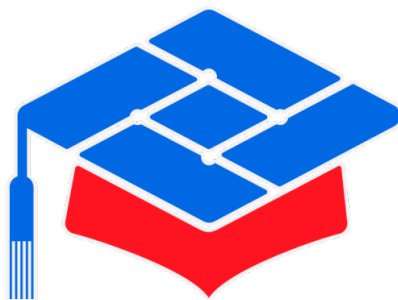




第十届全国大学生集成电路创新创业大赛

# 基于龙芯 SoC 的后量子密码与矩阵运算 加速系统

系统设计与验证报告

|      |                           |
|------|---------------------------|
| 参赛杯赛 | 龙芯中科杯                     |
| 队伍编号 | CICC1008439               |
| 团队名称 | 碳基生物                      |
| 目标器件 | Xilinx Artix-7 XC7A200T-1 |
| 证据版本 | 1b51f28（历史报告，待复测）         |
| 报告日期 | 2026-07-15                |

## 摘要

本项目在赛事提供的 LoongArch32r SoC 框架上集成 ML-KEM、1024 点 FFT 和  $4 \times 4$  矩阵乘法三类专用加速器，并以四通道 DMA 和矩阵乘法批处理序列器降低 CPU 搬运与调度开销。系统采用 CPU、系统总线和 ML-KEM 三时钟域架构；跨域控制、状态和数据通过显式 CDC 结构传递。矩阵乘法器使用四路 LUT 乘法器，每拍并行产生一个输出元素点积所需的四项乘积，并用双 Bank、流水 CRC32 与本地 DMA endpoint 支撑 5000 组连续任务。报告重点说明设计边界、关键决策及可复现证据，不把参考 SoC 基础设施误列为本项目贡献。

现有证据显示，历史实现的总体 WNS 为  $+0.070\text{ ns}$ ，资源占用为 55,066 LUT、31,000 FF、29.5 BRAM Tile 和 0 DSP；板级 5000 组矩阵乘法耗时为  $5.213592\text{ ms}$ ，CRC 校验通过。由于当前分支经历 PLL 与实现报告合并，所有 PPA 和板级性能必须按附带清单重新生成后再作为最终结论。

**关键词：** LoongArch32r；ML-KEM；DMA；矩阵乘法；FFT；跨时钟域；FPGA

## 主要贡献

- 在固定 AXI 端口预算内完成三类加速器集成，并用二级地址译码共享 FFT/ML-KEM 窗口；
- 构建 mem2mem、mem2ep、ep2mem 三类 DMA 路径及矩阵乘法硬件批处理序列器；
- 将 ML-KEM 计算域提升至独立时钟，并为命令、数据、状态和复位建立明确 CDC 边界；
- 建立从 RTL 静态检查、SoC 仿真、布局布线到板级评测的版本化证据链。

## 目录

|                     |          |
|---------------------|----------|
| 摘要                  | I        |
| 主要贡献                | I        |
| <b>1 需求、约束与设计边界</b> | <b>1</b> |
| 1.1 需求分解            | 1        |
| 1.2 软硬件边界           | 1        |
| <b>2 系统架构</b>       | <b>2</b> |
| 2.1 互连与地址空间         | 2        |
| 2.2 时钟、复位与 CDC      | 2        |
| 2.3 中断与完成通知         | 3        |

|          |                                   |          |
|----------|-----------------------------------|----------|
| <b>3</b> | <b>关键设计决策</b>                     | <b>3</b> |
| 3.1      | DMA 本地 endpoint 与批处理序列器 . . . . . | 3        |
| 3.2      | 共享 FFT/ML-KEM AXI 窗口 . . . . .    | 3        |
| 3.3      | ML-KEM 独立计算时钟 . . . . .           | 4        |
| <b>4</b> | <b>加速器实现</b>                      | <b>4</b> |
| 4.1      | 矩阵乘法与流水 CRC . . . . .             | 4        |
| 4.2      | 1024 点 FFT . . . . .              | 4        |
| 4.3      | ML-KEM . . . . .                  | 5        |
| <b>5</b> | <b>验证策略与判定标准</b>                  | <b>5</b> |
| 5.1      | 验证层次 . . . . .                    | 5        |
| 5.2      | ML-KEM 样本数量解释 . . . . .           | 5        |
| <b>6</b> | <b>已有结果与复测状态</b>                  | <b>7</b> |
| 6.1      | 矩阵乘法板级结果 . . . . .                | 7        |
| 6.2      | ML-KEM 性能与正确性 . . . . .           | 7        |
| 6.3      | FFT 结果 . . . . .                  | 7        |
| 6.4      | 时序、资源与功耗 . . . . .                | 8        |
| <b>7</b> | <b>局限、风险与下一步</b>                  | <b>8</b> |
| <b>8</b> | <b>结论</b>                         | <b>9</b> |
| <b>A</b> | <b>关键寄存器摘要</b>                    | <b>9</b> |
| <b>B</b> | <b>复现入口</b>                       | <b>9</b> |

# 1 需求、约束与设计边界

## 1.1 需求分解

表 1把竞赛目标转化为可验证的工程要求。每项要求都绑定明确证据，避免以“能够综合”替代功能、时序或板级行为验证。

表 1: 系统需求与验收证据

| 编号 | 要求                                            | 验收证据                                | 当前状态      |
|----|-----------------------------------------------|-------------------------------------|-----------|
| R1 | 4×4 无符号矩阵乘法，5000 组连续输入，结果 CRC32 正确            | 官方板级评测日志：CRC 匹配且收到 MATMUL_DONE      | 已有基线，待复测  |
| R2 | 复位释放前不得产生 UART、DMA 或加速器事务                     | ILA/测试平台断言与板级观察                     | 待当前版本复测   |
| R3 | 矩阵乘法路径不使用 DSP48E1                             | 布局布线后 utilization 与自动 DSP 检查        | 已有基线，待复测  |
| R4 | ML-KEM-512/768/1024 的 KeyGen、Encaps、Decaps 正确 | 3 个固定 KAT 向量，SW/CPU/DMA 三路径各 100 次  | 已有基线，待复测  |
| R5 | FFT CPU 与 DMA 路径结果和软件参考一致                     | 1024 点逐点最大误差及端到端耗时                  | 需补全最大误差日志 |
| R6 | 所有同步时钟域 setup/hold 收敛，CDC 结构可审计               | timing、clock interaction、CDC、DRC 报告 | 已有基线，待复测  |

## 1.2 软硬件边界

赛事框架提供 CPU、基础存储控制器、UART、DVI、confreg 和 AXI 基础互连。本项目的主要工作是加速器 RTL、DMA 数据路径、批处理控制、共享地址窗口、CDC 集成以及配套软件。CPU 只负责配置、启动、完成检测和结果判定；NTT、Keccak、多项式乘法、FFT 蝶形和矩阵乘积累加均在硬件内部执行。

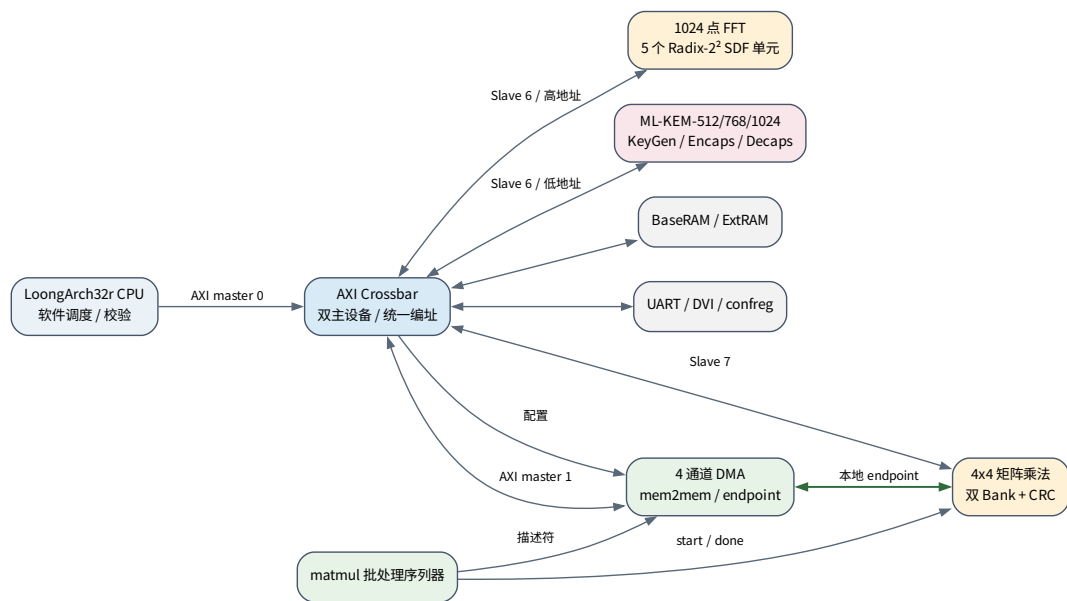


图 1: 系统组成、贡献边界与主要数据通路

## 2 系统架构

### 2.1 互连与地址空间

如图 1, CPU 和 DMA 分别作为 AXI Master 0/1 接入 Crossbar。DMA 通过 AXI 访问存储器, 同时通过本地 endpoint 直接访问矩阵乘法双 Bank, 避免把内部缓冲搬运再次送入 Crossbar。关键地址窗口见表 2; 完整寄存器定义以 RTL 和 BSP 头文件为准。

### 2.2 时钟、复位与 CDC

系统由 50 MHz 输入时钟经 PLL 产生三个同步逻辑域: CPU 域、系统总线域和 ML-KEM 计算域, 拓扑见图 2。CPU-SYS 之间跨越完整 AXI 通道; SYS-ML-KEM 之间跨越命令、窗口数据、完成状态和中断。设计原则是: 多位数据只随握手或稳定窗口跨域, 单比特状态经过同步器, 脉冲采用保持/应答结构, 复位采用异步置位、同步释放。

表 2: 关键物理地址窗口

| 物理地址                     | 模块             | 说明                                      |
|--------------------------|----------------|-----------------------------------------|
| 0x1C00_0000--0x1C7F_FFFF | BaseRAM/ExtRAM | 程序、源数据和结果存储                             |
| 0x1F00_0000              | UART 16550     | 串口输出                                    |
| 0x1F20_0000              | confreg        | GPIO、定时器与中断                             |
| 0x1F30_0000              | DMA / 序列器      | DMA 通道位于低 0x200 B; matmul 序列器从 0x200 开始 |
| 0x1F40_0000              | ML-KEM / FFT   | 由地址位二次译码的共享 Slave 6 窗口                  |
| 0x1F50_0000              | matmul         | 控制、CRC 和双 Bank 数据窗口                     |

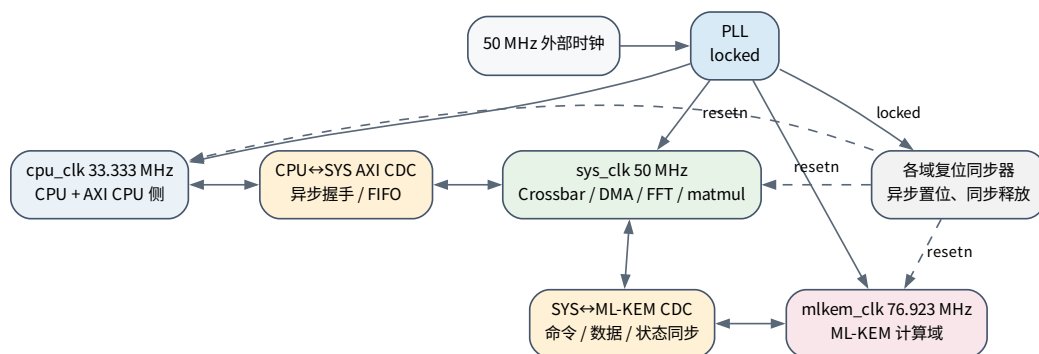


图 2: 时钟、复位与跨时钟域边界

复位约束比普通功能复位更严格：PLL 锁定且各域同步器完成释放前，AXI valid、DMA 请求、加速器 start 和 UART 写使能必须保持为零。这一约束需要通过专门的复位前监控测试确认，不能由“复位后程序正常运行”间接推出。

### 2.3 中断与完成通知

DMA 完成、matmul 完成和 ML-KEM 完成信号经同步后接入 CPU 中断系统。当前性能程序在短热路径中采用状态轮询，减少中断现场开销；功能回归仍需验证 done 的置位、W1C 清除、连续任务不丢失以及错误状态可恢复。

## 3 关键设计决策

### 3.1 DMA 本地 endpoint 与批处理序列器

通用 DMA 提供 4 个独立通道，每通道包含源地址、目的地址、长度、模式和 endpoint 配置。对普通内存或 FFT/ML-KEM 窗口使用 AXI burst；对 matmul 双 Bank 使用本地 valid/ready endpoint。后者减少地址译码和总线仲裁，但必须单独验证 backpressure、Bank 选择和非满 burst 长度。

矩阵乘法序列器把“读 128 B 输入、启动计算、等待完成、写回 192 B 结果、递增组号”固化为硬件循环。CPU 仅配置源/目的基址、组数和 stride。图 3 描述单组事务；双 Bank 允许搬运与计算在相邻组之间重叠。当前竞赛快速路径可关闭结果写回，改由硬件流水 CRC 对整批结果判定，从而消除不参与最终评分的存储流量。

### 3.2 共享 FFT/ML-KEM AXI 窗口

Crossbar Slave 6 内部按地址高位将事务分发到 ML-KEM 或 FFT。该方案节省一个顶层从端口，但会把二级译码放入共同路径。读写响应必须与请求目标绑定，不能在 AW/W 独立握手时重新按瞬时地址选择从机。验证重点包括 AW/W 不同拍到达、读写并发、从机反压及非法空洞地址响应。

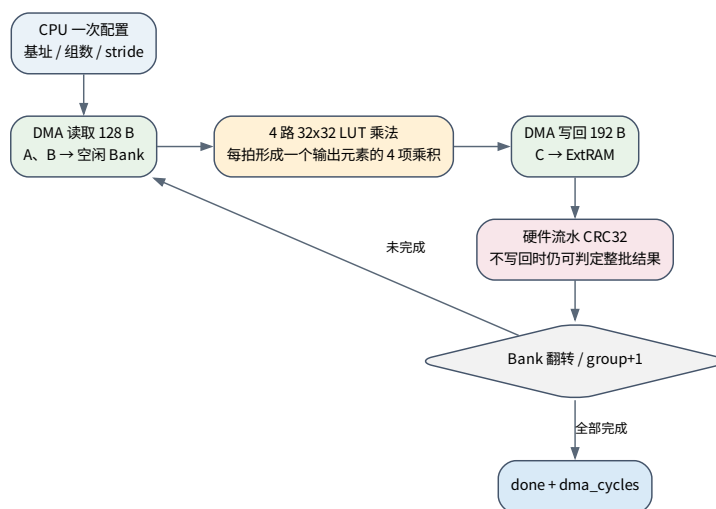


图 3: 矩阵乘法批处理事务与 CRC 路径

### 3.3 ML-KEM 独立计算时钟

ML-KEM 的关键通路 with 系统 AXI 通路性质不同。将计算核心放入独立的 76.923 MHz 时钟域，可在不提高 CPU 和 SRAM 压力的情况下缩短算法计算段。代价是增加 CDC 延迟与验证复杂度，因此性能比较必须报告端到端 CPU 周期，而不是仅报告核心 busy 周期；CDC 报告也必须作为实现准入条件。

## 4 加速器实现

### 4.1 矩阵乘法与流水 CRC

加速器支持最大  $4 \times 4$  的 32-bit 无符号矩阵，66-bit 结果元素按三个 32-bit 字输出。四个  $32 \times 32$ -bit LUT 乘法器在同一拍分别计算一个输出元素点积中的四项乘积，之后进入累加和写回流水；它们并非分别负责四行。A/B/C 缓冲采用双 Bank，控制器通过 Bank 所有权避免 DMA 与计算阵列同时写同一缓冲。

CRC32 与结果产生流水同步更新。对 5000 组评测，软件只输出起止标记和最终 CRC，减少 UART 对计时窗口的扰动。硬件 `dma_cycles` 记录批处理内部周期，板级主机计时则覆盖复位释放后协议观察与外部通信，两者用途不同，不应混为一个指标。

### 4.2 1024 点 FFT

FFT 数据通路由 5 个 Radix-2<sup>2</sup> SDF 单元串联，每个单元包含两个蝶形阶段，而不是“五级普通 radix-2 FFT”。延迟反馈结构以 BRAM/移位存储保存历史样本，旋转因子由 ROM 提供。输入输出窗口按 32-bit AXI word 编址，低 16 bit 承载有符号定点样本。CPU 路径用于接口诊断，DMA 路径反映系统实际吞吐。

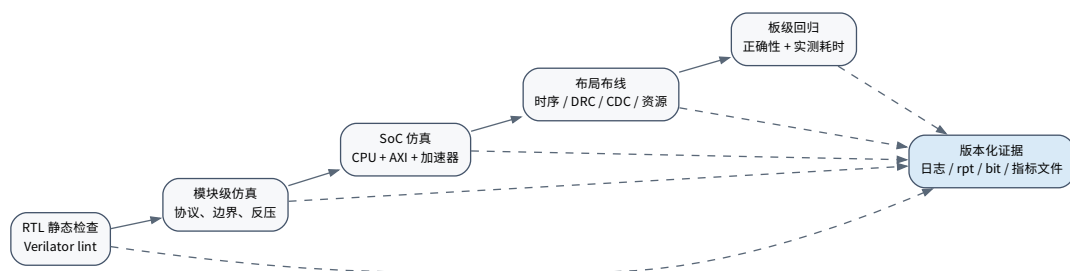


图 4: 从 RTL 到板级结果的分层证据链

### 4.3 ML-KEM

ML-KEM 核支持 FIPS 203 的三个参数集及 KeyGen、Encaps、Decaps。主要计算模块包括 Keccak/SHA3、采样、NTT/INTT、多项式乘法、压缩/解压和片上存储。AXI 包装器提供控制寄存器与 EK、DK、CT 数据窗口；DMA 路径批量搬运密钥和密文，CPU 路径保留为可观测的接口基线。

## 5 验证策略与判定标准

### 5.1 验证层次

验证按图 4 推进。静态检查发现语法、位宽和不可达集成；模块级仿真覆盖握手和边界；SoC 仿真覆盖 CPU、缓存、AXI 和软件驱动；实现报告检查时序、CDC、DRC、DSP 与资源；板级测试确认真实 PLL、SRAM、串口和复位行为。任何一层通过都不能替代后一层。

### 5.2 ML-KEM 样本数量解释

当前日志的 `STATIC_KAT_COUNT_0` 表示每个参数集使用一个固定向量，因此是 3 个不同 KAT 向量。每个向量在 SW、CPU、DMA 三条路径各运行 100 次，总计  $3 \times 3 \times 100 = 900$  次执行；不能表述为“360 个不同 KAT”。现有日志没有注入篡改密文，故本报告不宣称已覆盖隐式拒绝。若要加入该结论，应新增负向用例并保存独立证据。

表 3: 必须重跑的验证矩阵

| 测试             | 判定标准                                                       | 证据文件                               |
|----------------|------------------------------------------------------------|------------------------------------|
| Verilator lint | 无 Error; 集成顶层可达性检查通过                                       | evidence/lint.log                  |
| 指定软件构建         | asm、fft、fft_dma、mlkem_test、mlkem_hello 均成功链接               | evidence/build/*.log               |
| FFT CPU 路径     | 1024 点完成; 记录端到端 ns 与逐点最大误差                                 | evidence/soc/fft.log               |
| FFT DMA 路径     | 四次搬运无死锁; 最大误差和性能均输出                                        | evidence/soc/fft_dma.log           |
| ML-KEM 三路径     | k=2/3/4、SW/CPU/DMA 各 100 次, FAILS=0                        | evidence/soc/mlkem_test.log        |
| ML-KEM 示例      | 共享密钥一致且明文恢复, 输出 RESULT=PASS                                | evidence/soc/mlkem_hello.log       |
| 完整实现           | setup/hold WNS $\geq$ 0, unconstrained=0, DSP=0, DRC 无关键错误 | evidence/vivado/*.rpt              |
| CDC 审计         | 预期跨域结构已约束; 无未知组合跨域或未同步脉冲                                   | evidence/vivado/cdc.rpt            |
| matmul 官方板测    | 5000 组 CRC 匹配、收到 DONE、记录耗时和分数                              | evidence/board/matmul_official.log |
| 复位前零操作         | 复位释放前 UART/AXI/DMA/start 均无事务                              | ILA 截图或断言日志                        |

## 6 已有结果与复测状态

**证据状态：**以下数值由脚本从当前仓库可用日志/报告提取，证据版本为 1b51f28（历史报告，待复测）。PLL 与实现产物合并后必须执行完整复测；在新 evidence/ 生成前，这些数字只作为历史基线，不作为最终提交结论。

### 6.1 矩阵乘法板级结果

现有官方评测基线见表 4。外部耗时包含评测端从观察到 MATMUL\_START 至收到 MATMUL\_DONE 的时间；内部 dma\_cycles 应在复测日志中同时记录，用于区分硬件批处理时间与 UART/软件尾部时间。

表 4: 5000 组矩阵乘法板级基线

| 指标    | 结果          |
|-------|-------------|
| 组数    | 5,000       |
| 外部计时  | 5.213592 ms |
| CRC32 | PASS        |
| 评分    | 99.56%      |
| 满分目标  | 4.9 ms      |

### 6.2 ML-KEM 性能与正确性

表 5给出一次 KeyGen+Encaps+Decaps 往返的平均 CPU 周期。每个单元格对应 100 次执行且现有日志 FAILS=0。DMA 相比寄存器路径的收益来自批量窗口搬运；它不表示密码核心本身获得相同比例的加速。

表 5: ML-KEM 三路径平均周期（历史基线）

| 参数集         | 软件参考      | CPU 寄存器 | DMA 窗口  |
|-------------|-----------|---------|---------|
| ML-KEM-512  | 2,594,034 | 200,471 | 52,995  |
| ML-KEM-768  | 4,065,123 | 296,193 | 80,648  |
| ML-KEM-1024 | 6,041,204 | 404,476 | 113,273 |

### 6.3 FFT 结果

现有 CPU 与 DMA 日志使用相同 1024 点多音输入。表 6中的最大误差尚未由旧日志完整输出；当前测试程序已增加全 1024 点统计，复测后指标脚本会自动更新。

表 6: FFT 端到端性能（历史基线）

| 路径      | 硬件端到端时间        | 相对软件加速比 | 正确性                           |
|---------|----------------|---------|-------------------------------|
| CPU 寄存器 | 1,695,360 ns   | 161.94× | 已显示能量 bin 一致；需重跑全量误差          |
| DMA     | 407,310 ns     | 674.64× | Re=7, Im=6（按现有日志逐点估计，待程序直接输出） |
| 软件参考    | 274,786,920 ns | 1×      | 浮点参考                          |

#### 6.4 时序、资源与功耗

当前可用布局布线报告的汇总见表 7。三个域的 setup WNS 均为正；最终复测还必须检查 hold、未约束路径、clock interaction 和 CDC，而不能只引用一行 WNS。

表 7: 实现结果基线

| 类别 | 指标              | 当前可用值                   |
|----|-----------------|-------------------------|
| 时序 | cpu_clk / WNS   | 33.333 MHz / +0.737 ns  |
| 时序 | sys_clk / WNS   | 50.000 MHz / +0.125 ns  |
| 时序 | mlkem_clk / WNS | 76.923 MHz / +0.070 ns  |
| 资源 | Slice LUT       | 55,066 (41.16%)         |
| 资源 | Slice Register  | 31,000 (11.58%)         |
| 资源 | BRAM Tile       | 29.5 (8.08%)            |
| 资源 | DSP48E1         | 0                       |
| 功耗 | 动态 / 静态 / 总计    | 0.637 / 0.142 / 0.779 W |

Vivado 功耗置信度为 Low（未回标活动率）。由于没有回标 SAIF/VCD 活动率，总功耗只能作为向量无关估算，不宜用于细粒度模块功耗排名。若评审要求功耗结论，应以代表性 ML-KEM 和 matmul 负载生成活动文件后重新运行 `report_power`。

## 7 局限、风险与下一步

- 当前 PPA 与板级日志不是 PLL/CDC 合并后的同一提交证据，必须按清单重跑并记录 commit；
- 通用 SoC 仿真 testbench 以程序返回地址结束，尚未对复位前零事务写成自动断言；
- ML-KEM 固定向量覆盖正确性和重复稳定性，但不等价于大量独立随机 KAT，也未覆盖隐式拒绝；
- FFT 旧日志仅打印有能量的 bin，当前程序虽已统计全量最大误差，仍需生成新日志；
- 功耗报告没有活动率回标，置信度低；模块级功耗百分比不作为本报告结论；

- 层次化资源表容易因层次重叠而重复计数，因此正文只报告顶层非重叠总量。若需模块占比，必须用同一层次深度并增加“其他/未归属”项。

## 8 结论

本项目的核心价值不只是增加三个 AXI 外设，而是形成了可持续工作的异构数据路径：CPU 负责控制，DMA 负责搬运，序列器负责批处理，加速器负责计算，CDC 负责隔离不同性能目标的时钟域。共享窗口和本地 endpoint 在端口与延迟约束之间取得平衡；双 Bank 和流水 CRC 则面向 5000 组矩阵任务消除了不必要的软件与写回开销。

已有基线证明方案具备功能与性能潜力，但最终报告应只引用同一 commit 下重新生成的功能、时序、资源、CDC、功耗和板级证据。附带脚本已将这一过程标准化：执行完整回归、保存日志并运行 `make metrics all` 后，正文表格和 PDF 会随证据同步更新。

## A 关键寄存器摘要

表 8: DMA 与 matmul 序列器寄存器

| 偏移                        | 名称         | 含义                         |
|---------------------------|------------|----------------------------|
| <code>n*0x40+0x00</code>  | SRC_ADDR   | DMA 通道 $n$ 源物理地址           |
| <code>n*0x40+0x04</code>  | DST_ADDR   | DMA 通道 $n$ 目的物理地址          |
| <code>n*0x40+0x08</code>  | LENGTH     | 字节数                        |
| <code>n*0x40+0x10</code>  | CTRL       | start、AxSIZE、AxLEN         |
| <code>n*0x40+0x14</code>  | STATUS     | busy、done(W1C)             |
| <code>n*0x40+0x18</code>  | MODE       | mem2mem / mem2ep / ep2mem  |
| <code>0x200</code>        | SEQ_CTRL   | 批处理 start                  |
| <code>0x204</code>        | SEQ_STATUS | busy、done(W1C)             |
| <code>0x208--0x21C</code> | SEQ 参数     | 源/目的基址、组数、matmul 基址、stride |

## B 复现入口

详细顺序、预期输出和文件放置方式见同目录 `REPORT_RERUN_CHECKLIST.md`。常用入口如下：

```
cd doc/reports/system
./tools/run_report_tests.sh lint
./tools/run_report_tests.sh soc
./tools/run_report_tests.sh implement
./tools/run_report_tests.sh refresh
```

## 参考文献

- [1] National Institute of Standards and Technology, *FIPS 203: Module-Lattice-Based Key-Encapsulation Mechanism Standard*, 2024.
- [2] 龙芯中科技术有限公司, 龙芯 SoC 实验平台开发手册, 2024。
- [3] AMD Xilinx, *UG906: Vivado Design Analysis and Closure Techniques*.
- [4] S. He and M. Torkelson, “Designing pipeline FFT processor for OFDM,” URSI International Symposium on Signals, Systems, and Electronics, 1998.