

DPUで5GC UPFを実装してみた ～Re: DPUから始めるUPF内製化～

NTTドコモ
宮本 克真

日本電信電話
平井 志久

2024/01/24 JANOG55@京都

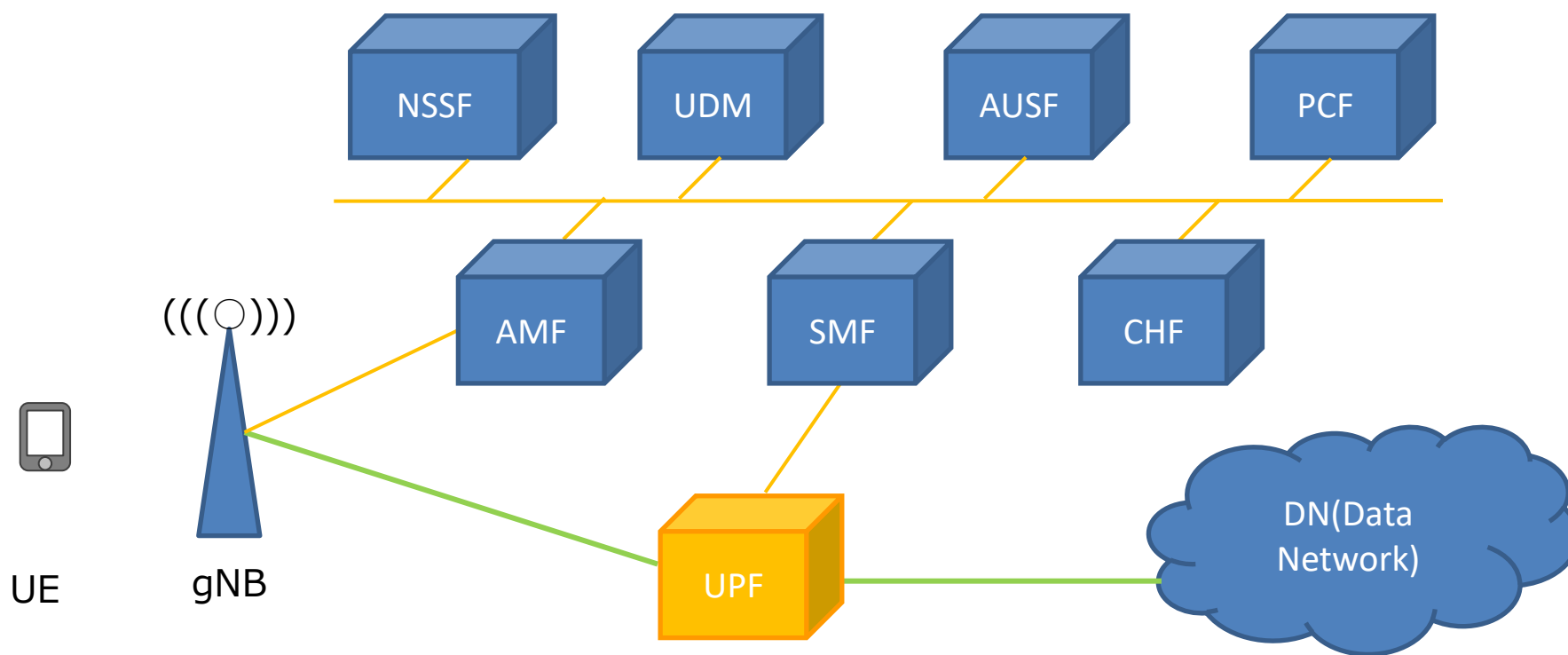
目次

- ✓ 取り組みの背景
- ✓ UPFの実装
- ✓ 商用化に向けた検討
- ✓ ディスカッション

取り組みの背景

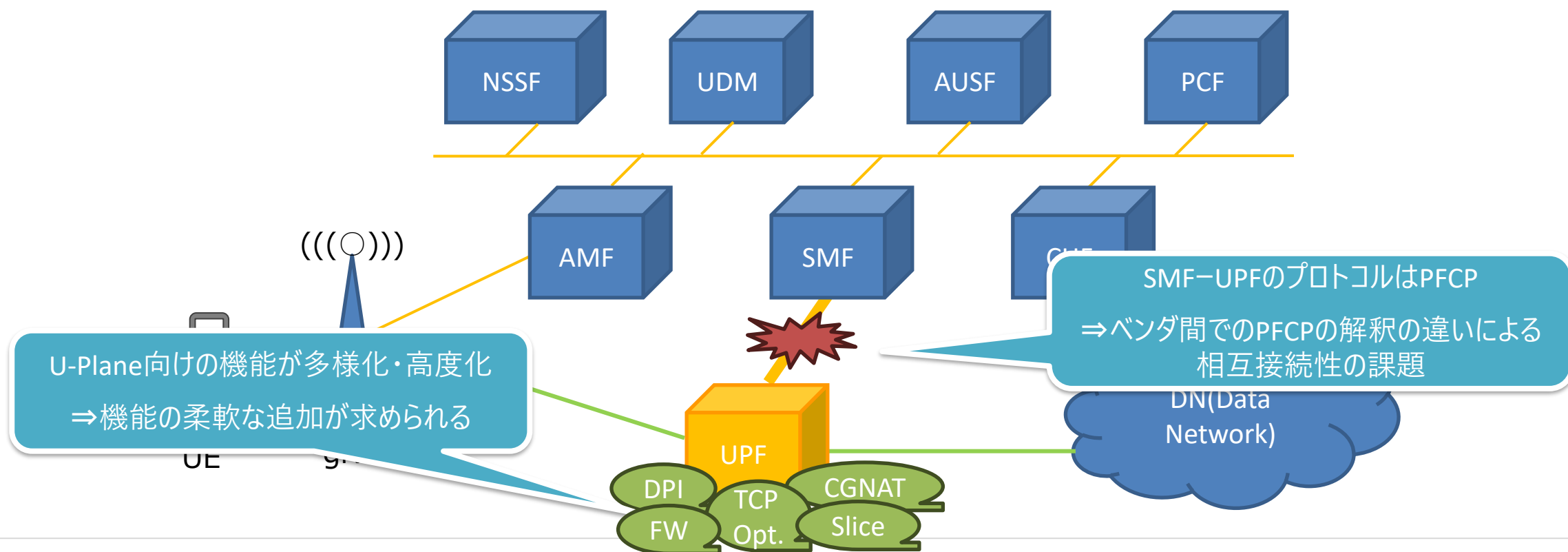
5GコアにおけるUPF

- 5Gコアはコントロールプレーンとユーザプレーンに分かれている
- UPF（User Plane Function）はユーザ通信を扱う機能部
- PDUセッションを確立することで、UEはDN（インターネット）に通信できる



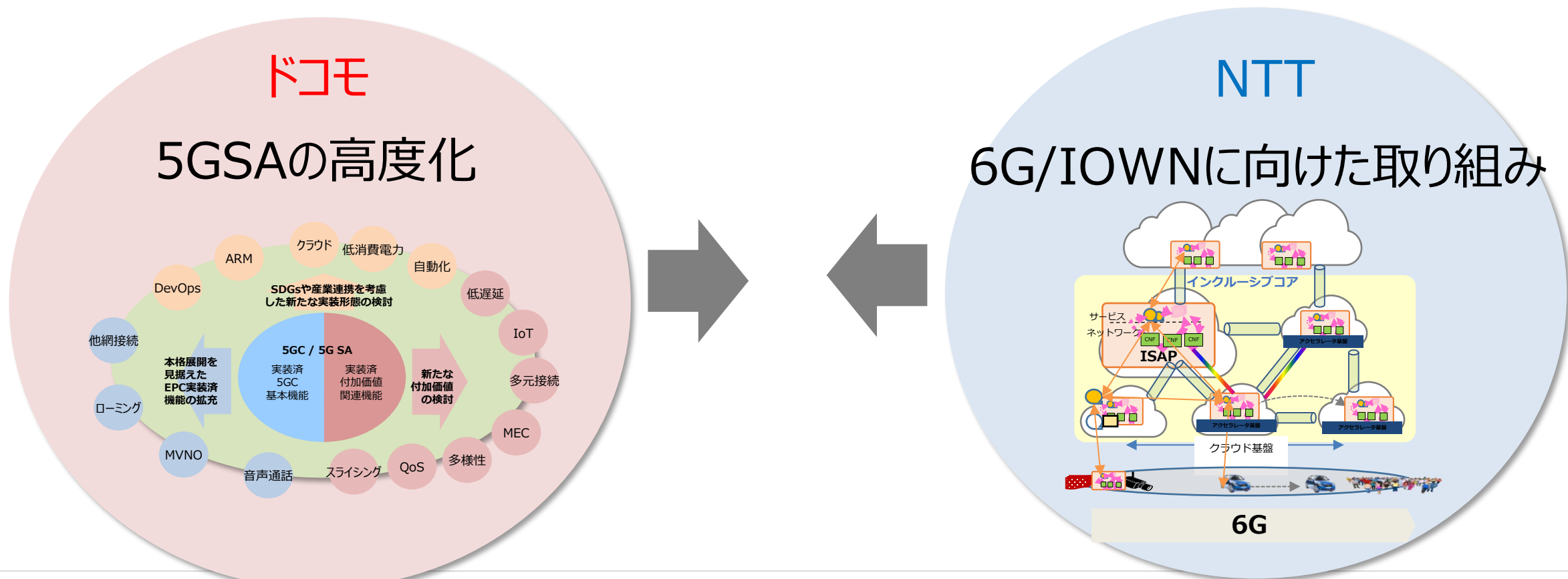
UPFの特徴・課題

- UPFには以下の特徴・課題がある
 - PFCP（SMF-UPF間のプロトコル）がベンダ間で解釈が異なる
 - 内部や周辺にさまざまな機能や要件が求められる
- 相互接続性や高機能化という点で、UPFには実装の柔軟性が求められる



UPFの内製化とIOWNの実現

SMFとの接続対応や高機能化へ対応するための一手段として、
ハードウェアアクセラレータであるDPUを用いたUPFの内製化に取り組む
内製UPFの導入を皮切りに、モバイルネットワーク全体にIOWNを適用していく



- DPU上にUPFを実装し、ShowNetのローカル5Gに接続して使ってみた
- PoCレベルではある程度動くところまで確認することができた



前回の議論で得られたこと

NVIDIA BlueField3を選んだ理由は？
他のDPUはどうか？

機能がいろいろあるが、
どこまで機能実装するの？

P-GWの内製化に取り組んだことがある
GW装置を手の内化できるのは強い

UPFのような台数が多いものは
内製によりコスト削減できそう

(内製する大変さはありつつも) オペレータとしてUPFを内製化することには大いに価値がありそう



今回、商用化に向けてさらなる実装改善や性能検証を行いました

得られた技術課題や商用化に向けた取り組みについて共有・議論させてください

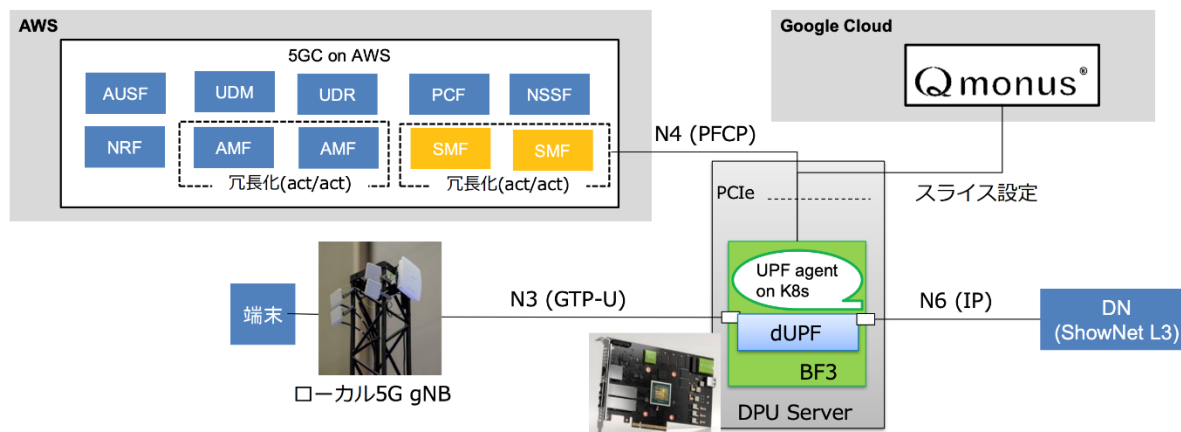
UPFの実装

前回(JANOG54)振り返り

Interop Tokyo/ShowNet 2024への出展



- ドコモの5GC on AWS上のSMFとのN4 (PFCP)相互接続
- コミュニケーションズのQmonus (E2Eオーケストレータ)からのスライス設定・制御

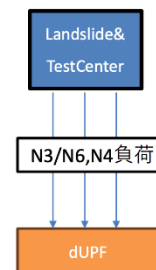


Copyright 2024 NTT CORPORATION

25

キャリアを想定した性能測定結果

- ShowNet期間においてSpirent様のTestCenter/Landslideを用いて各種性能評価を実施
- キャリアグレード性能検証の前段として、十分な性能を確認することが出来た。



項目	結果
最大セッション数	54,000セッション
PFCPトランザクション数	640TPS
4000基地局収容 (GTP-Path 4000)	双方向(パケット長650byte) UL 56Gbps DL 96Gbps
	Uplink(パケット長120byte) 66Gbps Downlink(パケット長1200byte) 96Gbps
1ユーザ収容	Uplink(パケット長120byte) 94Gbps Downlink(パケット長1200byte) 100Gbps

※dUPF1カードあたりの性能

©2024 NTT DOCOMO, INC. All Rights Reserved.

41

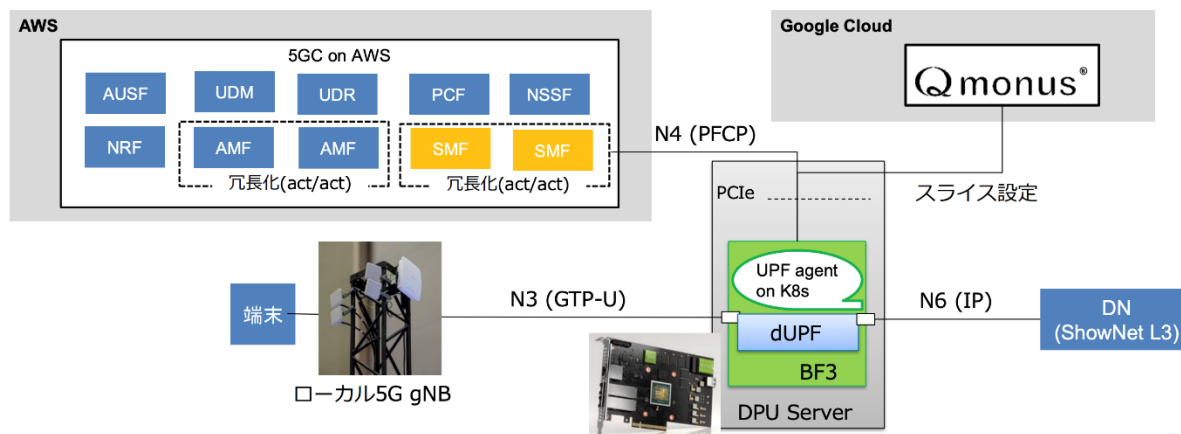
<https://www.janog.gr.jp/meeting/janog54/dpuupf/>

前回(JANOG54)振り返り

Interop Tokyo/ShowNet 2024への出展



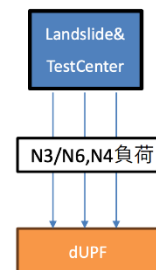
- ドコモの5GC on AWS上のSMFとのN4 (PFCP)相互接続
- コミュニケーションズのQmonus (E2Eオーケストレータ)からのスライス設定・制御



25

キャリアを想定した性能測定結果

- ShowNet期間においてSpirent様のTestCenter/Landslideを用いて各種性能評価を実施
- キャリアグレード性能検証の前段として、十分な性能を確認することが出来た。



項目	結果
最大セッション数	54,000セッション
PFCPトランザクション数	640TPS
4000基地局収容 (GTP-Path 4000)	双方向(パケット長650byte) UL 56Gbps DL 96Gbps
	Uplink(パケット長120byte) 66Gbps Downlink(パケット長1200byte) 96Gbps
1ユーザ収容	Uplink(パケット長120byte) 94Gbps Downlink(パケット長1200byte) 100Gbps

※dUPF1カードあたりの性能

©2024 NTT DOCOMO, INC. All Rights Reserved.

41

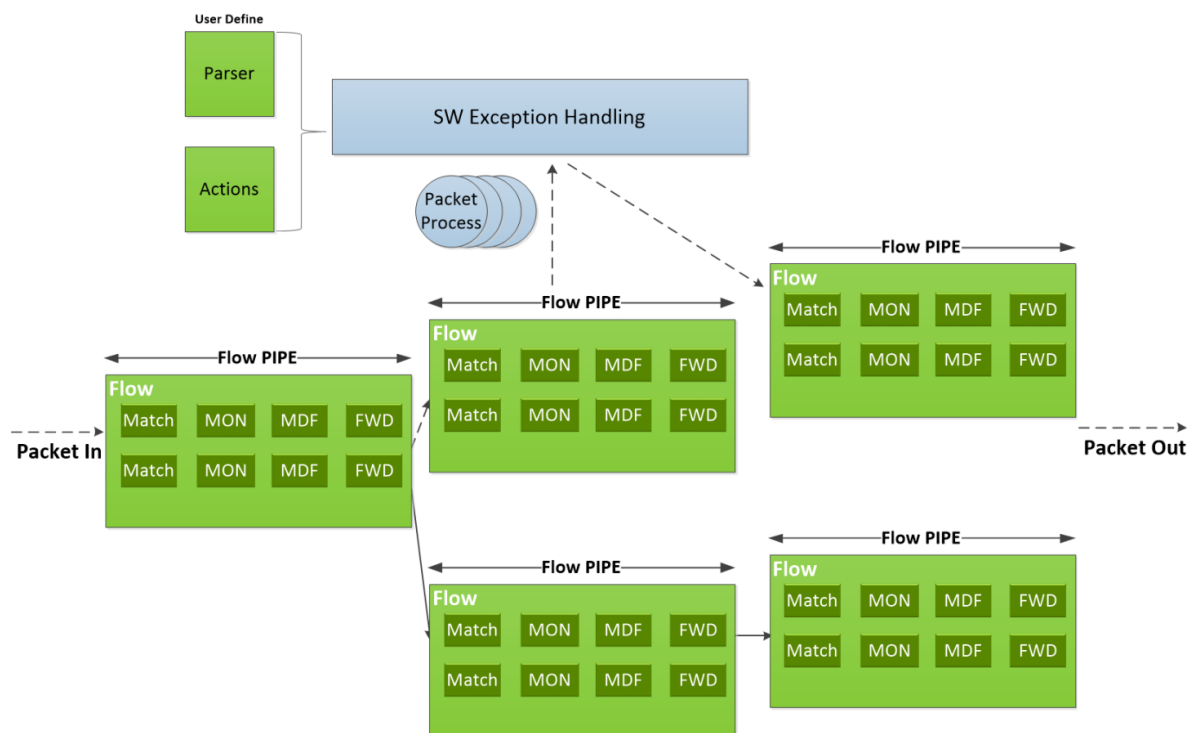
今回の発表では特に、、、

□ 一部実装を見直してキャリアグレードのより大規模な性能試験にチャレンジ！

□ ドコモ商用要件の細部に基づいて商用導入に向けた技術課題を網羅的に分析！！

実装見直しポイント

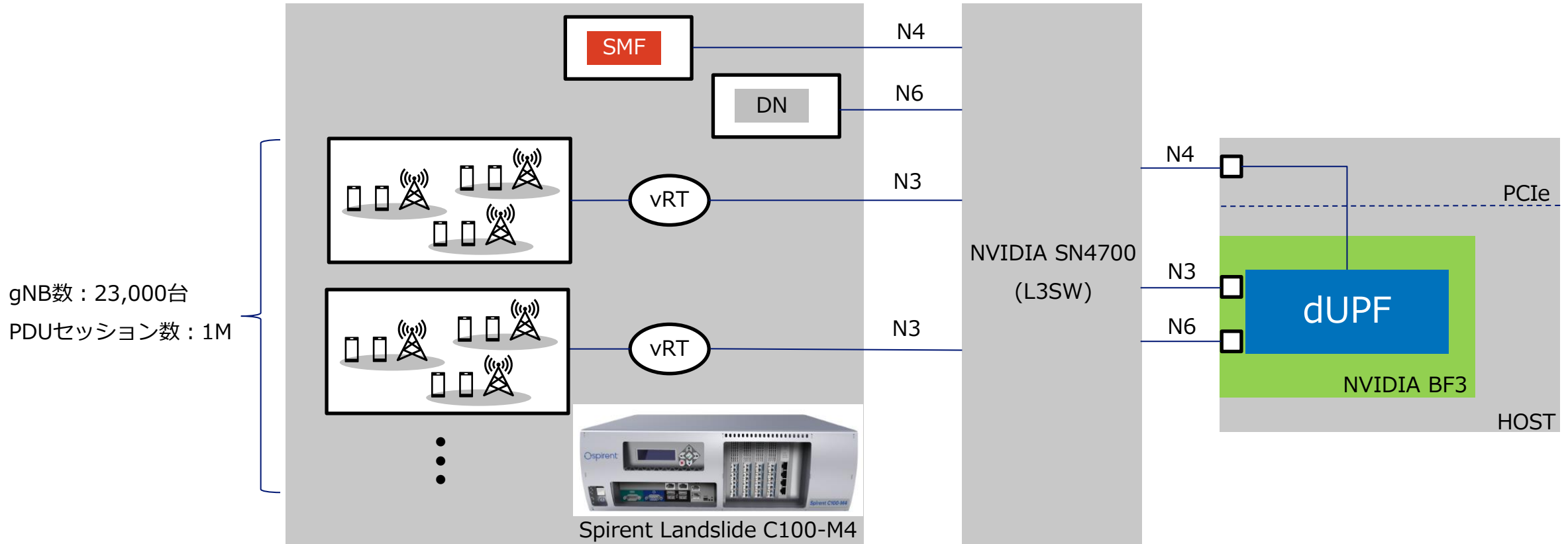
- ユーザセッション(PDUセッション)毎にフローエントリを投入する部分を、NVIDIA BlueField-3 (BF3)上のHW構造を考慮、フロー構成(各パイプでの処理定義と組合せ)を最適化
→ PDUセッション容量の拡大とトランザクション性能の向上を確認



DOCA Flowパイプライン・アーキテクチャ

(https://docs.nvidia.com/doca/sdk/doca+flow/index.html#src-3354484368_id-.DOCAFlowv2.9.1-Architecture)

- 今回もSpirent様/東陽テクニカ様のご支援の下、NTT武蔵野研究開発センタ内に環境構築
 - Landslide Virtual Test Manager x 1台
 - Landslide C100-M4 Test Server x 6台 (10Gポート x 36、40Gポート x 4、100Gポート x 8)



https://www.spirent.com/assets/u/landslide_c100-m4_datasheet

項目		結果
最大セッション数		1,000,000 セッション
PFCPTランザクション数		最大 1,000 TPS (500 session activation/秒)
23,000基地局収容 1M PDUセッション	Downlink only (パケットサイズ: 1200B)	11.7 Mpps 112 Gbps
	Uplink only (パケットサイズ: 120B)	? Mpps ? Gbps
	Bi-directional (パケットサイズ: 650B)	? Mpps ? Gbps

※DPUカード(BF3)1枚あたりの性能値

項目		結果
最大セッション数		1,000,000 セッション
PFCPTランザクション数		最大 1,000 TPS (500 session activation/秒)
23,000基地局収容 1M PDUセッション	Downlink only (パケットサイズ: 1200B)	11.7 Mpps 112 Gbps
	Uplink only (パケットサイズ: 120B)	? Mpps ? Gbps
	Bi-directional (パケットサイズ: 650B)	? Mpps ? Gbps

※DPUカード(BF3)1枚あたりの性能値

あくまでPoCレベルの実装の結果
→ 商用要件を考慮した設計・実装で
どこまで性能が担保されるか？

試験結果



項目		結果
最大セッション数		1,000,000 セッション
PFCPTランザクション数		最大 1,000 TPS (500 session activation/秒)
23,000基地局収容 1M PDUセッション	Downlink only (パケットサイズ: 1200B)	11.7 Mpps 112 Gbps
	Uplink only (パケットサイズ: 120B)	? Mpps ? Gbps
	Bi-directional (パケットサイズ: 650B)	? Mpps ? Gbps

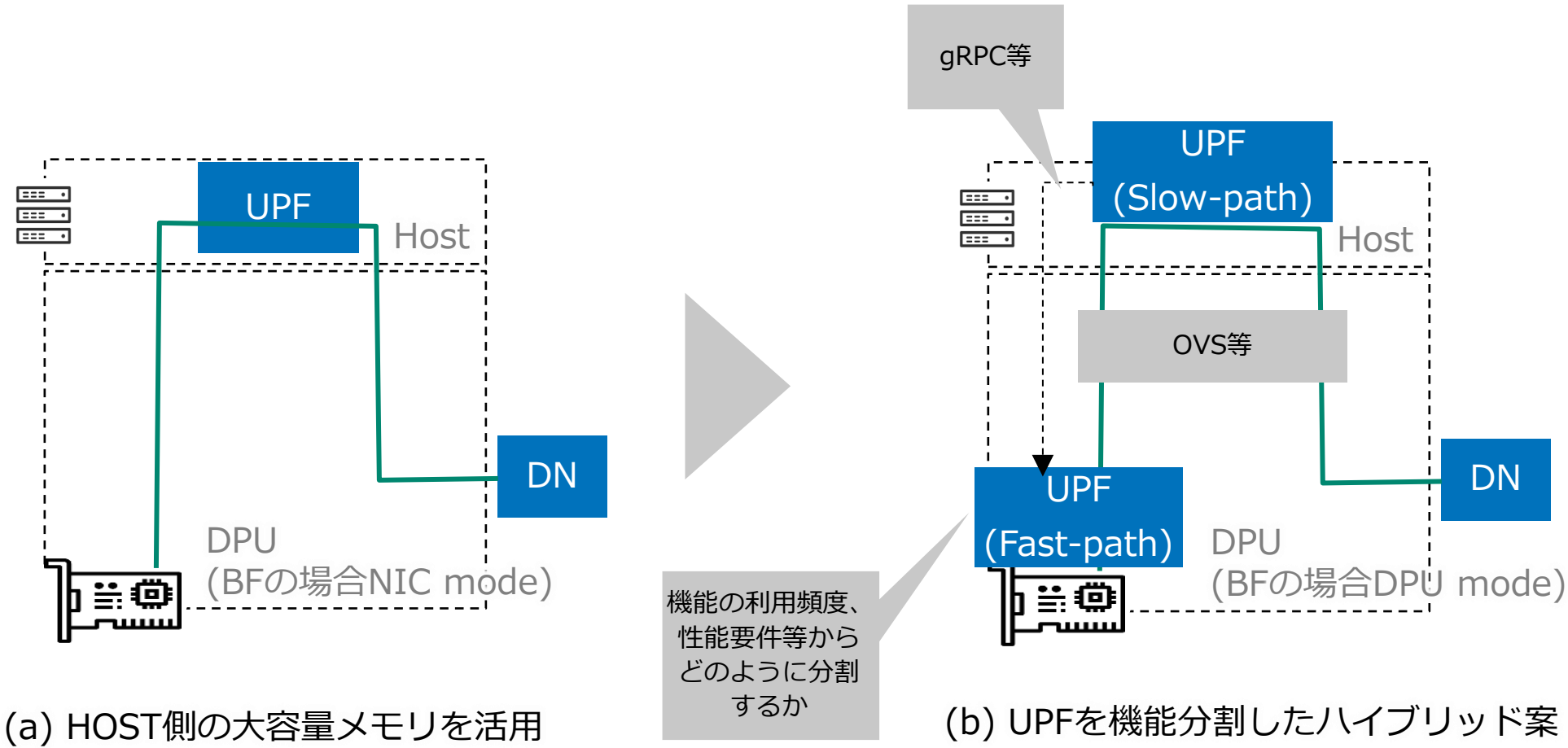
※DPUカード(BF3)1枚あたりの性能値

現在のHWオフロードしている
総フローエントリ数：約3M
(Uplink/Downlink合計)
DPUメモリ使用量：約13GB



商用要件を考慮したフロー構成では、
メモリ32GB(BF3の場合)を上回る見込み
→ DPUカード1枚では1Mセッション
乗せきれない可能性あり

検討例：DPU-HOSTの処理分割・連携



項目		結果
最大セッション数		1,000,000 セッション
PFCPTランザクション数		最大 1,000 TPS (500 session activation/秒)
23,000基地局収容 1M PDUセッション	Downlink only (パケットサイズ: 1200B)	11.7 Mpps 112 Gbps
	Uplink only (パケットサイズ: 120B)	? Mpps ? Gbps
	Bi-directional (パケットサイズ: 650B)	? Mpps ? Gbps

※DPUカード(BF3)1枚あたりの性能値

現在のHWオフロードしている
フローパイプライン段数：15段程度
(Uplink/Downlink合計)



商用要件を考慮したフロー構成では、
より多段構成となる見込み
→ Uプレーン性能が低下する可能性あり

参考：多段構成によるUP性能低下の例

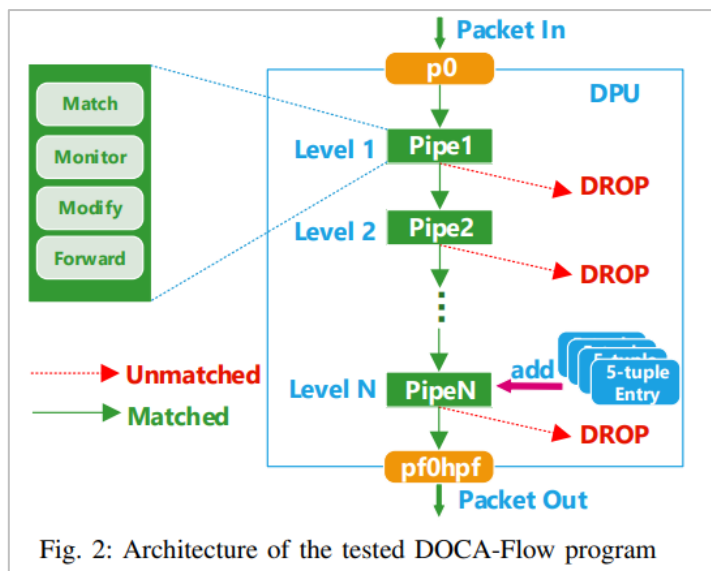
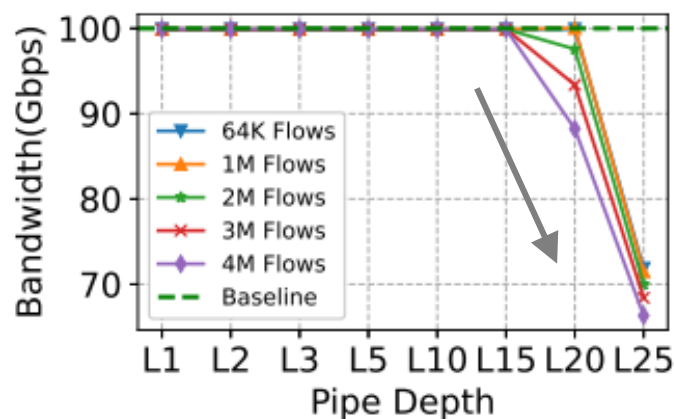
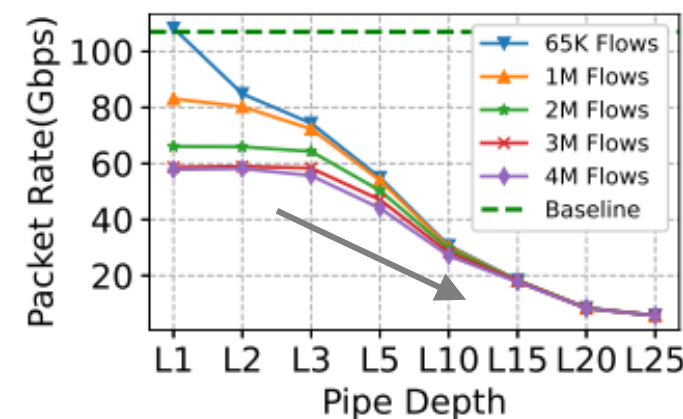


Fig. 2: Architecture of the tested DOCA-Flow program



(b) Bandwidth(Gbps)



(c) Packet Rate(Mpps)

Fig. 3: Bandwidth, packet rate and latency of DOCA-Flow under the different numbers of pre-installed entries and different pipe depths when using BlueField-2 DPU, (a) installed different number entries in the pipe at the first level while (b) and (c) installed 4M entries under different pipe depths.

Fuliang Li et al, "Performance Characteristics and Guidelines of Offloading Middleboxes onto BlueField-2 DPU,"

IEEE Transactions on Computers, 2024.

(<https://ieeexplore.ieee.org/document/10756527>)

主要課題リスト(DPU活用時のUPF実装課題の例)



■ 商用要件、3GPP標準仕様等を網羅的に調査し課題リスト化

		課題事項	対応・ワークアラウンド例
①	PFCPルール対応 (PDR/FAR/BAR/QER/URR)	ルール数が多い場合に、セッション確立速度とUプレーン性能に影響を与える恐れ	ルール数を抑えた運用対応
②	ハンドオーバー・着信時の パケットバッファリング	DPU単体ではメモリが不足する恐れ	収容ユーザ数を抑える
③	DPU-HOSTの処理分割・連携	DPUとHOST間の通信を担う手段・通信制御の確立	収容数を抑えて、DPU-HOST間の 連携を極力させない設計とする
④	フロールール制御APIの 処理性能不足	商用要件(ex. 10K PDUセッション確立/秒)を想定の場合、 DPUへのフロールール投入の処理速度が不足する恐れ	収容ユーザ数を抑える
⑤	3GPP標準機能追従 (PFCPルール対応以外)	N4-GTP等の他経路制御、N9接続(UL-CL)、バッファリング処理、 Status Report発行、帯域制限・帯域保証 等	—
⑥	冗長化構成	HOST-DPU物理構成、デプロイ構成 LAG等のインタフェース冗長構成	—
⑦	テレメトリ取得	PDUセッション毎にDPU上でパケットカウント可能な機能が必要 あったとしても商用要件を想定して処理性能が十分とは限らない	—
⑧	IPフラグメンテーション・ リアセンブル	DPUへのHWオフロード前提でパケットの分割・結合が行えるとは限らない	ソフトウェア処理で実現 但しDPUコア(ARM)がネックとなる恐れ
⑨	ルーティング	MACアドレス解決機構が標準で各DPUのSDKに備わっているとは限らない	ソフトウェア処理で独自実装 または子ルータ等に対応
⑩	ミラーリング	DPUへのHWオフロード前提でパケットのミラーリングが行えるとは限らない	ソフトウェア処理で独自実装 または外部スイッチ等に対応
⑪	L7カウント・インスペクション	DPUへのHWオフロードができない場合、DPUコア(ARM)/HOSTコア活用が発生 スローパスとなりUプレーン性能への影響の懸念	L7カウントを使わないサービスに限定 するなどの運用対応

他DPU (Marvell)での取り組み状況



- Marvell様/Hawkeye様にUPF実装に必須の機能要件への対応のご支援いただきながら進行中
- 既存実装との性能比較(w/ x86アーキテクチャ, NVIDIA BF3)の結果を含め、次回JANOGでのお披露目に乞うご期待！！（再びShowNetにもコントリブするかも？）



項目	仕様
製品名	HSC-5360
ベンダ	HAWKEYE TECH, CO., LTD.
プロセッサ	1x Marvell CN10624 ARM Neoverse N2 cores (24 Cores, 2.5 GHz)
システムメモリ	48GB ECC DDR5 (3x DDR Channels)
チップセット	1x MCU (NXP MIMXRT1176)
ネットワーク	2x 100GbE QSFP56 (w/ Break out Cable)
ストレージ	1x M.2 M Key for SSD (PCIe 5.0 x2; FF: 2230)
電力供給	Goldfinger (Max.: 75w) as primary ATX connector for additional power input (75W)

前回資料(<https://www.janog.gr.jp/meeting/janog54/dpuupf/>)

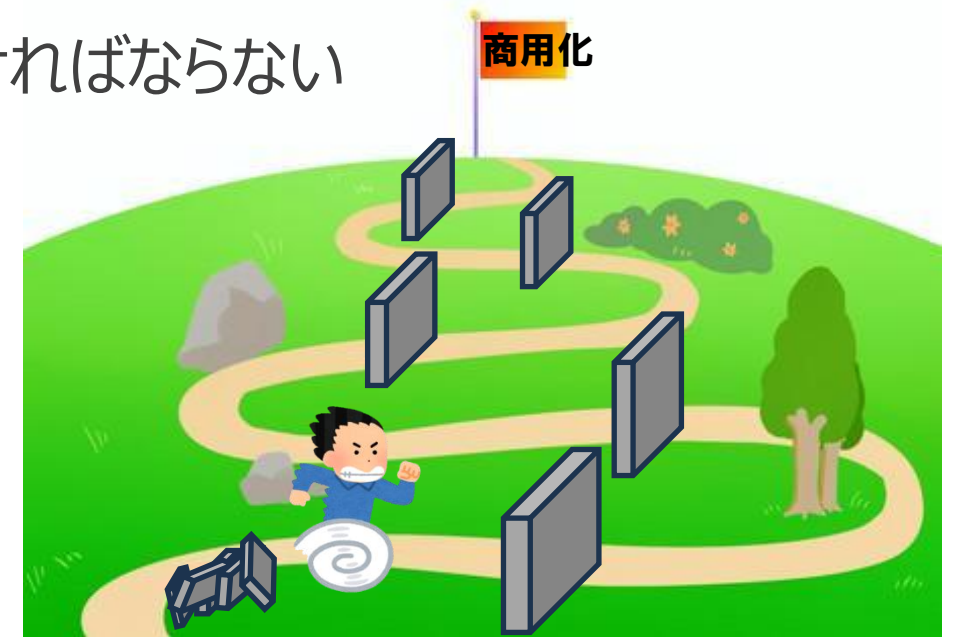
商用化に向けて

商用化への険しい道のり

PoCレベルとしては完成したが、商用化にはまだまだ超えるべき壁が山ほどある

- オペレーションどうするの？
- DPUは安定して調達できるの？
- 商用レベルの試験はしっかりできているの？
- 無中断アップデートできるの？
- 可用性は問題ない？

内製するからにはこれらを自分たちで解決していかなければならない



まだPoCの壁を越えたばかり

ハードウェアアクセラレータの扱いの難しさ

- DPUのようなアクセラレータを商用網に導入する際には考慮することがたくさんある
 - 価格面
 - 調達性
 - ベンダのハードウェアサポート
 - ライフサイクル（この手の製品はライフサイクルが短い…突然のディスコンなんてよくある…）
- どのベンダのDPUを使っていくか、などもしっかりと考えないといけない
 - ベンダによってハードウェアのスペック・制約が異なる
 - SDKもベンダによって違うため、開発のハードルに差がありそう

商用に求められる可用性の高さ

- 商用環境に求められる可用性は非常に高い
 - 特にUPFはユーザ通信を直接扱うノードであるため、障害による通信断を避けなければならない
- 内製するからには、当然LAGなどの冗長化の仕組みも自ら実装しないといけない



よかれと思ってNW機器の 4重化構成やってみた

NTTドコモ
石神美穂
2025/1/24

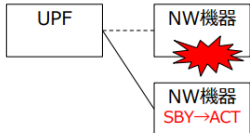
Copyright© 2025 NTT DOCOMO, INC. All rights reserved. NTT DOCOMO Confidential 1



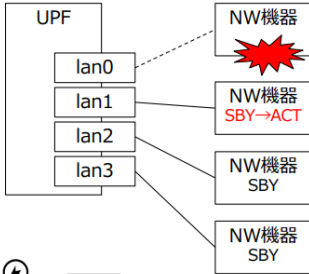
ドコモの計画保守と冗長構成の考え方

- UPFは数百万加入者収容のため、高い信頼性が必要
- ドコモ内部では、人材不足や効率UPのために即時駆けつけ不要な計画保守へ移行したいという要望が上がった
- 従来の2重化構成から4重化構成へとすることで、信頼性も確保しつつ即時の駆けつけ対応の必要性を無くし、計画的な保守を実現しよう

従来



4重化構成



24時間365日
故障時**即**駆けつけ対応 

  計画的に故障対応

UPF: User Plane Function
Uプレーン通信をつかさどる通信ノード

Copyright© 2025 NTT DOCOMO, INC. All rights reserved. NTT DOCOMO Confidential 3

午後のライトニングトークで冗長構成の共有があります（宣伝）

まとめ

- DPUを使ってUPFを実装してみた
 - 商用導入に向けた実装改善
 - 商用規模を想定した性能試験
- 内製したUPFを商用化するには、まだまだ検討事項が山積み
 - 商用ならではの機能や性能・品質の実現は想像以上にハードルが高い
 - 内製するからにはこうしたことも自分たちで検討しないといけない（覚悟が必要）
 - ベンダさんや他のオペレータさんの知見・ノウハウが不可欠な段階に来ているかも



自分たちだけで使うUPFを作りたいわけではないので、
ぜひ一緒にUPF作りませんか？

ディスカッション

Discussionポイント

トピック	説明
ネットワーク装置の内製化	そもそもネットワーク装置をキャリアが内製すべきなのか？ 餅は餅屋、やはりベンダに作ってもらうべきなのか？
実装上の技術課題解決に向けたアプローチ	ハードウェア制約によりセッション数の上限や処理性能の課題に直面している これらの解決に向けたアプローチとしてどういう策がありそうか？
内製装置の商用導入について	内製したコアNW装置の商用化の壁を乗り越えていくには？ (実例などがあれば教えていただきたいです)
UPFの内製を他社と一緒にやっていきたい話	UPFはモバイルNWの基本装置であるため、ベンダ・キャリアを跨いで開発するほうが良い？ 一緒に内製するメリット・デメリットなどを話し合いたい