BGP Security Vulnerabilities Analysis

Research Paper by Andreas Lorisch
28.07.2004
Fachbereich Informatik
GSO-FH-Nürnberg

Internet Draft by Sandra Murphy
draft-ietf-idr-bgp-vuln-00.txt
NAI Labs, June 2003

Agenda

1.Introduction
2.Background
3.Attacks
4.Vulnerabilities and Risks
5.Conclusion

Since BGP, as most other infrastructure protocols, was designed before the Internet became dangerous, it was created with little consideration for protection of the information it carries, or any other security measurements. There are no mechanisms internal to the BGP protocol to protect against attacks that modify, delete, forge, or replay data, any of which has the potential to disrupt overall network routing behavior. That this lack of protection is critical to our day to day internet usage was recently proofed by an article on "heise online", (http://www.heise.de/newsticker/meldung/48319) posted on 6th of June 2004.

1. Introduction

The Border Gateway Protocol, is the most important inter-domain routing protocol. The Internet is comprised of thousands of Autonomous Systems (AS), where each AS is a set of routers under a single technical administration and has its own routing policies. Most autonomous systems exchange routing information through the Border Gateway Protocol (BGP). Therefore as a widely deployed inter-domain routing protocol, BGP plays a critical role in the operation of the Internet.
The current implementation of BGP, the BGP-4 protocol [1], has a variety of vulnerabilities and intrinsic weaknesses. First of all, BGP packets are transmitted over TCP/IP without any encryption and authentication mechanisms. Communications between BGP peers are subject to active and passive wiretapping attacks. Transmission of fictitious BGP messages, modification or replay of valid messages could occur during the routing information exchange process. A false BGP route may cause deny of service to a destination or redirect that destination's traffic to another insecure location. Secure-BGP (S-BGP) [2] was proposed to enhance the security of BGP by verifying the authenticity and authorization of the BGP control traffic. However, the processing and bandwidth overhead coupled with the storage requirement poses nontrivial challenges to the adoption of SBGP to the Internet. Second, as a path vector protocol, BGP limits the distribution of a router's reachability information to its neighbor routers. A large number of temporary routing table fluctuations may be generated in response to a single link failure, change in AS topology or change in routing policy. Slow routing convergence can cause the delay or drop of data packets, and thus degrade the efficiency and reliability of the Internet infrastructure.

2. Background

BGP is the de facto inter-autonomous system routing protocol. An AS uses an interior gateway protocol and common metrics to route packets within itself. At the boundary of each autonomous system, peer border routers exchange network reachability information with other autonomous systems through BGP. BGP uses TCP as its transport protocol. Two BGP speakers form a transport protocol connection between one another. They exchange messages to open and confirm the connection parameters. When connection is first established, a BGP speaker sends its entire routing table to the peer. During the following BGP session, the incremental updates are sent as the routing table changes. Newly acquired routes are stored in the Routing Information Base (RIB) in BGP speakers.
Two types of BGP messages are important to BGP operation. First is KeepAlive message, which is sent periodically to ensure the aliveness of the connection. If the peers can't receive KeepAlive messages in a preset period of time, the BGP connection has to be closed. Physical connectivity failure (link failure, router crash), transient connectivity problems due to congestion, or even manual reboots, may result in the delay of KeepAlive message to the peers. Sequentially, when BGP sessions restart, the peers have to send the full routing table again.
Second is the Update message. Update messages are used to exchange routing information change between two peers. Usually it has two forms: *announcement* and *withdrawal*. A*nnouncement* messages carry on the prefix (destination) and ASPATH. Upon receipt of a new *Announcement*, each router evaluates the path vector and use local decision algorithms to

select the best route among all of the backup routes to that prefix. If the router is a transit router, it will append its unique AS number to the ASPATH and propagate to the downstream BGP speakers. Route *withdrawals* are sent when a router makes a new local decision that a network is no longer reachable.
In an optimal, stable wide-area network, routers should only generate routing updates for relatively infrequent policy changes and the addition of new physical networks.
Frequent BGP updates indicate the network routing instability. Routing instability has a number of possible origins, including router failures, network congestion, software implementation and configuration errors. After one or more of these problems affects the availability of a path to a set of prefix, the routers topologically close to the failure will detect the fault, withdraw the route and make a new local decision to find a new route, to the set of prefix. These routers will propagate the topology update information to the routers within the same autonomous systems. The boundary routers in the network may also propagate the updated information to the other AS routers. During the updated information propagating in the Internet, all affected routers will suffer a *convergence* problem, normally of short duration. Also, there is another kind of routing instability, so called *divergence*. As mentioned before, BGP is a policy based routing protocol, which allows administrator of an autonomous system to specify complex policies. In [6], it was showed that it is possible for autonomous systems to implement "unsafe" or mutually unsatisfiable policies, which will result in persistent route oscillations.

It is clear that the Internet-infrastructure is vulnerable to attacks through its routing protocols and BGP is no exception. Overall Internet behavior can be influenced by injecting bogus routing information into the BGP distributed routing database, intentionally or not. Attackers could modify, forge or replay BGP packets in order to alter that database. The same methods can also be used to disrupt local and overall network behavior by breaking the distributed communication of information between BGP peers.
As cryptographic authentication of the peer-peer communication is no integral part of the BGP protocol, it is subject to all the TCP/IP attacks, like IP spoofing, session stealing, etc. Therefore a moderately skilled attacker could inject BGP messages into the communication between BGP peers and thereby inject bogus routing information, or break the peer to peer connection, which could also be achieved by simple packet spoofing. Any break in the peer to peer communication has a ripple effect on routing that can be wide spread.
In order to make these kinds of attacks more difficult, current BGP implementations must support the authentication mechanism specified in [5], which doesn't imply, that this mechanism is also used. It is based on a pre-installed shared secret, which does not have the capability of IPSEC [4] to agree on a shared secret dynamically, therefore it is not used as in a default configuration, but should be used intentionally.
The cryptographic protections of [5] still don't protect against legitimate BGP speakers, misconfigured or masqueraded, injecting faulty routing information.
Historically, misconfigured and faulty routers have been responsible for widespread disruptions in the Internet.

Counterfeit routing information, after having spread through a portion of the network, can have many different effects on routing behavior.
For example, *starvation* occurs, if data traffic destined for a special node, is forwarded to a part of the network, which is not able to deliver it.
Some portion of the network could believe that it is *cut* from the rest of the network, or data could be routed along a *looping* path, and therefore would never be delivered.
Performance problems might occur if more data traffic is forwarded through some portion of the network than would otherwise need to carry the traffic (*network congestion*), if the traffic forwarding pattern changes at a raid rate (*churn*), if the resources of a particular router are exhausted by BGP messages themselves (resource exhaustion) or if they become a significant amount of the traffic (*overload*).
Also, misleaded traffic could take a much longer path to its destination (*delay*), or pass a part of the network now being able to inspect data (*eavesdrop*), that would otherwise never have passed that part.
Instability occurs if convergence on a global forwarding state is not achieved, and a *blackhole* means that great amounts of packets are dropped, if the increased traffic exceeds the capabilities of one router.

3. Attacks

The formerly described effects can be caused by different kinds of attacks.
On the first, since the routing data transported in BGP is carried in cleartext *eavesdropping* is a possible attack against routing data confidentiality, which is not a common requirement in the BGP specification.
Secondary, BGP doesn't provide any protection against *replay*, *man-in-the-middle* and *message modification* attacks, tough the latter one would be detected if the length of the TCP payload was altered.
BGP also doesn't provide protection against *message insertion* or *message deletion* attacks, but both are difficult to perform, because the correct guessing of the TCP sequence number or a *session stealing* attack becomes necessary.
While bogus routing data can present a *denial of service* attack on the end systems that are trying to transmit data through the network and on the network infrastructure itself, certain bogus information can represent a *denial of service* on the BGP routing protocol. For example, advertising large numbers of more specific routes (longer prefixes) can cause BGP traffic and router table size to increase, even explode.

The mechanism of [5] will counter the message insertion, deletion, and modification, man-in-the-middle attacks and the denial of service attacks from outsiders, but doesn't protect against eavesdropping attacks. It does not protect against replay attacks, which can only be successful under carefully timed circumstances, so the only protection against replay is provided by the TCP sequence number processing. The mechanism of [5] cannot protect against bogus routing information originated by a legitimate peer.

4. Vulnerabilities and Risks

The risks in BGP arise from three fundamental vulnerabilities.
BGP has no internal mechanism that provides strong protection of the integrity, freshness and peer entity authenticity of the messages in peer-peer BGP communications,
no mechanism has been specified within BGP to validate the authority of an AS to announce network layer reachability information (NLRI),
and no mechanism has been specified to ensure the authenticity of the path attributes announced by an AS.
The first fundamental vulnerability motivated the mandated support of [5] in the BGP specification. When that is employed, message integrity and peer entity authentication is provided. The mechanism of [5] assumes that the MD5 algorithm is secure and that the shared secret is protected and chosen to be difficult to guess.

In order to understand the risks of an attacker modifying different fields in the BGP messages, we should take a quick look at how BGP makes its routing decisions.
BGP could possibly receive multiple advertisements for the same route from multiple sources. BGP selects only one path as the best path. When the path is selected, BGP puts the selected path in the IP routing table and propagates the path to its neighbors. BGP uses the following criteria, in the order presented, to select a path for a destination:

- If the path specifies a next hop that is inaccessible, drop the update.
- Prefer the path with the largest weight.
- If the weights are the same, prefer the path with the largest local preference.
- If the local preferences are the same, prefer the path that was originated by BGP running on this router.
- If no route was originated, prefer the route that has the shortest AS_path.
- If all paths have the same AS_path length, prefer the path with the lowest origin type (where IGP is lower than EGP, and EGP is lower than incomplete).
- If the origin codes are the same, prefer the path with the lowest MED attribute.
- If the paths have the same MED, prefer the external path over the internal path.
- If the paths are still the same, prefer the path through the closest IGP neighbor.
- Prefer the path with the lowest IP address, as specified by the BGP router ID.

There are four different BGP message types - OPEN, KEEPALIVE, NOTIFICATION, and UPDATE.
BGP peers themselves are permitted to break peer-peer connections at any time, and so they cannot be said to be issuing "bogus" OPEN, KEEPALIVE or NOTIFICATION messages. However, BGP peers can disrupt routing by issuing bogus UPDATE messages.
In the following, the vulnerabilities of each type of BGP message is discussed, concentrating on these, having the greatest impact.

Message Header:
Each BGP message starts with a standard header. In all cases, syntactic errors in the message header will cause the BGP speaker to close the connection, release all associated BGP resources, delete all routes learned through that connection and run its decision process to decide on new routes.

Breaking up connections:
If a BGP peer receives an OPEN message in state Connect, Active or Established, this will cause the BGP speaker to bring down the connection, release all associated BGP resources, delete all associated routes, run its decision process and cause the state to return to Idle.
The deletion of routes can cause a cascading effect of routing changes propagating through other peers. Also, optionally, an implementation specific peer oscillation damping may be performed. The peer oscillation damping process can affect how soon the connection can be restarted. This damping is sometimes implemented to keep a faulty router from shutting down and restarting afterwards again and again. This 'flapping' would cause an increase in traffic, because the newly started BGP peer would each time consequently ask his neighbors to update his routing tables.
The same effects are caused if erroneous OPEN messages (e.g., unacceptable Hold state, malformed Optional Parameter, unsupported version, etc.), mal-formed UPDATE messages (Withdrawn Routes Length, Total Attribute Length, or Attribute Length that are improper, missing mandatory well-known attributes, Attribute Flags that conflict with the Attribute Type Codes, syntactic errors in the ORIGIN, NEXT_HOP or AS_PATH, etc.) or if a NOTIFICATION message in any state are received.
Consequently, the ability to spoof these messages can lead to a severe disruption of routing over a wide area.

OPEN:
In state OpenSent, the arrival of an OPEN message will cause the BGP speaker to transition to the OpenConfirm state. The later arrival of the legitimate peer's OPEN message will lead the BGP speaker to declare a connection collision. The collision detection procedure may cause the legitimate connection to be dropped, causing a widespread effect.

KEEPALIVE:
Receipt of a KEEPALIVE message when the peering connection is in the Connect, Active, and OpenSent states would cause the BGP speaker to transition to the Idle state and fail to establish a connection. The ability to spoof this message is a vulnerability. To exploit this vulnerability deliberately, the KEEPALIVE must be carefully timed in the sequence of messages exchanged between the peers; otherwise, it causes no damage.

UPDATE:
The Update message carries the routing information, therefore the ability to spoof any part of this message has the greatest effect on the routing tables.

Withdrawn Routes field:
An outsider could cause the elimination of existing legitimate routes by forging or modifying this field, but since the receiving BGP speaker will only withdraw routes associated with the sending BGP speaker, there is no opportunity for a BGP speaker to withdraw another BGP speaker's routes. An outsider could also cause the elimination of reestablished routes by replaying this withdrawal information from earlier packets.

Path Attributes field:
The path attributes (including *Attribute Flags, Attribute Type Codes, Attribute Length; Origin; As_path; Next_hop and Atomic_aggregate*) present many different vulnerabilities and risks.

Attribute Flags, Attribute Type Codes, Attribute Length:
A BGP peer or an outsider could modify the attribute length or attribute type (flags and type codes) so they did not reflect the attribute values that followed. If the flags were modified, the flags and type code could become incompatible (i.e., a mandatory attribute marked as partial), or a optional attribute could be interpreted as a mandatory attribute or vice versa. If the type code were modified, the attribute value could be interpreted as if it were the data type and value of a different attribute.
The most likely result from modifying the attribute length, flags, or type code would be a parse error of the UPDATE message. A parse error would cause the transmission of a NOTIFICATION message and the close of the connection. As a true BGP speaker is always able to close a connection at any time, this vulnerability represents an additional risk only when the source is an outsider, i.e., it presents no additional risk from a BGP peer.

Origin:
This field indicates whether the information was learned from IGP or EGP information. This field is used in making routing decisions, so there is some small vulnerability in being able to affect the receiving BGP speaker's routing decision by modifying this field.

As_path:
A BGP peer or outsider could announce an AS_PATH that was not accurate for the associated NLRI. As it is legitimate for a BGP peer not to verify that a received AS_PATH begins with the AS number of its peer, a malicious BGP peer could announce a path that begins with the AS of any BGP speaker with little impact on itself. This could affect the receiving BGP speaker's decision procedure and choice of installed route. The malicious peer could considerably shorten the AS_PATH, which will increase that route's chances of being chosen, possibly giving the malicious peer access to traffic it would otherwise not receive. The shortened AS_PATH also could result in routing loops, as it does not contain the information needed to prevent loops.
Coupled with the ability to use any value for the NEXT_HOP, this gives a malicious BGP speaker considerable control over the path traffic will take.

Originating Routes:
A special case of announcing a false AS_PATH occurs when the AS_PATH advertises a direct connection to a specific network address. A BGP peer or outsider could disrupt routing to the network(s) listed in the NLRI field by falsely advertising a direct connection to the network. The NLRI would become unreachable to the portion of the network that accepted this false route, unless the ultimate AS on the AS_PATH undertook to tunnel the packets it was forwarded for this NLRI on toward their true destination AS by a valid path. But even when the packets are tunneled to the correct destination AS, the route followed may not be optimal or may not follow the intended policy. Additionally, routing for other networks in the Internet could be affected if the false advertisement fragmented an aggregated address block, forcing the routers to handle (issue UPDATES, store, manage) the multiple fragments rather than the single aggregate. False originations for multiple addresses can result in routers and transit networks along the announced route to become flooded with mis-directed traffic.

Next_hop:
The NEXT_HOP attribute defines the IP address of the border router that should be used as the next hop when forwarding the NLRI listed in the UPDATE message. If the recipient is an external peer, then the recipient and the NEXT_HOP address must share a subnet. It is clear that an outsider modifying this field could disrupt the forwarding of traffic between the two AS's.
In the case that the recipient of the message is an external peer of an AS and the route was learned from another peer AS (this is one of two forms of "third party" NEXT_HOP), then the BGP speaker advertising the route has the opportunity to direct the recipient to forward traffic to a BGP speaker at the NEXT_HOP address. This affords the opportunity to direct traffic at a router that may not be able to continue forwarding the traffic. A malicious BGP speaker can also use this technique to force another AS to carry traffic it would otherwise not have to carry. In some cases, this could be to the malicious BGP speaker's benefit, as it could cause traffic to be carried long-haul by the victim AS to some other peering point it shared with the victim.

Local_pref:
The LOCAL_PREF attribute must be included in all messages with internal peers and excluded from messages with external peers. Consequently, modification of the LOCAL_PREF could effect the routing process within the AS only. Note that there is no requirement in the BGP RFC that the LOCAL_PREF be consistent among the internal BGP speakers of an AS. As BGP peers are free to choose the LOCAL_PREF as they wish, modification of this field is a vulnerability with respect to outsiders only.

Atomic_aggregate:
The ATOMIC_AGGREGATE field indicates that an AS somewhere along the way has aggregated several routes and advertised the aggregate NLRI without the AS_SET formed as usual from the AS's in the aggregated routes' AS_PATHs. BGP speakers receiving a route with ATOMIC_AGGREGATE are restricted from making the NLRI any more specific. Removing the ATOMIC_AGGREGATE attribute would remove the restriction, possibly causing traffic intended for the more specific NLRI to be routed incorrectly. Adding the ATOMIC_AGGREGATE attribute when no aggregation was done would have little effect, beyond restricting the un-aggregated NLRI from being made more specific. This vulnerability exists whether the source is a BGP peer or an outsider.

NLRI:
By modifying or forging this field, either an outsider or BGP peer source could cause disruption of routing to the announced network, overwhelm a router along the announced route, cause data loss when the announced route will not forward traffic to the announced network, route traffic by a sub-optimal route, etc.

TCP messages:
Since BGP runs over TCP, listening on port 179, BGP is subject to attacks on TCP. This includes all the widely known TCP attacks like TCP SYN, TCP SYN ACK, TCP ACK, TCP RST/FIN/FIN-ACK. Some of these attacks are countered through the use of [5], most of them are difficult to perform, because the correct sequence number has to be predicted.
If still successful, an attacker could break the legitimate connection, or in the worst case take over the connection.

DoS and DDos:
Because the packet directed to TCP port 179 are passed to the BGP process, that potentially resides on a slower processor in the router, flooding a router with TCP port 179 packets is a good way to perform DoS attacks against the router. No BGP protocol mechanism can defeat such attacks; other mechanisms must be employed.

5. Conclusion

The use of the mandatory-to-support mechanisms of [5] counter the message insertion, deletion, and modification attacks and man-in-the-middle attacks from outsiders. If routing data confidentiality were desired, the use of IPSEC ESP could provide that service.

As cryptographic-based mechanisms, both [5] and IPSEC assume that the cryptographic algorithms are secure, that secrets used are protected from exposure and are chosen well so as not to be guessable, that the platforms are securely managed and operated to prevent break-ins, etc.

These mechanisms do not prevent attacks that arise from a router's legitimate BGP peers. There are several possible solutions to prevent a BGP speaker from inserting bogus information in its advertisements to its peers, i.e., from mounting an attack on a network's origination or AS-PATH.

(1) Origination Protection: sign the originating AS.

(2) Origination and Adjacency Protection: sign the originating AS and predecessor information ([7])

(3) Origination and Route Protection: sign the originating AS, and nest signatures of AS_PATHs to the number of consecutive bad routers you want to prevent from causing damage. ([8])

(4) Filtering: rely on a registry to verify the AS_PATH and NLRI originating AS ([3]).

Filtering is in use near some customer attachment points, but is not effective near the Internet center. The other mechanisms are still controversial and are not yet in common use.

Current Situation:
The primary usage of BGP is as a means to provide reachability information to Autonomous Systems (AS) and to distribute external reachability internally within an AS. BGP is the routing protocol used to distribute global routing information in the Internet. BGP is therefore used by all major Internet Service Providers (ISP) and many smaller providers and other organizations.
The backbone routers of the major ISP's have a route to every reachable IP address. Analysis of BGP UPDATE's recorded in 1999 showed routing databases that contained about 61000 IPv4 address prefixes. Each (nonleaf) BGP speaker maintains a full routing table, and sends its best route for each prefix to each neighbor speaker. When a BGP speaker reboots, it receives complete routing tables (via UPDATE's) from each of its neighbors. The worst case arises at Network Access Points (NAP's), where ISP's are connected. A BGP speaker at a NAP might have about 30 peers.
On a daily basis, a BGP speaker at a NAP receives about 1500 UPDATE's from each peer. This rate is affected somewhat by Internet growth, but is mostly a function of UPDATE's sent due to link, component or congestive errors and recoveries. About 50% of all UPDATE's are sent as a result of route 'flaps' (getting a new route and then returning to the former).

Operational Protections:
The role which BGP plays in the Internet puts BGP implementations in unique conditions and places unique security requirements on BGP. BGP is operated over interprovider interfaces in which traffic levels push the state of the art in specialized packet forwarding hardware and exceed the performance capabilities of hardware implementation of decryption by many orders of magnitude. The capability of an attacker using a single workstation with high speed interface to generate false traffic for denial of service (DoS) far exceeds the capability of software based decryption or appropriately priced cryptographic hardware to detect the false traffic. One means to protect the network elements from DoS attacks under such conditions is to use packet based filtering techniques based on relatively simple inspections of packets. As a result, for an ISP carrying large volumes of traffic, the ability to packet filter on the basis of port numbers is an important protection against DoS attacks, and a necessary adjunct to cryptographic strength in encapsulation.

Current practice in ISP operation is to use certain common filtering techniques to reduce the exposure to attacks from outside the ISP. To protect Internal BGP (IBGP) sessions, filters are applied at all borders to an ISP network which remove all traffic destined for addresses of network elements internal addresses (typically contained within a single prefix) and the BGP port number (179). Packets from within an ISP are not forwarded from an internal interface to the BGP speaker's address on which External BGP (EBGP) sessions are supported, or to a peer's EBGP address if the BGP port number is found. With appropriate consideration in router design, in the event of failure of a BGP peer to provide the equivalent filtering, the risk of compromise can be limited to the peering session on which filtering is not performed by the peer or the interface or line card on which the peering is supported. There is substantial motivation and little effort for ISPs to maintain such filters.

These operational practices can considerably raise the difficulty for an outsider to launch a DoS attack against an ISP. Prevented from injecting sufficient traffic from outside a network to effect a DoS attack, the attacker would have to undertake much more difficult tasks, such as compromise of the ISP network elements or undetected tapping into physical media.

References:

[1] Y. Rekhter and T. Li, “a border Gateway Protocol 4 (BGP-4)”, RFC 1771, March 1995.

[2] S. Kent, C. Lynn, J. Mikkelson and K. Seo, “Secure Border Gateway Protocol (S-BGP) – Real World Performance and Deployment Issues”, California, February 2000.

[3] B. Smith and J.J. Garcia-Luna-Aceves, "Securing the Border Gateway Routing Protocol", Proc. Global Internet'96, London, UK, 20-21 November 1996.

[4] S. Kent and R.Atkinson, "Security Architecture for the Internet Protocol", RFC2401, November 1998.

[5] A. Heffernan, "Protection of BGP Sessions via the TCP MD5 Signature Option", RFC2385, August 1998.

[6] T. Griffin, F. Shepard, and G. Wilfong. “An Analysis of BGP Convergence Properties”, August 1999

[7] C. Villamizar, C. Alaettinoglu, D. Meyer, S. Murphy and C. Orange, "Routing Policy System Security", RFC 2725, December, 1999.

[8] S.Kent, C.Lynn, and K. Seo, "Secure Border Gateway Protocol (Secure-BGP)", IEEE Journal on Selected Areas in Communications, Vol. 18, No. 4, April 2000, pp. 582-592.

S.Murphy, “draft-ietf-idr-bgp-vuln-00.txt”, NAI Labs, June 2003

“Internet Working Technologies Handbook”, Ciscopress.com

Y.Liao, K. Zhang, “BGP Behavior Monitoring and Analysis”, Dep. of Computer Science, UoC