facebook

LISP Deployment at Facebook

Donn Lee
Network Engineering Team
3 October 2010
NANOG 50
Atlanta

Agenda

1. Background
2. LISP for v4
3. LISP for v6

Background

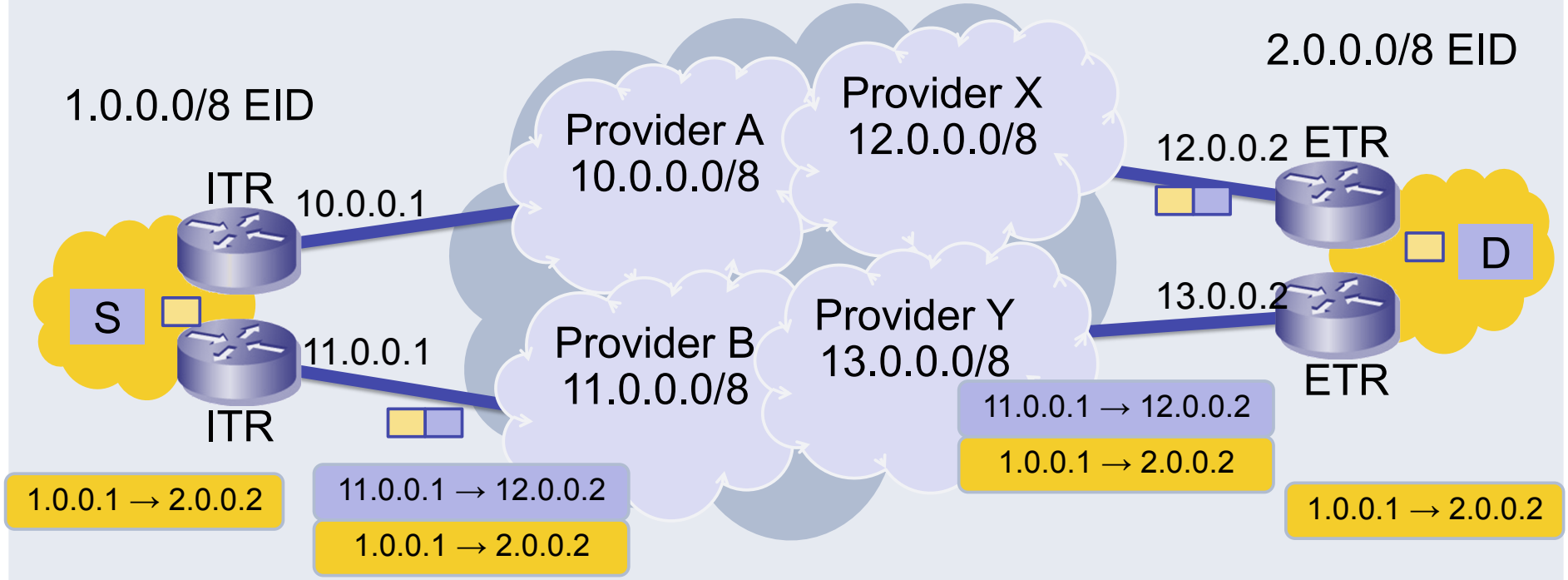
Locator/ID Separation Protocol (LISP)

- Scalable routing and addressing
- IETF working group
- draft-ietf-lisp-*.txt
- Decoupling of client's identifier and location
- IP encapsulation
  - Eg. v4 in v4, v6 in v4, v6 in v6
- Map and encap

Locator/ID Separation Protocol (LISP)

- No changes to end systems
- Incrementally deployable
- Multi-homing controls for stub AS
- Facebook was the first major website on LISP (v4 and v6)

LISP forwarding

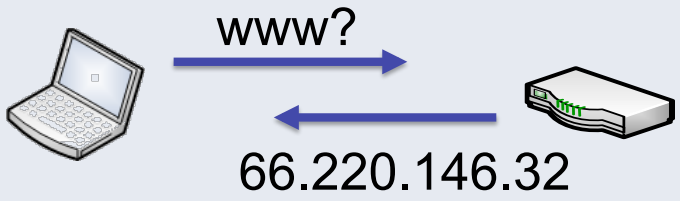


Glossary

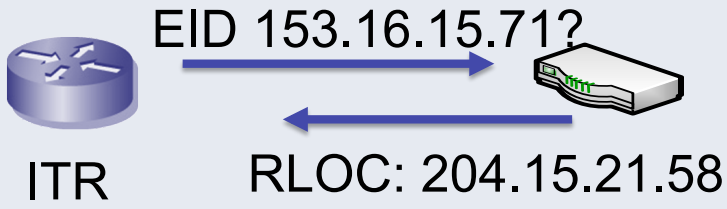
- **EID** = Endpoint Identifier
  - Portable ID
  - Inner header address
  - What you want to connect to (service, user)
- **Locator** = Routing Locator = RLOC
  - Globally routable address
  - Outer header address
  - How you reach an EID

My simplistic analogy

- DNS resolver provides IP address when you want to reach www.facebook.com



- LISP resolver provides **locator** when an ITR wants to reach (encap for) an **EID**



Why

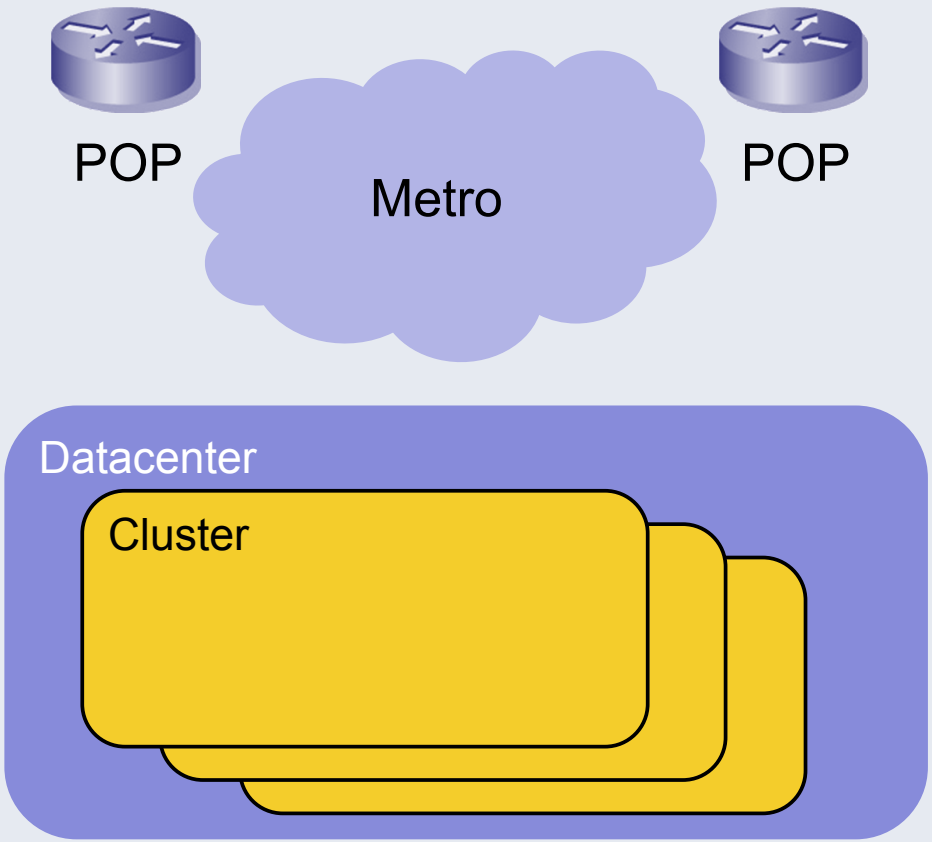
- Curious
  - How would it work at a large content provider?
- Deployment experience
- Provide real-world feedback, Influence spec
- Work with Dino again
- Easy

LISP goals

- Serve Facebook over LISP
- Model real-world, full-scale deployment
- Reach v6 users using LISP

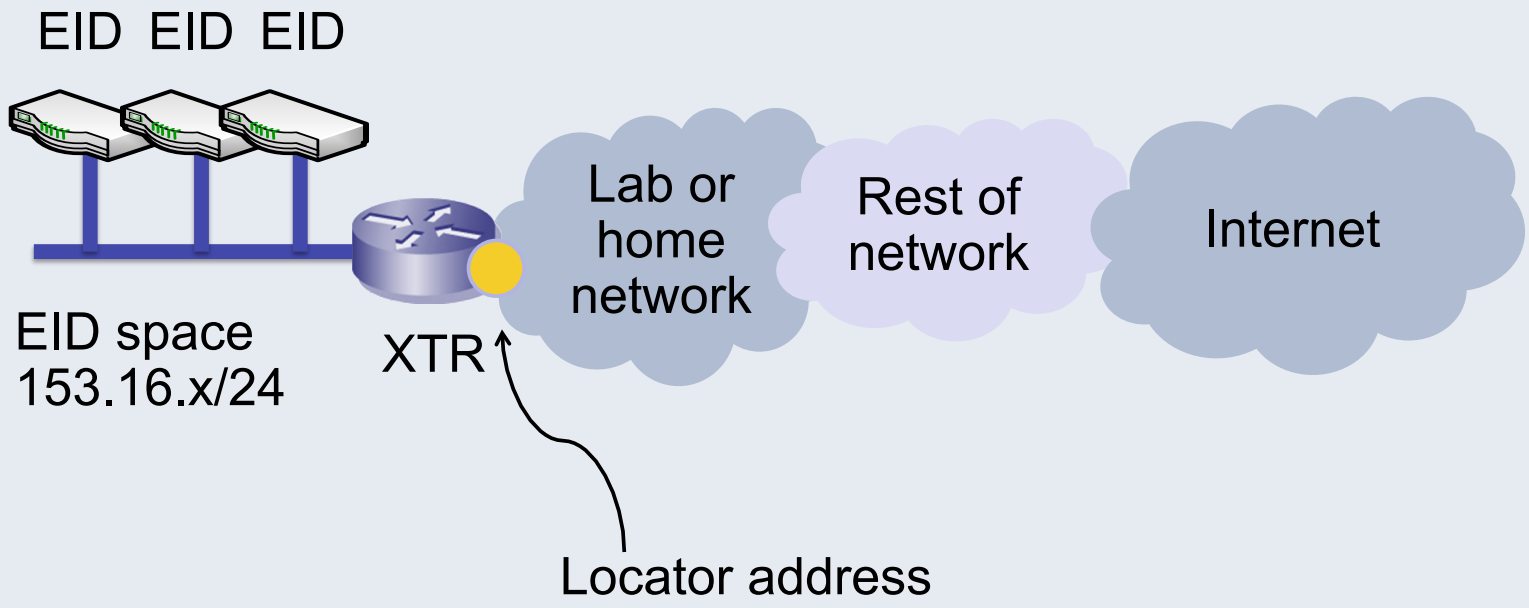
LISP for v4

XTR placement options



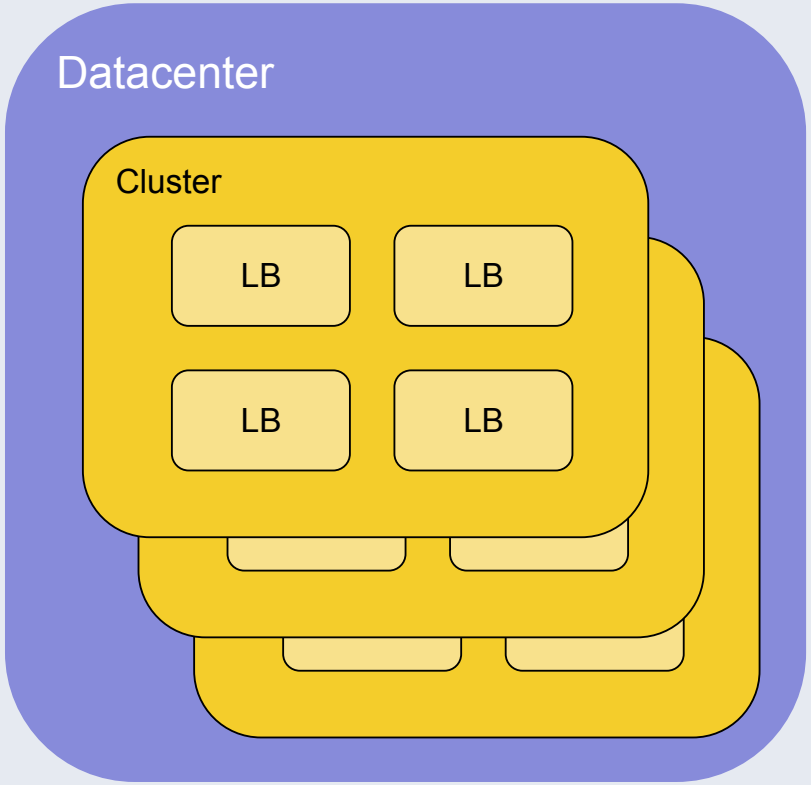
Typical beta setup

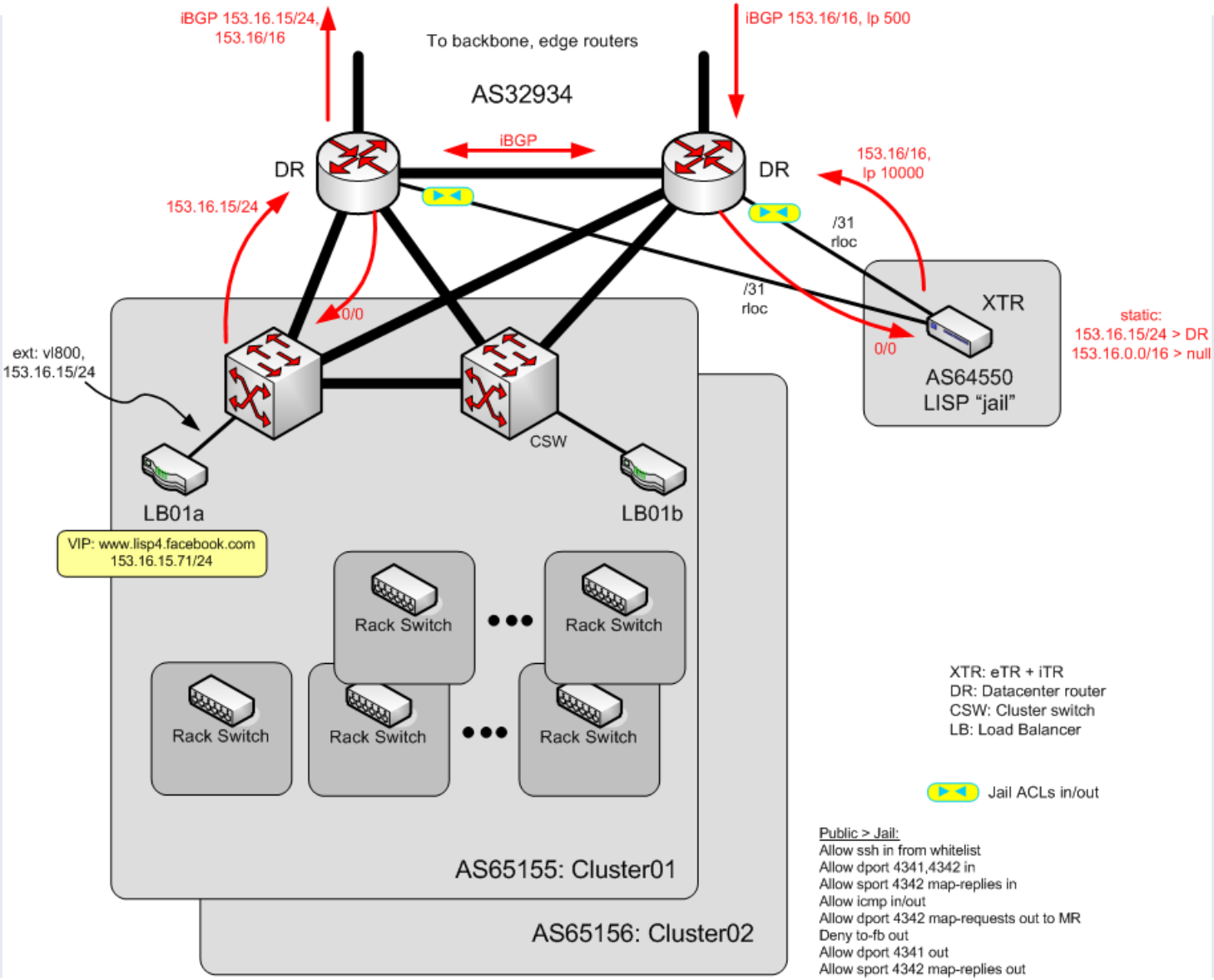
- XTR adjacent with EIDs/hosts



XTR placement

- But for us, EID = VIP
  - Beta user approach won't scale with large number of VIPs, clusters
- Choose to deploy at cluster agg (datacenter) layer
- Could also be viable at POP



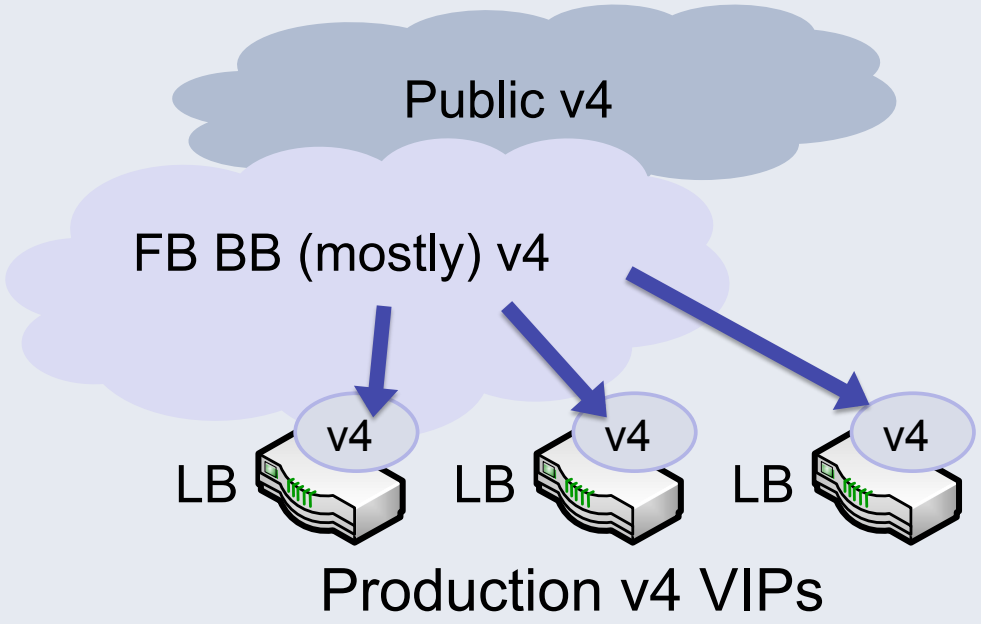

iBGP 153.16.15/24, 153.16/16
To backbone, edge routers
iBGP 153.16/16, lp 500
AS32934
iBGP
DR
DR
153.16/16, lp 10000
153.16.15/24
/31 rloc
/31 rloc
XTR
0/0
0/0
static:
153.16.15/24 > DR
153.16.0.0/16 > null
AS64550
LISP "jail"
ext: vl800, 153.16.15/24
CSW
LB01a
LB01b
VIP: www.lisp4.facebook.com
153.16.15.71/24
Rack Switch
Rack Switch
Rack Switch
Rack Switch
Rack Switch
AS65155: Cluster01
AS65156: Cluster02
XTR: eTR + iTR
DR: Datacenter router
CSW: Cluster switch
LB: Load Balancer
Jail ACLs in/out
Public > Jail:
Allow ssh in from whitelist
Allow dport 4341,4342 in
Allow sport 4342 map-replies in
Allow icmp in/out
Allow dport 4342 map-requests out to MR
Deny to-fb out
Allow dport 4341 out
Allow sport 4342 map-replies out
Allow 80,443 out

LISP for v4 results

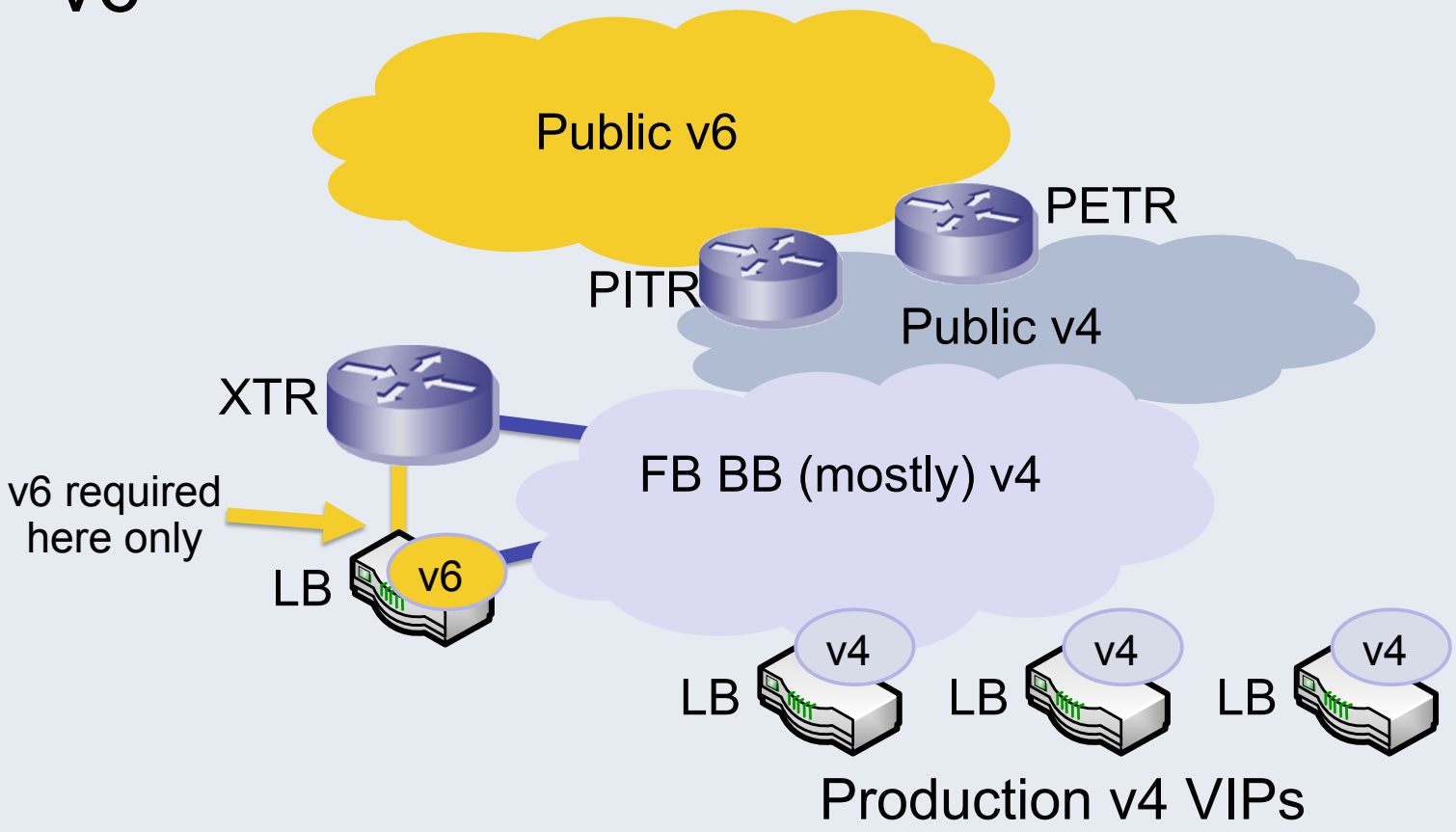
- Easy deployment, config
- Most effort: ACL authoring
  - To strictly filter LISP protocols, encap
- www.lisp4.facebook.com

Project Cakewalk
LISP for v6

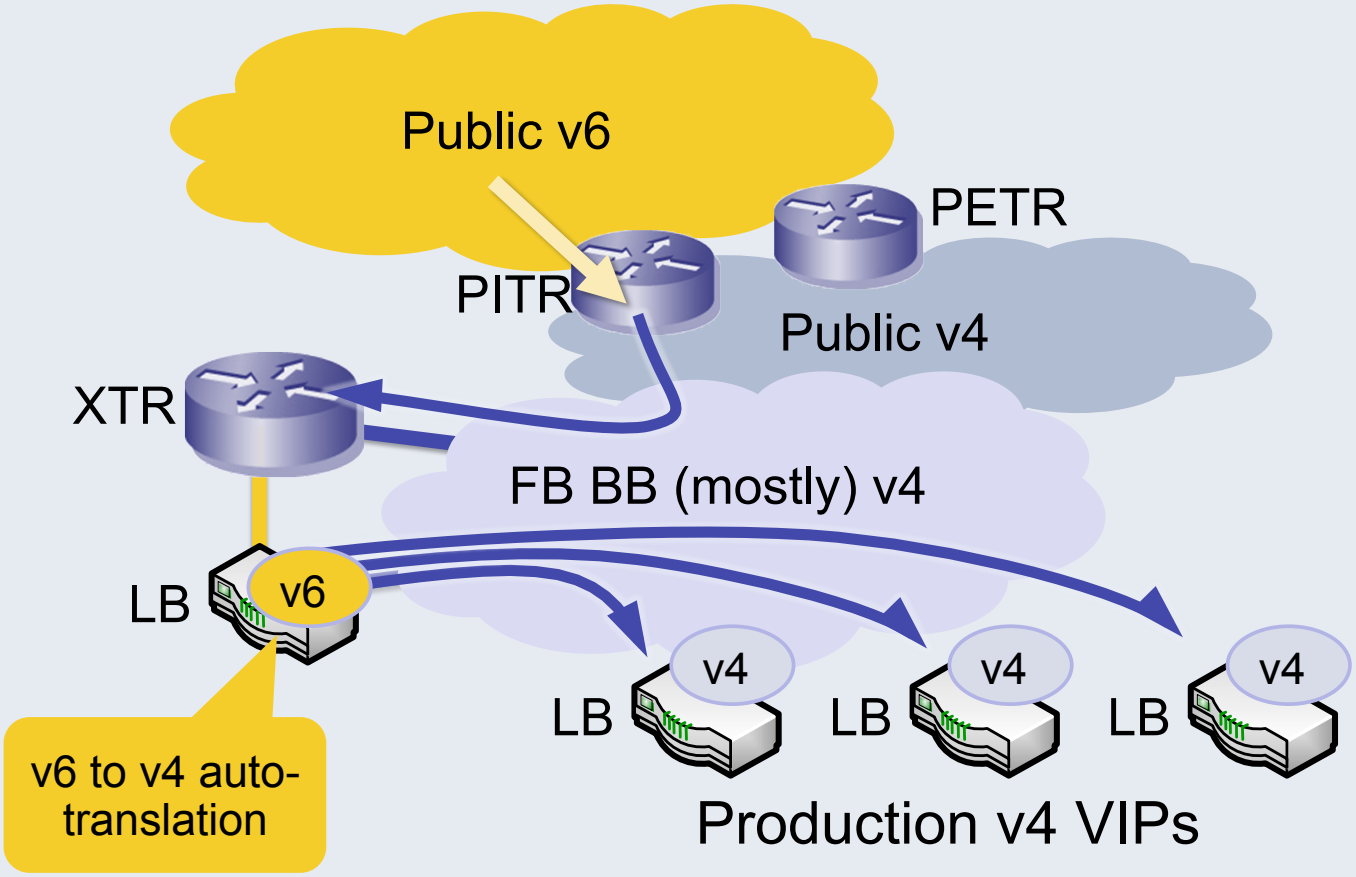
Production v4



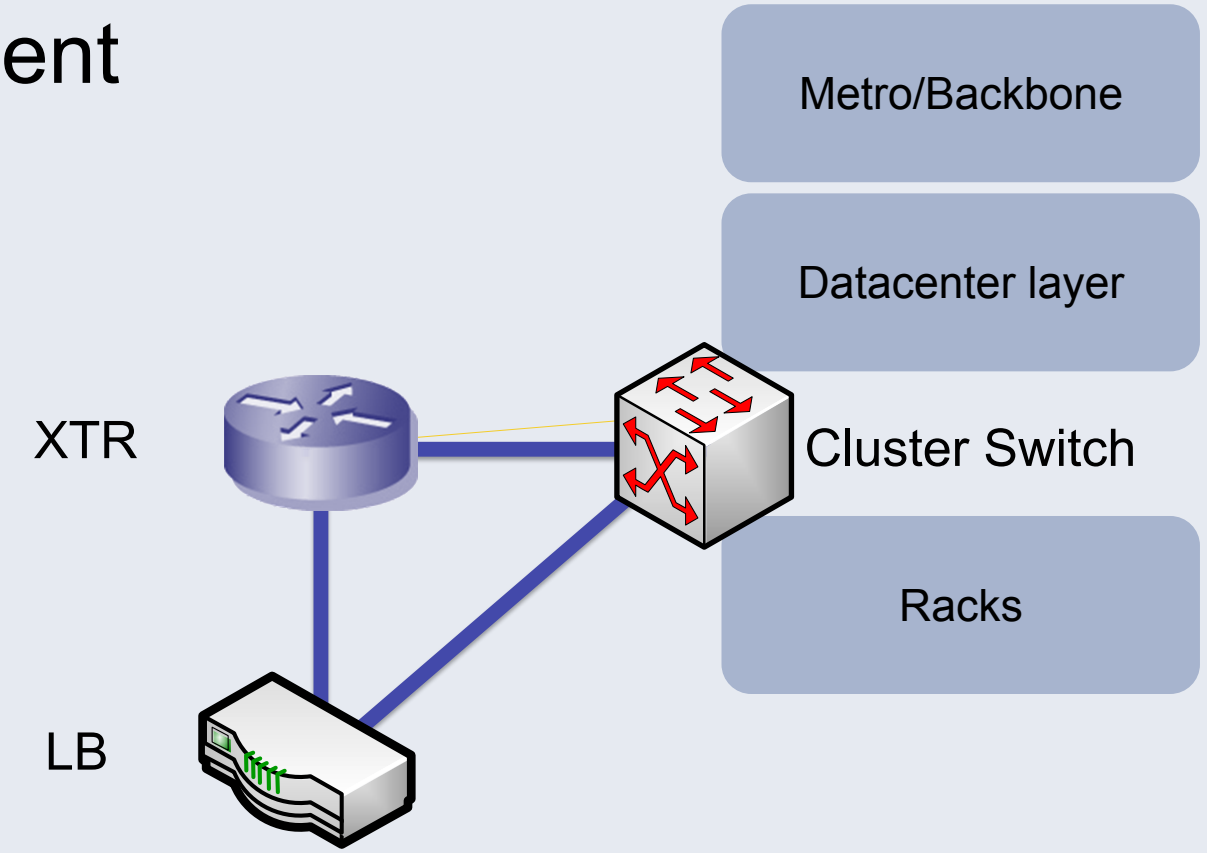
v6


Public v6
PETR
PITR
Public v4
XTR
FB BB (mostly) v4
v6 required
here only
LB
v6
LB
v4
LB
v4
LB
v4
Production v4 VIPs

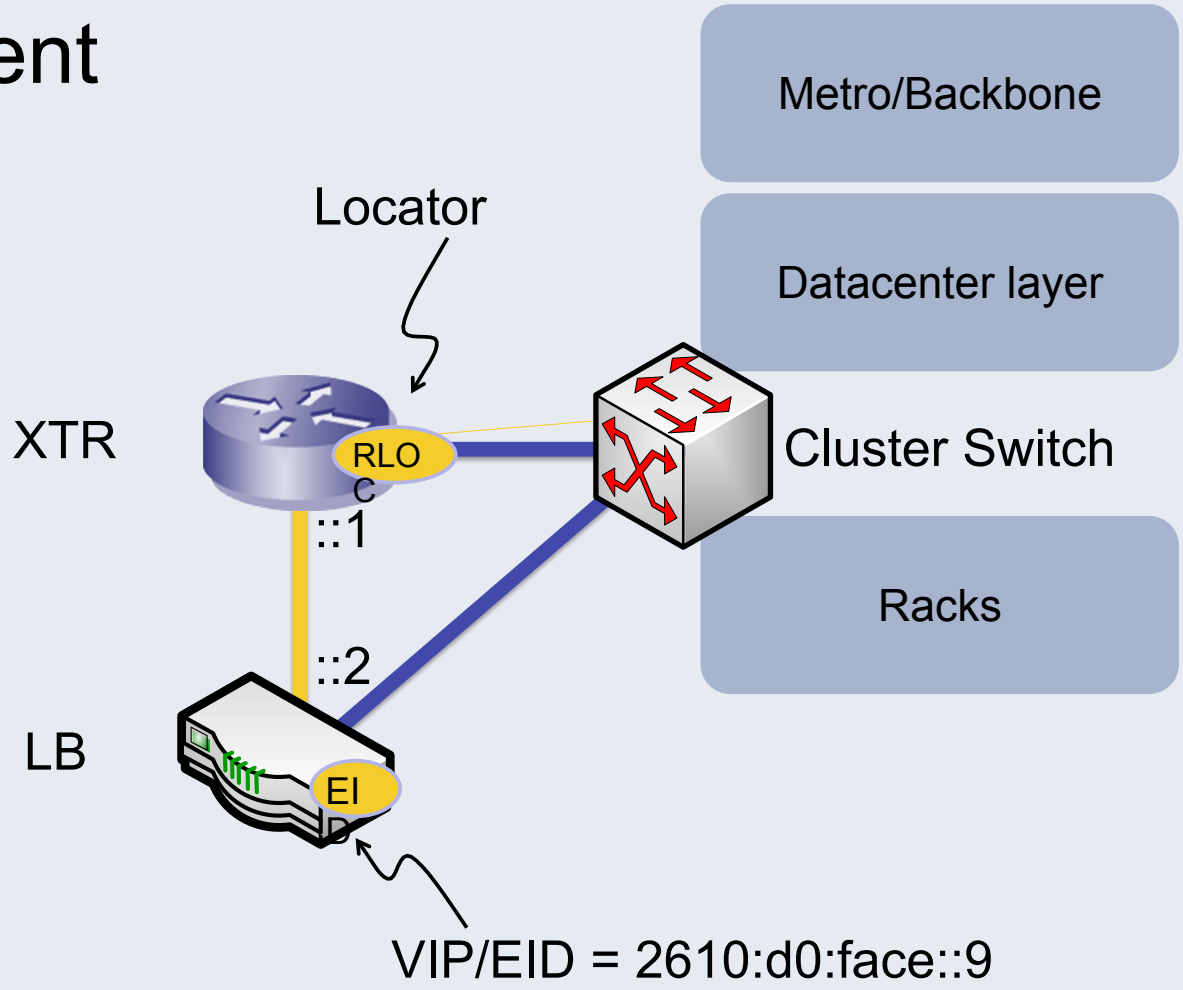
v6


Public v6
PETR
PITR
Public v4
XTR
FB BB (mostly) v4
LB
v6
v4
LB
v4
LB
v4
LB
v6 to v4 auto-translation
Production v4 VIPs

Deployment


Metro/Backbone
Datacenter layer
XTR
Cluster Switch
Racks
LB

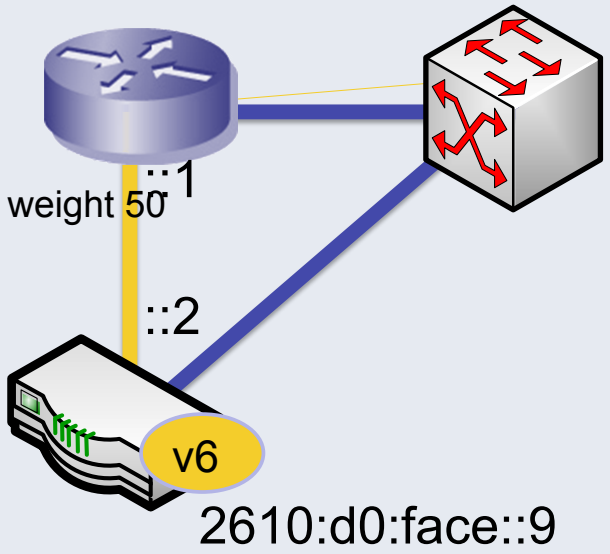
Deployment


Metro/Backbone
Datacenter layer
Racks
Locator
XTR
RLO
C
::1
Cluster Switch
::2
LB
EI
D
VIP/EID = 2610:d0:face::9

Config

XTR

```
ipv6 lisp use-petr 149.20.48.60
ipv6 lisp database-mapping 2610:D0:FACE::/48  74.119.77.125 priority 1 weight 50
ipv6 lisp itr map-resolver 128.223.156.35
ipv6 lisp etr map-server 206.223.132.89 key f00bar
```



LB

```
virtual www.lisp6.facebook.com_vs  {
   snat automap
   pool www.http.vips.pool
   destination 2610:d0:face::9.http
   ip protocol tcp
   profiles {
      http {}
      tcp-wan-optimized {}
   }
}
```

DNS

```
www.lisp6      IN    AAAA     2610:d0:face::9
*.lisp6        IN    CNAME    www.lisp6.facebook.com.
```

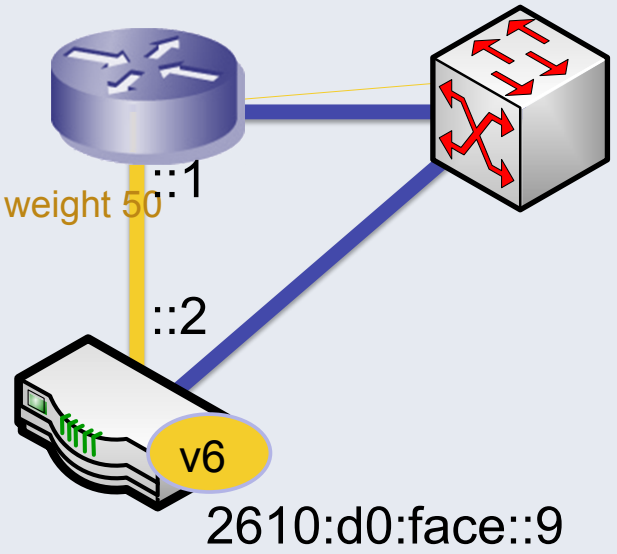
Config

XTR

```
ipv6 lisp use-petr 149.20.48.60
ipv6 lisp database-mapping 2610:D0:FACE::/48 74.119.77.125 priority 1 weight 50
ipv6 lisp itr map-resolver 128.223.156.35
ipv6 lisp etr map-server 206.223.132.89 key f00bar
```

LB

```
virtual www.lisp6.facebook.com_vs  {
   snat automap
   pool www.http.vips.pool
   destination 2610:d0:face::9.http
   ip protocol tcp
   profiles {
      http {}
      tcp-wan-optimized {}
   }
}
```

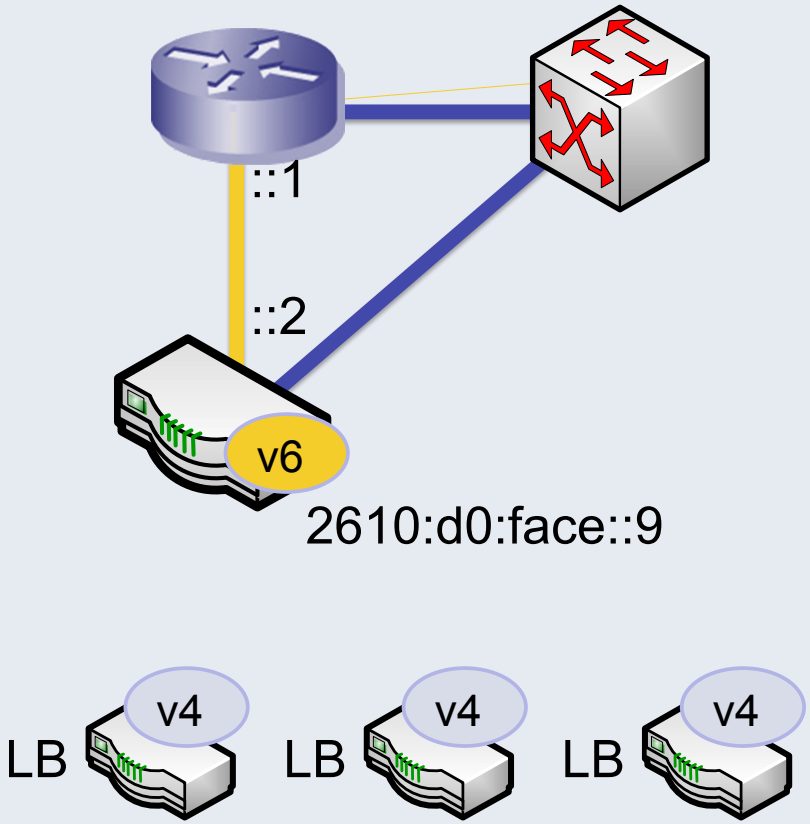


DNS

```
www.lisp6       IN    AAAA     2610:d0:face::9
*.lisp6         IN    CNAME    www.lisp6.facebook.com.
```

Cakewalk results

- Same FB codebase as v4
- Deployed in 4 hours
- Added one router (LISP XTR)
- Required three v6 addresses
- $0 cost
- LISP is transparent to all v6 users



www.lisp6.facebook.com
m.lisp6.facebook.com

Experimental, non-production

Make your next status update over IPv6