

FURUKAWA ELECTRIC CO., LTD.

# JANOG 51

## 仮想ルータの開発話

2023/1/26

古河電気工業株式会社

古河ネットワークソリューション株式会社

宇高 正憲



- 宇高 正憲 (うだか まさのり)
- 古河ネットワークソリューション株式会社 IP開発部
- 2009年～古河ネットワークソリューションに入社  
主にHW,SWルータのデータプレーン開発に携わる  
ネットワークプロセッサのマイクロコード(アセンブラっぽい)から  
DPDKを利用したC言語でのデータプレーンの開発を実施中

※最近のjanogだとSRv6 MUPあたりのデモなんかのデータプレーンを作成

# FITELnetシリーズ ラインアップ

- 拠点からセンタまで、専用機と仮想化製品
- キャリア設備、大規模企業ネットワークで培った機能・性能と品質
- ルータ統合管理クラウドサービス「Fらくねっと」



ハードウェア製品

仮想化製品

製品名	F70 / F71	F220 / F221	F2200	F2500	F220EX / F221EX	FX1	FX201	FX2
VPN拠点数	64	128	2,000	3,000	128	4,000~20,000	4,000~20,000	
VPN中継性能	1Gbps	2Gbps	2Gbps	4Gbps	2Gbps	13Gbps	16Gbps	<b>200Gbps</b>
適用	拠点エッジ サービスアダプタ	拠点エッジ 小規模センタ	中規模センタ、大規模エッジ DC間通信設備		キャリアサービス設備 エッジ装置	大規模センタ、キャリアサービス設備 DC間通信設備、大規模企業ネットワーク		
	エンタープライズ向け				キャリア/DC設備向け			
ユーザ収容	拠点エッジ/センタ(NFV)、キャリア設備(スモールスタート用)、クラウドGW シングルテナント・サーバ共有型				キャリア/DC設備 VRFによるマルチテナント・サーバ占有型			
動作コア数	2~3コアを想定				4コア以上を想定			
スループット	10Mbps~5Gbps (ライセンス制)				vFX:1Gbps~100Gbps(ライセンス制)、vFX-R : VPNライセンス			
機能ライセンス	なし				vFX : Routing, VRF, L2VPN, IPsec vFX-R : AFオプション			
スケールOption	なし(IPsec:3000, VRF:32, FIB300,000 に固定)				IPsec(4K-20K), VRF(100-4K), FIB(64K-1M) vFX-RはIPsecなし			
製品名	<b>vFX-S</b>				<b>vFX/vFX-R</b>			



# ほぼルータとして販売したお話

## 背景

100G IFを持つルータの開発が必要とされていた。

(2016年くらい)

HWから開発すると工数も長くなるため、DPDKなどで100G IFも対応されてきており、仮想ルータでも実現できるかチャレンジしてみた。

## 仮想ネットワークアプライアンス「vFX」の特徴

ハードウェア製品と同じアーキテクチャ・機能・品質

高スケール・高性能 – 暗号中継100Gbpsを実現

### アプライアンス製品 (FX1)

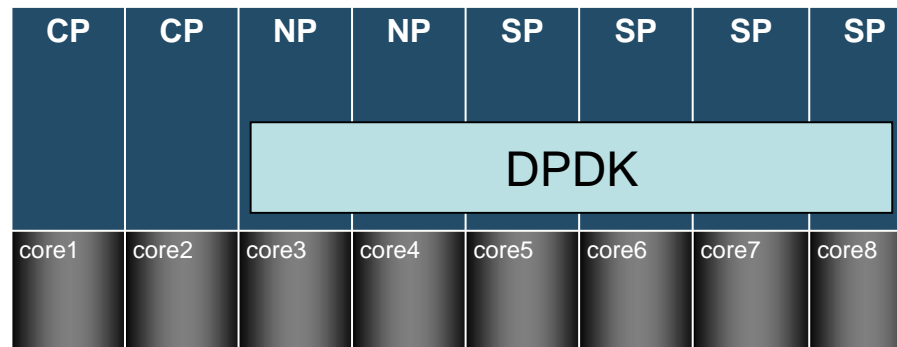
機能毎に専用プロセッサを持つ分散配置

Operation	MAP	offload
IKE	QoS	tunneling
routing	forwarding	fragment
CP	NP	SP



### 仮想化製品 (vFX)

実績のある3プロセッサアーキテクチャを継承



Hypervisor

Host OS



- 性能

IMIXで100Gbps

→コア数を増やしても性能が上がらない...

→トンネルでカプセル化すると性能劣化する...

- 安定性

長期間動作でパケロス無し

→showコマンド打つとパケロスする...

→未使用ifのshutdownでパケロスする...

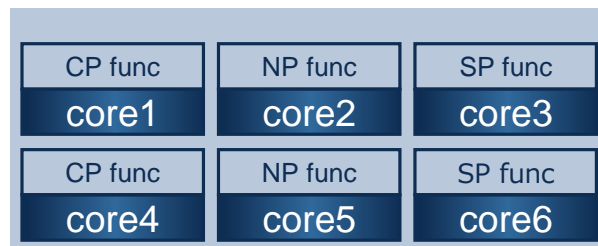
→ディスクへの書き込みでパケロスする...

→ごくまれにパケロスする...

- 売り方

- コア数を増やしても性能が上がらない...
  - lock lessな機構
  - 同じメモリ領域への書き込みをさせない
  - アーキテクチャをしっかりと考える  
(例えば物理ポート単位でコアを用意できるようにするとか)
- トンネルでカプセル化すると性能劣化する...
  - キャッシュやメモリのアライメントを意識する

## ルータソフトウェアの 開発範囲



仮想化基盤 (KVM, VMware)  
Linux

多くのボトルネックが存在する





## ● 仮想ルータの性能・品質に関して、様々な施策を実施

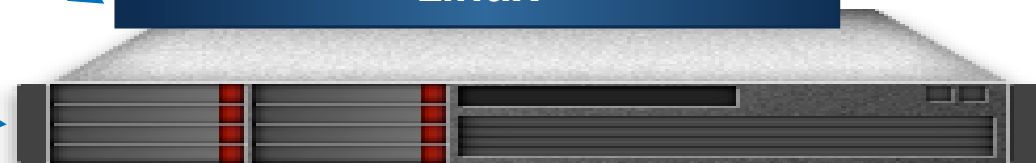
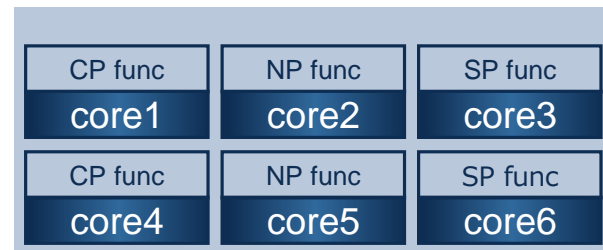
- CPUの設定変更

- 仮想化基盤の最適化

- Linuxの最適化

- PCIeの性能の最適化

- NICの設定変更



- CPUの最適化

  - BIOSの設定変更

  - perfなどでパフォーマンスカウンタを眺めながら仮想ルータのワークロードに適した設定をしていく

- Linuxの最適化

  - kernel thread, アプリ等のisolate

  - だいたいkernelの起動パラメータを追加していくと問題なくなる

- NICの最適化

  - DPDKやメーカーのサイトをしっかり読んで設定する

  - 送受信ディスクリプタ数などを変更

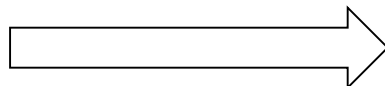
  - 実際に性能を測りながらよさそうな値を探す

- 既に仮想化基盤を運用中のお客様
  - ソフトウェアの購入で、すぐにvFXを使用可能
- 仮想化基盤を持たないお客様、ハードウェアルータと同様に使いたいお客様
  - 仮想化基盤の導入ってどうやればいいのか？
  - 仮想ルータって、どうやって動かせばいいのか？
  - サーバーは何を買えばいいのか？
  - サーバーの設定って何をすればいいのか？
  - OSは何を使えばいいのか？

ルータ機能・高速化のノウハウはあるが、  
汎用サーバ・仮想化基盤の保守・サポート体制を持っていない



キャリア向けルータ・  
性能向上に関するノウハウ



高品質なソフトウェアサポートを提供



お客様



Slerさん

ハードウェア・仮想化基盤運用  
に関するノウハウ

仮想化基盤



	開発目線	お客様目線
メリット	CPUが新しくなっても使用可能 その辺のPCで開発・検証可能 検証環境の保存が楽 開発期間が短い	使い方が変わらない 最新HW導入でパフォーマンスがUP
デメリット	PCIパススルー利用時はNICが変わると修正が必要 新しいサーバで検証が必要	HWルータに比べてEOSまでの期間が短い 古いホストOSがサポート対象外

- 100G~200Gくらいまでなら物理ルータと同じように使用可能  
→VM管理ツールと組み合わせれば使い勝手UP(ここが難しい気もしますが)
- vFXが現在利用されているサービス
  - 法人VPNサービス向け IPsec GW  
マルチテナント・シングルテナント
  - インターネット接続サービスのGW(法人/コンシューマ)
  - RR

仮想ルータを活用していきましょう