

ГЕТЕРОГЕННЫЙ ВЫЧИСЛИТЕЛЬНЫЙ КОМПЛЕКС ДЛЯ РЕШЕНИЯ ИНЖЕНЕРНЫХ И НАУЧНЫХ ЗАДАЧ

Попков А.Ю., Соченков И.В., Хачумов В.М.

ФГБУН Институт системного анализа Российской академии наук (ИСА РАН)

Аннотация. Представлены основные характеристики экспериментального программно-аппаратного комплекса. Комплекс предназначен для решения задач моделирования систем управления, параллельной обработки видеоданных и данных широкого назначения.

Введение

Высокопроизводительный вычислительный комплекс создан и развертывается в ИСА РАН в рамках Программы № 18 фундаментальных исследований Президиума РАН «Алгоритмы и математическое обеспечение для вычислительных систем сверхвысокой производительности». Назначение – поддержка выполнения научной работы и инженерных расчетов.

1. Архитектура вычислительного комплекса ИСА РАН

Комплекс представляет собой гетерогенную вычислительную систему с распределенной памятью. Общая архитектура системы представлена на рисунке 1.

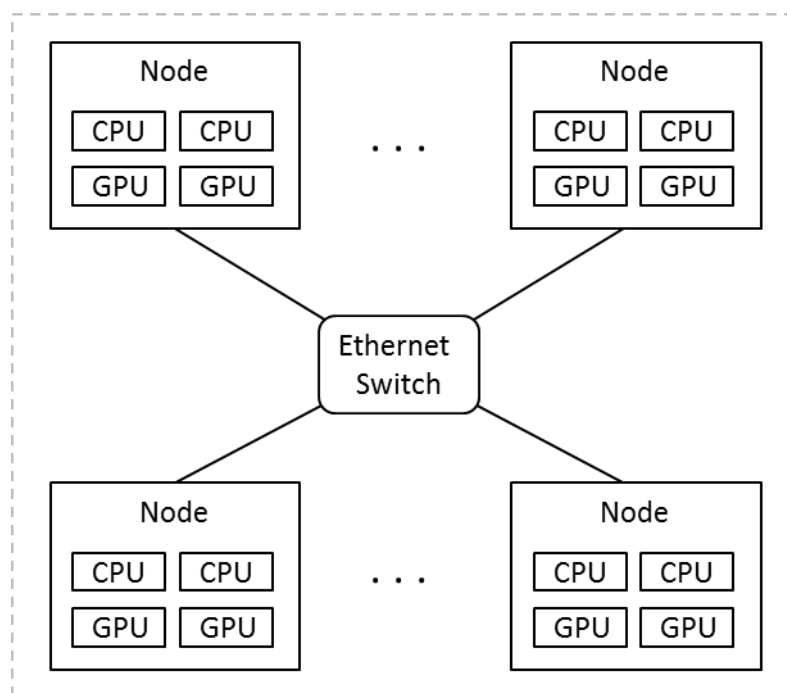


Рисунок 1 – Архитектура вычислительной системы

Вычислительные узлы имеют одинаковые аппаратные характеристики, что предполагает легкое горизонтальное масштабирование системы с целью увеличения общей производительности при решении требуемых задач. В настоящее время собрана и функционирует первая очередь программно-аппаратного гетерогенного комплекса (один вычислительный узел). В таблице 1 представлены основные характеристики узла системы.

Таблица 1 – Характеристики вычислительного узла системы

Наименование	Значение
Конструктив узла	1U RackMount
Тип процессора	Intel®Xeon® X5650 2.66 ГГц
Тактовая частота внешней шины	1333 МГц
Количество процессоров	2
Тип коммуникационной сети	Gigabit Ethernet
Оперативная память	32 Гб ECC
Блок питания	2 x 1400 Вт
Дисковая память	2 x 300 Гб
Система охлаждения	Воздушная

Каждый вычислительный узел оснащен двумя ускорителями NVidia Tesla M2050, характеристики которого представлены в таблице 2.

В настоящее время введена в строй первая очередь (один кластерный узел) системы.

Таблица 2 – Характеристики процессора NVidia Tesla M2050.

Наименование	Значение
Количество ядер CUDA	448
Объем памяти	3 Гб GDDR5
Пиковая производительность операций с плавающей запятой (двойная точность)	515 GFLOP
Пиковая производительность операций с плавающей запятой (одинарная точность)	1.03 TFLOP
Макс. потребление энергии	225 Вт
Интерфейс	PCIe x16 Gen2

2. Математическое обеспечение параллельной обработки видеоданных и данных широкого назначения

Вычислительная система работает под управлением 64-разрядной операционной системы Debian GNU/Linux 7.3. Для выполнения прикладных работ, проводимых сотрудниками ИСА РАН, ИПС им. А.К.Айламазяна РАН, РГААТА, комплекс оснащен прикладным программным обеспечением. В него входит несколько прикладных

пользовательских программ (таблица 3). В том числе установлена программная платформа для автоматизации проектирования интеллектуальных систем (разработана в ИПС им. А.К.Айламазяна РАН), которая позволяет эффективно решать как вычислительные задачи, допускающие внутренний (поточковый) параллелизм в рамках одного процесса, так и задачи, требующие распределения по данным между несколькими серверами КВУ.

Таблица 3 –Прикладное обеспечение (системы высокоуровневого программирования)

Наименование	Версия
CUDA	5.5
Atlas	3.10.1
Gdal	1.9.2
Libpng	1.5.12
OpenCV	2.4.9
Boost	1.55
Lapack	3.5.0
MPICH	1.4.1
Libxml2	2.8
GCC	4.7
MATLAB	Distributed Computing Server 8.0

Эффективный параллелизм между потоками выполнения на одном узле достигается за счет использования высокопроизводительных многоядерных процессоров. Возможные области применения системы: высокопроизводительная обработка потоков космических снимков, задачи диагностики и прогнозирования состояния сложных технических систем.

3. Тестирование работы комплекса на примере решения прикладной задачи

В качестве тестирующей программы был использован метод спектрографической «закраски», основывающийся на анализе текстур изображений [1]. Для тестирования использовалась выборка из одной тысячи изображений размером 512x768 пикселей. Размер сканирующего окна – 5x5, 9x9 и 15x15 пикселей соответственно. Количество вычислительных потоков варьировалось от 1 до 8. Результаты тестирования (исследования выполнены к.т.н. Фраленко В.П.) представлены на рисунке 2. Тестирование в эксперименте со сканирующими окнами позволило выбрать оптимальное число потоков. Эксперимент показал, что чем больше размер сканирующего окна, тем лучше масштабируется алгоритм.

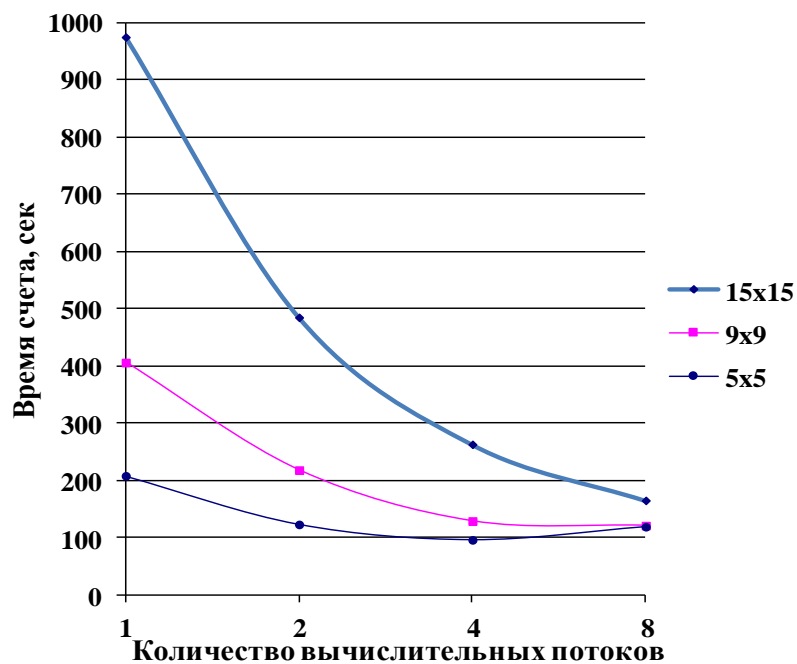


Рисунок 2 – Масштабируемость алгоритма для процессоров общего назначения

Проведенное исследование показало высокую эффективность работы классификатора на установке. Метод отличается низким количеством ложных срабатываний и может быть рекомендован для повсеместного использования.

Комплекс оказался удобным для моделирования процессов полета беспилотных летательных аппаратов (БПЛА) и управления видеокамерой в системе MATLAB [2]. На рисунке 3 представляет график отклонения луча видеокамеры, установленной на БПЛА от цели по оси абсцисс во времени (моделирование выполнено к.т.н. Абрамовым Н.С. и к.ф.-м.н. Хачумовым М.В.).

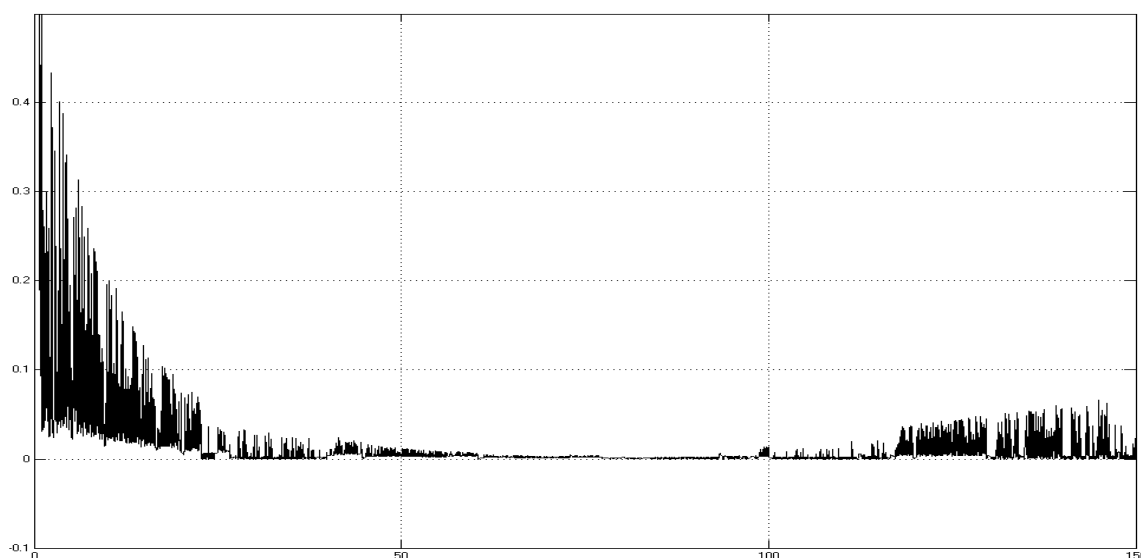


Рисунок 2 – График отклонения луча видеокамеры от цели по оси абсцисс во времени

Видно, что чем дальше БПЛА от объекта, тем больше ошибка наблюдения. При этом на всем интервале величина СКВО составила 0.056 м (первые две секунды не учитываются).

Цикл актуальных исследований на установке выполнен аспирантом РГААТА Лебедевым А.С. Целью исследований является разработка и исследование методов и инструментальных программных средств обработки интенсивных потоков данных на высокопроизводительных вычислительных системах на основе универсальных многоядерных и графических процессоров. В настоящее время с применением гетерогенного вычислительного комплекса ИСА РАН решаются следующие актуальные задачи: 1) разработка и отладка методов автоматического распараллеливания программ для современных вычислительных микроархитектур – универсальных многоядерных и графических процессоров; 2) исследование возможности расширения существующих подходов к распараллеливанию программ средствами распараллеливания для устройств с динамическим параллелизмом; 3) исследование и разработка методов динамического распараллеливания программ (just-in-time), направленных на расширение границ применимости существующих подходов к распараллеливанию.

Работа по построению вычислительного комплекса выполнена в рамках Программы фундаментальных исследований Президиума РАН №18 «Алгоритмы и математическое обеспечение для вычислительных систем сверхвысокой производительности». Тема «Исследование и разработка математического обеспечения и инструментальных программных средств параллельной обработки видеоданных и данных широкого назначения для интеллектуального управления видеокамерами и обеспечения на этой основе моделирования проводки беспилотных летательных аппаратов с применением параллельной среды программирования» (научн. рук.акад. Емельянов С.В., отв. исп. д.т.н. Хачумов В.М.). Комплекс используется в настоящее время для выполнения проектов РФФИ (13-07-00025 А, 14-07-31020 мол_а, 14-37-50316 мол_нр) и проекта ОНИТ 1 РАН 2.10 Развитие методов интеллектуального анализа данных и управления робототехническими системами с применением бесконтактных человеко-машинных интерфейсов

Литература

1. Фраленко В.П. Спектрографическая «закраска» полноцветных и мультиспектральных изображений (ПС "Emarking"). Свидетельство о государственной регистрации программы для ЭВМ № 2014611689 (дата подачи заявки: 11.12.2013, дата регистрации: 07.02.2014).
2. Абрамов Н.С., Хачумов М.В. Моделирование проводки по маршруту беспилотного летательного аппарата как задачи преследования цели. — Авиакосмическое приборостроение, № 9, 2013, с. 9–22.