



(12) 发明专利

(10) 授权公告号 CN 101505263 B

(45) 授权公告日 2011. 10. 26

(21) 申请号 200810065364. 2

CN 1878118 A, 2006. 12. 13,

(22) 申请日 2008. 02. 05

US 2006171332 A1, 2006. 08. 03,

(73) 专利权人 华为技术有限公司

刘庆江. ADHSNT 自发现层次存储网络拓扑模型设计. 《计算机科学》. 2007, 第 34 卷 (第 6 期), 61-63.

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

审查员 王晓玲

(72) 发明人 施广宇 陈坚 龚皓

(74) 专利代理机构 北京三高永信知识产权代理有限公司 11138

代理人 何文彬

(51) Int. Cl.

H04L 12/56 (2006. 01)

H04L 12/58 (2006. 01)

(56) 对比文件

CN 1939003 A, 2007. 03. 28,

CN 101079783 A, 2007. 11. 28,

CN 101064649 A, 2007. 10. 31,

WO 2006083495 A2, 2006. 08. 10,

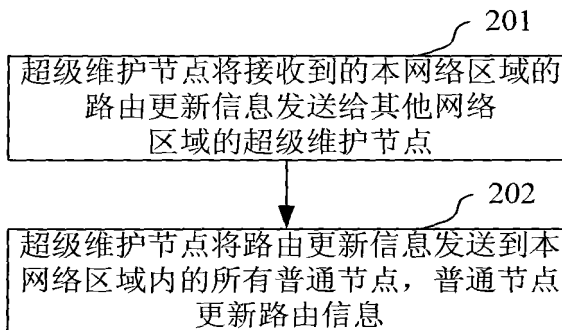
权利要求书 1 页 说明书 7 页 附图 4 页

(54) 发明名称

一种选择超级维护节点的方法及装置

(57) 摘要

本发明实施例提供一种结构化对等网络中维护路由信息的方法及装置,包括:选择处理能力强并且位于网络区域边界的节点作为超级维护节点,其他节点作为普通节点;将那些与超级维护节点网络距离较近的节点划分为一个网络区域;当节点加入或者失效时,普通节点将检测到的路由更新消息只发送给本网络区域的超级维护节点,超级维护节点再将更新消息转发给其他网络区域的超级维护节点,每个超级维护节点负责将收到的路由更新信息通知本网络区域内的普通节点。从而有效降低了 P2P 网络中跨网络区域的路由表维护开销。



1. 一种选择超级维护节点的方法,其特征在于,包括:

获得系统中节点之间的路由路径和所经过的路由器之间的时延信息;

对节点之间的路径信息和路由器之间的时延信息进行二分聚类,获得时延大和小的两个集合,把从节点出发沿着所述路径时所有上一跳时延为小,下一跳时延为大的路由器的地址作为所述节点的归属区域的标识,通过将具有相同区域标识的节点划分为同一个聚类,对全部节点划分得到多个聚类;

从节点形成的每一个聚类中,选择一个或者多个能力强的节点,作为超级维护节点。

2. 根据权利要求1所述的方法,其特征在于,所述获得系统中节点之间的路由路径和所经过的路由器之间的时延信息包括:

从系统中选择若干节点,执行 TraceRoute 命令,并对返回的结果进行处理,获得节点之间的路由信息和所经过的路由器之间的时延信息。

3. 根据权利要求2所述的方法,其特征在于,从系统中选择若干节点,执行 TraceRoute 命令包括:

由系统中的每一个节点选择若干其他节点作为目的节点执行 TraceRoute ;或者

由能力强的节点或者几个专门的计算机或者服务器,以系统中的所有节点作为目的节点,执行 TraceRoute 命令。

4. 一种选择超级维护节点的装置,其特征在于,包括:

探测单元,用于获得系统中节点之间的路由路径和所经过的路由器之间的时延信息;

计算单元,用于对节点之间的路径信息和路由器之间的时延信息进行二分聚类,获得时延大和小的两个集合,把从节点出发沿着所述路径时所有上一跳时延为小,下一跳时延为大的路由器的地址作为所述节点的归属区域的标识,通过将具有相同区域标识的节点划分为同一个聚类,对全部节点划分得到多个聚类;

选择单元,用于从节点形成的每一个聚类中,选择一个或者多个能力强的节点,作为超级维护节点。

5. 根据权利要求4所述的装置,其特征在于,所述探测单元包括:

路由执行模块,用于选择节点执行 TraceRoute 命令;

分析处理模块,用于处理返回的结果,获得节点之间的路由信息和所经过的路由器之间的时延信息。

一种选择超级维护节点的方法及装置

技术领域

[0001] 本发明涉及对等网 (P2P, Peer-to-Peer) 技术领域, 尤其涉及一种维护路由信息的方法及装置。

[0002] 背景技术

[0003] 由于 P2P 网络是一种自组织形态的网络系统, 该网络中, 每个节点加入网络或从网络中退出的行为均是随机性的。因此, 当节点加入或退出系统时, 系统需要通过采用更新机制对每个节点维护的路由信息进行更新, 才能够及时恢复路由关系, 使得查询可以可靠地进行。

[0004] 现有技术一般基于广播机制, 当 P2P 系统中某个节点加入或者失效时, 系统发送广播消息通知网络中所有节点更新路由信息, 这种机制虽然简单, 但是缺点也是非常明显的, 对系统中带宽要求很高, 当系统中同时失效的节点达到一定数量时, 容易产生网络风暴, 导致系统崩溃。

[0005] 发明内容

[0006] 本发明的实施例提供一种维护路由信息的方法及装置, 能够降低维护路由过程中产生的开销。

[0007] 本发明实施例提供的一种选择超级维护节点的方法, 包括:

[0008] 获得系统中节点之间的路由路径和所经过的路由器之间的时延信息;

[0009] 根据节点之间的路径信息和路由器之间的时延信息, 采用二分聚类的方法, 对节点划分形成多个聚类;

[0010] 从节点形成的每一个聚类中, 选择一个或者多个能力强的节点, 作为超级维护节点。

[0011] 本发明实施例提供的一种选择超级维护节点的装置, 包括:

[0012] 探测单元, 用于获得系统中节点之间的路由路径和所经过的路由器之间的时延信息;

[0013] 计算单元, 用于根据节点之间的路径信息和路由器之间的时延信息, 采用二分聚类的方法, 对多个节点划分形成多个聚类;

[0014] 选择单元, 用于从节点形成的每一个聚类中, 选择一个或者多个能力强的节点, 作为超级维护节点。

[0015] 本发明实施例提供的一种维护结构化对等网络中对等体路由表的方法及装置, 充分利用 P2P 系统中处理能力强, 并且位于网络区域边界的节点作为路由表更新维护节点, 负责将收到的路由更新信息通知本网络区域内的所有节点, 并将发生在本领域内的路由更新通知转发给其他区域的超级维护节点。这样, 不同区域的路由更新信息只在超级维护节点中相互传播, 并最终通过超级维护节点转发到网络中的所有节点, 从而有效降低了 P2P 网络中的由于节点变更产生的跨网络区域的路由表维护开销。

[0016] 图 1 是本发明实施例中选择超级维护节点的方法的流程图;

[0017] 图 2 是本发明实施例中自动形成超级维护节点的示意图;

- [0018] 图 3 是本发明实施例中一种超级维护节点产生装置的示意图；
- [0019] 图 4 是本发明实施例中一种构建多层次超级维护节点的示意图；
- [0020] 图 5 是本发明实施例中超级维护节点维护路由信息的方法流程图；
- [0021] 图 6 是本发明另一实施例中超级维护节点维护路由信息的方法流程图；
- [0022] 图 7 是本发明实施例中采用的条带分割方法划分网络区域的算法示意图；
- [0023] 图 8 是本发明实施例中一种维护路由信息的系统示意图；
- [0024] 图 9 是本发明实施例中一种超级维护节点的装置示意图。
- [0025] 下面将结合附图对本发明实施例的技术方案作进一步详细描述。
- [0026] 本发明实施例中,利用 P2P 网络中节点处理能力的差异性,选出处理能力强,并且位于网络边缘的节点作为超级维护节点,负责将本网络域的路由更新消息转发到其他网络域,从而在保证路由表及时更新的同时,能够有效降低路由消息的跨域流量。
- [0027] 附图说明
- [0028] 根据本发明的一个实施例,系统选择处理能力强并且位于网络区域边界的节点作为超级维护节点,其他节点作为普通节点;根据节点的地域信息,将那些与超级维护节点网络距离较近的节点划分为一个网络区域;当节点加入或者失效时,普通节点将检测到的路由更新消息只发送给本网络区域的超级维护节点,超级维护节点再将更新消息转发给其他网络区域的超级维护节点,每个超级维护节点负责将收到的路由更新信息通知本网络区域内的所有节点。这样,不同网络区域的路由更新信息只在超级维护节点中相互传播,并最终通过超级维护节点转发到网络中的所有节点,从而有效降低了 P2P 网络中跨网络区域的路由表维护开销。
- [0029] 本发明实施例中所指的超级维护节点可以为:处理能力强,并且位于网络区域边界的节点;所述的处理能力强的判断标准可以为:出口带宽,或者计算能力,或者硬盘大小,或者内存大小等,例如内存大于 4G 的节点即成为超级维护节点。
- [0030] 所述超级维护节点所负责维护的节点可以为:
- [0031] 所述超级维护节点所在的地域内的普通节点;
- [0032] 与所述超级维护节点时延小于 k 毫秒的普通节点,所述 K 值可以根据该超级维护节点所维护的普通节点的数量灵活设置,一般来讲, K 值越大,所维护的普通节点的数量越多;
- [0033] 或者也可以任意选择。
- [0034] 例如:根据节点的地域信息,将与超级维护节点网络距离较近的节点划分为一个网络区域,超级维护节点负责该网络区域内所有节点的路由信息更新。
- [0035] 根据地域关系来划分超级维护节点所维护的区域。比如将某个省的超级维护节点和普通节点划分为到一个网络区域中。也可以根据普通节点与超级维护节点间的时延信息划分网络区域,比如将与超级维护节点时延小于 20 毫秒的普通节点划分到一个网络区域中。
- [0036] 一个网络区域中的超级维护节点可以是 1 个或者多个。
- [0037] 参见图 1,本发明实施例提供一种选择超级维护节点的方法包括以下几个步骤:
- [0038] 步骤 101,获得系统中节点之间的路由路径和所经过的路由器之间的时延信息。
- [0039] 可以在系统中选择若干节点上,确定一定数量的目的节点,通过执行 TraceRoute

命令,并对返回的结果进行处理,获得节点之间的路由信息和所经过的路由器之间的时延信息。所述 TraceRoute 是计算机操作系统提供的工具,用于获得数据包到达目的节点所经过的中各路由器的地址清单和到达时间。

[0040] 可以由系统中的每一个节点以若干其他节点作为目的节点执行 TraceRoute,获得节点之间的路由路径和所经过的路由器之间的时延信息。例如,系统中每一个节点选取 2 个其他节点作为目的节点,每一个节点都执行 TraceRoute。具体选取的数量可以根据情况改变,只要能覆盖到整个系统就可以。

[0041] 也是可以由能力强的节点或者几个专门的计算机或者服务器,以系统中的所有节点作为目的节点,执行 TraceRoute 命令。

[0042] 图 2 是本发明实施例中自动形成超级维护节点的示意图,例如,在图 2 中,通过执行 TraceRoute 命令获得路由器 R1 与路由器 R2 之间的时延为 5 毫秒,路由器 R3 与路由器 R4 之间的时延为 100 毫秒。

[0043] 步骤 102,根据节点之间的路径信息和路由器之间的时延信息,采用二分聚类的方法,对多个节点划分形成多个聚类(Cluster)。

[0044] 节点将获得的路径信息和路由器之间的时延信息进行二分聚类,获得时延大和小的两个集合,把从本节点出发沿着该路径信息,所有上一跳时延为小但下一跳变为大的路由器的地址作为本节点归属区域的标识,称为位于网络区域边界的路由器。节点把这些标识和节点自己的信息注册到分布式哈希表或者某个数据存放位置,系统再把具有相同区域标识的节点形成聚类。通过这种方式,系统中的多个节点就会被划分成多个聚类,每个聚类组成的网络就是一个网络区域。

[0045] 例如,在图 2 中,R3 是节点的第一层归属区域标识(类比为“县”),那么所有以 R3 的地址作为归属路区域标识的节点形成一个聚类,R5 是节点的第二层归属区域标识(类比为“省”),那所有以 R5 的地址作为归属路区域标识的节点形成一个聚类。

[0046] 步骤 103,对于节点形成的每一个聚类中,选择一个或者多个能力强的节点,作为超级维护节点。

[0047] 对应于上述本发明实施例中选择超级维护节点的方法,本发明实施例还提供一种选择超级维护节点的装置,所述装置基于前面所述的方法实现,可以设置在节点,或者服务器,或者其他电信设备上。

[0048] 参见图 3,本发明实例提供一种自动产生超级维护节点的装置,包括:

[0049] 探测单元,用于获得系统中节点之间的路由路径和所经过的路由器之间的时延信息;

[0050] 计算单元,用于根据节点之间的路径信息和路由器之间的时延信息,采用二分聚类的方法,对多个节点划分形成多个聚类;

[0051] 选择单元,用于从节点形成的每一个聚类中,选择一个或者多个能力强的节点,作为超级维护节点。

[0052] 所述探测单元包括:

[0053] 路由执行模块,用于选择节点执行 TraceRoute 命令;

[0054] 分析处理模块,用于处理返回的结果,获得节点之间的路由信息和所经过的路由器之间的时延信息。

[0055] 另外,可以选择管理多个超级维护节点的级别较高的超级维护节点,形成等级关系,级别高的超级维护节点管辖多个级别低的超级维护节点。

[0056] 参见图 4,图 4 是一种构建多层次超级维护节点的示意图。第一网络区域中包括第一超级维护节点 A 和第一普通节点 B,第二网络区域中包括第二超级维护节点 C 和第二普通节点 D,第三网络区域包含第一网络区域和第二网络区域,第三网络区域包括第三超级维护节点 E 和第三普通节点 F,第四网络区域中包括第四超级维护节点 G 和第四普通节点 H。第三超级维护节点 E 和第一超级维护节点 A,第二超级维护节点 C 之间形成一种树状层次关系,并互相建立电信连接,第三超级维护节点 E 与第四超级维护节点 G 之间建立电信连接。

[0057] 本发明实施例提供一种利用超级维护节点维护路由信息的方法,包括以下步骤:

[0058] 参考图 5,图 5 是本发明实施例中超级维护节点维护路由信息的方法流程图。

[0059] 步骤 201、超级维护节点获得所属网络区域内的普通节点的路由更新信息,该超级维护节点根据所述路由更新信息,向其他网络区域的超级维护节点发送所述路由更新信息。

[0060] 所述普通节点可以将更新的路由信息发送给本区域的超级维护节点,所述超级维护节点接收所述路由更新信息;也可以由超级维护节点接收到普通节点发送的路由更新通知后,从所述普通节点获取所述节点的路由更新信息。

[0061] 当需要向所述超级维护节点发送消息时,普通节点获得所属区域超级维护节点的地址。普通节点可以通过多种手段获得所属区域超级维护节点的地址。超级维护节点采用分散式哈希表的方式将节点标识,IP 地址,层次关系等信息注册到 DHT 中,普通节点使用 DHT 的查询方法可以获得 P2P 系统中超级维护节点信息,包括超级维护节点的地址。

[0062] 所述超级维护节点也可以通过其他方式注册到公共地址中,比如注册到域名解析服务器(DNS, Domain Name System)或者数据库中。所述普通节点根据注册名称从 DNS 或者数据库中获得超级维护节点的地址。

[0063] 超级维护节点可以注册为 IP 任意播(Anycast)组成员,普通节点发起 Anycast 请求,则路由器返回它所在区域的超级维护节点的信息,包括超级维护节点的信息。也可以通过手工设置等方式为普通节点配置其所在网络区域的超级维护节点的地址。

[0064] 步骤 202、接收到所述更新路由信息的超级维护节点通知本区域内的普通节点更新路由信息。

[0065] 所述超级维护节点可以将新的路由信息发送给本区域的普通节点;也可以根据所述路由更新信息向普通节点发送路由更新通知,普通节点接收到路由更新通知后,主动到超级维护节点获取新的路由信息。

[0066] 所述超级维护节点可通过广播的形式发送路由更新信息,也可以通过并行多播的方式将路由更新信息逐步扩散发送到所属网络区域的普通节点。

[0067] 参见图 6,本发明的实施例还提供另外一种维护路由信息的方法,包括:

[0068] 步骤 301,当第一网络区域中的超级维护节点获知第二网络区域的所有超级维护节点失效时,第一网络区域的超级维护节点,根据构造带有地理位置标识的节点 ID 的方法,计算出第二网络区域的节点标识的范围,产生一条路由更新信息。

[0069] 所述路由更新信息包括失效节点标识的范围。

[0070] 步骤 302,第一网络区域的超级维护节点通知本区域内的普通节点更新路由信息。

[0071] 所述超级维护节点可以将新的路由信息发送给本区域的普通节点；也可以根据所述路由更新信息向普通节点发送路由更新通知，普通节点接收到路由更新通知后，主动到超级维护节点获取新的路由信息。

[0072] 所述超级维护节点可通过广播的形式发送路由更新信息，也可以通过并行多播的方式将路由更新信息逐步扩散发送到所属网络区域的普通节点。

[0073] 例如：当其他网络区域的超级维护节点获知某个网络区域的超级维护节点失效时，可根据所述构造带有地理位置标识的节点 ID 的方法获得失效超级维护节点所属区域的节点标识的范围，并将这个范围内的节点失效路由信息一次发送给所属网络区域的普通节点。大大减少了因逐个发送失效网络区域的节点路由更新信息而导致的维护开销。

[0074] 本发明实施例还提供一种构造带有地理位置标识的节点 ID 的方法，其方法包括以下步骤。

[0075] 步骤 401、获取节点的地理位置信息。

[0076] 节点在加入网络时会公布自己的地理位置信息，可以获取该节点的地理位置信息。

[0077] 步骤 402、利用条带分割的方法确定所述节点的 ID 哈希空间。

[0078] 步骤 403、在所述哈希空间中随机选取一个哈希值，并结合节点的其他属性信息，构造出节点的 ID 值。

[0079] 所述在哈希空间中随机选取的哈希值，可以作为节点 ID 的一部分（例如：前缀或者后缀，或者其中某关键字段）。

[0080] 步骤 402 和 403 中，参见图 7，采用条带分割选择节点 ID 的方法，每个区域各为图中的一个不同颜色的条块，整个哈希空间可以划分为 N 个条带，每个条带中再划分为 m 个（m 的数目为区域的数目大小）条目，每个区域的节点随机在属于该区域的条目中选择一个作为自己的节点 ID 的前缀（或者后缀，或者其中某关键字段），并结合节点的其他属性，构造出最终的节点 ID 值。这样就能很好地实现一种按地理区域位置平均划分哈希空间的节点 ID，大区域内的节点都被条带近似平均的分配到了各条目中。区域越大，分得也就越散。

[0081] 例如，ID 设定规则中，深圳市属于区域 B，那么一个位于深圳市的节点加入到网络中时，它会随机从哈希空间中选取一个条带，再从这个条带中找到属于该区域 B 的哈希数范围条目，并从此哈希范围内随机选择一个哈希数作为自己的节点 ID 的前缀（或者后缀，或者其中某关键字段），并结合节点的其他属性，构造出最终的节点 ID 值。通过这样一种条带分割选择 ID 的机制，由一个节点 ID 里的某关键字段再结合条带分割的规则，就能反推出该节点的详细地理位置，从而达到从节点 ID 中得知用户位置信息的目的。

[0082] 本发明实施例中，节点计算构造自己的 ID 的过程，既可以由节点自身完成，也可以统一由中心服务器完成，再由节点向中心服务器请求分配。

[0083] 对应于上述本发明实施例中基于超级节点维护路由信息的方法，本发明实施例还提供一种维护路由信息的系统，所述系统基于前面所述的方法实现，参见图 8，图 8 是该系统的示意图，该系统包括：第一普通节点，第一超级维护节点，第二普通节点，第三超级维护节点，第四普通节点，第四超级维护节点，第五普通节点。所述第一超级维护节点为，所述第一普通节点和第二普通节点所属网络区域的超级维护节点；所述第三超级维护节点为，所述第四普通节点所属网络区域的超级维护节点；所述第四超级维护节点为，所述第五普通

节点所属网络区域的超级维护节点。

[0084] 所述第一普通节点,产生路由更新信息,向第一超级维护节点发送路由更新信息;

[0085] 第一超级维护节点,接收来自第一普通节点的路由更新信息,根据所述更新路由信息,向第三超级维护节点和第二普通节点发送路由更新信息;

[0086] 第二普通节点,接收来自第一超级维护节点的路由更新信息,并更新路由信息;

[0087] 第三超级维护节点,接收来自第一超级维护节点的路由更新信息,根据所述路由更新消息,向第四普通节点和第四超级维护节点发送路由更新信息。

[0088] 第四普通节点,接收来自第三超级维护节点的路由更新信息,并更新路由信息。

[0089] 第四超级维护节点,接收来自第三超级维护节点的路由更新信息,并根据所述路由更新信息,向第五普通节点发送路由更新信息。

[0090] 第五普通节点,接收来自第四超级维护节点的路由更新信息,并更新路由信息。

[0091] 对应于上述本发明实施例中基于超级节点维护路由信息的方法,本发明实施例还提供一种维护路由信息的系统,所述系统基于前面所述的方法实现,参见图 8,图 8 是该系统的示意图,该系统包括:第一普通节点,第一超级维护节点,第二普通节点,第二超级维护节点,第三普通节点。所述第一超级维护节点为,所述第一普通节点和第二普通节点所属网络区域的超级维护节点,所述第二超级维护节点为,所述第三普通节点所属网络区域的超级维护节点。

[0092] 第一普通节点,用于产生路由更新信息,向第一超级维护节点发送路由更新信息;

[0093] 第一超级维护节点,接收来自第一普通节点的路由更新信息,根据所述路由更新信息,向第二超级维护节点和第二普通节点发送路由更新信息;

[0094] 第二普通节点,接收来自第一超级维护节点的路由更新信息,并更新路由信息;

[0095] 第二超级维护节点,接收来自第一超级维护节点的路由更新信息,并根据所述路由更新信息向第三普通节点发送路由更新信息;

[0096] 所述第一普通节点,产生路由更新信息,向第一超级维护节点发送路由更新信息;

[0097] 对应于上述本发明实施例中基于超级节点维护路由信息的方法,本发明实施例提供一种超级维护节点,包括:

[0098] 消息接收单元,用于获得路由更新信息;

[0099] 消息发送单元,用于根据所述的路由更新信息,向本网络区域普通节点和其他网络区域的超级维护节点发送路由更新信息。

[0100] 通过所述超级维护节点对网络区域内的路由更新信息统一转发给其他区域的超级维护节点或者网络区域内的普通节点。这样,不同区域的路由更新信息只在超级维护节点中相互传播从而有效降低了 P2P 网络中的由于节点变更产生的跨网络区域的路由表维护开销。

[0101] 参见图 9,本发明实施例提供的另一种超级维护节点,包括:

[0102] 消息接收单元,用于获得路由更新信息;

[0103] 监控单元,用于监控与其他超级维护节点的电信连接状态,并在获知第二区域的

所有维护节点都失效时,输出计算路由信息的指示;

[0104] 计算单元,用于接收所述计算路由信息的指示,根据构造带有地理位置标识的节点 ID 的方法计算出失效超级维护节点所属网络区域的节点标识的范围,产生路由更新信息;

[0105] 消息发送单元,用于根据所述的路由更新信息,向本网络区域普通节点和其他网络区域的超级维护节点发送路由更新信息。

[0106] 本实施例中超级维护节点对其他网络区域的超级维护节点进行监控,在某个网络区域的超级维护节点都失效的情况下,能够产生包括失效节点标识的范围的路由更新信息,并将这个范围内的节点失效路由信息一次发送给所属网络区域的普通节点。大大减少了因逐个发送失效网络区域的节点路由更新信息而导致的维护开销。

[0107] 本发明实施例提供的一种维护路由信息的方法及装置,充分利用 P2P 系统中处理能力强,并且位于网络区域边界的节点作为路由表更新维护节点,负责将收到的路由更新信息通知本网络区域内的普通节点,并将发生在 本网络区域内的路由更新信息转发给其他区域的超级维护节点。这样,不同区域的路由更新信息只在超级维护节点中相互传播,并最终通过超级维护节点转发到网络中的所有节点,从而有效降低了 P2P 网络中的由于节点变更产生的跨网络区域的路由表维护开销。

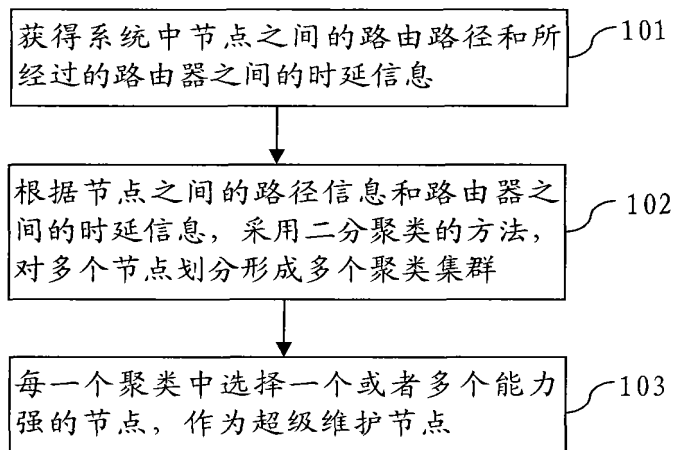


图 1

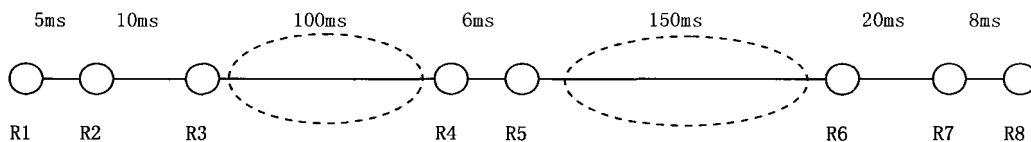


图 2

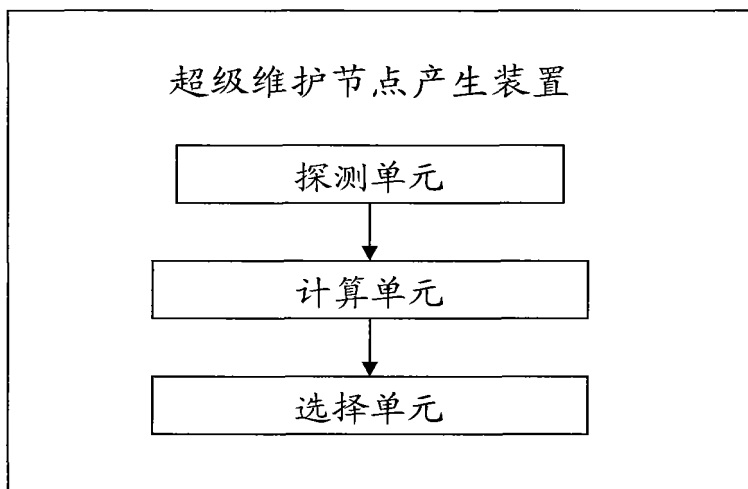


图 3

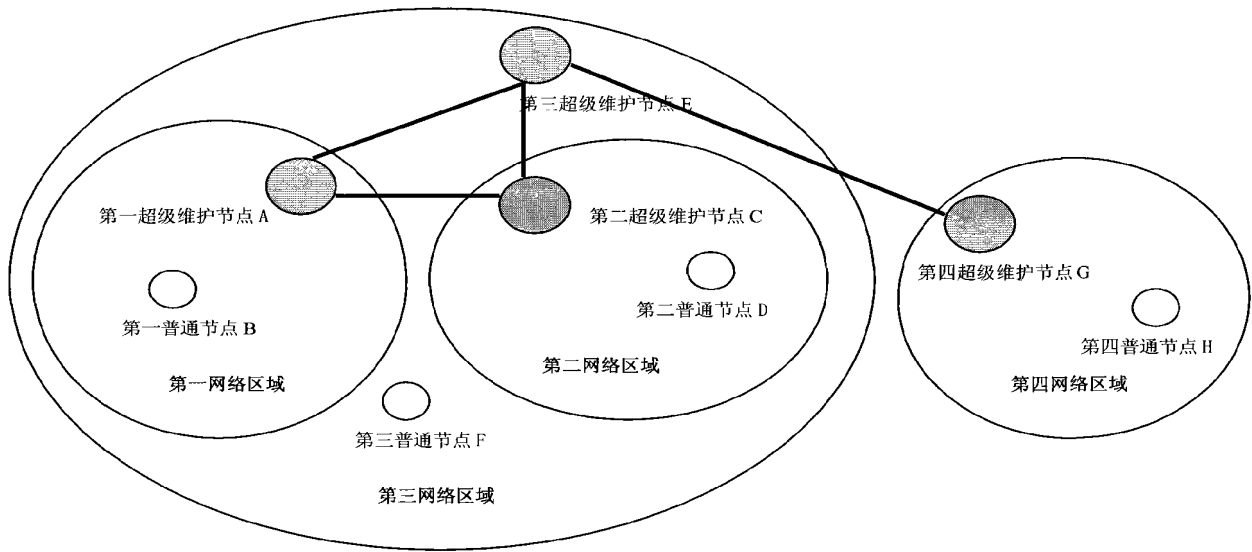


图 4

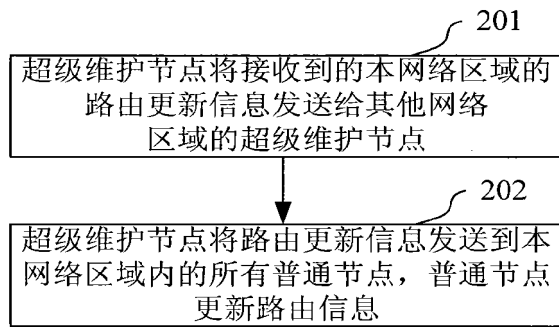


图 5

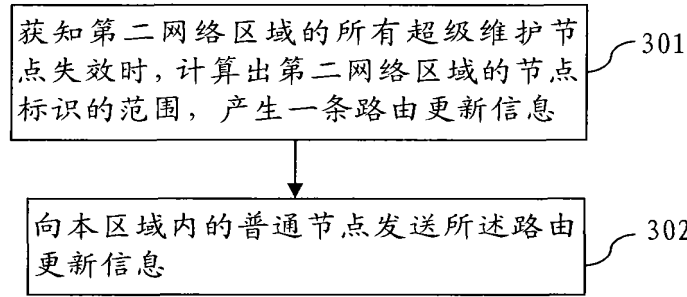


图 6

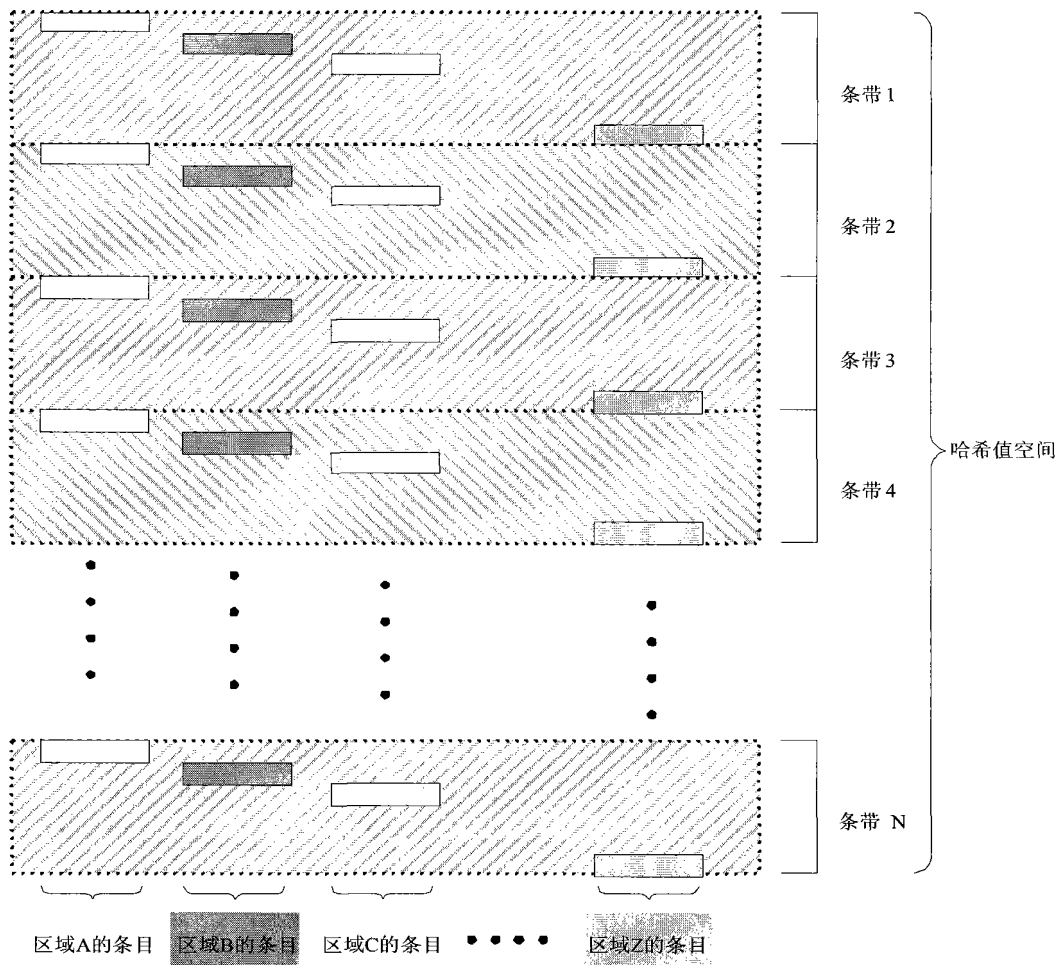


图 7

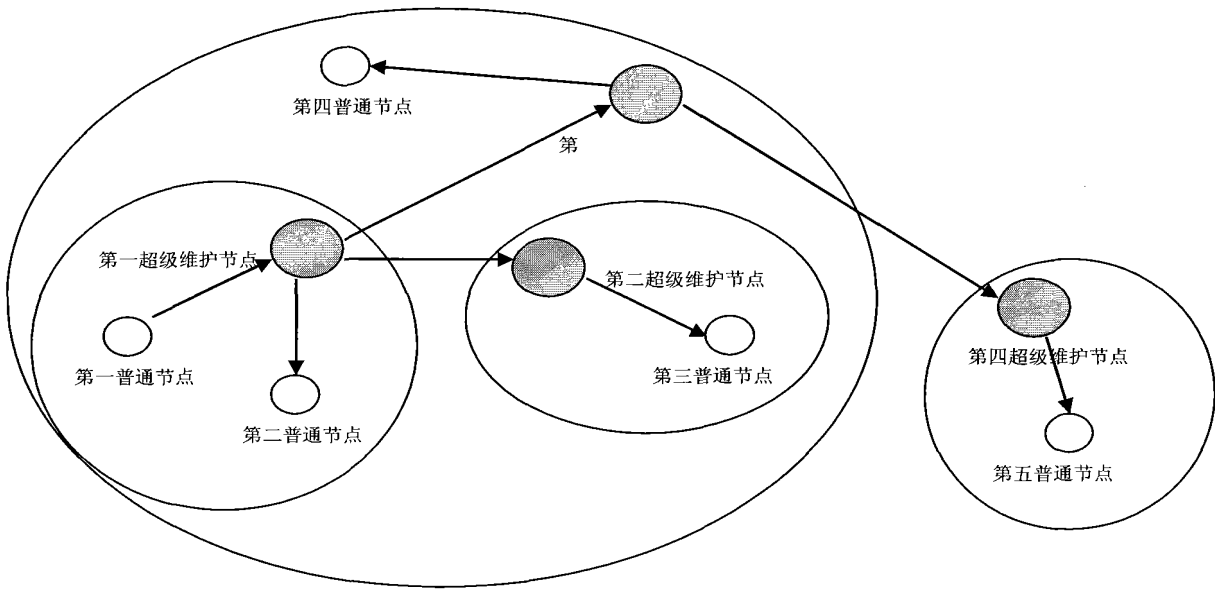


图 8

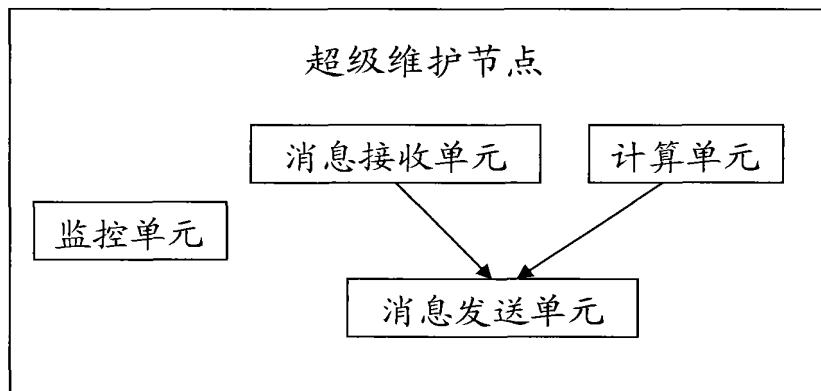


图 9