Abstract

The BGP specification describes a Decision Process for selecting the best route. This process uses a series of steps, made up of path attributes and other values, to first determine the Degree of Preference of a route and later as tie breakers. While existing mechanisms may achieve some of the same results described in this document, they can only do so through extensive configuration such as matching communities to explicit policy and/or route preference configurations present on each BGP speaker within their administrative domain (autonomous system). Implementing some specific fine grained policies through such mechanisms is cumbersome, if even possible.

This document defines a new Extended Community, called the Cost Community, which may be used as part of the Decision Process. The end result is a local Custom Decision Process.

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 http://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 August 7, 2017.
1. Introduction

The BGP specification defines a Decision Process [RFC4271] for selecting the best route. This process uses a series of steps, made up of path attributes and other values, to first determine the Degree of Preference of a route and later as tie breakers. While existing mechanisms may achieve some of the same results described in this document, they can only do so through extensive configuration such as
matching communities to explicit policy and/or route preference configurations present on each BGP speaker within their administrative domain (autonomous system). Implementing some specific fine grained policies through such mechanisms is cumbersome, if even possible. For example:

- **Local Preference:** The LOCAL_PREF is an attribute used to calculate the Degree of Preference in the Decision Process. There is no secondary degree of preference indicator available that can be considered after the MED, IGP metric, or other attributes.

- **Multi-Exit Discriminator (MED):** The MULTI_EXIT_DISC is an indicator of which local entrance point an AS would like a peering AS to use. As the MED is compared before the IGP metric, there is no way to set the MED so a route with a higher IGP metric is preferred over one with a lower IGP metric.

- **IGP Metric:** It is possible, to influence individual routes, but only by changing the next hops, and configuring the IGP metric for reaching them. This method is cumbersome and prone to confusion and error.

This document defines a new Extended Community [RFC4360], called the Cost Community, which may be used as part of the Decision Process. The end result is a local Custom Decision Process.

### 2. 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 [RFC2119].

### 3. The BGP Cost Community

The BGP Cost Community is an Opaque Extended Community [RFC4360] defined as follows:

**Type Field:**

The value of the high-order octet is determined by provisioning [RFC4360]. IANA has assigned the codepoint value 0x01 to 'Cost Community' in both the Transitive Opaque Extended Community Sub-Types registry and the Non-Transitive Opaque Extended Community Sub-Types registry.

**Value Field:**
The Value field contains four distinct sub-fields, described below:

```
+--------------------------------------+
| Point of Insertion (1 octet)         |
+--------------------------------------+
| R | Community-ID (7 bits)              |
+--------------------------------------+
| Cost (4 octets)                      |
|                                      |
|                                      |
+--------------------------------------+
```

The Point of Insertion (POI) sub-field indicates *after*, or *instead of*, which step in the Decision Process the Cost Community MUST be considered. The Point of Application (POA) refers to the step in the Decision Process where the Cost Community is effectively considered -- the relationship between the POI and the POA is determined by the Replace Bit (see below).

The value of the POI may reference a path attribute used in the Decision Process. In this case the POA is related to the step where the path attribute is considered.

The Decision Process includes some steps that do not correspond to any path attribute; the following values are defined:

128  ABSOLUTE_VALUE - Indicates that the Cost Community MUST be considered as the first step in determining the Degree of Preference of a route (Section 9.1.1 of [RFC4271]). If two routes have the same Cost, then the rest of the Calculation of Degree of Preference is to be followed.

129  IGP_COST - Indicates that the Cost Community MUST be considered after the interior (IGP) distance to the next-hop has been compared.

130  EXTERNAL_INTERNAL - Indicates that the Cost Community MUST be considered after Paragraph d of Section 9.1.2.2 of [RFC4271].

131  BGP_ID - Indicates that the Cost Community MUST be considered after the BGP Identifier (or ORIGINATOR_ID [RFC4456]) has been compared.
This document creates a new Cost Community Point of Insertion Registry (Section 7.1) that includes the relevant path attributes and these other values.

The Replace Bit (R-bit) is a single-bit field, that when set indicates that the Cost Community MUST replace the step indicated by the POI in the Decision Process.

If the R-bit is not set, then the POA is after the step in the Decision Process indicated by the POI, which may result in an additional step. If the R-bit is set, then the POA is at the step identified by the POI.

If the R-bit is set, the Cost in the Cost Community replaces the value of a path attribute at a specific step in the Decision Process, but not the attribute itself. For example, if the R-bit is set with the AS_PATH POI, the AS_PATH attribute would still be used for loop detection [RFC4271], but the Cost would replace its length in the Decision Process.

The R-bit MUST be ignored when used with the ABSOLUTE_VALUE POI.

If the Accumulated IGP Metric Attribute (AIGP) [RFC7311] is used such that the "AIGP-enhanced interior cost" replaces the "interior cost" tie breaker in the Decision Process, and the R-bit is set with the IGP_COST POI, then the Cost Community SHOULD be ignored in favor of the process described in Section 4.2 of [RFC7311].

The Community-ID sub-field contains an identifier to distinguish between multiple instances of the Cost Community.

The Cost sub-field is a 32-bit unsigned integer. It contains a value assigned by the network administrator that is significant to their administrative domain. The default Cost is 0x7FFFFFFF (half the maximum).

If the Cost Community is inserted after a step in the Decision Process, and is therefore only compared to other Cost Communities, the lower Cost MUST be preferred.

If the Cost Community replaces a step in the Decision Process, it MUST be treated exactly as the value it is replacing would be treated. It is up to the network administrator to select the appropriate Cost to use when replacing a specific step; the method to do that is outside the scope of this document.
4. Operation

The network administrator may use the Cost Community to assign a Cost to a route originated, or learned from a peer, in any part of their administrative domain. The POA MUST also be specified by the combination of the POI and the R-bit.

If a BGP speaker receives a route that contains the Cost Community, it MUST consider its Cost at the POA specified, during the Decision Process.

If the POI is not valid for the local Decision Process implementation, then the Cost Community SHOULD be silently ignored.

Multiple Cost Communities may indicate the same POA. All the Cost Communities for a specific POA MUST be considered starting with the one(s) with the lowest Community-ID. If multiple Cost Communities, for the same POA, with the same Community-ID are received for the same route from the same peer, then all except the one with the lowest Cost MUST be silently ignored.

Routes that do not contain the Cost Community (for a valid, particular POA), or a Community-ID present in a route from another peer, MUST be considered to have the default Cost.

If a range of routes is to be aggregated and the resultant aggregate path attributes do not carry the ATOMIC_AGGREGATE attribute, then the resulting aggregate SHOULD have an Extended Communities path attribute which contains the set union of all the Cost Communities from all of the aggregated routes. If multiple Cost Communities for the same POA (and with the same Community-ID) exist, then only the ones with the highest Cost SHOULD be included.

If the non-transitive version of a Cost Community is received across an Autonomous System boundary, then the receiver MUST strip it off the BGP update, and ignore it during the Decision Process.

5. Deployment Considerations

The mechanisms described in this document may be used to modify the Decision Process arbitrarily. It is important that a consistent Decision Process be maintained across the local Autonomous System to avoid potential routing loops. In other words, all the nodes in the AS that may have to consider the Cost Community MUST support the mechanisms described in this document.

Any mechanism which allows the modification of the Decision Process is capable of forming persistent routing loops in the control plane.
Network administrators deploying the Cost Community MUST ensure that each impacted router supports them mechanisms in this document for the POIs deployed within their network. This is similar in scope to a network administrator who uses communities [RFC1997] combined with filters or other policies to modify the Decision Process of BGP speakers. Consistency must be enforced at an administrative level.

6. Security Considerations

This document defines a new Extended Community, called the Cost Community, which may be used to customize the Decision Process. As such, the considerations outlined in [RFC4360] and [RFC4271] do not change.

To minimize the potential of creating routing loops (Section 5) or otherwise affecting the Decision Process in unintended ways, the propagation of Cost Communities MUST be disabled by default and MUST be explicitly enabled by the network administrator. Furthermore, all Cost Communities received across an Autonomous System boundary without explicitly being enabled MUST be stripped off the BGP update, and ignored during the Decision Process.

An ill-designed policy deployment using the Cost Community (e.g. one where there is no consistent POI support throughout the AS) may result in persistent routing loops that could result in loss of traffic. The design and implementation of policies for best route selection are outside the scope of this document.

7. IANA Considerations

IANA has assigned the codepoint value 0x01 to 'Cost Community' in both the Transitive Opaque Extended Community Sub-Types registry and the Non-Transitive Opaque Extended Community Sub-Types registry.

Section 3 also defines a series of values to be used to indicate steps in the Decision Process that do not map directly to a path attribute. IANA is expected to maintain a registry for the Cost Community POI values.

- Values 1 through 127 are to be assigned using the "Standards Action" policy or the Early Allocation process [RFC7120].
- Values 128 through 191 are to be assigned using the "IETF Consensus" policy.
- Values 192 through 254 are to be assigned using the "First Come First Served" policy.
Values 0 and 255 are reserved.

All the policies mentioned are documented in [RFC5226].

The table in Section 7.1 shows the initial allocations for the new Cost Community Point of Insertion registry.

### 7.1. Cost Community Point of Insertion Registry

The tables below document the initial Cost Community Point of Insertion Registry:

<table>
<thead>
<tr>
<th>Range</th>
<th>Registration Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Reserved</td>
</tr>
<tr>
<td>1-127</td>
<td>Standards Action</td>
</tr>
<tr>
<td>128-191</td>
<td>IETF Consensus</td>
</tr>
<tr>
<td>192-254</td>
<td>First Come First Served</td>
</tr>
<tr>
<td>255</td>
<td>Reserved</td>
</tr>
</tbody>
</table>

#### Registration Procedure

<table>
<thead>
<tr>
<th>Value</th>
<th>Code</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ORIGIN</td>
<td>[RFC4271]</td>
</tr>
<tr>
<td>2</td>
<td>AS_PATH</td>
<td>[RFC4271]</td>
</tr>
<tr>
<td>3</td>
<td>Unassigned</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>MULTI_EXIT_DISC</td>
<td>[RFC4271]</td>
</tr>
<tr>
<td>5</td>
<td>LOCAL_PREF</td>
<td>[RFC4271]</td>
</tr>
<tr>
<td>6-25</td>
<td>Unassigned</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>AIGP</td>
<td>[RFC7311]</td>
</tr>
<tr>
<td>27-127</td>
<td>Unassigned</td>
<td></td>
</tr>
<tr>
<td>128</td>
<td>ABSOLUTE_VALUE</td>
<td>[draft-ietf-idr-custom-decision]</td>
</tr>
<tr>
<td>129</td>
<td>IGP_COST</td>
<td>[draft-ietf-idr-custom-decision]</td>
</tr>
<tr>
<td>130</td>
<td>EXTERNAL_INTERNAL</td>
<td>[draft-ietf-idr-custom-decision]</td>
</tr>
<tr>
<td>131</td>
<td>BGP_ID</td>
<td>[draft-ietf-idr-custom-decision]</td>
</tr>
</tbody>
</table>

8. Acknowledgements

There have been many people who have shown their support and provided valuable input, comments and implementations -- the authors would like to thank all of them! We would like to also thank Dan Tappan.
for the Opaque Extended Community type. Bruno Decraene and Eric Rosen thoroughly reviewed this document and helped improved its quality significantly.

9. References

9.1. Normative References


9.2. Informative References


Appendix A. Change Log

This section is to be removed before publication.

A.1  Changes between the -00 and -01 versions.
    o  Updated authors' contact information.
    o  Editorial changes in the "Operations" and "Acknowledgement" sections.

A.2  Changes between the -01 and -02 versions.
    o  Updated authors' contact information.
    o  Added text to replace a step in the selection process.
    o  Minor edits.

A.3  Changes between the -02 and -03 versions.
    o  No changes; just a refresh.

A.4  Changes between the -03 and -04 versions.
    o  Updated authors' contact information.

A.5  Changes between the -04 and -05 versions.
    o  Updated authors' contact information.

A.6  Changes between the -05 and -06 versions.
    o  Updated RFC 7120 reference (from RFC 4020).

A.7  Changes between the -06 and -07 versions
    o  The review from Bruno Decraene and Eric Rosen resulted in several important changes related to the clarity and consistency of the document.
    o  Added considerations for co-existence with AIGP.
    o  Security Considerations.
A.8. Changes between the -07 and -08 versions

- Clarified the Security Considerations to ensure that routers don't apply the Cost Community by default.
- Separated the high-order bit in the Community-ID into its own field (for clarity). Called it the Replace Bit (R-bit).
- Introduced the POA concept.

Authors' Addresses

Alvaro Retana
Cisco Systems, Inc.
7025 Kit Creek Rd.
Research Triangle Park, NC 27709
USA

Email: aretana@cisco.com

Russ White
LinkedIn
Apex, NC 27539
USA

Email: russ@riw.us