ERIC BOYD

Senior Director, Strategic Projects

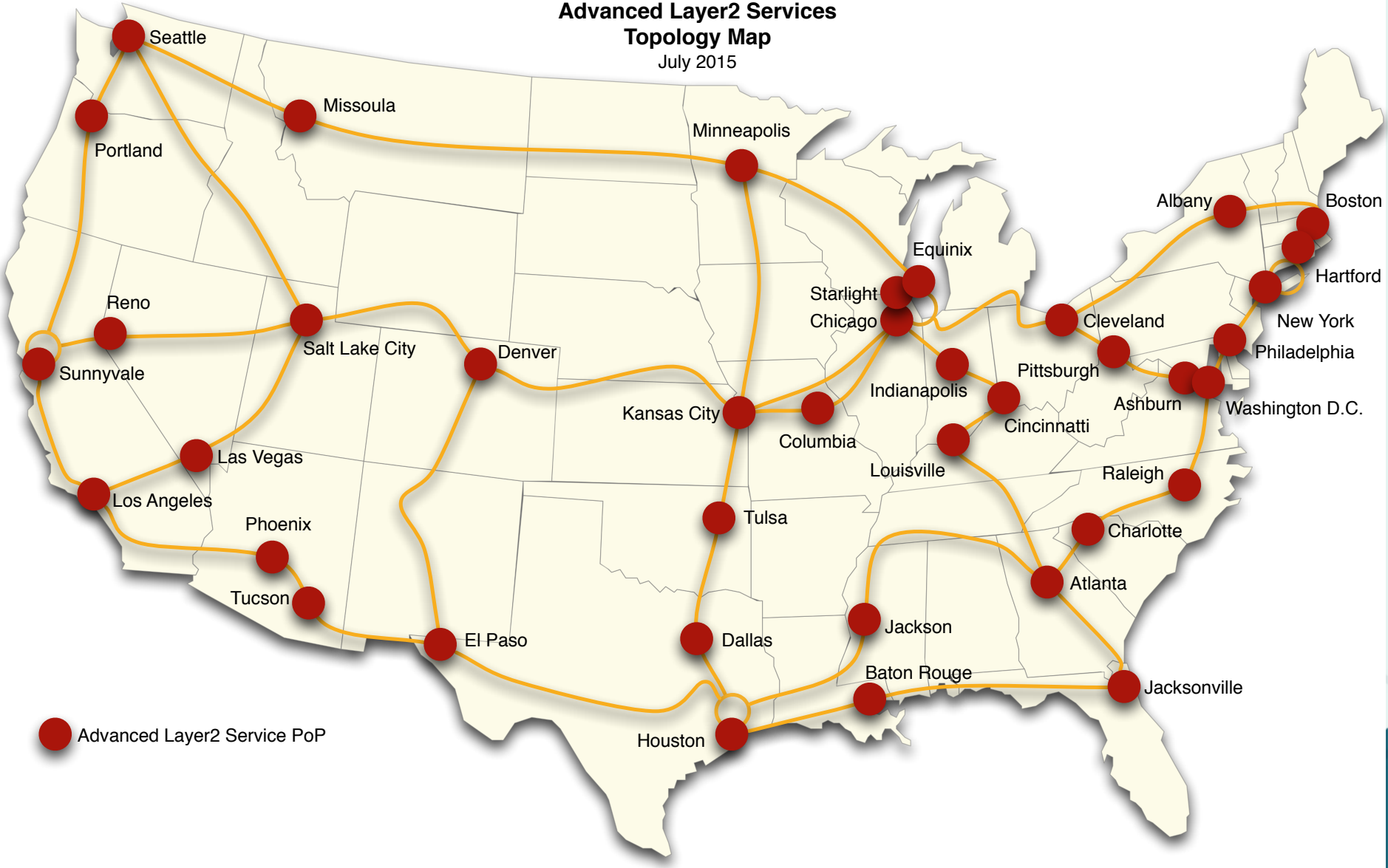
Growing The Internet2 Network above 100G: Choices and Challenges

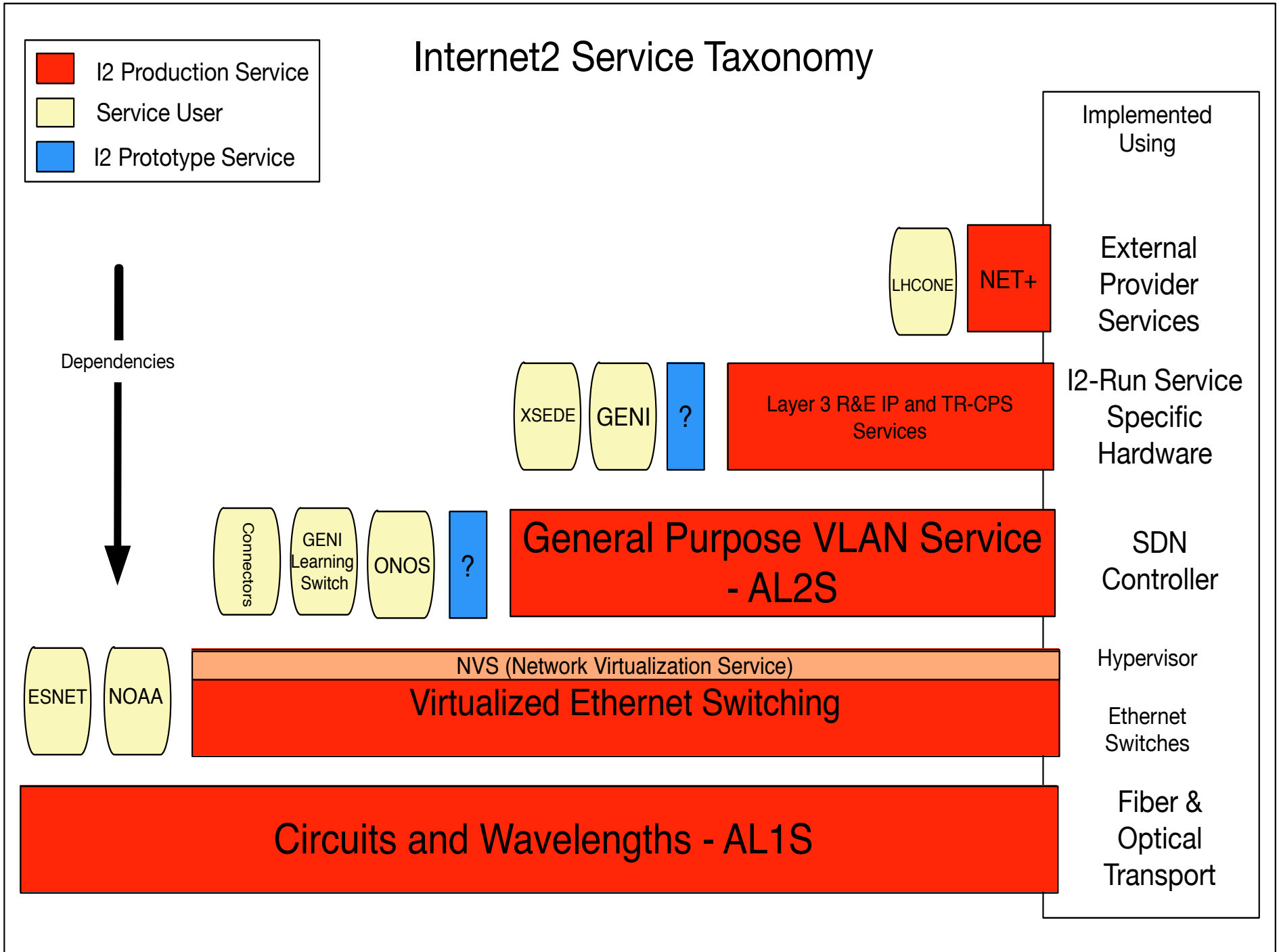
With help from Dale Finkelson, Ron Milford, Tom Johnson, and Chris Wilkinson

JULY, 2015

Internet2 Network

Advanced Layer2 Services
Topology Map
July 2015

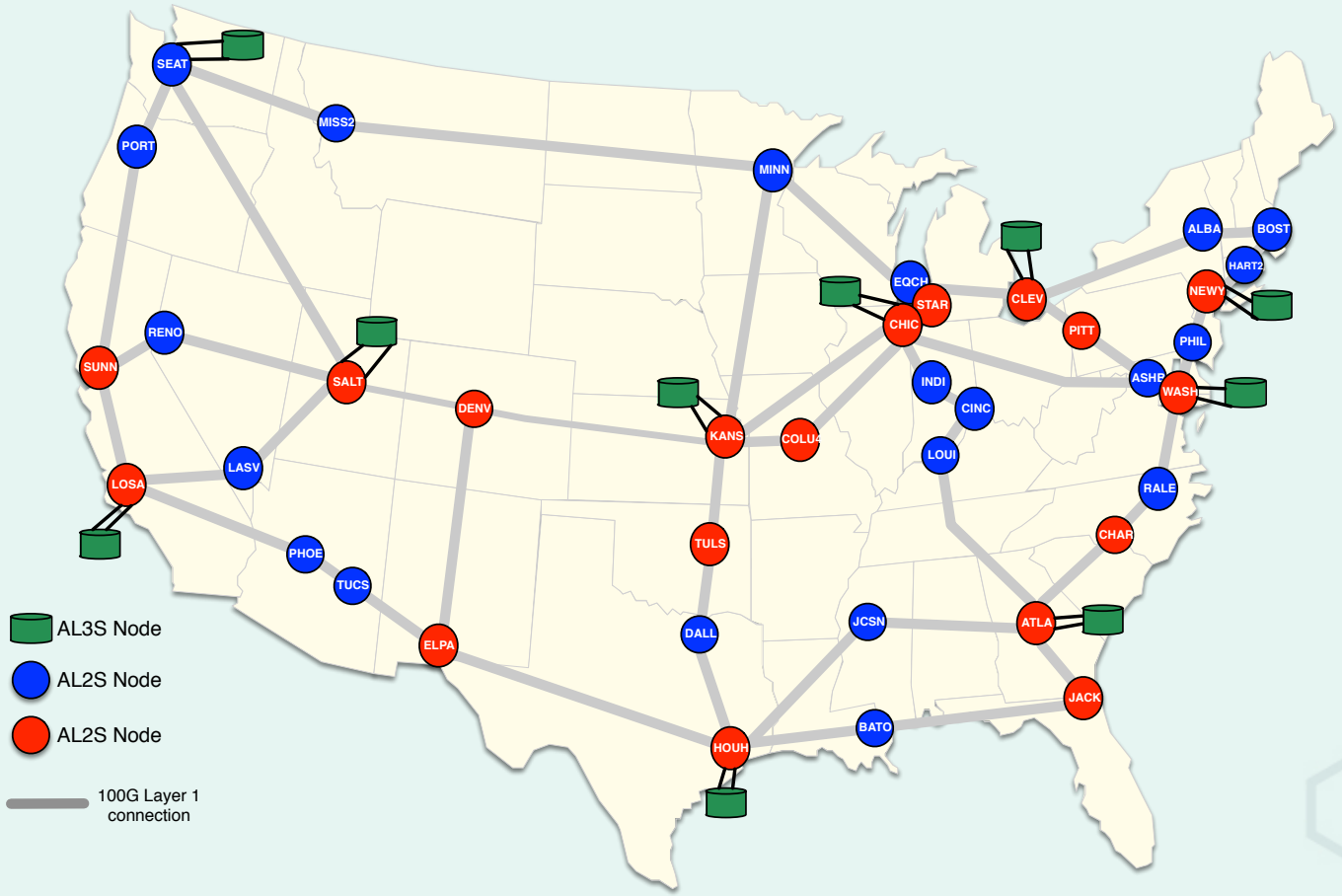



I2 Production Service
Service User
I2 Prototype Service
Internet2 Service Taxonomy
Implemented Using
Dependencies
LHCONE
NET+
External Provider Services
XSEDE
GENI
?
Layer 3 R&E IP and TR-CPS Services
I2-Run Service Specific Hardware
Connectors
GENI Learning Switch
ONOS
?
General Purpose VLAN Service - AL2S
SDN Controller
ESNET
NOAA
NVS (Network Virtualization Service)
Virtualized Ethernet Switching
Hypervisor
Ethernet Switches
Circuits and Wavelengths - AL1S
Fiber & Optical Transport

AL2S / AL3S Architecture and Capacity Today

- The initial capacity of AL2S was constructed as follows:
  - 1x100G connection between AL2S switches
  - 2x100G connection between AL2S switches and AL3S routers
- Routers are connected by a partial (soon to be full) mesh of VLANs.
- Pros of a full mesh:
  - Avoid unnecessary use of BW between switches and routers.
  - Segregate traffic to understand traffic patterns
- Cons of a full mesh:
  - Loss of transparency to network engineers …
  - … but fixable by SDN trace, clever DNS, or well-known routing

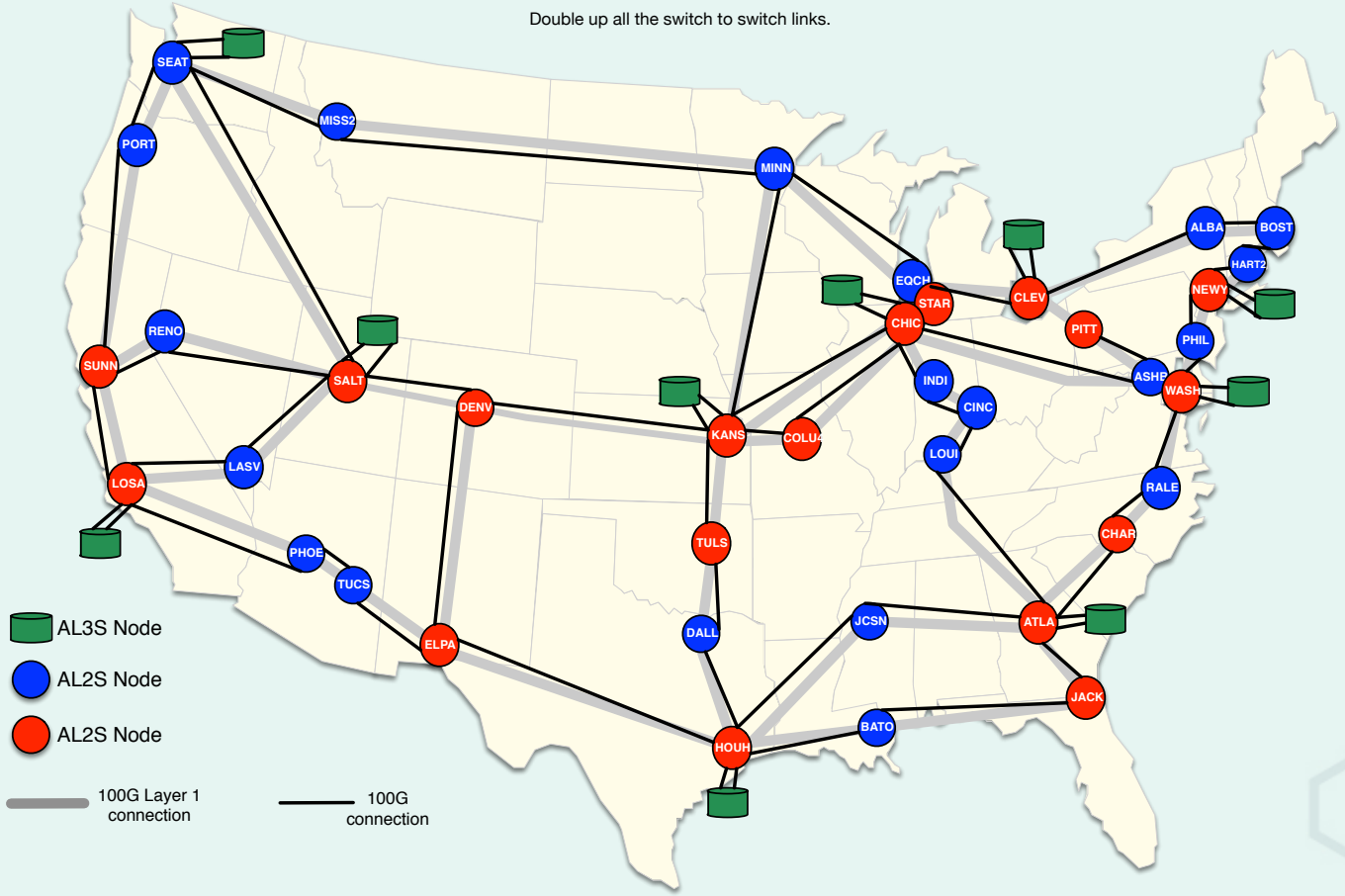
Initial Deployment



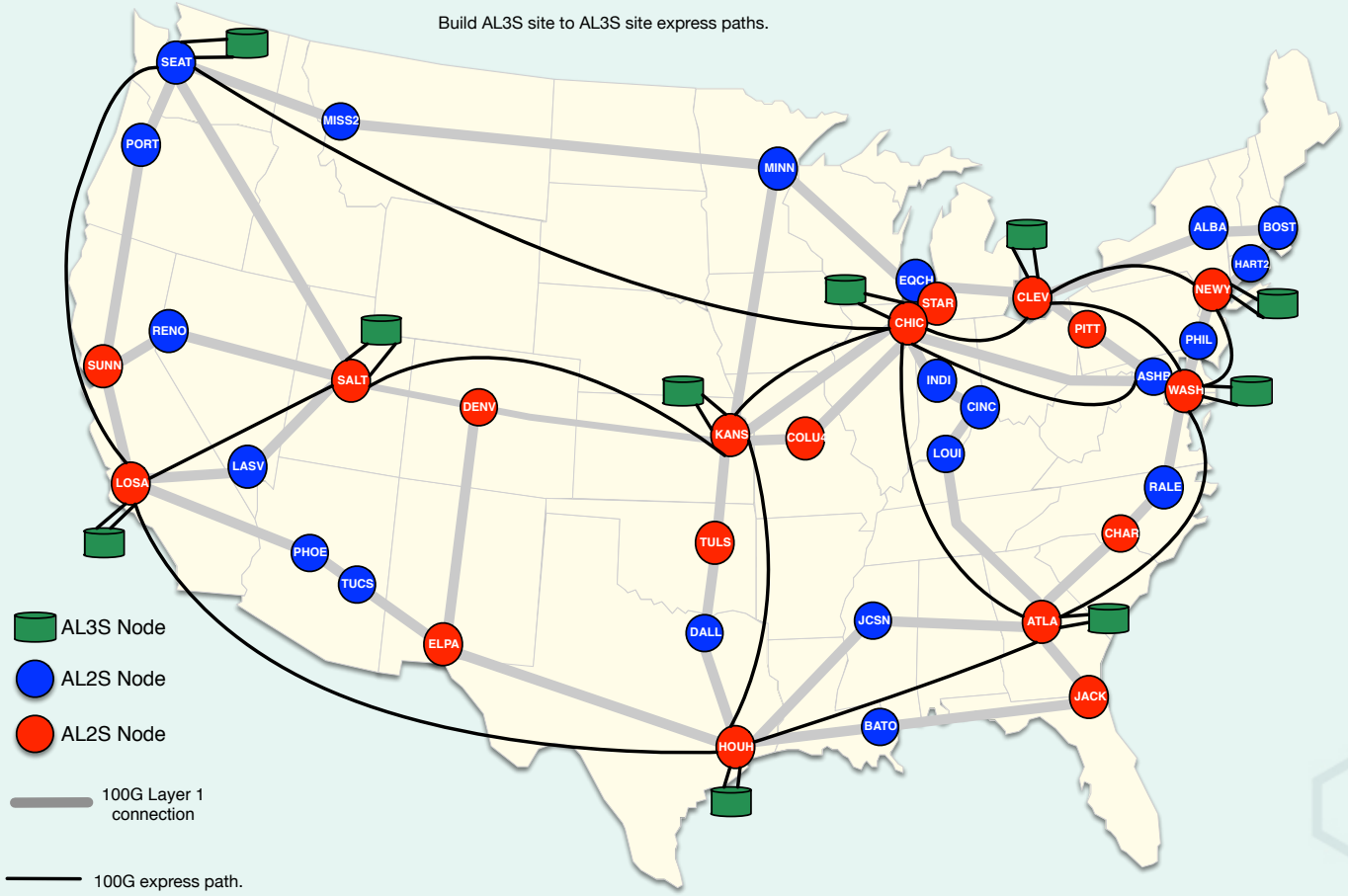
Upgrading Network Links

- When should we upgrade network links?
  - Upgrade threshold for AL2S is (currently) ~40% capacity on a specific AL2S link for 3 weeks in succession
  - Often see a drop-off in usage at the end of the school year and over the summer months
- How should we solve capacity ?
  - Should we upgrade specific hops between adjacent AL2S switches?
  - Should we build express lanes between switches adjacent to AL3S routers?
  - Should we build express lanes between switches adjacent to exchange points?
  - Should we build direct connections between AL3S routers?
  - Should we upgrade links between AL2S switches and AL3S routers?
- Are there (delaying) alternatives to upgrading network links?
  - Assuming a full mesh of VLANs between routers …
  - … automatically or semi-automatically migrate VLANs over different routes
  - Look for additional peerings for key Net+ services (e.g. AWS)
  - Work with connectors to optimize network usage

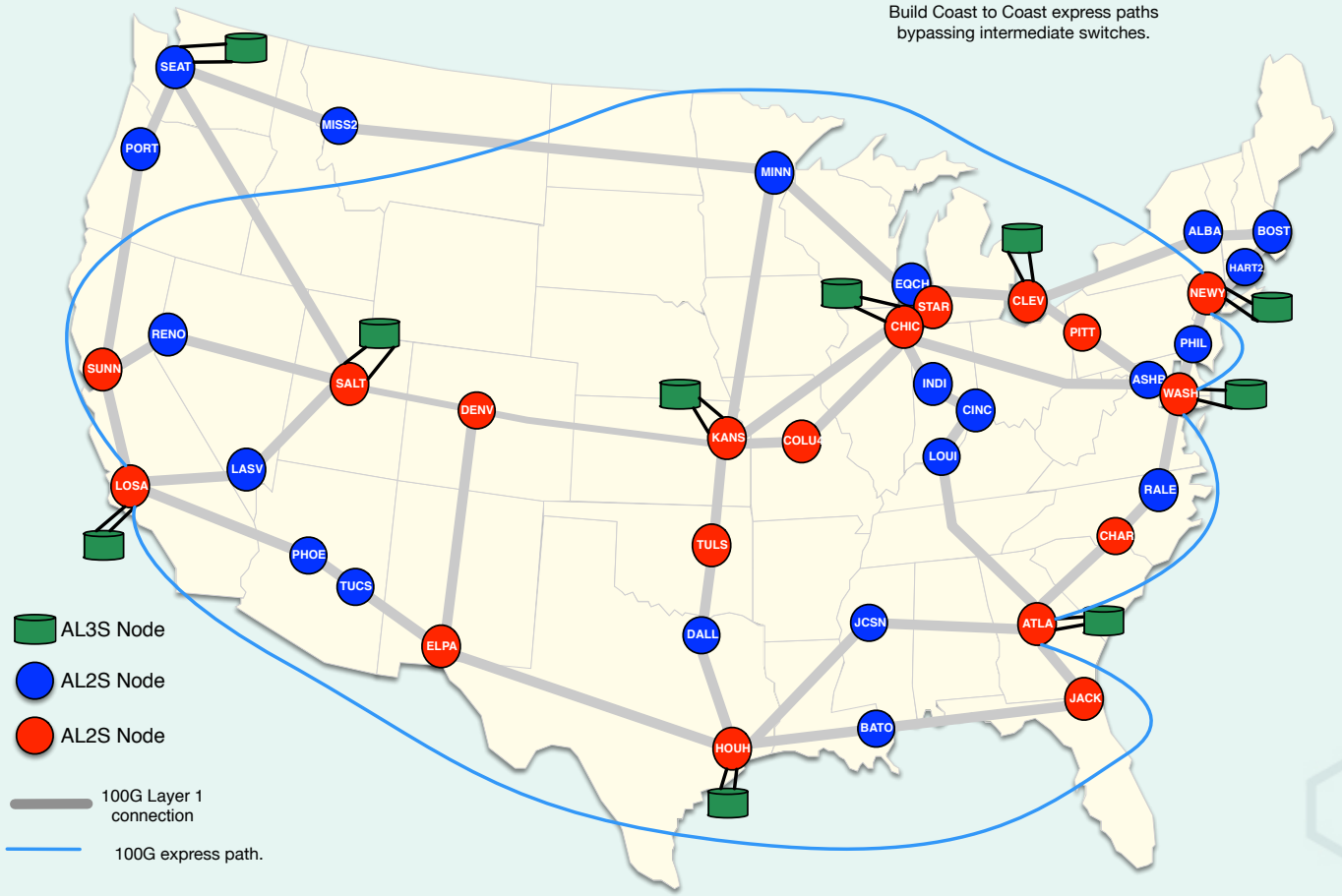
Switch-to-Switch Local Capacity Upgrades



Hop-to-Hop Express Capacity Upgrades



Coast-to-Coast Express Capacity Upgrades



Case Study: Raleigh to Washington DC

Circuit	6/26/15		6/19/15		6/12/15		6/5/15		5/29/15		5/22/15		5/15/15		5/8/15		5/1/15		4/24/15		4/17/15	
	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out
I2-ATLA-CHAR-100GE-07738						36%		40%		37%		38%					32%	32%	33%	33%	32%	31%
I2-CHIC-INDI-100GE-50993								32%		31%												
I2-ATLA-CINC-100GE-07740							32%		31%													
I2-CINC-INDI-100GE-51692							32%		30%													
I2-ASHB-WASH-100GE-09106										32%		35%				34%		51%		35%		48%
I2-ASHB-WASH-100GE-11823						34%								43%		43%		36%		41%		44%
I2-HART2-NEWY32AOA-100GE-12947									32%		34%						38%				33%	
I2-ATLA-ATLA-100GE-09159										31%												
I2-CHIC-CHIC-100GE-07769										31%	31%										44%	38%
I2-CHIC-CHIC-100GE-09166													33%		33%	32%	34%	36%	38%	32%		
I2-CHIC-STAR-100GE-07743						34%				30%		35%		41%		41%	33%	40%	32%	39%	41%	39%
I2-RALE-WASH-100GE-08888			32%		40%								45%		44%		43%		43%		44%	
I2-CHAR-RALE-100GE-10633					33%								34%		34%		30%		31%		30%	
I2-CHIC-KANS-100GE-07745																35%		36%		37%		38%
I2-ASHB-PITT-100GE-07737																31%		31%		39%		46%
I2-PHIL-WASH-100GE-10867																	34%				33%	
I2-WASH-WASH-100GE-07774																			30%			
I2-WASH-WASH-100GE-48338																	31%					
I2-CLEV-PITT-100GE-10334																	30%		36%		44%	
I2-LOSA-PHOE-100GE-09190																				30%		31%

Case Study: Raleigh to Washington DC

Circuit	6/26/15		6/19/15		6/12/15		6/5/15		5/29/15		5/2								4/24/15		4/17/15	
	In	Out	In	Out	In	Out	In	Out	In	Out	In								In	Out	In	Out
I2-ATLA-CHAR-100GE-07738						36%		40%		37%									33%	33%	32%	31%
I2-CHIC-INDI-100GE-50993								32%		31%												
I2-ATLA-CINC-100GE-07740							32%		31%													
I2-CINC-INDI-100GE-51692							32%		30%													
I2-ASHB-WASH-100GE-09106										32%										35%		48%
I2-ASHB-WASH-100GE-11823						34%														41%		44%
I2-HART2-NEWY32AOA-100GE-12947									32%		34%										33%	
I2-ATLA-ATLA-100GE-09159										31%												
I2-CHIC-CHIC-100GE-07769										31%	31%										44%	38%
I2-CHIC-CHIC-100GE-09166													33		33%	32%	34%	36%	38%	32%		
I2-CHIC-STAR-100GE-07743						34%				30%		35%		41%		41%	33%	40%	32%	39%	41%	39%
I2-RALE-WASH-100GE-08888			32%		40%								45%		44%		43%		43%		44%	
I2-CHAR-RALE-100GE-10633					33%								34%		34%		30%		31%		30%	
I2-CHIC-KANS-100GE-07745																35%		36%		37%		38%
I2-ASHB-PITT-100GE-07737																31%		31%		39%		46%
I2-PHIL-WASH-100GE-10867																	34%				33%	
I2-WASH-WASH-100GE-07774																			30%			
I2-WASH-WASH-100GE-48338																	31%					
I2-CLEV-PITT-100GE-10334																	30%		36%		44%	
I2-LOSA-PHOE-100GE-09190																				30%		31%

Raleigh to DC is heating up

INTERNET 2

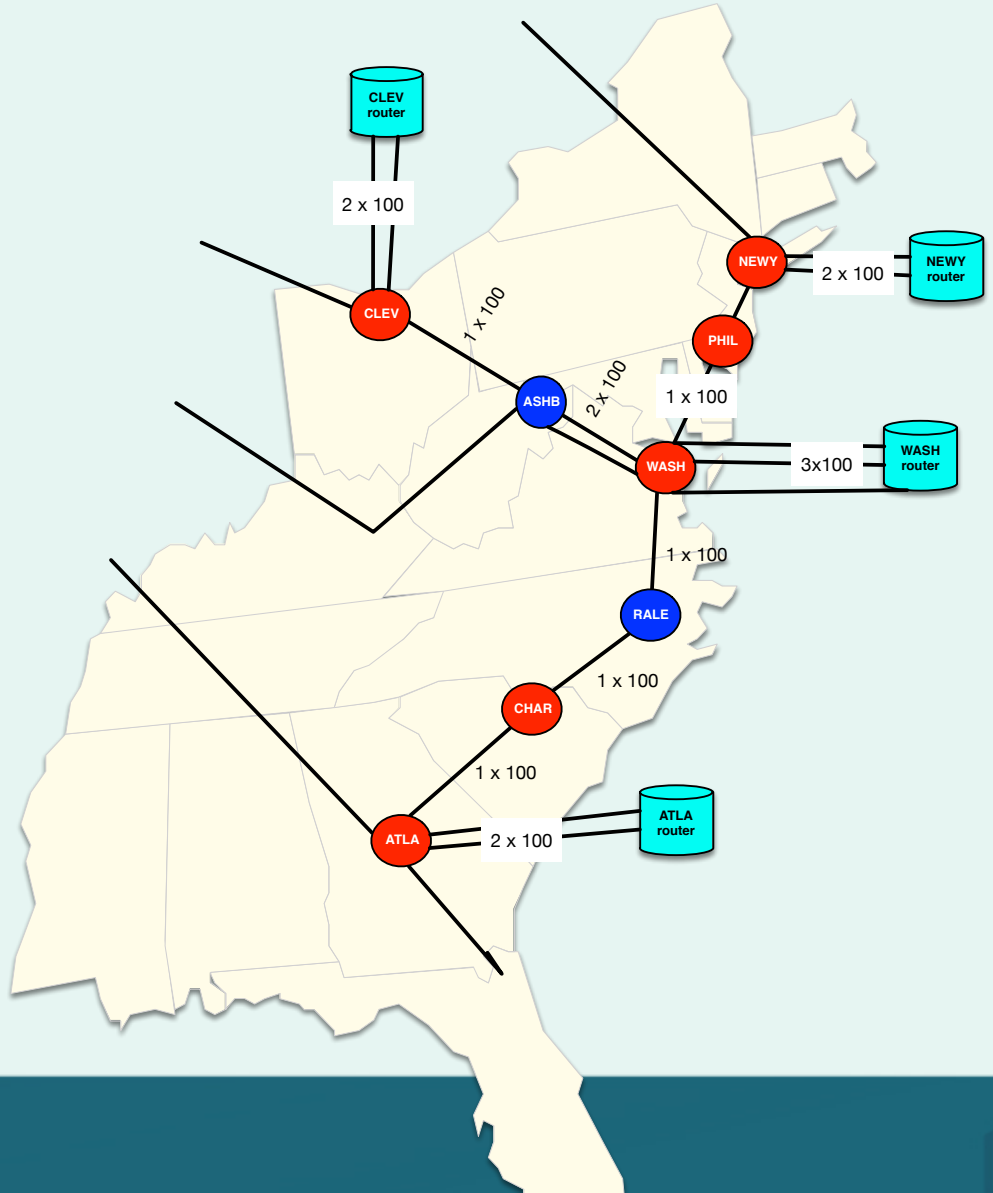
Case Study: Raleigh to Washington DC

Charlotte to Atlanta is close behind

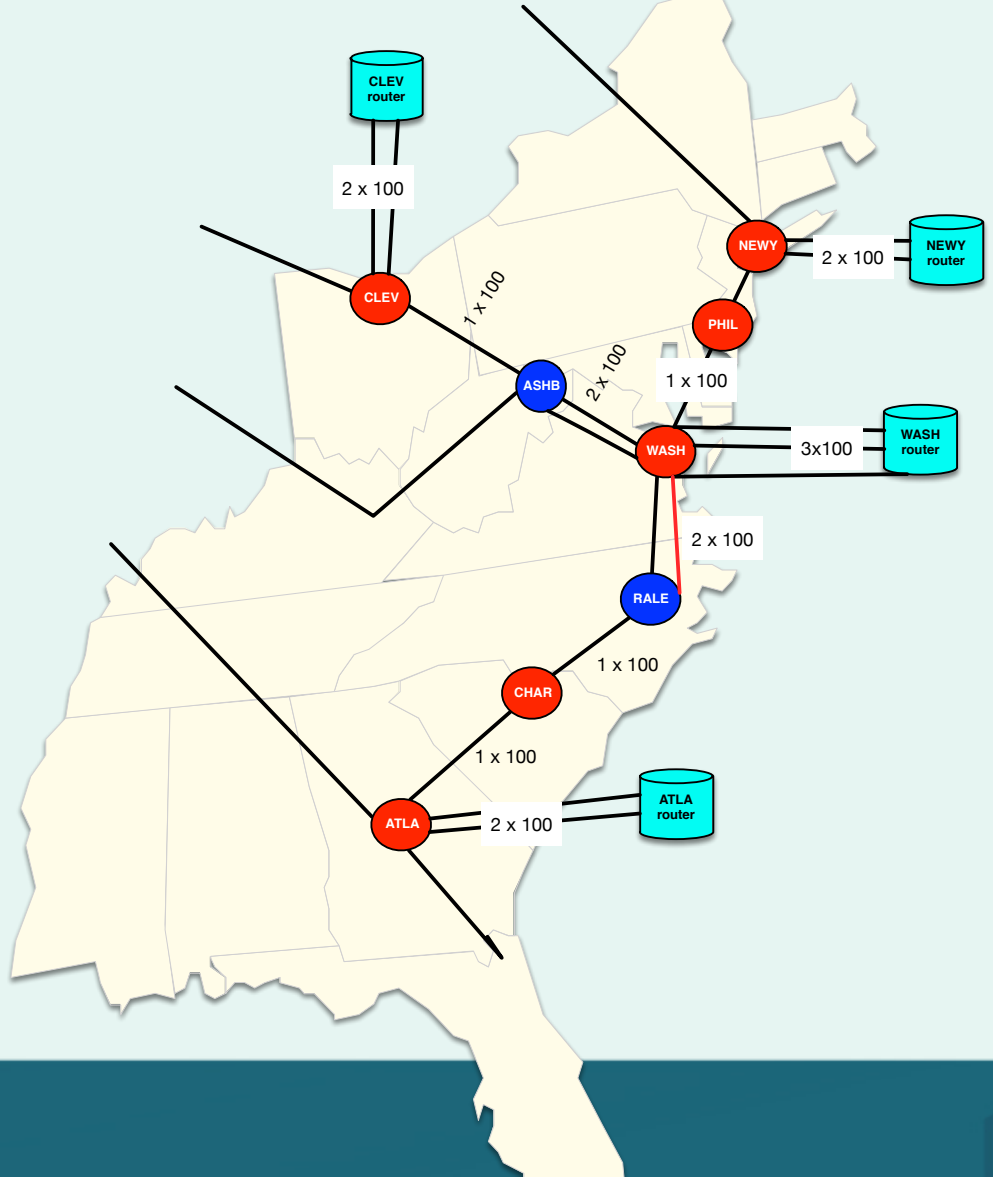
Circuit	6/26/15		6/19/15		6/12/15		6/5/15		[illegible]		5/22/15		5/15/15		5/8/15		5/1/15		4/24/15		4/17/15	
	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out
I2-ATLA-CHAR-100GE-07738						36%		40%		37%		38%					32%	32%	33%	33%	32%	31%
I2-CHIC-INDI-100GE-50993								32%		31%												
I2-ATLA-CINC-100GE-07740							32%		31%													
I2-CINC-INDI-100GE-51692							32%		30%													
I2-ASHB-WASH-100GE-09106										32%		35%				34%		51%		35%		48%
I2-ASHB-WASH-100GE-11823						34%								43%		43%		36%		41%		44%
I2-HART2-NEWY32AOA-100GE-12947									32%		34%						38%				33%	
I2-ATLA-ATLA-100GE-09159										31%												
I2-CHIC-CHIC-100GE-07769										31%	31%										44%	38%
I2-CHIC-CHIC-100GE-09166													33%		33%	32%	34%	36%	38%	32%		
I2-CHIC-STAR-100GE-07743						34%				30%		35%		41%		41%	33%	40%	32%	39%	41%	39%
I2-RALE-WASH-100GE-08888			32%		40%								45%		44%		43%		43%		44%	
I2-CHAR-RALE-100GE-10633					33%								34%		34%		30%		31%		30%	
I2-CHIC-KANS-100GE-07745																35%		36%		37%		38%
I2-ASHB-PITT-100GE-07737																31%		31%		39%		46%
I2-PHIL-WASH-100GE-10867																	34%				33%	
I2-WASH-WASH-100GE-07774																			30%			
I2-WASH-WASH-100GE-48338																	31%					
I2-CLEV-PITT-100GE-10334																	30%		36%		44%	
I2-LOSA-PHOE-100GE-09190																				30%		31%

INTERNET 2

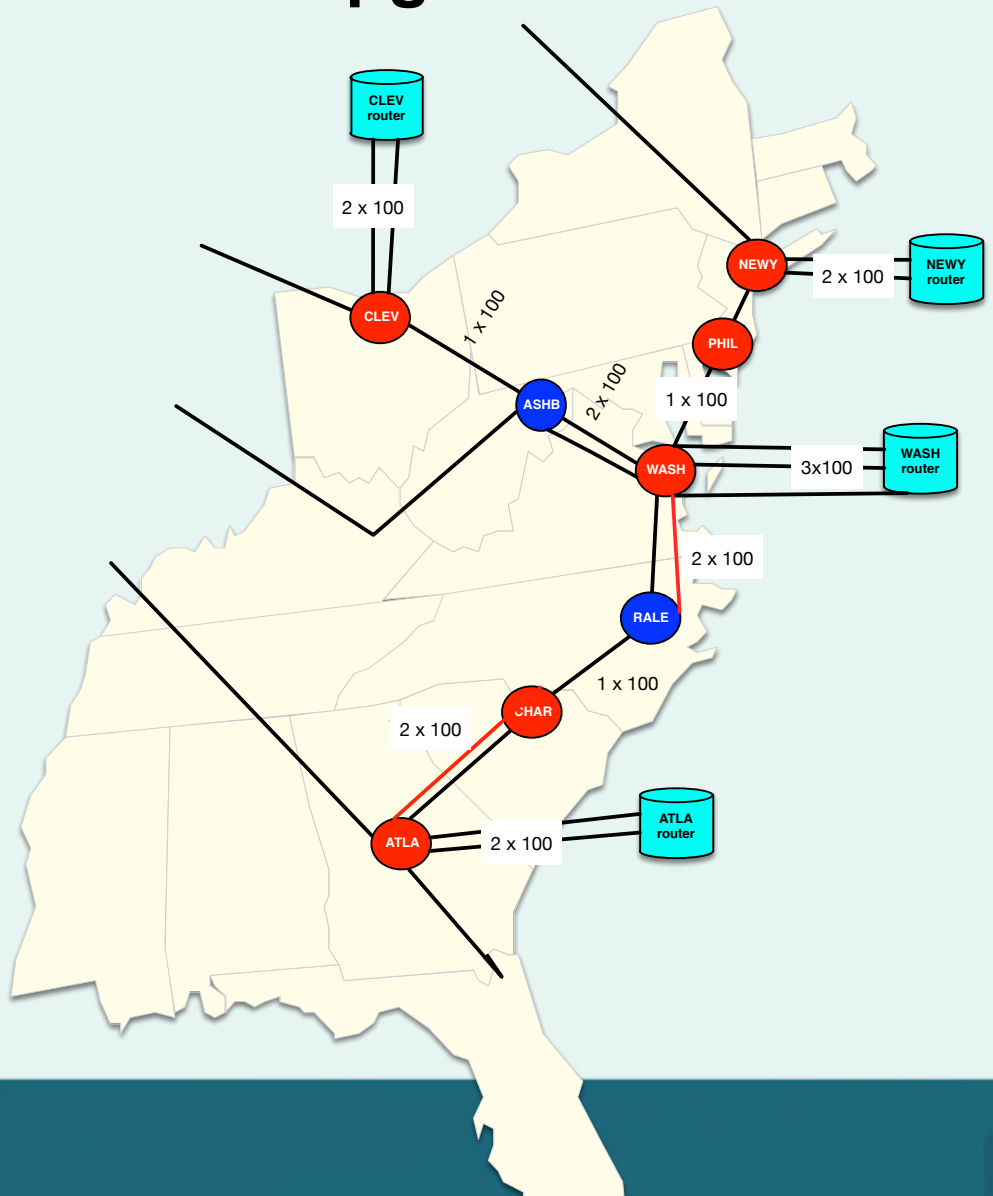
Southeastern AL2S – Where should we upgrade?



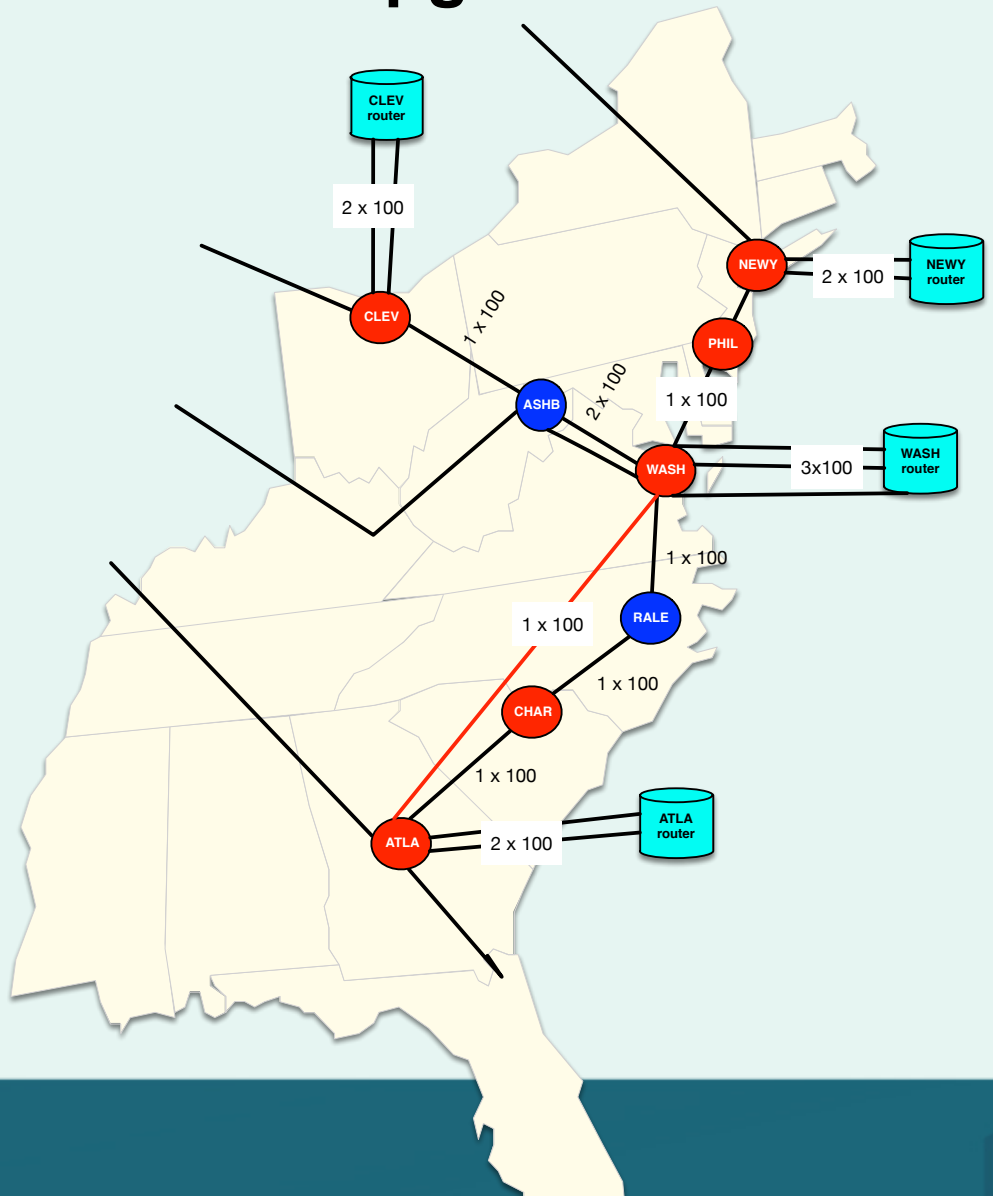
1A) Local AL2S Upgrade Raleigh <-> DC



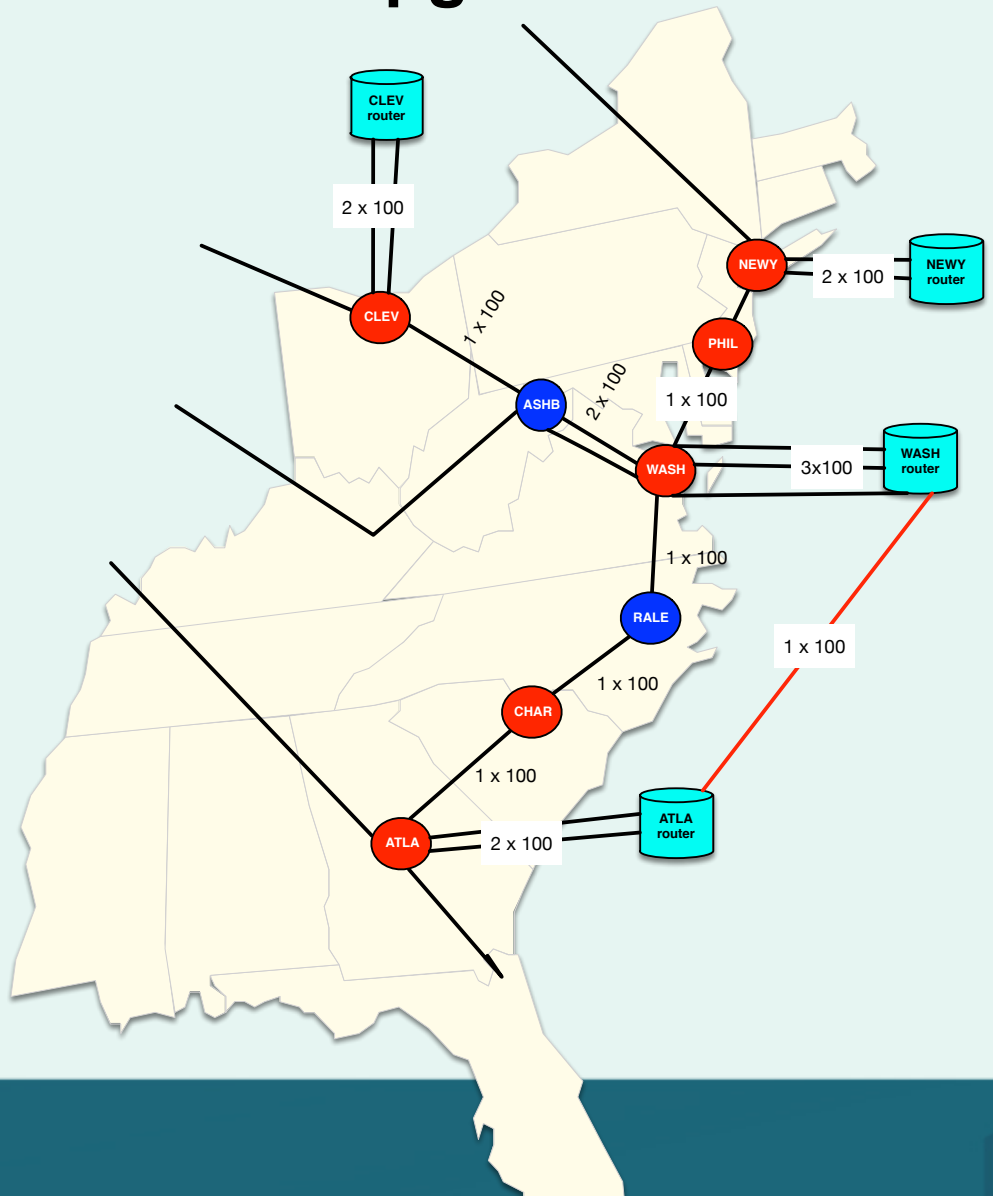
1B) Local AL2S Upgrade Atlanta <-> Charlotte



2) Express AL2S Upgrade Atlanta <-> DC



3) Express AL3S Upgrade Atlanta <-> DC



Evaluation

- 1) Local AL2S Upgrade Raleigh <-> DC
  - Pro: Adds capacity on hot L2 link minimizing immediate expense.
  - Con: If we eventually want to upgrade capacity of Atlanta router to DC router, link by link
  - Cost: 2 interfaces plus optical cost (short-term); 6 interfaces plus optical cost (long-term) (10 to reach the routers)
- 2) Express AL2S Upgrade Atlanta <-> DC
  - Pro: Adds capacity on hot L2 segment (multiple links) reducing long-term cost while preserving maximum flexibility
  - Cost: 2 interfaces plus optical cost (6 interfaces to reach router)
- 3) Express AL3S Upgrade Atlanta <-> DC
  - Pro: Adds capacity on hot L3 segment, minimizing long-term cost
  - Con: Only assists L3 traffic between Atlanta and DC(not Atlanta and NYC, for example, or DC and Houston)
  - Cost: 2 interfaces plus optical cost plus might force upgrade of link between Atlanta router and switch or DC router and switch

Recommendation: AL2S Atlanta <-> DC Upgrade

- Too many problems with AL3S upgrade
  - Loss of flexibility (only addresses subset of L3 traffic)
  - May not solve problem (e.g. if bulk of traffic is LA <-> NYC)
  - Insufficient data (e.g. we don't know if CENIC <-> WIX is the real problem)
- Raleigh <-> DC upgrade is too costly in the long run
- Express route between AL2S switch in Atlanta and AL2S switch in DC is preferred
  - Supports *all* AL3S traffic passing between Atlanta and DC
  - Minimizes interface costs
  - Creates no additional load on switch <-> router links
  - Preserves flexibility (e.g. ONOS or VLAN migration)

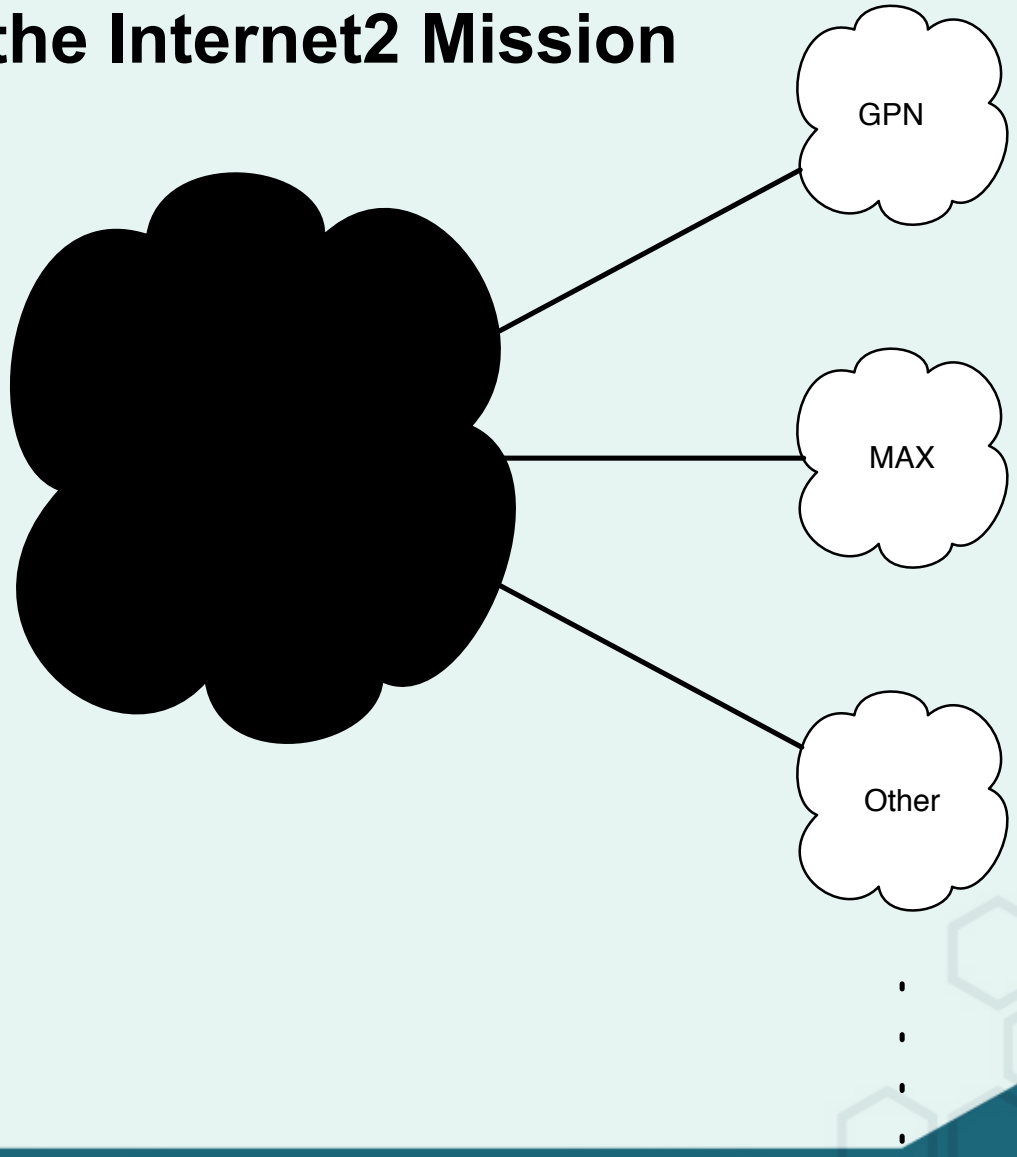
Case Study: Traffic into Ashburn (Home of AWS)

Ashburne to DC is heating up

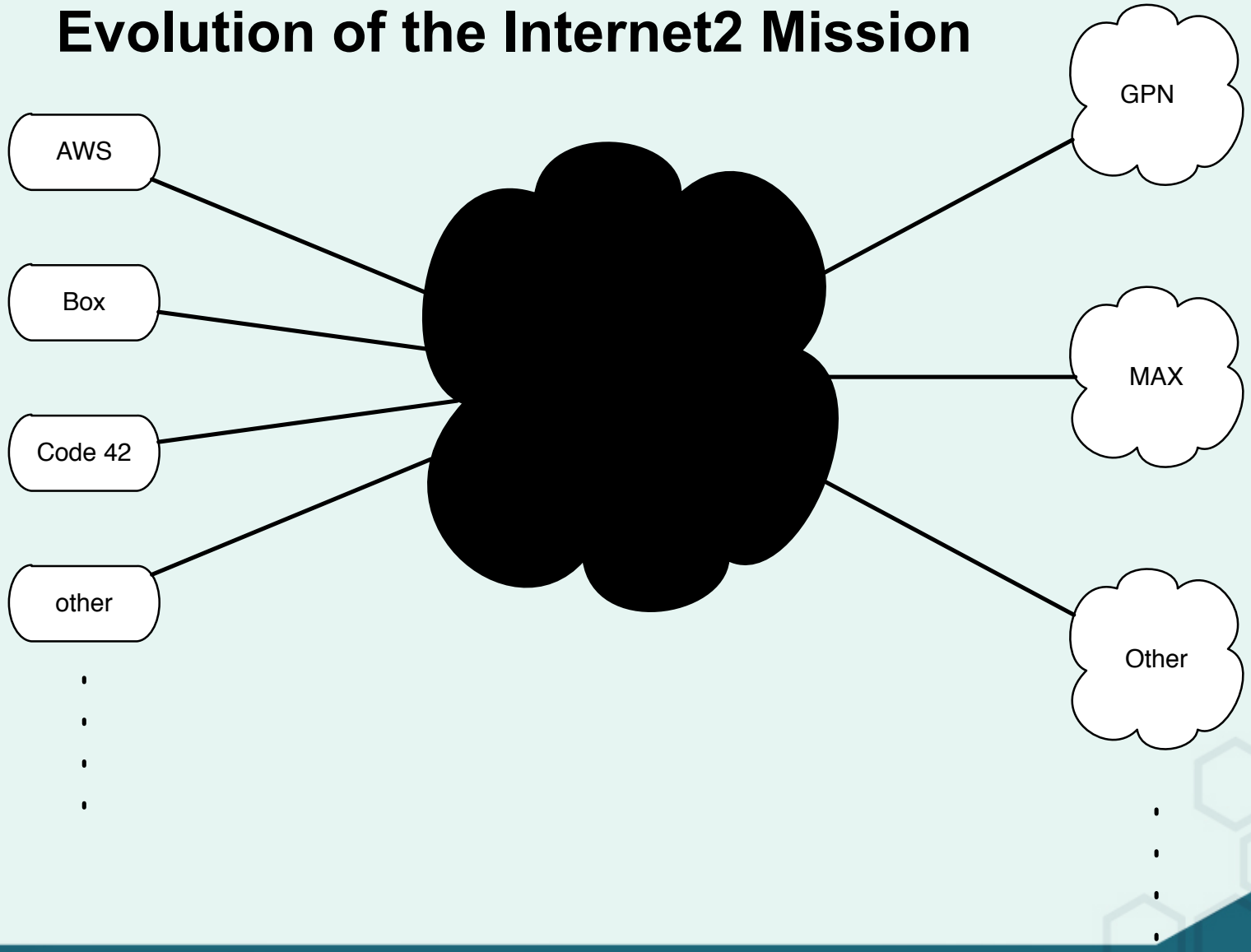
Circuit	6/26/15		6/19/15		6/12/15		6/5/15		5/29/15		5/22/15								4/24/15		4/17/15	
	In	Out	In	Out	In	Out	In	Out	In	Out	In	Out							In	Out	In	Out
I2-ATLA-CHAR-100GE-07738						36%		40%		37%		38%							33%	33%	32%	31%
I2-CHIC-INDI-100GE-50993								32%		31%												
I2-ATLA-CINC-100GE-07740							32%		31%													
I2-CINC-INDI-100GE-51692							32%		30%													
I2-ASHB-WASH-100GE-09106										32%		35%				34%		51%		35%		48%
I2-ASHB-WASH-100GE-11823						34%								43%		43%		36%		41%		44%
I2-HART2-NEWY32AOA-100GE-12947									32%		34%						38%				33%	
I2-ATLA-ATLA-100GE-09159										31%												
I2-CHIC-CHIC-100GE-07769										31%	31%										44%	38%
I2-CHIC-CHIC-100GE-09166													33%		33%	32%	34%	36%	38%	32%		
I2-CHIC-STAR-100GE-07743						34%				30%		35%		41%		41%	33%	40%	32%	39%	41%	39%
I2-RALE-WASH-100GE-08888			32%		40%								45%		44%		43%		43%		44%	
I2-CHAR-RALE-100GE-10633					33%								34%		34%		30%		31%		30%	
I2-CHIC-KANS-100GE-07745																35%		36%		37%		38%
I2-ASHB-PITT-100GE-07737																31%		31%		39%		46%
I2-PHIL-WASH-100GE-10867																	34%				33%	
I2-WASH-WASH-100GE-07774																			30%			
I2-WASH-WASH-100GE-48338																	31%					
I2-CLEV-PITT-100GE-10334																	30%		36%		44%	
I2-LOSA-PHOE-100GE-09190																				30%		31%

INTERNET 2

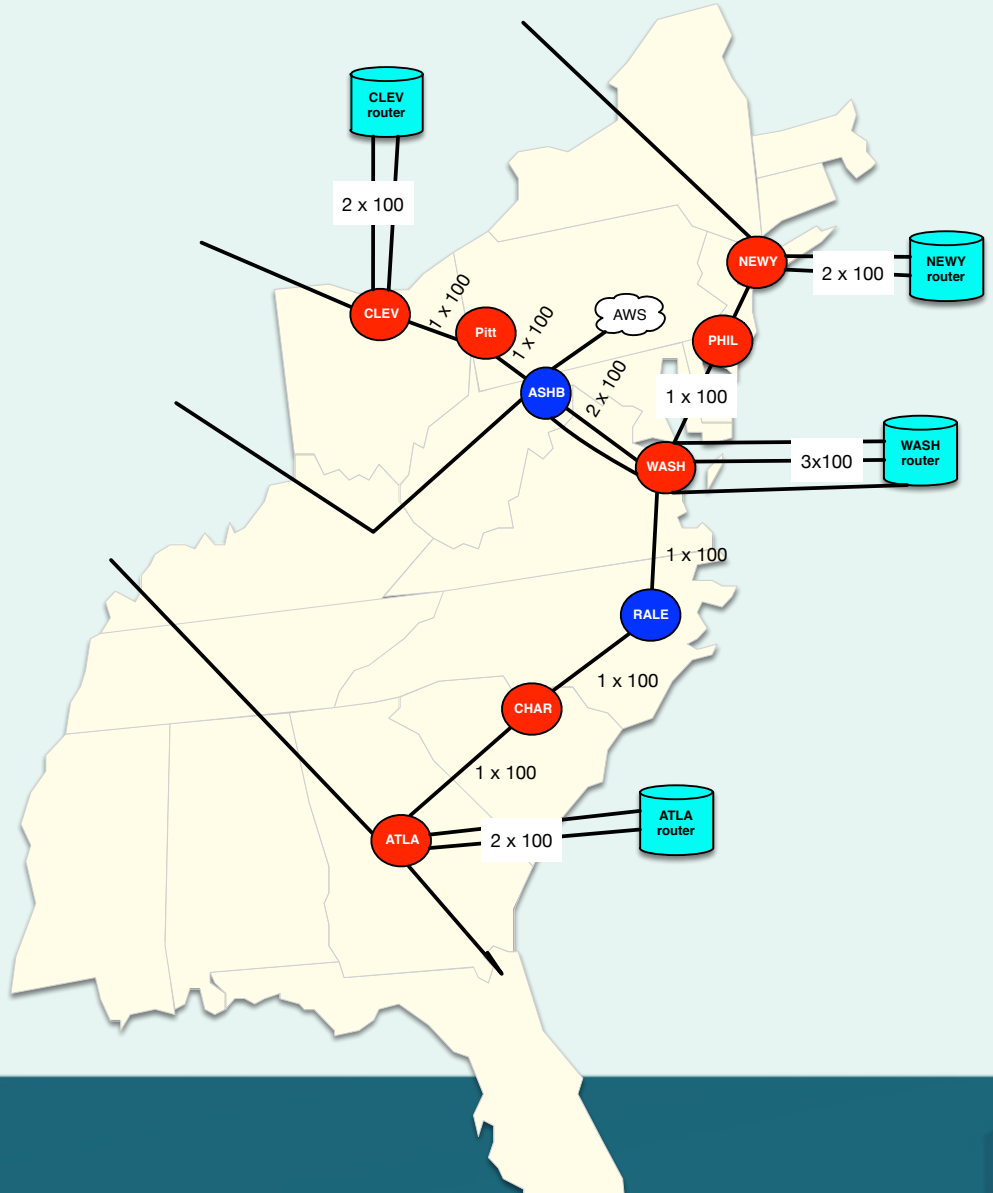
Evolution of the Internet2 Mission



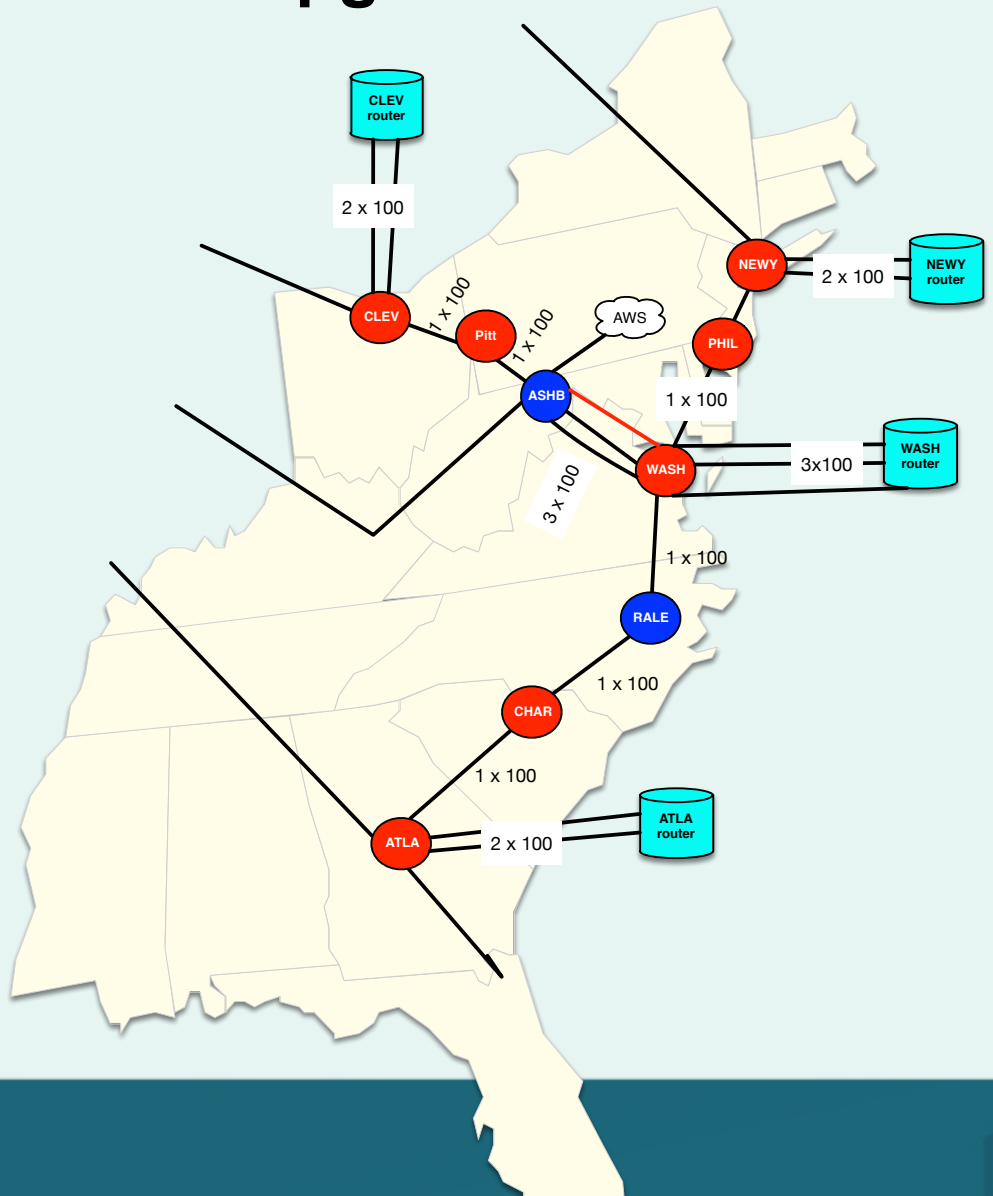
Evolution of the Internet2 Mission



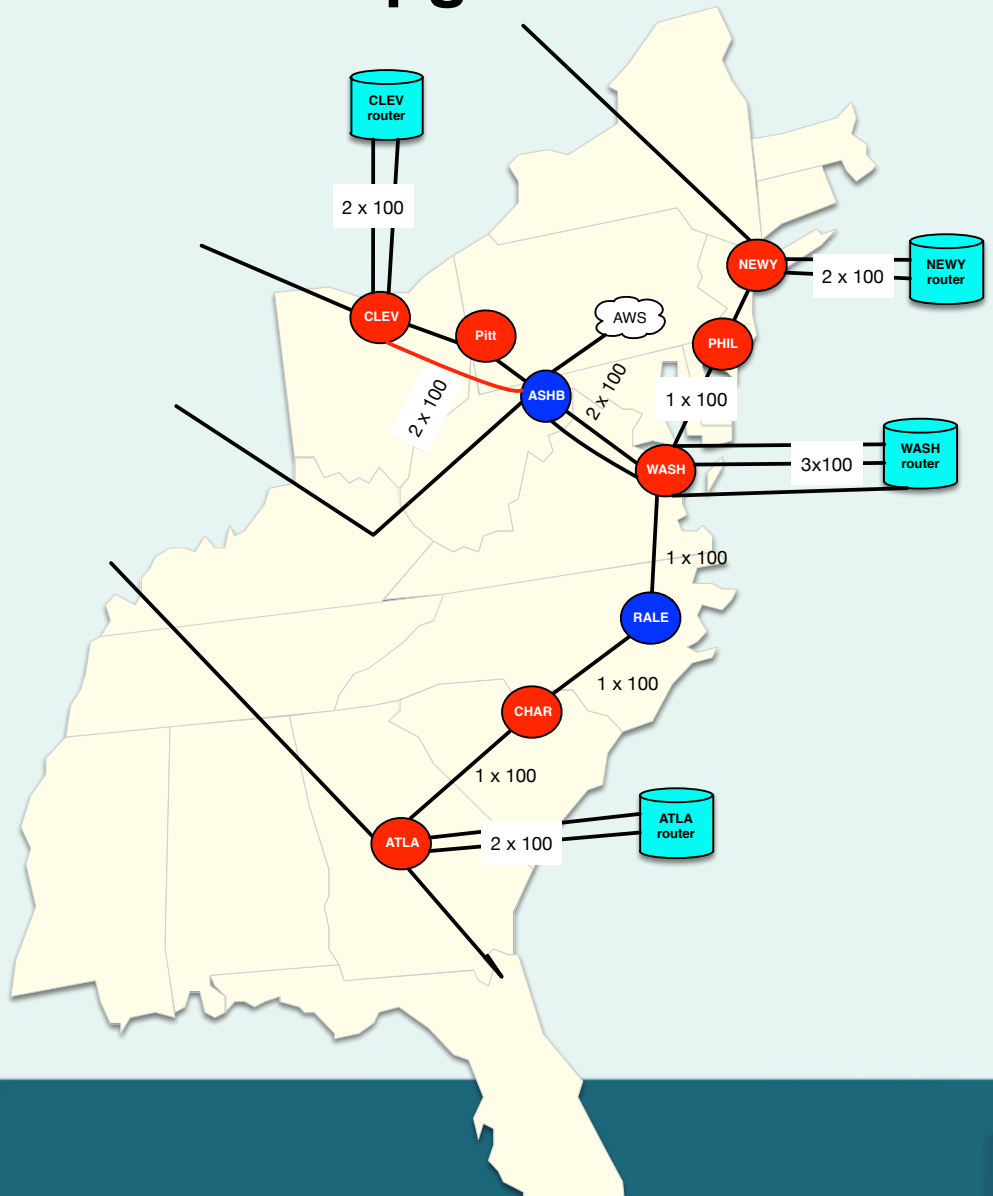
Amazon Web Services – How should we address?



1) Local AL2S Upgrade Ashburn <-> DC



2) Express AL2S Upgrade Ashburn <-> Cleveland



3) No upgrade – Rebalance Links

- Look for links below 40% capacity that can serve as alternate routes to links at or above 40% capacity.
- Automatically or semi-automatically move VLANs from one path to another.
- Potentially incurs the “too much transparency” problem.
  - See follow-on slide.

4) No upgrade – Additional AWS peering point(s)

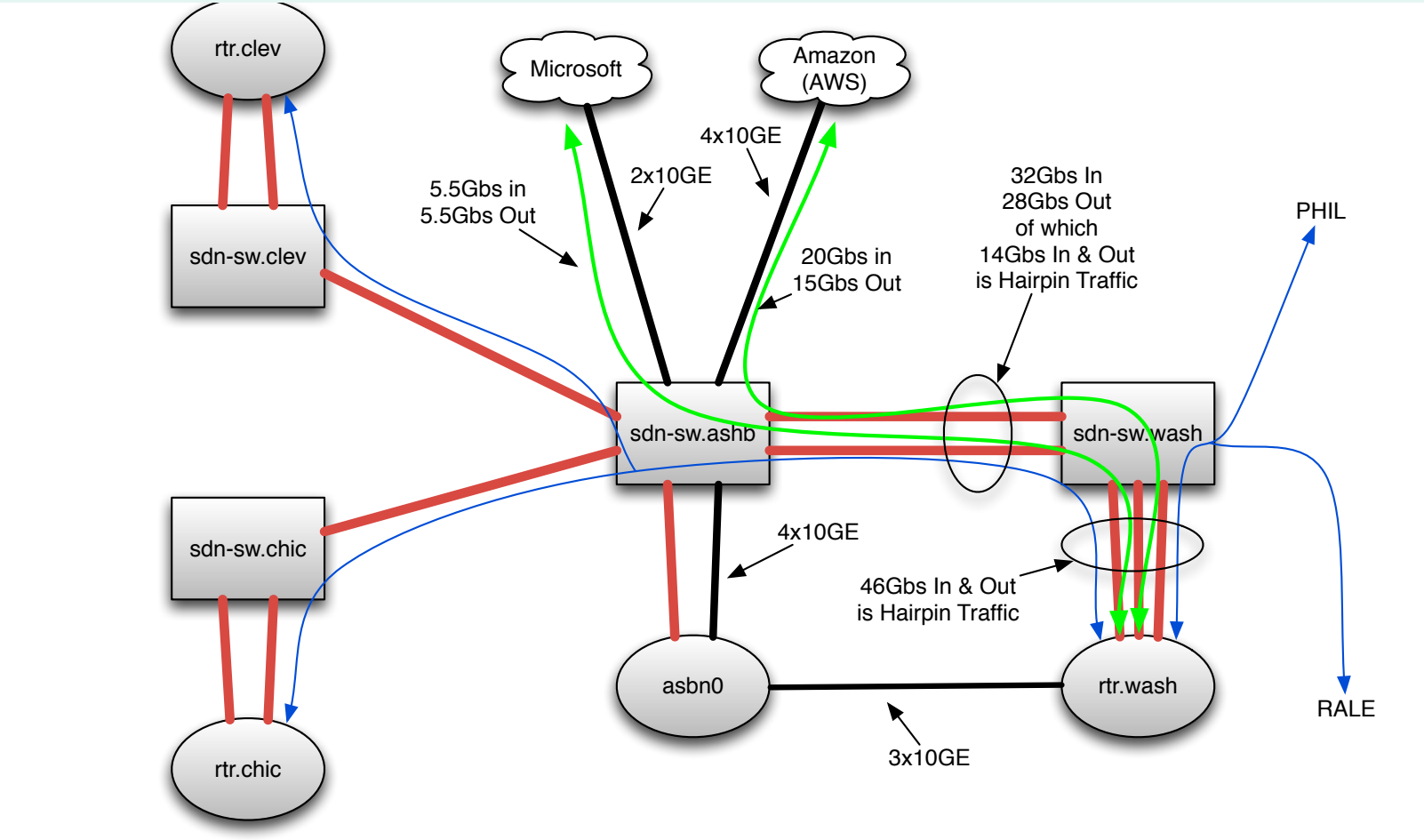
- Add one (or more) AWS peering point(s) (e.g. Atlanta, Chicago, or NYC)
- Potentially drain traffic from west coast to Ashburn.
- Additional resiliency benefit.
- Needs agreement from Amazon.

Recommendation: Wait til September, 2015

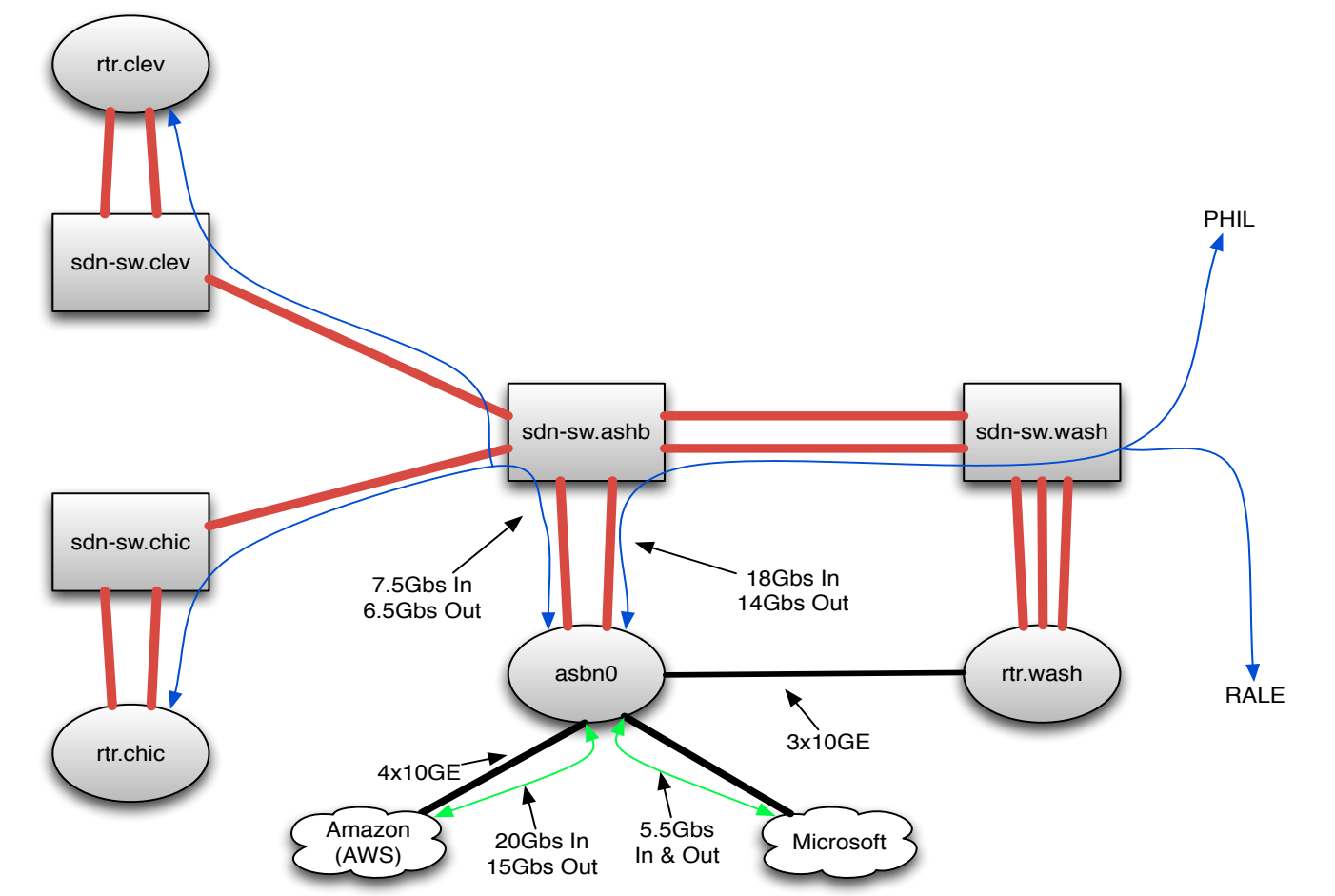
- Need more data
  - Implement full mesh of VLANs between routers (In progress)
  - Improve data quality by differentiating backup paths in software (Complete)
- Explore additional peering points with AWS (In progress)
- Revisit issue in early September

- Long Term Idea: Through software, implement or recommend implementation of alternate VLAN paths through network to reduce tendency of links to overheat

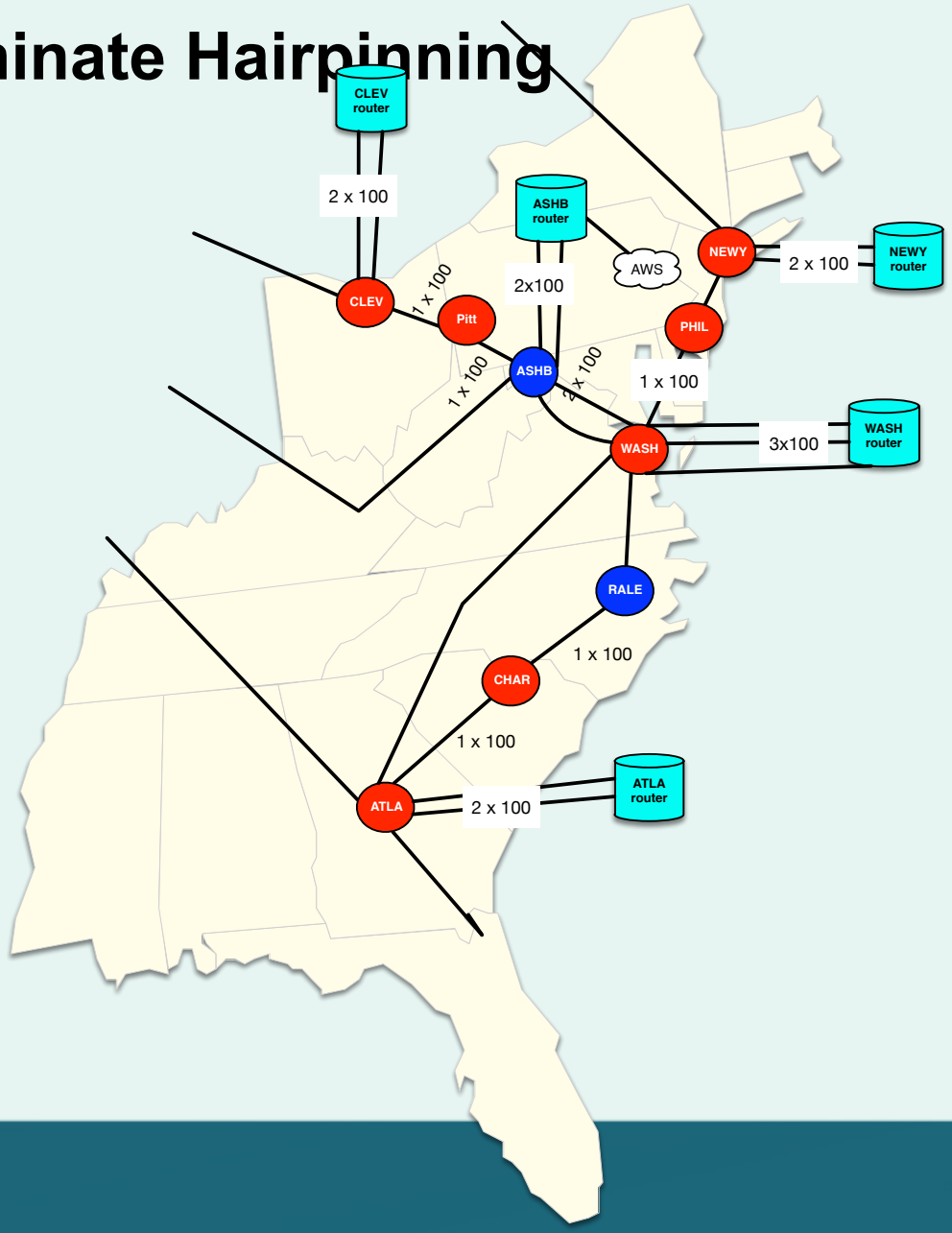
September Update



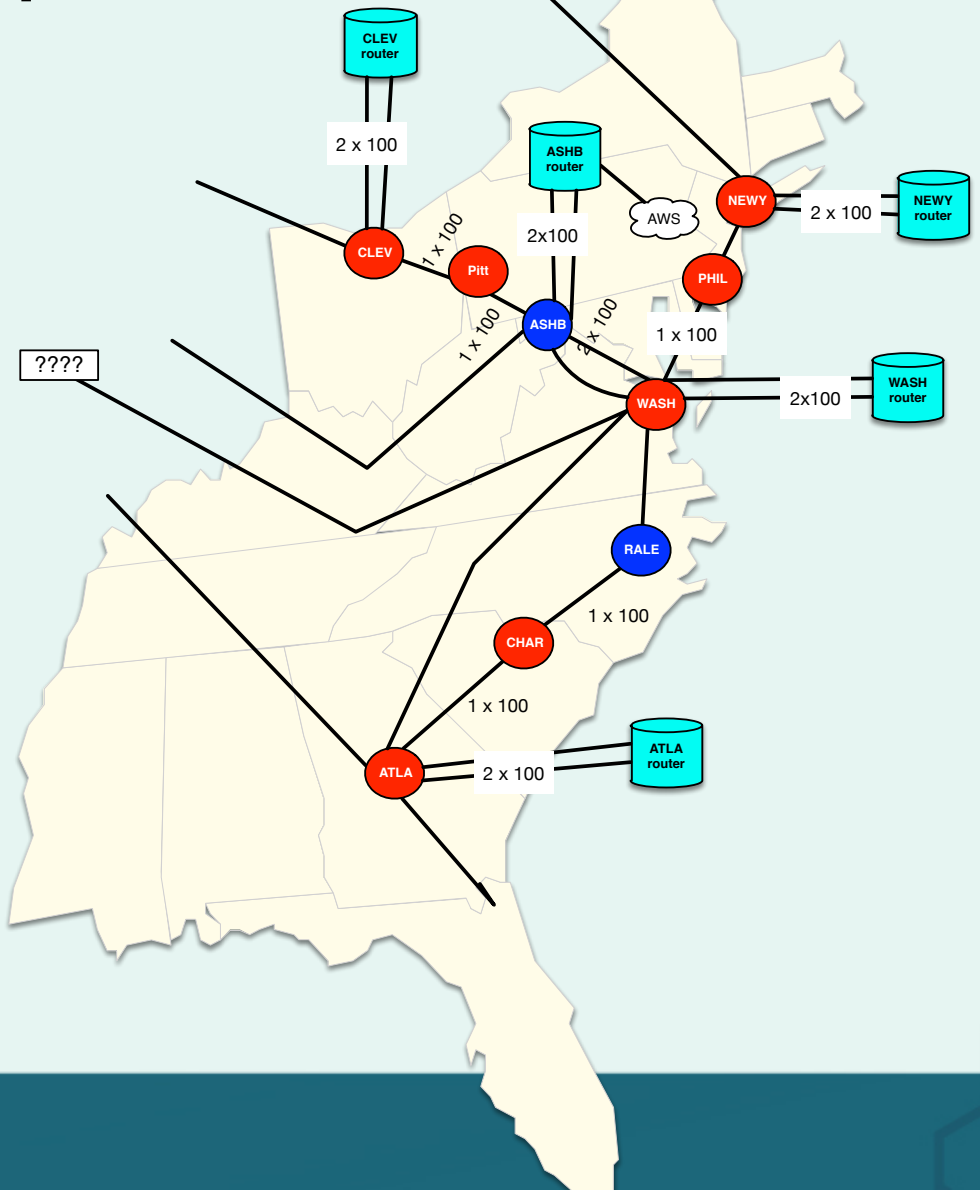
September Idea



5) Eliminate Hairpinning



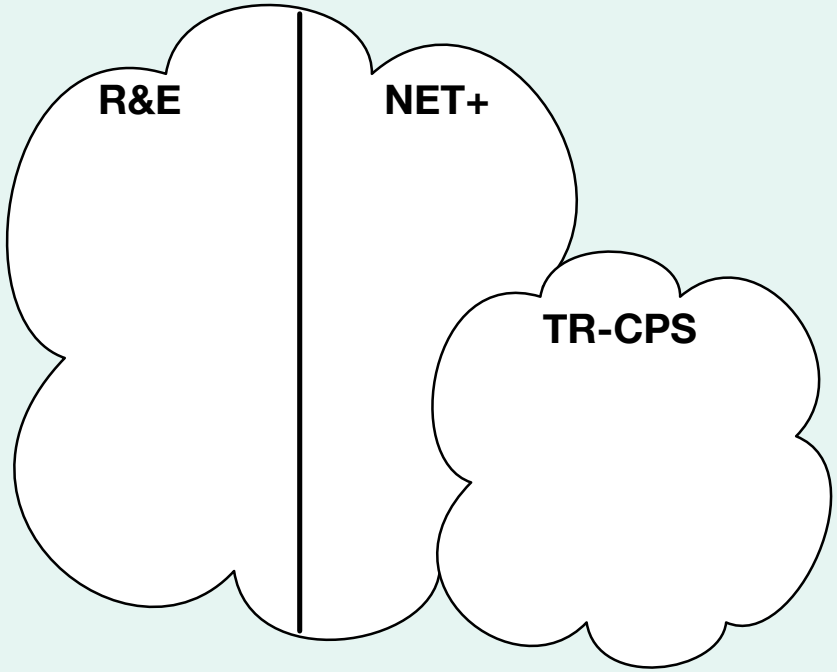
6) Free Up Port for Future Augments



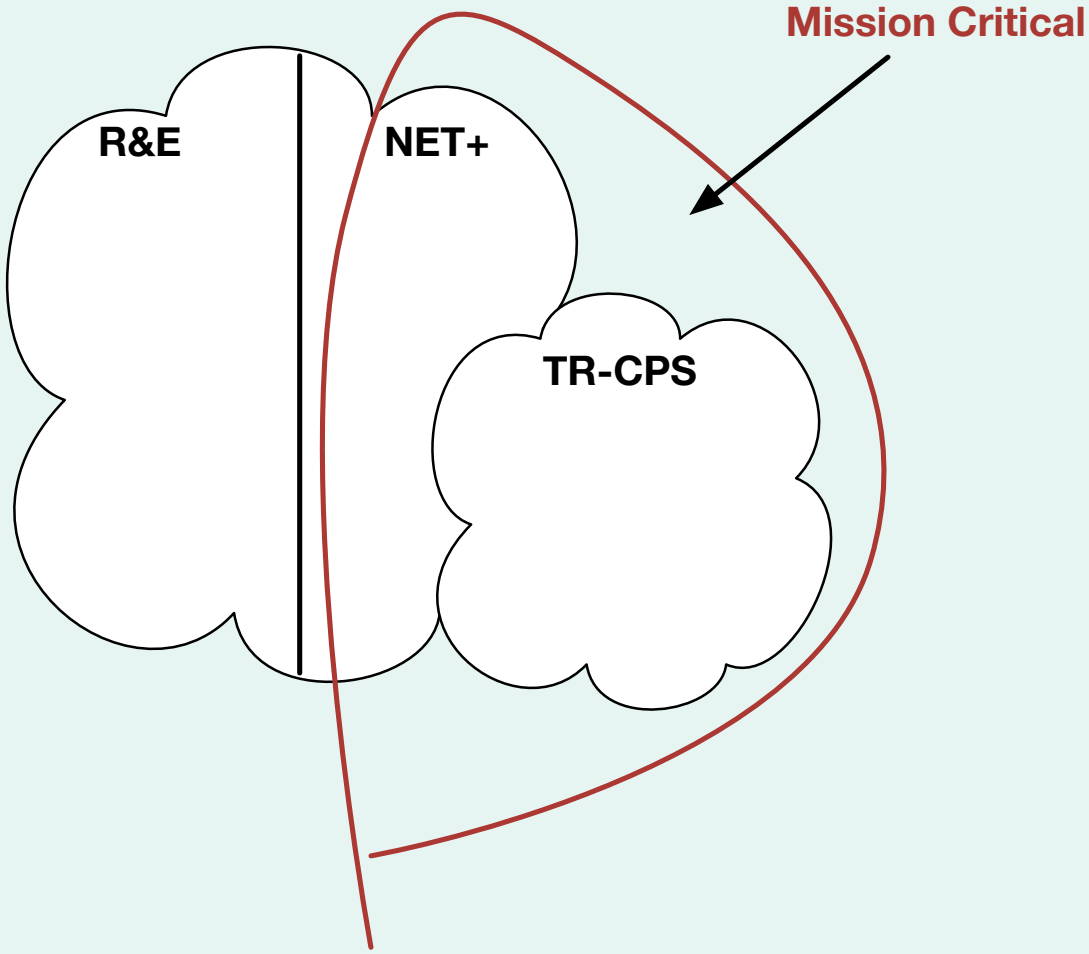
Challenge: Too Much Transparency?

- Internet2 AL2S is operated in an open and transparent fashion.
  - Users of AL2S must pick their end points
  - Users can, if they want, pick their exact path (e.g. southern tier, through the middle, or northern tier)
  - Even altruistic users do not have data guide their decision
- Side Effect: What happens when the lowest latency path (through the middle) overheats? Especially if we have software rebalancing the VLANs between routers to prevent any one link form overheating?
- Option 1: Augment capacity immediately?
  - Raises total cost to operate Internet2 Network
- Option 2: Reduce Transparency?
  - Users only get to pick their end points, not their path
- Option 3: Reduce user control?
  - Allow Internet2 to move any VLAN to alternate path if it helps keep us below capacity thresholds
- Option 4: Move AL3S VLANs to less used paths?
  - Effectively allows AL2S users priority over AL3S users
- Option 5: Move AL2S VLANs to less used paths?
  - Probably fine for data intensive science traffic (changes latency)
  - Probably not fine for users seeking to ensure non-redundant alternate path.

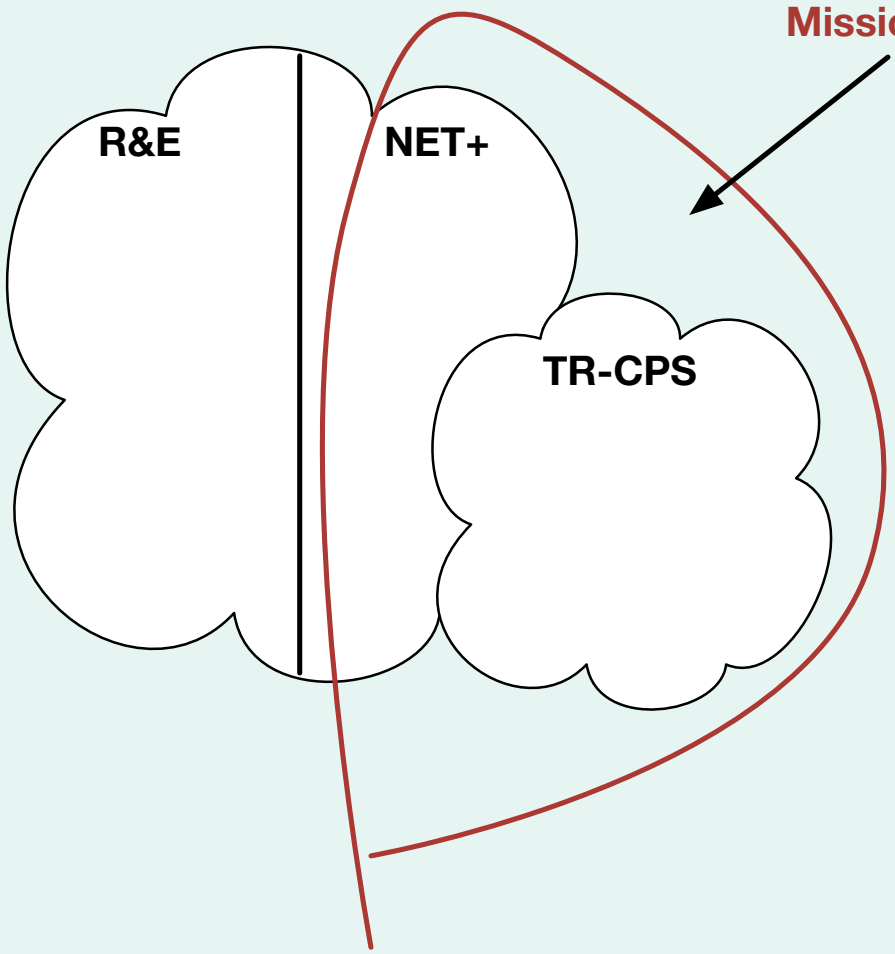
Internet2 Network Taxonomy



Internet2 Network Taxonomy



Internet2 Network Taxonomy



But does that architecture reflect the intent?

Network Architecture & Augmentation Principles (DRAFT)

Wasted Capacity

- Eliminate hair-pinning wherever possible to eliminate wasted capacity.
  - Add new routers, move peers from a switch to a router, etc.
- Egress traffic at the closest point

Routers

- Merge / hybridize AL3S and TR-CP routers.
- New routers should go into 3+-way junctions.
- New routers should maintain the full mesh of AL3S VLANs.

AL2s Augmentation

- When an upgrade to an AL2S circuit is called for, upgrade the express route between AL2S switches adjacent to routers (as opposed to the local route between adjacent AL2S switches or directly connecting routers).
- Handle TR-CPS upgrade needs by using AL2S.
- Land international connections at exchange points (because that's what we want other countries to do).
- Enable software programmability and automation (where possible).

ERIC BOYD

Senior Director, Strategic Projects

JULY, 2015

GROWING THE INTERNET2 NETWORK ABOVE 100G: CHOICES AND CHALLENGES

eboyd@internet2.edu

