It's kind of fun to do the impossible with DPDK

Yoshihiro Nakajima, Hirokazu Takahashi, Kunihiro Ishiguro, Koji Yamazaki

NTT Labs

NTT

Agenda

- **Motivation for fun ☺**
- **Fun with Lagopus SDN switch ☺**
- **Fun with speed ☺**
  - Smart FPGA for software dataplane
- **Fun with experience ☺**
  - SDN IX @ Interop Tokyo 2015 ShowNet

NTT

lagopus

Innovative R&D by NTT

Motivation for fun ☺

Trend

Trend shift in networking

- ✓ Closed (Vender lock-in)
- ✓ Yearly dev cycle
- ✓ Waterfall dev
- ✓ Standardization
- ✓ Protocol
- ✓ Special purpose HW / appliance
- ✓ Distributed cntrl
- ✓ Custom ASIC / FPGA
- ✓ Wired logic dataplane

- ✓ Open (lock-in free)
- ✓ Monthly dev cycle
- ✓ Agile dev
- ✓ DE fact standard
- ✓ API
- ✓ Commodity HW/ Server
- ✓ Logically centralized cntrl
- ✓ Merchant Chip
- ✓ Software dataplane

NTT

lagopus

Evaluate the benefits of SDN
by implementing control plane and switch
for fun ☺

4

NTT

Lagopus SDN switch project

NTT

Goal of Lagopus project

Provide NFV/SDN-aware switch framework

- SDN switch agent (OpenFlow, REST)
- 100Gbps high-performance soft dataplane
- Flexible/extensible switch configuration datastore
- DPDK extension (library, FPGA NIC, vNIC)
- Cloud middleware integration

Expand software-based packet processing to carrier networks

- Hardware acceleration and processing offload for scalable software dataplane

NTT

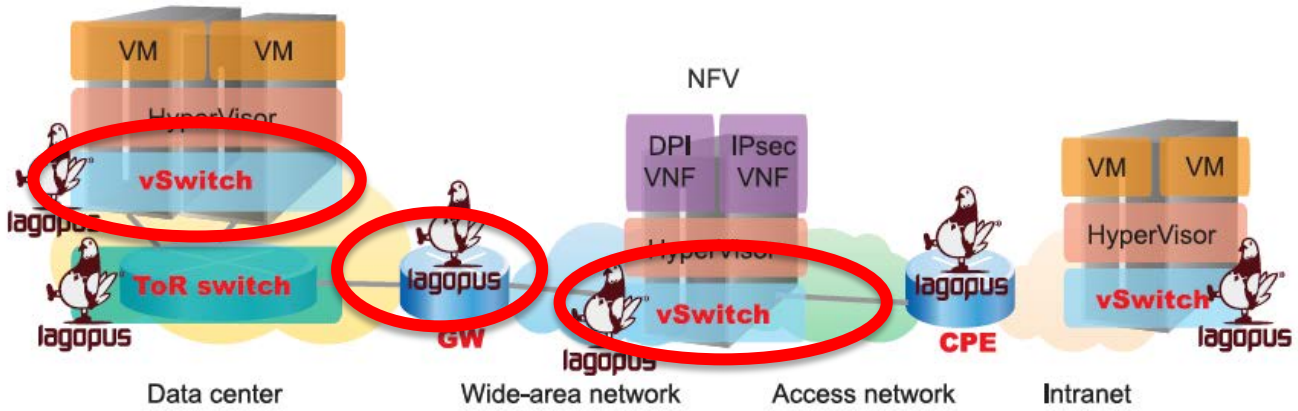
Target

High-performance SDN/OF software switch

- 40-Gbps packet processing throughput / port
- Mega-class flow entries support
- Low-latency and wire-rate speed in smaller packet size

Expands SDN to WAN, GW and NFV

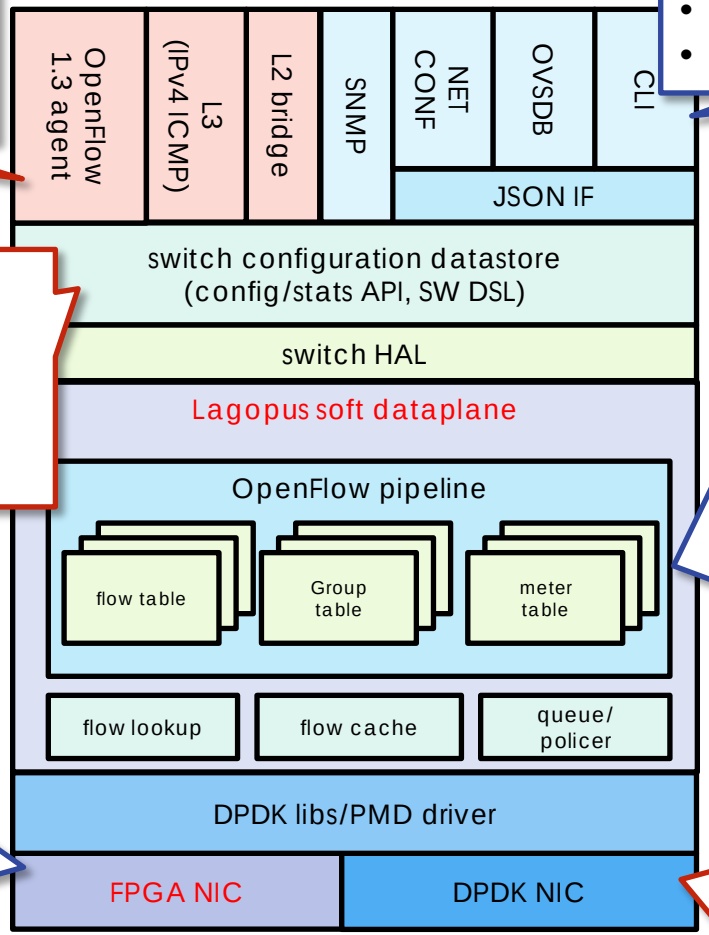
- Multiple frame format support
  - MPLS, PBB, MACinMAC, IPv4, IPv6
- Hybrid SDN support
  - REST API, OpenFlow 1.3, OVSDB, NETCONF
  - Legacy protocol support
- vSwtich for hypervisor, container virtualization



Lagopus vSwitch

SDN switch Agent
- Full OpenFlow 1.3.4 support
- Controller-less basic L2 and L3 support

SDN-aware management API
- OVSDB, REST
- Ansible support



Switch configuration datastore
- Pub/sub mechanism
- Switch config DSL
- JSON IF support

DPDK-enabled soft dataplane
- Over-10-Gbps performance
- Low latency packet processing
- high performance multi-layer flow lookup

Soft-dataplane aware FPGA NIC

DPDK-enabled vNIC for NFV (virtq-pmd)

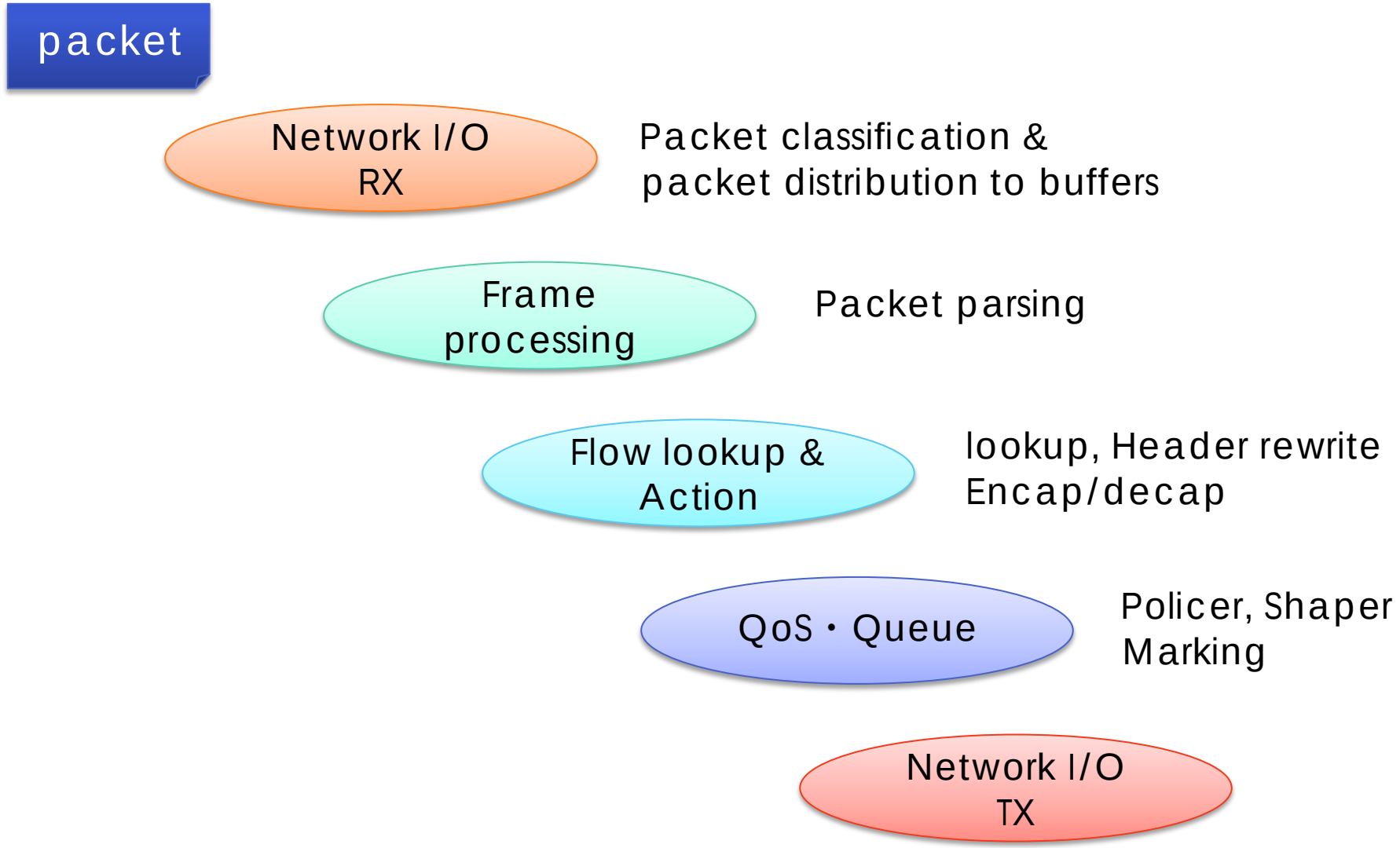
Fun with speed

- **Performance Improvement**
- **Smart FPGA NIC for software dataplane (collaboration with Xilinx)**

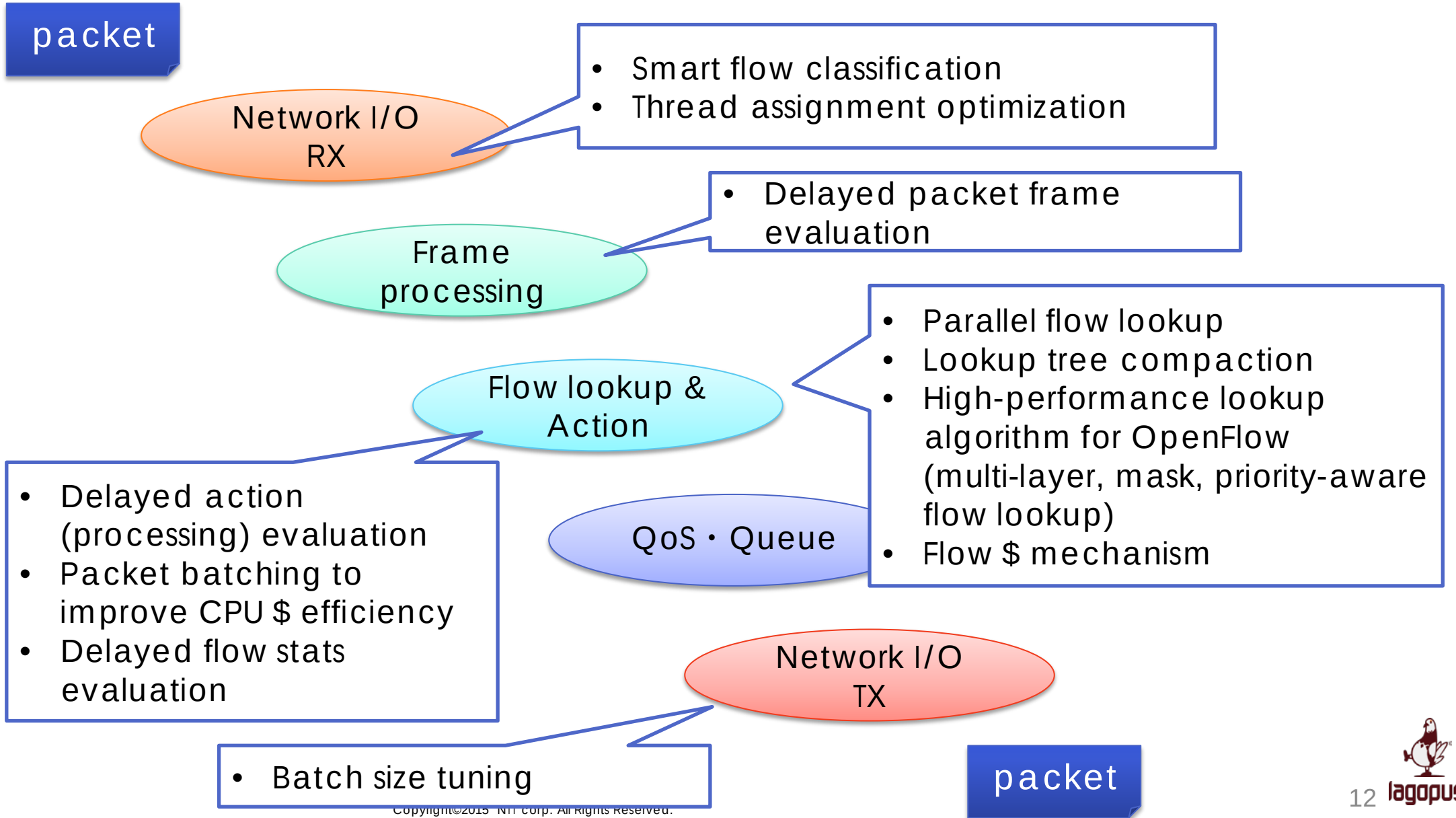
NTT

Performance Improvement

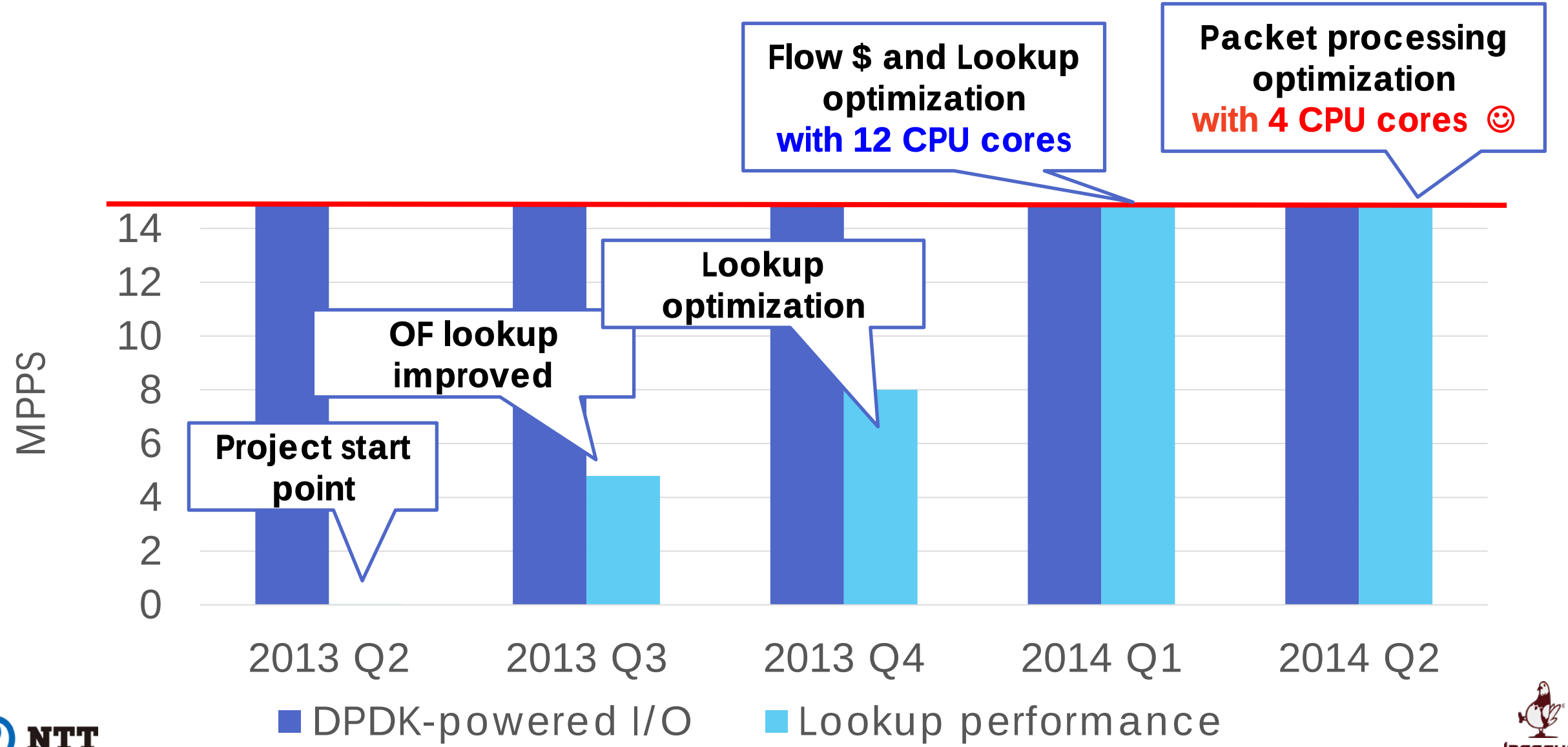
Basic packet processing



What we did for performance

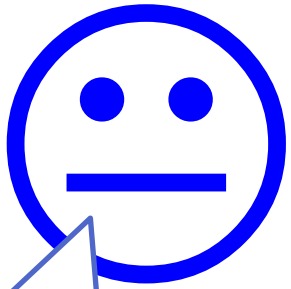


Road to 10Gbps packet processing with 1M OpenFlow flow entries



Big change from Y2013

Before project



10Gpbs by software dataplane? Impossible!!

Now

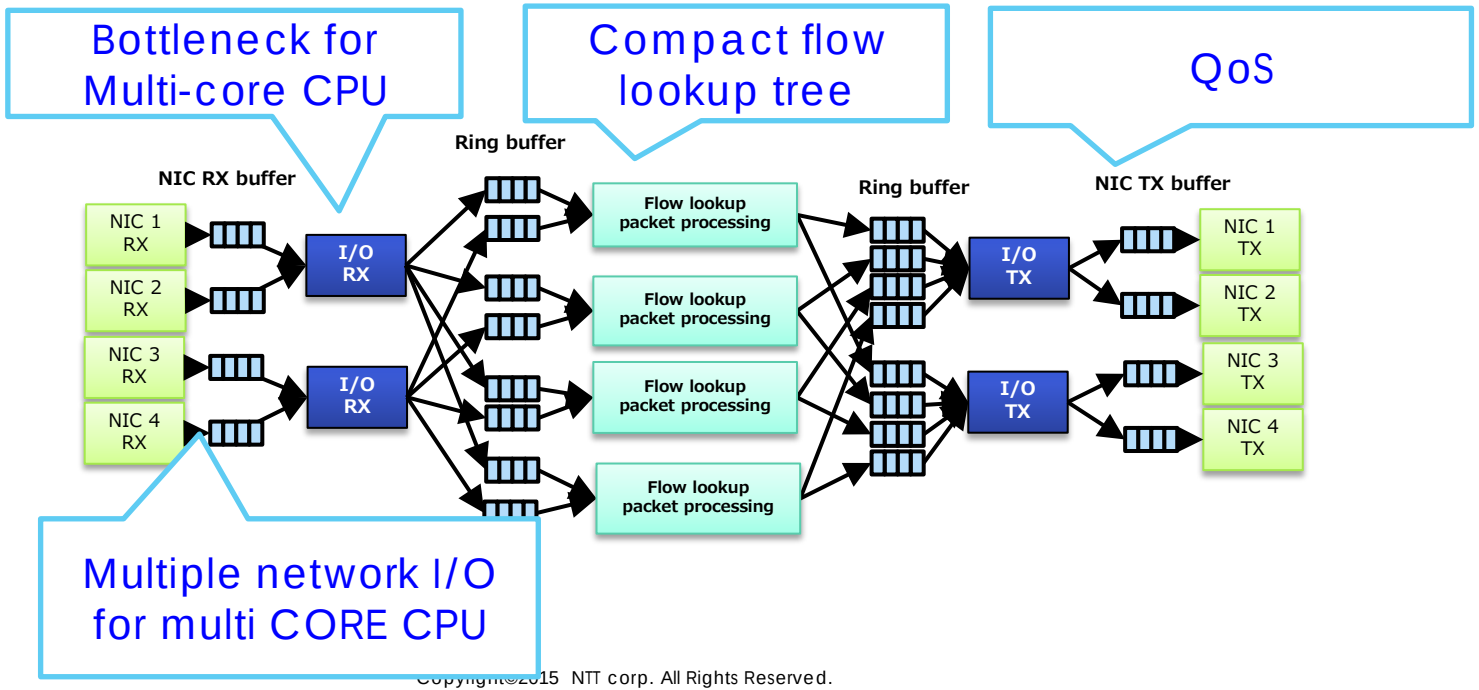
Software dataplane becomes great performance.
We try vSwitch for our usecases.

Smart FPGA NIC for software dataplane

NTT

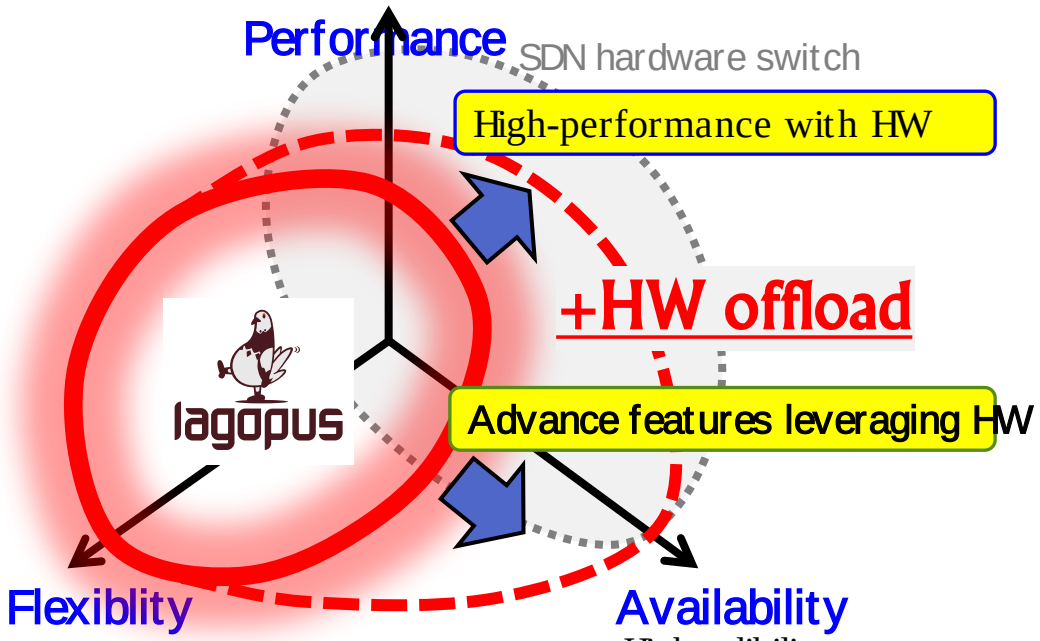
Motivation

- **Network I/O is not optimized for multi-core CPU**
  - Std NIC does not support RSS for WAN protocol
- **Software-based processing are heavy**
  - Packet classifier
  - Packet dispatcher are heavy
  - QoS and needs lots of CPU cycles

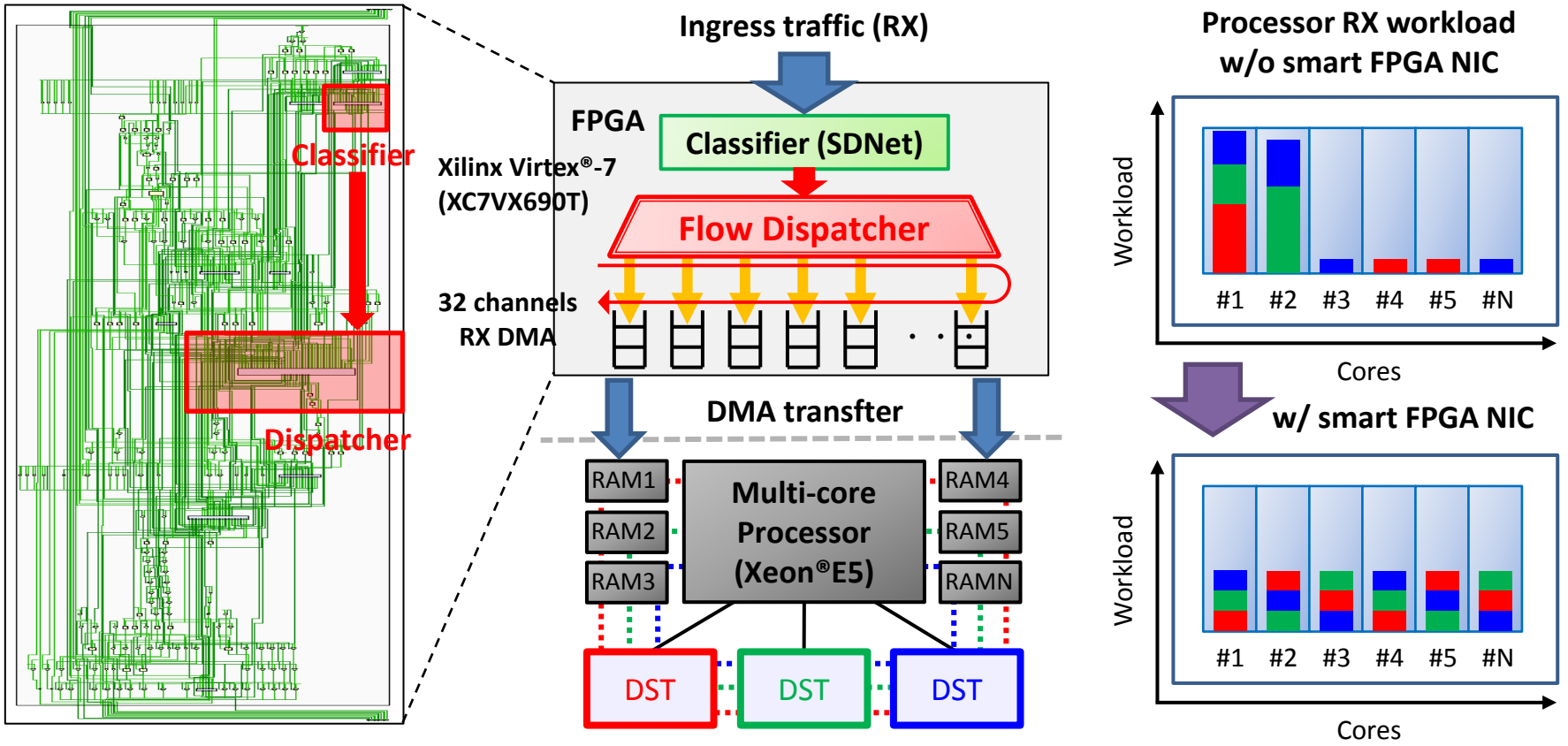


Co-design approach for performance

- **Leverage hardware offload processing of smart FPGA NIC**
  - Flexible hardware-based packet classifier & dispatcher
  - Hardware-based packet marking for post-packet-processing
- **Optimized to multi-core CPU**
  - Efficient packet processing for multi-thread



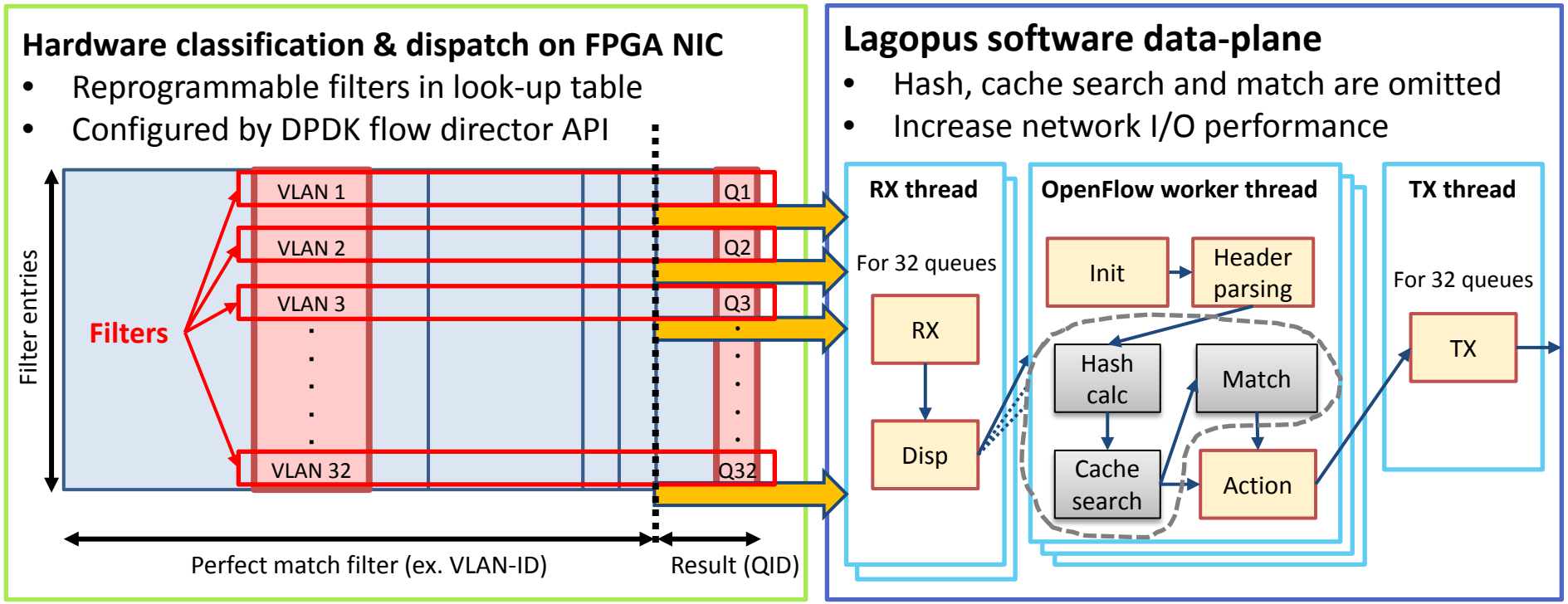
FPGA Flow Classification & Dispatch



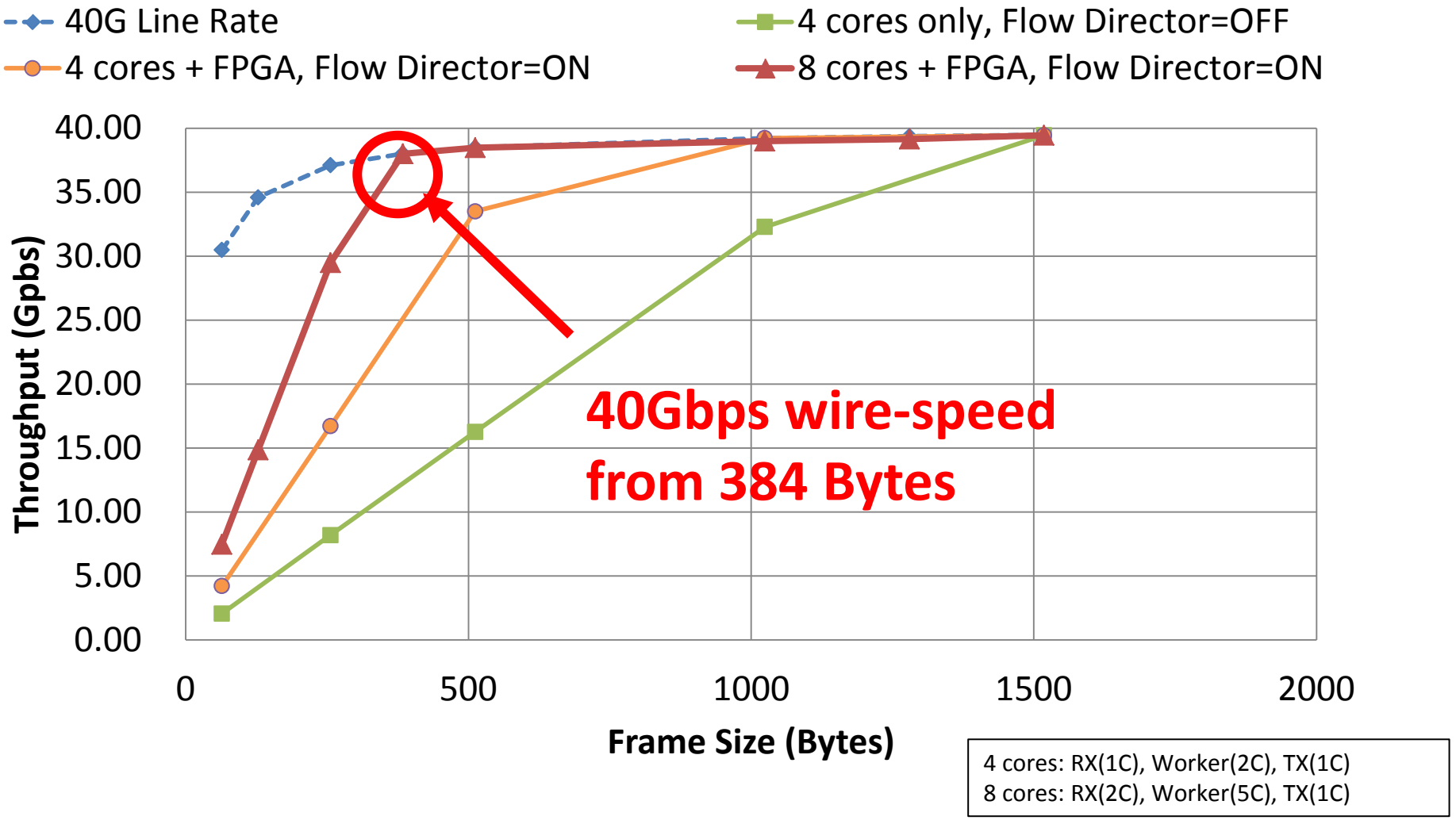
FPGA NIC and soft dataplane

Front-end hardware-based packet classification and packet dispatcher

- Reconfigurable packet filter for dynamic load balance for worker thread
- DPDK flow director API compatible



Improved performance with Smart FPGA NIC



ETH/VLAN/IPV4, 1 Flow, Intel Xeon® E5-2680 v2 (Ivy Bridge-EP)@ 2.8GHz, 10C/20T, Xilinx FPGA NIC, DPDK-1.7.1

Summery

- **Small extension of NIC give great benefit for software dataplane**
  - Performance improvement
  - Save CPU cycle for important processing
  - Save CPU cores for the same processing with standard NIC

NTT

lagopus

Fun with experience

SDN IX

NECOMA

SDN IX
@ Interop Tokyo 2015 ShowNet

Interop Tokyo is
the biggest Internet-related technology show in Japan.

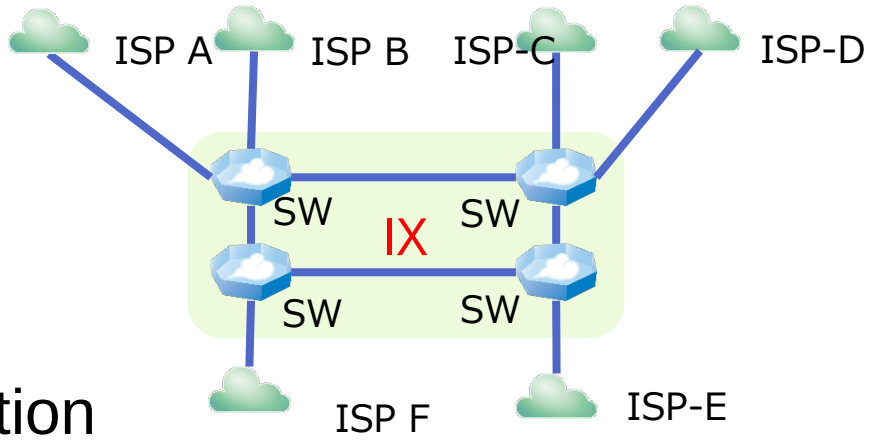
This trial was collaboration with NECOMA project
(NAIST & University of Tokyo)

NTT

Motivation

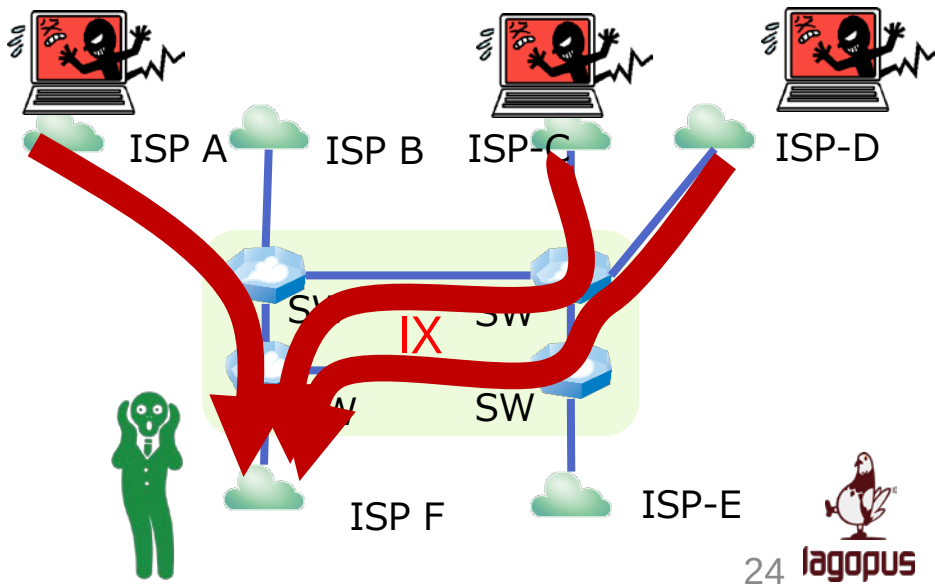
IX (Internet eXchange)

- Packet exchange point between ISP and DC-SP
- Boarder router of ISP exchanges route information



Issue

- Enhance automation in provisioning and configuration
- DDoS attack is one of the most critical issues
  - ISP wants to reduce DDoS-related traffic in origin
  - DDoS traffic occupies link bandwidth



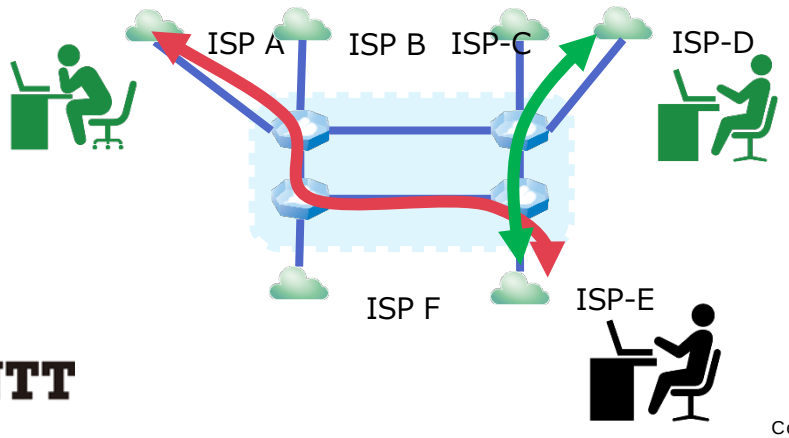
What is SDN IX?

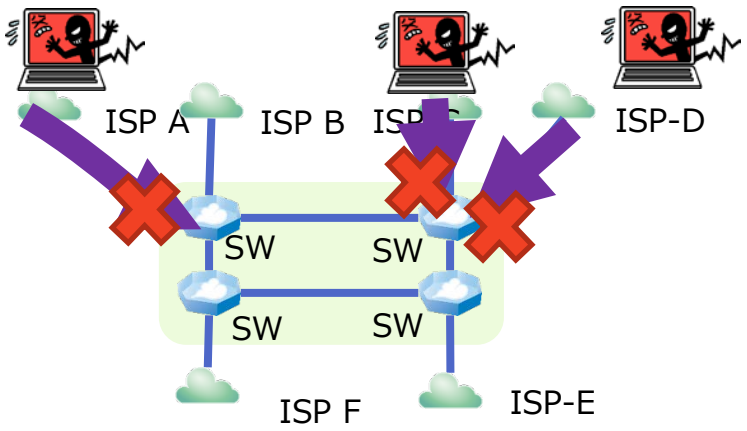
NECOMA

Developed by NECOMA project
(NAIST and University of Tokyo)

Next generation IX with SDN technology

- Web portal-based path provisioning between ISPs
  - Inter-AS L2 connectivity
    - VLAN-based path provisioning
    - Private peer provisioning
- Protect network from DDoS attack
  - On-demand 5-tuple-baesd packet filtering
- SDN IX controller and distributed SDN/OpenFlow IX core switch



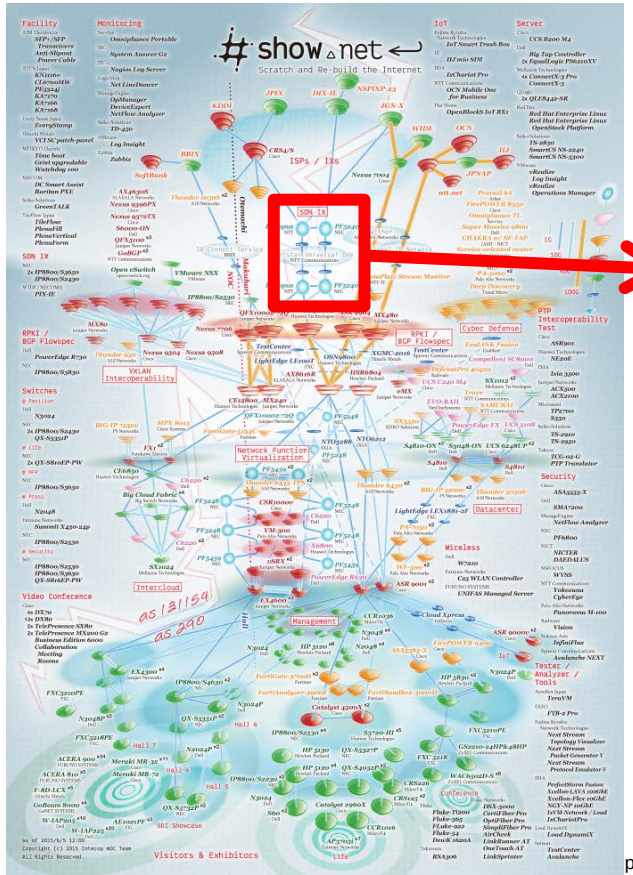


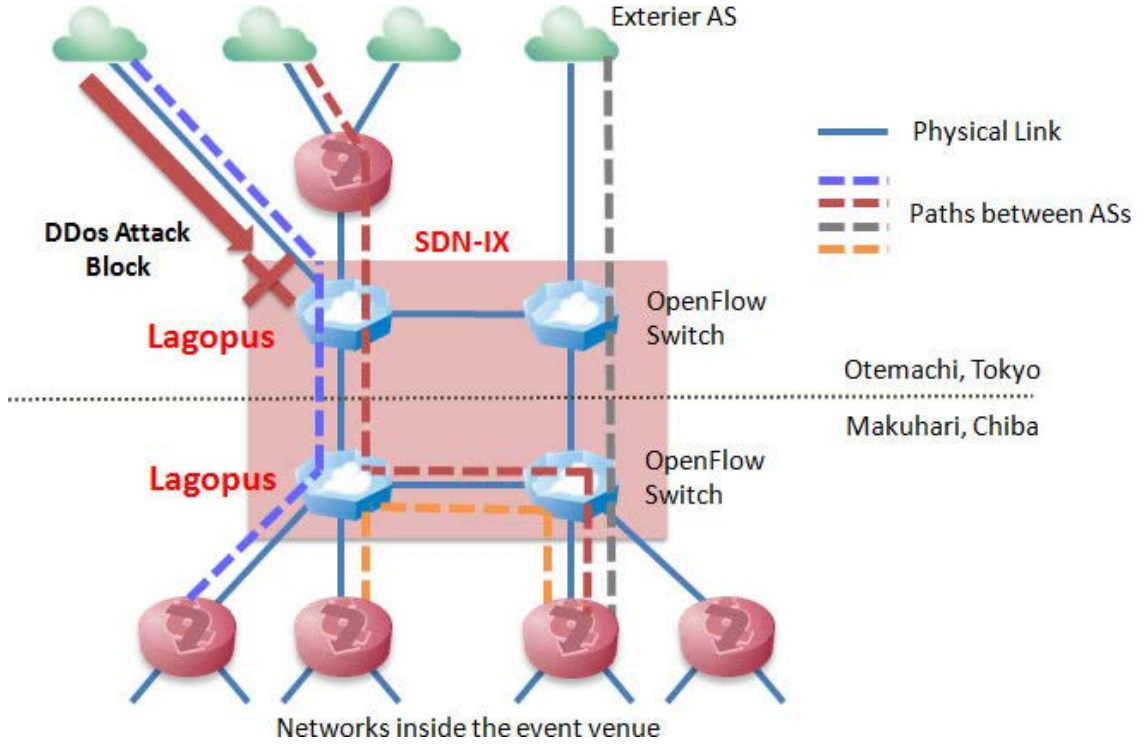
NTT

lagopus

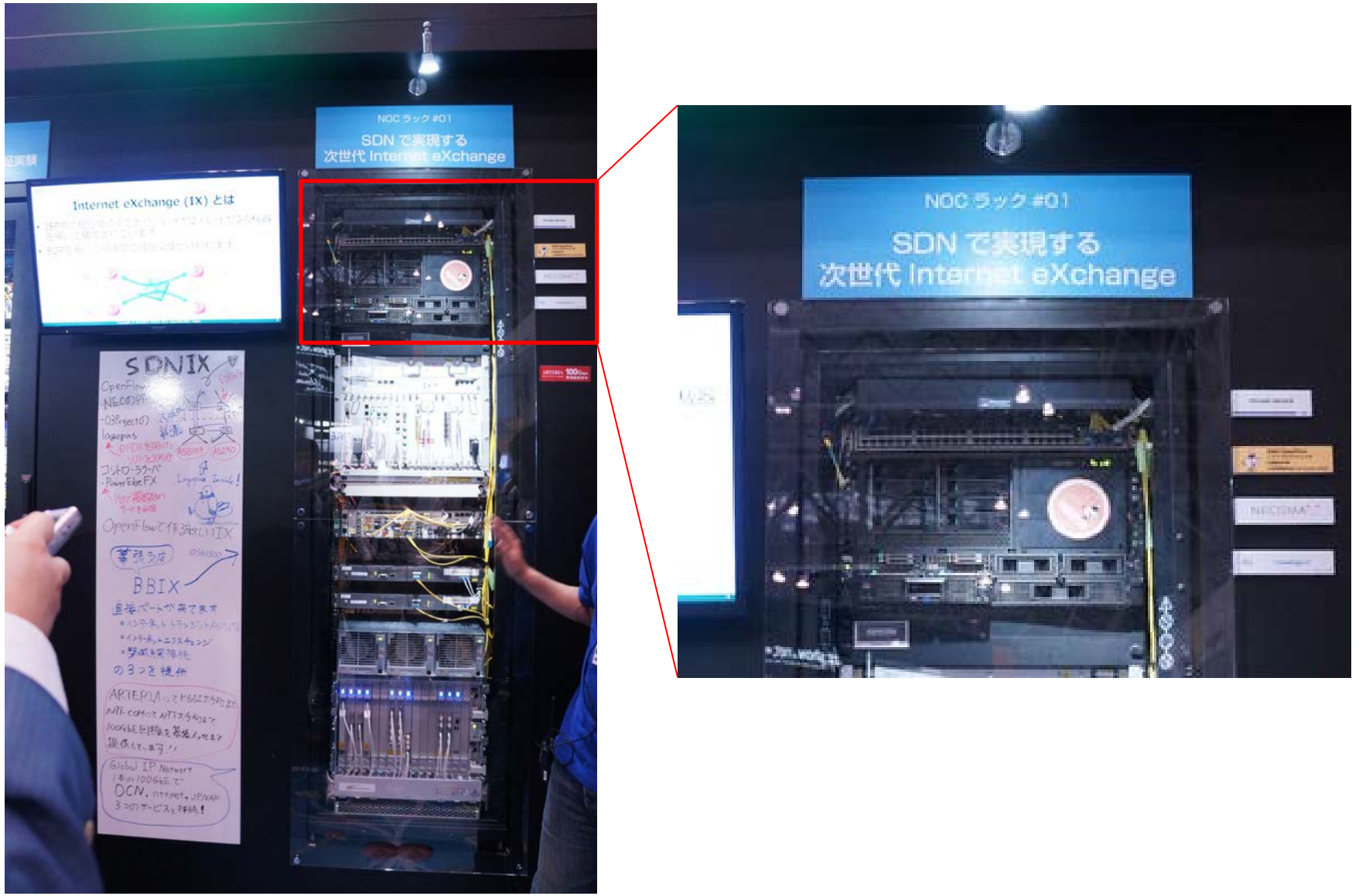
Lagopus @ ShowNet 2015

- **Two Lagopus (soft switch) are deployed for SDN-IX core switch**
  - Multiple 10Gbps links
  - Dual Xeon E5 8core CPUs

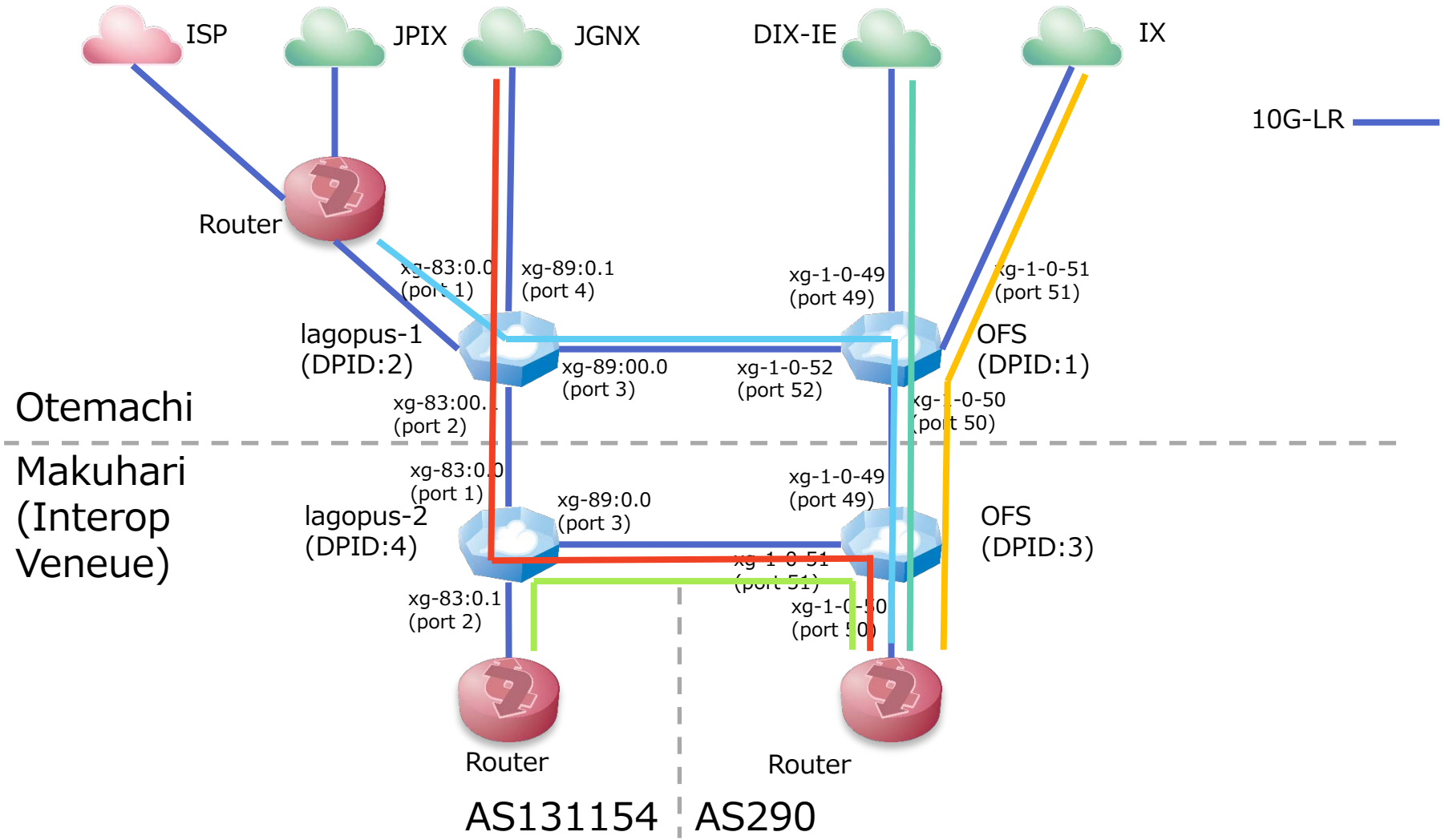




Lagopus @ ShowNet rack



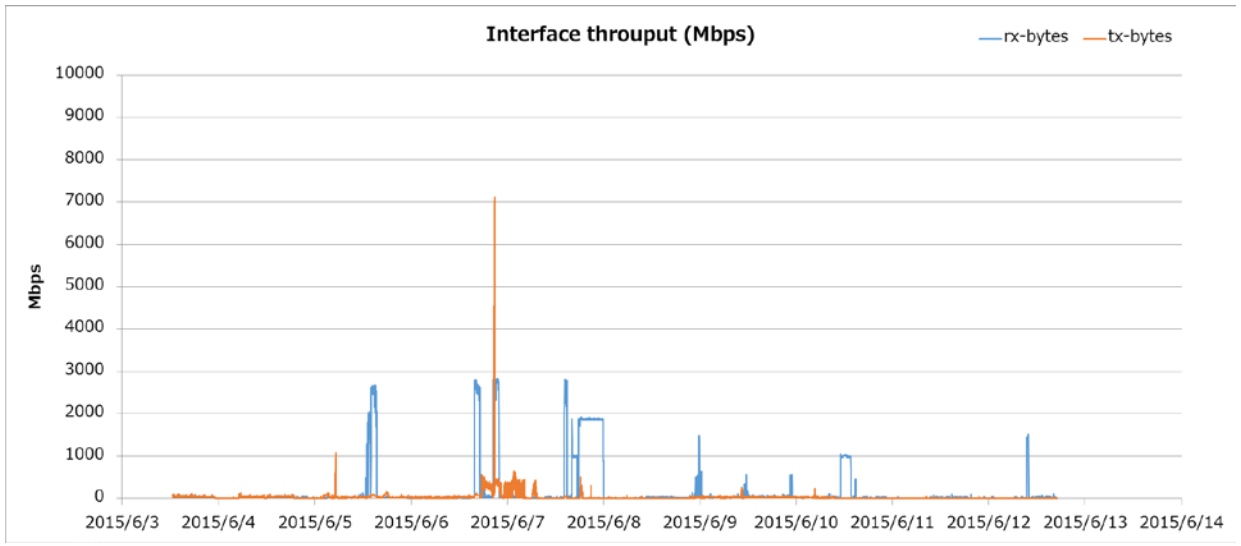
Path provisioning

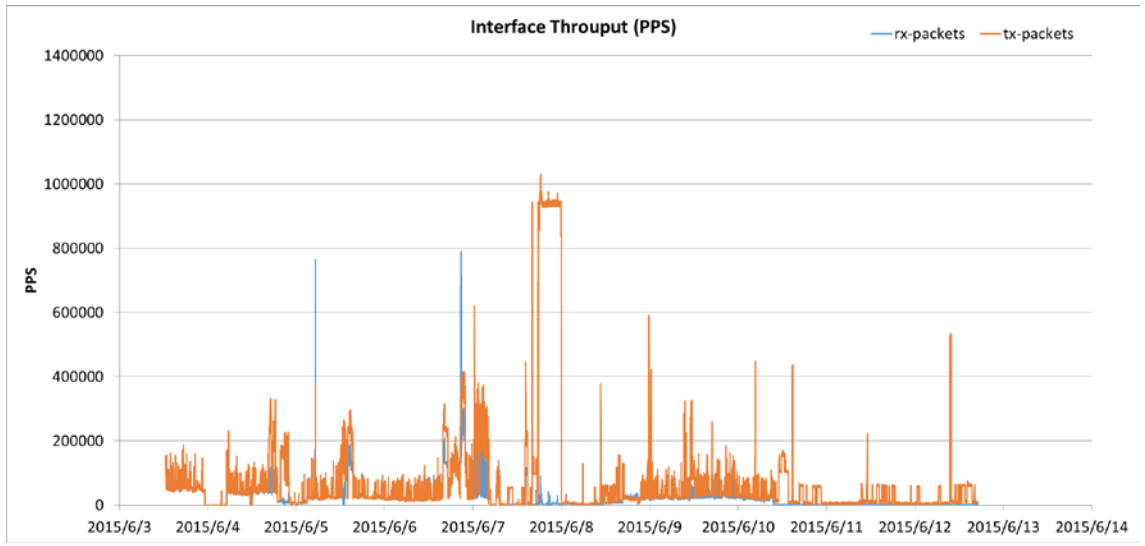


Traffic on Lagopus @Makuhari

Average 2Gbps throughput

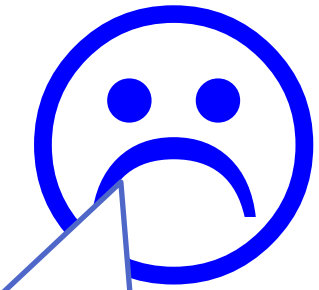
- No packet drop
- No reboot & no trouble for 1 week during Interop Tokyo
- Sometimes 10Gbps burst traffic





Big change happened

Before



vSwitch has lots of issues on performance, scalability, stability, ...

INTEROP TOKYO | 10 – 12 JUNE, 2015

After

vSwitch works well without any trouble!

Good performance, Good stability.

NTT

lagopus

Conclusion

- **It's kind of fun to do the impossible with DPDK**
  - Enjoy hacking with DPDK for your networking!
  - Performance optimization is fun ☺
- **Lagopus project commit to high-performance vswitch development for fun ☺**
  - We still have lots of issues for fun ☺
    - Lookup optimization, performance improvement, ….
- **Changing one's mind is great fun ☺**
  - Real experience change their mind ☺

NTT

lagopus

Visit our booth #172 in IDF15SFO

Lagopus demonstration

- vSwitch performance benchmark
  - Haswell-EP and Fortville
  - Carrier usecase
- MPLS-based segment routing (source routing) and NFV integration

https://github.com/lagopus/

vSwitch, DPDK extension, and more…

NTT