大阪大学
OSAKA UNIVERSITY

Near Real-time Failover Model for Continuous Inter-Domain Communication

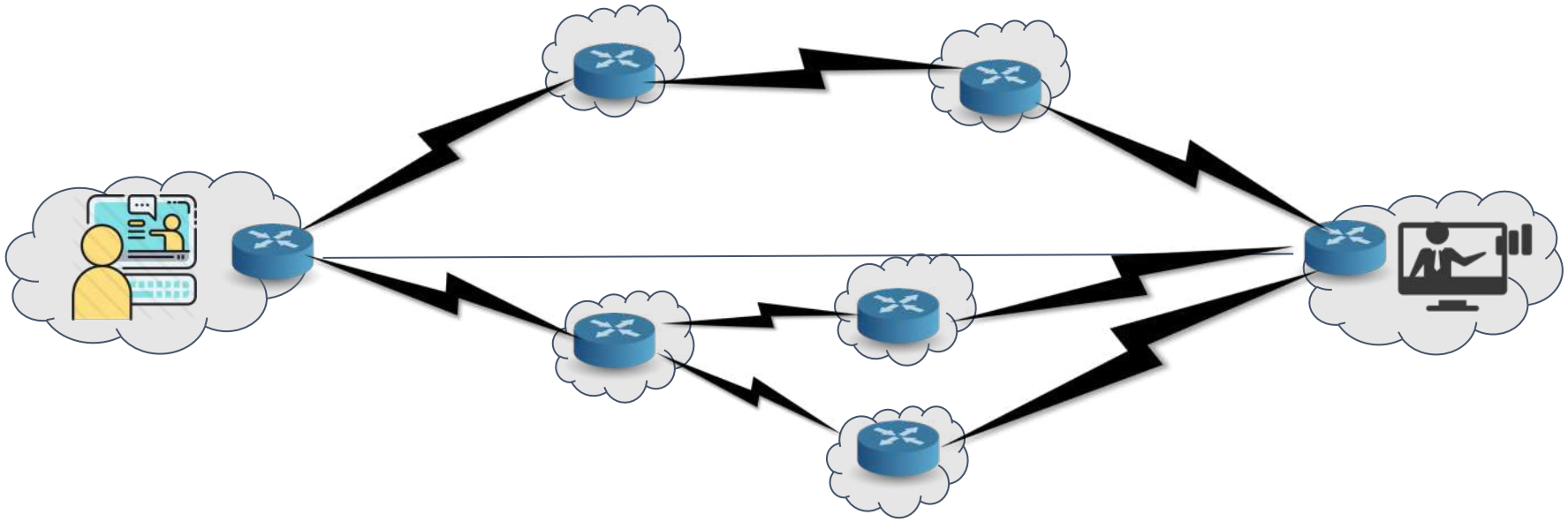
Sebastian Aguirre, Yoshiyuki Kido, Susumu Date, Shinji Shimojo

Contents

- Background
- Implementation
- Demonstration
- Conclusions
- Future Plan

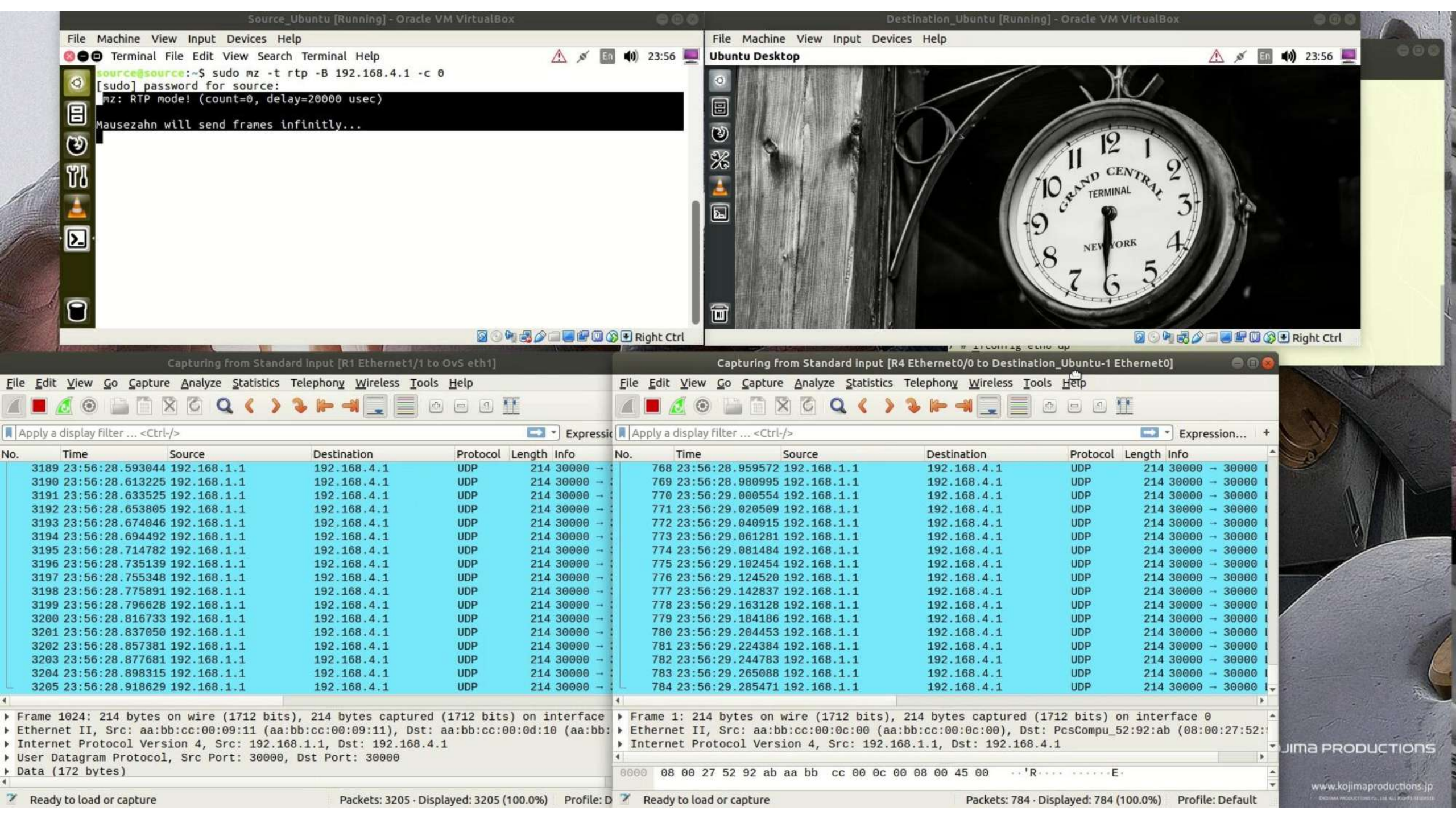
Background

- Motivation = continuous inter-domain communication
  - Increase availability between Autonomous Systems with redundancy



Background

- Border Gateway Protocol - Failover between Internet domains
  - Based on Keepalive and Hold timers
    - seconds to recover
    - Not all available next hops are considered
    - Traffic from all applications is forwarded to the same next-hop


Source_Ubuntu [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Terminal File Edit View Search Terminal Help
23:56
source@source:~$ sudo mz -t rtp -B 192.168.4.1 -c 0
[sudo] password for source:
mz: RTP mode! (count=0, delay=20000 usec)
Mausezahn will send frames infinitly...
Right Ctrl
Destination_Ubuntu [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Ubuntu Desktop
23:56
GRAND CENTRAL TERMINAL
NEW YORK
Right Ctrl
Capturing from Standard input [R1 Ethernet1/1 to OvS eth1]
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
Apply a display filter ... <Ctrl-/>
No. Time Source Destination Protocol Length Info
3189 23:56:28.593044 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3190 23:56:28.613225 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3191 23:56:28.633525 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3192 23:56:28.653805 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3193 23:56:28.674046 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3194 23:56:28.694492 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3195 23:56:28.714782 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3196 23:56:28.735139 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3197 23:56:28.755348 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3198 23:56:28.775891 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3199 23:56:28.796628 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3200 23:56:28.816733 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3201 23:56:28.837050 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3202 23:56:28.857381 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3203 23:56:28.877681 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3204 23:56:28.898315 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
3205 23:56:28.918629 192.168.1.1 192.168.4.1 UDP 214 30000 → 3
Frame 1024: 214 bytes on wire (1712 bits), 214 bytes captured (1712 bits) on interface
Ethernet II, Src: aa:bb:cc:00:09:11 (aa:bb:cc:00:09:11), Dst: aa:bb:cc:00:0d:10 (aa:bb:
Internet Protocol Version 4, Src: 192.168.1.1, Dst: 192.168.4.1
User Datagram Protocol, Src Port: 30000, Dst Port: 30000
Data (172 bytes)
Ready to load or capture
Packets: 3205 · Displayed: 3205 (100.0%)
Profile: D
Capturing from Standard input [R4 Ethernet0/0 to Destination_Ubuntu-1 Ethernet0]
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
Apply a display filter ... <Ctrl-/>
Expression...
No. Time Source Destination Protocol Length Info
768 23:56:28.959572 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
769 23:56:28.980995 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
770 23:56:29.000554 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
771 23:56:29.020509 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
772 23:56:29.040915 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
773 23:56:29.061281 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
774 23:56:29.081484 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
775 23:56:29.102454 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
776 23:56:29.124520 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
777 23:56:29.142837 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
778 23:56:29.163128 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
779 23:56:29.184186 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
780 23:56:29.204453 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
781 23:56:29.224384 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
782 23:56:29.244783 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
783 23:56:29.265088 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
784 23:56:29.285471 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
Frame 1: 214 bytes on wire (1712 bits), 214 bytes captured (1712 bits) on interface 0
Ethernet II, Src: aa:bb:cc:00:0c:00 (aa:bb:cc:00:0c:00), Dst: PcsCompu_52:92:ab (08:00:27:52:
Internet Protocol Version 4, Src: 192.168.1.1, Dst: 192.168.4.1
0000 08 00 27 52 92 ab aa bb cc 00 0c 00 08 00 45 00
Ready to load or capture
Packets: 784 · Displayed: 784 (100.0%)
Profile: Default
KOJIMA PRODUCTIONS
www.kojimaproductions.jp

Background

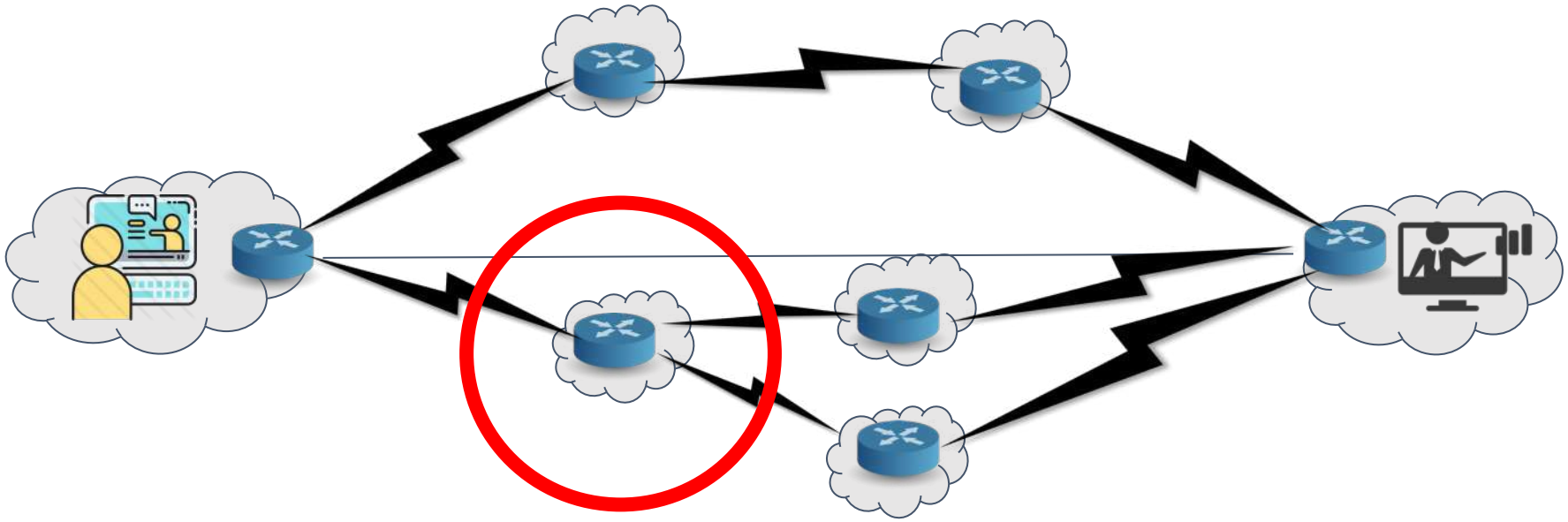
- Near Real-time
  - Process the communication interruption as soon as possible
  - **Stream processing**
    - Per-event processing of records as soon as they become available
  - High availability
    - Multiple “valid” next-hops
    - **BGP Monitoring Protocol** (RFC 7854) = learn all possible next-hops to reach a network
  - Define Services
    - Steer packet flows of specific applications
    - **SDN/OpenFlow** - configure the flow tables of network devices

Background

- How processing connectivity failure events as a stream of data enables continuous inter-domain communication?
  - Processing = update data-plane
  - Scalability
    - update time vs. available next-hops
  - Packet forwarding correctness
    - Next-hops should be valid according to BGP

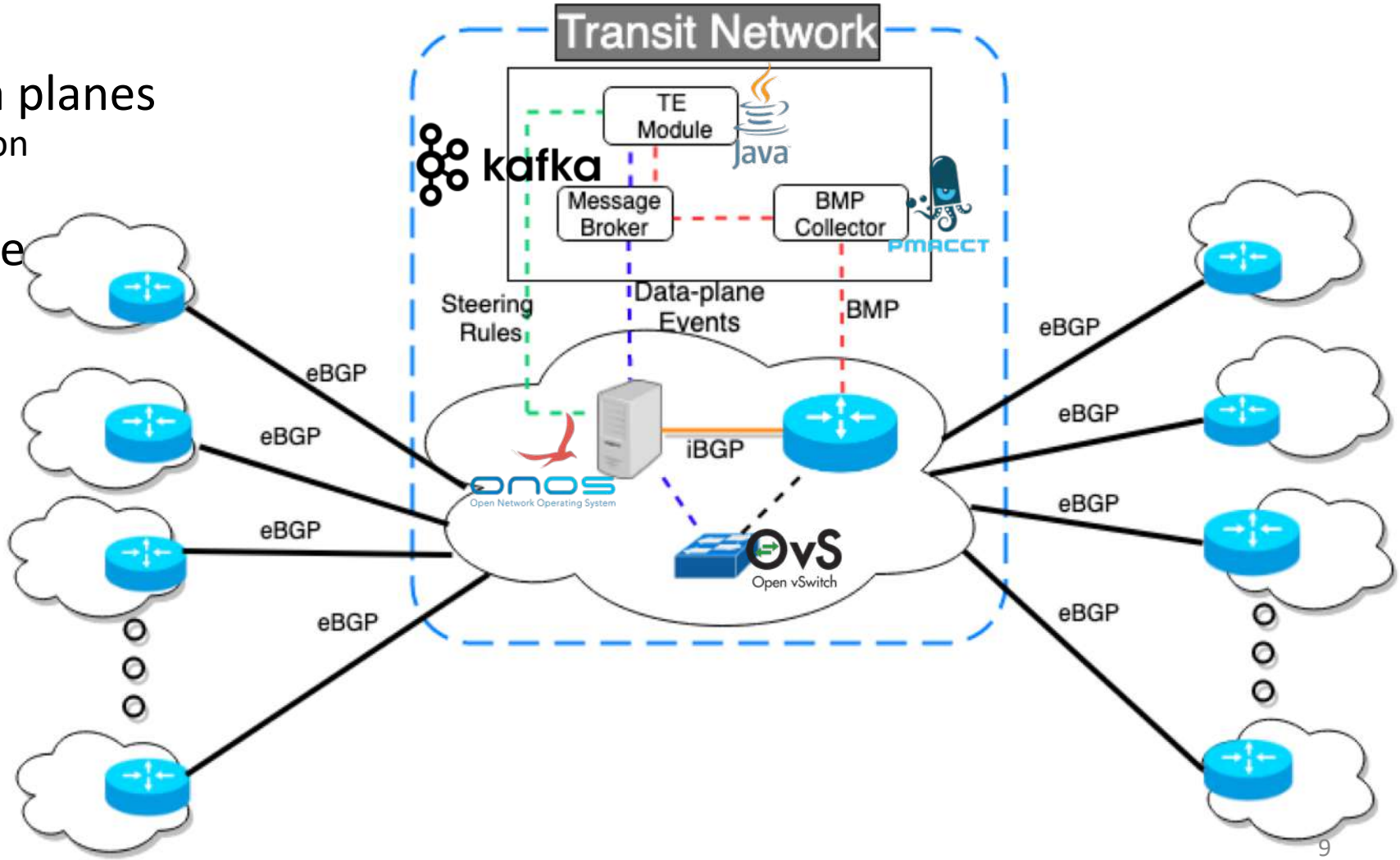
Implementation

- Transit Network
  - Bridges a connection between two or more networks (e.g. Tier-2 ISP)



Implementation

- Separate control and data planes
  - BGP routing as SDN application
- Traffic Engineering module
  - Handles failover process
  - Override BGP flow rules
- BMP Collector
  - Learns valid BGP routes
  - Publishes BGP Updates
- Message Broker
  - Intermediary



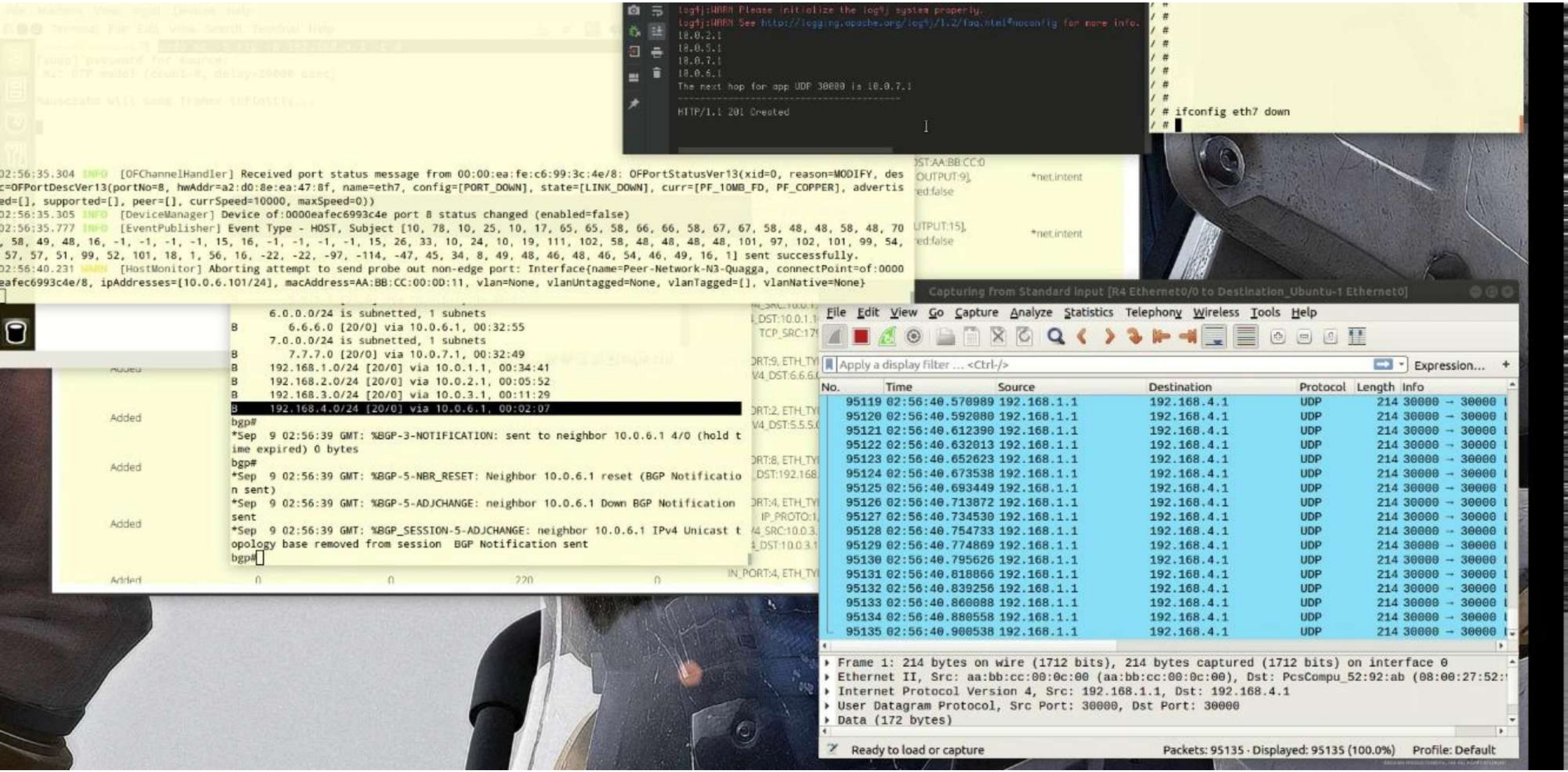
Demo

- Paket generation - Mausezahn
  - G.711 codec Real Time Protocol
  - 20 ms segment size, 160 bytes
  - UDP port 30000

- Testbed Topology
  - GNS3 - Network Software Emulator

Demo #1 - Connectivity Failure

- Data-plane port shutdown
- Convergence Time
  - 4 valid next-hops (Round-robin selection)
  - Packet loss

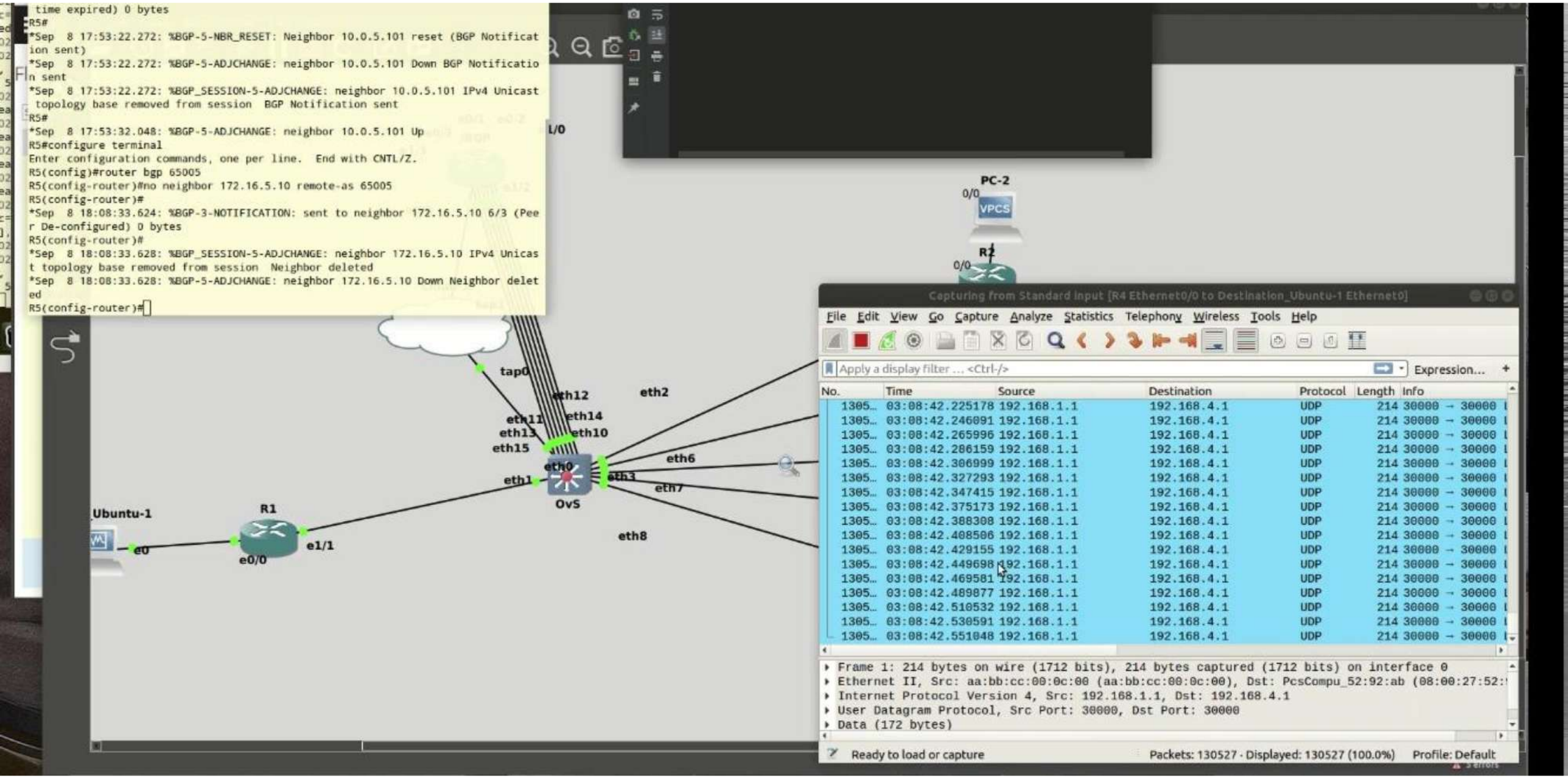

log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more info.
10.0.2.1
10.0.5.1
10.0.7.1
10.0.6.1
The next hop for app UDP 30000 is 10.0.7.1
HTTP/1.1 201 Created
/ # ifconfig eth7 down
/ #
02:56:35.304 INFO [OFChannelHandler] Received port status message from 00:00:ea:fe:c6:99:3c:4e/8: OFPortStatusVer13(xid=0, reason=MODIFY, des
c=OFPortDescVer13(portNo=8, hwAddr=a2:d0:8e:ea:47:8f, name=eth7, config=[PORT_DOWN], state=[LINK_DOWN], curr=[PF_10MB_FD, PF_COPPER], advertis
ed=[], supported=[], peer=[], currSpeed=10000, maxSpeed=0))
02:56:35.305 INFO [DeviceManager] Device of:000eafec6993c4e port 8 status changed (enabled=false)
02:56:35.777 INFO [EventPublisher] Event Type - HOST, Subject [10, 78, 10, 25, 10, 17, 65, 65, 58, 66, 66, 58, 67, 67, 58, 48, 48, 58, 48, 70
, 58, 49, 48, 16, -1, -1, -1, -1, -1, 15, 16, -1, -1, -1, -1, 15, 26, 33, 10, 24, 10, 19, 111, 102, 58, 48, 48, 48, 48, 101, 97, 102, 101, 99, 54,
57, 57, 51, 99, 52, 101, 18, 1, 56, 16, -22, -22, -97, -114, -47, 45, 34, 8, 49, 48, 46, 48, 46, 54, 46, 49, 16, 1] sent successfully.
02:56:40.231 WARN [HostMonitor] Aborting attempt to send probe out non-edge port: Interface{name=Peer-Network-N3-Quagga, connectPoint=of:0000
eafec6993c4e/8, ipAddresses=[10.0.6.101/24], macAddress=AA:BB:CC:00:0D:11, vlan=None, vlanUntagged=None, vlanTagged=[], vlanNative=None}
6.0.0.0/24 is subnetted, 1 subnets
B 6.6.6.0 [20/0] via 10.0.6.1, 00:32:55
7.0.0.0/24 is subnetted, 1 subnets
B 7.7.7.0 [20/0] via 10.0.7.1, 00:32:49
B 192.168.1.0/24 [20/0] via 10.0.1.1, 00:34:41
B 192.168.2.0/24 [20/0] via 10.0.2.1, 00:05:52
B 192.168.3.0/24 [20/0] via 10.0.3.1, 00:11:29
B 192.168.4.0/24 [20/0] via 10.0.6.1, 00:02:07
bgp#
*Sep 9 02:56:39 GMT: %BGP-3-NOTIFICATION: sent to neighbor 10.0.6.1 4/0 (hold t
ime expired) 0 bytes
bgp#
*Sep 9 02:56:39 GMT: %BGP-5-NBR_RESET: Neighbor 10.0.6.1 reset (BGP Notificatio
n sent)
*Sep 9 02:56:39 GMT: %BGP-5-ADJCHANGE: neighbor 10.0.6.1 Down BGP Notification
sent
*Sep 9 02:56:39 GMT: %BGP_SESSION-5-ADJCHANGE: neighbor 10.0.6.1 IPv4 Unicast t
opology base removed from session BGP Notification sent
bgp#
Capturing from Standard Input [R4 Ethernet0/0 to Destination_Ubuntu-1 Ethernet0]
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
Apply a display filter ... <Ctrl-/>
Expression...
No. Time Source Destination Protocol Length Info
95119 02:56:40.570989 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95120 02:56:40.592080 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95121 02:56:40.612390 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95122 02:56:40.632013 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95123 02:56:40.652623 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95124 02:56:40.673538 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95125 02:56:40.693449 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95126 02:56:40.713872 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95127 02:56:40.734530 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95128 02:56:40.754733 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95129 02:56:40.774869 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95130 02:56:40.795626 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95131 02:56:40.818866 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95132 02:56:40.839256 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95133 02:56:40.860088 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95134 02:56:40.880558 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
95135 02:56:40.900538 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
Frame 1: 214 bytes on wire (1712 bits), 214 bytes captured (1712 bits) on interface 0
Ethernet II, Src: aa:bb:cc:00:0c:00 (aa:bb:cc:00:0c:00), Dst: PcsCompu_52:92:ab (08:00:27:52:
Internet Protocol Version 4, Src: 192.168.1.1, Dst: 192.168.4.1
User Datagram Protocol, Src Port: 30000, Dst Port: 30000
Data (172 bytes)
Ready to load or capture
Packets: 95135 · Displayed: 95135 (100.0%)
Profile: Default

Demo #1 - Connectivity Failure

- Average communication loss ≈ 807.79 ms
  - Packet drop = 40

Demo #2 - BGP Update

- Packet forwarding correctness
  - Route withdrawals MUST redirect traffic back to valid BGP next-hops


time expired) 0 bytes
R5#
*Sep 8 17:53:22.272: %BGP-5-NBR_RESET: Neighbor 10.0.5.101 reset (BGP Notification sent)
*Sep 8 17:53:22.272: %BGP-5-ADJCHANGE: neighbor 10.0.5.101 Down BGP Notification sent
*Sep 8 17:53:22.272: %BGP_SESSION-5-ADJCHANGE: neighbor 10.0.5.101 IPv4 Unicast topology base removed from session BGP Notification sent
R5#
*Sep 8 17:53:32.048: %BGP-5-ADJCHANGE: neighbor 10.0.5.101 Up
R5#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
R5(config)#router bgp 65005
R5(config-router)#no neighbor 172.16.5.10 remote-as 65005
R5(config-router)#
*Sep 8 18:08:33.624: %BGP-3-NOTIFICATION: sent to neighbor 172.16.5.10 6/3 (Peer De-configured) 0 bytes
R5(config-router)#
*Sep 8 18:08:33.628: %BGP_SESSION-5-ADJCHANGE: neighbor 172.16.5.10 IPv4 Unicast topology base removed from session Neighbor deleted
*Sep 8 18:08:33.628: %BGP-5-ADJCHANGE: neighbor 172.16.5.10 Down Neighbor deleted
R5(config-router)#
PC-2
0/0
VPCS
R2
0/0
tap0
eth12
eth2
eth11
eth14
eth13
eth10
eth15
eth6
eth0
eth1
eth3
eth7
OvS
eth8
Ubuntu-1
R1
e0
e0/0
e1/1
Capturing from Standard Input [R4 Ethernet0/0 to Destination_Ubuntu-1 Ethernet0]
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
Apply a display filter ... <Ctrl-/>
Expression...
No. Time Source Destination Protocol Length Info
1305... 03:08:42.225178 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.246091 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.265996 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.286159 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.306999 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.327293 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.347415 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.375173 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.388308 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.408506 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.429155 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.449698 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.469581 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.489877 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.510532 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.530591 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
1305... 03:08:42.551048 192.168.1.1 192.168.4.1 UDP 214 30000 → 30000
Frame 1: 214 bytes on wire (1712 bits), 214 bytes captured (1712 bits) on interface 0
Ethernet II, Src: aa:bb:cc:00:0c:00 (aa:bb:cc:00:0c:00), Dst: PcsCompu_52:92:ab (08:00:27:52:
Internet Protocol Version 4, Src: 192.168.1.1, Dst: 192.168.4.1
User Datagram Protocol, Src Port: 30000, Dst Port: 30000
Data (172 bytes)
Ready to load or capture
Packets: 130527 · Displayed: 130527 (100.0%)
Profile: Default

Demo #2 - BGP Update

- Average communication loss ≈ 526.21 ms
  - Packet drop = 26

- BGP route withdrawal and traffic redirection ≈ 21.30 ms
  - Invalid forwarding = 1 packet

Preliminary Conclusions and Future Plans

- Failover time driven by stream processing is faster than the minimum configurable value of the BGP Hold Timer.
  - Simulation of complex inter-domain network topologies

- The stream processing task introduces incorrect packet forwarding.
  - Increase number of flows and BGP route table

Thanks!