USC Viterbi
School of Engineering
Department of Computer Science

Impactful Routing Research with the PEERING Testbed

Combining intradomain emulation with *real* BGP connectivity

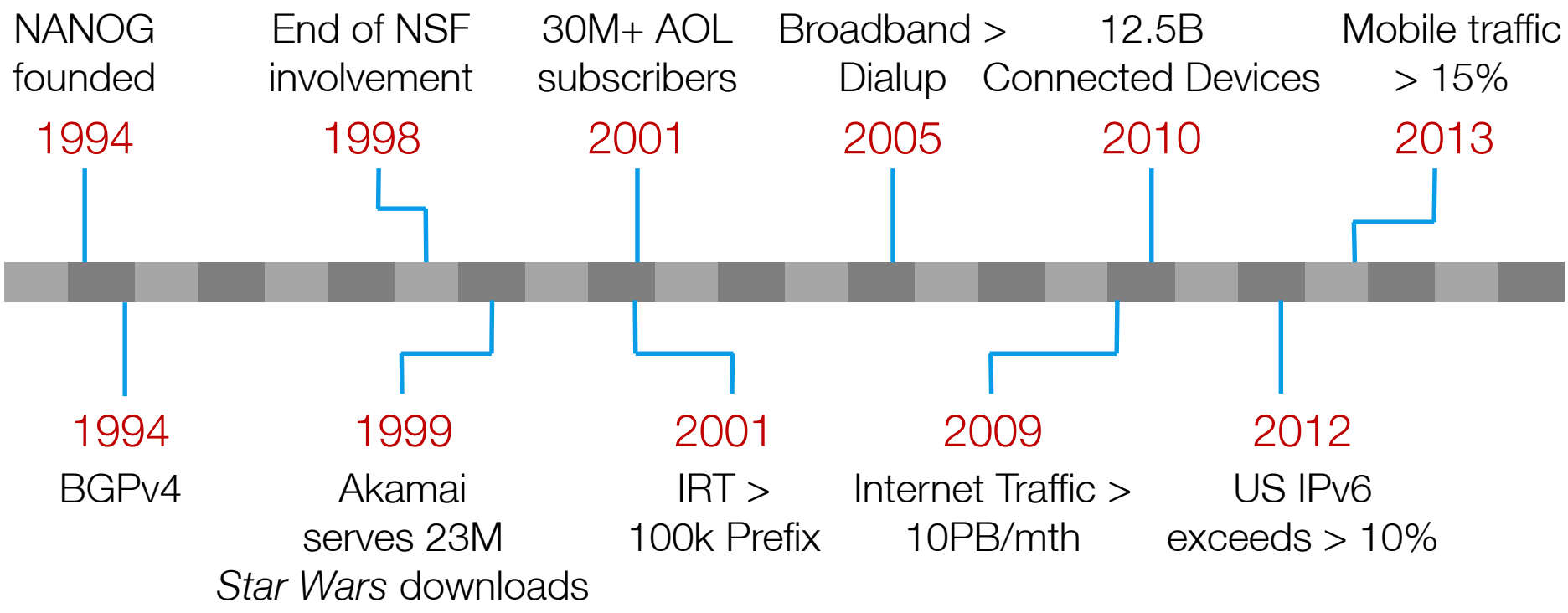
Brandon Schlinker, Kyriakos Zarifis, Ethan Katz-Bassett, and Minlan Yu
University of Southern California, California, USA

Italo Cunha, Universidade Federal de Minas Gerais, Minas Gerais, Brazil
Nick Feamster, Georgia Institute of Technology, Georgia, USA

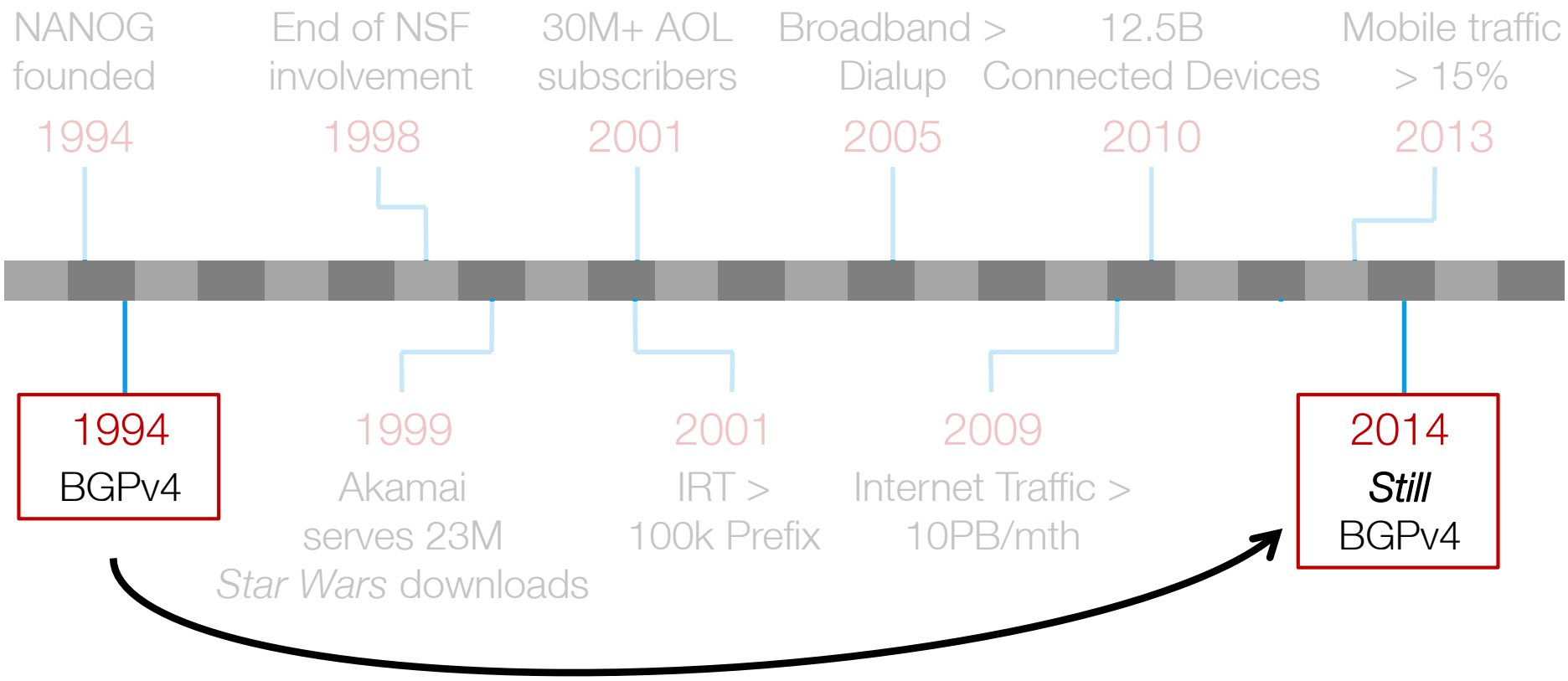
NANOG on the ROAD
Los Angeles, California, USA
May 2014

Funded By: Google NSF

20+ Years of Internet Innovation



Yet technology from 1994 *still* used today



What's so bad about BGP?

BGP contributes to many of the Internet's
fundamental problems

Examples of problems created by BGP

BGP contributes to many of the Internet's **fundamental problems**

BGP design results in:

- Poor performance *(inflated routes)*
- Security vulnerabilities *(route hijacking)*
- Longer outages (*lengthy convergence times*)
- Routing failures (*route redistribution issues*)
- QoS problems in gaming, VoIP (*path oscillations*)

(the list goes on…)

Examples of problems created by BGP

BGP contributes to many of the Internet's **fundamental problems**

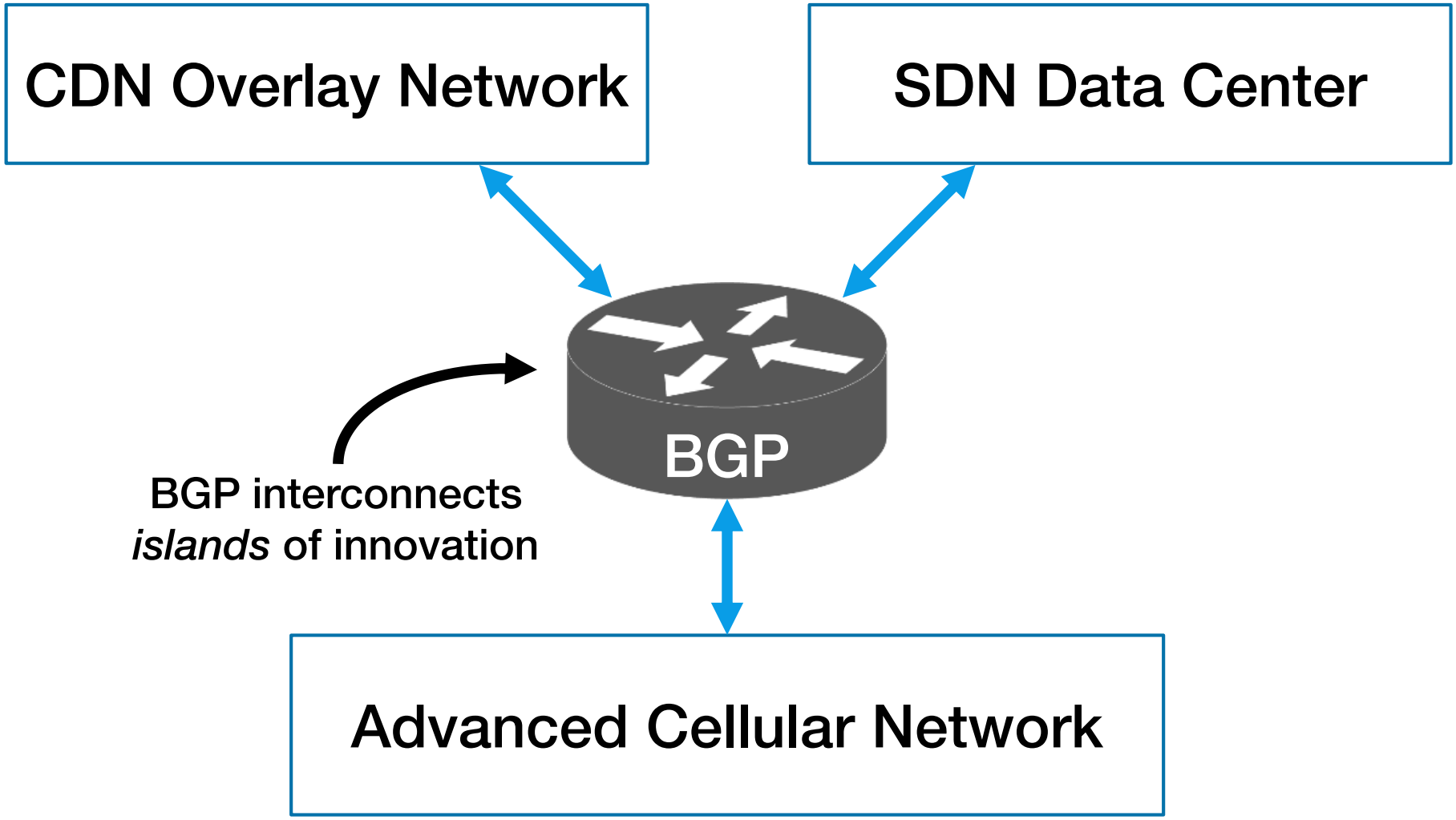
BGP design results in:

- Poor performance *(inflated routes)*
- Security vulnerabilities *(route hijacking)*
- Longer outages (*lengthy convergence times*)
- Routing failures (*route redistribution issues*)
- QoS problems in gaming, VoIP (*path oscillations*)

(the list goes on…)

We need research to ***understand*** and ***improve*** BGP

BGP limits capabilities of today's networks



How do we improve BGP?

Remainder of Talk:

- Why is ***impactful*** BGP research and innovation so difficult?
  Impactful ⇨ more than just a paper or RFC

How do we improve BGP?

Remainder of Talk:

- Why is ***impactful*** BGP research and innovation so difficult?
  Impactful ⇨ more than just a paper or RFC

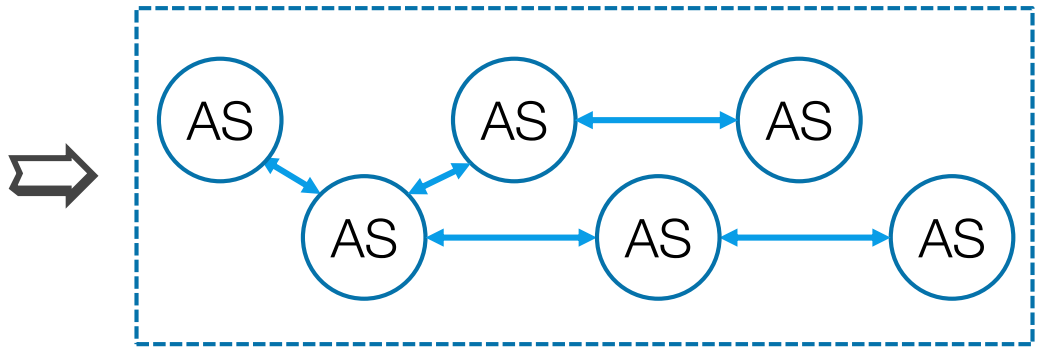
- How our PEERING testbed enables ***impactful*** BGP research
  Providing the ***control*** and ***realism*** needed to tackle ***key*** BGP problems

BGP interactions make research difficult

Interdomain problems often defined by *interactions* between ASes

- ***Investigate*** interactions when researching / defining a problem
- ***Incorporate*** these interactions when evaluating new system / technique

understanding interactions is ***key*** to productive research!

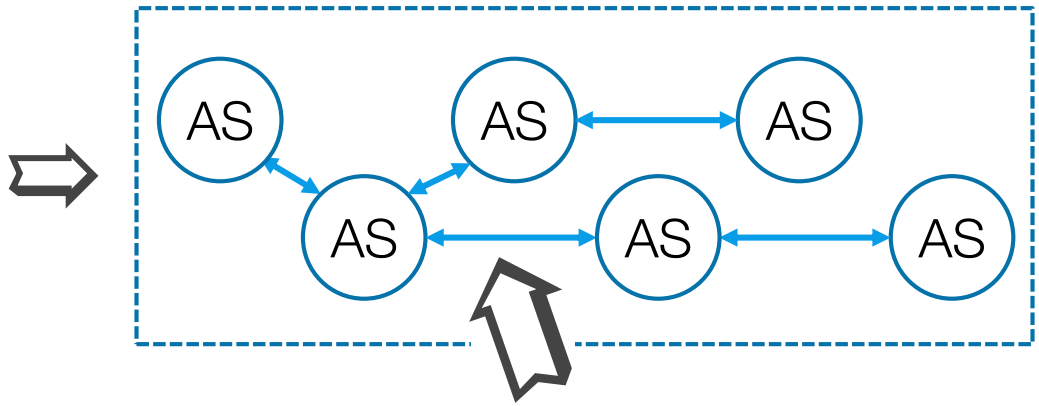


BGP interactions make research difficult

Interactions *cannot* be predicted / modeled

- ***Defined*** by the ***unknown*** policies of other ASes
- ***Driven*** by network conditions and operator updates
- Makes defining problems and ***realistic*** evaluation difficult!

understanding interactions is ***key*** to productive research!



interactions ***driven*** and ***defined*** by AS policy

Existing tools for BGP research

Route monitors

- Traceroutes, route collectors, BGP beacons, looking glasses

Existing tools for BGP research

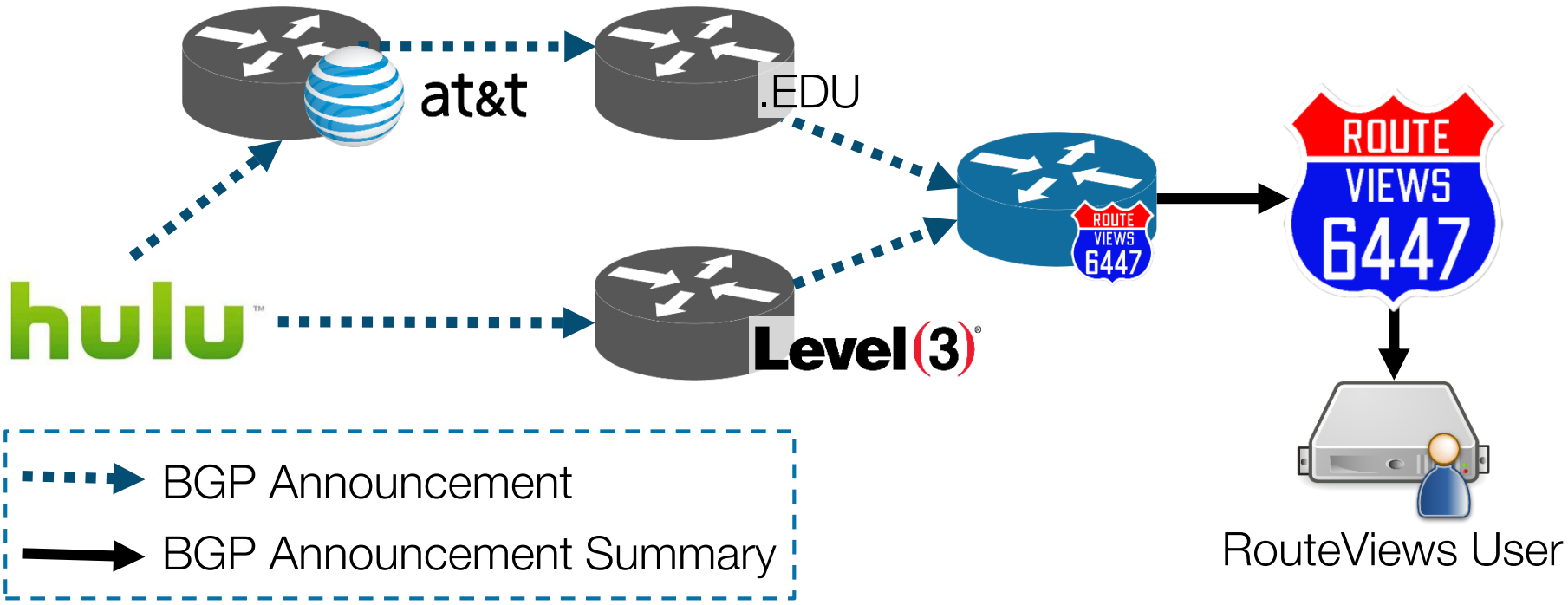
Route monitors

- Traceroutes, route collectors, BGP beacons, looking glasses

Simulation

- Provides complete control and visibility of:
  - AS interactions based on simulated policies
  - Interactions between routing protocols

Limitations of Route Monitors



- *Route monitors* do *not* support *interaction*
- Example: although RouteViews provides ***realism***, it lacks ***control***
  - Cannot advertise routes with path poisoning to find alternative paths
  - Cannot pass traffic across upstream links to evaluate performance
  - Cannot observe interactions between ***inter*** and ***intra***domain

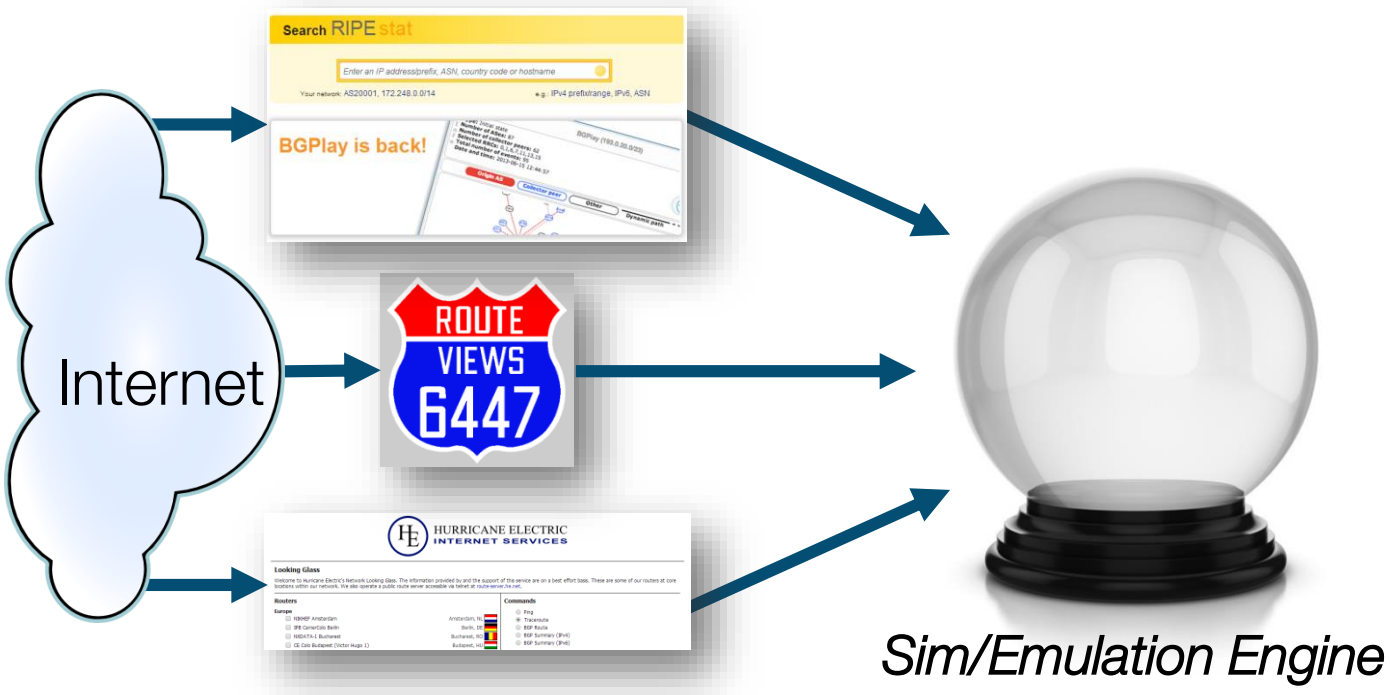
Limitations of Simulation

System Inputs:
Intradomain topology
Internet topology
BGP policies of all Internet's ASes

Sim/Emulation Engine

- Cannot realistically simulate interactions ***between*** AS
  - Because we don't ***know*** Internet's topology, nor every AS's policies
  - ***Accuracy*** of simulation engine bounded by accurate of its inputs
- Thus, while simulation provides full control, it ***lacks realism***

Route Monitors + Simulation?



- What about using route monitor data as input to simulation?
- Route monitors only provide ***snapshot*** of AS's state at time
  - Can't tell how an AS will ***react*** to events (neighbor update, failures)
  - Thus, cannot ***interact*** with the AS in a simulation environment

How does PEERING help?

Remainder of Talk:

- How *our* PEERING testbed enables ***impactful*** BGP research

  Providing the ***control*** and ***realism*** needed to tackle ***key*** BGP problems

PEERING Testbed Enables BGP Research

PEERING:

Pairing Emulated Experiments with Real Interdomain Network Gateways

PEERING is a BGP Testbed for Researchers *and* Network Operators

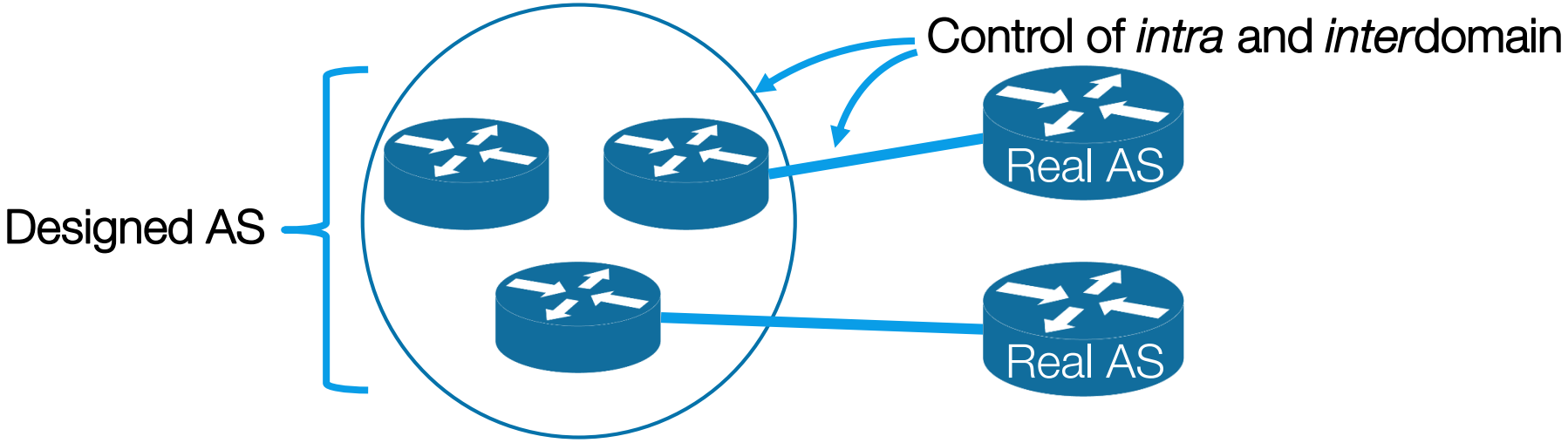
Contains tools needed to execute ***impactful*** routing experiments

- Can help ***analyze*** existing systems, policies, and topologies
- Can help ***evaluate*** new technologies, routing engines, topologies, etc.

How PEERING Testbed Works

With PEERING, a researcher or network operator:

- *Designs* an AS, including its topology and routing policies
- *Connects* the designed AS to *real* ASes on the Internet via BGP
- *Controls* the AS, including its exchange of traffic and routes



PEERING's components

Combines two components:

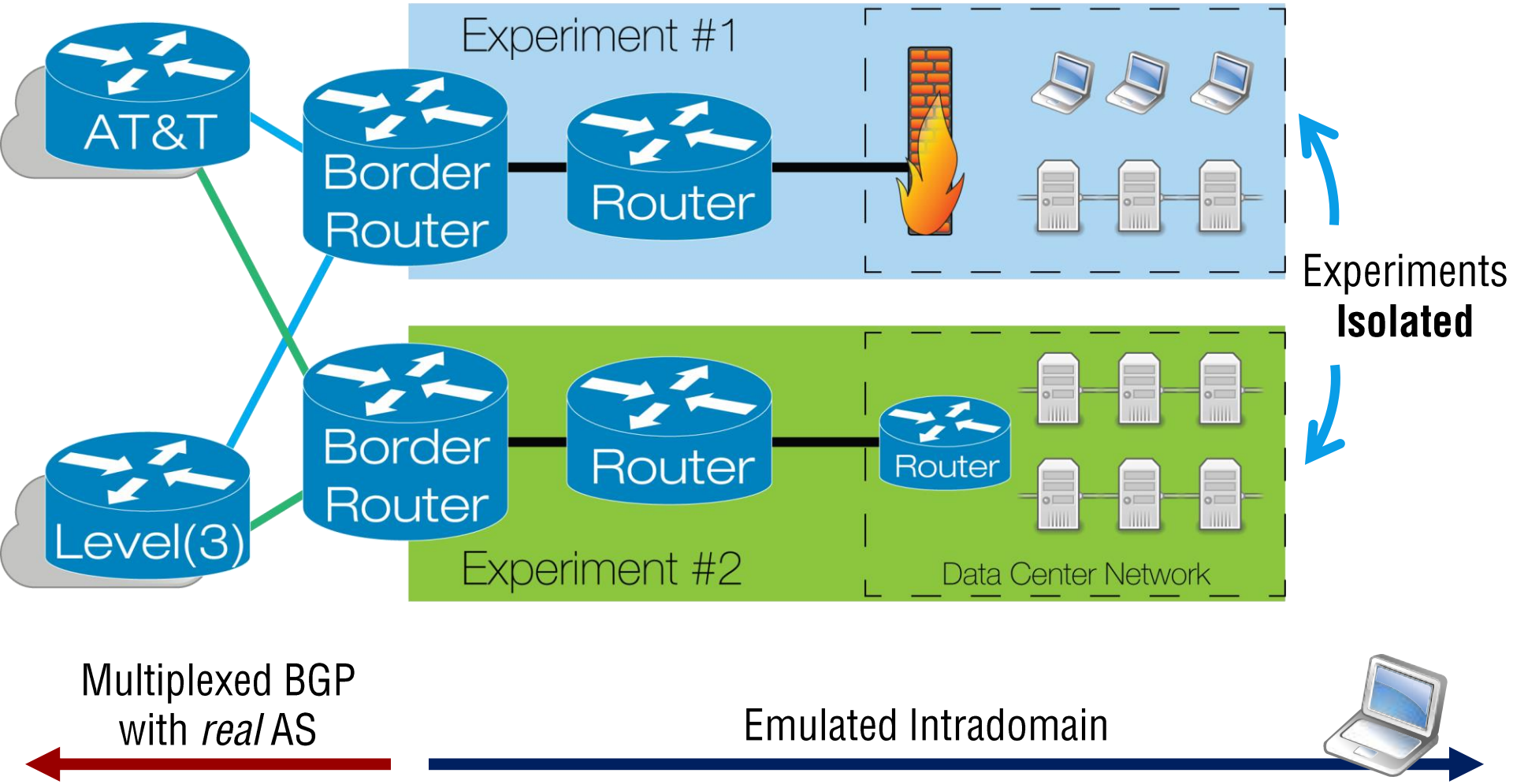
- Extended version of Mininet
  - Mininet enables *highly scalable* emulation of SDN networks
  - Extended Mininet to make easier to emulate non-SDN infrastructure
    - Extensions open-sourced as Mini**NExT**

PEERING's components

Combines two components:

- Extended version of Mininet
  - Mininet enables *highly scalable* emulation of SDN networks
  - Extended Mininet to make easier to emulate non-SDN infrastructure
    - Extensions open-sourced as Mini**NExT**
- Transit Portal
  - BGP multiplexing service and autonomous system (AS 47065)
  - US and international points of presence

Experiment isolation with PEERING



Advantages of BGP via PEERING

With previous approaches, experiments:

- Could only *monitor* changes on routes
- Could not interact with other ASes (such as exchanging routes, traffic)
- Provided only partial insight into *inter-AS interactions*

Advantages of BGP via PEERING

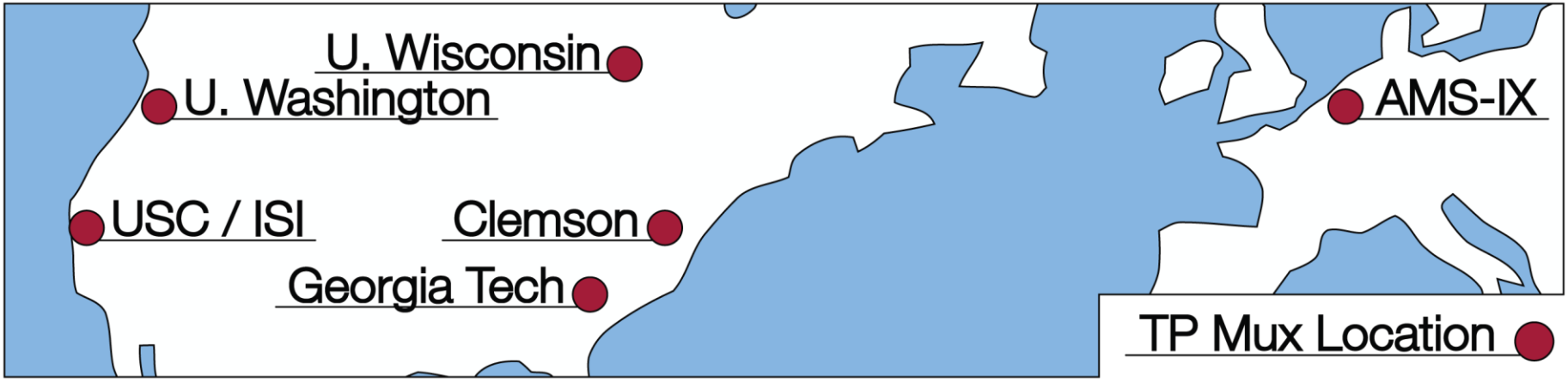
With previous approaches, experiments:

- Could only ***monitor*** changes on routes
- Could not interact with other ASes (such as exchanging routes, traffic)
- Provided only partial insight into ***inter-AS interactions***

With PEERING, experiments:

- Have access to a ***real*** AS number (47065) via multiplexing
- Can interact with ***real*** ASes on the Internet
  - ***Real*** ASes = { ISPs, hosting / content networks, academic networks, etc. }
  - Experiments **exchange routes, traffic** as if directly connected to upstream
  - Announcements propagated to the ***real*** Internet

PEERING PoPs

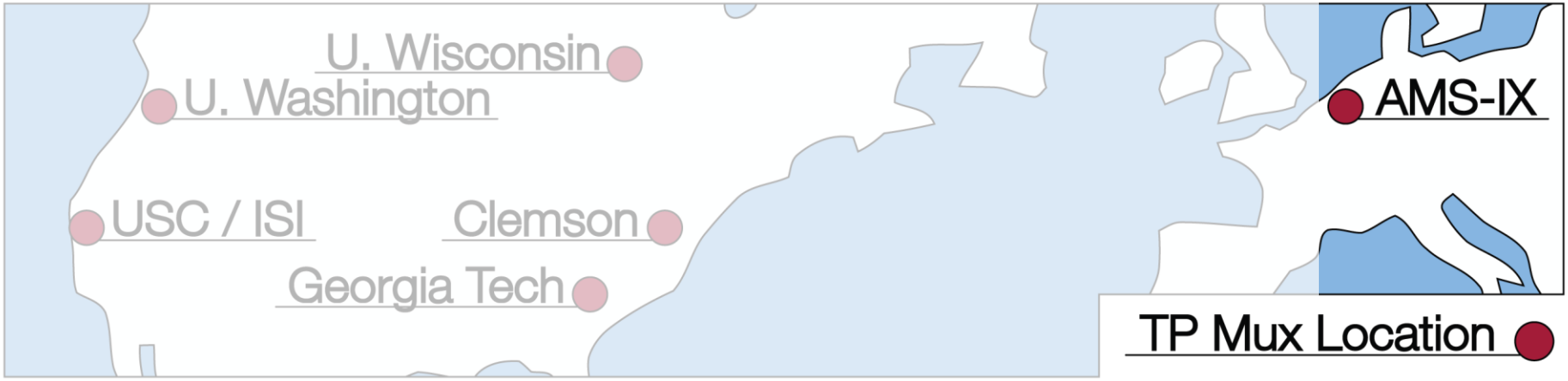


Infrastructure

- Dedicated AS (47065) and IP space (184.164.224.0/19)
- 7 points of presence across three continents (Brazil just added)
- Connectivity to over 600 ***real*** ISPs via BGP
- Infrastructure is growing into more IXPs, CDNs, and ISPs

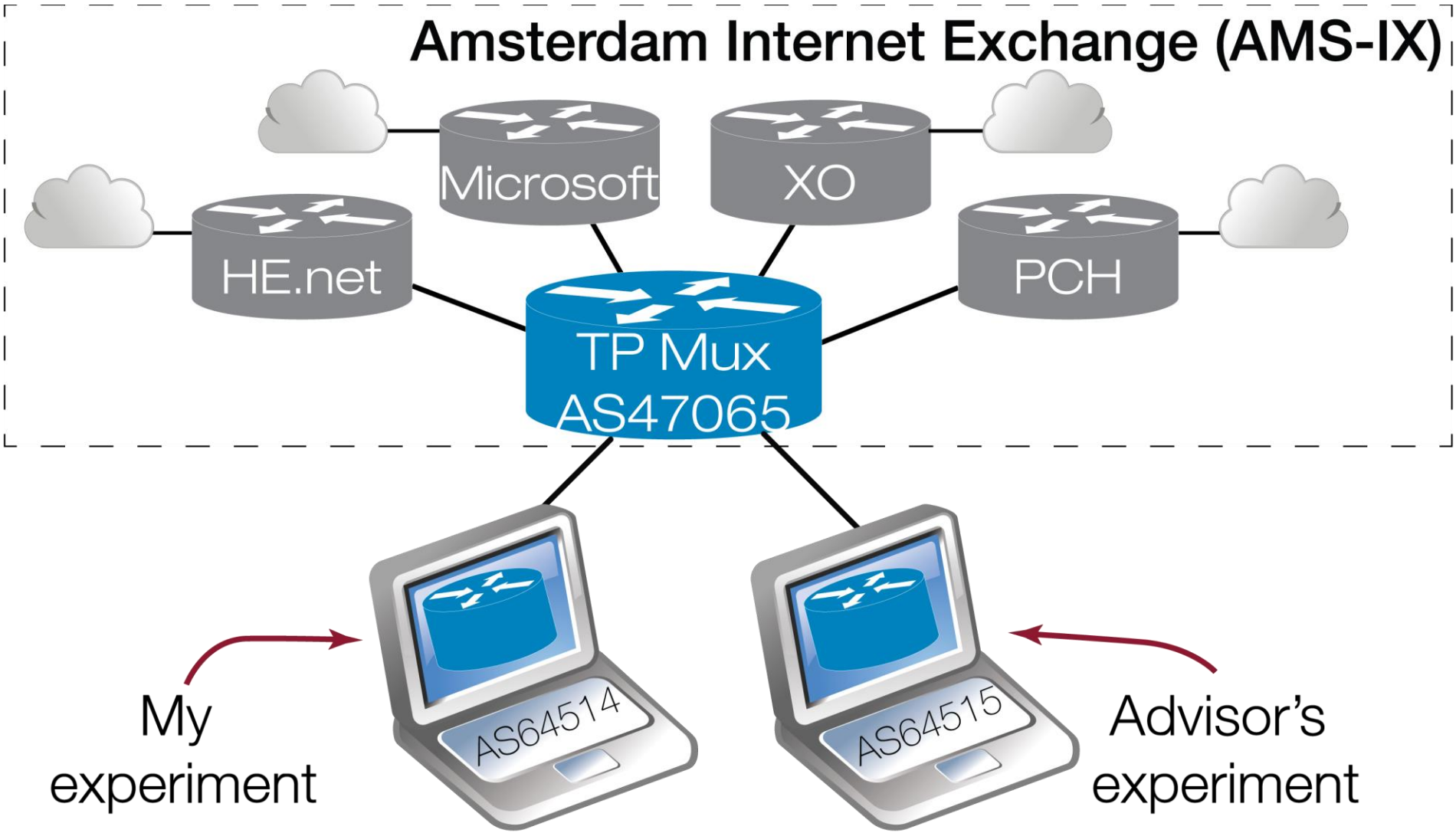
Experiment can connect to multiple PoPs simultaneously

PEERING PoPs

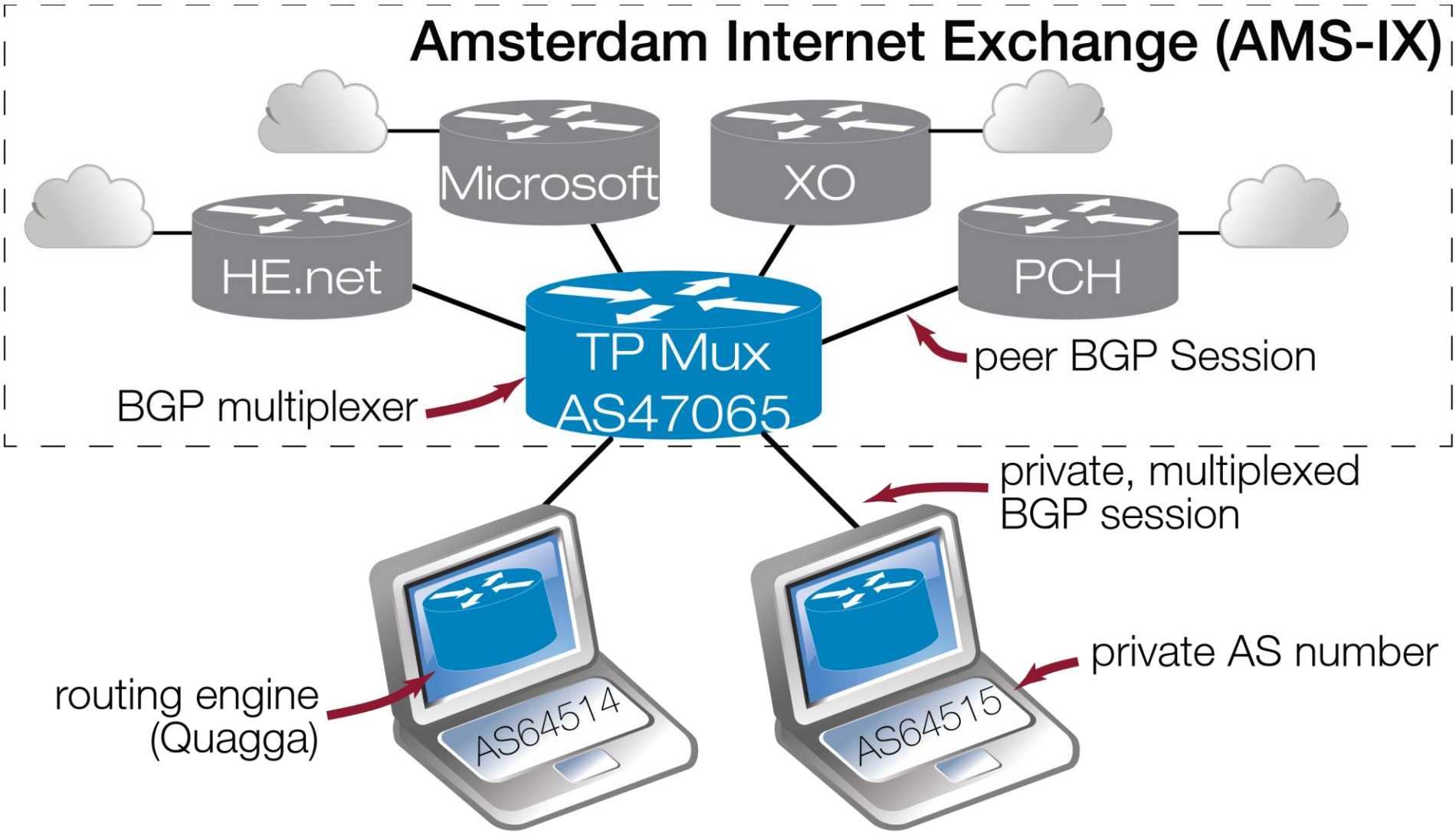


Let's take a look at the PEERING PoP at AMS-IX

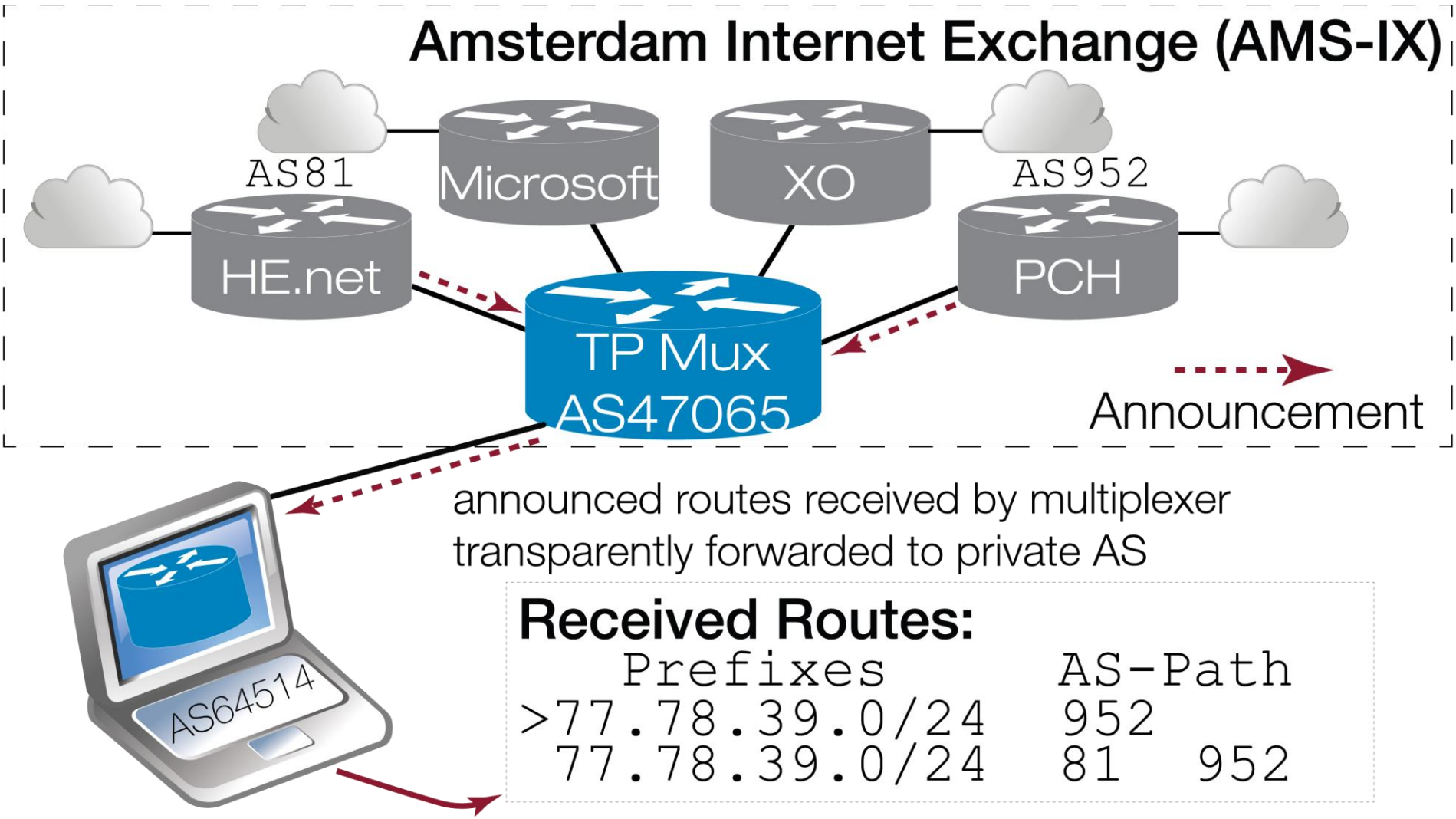
PEERING at AMS-IX



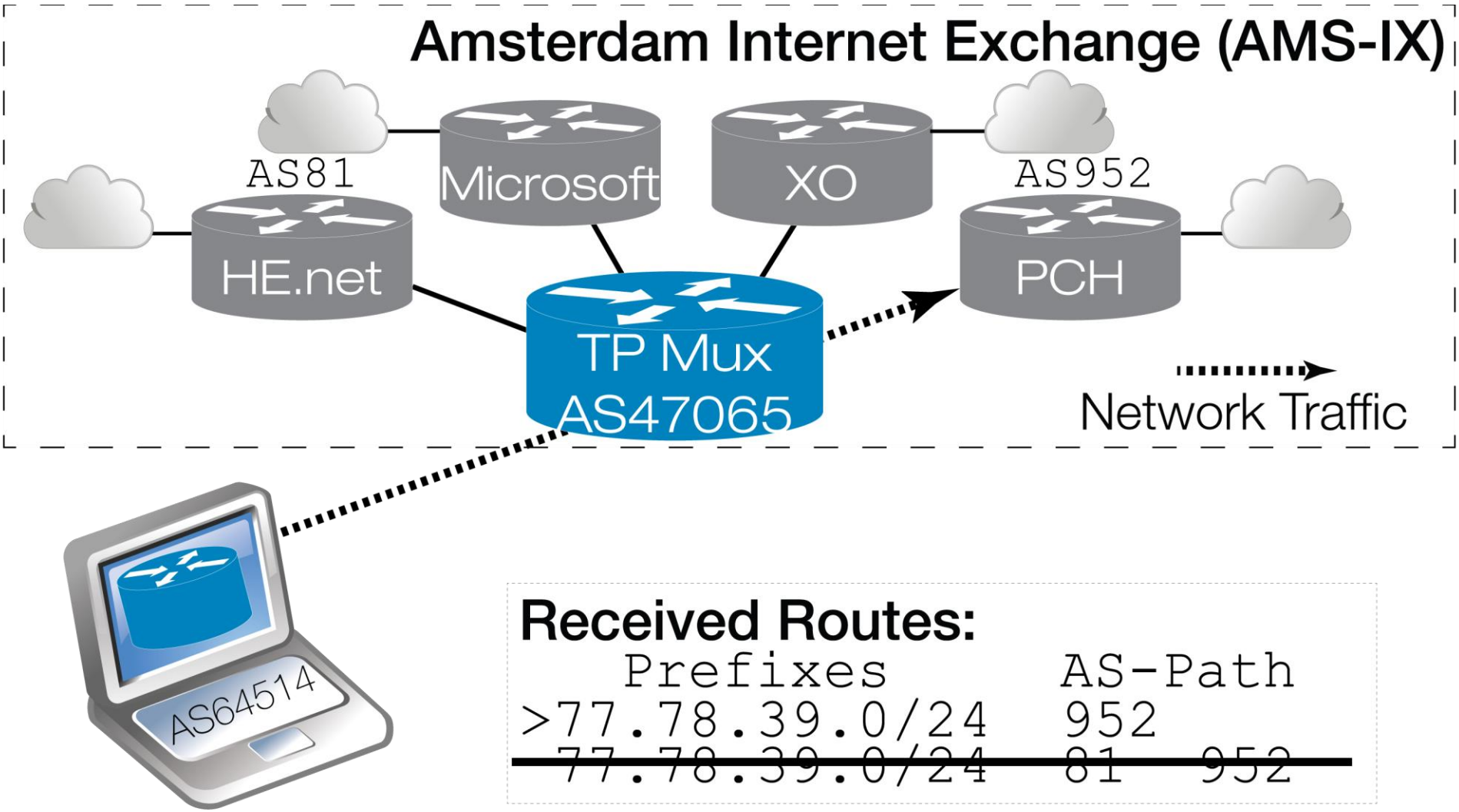
PEERING at AMS-IX



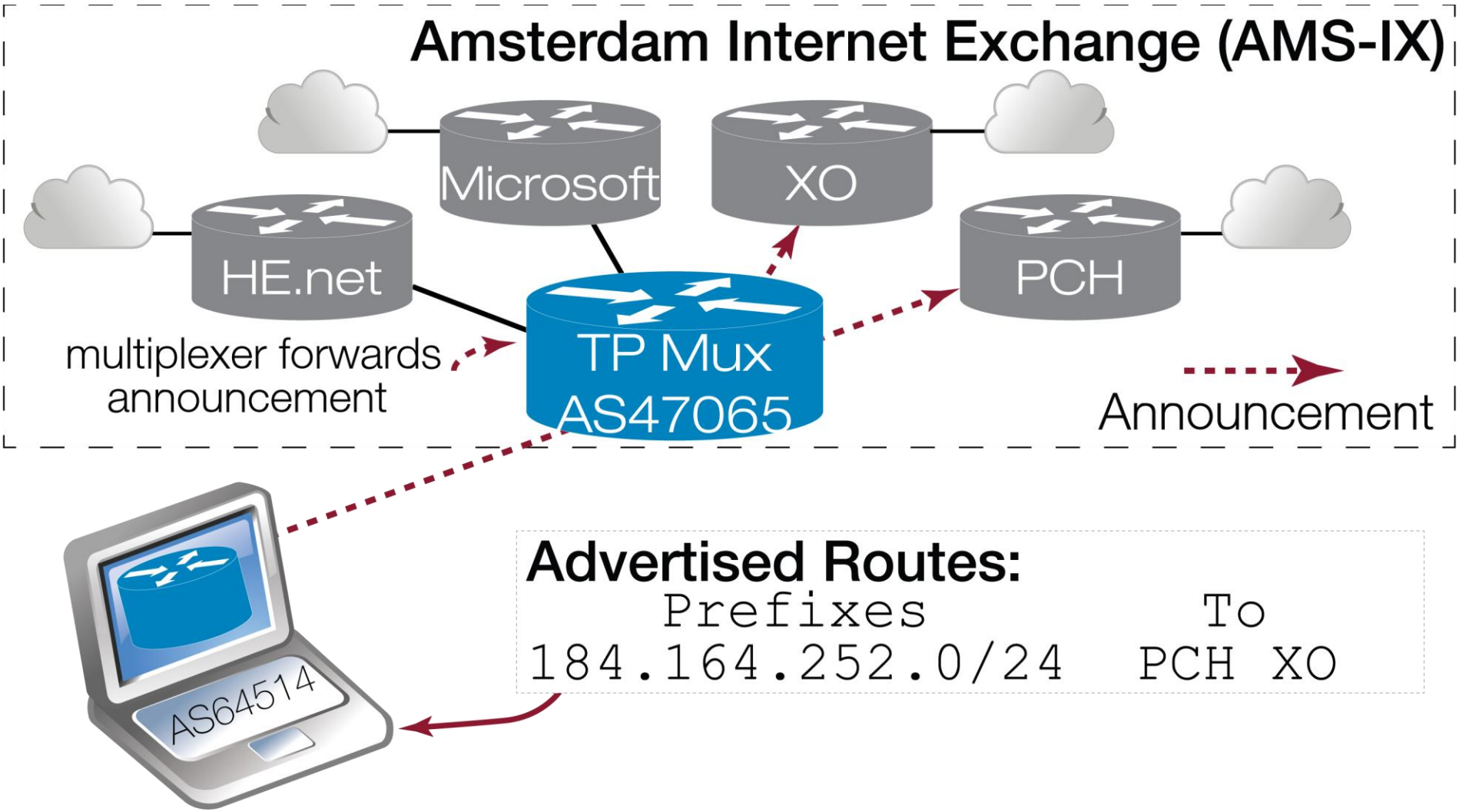
Receiving Routes via AMS-IX



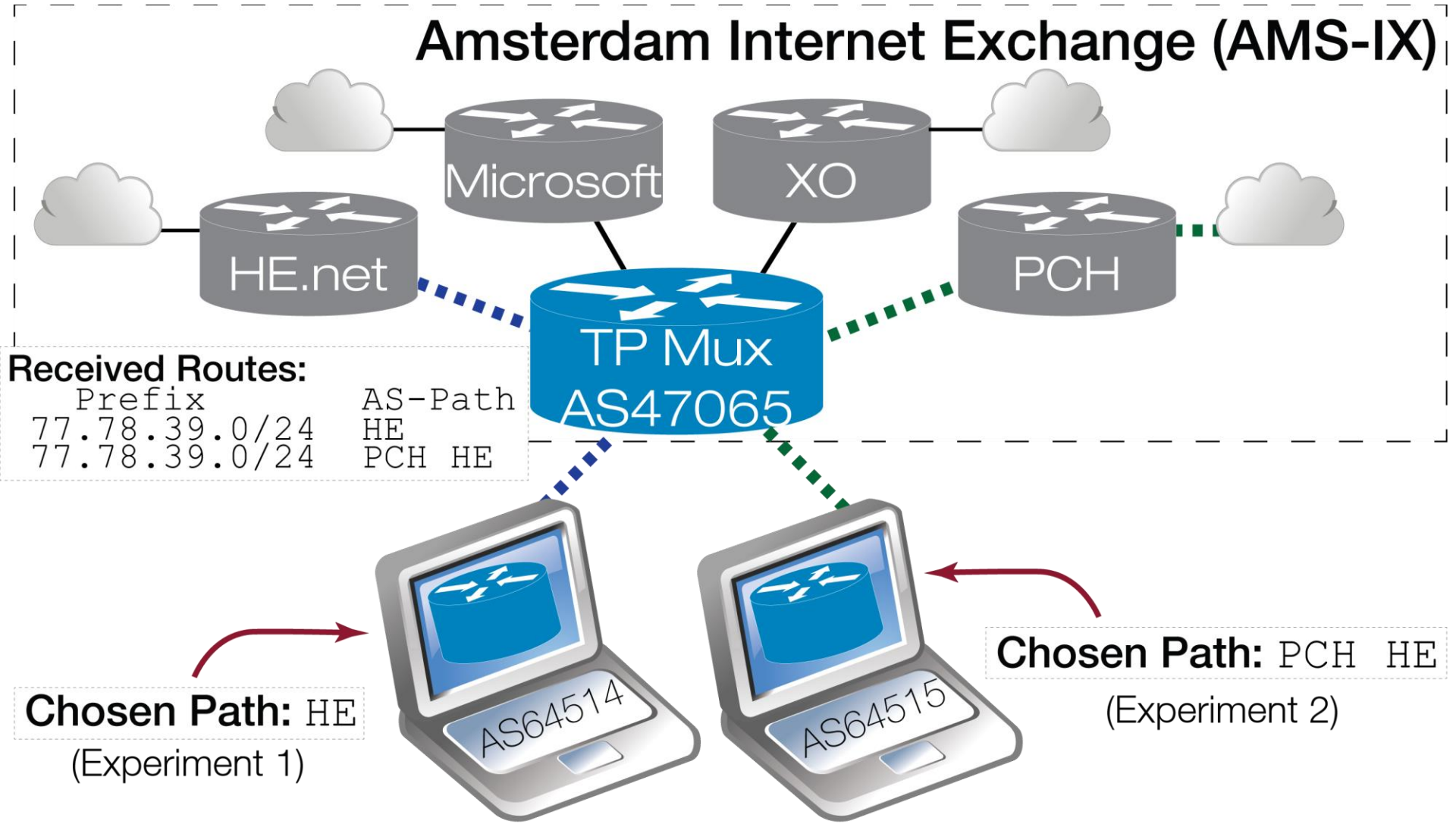
Exchanging Traffic via AMS-IX



Making Announcements via AMS-IX



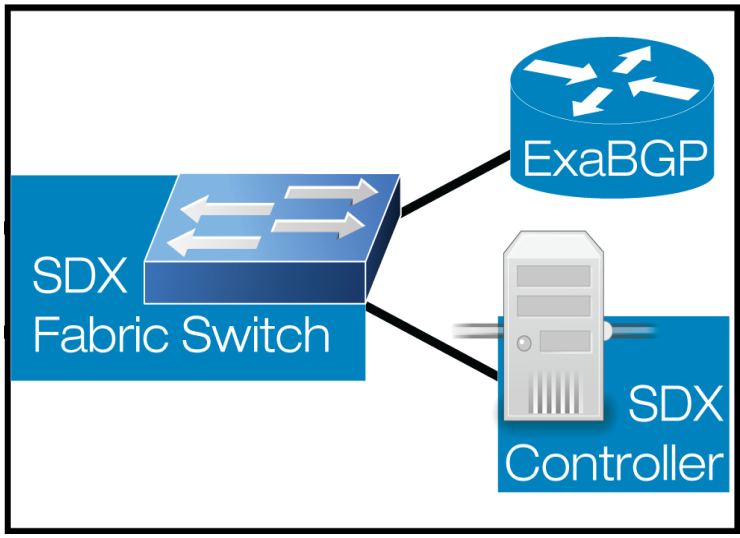
Isolation of Experiments at AMS-IX



Example PEERING Experiment

Testing Scalability of Software Defined Internet Exchange (SDX)

- SDX is essentially an advanced ***route server*** for IXPs providing control
  - Enables IXP peers to setup advanced policies (application specific peering)
  - BGP updates and controller policies define SDN fabric switch rules
- Built the SDX components – how do conduct ***realistic*** evaluation?



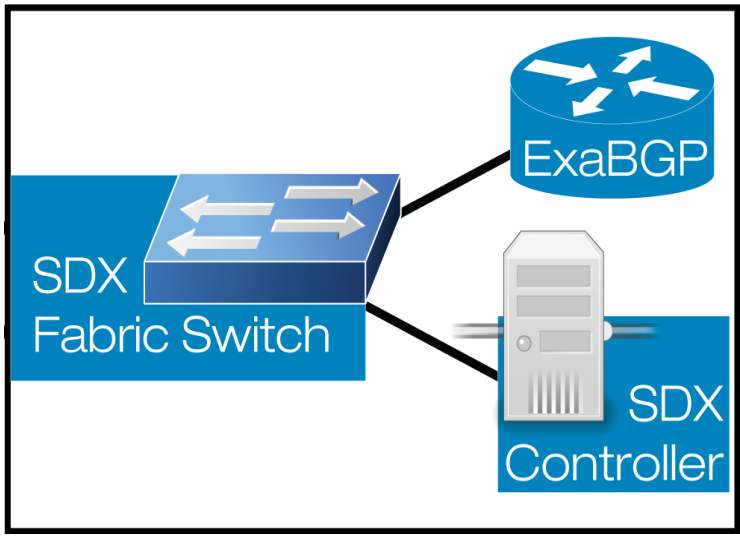
SDX Fabric

Example Peering Experiment

Testing Scalability of Software Defined Internet Exchange (SDX)

- SDX is essentially an advanced ***route server*** for IXPs providing control
  - Enables IXP peers to setup advanced policies (application specific peering)
  - BGP updates and controller policies define SDN fabric switch rules
- Built the SDX components – how do conduct ***realistic*** evaluation?

Can we test how SDX would operate in *real* IXP exchange?



SDX Fabric

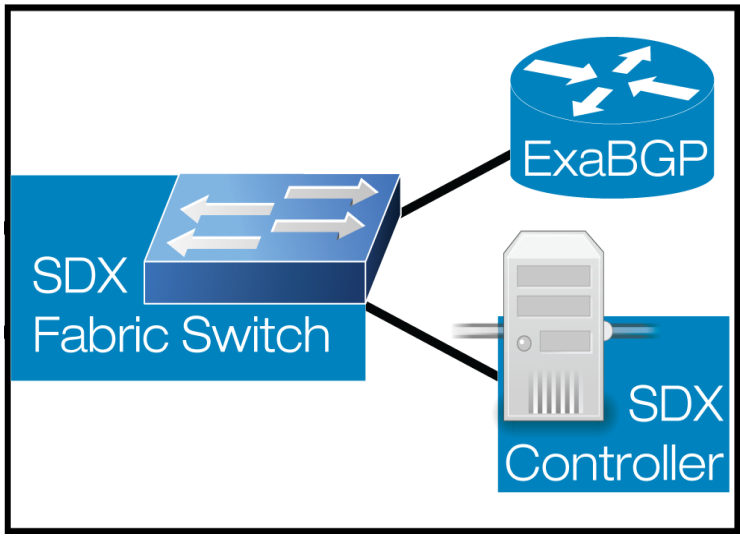
Example PEERING Experiment

Testing Scalability of Software Defined Internet Exchange (SDX)

- SDX is essentially an advanced ***route server*** for IXPs providing control
  - Enables IXP peers to setup advanced policies (application specific peering)
  - BGP updates and controller policies define SDN fabric switch rules
- Built the SDX components – how do conduct ***realistic*** evaluation?

Can we test how SDX would operate in *real* IXP exchange?

Need to add *real* peers from *real* IXP!

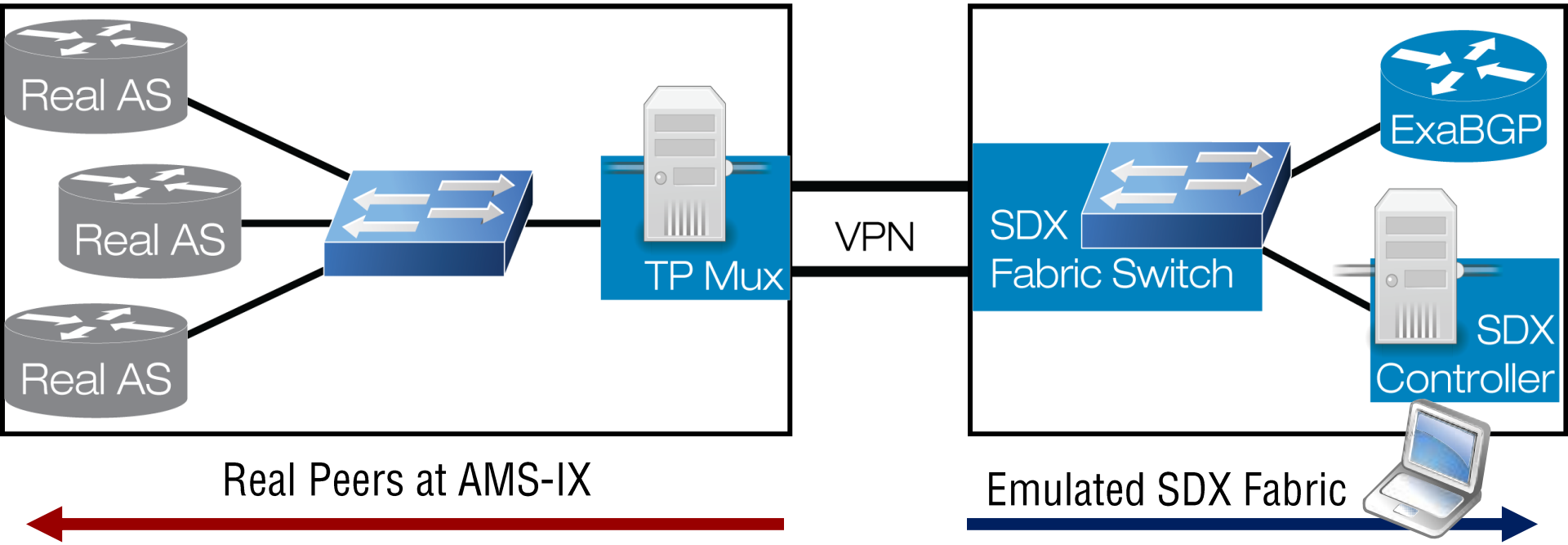


SDX Fabric

Example PEERING Experiment

Testing Scalability of Software Defined Internet Exchange (SDX)

- Real AS located at AMS-IX ***virtually*** added to emulated SDX fabric
  - TP Mux transparently tunneled BGP sessions over VPN connection
  - ExaBGP maintains BGP sessions with peers, forwards to controller
- Controller updates SDX fabric in response to ***real*** BGP messages



PEERING's use in prior research

- **Software Defined Internet Exchange (SDX)**
  - Enabled SDX experiments to exchange Internet traffic via diverse paths
  - Can also be used for emulation of fabric and scalability testing

PEERING's use in prior research

- **Software Defined Internet Exchange (SDX)**
  - Enabled SDX experiments to exchange Internet traffic via diverse paths
  - Can also be used for emulation of fabric and scalability testing

- **LIFEGUARD: Practical Repair of Persistent Route Failures**
  - Leveraged BGP path poisoning BGP to route around routing failures
  - Used Transit Portal for control and interactions with other ASes
    - Enabled feasibility of scheme on real Internet to be evaluate (path availability, receptiveness to poisoned announcements, etc.)

PEERING's use in prior research

- **Software Defined Internet Exchange (SDX)**
  - Enabled SDX experiments to exchange Internet traffic via diverse paths
  - Can also be used for emulation of fabric and scalability testing

- **LIFEGUARD: Practical Repair of Persistent Route Failures**
  - Leveraged BGP path poisoning BGP to route around routing failures
  - Used Transit Portal for control and interactions with other ASes
    - Enabled feasibility of scheme on real Internet to be evaluate (path availability, receptiveness to poisoned announcements, etc.)

- **PoiRoot: Investigating the Root Cause of Interdomain Path Changes**
  - Investigated triggers of route changes on the Internet
  - Made announcements via Transit Portal to partially infer AS's policies

PEERING's Infrastructure

Existing Infrastructure

- 5 US PoPs at universities, 2 International PoPs (AMS-IX, Brazil)
- /19 dedicated address space, dedicated ASN

PEERING's Infrastructure

Existing Infrastructure

- 5 US PoPs at universities, 2 International PoPs (AMS-IX, Brazil)
- /19 dedicated address space, dedicated ASN

Future Expansion

- BGP peerings at IXP locations and with ISPs + CDNs to provide:
  - greater path diversity
  - direct interaction with more ASes to enable richer experiments
  - Interaction with ASes from different vantage points to get more data
- Backbone between PoPs via Internet2 to support experiments
- Using Emulab to support easier emulation of datacenters
- Increased automation for experiment setup and management
  - Make PEERING's infrastructure easily accessible to more experiments

Future of BGP Research with PEERING

We'd like *your* input to help shape the future of PEERING

- Experiments you can envision being run with PEERING

Future of BGP Research with PEERING

We'd like *your* input to help shape the future of PEERING

- Experiments you can envision being run with PEERING
- Important open problems in Internet routing
  - The best problems come straight from the source – operators!

Future of BGP Research with PEERING

We'd like *your* input to help shape the future of PEERING

- Experiments you can envision being run with PEERING
- Important open problems in Internet routing
  - The best problems come straight from the source – operators!
- Help us expand PEERING's connectivity into more networks
  - Peer with us or help host us!
  - Looking for contacts at IXPs, ISPs, remote peering

Conclusion

Little innovation in BGP in past 20 years

- Yet BGP is the root of some of the Internet's most fundamental problems
- Researchers have long lacked the tools needed to run BGP experiments

PEERING changes the game in BGP research

- Connects emulated ASes with ***real*** ASes on the Internet

Looking to gain feedback and insight from network operators

- What problems would you like to see PEERING tackle?
- How can we best expand PEERING to help network operators?
- Interested in peering with us or hosting us?
- **Contact:** bgpmux-noc@gtnoise.net