A case study of IPv6 deployment tcd.ie

David Malone

<dwmalone@{maths.tcd,cnri.dit}.ie>

22 May 2003

Introduction

- TCD: University $\approx$ 17000 students/staff.
- Focused on central campus.
- Maths: $\approx$ 1000 students/staff.
- Traditional Unix service (10 servers, 60 xterms).
- Deployment 'plan' for IPv6.

Example IPv6 Addresses

- 2001:0770:0010:0300:0000:0000:86e2:510b
- 2001:770:10:300:0:0:86e2:510b
- 2001:770:10:300::86e2:510b
- 2001:770:10:300::134.226.81.11

Special Addresses

::	Unspec
::1	localhost
fe80:: block	link-local
fec0:: block	site-local
ff00:: block	multicast

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

- Use 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
- Use `rtadvd fxp0` on FreeBSD.
- Advertises prefix, router, lifetimes, MTU...
- Can cause problems: Rapid prefix changes, Misplugged routers...

Problems with routing

- Scenic routing between {cs,maths}.tcd.ie.
- /48 from BT. Tunnel: BT to CS, CS to Maths.
- Addressing 'plan'
  2001:618:409:100:: — ISS,
  2001:618:409:200:: — CS,
  2001:618:409:300:: — maths,
  2001:618:409:400:: — reserved.
- Renumbered when allocated 2001:770:10::/48 by HEAnet.

Stage 3: Services

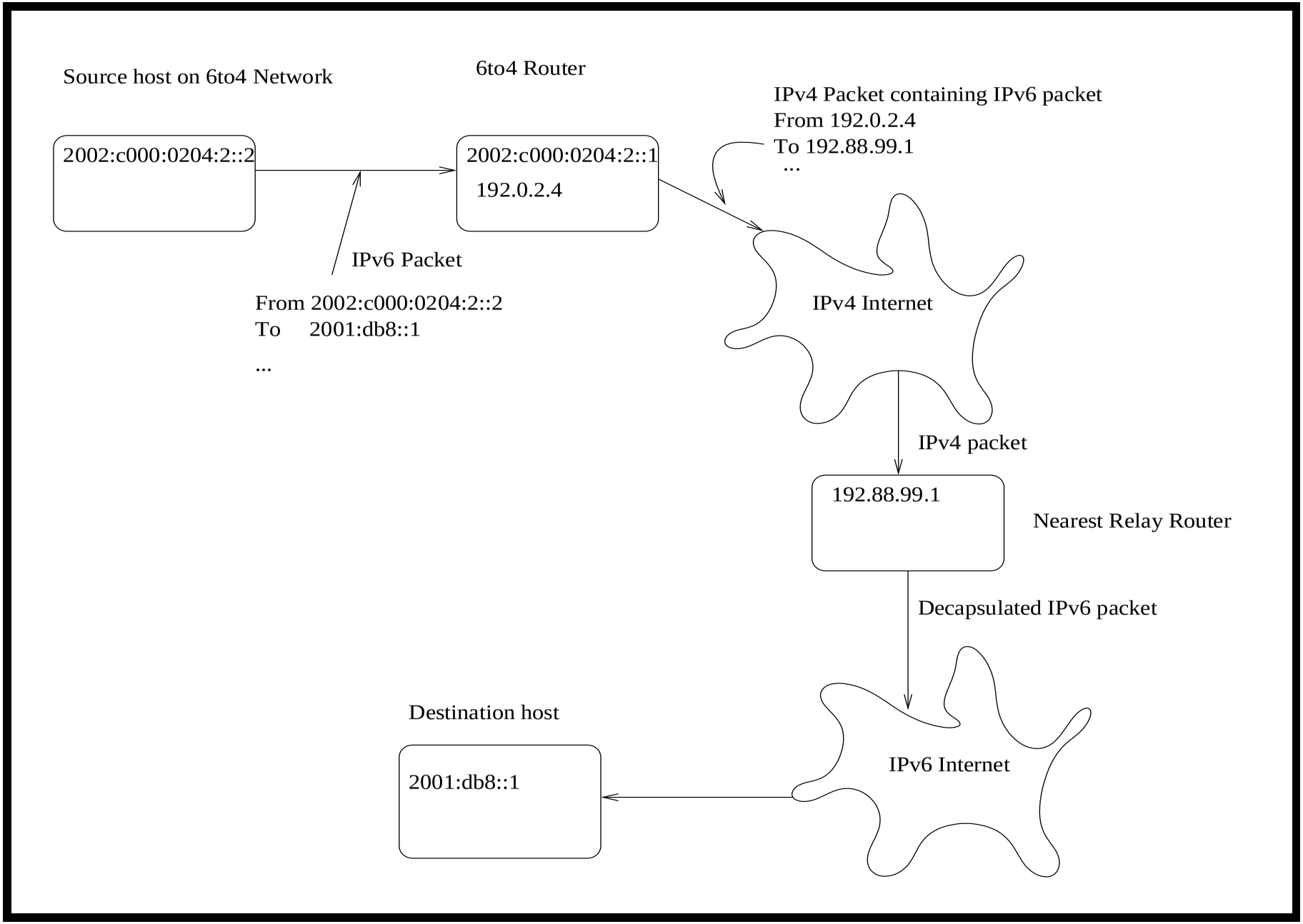
- To be useful need DNS info:
- A: salmon →134.226.81.11.
- PTR: 11.81.226.134.in-addr.arpa → salmon
- AAAA: salmon.ipv6 → 2001:770:10:300::...
- PTR: ...0.7.7.0.1.0.0.2.ip6.arpa → salmon.ip6
- New reverse zone delegated from upstream.

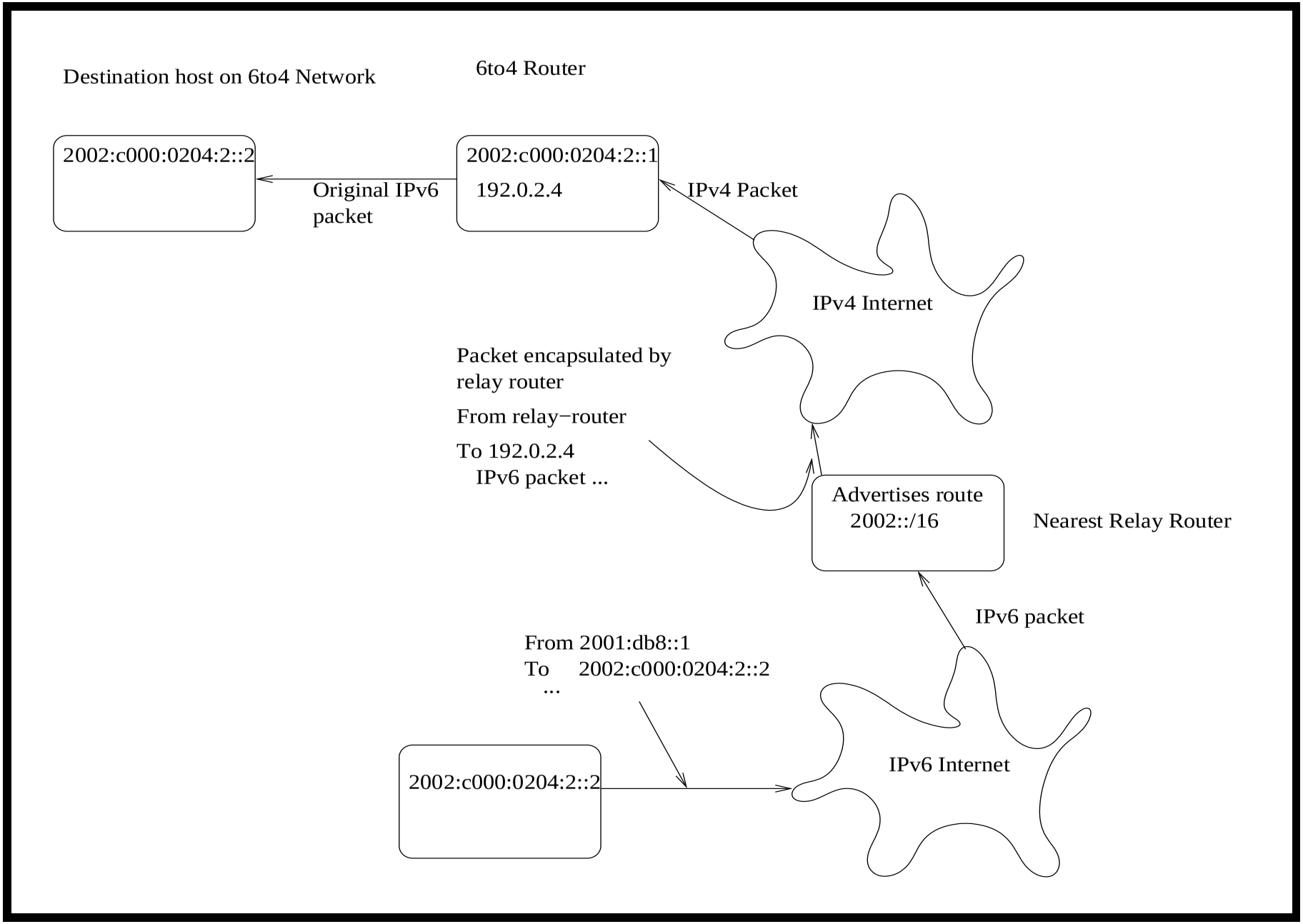
Testing

- 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.

Connecting at home

- 6to4 — very easy, easy config.
- www.xxx.yyy.zzz → 2002:WWXX:YYZZ::
- Eircom, Esat and HEAnet offer relays.
- Tunnels now available to those with fixed IPv4.


Source host on 6to4 Network
6to4 Router
IPv4 Packet containing IPv6 packet
From 192.0.2.4
To 192.88.99.1
...
2002:c000:0204:2::2
2002:c000:0204:2::1
192.0.2.4
IPv6 Packet
From 2002:c000:0204:2::2
To 2001:db8::1
...
IPv4 Internet
IPv4 packet
192.88.99.1
Nearest Relay Router
Decapsulated IPv6 packet
Destination host
2001:db8::1
IPv6 Internet


Destination host on 6to4 Network
6to4 Router
2002:c000:0204:2::2
2002:c000:0204:2::1
192.0.2.4
Original IPv6
packet
IPv4 Packet
IPv4 Internet
Packet encapsulated by
relay router
From relay−router
To 192.0.2.4
IPv6 packet ...
Advertises route
2002::/16
Nearest Relay Router
IPv6 packet
From 2001:db8::1
To 2002:c000:0204:2::2
...
2002:c000:0204:2::2
IPv6 Internet

Web Server

Upgrade to Apache2

- s/Port 80/Listen 80/
- Add IPv6 addrs to `.htaccess` files.
- Update log processing scripts.
- PHP problems? Run as CGI.
- Linux sendfile & checksum offloading.

DNS & Bind 9 (or 8.4)

Having AAAA and doing DNS *over* v6 different.

named.conf:

```
+       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. Autoconf?

Mail

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

We use MMDF — challenge by Dave Wilson.

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

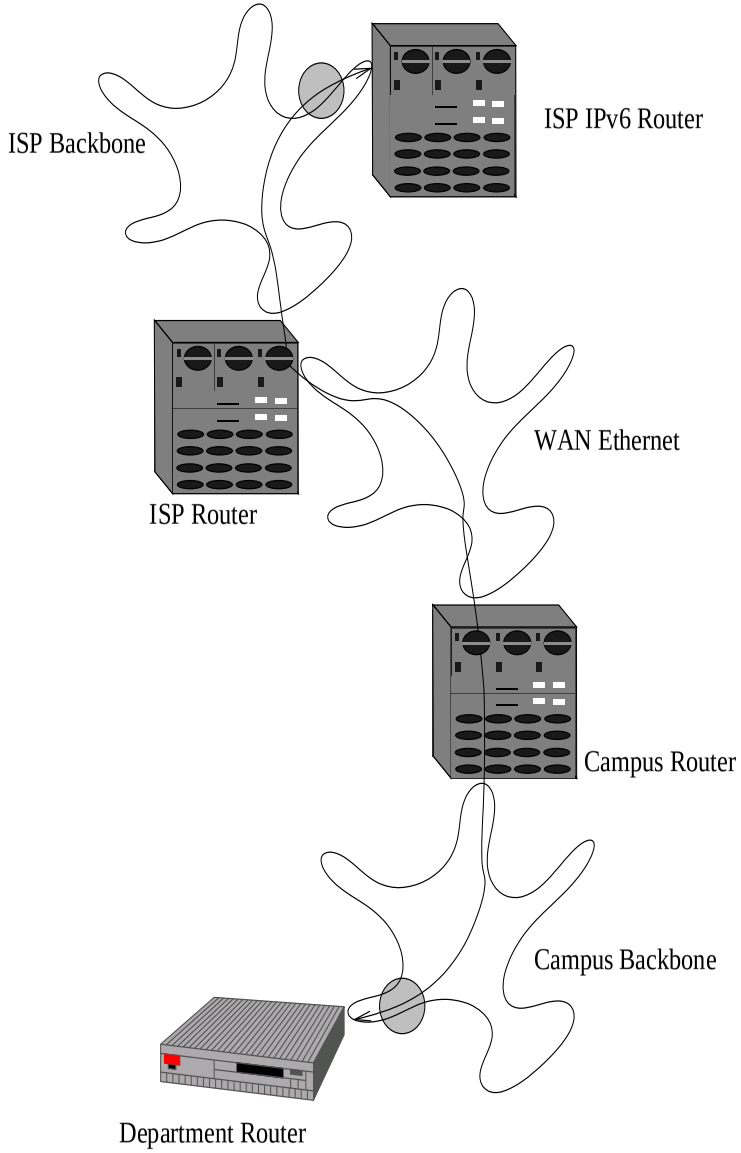
Converting Apps

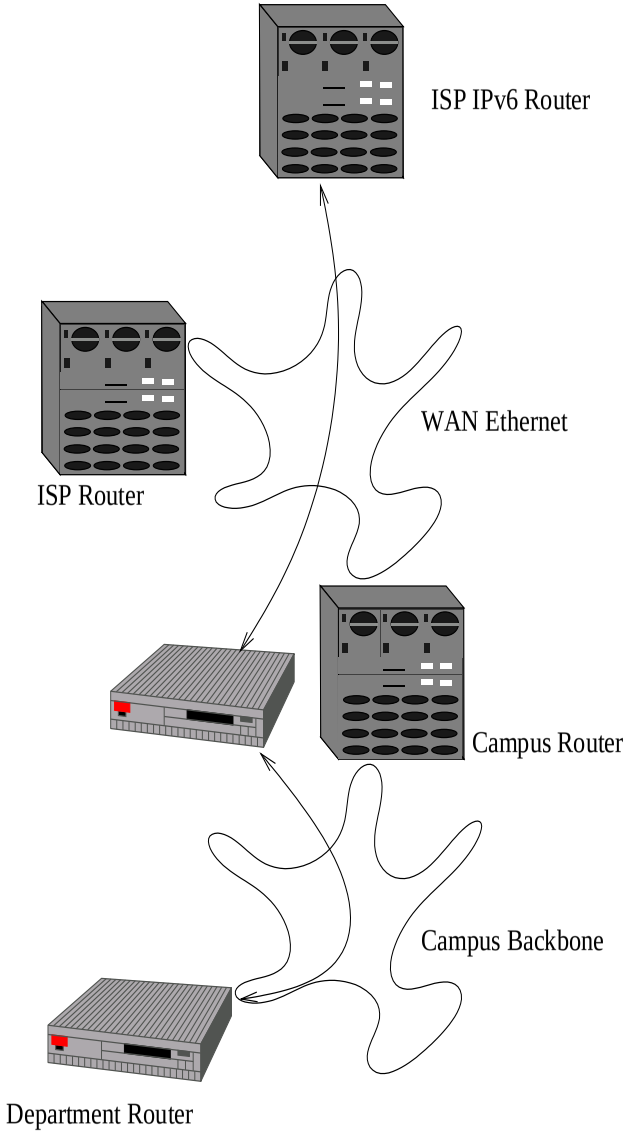
- Sockets pretty agnostic: `s/AF_INET/AF_INET6/`
- Need to look up A and AAAA records.
- New functions getaddrinfo and getnameinfo.

```
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);
```

Wait and see

- HTTP ≈500/d (max 96K!),
  SMTP ≈50/d each way.
- Use multicast to monitor population.
- Increasing numbers: OS X, Linux, *BSD.
- Link moved to native Ethernet.
- Protocol redundancy during SQLslammer!


ISP Backbone
ISP IPv6 Router
WAN Ethernet
ISP Router
Campus Router
Campus Backbone
Department Router


ISP IPv6 Router
WAN Ethernet
ISP Router
Campus Router
Campus Backbone
Department Router

Other services

- NNTP: innd+patch, Diablo-CVS.
- FTP: wu-ftpd+patch, tweaks for reverse DNS.
- ident, lpd, courier-imap.
- cvsup: netcat fudge.
- NTP: 4.1.80 problems with fall-back.
- WiFi: router + switch + APs.
  Need short lifetimes.

Stage 4: Future

- Working on college host.db, DNS, HTTP.
- Use FreeBSD/Linux vifs to connect all vlans?
- Need to IPv6 finger and *web proxy*.
- Bad doubleclick!
- Internal services: NFS, X11, phone, samba, snmp, diskless xterms...

Summary

1. link-local setup
2. connectivity, filtering, route advertisment.
3. test DNS, services, real DNS.