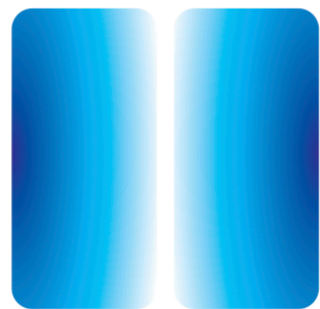


Cost Effective SDN IXP

Marc Bruyere - Christoffel Visser
24th Oct 2019 - UCSB



IIJ INNOVATION INSTITUTE



東京大学
THE UNIVERSITY OF TOKYO

Internet eXchange Point

- An Internet exchange points (IXP) are physical locations through which Internet infrastructure companies such as Internet Service Providers (ISPs) and CDNs connect with each other
- The core service is a **Layer2 switching fabric**
- Layer3 like Route Server are proposed service
- **Very stable and predictable environment**

IXP Costs

**Administrative
Commercial**



**Engineering
Technical**



**Hardware
Infrastructure**



IXP domains actions



Member/Customer Domain

Manage Member connection

Cost level to connect a new member

	Request to Connect	Provisioning	Setup	Validation	Maintenance
Admin	Medium	Low	None	None	None
Technical Human cost	Low to none	Medium	High	High	Medium to High
Physical Equipment	None	Medium to High	Medium	Low to none	Low to none

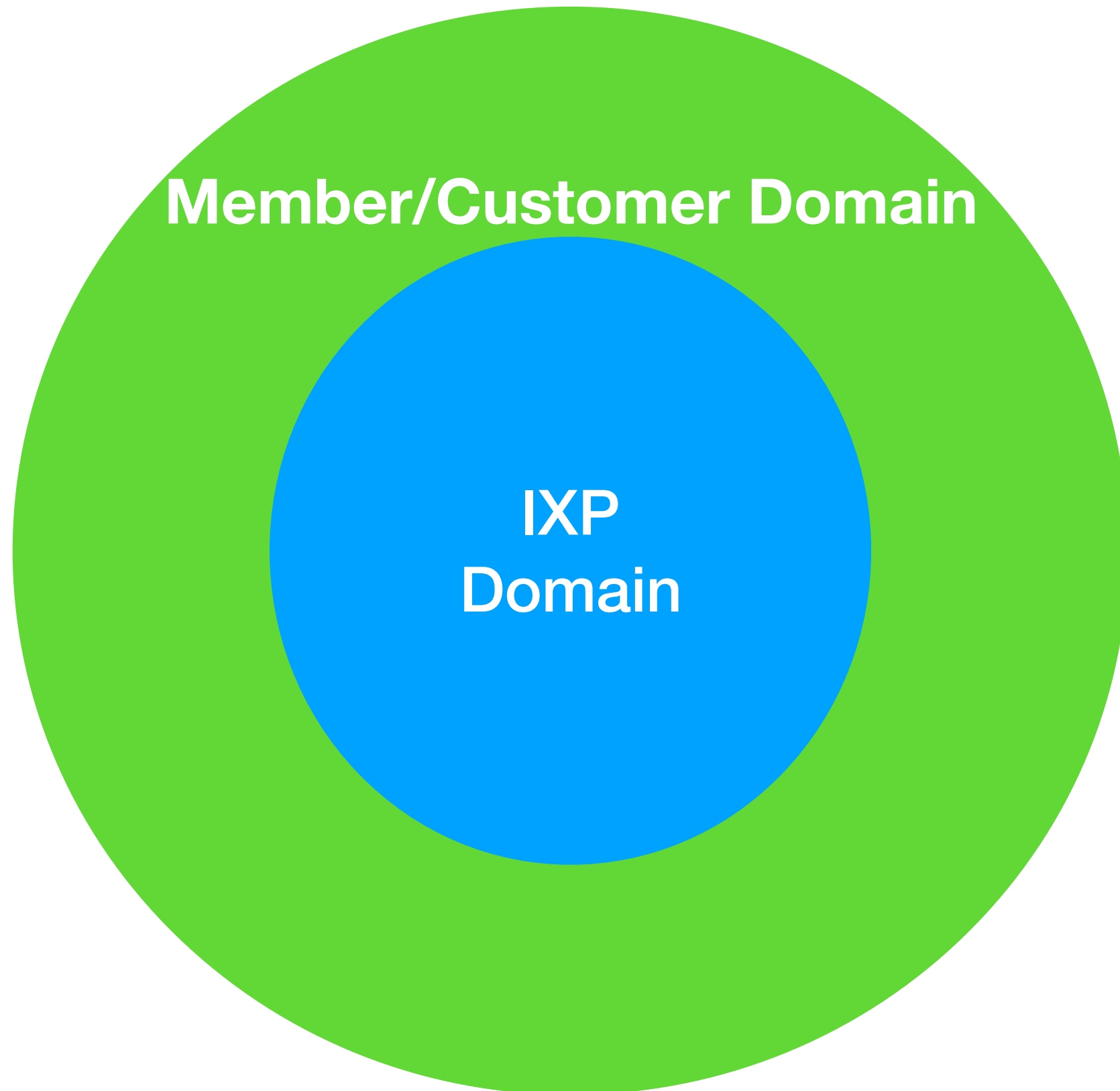
Cost level to connect a new member

	Request to Connect	Provisioning	Setup	Validation	Maintenance
Admin	Medium	Low	None	None	None
Technical Human cost	Low to none	Medium	High	High	Medium to High
Physical Equipment	None	Medium to High	Medium	Low to none	Low to none

Reduction goals

	Request to Connect	Provisioning	Setup	Validation	Maintenance
Admin	Medium	Low	None	None	None
Technical Human cost	Low to none	Low	Low	Low	Medium to low
Physical Equipment	None	Medium to High	Medium	Low to none	Low to none

IXP domains actions



IXP infrastructure cost level

	Design	Stagging	Validation Pre production	Maintenance
Admin	Medium	None	None	None
Technical Human cost	High	High	High	Medium to High
Physical Equipement	High	Medium	Low to none	None

IXP infrastructure cost level

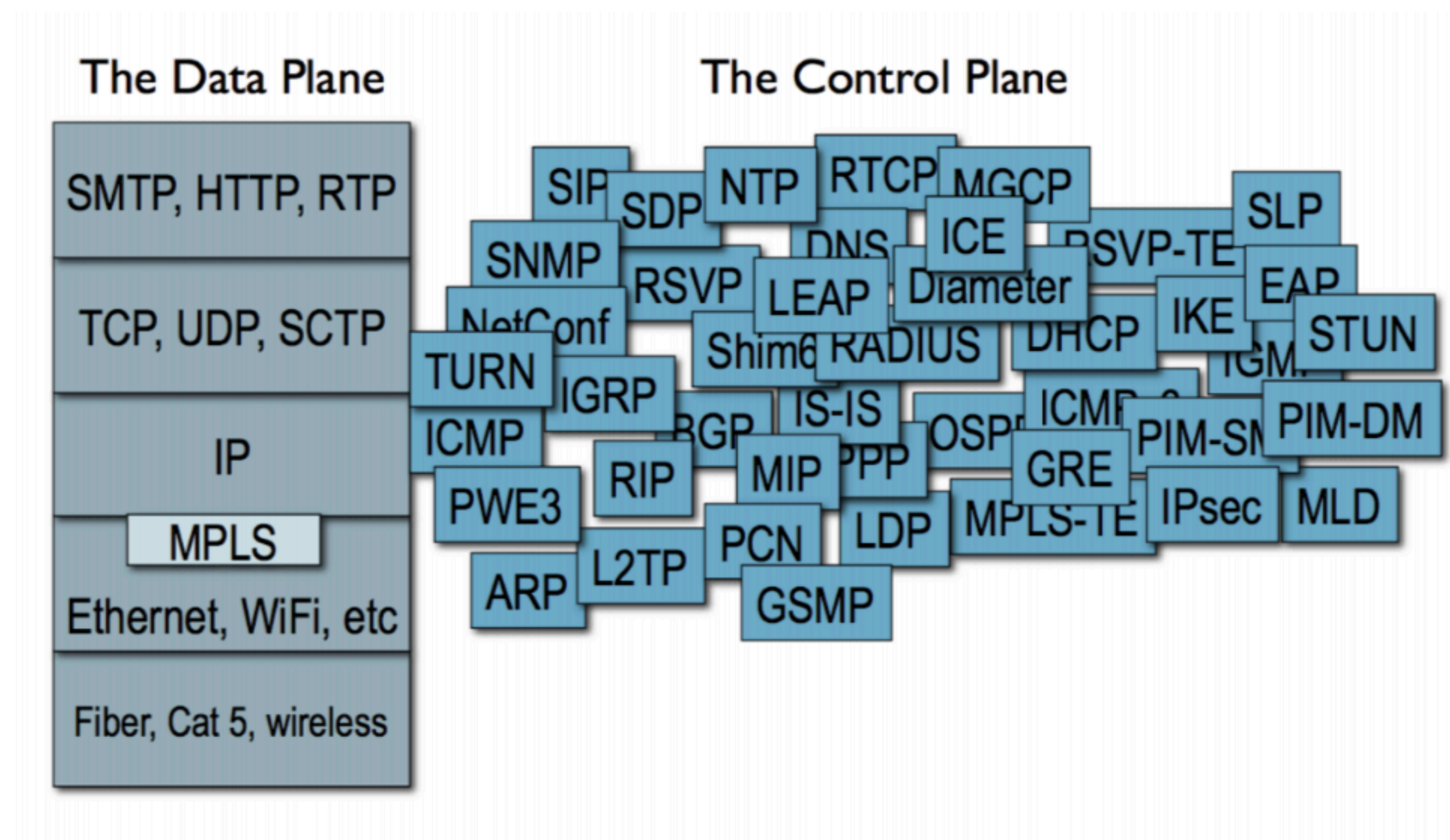
	Design	Stagging	Validation Pre production	Maintenance
Admin	Medium	None	None	None
Technical Human cost	High	High	High	Medium to High
Physical Equipment	High	Medium	Medium	None

Reduction goals

	Design	Stagging	Validation Pre production	Maintenance
Admin	Medium	None	None	None
Technical Human cost	Low	Low	Low	Medium to Low
Physical Equipement	Medium to Low	Low	None	None

Why the costs are high ?

It's not only about the problem of configuring... the level of complexity in networks is staggering



Source

Mark Handley. Re-thinking the control architecture of the internet.
Keynote talk. REARCH. December 2009.

Configuring each element is often done manually, using arcane low-level, vendor-specific “languages”

Cisco IOS

```
!  
ip multicast-routing  
!  
interface Loopback0  
  ip address 120.1.7.7 255.255.255.255  
  ip ospf 1 area 0  
!  
!  
interface Ethernet0/0  
  no ip address  
!  
interface Ethernet0/0.17  
  encapsulation dot1Q 17  
  ip address 125.1.17.7 255.255.255.0  
  ip pim bsr-border  
  ip pim sparse-mode  
!  
!  
router ospf 1  
  router-id 120.1.7.7  
  redistribute bgp 700 subnets  
!  
router bgp 700  
  neighbor 125.1.17.1 remote-as 100  
!  
  address-family ipv4  
    redistribute ospf 1 match internal external 1 external 2  
    neighbor 125.1.17.1 activate  
!  
  address-family ipv4 multicast  
    network 125.1.79.0 mask 255.255.255.0  
    redistribute ospf 1 match internal external 1 external 2
```

Juniper JunOS

```
interfaces {  
  so-0/0/0 {  
    unit 0 {  
      family inet {  
        address 10.12.1.2/24;  
      }  
      family mpls;  
    }  
  }  
  ge-0/1/0 {  
    vlan-tagging;  
    unit 0 {  
      vlan-id 100;  
      family inet {  
        address 10.108.1.1/24;  
      }  
      family mpls;  
    }  
    unit 1 {  
      vlan-id 200;  
      family inet {  
        address 10.208.1.1/24;  
      }  
    }  
  }  
  ...  
}  
protocols {  
  mpls {  
    interface all;  
  }  
  bgp {
```

A single mistyped line is enough to bring down the entire network

Cisco IOS

```
!  
ip multicast-routing  
!  
interface Loopback0  
ip address 120.1.1.1 255.255.255.0  
ip ospf 1 area 0  
!  
!  
interface Ethernet0/0  
no ip address  
!  
interface Ethernet0/0.17  
encapsulation dot1Q 17  
ip address 125.1.17.7 255.255.255.0  
ip pim bsr-border  
ip pim sparse-mode  
!  
!  
router ospf 1  
router-id 120.1.7.7  
redistribute bgp 700 subnets  
!  
router bgp 700  
neighbor 125.1.17.1 remote-as 100  
!  
address-family ipv4  
redistribute ospf 1 match internal external 1 external 2  
neighbor 125.1.17.1 activate  
!  
address-family ipv4 multicast  
network 125.1.79.0 mask 255.255.255.0  
redistribute ospf 1 match internal external 1 external 2
```

Juniper JunOS

```
interfaces {  
  so-0/0/0 {  
    unit 0 {  
      family inet {  
        address 10.0.0.1/24;  
      }  
      family mpls;  
    }  
  }  
  ge-0/1/0 {  
    vlan-tagging;  
    unit 0 {  
      vlan-id 100;  
      family inet {  
        address 10.108.1.1/24;  
      }  
      family mpls;  
    }  
    unit 1 {  
      vlan-id 200;  
      family inet {  
        address 10.208.1.1/24;  
      }  
    }  
  }  
  ...  
}  
protocols {  
  mpls {  
    interface all;  
  }  
  bgp {
```

Complexity + Low-level Management

=

**Problems
And
Higher Cost**

Anything else than 700 creates blackholes

How to reduce cost

Full Zero-Touch

Full control plane testing

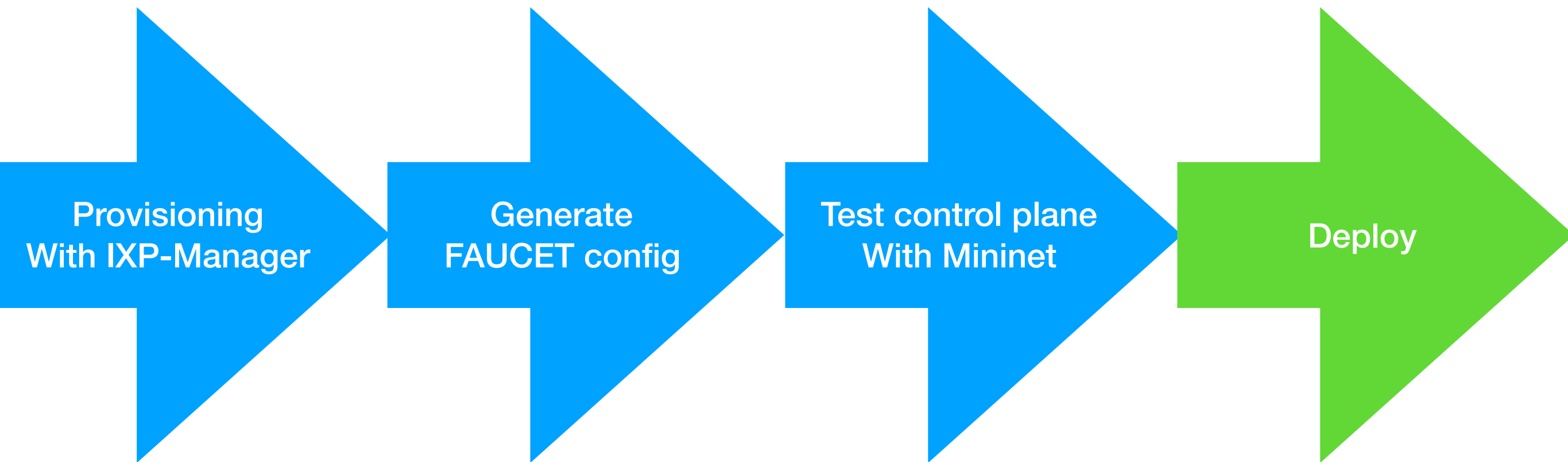
Reduce complexity

IXP-Manager | MX-Graph

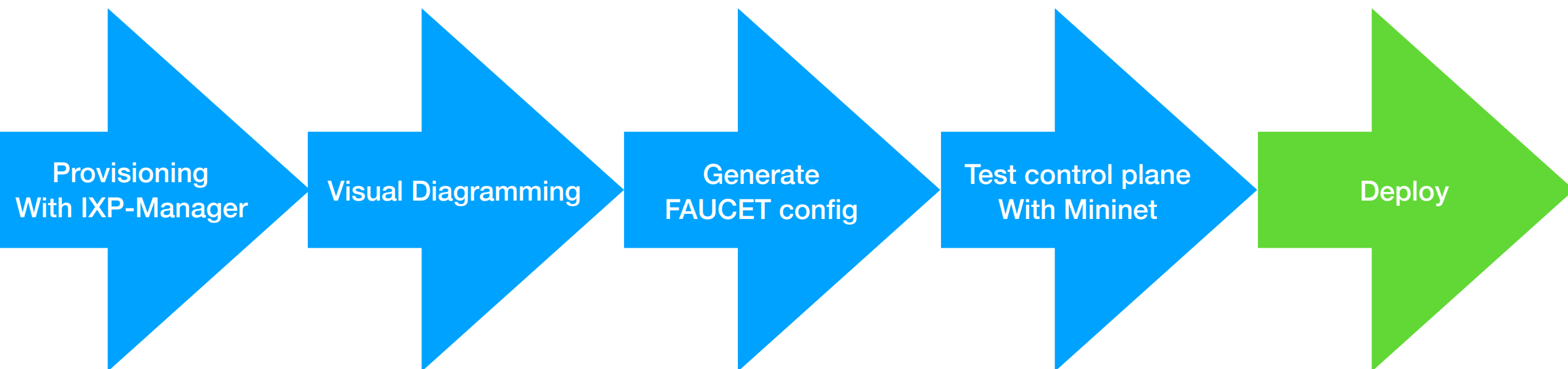
Validation phase with Mininet

Umbrella switching fabric

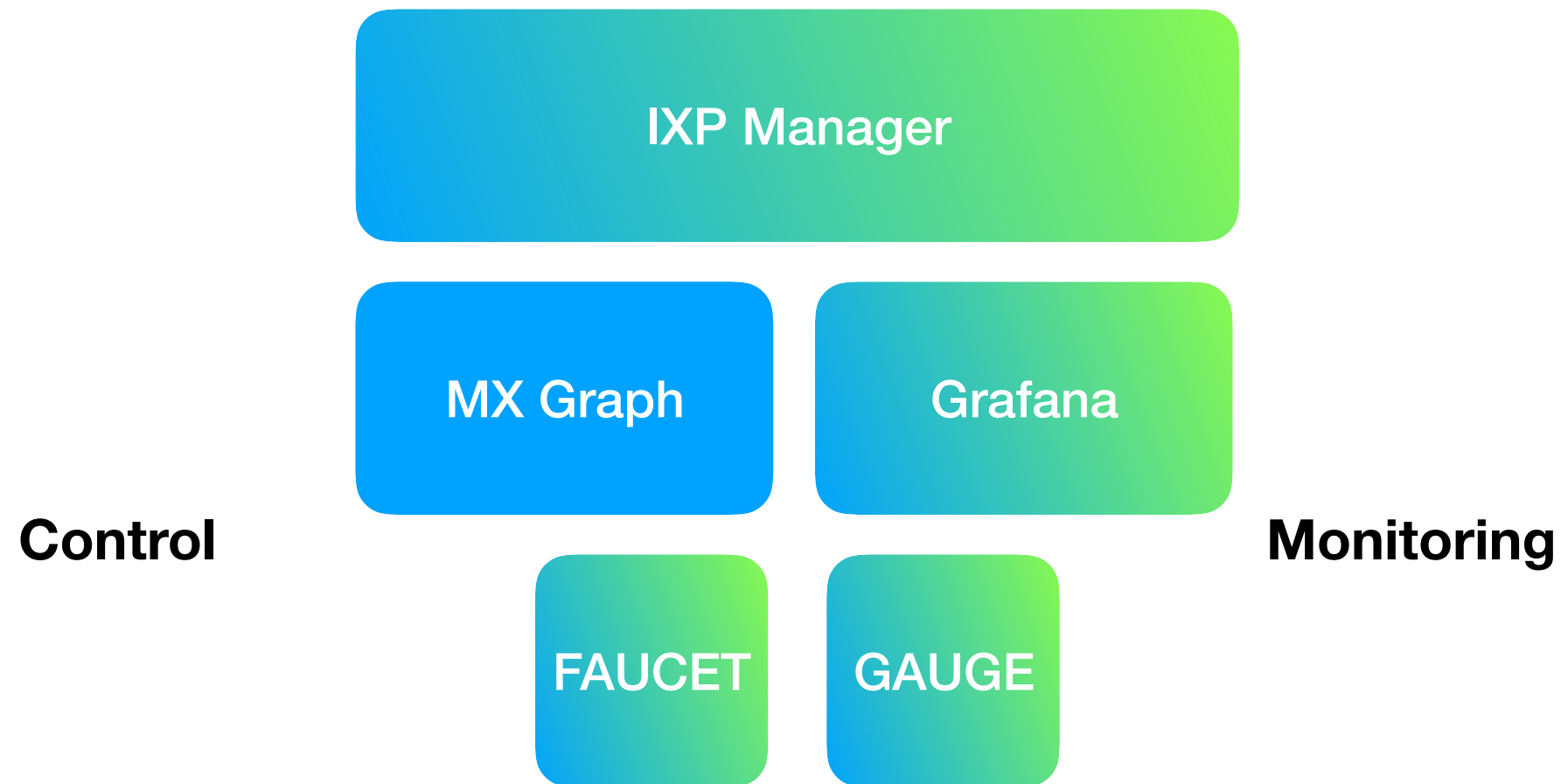
Connecting a new member



Changing the switching fabric



Proposed Architecture



Umbrella switching fabric

Arista

Cisco

Allied
Telesis

Noviflow

HPE Aruba

Umbrella
Core
P4 - Tofino

IXP-Manager



IXP Manager is a full stack management system for Internet eXchange Points (IXPs) which includes an administration and customer portal; provides end to end provisioning; and both teaches and implements best practice.

Powering 81 IXPs

GitHub

@ixpmanager

IXP-Manager

- Best practice
- Zero Touch provisioning
- Advanced Route Server configuration generation
- Excellent service to it's members and IXP operator

Visual Interface to program the data plane with MXGraph

- draw.io Open source core engine MX-Graph Editor
- Collect information through the IXP-Manager API
- Generate shortest path source routing layer2 encoding
- Failure and recovery mechanisms placement

Umbrella switching fabric

- No more Broadcast and perfect edge filtering
- Source routing layer2
- No OpenFlow Core switch **Management-less**
- Scale to thousands of hosts - multi Terabits

Core Umbrella

- P4 programmed - Tofino ASIC
- No CLI
- No Management
- Three Steps process:
 - Check 1st Byte @DST_MAC
 - Shift 1Byte to the left
 - Forward to the port number from the step1 information

Demo

Work in progress

- 2 IXP migration - ToulX @Toulouse and DIXIE @Tokyo
 - Small IXP - triangle topology - 10 operators
- Failure detection and recovery mechanism placement
- IXP-Manager Faucet meta data
- Testing module

Conclusion

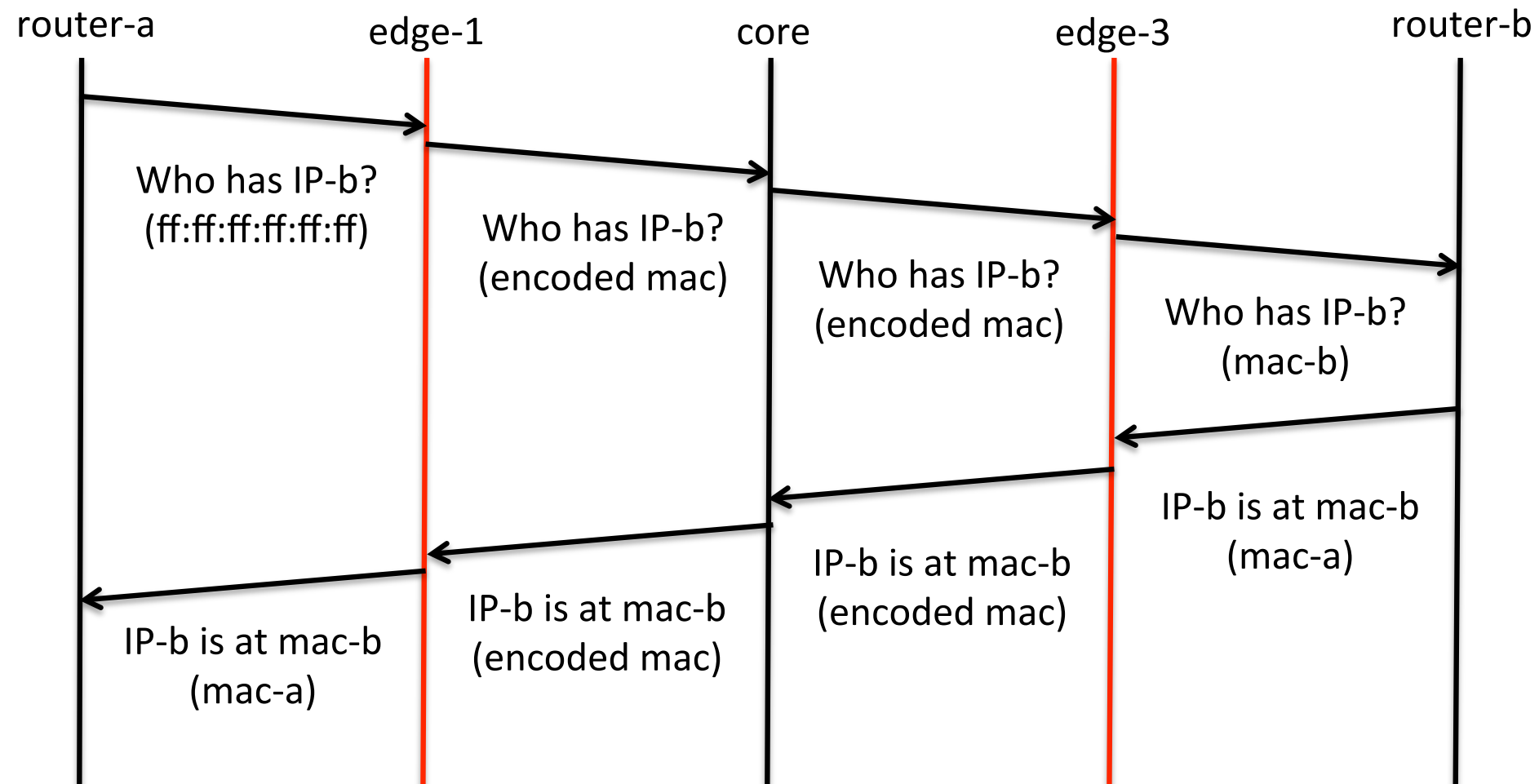
- Simplified management & No high skills requirement
- Zero-Touch Automation & Push-on-green
- Multi-vendor & Management-less Core Switch





marc@ij.ad.jp

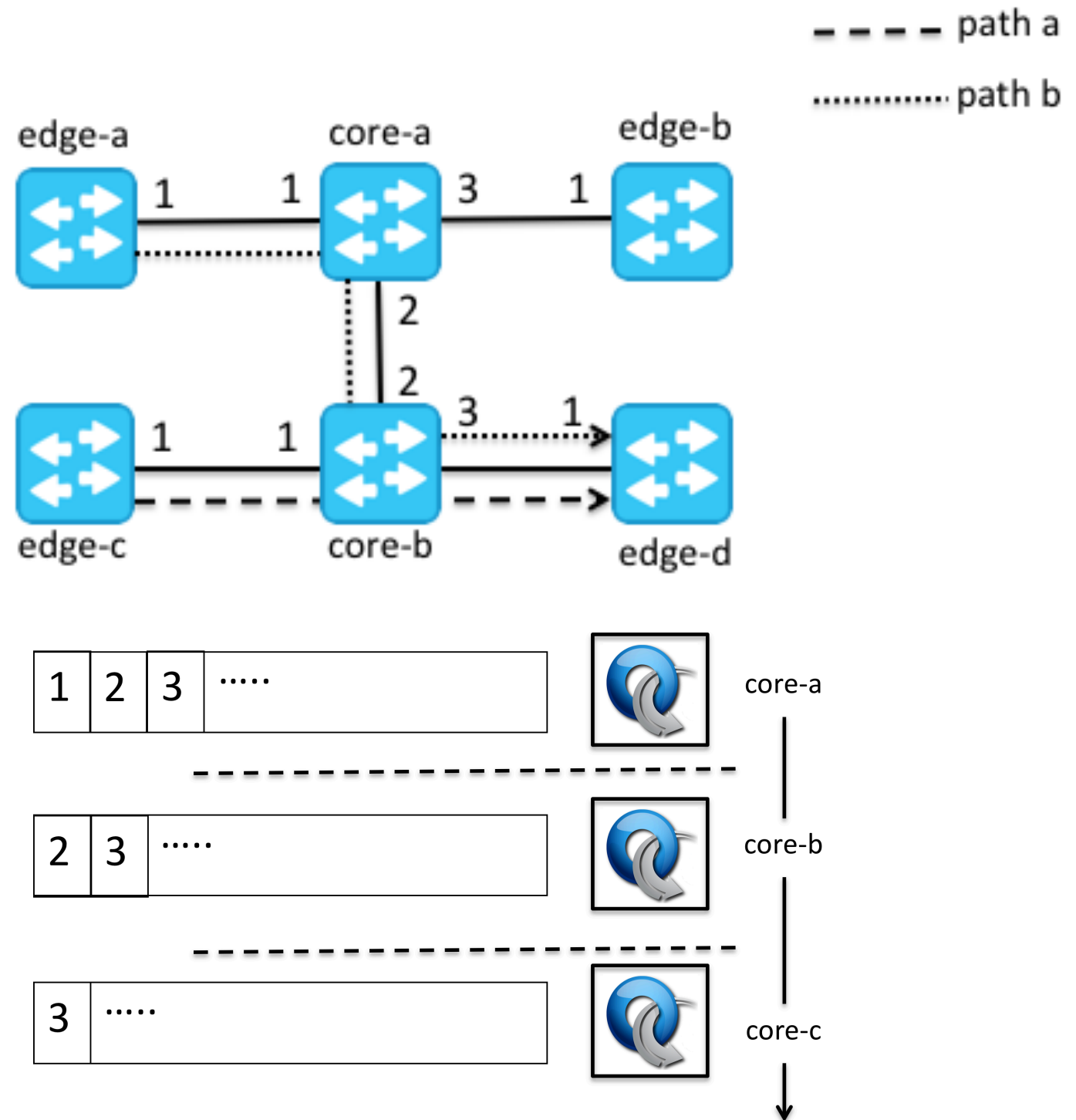
Encoded neighbor discovery in the data plane



Destination broadcast mac
are rewritten in an
encoded mac

Core Umbrella management-less switch

- P4 programmed - Tofino ASIC
- No CLI
- No Management
- Three Steps process:
 - Check 1st Byte @DST_MAC
 - Shift 1Byte to the left
 - Forward to the port number from the step1 information



The FAUCET OpenFlow Controller

- FAUCET is a small and production ready OpenFlow controller
- Very active community a weekly release
 - <https://github.com/faucetsdn/faucet/releases>
- Umbrella architecture support
- A Single YAML file to control all the domain
- FAUCET includes GAUGE a read only monitoring with Grafana dashboards