

特色专题

从标准到实践: 芯粒库如何破解 chiplet 产业化困局

郝沁汾¹, 蒲波², 孙凝晖³

摘要 随着先进工艺节点设计成本呈指数增长, chiplet(芯粒)架构被视为突破单片集成经济瓶颈的关键路径。然而, 当前 chiplet 产业面临“标准悬空、门槛高企、生态稀少”的三重困境: UCIE(Universal Chiplet Interconnect Express)联盟等互连标准虽快速演进, 但从标准规范到可量产物理实现之间存在巨大工程鸿沟; 头部厂商普遍依赖自研专有方案, 标准缺乏真实产品的流片验证而难以完善; 高昂的开发成本使中小企业望而却步, 生态参与者极度稀少。提出“芯粒库”(chiplet library)概念, 即一种基于“FPGA 原型验证平台+chiplet 互连协议物理层验证卡+EDA 仿真补偿”的混合快速验证体系。现场可编程门阵列(field programmable gate array, FPGA)平台承载芯粒功能逻辑与互连协议数字层的可编程验证, 物理层验证卡以真实先进封装工艺提供封装内互连信道的直接测量, 电子设计自动化(electronic design automation, EDA)仿真工具补偿验证卡上缺失的完整 OC(on-chip)物理场环境(电源分配网络(power distribution network, PDN)噪声、热耦合、跨 die(裸片/晶粒)干扰), 并支持封装工艺变体推算, 三者互补构成完整的验证链路。分析了该方案在运行频率差异、接口互连协议(Interconnect protocol, IP)分割、缺乏完整系统级芯片(system on chip, SoC)物理环境等方面的已知局限并给出应对思路, 进而探讨了芯粒库在推动标准实践迭代、盘活存量芯片资源和催熟芯粒架构增量开发等方面的产业意义。

关键词 chiplet; 芯粒互连标准; UCIE; 芯粒库; FPGA 原型验证; 先进封装; EDA

根据 ARM(Advanced RISC Machines)公司招股书援引 IBS(International Business Strategise)公司数据^[1], 28 nm 节点的全流程芯片设计成本约为 4800 万美元, 到 7 nm 节点上升至约 2.49 亿美元, 5 nm 节点进一步达到约 4.49 亿美元, 3 nm 节点则高达约 5.81 亿美元, 而 2 nm 节点预计将突破 7.25 亿美元(图 1)。先进节点的掩模成本、设计规则复杂度和验证工作量均呈指

数增长, 使得单片集成(monolithic SoC(system on chip))路线对绝大多数芯片设计公司而言已不具经济可行性。

chiplet(芯粒)架构正是在此背景下被提出的。其核心思想是将传统的大型 SoC 拆分为若干个功能相对独立的小芯片, 各芯片可采用不同的工艺节点独立制造, 最终通过先进封装技术(如 2.5D 硅中介层、3D 堆叠或扇出型封装)集成为完整的系统。这种“异构集成”(heterogeneous integration)理念最早可以追溯到美国国防部高级研究计划局(Defense Advanced Research Projects Agency, DARPA)在 2017 年启动的

1. 中国科学院微电子研究所, 北京 100086

2. 宁波德图科技有限公司, 宁波 315800

3. 中国科学院计算技术研究所, 北京 100190

收稿日期: 2026-04-10; 修回日期: 2026-05-22

作者简介: 郝沁汾, 研究员, 研究方向为计算机系统结构, 电子邮箱: haoqinfen@ict.ac.cn

引用格式: 郝沁汾, 蒲波, 孙凝晖. 从标准到实践: 芯粒库如何破解 chiplet 产业化困局[J]. 科技导报, 2026, 44(12): 54-67; doi:10.3981/j.issn.1000-7857.2026.04.00030

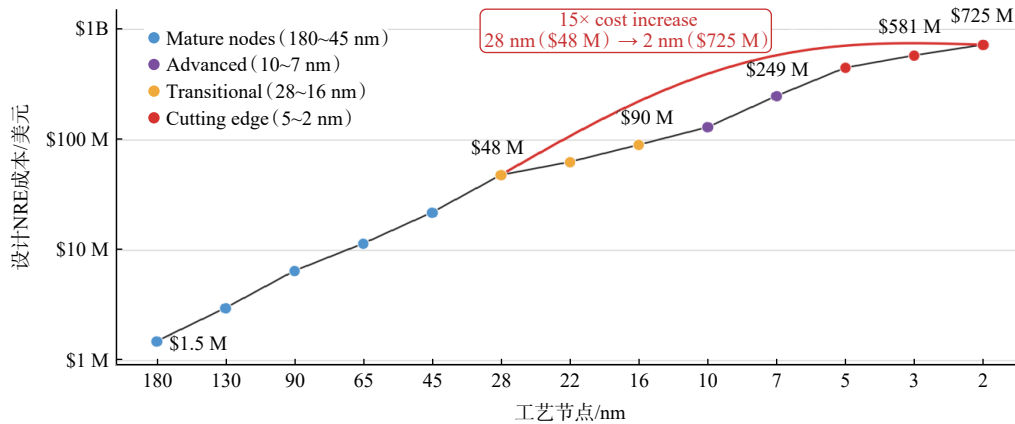


图1 芯片设计成本随工艺节点演进的增长趋势

Common Heterogeneous Integration and Intellectual Property Reuse Strategies(CHIPS)项目,该项目首次系统性地提出了 chiplet 的可复用性和标准化互连愿景^[2]。超威半导体公司(Advanced Micro Devices, Inc, AMD)于2019年发布的 Zen 2 架构是 chiplet 技术大规模商用的标志性事件——其采用 7 nm CCD(core complex die)与 12 nm IOD(I/O Die)的异构组合^[3],不仅降低了制造成本,而且显著提升了良率。AMD 在后续的芯片设计中都沿用了类似架构^[4-5]。

从标准层面看,chiplet 生态近年来发展迅速。通用芯粒互连技术(universal chiplet interconnect express, UCIe)联盟于2022年由英特尔、AMD、ARM、台积电等公司共同发起成立,旨在定义芯粒间的统一物理层和协议层互连标准。截至2025年,UCIe 规范已演进至 3.0 版本,支持的数据速率从 1.0 版本的 4 GT/s 跃升至 64 GT/s^[6],封装内互连密度也大幅提升。国内方面,计算机互连技术联盟(China Computer Interconnect Technology Alliance, CCITA)于2023年发布了首个面向国产生态的原生 chiplet 互连标准,为国产芯粒的互操作性奠定了规范基础^[7],后续国内连续制订了多个类似标准。

然而,标准的繁荣掩盖不了产业落地的滞后。产业实践表明,标准定义了理想的接口行为,但从标准规范到可量产的物理实现之间存在巨大的工程鸿沟。这一鸿沟的核心在于缺乏一个低成本、可快速迭代的实践平台——本文将其称为“芯粒库”(chiplet library),旨在通过构建快速验证体系,使标准在实践中得到检验与完善,打破 chiplet“有共识、无生态”的产业僵局。

1 chiplet 产业化的现实困境

1.1 标准多但实现少,落地遥遥无期

UCIe 联盟自2022年成立以来,成员已涵盖英特尔、AMD、台积电、三星、日月光、ARM、高通、Google 等全球主要半导体企业和云计算巨头。标准版本从 1.0 迅速演进至 3.0,覆盖了物理层(PHY)、die-to-die 适配层(D2D adapter)和协议层(protocol layer)3个层次(图2^[6])。

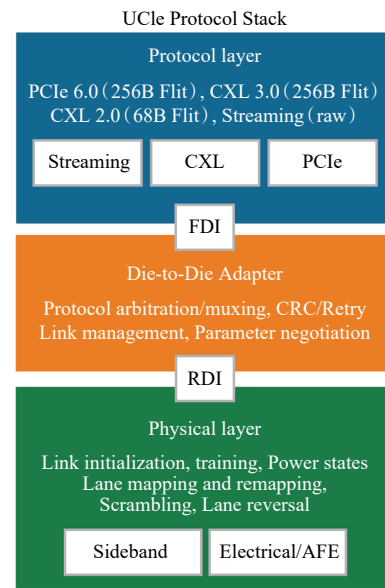


图2 UCIe 协议栈层次结构

但值得注意的是,chiplet 互连技术的复杂度远超传统的芯片间接口标准。以 UCIe 为例说明,UCIe PHY 的实现需要在极小的封装面积内集成高速收发电路,同时满足严格的时序裕量、眼图张开度和功耗约束,以实现优于片外 SerDes 约 20 倍的能效目标^[8]。

这意味着从纸面规范到硅片实现,需要经历复杂的电路设计、物理版图、信号完整性仿真和多次流片验证的漫长过程。对于大多数中小型芯片公司而言,独立完成 UCle PHY 的设计和验证几乎不可能——这需要专门的高速模拟电路团队和高昂的 chiplet 互连协议实现开发投入。

然而,标准在物理层定义上的完整性,并不等同于物理层实现的可落地性。根本问题在于:标准制定是“自顶向下”的过程,而产业化落地需要“自底向上”的工程实践——二者之间存在巨大鸿沟。标准缺乏足够的流片和封装实践来验证其物理层规定的合理性,迭代反馈机制尚不完善;与此同时,商业模式和 IP 许可体系尚未理顺,导致标准长期停留在纸面。

1.2 标准落地存在工程鸿沟,大厂依赖自有方案

积极参与标准制定的头部厂商,在自家产品中却普遍采用自研的专有互连方案。在英特尔的消费级和服务中央处理器(central processing unit, CPU)产品线(如 meteor lake、lunar lake)中,chiplet 互连协议仍为专有实现,尚未采用 UCle;AMD 同样如此——从 Zen 2 到 Zen 4 系列处理器均采用 Infinity Fabric 作为 chiplet 间的互连协议,该架构在带宽、延迟和一致性协议方面进行了深度定制优化,与 UCle 标准存在显著差异^[9-10]。在国内,大部分具有一定规模的芯片厂商也采用类似策略,在产品中使用自研的 chiplet 互连技术。这种局面的根源是一个深层博弈:头部厂商希望通过推动标准锁定生态(吸引更多参与者为其平台开发 chiplet),但又担心标准的不成熟给自己的产品带来影响和风险,结果形成恶性循环:标准因缺乏大厂真实产品的流片验证而无法完善,而标准不成熟又反过来加剧了大厂的观望情绪。

chiplet 互连标准在物理层之上的落地面临一个客观的技术困境——遵循同一份标准文本,并不等价于能够成功互操作。这一困境来自于标准中大量存在的“实现自定义”空间:UCle 1.x 对系统 bring-up 的管理协议没有标准定义,不同企业 PHY IP 在链路训练状态机(LTSM)的握手序列、PHY lane 检查顺序等实现细节上各有不同^[11];均衡电路(TX FFE、RX CTLE/DFE)的参数选取和训练算法同样允许厂商自行决定,不同 PHY 的均衡策略可能导致双方均衡协商无法收敛到有效工作点^[12]。正因如此,Intel 公司与 Ca-

dence 公司在 2023 年专门发起了一个寄存器传输级(register transfer level, RTL)互操作性仿真项目,对 2 家各自独立实现的 UCle PHY 进行逐状态的 LTSM 联合验证,并通过这一过程发现和修正了多处实现差异——这件事本身即说明“遵循标准”与“能够互操作”之间存在不可忽视的工程鸿沟^[11]。Siemens Digital Industries Software 的工程负责人 Rodriguez 也明确指出,UCle 1.1 仅能在 PHY 和 die-to-die 层实现互操作,软件与管理层的互操作直到 UCle 2.0 才被纳入标准^[12]。

IEEE P3468 标准工作组会议期间,某国际科技企业高级负责人曾就此问题表示,即便 2 颗芯片均严格遵循 UCle 规范,在实际对接中仍可能出现失败——这一判断与前述调研结论一致。此外,即使在封装工艺已经选定的情况下,整个生态链(IP、工具、流程、参考设计)需要高度对齐才能实现混搭互连,目前封装技术和配套工具链的成熟度仍不足以支撑真正的 chiplet 开放互连市场^[12]。

1.3 开发门槛极高,生态参与者过少

即使一家芯片设计公司决心采用 chiplet 架构,其面临的工程门槛也是极其高昂的。首先是设计复杂度:一颗 chiplet 芯片不仅需要完成核心功能的 RTL 设计、综合和布局布线,还需要额外设计 chiplet 互连接口(PHY+控制器),该接口的设计难度堪比一个独立的 IP 模块。其次是验证工作量:传统 SoC 的验证已经是芯片开发中最耗时的环节(通常占整个开发周期的 60%~70%),而多芯粒集成的系统需要额外验证芯粒间通信的功能正确性、协议一致性和物理层信号质量,验证工作量成倍增加^[13]。

多芯粒集成系统的验证工作量来源于互连电路与封装物理环境的强耦合。高速外设组件互连(Peripheral Component Interconnect Express, PCIe)PHY 工作在印制电路板(printed-circuit board, PCB)板级,信号以差分对形式在毫米至厘米量级的走线上传输,信号摆幅通常在数百毫伏,板级设计有充足的裕量容忍阻抗不连续和寄生效应;电源通过板级去耦电容就可以有效滤除噪声,热量由整块 PCB 和外壳散热,各物理效应之间相对独立,可以逐一分析和修正。UCle PHY 工作在封装内,物理环境的复杂性与 PCB 板级存在本质差异:接收端均衡手段弱,损耗造

成的裕量损失需靠发端驱动与均衡补偿,焊料凸点(bump)间距在微米至数十微米量级,相邻信号之间的串扰耦合极强,任何 bump 排布、走线间距或介质层厚度的细微偏差都可能导致信号完整性严重劣化^[14];封装内没有板级去耦电容的放置空间,电源分配网络(power distribution network, PDN)的去耦完全依赖极小面积内的片上电容和封装走线寄生电感之间的谐振特性,电源完整性分析和同步开关噪声(synchronous switching noise, SSN)抑制的设计难度远超板级^[15];同时,多颗裸片在极小封装面积内密集排布,热流密度极高,die 间热耦合效应显著,温度分布直接影响 PHY 模拟电路的偏置点和时序特性。信号完整性、电源完整性和热管理三者封装内高度耦合、相互影响,必须作为整体进行多物理场协同设计,而非像 PCB 设计那样可以分层独立处理。PHY 的模拟收发电路必须与特定封装工艺的上述全部物理参数协同设计,一旦封装参数发生细微变化,信号摆幅裕量、PDN 噪声容限或热偏置点就可能同时偏出设计窗口,需要回到流片层面重新调整——而不是像 PCIe 那样通过调整板级元件或更换走线拓扑即可修正。这种物理环境的本质差异,使得 UCIe PHY IP 的集成难度在起点上就与 PCIe 存在数量级的差距,而非仅仅是经验积累多寡的问题。

从经济角度看,单次先进节点流片(tape-out)的成本已相当可观——7 nm 节点约 1500 万~7500 万美元、5 nm 约 4000 万~1 亿美元^[1]。如果采用 chiplet 架构,虽然单颗 chiplet 可以使用更成熟的工艺节点以降低单次流片成本,但加上先进封装的额外费用,总体投入仍在千万至数亿美元量级。对于国内大量中小型 RISC-V 芯片设计团队而言(典型团队规模 20~50 人,年度研发预算在千万元人民币量级),“不敢试、试不起”成为常态。

这种高门槛导致了一个“马太效应”:只有英特尔、AMD、苹果、Google 等少数规模大的公司有能力和投入 chiplet 开发,生态参与者极度稀少。而参与者少又反过来抑制了 EDA 工具、封装工艺、测试方法等配套生态的发展——工具厂商和封测厂商缺乏足够的客户需求来推动工具链成熟和量产工艺优化,chiplet 技术的普惠价值无法释放,发展陷入停滞。

1.4 破局之道:构建低成本实践平台

综合以上三重困境,当前 chiplet 产业的核心矛盾可以归纳为一个相互强化的死循环:标准需要大量工程实践来暴露问题并迭代完善,但标准落地的高门槛导致实践者太少,标准长期停留在纸面;门槛高企又使生态参与者极少,工具链和封装工艺的配套成熟度难以提升,反过来进一步抬高门槛。大厂凭借自有技术体系闭门造车,中小企业望而却步,导致整个 chiplet 生态活力不足、普惠价值难以释放。

如何打破这个死循环?一种直觉上的回答是“再制定一个更好的标准”,但这恰恰是问题的根源所在——标准本身不是缺位的,缺位的是让标准落地的工程实践土壤。另一种直觉是“等待某个杀手级应用拉动需求”,但在门槛未降低之前,这种需求拉动只会进一步强化大厂垄断,而非激活中小企业的参与。

本文认为,真正的突破口在于构建一个低成本的实践平台:将分散的芯粒设计资源纳入共享验证环境,在流片前完成功能层与物理层的联合验证;让标准在真实使用中暴露问题并持续完善;让中小团队以极低门槛参与 chiplet 开发,积累工程经验,形成反哺标准和生态的正向循环。这一平台需要同时解决 3 个层面的问题:在技术层面,提供覆盖协议逻辑与封装内物理层的完整验证能力;在资源层面,提供复用社会存量芯片资源的接入机制;在生态层面,提供可持续积累和共享的芯粒 IP 资源池。本文将这样一个平台定义为“芯粒库”。其核心构成是“FPGA 原型验证平台+chiplet 互连协议物理层验证卡+EDA 仿真补偿”的混合快速验证体系。

2 芯粒库的可能实现路径探讨

2.1 路径一:统一标准先行,再开发标准化芯粒

理想化的路径是“标准先行”:全行业围绕统一的互连规范(如 UCIe)开发标准化 chiplet,自然形成一个可互换、可组合的开放市场——类似于 PC 行业的 USB 或 PCIe 标准所创造的即插即用生态。这也是 Intel 公司现在推动 UCIe 的主要思路。

这条路径在技术理念上无可挑剔,但忽视了半导体行业的基本现实。首先,chiplet 的互连涉及物理层、链路层和协议层 3 个层次的标准化,其中物理层

与封装工艺强耦合,不同封测厂商的工艺参数差异导致物理层标准的“一刀切”极为困难。芯片间互连标准化的难度远大于板级互连——后者只需定义连接器和信号规范,前者还需要约束封装内的微米级走线参数和热-电-力多物理场行为。

其次,标准的推广需要配套的 IP 生态。以 PCIe 标准为例,从 PCIe1.0 发布(2003 年)到被广泛采用花费近 5 年时间,期间需要 Synopsys、Cadence 等 EDA/IP 厂商开发成熟的 PHY IP 和控制器 IP,还需要芯片厂商在实际产品中验证这些 IP。UCIe 面临的挑战更大——它需要在封装内而非 PCB 板级实现,对 PHY IP 的面积、功耗和信号完整性的要求远高于板级标准。

最后,参与标准化需要企业投入大量资源。每家公司要基于新标准重新设计互连接口、进行新的流片验证,这意味着数百万到数千万美元的额外投入和 1~2 年的开发周期,意味着为了遵循一个标准,所有的公司要把以前的技术积累全部废弃。对于大多数公司(特别是中小企业)而言,在标准前景不明确的情况下进行如此大的前期投入是不可接受的。

2.2 路径二:纯 EDA 仿真验证

另一种看似可行的方案是完全在软件层面构建虚拟芯粒库:为每颗已有芯片建立行为级模型(SystemC/TLM)和物理级模型(SPICE/IBIS-AMI),通过仿真环境完成芯粒间互连的功能验证和信号完整性分析。

这条路径的局限性主要体现在 3 个层面。

1) 仿真精度与运行速度的根本矛盾。行为级仿真(TLM 事务级建模)运行速度快,可以在秒级时间内完成百万级事务的功能验证,但它忽略了所有物理层效应——时序、信号完整性、功耗等均未建模。物理级仿真(SPICE 电路仿真)精度高,可以逼近真实芯片的电气行为,但运行速度极慢——即使使用最先进的商用 EDA 工具和大规模计算集群,一个包含数亿晶体管 chiplet 的完整集成电路模拟仿真程序(simulation program with integrated circuit emphasis, SPICE)仿真,可能需要数周甚至数月的时间。

2) 模型获取的现实障碍。要为一颗芯片建立精确的仿真模型,通常需要获取该芯片的详细设计数据(网表、版图、工艺参数等)。但这些数据属于芯片公司的核心商业机密,几乎不可能共享给第三方平台。即使采用黑盒建模方法(基于芯片的 I/O 测量数据构建 IBIS-AMI 模型),模型精度也受限于测量条件和建模算法的能力。

3) 真实芯片行为的不可完全仿真性。真实芯片在运行中表现出大量难以在仿真中精确捕获的效应:工艺偏差(process variation)导致同一批次芯片之间的性能差异、老化效应(aging)导致性能随时间退化、温度依赖性导致芯片在不同热环境下行为不同。纯仿真环境固然可以参数化地考虑部分效应,但无法涵盖所有真实世界的复杂场景,验证结果与实际表现之间始终存在不可忽视的“仿真-硅片”(SIM-to-silicon)差距(图 3)。

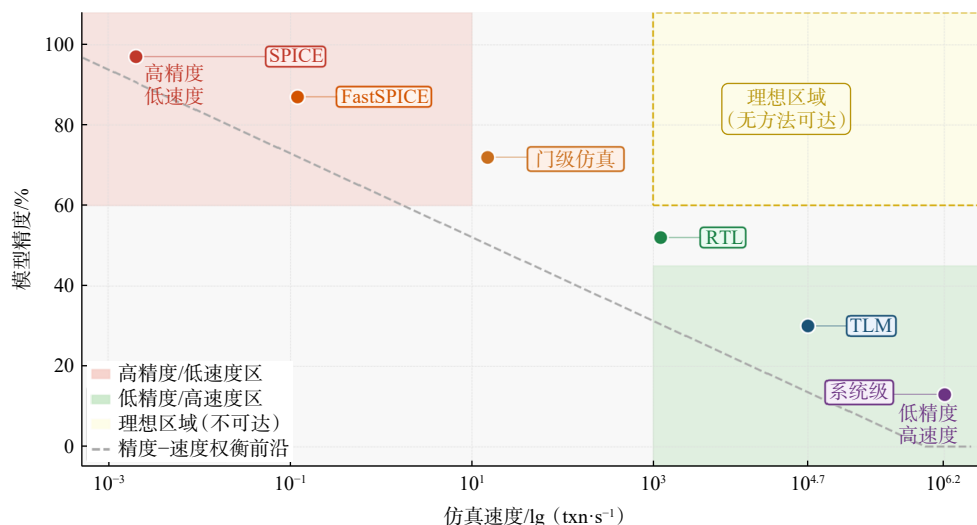


图 3 仿真精度和仿真速度之间的关系示意

2.3 路径三:一种基于混合验证思路的芯粒库实现方案

针对路径一“等标准太慢”和路径二“仿真不够真”2大问题,本文提出一种方案。其核心思想是“三位一体”:以FPGA平台提供功能层可编程验证环境、以chiplet互连协议物理层验证卡提供封装内物理层真实测量、以EDA仿真工具补偿其余物理层差异,三者互补构成完整的验证链路。

本文提出的混合验证方案在以下3个方面区别于已有工作。

1) 验证架构创新——“三位一体”联合验证体系。已有工作要么依赖纯FPGA原型(忽略物理层真实性),要么依赖纯EDA仿真(受限于模型精度和计算速度的矛盾),要么依赖单次定制流片的验证卡(成本过高、不可复用)。本文首次提出将三者协同工作:FPGA提供数字层快速可编程验证,验证卡以公共基础设施形式提供封装内物理层真实信道测量,EDA仿真提供验证卡上缺失的完整SoC物理环境补偿。三者的数据流构成完整的验证链路,而非独立运行后简单比较结果。

2) 工程方法创新——可复用验证基础设施。已有chiplet验证方案通常为特定产品一次性定制,验证成本由单个项目独立承担。本文提出的验证卡作为“芯粒库”的公共组件,一旦完成流片,所有使用相同chiplet互连标准的芯粒设计均可共用同一张卡进行物理层验证,单次流片成本由整个生态共同摊薄。

3) 方法论创新——将多物理场耦合为统一校正公式。已有研究分别从3个独立物理域分析PHY裕量劣化:PDN噪声与电源完整性^[15-16]关注芯片自身开关电流通过PDN阻抗引起的电源电压波动,热耦合与电热协同^[17-18]关注芯片功耗通过热阻累积引起的温升,跨die互连信号完整性^[14]关注相邻chiplet通过中介层电气与热路径的交叉干扰——但三者各自独立,尚未耦合为一个工程可用的定量模型。本文提出将这3个物理域的效应整合到统一的PHY眼高劣化校正框架中,分别对应3个劣化项:PDN电源噪声项、芯片自热温升项、跨die耦合干扰项。每一项的物理机制与前述3个已有研究方向一一对应。该校正框架使验证卡上孤立的PHY测量结果能够外推至完整SoC物理条件;本文在2.3.4节末段对该框架中跨

die热耦合系数 κ_{cross} 给出初步的数值层验证(结果落入Fengetal^[18]报告的硅中介层文献先验区间),其余分项的具体函数形式、参数标定与端到端实证将在后续论文中系统展开。

芯粒库混合验证体系自上而下分为3个功能域。

1) FPGA原型验证域。包含chiplet A和chiplet B的RTL逻辑、UCIe协议层和适配层。FPGA通过板级扩展连接器(FMC/HapsTrak子卡接口)与物理层验证卡连接。

2) 物理层验证域。验证卡内2颗die的chiplet互连PHY通过封装内bump互连信道通信,该信道为真实被测对象,提供眼图、时序裕量和误码率的外部测量点。

3) EDA仿真补偿域。根据FPGA的翻转活动率(VCD/SAIF),经功耗估算和多物理场联合仿真(PDN+热+眼图),将校正后的PHY裕量反馈至验证流程。

数据流向:FPGA翻转率—驱动EDA仿真—验证卡实测眼图校准—校正结果反馈(图4)。

2.3.1 FPGA原型验证平台:承载协议数字逻辑与功能验证并提供扩展通道

本文所提出的芯粒库混合验证方案的第1个支柱是FPGA原型验证平台。FPGA承载2类数字逻辑:其一是芯粒本身的SoC功能逻辑(RTL形式);其二是chiplet互连协议的数字层实现(协议层、适配层、链路训练状态机)。将待验证的芯粒逻辑以RTL形式部署在大容量FPGA(如AMD/XilinxVU19P)上^[19],任何互连协议参数的调整只需修改RTL代码并重新烧录比特流。这种可编程性使得发现协议问题—修改—重新验证的闭环从月级压缩到天级,为标准在实践中快速迭代奠定了工程基础。

这种验证方式相比纯EDA仿真具有三重决定性优势。

1) 真实性:互连协议的上层以完整的RTL代码在FPGA上运行,其行为与最终专用集成电路(application-specific integrated circuit, ASIC)实现的数字逻辑完全一致,而非简化的行为级模型。

2) 速度:FPGA的运行速度虽然低于ASIC(通常在10~200 MHz),但远高于RTL仿真(数赫兹到数十赫兹),速度差距可达4~6个数量级。这使得运行完整的操作系统启动、复杂软件负载和长时间压力测试

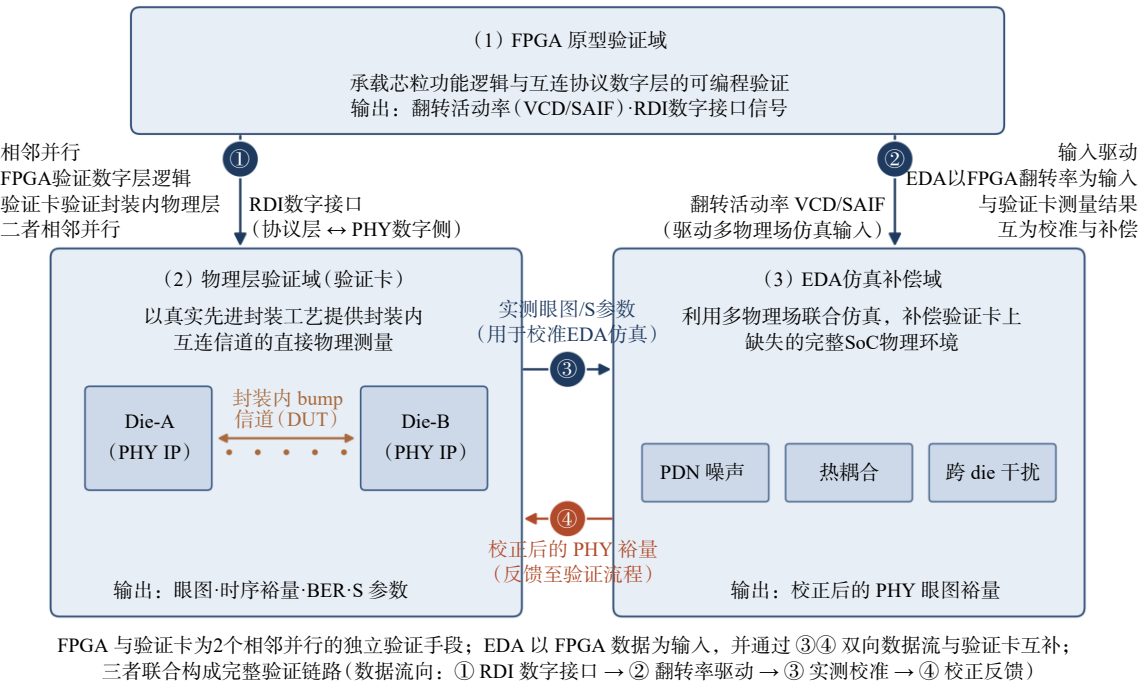


图4 芯粒库混合验证体系架构与数据流

成为可能。

3) 可迭代性: 芯粒逻辑修改后重新烧录 FPGA 只需数分钟到数十分钟, 而重新流片需要数月。这使得“发现问题—修改设计—重新验证”的迭代循环从月级缩短到天级。

另外, 在 FPGA 验证平台的硬件设计中, 一般具备扩展接口, 这为通过自定义协议和已经开发好的芯片资源对接提供了进一步的可能性。

第一类是标准协议外设接入接口, 以 PCIe“金手指”或线缆连接器为代表。几乎所有面向 SoC/ASIC 原型验证的高端 FPGA 平台均在板级提供 PCIeGen3/Gen4 标准接口。这类接口的意义在于: 用户在 FPGA 上开发自研芯片逻辑的同时, 可以将市场上已经量产并经过充分验证的 PCIe 外设(如网卡、AI 加速卡、存储控制卡等)直接引入验证环境, 在真实的业务流量和驱动栈下完成系统级联调, 而无需等待所有芯片组件均自主开发完毕再进行整合。

第二类是用户自定义协议扩展接口, 以高密度子卡连接器为物理载体, 允许用户通过子卡将任意自定义协议引入验证平台。本文后续提到的 chiplet 互连协议物理层验证卡, 即利用第二类接口中的单端并行通道接入 FPGA 验证平台, 实现对封装内基于源同步协议的 chiplet 互连接口的板级物理层验证。

2.3.2 用 chiplet 互连协议物理层验证卡覆盖封装内互连物理层的验证

FPGA 平台解决了“协议逻辑对不对”的问题, 但无法回答“封装内物理层行不行”——因为 FPGA 上的数字逻辑与真实封装内的 bump 阻抗、时钟偏斜分布、CTE 热应力之间存在本质断层。为填补这一盲区, 引入第 2 个支柱: chiplet 互连协议物理层验证卡(简称验证卡)。

验证卡的物理结构是一颗双裸片封装器件: 将 1 对 die-to-die PHY IP 以真实的先进封装工艺(有机中介层或 CoWoS)封装在一起, 2 颗 die 之间的封装内互连即为被测真实物理信道。验证卡以子卡形式插入 FPGA 平台: FPGA 上的 SoC-A 逻辑通过 RDI 数字接口连接验证卡上的 Die-A, SoC-B 逻辑连接验证卡上的 Die-B, Die-A 与 Die-B 之间的封装内 bump 信道承载真实的 die-to-die 数据流。工程师可以直接在 Die-A/Die-B 的外部接口测量眼图、时序裕量和比特误码率(bit error rate, BER), 所有数据来自真实封装环境, 不是仿真估算。理论上, 这个验证卡上可以封装多颗裸片, 以模拟多个 die 之间通过 chiplet 互连协议互连通信的场景, 本文中只讨论 2 个互连 IP 集成在一起的场景。

验证卡的核心价值在于可复用性: 它是芯粒库平

台的公共验证基础设施,一旦完成流片,所有使用相同 chiplet 互连标准的芯粒均可共用同一张卡进行物理层验证。单次流片成本由整个生态共同摊薄,而非由每家参与公司独立承担——这与 PCIe 合规测试设备在行业内共享使用的逻辑完全一致。芯粒库可以按主流封装工艺(有机中介层、CoWoS)分别维护对应版本的验证卡,覆盖主要的封装工艺体系。

验证卡以扩展子卡形式,可通过 3.3.1 节所述的第二类用户自定义协议扩展接口插入 FPGA 验证平台。

2.3.3 用 EDA 仿真工具补偿物理层差异,形成完整验证链路

FPGA 平台覆盖协议层和功能层,chiplet 互连协议物理层验证卡覆盖封装内物理层的直接验证,EDA 仿真工具的作用则是补偿前两者共同的物理盲区:FPGA 和验证卡均无法模拟完整 SoC 裸片中其他功能模块(CPU 核、缓存、内存控制器等)引入的封装内热耦合、PDN IR-drop 和 SSN,而 EDA 多物理场联合仿真正是填补这一缺失——将 FPGA 采集的翻转活动率作为输入,经功耗估算、PDN 瞬态仿真、热仿真和眼图仿真 4 步链路,将验证卡的测量结果外推至完整 SoC 物理环境下的 PHY 裕量估计。另外,如果已经存在验证卡的封装方式和客户的封装需求不同,有 2 种方法解决问题:一是围绕不同的封装工艺,做多种不同的芯粒互连协议物理层验证卡,有一定的代价;二是封装工艺变体推算——验证卡基于某种封装工艺(如 RDL 中介层)流片,EDA 的去嵌入-再嵌入算法用于将该测量结果推算至其他工艺变体(如 CoWoS),扩展覆盖范围。Bockelman 等^[20]在 1995 年提出的混合模式 S 参数理论,为差分信道的 S 参数表征提供了数学框架,是去嵌入-再嵌入算法在差分互连场景下的重要理论基础。类似的方法也被应用在第 5 代双倍数据率同步动态随机存取存储器(double data rate 5 synchronous dynamic random-access memory, DDR5)的芯片设计里^[21]。

综上,EDA 多物理场仿真在整个芯粒库架构中的角色可概括为:以功耗与热管理预测为手段,以补偿验证卡上缺失的完整 SoC 多物理场环境(PDN 噪声、热耦合、跨 die 干扰)为目的。

2.3.4 EDA 多物理场联合仿真的量化框架

EDA 仿真补偿的核心任务,是将 FPGA 采集的

翻转活动率映射为完整 SoC 物理环境下 PHY 眼图裕量的定量估计,该映射包含以下 4 个步骤。

第 1 步,翻转率—功耗。FPGA 运行负载时以 VCD 格式记录各模块的翻转活动率。经 glitch_factor(补偿 FPGA 的 LUT 结构吸收毛刺导致 ASIC 毛刺低估^[22])和 cg_factor(补偿 RTL 仿真无时钟门控与门级实现的差异^[23])2 项校正后,换算为 ASIC 条件下的模块级动态功耗分布,静态功耗由工艺库漏电模型另行补充。

第 2 步,功耗—PDN 噪声。将功耗分布经 $I=P/V_{dd}$ 转化为时变电流激励,输入双 chiplet+中介层的 PDN 网表(on-die、中介层、package 3 层 RLC 网络),由 ngspice 瞬态仿真求解 UCle PHY 电源引脚处的电压噪声波形,其中跨 die 耦合由中介层互阻抗 Z_{AB} 表征,含 SSN 在内的瞬态电源噪声分量由该瞬态仿真直接覆盖。

第 3 步,功耗—温度场。将功耗分布作为热源,采用有限差分法(finite difference method, FDM)求解稳态热传导方程,双 die 和中介层分别建模,通过 microbump 界面热耦合,输出双 chiplet 的温度场分布和跨 die 热耦合系数 κ_{cross} 。

第 4 步,PDN 噪声+温度—眼图裕量。将 PDN 噪声转化为 PHY 电源抖动,温度场转化为器件参数偏移(迁移率、阈值电压、互连电阻),采用 Volterra-Wiener 统计眼图方法分别评估 Level A 与 Level B 的眼图裕量。其中,本地 PDN 噪声与跨 die 互阻抗共同对应“PDN 噪声”物理场,分别构成 Level A 与 Level B 中的电源完整性劣化;芯片自热温升与跨 die 热耦合共同对应“热耦合”物理场,分别构成两级链路中的温度致劣化;而跨 die 互阻抗与热耦合在 Level B 中的耦合效应即对应“跨 die 干扰”物理场。3 类物理场在 Level A 与 Level B 2 级眼图裕量评估中的具体落点,正与 2.3 节创新点(3)所述的 PDN 噪声、热耦合、跨 die 干扰 3 项劣化——对应。

上述 4 步链路的每一步均有成熟的组件级工具(Verilator、ngspice、有限差分热求解器、scikit-rf 等)支撑,其输入输出关系、工具链选型和预期精度范围已在本团队的前期方法论研究中系统论证。本团队已设计端到端原型仿真链路,对典型双 chiplet 2.5D 硅中介层配置(10 mm×10 mm 中介层、2 mm 间距、硅热导率 $\kappa=148\text{ W/(m}\cdot\text{K)}$)执行了 6 个负载工况的有限

差分热场仿真,第3步求解器输出的跨 die 热耦合系数 κ_{cross} 区间为 0.566~0.583 K/W,落入 Feng 等^[18]报告的 2.5D 硅中介层文献先验区间 0.4~0.8 K/W,验证了热场仿真链路在数值层的有效性;6 个工况的稳态升温范围为 0.93~17.11 K。完整的多物理场参数标定(含 PDN 噪声谱、跨 die 信号耦合的定量验证以及温度扫描下的参数解耦)受当前行为级仿真模型的物理保真度限制,将作为后续工作在含真实 EM 提取与硬件测量的 sign-off 级验证平台上系统展开。

通过上述链路,FPGA 平台验证“逻辑对不对”,验证卡测量“当前封装工艺物理层行不行”,EDA 仿真则补偿完整 SoC 多物理场环境下的 PDN 噪声、热与跨 die 耦合差异,三者互补构成完整的验证链路。这种方案既避免了纯仿真的失真问题,又通过物理层验证卡获得了真实的封装内物理层数据作为校准基础。

2.4 已知局限与应对思路

任何工程方案都不可能完美无缺。坦诚面对混合验证方案(即路径三)的已知局限并提出务实的应对思路,比夸大其能力更有利于方案的推广和正确使用。

2.4.1 FPGA 运行频率与 SoC 存在差异

FPGA 平台与真实 ASIC 之间最显著的差异在于

运行频率。当代高性能处理器和加速器的工作频率普遍在 1~5 GHz 范围内,而 FPGA 上实现同等功能的设计通常只能运行在 10~200 MHz,频率差距可达 1~2 个数量级。

这一差异带来的直接后果:与绝对时钟频率强相关的性能指标——如实时吞吐量(GB/s)、单次操作延迟的绝对值(ns)——在 FPGA 平台上的测量结果无法直接映射为 ASIC 上的真实性能。例如,FPGA 上测得某数据路径的延迟为 50 个时钟周期,在 100 MHz 下对应 500 ns;但在真实 ASIC 上以 2 GHz 运行时,该延迟为 25 ns——二者的绝对值相差 20 倍。

然而,功能正确性验证不受频率差异的影响。数字逻辑的功能行为由 RTL 代码决定,与运行频率无关——在 100 MHz 下功能正确的设计,在 2 GHz 下同样功能正确(只要时序收敛)。此外,相对性能比例在很大程度上也与频率无关的。例如,方案 A 在 FPGA 上处理同一任务需要 100 万个时钟周期,方案 B 需要 80 万个时钟周期,二者的 20% 性能差距在 ASIC 上同样成立。因此,FPGA 平台完全可以用于方案比选和性能排序,只需通过频率比例换算即可推算绝对性能(表 1)。

表 1 FPGA 原型验证平台与 ASIC 频率差异及其对验证能力的影响矩阵表

验证维度	FPGA平台适用性	FPGA能力详情	ASIC参考	关键影响/缓解措施
功能正确性	✓ 适用	完整RTL/门级逻辑验证; 近乎完整的覆盖率	✓ Full	FPGA高度有效; 频率差距在此无影响
相对性能	✓ 适用	瓶颈识别; 相对吞吐量比较	✓ Full	缩放时序仍能按比例 反映瓶颈
绝对性能	✗ 不适用	FPGA: 50~200 MHz ASIC目标: 1~5 GHz (慢至1/5~1/100)	✓ Full	频率差距妨碍 直接时序签核
时序收敛	△ 部分适用	FPGA路径时序仅对FPGA 结构有效;不可移植到ASIC	✓ Full	ASIC需独立RTL综合 及STA
信号完整性 (SI)覆盖	✗ 不适用	仅PCB级互连;无倒装芯片 微凸点SI(需EDA工具)	✓ Full	由混合框架中的 EDA仿真覆盖
功耗/热分析	△ 部分适用	FPGA功耗≠ASIC功耗; 仅提供数量级估计	✓ Full	需EDA工具(RedHawk-SC / Cadence Celsius)
软件预加载 及驱动验证	✓ 适用	完整OS启动、驱动调试、 固件预加载 (FPGA的主要优势)	✓ Full	唯一支持预硅 SW验证的平台
调试可观测性	✓ 适用	ILA: 128通道, 512 K采样; JTAG协议分析器; 运行时波形捕获	✓ Full	FPGA比后硅提供更丰富 的运行时可观测性
协议合规性	✓ 适用	UClc/CCITA IP仿真; PCIe Gen4/5合规; 功能互操作测试	✓ Full	协议合规性可在 预硅阶段完整验证
✓ 适用—FPGA完全覆盖此维度 △ 部分适用—FPGA提供部分覆盖,需EDA工具补充 ✗ 不适用—无法在FPGA平台上单独验证				

2.4.2 芯粒互连接口 IP 在验证平台上被分割与真实封装中存在差异

在真实的 chiplet 封装中,芯粒互连接口 IP(包括协议层、适配层与物理层)是 SoC 裸片不可分割的组成部分,数字层与物理层在同一颗裸片上紧密集成,二者之间的内部接口(如 RDI)仅是片内的数字信号路径,无需跨越任何封装边界。

然而在本方案的验证架构中,这一原本一体化的接口 IP 被分割为 2 个部分:协议层与适配层等数字部分以 RTL 形式运行在 FPGA 之内,而物理层则位于 2.3.2 节所述的验证卡内裸片上,二者通过 FPGA 验证平台的板级扩展连接器(如 FMC/HapsTrak 子卡接口)相连。这意味着,原本在真实 SoC 裸片内部实现连接的数字-物理层接口,在验证平台上变成了需要穿越 PCB 走线和板级连接器的板级路径(长度在厘米量级),引入了真实片内路径所没有的寄生参数、阻抗失配和信号衰减。

这一切割带来的核心影响在于:PHY 数字侧接口上的信号时序裕量、驱动强度与终端匹配,在验证卡设计时均需针对这段额外的板级路径进行专门补偿,而非按照真实片内路径设计。若补偿不当,可能导致 PHY 链路训练异常或误码率偏高,产生与量产裸片行为不一致的验证结论。值得强调的是,这一局限仅存在于 FPGA 到验证卡之间的接口段,验证卡内部 2 颗裸片之间的封装内互连信道本身仍处于真实物理环境中,其测量结果不受影响。

应对思路是在验证卡与 FPGA 扩展接口的联合设计中,对板级路径进行阻抗匹配和信号完整性优化,并在 PHY 数字侧增加可配置的驱动强度与时序补偿参数,使板级路径的等效电气负载尽量逼近真实片内路径,将该差异对 PHY 链路行为的影响压缩至可接受范围之内。该问题事实上转化为了对 chiplet 互连协议物理层验证卡的进一步优化工作。

2.4.3 缺乏 SoC 裸片的完整物理环境

验证卡的封装体内仅集成了 die-to-die PHY IP 本身,其裸片面积和功能范围被刻意精简为只包含互连物理层所需的 I/O 电路。然而在真实的 SoC 裸片中,die-to-die PHY 只是众多功能模块之一,它与片上的 CPU 核、缓存、内存控制器、PCIe 控制器等模块共享同一颗裸片的 PDN、时钟树和基板。这种共享意

味着 PHY 的实际工作环境远比验证卡上的孤立环境复杂:其他模块的瞬态电流变化会通过共享 PDN 在 PHY 的电源引脚上产生噪声,多模块同时切换的 SSN(同步开关噪声)会叠加到 PHY 的参考电压和时钟上,片上其他高功耗模块的发热也会通过基板传导至 PHY 区域并改变其工作温度。

验证卡由于裸片上只有 PHY 电路,不存在其他功能模块,上述干扰源在验证卡上完全缺失。这导致验证卡上观测到的 PHY 链路质量指标——眼图张开度、时序裕量、误码率——实际上是在一个比真实 SoC 更为理想的物理环境下测得的,可能系统性地偏于乐观。在真实 SoC 中,当 CPU 处于高负载状态时,PHY 的实际误码率和时序裕量可能明显劣化于验证卡上的测量结果。

应对思路有 2 个层面。在验证卡设计层面,可以在裸片或 PCB 板上预置可编程的电流注入电路(如伪随机切换的负载电路),人为模拟其他模块对 PDN 和时钟的扰动,使 PHY 的工作噪声环境向真实 SoC 场景靠拢。在系统验证层面,则需依赖 2.3.3 节所述的 EDA 多物理场仿真工具:将 FPGA 运行典型负载时采集到的各模块翻转活动率(switching activity)以 VCD/SAIF 格式导出,经功耗估算工具结合目标工艺的单元模型换算为各模块的功耗分布,再以此为激励对完整 SoC 封装环境下的 PDN 噪声和热分布进行仿真,与验证卡的测量结果联合评估 PHY 链路在最坏工况下的裕量,形成完整的验证结论。

上述 2 条应对路径关系为互补而非择一:电流注入电路用于在线快速筛查异常工况,EDA 多物理场仿真用于离线精细评估完整工作负载下的裕量,二者共同构成分级验证体系。该联合评估的定量形式即为 2.3 节所述的多物理场校正方法,本文方法论创新由此在 SoC 物理环境补偿场景中得到落地。

2.4.4 热耦合对验证结果的影响与校正

本节聚焦于跨 die(chiplet 间)的热耦合,与 2.4.3 节中讨论的片上模块间热扰动互为补充。在 2.5D/3D 多芯粒堆叠中,热耦合对 PHY 链路裕量的影响不可忽视。Son 等^[7]指出 80℃ 下热效应对 UCIe32Gb/s 接口信号完整性的影响超过串扰。验证卡裸片仅含 PHY 电路(功耗<2 W),而在真实 SoC 中 PHY 与 CPU、缓存等高功耗模块共享基板,跨 die 温度梯度可达数

十摄氏度。

跨 die 热耦合可简化为 2 步因果链。第 1 步, Chiplet A 的功耗通过热传导引起 Chiplet B 的温升

$$\Delta T = k \times P_A \quad (1)$$

式中, k 为跨 die 热耦合系数(在稳态下等于等效热阻), 对于 2.5D 硅中介层典型值 0.05~0.3 K/W^[18]。

第 2 步, 温升通过 PHY 电路的温度敏感性转化为眼高劣化

$$\Delta E_H = S_T \times \Delta T \quad (2)$$

式中, S_T 为 PHY 眼高温度灵敏度, 典型值为 0.5~2.0 mV/K^[17](25~85℃ 近似常数)。将 2 步合并, 热耦合引起的眼高劣化为 $S_T \times k \times P_A$ 。

2.4.5 仿真模型保真度边界

2.3.3 与 2.3.4 节所述的 EDA 多物理场仿真链路当前在行为级模型上实现, 存在以下保真度边界, 本文据此明确区分“已验证项”与“后续工作项”。

1) 已验证项: 跨 die 热耦合系数 κ_{cross} 的数值层一致性已通过有限差分热仿真在 6 个负载工况上得到验证(实测区间 0.566~0.583 K/W, 落入 Feng 等^[18]报告的硅中介层文献先验区间 0.4~0.8 K/W), 表明热场仿真链路在数值层与已发表的硅基中介层热建模结果一致。

2) 后续工作项: (1) 完整的 PHY 眼图统计当前基于 Volterra-Wiener 行为级模型, 部分温度依赖项以固定公式参数注入而非完全从噪声/抖动统计推导, 后续需将温度依赖项纳入统一的噪声/抖动统计框架, 从底层物理量直接推导; (2) PDN 网表采用 lumped RLC 与解析跨 die 互阻抗, 未含分布式电磁提取, 后续需引入三维电磁场提取, 建立分布式 PDN 模型以提升高频精度; (3) chiplet 互连 PHY 以行为级功耗模型注入, 可综合 RTL 受专有 IP 限制不可公开获取, 后续将基于自研的 chiplet PHY IP, 用其可综合 RTL 替换行为级模型; (4) 跨 die 互连信号完整性与 PHY 眼图裕量的端到端定量验证尚未完成, 后续需搭建该端到端验证平台, 完成温度扫描参数解耦并与 sign-off 级仿真交叉比对。

上述保真度边界源于当前行为级仿真平台的方法论起点, 因此, 在 2.3 节创新点(3)所述的“统一校正框架”中, 仅声称框架的物理形式定义清晰、跨 die 热耦合一项已获文献先验数值层一致性验证; 完整的模

型参数实证、跨工艺节点泛化评估及与硅前 sign-off 工具链(如真实 EM 提取的 S 参数模型与 IBIS-AMI 接收机模型)的对标, 将作为后续论文工作展开。

3 芯粒库对产业发展的意义

3.1 让芯粒互连标准在实践中得到检验和完善

芯粒库的一个关键产业价值在于为标准演进提供实践土壤。当前 chiplet 标准面临的核心问题是“制定—无人使用—无法完善”的恶性循环。芯粒库平台通过降低实践门槛, 希望打破这一循环。

当大量开发者在芯粒库验证平台上使用这些标准协议进行实际的芯片互连验证时, 标准中的各类问题——协议状态机的边界条件遗漏、链路训练序列的鲁棒性不足、错误处理机制的覆盖不全等——就会在真实工作负载下快速暴露出来。这些实践中发现的问题可以直接反馈给标准制定组织, 推动标准进行有针对性的修订和完善。

3.2 芯粒库的双重实践路径: 盘活存量与催熟增量

从“如何让芯粒库真正利用社会上已有芯片资源”的视角看, FPGA 验证平台所独有的扩展卡体系, 是把抽象的芯粒库能力落到具体工程场景中的关键抓手。在这一框架下, 可以明确区分 2 类相互独立的应用场景: 其一是通过 2.3.1 节中的大部分 FPGA 验证平台厂商所具有的第一类接口有效利用现有已完成设计的芯片资源, 盘活市场存量; 其二是通过 2.3.1 节中的第二类接口在芯粒架构下开发更大规模的新芯片, 催熟未来增量。两类场景在工程路径上彼此独立, 都依托同一套 FPGA 加扩展卡基础设施, 但各自解决不同的工程问题。

第 1 类场景面向的是当前已经量产并广泛部署的标准芯片资源, 其核心是把社会上已经存在的并通过 PCIe 接口提供连接的芯片——例如, 以太网网卡、AI 加速卡、存储控制卡等——通过 FPGA 验证平台上的标准扩展卡接口, 引入到正在开发中的芯片验证环境。这种方式使得大量已量产的 PCIe 设备, 在芯粒库驱动的平台被自然地当作可调用的外部功能模块反复复用, 从而以极低的工程成本盘活了存量芯片资源, 在芯粒库上完成功能验证后, 拟开发的芯片可以完成流片并通过封装技术和已经开发好的芯片

实现对接。

第2类场景则完全面向芯粒架构自身的演进需求,解决的是“如何在立即流片的前提下开发和验证更大规模的芯粒化系统”这一核心问题。例如,一个团队已经拥有可工作的四核 RISC-V 芯片,希望在下一代产品中将规模扩展至八核并引入芯粒架构。团队可以在 FPGA 验证平台上开发 2 个 4 核的功能芯粒,并通过验证卡实现 2 个功能芯粒之间的互连验证。在这样的环境中,团队可以在无需进行先进封装流片的前提下,对“4+4 核”的芯粒化架构进行完整的功能验证和架构评估,使 SoC 芯片从单芯粒到多芯粒的架构演进,成为可以在实验室中反复迭代打磨的工程过程,而非“一次性、不可回头”的流片赌博。

4 结论

chiplet 技术不缺标准、不缺方向,缺的是一个让更多人低成本参与实践的入口。当前产业陷入“标准悬空、门槛高企、生态稀少”的三重困境,根本原因在于缺乏一个将标准、芯片和开发者有效连接起来的实践平台。

芯粒库正是这样一个平台,采用“FPGA 原型验证+chiplet 互连协议物理层验证卡+EDA 物理仿真补偿”的混合验证体系,在技术层面实现了“逻辑验证+物理预测”的完整链路;在产业层面,通过“盘活存量”和“催熟增量”双重路径,将 Chiplet 架构从少数规模较大公司的专属能力,变成全行业可参与的共享基础设施。虽然混合方案在运行频率、接口 IP 被分割、缺乏完整物理环境等维度存在已知局限,但一方面这些工作可以成为进一步的研究工作,并且问题的解决具备初步可行性;另一方面它在流片前提供了当前条件下高性价比的验证手段,能够将大部分风险前置化解。

展望未来,随着芯粒库平台上用户和芯片资源的不断积累,以及 UCIE、国内的芯粒互连标准等在实践中持续完善,chiplet 技术有望真正从“有共识”走向“有生态”,从“空中楼阁”走向产业落地。

参考文献 (References)

[1] Arm Holdings plc. Form F-1注册声明[EB/OL]. (2023-08-

21) [2026-06-14]. <https://www.sec.gov/Archives/edgar/data/1973239/000119312523216983/d393891dfl.htm>.

- [2] Olofsson A, Green D S, Demmin J. Enabling high-performance heterogeneous integration *via* interface standards, IP reuse, and modular design[C]. International Symposium on Microelectronics, 2018, 2018(1): 246-251.
- [3] Naffziger S, Lepak K, Paraschou M, et al. AMD chiplet architecture for high-performance server and desktop products [C]//Proceedings of IEEE International SolidState Circuits Conference (ISSCC). Piscataway, NJ: IEEE, 2020: 44-45.
- [4] Munger B, Wilcox K, Sniderman J, et al. "Zen 4": The AMD 5nm 5.7GHz x86-64 microprocessor core[C]//Proceedings of IEEE International Solid-State Circuits Conference (ISSCC). San Francisco: IEEE, 2023: 38-39.
- [5] Cohen B, Subramony M, Clark M. Next generation "Zen 5" core[C]//Proceedings of IEEE Hot Chips 36 Symposium (HCS). Piscataway, NJ: IEEE, 2024: 1-27.
- [6] UCIE Consortium, Universal Chiplet Interconnect Express (UCIE) Specification, Revision 3.0, 2025[EB/OL]. [2026-06-14]. <https://www.uciexpress.org/specifications>.
- [7] 中国电子工业标准化技术协会. 小芯片接口总线技术要求: T/CESA 1248—2023[S]. 北京: 中国电子工业标准化技术协会, 2023.
- [8] Das Sharma D, Pasdast G, Qian Z G, et al. Universal chiplet interconnect express (UCIE): An open industry standard for innovations with chiplets at package level[J]. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2022, 12(9): 1423-1431.
- [9] AMD Instinct MI200 series accelerator and node architectures: 3rd generation AMD Infinity architecture[EB/OL]. [2026-06-14]. <https://hc34.hotchips.org/assets/program/conference/day1/GPU%20HPC/HC2022.AMD.AlanSmith.v13.Final.20220818.pdf>.
- [10] Smith A, Alla V. AMD instinct MI300X generative AI accelerator and platform architecture[C]//Proceedings of IEEE Hot Chips 36 Symposium (HCS). Piscataway, NJ: IEEE, 2024: 1-22.
- [11] Cadence Design Systems. "UCIE interoperability between intel and cadence." Cadence community blog, IP Solutions [EB/OL]. [2026-02-25]. https://community.cadence.com/cadence_blogs_8/b/ip/posts/ucie-interoperability-between

- intel-and-cadence.
- [12] Bailey B. "Chiplets still a challenge with UCIE 2.0." [EB/OL]. [2026-02-25]. <https://semiengineering.com/chiplets-still-a-challenge-with-ucie-2-0>.
- [13] McLaurin T, Cron A. Applying IEEE test standards to multi-die designs[J]. *IEEE Design & Test*, 2022, 39(5): 7-16.
- [14] Shi B, Zhou Y, Sun H F, et al. Statistical method for eye diagram simulation in high-speed link nonlinear system applications[J]. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2024, 14(11): 2050-2061.
- [15] Totorica N, Li F. Signal and power integrity challenges for high density system-on-package[J]. *Semiconductor Science and Information Devices*, 2022, 4(2): 1-9.
- [16] Zhi C L, Dong G, Yang D G, et al. Electrical and thermal characteristics optimization in interposer-based 2.5-D integrated circuits[J]. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2025, 33(3): 627-637.
- [17] Son K Y, Kim K W, et al. The significance of thermal-aware universal chiplet interconnect express (UCIE) interface design in 2.5D/3D ICs[EB/OL]. [2026-02-25]. <https://doi.org/10.1109/EDAPS58880.2023.10468234>.
- [18] Feng J Y, Chen C, Fu R, et al. The model and influence factors of thermal interaction in chiplet 2.5D integration [C]//Proceedings of 23rd IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm). Piscataway, NJ: IEEE, 2024: 1-5.
- [19] Xilinx. Virtex UltraScale+FPGA data sheet: DC and AC switching characteristics: DS923[EB/OL]. [2025-06-14]. <https://docs.amd.com/r/en-US/ds923-virtex-ultrascale-plus>.
- [20] Bockelman D E, Eisenstadt W R. Combined differential and common-mode scattering parameters: Theory and simulation[J]. *IEEE Transactions on Microwave Theory and Techniques*, 1995, 43(7): 1530-1539.
- [21] Li A, Wang J, Xu Y, et al. A DDR5 interposer de-embedding method based on transfer function[C]//Proceedings of 2024 IEEE 33rd Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS). Piscataway, NJ: IEEE, 2024.
- [22] Najm F N. Transition density: A new measure of activity in digital circuits[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 1993, 12(2): 310-323.
- [23] Synopsys Inc. SpyGlass Power: RTL power optimization (CGR/CGE metrics)[EB/OL]. [2026-06-14]. <https://www.synopsys.com/verification/static-and-formal-verification/spyglass/spyglass-power.html>.

From standard to practice: How a chiplet library can break the industrialization dilemma of chiplets

HAO Qinfen¹, PU Bo², SUN Ninghui³

1. Institute of Microelectronics, Chinese Academy of Sciences, Beijing 100086, China

2. Ningbo Detu Technology Co., Ltd., Ningbo 315800, China

3. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

Abstract As design costs at advanced process nodes grow exponentially, the chiplet architecture is regarded as a key path to break through the economic bottleneck of monolithic integration. However, the current chiplet industry faces a threefold dilemma of "suspended standards, high barriers, and a sparse ecosystem": although interconnect standards such as the Universal Chiplet Interconnect Express (UCIe) consortium are evolving rapidly, a vast engineering gap remains between the standard specification and a manufacturable physical implementation; leading vendors generally rely on in-house proprietary solutions, so the standards lack tape-out validation from real products and are difficult to refine; and the high development cost deters small and medium-sized enterprises, leaving ecosystem participants extremely scarce. This paper proposes the concept of a "chiplet library", a hybrid rapid-verification system based on an FPGA prototyping platform, a chiplet-interconnect physical-layer (PHY) verification card, and EDA simulation compensation. The field-programmable gate array (FPGA) platform carries the programmable verification of chiplet functional logic and the digital layer of the interconnect protocol; the physical-layer verification card uses a real advanced-packaging process to provide direct measurement of in-package interconnect channels; and electronic design automation (EDA) simulation tools compensate for the complete on-chip (OC) physical-field environment missing on the verification card—including power distribution network (PDN) noise, thermal coupling, and cross-die interference—while supporting extrapolation to packaging-process variants. Together the three form a complete verification chain. The known limitations of this approach, in terms of operating-frequency differences, interface intellectual property (IP) partitioning, and the lack of a complete system-on-chip (SoC) physical environment, are analyzed and corresponding strategies are proposed. The industrial significance of the chiplet library is then discussed with respect to driving the iterative practice of standards, revitalizing existing chip resources, and accelerating the maturation of incremental chiplet-architecture development.

Keywords chiplet; chiplet interconnect standard; UCIe; chiplet library; FPGA prototyping verification; advanced packaging; EDA ●



(责任编辑 王微)