

Internet Engineering Task Force
Internet Draft
Intended status: Standards Track
Expiration Date: April 21, 2022

E. Chen
Palo Alto Networks
S. Sangli
Juniper Networks
October 20, 2021

Dynamic Capability for BGP-4
draft-ietf-idr-dynamic-cap-16.txt

Abstract

This document defines a new BGP capability termed "Dynamic Capability", which would allow the dynamic update of capabilities over an established BGP session. This capability would facilitate non-disruptive capability changes by BGP speakers.

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 April 21, 2022.

Copyright Notice

Copyright (c) 2021 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 Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Internet Draft

[draft-ietf-idr-dynamic-cap-16.txt](#)

October 2021

1. Introduction

Currently BGP capabilities [[RFC5492](#)] are only advertised in the BGP OPEN message [[RFC4271](#)] during the session initialization. In order to enable or disable a capability (such as the Address Family support [[RFC4760](#)]), an established session would need to be reset, which may disrupt other services running over the session. In addition, currently an advertised capability can not be updated on-demand over an established session. One example of such a requirement is for adjusting the "Restart Time" in the Graceful Restart Capability [[RFC4724](#)]) when performing certain planned maintenance in a network.

This document defines a new BGP capability termed "Dynamic Capability", which would allow the dynamic update of capabilities over an established BGP session. This capability would facilitate non-disruptive capability changes by BGP speakers.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "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.

2. Dynamic Capability

The Dynamic Capability is a new BGP capability [[RFC5492](#)]. The Capability Code for this capability is specified in the "IANA Considerations" section of this document. The Capability Value field consists of a list of capability codes (one-octet for each) that specify the capabilities that MAY be revised dynamically by the remote speaker.

By advertising the Dynamic Capability to a peer in the OPEN, a BGP speaker conveys to the peer that the speaker is capable of receiving and properly handling the CAPABILITY message (as defined in the next Section) from the peer after the BGP session has been established.

3. Capability Message

The CAPABILITY Message is a new BGP message type with type code 6. In addition to the fixed-size BGP header [[RFC4271](#)], the CAPABILITY message contains one or more of the following tuples of capability revisions:

```
+-----+
| Init/Ack (1 bit)           |
+-----+
| Ack Request (1 bit)       |
+-----+
| Reserved (5 bits)         |
+-----+
| Action (1 bit)           |
+-----+
| Sequence Number (4 octets) |
+-----+
| Capability Code (1 octet)  |
+-----+
| Capability Length (2 octets)|
+-----+
| Capability Value (variable)|
+-----+
```

The Init/Ack bit indicates whether a capability revision is being initiated (when set to 0), or being acknowledged (when set to 1).

The Ack Request bit indicates whether an acknowledgment is requested (when set to 1), or not (when set to 0) for a capability revision being initiated.

The Reserved bits should be set to zero by the sender and ignored by the receiver.

The Action bit is 0 for advertising a capability, and 1 for removing a capability.

The Sequence Number field can be used by a BGP speaker to match an acknowledgment with a capability revision that the speaker initiated previously.

Conceptually the triple <Capability Code, Capability Length, Capability Value> is the same as the one defined in [RFC5492], and it specifies a capability for which the "Action" shall be applied. The Capability Length field, though, is larger than the one specified in

[RFC5492].

If multiple capability instances (as described in [RFC5492]) are defined for the capability code, then each capability instance SHALL be revised individually. The triple <Capability Code, Capability Length, Capability Value> in the CAPABILITY message SHALL contain only one instance of the capability. The Multiprotocol Extensions Capability specified in [RFC4760] is an example of such a capability that has multiple instances defined.

If multiple capability instances (as described in [RFC5492]) are not defined for the capability code, then the "Action" specified applies to the whole capability identified by the capability code. Furthermore, if the "Action" is to remove a capability, then the Capability Length field SHOULD be set to zero by the sender and the Capability Value field MUST be ignored by the receiver even when the Capability Length field has a non-zero value.

If the "Action" is to remove a capability and the Capability Length field is zero, then the whole capability identified by the capability code is removed regardless whether multiple capability instances are defined for the capability code.

4. Operation

A BGP speaker that is willing to receive the CAPABILITY message (for one or more capability codes) from its peer SHOULD use the BGP Capabilities Advertisement [RFC5492] to advertise the Dynamic

Capability for these capability codes.

A BGP speaker MAY send to its peer a CAPABILITY message to initiate revisions for one or more capability codes only if these capability codes are listed in the Dynamic Capability of the OPEN message received from its peer.

A CAPABILITY message MAY be received only in the Established state. Receiving a CAPABILITY message in any other state is a Finite State Machine Error as defined in [RFC4271]. A BGP speaker SHOULD reset the HoldTimer upon receiving a CAPABILITY message from its peer.

When a BGP speaker sends a CAPABILITY message to its peer to initiate a capability revision, the Init/Ack bit for the capability revision in the message MUST be set to 0. The setting of the Ack Request bit is capability specific. The assignment of the Sequence Number is a local matter, but MUST allow the BGP speaker to unambiguously identify a capability revision it initiated previously based on the Sequence Number carried in the acknowledgment from the peer.

If the Init/Ack bit is set to 1 for a capability revision in a CAPABILITY message received by a BGP speaker, then the BGP speaker SHALL treat the capability revision as an acknowledgment of the receipt of a capability revision initiated by the BGP speaker. The BGP speaker MUST ignore the Ack Request bit, and SHALL use the Sequence Number carried in the capability revision to match with the capability revision previously initiated. The BGP speaker SHALL ignore an acknowledgment for a capability revision in which an acknowledgment was not requested by the BGP speaker. If the Sequence Number carried in the capability revision does not match any of the the Sequence Numbers used in the capability revisions initiated by the BGP speaker, then the BGP speaker SHOULD send a NOTIFICATION message as specified in the Error Handling section.

If the Init/Ack bit is set to 0 for a capability revision in a CAPABILITY message received by a BGP speaker, then the BGP speaker SHOULD first validate the capability code in the message. If the capability code is not listed in the Dynamic Capability advertised by the speaker to the peer, the BGP speaker SHOULD send a NOTIFICATION message as specified in the Error Handling section. For a valid capability code, if the Ack Request bit is set to 1, the BGP speaker MUST first send a CAPABILITY message to acknowledge the receipt of

the capability revision. The Init/Ack bit in the acknowledgment MUST be set to 1, and all the other fields in the capability revision MUST be kept unchanged.

After receiving a capability revision initiated by a peer, the BGP speaker SHALL update the capability previously received from that peer based on the Action bit in the message, and then function in accordance with the revised capability for the peer. The BGP speaker SHALL ignore such a capability revision that either results in no change to an existing capability, or removes a capability that was not advertised previously. The procedures specified in the "Error Handling" section SHOULD be followed when an error is detected in processing the CAPABILITY message.

In order to avoid ambiguities in sending and processing UPDATE messages, certain capability revisions may require close coordination between the BGP speaker (the Initiator) that initiates the capability revisions and another BGP speaker (the Receiver) that receives the capability revisions. The mechanism of acknowledgment defined in this document SHALL be used for the revision of such a capability. For the Initiator, the capability revision SHALL take effect (for the purpose of sending updates) immediately after the capability revision is sent, and the capability revision SHALL take effect (for the purpose of receiving updates) immediately after an acknowledgment is received from the Receiver. For the Receiver, the capability revision SHALL take effect (for the purpose of receiving updates)

immediately after the capability revision is received from the Initiator, and the capability revision SHALL take effect (for the purpose of sending updates) immediately after an acknowledgment is sent.

5. Error Handling

This document defines a new NOTIFICATION error code:

Error Code	Symbolic Name
7	CAPABILITY Message Error

The following error subcodes are defined as well:

Subcode	Symbolic Name
1	Unknown Sequence Number
2	Invalid Capability Length
3	Malformed Capability Value
4	Unsupported Capability Code

If a BGP speaker detects an error while processing a CAPABILITY message, it MUST send a NOTIFICATION message with Error Code CAPABILITY Message Error. If any of the defined error subcode is applicable, the Data field of the NOTIFICATION message MUST contain the tuple for the capability revision that causes the speaker to send the message.

If the Sequence Number carried in a capability revision marked as acknowledgment does not match any of the the Sequence Numbers used in the capability revisions initiated by the BGP speaker, then the error subcode is set to Unknown Sequence Number.

If the Capability Length field in the CAPABILITY message is incorrect for a Capability Code, then the error subcode is set to Invalid Capability Length.

If the Capability Value field in the CAPABILITY message is malformed (the definition of "malformed" depends on the Capability Code), then the error subcode is set to Malformed Capability Value.

If the Capability Code in the CAPABILITY message is not any of the capability codes advertised in the Dynamic Capability by the speaker, then the error subcode is set to Unsupported Capability Code.

6. Implementation Considerations

The extension specified in this document is designed for BGP capabilities in general. It can be used for a simple capability revision (e.g., a parameter change), as well as for a more complex revision that may involve changes to the encoding of BGP messages.

However, that does not mean all BGP capabilities warrant the support

of dynamic revisions. For a given capability, one should carefully consider the tradeoffs between the complexities in its implementation and the potential benefits when deciding whether to support its dynamic revision. For example, the tradeoff considerations could be more favorable for the Address Family Capability [[RFC4760](#)] and the Graceful Restart Capability [[RFC4724](#)] than for the ADD-PATH Capability [[RFC7911](#)].

[7.](#) IANA Considerations

This document defines the CAPABILITY message type for BGP with type code 6, and a NOTIFICATION error code and subcodes for the errors in a CAPABILITY message.

This document uses a BGP capability code to indicate that a BGP speaker supports the Dynamic Capability. The capability code 67 has been assigned by IANA.

[8.](#) Security Considerations

The extension proposed in this document does not change the underlying security or confidentiality issues inherent in the existing BGP [[RFC4271](#)].

[9.](#) Acknowledgments

The authors would like to thank Yakov Rekhter, Ravi Chandra, Dino Farinacci, Pedro Marques, Chandrashekhar Appanna, Derek Yeung, Bruno Rijnsman, John Scudder, Jeffrey Haas and Heidi Ou for their review and comments.

[10.](#) References

10.1. 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, <<http://www.rfc-editor.org/info/rfc2119>>.
- [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, <<http://www.rfc-editor.org/info/rfc4271>>.
- [RFC4760] Bates, T., Chandra, R., Rekhter, Y., and D. Katz, "Multiprotocol Extensions for BGP-4", [RFC 4760](#), DOI 10.17487/RFC4760, January 2007, <<http://www.rfc-editor.org/info/rfc4760>>.
- [RFC5492] Scudder, J. and R. Chandra, "Capabilities Advertisement with BGP-4", [RFC 5492](#), DOI 10.17487/RFC5492, February 2009, <<http://www.rfc-editor.org/info/rfc5492>>.
- [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>>.

10.2. Informative References

- [RFC4724] Sangli, S., Chen, E., Fernando, R., Scudder, J. and Y. Rekhter, "Graceful Restart Mechanism for BGP", [RFC 4724](#), DOI 10.17487/RFC4724, January 2007, <<http://www.rfc-editor.org/info/rfc4724>>.
- [RFC7911] Walton, D., Retana, A., Chen, E., and J. Scudder, "Advertisement of Multiple Paths in BGP", [RFC 7911](#), DOI 10.17487/RFC7911, July 2016, <<http://www.rfc-editor.org/info/rfc7911>>.

11. Authors' Addresses

Enke Chen
Palo Alto Networks, Inc.

Email: enchen@paloaltonetworks.com

Srihari R. Sangli
Juniper Networks, Inc.

Email: ssangli@juniper.net

