

PennNet and MAGPI

Shumon Huque
University of Pennsylvania
April 1st 2009

This presentation was prepared for Roch Guerin's TCOM 502 course (Advanced Networking) at Penn in April 2009. I was invited to speak to the class about the architecture and design of the campus network.

PennNet: at a glance

- University of Penn's campus data network
- 40,000 nodes
- > 200 buildings
- 400 subnets
- Primarily Switched Ethernet (10, 100, 1000, 10000)
- Distributed core (5 locations or NAPs)
- 2 commercial ISPs
- 1 connection to Internet2 (via MAGPI)

Penn address space

- IPv4 (AS 55):

- 165.123.0.0/16
- 128.91.0.0/16
- 130.91.0.0/16
- 158.130.0.0/16
- 192.84.2.0/24

- IPv6 (AS 55):

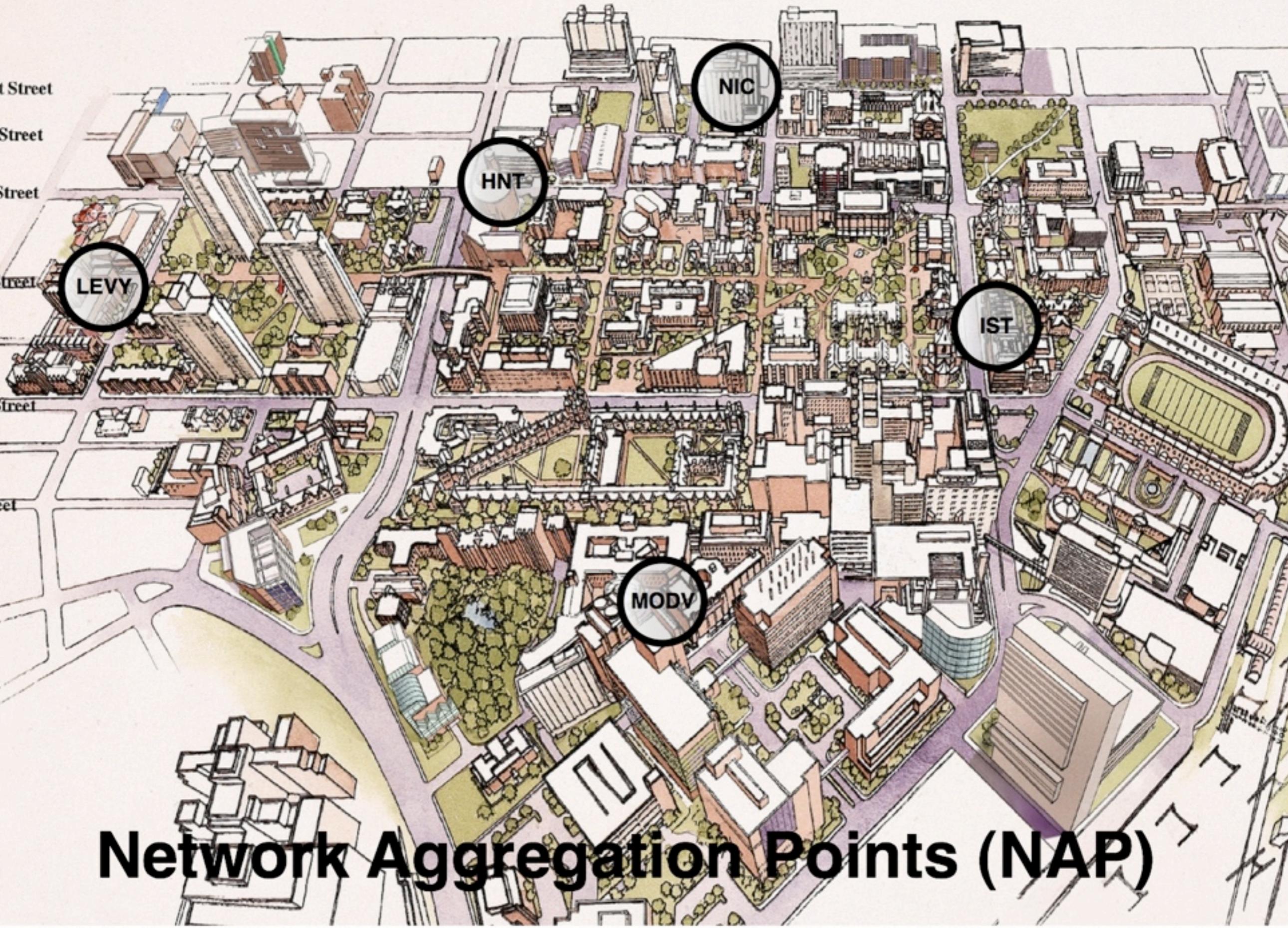
- 2001:468:1802::/48
- 2607:f470::/32
- **MAGPI (AS 10466)**
- 216.27.96.0/20
- 2001:468:1800::/40

Routing Protocols

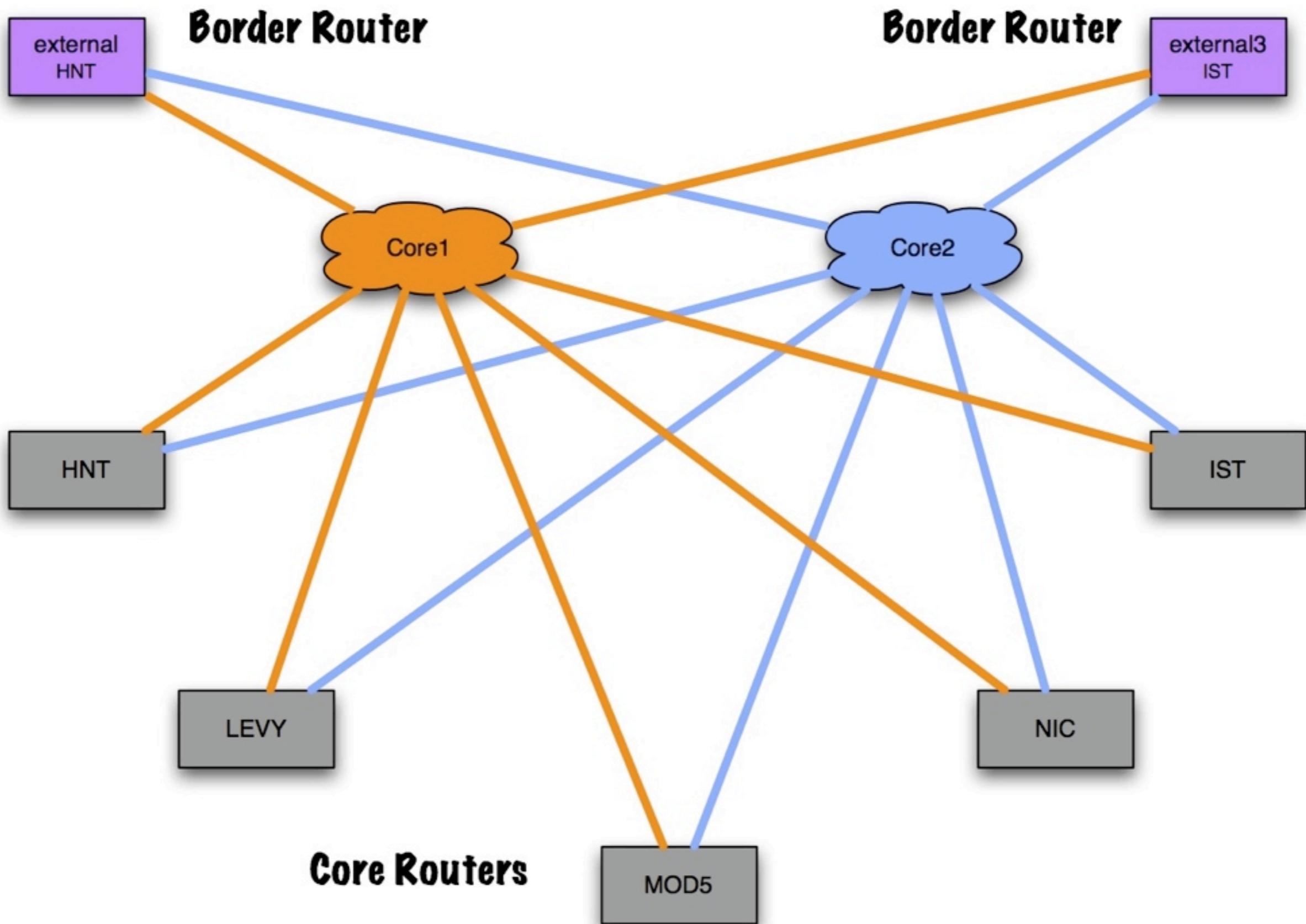
- Interior Routing
 - OSPF version 2 (IPv4)
 - IS-IS (IPv6)
- Exterior Routing
 - BGP4
- Multicast Routing
 - IGMP, PIM Sparse Mode, MSDP, M-BGP

40th Street 39th Street 38th Street 37th Street 36th Street 34th Street 33rd Street 32nd Street

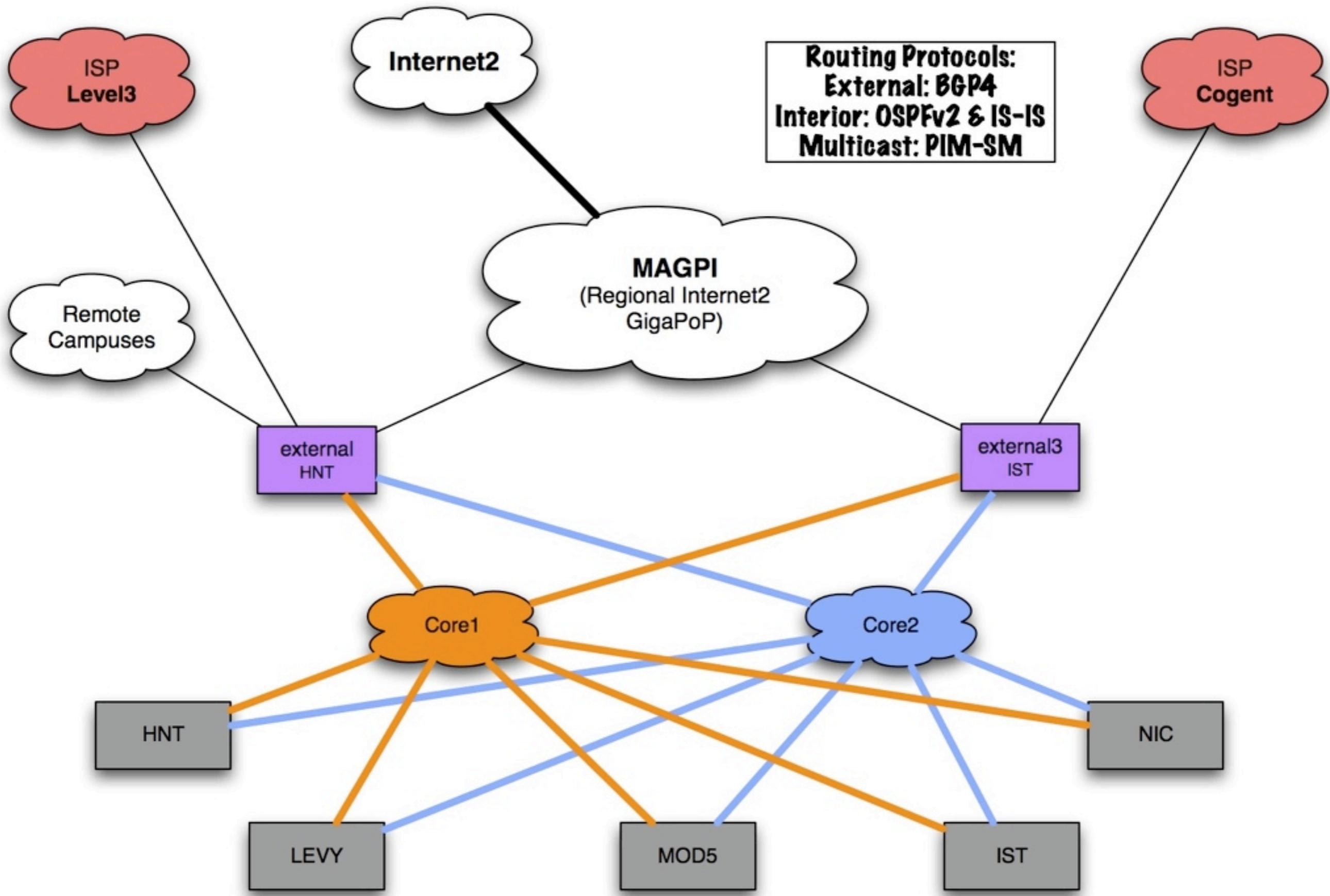
Chestnut Street
Sansom Street
Walnut Street
Locust Street
Spruce Street
Pine Street



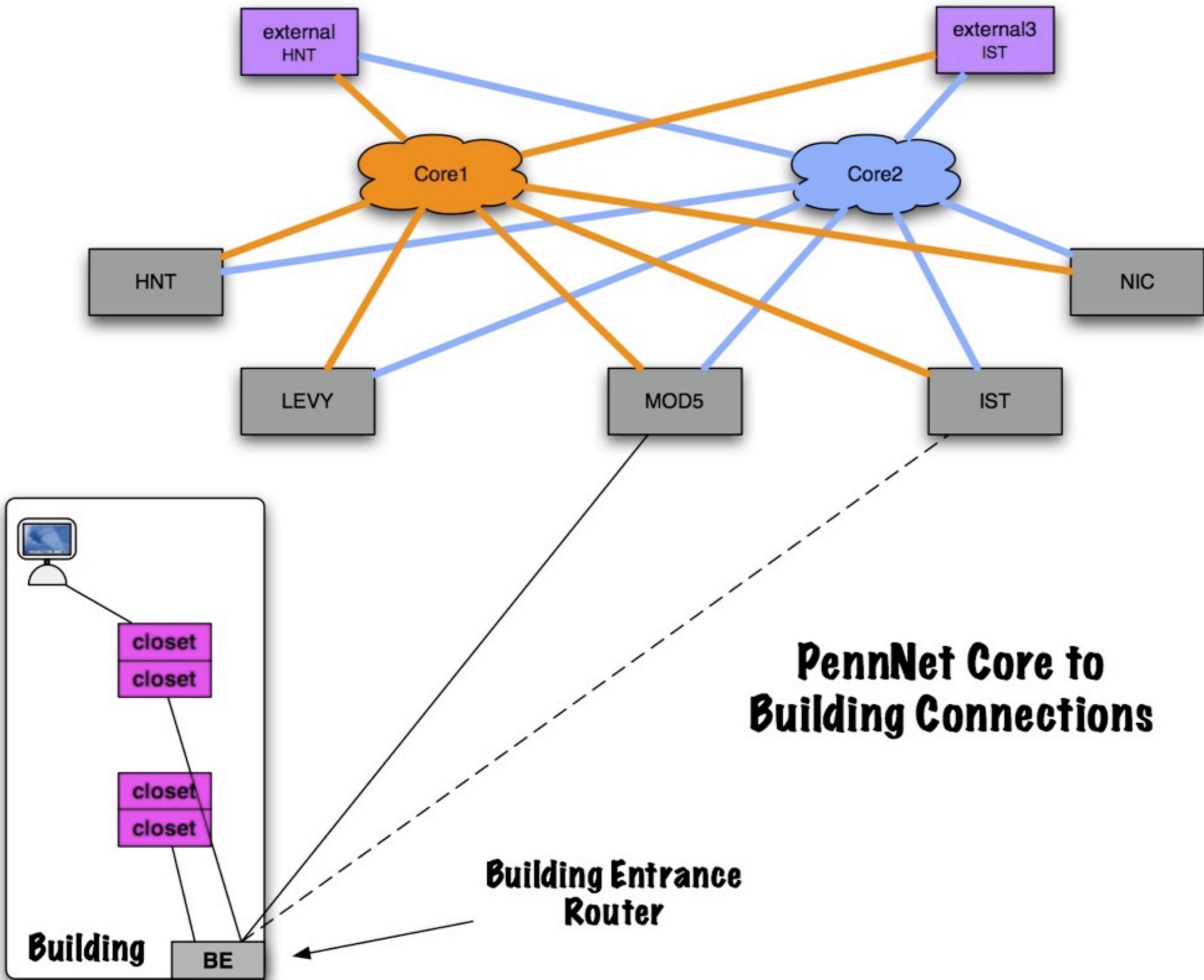
Network Aggregation Points (NAP)



PennNet 10Gig Routing Core

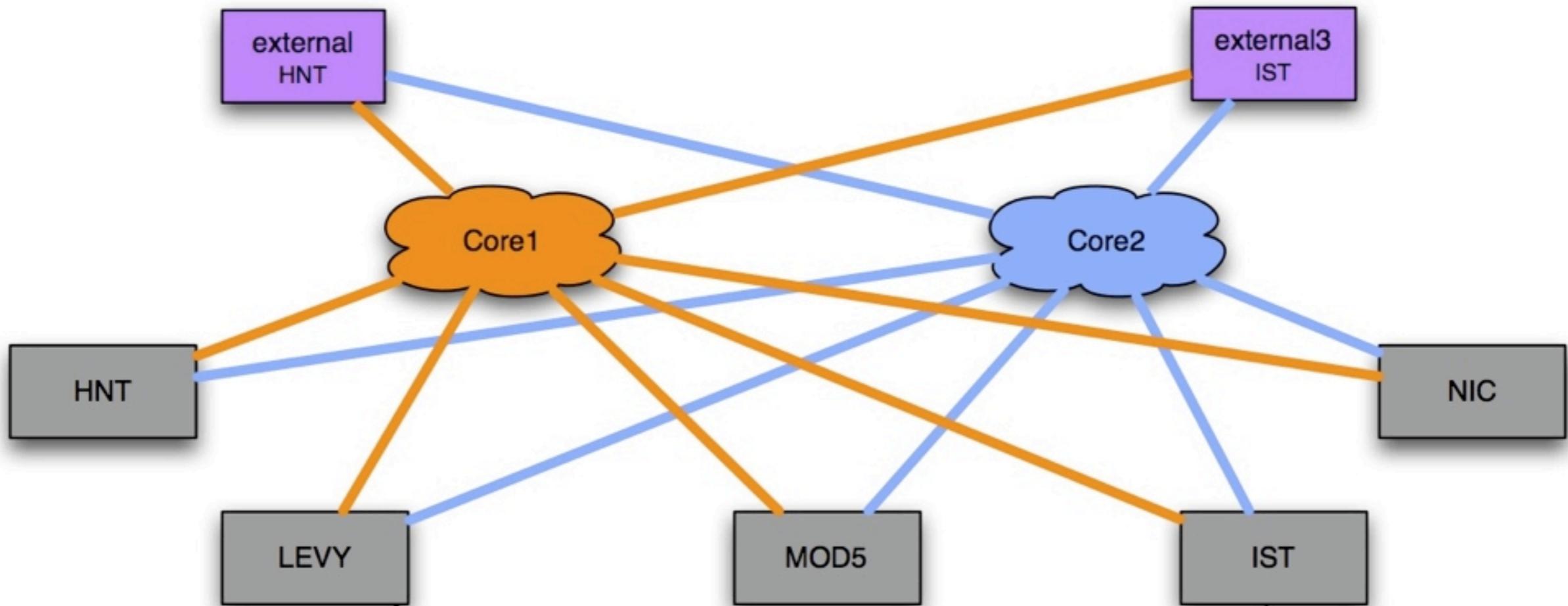


PennNet External Connectivity



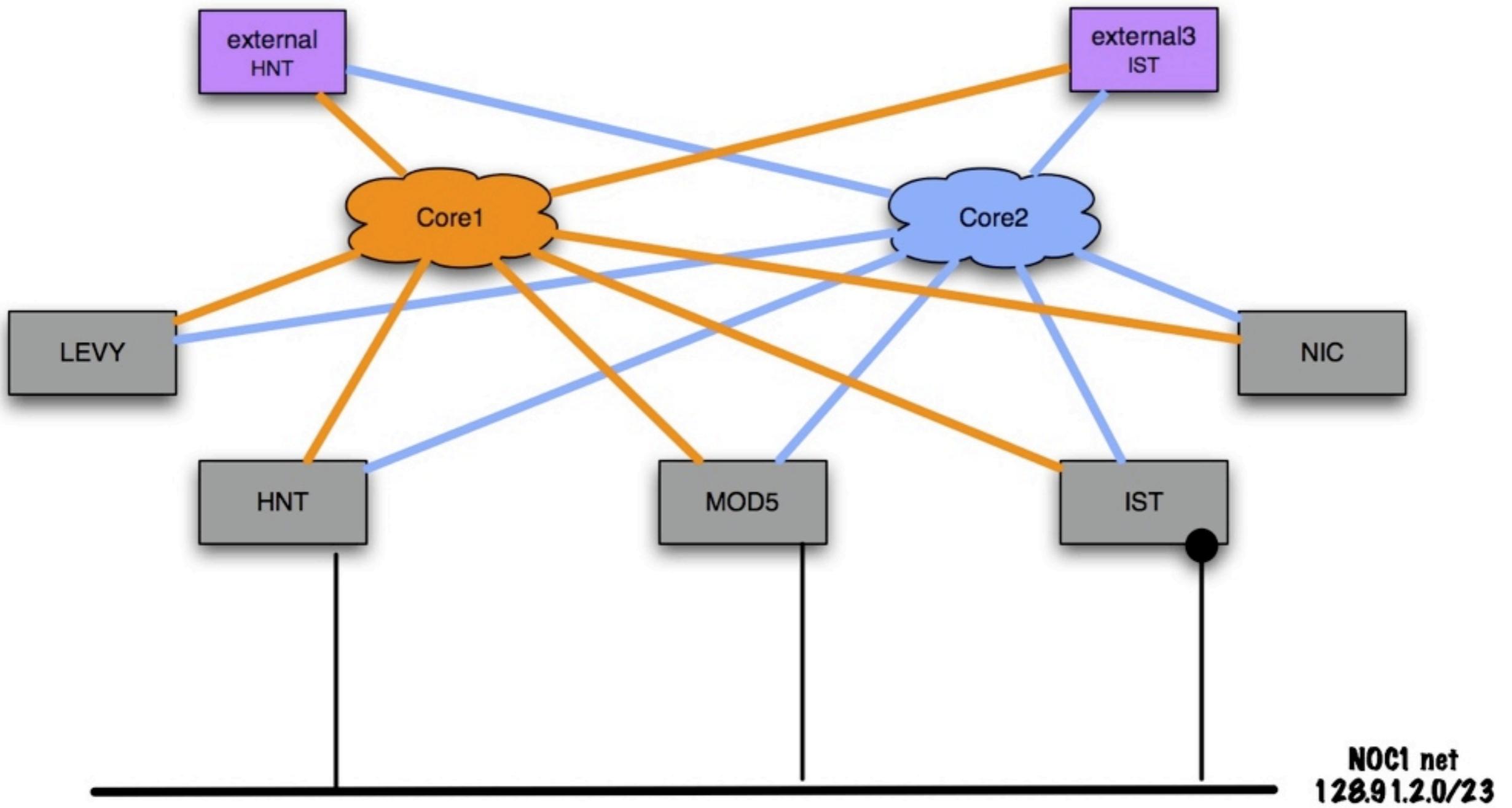
PennNet Core to Building Connections

Building Entrance Router

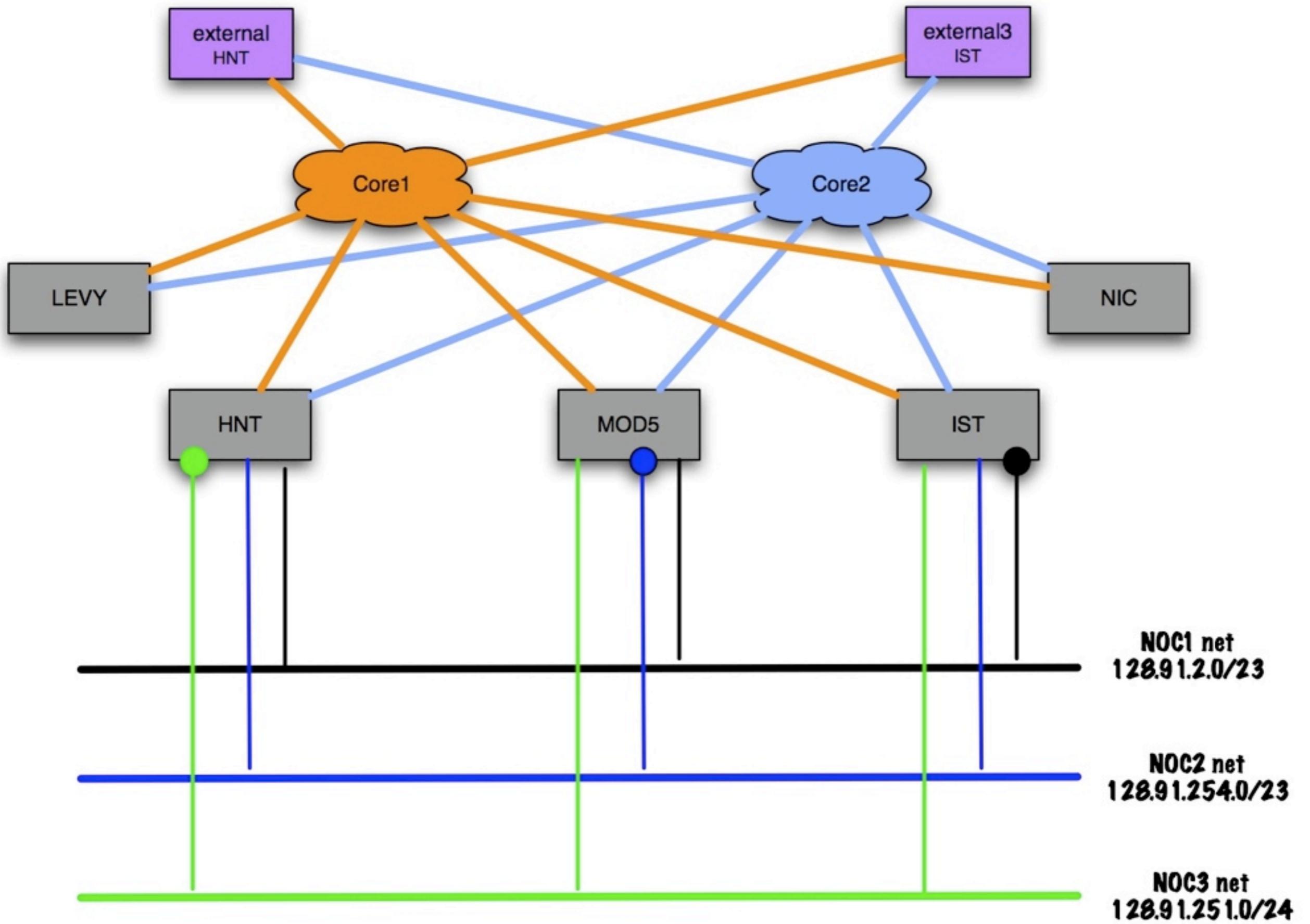


PennNet Core to School Connections

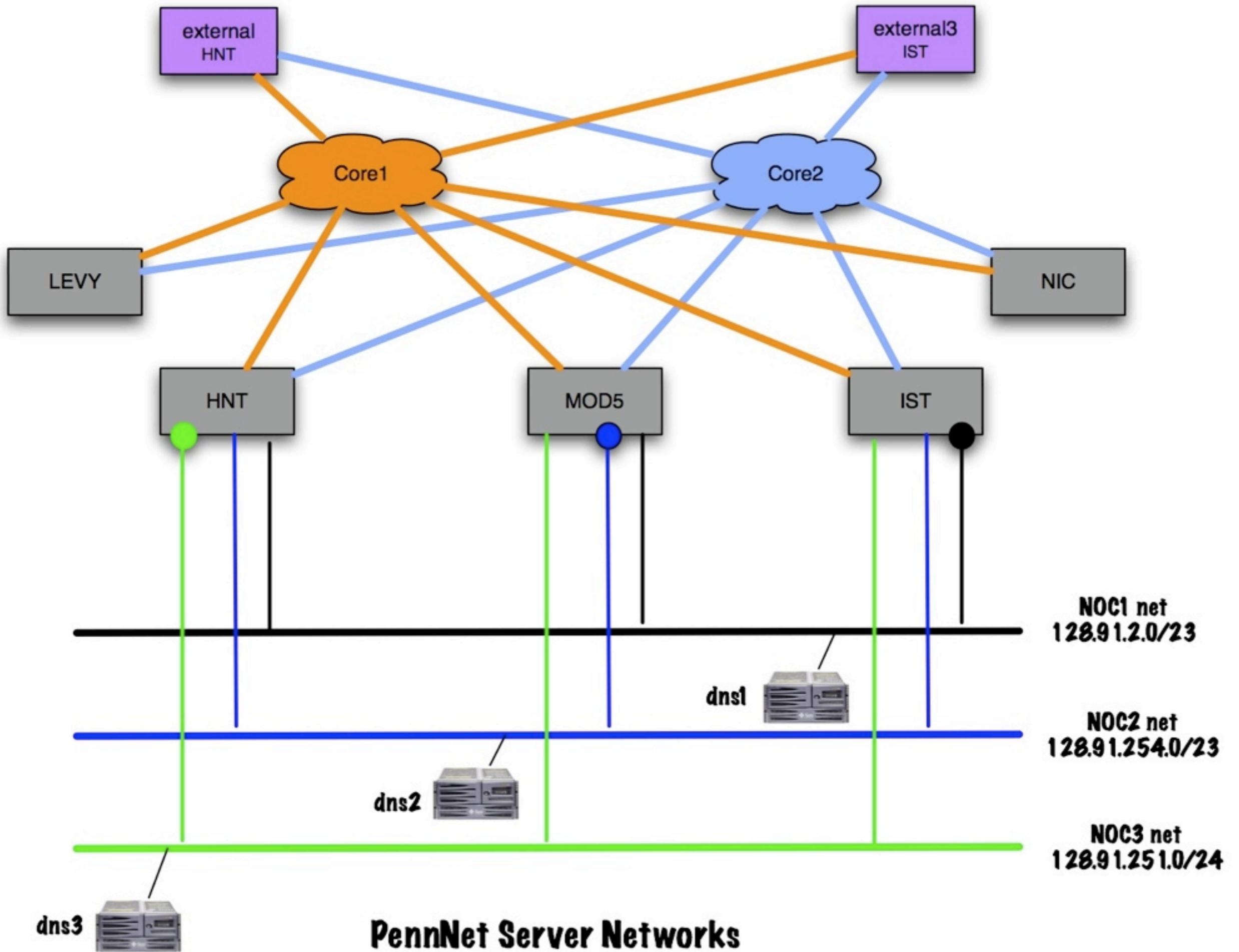


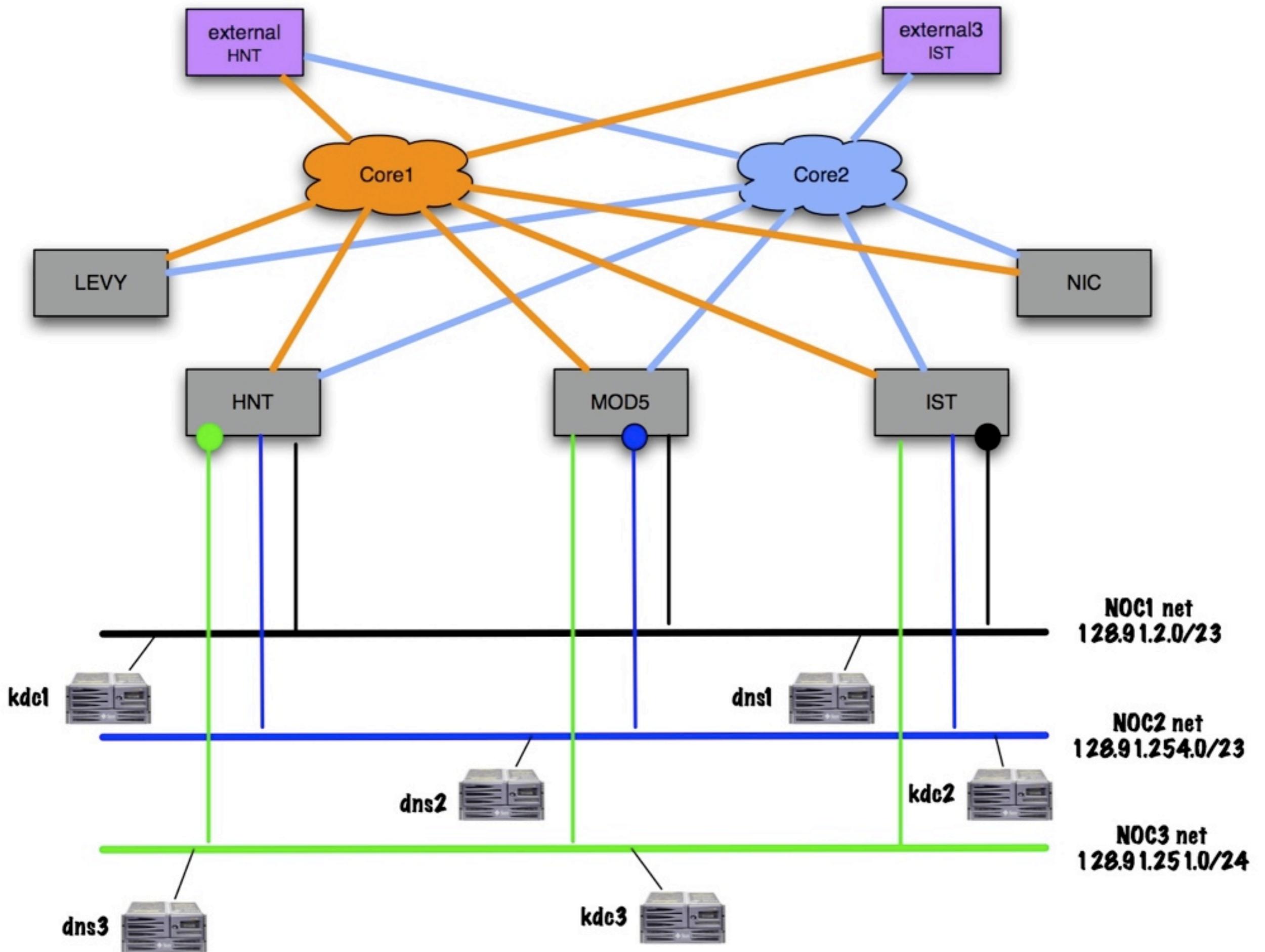


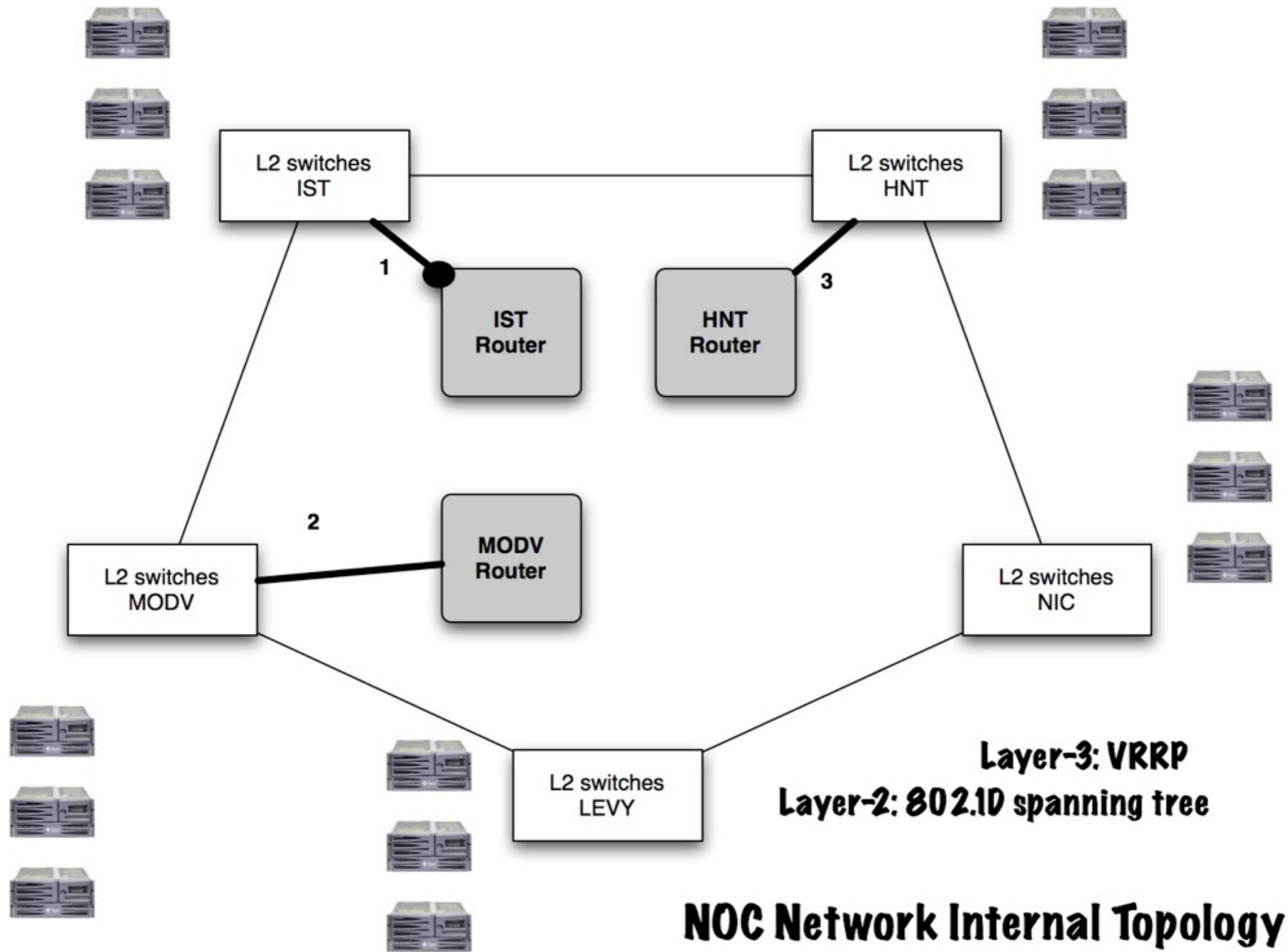
PennNet Server Networks

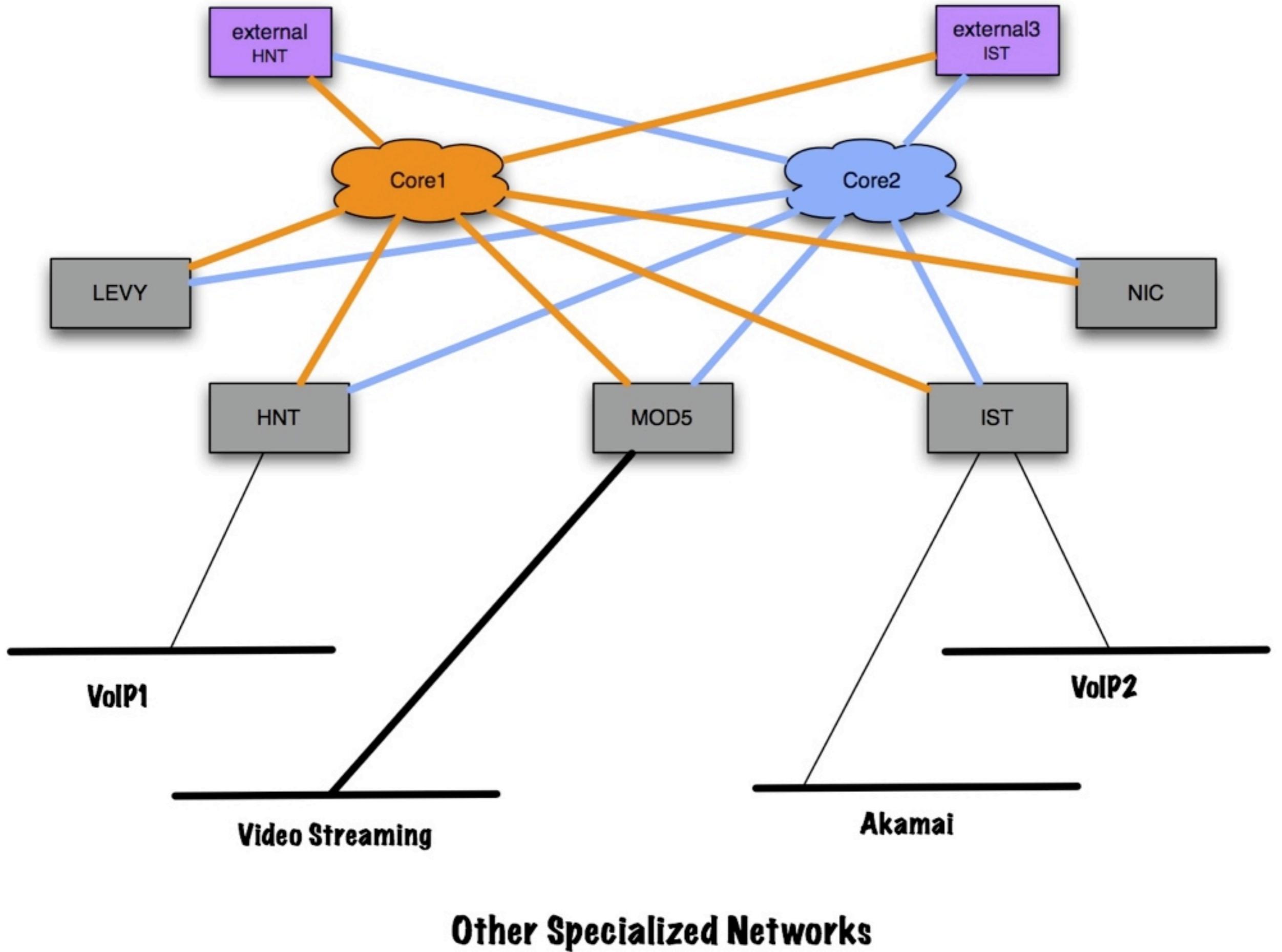


PennNet Server Networks

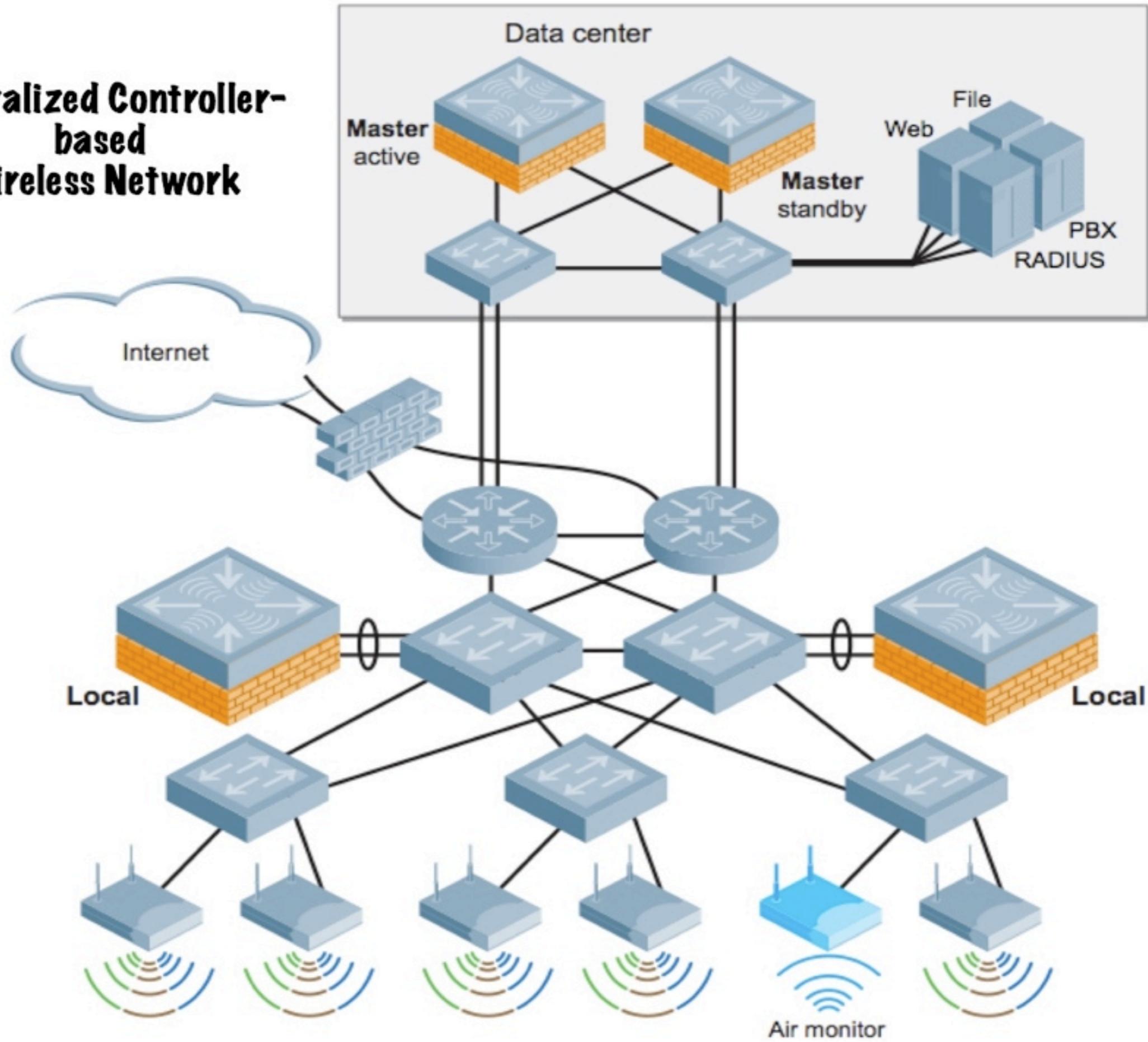








Centralized Controller-based Wireless Network

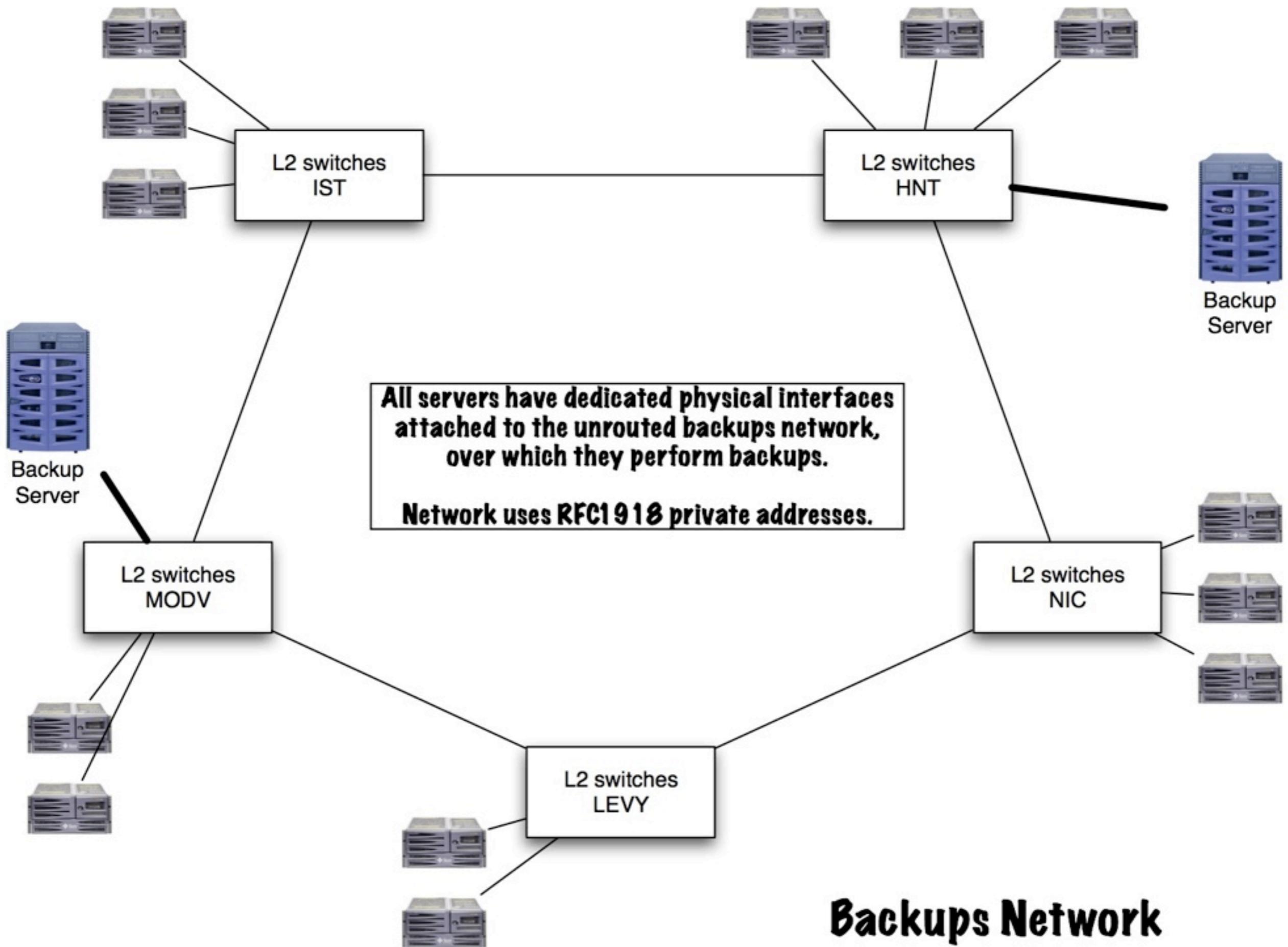


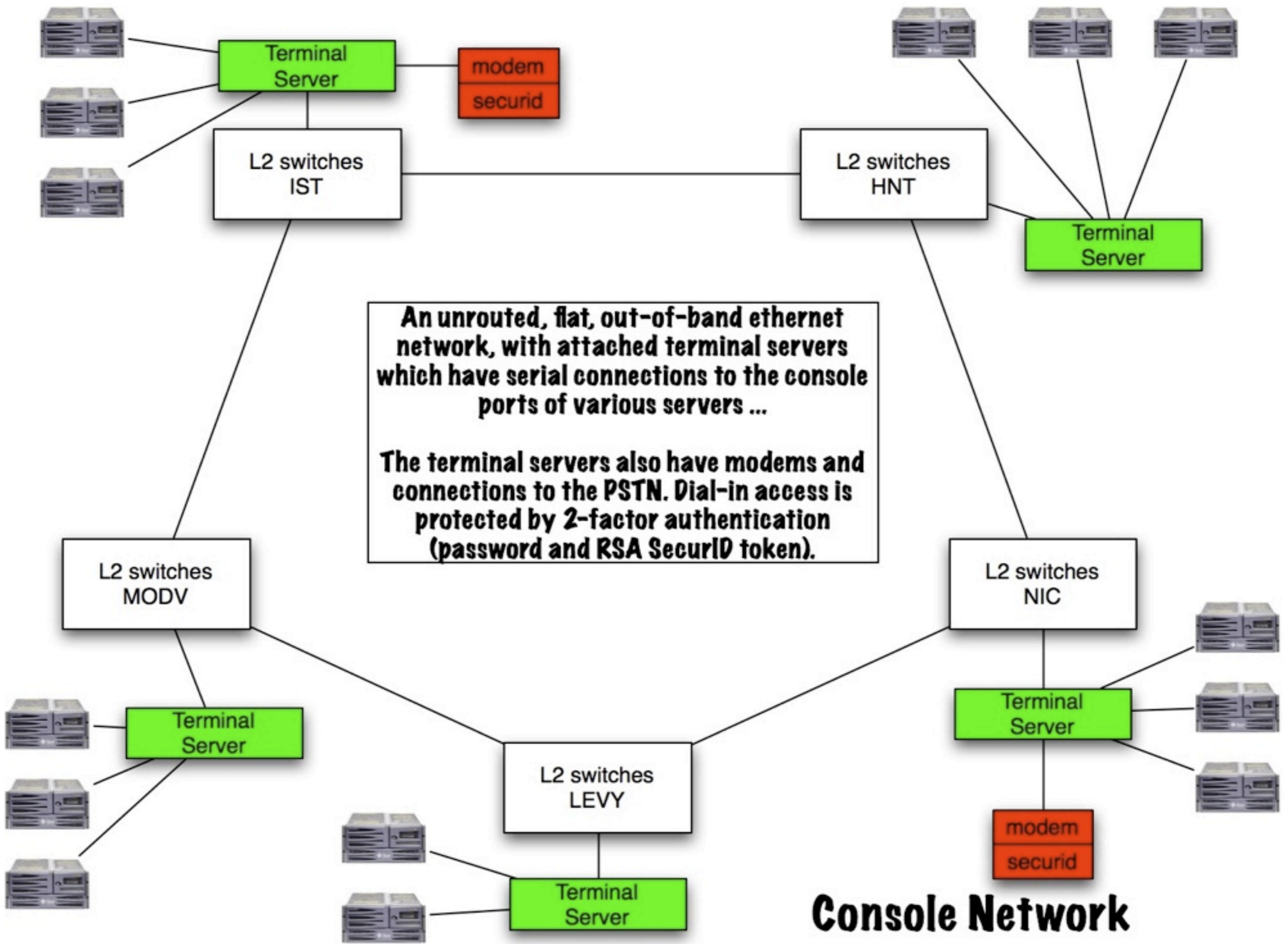
Wireless Network

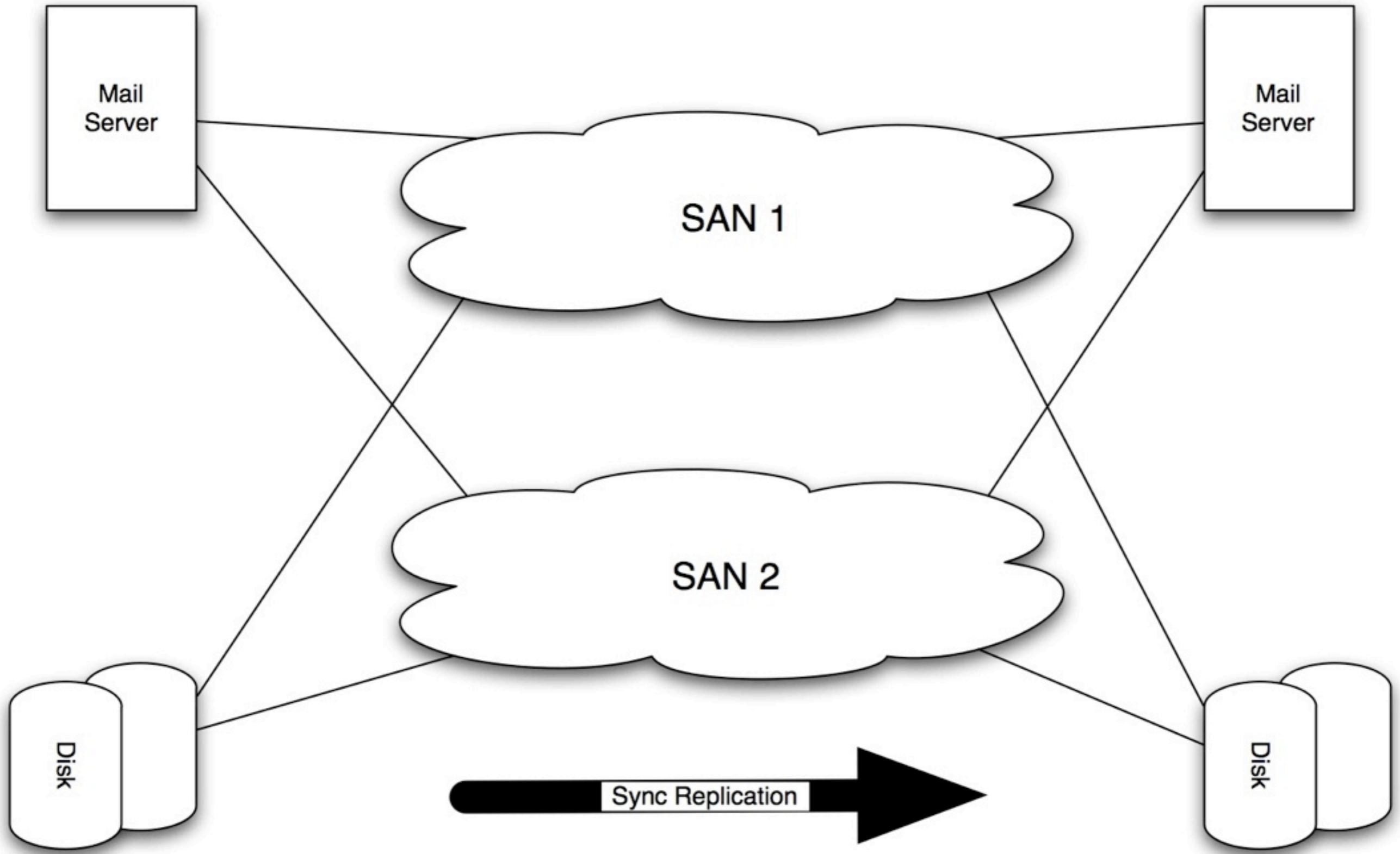
- Centralized Controller based architecture
- Lightweight Access Points
- Security: WPA (interim)
- Full Wireless Sec standard coming (802.11i or WPA2)
- 802.1x authentication (method: EAP-TTLS)
- Controllers act as pass through EAP authenticators
- EAP server: RADIUS
- RADIUS servers perform Kerberos 5 password verification

Out-of-band networks

- Several out-of-band networks exist
- Not connected to the Internet in any way
- Usually unrouted, flat Layer-2 networks
- Physically separate fiber infrastructure
- Private address space (RFC 1918)
- Uses: storage area network, server backups, console access, netflow collection, etc ...







FibreChannel Storage Area Networks

MAGPI

- Mid-Atlantic GigaPoP in Philadelphia for Internet2
- A regional network offering advanced services and connectivity to national and international R&E backbone networks, like Internet2
- Clients include many regional universities, colleges, hospitals, museums, and other non-profits
- Advanced services include IP Multicast and IPv6
- Also offers commercial Internet connectivity

Internet2

- Consortium of American universities
- Jointly fund and operate a high performance R&E backbone network
- Access model is via regional aggregation networks, called GigaPoPs (or RONs)

MAGPI

- Two points of presence:
 - Penn Campus
 - Carrier Hotel at 401 North Broad St
- Locations are connected by dark fiber ring
- DWDM nodes at each end carry misc circuits and services

Research & Education
Backbone Networks

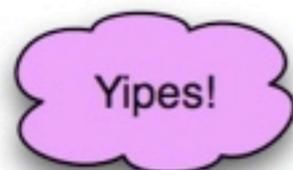


MAGPI GigaPop Peerings

October 2004

Commercial ISPs

Private peerings



Institutional Participants

Regular I2 members:

University of Pennsylvania
Princeton University
University of Delaware
Rutgers University
Temple University
Lehigh University
NJ Institute of Technology
Seton Hall University
Widener University
Fox Chase Cancer Center
University of Med & Dentistry of NJ

Sponsored participants/Affiliates:

Arcadia University
Franklin Institute Science Museum
Children's Hospital of Philadelphia
Thomas Jefferson University
Cermusa/St. Francis University
C.A.P.E.

Other institutions:

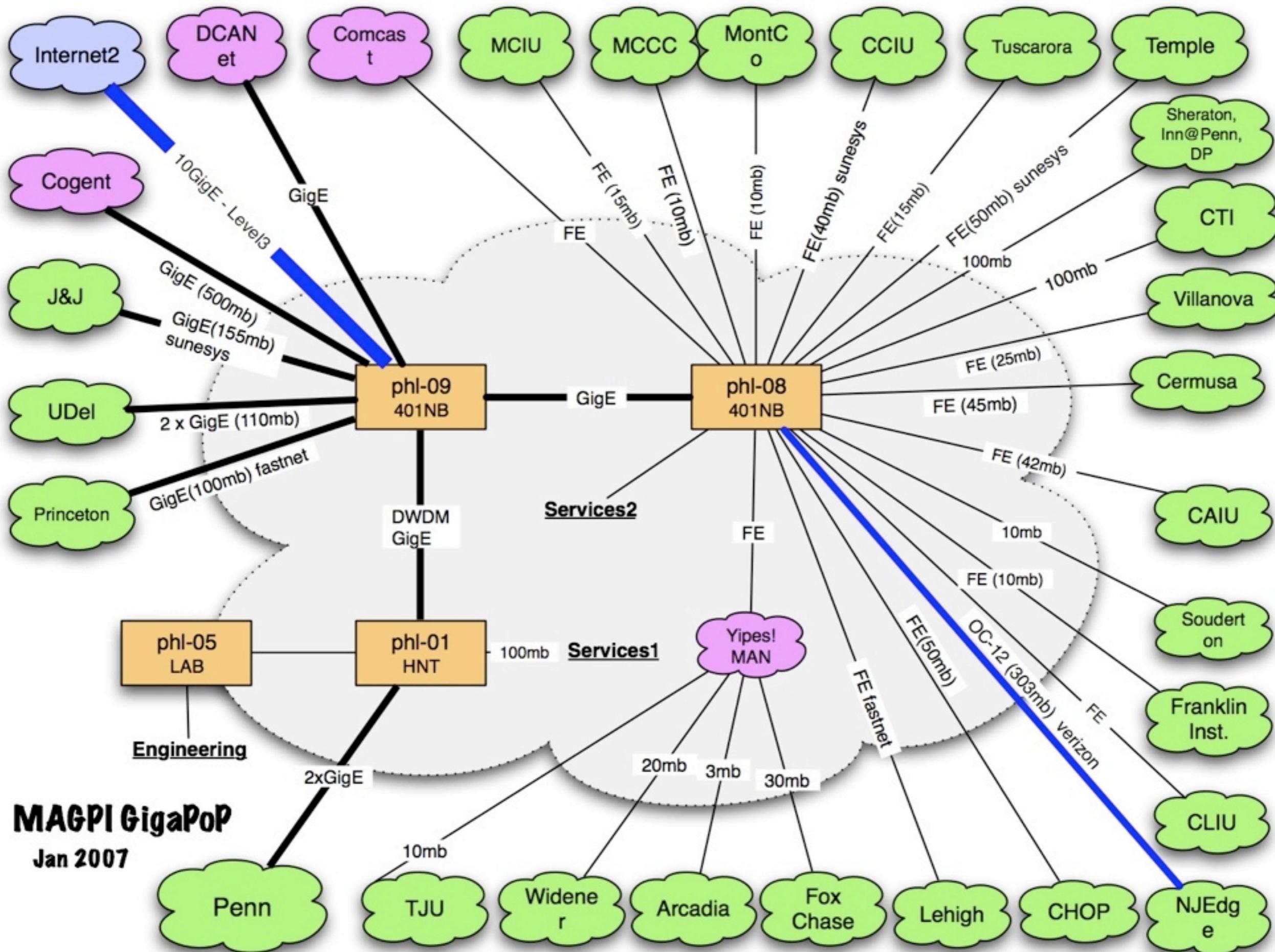
The Daily Pennsylvanian
University City Sheraton Hotel
The Inn at Penn
Campus Technologies Inc.

SEGP and K-12 units:

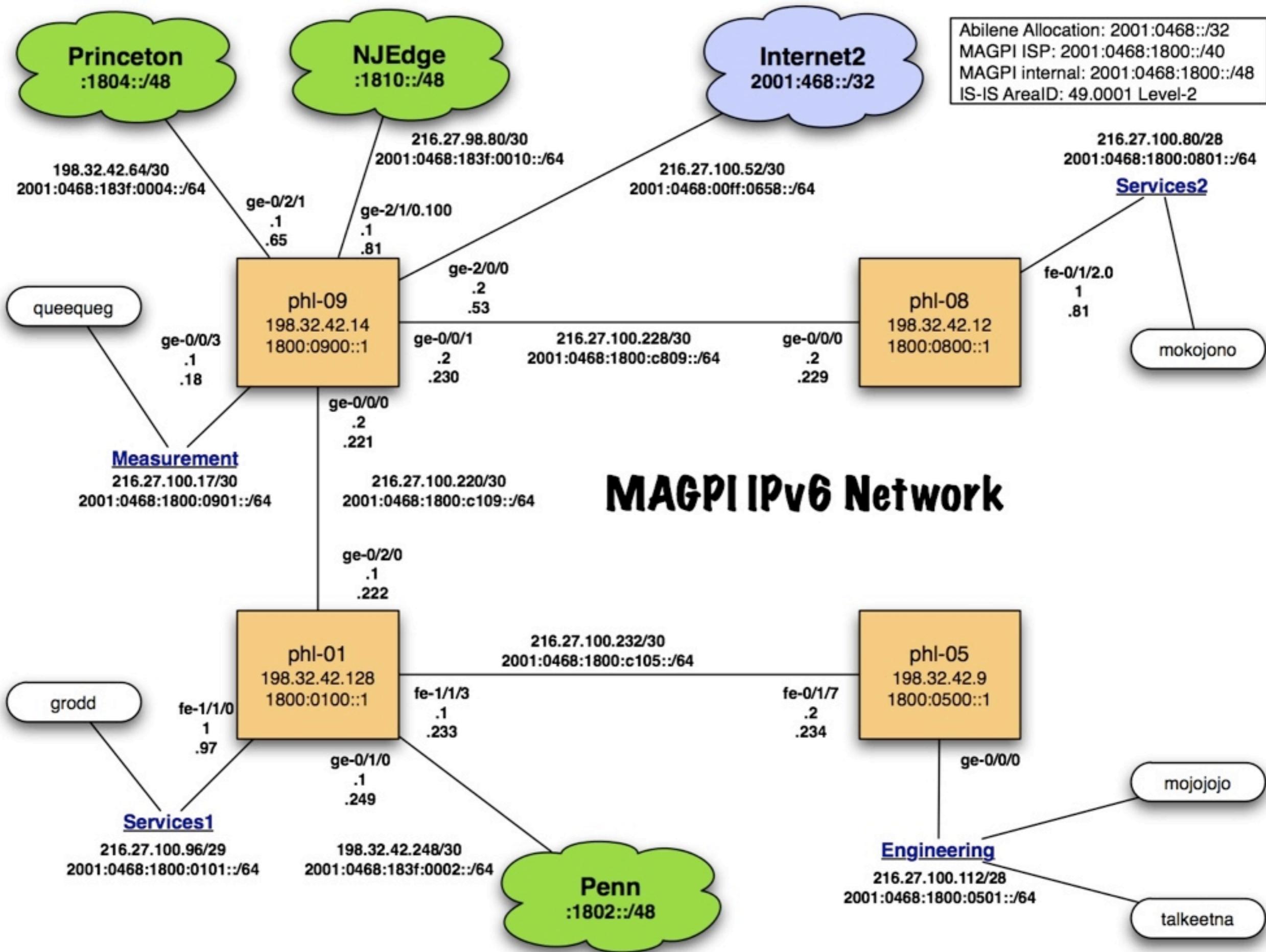
NJEdge
Chester County Intermediate Unit
Carbon Lehigh Intermediate Unit
Capital Area Intermediate Unit
Delaware County Intermediate Unit
Montgomery County Intermediate Unit
Montgomery County Community Coll.
(75 PA K-12 school districts)

Corporate members

Johnson & Johnson Pharmaceuticals



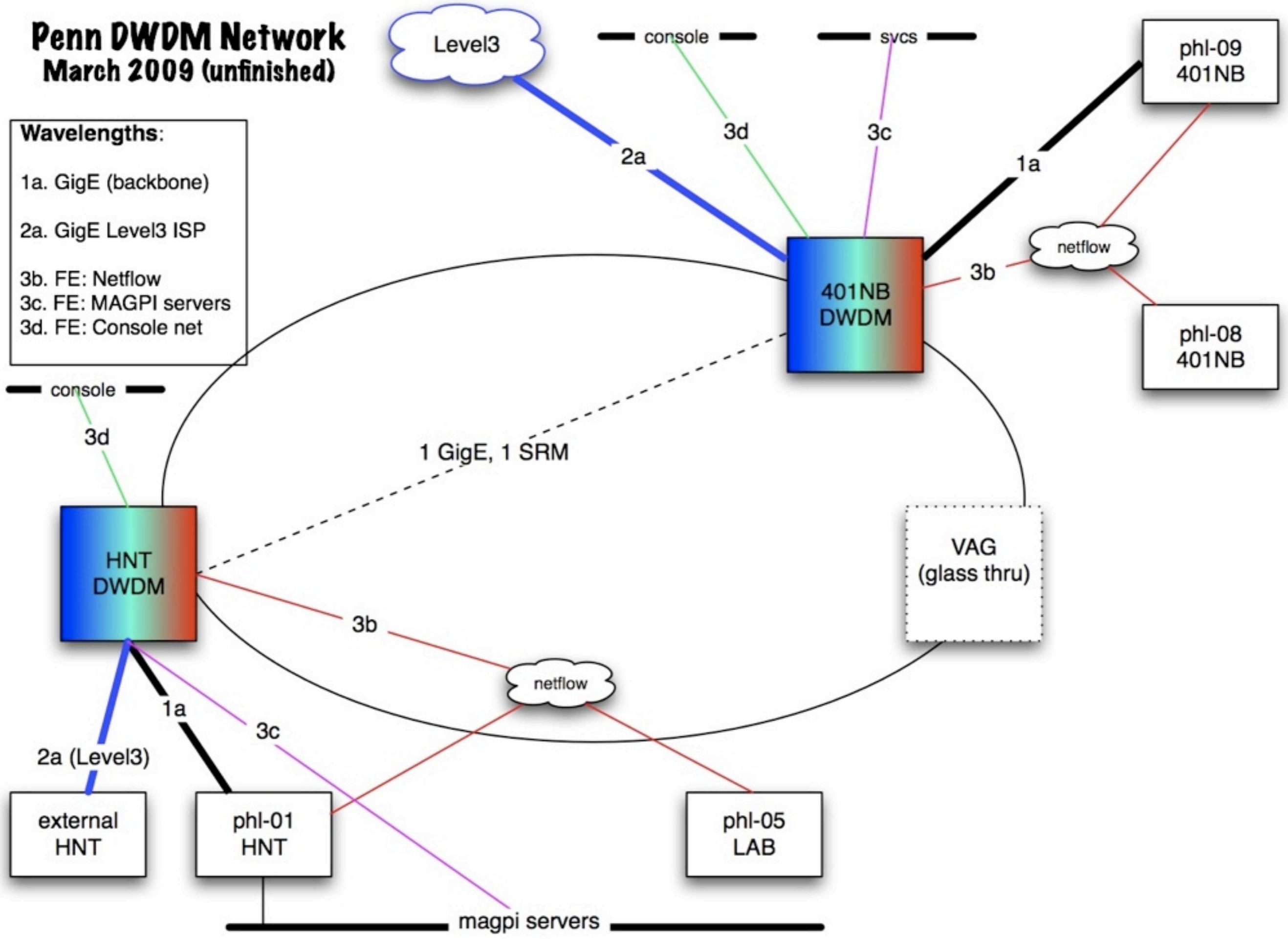
MAGPI GigaPop
Jan 2007



Penn DWDM Network

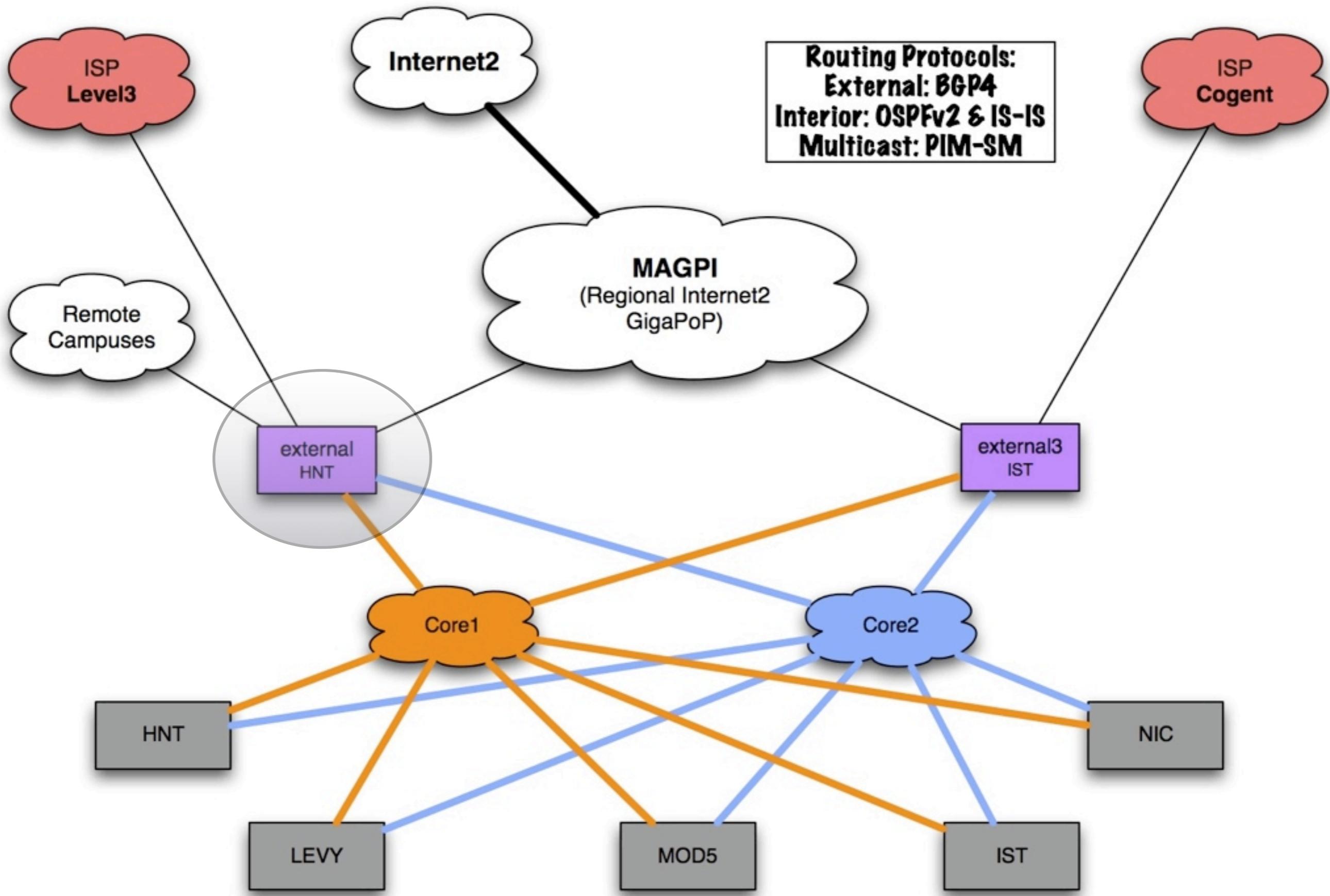
March 2009 (unfinished)

- Wavelengths:**
- 1a. GigE (backbone)
 - 2a. GigE Level3 ISP
 - 3b. FE: Netflow
 - 3c. FE: MAGPI servers
 - 3d. FE: Console net



Back to BGP ...

Let's look at the BGP configuration and state at a Penn border router



PennNet External Connectivity

shuque@external> show bgp summary

Groups: 6 Peers: 7 Down peers: 0

Table	Tot Paths	Act Paths	Suppressed	History	Damp	State	Pending
inet6.0	1730	1730	0	0	0	0	0
inet6.2	63	63	0	0	0	0	0
inet.0	682626	280778	0	0	0	0	0
inet.2	4491	4491	0	0	0	0	0

Peer	AS	InPkt	OutPkt	OutQ	Flaps	Last Up/Dwn	State #A
------	----	-------	--------	------	-------	-------------	----------

ctive/Received/Damped...

4.78.144.21 3356 416741 16558 0 0 5d 6:27:56 Establ

inet.0: 201708/276158/0

128.91.9.2 55 401562 217835 0 0 5d 6:28:01 Establ

inet.0: 66848/197123/0

128.91.10.2 55 404571 217863 0 0 5d 6:28:14 Establ

inet.0: 0/197123/0

128.91.63.8 55 11823 326457 0 0 4d 3:01:58 Establ

inet.0: 0/0/0

193.108.155.102 21357 12414 13547 0 2 4d 7:26:10 Establ

inet.0: 0/0/0

216.27.100.73 10466 41134 16560 0 0 5d 6:28:39 Establ

inet.0: 12222/12222/0

inet.2: 4491/4491/0

2001:468:183f:2::1 10466 34599 16558 0 0 5d 6:27:53 Estab

l

inet6.0: 1730/1730/0

inet6.2: 63/63/0

Level3 ebgp peering

shuque@external> show bgp neighbor 4.78.144.21

Peer: 4.78.144.21+48812 AS 3356 Local: 4.78.144.22+179 AS 55

Description: Level3 ISP

Type: External State: Established Flags: <Sync>

Last State: OpenConfirm Last Event: RecvKeepAlive

Last Error: None

Export: [OUTBOUND-LEVEL3-PREPEND-BGP] Import: [INBOUND-LEVEL3-BGP]

Options: <Preference PeerAS Refresh>

Holdtime: 90 Preference: 170

Number of flaps: 0

Peer ID: 4.68.0.60 Local ID: 192.84.2.254 Active Holdtime: 90

Keepalive Interval: 30 Peer index: 0

BFD: disabled, down

Local Interface: ge-2/1/2.0

NLRI advertised by peer: inet-unicast

NLRI for this session: inet-unicast

Peer supports Refresh capability (2)

Table inet.0 Bit: 30004

RIB State: BGP restart is complete

Send state: in sync

Active prefixes: 201713

Received prefixes: 276158

Suppressed due to damping: 0

Advertised prefixes: 5

Last traffic (seconds): Received 24 Sent 21 Checked 39

Input messages: Total 416832 Updates 401646 Refreshes 0 Octets 28637686

Output messages: Total 16570 Updates 0 Refreshes 0 Octets 315132

Output Queue[2]: 0

> show route **advertising**-protocol bgp 4.78.144.21

inet.0: 281817 destinations, 683677 routes (281816 active, 0 holddown, 1 hidden)

Prefix	Nexthop	MED	Lclpref	AS path
* 128.91.0.0/16	Self			55 [55] I
* 130.91.0.0/16	Self			55 [55] I
* 158.130.0.0/16	Self			55 [55] I
* 165.123.0.0/16	Self			55 [55] I
* 192.84.2.0/24	Self			55 [55] I

> show route **receive**-protocol bgp 4.78.144.21

inet.0: 281825 destinations, 683679 routes (281816 active, 8 holddown, 1 hidden)

Prefix	Nexthop	MED	Lclpref	AS path
* 3.0.0.0/8	4.78.144.21	0		3356 15412 930
4 80 I				
* 4.0.0.0/8	4.78.144.21	0		3356 I
* 4.0.0.0/9	4.78.144.21	0		3356 I
* 4.128.0.0/9	4.78.144.21	0		3356 I
6.1.0.0/16	4.78.144.21	0		3356 701 668 I
6.2.0.0/22	4.78.144.21	0		3356 701 668 I
6.3.0.0/18	4.78.144.21	0		3356 701 668 I
6.4.0.0/16	4.78.144.21	0		3356 701 668 I

[many more lines deleted]

MAGPI IPv4 Peering

> show bgp neighbor 216.27.100.73

Peer: 216.27.100.73+2667 AS 10466 Local: 216.27.100.74+179 AS 55

Type: External State: Established Flags: <Sync>

Last State: OpenConfirm Last Event: RecvKeepAlive

Last Error: None

Export: [export-penn-routes] Import: [INBOUND-MAGPI-BGP]

Options: <Preference AuthKey AddressFamily PeerAS Refresh>

Authentication key is configured

Address families configured: inet-unicast inet-multicast

Holdtime: 90 Preference: 170

Number of flaps: 0

Peer ID: 216.27.100.246 Local ID: 192.84.2.254 Active Holdtime: 90

Keepalive Interval: 30 Peer index: 0

BFD: disabled, down

Local Interface: ge-2/1/3.550

NLRI advertised by peer: inet-unicast inet-multicast

NLRI for this session: inet-unicast inet-multicast

Peer supports Refresh capability (2)

Table inet.0 Bit: 30001

RIB State: BGP restart is complete

Send state: in sync

Active prefixes: 12222

Received prefixes: 12222

Suppressed due to damping: 0

Advertised prefixes: 5

Table inet.2 Bit: 40000

RIB State: BGP restart is complete

Send state: in sync

Active prefixes: 4491

Received prefixes: 4491

Suppressed due to damping: 0

Advertised prefixes: 4

```
> show route advertising-protocol bgp 216.27.100.73
inet.0: 281826 destinations, 683687 routes (281817 active, 8 holddown, 1 hidden)
  Prefix                Nexthop                MED    Lclpref    AS path
* 128.91.0.0/16        Self                    I
* 130.91.0.0/16        Self                    I
* 158.130.0.0/16       Self                    I
* 165.123.0.0/16       Self                    I
* 192.84.2.0/24        Self                    I
```

```
inet.2: 4495 destinations, 4495 routes (4495 active, 0 holddown, 0 hidden)
  Prefix                Nexthop                MED    Lclpref    AS path
* 128.91.0.0/16        Self                    I
* 130.91.0.0/16        Self                    I
* 165.123.0.0/16       Self                    I
* 192.84.2.0/24        Self                    I
```

```
shuque@external> show route receive-protocol bgp 216.27.100.73
inet.0: 281826 destinations, 683682 routes (281815 active, 10 holddown, 1 hidden
)
  Prefix                Nexthop                MED    Lclpref    AS path
* 4.79.201.0/26        216.27.100.73          I      I          10466 11537 10
886 40220 I
* 6.1.0.0/16           216.27.100.73          I      I          10466 11537 66
8 I
* 6.2.0.0/22           216.27.100.73          I      I          10466 11537 66
8 I
* 6.3.0.0/18           216.27.100.73          I      I          10466 11537 66
8 I
* 6.4.0.0/16           216.27.100.73          I      I          10466 11537 66
8 I
[many more lines deleted ..]
```

shuque@external> show bgp summary

Groups: 6 Peers: 7 Down peers: 0

Table	Tot Paths	Act Paths	Suppressed	History	Damp	State	Pending
inet6.0	1730	1730	0	0	0	0	0
inet6.2	63	63	0	0	0	0	0
inet.0	682626	280778	0	0	0	0	0
inet.2	4491	4491	0	0	0	0	0

Peer	AS	InPkt	OutPkt	OutQ	Flaps	Last Up/Dwn	State #A
4.78.144.21	3356	416741	16558	0	0	5d 6:27:56	Establ
inet.0: 201708/276158/0							
128.91.9.2	55	401562	217835	0	0	5d 6:28:01	Establ
inet.0: 66848/197123/0							
128.91.10.2	55	404571	217863	0	0	5d 6:28:14	Establ
inet.0: 0/197123/0							
128.91.63.8	55	11823	326457	0	0	4d 3:01:58	Establ
inet.0: 0/0/0							
193.108.155.102	21357	12414	13547	0	2	4d 7:26:10	Establ
inet.0: 0/0/0							
216.27.100.73	10466	41134	16560	0	0	5d 6:28:39	Establ
inet.0: 12222/12222/0							
inet.2: 4491/4491/0							
2001:468:183f:2::1	10466	34599	16558	0	0	5d 6:27:53	Establ
inet6.0: 1730/1730/0							
inet6.2: 63/63/0							

```
> show route www.mit.edu
```

```
inet.0: 281832 destinations, 683694 routes (281831 active, 0 holddown, 1 hidden)
```

```
@ = Routing Use Only, # = Forwarding Use Only
```

```
+ = Active Route, - = Last Active, * = Both
```

```
18.0.0.0/8
```

```
*[BGP/170] 5d 06:54:12, localpref 220
```

```
AS path: 10466 11537 10578 3 I
```

```
> to 216.27.100.73 via ge-2/1/3.550
```

```
[BGP/170] 5d 06:53:14, MED 0, localpref 100
```

```
AS path: 3356 3 I
```

```
> to 4.78.144.21 via ge-2/1/2.0
```

```
[BGP/170] 1w3d 07:21:09, MED 13011, localpref 100
```

```
AS path: 174 3 3 3 I
```

```
to 128.91.9.2 via ge-2/0/0.0
```

```
> to 128.91.10.2 via ge-2/0/1.0
```

Active path (*) was selected by BGP's path selection algorithm, because it has a higher Local pref value. This path goes through the MAGPI peering (AS 10466, next hop 216.27.100.73). The path to www.mit.edu traverses 4 Autonomous systems (10466 MAGPI, 11537 Internet2, 10578 NOX/Harvard, and 3 MIT).

The 2nd path that was not preferred through Level3, actually has fewer AS hops. M.I.T. appears to be directly connected to Level3. The 3rd path is learned by internal BGP from the order border router through Cogent. MIT has prepended it's AS to this advertisement probably for traffic engineering reasons.

Questions?

Shumon Huque

Networking & Telecommunications

Information Systems & Computing

University of Pennsylvania

E-mail: [shuque -@- upenn.edu](mailto:shuque@upenn.edu)

Web: <http://www.huque.com/~shuque/>