

И. Г. Мороз-Подворчан

К вопросу о реальной эффективности параллельных вычислений в проектируемых системах реального времени

(Представлено академиком НАН Украины В. И. Скурихиным)

Ускорение вычислений в системах зависит от числа параллельно работающих процессоров. Рассматривается реальная оценка скоростных параметров в проектируемых системах реального времени при параллелизации вычислений.

К технической основе разносторонней информатизации относятся разнообразные средства вычислительной техники, в том числе обладающие системной сложностью. Здесь выделяются два типа вычислительных систем, характеризующиеся особенностями переработки информации во времени, — вычислительные системы свободного времени (ВС СВ) и вычислительные системы реального времени (ВС РВ), существенно различающиеся областями практического использования и, соответственно, подходами к проектированию. Если ВС СВ преимущественно ориентированы на обслуживание потока расчетных задач без особой дифференциации последних по времени их индивидуального решения на основании соображений эффективной переработки информации “в общем и целом”, то ВС РВ, напротив, больше направлены на индивидуальное обслуживание задач со строгим соблюдением времени решения, например при автоматизированном управлении объектами новой техники. И если процесс решения задач в ВС СВ во многих случаях можно рассматривать одноактно — от момента ввода исходных данных до момента получения окончательного результата, то процесс решения указанных задач управления в ВС РВ в своем большинстве представляет собой развернутую во времени процедуру оперирования получаемыми и выдаваемыми текущими данными в режиме функционирования автоматизируемого объекта — реализацию операторов управления “в нужное время и в нужном месте”. Контурно это показано на схеме (рис. 1), которая традиционно используется для иллюстрации качественного отличия ВС РВ от ВС СВ во временном аспекте.

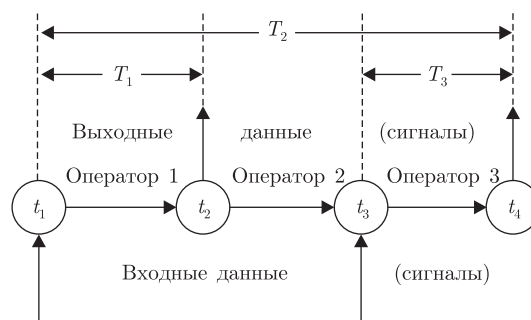


Рис. 1

В рамках данной работы определим процесс проектирования ВС РВ — специализированных вычислительных управляющих систем реального времени — как процесс придания проектируемой системе действительно требуемых свойств (насколько об этом можно судить на уровне имеющихся знаний и принятых положений, понимания условий данной ситуации и планируемых результатов), в первую очередь, в виде временных параметров переработки информации в проектируемой системе — реализации функций управления. Соответственно определим возможные получающиеся отклонения от требуемых значений параметров как допущенные ошибки в проектировании, которые в ряде случаев могут приводить к значительным негативным последствиям в управлении вплоть до катастрофических. Например, в областях ракетно- и авиастроения (особенно это касается прецизионной робототехники для высокоответственных применений).

Выделим класс специфических ошибок, происхождение которых объясняется неполнотой знаний о реалиях переработки информации в конкретных вычислительных схемах, в том числе при параллельной организации вычислений, в результате чего при оптимистической позиции проектировщиков оценки скоростных параметров переработки могут завышаться, что ощутимо искажает паспортно-производственные представления о проектируемой системе с понятными возможными последствиями, в частности, коммерческого характера. Такого рода ошибки можно назвать потенциально исправимыми при условии проведения дополнительных более глубоких “устанавливающих истину” исследований, в первую очередь имеющих вид точных доказательств. (В этой связи представляется методологически уместно обратить внимание на многочисленные факты использования стереотипного выражения “Дьявол находится в деталях”, которым некоторые научные работники, по сути, оправдывают отказ от необходимости проведения глубоких исследований там, где полученные результаты могут иметь решающее значение, возможно, приводя к выводам, оказывающимся прямо противоположными начальным.)

Приведем в качестве иллюстрирующего примера простейшую, но сохраняющую представительность, схему вычислений, реализуемую в блоке параллельно работающих однородных арифметических процессоров, и исследуем, что может обеспечить данная схема в вычислительном плане. Данный случай актуализируется вопросом о возможности/невозможности осуществления сверхлинейного ускорения переработки информации с ростом числа процессоров в многопроцессорной системе (машине) макроконвейерного типа [1–6]. Рассмотрим и оценим аналитически, что является необходимым условием для получения точных количественных оценок реально получающегося ускорения вычислений в блоке из p процессоров, в сравнении с тем же вычислением (решением той же задачи) лишь в одном из таких же процессоров, что принято выражать коэффициентом ускорения $K_y = T_1/T_p$, где T_1 — время решения данной задачи на одном процессоре, а T_p — время решения ее же в указанном блоке.

Допустим, речь идет о текущем вычислении суммы n однотипных чисел, плотным потоком поступающих в блок из однородных процессоров (рис. 2). Оперирование здесь заключается в вычислении в каждом из процессоров $m = n/p$ слагаемых промежуточных сумм, которые затем суммируются окончательно в одном из этих процессоров. Полное реальное время решения этой задачи в тактах τ составляет $T_p^* = \tau n/p + \tau \log_2 p$. (Для технического удобства параметрам придаются значения степеней 2. Так, $T_p^* = 32\tau + \tau + \tau$ при $n = 128$ и $p = 4$.)

Таким образом, реальная длительность процесса решения данной задачи в рамках данной схемы определяется как сумма двух компонентов, зависящих от p , причем с ростом p

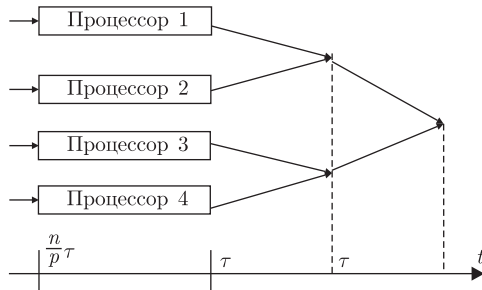


Рис. 2

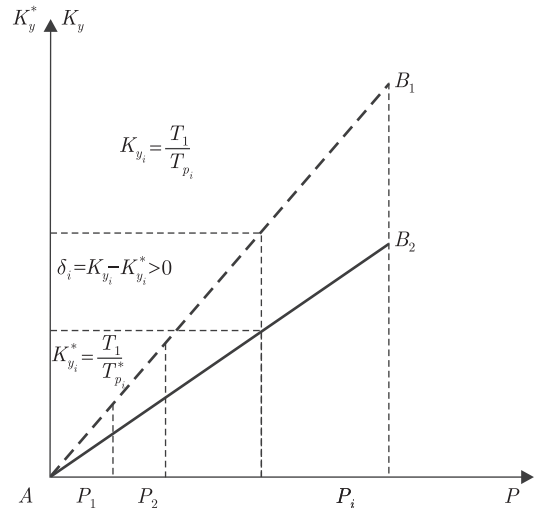


Рис. 3

первый компонент уменьшается линейно, а второй — увеличивается логарифмически. Отсюда следует, что, по меньшей мере применительно к задачам и схемам такого рода, выражение ускорения процесса вычислений в блоке из множества параллельных процессоров в сравнении с процессом решения той же задачи в одном процессоре приобретает следующий вид:

$$K_y^* = \frac{T_1}{T_p^*} = \frac{\tau n}{\tau \frac{n}{p} + \tau \log_2 p} < p,$$

в то время как

$$K_y = \frac{T_1}{T_p} = \frac{\tau n}{\tau \frac{n}{p}} = p,$$

где K_y^* имеет смысл реального ускорения, а K_y — смысл гипотетического линейного ускорения.

Этот факт стилизованно изображен на рис. 3, где $K_{y_i}^*$ обозначается сплошной линией B_2 , а K_{y_i} — штриховой B_1 . И хотя $K_{y_{i+1}}^* - K_{y_i}^* < K_{y_i} - K_{y_{i-1}}$, все равно

$$K_{y_{i+1}} - K_{y_{i+1}}^* > K_{y_i} - K_{y_i}^*.$$

Иначе говоря, здесь с увеличением числа процессоров скорость переработки задачной информации растет, а ускорение в рассматриваемом смысле падает.

Правомочно полагать, что решение более сложных в вычислительном отношении задач не приведет к получению более высокого ускорения. Если также полагать, что данная задача не сложнее любой другой, то приведенное можно считать доказательством невозможности достижения линейного ускорения в любой вычислительной системе с параллельной организацией вычислений, в том числе в макроконвейерной ЭВМ с увеличением числа параллельно работающих процессов в арифметическом блоке, тем более ускорения “более чем линейного” [4].

Разумеется, изложенное никоим образом не следует понимать как тенденциозную критику безусловно продуктивного метода параллельной организации вычислений в системах, тем более в многоканальных и многофункциональных. Мотивировка данной работы — формирование реалистических представлений о точных значениях действующих скоростных параметров, что особенно важно при проектировании специализированных ВС РВ.

1. Михалевиц В. С., Капитонова Ю. В., Летичевский А. А. Организация вычислений в многопроцессорных вычислительных системах // Кибернетика. – 1984. – № 3. – С. 1–10.
2. Михалевиц В. С., Капитонова Ю. В., Летичевский А. А. О методах организации макроконвейерных вычислений // Там же. – 1986. – № 3. – С. 3–10.
3. Молчанов И. Н. Введение в алгоритмы параллельных вычислений. – Киев: Наук. думка, 1990. – 128 с.
4. Капитонова Ю. В., Летичевский А. А. Математическая теория проектирования вычислительных систем. – Москва: Наука, 1988. – 296 с.
5. Мороз-Подворчан И. Г. К вопросу о верхней границе ускорения в макроконвейерной ЭВМ // Мат. машины и системы. – 2008. – № 2. – С. 70–72.
6. Мороз-Подворчан И. Г. О полноте описания скоростных возможностей макроконвейерной ЭВМ // Там же. – 2008. – № 4. – С. 57–60.

*Институт кибернетики им. В. М. Глушкова
НАН Украины, Киев*

Поступило в редакцию 09.10.2012

І. Г. Мороз-Подворчан

До питання про реальну ефективність паралельних обчислень у проєктованих системах реального часу

Прискорення обчислень у системах залежить від кількості процесорів, що працюють паралельно. Розглядається реальна оцінка швидкісних параметрів у проєктованих системах реального часу при паралелізації обчислень.

I. G. Moroz-Podvorchan

To the question concerning the real efficiency of parallel calculations in designed real-time computer systems

The acceleration of calculations in computer systems depends on the number of parallel processors. The real estimation of speed-parameters in designed computer systems under the calculation parallelism is considered.