Network Research Exhibition: the Future of Networking and Computing with Big Data Streams
November 11th, 2022

Demonstrating PolKA routing approach to support traffic engineering for data-intensive science

Magnos Martinello[2], *Cristina Klippel Dominicini*[1], *Rafael S. Guimarães*[1], *Everson Borges*[2], *Moises R. N. Ribeiro*[2], *and Edgard Cunha*[2]

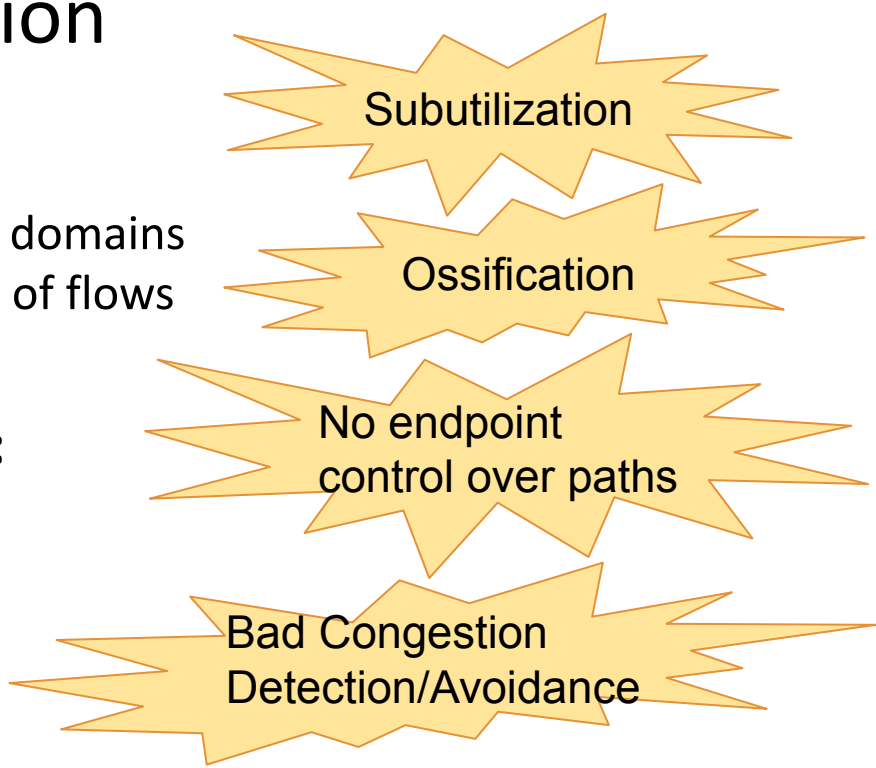
[1]*Federal Institute of Espírito Santo,* [2]*Federal University of Espírito Santo*

INSTITUTO FEDERAL
Espírito Santo

UFES

Motivation

- **Scientific Applications**:
  - High-speed WAN networks & Multiple domains
  - Massive data transfer & Large number of flows
  - E2E reliability
- **Table-based forwarding bottlenecks:**
  - Large number of states
  - Set of shortest paths
  - Latency for path configuration
- Alternative: **Source Routing (SR)**
  - A source specifies a path and adds a route label to the packet header.



Major Research and Development Challenges

- A Source Routing approach that meets the requirements:

open source/ interoperable	no tables in the core	support in prog. switches	fixed length header	topology agnostic multipath routing

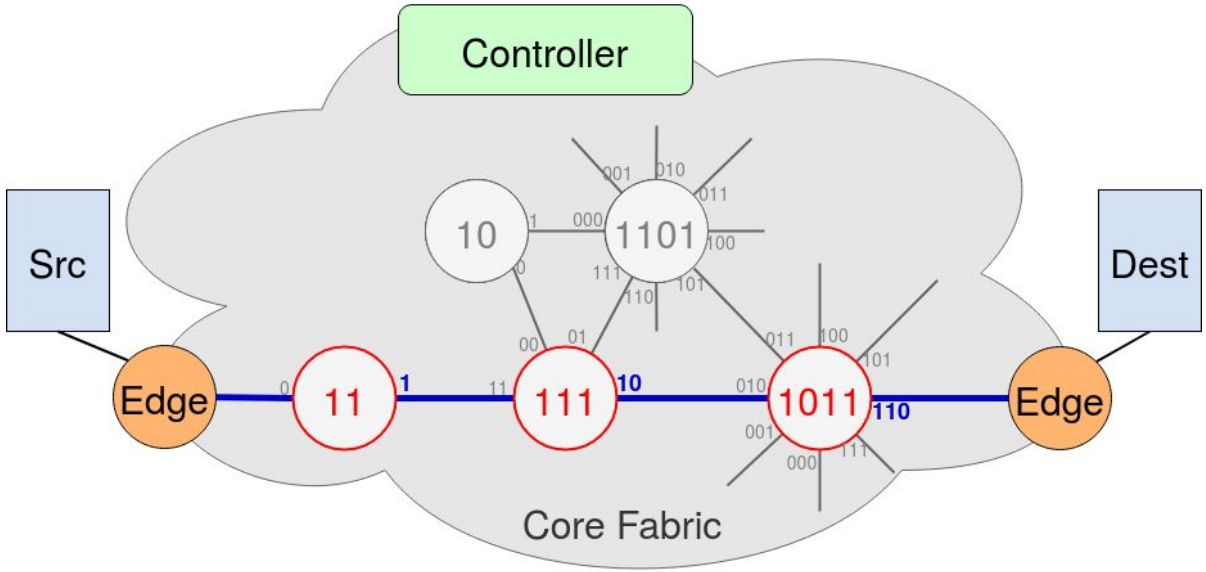
- PolKA: Polynomial Key-based Architecture for Source Routing
  - Polynomial Residue Number System (**RNS**)
  - Chinese Remainder Theorem (**CRT**)
  - Packet forwarding based on mod operation: **remainder of division**

$$portID = < routeID >_{nodeID}$$

  - mod computation reuses CRC operation (custom in Tofino switch)

How does PolKA work?

- The **Controller** chooses a **path** for a specific flow:
  - A set of switches: **nodeIDs**: {0011,0111,1011}
  - and their output ports: **portIDs**: {1 , 10, 110}



nodeID polynomials

$$s_1(t) = t + 1 = 11$$

$$s_2(t) = t^2 + t + 1 = 111$$

$$s_3(t) = t^3 + t + 1 = 1011$$

portID polynomials

$$o_1(t) = 1$$

$$o_2(t) = t = 10$$

$$o_3(t) = t^2 + t = 110$$

Calculate routeID with CRT

$$t^4 \equiv 1 \mod (t+1)$$

$$t^4 \equiv t \mod (t^2+t+1)$$

$$t^4 \equiv (t^2+t) \mod (t^3+t+1)$$

$$t^4 = 10000$$

portID = < routeID >$_{\text{nodeID}}$

$1 = \langle 10000 \rangle_{0011}$

$10 = \langle 10000 \rangle_{0111}$

$110 = \langle 10000 \rangle_{1011}$

Innovations to be demonstrated

- Data plane
  - Source Routing with Stateless Core
  - Forwarding at line rate by **reusing CRC in P4** programmable switches

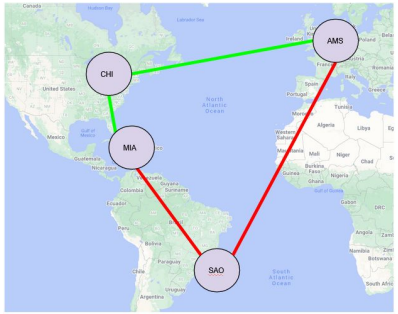
- Control plane
  - Easy to configure tunnels
  - **Integrated** in the Freerouter platform

- Potential to support:
  - Transfer of big data streams with aggregation of multiple flows
  - Big pipes/tunnels dynamically configured in the underlay network

Come to see our demonstrations !!

- **Big data streams at 100 Gbps**
  - PolKA@ Caltech P4 lab testbed
  - Multiple aggregated TCP flows steered to pre-configured tunnels
    - A route label represents paths in the underlay network
  - Comparison with Segment Routing: Both achieving line rate (100 Gbps)

- **Dynamic Traffic Steering at Intercontinental Testbed**
  - PolKA@ Global P4 lab testbed
  - Tunnel Setup for Traffic Engineering
    - Stateless Core
    - Define a explicit path at the Edge (route label)
    - Migration to another tunnel requires a single update at the Edge



Thank you for attention !

Magnos Martinello

magnos.martinello@ufes.br

** This work was a recipient of the 2021 Google Research Scholar and the 2022 Intel Connectivity Research Grant (Fast Forward Initiative) Awards, and Received funds from CAPES, CNPq, FAPESP, FAPES, CTIC, and RNP.*

INSTITUTO FEDERAL
Espírito Santo

UFES