

IPv6 in TCD

David Malone

`<dwmalone@maths.tcd.ie>`

`<David.Malone@nuim.ie>`

10 June 2008

Introduction

- TCD: \approx 17000 students/staff.
- Focused on central campus.
- Network — General: ISS, specialised: CS, ...
- Maths: \approx 1000 students/staff.
- Traditional Unix service (15 servers, 60 xterms).
- Deployment 'plan' for IPv6.

Stage 1: Getting Started

```
% ifconfig -a
lo0: flags=8049<UP,LOOPBACK,RUNNING,MULTICAST> mtu 16384
    inet6 ::1 prefixlen 128
    inet6 fe80::1%lo0 prefixlen 64 scopeid 0x1
    inet 127.0.0.1 netmask 0xff000000
en0: flags=8863<UP,BROADCAST,SMART,RUNNING,SIMPLEX,MULTICAST> mtu 1500
    inet6 fe80::203:93ff:fe46:17a6%en0 prefixlen 64 scopeid 0x4
    inet 10.0.0.1 netmask 0xff000000 broadcast 10.255.255.255
    ether 00:03:93:46:17:a6
```

Pinging & ssh

```
% ping6 ::1
PING6(56=40+8+8 bytes) ::1 --> ::1
16 bytes from ::1, icmp_seq=0 hlim=64 time=0.392 ms
% ping6 -I en0 fe80::230:65ff:fe03:d972
16 bytes from fe80::230:65ff:fe03:d972%en0, icmp_seq=0 hlim=64 time=1.373 ms
% ping6 -I en0 ff02::1
PING ff02::1(ff02::1) from fe80::2b0:d0ff:fed7:741d en0: 56 data bytes
64 bytes from ::1: icmp_seq=1 ttl=64 time=0.062 ms
64 bytes from fe80::2b0:d0ff:fe05:fc06: icmp_seq=1 ttl=64 time=0.194 ms (DUP!)
64 bytes from fe80::206:5bff:fe68:249b: icmp_seq=1 ttl=64 time=0.224 ms (DUP!)
64 bytes from fe80::202:b3ff:fe65:604b: icmp_seq=1 ttl=64 time=0.256 ms (DUP!)
64 bytes from fe80::2b0:d0ff:fef4:c6c5: icmp_seq=1 ttl=64 time=0.334 ms (DUP!)
64 bytes from fe80::203:93ff:fe46:17a6: icmp_seq=1 ttl=64 time=0.384 ms (DUP!)
...
% ssh fe80::2b0:d0ff:fef4:c6c5%en1
dwmalone@fe80::2b0:d0ff:fef4:c6c5%en1's password:
```

As good as your Ethernet multicast.

Stage 2: IPv6 connectivity

- Originally used tunnel (Protocol 41).
- 2001:618:400:e::/64 address space from BT.
- Use ip6fw for filtering.

```
% ifconfig gif0 create
% ifconfig gif0 tunnel 134.226.10.51 193.113.58.80
% ifconfig gif0 inet6 2001:618:400::1:DC4:1467 2001:618:400::1:DC4:1466 prefixlen 128
% route add -inet6 default 2001:618:400::1:DC4:1466
% ifconfig fxp0 inet6 2001:618:400:e::1 prefixle 64
% ip6fw add 100 deny all from 2001:618:400:e::/64 to any in via gif0
% ip6fw add 200 allow all from 2001:618:400:e::/64 to any in via fxp0
% ip6fw add 300 allow all from fe80::/10 to any in via fxp0
% ip6fw add 400 allow all from :: to ff02::/16 in via fxp0
% ip6fw add 500 deny all from any to any in via fxp0
% sysctl net.inet6.ip6.forwarding=1
```

Rtadvd

- Configure prefix 2001:618:400:e::
- Node gets 2001:618:400:e:2a0:c9ff:feb1:34e7
- (We manually assign some host IDs
::134.226.81.3)
- Use `rtadvd fxp0` on FreeBSD.
- Advertises prefix, router, lifetimes, MTU...
- Can cause problems: Rapid prefix changes,
Misplugged routers...

Planning and routing

- Originally two tunnels to {cs,maths}.tcd.ie.
- Merged to produce Addressing plan:
 - 2001:770:10:000:: — infrastructure/tunnels,
 - 2001:770:10:100:: — ISS,
 - 2001:770:10:200:: — CS,
 - 2001:770:10:300:: — maths,
 - 2001:770:10:400:: — Netsoc,
 - 2001:770:10:500:: — HPC Centre.
 - 2001:770:10:600:: — ISS RFC 1918.
- Renumbered to 2001:770:10::/48.

Stage 3: Services

- To be useful need DNS info:
- A: salmon → 134.226.81.11.
- PTR: 11.81.226.134.in-addr.arpa → salmon
- AAAA: salmon.ipv6 → 2001:770:10:300::...
- PTR: ...0.7.7.0.1.0.0.2.ip6.arpa → salmon.ip6
- New reverse zone delegated from upstream.

Testing

- After updating known hosts, seems good.
- Further tests to before AAAA for normal names.
- IPv6 at home. Search `ipv6.maths.tcd.ie` first.
- All still good, add AAAA for normal names.
- Six years since added AAAA for `www.maths` — no problems!

Connecting at home

- Public IPv4 \Rightarrow IPv6.
- Two main options: tunnel or 6to4.
- Get a tunnel from `www.sixxs.net`.
- 6to4: `www.xxx.yyy.zzz` \rightarrow `2002:WWXX:YYZZ::`
- Eircom and Esat both offer experimental services.

Web Server

Upgrade to Apache2

- s/Port 80/Listen 80/
- Update virtual servers that give IP.
- Add IPv6 addrs to .htaccess files.
- Update log processing scripts.
- Initially some software not ready: mod_php, sendfile.

DNS & Bind 9 (or 8.4)

Having AAAA and doing DNS *over* v6 different.

```
+ listen-on { any; };
+ listen-on-v6 { any; };
  query-source address * port 53;
+ query-source-v6 address * port 53;
- acl tcd { 134.226.0.0/16; }
+ acl tcd { 134.226.0.0/16; 2001:770:10::/48; }
```

zone file:

```
; Master nameserver for maths.tcd.ie.
ns                IN      A       134.226.81.11
+                IN      AAAA   2001:770:10:300::86e2:510b
```

Glue in .tcd.ie and .ie zone.

In maths zone file is authoritative, autoconf hard.

CS generate zone file, autoconf has proven easy.

Mail

IPv6: sendmail, postfix, qmail (patch), exim.

We use MMDF — challenge by Dave Wilson.

- Receive mail from network.
- Look up addresses for MX.
- Send mail to network.

Working fine.

Adjusted Received headers for SpamAssassin.

Converting Apps

- Sockets pretty agnostic: `s / AF_INET / AF_INET6 /`
- Need to look up A and AAAA records.
- New functions `getaddrinfo` and `getnameinfo`.

```
struct addrinfo hints, *res, *res0;
int s;
memset(&hints, 0, sizeof(hints));
hints.ai_family = PF_UNSPEC;
hints.ai_socktype = SOCK_STREAM;
getaddrinfo("www.kame.net", "http", &hints, &res0);
for (res = res0; res; res = res->ai_next) {
    s = socket(res->ai_family, res->ai_socktype,
              res->ai_protocol);
    if (connect(s, res->ai_addr, res->ai_addrlen) < 0)
        continue;
    break;
}
freeaddrinfo(res0);
```

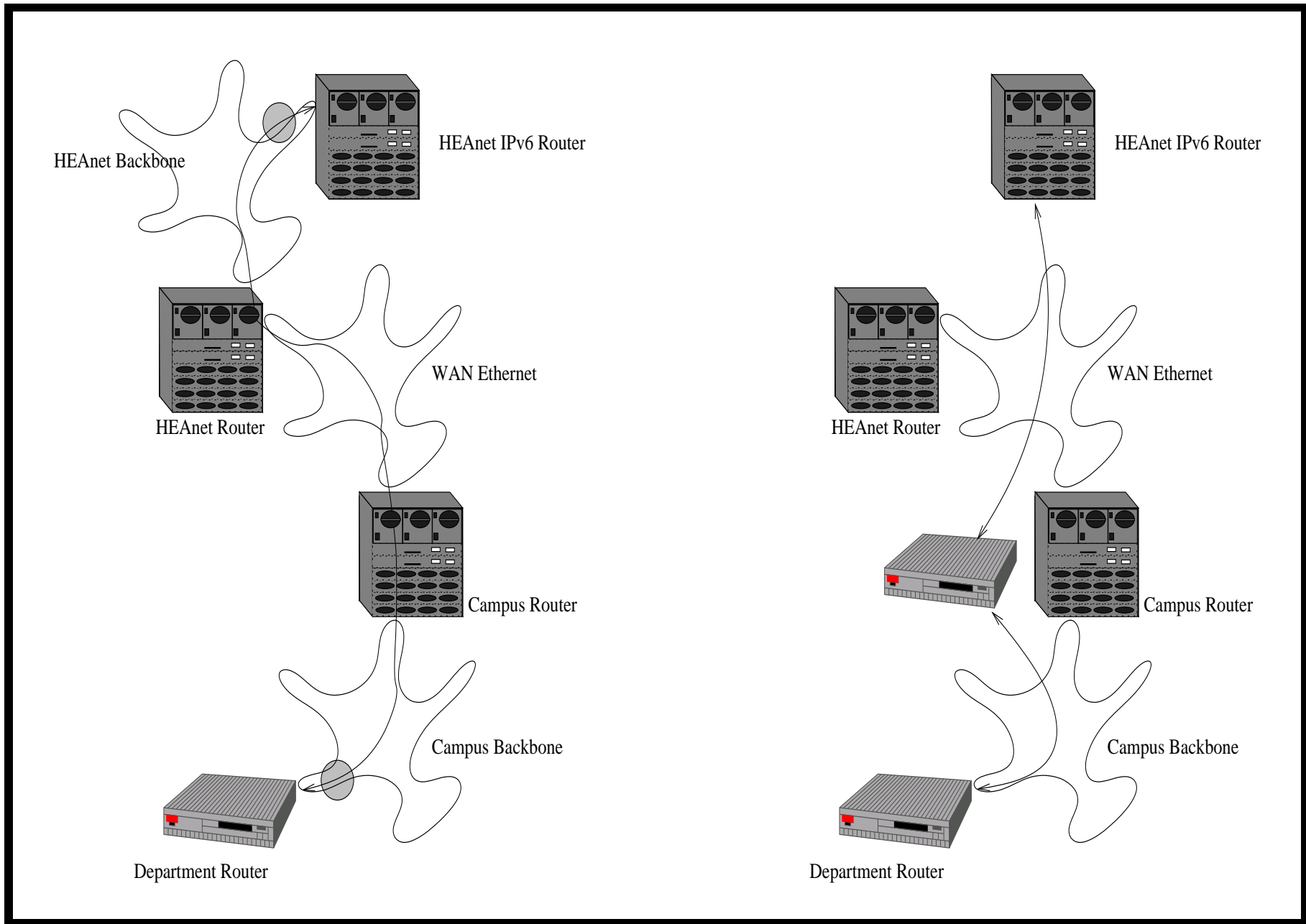
Other services

- NNTP: innd+patch.
- FTP: wu-ftpd+patch, tweaks for reverse DNS.
- ident, lpd, courier-imap.
- cvsup: netcat fudge.
- NTP/NFS: give -4 flag when no server support.
- WiFi: router(rtadv) + switch + APs.
Needed short lifetimes.

http://www.deepspace6.net/docs/ipv6_status_page_apps.html

Wait and see

- HTTP \approx 2000/d, SMTP \approx 300/d.
- Use multicast to monitor population.
- Increasing numbers: OS X, Linux, *BSD.
- Link moved to native Ethernet.
- Not doing IPv6 BGP yet.
- Static routes with ICMP redirect.



Campus IPv6 router

- Currently FreeBSD box.
- Basic IPv6 firewalling.
- IPv4 firewall drops all.
- Tunnels via internal dualstack.
- Static routing.
- Just works (modulo hardware).
- Move to Cisco 65XX.

6509 Plan

- Manually assign address to 6509.
- IPv6 on VLAN served by 6509.
- Bring up IPv6 on second 6509.
- Bring up IPv6 on both sides of PIX.
- Copy ACLs to PIX, enable routing.
- Finally route through PIX and 6509.
- Addressing: 134.226.xxx to 2001:770:10:1XX.
- Addressing: 10.xxx to 2001:770:10:6XX.

Stage 4: Future

- Basic deployment done, new services OK.
- Connecting various VLANs.
- Need to IPv6 finger and *web proxy*.
- Doubleclick now fixed!

- 2000/08/03** Maths upgraded to FreeBSD 4.1.
- 2000/08/09** I harass davew re IPv6 address.
- 2001/03/29** With mknell merge IPv6 networks.
- 2002/02/06** DNS/addresses allocated by HEAnet.
- 2002/06/04** AAAA added to production names.
- 2002/12/20** Natively over Gigabit Ethernet.
- 2003/07/21** AAAA glue in .ie zone.
- 2004/09/06** IPv6 capable .ie server.
- 2005/03/14** Google get a block of IPv6 addresses.

Interesting Events

- Protocol redundancy during SQLslammer!
- Weird etherchannel problems, IPv6 unaffected.
- Recovered correctly after network merge.
- Link-local & redirected traffic.
- Some X weirdness.
- Delay when IPv6 is down/unconfigured.

Summary

1. Link-local setup.
2. Connectivity, filtering, route advertisement.
3. Test DNS, test services, real DNS.
4. Integrate with other infrastructure.
5. Enjoy benefits and (occasional) quirks.