Раздел VI. Аппаратно-программные средства обеспечения безопасности

УДК 004.432

А.И. Дордопуло, Д.А. Сорокин

МЕТОДИКА СОКРАЩЕНИЯ АППАРАТНЫХ ЗАТРАТ В СЛОЖНЫХ СИСТЕМАХ ПРИ РЕШЕНИИ ЗАДАЧ С СУЩЕСТВЕННО-ПЕРЕМЕННОЙ ИНТЕНСИВНОСТЬЮ ПОТОКОВ ДАННЫХ*

Предложена методика многокритериальной редукции аппаратных затрат, основанная на известных методах редукции аппаратных затрат по числу выполняемых операций, по разрядности обрабатываемых операндов, по скважности и частоте, применение которой обеспечивает принципиальную возможность решения задач с существенно переменной интенсивностью потоков данных в едином вычислительном контуре. Применение разработанной методики позволяет достигать многократного выигрыша в скорости решения задач с существенно переменной интенсивностью потоков данных при их реализации на реконфигурируемых вычислительных системах.

Реконфигурируемые вычислительные системы; сокращение аппаратных затрат; редукция производительности.

A.I. Dordopulo, D.A. Sorokin

A METHODOLOGY OF HARDWARE OVERHEAD DECREASE IN COMPLEX SYSTEMS WHILE SOLVING TASKS WITH CONSIDERABLY VARIABLE DATA FLOW DENSITY

The article covers a methodology of multicriterion reduction of hardware overheads based on well-known methods of overhead reduction by number of performed operations, by processed operands capacity, by off-duty factor and frequency. The developed methodology, in principle, makes possible to solve problems with considerably variable data flow density in a single computational resource. Usage of this methodology provides a considerable multiple gain in speed for the problem with considerably variable data flow density for reconfigurable computer systems.

Reconfigurable computer systems; hardware overhead decrease; performance reduction.

Введение. Реконфигурируемые вычислительные системы (РВС), архитектура которых может быть адаптирована под структуру решаемой задачи, применяются для решения вычислительно трудоемких задач различных проблемных областей науки и техники. Применение РВС позволяет обеспечивать высокую реальную производительность задачи, в том числе при решении сильносвязанных задач, а также рост производительности системы, близкий к линейному, при увеличении аппаратных ресурсов. В настоящее время на РВС эффективно решаются задачи, для которых размер потока данных в вычислительной структуре задачи практически одинаков в любом месте аппаратно-реализованного информационного графа при условии представления в качестве вершин информационного графа функцио-

* Исследования выполнены при финансовой поддержке Министерства образования и науки РФ, госконтракт № 14.527.12.0004 от 03 октября 2011 г.

нально законченных преобразований. В то же время существует большой класс задач, в которых размер потоков данных в разных местах вычислительной структуры отличается на 2-4 десятичных порядка. К числу таких задач относятся задачи обеспечения безопасности сложных систем, транскодирования видеоизображений, молекулярного конструирования лекарств, автоматизированного размещения элементов и трассировки электрических соединений устройств электронной техники.

Главная особенность этих задач состоит в том, что размер потока данных в их вычислительной структуре заранее не определен и зависит от самих обрабатываемых данных.

Доказано, что информационный граф задачи, структурно реализуемой на РВС, должен быть представлен в виде множества изоморфных базовых подграфов [2], которые мультиплицируются по данным и по итерациям. Для эффективного решения задач на РВС необходимо обеспечить обработку и передачу данных в едином максимально возможном темпе, что, в свою очередь, требует аппаратной реализации всего информационного графа или, по крайней мере, базового подграфа задачи. При этом для ряда задач аппаратная реализация базового подграфа может потребовать ресурса большего, чем весь ресурс РВС, вследствие чего структурная реализация задачи на РВС кажется невозможной. В противном случае при последовательной реализации подграфов задачи невозможно организовать вычислительный конвейер, поскольку после прохождения операнда через вычислительную структуру подграфа РВС должна перестроиться на реализацию следующего подграфа. Скорость обработки данных, как следствие, снижается в десятки раз и становится значительно меньшей, чем при процедурной обработке.

Поэтому для решения на РВС задач с существенно изменяющимися размерами потоков данных в вычислительной структуре необходима разработка методики сокращения аппаратных затрат, позволяющей при своем применении эффективно реализовать такие задачи в условиях ограниченного аппаратного ресурса РВС.

Постановка проблемы. Для реализации задачи на РВС необходимо оценить возможность построения вычислительной структуры базового подграфа задачи на имеющемся аппаратном ресурсе РВС. При наличии в РВС аппаратного ресурса объемом L логических вентилей для успешной реализации базового подграфа задачи, состоящей из h подзадач P, должно выполняться условие $\sum_{i=1}^{h} N(P_i) \le L$, где $N(P_i)$ – объем ресурса, затрачиваемого на реализацию подзадачи P_i. Зачастую объем аппаратного ресурса РВС много больше объема ресурса вычислительной структуры базового подграфа задачи. Можно рассчитать максимальную степень распараллеливания базового подграфа $l = \left[L / \sum_{i=1}^{h} N(P_i) \right]$.

Если ресурса РВС недостаточно для реализации даже одного базового подграфа (l<1), то, согласно традиционным методам программирования РВС, такую задачу структурно реализовать невозможно. Однако это умозаключение справедливо тогда, когда вершины информационного графа задачи (базового подграфа) рассматриваются как атомарные (программно неделимые).

Для того чтобы реализовать вычислительную структуру со степенью распараллеливания l<1, необходимо провести сокращение аппаратных затрат на её реализацию. Можно определить требуемый коэффициент сокращения $r = l^{-1}$. Объем аппаратных затрат на реализацию базового подграфа задачи составит

$A = \sum_{i=1}^{n} N(P_i) = \sum_{i=1}^{n} f_i \cdot N(p_i)$, где $N(p_i)$ – объем ресурса, затрачиваемого на реализацию базового подграфа p_i подзадачи P_i, f_i – число базовых подграфов подзадачи P_i.

Традиционно для сокращения аппаратных затрат A на реализацию базового подграфа задачи выполнялась r-кратная редукция числа базовых подграфов f_i всех подзадач P_i: $A^r = \sum_{i=1}^{n} \left[f / r_i \right] \cdot N(p_i)$. Этот прием применим для задач, у которых выполняется условие $\forall i : f_i \geq r$. В противном случае редукция по числу базовых подграфов подзадачи P_i невозможна. Чтобы преодолеть это ограничение, необходимо редуцировать базовый подграф p_i подзадачи P_i, который в традиционных методах рассматривался как неделимый.

Производительность вычислительной структуры W прямо пропорциональна общему количеству операций задачи: $W = \sum_{i=1}^{n} K_i \cdot Ш_i \cdot V_i$, где K_i – число операций базового подграфа подзадачи P_i, $Ш_i$ – ширина потока данных подзадачи P_i, V_i – скорость обработки потока данных в подзадаче P_i. Ширина потока данных рассчитывается по формуле $Ш_i = f_i \cdot \rho_i$, где ρ_i – разрядность обрабатываемых операндов. Скорость обработки потока данных в подзадаче рассчитывается по формуле $V_i = \nu_i / S_i$, где ν_i – тактовая частота работы вычислительной структуры подзадачи P_i, S_i – скважность потока данных на входе подзадачи базового подграфа.

Скважность представляет собой натуральное число и является отношением количества тактов работы вычислительной структуры к количеству данных, которые можно подать на вход вычислительной структуры за это количество тактов, т.е. через какое количество тактов можно подать на вход вычислительной структуры следующее данное. Скважность вычислительной структуры определяется наличием обратных связей в базовом подграфе задачи или необходимостью выполнять чтение или запись данных в память типа DDR по случайным адресам.

Существует корреляция между производительностью вычислительной структуры задачи и аппаратными затратами на реализацию базового подграфа задачи, поэтому возможно при снижении производительности добиться уменьшения аппаратных затрат. В общем случае снижение производительности вычислительной структуры может быть осуществлено на коэффициент редукции $r' \geq r$. Несложно заметить, что редукцию производительности можно выполнять не только уменьшением числа базовых подграфов подзадачи f_i, но и сокращением числа операций базового подграфа подзадачи K_i, разрядности обрабатываемых операндов ρ_i, тактовой частоты работы вычислительной структуры подзадачи ν_i и увеличением скважности потока данных на входе подзадачи базового подграфа S_i. Следует отметить, что редукции производительностей каждой подзадачи должны быть осуществлены на одинаковую величину, но не обязательно должны быть редуцированы одинаковые параметры в разных подзадачах.

В [3] подробно описаны методы сокращения аппаратных затрат при решении задач на РВС, среди которых можно отметить следующие наиболее часто применяемые на практике методы:

- редукция по числу выполняемых операций;
- редукция по разрядности обрабатываемых операндов;
- редукция по скважности и частоте.

Заметим, что в общем случае невозможно эффективно сократить аппаратные затраты РВС при структурном решении задачи по единому критерию. Особенно это характерно для задач с существенно-переменной интенсивностью потоков данных. Для того чтобы выполнить необходимую редукцию аппаратных затрат в таких задачах, следует использовать методику многокритериальной редукции.

Многокритериальная редукция. Предварительно счетчик итераций α устанавливается в начальное состояние (α=0).

На первом этапе необходимо оценить аппаратные затраты A^{α} на реализацию базового подграфа задачи. Если аппаратные затраты не превышают имеющийся ресурс РВС ($A^{\alpha} \le L$), то можно выполнять её структурную реализацию. В противном случае необходимо выполнять сокращение аппаратных затрат на реализацию базового подграфа задачи. Для этого рассчитывается начальное значение коэффициента редукции $r^{\alpha} = \lfloor A^{\alpha} / L \rfloor$.

На втором этапе необходимо осуществить редукцию аппаратных затрат для всех i подзадач исходного базового подграфа задачи с коэффициентом r^{α}. Для i-й подзадачи сначала пытаемся осуществить редукцию по числу базовых подграфов подзадачи f_i. Если $r^{\alpha} \le f_i$, то значение числа базовых подграфов i-й редуцированной подзадачи составит $f_i^{'} = f_i / r^{\alpha}$. В противном случае при $r^{\alpha} > f_i$ редукция i-й подзадачи по f_i с коэффициентом r^{α} невозможна, поэтому выполняем редукцию с коэффициентом, равным f_i, а затем переходим к процессу редукции следующими методами: по разрядности выполняемых операций и количеству выполняемых операций, при этом рассчитывается новое значение коэффициента редукции $r_f^{\alpha} = r^{\alpha} / f_i$.

В общем случае неизвестно, каким из двух данных методов эффективнее осуществлять редукцию. Если в задаче обрабатываются данные с фиксированной запятой, то, как правило, целесообразно сначала осуществлять редукцию по разрядности выполняемых операций ρ_i, а потом по количеству выполняемых операций K_i. Если обрабатываются данные с плавающей запятой, то порядок применения этих методов может быть обратным. Для простоты рассуждений будем считать, что приоритет для редукции по ρ_i. Тогда при $r_f^{\alpha} \le \rho_i$ выполняем редукцию по разрядности, которая составит $\rho'_i = \rho_i / r_f^{\alpha}$. При $r_f^{\alpha} > \rho_i$ редукция i-й подзадачи по ρ_i с коэффициентом r_f^{α} невозможна, поэтому переходим к редукции по K_i. Снова рассчитывается новое значение коэффициента редукции $r_{\rho}^{\alpha} = r_f^{\alpha} / \rho_i$. При $r_{\rho}^{\alpha} \le K_i$ выполняем редукцию по количеству операций, которое составит $K'_i = K_i / r_{\rho}^{\alpha}$. При $r_{\rho}^{\alpha} > K_i$ редукция i-й подзадачи по количеству операций с коэффициентом r_{ρ}^{α} невозможна.

Стоит отметить, что в отличие от редукции по числу базовых подграфов подзадачи снижение аппаратных затрат при редукции по разрядности обрабатываемых операндов и редукции по числу операций происходит нелинейно. Поэтому параметры ρ_i и K_i редуцируются не максимально, а до каких-то значений $\rho'' = \phi(\rho_i, r_f^{\alpha})$ и $K'' = \phi(K_i, r_{\rho}^{\alpha})$, при которых аппаратные затраты минимальны.

Если редукции по разрядности выполняемых операций, по количеству выполняемых операций и по числу базовых подграфов подзадачи недостаточно, то есть для остаточного коэффициента редукции $r_K^{\alpha} = r_{\rho}^{\alpha} / K_i$ выполняется условие $r_K^{\alpha} > 1$, то применяем редукцию по скважности или частоте с этим коэффициентом. При этом сугубо эмпирическим путем определяем, по какому именно из этих па-

раметров применяем редукцию, исходя из свойств решаемой задачи (подзадачи). Затем переходим к редукции следующей подзадачи, и процесс повторяется до тех пор, пока не будут редуцированы все подзадачи базового подграфа.

На третьем этапе после выполнения редукции всех i подзадач необходимо оценить аппаратные затраты A^r на реализацию редуцированного базового подграфа задачи. Если $A^r \le L$, то процесс редукции базового подграфа задачи заканчивается, иначе счетчик итераций α наращивается на 1. Вычисляется значение модифицированного коэффициента редукции $r^{\alpha+1} = r^{\alpha} + \Delta r$, где Δr – заранее заданное приращение коэффициента редукции. Для каждой i-й подзадачи проверяется условие

$$\forall i : f_i \cdot K_i \cdot \rho_i \ge r^{\alpha+1}. \quad (1)$$

Если условие (3) не выполняется для всех подзадач, то редукция аппаратных затрат невозможна и соответственно, структурная реализация задачи на РВС с ресурсом L не может быть осуществлена. В этом случае целесообразно увеличивать аппаратный ресурс РВС.

Если хотя бы для одной подзадачи условие (1) выполняется, то можно продолжать выполнять редукцию аппаратных затрат, начиная со второго этапа. Процесс повторяется до тех пор, пока не выполнится условие $A^r \le L$ или не будет превышено наперед заданное число итераций.

Стоит отметить, что для достижения наибольшей эффективности использования аппаратного ресурса РВС при решении задачи необходимо стремиться к тому, чтобы производительность вычислительной структуры редуцированного базового подграфа стремилась к максимуму при $A^r \approx L$.

Заключение. Применение разработанной методики многокритериальной редукции позволяет сокращать аппаратные затраты при решении задач на сложных вычислительных системах и обеспечивает принципиальную возможность решения задач с существенно переменной интенсивностью потоков данных в едином вычислительном контуре. При этом, как показано в [3], обеспечивается многократный выигрыш в скорости решения таких задач по сравнению с традиционными многопроцессорными вычислительными системами.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Каляев И.А., Левин И.И., Семерников Е.А., Шмойлов В.И.* Реконфигурируемые мультиконвейерные вычислительные структуры / Под общ. ред. И.А. Каляева. – Ростов-на-Дону: Изд-во ЮНЦ РАН, 2008. – 320 с.
2. *Каляев А.В., Левин И.И.* Модульно-наращиваемые многопроцессорные системы со структурно-процедурной организацией вычислений. – М.: ООО «Изд-во Янус-К», 2003. – 380 с.
3. *Сорокин Д.А., Левин И.И., Дордопуло А.И., Мельников А.К.* Решение задач с существенно-переменной интенсивностью потоков данных на реконфигурируемых вычислительных системах // Вестник компьютерных и информационных технологий. – М.: Машиностроение, 2012. – № 2. – С. 24.

Статью рекомендовал к опубликованию д.т.н., профессор Я.Е. Ромм.

Дордопуло Алексей Игоревич – Учреждение Российской академии наук «Южный научный центр РАН»; e-mail: scorpio@mvs.tsure.ru; 344006, г. Ростов-на-Дону, пр. Чехова, 41; тел.: 88634315491; отдел информационных технологий и процессов управления; к.т.н.; старший научный сотрудник.

Сорокин Дмитрий Анатольевич – Научно-исследовательский институт многопроцессорных вычислительных систем им. А.В. Каляева Южного федерального университета; e-mail: jotun@inbox.ru; 347928, г. Таганрог, ул. Чехова, 2; тел.: 88634315491; научный сотрудник.

Dordopulo Alexey Igorevich – Southern Scientific Centre of the Russian Academy of Sciences; e-mail: scorpio@mvs.tsure.ru; 41, Chekhov aven.; Rostov-on-Don, 344006, Russia; phone: +78634315491; cand. of eng. sc.; senior scientist.

Sorokin Dmitry Anatolievich – Scientific-Research Institute Multiprocessing Computing Systems after Kalyaev of South Federal University; e-mail: jotun@inbox.ru; 2, Chekhov street, Taganrog, 347928, Russia; phone: +78634315491; scientific associate.

УДК 004.272.22

Ю.И. Доронченко

ПРОЦЕДУРА МЕЖИТЕРАЦИОННОЙ РЕДУКЦИИ ЗАДЕРЖЕК ДЛЯ КОНВЕЙЕРНОЙ РЕАЛИЗАЦИИ РЕКУРСИВНЫХ ВЫЧИСЛЕНИЙ*

Предложен способ сокращения аппаратных затрат при конвейерной реализации рекурсивных вычислений в вычислительно трудоемких алгоритмах защиты информации. Впервые показана эффективность применения предложенного метода для реконфигурируемых вычислительных систем. Применение предложенной процедуры межитерационной редукции обеспечивает минимальные рекурсивные связи при распределении вычислений во времени, что позволяет сократить затраты на синхронизацию от 1,3 до 6 раз и повысить удельную производительность конвейерной реализации рекурсивных вычислений на реконфигурируемой вычислительной системе.

Конвейерные вычисления; защита информации; рекурсивные вычисления; редукция, синхронизация.

Y.I. Doronchenko

PROCEDURE OF INTER-ITERATION REDUCTION OF DELAYS FOR PIPELINE RECURSIVE CALCULATIONS

The paper covers the method of reduction of hardware resource needed for pipeline recursive calculations in computationally laborious algorithms of information security. For the first time the author has shown effectiveness of application of the suggested method in reconfigurable computer systems. The developed procedure of inter-iteration reduction provides minimal recursive constraints for time distributed calculations and reduction of synchronization costs from 1.3 up to 6 times. Besides, it helps to increase specific performance of pipeline recursive calculations for reconfigurable computer system.

Pipeline calculations; information security; recursive calculations; reduction; synchronization.

Вопросы безопасности больших вычислительных и управляющих систем требуют оперативной обработки больших массивов данных с помощью вычислительно-трудоемких алгоритмов, что обусловливает применение высокопроизводительных вычислительных комплексов, реализующих методы блокирования несанкционированного доступа к информации. Реализация подобных методов защиты зачастую требует существенного аппаратного ресурса, поэтому актуальны методы, позволяющие снизить требования к ресурсу а, следовательно, и к массогабаритным характеристикам и стоимости.

При построении конвейерных вычислительных структур [1] существенные затраты аппаратного ресурса требуют вопросы синхронизации информационных потоков. В этой связи предлагается метод, позволяющий для рекурсивных вычис-

* Исследования выполнены при финансовой поддержке Министерства образования и науки РФ, госконтракт № 14.527.12.0004 от 03 октября 2011 г.