信息中。

[0059] 请求分组、答复分组、设立分组和隧道分组被称作特定分组。交换机 20 对特定分组执行的处理不同于对在交换机之间传送的数据分组执行的处理。在接收到在终端（未示出）之间交换的数据分组时，交换机 20 根据流条目信息传送该数据分组，并且如果该数据分组的流在流条目信息中未限定，则将此事实通知给控制单元 10。交换机 20 对于特定分组不进行这种处理。取而代之，交换机 20 依据接收到的分组是请求分组、答复分组，设立分组还是隧道分组来进行不同的处理。相对于数据分组处理，每个交换机向特定分组处理赋予优先权。特定分组处理的优先保证了通信系统的 QoS（服务质量）。

[0060] 特定分组还包括与从该分组的发送源起的具体字符对应的适当字数值。此计数值被称为端口计数。

[0061] 图 2 是示出本发明的通信系统中配备的控制单元和交换机的配置的示例的框图。虽然图 2 示出了通信系统具有四个交换机 20a 至 20d 的示例，但此示例不限制交换机的数量及其拓扑。此外，虽然在图 2 的示例中两个交换机 20a 和 20b 连接到控制单元 10，但连接到控制单元 10 的交换机的数目可以是一个或三个或更多个；控制单元 10 只需至少有一个相连的交换机。

[0062] 控制单元 10 包括控制单元侧特定分组发送 / 接收部 11、拓扑搜索部 12、拓扑存储部 13、控制信道路径确定部 14 和控制单元侧控制消息发送/接收部 15。

[0063] 控制单元侧特定分组发送 / 接收部 11（以下简称为特定分组发送 / 接收部 11）发送和接收特定分组。更具体而言，特定分组发送 / 接收部 11 发送请求分组和设立分组，并且接收答复分组。特定分组发送 / 接收部 11 还发送和接收隧道分组。

[0064] 特定分组发送 / 接收部 11 在接收到来自交换机的答复分组时，将该答复分组输出到拓扑搜索部 12。特定分组发送 / 接收部 11 在接收到来自交换机的隧道分组时，将该隧道分组输出到控制单元侧控制消息发送/接收部 15。

[0065] 特定分组发送 / 接收部 11 发送的求分组、设立分组和隧道分组通过源路由被顺次传送到路径上的交换机，并且到达该分组所寻址到的交换机。生成请求分组的拓扑搜索部 12 在请求分组中顺次记述控制单元 10 自身的输出端口和直到目的变换机的路径上的每个交换机的输出端口，并且特定分组发送 / 接收部 11 从该输出端口的列表中第一个记述的控制单元 10 自身的端口发送请求分组。生成设立分组的控制信道路径确定部 14 在设立分组中顺次记述控制单元 10 自身的输出端口和直到目的变换机的控制信道路上的每个交换机的输出端口，并且特定分组发送 / 接收部 11 从该输出端口的列表中第一个记述的控制单元 10 自身的端口发送设立分组。类似地，生成要发送到交换机的隧道分组的控制单元侧控制消息发送/接收部 15 在隧道分组中顺次记述控制单元 10 自身的输出端口和直到目的变换机的控制信道路上的每个交换机的输出端口，并且特定分组发送 / 接收部 11 从该输出端口的列表中第一个记述的控制单元 10 自身的端口发送隧道分组。后文将描述源路由。

[0066] 以下，为了方便起见将描述在特定分组中记述端口。然而，更具体而言，在特定分组中记述了标识控制单元 10 和每个交换机的端口的信息，并且端口标识号即是该信息。输出端口列表因此是端口标识号的列表。在以下描述中以使用号码（端口标识号）作为端口
标识信息为例。

【0067】拓扑搜索部 12 确定未检测连接目的地端口，然后生成请求分组并使得特定分组发送/接收部 11 发送请求分组。当从交换机发送的答复分组作为对请求分组的答复被输入时，拓扑搜索部 12 还从答复分组生成拓扑信息并将该信息存储到拓扑存储部 13 中。拓扑搜索部 12 重复此序列，直到搜索部 12 变得不能够从拓扑信息确定未检测连接目的地端口为止。

【0068】拓扑搜索部 12 最初选择控制单元 10 自身的被链接起来的端口作为未检测连接目的地端口。拓扑搜索部 12 接下来在请求分组中记述被选择为未检测连接目的地端口的端口作为输出端口。拓扑搜索部 12 还记述输出端口列表的长度（即记述的输出端口的数目）和端口计数的初始值。在该情况下，输出端口的长度是 1，因为只有控制单元 10 自身的链接起来的端口是输出端口。如果存在多个端口作为未检测连接目的地端口，则拓扑搜索部 12 为每个未检测连接目的地端口生成请求分组。即，一个未检测连接目的地端口对应于一个请求分组。特定分组发送/接收部 11 将每个请求分组从该请求分组中记述的端口发送。

【0069】此外，在接收到答复分组并生成拓扑信息时，拓扑搜索部 12 参考拓扑信息并且确定如下端口作为未检测连接目的地端口：不知道该端口链接到拓扑信息中记述的任何链接起来的交换机端口中的哪一个。拓扑搜索部 12 还从生成的拓扑信息确定从控制单元 10 到具有所关注的未检测连接目的地端口的交换机的路径。拓扑搜索部 12 根据该具体路径在请求分组中记述控制单元 10 自身和每个交换机的适当输出端口。拓扑搜索部 12 还在请求分组中记述输出端口列表的长度（即，记述的输出端口的数目）和端口计数的初始值。在此情况下，拓扑搜索部 12 还为每个未检测连接目的地端口生成请求分组。特定分组发送/接收部 11 将每个请求分组从该请求分组中记述的第一端口发送。

【0070】拓扑搜索部 12 基于作为对每个请求分组的答复从每个交换机获得的答复分组来生成拓扑信息。每个答复分组除了包括指示出发送了该答复分组的交换机的链接起来的端口的信息以外，还包括标识该交换机的信息以及接收到引起答复分组的发送的请求分组的交换机端口信息。拓扑搜索部 12 识别出答复分组发送源交换机的端口（此端口是该交换机的用于接收请求分组的端口）连接到与请求分组相对应的未检测连接目的地端口，并且将这两个端口之间的互连关系添加到拓扑信息。

【0071】拓扑搜索部 12 的请求分组创建的开始启动了控制单元 10 的一系列处理，即请求分组发送、答复分组接收、拓扑信息创建和设立分组发送。拓扑搜索部 12 周期性地删除拓扑信息，然后开始请求分组的创建，并从而开始该一系列处理。

【0072】拓扑存储部 13 是存储拓扑信息的装置。

【0073】控制信道路径确定部 14 参考完成的拓扑信息并且逐交换机地确定独立控制信道。对于此控制信道确定方法没有限制。例如，控制信道路径确定部 14 可使用迪杰斯特拉 (Dijkstra) 算法来计算到交换机的最短路径并将此路径确定为控制信道。控制信道路径确定部 14 还为每个交换机生成设立分组。此时，控制信道路径确定部 14 在设立分组中顺次记述控制单元 10 自身的输出端口和到设立分组目的地交换机的控制信道上的每个交换机的输出端口。控制信道路径确定部 14 还在设立分组中记述输出端口列表的长度（即记述的输出端口的数目）和端口计数的初始值。然后，控制信道路径确定部 14 使得特定分组发送/接收部 11 发送设立分组。
控制单元侧控制消息隧道传输部 15（以下简称为控制消息隧道传输部 15）利用隧道传输发送和接收在控制单元 10 和交换机之间交换的控制消息。即，控制消息隧道传输部 15 将发送到交换机的控制消息（例如流条目信息）编码，并且生成经过封装的隧道分组。此时，控制消息隧道传输部 15 在隧道分组中顺次记述控制单元 10 自身的输出端口和到隧道分组目的地交换机的控制信道上的每个交换机的输出端口。控制消息隧道传输部 15 还在隧道分组中记述输出端口列表和端口计数的初始值。然后，控制消息隧道传输部 15 使得特定分组发送 / 接收部 11 发送隧道分组。

此外，如果特定分组发送 / 接收部 11 接收到来自交换机的隧道分组，然后将隧道分组输出到控制消息隧道传输部 15，则控制消息隧道传输部 15 解封出隧道分组中包括的控制消息并进一步对控制消息解码。

虽然下面在本示例性实施例中描述初始端口计数值为 0 的示例，但端口计数的初始值可以不是 0。

特定分组发送 / 接收部 11、拓扑搜索部 12、拓扑存储部 13、控制信道路径确定部 14 和控制消息隧道传输部 15 可各自独立实现。

特定分组发送 / 接收部 11、拓扑搜索部 12、控制信道路径确定部 14 和控制消息隧道传输部 15 同样可由根据用于控制单元的程序存储的计算机的 CPU 来实现。在此情况下，例如计算机的程序存储装置（未示出）可存储控制单元用程序，并且 CPU 可加载程序并根据该程序充当特定分组发送 / 接收部 11、拓扑搜索部 12、控制信道路径确定部 14 和控制消息隧道传输部 15。

接下来，下面描述交换机配置。交换机 20a 至 20d 基本上具有相同的配置。以下描述交换机 20a 作为示例。

交换机 20a 包括交换机侧特定分组发送 / 接收部 21，信息提供部 23，路径存储部 24 和交换机侧控制消息隧道传输部 25。

交换机侧特定分组发送 / 接收部 21（以下简称为特定分组发送 / 接收部 21）传送、接收和发送特定分组。

特定分组发送 / 接收部 21 在接收到特定分组时将该特定分组中记述的输出端口列表的长度与端口计数相比较，并且判定交换机 20a 自身是否是该特定分组的目的地。在本示例性实施例中，由于端口计数的初始值是 0，所以如果通过向接收到的特定分组中的端口计数加 1 获得的值小于输出端口列表的长度，则判定交换机 20a 自身是该特定分组的目的地。如果通过向接收到的特定分组中的端口计数加 1 获得的值小于输出端口列表的长度，则判定交换机 20a 自身是该特定分组的目的地。输出端口列表的长度是特定分组中记述的输出端口的数目。在判定交换机 20a 自身是特定分组的目的地时，特定分组发送 / 接收部 21 在该特定分组的输出端口列表中记述的所有端口之中只将从端口计数决定的端口更新成交换机 20a 自身的用于接收到该特定分组的端口。在本示例性实施例中，如果特定分组中记述的端口计数的值是 “k”，则输出端口列表中记述的第 (k+1) 个端口被更新成交换机 20a 自身的用于接收到该特定分组的端口。此外，特定分组发送 / 接收部 21 识别输出端口列表中记述的第 k+2 端口，即，紧接被更新的那个之后的端口。此端口是要从交换机 20a 自身发送该特定分组之处。另外，特定分组发送 / 接收部 21 向特定分组中记述的端口计数的值加 1。特定分组发送 / 接
收部 21 接下来从所识别的端口发送特定分组。结果，具有递增的端口计数的特定分组被传送。

[0083] 在判定交换机 20a 自身是特定分组的目的地后，如果该特定分组是请求分组，则特定分组发送 / 接收部 21 将请求分组输出到信息提供部 23；如果该特定分组是设立分组，则特定分组发送 / 接收部 21 将设立分组输出到路由存储部 24；或者如果该特定分组是隧道分组，则特定分组发送 / 接收部 21 将隧道分组输出到交换机侧控制消息隧道传输部 25。

[0084] 另外，特定分组发送 / 接收部 21 发送答复分组。如后文所述，生成答复分组的信息提供部 23 在答复分组中顺次记述到控制单元 10 的路径上的每个交换机的输出端口。特定分组发送 / 接收部 21 从输出端口列表中第一个记述的交换机 20a 自身的端口发送该答复分组。

[0085] 信息提供部 23 在请求分组被从特定分组发送 / 接收部 21 输入到它时，创建包括指示出交换机 20a 自身的连接起来的端口的信息、指示出交换机 20a 的用于接收请求分组的端口的信息和标识交换机 20a 的信息的分组作为答复分组。信息提供部 23 在答复分组中顺次记述从交换机自身延伸到控制单元 10 的路径上的每个交换机的输出端口。以下描述如何决定输出端口列表。即，信息提供部 23 将输入的请求分组中记述的输出端口列表中的最末端端口更新成该交换机的用以接收到该请求分组的端口。此外，信息提供部 23 逆转在更新后的输出端口列表中重排列端口的顺序，并且在答复分组中记述经重排的输出端口列表。信息提供部 23 还在答复分组中记述输出端口列表的长度和端口计数的初始值，然后将特定分组发送 / 接收部 21 发送答复分组。结果，交换机从接收到请求分组的端口发送作为对该请求分组的答复的答复分组。

[0086] 路径存储部 24 在设立分组被从特定分组发送 / 接收部 21 输入到其中时存储充当从该设立分组延伸到控制单元 10 的控制信道的路径。更具体而言，路径存储部 24 将输入的设立分组中记述的输出端口列表中的最末端端口更新成交换机的用于接收到该设立分组的端口。此外，路径存储部 24 逆转在更新后的输出端口列表中重排列端口的顺序，并且存储经重排的输出端口列表作为指示出作为控制信道的路径信息的信息。

[0087] 交换机侧控制消息隧道传输部 25（以下简称为控制消息隧道传输部 25）通过隧道传输来发送和接收在控制单元 10 与交换机之间交换的控制消息。换言之，控制消息隧道传输部 25 对要发送到控制单元 10 的控制消息（例如，指示出接收到在流条目信息中未限定的分组的消息）编码，并且进行经封装的隧道分组。此时，控制消息隧道传输部 25 从路径存储部 24 中读取表示到控制单元 10 的控制信道的输出端口列表，并且将此输出端口列表记述在隧道分组中。控制消息隧道传输部 25 还在隧道分组中记述输出端口列表的长度和端口计数的初始值，然后，控制消息隧道传输部 25 使得特定分组发送 / 接收部 21 发送隧道分组。

[0088] 此外，在隧道分组被从特定分组发送 / 接收部 21 输入到控制消息隧道传输部 25 时，控制消息隧道传输部 25 解封出隧道分组中包括的控制消息（例如流条目信息）并对控制消息解码。

[0089] 特定分组发送 / 接收部 21、信息提供部 23、路径存储部 24 和控制消息隧道传输部 25 可各自独立实现。

[0090] 特定分组发送 / 接收部 21、信息提供部 23、路径存储部 24 和控制消息隧道传输部
同样可由根据用于分组传送单元的程序操作的计算机的 CPU 来实现。在此情况下，例如计算机的程序存储装置（未示出）可存储分组传送单元程序，并且 CPU 可加载程序并根据该程序充当特定分组发送/接收部 21、信息提供部 23、路径存储部 24 和控制消息传输部 25。

接下来，下面描述本发明中的源路由。源路由是一种分组传送方法，其中，要充当发送源的节点在分组中记述通信路径并且中继节点根据分组中记述的通信路径来传送分组。在源路由中，中继节点不需要存储预先提供的路径信息。本发明的通信系统在特定分组中记述输出端口列表，该输出端口列表表示各个节点要用来发送该特定分组的输出端口的顺序阵列。控制单元 10 和每个交换机通过从输出端口列表中记述的端口发送特定分组来实现源路由。控制单元 10 和每个交换机利用端口计数来判定与本地节点有关的输出端口被记述在输出端口列表中的什么位置。以下描述源路由的更具体示例。

图 3 是示出在控制单元向交换机发送设立分组时应使用的源路由的示例的说明图。然而，为了描述起来简单，使图 3 中所示的交换机的拓扑与图 2 所示的拓扑不同。图 3 中的交换机 20A 至 20D 与图 2 中的交换机 20a 至 20d 具有基本上相同的配置。此外，在图 3 中的控制单元 10 和交换机 20A 至 20D 旁边示出的标号表示赋于每个单元的端口标识号。在以下描述中用“#n”表示端口标识号为“n”的端口。

在此示例中，控制单元 10 向作为目的地的交换机 20D 发送设立分组。在该示例中，控制单元 10 首先将端口计数设定为具有初始值 0 并且发送其中记述了输出端口列表 [2, 3, 2, 1] 的设立分组。由于此输出端口列表中的第一端口号是“2”，所以控制单元 10 从控制单元 10 的端口 #2 发送设立分组。此状态在图 3（a）中示出。

交换机 20A 使用端口 #1 来接收该设立分组。由于端口计数的值是 0，所以交换机 20A 接下来将输出端口列表中的第一端口标识号更新成接收到设立分组的端口 #1 的端口标识号。此外，由于端口计数的值是 0，所以交换机 20A 还判定要用来发送设立分组的端口被记述在输出端口列表中的第二位置，并且识别出在第二位置记述的端口标识号“3”。然后，交换机 20A 将端口计数的值从 0 更新到 1，并且从交换机 20A 的端口 #3 发送设立分组。这样在设立分组中记述的输出端口列表是 [1, 3, 2, 1]。此状态在图 3（b）中示出。

交换机 20B 使用端口 #1 来接收此设立分组。由于端口计数的值是 1，所以交换机 20B 接下来将输出端口列表中的第二端口标识号更新成接收到设立分组的端口 #1 的端口标识号。此外，由于端口计数的值是 1，所以交换机 20B 还判定要用来发送设立分组的端口被记述在输出端口列表中的第三位置，并且识别出在第三位置记述的端口标识号“2”。然后，交换机 20B 将端口计数的值从 1 更新到 2，并且从交换机 20B 的端口 #2 发送设立分组。这样在设立分组中记述的输出端口列表是 [1, 1, 2, 1]。此状态在图 3（c）中示出。

交换机 20C 使用端口 #3 来接收此设立分组。由于端口计数的值是 2，所以交换机 20C 接下来将输出端口列表中的第三端口标识号更新成接收到设立分组的端口 #3 的端口标识号。此外，由于端口计数的值是 2，所以交换机 20C 还判定要用来发送设立分组的端口被记述在输出端口列表中的第四位置，并且识别出在第四位置记述的端口标识号“1”。然后，交换机 20C 将端口计数的值从 2 更新到 3，并且从交换机 20C 的端口 #1 发送设立分组。这样在设立分组中记述的输出端口列表是 [1, 1, 3, 1]。此状态在图 3（d）中示出。

交换机 20D 使用端口 #2 来接收此设立分组。由于端口计数的值是 3 并且输出端
口列表中的端口标识号的数目是四，所以交换机 20D 下接判定交换机 20D 自身是目的地。
交换机 20D 还将设立分组的 {1, 1, 3, 1} 输出端口列表中的最末端口标识号“1”更新成接收到
设立分组的端口 #2 的端口标识号。输出端口列表结果变成了 {1, 1, 3, 2}。交换机 20D 按
相反的顺序重排列输出端口列表，并且存储经重列的输出端口列表 {2, 3, 1, 1} 作为表示
通向控制单元 10 的控制信道的输出端口列表。

[0098] 虽然描述了控制单元向交换机发送设立分组的示例，但当控制单元仅向交换机
发送隧道分组或请求分组时，交换机 20A 至 20C 的操作也是基本上相同的。其结果是交换
机 20D 利用端口 #2 接收其中记述了输出端口列表 {1, 1, 3, 1} 的隧道分组或请求分组。

[0099] 然而，在接收到隧道分组时，交换机 20D 可跳过对隧道分组中的输出端口列表的处
理。

[0100] 以下描述在接收到请求分组时交换机 20D 的操作方式。由于端口计数的值是 3 并
且输出端口列表中的端口标识号的数目是四，所以交换机 20D 确定交换机 20D 自身是目的
地。交换机 20D 还将请求分组中记述的 {1, 1, 3, 1} 输出端口列表中的最末端口标识号更新
成接收到请求分组的端口 #2 的端口标识号。输出端口列表结果变成了 {1, 1, 3, 2}。交换机
20D 按相反的顺序重排列输出端口列表，并且在将经重排的输出端口列表 {2, 3, 1, 1} 记
述在答复分组后，与 {2, 3, 1, 1} 中的第一位置相对应的端口 #2 发送答复分组。交换机在
接收到答复分组时的操作与在接收到设立分组或隧道分组时的基本相同，并且此答复分组
到达控制单元 10。

[0101] 已经描述和示出了控制单元 10 发送寻址到交换机 20D 的特定分组的示例。对于
寻址到其他交换机的分组，每个交换机也按与上述基本相同的方式操作。

[0102] 此外，当每个交换机在隧道分组中记述作为通向控制单元 10 的控制信道预先存
储的输出端口列表并发送该隧道分组时，其他交换机与上述类似地操作并且仅次将该隧道
分组传送到控制单元 10。

[0103] 在图 3 中以初始端口计数值 0 为例进行了描述。每个交换机判定要更新输出端口
列表中的第几个端口标识号的方法或者每个交换机判定从哪个端口传送特定分组的方法
可根据端口计数的具体初始值来预设。接收到了特定分组的交换机判定该交换机自身是否
是目的地的方法也可根据端口计数的具体初始值来预设。

[0104] 接下来，下面描述特定分组的格式的示例。图 4 示出了特定分组格式的示例。图 4
中所示的示例涉及按以太网帧（以太网；注册商标）的格式实现特定分组。图 4 中以括号
形式示出的值表示字节数。按以太网帧格式提供了目的地 MAC 地址 51 和源 MAC 地址 52。然
而，本发明的通信系统不使用这些地址来发送/接收特定分组。取而代之，本发明的通信系
统如本示例性实施例中所述通过源路由来发送特定分组。在以太网帧格式的特定分组中，
指示出该分组是特定分组的值被存储到类型信息 53 中。当控制单元 10 或每个交换机生成
特定分组时，指示出该分组是特定分组的值被存储到类型信息 53 中。此外，当控制单元 10
或每个交换机接收到分组时，它们各自以在类型信息 53 内存储了指示出该分组是特定分
组的值这个事实为条件判定接收到的分组是特定分组。

[0105] 以太网帧格式的数据 54 包括标识特定分组的种类的类型代码 61。数据 54 还包括
第一数据长度 62、输出端口列表的长度 63、端口计数 64、输出端口列表 65、第二数据长度 66
以及特定分组的数据 67。为了与以太网帧格式中的类型信息 53 相区分，以下将把标识特定
分组的种类的类型代码 61 简称为类型 61。此外，为了与以太网帧格式中的数据 54 相区分，
以下将把特定分组的数据 67 称为特定分组数据 67。第一数据长度 62（在图 4 中示为“长
度”）以后的全部信息的数据长度被存储到第一数据长度 62 中。第二数据长度 66（在图 4
中示为“数据长度”）以后的全部信息的数据长度被存储到第二数据长度 66 中。
【0106】 输出端口列表的长度 63（在图 4 中示为“端口长度”）是输出端口列表中包括的端
口的数目。更具体而言，长度 63 是输出端口列表中包括的端口标识号的数目。
【0107】 端口计数 64（在图 4 中示为“端口计数”）是由特定分组的发送源的跳数所决定的
计数值。
【0108】 端口标识号以排列形式被存储在输出端口列表 65 内。
【0109】 下面对于每种特定分组描述以太网帧格式的数据 54 中包括的信息的示例。然而，
省略对第一数据长度 62、输出端口列表的长度 63、端口计数 64、输出端口列表 65 和第二数
据长度 66 的描述。
【0110】 图 5 是示出请求分组中包括的信息的示例的说明图。指出分组是请求分组的值
被存储到类型 61 中。控制单元 10 在生成请求分组时将指示出该分组是请求分组的值存储
到类型 61 中。此外，每个交换机在接收分组时以在类型 61 内存储了指出该分组是请求
分组的值这个事实为条件判定接收的分组是请求分组。
【0111】 对于请求分组，特定分组数据 67 是空（NULL）。
【0112】 图 6 是示出答复分组中包括的信息的示例的说明图。指出分组是答复分组的值
被存储到类型代码 61 中。每个交换机在生成答复分组时将指示出该分组是答复分组的值
存储到类型代码 61 中。此外，控制单元 10 在接收到分组时以在类型代码 61 内存储了指出
该分组是答复分组的值这个事实为条件判定接收到的分组是答复分组。
【0113】 对于答复分组，表示被连接起来的端口的信息、表示交换机的用于接收请求分组
的端口的信息以及标识要发送答复分组的交换机的信息被存储到特定分组数据 67 中。在
答复分组创建期间，要发送答复分组的交换机将这三种信息存储到特定分组数据 67 中。在
图 6 的示例中，“数据路径 ID”相当于交换机标识信息。另外，“请求端口”相当于请求分组
接收端口。此外，“端口 #n 状态”表示具有端口标识号 “n” 的端口是否被链接起来。例如，
交换机可为每个端口记述“端口 #n 状态”的值，使得对于被链接起来的端口“端口 #n 状态”
等于 1，并且使得对于未被链接起来的其他端口“端口 #n 状态”等于 0。于是可从逐端口记述
的“端口 #n 状态”值的集合来定位被链接起来的端口。然而，可以使用其他表示方法来
记述被链接起来的端口。
【0114】 图 7 是示出设立分组中包括的信息的示例的说明图。指出分组是设立分组的值
被存储到类型代码 61 中。控制单元 10 在生成设立分组时将指示出该分组是设立分组的值
存储到类型代码 61 中。此外，每个交换机在接收到分组时以在类型代码 61 内存储了指出
该分组是设立分组的值这个事实为条件判定接收到的分组是设立分组。
【0115】 对于设立分组，特定分组数据 67 是空。
【0116】 图 8 是示出隧道分组中包括的信息的示例的说明图。指出分组是隧道分组的值
被存储到类型 61 中。控制单元 10 或每个交换机在生成隧道分组时将指示出该分组是隧道
分组的值存储到类型代码 61 中。此外，控制单元 10 或每个交换机在接收到分组时以在类
型代码 61 内存储了指出该分组是隧道分组的值这个事实为条件判定接收到的分组是隧
隧道分组在特定分组数据67中存储了控制消息。具体而言，在OpenFlow方案的一般通信系统中作为控制消息交换的控制分组被存储到分组数据67中。在本发明的通信系统中，与一般通信系统中一样，也在控制单元与一个交换机之间交换控制消息。

以下描述操作。

图9至13是示意性地示出控制单元生成拓扑信息，并形成控制信道的处理步骤的说明图。参考图9至13的示图，为了方便起见用标号A来标记控制单元，并且同样地用标号B至E之一来标记每个交换机。控制单元A和每个交换机B至E的配置与图2中所示的相应配置基本相同。另外，从图9到图13的每一幅图中在控制单元A和每个交换机B旁边示出的数字表示端口标识号。此外，从图9到图13的每一幅图中判定为未检测到目的端的端口的端口标识号被各自包围在圆圈中。另外，在下面给出的描述中，连同控制单元A和交换机B至E的标号和端口标识号，控制单元A和交换机B至E的端口可被表述为“#n@A”, “#n@B”等等。跟在“#”后的数字表示端口标识号，并且跟在“@”后的符号标识出控制单元或交换机。例如，“#1@A”指的是控制单元10的端口1，“#2@B”指的是交换机B的端口2。

图14是示意性地表示控制单元A和交换机B至D交换请求分组和答复分组的处理步骤的实例的说明图。参考图14，带箭头的虚线表示请求分组交换，并且带箭头的实线表示答复分组交换。图14中每个标有箭头的线两端的与“#”一起示出的数字表示发送或接收请求或答复分组的端口的标识号。图14只示出了正在进行的请求分组和答复分组交换的示例，而每个分组的发送序列并不限于图14中所示的那些。例如，在图14中，控制单元A在与交换机B交换请求分组和答复分组之前，与交换机D交换请求分组和答复分组，然而此序列可被逆转。否则，控制单元A可并行地向交换机B和D两者发送请求分组，然后从两个交换机接收答复分组。

在其初始状态下，控制单元A的拓扑搜索部12选择控制单元A的连接起来的端口作为未检测连接目的地端口，这些端口是不知道它们是否已经连接到哪个交换机的哪个端口的。本示例中，拓扑搜索部12选择#1@A（控制单元A的端口1）和#2@A（控制单元A的端口2），如图9中所示。控制单元A的特定分组发送/接收部11随后从两个未检测连接目的地端口#1@A和#2@A中的每一个发送请求分组，如图9和14中所示。拓扑搜索部12在从#1@A发送的请求分组中记述(#1)作为源路由输出端口列表，并且类似地在从#2@A发送的请求分组的输出端口列表中记述(#2)。

交换机D的特定分组发送/接收部21使用#1@D（交换机D的端口1）来接收从端口#1@A发送的请求分组。交换机D的信息提供部23随后生成包括标识交换机D自身的信息、指示出用来接收到请求分组的端口#1的信息和与交换机D中的连接起来的端口有关的信息的答复分组。此时，信息提供部23在答复分组中记述#1, #2和#3作为与交换机D中的连接起来的端口有关的信息。信息提供部23还在答复分组中记述#1作为源路由输出端口列表。此发送状态在图10中示出。此答复分组在图14中被示为“答复(#1@D,端口:#1, #2, #3)。”

类似地，交换机B的特定分组发送/接收部21使用#2@B（交换机B的端口2）来接收从端口#2@A发送的请求分组。交换机B的信息提供部23随后生成包括标识交换机B
自身的信息，指示出用来接收到请求分组的端口 
#2 的信息和与交换机 B 中的链路连接的端口 
有关的信息的答复分组。此时，信息提供部 23 在答复分组中记述 
#1, #2 和 #3 作为与交换机 B 中的链路连接的端口有关的信息。信息提供部 23 还在答复分组中记述 
(#2) 作为源路由输出端口列表。这使得交换机 B 的特定分组发送 / 接收部 21 从用来接收到请求分组的 
端口 #2@B 发送答复分组。此发送状态在图 10 中出示。此答复分组在图 14 中被示为“答复 
（源 :#2@B，端口 :#1, #2, #3）”。

[0124] 通过控制单元 A 的用于发送请求分组的端口，控制单元 A 的特定分组发送 / 接收部 11 接收作为对该请求分组的答复从交换机 B 发送来的答复分组。控制单元 A 的拓扑搜索部 12 随后判定答复分组发送源交换机的，即交换机用于接收到请求分组的端口将连接到在发送该请求分组前选择的未检测连接目的地端口。控制单元 A 的拓扑搜索部 12 接下来将这两个端口之间的连线关系作为拓扑信息存储到拓扑存储部 13 中。此外，控制单元 A 的拓扑搜索部 12 选择答复分组发送源交换机的端口之一作为未检测连接目的地端口，并且生成从该未检测连接目的地端口发送的请求分组。拓扑搜索部 12 只需在此请求分组中记述 
控制单元 A 的输出端口和形成到所选择的未检测连接目的地端口的路径的每个交换机的 
输出端口。其结果是控制单元 A 的特定分组发送 / 接收部 11 发送该请求分组。

[0125] 在上述示例中，控制单元 A 的特定分组发送 / 接收部 11 在端口 #1@A 接收从端口 
#1@D 发送的答复分组。控制单元 A 的拓扑搜索部 12 随后判定端口 #1@D 连接到在发送请求 
分组前选择的未检测连接目的地端口 #1@A。控制单元 A 的拓扑搜索部 12 随后将其连接状态 
存储到拓扑存储部 13 中。此外，控制单元 A 的拓扑搜索部 12 从发送该答复分组的交换机 
D 的端口 #1, #2, #3 之中选择端口 #2 和 #3 作为未检测连接目的地端口（参见图 11）。拓扑搜索部 12 生成从未检测连接目的地端口 #2@D 发送的请求分组。拓扑搜索部 12 在此请 
求分组中记述 (#1, #2) 作为用于源路由的输出端口列表，并且控制单元 A 的特定分组发送 / 接收部 11 从在该输出端口列表中的第一位置处记述的端口 #1 发送该请求分组。交换机 D 
的特定分组发送 / 接收部 21 接收该请求分组并从在该输出端口列表中的第二位置处记述的 
端口 #2 传送该请求分组（参见图 11 和 14）。控制单元 A 的拓扑搜索部 12 还生成从未检测 
连接目的地端口 #3@D 发送的请求分组。拓扑搜索部 12 在此请求分组中记述 (#1, #3) 作 
为用于源路由的输出端口列表，并且控制单元 A 的特定分组发送 / 接收部 11 从在该输出端 
口列表中的第一位置处记述的端口 #1 发送该请求分组。交换机 D 的特定分组发送 / 接收 
部 21 接收该请求分组并从在该输出端口列表中的第二位置处记述的端口 #3 传送该请求分 
组（参见图 11 和 14）。

[0126] 类似地，控制单元 A 在端口 #2@A 接收从端口 #2@B 发送的答复分组。控制单元 A 接 
下来判定端口 #2@B 连接到未检测连接目的地端口 #2@A，并且存储此连接状态。此外，控制 
单元 A 从发送该答复分组的交换机 B 的端口 #1, #2, #3 之中选择端口 #1 和 #3 作为未检 
测连接目的地端口（参见图 11）。控制单元 A 生成从未检测连接目的地端口 #1@B 发送的请 
求分组。控制单元 A 在此请求分组中记述 (#2, #1) 作为用于源路由的输出端口列表，并且从在该输出端口列表中的第一位置处记述的端口 #2 发送该请求分组。交换机 B 接收该请 
求分组并从在该输出端口列表中的第二位置处记述的端口 #1 传送该请求分组（参见图 11 
和 14）。控制单元 A 还生成从未检测连接目的地端口 #3@B 发送的请求分组。控制单元 A 在 
此请求分组中记述 (#2, #3) 作为用于源路由的输出端口列表，并且从在该输出端口列表中
第一位置处记述的端口 #2 发送该请求分组。交换机 B 接收该请求分组并从在该输出端口列表中的第二位置处记述的端口 #3 传送该请求分组（参见图 11 和 14）。

[0127] 交换机 C 的特定分组发送 / 接收部 21 使用端口 #1OC 来接收从端口 #20D 传来的请求分组，并且使用端口 #20C 来接收从端口 #30B 传来的请求分组。交换机 C 的信息提供部 23 随后为用于接收从接收的每个端口生成包括标识交换机 C 自身的信息、指示出用于接收请求分组的端口的信息和关于交换机 C 中的连接起来的端口的信息的答复分组。此时，信息提供部 23 在每个答复分组中记述 #1 和 #2 作为与交换机 C 中的连接起来的端口有关的信息。在本示例中，信息提供部 23 在与从端口 #20D 传来的请求分组相对应的答复分组中记述 #1,1 作为输出端口列表。交换机 C 的特定分组发送 / 接收部 21 从端口 #1 发送该答复分组，然后交换机 B 从端口 #1 发送该答复分组，并且该答复分组到达控制单元 A（参见图 12 和 14）。此答复分组在图 14 中被示为“答复（源：#10C，端口：#1,1）”。类似地，交换机 C 在与从端口 #30B 传来的请求分组相对应的答复分组中记述 #2,2 作为输出端口列表。交换机 C 从端口 #2 发送该答复分组，然后交换机 B 从端口 #2 发送该答复分组，并且该答复分组到达控制单元 A（参见图 12 和 14）。此答复分组在图 14 中被示为“答复（源：#20C，端口：#1,2）”。

[0128] 控制单元 A 接收到的两个答复分组“答复（源：#10C，端口：#1,2）”和“答复（源：#20C，端口：#1,2）”都包括指示出存在端口 #1,2 作为交换机 C 的端口的信息。控制 A 接收到的答复分组因此在内容上有重复。控制单元 A 可以在留下这种重复信息的情况下生成拓扑信息。或者，控制单元 A 在生成拓扑信息之前可删除重复信息的部可以去除重叠。

[0129] 交换机 E 与交换机 C 类似地操作。即，交换机 E 的特定分组发送 / 接收部 21 使用端口 #10E 来接收从端口 #30D 传来的请求分组，并且使用端口 #20E 来接收从端口 #10B 传来的请求分组。交换机 E 的信息提供部 23 随后为请求分组的每个接收端口生成包括标识交换机 E 自身的信息、指示出请求分组接收端口的信息和关于交换机 C 中的连接起来的端口的信息的答复分组。此时，信息提供部 23 在每个答复分组中记述 #1 和 #2 作为与交换机 E 中的连接起来的端口有关的信息。在本示例中，信息提供部 23 在与从端口 #30B 传来的请求分组相对应的答复分组中记述 #1,1 作为输出端口列表。交换机 E 的特定分组发送 / 接收部 21 从端口 #1 发送该答复分组，然后交换机 B 从端口 #1 发送该答复分组，并且该答复分组到达控制单元 A （参见图 12 和 14）。此答复分组在图 14 中被示为“答复（源：#10E，端口：#1,2）”。类似地，交换机 E 在与从端口 #10B 传来的请求分组相对应的答复分组中记述 #2,2 作为输出端口列表。交换机 E 从端口 #2 发送该答复分组，然后交换机 B 从端口 #2 发送该答复分组，并且该答复分组到达控制单元 A（参见图 12 和 14）。此答复分组在图 14 中被示为“答复（源：#20E，端口：#1,2）”。

[0130] 控制单元 A 的特定分组发送 / 接收部 21 使用端口 #10A 来接收从端口 #10C 发送来的答复分组。控制单元 A 的拓扑搜索部 12 判定端口 #10C 连接到在发送请求分组前选择的未检测连接目的地端口 #20D，并将此信息添加到拓扑信息。类似地，控制单元 A 使用端口 #20A 来接收从端口 #20C 发送来的答复分组。控制单元 A 判定端口 #20C 连接到在发送请求分组前选择的未检测连接目的地端口 #30B，并将此信息添加到拓扑信息。

[0131] 控制单元 A 的特定分组发送 / 接收部 11 还使用端口 #10A 来接收从端口 #10E 发送来的答复分组。控制单元 A 的拓扑搜索部 12 判定端口 #10E 连接到在发送请求分组前选
择的未检测连接目的地端口 #3©D，并将此信息添加到拓扑信息。类似地，控制单元A使用端口 #2©A 来接收从端口 #2©E 发送来的答复分组。控制单元A判定端口 #2©E 连接到在发送请求分组前选择的未检测连接目的地端口 #1©B，并将此信息添加到拓扑信息。

【0132】结果，所有未检测连接目的地端口从拓扑信息中消失，并且控制单元A的拓扑搜索部12判定拓扑信息已完成。控制单元A的控制信路径确定部14为每个交换机设定独立控制信道。例如，控制信路径确定部14使用迪杰斯特拉的算法来计算到交换机的最短路径，并将此路径设定为用于该交换机的控制信道（参见图13）。此外，控制单元A的控制信路径确定部14为每个交换机B至E创造设立分组，并且特定分组发送/接收部11发送寻找从交换机B至E的设立分组。每个交换机B至E的路径存储部24在收到设立分组后，设立分组存储去到控制单元A的路径作为控制信道。虽然在图13中作为示例以粗线链路的形式示出了控制单元A为交换机C、E设定的控制信道，但这些控制信道并不限于图13中所示的那些。

【0133】控制信道A和交换机B至E随后经由控制信道通过隧道传输到交换机类控制消息。

【0134】下面描述在上述操作期间生成的拓扑信息的数据结构的示例。拓扑信息描述哪个交换机的哪个端口连接到哪个交换机的哪个端口。拓扑搜索部12利用其中互连的交换机的端口被清楚指示为诸如 (#1©A, #1©D) 这样的一对的数据结构来生成拓扑信息。可与意指“未知端口”的数据（此数据在本示例中称为“?”）相组合来指明未检测连接目的地端口。例如，(#3©D, ?) 的组合可用于表示端口 #3©D 是未检测连接目的地端口。或者，可利用如下数据结构来指明未检测连接目的地端口：在该数据结构中，不配对或组合未检测连接目的地端口，而是只包括指示出这些未检测连接目的地端口的信息。以下，与“?”相组合指明未检测连接目的地端口的数据结构被称为第一数据结构。不配对或组合未检测连接目的地端口而是只包括指示出这些未检测连接目的地端口的信息的数据结构被称为第二数据结构。可按第一数据结构或第二数据结构或者按任何其他描述格式的数据结构来记录拓扑信息。

【0135】以下描述作为图14中所示的请求分组和答复分组交换的结果生成拓扑信息的处理步骤。在下面每个示例中都描述按第一数据结构记录拓扑信息和按第二数据结构记录拓扑信息两者。

【0136】图 15 示出了在通信系统的初始状态下创建的拓扑信息的示例。在以第一数据结构记录时，初始状态下的拓扑信息被表述为一对端口 (#1©A, ?), (#2©A, ?)。在以第二数据结构记录时，不存在端口的对或组合，并且端口 #1©A, #2©A 被表述为未检测连接目的地端口。

【0137】图 16 示出了在接收到从端口 #1©D 发送来的答复分组时控制单元A创建的拓扑信息的示例。基于此答复分组，添加端口 #1©A 和 #1©D 相连接的信息。此外，添加端口 #2©D, #3©D 作为未检测连接目的地端口。

【0138】图 17 示出了在接收到从端口 #2©B 发送来的答复分组时控制单元A创建的拓扑信息的示例。基于此答复分组，添加端口 #2©A 和 #2©B 相连接的信息。此外，添加端口 #1©B, #3©B 作为未检测连接目的地端口。

【0139】图 18 示出了在接收到从端口 #2©C 发送来的答复分组时控制单元A创建的拓扑信息的示例。基于此答复分组，添加端口 #2©D 和 #1©C 相连接的信息。此外，添加端口 #2©C 作为未检测连接目的地端口。

【0140】图 19 示出了在接收到从端口 #1©E 发送来的答复分组时控制单元A创建的拓扑信
息的示例。基于此答复分组，添加端口 #3@D 和 #1@E 相连接的信息。此外，添加端口 #2@E 作为未检测连接目的地端口。

[0141] 图 20 展示了在接收到从端口 #2@E 发送来的答复分组时控制单元 A 创建的拓扑信息的示例。基于此答复分组，添加端口 #1@B 和 #2@E 相连接的信息。

[0142] 图 21 展示了在接收到从端口 #2@C 发送来的答复分组时控制单元 A 创建的拓扑信息的示例。基于此答复分组，添加端口 #3@B 和 #2@C 相连接的信息。在一个添加后不存在未检测连接目的地端口。

[0143] 通信系统周期性地通过从开始重新生成拓扑信息来为交换机的最新拓扑创建适当的拓扑信息，并且基于新的拓扑信息来设定控制信号。例如，控制单元 A 的拓扑搜索部 12 周期性地删除拓扑存储部 13 内存储的拓扑信息。该删除使处理器返回到其初始状态。然后，拓扑搜索部 12 重新开始从生成请求分组起的处理。结果，即使拓扑信息创建之后交换机的拓扑有变化，也可通过重新创建包含了新交换机拓扑的拓扑信息或者通过包含新交换机拓扑来设定控制信号。

[0144] 下面参考图 2, 22 和 23 来详述控制单元和交换机的操作。

[0145] 图 22 是示出控制单元 10 中的处理流程的示例的流程图。此流程图假定控制单元 10 处于其初始状态中，或者在控制单元中未存储拓扑信息。控制单元 10 的拓扑搜索部 12 选择未检测连接目的地端口 (步骤 S1)。拓扑搜索部 12 最初选择控制单元 10 的链接起来的端口。如果在拓扑存储部 13 内存储有拓扑信息，则拓扑搜索部 12 参考拓扑信息并且在每个交换机的链接起来的所有端口之中只选择不知道该端口连接到哪个交换机的哪个端口的端口。

[0146] 拓扑搜索部 12 判定在步骤 S1 中对未检测连接目的地端口的选择是否成功 (步骤 S2)。如果对未检测连接目的地端口的选择成功 (步骤 S2 中的判定结果为“是”)，则拓扑搜索部 12 生成从未检测连接目的地端口发送的请求分组，并且使得特定分组发送 / 接收部 11 发送该分组 (步骤 S3)。如果在步骤 S1 中选择了多个未检测连接目的地端口，则拓扑搜索部 12 为每个未检测连接目的地端口生成请求分组。如果未检测连接目的地端口是控制单元 10 的，则拓扑搜索部 12 在请求分组中记述该端口的端口标识号作为输出端口列表，然后记述 1 作为输出端口列表的长度。此外，控制单元 10 重复步骤 S1 至 S3 的循环处理，以按从具有离控制单元 10 的最小跳数的 (一个或多个) 交换机起的顺序判定端口间连接关系，该判定是以控制单元 10 的端口作为其起点进行的。因此，如果未检测连接目的地端口是交换机的，则拓扑搜索部 12 可识别从控制单元 10 的端口到未检测连接目的地端口的路径上的每个交换机的输出端口。因此，拓扑搜索部 12 在输出端口列表中顺序记述从控制单元 10 的该具体端口到未检测连接目的地端口的这些输出端口的端口标识号。拓扑搜索部 12 在请求分组中记述输出端口的数目作为输出端口列表的长度。此外，拓扑搜索部 12 在所有请求分组中记述初始端口计数值 0。特定分组发送 / 接收部 11 将每个请求分组从在该请求分组的输出端口列表中记述的第一端口发送。

[0147] 每个请求分组在在步骤 S3 中被发送之后到达连接到在请求分组的输出端口列表中记述的最末端端口也就是未检测连接目的地端口的交换机，并且该交换机从该交换机接收到该请求分组的端口发送寻址到控制单元 10 的答复分组。该答复分组包括标识出发送了该答复分组的交换机的信息、指示出该交换机中的链接起来的端口的信息和指示出该交换
机用来接收到该请求分组的端口的信息。

在步骤 S4 中，控制单元 10 的控制分组发送/接收部 11 接收从请求分组目的地交换机（即连接到未检测连接目的地端口的交换机）发送的答复分组。特定分组发送/接收部 11 在判定接收到的特定分组是答复分组时将该答复分组输出到拓扑搜索部 12。

拓扑搜索部 12 基于从特定分组发送/接收部 11 输入的答复分组来更新拓扑信息（步骤 S5）。即，拓扑搜索部 12 通过将这些关系存储到拓扑存储部 13 中来向拓扑信息添加从输入的答复信息明确的新的端口间连接关系。

如下所述，在步骤 S5 中只需要拓扑搜索部 12 基于输入的答复信息识别端口间连接关系。由于未检测连接目的地端口的端口标识号被记述在请求分组的输出端口列表中的最末位置处，所以此分组以控制单元 10 的端口作为其起点被顺次传送到每个交换机，并且到达连接到未检测连接目的地端口的交换机。从此交换机发送的答复分组包括标志该交换机自身的端口的信息、指示出该交换机中的链路起来的端口的信息和指示出该交换机用来接收到该请求分组的端口的信息。拓扑搜索部 12 因此查看通过与请求分组的发送路径相同的路径传送的答复分组。拓扑搜索部 12 还判定答复分组的发送源交换机连接到在请求分组的输出端口列表中的最末位置处记述的未检测连接目的地端口，并且判定该具体交换机连接到该交换机用来接收到请求分组的端口。这明确了未检测连接目的地端口连接到哪个交换机的哪个端口。拓扑搜索部 12 向拓扑信息添加未检测连接目的地端口和被判定与其相连/交换机的端口的连接的连接的端口之间的连接关系。拓扑搜索部 12 还拓扑信息添加与被判定连接到未检测连接目的地端口的交换机（即答复分组的发送源交换机）的每个连接起来的端口有关的信息。下一次进行步骤 S1 的操作时，拓扑搜索部 12 在被判定连接到未检测连接目的地端口的交换机的每台交换机的所有连接起来的端口之中只选择不知道该端口连接到哪个交换机的哪个端口的端口作为未检测连接目的地端口。

在步骤 S5 之后，控制单元 10 重复从步骤 S1 起往后的处理。

在拓扑信息中不存在未检测连接目的地端口并且对未检测连接目的地端点的选择失败（步骤 S2 中的“否”），则这意味着拓扑信息已完成。在此情况下，控制信道路径确定部 14 查看拓扑存储部 13 内存储的已完成的拓扑信息，并且设定通过每个交换机的控制信道。对于设定信道的方法没有限制。控制信道路径确定部 14 可计算到每个交换机的最短路径，并且将计算出的最短路径设定为控制信道。控制信道路径确定部 14 为每个交换机生成其中记录了指示出通向该交换机的控制信道的输出端口列表的设立分组。控制信道路径确定部 14 还在设立分组中记述初始端口计数值 0。控制信道路径确定部 14 还记述输出端口列表的长度。控制信道路径确定部 14 使得特定分组发送/接收部 11 发送逐交换机生成的设立分组（步骤 S6）。此时，特定分组发送/接收部 11 从设立分组的输出端口列表中记述的第一个端口发送每个设立分组。

接收设立分组的交换机基于设立分组来确定去到控制单元 10 的控制信道，并且存储该信道。这使得控制单元 10 和每个交换机能够认识到控制信道。当控制单元 10 发送控制信息时，控制信道路由传输部 15 通过编码和封装该控制信息来生成隧道分组。此时，控制信道路由传输部 15 在隧道分组中记述表示控制信道的输出端口列表、输出端口列表的长度和初始端口计数值 0。然后控制信道路由传输部 15 使得特定分组发送/接收部 11 发送隧道分组。特定分组发送/接收部 11 从隧道分组的输出端口列表中记述的第一个
端口发送隧道分组。在交换机发送控制消息的情况下，控制消息隧道传输部 25 和特定分组发送 / 接收部 21 与上述类似地操作。

[0154] 此外，在判定接收到来自交换机的隧道分组时，控制单元 10 的特定分组发送 /接收部 11 将隧道分组输出到控制消息隧道传输部 15。控制消息隧道传输部 15 解封出控制消息并且对解封出的控制消息解码。控制单元 10 执行作为解码出的控制消息的函数的适当处理。

[0155] 拓扑搜索部 12 周期性地删除拓扑存储部 13 内存储的拓扑信息。该删除使处理返回到其初始状态。然后，从初始状态起重新开始从步骤 S1 起往后的处理。结果，控制单元 10 周期性地再一次生成拓扑信息。因此，在创建拓扑信息后，即使交换机的拓扑变化了，控制单元 10 也可将新拓扑包含到已创建的新拓扑信息中。控制单元 10 还可设定包含了新拓扑的控制信息。

[0156] 图 23 是示出交换机中的处理流程的示例的流程图。在交换机接收到特定分组时，特定分组发送 / 接收部 21 判定接收到该特定分组的交换机自身是否是其目的地（步骤 S21）。如果交换机自身是目的地（步骤 S21 中的“是”），则特定分组发送 / 接收部 21 辨别特定分组的类型（步骤 S22）。

[0157] 如果特定分组是请求分组，则特定分组发送 / 接收部 21 将请求分组输出到信息提供部 23（步骤 S23）。信息提供部 23 还将生成包括与交换机自身的链路起的端口有关的信息、标识交换机的信息和用来接收信号请求分组的端口的标识号的答复分组。信息提供部 23 还将接收的请求分组的输出端口列表中记录的最末端口标识号更新成用来接收到请求分组的端口的端口标识号。此外，输出端口列表中的端口标识号被按相反的顺序重排列。信息提供部 23 将这个重排列的输出端口列表记录在答复分组中。信息提供部 23 还在答复分组中记录输出端口列表的长度和端口计数的初始值，并且使得特定分组发送 / 接收部 21 发送答复分组（步骤 S24）。在步骤 S24 中，特定分组发送 / 接收部 21 从输出端口列表中写入的第一端口标识号所表示的端口发送答复分组。结果，答复分组通过请求分组的传送路径按相反的方向被传送并到达控制单元 10。

[0158] 如果特定分组是设立分组，则特定分组发送 / 接收部 21 将设立分组输出到路径存储部 24（步骤 S25）。路径存储部 24 将设立分组的输出端口列表中记录的最末端口标识号更新成用来接收到设立分组的端口的端口标识号。路径存储部 24 还按相反的顺序重排列输出端口列表的端口标识号，并且存储经重排列的输出端口列表作为指示出成为控制信息的路径的信息（步骤 S26）。

[0159] 如果特定分组是请求分组，则特定分组发送 / 接收部 21 将隧道分组输出到控制消息隧道传输部 25（步骤 S27）。控制消息隧道传输部 25 解封出隧道分组中包括的控制消息并且对解封出的控制消息解码（步骤 S33）。交换机执行作为解码出的控制消息的函数的适当处理。

[0160] 如果接收到特定分组的交换机自身不是其目的地（步骤 S21 中的“否”），则特定分组发送 / 接收部 21 在输出端口列表中的所有端口标识号之中仅将按端口计数确定的位置处的端口标识号替换为用来接收到该特定分组的端口的端口标识号。特定分组发送 / 接收部 21 还向特定分组中的端口计数加 1（步骤 S29）。此外，特定分组发送 / 接收部 21 发送该特定分组（步骤 S30）。
在执行步骤 S24、S26、S28 或 S30 后，特定分组接收处理完成。

根据本发明，控制单元与交换机之间的控制信道是经由每个交换机用来传送数据分组的通信网络设定的。换言之，交换机用来传送数据分组的通信网络是利用一种通信网络实现的。这也产生了下面所述的额外有利效果。

第一，冗余性和容错性改善了。这是因为，即使在交换机用来传送数据分组的通信网络的一部分中发生故障的情况下，如果确保了从控制单元 10 到交换机的路径，则也还是能够设定控制信道。例如，在图 2 所示的拓扑中，即使控制单元 10 与交换机 20b 之间的链路故障清除了在两者之间建立的链接起来状态，控制单元 10 也从交换机 20a 按顺序向交换机 20a 至 20d 发送请求分组。这建立了控制单元 10 与每个交换机之间的控制信道。

下面论述与图 27 中作为示例示出的一般通信系统的比较。如图 27 中所示，在独立于分组传送通信网络 94 提供控制通信网络 93 的配置中，即使物理上存在使分组能够从控制单元 91 到达交换机 92 的路径，这些路径也不用于控制。从而，控制通信网络 93 中的故障导致了控制单元 91 中的控制故障，从而降低了容错性。反之不同，根据本发明，由于前述原因，冗余性和容错性改善了。

第二，因为不需为控制信道提供专门的通信网络，所以不需要管理这样的专门通信网络，从而可以降低通信系统的管理成本。如果交换机的数目增加或减少，控制单元 10 也周期性地删除拓扑信息并处理返回到其初始状态。此外，重新开始从步骤 S1(参见图 22) 起的处理使得能够创建特定时间点的最新拓扑信息，因此能够设定通向每个交换机的控制信道。

下面论述与图 27 中作为示例示出的一般通信系统的比较。在图 27 中作为示例示出的通信系统中，由于控制通信网络 93 和分组传送通信网络 94 两者都需要管理，所以管理成本增大了。例如，如果通过添加交换机的群组的拓扑，则必须向控制通信网络 93 和分组传送通信网络 94 两者添加链路以便适当变化。然而，在本发明中，对前述序列的使用降低了管理通信系统所需的成本。

第三，即使在交换机的数目增加之后，如上所述，控制单元 10 也可通过发送请求分组来创建最新的拓扑信息。此点，由于只有控制单元 10 是请求分组的发送源，所以与诸如采用 STP 中的 BPDU 之类的情况相比，与交换机数目增加时的特定分组交换相关联的流量可被抑制到更低的水平。此优点进而使得可以防止通信系统在开始拓扑信息的创建之前花费太多时间，并且防止通信线路施加沉重的负担。

下面论述与图 27 中作为示例示出的一般通信系统的比较。在图 27 中作为示例示出的一般通信系统中，如果交换机的拓扑变化，则由于通信系统用于重新拓扑控制帧（例如 STP 中的 BPDU）被在交换机之间交换，所以使用太多交换机将引起该控制帧的频繁生成并且改善此状态需要大量的时间或者导致通信线路被施加沉重的负担。然而，在本发明中，由于与交换机数目增加时的特定分组交换相关联的流量如上所述与诸如采用 STP 中的 BPDU 之类的情况相比能够被抑制到更低的水平，所以对于通信系统的规模能够实现可扩展性。

接下来，下面描述本发明的最低限度配置的示例。图 24 是示出根据本发明的通信系统的最低限度配置的示例的图。图 25 是示出本发明中的最低限度控制单元配置的示例的框图。图 26 是示出本发明中的最低限度分组传送单元配置的示例的框架。
[0170] 根据本发明的通信系统包括多个分组传送装置 90（例如交换机 20a 至 20d）和控制每个分组传送单元 90 的控制单元 80（例如控制单元 10）。

[0171] 控制单元 80 包括拓扑信息存储装置 81、未检测连接目的地端口确定装置 82、答复请求发送装置 83 和答复请求接收装置 84（参见图 24 和 25）。

[0172] 拓扑信息存储装置 81（例如拓扑存储部 13）存储包括关于控制单元 80 或分组传送单元 90 的端口的互连信息的拓扑信息。

[0173] 未检测连接目的地端口确定装置 82（例如拓扑搜索部 12 中的执行步骤 S1 的部分）根据拓扑信息在控制单元 80 或分组传送单元 90 的所有端口之中仅确定未检测连接目的地端口，未检测连接目的地端口是在其连接目的地搜索期间尚未成功检测到其连接目的的端口。

[0174] 答复请求发送装置 83（例如拓扑搜索部 12 和 /或控制单元侧特定分组发送 /接收部 11）经由未检测连接目的地端口向未检测连接目的地端口的连接目的地发送答复请求（例如请求分组）。

[0175] 答复接收装置 84（例如控制单元侧特定分组发送 /接收部 11）从成为未检测连接目的地端口的连接目的地的分组传送单元接收包括关于该分组传送单元的端口的信息的答复（例如答复分组）。

[0176] 分组传送单元 90 包括传送判定装置 91、传送装置 92 和答复发送装置 93。

[0177] 传送判定装置 91（例如交换机侧特定分组发送 /接收部 21 中的用于执行步骤 S21 的部分）根据接收到的答复请求的内容来判定该答复请求是否要被传送到另外的分组传送单元。

[0178] 传送装置 92（例如交换机侧特定分组发送 /接收部 21 中的用于执行步骤 S30 的部分）在答复请求被判定要被传送到任何其他分组传送单元时，根据答复请求的内容将答复请求传送到该其他分组传送单元。

[0179] 答复发送装置 93（例如交换机侧特定分组发送 /接收部 21 和 /或信息提供部 23）在答复请求被判定不要被传送到任何其他分组传送单元时，返回包括关于该分组传送单元的端口的信息的答复，该答复被沿着通向控制单元的路径发送。

[0180] 上述示例性实施例包括在下面第一 (1) 至 (10) 项中描述的那些作为特征的通信系统配置。也包括在随后的第 (11) 至 (13) 项中描述的那些作为特征的控制单元配置。还包括在随后的第 (14) 和 (15) 项中描述的那些作为特征的分组传送单元配置。

[0181] 本发明的通信系统包括多个分组传送装置（例如交换机 20a 至 20d）和控制每个分组传送单元的控制单元（例如控制单元 10），其中：控制单元包括：拓扑信息存储装置（例如拓扑存储部 13），该拓扑信息存储装置存储包括关于控制单元或每个分组传送单元的端口的互连信息的拓扑信息；未检测连接目的地端口确定装置（例如拓扑搜索部 12 中的用于执行步骤 S1 的部分），该未检测连接目的地端口确定装置根据拓扑信息在控制单元或每个分组传送单元的所有端口之中仅确定未检测连接目的地端口，未检测连接目的地端口是连接目的地未被检测的端口；答复请求发送装置（例如拓扑搜索部 12 和 /或控制单元侧特定分组发送 /接收部 11），该答复请求发送装置经由未检测连接目的地端口向未检测连接目的地端口的连接目的地发送答复请求（例如请求分组）；以及答复接收装置（例如控制单元侧特定分组发送 /接收部 11），该答复接收装置从成为未检测连接目的地端口的连接目的
的组传送单元接收包括关于该组传送单元的端口的信息的答复；并且该组传送单元包括；传送判定装置（例如交换机侧特定组传送 / 接收部 21 中的用于执行步骤 S21 的部分），该传送判定装置根据接收到的答复请求的内容来判定该答复请求是否要被传送到另外的该组传送单元；传送装置（例如交换机侧特定组传送 / 接收部 21 中的用于执行步骤 S30 的部分），该传送装置在答复请求被判定要被传送到任何其他组传送单元时，根据答复请求的内容将答复请求传送到任何其他组传送单元；以及答复发送装置（例如交换机侧特定组传送 / 接收部 21 和 / 或信息提供部 23），该答复发送装置在答复请求被判定不要被传送到任何其他组传送单元时，返回包括关于该组传送单元的端口的信息的答复，该答复被沿着通过控制单元的路径发送。

【0182】（2）此外，例如，控制单元还包括拓扑信息添加装置（例如拓扑搜索部 12 中的用于执行步骤 S5 的部分），该拓扑信息添加装置向拓扑信息添加与成为未检测连接目的地端口的连接目的地的组传送单元的端口有关并被包括在从该组传送单元的接收的答复中的信息，并且如果该组传送单元的端口之一是连接目的地未被检测的端口并且不是该组传送单元用来接收到答复请求的端口，则拓扑信息添加装置尚向拓扑信息添加指示出该端口未被检测的信息。

【0183】（3）另外，例如，如果没有检测连接目的地端口的端口判定不接收未检测连接目的地端口，则答复请求发送装置停止答复请求的发送。

【0184】（4）另外，例如，答复请求发送装置在答复请求中记述一列表作为答复请求传送路径，该列表是关于从控制单元的端口延伸到未检测连接目的地端口的路径的每个分组传送单元要用来输出答复请求的端口的信息的顺序序列，传送装置沿着答复请求中记述的传送路径传送答复请求，并且答复发送装置基于答复请求中记述的传送路径来辨识去到控制单元的答复发送路径。

【0185】（5）此外，例如，控制单元包括：控制信道确定装置（例如控制信道路径确定部 14），该控制信道确定装置使用拓扑信息来设定作为控制单元和每个分组传送单元之间的路径的控制信道；以及控制信道通知发送装置（例如控制单元侧特定组传送 / 接收部 11），该控制信道通知发送装置发送控制信道通知以将控制信道通知给每个分组传送单元，并且分组传送单元包括控制信道存储装置（例如路径存储部 24），该控制信道存储装置在于接收到控制信道通知时，使用控制信道通知来存储互连控制单元和该组传送单元的控制信道。

【0186】（6）本发明的通信系统包括多个分组传送装置和控制每个分组传送单元的控制单元，其中：控制单元包括；拓扑信息存储部，该拓扑信息存储部存储包括关于控制单元或分组传送单元的端口的互连信息的拓扑信息；未检测连接目的地端口确定部，该未检测连接目的地端口确定部根据拓扑信息在控制单元或分组传送单元的所有端口之中仅确定未检测连接目的地端口；未检测连接目的地端口是连接目的地未被检测的端口；答复请求发送部，该答复请求发送部由未检测连接目的地端口向未检测连接目的地端口的连接目的地发送答复请求；以及答复接收部，该答复接收部从成为未检测连接目的地端口的连接目的地的分组传送单元接收包括关于该分组传送单元的端口的信息的答复，并且分组传送单元包括；传送判定部，该传送判定部根据接收到的答复请求的内容来判定该答复请求是否要被传送到另外的分组传送单元；传送部，该传送部在答复请求被判定要被传送到任
何其他分组传送单元时，根据答复内容将答复请求传送到任何其他分组传送单元；
以及答复发送部，该答复发送部在答复请求被判定不要被传送到任何其他分组传送单元
时，返回包括关于该分组传送单元的端口的信息的答复，该答复被沿着并行到控制单元的路
径发送。

（7）此外，例如，控制单元还包括拓扑信息添加部。该拓扑信息添加部将拓扑信息
添加与控制单元的端口连接时的分组传送单元的端口有关并被包括在
从该分组传送单元接收的答复中的信息，并且如果该分组传送单元的端口之一是连接目的
地未被检测的端口并且不是该分组传送单元用以接收到来自答复请求的端口，则该拓扑信息添加
部还向拓扑信息添加部指示出该端口未被检测的信息。

（8）另外，例如，如果未检测连接目的地的端口确定部判定不存在未检测连接目的地
端口，则答复请求发送部停止答复请求的发送。

（9）另外，例如，答复请求发送部在答复请求中记述一列表作为答复请求传送路
径，该列表是关于从控制单元的端口延伸到未检测连接目的地端口的路径上的每个分组传
送单元要用来输出答复请求的端口的信息的顺序排列，传送部通过答复请求中记述的传送
路径传送答复请求，并且答复发送部基于答复请求中记述的传送路径来辨识去到控制单元
的答复发送路径。

（10）另外，例如，控制单元包括控制信息确定部。该控制信息确定部使用控制信
息来设定作为控制单元和每个分组传送单元之间的路径的控制信息；以及控制信息通知发
送部。该控制信息发送部发送控制信息通知以将控制信息通知给每个分组传送单元，并
且分组传送单元包括控制信息存储部。该控制信息存储部在接收到控制信息通知时，使
用控制信息通知来存储互连控制单元和该分组传送单元的控制信息。

（11）根据本发明的另一方面，控制单元包括控制单元。该控制单元包
括控制信息存储部。该控制信息存储部存储包括关于控制单元或每个分组传送单元的端
口的互连信息的拓扑信息。未检测连接目的地端口确定部，该未检测连接目的地端口确定
部根据拓扑信息在控制单元或每个分组传送单元的所有端口之中仅确定未检测连接目的
地端口，未检测连接目的地端口是连接目的地未被检测的端口。答复请求发送部、该答复请
求发送部经由未检测连接目的地端口并向未检测连接目的地端口的连接目的地发送答复请
求，并且答复接收部，该答复接收部从未检测连接目的地端口的连接目的地的分组传
送单元接收包括关于该分组传送单元的端口的信息的答复。

（12）此外，控制单元还可包括拓扑信息添加部。该拓扑信息添加部向拓扑信息添加
与未检测连接目的地端口的连接目的地的分组传送单元的端口有关并被包括在从
该分组传送单元接收的答复中的信息，并且如果该分组传送单元的端口之一是连接目的地
未被检测的端口并且不是该分组传送单元用以接收到来自答复请求的端口，则拓扑信息添加部
还向拓扑信息添加部指示出该端口未被检测的信息。

（13）本发明的被配置为控制多个分组传送装置控制单元可包括控制
信息存储装置。该控制信息存储装置存储包括关于控制单元或每个分组传送单元的端口的互连信息
的拓扑信息。未检测连接目的地端口确定装置，该未检测连接目的地端口确定装置根据拓
扑信息在控制单元或分组传送单元的所有端口之中仅确定未检测连接目的地端口，未检测
连接目的地端口是连接目的地未被检测的端口。答复请求发送装置，该答复请求发送装置
经由未检测连接目的地端口向未检测连接目的地端口的连接目的地发送回答请求，以及回答接收装置，该回答接收装置从成为未检测连接目的地端口的连接目的地的分组传送单元接收包括关于该分组传送单元的端口的信息的回答。

【0194】（14）根据本发明的另一方面的分组传送单元受控制单元控制，该分组传送单元包括：传送判定部，该传送判定部在从控制单元接收到被寻址到连接目的地未被检测的未检测连接目的地端口的连接目的地的答复请求时，根据该答复请求的内容来判定该答复请求是否要被传送到另外的分组传送单元；发送部，该发送部在答复请求被判定要被传送到任何其他分组传送单元时，根据答复请求的内容将答复请求传送到任何其他分组传送单元；以及答复发送部，该答复发送部在答复请求被判定不要被传送到任何其他分组传送单元时，返回包括关于该分组传送单元的端口的信息的答复，该答复被沿着通向控制单元的路径发送。

【0195】（15）此外，本发明的受控制单元控制的分组传送单元可包括：传送判定装置，该传送判定装置在从控制单元接收到被寻址到连接目的地未被检测的未检测连接目的地端口的连接目的地的答复请求时，根据该答复请求的内容来判定该答复请求是否要被传送到另外的分组传送单元；传送装置，该传送装置在答复请求被判定要被传送到任何其他分组传送单元时，根据答复请求的内容将答复请求传送到任何其他分组传送单元；以及答复发送装置，该答复发送装置在答复请求被判定不要被传送到任何其他分组传送单元时，返回包括关于该分组传送单元的端口的信息的答复，该答复被沿着通向控制单元的路径发送。

【0196】虽然以上已参考示例性实施例描述了本发明，但本发明不限于示例性实施例。在本发明的范围内可对本发明的配置和细节做出本领域的技术人员可理解的各种改变。

【0197】本申请基于2009年12月28日提交的日本专利申请No.2009-298852要求优先权，特此通过引用将该日本专利申请的全部公开内容并入本申请中。

【0198】工业应用

【0199】本发明适用于包括多个分组传送装置和控制每个分组传送单元的控制单元的通信系统。

【0200】标号列表

【0201】10 控制单元

【0202】11 控制单元侧特定分组发送 / 接收部

【0203】12 拓扑搜索部

【0204】13 拓扑存储部

【0205】14 控制信道路径确定部

【0206】15 控制单元侧控制消息隧道传输部

【0207】20a 至 20d, 20A 至 20D 交换机（分组传送单元）

【0208】21 交换机侧特定分组发送 / 接收部

【0209】23 信息提供部

【0210】24 路径存储部

【0211】25 交换机侧控制消息隧道传输部
图 1
图 3
<table>
<thead>
<tr>
<th>字段 (2)</th>
<th>数据长度 (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>类型 (2)</td>
<td>端口长度 (1)</td>
</tr>
<tr>
<td>数据 = [数据类型 ID (6) 字节数据1, 端口#1 状态 (3) 字节数据1, ...]</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>字段 (2)</th>
<th>数据长度 (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>类型 (2)</td>
<td>端口长度 (1)</td>
</tr>
<tr>
<td>数据 = [数据类型 ID (6) 字节数据1, 端口#1 状态 (3) 字节数据1, ...]</td>
<td></td>
</tr>
</tbody>
</table>

图 6

图 7
<table>
<thead>
<tr>
<th>类型 (2)</th>
<th>数据长度 (2)</th>
<th>数据 = [安全通信分组 (控制分组)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>端口长度 (1)</td>
<td>端口 #1(2)</td>
<td></td>
</tr>
<tr>
<td>端口 #2(2)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

图 8
图 11

图 12
图13
图 14
图15

图16
图 17

图 18
图 23
图 27