

ひとつのRFCがプライベートクラウドの 仮想ネットワークを3倍速くした話

KDDI株式会社

山口雄翔

自己紹介

■ 基本情報

- 名前: 山口 雄翔 (やまぐち ゆうしょう)
- 2025年 途中でKDDI入社

■ 業務内容

- プライベートクラウドMShip3の開発運用

■ 好きなもの

- ソフトウェアパケット処理 / OSS



JANOG53 ミーティング
フォトアルバムより

MShip3

■ プライベートクラウド

- NFV含む種々のVMが動く
 - ・ 高可用性/高性能が求められる
- 商用環境は以下の規模
 - ・ 6リージョン
 - ・ ハイパーバイザ4000台以上
 - ・ VM20000台以上

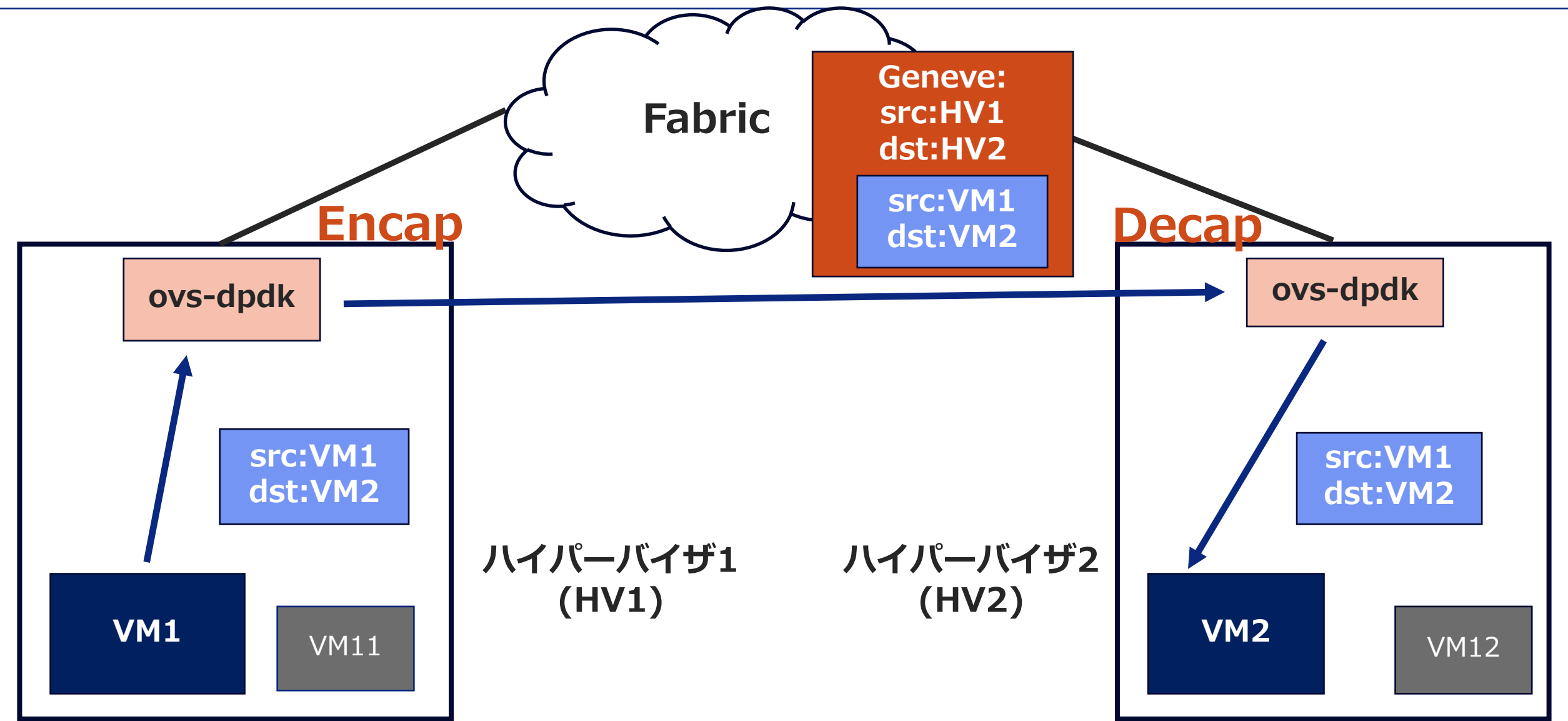


■ 仮想化基盤にはOpenStackを利用

- マルチテナントネットワークを実現
 - ・ OVN/OpenvSwitch(OVS-DPDK)による高性能仮想L2ネットワーク



MShip3のVM間仮想L2ネットワーク

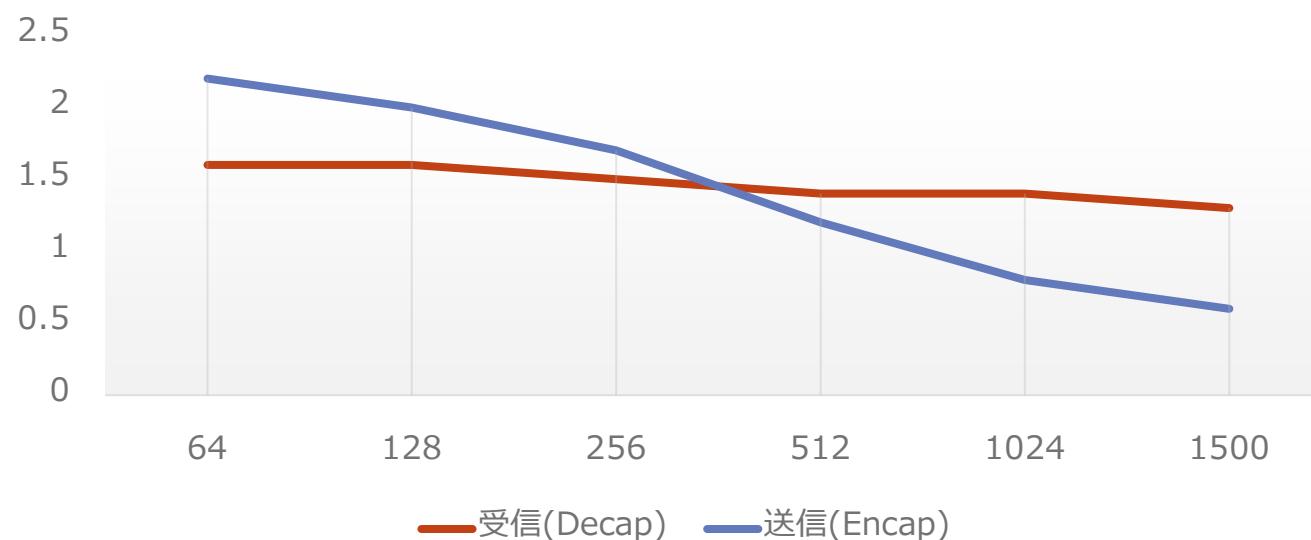


性能課題

- 送信(Encap)側のovs-dpdkが処理性能上限に達しているアラームが頻発
- 性能試験を実施

パケットサイズを変化させた時の

Encap/Decap性能

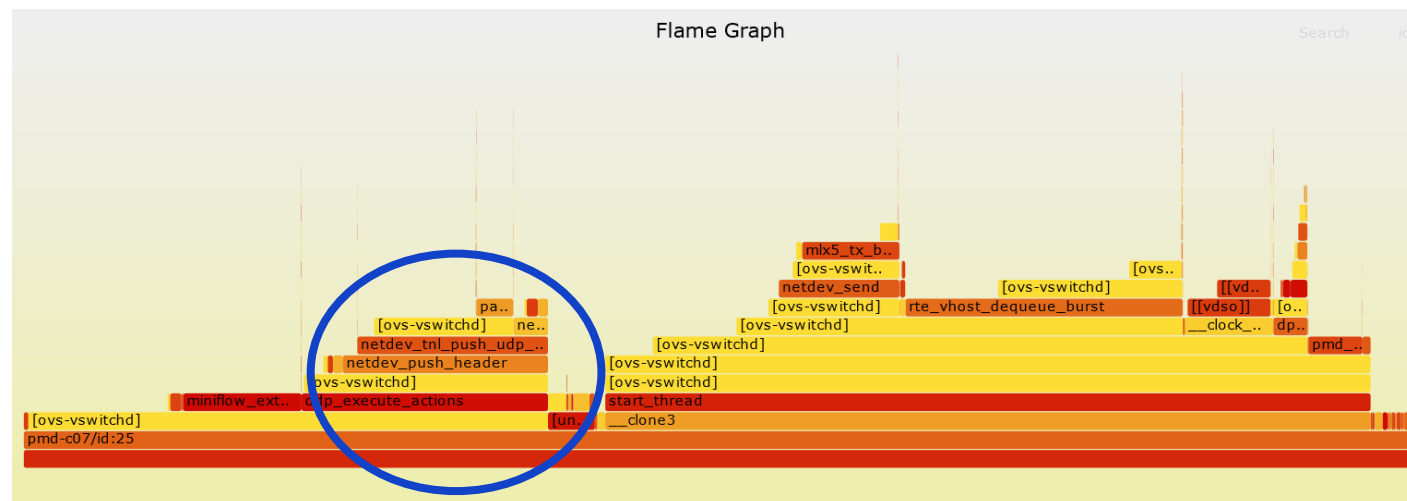


パケットサイズが大きくなるにつれて、
送信(Encap)側が顕著に性能劣化している

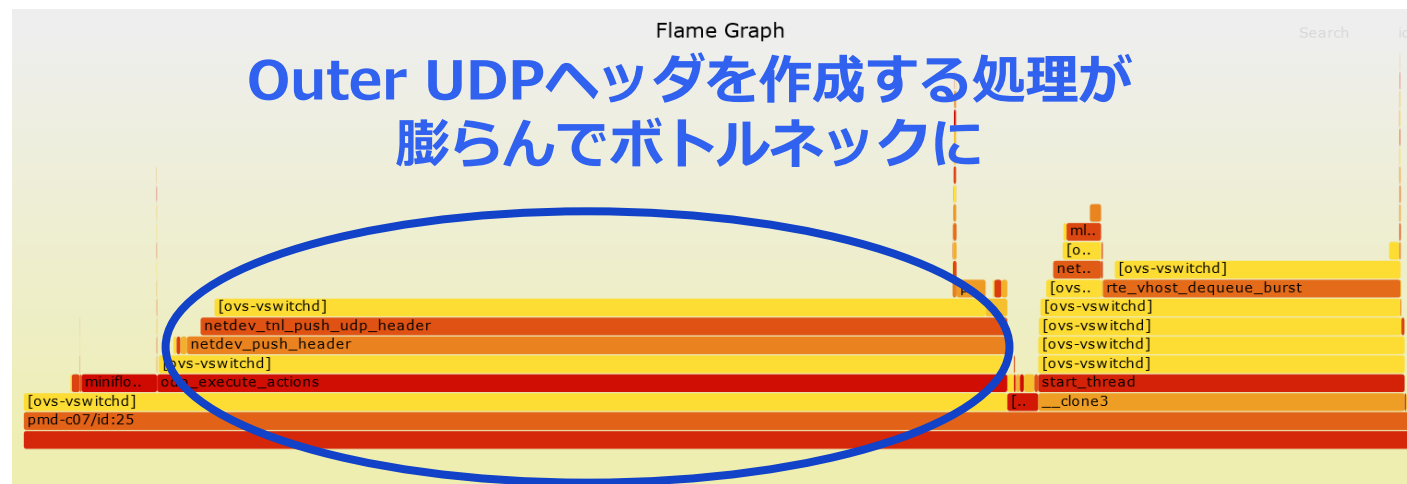
OVSに負荷をかけた状態で処理のボトルネックを分析

負荷パケットの
サイズ

64byte



1000byte



■ さらなる調査の結果、Outer UDPヘッダの**チェックサム計算処理**がボトルネックであると推定された

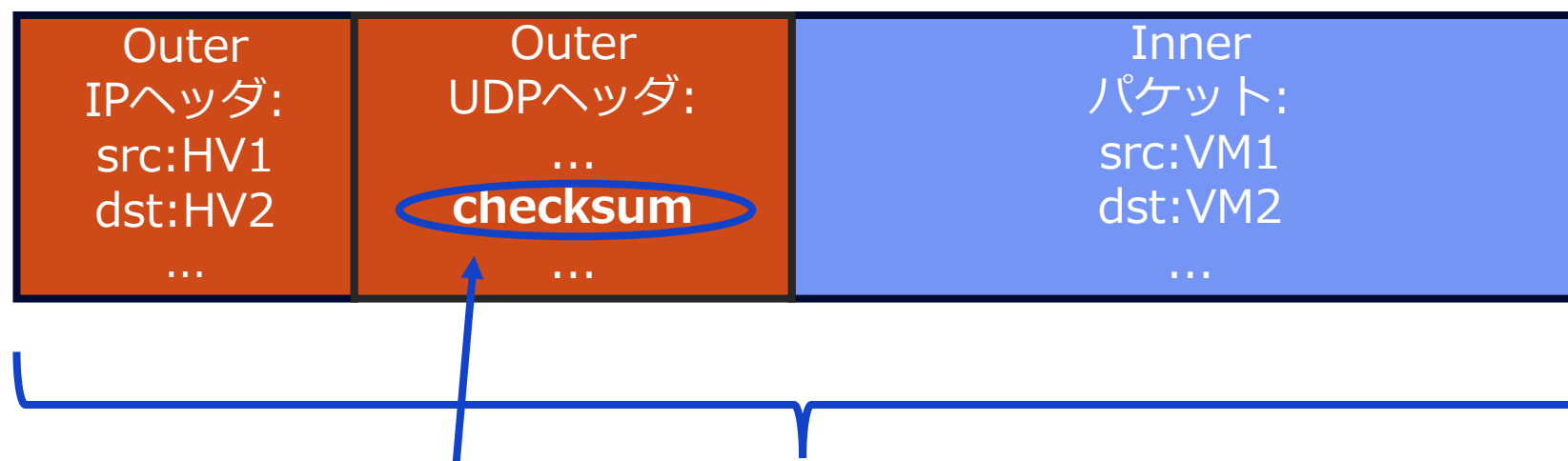
チェックサム計算について

■ チェックサムとは

- データ破損を検出するためにパケットに埋め込む検査値

■ 今回問題になったチェックサム計算

- Geneve のOuter UDPヘッダのチェックサム計算→書き込みの処理



この計算
パケット全体を走査して行う実装になっている

(図は厳密ではなくイメージです)

このチェックサム計算は必要か？

■ RFC2460 (現在のRFC8200)

- UDP over IPv6ではUDPチェックサム必須
 - ・ IPv6ヘッダ自体がチェックサムを含まないため
- IPv4では省略することが認められていて、運用でも使われている

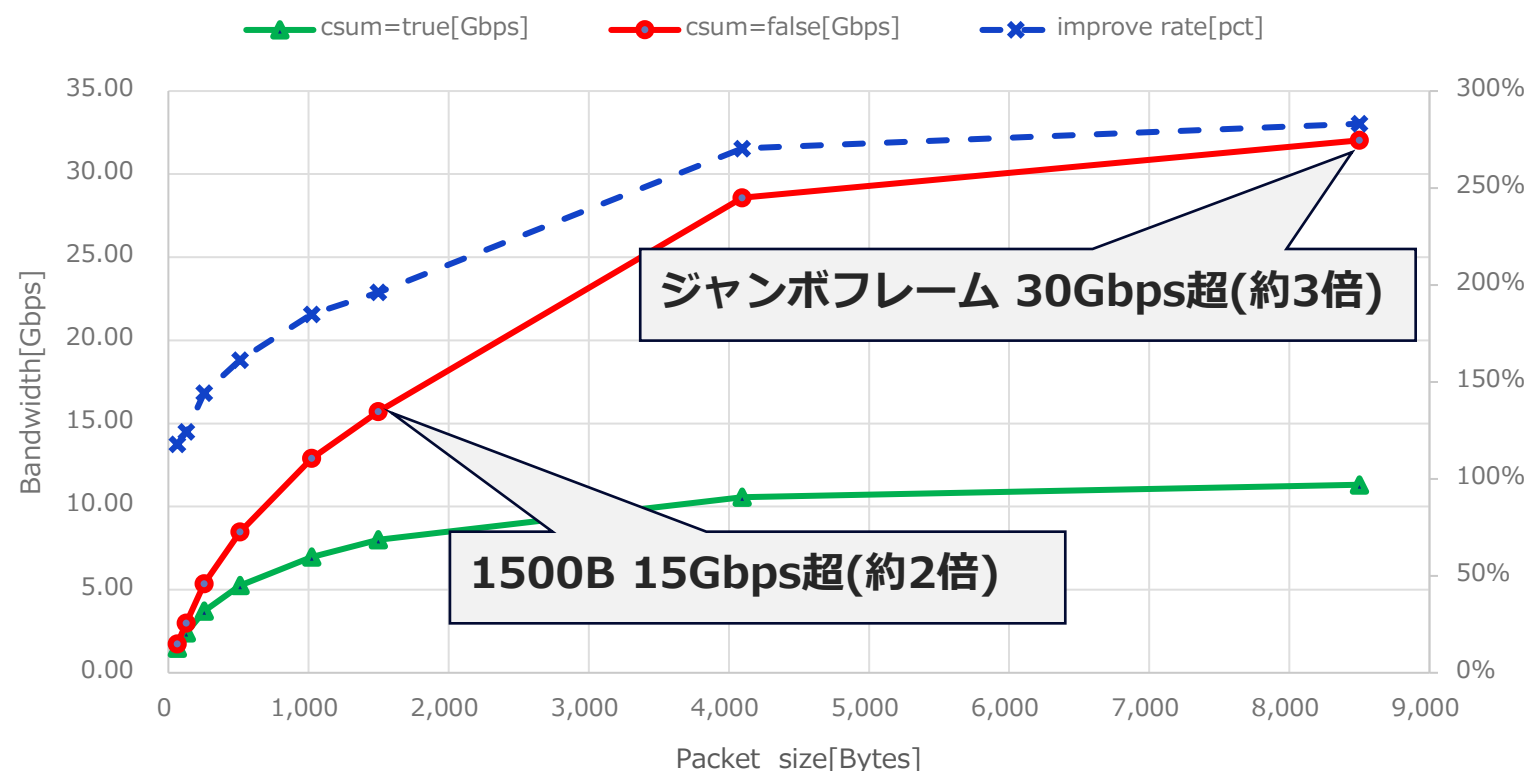
----- このオーバーヘッドは受け入れるしかないのか... ところが

■ RFC6935

- トンネルのOuterのUDPヘッダに限り、チェックサム計算を無効化することを認める
 - データ破損の検知はオーバーレイ側で行う前提
- このRFCをサポートする実装がOVS-DPDKにあるか...？

対応

- 2024年に、OVSにチェックサム無効をIPv6でサポートするパッチがマージ
- その機能を取り込んで試験したところ、大幅な性能改善を確認



青点線:改善率(第二軸)

赤:チェックサム無効
(新バージョン)

緑:チェックサム有効
(従来)

9000byte近くのジャンボフレームについては、
約3倍の送信性能改善を確認 (近日商用リリース予定)

最後に

- 本発表の内容の詳細は、KDDI公式Tech noteで公開済み
 - <https://tech-note.kddi.com/n/n74ad198da05c>
- “KDDI Tech note” でぜひ検索してください！

「つなぐチカラ」を進化させ、
誰もが思いを実現できる社会をつくる。

KDDI VISION 2030

