



南京邮电大学
Nanjing University of Posts and Telecommunications

基于GPU光线追踪引擎的 树遍历加速器设计与实现

—— 毕业论文答辩 ——

汇报人：B21030116 罗方喆

汇报时间：2025年5月29日



目录

CONTENTS

01

研究背景与目标

02

研究过程与设计

03

实验成果及分析

04

总结与展望



研究背景与目标

研究过程与设计

实验成果及分析

总结与展望

01

PART ONE

研究背景与目标

研究背景与目标

研究过程与设计

实验成果及分析

总结与展望

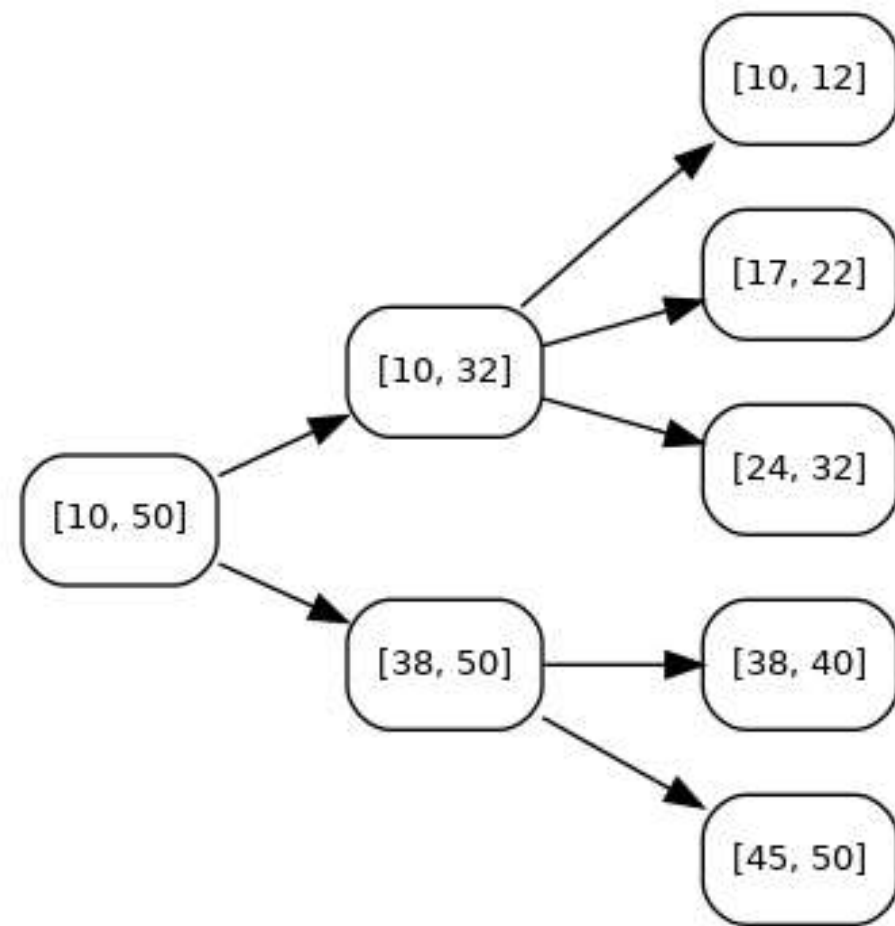
树结构的应用场景？



树是计算机科学领域的各类研究中
常用到的一种数据结构

B树及其变体（如B+树和R树）

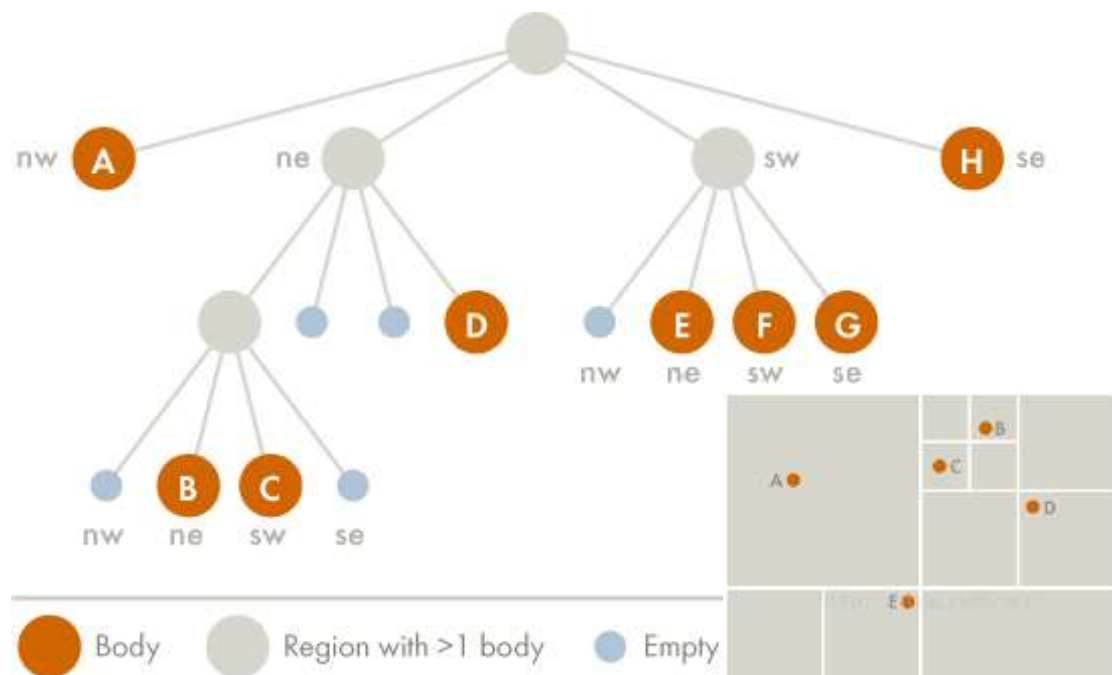
在**数据库系统**中被广泛用于索引数据，
支持快速的插入、删除和查询操作。



树结构在多领域的广泛应用

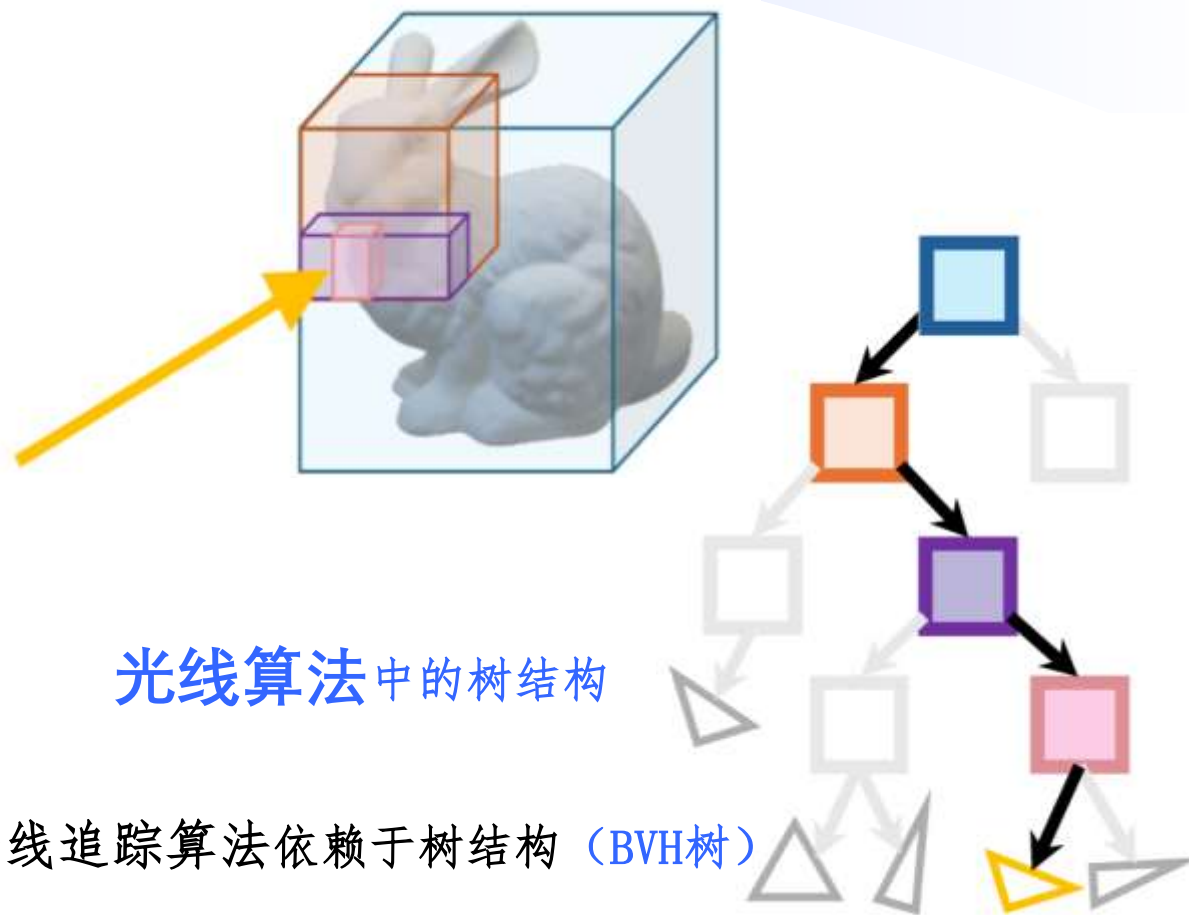


南京邮电大学
Nanjing University of Posts and Telecommunications



物理模拟中的树结构

N体模拟中，Barnes-Hut算法利用树结构，将计算复杂度从 $O(n^2)$ 降低到 $O(n \log n)$ ，显著提升了模拟效率。



光线算法中的树结构

光线追踪算法依赖于树结构（BVH树）来加速光线与场景几何体的相交测试，生成高质量的图像。

研究背景与目标

研究过程与设计

实验成果及分析

总结与展望

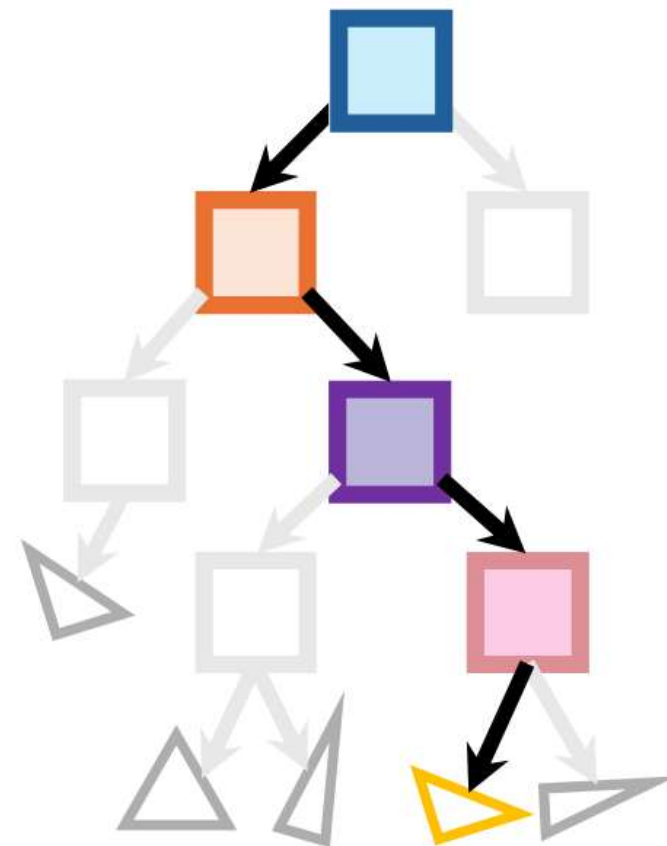
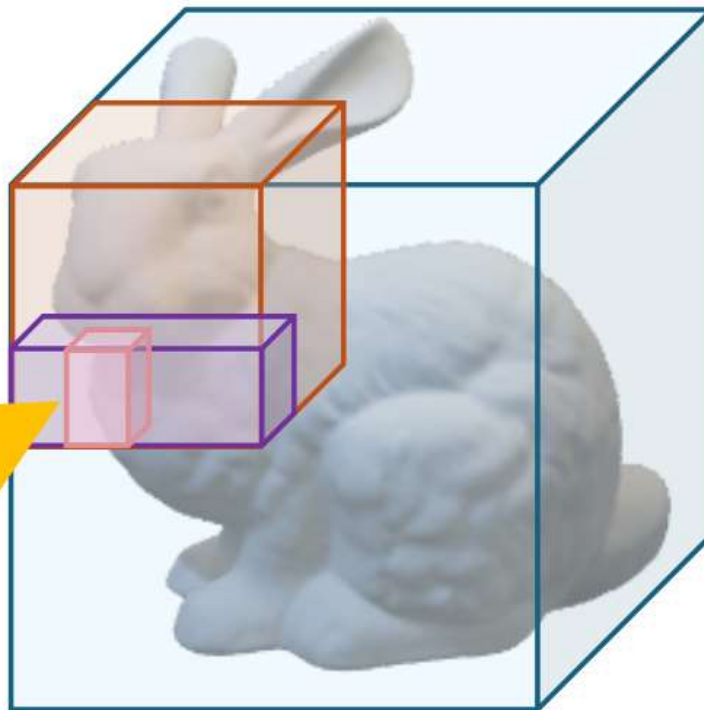
光追当中的树遍历加速

计算机图形学中的树结构

在计算机图形学中，光线追踪算法依赖于树结构（典型如k-d树，**BVH树**）

BVH树 (Bounding Volume Hierarchy)，即层次包围体

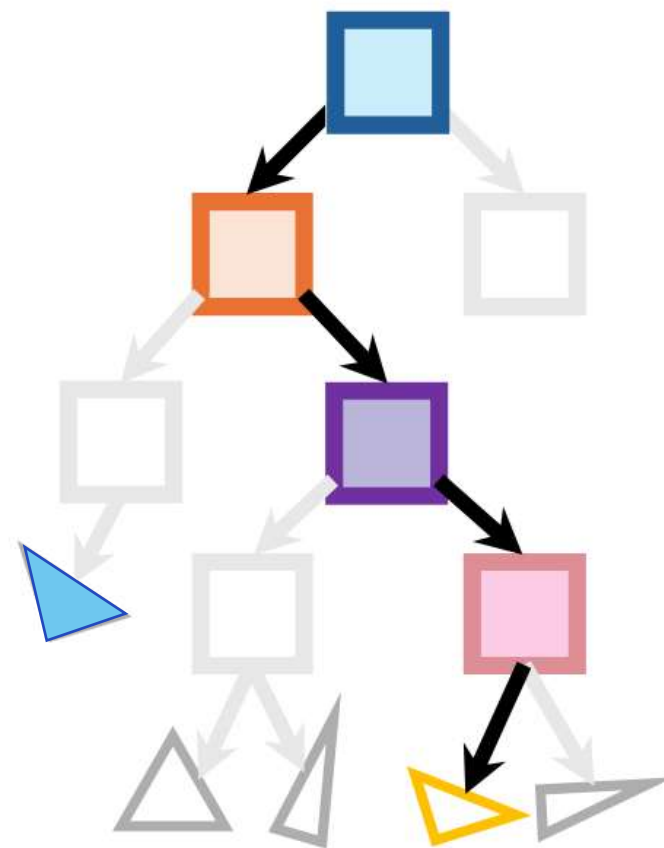
BVH通过层次化的包围盒结构，能够减少光线与几何体的交集测试次数，显著提升了渲染效率。



光追当中的树遍历加速

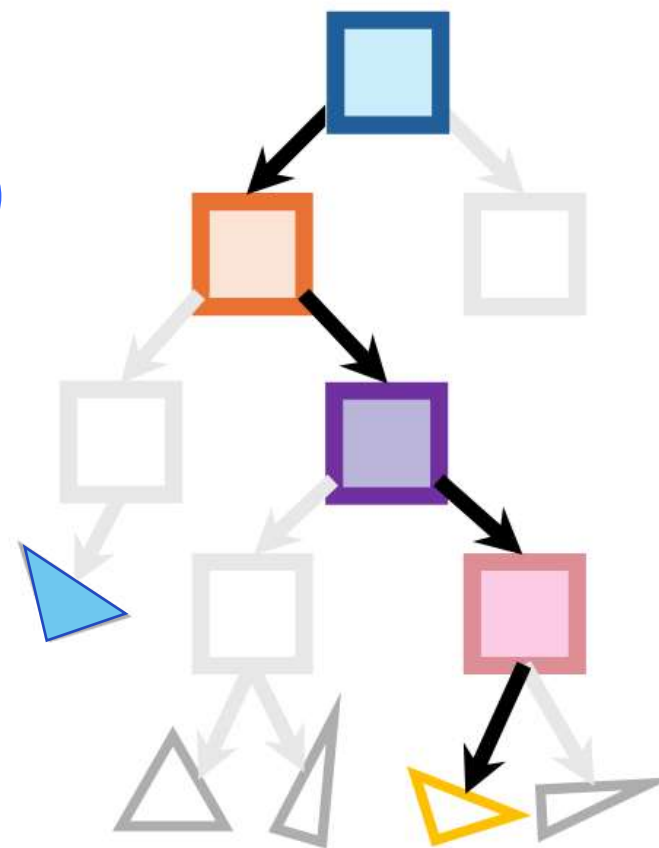
- 为了提升光线追踪当中BVH树遍历的性能，主流GPU厂商在硬件中集成了**光线追踪加速器** (Ray Tracing Accelerator, RTA)
- RTA通过**专有化的固定功能的光追单元设计**，**显著提升了BVH树结构的遍历效率**

然而，由于RTA设计的固定化及专有化除光追外，**其他的树结构**并不能简单的利用RTA硬件来加速自身遍历



本研究关注的核心问题是：

是否可以将“**现有RTA架构中，仅用于BVH树的专用加速**”，
应用到其他树遍历算法的加速中？



然而，由于RTA设计的固定化及专有化
除光追外，**其他的树结构**并不能简单的利用RTA硬件来
加速自身遍历

02

PART TWO

研究过程与设计

研究背景与目标

研究过程与设计

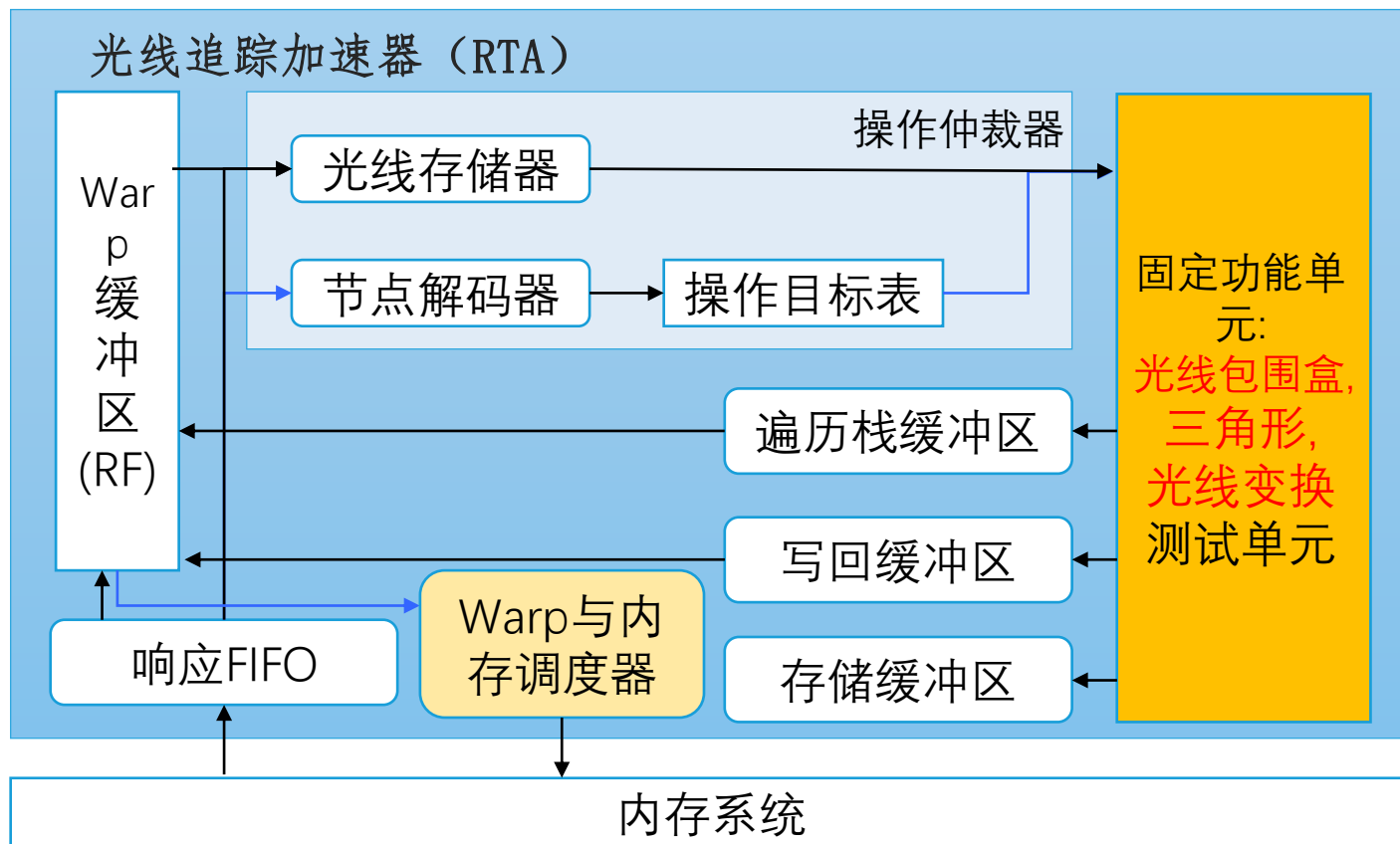
实验成果及分析

总结与展望

本研究尝试构建一种更具普适性的树遍历加速方案，在**保留RTA硬件优势**的基础上，探索其不同空间结构中的**适用性**与改造路径。

- 提出了基于**RTA硬件复用**进行修改的**TTA (Tree Traversal Accelerator)** 树遍历加速器
- 其在保留原有RTA硬件逻辑设计的基础上，还能**为B树搜索、N体模拟等树遍历操作**提供支持。

TTA: RTA功能模块的复用扩展



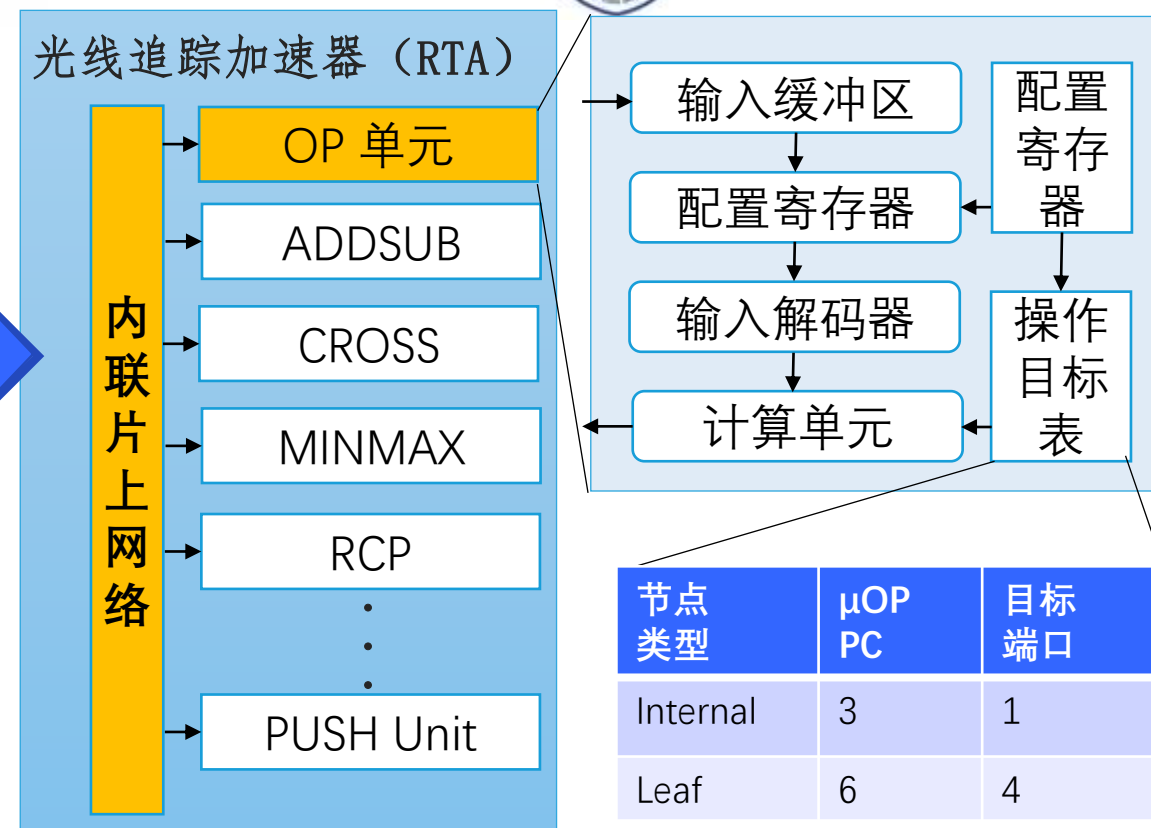
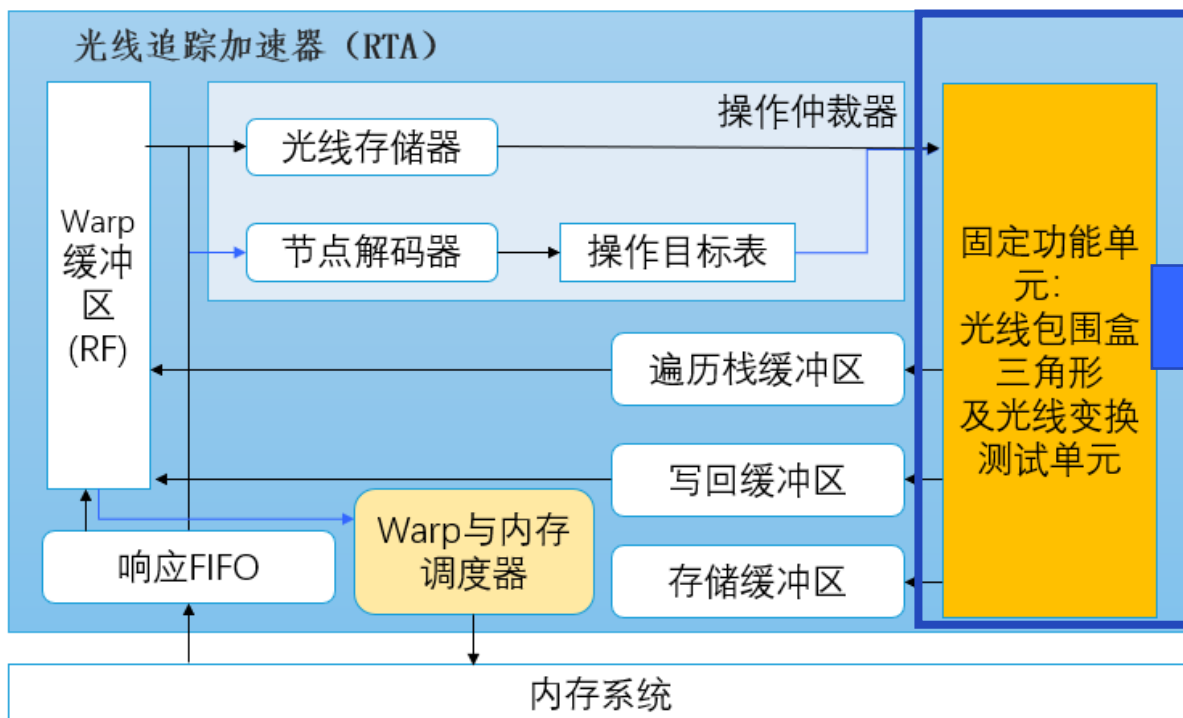
如图所示, 图为Vulkan-sim当中模拟的GPU光线追踪加速单元 (RT unit)

- 研究针对其**三类固定功能单元**的api设计进行了扩展, 使其在**硬件层面**也能够**额外支持两类**树遍历算法当中常用的关键操作:

- 查询键值比较 (Query-Key Value Comparison) : 索引搜索
- 点对点距离计算 (Point-to-Point Distance Calculation) : 半径搜索

- 然而**, 由于其仍采用的是**原本固有的运算单元**, 对于更复杂逻辑的树结构**适配较差!**

TTA+: 可编程操作单元设计



更灵活的设计方案——TTA+架构

- 通过把原本的固定功能管线分解成一个个独立的计算单元，通过可编程OP单元设计，以此完成更细化的树遍历逻辑和数据结构适配。
- 尽管该设计在效率方面稍低于TTA，然而在可扩展性以及灵活性方面有较大优势，可适配更多样化的应用需求（比如Ray-Sphere（光线与球体）的相交测试）

03

PART THREE

实验成果及分析

研究背景与目标

研究过程与设计

实验成果及分析

总结与展望



实验测试：比较多种树遍历结构的运行效果

实验平台： **Vulkan-Sim**——光线追踪硬件仿真平台

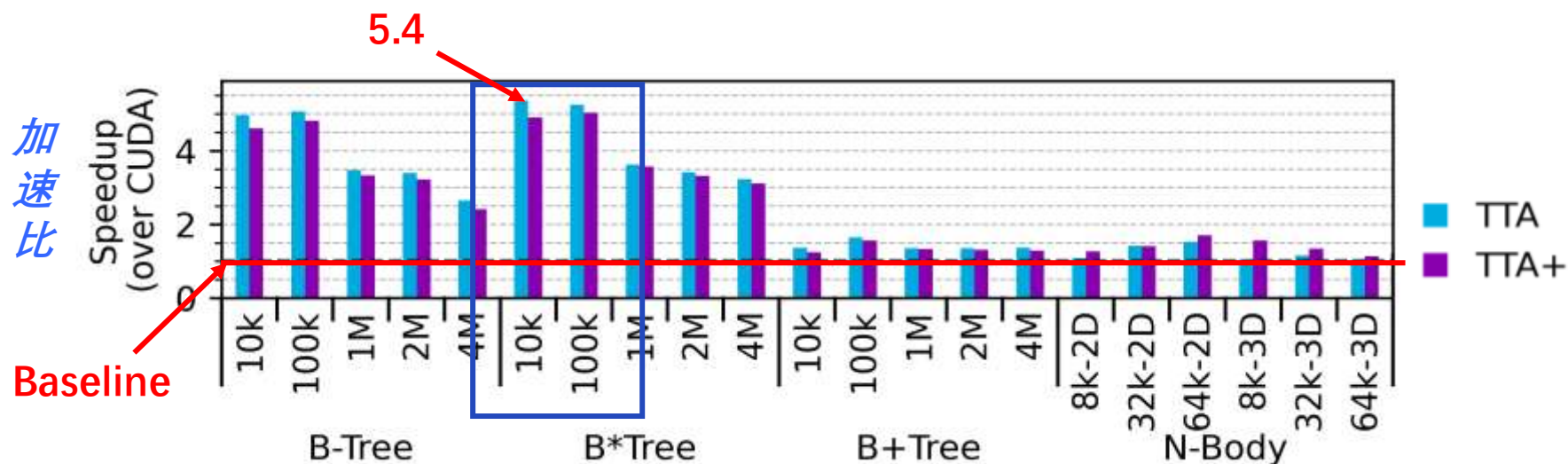
Benchmark:

- Btree及其变体：索引查询，搜索
- N-body：多体模拟仿真
- RTNN：半径搜索算法
- WKND-PT：复杂场景（球体交集）
测试

Comparison:

- 基准测试: Cuda单元 (Vulkan-sim RTA vkCmdTraceRaysKHR() api)
- *TTA*: 可复用硬件单元设计，低损耗扩展应用场景
- *TTA+*: 可编程性运算单元设计，拆解固定单元，进一步支持复杂场景逻辑

TTA: 最优仿真加速效果展示



图中展示的是B树的变种及N体模拟在TTA架构上运行的**加速效果**，纵轴是**加速比**

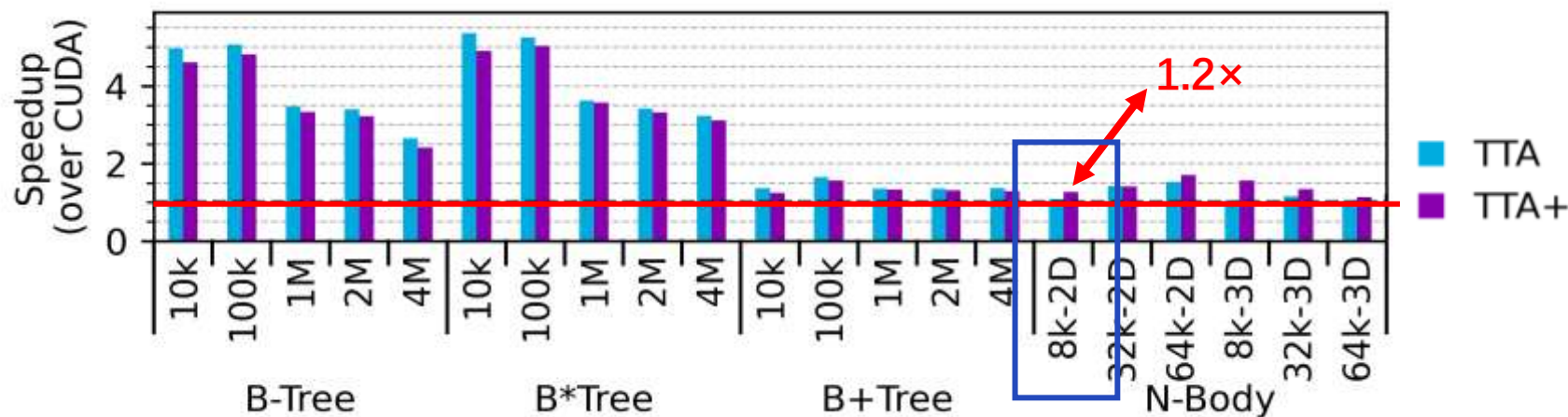
从实验结果来看，**TTA架构对树结构查询的加速效果显著**：

- 如图所示，在B-Tree及其变种（B*Tree、B+Tree）的查询任务中，TTA相比纯GPU通用核心的基线方案，最高实现了**5.4倍**的性能提升。

TTA与TTA+仿真加速效果展示：B-tree和N-body仿真



B树及其变种及N体模拟加速效果图



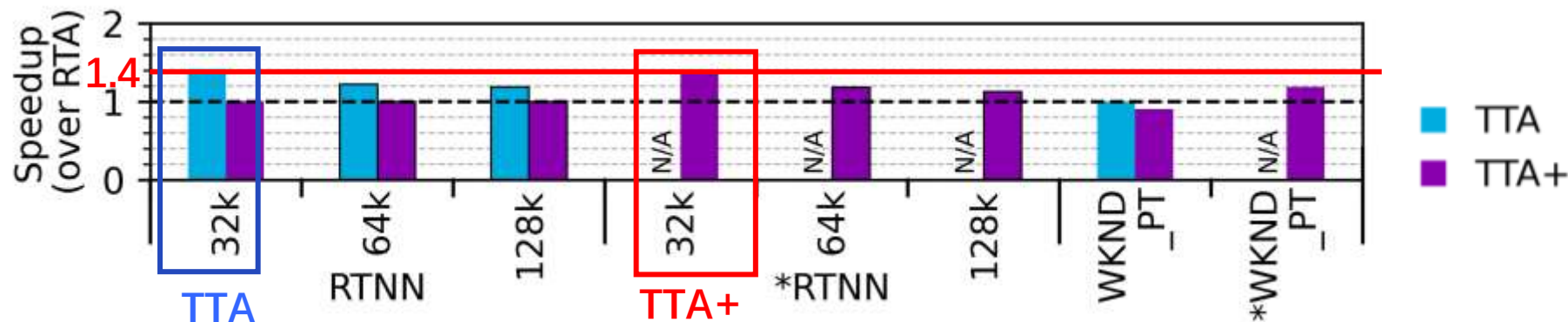
效果：

- 如图所示，我们观察到N-Body仿真的性能进一步提高了 **1.2x**，这使得TTA+ 的整体加速提高到 **1.9x**。
- **这一结果表明，有机会通过利用 TTA 和通用内核之间的并行计算来优化 TTA+。**
 - **B 树和B*Tree 查询** 在传统 GPU 上受到明显发散的影响，TTA 中的专用遍历功能缓解了这种情况。
 - 由于 **N-Body 仿真**在树遍历后进行了大量计算，因此我们还评估了合并树遍历内核和后处理内核的性能，这使得 TTA 和通用内核能够**并行工作**。

TTA与TTA+加速效果展示：半径搜索和复杂场景仿真



半径搜索和Lumibench中球体交集仿真加速效果图



效果：

- 用 TTA 替换了昂贵的交集着色器后，RTNN 半径搜索性能就提高了 **1.4x**。
- 而对于 TTA+，即便光线包围盒（Ray-box）着色器相交延迟较长，优化替换相交着色器（*RTNN）后，我们同样观察到速度提高了 **1.4x**。
 - RTNN 和 WKND PT 应用程序已在其基准测试中使用 RTA。然而，只需用 TTA 替换昂贵的交集着色器，RTNN 半径搜索性能就提高了 **1.4x**，此外，现有的 RTNN 速度比 CUDA 半径搜索的实现更快。
 - 对于 TTA+，由于 Ray-Box 相交延迟较长，RTNN 的基线实现速度会变慢，但通过优化替换相交着色器（*RTNN），我们同样观察到速度提高了 **1.4x**。

04

PART FOUR

总结与展望

研究背景与目标

研究过程与设计

实验成果及分析

总结与展望

课题总结



南京邮电大学
Nanjing University of Posts and Telecommunications

研究目标

本研究尝试构建一种更具普适性的树遍历加速方案，在保留RTA硬件优势的基础上，探索其不同空间结构中的适用性与改造路径。

设计方案

提出了TTA和TTA+两种设计方案，通过研究RTA的架构原理和优化策略，将其成功经验推广到更广泛的树遍历应用中，验证了RTA在其他领域的应用潜力。

研究成果

显著提升了树遍历操作的性能和能效。TTA实现了B树搜索高达5.4×的加速，能耗降低62%，TTA+支持更多树遍历应用，如N体模拟和半径搜索。为硬件加速器设计提供了新思路。

展望未来：我们可以



1. **进一步优化TTA和TTA+的设计**，提升性能和灵活性。通过持续改进硬件设计和编程接口，使其能够更好地适应不同树遍历应用的需求。
2. 推广TTA在更多领域的应用，如：
 - 在机器学习中，TTA可以用于加速决策树的遍历操作，提高模型训练和预测的效率。
 - 在数据挖掘中，TTA可以用于加速数据索引和查询操作，提高数据处理效率。
3. **持续创新**：随着人工智能和大数据技术的快速发展，树遍历操作的应用将更加广泛。
 - 通过持续创新，推动相关领域的技术进步。继续将TTA的技术优势与这些领域的应用需求相结合，
 - 通过与企业合作，将TTA的技术成果应用到实际产品中，为行业发展提供支持。



南京邮电大学
Nanjing University of Posts and Telecommunications

谢谢！

恳请各位指导老师批评指正！

汇报人：罗方喆 汇报时间：2025.4.10