



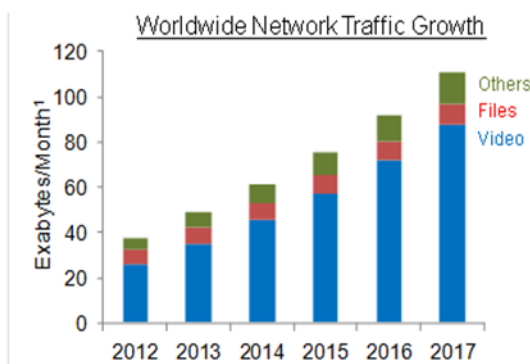
## QDR®-IV: 领先的网络系统中下一代高性能标准存储器

### 摘要

根据网络和计算设备供应商的规模，供应商所提供的各种方案要满足因特网以及存储特性，以支持系统性能提升的需要。所有路由器、转换器和高性能计算系统都需要更高速率的存储器来协助处理。因此，随机事务处理速率（RTR）——存储器可以提供的随机访问（读或写）次数，成为存储器的关键度量标准。本白皮书介绍的是赛普拉斯的 QDR®-IV 如何提供最高的 RTR 和标准的方案，以满足这些系统要求。

### 网络系统的演进

据预测，随着互联网连接设备和宽带视频服务的发展，在接下来的四年中，全球网络流量将增加三倍。到2017年，36亿互联网用户将使用190亿结点的网络连接。因此，网络设备供应商正在部署转换器和路由器线卡，为了每一代产品更新后使性能增加两倍（该性能以线卡速率（Gbit/s）或数据包速率（MP/s）为测量单位）。图1和图2描述了这种发展趋势。请注意，数据包速率与线卡速率成正比：数据包速率 = 线卡速率 / （最小数据包大小 + 数据包之间的间隙）。最小的数据包大小为64个字节，其中帧之间的间隔为20个字节。因此，100-Gbit/s线卡的数据包速率大概为150 MP/s。



<sup>1</sup>网络流量的测量单位为艾字节（10<sup>18</sup>字节）/月

图 1. 全球网络流量的增长趋势

来源: Cisco 视觉网络指数 (VNI) 预测 (2012–2017)

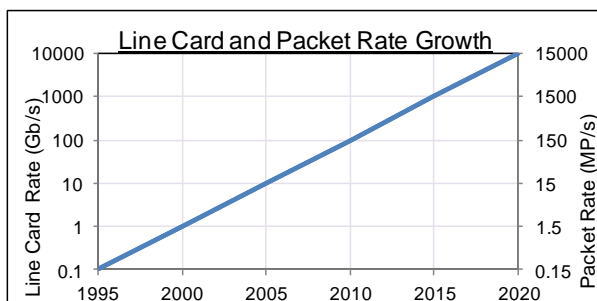


图 2. 线卡和数据包速率的增长趋势

来源: 以太网技术峰会

## 随机事务处理速率

随机事务处理速率（RTR）是指在给定的时间内，存储器可支持的随机访问（读或写）的次数。它的测量单位为事务处理/秒（例如，MT/s或GT/s）。网络系统需要的RTR由数据包速率和数据包要求的随机存储器访问次数确定，该数值会根据存储器支持的线卡功能而改变。

## 线卡存储器功能

当使用线卡上存储器的六个关键功能时，RTR的重要性是最明显的。如上所述，这些功能决定了整个系统的性能。主性能的瓶颈并不是切换速度而是通过检查数据包包头确定其类型和目标，查找转发表中的布线，分类数据包来进行处理以及更新各种统计和状态等操作所使用的时间。

## 数据包缓存

路由器上的每个线卡都有一个入口缓冲区和一个出口缓冲区。当数据包等待处理和调度到切换器结构时，入口缓冲区会保持输入物理链接中的数据包（这些数据包已经被分类和转发）。出口缓冲区存储切换器结构中的数据包，这些数据包将被处理并传输给输出物理链接。数据包速率越高，需要的数据包缓冲区的大小也越大，因此数据包缓冲区存储器主要由容量决定。如果大容量存储器没有所需要的RTR以匹配上数据包的速率，系统制造商通常会设计一个分级结构，其中使用一个“head-tail cache”来存储数据包缓冲区的包头和包尾。对于每个数据包，“head-tail cache”需要四到六个数据操作。

## 转发信息库查找表

转发信息库（FIB）查找表将出口地址存储在与目标IP地址相对应的切换器/路由器上。寻找是一个重复过程，包括对存储器进行的各种访问。可以在网络堆栈的第二层和第三层上执行查找过程。第二层查找通常通过散列算法实现。这些算法寻找与搜索索引相匹配的结果，而且，对于每个数据包，它们通常需要进行四次随机访问。对于每个数据包，使用最长前缀匹配的第三层查找需要8次随机访问。

## 数据包的分类或访问控制列表

在该处理阶段，处理器检查输入数据包的特性（如源地址、目标地址、源端口、目标端口和协议），并确定是否允许数据包通过。一旦数据包被分类或监管，它将暂时被存储在缓冲区子系统内，以进行处理。对于每个数据包，访问控制列表（ACL）查找的算法方法也需要10次以上的随机访问。

## 处理

处理过程包括确定将数据包发送到切换器结构的时间。该过程取决于数据包的目标和所需要的服务质量（QoS）或服务等级（CoS）。这些数据包被分为几种类型，每一类都与分层服务供应商（服务供应商的收入段）相关联。对于每个数据包，处理应用至少要求进行一个读数据操作和一个写数据操作，因此RTR非常重要。

## 统计与状态

路由器保持每个数据包和每个数据流（相关数据包的流）基础的统计量。该操作由计数器实现（这些计数器用于存储前缀、数据流和数据包类型等信息）。更新一个计数器时需要进行一个读取-修改-写入操作（两个随机存储器事务处理）。RTR要求同数据包速率以及所需要的计数器更新数量成正比。线卡中的相同存储器可以使用于统计和状态。

## RTR — 关键度量

所有功能都需要高RTR，而其他存储器度量不太重要。对于任意的数据宽度，随着RTR的增加，带宽也会同比增加。请注意，随着数据宽度增加，带宽也要增加。容量是数据包缓冲区的关键属性，但“head-tail cache”指出该功能还需要高RTR。

## 通过数据包速率和存储器功能确定RTR的要求

表1显示的是100 Gbit/s（150-MP/s）和200 Gbit/s（300-MP/s）线卡的RTR要求。

存储器功能	每个数据包的数据操作	100 Gbit/s 150 MP/s	200 Gbit/s 300 MP/s
“head-tail cache” /程序	4	600 MT/s	1200 MT/s
查找	4–8	600–1200 MT/s	1200–2400 MT/s
统计/状态	8–6（4–8计数器）	1200–2400 MT/s	2400–4800 MT/s
数据包缓存	4–6	600–900 MT/s	1200–1800 MT/s

表 1. 线卡的 RTR 要求

### RTR: 不受网络限制

RTR性能的值不受网络系统的限制。在高性能计数、通用服务器和图像处理应用中，RTR是关键的要求。高性能的计算系统以及解决复杂问题的计算集群和存储资源都要求存储器具有高RTR，以处理数据流中的数据。通用服务器共同使用最后一级缓存的“多核”和“众多核”处理器，它需要高RTR存储器以满足CPU的要求。另外，高性能的图像处理系统实现数字信号处理算法（这些算法要求对图像数据进行快速随机访问），因此它们也受益于高RTR存储器。

### 选择具有高RTR的存储器

设计师必须了解各种存储器类型所提供的RTR的区别，以优化系统性能。所有存储器的RTR都由内核存储器技术、存储器的结构（分组）和存储器接口确定。

#### 存储器技术

影响存储器RTR的最重要因素是内核存储器技术，例如，同步DRAM（SDRAM）、较小延迟DRAM（RLDRAM）或QDR<sup>®</sup>同步SRAM（QDR SRAM）。在SDRAM和RLDRAM中，实时随机访问主要由随机周期的时间延迟（tRC）限制。最大的RTR大约等于tRC（1/tRC）的倒数。十年以来，SDRAM tRC没有任何显著发展（预期将来也没有进展），仍然维持在~48 ns，相当于21 MT/s RTR。基于DRAM的其他存储器会降低tRC时间，同时需要提高容量。例如，RLDRAM 3的tRC为8 ns，相应于125-MT/s RTR。QDR SRAM被特殊优化，以用于随机访问。QDR-II+ Xtreme系列（是QDR-IV的一代）的RTR可达900-MT/s。

#### 存储器的结构（分组）

为了消除DRAM在带宽方面的基本tRC限制，DRAM器件供应商使用了“分组”概念，表示将器件分为多个组，每个组可以被独立访问。所有组中的RTR仍由存储器技术控制（如上所述）；但由于可以对每一组进行独立访问，因此器件的总事务处理速率（包括存储器接口限制的影响）是某时间内可用的组数量乘以某个组中的RTR。

基于DRAM的器件对组访问要求严格的限制，将存储器分为各组也表示器件的总事务处理速率不完全是随意指定的（仅在每个组中进行随机访问）。在纯粹的随机读取访问时会在多个组中复制相同的数据，这样通过复制因素（复制数据的时间量）可提高读操作的RTR。

然而，复制也会引起显著的权衡问题。复制因素会降低器件的有效容量。并且，在需要进行随机读和写访问的应用中不能使用复制。在这些情况下，需要针对每个数据拷贝重复进行已给的写访问，该过程需要执行多个数据操作，因此会抵消复制的效果。通过实现“乒乓”方法可以使读和写操作的RTR提高一倍，其中可以直接对两个组中的一组进行写访问，并且保存了一个映射用以指出需要保存最新数据的组。请注意，由于设置在每一组中的数据不是精确的副本，因此该方法的结果也不是真正的复

制。另外，该方法也不能将RTR提高到4倍（使用四个组），因为这样会在每个写操作后至少要更新两个组（再次留意该过程需要执行多个数据操作，因此会抵消复制的意义）。

### 存储器接口

影响存储器中RTR的最后因素是接口，该接口不用展示存储器的所有特性。例如，RLDRAM 3器件配置为16个分组，理论上能够实现2000 MT/s的总事务处理速率（16个组\* 125 MT/s）。然而，该接口使用了一个支持最大为1066百万条随机读或写指令（1066 MT/s）的单端口。因此，该器件的RTR由该接口限制。

本节使用下列术语介绍RTR：“每组的RTR”表示存储器技术中的RTR；“接口RTR”表示接口允许的事务处理速率；“存储器RTR”表示总随机事务处理速率。存储器的RTR是下面两组中较小的那个：1.接口RTR；2.每组的RTR乘以有效组数量得到的值。

### QDR-IV：性能最高的网络存储器

QDR-IV（在QDR联盟中定义）是第一个优化RTR的存储器，针对要求最为严苛的线卡功能使用：FIB查找表、统计、处理、分类和“head-tail cache”。

#### QDR-IV RTR

QDR-IV系列提供了两个主要器件选项：QDR-IV HP（高性能）和QDR-IV XP（“Xtreme”性能）。QDR-IV HP的最大工作频率为667 MHz，最大的RTR为1334 MT/s（每个端口的地址速率为667 MT/s）。QDR-IV XP（包括分组）的最大工作频率为1066 MHz，最大RTR为2132 MT/s。上述信息表示QDR-II+ Xtreme的性能已提高了2.4倍，作为当前最高性能的标准存储器方案。

QDR-IV还拥有其他特性，包括优化和简单化系统设计，如下面内容所述。

#### 双向端口

传统情况下，QDR存储器支持两个单向端口：一个读端口和一个写端口。它服务于平衡读写的网络功能如统计/状态、处理和数据包缓冲区。QDR-IV包含两个双向端口，这样可以使两个端口同时进行读操作。因此，RTR增加了两倍的读优势功能（如查找）。

#### ECC

QDR-IV是包含了ECC的第一个QDR SRAM，可以使软错误速率下降到0.01 FIT/Mbit。

#### 信号完整性

QDR-IV支持地址和数据总线反转，通过减少同时切换输出的噪声，降低功耗并提高信号的完整性。QDR-IV还包含片内终端和地址奇偶校验检测，用于增强信号完整性。

#### 时钟

QDR-IV包括一对差分指令和地址时钟（CK和CK#），以及自由运行的差分输出时钟（QK和QK#），该时钟用于读取数据（每9个或每18个I/O使用一对，随器件而异）。QDR-IV添加了用于写数据的差分输入时钟（DK和DK#，每9个或每18个I/O使用一对）。这样会扩大读操作的数据有效窗口。

#### 分组（仅针对QDR-IV XP）

传统情况下，QDR存储器不支持任何分组功能。QDR-IV通过支持8个分组脱离了该趋势。与DRAM的解决方案（DDR3/4 SDRAM或RLDRAM 3）中的分组相比，QDR-IV分组受限更少。每个时钟周期可以对QDR-IV的8个分组其中一个进行访问。每个时钟周期端口A会访问任意分组，但在同一个时钟周期内，端口B必须与端口A访问的分组不同。因此，QDR-IV XP很适合实现将搜索或计数器数据库拆分为两个不同组的算法，并可以为这种算法提供2132 MT/s而不会丢失任何访问。

### 灵活的I/O信令

QDR-IV支持HSTL/SSTL (1.2 V/1.25 V) 和伪开漏 (POD) (1.1 V/1.2 V) 的信令。其他存储器也支持这些信令标准 (DDR4 SDRAM使用POD, DDR3 SDRAM使用SSTL, RDRAM 3使用HSTL)。

### QDR-IV与其他存储器的性能比较

图3显示的是QDR-IV与其他存储器的RTR性能比较。QDR-IV可提供的RTR性能是其他存储器的两倍。该图表还指出满足最小RTR的唯一QDR-IV (两个数据包速率150-MP/s (100-Gbit/s) 和300-MP/s (200-Gbit/s) 都需要该RTR)。

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

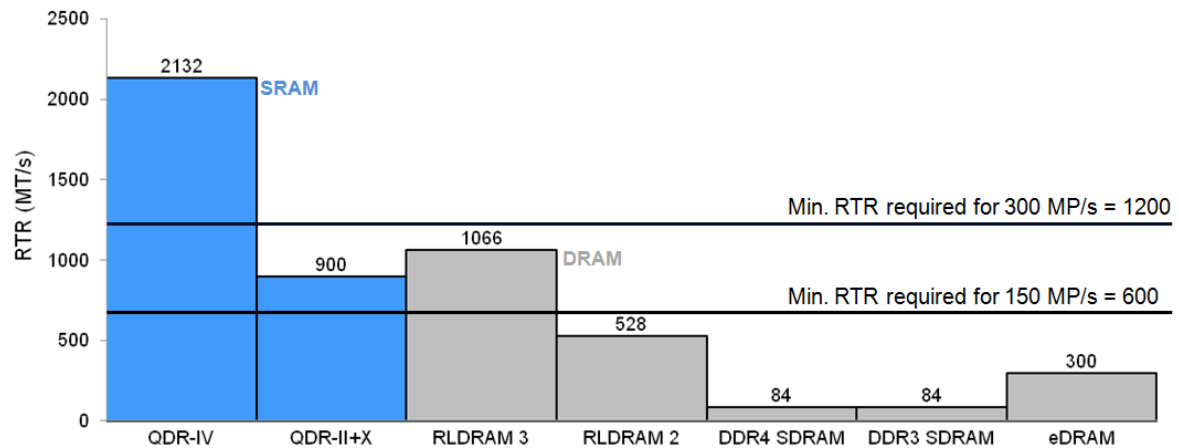


图 3. QDR-IV 性能比较

### 结论

QDR-IV (RTR的最大速度为2132 MT/s) 是市场上具有最高性能的标准存储器方案。它的高RTR与独特特性 (如对称双向端口、ECC、总线反转、ODT和地址奇偶校验) 结合时, 将可以为网络系统提供最佳的解决方案。QDR-IV的优势还适用于需要高RTR性能和信号完整性 (如高性能计算和图像处理) 的其他系统。

赛普拉斯半导体  
198 Champion Court  
San Jose, CA 95134-1709  
电话: 408-943-2600  
传真: 408-943-4730  
<http://www.cypress.com>

©赛普拉斯半导体公司, 2014。此处, 所包含的信息可能会随时更改, 恕不另行通知。除赛普拉斯产品内嵌的电路以外, 赛普拉斯半导体公司不对任何其他电路的使用承担任何责任。也不根据专利或其他权利以明示或暗示的方式授予任何许可。除非与赛普拉斯签订明确的书面协议, 否则赛普拉斯产品不保证能够用于或适用于医疗、生命支持、救生、关键控制或安全应用领域。此外, 对于合理预计会发生运行异常和故障并对用户造成严重伤害的生命支持系统, 赛普拉斯将不批准将其产品用作此类系统的关键组件。若将赛普拉斯产品用于生命支持系统, 则表示制造商将承担因此类使用而招致的所有风险, 并确保赛普拉斯免于因此而受到任何指控。

QDR 是赛普拉斯半导体公司的注册商标。此处引用的所有其他商标或注册商标归其各自所有者所有。

该源代码(软件和/或固件)均归赛普拉斯半导体公司(赛普拉斯)所有, 并受全球专利法规(美国和美国以外的专利法规)、美国版权法以及国际条约规定的保护和约束。赛普拉斯据此向获许可者授予适用于个人的、非独占性、不可转让的许可, 用以复制、使用、修改、创建赛普拉斯源代码的派生作品、编译赛普拉斯源代码和派生作品, 并且其目的只能是创建自定义软件和/或固件, 以支持获许可者仅将其获得的产品依照适用协议规定的方式与赛普拉斯集成电路配合使用。除上述指定用途外, 未经赛普拉斯的明确书面许可, 不得对此类源代码进行任何复制、修改、转换、编译或演示。

免责声明: 赛普拉斯不针对该材料提供任何类型的明示或暗示保证, 包括(但不限于)针对特定用途的适销性和适用性的暗示保证。赛普拉斯保留在不另行通知的情况下对此处所述材料进行更改的权利。赛普拉斯不对此处所述之任何产品或电路的应用或使用承担任何责任。对于合理预计可能发生运转异常和故障, 并对用户造成严重伤害的生命支持系统, 赛普拉斯不授权将其产品用作此类系统的关键组件。若将赛普拉斯产品用于生命支持系统, 则表示制造商将承担因此类使用而导致的所有风险, 并确保赛普拉斯免于因此而受到任何指控。

产品使用可能受适用的赛普拉斯软件许可协议限制。