

三维微处理器设计基本方法及前景分析

怡 磊,单光宝,肖建青,刘 松

(西安微电子技术研究所 陕西 西安 710054)

摘 要: 介绍了三维处理器设计的基本内容,详细阐述了用于三维处理器中三维存储系统的设计方法及发展趋势,最后概括描述了目前三维处理器、三维存储系统设计中存在的问题,这将为国内三维处理器设计提供一定的借鉴。

关键词: 三维集成技术;三维处理器;3D 存储;TSV

中图分类号: TN4

文献标识码: A

文章编号: 1000—7180(2016)08—0068—04

Analysis of Basic Methods and Prospect of 3D Microprocessor Design

YI Lei, SHAN Guang-bao, XIAO Jian-qing, LIU Song

(Xi'an Microelectronics Technology Institute, Xi'an 710054, China)

Abstract: This paper introduces the fundamental contents of the three dimensional processor, describes the design methodology and development of the three dimensional storage system and discusses the development of the future development of three-dimensional processor. Finally, the paper concludes the existing problems in the design of 3D processors and 3D memory.

Key words: three-dimensional integration technology; 3D processor; 3D memory; TSV

1 引言

三维集成技术是一种新兴的集成电路制造工艺技术^[1-4],它通过采用垂直互连硅通孔(through silicon via, TSV)技术立体集成多个设备层(如图 1(a)),相比传统平面工艺,三维集成技术能有效减少全局互连长度,降低芯片单位面积半导体器件密度,为微处理器设计提供了高密度、低延迟的层间互连结构,使其达到高性能、低功耗的设计需求,图 1(b)给出了一种三维多核处理器基本结构.由于存储系统作为微处理器重要组成部分,占据了其面积的 50%左右,同时也是限制处理器访问速度及成本的重要因素,因此目前三维处理器设计主要是针对三维存储系统进行研究。

本文旨在阐述目前三维处理器及三维存储系统的基本设计方法,以及三微处理器的发展趋势,并简要介绍三维处理器设计所面临的问题,还对三维处

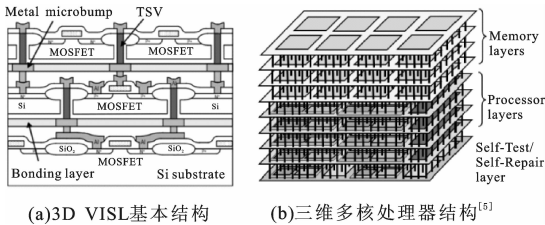


图 1 3D VISL 基本结构与三维多核处理器结构^[5]

理器设计及三维存储系统设计方法以及发展前景进行了分析,旨在为国内三维处理器研究提供一定参考。

2 三维处理器设计方法

2.1 三维处理器优点

三维处理器可以解决平面处理器所面临的全局互连问题,同时降低互连功耗和延迟,提升处理器性能.文献[1]以三维设计逻辑单元结构为例,详细说明了由于互连线长缩减,算术逻辑单元延迟减少了

大约 6%~30%;文献[2]将 Intel 奔腾 4 处理器进行三维设计,三维奔腾 4 处理器结构由于全局互连减少,系统中中继器及中继器锁存器减少 50%左右,除此之外,三维钟控网络的 RC 参数也减少了约 50%,从而降低处理器功耗,三维奔腾 4 处理器设计使处理器温升在 14 度时,处理器性能仍可上升 15%,同时功耗降低 15%。李丽^[3]等人研究了一种三维众核处理器存储架构,验证了三维结构对于处理器性能的提升效果。总的说来,目前国内关于三维处理器的研究仍旧很少,并且基本停留在比较初级的阶段。

2.2 三维存储系统的划分方式

目前三维存储系统的研究主要集中在由 DRAM/SRAM 构成 Cache 的三维划分。这部分主要介绍传统三维存储结构的划分方法。

2.2.1 直接划分

直接划分是将传统平面存储直接分割,将分割所得结构进行三维集成,增加单位面积存储容量。直接划分可以缩减等容量存储系统所占的芯片面积,提升了存储系统单位面积的存储容量,缩减处理器访问时间,但不能改变最坏路径的长度,对处理器的延迟以及功耗影响甚微。研究者提出将处理器核与缓存分别立体集成。

2.2.2 体划分

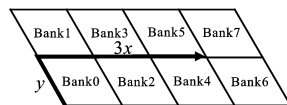
体划分是按存储体模块进行分割,可以缩减最长互连线长,降低线上 RC 参数、延迟和功耗。图 2(a)所示为有八个体模块的存储结构,图中给出其最长互连线,利用体划分对存储结构进行划分,如图 3(b)所示。由图 2 可知,水平方向最长互连线长由原先 $4 \times x$ 变为 x ,线长减少了 67%;假设图中 $x=y$,互连线总长将会减少了 50%。体划分方式减少互连线长多少的程度与划分方向密切相关,图中给出是可以最大限度降低最长互连线的划分方法。

2.2.3 阵列划分

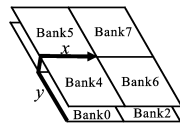
阵列划分是一种较精细的划分方式,即对 SRAM/DRAM 阵列进行划分。阵列划分可减少字线或位线长度,按划分方式不同,一般可分为字线划分和位线划分两种。

(1) 字线划分

字线划分是指将存储系统沿位线方向进行分割,将所得结构进行三维集成。如图 3(b)所示,字线划分可减少存储系统字线长度,降低字线电阻、电容,减少线上延时和功耗,提升存储系统性能。;图 3(b)中字线译码器置于下层,上层字线均以 TSV 孔



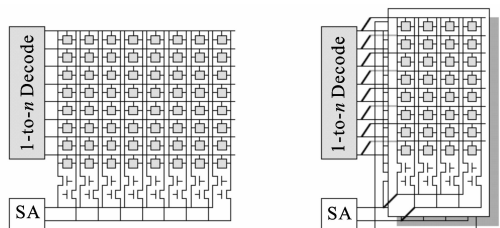
(a) 平铺八个存储体模块及其一条最坏传输路径



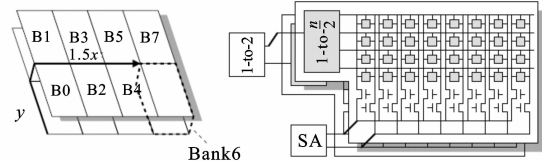
(b) 3D 存储体立体集成

图 2 平铺八个存储体模块及其一条最坏传输路径与 3D 存储体立体集成^[4]

与字线译码器互连,进行数据传送。对于单个字线而言,线上电阻、电容均减少为原先一半。但每个字线均需布置 TSV 孔与字线译码器互连,孔密度及数量庞大,对 TSV 工艺构成极大挑战。同时受限于目前工艺水平,TSV 孔尺寸相对较大,引入大量 TSV 孔将引起诸多热、机械应力等问题,这将削弱三维集成所带来的优势。



(a) 传统二维平面 SRAM 阵列 (b) 按字线划分存储阵列三维集成



(c) 存储体模块层次三维结构 (d) 按位线划分存储阵列三维集成

图 3 传统二维平面 SRAM 阵列、按字线划分存储阵列三维集成、存储体模块层次三维结构及按位线划分存储阵列三维集成^[9]

(2) 位线划分

位线划分是将存储系统沿字线方向进行划分,将所得结构进行三维集成。如图 3(d)所示,位线划分可减少位线长度,降低存储系统门负载及线上负载、功耗及延时,提升存储系统性能。采用位线划分进行三维集成时,将字线译码器减半,并置于两层,增加 1-to-2 选择器判断上层或下层的行译码器选通。尽管增加 1-to-2 选择器会引入额外延迟,但相比位线延迟减少可以忽略不计,并且位线划分 TSV 孔

数目较少,密度较低,相比字线划分易于实现。

2.3 三维存储的发展趋势

随着三维设计技术逐渐成熟,开始研究如何更好地利用三维存储系统来提高多核处理器性能,从而使三维多核处理器存储系统逐渐成为研究热点,目前,三维存储系统发展趋势主要有以下三点。

(1) 立体集成存储

目前研究将三维存储系统直接立体集成于多核处理器之上,构建高性能多核处理器结构。文献[5]基于 ELS 平台将三维存储结构应用于多核处理器,将 512K SRAM 立体集成于逻辑单元之上,其中 SRAM 和逻辑单元采用 TSMC 90 nm CMOS 工艺制作,整体异质集成于多核处理器之上。相比平面处理器,三维多核处理器系统性能提升了 66.4%,但存储系统和处理器核的立体集成并不能完全发挥三维集成技术以及三维互连的优点。文献[6]设计了一个高性能多端口存储 IP,将三维 DRAM 存储模块与多核处理器通过片上网络(Network on Chip, NOC)进行互连,利用三维集成技术以及三维互连的优点,极大地开发了三维 DRAM 存储模块潜力,利用三维 TSV 互连构成具有大带宽,低访问时间的 DRAM 存储模块。并且进一步采用基于对数互连的低延迟 3D NUMA(非一致性内存访问结构)与多核处理器立体集成,使得系统平均性能提升了 34%,同时将 4M 3D NUMA 分割为八层时,存储密度仅丢失 16%。

(2) 共享存储

共享存储也是三维多核处理器设计方式之一。共享存储是多个处理器核共享某一存储模块,提升存储系统利用率,降低访问时间,同时大幅提升带宽。文献[7]基于 TSV 互连结构,采用 H-trees 进行数据传输,构建多面缓存结构,将 TSV 互连分布在整个存储阵列之上构成两个垂直相交界面,如图 4 所示。其中,垂直面提供大的带宽,水平面虽然无法提供大带宽,但提供更大的体模块访问端口,减少存储体访问竞争。共享缓存可以大幅提升处理器性能,减少处理器所占面积。

(3) 分布式存储

分布式缓存是将细粒度的三维缓存结构嵌入片上存储,从而减少片外缓存访问时间。文献[8]将分布式缓存放置于 TSAR 多核处理器之中,对于 256 核的 TSAR,采用 1:1 缓存和互连比例,可以达到 2 Tb/s 的带宽传输速度。

以上介绍了三维立体集成处理器的设计思路和

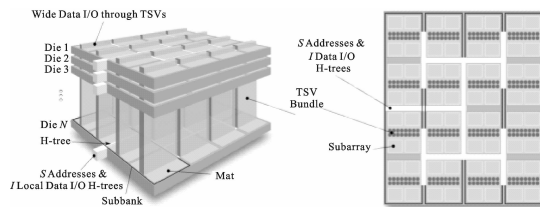


图 4 三维共享存储阵列

方法。三维立体集成技术作为一种新兴技术,可以解决后摩尔时代所面临的互连问题。随着处理器性能需求越来越高,诸如延时,功耗等问题将逐渐凸显,将三维技术应用于微处理器设计中是解决这些问题的有效方法。

未来关于立体集成处理器的研究,必然向着更高水平、更精细粒度以及更高集成部件的层面发展,依托于异质集成技术,三维计算机系统结构也将产生巨大的变化(如图 5 所示)。三维处理器架构的整体设计,包括存储系统,控制结构以及算术逻辑结构的设计,也将成为三维处理器设计的主要方向。

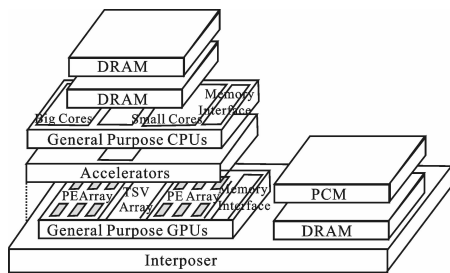


图 5 未来的 3D 计算系统结构概述-包括 CPU、GPU、加速器、片上存储器和片外存储器^[15]

3 困难与挑战

首先,互连结构的优劣直接影响三维集成设计对系统性能的影响。目前,3D 集成技术主要是采用 TSV 纵向互连结构,TSV 孔可缩短互连线长,降低功耗,但会引起较大寄生电容、电阻,同时由于制造工艺的限制,TSV 互连结构中 TSV 孔的尺寸较大,同时在制备过程中,需在 TSV 外围构建阻挡区域,极大地减少了芯片上可利用的面积,而且会带来应力集中等问题。并且当 TSV 孔数目过多时,TSV 孔内、孔间寄生参数将对处理器系统性能造成巨大的影响。针对这些问题,文献[9]提出采用光学互连结构取代 TSV 结构进行三维集成电路数据的传输,光学互连结构相比于 TSV 结构具有更小的寄生参数以及更快的传输速度,但光学互连结构设计制造方面仍存在困难,随着工艺制造技术进步,采用光学互

连结构取代 TSV 结构,将可能会改变整个计算机结构。

其次,三维处理器立体集成时,热源的布局非常重要,目前诸多学者针对三维处理器热设计进行了广泛的研究,Dong Oh Son^[10]等依据处理器核和 L2 高速缓存的相互布局,分析了三维处理器在立体集成过程中布局对于处理器热设计的影响,说明了三维处理器热问题的危害,同时给出了一种较为合理的布局方式, Kim^[11]等对 3D 集成 DRAM 结构热的问题进行了分析,他们分别采用了三种堆叠模型,并对其热情况进行了详细的分析,对三种结构所引发的热问题进行了仿真、比较,并得到最优化的结果。

尽管目前对 3D 处理器的热分析、热管理进行了很多研究,但 3D 处理器的热问题还远未得到解决。随着三维处理器设计中热问题越发不可忽视,热管理、热分析将是三维处理器设计所必须解决的问题。

此外,三维处理器设计缺乏系统的设计工具及方法学,后期测试手段也不够健全,而且生产成本较高、生产难度较大等缺点也限制着它的进步。Synopsys 公司 Monolithic 3D 公司对三维系统级设计规划工具进行了研究,构建了 3DSim 仿真器处理系统级设计的各项参数及模型。但这些工具都只是从一个或者数个方面解决三维设计中所面临的问题,并不能完整的对三维设计进行分析。

在国内,三维集成技术还处于初级阶段,存在着诸多问题与挑战。

4 结束语

本文总结了三维立体集成处理器的设计思路和研究进展,以及三维处理器设计所面临的若干问题。三维立体集成技术作为一种解决目前工艺设计瓶颈的方法,有着较好的发展前景。处理器作为系统的核心部件,对于系统性能、延时、数据交互起着决定性的作用。三维立体集成处理器设计将两者结合起来,用三维集成技术实现处理器性能提升。随着工艺不断发展,三维立体集成处理器设计必将取得更大的进步,并提供更优良的性能、更低的功耗、更小的延迟。尽管在设计中,依旧存在亟需解决诸如热分布、机械应力、物理排布和系统封装等各方面问题,但是随研究不断深入,这些问题也将会得以解决。

三维立体集成处理器设计是高性能、低功耗处理器设计的一个发展趋势,随着研究不断深入,三维处理器将提供更小面积、更高性能的处理器结构。

参考文献:

- [1] R Egawa, J Tada, H Kobayashi, et al. Evaluation of fine grain 3D integrated arithmetic units. In IEEE International 3D System Integration Conference, San Francisco, CA; IEEE, 2009;1-8.
- [2] Xie Y. Processor architecture design using 3D integration technology[C]// VLSI Design, 2010. VLSID'10. 23rd International Conference on. India; IEEE, 2010; 446-451.
- [3] 李丽, 张宇昂, 傅玉祥, 等. 三维众核片上处理器存储架构研究[J]. 南京大学学报(自然科学版), 2014, 50(3):330-334.
- [4] Puttaswamy K, Loh G H. Implementing caches in a 3D technology for high performance processors[C]// Computer Design: VLSI in Computers and Processors, 2005. ICCD 2005. Proceedings. 2005 IEEE International Conference on. South Korea; IEEE, 2005; 525-532.
- [5] Hsieh H C, Sun Y F, Yeh J C, et al. 3-D stacked memory system architecture exploration by esl virtual platform and reconfigurable stacking memory architecture in 3D-DSP SoC system[C]// Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on. Italy; IEEE, 2014;5012-5016.
- [6] Azarkhish E, Loi I, Benini L. A high-performance multiported L2 memory IP for scalable three-dimensional integration[C]// 3D Systems Integration Conference (3DIC), 2013 IEEE International. San Francisco, California; IEEE, 2013;1-8.
- [7] Lefter M, Voicu G R, Cotofana S D. A shared polyhedral cache for 3D wide-I/O multi-core computing platforms[C]// Circuits and Systems (ISCAS), 2015 IEEE International Symposium on. Portugal; IEEE, 2015;425-428.
- [8] Guthmuller E, Miro-Panades I, Greiner A. Architectural exploration of a fine-grained 3D cache for high performance in a manycore context[C]// Very Large Scale Integration (VLSI-SoC), 2013 IFIP/IEEE 21st International Conference on. Istanbul; IEEE, 2013;302-307.
- [9] Xie Y. Future memory and interconnect technologies [C]// Proceedings of the Conference on Design, Automation and Test in Europe EDA Consortium, San Jose; IEEE, 2013;964-969.
- [10] D O Son, H J Choi, H G Jeon, et al. Thermal-aware 3D Multi-core Process-or Design using Core and Level-2 Cache Placement[J]. International Journal of Control and Automation, 2013, 25(6):25-28.

的性能几乎相同的情况下,在低压下工作时的延时分别减少了 14.6%和 19.9%.同时,优化结构所引入的面积和功耗开销是相对较小的.

表 1 DFFASR 优化前后指标对比

	标准电压 下延时/ns	低电压 下延时/ns	功耗/ μW	面积
优化前	67.1	272.5	4.99	1
优化后	66.2	232.7	1.55 ~ 5.83	1.09
相对变化	-1.3%	-14.6%	-69%~+17%	+9%

表 2 LSUP 优化前后指标对比

	标准电压 下延时/ns	低电压 下延时/ns	功耗/ μW	面积
优化前	38.6	70.5	4.15	1
优化后	39.2	56.5	4.14 ~ 4.35	1.15
相对变化	+1.6%	-19.9%	-0.2%~+4.8%	+15%

而在最坏情况(SS,-40℃)下,DFFASR 和 LSUP 的优化结构在低压下的延时减少的幅度均更大,分别为 18.3%和 79.7%.

5 结束语

本文围绕将普通标准单元库改进为支持宽电压工作的标准单元库的方法,研究了剔除大扇入逻辑门对于高电源电压下的电路指标的影响,DC 综合实验表明剔除大扇入逻辑门对电路在高电压下的性能和功耗没有明显的负面影响,主要影响是在一定程度上增大电路的面积.分别对 32 nm 标准单元库中的异步置位复位 D 触发器单元和电平移位器单元进行了电路优化,在标准条件(TT,25℃)下,使得它们在低电压下的延时分别减少了 14.6%和 19.9%,解决了电源电压降至近阈值时性能恶化过于严重的问题.

参考文献:

[1] S J Surhud Khare. Prospects of Near-Threshold Voltage Design for Green Computing [C]// proceedings of the 26th International Conference on VLSI Design. Pune;IEEE, 2013. 120-124.

[2] A Steven K Hsu, Mark A Anders, Sanu K Mathew, et al. A 280 mV-to-1.1 V 256b reconfigurable SIMD vector permutation engine with 2-dimensional shuffle in 22 nm tri-gate CMOS [J]. JSSC, 2012; 48(1): 118-127.

[3] M A Himanshu Kaul, Steven Hsu, Amit Agarwal, et al. Near-threshold voltage (NTV) design—opportunities and challenges[C]// DAC. San Francisco, CA; IEEE, 2012: 1149-1154.

[4] Nitin Borkar, Vivek De, Shekhar Borkar et al. A 280mV-to-1.2V wide-operating-range IA-32 processor in 32nm CMOS[C]// ISSCC. San Francisco, CA; IEEE, 2012; 66-68.

[5] J P d G Yu Pu, Henk Corporaal, Yajun Ha. An ultra-low-energy multi-standard JPEG co-processor in 65 nm CMOS with sub/Near threshold supply voltage [J]. JSSCC, 2009; 45(3): 668-680.

[6] S L Wei Jin, Weifeng He, Zhigang Mao. A 230mV 8-bit sub-threshold microprocessor for wireless sensor network[C]// VLSISoC. Hong Kong; IEEE, 2011; 126-129.

作者简介:

王京睿 女,(1991-),硕士研究生.研究方向为低功耗数字电路设计. E-mail:wjr13@mails.tsinghua.edu.cn.

李翔宇 男,(1977-),博士,副研究员.研究方向是高效 VLSI/SOC,信息安全芯片等.

(上接第 71 页)

[11] Kim Y, Song Y H. Analysis of thermal behavior for 3D integration of DRAM[C]// Consumer Electronics (ISCE 2014), The 18th IEEE International Symposium on. Madrid Spain;IEEE, 2014;1-2.

作者简介:

怡 磊 男,(1991-),硕士研究生.研究方向为基于 TSV(硅

穿孔)的三维处理器结构以及三维电路设计.

E-mail:1311970283@qq.com.

单光宝 男,(1971-),研究员.研究方向为 TSV/MEMS 设备及其工艺技术.

肖建青 男,(1987-),博士研究生.研究方向为处理器架构研究.

刘 松 男,(1988-),博士研究生.研究方向为 TSV 三维串路设计.