

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号

WO 2018/233372 A1

(43) 国际公布日

2018 年 12 月 27 日 (27.12.2018)

WIPO | PCT

(51) 国际专利分类号:

H04L 29/08 (2006.01)

(21) 国际申请号:

PCT/CN2018/084470

(22) 国际申请日:

2018 年 4 月 25 日 (25.04.2018)

(25) 申请语言:

中文

(26) 公布语言:

中文

(30) 优先权:

201710481690.0 2017年6月22日 (22.06.2017) CN

(71) 申请人: 平安科技 (深圳) 有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];

中国广东省深圳市福田区八卦岭八卦三路平安大厦六楼, Guangdong 518000 (CN)。

(72) 发明人: 刘仁仕 (LIU, Renshi); 中国广东省深圳市福田区八卦岭八卦三路平安大厦六楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市世纪恒程知识产权代理事务所 (CENFO INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市南山区南山大道 3838 号设计产业园金栋二层 210-212 (原南头城工业村 11 栋), Guangdong 518052 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,

(54) Title: LVS SCHEDULING METHOD, DEVICE, AND APPARATUS SUPPORTING HASH SCHEDULING POLICY, AND STORAGE MEDIUM

(54) 发明名称: 支持哈希调度策略的lvs调度方法、设备、装置及存储介质

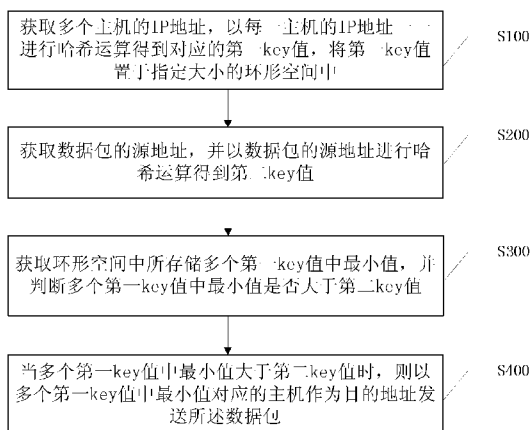


图 1

S100 Obtain IPs of a plurality of hosts, perform a hash operation using the IP of each host to obtain a corresponding first key value, and place the first key value in an annular space having a specified size
S200 Obtain a source address of a data packet, and perform a hash operation using the source address to obtain a second key value
S300 Obtain a minimum value of a plurality of first key values in the annular space, and determine whether the minimum value of the plurality of first key values is greater than the second key value
S400 If the minimum value of the plurality of first key values is greater than the second key value, transmit the data packet

(57) Abstract: The application provides an LVS scheduling method, device, and apparatus supporting a hash scheduling policy, and a storage medium. The method comprises: obtaining IPs of a plurality of hosts, performing a hash operation using the IP of each host to obtain a corresponding first key value, and placing the first key value in an annular space having a specified size; obtaining a source address of a data packet, and performing a hash operation using the source address to obtain a second key value; obtaining a minimum value of a plurality of first key values in the annular space, and determining whether the minimum value is greater than the second key value; and if so, transmitting the data packet using a host corresponding to the minimum value as a destination address. The application realizes reduced session drift and improves load balancing service quality.

(57) 摘要: 本申请提供了支持哈希调度策略的lvs调度方法、设备、装置及存储介质, 方法包括: 获取多个主机的IP, 以每一主机的IP一一进行哈希运算得到对应的第一key值, 将第一key值置于指定大小的环形空间中; 获取数据包的源地址, 并以该源地址进行哈希运算得到第二key值; 获取环形空间中多个第一key值中最小值, 并判断该最小值是否大于第二key值; 当是时, 则以该最小值对应的主机作为目的地址发送所述数据包。本申请实现了降低会话漂移, 提升负载均衡服务质量。

NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

支持哈希调度策略的lvs调度方法、设备、装置及存储介质

- [1] 本申请要求于2017年06月22日提交中国专利局、申请号为201710481690.0、发明名称为“支持哈希调度策略的lvs调度方法、设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。
- [2] 技术领域
- [3] 本申请涉及计算机网络技术领域，尤其涉及支持哈希调度策略的lvs调度方法、装置、设备及存储介质。
- [4] 背景技术
- [5] lvs（即Linux Virtual Server，表示Linux虚拟服务器）的sh（source address，即源地址）调度方法不支持源地址一致性hash调度，仅通过普通的hash算法根据源地址选择相应real server进行调度。当某台real server（即主机）宕机时，需要将real server从转发表中删除，否则将出现无法为部分客户端提供服务。当新添加real server时，需要将该real server加到转发表中。当前的sh调度方法下，real server的删除或者添加会引起大量的session（即会话）漂移。
- [6] 因此，现有技术还有待于改进和发展。
- [7] 发明内容
- [8] 鉴于上述现有技术的不足之处，本申请的目的在于提供支持哈希调度策略的lvs调度方法、装置、设备及存储介质，旨在解决现有技术中采用sh调度方法时，主机的删除或者添加会引起大量的会话漂移的问题。
- [9] 为了达到上述目的，本申请采取了以下技术方案：
- [10] 一种支持哈希调度策略的lvs调度方法，其中，所述方法包括以下步骤：
- [11] 获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；
- [12] 获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；

- [13] 获取环形空间中所存储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；
- [14] 当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。
- [15] 一种支持哈希调度策略的lvs调度设备，其中，所述支持哈希调度策略的lvs调度设备包括处理器、存储器及通信总线；
- [16] 所述通信总线用于实现处理器和存储器之间的连接通信；
- [17] 所述处理器用于执行存储器中存储的支持哈希调度策略的lvs调度程序，以实现以下步骤：
- [18] 获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；
- [19] 获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；
- [20] 获取环形空间中所存储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；
- [21] 当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。
- [22] 一种存储介质，其中，所述存储介质存储有一个或者多个程序，所述一个或者多个程序可被一个或者多个处理器执行，以实现所述的支持哈希调度策略的lvs调度方法的步骤。
- [23] 一种支持哈希调度策略的lvs调度装置，所述支持哈希调度策略的lvs调度装置包括：
- [24] 第一哈希模块，用于获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；
- [25] 第二哈希模块，用于获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；
- [26] 判断模块，用于获取环形空间中所存储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；
- [27] 发送模块，用于当多个第一key值中最小值大于第二key值时，则以多个第一ke

y值中最小值对应的主机作为目的地址发送所述数据包。

[28] 有益效果：本申请提供的支持哈希调度策略的lvs调度方法、设备及存储介质，当有某个主机连接到Nginx物理服务器上时，调度到相邻hash节点上的连接漂移到加入的主机；当有某个主机退出时，调度到其他主机的会话不会漂移，只有调度到退出的主机上的会话会重新分布，实现了降低会话漂移，提升负载均衡服务质量。

[29] 附图说明

[30] 图1为本申请所述的支持哈希调度策略的lvs调度方法较佳实施例的流程图。

[31] 图2为本申请所述的支持哈希调度策略的lvs调度方法中步骤S100的流程图。

[32] 图3为本申请所述的支持哈希调度策略的lvs调度方法较佳实施例中步骤S400的流程图。

[33] 图4为本申请所述的支持哈希调度策略的lvs调度程序较佳实施例的运行环境示意图。

[34] 图5为本申请所述的支持哈希调度策略的lvs调度程序较佳实施例的功能模块图。

[35] 具体实施方式

[36] 本申请提供支持哈希调度策略的lvs调度方法、设备及存储介质，为使本申请的目的、技术方案及效果更加清楚、明确，以下参照附图并举实施例对本申请进一步详细说明。应当理解，此处所描述的具体实施例仅用以解释本申请，并不用于限定本申请。

[37] 请参阅图1，是本申请所述的支持哈希调度策略的lvs调度方法较佳实施例的流程图。如图1所示，所述支持哈希调度策略的lvs调度方法，包括以下步骤：

[38] 步骤S100、获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中。

[39] 本实施例中，获取了主机的IP地址后，并不是与现有sh调度方法一样先根据源地址选择相应主机进行调度，而是先对主机的IP地址进行哈希运算，得到第一key值。例如有4台主机，分别记为主机1、主机2、主机3和主机4，其中主机1的IP地址为192.168.1.100，主机2的IP地址为192.168.1.101，主机3的IP地址为192.168.

1.102, 主机4的IP地址为192.168.1.104; 分别对3台主机的IP地址进行哈希运算, 并得到对应的第一key值, 具体的Hash (“192.168.1.100”) =key1、Hash (“192.168.1.101”) =key2、Hash (“192.168.1.102”) =key3、Hash (“192.168.1.103”) =key4, 得到的key1、key2、key3和key4都为第一key值; 将上述key1、key2、key3和key4都映射到一个环形空间中, 然后以顺时针的方向计算将所有的第一key值存储到离自己最近的主机中。

[40] 可以看出, 第一key值与主机处于同一个哈希空间中。在这样的部署环境中, 哈希环是不会更改的, 只要算出了对象的hash值就能快速的定位到对应的机器中, 这样就能找到对象真正的存储位置了。

[41] 步骤S200、获取数据包的源地址, 并以数据包的源地址进行哈希运算得到第二key值;

[42] 同样的, 本实施例中, 在获取了数据包的源地址之后, 直接对数据包的源地址进行哈希运算, 即进行hash(Packet.source_ip)的运算, 得到第二key值。获取的第二key值并不是部署到环形空间 (即哈希环) 中, 而是作为一个寻址用的中间值。

[43] 步骤S300、获取环形空间中所存储多个第一key值中最小值, 并判断多个第一key值中最小值是否大于第二key值。

[44] 在本实施例中, 先获取环形空间中所存储多个第一key值中最小值, 即进行MIN{hash(RealServer.ip)}的运算, 当得到了MIN{hash(RealServer.ip)}后与hash(Packet.source_ip) (即第二key值) 进行比较, 这一比较结果决定了目的地址的选定方向。

[45] 步骤S400、当多个第一key值中最小值大于第二key值时, 则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。

[46] 在步骤S400中, 若判断了多个第一key值中最小值大于第二key值时, 则获取MIN{hash(RealServer.ip)}对应的主机, 并以该主机对应的IP地址作为目的地址, 数据包根据该目的地址向此处发送。这样, 实现了当有某个主机连接到Nginx物理服务器上时, 调度到相邻hash节点上的连接漂移到加入的主机, 不会引发大量的会话漂移, 提升负载均衡服务质量。

[47] 优选的，如图2所示，所述步骤S100包括：

[48] 步骤S101、获取多个主机的IP地址，并以每一主机的IP地址一一进行哈希运算得到对应的第一key值；

[49] 步骤S102、获取ipv4的地址个数，并以ipv4的地址个数的两倍值作为空间大小设置一环形空间；

[50] 步骤S103、将第一key值置于所述环形空间中。

[51] 在步骤S101中，所获取的主机个数可能为0，也可能是大于0的正整数。当有至少一个主机与Nginx物理服务器连接时，则获取每一主机的IP地址，并将每一主机的IP地址都进行哈希运算得到第一key值。这样，多个主机的IP地址经哈希运算得到的key值均不一样，也就是每个key值都是唯一的，确保了后面的寻址不会出错。

[52] 在步骤S102中，为了确保环形空间中有足够的节点，需至少保证环形空间的大小是ipv4的地址个数的两倍。这与现有技术中采用sh调度时采用232-1的环形空间相比，节省了很多空间大小，以最小的空间大小满足了调度的基本需求。本实施中选择ipv4的地址个数的两倍值作为环形空间的空间大小，是因为目前主机的IP地址格式是ipv4。显然，当主机的IP地址为ipv6的格式，同样可以选择ipv6的地址个数的两倍值作为环形空间的空间大小。

[53] 而在步骤S103中，当有多个第一key值需置于所述环形空间中时，是按照第一key值得到的先后顺序，按照顺时针方向依次插入所述环形空间。

[54] 普通的sh调度方法中，采用的是hash求余算法，其最为不妥的地方就是在有机器的添加或者删除之后会照成大量的对象存储位置失效，这样就大大的不满足单调性了。基于步骤S101-S103的所采用的基于一致性哈希算法的处理方式，与普通的sh调度方法相比，能有效满足单独性。

[55] 以上面4台主机的例子继续举例说明。如果按照顺时针的方向，将key1存储到哈希环中的NODE1，将key3存储到哈希环中的NODE2，将key2和key4存储到哈希环中的NODE3，如果NODE2出现故障被删除了，那么按照顺时针迁移的方法，object3将会被迁移到NODE3中，这样仅仅是object3的映射位置发生了变化，其它的对象没有任何的改动。

- [56] 如果往集群中添加一个新的节点NODE4，通过对应的哈希算法得到KEY4，并映射到哈希环中。通过按顺时针迁移的规则，那么object2被迁移到了NODE4中，其它对象还保持这原有的存储位置。通过对节点的添加和删除的分析，一致性哈希算法在保持了单调性的同时，还是数据的迁移达到了最小，这样的算法对分布式集群来说是非常合适的，避免了大量数据迁移，减小了服务器的压力。
- [57] 优选的，如图3所示，所示步骤S400包括：
- [58] 步骤S401、当多个第一key值中最小值大于第二key值时，则将多个第一key值中最小值记为第一最小key值，并获取与第一最小key值对应主机；
- [59] 步骤S402、获取与第一最小key值对应主机的IP地址；
- [60] 步骤S403、以与第一最小key值对应主机的IP地址作为目的地址，将所述数据包发送至所述目的地址，并由该主机转发所述数据包。
- [61] 优选的，在所述支持哈希调度策略的lvs调度方法中，所述步骤S100之前还包括：
- [62] 步骤S10、主机通过Http和反向代理方式与Nginx物理服务器连接；其中，所述Nginx物理服务器为IMAP/POP3/SMTP 服务器。
- [63] 更具体的，所述步骤S10包括：
- [64] 1) Nginx物理服务器（即高性能的HTTP和反向代理服务器）中的LXC1-LXC8分别使用VF1-VF8连接至PF1-1，LXC1-LXC8还分别使用VF1-VF8连接只PF1-2；LXC9-LXC16分别使用VF1-VF8连接至PF2-1，LXC9-LXC16还分别使用VF1-VF8连接只PF2-2；其中PF1-1和PF2-1为VLAN External接口（即外部虚拟局域网接口）、PF1-2、PF2-2均为VLAN Internal接口（即内部虚拟局域网接口）
- [65] 2) Nginx物理服务器与IPSV（即IP虚拟服务器）通过后端服务器负载均衡器IPVS-DR连接；
- [66] 3、) IPSV通过Real IP与Tor建立BGP Peer，并发布VIP的主机路由，形成ECMP（即等价路由）后与CXP交换机连接，CXP交换机接收来自外网客户端的连接请求。通过上述方式，来实现lvs负载均衡。

- [67] 优选的，在所述支持哈希调度策略的lvs调度方法中，所述环形空间中每一ipv4的地址对应一个节点，每一节点均复制得到一个对应的虚拟节点。在一致性哈希算法中，为了尽可能的满足平衡性，其引入了虚拟节点。“虚拟节点”（virtual node）是实际节点（机器）在hash空间的复制（replica），一实际个节点（机器）对应了若干个“虚拟节点”，这个对应个数也成为“复制个数”，“虚拟节点”在hash空间中以hash值排列。通过虚拟节点的引入，对象的分布就比较均衡了。虚拟节点的hash计算可以采用对应节点的IP地址加数字后缀的方式。例如假设NODE1的IP地址为192.168.1.100。引入“虚拟节点”前，计算cache A的hash值：Hash(“192.168.1.100”);
- [68] 引入“虚拟节点”后，计算“虚拟节点”NODE1-1和NODE1-2的hash值：Hash(“192.168.1.100#1”); // NODE1-1，Hash(“192.168.1.100#2”); // NODE1-2。
- [69] 可见，本申请所述的支持哈希调度策略的lvs调度方法中，当有某个主机连接到Nginx物理服务器上时，调度到相邻hash节点上的连接漂移到加入的主机；当有某个主机退出时，调度到其他主机的会话不会漂移，只有调度到退出的主机上的会话会重新分布，实现了降低会话漂移，提升负载均衡服务质量。
- [70] 基于上述支持哈希调度策略的lvs调度方法，本申请还提供了一种支持哈希调度策略的lvs调度设备。如图4所示，所述支持哈希调度策略的lvs调度设备包括处理器11、存储器12及通信总线；
- [71] 所述通信总线用于实现处理器和存储器之间的连接通信；
- [72] 所述处理器用于执行存储器中存储的支持哈希调度策略的lvs调度程序，以实现以下步骤：
- [73] 获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；
- [74] 获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；
- [75] 获取环形空间中所存储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；
- [76] 当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。

- [77] 在本实施例中，所述的支持哈希调度策略的lvs调度程序10安装并运行于电子装置1中。所述电子装置1可以是桌上型计算机、笔记本、掌上电脑及服务器等计算设备。该电子装置1可包括，但不仅限于，存储器11、处理器12及显示器13。图4仅示出了具有组件11-13的电子装置1，但是应理解的是，并不要求实施所有示出的组件，可以替代的实施更多或者更少的组件。
- [78] 所述存储器11在一些实施例中可以是所述电子装置1的内部存储单元，例如该电子装置1的硬盘或内存。所述存储器11在另一些实施例中也可以是所述电子装置1的外部存储设备，例如所述电子装置1上配备的插接式硬盘，智能存储卡（Smart Media Card, SMC），安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）等。进一步地，所述存储器11还可以既包括所述电子装置1的内部存储单元也包括外部存储设备。所述存储器11用于存储安装于所述电子装置1的应用软件及各类数据，例如所述支持哈希调度策略的lvs调度程序的程序代码等。所述存储器11还可以用于暂时地存储已经输出或者将要输出的数据。
- [79] 所述处理器12在一些实施例中可以是一中央处理器（Central Processing Unit, CPU），微处理器或其他数据处理芯片，用于运行所述存储器11中存储的程序代码或处理数据，例如执行所述支持哈希调度策略的lvs调度程序10等。
- [80] 所述显示器13在一些实施例中可以是LED显示器、液晶显示器、触控式液晶显示器以及OLED（Organic Light-Emitting Diode，有机发光二极管）触摸器等。所述显示器13用于显示在所述电子装置1中处理的信息以及用于显示可视化的用户界面，例如应用菜单界面、应用图标界面等。所述电子装置1的部件11-13通过系统总线相互通信。
- [81] 进一步的，在所述支持哈希调度策略的lvs调度设备中，所述获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中的步骤中，所述处理器还用于执行所述支持哈希调度策略的lvs调度程序10，以实现以下步骤：
- [82] 获取多个主机的IP地址，并以每一主机的IP地址一一进行哈希运算得到对应的第一key值；
- [83] 获取ipv4的地址个数，并以ipv4的地址个数的两倍值作为空间大小设置一环形

空间；

[84] 将第一key值置于所述环形空间中。

[85] 进一步的，在所述支持哈希调度策略的lvs调度设备中，所述当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包的步骤中，所述处理器还用于执行所述支持哈希调度策略的lvs调度程序10，以实现以下步骤：

[86] 当多个第一key值中最小值大于第二key值时，则将多个第一key值中最小值记为第一最小key值，并获取与第一最小key值对应主机；

[87] 获取与第一最小key值对应主机的IP地址；

[88] 以与第一最小key值对应主机的IP地址作为目的地址，将所述数据包发送至所述目的地址，并由该主机转发所述数据包。

[89] 进一步的，在所述支持哈希调度策略的lvs调度设备中，所述获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中的步骤之前，所述处理器还用于执行所述支持哈希调度策略的lvs调度程序10，以实现以下步骤：

[90] 主机通过Http和反向代理方式与Nginx物理服务器连接；其中，所述Nginx物理服务器为IMAP/POP3/SMTP 服务器。

[91] 请参阅图5，是本申请安装支持哈希调度策略的lvs调度程序10较佳实施例的功能模块图。在本实施例中，所述的支持哈希调度策略的lvs调度程序10包括至少一个计算机程序指令段，该计算机程序指令段基于各部分实现的功能不同，可以分成一个或多个模块，所述一个或者多个模块被存储于所述存储器11中，并由一个或多个处理器（本实施例为所述处理器12）所执行，以完成本申请各实施例的支持哈希调度策略的lvs调度方法。例如，在图5中，所述的支持哈希调度策略的lvs调度程序10包括第一获取模块21、第二获取模块22、判断模块23及目的地址获取模块24；其中，第一获取模块21，用于获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；第二获取模块22，用于获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；判断模块23，用于获取环形空间中所存

储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；目的地址获取模块24，用于当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。

[92] 这些模块由处理器12执行，从而实现以下步骤：

[93] 获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；

[94] 获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；

[95] 获取环形空间中所存储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；

[96] 当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。具体实现与前面的实施例相同，在此不再赘述。

[97] 所述获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中的步骤包括：

[98] 获取多个主机的IP地址，并以每一主机的IP地址一一进行哈希运算得到对应的第一key值；

[99] 获取ipv4的地址个数，并以ipv4的地址个数的两倍值作为空间大小设置一环形空间；

[100] 将第一key值置于所述环形空间中。具体实现与前面的实施例相同，在此不再赘述。

[101] 所述当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包的步骤包括：

[102] 当多个第一key值中最小值大于第二key值时，则将多个第一key值中最小值记为第一最小key值，并获取与第一最小key值对应主机；

[103] 获取与第一最小key值对应主机的IP地址；

[104] 以与第一最小key值对应主机的IP地址作为目的地址，将所述数据包发送至所述目的地址，并由该主机转发所述数据包。具体实现与前面的实施例相同，在此不再赘述。

- [105] 所述获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中的步骤之前还包括：
- [106] 主机通过Http和反向代理方式与Nginx物理服务器连接；其中，所述Nginx物理服务器为IMAP/POP3/SMTP 服务器。具体实现与前面的实施例相同，在此不再赘述。
- [107] 所述环形空间中每一ipv4的地址对应一个节点，每一节点均复制得到一个对应的虚拟节点。具体实现与前面的实施例相同，在此不再赘述。
- [108] 基于上述支持哈希调度策略的lvs调度方法，本申请还提供了一种存储介质（及计算机可读存储介质）。所述存储介质存储有一个或者多个程序，所述一个或者多个程序可被一个或者多个处理器执行，以实现所述的支持哈希调度策略的lvs调度方法的步骤。
- [109] 基于上述支持哈希调度策略的lvs调度方法，本申请还提供了一种支持哈希调度策略的lvs调度装置，其特征在于，所述支持哈希调度策略的lvs调度装置包括：
- [110] 第一获取模块，用于获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；
- [111] 第二获取模块，用于获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；
- [112] 判断模块，用于获取环形空间中所存储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；
- [113] 目的地址获取模块，用于当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。
- [114] 进一步地，所述第一获取模块包括：
- [115] 哈希单元，用于获取多个主机的IP地址，并以每一主机的IP地址一一进行哈希运算得到对应的第一key值；
- [116] 第一获取单元，用于获取ipv4的地址个数，并以ipv4的地址个数的两倍值作为空间大小设置一环形空间；
- [117] 设置单元，用于将第一key值置于所述环形空间中。
- [118] 进一步地，所述目的地址获取模块包括：

- [119] 第二获取单元，用于当多个第一key值中最小值大于第二key值时，则将多个第一key值中最小值记为第一最小key值，并获取与第一最小key值对应主机；
- [120] 第三获取单元，用于获取与第一最小key值对应主机的IP地址；
- [121] 发送单元，用于以与第一最小key值对应主机的IP地址作为目的地址，将所述数据包发送至所述目的地址。
- [122] 进一步地，所述支持哈希调度策略的lvs调度装置还包括：
- [123] 连接模块，用于主机通过Http和反向代理方式与Nginx物理服务器连接；其中，所述Nginx物理服务器为IMAP/POP3/SMTP 服务器。
- [124] 其中，该支持哈希调度策略的lvs调度装置中各模块对应的实施例与上述基于支持哈希调度策略的lvs调度方法对应的各个实施例基本相同，故在此不再赘述。
- [125] 综上所述，本申请所提供的支持哈希调度策略的lvs调度方法、装置、设备及存储介质，当有某个主机连接到Nginx物理服务器上时，调度到相邻hash节点上的连接漂移到加入的主机；当有某个主机退出时，调度到其他主机的会话不会漂移，只有调度到退出的主机上的会话会重新分布，实现了降低会话漂移，提升负载均衡服务质量。
- [126] 可以理解的是，对本领域普通技术人员来说，可以根据本申请的技术方案及本申请构思加以等同替换或改变，而所有这些改变或替换都应属于本申请所附的权利要求的保护范围。

权利要求书

- [权利要求 1] 一种支持哈希调度策略的lvs调度方法，其特征在于，所述方法包括以下步骤：
- 获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；
- 获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；
- 获取环形空间中所存储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；
- 当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。
- [权利要求 2] 根据权利要求1所述支持哈希调度策略的lvs调度方法，其特征在于，所述获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中的步骤包括：
- 获取多个主机的IP地址，并以每一主机的IP地址一一进行哈希运算得到对应的第一key值；
- 获取ipv4的地址个数，并以ipv4的地址个数的两倍值作为空间大小设置一环形空间；
- 将第一key值置于所述环形空间中。
- [权利要求 3] 根据权利要求1所述支持哈希调度策略的lvs调度方法，其特征在于，所述当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包的步骤包括：
- 当多个第一key值中最小值大于第二key值时，则将多个第一key值中最小值记为第一最小key值，并获取与第一最小key值对应主机；
- 获取与第一最小key值对应主机的IP地址；
- 以与第一最小key值对应主机的IP地址作为目的地址，将所述数据包发送至所述目的地址。

- [权利要求 4] 根据权利要求2所述支持哈希调度策略的lvs调度方法，其特征在于，所述当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包的步骤包括：当多个第一key值中最小值大于第二key值时，则将多个第一key值中最小值记为第一最小key值，并获取与第一最小key值对应主机；获取与第一最小key值对应主机的IP地址；以与第一最小key值对应主机的IP地址作为目的地址，将所述数据包发送至所述目的地址。
- [权利要求 5] 根据权利要求1所述支持哈希调度策略的lvs调度方法，其特征在于，所述获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中的步骤之前还包括：
主机通过Http和反向代理方式与Nginx物理服务器连接；其中，所述Nginx物理服务器为IMAP/POP3/SMTP 服务器。
- [权利要求 6] 根据权利要求2所述支持哈希调度策略的lvs调度方法，其特征在于，所述获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中的步骤之前还包括：
主机通过Http和反向代理方式与Nginx物理服务器连接；其中，所述Nginx物理服务器为IMAP/POP3/SMTP 服务器。
- [权利要求 7] 根据权利要求2所述支持哈希调度策略的lvs调度方法，其特征在于，所述环形空间中每一ipv4的地址对应一个节点，每一节点均复制得到一个对应的虚拟节点。
- [权利要求 8] 一种支持哈希调度策略的lvs调度设备，其特征在于，所述支持哈希调度策略的lvs调度设备包括处理器、存储器及通信总线；
所述通信总线用于实现处理器和存储器之间的连接通信；
所述处理器用于执行存储器中存储的支持哈希调度策略的lvs调度程序，以实现以下步骤：

获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；
获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；
获取环形空间中所存储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；
当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。

[权利要求 9] 根据权利要求8所述支持哈希调度策略的lvs调度设备，其特征在于，所述获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中的步骤中，所述处理器还用于执行所述支持哈希调度策略的lvs调度程序，以实现以下步骤：

获取多个主机的IP地址，并以每一主机的IP地址一一进行哈希运算得到对应的第一key值；
获取ipv4的地址个数，并以ipv4的地址个数的两倍值作为空间大小设置一环形空间；
将第一key值置于所述环形空间中。

[权利要求 10] 根据权利要求8所述支持哈希调度策略的lvs调度设备，其特征在于，所述当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包的步骤中，所述处理器还用于执行所述支持哈希调度策略的lvs调度程序，以实现以下步骤：

当多个第一key值中最小值大于第二key值时，则将多个第一key值中最小值记为第一最小key值，并获取与第一最小key值对应主机；
获取与第一最小key值对应主机的IP地址；
以与第一最小key值对应主机的IP地址作为目的地址，将所述数据包发送至所述目的地址，并由该主机转发所述数据包。

- [权利要求 11] 根据权利要求8所述支持哈希调度策略的lvs调度设备，其特征在于，所述获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中的步骤之前还包括以下步骤：
主机通过Http和反向代理方式与Nginx物理服务器连接；其中，所述Nginx物理服务器为IMAP/POP3/SMTP 服务器。
- [权利要求 12] 一种存储介质，其特征在于，所述存储介质存储有一个或者多个程序，所述一个或者多个程序可被一个或者多个处理器执行，以实现如下步骤：
获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；
获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；
获取环形空间中所存储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；
当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。
- [权利要求 13] 根据权利要求12所述的存储介质，其特征在于，所述一个或者多个程序可被一个或者多个处理器执行，还实现如下步骤：
获取多个主机的IP地址，并以每一主机的IP地址一一进行哈希运算得到对应的第一key值；
获取ipv4的地址个数，并以ipv4的地址个数的两倍值作为空间大小设置一环形空间；
将第一key值置于所述环形空间中。
- [权利要求 14] 根据权利要求12所述的存储介质，其特征在于，所述一个或者多个程序可被一个或者多个处理器执行，还实现如下步骤：
当多个第一key值中最小值大于第二key值时，则将多个第一key值中最小值记为第一最小key值，并获取与第一最小key值对应主机；

获取与第一最小key值对应主机的IP地址；

以与第一最小key值对应主机的IP地址作为目的地址，将所述数据包发送至所述目的地址。

[权利要求 15] 根据权利要求12所述的存储介质，其特征在于，所述一个或者多个程序可被一个或者多个处理器执行，还实现如下步骤：

主机通过Http和反向代理方式与Nginx物理服务器连接；其中，所述Nginx物理服务器为IMAP/POP3/SMTP 服务器。

[权利要求 16] 根据权利要求13所述的存储介质，其特征在于，所述环形空间中每一ipv4的地址对应一个节点，每一节点均复制得到一个对应的虚拟节点。

[权利要求 17] 一种支持哈希调度策略的lvs调度装置，其特征在于，所述支持哈希调度策略的lvs调度装置包括：

第一获取模块，用于获取多个主机的IP地址，以每一主机的IP地址一一进行哈希运算得到对应的第一key值，将第一key值置于指定大小的环形空间中；

第二获取模块，用于获取数据包的源地址，并以数据包的源地址进行哈希运算得到第二key值；

判断模块，用于获取环形空间中所存储多个第一key值中最小值，并判断多个第一key值中最小值是否大于第二key值；

目的地址获取模块，用于当多个第一key值中最小值大于第二key值时，则以多个第一key值中最小值对应的主机作为目的地址发送所述数据包。

[权利要求 18] 根据权利要求17所述支持哈希调度策略的lvs调度装置，其特征在于，所述第一获取模块包括：

哈希单元，用于获取多个主机的IP地址，并以每一主机的IP地址一一进行哈希运算得到对应的第一key值；

第一获取单元，用于获取ipv4的地址个数，并以ipv4的地址个数的两倍值作为空间大小设置一环形空间；

设置单元，用于将第一key值置于所述环形空间中。

[权利要求 19] 根据权利要求17所述支持哈希调度策略的lvs调度装置，其特征在于，所述目的地址获取模块包括：

第二获取单元，用于当多个第一key值中最小值大于第二key值时，则将多个第一key值中最小值记为第一最小key值，并获取与第一最小key值对应主机；

第三获取单元，用于获取与第一最小key值对应主机的IP地址；

发送单元，用于以与第一最小key值对应主机的IP地址作为目的地址，将所述数据包发送至所述目的地址。

[权利要求 20] 根据权利要求17所述支持哈希调度策略的lvs调度装置，其特征在于，所述支持哈希调度策略的lvs调度装置还包括：

连接模块，用于主机通过Http和反向代理方式与Nginx物理服务器连接；其中，所述Nginx物理服务器为IMAP/POP3/SMTP 服务器。

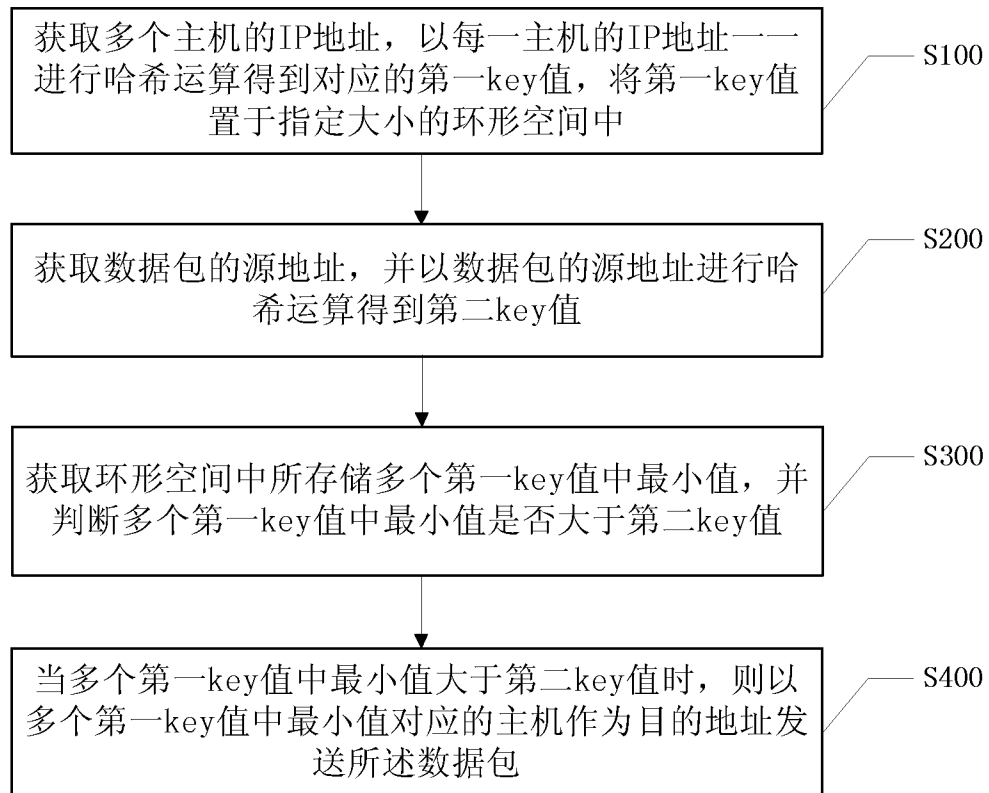


图 1

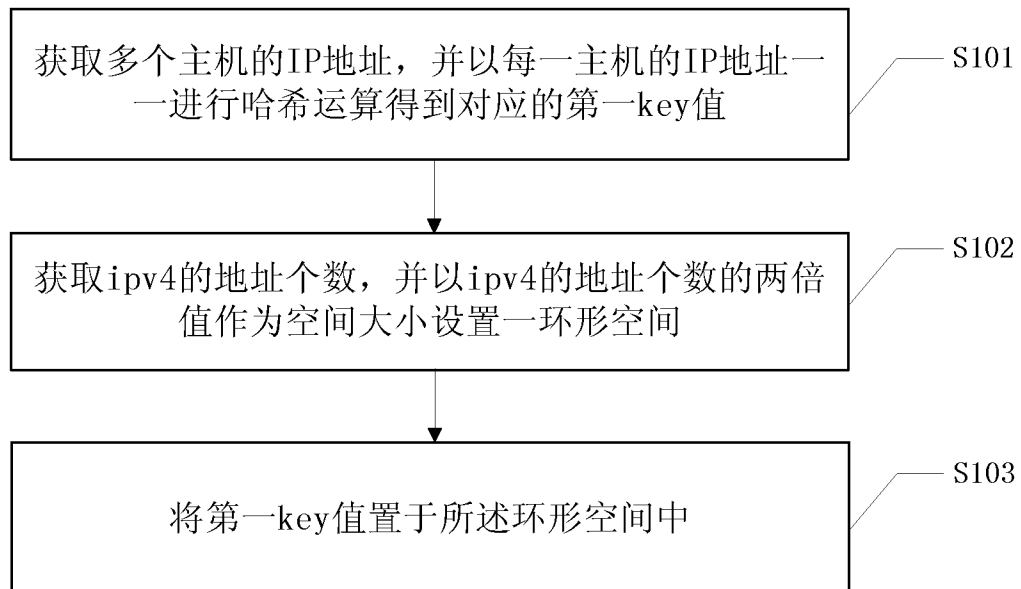


图 2

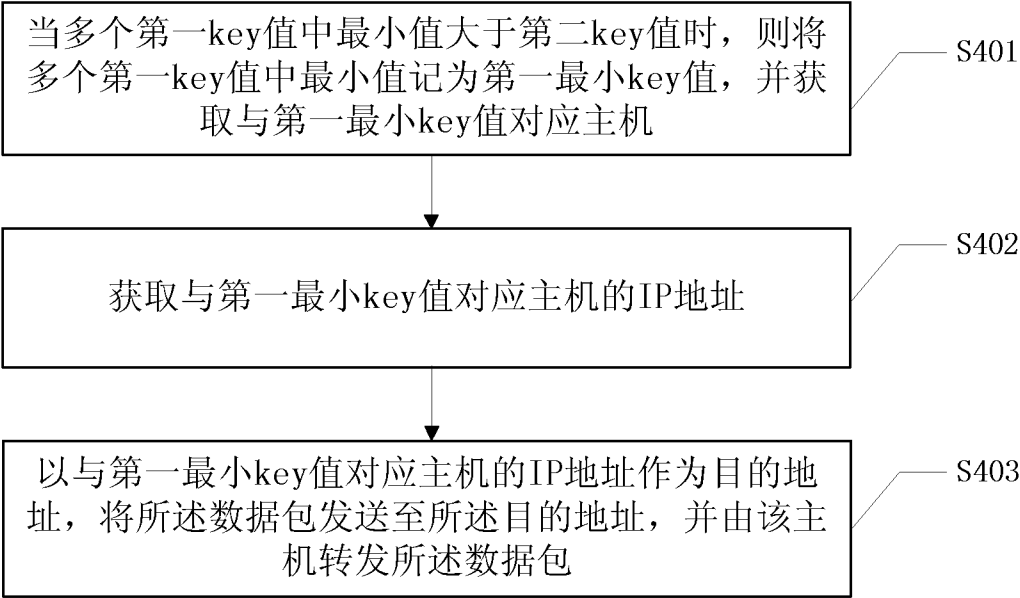


图 3

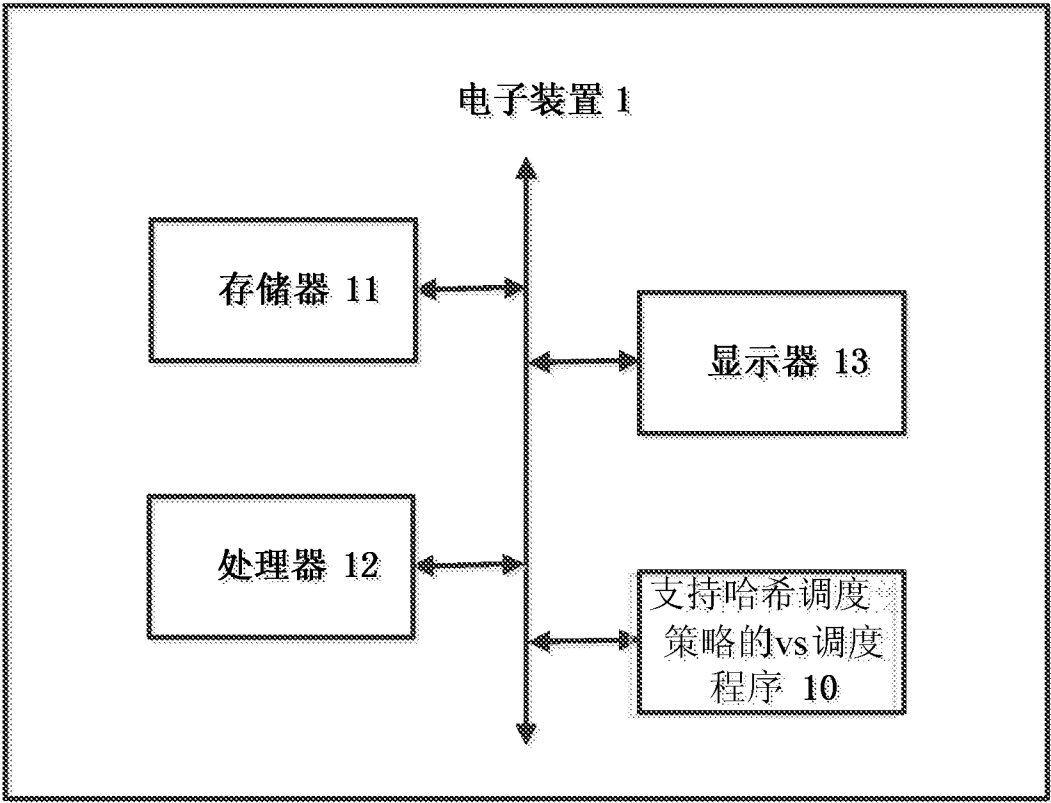
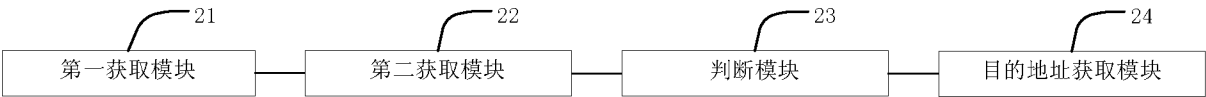


图 4



5

图 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/084470

A. CLASSIFICATION OF SUBJECT MATTER

H04L 29/08 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L; G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTEXT, CNABS, CNKI, DWPI, SIPOABS: 散列, 哈希, 虚拟, 复制, 环, 服务器, 主机, 地址, lvs, hash, schedul+, annular, ring, server, host

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104780201 A (SINA CORPORATION), 15 July 2015 (15.07.2015), description, paragraphs [0026]-[0059], and figures 1-6	1-20
X	CN 104852934 A (ALIBABA GROUP HOLDING LIMITED), 19 August 2015 (19.08.2015), description, paragraphs [0078]-[0123], and figures 1-2	1-20
X	CN 106856505 A (UESTC COMSYS INFORMATION CO., LTD.), 16 June 2017 (16.06.2017), description, paragraphs [0020]-[0042], and figures 1-6	1-20
PX	CN 107329827 A (PING AN TECH SHENZHEN CO., LTD.), 07 November 2017 (07.11.2017), description, paragraphs [0004]-[0058]	1-20

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 12 July 2018	Date of mailing of the international search report 18 July 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer WANG, Yue Telephone No. 62089109

International application No.
PCT/CN2018/084470

Form PCT/ISA/210 (patent family annex) (January 2015)

国际检索报告

国际申请号

PCT/CN2018/084470

A. 主题的分类 H04L 29/08 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类																	
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) H04L; G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNTXT, CNABS, CNKI, DWPI, SIPOABS: 散列, 哈希, 虚拟, 复制, 环, 服务器, 主机, 地址, lvs, hash, schedul+, annular, ring, server, host																	
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 104780201 A (新浪网技术中国有限公司) 2015年 7月 15日 (2015 - 07 - 15) 说明书第[0026]-[0059]段、图1-6</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 104852934 A (阿里巴巴集团控股有限公司) 2015年 8月 19日 (2015 - 08 - 19) 说明书第[0078]-[0123]段、图1-2</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>CN 106856505 A (成都康赛信息技术有限公司) 2017年 6月 16日 (2017 - 06 - 16) 说明书第[0020]-[0042]段、图1-6</td> <td>1-20</td> </tr> <tr> <td>PX</td> <td>CN 107329827 A (平安科技深圳有限公司) 2017年 11月 7日 (2017 - 11 - 07) 说明书第[0004]-[0058]段</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 104780201 A (新浪网技术中国有限公司) 2015年 7月 15日 (2015 - 07 - 15) 说明书第[0026]-[0059]段、图1-6	1-20	X	CN 104852934 A (阿里巴巴集团控股有限公司) 2015年 8月 19日 (2015 - 08 - 19) 说明书第[0078]-[0123]段、图1-2	1-20	X	CN 106856505 A (成都康赛信息技术有限公司) 2017年 6月 16日 (2017 - 06 - 16) 说明书第[0020]-[0042]段、图1-6	1-20	PX	CN 107329827 A (平安科技深圳有限公司) 2017年 11月 7日 (2017 - 11 - 07) 说明书第[0004]-[0058]段	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	CN 104780201 A (新浪网技术中国有限公司) 2015年 7月 15日 (2015 - 07 - 15) 说明书第[0026]-[0059]段、图1-6	1-20															
X	CN 104852934 A (阿里巴巴集团控股有限公司) 2015年 8月 19日 (2015 - 08 - 19) 说明书第[0078]-[0123]段、图1-2	1-20															
X	CN 106856505 A (成都康赛信息技术有限公司) 2017年 6月 16日 (2017 - 06 - 16) 说明书第[0020]-[0042]段、图1-6	1-20															
PX	CN 107329827 A (平安科技深圳有限公司) 2017年 11月 7日 (2017 - 11 - 07) 说明书第[0004]-[0058]段	1-20															
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																	
<table border="0"> <tr> <td> * 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 </td> <td> “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件 </td> </tr> </table>			* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件													
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																
国际检索实际完成的日期 2018年 7月 12日		国际检索报告邮寄日期 2018年 7月 18日															
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 王越 电话号码 62089109															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/084470

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	104780201	A	2015年 7月 15日	无	
CN	104852934	A	2015年 8月 19日	无	
CN	106856505	A	2017年 6月 16日	无	
CN	107329827	A	2017年 11月 7日	无	