



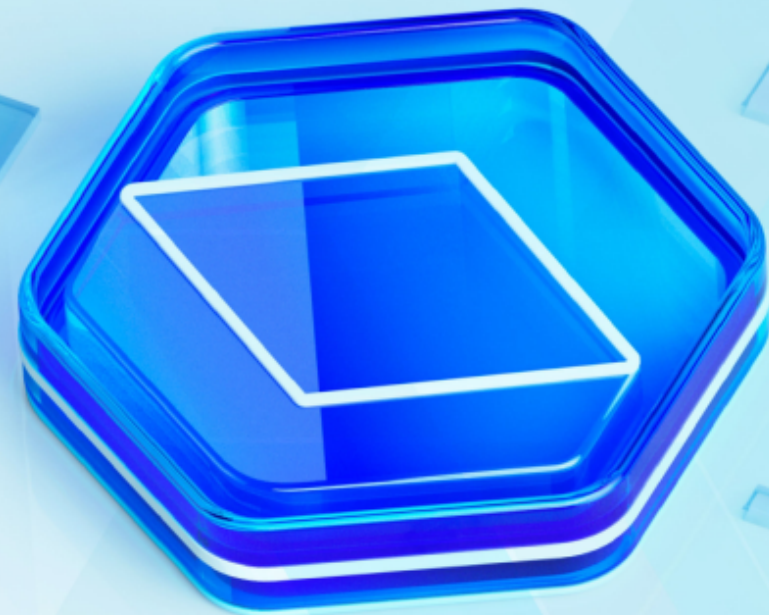
面向高性能RISC-V多核处理器的大 型级联FPGA验证方法学

—香山昆明湖16核CPU系统实践

李贤飞, 高级工程师, 北京开源芯片研究院

曹梦侠, 验证产品总监, 合见工软

2025.06.26



北京开源芯片研究院
BEIJING INSTITUTE OF OPEN SOURCE CHIP

X



合见工软
UNIVISTA

- 香山处理器发展介绍
- 昆明湖多核CPU验证的三大挑战
- 多核验证方法学
- 实践方法论与技术创新点
- 项目成果
- 结论与展望

"香山"系列处理器的技术进阶之路

"香山"系列处理器的技术演进，是一条持续向高性能计算迈进的清晰路径，体现了我国自主可控高性能RISC-V处理器的发展历程。

第一代：雁栖湖

架构探索与奠基的阶段，成功实现了乱序执行架构，验证了高性能RISC-V处理器的设计方法学，为后续迭代打下了坚实基础。

第二代：南湖

性能实现飞跃，对标ARM Cortex-A76级别，是一款业界公认的高性能RISC-V处理器核，标志着"香山"进入了可用、好用的阶段。

第三代：昆明湖

迈向数据中心级别，性能对标ARM Neoverse N2，是为服务器、数据中心等高性能计算场景设计的全新架构。

🔦 技术里程碑

香山系列处理器从首代验证高性能RISC-V设计可行性，到第三代面向数据中心级应用，实现了从单核到多核架构的跨越，展现了我国处理器技术的自主创新能力。

- ✓ 架构探索
- ✓ 乱序执行

第一代：雁栖湖

- ✓ 性能飞跃
- ✓ 对标A76
- ✓ 面向终端设备

第二代：南湖

- ✓ 对标N2高性能核心
- ✓ 面向数据中心应用

第三代：昆明湖

数据中心级RISC-V SoC的新征程：“昆明湖”设计目标与温榆河NoC规划

设计目标：“昆明湖”不仅仅是核心性能的提升，更是面向高性能计算场景的系统级SoC架构。昆明湖系列支持64核高性能SoC，面向数据中心级应用。

关键技术规划：

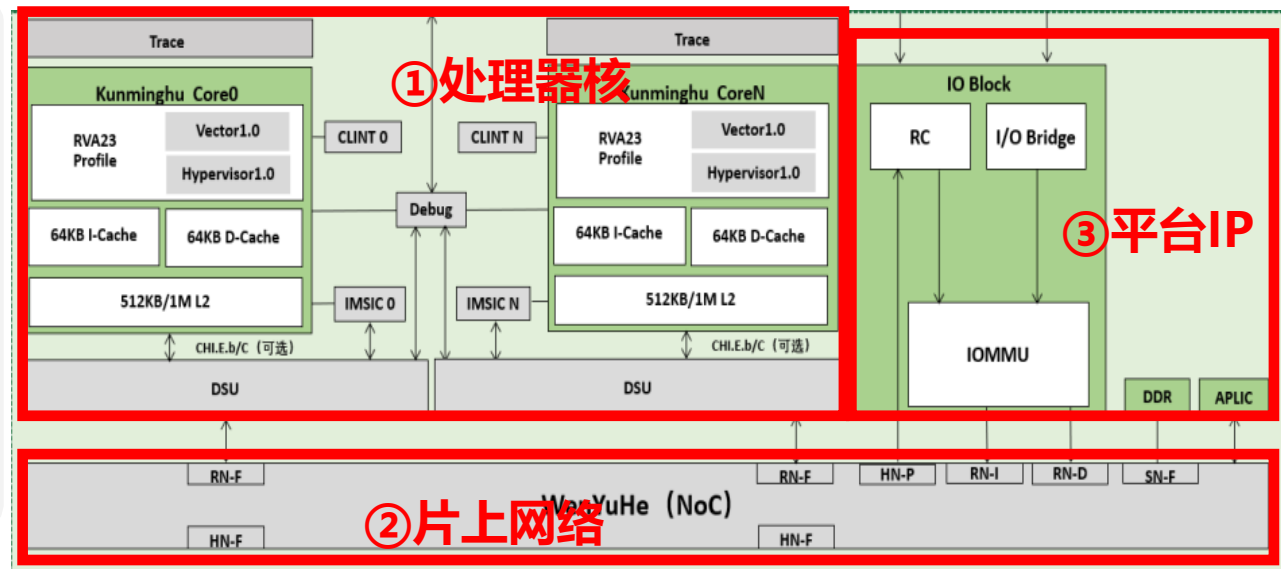
- **多核可扩展性：**架构设计上为未来扩展到64-256核预留空间
- **高速互联总线：**全新的高带宽、低延迟NoC支持高效数据交互
- **强一致性内存系统：**大规模多级缓存与目录式缓存一致性协议

温榆河NoC突破：2024年5月，开芯院发布了全球首个开源大规模片上互联网络IP——“温榆河”。经过18个月开发，成功完成支持64核互联的NoC IP开发和验证。

关键指标

- ✓ 已验证64核互联能力
- ✓ 设计支持12x12网格最大支持256个处理器核互联
- ✓ 采用标准CHI总线接口，打破ARM垄断
- ✓ 全球首个开源大规模片上互联网络IP

服务器场景交付



64个处理器核心 温榆河NoC 平台IP

温榆河NoC：高带宽片上互联网络 | 支持64-256核 | 标准CHI接口 | 可预留IO接口

高性能

低延迟高吞吐设计，支持Xeon级处理能力

安全可控

开源架构，自主可控，无国际限制

三大挑战

❖ 规模大 (Scale):

- 多核CPU系统，包含复杂的总线、多级缓存和DDR4、UART等接口，逻辑规模远超单片FPGA容量。
- 传统FPGA平台资源不足，且手动分割设计的方式困难且极易出错，Porting周期极长；

❖ 性能要求高 (Performance):

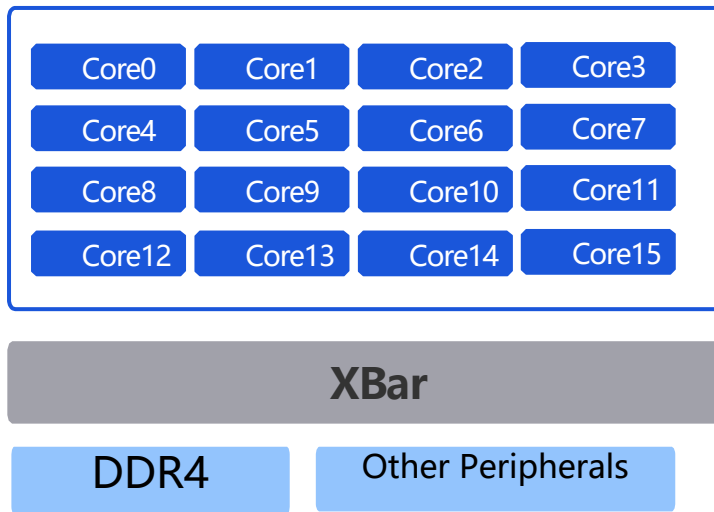
- 目标是在FPGA上实现足够高的运行速度（最终目标10MHz），以运行操作系统并进行有意义的软硬件协同验证。

❖ 调试困难 (Debug):

- 多核协同复杂：需保障 16 核缓存一致性、总线完整及高负载下调度优化。
- 在多核系统中，如何快速定位和复现因跨核交互、缓存一致性等引起的深层次Bug。缺乏专用调试工具，问题定位慢。

与合见合作之前，香山系列多核CPU验证亟需大规模 FPGA 平台与自动化工具链。

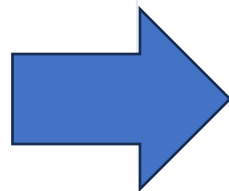
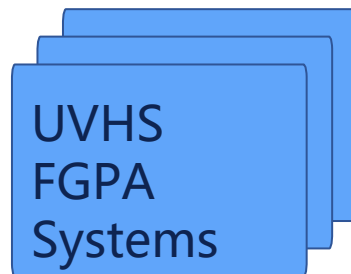
开芯院



合见

目标:

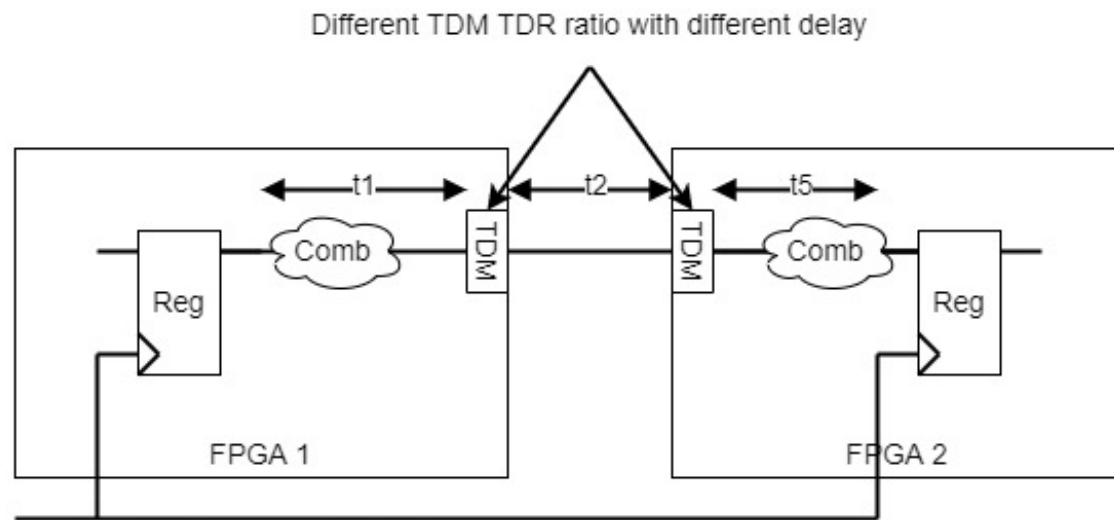
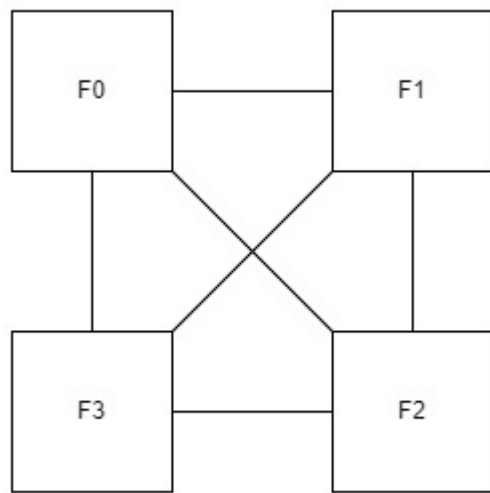
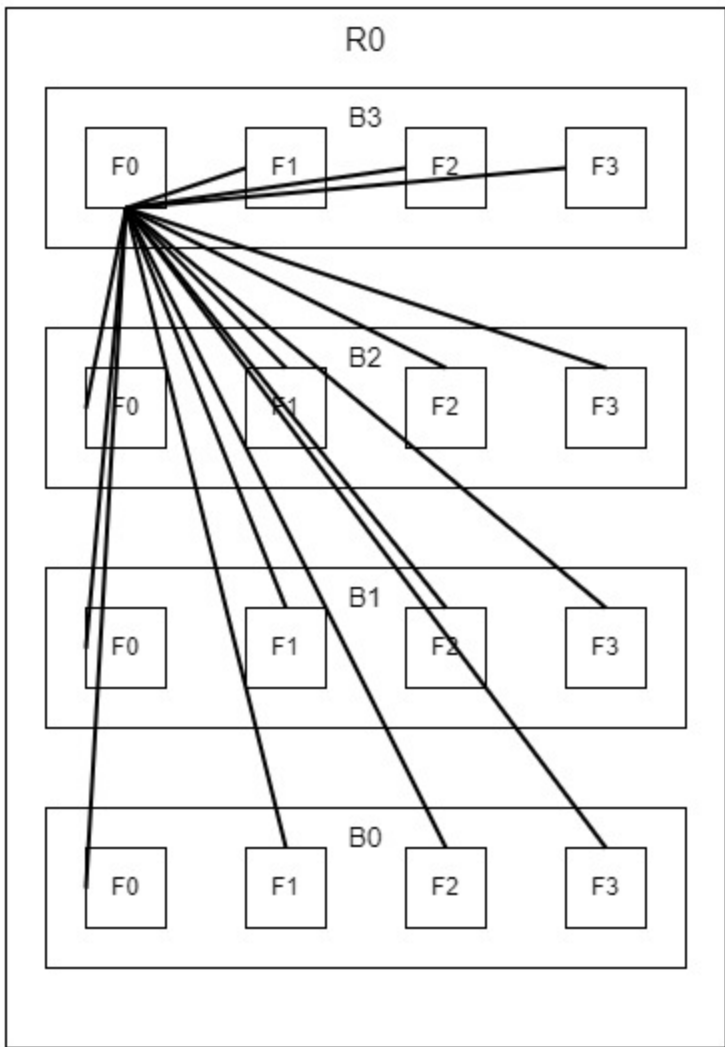
- 16核完整部署
- 1/2/4/16 CPU核多版本OS启动
- 全自动化分割及编译流程
- FPGA跨片频率达到10MHz性能



大型芯片原型系统的规划和设计 - 全互联架构

大型芯片原型系统的互联结构随着设计规模的增加需要做调整

- 任意两颗FPGA之间至少需要一根连线
- 能支持的最多FPGA的数量是FPGA互联连接器数量+1
- UVHS系统中单颗FPGA拥有24个LVDS 连接器以及8个GT连接器
- 全互联系统的Timing分析相对简单
- 随着两颗FPGA之间能分到的连线数量的减少, TDM的比例会增大, 系统运行的速度会降低



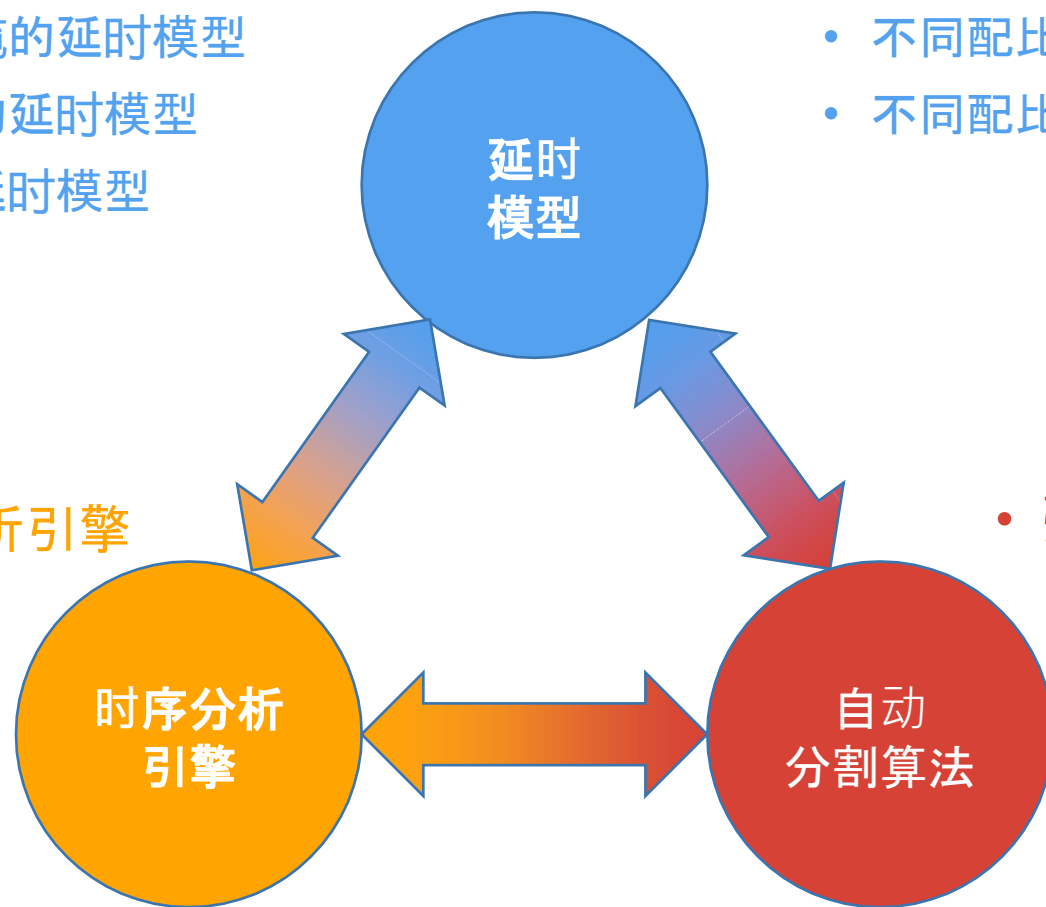
自动分割技术是大型硬件验证平台解决大设计验证的关键

- 硬件验证平台需要精确的延时模型
 - 不同长度全局时钟线缆的延时模型
 - 不同长度的互联线缆的延时模型
 - 各种主板和子板上的延时模型

- 不同种类TDM IP的延时模型
 - 不同配比 BitSlice TDM IP的延时模型
 - 不同配比 GTY TDM IP的延时模型

- ASIC流片级别的时序分析引擎

- 强大的自动分割算法
 - 时序驱动自动分割算法
 - 自动绕线TDM配比算法



方法学概述:

- 基于本项目经验总结, 提出了一套系统化的多核处理器 FPGA验证方法论, 涵盖从设计移植到系统调试的完整流程, 为 RISC-V多核验证提供了标准化路径。

四步验证方法论

- **设计移植与适配 (Porting):**
 - 建立了ASIC-to-FPGA的自动化适配流程, 包括时钟树、存储模型、接口IP转换的标准化处理方法。
- **编译迭代效率与资源优化使用的最大化平衡:**
 - 面向大规模设计, 在资源占用率与高效率编译迭代之间动态取舍, 寻找最大化的收益平衡点。
- **渐进式启动策略 (Bring-up):**
 - 创新提出了核数逐步扩展的验证策略, 结合单核到多核的系统复杂度梯度, 有效降低了系统调试难度。
- **软硬协同调试技术 (Debug):**
 - 基于硬件仿真与后门加载技术的协同调试方案, 精准解决系统启动故障, 并实现内核加载速度量级突破

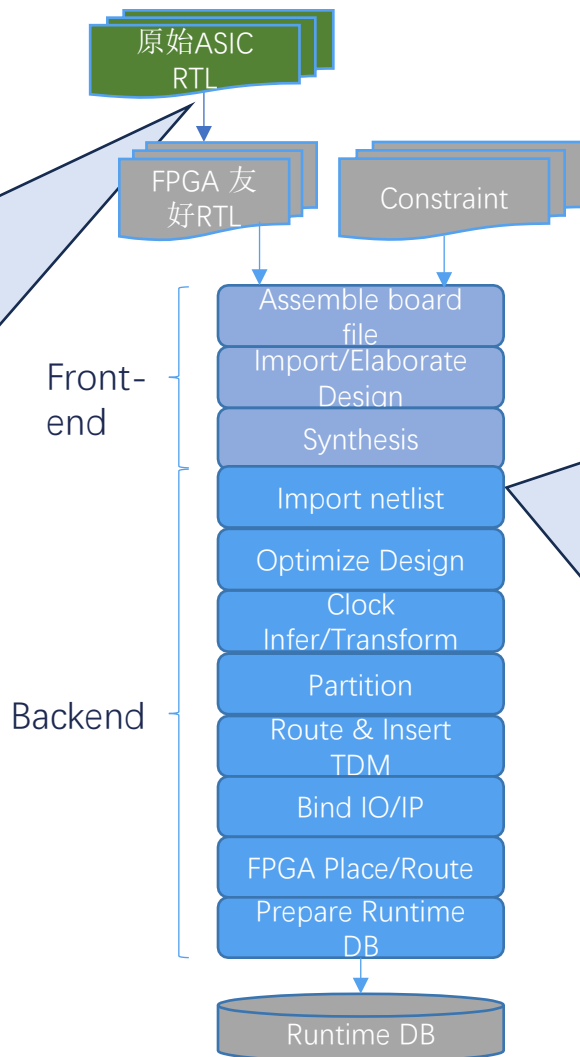
实践方法论1 - 快速高效的设计移植 (Porting)

ASIC到FPGA的高效移植挑战

- 将为ASIC设计的RTL代码高效、可靠地迁移到FPGA平台上，是验证过程中的首要任务。

我们的工作:

1. 时钟/复位结构修改
将CPU主时钟和VIO复位替换为FPGA友好的全局时钟/复位。
2. 存储模型替换:
将ASIC Memory替换为工具可识别处理的Memory Model。



我们的工作:

3. IP核配置适配:
重新配置Xilinx MIG IP以适配UVHS的PDDR4子卡, 并处理为DCP。
4. Boot ROM处理:
将客户的BRAM IP同样处理为DCP, 简化流程。

实践方法论2 – 灵活应对布局布线 (PnR) 难题

- FPGA后端Place&Routing挑战
 - 在16核高资源占用率 (16片FPGA方案下平均每片75%) 的情况下, 有2片FPGA难以生成bitstream。

- 解决方案 - 采用双管齐下策略:

- ✓ **方案A:** 让工具反复尝试不同的PnR策略迭代优化, 最终仍可以生成Bitfile, 但耗时较长;



FPGA
Utilization
75%

- 适合FPGA资源有限的情况, 能够接受较长P&R迭代时间。
- 但一旦架构稳定后, P&R迭代时间会减少

- ✓ **方案B:** 在资源允许的情况下, 增加FPGA资源到20片(5-box), 降低单片资源压力;



FPGA
Utilization
55%

- 适合FPGA资源充足的情况, 以FPGA资源换取P&R效率, 适合早期RTL快速迭代的场景

实战策略

通过灵活应用编译方案, 在项目进度和FPGA可用资源直接取舍平衡, 最终实现了稳定的bitstream生成。

实践方法论3 - 循序渐进的系统启动 (Bring-up)

策略选择：分而治之，循序渐进

- 从单核开始逐步增加核心数量，降低系统调试复杂度。

分解难题：

- 将复杂的16核系统验证分解为多个里程碑，每个阶段聚焦解决特定难题：
 - 单核：基础功能与接口验证
 - 双核：缓存一致性、核间通信基础验证
 - 4/8核：中等规模系统及调度验证
 - 16核：完整系统与极限性能验证

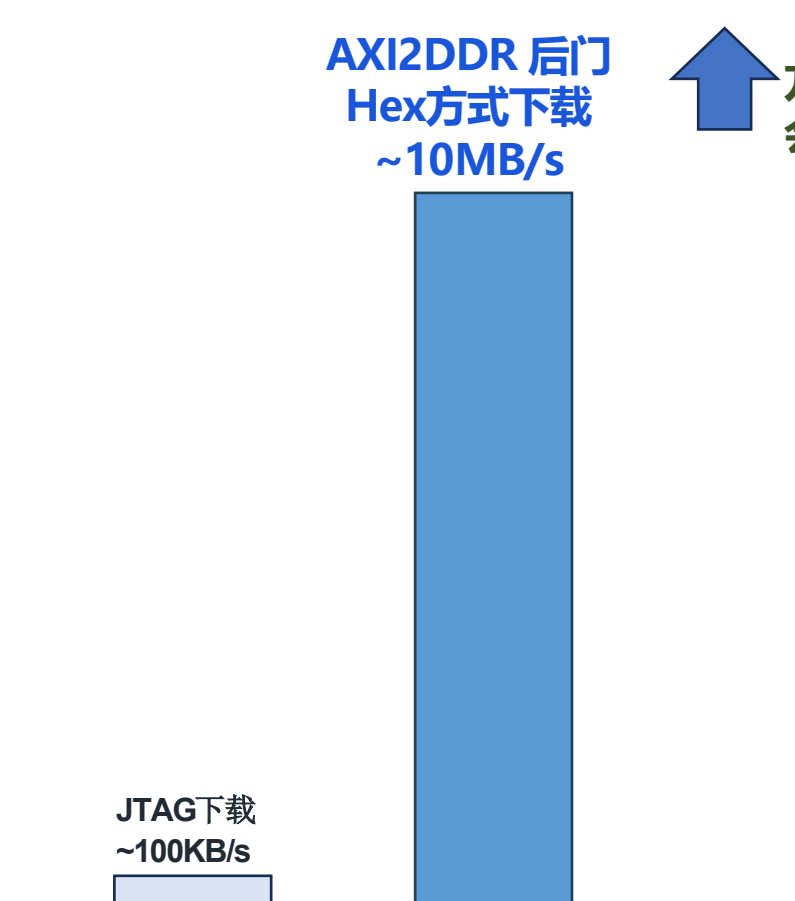


遇到问题

- Boot后卡死：系统正常启动Boot后，在读取Kernel时卡死。
- JTAG下载慢：初期使用JTAG下载Kernel，速度极慢，严重影响调试迭代效率

高效解决方案

- 效率优化：快速切换为AXI2DDR后门加载方案，Kernel下载速度得到质的飞跃。
- 功能Bug：合见平台快速暴露问题，通过仿真复现，最终定位为boot文件问题，修改后解决。



AXI2DDR方案使Kernel下载速度
提升100倍，调试效率质的飞跃

项目成果 - 多版本全自动化编译流程以及性能全面达成

1. 自动化流程

- 在项目初期移植阶段做完基础的移植工作后，后续的编译流程由UVHS工具自动化完成，包括时钟转换、自动分割、TDM IP 绑定、等等关键工作。
- 后续设计进版迭代，复用前套的自动化流程，省掉很多额外的工作。

2. 性能目标全面达成

- 单核/双核/4核/8核版本：CPU主时钟均能稳定运行在 15 MHz，为针对不同应用场景的性能评估和软件开发提供了高速验证平台。
- 16核完整版本：CPU主时钟稳定运行在 10.2 MHz，超越原定 10MHz的目标。
- 项目首次：在FPGA原型验证平台上，完整运行一个未经裁剪的、可启动OS的16核高性能昆明湖RISC-V处理器系统。
- 效率提升：相比传统方法，整体验证效率提升约40%，大幅缩短产品上市时间。



3. 多团队并行验证的需求与解决方案

• 实现方案

- 开发了兼容16核DB和多个4核DB的boardfile, 使验证平台能够灵活切换配置。

• 多用户支持

- 验证平台可以被灵活划分为多个独立的4核系统, 供4个不同软件或算法团队同时并行使用, 大大提高了开发效率。

• 资源利用率提升

- 一套硬件, 既能用于16核的系统级攻关, 也能用于多团队的日常开发, 极大提升了硬件资源的利用率。



综合工具评估：UVHS自带FPGA综合器UVSyn的各项指标比较



合见自研综合器UVsyn技术指标对比 (vs. FPGA自带综合器)

资源优化

LUT总量从4840万降低到4264万，节省11.9%关键资源，使设计布局更加灵活

性能提升

时钟频率从8.8 MHz提升至10.2 MHz，增加15.0%，成功突破项目目标

编译效率

编译时间从41小时缩短至35小时，效率提升14.6%

*测试条件说明：同等约束、同版本16core CPU设计，使用20块 VU19P FPGAs编译，对比FPGA自带综合工具与合见UVSyn综合器的综合优化效果

| 指标 | 第三方综合器 | UVSyn | 提升幅度 |
|-------|---------|----------|---------|
| LUT总量 | 4840万 | 4264万 | ↓ 11.9% |
| 时钟频率 | 8.8 MHz | 10.2 MHz | ↑ 15.0% |
| 编译时间 | 41小时 | 35小时 | ↓ 14.6% |

性能目标达成对比

3rd Party tool: 8.8 MHz

UVSyn : 10.2 MHz

10.2 MHz



后续项目工作展望

1. 验证效能深化

- 打通原型验证平台与Emulation的协同验证的相互转化路径, 提升调试效率
- Ram数据的debug读写, 构建内存分析平台
- 部署UVHS平台上其他更高效的debug方法, 帮助CPU问题定位

2. 更大规模的设备级联扩展, 更多核心验证场景

- 基于温榆河NOC的32/64核架构部署及验证
- 256核NoC互联压力测试
- 异构验证场景
 - RISC-V + AI加速核协同验证框架
 - 高带宽存储验证平台搭建

3. 希望EDA厂商能开发更多支持CPU核心之外验证场景的特性

- UPF低功耗验证
- DPA动态功耗验证

未来展望 - EDA协同推动RISC-V验证方法学标准化

1. 方法学沉淀

- 构建基于香山CPU的多核验证方法学：将本次项目的宝贵经验，沉淀为可复制、可扩展的成功范例参考流程，服务更多多核CPU设计的标准流程，造福业界和开源社区。

2. 工具链进化

- 软硬件协同验证再加速：探索Emulator（硬件仿真器）与UVHS（原型验证）的混合验证模式
- 多核验证方法学演进方向：
 - 构建缓存一致性协议验证方法学；
 - 探索跨核事件追踪框架，实现微架构级问题溯源。
- 基于合见即将推出的CHI XTOR方案探索温榆河NOC的相关应用，诸如Transaction硬件仿真加速以及System Performance分析工具等
- 开发基于Emulator的CPU Tracer类的调试工具
- 增加功耗-性能联合分析能力：为SoC设计提供更全面的评估维度

3. 生态建设

- 开放接口标准：推动验证工具接口标准化，促进产业链各方高效协作，生态共建加速RISC-V产业的健康发展。
- 系统级方案建设：为CPU设计和集成厂商，提供从工具、方法学到技术支持的系统级配套解决方案。
- 降低行业门槛：强大的EDA工具和服务，能显著降低高性能CPU的验证门槛，吸引更多伙伴加入RISC-V生态。

谢谢！



上海合见工业软件集团有限公司
Shanghai UniVista Industrial Software Group Co., Ltd.