

## **ОТЗЫВ**

**официального оппонента на диссертационную работу**

**Колганова Александра Сергеевича**

**«Автоматизация распараллеливания Фортран-программ для гетерогенных кластеров», представленную к защите на соискание ученой степени кандидата физико-математических наук по специальности 05.13.11 – «математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей»**

### **Актуальность**

Сопровождение унаследованных последовательных программ, их модификация для эффективного выполнения на гетерогенных высокопроизводительных кластерах является сложным делом.

Полностью автоматическое распараллеливание на кластер для многих программ практически невозможно, потому что при переходе от последовательной программы к параллельной в большинстве случаев требуется ее серьезное преобразование или даже изменение алгоритма. С целью автоматизации распараллеливания программ создаются высокоуровневые языки параллельного программирования, а также системы автоматизации распараллеливания программ.

Автором работы предпринята попытка доработать существующую систему автоматизированного распараллеливания программ, написанных на языке Фортран, для повышения эффективности программ, выполняемых на сложных гетерогенных кластерах, в узлах которых могут быть как многоядерные процессоры общего назначения, так и графические ускорители. Это, безусловно, является актуальной задачей.

### **Структура работы**

Диссертационная работа состоит из введения, шести глав, заключения, списка литературы из 52 наименований и двух приложений. Объем диссертации составляет 135 страниц.

**Во введении** обосновывается актуальность темы исследования, формулируются цель работы и поставленные для ее достижения задачи, определяется научная новизна и практическая значимость.

**В первой главе** описываются проблемы отображения на кластер и причины возникновения автоматизированных и автоматических систем распараллеливания для кластеров.

**Во второй главе** приводится схема работы и структура разработанной в рамках диссертационного исследования системы SAPFOR 2. Новая система была призвана устранить известные недостатки предыдущей системы автоматизации распараллеливания САПФОР и расширить класс программ, использующих структурные сетки, к которым можно было бы ее применить.

**В третьей главе** описываются основные объекты анализа программы – циклы и функции. Рассматриваются алгоритмы заполнения дерева циклов и графа вызовов функций. Рассматриваются проблемы межпроцедурного анализа, которые возникают в процессе распараллеливания циклов и программы в целом. Дается описание решения данных проблем в виде преобразования «подстановка функции» и «дублирование цепочек вызовов функций».

**В четвертой главе** вводится описание механизма областей распараллеливания. Под областью распараллеливания понимается совокупность исполняемых операторов исходного кода программы, описываемых с помощью фрагментов. В свою очередь, под фрагментом понимается множество исполняемых операторов в рамках одной области вложенности (функция, процедура, цикл, условный оператор и т.д.) с одним входом и несколькими выходами. Вводится понятие конфликта областей и описываются алгоритмы разрешения данных конфликтов.

**Глава пять** посвящена алгоритмам построения распределения данных и вычислений, отображению на гетерогенный кластер. В разделе 5.1 дается описание способов отображения на кластер и их характеристика. В разделе 5.2 описывается алгоритм связи массивов в программе между собой. В разделе 5.3 описывается алгоритм построения распределения вычислений и доступа к удаленным данным. В разделе 5.4 описывается алгоритм расстановки вычислительных регионов – фрагментов программы, которые целесообразно выполнять на ускорителях. В разделе 5.5 описывается алгоритм выбора эффективных схем распараллеливания программы посредством грубого подсчета характеристик ее параллельного выполнения.

**В шестой главе** приводятся основные результаты исследования реализованных алгоритмов автоматизированного распараллеливания в системе SAPFOR 2 на тестовых и прикладных программах.

**В заключении** формулируются основные результаты диссертационной работы и возможные направления дальнейших исследований.

### **Научная новизна и практическая значимость**

В работе получены следующие новые научные результаты:

- Разработаны алгоритмы построения распределения данных и вычислений.
- Разработан и реализован механизм областей распараллеливания.
- Разработана и реализована новая система (SAPFOR 2) автоматизации отображения Фортран-программ на гетерогенный кластер с поддержкой инкрементального распараллеливания.

Проведенные исследования применимости системы SAPFOR 2 для тестов и реальных приложений показали эффективность предложенного подхода, позволяющего упростить разработку параллельных программ для гетерогенных кластеров, а также сопровождение существующих последовательных программ.

### **Достоверность и обоснованность научных положений и выводов работы**

Результаты исследования опубликованы в 8 печатных работах (в том числе в 4 журналах из перечня ВАК, среди них 2 индексируются системами Scopus и/или Web of Science), докладывались на российских и международных конференциях.

### **Замечания**

К работе имеются следующие замечания.

1. В работе ставятся жесткие ограничения на класс программ, которые могут быть распараллелены системой SAPFOR 2 – нет поддержки указателей, производных типов данных, интерфейсных функций, функций с переменным числом параметров. Таким образом, можно считать, что разработанная система SAPFOR 2 прежде всего рассчитана на распараллеливание «старых» Фортран-программ. Новые программы, которые используют новые возможности языка Фортран, для своего распараллеливания потребуют дополнительных преобразований.
2. Динамический анализ используется очень ограниченно. Он мог бы упростить работу программиста по спецификации тех свойств

последовательной программы, которые не может определить статический анализ.

3. Распараллеливание работы с разреженными матрицами и неструктурными сетками не поддерживается.

4. В тестовой программе в листинге 6.1 есть ошибка. Массив A состоит из 10000 элементов. Но в программе есть цикл, который выводит за границу массива:

```
DO I=1, 10000
```

```
  A(I+1) = C(I)
```

```
ENDDO
```

Аналогично - в программе идут обращения к элементам массива B(I-25), которые также приводят к ошибкам. При этом SAPFOR 2 строит для данной программы некорректную параллельную программу.

5. Не реализовано разбиение массивов и циклов на блоки.

6. Вместо деклараций вида "показали высокую эффективность", "позволили значительно упростить и существенно ускорить разработку" желательны количественные экспериментальные оценки.

7. В работе отсутствует сравнение показателей эффективности созданной автором системы SAPFOR 2 и уже существующих систем, использующих, например, стандарт OpenACC.

8. Отсутствует обзор литературы по теме диссертации. Глава 1 обзором не является. Библиография нуждается в пополнении и актуализации. Две ссылки на англоязычные статьи 2012 года ([6] и [8]) - это явно мало.

9. Ссылки на электронные ресурсы неактуальны. Единственная относительно свежая система ParaWise датируется 2016 годом.

10. Неудачна формулировка результатов работы. Новизна сама по себе не является достоинством. Необходимо указать, чем новое лучше старого.

11. Фигурирующее в результате 2 инкрементальное распараллеливание не является чем-то новым. См., например, университетский курс А.О. Лациса "Параллельная обработка данных", с. 185, а также статьи Lei Pan, Wenhui Zhang, Arthur Asuncion, Ming Kin Lai, Michael B. Dillencourt, Lubomir F. Bic. Incremental Parallelization Using Navigational Programming: A Case Study. - Proceedings of the 2005 International Conference on Parallel Processing (ICPP'05), 10 pp. и Wenhui Zhang, Lei Pan, Qinghong Shang, Lubomir F. Bic, Michael B.

Dillencourt. Incremental Parallelization with Migration. - 2012 10th IEEE International Symposium on Parallel and Distributed Processing with Applications, pp. 223-230.

12. В работе имеются расхождения в использовании терминов, например, на с. 35 говорится о «витке» цикла, а на с. 106 – об «итерации» цикла.
13. Имеются грамматические ошибки в тексте диссертации (см., например, с. 14, 15, 29) и автореферата (с. 7, 24).

Приведенные замечания не влияют на общую положительную оценку работы.

### Заключение

Диссертация Колганова Александра Сергеевича «Автоматизация распараллеливания Фортран-программ для гетерогенных кластеров» является завершённой работой, в которой автору удалось решить поставленные перед диссертационным исследованием задачи. Основные результаты диссертации были полностью и своевременно опубликованы и докладывались на российских и международных конференциях. Автореферат диссертации верно отражает ее основное содержание.

Диссертация отвечает требованиям Положения ВАК РФ о порядке присуждения ученых степеней, а ее автор, Колганов Александр Сергеевич, заслуживает присуждения ему ученой степени кандидата физико-математических наук по специальности 05.13.11 – «математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Доктор физико-математических наук, заведующий сектором Федерального государственного учреждения «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук»



В. А. Галатенко

Подпись руки В.А. Галатенко заверяю  
Начальник отдела кадров Юлия Толстикова

«20» августа 2020 г.

