February 2022

SYSTEMS AND METHODS TO PROVIDE ANYCAST MPLS PING UTILITY IN AN IGP-SR NETWORK

Anonymous

SYSTEMS AND METHODS TO PROVIDE ANYCAST MPLS PING UTILITY IN AN IGP-SR NETWORK

ABSTRACT

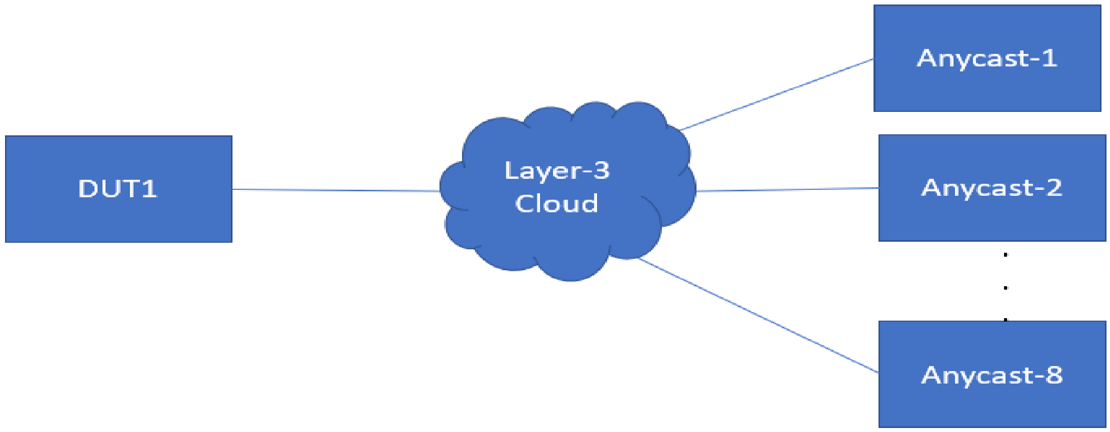
The present disclosure relates to systems and methods to provide Anycast MPLS Ping utility in an IGP-SR network. More particularly, to provide Anycast MPLS Ping utility in IGP-SR to detect the reachability/liveliness of the anycast prefixes. The present disclosure introduces the anycast ping, where network operators can validate the reachability to all anycast routers in one go by initiating the ping to anycast IP. As a part of the solution, the initiating router will walk through its own IGP database to determine the router IDs of all the routers which are advertising the destination prefix. If the destination prefix is available from more than one advertising routers then it will be marked as qualified for anycast ping, otherwise the reachability to the destination IP will be determined by normal ping.

DETAILED DESCRIPTION

Anycast IP, is a networking technique that allows for multiple machines to share the same IP address. Based on the location of the user request, the routers send it to the machine in the network that is closest. In other words, instead of forwarding to a specific device or to all devices in a group of, anycast addresses, network devices forward a packet to (or steer it through) one or more topologically nearest devices in a specific group of network devices. This is beneficial since, among other things, it reduces latency and increases redundancy. If one of the routers were to go offline, an anycast IP would choose the best path for users and automatically redirect them to the next closest router. Anycast SID (segment identifier), the use of anycast IP addresses has been extended to the Segment Routing (SR) network, wherein a group of SR-capable devices can be treated as anycast nodes, if they have the same IP and segment identifier.

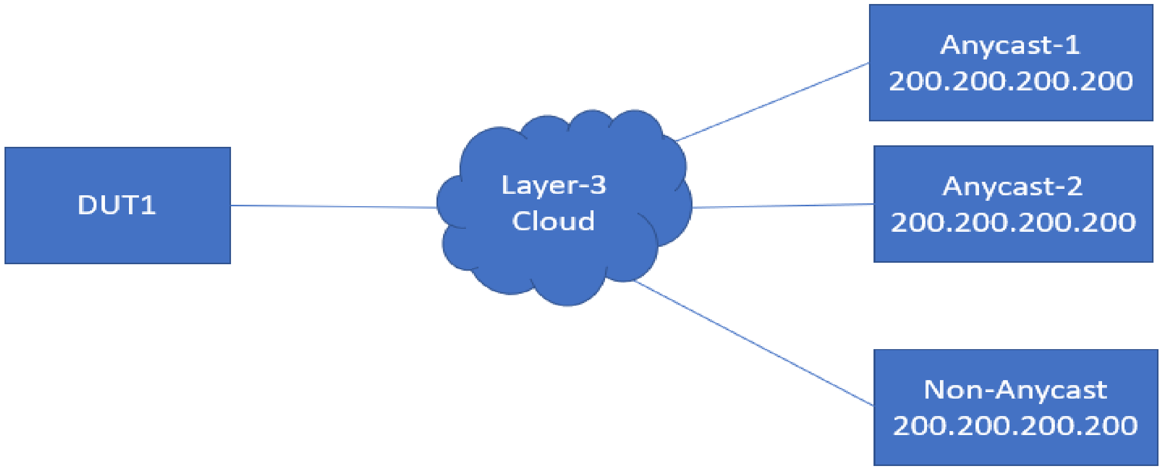
Same IP address on multiple routers to provide redundancy and load sharing is the main use case of Anycast feature. In a large deployment, Network operator can configure multiple anycast routers. In such a large deployment without having this invention, operator can face following problems/challenges. There are no methods defined to simultaneously determine the reachability/liveliness of all the anycast routers of an anycast group. So, if there are say 8 anycast routers in an anycast group (assume ~5k routers in a network) then network operators will have to individually check the reachability to each of the routers, which may be very time consuming (keeping in mind the size of network) while troubleshooting the network issues. Verification of the individual router reachability may not deliver the accurate results due to intermittent network issues. If there is any misconfiguration in the network due to which multiple routers are advertising the same prefix and SID, then there is no easy way to find out the problematic routers, other than going through the entire ISIS database (having thousands of routes) to filter out the advertising routers. Because of above scenario's/limitation and lack of this invention, a network operator can face a lot of problem/difficulty while debugging the Anycast network.

Figure 1: Explination of problems



In figure 1, DUT1 is connected in a topology which has 8 anycast routers in the anycast group. If the operator wants to determine the reachability from DUT1 to anycast routers, then he will be able to determine the reachability to only two anycast routers (primary and backup anycast routers) by initiating the mpls ping. There is no easy way to determine if other anycast routers are alive or not. In figure 1, DUT1 is connected in a topology which has 8 anycast routers in the anycast group and traffic is getting load shared across all anycast routers. Consider that customer is getting drops to the destination while going through the anycast routers. Now to determine the problematic anycast routers, he will have to check ping to each of the anycast routers individually. If the packet drop is intermittent, then operator might not be able to find the problematic router in the 1st instance and he will have to repeat the ping process to all anycast routers multiple times, which is very time consuming and tedious task.

Figure 2: Explination of problems continued



In figure 2, DUT1 is connected in a topology which has 2 anycast routers in the anycast group which are advertising the anycast IP 200.200.200.200/32 and one non anycast router also started advertising the same prefix 200.200.200.200/32. This will result into IP duplicity in the network, but there is no easy way to detect this other than going through the entire IGP database which may contain thousands of routes.

The present disclosure introduces the anycast ping, where network operators can validate the reachability to all anycast routers in one go by initiating the ping to anycast IP. As a part of the solution, the initiating router will walk through its own IGP database to determine the router IDs of all the routers which are advertising the destination prefix. If the destination prefix is available from more than one advertising routers then it will be marked as qualified for anycast ping, otherwise the reachability to the destination IP will be determined by normal ping. As part of the anycast ping solution, the initiating router will construct an ICMP Echo Request (MPLS Echo Request) packet for each of the advertising routers with destination IP as anycast IP and outgoing label as the Node SID of the advertising routers. So, Operators will be able to initiate ping test to all anycast routers at the same time. Based on the ICMP Echo Response (MPLS Echo Response) from the anycast routers, operators can determine the liveliness of all anycast routers which are part of the anycast group. Also, in case of misconfiguration where same prefix and SID are advertised from multiple routers, Anycast ping can help to determine the problematic routers by logging the router IDs of the advertising routers in the Anycast ping response output. By default, when an MPLS ECHO (ICMP) packet is created by a device, it looks up the FTN table and determines the outgoing label to reach that destination. That MPLS label is appended to the ICMP Packet and packet is forwarded through the MPLS domain from source to the destination.

Since Anycast IP is configured on many routers, it is not possible to determine the reachability of all anycast routers by generating the MPLS ECHO packet just by referring the FTN table of the originating node. So, as part of the present disclosure, the MPLS ping method will be modified for anycast destination, which will determine the Node SID of all the anycast routers (from which anycast destination is reachable) and construct the MPLS ECHO packets with destination IP as anycast address and outing going label as the Node SID of the advertising router. This will ensure that ICMP ECHO packet is delivered to the desired destination router. This problem cannot be solved by any other presently known ICMP protocol techniques.

The present invention eends ICMP requests to all the anycast routers and will be receiving the ICMP replies from all the anycast routers in an IGP network, ICMP identification number is used to match the ICMP request and ICMP reply.

According to the present disclosure, when the present invention IP pings an Anycast IP, it will generate a ping to all the Anycast nodes from where Anycast prefix has been learned. Ping primary Anycast router via primary next hop (Existing). Ping Secondary Anycast router via secondary next hop (Existing). Procedure to Ping all other Anycast routers except primary in IGP-SR domain: ICMP Packet is sent destined to Anycast IP address, ICMP packet is encapsulated within node SID of destination Anycast router, ICMP packet is forwarded using MPLS label switching, ICMP packet encapsulated within Destination Anycast router’s Node-SID will reach the destination Anycast router, Destination Anycast router will send the reply to the source using source ip in ICMP packet.

Figure 3: Scenario

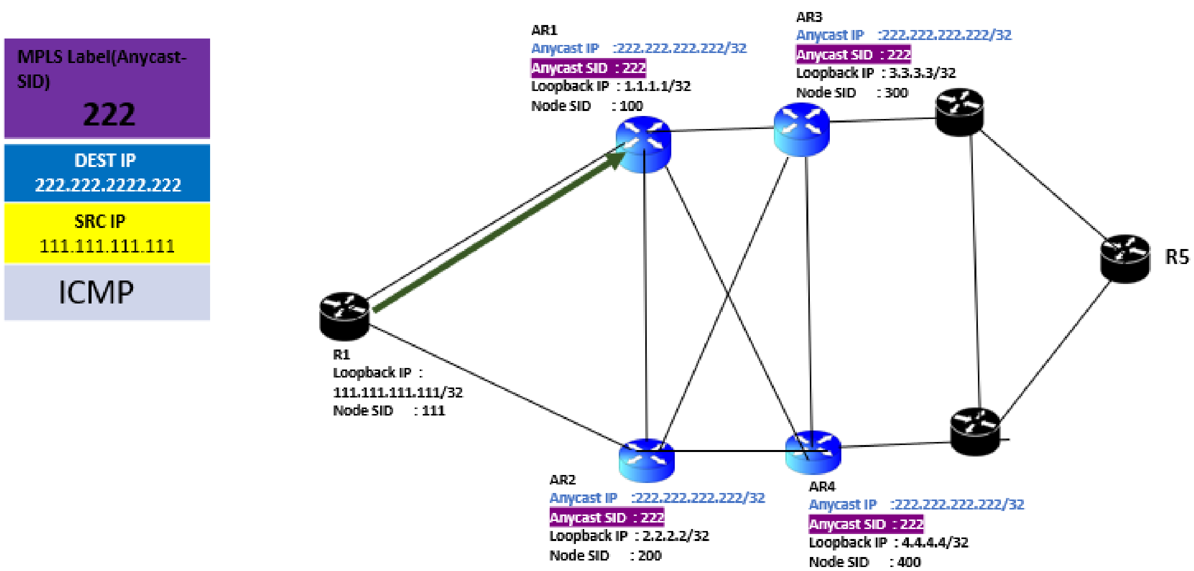
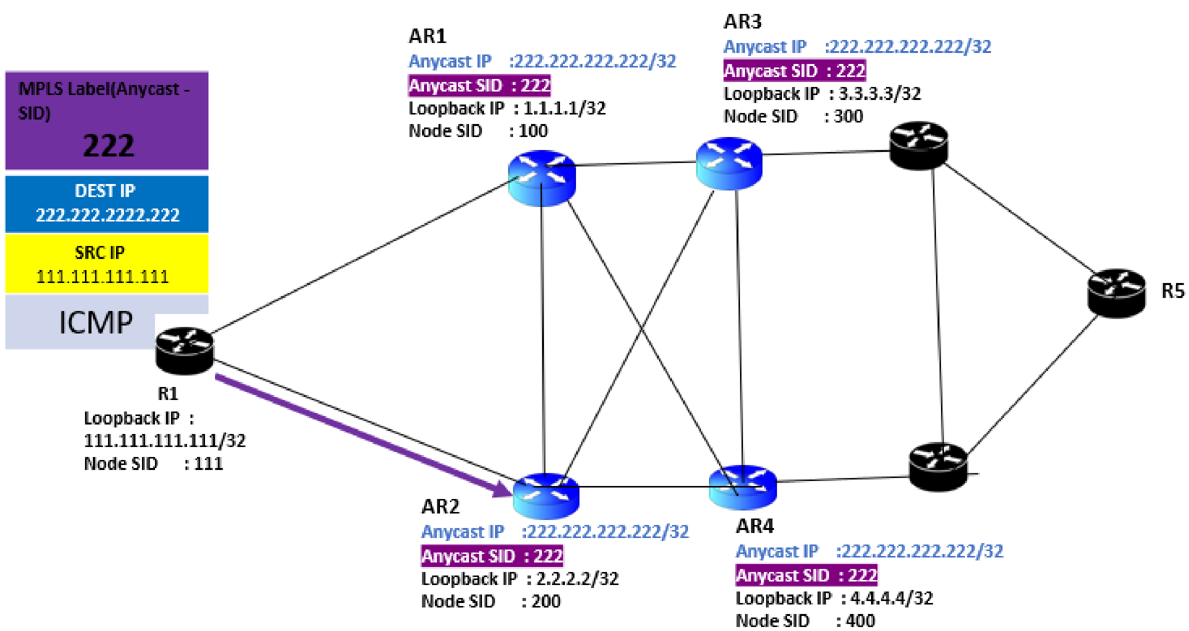


Figure 4: Scenario continued



Considering the scenario in figure 3 and figure 4, AR1,AR2,AR3 and AR4 are Anycast routers configured with same loopback ip 222.222.222.222/32 and Sid 222. For R1, AR1 is primary and AR2 is backup Anycast Router to reach the Network 222.222.222.222/32.

From R1 we can check the IP/MPLS connectivity of AR1 and AR2 by pinging the anycast ip 222.222.222.222/32 on primary and backup anycast routers.

Figure 5: Control plane

MPLS header	IP header	ICMP header	ICMP Payload

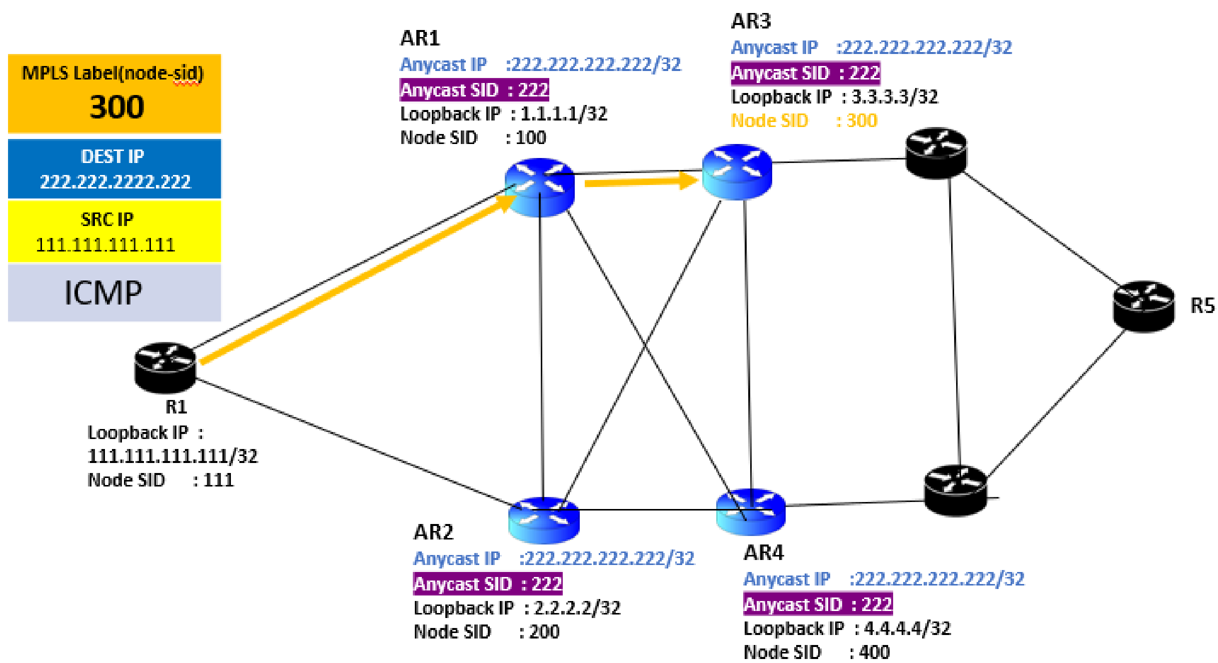
A complete request packet will be constructed in a control plane as shown in figure 5. Information of anycast routers can be derived from routing database of the IGP. An implementation can also maintain Anycast-DB in control plane. Anycast-DB can be used to get the information about all the anycast routers i.e., node-sid, Anycast IP, advertising router-id etc. as shown in figure 6.

Figure 6: Anycast ping packet format

MPLS Label(Node-SID)	DEST IP(Anycast IP)	SRC IP(Loopback IP)	ICMP header	ICMP Payload

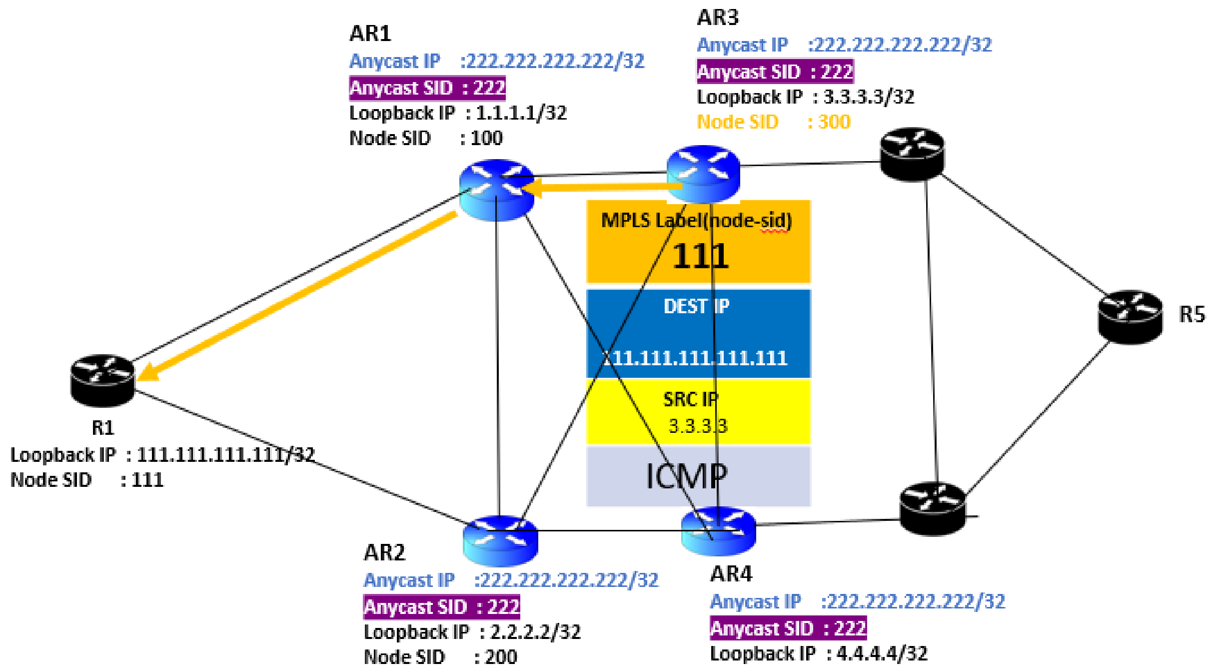
This packet will be injected into the data plane and the data plane will forward the packet as per MPLS Label(node-SID).

Figure 7: ICMP request packet construction and flow



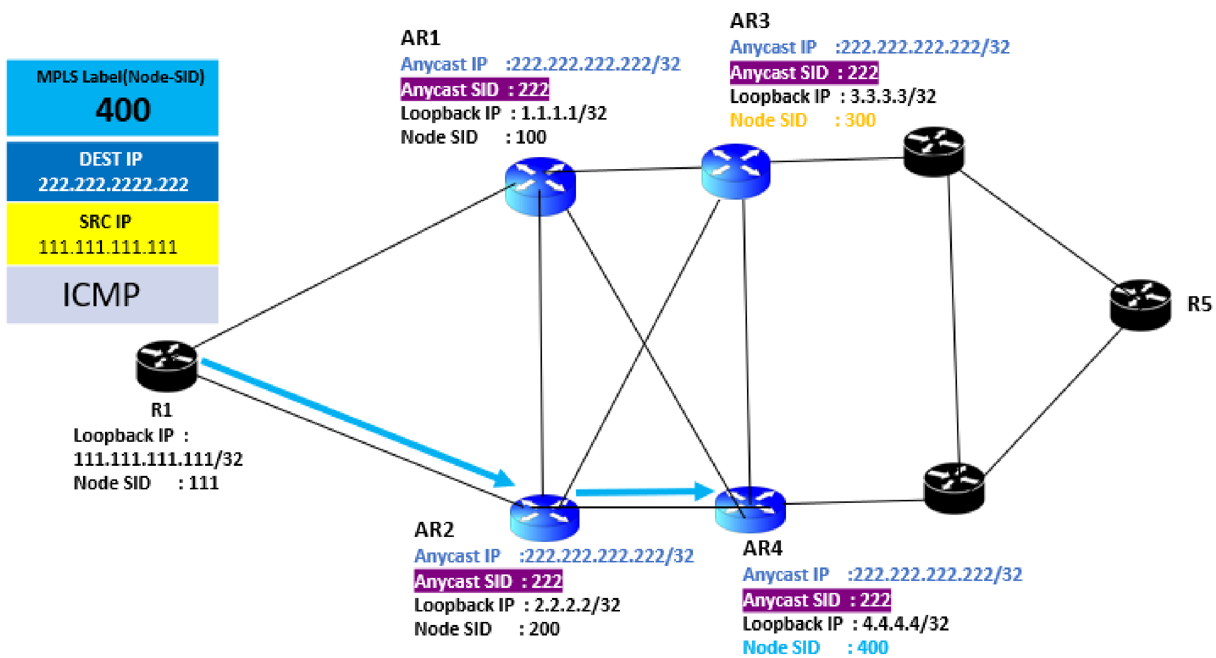
If it is desired to ensure connectivity to Anycast prefix on AR3 from AR1 the flow in figure 7 is followed. The present invention will ensure the AR3 node and Anycast IP reachability in two ways. The first, ICMP Request Packet should be destined to 222.222.222.222/32. And the second, ICMP Request packet should be encapsulated within node SID of AR3.

Figure 8: ICMP request packet construction and flow continued



Shown in figure 8, ICMP reply packet will be destined to src ip(111.111.111.111) in ICMP request packet. No special handling is required in construction of ICMP reply. Therefore, ICMP reply will be sent using the standard way. ICMP Reply Packet should be destined to 111.111.111.111/32. ICMP Request packet should be encapsulated within node SID i.e., 111 of R1.

Figure 9: Connectivity to Anycast prefix on AR4 from R1



Shown in figure 9, if it is desired to ensure connectivity to Anycast prefix on AR4 from R1, the AR4 node and Anycast IP will have reachability in 2 ways. The first, ICMP Packet should be destined to 222.222.222.222/32. The second, ICMP packet should be encapsulated within node SID of AR4.

As anycast ping will internally ping every anycast router in the IGP-SR domain, it has the below Advantages:

- It will determine the end-to-end fiber connectivity to every Anycast router.
- It will determine that all the anycast routers are in good state.
- It will determine that all the anycast router's DP is programmed correctly for Anycast route.
- If same ip is configured on 2 routers mistakenly, it can be determined with Anycast ping.

Additionally, Anycast ping is a good debugging utility to troubleshoot Anycast related issues in the field.

It will be appreciated that some embodiments described herein may include one or more generic or specialized processors ("one or more processors") such as microprocessors, digital signal processors, customized processors, and Field-Programmable Gate Arrays (FPGAs) and unique stored program instructions (including both software and firmware) that control the one or more processors to implement, in conjunction with certain non-processor circuits, some, most, or all of the functions of the methods and/or systems described herein. Alternatively, some or all functions may be implemented by a state machine that has no stored program instructions, or in one or more Application-Specific Integrated Circuits (ASICs), in which each function or some combinations of certain of the functions are implemented as custom logic. Of course, a combination of the aforementioned approaches may be used. Moreover, some embodiments may be implemented as a non-transitory computer-readable storage medium having computer-readable code stored thereon for programming a computer, server, appliance, device, etc. each of which may include a processor to perform methods as described and claimed herein. Examples of such computer-readable storage mediums include, but are not limited to, a hard disk, an optical storage device, a magnetic storage device, a ROM (Read Only Memory), a PROM (Programmable Read-Only Memory), an EPROM (Erasable Programmable Read-Only Memory), an EEPROM (Electrically Erasable Programmable Read-Only Memory), Flash memory, and the like. When stored in the non-transitory computer-readable medium, the software can include instructions executable by a processor that, in response to such execution, cause a processor or any other circuitry to perform a set of operations, steps, methods, processes, algorithms, etc.

Although the present disclosure has been illustrated and described herein with reference to preferred embodiments and specific examples thereof, it will be readily apparent to those of ordinary skill in the art that other embodiments and examples may perform similar functions and/or achieve like results. All such equivalent embodiments and examples are within the spirit and scope of the present disclosure.