

# SATA-IP デバイス向けリファレンス・デザイン説明書

Rev1.2 2009/06/05

このドキュメントは Xilinx 製 ML506/505 評価ボードで動作する SATA-IP デバイス向けのリファレンス・デザインに関して説明したものです。

## 1. SATA について

シリアル ATA (SATA)は従来のパラレル ATA(PATA)に替わる革新的なストレージ・インターフェイスです。また、最新の SATA インターフェイスにおける転送速度は、SATA-I 規格の 1.5Gbps から SATA-II 規格の 3.0Gbps に高速化されています。SATA プロトコルによる通信システム全体としては、図1に示すように、アプリケーション・レイヤ、トランスポート・レイヤ、リンク・レイヤ、物理(PHY)レイヤ、の4レイヤにより実装されるアーキテクチャとなります

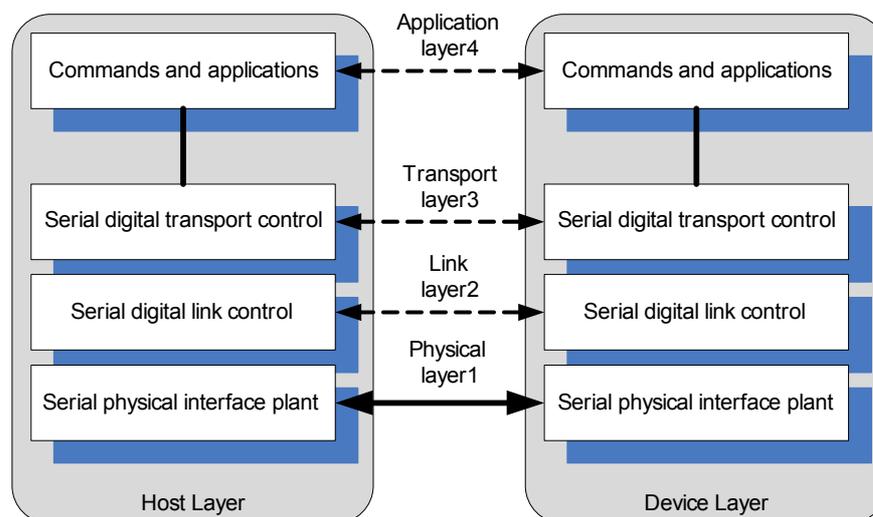


図 1: SATA のレイヤ構造

アプリケーション・レイヤはコマンド・ブロック・レジスタの制御を含む ATA コマンドの実行を担当します。トランスポート・レイヤではパケットや FIS(Frame Information Structure)と呼ばれるフレームによってホスト～デバイス間で転送される制御情報やデータを管理します。リンク・レイヤにおいては、生成されたフレームをもとにバイトごとの 8b/10b エンコード/デコードの実行や、10 ビットのデータ・ストリームが受信側で正しくデコードされるよう制御キャラクタの挿入を行います。PHY(物理)レイヤは、シリアル・データとして外部信号線上に流れるエンコード情報を送受信します。

本リファレンス・デザインでは、デバイス側において SATA-IP を含めた全 SATA 通信レイヤの具体的な実装方法例を紹介します。この評価システムは RAM ディスクをエミュレートし、SATA-II をサポートするチップセットが搭載されたホスト PC と接続することで SATA による高速転送を評価できます。

SATA-IP コアは Virtex-5 デバイスの GTP トランシーバと組み合わせて動作し、本リファレンス・デザインは Xilinx 製 ML506/505 評価ボード上で実装されます。より詳細については以下で説明します。

## 2. 動作環境

本リファレンス・デザインは図 2 に示される以下の環境で動作します。

- Xilinx 製 ML506/505 評価ボード(ML505 で動作するためにはデバイスを変更しての再コンパイルが必要な場合があります。)
- ISE 10.1.03 / EDK 10.1.03
- SATA インターフェイスを持つホスト PC (ML506/505 ボードのキットに付属する SATA クロスオーバー・ケーブルを使って、ボード上の J40 にてホスト PC と接続します。)
- シリアル(RS232C)ターミナル (ML506/505 ボードの P3 にて RS232C クロスケーブルで接続します) ターミナルの設定は、ボーレート=115,200 / データ=8bit / Non-Parity / Stop ビット=1bit としてください。

シリアル通信は必ずしも必須ではありませんが開発デバッグ時には便利です。

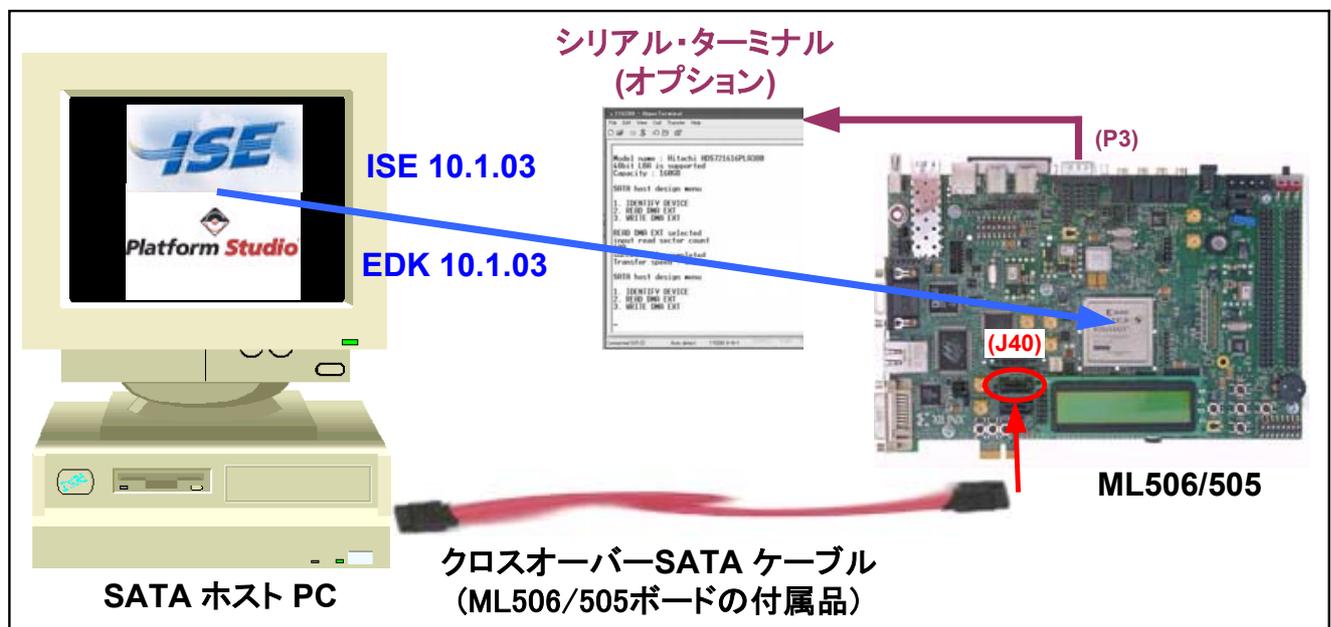


図 2: リファレンス・デザインの動作環境

本リファレンス・デザインの具体的な操作手順につきましては、「SATA\_IP デバイス・デモ手順書」を参照してください。また、リファレンス・デザインの SATA-IP コアは動作制限があり1時間経過後にデータ転送を停止します。

### 3. ハードウェアの説明

#### ● Virtex5 FPGA 上で実装される SATA IP デバイスデザイン

下図 3 のブロック接続図に示される通り、SATA-IP コアはトランスポート・レイヤの一部を含みますがほとんどがリンク・レイヤで構成されます。従って、PHY(物理)レイヤとトランスポート・レイヤはユーザ自身で用意しなくてはなりません。本リファレンス・デザインは Xilinx 製 ML506/505 評価ボードを対象として、トランスポート・レイヤと PHY レイヤの実装例を解説したものです。

本リファレンス・デザインにおいて ML-506 はボード上の DDR2-SDRAM を使った SATA の RAMDISK デバイスとして動作します。ホスト PC 上の WindowsXP からはこの RAMDISK が SATA の周辺デバイスとして認識され、リードやライト動作を行うことが可能です。

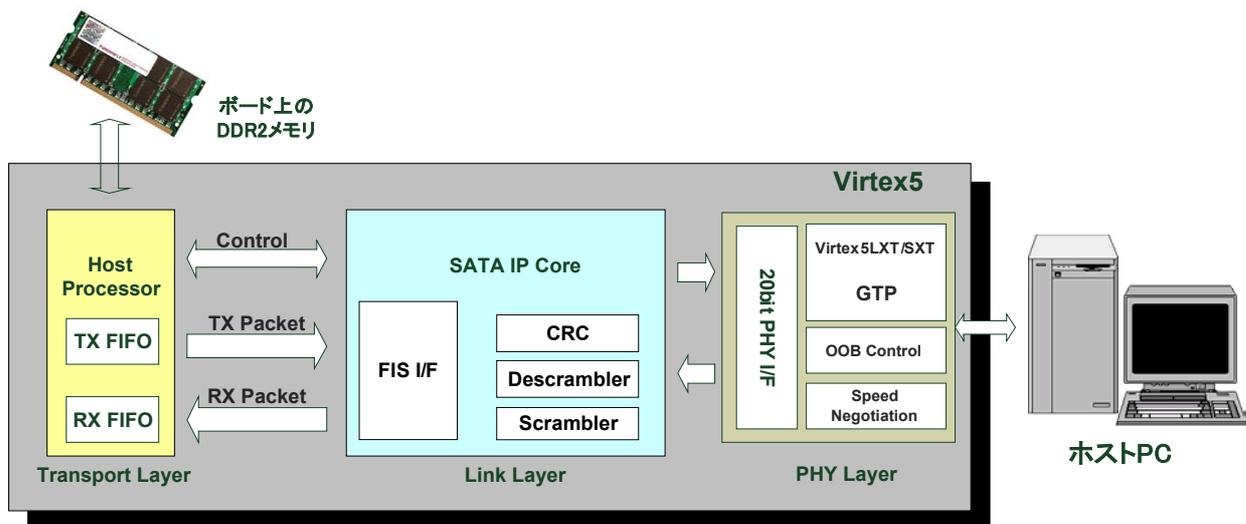


図 3: リンク・レイヤとトランスポート/PHY レイヤのブロック接続図

#### ● PHY レイヤ

Virtex-5 LXT/SXTには GTPブロック内に高速シリアル通信回路が内蔵されており、SATAのPHYレイヤはこのGTPブロックで実装されます。PHYレイヤはOOB(Out-of-Band)回路、SATA-I/IIの双方をサポートするための自動ネゴシエーション機能、そして20ビット幅でリンク・レイヤと通信するインターフェイス等が含まれます。

本リファレンス・デザインは Xilinx のウェブサイトからダウンロードできるアプリケーション・ノート(XAPP870)をもとに編集されたため、PHYレイヤは Virtex5 の GTPリソースに最適化されています。XAPP870 は SATA ホスト用のため、本リファレンス・デザインにおいては OOB 制御を SATA デバイス向けに修正しています。なお、SATA の特性情報については同じく Xilinx のウェブサイトより特性評価レポート(RPT087)がダウンロードできます。

ユーザ基板を新規に設計する場合、基板での高速シリアル特性を維持するため、UG196(Virtex-5 FPGA RocketIO GTP Transceiver User Guide)に記載のデザインガイドラインを遵守する必要があります。リファレンス・デザインにおいて PHY レイヤ回路のソースコードは“sata2phy\_ml505.v”内に記載されていますが、そこには“dev\_oob\_control.v”モジュールおよび“speed\_neg\_control.v”のインスタンスを含みます。

この“sata2phy\_ml505.v”は、自動ネゴシエーション機能、SATA-II 固定、SATA-I 固定、の3つの動作モードから選択できるよう設計されています。動作モードは、“sata2phy\_ml505.v”の 20～22 行目の3つの“define”文のうち、どれか1行を有効とし他の2行を“//”によるコメントアウトとすることで選択できます。リファレンス・デザインのデフォルトは下図 4 に示すように自動ネゴシエーション機能を有効としています。

```

14:
15: // Select
16: // - FIXI    for Fixed-SATAI speed PHY design
17: // - FIXII   for Fixed-SATAII speed PHY design
18: // - AUTONEG for Auto speed negotiation PHY design
19:
20: //`define FIXI
21: //`define FIXII
22: `define AUTONEG // Auto speed negotiation
23:

```

図 4: “sata2phy\_ml505.v”での PHY モード設定方法

自動ネゴシエーション機能を有効とした場合、J40 の SATA コネクタには SATA-I あるいは SATA-II のどちらのホスト PC も接続可能ですが、SATA-I 固定あるいは SATA-II 固定とした場合は接続可能なホスト PC が限定されます。

Virtex5 デバイスでは1個の GTP\_tile 内に1個の共有 PLLと2個の GTP チャンネルを格納しますが、SATA デバイスデザインにおいては、GTP\_Tile 内の 2 個の GTP チャンネルのうち1個の GTP しか使うことができません。他方の GTP チャンネルは必ず未使用の状態のままにしてください。

SATA-I あるいは SATA-II 固定モードの場合、PHY デザイン中で消費する DCM 数が1個で済むため、DCM が 2 個必要な自動速度ネゴシエーションより DCM リソースを節約できる利点があります。

(注意) PHY パラメータの変更はリファレンス・デザインのハードウェアを変更して再コンパイルする必要がありますため、製品版でのみ可能です。

#### ● トランスポート・レイヤ

トランスポート・レイヤの構造はハードウェアのアーキテクチャやユーザのアプリケーションに依存するため、ユーザ自身で設計する必要があります。本リファレンス・デザインにおいては、トランスポート・レイヤに MicroBlaze で動作させるために NPI インターフェイスを使っています。一般的な SATA コントローラと同様に、本リファレンス・デザインもメインメモリ上に FIS データを作成し、DMA によってリンク・レイヤと通信します。

本リファレンス・デザインにおいてトランスポート・レイヤのソースコードは SATA-IP コアと PHY レイヤのインスタンスを含む“npi\_sata\_target.vhd”内に記載されています。トランスポート・レイヤにおいてはホスト向けリファレンス・デザインと全く変更点はありません。従ってホスト向けデザインとの違いはソフトウェアのみとなります。ホスト向けリファレンス・デザインと同様にトランスポート・レイヤにおいては MicroBlaze をローカル・プロセッサとして使用し、メインメモリのコントローラとして MPMC (Multi-port memory controller)を使います。

● メモリコントローラ～IP コア間の接続回路

SATA-IP コアのインターフェイス信号定義を表 1 に示します。トランスポート・レイヤとのインターフェイス信号は大きくは送信系と受信系の2グループに分かれます。データ送信時および受信時のトランスポート信号波形をそれぞれ図 5 と図 6 に示します。図 7 はロジック接続のブロック図です。表2 は MicroBlaze からみたレジスタ・マッピングです。

MicroBlaze 上のソフトウェアは DMA を使いメインメモリ上の FIS データを SATA-IP のリンク・レイヤに送信あるいはリンク・レイヤからのデータをメインメモリ上に受信します。リンク・レイヤからのデータ受信時は、FIS のヘッダ情報をチェックしデータ FIS だけは別に指定したアドレスに転送するようになっています。また、送信時もデータ FIS のヘッダをデータパケットに自動的に追加して送信します。このメカニズムにより、ユーザ回路はヘッダ情報を意識せずに送受信データを管理できます。

信号名	方向	説明
共通インターフェイス信号		
trn_reset	In	コアをリセットする Low アクティブのリセット信号、正論理。
trn_link_up	Out	コアと SATA-PHY との通信が確立されると本トランザクション・リンク・アップ信号がアサートされる。
trn_clk	In	コアに対して供給するホストとのトランザクション・インターフェイス信号(trn_XXX)用のクロック信号。コア外部にて GCLK によるグローバル・クロック・バッファの挿入が必要。trn_clk の周波数は core_clk と同じかそれ以上とする必要がある。
core_clk	In	IP コアの動作クロック。(SATA-I の場合 37.5MHz で SATA-II の場合 75.0MHz) 本クロックは SATA PHY にて生成される。
dev_host_n	In	コアが SATA Host と SATA Device のどちらとして使われるかを指定する。SATA Host の場合'0'とし SATA Device の場合'1'とする。

信号名	方向	説明
送信トランザクション・インターフェイス信号		
trn_tsof_n	In	Transmit Start-Of-Frame (SOF): 送信 SATA FIS パケットの開始信号、負論理。
trn_teof_n	In	Transmit End-Of-Frame (EOF): 送信 SATA FIS パケットの終了信号、負論理。
trn_td[31:0]	In	Transmit Data: 送信 FIS パケットの 32 ビット・データ信号。
trn_tsrc_rdy_n	In	Transmit Source Ready: Host は trn_td[31:0] に有効な送信データを用意し本信号を Low とすることで転送を要求する、負論理。
trn_tdst_rdy_n	Out	Transmit Destination Ready: コアは Host から送られる送信データを trn_td[31:0] で受け取ることができる状態を示す信号、負論理。 trn_tsrc_rdy_n は本信号がネゲートされてから4trn_clk 期間以内にネゲートする必要がある。すなわち IP コアは本信号をネゲートしてから 4DWORD 分までの送信データ(trn_td[31:0])を受け取ることが可能。
trn_tsrc_dsc_n	In	Transmit Source Abort: Host は現在の SATA FIS パケット送信を中断したことを示す、負論理。SOF から EOF までの期間内であれば Host は本信号をいつアサートしても良い。Host が本信号をアサートした場合、コアは現在の転送を中断するために SYNC プリミティブを送信する。
trn_tdst_dsc_n	Out	Transmit Destination Abort: コアは現在の SATA FIS パケットを中断していることを示す、負論理。物理的なリンクがリセット状態となった場合にアサートされる。

表 1: SATA IP コアのインターフェイス信号

信号名	方向	説明
受信トランザクション・インターフェイス信号		
trn_rsuf_n	Out	Receive Start-Of-Frame (SOF): 受信 SATA FIS パケットの開始信号、負論理。
trn_reof_n	Out	Receive End-Of-Frame (EOF): 受信 SATA FIS パケットの終了信号、負論理。
trn_rd[31:0]	Out	Receive Data: 受信 FIS パケットの 32 ビット・データ信号。
trn_rsrc_rdy_n	Out	Receive Source Ready: コアが有効な受信データを trn_rd[31:0]に出力されている状態を示す、負論理。
trn_rdst_rdy_n	In	Receive Destination Ready: Host が trn_rd[31:0]で受信データを受け取ることができる状態を示す信号、負論理。 trn_rsrc_rdy_n は本信号がネゲートされてから 4trn_clk 期間以内にコアによってネゲートされる。従って Host は本信号をネゲートして以降にコアから送られてくる最大 4DWORD 分の受信データ(trn_rd[31:0])を受け取ることが可能な回路を実装しなくてはならない。
trn_rsrc_dsc_n	Out	Receive Source Abort: コアは現在の SATA FIS パケットを中断したことを示す、負論理。 SOF から EOF までの期間内であればいつでもアサートされる可能性がある。
trn_rdst_dsc_n	In	Receive Destination Abort: Host は現在の SATA FIS パケット受信を中断したことを示す、負論理。 Host が本信号をアサートした場合、コアは現在の転送を中断するために SYNC プリミティブを送信する。

信号名	方向	説明
Virtex5 GTP の SATA PHY インターフェイス信号		
PHYRESET	In	コアに対する PHY のリセット信号、正論理。本信号は PHY のリセット時に SATA PHY とインターフェイスをとるデータバッファをクリアするために使われる。
PHYCLK	In	Virtex5 GTP の SATA-PHY と通信する 16 ビット・データ用のリファレンスクロック。 - SATA-I の場合 75MHz - SATA-II の場合 150MHz Virtex5 GTP の SATAI/II PHY との通信には送信(TX)と受信(RX)では単一のクロックを共通して使うことができる。その理由として TX/RX データは実転送速度とリファレンス・クロック周波数の誤差を吸収するために GTP 内部に設けられている elastic バッファが使えるためである。本クロックは Virtex5 内部 DCM によって生成する必要があり SATA 速度ネゴシエーション・ロジックにより動作周波数が選択される。
TXDATA[15:0]	Out	コアから GTP に対して出力される 16 ビットの送信データ
TXDATAK[1:0]	Out	送信データのデータ/制御信号の認識シンボルとして使われる 2 ビット信号。 ("00":データ・バイト,"01":制御バイト,"1X":未定義)
RXDATA[15:0]	In	GTP からコアに出力される 16 ビットの受信データ
RXDATAK[1:0]	In	受信データのデータ/制御信号の認識シンボルとして使われる 2 ビット信号。 ("00":データ・バイト,"01":制御バイト,"1X":未定義)
RXDATAVALID	In	RXDATA[15:0]と RXDATAK[1:0]上のデータや認識シンボルが有効であることを示す。
LINKUP	In	SATA リンクの通信が確立されていることを示す、正論理。
PLLLOCK	In	GTP 用の DCM がロックできていることを示す、正論理。

表 1: SATA IP コアのインターフェイス信号(続き)

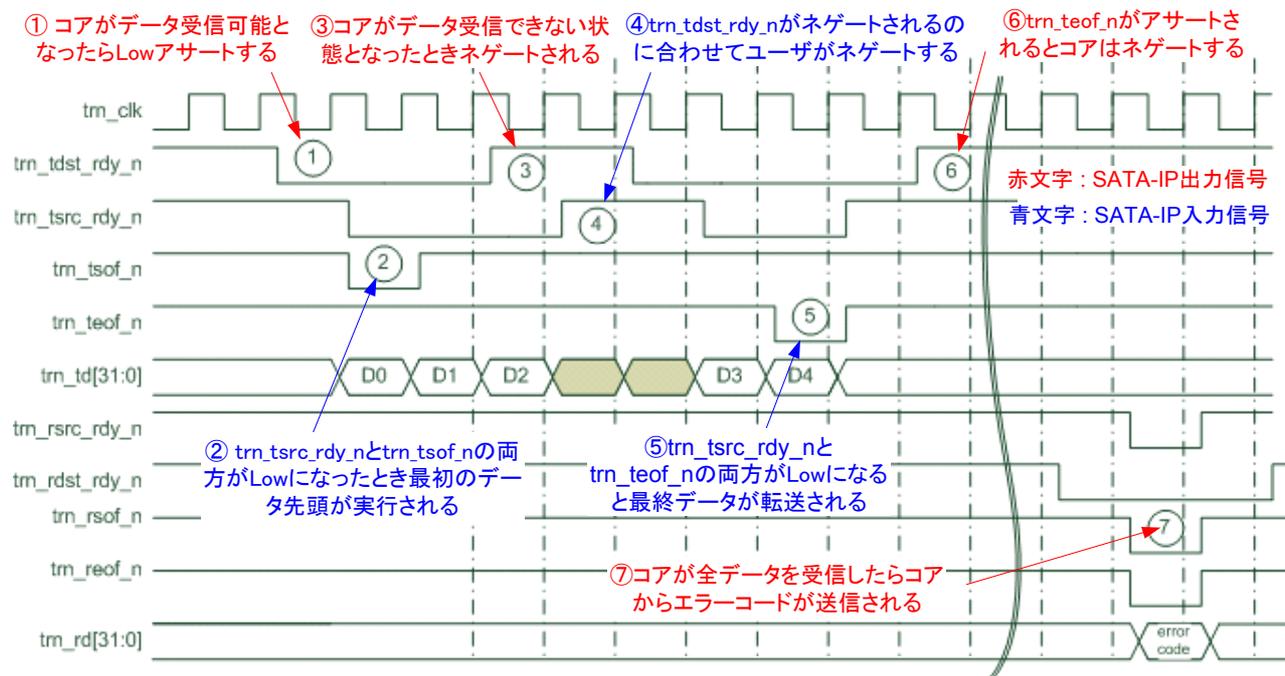


図 5: 送信トランザクションの信号波形

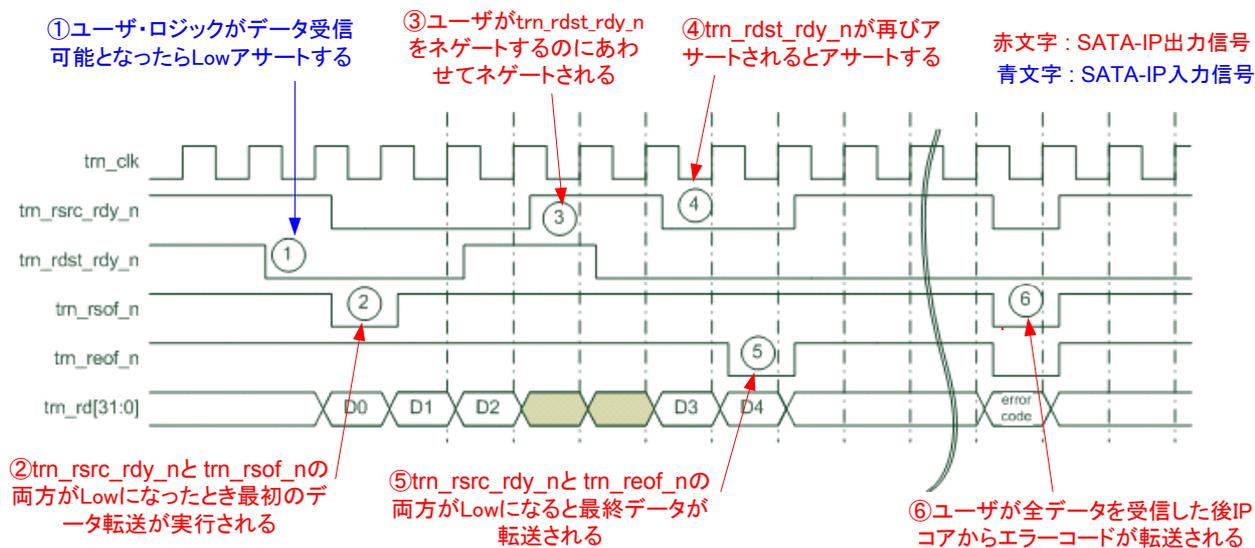


図 6: 受信トランザクションの信号波形

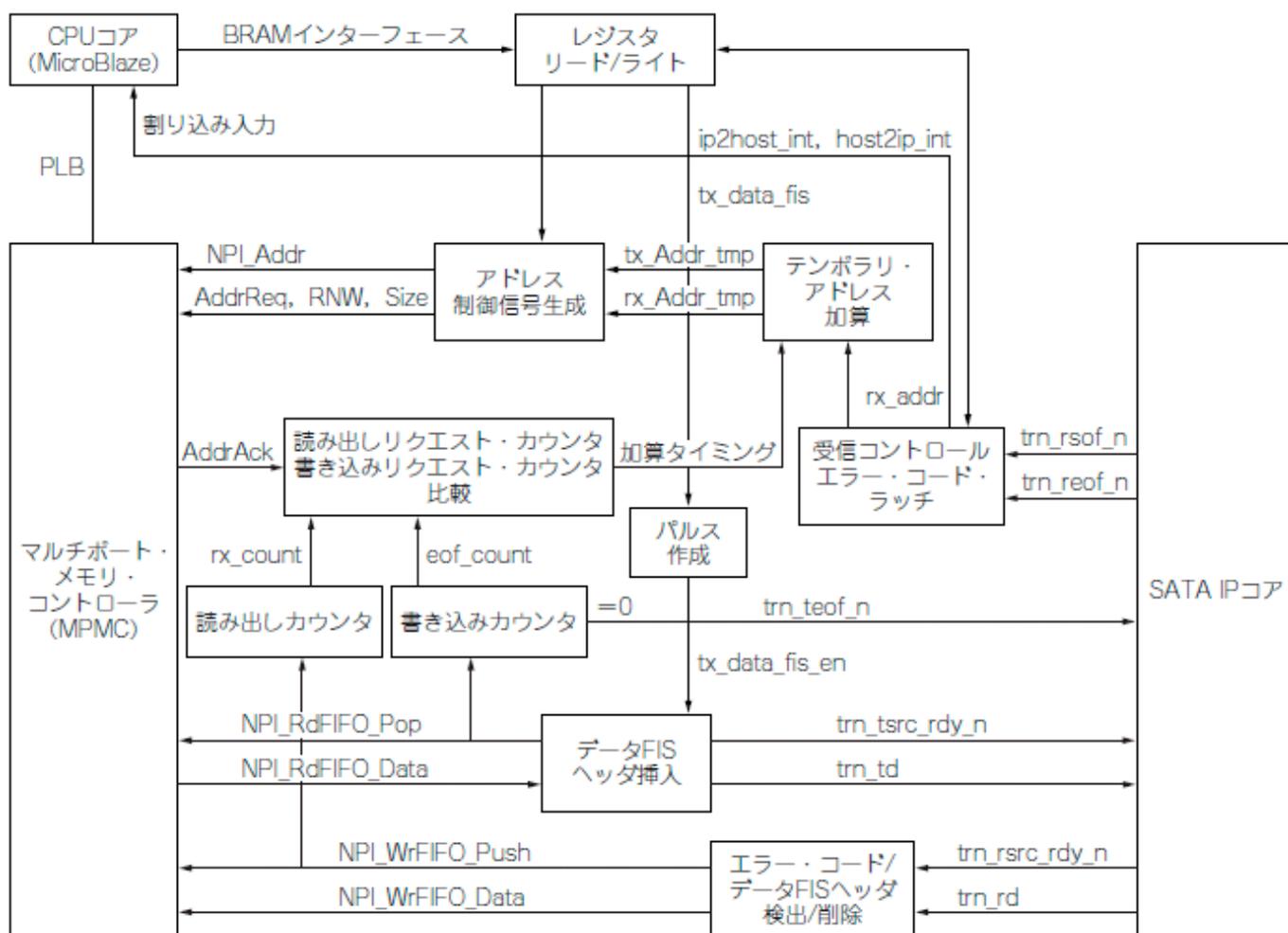


図 7: ロジック接続のブロック図

アドレス Rd/Wr	レジスタ名 ("sata_host.c"でのラベル名)	説明 (Bit ならばはリトル・エンディアンで表記)
BA+0x00 Rd	ステータス・レジスタ (STATUS)	[8]MPMC 準備完了 [7:4]OOB 状態データ [3]GTP PLL がロック [1]Gen2 でリンク [0]SATA IP コアがリンク
BA+0x04 Rd	エラー・コード・レジスタ (ERROR_CODE)	SATA IP のエラーコード。送信完了および受信完了時にセットされる。CRC エラーや FIS エラーはここで判別できる。
BA+0x08 Rd	割り込みクリア・レジスタ (INT_CLEAR)	空読みすると割り込みが解除される。
BA+0x0C Rd	受信ワード数レジスタ (RX_COUNT)	受信したワード数、クリアするまで全ての FIS データの受信数が積算される。
BA+0x00 Wr	送信データ格納アドレス (TX_ADDR)	送信するデータを格納している先頭のアドレスを指定する。
BA+0x04 Wr	受信データ格納アドレス 1 (RX_ADDR)	受信したデータのうち、データ FIS 以外のものが格納されるアドレス。1つ受信するたびに+100h アドレスが進む。
BA+0x08 Wr	コントロール・レジスタ (CONTROL)	[31]SATA リセット [30]送信要求 [29]データ FIS を送る [15:0]送信データワード数。本レジスタの書き込みで RX_COUNT レジスタがリセットされる。
BA+0x0C Wr	受信データ格納アドレス 2 (RX2_ADDR)	データ FIS を受信したときに格納される先頭アドレス。

(BA: ベース・アドレス)

表 2: MicroBlaze 側からのレジスタ・マップ

## 4. ソフトウェアの説明

### ● SATA デバイス動作

基本的に SATA 周辺デバイスは SATA ホストから送られてくる全ての必須(mandatory)コマンドをサポートしてはなりません。しかし本リファレンス・デザインにおいては、ユーザが SATA デバイスのソフトウェア制御の基本を理解しやすくするため、ソフトウェアを極力簡素化し動作に必要な最小限のコマンドのみサポートしています。

本デザインもホスト向けリファレンス・デザインと同様、SATA の通信は FIS (Frame Information Structure)データ構造を介して実行されます。デバイスデザインの MicroBlaze はメインメモリ上に FIS データを構築し、バス・マスタとなる DMA によって SATA ホストに転送されます。また、SATA ホストからの FIS データも同じように DMA によってメインメモリに転送されます。

従って SATA デバイスにおいて MicroBlaze は以下の手順でソフトウェアを実行します。

- (1) ブートアップが完了すると、SATA ホストに対して RegD2H の FIS データ・ストラクチャを送信します。
- (2) SATA ホストからのコマンドを待ちます。
- (3) 受信したコマンドを実行します。
- (4) FIS データを SATA ホストに送信します。
- (5) 必要に応じて追加の FIS データの送受信を行います。

## ● リファレンス・デザインのソフトウェア

本リファレンス・デザインのソースコードは“sata\_device.c”内に格納されています。最小限の実装となるため、ソースコードには以下のような制約があります。

- S.M.A.R.T (Self-monitoring, Analysis and Reporting Technology)のようなオプション機能はサポートされない。
- 48bit LBA はサポートされない。
- キャッシュはサポートされない。(リファレンス・デザイン事態が RAMDISK の周辺デバイスをエミュレートしているため、もともと不要。)
- UltraDMA は Mode5 およびそれ以下のモードのみサポート。

リンクが確立された後、ソフトウェアは RegD2H (Device to Host FIS)を送信し、SATA ホストからのコマンドに従って処理を実行します。本リファレンス・デザインでは SATA ホストから送られるコマンドを制限するため、独自の IDENTIFY DEVICE データが用意されています。ML506/505 ボードに標準装備される DDR メモリ容量が 256M バイトのため、ソフトウェアでは 128M バイトの容量と宣言しています。

実装を簡略化するため、SET MULTIPLE MODE においての上限は 1 とセットされるため、READ/WRITE SECTOR コマンドと READ/WRITE MULTIPLE コマンドのコマンド処理シーケンスは同一となります。この場合 PIO 転送パフォーマンスはあまりよくありませんが、実際には SATA ホストは PIO モードではなく UDMA モードを使って転送を行うため、実用上の問題はありません。

READ/WRITE においてアドレス・モードは CHS (Cylinder/Head/Sector) モードと LBA (Logical Block Address)モードの2種類があります。一般的に CHS モードは小容量の HDD に対して使われます。本リファレンス・デザインでは CHS モードは LBA モードのアクセスとして変換されます。

実装されるコマンドは、READ SECTOR、READ MULTIPLE、WRITE SECTOR、WRITE MULTIPLE、READ DMA、WRITE DMA です。これ以外の必須(mandatory)コマンドにおいては、何もせず単にエラーなしの RegD2H を返送します。必須でないコマンドにおいては、コマンドをサポートしていないことを示すため、エラーありの RegD2H を返送します。

本リファレンス・デザインではボードでの実動作チェックに必要とされる最小限のコマンドのみ実装されているため、ユーザのアプリケーション開発時には実際に ATA 規格が定める必須コマンドを全てサポートする必要があります。

## 5. ML506/505 における実ボードでの動作

### ● OS のブートアップ

本リファレンス・デザインを実装した ML506/505 ボードをホスト PC と接続した状態で、ホスト PC の WindowsXP ブートアップが完了すると、ユーザは図 8 に示すように OS から UltraDMA Mode5 転送モードの SATA 周辺デバイスとして認識されます。

OS ブートアップ時に発行される必須コマンドの例を表 3 に示します。ただし、OS ブート時に発行されるコマンドの種類やコマンド発行シーケンスは PC ・デバイス・ OS 等に依存するため、本リファレンス・デザインでは必ずしもどの PC でも認識されるという保証はありません。ユーザのホスト PC にて認識されない場合、実際に PC から発行されるコマンドに応じてリファレンス・デザインのソフトウェアを編集する必要があります。

コマンド・コード	コマンド種類
ECh	IDENTIFY DEVICE
E3h	IDLE
B0h	SMART
EFh	SET FEATURES (Set transfer mode)
C8h	READ DMA
C6h	SET MULTIPLE MODE
E1h	IDLE IMMEDIATE

表 3: WindowsXP ブート時に発行されるコマンドのサンプル

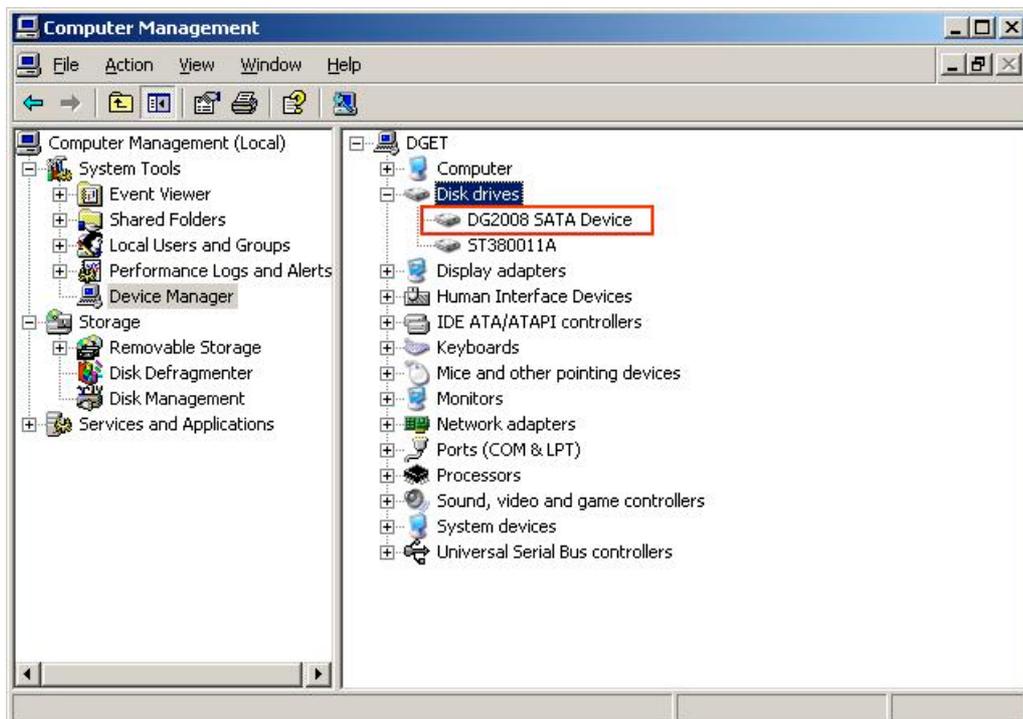


図 8: WindowsXP が ML506 ボードを SATA デバイスとして認識した画面例

SATA ホスト～デバイス間の転送モードは IDENTIFY DEVICE コマンドの後に設定されます。本リファレンス・デザインではオプションの S.M.A.R.T コマンドはサポートしていませんが、未サポートのコマンドに対してエラーの RegD2H で応答することにより OS は正しくブートします。また、IDLE や IDLE IMMEDIATE コマンドに対しては、エラーなしの RegD2H を戻すだけの処理を行います。また、SET MULTIPLE MODE の上限を 1 としているため、SATA ホストは SET MULTIPLE MODE コマンドにおいて 1 をセットします。

- OS によるディスクのフォーマット動作

WindowsXP から認識された RAMDISK へのフォーマットを実行したときの発行コマンド例を表 4 に示します。また、その結果を図 9 に示します。

コマンド・コード	コマンド種類
C8h	READ DMA
40h	READ VERIFY
CAh	WRITE DMA

表 4: フォーマット実行時に発行されるコマンドのサンプル

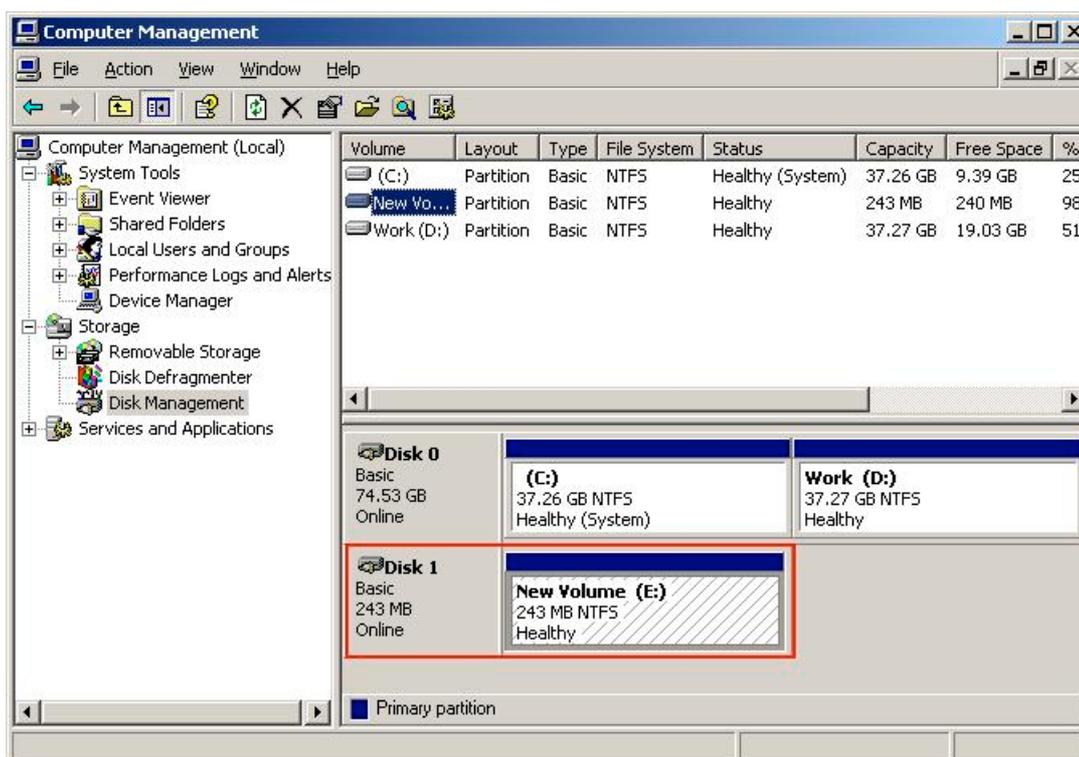


図 9: フォーマットにより新しいドライブとして認識した画面例

フォーマット実行時に WindowsXP は READ VERIFY コマンドを発行します。このコマンドにおいてはホストへのリード・データの転送が発生しないため、SATA デバイスは実メディアからのリード動作を実行せずにコマンドを完了させることが可能です。このため本デザインでは単にエラーなしの RegD2H をホストに返すのみの処理としています。

### ● OS のシャットダウン動作

表 5 に WindowsXP シャットダウン時に発行されるコマンド例を示します。WindowsXP は PC 内に残っているライト・データを転送し、続いて FLUSH CACHE と STANDBY IMMEDIATE コマンドを発行します。

コマンド・コード	コマンド種類
CAh	WRITE DMA
E7h	FLUSH CACHE
E0h	STANDBY IMMEDIATE

表 5: WindowsXP シャットダウン時に発行されるコマンドのサンプル

### ● パフォーマンス実測例

本リファレンス・デザインを実装した ML506 と市販の SATA HDD のベンチマークテスト結果例を表 10 に示します。市販の HDD と比較すると ML506 は 3 倍以上のパフォーマンスを達成します。

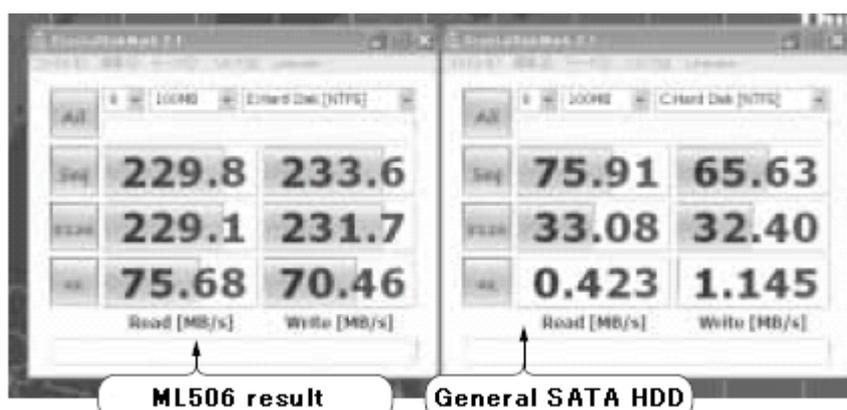


図 10: リード/ライトのパフォーマンス結果サンプル

## 6. サポート方針

本リファレンス・デザインは SATA-IP コアの設計技術の解説を目的とした参考情報のため、サポートは提供されません。

## 7. 改版履歴

リビジョン	日付	内容
0.1	2008/10/08	ドラフト版のリリース
1.0	2008/11/27	実ボードの動作記述を追加した第1版を発行
1.1	2008/12/12	dev_host_n 信号(コアの SATA Host/Device 区別信号)の追加
1.2	2009/06/05	trn_clk の周波数が core_clk 以上とする必要がある制約の記述を追加