



(12)发明专利

(10)授权公告号 CN 105900403 B

(45)授权公告日 2019.10.22

(21)申请号 201580004178.6

(22)申请日 2015.01.08

(65)同一申请的已公布的文献号
申请公布号 CN 105900403 A

(43)申请公布日 2016.08.24

(30)优先权数据

61/926,099 2014.01.10 US

61/973,013 2014.03.31 US

14/530,332 2014.10.31 US

(85)PCT国际申请进入国家阶段日
2016.07.08

(86)PCT国际申请的申请数据
PCT/US2015/010638 2015.01.08

(87)PCT国际申请的公布数据
W02015/105987 EN 2015.07.16

(73)专利权人 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 李頔 贝塔·朱克 张航

(51)Int.Cl.
H04L 29/06(2006.01)

(56)对比文件

W0 2013187054 A1,2013.12.19,

US 2013283374 A1,2013.10.24,

EP 2675119 A1,2013.12.18,

审查员 朱文君

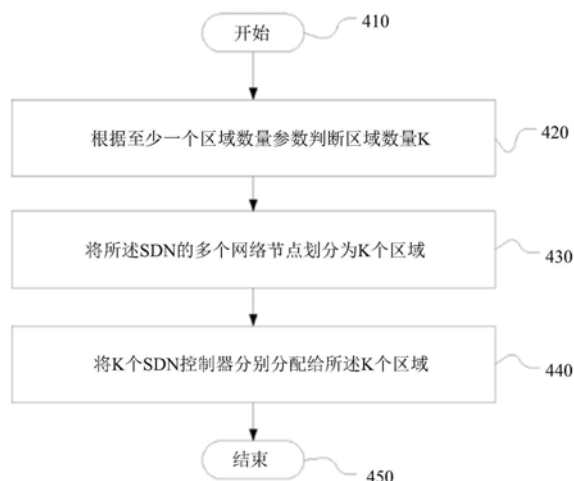
权利要求书2页 说明书11页 附图11页

(54)发明名称

软件定义网络中的分区系统和方法

(57)摘要

在一实施例中,软件定义网络(简称SDN)网络分区的方法包括:根据至少一个区域数量参数判断区域数量K;指定K,将所述SDN的多个网络节点划分为K个区域;将所述K个区域被分别分配给K个SDN控制器,其中,所述K个SDN控制器用于为所述多个网络节点中被分别分配的网络节点做出流量工程决策,并且对其进行分布式网络优化。



1. 一种有多个网络节点的软件定义网络(简称SDN)的网络分区方法,其特征在于,包括:

根据至少一个区域数量参数判断区域数量K;所述至少一个区域数量参数包括流量工程复杂度、网络规模、业务负荷,以及业务流信息中的至少一个;所述判断考虑了流量工程复杂度、网络规模、业务负荷之间的关系;

将所述SDN的所述多个网络节点划分为K个区域;所述K个区域为地理区域或逻辑区域;其中,地理分区支持通过由弧流模型制定的流量工程优化,逻辑分区支持通过由路径-流量模型制定的流量工程优化;

将K个SDN控制器分别分配给所述K个区域,其中,所述K个SDN控制器中的每个SDN控制器用于为划分到各个区域的网络节点做出流量工程决策,并且对其进行分布式网络优化;

所述方法还包括:选择K个SDN控制器的第一组,所述选择与所述分别分配同时进行或所述选择在所述划分和所述分别分配前进行;

具体地,所述选择K个SDN控制器的第一组包括根据多个候选SDN控制器相互之间的控制器间距从所述多个候选SDN控制器中选出所述K个SDN控制器的第一组,其中,所述选择还包括将所述K个SDN控制器的第一组相互之间的控制器间距最大化;为了实现所述间距最大化,当所述各个区域的网络节点密度相差较大时,K个SDN控制器的第一组相互之间的控制器间距被加权。

2. 根据权利要求1所述的方法,其特征在于,还包括:在所述分别分配的同时重新选择所述K个SDN控制器的第二组来替换所述K个SDN控制器的第一组。

3. 根据权利要求1所述的方法,其特征在于,所述相互之间的控制器间距包括跳数。

4. 根据权利要求1所述的方法,其特征在于,所述划分包括根据启发式算法将所述SDN分为地理区域,其中,所述启发式算法利用物理拓扑信息形成所述地理区域。

5. 根据权利要求4所述的方法,其特征在于,所述划分还包括根据聚类启发式算法形成所述地理区域。

6. 根据权利要求4所述的方法,其特征在于,所述划分还包括根据分区启发式算法形成所述地理区域。

7. 根据权利要求4所述的方法,其特征在于,所述划分还包括根据任务分配启发式算法形成所述地理区域。

8. 根据权利要求1所述的方法,其特征在于,所述划分包括根据启发式算法将所述SDN分为逻辑区域,其中,所述启发式算法利用数据流信息形成所述逻辑区域。

9. 根据权利要求8所述的方法,其特征在于,所述划分还包括根据任务分配启发式算法形成所述逻辑区域。

10. 根据权利要求8所述的方法,其特征在于,所述划分还包括根据图形分区启发式算法形成所述逻辑区域。

11. 根据权利要求1所述的方法,其特征在于,所述划分在所述分别分配前进行。

12. 根据权利要求1所述的方法,其特征在于,所述划分与所述分别分配同时进行。

13. 一种通信系统,其特征在于,包括:

多个网络节点,所述网络节点可被链路上的多条业务流穿越,且定义了流量工程问题;
多个控制器,可从所述多个控制器中选择区域控制器子集;

分区控制器,用于:

判断所述区域控制器子集的大小;

选择所述区域控制器子集;

启发式地将所述多个网络节点划分为区域,并将所述区域控制器子集分别分配给所述区域;具体地,位于控制面的所述区域控制器子集从所述网络节点中采集状态信息和业务需求以阐述流量工程问题;所述区域控制器子集向数据面硬件提供控制决策以引导网络的操作;

定期进行重新网络分区;

主控制器,用于将所述流量工程问题分解为与所述区域对应的流量工程子问题,其中,所述区域控制器子集用于为所述区域分别解决所述流量工程子问题;还用于控制和协调分区控制器,并对划分后的区域进行重新优化。

14. 根据权利要求13所述的通信系统,其特征在于,所述区域控制器子集用于根据所述流量工程子问题的各个解决方案为所述多个网络节点配置网络资源。

15. 根据权利要求13所述的通信系统,其特征在于,所述控制器还用于与所述区域控制器子集通信以调整所述流量工程子问题的状态。

软件定义网络中的分区系统和方法

[0001] 本申请要求于2014年10月31日由Li等人递交的发明名称为“软件定义网络中的分区系统和方法”的第14/530,332号美国专利申请案的在先申请优先权,该在先申请要求于2014年3月31日由Li等人递交的发明名称为“为软件定义网络中的分级流量工程生成逻辑区域”的第61/973,013号美国临时申请案,以及2014年1月10日由Li等人递交的发明名称为“软件定义网络中的分区系统和方法”的第61/926,099号美国临时申请案的在先申请的权益,这些在先申请的内容以引入的方式并入本文。

[0002] 相关申请案交叉申请

[0003] 本发明涉及于2013年12月27日递交的已转让给本发明受让人的,发明名称为“SDN中分布式流量工程的流量行为驱使动态分区”的第14/142,344号共同未决美国专利申请案,以及涉及于2013年12月26日递交的已转让给本发明受让人的,发明名称为“分级软件定义网络流量工程控制器”的第14/141,077号共同未决美国专利申请案,这些在先申请的内容以引入的方式并入本文。

技术领域

[0004] 本发明通常涉及软件定义网络(简称SDN)中的分区方法,在特定的实施例中,涉及一种分区的系统和方法。

背景技术

[0005] 受智能手机、平板电脑和视频流业务的驱动,无线网络处理的无线数据量显著增加并预计将持续增长。除了数据量的增长,连接无线网络的设备的数量也将持续以指数级增长,有望达到数十亿台设备。不同的应用将对未来无线网络的性能提出不同的需求。预计未来的无线网络非常灵活、高效、开放,且可为用户和消费者进行定制。

发明内容

[0006] 在一实施例中,一种有多个网络节点的软件定义网络(简称SDN)网络分区的方法包括:根据至少一个区域数量参数判断区域数量K。指定K,所述SDN的所述多个网络节点被划分为K个区域。所述K个区域被分别分配给K个SDN控制器。所述K个SDN控制器中的每个SDN控制器用于为划分到各自区域的网络节点做出流量工程决策,并对其进行分布式网络优化。

[0007] 在一实施例中,通信系统包括多个网络节点,多个控制器,分区控制器,和主控制器。所述多个网络节点可被链路上的多条业务流穿越。针对所述业务流、链路,以及网络节点定义了流量工程(简称TE)问题。所述分区控制器用于判断能从所述多个控制器中选出的区域控制器子集的大小。所述分区控制器还用于选择所述区域控制器子集。所述分区控制器还用于启发式地将所述多个网络节点划分为区域,并将所述区域控制器子集分别分配给所述区域。所述主控制器用于将所述TE问题分解为与所述区域对应的TE子问题。所述区域控制器子集用于为所述区域分别解决所述TE子问题。

[0008] 在一实施例中,分区控制器包括网络接口控制器(简称NIC)和处理器。所述NIC能通过控制面与多个SDN控制器以及多个网络节点耦合。所述多个网络节点可通过数据面被链路上的业务流穿越。可针对所述多个网络节点、所述业务流,和所述链路定义TE问题。所述处理器用于根据预期的TE复杂度判断区域控制器子集的大小,判断包括所述多个网络节点的网络的大小和包括所述业务流的业务负荷。所述处理器还用于根据控制器拓扑、网络拓扑,和所述业务负荷从所述多个SDN控制器中选出所述区域控制器子集,所述处理器还用于将所述多个网络节点划分为区域,并将所述区域控制器子集分别分配给所述区域,其中,在一些实施例中,所述分区启发式地完成。

附图说明

[0009] 为了更完整地理解本发明及其优点,现在参考下文结合附图进行的描述,其中:

[0010] 图1为一实施例的分区网络图;

[0011] 图2为一实施例的拓扑图;

[0012] 图3为一实施例的流向图;

[0013] 图4为一实施例的SDN分区方法的流向图;

[0014] 图5为一实施例的运算系统的框图;

[0015] 图6为另一实施例的拓扑图;

[0016] 图7示出了利用聚类启发式方法的地理分区结果;

[0017] 图8示出了利用划分启发式方法的地理分区结果;

[0018] 图9示出了利用任务分配启发式方法的地理分区结果;

[0019] 图10示出了基于区域的TE优化问题;

[0020] 图11示出了流向图构造。

具体实施方式

[0021] 下文将详细论述实施例的制作和使用。但应了解,本发明提供的许多适用发明概念可实施在多种具体环境中。所论述的具体实施例仅仅说明用以实施和使用本发明的具体方式,而不限本发明的范围。

[0022] 未来的无线网络有望满足对更高吞吐量,更低时延,更低耗能,更低成本,以及更多连接的需求。网络功能虚拟化(简称NFV)以及软件定义网络(简称SDN)等技术在建立未来无线网络中变得越来越重要。NFV使得传统意义上绑定专用硬件的网络功能可以在数据中心的云计算基础架构中运行。数据面、控制面和管理面网络功能从专用硬件基础设施中的分离很有可能成为未来无线网络架构的共同特征。一大益处在于提供弹性支持网络功能要求的能力。SDN是创建智能可编程网络的架构框架,其中,控制面以及数据面是分离的,网络智能与状态逻辑集中,且底层网络架构是从应用中分离出来的。

[0023] 流量工程(简称TE)问题为多种业务流判断路由路径,并根据各自的服务质量(简称QoS)需求沿所述路径分配资源,比如带宽,以实现网络效用最大化。随着网络规模的增长,TE问题只会变得越来越复杂。SDN的特征在于无线网络能被动态地分为区域以提升性能。更具体地说,应意识到,集中控制会因网络的规模和负荷变得不切实际或不可行,而当将无线网络分为区域,即将网络分区,能够实现TE问题的分布式优化。区域包括其区域控制

器,即SDN控制器,以及其关联的网元,包括区域的节点和链路。还应意识到,无线网络能被分为地理区域或逻辑区域。地理区域是根据地区的考量因素,如用户设备(简称UE)、无线节点和路由器的位置和密度等形成的。例如,若单个SDN控制器为10英里范围内的所有网元提供服务,可将此10英里范围定义为一个区域。或者,逻辑网络区域根据网络业务流形成。例如,与特定的业务流关联的网元能被分给单个逻辑区域和单个SDN控制器。地理分区支持通过由弧流模型制定的TE优化。逻辑分区支持通过由路径-流量模型制定的TE优化。还应意识到,还可以通过主区域控制器或主控制器来控制 and 协调本地区域控制器。随着流量波动及用户从一个区域迁移到另一个区域,无线网络可通过分区控制器定期进行重新分区,且可通过主区域控制器进行重新优化。

[0024] 分区问题的优化是为了在将网元分为区域与将分布式的、基于区域的TE优化的开销最小化之间实现平衡。

[0025] 通过TE问题的分布式优化,TE问题能被分解为有待本地区域控制器,即SDN控制器,解决的针对多种区域的子问题。位于控制面的SDN控制器从网元中采集状态信息和业务需求以阐述TE问题。SDN控制器向数据面硬件提供控制决策以引导网络的操作。

[0026] 控制面开销有若干来源。这些来源包括TE问题初始化、网络状态报告、网络配置指令,以及SDN控制器间的交互。TE问题初始化开销是由于主控制器分配流信息,如来源、目的地、需求等来选择区域控制器而造成的。网络状态与节点附带的链路的状态以及节点的状态相对应。连接同一区域内的节点的链路为内部链路,连接不同区域的节点的链路为边界链路。将SDN控制器分配给不同的区域可以将网络状态报告以及网络配置限定于一个指定的区域,但也会在协调多个区域时产生间接成本。非最佳少量区域在削减协调开销的同时会增加各个SDN控制器上的网络状态报告和网络配置负荷;另外,非最佳大量区域在削减各个SDN控制器的网络状态报告和网络配置负荷的同时会增加协调开销。所以,判断区域的数量是分区优化中的平衡考验。区域的数量判断由分区控制器根据预期的区域TE复杂度进行,区域TE复杂度通常体现为时间、网络规模,以及业务负荷。预期的区域TE复杂度的一大要素在于特定的SDN控制器的处理能力。预期的区域TE复杂度的另一大要素在于执行TE优化的快慢程度。TE问题的解决方法可在由业务负荷和吞吐量决定的(至少是部分决定的)一段时间内得到。如果不能足够快地解决TE问题,不知如何进行业务路由的节点会造成拥塞。相应地,如果特定的SDN控制器的处理能力不足以解决其预期的TE子问题,应该增加区域的数量,以进一步分散处理负荷。可根据经验并通过多种系统测量值和数据点来观察TE复杂度、网络规模和业务负荷这三者之间的关系。例如,可以在管理信息库(简称MIB)中集中化这些测量值和数据点。一旦得知请求的区域的数量,可以将网络分为,或划分为区域。区域有时也被称为块,或分区,可以在特定的实施例中重叠,而在其他的实施例中不重叠。根据划分问题或分区问题的解决办法进行所述划分。能通过一些路径找到划分问题的解决方法,其中,一些方法用于地理区域,而其它的方法用于逻辑区域。

[0027] 无线网络可包括许多SDN控制器。事实上,任何网元都能成为SDN控制器。在对无线网络进行分区时,分区控制器是从网络中可用的SDN控制器中选出的,所以一些SDN控制器可能没有被选中。例如,没有与其他任何网元或业务流相关联的SDN控制器可能会落选。同样地,一些网元可能与任何SDN控制器都不关联。被选为分区控制器的SDN控制器的数量取决于区域的数量。例如,可以配置网络使其为给每个区域分配一个SDN控制器。区域的数量

通常不超过可用的SDN控制器的数量。

[0028] 选择区域控制器的目的在于使控制器最小间距最大化,控制器最小间距通常用跳数来表示。在可替代实施例中,可使用其他标准来衡量该距离,如欧氏距离、通信时延,以及其他本领域技术人员可理解的衡量方法。选择分区控制器也是为了实现控制器的统一分配,这有利于控制器负荷平衡。控制器间距会受到指定地区中节点的密度的影响。在密集地区放置许多SDN控制器虽然更有利于获得小的控制器间距,但有悖于将控制器最小间距最大化的目标。为了平衡此趋势,可根据指定的节点分配状况下特定SDN控制器的集中或偏远程度来为一个SDN控制器的控制器间距进行加权。当SDN控制器的数量较小时,可以通过穷举搜索来进行选择,特别是要评估SDN控制器每一种可能的组合否则,启发式地进行选择。一旦SDN控制器被选出,每个SDN控制器被分派给一个区域,以便高效地为该SDN控制器分派解决其TE子问题的任务。进行SDN控制器的分派以便寻找分派问题的解决方法。

[0029] 对于有N个节点的网络来说,所述N个节点通过将其与其他节点相连接的链路转发数据面的数据包。节点的子集S用做控制器。分区过程通过将数据面节点的子集和控制器节点耦合来创建|S|区域。每个数据面节点仅与一个控制器相关联,该控制器采集其网络状态,并为其提供转发表。网络状态与节点附带的链路的状态相对应控制器在控制面进行交互,并参与分布式网络优化(简称DNO)。控制器合力解决DNO问题,DNO问题可被分解为与每个区域和协调主问题相关的多个子问题。分布式优化可迭代进行。在每次迭代时,可根据针对区域采集到的网络状态信息和使用边界链路的开销成本来解决区域子问题,以得到针对区域的资源分配。可利用解决主TE问题的主SDN控制器来协调针对边界链路的资源分配。

[0030] 针对地理分区,划分问题可用以下公式加以阐述(其中,未考虑TE问题初始化开销):

$$[0031] \quad \min \sum_{s \in S} (d_s^{coll} + d_s^{coor} + d_s^{prov})$$

[0032] 则:

$$[0033] \quad d_s^{coll} = \sum_{i \in N} a_{is} h_{is} \left(\alpha + \beta \sum_{j: j \neq i, j \in N} l_{ij} \right), \quad \forall s \in S$$

$$[0034] \quad d_s^{coor} = \eta \gamma \sum_{i \in N} \sum_{j: j \neq i, j \in N} 2(h_{is} + h_{sm}) l_{ij} (a_{is} + a_{js} - 2a_{is} a_{js}), \quad \forall s \in S$$

$$[0035] \quad d_s^{prov} = \sum_{i \in N} a_{is} h_{si} \left(\mu + v \sum_{j: j \neq i, j \in N} l_{ij} \right), \quad \forall s \in S$$

$$[0036] \quad -\tau \leq \sum_{i \in N} (a_{is} - a_{is'}) \left(1 + \sum_{j: j \neq i, j \in N} l_{ij} \right) \leq \tau, \quad \forall s, s' \in S$$

$$[0037] \quad a_{is} \in \{0, 1\}, \quad \forall i \in N, \forall s \in S$$

$$[0038] \quad \sum_{s \in S} a_{is} = 1, \quad \forall i \in N$$

[0039] 其中:

[0040] N是网络中节点的集合;

[0041] S 是控制器的集合,且 $S \subseteq N$;

[0042] m 是主控制器,且 $m \in S$;

[0043] α 是每个节点报告的数据单元数;

[0044] β 是每个链路报告的数据单元数;

[0045] γ 是每个链路的辅助数据单元数;

[0046] η 是分布式网络优化迭代的数量;

[0047] μ 是每个节点配置的数据单元数;

[0048] ν 是每个链路配置的数据单元数;

[0049] τ 是区域失衡容忍因子;

[0050] l_{ij} 是 i 和 j 的连接指示器,其中, $\forall i, j \in N$;

[0051] h_{ij} 是 i 到 j 的跳距,其中, $\forall i, j \in N$ 。

[0052] 公式的第一条件规定了每区采集开销 d_s^{coll} 。每个节点将关于自身及其附带的链路的信息报告给其控制器。所述报告在各个区域中进行,传输的比特总量被算作采集开销。第二条件描述了每区协调开销 d_s^{coord} ,其包括每个从控制器报告给主控制器并从主控制器中接收的比特总量。它们中的每一个都申请了各个区域共享的链路的数量和协调迭代的数量。第三条件规定了每区配置开销, d_s^{prov} ,其计算方法与每区采集开销类似。针对节点和链路进行配置。每条链路配置两次:每个端节点各配置一次。第四条件是区域平衡约束,其中,选择 τ 来确保问题的可行性。第五条件限制针对节点-控制器关联的决策必须为二元的。最后的限制条件指示每个节点关联一个控制器。 a_{is} 是表示 i 是否与 s 关联的决策变量。

[0053] 如上所述,有数种启发式方法可解决地理区域的划分,即分区,问题,包括聚类、划分,以及任务分配。聚类将每一个节点与基于跳距或采用指定的距离衡量标准时最近的控制器相关联。分区过程将网络划分为 S 个连接的组件,接着分配控制器。任务分配应用了MinMin原理。这些启发式方法利用物理拓扑信息形成地理区域。物理拓扑信息包括控制面拓扑和数据面拓扑。控制面拓扑信息包括关于SDN控制器、路由器、UE、SDN控制器间的连接,各个网络组件/SDN控制器对间的连接,每个UE/SDN控制器对间的连接,以及与节点和连接有关的参数的信息。参数可包括连接成本、节点能力等。连接可以是实际的,也可以是虚拟的。数据面拓扑信息包括以下信息:路由器;UE;物理链路,包括有线和无线的物理链路;以及与节点和链路有关的参数。这些参数可以包括引入到TE问题的计算负荷。

[0054] 通过聚类,控制器首先形成一个区域,该区域通过吸收拥有通向该区域的一个内部节点的链路的外部节点而不断扩大其范围。当节点能被多个区域吸收时,该节点被跳距最近的区域吸收。当多个区域与节点和控制器间的距离相同时,该决策被随机作出。在本方法中,控制器在地理上一般包含在区域中。

[0055] 通过划分,网络能被建模为利用最小总链路权重的链路集划分的 k 个大小平衡的相连接的组件的网络图形。这有时被称为平衡最小 k 链路集问题。划分启发法基于对网络图形的平衡最小 k 链路集的发现,其中 $k = |S|$ 。所述方法分两步进行。首先是解决自然加权的网络图形的平衡最小 k 链路集问题,其中,每个节点 i 通过 $\alpha + \beta \sum_{j \in N} l_{ij}$ 加权,并且所有链路都有相等的加权 γ 。通过采用已知的启发法例如多级划分来寻找解决方法。此时,控制器被分派给每个分区以实现将节点和控制器间以及主、从控制器间的总通信成本最小化的目标。

[0056] 分派问题被建模为任务分配问题。假定 Π 代表包含在划分解决方法步骤一中的分

区的集合。 Π 包含分区和控制器间的匹配数字,即 $|\Pi| = |S|$ 。假定 $B(P)$ 为分区 $P' \in \Pi$ 的边界,所述分区包括边的集合,所述边的一端为 P ,另一端为 P' ,且 $P' \neq P$ 。对于每个表示为 e^P 的边 $e \in P$,端顶点位于 P 上。控制器分派问题可用公式表示为:

$$\min \sum_{s \in S} \sum_{P \in \Pi} c_{s,P} (d_{s,P}^{\text{coll}} + d_{s,P}^{\text{coord}} + d_{s,P}^{\text{prov}})$$

[0057] 则:

$$[0058] \quad d_{s,P}^{\text{coll}} = \sum_{i \in P} h_{is} \left(\alpha + \beta \sum_{j: i \neq j, j \in N} l_{ij} \right), \forall s \in S, P \in \Pi$$

$$[0059] \quad d_{s,P}^{\text{coord}} = \sum_{e \in B(P)} 2\eta\gamma (h_{e^P_s} + h_{sm}), \forall s \in S, P \in \Pi$$

$$[0060] \quad d_{s,P}^{\text{prov}} = \sum_{i \in P} h_{si} \left(\mu + v \sum_{j: i \neq j, j \in N} l_{ij} \right), \forall s \in S, P \in \Pi$$

$$[0061] \quad c_{s,P} \in \{0, 1\}, \forall s \in S, P \in \Pi$$

$$[0062] \quad \sum_{s \in S} c_{s,P} = 1, \forall P \in \Pi$$

$$[0063] \quad \sum_{P \in \Pi} c_{s,P} \leq 1, \forall s \in S$$

[0064] 其中, $c_{s,P}$ 为决策变量,表示控制器是否被分派给分区 P 。前三个条件规定了由于分配各个控制器而给各分区而产生的采集开销、协调开销,以及配置开销。第四个限制条件要求决策必须为二元的。第五个限制条件强制只为分区分派一个控制器。第六个限制条件要求每个控制器不得与超过一个分区相连。第六个限制条件允许控制器多于分区,以提供灵活性。控制器并不一定要包含在为其分配的地理区域中。

[0065] 通过任务分配,分区问题被结构化,使得每个节点成为分配给一个控制器的任务,以将总成本最小化。分区配置 Z 的成本 $C(Z)$ 由以下公式表示:

$$[0066] \quad C(Z) = \sum_{s \in S} \left(\omega (d_s^{\text{coll}}(Z) + d_s^{\text{coord}}(Z) + d_s^{\text{prov}}(Z)) + (1 - \omega)p(Z) \right)$$

[0067] 其中, $d_s^{\text{coll}}(Z)$, $d_s^{\text{coord}}(Z)$, 和 $d_s^{\text{prov}}(Z)$ 是分区内信息采集成本、主-从协调成本,和配置成本。上述成本根据上述原始问题阐述计算。 p 是参考原始问题的第四分区大小平衡限制计算的分区失衡成本, ω 是加权因子。

[0068] 根据分配启发法, Z_s 是分配给控制器 s 的区域,其中, $s \in S$ 。在初始的分区配置 Z^0 中,每个控制器被分给将其当作自身一员的区域。所有非控制器节点都属于非关联节点集合 U 。参考区域 Z_s 定义 U 的前端 F_s ,并对所有区域重复这一操作。前端 F_s 包括 U 中的节点的集合,所述 U 拥有通向 Z_s 中的节点的链路。在每次迭代时,前端节点 i 从 U 移动到边界区域 Z_s 。所述移动分两步实现。第一步,针对每个前端 F_s 中的每个节点 i ,给每个边界区域 Z_s 和前端节点 i 计算成本增幅 $\Delta(i, s)$ 。根据等式 $\Delta(i, s) = C(Z^t(i, s)) - C(Z^t)$ 并参考前一个分区配置 Z^t 来计算

新的分区配置 $Z^t(i, s)$ 的成本。在第二步中,针对前端节点识别最小成本增幅,并移动该节点。当集合 U 为一个空集时,该步骤停止。

[0069] 逻辑区域是成员流的候选路由路径的集合。由于流路径间的链路共享,多个逻辑区域可以在物理上重叠。如上所述,解决逻辑区域的分区问题有若干种方法,包括任务分配启发法和图形划分启发法。启发算法利用数据流信息形成逻辑区域。数据流信息包括每条流的候选路由路径和与每条流关联的参数。参数可以包括TE问题采用的计算负荷,相关虚拟网络等。图形划分方法利用将流间的链路共享建模为流向图,并通过图形原理来解决该问题。针对逻辑区域,网络被建模为拥有节点集合 N 和链路集合 L 。一个子集的节点 S 被配置成控制器,其中之一为主控制器 m 。TE问题用于解决一个集合的业务流 F 。假设各条流的路由路径已知。分区过程通过将流子集与控制器节点进行耦合来创建 k 个区域。每条流仅与一个控制器关联,该控制器采集其路由路径上的节点和链路的状态。该控制器将流来源的TE决策通知给流来源。逻辑区域的大小被定义为区域中的流的候选路径的总数。建立逻辑区域的目的在于将开销最小化和平衡区域规模。

[0070] 用流控制器关联矩阵 $Z = [a_{fs}]_{|F| \times |S|}$ 来描述针对逻辑分区的分区配置 Z ,其中,每一列的矢量 Z_s 定义被从SDN控制器控制的区域 s 。对于流 $f \in F$ 和从控制器 $s \in S$ 来说,若 f 与 s 关联, $a_{fs} = 1$;否则, $a_{fs} = 0$ 。根据定义,区域 Z_c 的大小 $|Z_s|$ 被计算为 $|Z_s| = \sum_{f \in F} a_{fs} |f|$,其中, $|f|$ 是流 f 的大小,也就是针对流 f 的候选路径的数量。

[0071] 指定区域配置为 Z 、包括初始化成本 C^{init} 的开销成本为 $C(Z)$ 、采集成本为 $(Z) C^{coll}(Z)$ 、配置成本为 $C^{prov}(Z)$,且协调成本为 $C^{coord}(Z)$,则: $C(Z) = C^{init}(Z) + C^{coll}(Z) + C^{prov}(Z) + C^{coord}(Z)$ 。协调成本在优化迭代的过程中产生,而其他三项成本在优化的开始或结尾时产生。

[0072] 初始化成本是由于主控制器为从控制器分配流信息,包括流来源、目的地、需求等而产生的。分配给指定从控制器的信息限制为与从控制器关联的流的信息。初始化成本由 $C^{init}(Z_s) = \mu \sum_{f \in F} a_{fs} h_{ms}$ 得出,其中 μ 是每流描述的数据单元的数量, h_{ms} 是针对主控制器到从控制器的通信时的每个数据单元的开销成本。

[0073] 采集成本是由于每个从控制器采集自身区域中的节点和链路而产生的。采集成本由 $C^{coll}(Z) = \alpha \sum_{i \in N} u_i^s h_{is} + \beta \sum_{l \in L} v_l^s h_{l_{min}s}$ 得出,其中, $u_i^s = \max_{f \in F} (\delta_i^f a_{fs})$ 和

$v_l^s = \max_{f \in F} (\delta_l^f a_{fs})$ 分别表示节点 i 或链路 l 是否与从控制器 s 相关联。 α 是每个节点状态报告的数据单元的数量, β 是每个链路状态报告的数据单元的数量, h_{is} 是每个针对端节点到从控制器的通信的数据单元的开销成本, $h_{l_{min}s}$ 是每个针对端节点和从控制器间的链路的数据单元的开销成本。

[0074] 配置成本是由于从控制器在其区域内分配TE决策而产生的。配置成本由 $C^{prov}(Z) = v \sum_{f \in F} a_{fs} h_{sfsrc}$ 得出,其中, v 是每条流路径中的TE决策中的数据单元的数量, h_{sfsrc} 是每个针对从从控制器到流的来源节点的通信的数据单元的开销成本。

[0075] 协调成本是由于主控制器采集与更新针对从控制器的TE子问题的状态而产生的。交换的数据量与区域共享的链路的数量成正比。协调成本由

$C^{coord}(Z) = \eta \gamma (h_{sm} + h_{ms}) \sum_{f, g \in F, f \neq g} n_{fg} b_{fg}^s$ 得出,其中, $b_{fg}^s = a_{fs} + a_{gs} - 2a_{fs}a_{gs}$ 表示 f 和

g 之一是否属于区域 Z_s , η 是TE优化迭代的数量, γ 是每个区域间共享的链路的数据单元的数量, h_{sm} 和 h_{ms} 是针对主控制器和从控制器间的通信的开销成本, n_f 是 f 和 g 共享的链路的数量, 其中, $f \in F$ 且 $g \in F$ 。

[0076] 如上所述, 划分问题是对开销成本 $C(Z)$ 最小化的实践, 则:

$$[0077] \quad |Z_s| - |Z_{s'}| \leq \epsilon, \quad \forall s, s' \in S: s \neq s',$$

$$[0078] \quad a_{fs} \in \{0, 1\}, \quad \forall f \in F, \forall s \in S,$$

$$[0079] \quad \sum_{s \in S} a_{fs} = 1, \quad \forall f \in F,$$

[0080] 其中, ϵ 是负荷失衡容忍因子。

[0081] 为解决利用任务分配方法划分逻辑区域的划分问题, 指定失衡成本因子(简称ICF)为 ω , 通过移除规模平衡限制和增加惩罚项 $p(Z)$, 分区问题得到了缓解。新的成本等式和惩罚项由以下公式得出:

$$[0082] \quad C'(Z) = (1 - \omega) C(Z) + \omega p(Z)$$

$$[0083] \quad p(Z) = \max_{s, s' \in S: s \neq s'} (|Z_s| - |Z_{s'}|)$$

[0084] 相应地, 分区问题变为:

$$[0085] \quad \min_Z C'(Z)$$

[0086] 则: $a_{fs} \in \{0, 1\}, \forall f \in F, \forall s \in S$, 且 $\sum_{s \in S} a_{fs} = 1, \forall f \in F$ 。通过使用任务分配方法, 每条流, 即任务, 被分配给一个控制器, 即处理器, 从而将目标函数值最小化。例如MinMin启发法等算法, 迭代并递增地建立区域配置 Z 。

[0087] 在任务分配方法中, U 代表没有为其做分配决策的流的集合。首先, 区域配置 Z 是一个空集, 且 $U = F$ 。每次迭代都会为一条流做分配决策, 即给从控制器 s 分配流 f 。通过设置 $a_{fs} = 1$ 来表示。这时, 流从 U 中移出。

[0088] 在图形划分启发式方法中, 根据流的来源和目的地将流进行聚合, 从而将具有同一来源和同一目的地的流视为单流。然后, 通过将每条流表示成顶点, 并针对每个顶点共享的网络链路将顶点的子集与边相连, 以构造流向图。给每个顶点加权 $t \times |f|$, 其中, t 是原始流的数量, 即在聚合前顶点代表的流的数量, $|f|$ 为流的大小。也根据所通过的不同的流路径的数量为边加权。所得的流向图是超图, 其中, 每条边意味着链路共享并仅与一条网络链路相对应。由于 $k = |S|$, 因此流向图被划分为 k 个连接的组件, 每个组件定义一个逻辑区域, 且仅分配一个控制器给该组件。

[0089] 图1示出了一实施例的分区网络100。分区网络100包括: 四个控制器, 即控制器110-1到110-4; 主控制器120; 分区控制器150; 多个网节点140。多个网节点140分布在分区网络100的各处且分布方式多样。网络节点有时被称为接入点、基站(简称NodeB), 或演进型基站(简称eNB)。控制器110-1到110-4和主控制器120也是有附加能力和配置, 以用作控制器的网络节点。一般来说, 任何网络节点都能体现控制器。控制器110-1到110-4有时被称为SDN控制器、区域控制器, 或从控制器。主控制器120有时也被称为SDN控制器。

[0090] 分区控制器150将分区网络100划分为四个区域, 即区域130-1到130-4。为四个区域中的每个区域分配控制器110-1到110-4中的一个。为区域130-1分配控制器110-1, 以便

为其中的三个网络节点服务。为区域130-2分配控制器110-2,以便为其中的四个网络节点服务。为区域130-3分配控制器110-3,以便为其中的四个网络节点服务。为区域130-4分配控制器110-4,以便为其中的三个网络节点服务。主控制器120在控制器110-1到110-4之间协调TE优化。主控制器120和分区控制器150是单独的逻辑网元。主控制器120和分区控制器150可体现在同一计算系统或分开的计算系统中。

[0091] 图2为一实施例的拓扑图200。拓扑图200包括400个网络节点220,其中10个为控制器210。在拓扑图200中,网络节点220用白色圆圈表示,控制器210用黑色圆圈表示。任意控制器210也都能被指定为主控制器。网络节点220与代表链路的边230,或跳相连接。

[0092] 图3为一实施例的流向图300。流向图300包括用白框表示的业务流310,用灰色圆圈表示的超图边320,以及线330。业务流310有各自的加权。用业务流的流大小和业务流的流数量的乘积来计算业务流的加权。流大小是指针对该流的候选路径的数量。流数量指未被聚合的流的数量,也就是把来源和目的地相同的流当作一条聚合流。例如,在流向图300中,行数较低的流的加权从左到右依次为:4.0、5.0、10.0、8.0、9.0和1.0。超图边320有时被简称为边,表示物理链路。线330用多种方式连接业务流310和超图边320,表示特定的业务流共享物理链路。同样给各个超图边320分配加权,用以表示经过的流路径的数量。

[0093] 图4为一实施例的SDN分区方法的流向图。SDN包含多个网络节点和多个SDN控制器。该方法从起始步骤410开始。在区域数量步骤420中,决定网络需被划分为的区域数。根据至少一个区域数量参数判断该数量K。区域数量参数包括预期区域TE复杂度、网络规模、业务负荷,以及流信息。更具体地,该判断考虑了TE复杂度和网络规模、业务负荷,以及流信息三者间的关系。TE复杂度是控制器处理能力的功能,也是对于执行TE优化的快慢程度的操作需求。

[0094] 在特定的实施例中,该方法还包括一个选择的步骤,在此步骤中,可以从可用的SDN控制器中选出至少K个区域控制器。选择区域控制器的目的在于使控制器最小间距最大化。控制器间距通常用跳数表示。选出区域控制器的目的是为了实现区域控制器的统一分配。统一分配有助于负荷平衡。此外,在特定的实施例中,当网络节点密度相差较大时,给控制器间距加权以平衡各区域。当控制器数量K较小时,可以进行穷举选择,并有效地评估所有可能的组合。否则,可利用启发式方法来选择K个控制器。

[0095] 继续图4的实施例,在分区步骤430中,所述多个网络节点被划分为K个区域。在一些实施例中,K个区域为地理区域。在其他的实施例中,K个区域为逻辑区域。划分的目的是将TE优化过程中的开销成本最小化。TE优化的开销成本包括采集节点和链路的状态时产生的通信成本、进行配置时产生的通信成本,以及控制器间进行协调时产生的成本。根据启发式方法进行划分。可采用若干个启发式方法,包括:针对地理区域的聚类、划分,以及任务分配方法;以及针对逻辑区域的任务分配方法和图形划分方法。在分配步骤440中,将控制器分配给各个区域。该方法在结尾步骤450处结束。

[0096] 图5是可以用于实现在此公开的设备和方法的计算系统500的框图。特定设备可以使用示出的所有组件或其子集,且设备与设备间的集成程度不同。进一步,设备可以包含组件的多个实例,如多个处理单元、处理器、储存器、发射器,和接收器等。计算系统500可以包括处理单元502,所述处理单元502带有一个或多个输入/输出设备,比如扬声器、麦克风、鼠标、触摸屏、按键、键盘,以及打印机等。所述处理单元可以包括:中央处理器(简称CPU) 514、

存储器508、大容量存储器504、视频适配器510,以及用于连接总线520的I/O接口512。

[0097] 所述总线520可以为包括存储器总线或内存控制器、外围总线,或视频总线等几种任一类型的总线架构中的一个或多个。CPU514可以包括任意种类的电子数据处理器。存储器508可以包括任意类型的非瞬时性系统存储器,如静态随机存取存储器(简称SRAM)、动态随机存取存储器(简称DRAM)、同步动态随机存取存储器(简称SDRAM)、只读存储器(简称ROM),或其组合等。在一实施例中,存储器508可以包括启动时使用的ROM和执行程序时使用的用于存储程序 and 数据的DRAM。

[0098] 大容量存储器504可以包括任意类型的非瞬时性存储设备,所述非瞬时性存储设备用于存储数据、程序及其他信息的,并能使所述数据、程序,及其他信息可通过总线520进行访问。大容量存储器504可以包括,例如,固态硬盘、硬盘驱动器、磁盘驱动器,或光盘驱动器等中的一个或多个。

[0099] 视频适配器510和I/O接口512提供将外部输入输出设备耦合到处理单元502的接口。如图所示,输入和输出设备的例子包括与视频适配器510耦合的显示器518,以及和I/O接口512耦合的鼠标/键盘/打印机516。其他设备可以与处理单元502耦合,还可能会用上额外的或更少的接口卡。例如,串行接口,如通用串行总线(简称USB)(未图示),可以用于为打印机提供接口。

[0100] 处理单元502还包括一个或多个网络接口506,所述网络接口506可以包括有线链路如以太网线等,和/或接入接点或不同网络的无线链路。网络接口506允许处理单元502通过网络与远程单元通信。例如,网络接口506可以通过一个或多个发射器/发射天线以及一个或多个接收器/接收天线来提供无线通信。在一实施例中,将处理单元502与局域网522或广域网耦合来进行数据处理并与远程设备通信,如其他的处理单元、互联网,或远程存储设施等。

[0101] 图6为另一实施例的拓扑图600;拓扑图600包括网络节点610,其中,有八个被指定为控制器620。与图2相同,网络节点610用白色圆圈表示,控制器620用黑色圆圈表示。线630代表多个网络节点间的链路。

[0102] 图7示出了利用一实施例中的聚类启发式方法生成的地理分区结果700。分区结果700包括八个地理区域:710、720、730、740、750、760、770,和780。这八个区域中包含了图6拓扑图中的所有网络节点。图6中的控制器620也被分配给这八个地理区域。区域用形状各异的圆圈的组合表示,代表网络节点和其附带的链路。分区结果700还包括控制器620间的虚线,以及其各个区域中的每个网络节点。

[0103] 图8示出了利用一实施例中的划分启发式方法生成的地理分区结果。分区结果800包括八个地理区域:810、820、830、840、850、860、870,和880。这八个区域中包含了图6拓扑图中的所有网络节点。图6中的控制器620也被分配给这八个地理区域。区域用形状各异的圆圈的组合表示,代表网络节点和其附带的链路。分区结果800还包括控制器620间的虚线,以及其各个区域中的各个网络节点。

[0104] 图9是利用一实施例中的任务分配启发式方法生成的地理分区结果。分区结果900包括八个地理区域:910、920、930、940、950、960、970,和980。这八个区域中包含了图6拓扑图中的所有网络节点。图6中的控制器620也被分配给这八个地理区域。区域用形状各异的圆圈的组合表示,代表网络节点和其附带的链路。分区结果900还包括控制器620间的虚线,

以及其各个区域中的每个网络节点。

[0105] 图10示出了针对SDN1000的基于区域的TE优化问题。SDN1000包括主控制器1010、从控制器1020,以及网络节点1030。TE优化问题并非由主控制器1010解决,而是将所述TE优化问题分解成能被从控制器1020解决的子问题。从控制器1020中的每一个从控制器都被分配给区域1040。形成区域1040的目的在于使开销成本最小化。特定的成本只限定于特定的区域。例如,节点和链路状态的采集和配置,以虚线1050表示,只限定于每个区域。尽管额外的区域和从控制器能摊薄这些成本,与主-从协调,以实线1060表示,的关联的开销成本会随着从控制器的增加而增加。

[0106] 图11示出了根据拓扑1100-A的流向图1100-B的构造。拓扑1100-A包括三条业务流,为1110、1120和1130。业务流以各种方式穿越九个网络节点1140,所述九个网络节点1140的编号为1到9。第一业务流1110来源于节点1,其目的地为节点8。业务流1110有两条路由路径:1-2-5-7-8和1-3-5-7-8。第二业务流1120源于节点4,其目的地为节点8。业务流1120只有一条路由路径:4-5-7-8。第三业务流1130源于节点6,其目的地为节点9。业务流1130有两条路由路径:6-7-8-9和6-8-9。

[0107] 在流向图1100-B中,白色矩形表示业务流1110、1120,以及1130。给每条业务流分配加权,如括号中所示。加权是候选路由路径的数量,即流的大小与流的数量的乘积。在这个例子中,业务流1110、1120,以及1130的流的数量分别为1。因此,业务流各自的加权反映了候选路由路径的数量。流向图1100-B还示出了超图边1150和超图边1160。超图边1150代表节点7和节点8间的物理链路,该链路也在全部的所述三条业务流间共享。超图边1160代表节点5和7间的物理链路,该链路也在业务流1110和1120间共享。在这个例子中,其他物理链路未被共享,也未在流向图1100-B中示出。给超图边1150和1160也分配加权,该加权代表通过被示出的物理链路的候选路由路径的数量。超图边1150的加权为4,对应于分给业务流1110的两条路由路径、分给业务流1120的一条路径,以及分给业务流1130的路径6-7-8-9。相似地,超图边1160的加权为3,对应于分给业务流1110的两条路由路径和分给业务流1120的一条路由路径。

[0108] 虽然已参考说明性实施例描述了本发明,但此描述并不意图限制本发明。所属领域的技术人员在参考该描述后,将会明白说明性实施例的各种修改和组合,以及本发明其他实施例。因此,所附权利要求书意图涵盖任何此类修改或实施例。

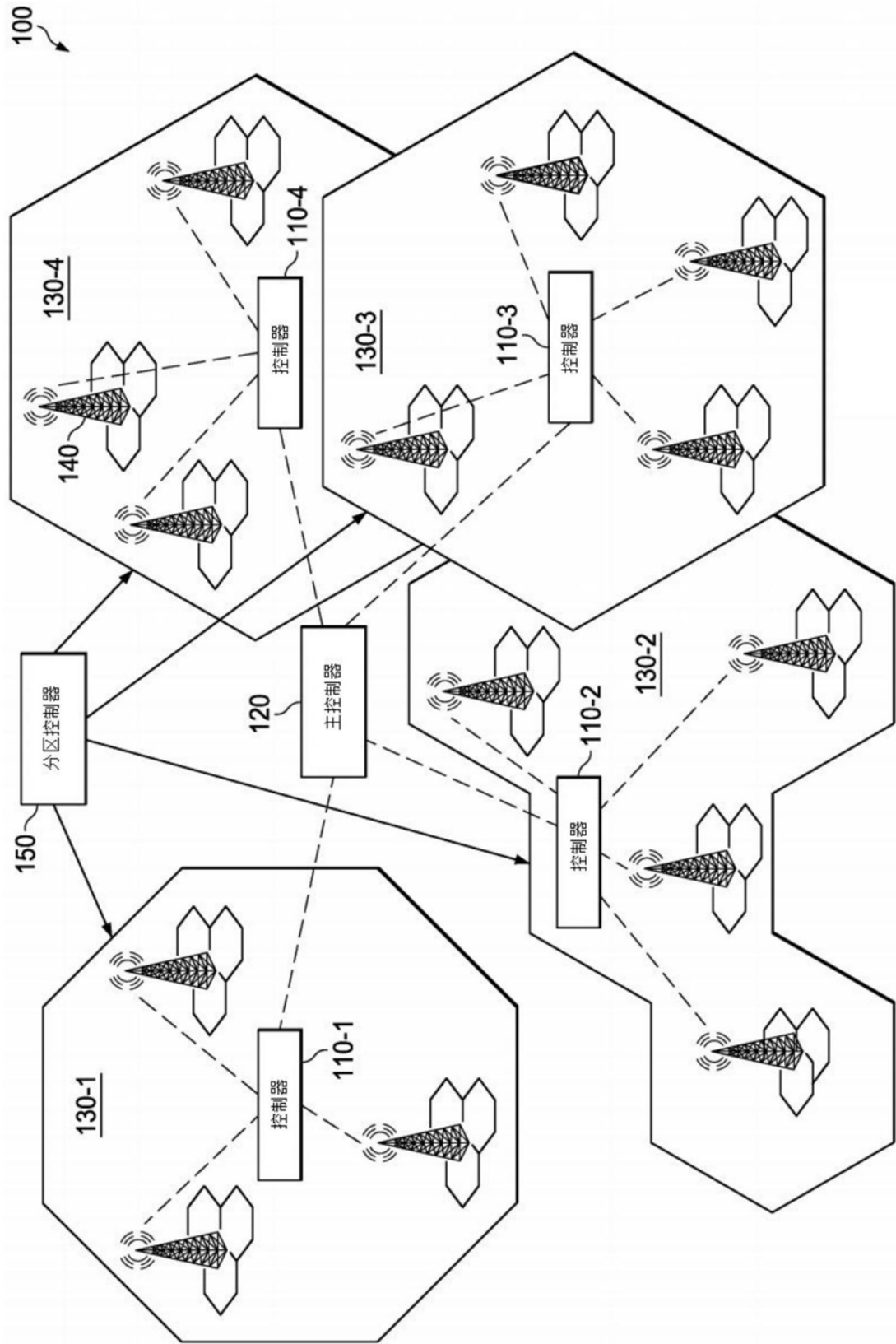


图1

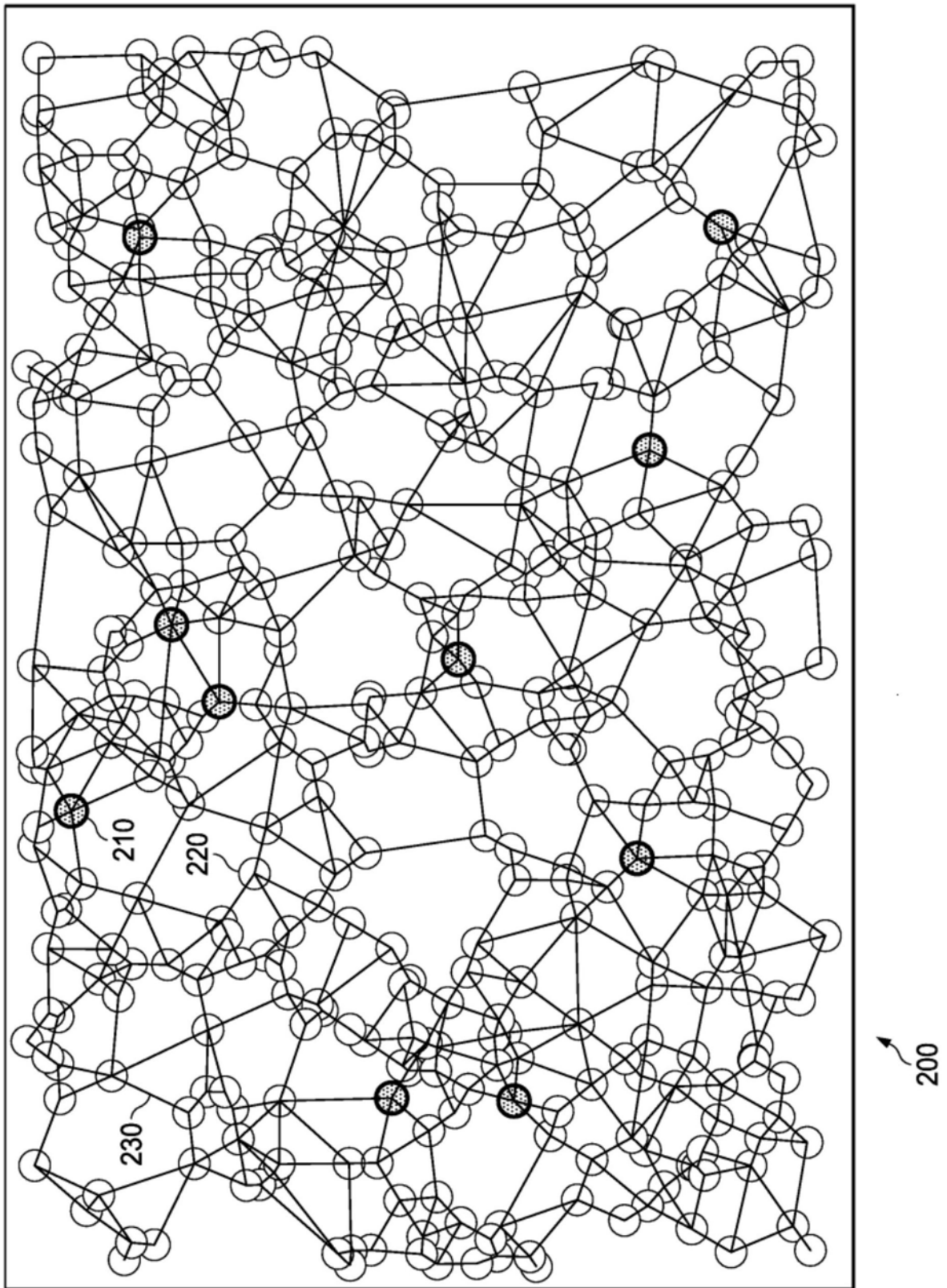


图2

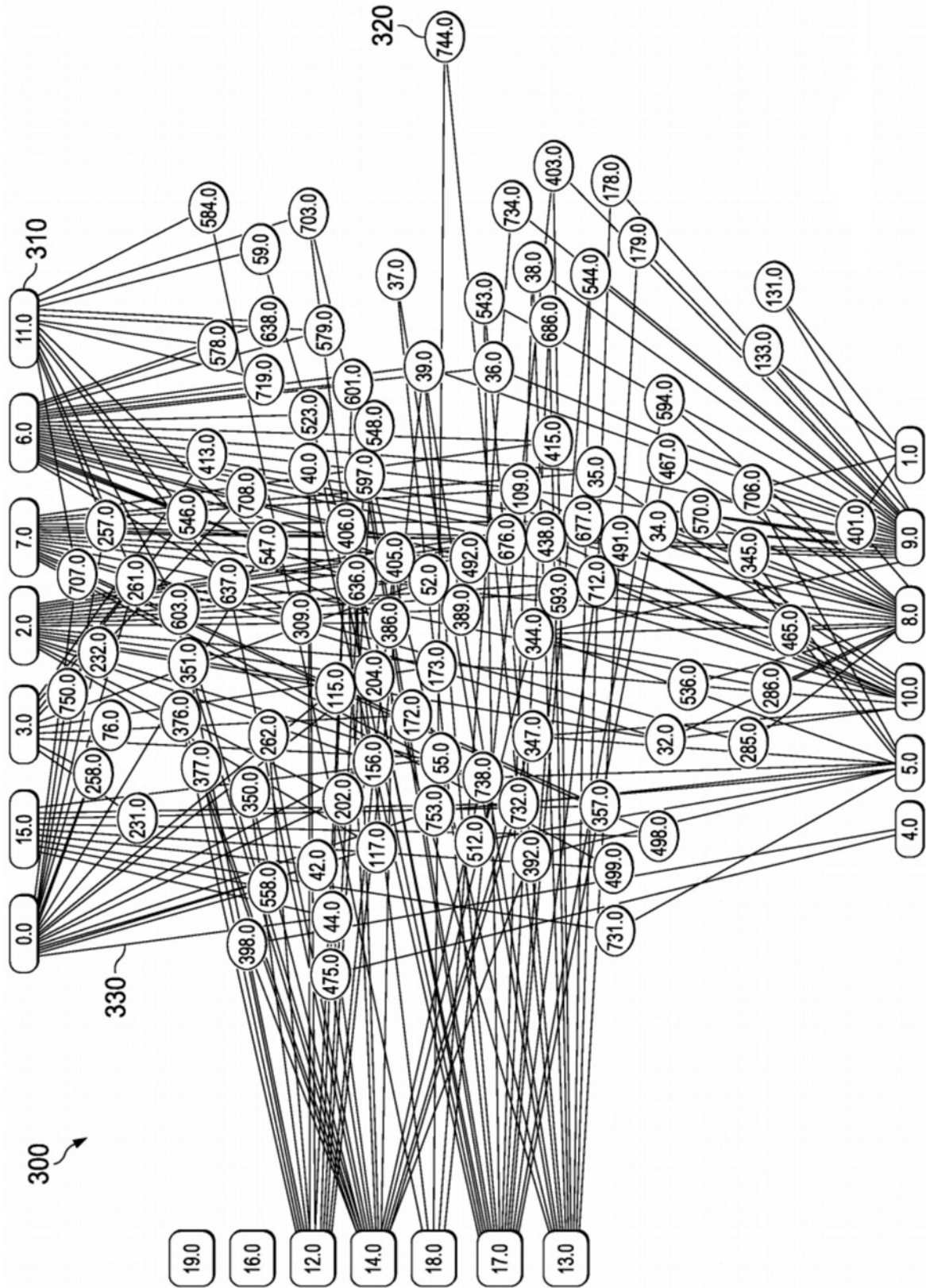


图3

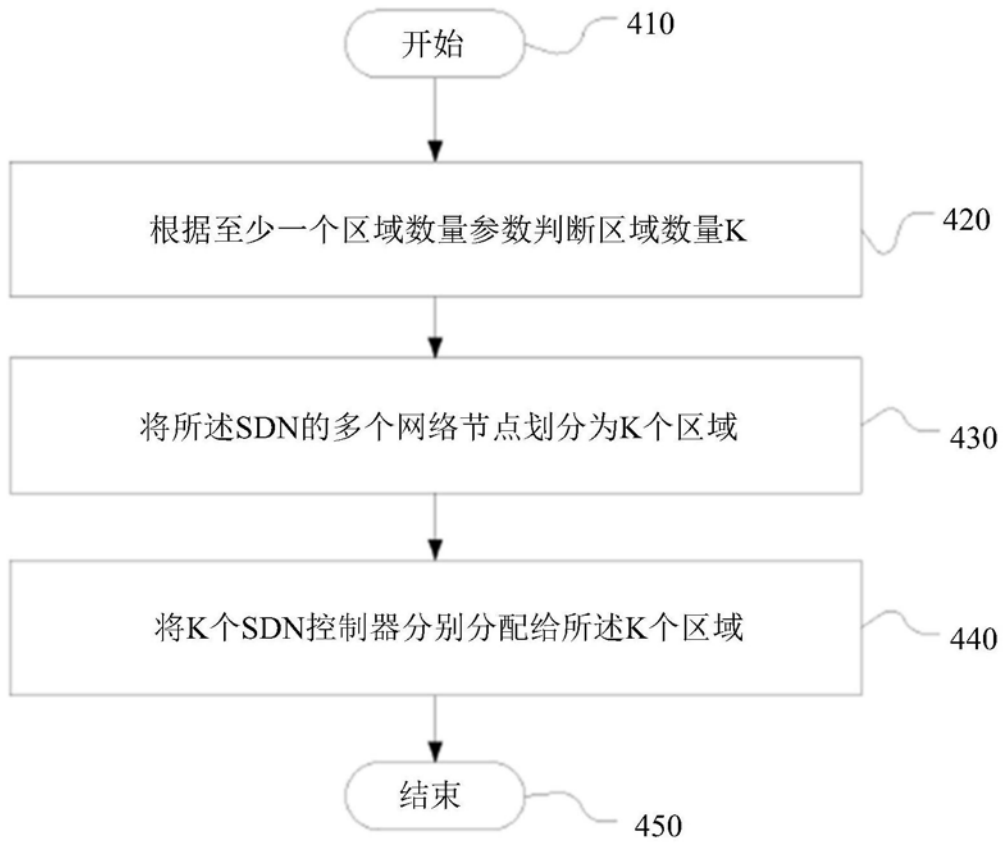


图4

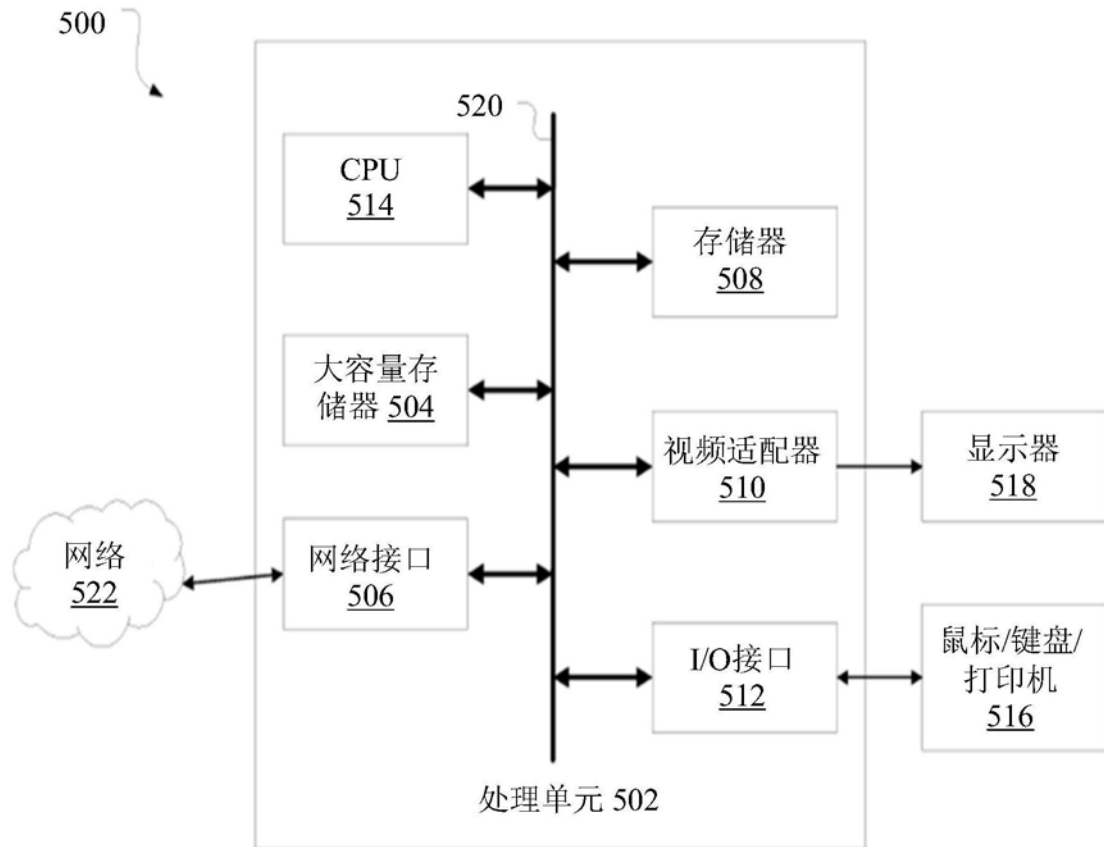


图5

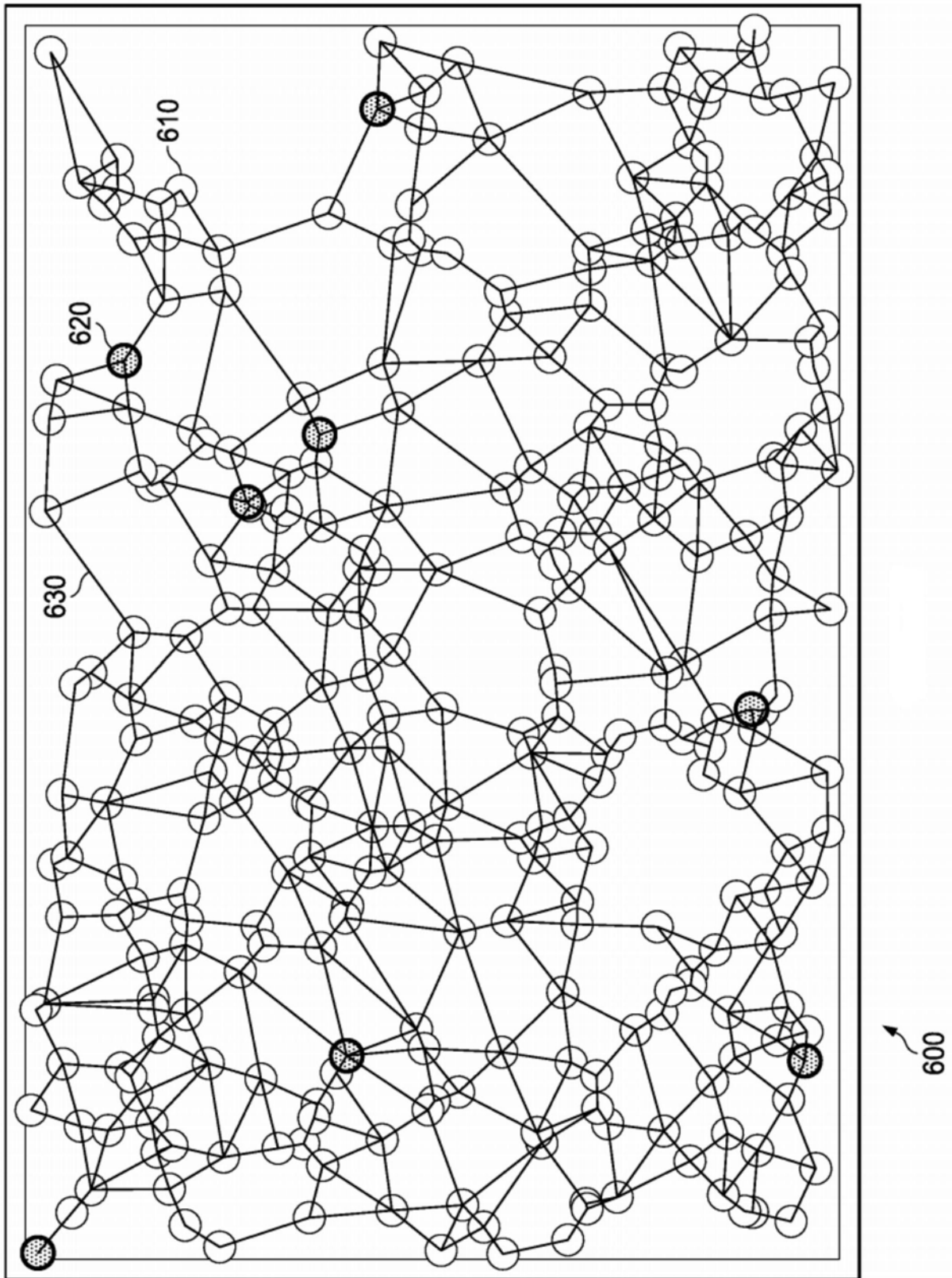


图6

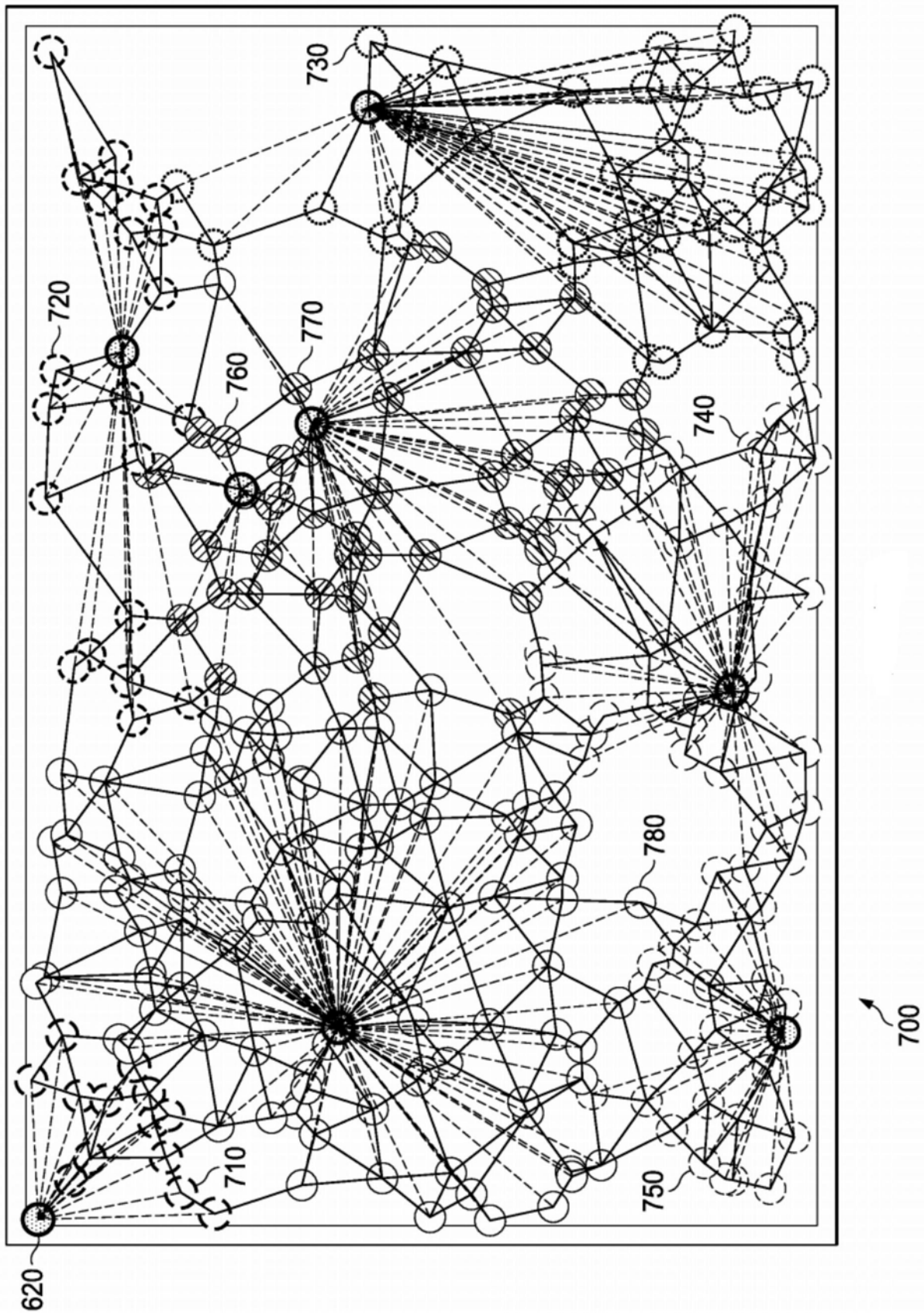


图7

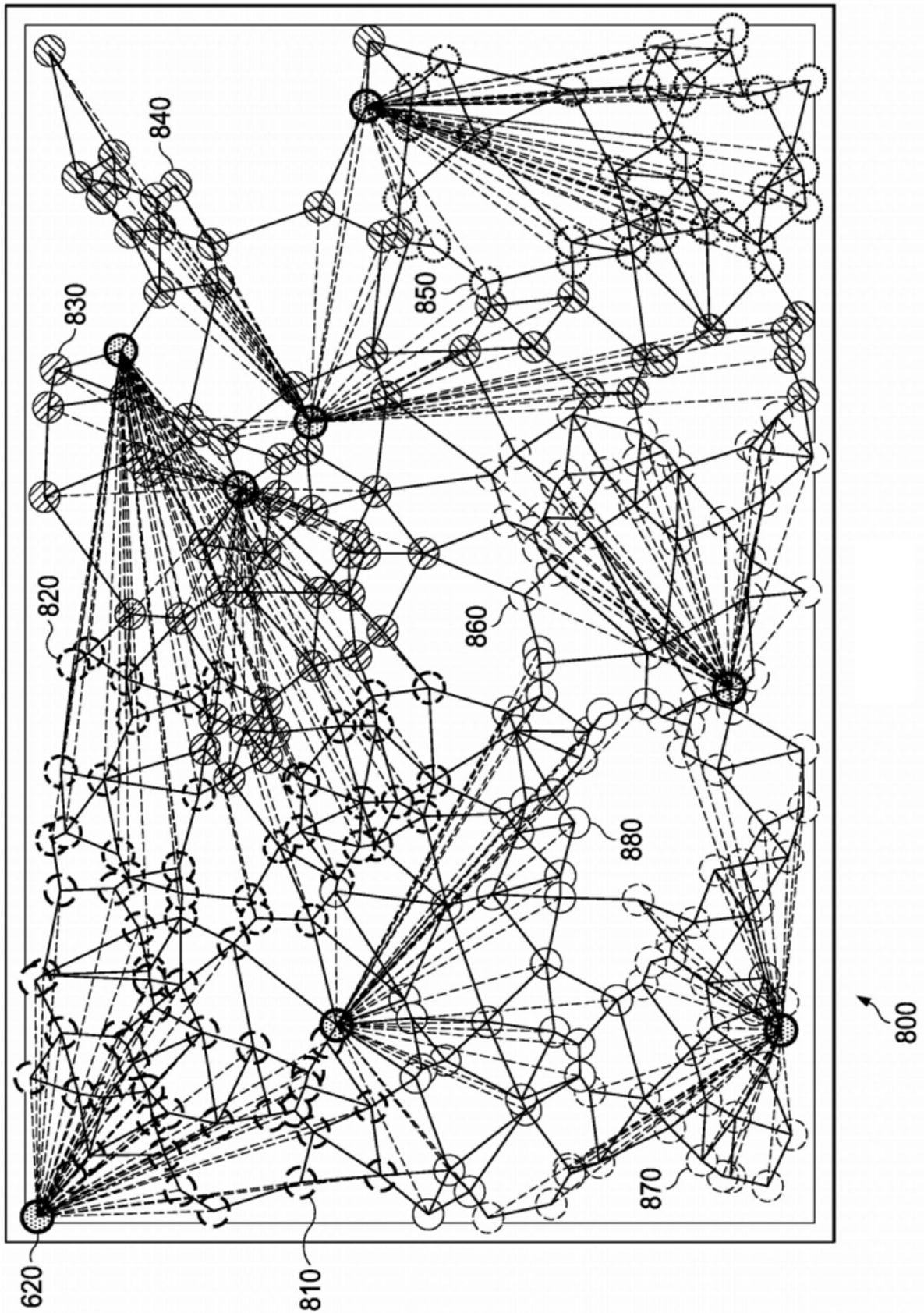


图8

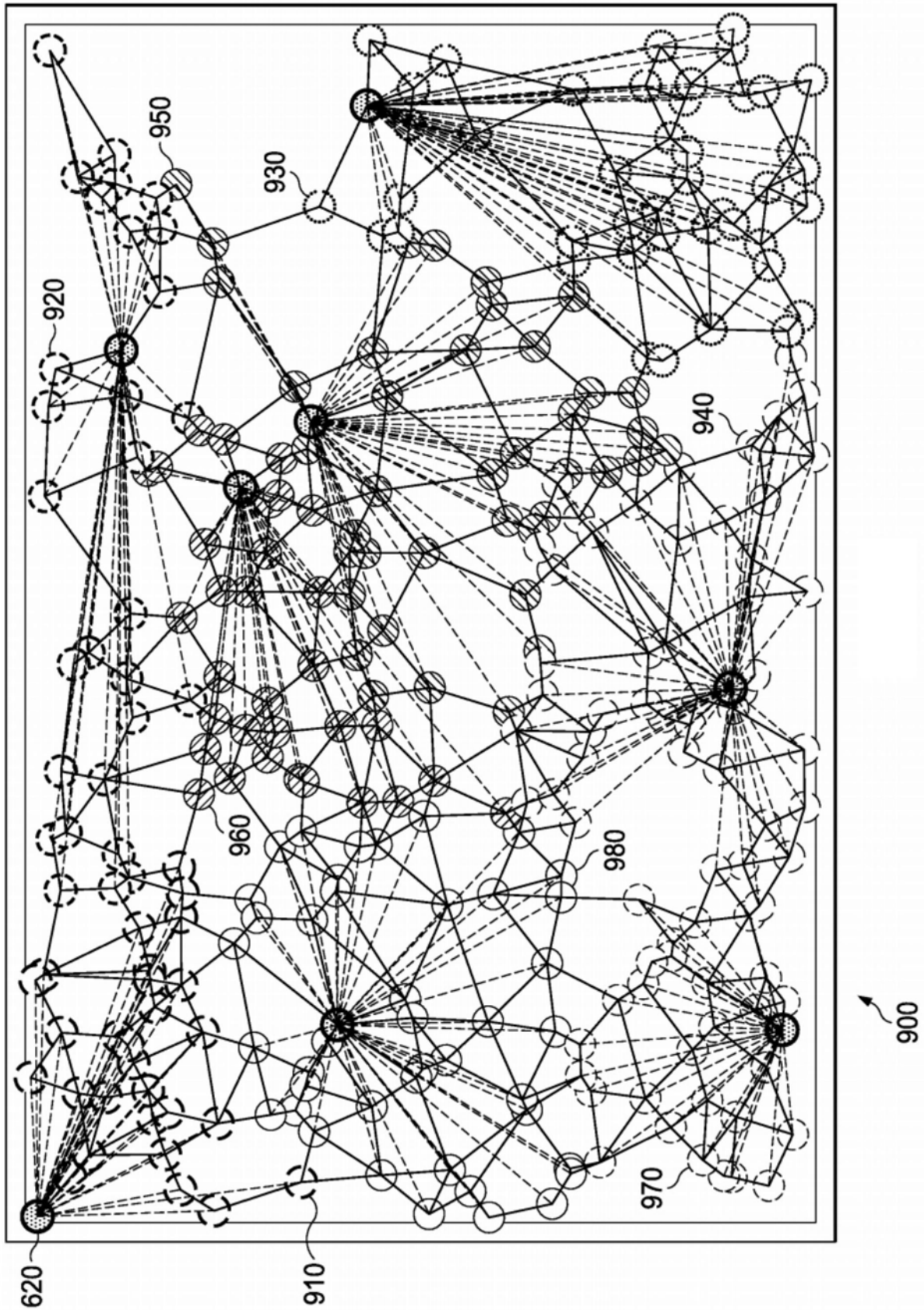


图9

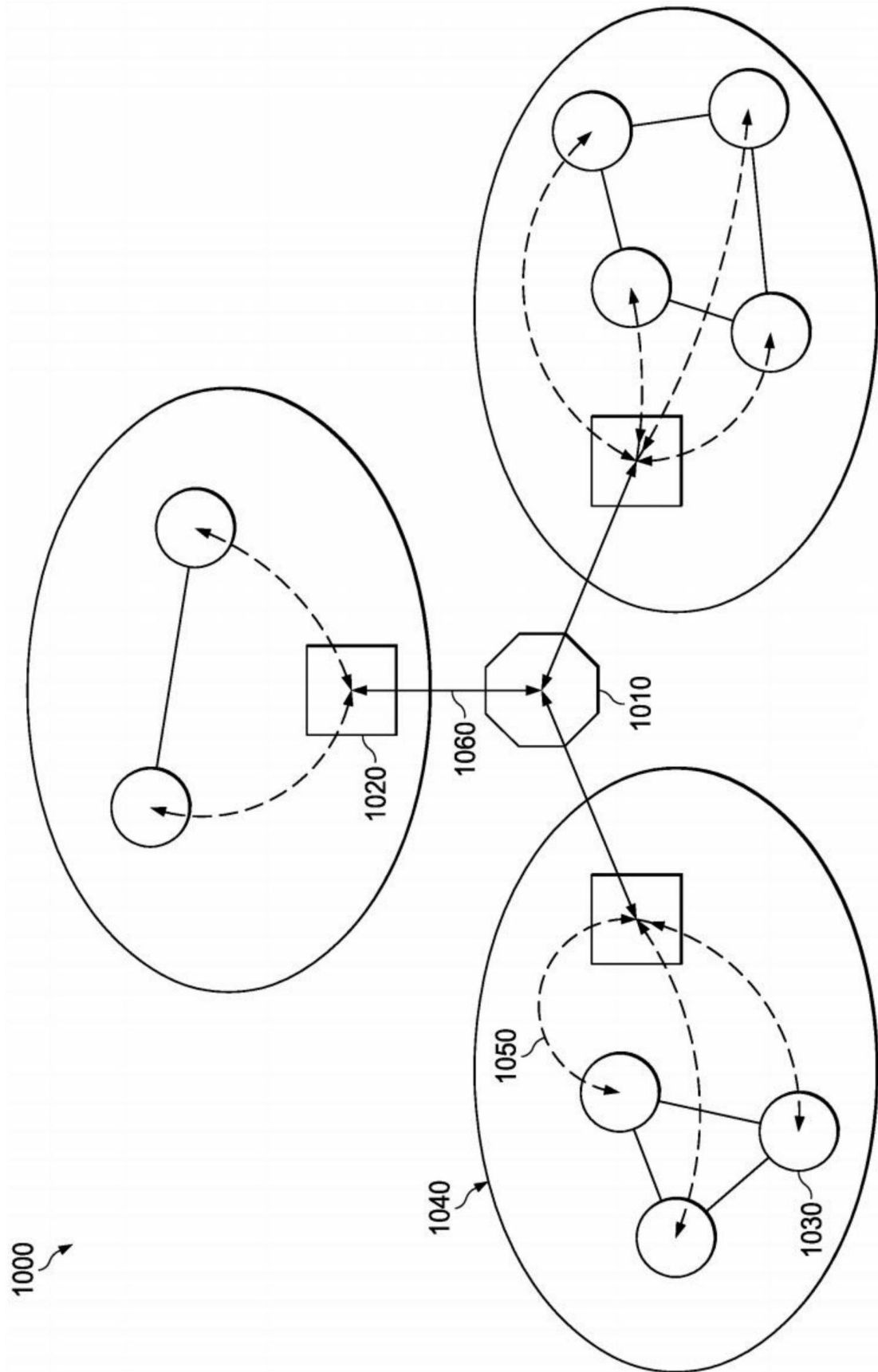


图10

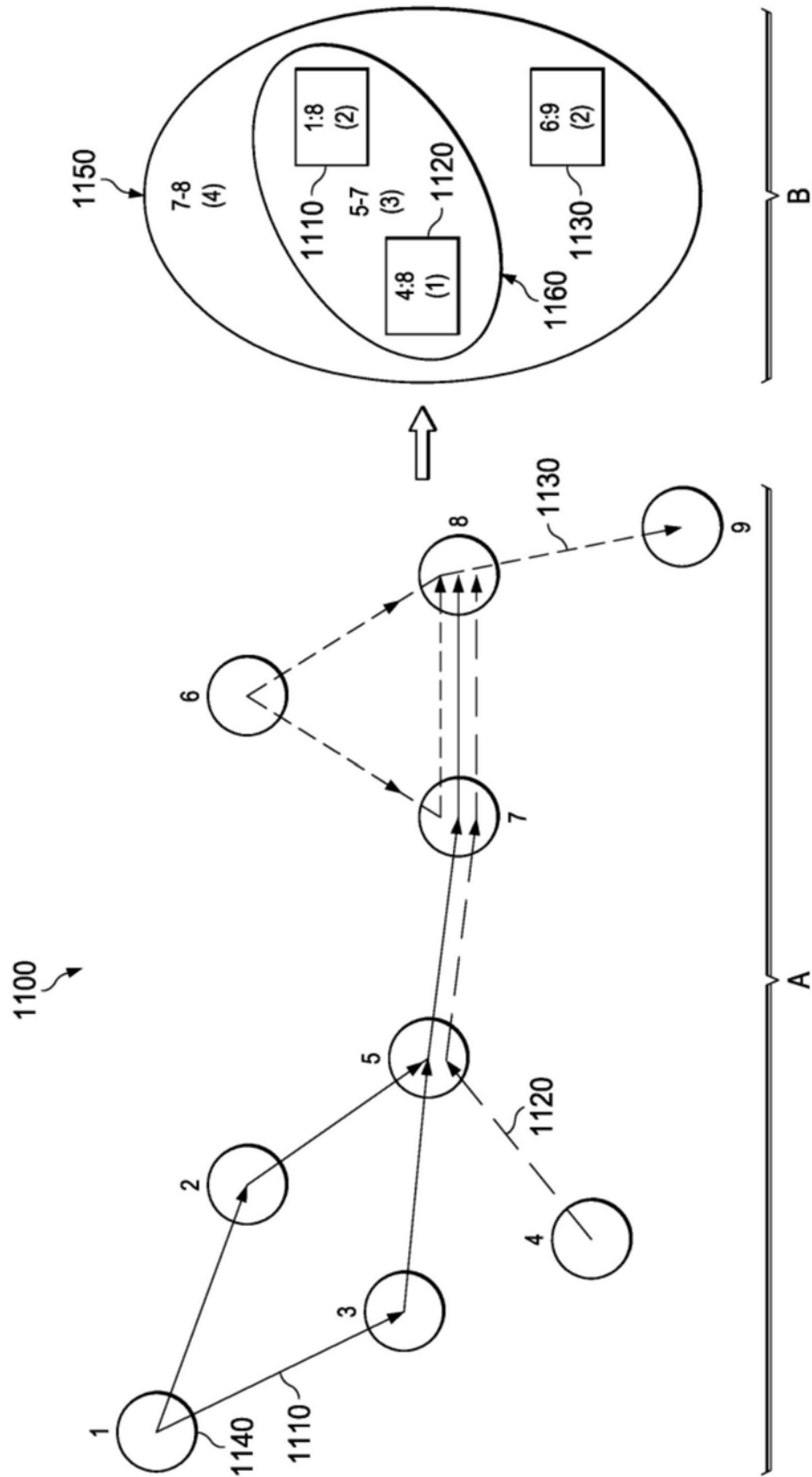


图11