

Московский физико-технический институт
(государственный университет)
Факультет радиотехники и кибернетики
Кафедра информатики и вычислительной техники

**Блок аппаратной предвыборки данных
в кэш первого уровня МП МЦСТ-R2000**

Студент: Аверьянов В. С. гр. 113
Научное руководство: Черепанов С. А.
Ганжа Т.В. к.т.н.

Проблема

Подсистема памяти работает недостаточно быстро

- Время доступа в L1 — 3 такта,
в L2 — 20 тактов
- Статическое планирование может не справиться с задержкой load → use в 20 тактов

Решение

Уменьшение количества промахов в L1

Цель работы

Разработать блок предвыборки данных в L1-кэш

Выбранный алгоритм

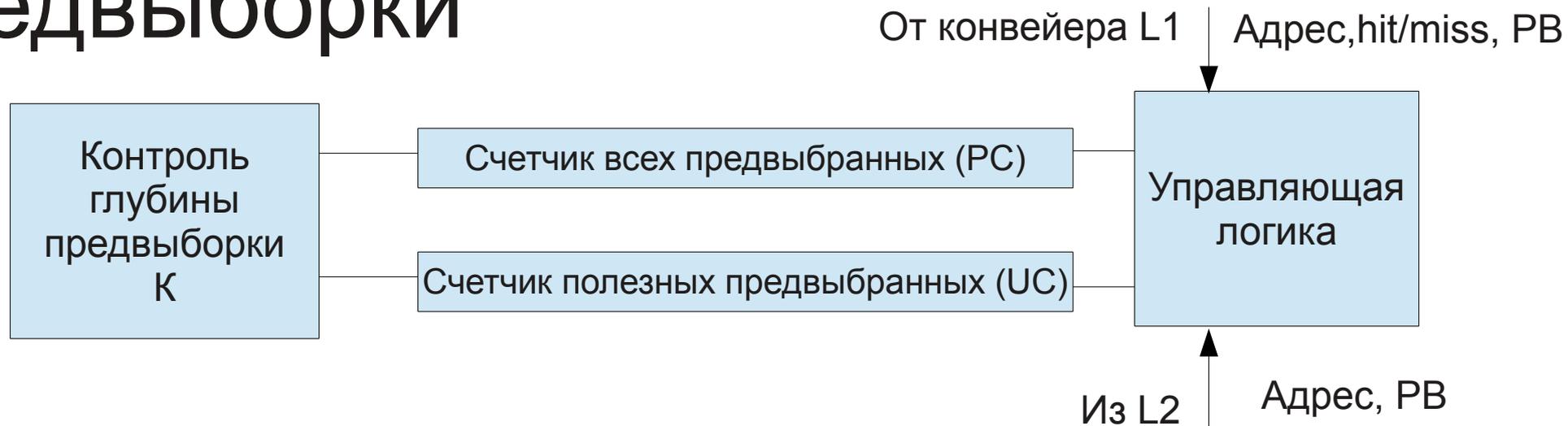
На основе результатов бакалаврской работы выбран алгоритм последовательной предвыборки с постоянным шагом и переменной глубиной Adaptive sequential prefetching

В случае промаха при обращении к строке памяти по адресу n , блок предвыборки запросит строки по адресам $n+1, n+2, \dots, n+K$. Где K — глубина предвыборки меняется по ходу исполнения программы

Достоинства алгоритма

- Малое количество аппаратуры
- Эффективен на задачах с плавающей точкой
- Возможность полного отключения предвыборки

Управление глубиной предвыборки



- Для каждой ячейки L1 вводится дополнительный бит состояния (PV) Для строки пришедшей из L2, если она была запрошена блоком предвыборки PV устанавливается 1, в противном случае PV = 0
- При использовании ядром блока с PV = 1, счетчик UC увеличивается на 1, PV блока устанавливается в 0
- На основе соотношения количества всех предвыбранных и полезных предвыбранных изменяется глубина

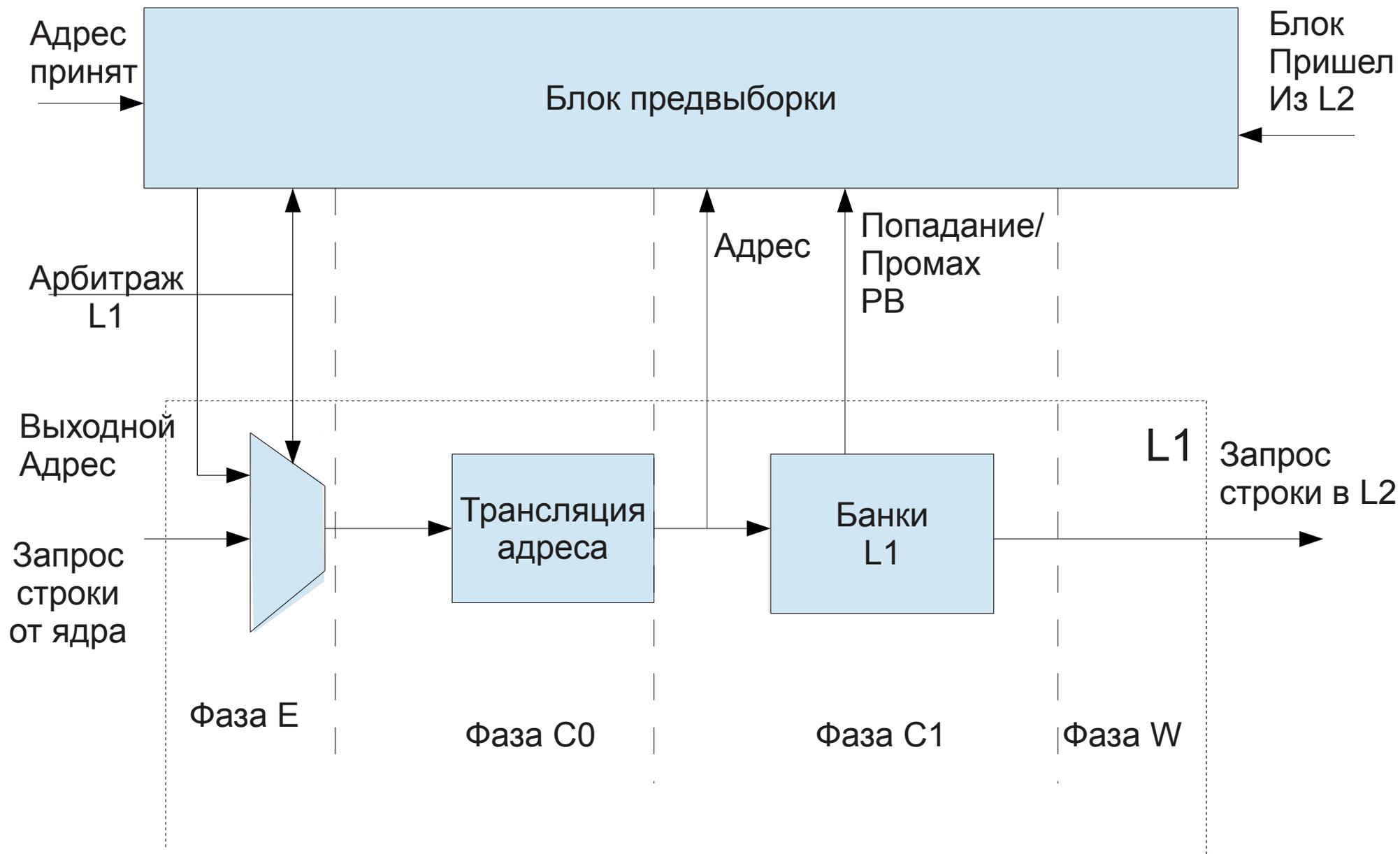
Управление глубиной предвыборки

- Вводятся параметры UT и LT (верхний и нижний пороги изменения глубины предвыборки)
- При переполнении РС происходит обновление значения K в зависимости от значения UC в данный момент:
 - $UC > UT \rightarrow K++$
 - $LT \leq UC \leq UT \rightarrow K$
 - $UC < LT \rightarrow K--$
- После обновления значения глубины РС и UC обнуляются

Требования

- Разработать блок предвыборки как отдельный модуль L1D-кэша
- Малый объем аппаратуры
- Работа на частоте 2ГГц
- Возможность полного отключения предвыборки

Взаимодействие с конвейером L1



Особенности реализации

Механизм отмены

- Исходная схема не учитывала механизм отмены в конвейере
- Решение — добавить дополнительный входной сигнал
- Выдается при отмене в конвейере либо при промахе в TLV
- Получив данный сигнал, блок предвыборки перестает выдавать адреса до следующего промаха

Физический дизайн

- Для достижения требуемой частоты в 2 ГГц выдача адресов в блок перенесена на фазу W

Взаимодействие с конвейером L1

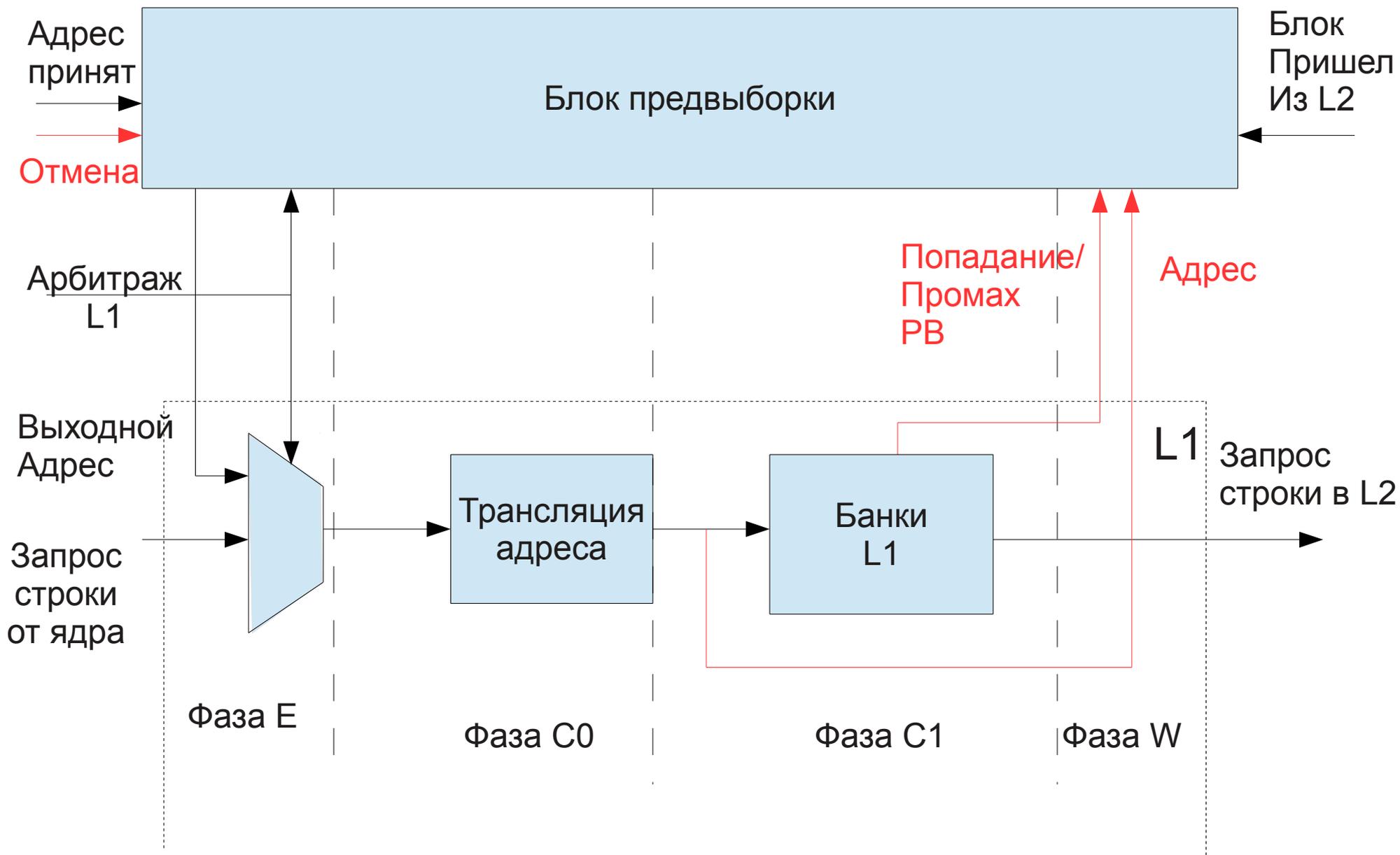
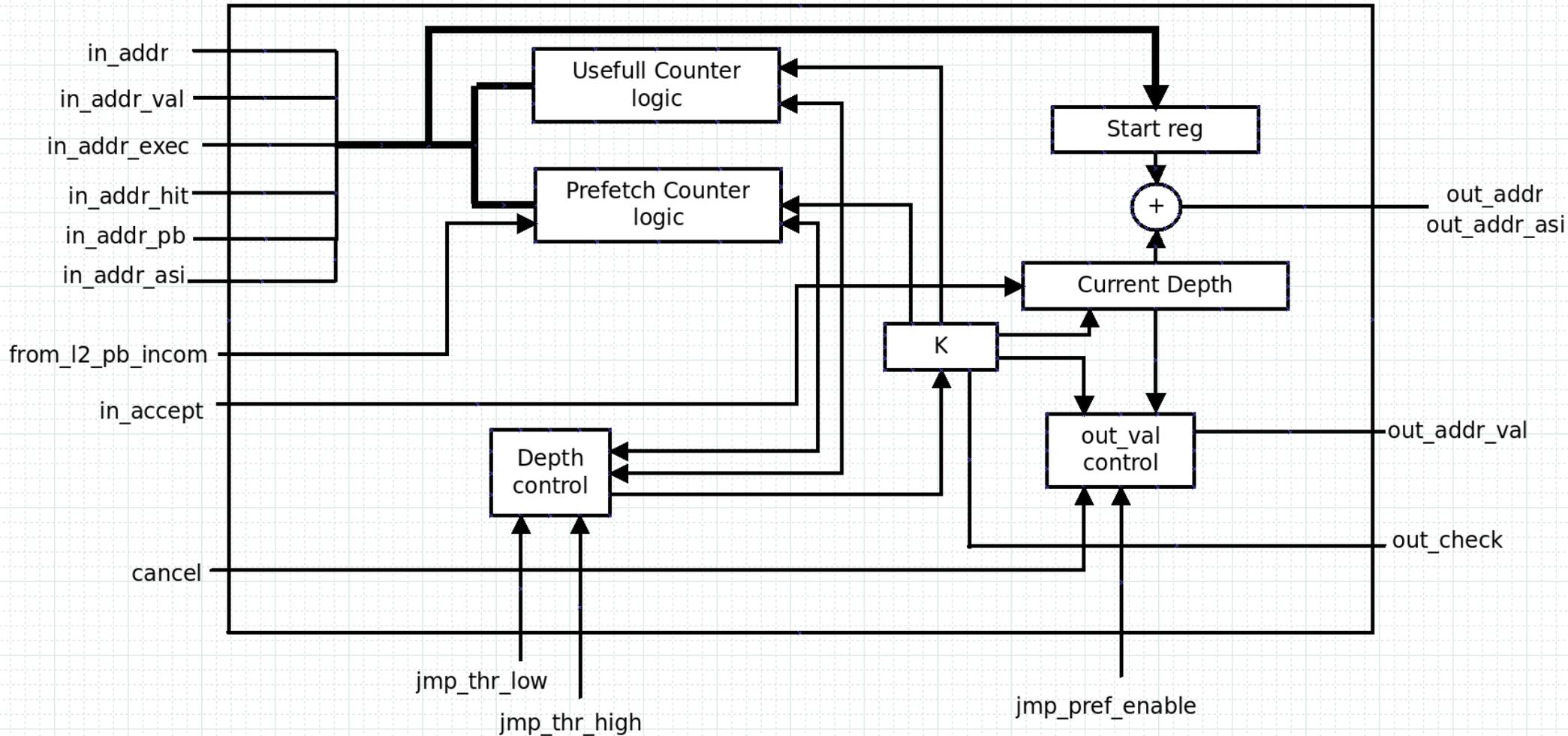


Схема блока предвыборки



- `jmp_pref_enable` — полное отключение предвыборки
- `jmp_thr_low` и `jmp_thr_high` задают верхний и нижний пороги изменения глубины предвыборки

Апробация

	mgrid	equake	ammp	parser	total
Предвыборка включена, сек	04:11:41	02:10:08	01:25:56	01:40:51	43:58:57
Предвыборка отключена,сек	02:16:55	00:57:25	01:26:23	01:46:32	38:53:22
Прирост производительности	45,60%	55,88%	-0,52%	-5,64%	11,58%
Уменьшение промахов на модели	90,62%	86,89%	2,26%	11,76%	30,37%

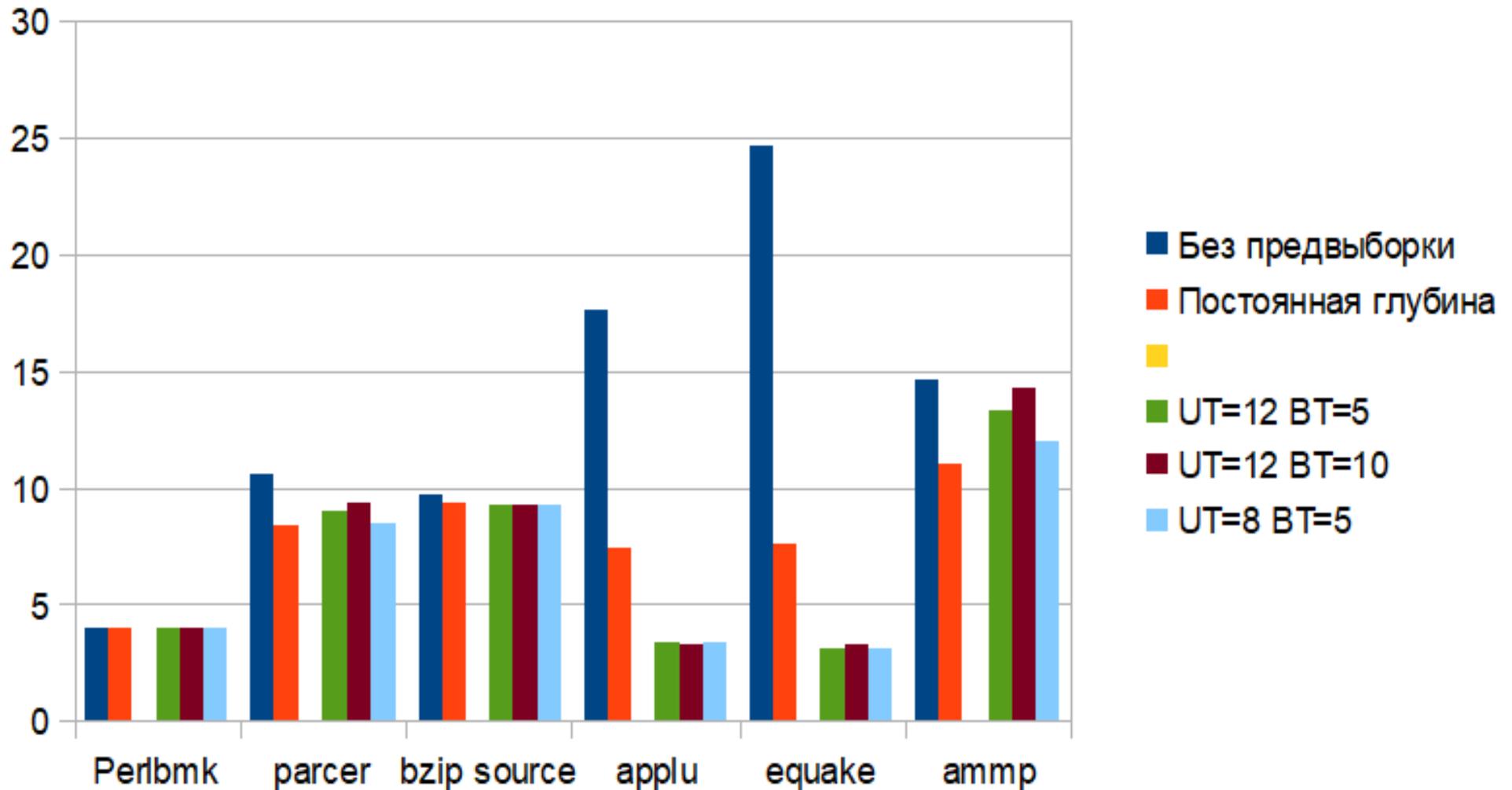
Исполнение тестов из пакета SPEC2000
на прототипе и модели

Результаты работы

- Разработано RTL-описание блока предвыборки данных в кэш первого уровня
- Блок предвыборки данных включен в состав МП R-2000
- Включение блока существенно ускорило нахождение логических ошибок в L1 на прототипе МП
- На прототипе МП время исполнения тестов пакета SPEC2000 уменьшилось с 44 часов до 39 часов

Корреляция с моделью

Промахов на 1000 инструкций



Результаты работы (на C модели)

