

# (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

(43) 国际公布日  
2015年7月30日 (30.07.2015)



(10) 国际公布号  
WO 2015/109803 A1

- (51) 国际专利分类号:  
H04L 12/803 (2013.01)
- (21) 国际申请号: PCT/CN2014/083087
- (22) 国际申请日: 2014年7月25日 (25.07.2014)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
201410033741.X 2014年1月23日 (23.01.2014) CN
- (71) 申请人: 中兴通讯股份有限公司 (ZTE CORPORATION) [CN/CN]; 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦, Guangdong 518057 (CN)。
- (72) 发明人: 陶峯郡 (TAO, Quanjun); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦中兴通讯股份有限公司转交, Guangdong 518057 (CN)。 田甜 (TIAN, Tian); 中国广东省深圳市南山区高新技术产业园科技南路中兴通讯大厦中兴通讯股份有限公司转交, Guangdong 518057 (CN)。
- (74) 代理人: 北京安信方达知识产权代理有限公司 (AFD CHINA INTELLECTUAL PROPERTY LAW OFFICE); 中国北京市海淀区学清路8号B座1601A, Beijing 100192 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, NA, RW, SD, SL, 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)。

[见续页]

(54) Title: LOAD BALANCING METHOD AND SYSTEM

(54) 发明名称: 一种负荷均衡的方法及系统

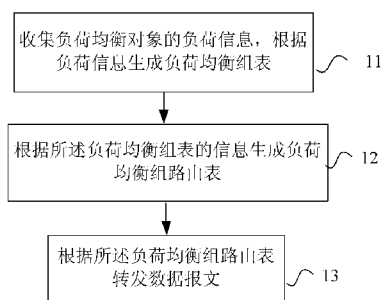


图 7 / FIG. 7

(57) Abstract: Provided are a load balancing method and system. The method comprises: collecting load information about a load balancing object, and according to the load information, generating a load balancing group table; according to information about the load balancing group table, generating a load balancing group routing table; and according to the load balancing group routing table, forwarding a data packet.

(57) 摘要: 本发明实施例提供一种负荷均衡的方法及系统, 该方法包括收集负荷均衡对象的负荷信息, 根据负荷信息生成负荷均衡组表; 根据所述负荷均衡组表的信息生成负荷均衡组路由表; 根据所述负荷均衡组路由表转发数据报文。

- 11 COLLECTING LOAD INFORMATION ABOUT A LOAD BALANCING OBJECT, AND ACCORDING TO THE LOAD INFORMATION, GENERATING A LOAD BALANCING GROUP TABLE
- 12 ACCORDING TO INFORMATION ABOUT THE LOAD BALANCING GROUP TABLE, GENERATING A LOAD BALANCING GROUP ROUTING TABLE
- 13 ACCORDING TO THE LOAD BALANCING GROUP ROUTING TABLE, FORWARDING A DATA PACKET



WO 2015/109803 A1

**根据细则 4.17 的声明:**

- 关于申请人有权申请并被授予专利(细则 4.17(ii))
- 发明人资格(细则 4.17(iv))

**本国际公布:**

- 包括国际检索报告(条约第 21 条(3))。

## 一种负荷均衡的方法及系统

### 技术领域

5 本发明涉及通讯领域，尤其是涉及一种通讯网元虚拟化负荷均衡控制方法。

### 背景技术

10 电信运营商的网络中有大量以及还在密集增长的各种专有的硬件设备，推出一个新的网络服务往往还需要增加更多的新的专属硬件设备，而放置这些设备的空间以及为它们进行供电变得越来越困难，相伴的困难同时还有能源成本的加剧上升，资本投资的挑战以及必要设计技能的缺乏，集成和操作变得日益复杂的硬件设备。此外，硬件设备的生命周期很短，这就需要大量的采购-设计-集成-部署进行几乎对收入无益的重复。更糟糕的是，硬件的生命周期开始变短，而服务的创新速度加快，这就抑制了从网络服务创新获得  
15 新增收入，且约束了日益网络化连接的世界的创新。

网络功能虚拟化（Network Function Virtualization，简称 NFV）的目的是利用 IT（Information Technology，信息技术）标准化的虚拟化技术，来将许多网络设备类型统一到能放置在数据中心、网络节点以及终端处的工业标准的大容量服务器、交换机和存储中来解决以上问题。

20 图 1 是相关技术中得到大多数研究者与研究机构认同的一种网络虚拟化架构示意图，其中：

网络管理和编排平台（Management & Orchestration Platform，简称 MOP）负责网络虚拟化运行环境的创建和管理，包括对网络功能虚拟化基础设施（Network Function Virtualization Infrastructure，简称 NFVI）域和网络功能虚  
25 拟化业务（Network Function Virtualization Service，简称 NVFS）域的协调管理；它包括网络功能虚拟化编排器（Network Function Virtualization Orchestration，简称 NFVO），及虚拟网络功能管理器（Virtual Network Function Management，简称 VNFM），其中 NFVO 主要负责全网业务编排，及网络资源管理，VNFM 主要负责虚拟网络功能（Virtual Network Function，简称 VNF）

资源管理;

云管理系统 (Cloud Management System, 简称 CMS), 主要功能是实现云的管理和监控;

5 传输网络管理系统(Transport Network Management System, 简称 TNMS), 主要功能对传输 IP 层网络管理和监控;

Hypervisor (程序管理)/虚拟机管理 (Virtual Machine Management, 简称 VMM) 主要功能是实现虚拟化能力, 将硬件资源虚拟为多个虚拟机 (Virtual Machine, 简称 VM);

10 运营支撑系统 (Operate Support System, 简称 OSS) /网元管理系统 (Element Management System, 简称 EMS) 为运营商网元管理后台, 负责对虚拟网元业务逻辑进行管理和监控;

虚拟网元功能 (Virtual Network Function, 简称 VNF), 通过虚拟机加载相应网元软件版本, 虚拟运营商网络中物理网元。

15 其中, 操作管理单元 (Operator Management Unit, 简称 OMU) 是网元运维管理单元, 在虚拟化部署场景下, OMU 需要增加和网管系统/OSS 的接口, 实现对所属的虚拟网元的虚拟机进行监控管理;

虚拟网元子功能组件 (Virtual Network Sub Function, 简称 VNSF) 为虚拟网元的软件子功能模块, 一个网元可以分解为多个子功能组件, 每个功能组件完成一个功能, 同类组件可以根据处理能力要求映射为一到多个 VM;

20 服务器, 存储设备以及传输网络设备为网络虚拟化基础设施, 主要为网络虚拟化提供硬件资源池。

在网元虚拟化场景, 为了提高 VNF 性能及可靠性, 通常需要采用多个 VNF 或者 VNF 内多个 VNFC 并发处理方式, 负载均衡 (Load Balance, 简称 LB) 解决问题的关键, 相关技术的 LB 机制, 基本架构如图 2 所示。

25 相关技术的架构中, 流量基本先汇总到 LB, 由 LB 根据负载均衡策略进行分流, 在未来云化网络模型中, VNF 可能无限扩容, 这样会导致 LB 成为网络的瓶颈。尤其后期 4G 移动网络转发面设备也进行虚拟化时, 转发面数据也要经过 LB。当然 LB 可以采用分层的方式, 这一方面会增加网络拓扑的

复杂度，另外也增加网络投资成本。

## 发明内容

5 本发明实施例要解决的技术问题是提供一种负荷均衡的方法及系统，以降低负荷均衡的负荷。

一种负荷均衡的方法，包括：

收集负荷均衡对象的负荷信息，根据负荷信息生成负荷均衡组表；

根据所述负荷均衡组表的信息生成负荷均衡组路由表；

根据所述负荷均衡组路由表转发数据报文。

10 较佳地，上述方法还可具有下面特点：所述负荷信息至少包括：

负荷均衡分组分支对象的处理能力信息或负荷权重信息；

负荷均衡分组分支对象数量；

负荷均衡分组分支对象的地址信息；

负荷均衡分组的数目；

15 负荷均衡分组的分组识别信息。

较佳地，上述方法还可具有下面特点：所述收集负荷均衡对象的负荷信息包括：

在分层组织架构模型下，分别收集各负荷均衡分组的负荷信息；

20 再汇总所有负荷均衡分组的负荷信息，并确定各负荷均衡分组的负荷权重信息。

较佳地，上述方法还可具有下面特点：所述收集负荷均衡对象的负荷信息包括：

在虚拟网络功能架构模型下，收集网络功能虚拟化编排器设置的虚拟网元间的负荷信息；

25 收集虚拟网元管理器设置的虚拟网元内各虚拟机间的负荷信息。

较佳地，上述方法还可具有下面特点：所述负荷均衡组路由表中的信息

至少包括:

分组标识、分组识别信息、子分支对象负载权重信息、子分组对象交换出口端口。

较佳地,上述方法还可具有下面特点:

- 5 所述分支对象负载权重信息由承载空间值表示,所述负荷均衡组路由表中的各分支对象按承载空间值从小至大的顺序排列,所述分支对象对应的负荷比例等于其对应的承载空间值减上一个分支对象对应的承载空间值,然后除以分组中最大的分支承载空间值。

较佳地,上述方法还可具有下面特点:

- 10 所述根据所述负荷均衡组路由表转发数据报文,包括:  
获取所述数据报文的用户标识;  
根据所述用户标识确定所述数据报文所属的承载空间值;  
根据所述承载空间值从所述负荷均衡路由表中选择对应的分组子分支对象转发所述数据报文。

- 15 较佳地,上述方法还可具有下面特点:所述根据所述用户标识确定所述数据报文所属的承载空间值,包括:

- 20 对所述用户标识进行哈希运算,取指定位置处与所述承载空间值相同位数的哈希值,并与所对应分组中各承载空间值进行比较,当所述哈希值小于第*i*分支对象的承载空间值,而大于第*i-1*分支对象的承载空间值,则确定所述数据报文所属的承载空间值为第*i*分支对象的承载空间值,*i*为正整数,*i=1*时,第0分支对象的承载空间值为0。

一种负荷均衡的系统,包括负荷均衡模块,网络控制器以及网络交换模块,其中:

- 25 负荷均衡模块设置为:收集负荷均衡对象的负荷信息,根据负荷信息生成负荷均衡组表;

网络控制器设置为:根据所述负荷均衡组表的信息生成负荷均衡组路由表,将所述负荷均衡路由表下发给交换模块;以及

网络交换模块设置为：根据所述负荷均衡组路由表转发数据报文。

较佳地，上述系统还可具有下面特点：所述负荷均衡模块收集的负荷信息至少包括：

负荷均衡分组分支对象的处理能力信息或负荷权重信息；

5 负荷均衡分组分支对象数量；

负荷均衡分组分支对象的地址信息；

负荷均衡分组的数目；和/或

负荷均衡分组的分组识别信息。

10 较佳地，上述系统还可具有下面特点：在在分层组织架构模型下，所述负荷均衡模块包括多个第一负荷均衡模块和第二负载均衡模块，

所述第一负荷均衡模块分别设置为：收集各负荷均衡分组的负荷信息，将收集到的负荷信息上报给第二负荷均衡模块；

所述第二负荷均衡模块设置为：汇总所有负荷均衡分组的负荷信息，并确定各负荷均衡分组的负荷权重信息。

15 较佳地，上述系统还可具有下面特点：在虚拟网络功能架构模型下，所述负荷均衡模块包括：

第一负荷均衡模块，设置为：收集网络功能虚拟化编排器设置的虚拟网元间的负荷信息；

20 第二负荷均衡模块，设置为：收集虚拟网元管理器设置的虚拟网元内各虚拟机间的负荷信息。

较佳地，上述系统还可具有下面特点：所述控制模块生成负荷均衡组路由表中的信息至少包括：

分组标识、分组识别信息、子分支对象负载权重信息和子分组对象交换出口端口，其中，

25 所述分支对象负载权重信息由承载空间值表示，所述负荷均衡组路由表中的各分支对象按承载空间值从小至大的顺序排列，所述分支对象对应的负荷比例等于其对应的承载空间值减去上一个分支对象对应的承载空间值，然

后除以分组中最大的分支承载空间值。

较佳地，上述系统还可具有下面特点：

所述网络交换模块，是设置为：获取所述数据报文的用户标识；根据所述用户标识确定所述数据报文所属的承载空间值；根据所述承载空间值从所述负荷均衡路由表中选择对应的分组子分支对象转发所述数据报文。

较佳地，上述系统还可具有下面特点：

所述网络交换模块是设置为以如下方式根据所述用户标识确定所述数据报文所属的承载空间值：对所述用户标识进行哈希运算，取指定位置处与所述承载空间值相同位数的哈希值，并与所对应分组中各承载空间值进行比较，当所述哈希值小于第  $i$  分支对象的承载空间值，而大于第  $i-1$  分支对象的承载空间值，则确定所述数据报文所属的承载空间值为第  $i$  分支对象的承载空间值， $i$  为正整数， $i=1$  时，第 0 分支对象的承载空间值为 0。

本发明实施例还提供一种计算机程序，包括程序指令，当该程序指令被负荷均衡系统执行时，使得该负荷均衡系统可执行上述方法。

本发明实施例还提供一种载有上述计算机程序的载体。

综上，本发明实施例提供的一种负荷均衡的方法及系统，引入转发和控制分离思想，数据流旁路 LB，LB 负责收集 VNF 的性能数据，并通过网络控制器北向接口动态更改交换机中数据流权重，进而降低 LB 负荷，避免 LB 成为网络瓶颈。

## 附图概述

图 1 是相关技术的网络虚拟化架构示意图；

图 2 是相关技术的负荷均衡系统的基本架构图；

图 3 是本发明实施例的负荷均衡系统的基本架构的示意图；

图 4 是本发明实施例的负荷均衡系统的示意图；

图 5 是本发明实施例的分层架构模型的示意图；

图 6 是本发明实施例的网络虚拟化架构模型的示意图；

图 7 为本发明实施例的负荷均衡的方法的流程图；

图 8 为本发明实施例的对应图 3 架构模型下的创建负荷均衡分组路由表的流程图；

5 图 9 为本发明实施例的对应图 5 架构模型下的创建负荷均衡分组路由表的流程图；

图 10 为本发明实施例的对应图 6 架构模型下的创建负荷均衡分组路由表的流程图；

图 11 是本发明一实施例的根据 LB 分组路由表路由的流程图；

10 图 12 是本发明另一实施例的根据 LB 分组路由表路由的流程图。

### 本发明的较佳实施方式

下文中将结合附图对本发明的实施例进行详细说明。需要说明的是，在不冲突的情况下，本申请中的实施例及实施例中的特征可以相互任意组合。

15 图 3 是本发明实施例的负荷均衡系统的基本架构的示意图，对比相关技术图 2，本发明实施例利用转发控制分离思想，使数据流旁路负荷均衡模块 LB，引入网络控制器（Controller），以降低负荷。

图 4 是本发明实施例的负荷均衡系统的示意图，如图 4 所示，本实施例的负荷均衡系统包括：负荷均衡模块、网络控制器、及网络交换模块。其中：

20 负荷均衡模块包括：数据收集单元、负荷均衡组组表生成单元以及负荷均衡组表修改单元，其中，

数据收集单元，设置为收集负荷均衡对象的负荷信息；

负荷均衡组表生成单元，设置为根据负荷信息生成所述负荷均衡组表，负荷均衡组表中包含：负荷均衡对象地址信息、权重信息及分组识别信息。

25 网络控制器，设置为根据收到的所述负荷均衡组表信息生成负荷均衡组路由表，并把所述负荷均衡组路由表下发至所述交换模块；

网络交换模块，设置为根据所述负荷均衡组路由表转发数据报文。

其中，所述收集单元所搜集的负荷信息包括，但不限于：  
所述负荷均衡分组分支对象的最大和最小处理能力信息；  
所述负荷均衡分组分支对象数量；  
所述负荷均衡分组分支对象的地址信息  
5 所述负荷均衡分组的分组数目；和/或  
所述负荷均衡分组识别信息。

其中，所述负荷均衡组路由表信息包括：分组标识，分组识别信息，和/或分组执行信息，所述分组执行信息可以包括：分组子分支对象的目标地址、子分支对象负载权重信息、子分支对象负载均衡策略、和/或子分组对象交换  
10 出口端口。

在一优选实施例中，在分层组织架构模型下，负荷均衡模块可以包括多个第一负荷均衡模块和一第二负荷均衡模块，其中，

所述第一负荷均衡模块分别设置为收集各负荷均衡分组的负荷信息，将收集到的负荷信息上报给第二负荷均衡模块；

15 所述第二负荷均衡模块，设置为汇总所有负荷均衡分组的负荷信息，并确定各负荷均衡分组的负荷权重信息。

图 5 是本发明实施例的分层架构模型的示意图，如图 5 所示，架构中负荷均衡模块分为两层，一层网络层负荷均衡模块（Net-LB），一层网元层负荷均衡模块（Vnf-LB），Net-LB 主要负责对虚拟网元进行负荷均衡操作，  
20 Vnf-LB 负责对虚拟网元内虚拟机 VM 的负荷均衡操作。Vnf-LB 负责对 VNF 内部的虚拟机负荷信息收集，Net-LB 负责对 VNF 整体负荷信息进行收集（如 VNF 内虚拟机的平均负荷）。分层架构模型中，两层的均衡模块进行数据收集有如下优点：Vnf-LB 可以降低 Net-LB 的负荷分担；Vnf-LB 可以屏蔽 VNF 内的虚拟机网络拓扑；单 VNF 内虚拟机负荷调整，不会波及到其他 Vnf-LB。

25 在另一优选实施例中，在虚拟网络功能架构模型下，所述负荷均衡模块包括：第一负荷均衡模块，设置为接收网络功能虚拟化编排器设置的虚拟网元间负荷均衡负荷信息，该负荷信息至少包含：网络分组数目、各分组中包含的虚拟化网元功能的数目、及各虚拟网元的能力信息或者负荷权重信息；

第二负荷均衡模块，设置为接收虚拟网元功能管理器设置的虚拟网元内各虚拟机间的负荷均衡负荷信息，所述负荷信息至少包含：虚拟网元功能中分组数目、各分组中包含的虚拟机数目、及各虚拟机的能力信息或者负荷权重信息。

5 图 6 是本发明实施例的网络虚拟化架构模型的示意图，如图 6 所示，Net-LB 与网络功能虚拟化编排器（NFVO）相连，Vnf-LB 与虚拟网元管理器（VNFM）相连，NFVO 负责向 Net-LB 设置 VNF 间的负荷均衡，VNFM 负责向 Vnf-LB 设置 VNF 内虚拟机间的负荷均衡。

10 图 7 为本发明实施例的负荷均衡的方法的流程图，如图 7 所示，本实施例的方法包括：

步骤 11、收集负荷均衡对象的负荷信息，根据负荷信息生成负荷均衡组表；

步骤 12、根据所述负荷均衡组表的信息生成负荷均衡组路由表；

步骤 13、根据所述负荷均衡组路由表转发数据报文。

15

本发明实施例还提供一种计算机程序，包括程序指令，当该程序指令被负荷均衡系统执行时，使得该负荷均衡系统可执行上述方法。

本发明实施例还提供一种载有上述计算机程序的载体。

20 图 8 是对应图 3 架构模型下的创建负荷均衡分组路由表的流程图，如图 8 所示，包括以下步骤：

步骤 701、负荷均衡模块（LB）收集负荷信息，该负荷信息可以包括：

1) 所需均衡的虚拟网元 VNF 支持最大和最小实例数（支持最大和最小实例数是处理能力一种体现）；

25 2) 同种 VNF 数量；

3) 所需均衡的 VNF 分组数量；

4) 所需均衡的 VNF 地址信息，如 IP 地址、MAC 地址等；

5) 分组识别信息, 主要用于交换模块收到数据报文根据所述分组识别信息判断其所属的 LB 分组, 分组识别信息根据所服务的业务类型, 可以设置为目的 IP 地址字段, 或目的 MAC (Media Access Control, 媒体接入控制) 地址或等等, 分组识别信息可以由网络运维管理者通过网管后台设置。

- 5 步骤 702、负荷均衡模块 (LB) 根据数据收集单元收集的负荷信息, 构建负荷均衡组表, 组表包括: 组标识、分组识别信息、分组分支选择策略、分组分支对象的地址信息 (包括 IP 地址和 MAC 地址信息)、分组分支的权重信息, 每条分支权重信息均可以包括一个承载空间值。承载空间值可以由 1 位或者多位数字构成, 并且分组内的各分支的承载空间按照递增的顺序进行排列, 比如某分组共有  $n$  条分支, 其中, 第  $i$  条分支承载空间值为  $L_i$ , 其代表  $L_{i-1}$  至  $L_i$  空间范围, 第一条分支的空间范围指 0 至  $L_1$  的空间范围。

本发明实施例的权重信息除了可以以承载空间方式体现, 也可以使用其他如百分比格式, 为了简化描述, 在本发明实施例中后续文档中选用了承载空间的方式;

- 15 LB 组表生成单元向网络控制器 Controller 请求创建 LB 分组路由表, 请求中携带负荷均衡组表信息。

步骤 703、网络控制器 (Controller) 收到请求后, 根据所述负荷均衡组表的信息生成负荷均衡组路由表;

- 20 网络控制器根据当前的网络拓扑信息和负荷均衡组表中 VNF 地址信息, 查找与 VNF 相关联的交换机 (Switch), 并为此交换机配置负荷均衡分组路由表, 路由表格式可以如表 1 所示。

- 25 负荷均衡分组路由表可以包括: 组标识信息、组分支选择策略、LB 组表分组识别信息、和/或 LB 组表执行出口信息。其中, LB 分组识别信息可以包括: 原目标 MAC 地址、原目标 IP 地址、和/或入口端口等, 这些信息主要被用于交换机识别分组; LB 分组执行出口信息包括: 分支目标 IP 地址、目标 MAC 地址、分支出口端口及权重信息, 这里权重信息以负载空间为例, 其中, 分支目标 IP 地址和 MAC 地址主要用于当 Switch 根据负荷均衡策略选择分支转发数据报文时, 根据需要可以把分支目标 IP 地址和 MAC 地址替换原报文中的目的 IP 地址和目的 MAC 地址。

LB 分组路由表中，每个分组可以对应 1 条或者多条分组识别信息，同时同一个组可以对应多条分支记录。

表 1

组标识	分支选择策略	LB 分组识别信息			LB 分组执行出口信息			
		原目标 MAC 地址	原目标 IP 地址	入口端口	目标 MAC 地址	目标 IP 地址	出口端口	负载空间
G1	1	00-1c-52-7b-3b-37	10.70.25.34	112	00-1c-52-7b-3b-01	192.0.0.1	1	30
					00-1c-52-7b-3b-02	192.0.0.2	2	50
					00-1c-52-7b-3b-03	192.0.0.3	3	70
					00-1c-52-7b-3b-04	192.0.0.4	4	100
G2	1	01-00-52-7b-3b-01	10.70.25.34	112	01-00-52-7b-3b-05	192.0.0.5	8	40
					01-00-52-7b-3b-06	192.0.0.6	7	60
		01-00-52-7b-3b-02	10.70.25.35	111	01-00-52-7b-3b-07	192.0.0.7	6	80
					01-00-52-7b-3b-08	192.0.0.7	5	100
...	...	...	...	...	...	...	...	

5 Controller 下发 LB 分组路由表至目标交换机 Switch 上。

步骤 704、交换模块存储收到的 LB 分组路由表，并向 Controller 返回 LB 分组路由表下发完成应答。

步骤 705、Controller 向 LB 返回 LB 分组路由表创建完成。

10 图 9 是对应图 5 架构模型下的创建负荷均衡分组路由表的流程图，包括如下步骤：

步骤 801、网元级负荷均衡模块 (Vnf-LB) 负责收集 VNF 网元内部各虚拟机 LB 负荷信息，包含的内容如步骤 701 所述，Vnf-LB 负责生成 VNF 网元内的 LB 分组表，表包含的信息参考步骤 702 所述；

Vnf-LB 向 Net-LB 上报其所收集的 LB 负荷信息。

15 步骤 802、Net-LB 收到其子级 Vnf-LB 上报的信息进行汇总，同种类型分组进行合并，并重新分配权重，如 VNF1 和 VNF2 之间属于负荷分担关系，Net-LB 把 VNF1 和 VNF2 内功能相同的分组进行合并，生成一个更大的分组。

步骤 803-步骤 806 同步骤 702-步骤 705。

步骤 807、Net-LB 向 Vnf-LB 返回 LB 分组设置成功。

图 10 为本发明实施例的对应图 6 架构模型下的创建负载均衡分组路由表的流程图，如图 10 所示，包括以下步骤：

- 5            步骤 901、NFVO 向 Net-LB 设置 VNF 间负载均衡负荷信息，包含 VNF 的分组，每个负荷信息包含：VNF 信息、各 VNF 支持的能力信息（如最大支持实例数、最大接入会话数等）及 VNF 对外呈现的地址信息等；

步骤 902-步骤 905 同步骤 702-步骤 705；

步骤 906、Net-LB 向 NFVO 返回 LB 设置完成信息；

- 10           步骤 907、VNFM 向 Vnf-LB 设置 VNF 内各虚拟机 VM 间负载均衡负荷信息，信息中包含：VM 的分组信息、每个分组包含的 VM 信息、各 VM 支持的能力信息及各 VM 的地址信息等；

步骤 908a、Vnf-Net 可以直接向 Controller 请求创建 LB 分组路由表；或者

- 15           步骤 908b-a-步骤 908b-b、Vnf-Net 可以经过 Net-LB 转发向 Controller 请求创建 LB 分组路由表；

步骤 909-步骤 910 同步骤 703-步骤 704；

步骤 911a、Controller 向 Vnf-LB 返回创建 LB 分组路由表完成；或者

- 20           步骤 911b-a-步骤 911b-b、Controller 向 Net-LB 返回创建 LB 分组路由表完成，然后 Net-LB 向 Vnf-LB 返回创建 LB 分组路由表完成；

步骤 912、Vnf-LB 向 VNFM 返回 LB 分组信息设置完成。

图 11 是本发明实施例的根据 LB 分组路由表路由的流程图，这里的 Switch 对应图 4 的交换模块，如图 11 所示，包括以下步骤：

步骤 1000、Switch 收到数据报文触发 LB 流程；

- 25           步骤 1001、switch 收到数据报文，首先根据 LB 分组路由表中的分组识别信息，判断当前数据报文所属的分组，如果没有找到相应的分组，则返回，退出 LB 处理流程；

分组识别信息包含所需检测的关键字段，如根据目标地址信息进行判断，或者根据目标端口等信息，switch 获取数据报文中对应的关键字段值与 LB 分组路由表中分组识别信息进行匹配，从而判断是否有所属的分组。

5 步骤 1002、Switch 根据分组负载均衡策略从 LB 分组路由表中选择分组子分支对象记录（即对应的路由），替换数据报文中的原目标地址为所述分组子分支对象记录中记载的目的地址信息，并按该记录中的交换出口端口转发所述数据报文。

替换目的地址信息这里可选，在特殊场景下可以不替换。

10 图 12 为本发明另一实施例的根据 LB 分组路由表路由的流程图，本实施例基于承载空间负载均衡算法确定路由，利用该算法可以用于交换模块保证同个用户的多个消息均从同一端口转发，本实施例中，交换模块收到数据报文，获取所述数据报文中的能区分用户的标识，并根据所述的标识计算其所属的承载空间，所述转发模块并根据承载空间选择分组分支对象的交换出口端口转发所述数据报文。

15 其中，所述能区分用户的标识可以由所述数据报文中的源 IP 地址，或者目的 IP 地址，或者报文载荷特点位置数据片段构成。

其中，承载空间是用于均衡分组中各子分支对象的负荷比例，由一位或者多位数字组成，所述分组分支对象对应的负荷比例等于其对应的承载空间值减去上一个分支的承载空间值，然后除以分组中最大的分支承载空间值。  
20 所述分组中，各分支对象的承载空间值是按从小至大的顺序排列，所述分组中的第一分支的负荷比例等于其对应的承载空间值减 0，除以承载空间总和。

例如，一负荷均衡分组中有 3 支分支对象，负荷比例分别为 20%、30%、50%，那么这 3 支分支对象的承载空间值分别为：20、50、100。

25 其中，确定所述数据报文所属的承载空间值包括，对所述用户标识进行哈希运算，取哈希值末尾与承载空间相同位数，并与所对应的分组中各承载空间值进行比较，当所述用户标识的哈希末尾值小于  $i$  分支对象的承载空间值  $LBS_i$ ，而大于所述  $i-1$  分支对象的承载空间值  $LBS_{i-1}$ ，则确定所述数据报文所属的承载空间值为  $i$  分支对象的承载空间值， $i=1$  时， $LBS_0=0$ 。

如图 12 所示，可以包括以下步骤：

步骤 1100-步骤 1101 同步骤 1000-步骤 1001；

步骤 1102、Switch 选择根据承载空间算法选择分组分支；

5 步骤 1103、Switch 获取数据报文中源 IP 地址信息，并对其进行哈希取模，  
取模的长度等于承载空间的长度；

步骤 1104、比较 LB 分组路由表对应组记录中分支承载空间值 ( $LBS_i$ )，  
如果取模值大于  $LBS_{i-1}$ ，而小于  $LBS_i$ ，则选择  $i$  分支， $i=1$  时， $LBS_0=0$ ；

10 步骤 1105、替换数据报文中的原目标地址为所述分支对象记录中记载的  
目的地址信息，并按该记录中的交换出口端口转发所述数据报文。所述目的  
地址信息，可以包括目的 IP 地址及目的 MAC 地址。

本领域普通技术人员可以理解上述方法中的全部或部分步骤可通过程序  
来指令相关硬件完成，所述程序可以存储于计算机可读存储介质中，如只读  
存储器、磁盘或光盘等。可选地，上述实施例的全部或部分步骤也可以使用  
15 一个或多个集成电路来实现。相应地，上述实施例中的各模块/单元可以采用  
硬件的形式实现，也可以采用软件功能模块的形式实现。本发明不限制于任  
何特定形式的硬件和软件的结合。

20 以上仅为本发明的优选实施例，当然，本发明还可有其他多种实施例，  
在不背离本发明精神及其实质的情况下，熟悉本领域的技术人员当可根据本  
发明作出各种相应的改变和变形，但这些相应的改变和变形都应属于本发明  
所附的权利要求的保护范围。

### 工业实用性

25 本发明实施例提供了一种负荷均衡的方法及系统，引入转发和控制分离  
思想，数据流旁路 LB，LB 负责收集 VNF 的性能数据，并通过网络控制器北  
向接口动态更改交换机中数据流权重，进而降低 LB 负荷，避免 LB 成为网络  
瓶颈进。

## 权 利 要 求 书

1、一种负荷均衡的方法，包括：

收集负荷均衡对象的负荷信息，根据所述负荷信息生成负荷均衡组表；

根据所述负荷均衡组表的信息生成负荷均衡组路由表；

5 根据所述负荷均衡组路由表转发数据报文。

2、如权利要求 1 所述的方法，其中：所述负荷信息至少包括：

负荷均衡分组分支对象的处理能力信息或负荷权重信息；

负荷均衡分组分支对象数量；

负荷均衡分组分支对象的地址信息；

10 负荷均衡分组的数目；和/或

负荷均衡分组的分组识别信息。

3、如权利要求 1 所述的方法，其中：所述收集负荷均衡对象的负荷信息包括：

在分层组织架构模型下，分别收集各负荷均衡分组的负荷信息；

15 再汇总所有负荷均衡分组的负荷信息，并确定各负荷均衡分组的负荷权重信息。

4、如权利要求 1 所述的方法，其中：所述收集负荷均衡对象的负荷信息包括：

20 在虚拟网络功能架构模型下，收集网络功能虚拟化编排器设置的虚拟网元间的负荷信息；

收集虚拟网元管理器设置的虚拟网元内虚拟机间的负荷信息。

5、如权利要求 1-4 任一项所述的方法，其中：所述负荷均衡组路由表中的信息至少包括：

25 分组标识、分组识别信息、子分支对象负载权重信息，和/或子分组对象交换出口端口。

6、如权利要求 5 所述的方法，其中：

所述分支对象负载权重信息由承载空间值表示，所述负荷均衡组路由表中的分支对象按承载空间值从小至大的顺序排列，所述分支对象对应的负荷比例等于所述分支对象对应的承载空间值减去上一个分支对象对应的承载空间值，然后除以分组中最大的分支承载空间值。

5 7、如权利要求6所述的方法，其中：所述根据所述负荷均衡组路由表转发数据报文，包括：

获取所述数据报文的用户标识；

根据所述用户标识确定所述数据报文所属的承载空间值；

10 根据所述承载空间值从所述负荷均衡路由表中选择对应的分组子分支对象转发所述数据报文。

8、如权利要求7所述的方法，其中：所述根据所述用户标识确定所述数据报文所属的承载空间值，包括：

15 对所述用户标识进行哈希运算，取指定位置处与所述承载空间值相同位数的哈希值，并与所对应分组中各承载空间值进行比较，当所述哈希值小于第*i*分支对象的承载空间值，而大于第*i-1*分支对象的承载空间值，则确定所述数据报文所属的承载空间值为第*i*分支对象的承载空间值，*i*为正整数，*i=1*时，第0分支对象的承载空间值为0。

9、一种负荷均衡的系统，包括负荷均衡模块，网络控制器以及网络交换模块，其中：

20 所述负荷均衡模块设置为：收集负荷均衡对象的负荷信息，根据负荷信息生成负荷均衡组表；

所述网络控制器设置为：根据所述负荷均衡组表的信息生成负荷均衡组路由表，将所述负荷均衡路由表下发给交换模块；所述网络交换模块设置为：根据所述负荷均衡组路由表转发数据报文。

25 10、如权利要求9所述的系统，其中：所述负荷均衡模块收集的负荷信息至少包括：

负荷均衡分组分支对象的处理能力信息或负荷权重信息；

负荷均衡分组分支对象数量；

负荷均衡分组分支对象的地址信息;

负荷均衡分组的数目; 和/或

负荷均衡分组的分组识别信息。

11、如权利要求 9 所述的系统, 其中: 在在分层组织架构模型下, 所述  
5 负荷均衡模块包括多个第一负荷均衡模块和第二负载均衡模块,

所述第一负荷均衡模块分别设置为收集各负荷均衡分组的负荷信息, 将收集到的负荷信息上报给所述第二负荷均衡模块;

所述第二负荷均衡模块设置为: 汇总所有负荷均衡分组的负荷信息, 并确定各负荷均衡分组的负荷权重信息。

10 12、如权利要求 9 所述的系统, 其中: 在虚拟网络功能架构模型下, 所述负荷均衡模块包括:

第一负荷均衡模块, 其设置为: 收集网络功能虚拟化编排器设置的虚拟网元间的负荷信息; 以及

15 第二负荷均衡模块, 其设置为: 收集虚拟网元管理器设置的虚拟网元内各虚拟机间的负荷信息。

13、如权利要求 9-12 任一项所述的系统, 其中: 所述网络控制器生成负荷均衡组路由表中的信息至少包括:

分组标识、分组识别信息、子分支对象负载权重信息和子分组对象交换出口端口, 其中,

20 所述分支对象负载权重信息由承载空间值表示, 所述负荷均衡组路由表中的分支对象按承载空间值从小至大的顺序排列, 所述分支对象对应的负荷比例等于所述分支对象对应的承载空间值减去上一个分支对象对应的承载空间值, 然后除以分组中最大的分支承载空间值。

14、如权利要求 13 所述的系统, 其中:

25 所述网络交换模块, 是设置为: 获取所述数据报文的用户标识; 根据所述用户标识确定所述数据报文所属的承载空间值; 根据所述承载空间值从所述负荷均衡路由表中选择对应的分组子分支对象转发所述数据报文。

15、如权利要求 14 所述的系统，其中：

所述网络交换模块是设置为以如下方式根据所述用户标识确定所述数据报文所属的承载空间值：对所述用户标识进行哈希运算，取指定位置处与所述承载空间值相同位数的哈希值，并与所对应分组中各承载空间值进行比较，  
5 当所述哈希值小于第  $i$  分支对象的承载空间值，而大于第  $i-1$  分支对象的承载空间值，则确定所述数据报文所属的承载空间值为第  $i$  分支对象的承载空间值， $i$  为正整数， $i=1$  时，第 0 分支对象的承载空间值为 0。

16、一种计算机程序，包括程序指令，当该程序指令被负荷均衡系统执行时，使得该负荷均衡系统可执行权利要求 1-8 任一项所述的方法。

10 17、一种载有权利要求 16 所述计算机程序的载体。

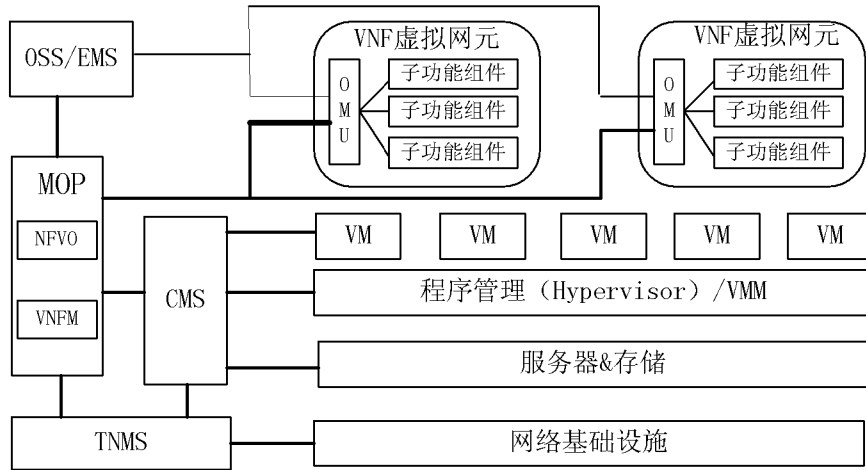


图 1

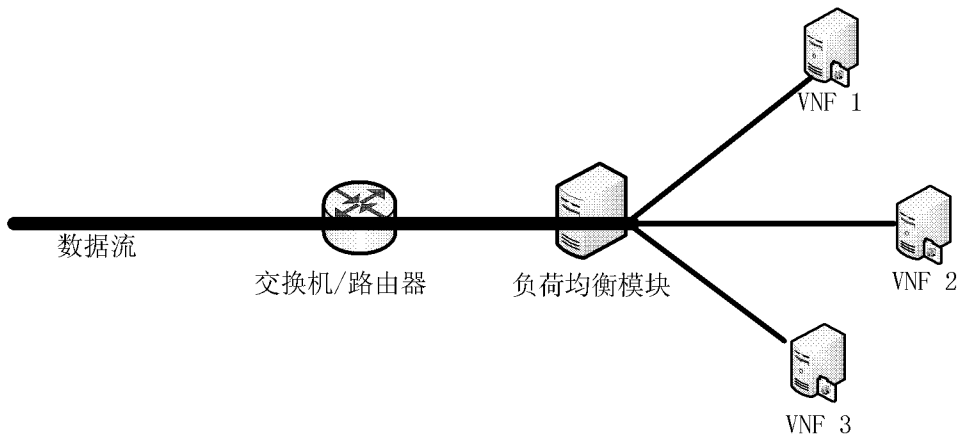


图 2

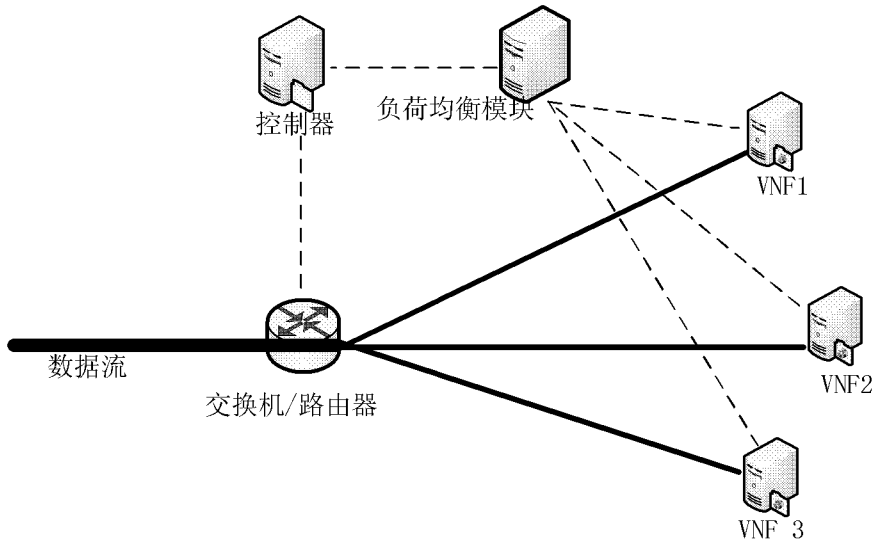


图 3

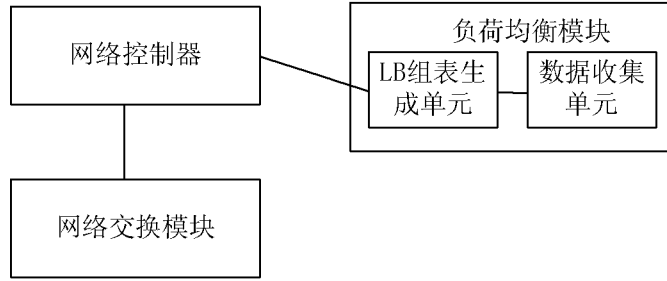


图 4

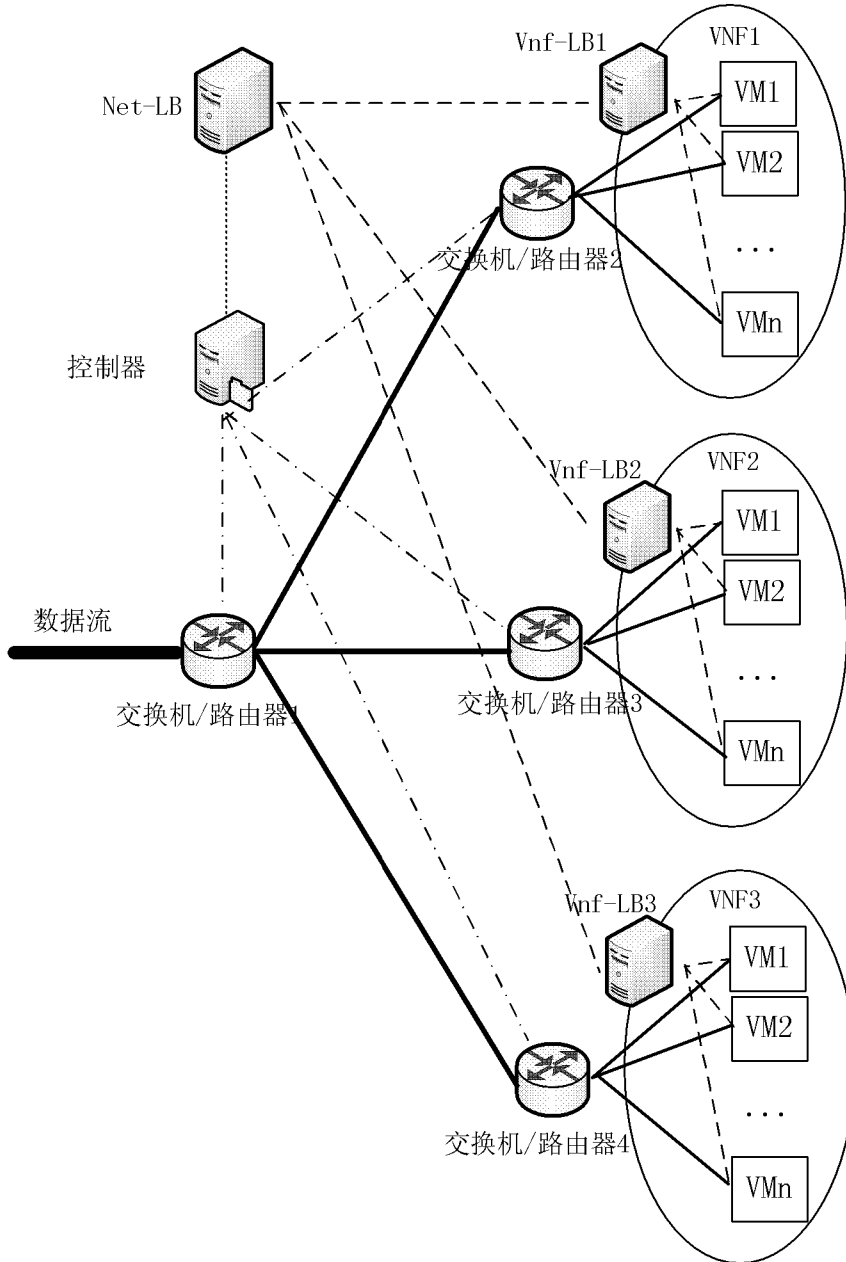


图 5

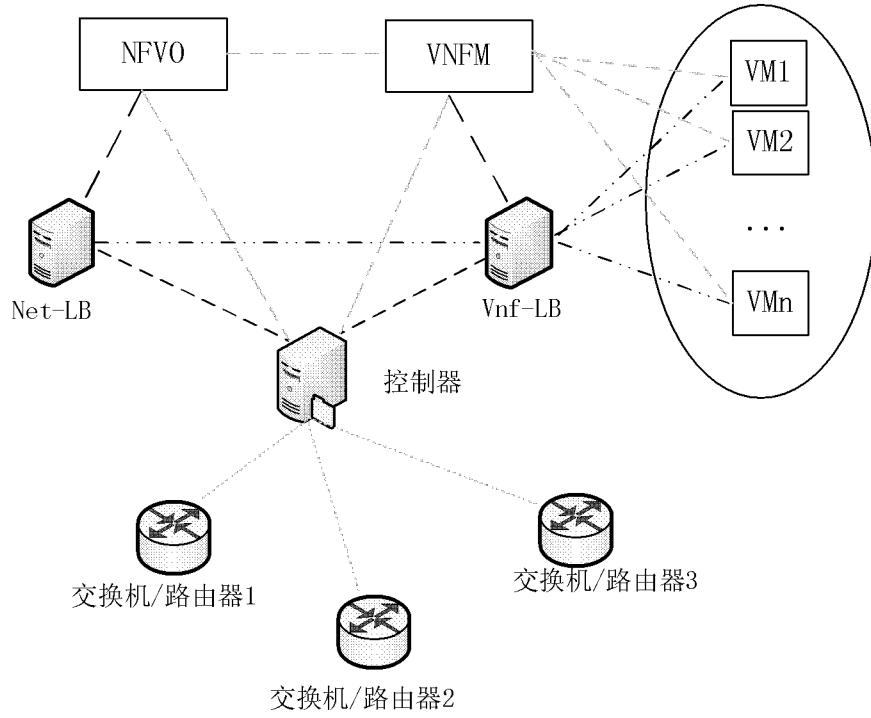


图 6

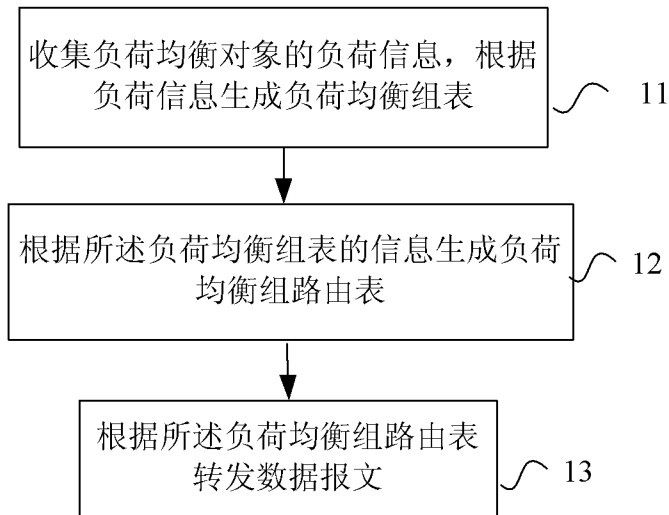


图 7

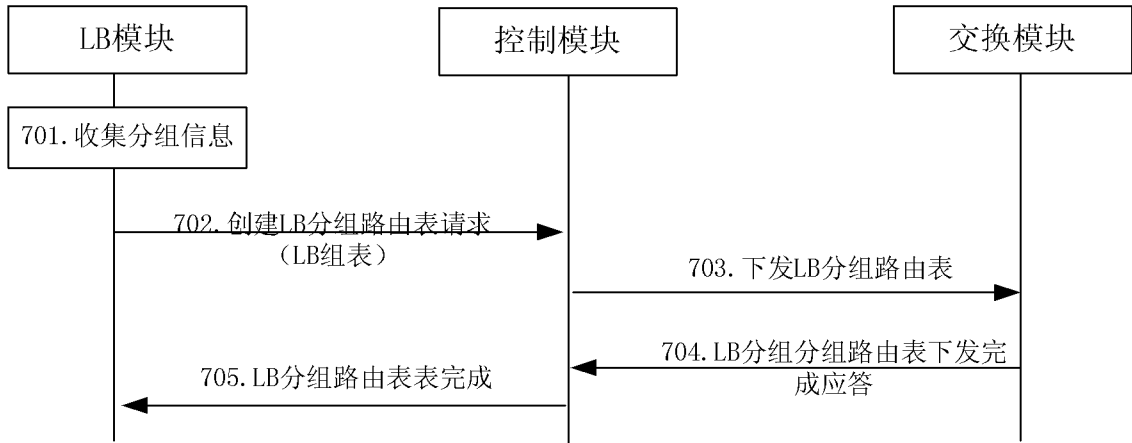


图 8

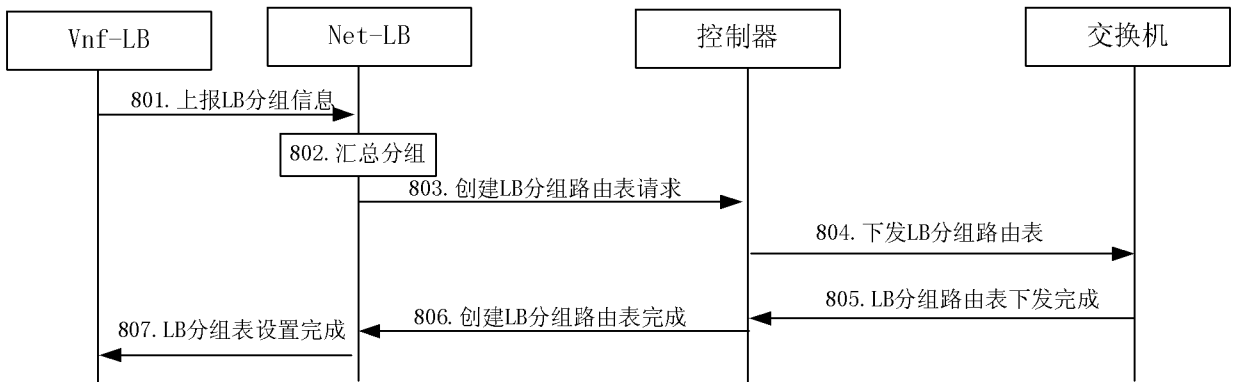


图 9

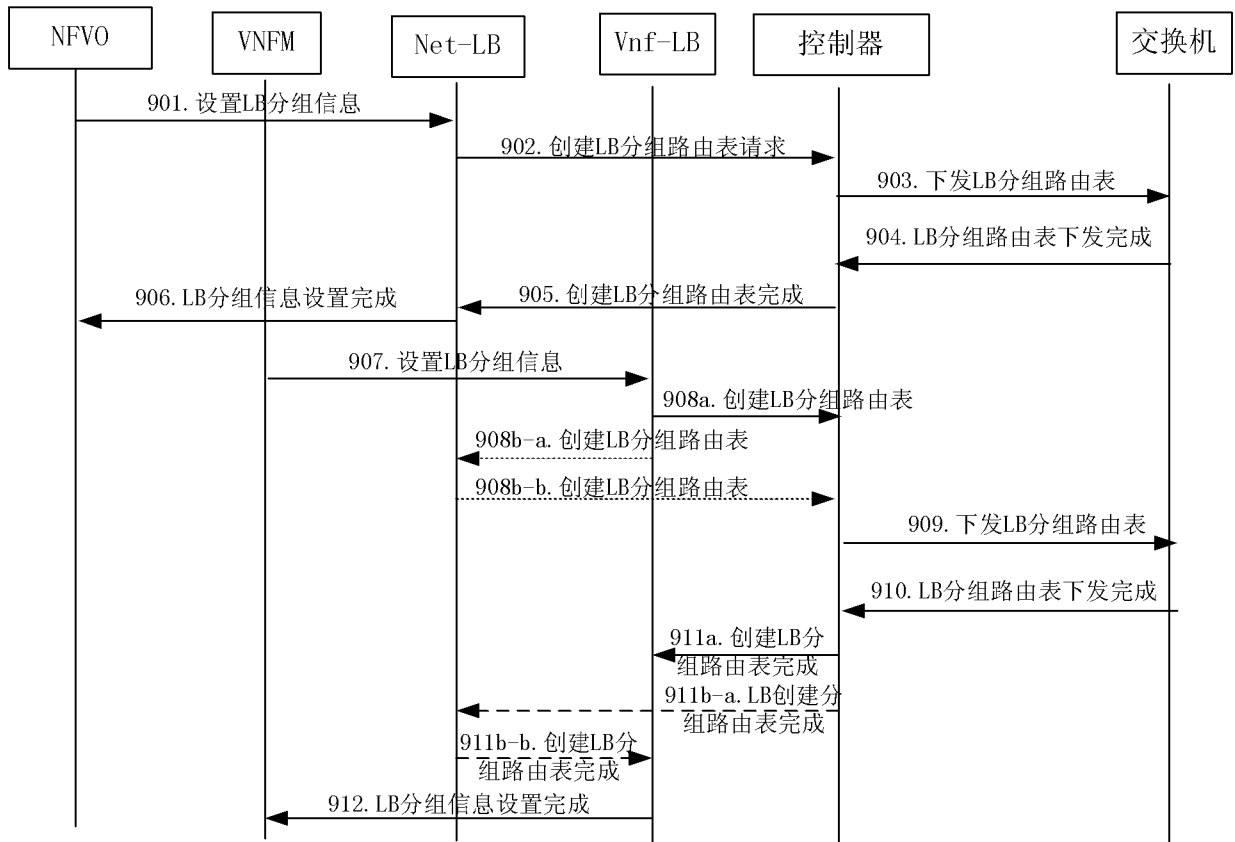


图 10

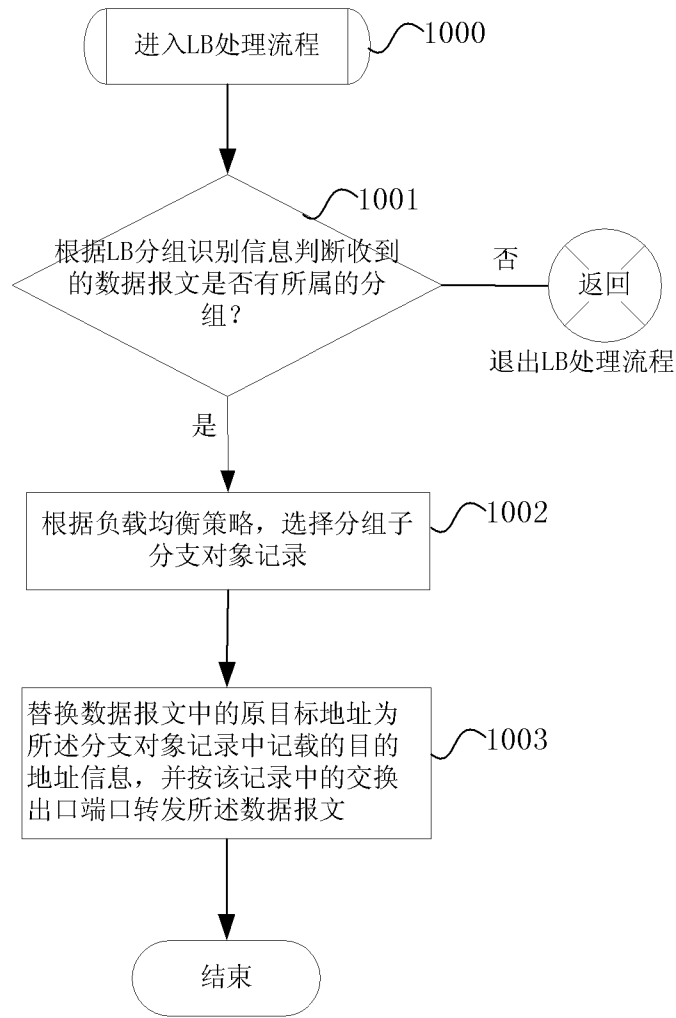


图 11

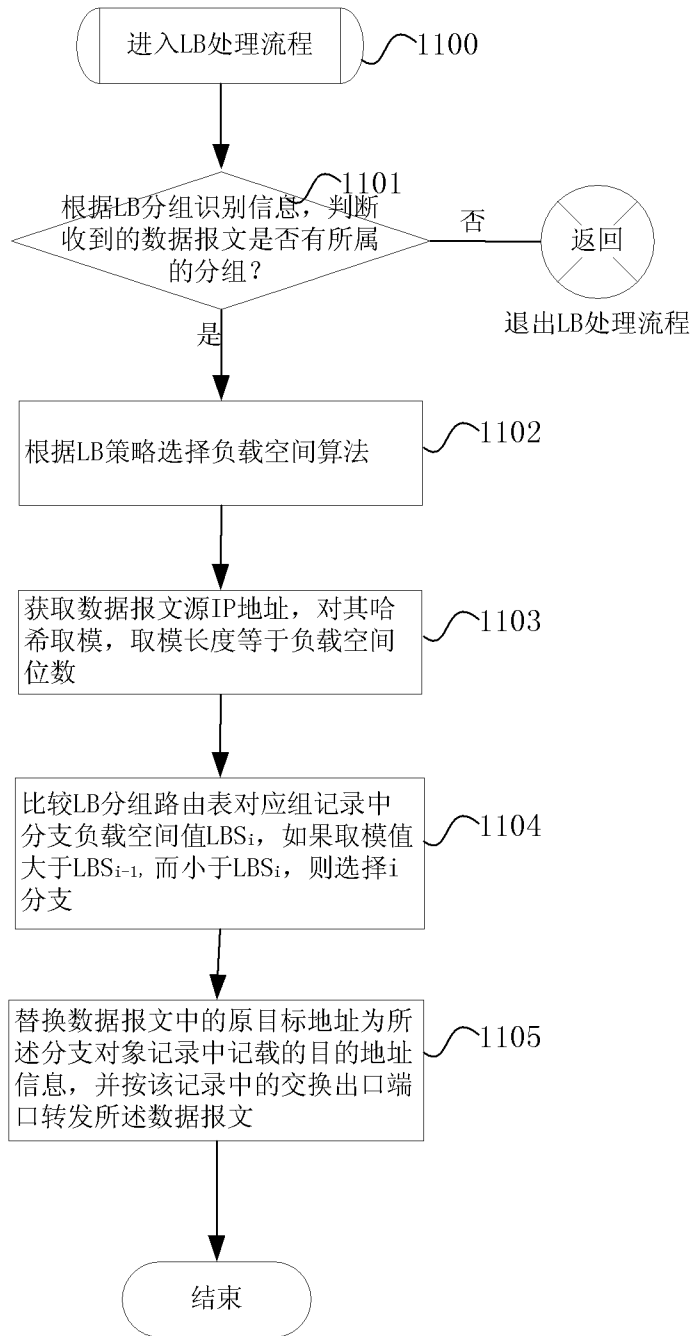


图 12

# INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2014/083087**

## A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/803 (2013.01) 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; H04W

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)

CPRSABS, CNTXT, CNKI, DWPI, SIPOABS: load, balance, collect, information, routing, route, select

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 1738291 A (UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA), 22 February 2006 (22.02.2006), description, page 3, line 9 to page 10, line 6, and figure 5	1-5, 9-12, 16-17
A	CN 1738291 A (UNIVERSITY OF ELECTRONIC SCIENCE AND TECHNOLOGY OF CHINA), 22 February 2006 (22.02.2006), the whole document	6-8, 13-15
A	CN 101447929 A (HUAWEI TECHNOLOGIES CO., LTD.), 03 June 2009 (03.06.2009), the whole document	1-17
A	CN 101478484 A (HUAWEI TECHNOLOGIES CO., LTD.), 08 July 2009 (08.07.2009), the whole document	1-17

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“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</p> <p>“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</p> <p>“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</p> <p>“&amp;” document member of the same patent family</p>
---	---

Date of the actual completion of the international search  
09 October 2014 (09.10.2014)

Date of mailing of the international search report  
**09 December 2014 (09.12.2014)**

Name and mailing address of the ISA/CN:  
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  
**LI, Yanxin**  
Telephone No.: (86-10) **62089375**

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.

**PCT/CN2014/083087**

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 1738291 A	22 February 2006	CN 100399772 C	02 July 2008
CN 101447929 A	03 June 2009	CN 101447929 B	08 June 2011
CN 101478484 A	08 July 2009	None	

国际检索报告

国际申请号

PCT/CN2014/083087

<p>A. 主题的分类</p> <p>H04L 12/803(2013.01)i</p> <p>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</p>																	
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04L; H04W</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CPRSABS, CNTXT, CNKI, DWPI, SIPOABS: 负荷, 负载, 均衡, 收集, 信息, 路由, 选择, load, balance, collect, information, routing, route, select</p>																	
<p>C. 相关文件</p> <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 1738291 A (电子科技大学) 2006年 2月 22日 (2006 - 02 - 22) 说明书第3页第9行至第10页第6行, 图5</td> <td>1-5, 9-12, 16-17</td> </tr> <tr> <td>A</td> <td>CN 1738291 A (电子科技大学) 2006年 2月 22日 (2006 - 02 - 22) 全文</td> <td>6-8, 13-15</td> </tr> <tr> <td>A</td> <td>CN 101447929 A (华为技术有限公司) 2009年 6月 03日 (2009 - 06 - 03) 全文</td> <td>1-17</td> </tr> <tr> <td>A</td> <td>CN 101478484 A (华为技术有限公司) 2009年 7月 08日 (2009 - 07 - 08) 全文</td> <td>1-17</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 1738291 A (电子科技大学) 2006年 2月 22日 (2006 - 02 - 22) 说明书第3页第9行至第10页第6行, 图5	1-5, 9-12, 16-17	A	CN 1738291 A (电子科技大学) 2006年 2月 22日 (2006 - 02 - 22) 全文	6-8, 13-15	A	CN 101447929 A (华为技术有限公司) 2009年 6月 03日 (2009 - 06 - 03) 全文	1-17	A	CN 101478484 A (华为技术有限公司) 2009年 7月 08日 (2009 - 07 - 08) 全文	1-17
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	CN 1738291 A (电子科技大学) 2006年 2月 22日 (2006 - 02 - 22) 说明书第3页第9行至第10页第6行, 图5	1-5, 9-12, 16-17															
A	CN 1738291 A (电子科技大学) 2006年 2月 22日 (2006 - 02 - 22) 全文	6-8, 13-15															
A	CN 101447929 A (华为技术有限公司) 2009年 6月 03日 (2009 - 06 - 03) 全文	1-17															
A	CN 101478484 A (华为技术有限公司) 2009年 7月 08日 (2009 - 07 - 08) 全文	1-17															
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。</p> <p><input checked="" type="checkbox"/> 见同族专利附件。</p>																	
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&amp;” 同族专利的文件</p>																	
<p>国际检索实际完成的日期</p> <p>2014年 10月 09日</p>		<p>国际检索报告邮寄日期</p> <p>2014年 12月 09日</p>															
<p>ISA/CN的名称和邮寄地址</p> <p>中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国</p> <p>传真号 (86-10)62019451</p>		<p>受权官员</p> <p>李彦欣</p> <p>电话号码 (86-10)62089375</p>															

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2014/083087

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	1738291	A	2006年 2月 22日	CN	100399772	C	2008年 7月 02日
CN	101447929	A	2009年 6月 03日	CN	101447929	B	2011年 6月 08日
CN	101478484	A	2009年 7月 08日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)