Experience with implementing VNF chains with Segment Routing and PCEP

CEES PORTEGIES[1]

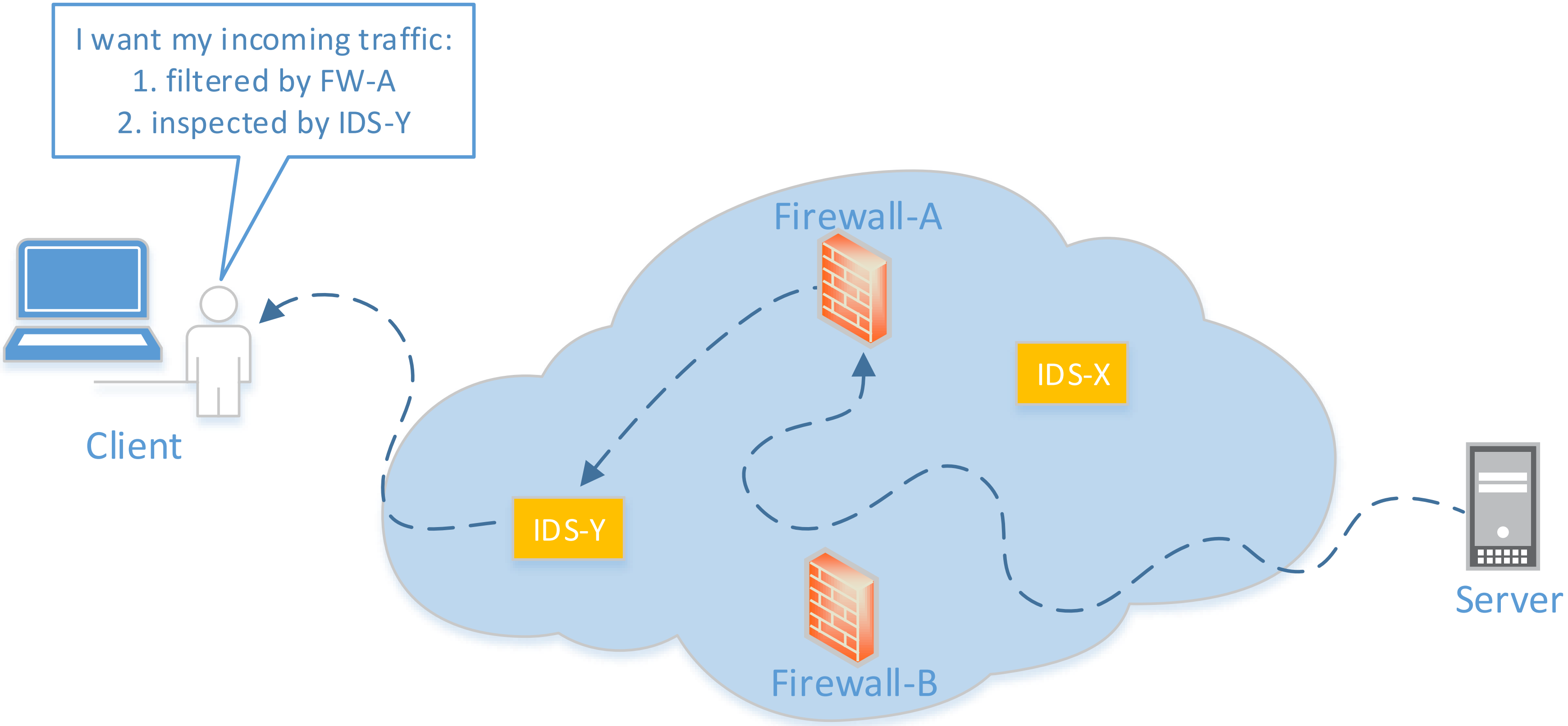
[1]Multiscale Networked Systems group, University of Amsterdam (mns-research.nl)

UNIVERSITY OF AMSTERDAM

SURF

Motivation

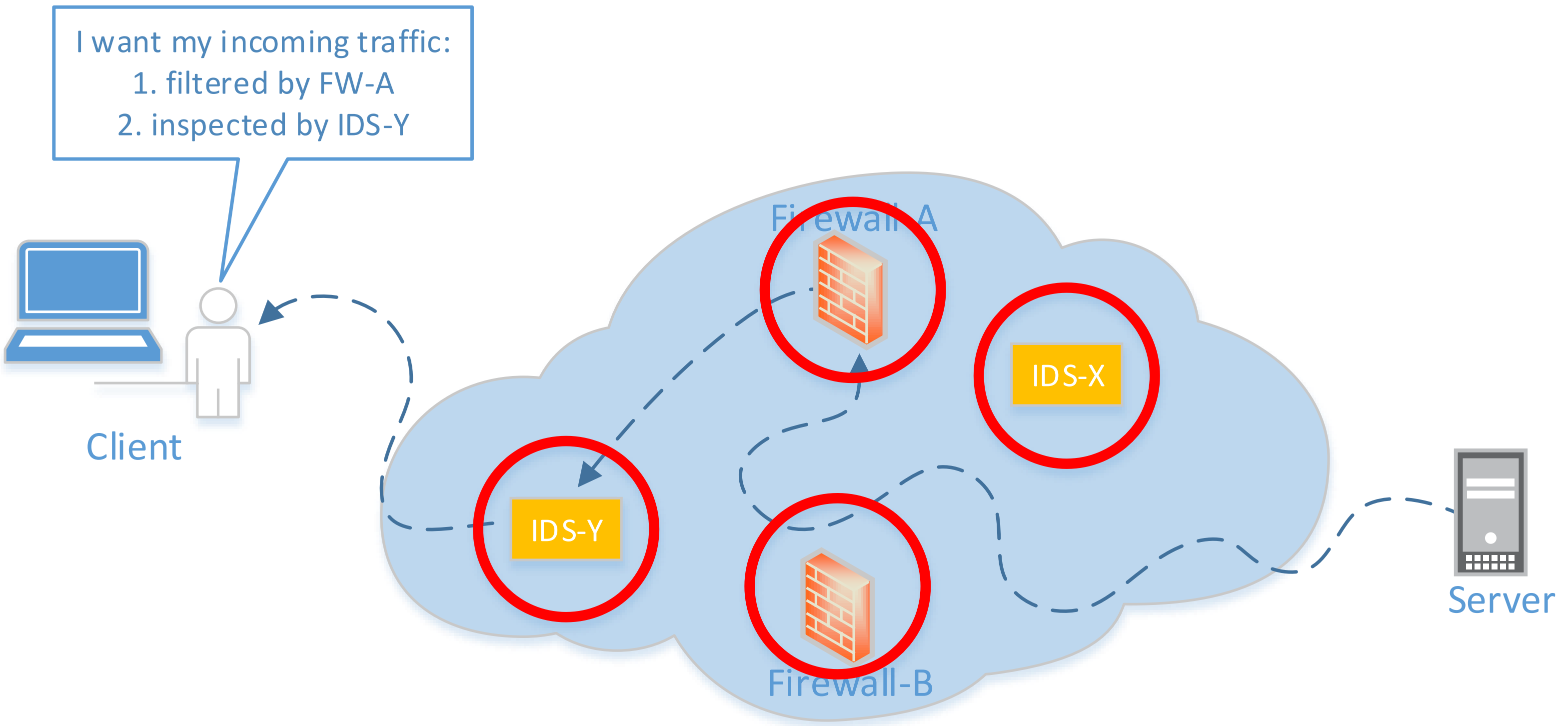
"How can we build network paths through network devices wherever they may be in the network?"



How does the network operator build the path?

Problem 1: Service deployment

- How do we deploy new services/functions for clients?



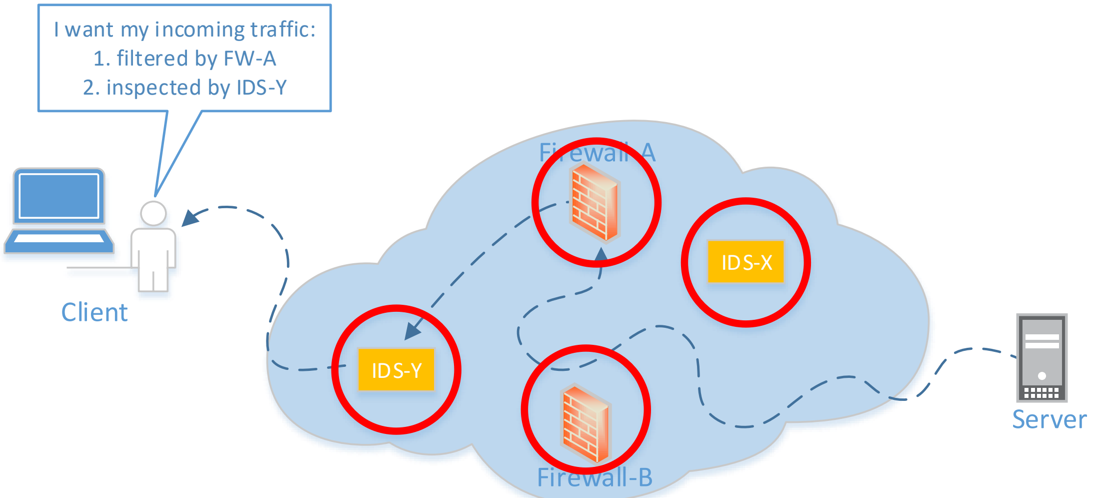
How does the network operator deploy the services

Network Function Virtualization

- Hardware boxes become virtual box
  - Firewall ‘function’ box becomes Virtual Network Function (VNF) Firewall

- Can be deployed on generic compute

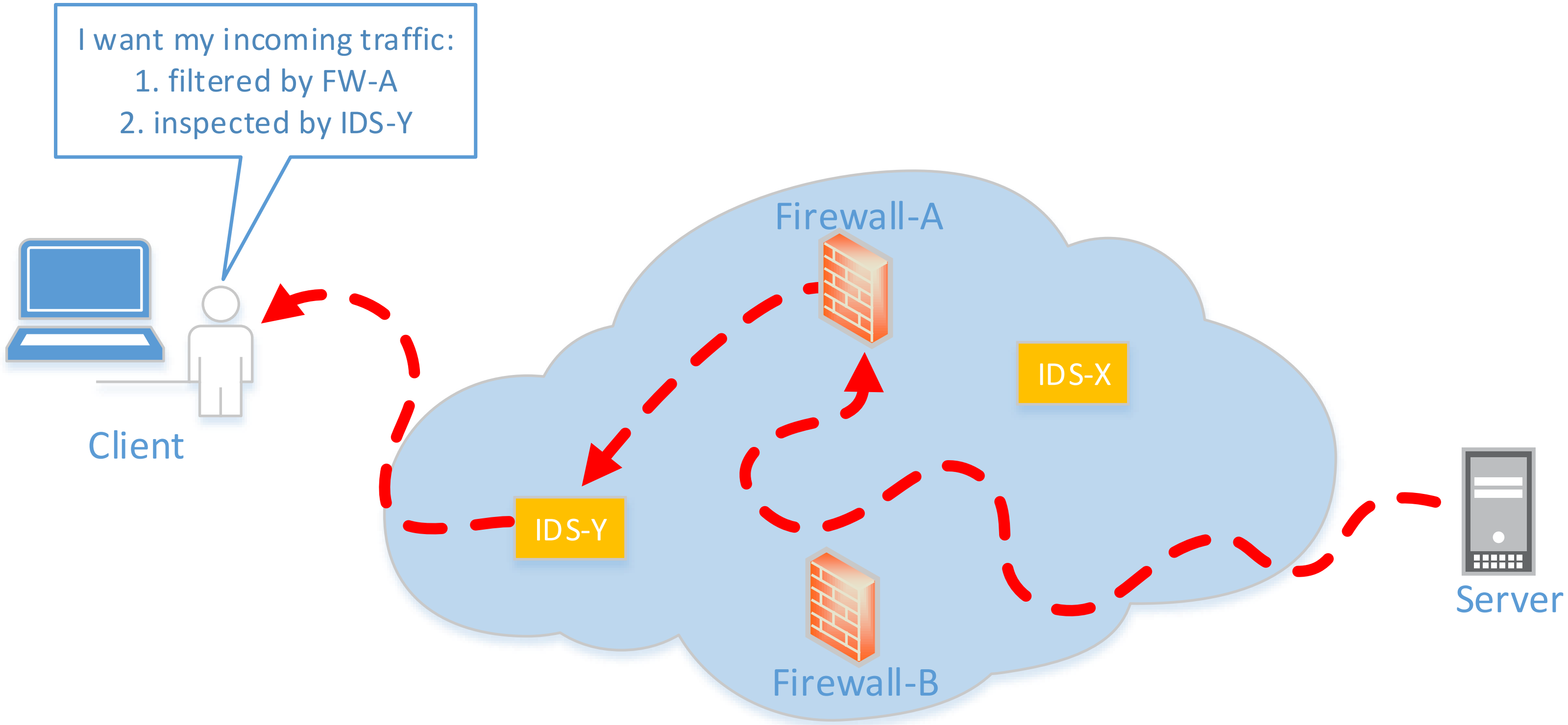
- Service Function Chain: multiple VNFs after each other
  - “First I want to pass my traffic through a Firewall and then through an IDS”



How does the network operator deploy the services

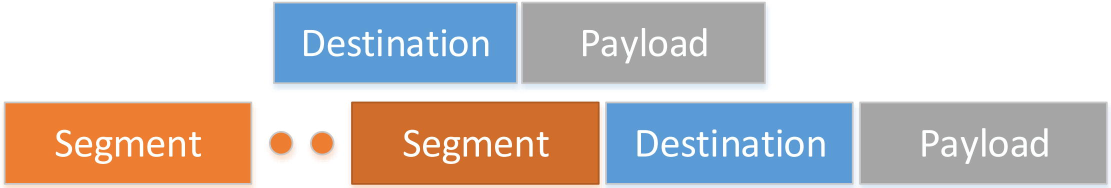
Problem 2: Steering the traffic

- How do we steer traffic to each of those deployed services?

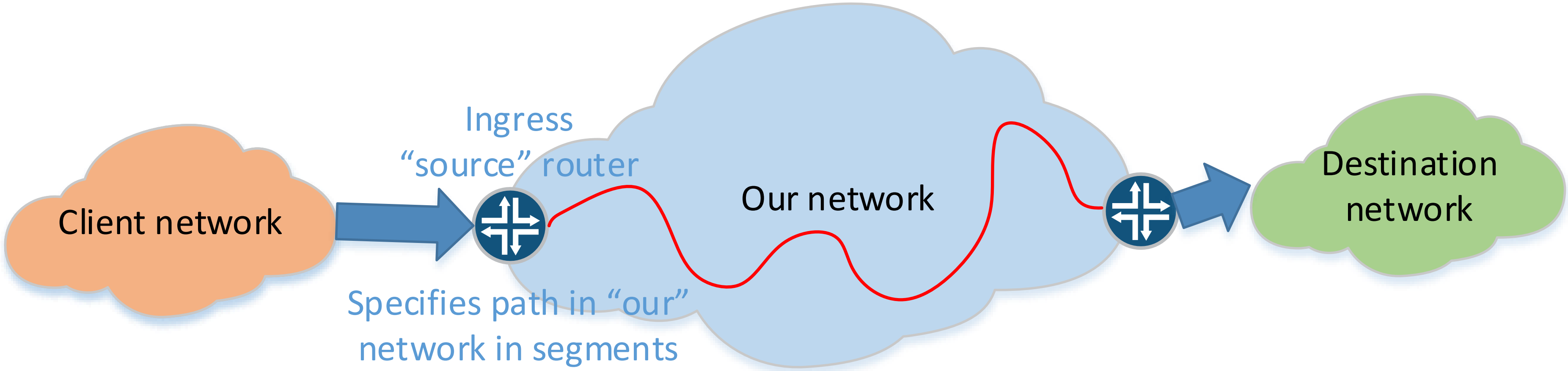


How does the network operator steer traffic to the services

Segment Routing

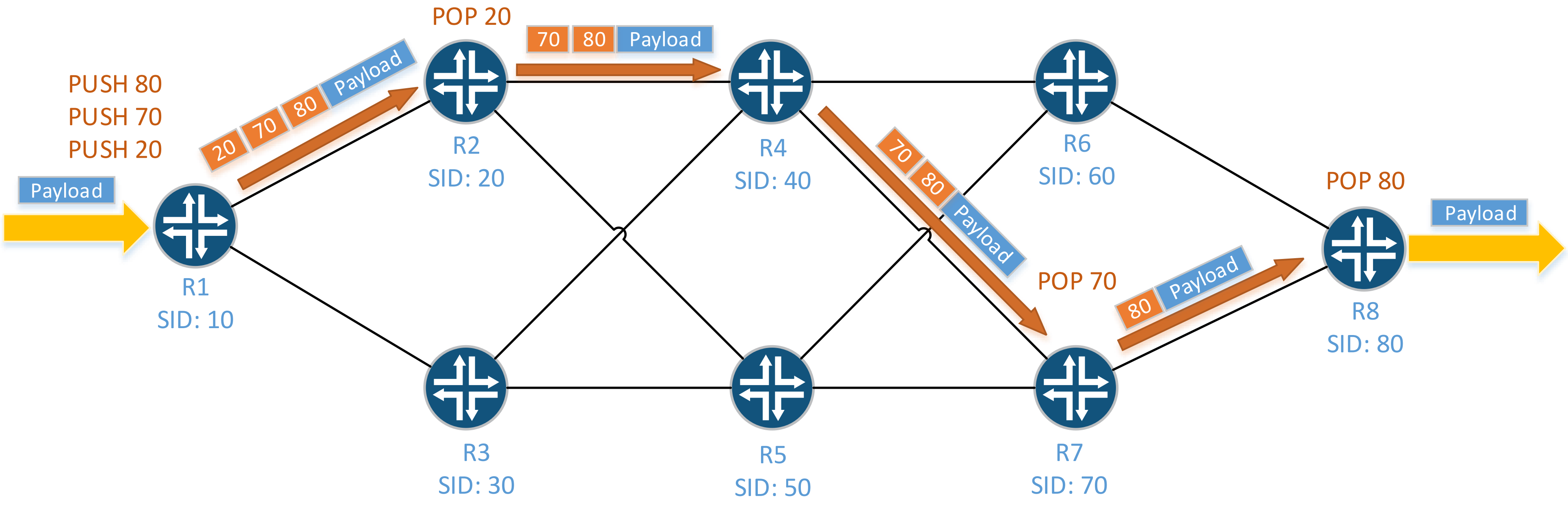


- IP Routing: Destination based
- Segment routing: Source based
  - "Source" determines the path to be taken
  - Segment IDentifier path ('stack')
  - Computation overhead at ingress
- Node, prefix, adjacency and anycast segments



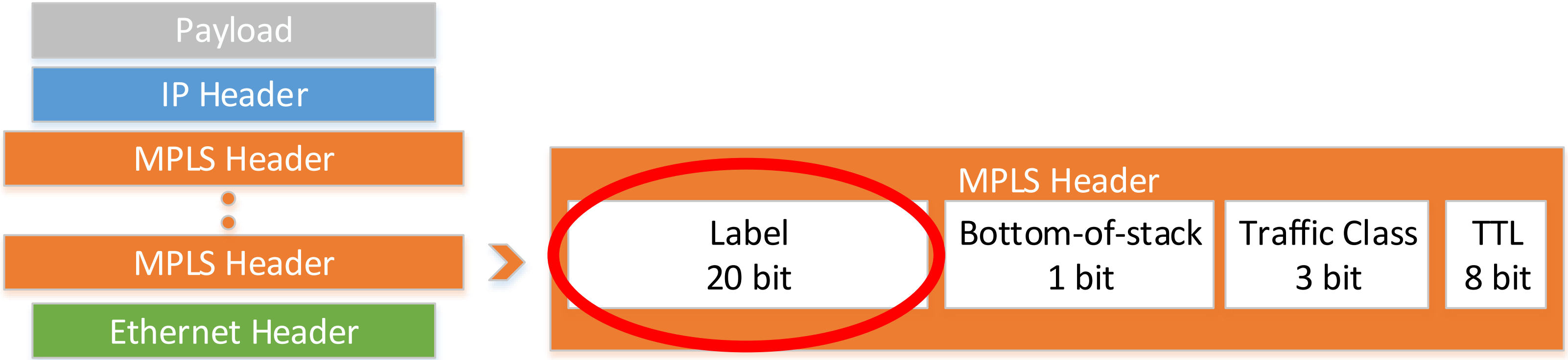
Segment Routing

- Traffic incoming on R1 should always go through R2 and R7
  - The rest of the path is up to the IGP
- Labels explicitly targeting R2 and R7 are pushed



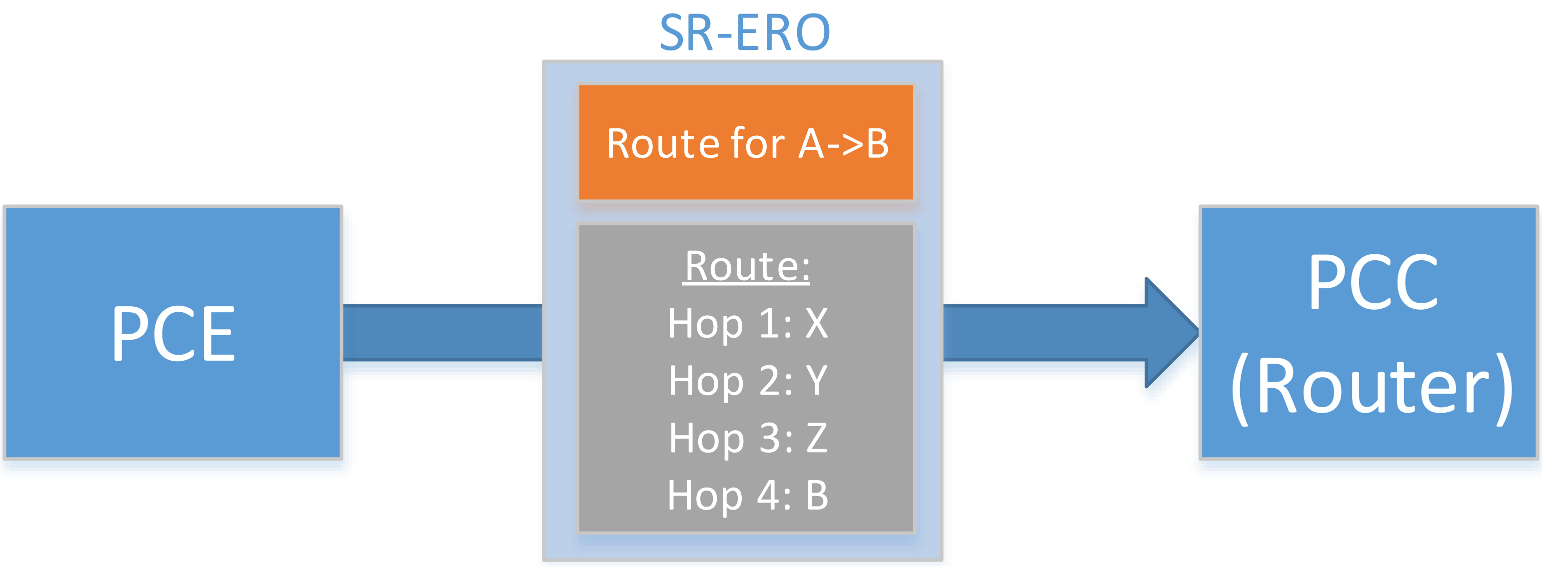
SR-MPLS

- SR-MPLS re-uses Multi Protocol Label Switching data plane
  - MPLS Label -> Segment Identifier (SID)
  - Label in MPLS: Locally significant
  - Label in SR-MPLS: Globally significant
- Paths
  - MPLS: Label Switched Path (LSP)
  - SR-MPLS: Segment Routed Label Switched Path (SR-LSP)
- Label distribution
  - SR-MPLS: IGP
- IGPs with SR Support
  - IS-IS
  - OSPF



Path Computation Element Protocol (PCEP)

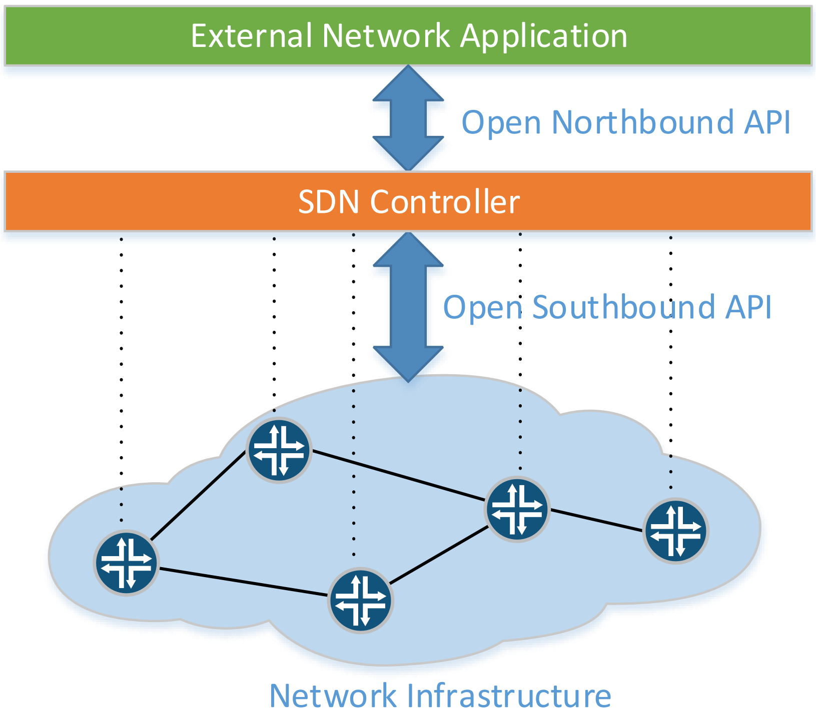
- "How to signal the SR-LSPs?"
  - Paths as Explicit Route Objects (ERO)
    - For segment routing this becomes Segment Routing ERO (SR-ERO)
  - Consists of Path Computation Client (PCC) and Path Computation Element (PCE)
    - The PCE pushes out the SR-EROs
    - The PCC receives SR-EROs



SDN Controller with PCEP

- SDN Controller: abstraction over the network

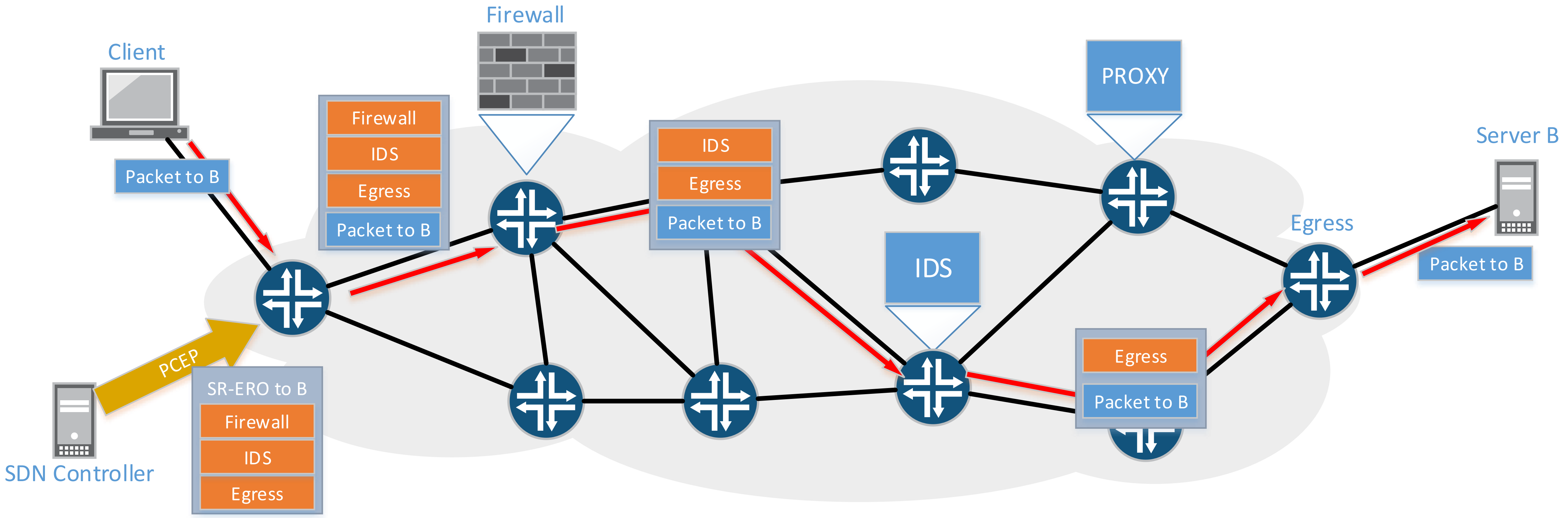
- Northbound API
  - External coordination
- Southbound API
  - Controlling Paths
    - PCEP
  - Topological information
    - BGP Link State (BGP-LS)



Research Question

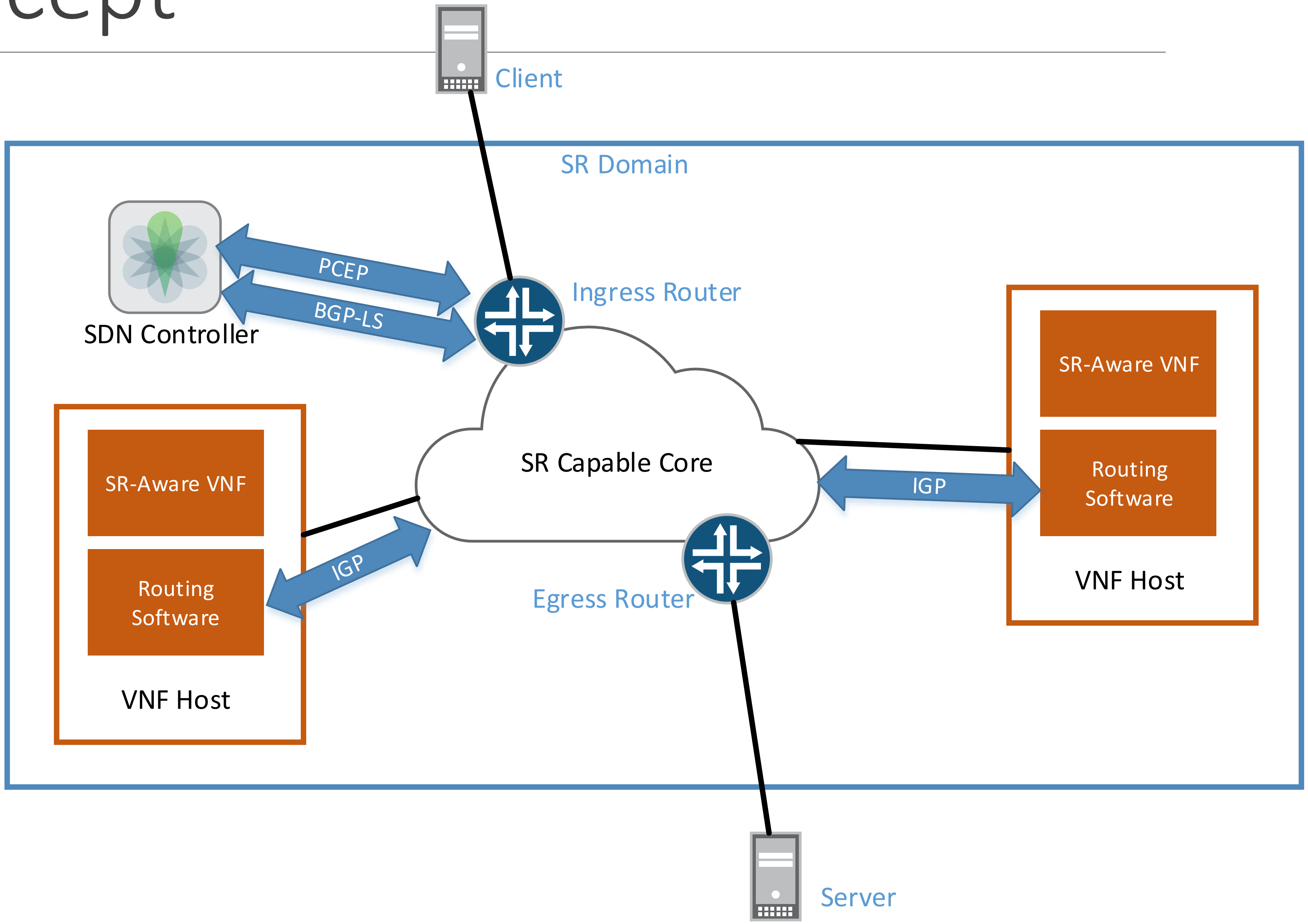
"Can PCEP be used to create SR-MPLS network paths to assist the network integration of VNFs?"

- VNFs compatible with SR: SR-Aware
- PCEP controlling SR-LSPs in the SR-MPLS data plane



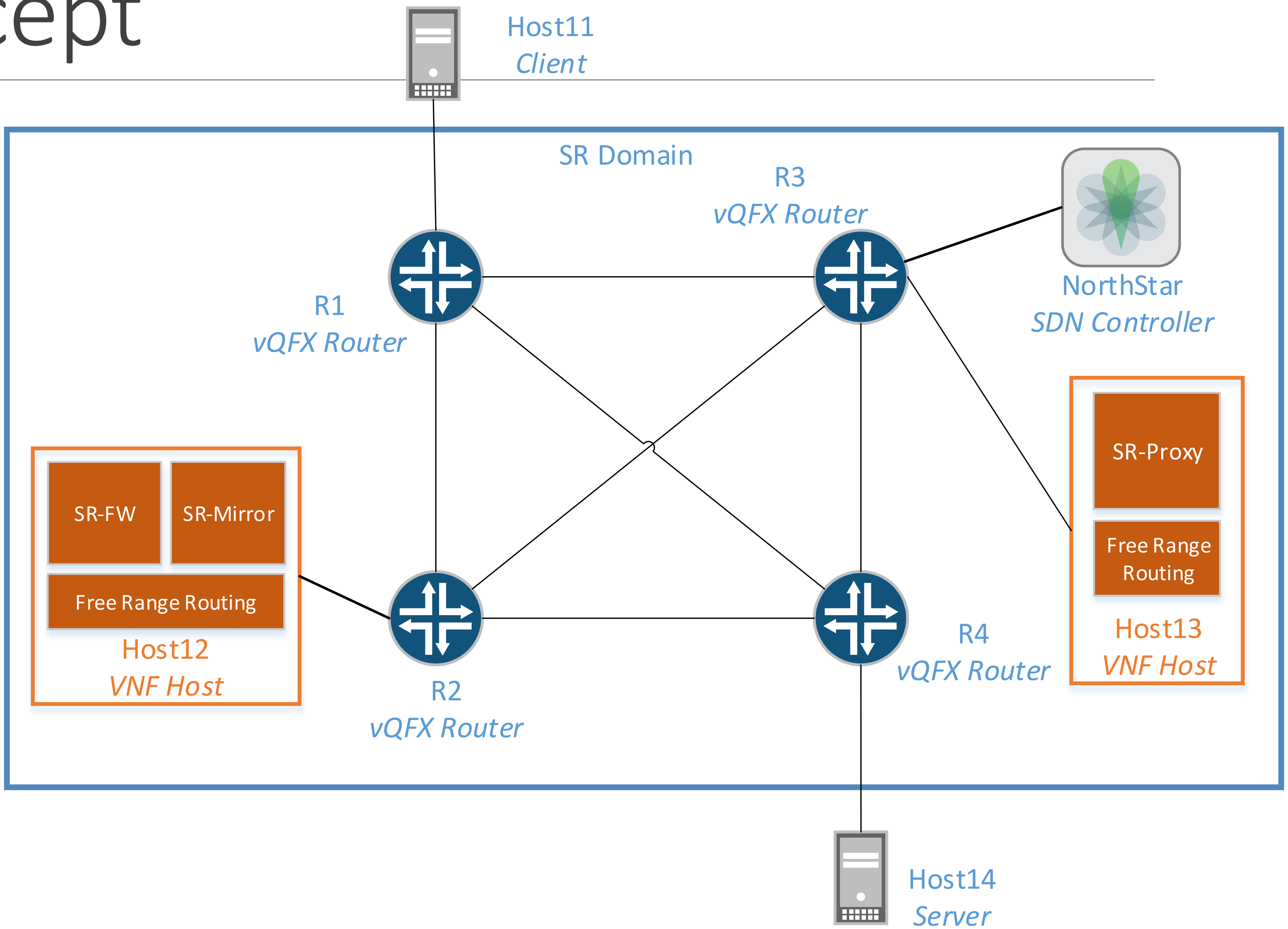
Proof of concept

- VNFs
  - SR-Aware
  - Migration
  - Chaining
- SR-MPLS data plane
- SDN controller
  - PCEP
  - BGP-LS
- Coordinating
  - VNFs & SR-LSPs



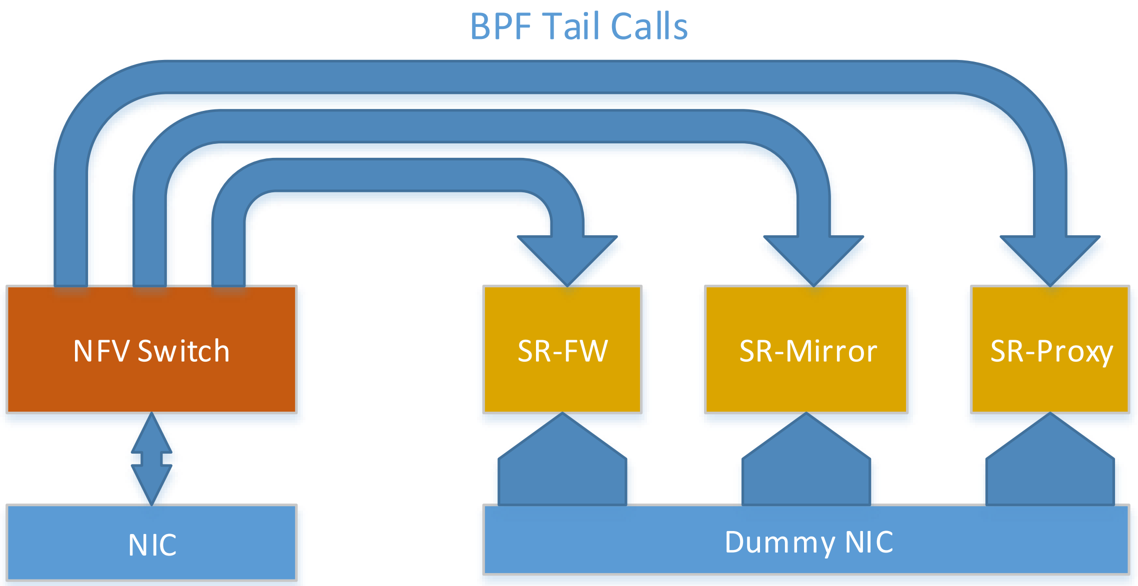
Proof of concept

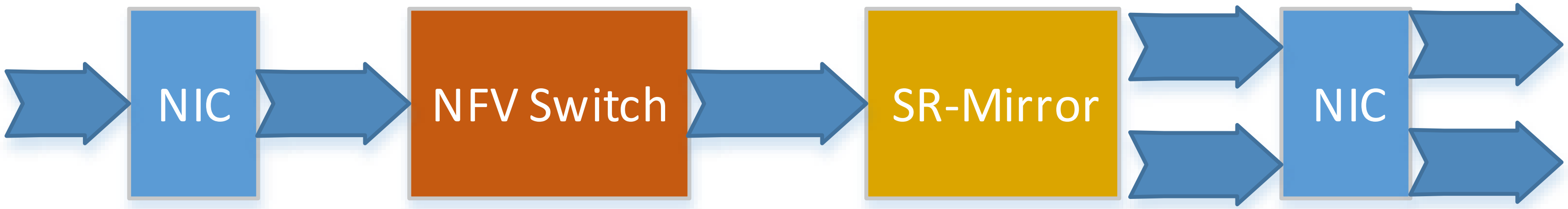
- BPF based VNFs
  - Ubuntu 18.04 VNF hosts
  - FRR 7.4-dev
  - Host12 & Host13
- SR Capable Routers
  - R1 - R4
  - Juniper vQFX 19.4R1.10
- IGP: IS-IS
- Client & Server
  - Host11 & Host14
  - Ubuntu 18.04
- NorthStar SDN controller
  - Version 6.0.0



BPF based VNFs

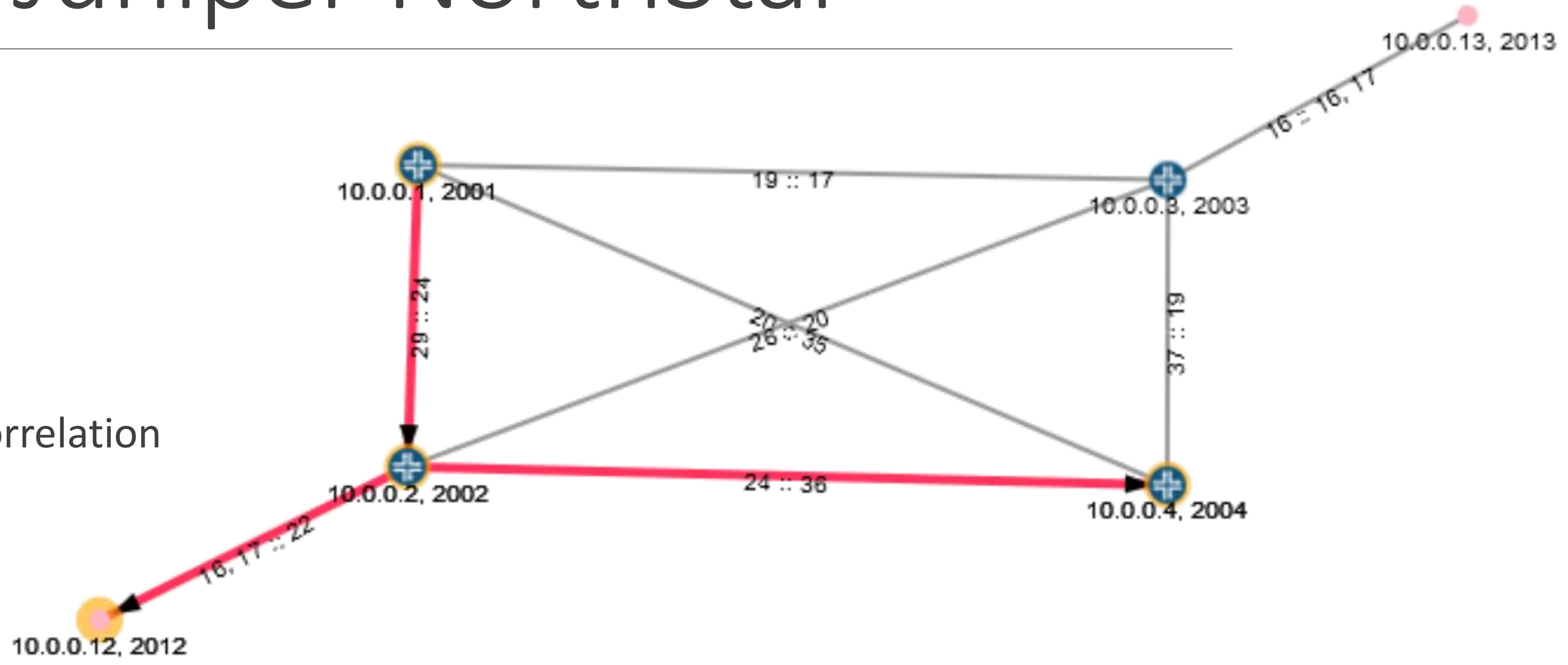
- University of Amsterdam developed BPF TC classifiers
  - RON19, Łukasz Makowski
- NFV Switch
  - Switching layer
- 3 Types of VNFs
  - SR-Firewall
  - SR-Mirror
  - SR-Proxy
- BPF Tail Call:
  - Non-returning context switch





SDN Controller: Juniper NorthStar

- North Bound:
  - REST API
- South Bound:
  - BGP-LS: Topology acquisition
  - NETCONF: Juniper configuration correlation
  - PCEP: Path construction
- No arbitrary SIDs in SR-LSPs



Node | Link | Tunnel | +

Name ↑	Hostname	IP Address	Type	NETCONF Status	PCEP Status	SID	SR	SRGB	PCE-SPRING	Last Update
0010.0000.0001	vqfx1	10.0.0.1	JUNIPER	Up	Up	2001	✓	1000-8001	✓	
0010.0000.0002	vqfx2	10.0.0.2	JUNIPER	Up	Up	2002	✓	1000-8001	✓	
0010.0000.0003	vqfx3	10.0.0.3	JUNIPER	Up	Up	2003	✓	1000-8001	✓	
0010.0000.0004	vqfx4	10.0.0.4	JUNIPER	Up	Up	2004	✓	1000-8001	✓	
0010.0000.0012	host12	10.0.0.12	GENERIC			2012	✓	1000-8001	⊘	
0010.0000.0013	host13	10.0.0.13	GENERIC			2013	✓	1000-8001	⊘	

Coordination for VNF migration

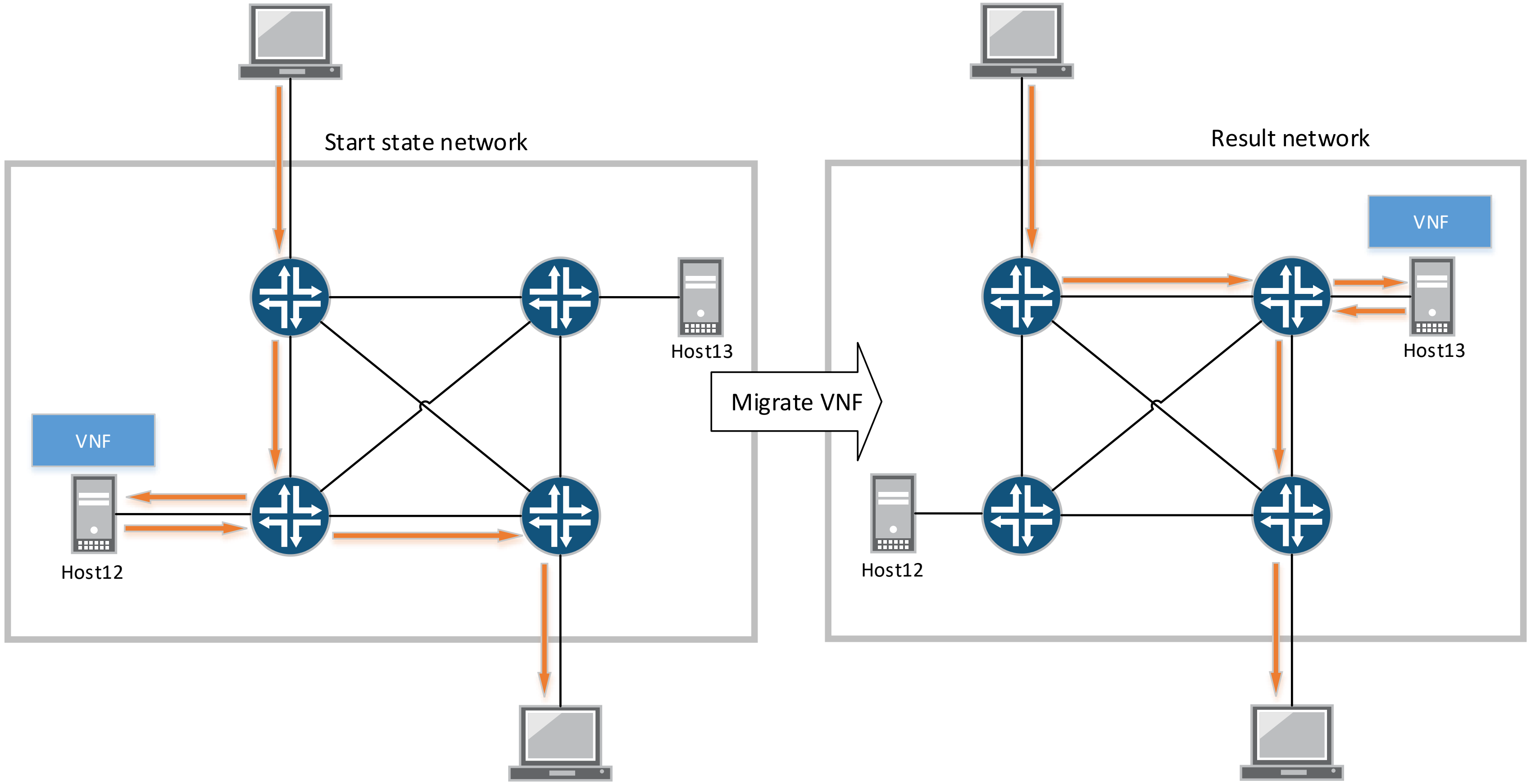
- Custom tool:
- Paths: SDN SR-LSPs
- VNFs: BPF system

Migrate VNF from Host 1 to Host 2

	Interaction		Description	Protocol
Step 1	Start VNF	VNF Host 2	Attach VNF to NFV Switch on Host 2	BPF
Step 2	Change SR-LSP		Change the NorthStar LSP from Host 1 to Host 2	NorthStar REST API
Step 3	Push new SR-LSP	Router 1	NorthStar pushes the new LSP via PCEP to R1	PCEP
Step 4	Stop VNF	VNF Host 1	Detach VNF from NFV Switch on Host 1	BPF

Experiment 1: Re-instantiation

- The VNF is moved from host12 to host13

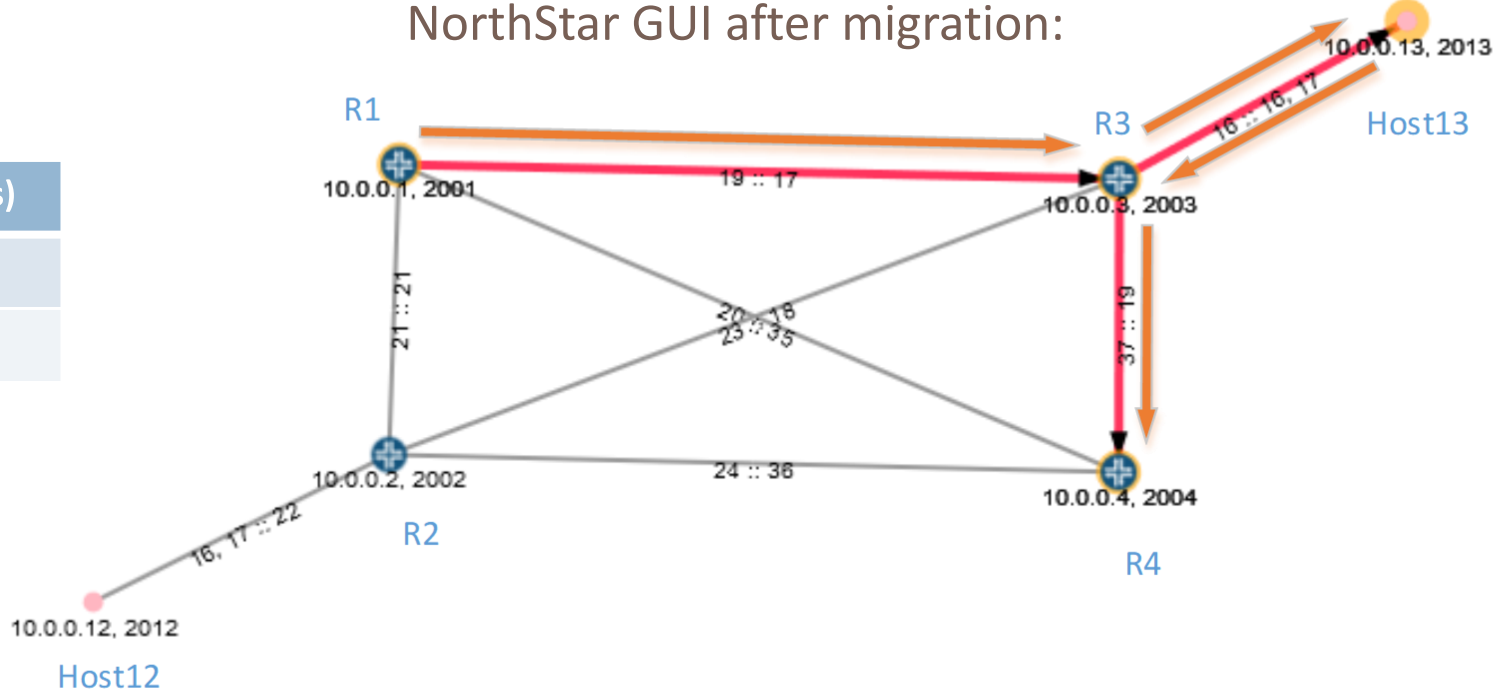


Experiment 1: Results

- No packet loss observed
- No increase in latency
  - The high latency is the result of the vQFX routers

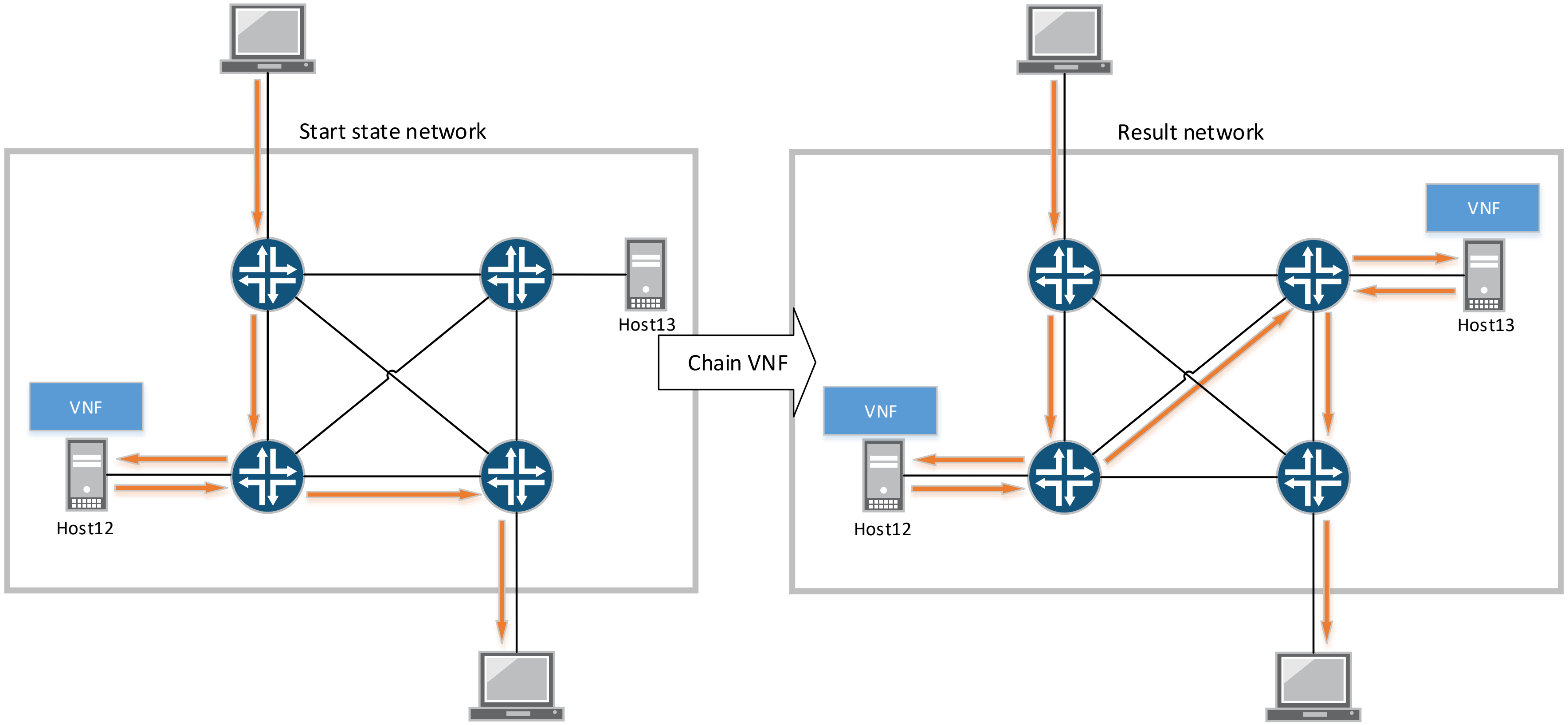
	Latency (ms)	Std. Dev. (ms)
Before Migration	266	49
After Migration	275	44

NorthStar GUI after migration:



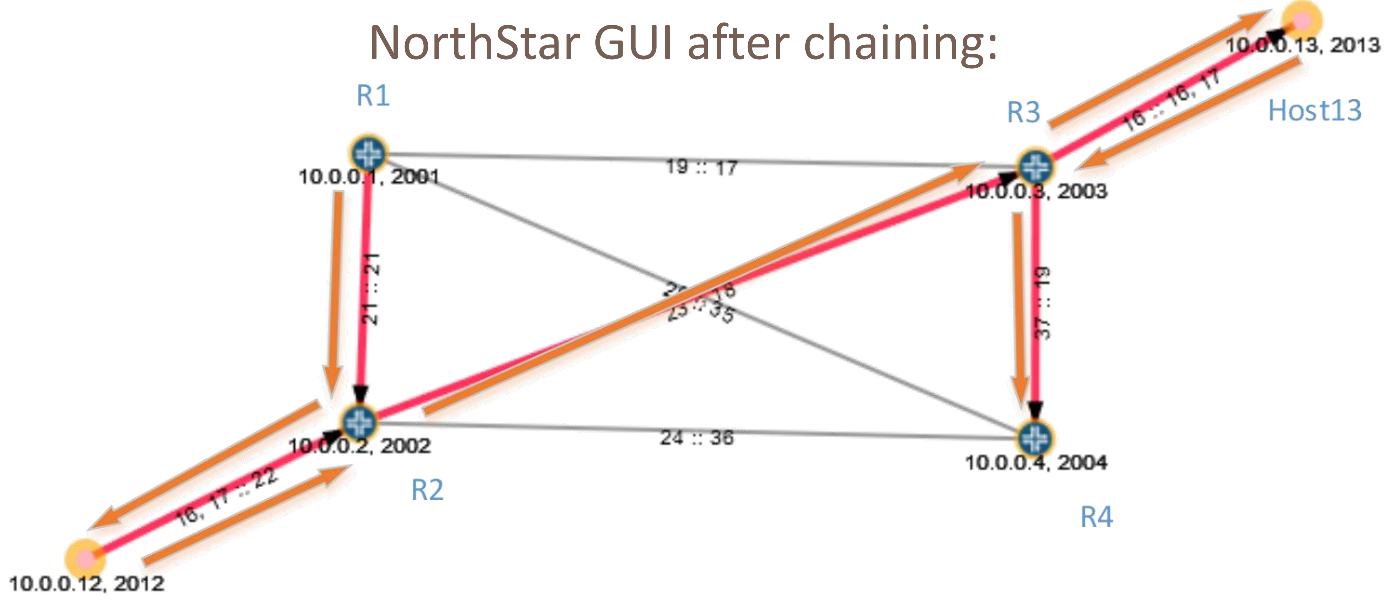
Experiment 2: Chaining

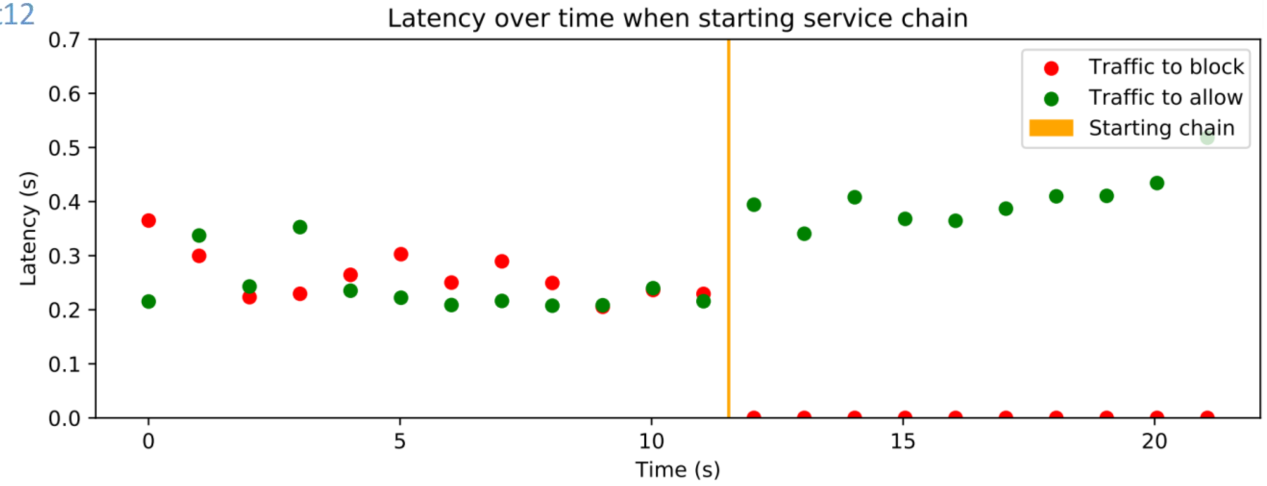
- VNF on host12 allows traffic towards port 80
- VNF on host13 drops traffic towards port 80



Experiment 2: Results

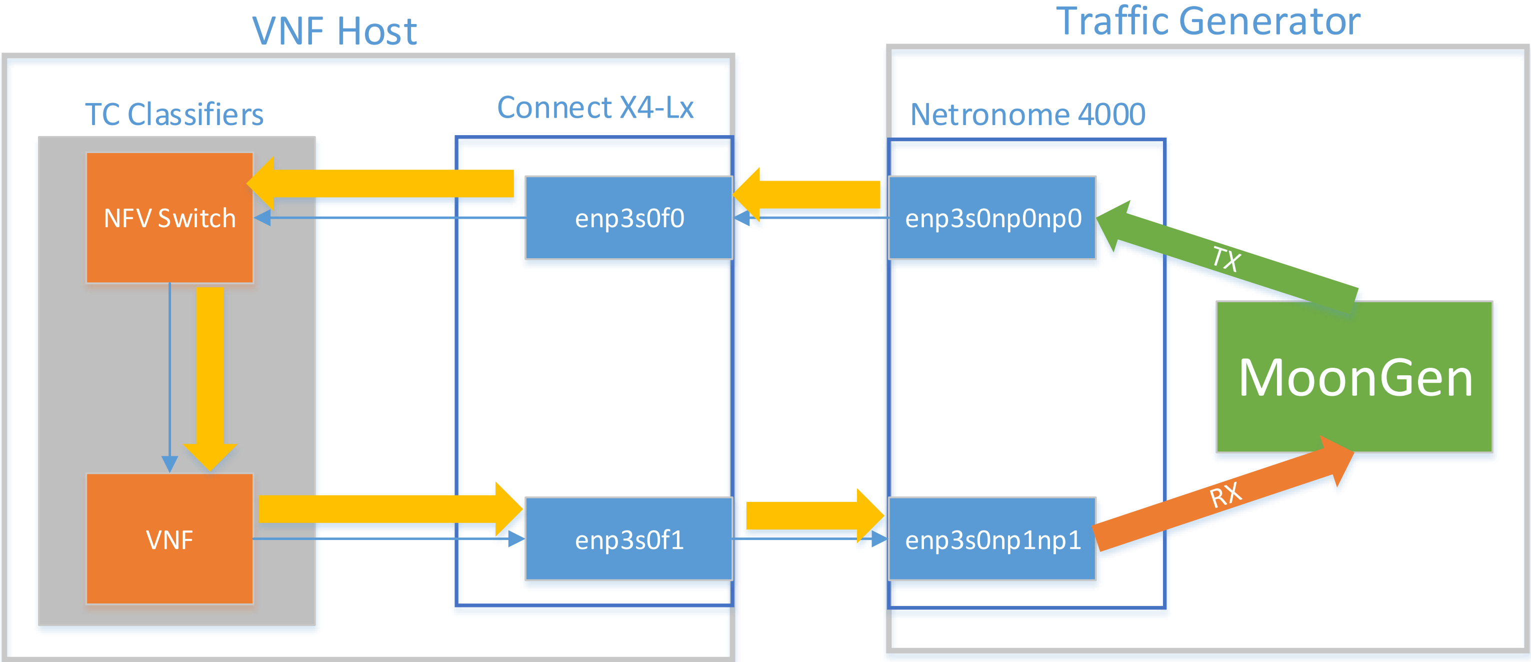
NorthStar GUI after chaining:





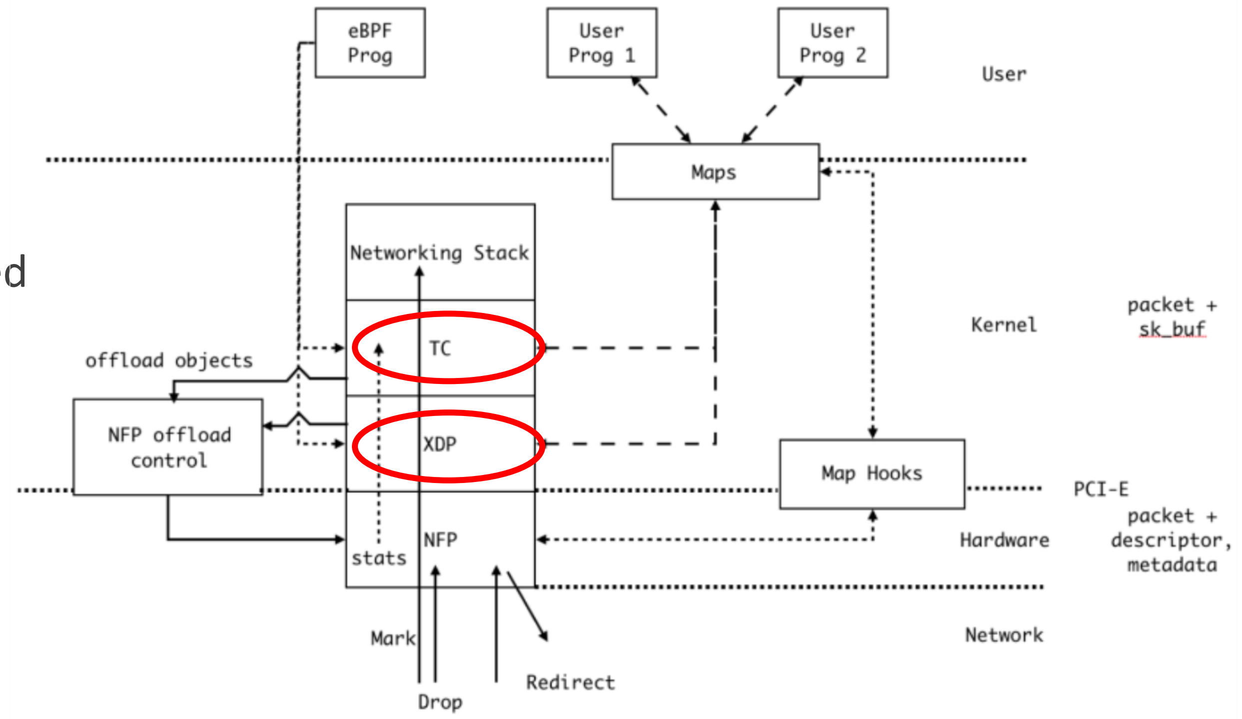
Performance

- Some preliminary performance testing has been performed
- Focused on NFV Switch + VNF performance
- Traffic generation
  - SR-MPLS tagged traffic, not widely supported
- MoonGen
  - Can do at least 25Gbit/s
- Current findings:
  - Mostly independent of packet size
  - ~ 0.71 Mpps throughput
  - Fully CPU bound
    - On Xeon E5-1620 @ 3.5Ghz v4



Performance: Hardware Offload

- VNFs are BPF TC Classifiers
  - 'High' in the driver stack
- TC Classifier offload is limited
  - Very dependent on NIC, Netronome has most features
  - Pushing/popping labels (resizing packets) not supported
- XDP as a potential alternative
  - Lower level
  - Requires rewriting of the MPLS header parsing library
- DPDK can support BPF Programs
  - Very limited support, no maps
  - Not viable (currently)



Netronome Driver (NFP) overview

Conclusions and Future work

"Can PCEP be used to create SR-MPLS network paths to assist the network integration of VNFs?"

- Yes. Experiments were successful
  - Migration occurred fully transparent to the user with no packet loss
  - A service chain could be constructed of two VNFs
- Issues:
  - IGP SR support still heavily in flux, support not equal between vendors
  - Varying support for SDN protocols
- Future work
  - More SR-MPLS SR-Aware VNFs to construct more varied service chains
  - More in depth performance testing
  - Evaluation for scalability/viability in a production network

University of Amsterdam

SURF