Secure Border Gateway Protocol (Secure-BGP)

Stephen Kent, Charles Lynn, and Karen Seo

Published in IEEE Journal on Selected Areas in Communications Vol. 18, No. 4, April 2000, pp. 582-592

Abstract-- The Border Gateway Protocol (BGP), which is used to distribute routing information between autonomous systems (ASes), is a critical component of the Internet's routing infrastructure. It is highly vulnerable to a variety of malicious attacks, due to the lack of a secure means of verifying the authenticity and legitimacy of BGP control traffic. This document describes a secure, scalable, deployable architecture (S-BGP) for an authorization and authentication system that addresses most of the security problems associated with BGP. The paper discusses the vulnerabilities and security requirements associated with BGP, describes the S-BGP countermeasures, and explains how they address these vulnerabilities and requirements. In addition, this paper provides a comparison of this architecture with other approaches that have been proposed, analyzes the performance implications of the proposed countermeasures, and addresses operational issues.

Index Terms–security, public-key cryptography, routing, digital signatures, denial of service

I. Problem Description

Internet routing is based on a distributed system composed of many routers, grouped into management domains called Autonomous Systems (ASes). Routing information is exchanged between ASes in Border Gateway Protocol (BGP) [1] UPDATE messages. BGP has proven to be highly vulnerable to a variety of attacks [2], due to the lack of a scalable means of verifying the authenticity and legitimacy of BGP control traffic. In April 1997, we began work on the security architecture described in this paper. In this section we describe the problem–how the protocol works, the nature of observed BGP traffic in the Internet, the correct operation of BGP, the threat model and BGP vulnerabilities, and the goals, constraints and assumptions that apply to the proposed countermeasures.

A. Overview of BGP

The BGP-4 protocol, both message syntax and the route propagation algorithm, is described in [1]. Routers implementing BGP, BGP "speakers," exchange routing information via UPDATE messages. An UPDATE message consists of three parts: a list of address prefixes[1] for destinations that are no longer reachable (via the previously specified route); a list of prefixes that are reachable; and the characteristics of the cumulative path and current inter-AS hop, contained in path attributes, that can be used to reach the address prefixes. The attribute used to specify the inter-AS path, the AS_PATH attribute, specifies a sequence of Autonomous Systems (ASes) along the path, each identified by its AS number.

When propagating an UPDATE to a neighboring AS, the BGP speaker prepends its AS number to the sequence, and updates certain other path attributes. Since an UPDATE can specify only one path, only prefixes that share that path may be aggregated into the UPDATE.

The backbone routers of the major internet service providers (ISPs) have a route to every reachable IP address. Analysis of BGP UPDATEs recorded during January 1999, showed routing databases that contained about 61,000 IPv4 address prefixes. Each (non-leaf) 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 UPDATEs) from each of its neighbors. The worst case arises at Network Access Points (NAPs), where ISPs are connected together via a high speed (100Mb/s) LAN. A BGP speaker at a NAP might have about 30 peers.

On a daily basis, a BGP speaker at a NAP receives about 1425 UPDATEs from each peer, an average UPDATE rate of about 1 per minute per peer. This rate is affected somewhat by Internet growth (about 25 network prefixes are added each day), but is mostly a function of UPDATEs sent due to link, component, or congestive failures and recoveries. Analysis shows that about 50% of all UPDATEs are sent as a result of route "flaps," i.e., transient communication failures that, when remedied, result in a return to the original route. This sort of routing behavior has long been characteristic of the Internet[2] [3] and the proposed security mechanisms take advantage of this behavior to achieve acceptable performance, as discussed in Section VI.

B. Correct Operation of BGP

Security for BGP is defined by the correct operation of BGP speakers (Byzantine failures). This definition is based on the observation that any successful attack against BGP should result in other than correct operation, presumably yielding degraded operation. Correct operation of BGP depends upon the integrity, authenticity, and timeliness of the routing information it distributes as well as each BGP speaker's processing, storing, and distribution of this information in accordance with both the BGP specification and with the (local) routing policies of the BGP speaker's AS. The following statements characterize the primary correct operation features of BGP.

- Each UPDATE received by a BGP speaker from a peer was sent by the indicated peer, was not modified en route from the peer, and contains routing information no less recent than the routing information previously received for the indicated prefixes from that peer.
- The UPDATE was intended for receipt by the peer that received it.
- The peer that sent the UPDATE was authorized to act on behalf of its AS to advertise the routing information contained within the UPDATE to BGP peers in the recipient AS.
- The owner of an address space corresponding to a reachable prefix advertised in an UPDATE was authorized by its parent organization to own that address space.
- The first AS in the route was authorized, by the owners of the address space corresponding to the set of reachable prefixes, to advertise those prefixes.
- If the UPDATE indicates a withdrawn route, then the peer withdrawing the route was a legitimate advertiser for that route, prior to its withdrawal.
- The peer that sent the UPDATE correctly applied the BGP rules and its AS's routing policies in modifying, storing, and distributing the UPDATE, in selecting the route, and in deriving forwarding information from it.
- The BGP speaker that received the UPDATE correctly applied the BGP rules and its AS's routing policies in determining whether to accept the UPDATE.

The countermeasures developed for S-BGP meet the first six of these criteria, even in the face of subversion of BGP speakers (Byzantine failures). Section IV provides a detailed analysis of how each countermeasure contributes to correct operation. However, because the local policy features of BGP allows a speaker considerable latitude in determining how to process an UPDATE, these countermeasures cannot meet the last two criteria, i.e., such attacks could be attributed to local policies not visible outside an AS. To address such attacks, the semantics of BGP itself would have to change. Moreover, because UPDATEs do not carry sequence numbers, a BGP speaker can generate an UPDATE based on old information, e.g., withdrawing or reasserting a route based on outdated information. Thus the temporal accuracy of UPDATEs, in the face of Byzantine failures, is enforced only very coarsely by these countermeasures. (Section V provides more details on residual vulnerabilities.)

C. Threat Model and BGP Vulnerabilities

BGP has a number of vulnerabilities that 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. BGP uses TCP/IP for transport and this protocol, and its payload, can be attacked. 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. Most of these attacks transform routers into hostile insiders. Effective security measures must address such Byzantine attacks.

Exploitation of these vulnerabilities allows a variety of attacks. For example, fictitious BGP messages might be injected into a link (spoofing). Authentic BGP messages might be captured and either modified and re-injected into the link, combined incorrectly, or suppressed altogether. A compromised BGP speaker could generate UPDATEs for routes that do not, legitimately, pass through that speaker. All of these attacks are countered by the mechanisms described in Section III.

UPDATE messages could be generated too frequently by a compromised BGP speaker, or the selection of routes and distribution of UPDATEs could violate the local routing policies. These failures are not addressed by the proposed countermeasures.

Better physical and procedural security for network management facilities, BGP speakers, and communication links; link-level encryption of inter-router (BGP speaker) traffic; and end-to-end encryption of management information would reduce some of these vulnerabilities. However, some aspects of such security approaches are economically unattractive or infeasible. Moreover, accidental (vs. malicious) misconfiguration would not be prevented by such measures and such misconfiguration has proved to be a source of several significant Internet outages in the past. Any security approach that leaves BGP vulnerable to such benign "attacks" violates the "principle of least privilege" and leaves the Internet routing system vulnerable at its weakest link. In contrast, the security approach described here satisfies this principle, so that any attack on any component of the routing system is limited in its impact on the Internet as a whole.

D. Goals, Constraints, and Assumptions

In order to create countermeasures that are both effective and practical, the S-BGP architecture is based on the following goals, constraints, and assumptions.

The S-BGP architecture must handle the projected growth and usage of the Internet in terms of performance (storage, processing, network bandwidth). It should be dynamic (responding automatically to topology changes, including the addition of new networks, routers and ASes) and scalable (able to handle the growth of the Internet in terms of addresses, routes, BGP control traffic, etc.).

The countermeasures must be consistent with the BGP protocol standards and with the likely evolution of these standards. This includes packet size limits, e.g., 4096 byte maximum for UPDATEs, and BGP features, e.g., path aggregation, communities, and multi-protocol support, e.g., multi-protocol label switching (MPLS). For example, to avoid modifying BGP packet formats, we have chosen to employ the BGP optional, transitive path attribute as a mechanism for distributing countermeasures information. Non S-BGP routers should pass this attribute type transparently, without understanding it.

The S-BGP architecture must be deployable. A primary goal of this work is to not only find countermeasures for BGP vulnerabilities but to cause them to be adopted by ISPs and router vendors. To accomplish this, the countermeasures must use available technology that can be incrementally deployed and they should leverage off of the existing infrastructure, e.g., the Internet Corporation for Assigned Names and Numbers (ICANN), and routing registries. In addition, they must avoid dependency loops, i.e., the architecture cannot depend on correct operation of inter-AS routing during initialization, e.g., it cannot rely on non-local databases.

II. Prior Work

The earliest significant work published on the topic of routing protocol security is Perlman's doctoral dissertation [21]. The S-BGP design shares several features of that work, e.g., we address Byzantine failures and make extensive use of digital signatures. However we differ in many other respects, e.g., our design applies to a standard exterior (vs. interior) routing protocol, and we pay considerable attention to infrastructure and performance implications.

At the time that we began this work, previously published work on improving the security of BGP, and more generally distance-vector protocols, included proposals for adding sequence numbers to BGP messages [4,5,6], authentication of BGP messages [1,5,6], neighbor-to-neighbor encryption of BGP messages [4], and adding information to UPDATE messages to protect against tampering as the UPDATE propagates around the Internet [4,5,7].

None of this work proposed a comprehensive solution to the BGP security problems described above; each focused on one or more aspects of the problem without considering the full range of issues that are critical to a viable solution. For example, none addressed issues associated with the generation and distribution of public key certificates and certificate revocation lists (CRLs) needed to support validation of signed UPDATEs. Some proposals made changes to BGP that are inconsistent with the protocol standards, a reasonable approach only if one were presented with a "clean slate." None of the prior work examined the statistics of BGP operating in the Internet; this sometimes led authors to focus on performance concerns that are not the major impediment to deploying viable solutions. Some of the work developed solutions for distance vector protocols, but erroneously claimed applicability to BGP, which is described as a path vector protocol.

In contrast, the BGP security architecture reported in this paper is comprehensive, including a design for the infrastructure needed to establish and maintain the system. The optional transitive path attribute it employs is consistent with BGP standards and can be safely carried through routers not implementing S-BGP. This architecture incorporates the notion of an address attestation, which establishes that a "first hop" BGP speaker is authorized to advertise a route to a destination. No prior work includes an equivalent notion. Finally, the performance of the design presented here has been modeled based on actual BGP statistics. No other work has been so rigorously analyzed from a performance perspective.

In [7], the scheme proposed is similar to our route attestations (RAs) in that before distributing an UPDATE to an external neighbor, the BGP speaker signs the route. Our approach differs from the scheme proposed in [7] in that we sign a routing data structure that specifies the next hop AS, explicitly indicating that this AS is authorized to advertise the route in question to the identified neighbor. Hence, our approach avoids a vulnerability not addressed in [7], i.e., provides protection against "cut and paste attacks" in which a BGP speaker inserts itself into a route (using a valid UPDATE containing a route which the speaker was not authorized to use).

The approach proposed in [4,5] was developed in response to the perceived communication and computation overhead of schemes such as [7]. In the case of [7] (and our route attestations), each of the signatures must be carried in each UPDATE and verified by each recipient to validate each received UPDATE. The approach proposed in [4,5] includes only a single signature for a route in an UPDATE; this signature covers only the destination and penultimate ASes listed in the path and is generated by a BGP speaker in the destination AS. While the overhead of [4,5] is less than that of [7] or our route attestations, it fails to provide the protection afforded by the iterated signature schemes in an environment with more sophisticated routing policies (e.g., policies other than shortest path), such as those typically supported by BGP. Moreover, our analysis shows that generation, validation, and transmission of digital signatures does not impose an unacceptable computational or communication burden.

III. Proposed Countermeasures

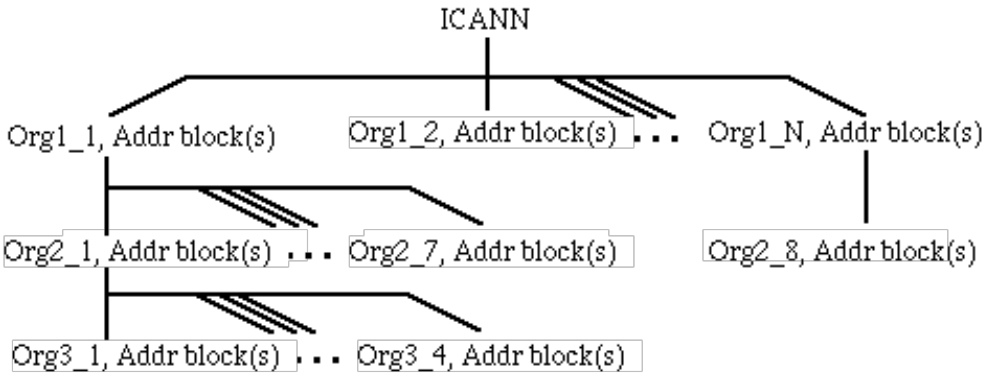
The approach adopted to securing BGP route distribution involves two Public Key Infrastructures[3] (PKIs), a new path attribute containing "attestations", and the use of IPsec. These components are used by a BGP speaker to validate the authenticity and data integrity of BGP UPDATEs that it receives and to verify the identity and authorization of the senders. This section discusses in more detail the PKIs and certificates, the attestations, the use of IPsec, and the distribution of this countermeasures information.

A. Public Key Infrastructures (PKIs) and Certificates

S-BGP uses two PKIs, based on X.509 (v3) certificates, to enable BGP speakers to validate the identities and authorization of BGP speakers and of owners of ASes and of portions of the IP address space. These PKIs parallel the existing IP address and AS number assignment delegation system and take advantage of this extant infrastructure. Because these PKIs mirror existing infrastructure, their creation avoids many of the "trust" issues that often complicate the creation of a PKI. The two PKIs involve four types of certificates, as illustrated below. In the diagrams:

- The higher node is the issuer for the certificates defined in the tier below it.
- The name of the current tree node (organization, AS, router, etc.) is the subject of the certificate.
- Any additional fields shown in the node, e.g., address block(s), are in an extension in the certificate.
- Other X.509 certificate fields are assumed, but not shown–sequence number, subject public key, signature, validity period, etc.

Note that the organizations that assign addresses (Registries, ISPs, DSPs, etc.) and the organizations that obtain autonomous system numbers from a Registry may be different. An organization could receive its AS number from a registry and its address block from an ISP. So the Org3_4 shown in the first PKI hierarchy (Figure 1) could correspond to the Org 1 shown in the second PKI hierarchy (Figure 2).



Subject	Description
Org1_x	a 1st tier organization (usually a registry)
Org2_x	a 2nd tier organization (usually an ISP or DSP)
Org3_x	a 3rd tier organization (usually a DSP or user organization)

Extension	Description
Addr blocks(s)	one or more IP address blocks assigned to the organization

Figure 1: Address Allocation PKI Structure

1) A PKI for Address Allocation: This architecture calls for a certificate to be issued to each organization that is granted "ownership" of a portion of the IP address space. This certificate is issued through the same chain of entities that, in the existing environment, is responsible for address allocation. The root of this chain is the ICANN, followed by regional address space allocation authorities (e.g., ARIN and RIPE), ISPs, DSPs, and end users. Note that the proposed system does not require that address assignments be certified all the way to the subscriber. If a subscriber's address is allocated from that of a DSP or an ISP with which it is currently affiliated, then the certification process need only be effected as far as the ISP/DSP. The same applies to DSPs that receive their address-space assignments from ISPs. Also note that a subscriber (or a DSP) who does not participate in BGP exchanges (e.g., is singly-homed) need not be issued a certificate if the subscriber's address space is derived from that of an encompassing ISP or DSP.[4] Finally, if an organization owns multiple ranges of addresses, this design calls for assigning a single certificate[5] containing a list of address blocks, so as to minimize the number of certificates needed to validate an UPDATE.

This PKI reflects the assignment of address blocks to organizations by binding address block(s) to a public key belonging to the organization to which the addresses are being assigned. Unlike a typical X.509 certificate, the identity of the subject is not the primary focus of these certificates; instead, these certificates are used to prove ownership of a block of addresses.[6] Each certificate in this PKI contains a (private) extension that specifies the set of address blocks that have been allocated to the organization. The subject alternate name in each certificate is the DNS name of an organization: an ISP, DSP, or a subscriber. The ICANN, as root, is represented nominally by a self-signed certificate that contains an extension expressing ownership of the entire address space.

In Figure 1 we note the following.

a. ICANN is the root and issues certificates to the first tier of organizations (Org1_x). Under current practice, Org1_x would be an Internet Registry, although historically it could have been an ISP, an organization, etc. ICANN signs the tier 1 certificates using its private key.

b. Org1_x then assigns sub-blocks of its address space to ISPs or DSPs. In the diagram, for example, Org1_1 issues a certificate to each of Org2_1 through Org2_7 and Org1_N issues a certificate to Org2_8. Org1_x signs the certificate using the private key corresponding to the public key in the certificate it received in (a).

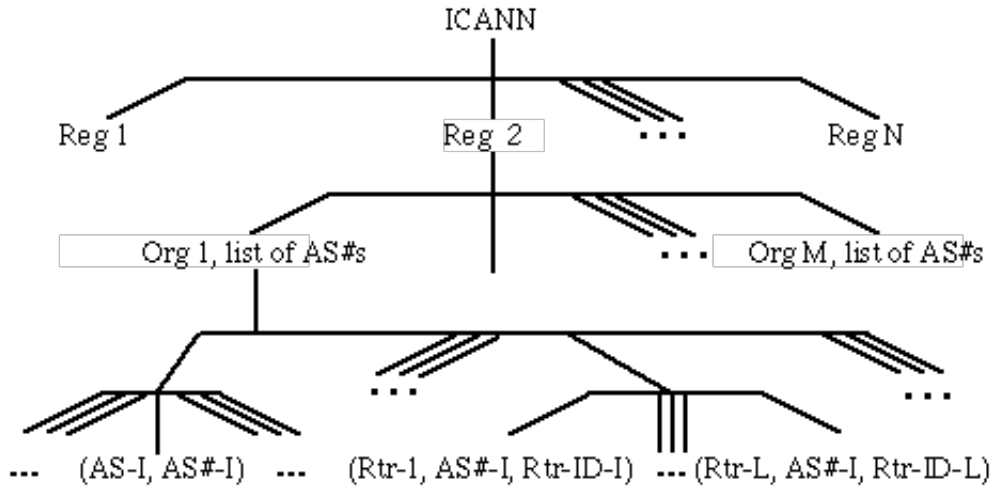
c. Org2_x then assigns sub-blocks of its address space to customers, DSPs, etc. In the diagram, Org2_1 issues a certificate to each of Org3_1 through Org3_4. Org2_x signs the certificate using the private key corresponding to the public key in the certificate it received in (b).

d. And so on….

Table 1 summarizes the Issuer/Subject relationships for the certificates in this PKI.

Certificate Type	Issuer	Subject
Root	ICANN	ICANN
Registry	ICANN	Registry
ISP/DSP	Registry (or ICANN or ISP)	ISP/DSP
Subscriber	ISP/DSP (or Registry or ICANN)	Subscriber

Table 1: Address Allocation PKI Certificate Overview



Subject	Description
Reg n	DNS name of registry n
Org m	DNS name of ISP/DSP/organization m
AS-i	DNS name of autonomous system i
Rtr-j	DNS name of router j

Extension	Description
AS#-i	autonomous system number for AS-i
Rtr-ID-k	router identifier (IP address) for router k

Figure 2: Autonomous System Identification and BGP Speaker PKI

2) A PKI for Assignment of ASes and Router Associations: Three types of certificates will be used to support the authentication of ASes and BGP speakers, and the relationship between speakers and ASes. Here too the ICANN is the root and the next tier consists of registries, but the third tier consists of organizations that own ASes, followed by a tier of AS numbers and routers. The result is a broader, shallower certification tree. As before, this tree parallels existing "trust relationships," i.e., the ICANN assigns AS numbers to registries, which in turn assign one or more AS numbers to organizations (e.g., ISPs/DSPs) that run BGP. Each of these organizations is authoritative for identifying routers as representatives (BGP speakers) for the AS(es) that the organization owns. In order to express the ownership of an AS by an organization, each third tier certificate carries an extension that enumerates the ASes assigned to that organization. Validation of fourth tier certificates requires matching asserted AS numbers against these extensions. For each fourth tier AS certificate (b, below) there are typically several router (BGP speaker) certificates (c, below), each specifying the same AS number. (Note that there could be more than one certificate assigned to a BGP speaker if the speaker acts as a proxy for another AS.) As shown in the Figure 2, these 3 types of certificates bind together:

a. AS numbers and an organization's public key–a registry issues these to organizations and signs them using its private key. The alternate name in the certificate is the DNS name of the organization. An extension contains the (list of ranges of) AS number(s).

b. An AS number and its public key–An organization issues these and signs them using the private key corresponding to the public key in the certificate described in (a). The issuer alternate name in the certificate is the DNS name of the organization. The subject alternate name is the AS number.

c. A router (DNS) name, a router id, an AS number, and the router's public key–An organization issues these and signs them using the private key corresponding to the public key in the certificate described in (a). Both the router id (an IP address) and the AS number are extensions in this certificate, and the binding of three items is a critical aspect of this certificate. The alternate name in the certificate is the DNS name of the router corresponding to the router id.

Certificate Type	Issuer	Subject	Extensions
Root	ICANN	ICANN	All AS #s
Registry	ICANN	Registry	AS #s
AS Owner	Registry (or ICANN)	ISP/DSP or Subscriber	AS #s
AS	ISP/DSP or Subscriber	AS number (in DNS format)	AS #
BGP Speaker	ISP/DSP or Subscriber	BGP Speaker DNS name	AS #, Router ID

Table 2: AS and BGP Speaker PKI Certificate Overview

B. Attestations

An attestation establishes that the subject of the attestation (an AS) is authorized by the issuer to advertise a path to the specified blocks of address space. There are two classes of attestations, address and route, although a single format is employed to represent both. Route attestations are carried in a new type of optional BGP path attribute as part of UPDATE messages:

- *Address attestations.* Here the issuer is the organization that owns the address space and the subject is an AS that may originate it, e.g., the organization's provider. The issuer signs an address attestation using the private key that corresponds to the public key in the certificate (see Figure 1) assigning this address space to the issuer.
- *Route attestations.* Here the subject is a transit AS. A route attestation is signed by the S-BGP speaker (or offline by the management of the AS). The signer uses the private key that corresponds to the public key in the certificate that binds the speaker to the subject AS (see Figure 2).

If an organization has more than one AS, there are separate attestations for each AS rather than just one attestation containing multiple AS numbers. Each AS will have its own set of BGP speakers and its own authentication certificate(s) as well. This applies to both the stub and transit AS cases. Figure 3 summarizes the structure for address and route attestations.

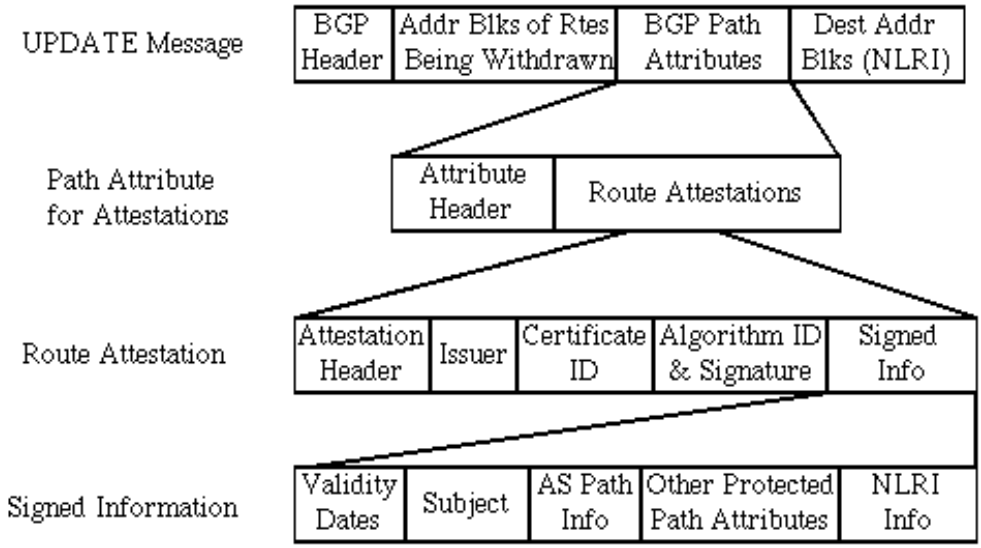


Figure 3: UPDATE Format with Route Attestations

C. Route Validation

Attestations and certificates are used by BGP speakers to validate routes asserted in UPDATE messages, i.e., to verify that the first AS in the route has been authorized to advertise the address block(s) by the address block owner(s), and that each subsequent AS has been authorized to advertise the route for the address block(s) by the preceding AS in the route. To validate a route received from AS_n, AS_{n+1} needs:

- 1 address attestation from each organization owning an address block or blocks in the NLRI
- 1 address allocation certificate from each organization owning an address block or blocks in the NLRI
- 1 route attestation from every S-BGP speaker (or its AS) along the path (AS_n to AS_1), where the route attestation generated and signed by router $_x$ (or AS_x) specifies the NLRI and the AS_PATH from AS_{x+1} through AS_1
- 1 certificate for each S-BGP speaker along the path (AS_n to AS_1) to check the signatures on the route attestations

and, of course, all the relevant CRLs must have been verified.

This means that for each UPDATE, there must be attestations confirming that all the ASes in the BGP UPDATE are authorized to advertise routes to the destination IP address block(s). This includes ASes that are providing third party advertisements for ASes that are not running BGP.

The attestations are not used for checking withdrawn routes because the authorization of the BGP speaker to advertise those routes was verified at the time they were installed into the local routing information base (Loc-RIB). Moreover, if the BGP speaker has lost the authorization to advertise that route, then the route is by definition no longer valid and should be withdrawn.

Use of IPsec on inter-router communication paths prevents an active wiretapper from spoofing route withdrawals, or replaying valid UPDATEs at times when a BGP speaker would not transmit them, e.g., after a route has been withdrawn and prior to advertisement of the same or a different route.

D. Distribution of Countermeasures Information

This section discusses the mechanisms used to distribute certificates, CRLs, and address and route attestations to the relevant devices performing route validation: S-BGP speakers, route servers, etc.

1) Distribution of Certificates, CRLs and Address Attestations: Each S-BGP speaker must have access to the public keys required to validate UPDATEs. For non-leaf S-BGP speakers, this amounts to a full set of certificates encompassing all address space owners, AS owners, and some number of S-BGP speakers (plus the ICANN and registry certificates). An X.509 certificate used in this environment is about 450 bytes long, depending on naming conventions and extensions. In the current (June 1999) Internet environment, there are approximately 5,300 autonomous systems, 44,000 organizations that own address prefixes, and 7,500 BGP speakers. The resulting certificate database comprises about 32 Mbytes, and it can be expected to grow each year as more address prefixes, autonomous systems, and BGP speakers are added. The CRL database associated with these certificates would add to this total, though probably not significantly, since these certificates are issued to organizations and devices, not people

At first glance it might seem appropriate to transmit certificates as part of each UPDATE. This would ensure that each receiving BGP speaker would receive all the data needed to validate the route attestations in an UPDATE, and it would be easy for each BGP speaker to include its own certificate as part of the forwarding process. However, this would be very wasteful of bandwidth, as each BGP speaker would receive many redundant copies of certificates.[7] More importantly, this approach is infeasible, because BGP UPDATEs are limited in length to 4096 bytes and thus are too small to carry the necessary certificates for most UPDATEs.[8] The introduction of a new type of BGP message for transmission of certificates (and CRLs) could address the packet size problem, but would still tend to be very wasteful of bandwidth and would not be backward compatible.

Instead, this architecture uses out-of-band distribution of certificates and CRLs to all S-BGP speakers. This is an attractive approach to the distribution problem for several reasons. This database is relatively static and thus a good candidate for caching and incremental update. Moreover, the certificates can be validated (and processed against CRLs) and reduced to a more compact format by ISPs/DSPs prior to distribution to S-BGP speakers. (Only the public key, subject, and selected extensions need be retained.) This avoids the need for each speaker to perform this processing (entailing tens of thousands of signature validations), and it saves both bandwidth and storage space. Although memory is inexpensive, most currently deployed commercial routers do not possess sufficient memory to store all of these certificates so either additional memory or auxiliary systems will be needed, even with preprocessing.

To address the distribution problem we make use of two tiers of repositories from which one can download the entire certificate and CRL database. The top tier consists of several replicated, easy to access storage sites, e.g., the NAP route servers. The second tier of repositories are operated by the ISPs/DSPs, to provide local access for

the S-BGP speakers within each AS. (Per-AS repositories avoid the dependency loop that would occur if one required inter-AS routing in order to access this database.) Bulk transfer of the whole certificate or CRL database, from the first to second tier repositories, can be effected via FTP or TFTP; since certificates and CRLs are signed and carry validity interval information, there is no need for additional integrity mechanisms in the transfer. The second tier repositories also will query top tier repositories to get new CRLs, based on the CRL NEXT UPDATE field, and to get new certificates, based on certificate expiration, e.g., using the Lightweight Directory Access Protocol (LDAP). Retrieval of newly issued certificates (in between downloads of the complete database) for newly created ASes, organizations with new address space ownership, etc., could be effect in the same manner, from daily incremental update files.

This allows an ISP/DSP to operate in an anticipatory fashion, retrieving certificates before it needs them. Each second tier repository will validate all of the certificates and CRLs it retrieves, and produce a more compact (locally signed) database ready for consumption by the S-BGP speakers within its administrative purview. If deemed necessary, the top tier repositories also can push CRLs issued prior to scheduled dates.

The same analysis applies to address space attestations. Each address attestation requires about 110 bytes, and this amounts to approximately 4 Mbytes. Carriage of these data items in UPDATEs would usually be redundant and thus is more effectively handled via the same, out-of-band distribution mechanism. Here too, this distribution model allows pre-processing by ISP/DSP NOCs, further eliminating significant signature validation overhead (there would be roughly 44,000 such attestations currently). This model also simplifies revocation of address attestations, i.e., an address space owner can issue a CRL-like, signed data structure and include it in the database for downloading and pre-processing by each ISP/DSP. Preprocessing reduces the total space required for the certificate and address attestation databases from about 36 Mbytes to about 11 Mbytes.

2) Distribution of Route Attestations: Route attestations are distributed with BGP UPDATEs in a newly defined, optional, transitive path attribute. This approach requires BGP speakers to be upgraded to a BGP release with these countermeasures. When an S-BGP speaker opens a BGP session with a peer, transmitting the advertisable portion of its routing information database via UPDATEs, relevant route attestations are sent with each UPDATE. These attestations employ a compact encoding scheme to help ensure that they fit within the BGP packet size limits, even when route or address aggregation is employed. (We preserve S-BGP security guarantees in the face of aggregation by explicitly including signed attribute data that otherwise would be lost when aggregation occurs.) The S-BGP speaker receiving an UPDATE[9] caches the associated attestations with the route in its routing information database. Each BGP speaker generates route attestations based on receipt of UPDATEs from its neighbors, as described in Section IV.B, thus in-band distribution is appropriate. As noted below in Section VI.B, the bandwidth required to support in-band distribution of route attestations is negligible (compared to user traffic).

E. IPsec and Router Authentication

BGP is transported over TCP and thus is protected against misordered, lost, or replayed packets, to the extent that the TCP sequence number management facility is secure. BGP-4 provides a means for carrying authentication information in BGP messages, but there is no prescribed key management scheme and there is no facility for sequence numbering of BGP messages, hence this facility is not employed here. Instead, we use the Encapsulating Security Payload (ESP) protocol (with NULL encryption), from the IP security protocol suite (IPsec) [8,9,10,11,12] to provide authentication, data integrity, and anti-replay on a point-to-point basis, i.e., between BGP speakers. The Internet Key Exchange (IKE) protocol is used for key management services in support of ESP. (Although the Authentication Heahder protocol from the IPsec suite could be used here, it is less efficient and thus was not selected.) The PKI established for router and AS authentication provides the necessary certificates (see Section III.A.2).

F. Other Issues

BGP Path Attributes are being standardized to support Communities (to support policy)[13], Confederations (used to transparently split a large AS into several smaller ASes to reduce the n^2 squared peering requirements)[14], and to support additional protocols (IPV6, MPLS, and multicast[15]). The Route Attestation mechanism is designed to provide protection for these and other new path attributes, in the same manner in which it protects the current path attributes.

IV. How These Countermeasures Address BGP Vulnerabilities

This section describes how the proposed countermeasures reduce the vulnerability of BGP to the attacks described earlier and provide much of the functionality necessary for ensuring the correct operation of BGP.

A. Certificates

The certificates described above are used to enable verification of:

- An AS's authorization to "advertise" a block of addresses–The certificates from III.A.1 are used to verify that an AS is authorized to "advertise" a block of addresses. Specifically, the signature on an address attestation must be verifiable using the public key in a certificate containing the address block(s) that include the address block(s) in the address attestation.
- An organization's ownership of an AS number–The certificates from III.A.2.a are used to verify that an AS has been assigned to the holder of a particular public key, i.e., an ISP, DSP, or subscriber organization. They are used to validate III.A.2.b or III.A.2.c certificates through the AS number linkage.
- An AS's identity–The certificates from III.A.2.b (or certificate data pre-processed by a NOC and distributed to routers) are used to verify the signature of an AS on a route attestation.
- A BGP speaker's identity and its association with an AS–The certificates from III.A.2.c are used to verify the signature of a speaker on a route attestation, and in conjunction with III.A.2.a to make sure that the speaker is authorized to act on behalf of the AS.
- Identity and authorization of a BGP peer–The certificates from III.A.2.c are used by the BGP speakers when establishing peering sessions, to authenticate each other.

B. Address and Route Attestations

These two countermeasures support validation of the address prefixes and path information in an UPDATE.

Address attestations protect BGP against misbehaving BGP speakers that originate or distribute erroneous UPDATEs and BGP speakers whose advertisable destination addresses have been misconfigured. Whenever an S-BGP speaker advertises itself as the starting point of a route for some address prefix, other S-BGP speakers will verify that the AS represented by that speaker is the subject of an address attestation signed by the owner of the address prefix. Since only the organization that owns the prefix can sign such an attestation, no S-BGP speaker can falsify such an advertisement.

Each S-BGP UPDATE will include a set of route attestations, one per AS listed in the UPDATE, each of which is added to the UPDATE as it propagates among ASes. A route attestation indicates that the signing BGP speaker is authorized to advertise to the neighbor the route constructed thus far, by the organization owning the AS (in which the speaker resides). The route attestation is digitally signed by the S-BGP speaker distributing the UPDATE. It includes the identification of the S-BGP speaker's certificate issued by the owner of the AS, the destination addresses in the route, the list of identifiers of ASes in the route, the identifier of the AS to which the UPDATE is directed, a maximum lifetime, and other transitive data requiring protection. Each recipient of an UPDATE verifies the route attestations contained within it before deciding whether to accept and distribute the UPDATE. Route attestations protect BGP against misbehaving BGP speakers that distribute erroneous UPDATEs, and against misconfigured local routing policies.

C. IPsec

IPsec (specifically ESP and IKE) provides the security services needed by the receiving BGP speaker to verify message integrity, the identity of the sender, and the fact that it (the receiver) is the intended recipient of every message. Although the attestations in UPDATE messages protect against a wide range of active wiretap attacks, use of ESP provides protection for all BGP traffic, prevents replay of messages across a link, and protects TCP against various forms of attack, including SYN flooding

and spoofed RSTs (resets).

V. Residual Vulnerabilities

The S-BGP system (address attestations, route attestations, PKIs, and IPsec for all BGP messages) addresses many of the vulnerabilities of BGP-4. Nevertheless, there exist vulnerabilities that are not eliminated by this system, including the following:

- Suppression of BGP messages by a misbehaving BGP speaker is not addressed. Use of IPsec (and TCP) will detect active wiretap attacks that result in lost or reordered BGP packets. However, a compromised BGP speaker can elect to not transmit BGP messages, even when local policy would call for such transmission, e.g., for route withdrawal. This is undetectable by the proposed countermeasures, although coordinated network monitoring might be able to detect such misbehavior. The substantial flexibility afforded by local policies in BGP appears to preclude countering this vulnerability if such policies are to remain private to ASes (as allowed by the BGP specification and as currently practiced).
- Passive wiretapping to discover network connectivity information is not addressed. These attacks could be countered by enabling the confidentiality feature of ESP, if the risk exceeds the cost to encrypt and decrypt UPDATE messages.
- A BGP speaker may reassert a route that was withdrawn earlier, even if the route has not been re-advertised. This vulnerability exists because BGP UPDATEs do not carry sequence numbers or timestamps that could be used to determine the currency of UPDATEs. However, route attestations do expire, so there is a limit on how long an old attestation can be used for such purposes. The possibility also exists to add a CRL-like function for route attestation revocation, a possibility that will be explored later in our work.
- Verification that the BGP peers that exchanged the UPDATE, correctly applied BGP rules, local policies, etc., is not addressed. As above, BGP affords speakers considerable latitude with regard to local policy and ASes do not usually make public their local routing policies, hence it appears difficult to counter such problems. S-BGP restricts malicious behavior to the set of actions for which the speaker (or AS) is authorized, based on externally verifiable constraints.

VI. Performance and Operational Issues

In developing the S-BGP architecture, we have paid close attention to the performance and operational impact of the proposed countermeasures. Previous work in the area of routing security has often focused almost exclusively on the costs of generating and validating digital signatures. While such costs are an important factor, our analysis suggests that the bandwidth and storage requirements associated with signatures, certificates and CRLs are much bigger problems. The following analysis is based on examination of actual Internet routing data plus simulation of the effects of S-BGP countermeasures.

A. Processing

The computation burden for signature generation and validation appears to be tractable in this proposed architecture. We have selected the Digital Signature Algorithm (DSA) to minimize the size of the signatures, specifically for route attestations.[10] DSA yields only a 40-byte signature, vs. the 128-byte signature typical for RSA (using 1024-bit keys). DSA also allows for pre-computation, which permits lower latency in signature generation by S-BGP speakers. Other (S-BGP-specific) techniques (described below) significantly reduce the need for signature validation operations, so the processing asymmetry exhibited by DSA is not a concern here.

The rate at which new UPDATEs are created is not so great that signature generation and validation of route attestations is expected to pose a bottleneck. A BGP speaker at a NAP, peering with about 30 other BGP speakers, receives an average total of about .5 UPDATEs per second. Each route contains an average of 3.6 ASes, and there is one route attestation per AS, yielding a rate of about 1.8 signature validations per second. In contrast, each UPDATE generated by an S-BGP speaker requires just one signature (added by the speaker), thus the signature generation rate is less than one third (1/3.6) of this value. Peak load figures may be about a factor of ten greater, yielding a peak load of about 18 signatures per second. However, analysis of data from NAPs shows that 50% or more of all UPDATEs repeat routes already known to a BGP speaker (e.g., due to link flapping). Thus caching just one route for each address prefix enables a speaker to avoid the need to validate signatures on the vast majority of UPDATEs, reducing the peak load to about 9 signatures per second, a tolerably low computational burden.[11]

Upon initialization (reboot), a BGP speaker receives complete routing table updates from each peer. This means that an S-BGP speaker will receive a very large number of route attestations requiring validation, i.e., about 220,000 per peer, in a short time interval. This sort of initialization transient would be unacceptable, even though reboots and installation of new speakers is relatively infrequent.[12] To avoid this problem, we propose the addition of non-volatile storage for validated route attestations, to preserve the cache across reboots. When installing a new BGP speaker in an AS, a NOC could dump the cache from another speaker in the AS and reload it into the new BGP speaker, to seed the cache in the newly installed device.

B. Transmission Bandwidth

The transmission of countermeasures data in UPDATEs increases the size of these messages to approximately 450 bytes (based on an average of 3.6 route attestations per path), for a typical UPDATE that previously required only 63 bytes. This represents a significant percentage increase in BGP overhead (over 700%), but the transmission of UPDATEs represents a very, very small amount of data relative to subscriber traffic. As noted above, even a speaker at a NAP sees an average rate of less than one UPDATE per second, and such speakers are connected via 100 Mb/s interfaces today, with plans to transition to multi-Gb/s interfaces in the future.

Downloading the certificate, CRL, and address attestation databases contributes an insignificant increment to this overhead. Full database transmission, from a top tier to a second tier repository entails about a 36 Mbyte file transfer for each ISP/DSP. Even if performed on a daily basis, this traffic is swamped by subscriber file transfers. Transfers from a second tier repository to each BGP speaker in its AS are smaller, about 11 Mbytes (due to certificate extraction). Here too, even if performed daily, this would be but a drop in the ocean of subscriber traffic, and use of incremental transfers is a more likely scenario. So the impact on utilization of Internet bandwidth due to transmission of all of the countermeasures data is minimal. Also, the speed of inter-router circuits continues to increasing substantially, further minimizing the impact of transmission of additional control traffic.

C. Storage/Memory

UPDATEs received from neighbors are held by a BGP router in Routing Information Bases (RIBs) and used to generate new UPDATEs for transmission to other BGP routers. The additional memory required for pre-processed certificates and address attestations amounts to about 11 Mbytes, as noted earlier. The space required for route attestations is about 26 Mbytes per peer, a modest amount for a typical speaker with 2 or 3 peers, but a significant amount of storage for speakers at NAPs, where each speaker has about 30 peers. This amounts to a large but feasible amount of additional RAM by current standards, where a high end workstation can be configured with hundreds of megabytes of RAM. The routers that act as BGP speakers at NAPs are large, very expensive devices. However the storage capacity of the routers currently used by ISPs/DSPs would not permit storage of S-BGP UPDATEs in their RIBs, if S-BGP were deployed. Thus additional, non-volatile storage is needed in BGP speakers to support these databases. It may be feasible to significantly reduce the storage required here, since the routes (and thus route attestations) received from different peers tend to exhibit significant overlap in their suffixes.

D. Deployment and Transition Issues

Deploying S-BGP raises a number of other issues:

- *Adoption of S-BGP by several groups.* The ISPs, DSPs, and subscriber organizations running BGP will need to cooperate in the generation and distribution of attestations. The first tier ISPs (those connected to the NAPs) must implement the S-BGP security mechanisms in order offer significant benefit to the Internet community. (Lower level ISPs, DSPs, and subscriber organizations will need to implement the S-BGP security mechanisms only if the expense can be justified.) ICANN and registration authorities will need to expand their operational procedures to support generation of address space and AS number delegation certificates. Finally, router vendors need to provide additional storage in next generation products, or offer ancillary devices for use with existing router products, and revise BGP software to support S-BGP.

- *S-BGP interaction with other exterior and interior routing protocols.* External routes received from external peers need to be redistributed within the AS in order to maintain a consistent and stable view of the exterior routes across the AS. Interior routing protocols will not propagate S-BGP attestations, but if each border S-BGP speaker maintains an iBGP[13] connection with all other transit and border routers within the AS, this problem will be averted.
- *BGP-4 to S-BGP Transition.* The route attestation path attribute is optional for both external and internal BGP exchanges. This allows extensive regression testing before deploying S-BGP on production equipment.

VII. Subsequent Other Work

Since we began this work, there have been several other efforts towards securing BGP. These include an assessment of BGP vulnerabilities[14] and several approaches to addressing some of these vulnerabilities.

- *TCP/MD5*[16]. This defines a TCP option for carrying an MD5 digest. This mechanism offers data origin authentication and data integrity, on a point-to-point basis. It protects the TCP connection used to transport BGP traffic from spoofing attacks and connection hijacking. Lack of an automated key distribution protocol complicates management and encourages overly long term use of symmetric keys. Moreover, because it fails to protect against any attacks that subvert routers or the management of routers, its overall security efficacy is quite limited.
- *NLRI Origin Verification*[17]. This mechanism proposes adding an address prefix delegation tree to Secure DNS [18]. For each prefix that has been authorized for use, a new resource record specifies the number of the AS that is authorized to originate that prefix. This mechanism does not address route authorization, nor does the proposal describe in detail how this data would be distributed to BGP speakers. It does represent an alternative format and database option for the address attestations developed in our architecture.
- *Routing Policy System Security*[19]. This approach places authorization information into a small number of databases (Internet Routing Registries) accessible by routers. The information is placed into a database by the organization responsible for the authorization, using some secure access method, e.g., SSL. It proposes that BGP speakers compare the routes that they receive to the routes listed in the database, rejecting any routes not found. No changes to BGP are required. This approach does not appear to be very dynamic, although details of how, and how often, the registries are to be accessed are omitted. Use of such registries would require ISPs/DSPs to publicize what is now local policy information, which most have refused to do. Moreover, the routes stored in the registries are not signed, so attacks against these databases, including malicious or benign errors by an ISP/DSP, could compromise security.
- *Hash Chain Signatures*[20]. This work describes two protocols, COSP and IOSP, based on hash chain signatures, that offer very rapid signature generation and validation in a routing protocol context. COSP is not applicable because it requires signing messages at fixed time intervals, whereas BGP generates UPDATEs on demand, as a result of topology changes; IOSP is not applicable because its efficiency depends on each router receiving essentially all routing updates, which is not characteristic of the operation of BGP.

VIII. Future Work

Deploying this technology into the Internet will require the creation of the supporting Public Key Infrastructures, rooted at the ICANN, and convincing ICANN, the Internet Registries, the major ISPs, the owners of IP address blocks and the router vendors of the benefits and supportability of these security mechanisms. To facilitate this transfer of technology, we are currently building a proof-of-concept prototype of S-BGP to demonstrate the viability and feasibility of deploying this technology into the Internet. This involves deploying the technology in DARPA's CAIRN testbed and running experiments with current Internet BGP traffic and Merit's historical BGP data.

The results of this work will include the prototype software (in GateD) and a report on the results of these tests, e.g., analysis of S-BGP performance and overhead costs with various optimizations. We hope to be able to pursue additional technology transfer activities to facilitate adoption of S-BGP.

IX. Summary

BGP is a critical component of the Internet's routing infrastructure and highly vulnerable to a variety of attacks. The S-BGP countermeasures use IPsec, Public Key Infrastructure (PKI) technology, and a new BGP path attribute ("attestations") to ensure the authenticity and integrity of BGP communication on a point-to-point basis, and to validate BGP routing UPDATEs on a source to (multicast) destination basis. These enhancements will allow Internet Service Providers (ISPs) and their customers to verify that:

- reachability information they receive is from an authentic and authorized BGP peering relationship and has not been modified without authorization;
- the authorization of an organization to claim ownership of a block of IP addresses (a (sub)network) is substantiated by a chain of authorizations rooted at the Internet Corporation for Assigned Names and Numbers;
- an originating AS is authorized to advertise reachability to a block of IP addresses by the organization owning that address block;
- the ASes that processed the routing information en-route from the originating AS are not, either through mis-configuration, internal error, or compromise, advertising reachability information that is inconsistent with nominal topology;
- each AS, and its BGP speakers, that advertise a given route are identifiable and authorized to participate in global Internet routing, by a chain of authorizations rooted at the ICANN.

Acknowledgement

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 provided later support, under Dr. Hilarie Orman, enabling us to refine, implement and test the design. S-BGP has 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 and the associated evaluations of other approaches and performance and operational issues. The authors would also like to thank Michelle Casagni, for her work on the performance analysis and Joanne Mikkelson, Dennis Rockwell, and Nicholas Shectman for their efforts during the ongoing implementation and experimentation phase.

References

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

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

[3] C. Villamizar, R. Chandra, R. Govindan., "BGP Route Flap Damping," RFC 2439, November 1998.

[4] Smith, B.R. and Garcia-Luna-Aceves, J.J., "Securing the Border Gateway Routing Protocol," Proceedings of Global Internet '96, November 1996.

[5] Smith, B.R, Murphy, S., and Garcia-Luna-Aceves, J.J., "Securing Distance-Vector Routing Protocols," Symposium on Network and Distributed System Security, February 1997.

[6] Kumar, B., "Integration of Security in Network Routing Protocols," ACM SIGSAC Review, vol.11, no.2, Spring 1993.

[7] Murphy, S., panel presentation (no text) on "Security Architecture for the Internet Infrastructure," Symposium on Network and Distributed System Security, April 1995.

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

[9] R. Glenn & S. Kent, "The NULL Encryption Algorithm and its Use with IPsec," RFC 2410, November 1998.

[10] S. Kent, R. Atkinson, "IP Encapsulating Security Payload (ESP)," RFC 2406, November 1998.

[11] D. Maughan, M. Schertler, M. Schneider, J. Turner, "Internet Security Association and Key Management Protocol (ISAKMP)," RFC 2408, November 1998.

[12] D. Harkins, D. Carrel, "The Internet Key Exchange (IKE)," RFC 2406, November 1998.

[13] R. Chandra, P. Traina, T. Li, "BGP Communities Attribute", RFC 1997, August 1996.

[14] P. Traina, "Autonomous System Confederations for BGP," RFC 1965, June 1996.

[15] T. Bates, R. Chandra, D. Katz, Y. Rekhter, "Multiprotocol Extensions for BGP-4," RFC 2283, February 1998.

[16] A. Heffernan, "Protection of BGP Sessions via the TCP MD5 Signature Option," RFC 2385, August 1998.

[17] T. Bates, R. Bush, T. Li, Y. Rekhter, "DNS-based NLRI origin AS verification in BGP," presentation at NANOG 12, February 1998, http://www.nanog.org/mtg-9802.

[18] D. Eastlake, 3rd, C. Kaufman, "Domain Name System Security Extensions," RFC 2065, January 1997.

[19] C. Alaettinoglu, T. Bates, E. Gerich, D. Karrenberg, D. Meyer, M. Terpstra, C. Villamizar, "Routing Policy Specification Language (RPSL)," RFC 2280, January 1998.

[20] K. Zhang, "Efficient Protocols for Signing Routing Messages," Network and Distributed System Security Symposium, March 1998.

[21] R. Perlman, "Network Layer Protocols With Byzantine Robustness," MIT/LCS/TR-429, October, 1988.

Footnotes

1. A prefix specifies an IP address block, and consists of a count of the most significant bits in an IP address, and the value of those bits.
2. In a discussion with David Mills, an architect of the NSFNET, he confirmed that route flapping has been a characteristic of the Internet since the mid 1980's.
3. Technically, two certification hierarchies are employed, but they might employ common procedures etc.and thus be considered a single PKI.
4. For historical reasons, the chain of issuance described above was sometimes short circuited. A saubscriber (or DSP) who has an address-space assignment that has bypassed the normal allocation procedure, who has changed DSP's/ISP's and retained the originally assigned address, must also be certified, even if not a BGP user.
5. If an Organization acquires additional address blocks, a new certificate is issued to reflect the increased scope of ownership.
6. One could place the address block in an X.509 attribute certificate, linked to an X.509 public key certificate, in a more elegant approach to representing this data. However, because the number of certificates involved is so great, and because attribute certificates are not yet widely suppoted, we have chosen to add the address block information as a private, v3 extension to a public key certificate.
7. The redundancy arises from several factors. A BGP speaker tends to receive routes to the same destination, via each interface, with considerable overlap of ASes in each route. Withdrawl and later re-advertisement of the same routes via the same interface results in additional redundancy.
8. Many UPDATE's contain routes for multiple address blocks and some routes contain many AS numbers, and each requires its own certificate.
9. We have not yet determined if a mechanism for revoking route attestations is required, or if a modest attestation lifetime will suffice.
10. Since certificates, CRL's, and address attestations are stripped of signatures in preprocessing, the choice of signature algorithm is not so critical for these data structures.
11. For example, SSLeay software can perform about 40 1024-bit DSA signatures per second on a 450 MHz Pentium II processor.
12. Because of the important service they provide, BGP speakers are usually afforded UPS protection and new software is deployed only after extensive testing, to minimize the likelihood of crashes.
13. The acronym "iBGP" denotes usr of intra-AS use of BGP, in contrasst to the common inter-AS use of BGP, sometimes referred to as eBGP.
14. S. Murphy's "BGP Security Analysis surveys this topic. At the time this paper was prepared, the most recent version was , June 1999.

Manuscript received February 1, 1999; revised November 11, 1999. This work was supported by NSA in 1997, and by DARPA.

The authors are with BBN Technologies, Cambridge, MA 02138 USA.

Publisher Item Identifier S 0733-8716(00)01519-5.