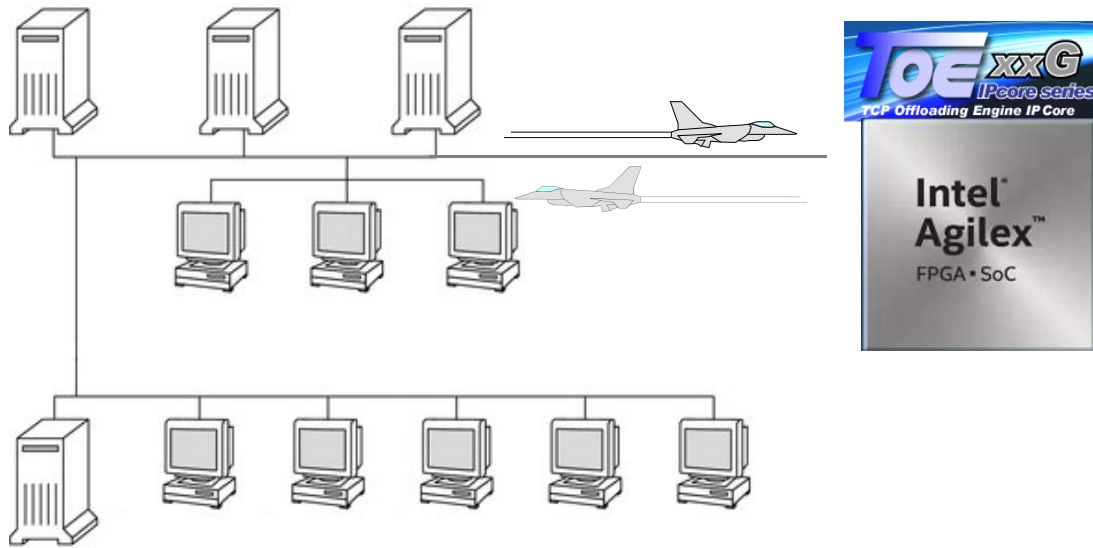


Intel版TOE_{xxxG}-IPコアのご紹介

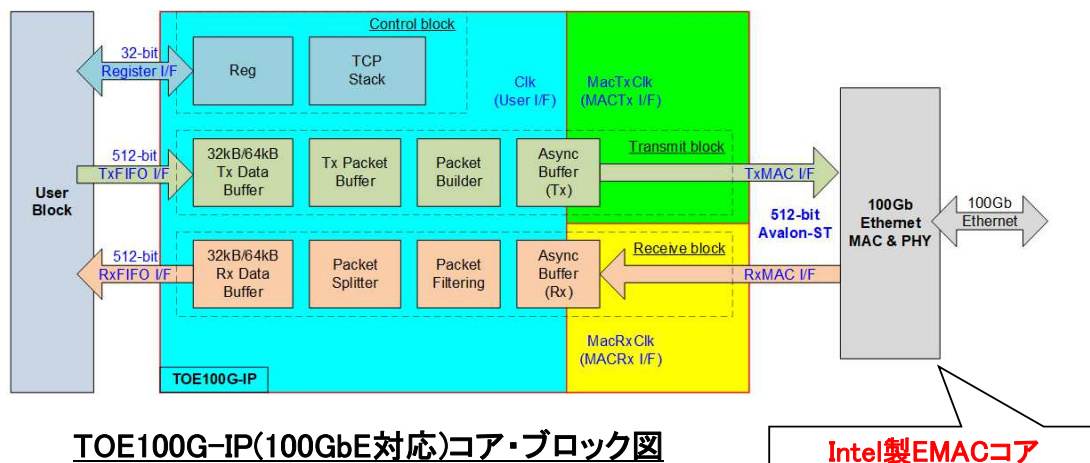
Ver2.0AJ



純ロジックのIPコアで超高速TCP実装

TOE_{xxxG}-IPコアの概要

- ・ 100G/40G/25G/10G/1GbE対応のTCP/IPオフロード・エンジン・コア
- ・ ユーザ回路とIntel製EMACコアの間に挿入
- ・ TCPの送信・受信の両方向の処理を完全HW・全自動化
- ・ Full Duplex(送受信同時)通信をサポート



TOE100G-IP(100GbE対応)コア・ブロック図

Intel製EMACコア

TOE_{xxxG}-IPコアのラインナップ

ファミリ ラインレート	GbE	10GbE	25GbE	40GbE	100GbE
Cyclone V	対応済				
Arria V	対応済				
Cyclone 10 GX	対応済	対応済			
Arria 10	対応済	対応済		対応済	
Stratix 10	対応可	対応済	対応済	対応可	対応済
Agilex F		対応可	対応可	対応可	対応済

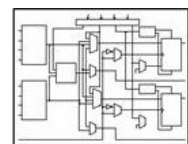
TOE_{xxxG}-IPコア・ラインナップ (2021/8/1時点)

対応済: 即時納品可

対応可: 注文可能

TOE_{xxxG}-IPコアの特長1

- ・ TCP/IP処理を完全ロジック・ハードウェア化
 - CPU負荷がゼロ、理論上最小の処理レイテンシ
- ・ 送信のみ/受信のみ/同時送受信を高速転送
 - ラインレートの90%程度の実転送速度
 - 同時送受信(全二重通信)でも80%程度を達成
- ・ 転送データの信頼性を確保
 - 送信時: ACK抜け/重複/タイムアウト等で自動リトライ
 - 受信時: シーケンス番号を評価し返信ACKを自動制御
 - フロー制御: TCPウィンドウ・アップデート・パケット自動送受信



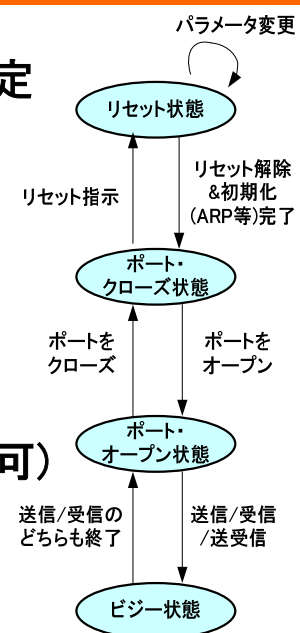
TOE_{xxxG}-IPコアの特長2

- ・ データバッファ容量を選択可能
 - FPGAメモリ・リソースとパフォーマンスからユーザが選択
- ・ Intel製EMACコア対応
 - 弊社製低価格EMAC-IPコアとの接続も可能
- ・ 実機動作リファレンス・デザインを用意
 - Intel評価ボードで動作するプロジェクト
 - 購入前にパフォーマンス・信頼性の実機評価を検証可能
 - 製品のリファレンスはコア以外の全回路をソースで添付



TOE_{xxxG}-IPコアの動作概要

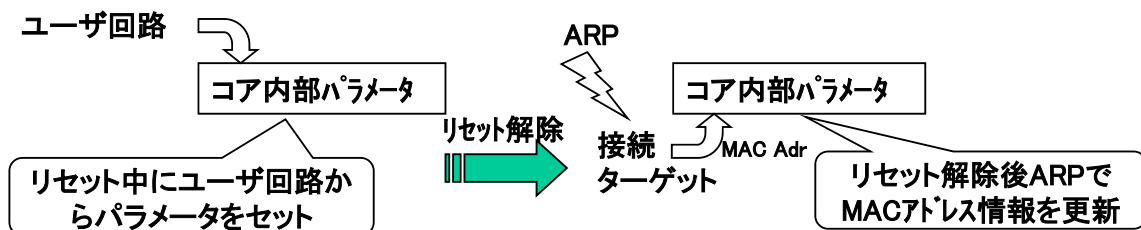
- ・ リセット状態でパラメータ(IP&MACアドレス等)を設定
- ・ リセット解除で初期化(ARP等)を実行
- ・ 初期化完了後ポートをオープン(アクティブ/パッシブ)
- ・ 送信・受信は独立して動作(同時送受信可)
- ・ パラメータ変更はリセット状態で実施
(転送長/パケット長はビジー以外の状態で変更可)



コアの状態遷移図

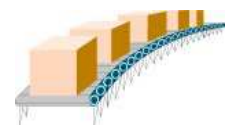
初期化動作

- ・ パラメータの初期値設定
 - コアのリセット維持中にユーザ回路より設定
 - IPおよびMACアドレス・ポート番号を指定
 - 設定を完了するとリセットを解除
- ・ リセット解除後ARP実行(相手MAC Adr手動設定も可能)
 - クライアント・モード: 接続ターゲットに対してARPを発行
 - サーバー・モード: 接続ターゲットからのARPを待機



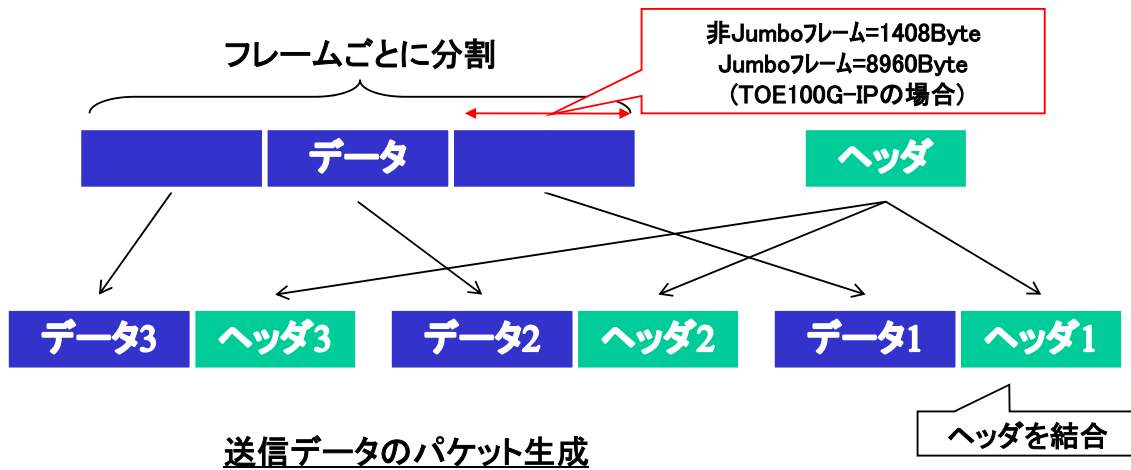
高速送信

- ・ 送信パケットの生成
 - ユーザ回路は送信データをFIFO I/Fで書込み
 - 送信データをフレームサイズで分割
 - ヘッダと送信データを結合しEMACへ出力
- ・ 自動再送機能
 - ターゲットからのACKを常時チェック
 - ACK抜け/重複/タイムアウト等の異常ACKを検出
 - 異常ACKの種類に応じて適切な再送データを用意



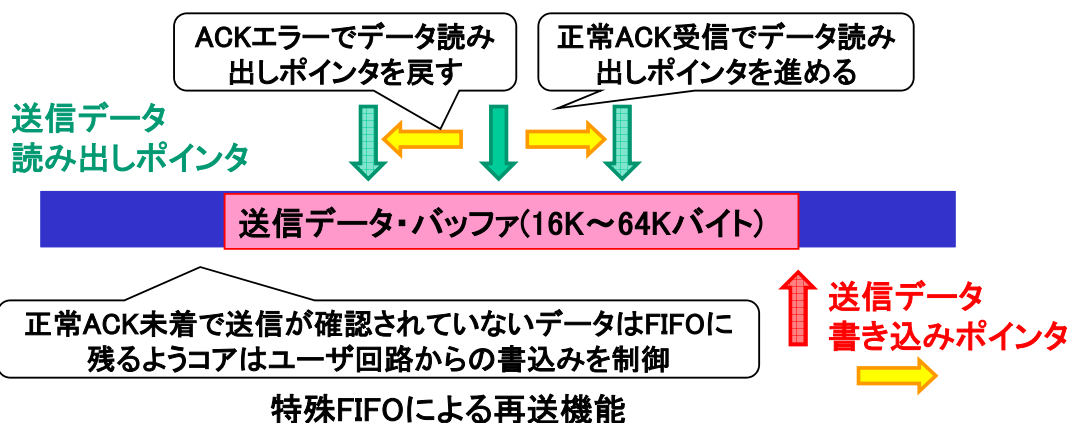
送信パケットの生成

- 送信データにヘッダを自動的に付加しEMACへ転送
 - データはコア内にてフレーム・サイズで分割
 - チェックサムやシーケンス番号などもコアが自動生成






自動再送機能

- 専用設計された特殊なFIFOで再送を実装
 - 正常ACKで読み出しポインタを進める
 - 異常ACKの場合コアは適切な位置にポインタを戻す
 - ポインタ制御・再送の実行をコアが自動制御



高速受信

- ・ 受信パケットのヘッダ・チェック 
 - 受信対象でない場合やチェックサム・エラーの場合は破棄
- ・ データ並び替え 
 - シーケンス番号の入れ替え発生時に並び替え
 - 受信済みデータの再送要求を防ぎ転送効率を維持
 - 並び替え処理できない場合は重複ACKで再送要求
- ・ 重複データの結合 
 - 受信データが前回受信パケットと重複する部分を検出
 - 重複箇所のみ破棄し連続データとして復元

受信パケットのヘッダ・チェック

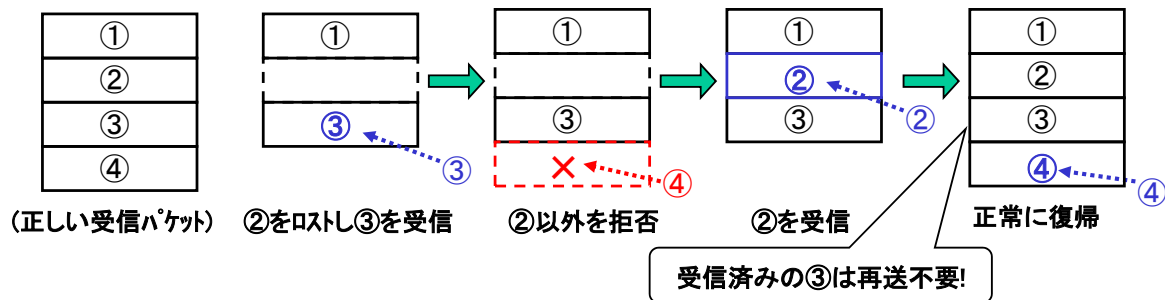
- ・ 受信ヘッダのチェック・サムが正しいことを確認
 - 更に以下の条件を満たすこともコアが確認

Byteオフセット	プロトコル	データ内容	確認条件
0-5	ICMP	あて先MACアドレス	SML/SMHLレジスタで設定したコアのMACアドレスと一致
6-11	ICMP	送信元MACアドレス	ARPで検出した通信ターゲットのMACアドレスと一致
12-13	ICMP	タイプ	0x0800 (IPパケット)であること
14	IP	バージョン/ヘッダ長	0x45 (IPv4, IPヘッダ長=20)であること
20	IP	フラグ/フラグメントOFS	下位6ビットがゼロ(フラグメントの有無チェック)であること
23	IP	プロトコル番号	0x06(TCPパケット)であること
26-29	IP	送信元IPアドレス	DIPLレジスタで設定した通信ターゲットのIPアドレスと一致
30-33	IP	あて先IPアドレス	SIPLレジスタで設定したコアのIPアドレスと一致
34-35	TCP	送信元ポート番号	DPNLレジスタで設定/パッシブオープンで取得した通信ターゲットのポート番号と一致
36-37	TCP	あて先ポート番号	SPNLレジスタで設定したコアのポート番号と一致
38-41	TCP	シーケンス番号	前回処理時のシーケンス番号との差分がコアで処理可能範囲内

受信パケット・ヘッダの確認条件

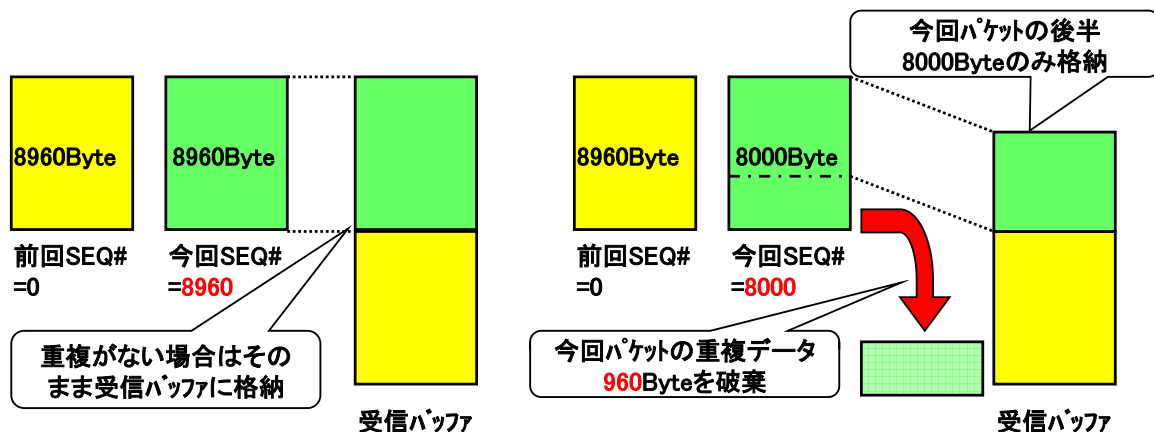
データ並び替え

- シーケンス番号がスキップした(ロスト)場合に機能
 - ロスト状態を解消するパケット以外は受信しない
 - スキップ部のデータを回復するロスト解消パケットのみ受信
- データ並び替え
 - ロスト解消パケットからデータの連続性を完全に復元
 - 受信済みデータは再送要求せずパフォーマンスを維持



重複データの結合

- データの重複を検出し自動的に補正
 - シーケンス番号から受信データの重複を検出
 - 重複箇所のみ破棄し連続データとして復元



フロー制御 (Window Update自動送信機能)

- ・ TCP Window Update (ACK)パケット生成
 - 受信データ・バッファの空きスペース復活を検出
 - 設定した閾値でWindow Updateパケットを自動送信
 - 相手送信側ではWindowサイズ復活で送信が再開可能

IP末尾42=FPGA IP末尾25=PC

Source	Destination	Protocol	Length	Info
192.168.11.42	192.168.11.25	TCP	60	4000-50223 [ACK] Seq=1 Ack=61321 win=4213 Len=0
192.168.11.42	192.168.11.25	TCP	60	4000-50223 [ACK] Seq=1 Ack=62781 win=2753 Len=0
192.168.11.42	192.168.11.25	TCP	60	4000-50223 [ACK] Seq=1 Ack=64241 win=1298 Len=0
192.168.11.42	192.168.11.25	TCP	60	[TCP window Update] 4000-50223 [ACK] Seq=1 Ack=64241 win=3352 Len=0
192.168.11.42	192.168.11.25	TCP	60	[TCP window Update] 4000-50223 [ACK] Seq=1 Ack=64241 win=5406 Len=0
192.168.11.42	192.168.11.25	TCP	60	[TCP window Update] 4000-50223 [ACK] Seq=1 Ack=64241 win=7460 Len=0
192.168.11.42	192.168.11.25	TCP	60	[TCP window Update] 4000-50223 [ACK] Seq=1 Ack=64241 win=9514 Len=0
192.168.11.25	192.168.11.42	TCP	1514	50223-4000 [PSH, ACK] Seq=64241 Ack=1 win=256960 Len=1460
192.168.11.25	192.168.11.42	TCP	1514	50223-4000 [ACK] Seq=65701 Ack=1 win=256960 Len=1460
192.168.11.25	192.168.11.42	TCP	1514	50223-4000 [ACK] Seq=67161 Ack=1 win=256960 Len=1460

(PC->FPGAの受信に対する通常ACK)

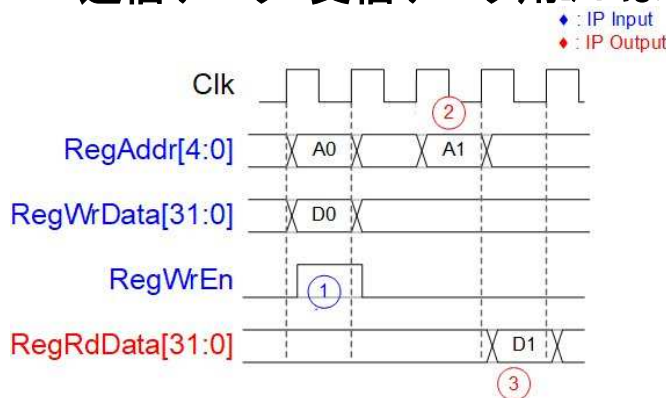
Windowサイズ復活後PCからのデータ送信を再開

Window Updateパケットの自動送信
(TOE1G-IPコアの実動作例)

Windowサイズ復活でコアからWindow Updateパケットを自動発行

ユーザ・インターフェース(制御)

- ・ レジスタI/F、送信FIFO I/F、受信FIFO I/Fの3種類
 - レジスタI/Fは初期パラメータの設定、方向切り替え指示
 - 送信データ・受信データ用I/Fは標準的なFIFO I/F

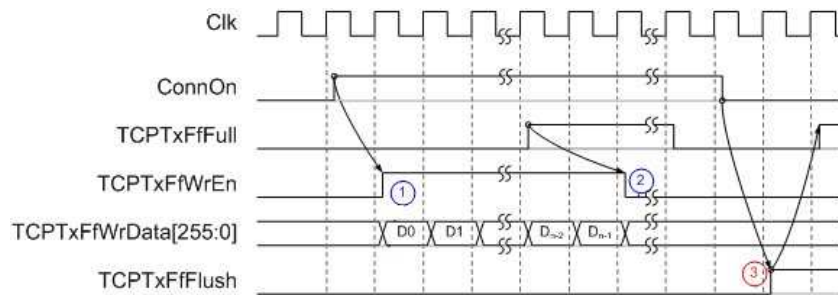


[レジスタの書き込み]
①アドレスとデータを設定しWrEnで書き込み

[レジスタの読出し]
②アドレスを与えたる
③次クロックで有効データが出力

レジスタI/Fのタイムチャート

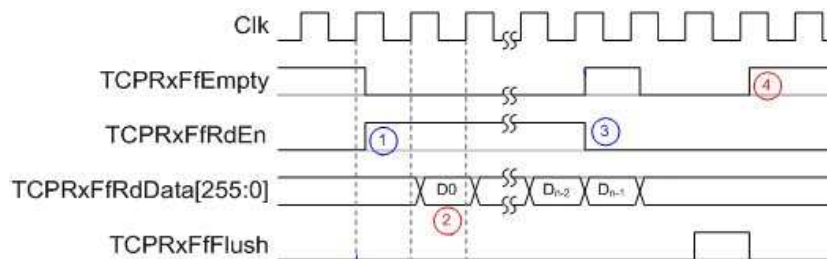
ユーザ・インターフェース(データ)



送信FIFO I/Fのタイムチャート

[送信データの書込み]

- ①データをWrEnで書込み
- ②Fullになってから4クロック以内にライト中断
- ③FlushでFIFOクリア



受信FIFO I/Fのタイムチャート

[受信データの読み出し]

- ①非EmptyでRdEnにて読出し
- ②次のクロックでデータ出力
- ③Emptyではリード禁止
- ④FlushでFIFOクリア

データ・バッファ容量の設定

- ・ 送受信のデータ・バッファをパラメタライズで設定可能
 - ① 送信データ・バッファ: 16K/32K/64KByteから選択
 - ② 受信データ・バッファ: 16K/32K/64KByteから選択
- ・ リソースとパフォーマンスの最適点を調整できる

ジェネリック名	設定範囲	説明
TxBufBitWidth	9-11	送信データ・バッファ・サイズをアドレス・ビット幅で設定します。例えば9の場合16Kバイト、11の場合64Kバイトとなります。
RxBufBitWidth	9-11	受信データ・バッファ・サイズをアドレス・ビット幅で設定します。例えば9の場合16Kバイト、11の場合64Kバイトとなります。

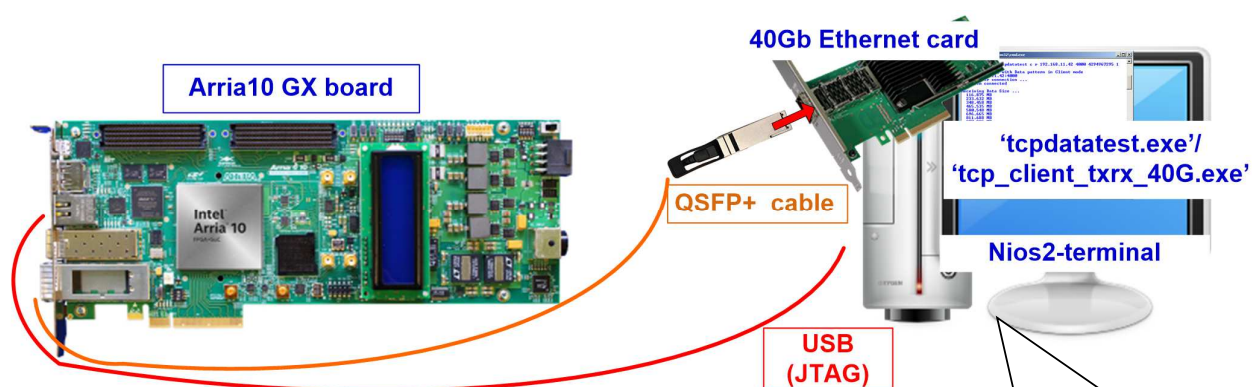
各データ・バッファはパラメタライズで設定できる

評価用SOFファイル

- ・ 2種類の評価用SOFファイル
 - PC～FPGA間または2枚のFPGAボード間で通信
 - ただしPC評価の場合PC側がボトルネックになる
(汎用PCで高ラインレートのTCPデータ実時間処理は困難)
- ・ 転送パフォーマンス測定/データ・ベリファイ
 - 半二重通信および全二重通信をサポート
 - ベリファイによりデータ信頼性の確認が可能
- ・ SOFファイルでコア購入前に実機評価可能

評価環境A: PC～FPGA間通信

- ・ PC側は評価用ソフトウェアを使用し評価
 - 注意点: 1GbE版以外ではPC側がボトルネックで最高パフォーマンスは出せない

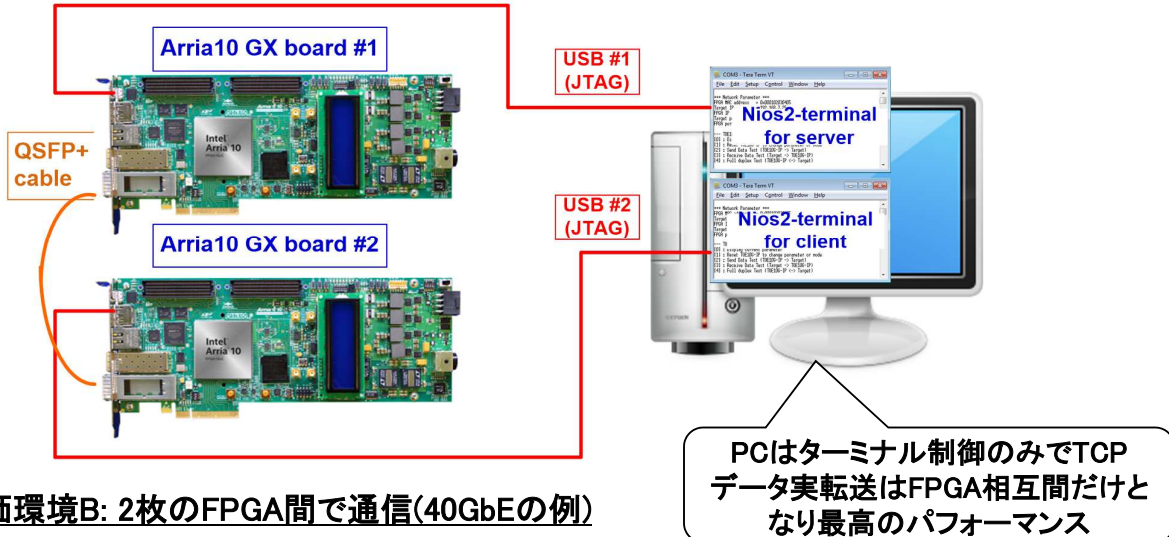


評価環境A: PC～FPGA間通信(40GbEの例)

注: 実際のデータ転送ではPC側のTCP処理がボトルネックとなり転送速度は制限されます

評価環境B: 2枚のFPGAボード間で通信

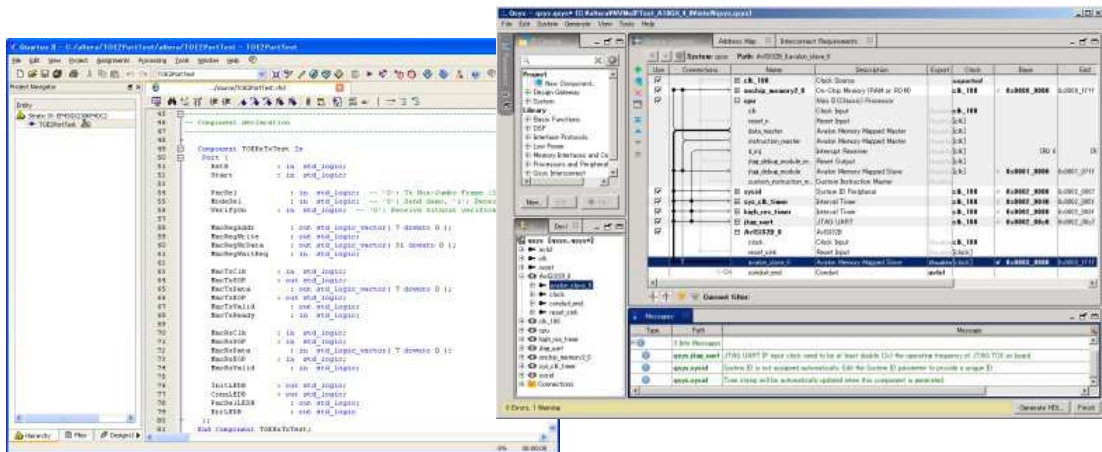
- ・ FPGAボード2枚間でTCP相互通信
 - 外部PCのNios2-Terminalで両ボードを制御



評価環境B: 2枚のFPGA間で通信(40GbEの例)

リファレンス・デザイン

- ・ 実機動作するQuartusデザイン・プロジェクト
 - コア部以外の全回路をソースコードで製品に標準添付



評価ボードで実動作するQuartus/Qsysプロジェクト

リファレンスと実機評価による開発

- ・ リファレンス+評価ボードによる確実な開発
 - まず最初に製品添付のリファレンスで実機動作を確認
 - そこからユーザ製品に向け少しずつ編集
 - 編集ごとに実機動作をStep by Stepで確認
 - 問題があれば1ステップ前に戻るだけで動く状態にすぐ復帰できる



大きな後戻りがなく確実に短期間での製品開発が可能!

消費リソース

- ・ TOE_{xx}G-IPコア単体消費リソース
 - 送信/受信ともデータ・バッファを最大の64Kバイトに設定した場合のリソース



ラインレート (ファミリ)	クロック周波数	消費ロジック	最大消費メモリ
GbE (CycloneV E)	125MHz	2047 ALMs	1181696 bit
10GbE (Arria10 GX)	156.25MHz	2453 ALMs	1179648 bit
25GbE (Stratix10 GX)	350MHz	3273 ALMs	1179648 bit
40GbE (Arria10 GX)	322MHz	3656 ALMs	1179648 bit
100GbE (Agilex F)	350MHz	7713 ALMs	1837056 bit

TOE_{xx}G-IPコア単体コンパイル結果

送受信ともバッファを最大に設定した場合です。
バッファ容量を削減すればBlockメモリ消費リソースを節約できます。

パフォーマンス測定例(100GbE版)

```

+++ TOE100G-IP Send Mode +++
Enter transfer size (aligned to 512-bit): 64 - 0x3FFFFFFC0 => 0x3FFFFFFC0
Enter packet size (aligned to 512-bit) : 64 - 8960 => 8960
Input mode : [0] Client [1] Server => 1
Run test application on PC by following command
tcpdatatest c r 192.168.100.25 60001 274877906880 1
  
```

```

Wait Open connection ...
Connection opened
Start data sending
Send 12392 MByte Recv 0 Byte
Send 24784 MByte Recv 0 Byte
Send 37176 MByte Recv 0 Byte

Send 247839 MByte Recv 0 Byte
Send 260231 MByte Recv 0 Byte
Send 272623 MByte Recv 0 Byte
Send data complete

Close connection
Connection closed
  
```

```

Total tx transfer size = 4294967295 512-bit
Total = 274.877[GB] , Time = 22186[ms] , Transfer speed = 12389[MB/s]
  
```

256Gバイト・データを
ジャンボフレーム(8960バイト)で
他方のFPGAへ送信

転送パフォーマンス
12,389MByte/sec!

2枚のFPGAボード間での片方向(送信→受信)評価結果例 (100GbE)

パフォーマンス実測値(各ラインレート)

ラインレート	半二重通信	全二重通信	測定条件
1GbE	116MByte/s	103MByte/s	FPGA-PC間転送
10GbE	1,200MByte/s	924MByte/s	FPGA-PC間転送
25GbE	3,098MByte/s	3,096MByte/s	2枚のFPGAボード間
40GbE	4,907MByte/s	4,858MByte/s	2枚のFPGAボード間
100GbE	12,389MByte/s	11,177MByte/s	2枚のFPGAボード間

各ラインレート対応版のTOE_{xx}G-IPコア・パフォーマンス実測値

結論

- ・ TOE_{xxxG}-IPコアで接続ターゲットと高速データ転送
 - 保証されたTCPデータでラインレート上限に迫る転送速度
- ・ 再送を含めた複雑なTCP処理を完全にハード化
 - データ並び替えや重複データ削除もコアが自動的に処理
- ・ レジスタやFIFOにより接続容易なユーザI/F
 - CPUなしのシステム実装を可能とするシンプルなI/F
- ・ 実機動作を確認できるリファレンス・デザイン
 - デザインと動作確認の同時並行が可能

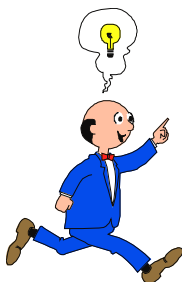


ネットワーク帯域を最大限に活用する
製品開発を強力に支援します!



問い合わせ

- ・ ホームページに詳細な技術資料を用意
 - https://dgway.com/TOE-IP_A.html
- ・ 問い合わせ
 - 株式会社Design Gateway
 - E-mail : info@dgway.com
 - FAX : 050-3588-7915



改版履歴

Rev.	日時	履歴
2.0AJ	8/2/2021	TOExxG-IPシリーズを統合した日本語プレゼン初版リリース