

# ネットワーク機器とFPGA

名古屋大学 情報基盤センター  
情報基盤ネットワーク研究部門  
基盤ネットワーク研究グループ

嶋田 創

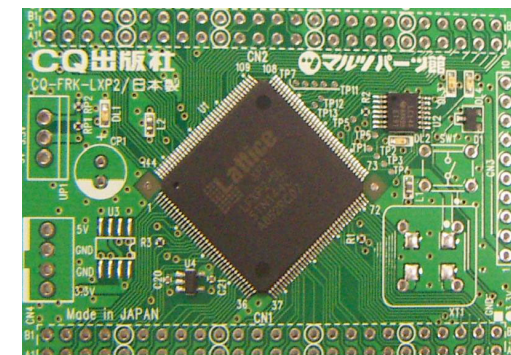
# ネットワークのハードウェア周りを実装するには?

1

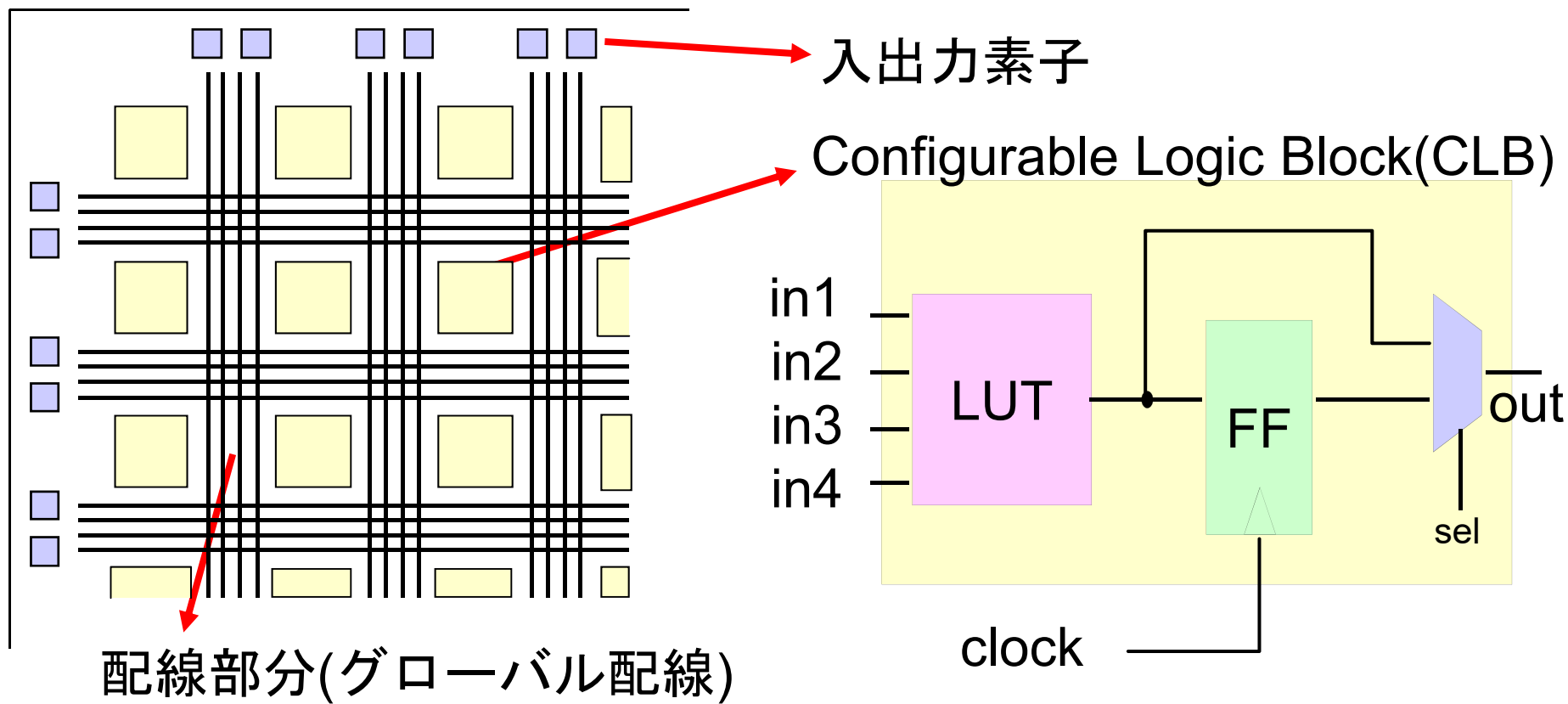
- 今までネットワークに関連するL1,L2,(L3)の世界とハードウェアの関係を見てきた
- 中身のよくわからない部分としてASICで構成されている部分がある
  - 高速化の要となっているようだが中身は細かく分からない
    - 他の企業に真似されると嫌なので、特に最近では公開されない
- ASICの部分は自分で細かく見たりすることはできない?  
→FPGAで実装することで確認できるかもしれない

# FPGA: Field Programmable Gate Array

- 近年多用される再構成可能ハードウェア
- LUTを使った構成が主流
  - LUT(Look-Up Table): 任意の3-8入力の信号に対して任意の値を出力する論理素子
- プロトタイピングで多用される
  - もしくは少量生産
    - ネットワーク機器ではよくある
  - もしくはASICが来るまでのつなぎ

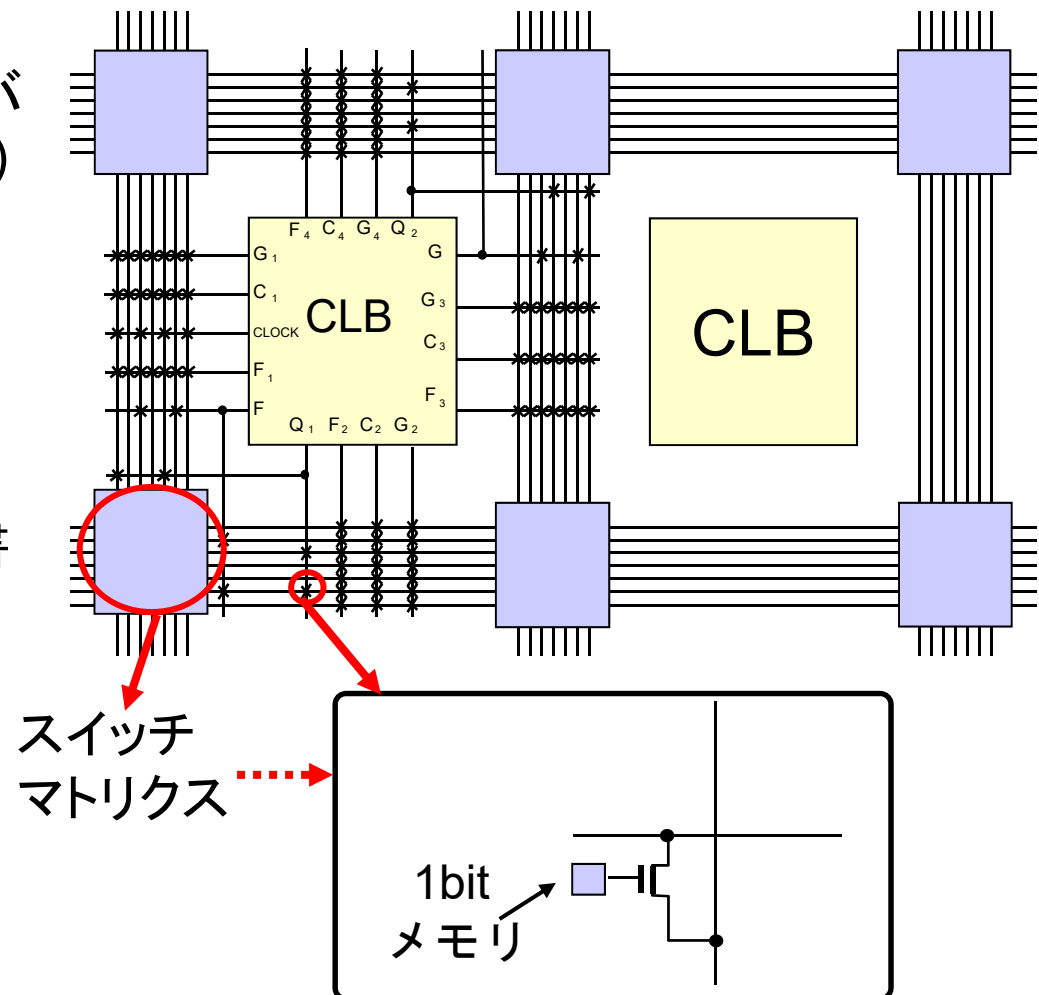


# LUTを使うFPGAの概観



# グローバル配線の構成

- 接続部分は2箇所
  - グローバル配線とグローバル配線(スイッチマトリクス)
  - グローバル配線とCLB
- 配線の接続はパストランジスタで制御される
  - パストランジスタに接続されたメモリに接続情報を書き込む



# LUT(Look-Up Table): 任意の論理値を出力できる論理素子

- RAMベースのLUTを考えると考えやすい
  - e.g. 4bit入力アドレスに対して1bitを出力するRAM
- LUTはマルチプレクサやROMなどでも実現される

RAMの値

ABCD	Q
0000	0
1000	1
0100	1
1100	0
0010	1
1010	1
0110	0
1110	1
0001	1
1001	0
0101	1
1101	1
0011	0
1011	1
0111	1
1111	0

入力  
(=アドレス)

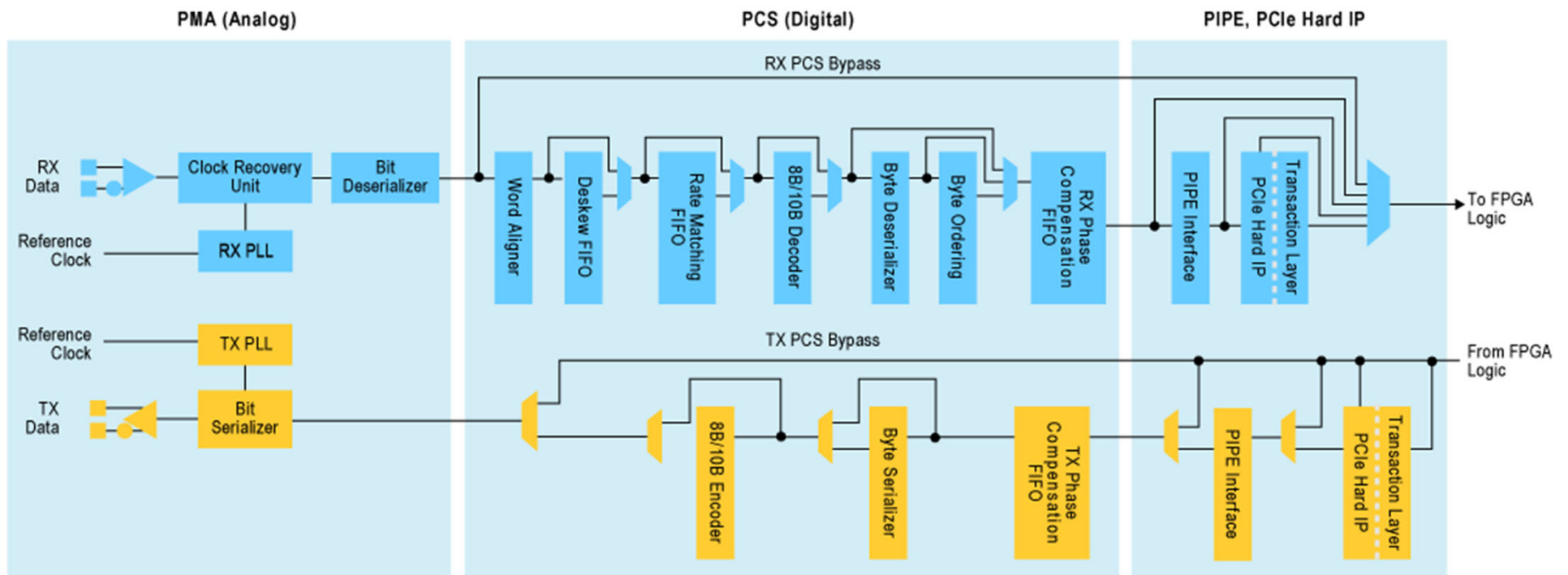


# 最近のFPGAはCLB以外もいろいろ搭載している

- ブロックSRAM
  - 容量重視、速度重視などバリエーションあり
- 全加算器(高速キャリー線付き)
- 乗算器
- 組み込みプロセッサ(2-3GHzのARMコアなど)
- DSPコア(可変精度)
  - 最近だとBF16などAI向け特殊精度にも対応
- 高速I/O
  - おおむね6Gbps以上
- メモリコントローラ(DDR3/4/5-SDRAM, HBMなど)
- PCI-Express IPコア
- 暗号化/復号処理コア(AES, SHAなど)

# Alteraの高速I/Oの物理構造

- PMS/PCSなどはUTP利用イーサネットと似た構成
- PMAは光ファイバ利用イーサネットと似た構成



# FPGAメーカー

- AlteraとXilinxが業界大手
  - 10G以上を実用的に使おうとすると実質この2社
    - 高速IO付きFPGAでないと利用ピン数が多くなりすぎる
- Altera(2015/6にIntelに買収された...が2025/4に売却)
  - 高速IO付きFPGAのバリエーションが多い
  - Intelの高性能プロセス(7/10/14nm)を利用したAgilexシリーズ
  - Intel XeonとのMulti Chip Module版も発表された(2016/4)
- Xilinx(2022/2にAMDに買収された)
  - 無料で使えるIPコアはこちらの方が多かった印象
- その他: 1000BASEあたりまでは対応できる
  - Actel: アンチヒューズ型で著名(高速だが書き換え回数1回をラインアップ)
  - Quicklogic, Lattice: 小規模FPGAで強い

# 高速IOを持つFPGA(Altera)

買収後に開発が滞っていた感じが、ようやくローエンドまで揃ってきた所でまた売却...

- Agilex 7-F (10nm): 58Gbps(PAM4) x48(最大)
- Agilex 7-I(10nm): 116Gbps(PAM4) x72(最大), 58Gbps (PAM4) x120(最大), XeonへのCompute Express Link
- Agilex 7-M(7nm): 116Gbps(PAM4) x4(最大), 58Gbps(PAM4) x12(最大), CXL, HBM2
- Agilex 5-E(10nm): 28Gbps x24(最大)
- Agilex 5-D(10nm): 28Gbps x32(最大)
- (Agilex I): 32Gbps x72(最大), 58Gbps(PAM4) x56(最大), 116Gbps(PAM4) x8(最大)
- Agilex 9なる無線向けADC/DAC付きFPGAも

# 高速IOを持つFPGA(Altera)

(Altera時代の旧モデル)

- Stratix

- Stratix V GT(28nm): 28.05Gbps x4, 12.5Gbps x32(最大)
  - Stratix V GX(28nm): 14.1Gbps x66(最大)
- Stratix 10 TX(14nm): 56Gbps(PAM4) x60(最大)
  - 普通のトランシーバとして使う場合は30Gbps x120(最大)

- Arria

- Arria V GZ(28nm): 12.5Gbps x36(最大)
- Arria 10 GT(20nm): 17.78Gbps x96(最大)
  - Arria 10 GT(20nm): 25.8Gbps x6(最大)

- Cyclone

- Cyclone V GT(28nm): 6.144Gbps x12(最大)
- Cyclone 10 GX(20nm): 12.5Gbps x12(最大)

# 高速IOを持つFPGA(Xilinx(AMD))

まだ買収に伴う製品開発停滞下にあるように見える

- Virtex

- Virtex UltraScale(20nm): 30.5Gbps x60(最大)
- Virtex UltraScale+(16nm): 32.75Gbps x128(最大)、58Gbps(PAM4) x48(最大)

- Kintex

- Kintex UltraScale(20nm): 16.3Gbps x64(最大)
- Kintex UltraScale+(16nm): 32.75Gbps x32(最大)

- Artix

- Artix Ultracale+(16nm): 16.3Gbps x12(最大)
- Artix-7(28nm): 6.6Gbps x16(最大)

- Spartan

- Spartan Ultrascale+(16nm): 16.3Gbps x8(最大)

# 高速I/Oを使ったイーサネットのMAC層

- 通常、FPGAメーカーから汎用バスインタフェースを持つMAC層がIPコアとして提供されている
- 自社の各FPGAモデルのリソースを最大限に利用した実装のIPコアを提供している
  - 10Gbps超の高速I/Oを持たないFPGAでも100GBASEを実装したりとか
- IP Coreもアップグレードできる(される)
  - 例: Alteraは2013/11に10G/40G/100G Ethernet IP Coreを更新
    - 100Gは55%小型化、70%低レイテンシ
    - 40Gは40%小型化、60%低レイテンシ
    - 10Gは20%小型化、24%低レイテンシ
  - OpenCores[1]とかでも有志による新しいオープンな実装が出ることはある

[1] <https://opencores.org/>

# Ethernet IP Core



- Altera(Intel)[1]
  - Altera時代からあったTriple Speed Ethernet (10/100/1000BASE-T)から400GBASEまで
  - 25GBASEとか50GBASEとかも
- Xilinx(AMD)[2]
  - ([2]は検索ベースなので一覧性が悪い)
  - Altera同様、1000BASEや10/25GBASE[2]から400GBASE[4]まで

[1] <https://www.intel.com/content/www/us/en/products/details/fpga/intellectual-property/interface-protocols/ethernet.html>

[2] <https://www.xilinx.com/products/intellectual-property/nav-interface-interconnect/nav-wired.html>

[3] <https://www.xilinx.com/products/intellectual-property/em-di-400gemac.html>

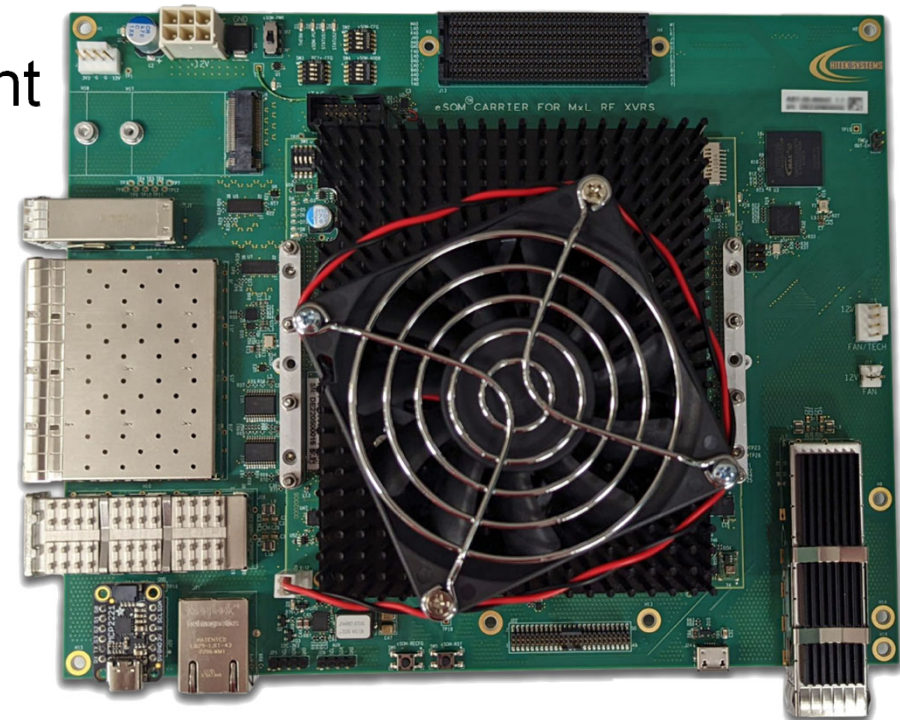
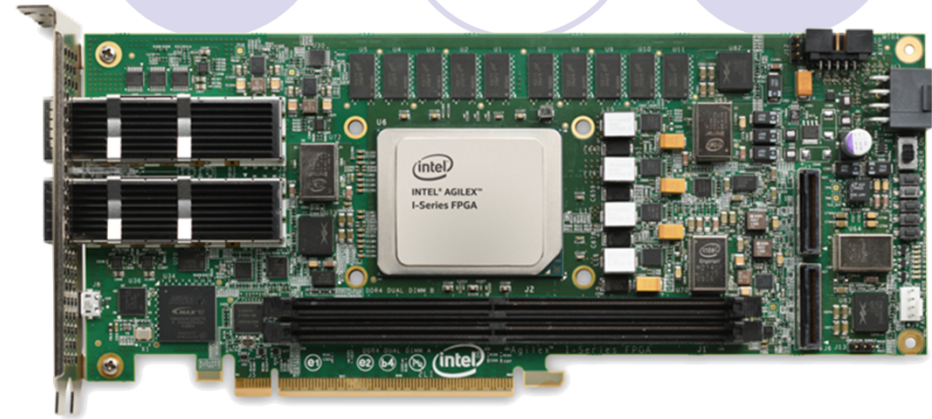
[4] <https://www.xilinx.com/products/intellectual-property/ef-di-25gemac.html>

# FPGAの高速I/Oを利用して実装できる 他の高速通信規格

- Fibre Channel
  - Storage Area Networkで多用される
  - 1.0625/2.125/4.25/8.5Gbps x n
  - 10.3125/14.025/28.05/28.9/57.8Gbps x n
    - 東ねて「32/64/128/256Gbps級」の規格での利用が一般的
- OTN(Optical Transport Network)
  - 大規模バックボーンネットワークにて利用される
  - 10.4/26.4/40.3/52.8/104Gbps x n
    - こちらも東ねたものに規格がついている
- Interlaken: 3.125-6.375Gbps x n
  - 他の規格に押されて後継規格が出ていない

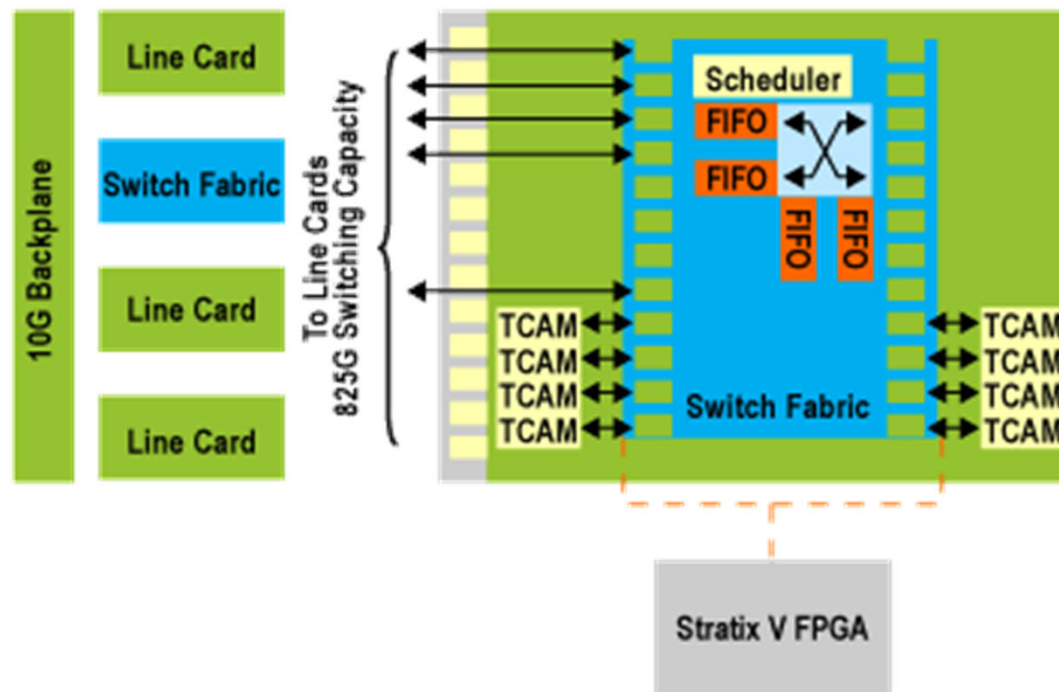
# 高速通信が使えるFPGAボードの例

- Agilex 7 FPGA I series Development Kit(上写真)
  - QSFP-DD x2
  - MCIOx8コネクタ x2
  - PCIe Gen5 x16
- Agilex eSOM7 Development Platforms(下写真)
  - QSFP-DD x1
  - QSFP28 x1
  - SFP28 x4
  - M.2 x4 PCIe
  - JESD-204C/B x8



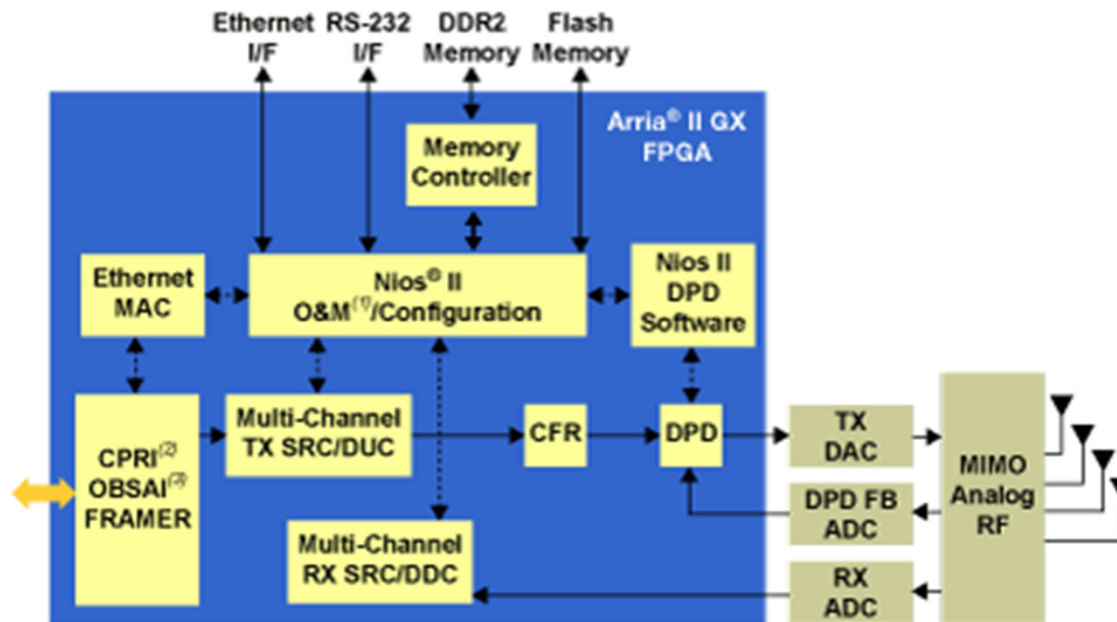
# Alteraによるクロスバススイッチの実装例

- Stratix V GXを利用
- 14.1Gbpsのトランシーバ x66間の通信のスイッチング
- ルーティングのためのTCAMを併用



# FPGAによる無線ネットワークのデジタル信号処理

- Arria II GXによる無線のデジタル信号処理部実装
  - DSPブロックなどを活用
- ソフトウェア無線の実装などで活用できそう

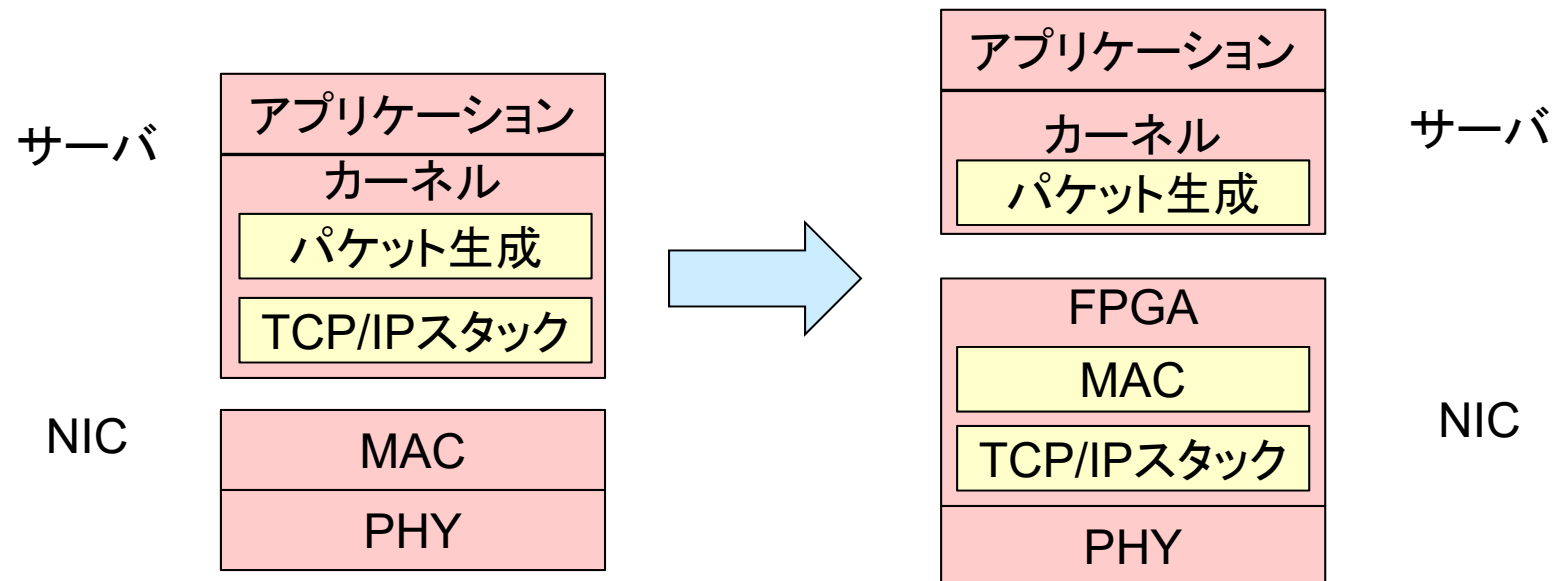


# 高頻度トレードに見る高速送受信処理部におけるFPGA利用

- HFT: High Frequency Trading
  - アルゴリズムによる(株式)取引方法の1つ
  - 取引時のマージンを低くするが、高頻度で取引をすることで
  - ミリ秒単位の高速(株式)取引が重要になる
    - “2005円で売り”と”2010円で買い”が出そうならば、“2006円で買って2009円で売る”という
    - 最近だとマイクロ秒とかのオーダーに...
- このような取引では取引依頼の少しの遅延が大きな損失に  
→FPGAによる取引依頼部ハードウェア化
  - アルゴリズムの部分は引き続きサーバ部分

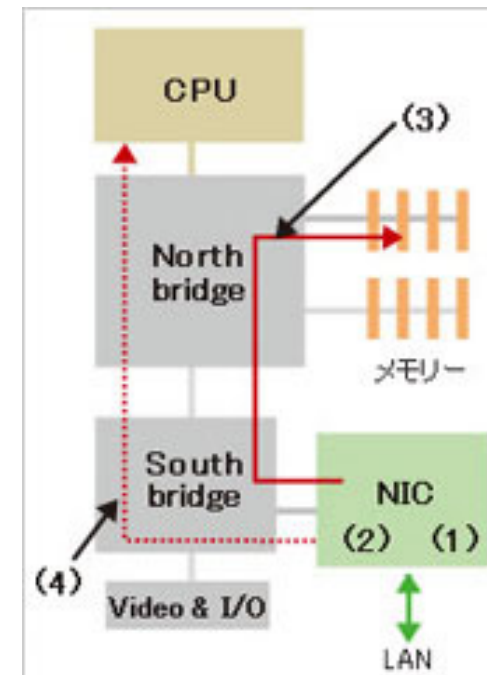
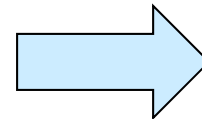
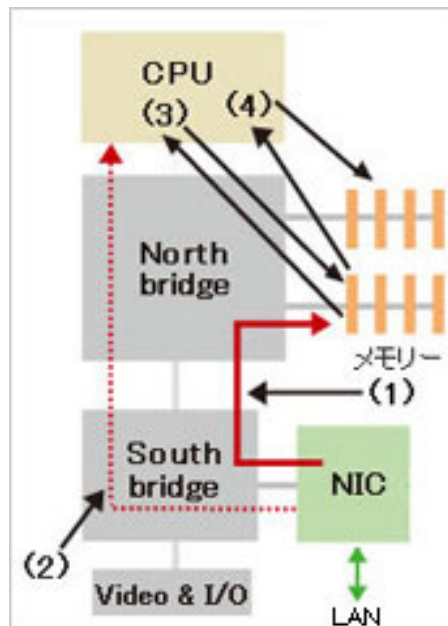
# HFTのネットワークにおけるFPGA利用 (1/3)

- 初期: FPGA付きNICによるTCPオフローディング
  - TCPオフローディング: TCP/IPスタックをFPGA側で実行することでサーバ側の負荷を軽減
  - サーバで生成した取引発注の通信内容をFPGA側のTCP/IPスタックにて送信
    - FIXプロトコル: 金融取引の標準プロトコル



# TCPオフローディング

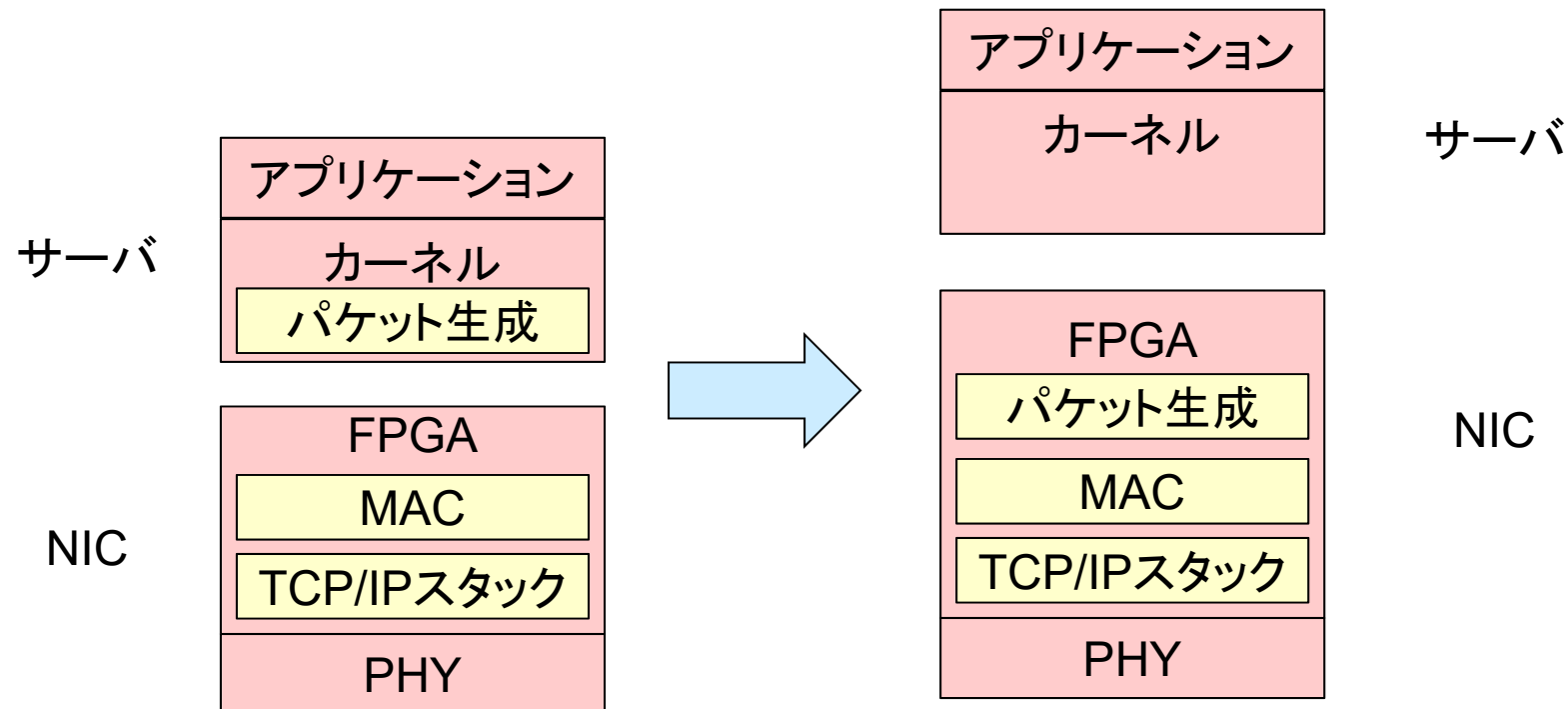
- 従来
  - パケットデータのメモリへの読み書きにCPUが介在
- TCPオフローディング
  - パケットデータはメモリに書き込まれてから受信通知が来る
  - メモリ上のパケットデータに対して送信依頼ができる



# HFTのネットワークにおけるFPGA利用 (2/3)

- 中期:

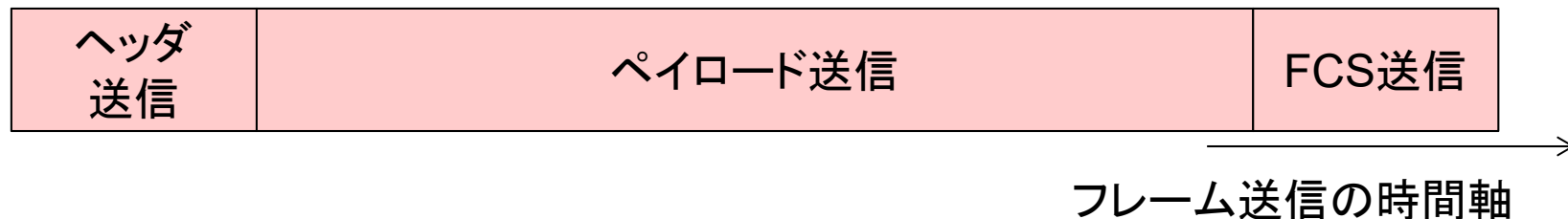
- 発注の通信をFPGA内部で生成
- サーバ側は取引発注内容自体のリクエスト処理のみ



# HFTのネットワークにおけるFPGA利用 (3/3)

- 最終形: 投機的な取引リクエスト
  - 過去の値動きを元に発注すべき取引内容を予測
  - 最新の値動き結果が来る前に取引内容(のイーサネットフレーム)を送信開始
  - 予定通りの値動き: そのまま送信
  - 予定とは異なる値動き: イーサネットフレームの送信をキャンセル
    - フレーム最後のFCSに誤った値を付与
    - 非常に迷惑な行為なので、当然、証券会社側の確認は取っているはず

→あまりにもえげつないのでHFTは規制される傾向



# 高速トレードにおけるFPGA(小ネタ)+FPGA関連小ネタ

- J.P.MorganがポートフォリオのリスクシミュレーションにFPGAアクセラレータ利用(2011)
  - x86サーバ数千台並列で8-12時間
    - アクセラレータ付属サーバ40台で4分(120倍の高速化!)
      - 途中でGPUで14-15倍の高速化も行った
- AristaがFPGA内蔵ネットワークスイッチを出してたことも
- Alteraを買収したIntelがFPGAを搭載したSmartNICを売り出したことも
  - XilinxもAlveo U50とか出してきた
- 最近だと3Dスタック実装を利用したFPGAも出てきている
  - LUTを積んだチップをTSV(Through Silicon Via)経由でスタックして回路面積/遅延を削減
  - HBM(High Bandwidth Memory)を積んでメモリボトルネックを緩和

# FPGAとネットワークの研究(1/2)

- NICTのNTPサーバのFPGA化[1]
  - サーバ間の時間がずれるとファイル管理やログが大変なことになるので、クラウドサービス事業でもNTPは重要
    - さくらインターネットなどもNTPサーバのFPGA化やっている[2]
      - 大本の時刻供給源をGNSS衛星にしている
- FPGA NICを使った高速Key Value Storeサーバ[3]
- FPGA NICへのChangeFinder実装[4]

[1] <https://www.nict.go.jp/publication/NICT-News/0610/research/02.html>

[2] <https://elab.sakura.ad.jp/ntp-trial/>

[3] 徳差ら, "多様なデータ構造を有するKey-Value Storeアプリケーションの設計," 情報処理学会論文誌, 2016.

[4] Iwata et al. "An FPGA-Based Change-Point Detection for 10Gbps Packet Stream," 電子情報通信学会論文誌, 2019.

# FPGAとネットワークの研究(2/2)

- 嶋田研話

- パケットのペイロードの1-gram特徴量の抽出[1]
- TCAMを用いたTCPセッション再構成とペイロード特徴量抽出[2]
- オンチップでのTCPセッション再構成とNFAによる悪性通信検知[3]
- FPGAによるペイロードの周波数変換処理とサーバ側の機械学習による悪性通信検知[4]

[1] Yanase et al., "Network Access Control by FPGA-Based Network Switch using HW/SW Cooperated IDS," IEICE Tech. Rep., 2014.

[2] Fuchigami et al, "FPGA Base TCP Session Features Extraction Utilizing Off-Chip Memories," INTERNET 2015, 2015.

[3] Hu et al., "Realtime Malicious Traffic Detection targeted for TCP Out-of-Order Packets based on FPGA," IEEE Access, 2023.

[4] Hu et al., "Enhancing Detection of Malicious Traffic through FPGA-based Frequency Transformation and Machine Learning," IEEE Access, 2024.