Network Working Group
Internet-Draft
Intended status: Informational
Expires: May 3, 2018

F. Templin, Ed.
Boeing Research & Technology
G. Dawra
A. Lindem
Cisco Systems, Inc.
October 30, 2017

A Simple BGP-based Mobile Routing System for the Aeronautical Telecommunications Network
draft-templin-atn-bgp-04.txt

Abstract

The International Civil Aviation Organization (ICAO) is investigating mobile routing solutions for a worldwide Aeronautical Telecommunications Network with Internet Protocol Services (ATN/IPS). The ATN/IPS will eventually replace existing communication services with an IPv6-based service supporting pervasive Air Traffic Management (ATM) for Air Traffic Controllers (ATC), Airline Operations Controllers (AOC), and all commercial aircraft worldwide. This informational document describes a simple mobile routing service based on mature industry standards to address the ATN/IPS requirements.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 3, 2018.

Copyright Notice

Table of Contents

1. Introduction

The International Civil Aviation Organization [ICAO] is investigating mobile routing solutions for a worldwide Aeronautical Telecommunications Network with Internet Protocol Services (ATN/IPS). The ATN/IPS will eventually replace existing communication services with an IPv6-based service supporting pervasive Air Traffic Management (ATM) for Air Traffic Controllers (ATC), Airline Operations Controllers (AOC), and all commercial aircraft worldwide. This informational document describes a simple mobile routing service based on mature industry standards to address the ATN/IPS requirements.

Aircraft communicate via wireless aviation data links that typically support much lower data rates than terrestrial wireless and wired-line communications. For example, VHF-based data links only support data rates on the order of 32Kbps and an emerging L-Band data link

that is expected to play a key role in future aeronautical communications only supports rates on the order of 1Mbps. Although satellite data links can provide much higher data rates during optimal conditions, they (like all other aviation data links) are subject to errors, delay, disruption, signal intermittence,

degradation due to atmospheric conditions, etc. The well-connected ground domain ATN/IPS network should therefore treat each safety-of-flight critical packet produced by (or destined to) an aircraft as a precious commodity and strive for a "better-than-best-effort" service that provides the highest possible degree of reliability.

The ATN/IPS assumes a worldwide connected Internetwork for carrying ATM communications. The Internetwork could be manifested as a private collection of long-haul backbone links (e.g., fiberoptics, copper, SATCOM, etc.) interconnected by high-performance networking gear such as bridges, switches and routers. Such a private Internetwork would need to connect all ATN/IPS participants worldwide, and could therefore present a considerable cost for a large-scale deployment of new infrastructure. Alternatively, the ATN/IPS could be deployed as an overlay over the existing global public Internet itself as long as sufficient security and reliability provisions are met. For example, ATN/IPS nodes could be deployed as part of an SD-WAN or an MPLS-WAN that rides over the public Internet via secured tunnels.

The ATN/IPS further assumes that each aircraft will receive an IPv6 Mobile Network Prefix (MNP) that accompanies the aircraft wherever it travels. ATCs and AOCs will likewise receive IPv6 prefixes, but they would typically appear in static (not mobile) deployments. Throughout the rest of this document, we therefore use the term "MNP" when discussing an IPv6 prefix that is delegated to any ATN/IPS end system, including ATCs, AOCs and aircraft. We also use the term Mobility Service Prefix (MSP) to refer to an aggregated prefix assigned to the ATN/IPS by an Internet assigned numbers authority, and from which all MNPs are delegated (e.g., up to 2**32 IPv6 /64 MNPs could be delegated from the MSP 2001:db8::/32).

[CBB] describes an aviation mobile routing service based on dynamic updates in the global public Internet Border Gateway Protocol (BGP) [RFC4271] routing system. Practical experience with the approach has shown that frequent injections and withdrawals of MNPs in the

Internet routing system results in excessive BGP update messaging, slow routing table convergence times, and extended outages when no route is available. This is due to both conservative default BGP protocol timing parameters (see Section 5) and the complex peering interconnections of BGP routers within the global Internet infrastructure. The situation is further exacerbated by frequent aircraft mobility events that each result in BGP updates that must be propagated to all BGP routers in the Internet that carry a full routing table.

We therefore consider an approach using a BGP overlay network routing system where a private BGP routing protocol instance is maintained

between ATN/IPS Autonomous System (AS) Border Routers (ASBRs). The private BGP instance does not interact with the Internetwork BGP routing system, and BGP updates are unidirectional from "stub" ASBRs (s-ASBRs) to a very small set of "core" ASBRs (c-ASBRs) in a hub-and-spokes arrangement. For the AERO proposal [I-D.templin-aerolink], the s-ASBRs correspond to AERO Servers. For the LISP proposal [I-D.ietf-lisp-rfc6830bis], the s-ASBRs correspond to xTRs that connect directly to the BGP system instead of via Map Servers and Resolvers. No non-standard extensions of the BGP protocol are necessary.

The s-ASBRs for each stub AS connect to a small number of c-ASBRs via dedicated high speed links and/or tunnels across the Internetwork using industry-standard encapsulations (e.g., Generic Routing Encapsulation (GRE) [RFC2784], IPsec [RFC4301] etc.). The s-ASBRs engage in external BGP (eBGP) peerings with their respective c-ASBRs, and only maintain routing table entries for the MNPs currently active within the stub AS. A stub AS may connect to the core via multiple s-ASBRs, in which case the s-ASBRs would engage in an Interior Gateway Protocol (IGP) among themselves to maintain a common view of the stub AS MNPs. (The s-ASBRs need not engage in internal BGP (iBGP) peerings, since they do not receive any BGP updates from c-ASBRs and therefore have no BGP information to share with each other.) Finally, the s-ASBRs also maintain default routes with their c-ASBRs as the next hop, and therefore hold only partial topology information.

The c-ASBRs connect to other c-ASBRs using iBGP peerings over which they collaboratively maintain a full routing table for all active

MNPs currently in service. Therefore, only the c-ASBRs maintain a full BGP routing table and never send any BGP updates to s-ASBRs. This simple arrangement therefore greatly reduces the number of BGP updates that need to be synchronized among peers, and the number is reduced further still when localized mobility events within stub ASes (i.e., "intradomain" mobility events) are mitigated within the AS instead of being propagated to the core.

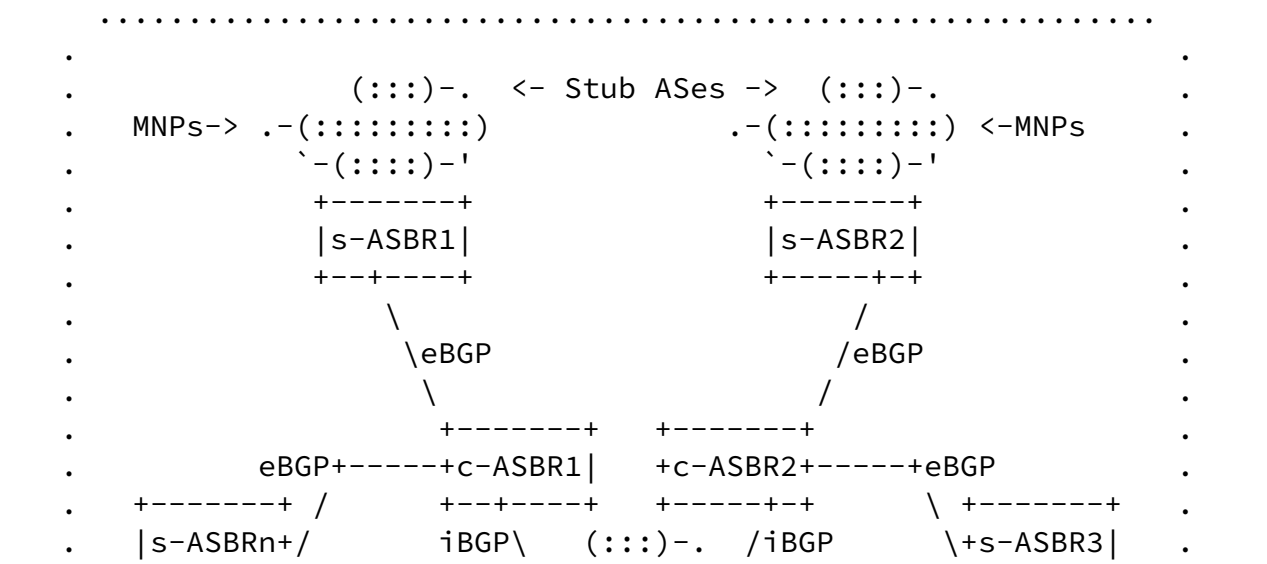
The following section provides a detailed discussion of the proposed BGP-based ATN/IPS routing system.

2. Proposed BGP-based ATN/IPS Routing System

The proposed ATN/IPS routing system comprises a private BGP instance coordinated between ASBRs in an overlay network. The overlay does not interact with the native Internetwork BGP routing system, and each c-ASBR advertises only a small and unchanging set of MSPs into the Internetwork instead of the full dynamically changing set of MNPs. The system corresponds to the framework first specified by the

LISP+ALT proposal [RFC6836] and later also adopted by the AERO proposal [I-D.templin-aerolink]. The system differs from the LISP Delegated Database Tree (DDT) proposal [I-D.ietf-lisp-ddt] that is designed with scalability as the primary consideration.

In a reference deployment, one or more s-ASBRs connect each stub AS to the overlay using a shared stub AS Number (ASN). Each s-ASBR further uses eBGP to peer with one or more c-ASBRs. All c-ASBRs are members of the same core AS, and use a shared core ASN. The c-ASBRs further use iBGP to maintain a synchronized consistent view of all active MNPs currently in service. Figure 1 below represents the reference deployment. Note that in the figure only two s-ASBRs show detail, but similar arrangements are implied for all other s-ASBRs. Note also that each stub AS shows only a single s-ASBR with a single c-ASBR connection, but in practical deployments each stub AS may have multiple s-ASBRs that peer with multiple c-ASBRs via eBGP, e.g., for fault tolerance.


(:::)-. <- Stub ASes -> (:::)-.
MNPs-> .-(::::::::::) .-(::::::::::) <-MNPs
`-(::::)-'  `-(::::)-'
+-------+ +-------+
|s-ASBR1| |s-ASBR2|
+--+----+ +-----+-+
\ /
\eBGP /eBGP
\ /
+-------+ +-------+
eBGP+-----+c-ASBR1| +c-ASBR2+-----+eBGP
+-------+ / +--+----+ +-----+-+ \ +-------+
|s-ASBRn+/ iBGP\ (:::)-. /iBGP \+s-ASBR3|

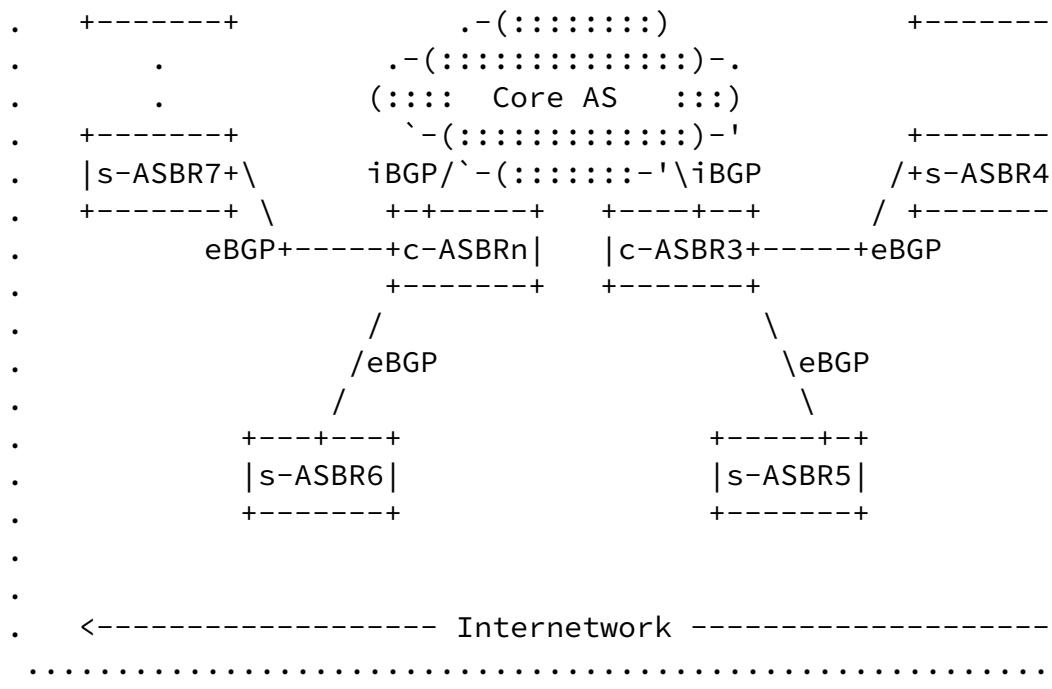


Figure 1: Reference Deployment

In the reference deployment, each s-ASBR maintains routes for active MNPs that currently belong to its stub AS, and dynamically announces new MNPs and withdraws departed MNPs in its eBGP updates to c-ASBRs in response to "interdomain" mobility events. Since ATN/IPS end systems are expected to remain within the same stub AS for extended timeframes, however, intradomain mobility events (such as an aircraft handing off between cell towers) are handled locally within the stub AS instead of being propagated as interdomain eBGP updates.

Each c-ASBR configures a black-hole route for each of its MSPs. By black-holing the MSPs, the c-ASBR will maintain forwarding table entries only for the MNPs that are currently active, and packets

destined to all other MNPs will correctly incur ICMPv6 Destination Unreachable messages [RFC4443] due to the black hole route. The c-ASBRs do not send eBGP updates for MNPs to s-ASBRs, but instead originate a default route. In this way, s-ASBRs have only partial topology knowledge (i.e., they know only about the active MNPs currently within their stub ASes) and they forward all other packets to c-ASBRs which have full topology knowledge.

Scaling properties of this ATN/IPS routing system are limited by the number of BGP routes that can be carried by the c-ASBRs. A 2015 study showed that BGP routers in the global public Internet at that time carried more than 500K routes with linear growth and no signs of router resource exhaustion [BGP]. A more recent network emulation study also showed that a single c-ASBR can accommodate at least 1M dynamically changing BGP routes even on a lightweight virtual machine, with the expectation that high-performance dedicated router hardware can support even more.

Therefore, assuming each c-ASBR can carry 1M or more routes, this means that at least 1M ATN/IPS end system MNPs can be serviced by a single set of c-ASBRs. A means of increasing scaling would be to assign a different set of c-ASBRs for each set of MSPs. In that case, each s-ASBR still peers with one or more c-ASBRs from each set of c-ASBRs, but the s-ASBR institutes route filters so that it only sends BGP updates to the specific set of c-ASBRs that aggregate the MSP. For example, if the MSP for the ATN/IPS deployment is 2001:db8::/32, a first set of c-ASBRs could service the MSP segment 2001:db8::/40, a second set could service 2001:db8:0100::/40, a third set could service 2001:db8:0200::/40, etc.

Assuming a sufficient number of c-ASBR sets, the ATN/IPS routing system can then accommodate 1B or more MNPs. In this way, each set of c-ASBRs services a specific set of MSPs that they advertise to the native Internetwork routing system, and each s-ASBR configures MSP-specific routes that list the correct set of c-ASBRs as next hops. This arrangement also allows for natural incremental deployment, and can support small scale initial deployments followed by dynamic deployment of additional ATN/IPS infrastructure elements without disturbing the already-deployed base.

Finally, c-ASBRs may have multiple routing table entries for a single MNP advertised by multiple s-ASBRs. Each s-ASBR can be assigned a MULTI_EXIT_DISC (MED) metric for routes that it originates in its eBGP peering configurations [RFC4451] so that c-ASBRs can determine preferences for MNPs learned from multiple s-ASBRs. In this way, c-ASBRs can select the neighboring s-ASBR with the lowest MED value, i.e., even if it is not on the shortest path. The c-ASBR can then fail over to a s-ASBR with a larger MED value in case of MNP

withdrawal or s-ASBR failure. Such an event could correspond to an

aviation data link handover, e.g., when an aircraft switches over from a satellite link to an L-Band link.

3. Route Optimization

ATN/IPS end systems will frequently need to communicate with correspondents located in other stub ASes. In the ASBR peering arrangement discussed in Section 2, this can initially only be accommodated by having the source s-ASBR forward packets to a c-ASBR which then forwards the packets toward the destination s-ASBR where the destination ATN/IPS end system resides. In many cases, it would be desirable to eliminate c-ASBRs from this "dogleg" route so that the source s-ASBR can send packets directly to the destination s-ASBR through tunneling across the Internetwork. This can be accomplished using a mapping resolution service such as proposed in AERO [I-D.templin-aerolink] or LISP [I-D.ietf-lisp-rfc6830bis][I-D.ietf-lisp-rfc6833bis]. Employment of the mapping resolution service results in a condition known as route optimization.

A route optimization example is shown in Figure 2 and Figure 3 below. In the first figure, the dogleg route between correspondents in the stub ASes traverses the path from s-ASBR1 to c-ASBR1 to c-ASBR2 to S-ASBR2. In the second figure, the optimized route goes directly from s-ASBR1 to s-ASBR2, i.e., the c-ASBRs are not included in the path.

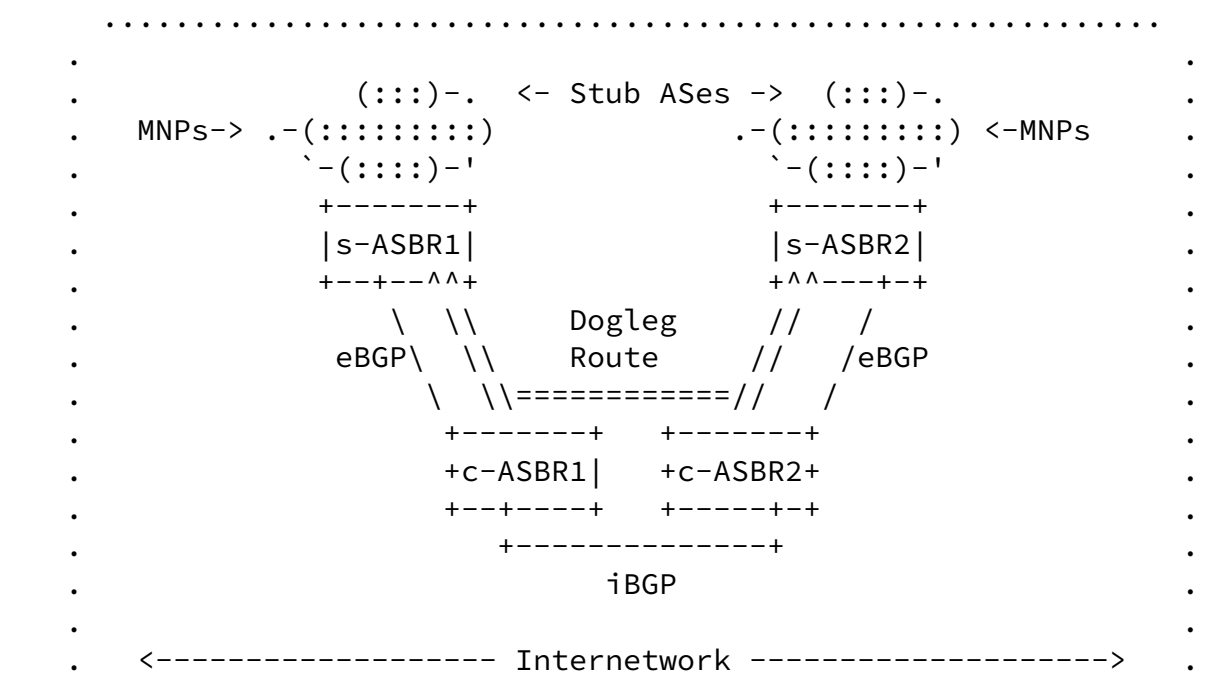


Figure 2: Dogleg Route Before Optimization

Figure 3: Direct Route Following Optimization

Note that route optimization can fail if the source s-ASBR cannot tunnel packets directly to the destination s-ASBR due to some form of

Internetwork blockage such as filtering middleboxes. It is also necessary for the source s-ASBR to quickly detect and adjust to

failure of the destination s-ASBR and/or movement of the destination MNP. In both of these cases, significant packet loss could occur before the source s-ASBR can detect that the destination MNP is no longer reachable via the route-optimized path. This implies that route optimized paths may not always be the best choice for carrying safety-of-flight critical packets with high reliability requirements.

In all cases, s-ASBRs do not advertise MNPs discovered via route optimization to c-ASBRs. Instead, s-ASBRs keep MNPs discovered via route optimization in a local table that is kept separate from the MNPs of ATN/IPS end systems within their own stub AS.

4. Route Availability

In the ATN/IPS BGP-based routing system proposed in this document, each s-ASBR always has a default route and can therefore always send packets via the dogleg route through a c-ASBR even if a route optimized path has been established. The direct paths between s-ASBRs and c-ASBRs are maintained by BGP peering session keepalives such that, if a link or an ASBR goes down, BGP will detect the failure and readjust the routing tables. However, ASBRs and the links that interconnect them are expected to be secured as highly-available and fault tolerant critical infrastructure such that peering session failures should be extremely rare.

This represents a distinct architectural difference from other approaches that only operate over route optimized paths. With the approach described herein the source s-ASBR will always have a working route, even if only via a dogleg path through a c-ASBR. This gives rise to the possibility of sending {high-priority, low-data-rate} packets via the assured dogleg route and {low-priority, high-data-rate} packets via the optimized route, e.g., based on per-packet quality of service indications. This could also give rise to a fair pricing model that would charge more for use of the high-assurance dogleg path and less for use of the lesser-assured route-optimized path.

This distinction is important to aviation networking, where isolated safety-of-flight critical packets such as produced by the Controller

Pilot Data Link Communications (CPDLC) facility may not be eligible for retransmission, e.g., if an aviation data link is failing. If there is no route available, the packet can be dropped or delayed and safety-of-flight parameters could be lost. Even when an optimized route is discovered on-demand, the route may not work and again safety-of-flight critical packets could be lost.

5. BGP Protocol Considerations

The number of eBGP peering sessions that each c-ASBR must service is proportional to the number of s-ASBRs in the system. Network emulations with lightweight virtual machines have shown that a single c-ASBR can service at least 100 eBGP peerings from s-ASBRs that each advertise 10K MNP routes (i.e., 1M total). It is expected that robust c-ASBRs can service many more peerings than this - possibly by multiple orders of magnitude. But even assuming a conservative limit, the number of s-ASBRs could be increased by also increasing the number of c-ASBRs. Since c-ASBRs also peer with each other using iBGP, however, larger-scale c-ASBR deployments may need to employ an adjunct facility such as BGP route reflectors [RFC4456].

The number of aircraft in operation at a given time wordlwide is likely to be significantly less than 1M, but we will assume this number for a worst-case analysis. Assuming an average 1hour flight profile from gate-to-gate, and 10 data link transitions per flight, the entire system will need to service at most 10M BGP updates per hour (2778 updates per second). This number is within the realm of the peak BGP update messaging seen in the global public Internet today [BGP2]. Assuming a BGP update message size of 100 bytes (800bits), the total amount of BGP control message traffic to a single c-ASBR will be less than 2.5Mbps which is a nominal rate for modern data links.

Industry standard BGP routers provide configurable parameters with conservative default values. For example, the default hold time is 90 seconds, the default keepalive time is 1/3 of the hold time, and the default MinRouteAdvertisementinterval is 30 seconds for eBGP peers and 5 seconds for iBGP peers (see Section 10 of [RFC4271]). For the simple mobile routing system described herein, these

parameters can and should be set to more aggressive values to support faster neighbor/link failure detection and faster routing protocol convergence times. For example, a hold time of 3 seconds and a MinRouteAdvertisementinterval of 0 seconds for both iBGP and eBGP.

By default, MED only compares metrics that originate from multiple neighbors within the same AS [RFC4451]. In order to compare MED metrics that come from different ASes, a router configuration file entry may be needed (e.g., Cisco routers require the configuration file entry "bgp always-compare-med"). Furthermore, in order for the MED discriminator to be applied correctly, the AS_PATH phase in the BGP route selection process must be disabled (e.g., Cisco routers use the configuration file entry "bgp bestpath as-path ignore").

6. Implementation Status

The BGP routing arrangement described in this document has been modeled in realistic network emulations showing that the MED process results in selection of the best peer when multiple peers advertise the same MNP. Modeling has also shown that at least 1 million MNPs can be propagated to each c-ASBR even on lightweight virtual machines.

7. IANA Considerations

This document does not introduce any IANA considerations.

8. Security Considerations

ATN/IPS ASBRs on the open Internet are susceptible to the same attack profiles as for any Internet nodes. For this reason, ASBRs should employ physical security and/or IP securing mechanisms such as IPsec [RFC4301], TLS [RFC5246], etc.

ATN/IPS ASBRs present targets for Distributed Denial of Service (DDoS) attacks. This concern is no different than for any node on the open Internet, where attackers could send spoofed packets to the node at high data rates. This can be mitigated by connecting ATN/IPS ASBRs over dedicated links with no connections to the Internet and/or

when ASBR connections to the Internet are only permitted through well-managed firewalls.

ATN/IPS s-ASBRs should institute rate limits to protect low data rate aviation data links from receiving DDoS packet floods.

9. Acknowledgements

This work is aligned with the FAA as per the SE2025 contract number DTFAWA-15-D-00030.

This work is aligned with the NASA Safe Autonomous Systems Operation (SASO) program under NASA contract number NNA16BD84C.

This work is aligned with the Boeing Information Technology (BIT) MobileNet program.

10. References

10.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/info/rfc791>.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <https://www.rfc-editor.org/info/rfc2460>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC4443] Conta, A., Deering, S., and M. Gupta, Ed., "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", STD 89,

RFC 4443, DOI 10.17487/RFC4443, March 2006, <https://www.rfc-editor.org/info/rfc4443>.

[RFC4451] McPherson, D. and V. Gill, "BGP MULTI_EXIT_DISC (MED) Considerations", RFC 4451, DOI 10.17487/RFC4451, March 2006, <https://www.rfc-editor.org/info/rfc4451>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <https://www.rfc-editor.org/info/rfc4456>.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <https://www.rfc-editor.org/info/rfc4861>.

10.2. Informative References

[BGP] Huston, G., "BGP in 2015, http://potaroo.net", January 2016.

[BGP2] Huston, G., "BGP Instability Report, http://bgpupdates.potaroo.net/instability/bgpupd.html", May 2017.

[CBB] Dul, A., "Global IP Network Mobility using Border Gateway Protocol (BGP), http://www.quark.net/docs/Global_IP_Network_Mobility_using_BGP.pdf", March 2006.

[I-D.ietf-lisp-ddt]
Fuller, V., Lewis, D., Ermagan, V., Jain, A., and A. Smirnov, "LISP Delegated Database Tree", draft-ietf-lisp-ddt-09 (work in progress), January 2017.

[I-D.ietf-lisp-rfc6830bis]
Farinacci, D., Fuller, V., Meyer, D., Lewis, D., and A. Cabellos-Aparicio, "The Locator/ID Separation Protocol (LISP)", draft-ietf-lisp-rfc6830bis-06 (work in progress), October 2017.

[I-D.ietf-lisp-rfc6833bis]
Fuller, V., Farinacci, D., and A. Cabellos-Aparicio, "Locator/ID Separation Protocol (LISP) Control-Plane", draft-ietf-lisp-rfc6833bis-06 (work in progress), October 2017.

[I-D.templin-aerolink]
Templin, F., "Asymmetric Extended Route Optimization (AERO)", draft-templin-aerolink-75 (work in progress), May 2017.

[ICAO] ICAO, I., "http://www.icao.int/Pages/default.aspx", February 2017.

[RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, DOI 10.17487/RFC2784, March 2000, <https://www.rfc-editor.org/info/rfc2784>.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI 10.17487/RFC4301, December 2005, <https://www.rfc-editor.org/info/rfc4301>.

[RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, DOI 10.17487/RFC5246, August 2008, <https://www.rfc-editor.org/info/rfc5246>.

[RFC6836] Fuller, V., Farinacci, D., Meyer, D., and D. Lewis, "Locator/ID Separation Protocol Alternative Logical Topology (LISP+ALT)", RFC 6836, DOI 10.17487/RFC6836, January 2013, <https://www.rfc-editor.org/info/rfc6836>.

Authors' Addresses

Fred L. Templin (editor)
Boeing Research & Technology

P.O. Box 3707
Seattle, WA 98124
USA

Email: fltemplin@acm.org

Gaurav Dawra
Cisco Systems, Inc.
USA

Email: gdawra@cisco.com

Acee Lindem
Cisco Systems, Inc.
USA

Email: acee@cisco.com