Multihoming Without BGP

Paul Vixie <paul@vix.com>

Vixie Enterprises

Multihomed Server Implications

- Has only provider assigned address space.
- Needs multiple full routing tables.
- Depends on DNS “round robin”.
- Motives: performance or reliability?

Outbound Route Balancing

TCP is ACK-timed $\rightarrow$ path symmetry is good.

Symmetry in the intermediate system is really hard.

First hop symmetry is most of what you want anyway.

BGP path selection is hard to defeat.

Per-interface Default Route

- BSD TCP binds outbound route to PCB on SYN-ACK.
- Our trick: remember the inbound interface identity from the SYN.
- Each interface has its own “default route”.
- For outbound TCP and all UDP, a normal default is also needed.

netstat -r -n

Destination	Gateway	Flags	Refs	Use	Interface
default:de1	137.39.63.225	UGS	475	989069	de1
default:de2	204.74.120.1	UGS	243	294363	de2
default	137.39.63.225	UGS	63	51089	de1
127	127.0.0.1	UGRS	0	0	lo0
127.0.0.1	127.0.0.1	UH	11	643	lo0
137.39.63.224/27	link#2	UC	0	0	de1
137.39.63.225	0:0:c:35:29:a0	UHL	2	56	de1
137.39.63.255	link#2	UHL	0	1	de1
192.168.1	link#1	UC	0	0	de0
192.168.1.1	0:0:f8:2:b3:66	UHL	1	20	lo0
192.168.1.2	8:0:69:2:65:e7	UHL	13	116768	de0
192.168.1.255	link#1	UHL	1	36	de0
204.74.120/27	link#3	UC	0	0	de2
204.74.120.1	0:e0:14:3d:bd:1	UHL	1	0	de2
204.74.120.31	link#3	UHL	0	1	de2
224/8	link#1	UC	0	0	de0

Observed Results

- Incoming connections rotate per DNS round robin.
- Users stopped complaining about slow/hung connections.
- Providers like it better than sending a full BGP table.

Complications

- Used `socket` utility to forward `telnet` and `smtp`.
- Some `Squid` versions leak memory, need periodic restarts.
- A full NAT solution would be easier to manage.
- DEC PAIX may be the only place where this is useful.

Future Work

- Round robin for outgoing TCP.
- Do more thinking about UDP.
- Get a WWW server vendor interested.
- Probably a dead end for Vixie.

Resources

- In `ftp://ftp.vix.com/pub/vixie/ifdefault/` we see...
- ```
  -rw-rw-r-  1 vixie  vix  1731 Jan 30 22:15 ifconfig-diffs
  -rw-rw-r-  1 vixie  vix  5386 Jan 30 21:59 kernel-diffs
  -rw-rw-r-  1 vixie  vix  3696 Jan 30 22:23 netstat-diffs
  ```
- Developed, running in production on BSD/OS 2.1.
- Should port easily to other BSD derived systems.
- Use at your own risk, no warranty expressed or implied.
  ```