

raNVMe-IPのご紹介 (Intel版)

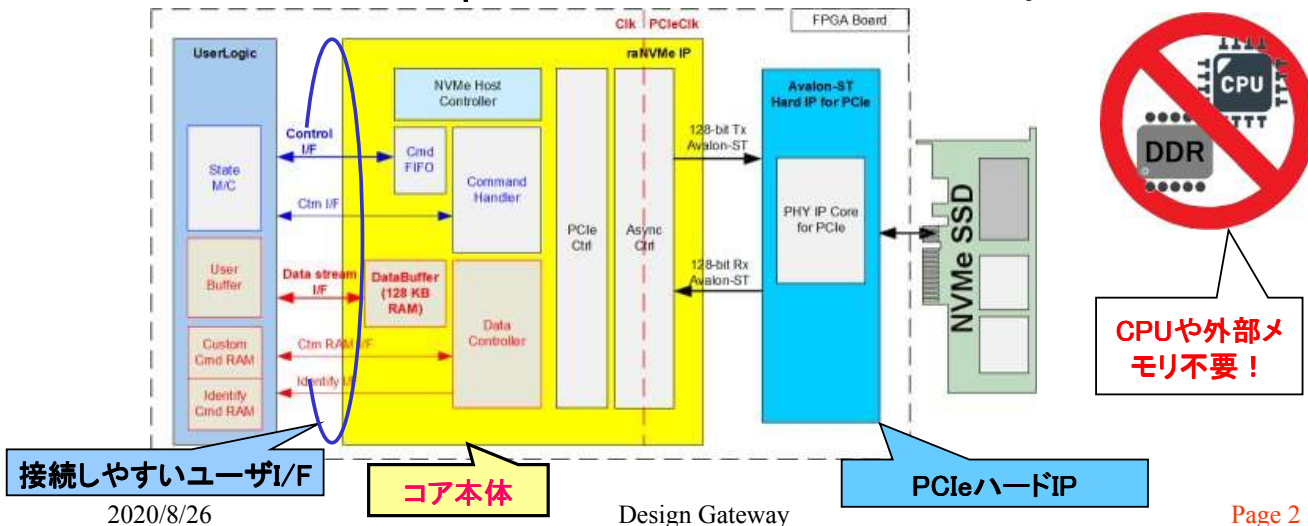
SSDへのランダム・アクセスを
CPUなしで実現



SSD高速データ・ベース検索システムに最適

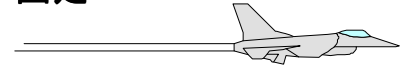
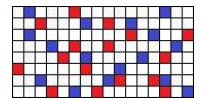
raNVMe-IPとは

- 概要: NVMe SSDとFPGAを直結しランダムRead/Writeを自動実行
- 特長: CPUや外部メモリ不要,純ハードロジック・コアによる超高速動作
- アプリ例1: SSDにデータベースを格納する検索システム
- アプリ例2: Start/Stopによる画像/センサデータ収録システム



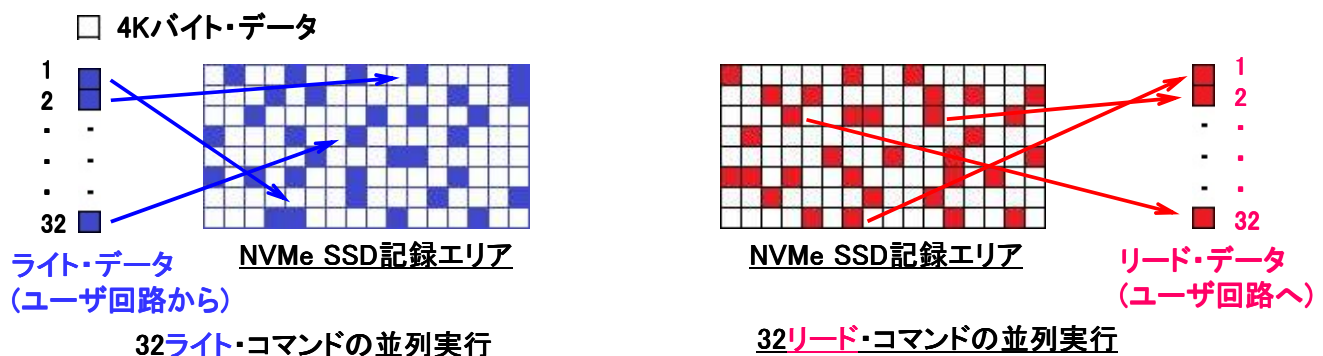
raNVMe-IPの特長

1. コア動作:ランダム・ライト/リードを自動実行
 - ライト又はリードいずれかで最大32コマンドを並列同時に実行可能
 - 1コマンド当りのライトまたはリード・データ量は4Kバイト固定
2. 高性能:ハイ・パフォーマンスかつコンパクト
 - Write=約590KIOPs、Read=約220KIOPs
 - コア単体サイズ=2016ALM, 必要メモリ=約136KBytes
3. インターフェイス:接続容易なユーザI/F
 - コマンド、データともFIFO接続(コマンドFIFOは最大32コマンド分)
 - CPUや外部DDR不要(データ・バッファに内部ブロック・メモリ使用)
4. 多機能:リード/ライトに加えて各カスタムコマンドをサポート
 - SMART/FLUSH/Shutdownなどカスタム・コマンドを実装
5. 環境: Intel評価ボードでの試作開発環境
 - 実機動作するリファレンス・デザインをコアに添付



特長1:コア動作

- ・ ライトコマンドまたはリードコマンドのどちらかを選択
- ・ アクセス(アドレス)先の異なる最大32コマンドを並列実行
- ・ 1コマンド当りのデータ量はライト・リードとも4Kバイト固定



raNVMe-IPコアによるコマンド並列実行イメージ

特長2: パフォーマンス(ライト)

```

+++ Write Command selected +++

Please input [Start Address] and [Length] in unit of 8
Select Transfer Mode      : [0] Sequential [1] Random => [0]
Enable Data Verification  : [0] Disable   [1] Enable  => [1]
Enter Start Address <512 Byte> : 0x0 - 0x3B9E12A8 => [0]
Enter Length <512 Byte>      : 0x8 - 0x3B9E12A8 => [0x40000000]
Selected Pattern [0]Inc32 [1]Dec32 [2]A11_0 [3]A11_1 [4]LFSR => [4]

In Progress Progress status
100%

Total = 34.359 [GB] , Time = 14663 [ms]
Transfer speed = 2343 [MB/s], 599K [IOPS]

+++ Write Command selected +++

Please input [Start Address] and [Length] in unit of 8
Select Transfer Mode      : [0] Sequential [1] Random => [1]
Enable Data Verification  : [0] Disable   [1] Enable  => [0]
Enter Start Address <512 Byte> : 0x0 - 0x3B9E12A8 => [0]
Enter Length <512 Byte>      : 0x8 - 0x3B9E12A8 => [0x40000000]

In Progress Progress status
100%

Total = 34.359 [GB] , Time = 14839 [ms]
Transfer speed = 2315 [MB/s], 592K [IOPS]

```

シーケンシャル・ライト
2343MB/s

ランダム・ライト
592,000IOPS

評価条件:
評価SSD: Samsung SSD 970PRO 512GB

raNVMe-IPコア実機ライト・パフォーマンス評価結果例

特長2: パフォーマンス(リード)

```

+++ Read Command selected +++

Please input [Start Address] and [Length] in unit of 8
Select Transfer Mode      : [0] Sequential [1] Random => [0]
Enable Data Verification  : [0] Disable   [1] Enable  => [1]
Enter Start Address <512 Byte> : 0x0 - 0x3B9E12A8 => [0]
Enter Length <512 Byte>      : 0x8 - 0x3B9E12A8 => [0x40000000]
Selected Pattern [0]Inc32 [1]Dec32 [2]A11_0 [3]A11_1 [4]LFSR => [4]

In Progress Progress status
100%

Total = 34.359 [GB] , Time = 14969 [ms]
Transfer speed = 2295 [MB/s], 587K [IOPS]

+++ Read Command selected +++

Please input [Start Address] and [Length] in unit of 8
Select Transfer Mode      : [0] Sequential [1] Random => [1]
Enable Data Verification  : [0] Disable   [1] Enable  => [0]
Enter Start Address <512 Byte> : 0x0 - 0x3B9E12A8 => [0]
Enter Length <512 Byte>      : 0x8 - 0x3B9E12A8 => [0x40000000]

In Progress Progress status
100%

Total = 34.359 [GB] , Time = 38893 [ms]
Transfer speed = 883 [MB/s], 226K [IOPS]

```

シーケンシャル・リード
2295MB/s

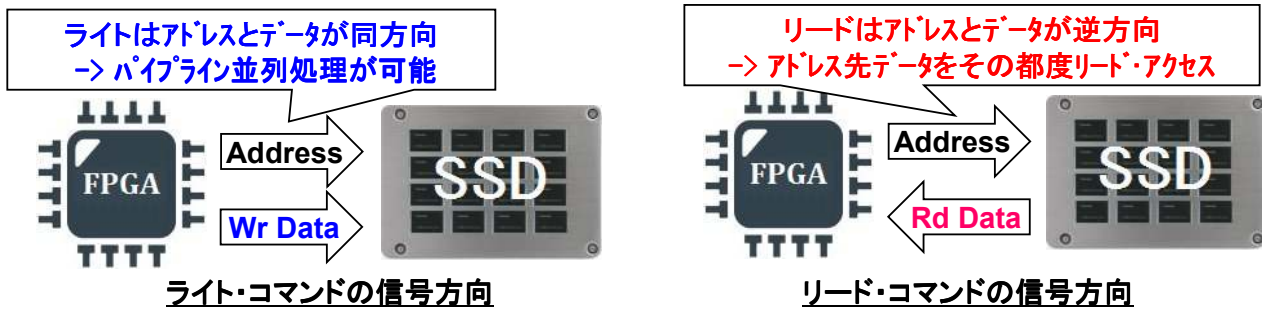
ランダム・リード
226,000IOPS

評価条件:
評価SSD: Samsung SSD 970PRO 512GB

raNVMe-IPコア実機リード・パフォーマンス評価結果例

特長2: パフォーマンス(SSD特性)説明

1. ライトはシーケンシャル・ランダムとも速い
 - ライトではSSD内キャッシュを使って見かけ上の書き込み速度向上が可能
 - データのNANDメモリ書き込みが終わらなくてもSSDは完了通知(ライトバック)
 - ライト先をまとめてNANDメモリのブロック消去回数を削減可能
2. リードはシーケンシャルは速くランダムは多少遅くなる
 - シーケンシャル・リードの場合SSD側でプリフェッチ(先読み)が可能
 - ランダム・リードではリード先がSSD側で予測できず逐次リード動作が必要
 - SSDではコマンドでアクセス先情報を受信してから対象メモリのリード開始



特長2: コンパクト

- ・ コア消費リソースを最小化 (2000ALM程度)
 - NVMe SSD制御に特化・最適化した制御ロジックを実装
- ・ ブロックRAMによるデータ・バッファ
 - 内部メモリにより最小のオーバーヘッドでランダム・アクセスを実現

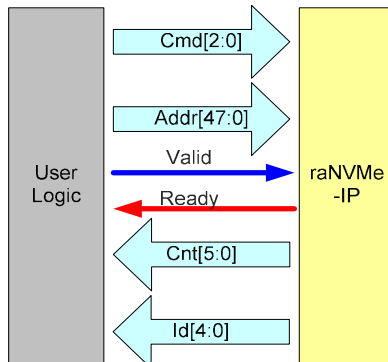
Family	Example Device	Fmax (MHz)	ALMs	Registers	Block Memory bit	Design Tools
Arria10 GX	10AX115S2F45I2SG	300	2016	3437	1,114,176	QuartusII 16.0

raNVMe-IPコア単体の消費リソース

A10AX1150全部のALM数
427,200個に対して僅か2016
個=0.47%のロジック消費量

特長3: ユーザI/F (コマンド)

- ・ 接続容易なインターフェイス
 - コマンド種類とアドレスをセット
 - ReadyをモニタしValidパルスでコマンド実行開始
- ・ コマンド進捗はCntとId値で確認
 - Cnt値で並列実行中のコマンド数(32以下)を確認
 - Id値で実行中コマンド用のデータを区別

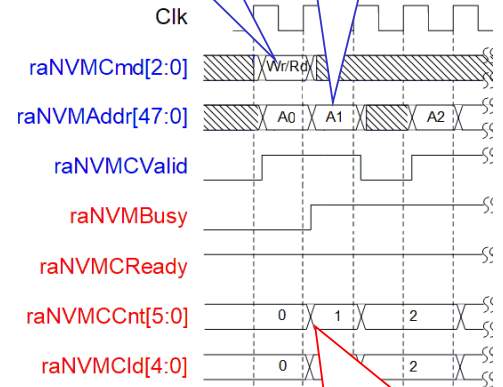


コマンド基本I/Fの信号線

2020/8/26

Design Gateway

コマンドとアドレスをセットし Validで命令発行
 Busy=1でもReady=1なら次コマンドを連続して発行可



コアは新コマンドを受けるとCnt値をインクリメント,32コマンドまで並列実行可能

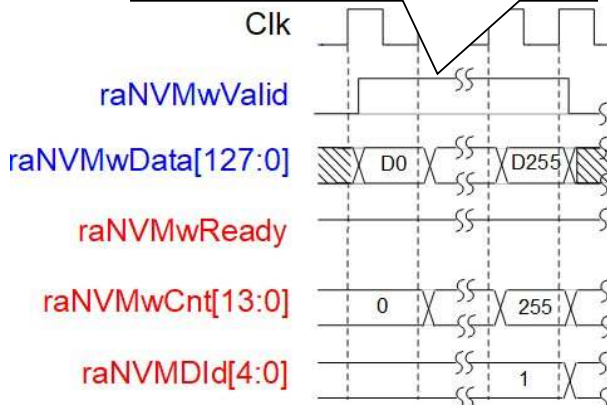
コマンドI/Fのタイムチャート

Page 9

特長3: データI/F

- ・ ユーザ回路I/Fはリード/ライト別の128bit幅データ・バス
 - 256クロック間Valid信号に同期して4KByte分連続ライト/リード

ユーザ回路は256クロック連続して全4Kバイトのライト・データをコアへ出力する

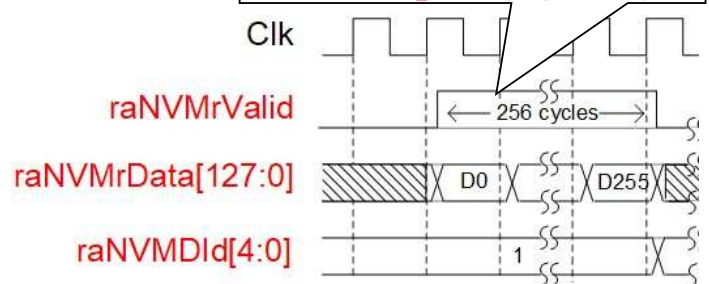


ライト・データの書き込み

2020/8/26

Design Gateway

コアは256クロック連続して全4Kバイトのリード・データをユーザ回路へ出力する



リード・データの読み出し

Page 10

特長4: 多機能

- SSDの状態を取得するSMARTコマンドに対応
 - SSD内部温度や総書き込み数などでSSDの健康状態をモニタリング
- FLUSHコマンドでキャッシュを任意のタイミングでフラッシュ
 - パフォーマンス維持と書き込みデータ退避のトレードオフをユーザが制御可
- Shutdownコマンドをコアが実行
 - 電源停止時の処理プロセスをユーザ指示でコアが実行

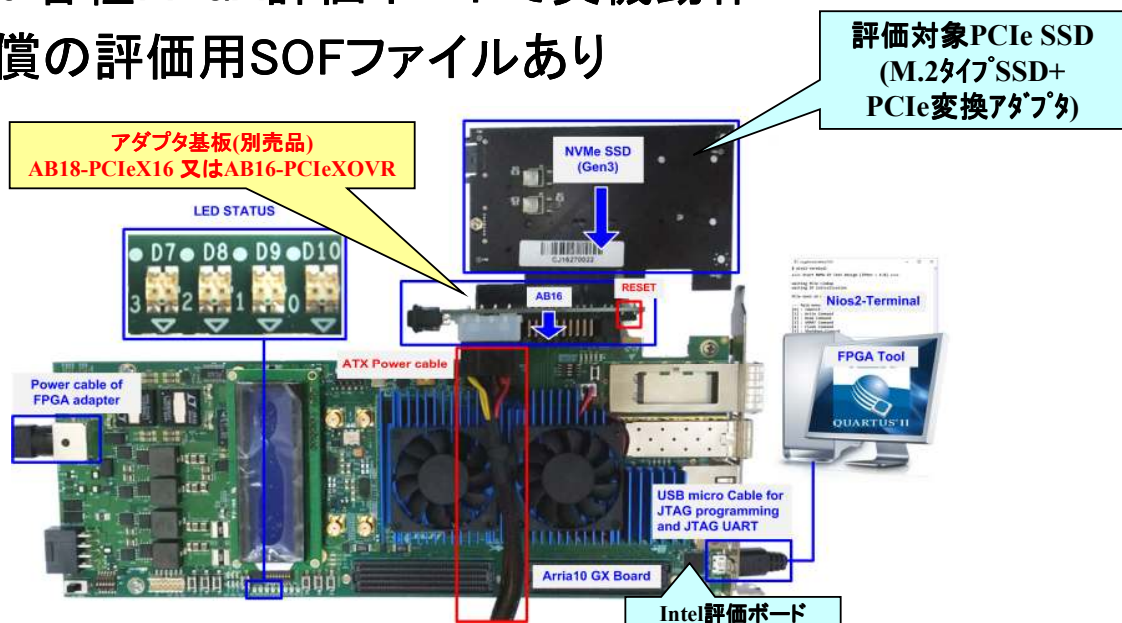
SSD温度やライト・サイクル数などの現状態を把握可能

```
<< SMART Log Information >>
Temperature           : 32 Degree Celsius
Total Data Read       : 47469 GB
Total Data Written    : 65373 GB
Power On Cycles       : 3991 Times
Power On Hours        : 79 Hours
Unsafe Shutdowns     : 220 Times
```

SMARTコマンドの実行結果例

特長5: 評価/開発環境

- Intel各種FPGA評価ボードで実機動作
- 無償の評価用SOFファイルあり



raNVMe-IPコア実機評価環境 (Arria10GXの評価環境例)

特長5: リファレンス・デザイン

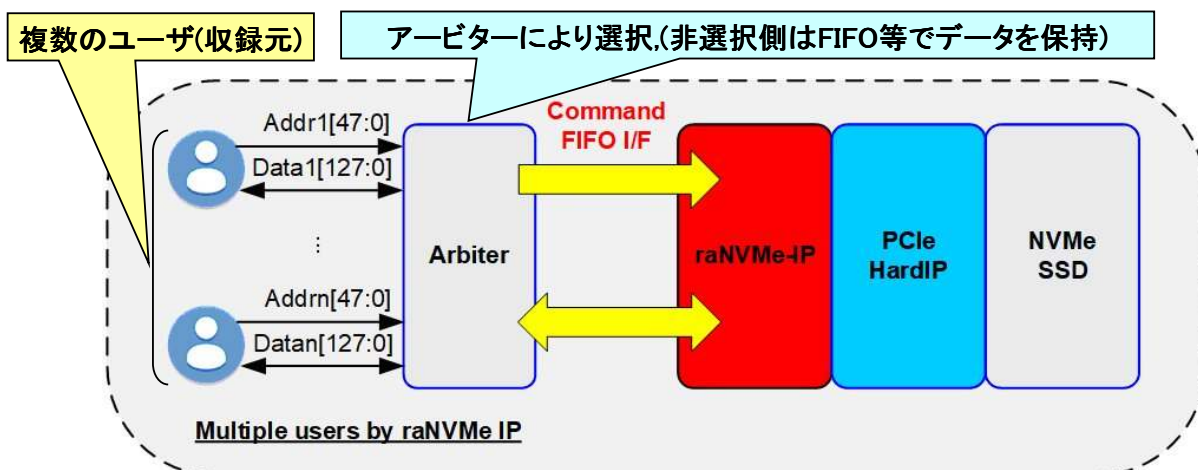
- ・ 評価用ビットファイルのQuartusプロジェクト
- ・ IPコア以外の全ソースコード(ファーム含む)を提供
- ・ ユーザ製品の開発期間短縮に貢献
 - まず最初に元のリファレンスで実機動作を確認
 - そこからユーザ製品に向け少しずつ編集
 - 編集ごとに実機動作をStep by Stepで確認



大きな後戻りがなく確実に短期間での製品開発が可能!

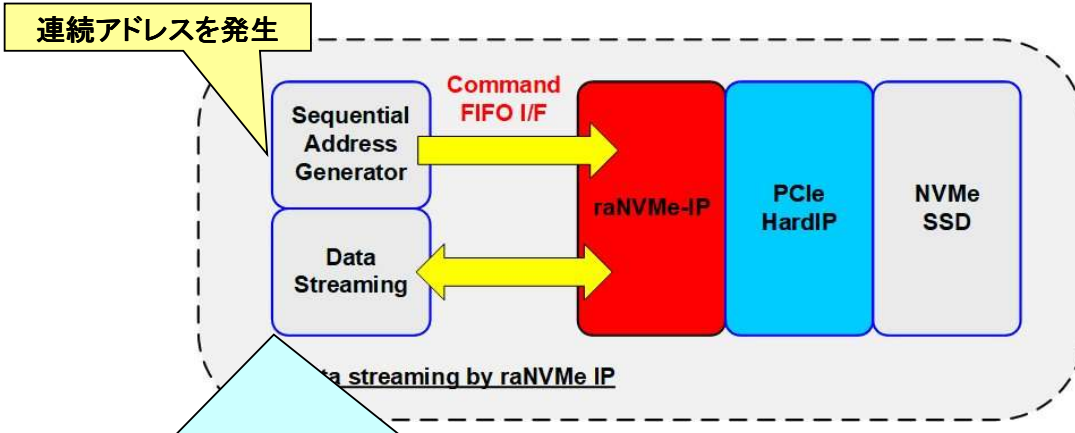
raNVMe-IPアプリケーション1

- ・ 複数ユーザ(ホスト)間によるSSDの共有システム
 - 複数のユーザ回路(例: センサ・データ)で単一のSSDを共有
 - 複数ユーザ回路とraNVMe-IPコア間にアービター回路を挿入



raNVMe-IPアプリケーション2

- ・ データ・ストリーム系アプリ(動画収録等の開始～停止タイプ)
 - 総収録データ量が開始/停止次第で事前に決まらないアプリ向け
 - 4Kバイト単位で収録,余剰データ量も最大4Kバイトと節約できる



収録データ発生元(4Kバイト単位で停止後の余剰分は0-Fill等のダミーデータで充填)

raNVMe-IP紹介: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/raNVMe-IP_A.html
- ・ 問い合わせ
 - 株式会社Design Gateway
 - E-mail : info@dgway.com
 - FAX : 050-3588-7915

