(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
02 July 2020 (02.07.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/135219 A1

(51) International Patent Classification:
H04L 29/06 (2006.01) *G06F 9/455* (2006.01)
H04L 12/741 (2013.01)

(21) International Application Number:
PCT/CN2019/126620

(22) International Filing Date:
19 December 2019 (19.12.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16/236,232 28 December 2018 (28.12.2018) US

(71) Applicant: **ALIBABA GROUP HOLDING LIMITED** [—/CN]; Fourth Floor, One Capital Place, P.O. Box 847, George Town, Grand Cayman (KY).

(72) Inventor: **CHENG, Gang**; Alibaba Group Legal Department 5/F, Building 3, No. 969 West Wen Yi Road, Yu Hang District, Hangzhou, Zhejiang 311121 (CN).

(74) Agent: **BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.**; 16th Fl., Block A, Corporate Square, No. 35 Jinrong Street, Beijing 100033 (CN).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: 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) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (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,

(54) Title: METHOD, APPARATUS, AND COMPUTER-READABLE STORAGE MEDIUM FOR NETWORK CONTROL

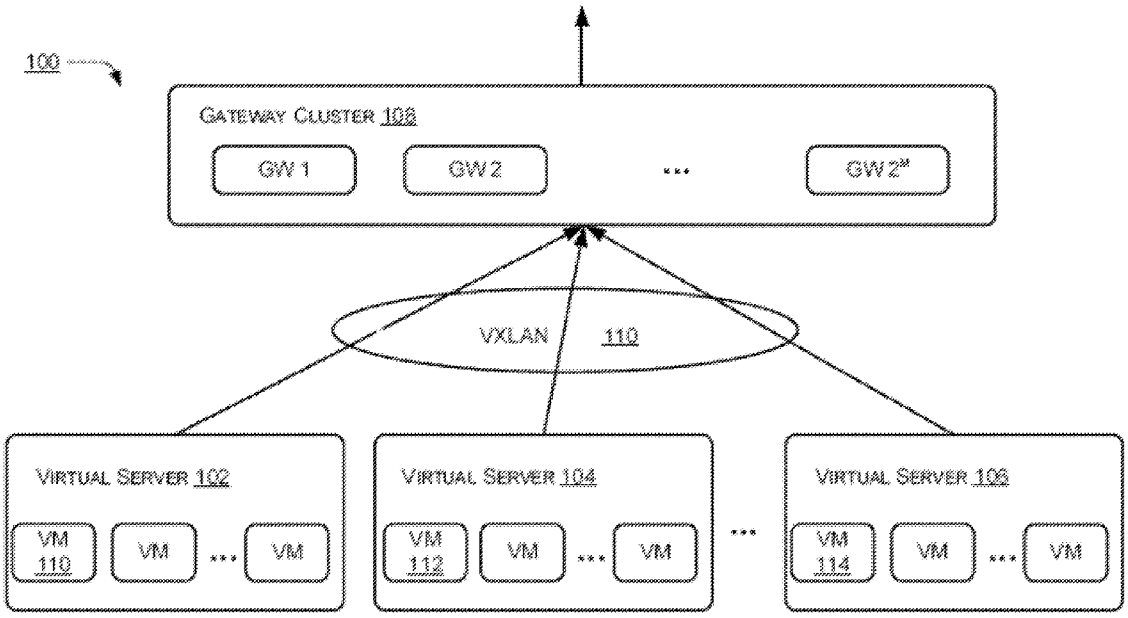

FIG. 1

(57) Abstract: Apparatuses and methods comprise receiving traffic from multiple virtual machines (VMs) associated with a same identification (ID); mapping the ID into a destination Internet Protocol (IP) address; and forwarding the traffic from the multiple VMs to a first gateway instance announcing the destination IP address.

WO 2020/135219 A1

[Continued on next page]

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

METHOD, APPARATUS, AND COMPUTER-READABLE STORAGE MEDIUM FOR NETWORK CONTROL

CROSS REFERENCE TO RELATED APPLICATION

[1] This disclosure claims the benefits of priority to United States application number 16/236,232, filed December 28, 2018, which is incorporated herein by reference in its entirety.

BACKGROUND

[2] A cloud computing architecture may be built up for multiple tenants. Each tenant is assigned its own network identification (tenant ID). Traditionally, virtual local area networks (VLANs) are used to isolate tenants in a cloud computing environment. However, VLAN specifications only allow for up to 4,096 network IDs to be assigned, which may not be sufficient for a large cloud computing environment.

[3] Virtual Extensible Local Area Network (VXLAN) is an encapsulation protocol that enables the cloud computing environment while logically isolating tenants. The primary goal of VXLAN is to extend the virtual VLAN address space by adding a 24-bit segment ID and increasing the number of available IDs to 16 million. VXLAN also allows a server to instruct network switches where to send packets. Each switch has software to receive and process instructions from the server. Packet-moving decisions and network traffic flow may be programmed for individual switches.

[4] VXLAN is also referred to as an overlay technology because it allows stretching Layer 2 connections over an intervening Layer 3 network by encapsulating (tunneling) data into a VXLAN packet. Devices that support VXLANs are called virtual tunnel endpoints (VTEPs), which encapsulate and de-encapsulate VXLAN traffic. The migration of virtual machines is enabled between servers that exist in separate Layer 2 domains by tunneling the

traffic over Layer 3 networks. This functionality allows the user to dynamically allocate resources within or between data centers without being constrained by Layer 2 boundaries or being forced to create large or geographically stretched Layer 2 domains. Using routing protocols to connect Layer 2 domains also allows load-balance of the traffic to ensure that the user gets the best use of available bandwidth.

[5] Each virtual private cloud (VPC) user has a unique VPC ID, and different users have different VPC IDs. Therefore, network traffics of different users may be isolated and secured. A user may have more than one virtual machines (VMs) hosted on different servers. A service provider needs to ensure that each VPC user does not consume resource such as bandwidth or dataflow more than purchased.

[6] In a gateway cluster, each gateway instance announces a range of Virtual Internet protocol (VIP) addresses to the network so that a packet with a destination IP address may be forwarded to the gateway instance based on in which range of VIP addresses the destination VIP address falls. Due to network load balancing, the traffic from a single VPC is distributed among different gateway instances. When one gateway instance is down, the packet may be randomly rerouted to a different gateway instance in the gateway cluster. Therefore, how to perform the traffic limitation of the same user becomes a problem.

BRIEF DESCRIPTION OF THE DRAWINGS

[7] The detailed description is set forth with reference to the accompanying figures. In the figures, the left-most digit(s) of a reference number identifies the figure in which the reference number first appears. The use of the same reference numbers in different figures indicates similar or identical items or features.

[8] FIG. 1 illustrates an example block diagram of multiple VMs of the same/single user hosted on different servers.

[9] FIG. 2 illustrates an example diagram of a gateway cluster.

[10] FIG. 3A illustrates an example process for throttling traffic from multiple VMs of a same/single user.

[11] FIG. 3B illustrates an example process for rerouting traffic when one or more gateway instances are down.

[12] FIG. 4 illustrates an example process for establishing the gateway cluster.

[13] FIG. 5 illustrates an example block diagrams of an apparatus for throttling traffic from multiple VMs of the same/single user.

DETAILED DESCRIPTION

[14] Apparatuses and methods discussed herein are directed to improving traffic control of cloud service, and more specifically to throttling traffic of multiple VMs of a same/single user.

[15] Apparatuses and methods discussed herein may be usable to design a highly scalable and available solution in which the total resource occupied by multiple VMs belonging to a same/single user is limited based on how much resource the user has purchased. Throughout the context, the user may not be limited to a person. The user may be an entity who purchases resources, such as a company, an institution, an organization, a human person, etc. Moreover, in case of a single or multiple gateway instances failure, the traffic may be rerouted, and thus the availability of the gateway cluster is improved.

[16] Traffic from multiple VMs associated with a same ID may be received. Traffic may refer to the amount of data moving across a network at a given point of time. Network data in networks may be encapsulated in packets, which provide the load in the network. The ID may be mapped into a destination IP address. The traffic from the multiple VMs may be encapsulated using the destination IP address. The traffic from the multiple VMs may be forwarded to a first gateway instance announcing the destination IP address. The traffic from the multiple VMs may be monitored by determining whether a traffic indicator exceeds a threshold. Upon determining that the traffic indicator exceeds the threshold, the traffic from the multiple VMs may be throttled.

[17] The traffic indicator may indicate a bandwidth occupied by the traffic of the multiple VMs.

[18] The traffic indicator may indicate a dataflow of the multiple VMs.

[19] Mapping the ID into the destination IP address may include hashing the ID into a whole IP address range of a gateway cluster. The gateway cluster may include the first gateway instance.

[20] The whole IP address range may include a plurality of IP address sub-ranges. The gateway cluster may include a plurality of gateway instances. Each gateway instance of the plurality of gateway instance may announce an IP sub-range of the plurality of IP address sub-ranges.

[21] Whether the first gateway instance is down may be determined. For example, the gateway instance may fail, crash, be busy, etc. If the attempt of routing the traffic to the first gateway instance is unsuccessful, this indicates that the first gateway instance may be down. Upon determining that the first gateway instance is down, the traffic from the multiple VMs may be rerouted to a second gateway instance in the gateway cluster.

[22] FIG. 1 illustrates an example block diagram 100 showing multiple VMs of the same/single user hosted on different servers. Referring to FIG. 1, each of the three virtual servers 102, 104, and 106 shown in this example, hosts multiple virtual machines (VMs). Though FIG. 1 shows three servers, other numbers of servers may be provided. Each of the three virtual servers 102, 104, and 106 may serve a plurality of VMs. Each of the virtual servers 102, 104, and 106 may route traffic from different VMs to a gateway cluster 108 through the VXLAN 110. A user may have more than one VMs hosted on different servers. For example, the VM 110 hosted on the virtual server 102, the VM 112 hosted on the virtual server 104, and the VM 114 hosted on the virtual server 106 may belong to the same/single user with a single user/tenant ID. The user may be a VPC user, and the user/tenant ID may be a VPC ID.

[23] FIG. 2 illustrates an example diagram of the gateway cluster 108. The total number of gateway instances in the gateway cluster 108 may be of any positive value based on hardware and/or software configurations associated with the gateway cluster. By way of example and not limitation, FIG. 2 shows the total number of gateway instances in the gateway cluster 108 to be a power of 2, for example, 2^m, where the number m may be an integer. Gateway instances, GW1, GW2, GW 3, GW 4,... $GW2^{m-1}$, and GW $2^m$are shown in

this example, each of which may announce a part of the whole VIP address range. Though six gateway instances are shown in this example, other numbers of gateway instances may be provided. The structure of the (m+1) level full binary tree may be described as follows.

[24] A Level-0 root node may be provided. The Level-0 root node may have a corresponding IP address range which is the whole VIP address range of $[1, 2^k]$. The Level-0 root node 202 may be split into two Level-1 nodes 204 and 206. The left child Level-1 node 204 may have a corresponding IP address range of $[1, 2^{k-1}]$, while the right child Level-1 node 206 may have a corresponding IP address range of $[2^{k-1}+1, 2^k]$. The number k may be an integer, and 2^k may represent the IP address space to be accessed.

[25] The Level-1 node 204 may be split into two Level-2 nodes 208 and 210. The Level-1 node 206 may be split into two Level-2 nodes 212 and 214. The Level-2 node 208 may have a corresponding IP address range of $[1, 2^{k-2}]$. The Level-2 node 210 may have a corresponding IP address range of $[2^{k-2}+1, 2^{k-1}]$. The Level-2 node 212 may have a corresponding IP address range of $[2^{k-1}+1, 2^{k-1}+2^{k-2}]$. The Level-2 node 214 may have a corresponding IP address range of $[2^{k-1}+2^{k-2}+1, 2^k]$. The node splitting and the sub-range dividing may be repeated in an iterative manner.

[26] Each Level-(m-1) node may be split into two Level-m nodes. The Level-(m-1) node 216 may be split into two Level-m nodes GW1 and GW2. The Level-(m-1) node 218 may be split into two Level-m nodes GW3 and GW4. The Level-(m-1) node 220 may be split into two Level-m nodes $GW2^{m-1}$ and $GW2^m$. Each of the gateway instances, GW1, GW2, GW 3, GW 4, …,$GW2^{m-1}$,and GW 2^m, may be a leaf node of the binary tree. The gateway instance GW1 may have a corresponding IP address range of $[1, 2^{k-m}]$. The gateway instance GW2 may have a corresponding IP address range of $[2^{k-m}+1, 2^{k-m+1}]$. The gateway instance $GW2^m$may have a corresponding IP address range of $[2^{k-1}+2^{k-2}+\ldots+2^{k-m}+1, 2^k]$.

[27] Each of the gateway instances, GW1, GW2, GW 3, GW 4, …,$GW2^{m-1}$,and GW 2^m, may have a single path to reach the Level-0 root node 202. In such a path, there are(m+1) nodes, including the gateway instance node, representing (m+1) IP address ranges. Hence, each of the gateway instances, GW1, GW2, GW3, GW 4, …, $GW2^{m-1}$,and GW 2^m, may announce (m+1) IP address ranges to the network. For instance, the gateway instance GW1 may announce (m+1) address ranges $[1, 2^{k-m}]$, $[1, 2^{k-m+1}]$, $[1, 2^{k-m+2}]$, …, $[1, 2^k]$.

[28] In the above gateway cluster 108, m is an integer and may be selected based on the number of the gateway instances. K is an integer and may be selected based on the IP address range to be accessed. Because m and k may be selected based on actual needs, the gateway cluster may be scalable and expandable. Moreover, in case of failure of one or more gateway instances, the traffic may be rerouted to a gateway instance under the same upper-level parent node. Thus, the availability of the gateway cluster 108may be improved.

[29] The above manner of dividing the IP address range or splitting nodes is merely an example, and there may be other ways of division/splitting, for example, an upper-level node may be split into three, four, or other number of lower-level nodes.

[30] FIG. 3A illustrates an example process 300 for throttling traffic from multiple VMs of a same/single user.

[31] At block 302, one or more servers may receive traffic from multiple VMs of the same/single user, for example, a VPC user with the same VPC ID. For example, referring to FIG. 1, the server 102 may receive traffic from the VM 110, the server 104 may receive traffic from the VM 112, and the server 106 may receive traffic from the VM 114, where the VMs 110, 112, and 114 may be associated with the same ID. Though three VMs are used here, other numbers of VMs may be provided.

[32] At block 304,the one or more servers may obtain a destination IP address for the traffic by mapping the ID, for example, the VPC ID, into the whole VIP address range of the gateway cluster 108. Also referring to FIG. 1, the server 102 may obtain the destination IP address for the traffic from the VM 110.The server 104 may obtain the destination IP address for the traffic from the VM 112.The server 106 may obtain the destination IP address for the traffic from the VM 114.Because VMs 110, 112, and 114 are associated with the same ID, the destination IP address for the traffic from VMs 110, 112, and 114 may be the same. Though three VMs are used here, other numbers of VMs may be provided.

[33] The mapping may be performed by hashing the ID into the whole VIP address range of the gateway cluster 108. Other methods of mapping may also be used as long as the same ID corresponds to or is mapped to the same destination IP address after the mapping. For example, a table may be used to keep the corresponding relationship between IDs and destination IP addresses.

[34] As an alternative embodiment, the same ID may be mapped to the same address sub-range served by a same gateway instance. In that case, the ID may be mapped into one of a group of address sub-ranges, where each address sub-range correspond to a gateway instance.

[35] As another alternative embodiment, the same ID may be mapped to the same gateway addressor the same gateway instance ID. In that case, the ID may be mapped into one of a group of gateway instance IDs, where each gateway instance ID correspond to a gateway instance. Other methods of mapping between IDs and gateway instances may be used as long as the traffic from different VMs associated with the same ID may go through the same gateway instance such that the gateway instance may be able to monitor and throttle/control the traffic of the same user.

[36] By way of example and not limitation, the IDs may be mapped to a group of one or more gateway instances which can serve and handle the traffic of the same user (i.e., traffic having the same user ID) using one of the mapping methods (for example, hashing, etc.) described in the foregoing description, and one of the one or more gateway instances may be selected to serve and handle the traffic of the same user in a rotating manner.

[37] Alternatively, after the traffic of the same ID is mapped into a determinative group of one or more gateway instances, one of the one or more gateway instances within this particular group may be randomly selected to serve and handle the traffic of the same user.

[38] Using a determinative group of one or more gateway instances to serve and handle all the traffic having a same ID (e.g., a same user ID) helps alleviating the excessive workload that a single gateway instance may handle for the traffic coming from a same user, and improves the speed of transmitting and handling the traffic of the same user through the gateway, for example. This is especially true when the user is a large entity such as an enterprise, an institution, etc., in which the amount of traffic coming from the user is normally large at certain period of time.

[39] In implementations, whether traffic of a same user is mapped into a determinative group of gateway instances or a single gateway instance may depend on the amount of bandwidth that this user purchases, or the amount of traffic that the user is sending and/or receiving, etc. For example, a group of gateway instances may be used and mapped

into if the amount of bandwidth that this user purchases or the amount of traffic that the user is sending and/or receiving is larger than a certain threshold (such as 100MB, 500GB, 1GB, etc.)

[40] At block 306, the traffic from multiple VMs of the same/single user may be encapsulated using the destination IP address obtained at block 304. Packets carried in the traffic may be encapsulated based on tunneling protocols, for example, VXLAN/NvGRE/Genve. Referring to FIG. 1, the server 102 may encapsulate traffic from the VM 110, the server 104 may encapsulate traffic from the VM 112, and the server 106 may encapsulate traffic from the VM 114. Because the VMs 110, 112, and 114 are associated with the same ID, and the same ID may be mapped to the same destination IP address as discussed above, the traffic from the same user may reach the same gateway instance in the gateway cluster 108.Traffic from different VMs associated with different IDs may be routed to different gateway instances.

[41] As an alternative embodiment, the same ID may be mapped to the same address sub-range served by a same gateway instance. In that case, the traffic from multiple VMs of the same/single user may be encapsulated using the address sub-range.

[42] As another alternative embodiment, the same ID may be mapped to the same gateway address or the same gateway instance ID. In that case, the traffic from multiple VMs of the same/single user may be encapsulated using the same gateway instance ID.

[43] At block 308, the traffic from multiple VMs of the same/single user may be forwarded to a first gateway instance GW1 announcing the destination IP address or an IP address sub-range including the destination IP address in the gateway cluster 108. Also referring to FIG. 1, the server 102 may forward the traffic from the VM 110 to the first gateway instance GW1, the server 104 may forward the traffic from the VM 112 to the first gateway instance GW1, and the server 106 may forward the traffic from the VM 114 to the first gateway instance GW1. Therefore, the traffic from multiple VMs of the same/single user may reach the same gateway instance in the gateway cluster 108.

[44] As an alternative embodiment, the same ID may be mapped to the same address sub-range served by the same gateway instance. In that case, the ID may be mapped into one of a group of address sub-ranges, where each address sub-range correspond to a gateway

instance. The traffic from multiple VMs of the same/single user may be forwarded to the first gateway instance GW 1 announcing the address sub-range. Also referring to FIG. 1, the server 102 may forward the traffic from the VM 110 to the first gateway instance GW 1announcing the address sub-range, the server 104 may forward the traffic from the VM 112 to the first gateway instance GW 1announcing the same address sub-range, and the server 106 may forward the traffic from the VM 114 to the first gateway instance GW 1announcing the same address sub-range. Therefore, the traffic from multiple VMs of the same/single user may reach the same gateway instance in the gateway cluster 108.

[45] As another alternative embodiment, the same ID may be mapped to the same gateway addressor the same gateway instance ID. In that case, the ID may be mapped into one of a group of gateway instance ID, where each gateway instance ID correspond to a gateway instance. The traffic from multiple VMs of the same/single user may be forwarded to the first gateway instance GW 1 associated with the gateway instance ID. Also referring to FIG. 1, the server 102 may forward the traffic from the VM 110 to the first gateway instance GW 1associated with the gateway instance ID, the server 104 may forward the traffic from the VM 112 to the first gateway instance GW 1associated with the same gateway instance ID, and the server 106 may forward the traffic from the VM 114 to the first gateway instance GW 1associated with the same gateway instance ID. Therefore, the traffic from multiple VMs of the same/single user may reach the same gateway instance in the gateway cluster 108.

[46] At block 310,the first gateway instance GW 1 may monitor the traffic from multiple VMs of the same/single user by determining whether a traffic indicator exceeds a threshold. The gateway instances may be programmed to monitor the traffic indicator, which may indicate the bandwidth or the dataflow consumed by the traffic of the user. The traffic indicator may also be other suitable parameters representing resources consumed by the user. The threshold may be set based on how much resource the user has purchased.

[47] At block 312,the first gateway instance GW 1 may throttle the traffic of multiple VMs of the same/single user upon determining that the traffic indicator exceeds the threshold. For example, a user may have purchased 4GHzbandwidth. When the bandwidth occupied by the traffic of multiple VMs of the user exceeds 4GHz, the first gateway instance GW1 may throttle the traffic of multiple VMs of the same/single user such that the traffic of multiple

VMs of the same/single user occupies no more than 4GHz bandwidth. Also, the first gateway instance GW1 may send a warning or a message to the user indicating that the traffic of the user is approaching or has reached the purchased bandwidth before clipping or throttling the traffic of the user. In response to the warning or the message, the user may choose to purchase more resources or accept clipping or throttling of the traffic.

[48] In the above process 300, all traffic associated with the same ID may be routed to a single gateway instance to facilitate the measurement of the amount of traffic consumed by the same user. Also, traffic of the same user may be routed to one or more (or any one of these one or more) gateway instances that are selected in a determinative manner, rather than randomly, and the amount of the traffic of the same user may be measured/throttled by these one or more gateway instances. By way of example and not limitation, the IDs may be mapped to a group of one or more gateway instances which can serve and handle the traffic of the same user (i.e., traffic having the same user ID) using one of the mapping methods (for example, hashing, etc.) described in the foregoing description, and one of the one or more gateway instances may be selected to serve and handle the traffic of the same user in a rotating manner.

[49] Alternatively, after the traffic of the same ID is mapped into a determinative group of one or more gateway instances, one of the one or more gateway instances within this particular group may be randomly selected to serve and handle the traffic of the same user.

[50] Using a determinative group of one or more gateway instances to serve and handle all the traffic having a same ID (e.g., a same user ID) helps alleviating the excessive workload that a single gateway instance may handle for the traffic coming from a same user, and improves the speed of transmitting and handling the traffic of the same user through the gateway, for example. This is especially true when the user is a large entity such as an enterprise, an institution, etc., in which the amount of traffic coming from the user is normally large at certain period of time.

[51] In implementations, whether traffic of a same user is mapped into a determinative group of gateway instances or a single gateway instance may depend on the amount of bandwidth that this user purchases, or the amount of traffic that the user is sending and/or receiving, etc. For example, a group of gateway instances may be used and mapped

into if the amount of bandwidth that this user purchases or the amount of traffic that the user is sending and/or receiving is larger than a certain threshold (such as 100MB, 500GB, 1GB, etc.)

[52] FIG. 3B illustrates an example process for rerouting traffic when one or more gateway instances are down. Rerouting of the traffic illustrated in FIG. 3B may follow block 312 of FIG. 3A.

[53] At block 314,if the first gateway instance GW1 is down, the traffic from multiple VMs of the same/single user may be rerouted to a second gateway instance GW2. For example, if the attempt of routing the traffic to the first gateway instance GW 1 is unsuccessful, indicating that the first gateway instance GW 1 may be down, the traffic may be rerouted to the second gateway instance GW 2. If the first gateway instance GW1 is a Level-m node, the second gateway instance GW2 may share the same Level-(m-1) node with the first gateway instance GW1.

[54] At block 316,if the second gateway instance GW 2 is also down, the traffic from multiple VMs of the same/single user may be rerouted to a third gateway instance GW3. For example, if the attempt of rerouting the traffic to the second gateway instance GW 2 is unsuccessful, indicating that the second gateway instance GW2 is also down, the traffic may be rerouted to the third gateway instance GW3. The third gateway instance GW3 may share the same Level-(m-2) node with the first gateway instance GW1.

[55] At block 318,if all gateway instances under a Level-n node are down, the traffic from multiple VMs of the same/single user may be rerouted to the closest gateway instance sharing the same Level-(n-1) node with the first gateway instance GW 1.For example, if all previous attempts of rerouting the traffic to gateway instances under a Level-n node are unsuccessful, indicating that all gateway instances under the Level-n node are down, the traffic may be rerouted to the gateway instance sharing the same Level-(n-1) node with the first gateway instance GW1, where $1 \leq n \leq m-1$.

[56] As shown above, the availability and the fault-tolerant ability of the gateway cluster 108 may be improved while achieving the traffic throttling for the network traffic from multiple VMs of the same/single user.

[57] FIG. 4illustrates an example process 400 for establishing the gateway cluster

108.

[58] At block 402, a Level-0 root node, which may correspond to the whole VIP address range of the gateway cluster 108, may be provided. For example, the gateway cluster 108 may be constructed based on a (m+1) level full binary tree, and each gateway instance may be a leaf node. Also, the gateway cluster 108 may be constructed based on a binary tree, a ternary tree, a quaternary tree, other tree structures, or any combination thereof.

[59] At block 404, the Level-0 root node may be split into a first predetermined number ofLevel-1 nodes. For example, if the gateway cluster 108 is constructed based on a binary tree, the first predetermined number may be two. In that case, the Level-0 root node may be split into two Level-1 nodes. If the gateway cluster 108 is constructed based on a ternary tree or a quaternary tree, the first predetermined number may be three or four. In such cases, the Level-0 root node may be split into three or four Level-1 nodes. Also, other tree structures or any combination of different tree structures may be used.

[60] At block 406, the whole VIP address range may be divided into the first predetermined number ofLevel-1 sub-ranges. For example, if the gateway cluster 108 is constructed based on a binary tree, the first predetermined number may be two. In that case, the whole VIP address range may be divided into two Level-1 sub-ranges. If the gateway cluster 108 is constructed based on a ternary tree or a quaternary tree, the first predetermined number may be three or four. In such cases, the whole VIP address range may be divided into three or four Level-1 sub-ranges. Also, other tree structures or any combination of different tree structures may be used.

[61] At block 408, each Level-n node may be split into a second predetermined number of Level-(n+1) nodes. For example, if the gateway cluster 108 is constructed based on a binary tree, the second predetermined number may be two. In that case, the Level-n node may be split into two Level-(n+1) nodes. If the gateway cluster 108 is constructed based on a ternary tree or a quaternary tree, the second predetermined number may be three or four. In such cases, each Level-n node may be split into three or four Level-(n+1) nodes. Also, other tree structures or any combination of different tree structures may be used.

[62] At block 410, and each Level-n sub-range may be divided into the second predetermined number of Level-(n+1) sub-ranges. For example, if the gateway cluster 108 is

constructed based on a binary tree, the first predetermined number may be two. In that case, each Level-n sub-range may be divided into two Level-(n+1) sub-ranges. If the gateway cluster 108 is constructed based on a ternary tree or a quaternary tree, the first predetermined number may be three or four. In such cases, each Level-n sub-range may be divided into three or four Level-(n+1) sub-ranges. Also, other tree structures or any combination of different tree structures may be used.

[63] At block 412,splitting each Level-n node and dividing each Level-n sub-range may be repeated in an iterative manner, with n = 1 as an initial value, and n+1 acting as a value for a next iteration, till n = m-1.Nand m may be integers, where $1 \leqslant n \leqslant m-1$.For example, if the gateway cluster 108 is constructed based on a binary tree, 2^m may be a total number of gateway instances. If the gateway cluster 108 is constructed based on a ternary tree or a quaternary tree,3^mor 4^m may be a total number of gateway instances. Also, other tree structures or any combination of different tree structures may be used.

[64] FIG. 5 illustrates an example block diagrams of an apparatus for throttling traffic from multiple VMs of the same/single user.

[65] FIG. 5 is only one example of an apparatus 500 and is not intended to suggest any limitation as to the scope of use or functionality of any computing device utilized to perform the processes and/or procedures described above. Other well-known computing devices, apparatuses, environments and/or configurations that may be suitable for use with the embodiments include, but are not limited to, driver/passenger computers, server computers, hand-held or laptop devices, multiprocessor apparatuses, microprocessor-based apparatuses, set-top boxes, game consoles, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above apparatuses or devices, implementations using field programmable gate arrays ("FPGAs") and application specific integrated circuits ("ASICs"), and/or the like.

[66] The apparatus 500 may include one or more processors 502 and memory 504 communicatively coupled to the processor(s) 502. The processor(s) 502 may execute one or more modules and/or processes to cause the processor(s) 502 to perform a variety of functions. In some embodiments, the processor(s) 502 may include a central processing unit (CPU), a graphics processing unit (GPU), both CPU and GPU, or other processing units or components

known in the art. Additionally, each of the processor(s) 502 may possess its own local memory, which also may store program modules, program data, and/or one or more operating apparatuses.

[67] Depending on the exact configuration and type of the apparatus 500, the memory 504 may be volatile, such as RAM, non-volatile, such as ROM, flash memory, miniature hard drive, memory card, and the like, or some combination thereof. The memory 504 may include computer-executable instructions that are executable by the processor(s) 502, when executed by the processor(s) 502, cause the processor(s) 502 to implement systems and processes described with reference to FIGS. 1 – 4.

[68] The apparatus 500 may additionally include an input/output (I/O) interface 506 for receiving and outputting data. The apparatus 500 may also include a communication module 508 allowing the apparatus 500 to communicate with other devices (not shown) over a network (not shown). The network may include the Internet, wired media such as a wired network or direct-wired connections, and wireless media such as acoustic, radio frequency (RF), infrared, and other wireless media.

[69] With systems and processes discussed herein, a high available and fault-tolerant solution may be provided where traffic throttling for the network traffic from the multiple VMs of a same/single user may be achieved.

[70] Though systems and processes are discussed herein with regard to VMs and VPC user/tenant, processes and systems discussed herein may be used in any suitable scenarios where it is necessary to monitor and control traffic from different devices of the same/single user.

[71] Some or all operations of the methods described above can be performed by execution of computer-readable instructions stored on a computer-readable storage medium, as defined below. The term “computer-readable instructions” as used in the description and claims, include routines, applications, application modules, program modules, programs, components, data structures, algorithms, and the like. Computer-readable instructions can be implemented on various system configurations, including single-processor or multiprocessor systems, minicomputers, mainframe computers, personal computers, hand-held computing devices, microprocessor-based, programmable consumer electronics, combinations thereof,

and the like.

[72] The computer-readable storage media may include volatile memory (such as random access memory (RAM)) and/or non-volatile memory (such as read-only memory (ROM), flash memory, etc.). The computer-readable storage media may also include additional removable storage and/or non-removable storage including, but not limited to, flash memory, magnetic storage, optical storage, and/or tape storage that may provide non-volatile storage of computer-readable instructions, data structures, program modules, and the like.

[73] A non-transient computer-readable storage medium is an example of computer-readable media. Computer-readable media includes at least two types of computer-readable media, namely computer-readable storage media and communications media. Computer-readable storage media includes volatile and non-volatile, removable and non-removable media implemented in any process or technology for storage of information such as computer-readable instructions, data structures, program modules, or other data. Computer-readable storage media includes, but is not limited to, phase change memory (PRAM), static random-access memory (SRAM), dynamic random-access memory (DRAM), other types of random-access memory (RAM), read-only memory (ROM), electrically erasable programmable read-only memory (EEPROM), flash memory or other memory technology, compact disk read-only memory (CD-ROM), digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other non-transmission medium that can be used to store information for access by a computing device. In contrast, communication media may embody computer-readable instructions, data structures, program modules, or other data in a modulated data signal, such as a carrier wave, or other transmission mechanism. As defined herein, computer-readable storage media do not include communication media.

[74] The computer-readable instructions stored on one or more non-transitory computer-readable storage media that, when executed by one or more processors, may perform operations described above with reference to FIGS. 1-5. Generally, computer-readable instructions include routines, programs, objects, components, data structures, and the like that perform particular functions or implement particular abstract data types. The order in which the operations are described is not intended to be construed as a

limitation, and any number of the described operations can be combined in any order and/or in parallel to implement the processes.

Example Clauses

[75] Clause 1. A method, comprising: receiving traffic from multiple virtual machines (VMs) associated with a same identification (ID); mapping the ID into a destination Internet Protocol (IP) address; and forwarding the traffic from the multiple VMs to a first gateway instance announcing the destination IP address.

[76] Clause 2. The method of clause 1, wherein before forwarding the traffic from the multiple VMs to a first gateway instance, the method further comprises encapsulating the traffic from the multiple VMs using the destination IP address.

[77] Clause 3. The method of clause 1, further comprising: monitoring the traffic from the multiple VMs by determining whether a traffic indicator exceeds a threshold; and upon determining that the traffic indicator exceeds the threshold, throttling the traffic from the multiple VMs.

[78] Clause 4. The method of clause 3, wherein the traffic indicator indicates a bandwidth occupied by the traffic of the multiple VMs.

[79] Clause 5. The method of clause 3, wherein the traffic indicator indicates a dataflow of the multiple VMs.

[80] Clause 6. The method of clause 1, wherein mapping the ID into the destination IP address includes hashing the ID into a whole IP address range of a gateway cluster, the gateway cluster including a the first gateway instance.

[81] Clause 7. The method of clause 6, wherein the whole IP address range includes a plurality of IP address sub-ranges, the gateway cluster including a plurality of gateway instances, each gateway instance of the plurality of gateway instances announcing an IP sub-range of the plurality of IP address sub-ranges.

[82] Clause 8. The method of clause 7, further comprising: determining whether the first gateway instance is down; and upon determining that the first gateway instance is down, rerouting the traffic from the multiple VMs to a second gateway instance in the gateway cluster.

[83] Clause 9. An apparatus, comprising: one or more processors; a memory coupled to the one or more processors, the memory storing computer-readable instructions executable by the one or more processors, that when executed by the one or more processors, cause the one or more processors to perform operations including: receiving traffic from multiple virtual machines (VMs) associated with a same identification (ID); mapping the ID into a destination Internet Protocol (IP) address; and forwarding the traffic from the multiple VMs to a first gateway instance announcing the destination IP address.

[84] Clause 10. The apparatus of clause 9, wherein before forwarding the traffic from the multiple VMs to a first gateway instance, the operations further comprise encapsulating the traffic from the multiple VMs using the destination IP address.

[85] Clause 11. The apparatus of clause 9, wherein the operations further comprise: monitoring the traffic from the multiple VMs by determining whether a traffic indicator exceeds a threshold; and upon determining that the traffic indicator exceeds the threshold, throttling the traffic from the multiple VMs.

[86] Clause 12. The apparatus of clause 11, wherein the traffic indicator indicates a bandwidth occupied by the traffic of the multiple VMs.

[87] Clause 13. The apparatus of clause 11, wherein the traffic indicator indicates a dataflow of the multiple VMs.

[88] Clause 14. The apparatus of clause 9, wherein mapping the ID into the destination IP address includes hashing the ID into a whole IP address range of a gateway cluster, the gateway cluster including the first gateway instance.

[89] Clause 15. The apparatus of clause 14, wherein the whole IP address range includes a plurality of IP address sub-ranges, the gateway cluster including a plurality of gateway instances, each gateway instance of the plurality of gateway instances announcing an IP sub-range of the plurality of IP address sub-ranges.

[90] Clause 16. The apparatus of clause 15, wherein the operations further comprise: determining whether the first gateway instance is down; and upon determining that the first gateway instance is down, rerouting the traffic from the multiple VMs to a second gateway instance in the gateway cluster.

[91] Clause 17. A computer-readable storage medium storing computer-readable

instructions executable by one or more processors, that when executed by the one or more processors, cause the one or more processors to perform operations comprising: receiving traffic from multiple virtual machines (VMs) associated with a same identification (ID); mapping the ID into a destination Internet Protocol (IP) address; and forwarding the traffic from the multiple VMs to a first gateway instance announcing the destination IP address.

[92] Clause 18. The computer-readable storage medium of clause 17, wherein before forwarding the traffic from the multiple VMs to a first gateway instance, the operations further comprise encapsulating the traffic from the multiple VMs using the destination IP address.

[93] Clause 19. The computer-readable storage medium of clause 17, the operations further comprise: monitoring the traffic from the multiple VMs by determining whether a traffic indicator exceeds a threshold; and upon determining that the traffic indicator exceeds the threshold, throttling the traffic from the multiple VMs.

[94] Clause 20. The computer-readable storage medium of clause 17, wherein the traffic indicator indicates a bandwidth occupied by the traffic of the multiple VMs.

Conclusion

[95] Although the subject matter has been described in language specific to structural features and/or methodological acts, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features or acts described. Rather, the specific features and acts are disclosed as exemplary forms of implementing the claims.

CLAIMS

1. A method, comprising:

receiving traffic from multiple virtual machines (VMs) associated with a same identification (ID);

mapping the ID into a destination Internet Protocol (IP) address; and

forwarding the traffic from the multiple VMs to a first gateway instance announcing the destination IP address.

2. The method of claim 1, wherein before forwarding the traffic from the multiple VMs to a first gateway instance, the method further comprises encapsulating the traffic from the multiple VMs using the destination IP address.

3. The method of claim 1, further comprising:

monitoring the traffic from the multiple VMs by determining whether a traffic indicator exceeds a threshold; and

upon determining that the traffic indicator exceeds the threshold, throttling the traffic from the multiple VMs.

4. The method of claim 3, wherein the traffic indicator indicates a bandwidth occupied by the traffic of the multiple VMs.

5. The method of claim 3, wherein the traffic indicator indicates a dataflow of the multiple VMs.

6. The method of claim 1, wherein mapping the ID into the destination IP address includes hashing the ID into a whole IP address range of a gateway cluster, the gateway cluster including a the first gateway instance.

7. The method of claim 6, wherein the whole IP address range includes a plurality of IP address sub-ranges, the gateway cluster including a plurality of gateway instances, each gateway instance of the plurality of gateway instances announcing an IP sub-range of the

plurality of IP address sub-ranges.

8. The method of claim 7, further comprising:

determining whether the first gateway instance is down; and

upon determining that the first gateway instance is down, rerouting the traffic from the multiple VMs to a second gateway instance in the gateway cluster.

9. An apparatus, comprising:

one or more processors;

a memory coupled to the one or more processors, the memory storing computer-readable instructions executable by the one or more processors, that when executed by the one or more processors, cause the one or more processors to perform operations including:

receiving traffic from multiple virtual machines (VMs) associated with a same identification (ID);

mapping the ID into a destination Internet Protocol (IP) address; and

forwarding the traffic from the multiple VMs to a first gateway instance announcing the destination IP address.

10. The apparatus of claim 9, wherein before forwarding the traffic from the multiple VMs to a first gateway instance, the operations further comprise encapsulating the traffic from the multiple VMs using the destination IP address.

11. The apparatus of claim 9, wherein the operations further comprise:

monitoring the traffic from the multiple VMs by determining whether a traffic indicator exceeds a threshold; and

upon determining that the traffic indicator exceeds the threshold, throttling the traffic from the multiple VMs.

12. The apparatus of claim 11, wherein the traffic indicator indicates a bandwidth

occupied by the traffic of the multiple VMs.

13. The apparatus of claim 11,wherein the traffic indicator indicates a dataflow of the multiple VMs.

14. The apparatus of claim 9,wherein mapping the ID into the destination IP address includes hashing the ID into a whole IP address range of a gateway cluster, the gateway cluster including the first gateway instance.

15. The apparatus of claim 14, wherein the whole IP address range includes a plurality of IP address sub-ranges, the gateway cluster including a plurality of gateway instances, each gateway instance of the plurality of gateway instances announcing an IP sub-range of the plurality of IP address sub-ranges.

16. The apparatus of claim 15, wherein the operations further comprise:

determining whether the first gateway instance is down; and

upon determining that the first gateway instance is down, rerouting the traffic from the multiple VMs to a second gateway instance in the gateway cluster.

17. A computer-readable storage medium storing computer-readable instructions executable by one or more processors, that when executed by the one or more processors, cause the one or more processors to perform operations comprising:

receiving traffic from multiple virtual machines (VMs) associated with a same identification (ID);

mapping the ID into a destination Internet Protocol (IP) address; and

forwarding the traffic from the multiple VMs to a first gateway instance announcing the destination IP address.

18. The computer-readable storage medium of claim 17, wherein before forwarding the traffic from the multiple VMs to a first gateway instance, the operations further comprise encapsulating the traffic from the multiple VMs using the destination IP address.

19. The computer-readable storage medium of claim 17, the operations further comprise:

monitoring the traffic from the multiple VMs by determining whether a traffic indicator exceeds a threshold; and

upon determining that the traffic indicator exceeds the threshold, throttling the traffic from the multiple VMs.

20. The computer-readable storage medium of claim 17, wherein the traffic indicator indicates a bandwidth occupied by the traffic of the multiple VMs.

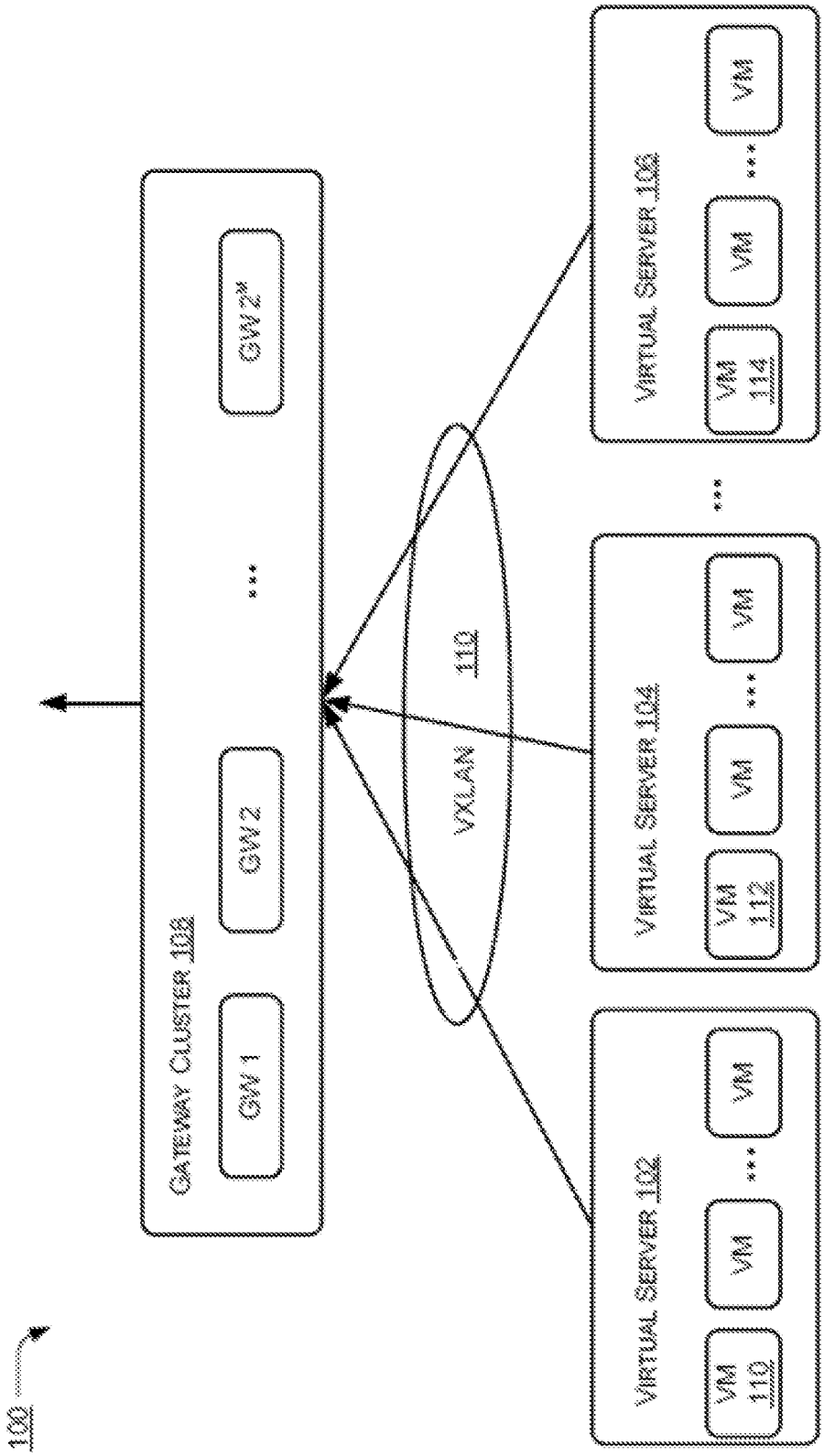
100
Gateway Cluster 106
GW 1
GW 2
GW 2"
VXLAN 110
Virtual Server 102
VM 110
VM
VM
Virtual Server 104
VM 112
VM
VM
Virtual Server 108
VM 114
VM
VM

FIG. 1

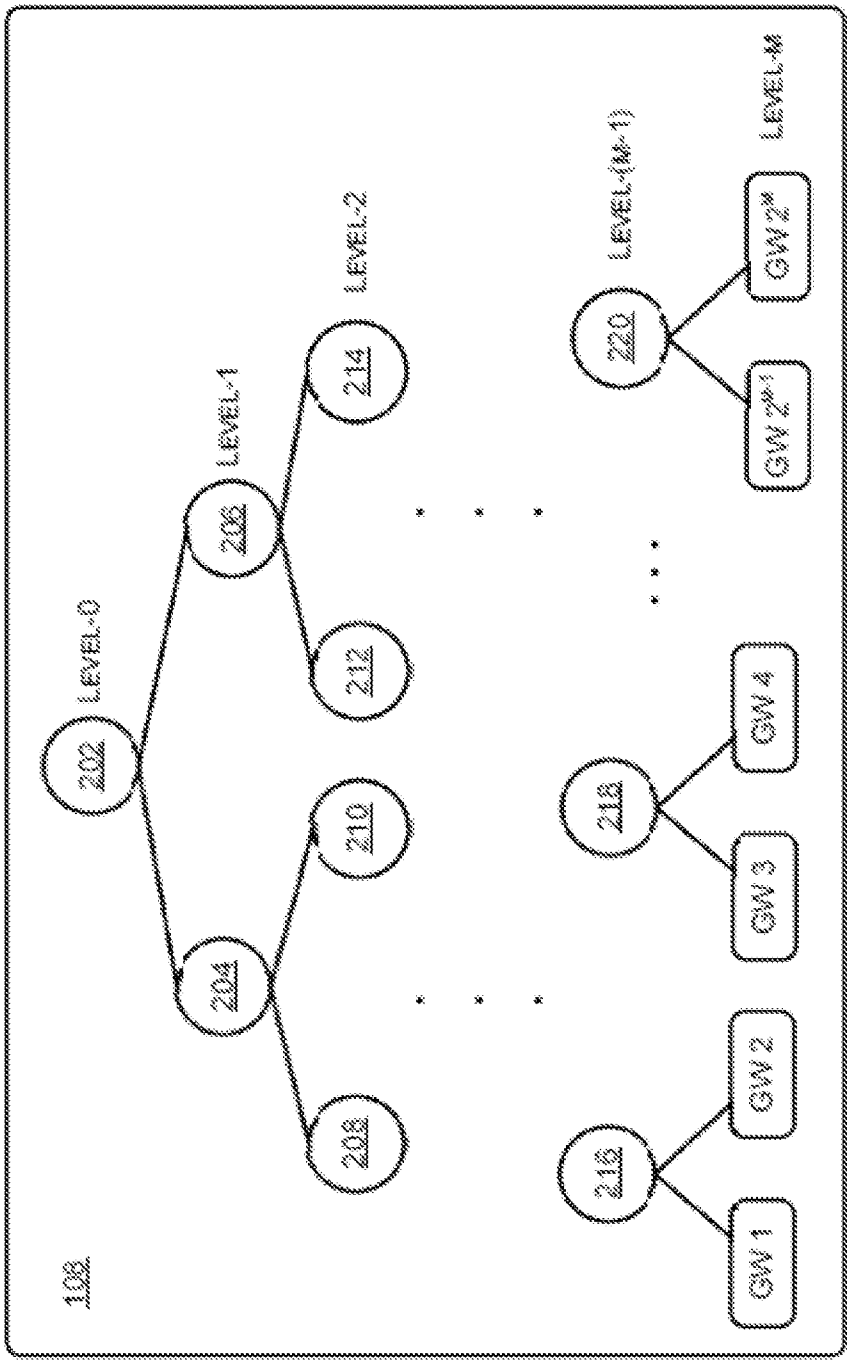
106
LEVEL-0
LEVEL-1
LEVEL-2
LEVEL-(M-1)
LEVEL-M
202
204
206
208
210
212
214
216
216
218
220
GW 1
GW 2
GW 3
GW 4
GW 2^M-1
GW 2^M

FIG. 2

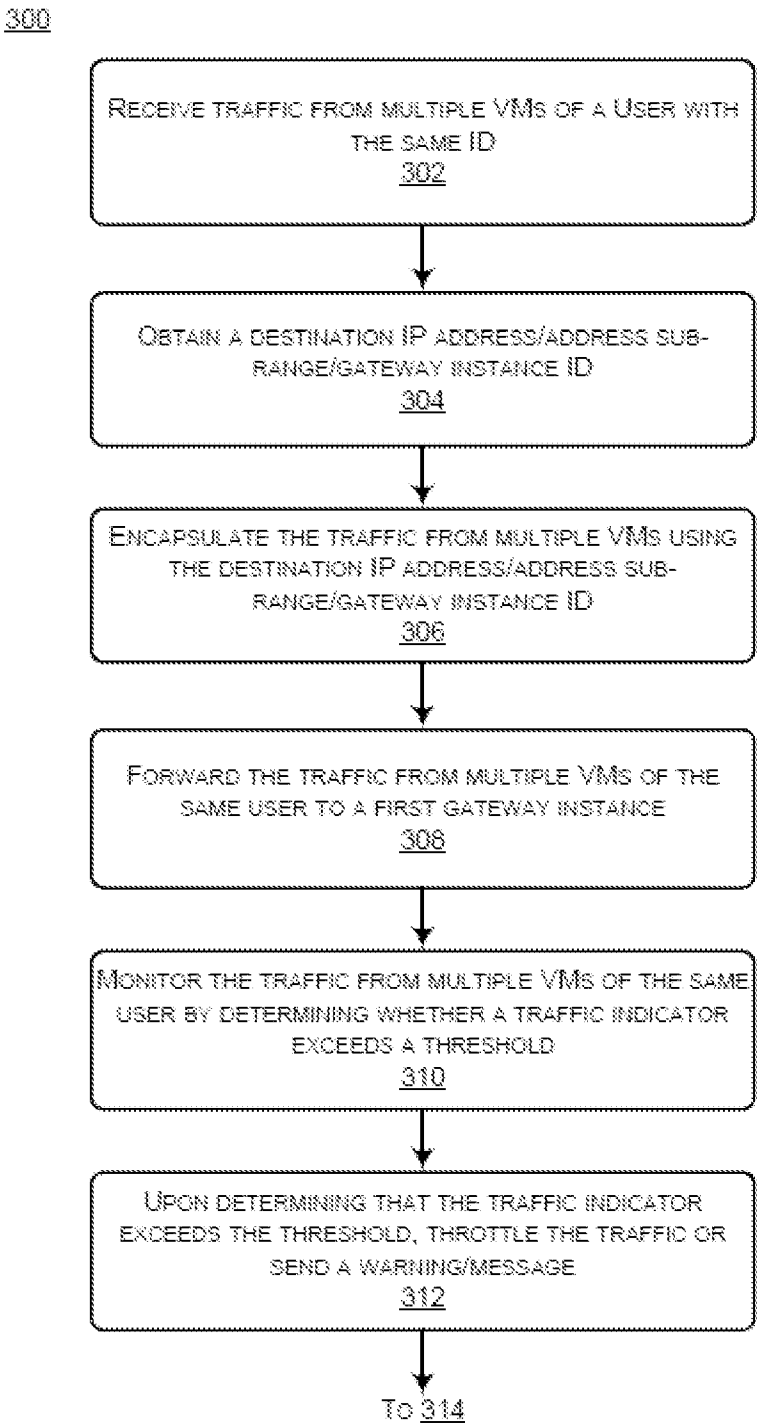
300
Receive traffic from multiple VMs of a User with the same ID
302
Obtain a destination IP address/address sub-range/gateway instance ID
304
Encapsulate the traffic from multiple VMs using the destination IP address/address sub-range/gateway instance ID
306
Forward the traffic from multiple VMs of the same user to a first gateway instance
308
Monitor the traffic from multiple VMs of the same user by determining whether a traffic indicator exceeds a threshold
310
Upon determining that the traffic indicator exceeds the threshold, throttle the traffic or send a warning/message
312
To 314

FIG. 3A

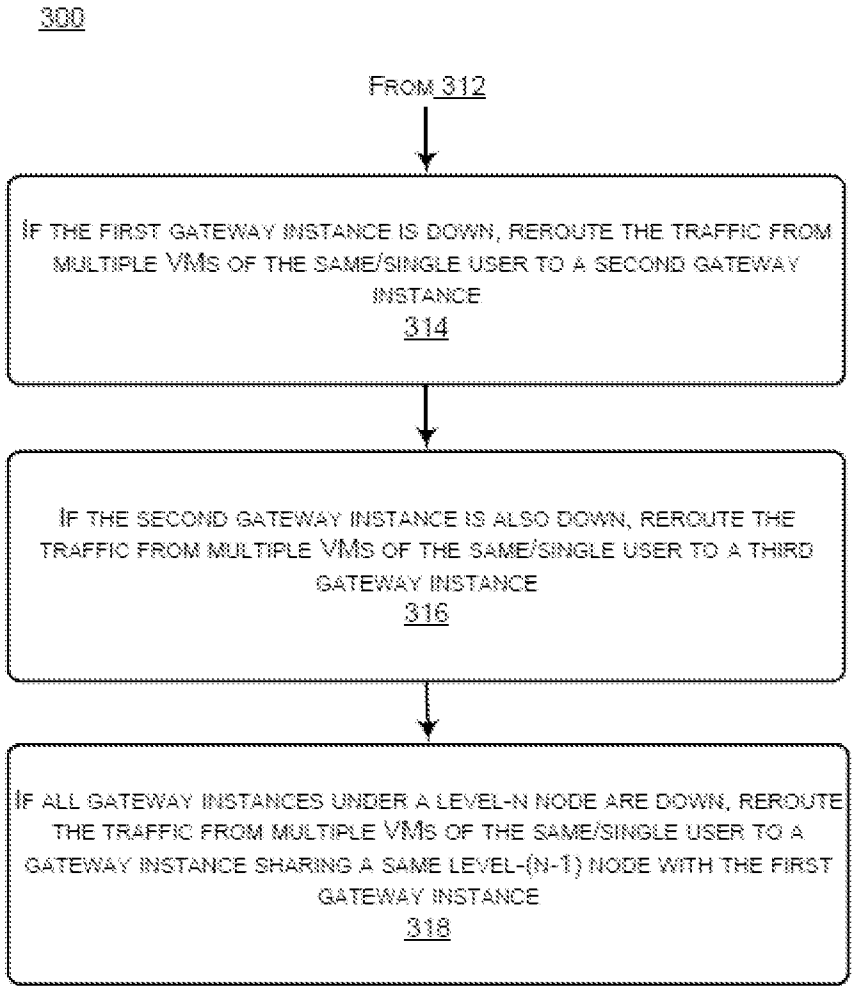
300
FROM 312
IF THE FIRST GATEWAY INSTANCE IS DOWN, REROUTE THE TRAFFIC FROM MULTIPLE VMS OF THE SAME/SINGLE USER TO A SECOND GATEWAY INSTANCE
314
IF THE SECOND GATEWAY INSTANCE IS ALSO DOWN, REROUTE THE TRAFFIC FROM MULTIPLE VMS OF THE SAME/SINGLE USER TO A THIRD GATEWAY INSTANCE
316
IF ALL GATEWAY INSTANCES UNDER A LEVEL-N NODE ARE DOWN, REROUTE THE TRAFFIC FROM MULTIPLE VMS OF THE SAME/SINGLE USER TO A GATEWAY INSTANCE SHARING A SAME LEVEL-(N-1) NODE WITH THE FIRST GATEWAY INSTANCE
318

FIG. 3B

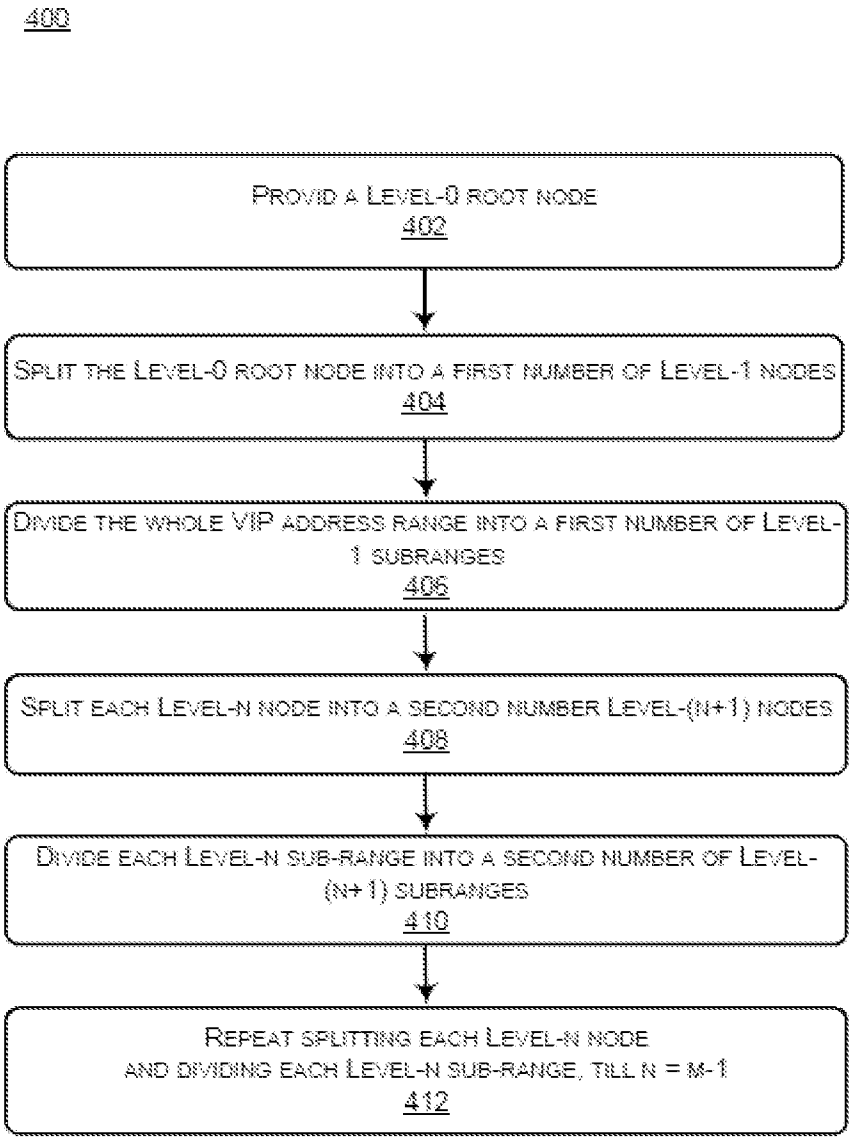
400
PROVID A LEVEL-0 ROOT NODE
402
SPLIT THE LEVEL-0 ROOT NODE INTO A FIRST NUMBER OF LEVEL-1 NODES
404
DIVIDE THE WHOLE VIP ADDRESS RANGE INTO A FIRST NUMBER OF LEVEL-1 SUBRANGES
406
SPLIT EACH LEVEL-N NODE INTO A SECOND NUMBER LEVEL-(N+1) NODES
408
DIVIDE EACH LEVEL-N SUB-RANGE INTO A SECOND NUMBER OF LEVEL-(N+1) SUBRANGES
410
REPEAT SPLITTING EACH LEVEL-N NODE AND DIVIDING EACH LEVEL-N SUB-RANGE, TILL N = M-1
412

FIG. 4

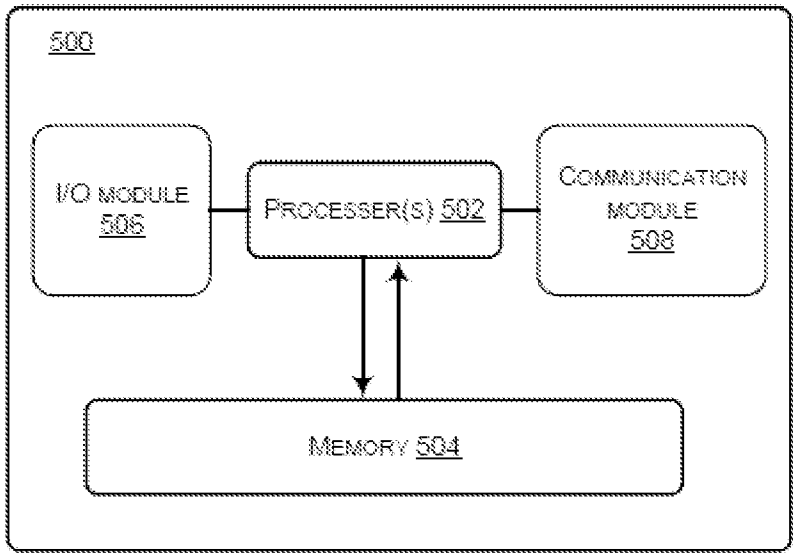
500
I/O MODULE 506
PROCESSER(S) 502
COMMUNICATION MODULE 508
MEMORY 504

FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2019/126620

A. CLASSIFICATION OF SUBJECT MATTER

H04L 29/06(2006.01)i; H04L 12/741(2013.01)i; G06F 9/455(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)

WPI, EPODOC, CNKI, CNPAT: VM, VMs, virtual machine, ID, identifier, map, IP, gateway, switch, instance, table, corresponding, indicator

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2016285826 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 29 September 2016 (2016-09-29) description, paragraphs [0014]-[0033], claim 1	1-20
A	CN 103701928 A (SHANDONG UNIVERSITY) 02 April 2014 (2014-04-02) the whole document	1-20
A	CN 104468775 A (G-CLOUD TECHNOLOGY CO., LTD.) 25 March 2015 (2015-03-25) the whole document	1-20
A	CN 106953943 A (CHINA UNITED NETWORK COMMUNICATIONS CORPORATION) 14 July 2017 (2017-07-14) the whole document	1-20

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"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)
"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
"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
"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
"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
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
02 March 2020	**27 March 2020**

Name and mailing address of the ISA/CN	Authorized officer
National Intellectual Property Administration, PRC **6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088** **China**	**LI,Jia**
Facsimile No. **(86-10)62019451**	Telephone No. **86-(10)-53961439**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2019/126620

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US 2016285826 A1	29 September 2016	None	
CN 103701928 A	02 April 2014	None	
CN 104468775 A	25 March 2015	None	
CN 106953943 A	14 July 2017	None	