



(12) 发明专利申请

(10) 申请公布号 CN 104094565 A

(43) 申请公布日 2014. 10. 08

(21) 申请号 201280068918. 9

代理人 李兰 孙志湧

(22) 申请日 2012. 10. 31

(51) Int. Cl.

(30) 优先权数据

H04L 12/803(2013. 01)

2012-020695 2012. 02. 02 JP

(85) PCT国际申请进入国家阶段日

2014. 08. 01

(86) PCT国际申请的申请数据

PCT/JP2012/006998 2012. 10. 31

(87) PCT国际申请的公布数据

W02013/114490 JA 2013. 08. 08

(71) 申请人 日本电气株式会社

地址 日本东京

(72) 发明人 河野正太郎

(74) 专利代理机构 中原信达知识产权代理有限

责任公司 11219

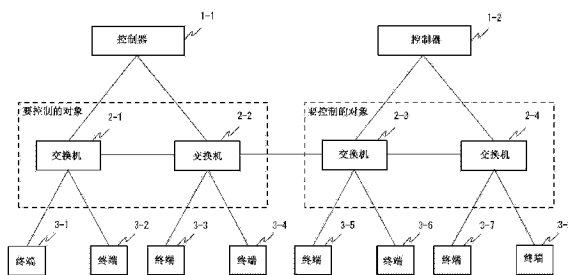
权利要求书3页 说明书12页 附图15页

(54) 发明名称

控制器、用于分配负载的方法、存储程序的非瞬时计算机可读介质、计算机系统和控制设备

(57) 摘要

开放流网络包括多个控制器 (1)。每一控制器 (1) 包括负载控制表 (153) 和负载控制单元 (14)。负载控制表 (153) 控制开放流网络中的交换机 (2) 的每一个在预定时间内已经传输到控制交换机 (2) 的控制器 (1) 的消息的数目和交换机 (2) 在预定时间内已经接收的分组中的至少一个。当预定事件发生时,负载控制单元 (14) 根据负载控制表 (153),检测具有高于或等于第一阈值的处理负载的控制器 (1),并且使由所检测的控制器 (1) 控制的交换机 (2) 的至少一个处于另一控制器 (1) 的控制下。



1. 一种控制器,所述控制器用于控制在开放流网络中的多个交换机中的一些,所述控制器包括:

负载控制表,所述负载控制表被配置成控制下述中的至少一个:开放流网络中的所述交换机中的每一个在预定时间中已经向控制所述交换机的控制器传输的消息的数目、以及所述交换机在预定时间中已经接收到的分组的数目;以及

负载控制装置,所述负载控制装置被配置成当预定事件发生时,基于所述负载控制表来检测处理负载应当被分配的控制器,并且使由所检测到的控制器所控制的交换机中的至少一个处于另一控制器的控制下。

2. 根据权利要求1所述的控制器,其中,

所述处理负载是数值信息,所述数值信息的较高值指示较高负载,并且

所述负载控制装置检测具有高于或等于第一阈值的处理负载的控制器。

3. 根据权利要求2所述的控制器,其中,所述负载控制装置将下述控制器选择为所述另一控制器:该控制器具有相对低的处理负载。

4. 根据权利要求2所述的控制器,其中,

所述负载控制表控制有关在开放流网络中的所述交换机中的每一个与控制所述交换机的控制器之间的跳数的信息,并且

所述负载控制装置将下述交换机选择为将处于所述另一控制器的控制下的至少一个交换机:该交换机具有对所述处理负载相对小的影响并且具有大于或等于第二阈值的跳数。

5. 根据权利要求2或3所述的控制器,其中,所述负载控制装置通过使所述消息的数目和所接收到的分组的数目相乘来计算所述处理负载。

6. 根据权利要求1至5中的任何一项所述的控制器,其中,所述预定事件包括下述中的一个:预定时间的度过、对开放流网络添加交换机、以及从开放流网络中删除交换机。

7. 根据权利要求1至6中的任何一项所述的控制器,其中,所述负载控制装置基于所接收到的消息的数目和所接收到的分组的数目、以及由每个控制器所控制的交换机的数目,来计算所述控制器的处理负载。

8. 根据权利要求1至7中的任何一项所述的控制器,进一步包括拓扑更新装置,所述拓扑更新装置被配置成控制关于开放流网络的拓扑的信息,

其中,当所述预定事件发生时,所述拓扑更新装置检测具有连接故障的受控交换机,并且在所述负载控制装置执行处理之前,消除所检测到的交换机的连接故障。

9. 一种用于分配控制器的负载的方法,所述控制器控制开放流网络中的多个交换机中的一些,所述方法包括:

控制步骤,所述控制步骤控制负载控制表,所述负载控制表包括下述中的至少一个:开放流网络中的所述交换机中的每一个在预定时间中已经向控制所述交换机的控制器传输的消息的数目、以及所述交换机在预定时间中已经接收到的分组的数目;以及

负载控制步骤,当预定事件发生时,基于所述负载控制表来检测处理负载应当被分配的控制器,并且使由所检测到的控制器所控制的交换机中的至少一个处于另一控制器的控制下。

10. 根据权利要求9所述的方法,其中,

所述处理负载是数值信息,所述数值信息的较高值指示较高负载,并且

所述负载控制步骤包括检测具有高于或等于第一阈值的处理负载的控制器。

11. 根据权利要求 10 所述的方法,其中,所述负载控制步骤包括将下述控制器选择为所述另一控制器:该控制器具有相对低的处理负载。

12. 根据权利要求 11 所述的方法,其中,

所述负载控制表控制有关在开放流网络中的所述交换机中的每一个与控制所述交换机的控制器之间的跳数的信息,并且

所述负载控制步骤包括将下述交换机选择为将处于所述另一控制器的控制下的至少一个交换机:该交换机具有对所述处理负载相对小的影响并且具有大于或等于第二阈值的跳数。

13. 根据权利要求 10 或 11 所述的方法,其中,所述负载控制步骤包括通过使所述消息的数目和所接收到的分组的数目相乘来计算所述处理负载。

14. 根据权利要求 9 至 13 中的任何一项所述的方法,其中,所述预定事件包括下述中的一个:预定时间的度过、对开放流网络添加交换机、以及从开放流网络中删除交换机。

15. 根据权利要求 10 至 14 中的任何一项所述的方法,其中,所述负载控制步骤包括基于所接收到的消息的数目和所接收到的分组的数目、以及由每个控制器所控制的交换机的数目,来计算每个控制器的处理负载。

16. 一种存储程序的非瞬时计算机可读介质,所述程序用于使得计算机执行用于分配控制器的负载的方法,所述控制器控制开放流网络中的多个交换机中的一些,所述方法包括:

控制步骤,所述控制步骤控制负载控制表,所述负载控制表包括下述中的至少一个:开放流网络中的所述交换机中的每一个在预定时间中已经向控制所述交换机的控制器传输的消息的数目、以及所述交换机在预定时间段中已经接收到的分组的数目;以及

负载控制步骤,当预定事件发生时,基于所述负载控制表来检测处理负载应当被分配的控制器,并且使由所检测到的控制器所控制的交换机中的至少一个处于另一控制器的控制下。

17. 根据权利要求 16 所述的非瞬时计算机可读介质,其中,

所述处理负载是数值信息,所述数值信息的较高值指示较高负载,并且

所述负载控制步骤包括检测具有高于或等于第一阈值的处理负载的控制器。

18. 根据权利要求 17 所述的非瞬时计算机可读介质,其中,所述负载控制步骤包括将下述控制器选择为所述另一控制器:该控制器具有相对低的处理负载。

19. 根据权利要求 18 所述的非瞬时计算机可读介质,其中,

所述负载控制表控制有关在开放流网络中的所述交换机中的每一个与控制所述交换机的控制器之间的跳数的信息,并且

所述负载控制步骤包括将下述交换机选择为将处于所述另一控制器的控制下的至少一个交换机:该交换机具有对所述处理负载相对小的影响并且具有大于或等于第二阈值的跳数。

20. 根据权利要求 17 或 18 所述的非瞬时计算机可读介质,其中,所述负载控制步骤包括通过使所述消息的数目和所接收到的分组的数目相乘来计算所述处理负载。

21. 根据权利要求 16 至 20 中的任何一项所述的非瞬时计算机可读介质,其中,所述预定事件包括下述中的一个:预定时间的度过、对开放流网络添加交换机、以及从开放流网络中删除交换机。

22. 根据权利要求 21 所述的非瞬时计算机可读介质,其中,所述负载控制步骤包括当作为所述预定事件,将交换机添加到开放流网络或从开放流网络中删除交换机时,更新所述负载控制表,并且随后与所述第一阈值进行比较。

23. 根据权利要求 16 至 22 的任何一项所述的非瞬时计算机可读介质,其中,所述负载控制步骤包括基于所接收到的消息的数目和所接收到的分组的数目、以及由每个控制器所控制的交换机的数目,来计算每个控制器的处理负载。

24. 一种计算机系统,包括:

多个交换机,所述多个交换机中的每一个被配置成基于流表来在开放流网络中转发分组;以及

多个控制器,所述多个控制器中的每一个被配置成控制所述交换机中的一些,

其中,所述控制器中的每一个包括:

负载控制表,所述负载控制表被配置成控制下述中的至少一个:开放流网络中的所述交换机中的每一个在预定时间中已经向控制所述交换机的控制器传输的消息的数目、以及所述交换机在预定时间中已经接收到的分组的数目;以及

负载控制装置,所述负载控制装置被配置成当预定事件发生时,基于所述负载控制表来检测处理负载应当被分配的控制器,并且使由所检测到的控制器所控制的交换机中的至少一个处于另一控制器的控制下。

25. 一种用于分配多个控制器的负载的控制设备,所述控制器控制开放流网络中的交换机,所述控制设备包括:

负载控制表,所述负载控制表被配置成控制下述中的至少一个:开放流网络中的所述交换机中的每一个在预定时间中已经向控制所述交换机的控制器传输的消息的数目、以及所述交换机在预定时间中已经接收到的分组的数目;以及

负载控制装置,所述负载控制装置被配置成当预定事件发生时,基于所述负载控制表来检测处理负载应当被分配的控制器,并且使由所检测到的控制器所控制的交换机中的至少一个处于另一控制器的控制下。

## 控制器、用于分配负载的方法、存储程序的非瞬时计算机可读介质、计算机系统和控制设备

### 技术领域

[0001] 本发明涉及控制器、分配负载方法、存储程序的非瞬时计算机可读介质、计算机系统和控制设备。特别地，本发明涉及与开放流 (OpenFlow) 网络有关的控制器、分配负载方法、存储程序的非瞬时计算机可读介质、计算机系统和控制设备。

### 背景技术

[0002] 开放流网络是将由介质访问控制 (MAC) 地址、IP 地址和端口号的组合确定的一系列通信定义为“流”并且对每一流执行路径控制的网络控制技术。开放流网络包括用于计算分组的路径的控制器 (OFC :OpenFlow 控制器)、用于转发分组的交换机 (OFS :OpenFlow 交换机) 以及连接到交换机的终端。

[0003] 每一交换机包括对每一流,描述分组转发路径、分组转发方法等等的流表。控制器在交换机的流表条目中设定分组转发规则。每一交换机根据在流表中设定的转发规则转发分组。

[0004] 经由称为安全信道的安全套接层 / 传输层安全 (SSL/TLS) 或传输控制协议 (TCP), 连接控制器和交换机。经由安全信道传输或接收开放流协议消息。

[0005] 当交换机从终端接收分组时,交换机参考流表的报头字段 (报头字段属性) 并且搜索具有与该分组的报头信息匹配的报头字段的条目。如果这种条目存在,交换机更新该条目的统计信息 (“计数器”属性) 并且执行指定处理 (“动作”属性)。如果这种条目不存在,交换机将该分组传输到控制器 (packet-in 消息)。

[0006] 控制器接收 packet-in 消息并且计算该分组的路径。然后,根据所计算的路径 (flow-mod 消息),控制器将对应于该分组的条目添加到交换机的流表中。然后,控制器将该分组传输到交换机 (packet-out 消息)。

[0007] 如果在开放流网络中,更多终端连接到交换机,交换机将更多 packet-in 消息传输到控制器。在这种情况下,单一控制器可能不能够处理该消息。为此,最好在开放流网络中安装多个控制器,以便分发控制器从交换机接收的消息。

[0008] 在开放流网络中安装多个控制器中,每一控制器被分派由该控制器控制的交换机。因此,每一控制器将消息仅传输到由此控制的交换机或仅从其接收消息。因此,可以分发控制器从交换机接收的消息。

[0009] 假定多个控制器安装在开放流网络中,当在每一交换机的流表中设定转发规则时,每一控制器要求如下所述的路径信息和拓扑信息。为此,控制器使这些信息集相互同步。

[0010] (1) 路径信息

[0011] 路径信息是指示开放流网络中的最短路径的信息。每一控制器由拓扑信息计算路径信息。

[0012] (2) 拓扑信息

[0013] 拓扑信息是有关开放流网络中的交换机的连接的信息。每一控制器通过将链路层发现协议 (LLDP) 的查询分组等等定期地传输到交换机, 获得有关由该控制器控制的每一交换机的拓扑信息

[0014] 引用清单

[0015] 专利文献

[0016] [专利文献 1] 日本未审专利申请公开 No. 2011-166692

[0017] 非专利文献

[0018] [非专利文献 1] OpenFlow Switch Specification Version 1.1.0 Implemented (Wire Protocol 0x02) 2011 年 2 月 28 日, [2012 年 1 月 16 日检索的], 网络

[0019] <URL:http:

[0020] //www.openflow.org/documents/openflow-spec-v1.1.0.pdf>

## 发明内容

[0021] 技术问题

[0022] 如上所述, 可以在开放流网络中安装多个控制器, 并且由此分配控制器的负载。然而, 将近似相同数目的交换机简单分派给每一控制器不允许均衡控制器的负载。对此, 存在下述原因 (1) 和 (2)。

[0023] (1) 交换机的使用状态

[0024] 由控制器控制的交换机传输到该控制器的消息的数目在交换机间改变。消息的数目还随时区改变。

[0025] (2) 网络配置 (拓扑) 的更新

[0026] 由控制器控制的交换机传输到该控制器的消息的数目随由交换机之间的连接故障、交换机的添加或删除等等引起的拓扑配置的变化而改变。

[0027] 鉴于上述问题, 做出了本发明, 并且其主要目的是在包括多个控制器的开放流网络中提供一种能均衡控制器的负载的控制器、分配负载方法、存储程序的非瞬时计算机可读介质、计算机系统和控制设备。

[0028] 技术方案

[0029] 根据本发明的方面, 一种用于控制开放流网络中的多个交换机中的一些的控制器, 该控制器包括:

[0030] 负载控制表, 该负载控制表被配置成控制控制下述中的至少一个: 开放流网络中的交换机中的每一个在预定时间中已经向控制交换机的控制器传输的消息的数目、以及交换机在预定时间段中已经接收到的分组的数目; 以及

[0031] 负载控制装置, 该负载控制装置被配置成当预定事件发生时, 基于负载控制表来检测处理负载应当被分配的控制器, 并且使由所检测到的控制器所控制的交换机中的至少一个处于另一控制器的控制下。

[0032] 根据本发明的方面, 一种用于分配控制器的负载的方法, 控制器控制开放流网络中的多个交换机中的一些, 方法包括:

[0033] 控制步骤, 该控制步骤控制负载控制表, 该负载控制表包括下述中的至少一个: 开

放流网络中的交换机中的每一个在预定时间中已经向控制交换机的控制器传输的消息的数目、以及交换机在预定时间段中已经接收到的分组的数目；以及

[0034] 负载控制步骤,当预定事件发生时,基于负载控制表来检测处理负载应当被分配的控制器,并且使由所检测到的控制器所控制的交换机中的至少一个处于另一控制器的控制下。

[0035] 根据本发明的方面,一种存储程序的非瞬时计算机可读介质,该程序使得计算机执行用于分配控制器的负载的方法,该控制器控制开放流网络中的多个交换机中的一些,方该法包括:

[0036] 控制步骤,该控制步骤控制负载控制表,该负载控制表包括下述中的至少一个:开放流网络中的交换机中的每一个在预定时间中已经向控制交换机的控制器传输的消息的数目、以及交换机在预定时间段中已经接收到的分组的数目;以及

[0037] 负载控制步骤,当预定事件发生时,基于负载控制表来检测处理负载应当被分配的控制器,并且使由所检测到的控制器所控制的交换机中的至少一个处于另一控制器的控制下。

[0038] 根据本发明的方面,一种计算机系统,包括:

[0039] 多个交换机,每一交换机被配置成根据流表,在开放流网络中转发分组;以及

[0040] 多个控制器,每一控制器被配置成控制交换机中的一些,

[0041] 其中,该控制器的每一个包括:

[0042] 负载控制表,该负载控制表被配置成控制控制下述中的至少一个:开放流网络中的交换机中的每一个在预定时间中已经向控制交换机的控制器传输的消息的数目、以及交换机在预定时间段中已经接收到的分组的数目;以及

[0043] 负载控制装置,该负载控制装置被配置成当预定事件发生时,基于负载控制表来检测处理负载应当被分配的控制器,并且使由所检测到的控制器所控制的交换机中的至少一个处于另一控制器的控制下。

[0044] 根据本发明的方面,一种控制用于分配多个控制器的负载的设备,该控制器控制开放流网络中的交换机,该控制设备包括:

[0045] 负载控制表,该负载控制表被配置成控制控制下述中的至少一个:开放流网络中的交换机中的每一个在预定时间中已经向控制交换机的控制器传输的消息的数目、以及交换机在预定时间段中已经接收到的分组的数目;以及

[0046] 负载控制装置,该负载控制装置被配置成当预定事件发生时,基于负载控制表来检测处理负载应当被分配的控制器,并且使由所检测到的控制器所控制的交换机中的至少一个处于另一控制器的控制下。

[0047] 有益效果

[0048] 根据本发明,可以在包括多个控制器的开放流网络中提供一种能均衡控制器的负载的控制器、负载分配方法、存储程序的非瞬时计算机可读介质、计算机系统和控制设备。

## 附图说明

[0049] 图 1 是示出根据第一实施例的开放流网络的结构框图;

[0050] 图 2 是示出根据第一实施例的控制器结构的框图;

- [0051] 图 3 是示出根据第一实施例的负载控制表的结构的原则图；
- [0052] 图 4 是示出根据第一实施例的交换机的结构的框图；
- [0053] 图 5 是示出根据第一实施例的控制器控制操作的流程图；
- [0054] 图 6 是示出由根据第一实施例的控制器执行的负载信息更新处理的流程图；
- [0055] 图 7 是示出由根据第一实施例的控制器执行的负载分配处理的流程图；
- [0056] 图 8 是示出将交换机添加到根据第一实施例的开放流网络的处理的流程图；
- [0057] 图 9 是示出由根据第一实施例的交换机执行的控制器信息获得处理的流程图；
- [0058] 图 10 是示出从根据第一实施例的开放流网络删除交换机的处理的流程图；
- [0059] 图 11 是示出根据第一实施例的开放流网络的结构的框图；
- [0060] 图 12 示出根据第一实施例的负载控制表的示例结构；
- [0061] 图 13 示出根据第一实施例的负载控制表的示例结构；
- [0062] 图 14 是示出根据第一实施例的开放流网络的结构的框图；
- [0063] 图 15 是示出根据第一实施例的控制器的硬件结构的框图；以及
- [0064] 图 16 是示出根据本发明的开放流网络的结构的框图。

## 具体实施方式

### [0065] 第一实施例

[0066] 现在,将参考附图,描述本发明的实施例。首先,将概述根据本实施例的开放流网络。图 1 是示出根据本实施例的开放流网络的结构的框图。

[0067] 开放流网络包括多个控制器 1(1-1, 1-2)、多个交换机 2(2-1 至 2-4) 和多个终端(3-1 至 3-8)。控制器 1-1 控制交换机 2-1 和 2-2。具体地,它控制这些交换机的流表等等。类似地,控制器 1-2 控制交换机 2-3 和 2-4。具体地,它控制这些交换机的流表等等。当然,控制器 1-1 和 1-2 能控制任意多个交换机。在下文中,控制器 1-1 和 1-2 将简称为“控制器 1”,除非需要相互区分这些控制器。

[0068] 交换机 2-1 包括流表。交换机 2-1 是从终端 3-1 和 3-2 或另一交换机 2-2 接收分组,以及根据流表转发所接收的分组的设备。当然,交换机 2-1 能从任意多个终端接收分组。交换机 2-2 至 2-4 的结构与交换机 2-1 类似。在下文中,交换机 2-1 至 2-4 将称为“交换机 2”,除非需要相互区分这些交换机。

[0069] 终端 3-1 传输或接收分组。终端 3-1 是例如通用个人计算机、移动电话等等。同样适用于终端 3-2 至 3-7。在下文中,终端 3-1 至 3-8 将称为“终端 3”,除非需要相互区分这些终端。

[0070] 接着,参考图 2,将描述控制器 1 的内部结构和操作。图 2 是示出控制器 1 的内部结构的框图。控制器 1 包括消息控制单元 11、路径控制单元 12、拓扑更新单元 13、负载控制单元 14 和网络信息存储单元 15。控制器 1 只需要是包括中央处理单元(CPU) 等等的计算机。

[0071] 网络信息存储单元 15 存储路径信息 151、拓扑信息 152 和负载控制表 153。路径信息 151 包括所谓的“流信息”,指示每一交换机 2 如何传输分组。拓扑信息 152 指示开放流网络的结构。负载控制表 153 是保存用于掌握每一控制器 1 的负载的状态的信息的表。稍后将参考图 3 描述负载控制表 153。



[0072] 在所有控制器 1 之间,同步网络信息存储单元 15 中的信息集(控制器 1 相互告知信息集的任何变化)。即,当更新一个控制器 1 的网络信息存储单元 15 中的任何信息时,同样地更新另一控制器 1 的每一个的网络信息存储单元 15 中的相应信息。

[0073] 消息控制单元 11 是与由控制器 1 控制的交换机 2 传输或接收由开放流网络协议定义的各种类型消息的处理单元。基于所接收的消息,消息控制单元 11 指示处理单元(路径控制单元 12、拓扑更新单元 13、负载控制单元 14) 执行处理。

[0074] 路径控制单元 12 通过消息控制单元 11,从交换机 2 接收 packet-in 消息。然后,路径控制单元 12 参考拓扑信息 152 并且根据 packet-in 消息计算路径。然后,路径控制单元 12 根据所计算的路径,更新路径信息 151。

[0075] 拓扑更新单元 13 通过将查询分组传输到交换机 2,获得有关由控制器 1 控制的交换机 2 的每一个的连接的信息。然后,拓扑更新单元 13 根据所获得的连接信息,更新拓扑信息 152。

[0076] 负载控制单元 14 参考并更新负载控制表 153 中的负载信息。当预定事件(预定时间度过、添加或删除交换机 2 等等)发生时,负载控制单元 14 计算每一控制器 1 的处理负载并且检测具有高处理负载的控制器 1。然后,负载控制单元 14 确定是否应当分配所检测的控制器 1 的处理负载。如果确定应当分配处理负载,负载控制单元 14 将由所检测的控制器 1 控制的交换机 2 的一些置于具有低处理负载的控制器 1 的控制下。稍后将参考图 6 和 7,描述处理负载分配处理的详情。

[0077] 图 3 是示出负载控制表 153 的结构的原则图。负载控制表 153 是具有作为属性的交换机 ID、控制器 ID、跳数、所接收的消息计数和所接收的分组计数的表。负载控制表 153 具有对应于开放流网络中的交换机的条目。

[0078] 交换机 ID 是用于唯一地识别开放流网络中的交换机 2 的属性并且是用作负载控制表 153 的主关键字的信息。交换机 ID 是例如由交换机 2 的 MAC 地址的 48 位组成的数据路径 ID,其 12 位取决于实施方式。控制器 1 能在建立安全信道后,通过与交换机 2 交换开放流协议 Features/ 应答消息,获得交换机 2 的数据路径 ID。

[0079] 控制器 ID 是用于唯一地识别开放流网络中的控制器 1 的属性。控制器 ID 可以是例如控制器 1 的网际协议(IP) 地址。

[0080] 跳数是指示控制器 1 和交换机 2 之间的跳数的属性(分组通过的通信设备的数目)。

[0081] 所接收的消息计数是指示在预定时间内,控制器 1 已经从交换机 2 接收的消息的数目的属性。每次控制器 1 从交换机 2 接收消息时,增加所接收的消息计数(+1)。

[0082] 所接收的分组计数是指示由控制器 1 控制的交换机 2 已经从终端 3 或相邻的交换机 2 接收的分组的数目的属性。控制器 1 能通过由交换机 2 传输或接收开放流协议统计请求/ 应答消息,掌握交换机 2 的每一端口在预定时间内已经接收的分组的数目。

[0083] 接着,参考图 4,描述根据本实施例的交换机 2 的内部结构和操作。图 4 是示出根据本实施例的交换机 2 的内部结构的框图。

[0084] 交换机 2 包括分组控制单元 21 和消息控制单元 22。交换机 2 存储流表 23、交换机 ID 24 和控制器信息 25。控制器信息 25 包括控制器 ID 251 和跳数 252。

[0085] 当分组控制单元 21 从终端 3 或相邻交换机 2 接收分组时,参考流表 23,确定所接

收的分组的目的地,然后将该分组转发到该目的地。当将新的交换机 2 添加到开放流网络时,分组控制单元 21 通过将查询分组传输到交换机 2,获得相邻交换机 2 中的控制器信息 25。

[0086] 消息控制单元 22 与控制器 1 的消息控制单元 11 传输或接收由开放流协议定义的消息。当消息控制单元 22 从控制器 1 接收 flow-mod 消息时,根据所接收的消息,更新流表 23。此外,当交换机 2 与控制器 1 建立安全信道时,消息控制单元 22 将交换机 ID 24 传输到控制器 1。

[0087] 流表 23 是由控制器 1 设定的信息。对流表 23 中的每一通信流,描述分组转发路径、分组转发方法等等。有关细节,请参见例如非专利文献 1。

[0088] 交换机 ID24 是唯一地识别交换机 2 的信息。控制器 ID 251 是指示控制流表 23 的控制器 1 的 ID 的信息。跳数 252 是交换机 2 和控制器 1 之间的跳数。

[0089] 接着,参考图 5 至 7,描述控制器 1 的操作。如上所述,当预定事件(预定时间度过等等)发生时,控制器 1 的负载控制单元 14 执行分配每一控制器 1 的负载的处理(负载分配处理)。在下文中,将详细地描述控制器 1 的负载分配处理及相关操作。

[0090] 图 5 是示出控制器 1 的控制操作的流程图。控制器 1 等待预定时间经过(步骤 A1)。在预定时间经过后,控制器 1 执行负载信息更新处理(图 6)(步骤 A2)。

[0091] 图 6 是示出由控制器 1 执行的负载信息更新处理的详情的流程图。首先,控制器 1 的拓扑更新单元 13 通过将查询分组传输到交换机 2,获得有关由控制器 1 控制的每一交换机 2 的每一端口的连接的信息(步骤 B1)。如果控制器 1 检测到具有连接故障的交换机 2(在下文中,称为交换机 S0)(步骤 B2:是),拓扑更新单元 13 将交换机 S0 告知负载控制单元 14。然后,负载控制单元 14 计算开放流网络中的每一控制器 1 的处理负载来获得具有最低处理负载的控制器 1(在下文中,称为控制器 C0)的 ID(步骤 B3)。注意,稍后将参考图 7,描述计算处理负载的方法。

[0092] 进一步,负载控制单元 14 在负载控制表 153 中,在交换机 S0 的条目的“控制器 ID”属性中设定控制器 C0 的 ID(步骤 B4)。控制器 C0 的负载控制单元 14 也通过安全信道,在交换机 S0 的控制器 ID 251 中设定控制器 C0 的 ID(步骤 B5)。

[0093] 控制器 1 的负载控制单元 14 还从由控制器 1 控制的交换机 2 的每一个获得所接收的分组计数并在负载控制表 153 中设定它们(步骤 B6, B7)。此时,负载控制单元 14 排除由于拓扑的变化未使用过的端口的所接收的分组数。由于上述步骤 B1 至 B7,已经消除了连接故障,并且已经获得所接收的分组计数。在这种状态下,执行控制器负载分配处理(步骤 B8)。图 7 示出了负载分配处理的详情(步骤 B8)。

[0094] 负载控制单元 14 参考负载控制表 153 并计算每一控制器 1 的处理负载。例如,负载控制单元 14 通过使所接收的消息计数和所接收的分组计数相乘,计算以数值信息形式的处理负载。然后,负载控制单元 14 获得具有最高处理负载的控制器 1(在下文中,称为控制器 C1)的 ID(步骤 C1)。

[0095] 注意,处理负载不一定是数值信息,也可以是基于预定阈值的诸如“高”、“中”或“低”的信息。

[0096] 然后,负载控制单元 14 确定控制器 C1 的处理负载是否小于或等于预定阈值(第一阈值)(步骤 C2)。如果处理负载小于或等于预定阈值(步骤 C2:是),负载控制单元 14

结束该处理。

[0097] 相反,如果控制器 C1 的处理负载不小于或等于预定阈值(步骤 C2:否),负载控制单元 14 选择具有最低处理负载的控制器 1(在下文中,称为控制器 C2)并且获得其 ID(步骤 C3)。

[0098] 然后,负载控制单元 14 从由控制器 C1 控制的交换机 2 中,选择对处理负载具有相对小影响的交换机 2(在下文中,称为交换机 S1)并且获得其 ID(步骤 C4)。换句话说,负载控制单元 14 根据对处理负载的影响大小的升序,选择交换机 2。例如,当负载控制单元 14 首次执行步骤 C4 时,将对处理负载具有最小影响的交换机 2 选择为交换机 S1。

[0099] 然后,负载控制单元 14 从负载控制表 153 获得交换机 S1 的跳数并且确定所获得的跳数是否大于或等于阈值(第二阈值)(步骤 C5)。如果跳数不大于或等于阈值(步骤 C5:否),负载控制单元 14 返回到步骤 C4。此时,负载控制单元 14 选择仅次于在上一步骤 C4 中选择的交换机 S1 的、具有最小处理负载的交换机 2。

[0100] 如果跳数大于或等于阈值(步骤 C5:是),负载控制单元 14 在负载控制表 153 的交换机 S1 的条目的“控制器 ID”属性中,设定控制器 C2 的 ID(步骤 C6)。负载控制单元 14 还通过安全信道,在交换机 S1 的控制器 ID 251 中,设置控制器 C2 的 ID(步骤 C7)。

[0101] 在步骤 C7 后,负载控制单元 14 返回到步骤 C1。

[0102] 尽管在上述处理中,负载控制单元 14 通过与预定阈值(第一阈值)进行比较,选择控制器 C1,但也可以以其他方式选择控制器 C1。即,负载控制单元 14 可以根据任何标准,选择应当分配其处理负载的控制器 1。例如,负载控制单元 14 可以将具有最高处理负载的控制器 1 无条件地选择为控制器 C1。

[0103] 尽管在上述处理中,对跳数做出了确定,但在不确定跳数的情况下,也可以将具有最低处理负载的交换机 2 确定为交换机 S1。然而,跳数的确定具有简单地分配处理负载的效果,还具有防止网络结构复杂化的效果。

[0104] 尽管在上述描述中,将具有最低处理负载的控制器 1 选择为控制器 1(C2)来交替地控制交换机 2,但可以选择具有小于或等于预定值的处理负载的任何控制器 1。然而,通过根据处理负载的大小的升序选择控制器 1,能最佳地均衡处理负载。

[0105] 接着,参考图 8 和 9,描述当新的交换机 2 添加到开放流网络时,由设备(控制器 1、交换机 2)执行的处理。图 8 是示出添加交换机 2 的处理的流程图。

[0106] 首先,待添加的交换机 2(在下文中,称为“对象交换机 2”)的分组控制单元 21 从相邻交换机 2 获得控制器信息 25(步骤 D1)。将参考图 9,描述该控制器信息获得处理。

[0107] 对象交换机 2 的分组控制单元 21 将查询分组广播到相邻交换机并且接收控制器信息 25(步骤 E1)。如果分组控制单元 21 从多个交换机 2 接收控制器信息 25 的集合(步骤 E2:是),它选择包括最小跳数 252 的控制器信息 25 的集合(步骤 E3)。然后,分组控制单元 21 通过消息控制单元 22 与控制器 1 通信并且在与控制器 1 联系前,获得跳数 252(步骤 E4)。

[0108] 再参考图 8,已经获得控制器信息 25(步骤 D1)的交换机 2 的消息控制单元 22 参考控制器信息 25 中的控制器 ID 251,并且与由控制器 ID 251 识别的控制器 1 建立安全信道(步骤 D2)。

[0109] 然后,消息控制单元 22 从控制器 1 的负载控制表 153 获得另一控制器 1 的 ID 并

且与那些控制器 1 建立安全信道（步骤 D3, D4）。

[0110] 开放流网络中的每一控制器 1 的负载控制单元 14 将对象交换机 2 的条目添加到负载控制表 153（步骤 D5）。此时，负载控制单元 14 在负载控制表 153 的“交换机 ID”属性中，设定对象交换机 2 的交换机 ID 24 的值。负载控制单元 14 还在负载控制表 153 的“跳数”属性中，设定对象交换机 2 的跳数 252 的值。负载控制单元 14 还在负载控制表 153 的“所接收的消息计数”和“所接收的分组计数”属性的每一个中设置 0。然后，负载控制单元 14 执行负载信息更新处理（图 6 的处理）（步骤 D6）。

[0111] 接着，参考图 10，将描述当从开放流网络删除交换机 2 时，由设备（控制器 1、交换机 2）执行的处理。图 10 是示出删除交换机 2 的处理的流程图。

[0112] 待删除的交换机 2（在下文中，称为“对象交换机 2”）的消息控制单元 22 关闭开放流网络中，与所有控制器 1 的安全信道（步骤 F1）。作为响应，每一控制器 1 的负载控制单元 14 从负载控制表 153 删除对象交换机 2 的条目（步骤 F2）。然后，负载控制单元 14 执行负载信息更新处理（图 6 的处理）（步骤 F3）。

[0113] 接着，参考图 11 至 14，描述图 7 中所示的、分配控制器 1 的处理负载的具体的处理示例。图 11 是示出在执行负载分配处理前，开放流网络的结构框图。图 12 是示出在执行负载分配处理前，负载控制表 153 的结构原理图。

[0114] 负载控制单元 14 计算开放流网络中的控制器（控制器 A 和控制器 B）的处理负载（步骤 C1）。在该实施例中，负载控制单元 14 使用下述公式，计算每一控制器的处理负载。

[0115] 待计算的控制器 C 的处理负载 = （控制器 C 的所接收的消息计数 / 所有控制器的总接收的消息计数）\*（由该控制器 C 控制的交换机的所接收的分组计数 / 所有交换机的总接收的分组计数）

[0116] 参考图 12，总接收的消息计数为 120 以及总接收的分组计数为 800。因此，控制器 A 的处理负载是  $(50/120) * (300/800) \approx 0.16$ 。类似地，控制器 B 的处理负载是  $(70/120) * (500/800) \approx 0.36$ 。因此，负载控制单元 14 将控制器 B 选择为具有最高处理负载的控制器（步骤 C1）。

[0117] 上述计算方法仅是说明性的，并且可以采用其他方法。例如，可以将乘法公式的左侧或右侧（或两侧）乘以加权因子。还可以通过加法而不是乘法计算处理负载。即，负载控制单元 14 只需要根据负载控制表 153 中的信息，计算每一控制器 1 的处理负载。注意，负载控制单元 14 能通过使用负载控制表 153 中的所接收的消息计数和所接收的分组计数的至少一个，掌握每一控制器 1 的处理负载的趋势，尽管会降低计算精度。

[0118] 然后，负载控制单元 14 确定所选控制器 B 的处理负载是否小于或等于预定阈值（步骤 C2）。假定阈值为 0.3，负载控制单元 14 确定控制器 B 的处理负载是否小于或等于该阈值（步骤 C2：否）。然后，负载控制单元 14 将控制器 A 选择为具有最低处理负载的控制器（步骤 C3）。

[0119] 然后，负载控制单元 14 获得对由控制器 B 控制的交换机的处理负载具有最小影响的交换机的 ID。参考图 12，交换机 C 对处理负载的影响为  $(5/120) * (100/800) \approx 0.01$ 。交换机 D 对处理负载的影响为  $(30/120) * (200/800) \approx 0.06$ 。交换机 E 对处理负载的影响为  $(15/120) * (100/800) \approx 0.02$ 。交换机 F 对处理负载的影响为  $(20/120) * (100/800) \approx 0.02$ 。因此，负载控制单元 14 将交换机 C 选择为对处理负载具有最小影响的交换机（步骤 C4）。

[0120] 然后,负载控制单元 14 参考负载控制表 153 并且确定交换机 C 的跳数是否是阈值 (步骤 C5)。假定阈值为 2,负载控制单元 14 确定交换机 C 的跳数大于或等于该阈值 (步骤 C5:是)。

[0121] 然后,负载控制单元 14 在负载控制表 153 的交换机 C 的条目的“控制器 ID”属性中设定控制器 A 的 ID(步骤 C6)。负载控制单元 14 还通过安全信道,在交换机 C 的控制器 ID 251 中设定控制器 A 的 ID(交换机 C7)。图 13 示出当该处理完成时的时间点的负载控制表 153。下划线部分表示控制器的变化。尽管跳数值可能随控制器的变化而改变,但在本实施例中,假定该跳数不改变。

[0122] 然后,负载控制单元 14 再次计算开放流网络中的控制器 (控制器 A 和控制器 B) 的处理负载 (步骤 C1)。

[0123] 参考图 13,总接收的消息计数为 120,以及总接收的分组计数为 800。因此,控制器 A 的处理负载是  $(55/120)*(400/800) \approx 0.23$ 。类似地,控制器 B 的处理负载是  $(65/120)*(400/800) \approx 0.27$ 。因此,负载控制单元 14 将控制器 B 选择为具有最高处理负载的控制器 (步骤 C1)。

[0124] 然后,负载控制单元 14 确定所选控制器 B 的处理负载是否小于或等于预定阈值 (步骤 C2)。由于该阈值为 0.3,因此,负载控制单元 14 确定控制器 B 的处理负载小于或等于该阈值 (步骤 C2:是),结束该处理。图 14 是示出在执行负载分配处理后,开放流网络的结构框图。

[0125] 接着,将描述根据本实施例的开放流网络的效果。负载控制单元 14 根据交换机 2 的使用状态或开放流网络的拓扑,分配控制器 1 的处理负载。由此,可以避免处理负载不均匀地施加在特定控制器 1 上以及准确地分配负载。

[0126] 此外,当将交换机 2 添加到开放流网络或从其删除时,控制器 1 在分配负载前更新拓扑信息 (执行上述步骤 B1 至 B7)。由此,能在更新开放流网络中的每一设备的连接状态后分配负载。

[0127] 第二实施例

[0128] 根据本发明的第二实施例的控制器 1 的特征在于控制器 1 在图 7 的负载分配处理中,考虑由每一控制器控制的交换机的数目。在下文中,将描述第二实施例与第一实施例之间的区别。

[0129] 在上述图 7 中,负载控制单元 14 使用下述公式,计算每一控制器 1 的处理负载。

[0130] 待计算的控制器 C 的处理负载 = (由该控制器 C 控制的交换机的数目 / 开放流网络中的所有交换机的数目) \* (控制器 C 的所接收的消息的计数 / 所有控制器的总的接收的消息计数) \* (由该控制器 C 控制的交换机的所接收的分组计数 / 所有交换机的总的接收的分组计数)

[0131] 接着,再参考图 11 至 14,将描述根据本实施例,分配控制器 1 的负载的处理的具体示例 (图 7)。如上所述,图 11 是示出在执行负载分配处理前,开放流网络的结构框图。图 12 是示出在执行负载分配处理前,负载控制表 153 的结构原理图。

[0132] 参考图 12,总交换机计数为 6 以及总接收的消息计数为 120。因此,控制器 A 的处理负载为  $(2/6)*(50/120)*(300/800) \approx 0.05$ 。类似地,控制器 B 的处理负载为  $(4/6)*(70/120)*(500/800) \approx 0.24$ 。因此,负载控制单元 14 将控制器 B 选择为具有最高

处理负载的控制器（步骤 C1）。

[0133] 然后，负载控制单元 14 确定所选控制器 B 的处理负载是否小于或等于预定阈值（步骤 C2）。假定阈值为 0.2，负载控制单元 14 确定控制器 B 的处理负载不小于或等于该阈值（步骤 C2：否）。然后，负载控制单元 14 将控制器 A 选择为具有最低处理负载的控制器（步骤 C3）。

[0134] 然后，负载控制单元 14 获得对由控制器 B 控制的交换机的处理负载具有最小影响的交换机的 ID。负载控制单元 14 根据图 12，计算对处理负载的影响如下。

[0135] 交换机 C： $(5/120) * (100/800)$

[0136] 交换机 D： $(30/120) * (200/800)$

[0137] 交换机 E： $(15/120) * (100/800)$

[0138] 交换机 F： $(20/120) * (100/800)$

[0139] 因此，负载控制单元 14 将交换机 C 选择为对处理负载具有最小影响的交换机（步骤 C4）。然后，负载控制单元 14 参考负载控制表 153 并且确定交换机 C 的跳数是否大于或等于阈值（步骤 C5）。假定阈值为 2，负载控制单元 14 确定交换机 C 的跳数大于或等于该阈值（步骤 C5：是）。

[0140] 然后，负载控制单元 14 在负载控制表 153 的交换机 C 的条目的“控制器 ID”属性中设定控制器 B 的 ID（步骤 C6）。负载控制单元 14 还通过安全信道，在交换机 C 的控制器 ID 251 中设定控制器 B 的 ID（交换机 C7）。图 13 示出当该处理完成时的时间点的负载控制表 153。下划线部分表示控制器的变化。

[0141] 负载控制单元 14 再次计算开放流网络中的控制器（控制器 A 和控制器 B）的处理负载（步骤 C1）。

[0142] 参考图 13，可知总交换机计数为 6，总的接收的消息计数为 120，以及总的接收的分组计数为 800。因此，控制器 A 的处理负载为  $(3/6) * (55/120) * (400/800) \approx 0.11$ 。类似地，控制器 B 的处理负载为  $(3/6) * (65/120) * (400/800) \approx 0.13$ 。因此，负载控制单元 14 将控制器 B 选择为具有最高处理负载的控制器（步骤 C1）。

[0143] 然后，负载控制单元 14 确定所选控制器 B 的处理负载是否小于或等于预定阈值（步骤 C2）。由于该阈值为 0.2，因此，负载控制单元 14 确定控制器 B 的处理负载小于或等于该阈值（步骤 C2：是），结束该处理。图 14 是示出在执行负载分配处理后，开放流网络的结构框图。

[0144] 同样在本实施例中，可以避免负载不均匀地施加在特定控制器 1 上以及准确地分配负载。如上可知，负载控制单元 14 还考虑总交换机计数和由每一控制器 1 控制的交换机的数目之间的关系，然后分配负载。由此，可以执行还考虑网络的连接状态的负载分配（即，避免由每一控制器控制的交换机的数目在网络中不均匀的负载分配）。

[0145] 尽管已经对于其实例描述了本发明，但本发明不限于其结构。当然，本发明包括本领域的技术人员在不背离本申请的权利要求的范围的情况下，对实施例所做的变化、修改或组合。

[0146] 由控制器 1 的处理单元（消息控制单元 11、路径控制单元 12、拓扑更新单元 13 和负载控制单元 14）执行的处理能体现为在任何计算机上运行的程序。该程序能存储在各种类型的非瞬时计算机可读介质中并提供给计算机。非瞬时计算机可读介质的示例包括各种

类型的有形存储介质。非瞬时计算机可读介质的示例包括磁存储介质（例如柔性盘、磁带、硬盘驱动）、磁光存储介质（例如，磁光盘）、光盘只读存储器（CD-ROM）、CD-R、CD-R/W、半导体存储器（例如，掩膜型ROM、可编程ROM(PROM)、可擦除PROM(EPROM)、闪速ROM和随机存取存储器(RAM)。可以通过各种类型的瞬时计算机可读介质，将程序提供给计算机。瞬时计算机可读介质的示例包括电信号、光信号和电磁波。瞬时计算机可读介质能经由有线通信路径，诸如电线或光纤，或经由无线通信路径将程序提供给计算机。

[0147] 图 15 是示出控制器 1 的示例硬件结构的图。控制器 1 包括中央处理单元 (CPU) 401 和存储器 402。CPU 401 和存储器 402 通过总线，连接到用作辅助存储设备的硬盘驱动器 (HDD) 403。存储介质，诸如 HDD 403 能存储用于协同操作系统，向 CPU 401 等等提供指令并且执行 CPU 401 等等的单元的功能的计算机程序。即，当该程序加载到存储器 402 中时，形成控制器 1 的块，然后，CPU 401 执行根据程序的处理并且与其他硬件组件合作。然后，当 CPU 401 执行预定程序时，控制器 1 执行处理。

[0148] 图 16 是示意性地示出本发明的框图。开放流网络包括多个控制器 1 和多个交换机 2。每一控制器 1 包括负载控制表 153 和负载控制单元 14。

[0149] 负载控制表 153 控制在预定时间内，开放流网络中的每一交换机 2 已经传输到控制该交换机 2 的控制器 1 的消息的数目和在预定时间内，每一交换机 2 已经接收的分组的数目的至少一个。即，负载控制表 153 保存专用于开放流网络的负载信息。当预定事件发生时，负载控制单元 14 参考负载控制表 153 并且检测具有高于或等于第一阈值的处理负载的控制器 1。例如，负载控制单元 14 通过将所接收的分组计数和所接收的消息计数相乘，计算处理负载。然后，负载控制单元 14 使由所检测的控制器 1 控制的交换机 2 的至少一个处于另一控制器 1 的控制下。

[0150] 根据上述结构，负载控制单元 14 能根据交换机 2 的使用状态或开放流网络的拓扑，分配控制器 1 的负载。即，可以避免负载不均匀地施加在特定控制器 1 上。

[0151] 尽管在上述描述中，每一控制器 1 包括负载控制表 153 和负载控制单元 14，其他结构也是可能的。例如，可以在开放流网络中提供用于指示每一控制器 1 控制交换机 2 的控制设备。这种控制设备包括负载控制表 153 和负载控制单元 14。控制设备计算每一控制器 1 的处理负载以及使由具有高处理负载的控制器 1 控制的交换机 2 处于另一控制器 1 的控制下。该结构也能产生上述效果。

[0152] 本申请要求基于 2012 年 2 月 2 日提交的、日本专利申请 No. 2012-020695 的优先权，其公开被整体结合于此。

[0153] 附图标记列表

- [0154] 1 控制器
- [0155] 11 消息控制单元
- [0156] 12 路径控制单元
- [0157] 13 拓扑更新单元
- [0158] 14 负载控制单元
- [0159] 15 网络信息存储单元
- [0160] 151 路径信息
- [0161] 152 拓扑信息

---

[0162]	153	负载控制表
[0163]	2	交换机
[0164]	21	分组控制单元
[0165]	22	消息控制单元
[0166]	23	流表
[0167]	24	交换机 ID
[0168]	25	控制器信息
[0169]	251	控制器 ID
[0170]	252	跳数
[0171]	3	终端
[0172]	401	CPU
[0173]	402	存储器
[0174]	403	HDD



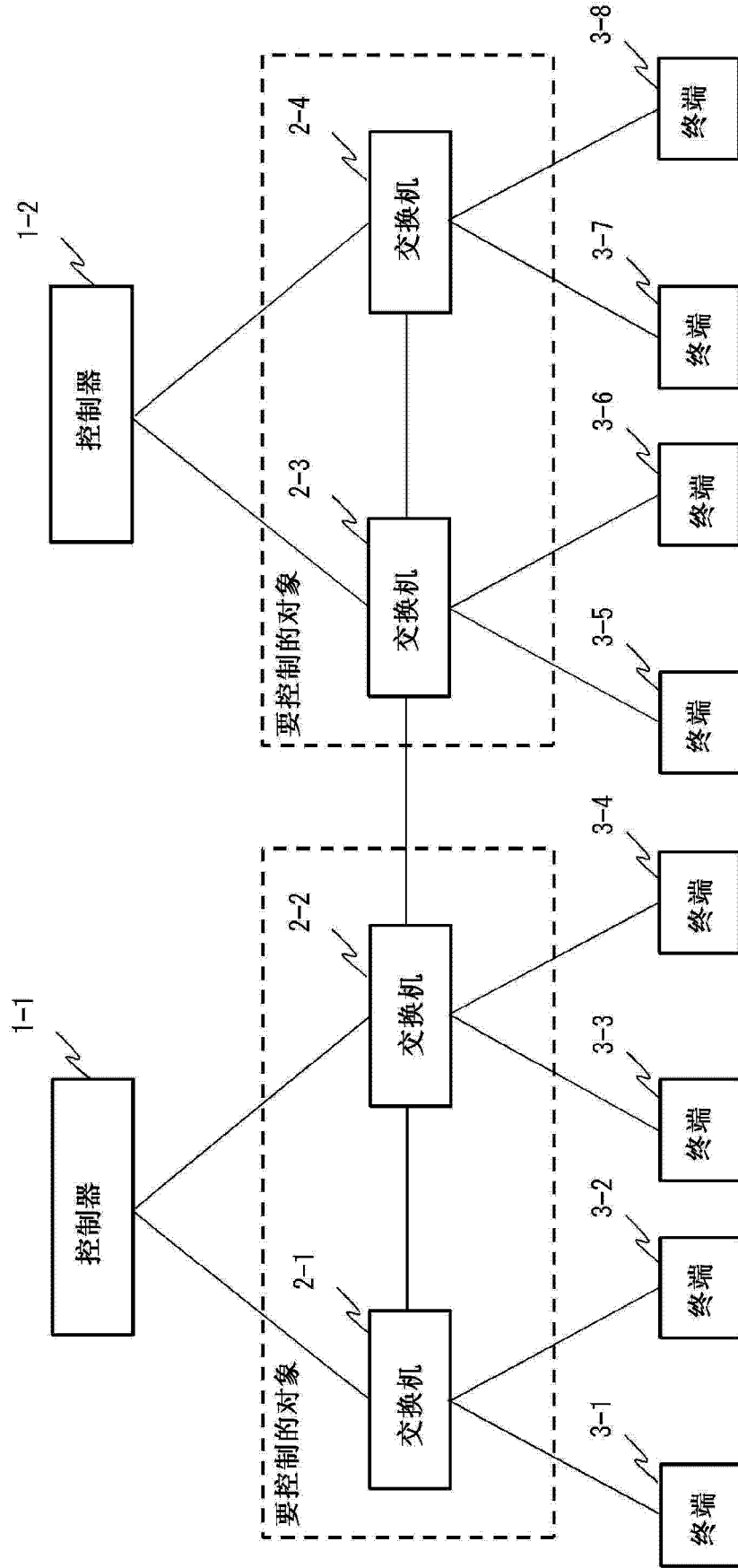


图 1

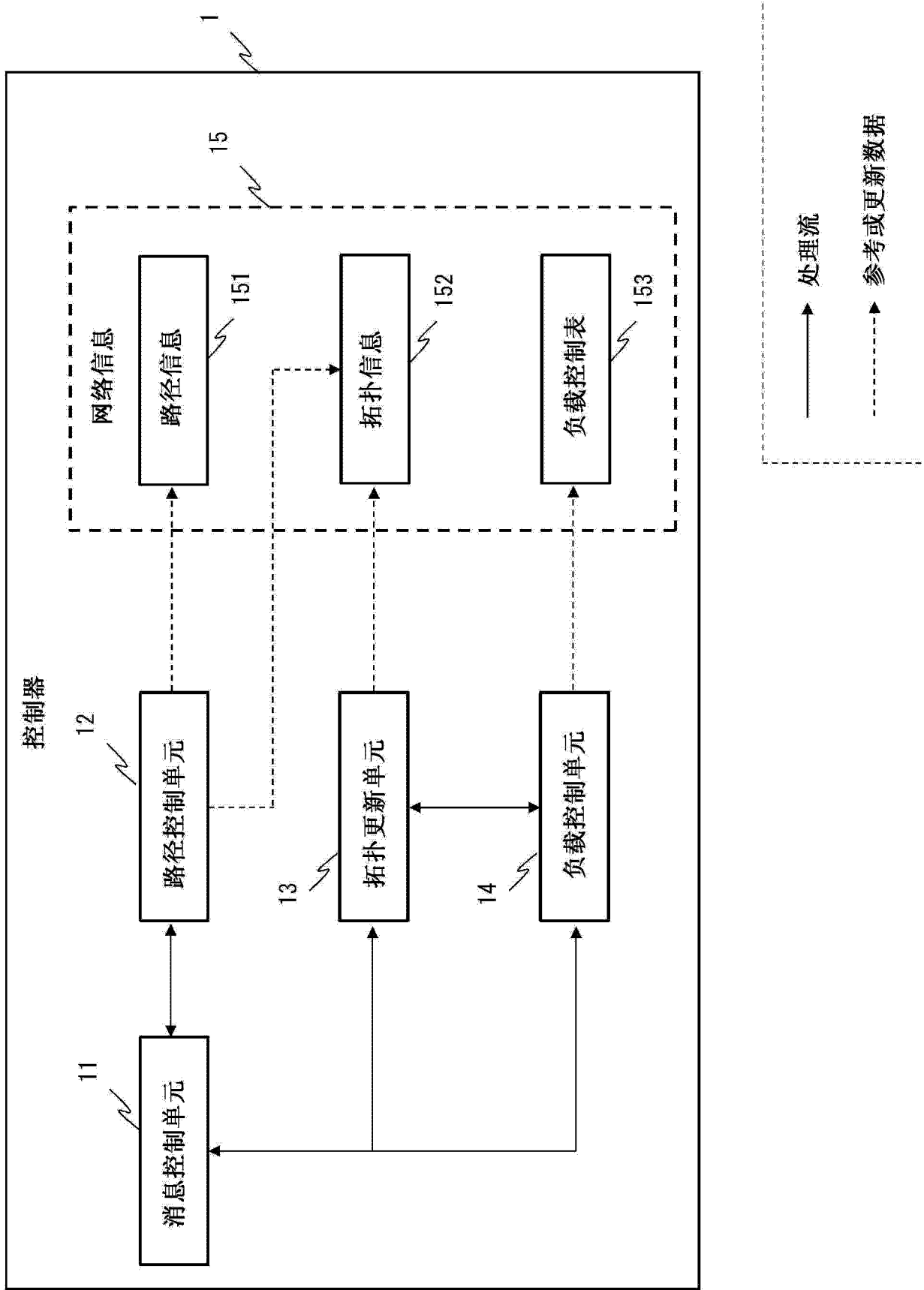


图 2

交换机ID	控制器ID	跳数	所接收的消息计数	所接收的分组计数
交换机A	控制器A	1	30	100
交换机B	控制器A	2	20	200
交换机C	控制器A	1	5	100
交换机D	控制器B	1	30	200
-	-	-	-	-

图 3

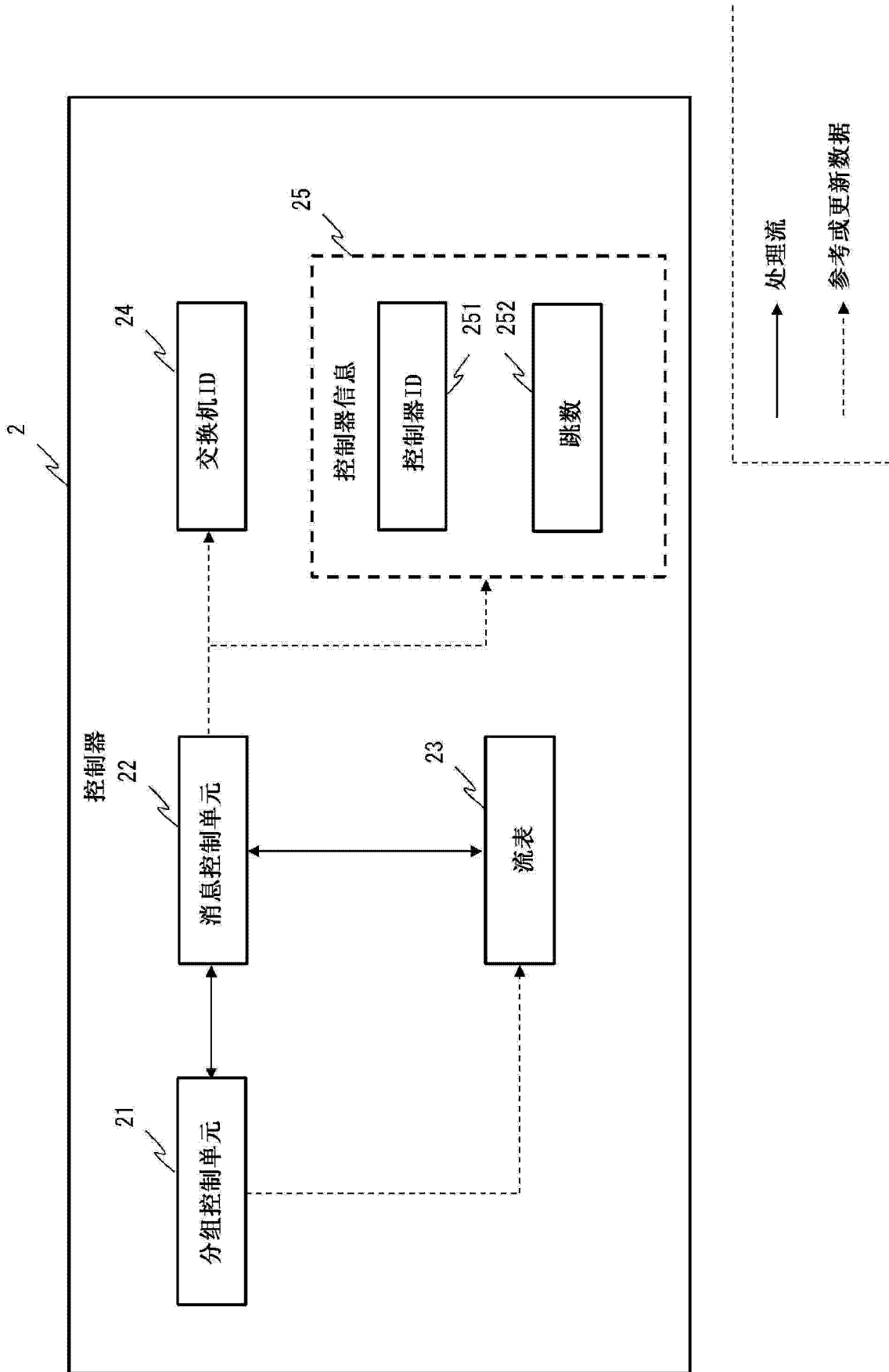


图 4

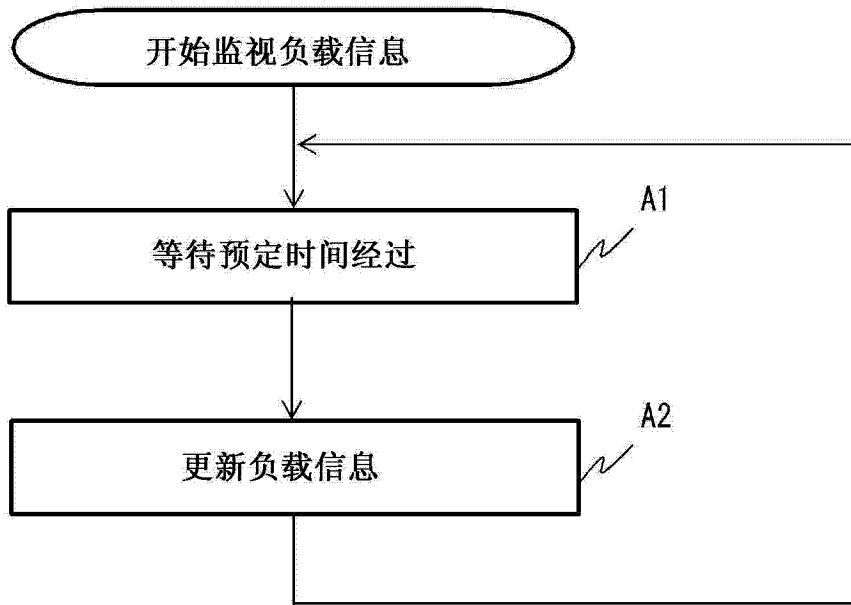


图 5

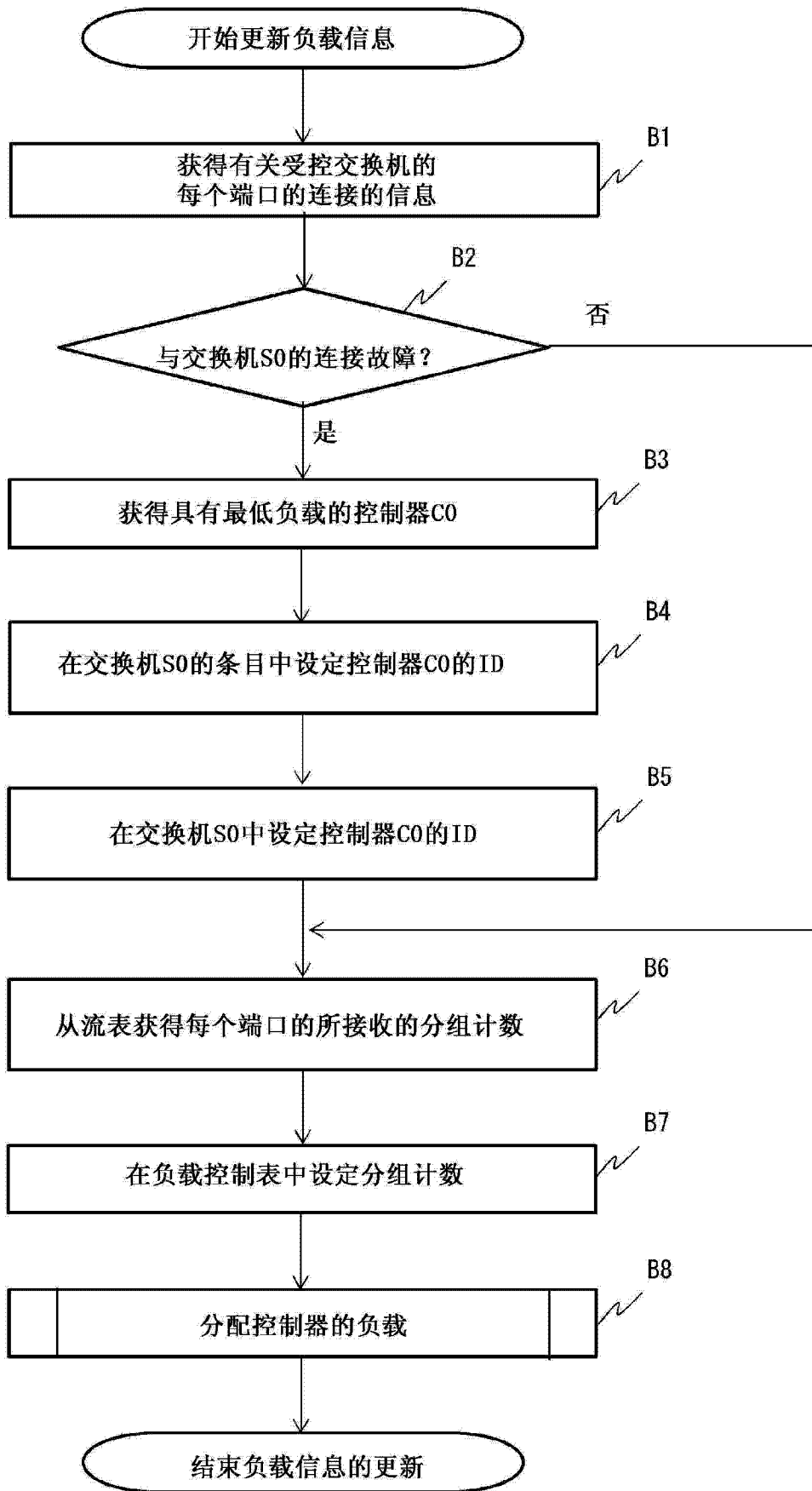


图 6

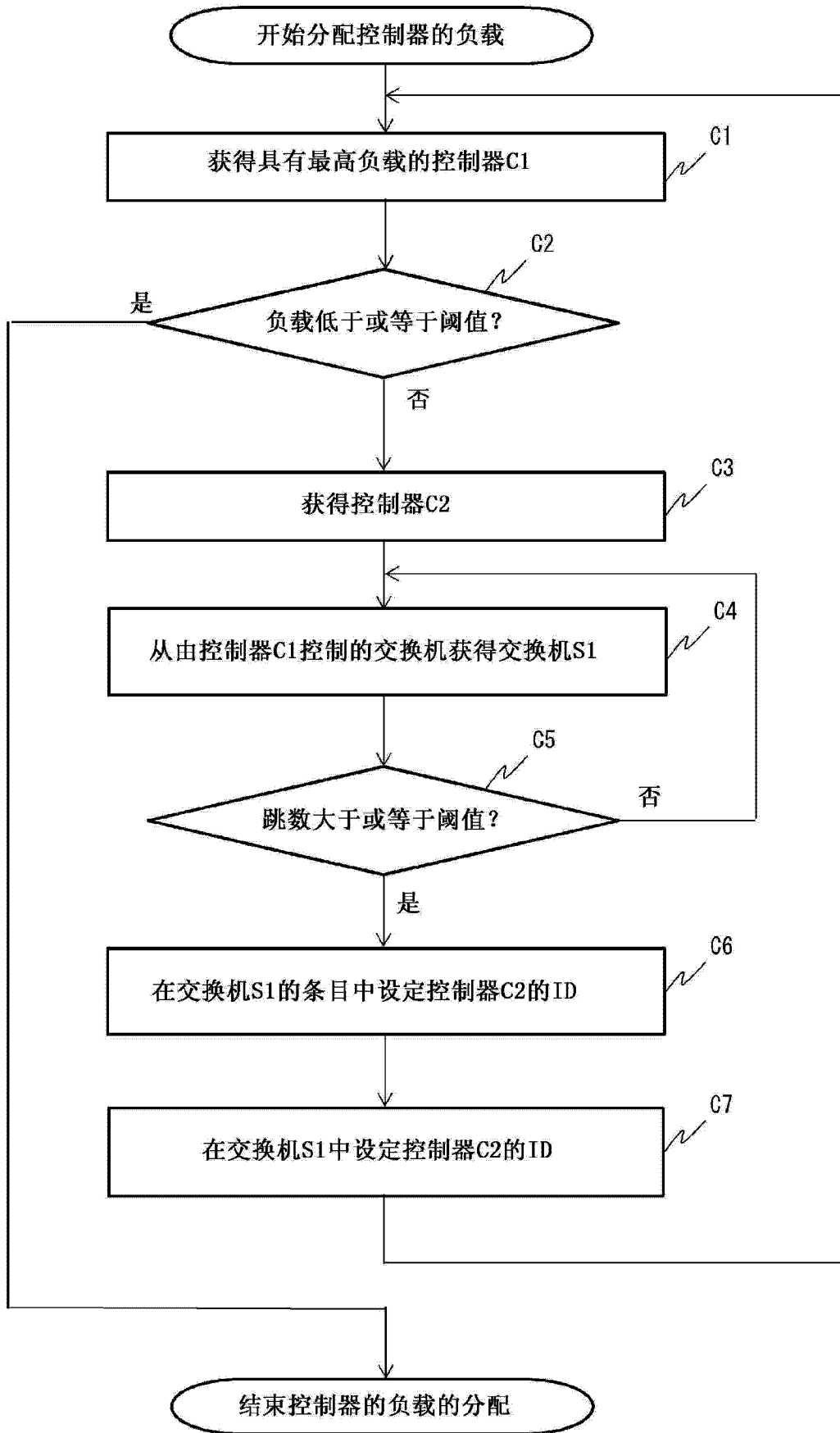


图 7

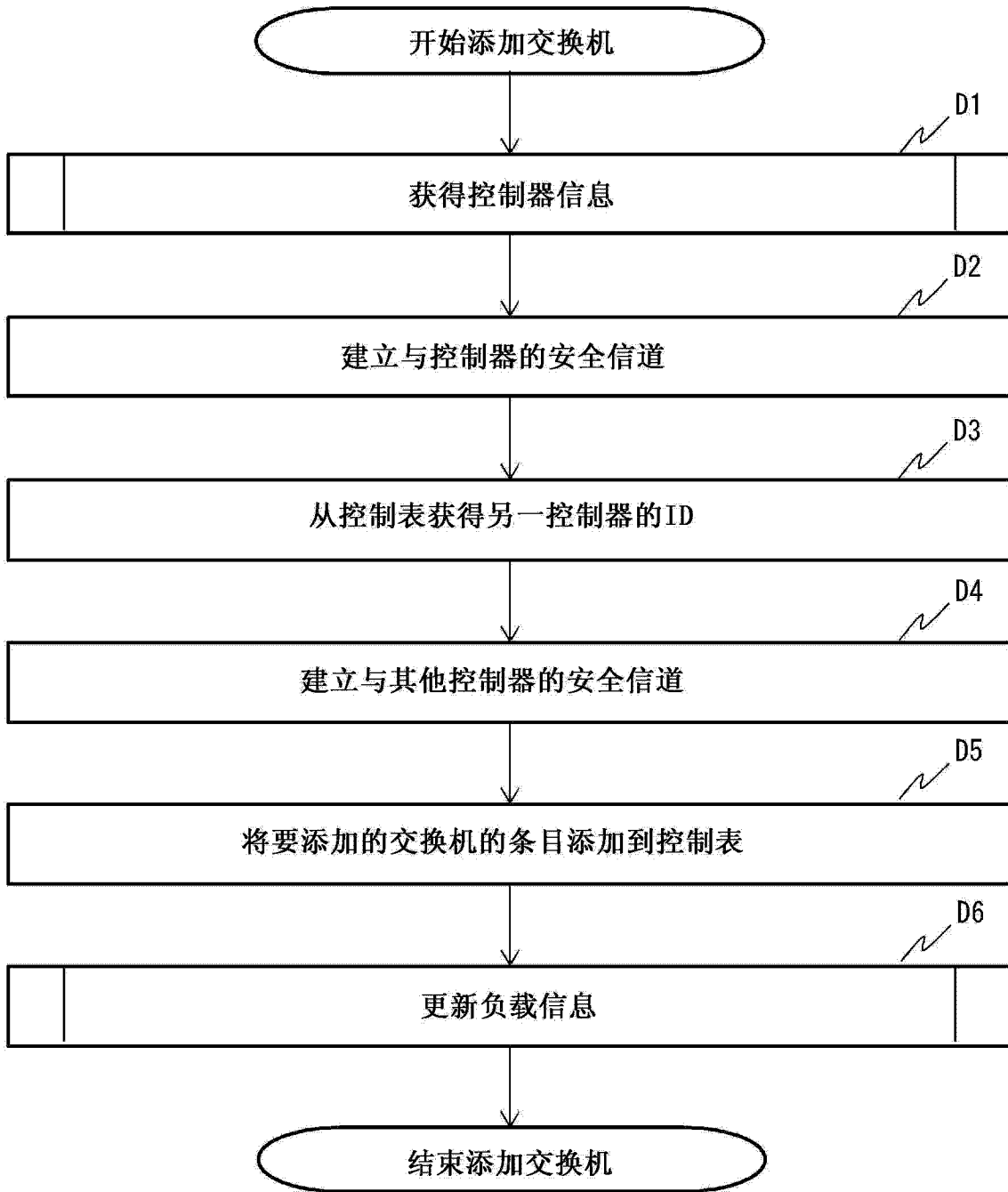


图 8



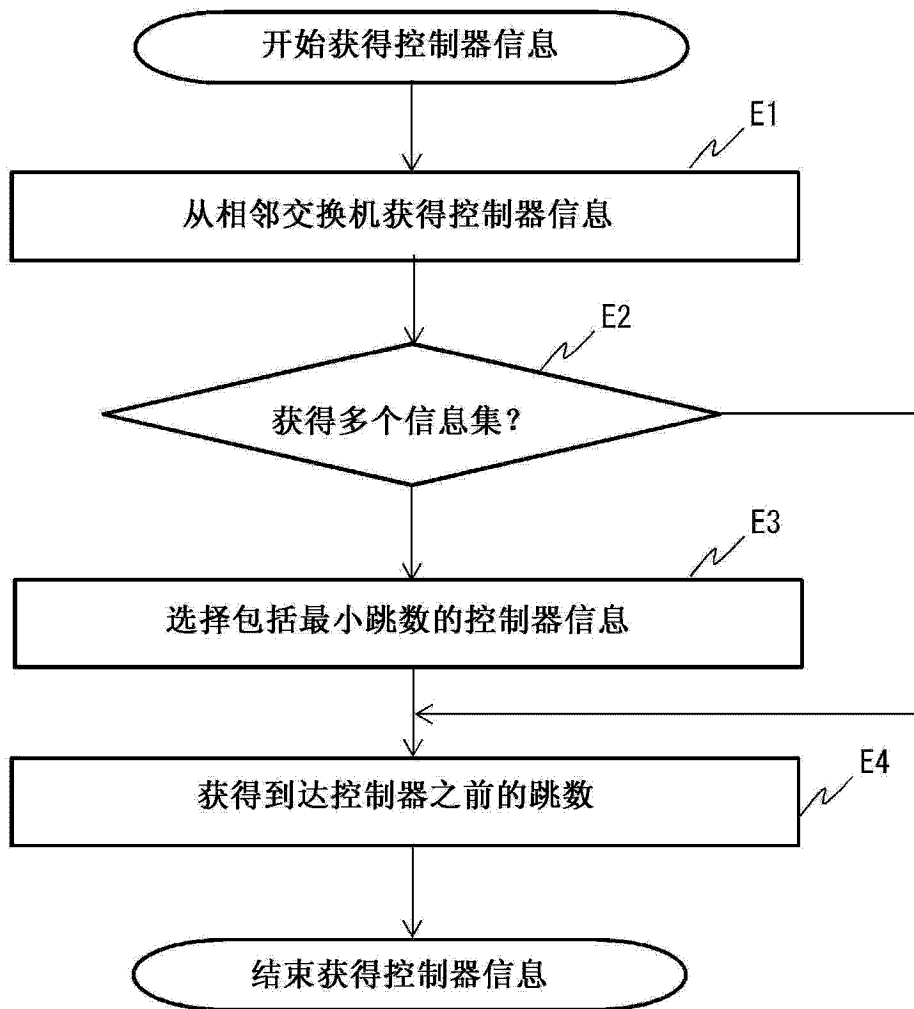


图 9

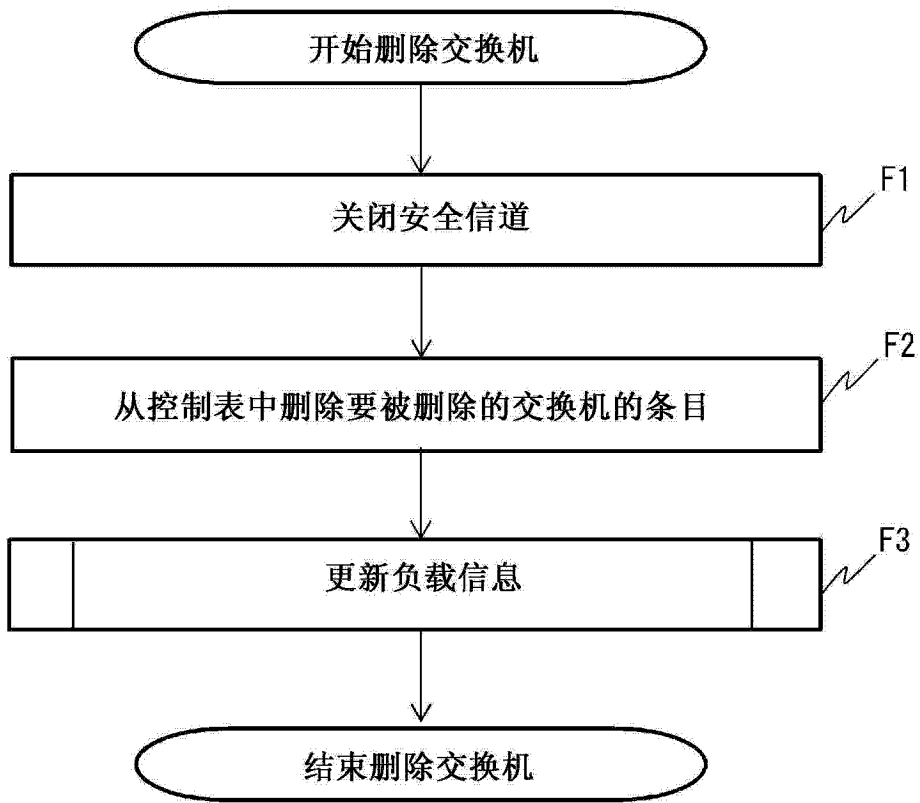


图 10

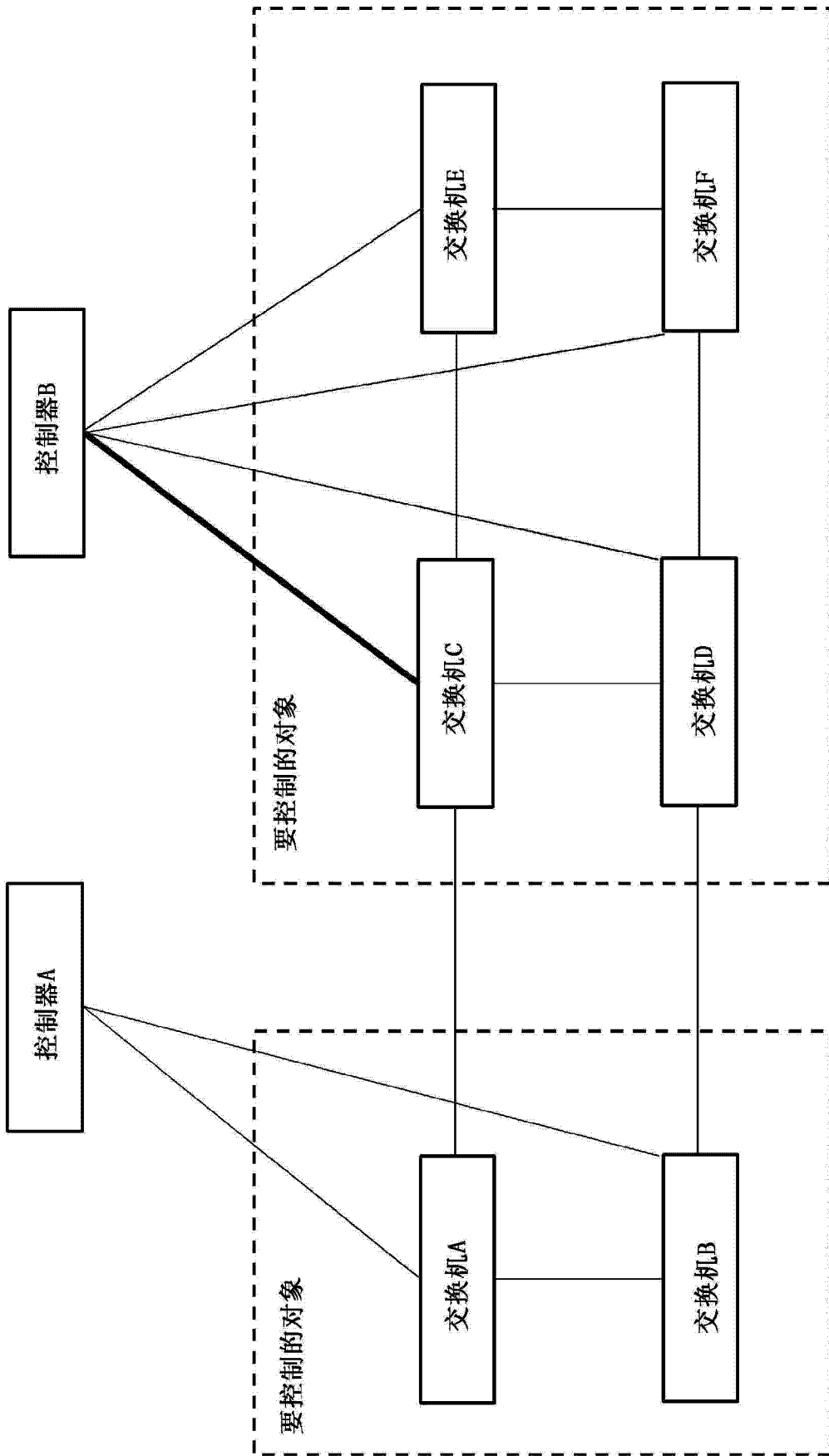


图 11

交换机ID	控制器ID	跳数	所接收的消息计数	所接收的分组计数
交换机A	控制器A	1	30	100
交换机B	控制器A	2	20	200
交换机C	控制器B	2	5	100
交换机D	控制器B	1	30	200
交换机E	控制器B	2	15	100
交换机F	控制器B	2	20	100

图 12

交换机ID	控制器ID	跳数	所接收的消息计数	所接收的分组计数
交换机A	控制器A	1	30	100
交换机B	控制器A	2	20	200
交换机C	控制器A	2	5	100
交换机D	控制器B	1	30	200
交换机E	控制器B	2	15	100
交换机F	控制器B	2	20	100

图 13

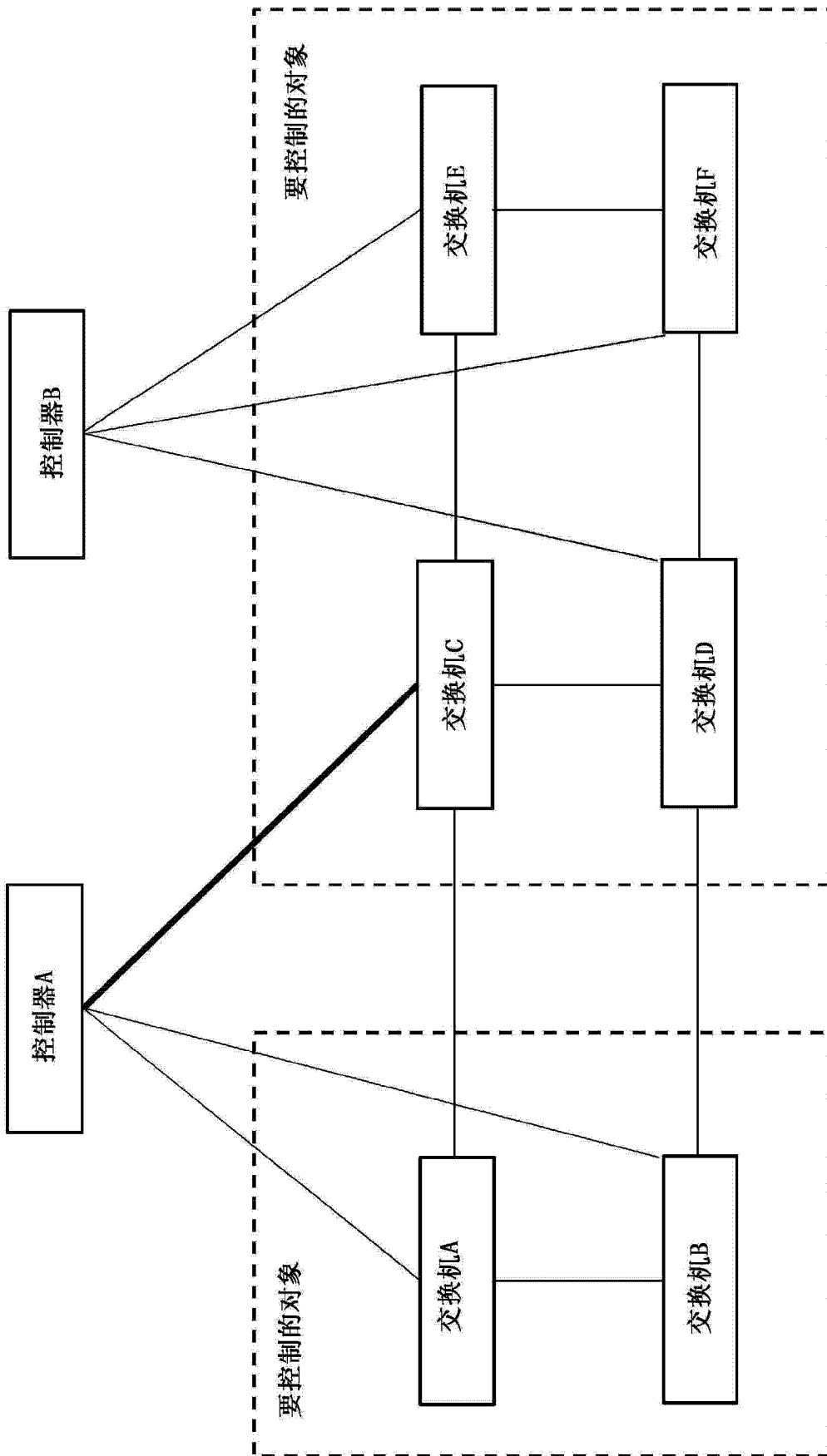


图 14

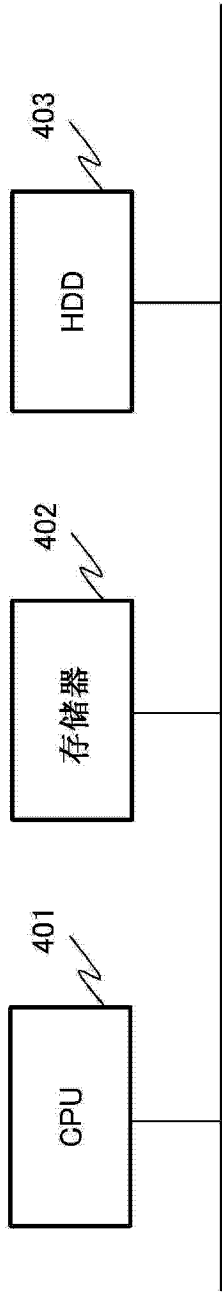


图 15

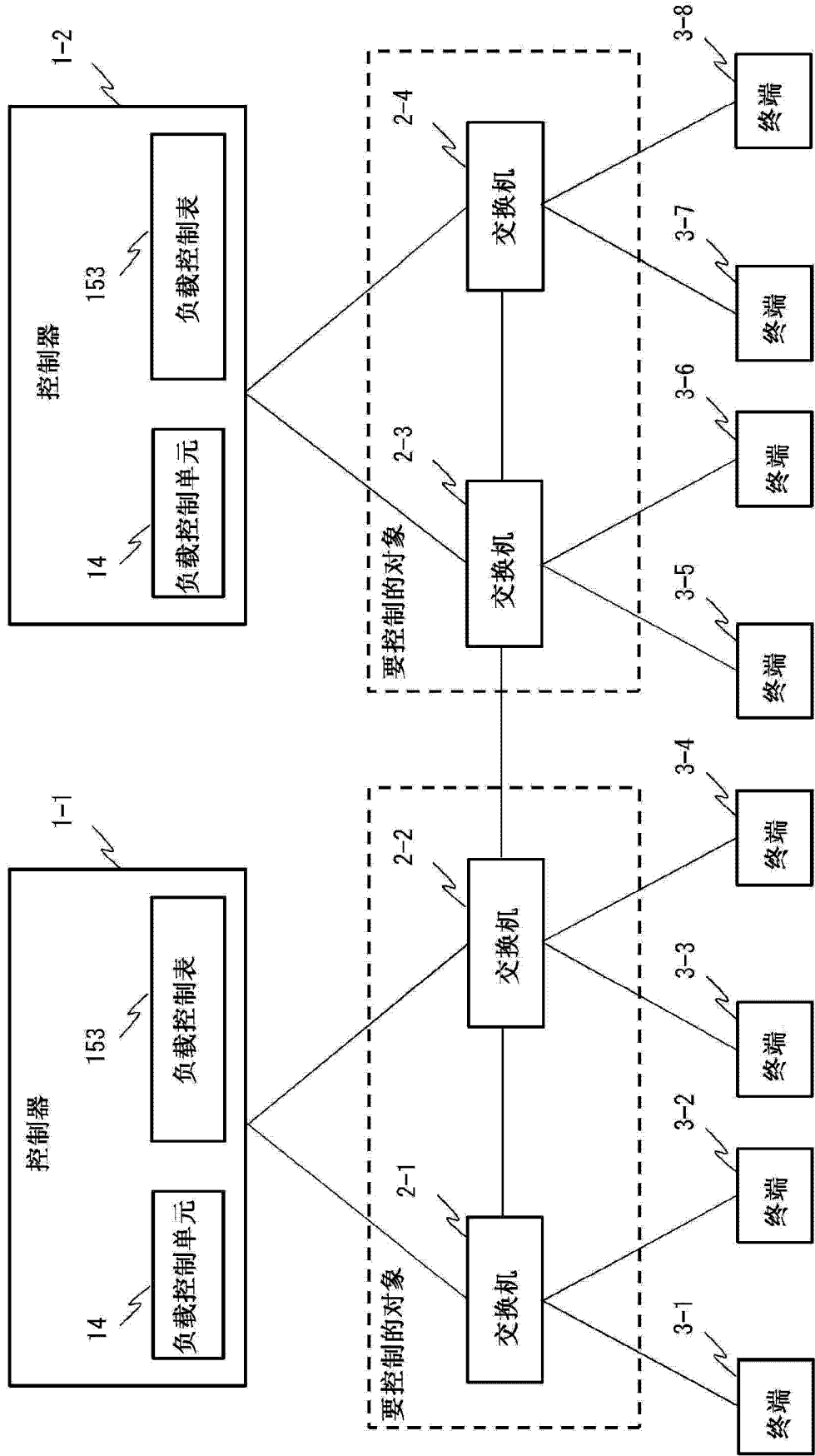


图 16