



## Routing and RFC-1998

AKA using BGP communities to influence routing

# Routing and RFC-1998

Jason Sinclair and Bob Purdon

PIPE Networks Ltd

## What is RFC-1998?

**RFC-1988 discusses the use of the BGP community attribute to provide hints to upstream providers on how you wish to have your traffic routed.**

**RFC-1988 provides a short example of manipulating local preference on an upstream provider's network based on the community attribute. In the real world, some providers do not support this functionality at all, and some offer quite rich functionality.**

**This can be handy for networks that are multi-homed and wish to influence the distribution of traffic over multiple upstream circuits.**

**From the upstream provider's perspective, a standard configuration can support varying customer needs, and customers can adjust routing themselves without requiring reconfiguration at the provider end.**

## Who supports this stuff?

Over the last month or so the guys at PIPE have been researching various transit providers to see who supports what...

Of those we researched, the following support the use of the BGP community attribute by customers:

Verizon - AS703

NTT - AS2914

VSNL - AS6453

Powertel - AS9837

Uecomm - AS10223

AAPT - AS2764

Vocus - AS4826

Optus - AS7474

Pacnet - AS10026

There are no doubt others that we didn't research.

## Who supports this stuff?

**A couple of providers didn't respond to us.**

**A couple cited a lack of customer demand for not implementing RFC-1998 style support (PIPE being one, although we're likely to do it for completeness in the not too distant future).**

**One cited added difficulty in capacity planning.**

**A couple gave no specific reason - they just don't support it ☹**

## Who supports this stuff?

### **Verizon - AS703**

**Supports altering of LOCAL\_PREF**

**Supports AS Prepend**

**Supports route black holing**

**Supports keeping prefix within regional network only.**

**Supports keeping prefix within entire network only.**

### **AAPT - AS2764**

**Supports altering of LOCAL\_PREF**

**Supports keeping prefix within local pop only.**

**Supports keeping prefix within regional network only.**

**Supports keeping prefix within entire network only.**

## Who supports this stuff?

### **NTT - AS2914**

**Supports altering of LOCAL\_PREF**

**Supports AS prepend**

**Supports selective prefix advertisement to peer.**

**Supports keeping prefix within entire network only.**

### **Vocus - AS4826**

**Supports AS prepend**

**Supports route black holing**

**Supports selective prefix advertisement to peer.**

## Who supports this stuff?

### **VSNL - AS6453**

**Supports altering of LOCAL\_PREF**

**Supports route black holing**

**Supports keeping prefix within entire network only.**

### **Optus - AS7474**

**Supports altering of LOCAL\_PREF**

### **Powertel - AS9837**

**Supports keeping prefix within entire network only.**



## Who supports this stuff?

### **Pacnet - AS10026**

**Supports altering of LOCAL\_PREF**

**Supports AS prepend**

**Supports selective prefix advertisement to peer.**

**Supports keeping prefix within entire network only.**

### **Uecomm - AS10223**

**Supports altering of LOCAL\_PREF**

**Supports route black holing**

## What does it look like?

### Taking Vocus as an example:

|           |                                       |
|-----------|---------------------------------------|
| 4826:1000 | Export to All                         |
| 4826:2100 | Export to Level 3                     |
| 4826:2101 | Prepend once to Level 3               |
| 4826:2102 | Prepend twice to Level 3              |
| 4826:2103 | Prepend 3x to Level 3                 |
| 4826:2108 | Set No-Export to Level 3 announcement |
| 4826:2109 | Do not Export to Level 3              |
| 4826:2200 | Export to Verizon                     |
| 4826:2201 | Prepend once to Verizon               |
| 4826:2202 | Prepend twice to Verizon              |
| 4826:2203 | Prepend 3x to Verizon                 |
| 4826:2208 | Set No-Export to Verizon announcement |
| 4826:2209 | Do not Export to Verizon              |
| 4826:666  | Black hole                            |

## How do I use this stuff?

**A pretty simplistic example:**

```
router bgp 24130
  network 192.168.1.0
  network 192.168.2.0
  network 192.168.3.0
  neighbor x.x.x.x remote-as 4826
  neighbor x.x.x.x route-map AS4826-EXPORT out
  neighbor x.x.x.x send-community
```

**We're going to advertise 3 x /24's to our upstream. The route-map is where the real work gets done...**

## How do I use this stuff?

The prefix-list will determine which prefixes we wish to set the community on:

```
ip prefix-list AS4826-PREFIXES-OUT-PREPEND seq 10 permit 192.168.1.0/24
ip prefix-list AS4826-PREFIXES-OUT-PREPEND seq 12 permit 192.168.2.0/24
ip prefix-list AS4826-PREFIXES-OUT-PREPEND seq 99 deny 0.0.0.0/0 le 32
```

The AS-path access list is how we'll permit everything else.

```
ip as-path access-list 1 permit .*
```

## How do I use this stuff?

First up we match against the prefix list. If we get a match, we set the communities required for our upstream (Vocus in this example) to prepend 3 times to both Level 3 and Verizon.

```
route-map AS4826-EXPORT permit 10
  match ip address prefix-list AS4826-PREFIXES-OUT-PREPEND
  set community 4826:2103 4826:2203
```

The next section of this route-map matches everything else and permits it un-altered.

```
route-map AS4826-EXPORT permit 20
  match as-path 1
```

## How do I use this stuff?

**A catch all at the end – doesn't do much in this example.**

```
route-map AS4826-EXPORT deny 30  
  match as-path 1
```

**This configuration will cause Vocus to prepend 3 times for 192.168.1.0/24 and 192.168.2.0/24, while leaving 192.168.3.0/24 untouched.**

**Of course in the real world you'd have better announcement filtering 😊**

## A community resource.

We're intending to put together a community resource, either hosted at PIPE, or on the AusNOG web site, that lists the communities supported by various providers within Australia.

If you represent a transit provider, and you support RFC-1998 like functionality (or you don't, and you want to be listed as such), then please e-mail [bobp@pipenetworks.com](mailto:bobp@pipenetworks.com) and we'll get your information added to the list.

## What else can I do with communities?

**This presentation has been all about sending community attributes to upstream providers, but of course you can use communities internally within your network.**

**On the transit network at PIPE we assign communities to prefixes we learn from upstream transit providers and customers. In doing so, at any point in our network we know where we learned a given prefix from, and by filtering, we can control where we advertise it to.**

**We can tell which transit provider a prefix came from, if we learned it from peering (and if so, from which state), and whether we learned it via BGP from a customer or as the result of a 'network' statement in the BGP config.**

**How you do this might be the subject of a future presentation ☺**





**Thank you**  
This concludes our presentation