



WP485 (v1.0) 2017 年 6 月 14 日

Virtex UltraScale+ HBM FPGA： 革命性提升存储器的性能

作者：Mike Wissolik、Darren Zacher、Anthony Torza 和 Brandon Day

数据中心、有线应用及其它带宽密集型应用所需的性能，远远高于传统的 DRAM 技术。和市场上已有的存储器相比，HBM 存储器在性能、功耗和尺寸上，能为系统架构师和 FPGA 设计人员带来前所未有的优势。

摘要

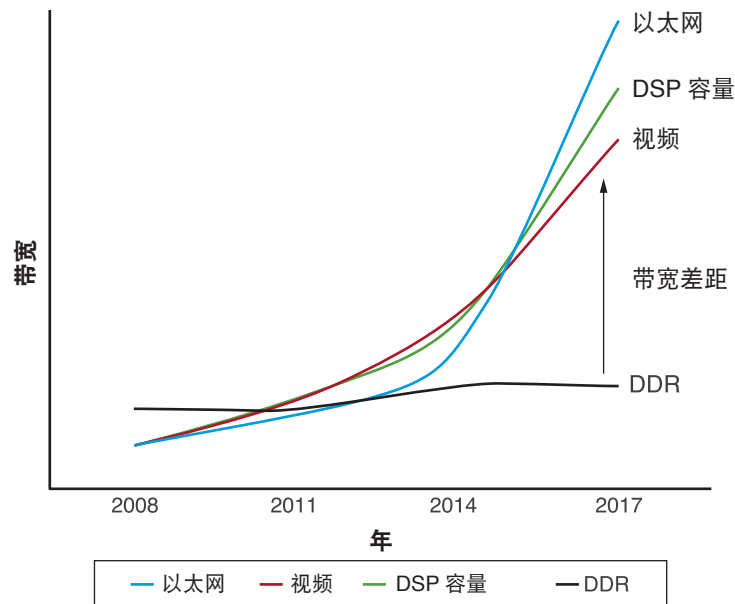
在过去的十年里，电子系统在计算带宽上呈现出指数级的增长。计算带宽的大幅提升，也显著提高了存储带宽要求，以满足计算需求。这类系统的设计人员经常发现市场上的并行存储器（例如 DDR4）再也无法满足应用的带宽需求。

赛灵思支持高带宽存储器 (HBM) 的 FPGA 能够以最低的功耗、尺寸和系统成本提供高带宽，显然能够轻松应对这类挑战。在设计这款 FPGA 的过程中，赛灵思与其他领先半导体厂商一样，选择了业界唯一经过证明的堆叠硅片互联技术（即台积电 (TSMC) 的 CoWoS 集成工艺）。这篇白皮书将介绍赛灵思 Virtex® UltraScale+™ HBM 器件如何满足大幅提升的系统存储带宽需求，同时保持功耗、尺寸和成本在限定范围内。

行业趋势：带宽和功耗

过去十年里，并行存储器接口的带宽功能进步缓慢——如今 FPGA 中支持的最大 DDR4 数据速率仍然不足 2008 年 DDR3 数据速率的 2 倍。但就在这段时间内，存储带宽需求增长远远超过了 DDR4 的性能。考虑以太网的发展趋势：从 DDR3 时代开始，以太网端口速度从 10Gb/s 提高到了 40Gb/s，然后提高到 100Gb/s，现在到了 400Gb/s——原始带宽增加了 10 倍以上。

类似的趋势也存在于高性能计算和视频广播市场。FPGA 机器学习 DSP 容量已从最大型 Virtex-6 FPGA 中的 2,000 个 DSP 增加到了现在的最大型 Virtex UltraScale+ 器件中的超过 12,000 个 DSP 元件。视频广播行业已经从标准清晰度过渡到 2K，现已达到 4K，很快将达到 8K。这些应用领域中，所需带宽与 DDR4 DIMM 能提供的带宽之间都存在明显的差距。见图 1。



WP485_01_051217

图 1：存储器带宽要求比较

为了弥补带宽差距，系统架构师若要在这些应用中使用 DDR4，就必须增加系统中的 DDR4 元件数量——这不是为了增加容量，而是为了在 FPGA 与存储器之间提供所需的传输带宽。四个 DDR4 DIMM 以 2,667Mb/s 的数据速率运行，所能实现的最高带宽为 85.2GB/s。如果应用所需的带宽超过这个值，那么 DDR 方案就会因功耗、PCB 尺寸和成本问题变得不可行。不难看出，这些高带宽应用中需要一种新的 DRAM 存储方案。

从能效的角度重新审视这十年时间，很显然这种不惜一切代价“提高性能”的时代已经结束。MDPI 发表的一篇文章中预测，到 2030 年，根据当时数据中心设备的实际能效来看，仅数据中心就能消耗 3%-13% 的全球能源供应 [参考资料 1]。设计人员极为重视能效性能，尤其在这个多兆瓦级数据中心

时代。他们还重视高效散热解决方案，因为可靠的通风和冷却需要很高运营支出——总能耗的三分之一 [参考资料 2]。因此，供应商如果能以较低散热量实现最高单位成本计算性能和单位功耗计算性能，则其解决方案会极富吸引力。

DDR4 DIMM 的替代方案

为了弥补带宽差距，半导体行业引入了多种巧妙方案来替代 DDR4。见表 1。最近，业内兴起了基于收发器的串行存储器技术，例如混合存储立方体 (HMC)。这些技术提供更高的存储器带宽，能够在单个芯片中提供相当于几个 DDR4 DIMM 的存储带宽——但需要将多达 64 个超高速串行收发器分配至存储器子系统中。

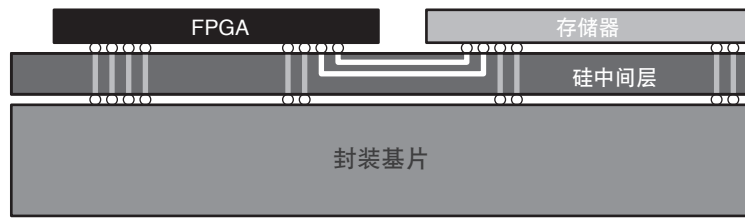
表 1 : 不同存储器解决方案的关键特性对比

	DDR-4 DIMM	RLDRAM-3	HMC	HBM
描述	服务器和 PC 中使用的标准商用存储器	用于数据包缓冲应用的低时延 DRAM	混合存储立方体串行 DRAM	集成在 FPGA 封装中的高带宽存储器 DRAM
带宽	21.3GB/s	12.8GB/s	160GB/s	460GB/s
典型深度	16GB	2GB	4GB	8GB
价格 /GB	\$	\$\$	\$\$\$	\$\$
PCB 要求	高	高	中	无
pJ/ 位	~27	~40	~30	~7
时延	中	低	高	中

高带宽存储器简介

通过移除 PCB，HBM 能以不同方式解决存储器带宽问题。HBM 利用硅片堆叠技术将 FPGA 和 DRAM 并排放在同一封装内。这样，采用相同封装的 DRAM 结构能够实现多个 Tb/s 的带宽。与其它存储器技术相比，该技术使设计人员能够实现带宽的大幅度跨越式提升。

HBM 器件采用台积电 (TSMC) 的业界标准 CoWoS(chip-on-wafer-on-substrate) 堆叠硅片组装工艺进行组装。赛灵思已经在过去三代高端 Virtex 器件中采用这种组装技术，因此该技术已得到了证明。CoWoS 起初由赛灵思率先采用，作为硅片堆叠互联技术应用于 28nm Virtex-7 FPGA 中。CoWoS 组装工艺将有源芯片放在无源硅中间层上。硅与硅的堆叠结构允许通过非常小、分布非常密集的微凸块来连接相邻的硅器件——这里是将 FPGA 连接到 DRAM，之间有成千上万的信号。见图 2。



WP485_02_051217

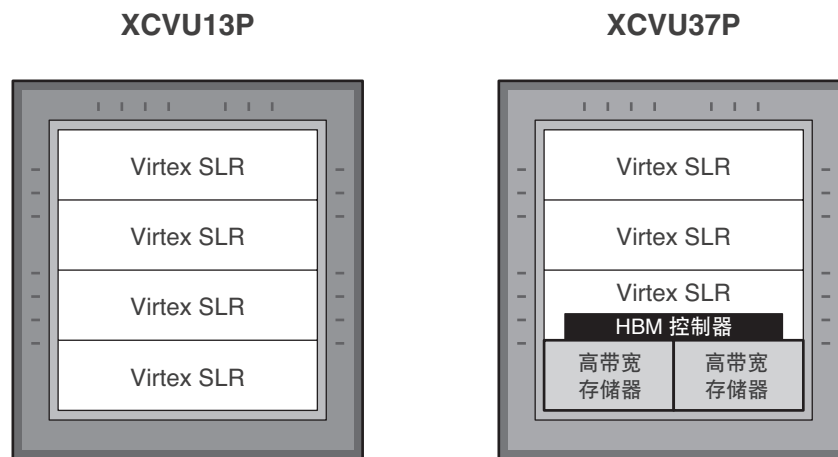
图 2 : TSMC CoWoS 组装工艺允许通过数千个非常小的线连接相邻晶片

采用 CoWoS 组装工艺，与典型的 DDR4 PCB 走线相比，不仅连接 HBM 的 DQ 走线总长度不足 3mm，而且电容和电感 (LC) 寄生效应极低。这样，HBM I/O 结构的芯片面积比典型外部 DDR4 I/O 结构的芯片面积小 20 倍。HBM 接口非常小，以致于单个 HBM 堆栈接口就包含 1,024 个 DQ 引脚，而且 I/O 芯片面积仅为单个 DDR4 DIMM 接口 I/O 芯片面积的一半。具有 1,024 个 DQ 引脚，而且低寄生效应，这样能实现非常高的 HBM 堆栈输入输出带宽，而时延与 DDR4 相近。

对于采用 HBM 的 FPGA，所使用的外部 DDR4 的数量与容量要求有关，与带宽要求无关。这样所用的 DDR4 组件数量大大减少，为设计人员节省了 PCB 空间和功耗。有些情况下根本不需要外部存储器。

赛灵思 HBM 解决方案简介

如图 3 所示，Virtex UltraScale+ HBM 器件通过集成赛灵思合作供应商提供的业经验证的 HBM 控制器和存储器堆栈，可使用与已投入量产的赛灵思 16nm UltraScale+ FPGA 系列相同的构建块来构建。利用经过量产验证的 CoWoS 组装工艺集成 HBM，通过标准的 Virtex FPGA 组装流程将基础 FPGA 组件与 HBM 简单地堆叠在一起。该方法消除了产能风险，因为基础 FPGA 系列器件中使用的所有芯片、IP 和软件都经过了量产质量级认证。



量产级 VU13P (SSI 技术)

支持 HBM 的 VU37P

WP485_03_060917

图 3 : SSI 技术与支持 HBM 的 XCVU37P

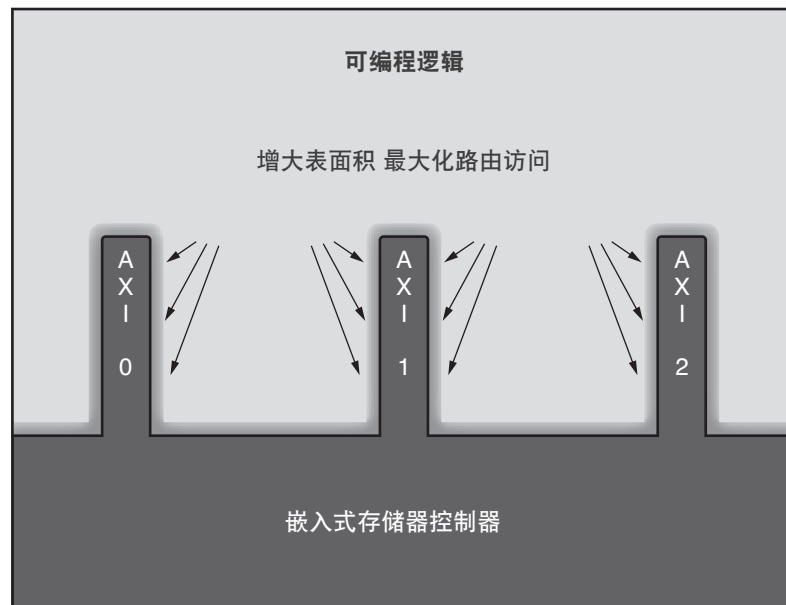
Virtex UltraScale+ HBM 器件中新增加的模块只有 HBM、控制器和加速器的缓存一致性互连 (CCIX) 模块。收发器、PCIe® 的集成模块、以太网、Vivado® Design Suite 等均已经得到量产质量级认证，使设计人员能够集中精力充分发挥 HBM 的特性与功能，使产品在市场上中脱颖而出。

时序收敛创新

由于 Virtex UltraScale+ HBM 器件的基础已经得到验证，因此赛灵思工程师可将创新工作重点放在优化 HBM 存储器控制器上。HBM 与 FPGA 集成的过程中，最明显的挑战在于有效利用 HBM 提供的所有存储器带宽。赛灵思在这些器件中提供了几大关键而独特的创新特性，以帮助客户获得 HBM 堆栈最高可用的输入输出带宽。

扩展的 AXI 接口

第一个创新是 AXI 接口，用户可通过该接口连接存储器控制器。典型的集成 IP 与紧挨着嵌入式 IP 模块的可编程逻辑进行接口连接。对于大部分模块来说这已经足够，因为本地路由的汇聚带宽足以从该模块输入输出数据。然而，进出 HBM 的带宽非常高，因而有必要创建新的接口结构类型，以便扩展到可编程互连。该结构明显增加接口表面积，极大提高用户 AXI 接口的可用互连能力，可实现 3.7Tb/s 的运行速度。见图 4。



WP485_04_061317

图 4 : 扩展的 AXI 接口

灵活寻址

第二个创新是 HBM 存储器控制器中包含的灵活寻址功能。HBM 堆栈将存储器地址空间分成伪通道。这意味着任何给定的 HBM DQ 位都被分配到特定的存储器地址区域。因此，如果设计人员想把数据写入存储地址，只能通过与该地址关联的伪通道来写入数据。

如果设计人员想把 HBM 堆栈视为单个连续存储器，或者跨伪通道边界将它们进行分区，那么这种限制并不理想。为了克服这种局限性，赛灵思在嵌入式存储器控制器中包含了一个 AXI 交换网络。这个交换网络能够根据地址从任意源 AXI 接口将存储器读和写路由到任意 HBM 伪通道。该功能称为灵活寻址，因为它允许任意用户 AXI 接口访问任意 HBM 存储器地址。

对于想针对特定存储器访问形式来优化存储器控制器的用户来说，也可以绕开灵活寻址。见图 5。

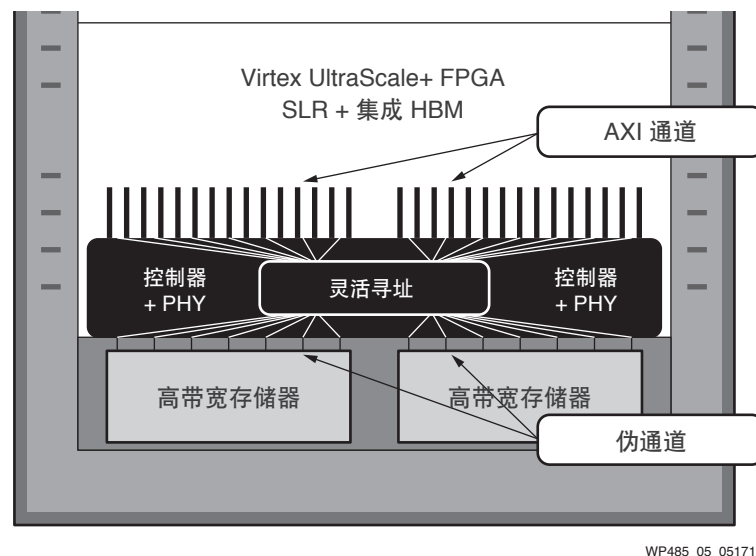
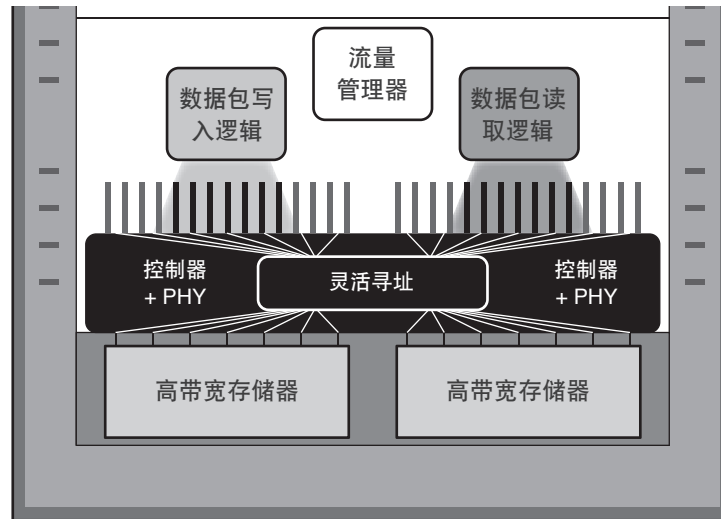


图 5 : AXI 接口 (到用户逻辑) 和 HBM 伪通道 (到 HBM 堆栈)

灵活寻址具备多个重要优势：

- 1 让用户能够完全控制对 HBM 堆栈的寻址。**由于交换网络能路由整个器件，因而用户无需遵守 HBM 固有的严格伪通道要求。32 个 AXI 接口都能读写任一 HBM 堆栈上的任意 HBM 伪通道，使用户能完全控制地址分区，无需考虑伪通道边界。
- 2 允许设计人员根据设计的时序收敛情况使用最方便的 AXI 接口进行连接。**例如，向存储器写入数据的逻辑无需与从存储器读取数据的逻辑处在相同位置。在基本流量管理器实例中，数据包写入和数据包读取模块的 AXI 接口均可处在距离模块最近的位置。见图 6。



WP485_06_051917

图 6：典型的以太网桥接设计

利用灵活寻址，可将数据包写入和数据包读取逻辑分开，以避免争夺路由资源。

- 4 大量减少可编程路由资源。**存储器控制器中有很多 AXI 总线，宽度均为 256 位。因此，如果集成存储器控制器中提供 32 条在器件中水平路由的通道，就能为 FPGA 设计人员释放可观的资源，以便用于具有更高价值的功能。AXI 交换网络如果完全在 FPGA 逻辑中实现，会占用 250,000 个 LUT。如果利用灵活寻址，则整个交换网络无需使用 LUT。
- 5 与伪通道方案相比，允许设计人员更高效地使用 AXI 接口。**HBM 伪通道具有典型的 DRAM 低效特征，例如激活、预充电和刷新延迟。尽管存储器控制器确实通过重新排序来提高效率，但是 DRAM 不可能达到 100% 高效。然而，单个 AXI 接口能够将多个伪通道访问实现流水线，因此获得高于 HBM 伪通道效率的 AXI 接口效率。在众多应用中，这意味着所需的 AXI 接口数量更少，能够释放更多可编程资源。

能效和热管理方面的创新

赛灵思客户非常重视能效性能。TSMC 16nm FinFET+ 工艺支持双电压运行，使客户能够选择最高绝对性能或者最高每瓦特性能。利用此工艺，赛灵思能够提供业界最低内核电压，将动态总功耗降低 30%，提供行业领先的收发器技术，以及在 FPGA 中混合多种集成模块，例如以太网、Interlaken 和 PCIe。

HBM 技术使赛灵思能够取消外部存储器接口，用中间层上的走线取而代之，从而将节能设计推向新高度。这样做可降低存储器接口电容，从而降低多 Tb/s 存储带宽所需的功耗，将每比特功耗降低 5 倍。

热管理方面，赛灵思提供多种独特技术，用以抵消在集成 HBM 之后产生的不可避免的热密度增加问题。赛灵思的 Virtex UltraScale+ HBM 器件采用散热片就绪型、无盖、裸芯片、倒装芯片封装，能显著改善散热性能，缓和更高热密度问题。这些无盖封装已用于其它 Virtex UltraScale+ FPGA，并在大多数使用案例中将散热设计改善 10° C 左右。这实现了更高的计算性能上限和 / 或更低的散热设计成本。如需了解更多信息，敬请阅读赛灵思应用指南《UltraScale+ FPGA D2104 无盖倒装芯片封装的机械和散热设计指南》[参考资料 3]。见图 7。

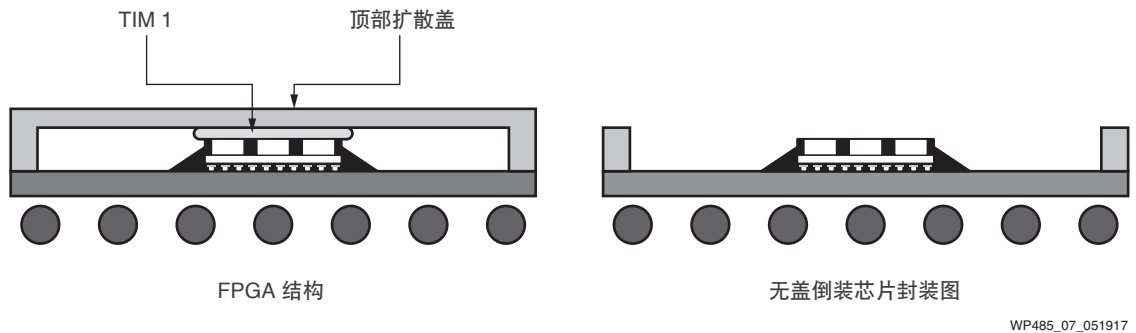


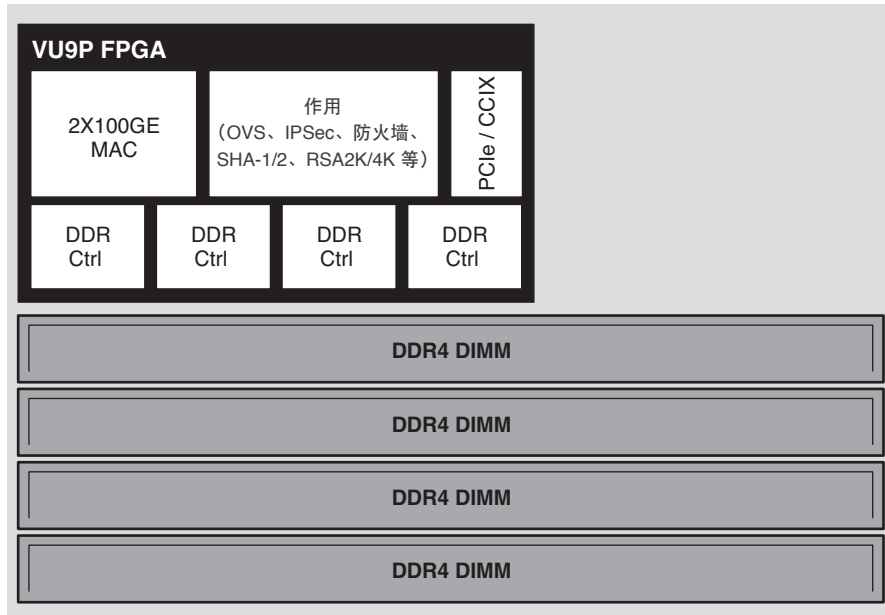
图 7：有盖 vs 无盖倒装芯片封装

应用实例：智能网络接口卡

HBM 与高端可编程逻辑的联姻在网络、数据中心、音频 / 视频广播、雷达、测试与测量等众多应用领域中发挥出巨大优势。其中一种应用是智能网络接口卡或智能 NIC。智能 NIC 包含：一个或多个网络端口，一个连接 CPU 的接口（例如 PCIe® 或 CCIX），要加速的网络功能（例如 OVS、GZIP、IPSec、SSL 等），以及用于数据包存储和键值查找的存储器。传统的智能 NIC 需要在 PCB 上安装四个 72 脚 DIMM，以提供足够的存储器带宽来服务两个 100G 端口。连接四个 DIMM 需要驱动 624 个 I/O，会显著增加总功耗。容纳四个 DIMM 需要全高全长 (FHFL) 尺寸，会带来一系列功耗和空间效率挑战。

如果在采用 HBM 的 VU35P 中实现相同方案，则尺寸能缩减到半高半长 (HHHL)，因为外部 DRAM 元器件被 HBM 堆栈取代（见图 8）。VU35P 方案（图 9）的功耗约为 50%，因为避免了 DIMM 接口的 I/O 功耗。若使用包含两个 HBM 堆栈的 VU35P 器件，那么得益于 HBM 带宽，查找速度可提升 3 倍，而且搜索条目是市场上的 TCAM 的 2 倍。除了最终解决方案的这些固有优势外，HBM 方案还能简化 PCB 并降低存储器子系统的复杂性，从而实现更简单、风险更低的设计流程。

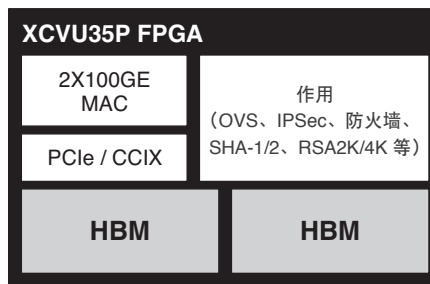
现有基础架构



WP485_08_051717

图 8 : 现有基础架构

Virtex UltraScale+ HBM FPGA



WP485_09_051917

图 9 : Virtex UltraScale+ HBM 解决方案

结论

未来很多系统会超出 DDR 所能提供的带宽，HBM 作为最佳选择，能大幅提高存储器带宽，并实现最佳的单位功耗性能。赛灵思 Virtex UltraScale+ HBM 器件提供恰当的存储器带宽和可编程计算性能组合。凭借这些器件，赛灵思重点帮助设计人员充分利用 HBM 的性能，同时将经过验证的芯片工艺和架构、组装技术以及设计工具作为设计开发的基础。设计人员和系统架构师都会领略通过 Virtex UltraScale+ HBM 器件将 HBM 功能引入系统所带来的优势。

参考资料

- 1 “On Global Electricity Usage of Communication Technology:Trends to 2030” , Anders S. G. Andrae and Tomas Edler, Apr 30, 2015.
<http://www.mdpi.com/2078-1547/6/1/117/htm>
- 2 “Data Center Specific Thermal and Energy Saving Techniques” , Tausif Muzaffar and Xiao Qin, Jun 18, 2015.
<http://www.slideshare.net/xqin74/data-center-specific-thermal-and-energy-saving-techniques>
- 3 Xilinx application note [XAPP1301](#), Mechanical and Thermal Design Guidelines for the UltraScale+ FPGA D2104 Lidless Flip-Chip Packages.

修订历史

下表列出了本文档的修订历史：

日期	版本	修订描述
2017 年 06 月 14 日	1.0	赛灵思初始版本

免责声明

本文向贵司 / 您所提供的信息（下称“资料”）仅在选择和使用赛灵思产品时供参考。在适用法律允许的最大范围内：(1) 资料均按“现状”提供，且不保证不存在任何瑕疵，赛灵思在此声明对资料及其状况不作任何保证或担保，无论是明示、暗示还是法定的保证，包括但不限于对适销性、非侵权性或任何特定用途的适用性的保证；且 (2) 赛灵思对任何因资料发生的或与资料有关的（含对资料的使用）任何损失或赔偿（包括任何直接、间接、特殊、附带或连带损失或赔偿，如数据、利润、商誉的损失或任何因第三方行为造成的任何类型的损失或赔偿），均不承担责任，不论该等损失或者赔偿是何种类或性质，也不论是基于合同、侵权、过失或是其他责任认定原理，即便该损失或赔偿可以合理预见或赛灵思事前被告知有发生该损失或赔偿的可能。赛灵思无义务纠正资料中包含的任何错误，也无义务对资料或产品说明书发生的更新进行通知。未经赛灵思公司的事先书面许可，贵司 / 您不得复制、修改、分发或公开展示本资料。部分产品受赛灵思有限保证条款的约束，请参阅赛灵思销售条款：<http://china.xilinx.com/legal.htm#tos>；IP 核可能受赛灵思向贵司 / 您签发的许可证中所包含的保证与支持条款的约束。赛灵思产品并非为故障安全保护目的而设计，也不具备此故障安全保护功能，不能用于任何需要专门故障安全保护性能的用途，如果把赛灵思产品应用于此类特殊用途，贵司 / 您将自行承担风险和责任。请参阅赛灵思销售条款：<http://china.xilinx.com/legal.htm#tos>。

汽车应用免责声明

汽车产品（产品部件号中标识为“XA”）不保证用于安全气囊的开发或用于影响车辆控制的应用（“安全应用”），除非在该赛灵思产品中具备故障安全保护或者额外功能，符合 ISO 26262 汽车安全标准（“安全设计”）。为安全起见，客户应在使用或分销任何集成有该产品的系统之前，对这些系统进行全面测试。在没有安全设计的安全应用中使用产品的风险完全由客户承担，仅受有关产品责任的适用法律和法规限制。