

Министерство образования и науки Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра дискретной математики и
информационных технологий

**Исследование эффективности решения вычислительно сложных задач в
параллельном режиме в зависимости от аппаратных особенностей
используемой платформы**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

Студента 2 курса 271 группы
направления 09.04.01 «Информатика и вычислительная техника»
факультета компьютерных наук и информационных технологий
Каценеленбогена Артёма Александровича

Научный руководитель
к. ф.-м.н., доцент

дата, подпись

А.Д. Панферов

Заведующий кафедрой
к. ф.-м.н., доцент

дата, подпись

Л.Б. Тяпаев

Саратов 2018 год

ВВЕДЕНИЕ

Многие отрасли современной промышленности уже невозможны без средств цифрового проектирования и моделирования поведения разрабатываемых изделий. Современная наука тоже в значительной части оперирует не физическими объектами, а математическими моделями процессов и явлений. При этом сложность моделей, обеспечивающих приемлемую достоверность воспроизведения характеристик механизмов, систем, процессов уже давно далеко превышает возможности обычных (персональных) компьютеров.

Направление решения этой проблемы было найдено на пути перехода к массово - параллельным вычислительным системам. Хотя создание таких систем сталкивается с определенными техническими проблемами и ограничениями, а вопросы адаптации алгоритмов к их возможностям чрезвычайно сложны, в отсутствие других альтернатив это направление быстро развивается.

Для практического использования параллельных вычислительных систем важно быть в курсе их технологических и архитектурных особенностей, уметь правильно конструировать и эксплуатировать их.

Передо мной была поставлена задача изучить архитектурные особенности современных суперкомпьютерных систем, способы и инструменты оценки их производительности. С тем, чтобы с учетом этих знаний, проанализировать архитектуру и производительность вычислительного кластера СГУ.

Практика показывает, что при решении практических задач на параллельных системах, существует очень много факторов, влияющих на эффективность их решения. Это могут быть как параметры самой обчислительной модели, так и предоставляемые в её распоряжение вычислительные ресурсы. Исследование таких особенностей поведения реальных задач мне было поручено выполнить на одной из версий кода, разработанного для моделирования процессов в сильных электрических

полях, описываемых уравнениями квантовой электродинамики. Эта задача ресурсоёмка и оценка её вычислительной сложности при фактически интересных значениях параметров и поиск путей сокращения времени её решения важны для достижения значимых результатов моделирования.

Магистерская работа состоит из введения, четырёх глав и одной под главы, заключения, списка использованных источников и приложения.

В первой главе поясняются архитектурные особенности современных высокопроизводительных вычислительных систем.

Во второй главе рассматриваются программные технологии реализации аппаратного параллелизма и оценки их эффективности.

В третьей главе описывается архитектура и особенности вычислительного кластера СГУ и проводится тестирование производительности двух узлов разного состава, с помощью теста HPL.

В четвёртой главе проведён вычислительный эксперимент, произведена оценка и оптимизация вычислительной задачи на примере.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Архитектурные особенности современных высокопроизводительных вычислительных систем. Применение параллельных вычислений является актуальным направлением развития ЭВМ. Это обосновано физическим ограничением максимально возможного быстродействия последовательных ЭВМ и существованием вычислительных задач для решения которых, возможностей существующей вычислительной техники недостаточно.

Необходимость параллельных вычислений объясняется запросами теории и практики при моделировании климата, в геномной инженерии, проектировании интегральных схем, анализе загрязнения окружающей среды, создании лекарственных препаратов и других, вычислительно сложных задачах. Эти задачи требуют для своего анализа ЭВМ с производительностью более 1000 миллиардов операций с плавающей запятой в секунду (*1 TFlops*). Чего невозможно достичь на существующих последовательных ЭВМ.

Знание современных тенденций развития ЭВМ и аппаратных средств реализации параллелизма, умения разрабатывать модели, методы и программы параллельного решения задач обработки данных, следует отнести к важным квалификационным характеристикам современного программиста.

Производительность суперкомпьютера, как правило, оценивают и выражают в количестве операций над числами с плавающей точкой в секунду, а именно Floating-point Operations Per Second (FLOPS). Это связано с задачами численного моделирования, для которых по большей части и создаются суперкомпьютеры. Чаще всего эти задачи требуют вычислений, связанных с вещественными числами, зачастую с высокой степенью точности, а не целыми числами. По этой причине к суперкомпьютерам неприменима мера быстродействия обычных компьютерных систем - количество миллионов операций в секунду. Хотя данная мера довольно неоднозначна и приближительна, оценка в FLOPS позволяет сравнивать

суперкомпьютерные системы друг с другом, опираясь на объективный критерий.

Первые ЭВМ имели производительность около 1 кфлопса, то есть 1000 операций с плавающей точкой в секунду. Компьютер, имевший производительность в 1 Мфлопс (CDC 6600), был создан в 1964 году в Соединённых Штатах. Производительность в 1 Гигафлопс была достигнута суперкомпьютерами NEC SX-2 в 1983 году с результатом 1.3 Гфлопс, и М-13 академика Карцева с результатом в 2,4 Гфлопс. Рубеж в 1 Тфлопс был преодолен в 1996 году суперкомпьютером ASCI Red. Наконец, производительность в 1 Петафлопс была достигнута в 2008 году суперкомпьютером IBM Roadrunner.

На данный момент целью является достижение значения производительности системы в 1 Экзафлопс. График роста производительности вычислительных систем показан на рисунке 1.

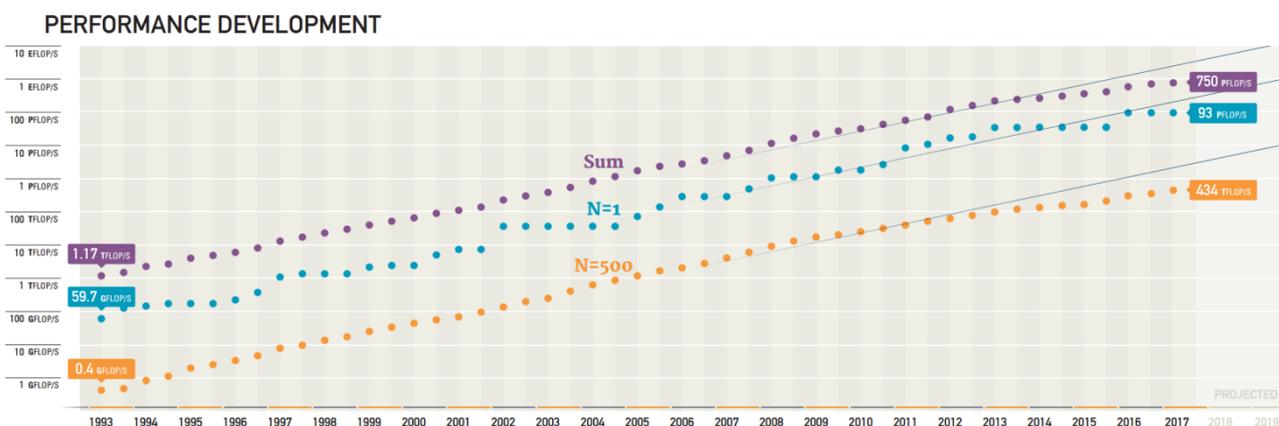


Рисунок 1 - Рост производительности по версии Top500 List.

Ресурсом для отслеживания роста производительности современных суперкомпьютеров является сайт www.top500.org. На данном сайте можно найти информацию о 500 мощнейших суперкомпьютерах мира.

По результатам на ноябрь 2017 года лидером в данной области является Китай с их суперкомпьютером Sunway TaihuLight. Его производительность по результатам тестирования с помощью Linpack равняется 93014,6 TFlops. Пиковая же его мощность оценивается в 125436

TFlops. В составе данного суперкомпьютера имеется 10649600 ядер и 1310720 ГБ памяти. В суперкомпьютере используются процессоры Sunway SW26010 260C 1.45GHz, разработанные Китаем специально для этой системы. Суперкомпьютер работает под управлением операционной системой Sunway RaiseOS 2.0.5. Узлы вычислительной системы связываются между собой посредством технологии, которая тоже разработана специально под этот проект и, соответственно, имеет название Sunway.

Большинство вычислительных систем на момент составления списка функционируют на процессорах Intel. Если быть точным, 86% суперкомпьютеров в данном списке используют процессоры Intel Xeon.

Стоит отметить, что основной технологией для связи узлов между собой в данном списке является проверенный временем Ethernet (около сорока шести процентов от общего количества систем). Чуть менее популярным средством связи узлов в суперкомпьютерах данного списка является Infiniband и он используется в тридцати двух процентах случаев от общего количества вычислительных систем в списке. Стоит отметить проприетарную технологию фирмы Cray - Aries interconnect, она используется в восьми процентах устройств в данном списке, разработанных и произведенных этой компанией. Оставшиеся четырнадцать процентов используют иные специализированные технологии (например, упомянутая Sunway).

Слабейшей вычислительной системой из данного списка является Discover SCU11, он расположен в США. Её тестовая производительность равняется 548 TFlops. Пиковая же мощность оценивается в 712 TFlops. В составе данного суперкомпьютера имеется 17136 ядер и 78336 ГБ памяти. В суперкомпьютере используются процессоры Xeon E5-2697v3 14C 2.6GHz. Суперкомпьютер работает под управлением операционной системой SUSE Linux Enterprise Server 11. Узлы системы связываются между собой Infiniband FDR.

Мощнейшим суперкомпьютером в СНГ является Ломоносов-2, он расположен в МГУ. Среди вычислительных систем данного списка он занимает 63 место. Его производительность по результатам тестирования с помощью Linpack оценивается в 2102 TFlop/c. Пиковая же его мощность оценивается в 2962.25 TFlop/c. В составе этого суперкомпьютера имеется 42688 ядер и 94208 ГБ памяти. В данном суперкомпьютере используются процессоры Xeon E5-2697v3 14C 2.6GHz. Помимо этого в суперкомпьютере Ломоносов-2 для вычислений используются графические процессоры Nvidia K40m. Суперкомпьютер работает под управлением операционной системой Linux. Узлы вычислительной системы связываются между собой посредством технологии Infiniband FDR.

OpenMP и MPI. OpenMP является самым распространённым средством программирования многопроцессорных систем с общей памятью. Эта технология позволяет легко добавить параллельные области в уже существующую последовательную программу.

Библиотека MPI же является самым распространённым средством программирования для параллельных вычислительных систем с разделяемой памятью, конкретнее, многопроцессорных систем (суперкомпьютеров) с разделяемой памятью или вычислительных кластеров.

Под параллельной программой в OpenMP понимается программа, для которой в специально указанных при помощи директив местах (параллельных фрагментах) исполняемый код может быть разделен на несколько отдельных командных потоков (threads). В общем виде программа представляется в виде набора последовательных и параллельных участков программного кода.

Важно, что разделение вычислений между потоками осуществляется под управлением соответствующих директив OpenMP. Равномерное распределение вычислительной нагрузки (балансировка) имеет принципиальное значение для максимально возможного ускорения выполнения параллельной программы.

Потоки могут выполняться на разных процессорах или процессорных ядрах, либо могут группироваться для исполнения на одном вычислительном элементе, в этом случае их исполнение осуществляется в режиме разделения времени. В предельном случае для выполнения параллельной программы может использоваться один процессор – как и в MPI такой способ применяется для начальной проверки правильности параллельной программы.

В начале выполнения параллельных фрагментов программы определяется количество потоков и обычно оно совпадает с количеством имеющихся вычислительных элементов в системе, изменение количества создаваемых потоков может быть выполнено при помощи целого ряда средств OpenMP. Все потоки в параллельных фрагментах программы последовательно пронумерованы от 0 до $n-1$, где n есть общее количество потоков. Номер потока может быть получен при помощи функции OpenMP.

Использование потоков в технологии OpenMP для организации параллелизма позволяет учитывать преимущества многопроцессорных вычислительных систем с общей памятью. Выполнение потоков одной и той же параллельной программы в общем адресном пространстве обеспечивает возможность использования общих данных в параллельно выполняемых потоках без каких-либо трудоемких межпроцессорных передач сообщений, в отличие от процессов в технологии MPI для систем с распределенной памятью. И, кроме того, управление потоками (создание, приостановка, активизация, завершение) требует меньше накладных расходов для операционной системы по сравнению с процессами.

В MPI же под параллельной программой понимается несколько процессов выполняемых одновременно. Они могут выполняться на разных процессорах, однако на одном процессоре может выполняться и несколько процессов. В предельном случае для выполнения параллельной программы может использоваться только один процессор – такой способ применяется при первичной проверке правильности параллельной программы.

Все процессы параллельной программы создаются на основе копии одного и того же программного кода (модель SPMP). Программный код, представленный в виде исполняемой программы, должен быть доступен во время запуска параллельной программы на всех используемых процессорах. Исходный программный код исполняемой программы разрабатывается на алгоритмических языках С или Fortran с использованием реализации библиотеки MPI.

Количество процессов и число используемых процессоров определяется средствами среды исполнения MPI-программ в момент запуска параллельной программы, который не может меняться в ходе вычислений. Однако возможность динамического изменения количества процессов предусмотрена в стандарте MPI-2. Все процессы программы последовательно перенумерованы от 0 до $n-1$, где n общее количество процессов. А номер процесса называется рангом процесса.

Архитектура и особенности вычислительного кластера СГУ.

Вычислительный кластер СГУ состоит в общей сложности из 18 узлов: одного управляющего и семнадцати счетных. Узлы кластера обладают следующими характеристиками: 10 узлов с RAM 16 Гб, Intel(R) Xeon(R) CPU E5345 2.33ГГц (8 потоков на узле), 2 узла с RAM 64 Гб, Intel(R) Xeon(R) CPU E5-2660v2 2.26 ГГц (40 потоков на узле), 2 узла с RAM 16 Гб, Intel(R) Xeon(R) CPU E5-2603v2 1.8 ГГц (4 потока на узле), 2 узла с RAM 32 Гб, Intel(R) Xeon(R) CPU E5-2603v2 1.8 ГГц (8 потоков на узле) и 2 узла с RAM 128 Гб, Intel(R) Xeon(R) CPU E5-2640v2 2.5 ГГц (32 потока на узле), OS Fedora 10, сеть 1000 Мбит/с. Также имеется два узла, на которых установлены ускорители Intel Xeon Phi 5110P, процессоры Intel Xeon E5-2603v2. Помимо этого в составе кластера СГУ имеется, два узла расположенных на факультете КНИИТ, которые имеют в своём составе ускорители Nvidia Tesla PNY K20. Пиковой производительностью кластера СГУ является значение в 1.2 TFlop/s.

За распределение заданий по счетным узлам кластера отвечает менеджер ресурсов Torque, в качестве средства для распараллеливания задач используется MPI в реализации MPICH2.

При тестировании вычислительного кластера СГУ мной был использован тест HPL 2017 года. Тестирование проводилось с максимально возможным для использования количеством физических ядер процессоров на двух узлах с процессорами Intel(R) Xeon(R) CPU E5345 2.33ГГц на первом узле и Intel(R) Xeon(R) CPU E5-2660v2 2.26 ГГц. Результаты тестирования приведены в виде графика на диаграмме 1 по оси y указаны размеры матрицы, а по оси x производительность в GFlops. По причине превышения максимального объема оперативной памяти для прохождения теста результаты тестирования для первого узла указаны до максимального размера матрицы в 62008. Максимальным размером матрицы для тестирования обоих узлов было выбрано значение в 85000. Как можно заметить на диаграмме 1 значение производительности узла практически не

изменялось после размера матрицы в 10000.

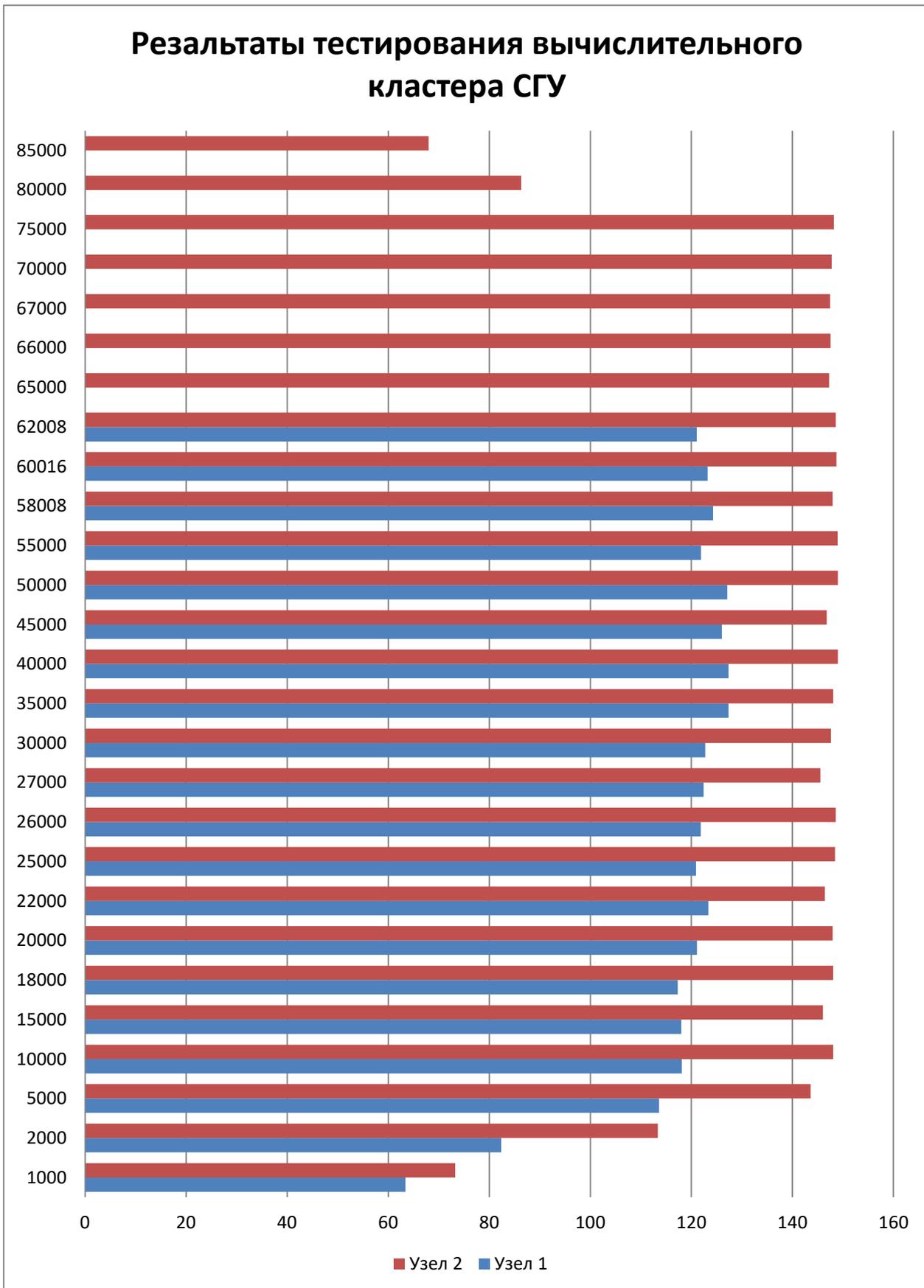


Диаграмма 1 - Результаты тестирования вычислительного кластера СГУ с помощью теста HPL.

Оценка и оптимизация вычислительной задачи на примере. В ходе работы требовалось решить систему уравнений для набора параметров, определенных физической задачей, и исследовать поведение функции плотности распределения на заданной области импульсного пространства.

σ - это параметр, которым задаётся временной интервал действия импульса поля. В таблице 1 указаны значения параметра σ и соответствующее им время решения задачи.

Таблица 1 - Зависимость времени решения задачи от параметра σ .

| Значение параметра σ | Время счёта в секундах |
|-----------------------------|------------------------|
| 4 | 4.308334 |
| 6 | 5.976070 |
| 8 | 7.859549 |
| 12 | 10.901475 |
| 16 | 10.996152 |
| 24 | 18.880059 |

ChastotaP – это параметр частоты электрического поля (его основной гармоники при сложной зависимости напряженности поля от времени). В таблице 2 указаны значения параметра chastotaP и соответствующее им время решения задачи.

Таблица 2 - Зависимость времени решения задачи от параметра chastotaP.

| | | | | | |
|------------------------------|-----------|-----------|------------|------------|-------------|
| Значение параметра ChastotaP | 0.01 | 0.005 | 0.002 | 0.001 | 0.0005 |
| Время счёта в секундах | 10.359967 | 40.496934 | 240.341460 | 918.734432 | 3278.643999 |

В таблице 3 отображена зависимость времени решения задачи от аппаратных особенностей платформы.

Таблица 3 - Зависимость времени решения задачи от аппаратных особенностей платформы

| | | | | | |
|---------------------------------------|-----------|-----------|-----------|-----------|----------|
| Количество узлов x количество потоков | 1x8 | 2x4 | 4x2 | 8x1 | 2x8 |
| Время счёта в секундах | 13.763526 | 13.354678 | 12.053225 | 13.369010 | 8.815925 |

ЗАКЛЮЧЕНИЕ

В представленной работе были изучены архитектурные особенности современных суперкомпьютерных систем, способы и инструменты оценки их производительности. Фактическим материалом для этого служили спецификации и описания наиболее производительных вычислительных систем, информация о которых представлена в открытых источниках. В них использованы самые современные технические и технологические решения.

Также мной был изучен гетерогенный вычислительный кластер СГУ. Был сделан вывод, что хотя его производительность относительно скромная, использованные в нем решения и технологии вполне репрезентативны и позволяют знакомиться с современными методами решения сложных вычислительных задач, с потенциалом масштабирования вплоть до самых мощных современных систем.

Мною на кластере СГУ выполнялись как стандартные тесты производительности, так и подробное исследование возможности оптимизации специализированного кода, разработанного для моделирования процессов в сильных электрических полях, описываемых уравнениями квантовой электродинамики. Это пример актуальной и перспективной научной задачи, сама возможность решения которой определяется наличием необходимых вычислительных ресурсов и умением их эффективно использовать.

Полученные результаты предоставили фактический материал для дальнейшего развития и совершенствования программы.