using routing domains / routing tables in a production network

Peter Hessler
phessler@openbsd.org

OpenBSD

27 September, 2014
rtable vs rdomain

- rtable
  - alternate routing table, usable with the same interfaces
  - ip addresses cannot overlap
  - multiple rtables can belong to a single rdomain
  - can be used for Policy Based Routing
rtable vs rdomain

- **rdomain**
  - completely independent routing table instance
  - assign 10.0.0.1/16 a dozen times
  - interfaces can be assigned to only one rdomain at a time
  - how we 'know' which one incoming packets should use
  - rdomains always contain at least one rtable
- first added in OpenBSD 4.9, released October 2009
- initially was IPv4 only
- IPv6 support added in OpenBSD 5.5, released May 2014
vrf-lite vs full vrf

- vrf-lite
  - multiple routing tables
  - done by hand
  - very common in smaller enterprises
  - only needs a single system
  - ...where most of my experience comes from

- vrf
vrf-lite vs full vrf

- vrf-lite
- vrf
  - also known as 'mpls'
  - requires bgp, ldpd and large networks
  - most frequently used to connect multiple sites in a single network
caveats

- default routes for all the domains!
  - seriously
    - the ’do we have a valid route’ check happens *before* pf
  - very common mistake

- debugging can be painful

- which route will be used?

- but, how do we send (some) traffic to a different rdomain?
Simple setup

```sh
$ ifconfig re0 rdomain 1
$ ifconfig re0 10.0.0.10/16
$ ifconfig lo1 rdomain 1
$ ifconfig lo1 127.0.0.1/8
$ route -T 1 add default 10.0.0.1
$ route -T 1 exec /usr/sbin/sshd
```
Simple setup

$ ifconfig em0
em0: flags=28843<UP,BROADCAST,...> rdomain 1 mtu 1500
  lladdr 28:d2:44:ac:5d:59
  priority: 0
  media: Ethernet autoselect (none)
  status: no carrier
  inet 10.0.0.1 netmask 0xfffff0000 broadcast 10.0.255.255

$ ifconfig lo1
lo1: flags=28049<UP,LOOPBACK,...> rdomain 1 mtu 32768
  priority: 0
  groups: lo
  inet 127.0.0.1 netmask 0xff000000
Simple setup

```
$ netstat -T1 -rnf inet
```

Routing tables

Internet:

<table>
<thead>
<tr>
<th>Destination</th>
<th>Gateway</th>
<th>Flags</th>
<th>Prio</th>
<th>Iface</th>
</tr>
</thead>
<tbody>
<tr>
<td>default</td>
<td>10.0.0.1</td>
<td>GS</td>
<td>8</td>
<td>em0</td>
</tr>
<tr>
<td>10.0/16</td>
<td>link#1</td>
<td>C</td>
<td>4</td>
<td>em0</td>
</tr>
<tr>
<td>10.0.0.1</td>
<td>28:d2:44:ac:5d:59</td>
<td>HLl</td>
<td>1</td>
<td>lo0</td>
</tr>
<tr>
<td>10.0.255.255</td>
<td>link#1</td>
<td>HLb</td>
<td>1</td>
<td>em0</td>
</tr>
<tr>
<td>127.0.0.1</td>
<td>127.0.0.1</td>
<td>UH</td>
<td>4</td>
<td>lo1</td>
</tr>
</tbody>
</table>
pf.conf:

anchor "cust1.example.com" on rdomain 15 {
    block
    pass proto icmp
    pass proto tcp from any to any port 80
}
pass in on rdomain 2 rtable 4
pass out from 10.0.0.0/16 to any nat-to (egress) rtable 20
shared infrastructure (vrf-lite)

- very common
- just a management network
- two rdomains, one pipe
- backup servers
- monitoring
- etc
full vrf

- ldpd
  - label distribution protocol daemon
  - distributes mpls label mappings
- bgpd
  - distribute our networks over the mpls ”tunnel”
production: discovering pitfalls

- route -T 1 exec
- adding rdomain to an interface
- ftp-proxy
- source and destination rdomains matter
- ntpd
- on rdomain
route -T 1 exec
  - originally for testing and hacking, turned out to be very useful
  - recommended method to start a daemon in a second rdomain
  - ...except a few network tools and a limited number of daemons

- adding rdomain to an interface
- ftp-proxy
- source and destination rdomains matter
- ntpd
- on rdomain
production: discovering pitfalls

• route -T 1 exec
• adding rdomain to an interface
  • erases IP address config
  • vlan vs parent interface
  • carp
• ftp-proxy
• source and destination rdomains matter
• ntpd
• on rdomain
route -T 1 exec
adding rdomain to an interface
ftp-proxy
  sometimes, you simply want to ftp from *and* to different rdomains
source and destination rdomains matter
ntpd
on rdomain
route -T 1 exec
adding rdomain to an interface
ftp-proxy
source and destination rdomains matter
ntpd
  normal solution to needing services in a second rdomain? run the daemon again
  running a second ntpd to provide time? Holy clock-skew Batman!
on rdomain
route -T 1 exec
adding rdomain to an interface
ftp-proxy
source and destination rd domains matter
ntpd
on rdomain
  you want to match packets traveling on an rdomain
best practices

- default routes for all the things
  - as I said, real common mistake
- pf.conf tricks
- spend extra time in the planning stages
very special thanks

- henning@ for adding the multiple routing table support
- claudio@ writing the code and for putting up with all of my asinine questions when we first tested
- reyk@ for lots of work in bringing this into the tree and funding this via his (former) company
Questions?