



(51) International Patent Classification:
H04L 12/28 (2006.01)

(21) International Application Number:
PCT/CN2019/126621

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
16/236,242 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 Depart-
ment 5/F, Building 3, No. 969 West Wen Yi Road, Yu Hang
District, Hangzhou, Zhejiang 311121 (CN).

(74) Agent: **BEIJING SANYOU INTELLECTUAL PROP-
ERTY 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,
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))

(54) Title: METHOD, APPARATUS, AND COMPUTER-READABLE STORAGE MEDIUM FOR NETWORK OPTIMIZATION
OF CLOUD STORAGE SERVICE

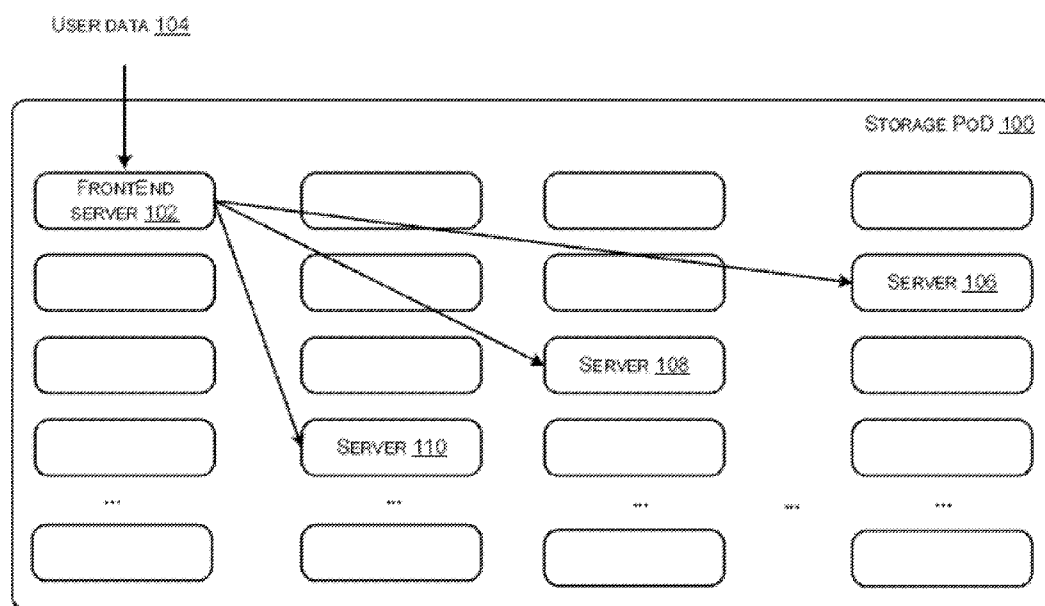


FIG. 1

(57) **Abstract:** Apparatuses and methods include receiving, by a switch, a packet; creating, by the switch, one or more copies of the packet; selecting, by the switch, a plurality of servers; and separately forwarding, by the switch, the packet and the one or more copies to the plurality of servers.

METHOD, APPARATUS, AND COMPUTER-READABLE STORAGE MEDIUM FOR NETWORK OPTIMIZATION OF CLOUD STORAGE SERVICE

5

CROSS REFERENCE TO RELATED APPLICATION

[1] This disclosure claims the benefits of priority to United States application number 16/236,242, filed December 28, 2018, which is incorporated herein by reference in its entirety.

10

BACKGROUND

[2] For most cloud service providers, the data center network is a Clos network, which is a kind of multistage circuit switching network. Generally, when receiving a copy of data from a user, the cloud service server creates multiple copies of the received data and distributes the multiple copies to different servers. However, because multiple copies of data need to be distributed through the network, a heavy pressure is posed on the network connecting the storage servers. In addition, creating and sending multiple copies to the network will consume a significant amount of CPU power of the Front End server.

15
20

BRIEF DESCRIPTION OF THE DRAWINGS

[3] 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.

25

[4] FIG. 1 illustrates an example block diagram of a storage Point of Delivery (PoD).

[5] FIG. 2 illustrates an example block diagram of a Clos data center network.

[6] FIG.3A illustrate an example process for network optimization of cloud storage service.

[7] FIG.3B illustrate an example process for network optimization of cloud storage service.

5 [8] FIG. 4 illustrates an example block diagram of an apparatus for optimizing cloud storage service.

DETAILED DESCRIPTION

[9] Apparatuses and methods discussed herein are directed to cloud storage service, and
10 more specifically to networking optimization of cloud storage service.

[10] A switch may receive a data packet from another party, for example, a sending server, a user device, a computing device, a process, an application, etc. The switch may determine whether a destination Internet Protocol (IP) address of the data packet is a reserved or designated IP address. Upon determining that the destination IP address of the data packet
15 is the reserved or designated IP address, the switch may create one or more copies of the data packet. The switch may select a plurality of servers and forward the data packet and the one or more copies of the data packet to the plurality servers separately.

[11] The one or more copies of the data packet may include a predetermined number (e.g., 1, 2, 3, 4, ... , etc.) of copies of the data packet that is set in advance according to
20 requirements of the cloud storage service. Example requirements include a data storage requirement, a data reliability requirement, a data availability requirement, etc.

[12] The predetermined number of copies may include a first copy having a first destination IP address, a second copy having a second destination IP address, a third copy having a third destination IP address, ..., and an n^{th} copy having an n^{th} destination IP address,
25 where n is the predetermined number.

[13] The plurality of servers may include the predetermined number of servers, for example, a first server having a first server IP address, a second server having a second server IP address, a third server having a third server IP address, ..., and an n^{th} server having an n^{th}

server IP address where n is the predetermined number, which may be any positive integer such as 1, 2, 3, etc.

[14] The switch may replace the first destination IP address of the first copy using the first server IP address, the second destination IP address of the second copy using the second server IP address of the second server, the third destination IP address of the third copy using the third server IP address, ..., and the n^{th} destination IP address of the n^{th} copy using the n^{th} server IP address.

[15] The selection of the plurality of servers may be performed randomly, in a rotating manner, using a hash function, or based on any other suitable criteria. For example, idle servers, servers with high capacities, servers that are not busy, etc. may be selected.

[16] The first server may receive the first copy from the switch and send an acknowledgement message (ACK) to the switch. The second server may receive the second copy from the switch and send an ACK to the switch. The n^{th} server may receive the n^{th} copy from the switch and send an ACK to the switch.

[17] The switch may create an $(n+1)^{\text{th}}$ copy of the packet if the switch does not receive the ACK from the any of the plurality servers within a predetermined period. The additional copy may include a $(n+1)^{\text{th}}$ destination IP address.

[18] The switch may select an $(n+1)^{\text{th}}$ server having a $(n+1)^{\text{th}}$ server IP address and replace the $(n+1)^{\text{th}}$ destination IP address of the $(n+1)^{\text{th}}$ copy using the $(n+1)^{\text{th}}$ server IP address and forward the $(n+1)^{\text{th}}$ copy to the $(n+1)^{\text{th}}$ server.

[19] FIG. 1 illustrates an example block diagram of a storage PoD100, which may be a module of network, compute, storage, and application components that work together to deliver networking services. Referring to FIG. 1, when a Front End server 102 for cloud storage service receives data104 from a user (not shown), the Front End server 102 may create a predetermined number of copies of the data 104 and pick a predetermined number of servers in the same storage PoD 100, for example, servers 106, 108, and 110. Though three servers are picked in this example, other number of servers may be picked. The Front End server 102 may save one copy on each of the selected servers 106, 108, and 110. If one server is down, the user data may be found on other servers. For example, the server 106 may fail, crash, be too busy to respond, etc., the user data may be available on servers 108 and 110. In

case of a single server or a single rack failure, the user data still can be recovered, thereby improving the availability of data.

[20] FIG. 2 illustrates an example block diagram of a data center network 200, where multiple layers of switches are used to connect all servers. Though FIG. 2 shows two layers of switches, i.e., a Tier-1 layer and a top of rack (TOR) layer, there may be other layers. The tier-1 layer may also be referred as the top layer. The data center network 200 may be a Clos network or any suitable types of network.

[21] The IP address range of storage servers in a PoD may be represented by a subnet A.B.C.D/E, where E may be the subnet mask length. The last IP address in the IP address range may be referred to as a special IP address, i.e., the replication IP address. For example, if the server subnet of the whole PoD is 192.168.0.0/24, the last IP address 192.168.0.255 may be reserved as the replication IP address. The replication IP address may be reserved internally without being assigned to any server. Additionally or alternatively, one or more other IP addresses within the IP address range may be reserved as replication IP address(es).

[22] Referring to FIG. 2, the Tier-1 layer may include a Tier-1 SW 202 and a Tier-1 SW 204. The Tier-1 SW 202 and the Tier-1 SW 204 may be implemented as programmable switches to introduce a new network protocol which may not be realized by an existing black box switch. The switch ASIC (not shown) in the Tier-1 SW 202 and the Tier-1 SW 204 may be able to create the predetermined number of copies of data packets without consuming the CPU resource. Though FIG. 2 shows two Tier-1 switches, the Tier-1 layer may include other number of switches.

[23] The TOR layer may include a TOR 206, a TOR 208, and a TOR 210. Though FIG. 2 shows three TORs, the TOR layer may include other number of TORs. The Tier-1 SW 202 may be coupled to the TOR 206, the TOR 208, and the TOR 210. The Tier-1 SW 204 may also be coupled to the TOR 206, the TOR 208, and the TOR 210.

[24] The TOR 206 may be coupled to each server of a rack 212. The TOR 208 may be coupled to each server of a rack 214. The TOR 210 may be coupled to each server of a rack 216. Another party, for example, a sending server, a user device, a computing device, a process, an application, etc. may send data to the system 200. A sending server 218 and a first server 220 may be in the rack 212. A second server 222 may be in the rack 214. A third server

224 and a fourth server 226 may be in the rack216.

[25] An arrow 228 presents that the sending server 218 may send one or more packets to the Tier-1 SW 202. For example, the sending server 218 in the rack 212 or a party outside the network 200 may receive data from a user (not shown), and the sending server 218 may encapsulate the user data into an original packet to be sent to the Tier-1 SW 202. Additionally or alternatively, the user data may be encapsulated by the sending party, for example, a user device, a computing device, a process, an application, etc. The original packet may include a source address and a destination address, where the source address may be the IP address of the sending server 218 and the destination address may be the reserved IP address, for example, 192.168.0.255, which may be the replication IP address. Because the reserved IP address 192.168.0.255 may not be assigned to any server, the original packet from the sending server 218 with the reserved IP address as the destination IP address may be routed to one of the top layer switches in the Clos network 200, for example, the Tier-1 SW 202.

[26] The Tier-1 SW 202 may identify the original packet as to be copied if the destination IP address of the packet is determined as the reserved IP address. The Tier-1 SW 202 may create a predetermined number of copies. The predetermined number may be set in advance according to requirements of the cloud storage service. Example requirements may include a data storage requirement, a data reliability requirement, a data availability requirement, etc.

[27] The predetermined number of copies may include a first copy, a second copy, a third copy, ..., and an n^{th} copy, where n is the predetermined number. Thus, each of the predetermined number of copies may be the same as the original packet sent from the sending party, for example, the sending server 218. Each of the predetermined number of copies may have the same source IP address and the same destination IP address as the original packet. The destination IP address of the first copy may be referred to as the first destination IP address. The destination IP address of the second copy may be referred to as the second destination IP address. The destination IP address of the third copy may be referred to as the third destination IP address. The destination IP address of the n^{th} copy may be referred to as the n^{th} destination IP address, where n is the predetermined number.

[28] The Tier-1 SW 202 may select the predetermined number of servers in the PoD. For example, the first server 220 in the rack 212, the second server 222 in the rack 214, and the third server 224 in the rack 216, ..., an n^{th} server (not shown) may be selected. The selected servers may be in different racks or the same rack. The selection of the predetermined number of servers may be performed by the Tier-1 SW 202 randomly, in a rotating manner, using a hash function, or based on any other suitable criteria. For example, idle servers, servers with high capacities, servers that are not busy, etc. may be selected. Moreover, the reserved IP address, for example, 192.168.0.255, may be exempted from the selection because the reserved IP address may be used as an indicator for the Tier-1 SW to generate predetermined number of copies of the original packet.

[29] The Tier-1 SW 202 may replace the destination IP address of each of the predetermined number of copies using the IP address of each of the selected predetermined number of servers. For example, the Tier-1 SW 202 may replace the first destination IP address of the first copy using an IP address of the first server 220. The Tier-1 SW 202 may replace the second destination IP address of the second copy using an IP address of the second server 222. The Tier-1 SW 202 may replace the third destination IP address of the third copy using an IP address of the third cloud storage server 224. The Tier-1 SW 202 may replace the n^{th} destination IP address of the n^{th} copy using an IP address of the n^{th} cloud storage server 224.

[30] An arrow 230 represents that the Tier-1 SW 202 may route the first copy to the first server 220, for example, through the TOR 206. The TOR 206 may receive the first copy from the Tier-1 SW 202 and identify the first destination IP address of the first copy as the destination IP address of the first server 220 and forward the first copy to the first server 220.

[31] The arrow 232 represents that the Tier-1 SW 202 may route the second copy to the second server 222, for example, through the TOR 208. The TOR 208 may receive the second copy from the Tier-1 SW 202 and identify the second destination IP address of the second copy as the destination IP address of the second server 222 and forward the second copy to the second server 222.

[32] The arrow 234 may represent that the Tier-1 SW 202 may route the third copy to the third server 224, for example, through the TOR 210. The TOR 210 may receive the third

copy from the Tier-1 SW 202 and identify the third destination IP address of the third copy as the destination IP address of the third server 224 and forward the third copy to the third server 224.

[33] Though three arrows 230, 232, and 234 are described in this example, other
5 arrows may be added to represent that the Tier-1 SW 202 may forward the n^{th} copy to the n^{th} server.

[34] In case of failure of any of the selected predetermined number of servers, for example, the first server 220 may fail, crash, be too busy to respond, etc., the attempt of routing the first copy to the first server may be unsuccessful. The Tier-1 SW 202 may
10 generate an $(n+1)^{\text{th}}$ copy of the original packet, select a $(n+1)^{\text{th}}$ server from the same PoD, and route the $(n+1)^{\text{th}}$ copy to a $(n+1)^{\text{th}}$ server 226, where n is the predetermined number. The $(n+1)^{\text{th}}$ copy may have the same source IP address and the same destination IP address as the original packet. The destination IP address of the $(n+1)^{\text{th}}$ copy may be referred to as the $(n+1)^{\text{th}}$ destination IP address. The selection of the $(n+1)^{\text{th}}$ server by the Tier-1 SW 202 may
15 be performed randomly, in a rotating manner, using a hash function, or based on any other suitable criteria. For example, idle servers, servers with high capacities, servers that are not busy, etc. may be selected. Moreover, the reserved IP address, for example, 192.168.0.255, may be exempted from the selection because the reserved IP address may be used as an indicator for the Tier-1 SW to generate predetermined number of copies of the original packet.
20 The Tier-1 SW 202 may replace the $(n+1)^{\text{th}}$ destination IP address of the $(n+1)^{\text{th}}$ copy using an IP address of the $(n+1)^{\text{th}}$ server 226.

[35] An arrow 236 represents that the Tier-1 SW 202 may route the $(n+1)^{\text{th}}$ copy to the $(n+1)^{\text{th}}$ server 226, for example, through the TOR 210. The TOR 210 may receive the $(n+1)^{\text{th}}$ copy from the Tier-1 SW 202 and identify the $(n+1)^{\text{th}}$ destination IP address of the
25 $(n+1)^{\text{th}}$ copy as the destination IP address of the $(n+1)^{\text{th}}$ server 224 and forward the $(n+1)^{\text{th}}$ copy to the $(n+1)^{\text{th}}$ server 226.

[36] FIG.3A and FIG.3B illustrate example flowcharts of a process 300 for network optimization of cloud storage service. Referring to FIG.3A and FIG.3B, the process 300 may include the following.

[37] At block 302, the Tier-1 SW 202 may receive the original packet from a sending
30

party, for example, the sending server 218, a user device, a computing device, a process, an application, etc. Also referring to FIG. 2, the sending server 218 in the rack 212 or a party outside the network 200 may receive data from a user (not shown), and the sending server 218 may encapsulate the user data into an original packet to be sent to the Tier-1 SW 202.

5 Additionally or alternatively, the user data may be encapsulated by the sending party, for example, a user device, a computing device, a process, an application, etc. The original packet may include a source address and a destination address, where the source address may be the IP address of the sending server 218 and the destination address may be the reserved IP address, for example, 192.168.0.255, which may be the replication IP address. Because the
10 reserved IP address 192.168.0.255 may not be assigned to any server, the original packet from the sending server 218 with the reserved IP address as the destination IP address may be routed to one of the top layer switches in the Clos network 200, for example, the Tier-1 SW 202.

[38] At block 304, the Tier-1 SW 202 may determine whether the destination IP
15 address of the original packet is a reserved IP address. The Tier-1 SW 202 may identify the original packet as to be copied if the destination IP address of the packet is determined as the reserved IP address.

[39] At block 306, the Tier-1 SW 202 may create a predetermined number of copies. The predetermined number may be set in advance according to requirements of the cloud
20 storage service. Example requirements may include a data storage requirement, a data reliability requirement, a data availability requirement, etc.

[40] The predetermined number of copies may include a first copy, a second copy, a third copy, ..., and an n^{th} copy, where n is the predetermined number. Thus, each of the predetermined number of copies may be the same as the original packet sent from the sending
25 party, for example, the sending server 218. Each of the predetermined number of copies may have the same source IP address and the same destination IP address as the original packet. The destination IP address of the first copy may be referred to as the first destination IP address. The destination IP address of the second copy may be referred to as the second destination IP address. The destination IP address of the third copy may be referred to as the
30 third destination IP address. The destination IP address of the n^{th} copy may be referred to as

the n^{th} destination IP address, where n is the predetermined number.

[41] At block 308, The Tier-1 SW 202 may select the predetermined number of servers. The selected servers may be in different racks or the same rack. The selection of servers may be performed by the Tier-1 SW 202 randomly, in a rotating manner, using a hash function, or based on any other suitable criteria. For example, idle servers, servers with high capacities, servers that are not busy, etc. may be selected. Moreover, the reserved IP address, for example, 192.168.0.255, may be exempted from the selection because the reserved IP address may be used as an indicator for the Tier-1 SW to generate predetermined number of copies of the original packet.

[42] At block 310, The Tier-1 SW 202 may replace the destination IP address of each of the predetermined number of copies using the IP address of each of the selected predetermined number of servers. Also referring to FIG. 2, the Tier-1 SW 202 may replace the first destination IP address of the first copy using an IP address of the first server 220. The Tier-1 SW 202 may replace the second destination IP address of the second copy using an IP address of the second server 222. The Tier-1 SW 202 may replace the third destination IP address of the third copy using an IP address of the third cloud storage server 224. The Tier-1 SW 202 may replace the n^{th} destination IP address of the n^{th} copy using an IP address of the n^{th} cloud storage server 224.

[43] At block 312, the Tier-1 SW 202 may forward the predetermined number of copies to the predetermined number of servers respectively. Also referring to FIG. 2, the Tier-1 SW 202 may route the first copy to the first server 220, for example, through the TOR 206. The TOR 206 may receive the first copy from the Tier-1 SW 202 and identify the first destination IP address of the first copy as the destination IP address of the first server 220 and forward the first copy to the first server 220. The Tier-1 SW 202 may route the second copy to the second server 222, for example, through the TOR 208. The TOR 208 may receive the second copy from the Tier-1 SW 202 and identify the second destination IP address of the second copy as the destination IP address of the second server 222 and forward the second copy to the second server 222. The Tier-1 SW 202 may route the third copy to the third server 224, for example, through the TOR 210. The TOR 210 may receive the third copy from the Tier-1 SW 202 and identify the third destination IP address of the third copy as the destination

IP address of the third server 224 and forward the third copy to the third server 224. The Tier-1 SW 202 may forward the n^{th} copy to the n^{th} server.

[44] After receiving each of the predetermined number of copies, each of the selected predetermined number of servers may send an acknowledgment message (ACK) back to the Tier-1 SW 202 to confirm that the copy has been received. Therefore, after the block 312, the process 300 may further include the following as shown in FIG. 3B.

[45] Referring to FIG. 3B, at block 314, each of the predetermined number of servers may receive the first copy, the second copy, the third copy, ..., or the n^{th} copy from the Tier-1 SW. For example, the first server 220 may receive the first copy from the Tier-1 SW 202. The second server 222 may receive the second copy from the Tier-1 SW 202. The third server 224 may receive the third copy from the Tier-1 SW 202. The n^{th} server may receive the n^{th} copy from the Tier-1 SW 202.

[46] At block 316, each of the predetermined number of servers may send an ACK to the Tier-1 SW 202. For example, the first server 220 may send an ACK to the Tier-1 SW 202. The second server 222 may send an ACK to the Tier-1 SW 202. The third server 224 may send an ACK to the Tier-1 SW 202. The n^{th} server may send an ACK to the Tier-1 SW 202.

[47] In case of failure of any of the selected predetermined number of servers, for example, the server may fail, crash, be too busy to respond, etc., the server may not be able to send the ACK back to the Tier-1 SW 202. Without receiving the ACK within a certain period, the Tier-1 SW 202 may fall back to a previous step of generating copies of the original packet and sending the copies to other servers.

[48] At block 318, the Tier-1 SW 202 may create an $(n+1)^{\text{th}}$ copy of the original packet if the Tier-1 SW 202 does not receive the ACK from any of the selected predetermined number of servers within a predetermined period. The predetermined period may be set as necessary. The $(n+1)^{\text{th}}$ copy may have the same source IP address and the same destination IP address as the original packet. The destination IP address of the fourth copy may be referred to as the $(n+1)^{\text{th}}$ destination IP address.

[49] At block 320, the Tier-1 SW 202 may select an $(n+1)^{\text{th}}$ server 226. The selection of the $(n+1)^{\text{th}}$ server by the Tier-1 SW 202 may be performed by the Tier-1 SW 202 randomly, in a rotating manner, using a hash function, or based on any other suitable criteria. For

example, idle servers, servers with high capacities, servers that are not busy, etc. may be selected. Moreover, the reserved IP address, for example, 192.168.0.255, may be exempted from the selection because the reserved IP address may be used as an indicator for the Tier-1 SW to generate predetermined number of copies of the original packet.

5 [50] At block 322, the Tier-1 SW 202 may replace the $(n+1)^{\text{th}}$ destination IP address of the $(n+1)^{\text{th}}$ copy using an IP address of the $(n+1)^{\text{th}}$ server 226.

[51] At block 324, the Tier-1 SW 202 may forward the $(n+1)^{\text{th}}$ copy to the $(n+1)^{\text{th}}$ server. Also referring to FIG. 2, the Tier-1 SW 202 may route the $(n+1)^{\text{th}}$ copy to the $(n+1)^{\text{th}}$ server 226, for example, through the TOR 210. The TOR 210 may receive the $(n+1)^{\text{th}}$ copy
10 from the Tier-1 SW 202 and identify the $(n+1)^{\text{th}}$ destination IP address of the $(n+1)^{\text{th}}$ copy as the destination IP address of the $(n+1)^{\text{th}}$ server 224 and forward the $(n+1)^{\text{th}}$ copy to the fourth server 226.

[52] With systems and processes discussed herein, if one server is down, since there are copies of the data saved in other servers, the data may be recovered, thereby improving
15 the availability of the cloud storage service.

[53] FIG. 4 illustrates an example block diagram of an apparatus 400 for optimizing cloud storage service.

[54] FIG. 4 is only one example of an apparatus 400 and is not intended to suggest any limitation as to the scope of use or functionality of any computing device utilized to
20 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,
25 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.

[55] The apparatus 400 may include one or more processors 402 and memory 404 communicatively coupled to the processor(s) 402. The processor(s) 402 may execute one or
30 more modules and/or processes to cause the processor(s) 402 to perform a variety of functions.

In some embodiments, the processor(s) 402 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) 402 may possess its own local memory, which also may store program modules, program data, and/or one or more operating
5 apparatuses.

[56] Depending on the exact configuration and type of the apparatus 400, the memory 404 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 404 may include computer-executable instructions that are executable by the processor(s) 402, when executed
10 by the processor(s) 402, cause the processor(s) 402 to implement systems and processes described with reference to FIG. 1 FIG. 2, FIG. 3A, FIG. 3B.

[57] The apparatus 400 may additionally include an input/output (I/O) interface 406 for receiving and outputting data. The apparatus 400 may also include a communication module 408 allowing the apparatus 400 to communicate with other devices (not shown) over
15 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.

[58] With systems and processes discussed herein, a sending server only needs to send out the original packet through the network once. Thus, the CPU power of the sending
20 server may be saved. Because the packet may be copied by a top layer switch, for example, Tier-1 SW rather than the sending server, the network load and cost may be reduced significantly.

[59] 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,
25 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
30 devices, microprocessor-based, programmable consumer electronics, combinations thereof,

and the like.

[60] 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
5 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.

[61] A non-transient computer-readable storage medium is an example of computer-readable media. Computer-readable media includes at least two types of
10 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
15 (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
20 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.

[62] The computer-readable instructions stored on one or more non-transitory
25 computer-readable storage media that, when executed by one or more processors, may perform operations described above with reference to FIG. 1- FIG. 4. Generally, computer-readable instructions include routines, programs, objects, components, data structures, and the like that perform particular functions or implement particular abstract data
30 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

[63] Clause 1. A method comprising: receiving, by a switch, a packet; creating, by
5 the switch, one or more copies of the packet; selecting, by the switch, a plurality of servers;
and separately forwarding, by the switch, the packet and the one or more copies to the
plurality of servers.

[64] Clause 2. The method of clause 1, wherein prior to creating by the switch the
one or more copies of the packet, the method further comprises determining, by the switch,
10 whether a destination Internet Protocol (IP) address of the packet is a reserved IP address, and
wherein the switch creates the one or more copies of the packet upon determining that the
destination IP address of the packet is the reserved IP address.

[65] Clause 3. The method of clause 1, wherein the plurality of servers includes a
predetermined number of servers, the predetermined number being determined based on a
15 particular policy.

[66] Clause 4. The method of clause 3, further comprising replacing, by the switch,
destination IP addresses of the packet and the one or more copies of the packet by respective
server IP addresses of the plurality of servers.

[67] Clause 5. The method of clause 1, wherein selecting by the switch the
20 plurality of servers comprises selecting the plurality of servers randomly.

[68] Clause 6. The method of clause 1, further comprising receiving, by the switch,
an acknowledgement message (ACK) from a first server of the plurality of servers.

[69] Clause 7. The method of clause 1, further comprising: creating, by the switch,
an additional copy of the packet if the switch does not receive an acknowledgement message
25 (ACK) from a first server of the plurality of servers within a predetermined period; selecting,
by the switch, a new server; and forwarding, by the switch, the additional copy to the new
server.

[70] Clause 8. 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 acts comprising: receiving, by a switch, a packet; creating, by the switch, one or more copies of the packet; selecting, by the switch, a plurality of servers; and separately forwarding, by the switch, the packet and the one or more copies to the plurality of servers.

[71] Clause 9. The apparatus of clause 8, wherein prior to creating by the switch the one or more copies of the packet, the acts further comprise determining, by the switch, whether a destination Internet Protocol (IP) address of the packet is a reserved IP address, and wherein the switch creates the one or more copies of the packet upon determining that the destination IP address of the packet is the reserved IP address.

[72] Clause 10. The apparatus of clause 8, wherein a predetermined number of servers, the predetermined number being determined based on a particular policy.

[73] Clause 11. The apparatus of clause 10, wherein the acts further comprise replacing, by the switch, destination IP addresses of the packet and the one or more copies of the packet by respective server IP addresses of the plurality of servers.

[74] Clause 12. The apparatus of clause 8, wherein selecting by the switch the plurality of servers comprises selecting the plurality of servers randomly.

[75] Clause 13. The apparatus of clause 8, wherein the acts further comprise receiving, by the switch, an acknowledgement message (ACK) from a first server of the plurality of servers.

[76] Clause 14. The apparatus of clause 8, wherein the acts further comprise: creating, by the switch, an additional copy of the packet if the switch does not receive an acknowledgement message (ACK) from a first server of the plurality of servers within a predetermined period; selecting, by the switch, a new server; and forwarding, by the switch, the additional copy to the new server.

[77] Clause 15. 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 acts comprising: receiving, by a switch, a packet; creating, by the switch, one or more copies of the packet; selecting, by the switch, a plurality of servers; and separately forwarding, by the switch, the packet and the one

or more copies to the plurality of servers.

[78] Clause 16. The computer-readable storage medium of clause 15, wherein prior to creating by the switch the one or more copies of the packet, the acts further comprise determining, by the switch, whether a destination Internet Protocol (IP) address of the packet is a reserved IP address, and wherein the switch creates the one or more copies of the packet upon determining that the destination IP address of the packet is the reserved IP address.

[79] Clause 17. The computer-readable storage medium of clause 15, wherein a predetermined number of servers, the predetermined number being determined based on a particular policy.

[80] Clause 18. The computer-readable storage medium of clause 17 wherein the acts further comprise replacing, by the switch, destination IP addresses of the packet and the one or more copies of the packet by respective server IP addresses of the plurality of servers.

[81] Clause 19. The computer-readable storage medium of clause 15, wherein selecting by the switch the plurality of servers comprises selecting the plurality of servers randomly.

[82] Clause 20. The computer-readable storage medium of clause 15, wherein the acts further comprise receiving, by the switch, an acknowledgement message (ACK) from a first server of the plurality of servers.

Conclusion

[83] 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, by a switch, a packet;
creating, by the switch, one or more copies of the packet;
5 selecting, by the switch, a plurality of servers; and
separately forwarding, by the switch, the packet and the one or more copies to the plurality of servers.
2. The method of claim 1, wherein prior to creating by the switch the one or more copies of the packet, the method further comprises determining, by the switch, whether a
10 destination Internet Protocol (IP) address of the packet is a reserved IP address, and wherein the switch creates the one or more copies of the packet upon determining that the destination IP address of the packet is the reserved IP address.
3. The method of claim 1, wherein the plurality of servers includes a predetermined number of servers, the predetermined number being determined based on a particular policy.
- 15 4. The method of claim 3, further comprising replacing, by the switch, destination IP addresses of the packet and the one or more copies of the packet by respective server IP addresses of the plurality of servers.
5. The method of claim 1, wherein selecting by the switch the plurality of servers comprises selecting the plurality of servers randomly.
- 20 6. The method of claim 1, further comprising receiving, by the switch, an acknowledgement message (ACK) from a first server of the plurality of servers.
7. The method of claim 1, further comprising:
creating, by the switch, an additional copy of the packet if the switch does not receive an acknowledgement message (ACK) from a first server of the plurality of servers within a
25 predetermined period;

selecting, by the switch, a new server; and

forwarding, by the switch, the additional copy to the new server.

8. An apparatus, comprising:

one or more processors;

5 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 acts comprising:

receiving, by a switch, a packet;

creating, by the switch, one or more copies of the packet;

10 selecting, by the switch, a plurality of servers; and

separately forwarding, by the switch, the packet and the one or more copies to the plurality of servers.

9. The apparatus of claim 8, wherein prior to creating by the switch the one or more copies of the packet, the acts further comprise

15 determining, by the switch, whether a destination Internet Protocol (IP) address of the packet is a reserved IP address, and wherein the switch creates the one or more copies of the packet upon determining that the destination IP address of the packet is the reserved IP address.

20 10. The apparatus of claim 8, wherein a predetermined number of servers, the predetermined number being determined based on a particular policy.

11. The apparatus of claim 10, wherein the acts further comprise replacing, by the switch, destination IP addresses of the packet and the one or more copies of the packet by respective server IP addresses of the plurality of servers.

25 12. The apparatus of claim 8, wherein selecting by the switch the plurality of servers comprises selecting the plurality of servers randomly.

13. The apparatus of claim 8, wherein the acts further comprise receiving, by the switch, an acknowledgement message (ACK) from a first server of the plurality of servers.

14. The apparatus of claim 8, wherein the acts further comprise:

creating, by the switch, an additional copy of the packet if the switch does not receive
5 an acknowledgement message (ACK) from a first server of the plurality of servers within a predetermined period;

selecting, by the switch, a new server; and

forwarding, by the switch, the additional copy to the new server.

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

receiving, by a switch, a packet;

creating, by the switch, one or more copies of the packet;

selecting, by the switch, a plurality of servers; and

15 separately forwarding, by the switch, the packet and the one or more copies to the plurality of servers.

16. The computer-readable storage medium of claim 15, wherein prior to creating by the switch the one or more copies of the packet, the acts further comprise:

determining, by the switch, whether a destination Internet Protocol (IP) address of the
20 packet is a reserved IP address, and wherein the switch creates the one or more copies of the packet upon determining that the destination IP address of the packet is the reserved IP address.

17. The computer-readable storage medium of claim 15, wherein a predetermined number of servers, the predetermined number being determined based on a particular policy.

25 18. The computer-readable storage medium of claim 17, wherein the acts further

comprise replacing, by the switch, destination IP addresses of the packet and the one or more copies of the packet by respective server IP addresses of the plurality of servers.

19. The computer-readable storage medium of claim 15, wherein selecting by the switch the plurality of servers comprises selecting the plurality of servers randomly.

5 20. The computer-readable storage medium of claim 15, wherein the acts further comprise receiving, by the switch, an acknowledgement message (ACK) from a first server of the plurality of servers.

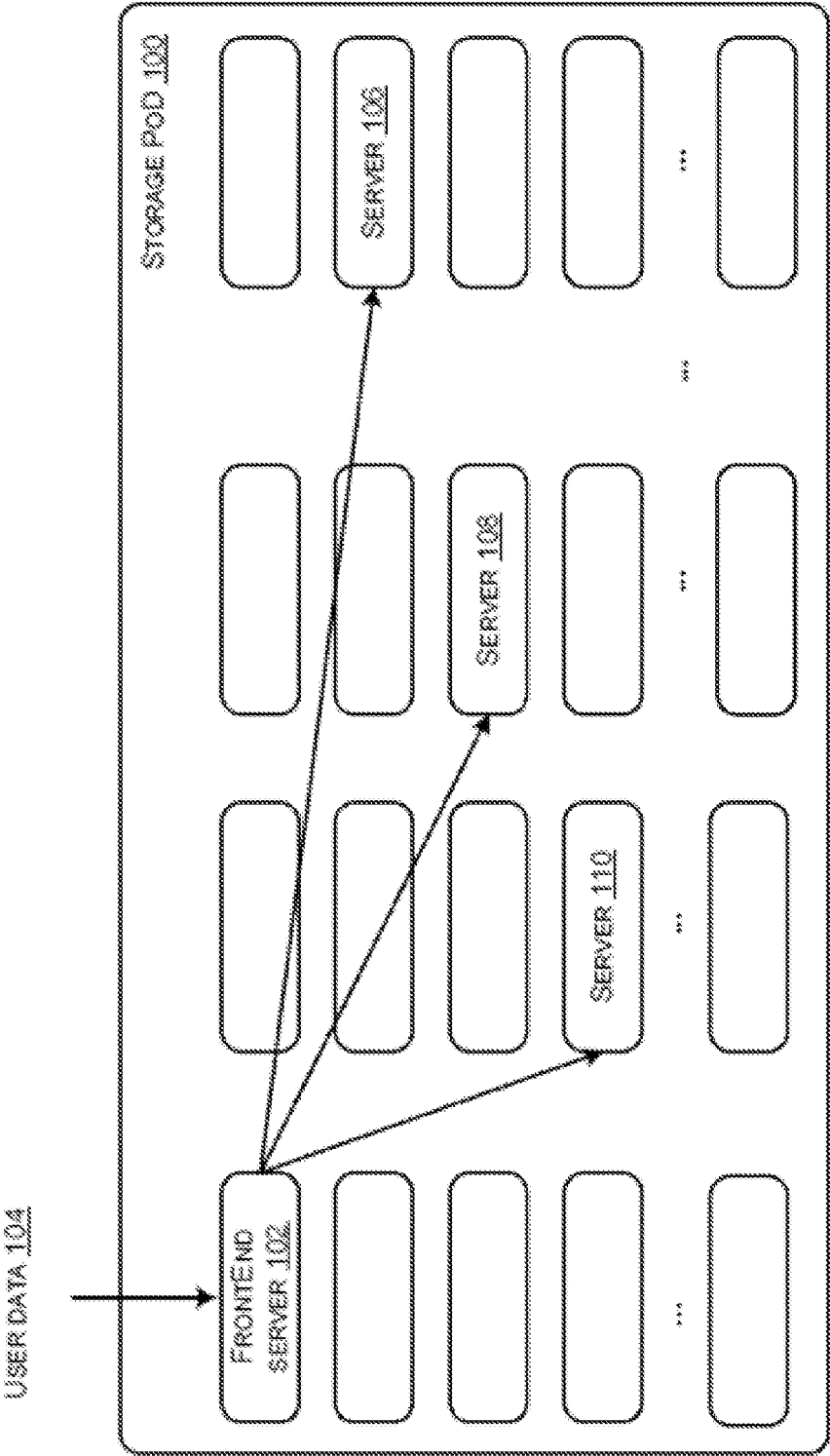


FIG. 1

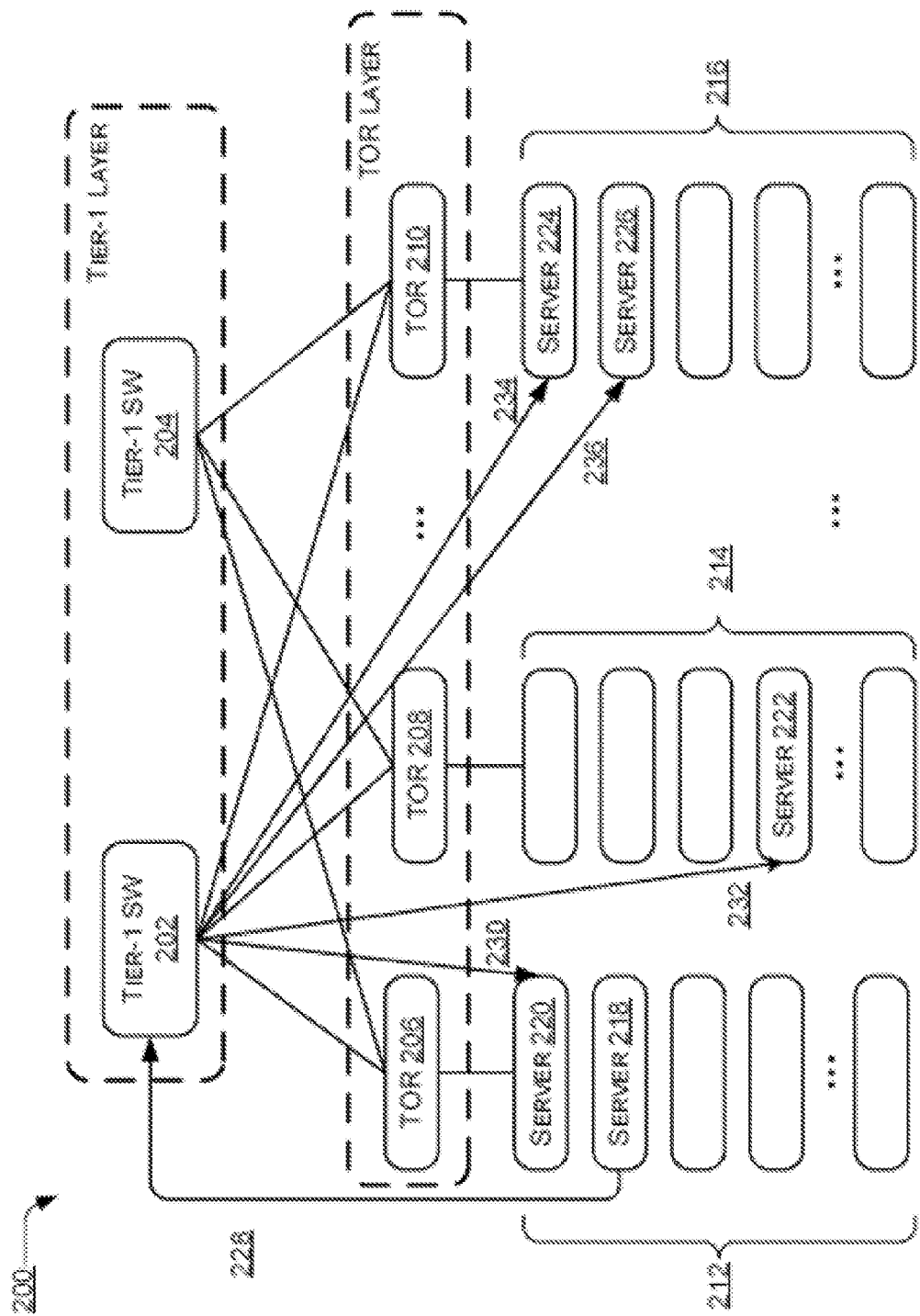
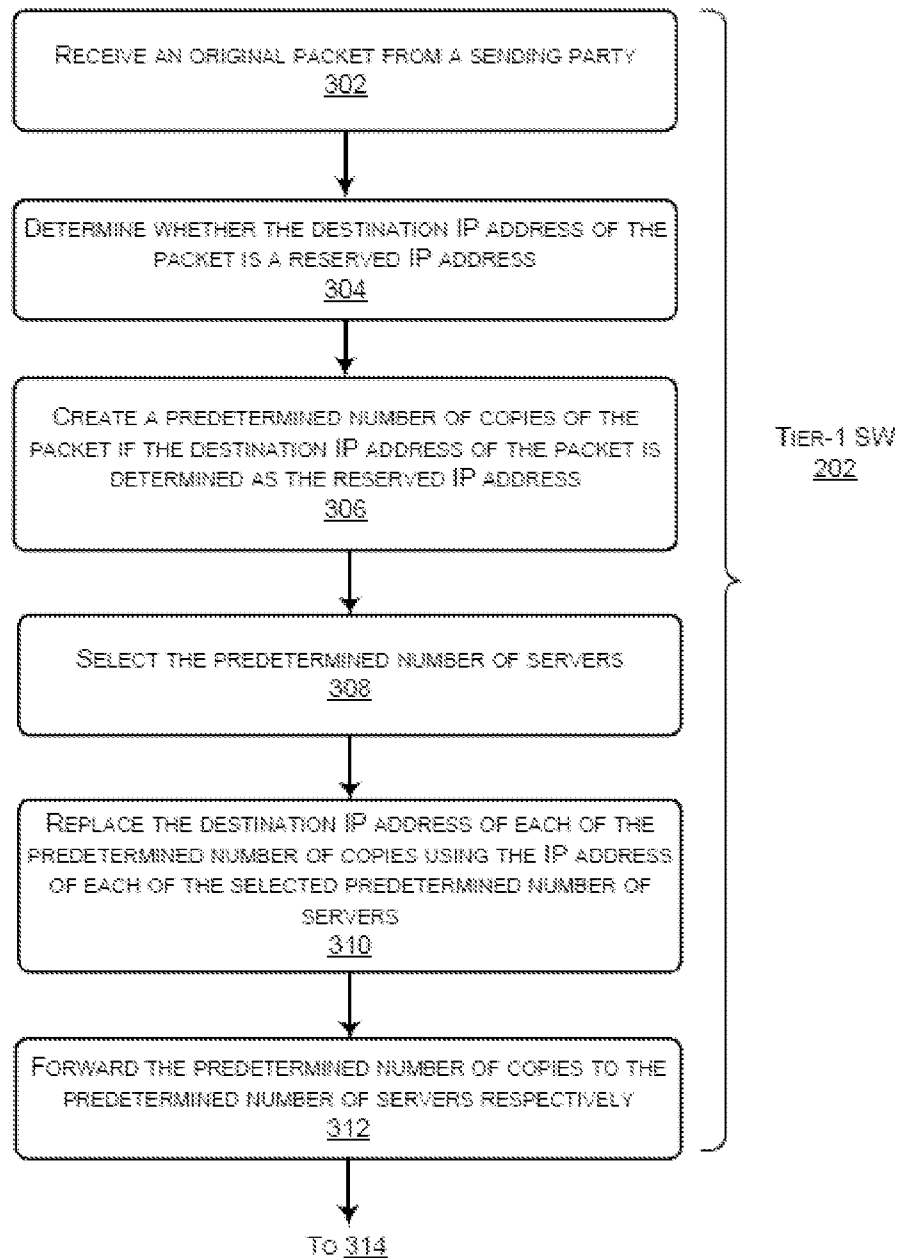


FIG. 2

3 / 5

300**FIG. 3A**

4 / 5

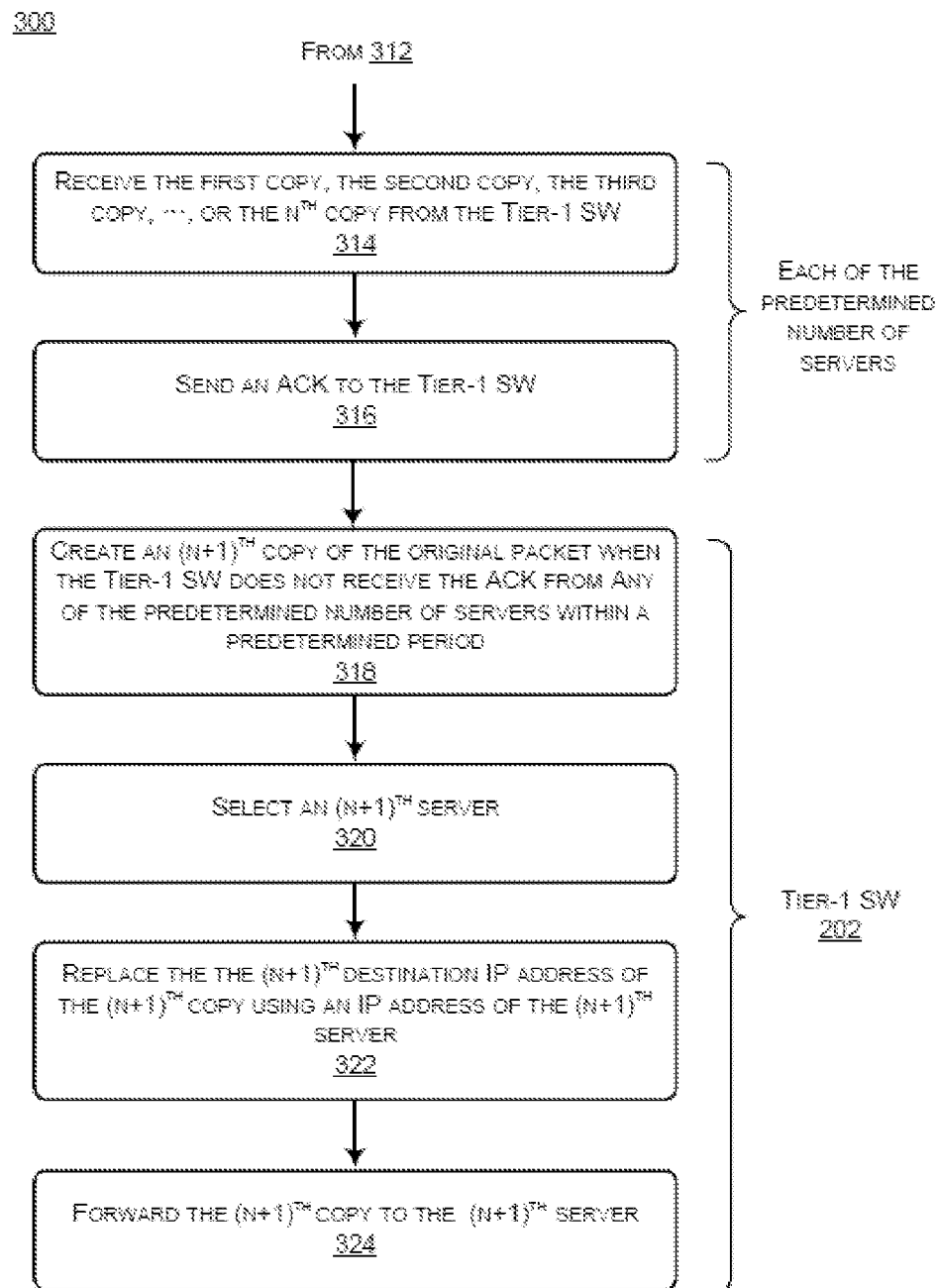
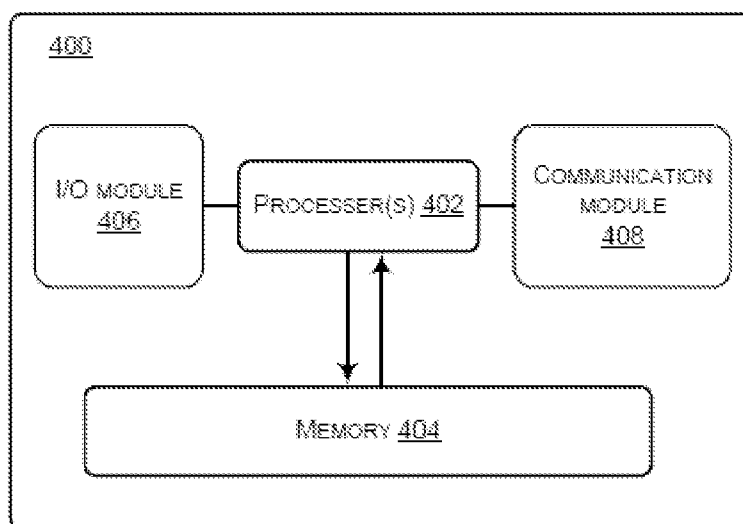


FIG. 3B

5 / 5

**FIG. 4**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/126621

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/28(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

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: switch, cloud, packet, copy, duplicate, create, server, forward, distribute

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 7830883 B1 (WORLD WIDE PACKETS, INC.) 09 November 2010 (2010-11-09) description, column 4 line 11 to column 6 line 21, column 11 line 48 to column 12 line 56, figures 2, 6 and claims 1-2	1-20
Y	US 2017353351 A1 (ALIBABA GROUP HOLDING LIMITED) 07 December 2017 (2017-12-07) description, paragraphs [0001]-[0002]	1-20
Y	US 8018938 B1 (WORLD WIDE PACKETS, INC.) 13 September 2011 (2011-09-13) description, column 12 line 44 to line 61 and figures 2, 6	1-20
Y	CN 202679422 U (BEIJING IDC TECHNOLOGY CO., LTD.) 16 January 2013 (2013-01-16) description, paragraph [0012] and figure 1	1-20
A	US 2015215405 A1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 30 July 2015 (2015-07-30) 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

06 March 2020

Date of mailing of the international search report

20 March 2020

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC
6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing
100088
China

Authorized officer

MA, Xiaoyu

Facsimile No. (86-10)62019451

Telephone No. 86-(10)-53961352

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/126621

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	7830883	B1	09 November 2010	None			
US	2017353351	A1	07 December 2017	CN	107465590	A	12 December 2017
US	8018938	B1	13 September 2011	None			
CN	202679422	U	16 January 2013	None			
US	2015215405	A1	30 July 2015	KR	20150088442	A	03 August 2015