

# WHITEPAPER

Dinesh Maheshwari,  
CTO, Memory Products Division  
Cypress Semiconductor Corp.



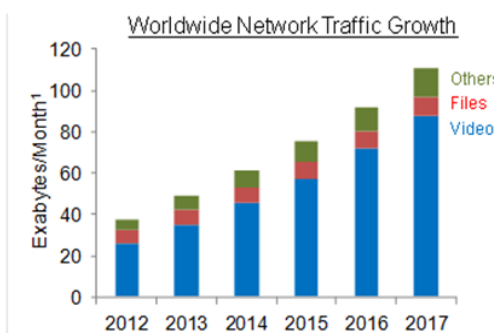
## QDR®-IV：最先端のネットワークシステムに 対応した次世代最高性能のメモリ標準

### 概要

インターネットの要求を満たすためにネットワークとコンピュータ機器のプロバイダがソリューションを増やすに当たって、メモリの諸機能が次第にシステム性能向上を左右するようになります。全ての、ルータ、スイッチ、および高性能のコンピューティングシステムは絶えず増大するレートで演算処理装置で処理可能なメモリを必要とします。従って、ランダムトランザクションレート(RTR)、即ちメモリが対応できるランダム読み込みまたは書き込みアクセス回数は重要なメモリ測定基準となります。このホワイトペーパーでは、サイプレスのQDR®-IVが、これらシステムの課題を満たすために、どのようにして最速のRTRおよび標準ベースのソリューションを提供するかについて説明します。

### 現在のネットワークシステムへの要求

世界規模のネットワーク通信量は、インターネットに接続されたデバイスや大量の帯域幅を使用するビデオサービスの普及にともなって、今後4年間で3倍になると予測されています。2017年までに、36億人のインターネットユーザが190億のネットワークに接続するでしょう。それに応えるために、ネットワークの機器プロバイダは、スイッチとルーターのラインカードを配備しています。あらゆる世代でラインカードレート(Gbit/s)またはパケットレート(MP/s)として測定されるラインカードの性能は倍増しています。図1と図2は、これらの傾向を示しています。パケット転送レートはラインカードレートに正比例していることに注意してください。パケットレート = ラインカードレート / (最小パケットサイズ + パケット間ギャップ)。最小パケットサイズは、20バイトのフレームギャップ付きで、64バイトです。それゆえ、100-Gbit/sのラインカードに対応したパケットレートは、150 MP/sです。



<sup>1</sup> ネットワーク通信量は、毎月エクサバイト(10<sup>18</sup> バイト)単位で測定

図 1. 世界的なネットワークトラフィックの増加

出典: Cisco Visual Networking Index (VNI) Forecast (2012–2017)

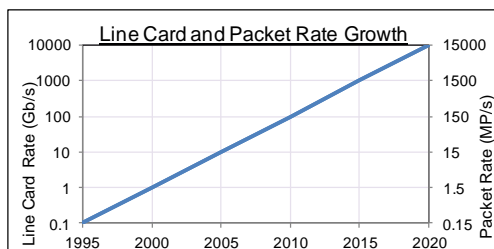


図 2. ラインカードとパケットレートの増加

出典: Ethernet Technology Summit

## ランダム トランザクション レート

ランダム トランザクション レート(RTR)は、メモリが所定の時間内に対応できるランダム読み出しまたは書き込みアクセスの数を表します。毎秒当たりのトランザクション数(例えばMT/s、GT/s)で測定されます。ネットワーク システムで必要なRTRは、パケット レートおよびメモリが対応するラインカードの機能に応じて変化する各パケット用に必要なランダム メモリ アクセスの回数によって決定されます。

## ラインカード メモリの機能

ネットワーク システムにおけるRTRの重要性は、ラインカード上メモリの6つの主要な機能との関連で見るのが最適です。以下のセクションで詳細に述べるように、これらの機能は全システムの性能を決定します。主要な性能のボトルネックは、スイッチング速度ではありません。ボトルネックになるのは、そのタイプと行先を決めるパケットのヘッダを調べ、転送テーブル内のルートを検索し、スケジュール用のパケットを分類し、様々な統計や状態を更新するために要する時間です。

## パケットバッファ

ルータ上の各ラインカードは、入力バッファと出力バッファを持っています。パッケージがスケジューリングされ、スイッチング ファブリックに送信される番を待っている間(分類と転送の決定がなされた後)、入力バッファは入力物理リンクからのパッケージを保持します。出力バッファは、パケットがスケジュールされ、外部物理リンクに配送されるように、スイッチング ファブリックからのパケットを格納します。パケット速度レートが高ければ高いほどバッファ サイズもより大きくならなければなりません。よって、パケット バッファ メモリの選択は、主に容量によって決定されます。大容量メモリが、パケットレートに対応するために必要なRTRを備えていないなら、システム メーカーは、多くの場合階層構造を設計します。この階層構造でメーカーはパケット バッファの先端部と終端部を格納するために「ヘッドテイル キャッシュ」を使用します。ヘッドテイル キャッシュは、パケット当たりで最大4~6のトランザクションを要求できます。

## 転送情報ベース ルックアップ テーブル

転送情報ベース(FIB)ルックアップ テーブルは、宛先IPアドレスに対応するスイッチ/ルータ上の終了ポートアドレスを格納します。ルックアップは、メモリへの複数アドレスを含む反復プロセスです。ルックアップは、ネットワーク スタック内の2層と3層の両方で行われます。2層ルックアップは、一般的には検索インデックスへの完全一致を期待し、通常はパケットごとに4つのランダムアクセスを必要とするハッシュ アルゴリズムを介して実装されています。最長のプレフィックス マッチを使うアルゴリズム3層ルックアップは、パケットごとに最大8ランダム アクセスを要求できます。

## パケット分類またはアクセス制御リスト

この処理ステップでは、プロセッサは、着信パケット(例えば、送信元アドレス、宛先アドレス、送信元ポート、宛先ポート、およびプロトコルからなる5組)の特性を調べ、パケットの通過を許可するか否かの決定を行います。パケットが、一旦分類または確認されると、それは一時的にスケジューリングのためにバッファ サブシステムに格納されます。アクセス コントロール リスト(ACL)へアクセスするためのアルゴリズム的アプローチは、パケットごとに10回以上のランダムアクセスを必要とします。

## スケジューリング

スケジューリングは、スイッチング ファブリックへパケットを送信するタイミングを決定するプロセスであり、パケットの宛先と必要なサービス品質(QoS)またはサービス クラス(CoS)に基づいて決定されます。パケットは、段階的なサービス提供(サービス プロバイダの収益区分)と関係のある、幾つかのクラスに分類されます。スケジューリング アプリケーションは、パケットごとに少なくとも1回の読み込みトランザクションと1回の書き込みトランザクションを必要とします。

### 統計と状態

ルータは、パケット単位とフロー（関連パケットのストリーム）単位の統計量を維持します。これは、プレフィックス、フロー、およびパケット分類の情報を格納するために使われるカウンタによって達成されます。カウンタの更新は、リード モディファイ ライト動作（2つのランダム メモリ トランザクション）を必要とします。RTRの要件は、パケット転送レートと必要なカウンタの更新数の両方に比例します。ライン カード内の同じメモリは、統計と状態に応じて共用することができます。

### RTR-主要なメトリック

各機能は、高いRTRを要求します。他のメモリ メトリックはそれほど重要ではありません。帯域幅は任意のデータ幅のRTRに比例して増加します。帯域幅は、データ幅の増加により増加する場合もあることに注意してください。メモリ容量は、パケットのバッファリングのための重要な属性ですが、ヘッドテイル キャッシュも同様に、この機能で高いRTRの必要性を明らかにしています。

### パケット レートとメモリ機能でのRTRの要件

表 1は、100Gbit/s (150MP/s)と200Gbit/s (300MP/s)ライン カード用のRTRの要件を示します。

メモリ機能	パケット当たりの トランザクション	100Gbit/s 150MP/s	200Gbit/s 300MP/s
ヘッドテイル キャッシュ/スケジューラ	4	600MT/s	1200MT/s
ルックアップ	4~8	600~1200MT/s	1200~2400MT/s
統計/ステート	8~6 (4~8カウンタ)	1200~2400MT/s	2400~4800MT/s
パケット バッファ	4~6	600~00MT/s	1200~1800MT/s

表 1. ライン カード用の RTR 要件

### RTR: ネットワーク システムを超えて

RTR性能の値は、ネットワーク システムに限定されるものではありません。RTRはまた、高性能のコンピューティング、汎用サーバ、および画像処理アプリケーションの重要要件でもあります。高性能コンピューティング システム（複雑な問題を解決するために集められた計算と記憶の供給源のグループであるクラスター）は、幾つかのストリームからのデータを処理するために高RTRのメモリを必要としています。共有された最終レベルのキャッシュ メモリを装備した「マルチコア」と「多コア」プロセッサを使って構築された汎用サーバでは、CPUの要求に対応する高RTRのメモリを必要とします。さらに、高性能の画像処理システムは、画像データへの高速なランダムアクセスを要求するデジタル信号処理アルゴリズムを実装しているので、このようなシステムに対して高RTRメモリは有用です。

### 高RTRのメモリ デバイスの選択

設計者は、システム性能を最適化するために、様々なメモリタイプで提供されるRTRの違いを理解する必要があります。コアメモリ技術、メモリ デバイス構造（バンク）、およびメモリ インターフェースは、任意のメモリ デバイスのRTRを決定します。

## メモリ技術

メモリ デバイスのRTRを決定する最も重要な要因は、コア メモリの技術です。例えば、シンクロナスDRAM (SDRAM)、レイテンシー短縮済みDRAM (RLDRAM)、またはQDR<sup>®</sup>シンクロナスSRAM (QDR SRAM)です。SDRAMとRLDRAMは真にランダムアクセスであり、主にランダムサイクル時間の待ち時間 (tRC) により制限されます。最大RTRは、およそtRCの逆数 (1/tRC) です。SDRAMのtRCは、過去10年間で実質的には進化していません (今後も進化することが期待されません)。そして、tRCは、21MT/s RTRに対応する値である~48nsです。他のDRAMベースのメモリデバイスは、容量を犠牲にしてtRCを改善するよう設計されています。例えば、RLDRAM 3は、8nsのtRCを持ち、その値は125-MT/s RTRに対応しています。QDR SRAMは、特にランダムアクセス用に最適化されています。QDR-II+ Xtremeファミリ (tQDR-IVの前身) は、900-MT/sのRTRを実現しています。

## メモリデバイス構造 (バンク)

帯域幅に関するDRAMのtRCの主要な制限を克服するために、DRAMデバイスのプロバイダは、デバイスをバンクに分割し、その各々のバンクを独立してアドレス指定することができる「バンキング」の概念を導入しています。如何なるバンク内のRTRもメモリ技術 (前に説明したように) によって未だに制限されています。各バンクは独立してアクセスできるので、デバイスの全トランザクションレート (メモリインタフェースの制約の影響を除く) は、任意の時間でアクティブにできるバンクの数にバンク内のRTRを掛けたものです。

DRAMベースのデバイスはバンクアクセスに大きな制限を課し、バンクへのメモリの分割は、デバイス全体の総トランザクション レートは純粋にランダム (アクセスは各バンク内でのみランダムです) ではないことを意味します。純粋にランダムな読み込みアクセスを有効にするためには、複数のバンクに同じデータを複製することが必要です。その結果、複製因数 (データが複製された回数) によってread RTRが増加します。

しかし、複製も重要なトレードオフになります。複製は、複製因子によってデバイスの有効容量を低減します。また、複製部分はランダム読み込みと書き込みアクセスを必要とするアプリケーションでは動作しません。これらの場合、任意の書き込みアクセスは、複数のトランザクションを必要とするデータコピーの各々に対して繰り返す必要があり、これは複製の目的に反します。「ピンポン」方式を採用して、読み込みと書き込み用に2x RTRを可能にすることができます。この方式では、書き込みが2個のバンクのどちらか一方に直接アクセスし、2個のバンクの中にどれが最新のデータを格納しているかを指定するためにマップが維持されます。この方式では、各バンク内のデータ セットが正確なコピーではないため、本当の複製とならないことに注意してください。また、この方式では、4x RTR (4バンクを使用する場合) に調整することはできません。理由は、各書き込みの後に少なくとも2個のバンクを更新する必要があるからです (これは、複数のトランザクションを意味しており、複製の目的に沿わないからです)。

## メモリ インターフェース

メモリ デバイスのRTRの最後の決定要因は、インターフェースの能力です。これがメモリの全能力を発揮させるとは限りません。例えば、RLDRAM 3を検討してみましょう。このデバイスは、16個のバンクに設定されており、2000MT/s (16バンク\*125MT/s) の理論上の総トランザクション レートを可能にします。しかし、インターフェースは、最大10億6600万回のランダム読み込みか書き込みコマンド (1066MT/s) に対応した1個のポートを使用しています。従って、デバイスのRTRはインターフェースにより制限されます。

本書では、RTRを説明するために、次の用語を使用しています: 「RTR per bank」(バンク毎のRTR) はメモリ技術のRTR、「interface RTR」(インターフェースのRTR) はインターフェースで許可されたトランザクションレート、「memory device RTR」(メモリデバイスのRTR) はランダム トランザクション レートの合計です。メモリ デバイスのRTRは、インターフェースのRTRまたはアクティブなバンク数にバンク毎のRTRを掛け合わせたものの、どちらか低い方となります。

## QDR-IV—最高性能ネットワーク メモリについて



QDR-IVは、QDRコンソーシアムにより定義され、RTRに完全に最適化された最初のメモリであり、FIBルックアップ テーブル、統計、ステート、スケジューリング、分類、およびヘッドテイル キャッシュの最も要求の厳しいライン カード機能に向いています。

### QDR-IV RTR

QDR-IVファミリは、QDR-IV HP(高性能)とQDR-IV XP(「Xtreme」性能)という2つの重要なデバイス オプションがあります。QDR-IV HPは、最大667MHzの周波数で動作し、最大1334MT/sのRTRを可能にします(2個のポートはそれぞれ667MT/sのアドレス レートに対応します)。QDR-IV XPは、バンキングを備えており、最大1066MHzの周波数で動作し、最大2132MT/sのRTRを可能にします。これは、今まで最高性能の標準ベースのメモリ ソリューションであったQDR-II+ Xtremeよりも2.4倍改善されていることを示しています。

QDR-IVは、以下の節で説明するようにシステム設計を最適化および簡略化する機能にも対応しています。

### 双方向ポート

従来のQDRメモリは、2個の一方方向ポート(1個の読み込みポートと1個の書き込みポート)をサポートしています。これは、統計/ステート、スケジューリング、パケット バッファリングなど、読み込みと書き込みのバランスを取ったネットワーク機能に適しています。QDR-IVは、2個の双方向ポートを備えており、両方のポートを同時に読み込み動作に使用することができます。これにより、ルックアップのような、読み込みにより支配された機能用にRTRが2倍増加します。

### ECC

QDR-IVは、ECCを内蔵した最初のQDR SRAMであり、ソフト エラー レートをMビットあたり、0.01 FIT(故障率)以下に低減します。

### 信号堅固性の機能

QDR-IVは、アドレスおよびデータ バス反転に対応しており、同時スイッチング出力ノイズを低減することで消費電力を減少させ、信号堅固性を向上させます。QDR-IVは、信号堅固性を向上させるダイ上の終端とアドレス パリティ検出にも対応しています。

### クロッキング

QDR-IVは、コマンドとアドレス クロックの差動対(CKとCK#)、および読み込みデータ用のフリーランニング差動出力クロック(QKとQK#)(デバイスによって、9または18 I/O毎に1対)を備えています。QDR-IVは、書き込みデータ用に差動入力クロック(DKとDK#、9または18 I/O毎に1対)を追加します。入力クロックの追加により、読み込み動作のデータ有効ウィンドウが広がります。

### バンキング(QDR-IV XPのみ)

従来のQDRメモリはバンキングに対応していません。QDR-IVはそれらとは違って、8個のバンクをサポートしています。QDR-IVバンキングは、DRAMベースのソリューション(DDR3/4 SDRAMまたはRLDRAM 3)よりもはるかに制約が弱いです。クロック サイクル毎に、QDR-IVの8バンクのいずれかがアクセスされます。ポートAはクロック サイクル毎にどのバンクにもアクセスできますが、ポートBは、同じサイクルでポートAがアクセスしたバンク以外のバンクのみにアクセスすることができます。このため、QDR-IV XPは、検索またはカウンタのデータベースを2つの異なったグループに分けるアルゴリズムに適しており、アクセスを逃さずにそのようなアルゴリズムに最大2132MT/sを提供しています。

### フレキシブルなI/O信号方式

QDR-IVは、HSTL/SSTL(1.2V/1.25V)と疑似オープン ドレイン(POD)(1.1V/1.2V)信号方式をサポートしています。その他のメモリ デバイスもこれらの信号方式標準をサポートしています(DDR4 SDRAMはPOD、DDR3 SDRAMはSSTL、RLDRAM 3はHSTL)。

### QDR-IVと他のメモリとの性能比較

図3では、他のメモリに対するQDR-IVの優れたRTR性能を数値で表しています。QDR-IVのRTR性能は、他のメモリよりも少なくとも2倍高いです。図に示されたように、150MP/s(100Gbit/s)と300MP/s(200Gbit/s)両方のパケット レートで要求された最小RTRを満たすのは、QDR-IVのみです。

Comparison of Memory Device RTR for 300 MP/s Line Card Packet Rates

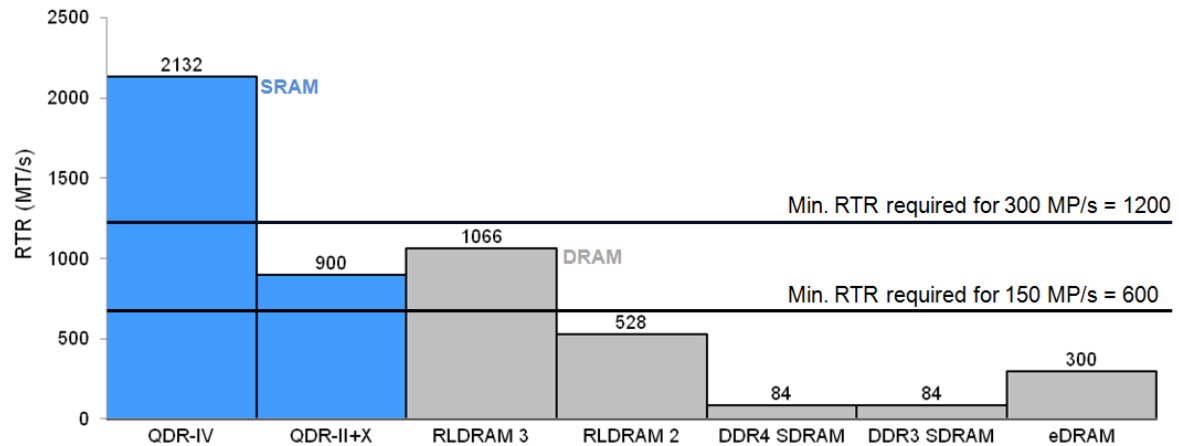


図 3. QDR-IV 性能比較

### 結論

2132MT/sの最大RTRに対応できるQDR-IVは、市場最高性能の標準ベースのメモリ ソリューションです。その高いRTRおよびデュアル双方向ポート、ECC、バス反転、ODT、アドレス パリティなどの異なる機能により、QDR-IVはネットワーク システムに最適なソリューションとなります。QDR-IVの利点は、高性能コンピューティングや画像処理などの高いRTR性能と信号堅固性を必要とした他のシステムにも適用することができます。

Cypress Semiconductor  
198 Champion Court  
San Jose, CA 95134-1709  
Phone: 408-943-2600  
Fax: 408-943-4730  
<http://www.cypress.com>

© Cypress Semiconductor Corporation, 2014. 本文書に記載される情報は、予告なく変更される場合があります。Cypress Semiconductor Corporation (サイプレス セミコンダクタ社) は、サイプレス製品に組み込まれた回路以外のいかなる回路を使用することに対して一切の責任を負いません。サイプレス セミコンダクタ社は、特許またはその他の権利に基づくライセンスを譲渡することも、または含意することはありません。サイプレス製品は、サイプレスとの書面による合意に基づくものでない限り、医療、生命維持、救命、重要な管理、または安全の用途のために使用することを保証するものではなく、また使用することを意図したものではありません。さらにサイプレスは、誤動作や故障によって使用者に重大な傷害をもたらすことが合理的に予想される生命維持システムの重要なコンポーネントとしてサイプレス製品を使用することを許可していません。生命維持システムの用途にサイプレス製品を供することは、製造者がそのような使用におけるあらゆるリスクを負うことを意味し、その結果サイプレスはあらゆる責任を免除されることを意味します。

QDR は、サイプレス セミコンダクタ社の登録商標です。本書で言及するその他全ての商標または登録商標は、各社の所有物です。

このソースコード(ソフトウェアおよび/またはファームウェア)はサイプレス セミコンダクタ社(以下「サイプレス」)が所有し、全世界の特許権保護(米国およびその他の国)、米国の著作権法ならびに国際協定の条項により保護され、かつそれらに従います。サイプレスが本書面によりライセンシーに付与するライセンスは、個人的、非独占的かつ譲渡不能のライセンスであり、適用される契約で指定されたサイプレスの集積回路と併用されるライセンシーの製品のみをサポートするカスタム ソフトウェアおよび/またはカスタム ファームウェアを作成する目的に限って、サイプレスのソースコードの派生著作物をコピー、使用、変更して作成するためのライセンス、ならびにサイプレスのソースコードおよび派生著作物をコンパイルするためのライセンスです。上記で指定された場合を除き、サイプレスの書面による明示的な許可なくして本ソースコードを複製、変更、変換、コンパイル、または表示することは全て禁止します。

免責条項: サイプレスは、明示的または黙示的を問わず、本資料に関するいかなる種類の保証も行いません。これには、商品性または特定目的への適合性の黙示的な保証が含まれますが、これに限定されません。サイプレスは、本文書に記載される資料に対して今後予告なく変更を加える権利を留保します。サイプレスは、本文書に記載されるいかなる製品または回路を適用または使用したことによって生ずるいかなる責任も負いません。サイプレスは、誤動作や故障によって使用者に重大な傷害をもたらすことが合理的に予想される生命維持システムの重要なコンポーネントとしてサイプレス製品を使用することを許可していません。生命維持システムの用途にサイプレス製品を供することは、製造者がそのような使用におけるあらゆるリスクを負うことを意味し、その結果サイプレスはあらゆる責任を免除されることを意味します。

ソフトウェアの使用は、適用されるサイプレス ソフトウェア ライセンス契約によって制限され、かつ制約される場合があります。