Network Working Group
Request for Comments: 3882
Category: Informational

D. Turk
Bell Canada

September 2004

Configuring BGP to Block Denial-of-Service Attacks

Status of this Memo

Copyright Notice

Abstract

This document describes an operational technique that uses BGP communities to remotely trigger black-holing of a particular destination network to block denial-of-service attacks. Black-holing can be applied on a selection of routers rather than all BGP-speaking routers in the network. The document also describes a sinkhole tunnel technique using BGP communities and tunnels to pull traffic into a sinkhole router for analysis.

Table of Contents

1. Existing BGP-Triggered Black-holing Techniques

Current BGP-triggered black-holing techniques rely on altering the BGP next hop address of a network targeted by an attack throughout the iBGP network. A customized iBGP advertisement is generated from a router participating in the destination/attacked AS where the next hop address for the targeted network or host is modified to point to an RFC 1918 [RFC1918] (private internet) address. Most routers on the Internet, especially edge routers, have static routes pointing RFC 1918 addresses to the null interface. Those static routes drive all traffic destined to the network under attack to the null interface.

When an iBGP-speaking router inside the destination AS receives the iBGP update, the advertised prefix will be added to the routing table with a next hop of one of the networks listed in RFC 1918. The router will then attempt to resolve the RFC 1918 next-hop in order to qualify the route and derive a forwarding interface. This process will return a valid next hop as the null interface. Assuming the router is properly configured to direct RFC 1918 destined traffic to a null interface, traffic destined to the attacked network gets dropped, making the attacked network unreachable to the attacker and everyone else.

While this technique shields the internal infrastructure from the attack, protecting a large number of devices, it has the undesirable side effect of rendering the targeted/attacked network unreachable throughout the entire destination AS. Even if a static route pointing an RFC 1918 address to a null interface is not configured on all routers within the destination AS, the modified next hop makes the traffic un-routable to its legitimate destination.

Network operators usually use the BGP-triggered black holes for a short period of time. The technique causes traffic drops on all ingress points of the AS for traffic destined to the attacked network. By default, routers dropping traffic into a null interface should send an "ICMP unreachable" message to the source address belonging to the origin/attacking AS.

Once the procedure reaches this point, one of the source addresses of the attack traffic is hijacked by introducing a device with the same source IP address into the BGP domain of the destination/attacked AS. The device hijacking the source address collects the ICMP unreachable packets. The source addresses of these ICMP unreachable packets reveal which edge routers within the destination/attacked AS the attack is coming from. The network operator may then opt to manually stop the traffic on the routers from which attack traffic is entering.

2. Enhanced BGP-Triggered Black-holing Technique

This paper describes a technique developed to instruct a selected set of routers to alter the next hop address of a particular prefix by use of the BGP protocol. The next hop can either be a null interface or, as discussed later on in this paper, a sinkhole tunnel interface. This technique does not invoke an access list or rate limiting statement to treat attack traffic, nor does it involve a network wide change of the attacked prefix next hop address. The next hop will only be changed on a selection of routers with the aid of BGP communities within the destination/attacked AS.

To prepare the network for this technique, the network operator needs to define a unique community value for each destination AS border router that could potentially drive attack traffic to the victim. For example, a network with a BGP autonomous system number 65001 has two border routers (R1 and R2). Community value 65001:1 is assigned to identify R1, community value 65001:2 is assigned to identify R2, and community value 65001:666 is assigned to identify both R1 and R2.

After the BGP community assignment, R1 and R2 must be configured with the following:

1. Static route pointing an RFC 1918 network to a null interface.

2. AS-Path access list that matches local BGP prefix advertisement.

3. BGP community access list to match the community value assigned by the network operator for the particular router (i.e., 65001:1 for R1).

4. BGP community access list to match the community value assigned by the network operator for all routers (i.e., 65001:666 for R1 and R2)

5. Under the BGP process, an iBGP import route policy should be applied on received iBGP advertisements to do the following logic. (Statements are in a logical AND order)

   a. A policy statement to permit routes that match the following criteria and apply the following changes.

      i. Match for a community specific to that router (i.e., 65001:1, for R1).

      ii. Match the AS-Path to locally generated BGP advertisements.

      iii. Set the BGP next hop to an RFC 1918 network.

iv. Overwrite the BGP community with the well-known community (no-advertise).

b. A policy statement to permit routes that match the following criteria and apply the following changes.

i. Match for a community that covers all routers (i.e., 65001:666).

ii. Match the AS-Path to locally generated BGP advertisements.

iii. Set the BGP next hop to an RFC 1918 network.

iv. Overwrite the BGP community with the well-known community (no-advertise).

After the policies have been configured on R1 and R2, the network operator can, in the case of an attack, advertise the targeted network that could be one or more /32 "host" routes into iBGP of the destination/attacked AS. The advertisement must contain the community value associated with the router(s) where the attack is arriving in addition to the well-known community (no-export). Using BGP communities preserves the original next hop address of the targeted network on all routers where the special route policy configuration is not present. iBGP will then carry the prefix advertisement to all routers in the destination/attacked AS. All routers within the destination AS, except the ones that match the community stamped on the prefix, will be oblivious to the community value and will install the network route with the legitimate next hop address. Routers that match the community will also install the network route into their routing table but will alter the next hop address to an RFC 1918 network and then to a null interface as per the route policies configuration and recursive route lookup. The reason for matching locally announced networks is to make sure that no eBGP peer can misuse this community to drive any network to a null interface. Blackholing the targeted/attacked hosts is recommended, but not the entire address block they belong to so that the blackhole effect has the minimum impact on the attacked network.

This technique stops traffic from getting forwarded to the legitimate destination on routers identified as transit routers for attack traffic and that have route map matches for the community value associated with the network advertisement. All other traffic on the network will still get forwarded to the legitimate destination thus minimizing the impact on the targeted network.

3. Sinkhole Tunnels

Following the "Enhanced BGP-Triggered Black-holing Technique", it may become a requirement to take a look at the attack traffic for further analysis. This requirement adds to the complexity of the exercise. Usually with broadcast interfaces, network operators install network sniffers on a spanned port of a switch for analysis of traffic. Another method would be to announce a network prefix that covers the attack host address into iBGP, altering the next hop into a sinkhole device that can log traffic for analysis. The latter technique results in taking down the services offered on the targeted/attacked IP addresses. Inter-AS traffic will be sucked into the sinkhole, along with Intra-AS traffic. Packet level analysis involves redirecting traffic away from the destination host to a sniffer or a router. As a result, if the traffic being examined includes legitimate traffic, that legitimate traffic will never make it to the destination host. This will result in denial of service for the legitimate traffic.

A better alternative would be to use a sinkhole tunnel. A sinkhole tunnel is implemented at all possible entry points from which attacks can pass into the destination/attacked AS. Using the BGP community technique, traffic destined to the attacked/targeted host could be re-routed to a special path (tunnel) where a sniffer could capture the traffic for analysis. After being analyzed, traffic will exit the tunnel and be routed normally to the destination host. In other words, the traffic will pass through the network to a sniffer without altering the next hop information of the destination network. All routers within the destination/attacked AS iBGP domain will have the proper next hop address. Only the entry point router will have the altered next hop information.

To detail the procedure, a sinkhole router with an optional sniffer attached to its interface is installed and configured to participate in the IGP and iBGP of the attacked AS. Next, a tunnel is created, using MPLS Traffic Engineering as an example, from all border routers attacks can potentially enter from (Inter-AS traffic) to the sinkhole router. When a host or network is under attack, a customized iBGP advertisement is sent to announce the network address of the attacked host(s) with the proper next hop that insures traffic will reach those hosts or networks. The customized advertisement will also have a community string value that matches the set of border routers the attack is entering from, as described in section 2. The new next hop address configured within the route policy section of all border routers should be the sinkhole IP address. This IP address belongs to the /30 subnet assigned to the tunnel connecting the border router to the sinkhole router.

Routers that do not have a match for the community string will do regular routing. Lack of a community string match on these routers will insure that the special route policy does not change the next hop address. Traffic entering from border routers that do not have a match to the special community will pass through regular router interfaces to the legitimate destination. It might also be required to allow the traffic to reach its destination after being captured. In this case, a default network route is configured to point to any interface attached and configured on the iBGP network. This would also include the same physical interface the tunnel is built on. Since the next hop address is not changed on the sinkhole device, traffic entering this device from the tunnel will be sent back to the network due to the presence of the default route. Routing protocols will then take care of properly routing the traffic to its original destination (attacked network).

It becomes apparent that this technique can also be used for purposes other than analyzing attack traffic. Legitimate traffic could also be pulled out of normal routing into a tunnel and then reinserted into the backbone without altering the next hop addressing scheme throughout the iBGP network.

MPLS Traffic Engineering with its many features, is a good method of sliding traffic to the sinkhole device. Features like QoS policies can be applied on the attack traffic, thus preventing it from competing on resources with legitimate traffic.

To be able to alter the next hop on the border router, a subnet of an RFC 1918 network is statically routed to the tunnel interface. An example of the static route is:

```
ip route 192.168.0.12 255.255.255.255 Tunnel0
```

Setting the next hop of the target IP address to 192.168.0.12/32 will force the traffic to go through the tunnel.

Traffic is received at the sinkhole interface via the TE tunnel. Subsequently, three methods could be installed, namely rate-limiting policies, QoS policies, and access lists. These policies could rate limit or drop traffic classified as attack traffic. This process would be completed on the interface of the sinkhole device. Another useful application for a sinkhole router is to pull in traffic via tunnels to an inbound interface and have a default route statement forwarding the traffic out to an Ethernet interface. The Ethernet interface is connected to the iBGP network and guarantees proper delivery of traffic however, it still allows the use of a packet sniffer to further analyze the attack traffic.

This becomes very useful when it is not feasible to apply an Access list or a rate limiting statement on the BGP border router or last hop router before the attacked host or network because of hardware or software limitations. Hence, instead of upgrading interfaces at the point of entry of attack traffic, the latter could be pulled into the sinkhole and treated on that device. Operational costs can be rendered minimal if the sinkhole router is a powerful device.

4. Security Considerations

It is very important to practice tight control over eBGP peering points before implementing the techniques described in this paper. eBGP customers might be able to blackhole a particular subnet using the Blackhole communities. To eliminate the risk, the match for locally generated BGP advertisements in the special route policy should not be neglected.

5. Acknowledgments

The author of this document would like to acknowledge the developers of the remotely triggered black-holing technique and the developers of the backscatter technique for collecting backscatter traffic. The author would also like to thank all members of the IP Engineering department for their help in verifying the functionality of this technique.

6. Informative References

[RFC1918] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

7. Author's Addresses

Doughan Turk
Bell Canada
100 Wynford Drive

EMail: doughan.turk@bell.ca

8. Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.