

Исследование зависимости производительности DSP-ядра от глубины его конвейера инструкций

А.А. Беляев

ГУП НПЦ «ЭЛВИС», bel@elvees.com

Аннотация — Анализируется зависимость производительности процессорного ядра сигнальной обработки (DSP-ядра) от глубины его конвейера инструкций при наличии в исполняемом приложении программных переходов и зависимостей по данным.

Ключевые слова — конвейер инструкций, DSP-ядро, зависимость по данным, оптимизация конвейера.

I. ВВЕДЕНИЕ

Входящие в состав систем на кристалле DSP-ядра (DSP - Digital Signal Processing) представляют собой процессорные IP-ядра, предназначенные для высокоскоростной обработки данных в реальном масштабе времени. Одним из основных архитектурных методов повышения производительности таких процессорных ядер является конвейерная обработка команд.

Вместе с тем, известно, что у конвейеризации есть определенные недостатки - при увеличении числа фаз на производительности конвейера начинают сказываться дополнительные такты задержки, возникающие при наличии в исполняемой программе зависимостей по данным между инструкциями и при отработке программных переходов. В связи с этим, при разработке новых процессоров, актуальной является задача построения оптимального по производительности конвейера инструкций с учетом статистических характеристик исполняемых программ.

Задача оптимизации конвейера становится ещё более актуальной при проектировании DSP-ядер для многоядерных систем на кристалле. Это связано не только с повышающимися требованиями по производительности, но и с постоянно возрастающей сложностью таких систем; а также с усложнением самого процесса проектирования при одновременном сокращении его сроков. При проектировании DSP-ядер для систем на кристалле должны учитываться также характеристики используемого технологического базиса (библиотеки элементов).

Данная работа посвящена анализу зависимости производительности DSP-ядра от глубины его конвейера инструкций, при наличии в исполняемом приложении программных переходов и зависимостей по данным. При этом, исследуется влияние на вид этой зависимости как статистических характеристик ис-

полняемой программной трассы, так и параметров конвейера и используемого технологического базиса.

Результаты данной работы основываются на исследованиях, проведенных при создании серии DSP-ядер ELcore-xxTM для сигнальных контроллеров семейства «Мультикор» [1,2].

II. ПРОИЗВОДИТЕЛЬНОСТЬ DSP-ЯДРА КАК ФУНКЦИЯ ГЛУБИНЫ ЕГО КОНВЕЙЕРА

Любой программный конвейер может быть разделен на две части – установочную (подготовительную) и исполнительную. Каждая из них, в свою очередь, может состоять из нескольких конвейерных фаз. Во время установочной части конвейера происходит выборка инструкции из программной памяти, ее декодирование и выборка операндов. Во время исполнительной части выполняется функциональная операция и результат заносится в ячейку памяти или регистр.

Потери, связанные с программными переходами, определяются длительностью установочной части конвейера, так как при каждом программном переходе происходит перезагрузка всей установочной части. Торможение же конвейера, вызванное наличием зависимостей по данным в исполняемом коде, напротив, зависит только от числа фаз его исполнительной части. Механизм влияния программных переходов и зависимостей по данным на производительность конвейера инструкций может быть проиллюстрирован на примере одного из DSP-ядер платформы «Мультикор» - DSP-ядра ELcore-18TM [3]. Программный конвейер этого DSP-ядра при выполнении вычислительной операции состоит из семи фаз:

- 1 фаза (A): формирование адреса памяти программ;
- 2 фаза (F): выборка инструкции из программной памяти;
- 3 фаза (D): декодирование инструкции;
- 4 фаза (E): чтение данных из регистрового файла;
- 5 фаза (E1): исполнение инструкции;
- 6 фаза (E2): исполнение инструкции;
- 7 фаза (E3): исполнение инструкции, запись данных в регистровый файл.

При такой организации программного конвейера выполнение инструкции программного перехода тре-

бует перезагрузки первых четырех (установочных) фаз. Работа такого конвейера, при выполнении инструкции программного перехода J #A, проиллюстрирована временной диаграммой на рис. 1.

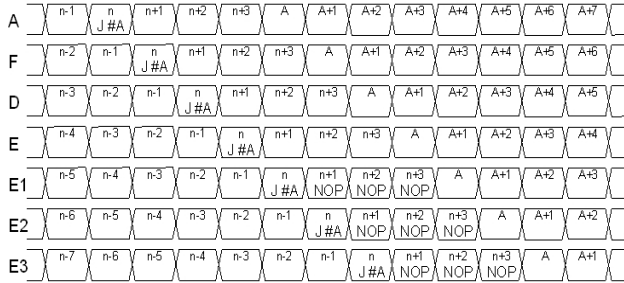


Рис. 1. Временная диаграмма работы программного конвейера DSP-ядра ELcore-18™ при выполнении инструкции программного перехода J #A

Другим источником временных потерь, возникающих при конвейеризации вычислительных операций, является наличие в программном коде рекурсивных зависимостей по данным между вычислительными инструкциями. Под данными, в общем случае, подразумеваются как числовые операнды, так и их признаки, используемые в качестве кодов условий (condition codes) при исполнении условных инструкций.

Зависимость по данным между вычислительными инструкциями имеется в тех случаях, когда текущая исполняемая инструкция использует в качестве входного операнда результат предыдущей инструкции. Пример программного кода на ассемблере ELcore-18™, с зависимостью по данным между вычислительными инструкциями, приводится ниже (индекс в скобках указывает номер инструкции):

```
(n-1)   ADD R1, R2, R3
(n)     MPF R3, R4, R5
(n+1)   ASL R6, R7, R7
(n+2)   SUB R5, R8, R8
```

В приведенном примере, инструкция (n) использует в качестве входного операнда значение, хранящееся в регистре R3, являющееся результатом инструкции (n-1), а инструкция (n+2) - значение R5, являющееся результатом инструкции (n).

Для получения правильного результата, при работе такой программы (рис.2) в конвейер вводятся дополнительные такты торможения (NOP) – два такта перед исполнением инструкция (n) и один такт перед исполнением инструкция (n+2). Длительность торможения конвейера зависит от расстояния в трассе между зависимыми инструкциями. Расстоянием зависимости между двумя инструкциями (разрешающей и зависимой) называется число инструкций, расположенных между ними плюс одна (то есть расстояние между инструкциями (n) и (n+k) равно k).

Если расстояние зависимости между двумя инструкциями i оказывается меньше, чем число фаз ис-

полнительной части конвейера N_E , то это приводит к торможению конвейера на $(N_E - i)$ тактов.

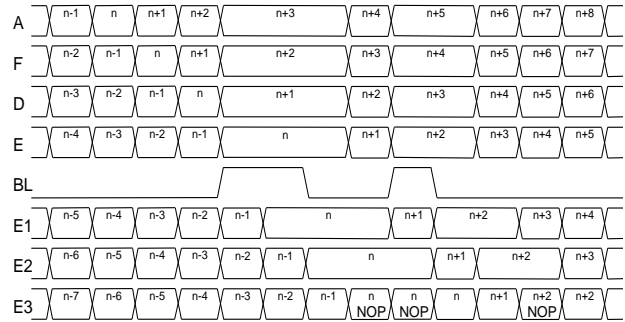


Рис. 2. Временная диаграмма работы конвейера DSP-ядра ELcore-18™ при выполнении фрагмента программы с наличием зависимости по данным

В общем виде влияние зависимостей по данным и программных переходов на производительность конвейера выражается формулой [4]:

$$BW^{-1} = 1 + p_b(N_S - 1) + \sum_{i=1}^{N_E-1} p_i(N_E - i) \quad (1)$$

где: BW^{-1} - величина, обратная производительности программного конвейера, имеющая смысл среднего числа тактов, затрачиваемых на исполнение одной инструкции; N_S и N_E – число фаз, соответственно, установочной и исполнительной части конвейера; p_b – вероятность исполненных инструкций программных переходов (т.е. отношение числа таких инструкций к общему числу исполняемых инструкций); p_i – вероятность вычислительных инструкций, зависимых по данным, с расстоянием зависимости i .

Переход к одному аргументу делается в предположении, что при увеличении глубины конвейера соотношение N_S и N_E остается постоянным.

Пусть E/S – отношение комбинационных задержек (а следовательно, и числа фаз) исполнительной и установочной части конвейера. Без ограничения общности, можно считать, что E и S – взаимно простые натуральные числа, одно из которых или оба могут быть равны единице. Тогда может быть введен один параметр – n , характеризующий глубину (или «гранулярность») конвейера, такой, что: $N_E = nE$, $N_S = nS$, и общая длина конвейера равна: $N = N_S + N_E = n(E+S)$.

Таким образом, при заданном соотношении E/S конвейер может быть охарактеризован единственным параметром n , определяющим его глубину. Формула (1) в этом случае может быть переписана в виде:

$$BW_{E,S}^{-1}(n) = 1 + p_b(nS - 1) + \sum_{i=1}^{nE-1} p_i(nE - i). \quad (2)$$

На основе полученной формулы может быть определена оптимальная глубина конвейера.

Пусть C - суммарная комбинационная задержка процессора (то есть полная комбинационная задержка

процессора при условии удаления из него всех триггеров), включая комбинационную задержку установочной C_S и исполнительной части C_E ; L - общие временные затраты на срабатывание триггера (включая setup time, hold time, clock skew, jitter и т.д.).

Тогда неконвейеризованная вычислительная машина будет выполнять каждую инструкцию за время:

$$T_{ser} = C + L = C_S + C_E + L.$$

При этом число тактов на инструкцию будет, по определению, равно единице. Процессор, конвейеризованный на N фаз, будет иметь рабочий период тактовой частоты:

$$T_{pip} = (C/N) + L.$$

Среднее число тактов, затрачиваемых на исполнение одной инструкции, для такого процессора будет определяться формулой (2).

Тогда, с учетом как изменения периода тактовой частоты, так и торможений конвейера из-за программных переходов и зависимостей по данным, общий эффект от конвейеризации, т.е. увеличение производительности (пропускной способности) конвейеризованного процессора по сравнению с неконвейеризованным, будет определяться согласно формуле:

$$\Psi_{E,S}(n) = \frac{T_{ser} BW_{ser}^{-1}}{T_{pip} BW_{pip}^{-1}} = \left(1 + p_b(nS - 1) + (n - k)E \sum_{i=1}^{kE-1} p_i + D_{E,S}(k) \right)^{-1} \left[\frac{n(S + E)(\gamma + 1)}{n(S + E) + \gamma} \right] \quad (3)$$

III. ПАРАМЕТРЫ, ВЛИЯЮЩИЕ НА ПРОИЗВОДИТЕЛЬНОСТЬ ПРОГРАММНОГО КОНВЕЙЕРА

Как было сказано выше, полученная функция определяет общий эффект от конвейеризации, т.е. увеличение производительности конвейеризованного процессора по сравнению с неконвейеризованным. Другими словами, данная функция представляет собой отношение производительности конвейеризованного процессора к производительности неконвейеризованного. Поскольку при проектировании DSP-ядер повышение производительности рассматривается как основная задача, именно эта функция является объектом анализа. Задача состоит не только в том, чтобы определить оптимальную глубину конвейера, но и исследовать влияние параметров конвейера и статистических характеристик исполняемых программ на достигаемую производительность.

Рассмотрим полученную формулу с целью определения качественного характера зависимости этой величины от входящих в формулу параметров. Вид рассматриваемой функции зависит от параметров: γ - отношение суммарной комбинационной задержки процессора к общему времени срабатывания триггера; от параметров конвейера E и S ; а также от статистических характеристик исполняемой программы.

Исследования были проведены на примере моделирования двух реальных программных приложений [5,6]: программы вычисления собственных значений матрицы - "Eigenvalue" (EIGEN) и программы, реализующей метод исключения Гаусса - "Gaussian elimination method" (GAUSS).

Для указанных приложений по их программным трассам были определены статистические характеристики p_b и p_i , значения которых приведены в табл.1.

Таблица 1

Статистические характеристики трассы для приложений EIGEN и GAUSS

Приложение	p_b	p_1	p_2	p_3	p_4	p_5
EIGEN	0.08	0.521	0.0186	0.0061	0.0068	0.0014
GAUSS	0.12	0.566	0.0554	0.0004	0.00035	0.0033

На приводимых ниже графиках проиллюстрирован характер рассматриваемых зависимостей.

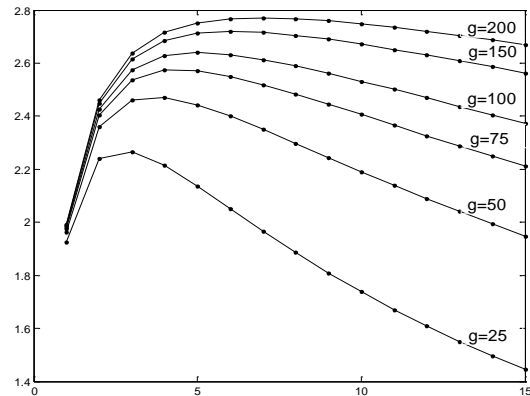


Рис. 3. Функция $\Psi_{E,S}(n)$ для приложения EIGEN при различных значениях γ ($\gamma = g$), $E=S=1$

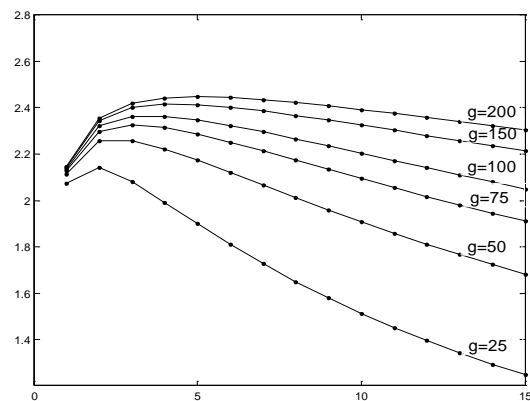


Рис. 4. Функция $\Psi_{E,S}(n)$ для приложения GAUSS при различных значениях γ ($\gamma = g$), $E=S=1$

На рисунках 3 и 4 представлена функция $\Psi_{E,S}(n)$ для приложения EIGEN (рис.3) и приложения GAUSS (рис.4) при различных значениях величины γ ($\gamma = g$) и параметрах конвейера $E=S=1$. На рисунках 5 и 6

представлена функция $\Psi_{E,S}(n)$ для приложения GAUSS при различных значениях вероятностной характеристики p_1 (для $i>1$ $p_i=0$; $p_S=0.12$), для двух значений величины γ : $\gamma=100$ (рис.5) и $\gamma=25$ (рис.6) и параметрах конвейера $E=S=1$.

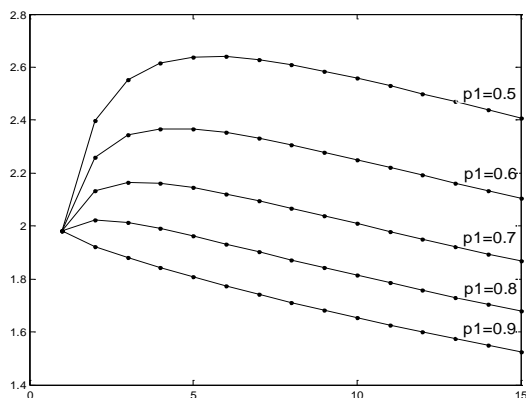


Рис. 5. Функция $\Psi_{E,S}(n)$ для приложения GAUSS при различных значениях p_1 , $\gamma=100$, $E=S=1$

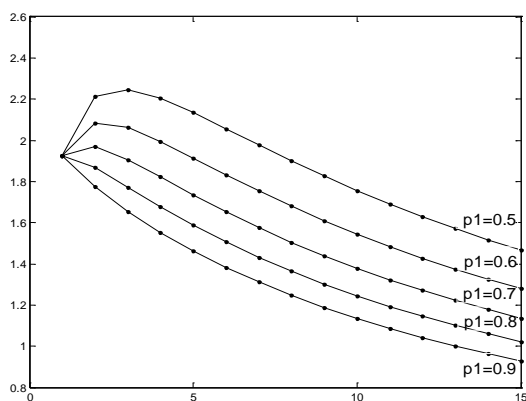


Рис. 6. Функция $\Psi_{E,S}(n)$ для приложения GAUSS при различных значениях p_1 , $\gamma=25$, $E=S=1$

IV. АНАЛИЗ РЕЗУЛЬТАТОВ

Как видно из представленных на рис.3 и 4 графиков, при увеличении отношения общей комбинационной задержки к времени срабатывания триггера γ , точка оптимума смещается, как и следовало ожидать, в сторону более глубокой конвейеризации.

Графики, приведенные на рис.5 и 6, позволяют оценить те граничные значения вероятностной характеристики p_1 (имеющей смысл вероятности дуг зависимости с единичным расстоянием), выше которой конвейеризация вообще теряет смысл. Например, если $\gamma=25$, для $p_1 > 0.7$ функция $\Psi_{E,S}(n)$ имеет убывающий характер и достигает максимума при $n=1$. Для процессора, характеризующегося отношением общей комбинационной задержки к времени срабатывания триггера $\gamma=100$, граница находится в пределах $p_1 = 0.8 - 0.9$.

V. ЗАКЛЮЧЕНИЕ

Создание высокопроизводительных сигнальных процессоров для систем на кристалле связано с выбором оптимальной, с точки зрения быстродействия, структуры программного конвейера для DSP-ядер с гарвардской архитектурой.

Проведенный в данной работе анализ наглядно показывает преимущества, которые дает математический подход к оптимизации структуры конвейера инструкций по сравнению с компьютерным моделированием, не подкрепленным теоретической моделью. Такой анализ дает возможность разработчику DSP-ядра ориентироваться в пространстве возможных решений по поиску оптимальной структуры его программного конвейера, тем самым ускоряя и облегчая процесс проектирования. Полученные аналитические выражения для производительности конвейера инструкций и для оптимальной его глубины, при выполнении приложений с программными переходами и зависимостями по данным, позволяет на архитектурном уровне осуществить выбор оптимальной структуры конвейера DSP-ядра с учетом статистических характеристик исполняемых программных трасс.

ЛИТЕРАТУРА

- [1] Петричкович Я.Я., Солохина Т.В. Цифровые сигнальные контроллеры «МУЛЬТИКОР» - новые отечественные серии систем на кристалле // Доклады конференции «Цифровая обработка сигналов и ее применения (DSPA-2004)». - 2004. - Т. 1. - С. 8—15.
- [2] Солохина Т.В., Петричкович Я.Я., Глушков А.В. Архитектура отечественных серий микросхем типа «система или сеть на кристалле» на базе IP – библиотек платформы «МУЛЬТИКОР» // Проблемы разработки перспективных микроэлектронных систем. Сб. научных трудов / под общ. ред. А.Л. Стемповского. - М.: ИППМ РАН. - 2005. - С. 419—426.
- [3] Беляев А.А. Организация программного конвейера DSP-ядер серии ELcore-xx™ IP-Библиотеки «МУЛЬТИКОР» // Проблемы разработки перспективных микроэлектронных систем. Сб. научных трудов / под общ. ред. А.Л. Стемповского. - М.: ИППМ РАН. - 2005. - С. 508—511.
- [4] Emma P.G., Davidson E.S. Characterization of Branch and Data Dependencies in Programs for Evaluating Pipeline Performance // IEEE Trans. On Computers, July 1987. V. C-36, NO.7. - P. 859—875.
- [5] Беляев А.А. Влияние программных переходов и зависимостей по данным в исполняемом программном коде на производительность конвейера DSP-ядра // Известия высших учебных заведений // Электроника. - М.: МИЭТ, 2009. - №3 (77). - С. 75—80.
- [6] Беляев А.А. Влияние глубины конвейера на производительность процессора // Известия высших учебных заведений // Электроника. - М.: МИЭТ, 2009. - №6 (80). - С. 50—53.