

无线领域的出色表现

**Xilinx                    20nm**  
**UltraScale 架构助推**  
**无线电应用发展**

新一代 5G 系统的设计十分复杂，而 UltraScale 器件内置的相关功能，能让这项工作变得更加简单。

Michel Pecot

赛灵思公司无线系统架构师

[Michel.pecot@xilinx.com](mailto:Michel.pecot@xilinx.com)

即将到来的 5G 无线通信系统似乎需要支持比目前使用的 4G 系统更大的带宽(200 MHz 及以上)，以及大型的天线阵列，以实现更高的载波频率，从而有可能构建小得多的天线元。这些所谓的大规模多输入多输出 (MIMO) 应用连同更加迫切的延迟需求可将设计复杂度提高一个数量级。

去年年底，赛灵思宣布推出 20nm UltraScale™ 系列，目前第一款器件已在运输途中[1,2,3]。这项新技术与之前的 28nm 7 系列相比具有众多优势，尤其是在无线通信方面。确实，这款新型芯片与赛灵思 Vivado®设计套件[4,5]工具的结合完美适用于新一代无线电应用等高性能信号处理设计。

我们来看看针对这类设计 UltraScale 器件有哪些优势，着重观察架构方面——尤其是当实现一些用于无线数字前端 (DFE) 应用的最常见功能时，这种增强功能会给 DSP48 Slice 和 Block RAM 带来哪些优势。与 7 系列相比，UltraScale 系列可提供更加密集的布线和时钟资源，能够实现更高的设备利用率，尤其针对高速设计。但是，这些特性通常不能对设计架构产生直接影响，因此我们在这里不做讨论。

## UltraScale 架构增强功能简介

UltraScale 20nm 架构不仅能随着几何节点的迁移提高集成能力，提升架构性能以及降低功率消耗，还包括一些显著增强的全新功能，可直接支持 DFE 应用。而这些功能对 UltraScale Kintex®器件而言尤其重要，为此赛灵思已根据此类设计的需求进行了重大调整。

首先，这些器件包含多达 5,520 个 DSP48 Slice，这几乎是 7 系列 FPGA 的最大数量 (1,920) 的三倍 (Zynq® - 700 All Programmable SoC 为 2,020)，因此，它可以实现高集成度。比如，一个中型 UltraScale FPGA 的瞬时带宽就能达到 80 到 100 MHz，您可以利用这一带宽实现完整的 8Tx/8Rx DFE 系统，而在 7 系列架构中，必须使用双芯片解决方案才能实现，因为每个芯片只能有效支持一个 4x4 系统。如欲了解有关这类

设计各项功能的详细信息，敬请参阅赛灵思白皮书 WP445 “采用赛灵思 All Programmable FPGAs 以及 SoC 实现高速无线电设计” [6]。

## SerDes 可在最低速度等级器件上支持 12.5 Gbps 流量，实现最大

### JESD204B 接口连接速度。

受远辐射被动降温的散热限制，将复杂设计集成到单个器件要求大幅降低功耗，以散发热量。UltraScale 系列在提供这项功能时，其静态功耗比同等规模的 7 系列器件低 10%-15%，动态功耗比类似设计低 20%-25%。此外，赛灵思还大幅降低了 UltraScale 产品线的 SerDes 功耗。

此外，它还存在性能优势。最低速度等级 UltraScale 器件支持时钟速率高于 500MHz 的设计，而 7 系列器件则要求达到中速等级。然而，即使这样，Block RAM 从计时角度来看要求仍然严苛，并且必须选择 WRITE\_FIRST 或 NO\_CHANGE 模式以达到这种性能。不能使用 READ\_FIRST，因为它的限值在 470MHz 左右，而另外两种模式可达到 530MHz。无论何时 NO\_CHANGE 总是您的最佳选择，因为它同时还能实现最低功耗。

同样，SerDes 可在最低 UltraScale 速度等级上支持高达 12.5 Gbps 流量，从而实现最大 JESD204B 接口连接速度，其应很快可在大多数 DAC 和 ADC 上实现。同样，最低的 UltraScale 速度等级还可支持两个最高 CPRI 等级（7 级和 8 级，其流量分别为 9.8304 和 10.1376Gbps）以及 10GE 接口，通常用于 DFE 系统。

此外，UltraScale Kintex 资源组合更适合无线电应用，它能够实现逻辑资源的最佳用法。该 DSP 逻辑比尤其符合 DFE 设计的典型需求。确切地说，UltraScale Kintex 器件拥有每千查找表（LUTs）8-8.5 个 DSP48 Slice，而 7 系列器件只有大概 6 个。

赛灵思还大幅增加了 UltraScale 架构的时钟和布线资源。这项增加提高了器件利用率，尤其针对高时钟速率设计。实际上，这样做减少了布线拥塞，设计人员可以实现更好的设计封装和 LUT 利用率，尤其是使 LUT/SRL 压缩变得更为高效。用户可以利用这项有意思的架构特性更好地打包设计，从而优化资源利用率以及动态功耗，其中相关逻辑的动态功耗下降系数可达 1.7。LUT/SRL 的压缩原理包括采用 LUT6 的两个输出在单个 LUT 内打包两个不同函数。这样，如果两个 LUT5 共享相同的输出或内存读取/写入地址，您可以将实现逻辑函数或内存的两个 LUT5 打包到一个 LUT6 中。同样，也可以将两个 SRL16 打包到一个 LUT6 中。

该特性对于数字无线电设计非常实用，该设计通常将共享同一地址的多个小内存（例如储存滤波系数的 ROM）和很多短延迟线（小于 16 个周期）集成到按时间排列的不同信号路径中。

数据复用功能，尤其是双输入复用器，也将受益于这项特性。但是要想获得较高的时钟速率，必须谨慎使用 LUT/SRL 压缩。首先，必须用连接到 O6/O5 LUT 输出的两个触发器以避免发生任何时序问题。基于相同的原因，建议仅对相关逻辑使用这项功能，该策略还能起到限制布线拥塞的作用。

在 UltraScale 器件中，时钟架构和可配置逻辑块（CLB）均有助于更好地利用器件。尽管 CLB 仍然以 7 系列架构为基础，但现在每个 CLB 拥有单个 Slice（而非两个），其中集成了 8 个 6 输入 LUT 和 16 个触发器。因此进位链有 8 位长，且提供更宽的输出复用器。另外，赛灵思还增加了控制类资源（也就是时钟，时钟使能和复位信号均共享于 CLB 中的存储组件）。

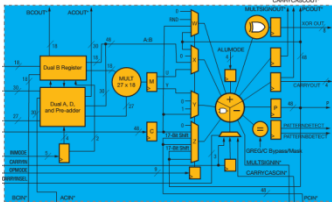
然而，从本质上来讲，还是 DSP48 Slice 和 Block RAM 的改进对无线电设计架构的影响最大。我们来详细了解一下。

### UltraScale DSP48

#### Slice 架构的优势

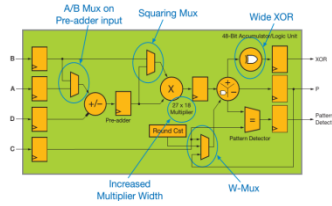
图 1 给出了 UltraScale DSP48Slice(DSP48E2)的视图。上面的原理图（图表“a”）显示了详细架构，下面的原理图（“b”）强调了与 7 系

列 Slice(DSP48E1)相比增强的功能。



\* 这些信号是专用于 DSP48E2 列的内部布线路径,其不可通过通用布线资源进行访问。

(a) 详细的 DSP48E2 架构



预加法器输入端上的 A/B 多路复用器  
 方形多路复用器  
 宽 XOR  
 增加的乘法器宽度  
 W-多路复用器

(b) DSP48E2 高层次功能视图

图 1 – UltraScale DSP48 Slice 架构

赛灵思用户指南 UG579 全面介绍了 DSP48E2 功能[7]。UltraScale 架构的主要增强功能为:

- 赛灵思将乘法器的宽度从  $25 \times 18$  增加到  $27 \times 18$ , 预加法器宽度也相应增加到 27 位。
- 您可以选择预加法器

输入端为 A 或 B, 输出端集成了一些多路复用器逻辑, 从而允许在任意乘法器输入端上 (27 位或 18 位输入) 馈送  $D \pm A$  或  $D \pm B$ 。

- 预加法器输出端可馈送两个乘法器输入 (在 18 位输入端上有适当的 MSB 截断), 因此允许计算高达 18 位数据的  $(D \pm A)^2$  或  $(D \pm B)^2$ 。
- 通过额外的 W-mux 多路复用器在算术逻辑单元 (ALU) 中添加了第四个操作数, 可将其看作输入端 C、P 或一个常量值 (在 FPGA 配置时定义)。这样, 使用乘法器时便可以执行一个三输入操作, 如  $A * B + C + P$  或  $A * B + P + PCIN$ 。值得注意的是, 只能在 ALU 中添加 W-mux 输出 (不允许减少)。
- 赛灵思集成了其它逻辑, 从而可在 X、Y 或 Z 多路复用器输出端中的任意两个之间执行 96 位宽异或。实际上此处可提供四个不同模式,  $1 \times 96$  位、 $2 \times 48$  位、 $4 \times 24$  位或  $8 \times 12$  位异或操作。

将乘法器尺寸从  $25 \times 18$  扩大到  $27 \times 18$ , 这对 DSP48

Slice 芯片面积的影响极小, 但会显著提高对浮点运算的支持。首先, 需要指出的是, DSP48E2 可有效支持高达  $28 \times 18$  位或  $27 \times 19$  位符号乘法运算。这是通过用输入端 C 处理额外的比特位来实现, 如图 2, 显示了 28 位操作数 X 和 18 位操作数 Y 之间的乘法运算。

46 位输出端的 45 个最高有效位 (MSBs) 的计算方式为:  $Z[45:1] = X[27:1] * Y[17:0] + X[0] * Y[17:1]$

X 的 27 个最高有效位和 Y 的 18 位可直接馈送到 DSP48E2 乘法器输入端, 而  $X[0] * Y[17:1]$  源自外部 17 位 AND 运算符, 并在一个流水线步骤后被发送到输入端 C 以匹配 DSP48E2 延迟。事实上 AND 运算符可通过由  $X[0]$  控制的复位引脚将  $Y[17:1]$  直接馈送到寄存器中。同样, 外部 1 位 AND 运算符和用于实现延迟平衡的三时钟周期延迟可用来计算 Z,  $Z[0]$  的 LSB。

因此您可以执行具有单个 DSP48E2 Slice 和 18 个 LUT/触发器对的  $28 \times 18$  位乘法器。这同样适用于使用其它 27 个 LUT/触发器对的  $27 \times 19$  位乘法器。这两种情况下, 均可通过 W-mux 支持运算结果的收敛舍入。

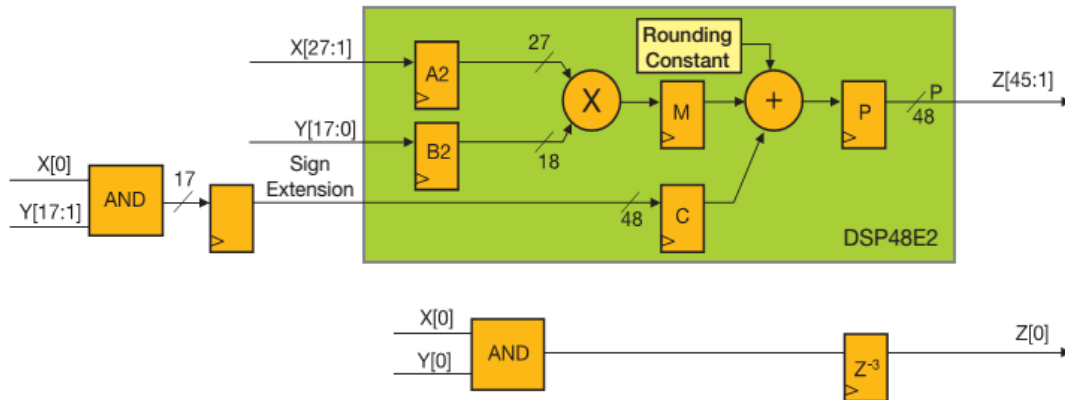


图 2 – 具备输出收敛舍入功能的 28×18 位符号乘法运算

双精度浮点乘法运算涉及两个运算符的 53 位非符号尾数的整数乘积。尽管双精度浮点计数法中储存有一个 52 位数值 ( $m$ )，但它代表的是非符号尾数的小数部分，而实际上其是标准化的  $1+m$  值，需要将这些值乘在一起；因此乘法运算将要求使用附加位。考虑到两个 53 位操作数的 MSB 都等于 1，并适当分解乘法运算从而以最佳方式利用 DSP48E2 26×27 位非符号乘法器及其改善后的各项功能（例如，由 W-mux 启动的真正的三输入 48 位加法器），可以看到，只需 6 个 DSP48E2 Slice 和极少的外部逻辑就可以构建 53×53 位非符号乘法运算。本文未涉及这类实现方式的所有细节，但在上一代 7 系列器件上，要实现类似的方法则需要 10 个 DSP48E1 Slice；因此 UltraScale 架构将需要的 Slice 数量减少了 40%。DSP48E2 的 27×18 乘法器对于以融合数据路径为基础的应用非常实用。最近，IEEE 浮点标准中加入了融

合乘累加运算符概念[8]。一般来说，这包括建立浮点运算  $A*B+C$ ，且乘法器和加法器之间无需明确舍入、标准化和非规范化的数据。当使用传统的浮点运算时，这些功能的代价非常高昂，且占用了最多的延迟时间。可推广这一概念以建立积和运算符，这在线性代数中非常常见（矩阵积，Cholesky 分解法）。因此，这种方法在成本或时间紧迫的情况下十分高效，同时要求具备浮点计数法的准确性和动态范围。在无线电 DFE 应用中，数字预失真功能通常要求一些硬件加速来支持，以提高非线性滤波器系数的更新速度。然后，您可以在 FPGA 架构中建立一个或多个浮点 MAC 引擎，以协助软件中运行的系数估算算法（例如，在 Zynq SoC 的一个 ARM®Cortex™-A9 核上）。对于这类算术结构，已经证明如果尾数宽度略有上升，从 23 位上升到 26 位，则与真正的单精度浮点实现相比，准确度会有所提高，但会减少延迟和空间占用。

UltraScale 架构非常适合该用途，因为它只需 2 个 DSP48 Slice 就可以打造单精度的融合乘法器，而 7 系列器件则需使用 3 个 Slice 以及其它架构逻辑。预加法器集成在位于乘法器前面的 DSP48 Slice 中，可提供一种高效的对称滤波器实现方式，这在 DFE 设计中很常见，可以实现数字上变频器(DUC)和数字下变频器(DDC)功能。对于 N 抽头对称滤波器来说，其输出实例的计算方式如下：

$$y(n) = \sum_{k=0}^{[N/2]-1} h(k) \cdot (x(n-k) + x(n-N+1+k))$$

其中  $x(n)$  代表输入信号， $h(n)$  代表滤波器脉冲响应，其中  $h(n) = h(N-1-n)$ 。

因此，成对的输入实例会被馈送入预加法器中，而输出则会进一步乘以相应的滤波器系数。在 7 系列架构上，预加法器必须使用 DSP48E1 的 30 位输入端 (A) 以及 25 位输入端 (D)，且其输出端需连接乘法器的 25 位输入端，同时将输入端 B 布线至 18 位乘法器输入端。

## 通过额外的 W-mux 多路复用器可将第四个输入操作数添加到

### ALU 中对无线电应用最为有益。

因此，在建立对称滤波器时，系数数值化后不得超过 18 位，这样可将阻带衰减限制在 85 到 90dB 左右。对于新一代 5G 无线电系统来说，这也许会成为问题，因为该系统很可能在干扰水平较高的环境下运行，因此也许需要更大的衰减。

UltraScale 架构解决了这个问题，因为可选择预加法器输入端作为 A 或 B，并在输出端集成了一些多路复用逻辑，从而支持将  $D \pm A$  或  $D \pm B$  馈送至任意乘法器输入端（27 位或 18 位输入端）。因此，可支持系数达 27 位的对称滤波器。赛灵思还为 DSP48E2 Slice 添加了另一项功能，将预加法器输出端连接至乘法器的两个输入端（在 18 位输入端上设有恰当的 MSB 截断）。这样便可以执行多达 18 位数据的  $(D \pm A)^2$  或  $(D \pm B)^2$  运算，从而可有效用于评估平方误差项的总和。此类运算在优化问题中十分常见，比如，在实施最小平方解决方案以获得调制解调器均衡器系数时，或按时间排列两个信号时。

毫无争议，通过额外的 W-mux 多路复用器将第四个输入操作数添加到 ALU 中对无线电应用最为有益。相比 7 系列器件上针对此类设计相同的实现方案，这个操作数通常可以将 DSP48 需求量减少 10%-20%。

只能将 W-mux 输出加入 ALU（不允许减少）中，并可将其动态设置为寄存器 C 或 P 内容，或在 FPGA 配置过程中将其设为常量值（比如用于 DSP48 输出端的收敛舍入或对称舍入的常量），或只需将其归零。通过这种方式可在使用乘法器时执行真正的三输入运算，如  $A*B+C+P$ ， $A*B+C+PCIN$ ， $A*B+P+PCIN$ ，而这在 7 系列架构中是不可能实现的。的确，使用乘法器阶段会生成最后两个部分乘积输出，然后将其添加到 ALU 中，以完成运算（见图 1）。因此，启动乘法器后，乘法器将使用 ALU 的两个输入端并执行一项三输入运算，而 7 系列器件则无法执行。

受益于这一额外 ALU 输入端的两个最主要实例是半并行滤波器和复合乘积累加（MAC）操作数。下面我们将对这两个实例进行详细讲解。

### 关于滤波器和 MAC

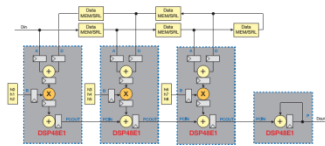
在任何 DFE 应用中，线性滤波器都是最常用的处理单元。在赛灵思 FPGA 上集成此项功能后，建议[6]尽可能实现多通道滤波器，因为它的复合采样率（定义为通道数量与每条通道一般信号采样频率的乘积）与该设计运行所用的时钟速率相同。在所谓的并行架构中，每个 DSP48 Slice 根据数据通道支

持单一滤波器系数，从而大幅简化了控制逻辑，并因此最大限度地降低设计资源利用率。

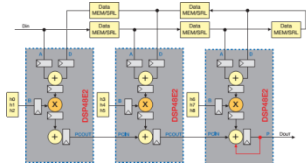
但是由于提高了时钟速率（例如，在最低速度等级 UltraScale 器件上时钟速率超过 500 MHz），且滤波器以相对较低的采样率运行，因此通常会选择时钟速率作为复合采样率的倍数。可尽量提高时钟速率，以进一步减小设计占用的空间，同时降低功耗。在这种情况下，将需建立一个半并行架构，其中每个 DSP48 会根据通道处理 K 系数，而 K 是时钟速率与复合采样率之间的比率。最高效的实现方案包括将滤波器拆分为多个 K 相位，每个 DSP48 可处理这些 K 相位的一个特定系数。

在每个时钟周期中计算出连续多个滤波器输出相位，并将其累加形成一个输出样本（每个 K 周期一次）。因此，与并行实现方案相比，滤波器输出需要增加一个累加器。这种全精度累加器可作用于较大的数据宽度，等于  $b_s+b_c+b_f$ ，其中  $b_s$  和  $b_c$  分别表示数据样本的位宽度和系数，而  $b_f=\log_2 N$  是滤波器的位增长，N 是系数总数。因此，常规做法是在 DSP48 Slice 内实现累加器，以保证支持最高时钟速率并最大限度地降低占用空间和功率消耗。

需要注意的是，各类滤波器都可以生成半并行架构，例如：单频率滤波器、整数型滤波器或分数率内插和抽取滤波器。图 3 是简化后的 7 系列和 UltraScale 实现方块图。图中清晰地突出了 UltraScale 解决方案的优势，由于其具备 W-mux 功能，相位累加器可被最后一个 DSP48 Slice 吸收。



(a) 7 系列实现方案



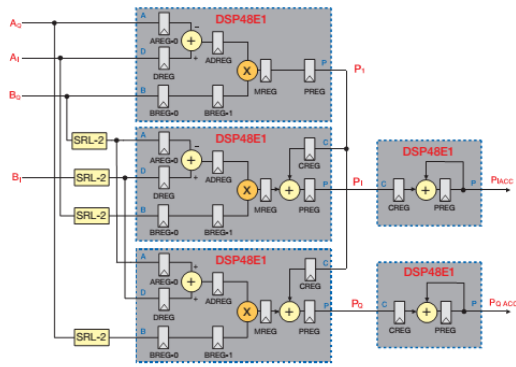
(b) UltraScale 实现方案

图 3 - 在 7 系列和 UltraScale 架构上实现半并行滤波器

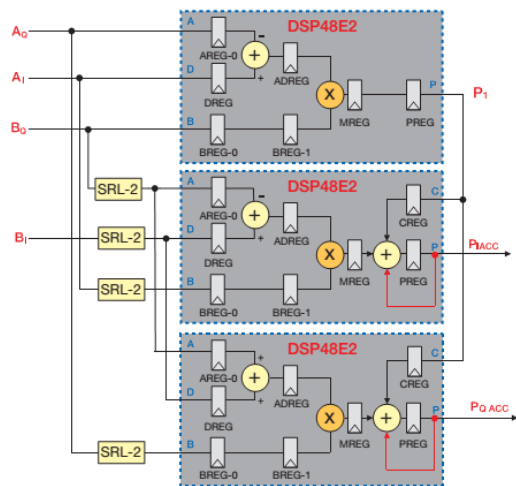
下面我们来思考一下如何实现每个时钟周期产生一个输出单元的全并行复合 MAC 运算符。众所周知，您可以重写复乘积的等式  $P_1+j.P_Q=(A_1+j.A_Q).(B_1+j.B_Q)$ ，从而根据以下等式仅使用三个真正的乘法运算：

- $P_1 = P_1 + A_1 \cdot (B_1 - B_Q)$
  - $P_Q = P_1 + A_Q \cdot (B_1 + B_Q)$
- 其中  $P_1 = B_Q \cdot (A_1 - A_Q)$ 。





(a) 7 系列实现方案



(b) UltraScale 实现方案

图 4 – 在 7 系列和 UltraScale 架构上实现复合 MAC

因此，利用内置加法器，您只需三个 DSP48（一个用于计算  $P_1$ ，另外两个用来处理  $P_1$  和  $P_0$  输出）就能实现复合乘法器。根据延迟需求以及其中表述的速度性能指令，需要增加一些逻辑以在不同数据路径间平衡延迟。为获取最大速度支持，必须对 DSP48 实行完全流水线化，从而让运算符整个延迟六个周期。随后在每个输入端增加一个两周延迟线，让实际数据路径和虚构数据路径保持一致。在每个输入位配备四个 SRL2，并利用 SRL 压缩功能将四个 SRL2 打包到两个 LUT 中。通过在每个  $P_1$  和  $P_0$  输出端添加累加器来最终完成复合 MAC。此外，这个累加器会作用于较大的数据宽度，因此能更好地集成到 DSP48 Slice 中。图 4 显示 7 系列和 UltraScale 器件相应的实现方案，从而再一次展示了 W-mux 集成的优势。 $P_1$  和  $P_0$  DSP48E2 Slice 在吸收累加器时可节省 40% 资源。值得一提的是，该方式还减少了延迟，对一些应用十分有益。

您可以用类似结构构建带有三个真实滤波器的复合滤波器（其中一个具备复合数据和系数），如图 5 所示。输入信号的真实部分和虚构部分会馈送至两个真实滤波器中，其系数分别为该滤波器系数虚构部分和真实部分的差值和总和。第三个滤波器用系数的真实部分处理并行的输入端真实与虚构部分的总和。当需要建立并行滤波器时，最终可将三个滤波器的输出端结合起来，以生成输出端的真实与虚构组件，其将再次受益于 W-mux，这也是 DFE 应用中使用均衡器的典型状况。

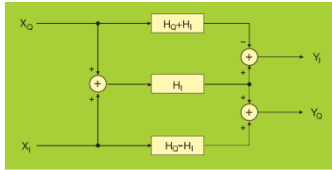


图 5 – 复合滤波器的实现架构

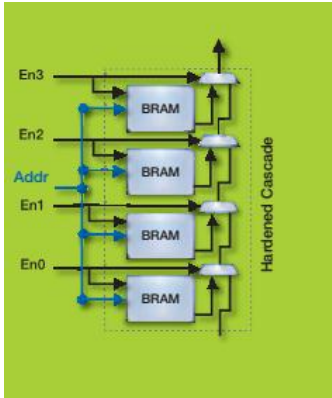


图 6 – UltraScale 器件上的 BRAM 级联

## UltraScale 内存架构 的优势

UltraScale 器件集成的 Block RAM 与 7 系列中集成的 Block RAM 基本相同，但新架构引入了硬件数据级联方案以及动态电源门控能力。图 6 显示了该级联，其中展示的是同一列中在每个上下相邻的 Block RAM 间嵌入的数据多路复用器。因此，无需使用其它逻辑资源便能以由下至上的方式建立更大内存。

该级联覆盖了器件上的所有整列，但最好将级联的使用限制在单个时钟区域（即 12 个连续 BRAM）以避免时钟偏移并最大限度提升时序性能。此外，它还提供足够的灵活性，可支持该级联功能的不同实现。实际上，您既可以将多路复用器应用到位于可选寄存器前后的 Block RAM 数据输入端，也可以用于数据输出端。

该级联使建立需要多个 BRAM 的大型内存成为可能，同时占用空间极少，支持最高时钟速率以及最低功率消耗，而这些在 7 系列器件上则无法实现。比如，在 7 Se 系列器件上，通过配置为  $16K \times 2$  位的八个 BRAM（36K）实现存储 16 位数据的 16K 内存，以避免外部数据复用，而这样则需要增加逻辑资源和延迟，并会影响计时与布线拥塞。从动态功率角度来说，这种方法效率较低，因为在所有读取或写入操作中都要启

动这八个 Block RAM。最佳解决方案包括使用 2Kx16 位配置，由于这样只需启动一个 BRAM，因此消耗的动态功率只有上述方法的八分之一。级联功能连同动态功率门控功能正好可以在 UltraScale 器件上实现这种效果。

Block RAM 级联的另一种直接应用与实现 I/Q 数据转换功能有关，通常与 DFE 系统的基带 CPRI 接口集成。图 7 展示了通常由一个 NxM 存储器阵列组成的高级交换架构。N 入口流上的连续数据会根据其输出目的地写入到相应的 Block RAM 中并形成一行，且会从相应的 Block RAM 以一系列的形式读取 M 出口流。因此，可以用 BRAM 级联有效实现每一列。

如欲了解有关 20nmUltraScale 系列的更多信息，敬请访问 <http://china.xilinx.com/products/silicon-devices/fpga/index.htm>。

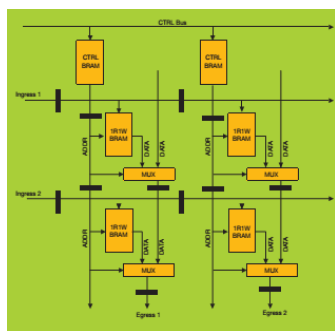


图 7- 数据交换高级架构

#### 参考文献

1. 赛灵思背景资料，2013 年 7 月，“推出 UltraScale 架构：业界首款 ASIC 级 All Programmable 架构”

2. 赛灵思白皮书 WP435，2013 年 8 月，“赛灵思 UltraScale：为您未来架构而打造的新一代架构”
3. 赛灵思数据手册 DS890，2014 年 2 月 6 日，“UltraScale 架构和产品简介”
4. 赛灵思背景资料，2013 年 7 月，“Vivado 设计套件加速设计生产力的九大理由”
5. 赛灵思用户指南 UG949，“Vivado 设计套件的设计方法指南”
6. 赛灵思白皮书 WP445，2014 年 1 月 20 日，“采用赛灵思 All Programmable FPGA 和 SoC 实现高速无线电设计”
7. 赛灵思用户指南 UG579，“UltraScale 架构——DSP Slice，高级规范用户指南”
8. IEEE 计算机社团，2008 年 8 月 29 日，“针对浮点架构的 IEEE 标准”