Campus IPv6

Addressing, Software Versions, Topology Issues, DNS Support, Traffic

Campus Addressing

- Most sites will receive /48 assignments:

16 bits left for subnetting - what to do with them?

Campus Addressing

1. Sequentially, e.g.
   0000
   0001
   …
   FFFF

   16 bits = 65535 subnets

Campus Addressing

1. Sequentially
2. Following existing IPv4:

   Subnets or combinations of nets & subnets, or VLANs, etc., e.g.

   - 128.8.**60**.0/24 003c
   - 128.8.**91**.0/24 005b
   - 128.8.**156**.0/24 009c
   - **156.56.60**.0/24 vs. **129.79.60**.0/24?
     - 013c or 383c or 9c3c vs. 023c or 4f3c or 813c

Campus Addressing

1. Sequentially
2. Following existing IPv4
3. Topological/aggregating

reflecting wiring plants, supernets, large broadcast domains, etc.

Main library = 0010/60

Floor in library = 001a/64

Computing center = 0020/55

Student servers = 002c/64

Medical school = 00c0/50

and so on. . .

New Things to Think About

- You're not limited to 254 hosts per subnet!
  - Switch-rich LANs allow for larger broadcast domains (with tiny collision domains), perhaps thousands of hosts/LAN…
- No "secondary subnets" (though >1 address/interface)
- No tiny subnets either (no /126, /127, /128) — plan for what you need for backbone blocks, loopbacks, etc.
  - Note RFC 3627: "Use of /127 Prefix Length Between Routers Considered Harmful"
- Subnet anycast
  - Cisco supports it
  - Juniper doesn't

New Things to Think About

- Every /64 subnet has far more than enough addresses to contain all of the computers on the planet, and with a /48 you have 65536 of those subnets - use this power wisely!
- With so many subnets your IGP may end up carrying thousands of routes — consider internal topology and aggregation to avoid future problems.

New Things to Think About

- Renumbering will likely be a fact of life. Although v6 does make it easier, it still isn't pretty. . .
  - Avoid using numeric addresses at all costs
  - Avoid hard-configured addresses on hosts except for servers
  - Anticipate that changing ISPs will mean renumbering

IPv6 addressing at Merit

- Merit is currently using Internet2 allocated space for IPv6 routing — 2001:468:1400::0/40
- OSPFv3 is currently used as the IGP within Merit.
- Deployment is currently limited to Merit and Michigan Tech University.
- Merit has received a direct /32 allocation from ARIN, 2001:48A8::0/32, but is not yet announcing the space.
- An addressing plan has been developed to deploy the /32 prefix.
- The addressing plan divides state into four regions:
  - Eastern, Central, and Western Lower Peninsula; Upper Peninsula
  - The Merit address space will be divided between the regions
- Merit's 12 member universities will initially receive /40 allocations.
  - Allocations will be made sparsely to allow expansion up to /38
- Merit affiliate members will receive /48 allocations by default unless they can justify larger allocations.

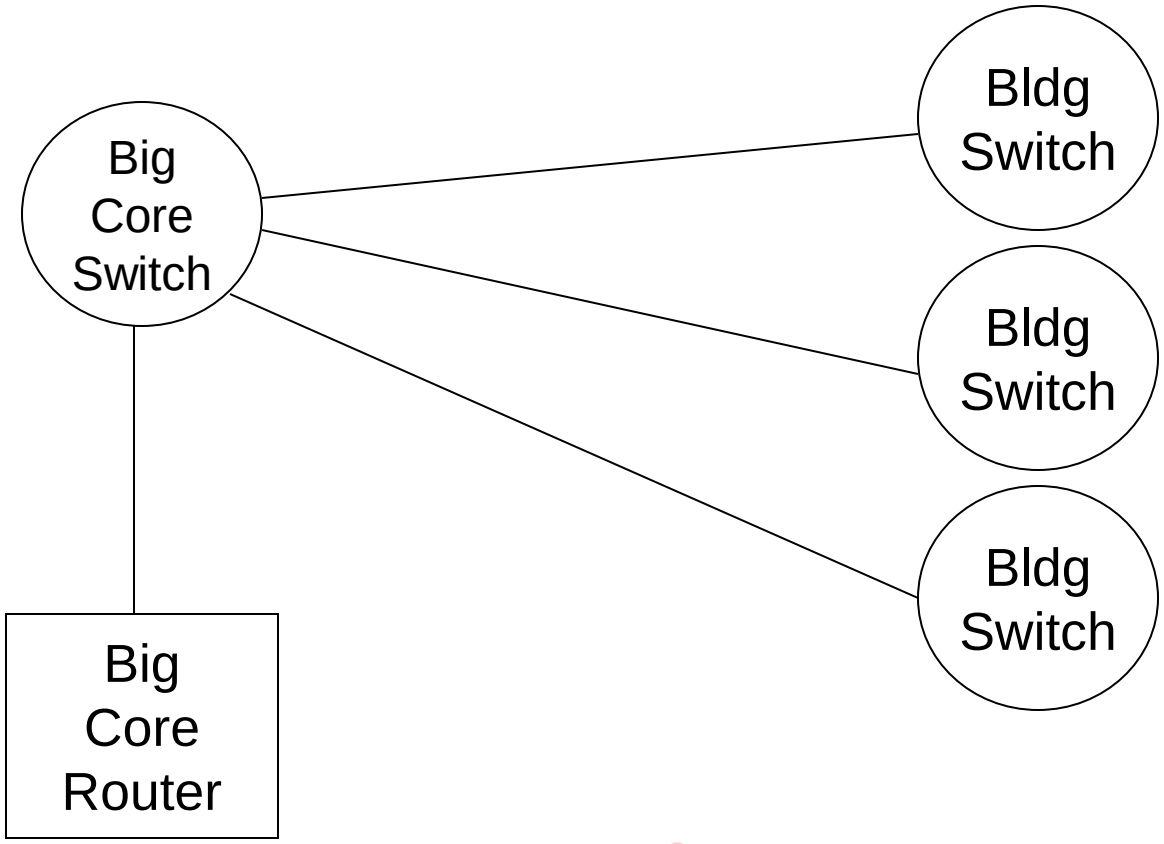
Router Software Versions

- JUNOS 5.1 and up — Line Rate v6 (just turn it on)
- IOS — Use Feature Navigator to find a version: http://tools.cisco.com/ITDIT/CFN/jsp/index.jsp
  - IOS 12.2T and 12.3(6a)(LD)
  - IOS 12.0(22)S6 and up — GSR only
  - 6500 with IOS 12.2(17a)SX
  - 7600 with SUP720 card 12.2(17d)SXB

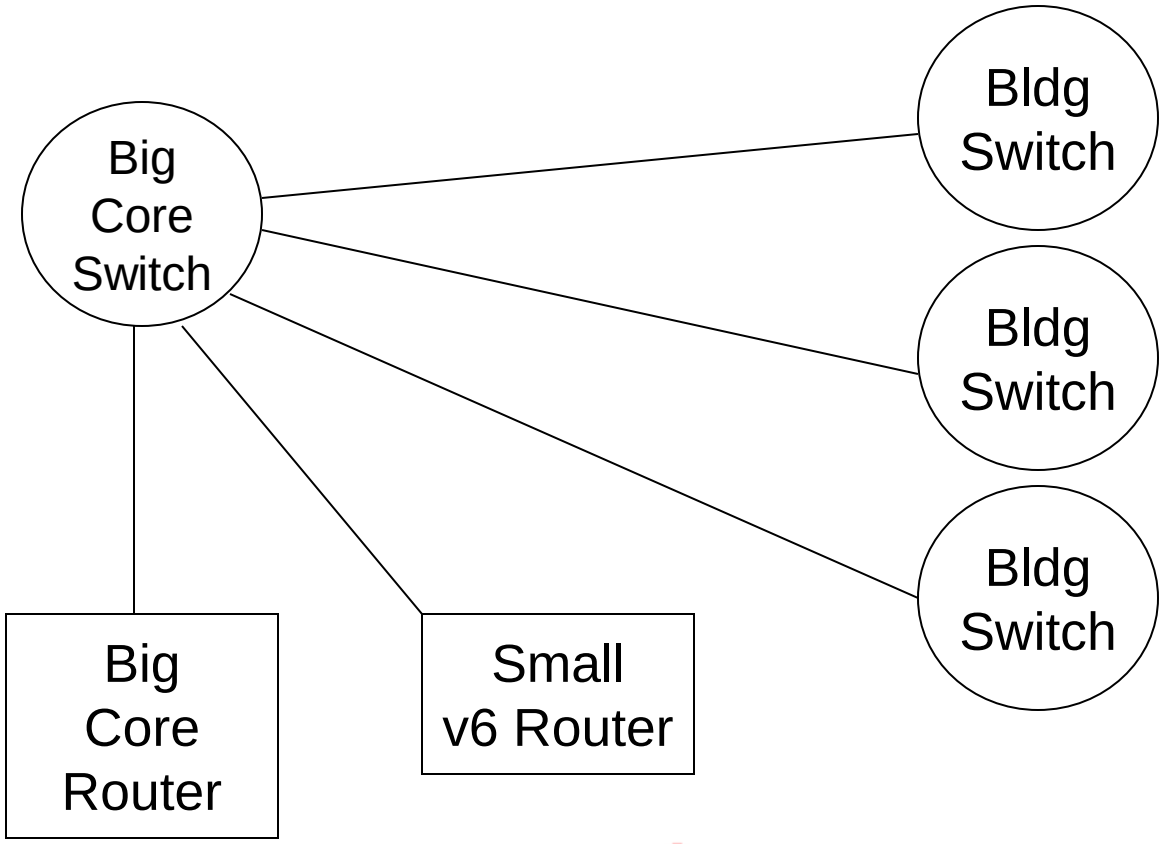
Topology Issues

- v6 in a production network

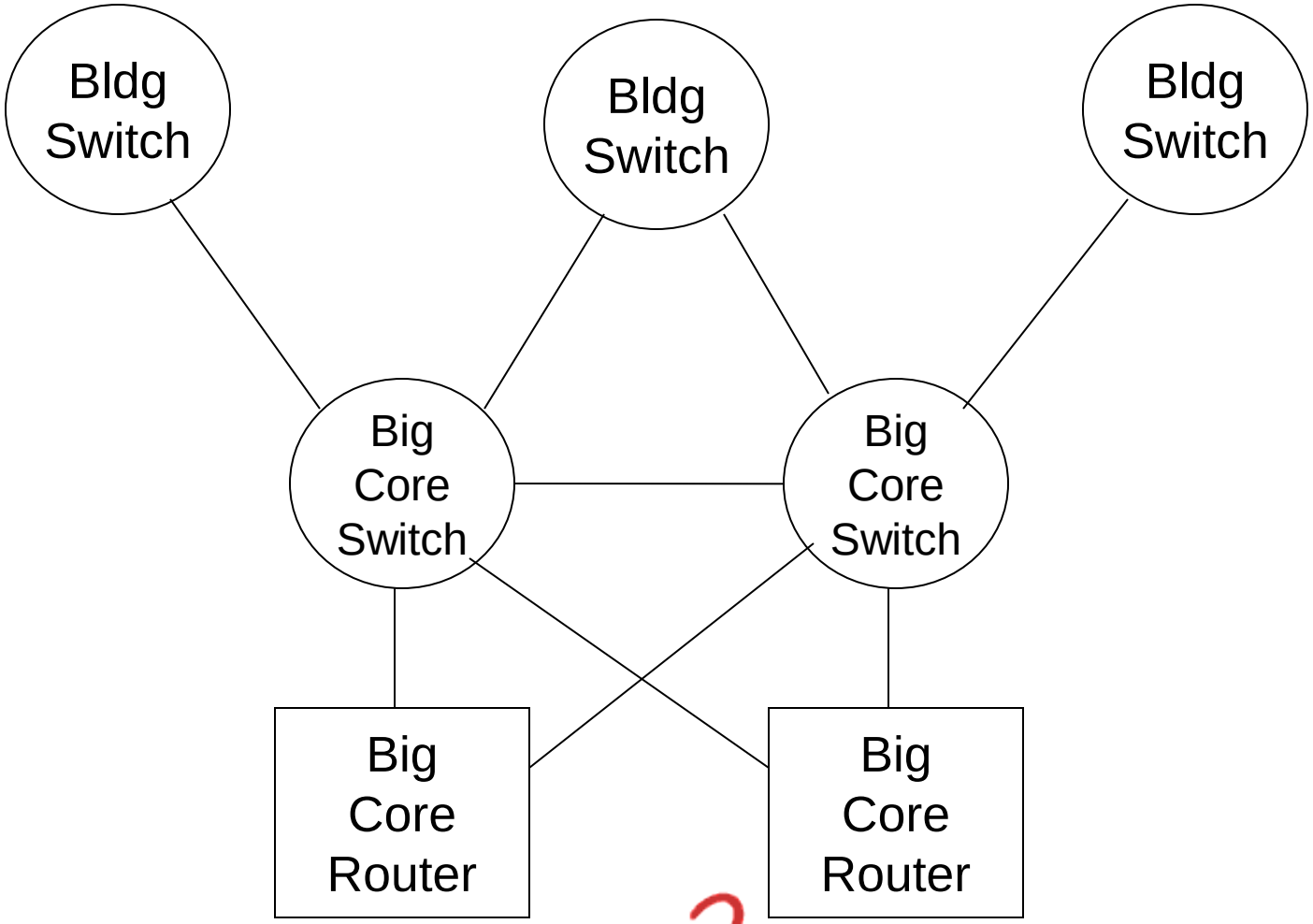
Layer-2 Campus
1 Switch



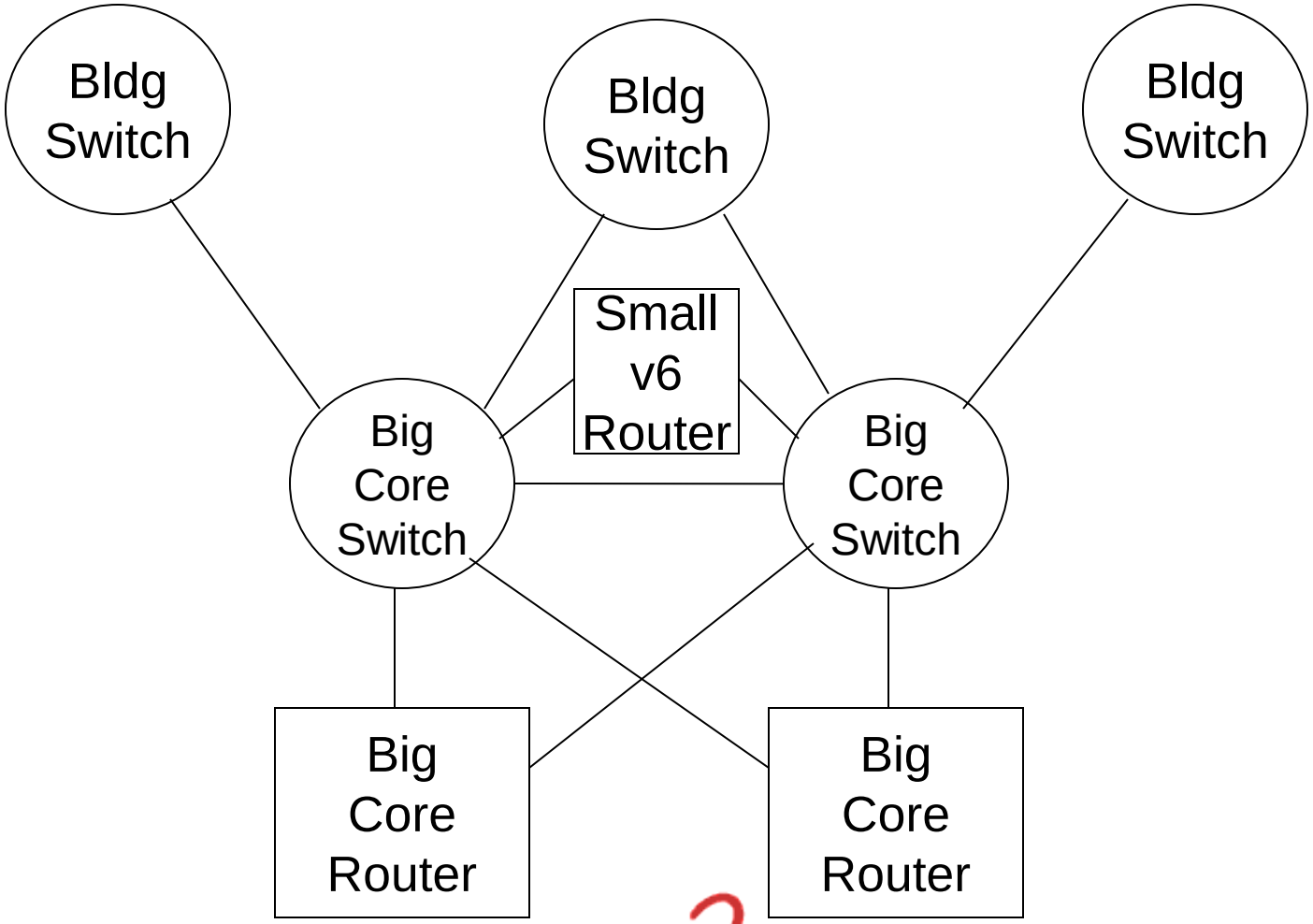
Layer-2 Campus
1 Switch



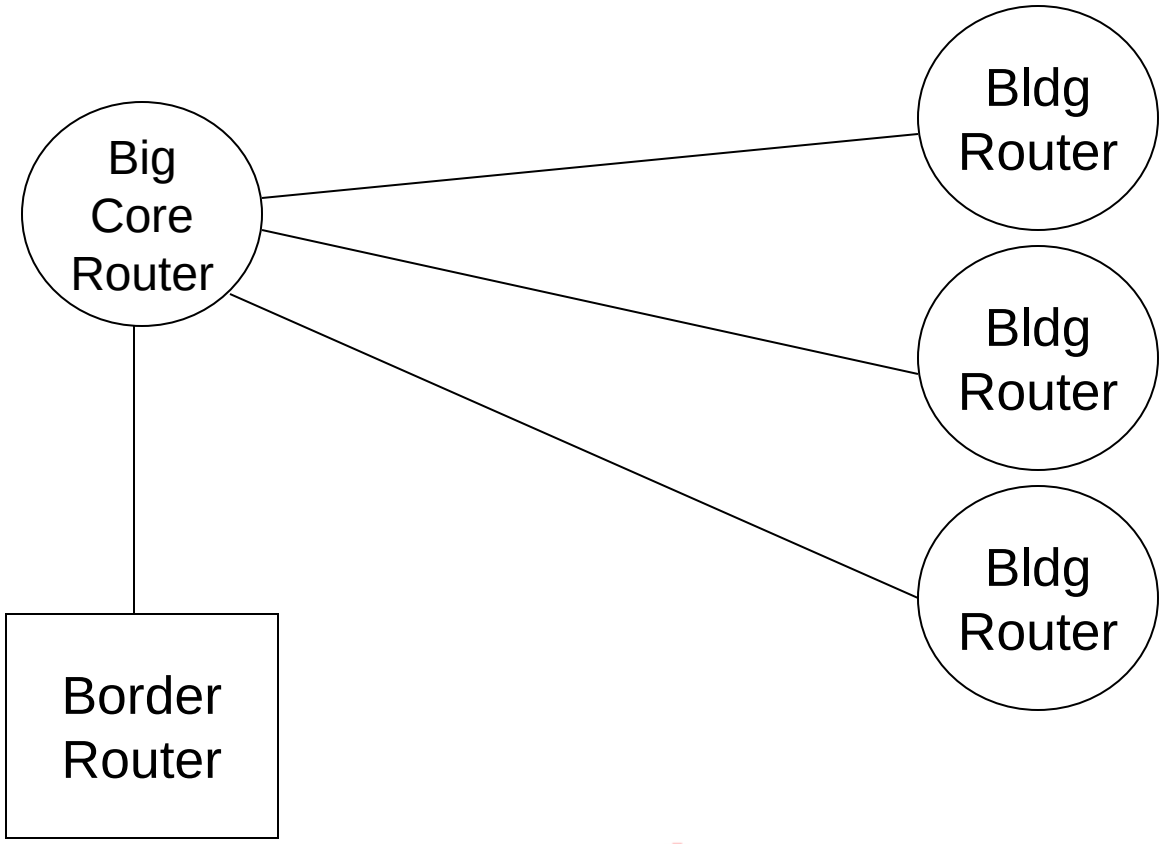
Layer-2 Campus
2 Core Switches



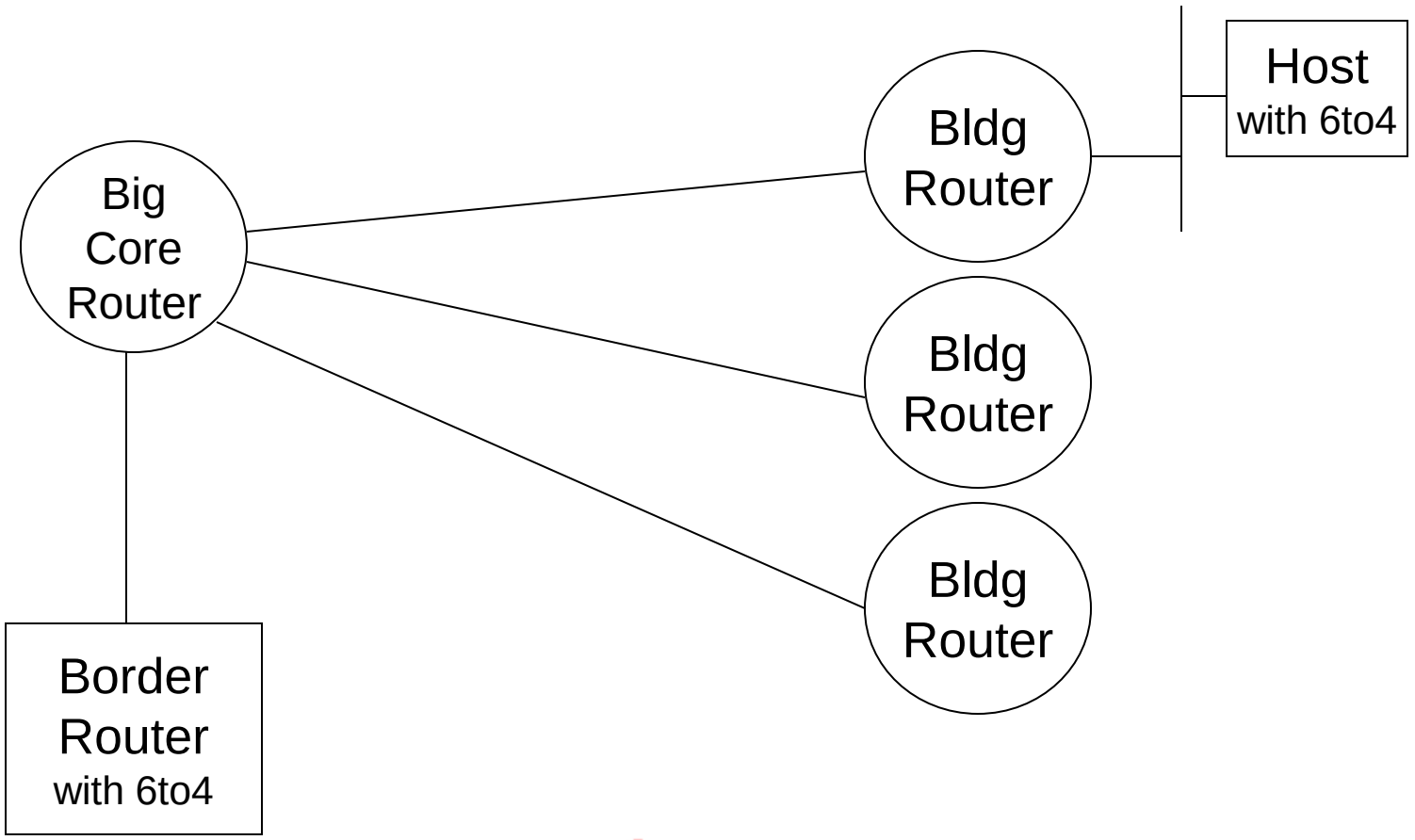
Layer-2 Campus
2 Core Switches



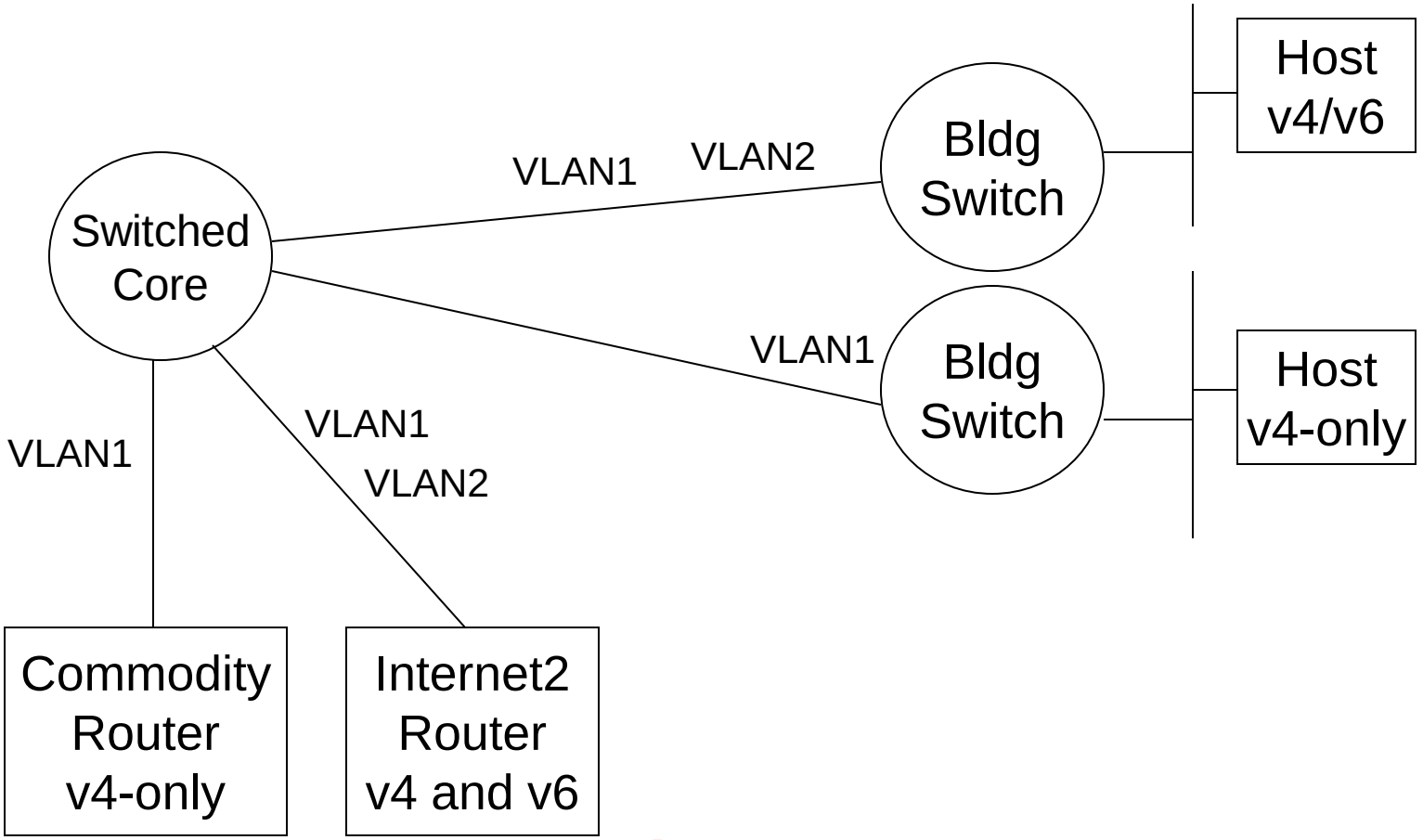
Layer-3 Campus



Layer-3 Campus



Edge Router Options



Routing Protocols

- iBGP and IGP (RIPng/IS-IS)
  - IPv6 iBGP sessions in parallel with IPv4
- Static Routing
  - all the obvious scaling problems, but works OK to get started, especially using a trunked v6 VLAN.
- OSPFv3 is available in IOS 12.3 and JUNOS.
  - It runs in a ships-in-the-night mode relative to OSPFv2 for IPv4 — neither knows about the other.

DNS Issues

- BIND Versions
  - All modern versions of BIND support AAAA
  - BIND9 can use IPv6 transport for queries
- An IPv6 root test project is underway; see www.rs.net for details.
- ip6.int vs. ip6.arpa
  - ip6.arpa is in the roots
- Some registrars and registries are now supporting IPv6 NS records.

Equipment Needs

- Tunnel Router (Cisco 2600) ~$2,000
  - A router with two Ethernet interfaces is best, to avoid one-armed routing.
- Workstation Linux Box ~$1,000
  - For testing and demonstrations, any old cast-off Pentium will get you going. . .

Future Needs

- Routers: more platform support, new features, speed, management
- Servers: dual-stack, application support
- Workstations: application support, address selection
- Topology: multihoming

DNS

Basic Ideas

- DNS in IPv6 is much like DNS in IPv4.
- It is impossible to remember IPv6 addresses — DNS is the ***only*** way to remain sane.
- Keep files and delegations as simple as possible.
- Can use IPv4 or IPv6 as transport for DNS traffic.
- Modern versions of BIND will work. BIND 9 is stable and works with IPv6 transport.
- There is work on dynamic DNS in progress, but we don't need to worry about that for now.

Forward Lookups

- Uses AAAA records to assign IPv6 addresses to names.
- Multiple addresses possible for any given name – for example, in a multi-homed situation.
- Can assign A records and AAAA records to a given name/domain.
- Can also assign separate domains for IPv6 and IPv4.
- Don’t be afraid to experiment!

Sample Forward Lookup File

```
;; domain.edu  (use your favorite naming scheme)
$TTL          86400
@    IN      SOA     ns1.domain.edu. root.domain.edu. (
                  2002093000        ; serial - YYYYMMDDXX
                  21600             ; refresh - 6 hours
                  1200              ; retry - 20 minutes
                  3600000           ; expire - long time
                  86400)            ; minimum TTL - 24 hours
;; Nameservers
         IN       NS       ns1.domain.edu.
         IN       NS       ns2.domain.edu.
;; Hosts with just A records
host1             IN       A        1.0.0.1
;; Hosts with both A and AAAA records
host2             IN       A        1.0.0.2
                  IN       AAAA     2001:468:100::2
:: Separate domain
$ORIGIN ip6.domain.edu
host1             IN       AAAA     2001:468:100::1
```

Reverse Lookups

- Reverses should be put in for both ip6.int and ip6.arpa domains.
- The ip6.int domains have been deprecated, but some hosts still use them.
- Can use same file for both – use the @ notation and point to the same file in the named.conf file.
- File uses nibble format – see examples on next slide.

Sample Reverse Lookup File

```
;; 0.0.0.0.0.0.1.0.8.6.4.0.1.0.0.2.rev  (use your favorite naming scheme
;; These are reverses for 2001:468:100::/64)
;; File can be used for both ip6.arpa and ip6.int.
$TTL          86400
@    IN      SOA     ns1.domain.edu. root.domain.edu. (
                  2002093000        ; serial - YYYYMMDDXX
                  21600             ; refresh - 6 hours
                  1200              ; retry - 20 minutes
                  3600000           ; expire - long time
                  86400)            ; minimum TTL - 24 hours
;; Nameservers
         IN       NS       ns1.domain.edu.
         IN       NS       ns2.domain.edu.

; This is the forward analog for address:
; host1.ip6.domain.edu. In aaaa 2001:468:100::1
;
1.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0   IN      PTR     host1.ip6.domain.edu.
2.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0   IN      PTR     host2.domain.edu.
;;
;; Can delegate to other nameservers in the usual way
;;
```

Sample Configuration File

```
// named.conf  (use your favorite naming scheme)

zone “domain.edu” {
   type master;
   file “master/domain.edu”;
}
zone “0.0.0.0.0.0.1.0.8.6.4.0.1.0.0.2.ip6.int" {
   type master;
   file "master/0.0.0.0.0.0.1.0.8.6.4.0.1.0.0.2.rev";
};
zone “0.0.0.0.0.0.1.0.8.6.4.0.1.0.0.2.ip6.arpa" {
   type master;
   file "master/0.0.0.0.0.0.1.0.8.6.4.0.1.0.0.2.rev";
};
```

DNS Notes

- Bind 8 can return a AAAA record using IPv4 transport.
- Bind 9 can use IPv6 transport.

- When the same name returns both an A and AAAA record, the AAAA is preferred.
- At least one application, Safari, explicitly does not follow this behavior.

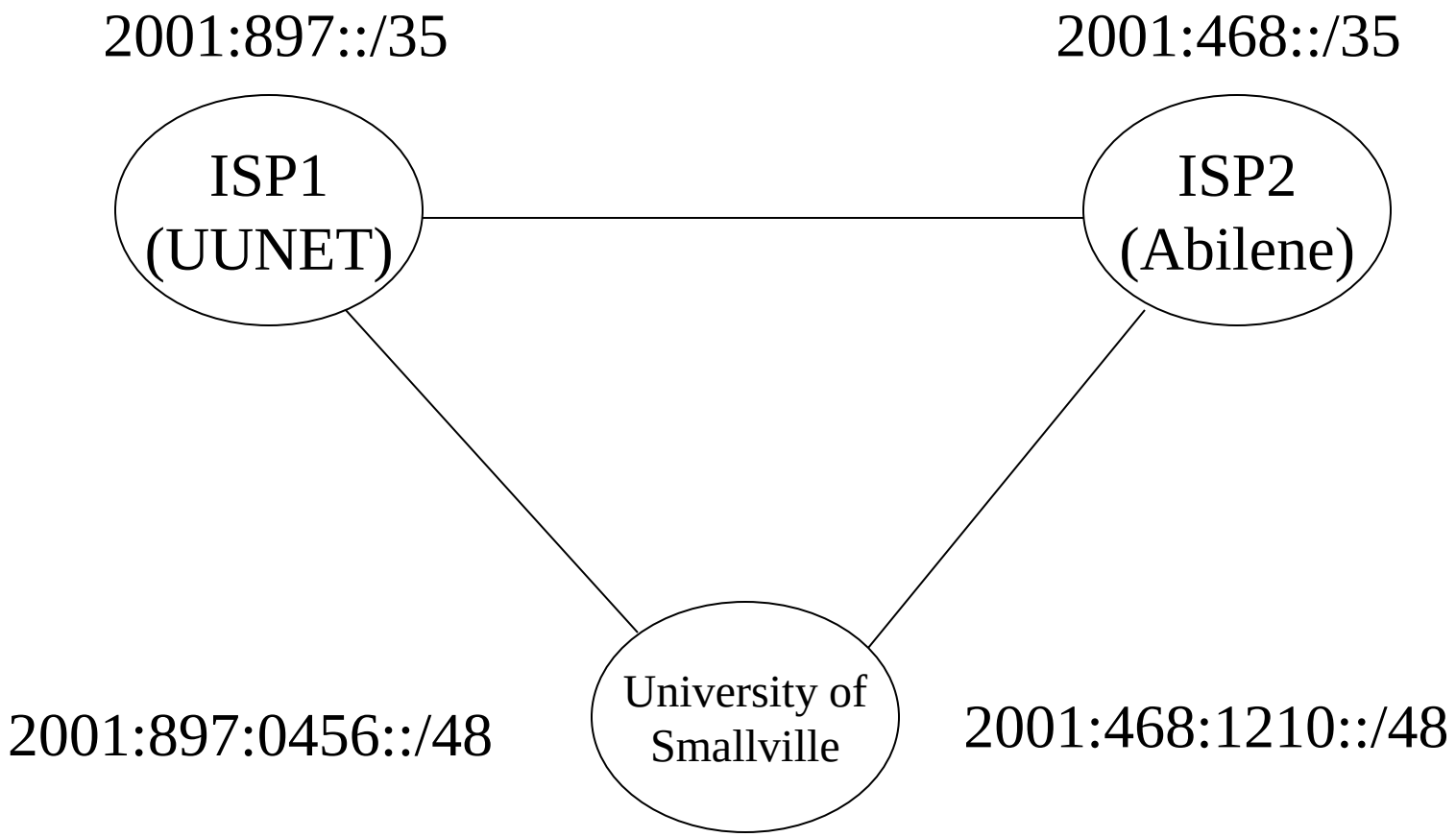
Multihoming

A Discussion

Multihoming Issues

- Many sites are multihomed in the current Internet
  - reliability
  - stability — which provider will stay in business?
  - competition
  - AUP — commodity vs. R&E
- In IPv4 we can use provider-independent addresses, or “poke holes” in the aggregation
- But all IPv6 addresses are provider-assigned!

Multihoming



Problems With Multiple Addresses

- If the host or app chooses from several global addresses, that choice overrides policy, may conflict with routing intentions and can break connectivity
- Address selection rules are complex and controversial; see RFC 3484
  - Other informational RFCs are RFC 3582, RFC 4116, RFC 4218, RFC 4219

Problems With PI Addressing

- Current protocols can only control routing table growth if routes are aggregated.
- Only about 12,000 sites are multihomed today, but that number is constantly increasing.
- The address space is so large that routing table growth could easily exceed the capability of the hardware and protocols.

What To Do?

- IPv6 can't be deployed on a large scale without multihoming support — nobody is disputing this.
- It seems likely that there will be short-term fixes to allow v6 deployment, and long-term solutions.
- IETF multi6 and shim6 working groups
- recent IAB workshop
  - http://www.1-4-5.net/~dmm/draft-iab-raws-report-00.txt
- two mailing lists that are discussing IPv6 multihoming options
  - https://www1.ietf.org/mailman/listinfo/ram
  - https://www1.ietf.org/mailman/listinfo/architecture-discuss
- see also
  - http://www3.tools.ietf.org/group/irtf/trac/wiki/RoutingResearchGroup

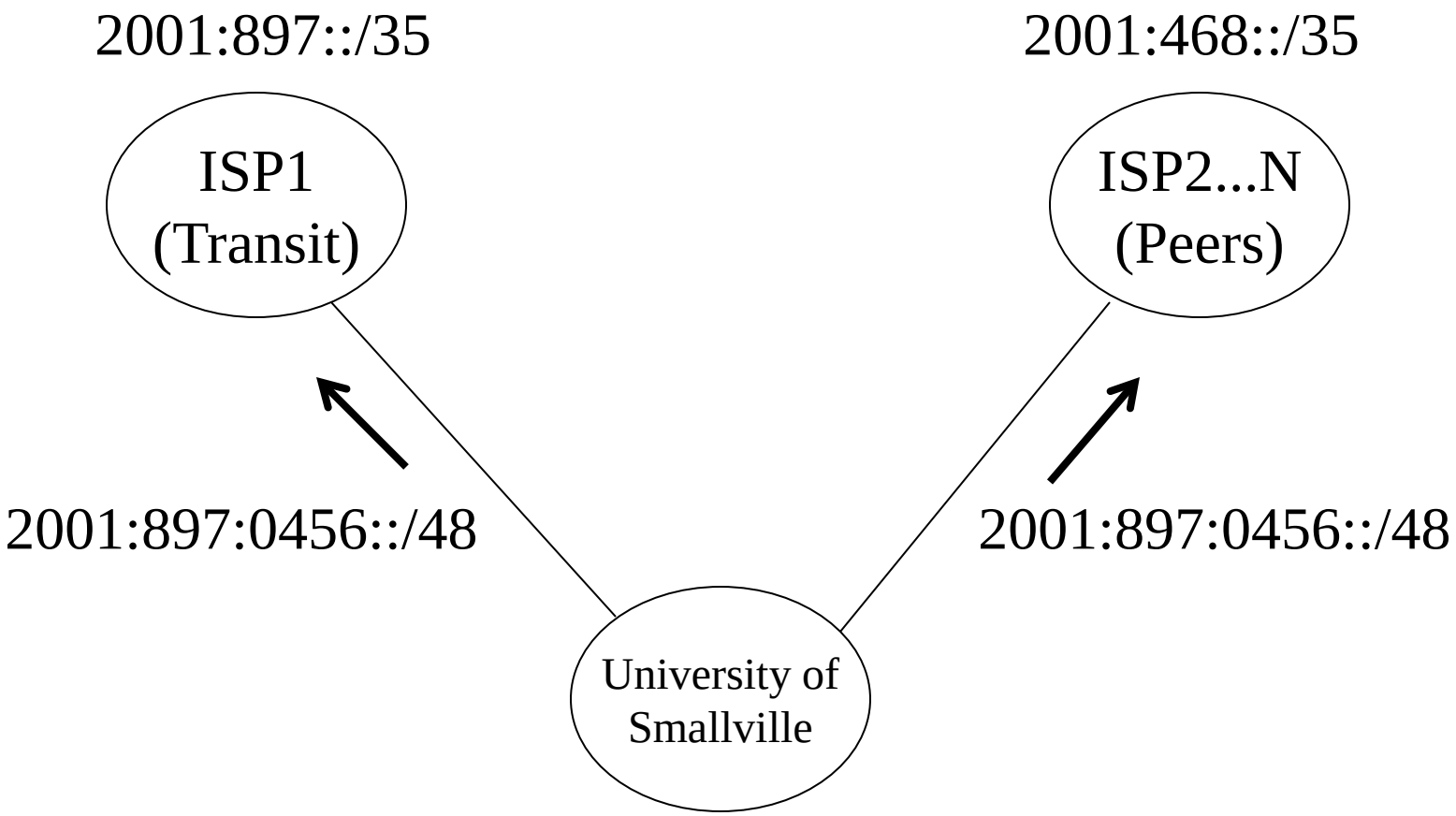
Get PI Space

- The RIRs have revised their rules for allocating PI space; the key is that you must plan to assign 200 /48s within 2 years.
- This isn't as hard as it sounds, but it is probably something only gigaPoPs or large university systems can do (exercise in creativity).
- This breaks when commodity providers start offering IPv6 (unless the gigaPoP aggregates all the commodity providers as well as R&E).

Poke Holes

- The standard practice in IPv4 is to get addresses from one ISP, and advertise that space to all of our providers, effectively making it a PI address.
- In the v6 world, most providers probably won't advertise a foreign prefix to their peers, but will carry it within their own network.
- Requires that one ISP be designated as the transit provider, and others are effectively peers.
- ARIN is now allocating /48s from 2620:0::/32

Poke Holes


2001:897::/35
2001:468::/35
ISP1
(Transit)
ISP2...N
(Peers)
2001:897:0456::/48
2001:897:0456::/48
University of
Smallville

Things to watch for in the BGP lab

- You have to be able to reach the peer's address for BGP to come up: static, OSPF, connected.
- Your source-address needs to be the same as the one they're trying to reach (and vice-versa).
- Remember that you have to have your /48 in your IGP.
  - IOS: `network` statement and `static-route-to-Null` or `aggregate-address ... summary-only`
  - JunOS: `routing-options static`
- Advertise your upstream's originating address into your IGP for your downstreams to be able to reach it, or set `next-hop-self`.
- iBGP members don't send iBGP-learned prefixes to other iBGP peers: they expect mesh. So, you should iBGP among all of A, B, and C.
- Best practice is to send only your aggregated prefix upstream.

BGP Lab

- Configure iBGP peerings between routers A, B and C, using loopback addresses
- Configure eBGP between pods, using interface addresses agreed to between each pair of pods
- Advertise your /40 aggregate to the other pods
- Verify intra-pod and inter-pod connectivity with ping and traceroute
- Can you see the other pods' BGP advertisements?

- Configure eBGP between router A and the external connection to the twenty-first router
- Verify receipt of BGP routes from the outside
- Verify external connectivity with ping and traceroute to ping-nycm.abilene.ucaid.edu
- Connect to http://www.kame.net and see the swimming turtle!