大阪大学
OSAKA UNIVERSITY

Architecture of Traffic Engineering module for reconfigurable data-plane routers

Sebastian Aguirre, Yoshiyuki Kido, Susumu Date, Shinji Shimojo

Osaka University

PRAGMA 36
Jeju, South Korea
April 24th, 2019

Contents

Research Background

- Strict hierarchy [3]
  - Global transit providers
  - Regional Providers
  - Consumer Networks / Hosting
- New Internet Routing Ecosystem [4-5]
  - Content Providers build own global backbone
  - "Flattening" of the Internet
    - Innovation in new designs for Traffic Engineering

WORLD INTERNET USAGE AND POPU
MARCH, 2019 - New U

World Regions	Population (2019 Est.)	Population % of World	Internet U 25 Mar 2
Africa	1,320,038,716	17.0 %	474,12
Asia	[illegible]	54.7 %	2,190,9
Europe	866,433,007	11.2 %	718,17
Latin America / Caribbean	658,345,826	8.5 %	438,24
Middle East	258,356,867	3.3 %	170,03
North America	366,496,802	4.7 %	326,50
Oceania / Australia	41,839,201	0.5 %	28,43
WORLD TOTAL	7,753,483,209	100.0 %	4,346,50

Figure 1. World Internet Usage [1]

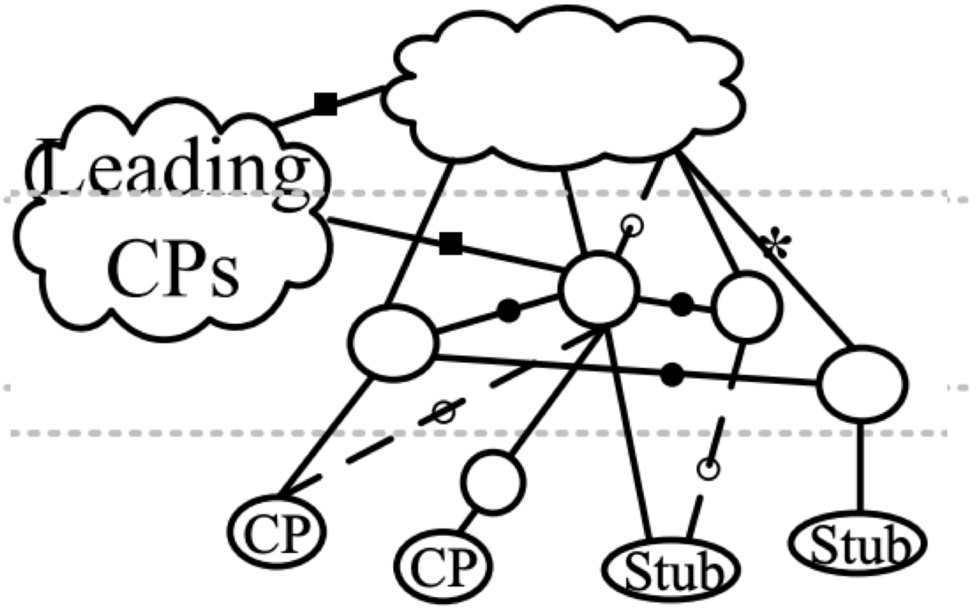

Figure 2. Internet Topology [2]

Research Problem

- Border Gateway Protocol - BGP
  - Routing based on network prefix
  - No application awareness [6]
  - Internet Giants own backbone [7]

- Alternatives
  - Complex to deploy [8]
  - Modified BGP versions [9]
    - Do not scale
  - Bypass BGP peering agreements

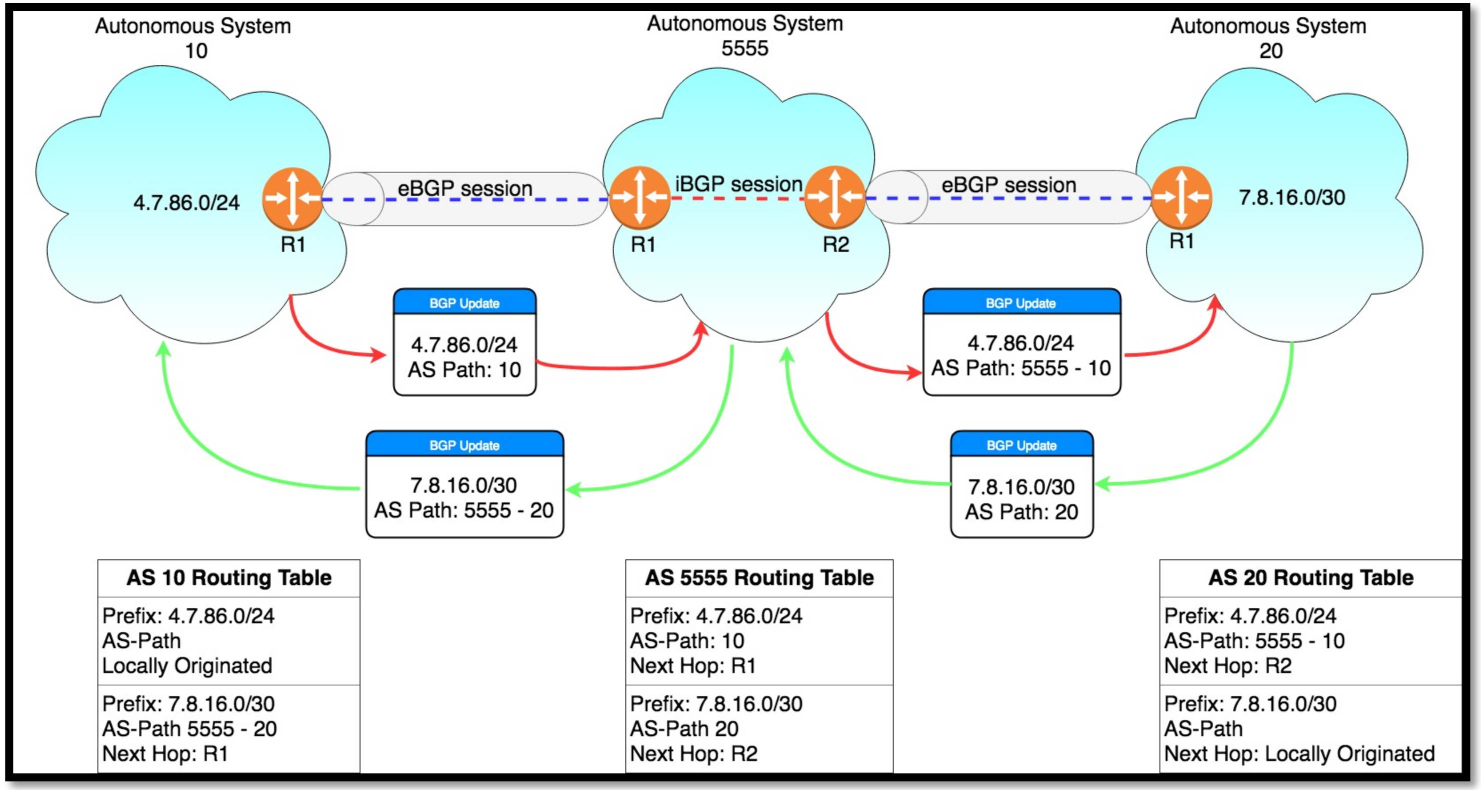

Figure 3. Inter-domain BGP route sharing

Research Question

How can we route inter-domain traffic considering the application layer information without bypassing BGP routing policies?

- Which Interdomain Traffic Engineering guidelines must be considered?

- Which services can benefit from such routing feature? How?

Research Proposal

- Leverage Software Defined Technologies [10]
  - Control plane - forwarding plane separation
  - Centralized Control
- Forwarding plane configuration
  - OpenFlow
  - P4Runtime
- Network Programmability

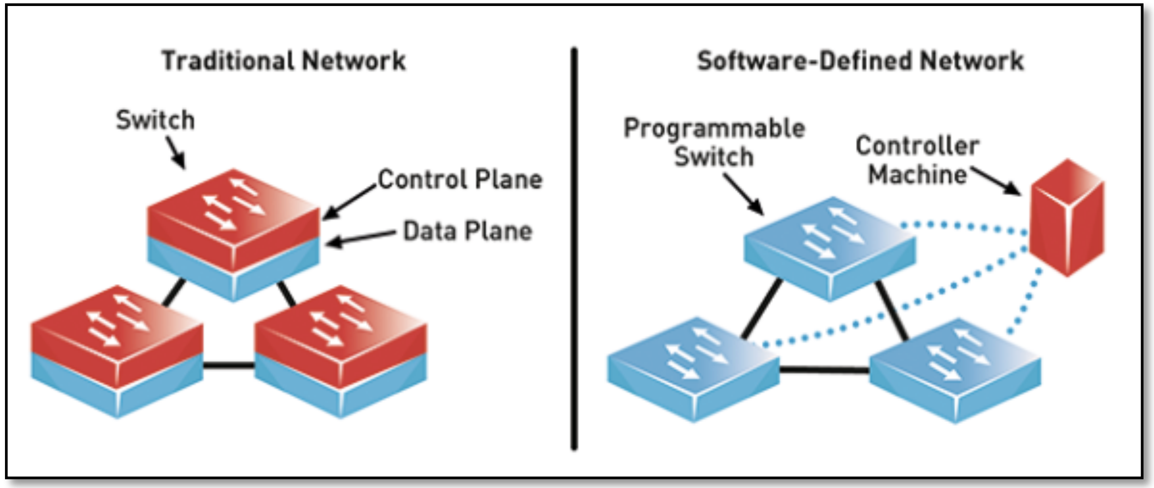

Figure 5. Traditional vs. Software defined network [11]

Traffic Engineering Module

Routing decision and forwarding process are decoupled.

- Software Defined Network
  - BGP Function Virtualization
  - Traffic Engineering Module

Packet Forwarding criteria

- Persist in the data plane
- “Line-rate“ forwarding decision
  - Minimum query to the routing table

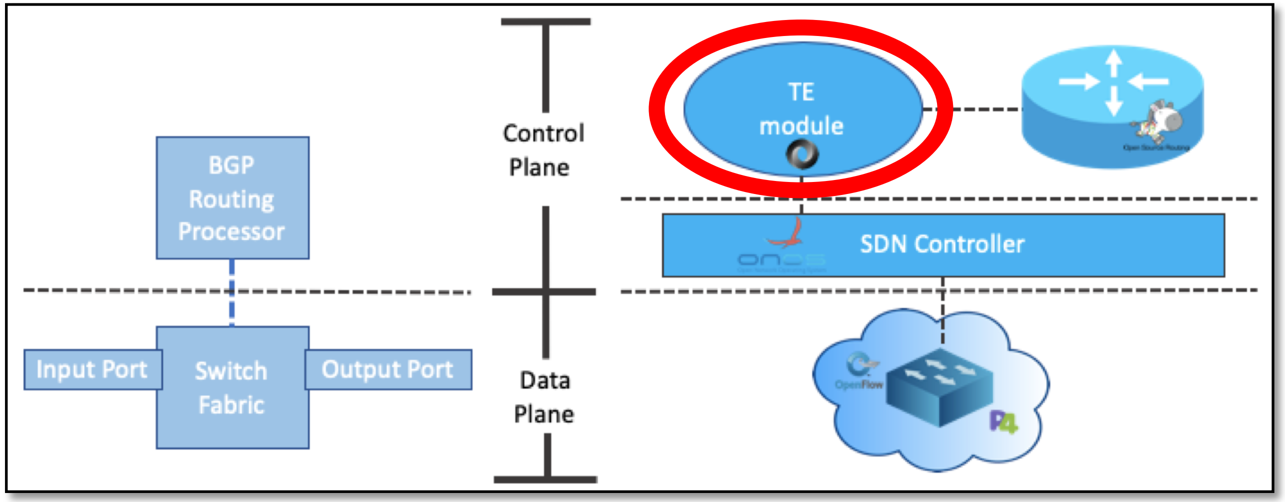

Figure 6. Traditional vs. proposed routing model

TE Module - Components and Interfaces

- BGP Aware!
  - iBGP peering with SDN-Controller
- REST API
  - Override BGP flows with higher priority flows

Flows for Device of:00000000000000a1 (30 Total)

Search | Search By

STATE	PACKETS	DURATION	FLOW PRIORITY	TABLE NAME	SELECTOR	TREATMENT	APP NAME
Added	0	3,335	220	0	IN_PORT:2, ETH_TYPE:ipv4, IPV4_DST:192.168.3.0/24	imm[ETH_DST:00:00:00:00:03:01, OUTPUT:3], cleared:false	*net.intent
Added	0	3,337	1000	0	IN_PORT:3, ETH_TYPE:ipv4, IP_PROTO:1, IPV4_SRC:10.0.3.1/32, IPV4_DST:10.0.3.101/32	imm[OUTPUT:4], cleared:false	*net.intent
Added	2,209	3,337	1000	0	IN_PORT:4, ETH_TYPE:ipv4, IP_PROTO:6, IPV4_SRC:10.0.1.101/32, IPV4_DST:10.0.1.1/32, TCP_DST:179	imm[OUTPUT:1], cleared:false	*net.intent
Added	0	3,337	1000	0	IN_PORT:4, ETH_TYPE:ipv4, IP_PROTO:1, IPV4_SRC:10.0.1.101/32, IPV4_DST:10.0.1.1/32	imm[OUTPUT:1], cleared:false	*net.intent
Added	1,837	3,337	1000	0	IN_PORT:4, ETH_TYPE:ipv4, IP_PROTO:6, IPV4_SRC:10.0.2.101/32, IPV4_DST:10.0.2.1/32, TCP_DST:179	imm[OUTPUT:2], cleared:false	*net.intent
Added	0	207	4000	0	ETH_TYPE:ipv4, IP_PROTO:6, TCP_DST:2000	imm[ETH_DST:00:00:00:00:03:01, OUTPUT:3], cleared:false	TE_Module

Figure 8. Flow Table

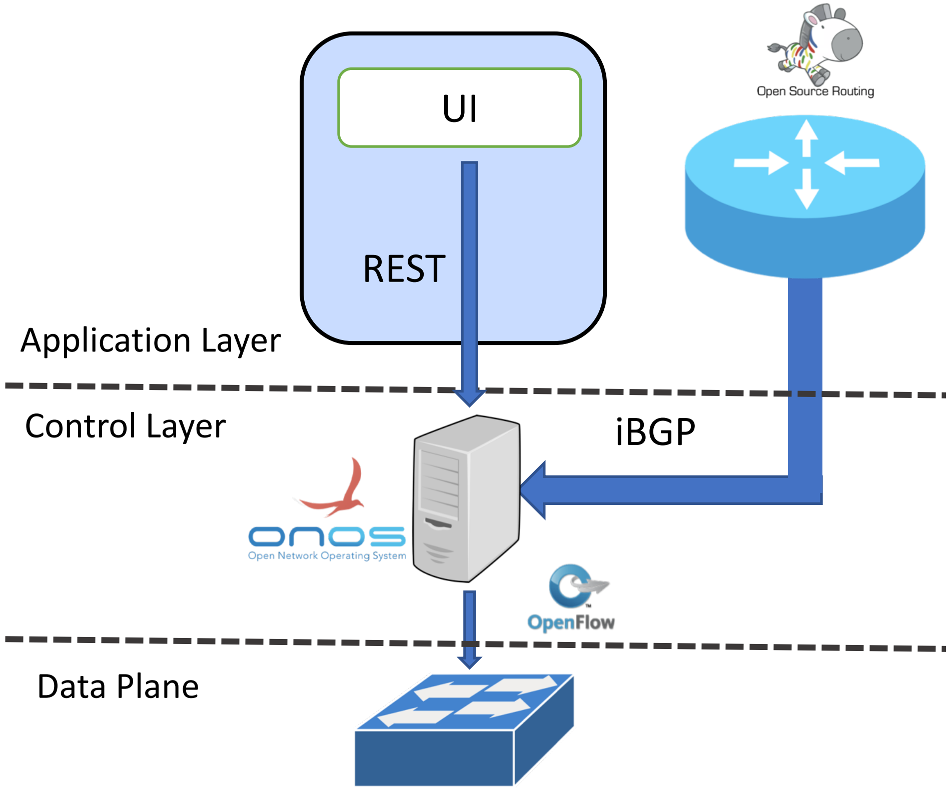

Figure 7. Traffic Engineering Module

Traffic Engineering Guidelines

- BGP peering agreements
  - Routing table updates
  - Incorrect installation of rules
  - Consider valid and best routes

- Learn entire BGP table!
  - Use all valid next hops

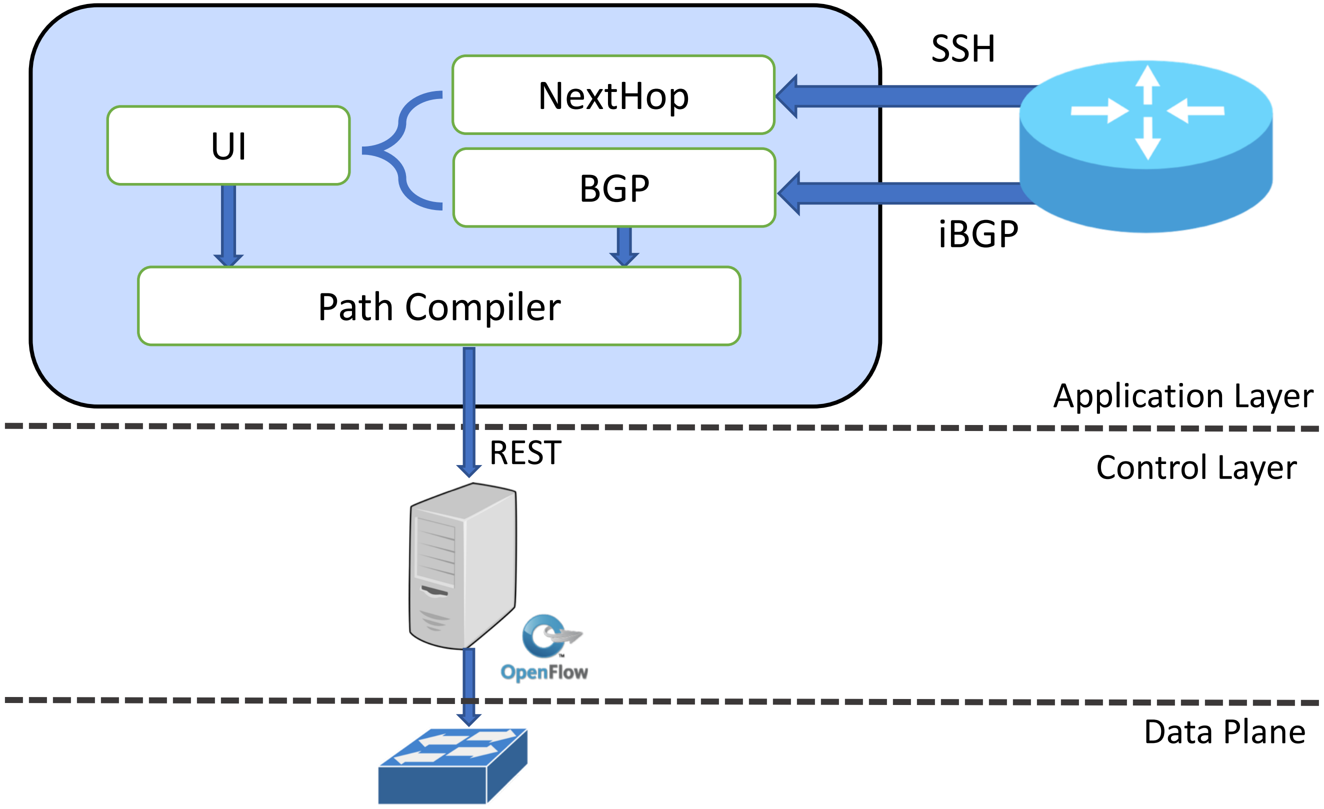

Figure 9. Traffic Engineering Module

Application Specific Peering/SD-WAN

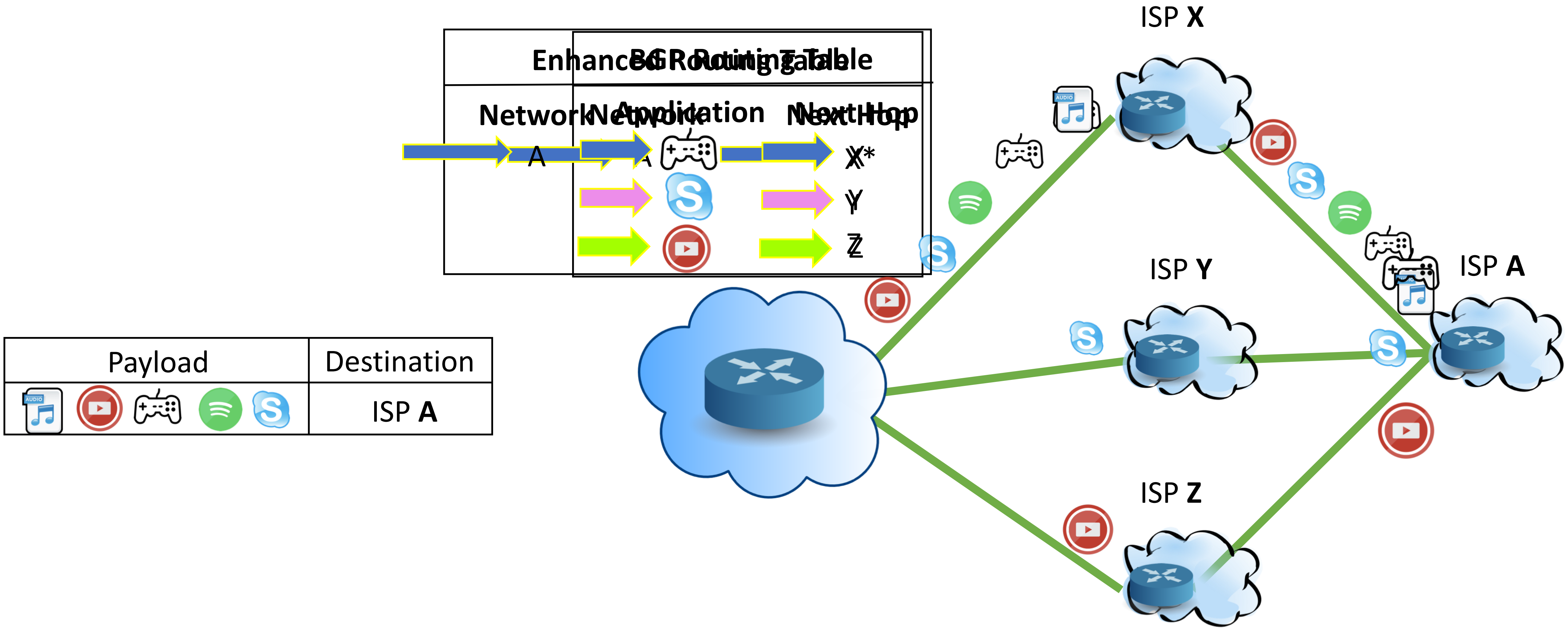

Figure 10. Application Specific Peering

Future Works

- Traffic Engineering Module
    - Link Congestion
    - Other guidelines for the routing decision
        - Upgraded link capacity
        - BGP parameters
        - Popular destinations
        - Traffic volumes
    - Interface with ML/Firewall classification technique
        - Knowledge-Based Network
- Deploy over Pragma-ENT
- Deploy over P4 compatible dataplane

Takeaways

- BGP does not provide mechanisms for traffic engineering considering specific applications.
- SDN allows to decouple the BGP routing decision from the forwarding process.
  - Packet Forwarding criteria persist in the data plane.
- Application specific peering is possible matching the TCP/UDP port number of packet header in data plane match-action tables.

References

[1] - Internet World Stats. WORLD INTERNET USAGE AND POPULATION STATISTICS MARCH, 2019 - New Update. Internet World Stats. www.internetworldstats.com/stats.htm

[2] - Y. Xiang, X. Yin, Z. Wang and J. Wu, "Internet Flattening: Monitoring and Analysis of Inter-Domain Routing," *2011 IEEE International Conference on Communications (ICC)*, Kyoto, 2011, pp. 1-6.

[3] - Craig Labovitz, Scott Iekel-Johnson, Danny McPherson, Jon Oberheide, and Farnam Jahanian. 2010. Internet inter-domain traffic. In Proceedings of the ACM SIGCOMM 2010 conference (SIGCOMM '10). ACM, New York, NY, USA, 75-86.

[4] - Brandon Schlinker, Hyojeong Kim, Timothy Cui, Ethan Katz-Bassett, Harsha V Madhyastha, Ìtalo Cunha, James Quinn, Saif Hasan, Petr Lapukhov, and Hongyi Zeng. Engineering Egress with Edge Fabric. In SIGCOMM, 2017.

[5] - Kok-Kiong Yap, Murtaza Motiwala, Jeremy Rahe, Steve Padgett, Matthew Holliman, Gary Baldus, Marcus Hines, Taeeun Kim, Ashok Narayanan, Ankur Jain, et al. Taking the Edge off with Espresso: Scale, Reliability and Programmability for Global Internet Peering. In SIGCOMM, 2017.

[6] - Arpit Gupta, Laurent Vanbever, Muhammad Shahbaz, Sean P. Donovan, Brandon Schlinker, Nick Feamster, Jennifer Rexford, Scott Shenker, Russ Clark, and Ethan Katz-Bassett. 2014. SDX: a software defined internet exchange. In Proceedings of the 2014 ACM conference on SIGCOMM(SIGCOMM '14). ACM, New York, NY, USA, 551-562.

References

[7] - Brandon Schlinker, Hyojeong Kim, Timothy Cui, Ethan Katz-Bassett, Harsha V Madhyastha, Ìtalo Cunha, James Quinn, Saif Hasan, Petr Lapukhov, and Hongyi Zeng. Engineering Egress with Edge Fabric. In SIGCOMM, 2017.

[8] - Ratul Mahajan, David Wetherall, and Thomas Anderson. 2007. Mutually controlled routing with independent ISPs. In Proceedings of the 4th USENIX conference on Networked systems design & implementation (NSDI'07). USENIX Association, Berkeley, CA, USA, 26-26.

[9] - Matthew Caesar, Donald Caldwell, Nick Feamster, Jennifer Rexford, Aman Shaikh, and Jacobus van der Merwe. 2005. Design and implementation of a routing control platform. In Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation - Volume 2(NSDI'05), Vol. 2. USENIX Association, Berkeley, CA, USA, 15-28.

[10] - Pat Bosshart, Glen Gibb, Hun-Seok Kim, George Varghese, Nick McKeown, Martin Izzard, Fernando Mujica, and Mark Horowitz. 2013. Forwarding metamorphosis: fast programmable match-action processing in hardware for SDN. SIGCOMM Comput. Commun. Rev. 43, 4 (August 2013), 99-110.

[11] – COMMSBUSINESS. Software Defined Networking (SDN) Explained, 2016. Insight. www.commsbusiness.co.uk/features/software-defined-networking-sdn-explained/

Thanks!
감사