ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

Δίκτυα Καθοριζόμενα από Λογισμικό

Ενότητα 4.2: Inter-domain SDN: Επισκόπηση έρευνας

Ξενοφώντας Δημητρόπουλος
Τμήμα Επιστήμης Υπολογιστών

Inter-domain SDN: A Research Overview

Vasileios (Vassilis) Kotronis

Outline

➔ Part I (~45-50 min)
- A brief recap of BGP and how it works
- Communication between SDN domain controllers
- Partial SDN deployment with BGP compatibility
- Outsourcing and centralizing inter-domain routing
- Control Exchange Points and end-to-end QoS
- Software Defined Internet Exchanges

➔ Part II (~40 min)
- SIREN: a hybrid SDN Inter-domain Routing EmulatioN framework
- Short demo of SIREN

➔ General directions for inter-domain SDN (~2-3 min)

PART I

Routing Hierarchies in the Internet*

→ The Internet = a network of networks/domains

→ How do we route packets within such an environment?

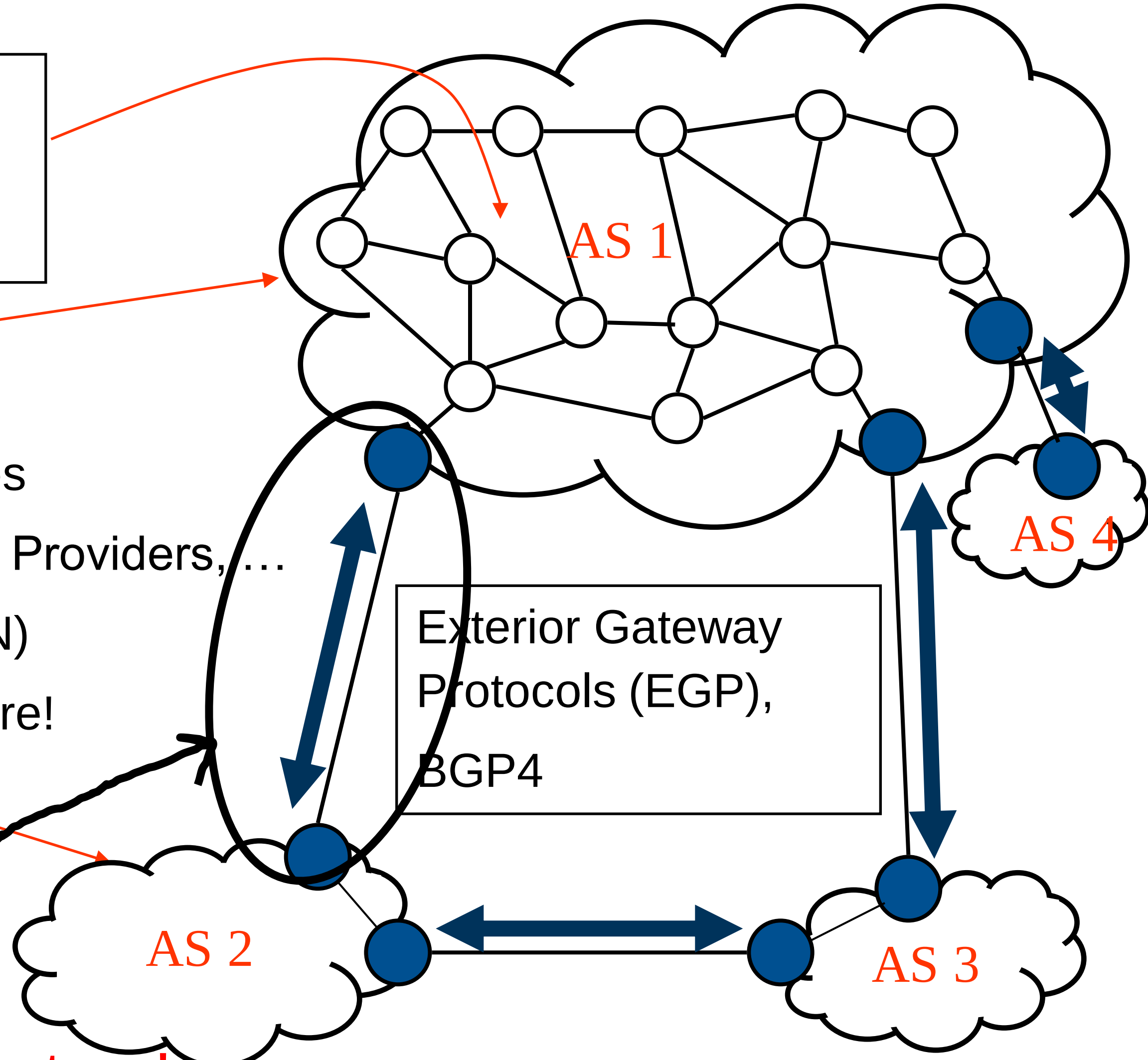
- **Level 1: Routing within a domain**
  - Use an Interior Gateway Protocol (IGP) for intra-domain routing
  - Based on Distance Vector or Link State → RIP, OSPF, IS-IS, …
  - Choice by administration of the routing domain
  - Example: HOL forms such a routing domain

- **Level 2: Routing between domains**
  - Use an Exterior Gateway Protocol (EGP) for inter-domain routing
  - Today‘s standard is a path vector protocol, supporting policies
    → Border Gateway Protocol (BGP), Version 4 (BGP4)
  - Routing domains = Autonomous Systems (ASes)

*Original slide from Dr. Xenofontas Dimitropoulos for the CN 2014 course, ETH Zurich

A basic view of the Internet*

Interior Gateway Protocols (IGP),
OSPF, IS-IS, ...



Autonomous Systems (AS):

- ➔ Managed by autonomous entities
  - ISPs, Governments, Content Providers, …
- ➔ Have a unique AS Number (ASN)
  - Around ~46000 ASes out there!

Border Router

Routing information

Focus of this Lecture!

*Original slide from Dr. Xenofontas Dimitropoulos for the CN 2014 course, ETH Zurich

Border Gateway Protocol (BGP)*

- Internet: Arbitrarily interconnected set of ASes
  - Not restricted to the Tier model or tree structures
  - Even denser than you might think (peering agreements at IXPs, etc.)

➔ BGP ‘‘is the glue that holds the Internet together“
  - Communicates **prefix reachability** information to ASes
  - Information collected by ASes is used to configure forwarding tables of border routers

➔ **Path vector** protocol
  - Exchange of routes to destinations in the form of AS path vectors (Dest_IP_Prefix, AS1→AS2→AS3→…)
  - No explicit distance metric exchanged!
  - ASes can detect routing loops by AS path analysis on route ads

➔ Extensive support for defining **routing policies**
  - Customers/Providers/Peers, TE, security, cost reduction, …

*Original slide from Dr. Xenofontas Dimitropoulos for the CN 2014 course, ETH Zurich

Why do we need a new "SDN-BGP"? The Internet works fine as is, right?

- **Current architecture does NOT support innovation**
  - Strong coupling between architecture (protocols etc.) and infrastructure (network devices)
  - Architectural rigidity → barrier to innovation, not sufficient modularity
- **Management, trouble-shooting and security are hard**
  - Manual configuration of several knobs
  - "Masters of complexity" paradigm
- **Large convergence times (~10s of sec up to minutes)**
- **Scalability issues due to routing table size and churn**
  - Controller has plenty of CPU power and capacity, routers do not
- **BGP routing inconsistencies can cause anomalies**
  - Black-holes, loops, routing disputes
- **Difficulties with enforcing policy (outdated BGP knobs)**

SDNi: SDN Controller Interconnection

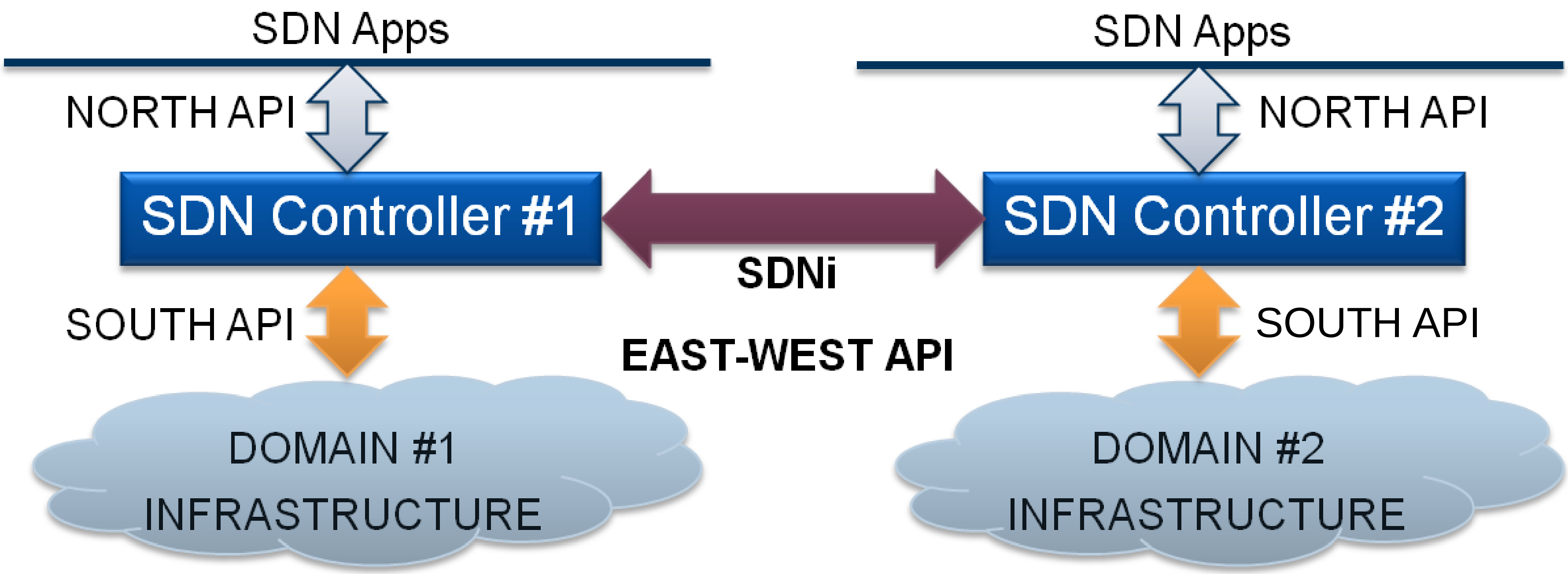
- **Main motivating factor: SDN Partitioning**
  - Scalability (devices/controller)
  - Manageability (separate responsibilities)
  - Privacy (each domain on its own)
  - Deployment (SDN islands within legacy networks)

- **Oriented to horizontal partitioning**
  - In contrast to vertical e.g., like FlowVisor/OpenVirtex does

- **Advocates interconnection between controllers**
  - Each SDN domain: controlled by one SDN Controller/NOS
  - SDNi is an interface mechanism between SDN domains
  - Relates to control plane cooperation

http://www.ietf.org/proceedings/84/slides/slides-84-sdnrg-5.pdf

SDNi main idea and exchanged state



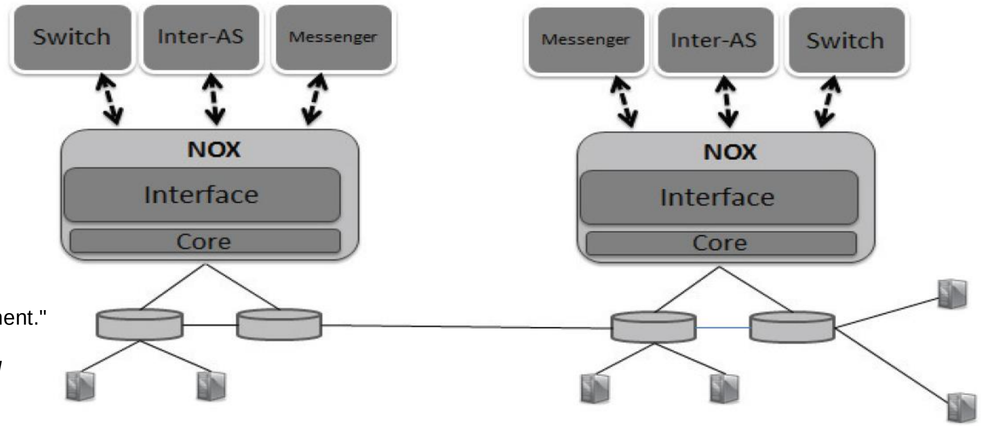
➔ What can be exchanged?

- Network topology (or “slices” thereof)
- Network events (e.g., “link-down”, or “DDoS congestion”)
- User-defined request information (e.g., “allocate 1Gbps now”)
- User app QoS requirements (e.g., “latency<40ms”)
- Infrastructure status (e.g., energy consumption)

http://www.ietf.org/proceedings/84/slides/slides-84-sdnrg-5.pdf

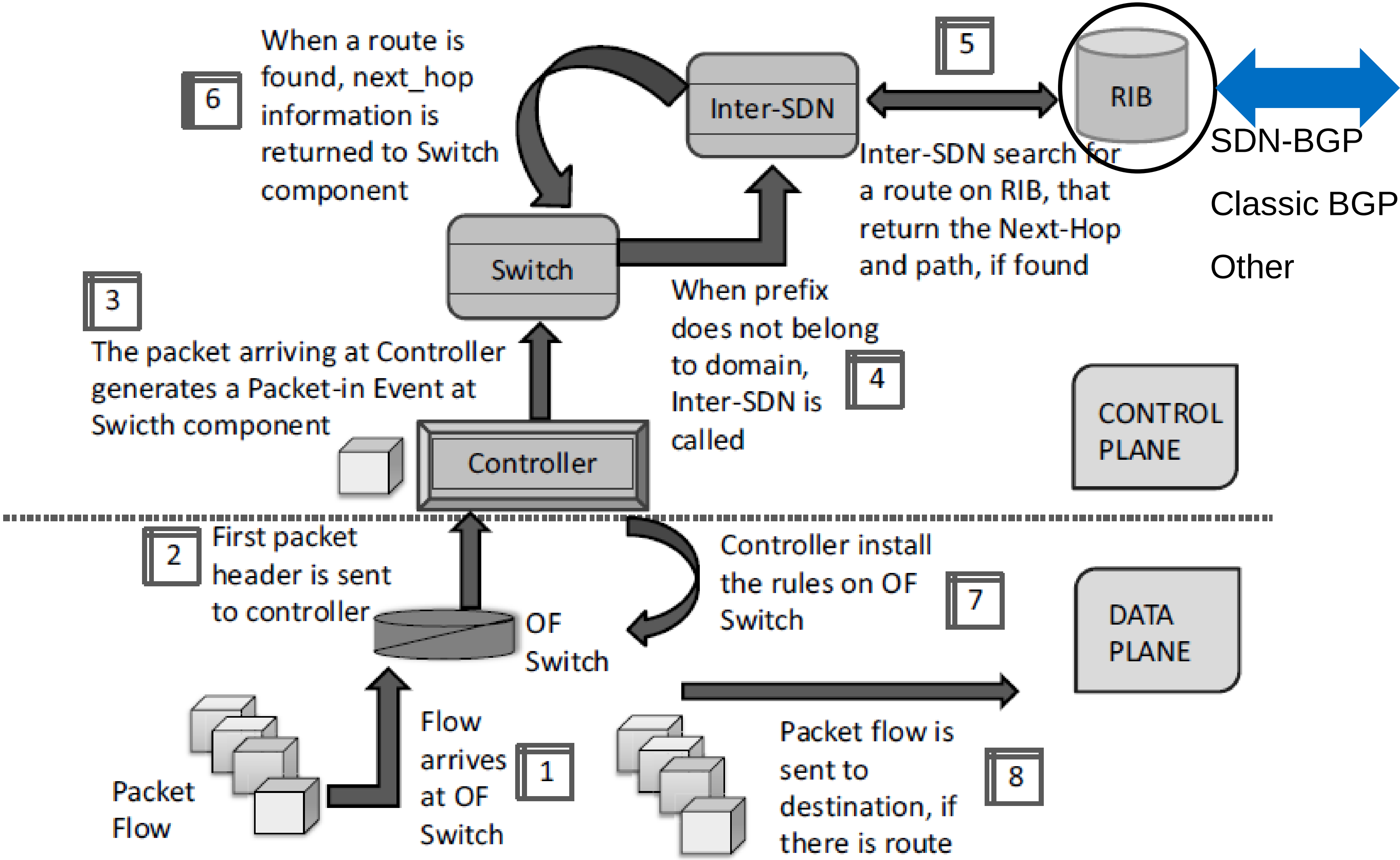
An Inter-AS Routing Component for SDN

- Component that runs on top of an SDN controller (NOX)
  - Extend “Switch” and “Messenger” modules of NOX
- Exchange of inter-AS prefix reachability information
- Routing based only on destination IP prefix as vanilla BGP
- Maintain DEST_IP_PREFIX, AS_PATH, NEXT_HOP, added NEXT_DPID and NEXT_DPID_PORT
- Loops are handled via AS_PATH checking (as in BGP)
- Essentially replicates BGP primitives over an inter-controller, SDN setup



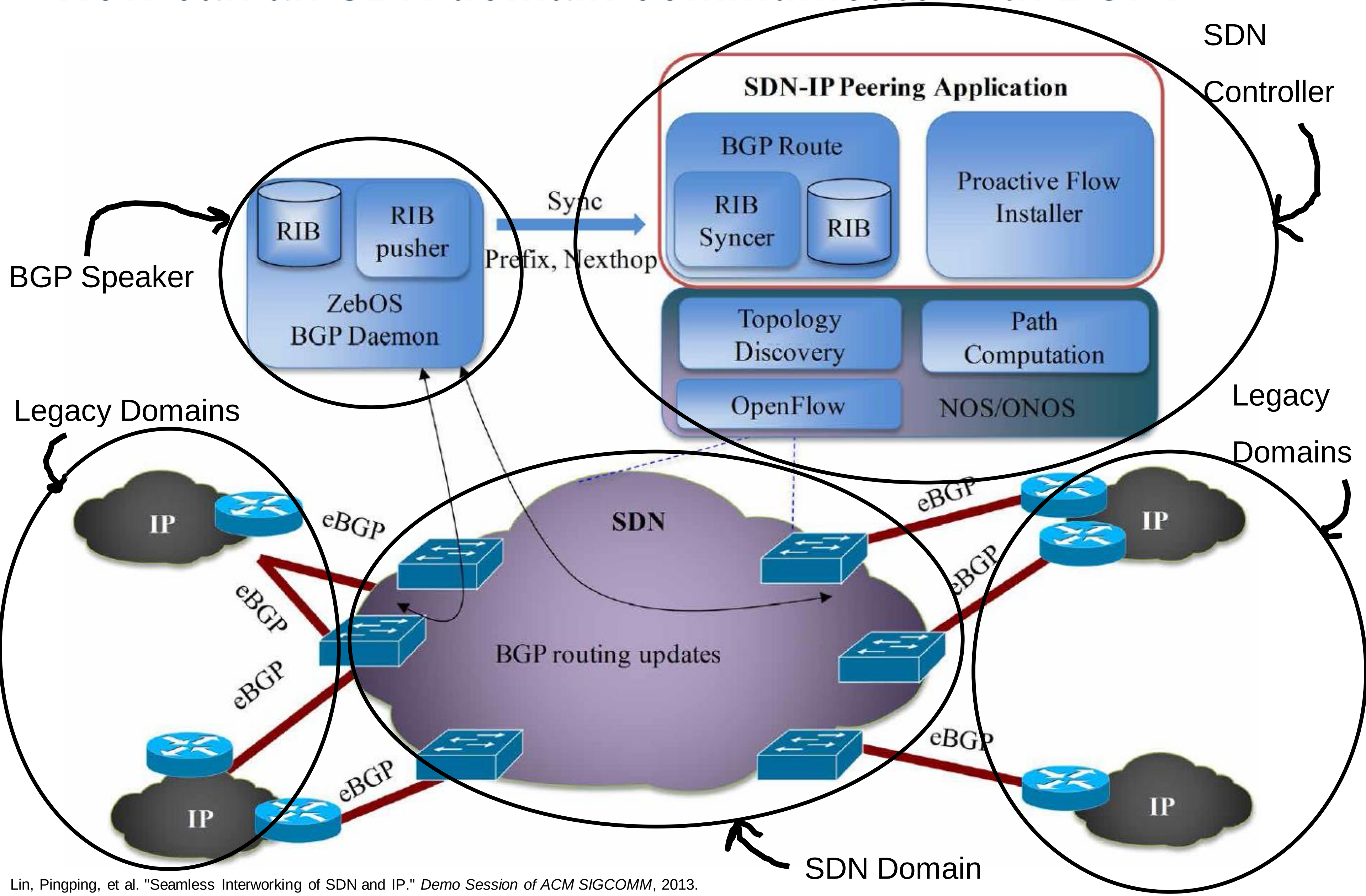
Bennesby, Ricardo, et al. "Innovating on Interdomain Routing with an Inter-SDN Component." *Advanced Information Networking and Applications (AINA), 2014 IEEE 28th International Conference on*. IEEE, 2014.

Example: Steps for inter-domain routing decision



Bennesby, Ricardo, et al. "Innovating on Interdomain Routing with an Inter-SDN Component." *Advanced Information Networking and Applications (AINA), 2014 IEEE 28th International Conference on*. IEEE, 2014.

How can an SDN domain communicate with BGP?



Lin, Pingping, et al. "Seamless Interworking of SDN and IP." *Demo Session of ACM SIGCOMM*, 2013.

Outsourcing the Routing Control Logic: Better Internet Routing Based on SDN Principles

Vasileios Kotronis, Xenofontas Dimitropoulos, and Bernhard Ager. "Outsourcing the Routing Control Logic: Better Internet Routing Based on SDN Principles." In *Proceedings of the 11th ACM Workshop on Hot Topics in Networks (HotNets-XI)*, 2012

Routing management and optimization is complex

Complex research problems

Diverse objectives

- BGP policies
- Over-the-top service guarantees
- SLAs with client networks
- Peering agreements
- Transit cost reduction
- Green TE
- Scalability
- Security
- Etc…

Complex multi-objective optimization

ideal routing

Compile into low-level configuration

Also: we are stuck with BGP

- Has kept the Internet working for decades
- But it is (almost) the same as decades ago
- Well-known technical drawbacks
  - Poor security, adoption of RPKI very slow
    → several prefix hijacking incidents
  - Slow convergence times
    → 30% of the packet loss is due to BGP
  - Policy disputes
  - No support for end-to-end circuits
  - No support for DoS attack mitigation

- It is very difficult to evolve → ossification

Routing and Shifts in the ISP industry

- Profits in pure transit drop ($/Mbps)
- Traffic increases, but so does management complexity
  - Increased load from Content Providers, CDNs
  - Do ISPs have incentives to upgrade their carrier networks for free?
  - Who should pay for the network and its management complexity?
- Bit pipe ISP model under heavy revision
- Pressure for reduced operational costs (OPEX)

➔ Focus on higher-margin services

- IPTV, VoIP, cloud-hosting (remember the "Cloud" lecture)

➔ Exploration of different financial paradigms

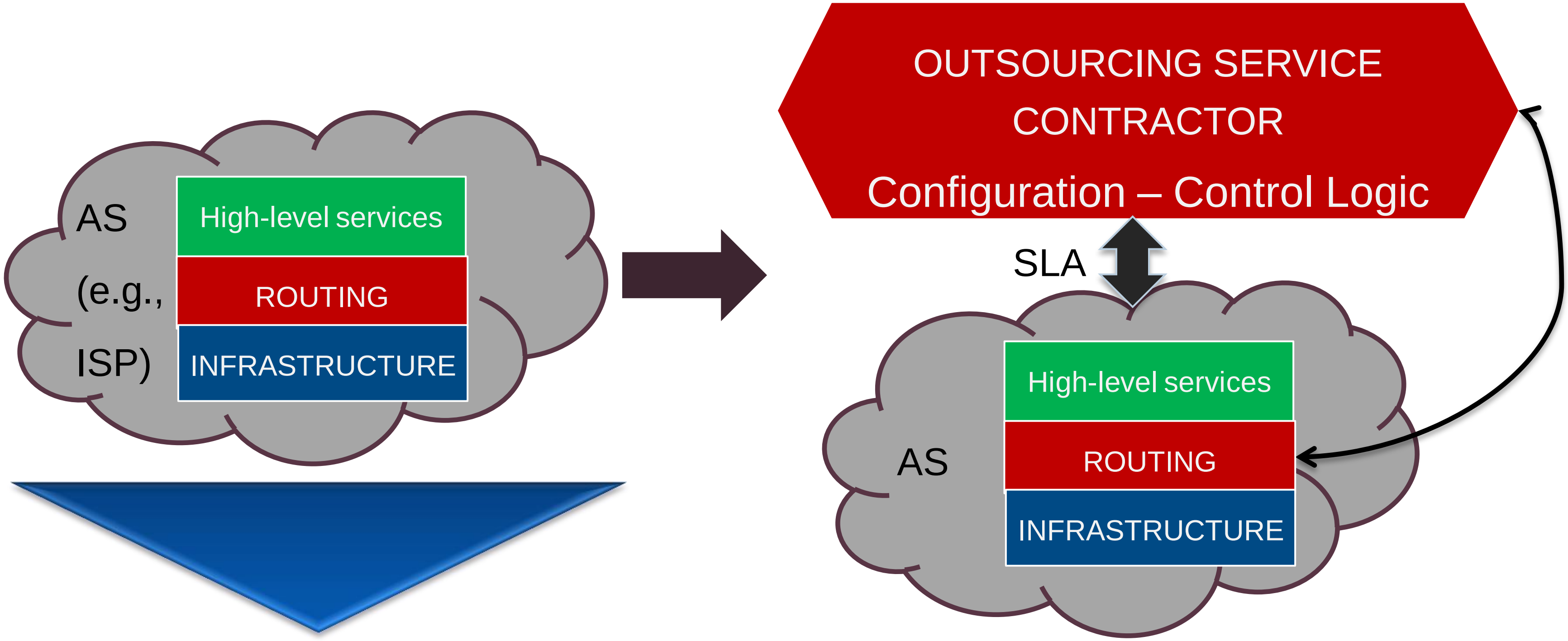
The case for Outsourcing

- Well-known practice to reduce-streamline OPEX
  - Benefits from economy of scale
  - Ecosystem of managed networking services, e.g., IBM outsources network management to AT&T

- Outsourcing makes sense for Internet routing:
  - Internet routing and optimization is hard
  - Gets harder as the service requirements grow
  - Large effort – Small payoff
  - Complexity hinders sophisticated routing

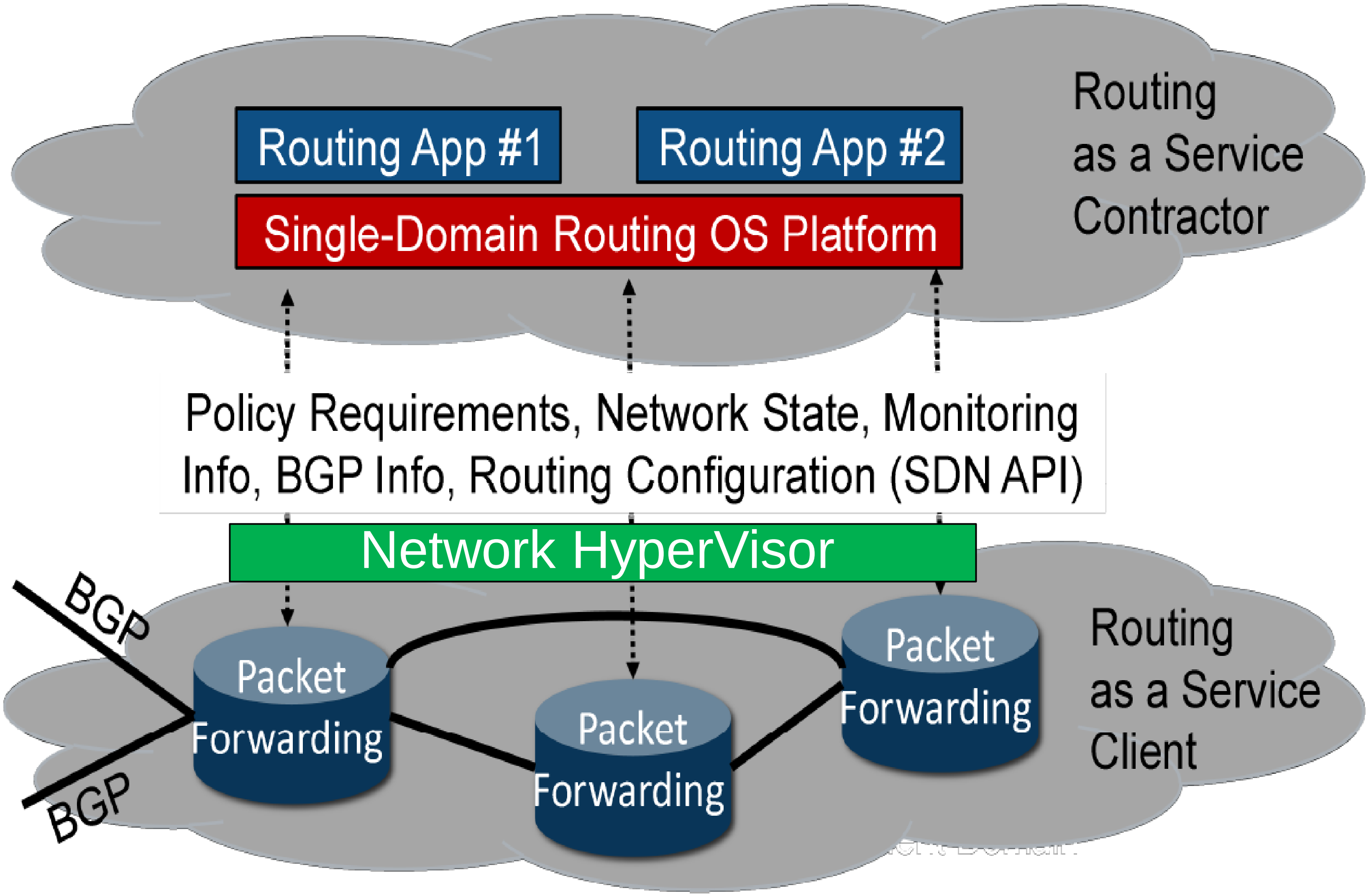
- Idea: Routing Logic Outsourcing

Outsourcing the Routing Logic



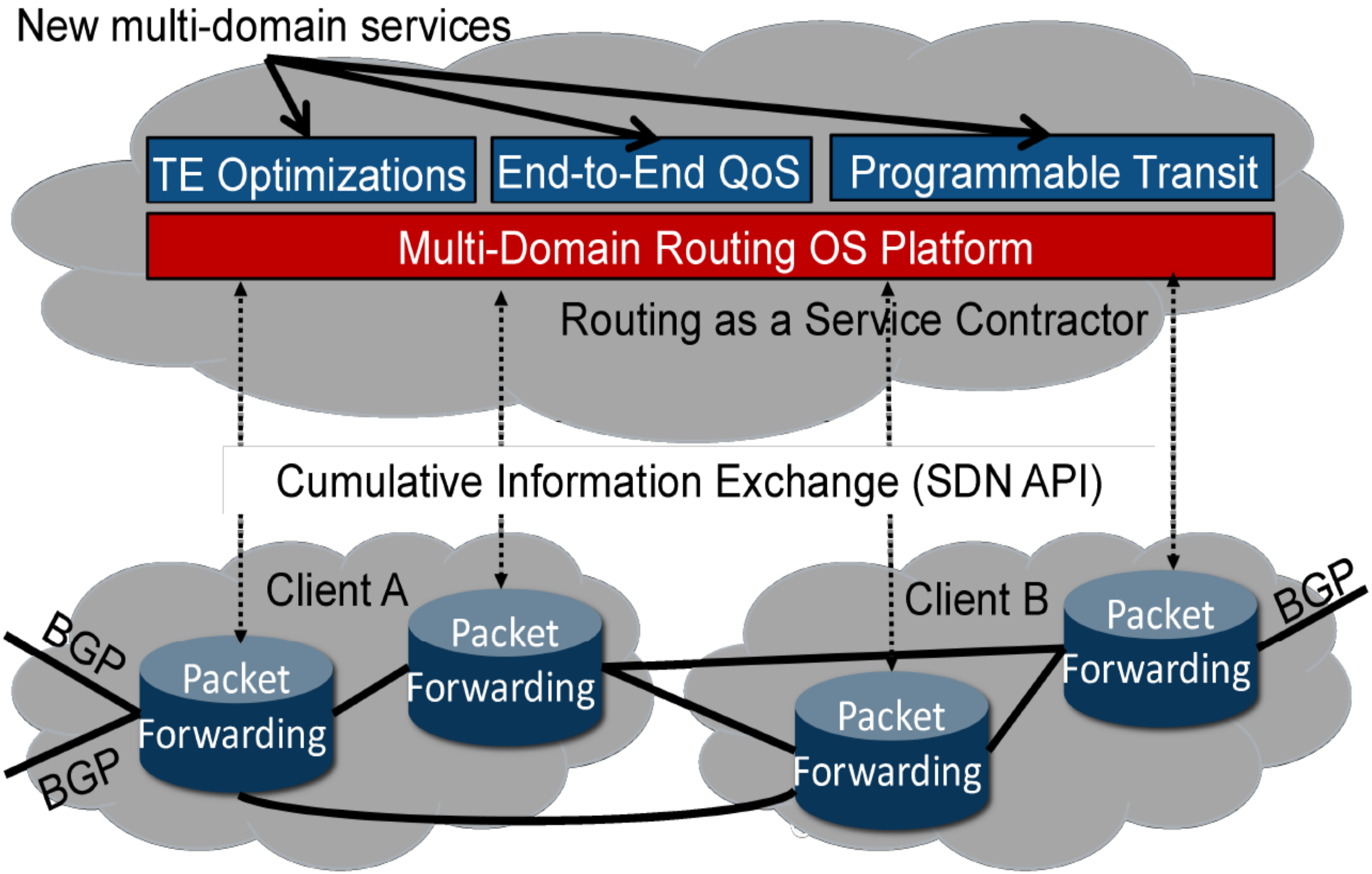
- Focus on profitable services on top of routing
- Buy expertise from specialized contractor
- Form interactive business relationship

SDN: simpler outsourcing of per-domain routing



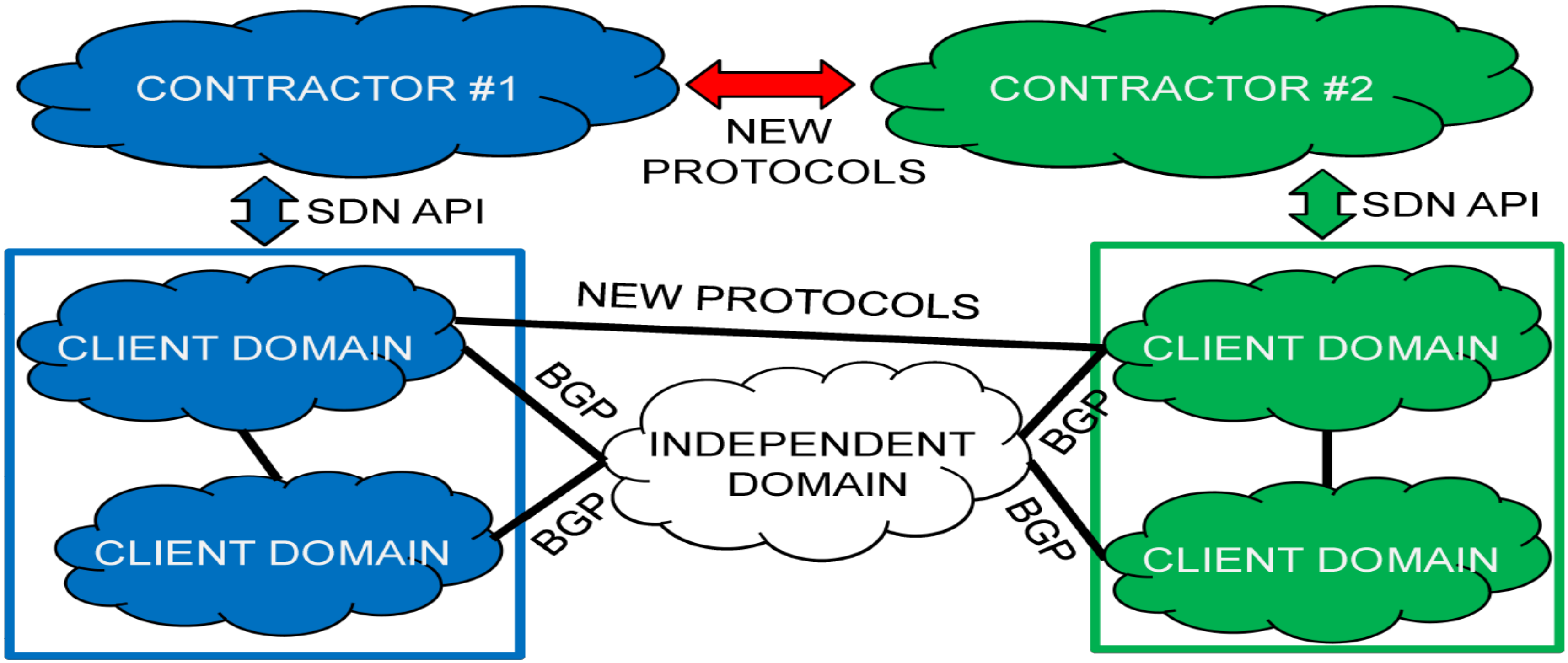
Step 1: Outsourcing the per-domain routing control plane

Thinking bigger: cumulative outsourcing



Step 2: Cumulative routing outsourcing enables new multi-domain services

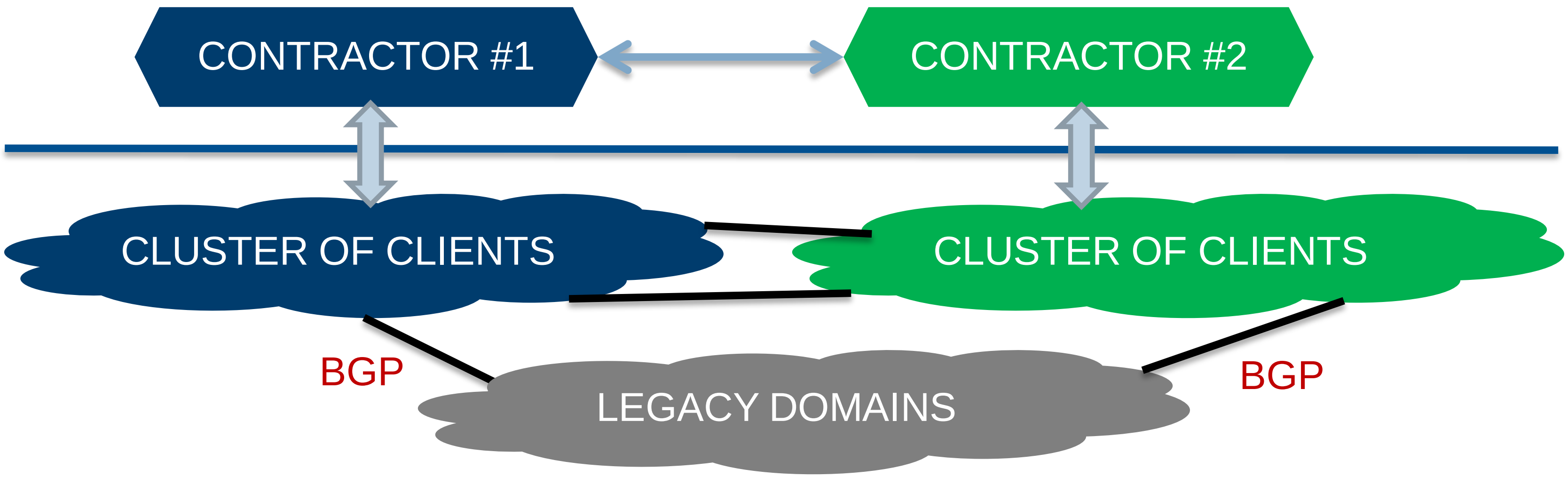
Additional benefit: legacy-compatible evolution



Step 3: Legacy-compatible evolution: multiple contractors, multiple clients

- **Ecosystem** of outsourcing service contractors - clients
- New routing-signaling protocols **within** the clusters
- New protocols for contractor **interoperability**
- **Legacy Compatibility** (BGP)

Recap: Routing Outsourcing



Benefits

- Legacy-compatible inter-domain control plane evolution
- Inter-domain optimizations
- Multi-domain TE
- Economy of Scale

Challenges

- Resiliency/scalability of multi-domain routing control platform
- Evaluation of viability of routing outsourcing business model
- Incentive-based optimizations

Control Exchange Points: Providing QoS-enabled End-to-End Services via SDN-based Inter-domain Routing Orchestration

Control Exchange Points: Providing QoS-enabled End-to-End Services via SDN-based Inter-domain Routing Orchestration

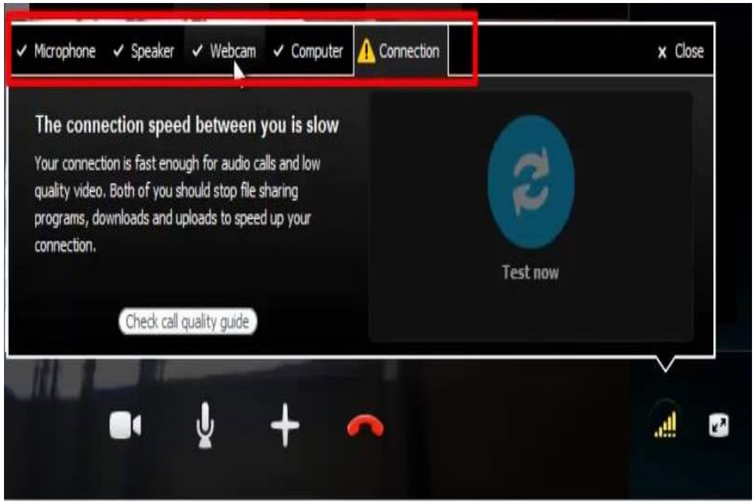
Vasileios Kotronis, Xenofontas Dimitropoulos, Rowan Klöti, Bernhard Ager, Panagiotis Georgopoulos and Stefan Schmid

Proceedings of the Research Track of the 3rd Open Networking Summit (ONS),

Santa Clara, CA, USA, March 2014

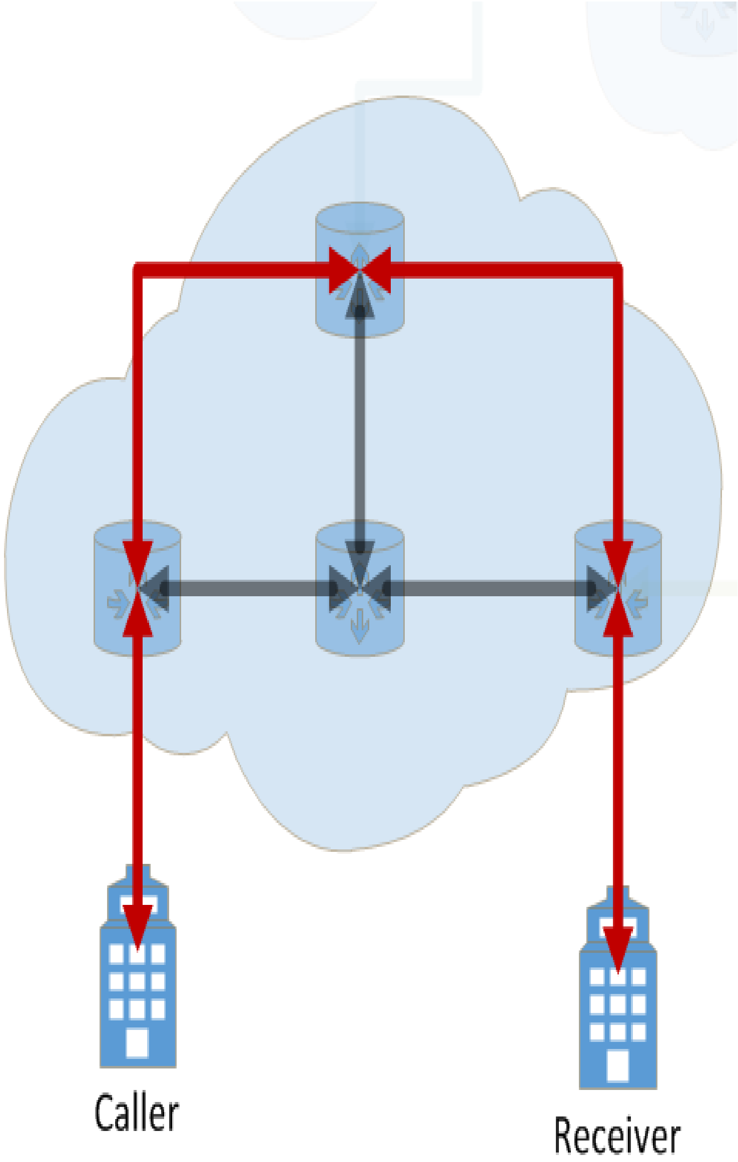
Motivation

- Internet works, but it is not as reliable and performant as we would like
- Besides classic apps, like skype video calls, we expect:
  - Telemusic
  - Telesurgery
  - Remote Control of Critical Infrastructure (e.g., energy plants)



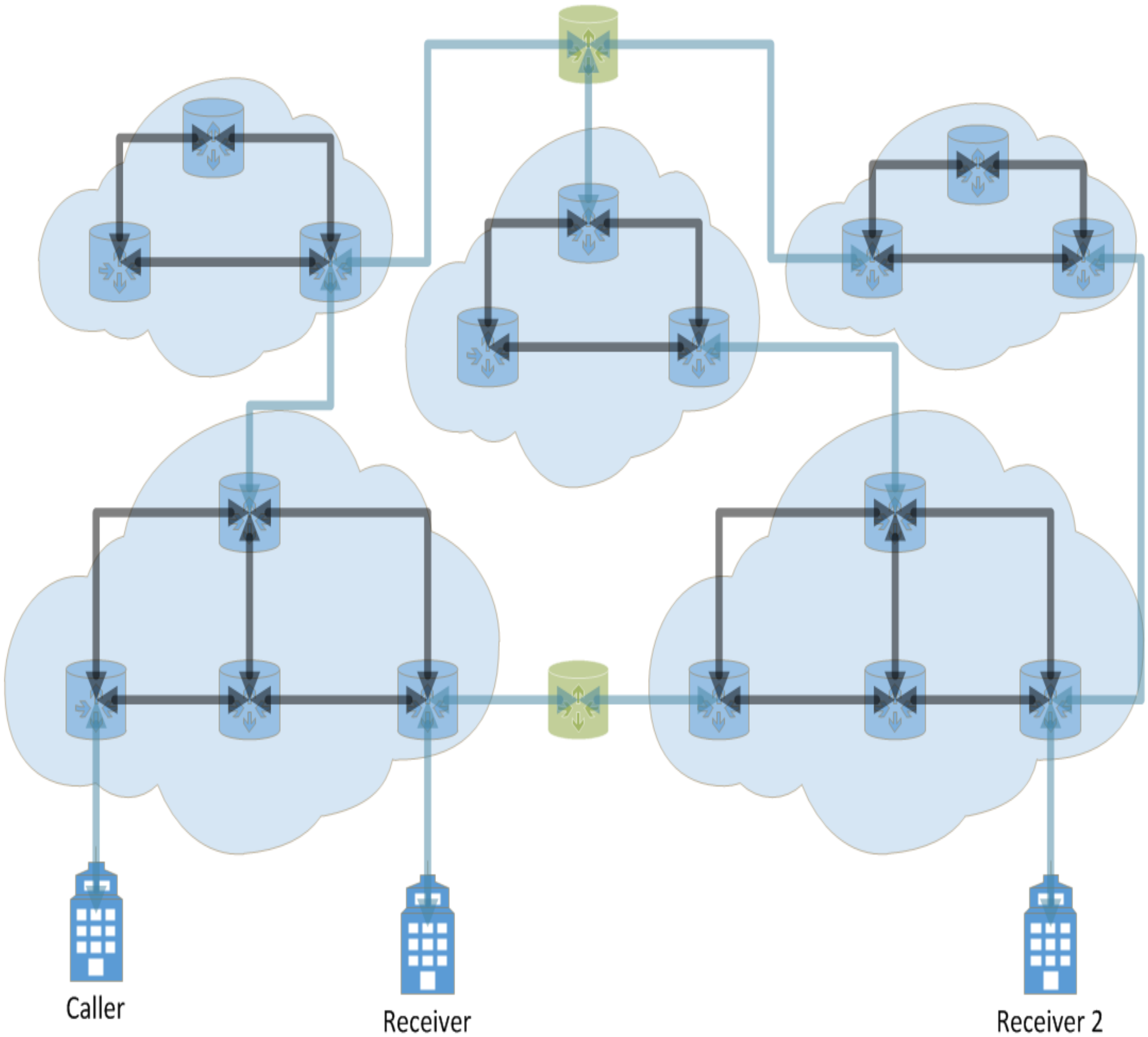
➔Question: Can today's Internet support those services properly?

Possible for a single provider



- An ISP can allocate resources within their domain
- ISP has full overview of link utilization
- ISP controls the embedding of its traffic matrix
- Common practice = dedicated lines with guaranteed:
  - Max latency
  - Min bandwidth
  - Max jitter
  - ...

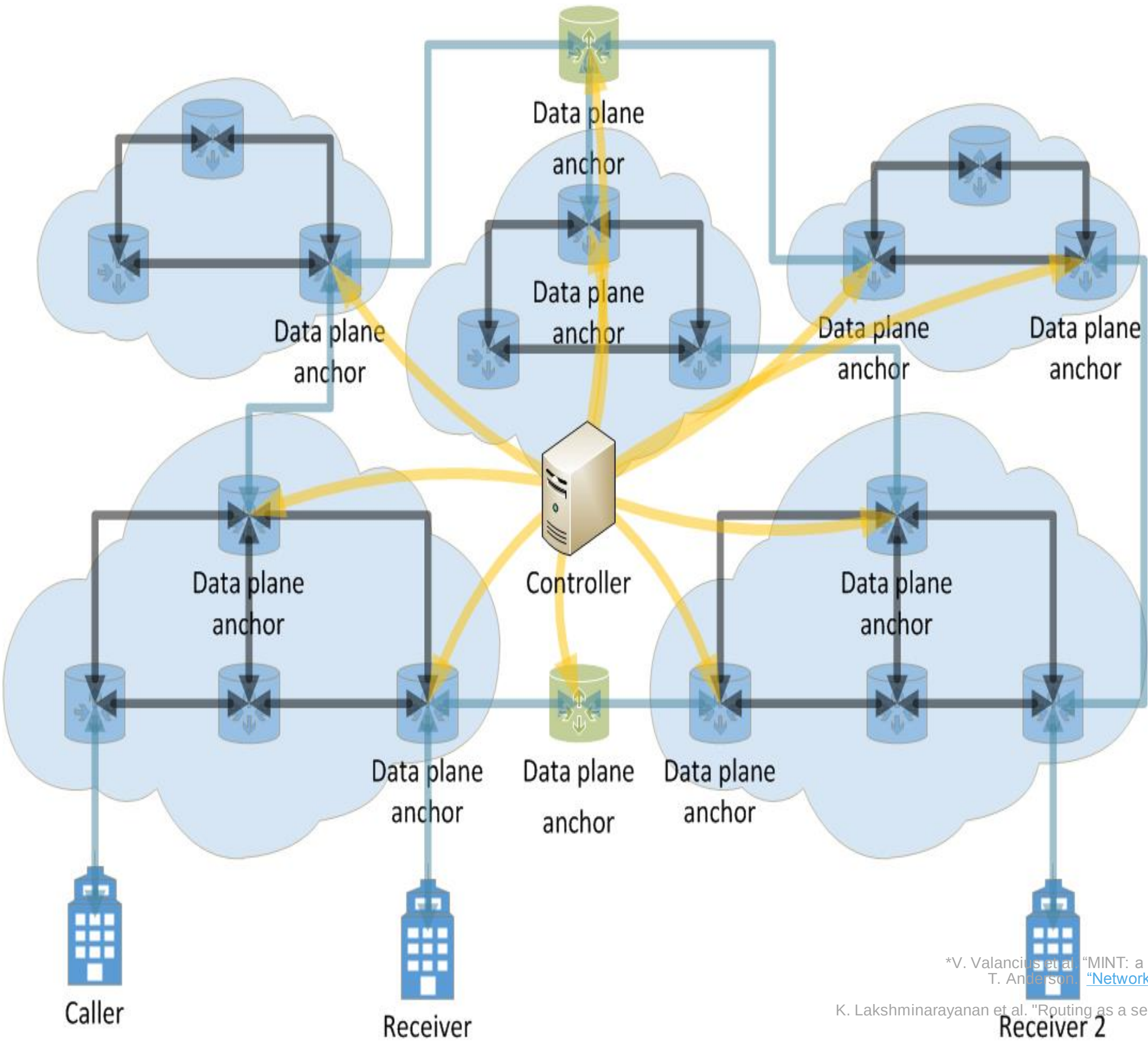
What if the endpoints have different providers?



Inter-domain routing limits us

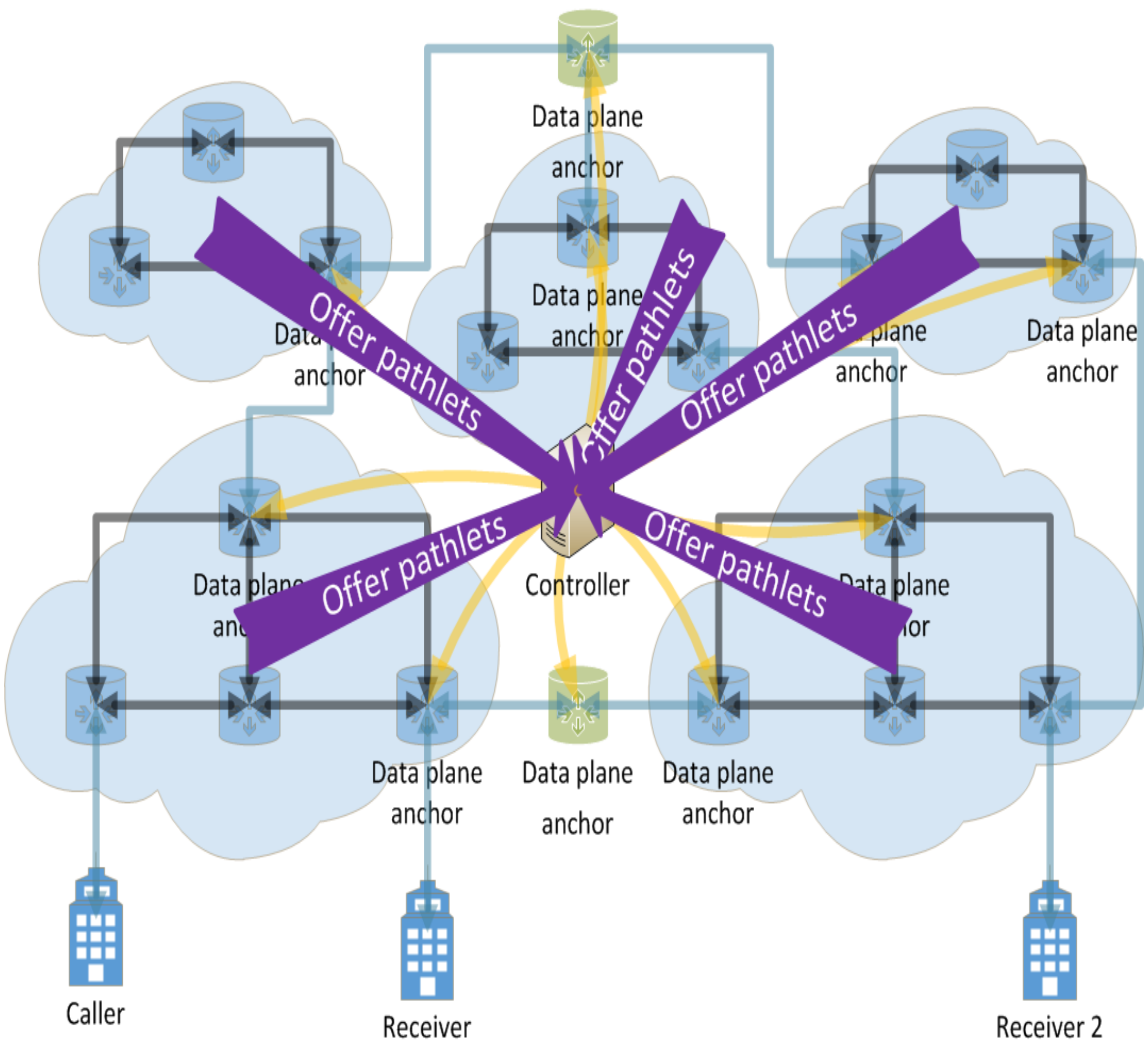
- No end-to-end guarantees for:
  - Availability
  - Latency
  - Bandwidth
  - ...
- Current inter-domain routing does not allow this
  - BGP focuses on *reachability*, not *QoS guarantees*
  - We can’t replace BGP (easily)

We propose Control Exchange Points (CXPs)

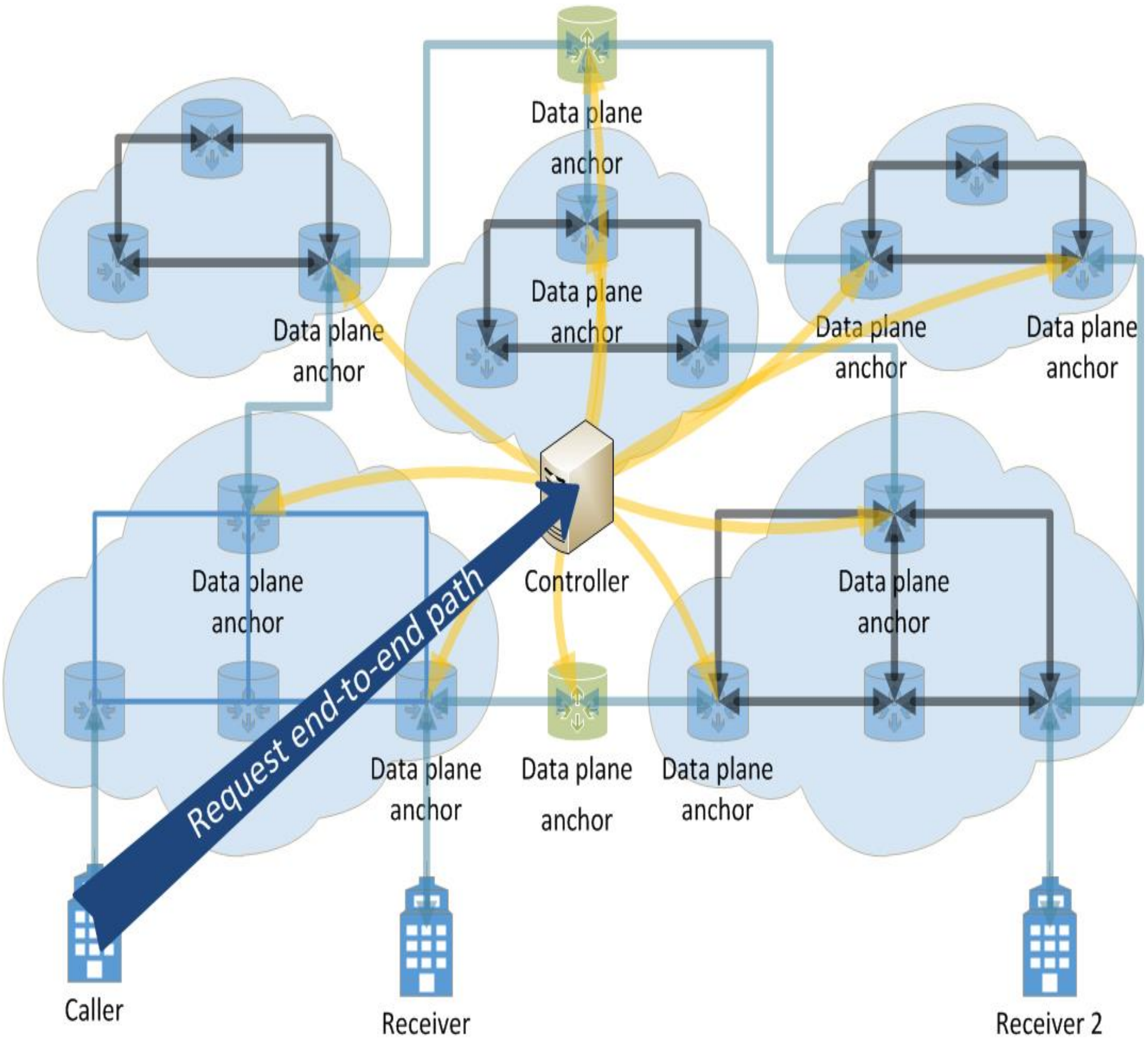


*V. Valancius et al. "MINT: a Market for INternet Transit", In *CoNEXT '08*
T. Anderson. "Networking as a Service", *HOTI-21 keynote. (2013)*
K. Lakshminarayanan et al. "Routing as a service". Tech. Rep. UCB-CS-04-1327 (2004)

ISPs announce pathlets

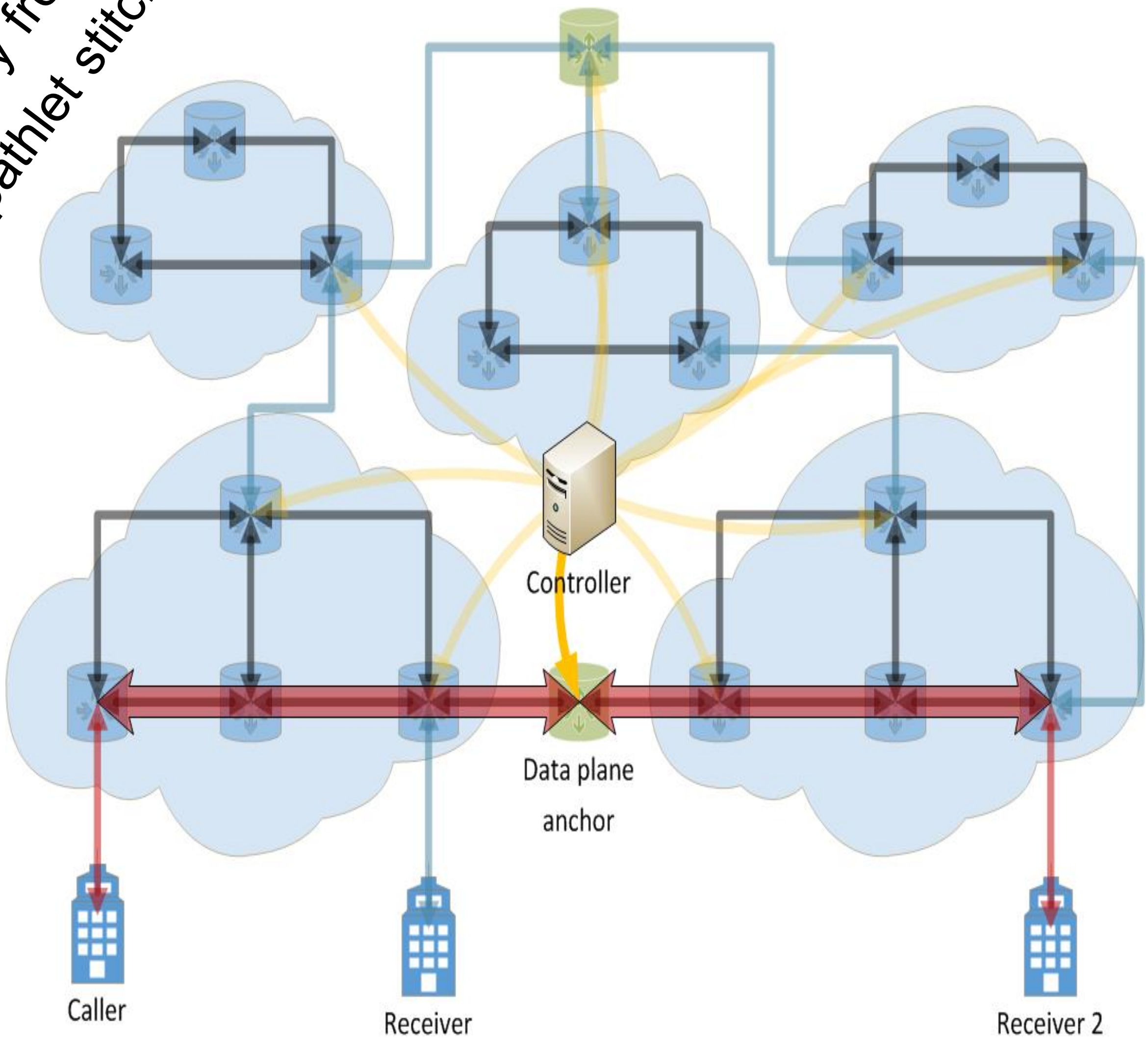


User requests end-to-end path

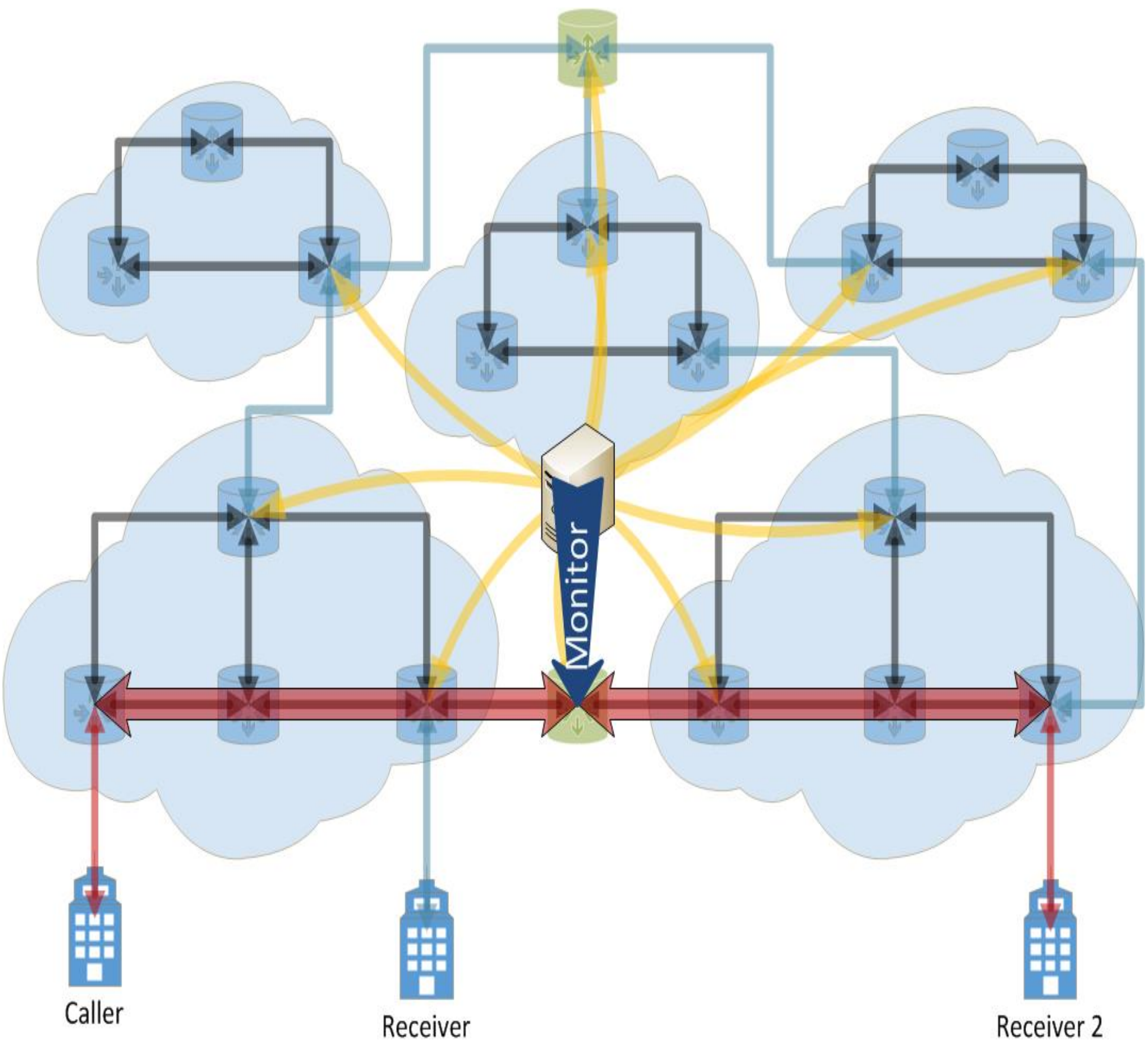


Controller stitches pathlets together

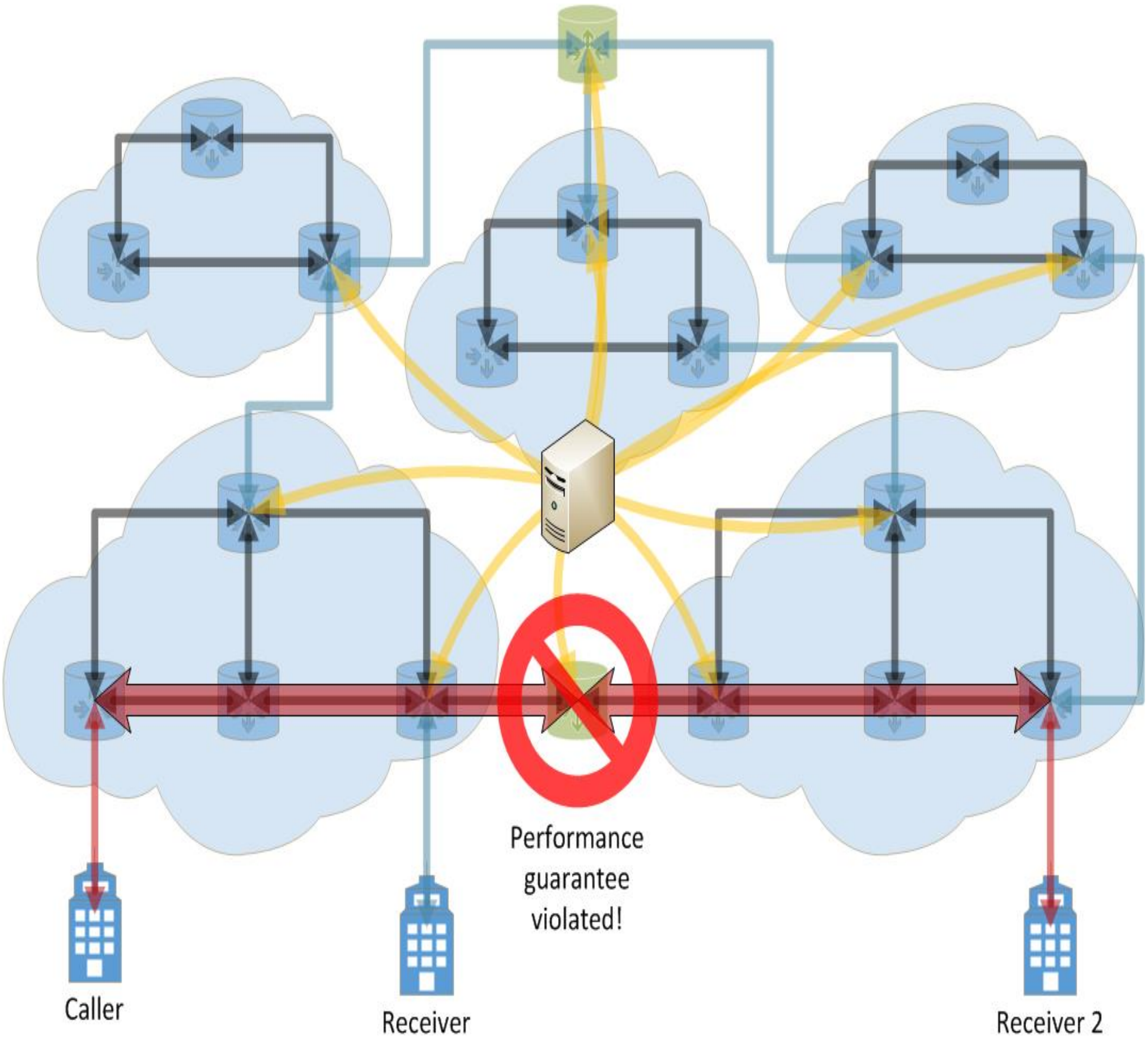
This works independently from BGP!
(end-to-end tunnel/pathlet stitching)



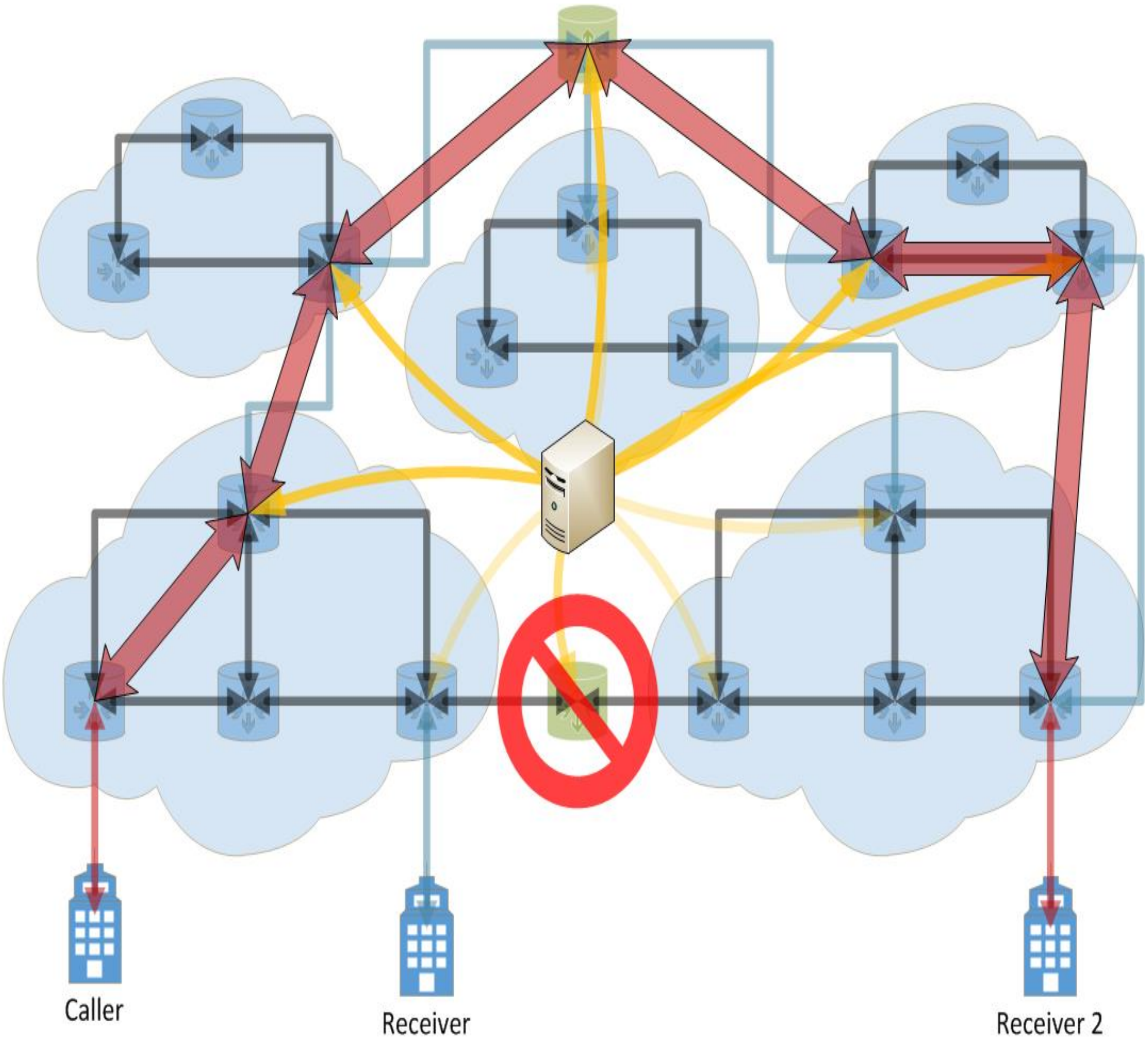
Controller monitors guarantees



Controller detects guarantee violations



Controller chooses alternative route



Best location for CXP data plane anchors?

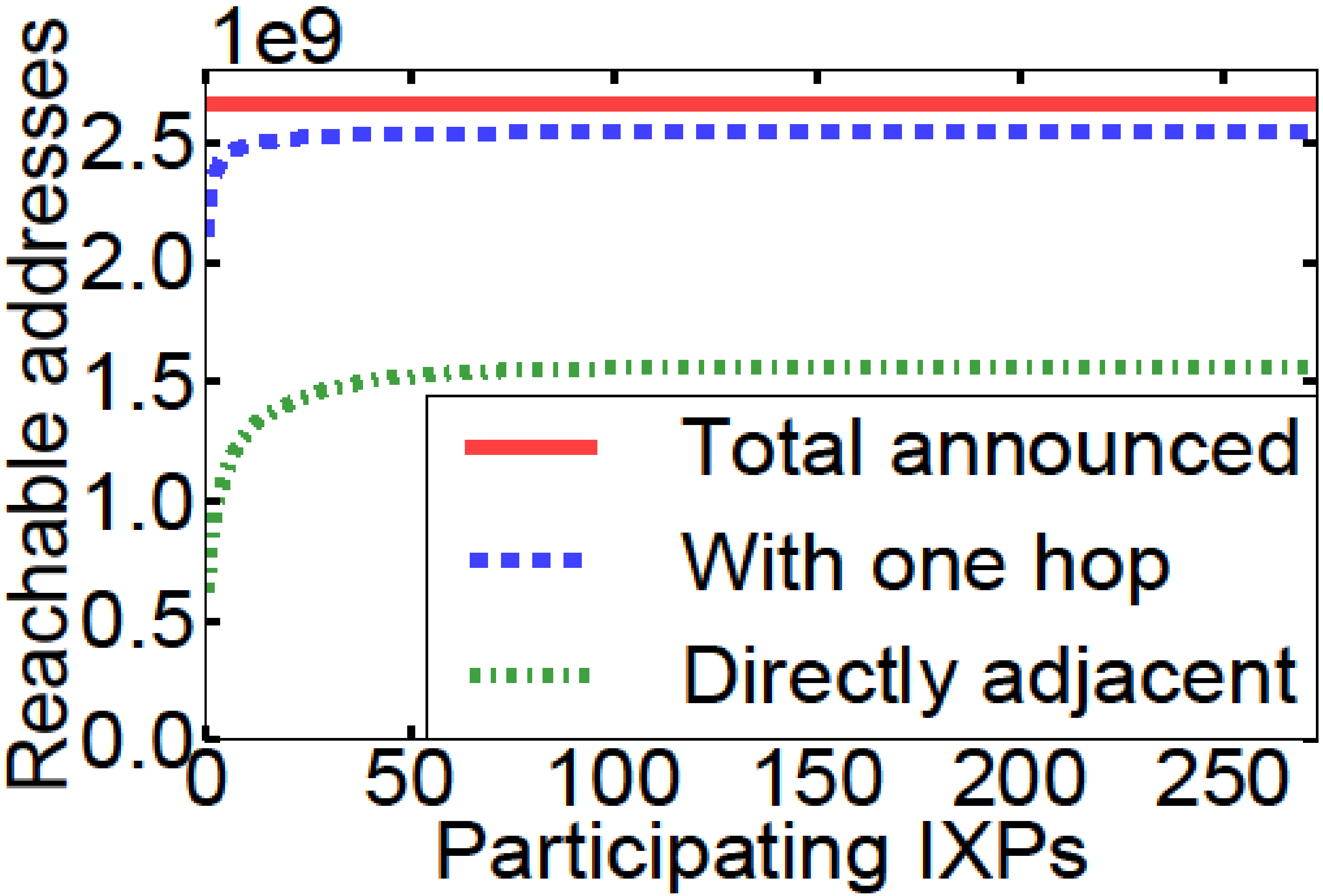
- Good path diversity
- Maximal coverage of potential users
- Well-connected deployments
- High bandwidth and availability
- Provider neutrality

IXPs have the desired properties!

- Internet Exchange Points are public peering points
- They can have hundreds of providers participating
- They exchange up to Tbps of traffic
- They are independent of individual members
- Ideal locations for having impact on inter-domain
  - (See SDX discussion later)

- But what about path diversity? Coverage?

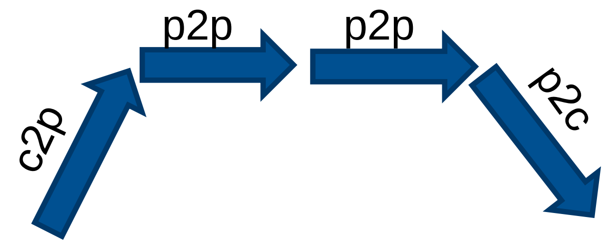
IP address coverage



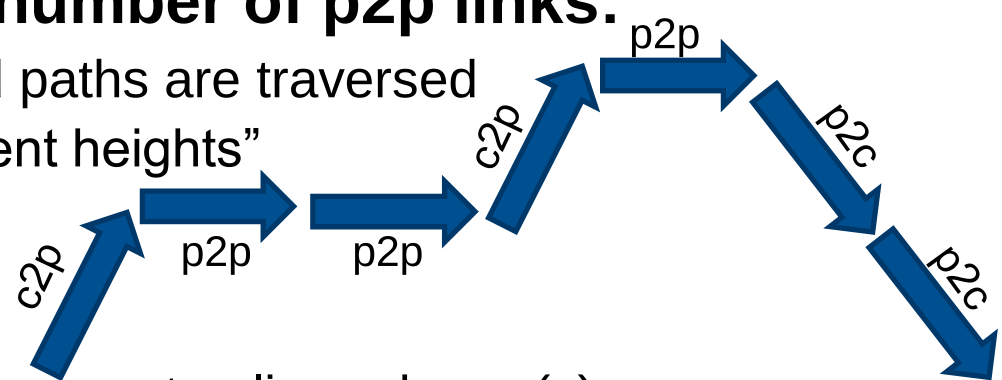
Based on Euro-IX dataset at https://euro-ix.net/ , snapshot of 9/4/2014

Rethinking path selection with CXPs for diversity

- **Scenario 1→ valley-free routing by allowing:**
  - arbitrary p2p hops between the uphill and the downhill path
  - one CXP-mediated path traversed
  - “mountain with wide peak”
  - pure valley-free allows a narrow peak

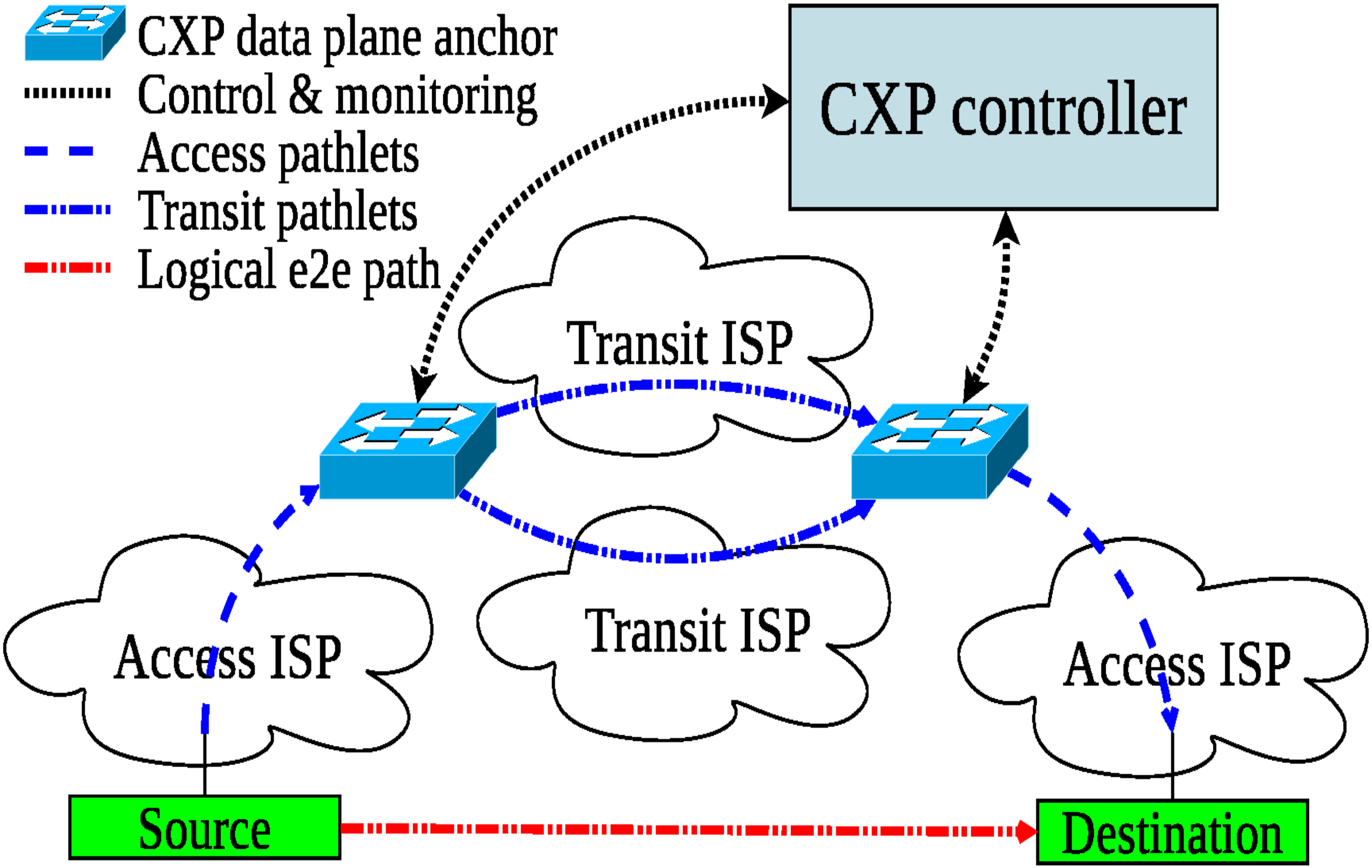


- **Scenario 2 → unlimited number of p2p links:**
  - any number of CXP-mediated paths are traversed
  - “mountain with steps at different heights”



- **Other Scenarios:**
  - freedom/incentives to extend current policy scheme(s)
  - exploit Internet’s path diversity, under policy compliance
  - experiments: suggested gains up to one order of magnitude

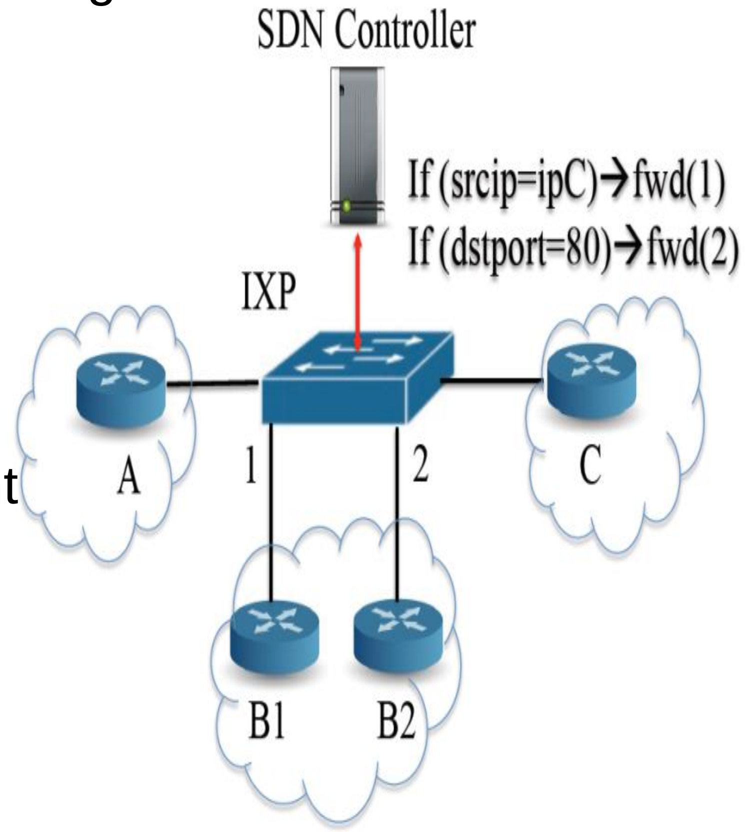
Recap of the CXP concept



SDX: a Software Defined Internet Exchange

- SDN in inter-domain routing: challenges of state-of-the-art
  - Very little done, high costs, high risk
  - Hard to deploy new solutions, hard to change BGP
  - Routing only on destination prefix
  - Influence only over direct neighbors
  - Indirect expression of policy

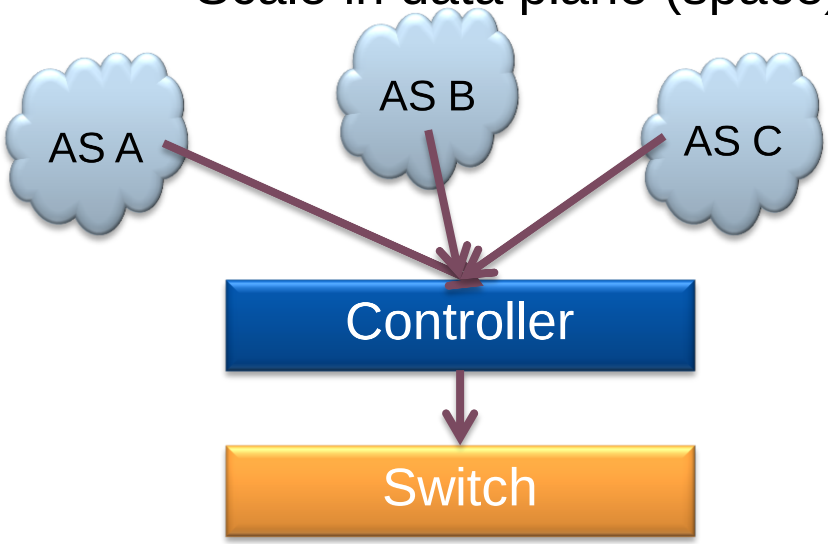
- Target IXPs for initial deployment
  - Offer structural advantage
  - Interoperate with current IXP equipment
  - Interoperate with BGP
  - Offload cumbersome BGP tasks to SDN controller



Gupta, Arpit, et al. "SDX: A Software Defined Internet Exchange.", ACM SIGCOMM 2014

SDX: SDN at IXPs

- **Opportunities:** Freedom from current BGP constraints
  - Matching in different packet header fields-flow space
  - Control messages from remote networks (e.g., content providers)
  - Direct control over data plane
- **Challenges:** no SDN control framework for inter-domain
  - Main issue: scaling to 100s-1000s ISPs present at an IXP
  - Scale in data plane (space) and control plane (time)



Inputs:

1. BGP Routes per IP prefix from all ASes at the exchange (including attributes)
2. Selection function (for routing and/or rewriting packets) based on participant policies

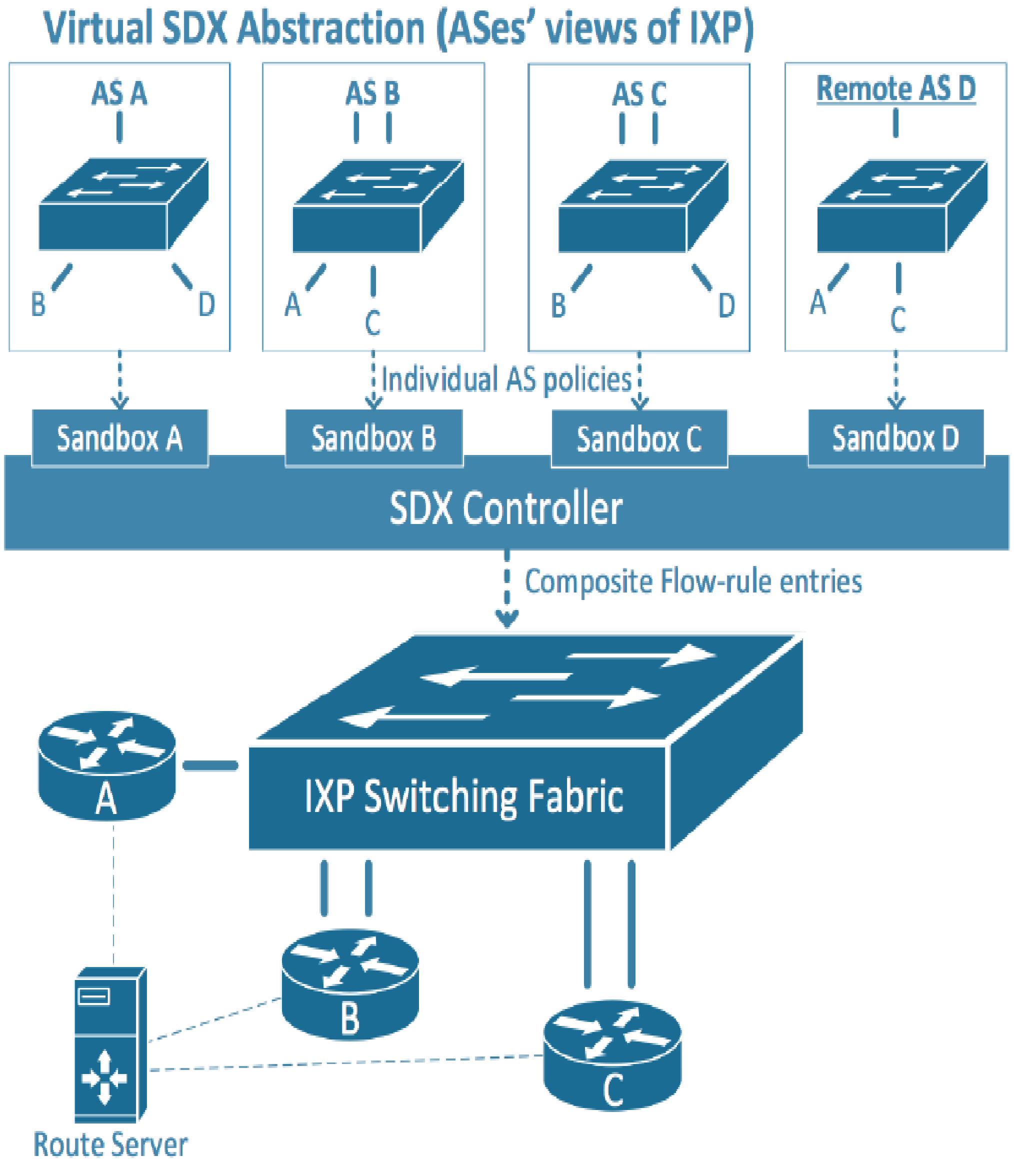
Outputs:

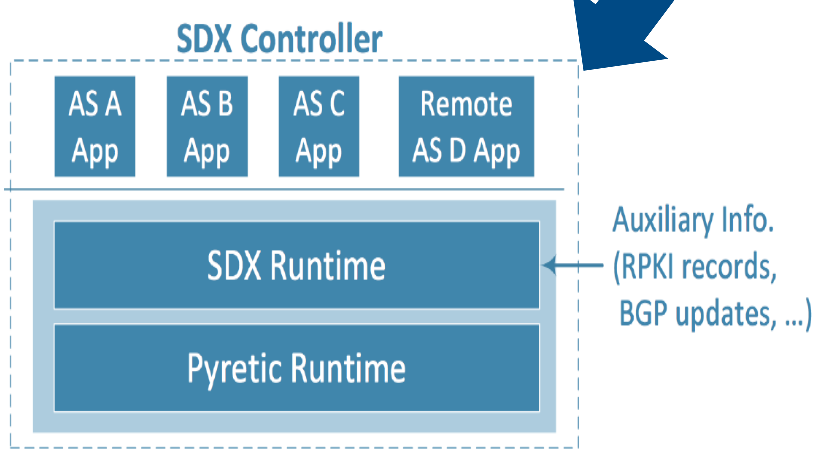
1. FIB entries in switch
2. Packet rewriting actions

Gupta, Arpit, et al. "SDX: A Software Defined Internet Exchange.", ACM SIGCOMM 2014

SDX: Virtualizing the IXP

- Each AS sees only its own virtual topology (isolation)
- ISPs that do not have business relationships do not see each other
- The SDX controller resolves conflicts using policy composition
- Symbolic execution at SDX
- Pyretic programming language/abstractions





Gupta, Arpit, et al. "SDX: A Software Defined Internet Exchange.", ACM SIGCOMM 2014

SDX: potential applications

- Traffic Offloading
- Inbound TE
- Application-specific Peering
- WAN load balancing
- Redirection to Middleboxes
- Fast convergence
- Prevent free-riding
- Upstream DoS blocking
- IXP fabric virtualization
- Policy validation (with RPKI cooperation)

Gupta, Arpit, et al. "SDX: A Software Defined Internet Exchange.", ACM SIGCOMM 2014

PART II

SIREN: a hybrid SDN Inter-domain Routing EmulatioN framework

Vasileios Kotronis

Collaborators: Adrian Gämperli, Fontas Dimitropoulos

Motivation

- BGP routing has problems
  - Can take several minutes to converge
  - Does not support end-to-end circuits
  - Is complex to manage
  - Is *very* difficult to change and evolve
- Can SDN help improve BGP?
  - SDN centralization on the inter-domain level [1]
  - Communication between SDN domain controllers [2]
  - Software Defined Internet Exchanges (SDX) [3]

➔ Need for hybrid BGP-SDN emulator to test new research ideas!

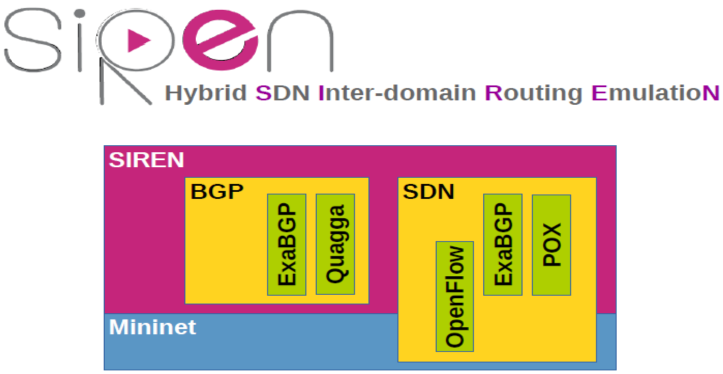
[1] Kotronis, V. et al. “Outsourcing the Routing Control Logic: Better Internet Routing Based on SDN Principles”. *In Proc. of ACM HotNets-XI*, 2012.

[2] Bennesby, Ricardo, et al. "Innovating on Interdomain Routing with an Inter-SDN Component“. *Advanced Information Networking and Applications (AINA), 2014 IEEE 28th International Conference on*, 2014.

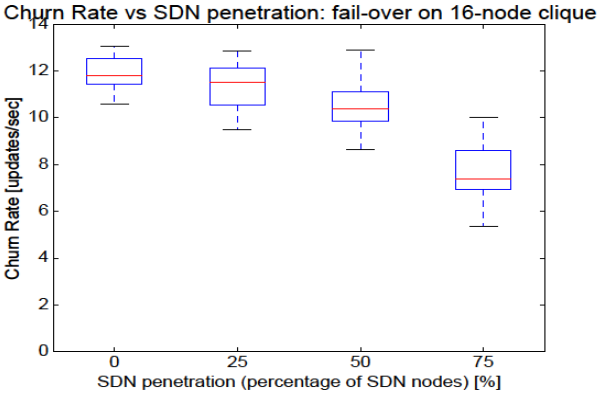
[3] Gupta, A., et al. “SDX: A Software Defined Internet Exchange”. *In Proc. of ACM SIGCOMM*, 2014.

Objectives

- **Develop hybrid BGP-SDN emulation framework**
  - Emulate multiple AS
  - Use real router software
  - Enable BGP-SDN interactions
  - Simplify experiment management
  - Easily visualize results



- **Evaluate effect of gradual SDN centralization on BGP convergence as a use case [1]**
  - Design centralized multi-AS controller
  - Implement using SDN mechanisms
  - Run over emulation framework



[1] Adrian Gämperli, Vasileios Kotronis and Xenofontas Dimitropoulos . "Evaluating the Effect of Centralization on Routing Convergence on a Hybrid BGP-SDN Emulation Framework". *Demo session of ACM SIGCOMM*, Chicago, Illinois, USA, August 2014.

SIREN features

- Based on Mininet, Quagga, POX and ExaBGP
  - Mininet : open-source SDN emulation framework
  - Quagga : BGP routing daemon for legacy emulation, BGP speaker
  - POX : open-source SDN OpenFlow-enabled controller (v1.0)
  - ExaBGP: Python-based BGP library and API
- Automated experiment management
  - Simple interface, comprehensive CLI
  - Automatic IP address assignment and configuration
  - Run experiment batches over multiple computing nodes
- Live visualization of routing changes
- Packet loss measurements between end-points
- Log collection and analysis

BGP-specific functions on SIREN

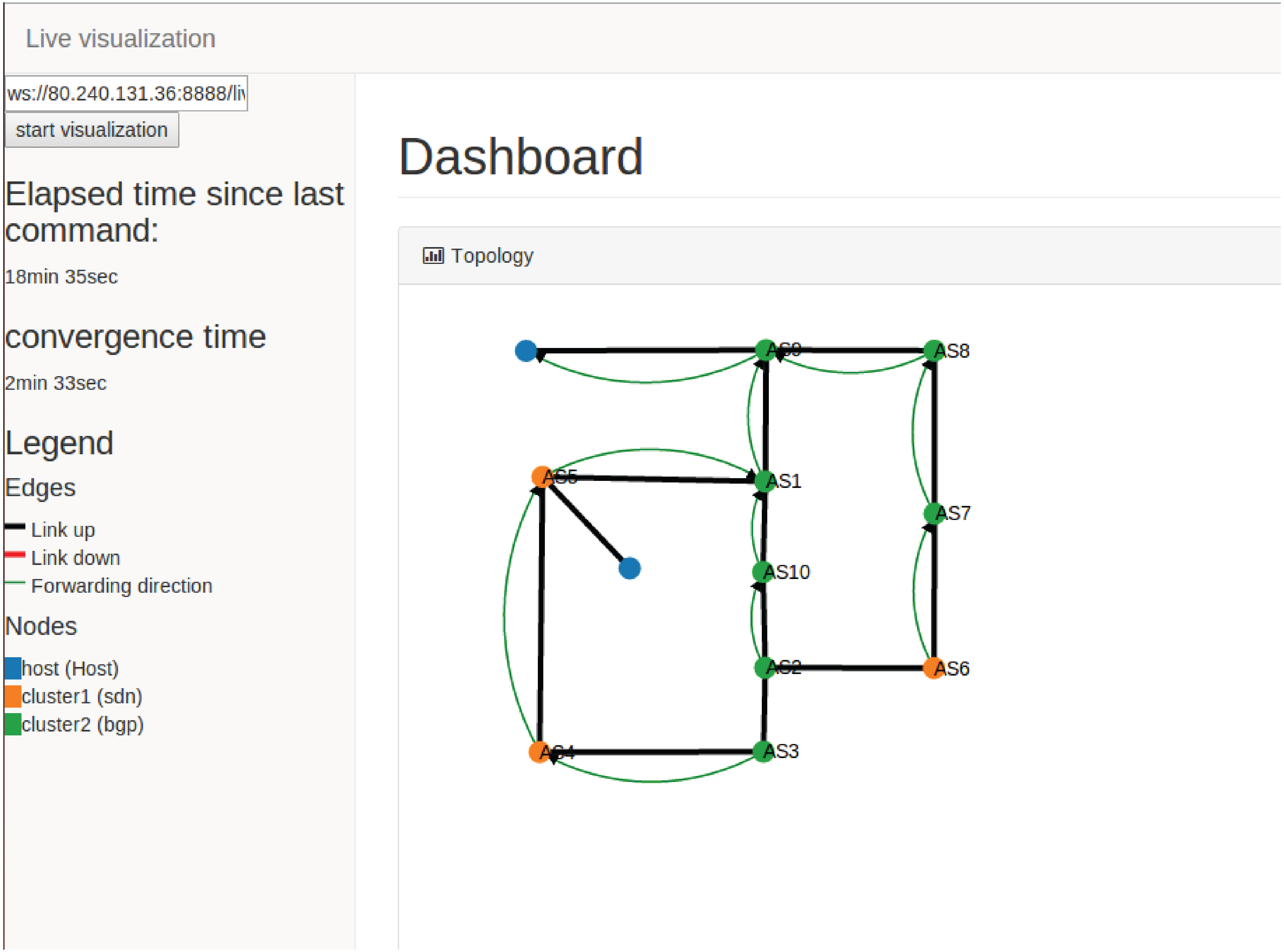
- Announce IPv4 prefixes
- Wait until convergence is successful (stable routing)
- Measure convergence times (from trigger to last update)
- Measure average routing update churn rates
- Detect failed routing setups (no all-to-all connectivity)
- Set valley-free policies on Quagga routers
  - Pre-defined template selection (BGP filters, etc.)
  - Path prepending for inbound TE
  - Disclaimer: this is implemented on Quagga but not SDN (yet)

➔ Suggestions for commands applying to your use case are welcome!

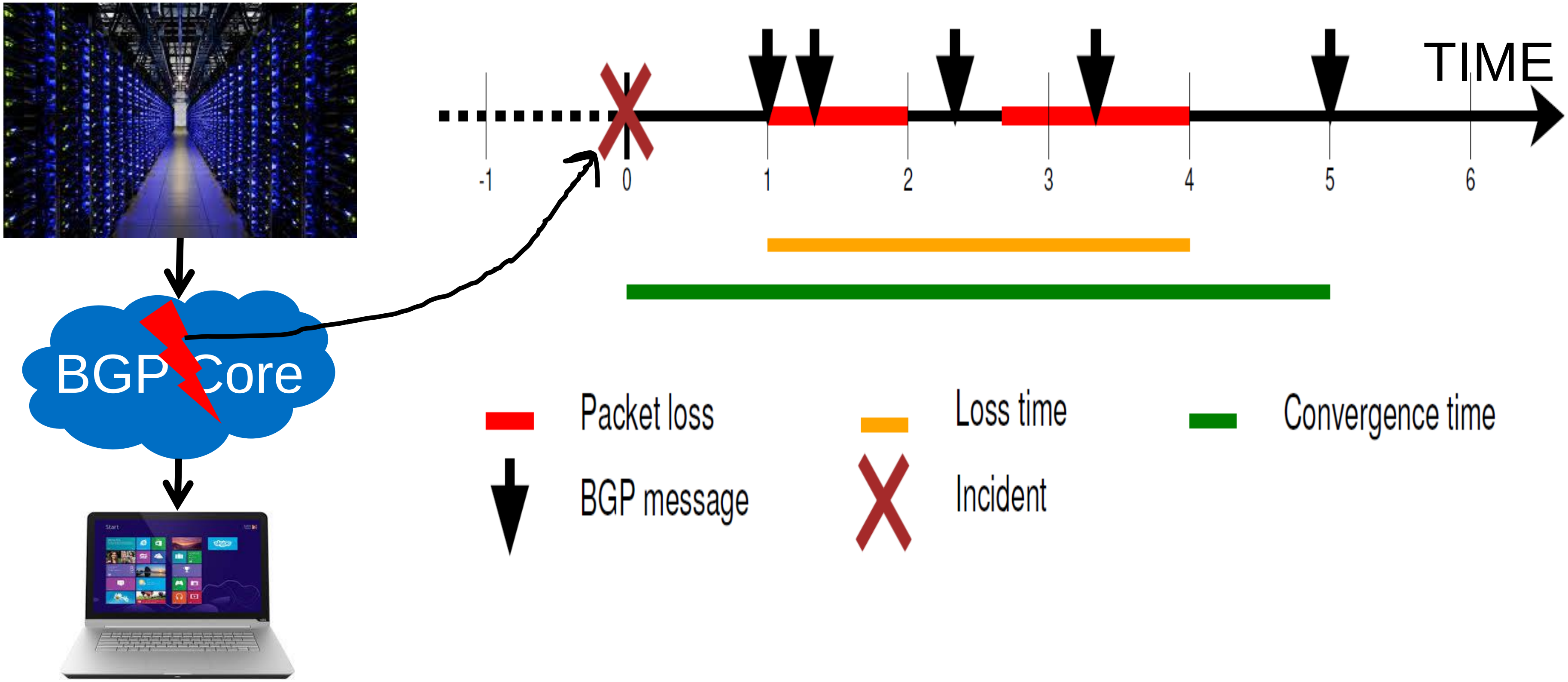
Usage modes

- Command Line Interface (CLI) like in Mininet
  - Classic terminal scripting
  - Write python scripts and invoke them, collect results
- Live visualization
  - Web interface for live interaction with the network
  - Bring links up down with a mouse click, check reactions
  - Depicts routing-related info (e.g., convergence time)
- Experiment manager
  - Distribute batches of experiments to multiple nodes
  - Disclaimer: useful for distributing a batch of small experiments, but not a single large experiment

Example 1: Live Visualization (cf. Demo in the end)



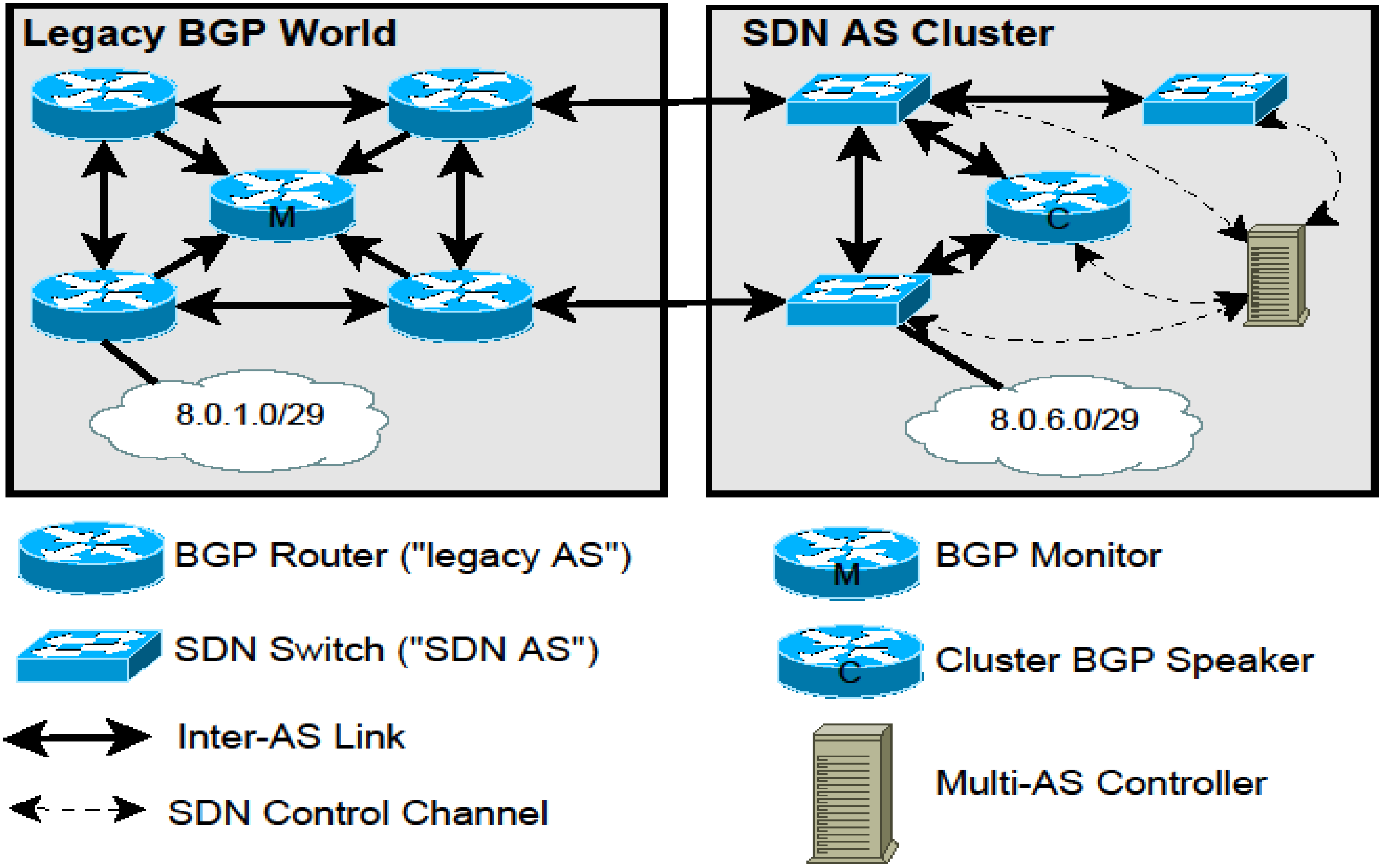
Example 2: Loss and Delays during convergence



- ➔ Sample experiment: Unicast UDP stream between client and server over BGP core
- ➔ Use buffers to hold traffic of X sec (user-defined), measure impact of loss/delay
- ➔ Sample incident: inter-domain link-down event

SIREN and the SDN Routing Centralization Use Case: Testing a “Routing-as-a-Service” multi-AS controller

Sample SIREN Setup



Controller Design Goals

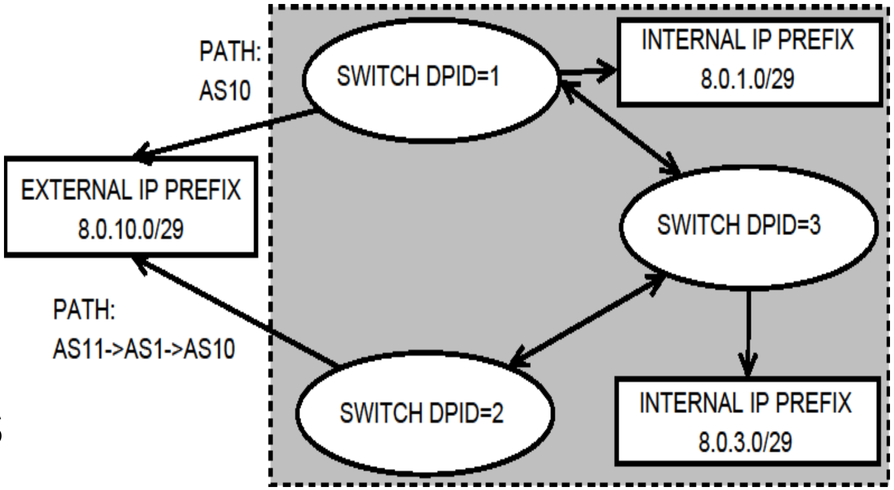
- **Exploit centralization on the AS level**
  - ➔ Help achieve more stable routing overall
- **Interoperation/compatibility with BGP**
- **No cluster lock-in**
  - ➔ Preserve ASNs in route ads
  - ➔ Each AS maintains its identity and policies
  - ➔ Mechanisms like AS-level ACLs, BGP communities do not change
- **Disjoint clusters**
  - ➔ AS paths may enter, exit and reenter the cluster at different points
  - ➔ Calculate paths using global view of cluster and legacy BGP info
- **Hybrid routing: link-state Dijkstra and path-vector BGP**
- **No loops: cannot simply use same mechanism as BGP**

Implemented Controller Features

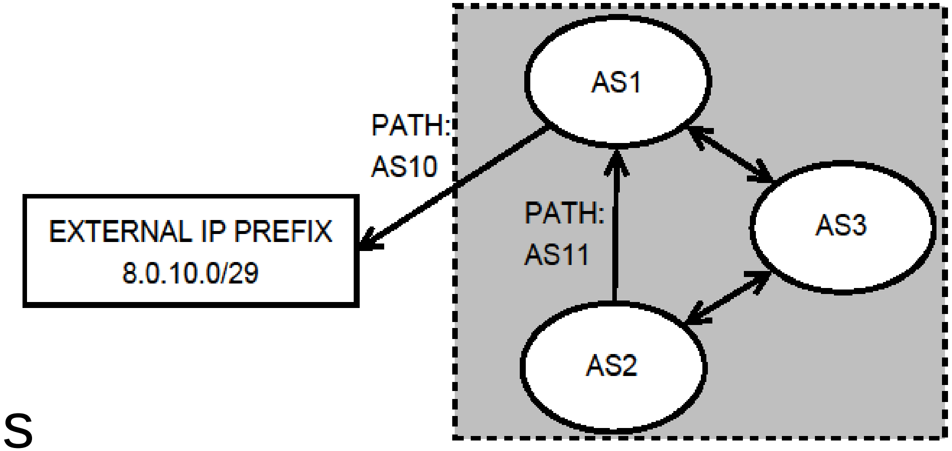
- Centralized controller which speaks:
  - BGP outside of cluster (east-west)
  - OpenFlow within cluster (southbound)
- Transparent to BGP
  - Outside ASes are oblivious to the multi-AS cluster architecture
  - Controller behaves as a BGP router
- Hybrid path-vector / link-state inter-domain routing
  - Shortest path, no AS-level loops
- Delayed Route ads and SDN reconfiguration
  - Cluster Reconfiguration Wait Interval (CRWI)
  - Similar concept with BGP MRAI

Switch and AS graphs: Per-prefix routing and loop avoidance

- Physical topology of cluster
  - Switches
  - Prefixes
  - Inter-switch connections
- BGP paths to external prefixes



- Map switches to ASNs
- AS-level view of the cluster
- Sanitize paths that exit and reenter cluster!
- Avoid loops using virtual links crossing external ASes



The path recomputation problem

- The SDN cluster controller controls inter-domain routing interactions on behalf of several ASes
- Each AS has several external peers

➔ Multiple BGP updates per second to the controller

- Each update → triggers changes in switch and AS graphs
  - Path recomputation throughout the cluster
  - Switch reconfiguration through manipulation of the flow tables
  - Can take 100s of ms (RTT + switch processing delays)
  - Expensive process!
- During reconfiguration → more updates are received
- Plus: controller's actions need to be advertised to external peers → further instability and processing overhead

Delayed path recomputation

- "Cluster Waiting Recomputation Interval" (CRWI)
- After CRWI timeout happens:
  - Compute and install locally rules associated with new paths
  - Directly advertise the changes over BGP to the outside world

- Benefits:
  - Avoid routing inconsistencies with neighbours due to outdated info
  - Make the network more stable by "rate-limiting" the cluster controller
  - Reduce number of required path changes
  - Leave some temporal slack for the forwarding rule installation

➔ In our experiments, we found that a CRWI of 1 sec is sufficient to avoid any problems with routing inconsistencies and flow rule installation delays.

Other implementation details

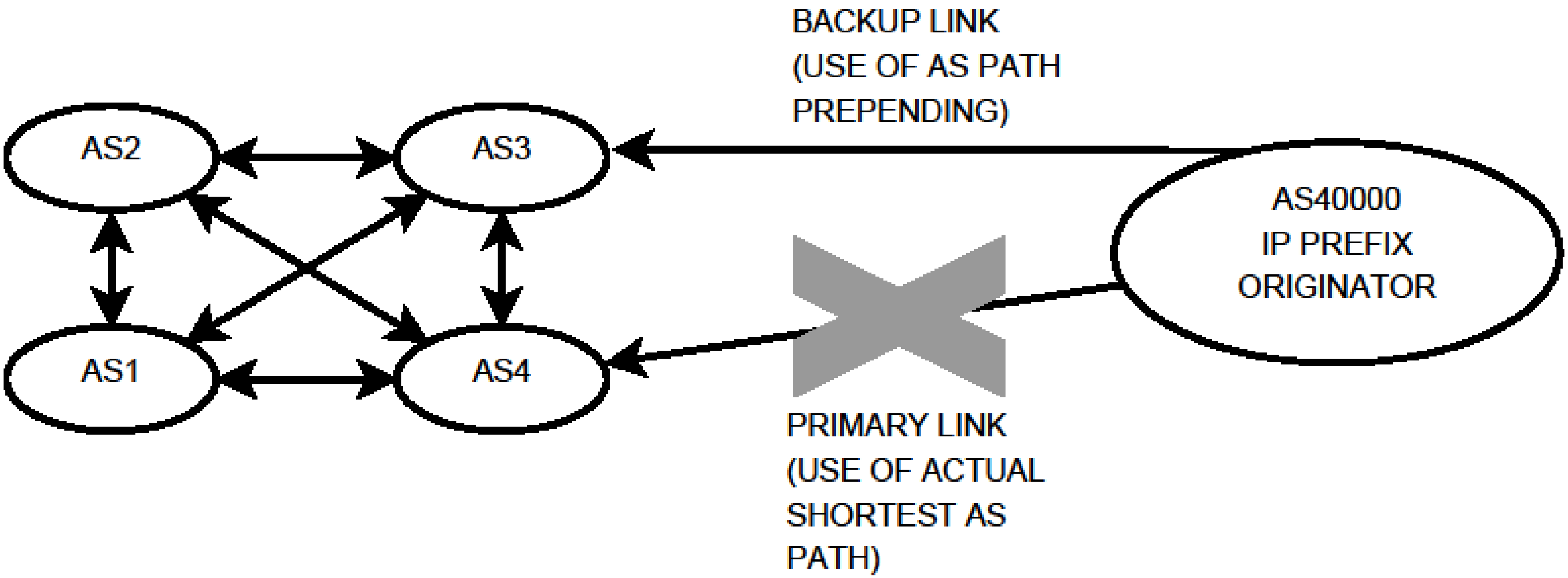
- Partial support for consistent state updates
  - During the reconfiguration of the cluster switches
  - Avoid OpenFlow barriers due to time requirements
  - Prefer installing paths in reverse order (dest to source switch)
- Proxied control traffic (BGP) handled via flow rules
  - BGP session “outsourcing” (from switch to controller)
- Direct data (ARP, IP) traffic handled via flow rules
  - Proactive strategy based on learned routing info
- More details on the implementation of the multi-AS controller (e.g., topology detection mechanisms) at [1]

[1] *Evaluating the Effect of SDN Centralization on Internet Routing Convergence*

author:Adrian Gämperli, advisors: Vasileios Kotronis, Xenofontas Dimitropoulos, supervisor: Prof. Bernhard Plattner

Master thesis at ETH Zurich, TIK institute, 2014.

Experiments: Route fail-over for multi-homed client



- Client upstreams are selected at random from an ISP core
- BGP on primary link: normal prefix ad propagation
- BGP on backup link: use path prepending in prefix ads
- Incident: primary link breaks, fail-over to backup

  ➔ Path exploration for new shortest path(s): ISPs to client

Experiments: What we measure and how

- How gradual SDN penetration on the inter-AS level affects:
  - Routing convergence times
  - Routing update churn rates

- Convergence time
  - Measurement starts at the time of the incident (e.g., link-down event)
  - Note that we are using fast keep-alive and hold-down timers

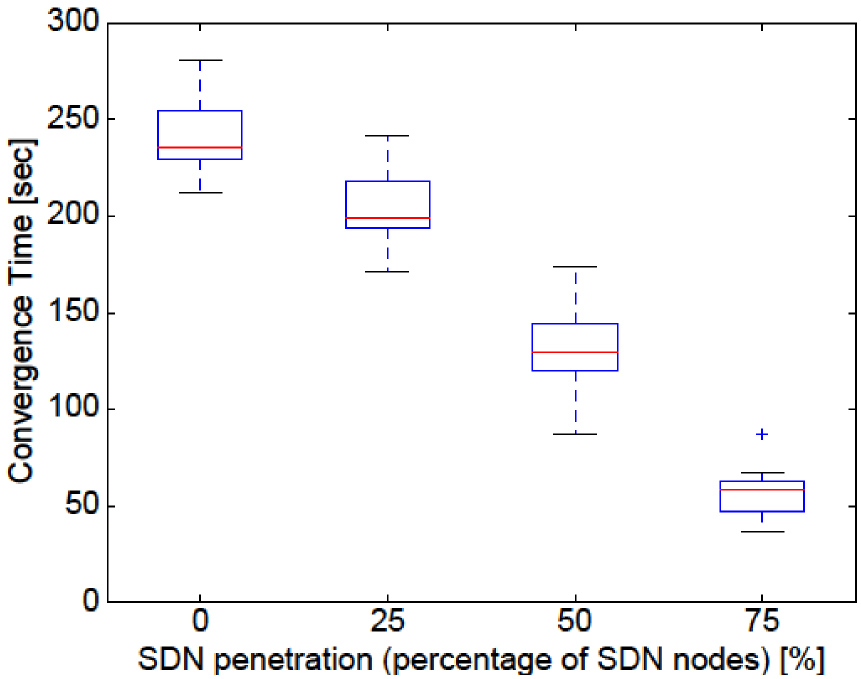
    ➔We want to explore what happens **after** the link-down detection
  - Measurement ends when the last update is received

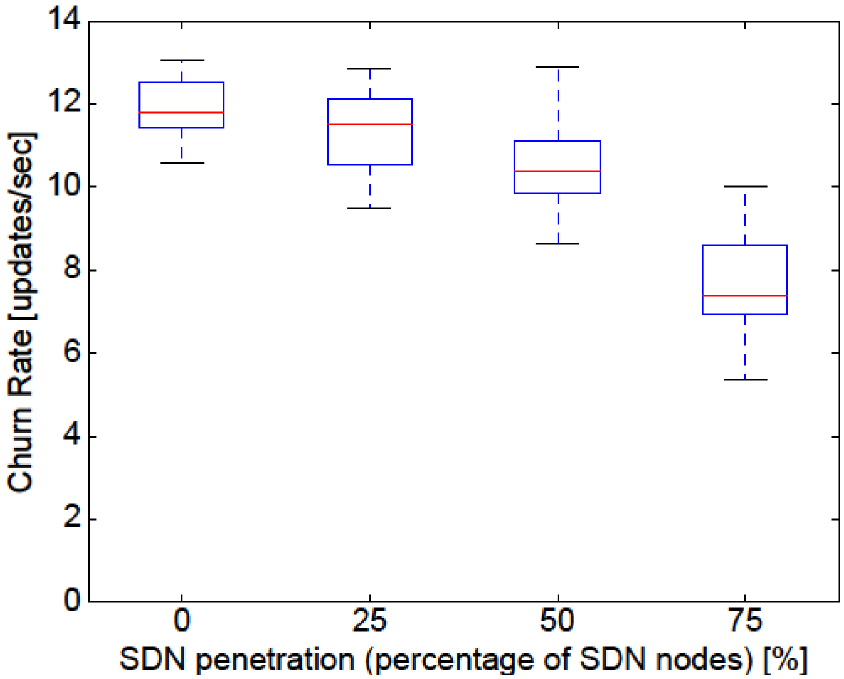
    ➔ Safe-guard intervals to make sure we have seen **all** updates

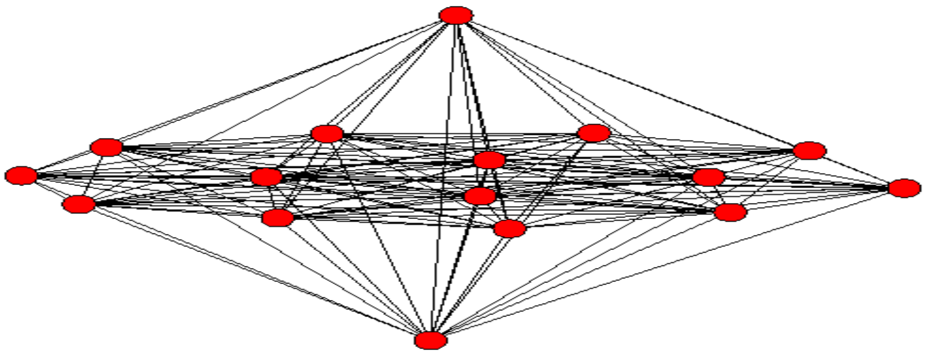
- Churn rate
  - Measure the total number of updates within convergence interval
  - Divide by convergence interval

  ➔ Average routing update churn rate

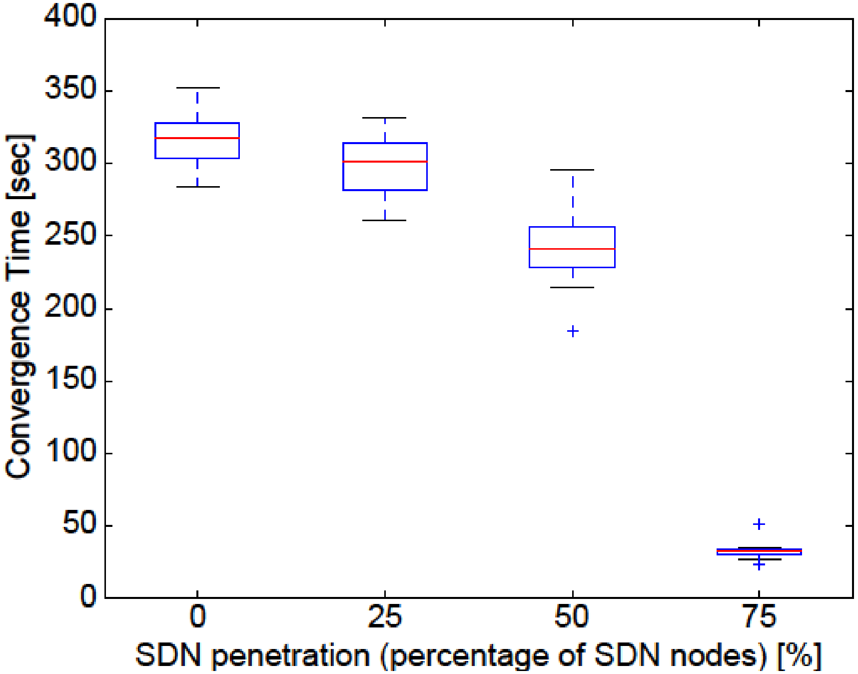
Example 1: 16-node clique (time and churn rate)

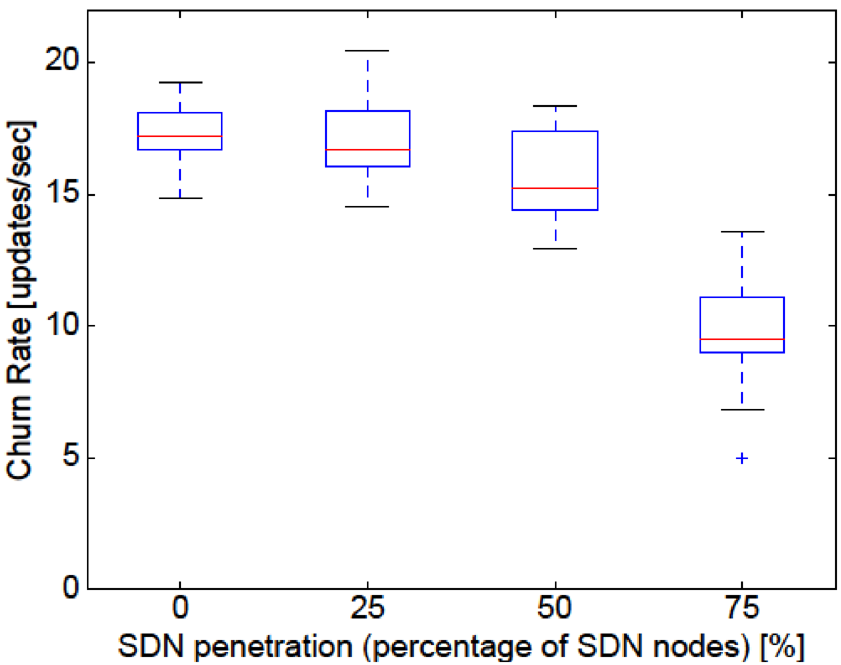


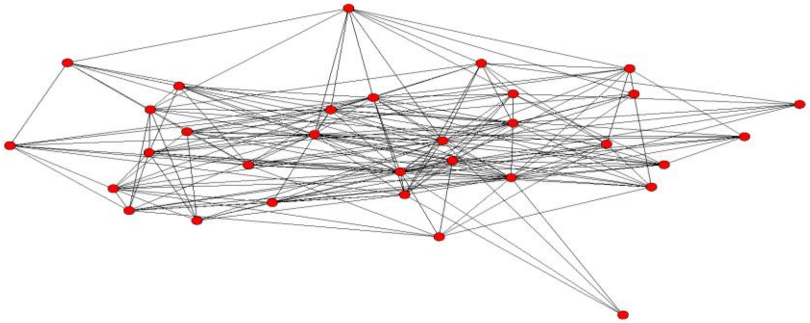




Example 2: 32-node scale-free (time and churn rate)







Observations and insights

- Gradual deployment of SDN might help stability
- Benefits in convergence times
  - Can be seen already with small penetration levels
  - Almost linear reductions
- Benefits in churn rates
  - Need larger deployments to be tangible
  - Are comparable (some times slightly worse) at small scales
- Critical mass for a Routing-as-a-Service contractor
  - Somewhere between 25% and 50%
  - Between these levels and at the 32-node scale
    ➔ conv. times can be reduced by 20%, while churn rates by 15%

➔ These results serve only as a Proof of Concept

➔ Encourage research along this direction

Take-away messages for the convergence use case

- **Difficult** to understand the dynamics between centrally controlled SDN clusters and the BGP world
  - Hybrid link state / path vector routing across domains!

- Emulation → helps gathering **meaningful** results

- With SIREN, we can experiment with **real** code
  - Focus on experiment rather than tool
  - Evaluate improvements to BGP via SDN

SIREN framework and multi-AS controller: Outlook

- Policy-compliant path calculation on controller
  - Valley-free shortest paths, c2p/p2c/p2p policies
  - Exploit bird's eye view and rich path diversity
  - Policy interactions between different services
- Controller trade-offs
  - Scalability, resiliency, centralization
  - Proper controller placement in a multi-domain setting?
  - Latency, fail-over, distribution trade-offs
- Abstractions and Services
  - Northbound interface for multi-domain services?
  - Virtualization/slicing abstractions?
  - Example service: defense against DDoS link-flooding attacks
- Packet loss and BGP/SDN convergence relationship

Download the code at:

https://bitbucket.org/gaadrian/siren/downloads

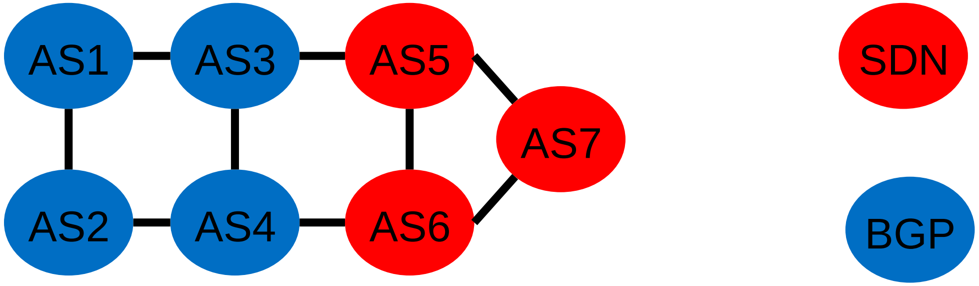
(Licenced under Apache 2.0)

YouTube video show-casing the framework

https://www.youtube.com/watch?v=Cbc8XlIp_C0&feature=youtu.be

LIVE DEMO

- Demonstrate visualization tool
- Simple hybrid BGP-SDN topology
  - SDN and BGP ASes
  - Each AS includes a test host
  - Check forwarding to AS1 host
- Start with stable routing, all-to-all pings ok
- Bring link down, check routing reactions
- Monitor how forwarding is affected from change
- Calculate convergence time on the fly



General directions for inter-domain SDN

- SDN testbed federation will be the first incubator
  - Common APIs (GENI*, NSI**)
  - Open mindset from administrators
  - Will learn about such testbeds in following lecture
- Need to better quantify benefits for ISP transition
  - Downtime minimization
  - Smooth migration schemes for the core
  - CAPEX and OPEX gains?
- SDX approach quite promising
  - Potential vehicle for other research ideas (CXPs)
- PCE-based path computation/installation and SDN
  - Path Computation Elements → No need to reinvent the wheel!
- We still have a long way towards standardization (SDNi, IETF drafts) → open design space!

* GENI: http://www.geni.net/

**NSI: http://www.terena.org/activities/e2e/ws2/slides2/11_NSI_Eduard.pdf

Any Questions 

For research collaborations please:

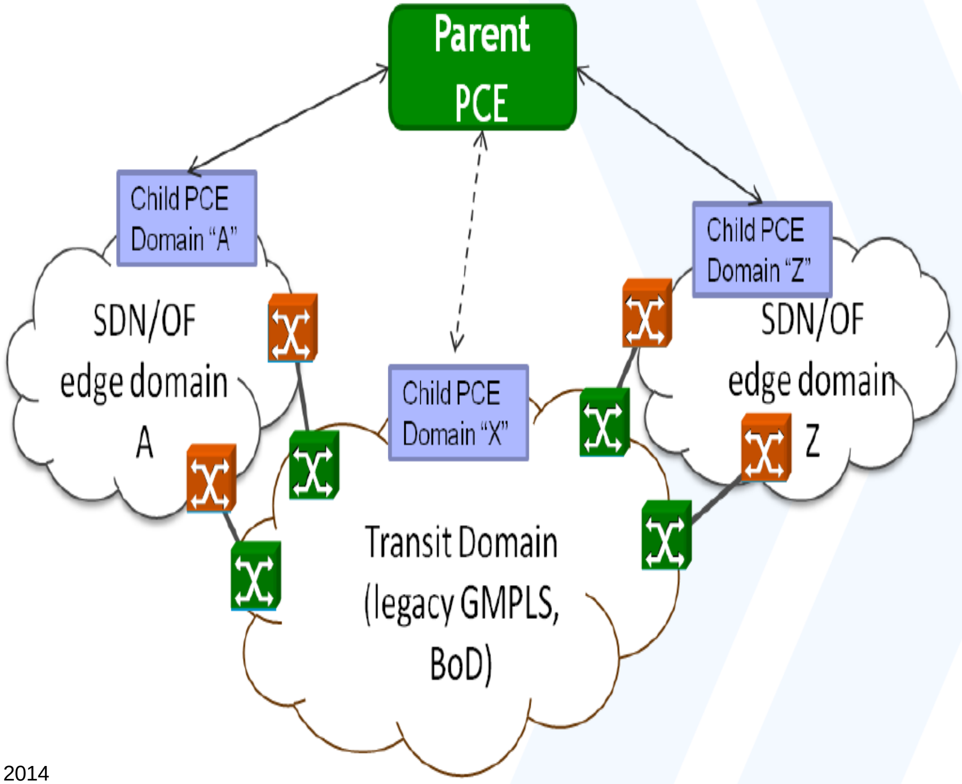
1) Check related projects on my webpage: http://www.csg.ethz.ch/people/vkotroni
2) Contact me via email: vasileios.kotronis@tik.ee.ethz.ch
3) Find me at Facebook or Linkedin (careful with “Vasileios” ☺)

BACKUP

Implementation avenue: Inter-domain transport SDN with PCE

BACKUP

- PCE is a mature concept with solid background on IETF RFCs
- Path computation and service provisioning in complex inter-domain TN
- Currently WAN is mostly static, configured on long time-scales
- SDN flexibility might be a good match for a PCE-enabled WAN
- Deployment scenario with SDN
  - SDN at the edge
    - → L2 HW with OpenFlow control
    - → Mice flows (L2/3/4 tuples)
  - GMPLS in the core
    - → HW with SNMP, PCEP etc. APIs
    - → Elephant flows (LSPs)
- Hierarchical PCEs feasible



"Inter-domain Transport SDN with PCE", from Giacomo Bernini, PACE workshop, 16th of June 2014

Candidate Clients

→ **Small or medium sized network providers**

- Why? Global trend:
  - Higher and higher interconnectivity, new services
  - “Flattening” of the AS topology graph*
  - Need for sophisticated Traffic Engineering

  → Complexity increases

  → **Who should handle the complexity?**

- Observation: Large number of potential clients out there (~10s of thousands)

BACKUP

*C. Labovitz et al. Internet Inter-Domain Traffic. SIGCOMM 2010.

Candidate Contractors

➔ **Larger and tier-1 ISPs (but: potential conflicts)**

➔ **External specialized parties (more objective)**

BACKUP

- Why?
  - Considerable expertise in routing
  - Incentive for a new service type provision (outsourcing)
  - Opportunity for an economy of scale

- Example: AT&T
  - Tier-1 ISP
  - Market leader in handling outsourced network services

Outsourcing: smooth transition

- Transition stages:

Consultation, Information about Best Practices → Routing Control Logic Path Computation + Configuration → Low-level control of FIBs, RIBs Routing Control Plane

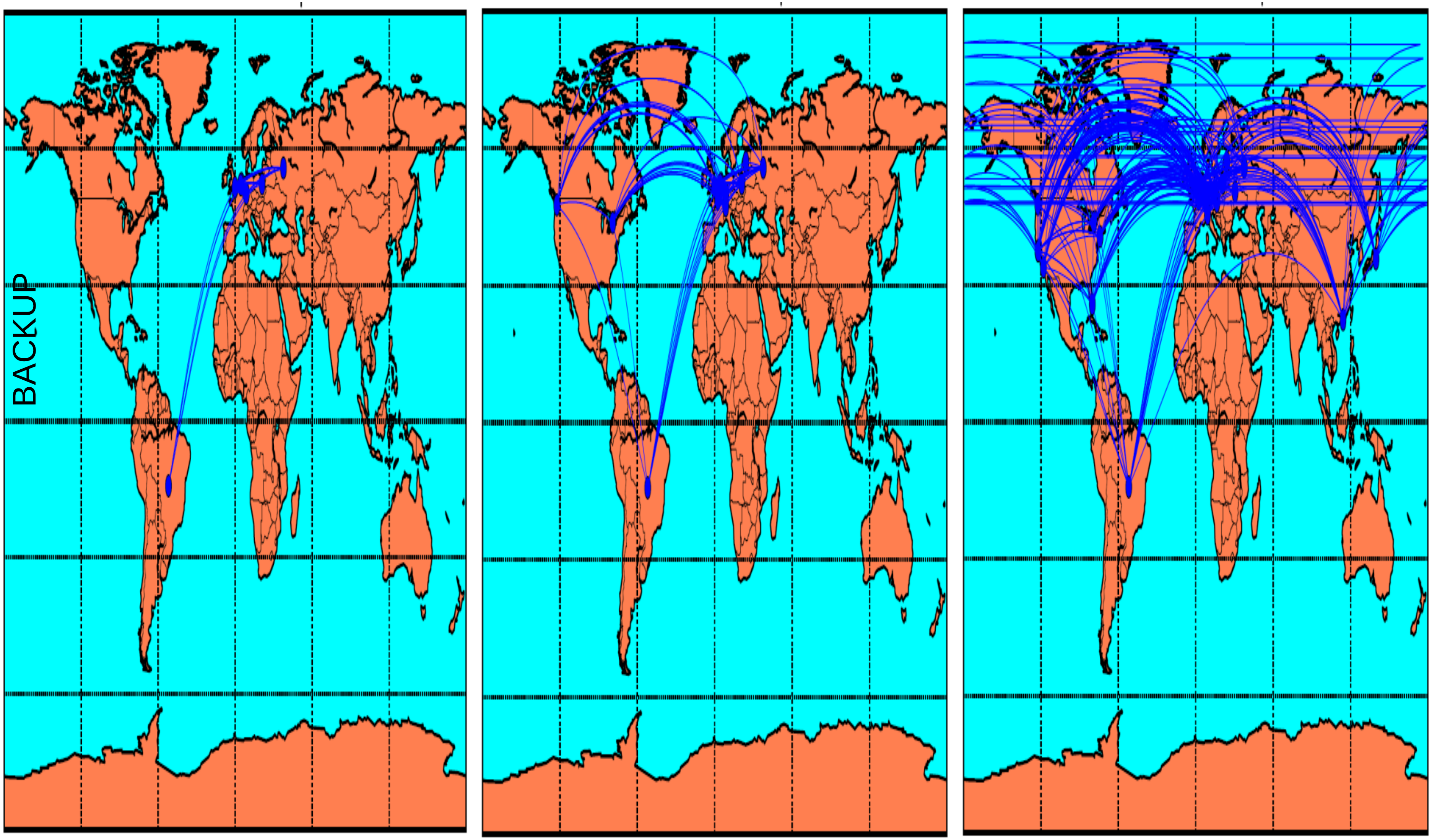
BACKUP

- During the transition the client:
  - Shapes **his own** policies (based on business model)
  - **Expresses his requirements** to the contractor
  - Maintains policy **privacy*** based on:
    - Trusted third party model
    - NDAs
- If not satisfied ➔ backtrack/change provider
  - Providers “behave” better when competition exists

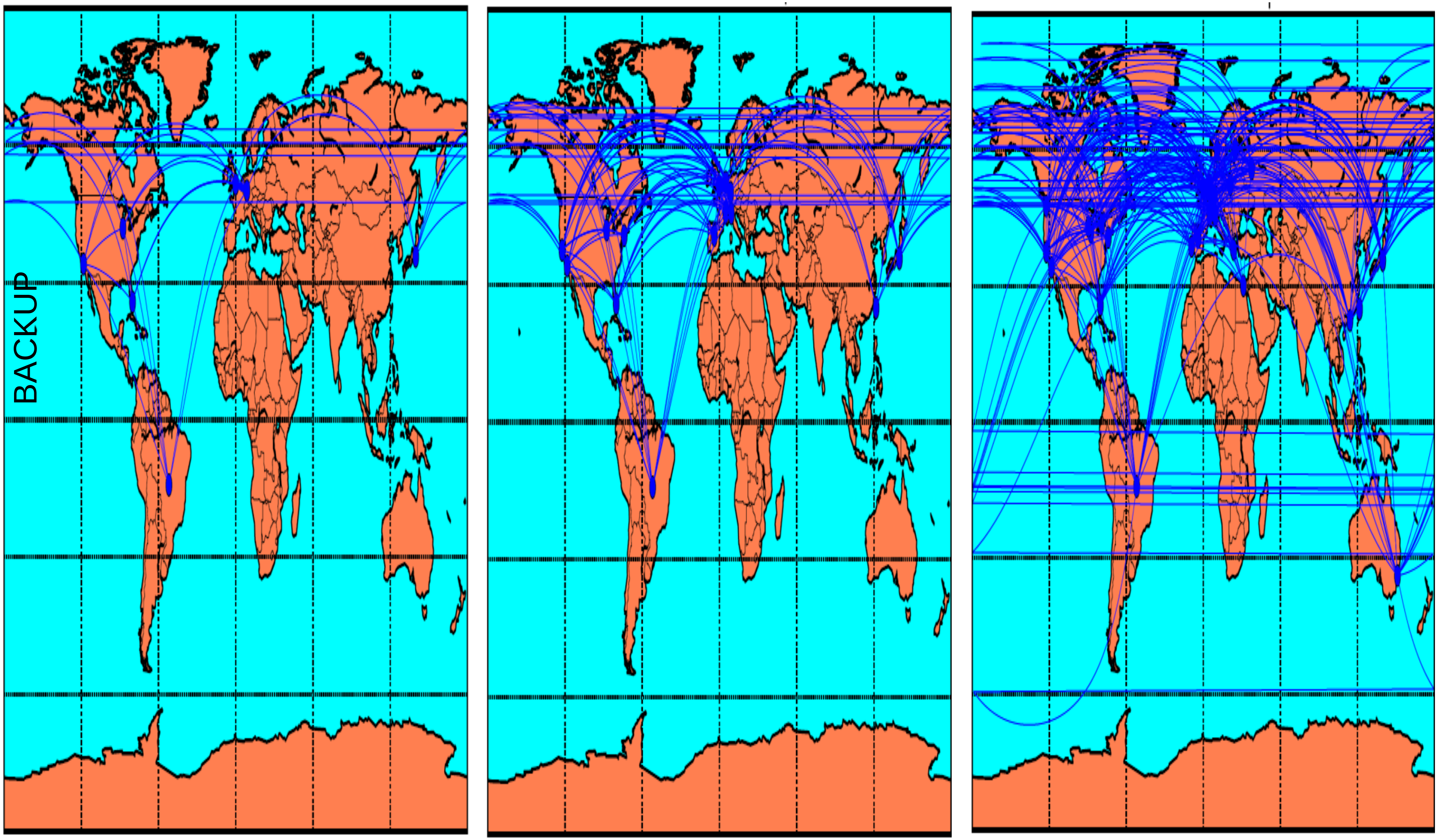
*No leakage to competitors

[CXPs] Evolution of inter-IXP graph based on IXP membership

BACKUP



[CXPs] Evolution of inter-IXP graph based on IXP IP coverage Maximum Utility Function → Much more diversity!



Τέλος Ενότητας

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Κρήτης**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σημειώματα

Σημείωμα αδειοδότησης

Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Κρήτης, Ξενοφώντας Δημητρόπουλος. «**Δίκτυα Καθοριζόμενα από Λογισμικό. Ενότητα 4.2: Inter-domain SDN: Επισκόπηση έρευνας**». Έκδοση: 1.0. Ηράκλειο/Ρέθυμνο 2015. Διαθέσιμο από τη δικτυακή διεύθυνση: http://www.csd.uoc.gr/~hy436/