

Network Working Group
Internet-Draft
Intended status: Informational
Expires: 22 July 2022

R. Bush
Arrcus, Inc. & Internet Initiative Japan
J. Dong
Huawei Technologies
J. Haas, Ed.
Juniper Networks
W. Kumari, Ed.
Google
18 January 2022

Requirements and Considerations in BGP Peer Auto-Configuration draft-ietf-idr-bgp-autoconf-considerations-02

Abstract

This draft is an exploration of the requirements, the alternatives, and trade-offs in BGP peer auto-discovery at various layers in the stack. It is based on discussions in the IDR Working Group BGP Autoconf Design Team. The current target environment is the datacenter.

This document is not intended to become an RFC.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [[RFC2119](#)] [[RFC8174](#)] when, and only when, they appear in all capitals, as shown here.

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 22 July 2022.

Internet-Draft

BGP Peer Auto-Config Reqs

January 2022

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the [Trust Legal Provisions](#) and are provided without warranty as described in the Revised BSD License.

Table of Contents

1.	Introduction	3
2.	Design Team Determinations	3
2.1.	Problem Scope	3
2.2.	Simplicity	3
2.3.	BGP Auto-Discovery Protocol State Requirements	3
2.3.1.	BGP Auto-Discovery Protocol State	4
2.3.2.	BGP Session Protocol State	4
2.4.	BGP Auto-Discovery Protocol Transport Requirements	4
2.5.	Operator Configuration	5
3.	Design Principle Considerations	5
3.1.	Transport Considerations	5
3.2.	Auto-Discovery Protocol Timing Considerations	6
3.3.	Relationship with BGP	6
3.4.	Session Selection Considerations	6
3.5.	Session Stability Considerations	7
3.6.	Operational Trust Considerations	7
3.7.	Error Handling Considerations	9
4.	IANA Considerations	9
5.	Security Considerations	10
5.1.	BGP Transport Security Considerations	10
5.2.	Auto-discovery Protocol Considerations	10
5.2.1.	Potential Scopes of an Auto-discovery Protocol	10
5.2.2.	Desired Security Properties of the Auto-discovery Protocols	11
6.	Acknowledgments	12
7.	References	12
7.1.	Normative References	12

7.2.	Informative References	12
Appendix A.	Analysis of Candidate Approaches	14
A.1.	BGP Peer Discovery at Layer Two	14
A.1.1.	LLDP based Approach	15
A.1.2.	L3DL based Approach	15

A.2.	Link-Local Discovery	16
A.3.	BGP peer Discovery at Layer Three	16
A.3.1.	New BGP Hello Message based Approach	17
A.3.2.	BGP OPEN Message based Approach	17
A.3.3.	Bootstrapping BGP via BGP	18
A.3.4.	Bootstrapping BGP via OSPF	18
	Authors' Addresses	18

[1.](#) Introduction

This draft is an exploration of the requirements, the alternatives, and trade-offs in BGP peer auto-discovery at various layers in the stack. It is based on discussions in the IDR Working Group BGP Autoconf Design Team. The current target environment is the datacenter.

[2.](#) Design Team Determinations

[2.1.](#) Problem Scope

The current target environment is BGP as used for the underlay routing protocol in data center networks. Other scenarios may be considered as part of the analysis for this work, but work on those environments will be deferred to other efforts.

[2.2.](#) Simplicity

The auto-discovery mechanism is designed to be simple.

The goal is to select BGP Speakers where a BGP session may be successfully negotiated for a particular purpose. The auto-discovery mechanism will not replace or conflict with data exchanged by the BGP FSM, including its OPEN message.

[2.3.](#) BGP Auto-Discovery Protocol State Requirements

The Auto-Discovery Protocol is used discover BGP Session end-points. In other words, enough information to for a BGP Speaker to initiate a connection in the BGP protocol.

The BGP Session Properties, used by the discovering client to determine acceptability of the discovered session, are "discovered at OPEN" by the client by initiating a BGP session with the discovered end-point.

The required state that MUST be carried by the BGP Auto-Discovery Protocol for a discovered session includes:

- * IP addresses
- * Transport security parameters
- * GTSM [[RFC5082](#)] configuration, if any
- * BGP Session Protocol State Version Number

BGP Session Protocol State, discovered at BGP OPEN:

- * AS Numbers
- * BGP Identifier
- * Supported AFI/SAFIs

[2.3.1.](#) BGP Auto-Discovery Protocol State

- * Support for IPv4 and IPv6 address families, but do not assume that both are available.
- * The ability to use directly attached interface addresses, or the device's Loopback address. When using the Loopback address, potentially exchange additional information to bootstrap forwarding to that address.
- * Discovery of BGP transport protocol end-points and essential properties such as IP addresses, transport security parameters, and support for GTSM.
- * Transport security parameters include protocol – such as plain TCP, TCP-AO [[RFC5925](#)], IPsec [[RFC4301](#)], TCP-MD5 [[RFC2385](#)] – and necessary configuration for that protocol. Some example considerations for this are represented in YANG Data Model for Key Chains [[RFC8177](#)].
- * A version number representing when the BGP Session Protocol State has last changed. This can be used as a hint by an auto-discovery

client to determine when the state has been updated from a prior version. This can reduce repeated connections from an auto-discovery client to the discovered BGP Speaker when information has not changed.

[2.3.2.](#) BGP Session Protocol State

- * Discovery of BGP peer session parameters relevant to peer selection such as Autonomous System (AS) Numbers, BGP Identifiers, supported address families/subsequent-address families (AFI/SAFIs).

[2.4.](#) BGP Auto-Discovery Protocol Transport Requirements

BGP Auto-Discovery Protocol State may be carried in multiple protocols operating in different transport layers.

Implementations supporting more than one protocol for this state must have a mechanism for consistently selecting discovered BGP sessions. The BGP Identifier, which is carried by the BGP OPEN message, can help detect sessions to the same BGP Speaker carried in multiple protocols.

[2.5.](#) Operator Configuration

With BGP auto-discovery, some configuration of BGP is still needed. Operator configuration should be able to decide at least the following:

- * Select or otherwise filter which peers to actually try to send BGP OPEN messages.
- * Decide the parameters to use. For example:
 - IP addressing: IPv4 or IPv6.
 - Interface for peering: Loopback, or Direct.
 - Any special forwarding or routing needed for reaching the prospective peer; for example, loopback.
 - AS numbering.
 - BGP Transport Security Parameters.
 - BGP Policy that is appropriate for the type of discovered

session.

In addition to actually forming the BGP sessions, a common deployment model may also be the so called "validation" model. In this model, the operator configures the BGP sessions manually, and uses the information collected/populated by the BGP Auto-Configuration mechanism to validate that the sessions are correct.

[3.](#) Design Principle Considerations

This section summarizes the considerations of possible criteria for the design of a BGP auto-discovery mechanism, which may need further discussion in a wider community than the design team; for example, the IDR Working Group.

[3.1.](#) Transport Considerations

The network layer of the discovery mechanism may impact the scoping of the deployment of the auto-discovery mechanism.

- * Layer 2: For example, based on Ethernet.
- * Layer 3: Which is generic for any link-layer protocol.

Potentially leveraging existing protocols deployed in the data center.

The length of messages supported by the protocol.

How extensible the protocol is to carry future state for BGP auto-configuration.

[3.2.](#) Auto-Discovery Protocol Timing Considerations

Establishing a reasonable expectation for the timeliness of auto-configuration is desirable. When a link is plugged-in, one shouldn't have to wait minutes for potential peers to be discovered and BGP session establishment attempted. For protocols crafted explicitly for BGP auto-configuration, the time for discovery should be a reasonable amount of time; for example ten seconds or less.

Since discovery mechanisms may become very chatty when utilized by a

number of devices on shared networks, the protocol should not impose undue burden on the devices on that network to process the discovery messages. New auto-discovery protocols MUST NOT transmit messages more than once a second.

When an auto-discovery mechanism is used for a point-to-point link, or with the expectation of establishing a BGP session with a single BGP Speaker on that network, the auto-discovery protocol MAY quiesce once the discovered BGP session has become Established.

In cases where the auto-discovery protocol is carried as state in another protocol, that protocol will have its own timeliness considerations. The auto-discovery mechanism SHOULD NOT interfere with the timing of the existing protocol.

[3.3.](#) Relationship with BGP

- * The auto-discovery mechanism should be independent from BGP session establishment.
- * Not affect on BGP session establishment and routing exchange, other than the interactions for triggering the setup/removal of peer sessions based on the discovery mechanism.
- * Potentially leveraging existing BGP protocol sessions for discovery of new BGP sessions.

[3.4.](#) Session Selection Considerations

Candidate BGP sessions to a given BGP Speaker may be discovered by one or more auto-discovery protocols. Even for a single protocol, multiple transport session endpoints may be discovered for the same BGP Speaker. These different sessions may be required for supporting different address families, such as IPv4/IPv6, depending on the BGP operational practices for that device. Examples include a distinct

and matching session for the IPv4/IPv6 address family, a unified session carrying IPv4 over IPv6 and vice-versa, etc.

The BGP Identifier (router-id), a required protocol component of BGP, can serve to identify the same instance of the BGP Speaker. This is a required element of the information to be carried in the auto-discovery protocol.

When multiple mechanisms exist to discovery the same BGP speaker in an implementation, that implementation MUST document the process by which it chooses discovered peers. Those implementations also MUST describe interactions with their protocol state machinery for each mechanism.

[3.5.](#) Session Stability Considerations

BFD [[RFC5880](#)] is often used to provide fast failure detection for the BGP protocol. To provide for maximum compatibility and ease of use for auto-discovered sessions, [[I-D.ietf-idr-bgp-bfd-strict-mode](#)] SHOULD be used to provide consistent BFD protection for an auto-discovered BGP session.

[3.6.](#) Operational Trust Considerations

Different deployment models will have different trust models and requirements. Some of this will be driven by the size, complexity and operational practices of the operator. For example, some operators have very strict physical protection of the datacenter, and their deployment model assumes that anything which plugs into devices in the datacenter is, by definition, trusted. Other operators take a very different approach, and assume the least possible amount of trust.

Much of this difference is also reflected in the operator's

bootstrapping solution. Some operators build individual configurations for each device, and manually provision the configuration into the non-volatile storage of the device before it is shipped. Other operators use solutions similar to PXE Boot to automatically load an operating system and configuration onto the device, based on a unique device identifier (such as management Ethernet MAC address). Some operators pre-configure devices with identical base configurations containing some bootstrapping policy logic (e.g., "If you are a Model-X device, and interface 23 is connected to a device of type Y, then you must be at Stage-2 in a Clos fabric") and allow the device to use this policy information to infer its role and position. A final set of datacenter operators, for example enterprises, would like to be able to simply unpack a new device, plug it in and have the device infer everything. (It is unclear if this is a deployment model that we want to support.)

Many datacenter operators already have a well-developed process for installing and bringing up a new datacenter network, complete with solutions to bootstrap and configure the network. These operators will want to be able to use the BGP Autoconf mechanism to perform validation of the datacenter fabric, and ongoing "sanity-checking" to confirm that the datacenter is correctly cabled, and that the BGP sessions which have been configured from the database match what the autodiscovered sessions would have created. Over time, if the BGP Autoconf solution proves to be successful, reliable, and scaleable, operators may begin using it as the primary source of record.

Closely related to these considerations is the "scope" of the discovery process. It is expected that many operators will wish to only perform discovery on "infrastructure" or "fabric" interfaces, and not interfaces to customers.

It is not clear that the solution that chosen will be able to meet all of the trust and deployment models, and we will need to prioritize which set(s) of deployment scenarios are the most important for the Working Group to solve.

Trust/Operational deployment driven requirements. The solution should:

- * Allow operators to determine which classes of interfaces the discovery protocol operates on (e.g: "Interfaces numbered 1-17" or "Only 100GE interfaces"). This is likely an implementation detail.
- * Allow operation in a "validation" or "verification" only mode, where the Autoconf solution populates a database or model showing what sessions it would bring up if allowed.

- * Ideally allow for different levels of "granularity" in pre-configuration. For example, if the protocol is capable of autoconfiguring everything, it should also support filtering or limiting the session according to configured policy. (Likely an implementation detail.)
- * Support preconfigured authentication systems. This is an area where more discussion is needed! The solution **MUST** also support a "no authentication" mode. Negotiated keying solutions, such as IKE, may be desirable but not mandatory for the solution.
- * Support Ethernet sub-interfaces such as VLANs.
- * Support non-Ethernet interfaces. This may include tunnels.

[3.7.](#) Error Handling Considerations

The purpose of the BGP auto-discovery protocol is to discover potential BGP sessions and provide enough information for a BGP Speaker to start a BGP session. It is possible for the information present in the auto-discovery protocol to not match the session's information. Such mis-matches will result in different classes of problems:

- * The BGP transport session may not connect. This could be the result of mismatches in IP addresses, GTSM configuration, BGP transport security configuration, etc. In these cases, a BGP Speaker attempts to establish a session and fails. Implementations **SHOULD** provide a way to clear such discovered sessions or exclude them from further connect attempts.
- * The BGP transport session connects, but the parameters in the BGP OPEN message do not match those in the auto-discovery protocol. In this case, the implementation may wish to disconnect from the BGP session and exclude it from further connection attempts. The implementation **SHOULD** raise a visible fault to the operator. The implementation **SHOULD** provide a mechanism to permit further attempts to connect to the discovered session.
- * The operator may choose to leverage the auto-discovery mode for validation purposes only. The implementation should provide access to the operator for discovered BGP sessions from the auto-discovery protocol; for example via the user-interface. The implementation **SHOULD** permit a manually configured BGP session to conflict with information present in the auto-discovery protocol, but **SHOULD** raise an alarm with the operator that this has been done.

[4.](#) IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

[5.](#) Security Considerations

There are two primary components to be secured in environments utilizing BGP auto-discovery: The BGP transport layer discovered via the protocol, and the auto-discovery protocol itself.

[5.1.](#) BGP Transport Security Considerations

The purpose of the auto-discovery protocol is to ease the setup of BGP sessions for various applications, including data-center fabrics. However, care must be taken to not permit sessions to be setup outside of trusted environments. It is RECOMMENDED that sessions advertised using BGP auto-discovery be protected at the transport layer using mechanisms such as TCP-AO, IPsec, or the deprecated TCP-MD5.

It is thus a requirement that the auto-discovery protocol carry sufficient information to determine what transport security is to be used when establishing a BGP session.

All Security Considerations from [[RFC4272](#)], BGP Security Vulnerabilities Analysis, continue to apply.

[5.2.](#) Auto-discovery Protocol Considerations

As noted in previous sections, BGP auto-discovery be scoped to different portions of the network dependent on the network layer at which it is deployed. The information present in the auto-discovery protocol is considered sensitive, since it identifies resources running the BGP protocol. Care should be exercised in avoiding inadvertent disclosure of BGP sessions that are configured to permit auto-configuration even when BGP session transport security is in use. The auto-discovery protocol sets the context for such inadvertent disclosure.

[5.2.1.](#) Potential Scopes of an Auto-discovery Protocol

A Layer 2 unicast protocol targets a known device, potentially discovered through other means. The targeted device receives the message. Depending on the Layer 2 environment, other devices on the same link may be able to observe the protocol messages. Point to point links may also fall into this category.

A Layer 2 multicast protocol targets a group of devices on that Layer 2 multicast domain. A set of devices in that domain receives the message. Such messages may cross a number of devices in the domain. An example of this includes a set of Ethernet switches.

A Layer 3 unicast protocol inherits the properties of the Layer 2 protocol, and is intended to address a specific address - typically one device. Layer 3 unicast protocols may leverage GTSM for their security.

A Layer 3 multicast protocol addresses a group of devices in a given multicast domain. Such domains may be scoped, such as a single link's "All-Routers" group or perhaps all devices subscribed to a specific multicast group in a network. In many cases, a Layer 3 multicast protocol inherits the properties of the Layer 2 multicast protocol. Link-local scoped multicast protocols may be able to leverage GTSM.

A Layer 7 protocol is scoped per the mechanism in the underlying protocol. IGPs such as OSPF and IS-IS provide an "internal" scoping. BGP, depending on the deployment of the underlying address family, may vary from a targeted advertisement, to Internet-wide.

Each of these scopes provide different opportunities for inadvertent disclosure. The auto-discovery protocol will need to address how the desired security properties interact with the protocol scope.

[5.2.2.](#) Desired Security Properties of the Auto-discovery Protocols

Data Integrity is a required property. The data that is transmitted by a speaker of the auto-configuration protocol should be able to pass among its speakers properly.

Peer Entity authentication is a required property for Layer 2 and Layer 3 implementations. In a Layer 7 protocol, that protocol may provide the necessary authentication.

Confidentiality is an optional property. There is a tension between the desire to provide for a simple auto-configuration protocol that is easy to diagnose and debug with inadvertent disclosure.

The auto-configuration protocol must be resistant to Denial of Service, and to causing Denial of Service to discovered BGP session end-points.

[6.](#) Acknowledgments

The IDR BGP Auto-Conf Design Team.

[7.](#) References

[7.1.](#) Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in [RFC 2119](#) Key Words", [BCP 14](#), [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

[7.2.](#) Informative References

[I-D.acee-idr-lldp-peer-discovery]
Lindem, A., Patel, K., Zandi, S., Haas, J., and X. Xu,
"BGP Logical Link Discovery Protocol (LLDP) Peer
Discovery", Work in Progress, Internet-Draft, [draft-acee-idr-lldp-peer-discovery-10](#), 8 August 2021,
<<https://www.ietf.org/archive/id/draft-acee-idr-lldp-peer->

[discovery-10.txt](#)>.

[I-D.acee-ospf-bgp-rr]

Lindem, A., Patel, K., Zandi, S., and R. Raszuk, "OSPF Extensions for Advertising/Signaling BGP Route Reflector Information", Work in Progress, Internet-Draft, [draft-acee-ospf-bgp-rr-01](#), 7 September 2017, <<https://www.ietf.org/archive/id/draft-acee-ospf-bgp-rr-01.txt>>.

[I-D.ietf-idr-bgp-bfd-strict-mode]

Zheng, M., Lindem, A., Haas, J., and A. Fu, "BGP BFD Strict-Mode", Work in Progress, Internet-Draft, [draft-ietf-idr-bgp-bfd-strict-mode-06](#), 8 November 2021, <<https://www.ietf.org/archive/id/draft-ietf-idr-bgp-bfd-strict-mode-06.txt>>.

[I-D.ietf-lsvr-l3dl]

Bush, R., Austein, R., and K. Patel, "Layer-3 Discovery and Liveness", Work in Progress, Internet-Draft, [draft-ietf-lsvr-l3dl-08](#), 14 October 2021, <<https://www.ietf.org/archive/id/draft-ietf-lsvr-l3dl-08.txt>>.

Bush, et al.

Expires 22 July 2022

[Page 12]

Internet-Draft

BGP Peer Auto-Config Reqs

January 2022

[I-D.ietf-lsvr-l3dl-signing]

Bush, R., Housley, R., and R. Austein, "Layer-3 Discovery and Liveness Signing", Work in Progress, Internet-Draft, [draft-ietf-lsvr-l3dl-signing-03](#), 14 October 2021, <<https://www.ietf.org/archive/id/draft-ietf-lsvr-l3dl-signing-03.txt>>.

[I-D.ietf-lsvr-l3dl-ulpc]

Bush, R. and K. Patel, "L3DL Upper-Layer Protocol Configuration", Work in Progress, Internet-Draft, [draft-ietf-lsvr-l3dl-ulpc-02](#), 14 October 2021, <<https://www.ietf.org/archive/id/draft-ietf-lsvr-l3dl-ulpc-02.txt>>.

[I-D.ietf-lsvr-lsoe]

Bush, R., Austein, R., and K. Patel, "Link State Over Ethernet", Work in Progress, Internet-Draft, [draft-ietf-lsvr-lsoe-01](#), 17 February 2019,

<<https://www.ietf.org/archive/id/draft-ietf-lsvr-lsoe-01.txt>>.

[I-D.raszuk-idr-bgp-auto-discovery]

Raszuk, R., Mitchell, J., Kumari, W., Patel, K., and J. Scudder, "BGP Auto Discovery", Work in Progress, Internet-Draft, [draft-raszuk-idr-bgp-auto-discovery-07](https://www.ietf.org/archive/id/draft-raszuk-idr-bgp-auto-discovery-07), 13 October 2021, <<https://www.ietf.org/archive/id/draft-raszuk-idr-bgp-auto-discovery-07.txt>>.

[I-D.raszuk-idr-bgp-auto-session-setup]

Raszuk, R., "BGP Automated Session Setup", Work in Progress, Internet-Draft, [draft-raszuk-idr-bgp-auto-session-setup-01](https://www.ietf.org/archive/id/draft-raszuk-idr-bgp-auto-session-setup-01), 11 December 2019, <<https://www.ietf.org/archive/id/draft-raszuk-idr-bgp-auto-session-setup-01.txt>>.

[I-D.xu-idr-neighbor-autodiscovery]

Xu, X., Talaulikar, K., Bi, K., Tantsura, J., and N. Triantafyllis, "BGP Neighbor Discovery", Work in Progress, Internet-Draft, [draft-xu-idr-neighbor-autodiscovery-12](https://www.ietf.org/archive/id/draft-xu-idr-neighbor-autodiscovery-12), 26 November 2019, <<https://www.ietf.org/archive/id/draft-xu-idr-neighbor-autodiscovery-12.txt>>.

[RFC0826] Plummer, D., "An Ethernet Address Resolution Protocol: Or Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware", STD 37, [RFC 826](https://www.rfc-editor.org/info/rfc826), DOI 10.17487/RFC0826, November 1982, <<https://www.rfc-editor.org/info/rfc826>>.

Bush, et al.

Expires 22 July 2022

[Page 13]

Internet-Draft

BGP Peer Auto-Config Reqs

January 2022

[RFC2385] Heffernan, A., "Protection of BGP Sessions via the TCP MD5 Signature Option", [RFC 2385](https://www.rfc-editor.org/info/rfc2385), DOI 10.17487/RFC2385, August 1998, <<https://www.rfc-editor.org/info/rfc2385>>.

[RFC4272] Murphy, S., "BGP Security Vulnerabilities Analysis", [RFC 4272](https://www.rfc-editor.org/info/rfc4272), DOI 10.17487/RFC4272, January 2006, <<https://www.rfc-editor.org/info/rfc4272>>.

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

- [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>>.
- [RFC5082] Gill, V., Heasley, J., Meyer, D., Savola, P., Ed., and C. Pignataro, "The Generalized TTL Security Mechanism (GTSM)", [RFC 5082](#), DOI 10.17487/RFC5082, October 2007, <<https://www.rfc-editor.org/info/rfc5082>>.
- [RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", [RFC 5880](#), DOI 10.17487/RFC5880, June 2010, <<https://www.rfc-editor.org/info/rfc5880>>.
- [RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", [RFC 5925](#), DOI 10.17487/RFC5925, June 2010, <<https://www.rfc-editor.org/info/rfc5925>>.
- [RFC8177] Lindem, A., Ed., Qu, Y., Yeung, D., Chen, I., and J. Zhang, "YANG Data Model for Key Chains", [RFC 8177](#), DOI 10.17487/RFC8177, June 2017, <<https://www.rfc-editor.org/info/rfc8177>>.

[Appendix A](#). Analysis of Candidate Approaches

As part of the work on distilling the requirements for BGP auto-discovery, the Design Team reviewed several proposals for implementing auto-discovery. The analysis of these proposals, including missing elements the Design Team decided were part of the requirements, follows.

[A.1](#). BGP Peer Discovery at Layer Two

BGP Discovery at Layer-2 would entail finding potential peers on a LAN or on Point-to-Point links and discovering their Layer-3 attributes, such as, IP addresses, etc.

There are two available candidates for peer discovery at Layer-2, one is based on Link Layer Discovery Protocol (LLDP) and the other is based on Layer 3 Discovery Protocol, L3DL [[I-D.ietf-lsvr-l3dl](#)].

[A.1.1.](#) LLDP based Approach

LLDP is a widely deployed protocol with implementations in most devices in data centers. Currently it only advertises the management Layer-3 address, but could presumably be extended to include the per-interface addresses.

LLDP has a limitation that all information must fit in a single PDU (it does not support fragmentation / a "session"). There is an early LLDPv2 development effort to extend this in the IEEE.

[I-D.acee-idr-lldp-peer-discovery] describes how to use the LLDP IETF Organizationally Specific TLV to augment the LLDP TLV set to exchange BGP Config Sub-TLVs signaling:

- * AFI
- * IP address (IPv4 or IPv6)
- * Local AS number
- * Local BGP Identifier (AKA, BGP Router ID)
- * Session Group-ID; i.e., the BGP Device Role
- * BGP Session Capabilities
- * Key Chain
- * Local Address (as BGP Next Hop).

[A.1.2.](#) L3DL based Approach

L3DL [[I-D.ietf-lsvr-l3dl](#)] is an ongoing development in the IETF LSVR Working Group with the goal of discovering IP Layer-3 attributes of links, such as neighbor IP addressing, logical link IP encapsulation abilities, and link liveness which may then be disseminated for the use of BGP-SPF and similar protocols.

L3DL Upper Layer Protocol Configuration, [[I-D.ietf-lsvr-l3dl-ulpc](#)], details signaling the minimal set of parameters needed to start a BGP session with a discovered peer. Details such as loopback peering are handled by attributes in the L3DL protocol itself. The information which can be discovered by L3DL is:

- * AS number
- * Local IP address, IPv4 or IPv6, and
- * BGP Authentication.

L3DL and L3DL-ULPC have well-specified security mechanisms, see [[I-D.ietf-lsvr-l3dl-signing](#)].

The functionality of L3DL-ULPC is similar but not quite the same as the needs of IDR Design Team. For example, L3DL is designed to meet more complex needs. L3DL's predecessor, LSOE [[I-D.ietf-lsvr-lsoe](#)], was simpler and might be a better candidate for adaptation. If needed, the design of LSOE may be customized for the needs of BGP peer auto-discovery.

Unlike LLDP, L3DL has only one implementation, and LSOE has only one open source implementation, and neither is significantly deployed.

[A.2.](#) Link-Local Discovery

Some existing BGP auto-configuration mechanisms leverage "point to point" addressing schemes to bootstrap BGP sessions. One example utilizes an IP subnet numbered such that it may contain only two hosts - for IPv4, a /30 or /31 network; for IPv6 a /127 network. An additional mechanism may leverage IPv4 ARP [[RFC0826](#)] or IPv6 Neighbor Discovery [[RFC4861](#)] to learn of hosts on a subnet.

Such existing mechanisms do not provide an auto-discovery protocol with necessary parameters. Rather, they simplify configuration by permitting BGP session configuration templates to be easily applied to interfaces without requiring addressing to be known a priori.

[A.3.](#) BGP peer Discovery at Layer Three

Discovery at Layer-3 can assume IP addressability, though the IP addresses of potential peers are not known a priori and need to be discovered before further negotiation. IP multicast may be a good choice to address the above concern.

The possible problem would appear to discovery at Layer-3 is that one may not know whether to use IPv4 or IPv6. This might be exacerbated by the possibility of a potential peer not being on the local subnet, and hence broadcast and similar techniques may not be applicable. While in data center network or networks in a single administrative domain, such issue could be easily solved.

If one can assume that the BGP session is based on point-to-point link, then discovery might try IPv6 link-local or even IPv4 link-local. A link broadcast or multicast protocol may also be used. For switched or bridged multi-point which is at least on the same subnet, VLAN, etc., multicast or broadcasts might be a viable approach.

There are four available candidates for BGP peer discovery at Layer-3: One is based on extending BGP with new Hello message for peer auto-discovery. One is based on reusing BGP OPEN message format for peer auto-discovery. One is based on bootstrapping BGP sessions via existing BGP sessions. One is based upon bootstrapping a BGP Route Reflector via the OSPF protocol.

[A.3.1.](#) New BGP Hello Message based Approach

[I-D.xu-idr-neighbor-autodiscovery] describes a BGP neighbor discovery mechanism which is based on a newly defined UDP based BGP Hello message. The BGP Hello message is sent in multicast to discover the directly connected BGP peers. According to the message header format and the TLVs carried in the message, the information which can be signaled is:

- * AS number
- * BGP Identifier
- * Accepted ASN list
- * Peering address (IPv4 or IPv6)
- * Local prefix (for loopback)
- * Link attributes
- * Neighbor state
- * BGP Authentication

The mechanisms in this draft do not currently handle fragmentation.

The mechanism in this draft is perhaps unique among the other proposals in requiring bi-directional state.

[A.3.2.](#) BGP OPEN Message based Approach

[I-D.raszuk-idr-bgp-auto-session-setup] describes a BGP neighbor discovery mechanism by reusing BGP OPEN message format with newly defined UDP port. The message is called BGP Session Explorer (BSE) packet and is sent in multicast. Since the message format is the same as BGP OPEN, the information which can be signaled is:

- * AS number
- * BGP Identifier

- * Peering address

The mechanism is currently under-specified with respect to a number of similar properties described elsewhere. A general implication is that those properties – and others providing for extensibility of the auto-discovery mechanism – would need to be added to the BGP OPEN message and deal with the related impacts on the BGP session's finite-state machine.

BGP PDUs, including the OPEN message, may be up to 4k in size. Since this mechanism leverages Layer 3 multicast, a PDU fragmentation mechanism may need to be described.

[A.3.3.](#) Bootstrapping BGP via BGP

[I-D.raszuk-idr-bgp-auto-discovery] describes a new BGP address family. The NLRI carries a Group ID + BGP Identifier as the key. A new BGP Path Attribute carries information about the sessions:

- * AS Number
- * AFI/SAFI
- * BGP Identifier
- * Peer Transport Address
- * Flags to declare a session for information only, to force a reset of a session on parameter changes, etc.

Since the BGP auto-discovery state is carried by BGP, it inherits the security implications of the underlying BGP session.

PDU size considerations are identical to those of a BGP UPDATE message.

Similarly, extensibility considerations would rely on either the new BGP Path Attribute, or one yet to be defined.

[A.3.4.](#) Bootstrapping BGP via OSPF

[I-D.acee-ospf-bgp-rr] describes a mechanism to learn BGP Route Reflectors via OSPFv2/OSPFv3 LSAs. Multiple types of scopes are defined for these LSAs to help constrain where they are advertised in an OSPF domain.

The BGP Route Reflector TLV contains:

- * Local AS Number
- * IPv4 or IPv6 Address of the Route Reflector
- * A list of AFI/SAFIs supported by the Route Reflector

The BGP Route Reflector TLV may be advertised more than once, potentially to describe different IP transport endpoints.

This mechanism does not provide for security properties of the BGP session or transport properties such as BFD or GTSM.

Authors' Addresses

Bush, et al.

Expires 22 July 2022

[Page 18]

Internet-Draft

BGP Peer Auto-Config Reqs

January 2022

Randy Bush
Arrcus, Inc. & Internet Initiative Japan
5147 Crystal Springs
Bainbridge Island, WA 98110
United States of America

Email: randy@psg.com

Jie Dong
Huawei Technologies
Huawei Campus, No. 156 Beiqing Road
Beijing
100095
China

Email: jie.dong@huawei.com

Jeffrey Haas (editor)
Juniper Networks
1133 Innovation Way
Sunnyvale, CA 94089
United States of America

Email: jhaas@juniper.net

Warren Kumari (editor)
Google
1600 Amphitheatre Parkway
Mountain View, CA 94043
United States of America

Email: warren@kumari.net