



VII Workshop de Tecnologias de Redes do

POP-BA

26 Set 2016

Monitoramento de enlaces de 100Gbps e consistência de regras OpenFlow

Humberto Galiza

Engenheiro de Redes Sênior - NEG AmLight RNP

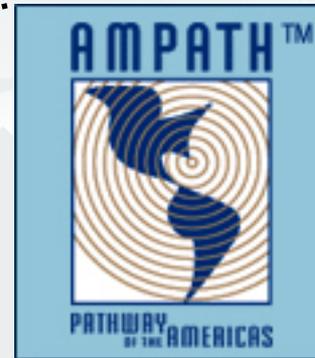
galiza@amlight.net



Quem somos?

AMPATH

- R&E IXP – **RXP** – baseado em Miami (NOTA)
- Principal ponto de encontro entre Redes Acadêmicas Latino Americanas e Norte-Americanas
- Apoio a projetos de pesquisa financiados pela National Science Foundation (NSF)



Quem somos [2]?

AMLIGHT

- Projeto cujo objetivo é instalar, operar, monitorar e evoluir um conjunto de enlaces internacionais interconectando **Pontos de Troca de Tráfego (em inglês, IXP) Acadêmico** em São Paulo (BR) e Santiago (CL) com a AMPATH em Miami

IXP Distribuído

PARCERIA ENTRE

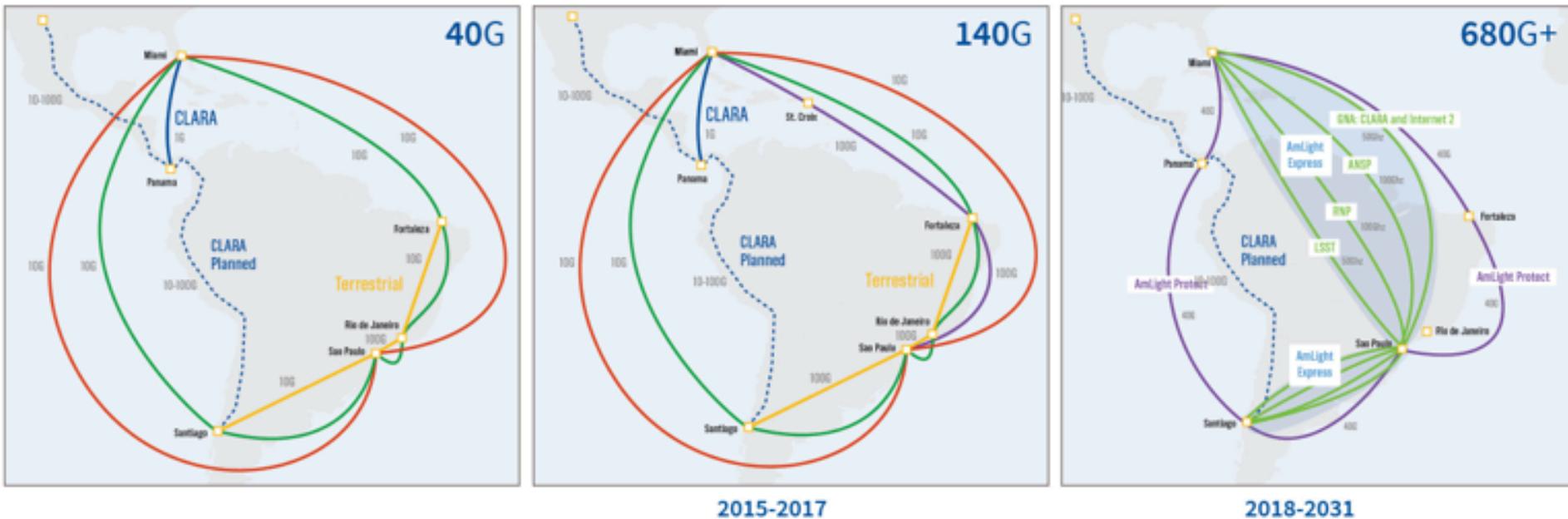
- **ANSP** – Rede Acadêmica de São Paulo
- **AURA** – Associação de Universidades para Pesquisa em Astronomia
- **FIU** - Florida International University
- **NSF** - National Science Foundation
- **RedClara** – Cooperación Latino Americana de Redes Avanzadas
- **RNP** - Rede Nacional de Ensino e Pesquisa
- **REUNA** - Red Universitaria Nacional



Evolução da AmLight

2015: AmLight-EXP: EXPRESS AND PROTECT

Backbone: AMLIGHT: Current to 2031



NSF support for [AmLight](#) Express & Protect is part of a scalable rational architecture, designed to support the needs of the U.S.-Western Hemisphere research and education community that supports the evolving nature of discovery and scholarships.

NSF Award# ACI-1451018

AmLight-Exp: evolução (1/3)

- De 2013 à Dezembro 2015: 6 x 10G
 - 40 Gbps entre America do Sul e Miami
 - 10 Gbps entre São Paulo e Fortaleza/Brasil
 - 10 Gbps entre Brasil e Chile
- Janeiro 2016: 2 x 10G adicionados
 - 60 Gbps entre America do Sul e Miami
 - 10 Gbps entre São Paulo e Fortaleza/Brasil
 - 10 Gbps entre Brasil e Chile

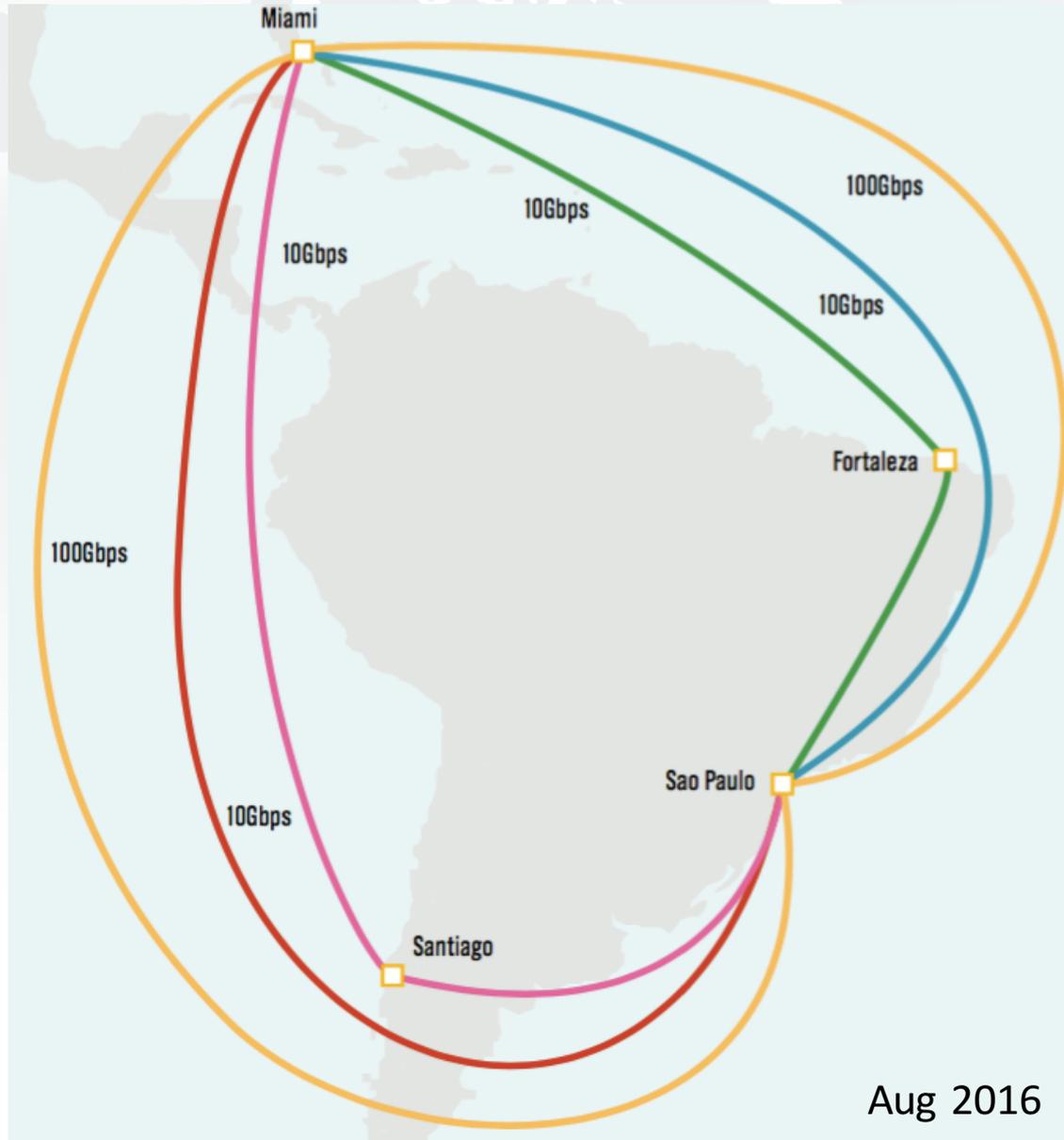
AmLight-Exp: evolução (2/3)

- Abril 2016: 1 x 100G “entregue”
 - Miami x Sao Paulo via Atlântico
 - Primeiro circuito de 100G no cabo SACS
 - Muito instável
 - Problemas com cross-conexões, switches and equipamentos ópticos
- Junho 2016: adicionado mais um lambda 100G
 - Miami x Sao Paulo via Pacífico
 - Segundo circuito de 100G no cabo SACS

AmLight-Exp: evolução (3/3)

- Final de Junho 2016: Finalmente, estabilidade!
 - Módulos 100G Brocade substituídos devido a problemas de interoperabilidade com a CIENA (sistema óptico submarino)
 - Placas da CIENA substituídas e sistema operacional atualizado
 - Cross-conexões no Terremark e cabos IFCs limpos (n-vezes)
- Julho 2016:
 - Upgrade de software nos switches da AMPATH: novos módulos e novo SO
 - 100G links adicionados à rede de produção da AmLight-Exp
 - 2 x 10G links entre São Paulo e Miami decomissionados
 - Capacidade Total Atual:
 - 230 Gbps entre América do Sul e Miami
 - 10 Gbps entre São Paulo e Fortaleza/Brasil
 - 10 Gbps entre Brasil e Chile

AmLight-ExP: Topologia Atual



Validação do 100G

CENÁRIO DE TESTES

OpenFlow + IXIA 100G Traffic Generator

GERADOR DE TRÁFEGO



3/1



MLX-USA

2 Flows: sentido EUA → Brasil

Flow ID: 241 Priority: 32768 Status: Active

Rule:

In Port: e3/1

Action: FORWARD

Out Port: e3/2

RECEPÇÃO DO TRÁFEGO

Flow ID: 242 Priority: 32768 Status: Active

Rule:

In Port: e3/2

Action: FORWARD

Out Port: e3/1

TESTE REFLEXIVO
"LOOP LÓGICO"

3/2

7/1



MLX-BRA

Validação do 100G [2]

CENÁRIO DE TESTES

OpenFlow + IXIA 100G Traffic Generator

GERADOR DE TRÁFEGO



3/1



MLX-USA

3/2

1 Flow: sentido Brasil → EUA

Flow ID: 100 Priority: 32768 Status: Active

Rule:

In Port: e7/1

Action: FORWARD

Out Port: e7/1

Statistics:

Total Pkts: 28816124846

Total Bytes: 14154014509497



TESTE REFLEXIVO
"LOOP LÓGICO"

7/1

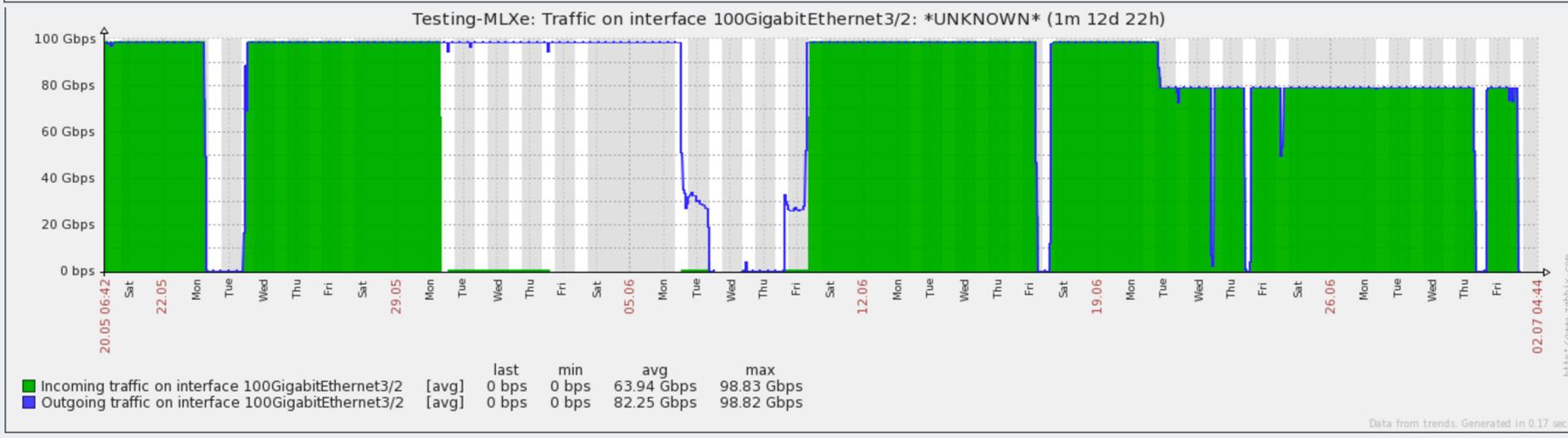
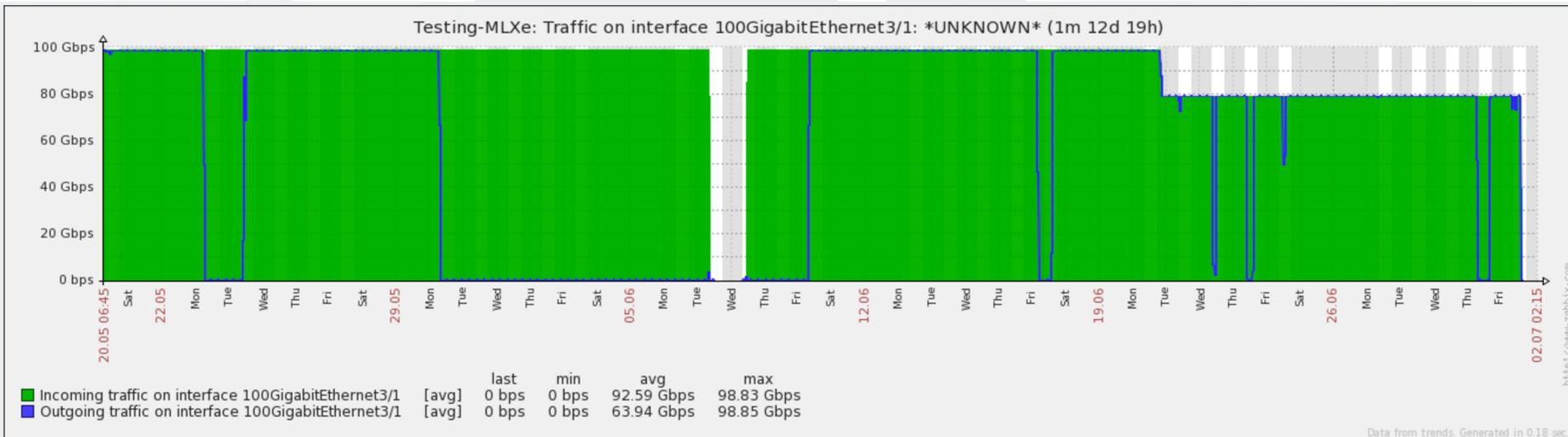


MLX-BRA

10

Validação do 100G [3]

CENÁRIO DE TESTES



Desafios na operação da AmLight

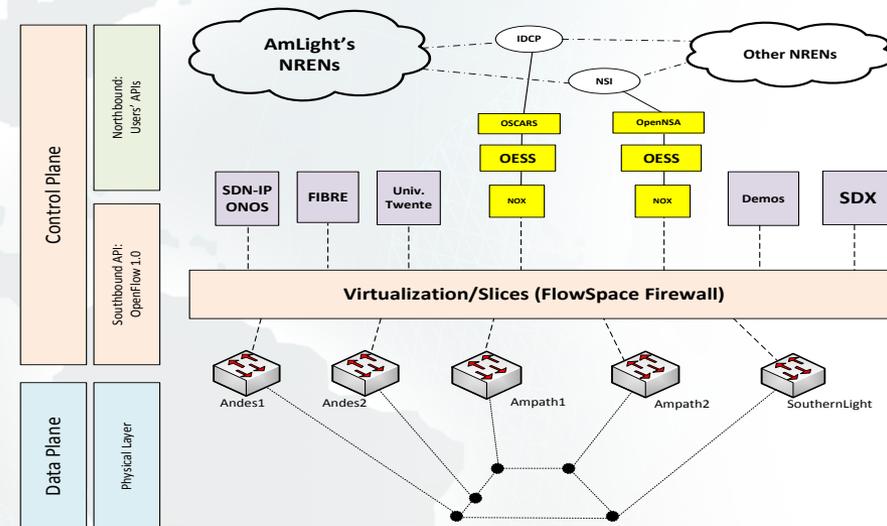
CAMADA FÍSICA

- OpenFlow + Portas 100G: Falta de suporte de tecnologias legadas para validação do caminho físico
 - WAN-PHY, LFM/CFM, UDLD, BFD, etc.
- Espelhamento de tráfego 100G ainda é custoso:
 - Host 100G, TAP, port mirror, etc.) → CAPEX alto

Desafios na operação da AmLight [2]

CAMADA LÓGICA

- Múltiplos usuários e múltiplos controladores em paralelo: dificuldade imensa de troubleshooting!
- Mesmo situações simples se tornam um pesadelo na hora do troubleshooting.
 - Ex: Flow em que o tráfego IPv4 funciona e o IPv6 não.



Como garantir a consistência dos flows e fazer troubleshooting nesse tipo de cenário?

Verificação de Consistência de regras OpenFlow

FERRAMENTAS DE TROUBLESHOOTING

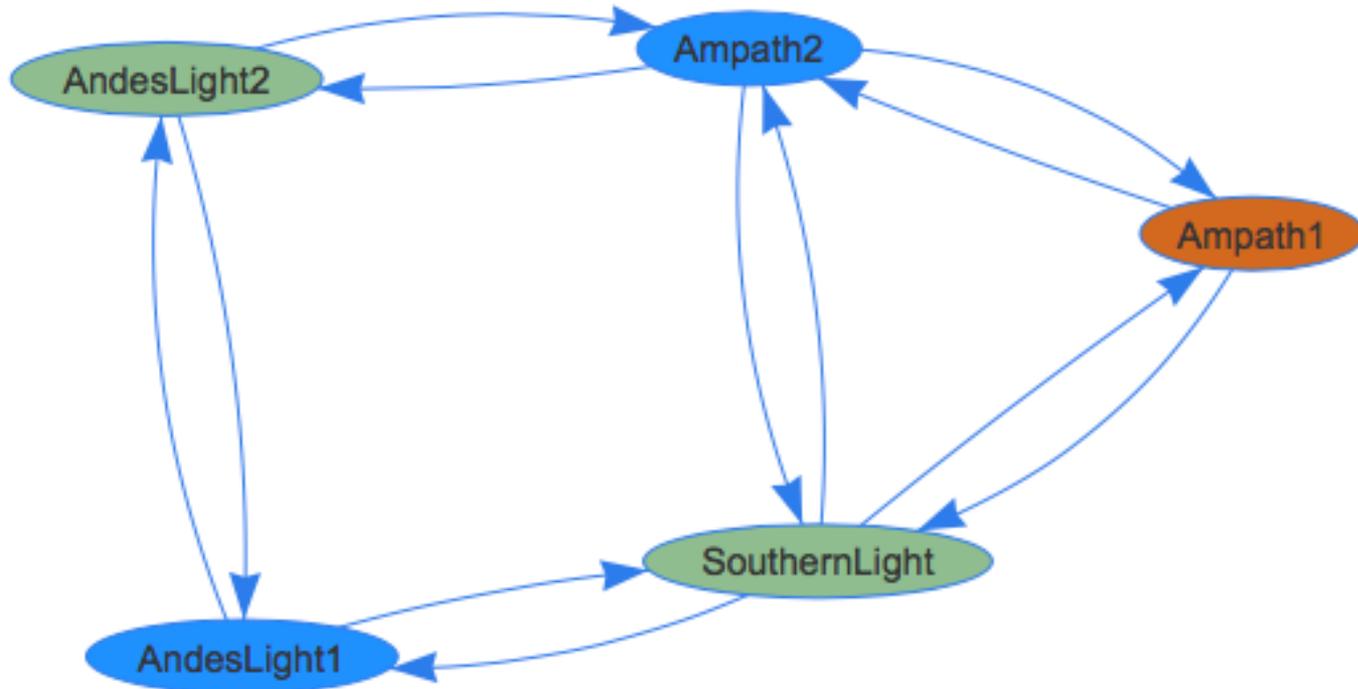
- Troubleshooting em SDN é uma tarefa complexa e ferramentas de visibilidade de rede são necessárias
 - Plano de Controle e Plano de Dados
- A ferramenta mais simples para se ter visibilidade da rede é o traceroute.
 - Ajuda a traçar o caminho que um determinado pacote escolheu pela rede.
- O traceroute convencional é limitado em redes SDN/OpenFlow
 - Seu princípio de funcionamento é baseado no campo TTL de pacotes IP.

Verificação de Consistência de regras OpenFlow [2]

- Solução baseada no artigo:
 - Agarwal, Kanak, et al. "SDN traceroute: Tracing SDN forwarding without changing network behavior." *Proceedings of the third workshop on Hot topics in software defined networking*. ACM, 2014.
 - Software não disponível!
- Características:
 - Não invasivo: não requer modificação nas regras de produção
 - Baixo overhead: requer poucas entradas na TCAM por switch
 - Generico: pode traçar pacotes Ethernet/IP
 - Prático: compatível com a maioria dos switches SDN existentes (ou vSwitch)
- Desenvolvido pela AmLight usando Python + JavaScript + Ryu + OpenFlow
 - Open source: <http://github.com/amlight/SDNTrace>
 - Documentação em progresso!

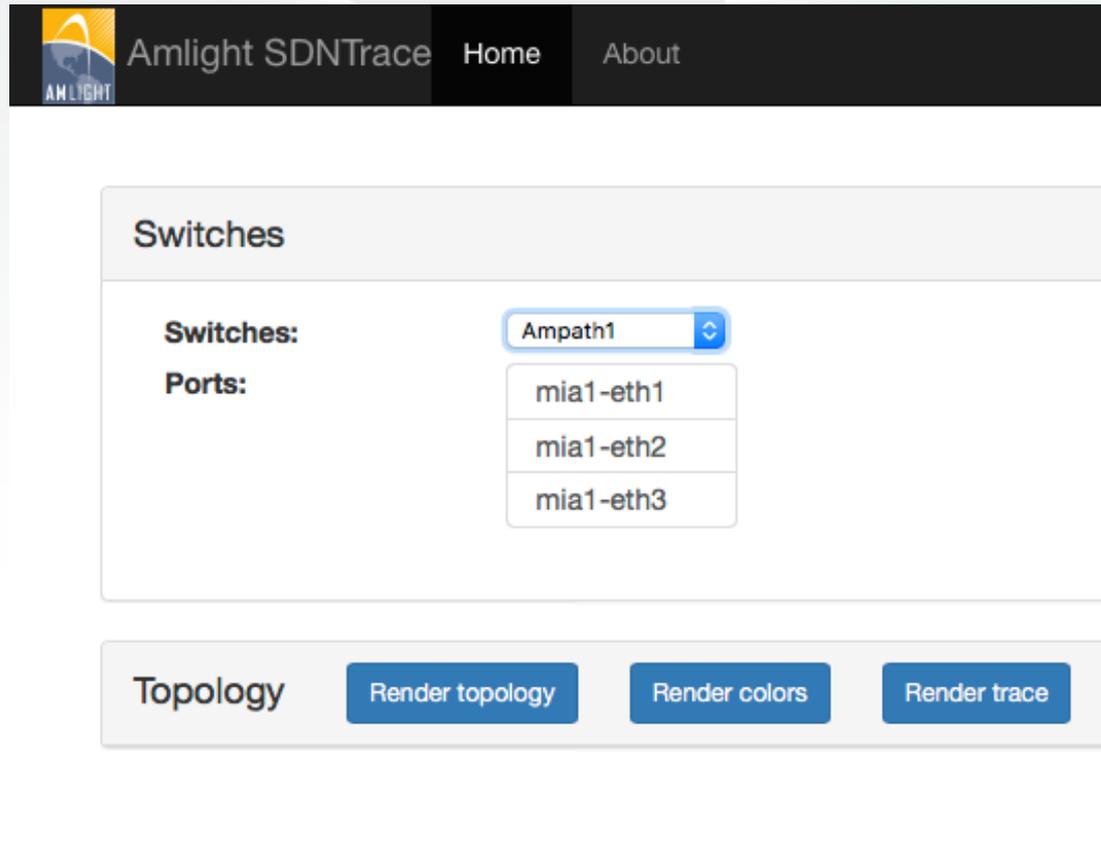
SDNTrace: visão geral

- Faz **coloração de vértices** (Algoritmo de Greedy) da topologia de rede
- Nessa implementação, o campo VLAN priority field (PCP - três bits) são usados para armazenar as cores: 1-green, 2-blue, 3-chocolate, etc.
 - 000 → reservado para o tráfego de produção (sem tagged)
- Instalação de regras OpenFlow nos switches: *envie para o controlador if tag == cor do switch adjacente*



SDN Trace: visão geral [2]

- Pesquisador/Engenheiro/Usuário acessa uma interface web
- Seleciona o switch e a porta de origem



The screenshot shows the Amlight SDNTrace web interface. At the top, there is a navigation bar with the Amlight logo and links for "Home" and "About". The main content area is titled "Switches" and contains a form for selecting a switch and its ports. The "Switches:" label is followed by a dropdown menu currently showing "Ampath1". Below this, the "Ports:" label is followed by a list of three options: "mia1-eth1", "mia1-eth2", and "mia1-eth3". At the bottom of the interface, there is a "Topology" section with three buttons: "Render topology", "Render colors", and "Render trace".

SDN Trace: visão geral [3]

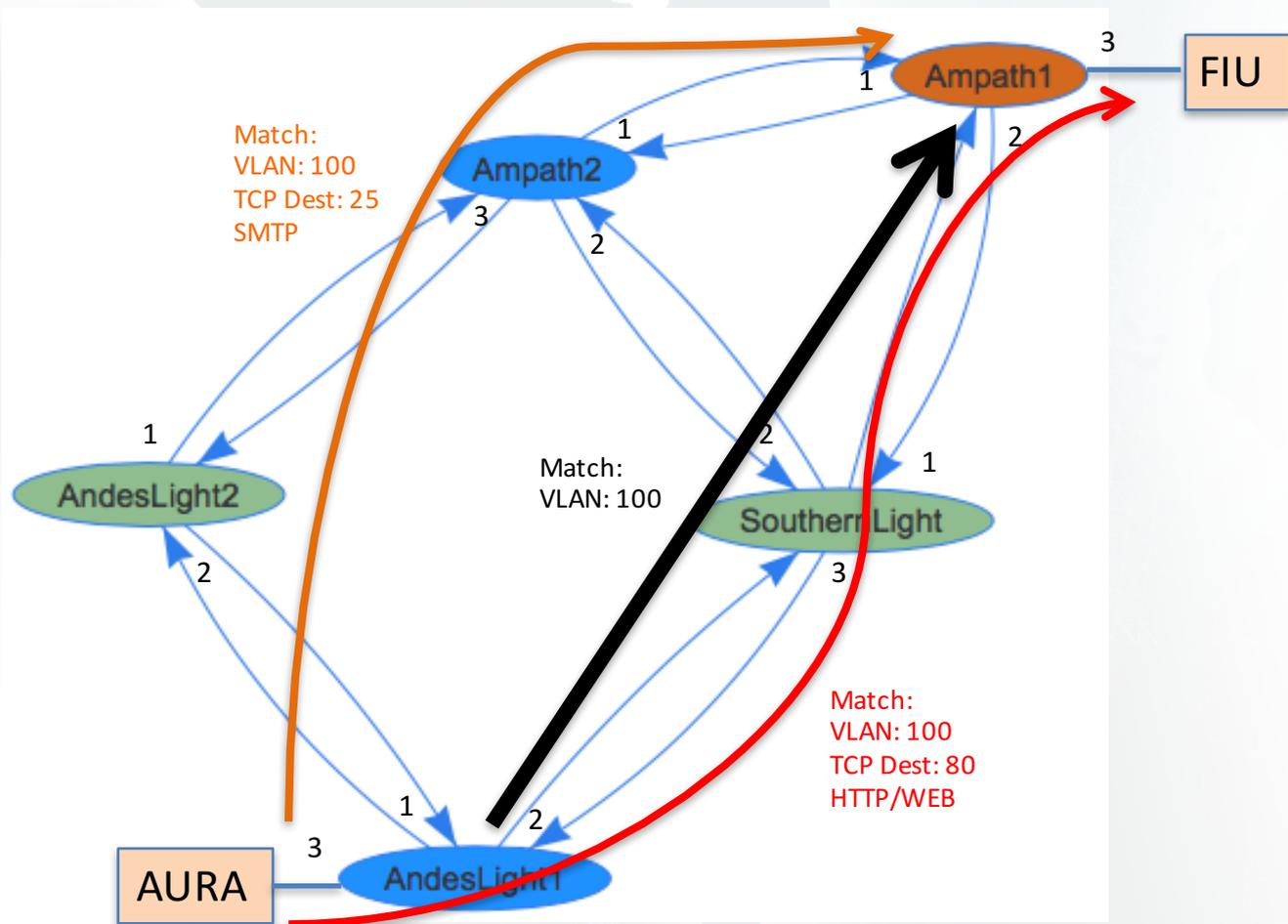
- Usuário cria um pacote "probe" arbitrário
- E escolhe o tipo de trace: Layer-2, Layer-3, ou ambos

Trace Layer 2	Trace Layer 3	Trace Full
MAC Origem	MAC Origem	
MAC Destino	MAC Destino	
VLAN	VLAN	
Ethertype	Ethertype	
	Trace Layer 2	

AmLight SDN Data Plane

Legado: VLAN – all traffic, single path

SDN: "matches" complexos: customização de perfis de tráfego



Integração com outras ferramentas de monitoramento

- SDNTrace API permite:
 - Rodar um trace a partir do seu próprio controlador
 - Obter a topologia da rede
- Roadmap:
 - Expandir o SDN Trace para se tornar um SDN Looking Glass
 - Traceroute, visualização, tráfego e utilização de porta, black hole detection, etc.
 - Lançamento: final deste ano

Conclusão

LIÇÕES APRENDIDAS: 100G

- Não espere que a ativação do seu primeiro 100G em uma rede de terceiros será fácil – porque não será!
- Higiene é uma OBRIGAÇÃO: fibras e conectores tem que estar perfeitamente limpos!
- Um gerador de tráfego torna a sua vida muito mais fácil
- Crie um "baseline" com pacotes de diferentes tamanhos para entender os limites do seu equipamento

Conclusão

LIÇÕES APRENDIDAS: Operação e OpenFlow

- Tecnologias já consolidadas ainda não são suportadas em portas 100G com OpenFlow: mais capacidade e mais desafios
- A combinação entre 100G e OpenFlow requer ainda mais atenção para a verificação de consistência de regras na rede
- O SDNTrace consegue debugar flows arbitrários e pacotes sem necessitar de nenhum tipo especial de suporte além do OpenFlow “básico.”
- A característica chave da ferramenta é que ela não requer nenhum tipo de modificação nas regras de produção.



VII Workshop de Tecnologias de Redes do

POP-BA

26 Set 2016

Obrigado! Perguntas?

Monitoramento de enlaces de 100G e consistência de regras OpenFlow

Humberto Galiza
galiza@amlight.net

