



香山昆明湖 V3 前端架构

李衍君 徐之皓

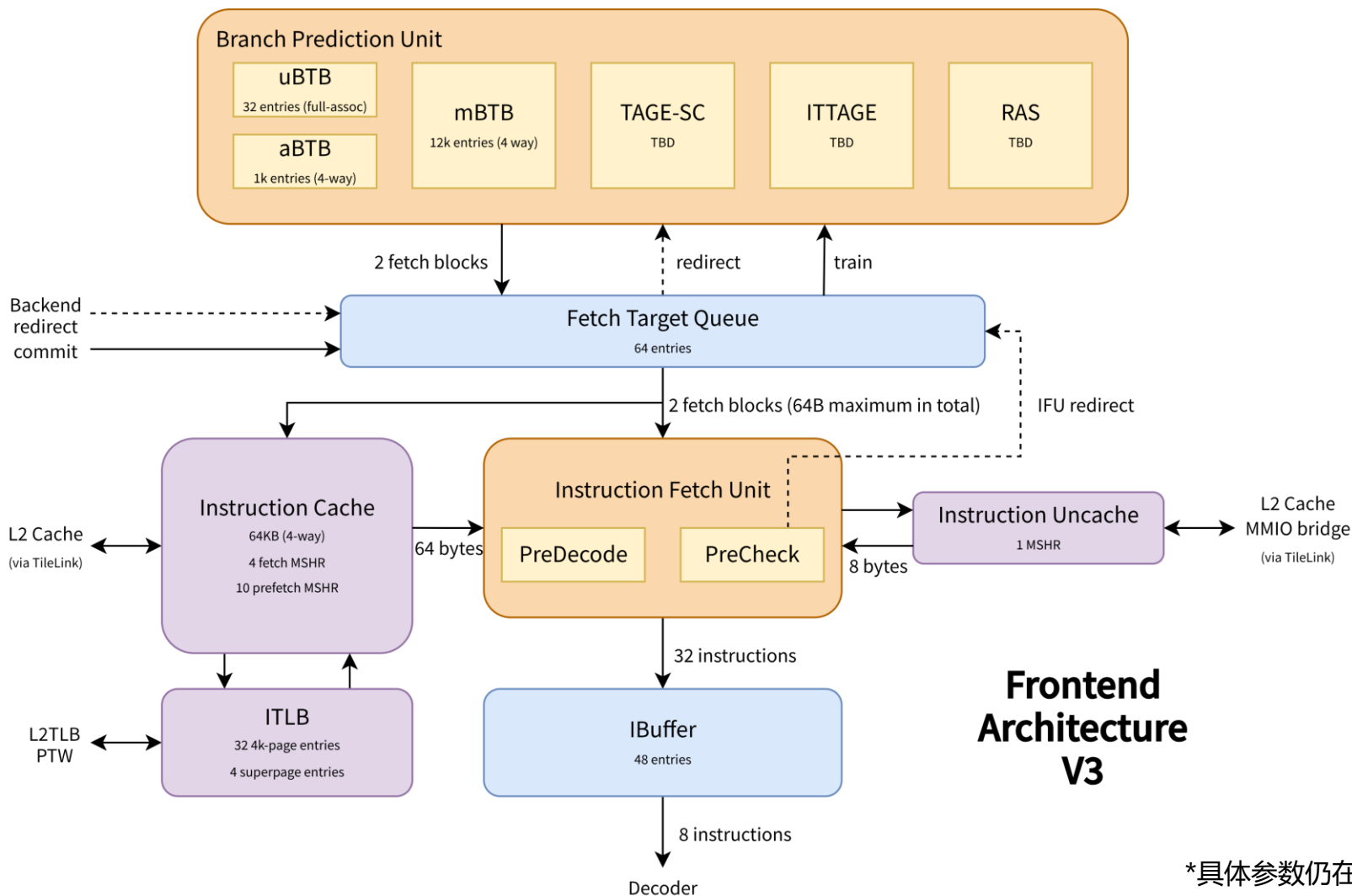
中国科学院计算技术研究所

2025 年 7 月 16 日

在一切开始之前

- Scalafmt 自动格式化
 - 代码缩进、对齐、折行等基础格式
 - `make reformat` 自动格式化, CI 加入格式检查
- 代码风格全面规范
 - 所有变量采用驼峰命名, 根据变量性质决定大驼峰/小驼峰
 - 添加 `EnumUInt` 等类, 进行常量自动检查
 - 尽量做到一个文件里只包含一个类
 - 尽量减少类长度, 尽可能使用函数以消除重复代码
 - 全新的传参系统
- 与开发同步的设计文档
- 欢迎大家加入香山开发!

昆明湖 V3 前端整体设计



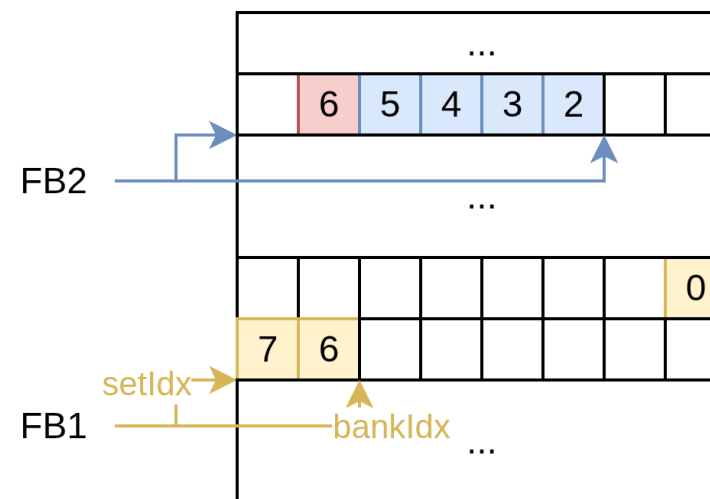
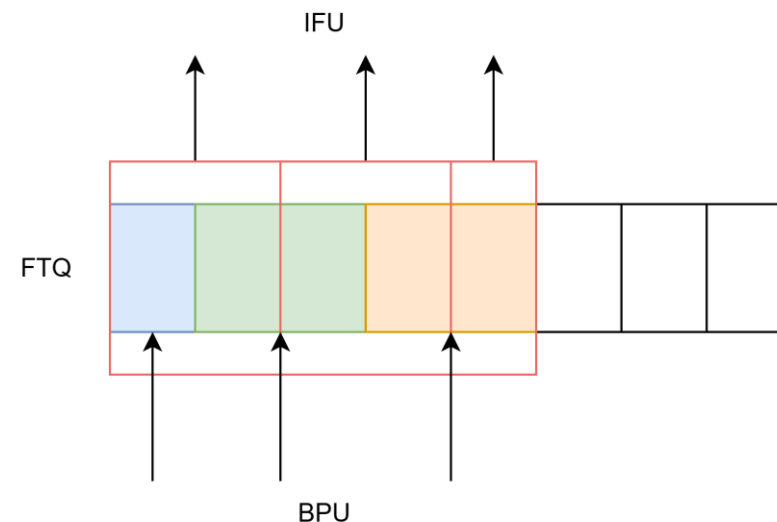
Frontend Architecture V3

*具体参数仍在探索, 仅供参考

昆明湖 V3 前端整体架构升级

- 目标：为 8 发射后端提供足够指令
- 保留解耦前端设计
- 支持 2-taken 与 2-fetch 的 FTQ
 - 提交时释放 → 分生命周期细粒度管理
- 支持 2-fetch 的 IFU 与 ICache
 - 单取指块 34B → 双取指块 64B
- 更强大的分支预测器
 - 32B 预测块 → 64B 预测块
 - FTB 结构 → BTB 结构
 - 32B 内 2 个分支 → 每个对齐 32B 内 4 个分支
 - 1-taken → 受限 2-taken

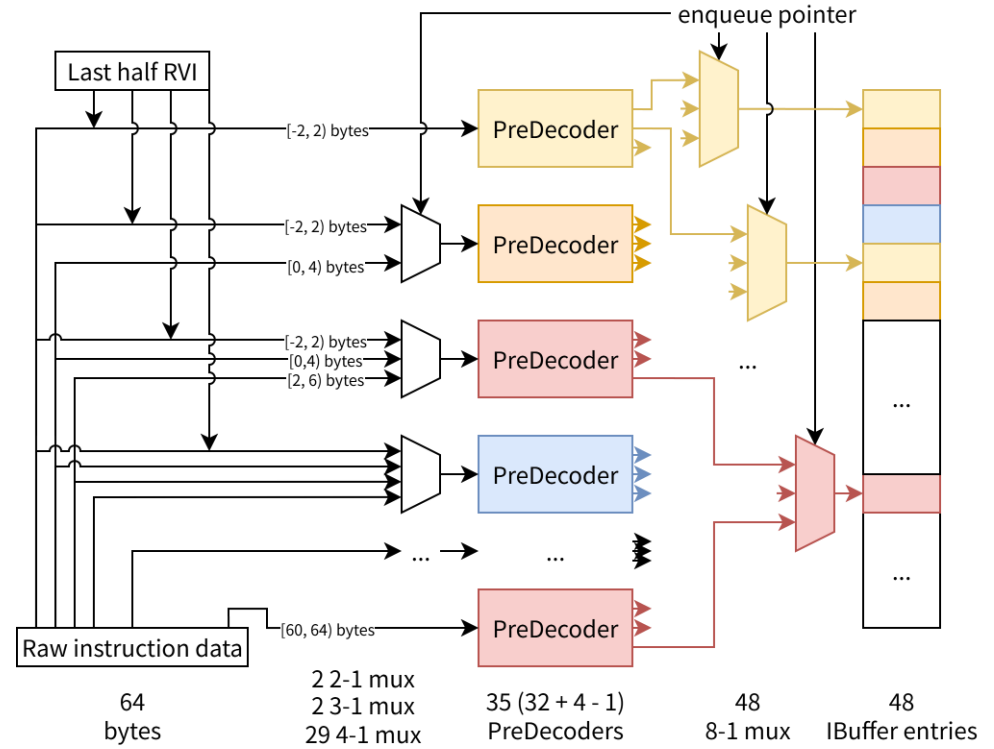
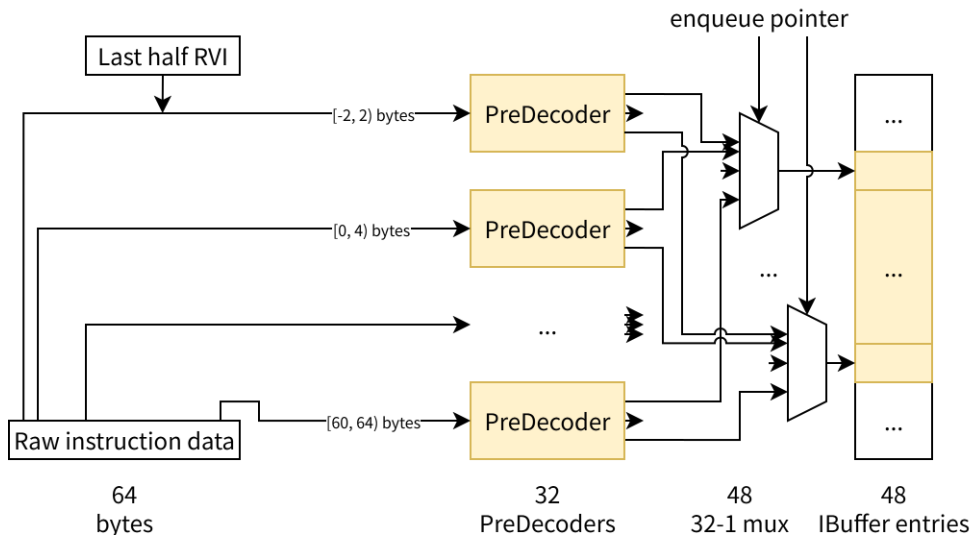
- 2-taken 与 2-fetch 支持
 - 入队和出队分开管理，互不干扰
 - 出队进行流量控制：没有冲突、总长小于 64B
- Redirect 延迟 3 → 2
 - 后端重定向后不经过 BPU，直接发往 IFU
 - 下一周期 BPU 预测旁路到 IFU
- 分生命周期管理队列
 - 预测信息在进入 IFU 后可立即释放
 - 训练信息需要在提交后送往 BPU 才能释放





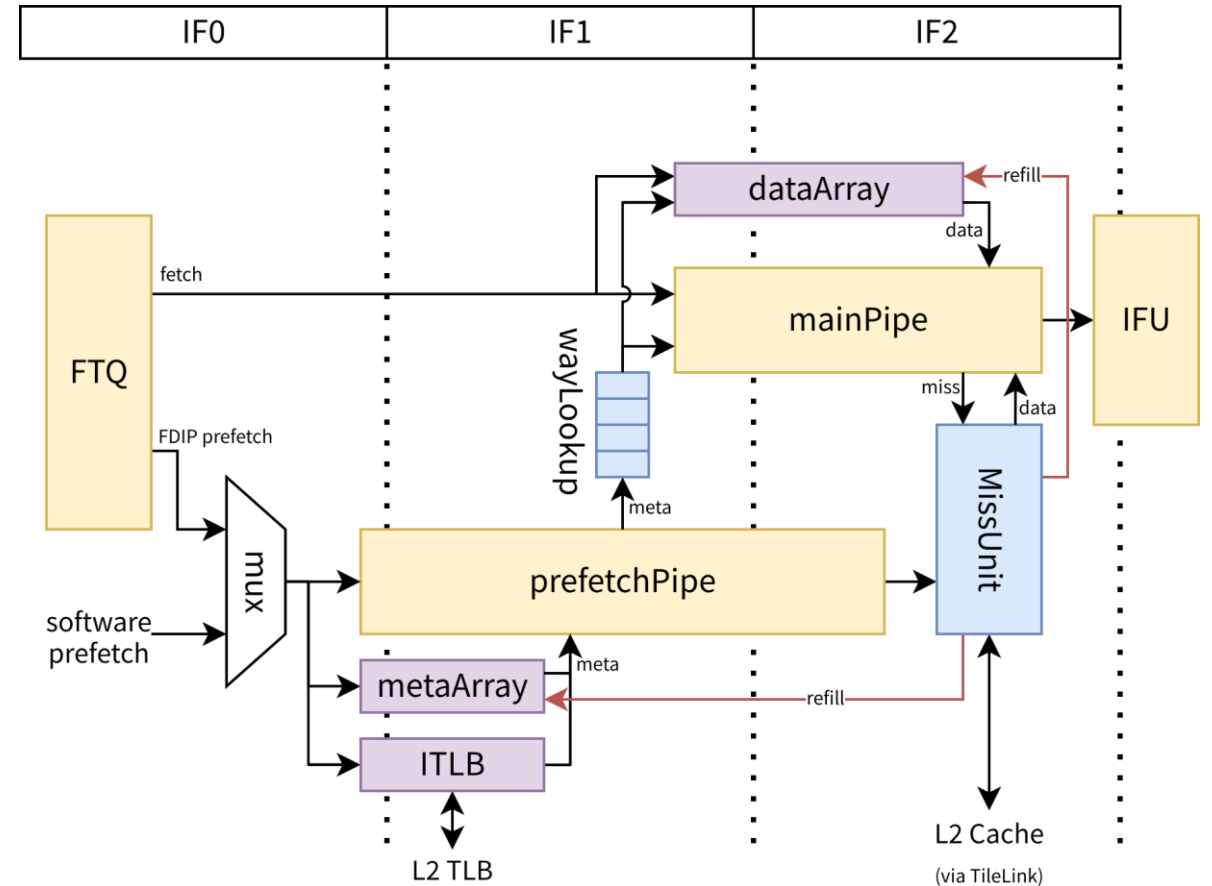
IFU & IBuffer

- 2-fetch 支持
- 单周期最大供指数 16 → 32: 时序困难
- IBuffer 入队预对齐 (左 → 右)



ICache

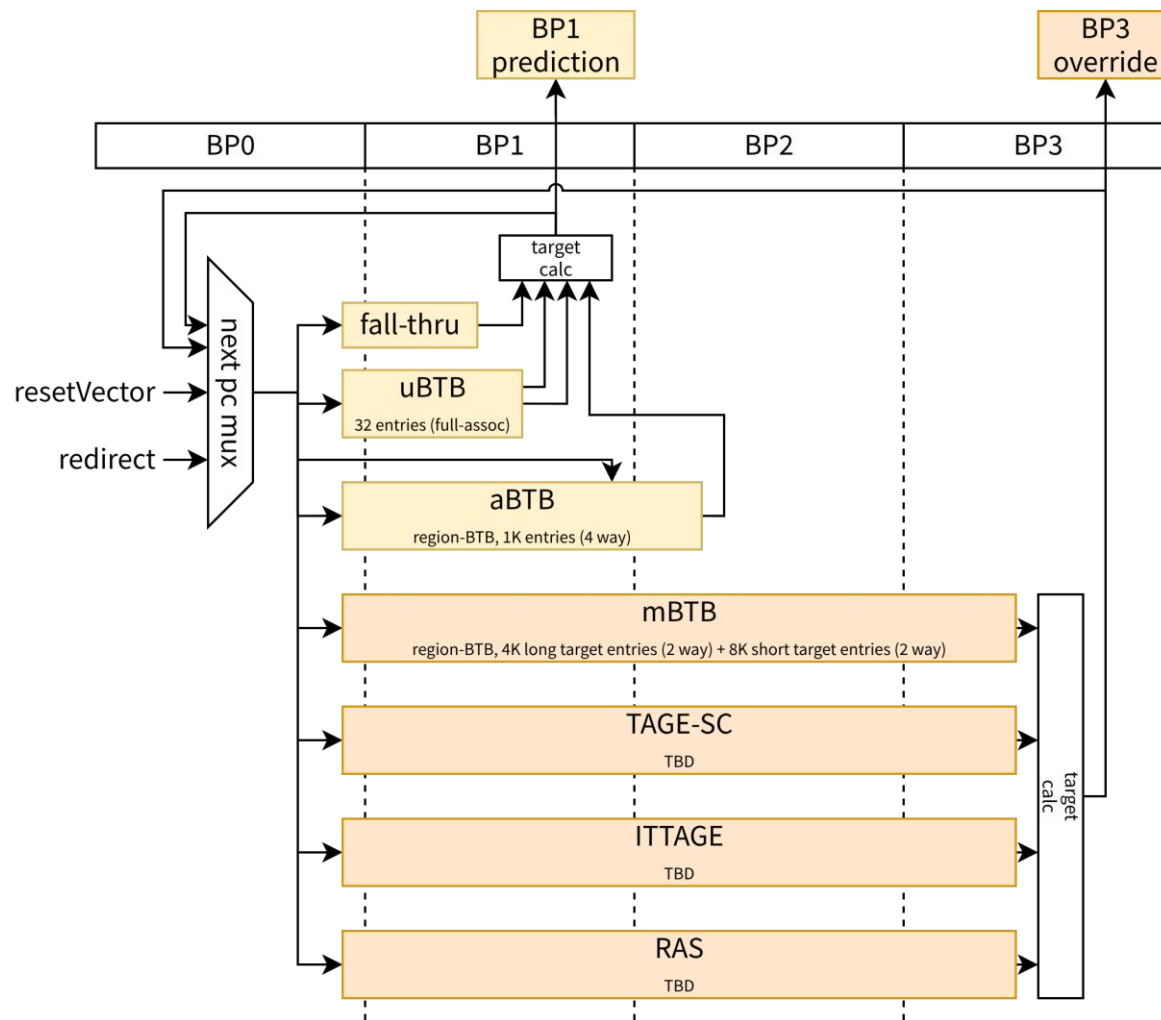
- 64KB (4 路组相联)
- 64B 缓存行
- 8B 细粒度存储
- 4 + 10 MSHR
- 支持只读命中数据的低功耗读
- 支持 Zicbop 扩展软件预取指令
- 支持 parity 校验
 - 支持出错自动恢复 (从L2重取)
- mainPipe 减拍 (new)





BPU 整体结构

- 三级预测流水 (overriding)
 - BP1: uBTB + aBTB
 - BP3: mBTB + RAS + TAGE + IT-TAGE
- 禁止预测块跨页
- Half-align
- 受限 2-taken



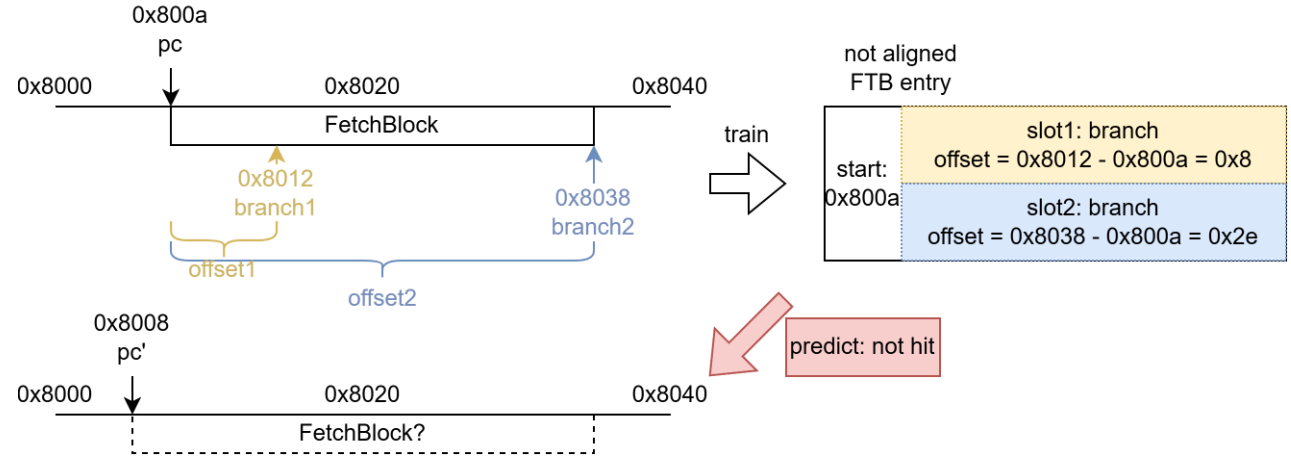
*具体参数仍在探索, 仅供参考



Half-align

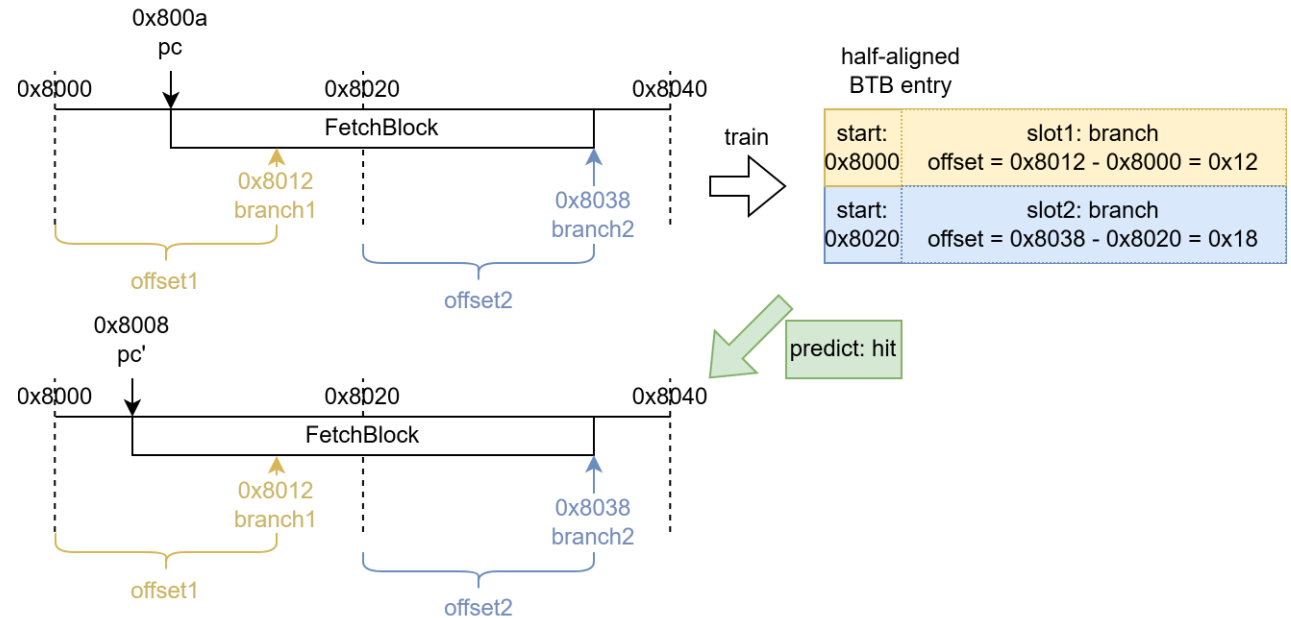
• 非对齐:

- 用起始地址索引预测器表
- 存储分支到取指块起始地址的偏移



• Half-align:

- 将起始地址对齐到取指块最大大小的一半
- 64B 取指块 → 32B 对齐
- 用对齐地址索引预测器表
- 存储分支地址低位



受限 2-taken

- 利用 10% 的面积换取 50% 的性能收益
- 仅使用 uBTB 在特定条件下生成 2-taken 预测
- 第一条为目标稳定的分支
 - 训练以来未见过不同目标的间接跳转
 - 直接跳转
 - 条件分支
- 第二条
 - 非条件：TAGE-SC 读口冲突
 - 非间接：ITTAGE 读口冲突
 - 非连续 call：RAS 写口冲突

分支1 \ 分支2	条件分支	call	其余直接跳转	ret	其余稳定间接跳转	其余不稳定间接跳转
条件分支						
稳定call						
不稳定call						
其余直接跳转						
Ret						
其余稳定间接跳转						
其余不稳定间接跳转						

Micro-BTB

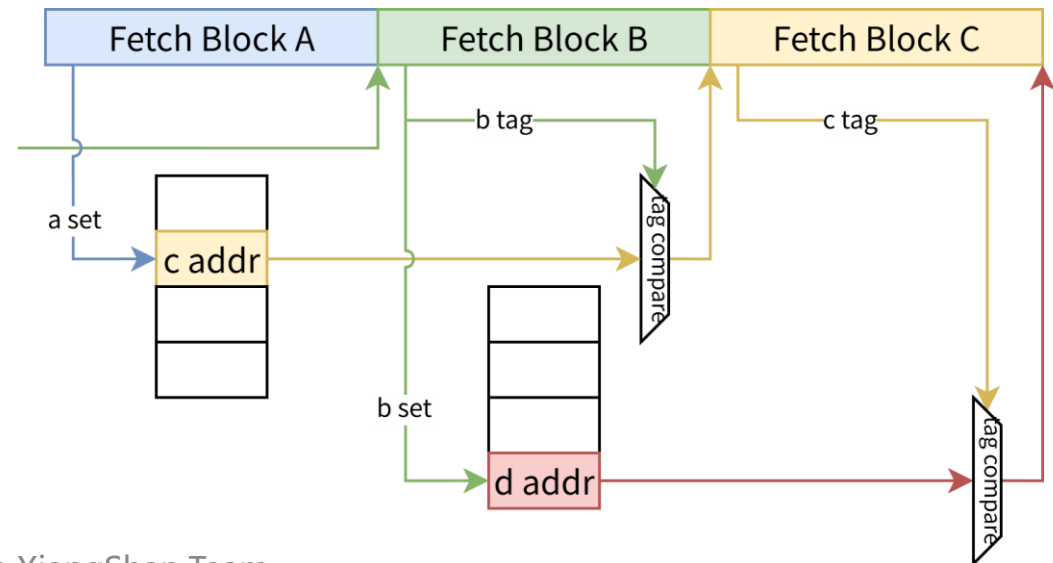
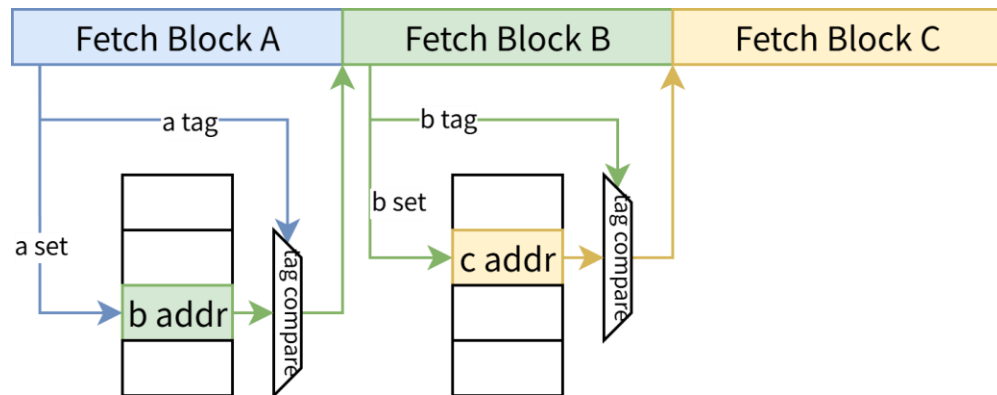
- S1 预测器
- 32 项全相联寄存器堆
- 支持受限 2-taken
- 2 bit useful counter
- Not-useful + PLRU 替换
- 2 bit taken counter

域	解释
valid	uBTB 项有效
tag	uBTB tag
usefulCnt	uBTB 项 useful 计数器, 用于替换
slot1.position	分支 1 位置
slot1.attribute	分支 1 类型
slot1.isStaticTarget	分支 1 跳转目标是否稳定
slot1.target	分支 1 跳转地址
slot1.takenCnt	分支 1 计数器
slot2.valid	分支 2 有效
slot2.position	分支 2 位置
slot2.attribute	分支 2 类型
slot2.target	分支 2 目标

🏔️ Ahead-BTB

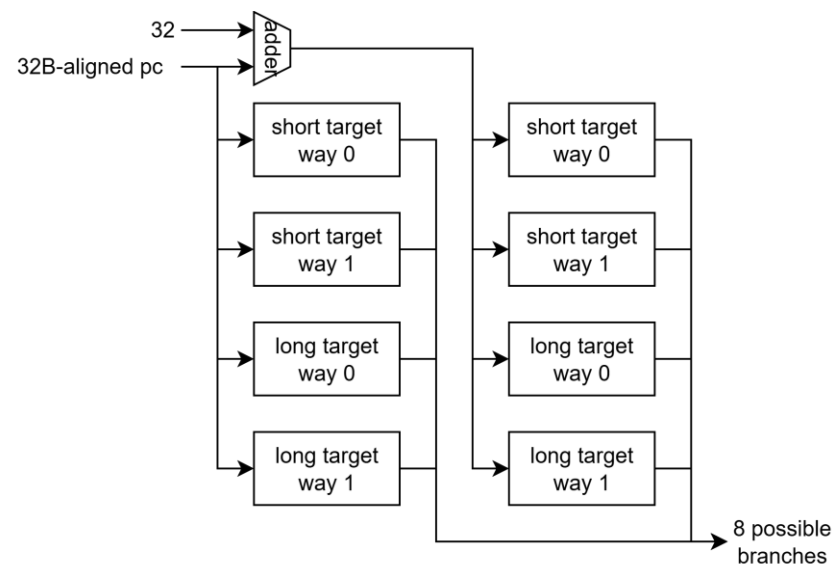
- 用 ahead pipeline 实现的大容量的 S1 预测器
- 1K 项 (4 路组相联)

预测器	索引 (查表)	Tag (判断命中)	预测
uBTB (左)	A	A	B
aBTB (右)	A	B	C



S3 预测器

- Main-BTB:
 - 4K 长 target 项, 8K 短 target 项 (2+2 路组相联, 奇偶 bank)
 - Half-align
 - 存储分支相对 32B 对齐 pc 的偏移量
 - 预测时同时查询当前对齐和下一个对齐的 block
- TAGE-SC & ITTAGE & RAS:
 - 基本保留 V2 设计
 - 适配 BTB 预测块结构
 - 参数调优
 - TAGE-SC 做 s2 预测改为做 s3 预测
 - 时序更宽松, 允许增大容量



感谢大家对香山前端的关注!