

組込みアプリケーションにおけるDDR4バンク・グループ

シノプシス Graham Allan

DDR4は、1990年代に非同期DRAMから最初のSDRAMへ移行して以来、最も複雑なDRAMの世代交代といえます。新しいDDR4規格がもたらすシステム消費電力の削減と全体的な帯域幅拡大という利点を最大限に活かしたSoCを開発するには、DDR4 SDRAMの細部まで十分に理解しておく必要があります。

DDR4規格は、基本的にはDDR3を進化 / 発展させたものですが、新しい機能も20以上追加されています。これらの新機能の多くは、今後の組込みシステム・アプリケーションにおけるメモリーの利用方法に大きな影響を与えます。本稿では、DDR4のバンク・グループを効果的に活用してデータ効率を最大限に高める方法についてご説明します。

DDR4 SDRAMのバンク・グループとは

DDR4 SDRAMのバンク・グループは、もともとGDDR5グラフィック・メモリーで使われていた機能を転用したもので、バンク・グループの必要性を理解するには、まずDDR SDRAMのプリフェッチの概念を理解しておく必要があります。プリフェッチとは、DDRメモリーに対して1回のカラム・コマンドを実行すると何ワードのデータがフェッチされるかを表す用語です。DRAMのコアの動作速度はインターフェイス速度よりも遅く、このギャップを埋めるため、データに並列にアクセスし、これらをシリアル変換してからインターフェイスに送出する方法をとっています。たとえばDDR3のプリフェッチは8ワードですが、これは1回のReadまたはWriteコマンドを8ワードのデータに対して実行し、4クロック・サイクルの立ち上がりエッジと立ち下がりエッジの両方を使用して8つの連続したクロック・エッジでSDRAMインターフェイスから（またはSDRAMインターフェイスへ）バースト転送します。基本的に、DDR3のプリフェッチが8ということは、インターフェイスの速度がDRAMコアの8倍高速と考えることができます。プリフェッチには1つの問題があります。それは、SDRAMの最小バースト長がプリフェッチ・サイズによって事実上決まってしまうという点です。たとえばプリフェッチ・サイズ8のDDR3の場合、バースト長を4ワードとすることは非常に困難です。そこで、プリフェッチ・サイズを据え置いたまま従来と同じペースでインターフェイス速度を高速化する手段としてDDR4に採用されたのが、バンク・グループ機能です。

DRAMの世代が進んでもDRAMコア自体はそれほど高速化しておらず、SDRAMインターフェイスの高速化はもっぱらプリフェッチ・サイズの拡大によって実現していました。この従来の方法でDDR4を高速化しようとするとプリフェッチ・サイズは16になりますが、そうすると必要な配線が増加してDRAMは大幅に大型化し、コストも上昇します。こうした点を嫌って、DDR4では16ワード・プリフェッチの導入が回避されました。もう1つ重要なのは、プリフェッチ・サイズを16ワードとすると現在のコンピュータで一般的な64ビットのキャッシュ・ライン・サイズとの相性が悪くなるという点です。64ビットまたは72ビット・インターフェイスの典型的なコンピューティング環境では64ビットのキャッシュ・ラインを使用しており、プリフェッチ・サイズとバースト長が両方とも8の方が好都合です。キャッシュ・ライン・サイズとバースト長の境界が揃っていないと、組込みシステムのパフォーマンスに悪影響を与える可能性があります。

DDR4のプリフェッチとパフォーマンス

プリフェッチとパフォーマンスの関係を理解するため、まずこれまでのSDRAMを振り返ってみましょう。図1は、SDRAMの最大データ・レートと最大カラム・サイクルを各世代別に比較したものです。この図から分かるように、コアの速度は非常に遅く、世代間でほとんど変化していません。一方、インターフェイス速度は世代が進むにつれ大幅に向上去っています。SDRAMのコアがなかなか高速化しないのは、SDRAMの大容量化が進むことによってプロセス微細化の恩恵が相殺されているのが大きな理由です。

結局、SDRAMは大容量化を最優先して開発が行われており、プロセスの微細化がそれを補っている形です。これに対し、I/Oの速度は世代が進むにつれ着実に高速化しています。どうしてこのようなことが可能になるのでしょうか。その理由はプリフェッチにあります。

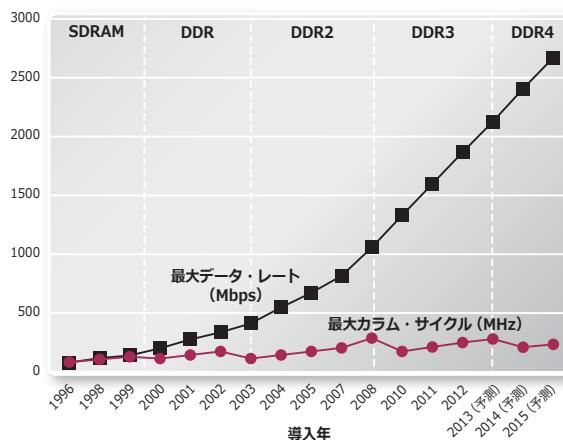


図1. SDRAMのパフォーマンスの向上

最初のSDR (Single Data Rate) SDRAMが登場した時は、プリフェッチの必要がありませんでした。1カラム・サイクルを実行するたびに1ワードのデータにアクセスし、それをそのままSDRAMから送出していました。実際には、SDRAMには「2Nルール」と呼ばれるものがあり、プリフェッチ・サイズを2とすることも可能でしたが、2Nが規格の制約となると判断され、最終的にコモディティ・デバイスとなったSDRAMはすべてプリフェッチ・サイズを1としていました。その後登場したDDR SDRAMでは、コアの速度がSDRAMで要求される帯域幅に追いつかなくなり、この時からインターフェイスの最大データ・レートと内部のカラム・サイクルにギャップが生じるようになりました (図2)。

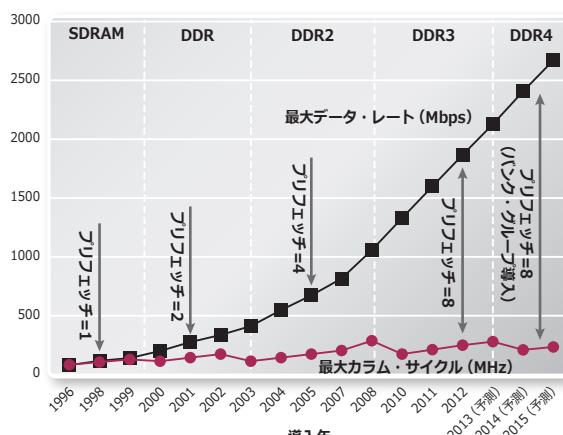


図2. プリフェッチによって実現するSDRAMのパフォーマンス向上

What's New in DesignWare IP?

組込みアプリケーションにおけるDDR4 バンク・グループ

前ページより続く▶

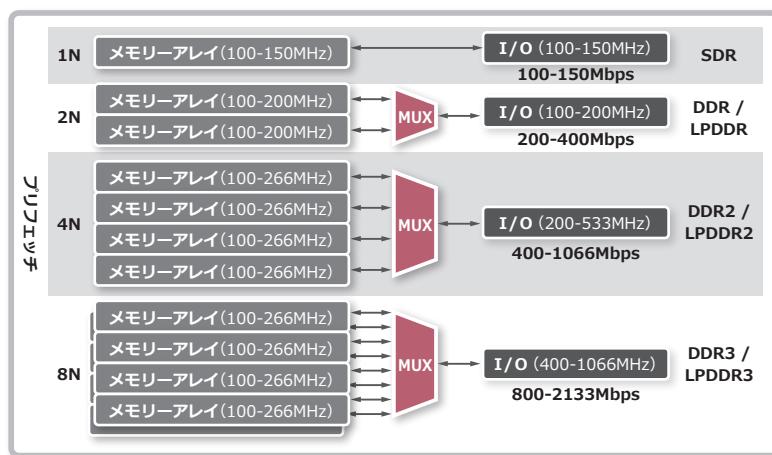


図3. これまでのDRAMのプリフェッチ

このギャップは世代が進むにつれて拡大しており、DDR2ではプリフェッチを4、DDR3ではプリフェッチを8とすることで対処しています。DDR4では、過大なプリフェッチ・サイズの弊害を避けるため、バンク・グループの概念を導入することによってプリフェッチ・サイズを8に据え置いています。

図3は、SDR SDRAMからDDR3 SDRAMまでの4世代のSDRAMにおけるプリフェッチの変遷を示したもので、1回のカラム・コマンド(Read / Write)でアクセスされるワード数をメモリーアレイの数で表すとともに、これらのワードをSDRAMインターフェイスで送受信するのに必要なおおよそのサイクル・タイム(MHz)も示しています。

一番上に示してあるのはSDR SDRAMです。1990年代中頃は、メモリーアレイの速度とI/Oの速度が一致していました。その後すぐに、SDRAMコアよりも高いデータ・レートを実現するためにプリフェッチ・サイズ2(2Nプリフェッチ)を採用したDDRが登場しました(LPDDRも同様のプリフェッチを採用)。図に示してあるマルチブレクサは2ワードにアクセスし、これらをクロックの立ち上がりエッジと立ち下がりエッジの両方でインターフェイスにシリアル出力します。このパターンが、8NプリフェッチのDDR3 / LPDDR3まで続きました。

DDR4とバンク・グループ

では、DDR4はどのようにして8Nプリフェッチのままインターフェイス速度を向上しているのでしょうか。図3に示したパターンをDDR4に当てはめると、DDR4のプリフェッチ・サイズは16になりますが、これが好ましくないのは前述したとおりです。そこで、DDR4ではバンク・グループという考え方を導入してこの問題を解決しています。これは、1つのバンク・グループで8Nプリフェッチを実行し、もう1つの独立したバンク・グループでも8Nプリフェッチを実行するというものです。各バンク・グループはそれぞれ独立しているため、カラム・サイクルはバンク・グループ内で完結し、別のバンク・グループの動作には影響を与えません。つまり、DDR4 SDRAMは内部のバンク・グループを時分割多重することによって、外部インターフェイスでバースト長8のデータを転送するのに必要な時間よりも内部コアのサイクル・タイムが長いという事実を隠蔽していることになります。図4は、組込みアプリケーションでよく使われるx16 DDR4 SDRAM(2バンク・グループ)におけるバンク・グループの構成を示したもので。

バンク・グループに関する新しい仕様項目

バンク・グループの導入に伴い、DDR4では新しい仕様項目が追加されています。特に重要なのが、tCCD_SとtCCD_Lの2つです。「CCD」は「Column to Column Delay」またはカラム・サイドの「Command to Command

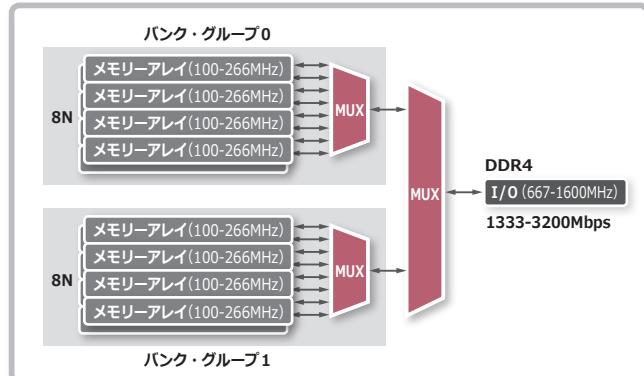


図4. DDR4の8Nプリフェッチ

Delay」の意味で、「_S」はshort、「_L」はlongを表します。

1つのバンク・グループ(8Nプリフェッチ)でカラム・コマンドを開始すると、そのバンク・グループの全サイクル・タイム内にプリフェッチを完了する必要があります。バンク・グループを切り替えてカラム・コマンドを実行する場合は、これらの仕様への依存は一切ありません。このような場合はtCCD_Sを使用します。この仕様値は4クロック・サイクルで、DDR3の場合と同様、アクセスに制約は生じません。

注意が必要なのは、同じバンク・グループ内でコマンドが連続するケースです。この場合、DDR4の新しい仕様項目であるtCCD_Lに注目する必要があります。この値は通常4クロック・サイクルより大きく、デザインのパフォーマンスに影響する可能性があるため、組込みアプリケーションでは特に注意が必要です。図5に、tCCD_SとtCCD_Lの違いを示します。

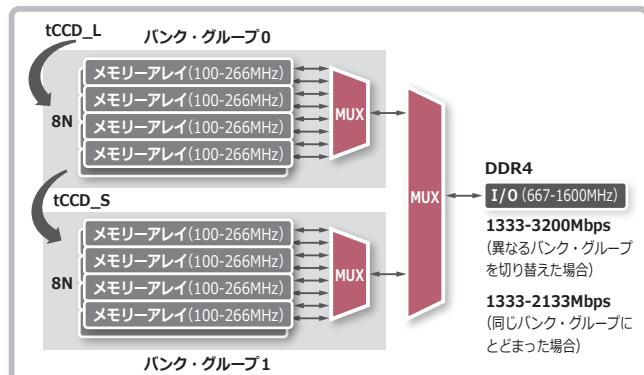


図5. DDR4のバンク・グループにおけるtCCD_LとtCCD_S

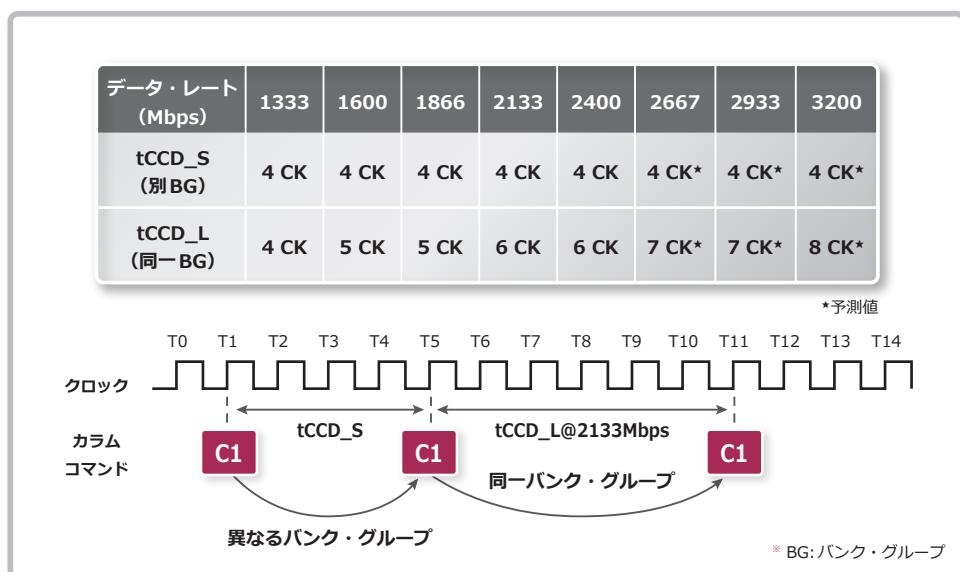


図6. 同一バンク・グループと異なるバンク・グループで2種類のtCCDパラメータを定義

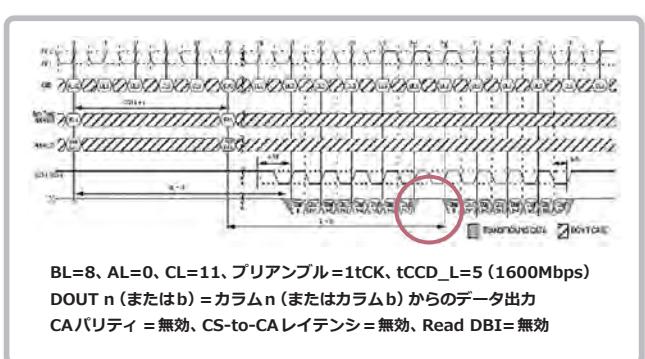
新しいtCCDが持つ大きな意味

図6に示すように、DDR4で異なるバンク・グループを切り替えるには4クロック・サイクルの遅延が必要です。4クロック・サイクルはちょうどバースト長8に一致します。4クロック・サイクルは立ち上がりと立ち下りの8クロック・エッジに相当するため、バースト長8なら4クロック・サイクルのすべてのクロック・エッジを利用して効率よくデータを送受信できます。これなら、バス帯域幅の無駄はありません。1つのコマンドを実行した後、データ転送のない無駄なバス・サイクルなしに次のコマンドを実行できます。

これに対し、同じバンク・グループでカラム・コマンドを連続実行した場合、1600Mbps以上ではtCCD_Lの値が4クロックを超えます(1600および1866Mbpsでは5クロック、2133Mbpsでは6クロック)。図6の表の下に示したタイミング図を見ると、異なるバンク・グループを切り替えた場合は4クロック・サイクルの間隔でメモリーにアクセスできています。これに対し、たとえば2133Mbpsで同じバンク・グループに対してカラム・コマンドを実行した場合は6クロック・サイクルの間隔が必要です。実際のデータ転送に必要なのは4クロック・サイクルのみで、残りの2クロック・サイクル、すなわちバス帯域幅の33%が無駄になります。

トラフィック・スケジューリングで バンク・グループの利用効率を向上

図7のタイミング図は、1600Mbpsの場合におけるDDR4のバンク・グループの影響を示したもので。tCCD_L=5とは、同じバンク・グループでカラム・コマンドを連続実行した場合、5クロック・サイクルの間隔が必要であることを意味します。その影響は、丸で囲んだ部分に見ることができます。5クロック・サイクルを使用した場合、1クロック・サイクルはデータの送受信が行われず無駄になります。データ・レートが高くなるとtCCD_Lの値も6、7、8クロック・サイクルと大きくなるため、この問題はさらに深刻になります。tCCD_L=8の場合、同じバンク・グループ内でカラム・コマンドを連続実行すると帯域幅の半分が無駄になります。

図7. カラム・コマンド間の無駄なクロック・サイクル^{※1}

DDR4には多くの新機能がありますが、組込みシステムのパフォーマンスと帯域幅を向上させるには適切なインプリメンテーションが必要です。シノプシスは、DDR4マルチPHY、拡張ユーバーサルDDRメモリー・コントローラ、検証用IPを含め、包括的なDDR4ソリューションをご提供しています。シノプシスのDesignWare DDR4ソリューションはDDR4とDDR3以外にローパワー・メモリー(LPDDR2、LPDDR3)もワンストップでサポートしています。シノプシスのIPが皆様のシステムにもたらすメリットについて、詳しくは<http://www.synopsys.com/IP/InterfaceIP/DDRn/Pages/default.aspx>をご参照ください。

著者紹介

Graham Allan : 2007年6月にシノプシス入社。以前はMOSAID Technologies社で半導体IPのマーケティング・ディレクタを担当。メモリー業界に23年以上従事し、業界カンファレンスでの講演経験多数。JEDECメモリー規格委員会にシノプシス代表として参加。SDRAM、DDR、DDR2メモリーの標準化で多くの助言を行う。DRAMデザインの分野で16件の特許を保有。

^{※1} 出典 JEDEC Solid State Technology Association (2012年9月)「DDR4 SDRAM STANDARD」(JESD79-4)