

ОБОБЩЕННАЯ МОДЕЛЬ СИНТЕЗА ПАРАЛЛЕЛЬНЫХ ВРЕМЯПАРАМЕТРИЗОВАННЫХ ПРОЦЕССОВ ДЛЯ КЛАСТЕРНЫХ ВС

Г.А. ПОЛЯКОВ¹
Е.Г. ТОЛСТОЛУЖСКАЯ¹
Ю.А. АРТИУХ²

¹Белгородский государственный национальный исследовательский университет

²Харьковский национальный университет им. В.Н. Каразина

e-mail:
tda_ua@pochtamt.ru
ulia_artiuh@mail.ru

В статье приведены основные этапы синтеза параллельных времяпараметризованных процессов для кластерных вычислительных систем (ВС). Разработанная модель позволяет повысить эффективность вычислительного процесса, реализуемого на кластерных ВС.

Ключевые слова: времяпараметризованный параллельный процесс, кластерные вычислительные системы (ВС), методы параллельной обработки данных, декомпозиция, модель управления.

Введение

В современном мире для параллельных вычислений наиболее часто используют кластерные ВС, т.к. они характеризуются относительно невысокой стоимостью, возможностью реинжиниринга и масштабирования оборудования, простотой развертывания и пр. [1, 2, 8].

Общепринятый подход к разработке параллельных алгоритмов задач состоит из следующих основных этапов [1, 2].

1. Декомпозиция задач на подзадачи.
2. Распределение подзадач по процессам.
3. Планирование коммуникаций.
4. Укрупнение.
5. Распределение ресурсов.

На данном этапе развития подхода к синтезу параллельных алгоритмов все вышеперечисленные этапы являются творческими и выполняются разработчиком. Таким образом, эффективность получаемых параллельных алгоритмов напрямую зависит от степени мастерства и профессиональной подготовки разработчика. Более того, при не автоматизированном создании параллельных алгоритмов используется только метод совмещения независимых операций. В литературе [1, 2] отмечается, что человек не способен разработать эффективные параллельные алгоритмы для более чем 6 процессов. При этом количественные оценки показателей эффективности параллельных программ (которые могут, как удовлетворять, так и не удовлетворять выдвигаемым требованиям) разработчик получает только после выполнения всех этапов разработки параллельного алгоритма и запуска параллельной программы на кластерной системе. Выходом из сложившейся ситуации является использование при синтезе параллельных программ для кластерных систем времяпараметризованных моделей параллельных процессов, содержащих в явном виде конкретные оценки показателей эффективности [3,4].

Постановка задачи.

Исходная информация:

- Си – программа решаемой задачи;
- класс параллельной ЭВМ – кластерные вычислительные системы;
- известные методы параллельной обработки – совмещение независимых операций, мультипараллельная смесь алгоритмов, кодово-матричная обработка, конвейерная обработка, декомпозиционная обработка [3-6];
- система требований и ограничений (время решения, ресурс средств).



Требуется разработать обобщенную модель синтеза параллельных времяпараметризованных процессов для кластерных ВС, удовлетворяющих заданным требованиям к показателям эффективности.

Этапы решения задачи

Обобщенная модель синтеза параллельных времяпараметризованных процессов на кластерных ВС представлена на рис. 1.

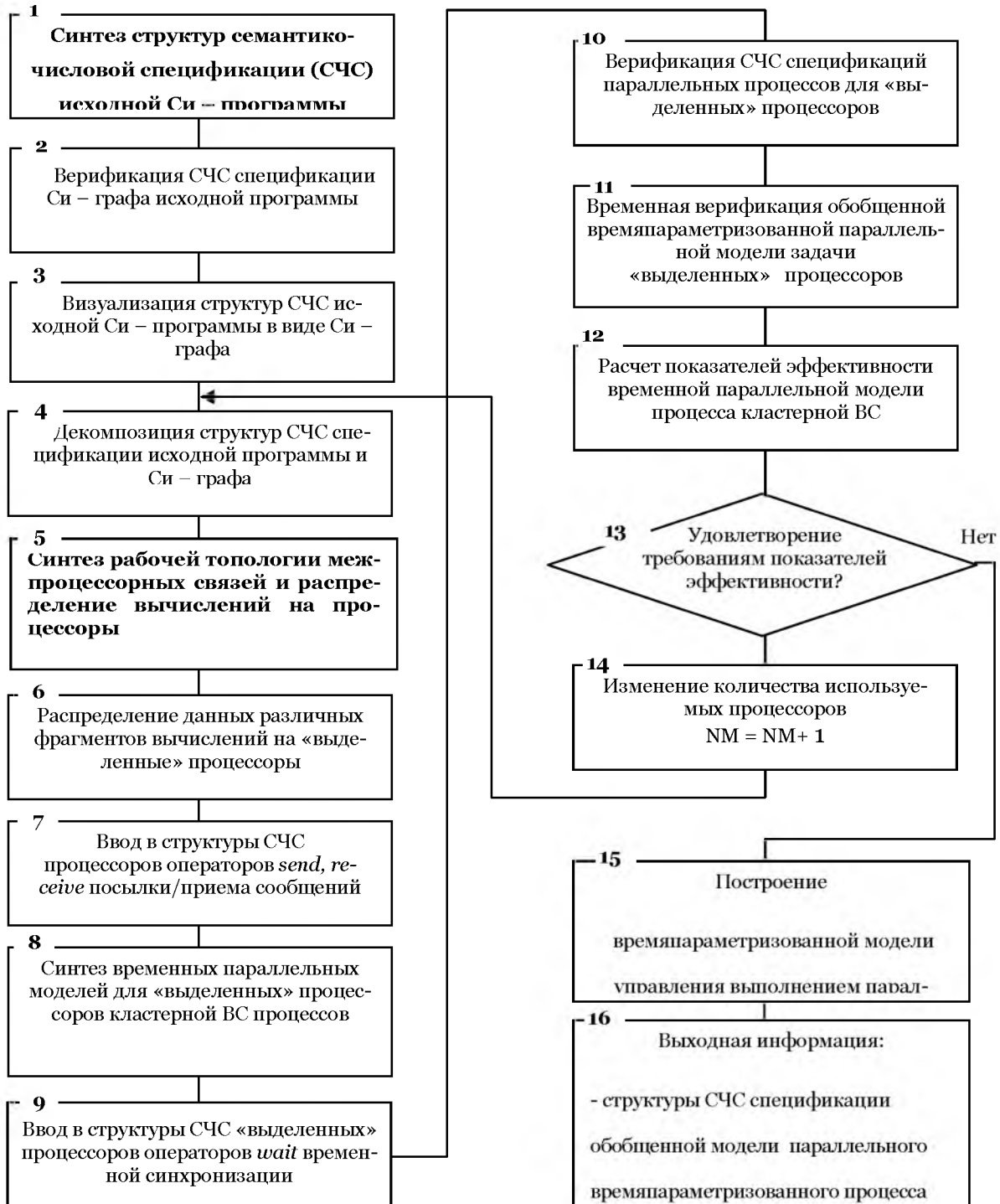


Рис. 1. Обобщенная модель синтеза параллельных времяпараметризованных процессов на кластерных ВС

Содержанием этапа 1 (блоки 1-3, рис. 1) является решение следующих задач:

- синтез для исходной Си – программы структур семантико-числовой спецификации (СЧС) [4,5];

- синтез, исходя из СЧС спецификации Си-программы, соответствующего графического представления в виде Си-графа [4];
- проверка корректности результатов синтеза СЧС спецификации (структур BF и CF) исходной Си – программы и соответствующего Си-графа [4].

Этап 2 (блок 4, рис. 1). На данном этапе выполняется разделение структур СЧС спецификации исходной программы и Си-графа задачи на временные фрагменты для целей их последующей параллельной реализации.

Этап 3 (блок 5, рис. 1) – определение в кластере подмножества процессоров и их коммуникационных связей, выделяемых для реализации временных фрагментов параллельного времяпараметризованного процесса [4].

Этап 4 (блок 6, рис. 1). На данном этапе происходит распределение данных различных временных фрагментов вычислений на «выделенные» процессоры.

Этап 5 (блок 7, рис. 1). Содержанием этапа является определение количества и состава межфрагментных обменов данными и ввод в структуры СЧС процессоров операторов $send, receive$ посылки/приема сообщений.

Содержанием этапа 6 (блок 8, рис. 1) является синтез временных параллельных моделей фрагментов для «выделенных» процессоров с учетом введенных операторов $send, receive$ посылки/приема сообщений [7].

Этап 7 (блок 9, рис. 1). Задачей этапа является ввод операторов $wait$ временной синхронизации в модели времяпараметризованных фрагментов «выделенных» процессоров на основе учета реальных связей между операторами и длительностей выполнения операторов каждого фрагмента.

Содержанием этапа 8 (блоки 10, 11, рис. 1) является проверка корректности синтезированных СЧС спецификаций моделей параллельных процессов для «выделенных» процессоров и временная верификация обобщенной времяпараметризованной параллельной модели задачи [4].

На этапе 9 (блоки 12 – 14, рис. 1) обеспечивается оценка показателей эффективности синтезированной обобщенной времяпараметризованной параллельной модели процесса [4].

Математическое ожидание времени реализации множества P операторов алгоритма

$$T(P) = \sum_{\xi=1}^w p_{\xi} T_{\xi}, \quad (c),$$

где w – число ветвей в алгоритме, p_{ξ} – вероятность реализации ξ -й ветви.

Время реализации ξ -й ветви параллельного алгоритма

$$T_{\xi} = \max_{P_j \in P(\xi)} (t_j^H + t_j), \quad (c),$$

где $P(\xi)$ – множество операторов ξ -й ветви, t_j^H и t_j – момент начала и относительная временная глубина оператора $P_j \in P(\xi)$.

Дисперсия времени реализации T_k ветвей различной длительности параллельного алгоритма

$$D = \sum_{\xi} (T_{\xi} - T(P))^2, \quad (c2), \quad \xi = 1, \dots, w.$$

Среднее быстродействие при реализации параллельных алгоритмов

$$B(P) = \frac{\sum_{\xi=1}^w p_{\xi} \sum_{t^H=1}^{T_{\xi}} H^{\xi}(t^H)}{\sum_{\xi=1}^w p_{\xi} T_{\xi}} \quad (\text{опер/с}),$$

где $H^{\xi}(t^H)$ – количество операторов ξ -й ветви алгоритма, реализация которых начинается в момент времени t^H .

Снижение временных затрат на выполнение алгоритмов за счет перехода к их параллельной реализации



$$DT = \frac{T_{noc}(P)}{T_{пар}(P)} \text{ (раз)},$$

где $T_{noc}(P)$ и $T_{пар}(P)$ – среднее время соответственно последовательной и параллельной реализации алгоритма.

Среднее значение полного показателя загрузки всех компонентов, входящих в состав параллельного устройства

$$S(P) = \sum_{\eta=1}^v S(P_{\eta}) = \frac{1}{T} \sum_{\eta=1}^v \frac{1}{n_{\eta}} \sum_{\xi=1}^w (t_{\eta}^0 \cdot K_{\eta\xi} \cdot p_{\xi}), \text{ (раз)}.$$

Среднее значение показателя загрузки компонентов каждого типа $\Theta_{\eta} \in \Theta$

$$S(P_{\eta}) = \frac{\sum_{\xi=1}^w (t_{\eta}^0 \cdot K_{\eta\xi} \cdot p_{\xi})}{n_{\eta} \cdot T}, \text{ (раз)},$$

где $P_{\eta} \subseteq P$ – подмножество операторов $P_j \in P$, имеющих тип Θ_{η} ,

$$K_{\eta\xi} = \sum_{\delta=1}^{n_{\eta}} K_{\eta\xi\delta}.$$

Важное прикладное значение имеет задача определения числа NM процессоров, применение которых для параллельной реализации алгоритма обеспечивает достаточно большое снижение временных затрат (либо заданное снижение временных затрат) за счет параллельного выполнения алгоритма, с одной стороны, и возможно более высокое значение коэффициента использования оборудования (либо достижение заданного значения этого коэффициента), с другой. В таких случаях можно в качестве производного показателя использовать аддитивный показатель эффективности распараллеливания

$$R(P) = K_T \frac{DT(NM)}{NM} + K_S S(NM),$$

где K_T и K_S являются весовыми коэффициентами, определяющими «пользовательскую» важность учета в эффективности распараллеливания величины сокращения времени реализации алгоритма ($K_T \leq 1$) и степени загрузки оборудования параллельным алгоритмом ($K_S \leq 1$).

На этапе 10 (блок 15, рис. 1) обеспечивается организация управления взаимодействием «выделенного» ресурса кластера (процессоров и линий передачи данных) при выполнении различных ветвей обобщенной времяпараметризованной параллельной модели процесса кластерной ВС [9].

Выводы.

1. Широкое применение кластерных систем в различных прикладных областях предъявляет высокие требования к эффективности их использования. Известные системы параллельного программирования, основанные, в значительной степени, на субъективном опыте и творчестве специалистов, не могут обеспечить качественное решение этой проблемы.

2. Одним из путей повышения эффективности кластерных ВС является использование времяпараметризованных моделей параллельных процессов.

3. Применение моделей данного класса в сочетании с формализацией всех основных этапов их синтеза обеспечивает потенциальные возможности повышения эффективности кластерных ВС за счет совмещения учета специфики решаемой задачи, архитектурных особенностей ВС и временных характеристик процессов параллельной обработки данных.

Список литературы

1. Воеводин, В.В., Воеводин Вл. В. Параллельные вычисления. – СПб.: БХВ-Петербург, 2002. – 608 с.
 2. Немнюгин, С.А. Параллельное программирование для многопроцессорных вычислительных систем. / С.А. Немнюгин, О.Л. Стесик. – СПб.: БХВ-Петербург, 2002. – 400 с.

3. Поляков, Г.А. Технология проектирования времяпараметризованных мультипараллельных программ как стратегия развития систем параллельного проектирования. / Г.О. Поляков, Е.Г. Толстолужская. // *Радіоелектронні і комп'ютерні системи – Х.*: Національний аерокосмічний університет ім. М.Є. Жуковського «Харківський авіаційний інститут», 2009. – Вип. 6(40). – С. 166-171.

4. Поляков, Г.А. Синтез и анализ параллельных процессов в адаптивных времяпараметризованных вычислительных системах / Г.А. Поляков, С.И. Шматов, Е.Г. Толстолужская, Д.А. Толстолужский: монография. – Х. : ХНУ имени В.Н. Каразина, 2012. –С. 434 – 575.

5. Поляков, Г.А. Синтез вычислительных подсистем параллельных неперестраиваемых спецпроцессоров с использованием аппарата структур семантико-числовой спецификации/ Г.А. Поляков, В.В. Толстолужская, В.В. Лысых // *Научные ведомости БелГУ. Сер. История. Политология. Экономика. Информатика.* –2012. – № 13 (132). –Вып. 23/1. –С. 142-150

6. Поляков, Г.А. Методы цифровой времяпараметризованной мультипараллельной обработки данных/ Г.А. Поляков, В.В. Толстолужская, Е.Г. Толстолужская. // *Прикладная математика, управление и информатика: сборник трудов междунар. молодеж. конф., Белгород, 3-5 октября 2012 г.: в 2 т.* – Белгород : ИД «Белгород», 2012. – Т. 2. – С. 511-519.

7. Поляков, Г.А. Разработка фрагментированной временной параллельной модели алгоритма Гаусса на основе формальных полиномов и структур семантико-числовой спецификации/ Г.А. Поляков, К.В. Лысых. // *Прикладная математика, управление и информатика: сборник трудов междунар. молодеж. конф., Белгород, 3-5 октября 2012г.: в 2 т.* – Белгород : ИД «Белгород», 2012. – Т. 2. – С. 469-472.

8. Гергель, В.П. Технологии построения и использования кластерных систем [Электронный ресурс] / Интернет университет информационных технологий ИНТУИТ.Р, 2009. – Режим доступа: <http://www.intuit.ru/department/supercomputing/tbucs/1/> свободный. – Загл. с экрана.

9. Толстолужская, Е.Г., Артюх, Ю.А. Методика синтеза временной модели управления параллельным вычислительным процессом. – *Системы обработки информации.* Харьков, 2011. Вип. № 8, 2011. – с. 248–253.

GENERALIZE MODEL OF PARALLEL TIMEPARAMETERIZED PROCESSES SYNTHESIS FOR CLUSTER COMPUTER SYSTEMS

G.A. POLYAKOV¹

E.G. TOLSTOLUJSKAIJA¹

I.U.A. ARTIUKH²

¹⁾ *Belgorod National Research University*

²⁾ *V.N. Karazin Kharkiv National University*

e-mail:

tda_ua@pochtamt.ru

ulia_artiuh@mail.ru

The paper presents the main stages of the parallel processes synthesis for cluster time-parameter computational systems (CS). The developed model can improve the efficiency of the computational process, implemented on clustered CS.

Keywords: time-parametr parallel algorithm, cluster computing (CC), the methods of parallel processing, defrag, management model.