BGP Guide

Deployment details and member-facing services guide

version 1.2 – May 2016

KanREN, INC 2029 Becker Drive, Suite 282, Lawrence, KS 66047 **T** 785-856-9800 **F** 785-856-1377 www.kanren.net

Introduction

What is KanREN?

The Kansas Research and Education Network (KanREN, Inc.) is a non-profit consortium of colleges, universities, school districts and other organizations in Kansas, organized for the purpose of facilitating communication among them, and providing themselves with connectivity to the Internet via a statewide TCP/IP network.

KanREN's Mission

The purpose of KanREN, Inc. is to provide innovative, cost-effective network technology and attention to individual member needs.

KanREN's Vision

Kansas educators, learners and researchers will be aided in attaining their goals through KanREN, Inc., an innovator and provider of premier network and technology services.

KanREN's Values

- **Integrity** - We practice fiscal,technical and professional integrity at all times.
- **Community** - The KanREN membership is an interconnected community that has created, and continues to support KanREN, and we will assist our members first, then other organizations.
- **Respect** - We treat members, one another and others with respect and consideration.
- **Solutions** - We provide sustainable and equitable technology solutions for members.
- **Partnerships** - We seek collaboration that furthers our vision.
- **Standards** - We adhere to industry standards and practices.

Important Facts

- KanREN is an independent, not-for-profit 501(c)3 Kansas corporation, Bylaws at http://www.kanren.net/about/KanREN%20Bylaws.pdf
- KanREN is not primarily a network – KanREN is an educational consortium. KanREN operates a state-wide TCP/IP network on behalf of its member institutions.
- Membership in KanREN is open to any college, university, library, or school district in the state of Kansas. Other non-profit organizations may join the consortium subject to the approval of the KanREN executive committee.
- KanREN is not a commercial Internet Service Provider (ISP), though we do provide dedicated Internet connectivity for most of our member sites.

- KanREN is not supported with any funding from the state or federal government. Though begun with funding from the National Science Foundation in 1993, today KanREN is completely supported by membership fees paid by its member institutions.
- KanREN is not an agency of the government of the state of Kansas.
- KanREN is an Internet2 Sponsored Education Group Participant (SEGP)
- KanREN is a Kan-Ed peering partner
- KanREN is affiliated with the Great Plains Network
- KanREN is a member of Net@EDU - The policy making division of Educause.
- KanREN provides E-rate discounts (SPIN: 143005645)

KanREN BGP Deployment

The KanREN core network features a robust BGP environment with an advanced featureset. At the center of its features, KanREN supports several BGP mechanisms that allow member-signaled traffic engineering. KanREN has invested significant resources in the design, initial deployment, and continuing maintenance of the features to provide advanced, stable features to its members.

Questions regarding KanREN's BGP environment should be forwarded to support@kanren.net if they are of a non-critical nature. Questions of a critical nature should be forwarded to KanREN's Network Operations Center at 785-856-9820.

Requests for new BGP features should be funneled through KanREN's Emerging Services Working Group (ESWG).

Types of BGP Peers

KanREN categorizes all BGP peers into groups for purposes of feature support, policy enforcement, and route selection. Categorization of peers allows for quicker and more consistent deployment of services. KanREN's peer groups are as follows:

1. **Paid Transit Peers and Commercial Peering Networks**
   Definition: Internet Service Providers and networks offer a significant portion of the global BGP route table
2. **Lateral, Settlement-Free Peers**
   Definition: Direct peering sessions with Content Delivery Networks or residential ISP networks
3. **Upstream R&E and CAI Peers**
   Definition: Upstream Research and Education Networks (regional, national, or international) as well as Community Anchor Institution Networks
4. **KanREN Member Peers**
   Definition: KanREN's members that originate and announce their own prefixes

Please note: Peers change frequently based upon availability, need, and economics. A list of current BGP peers can be found at http://www.kanren.net/services/bgp.shtml or via email to support@kanren.net.

Global Local Preference Values

KanREN applies the following Local Preference values to inbound routes. Additional Local Preference values may be signaled by members with properly formatted community strings. Please reference the Traffic Engineering Features section for more details.

Value	Description
100	Default value applied to "bulk" prefixes
200	Upstream R&E, and CAI prefixes
300	Lateral, Settlement-Free prefixes
400	Member-learned prefixes

Internet Route Registries (IRR) and Regional Internet Registries (RIR)

All networks wishing to BGP peer with KanREN must have a RIR-assigned, public ASN with proper whois information. For North America, most users should petition the American Registry for Internet Numbers (ARIN) for a public, registered ASN.

Additionally, KanREN reserves the right to create entries in the RADb Internet Route Registry (IRR) database maintained by Merit. This includes entries for member IP spaces. KanREN reserves the right to make similar entries in other IRRs as needed. IRR entries are required for some upstream ISPs before a prefix can be accepted and do not indicate any level of "ownership" over the numbered resource.

Address Family Support

As of January 2011, KanREN supports the following MP-BGP Address Families:

1. IPv4 unicast
2. IPv4 multicast
3. IPv6 unicast

IPv6 multicast support will be added summer 2011 while L2VPLS and L3VPN address families may be added at a later date depending on member need. Support for these advanced address families will be determined by KanREN's Emerging Services Working Group (ESWG).

Address Family Use Case and Configuration Examples

Any network wishing services beyond IPv4 Unicast, must configure multi-protocol BGP. KanREN suggests members configure BGP with an eye toward MP-BGP features.

Cisco IOS Configuration Example

```
!
router bgp 65000
 neighbor 2001:0DB8::1 remote-as 2495
 neighbor 2001:0DB8::1 description eBGP-to-KanREN
 neighbor 192.168.1.1 remote-as 2495
 neighbor 192.168.1.1 description eBGP-to-KanREN
 !
 address-family ipv4
  neighbor 192.168.1.1 activate
  network 172.16.20.0 mask 255.255.255.0
 exit-address-family
 !
 address-family ipv6
  neighbor 2001:0DB8::1 activate
  network 2001:0DB8::/32
 exit-address-family
!
```

Juniper JunOS Configuration Example

```
routing-options {
    autonomous-system 65000;
}
protocols {
    bgp {
        group to-KanREN {
             description eBGP-to-KanREN;
```

```
                type external;
                peer-as 2495;
                neighbor 192.168.1.1;
                neighbor 2001:0DB8::1;
            }
        }
    }
```

Supported BGP Timers

BGP-4 features user-configurable Hold and Keep-Alive timers. These values are negotiated during initial session setup along with configured address families and other capabilities. Timers applied to a BGP neighbor will be applied to all address families configured for the specific peer. Please note that Cisco and Juniper routers (and potentially others) immediately reset the peering session when new timers are configured.

Aggressive BGP timers are only useful during control plane events (software failure, abnormally high CPU, etc). Any event that results in an interface down (power outage, disconnected cable, failed optic, etc) causes the BGP session to be marked down and traffic rerouted immediately. These features were added many years ago; most vendors refer to it as Fast External Neighbor Loss Detection.

KanREN strongly suggests configuring timers with the following formula:

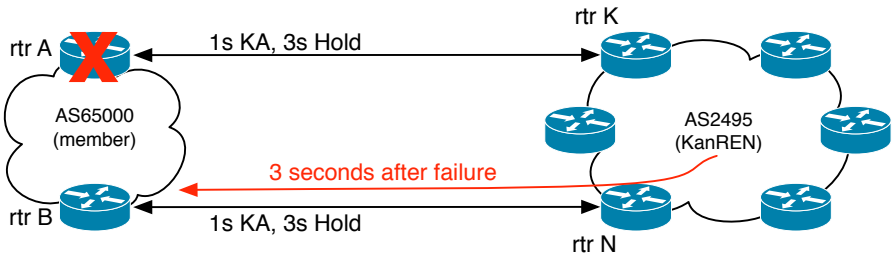
```
Hold Timer = (Keep-Alive Timer) * 3
```

KanREN supports the following BGP timers:

	Keep-Alive Timer	Hold Timer
KanREN Default	60 seconds	180 seconds
Cisco Default	60 seconds	180 seconds
Juniper Default	30 seconds	90 seconds
KanREN Minimum	1 second	3 seconds
Cisco Minimum	1 second	1 second
Juniper Minimum	1/3rd of negotiated hold timer	6 seconds

BGP Timer User Case

Members with multiple peering sessions into the KanREN core network may configure aggressive BGP timers to enable faster failure discovery and service restoration. In the diagram below, a KanREN member has multiple BGP sessions into the KanREN network from multiple routers on the member LAN. BGP Keep-Alive and Hold Timers have been configured (and negotiated) between the two peers. Router A on the members LAN has experienced a software failure and is no longer transmitting BGP control traffic. 3 seconds after the issue first appears, KanREN marks the BGP session to Router A down and moves traffic to the B<>N connection.



While service in this case was interrupted, a 3 second outage is typically favorable to a 180 second outage (default timers). Bi-Directional Forwarding Detection can identify and act on trouble MUCH faster and is outlined in the following section.

Aggressive BGP timers are typically most useful when the member maintains multiple BGP session into the KanREN network. Members with a single session do not gain any measurable benefit from aggressive BGP timers since a second path is not available for rerouting traffic.

BGP Timer Configuration Examples

Cisco IOS Example

```
!
router bgp 65000
!
neighbor 10.10.10.1 remote-as 2495
neighbor 10.10.10.1 description eBGP-to-KanREN
neighbor 10.10.10.1 timers <KA value> <Hold value>
!
```

Juniper JunOS Example

```
routing-options {
    autonomous-system 65000;
}
protocols {
    bgp {
        group to-KanREN {
             description eBGP-to-KanREN;
             type external;
             peer-as 2495;
             neighbor 192.168.1.1 {
                 hold-time 6;
             }
        }
    }
}
```

Bi-Directional Forwarding Detection (BFD) Support

For members that require the fastest possible failover, KanREN suggests use of Bi-Directional Forwarding Detection (BFD) on the individual BGP session(s). BFD is a more scaleable solution since many platforms perform BFD functions in hardware (as opposed to control plane CPUs). Because BFD configuration must be coordinated on both sides, members must work with KanREN to properly configure BFD features.

KanREN supports the following BFD timers for BGP sessions:

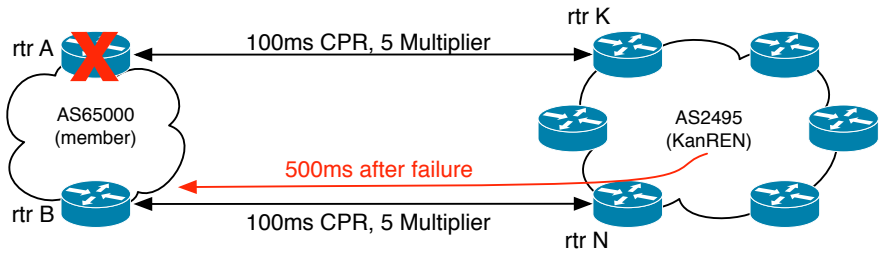
	Control Packet Rate	Multiplier
KanREN Default	100ms	5
KanREN Minimum	50ms	3

BFD is configured with a rate for control packets in milliseconds coupled with a multiplier. The packet rate indicates how quickly keep-alive packets will be generated. Multiplier values indicate how many intervals of missing control packets should pass before action is taken. For example, KanREN's default values of 100ms and 5 intervals will result in 500ms of service outage before both peers mark the session offline.

BFD is a “helper” protocol meaning it runs in parallel to routing protocols. The individual routing protocols must have BFD features enabled; thus allowing for a simple feature turnup. BFD only brings a session up with its neighbor after the higher-level routing protocol has established a neighbor / adjacency session. When a failure is detected, BFD signals the event to all relevant routing protocols so they may take action. When the failure is resolved, the routing protocol re-establishes its session then requests BFD bring its session online as well.

BFD Use Case

In the diagram below, a KanREN member has multiple BGP sessions into the KanREN network with BFD configured. A failure on the member’s LAN caused Router A to stop generating BFD control packets. After 500 miliseconds, the KanREN network detected the failure and moved traffic to the B<>N connection. In this case BFD offers the minimal service outage possible.



BFD Configuration Examples

Cisco IOS Example

```
!
interface GigabitEthernet 0/0
 description Ethernet to KanREN
 bfd interval 100 min_rx 100 multiplier 5
!
router bgp 65000
 neighbor 10.10.10.1 remote-as 2495
 neighbor 10.10.10.1 description eBGP-to-KanREN
 neighbor 10.10.10.1 fall-over bfd
!
```

Juniper JunOS Example

```
routing-options {
    autonomous-system 65000;
}
protocols {
    bgp {
        group to-KanREN {
             description eBGP-to-KanREN;
             type external;
             peer-as 2495;
             neighbor 10.10.10.1 {
                bfd-liveness-detection {
                    minimum-interval 100;
                    minimum-receive-interval 100;
                    multiplier 5;
                }
            }
        }
    }
}
```

Cooperative Route Filtering / Outbound Route Filtering Support

Outbound Route Filtering (ORF), sometimes referenced as Cooperative Route Filtering (CRF), is defined in RFC5292 and supported by KanREN for member BGP sessions. Lateral and upstream peering sessions do not support ORF features.

Outbound Route Filtering is negotiated as a capability during session establishment. Adding ORF features to an existing session may cause an immediate BGP session reset and should not be added outside a scheduled, coordinated maintenance window.

If ORF is supported on the session, the ingress route filter of the downstream router is transferred to the upstream router and applied as an outbound route filter. ORF allows the upstream router to process route filters before flooding routes to the downstream device. This saves CPU and memory resources on the downstream router. Changes to the downstream route filter are only transferred to the upstream after issuing a soft outbound reset on the downstream router's session.

KanREN supports ORF features; however, the following caveats should be noted:

a. Only prefix lists are transferred for ORF purposes; complex filters cannot currently (May 2011) be transferred
b. Some platforms must configure compatibility modes

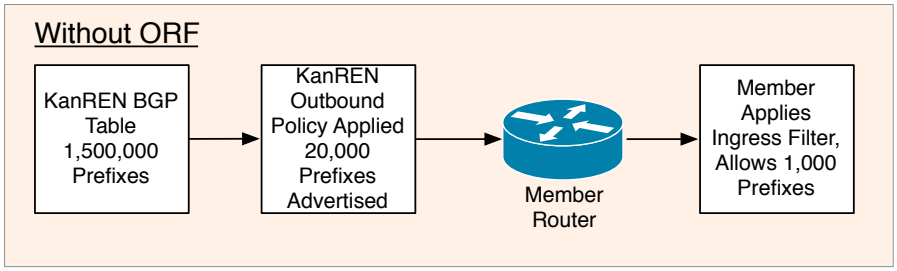
KanREN is happy to integrate site needs into its outbound route filters; making ORF unnecessary for most situations. If your site would like to work with KanREN to configure ORF features, please contact us via email at support@kanren.net.

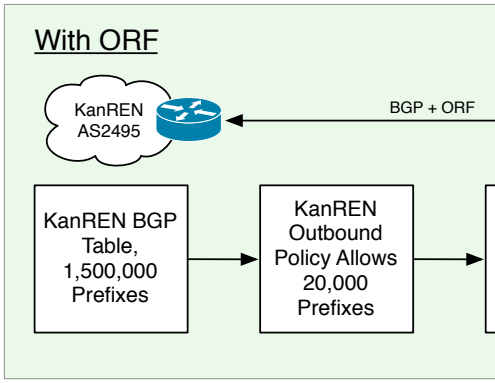
ORF Use Case

ORF can help save resources on member routers by applying a member's ingress filter before prefixes are advertised from the KanREN core.

In the diagram below, KanREN has a configured outbound route policy that permits 20,000 prefixes. The member router accepts these prefixes, stores them in local memory, THEN applies its ingress filter and installs 1,000 prefixes. This results in the router storing 19,000 unused and unnecessary prefixes in its memory.

By adding ORF capabilities to the BGP session, KanREN's core router applies the member's filter BEFORE advertising any prefixes. This allows for the greatest efficiency in resource utilization on the member router.





ORF Configuration Examples

Cisco IOS Example

```
!
router bgp 65000
 neighbor 10.10.10.1 remote-as 2495
 neighbor 10.10.10.1 description eBGP-to-KanREN
address-family ipv4
 neighbor 10.10.10.1 capability orf prefix-list send
 neighbor 10.10.10.1 prefix-list orf-example_in in
!
ip prefix-list orf-example seq 100 permit 192.168.1.0/24
ip prefix-list orf-example seq 110 deny 172.16.0.0/12
!
```

Juniper JunOS Example

```
routing-options {
    autonomous-system 65000;
}
protocols {
    bgp {
        group to-KanREN {
             description eBGP-to-KanREN;
             type external;
             peer-as 2495;
             neighbor 10.10.10.1 {
                capability orf {
                    prefix-list-cisco send;
                }
            }
        }
    }
}
```

Route Propagation Control Features

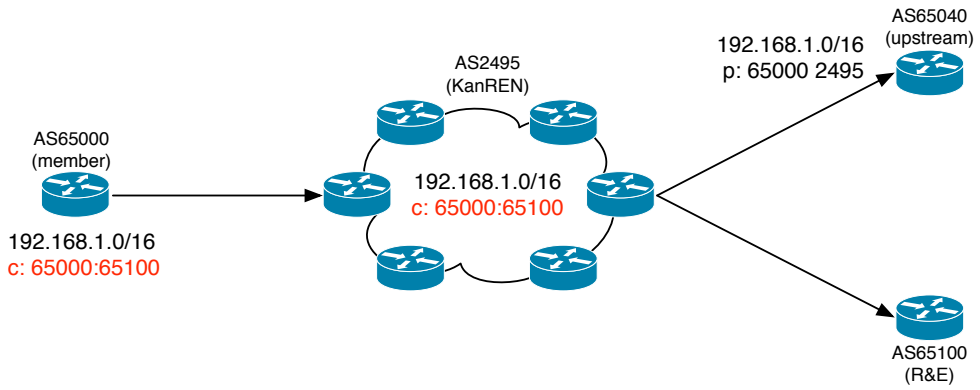
KanREN supports several Route Propagation BGP community strings for member prefixes. KanREN only accepts these community values on ingress from member BGP sessions. Lateral and upstream peers wishing access to KanREN's propagation BGP communities should communicate with KanREN via email to support@kanren.net.

The following community strings allow for propagation control of prefixes learned from member sessions

Value	Description
2495:1010	Do not advertise to upstream peers. (ex: Cox, Cogent, Hurricane Electric, CPS, TR-CPS)
2495:1020	Do not advertise to lateral peers. (ex: Nextech, KC Peering Exchange)
2495:1030	Do not advertise to commercial private peers. (ex: Google)
2495:1040	Do not advertise to other Kansas education networks. (ex: Kan-ed)
2495:1050	Do not advertise to State of Kansas agency networks. (ex: DISC)
2495:1060	Do not advertise to upstream R&E peers. (ex: Internet2, GPN)
2495:1070	Do not advertise to upstream non-commodity peers. (ex: CPS, TR-CPS)
2495:2000	Do not advertise outside as2495; ie: no-export
65000:####	Do not advertise to as#### – MUST be a direct peer to KanREN (as2495)

Route Propagation Use Case

In the diagram below a KanREN member advertises the 192.168.1.0/16 with a community value of 65000:65100. KanREN forwards the prefix to AS65040 but not AS65100; allowing per-AS propagation control.



Route Propagation Configuration Examples

Cisco IOS Example

```
!
ip prefix-list slash-16 seq 100 permit 192.168.1.0/16
!
route-map example_out permit 100
 match ip address prefix-list slash-16
 set community 65000:65100
!
router bgp 65000
!
neighbor 10.10.10.1 remote-as 2495
neighbor 10.10.10.1 description eBGP-to-KanREN
!
address-family ipv4
!
neighbor 10.10.10.1 route-map out example_out
!
```

Juniper JunOS Example

```
routing-options {
    autonomous-system 65000;
}
protocols {
    bgp {
        group kanren {
            description eBGP-to-KanREN;
            type external;
            export example_out;
            peer-as 2495;
            neighbor 10.10.10.1;
        }
    }
}
policy-options {
    prefix-list slash-16 {
        192.168.1.0/16;
    }
    policy-statement example_out {
        term accept-routes {
            from {
                prefix-list-filter slash-16;
            }
            then {
                community set no-65100;
                accept;
            }
        }
    }
    community no-65100 members 65000:65100;
}
```

Traffic Engineering Features

KanREN offers BGP-speaking members access to several Traffic Engineering (TE) features. KanREN's service offering includes a combination of BGP communities and ingress MED support. Lateral and upstream peers do not have access to these features by default. Any lateral or upstream peer wishing access to KanREN's traffic engineering features should communicate via email to support@kanren.net.

Please Note: All prepend actions are taken AFTER **and** IN ADDITION to KanREN's prepend policy for the BGP peer in question.

For example, KanREN's routing policy may dictate an additional prepend to commodity ISPs. If the community string `2495:4041` (ie: global, single prepend) is attached to a prefix, the advertised AS path to any upstream ISP will be `65000+2495+2495+2495` while advertisements to lateral peers will be `65000+2495+2495`.

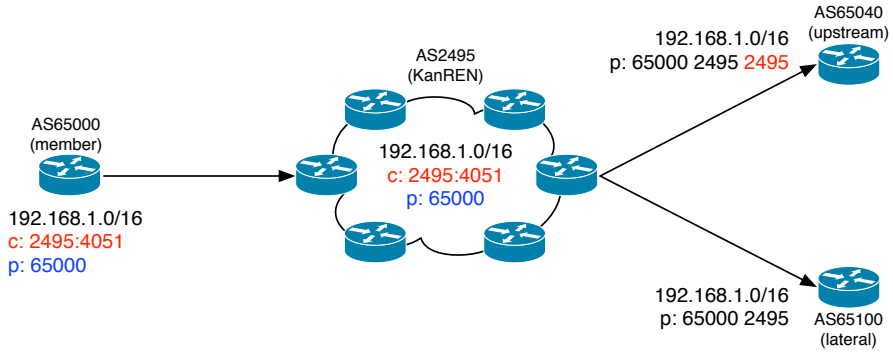
Community-Based TE

The following table outlines all traffic engineering BGP community strings:

Value	Description
2495:4041	Prepend as2495 1 time to all eBGP peers – Global egress prepend
2495:4042	Prepend as2495 2 times to all eBGP peers – Global egress prepend
2495:4043	Prepend as2495 3 times to all eBGP peers – Global egress prepend
2495:4044	Prepend as2495 4 times to all eBGP peers – Global egress prepend
2495:4051	Prepend as2495 1 time to all upstream peers
2495:4052	Prepend as2495 2 times to all upstream peers
2495:4053	Prepend as2495 3 times to all upstream peers
2495:4054	Prepend as2495 4 times to all upstream peers
2495:4061	Prepend as2495 1 time to all lateral peers
2495:4062	Prepend as2495 2 times to all lateral peers
2495:4063	Prepend as2495 3 times to all lateral peers
2495:4064	Prepend as2495 4 times to all lateral peers
2495:4071	Prepend as2495 1 time to all downstream peers
2495:4072	Prepend as2495 2 times to all downstream peers
2495:4073	Prepend as2495 3 times to all downstream peers
2495:4074	Prepend as2495 4 times to all downstream peers
2495:80	Set local preference to 80 (lower than default of 100)
2495:120	Set local preference to 120 (higher than default of 100)
2495:450	Set local preference to 450 (higher than member default of 400)
2495:900	Set local preference to 900 (higher than all other options)

Community-Based TE Use Case

In the example below, a KanREN member advertises `192.168.1.0/16` with a community value of `2495:4051`. KanREN accepts the route into its global routing table with an AS path of `2495+65000`. An additional `2495` AS hop (prepend) is added to advertisements to upstream networks but not other networks.



This allows fine grained control over how traffic destined to the member network is delivered to the KanREN network. Multiple BGP community strings can be attached to a single prefix announcement; allowing for rich Traffic Engineering capabilities.

Community-Based TE Configuration Examples

Cisco IOS Example

```
!
ip prefix-list slash-16 seq 100 permit 192.168.1.0/16
!
route-map example_out permit 100
 match ip address prefix-list slash-16
 set community 2495:4051
!
router bgp 65000
!
neighbor 10.10.10.1 remote-as 2495
neighbor 10.10.10.1 description eBGP-to-KanREN
!
address-family ipv4
!
neighbor 10.10.10.1 route-map out example_out
!
```

Juniper JunOS Example

```
routing-options {
    autonomous-system 65000;
}
protocols {
    bgp {
        group kanren {
            description eBGP-to-KanREN;
            type external;
            export example_out;
            peer-as 2495;
            neighbor 10.10.10.1;
        }
    }
}
```

```
policy-options {
    prefix-list slash-16 {
        192.168.1.0/16;
    }
    policy-statement example_out {
        term accept-routes {
            from {
                prefix-list-filter slash-16;
            }
            then {
                community set no-upstreams;
                accept;
            }
        }
    }
    community no-upstreams members 2495:4051;
}
```

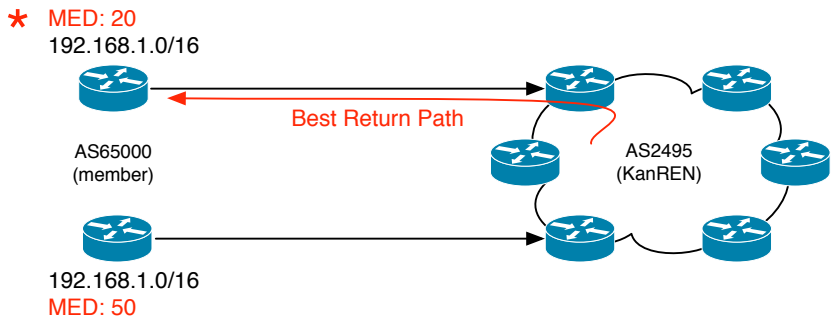
MED-Based Traffic Engineering

KanREN can accept and honor ingress Multi-Exit Discriminators (MEDs) values from members with multiple BGP sessions into the KanREN network. MED support is not configured by default but can be added upon request. The use of MEDs allows fine-grained control over KanREN's path selection across multiple BGP sessions. The use of MEDs will ONLY impact the path selection for traffic destined to the member's network. When comparing candidate BGP paths, the KanREN network considers paths with no MED value as least preferred.

KanREN suggests using BGP community signaled local preference values instead of BGP MEDs. MEDs can be valuable when extremely fine grained path selection control is required; however, most KanREN members will not require such fine grained control due to the limited number of prefixes most members announce. In the event that local preference signaling does not meet your site's needs, please contact KanREN via email to support@kanren.net and request MED support be added to your peering session(s).

MED TE Use Case

If a member is to announce many prefixes (100+) and requires extremely fine grained control over KanREN's best path selection process, use of MEDs in conjunction with various community-based TE tactics can be advantageous.



In this example, the member network (AS65000) originates the prefix `192.168.1.0/16` from two separate local routers to two separate KanREN routers. KanREN selects the advertisement with the lowest MED value and all traffic destined to the prefix is delivered to the top router.

MED TE Configuration Examples

Cisco IOS Example

```
!
ip prefix-list slash-16 seq 100 permit 192.168.1.0/16
!
route-map example-top_out permit 100
 match ip address prefix-list slash-16
 set metric 20
!
route-map example-bottom_out permit 100
 match ip address prefix-list slash-16
 set metric 50
!
router bgp 65000
!
neighbor 10.10.10.1 remote-as 2495
neighbor 10.10.10.1 description TOP-eBGP-to-KanREN
neighbor 10.10.80.1 remote-as 2495
neighbor 10.10.80.1 description BOTTOM-eBGP-to-KanREN
!
address-family ipv4
!
neighbor 10.10.10.1 route-map out example-top_out
neighbor 10.10.80.1 route-map out example-bottom_out
!
```

Juniper JunOS Example

```
routing-options {
    autonomous-system 65000;
}
protocols {
    bgp {
        group kanren-top {
            description eBGP-to-KanREN;
            type external;
            export example-top_out;
            peer-as 2495;
            neighbor 10.10.10.86;
        }
        group kanren-top {
            description eBGP-to-KanREN;
            type external;
            export example-bottom_out;
            peer-as 2495;
            neighbor 10.10.80.1;
        }
    }
}
policy-options {
    prefix-list slash-16 {
        192.168.1.0/16;
    }
    policy-statement example-top_out {
        term accept-routes {
            from {
                prefix-list-filter slash-16;
            }
```

```
                then {
                    metric 20;
                    accept;
                }
            }
        }
        policy-statement example-bottom_out {
            term accept-routes {
                from {
                    prefix-list-filter slash-16;
                }
                then {
                    metric 50;
                    accept;
                }
            }
        }
    }
```

BGP Security Features

KanREN supports several security-focused features on member BGP sessions. Some of the features are engineered to protect the BGP session itself while others are engineered to protect network hosts.

BGP MD5 Passphrase Support

KanREN supports but does not require MD5 passphrase on eBGP sessions. Sites may request MD5 passphrase support on BGP sessions either before or after session establishment. Please note that adding MD5 passphrase support to an existing, working BGP session will require the session be hard reset. Adding MD5 passphrase support to a BGP session should only be done during a scheduled maintenance window.

Members desiring MD5 passphrase support on new or existing BGP sessions should communicate with KanREN via email to support@kanren.net. KanREN will request the member select the passphrase that meets member security policy needs; KanREN does not dictate a passphrase minimum length or complexity.

BGP MD5 Passphrase Configuration Examples

Cisco IOS Example

```
!
router bgp 65000
 neighbor 10.10.10.1 remote-as 2495
 neighbor 10.10.10.1 description eBGP-to-KanREN
 neighbor 10.10.10.1 password <phrase>
!
```

Juniper JunOS Example

```
routing-options {
    autonomous-system 65000;
}
protocols {
    bgp {
        group kanren {
            description eBGP-to-KanREN;
            type external;
            peer-as 2495;
            neighbor 10.10.10.1 {
```

```
                authentication-key "$9$WZzx-bsY4DiqTz/Au1EhVwsgJGUjH"; ##
SECRET-DATA
            }
        }
    }
}
```

Generalized TTL Security Mechanism Support

KanREN's core network supports Generalized TTL Security Mechanism (GTSM) as defined in RFC3682. GTSM is not configured by default for any peer groups. Any peer may request GTSM support be added to their session(s) with KanREN; however, adding the configuration will cause existing BGP sessions to reset. KanREN strongly suggests GTSM only be added to existing sessions during a coordinated, scheduled maintenance window. Member sites requesting GTSM support should communicate with KanREN via email to support@kanren.net.

Generalized TTL Security Mechanism (GTSM) changes default behavior of control traffic on eBGP sessions. By default, eBGP sessions generate TCP traffic with a TTL (IPv4) or hop count (IPv6) of 1. Enabling GTSM causes the router to generate BGP TCP control traffic with a TTL (or hop count) value of 255. Once GTSM is enabled on a BGP session, any TCP control traffic that arrives to the local router without a TTL (or hop count) value of 255 or 254 is dropped.

For eBGP multi-hop enabled sessions, GTSM runs the following calculation:

```
Expected TTL = (255) - (configured multi-hop)
```

Any TCP BGP control traffic arriving at the local router that does not match the Expected TTL is dropped.

Please check vendor implementations before configuring GTSM. Many vendors either lack GTSM support or the feature must be "hacked" into place. Other platforms perform GTSM checks **after** MD5 checks; with all work happening in software. This leaves many platforms open to attacks that may be amplified with GTSM configured (2 software tasks with GTSM compared to 1 software task without GTSM).

GTSM Example Configurations

Cisco IOS Example

```
!
router bgp 65000
neighbor 10.10.10.1 remote-as 2594
neighbor 10.10.10.1 description eBGP-to-KanREN
neighbor 10.10.10.1 ttl-security hops <number>
!
```

BGP Blackhole Routing

Destination-Based Blackhole Routing (DB-BHR) can be a useful mitigation technique during (D)DoS events. If a small number of local hosts are targeted by a Denial of Service attack, DB-BHR prefixes can be signaled to the KanREN network and all packets destined to the host(s) will be re-routed in hardware within the KanREN network. Signaling a DB-BHR prefix can help remove unnecessary load from end hosts, security devices, and software-based routers; allowing legitimate traffic to flow as expected. Any DB-BHR prefix will effectively eliminate all communications to the host in question from the KanREN network.

KanREN supports member-signaled DB-BHR advertisements that meet the following criteria:

a. The advertised network MUST be part of a RIR-assigned block belonging to the site in question
b. The advertised network must be between /28 and /32 in size for IPv4 and /64 to /128 in size for IPv6
c. The total count of DB-BHR prefixes does not exceed 100

KanREN strongly suggests signaling DB-BHR based upon /32 networks for IPv4 and /128 networks for IPv6. This provides the highest degree of control during security events.

Traffic destined to the DB-BHR host prefixes is redirected to a capture server on the KanREN backbone. This allows traffic be captured for later analysis and characterization of the (D)DoS event. Accounts on the capture server can be requested via email to support@kanren.net and should be arranged before the event.

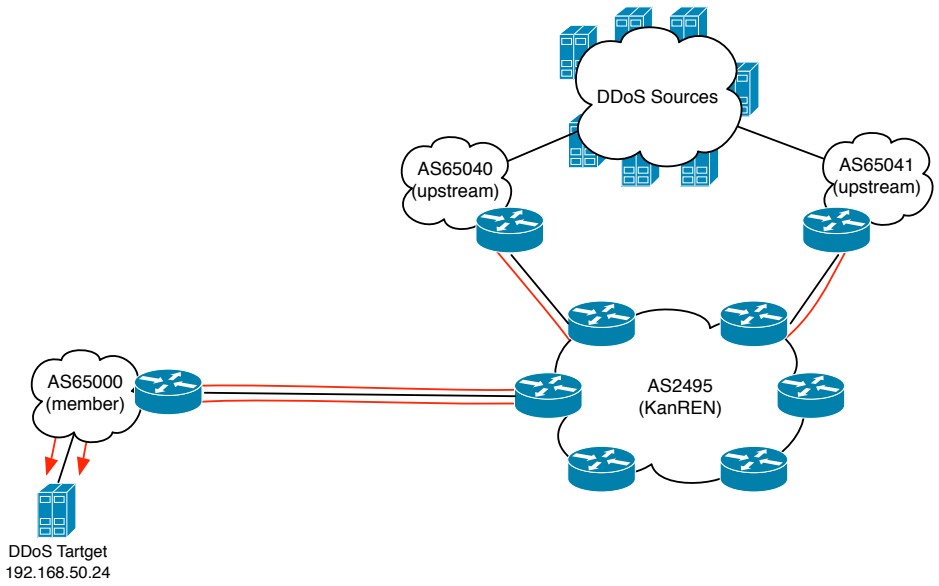
KanREN supports the following community strings for DB-BHR:

Value	Description
`2495:9999`	Signal local blackhole route – Prefix must be between /28 and /32
`2495:9998`	Signal local and remote blackhole route – Prefix must be between /28 and /32

Typically DB-BHR prefixes are not propagated beyond KanREN's AS borders, ie: KanREN does not allow such routes to leak into other member, lateral, or upstream networks. The `2495:9998` community string creates a DB-BHR within the KanREN network AND attempts to translate that value to lateral and upstream providers. Please note that `2495:9998` features are offered on a Best Effort basis; many lateral and upstream networks do not support such features.

DB-BHR Use Case

Destination-based Blackhole Routing (DB-BHR) can be an effective tool when used to fight (Distributed) Denial-of-Service attacks. The diagram below depicts a typical DoS attack scenario.



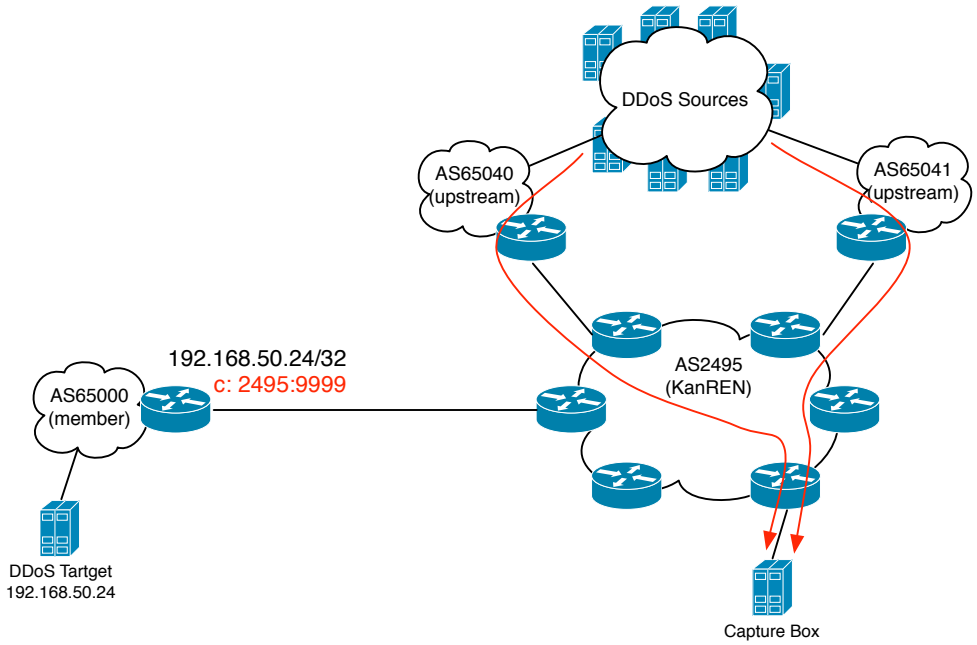
In this case, the host 192.168.50.24 is attacked with a large number of packets and/or traffic, sourced from multiple hosts. A surprisingly small DDoS attack can have a crippling effect on LAN services to many users. Packet and traffic level impact varies depending on internal architecture, but packet counts as low as 200,000 packets per second or traffic levels above 200Mbps can trigger wide-spread outages on the internal LAN for many users.

Sites with multiple, software-based systems may experience outages to all users connecting via the shared device (ex: all users behind a common firewall). Examples of typical software-based systems include:

a. Firewalls
b. IDP/IPS
c. Layer 4-7 aware traffic shapers
d. Load balancers
e. Individual, end hosts (desktops, laptops, printers, servers, etc)

DB-BHR can be signaled into the KanREN network to alleviate unnecessary load on shared software-based systems.

In the diagram below, a DB-BHR is signaled into the KanREN network, thus diverting all traffic destined to the targeted host to a capture server on the KanREN backbone.

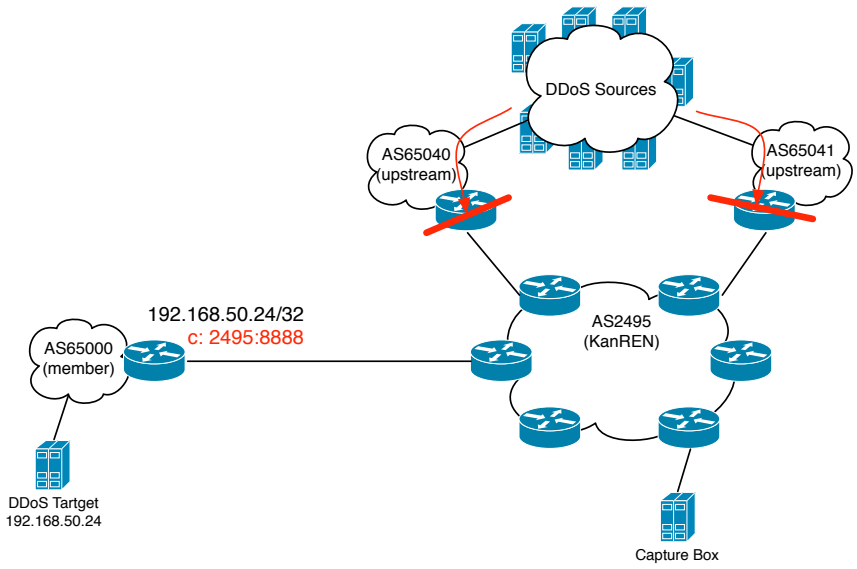


In the diagram below, a DB-BHR is signaled into the KanREN network with a community value of `2495:9999` (ie: local and remote DB-BHR). The result of the DB-BHR advertisement is all packets are stopped before event entering the KanREN network.

DB-BHR How-To

The following How-To assumes the following constants:

a. ARIN-assigned IP block in the example is 192.168.0.0/16
b. The LAN's internal core router is 192.168.10.1
c. ARIN-assigned ASN for the site is 300
d. Point-to-Point link with KanREN is 164.113.250.0/30
e. ARIN-assigned ASN for KanREN is 2495



Pre-Existing Configuration

```
!
interface GigabitEthernet0/0
 description Link to KanREN
 ip address 164.113.250.2 255.255.255.252
!
interface GigabitEthernet0/1
 description Link to LAN Core
 ip address 192.168.10.2 255.255.255.252
!
router bgp 300
 neighbor 164.113.250.1 remote-as 2495
 neighbor 164.113.250.1 description eBGP peering session with KanREN
 !
 address-family ipv4
  redistribute static
  neighbor 164.113.250.1 activate
  neighbor 164.113.250.1 route-map existing-route-map_in in
  neighbor 164.113.250.1 route-map existing-route-map_out out
  network 192.168.0.0 mask 255.255.0.0
 exit-address-family
!
ip route 192.168.0.0 255.255.0.0 192.168.10.1
!
ip prefix-list campus_aggregate seq 100 permit 192.168.0.0/16 le 24
!
ip prefix-list default_route seq 100 permit 0.0.0.0/0
!
route-map existing-route-map_out permit 10
  match ip address prefix-list campus_aggregate
!
route-map existing-route-map_in permit 10
  match ip address prefix-list default_route
!
```

Individual DB-BHR Configuration Steps

1. Build the basic architecture.
   We first need to prepare the existing BGP structure so that blackhole routes will not impact existing BGP advertisements.
   a. Build a prefix list to only allow very small network advertisements.
      KanREN will only accept blackhole routes for prefixes /28 or more specific (ie: /28 through /32). It is strongly suggested that blackhole routes be signaled with a length of /32 to avoid accidental service outages to hosts.

```
!
ip prefix-list blackhole_eligible seq 100 permit 192.168.0.0/16 ge 28 le 32
!
```

   b. Configure support for the community new format.
      BGP Communities originally did not support ##:## formatting. As of IOS 12.4(24), Cisco (and likely others) still defaults to the old style of BGP community format, thus providing seamless backwards compatibility. The KanREN network only supports the newer format and thus downstream peers utilizing Cisco routing platforms will need the following configuration line.

```
!
ip bgp-community new-format
!
```

   c. Update the outbound route map to allow matches to our new prefix list.
      Some users may choose to author completely new route maps to allow for easy fallback to an existing, known-good configuration. This step is not expressly required but is suggested

```
!
route-map new-route-map_out permit 10
 match ip address prefix-list campus_aggregate
!
route-map new-route-map_out permit 100
 match ip address prefix-list blackhole_eligible
 set community 2495:9999
!
```

   d. Begin sending KanREN BGP communities from the local network.
      By default, BGP does not send community values with its routing updates. Luckily, adding the feature is simple and does not result in a service outage.

```
!
router bgp 300
 address-family ipv4
  neighbor 164.113.250.1 send-community
!
```

   e. Filter redistribution into BGP.
      Filtering which static routes are redistributed into BGP can help eliminate extra "junk" in your BGP table. This is especially important for members running iBGP meshes internally. The following configuration will allow redistribution of the existing static route without any modification but will require new statics to be configured with a special tag.

```
!
route-map static-to-bgp permit 100
 match ip address prefix-list campus_aggregate
!
route-map static-to-bgp permit 200
 match ip address prefix-list blackhole_eligible
 match tag 9999
!
router bgp 300
 address-family ipv4
  redistribute static route-map static-to-bgp
!
```

f. Put new outbound route map into service.

NOTE: It is strongly suggested that this process be done during a maintenance window.

```
!
router bgp 300
 address-family ipv4
  no neighbor 164.113.250.1 route-map existing-route-map_out out
  neighbor 164.113.250.1 new-route-map_out out
!
```

2. Test configuration

Once all of the pieces are in place, it is suggested that some testing be performed. Locate a single IP on the LAN that is currently unused. We will signal a blackhole route and verify the route's presences in the KanREN network's core routing nodes.

a. Generate the route.

BGP needs a route to advertise. Our route map only allows potential routes with the following characteristics

i. the aggregate /16

ii. routes within the /16, between /28 and /32 in length, with a tag of 9999 – ie: BHR eligible

NOTE: The following will stop all traffic from the KanREN network from being forwarded to the host 192.168.80.1.

```
!
ip route 192.168.80.1 255.255.255.255 Null0 254 tag 9999
!
```

b. Call KanREN to verify our route table.

Additional DB-BHR Thoughts

This tutorial is extremely simple and makes assumptions that policy allows creation of static routes on border, BGP-speaking routing hardware. In many cases, border devices are maintained by a group of core network architects, engineers, and administrators. Often times, these groups are not responsible for security or aware of current security events. Some organizations may wish to deploy blackhole signaling via a non-core BGP speaker. This allows for a clear separation of responsibility.

Additionally, KanREN can generate a DB-BHR on a member's behalf from a secure, dedicated device on the KanREN core network. Any requests for KanREN-generated BHRs **must** come from the member's identified primary technical contact via either email to support@kanren.net or a phone call to the KanREN Network Operations Center. Members may also arrange for additional, pre-approved points of contact **before** an event. **Requests from non-approved sources will not be accepted.**

Members with multiple BGP sessions into the KanREN network may wish to deploy some features in their internal BGP architecture to provide more redundant services. If deployed properly, a blackhole route could be generated on either border router and signaled on both KanREN BGP sessions; thus making configuration simpler, faster, and more redundant.

Members redistributing static routes into an IGP (ie: OSPF, EIGRP, IS-IS, RIP, etc) should pay very close attention to where and how blackhole routes are generated. It is possible to inadvertently cut local access to a host if a route is injected that over-rides all other routes (ie: its more specific).

Members considering complex service deployments are encouraged to converse with KanREN about needs, goals, and potential solutions.

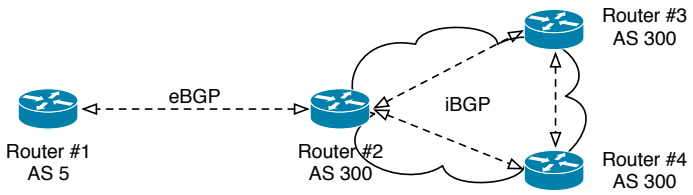
BGP General Concepts

What is BGP?

The Border Gateway Protocol (BGP) allows for scalable, feature-rich, protocol agnostic distribution of network reachability information between separate, autonomous networks. BGP is designed as a distance-vector protocol. Official BGP support started in June 1989 with the release of RFC1105. In March 1995, RFC1771 outlined BGP version 4 and added a major enhancement by allowing CIDR announcements. In January 2006 RFC4271 superseded RFC1771. BGP version 4 (BGP-4) is the current protocol suite deployed on networks worldwide.

eBGP versus iBGP

There are two different types of BGP peers; Exterior (eBGP) and Interior (iBGP). Peers with different ASNs are considered external while peers with the same ASN are considered internal. In the diagram below, Router #1 is eBGP peered with Router #2. Routers #2, #3, and #4 are iBGP peered. It can also be said that R1, R2, and R3 are “fully meshed”; meaning all routers have peering sessions with all other routers within the same autonomous system.



BGP Path Selection Process

As BGP receives updated route information, the protocol compares advertised paths of the same length using the following process:

- Does the locally configured ASN appear in the AS path? If yes, reject the route and end process.
- Is the advertised IP Next Hop reachable? If no, reject the route and end process.
- Prefer the path with the highest Local Preference.
- Prefer the path with the fewest AS hops.
- Prefer the path with the lowest origin type.
- Prefer the path with the lowest MED value.
- Prefer the path eBGP learned paths over iBGP learned paths.
- Prefer the path with the lowest IGP metric to the BGP next hop.
- Prefer the oldest path (ie: avoid route flapping).
- Prefer the path with the lowest router identification.
- Prefer the path with the lowest peer IP address.

Given the following two routes, #1 is selected because it is a “longer” prefix:

```
*  1
Prefix: 192.168.1.0/24
Local Pref: 200
AS Path: 65000 65005

   2
Prefix 192.168.1.0/16
Local Pref: 100
AS Path: 65000
```

In this case, the two prefixes are of different lengths. Additional data points are never compared.

Given the following three options, route #2 is selected:

```
   1
Prefix: 192.168.1.0/24
Local Pref: 200
AS Path: 65000
*  2
Prefix: 192.168.1.0/24
Local Pref: 600
AS Path: 65000 65000 65002 65001
   3
Prefix: 192.168.1.0/24
Local Pref: 100
AS Path: 65000 65001
```

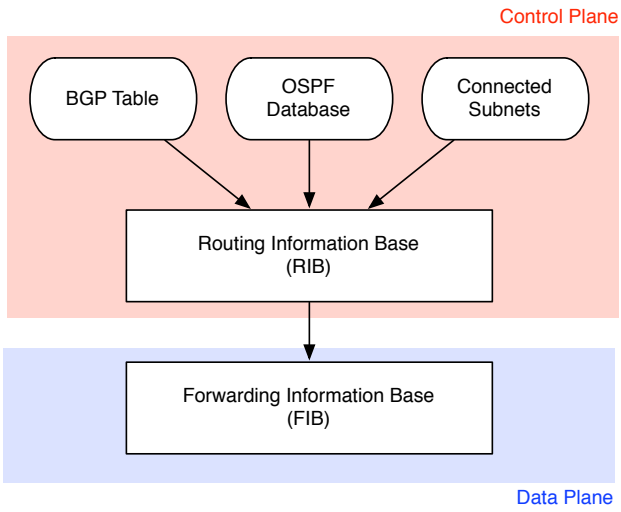
In this example, route #2 is selected eventhough its AS path is longer due to Local Preference.

BGP Table versus RIB versus FIB

A key component to understanding routing practice and operation is the distinction between a protocol-specific table (or database), a Routing Information Base (RIB), and the Forwarding Information Base (FIB). BGP maintains all routes received from all peers. When an advertisement is received and passes initial validity checks (next hop, AS path, etc) it is installed in the BGP table and sent to the RIB.

The RIB accepts candidate routes from many sources including: statically configured routes, directly connected routes, OSPF, IS-IS, BGP, etc. The RIB selects between candidate routes based upon vendor-specific processes. Router vendors use a weighting system for candidate routes learned from specific sources; Cisco and Juniper refer to this value as “Administrative Distance”. Once the RIB has selected a specific route, the information is sent to the FIB and packet routing is achieved.

It can also be helpful to consider BGP, OSPF, IS-IS, etc as part of the Control Plane while the FIB is part of the Data Plane. The RIB acts as the “glue” between the two discrete functions. The following diagram depicts a router with multiple protocols (route sources) all feeding into a single routing table (RIB). Many routers support separation of the individual protocols in additional to separation of the RIB (ie: VRF), allowing for extremely flexible routing service design.



Consider the following router CLI output:

```
telnet@example#show ip bgp routes 192.168.130.66
Number of BGP Routes matching display condition : 1
Status A:AGGREGATE B:BEST b:NOT-INSTALLED-BEST C:CONFED_EBGP D:DAMPED
       E:EBGP H:HISTORY I:IBGP L:LOCAL M:MULTIPATH m:NOT-INSTALLED-MULTIPATH
       S:SUPPRESSED F:FILTERED s:STALE
       Prefix             Next Hop        MED        LocPrf     Weight Status
1      192.168.128.0/19   192.168.199.2   0          100        0      BI
         AS_PATH:
       Last update to IP routing table: 39d10h57m14s, 1 path(s) installed:
       Route is advertised to 13 peers:
        192.168.200.242(2496)   192.168.200.246(2496)   192.168.200.22(30314)
        192.168.199.100(2495)   192.168.199.101(2495)   192.168.199.103(2495)
        192.168.199.104(2495)   192.168.199.105(2495)   192.168.199.107(2495)
        192.168.199.108(2495)   192.168.199.110(2495)   192.168.199.111(2495)
        24.124.7.237(12015)
telnet@example#
telnet@example#
telnet@example#show ip route 192.168.130.66
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
BGP  Codes - i:iBGP e:eBGP
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
       Destination        Gateway         Port        Cost          Type Uptime
1      192.168.130.64/26  192.168.200.10  ve 603      110/1001      O    13d12h
```

In the CLI output above, differences between the BGP table and RIB can be observed. In both cases, information for the destination host `192.168.130.66` is displayed. In this case, the BGP table has evaluated and installed the prefix `192.168.128.0/19` (as well as advertised the prefix to 13 other peers); however, the RIB installed a route learned via OSPF.

Appendix

Member BGP Communities Reference Table

Value	Description
2495:1010	Do not advertise to upstream peers. (ex: Cox, Cogent, Hurricane Electric, CPS, TR-CPS)
2495:1020	Do not advertise to lateral peers. (ex: Nextech, KC Peering Exchange)
2495:1030	Do not advertise to commercial private peers. (ex: Google)
2495:1040	Do not advertise to other Kansas education networks. (ex: Kan-ed)
2495:1050	Do not advertise to State of Kansas agency networks. (ex: DISC)
2495:1060	Do not advertise to upstream R&E peers. (ex: Internet2, GPN)
2495:1070	Do not advertise to upstream non-commodity peers. (ex: CPS, TRCPS)
2495:2000	Do not advertise outside as2495; ie: no-export
65000:####	Do not advertise to as### – MUST be a direct peer to KanREN (as2495)
2495:4041	Prepend as2495 1 time to all eBGP peers – Global egress prepend
2495:4042	Prepend as2495 2 times to all eBGP peers – Global egress prepend
2495:4043	Prepend as2495 3 times to all eBGP peers – Global egress prepend
2495:4044	Prepend as2495 4 times to all eBGP peers – Global egress prepend
2495:4051	Prepend as2495 1 time to all upstream peers
2495:4052	Prepend as2495 2 times to all upstream peers
2495:4053	Prepend as2495 3 times to all upstream peers
2495:4054	Prepend as2495 4 times to all upstream peers
2495:4061	Prepend as2495 1 time to all lateral peers
2495:4062	Prepend as2495 2 times to all lateral peers
2495:4063	Prepend as2495 3 times to all lateral peers
2495:4064	Prepend as2495 4 times to all lateral peers
2495:4071	Prepend as2495 1 time to all downstream peers
2495:4072	Prepend as2495 2 times to all downstream peers
2495:4073	Prepend as2495 3 times to all downstream peers
2495:4074	Prepend as2495 4 times to all downstream peers
2495:80	Set local preference to 80 (lower than default of 100)
2495:120	Set local preference to 120 (higher than default of 100)
2495:450	Set local preference to 450 (higher than member default of 400)
2495:900	Set local preference to 900 (higher than all other option)
2495:9999	Signal local blackhole route – Prefix must be between /28 and /32
2495:99998	Signal local and remote blackhole route – Prefix must be between /28 and /32

* Remote blackhole route translation offered on a best-effort basis. Most peers do not support this feature.

KanREN Internal BGP Communities Reference Table

The values in the table below are set by KanREN routing policy on ingress to the KanREN network. Members are not allowed to signal these values. These community values will be removed from any received route for subsequent prefix processing.

Value	Description
2495:1	Default route learned from upstream(s).
65100:100	Learned from a commodity Internet service provider OR commercial peering service. (ex: Cox, CPS)
65100‘200	Learned from a lateral OR direct commercial peer. (ex: Nextech, Google)
65100:300	Learned from an upstream R&E peer. (ex: Internet2, GPN)
65100:400	Learned from a KanREN member. (ex: KU, KSU)
65100:500	Reserved for future use.
65100:600	Learned from Kan-ed KAP network – KAP requesting l2 access.
65100:700	Learned from Kan-ed KAP network – KAP not requesting l2 access.
65200:*	Route learned from * ASN.
65301:1	Leak route to Google Global Cache in Lawrence.
65302:1	Leak route to Netflix cache in Lawrence.
65303:1	Leak route to Akamai CDN cache in Topeka.
65303:2	Leak route to Akamai CDN cache in Wichita.

By default, KanREN does not send communities to member organizations. Any member that would like to receive community values with routing updates should contact the KanREN NOC via email to support@kanren.net.

RFC Support

RFC 3392 – Capabilities Advertisement with BGP-4, November 2002

RFC 4271 – A Border Gateway Protocol 4 (BGP-4), January 2006

RFC 4760 – Multiprotocol Extensions for BGP-4, January 2007

RFC 4893 – BGP Support for Four-octect AS Number Space, May 2007

RFC 5082 – The Generalized TTL Security Mechanism, October 2007

RFC 5292 – Address-Prefix-Based ORF for BGP-4, August 2008

Glossary

- 4-Byte ASN – An extended capability for BGP-4 speakers that increases AS number space from 65,536 to 4,294,967,296 potential AS numbers.
- Address Family Identifier – Portion of BGP update messages that indicate what type of prefix is contained within the update. Examples: IPv4, IPv6, Appletalk, IPX.
- AS – Autonomous System. A group of devices with common or complementary administration, policy, and configuration. Typically the separation is made on an organization boundary (KanREN, KU, Internet2, etc).
- ASN – Autonomous System Number. Globally unique number assigned by a RIR to a specific AS or pulled from the range of “private” numbers.
- BGP – Border Gateway Protocol. Throughout this document, BGP will be used to refer to the suite of RFCs commonly referred to as BGP-4.
- CIDR – Classless Inter-Domain Routing. Allows for subnetting of larger networks beyond the original, classful designations by moving subnet mask boundaries from the byte-level to the bit-level. Sometimes referred to as “slash notation”.
- MED – Multi-Exit Discriminator. MEDs can indicate to external ASes a more or less preferred BGP path. MEDs allow for fine grained control over path selection of remote ASes.
- MP-BGP – Multiprotocol BGP. Extensions to BGP allowing services other than IPv4 unicast. Defined in RFC2858.
- NLRI – Network Layer Reachability Information. Allows BGP to tell neighbors about learned prefixes, their length, the IP next-hop, and other attributes (ie: community values, local preference, MEDs, etc). NLRI are unique to BGP version 4. Exchanged during update messages.
- Subsequent Address Family Identifier – An 8 bit value, signaled as part of NLRI update messages that allows for further segmentation of NLRI by dividing into Unicast, Multicast, or Unicast+Multicast groupings.

Change Log

Version 1.2 – May 2016 – Updated BHR communities. Fixed typos. Updated list of internal communities.

Version 1.1 – May 2011 – Fixed a few spelling and grammatical problems throughout.

Version 1.0 – January 2011 – Initial Release