salzburgresearch

Ferdinand von Tüllenburg

Layer-2 Failure Recovery Methods in Critical Communication Networks

Dependable Communication for Critical Infrastructures

Electricity

Health

Transport

Finance

Dependable Communication is required
Outages lead to failures, degraded services

In future rising complexity:

- Interconnection / growing of distinct CI
- Massive inclusion of sensors, actuators, mobile devices
- To create new services / businesses
- Also over long distances (WAN)

- Need for
  - Standardization of communication
  - Flexibility and programmability
  - Simpler maintainability / management
  - Enhance Dependability of Communication

What means Dependable Communication?

Reliability / Availability

- Perform required functionality for a period of time

Required functionality

- Quality of Server (QoS)
- Assured Service

Means of Dependable Communication

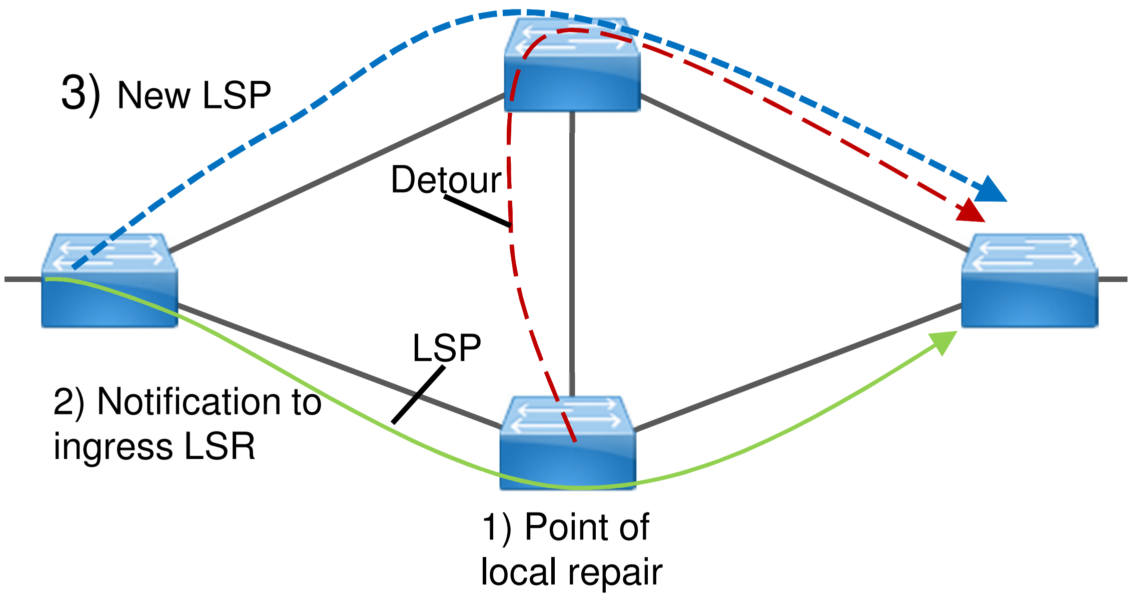
- Fault Tolerance
- Fault Detection/Isolation
- Fault Avoidance
- Fault Restoration

Focus on fault tolerance mechanisms

- Main idea: Reroute traffic quickly when fault occurs
- 3 Approaches:
  - RSVP-TE
  - RSTP
  - OpenFlow Fast Failover Groups

RSVP-TE Fast Reroute

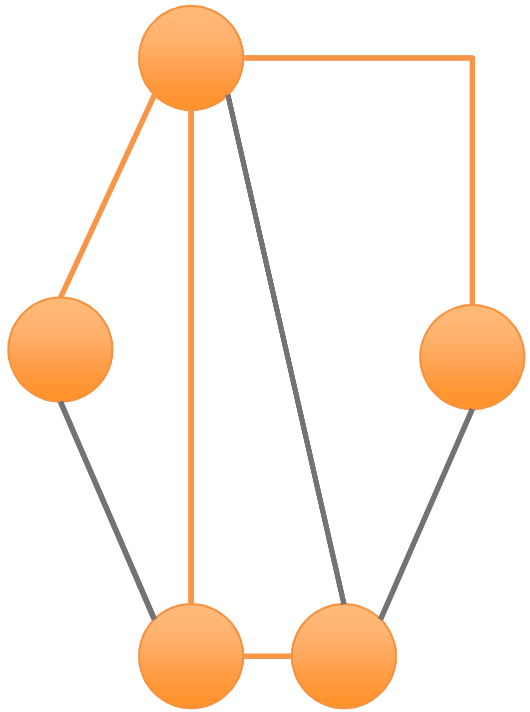
- MPLS Approach
  - Packets are labeled at ingress-routers
  - Labled packets are fast-switched at core routers via LSP
  - Including Resource Reservation
- Proposed for
  - Bandwidth separation, traffic separation, increasing reliability
- RSVP-TE Fast Reroute Operation
  - Pre-computation / pre-establishment of several detours
  - Detours provide local repair capabilities



- Performance
  - Local Repair: Several ms
  - Failure Detection: ms to s

Rapid Spanning Tree Protocol

- Layer-2 protocol used for loop avoidance
- Bridges build Minimal Spanning Tree (MST)
  - Cost based
- Redundant links are used as backup
  - Root Ports: Forwarding port. Best connection to root bridge
  - Designated Ports: Forwarding port to a network segment
  - Alternate / Backup Ports: blocked port to another network segment.
    - Can be quickly activated.
- In case of link failure:
  - Topology change message is generated (by detecting node)
  - New spanning tree is computed
  - After computation: Fast switch over



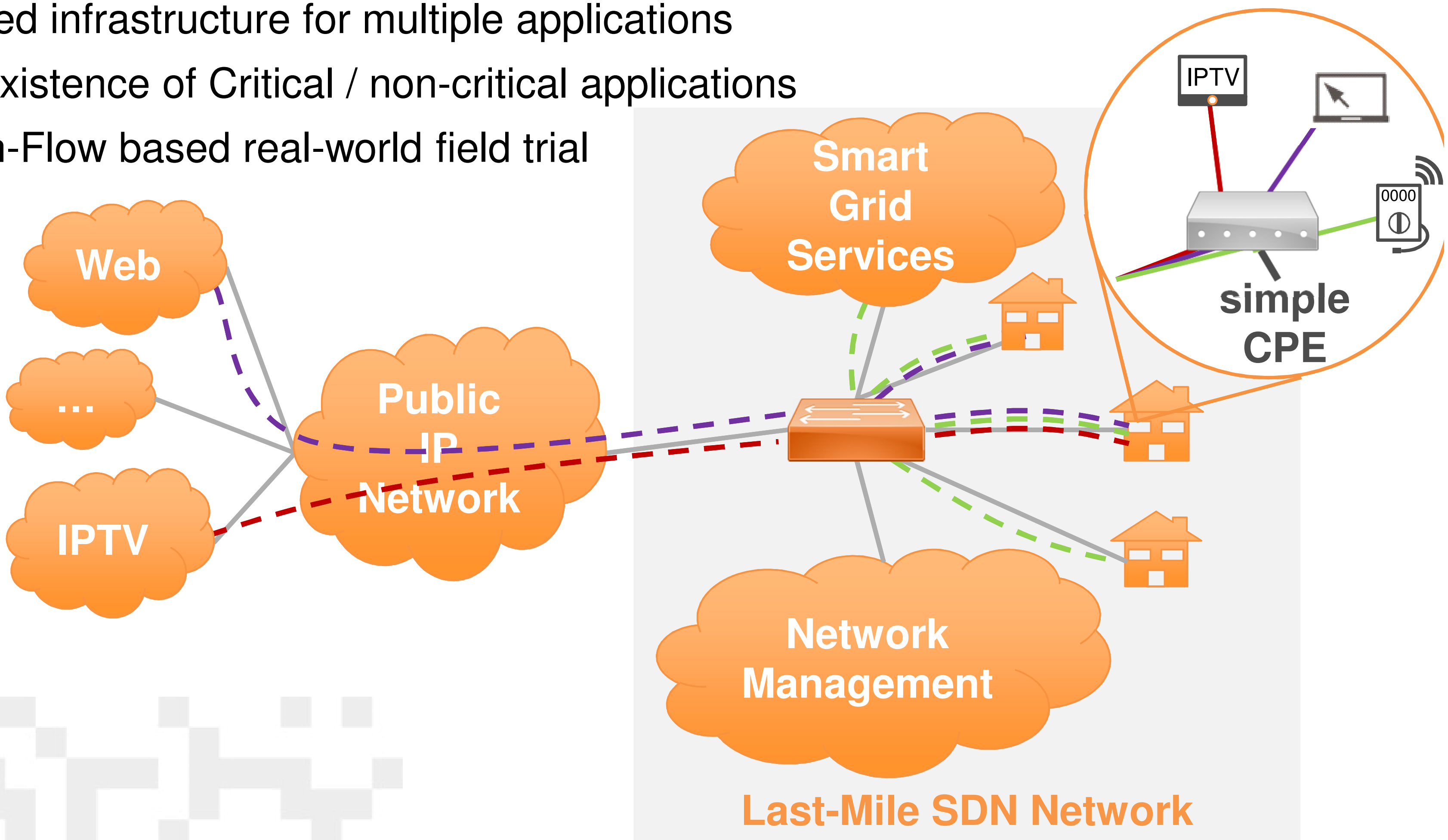
OpenFlow Fast Failover Groups

Incoming	Actions
Port:1	1 Group:ff1
Group:ff1	1 Watch_port:2;output:2
	2 Watch_port:3;output:3

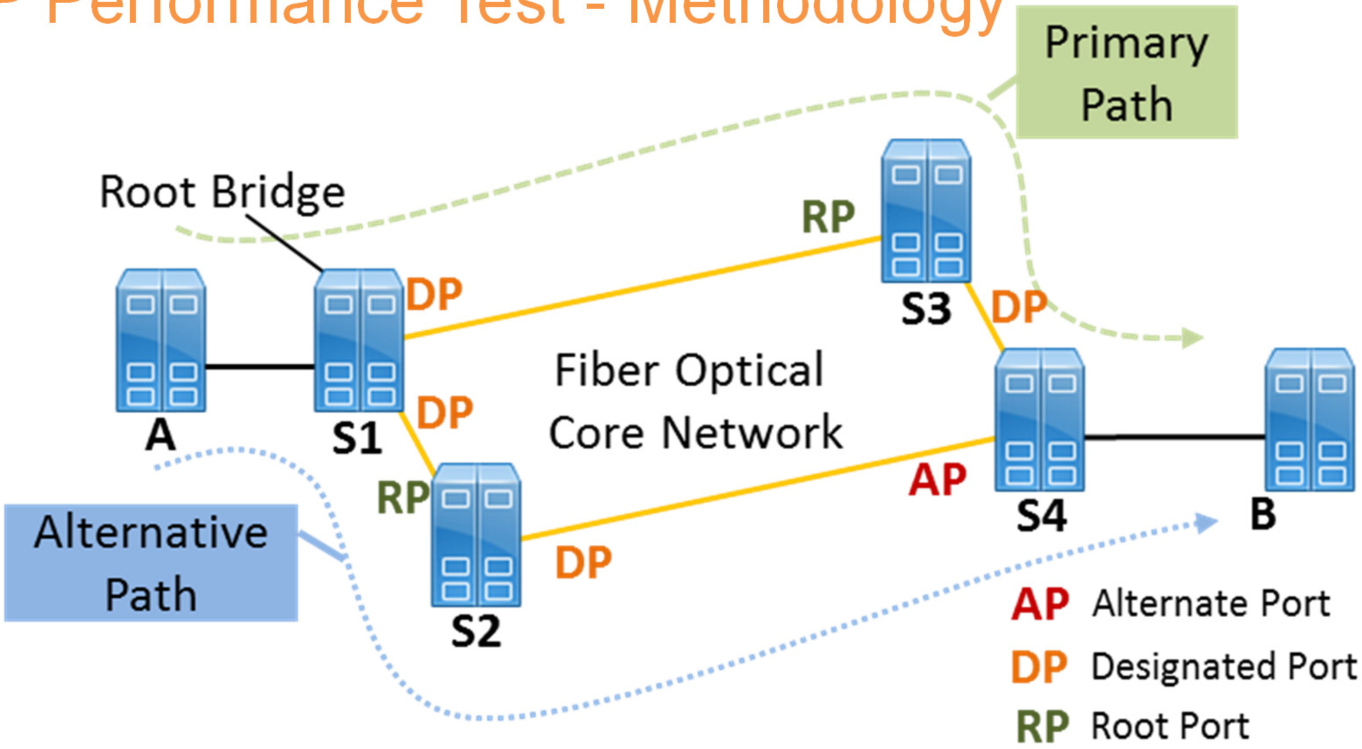
- Provides link redundancy
  - Watches port liveness/status.
  - Connect forwarding rules to the liveness/status of ports/links
  - Packets are sent via the first port with status ‘up’
- Data plane only forwarding decision / No distributed algorithm
- Good for time sensitive applications
- In larger networks
  - Control plane manages Failover Groups (e.g. NFV)
  - See also: Du, Pfeiffenberger, Bittencourt

OPOSSUM OpenFlow Testbed

- Testbed for SDN based Critical Infrastructure Communication
- Traffic Separation for critical and non-critical applications
- Shared infrastructure for multiple applications
- Co-Existence of Critical / non-critical applications
- Open-Flow based real-world field trial



RSTP Performance Test - Methodology



- Automatic Link (de)activation every 10 s at S3.
- After Disconnection
  - S3 selects itself as new root
  - Sends information replies with S1 as root bridge and enables AP
- Measurement
  - Interruption time is estimated based on lost sequential packets (next slide)

Measurement Methodology

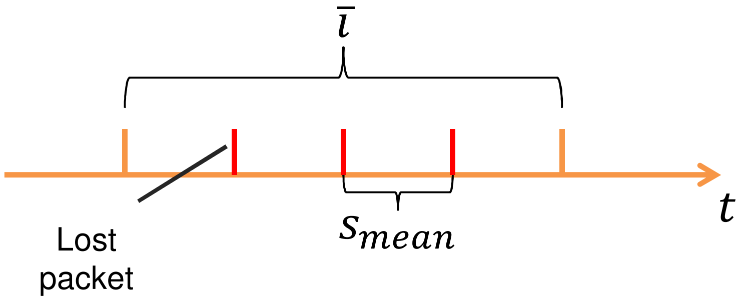
- UDP Sending Application
  - 288 Byte Ethernet Packets
  - 500 us mean sending interval
  - UDP Payload contains sending timestamp, packet sequence number (starting with 0)
- UDP Receiving Application
  - Evaluates lost, duplicated, and reordered packets
  - Computes one way delay (when time sync. Is well)
- Computing Interruption time

$$\bar{\iota} = s_{mean} \times (l_{seq} + 1)$$

$\bar{\iota}$: Upper bound interruption time

l_{seq}: Number of lost sequential packets

s_{mean}: Mean sending interval



RSTP Performance Test - Results

- 40 Actions
  - 20 Disconnections leads to **path repair** action
  - 20 Reconnections leads to **path restore** action

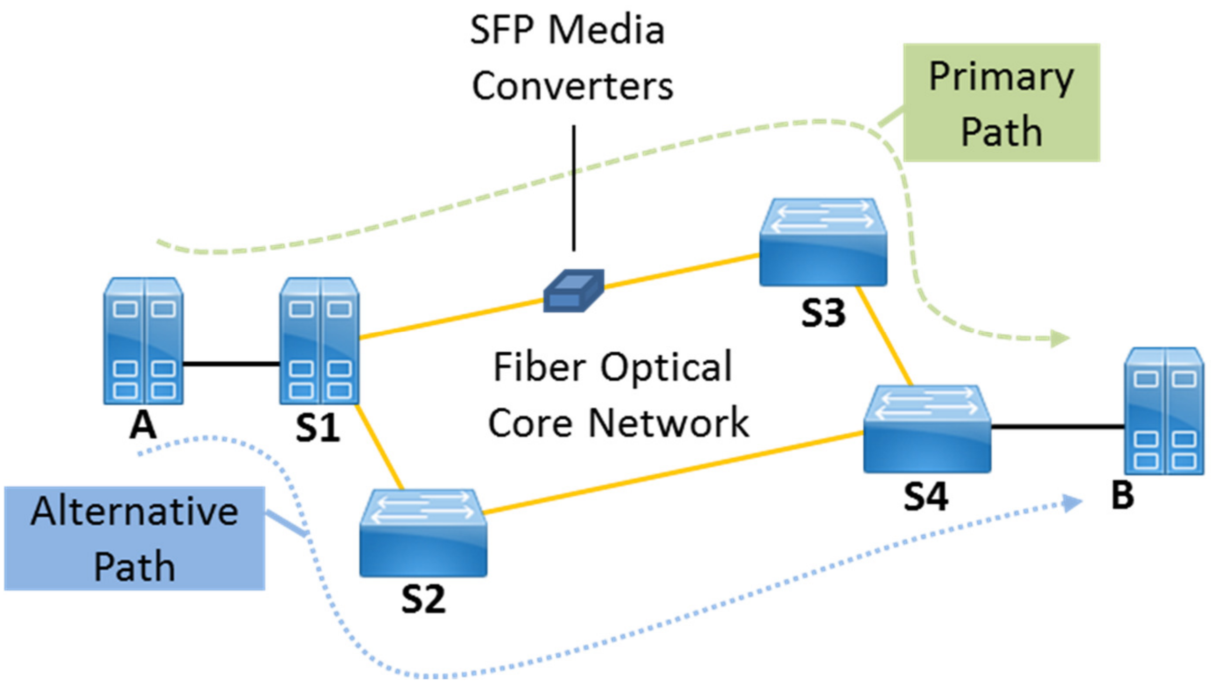
- Path Repair Performance
  - Minimum 3 ms
  - Maximum 65 ms
  - Average (mean): 26 ms

- Good Performance
  - 50 ms upper threshold for applications

- Path Restore Performance
  - Minimum: < 1 ms (no packet loss).
  - Maximum: 809 ms
  - Average (mean): 401 ms

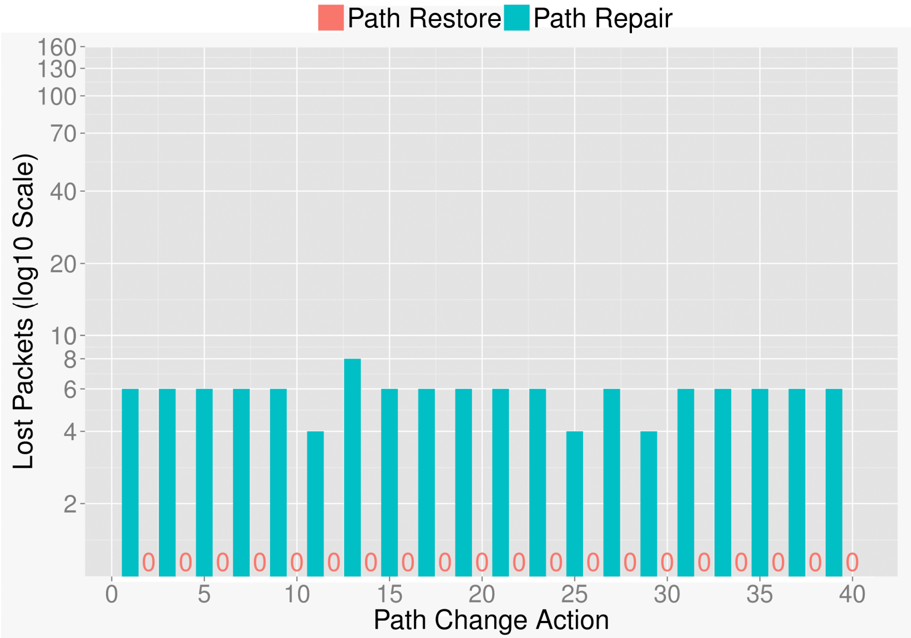
- Remarkable Behavior due to:
  - MAC Address Flushing
  - Inefficient Software implementation
  - Operating System Scheduling / Hardware Control at Host Computer

OpenFlow Fast-Failover Test - Methodology



- Two Test scenarios
  - Automated: Simulation of software failures
  - Manual: Simulation of link failures
    - Uses SFP Media Converter as interrupter to avoid contact chatter

OpenFlow Fast-Failover Test - Results

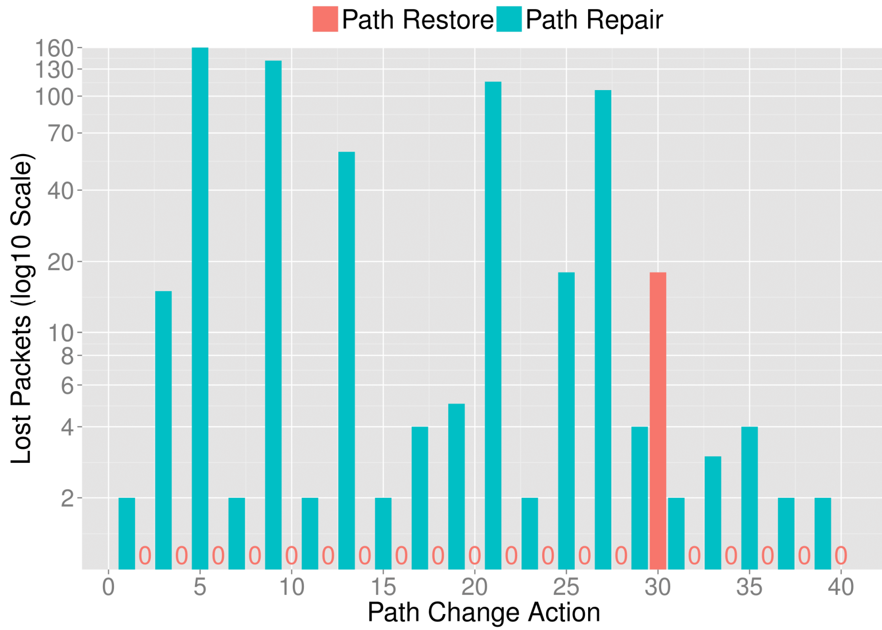


Automatic Switching

- Mean packet sending interval: 500 us
- Path Repair (green)
  - Min.: 4 Pkts, Max: 8 Pkts, Mean: 6 Pkts.
  - Interruption time $\bar{\iota}$: [3 , 5] ms (Avg: 3,5 ms)
- Path Restore (red)
  - No Interruption

Manual Switching

- Mean packet sending interval: 500 us
- Path Repair (green)
  - Min.: 2 Pkts, Max: 160 Pkts, Mean: 33 Pkts.
  - Mean Interruption time $\bar{\iota}$: [2, 81] ms (Avg.: 17 ms)
- Path Restore (red)
  - Action 30: 18 Pkts lost (interruption time ~10ms)



Results Comparision

Average Failover Times	OpenFlow Fast Failover		RSTP	RSTP [Siemens]
	Automatic	Manual		
Path Repair	< 5 ms	< 20 ms	< 30 ms	< 50 ms
Path Restore	No interruption	Mostly no Interruption	< 500 ms	n. a.

Conclusion

- Hard to verify RSTP results
  - Software Switches (OVS) influenced by
    - Host Hardware
    - Operating System (e. g. Scheduling)
    - Possibly algorithms not properly implemented
- OpenFlow Fast Failover
  - Contact Chatter when manually plugging optical cables
  - Remarkable Differences between manual and automatic Test scenarios
    - Degraded operation modes of NIC drivers
      - Takes long until OVS gets informed about lost link

Conclusion & Future Work

- Good Performance of OpenFlow FastFailover
  - sub 10 ms range possible.
  - Better then RSTP (?)
  - Simpler, cheaper as MPLS
  - Well suited for being integrated into OPOSSUM Testbed
- But: Performance depends
  - Software Switch implementation issues
  - Link failures vs. Software failures
- For the future:
  - Improving our measurements
    - Sub-millisecond accuracy (using PTP, Sync-E)
  - Study further:
    - RSVP-TE Fast Reroute with OpenFlow Fast Failover

Summary

Task: Buidling a Testbed for Critical Infrastructures Communication

Failover Mechanisms are important!

MPLS Approach

- Good Performance
- Complex to manage / install
- Needs IP based infrastructure
- Expensive

RSTP Approach

- Supported by almost all Network devices
- Layer-2

OpenFlow Fast Failover

- New SDN based approach
- How is the performance?

What has been shown:

- OpenFlow provides a pretty good performance
- Makes it a promising candidate for CI communication.

Thank you!

salzburgresearch

Ferdinand von Tüllenburg
Position
Salzburg Research Forschungsgesellschaft mbH
Jakob Haringer Straße 5/3 | 5020 Salzburg, Austria
T +43.662.2288-0 | F -222
vorname.nachname@salzburgresearch.at