



Innovative R&D by NTT



Lagopus: 高速なパケット処理を実現する OpenFlowソフトウェアスイッチ

中島佳宏・高橋宏和・益谷仁士・日比智也・
河井彩公子・石黒邦宏
NTT



Agenda

■ 背景と目的

■ Lagopus: SDN/OpenFlow ソフトウェアスイッチ

- 構成
- 実装の工夫
- 性能評価

■ Lagopusを使った柔軟なネットワーク

- SDN IX @ Interop Tokyo 2015
- Segment Routing with MPLS (Source Routing, SPRING)



Innovative R&D by NTT

Lagopus SDN/OpenFlow switch project

Lagopusプロジェクトの目的

■ NFVやSDNを実現するソフトウェアコンポーネントを提供

- スイッチエージェント (OpenFlow, REST)
- 高速なソフトウェアデータプレーン
- 柔軟・拡張可能なスイッチ設定機構
- 高速I/Oライブラリ (DPDK)
- 外部システム連携ライブラリ (クラウドミドルウェア・補助ツール)
- Network I/O ハードウェア拡張

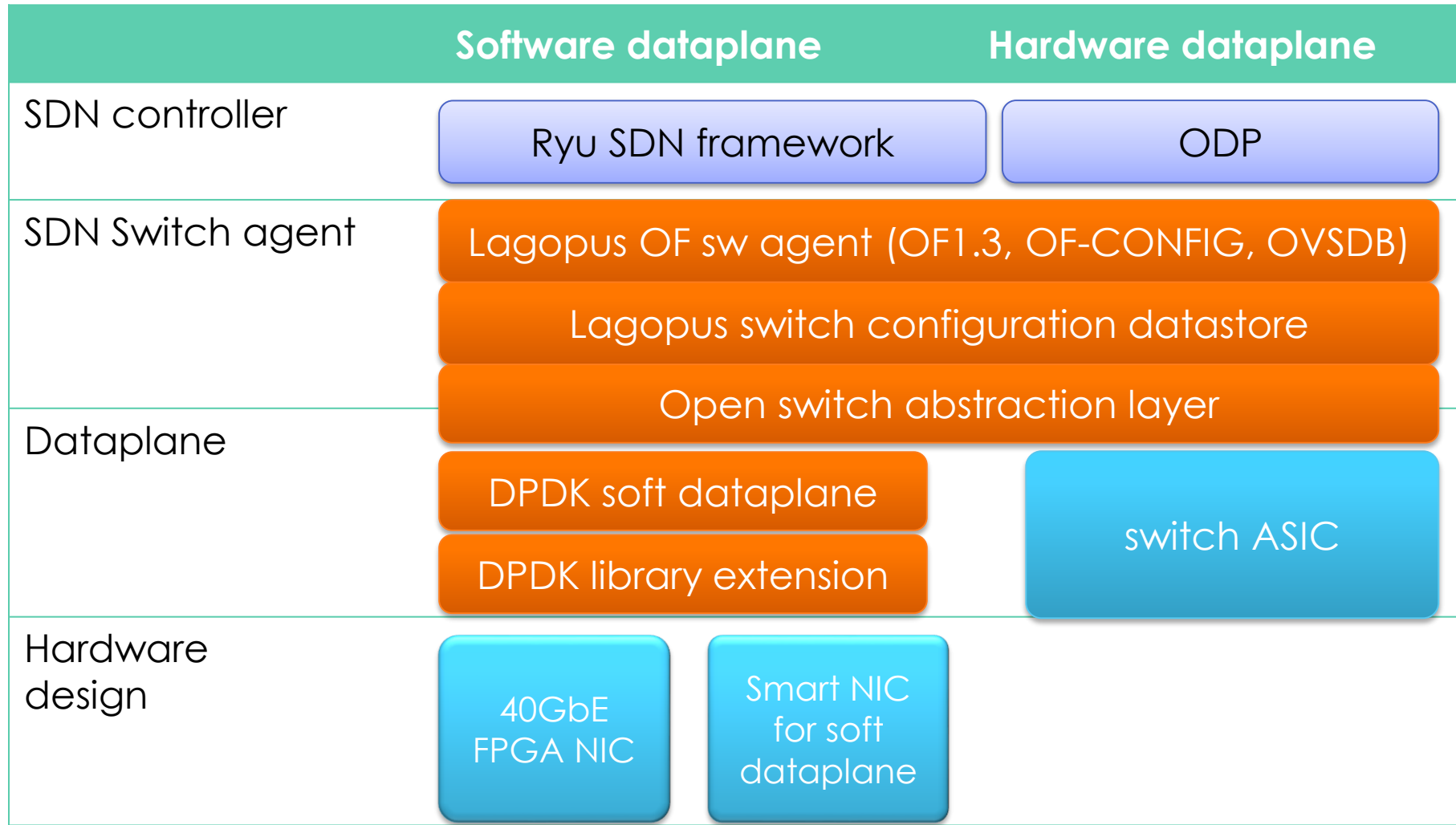
■ ソフトウェアパケット処理の適応領域の拡大

- キャリア・ネットワークでも活用可能な高いスケーラビリティ性
- 運用を容易にする高可用性・保守成性の向上

■ ネットワーク研究のツール・ライブラリとしての活用促進

- 高速検索、QoSコンポーネント、次世代で求められる拡張

Lagopus project overview



Lagopusの概要

■ OpenFlow 1.3に最も適応したソフトウェアスイッチ

- 複数のプロトコルフレーム処理の対応
 - Ethernet, Mac in MAC, MPLS, PBB, IPv4, IPv6...
- 複数Flow tableやグループテーブルへの対応
- QoS機能 (queue, policer, meter table)

■ DPDKを活用した高性能データプレーン

- 10Gbpsを超えるパケット処理・転送性能 (20MPPS以上)
- 100万フローエントリへの対応 (4K flow mod/sec)

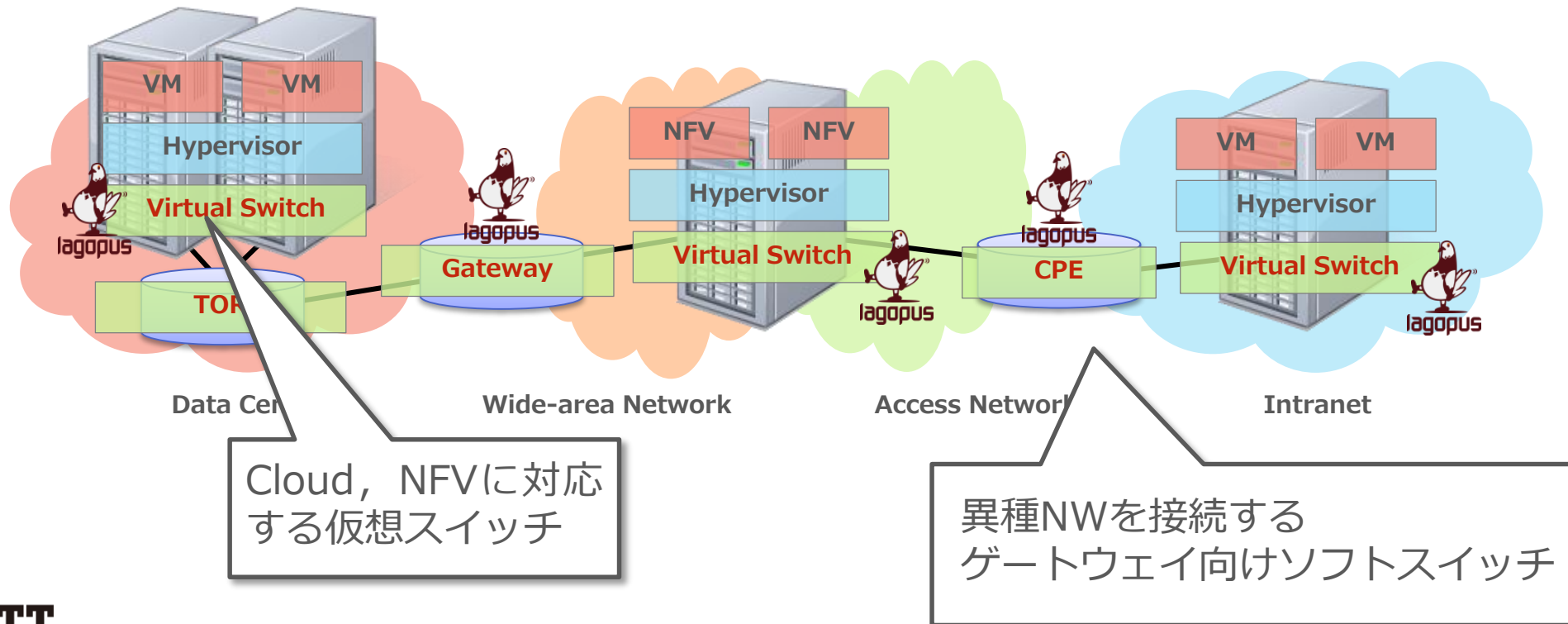
■ 外部連携API

- CLIやJSONによるスイッチ構成設定
 - Flow設定、統計情報・状況把握

Lagopusの適応先

■ ネットワークのエッジのノード

- 柔軟な制御や設定が求められるノード
- 多機能で大量のルート数やルールが求められるノード
- 求められるポート数が比較的少ないノード



Lagopus vSwitchの設計

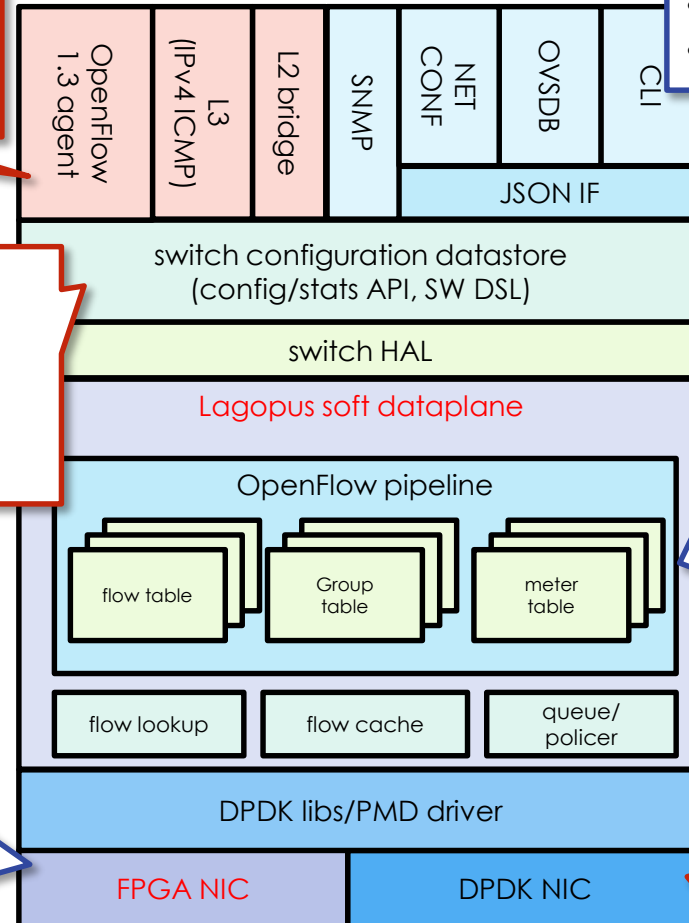
SDN switch Agent

- Full OpenFlow 1.3.4 support
- Controller-less basic L2 and L3 support

Switch configuration datastore

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

Soft-dataplane aware FPGA NIC



SDN-aware management API

- OVSDB, REST
- Ansible support

DPDK-enabled soft dataplane

- Over-10-Gbps performance
- Low latency packet processing
- high performance multi-layer flow lookup

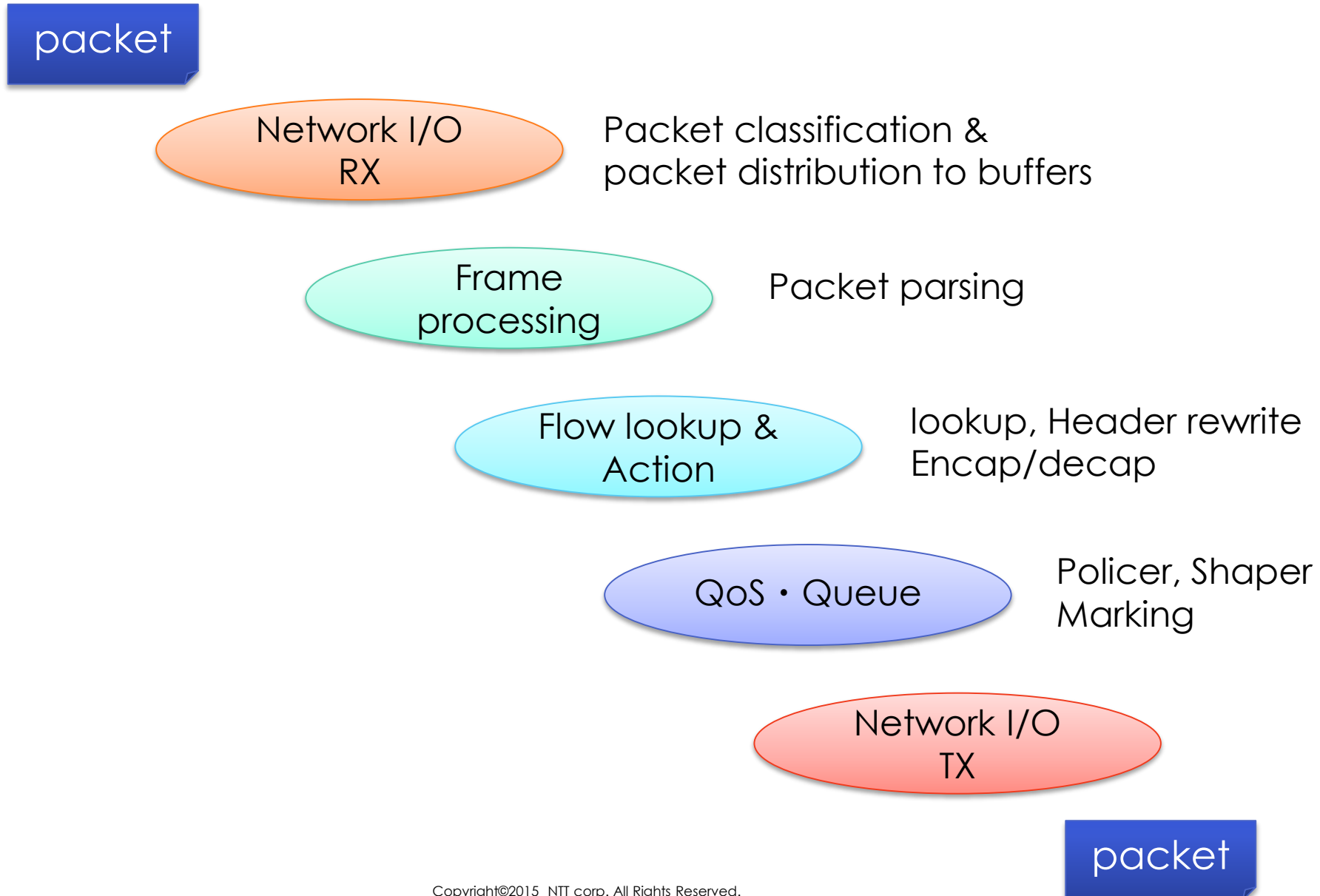
DPDK-enabled vNIC for NFV (virtq-pmd)



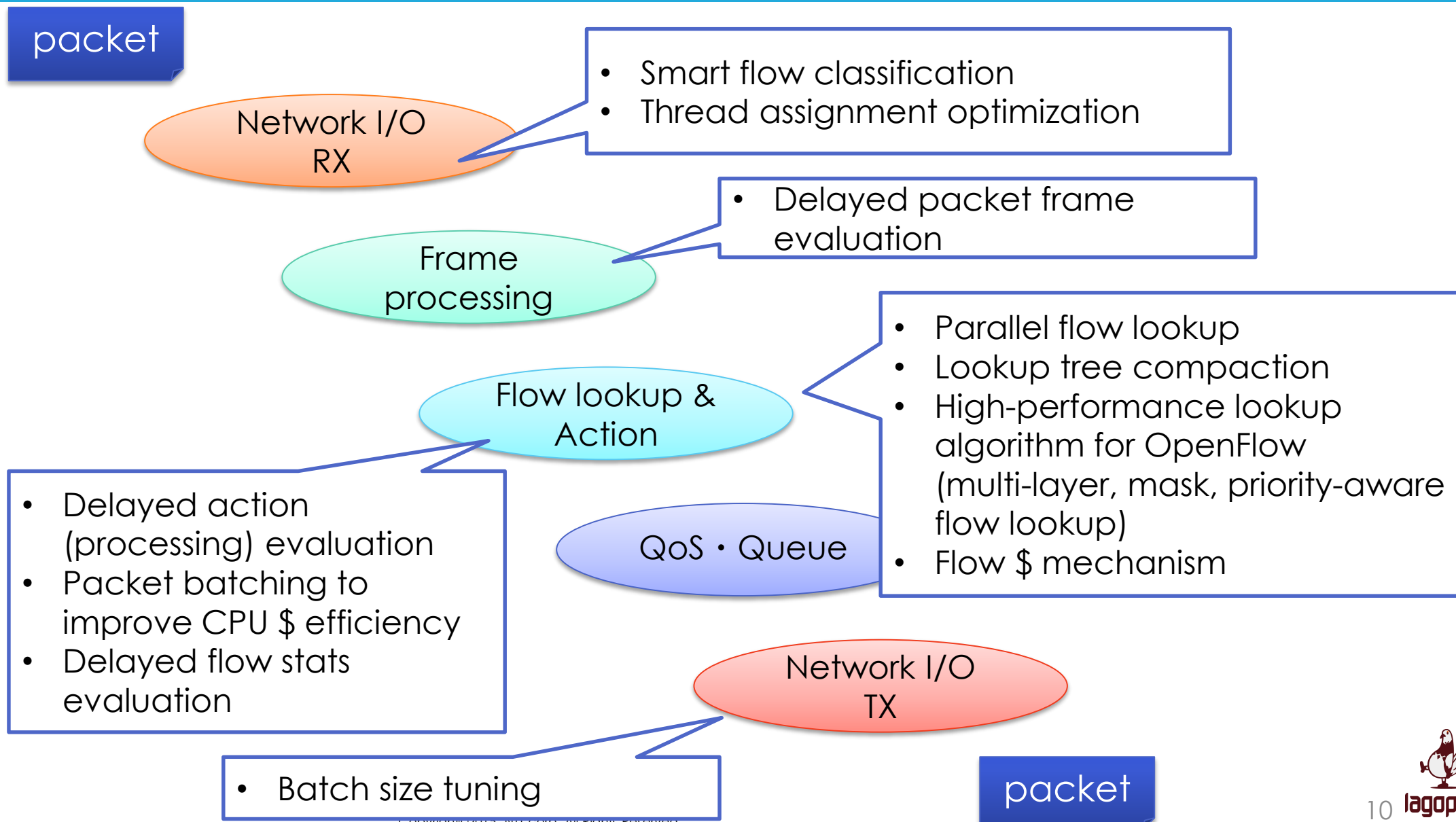
Innovative R&D by NTT

パケット処理の高速化

パケット処理の流れ



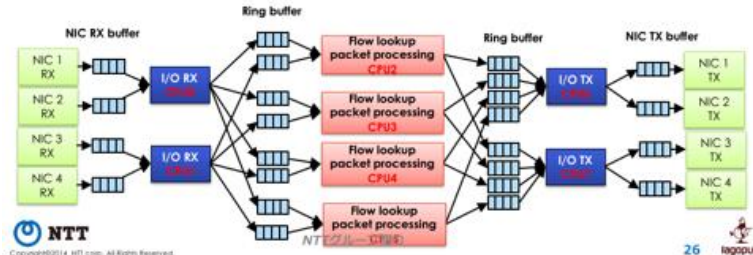
性能向上に向けての工夫や実装 (1/2)



性能向上に向けての工夫や実装 (2/2)

複数CPUコアによるパケット処理

- OpenFlow data plane processing
 - パケットはコピーせず参照渡し
 - パケット群に対するバッチ処理
- Exploit many core CPUs
 - I/O処理とFlow処理を分離
 - マルチコアCPUを利用したパイプライン化
 - CPU D-cache利用効率を向上

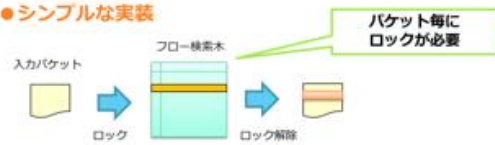


バッチ処理によるロック削減

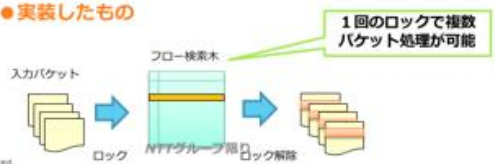
■ ロックを減らす実装

- AgentやFlow ruleのためのカウンタ管理のバッチ化
- 動的にフロールール設定が行われるのでstaticな構成は意味が無い

● シンプルな実装



● 実装したもの



バイパスするためのキャッシュ機構

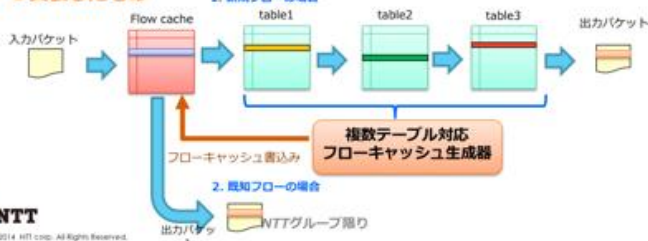
■ 毎回検索するのは資源の浪費

- とくに複数テーブル設定時
- 時間的局所性を利用 (流れたフローだけキャッシュ)

● シンプルな実装



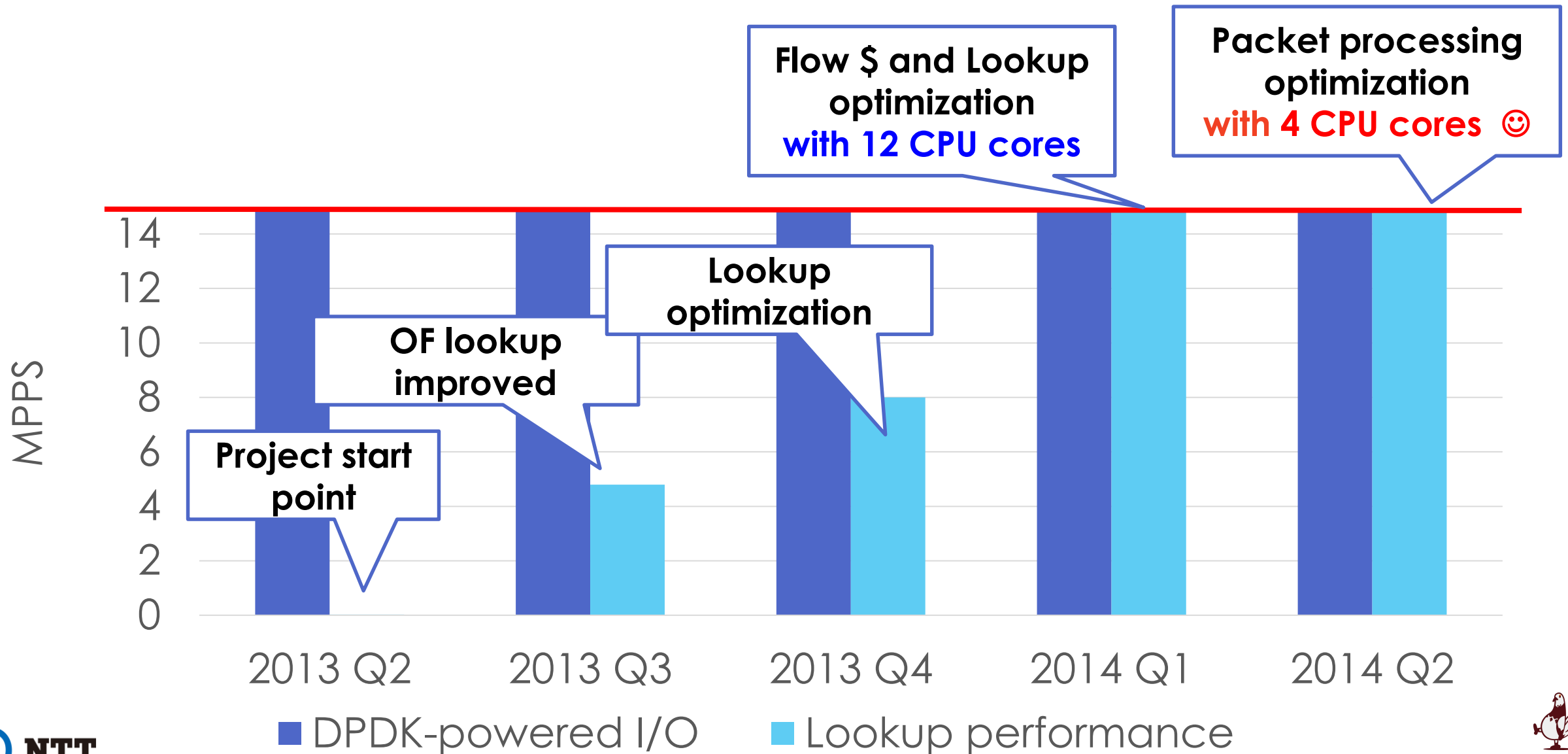
● 実装したもの



■ 詳しくはソースコードを!

- <https://github.com/lagopus/lagopus/>

10Gbpsパケット処理の実現にむけての進捗



Ryu OpenFlow 1.3 Conformance test Status (2015.02.12)

Type	Action	Set field	Match	Group	Meter	Total
# of test scenario (mandatory, optional)	56 (3, 53)	161 (0, 161)	714 (108, 606)	15 (3, 12)	36 (0, 36)	991 (114, 877)
Lagopus 2014.11.09	56 (3, 56)	161 (0, 161)	714 (108, 606)	15 (3, 12)	34 (0, 34)	980 (114, 866)
OVS (kernel) 2014.08.08	34 (3, 31)	96 (0, 96)	534 (108, 426)	6 (3, 3)	0 (0, 0)	670 (114, 556)
OVS (netdev) 2014.11.05	34 (3, 31)	102 (0, 102)	467 (93, 374)	8 (3, 5)	0 (0, 0)	611 (99, 556)
IVS 2015.02.11	17 (3, 14)	46 (0, 46)	323 (108, 229)	3 (0, 2)	0 (0, 0)	402 (111, 291)
ofswitch 2015.01.08	50 (3, 47)	100 (0, 100)	708 (108, 600)	15 (3, 12)	30 (0, 30)	962 (114, 848)
LINC 2015.01.29	24 (3, 21)	68 (0, 68)	428 (108, 320)	3 (3, 0)	4 (0, 4)	523 (114, 409)
Trema 2014.11.28	50 (3, 47)	159 (0, 159)	708 (108, 600)	15 (3, 12)	34 (0, 34)	966 (114, 854)



Innovative R&D by NTT

Lagopusを使った柔軟なネットワーク



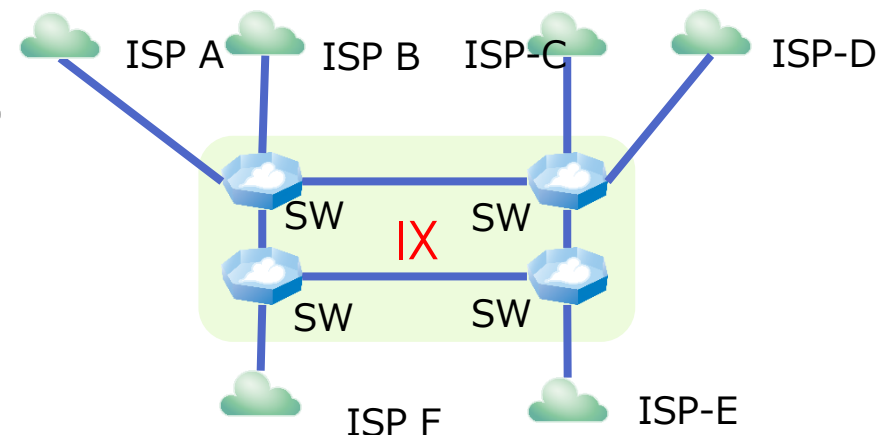
SDN IX @ Interop Tokyo 2015 ShowNet

**Thanks to NECOMA project
(NAIST & University of Tokyo)**

SDN IXの背景

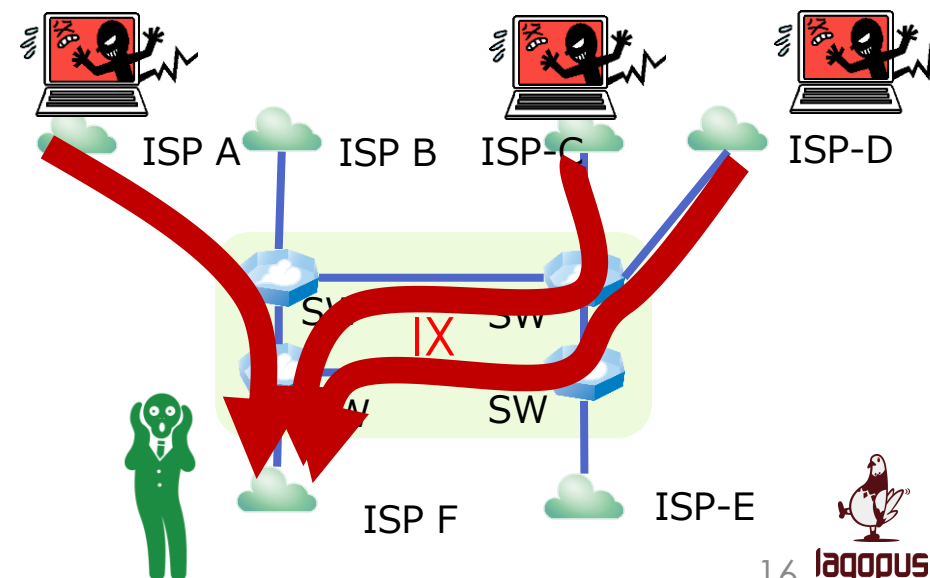
■ IX (Internet eXchange)

- インターネットサービス・プロバイダ (ISP) 間のL2での接続ポイント
 - ISP間のパブリックな接続・プライベート接続



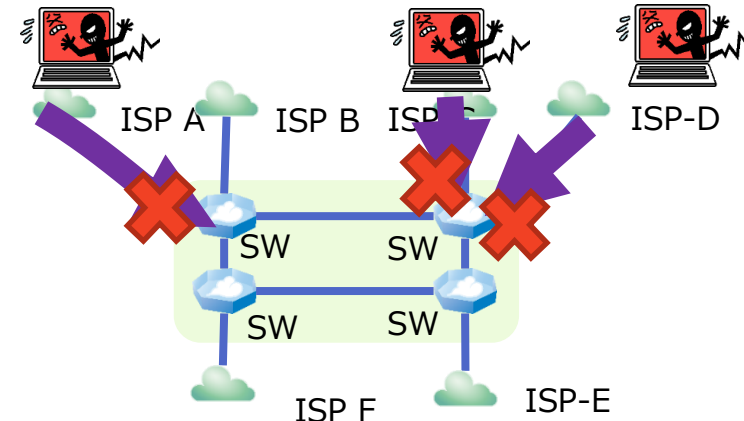
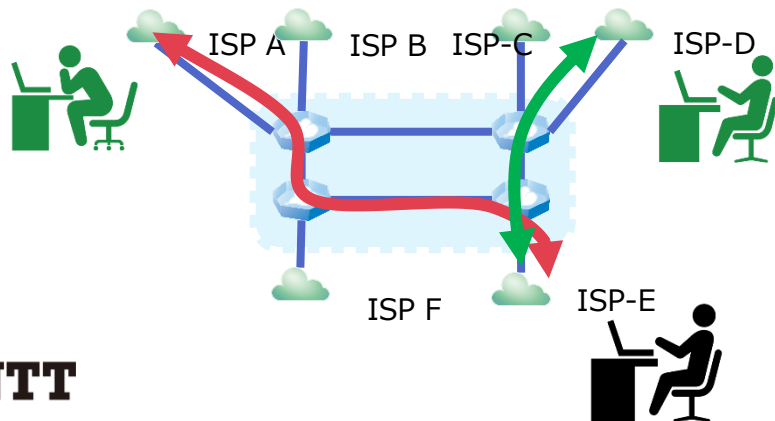
■ やりたいことや困っていること

- 接続設定の自動化とポータル化、API提供
- DDoS トラヒックを削減・止めたい
 - リンク帯域を食い尽くすことも
 - 入り口でDDoSトラヒックをdrop
 - 外部のDPI装置やFlowSpecとの連携



■ 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-based packet filtering
- SDN IX controller and distributed SDN/OpenFlow IX core switch

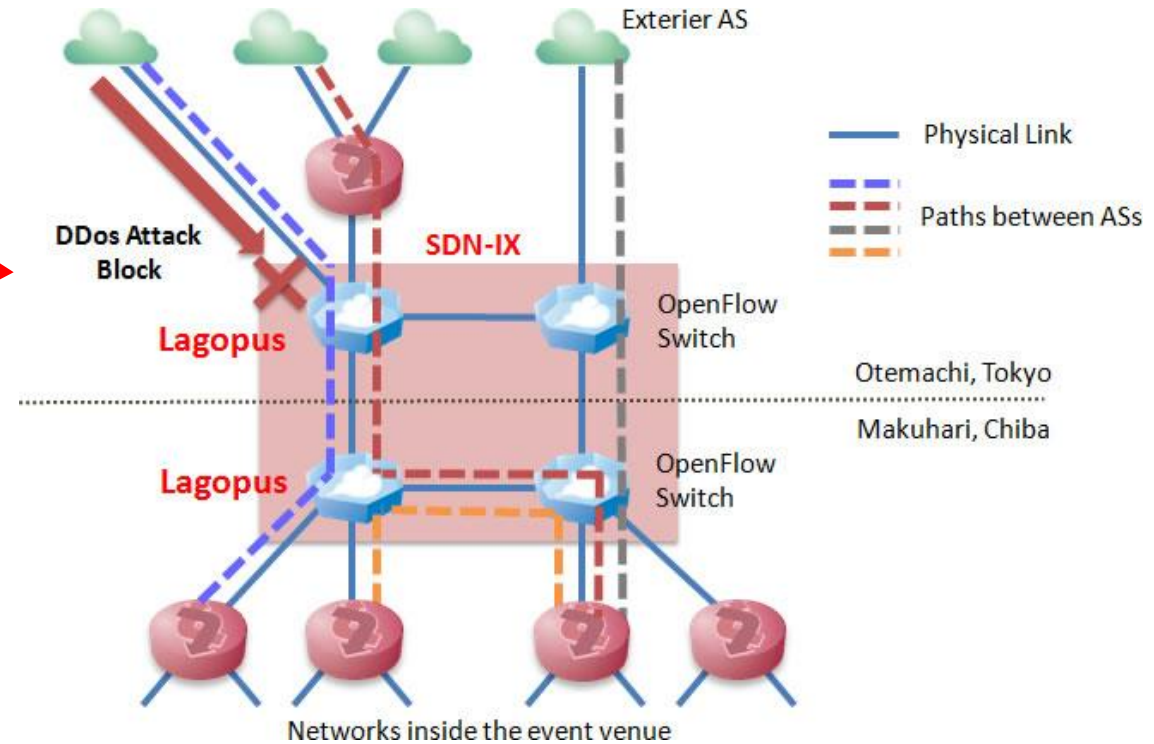
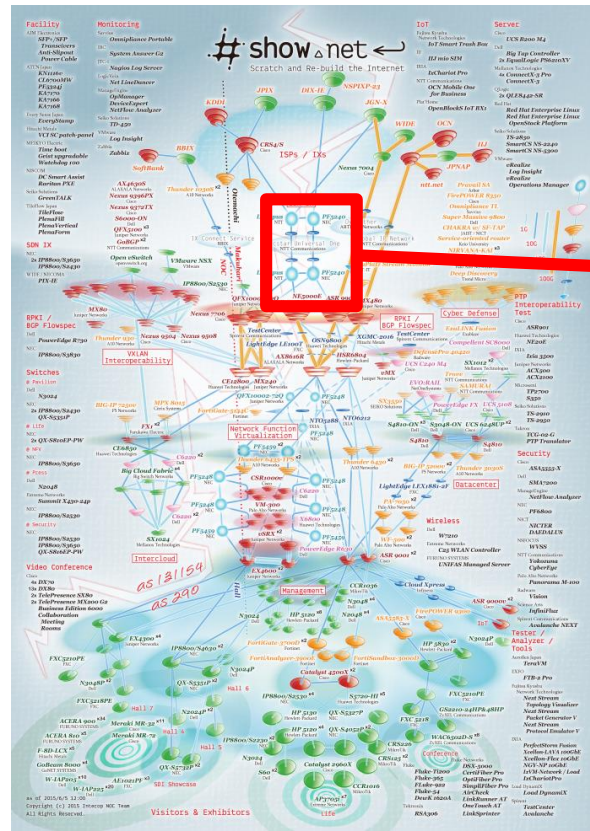


Lagopus @ ShowNet 2015

■ 2台のLagopusを設置@大手町と幕張

● アキバで買えるIntel Xeonサーバ

- 2 x Xeon E5 (Sandy-bridge) 8core CPU
- 6 x 10GbE NIC



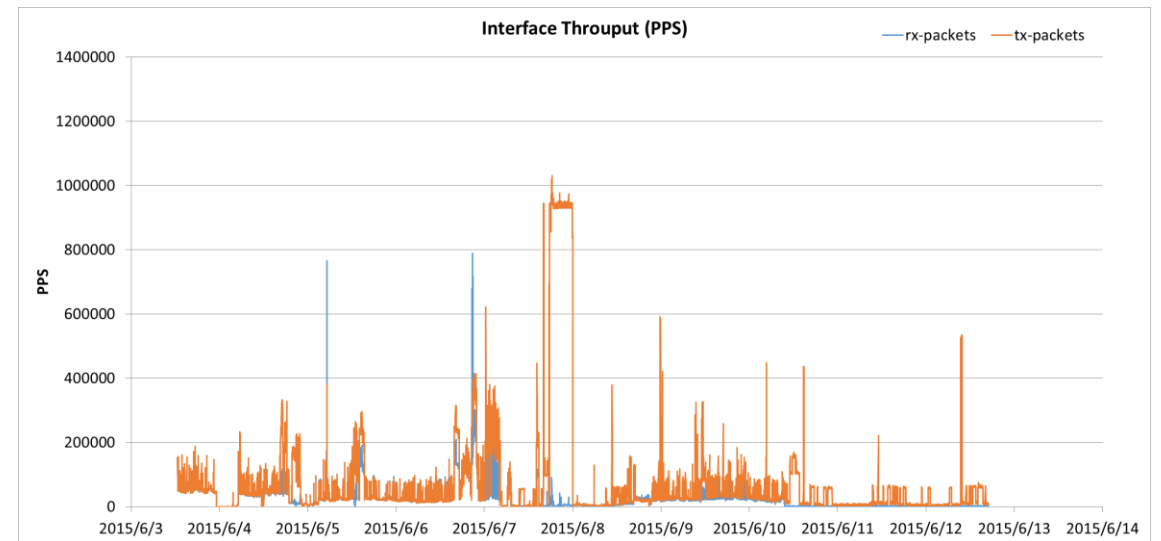
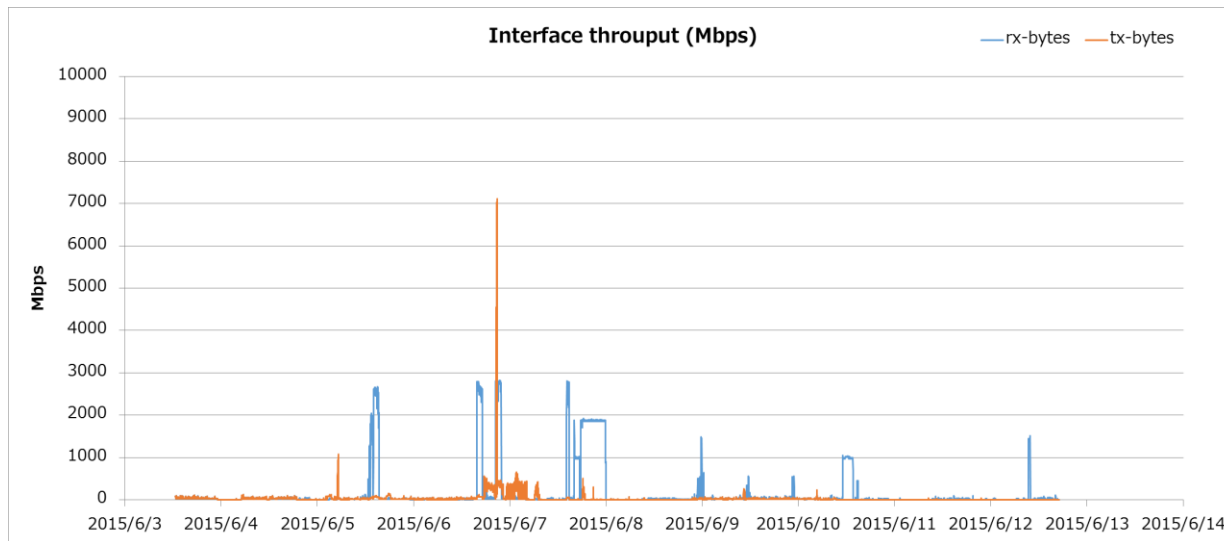
Lagopus @ ShowNet rack



Lagopus@幕張を通ったトラヒック

■ 平均2Gbps throughput

- パケットドロップは発生なし
- ときたま10Gbpsのトラヒックにも耐えた
- Interop Tokyo会期中リブートなし & トラブルなし





Innovative R&D by NTT

Segment Routing with MPLS by Lagopus + Routing Stack

**Source/location-aware traffic control with scalable
software-defined networking with MPLS and NFV
appliance**

SPRING using MPLS

■ 集中制御と分散制御のいいところどりのNWアーキテクチャ

- MPLSネットワークの簡易化 + ラベルスタックベースの転送
- ラベルスタックに基づいたパケット転送
- Segment Routing, Source routing
- NFVとの連携が可能

■ 何につかえるの？

- ポリシールーティング
 - ・ 最短ルートやパス制御との連携
- サービスチェイニング
 - ・ 特定のソースやロケーション及びフローに対しての適応

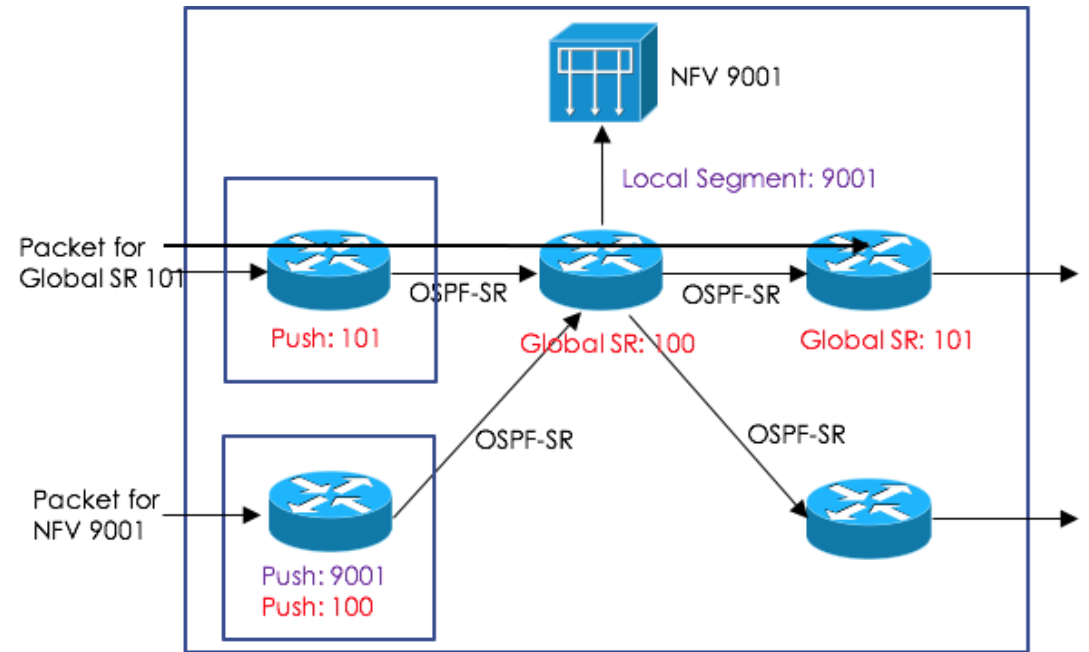
SPRINGの動作

■ エッジルータが転送パスとポリシーを表現するラベルスタックをプッシュ

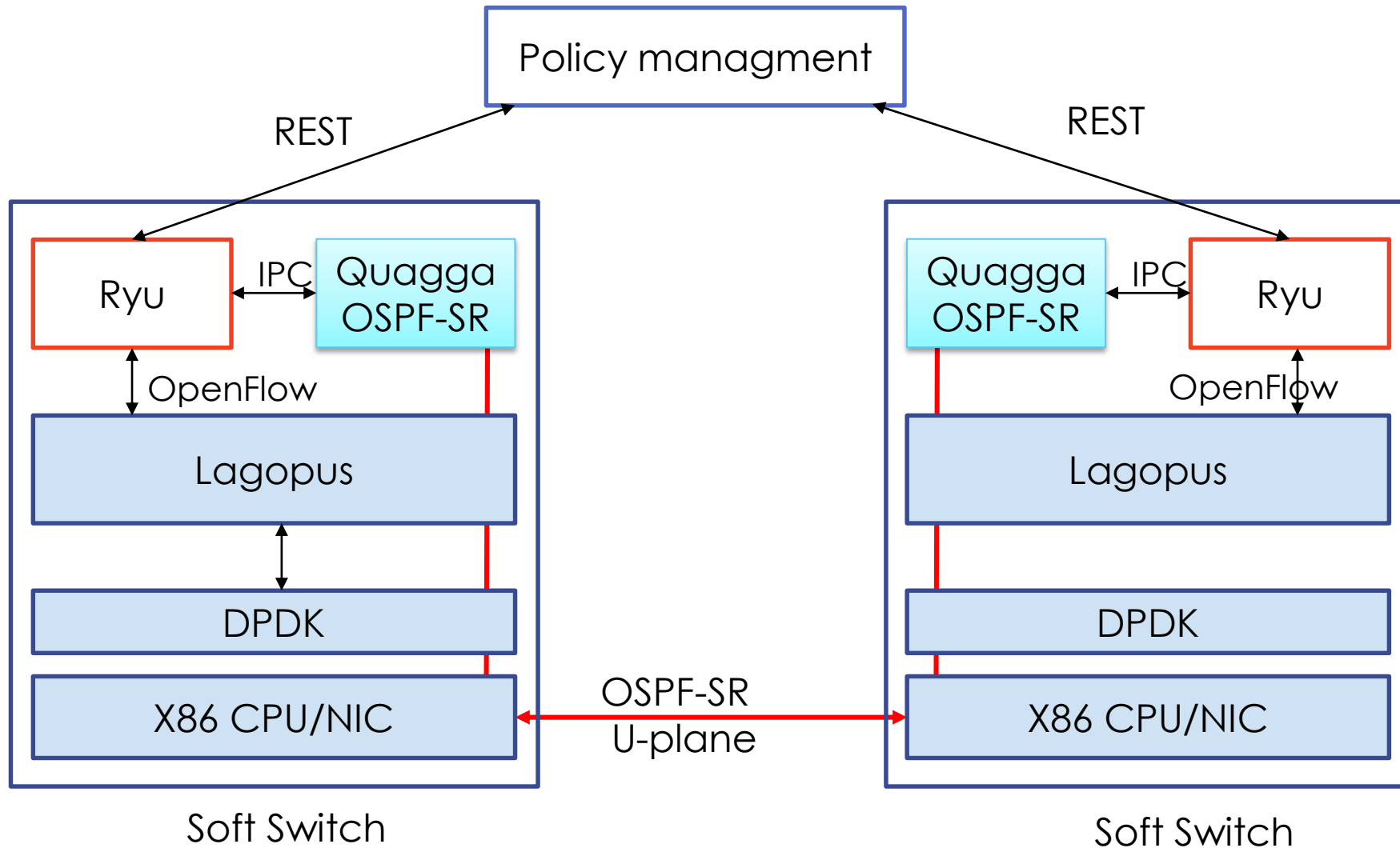
- ポリシー管理は上位のマネージメントツールに基づく制御

■ ルータはラベルに基づきパケット転送

- パス計算は各ルータにおけるCSPFベースによる最短経路転送
- リンクダウンへの対応は各時



SPRING implementation with Ryu & Lagopus

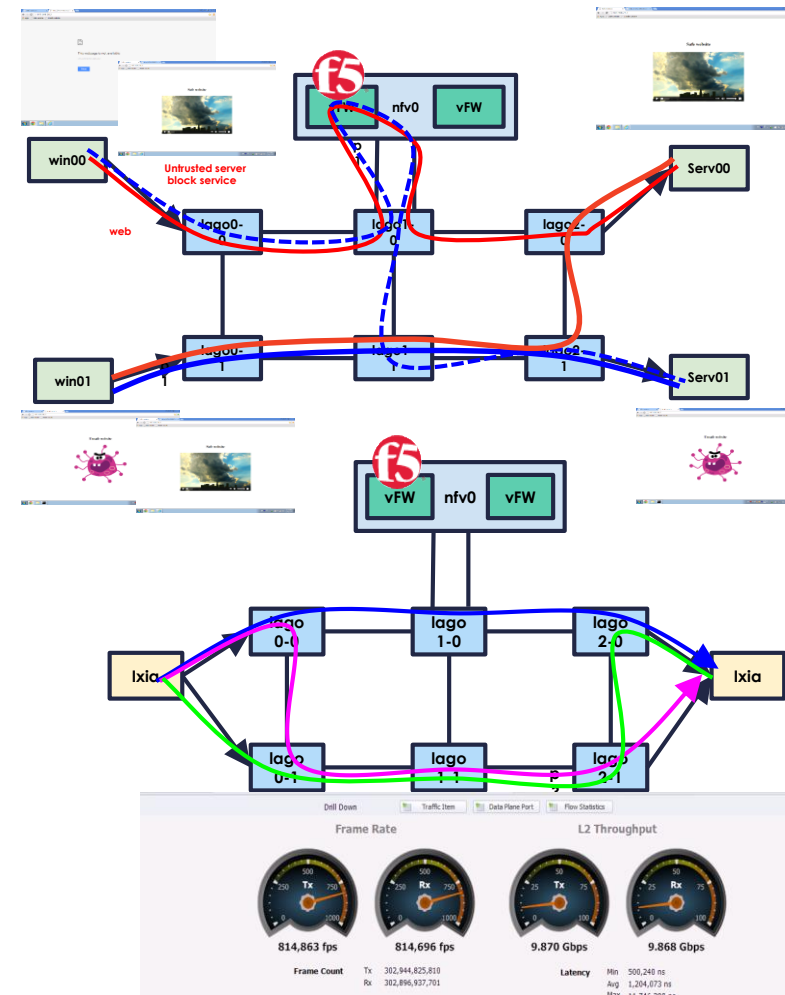


SPRINGの PoC

■ ロケーション指向の packets 転送 + サービスチェイニング

■ 複数の仮想NW

- Enterprise intra network
 - Web service and clients
 - Untrusted site blocking by NFV service
- Lab network
 - Policy management
 - Explicit routing for TE



まとめ

■ 高速なLagopusソフトスイッチをオープンソースとして提供

- みなさまの研究テーマとして活用してみませんか？
 - 既存のサーバで動作可能
 - 新しいNW制御研究や高速化の研究のツールとして
 - 相談にのります
- 学術系におけるスイッチの教育材料として利用してみませんか？



Web: <http://lagopus.github.io/>
Github: <https://github.com/lagopus/>

開発に参加してみませんか？
Lagopus本とかも後日公開 (予定)