Network Working Group
Internet Draft
Expiration Date: October 2004
File Name: draft-ng-sobgp-bgp-extensions-02.txt

James Ng
(editor)
Cisco Systems
April 2004

Extensions to BGP to Support Secure Origin BGP (soBGP)
draft-ng-sobgp-bgp-extensions-02.txt

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

Internet Drafts are working documents of the Internet Engineering Task Force (IETF), its Areas, and its Working Groups. Note that other groups may also distribute working documents as Internet Drafts.

Internet Drafts are draft documents valid for a maximum of six months. Internet Drafts may be updated, replaced, or obsoleted by other documents at any time. It is not appropriate to use Internet Drafts as reference material or to cite them other than as a "working draft" or "work in progress".

The list of current Internet-Drafts can be accessed at http//www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http//www.ietf.org/shadow.html.

1. Contributors

A large number of people contributed to or provided valuable feedback on this document; we've tried to include all of them here (in no particular order), but might have missed a few: Russ White, Alvaro Retana, Dave Cook, John Scudder, David Ward, Martin Djernaes, Chris Lonvick, Brian Weiss, Tim Gage, Scott Fanning, Barry Friedman, Jim Duncan, Yi Yang, Robert Adams, Tony Tauber, Iljitsch van Beijnum, and Jonathan Natale.

2. Abstract

There is a great deal of concern over the security of routing systems within the Internet, particularly in relation to the Border Gateway Protocol [BGP], which is used to provide routing information between autonomous systems. This document proposes a system where the origin of any advertisement within BGP can be verified and authenticated, preventing the advertisement of prefix blocks by unauthorized networks, and verifying that the final destination in the path is actually within the autonomous system to which the packets are being routed.

This document does not:

- Attempt to provide information on how such a security system could or should be deployed; readers are referenced to [SOBGP-DEPLOY] for this discussion.

- Attempt to determine what sorts of keys should be used within such a system, nor how any sort of trust relationship can or should be built between the entities cooperating within the routing system. These are considered in [SOBGP-CERTIFICATE].

- Attempt to analyse the performance, memory utilization, or other impacts on devices running this protocol; these are addressed in [SOBGP-DEPLOY].

- Attempt to analyse the security protection provided by the proposed security system.

This document primarily focuses on extensions to the BGP protocol itself to support such a security system. This document is to solve several vulnerabilities within a BGP routing system given the following constraints and assumptions:

- Any signalling which must take place to provide security should be specified in as flexible a way as possible. For instance, any certificates exchanged may be solely contained with the BGP protocol, or through some other transport/distribution mechanism. [SOBGP-DEPLOY] contains more discussion on this issue.

- The proposed solution should be incrementally deployable, and require minimal changes in a network to support it (software changes only along the path where support is required, for instance).

- The proposed solution should not rely on the routing system it is securing to reach information necessary to secure the routing

system (reliance on an IGP to reach destinations within a routing domain is acceptable, but reliance on BGP to reach destinations outside the routing domain is unacceptable).

- The operator may configure various levels of trust, preferring faster convergence over security validtion, or more immediate and stronger security validation over convergence times.

- All routing information received from peering relationships within an autonomous system is implicitly trusted.

- No current optimizations in implementations of the BGP protocol should be affected, if possible.

It's important to note that any given security system is a tradeoff between several factors, including the actual security provided and the difficulty entailed in deploying the system; different solutions may provide different levels of protection between these, leaving some parts of the problem unsolved.

This proposal does not protect against:

- Attacks through all forms of autonomous system spoofing. This proposal can be subverted if an attacker is able to masquerade as an autonomous system under some network conditions. It is assumed that management of inter-AS connections and policy implementation can resolve these possible attacks.

- Path Authentication. While this system allows the path any given update advertises to be checked against paths known to be valid, it does not ensure this particular update provides the best path. It is observed that there are many possible valid paths within a BGP routing system, and BGP itself is designed to find the correct path among the possible paths.

- Authentication of path attributes, such as the community and MED attributes passed with a prefix.

The IETF has been notified of intellectual property rights claimed in regard to some or all of the specification contained in this document. For more information consult the online list of claimed rights.

3. Definitions

o Entity: A participant in the internetwork routing system.

4. The Security Message

This document proposes a new message type, the SECURITY message, which is to be used for carrying security information within the BGP protocol. The SECURITY message is type 6. The SECURITY message is used to transport the certificates described in [SOBGP-CERTIFICATE].

4.1. Negotiating Security Capability

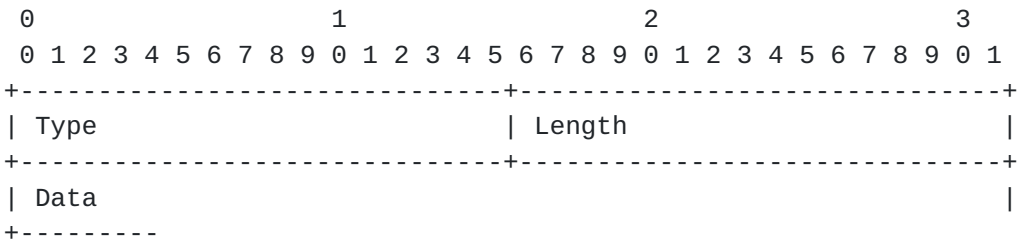
The ability to exchange SECURITY messages shall be negotiated at session startup, as described in [CAPABILITY]. The capability code is <to be assigned by IANA>. A pair of BGP speakers MAY negotiate to send messages in the SECURITY message type only (exchange no NLRI or forwarding information).

If two BGP speakers have negotiated to exchange SECURITY messages, they should exchange the certificates contained in their local databases.

If two speakers have negotiated to exchange SECURITY messages and NLRI messages (as described in [BGP]), no NLRI or SECURITY information SHOULD be transmitted by a speaker other than a SECURITY Option message until it receives a SECURITY Option message from its peer.

4.2. The Security Message Format

The SECURITY message is formatted as described in [BGP], with a type code of 6. Within each message is a series of TLVs, or security message blocks, formatted as:



o Type: A two octet unsigned integer describing the type of information contained within the data field.

- Length: A two octet unsigned integer describing the length of the data field, in octets.

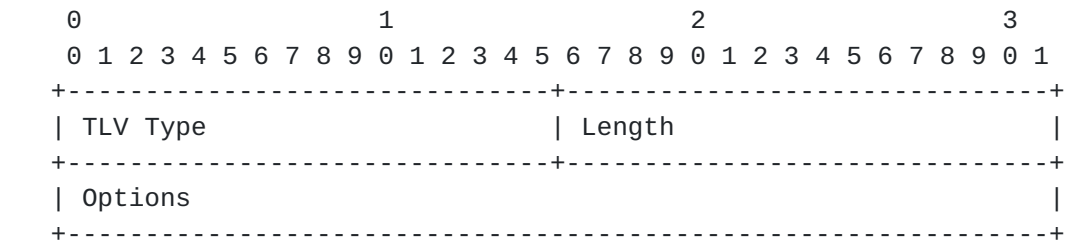
- Data: The data, as described within this and other documents which describe information to be carried within the SECURITY message type.

Two TLVs are currently defined within the SECURITY message. Further TLVs are defined for carrying certificates in [SOBGP-CERTIFICATE].

4.2.1. The SECURITY Option TLV

The SECURITY Option TLV provides a way for exchanging speakers to inform their peers about local configurations which may pertain to the peering session. SECURITY Option TLVs are encapsulated within a TLV Type 1, and transmitted within the SECURITY message type.

If SECURITY Option TLVs are transmitted, they MUST be transmitted before the transmission of any other SECURITY data.



- TLV type: (2 octets), 1 (0x0001)

- Length: (2 octets), set to 2

- Options: (4 octets), a bitfield, described below

The options field is a 32 bit bitfield, allowing up to 32 different options to be specified.

- Bit 0: If set, indicates that SECURITY information should be sent before NLRI information on this session; if cleared, indicates that NLRI information should be sent before SECURITY information.

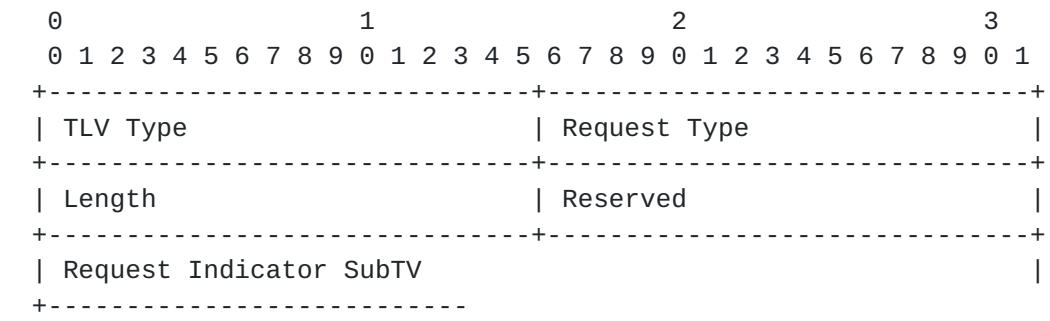
- Bit 1: If set, indicates that this peer will only transmit validated certificates of any type along this session (valid only on iBGP sessions).

o Bit 2: If set, indicates that this peer will only accept validated certificates of any type along this session (valid only on iBGP sessions).

Bit 0 in the option field allows the operator to configure the local device so it receives all prefixes first, decreasing convergence to the minimum time, or receives all SECURITY information first, allowing all prefixes to be validated before they are installed.

Bits 1 and 2 allow peers along an iBGP session to trust the certifications they receive without validating them. If bit 1 is set on the transmitting peer, and bit 2 is set on the receiving peer, the receiving peer may accept all received certificates from the transmitting peer as already validated.

4.2.2. The Request TLV



o TLV type: (2 octets), 2

o Request Type: (2 octets), treated as an unsigned integer indicating the type of the type of information requested.

o Length: (2 octets), set to the total length of the request in octets.

o Reserved: (2 octets), set to 0x0000.

o Request Indicator: The information indicated by the request type bit field.

The Request Type field indicates the type of certificates requested. Three request types are defined in this document.

1 Any certificate matching the Request Indicator are requested.

2 EntityCerts matching the Request Indicator are requested.

3 ASPolicyCerts matching the Request Indicator are requested.

4 PrefixPolicyCerts matching the Request Indicator are requested.

Request indicator SubTVs restrict the set of certificates returned; there may be one or more request indicator SubTVs included in a request. Each SubTV consists of a two octet type field, treated as an unsigned integer, and a fixed length field containing the request indicator.

- Type 1: A four octet origin/authorized AS Number; two octet AS numbers shall be right justified within this field (placed in the two least significant octets).
- Type 2: A four octet signer/authorizer AS Number; two octet AS numbers shall be right justified within this field (placed in the two least significant octets).
- Type 3: A four octet IPv4 address is included in the request indicator.
- Type 4: A sixteen octet IPv6 address is included in the request indicator.
- Type 5: An eight octet starting serial number is included in the request indicator.
- Type 6: An eight octet ending serial number is included in the request indicator.

5. Receiving and Processing SECURITY messages

The section sbelow describe the receipt and processing of SECURITY messages.

5.1. The Level of Certificate Processing

Each section below describes the processing of SECURITY messages based on the way in which soBGP is deployed on the BGP speaker. For more information on deployment options, see [SOBGP-DEPLOY].

5.1.1. Full Local Certificate Processing

If a BGP speaker is fully processing certificates locally, for each received certificate it must:

- The unique identifiers (described for each certificate type in [SOBGP-CERTIFICATE] of any certificate received MUST be compared to the unique identifiers of certificates already held in local databases. Any certificate which matches a certificate already held in a local database MUST be discarded.
- Certificates received from untrusted peers and certificates received without the VALIDATED bit set SHOULD be placed in a local certificate database, and processed according to [SOBGP-CERTIFICATE].
- Certificates received from iBGP peers which have negotiated trusted certificate exchange SHOULD placed in a local certificate database, and processed as though they were all sucessfully validated according to the process described in [SOBGP-CERTIFICATE].
- All certificates received and placed in the local certificate database, and marked as validated, SHOULD be readvertised to all iBGP peers which have a trusted peering relationship, marked as VALIDATED. Signatures SHOULD be stripped from any certificate advertised to an iBGP peer with a trusted peering relationship.
- All certificates received and placed in the local certificate database, and not marked as validated, should be readvertised to trusted iBGP peers; they MUST not be marked as VALIDATED.
- All certificates received and placed in the local certificate database should be readvertised to all BGP peers which have negotiated the untrusted exchange of soBGP certificates.

5.1.2. No Local Cryptographic Processing

If a BGP speaker is accepting certificates only from trusted peering relationships, for each received certificate, it must:

- The unique identifiers (described for each certificate type in [SOBGP-CERTIFICATE] of any certificate received MUST be compared to the unique identifiers of certificates already held in local databases. Any certificate which matches a certificate already

held in a local database MUST be discarded.

- Certificates received from iBGP peers which have negotiated trusted certificate exchange and marked as VALIDATED must be placed in a local certificate database, and processed as though they were all sucessfully validated according to the process described in [SOBGP-CERTIFICATE].

- Certificates received from any BGP peer which has not negotiated trusted certificate exchange, and any received certificate which is not marked as VALIDATED, should be readvertised to all BGP peers which have not negotiated trusted certificate exchange. These certificates MUST be discarded; they must not be stored locally.

5.1.3. No Local Certificate Processing

If a BGP speaker is validating routing information through direct interaction with another device performing soBGP processing, and is not processing or storing certificates locally, for each soBGP certificate received, it must readvertise all certificates received from any BGP peer to all other BGP peers which have negotiated soBGP certificate exchange through the SECURITY message. The BGP speaker MUST NOT mark the readvertised certificates as VALIDATED.

5.2. Filtering of Certificates

A BGP speaker may, for reasons of policy, filter soBGP certificates received from a peer.

- If a BGP speaker is part of a transit AS, it SHOULD NOT filter soBGP certificates.

- A BGP speaker MAY discard soBGP certificates which describe the authorization of address space which is being filtered out of the local routing information.

5.3. Receiving and Processing Requests

If a device receives a Request TLV, as described in the section "The Security Message," above, it should:

- Examine the request to ensure it is logically consistent. For instance, requesting an Entitycert based on an IPv4 address

range is not logically consistent, since these certificates only contain an AS and a Signer AS.

- If the request is not logically consistent, discard it.
- If the request is logically consistent, examine its local databases, and transmit the certificates requested which fulfill the conditions supplied in the request indicator SubTVs.
- If more than one of the same request indicator is included in a request message, they shall be treated as an OR condition; if any of the conditions match, the certificate shall match the set.

6. Operation and General Requirements

This section discusses the processing of routing updates, validation of the AS_PATH contained in routing updates, general BGP requirements for autonomous systems running soBGP, and requirements for BGP sessions exchanging soBGP certificates.

6.1. Receiving and Validating Prefixes

A local database of authorized prefix/origin AS/policy triples is built using the techniques described in [SOBGP-CERTIFICATE], and used to validate updates as they are received (or updates which exist in the AdjRIB-In), as described in this section. This local database is termed the "authorization database."

Since one of the goals of soBGP, stated at the beginning of this document, is to impact the performance of the BGP protocol as little as possible, there is no requirement to perform the steps outlined in this section at any particular time in the processing of received BGP updates. The validation of routes received may occur as these routes are received, before they are placed in the Adj-RIB-In, periodically, or some time after convergence has been attained, as determined by configuration on the device validating the routes received.

- The local authorization database is examined, and the entries with the longest prefix length which encompasses the range of addresses described by the prefix is chosen.
- This set of entries is examined to determine if the route is originated from one of the AS' listed in the set of entries.

- If the route is not originated from one of the AS' listed in the set of entries, the route is said to be invalid, and should be excluded from further consideration for installation in to the local RIB.

- If the authorization database entry indicates the Second Hop must be checked, the second hop of the AS PATH should be verified as described in the section "Verifying the Second Hop," below. If the second hop check fails, the route is said to be invalid, and should be removed from further consideration for installation into the local RIB.

- If the authorization database indicates the AS_PATH of the update must be checked, the AS PATH of the route should be verified, as described in the section "Verifying the AS PATH," below. If the AS_PATH is not validated, the route should be removbed from further consideration for installation in the local RIB.

- Other policies contained in the local authorization database should be applied as directed by the policy.

- If there is no entry in the local authorization database which encompasses the range of addresses described by the prefix, then the route is said to be unverified, and should be handled according to local policy (either discarded, or have its security preference lowered).

6.2. Aggregation

Aggregation is a difficult problem with any method which attempts to verify the origin of any given prefix, since aggregation removes the relationship between prefixes originated and originators. Prefixes may only be aggregated by an entity which is otherwise authorized to advertise the aggregated prefix.

6.3. Verifying the Path

In order to verify the path of any given advertisement, we propose to build a directed graph of all possible transit paths within the Internet, and verifying the path of each advertisement against this graph. Note that for any given prefix, this graph will contain a superset of all valid paths for the prefix. The BGP protocol itself is designed to choose the correct path out of the possible paths.

In order to build the directed graph, each entity within the routing

system may advertise the autonomous systems it is connected to using the attached AS' field carried in a certificate as described in [SOBGP-CERTIFICATE]. This graph is called the PATH database.

6.3.1. Building and Using the Directed Graph

Each device verifying routing information it is receiving may build a directed graph from the information contained in the Policy Certificate for each AS (as described above), which provides a list of all known valid paths through the internetwork. Each link between a pair of AS' may be verified from both ends of the link (using a two way connectivity check), thus ensuring that no single AS may advertise itself as connected to an entity it is not connected to.

As routes are validated, the AS PATH contained in the route may be checked against the PATH database for congruence with a known good path. If the AS PATH traverses a link which cannot be verified to exist in both directions (the link fails the two way connectivity check), it's security preference may be lowered or the route may be discarded, as local device configuration directs.

6.4. Verifying the Second Hop

As a prefix is processed by a receiving the device, the originator and second hop in the AS PATH may be checked against the authorized originator and the list of attached AS' advertised by that originator. If the second hop in the AS path (the second entry in the AS PATH) does not match one of the attached AS' advertised by the originator, the prefix should be treated as invalid, and discarded.

6.5. Requirements for Systems Running soBGP

There are three general requirements imposed on autonomous system border routers and devices running soBGP:

- Any peering session along the border of an autonomous system running soBGP SHOULD be authenticated through some means such as [BGP-MD5], IPsec ([ESP], [AH]), or through some other current, effective means of protecting BGP sessions from being hijaaked, or otherwise abused.

- Any peering session along which soBGP certificates are exchanged MUST be authenticated through some means such as [BGP-MD5], IPsec ([ESP, [AH]), or through some other current, effective

means of protecting BGP sessions from being hijaaked, or otherwise abused.

- The AS_PATH of any routing information received from any BGP peer outside the autonomous system MUST be checked to validate the next hop AS is the AS the update was received from. If the next hop AS in any received update does not match the configured AS the route is learned from, the update MUST be discarded.

7. The Security Preference

As indicated in other sections of this document, the user has a wide range of flexibility when determining the level of security to be applied to each prefix. The level of security may be translated into a Security Preference and used to influence the route selection process [BGP].

To determine the Secuirty Preference of a prefix, the following algorithm may be used:

1. A Security Preference is assigned to each prefix indicating neutral trust. The neutral value is locally significant to the autonomous system, and all routers in it MUST use the same value.

2. The Security Preference SHOULD be lowered for any of the following cases:

   - The AS_PATH verification failed at a link other than the second hop.
   - The Second Hop verification failed.
   - The prefix is unverified.
   - The prefix is invalid.

   Each of these cases represents a different degree of failure in the validation and verification process. The amount by which the Security Preference is lowered is locally significant to the autonomous system, and all routers in it MUST use the same value. If the local policy determines that a prefix may be used (even if it may have failed the validation and verification process), it is recommended that the Security Preference be lowered such that a lower value is assigned to invalid paths.

3. The Security Preference SHOULD be increased for any of the

following cases:

- o The prefix is validated.
- o The Second Hop verification passed.
- o The AS_PATH verification passed for the whole path.

The amount by which the Security Preference is increased is locally significant to the autonomous system, and all routers in it MUST use the same value. It is recommended that the Security Preference be increased such that a higher value is assigned to paths that pass all the validation and verification steps.

The resulting Security Preference value may be used to select among different routes for the same prefix; the higher value MUST be preferred. Any of the following methods may be used:

A Consider the Security Preference prior to calculating the degree of preference [BGP] for a prefix.

B Assign the value of the Security Preference to any of the attributes used in the Decision Process [BGP]. Care must be taken with attributes for which the lower value is preferred.

C Use a Cost Community [COST] and its associated methods to consider the Security Preference at any step in the Decision Process [BGP] without overloading other attributes. Care must be taken as the lowest value in a Cost Community is preferred.

The method selected MUST be consistent through the local Autonomous System.

8. Security Considerations

This document defines extensions to BGP that address specific security concerns for the protocol. While it adds functionality, the flexilibty allows it to not introduce any new security concerns.

9. IANA Considerations

This document defines the Security Message for BGP, which contains a series of TLVs. IANA is expected to maintain a registry of all the values defined, as follows:

The SECURITY message Type field :

- Type value 0 is reserved.
- Type values 1 and 2 are assigned in this document.
- Type values 3 through 16575 MUST be assigned using the "IETF Consensus" policy defined in RFC2434 [RFC2434].
- Type values 16576 through 32895 SHOULD be assigned using the "Specification Required" policy defined in RFC2434 [RFC2434].
- Type values 32896 through 65535 are for "Private Use" as defined in RFC2434 [RFC2434].

Request TLV Request Type Field:

- Bits 1 through 3 are assigned in this document.
- Bits 4 thru 8 MUST be assigned using the "IETF Consensus" policy defined in RFC2434 [RFC2434].
- Bits 9 thru 16 are for "Private Use" as defined in RFC2434 [RFC2434].

10. Normative References

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

[MULTIPROTOCOL-BGP]
Bates, T., Chandra, R., Katz, D., and Rekhter, Y., "Multiprotocol Extensions for BGP-4", RFC 2858, June 2000

[CAPABILITY]
Chandra, R., Scudder, J., "Capabilities Advertisement with BGP-4", RFC2842, May 2000

[SOBGP-DEPLOY]
White, R. (editor), "Architecture and Deployment Considerations for Secure Origin BGP (soBGP) Deployment", draft-white-sobgp-

deployment-02, April 2004

[SOBGP-CERTIFICATE]
Weiss, Brian (editor), "Secure Origin BGP (soBGP) Certificates", draft-weis-sobgp-certificates-01.txt, October 2003

11. Informative References

[COST]
Retana, A., White, R., "BGP Custom Decision Process", draft-retana-bgp-custom-decision-00, October 2002.

[RFC2434]
Narten, T., Alvestrand, H., "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 2434, October 1998.

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

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

[AH] Kent, S., and R. Atkinson, "IP Authentication Header", RFC 2402, November 1998.

[SOBGP-RADIUS]
Lovnick, C, "RADIUS Attributes for soBGP Support", draft-lonvick-sobgp-radius-04.txt, February 2004

12. Editor's Address

James Ng (Editor)
Cisco Systems
7025 Kit Creek Road
Research Triangle Park, NC 27709
jamng@cisco.com