Secure Border Gateway Protocol (S-BGP) — Real World Performance and Deployment Issues

Stephen Kent, Charles Lynn, Joanne Mikkelson, and Karen Seo
BBN Technologies

Abstract

The Border Gateway Protocol (BGP), which is used to distribute routing information between autonomous systems, is an important component of the Internet's routing infrastructure. Secure BGP (S-BGP) addresses critical BGP vulnerabilities by providing a scalable means of verifying the authenticity and authorization of BGP control traffic. To facilitate widespread adoption, S-BGP must avoid introducing undue overhead (processing, bandwidth, storage) and must be incrementally deployable, i.e., interoperable with BGP. To provide a proof of concept demonstration, we developed a prototype implementation of S-BGP and deployed it in DARPA's CAIRN testbed. Real Internet BGP traffic was fed to the testbed routers via replay of a recorded BGP peering session with an ISP's BGP router. This document describes the results of these experiments – examining interoperability, the efficacy of the S-BGP countermeasures in securing BGP control traffic, and their impact on BGP performance, and thus evaluating the feasibility of deployment in the Internet.

1. Border Gateway Protocol (BGP)

Internet routing is implemented using a distributed system composed of many routers, grouped into administrative domains called Autonomous Systems (ASes). Routing information is exchanged between ASes using Border Gateway Protocol (BGP) [2,3] UPDATE messages. BGP has a number of vulnerabilities [1,3,5] which can be exploited to cause problems such as misdelivery or non-delivery of user traffic, misuse of network resources, network congestion and packet delays, and violation of local routing policies.

Communication between BGP peers is subject to active and passive wiretapping attacks. BGP and the TCP/IP protocol used by it can be attacked. A BGP speaker can be compromised, e.g., a speaker's BGP-related software, configuration information, or routing databases may be modified or replaced illicitly via unauthorized access to a router, or to a server from which router software is downloaded, or via a spoofed distribution channel, etc. Such attacks could result in transmission of fictitious BGP messages, modification or replay of valid messages, or suppression of valid messages. If cryptographic keying material is used to secure BGP control traffic, that too may be compromised. We have developed security enhancements to BGP that address most of these vulnerabilities by providing a secure, scalable system: Secure-BGP (S-BGP) [1,3]. Better physical, procedural and basic communication security for BGP routers could address some of these attacks. However, such measures would not counter any of the many forms of attacks that compromise routers themselves. Experience with accidental mis-configurations, and the many vulnerabilities of management system components, strongly argue in favor of countering such Byzantine failures if we are to provide adequate protection for the Internet.

The BGP-4 protocol, including descriptions of the UPDATE message and the route propagation algorithm, is described in [2,3]. Briefly, the numbers that identify IP networks are specified by a prefix, which consists of a count of significant bits in an IP address and the value of those bits. An UPDATE consists of three parts: "withdrawals" - a list of prefixes for destinations that are no longer reachable via a previously specified route; "network layer reachability information (NLRI)" - a list of IPv4 address prefixes that are reachable; and "path attributes" - the characteristics of the cumulative path that can be used to reach the NLRI in this UPDATE. These path attributes include reachability information for IPv6 address prefixes in the "Multi-Protocol Reachable NLRI" and "Multi-Protocol Unreachable NLRI" path attributes. The attribute used to specify the path, the AS_PATH attribute, is basically a sequence of the Autonomous Systems (ASes) along the path, each identified by its AS number. When propagating an

Appeared in Proceedings of the Network and Distributed System Security Symposium (NDSS 2000), San Diego, California, February 2000.

UPDATE to a neighboring AS, the BGP speaker prepends its AS number to the sequence, and updates other path attributes as appropriate. Since an UPDATE can specify only one path, only address prefixes that share that path may be combined into the UPDATE.

We maintain that security for BGP should be defined as the correct operation of BGP speakers. This definition is based on the observation that any successful attack against BGP should result in other than correct operation, presumably yielding degraded routing. Correct operation of BGP depends upon the integrity, authenticity, and timeliness of the routing information that BGP distributes (via UPDATEs). It also depends on each BGP speaker's processing, storage, and distribution of this information in accordance with both the BGP specification and the routing policies of the BGP speaker's Autonomous System. The countermeasures being developed address the following points of correct (secure) operation of BGP:

1. Each UPDATE that a BGP speaker receives from a peer was sent by that peer, was not modified en-route from that peer, and contains routing information no less recent than the routing information previously received for the indicated destinations from that peer.

2. The UPDATE was intended for the BGP speaker or AS that received it.

3. The peer sending the UPDATE was authorized to act on behalf of its AS to advertise the routing information in the UPDATE to BGP speakers in the recipient AS.

4. The AS originating the route, i.e., that contains the BGP speaker that originally included the list of reachable destinations within the UPDATE, was authorized to represent those destinations by the organization(s) that owns them.

5. The organization *owning* the IP address space advertised in the UPDATE was allocated that address space through a chain of delegations originating at the ICANN (formerly IANA).

6. If the UPDATE indicates a withdrawn route, i.e., one that is no longer feasible, then the peer withdrawing the route was previously authorized to advertise that route.

7. Finally, both the BGP speaker that sent the UPDATE, and the peer that received the UPDATE, correctly applied BGP processing rules and the (local) routing policy specified by it's AS.

The security measures developed for BGP address the first six of these requirements, even if one or more BGP speakers have been subverted. The last requirement is not addressed by S-BGP, primarily because of the considerable latitude afforded BGP speakers by local routing policies, and because ASes generally do not publish details of the policies. Without knowledge of these local policies, other ASes cannot determine if local routing policies are being correctly applied. Moreover, because UPDATEs do not carry sequence numbers, a BGP speaker is free to generate an UPDATE based on old information, e.g., it may advertise a route that was formerly valid but which has been withdrawn. Thus, in the face of Byzantine failures (vs. active wiretapping attacks against inter-speaker links), the timeliness of UPDATEs is enforced only on a very coarse basis by these countermeasures.

1.1 Deployment of S-BGP in the Internet

In addition to the basic problem of how to secure BGP, there is the problem of how to deploy the solution in the Internet. Deploying S-BGP will require the adoption of this technology by ISPs and by router vendors, plus PKI support by the registries that allocate autonomous system numbers to ISPs and DSPs (down-stream providers), and address prefixes to customers. Because of the distributed management of the Internet infrastructure, the anticipated growth in the size and inter-connectivity of the Internet, the large volume of BGP UPDATE traffic, and resource limitations in the Internet's routers and circuits, it is crucial that S-BGP be scalable and incrementally deployable.

- Scalability – The impact of S-BGP on a router's CPU and storage utilization, and on network bandwidth must be within acceptable limits.

- Deployability – In order to successfully deploy S-BGP, two major issues need to be addressed. First, S-BGP countermeasure information must be forwarded between S-BGP routers in the same AS. Also, since S-BGP introduces a new BGP path attribute, one must provide backward compatibility between S-BGP and BGP-4 so that it is possible to incrementally deploy these countermeasures.

2. S-BGP countermeasures

This section provides a high-level overview of the S-BGP architecture. For further details see [1,3]. Document [1] also contains an appendix that describes a number of alternative approaches that were studied but not chosen, e.g., “path” and “neighbor” attestations (vs. the selected “route” attestation), because of vulnerabilities in these approaches or because of adverse performance implications.

2.1 Design constraints and assumptions

The design of S-BGP was based on several assumptions and constraints. First, the countermeasures must not violate the BGP specifications, e.g., any additional data carried in UPDATEs must make use of defined extension mechanisms and the maximum (BGP) packet size must not be exceeded. Thus, for example, the certificates and CRLs needed to support S-BGP are transmitted out-of-band and we have chosen to carry S-BGP data in-band, as an optional, transitive, path attribute. The countermeasures should be dynamic, responding quickly to topology changes, including the addition of a new AS, network, or router. The performance impact of the countermeasures must be minimal. Finally, the countermeasures should scale, as the Internet continues to grow at a well-documented, substantial pace.

2.2 Security mechanisms

The S-BGP architecture employs three primary security mechanisms: PKIs, attestations, and IPsec.

We make use of a Public Key Infrastructure (PKI), based on the use of X.509 v3 certificates, that supports the authentication of IP address block ownership, AS Number ownership, AS identification, and BGP router identification and authorization to represent an AS. This involves three kinds of certificates. The first type of certificate binds a public key to an organization and to a set of IP address prefixes. These certificates are used either to verify that an originating AS "owns" a specified portion of the IP address space, or that the owner has authorized the AS to advertise the address space. The certificates are arranged into a singly-rooted hierarchy that parallels the existing IP address allocation system. Thus the ICANN is the root, and the next tier generally will consist of registries such as ARIN and RIPE. (Because of historic allocation of addresses, this characterization is a simplification of the actual certification tree. For example, some ISPs and subscriber organizations hold addresses assigned directly by IANA, the predecessor to the ICANN.) The next tier generally consists of ISPs. An additional tier represents DSPs or subscribers, when these entities participate in BGP. Note that only those entities that execute BGP need these certificates. Finally, if an organization owns multiple ranges of addresses, this design calls for assigning a single certificate containing a list of address blocks, so as to minimize the number of certificates needed to validate an UPDATE. (If an organization acquires additional address blocks, a new certificate is issued to reflect the increased scope of ownership.) Figure 1 illustrates this certification tree, showing the organizations that are represented at each tier.

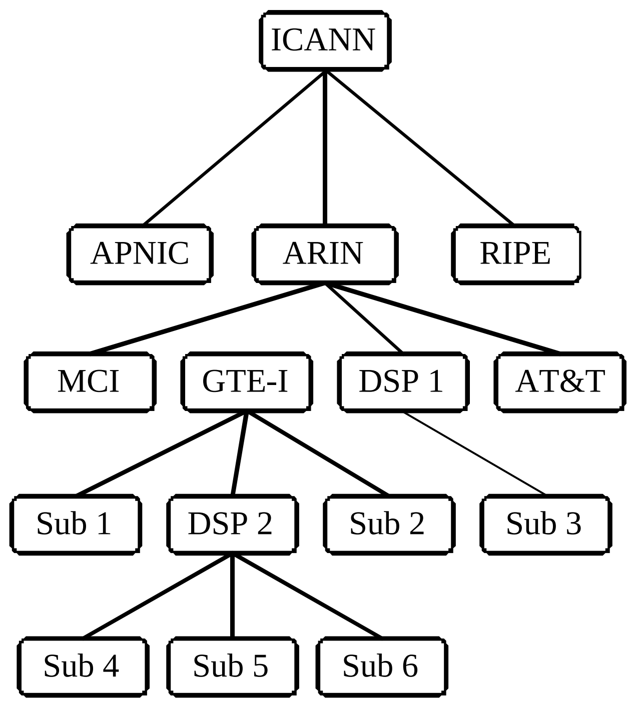

Figure 1: Address Space PKI Structure

The second type of certificate binds a public key to an organization and a set of AS numbers, and the third binds a public key to an AS number and to a BGP router ID. Together, these two types of certificates allow BGP speakers to authenticate one another, and to verify that a given speaker is authorized to represent a specified AS. Here too, the ICANN is the root of the hierarchy, and the second tier consists of registries. (The certificates issued by the ICANN to registries are conventional in format and are employed to permit use of a single root for all classes of certificates.) The third tier consists of

ISPs, DSPs, and subscribers. The second type of certificate is issued at the second tier, and the third type at the third tier. Lower tiers generally represent ASes and routers associated with higher tier organizations. Figure 2 illustrates a simple example of the tree structure for these two types of certificates, noting the organizations involved at the top three tiers, and the ASes and routers that populate the lower tiers.

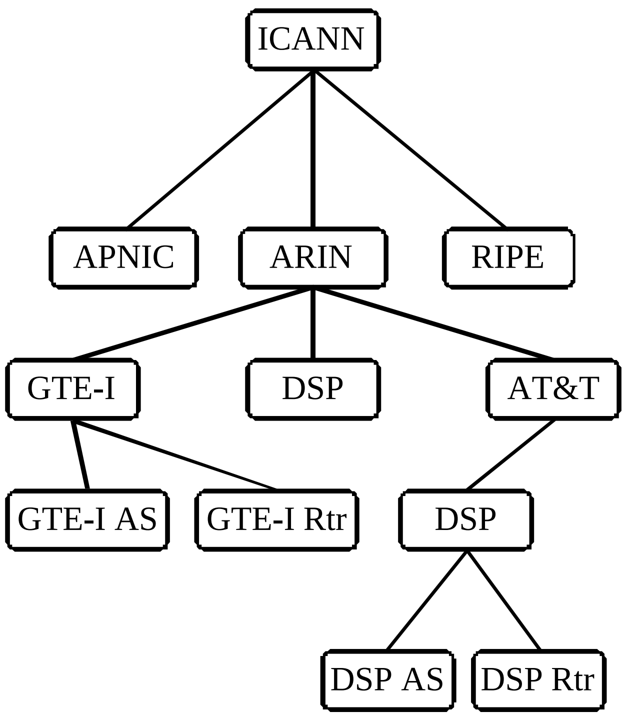

Figure 2: AS Ownership and Router ID PKI

There already exist procedures and personnel to manage the assignment of IP address prefixes and AS numbers. We propose to take advantage of this existing infrastructure to manage these certificates. The PKIs that must be created to support S-BGP will overlay the existing administrative framework, based on the ICANN, regional registries, ISPs, etc.

"Attestations" constitute the second major component of the architecture. They form the heart of S-BGP, and represent a conceptually straightforward means of achieving the critical security guarantees described above. It is the use of attestations, protected by digital signatures, that permits S-BGP to counter Byzantine attacks. Attestations are signed and validated using the keys and certificates from the PKI described above. They enable each BGP speaker that receives a route advertisement to verify that each AS along the route has been authorized by the preceding AS along the path to advertise the route, and that the originating AS has been authorized by the owner of each IP address prefix contained in the UPDATE to advertise these prefixes. The attestations are carried in a new, optional, BGP transitive path attribute that contains digital signatures covering the routing information. Conceptually, there are two types of attestations:

- Route attestations (RAs) — where the issuer is an AS (or a router authorized to represent the AS) and the subject is a transit AS or another AS providing third party advertisements for an AS that is not running BGP.
- Address attestations (AAs) — where the issuer is the organization that owns the address prefixes contained in the attestation and the subject is one or more ASes that are authorized to advertise these prefixes, e.g., the organization's Internet service provider(s).

Figure 3 illustrates the layout of the Attestation Path Attribute in an S-BGP UPDATE, and shows the structure of an RA and how it relates to the BGP UPDATE structure. Within an RA, there is a header, the name of the entity that signed the RA (the issuer), a back pointer to the certificate needed to validate the RA (the certificate ID), and an indication of the algorithm used to sign the RA. Finally, there is the data that is protected by the RA, i.e., covered by the digital signature. This includes a validity interval, the BGP speaker to which the RA applies (the subject), the AS path and NLRI, and any other path attributes that must be protected. Note that there is one RA for each AS in the path.

IPsec [4], specifically the encapsulating security payload (ESP) is used to provide data and partial sequence integrity, and peer entity authentication for BGP control traffic. Because it is implemented at the IP layer, IPsec protects the integrity of the TCP connections used between BGP speakers. Its anti-replay mechanisms detect and reject replayed packets more quickly than TCP, helping to reduce the effect of a denial of service attack. The IPsec anti-replay mechanisms plus TCP sequence numbers ensure the "no less recent" requirement for correct operation of BGP, relative to attacks mounted against inter-router links. If confidentiality of BGP control traffic becomes an issue, it will be easy to later enable the IPsec confidentiality mechanisms where needed, without any changes to BGP.

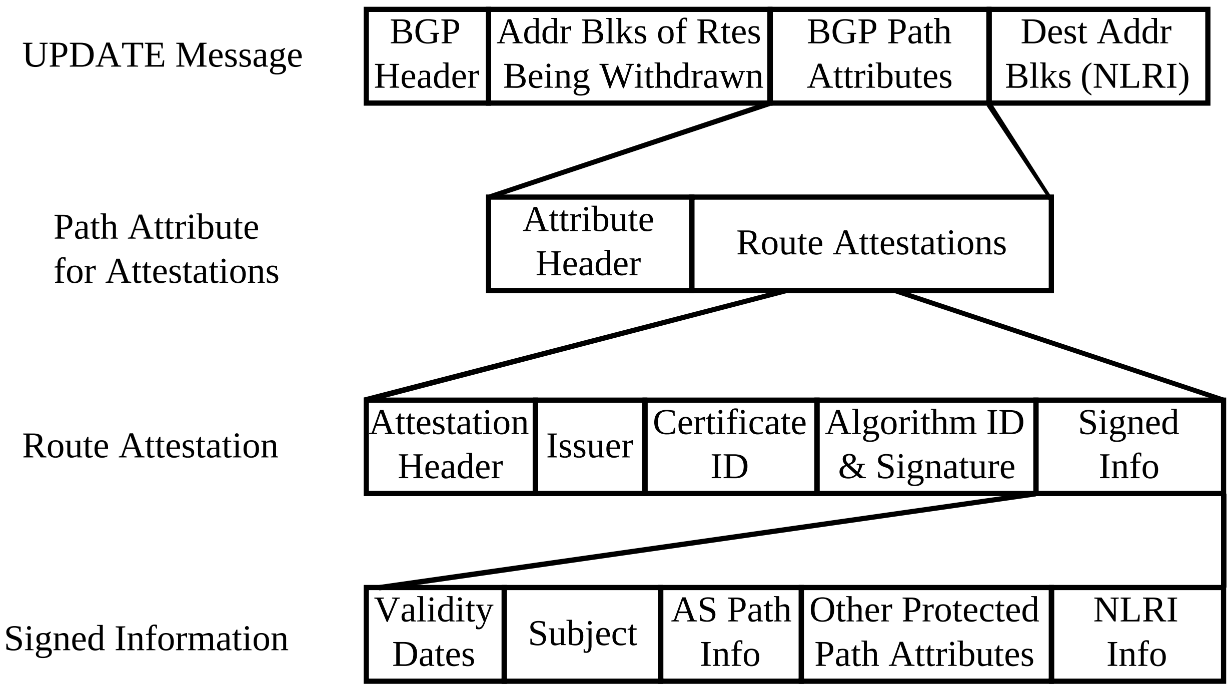

Figure 3: An S-BGP UPDATE Message

2.3 Distribution of S-BGP countermeasures information

BGP UPDATEs are limited in length to 4096 bytes. Certificates are relatively large (approx. 440 bytes using DSA keys). Also, the combined length of the Address Attestations needed to validate the NLRI in an UPDATE that contains many (e.g., hundreds of) address prefixes can be large, relative to the size limit of an UPDATE. Sending certificates and AAs with each UPDATE would often be redundant and wasteful of bandwidth. Since certificates and AAs are relatively stable (e.g., unchanging for months or maybe a year), this architecture uses out-of-band distribution of certificates and AAs to all S-BGP speakers. The S-BGP architecture makes use of two tiers of repositories from which an AS's management (first tier) or a router (second tier) can download the entire certificate and AA database. The first tier consists of several replicated, easy to access storage sites, e.g., the NAP route servers. The second tier of repositories is operated by ISPs/DSPs, to provide local access to certificates and AAs for the S-BGP speakers within their respective ASes.

Route Attestations are distributed with BGP UPDATEs in the newly defined optional, transitive "attestation" path attribute (see Section 2.1). The BGP speaker receiving an UPDATE caches the associated attestations with the route in its routing information database (RIB). These attestations, plus an RA for the local AS, are included in any UPDATEs that are subsequently sent to the speaker's peers.

2.4 Processing of S-BGP countermeasure information

Processing of S-BGP countermeasures data occurs as follows (see [1,3] for more detail):

- Generation and signing of certificates and CRLs are handled by the issuing Certification Authority, e.g., the ICANN, a registry, an ISP/DSP or a subscriber. These certificates and CRLs are posted to directories for retrieval by ISPs and DSPs.
- The validation of certificates and CRLs, and the generation/signing of AAs, is handled by the ISPs/DSPs and subscriber organizations, e.g., by a NOC for an AS. The second tier of the distribution system described in Section 2.2 produces pre-processed (validated, "extracted", etc.) certificates, CRLs and AAs thus saving the routers from having to do this work. Note: The option also exists to

distribute AAs in UPDATEs, if there is space in the UPDATE and if there is a reason not to wait for the standard out-of-band distribution mechanism. In this case, the router must validate the AA.
- The validation of RAs, and their use in validating the routing information in an UPDATE is handled by the S-BGP router that receives the UPDATE. To minimize the load during router initialization, verification of an UPDATE is done only after an UPDATE is selected for use in the routing table. In this fashion, the task of validating the very large number of UPDATEs that may be received during initialization is spread out over time, lessening the performance impact. S-BGP routers also effect the generation and signing of RAs needed for any UPDATEs being sent to peers.
- S-BGP peers implement IPsec in order to secure each TCP/IP session carrying the BGP control traffic.

2.5 Effectiveness

The countermeasures described above satisfy the first six security requirements described in Section 1.1. The following maps countermeasures to requirements, though not in exact correspondence with the list in 1.1:

- Each route attestation is signed by the AS that sends it, so a BGP speaker can verify that each UPDATE it receives from a peer was sent by that peer and was not modified en-route from that peer. The use of IPsec ensures the timeliness of UPDATEs, on a point-to-point basis.

- Because a route attestation specifies the next AS in a path, as part of the signed data, a receiver can detect if an UPDATE is received by other than the AS for which it was intended.

- The first certification hierarchy described above binds one or more address spaces to an organization, thus ensuring that the organization *owning* the IP address space advertised in the UPDATE was allocated that address space through a chain of delegations originating at the ICANN.

- The first route attestation, plus the address attestation, enables a receiver to verify that the BGP speaker that originated the list of reachable destinations within the UPDATE, was authorized to represent those destinations by the organization(s) that owns them.

- The second certification hierarchy described above binds a BGP ID to one or more AS numbers, to ensure that a BGP speaker sending an UPDATE was authorized to act on behalf of its AS to advertise the routing information in the UPDATE.

- The authentication and integrity offered by IPsec, plus simple matching by a receiver, ensures that an UPDATE indicating a withdrawn route will be accepted only from the peer that was previously authorized to advertise that route.

2.6 Deployability

We anticipate using internal BGP (iBGP) sessions to distribute the external routing information among the S-BGP speakers within a single AS. In these scenarios, the S-BGP attestation path attribute information would be forwarded along with the other UPDATE information, without modification.

S-BGP introduces a new attribute, the attestation path attribute. To enable backwards compatibility with BGP-4, this new transitive attribute will be defined as discretionary (optional) for both external and internal BGP exchanges. Interoperatibility between BGP-4 and S-BGP will then be possible because the BGP protocol specifies that implementations should accept unrecognized optional transitive attributes and retain them for propagation to other BGP/S-BGP speakers. This backwards compatibility will remove the need to require that all BGP implementations to support S-BGP from day one and make it possible to incrementally deploy these countermeasures — a necessity given the distributed nature of Internet ownership and operation. In an S-BGP AS that receives UPDATEs which have gone through BGP-4 ASes, it will be up to the local administrators to set the security policy as to how the AS's S-BGP speakers should handle UPDATEs that are missing route attestations (or for which address attestations or certificates are not available).

S-BGP imposes significant demands on routers, primarily in terms of memory requirements, and, to a lesser extent, in terms of CPU utilization. In the near term, this overhead can be ameliorated by co-locating a separate S-BGP speaker with each border router. The S-BGP speaker will perform all the typical BGP functions plus the new S-BGP functions such as signing/verifying attestations; fetching/storing extracted certificates/CRLs and AAs, and verifying routes. The border routers will be responsible only for forwarding data packets and for participating in intra-domain routing. As a local policy, the attestations could be filtered out of the BGP UPDATEs sent to the border router. This would eliminate the requirement for storage to hold these attestations. In the longer term, the additional CPU and memory resources will be more conveniently handled by adding hardware cryptographic support and memory to the router itself. This will eliminate the need for co-located boxes and make network management simpler.

3. Experimental background

3.1 Testbed setting

The results described in this document were collected from experiments run in DARPA's Collaborative Advanced Interagency Research Network (CAIRN) testbed. This testbed is used for research on and development of advanced computer network protocols. It is composed of programmable routers (mostly PCs) and multimedia hosts. The routers are interconnected by OC-3 and DS-3 ATM backbone links (configurable ATM virtual circuits) and T1 (and faster) tail circuits. The CAIRN testbed is connected to the Internet at several points. Experimenters can download software to routers and configure the topology as required. In this case, we downloaded S-BGP software to a number of routers and configured them into several interconnected ASes. We then established BGP peering sessions between these routers and between some of them and BGP routers in the external Internet.

The CAIRN routers were configured with FreeBSD 2.2.x as the OS and GateD as the BGP implementation. The hardware platform consisted of Pentium Pro (200-MHz 686-class CPU) with 64 to 256 Mbytes of memory.

3.2 Software

Our S-BGP countermeasures were added to GateD's BGP implementation. This software was instrumented to collect the performance statistics described below. In addition, we developed a man-in-the-middle (MITM) tool that was placed between a CAIRN BGP speaker and a BGP peer in the Internet. The MITM was used to intercept BGP control traffic coming from the BGP peer in the Internet, augment it with S-BGP countermeasure data, and then pass it to the S-BGP peer in the CAIRN testbed. In this fashion, we simulated the effect of having deployed S-BGP in the Internet. This tool also can be used to test the effectiveness and robustness of the S-BGP countermeasures. This can be done, for example, by injecting bad routes, making the link appear to go up and down, etc. We also developed a tool that accepts a recording of multiple sessions of BGP traffic and replays it to simulate the original order and arrival intervals for the BGP traffic. This latter tool allowed us to conduct tests with deterministic data, to facilitate comparison of results, e.g., for BGP-4 vs., S-BGP and for various optimizations of S-BGP.

3.3 Experiments

Several sets of experiments were run to measure the impact of S-BGP countermeasures. The questions of interest were:

- Do these countermeasures ensure the aspects of correct operation of BGP described in Section 1.1?
- How much overhead (bandwidth, CPU, storage) do they add to BGP-4? Are there optimizations that can reduce these costs?
- Is S-BGP readily deployable in the Internet? Does it interoperate with BGP-4 so that it can be deployed incrementally? Can iBGP be used to exchange S-BGP information between S-BGP routers in the same AS?

The scenario in which tests were conducted reflects a router in stable (steady state) operation. This is the case that reflects operation with S-BGP deployed. The S-BGP router processes UPDATEs sent/received due to changes in routing, but has (extracted) certificates, CRLs and AAs on hand. The frequency of UPDATE messages is dependent on the stability of the Internet. Routes fluctuate ("flap") for several reasons, e.g., link failure and re-establishment, router mis-configurations, software bugs, etc. Also, new network prefixes, ASes, and BGP speakers are added constantly to the network. In S-BGP, the resources consumed to transmit/send and generate/validate BGP UPDATE messages are a function of both the size and number of route attestations and the rate of routing changes in the Internet. The tables below show traffic characteristics observed during the month of May 1999.

We did not measure the effect of router initialization. The start up transient was not included in our tests because the results are completely predictable [1,3]. Also, the resources (bandwidth, CPU, memory) needed for establishing a BGP session (including IPsec costs) are negligible compared to those needed for other S-BGP traffic.

Experiments were performed with different values for several parameters. Tests were run with BGP-4 and with S-BGP, and several optimizations were compared for S-BGP. The optimizations examined included caching routes and attestations after verification (receiver) and after signing (transmitter). Cryptographic hardware was not employed, although the use of appropriate hardware would probably improve performance in many contexts. The tests made use of real Internet traffic, recorded and played back using the tool described above. In addition to normal operation, several attack scenarios were examined, e.g., injection of various forms of bad routes. The following sections describe these experiments in more detail.

4. Security

To test whether or not S-BGP delivers the desired improvement in BGP security, we tested S-BGP in the face of several types of attacks. We injected bad routes (NLRI and/or path attributes) that are not verifiable against the corresponding route and address attestations. These attacks were performed when the router was in stable operational mode and using playback of recorded traffic. The results showed that use of S-BGP detected and rejected spurious UPDATE messages, as expected.

5. Performance

We examined the impact of the S-BGP countermeasures on BGP performance. Bandwidth, storage, and CPU utilization in BGP speakers is driven by several factors. The rate of change in connectivity and routing policies determines the frequency of BGP UPDATEs. The length of an UPDATE is a function of the number of paths being withdrawn per UPDATE, the number of address prefixes being advertised per UPDATE and the path attributes included in the UPDATEs. The sizes of the routing tables (ADJ-RIB-IN, LOC-RIB, ADJ-RIB-OUT) are a function of the number of separately advertised destinations (address prefixes) and routes. Although aggregation of routes helps reduce the volume of UPDATEs and the number of entries in the routing tables, BGP speakers often experience heavy loads because of the large number of link flaps in the Internet, and because of the growth in the number of networks. Therefore it is crucial to understand the impact of S-BGP on the performance of these speakers.

The following sections discuss the overhead imposed by S-BGP countermeasures in terms of bandwidth, storage, and CPU utilization. The first three sections provide the background for this analysis. They describe the statistics that underlie these test results, summarize current BGP traffic characteristics, and discuss the factors that determine the overhead added to an UPDATE by an Attestation Path attribute. The last four sections discuss the results of the performance analysis. They analyze the volume of certificates and address attestations needed to support S-BGP, and discus the impact on BGP of implementing these countermeasures with and without optimizations.

5.1 Underlying statistics

The results in the later sections are driven by a variety of aspects of the Internet and its administration. Some of the key numbers are shown below. These were obtained via analysis of 1999 data from Merit BGP statistics and from other open sources.

- number of ASes - Of 6,262 AS numbers issued as of June 1999, only 5,275 were observed in dumps of actual BGP traffic.
- number of organizations assigned address prefixes - The 194,501 address blocks that have been issued as of February 1999, are assembled into contiguous ranges that are assigned to 43,931 organizations. Note: Each organization will be assigned one certificate attesting to its ownership of address blocks, not one certificate per address block. (This count overstates the number of certificates required, as an organization that has acquired multiple prefixes over time is counted more than once.)
- number of address prefixes in an Internet routing table - Of the 194,501 address blocks which have been issued as of February 1999, only 74,191 were visible in dumps of actual BGP traffic.
- number of BGP speakers for all ASes - This was estimated to be 7,455 by examining the routing registry databases and estimating that one could divide the 5,275 ASes into 3 groups: approximately 20 large ISPs with 50 speakers plus 300 ISPs with 5 speakers plus the remaining 4,955 ASes with 1 speaker.

4	Number of Internet address registries
1,781	Number of organizations with one or more ASes
5,275	Number of ASes
43,931	Number of organizations assigned address prefixes

74,191	Number of address prefixes in an Internet routing table (LOC-RIB)
7,455	Estimated number of BGP speakers

Table 1: Administrative and routing statistics

5.2 BGP traffic characteristics

In order to assess the impact of S-BGP countermeasures, it is essential to understand the nature of current BGP traffic, e.g., the number and kinds of BGP control messages that have to be protected. This section provides a summary of some current BGP traffic characteristics as received by a BGP router during a realtime peering session.

3,571	Number of KEEPALIVEs per day
68	Kbytes of KEEPALIVEs per day
1,426	Number of UPDATEs per day
89	Mbytes of UPDATEs per day (not including attestations)

Table 2: BGP daily traffic statistics

5.3 Attestations

Several factors determine the number of bytes of overhead added to an UPDATE by an Attestation path attribute. The length (bytes) of an Attestation path attribute is a function of the three components shown below. In the later sections, to be conservative, we examined the case that results in adding the most overhead, where all protected fields are assumed to have to be explicitly listed in the Route Attestations.

- Path attribute header (3 or 4 bytes)
- Route Attestations (RAs)
  - the number of Route Attestations (the number of ASes in the AS_PATH) – One route attestation is needed for each AS in the path.
  - the length of each Route Attestation – This is a function of some fixed length fields (header, issuer, signature, validity fields, subject, etc.) plus any UPDATE fields that are "explicitly" listed in the RA, as opposed to being "implicitly" listed (assumed to be the values in the body of the UPDATE). Which of the UPDATE fields are protected by the RA, is determined by the security policy of the AS(es) that generated the RAs in the UPDATE(s). (If route aggregation has occurred, then multiple ASes and UPDATEs will be involved.) For each RA in an UPDATE, the protected fields are listed explicitly only as necessary.
- Address Attestations (AAs) (if included in an UPDATE)
  - the number of Address Attestations (zero to a small number) – Typically, AAs are distributed to S-BGP speakers out of band rather than in UPDATEs. But an AA may be sent in an UPDATE if an AS wants to flood an AA through the Internet more quickly than would happen with the usual distribution mechanism, e.g., an organization has changed ISPs and the new ISP wants to quickly send out an AA to advertise that its AS now has the right to advertise the organization's address prefix and that the old AA is no longer valid.
  - the length of the Address Attestation – This is a function of some fixed length fields plus how many address prefixes are covered (one to a small number).

5.4 Certificates

A key component of the overhead of S-BGP is the certificates that are used to verify the binding between an organization and an AS, between an AS and a router, and between an organization and a set of IP address prefixes. This section lists the estimated number and kind of certificates that are required to support S-BGP. Note that because of the size limitation on a BGP UPDATE message, it generally would not be possible to fit all of the certificates (and attestations) needed to validate an UPDATE into the UPDATE message itself. Accordingly, subsequent sections assume that each S-BGP speaker validating UPDATEs will have access to pre-validated certificates at a repository for its AS. The speaker's AS, e.g., its NOC, will have already fetched the complete set of certificates (and AAs) from a higher level repository, e.g., at an Internet registry, validated them, extracted the needed fields, and signed the resulting digested data. This data is then uploaded to the S-BGP speakers. Although the certificates vary somewhat in length because they contain different extensions, etc., they are approximately the same size.

.

4	1 per Internet registry
43,931	1 per organization that has been assigned an address prefix

1,781	1 per organization that has been assigned an AS number
5,275	1 per AS (on the assumption that each AS has BGP routers)
7,455	1 per AS/BGP speaker
58,446	total certificates at present
550	average # bytes/certificate (using 1024-bit DSA keys)
32.2	total Mbytes
175	total number/day new certificates
96.3	total Kbytes for new certificates

Table 3: Certificate Statistics

5.5 Address Attestations

Another component of the overhead of S-BGP is the address attestations (AAs) that are used to verify the authorization of the originating AS to advertise the NLRI in an UPDATE. As noted above, because of the size limitation on a BGP UPDATE message, it is not possible to fit all of the certificates and attestations needed to validate an UPDATE into the UPDATE message itself. This analysis assumes that each S-BGP speaker will have access to pre-validated AAs at a repository for its AS, using the same distribution approach described for certificates. Although the AAs vary somewhat in length because they contain different address prefixes, etc., they are approximately the same size. Table 4 summarizes the computation of the AA storage requirements, exclusive of the overhead imposed by data structure formats for storage in GateD.

.

43,931	1 per organization which has been assigned an address prefix
96	bytes per AA
4.2	Mbytes total

Table 4: Address Attestation Storage

5.6 Performance without optimizations

We first tested the performance impact of the S-BGP countermeasures without employing any optimizations. In addition to the features of the basic architecture such as local caching of certificates, CRLs and AAs, we make several assumptions: every UPDATE with route advertisements has to be validated; all validation of UPDATEs must be performed upon receipt of the UPDATE; all cryptographic operations are implemented in software, and no validation results are cached for later re-use.

The results are presented on a per peering session basis. This allows one to estimate the overhead seen by a router with N peers by simply multiplying the processing time and RA storage requirements by N. Also, this provides a good approximation of how a multi-peer speaker behaves, since verification is done only for the “best” route. Under typical conditions, a BGP speaker receives one copy of a given UPDATE from every one of its external peers, plus one from every other BGP router in the same AS. In the absence of “best route only” validation, every one of the UPDATEs received via eBGP must be validated, whereas those received via iBGP are assumed to have been validated by the sender. For example, at a NAP, a router might have approximately 30 eBGP peers. So one would multiply the storage and CPU numbers for one peer by 30 plus the number of BGP peers in the router’s own AS to obtain the total RA storage and CPU needed. Additional bandwidth would be on a per link basis, so the bandwidth numbers shown below would not have to be multiplied by the number of peers. Since validation of a route is done only for the route that is selected, the CPU costs approach those of the single peer case. (The router may not receive the "best" route first and hence may have to verify more than one UPDATE, depending on the order of arrival of different paths to the same destination, when the best one arrives, how long one waits before deciding one has the "best" route, etc.)

For these tests, we compared BGP-4 to S-BGP, in stable operation, using playback of a realtime peering session with a BGP router in an ISP. We verified all UPDATEs immediately upon receipt (for one peer). Given the size of the routing table (LOC-RIB) and the daily volume of UPDATEs, these experiments showed a significant impact on CPU and storage. A number of optimizations (see Section 5.7) can be implemented that reduce the S-BGP overhead.

5.6.1 Bandwidth utilization: S-BGP consumes link bandwidth at three different stages: 1) during the initial routing exchange when starting up a BGP peering session (including use of IPsec), 2) whenever there are changes in routing information, and 3) at steady state with the exchange of KEEPALIVE messages to maintain the BGP peering session. Assuming that certificates and AAs are already present during stable operations, the bandwidth utilization for the S-BGP overhead needed for a sample peering session was 1.4 Kbs. During steady state operations, the S-BGP overhead is dwarfed by the user traffic and thus does not represent a significant adverse impact on router traffic.

5.6.2 CPU utilization: In the unoptimized case, a BGP speaker validates every route advertisement that it receives from an external peer. This requires validation of one route attestation for each unique AS in the AS_PATH attribute, and possibly (though with low frequency) a small number of address attestations. Each validation requires hashing the covered data and validation of the signature. In addition, the BGP speaker must sign one route attestation for every UPDATE it sends. The numbers below are based on the use of SHA-1 for hashing and DSA with a 1024 bit key for signing and verification.

	Steady state minutes/day
processing RAs (sending)	16.2
processing RAs (receiving)	123.7
total	139.9

5.6.3 Storage/Memory utilization: With regard to storage/memory utilization, the space needed for certificates (Section 5.4), AAs (Section 5.5), and RAs to support validation of UPDATEs during a sample peering session is shown below. Note that only the certificate and AA figures are per-router, whereas the RA figures are per-peer. (The AA storage figure here is considerably greater than that provided in Section 5.5, because of data structure format overhead.)

	Storage (Mbytes)
certificates	32.2
AAs	10.1
RAs	16.7
total S-BGP overhead	60.0
total BGP w/o S-BGP	19.9

5.7 Performance with optimizations

Our initial set of experiments (Section 5.6) indicated that the S-BGP countermeasures impose significant overhead on an ongoing basis (CPU, storage). However, we have identified a number of optimizations that greatly reduce the overhead costs of the BGP countermeasures. These are described in the following sections along with the savings that resulted.

5.7.1 Handling peak loads: Three optimizations address the problem of S-BGP costs by adding resources or deferring the work involved in S-BGP countermeasures. Although these approaches do not reduce the total amount of work that has to be performed, they address the problem of "peak loads" that the router might otherwise be unable to handle in a timely manner. They include:

- use of an auxiliary or upgraded processor to enable routers to handle processing of BGP countermeasures data
- prioritization of verification tasks to reduce the impact of topology changes due to hardware/software component failure and recovery, congestion, etc. This technique can also reduce the amount of work done, e.g., by deferring validation of a route until the route is needed and therefore avoiding validation of never-used routes. The router verifies only the "best" route chosen for LOC-RIB.
- background verification of alternate routes – In this case, the router spends idle processing time verifying any alternate routes so that if they are needed (e.g., the current best route is withdrawn), they have already been verified and are ready for use.

5.7.2 Caching routes and attestations: Examination of BGP traffic shows that most of the UPDATEs (withdrawals and advertisements) in the Internet appear to be caused by temporary changes in link status (link "flapping") rather than by actual changes to the topology (new networks, ASes, or links; genuine decommissioning of a link or network). Accordingly, when an UPDATE containing a "withdrawal" appears, an S-BGP speaker can keep the old already validated (but now "withdrawn") route and associated attestations in its cache and mark the route "withdrawn" rather than deleting it. The S-BGP speaker can then simply compare each newly received route to previously received routes for the same NLRI and validate only those routes that are not in its cache. Similarly, locally generated RAs can be cached for use in sending UPDATEs.

- Bandwidth – The proposed caching does not affect bandwidth, i.e., the same UPDATEs are received and transmitted.
- CPU – Analysis of BGP data from NAPs (from Merit) suggests that the proposed caching should result in significant savings in the steady state case. That analysis showed that caching one route per destination would enable a router to avoid re-validating about 53% of UPDATEs during a real

world BGP peering session. However, the BGP data used in the experiments was provided by an ISP, with our testbed acting like a DSP, not a peer ISP. As a result, most route flaps seem to have been removed and thus there was only a very small performance gain resulting from use of a cache.

- Storage – This caching affects storage requirements for routes and route attestations. The more distinct routes per destination per peer that are kept, the fewer routes that need to be re-validated, but the more storage that is required. The storage required for handling certificates and AAs is not affected by this optimization. Note: BGP already requires the router to keep one route per destination/peer (ADJ-RIB-IN). (The additional storage needed for a depth one cache arises for outbound RAs only.)

	Without caching	**With a cache of depth 1**
S-BGP storage	16.7 Mbytes	25.8 Mbytes
CPU	139.9 minutes/day	137.3 minutes/day

5.7.3 Using certificate and AA extracts: Instead of storing the entire certificate or AA, the S-BGP speakers and certificate/AA repositories can keep only the needed fields (e.g., issuer, public key, validity period, extensions) and discard unneeded fields (e.g., signature and ASN encoding). This does not affect CPU usage, but significantly reduces the amount of bandwidth needed to transmit certificates and AAs (both at initialization and on an ongoing basis) and reduces the space needed to store them. A typical certificate for this application would be about 550 bytes long; the average extract is 160 bytes. The average AA is about 96 bytes long; the average extract is 40 bytes long.

	At router initialization
using full certificates and AAs	36.4 Mbytes
using extracts	11.2 Mbytes

5.7.4 Use of cryptographic hardware: The use of cryptographic hardware provides better protection for keys and may provide better performance depending on the hardware selected and the software that it replaces. Significant CPU savings can be realized by the addition of cryptographic hardware, e.g., a PC card or special chip specifically designed to do cryptographic operations, though this is not necessarily true in all cases. For example, at present, there is no cryptographic hardware designed to speed up hashing. Therefore, use of general purpose processors may produce the same or better performance as use of cryptographic hardware for hashing. The speed of this operation depends heavily on the quality of the implementation and can vary by an order of magnitude.

With regard to signing and verification, the amount of savings will vary widely depending on the hardware approach chosen. PC cards such as the Spyrus LYNKS provide lower performance but may be easier to add to a router, i.e., just plug it in to the system. Specialized chips such as the HiFn can provide greater performance but require upgrading an existing router board or creating a new board. In this round of tests, we were unable to introduce cryptographic hardware to provide a comparison with the software used in the unoptimized tests, due to time limitations.

5.7.5 Where is S-BGP needed?: It should be observed that S-BGP does not need to be deployed in most of the routers in the Internet. It will achieve the desired results if deployed in just the BGP routers of the half a dozen or so major ISPs.

- If a router receives only one route to an address prefix, then if it has traffic to send to that destination, it has no choice but to use this route to reach that destination. So there is no need to validate the route. The two most likely scenarios where this will happen are when destinations have only one path to the router in question, e.g., singly homed "leaf" organizations and DSPs, or when a multi-homed DSP advertises one path to an address block.
- Today, most organizations obtain their address blocks from their ISP rather than from an Internet registry. In many cases, the ISP will handle advertisements of routes to its customer organizations and consequently will handle the certificates and route/address attestations associated with any address prefixes which it assigns to its customer organizations. These organizations do not generate or verify address attestations or need certificates for their address prefixes.
- Most DSPs and subscribers use default routing, not BGP and consequently they do not need to validate (or receive) attestations.

6. Deployability

In addition to performance that supports scalability, in order to be deployable, S-BGP must interoperate with BGP. To verify that S-BGP meets these requirements, we tested a critical feature that would affect deployment. Using a playback of recorded traffic (with a BGP router in the Internet), we verified that S-BGP speakers could distribute attestations between S-BGP speakers in the same AS. These tests were performed when the router was in stable operational mode. No problems were observed. UPDATEs were successfully processed.

7. Future Work

The next step toward deployment of S-BGP calls for working with the registries and ISPs to put prototype certification authorities in place and gain experience with the relevant policies and procedures. We also plan to work with router vendors to integrate S-BGP countermeasures into their products.

8. Conclusions

In addition to the security-related benefits to be gained, performance considerations are crucial in convincing users and vendors to adopt the S-BGP countermeasures and deploy them into the Internet. Our experiments with a prototype implementation and real-world BGP traffic supported prior analysis results which indicated that the overhead added by the S-BGP countermeasures needed the CPU/memory equivalent of a desktop PC.

- A number of techniques can be used to enable the BGP speakers to handle spikes in the UPDATE traffic. This includes use of auxiliary processors, deferral of route validation until a route is needed, and offloading of certificate and AA processing.
- CPU — Although caching of recent route data should enable a speaker to avoid the need to validate approximately 53% of UPDATEs, our testing showed that a DSP might see little benefit from caching. However, the UPDATE arrival rate in such circumstances is sufficiently low to mitigate CPU utilization concerns anyway. In contrast, we anticipate that an S-BGP speaker at a NAP would benefit considerably from caching, based on analysis of Merit data. Cryptographic hardware could be added to handle the signature and verification tasks, but high speed signature algorithm software, e.g., the OpenSSL distribution, provides very good performance. CPU requirements for processing certificates and AAs are addressed by having organizations/ASes handle this processing and using an out-of-band distribution of certificates and AAs to all S-BGP speakers.
- Bandwidth — The increased transmission bandwidth required by S-BGP on a steady-state basis represents a small amount of data relative to subscriber traffic. In addition, a number of optimizations have been adopted to minimize overhead – the encoding scheme used for attestations, the choice of signature algorithm, and the use of certificate and AA extracts. Even at initialization, the time required to transmit certificates and AAs for the full Internet routing table is minimal.
- Storage — Analysis of the storage capacity and utilization of the routers currently used by ISPs indicates that many would not be able to cache S-BGP route and route attestation data. Either router upgrades or auxiliary boxes are needed to provide this space.
- Operational issues — This architecture supports straight forward solutions for incremental deployment, i.e., maintains interoperability between S-BGP and BGP-4 and between eBGP and iBGP.

In the near term, S-BGP's resource requirements can be met by the addition of an auxiliary device, essentially a PC, to each router. By migrating BGP processing to this device, the storage problems cited above are averted. These devices can easily be provisioned with sufficient RAM to support the larger routing table size implied by the addition of the S-BGP countermeasures data. The space required for the certificate and AA databases is minor compared to typical disk drive capacities. Moreover, low cost cryptographic support for signature generation and validation can be provided through the use of cryptographic hardware, e.g., adding a fast modular arithmetic chip to a router, or plugging in a PCMCIA crypto-processor card. Use of hardware also offers improved security for the private keys employed by the routers to sign RAs and for IPsec (IKE) key management. In the longer term, the routers themselves could be upgraded with additional CPU power and memory.

9. Acknowledgements

Many individuals contributed to the design and development of S-BGP. Initial funding was provided by NSA, in April of 1997, yielding a first cut design. DARPA and NSA provided later support, enabling us to refine the design, implement it, and conduct the tests reported in this paper. S-BGP benefited significantly from the insight and efforts of Martha Steenstrup and Luis Sanchez. As members of the architecture team, their contributions were critical to the design of the attestation and PKI schemes, as well as the evaluation of other approaches and of performance and operational issues. The authors also would like to thank Michelle Casagni, for her work during the initial performance analysis phase, and Dennis Rockwell and Nicholas Shectman for their efforts during the implementation and experimentation phase.

10. References

[1] BBN Report 8217, "An Architecture for BGP Countermeasures," November, 1997.

[2] Y. Rekhter, T. Li, "A Border Gateway Protocol 4 (BGP-4)," RFC 1771, March, 1995.

[3] S. Kent, C. Lynn, and K. Seo, "Secure Border Gateway Protocol (S-BGP)," to appear in IEEE JSAC.

[4] S. Kent, R. Atkinson, "Security Architecture for the Internet Protocol," RFC 2401, November 1998.

[5] S. Murphy, "BGP Security Analysis," draft-murphy-bgp-secr-02.txt, November 1998.