



(12) 发明专利申请

(10) 申请公布号 CN 115589386 A

(43) 申请公布日 2023. 01. 10

(21) 申请号 202110700475.1

(22) 申请日 2021.06.23

(71) 申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72) 发明人 彭媛媛 赵世振 曹培睿 王景燕  
章春晖

(74) 专利代理机构 深圳市深佳知识产权代理事务  
所(普通合伙) 44285  
专利代理师 聂秀娜

(51) Int.Cl.

H04L 47/12 (2022.01)

H04L 47/10 (2022.01)

H04Q 11/00 (2006.01)

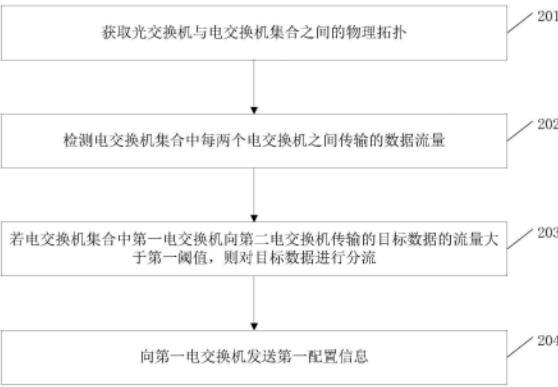
权利要求书3页 说明书9页 附图4页

(54) 发明名称

一种数据分流方法及相关设备

(57) 摘要

本申请实施例提供了一种数据分流方法及相关设备,避免了网络拥塞,提高了网络性能。本申请方法包括:控制器获取光交换机与电交换机集合之间的物理拓扑。控制器会实时检测电交换机集合中每两个电交换机之间传输数据的流量,其中,每两个电交换机之间的数据传输经过光交换机。若第一电交换机向第二电交换机传输的目标数据的流量大于第一阈值,则控制器确定目标数据中第一数据的第一传输路径,并确定目标数据中第二数据的第二传输路径。其中,第一数据传输所采用的跳数不同于第二数据传输所采用的跳数。进而,控制器向第一电交换机发送第一配置信息,以指示第一数据的第一传输路径和第二数据的第二传输路径。



1. 一种数据分流方法,其特征在于,包括:

获取光交换机与电交换机集合之间的物理拓扑,其中,所述光交换机包括多个光口,所述电交换机集合中每个电交换机连接至少一个所述光口;

检测所述电交换机集合中每两个电交换机之间传输数据的流量,其中,每两个电交换机之间传输的数据经过所述光交换机;

若所述电交换机集合中第一电交换机向第二电交换机传输的目标数据的流量大于第一阈值,则确定所述目标数据中第一数据从所述第一电交换机传输至所述第二电交换机的第一传输路径,并确定所述目标数据中第二数据从所述第一电交换机传输至所述第二电交换机的第二传输路径,其中,所述第一数据在所述第一传输路径上经过的电交换机的数量不同于所述第二数据在所述第二传输路径上经过的电交换机的数量;

向所述第一电交换机发送第一配置信息,所述第一配置信息用于指示所述第一数据的所述第一传输路径和所述第二数据的所述第二传输路径。

2. 根据权利要求1所述的方法,其特征在于,所述第一传输路径为所述第一数据从所述第一电交换机传输到所述第二电交换机的传输路径,所述第二传输路径为所述第二数据从所述第一电交换机先传输到第三电交换机再传输到所述第二电交换机的传输路径。

3. 根据权利要求1或2所述的方法,其特征在于,所述第一数据的流量小于或等于所述第一阈值。

4. 根据权利要求1至3中任一项所述的方法,其特征在于,确定所述第一传输路径和所述第二传输路径之前,所述方法还包括:

若所述目标数据上涨的流量大于预设门限,则调节所述第一电交换机与所述第二电交换机之间的物理拓扑,其中,物理拓扑调节后的所述第一电交换机与所述第二电交换机之间的光路数量大于物理拓扑调节前的所述第一电交换机与所述第二电交换机之间的光路数量。

5. 根据权利要求1至4中任一项所述的方法,其特征在于,若所述第一电交换机向所述第二电交换机传输的目标数据的流量大于所述第一阈值,所述方法还包括:

确定所述目标数据中第三数据从所述第一电交换机传输至所述第二电交换机的第三传输路径,其中,所述第三数据在所述第三传输路径上经过的电交换机的数量不同于所述第一数据在所述第一传输路径上经过的电交换机的数量;

向所述第三电交换机发送第二配置信息,所述第二配置信息用于指示所述第三数据的所述第三传输路径。

6. 根据权利要求1至4中任一项所述的方法,其特征在于,所述方法还包括:

若所述第一电交换机向所述第二电交换机传输的目标数据的流量大于所述第二阈值,则确定所述目标数据中第三数据从所述第一电交换机传输至所述第二电交换机的第三传输路径,其中,所述第二阈值大于所述第一阈值,所述第三数据在所述第三传输路径上经过的电交换机的数量不同于所述第一数据在所述第一传输路径上经过的电交换机的数量;

向所述第三电交换机发送第二配置信息,所述第二配置信息用于指示所述第三数据的所述第三传输路径。

7. 根据权利要求6所述的方法,其特征在于,所述第一数据流量等于所述第一阈值,所述第二数据流量等于所述第二阈值与所述第一阈值的差值。

8. 根据权利要求5至7中任一项所述的方法,其特征在于,所述第三传输路径为所述第三数据从所述第一电交换机先传输到第四电交换机再传输到所述第二电交换机的传输路径,或者,所述第三传输路径为所述第三数据从所述第一电交换机依次传输到所述第三电交换机、所述第四电交换机和所述第二电交换机的传输路径。

9. 根据权利要求1至8中任一项所述的方法,其特征在于,每个所述电交换机为汇聚层电交换机。

10. 根据权利要求9所述的方法,其特征在于,每个所述电交换机还与柜顶交换机连接。

11. 一种控制器,其特征在于,包括:处理器、存储器以及收发器,所述处理器、所述存储器以及所述收发器通过线路互相连接;

所述处理器用于:

获取光交换机与电交换机集合之间的物理拓扑,其中,所述光交换机包括多个光口,所述电交换机集合中每个电交换机连接至少一个所述光口;

检测所述电交换机集合中每两个电交换机之间传输数据的流量,其中,每两个电交换机之间传输的数据经过光交换机;

若所述电交换机集合中第一电交换机向第二电交换机传输的目标数据的流量大于第一阈值,则确定所述目标数据中第一数据从所述第一电交换机传输至所述第二电交换机的第一传输路径,并确定所述目标数据中第二数据从所述第一电交换机传输至所述第二电交换机的第二传输路径,其中,所述第一数据在所述第一传输路径上经过的电交换机的数量不同于所述第二数据在所述第二传输路径上经过的电交换机的数量;

所述收发器用于:

向所述第一电交换机发送第一配置信息,所述第一配置信息用于指示所述第一数据的所述第一传输路径和所述第二数据的所述第二传输路径。

12. 根据权利要求11所述的控制器,其特征在于,所述第一传输路径为所述第一数据从所述第一电交换机传输到所述第二电交换机的传输路径,所述第二传输路径为所述第二数据从所述第一电交换机先传输到第三电交换机再传输到所述第二电交换机的传输路径。

13. 根据权利要求11或12所述的控制器,其特征在于,所述第一数据的流量小于或等于所述第一阈值。

14. 根据权利要求11至13中任一项所述的控制器,其特征在于,确定所述第一传输路径和所述第二传输路径之前,所述处理器还用于:

若所述目标数据上涨的流量大于预设门限,则调节所述第一电交换机与所述第二电交换机之间的物理拓扑,其中,物理拓扑调节后的所述第一电交换机与所述第二电交换机之间的光路数量大于物理拓扑调节前的所述第一电交换机与所述第二电交换机之间的光路数量。

15. 根据权利要求11至14中任一项所述的控制器,其特征在于,若所述第一电交换机向所述第二电交换机传输的目标数据的流量大于所述第一阈值,所述处理器还用于:

确定所述目标数据中第三数据从所述第一电交换机传输至所述第二电交换机的第三传输路径,其中,所述第三数据在所述第三传输路径上经过的电交换机的数量不同于所述第一数据在所述第一传输路径上经过的电交换机的数量;

所述收发器还用于:

向所述第三电交换机发送第二配置信息,所述第二配置信息用于指示所述第三数据的所述第三传输路径。

16. 根据权利要求11至14中任一项所述的控制器,其特征在于,所述处理器还用于:

若所述第一电交换机向所述第二电交换机传输的目标数据的流量大于所述第二阈值,则确定所述目标数据中第三数据从所述第一电交换机传输至所述第二电交换机的第三传输路径,其中,所述第二阈值大于所述第一阈值,所述第三数据在所述第三传输路径上经过的电交换机的数量不同于所述第一数据在所述第一传输路径上经过的电交换机的数量;

所述收发器还用于:

向所述第三电交换机发送第二配置信息,所述第二配置信息用于指示所述第三数据的所述第三传输路径。

17. 根据权利要求16所述的控制器,其特征在于,所述第一数据流量等于所述第一阈值,所述第二数据流量等于所述第二阈值与所述第一阈值的差值。

18. 根据权利要求15至17中任一项所述的控制器,其特征在于,所述第三传输路径为所述第三数据从所述第一电交换机先传输到第四电交换机再传输到所述第二电交换机的传输路径,或者,所述第三传输路径为所述第三数据从所述第一电交换机依次传输到所述第三电交换机、所述第四电交换机和所述第二电交换机的传输路径。

19. 根据权利要求11至18中任一项所述的控制器,其特征在于,每个所述电交换机为汇聚层电交换机。

20. 根据权利要求11至19中任一项所述的控制器,其特征在于,每个所述电交换机还与柜顶交换机连接。

21. 一种通信系统,其特征在于,包括:控制器、光交换机和电交换机集合,其中,所述光交换机包括多个光口,所述电交换机集合中每个电交换机连接至少一个所述光口,每两个电交换机之间传输的数据经过光交换机,所述控制器用于执行权利要求1至10中任一项所述的方法。

## 一种数据分流方法及相关设备

### 技术领域

[0001] 本申请涉及通信领域,尤其涉及一种数据分流方法及相关设备。

### 背景技术

[0002] 云时代的到来促进了各种互联网应用蓬勃发展,极大地丰富了人们的生活。随之而来的是网络流量指数级增长,这对其背后的数据中心的带宽、时延、功耗、成本,以及可靠性、拓展性、灵活性等方面提出了更高的要求。

[0003] 为了获得高带宽、更低的延迟和更低的功耗,会考虑将光交换机应用到数据中心。不同的电交换机之间通过光交换机来实现通信。其中,在光交换机中,只有相互导通光路的端口之间能实现通信。为了根据实际需求建立光路,光交换机需要具有重配置能力,以实现动态切换拓扑。在实际应用中,电交换机之间的流量变化不均匀,虽然采用光交换机可以实现高频率的动态拓扑切换,但是遇到流量突发的情形,仍会出现网络拥塞,从而导致网络性能下降。

### 发明内容

[0004] 本申请实施例提供了一种数据分流方法及相关设备,在遇到流量突发时,可以基于设定的阈值对数据分流,通过多条传输路径来进行负载分担,避免了网络拥塞,提高了网络性能。

[0005] 第一方面,本申请提供了一种数据分流方法。该方法由控制器执行,具体包括多个步骤。控制器获取光交换机与电交换机集合之间的物理拓扑。其中,光交换机包括多个光口,电交换机集合中每个电交换机连接至少一个光口,每两个电交换机之间传输的数据经过光交换机。控制器会实时检测电交换机集合中每两个电交换机之间传输数据的流量。若电交换机集合中第一电交换机向第二电交换机传输的目标数据的流量大于第一阈值,则需要对目标数据进行分流。具体地,控制器确定目标数据中第一数据从第一电交换机传输至第二电交换机的第一传输路径,并确定目标数据中第二数据从第一电交换机传输至第二电交换机的第二传输路径。其中,第一数据在第一传输路径上经过的电交换机的数量不同于第二数据在第二传输路径上经过的电交换机的数量。即第一数据传输所采用的跳数不同于第二数据传输所采用的跳数。进而,控制器向第一电交换机发送第一配置信息,以指示第一数据的第一传输路径和第二数据的第二传输路径。

[0006] 在该实施方式中,如果两个电交换机之间传输数据的流量较大,控制器可以对该数据进行分流,并确定分流后多路数据各自的传输路径。其中,不同流向的数据选取的跳数也不同,例如,一部分数据走1跳路由,另一部分数据走2跳路由。进而,控制器基于分流后的传输路径对电交换机进行配置。通过上述方式,在遇到流量突发时,可以基于设定的阈值对数据分流,通过多条传输路径来进行负载分担,避免了网络拥塞,提高了网络性能。

[0007] 在一些可能的实施方式中,第一传输路径为第一数据从第一电交换机传输到第二电交换机的传输路径,第二传输路径为第二数据从第一电交换机先传输到第三电交换机再

传输到第二电交换机的传输路径。在该实施方式中,分流后的两路数据分别采用1跳和2跳的传输方式。应理解,如果每两个电交换机之间都采用1跳的传输方式,抗流量突发的能力较弱,并且难以实现负载分担。如果每两个电交换机之间都采用多跳的传输方式,会增加整个网络的负担。因此,通过上述方式即具备抗流量突发的能力,还可以兼顾网络效率。

[0008] 在一些可能的实施方式中,第一数据的流量小于或等于第一阈值,即采用1跳传输路径的数据流量需要小于或等于分流阈值,剩余的数据采用多跳传输。通过这种方式可以控制最短路径传输的数据流量,优先避免最短路径传输的数据发生拥塞。

[0009] 在一些可能的实施方式中,确定第一传输路径和第二传输路径之前,方法还包括:若目标数据上涨的流量大于预设门限,则调节第一电交换机与第二电交换机之间的物理拓扑。其中,物理拓扑调节后的第一电交换机与第二电交换机之间的光路数量大于物理拓扑调节前的第一电交换机与第二电交换机之间的光路数量。通过上述方式,若两个电交换机之间流量变化较大,可以通过光交换机重新配置两个电交换机之间的光路,以适应流量变化。

[0010] 在一些可能的实施方式中,若第一电交换机向第二电交换机传输的目标数据的流量大于第一阈值,方法还包括:确定目标数据中第三数据从第一电交换机传输至第二电交换机的第三传输路径。其中,第三数据在第三传输路径上经过的电交换机的数量不同于第一数据在第一传输路径上经过的电交换机的数量。向第三电交换机发送第二配置信息,第二配置信息用于指示第三数据的第三传输路径。在该实施方式中,这种基于阈值的分流方法不仅限于分成两路不同的传输路径,还可以支持更多路的分流,提高了本方案的扩展性。

[0011] 在一些可能的实施方式中,方法还包括:若第一电交换机向第二电交换机传输的目标数据的流量大于第二阈值,则确定目标数据中第三数据从第一电交换机传输至第二电交换机的第三传输路径。其中,第二阈值大于第一阈值,第三数据在第三传输路径上经过的电交换机的数量不同于第一数据在第一传输路径上经过的电交换机的数量。向第三电交换机发送第二配置信息,第二配置信息用于指示第三数据的第三传输路径。在该实施方式中,还可以设置多个分流阈值,以应用于更多路分流的场景中,分流的方式更灵活。

[0012] 在一些可能的实施方式中,第一数据流量等于第一阈值,第二数据流量等于第二阈值与第一阈值的差值。通过多个分流阈值来限制各路数据的流量,可以更合理地分配分流后各路数据的流量。

[0013] 在一些可能的实施方式中,第三传输路径为第三数据从第一电交换机先传输到第四电交换机再传输到第二电交换机的传输路径。或者,第三传输路径为第三数据从第一电交换机依次传输到第三电交换机、第四电交换机和第二电交换机的传输路径。在该实施方式中,分流后除了采用1跳传输的那一路数据外,其他各路数据可以采用多跳传输,具体的跳数可以是2跳,也可以是3跳或更多跳,进一步提高了本方案的扩展性。

[0014] 在一些可能的实施方式中,每个电交换机为汇聚层电交换机。

[0015] 在一些可能的实施方式中,每个电交换机还与柜顶交换机连接。

[0016] 第二方面,本申请提供了一种控制器,包括处理器、存储器以及收发器。其中,该处理器、该存储器以及该光收发器通过线路互相连接。处理器用于:获取光交换机与电交换机集合之间的物理拓扑。其中,光交换机包括多个光口,电交换机集合中每个电交换机连接至少一个光口。检测电交换机集合中每两个电交换机之间传输数据的流量,其中,每两个电交

交换机之间传输的数据经过光交换机。若电交换机集合中第一电交换机向第二电交换机传输的目标数据的流量大于第一阈值,则确定目标数据中第一数据从第一电交换机传输至第二电交换机的第一传输路径,并确定目标数据中第二数据从第一电交换机传输至第二电交换机的第二传输路径。其中,第一数据在第一传输路径上经过的电交换机的数量不同于第二数据在第二传输路径上经过的电交换机的数量。收发器用于:向第一电交换机发送第一配置信息,第一配置信息用于指示第一数据的第一传输路径和第二数据的第二传输路径。

[0017] 在一些可能的实施方式中,第一传输路径为第一数据从第一电交换机传输到第二电交换机的传输路径,第二传输路径为第二数据从第一电交换机先传输到第三电交换机再传输到第二电交换机的传输路径。

[0018] 在一些可能的实施方式中,第一数据的流量小于或等于第一阈值。

[0019] 在一些可能的实施方式中,在第一传输路径和第二传输路径之前,处理器还用于:若目标数据上涨的流量大于预设门限,则调节第一电交换机与第二电交换机之间的物理拓扑,其中,物理拓扑调节后的第一电交换机与第二电交换机之间的光路数量大于物理拓扑调节前的第一电交换机与第二电交换机之间的光路数量。

[0020] 在一些可能的实施方式中,若第一电交换机向第二电交换机传输的目标数据的流量大于第一阈值,处理器还用于:确定目标数据中第三数据从第一电交换机传输至第二电交换机的第三传输路径。其中,第三数据在第三传输路径上经过的电交换机的数量不同于第一数据在第一传输路径上经过的电交换机的数量。收发器还用于:向第三电交换机发送第二配置信息,第二配置信息用于指示第三数据的第三传输路径。

[0021] 在一些可能的实施方式中,处理器还用于:若第一电交换机向第二电交换机传输的目标数据的流量大于第二阈值,则确定目标数据中第三数据从第一电交换机传输至第二电交换机的第三传输路径。其中,第二阈值大于第一阈值,第三数据在第三传输路径上经过的电交换机的数量不同于第一数据在第一传输路径上经过的电交换机的数量。收发器还用于:向第三电交换机发送第二配置信息,第二配置信息用于指示第三数据的第三传输路径。

[0022] 在一些可能的实施方式中,第一数据流量等于第一阈值,第二数据流量等于第二阈值与第一阈值的差值。

[0023] 在一些可能的实施方式中,第三传输路径为第三数据从第一电交换机先传输到第四电交换机再传输到第二电交换机的传输路径。或者,第三传输路径为第三数据从第一电交换机依次传输到第三电交换机、第四电交换机和第二电交换机的传输路径。

[0024] 在一些可能的实施方式中,每个电交换机为汇聚层电交换机。

[0025] 在一些可能的实施方式中,每个电交换机还与柜顶交换机连接。

[0026] 第三方面,本申请提供了一种通信系统,包括:控制器、光交换机和电交换机集合。其中,光交换机包括多个光口,电交换机集合中每个电交换机连接至少一个光口,每两个电交换机之间传输的数据经过光交换机。控制器用于执行上述第一方面任意一种方法的部分或全部步骤。

[0027] 本申请实施例中,每两个电交换机之间经过光交换机来实现数据传输。如果两个电交换机之间传输数据的流量较大,控制器可以对该数据进行分流,并确定分流后多路数据各自的传输路径。其中,不同流向的数据选取的跳数也不同,例如,一部分数据走1跳路由,另一部分数据走2跳路由。进而,控制器基于分流后的传输路径对电交换机进行配置。通

过上述方式,在遇到流量突发时,可以基于设定的阈值对数据分流,通过多条传输路径来进行负载分担,避免了网络拥塞,提高了网络性能。另外,本申请提供的方法不用依赖快速光交换技术,只有在流量变化较大的情况下才需要调制拓扑,也可以避免通过光交换机频繁重配置带来的风险。并且,采用本申请提供的路由算法能够保证电交换机之间的流量只在最必要的情况下才被分流,否则尽可能走1跳路径。其优势在于:一方面,保证了网络效率,可以使整个网络流量的平均跳数尽可能小,降低时延。另一方面,兼顾了应对突发的处理能力,当有突发流量来临时,可以平滑分配到多跳路由,从其它路径中转到达目的地。

## 附图说明

- [0028] 图1(a)为本申请实施例中的一种网络结构示意图;
- [0029] 图1(b)为本申请实施例中的另一种网络结构示意图;
- [0030] 图2为本申请实施例中数据分流方法的一个实施例示意图;
- [0031] 图3(a)为光交换机重配置前的物理拓扑示意图;
- [0032] 图3(b)为光交换机重配置后的物理拓扑示意图;
- [0033] 图4为本申请实施例中数据分流的一个实施例示意图;
- [0034] 图5为本申请实施例中控制器的一种结构示意图;
- [0035] 图6为本申请实施例中通信系统的一种结构示意图。

## 具体实施方式

[0036] 本申请实施例提供了一种数据分流方法及相关设备,在遇到流量突发时,可以基于设定的阈值对数据分流,通过多条传输路径来进行负载分担,避免了网络拥塞,提高了网络性能。

[0037] 需要说明的是,本申请说明书和权利要求书及上述附图中的术语“第一”、“第二”和“第三”等用于区别类似的对象,而非限定特定的顺序或先后次序。应理解,上述术语在适当情况下可以互换,以便在本申请描述的实施例能够以除了在本申请描述的内容以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含。例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0038] 图1(a)为本申请实施例中的一种网络结构示意图。如图1(a)所示,本申请主要应用于电交换机与光交换机交互的场景中,每两个电交换机之间的数据通信通过光交换机来完成。其中,电交换机包括如图1(a)所示的电交换机1、电交换机2、电交换机3和电交换机4。光交换机包括如图1(a)所示的光交换机A、光交换机B和光交换机C。具体地,每两个电交换机之间通过光交换机建立连接。电交换机上的物理接口与光交换机上对应的一个光口通过光纤连接。

[0039] 图1(b)为本申请实施例中的另一种网络结构示意图。如图1(b)所示,在一些可能的实施方式中,本申请具体可以应用在全光数据中心的网络场景中。具体地,该全光数据中心包括多个转发点(Point of Delivery, PoD)集群,并且核心层全部采用光交换机。每个PoD集群包括多个柜顶(Top of Rack, ToR)交换机,并且,每个PoD集群还包括多个汇聚层的



电交换机,汇聚层的电交换机与ToR均匀互连。汇聚层的各电交换机之间可以通过核心层的光交换机实现相互通信。应理解,核心层的光交换机提供了重配置能力,可以让不同PoD集群间导通的链路数灵活可调,PoD集群间实际流量需求越多,光交换机为其分配的连通链路数也越多,从而可以在拓扑层面提供应对不均匀流量的能力。但是,若遇到流量突发的情形,只从拓扑层面来应对仍有可能会出现网络拥塞,从而导致网络性能下降。

[0040] 为此,本申请提供了一种数据分流方法,在遇到流量突发时,可以基于设定的阈值对数据分流,通过多条传输路径来进行负载分担,避免了网络拥塞,提高了网络性能。下面进行详细介绍。应理解,上述图1(a)和图1(b)只展示了转发面的网络结构,在实际应用中,无论是拓扑调整还是路由调整都是由控制面操作的。本申请提供的数据分流方法具体是由控制器对各交换机进行配置来实现的。

[0041] 图2为本申请实施例中数据分流方法的一个实施例示意图。在示例中,数据分流方法包括如下步骤。

[0042] 201、获取光交换机与电交换机集合之间的物理拓扑。

[0043] 本实施例中,控制器可以通过预配置或人工输入的方式获取光交换机与电交换机集合之间的物理拓扑。该物理拓扑具体包括电交换机上接口与光交换机上光口之间的物理连接关系。其中,每两个电交换机之间通过光交换机连接,即每两个电交换机之间传输的数据通过光交换机。具体地,每个光交换机包括多个用于光交换的光口,每个电交换机的接口与至少一个光口连接。应理解,本申请中光交换机的数量通常为多个,具体数量此处不做限定。另外,本申请不限定上述物理拓扑的具体应用场景,只要是一层电交换机和一层光交换机之间建立的连接即可,例如上述物理拓扑具体可应用于图1(b)所示的数据中心。

[0044] 202、检测电交换机集合中每两个电交换机之间传输的数据流量。

[0045] 控制器会实时监控每两个电交换机之间的数据流量并进行记录,若控制器检测到某两个电交换机之间的流量发生了较大变化,那么控制器就需要对这两个电交换机之间的物理拓扑进行调整。具体地,控制器可以控制光交换机重新配置这两个电交换机之间的光路。例如,第一电交换机与第二电交换机之间上涨的流量大于预设门限,则调节第一电交换机与第二电交换机之间的物理拓扑,以增加第一电交换机与第二电交换机之间的光路数量。需要说明的是,两个电交换机之间的光路数量=两个电交换机之间的流量/每条光路所能提供的带宽。为了便于理解,下面通过一个具体的示例进行介绍。

[0046] 图3(a)为光交换机重配置前的物理拓扑示意图。如图3(a)所示为电交换机1-电交换机4与光交换机之间的初始物理拓扑。其中,电交换机1与电交换机2之间以及电交换机3与电交换机4之间的光路数量较多。之后,控制器通过实时监控发现4个电交换机之间的流量矩阵发生较大变化,需要进行拓扑调整。图3(b)为光交换机重配置后的物理拓扑示意图。如图3(b)所示,进行拓扑调整后,电交换机1与电交换机4之间以及电交换机2与电交换机3之间的光路数量较多,以适应新的流量需求。

[0047] 在一种可能的实施方式中,本申请可以采用链路聚合(eth-trunk)的方式进行拓扑调整。应理解,采用链路聚合技术可以在不进行硬件升级的条件下,通过将多个物理接口捆绑为一个逻辑接口,这个逻辑接口称之为链路聚合接口或Eth-Trunk接口,从而达到增加链路带宽的目的。在实现增大带宽目的的同时,链路聚合采用备份链路的机制,可以有效的提高设备之间链路的可靠性。具体地,控制器可以根据流量的变化在已经配置好的Eth-

Trunk接口中删除部分物理接口,从而删除某两个电交换机之间的部分光路。进而,控制器在另外两个电交换机之间建立新的光路,并配置新的Eth-Trunk接口,从而实现拓扑调整。例如图3(a)所示,电交换机1与电交换机2之间相连的两组接口可以配置为一个Eth-Trunk接口,电交换机3与电交换机4之间相连的两组接口可以配置为一个Eth-Trunk接口。经过拓扑调整后,如图3(b)所示,电交换机1与电交换机4之间相连的两组接口可以配置为一个Eth-Trunk接口,电交换机2与电交换机3之间相连的两组接口可以配置为一个Eth-Trunk接口。

[0048] 203、若电交换机集合中第一电交换机向第二电交换机传输的目标数据的流量大于第一阈值,则对目标数据进行分流。

[0049] 本实施例中,除了基于流量变化进行拓扑调整外,还可以拓扑调整的基础上进一步对两个电交换机之间传输的数据进行分流以应对流量突发。其中,可以预先设置一个分流阈值,例如80Mbps。如果两个电交换机之间的流量大于分流阈值,则控制器对数据进行分流。应理解,具体的分流数量可以根据实际需求进行选择,本申请不做限定。并且,为了更好地进行负载分担,分流后不同路径所采用的跳数不同,下面结合一些具体的示例进行介绍。

[0050] 在一种可能的实施方式中,基于一个分流阈值对数据进行分流传输。具体地,若电交换机集合中第一电交换机向第二电交换机传输的目标数据的流量大于第一阈值,则控制器确定目标数据中第一数据从第一电交换机传输至第二电交换机的第一传输路径,并确定目标数据中第二数据从第一电交换机传输至第二电交换机的第二传输路径。其中,第一数据在第一传输路径上经过的电交换机的数量不同于第二数据在第二传输路径上经过的电交换机的数量。也就是说,第一数据经过的跳数不同于第二数据经过的跳数。

[0051] 图4为本申请实施例中数据分流的一个实施例示意图。如图4所示,电交换机4向交换机1传输的数据可以进行分流,具体可以分为2条传输路径。第一路从电交换机4直达电交换机1,其中只经过一次光交换,第一路就视为1跳的传输路径。第二路从电交换机4先传输到电交换机3,经过电交换机3中转后再传输到电交换机1,其中经过了两次光交换,第二路就视为2跳的传输路径。应理解,在上述2条传输路径的基础上,还可以进一步分出第3条等更多传输路径,具体此处不做限定。例如,第三路可以从电交换机4先传输到电交换机2,经过电交换机2中转后再传输到电交换机1,这种也视为2跳的传输路径。需要说明的是,在实际应用中,应当优先避免最短路径传输的数据发生拥塞,因此需要控制最短路径传输的数据流量。在一种可能的实施方式中,采用1跳传输路径的数据流量需要小于或等于分流阈值,剩余的数据采用多跳传输。

[0052] 需要说明的是,数据分流后所采用的跳数不仅限于1跳和2跳,在实际应用中,还可以采用3跳或者更多跳,具体此处不做限定。仍以上述图4为例,数据分流后的其中一路可以从电交换机4先传输到电交换机3,再从电交换机3传输到电交换机2,最后从电交换机2再传输到电交换机1,其中经过了三次光交换,即视为3跳的传输路径。应理解,如果每两个电交换机之间都采用1跳的传输方式,抗流量突发的能力较弱,并且难以实现负载分担。如果每两个电交换机之间都采用多跳的传输方式,会增加整个网络的负担。因此,具体可以根据实际需求来选择跳数,即要具备抗流量突发的能力,还要兼顾网络效率。

[0053] 在另一种可能的实施方式中,还可以基于多个分流阈值来对数据进行分流传输。例如,设置有两个分流阈值,分别为第一阈值和第二阈值,第二阈值大于第一阈值。如果第

一电交换机向第二电交换机传输的数据流量大于第一阈值且小于或等于第二阈值,则将数据分成两个不同的传输路径,具体方式与上述基于一个分流阈值进行分流的实施方式类似,此处不再赘述。如果第一电交换机向第二电交换机传输的数据流量大于第二阈值,则考虑将数据分成三个不同的传输路径。具体地,第一路采用1跳的传输方式,并且第一路的数据流量等于第一阈值,优先保证最短传输路径有足够的流量且不会拥塞。第二路采用多跳的传输方式,第二路数据的流量等于第二阈值与第一阈值的差值。除了第一路和第二路之外剩余的数据走第三路,第三路同样采用多跳的传输方式。其中,本申请不限定第二路和第三路所采用的具体跳数。应理解,在实际应用中,还可以设置更多个分流阈值,具体此处不做限定。通过多个分流阈值来限制各路数据的流量,可以更合理地分配分流后各路数据的流量。

[0054] 204、向第一电交换机发送第一配置信息。

[0055] 在控制器确定了分流后各路数据的传输路径后,控制器可以向第一电交换机发送第一配置信息,以指示用于传输第一数据的第一传输路径和用于传输第二数据的第二传输路径。在数据流到来时,第一电交换机即可根据第一配置信息对数据流进行分流转发。

[0056] 下面结合具体的应用场景对阈值分流的实现方式进行介绍。以上述图1(b)应用与数据中心的场景为例,对于支持多级流表的数据中心,汇聚层电交换机首先需要对不同流向的各路数据进行标记。具体地,可以修改数据流的差分服务代码点(Differentiated Services Code Point,DSCP)值,使得不同流向数据流的DSCP值不同。接下来,汇聚层电交换机将标记后不同流向的数据流按分流比例分配对应的端口,以使得不同标记的数据流按照不同的传输路径传输至目的地址。例如,设定的分流阈值是80Mbps,数据流默认的DSCP值为2,当有100Mbps的数据流量到来时,可以将其中20Mbps数据流的DSCP值修改为4。之后,在多级流表的作用下,DSCP值修改为4的数据流会采用2跳的传输方式,未被修改DSCP值的数据流采用1跳的传输方式。

[0057] 需要说明的是,对于不支持多级流表的数据中心,还可以通过ToR交换机对不同流向的各路数据进行标记。之后,再由汇聚层交换机识别各路数据的标记,进而将标记后不同流向的数据流按分流比例分配对应的端口,以使得不同标记的数据流按照不同的传输路径传输至目的地址。

[0058] 本申请实施例中,每两个电交换机之间经过光交换机来实现数据传输。如果两个电交换机之间传输数据的流量较大,控制器可以对该数据进行分流,并确定分流后多路数据各自的传输路径。其中,不同流向的数据选取的跳数也不同,例如,一部分数据走1跳路由,另一部分数据走2跳路由。进而,控制器基于分流后的传输路径对电交换机进行配置。通过上述方式,在遇到流量突发时,可以基于设定的阈值对数据分流,通过多条传输路径来进行负载分担,避免了网络拥塞,提高了网络性能。另外,本申请提供的方法不用依赖快速光交换技术,只有在流量变化较大的情况下才需要调制拓扑,也可以避免通过光交换机频繁重配置带来的风险。

[0059] 并且,采用本申请提供的路由算法能够保证电交换机之间的流量只在最必要的情况下才被分流,否则尽可能走1跳路径。其优势在于:一方面,保证了网络效率,可以使整个网络流量的平均跳数尽可能小,降低时延。另一方面,兼顾了应对突发的处理能力,当有突发流量来临时,可以平滑分配到多跳路由,从其它路径中转到达目的地。

[0060] 上面对本申请提供的数据分流方法进行了介绍,下面对本申请提供的控制器进行介绍。

[0061] 图5为本申请实施例中控制器的一种结构示意图。如图5所示,该控制器包括处理器501、存储器502和收发器503。该处理器501、存储器502和收发器503通过线路相互连接。其中,存储器502用于存储程序指令和数据。需要说明的是,收发器503用于执行上述图2所示实施例中信息的收发操作。处理器501用于执行上述图2所示实施例中除了信息收发外的其他操作。

[0062] 需要说明的是,上述图5中所示的处理器可以采用通用的中央处理器(Central Processing Unit,CPU),微处理器,应用专用集成电路ASIC,或者至少一个集成电路,用于执行相关程序,以实现本申请实施例所提供的技术方案。上述图5中所示的存储器可以存储操作系统和其他应用程序。在通过软件或者固件来实现本申请实施例提供的技术方案时,用于实现本申请实施例提供的技术方案的程序代码保存在存储器中,并由处理器来执行。在一实施例中,处理器内部可以包括存储器。在另一实施例中,处理器和存储器是两个独立的结构。

[0063] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统,装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0064] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,随机接入存储器等。具体地,例如:上述处理单元或处理器可以是中央处理器,通用处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件或者其任意组合。上述的这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本申请的范围。

[0065] 当使用软件实现时,上述实施例描述的方法步骤可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时,全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一个计算机可读存储介质传输,例如,所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线(DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质,(例如,软盘、硬盘、磁带)、光介质(例如,DVD)、或者半导体介质(例如固态硬盘Solid State Disk(SSD))等。

[0066] 下面对本申请提供的一种通信系统进行介绍。

[0067] 图6为本申请实施例中通信系统的一种结构示意图。如图6所示,通信系统包括控制器10、光交换机20和电交换机集合,电交换机集合包括电交换机301、电交换机302和电交换机303。其中,光交换机包括多个光口,电交换机集合中每个电交换机连接至少一个光口,

每两个电交换机之间传输的数据经过光交换机。具体地,控制器10可以控制光交换机进行光路配置,以调整每两个电交换机之间的物理拓扑。并且,控制器10还可以控制每个电交换机的转发路径,以实现路由调整。需要说明的是,控制器10用于执行上述图2所示实施例的数据分流方法,具体此处不再赘述。

[0068] 应理解,本申请不限定上述通信系统中光交换机的数量和电交换机集合中电交换机的数量。另外,在一些可能的实施方式中,通信系统中也可以包括不止一个控制器。例如,一个控制器用于控制光交换机,另一个控制器用于控制电交换机,具体此处不做限定。

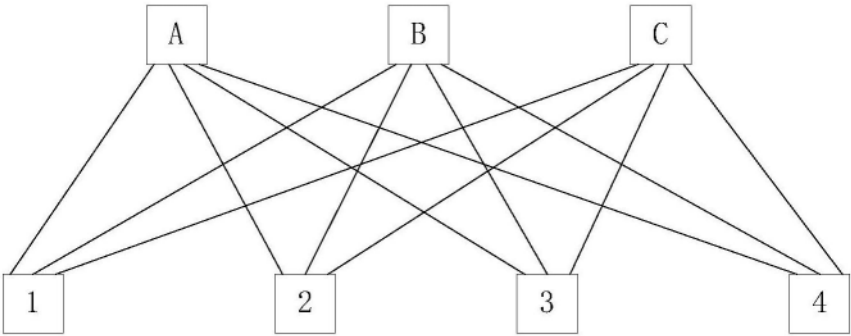


图1 (a)

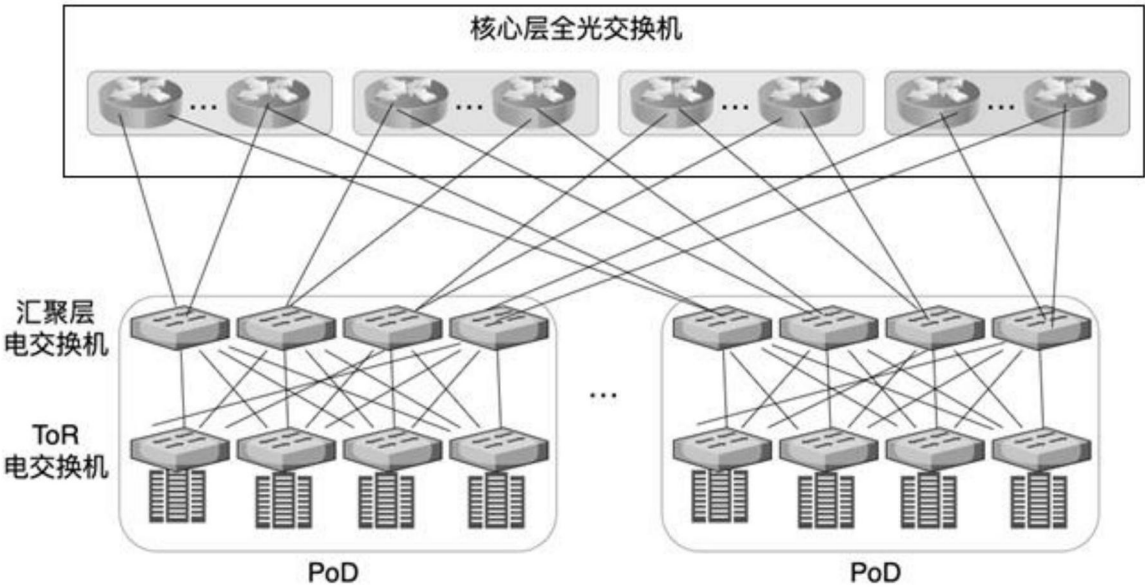


图1 (b)

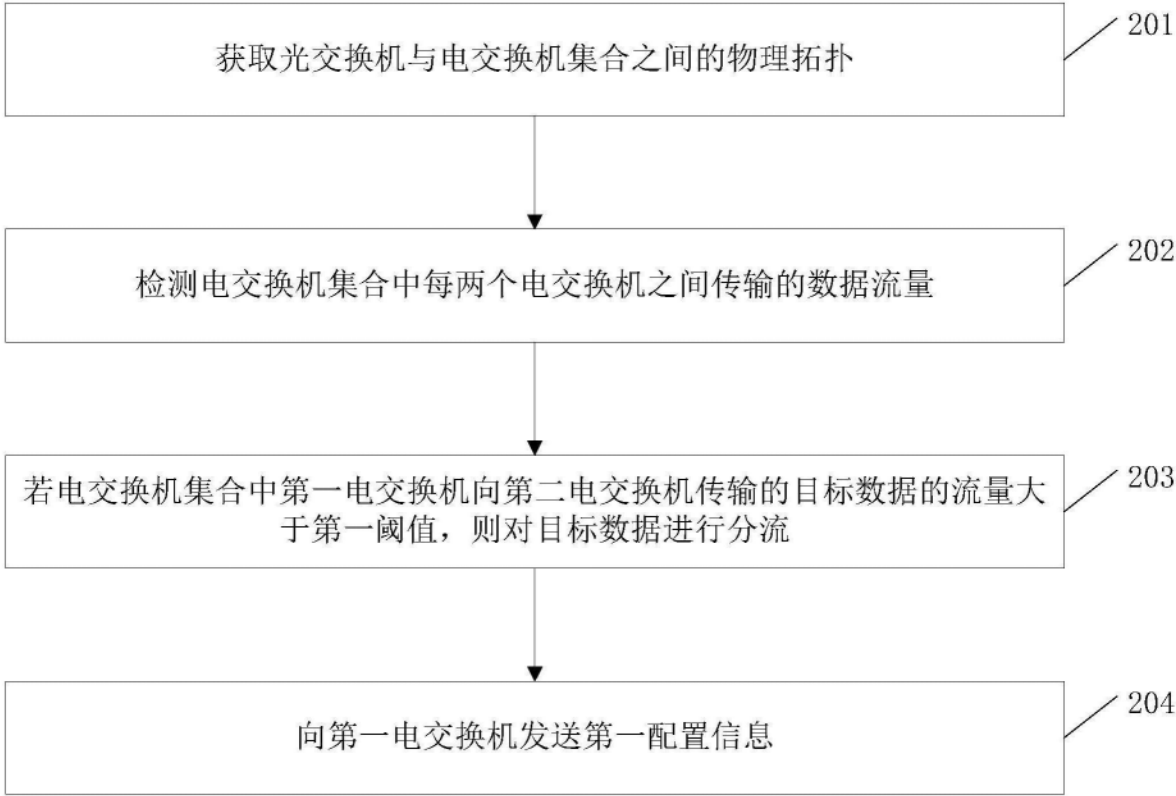


图2

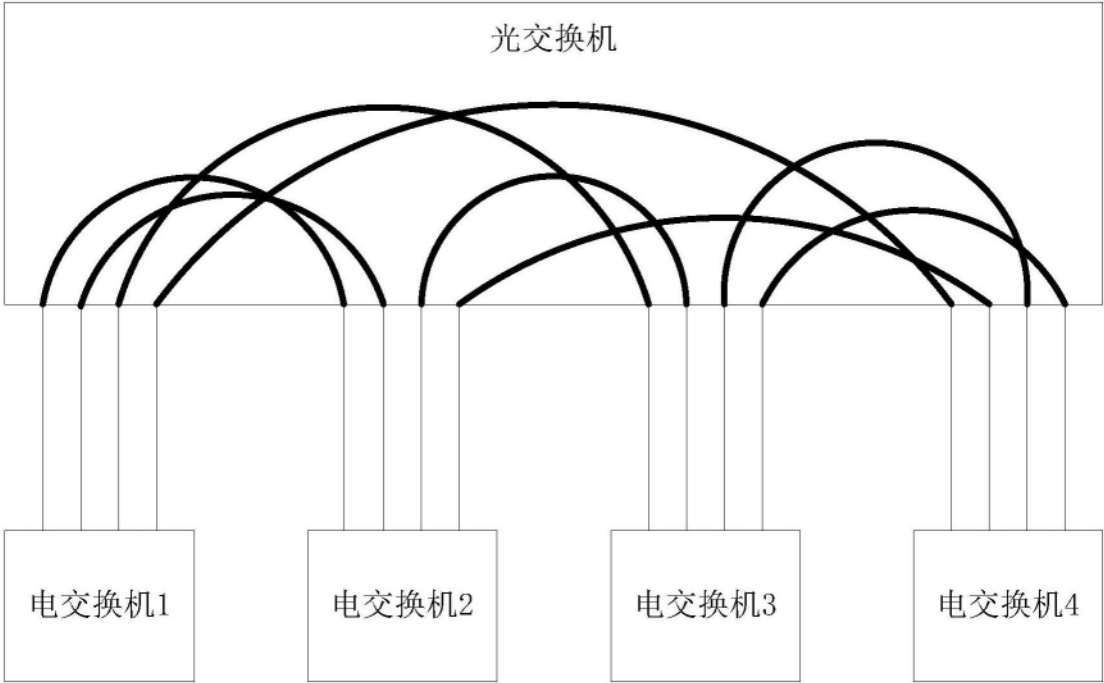


图3 (a)

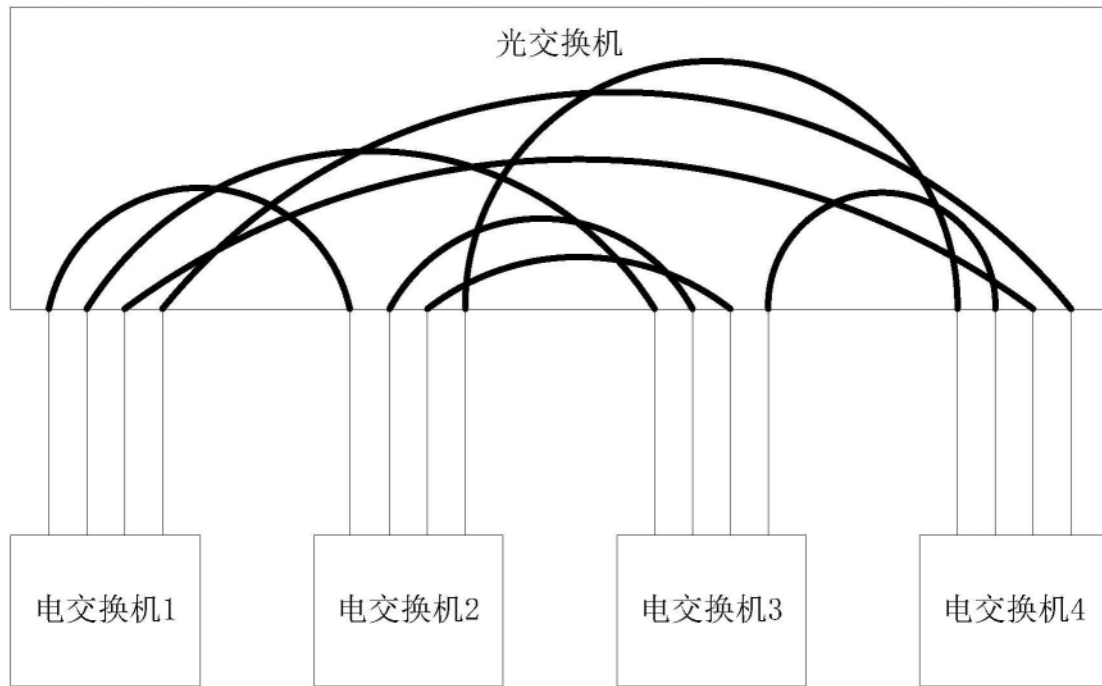


图3 (b)

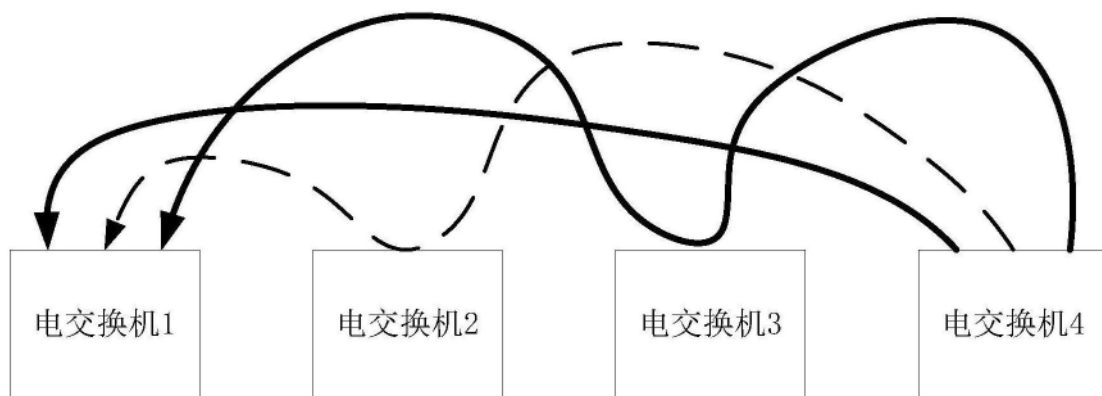


图4



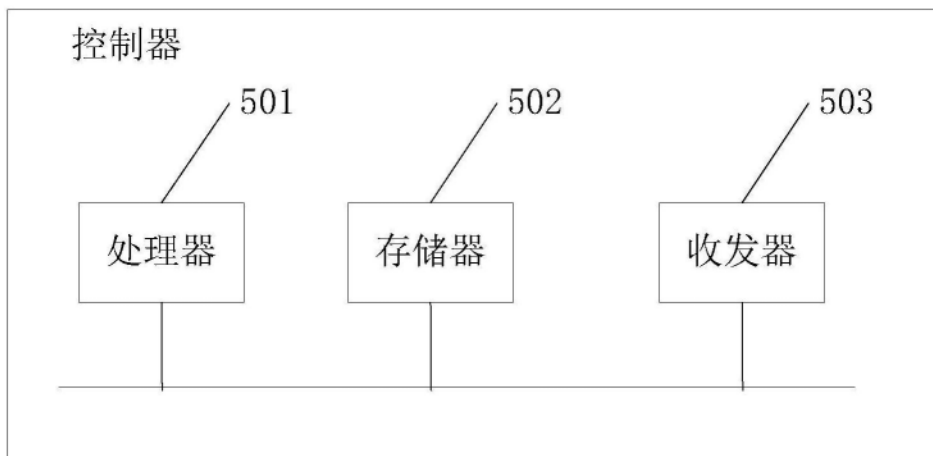


图5

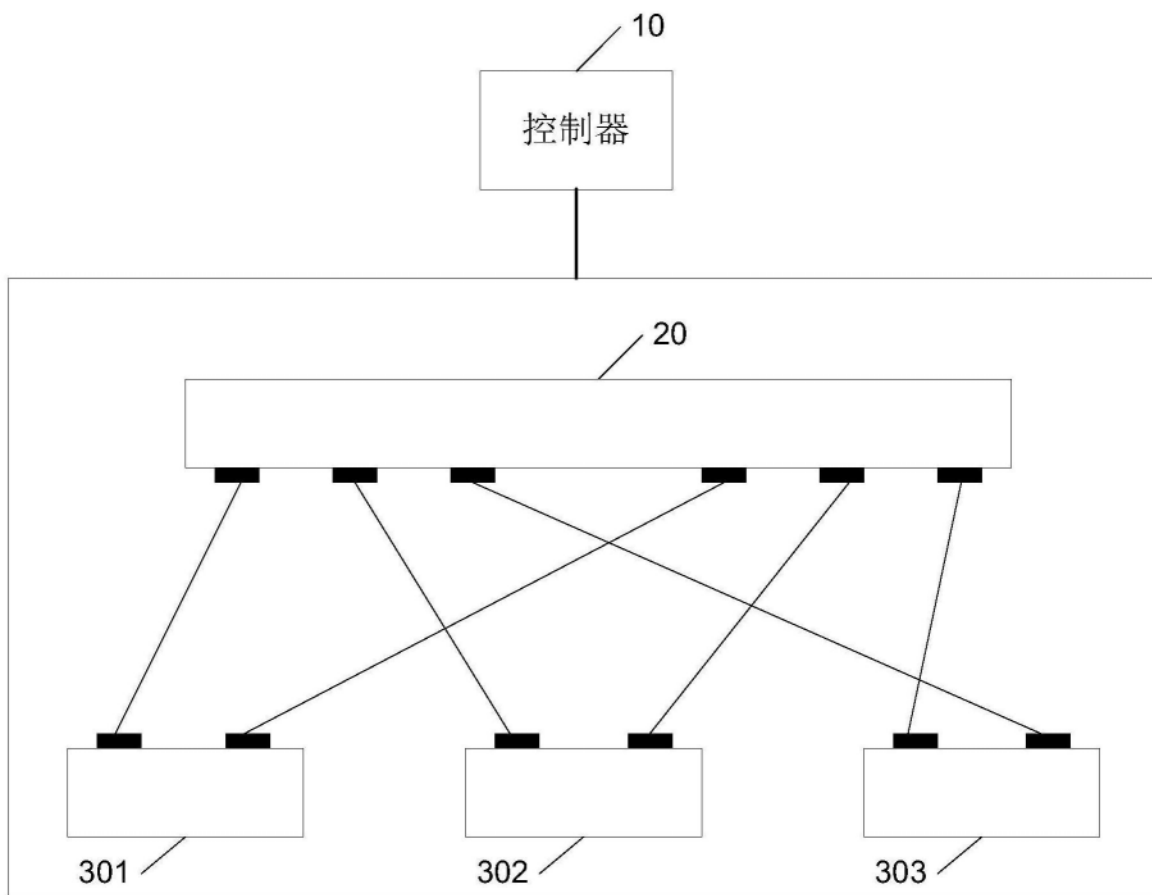


图6