

УДК 004.272.43.003.13

Юрий Иванович Сбитнев, вед. инж. ОАО «Ювелиры Урала»,
корпорация ЯВА, Екатеринбург, Россия,

Геннадий Григорьевич Швачич, канд. техн. наук, доцент,

Максим Александрович Ткач, зав. лаб.

Национальная металлургическая академия Украины,

Днепропетровск, Украина

sgg1@ukr.net

О ПРОБЛЕМЕ ИССЛЕДОВАНИЯ ЭФФЕКТИВНОСТИ МОДУЛЬНОЙ КЛАСТЕРНОЙ СИСТЕМЫ

Работа посвящена проблеме исследования оценок эффективности многопроцессорной вычислительной системы. Основное внимание уделяется исследованию эффективности и ускорения вычислений за счет увеличения количества узлов кластерной системы. Выведены аналитические соотношения основных оценок эффективности вычислений через параметры многопроцессорной системы.

Ключевые слова: *персональный вычислительный кластер, многопроцессорная вычислительная система, ускорение, эффективность, вычислительные узлы.*

Введение

Применение параллельных вычислительных систем представляет собой стратегическое направление развития вычислительной техники. Это обстоятельство вызвано не только принципиальным ограничением максимально возможного быстродействия обычных последовательных ЭВМ, но объясняется практически постоянным существованием вычислительных задач, для решения которых недостаточно возможностей существующих средств вычислительной техники. К таким задачам относятся, например, численное моделирование процессов гидродинамики и металлургической теплофизики [1,2,3], задачи распознавания образов, оптимизационные задачи с большим количеством параметров, моделирования климата, расчеты в генной инженерии, проектирования интегральных схем, анализ загрязнения окружающей среды [4], а также решение широкого круга многомерных нестационарных задач [5], и др.

Однако эффективность параллелизации вычислений существенно зависит от многих факторов, один из важнейших – это особенности пересылки данных между соседними узлами многопроцессорной системы, которая обычно является самой медленной частью алгоритма и может свести на нет эффект от увеличения числа используемых процессоров.

Практически одновременно с появлением компьютеров возникла необходимость в оценке их производительности и в последующем сравнении вычислительных систем, учитывая этот параметр. Конечно, хотелось бы иметь простую и единственную методику априорного сравнения вычислительных систем между собой. В идеальной ситуации это могло бы быть определение по некоторому закону одного числа для каждой системы, которое бы и отображало их общую характеристику. С этой точки зрения естественным признаком любого компьютера выступает его пиковая производительность. Этот показатель определяет тот максимум, на какой способный соответствующий компьютер. Для задач вычислительного характера во многих случаях операции над вещественными числами выполняются медленнее, чем, скажем, управляющие, или регистровые, операции. Это послужило причиной применения другого способа измерения пиковой производительности, когда учитывается количество действительных операций, которые выполняет компьютер за единицу времени. Для этого была введена единица измерения – флопс. Учет вычислительной сложности программы дает возможность пользователю определить только нижний предел оценки затраты времени на ее выполнение. Существуют программы, которые исполняют роль эталона, то есть они позволяют судить о возможностях вычислительной системы. На основании разных критериев создавали так называемые тестовые наборы программ. Например, таким тестом стал *LINPACK*, он предназначен для решения систем линейных алгебраических уравнений.

Характеристики компьютеров, полученные по результатам тестирования, всегда вызывали, и будут вызывать недоверие и критику. Одно из возможных направлений выхода из данной ситуации заключается в формировании набора тестов, каждый из которых дополнял бы друг друга. Таким образом, если одно число не может адекватно охарактеризовать вычислительную систему, то можно попробовать сделать это при помощи набора чисел.

Из этих позиций и будет оцениваться эффективность многопроцессорной кластерной системы, которая исследуется в данной работе. Эти исследования проведены в пределах выполнения научно-исследовательской работы «Математическое моделирование режимов термической обработки при скоростном циклическом нагревании и охлаждении длинномерного изделия».

Постановка проблемы исследований

Одна из основных проблем использования кластерной системы, для решения класса задач, которые рассматриваются в этой работе, может быть сформулирована таким образом: имеем разностную сетку размерности M , время вычисления задачи, которая развязывается с использованием однопроцессорной системы, определяется величиной t . Этот параметр

является ключевым. Необходимо существенно уменьшить время вычислений, сохраняя значение M .

Именно вопросом быстродействия и производительности уделяется основное внимание при конструировании кластерных систем. Такой подход ориентирован, например, на разработку новых технологических процессов (когда время вычислений является критической величиной) [6]. Кроме того, аналогичные задачи часто приходится решать в медицине, военной технике и др.

Анализ последних исследований и публикаций

подавляющее большинство функционирующих супервычислительных установок – это фактически многопроцессорные параллельные вычислительные системы архитектуры *MPP* (*Massively Parallel Processing*). Многопроцессорные вычислительные системы, сконструированные на базе локальных сетей, начали называть «кластерными системами» или просто «кластерами». Это объясняется тем, что логично упомянутая система *MPP* мало отличается от обычной локальной сети.

В данной работе рассматриваются так называемые «блейд» серверные решения многопроцессорных систем, при которых несколько однотипных материнских модулей устанавливаются в одном корпусе. Практика показывает, что блейд-системы более компактны и удобны в обслуживании, а их реализация не намного дороже по сравнению с многопроцессорными компьютерными комплексами. Но благодаря растущему спросу и предложению «блейд» конфигураций на нашем рынке, было принято решение о создании самой такой кластерной вычислительной системы для математического моделирования многомерных задач.

Организация блейд-кластера заключается в объединении лезвий в единую вычислительную сеть через коммутатор, который установлен в том же корпусе. Для блейд-кластера бывает достаточно одного жесткого диска, на котором находится образ загружаемой системы, при этом используют механизм сетевой загрузки *Network boot*. При включении системы *Master-узел* через сетевой коммутатор раздает *IP*-адреса всем узлам кластера, то есть происходит начальная инициализация, и кластер готов к работе.

Именно учитывая последние достижения в области конструирования многопроцессорных систем, и был создан персональный вычислительный кластер [7], блок-схема которого представлена на рис.1. К тому же в конфигурации кластера было избрано шесть лезвий и модульный принцип его реализации. Это обеспечивает в случае необходимости его расширение за счет установления дополнительных модулей. Каждый узел работает под управлением собственной копии операционной системы, причем чаще всего используют стандартные операционные системы: *Linux*, *NT*, *Solaris* и т.п. Состав и мощность узлов описанного кластера может меняться, что позволяет создавать неоднородные системы. Коммутирующая сеть соединяет процессоры друг с другом.

Особенность блок-схемы модуля многопроцессорной системы (рис. 1) заключается в том, что все его вычислительные узлы содержат процессор (1) *C7 CPU*, присоединенный шиной *FSB (Front Side Bus 533/400 МГц)* к южному мосту *Cn700* (2) с интегрированным видеоконтроллером *VIA UniChrome Pro* и видеовыходами *SVGA* (3), *TV* (4) и интерфейсо *AGP 8X* (5), а южный мост подключен к локальной памяти (6), которая работает по стандартам *DDR2 533/400* или *DDR 400/333/266*. Южный и северный мосты соединены в соответствии с модульной архитектурой платформ *VIA V-MAP* (7) (*Modular Architecture Platform*). Для соединения северного моста на чипсете *VT8237A* (8) и южного на чипсете *V-MAP* предусмотрено использование шины *Ultra V-Link*, которая работает со скоростью 533 МБ/с. К чипсету подключен контроллер *VIA Drive Station* (9), который поддерживает интерфейсы *SATA*, *PATA* и режим *RAID*, а также присоединена шина *PCI Bus* с двумя разъемными соединениями *PCI* (10,11), в которых установлены сетевые интерфейсы, которые поддерживают режимы *channel bonding* и *Gigabit Ethernet* (12, 13). К мосту *VIA VT8237A* подключен интегрированный аудиоконтроллер (14) *VIA Vinyl™ HD Audio*, контроллер клавиатуры и манипулятора мыши *PS/2* (15), а также восемь высокоскоростных портов стандарта *USB, 2.0* (16), и контроллер *VT1211* (17), который представляет собой полнофункциональный *Super I/O*-чип с контроллером дисководов гибких дисков, интерфейсом параллельного порта *IEEE-1284*, двумя последовательными портами *16C550-UART*, контроллером *VFIR* (скоростной инфракрасный порт), игровым портом, который поддерживается двумя джойстиком, *MIDI*-интерфейсом и интерфейсом *4M FLASH-ROM BIOS* (18), интегрированным сетевым интерфейсом, который может поддерживать режим сетевой загрузки и *Fast Ethernet* (19).

Основой такого кластера является, установлены в стойке плотно упакованные системы с процессорами лезвийного типа. Внутри стойки размещены узлы, аппаратура для эффективного соединения компонентов, аппаратура управления внутренней сетью системы и др.

Узлы кластера могут быть функционально объединены в две группы, а именно:

1. Вычислительные узлы, которые решают основные вычислительные задачи, для которых и спроектирована кластерная система.

2. Узлы инфраструктуры, в частности ввод – вывод, узлы управления и узлы запоминающих устройств. Они обеспечивают управление системами и заданными функциями, которые необходимы для объединения компьютерных узлов в цельный комплекс.

Упаковка вычислительных узлов, на сколько это, возможно, является плотной, и имеет необходимые условия для эффективного соединения компонентов. Главные узлы, узлы управления и узлы запоминающих устройств, обеспечивают особые функции управления кластером – загрузка, управление устройствами, внешний ввод – вывод т. д.

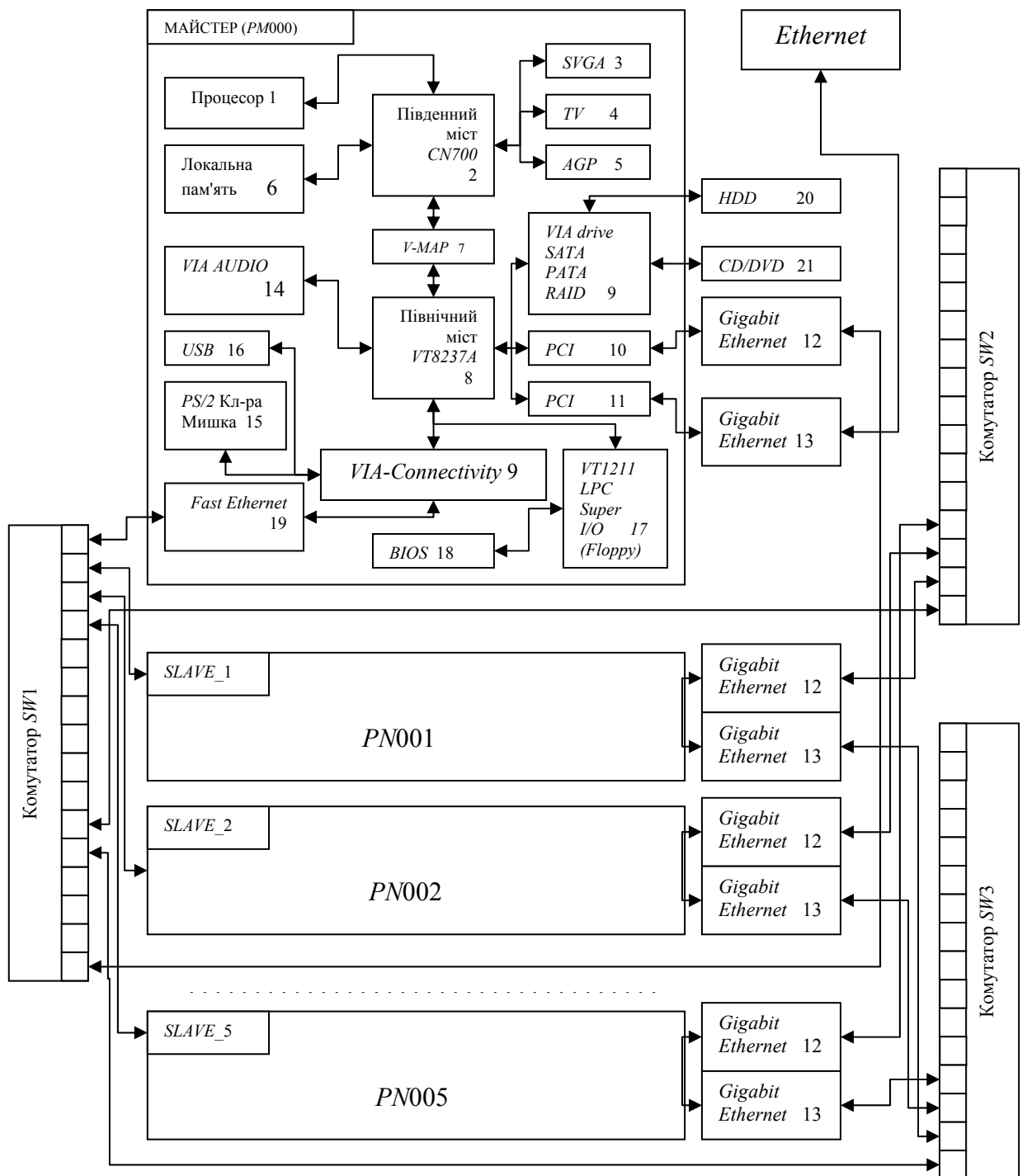


Рисунок 1 – Блок-схема модуля многопроцессорной системы

Персональний вичислительний кластер, блок-схема якого зображена на рис. 1, має такі розміри: ширина 19, висота 10,9, глибина 9 дюймів. Вес кластера становить приблизительно 8 кг.

Цель и задачи исследований

Задачи, на решение которых направлена эта работа, заключается в исследовании оценок эффективности разработанной многопроцессорной кластерной вычислительной системы.

При этом необходимо решить такие задачи:

- оценить эффективность кластерной системы в организации одностороннего и двустороннего режимов граничного обмена данными;
- оценить эффективность кластерной системы в организации полудуплексного и дуплексного режимов работы;
- оценить эффективности кластерной системы при организации многоканальных режимов функционирования вычислительной сети кластера.

Кроме того, необходимо провести исследования, направленные на определение загруженности линий связи кластерной системы. При этом надлежит вывести аналитическое соотношение для коэффициента загруженности линий связи кластерной системы через ее параметры.

Для удобства проведения исследований эффективности персонального вычислительного кластера требуется также установить аналитические соотношения для определения его основных числовых характеристик через параметры вычислительной системы. Проведенные исследования направлены на дальнейшее развитие подхода, освещенного авторами в [8].

Изложение основного материала исследований

Итак, рассматриваем задачу уменьшения времени расчетов путем увеличения числа узлов кластерной системы. При этом будем считать, что область вычислений равномерно распределяется между узлами кластерной системы. Для удобства исследований будем предполагать, что форма области, на которой проводятся вычисления, является квадратной.

В данном классе задач все вычисления выполняются на базе разностной сетки. К тому же для анализа эффективности многопроцессорной системы важнейшим параметром будет время счета одной итерации (T_{it}) относительно области вычислений. Тогда в условиях применения многопроцессорной системы общее время одной итерации будет определяться на основании такого соотношения:

$$T = T_n + T_{ex} , \quad (1)$$

Здесь $T_n = \frac{T_{it}}{N}$, представляет собой время счета одной итерации при использовании N вычислительных узлов, с., T_{ex} – время граничного обмена между узлами кластера, с. Заметим, что когда время счета итерации зависит лишь от мощности процессора, то время граничного обмена диктуется

размером разностной сетки, количеством узлов кластерной системы и пропускной способностью вычислительной сети.

Следовательно, величину T_{ex} можно определить из такого соотношения:

$$T_{ex} = \frac{E}{V}. \quad (2)$$

В выражении (2) E – объем данных в области граничного обмена (Гбит), V – пропускная способность сети кластера (Гбит/с). В условиях, когда область вычислений максимально загружена и равномерно распределена между узлами узла кластера, можно определить формулу для вычисления объема данных граничного обмена. Эта формула будет иметь такой вид:

$$E = m \cdot (N - 1) \sqrt{R}, \quad (3)$$

где N – количество узлов кластера, R – объем оперативной памяти узла кластера. Отметим, что когда $N = 1$, то значение величины E обратится в нуль, что является полностью очевидным. Значение m может равняться единице для одностороннего режима граничного обмена данными, или двум для двустороннего.

В отмеченных условиях можно оценить время одной итерации, в состав которого будет входить, собственно, время счета одной итерации при использовании N узлов кластерной системы и время граничного обмена данными в зависимости от количества узлов кластера N , то есть:

$$T = \frac{T_i}{N} + \frac{m \cdot (N - 1) \cdot \sqrt{R}}{V}. \quad (4)$$

Анализ соотношения (4) показывает, что, разделяя область вычислений между узлами кластера, уменьшаем объем вычислений для его каждого лезвия. В силу того, что узлы многопроцессорной системы работают параллельно, то и общее время итерации становится меньше. В то же время, с увеличением узлов кластера увеличивается также и объем граничных данных, и соответственно время обмена данными между узлами. Реализуя режим агрегации каналов вычислительной сети кластера, можно увеличить скорость обмена данными в сети в k раз, при этом величина k соответствует количеству каналов, которые работают параллельно. Тогда производительность сети будет определяться таким образом:

$$V = k \cdot d \cdot V_n, \quad (5)$$

где V_n – пропускная способность сети кластера, Гбит/с, k – количество каналов связи вычислительной сети, которые работают одновременно

(количество вычислительных сетей), d – полудуплексный ($d = 1$), или дуплексный ($d = 2$) режим работы вычислительной сети кластерной системы.

Для расчета ускорения и эффективности за основу были взяты общепринятые в теории параллельных вычислений формулы. Для эффективности распараллеливания:

$$Q = \frac{S}{N}, \quad (6)$$

при этом

$$S = \frac{t_1}{t_n}, \quad (7)$$

где S – ускорение; t_1 – время выполнения задачи на одном процессоре; t_n – время выполнения задачи на n процессорах.

Оценку эффективности кластерной системы можно получить через основные параметры персонального вычислительного кластера, в частности через временные параметры вычислительного процесса, а именно:

$$\eta = \frac{T_n}{T}, \quad (8)$$

Тогда с учетом соотношения (1) формулу (8) можно записать таким образом:

$$\eta = \frac{\frac{T_i}{N}}{\frac{T_i}{N} + T_{ex}}. \quad (9)$$

Принимая во внимание выражения (2), (3) (5) получаем окончательную оценку эффективности кластерной системы через ее параметры, то есть:

$$\eta = \frac{T_i \cdot k \cdot d \cdot V_n \cdot N}{T_i \cdot k \cdot d \cdot V_n \cdot N + N^2 \cdot m \cdot (N-1) \cdot \sqrt{R}}. \quad (10)$$

Расчеты эффективности кластерной системы, которые проводились по формулам (6) и (10), позволяют отметить совпадение результатов с точностью до одного знака, который объясняется характером вычислений.

В соответствии с этими соотношениями были проведены вычислительные эксперименты для многопроцессорной кластерной системы,

блок-схема которой изображена на рис. 1. Здесь выходными были приняты такие характеристики класса решаемых задач и самой кластерной системы:

$$T_{it} = 100 \text{ с}, V_n = 1 \text{ Гбит/с}, R = 8 \text{ Гбит.}$$

На первом этапе для отображения полной картины процессов, которые происходят в многопроцессорной вычислительной системе, были проведены сугубо теоретические исследования. Для полудуплексного программируемого режима передачи данных в вычислительной системе были рассмотрены режимы односторонней и двусторонней пересылки данных между ее соседними узлами.

Заметим, что полудуплексный режим позволяет передавать информацию каналами связи по очереди в разные моменты времени. Схема реализации такого режима изображена на рис. 2.

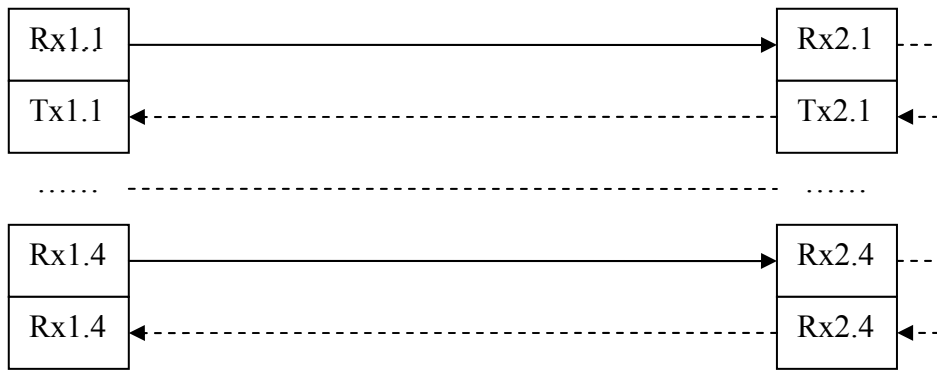


Рисунок 2 – Схема полудуплексного режима передачи данных в вычислительной системе

На рис. 2 пунктирными линиями показан процесс согласования приема и передачи данных.

Исследование особенностей организации граничного обмена данными

При решении заданного типа задач в течение каждой итерации процессоры обмениваются данными на стыках блоков, используя актуальные значения переменных. На первом этапе особенности пересылки данных между соседними узлами многопроцессорной системы исследуются в одном направлении.

Исходные данные для исследования представлены в табл. 1.

Таблица 1. Исходные даны для расчета характеристик системы при одностороннем режиме обмена информацией

V_n	1 Гбит/с
T_{it}	100 с
R	8 Гбит
m	1
d	1
k	1

Полученные результаты моделирования отображены в табл. 2.

Таблица 2. Результаты расчета основных характеристик системы при одностороннем режиме обмена информацией

колич. узлов	T_n	T_{ex}	T	USK	EF
1	100,00	0,00	100,00	1,00	1,00
2	50,00	2,83	52,83	1,89	0,95
3	33,33	5,66	38,99	2,56	0,85
4	25,00	8,49	33,49	2,99	0,75
5	20,00	11,31	31,31	3,19	0,64
6	16,67	14,14	30,81	3,25	0,54
7	14,29	16,97	31,26	3,20	0,46
8	12,50	19,80	32,30	3,10	0,39
9	11,11	22,63	33,74	2,96	0,33
10	10,00	25,46	35,46	2,82	0,28
11	9,09	28,28	37,38	2,68	0,24
12	8,33	31,11	39,45	2,54	0,21
13	7,69	33,94	41,63	2,40	0,18
14	7,14	36,77	43,91	2,28	0,16
15	6,67	39,60	46,26	2,16	0,14

Результаты моделирования представлены также в виде графических зависимостей (рис. 3, 4).

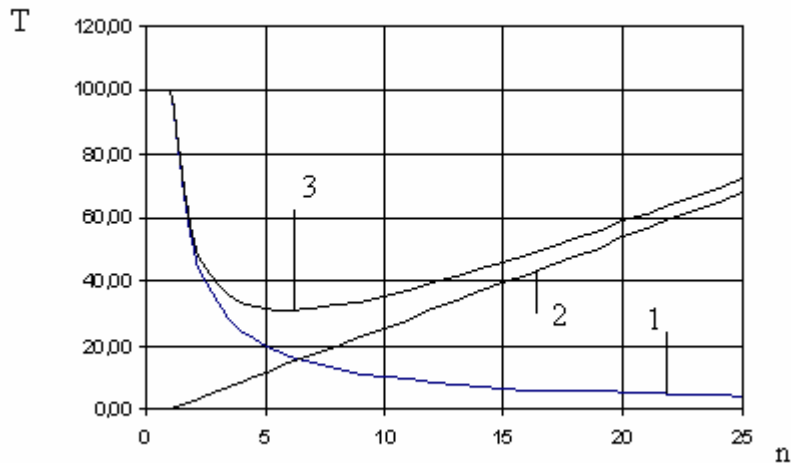


Рисунок 3 – Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы при одностороннем граничном обмене данными

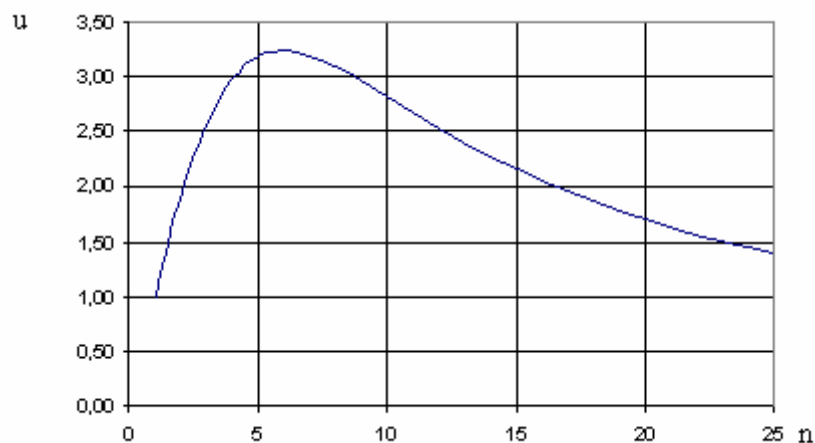


Рисунок 4 – Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы при одностороннем граничном обмене данными

Как видим, на рис.3, время счета одной итерации при увеличении числа узлов многопроцессорной системы уменьшается за гиперболической зависимостью (кривая 1). Наряду с этим время граничного обмена при увеличении числа узлов многопроцессорной системы увеличивается по линейному закону (линия 2). Общую картину изменения времени счета одной итерации в многопроцессорной системе иллюстрирует зависимость, отображенная кривой 3. Анализ такой кривой показывает, что время расчета на первом этапе уменьшается при увеличении количества узлов кластера. Подобный результат, казалось, и был предусмотрен. Однако уменьшение такого времени происходит до определенного предела. Если, например, количество узлов будет, превышает шести, то общее время расчета начинает расти. Происходит это на фоне увеличения объема данных, которые

пересылаются между узлами. Таким образом, можно отметить, что при постоянном размере сетки, в условиях данной задачи, нет смысла, чтобы количество лезвий в кластере были больше шести. Сравнительно невысокую эффективность вычислений (54 %) при минимальном времени вычислений одной итерации с использованием шести узлов вычислительной системы можно объяснить тем, что ресурсы кластера используются не в полном объеме.

При прочих равных условиях рассмотрим наиболее типичный вариант пересылки данных между соседними узлами системы – двусторонний. Исходные данные для исследования такого режима работы вычислительной системы представлены в табл. 3.

Таблица 3. Исходные данные для расчета характеристик системы при двустороннем режиме обмена информацией

V_n	1 Гбит/с
T_{it}	100 с
R	8 Гбит
m	2
d	1
k	1

Полученные результаты моделирования представлены в табл. 4.

Таблица 4. Результаты расчета основных характеристик системы при двустороннем режиме обмена информацией

колич. узлов	T_n	T_{ex}	T	USK	EF
1	100,00	0,00	100,00	1,00	1,00
2	50,00	5,66	55,66	1,80	0,90
3	33,33	11,31	44,65	2,24	0,75
4	25,00	16,97	41,97	2,38	0,60
5	20,00	22,63	42,63	2,35	0,47
6	16,67	28,28	44,95	2,22	0,37
7	14,29	33,94	48,23	2,07	0,30
8	12,50	39,60	52,10	1,92	0,24
9	11,11	45,25	56,37	1,77	0,20
10	10,00	50,91	60,91	1,64	0,16
11	9,09	56,57	65,66	1,52	0,14
12	8,33	62,23	70,56	1,42	0,12
13	7,69	67,88	75,57	1,32	0,10
14	7,14	73,54	80,68	1,24	0,09
15	6,67	79,20	85,86	1,16	0,08

Результаты моделирования представлены также в виде графических зависимостей (рис. 5 и 6).

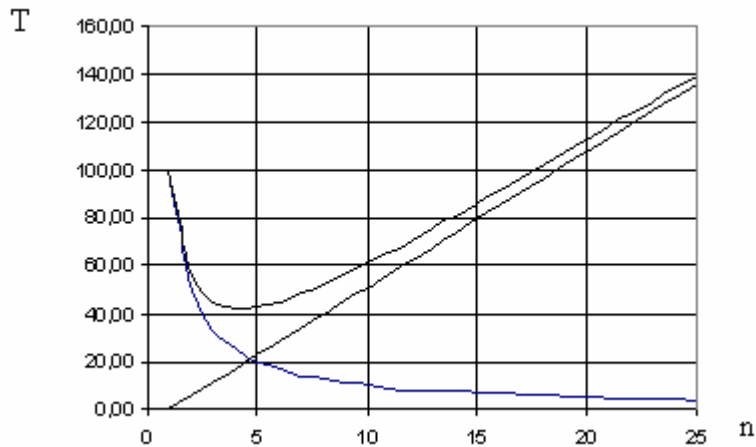


Рисунок 5 – Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы при двустороннем граничном обмене данными

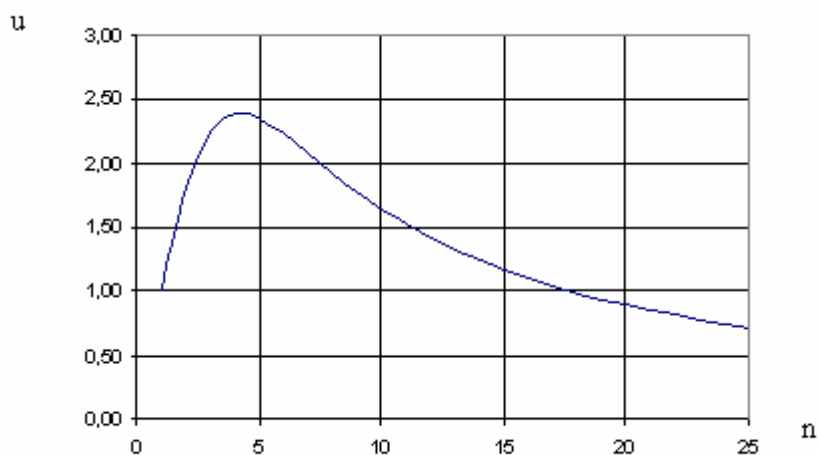


Рисунок 6 – Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы при двустороннем граничном обмене данными

Следовательно, имеем предпосылки для подведения промежуточных итогов исследований. С одной стороны, двусторонний режим обмена данными, при прочих равных условиях, позволяет в оптимальном варианте функционирование кластерной системы уменьшить количество узлов до четырех (рис. 6) и, что очевидно, повысить эффективность вычислений до 60 %. Однако, с другой стороны, как и ожидалось, при двустороннем режиме обмена данными уменьшится ускорение вычислений за счет увеличения времени граничного обмена (рис.4, 6). Очевидно, чтобы увеличить ускорение вычислений необходимо отдельно исследовать влияние способа передачи

данных между узлами кластерной системы. На решение этой проблемы и направлены последующие исследования.

Исследование особенностей реализации дуплексного режима работы системы

В спроектированной кластерной системе, блок-схема которой изображена на рис.1, для соединения внешних устройств используется шина *PCI* (англ. *Peripheral Component Interconnect*, что дословно переводится, как взаимосвязь периферийных компонентов). И здесь надлежит отметить важное нововведение: после физического подключения внешнего устройства к разъемному соединению шины *PCI* происходит автоматическое установление конфигурации этого устройства. Между прочим, шина *PCI* – это одно из основных средств для подключения самых разнообразных периферийных устройств к системной плате кластера. Среди причины успеха *PCI* – большая скорость и режим динамической конфигурации подключенных к ней периферийных устройств.

Кроме того, эта шина предусматривает дуплексный режим передачи данных. В режиме “дуплекс” устройства могут передавать и принимать информацию или данные одновременно двумя каналами связи, разделенными (разнесенными) физически. Схема организации такого режима изображена на рис. 7.

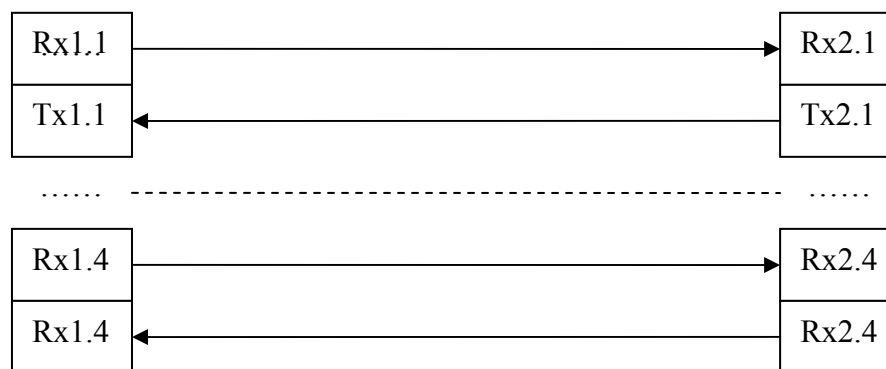


Рисунок 7 – Схема дуплексного режима передачи данных в вычислительной системе

Заметим, что дуплексный режим – наиболее скоростной в работе вычислительных систем. Он позволяет эффективно использовать вычислительные возможности кластерных комплексов в сочетании с высокой скоростью передачи данных каналами связи. Учитывая описанные обстоятельства, подчеркнем важность исследования характеристик

эффективности многопроцессорной системы при реализации дуплексного режима работы.

Исходные данные для этого режима исследований представлены в табл. 5.

Таблица 5. Исходные даны для расчета характеристик системы при реализации дуплексного режима работы кластера

V_n	1 Гбит/с
T_{it}	100 с
R	8 Гбит
m	2
d	2
k	1

Полученные результаты моделирования сведены в табл. 6.

Таблица 6. Результаты расчета основных характеристик системы при реализации дуплексного режима работы кластера

колич. узлов	T_n	T_{ex}	T	USK	EF
1	100,00	0,00	100,00	1,00	1,00
2	50,00	2,83	52,83	1,89	0,95
3	33,33	5,66	38,99	2,56	0,85
4	25,00	8,49	33,49	2,99	0,75
5	20,00	11,31	31,31	3,19	0,64
6	16,67	14,14	30,81	3,25	0,54
7	14,29	16,97	31,26	3,20	0,46
8	12,50	19,80	32,30	3,10	0,39
9	11,11	22,63	33,74	2,96	0,33
10	10,00	25,46	35,46	2,82	0,28
11	9,09	28,28	37,38	2,68	0,24
12	8,33	31,11	39,45	2,54	0,21
13	7,69	33,94	41,63	2,40	0,18
14	7,14	36,77	43,91	2,28	0,16
15	6,67	39,60	46,26	2,16	0,14

Результаты моделирования предствлены также в виде графических зависимостей (рис. 8, 9).

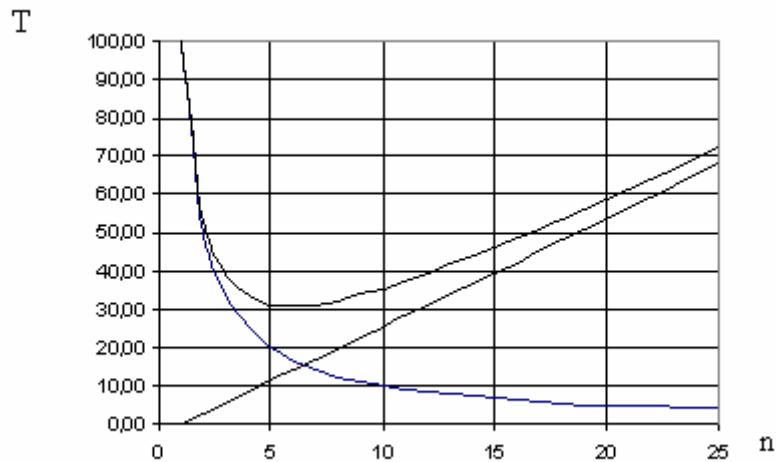


Рисунок 8 – Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы при дуплексном режиме работы

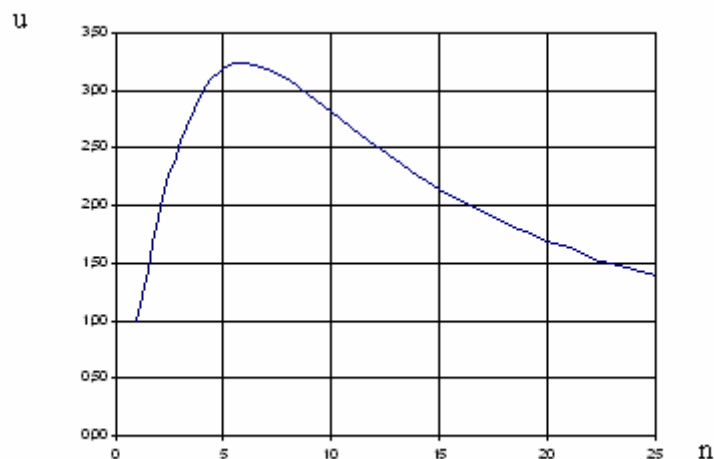


Рисунок 9 – Кривые зависимости ускорения вычислений от количества узлов многопроцессорной системы при дуплексном режиме работы

Сравнительный анализ полудуплексного (табл. 4) и дуплексного (табл. 6) режимов работы системы показывает, что в режиме “дуплекс” существенно уменьшилось время вычислений (рис. 4, 8). Кроме того, значительно выросло ускорение вычислений (см. рис. 5, 9). Однако, одновременно, необходимо отметить некоторое снижение эффективности вычислений (54 % против 60%).

Очевидно, что для повышения эффективности вычислений необходимо вместо одной коммутационной сети, где происходил обмен данными, вводить вторую, которая работает параллельно с первой, что позволит

уменьшить время граничного обмена данными. Следовательно, следующие исследования будут направлены на реализацию процедуры усовершенствования обмена данными между узлами кластерной системы на сетевом уровне.

Исследование особенностей формирования вычислительной сети кластерной системы

Основная особенность кластерной системы, которая рассматривается в этой работе (рис. 1), заключается в том, что обмен данных между вычислительными узлами вынесен в отдельную сеть, которая работает на канальном (втором) уровне с использованием механизмов *channel bonding*. Это обеспечило повышение скорости обмена данными и снизило загрузку канала, который соединяет узлы кластера. Такая реализация механизма *channel bonding* кластерной вычислительной системы позволяет повысить ее эффективность. Последующие исследования направлены на реализацию процедуры повышения эффективности распараллеливания при помощи введения дополнительных вычислительных сетей.

Исходные данные для изучения упомянутого режима работы многопроцессорной системы перечислены в табл. 7.

Таблица 7. Исходные данные для расчета характеристик системы при двухканальном режиме функционирования вычислительной сети кластера

V_n	1 Гбит/с
T_{it}	100 с
R	8 Гбит
m	2
d	2
k	2

Полученные результаты моделирования сведены в табл. 8.

Таблица 8. Результаты расчета основных характеристик системы при реализации двухканального режима функционирования вычислительной сети кластера

колич. узлов	T_n	T_{ex}	T	USK	EF
1	100,00	0,00	100,00	1,00	1,00
2	50,00	1,41	51,41	1,94	0,97
3	33,33	2,83	36,16	2,77	0,92
4	25,00	4,24	29,24	3,42	0,85
5	20,00	5,66	25,66	3,90	0,78
6	16,67	7,07	23,74	4,21	0,70
7	14,29	8,49	22,77	4,39	0,63
8	12,50	9,90	22,40	4,46	0,56
9	11,11	11,31	22,42	4,46	0,50
10	10,00	12,73	22,73	4,40	0,44
11	9,09	14,14	23,23	4,30	0,39
12	8,33	15,56	23,89	4,19	0,35
13	7,69	16,97	24,66	4,05	0,31
14	7,14	18,38	25,53	3,92	0,28
15	6,67	19,80	26,47	3,78	0,25

Результаты моделирования представлены также в виде графических зависимостей (рис. 10, 11).

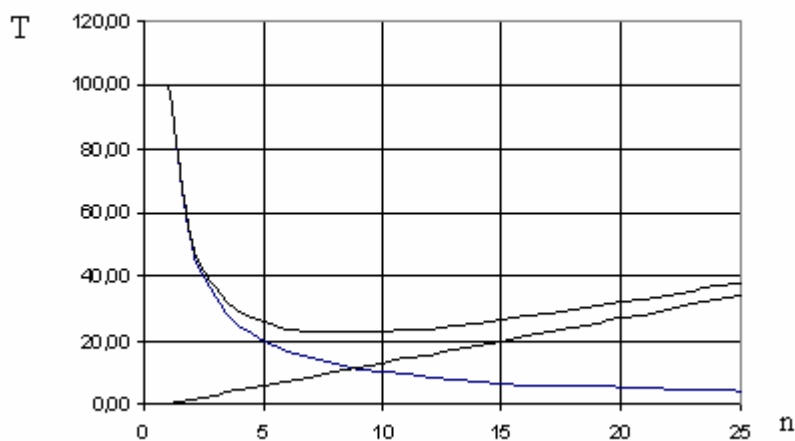


Рисунок 10 – Кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы при реализации двухканального режима функционирования вычислительной сети кластера

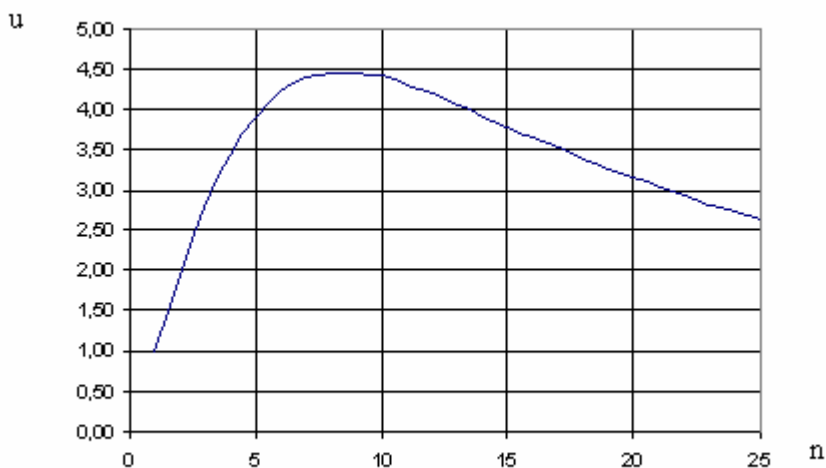


Рисунок 11 – Кривые зависимости ускорения вычислений от количества узлов многопроцессорной системы при реализации двухканального режима функционирования вычислительной сети кластера

Итак, имеем предпосылки для количественной оценки эффективности многопроцессорной системы при реализации двухканального режима функционирования вычислительной сети кластера ($\kappa = 2$). Как показывают расчетные данные (табл. 8), такой режим работы кластера позволил не только повысить эффективность системы (в оптимальном режиме это больше 70 %), но и существенно сократить время вычислений (см. рис. 8, 10), а также ускорить вычисления (см. рис. 9, 11). Таких результатов удалось достичь за счет уменьшения времени граничного обмена данными между вычислительными узлами кластерной системы. Заметим, что именно такой режим работы вычислительной системы и внедрен для кластера, блок-схема которого изображена на рис.1. Однако в ходе исследований возник вопрос: каким образом изменятся оценки эффективности кластерной системы при последующем расширении режима *channel bonding*? Например, если $\kappa = 4$. Ответ на этот вопрос можно получить, проведя дополнительные исследования.

Исходные данные для отмеченного режима работы многопроцессорной системы перечислены в табл. 9.

Таблица 9. Исходные данные для расчета характеристик системы при четырехканальном режиме функционирования вычислительной сети кластера

V_n	1 Гбит/с
T_{it}	100 с
R	8 Гбит
m	2
d	2
k	4

Полученные результаты моделирования сведены в табл. 10.

Таблица 10. Результаты расчета основных характеристик системы при реализации четырехканального режима функционирования вычислительной сети кластера

колич. узлов	T_n	T_{ex}	T	USK	EF
1	100,00	0,00	100,00	1,00	1,00
2	50,00	0,71	50,71	1,97	0,99
3	33,33	1,41	34,75	2,88	0,96
4	25,00	2,12	27,12	3,69	0,92
5	20,00	2,83	22,83	4,38	0,88
6	16,67	3,54	20,20	4,95	0,82
7	14,29	4,24	18,53	5,40	0,77
8	12,50	4,95	17,45	5,73	0,72
9	11,11	5,66	16,77	5,96	0,66
10	10,00	6,36	16,36	6,11	0,61
11	9,09	7,07	16,16	6,19	0,56
12	8,33	7,78	16,11	6,21	0,52
13	7,69	8,49	16,18	6,18	0,48
14	7,14	9,19	16,34	6,12	0,44
15	6,67	9,90	16,57	6,04	0,40

Результаты моделирования представлены также в виде графических зависимостей (рис. 12, 13).

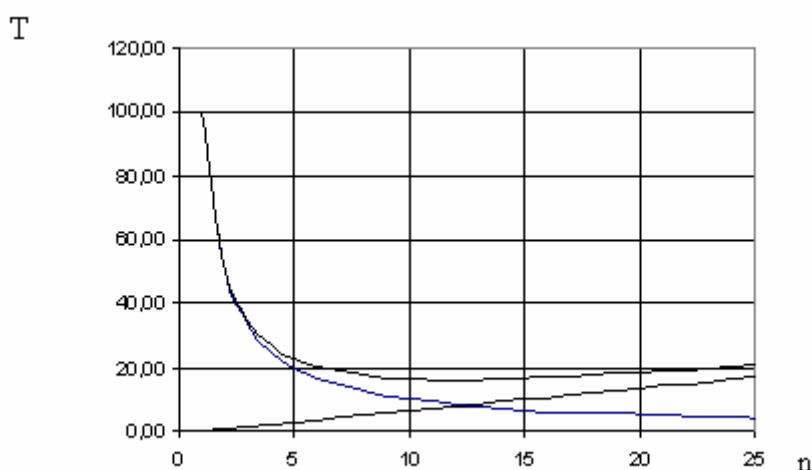


Рисунок 12 – кривые зависимости времени расчета одной итерации от количества узлов многопроцессорной системы при реализации четырехканального режима функционирования вычислительной сети кластера

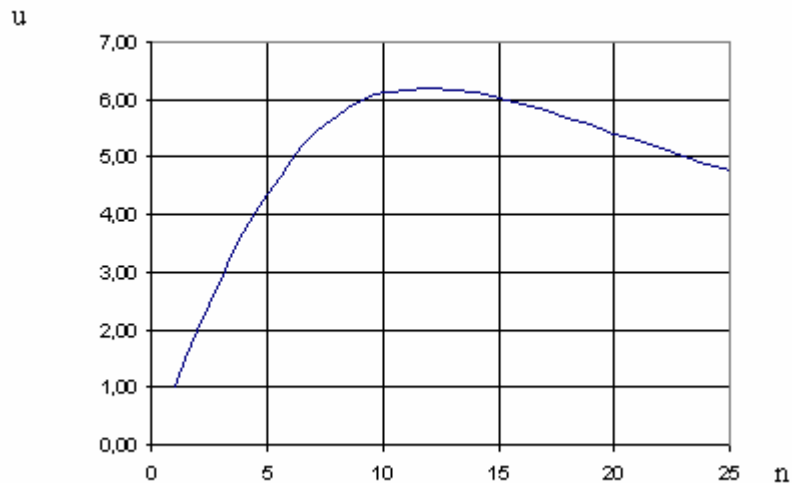


Рисунок 13 – Кривая зависимости ускорения вычислений от количества узлов многопроцессорной системы при реализации четырехканального режима функционирования вычислительной сети кластера

Таким образом, сравнительный анализ реализации режима *channel bonding* кластерной системы, когда $\kappa = 2$ и $\kappa = 4$, показывает, что при сравнительно одинаковой эффективности распараллеливания несколько уменьшается время вычислений (приблизительно на 5 с.) и значительно растет ускорение вычислений (см. рис. 11, 13). В практической работе кластерной системы с режимом формирования вычислительной сети, когда $\kappa = 4$, исследователю придется искать ответ на вопрос: будет ли оправданным уменьшение времени вычислений на 5 с. и увеличение ускорения вычислений при условии вкладывания сравнительно серьезных материальных средств в усложнение конструкции кластерной системы? В некоторой мере ответ на поставленный вопрос можно получить, если провести исследование коэффициента загрузки линий связи кластерной системы.

Исследование загрузки линий связи кластерной системы

Итак, рассмотрим характеристику коэффициента загрузки линий связи кластерной системы. С этой целью выведем соотношение для коэффициента загрузки линий связи через параметры кластерной системы, при этом принимается, что:

$$\xi = \frac{T_{ex}}{T}. \quad (11)$$

Учитывая выражения (2) – (5), окончательное значение коэффициента загрузки линий связи можно записать в виде аналитического соотношения, выраженного через параметры вычислительного кластера, т.е.:

$$\xi = \frac{m \cdot N \cdot (N - 1) \cdot \sqrt{R}}{T_i \cdot k \cdot d \cdot V_n + N \cdot m \cdot (N - 1) \cdot \sqrt{R}} \quad (12)$$

Результаты расчета коэффициента загрузки линий связи для дуплексного режима работы многопроцессорной системы без использования режима *channel bonding* (КЗС1), а также с его использованием, когда $\kappa = 2$ (КЗС2) и $\kappa = 4$ (КЗС3), приведены в табл. 9.

Таблица 9. Результаты расчета коэффициента загрузки линий связи кластера при разных режимах функционирования вычислительной сети.

колич. узлов	КЗС1	КЗС2	КЗС3
1	0,00	0,00	0,00
2	0,05	0,03	0,01
3	0,15	0,08	0,04
4	0,25	0,15	0,08
5	0,36	0,22	0,12
6	0,46	0,30	0,18
7	0,54	0,37	0,23
8	0,61	0,44	0,28
9	0,67	0,50	0,34
10	0,72	0,56	0,39
11	0,76	0,61	0,44
12	0,79	0,65	0,48
13	0,82	0,69	0,52
14	0,84	0,72	0,56
15	0,86	0,75	0,60
16	0,87	0,77	0,63
17	0,88	0,79	0,66
18	0,90	0,81	0,68
19	0,91	0,83	0,71
20	0,91	0,84	0,73
21	0,92	0,86	0,75
22	0,93	0,87	0,77
23	0,93	0,88	0,78
24	0,94	0,89	0,80
25	0,94	0,89	0,81

Полученные результаты позволяют сделать вывод, что, как и ожидалось, при увеличении числа узлов кластерной системы значения коэффициента загрузки линий связи будут расти. С другой стороны известно, что когда $\xi = 50 \%$, то оперативная память коммутатора будет

использоваться, приблизительно, на 70 %. Запас этой памяти (до 30 %) резервируется для устранения коллизий, которые могут возникать в результате загруженности вычислительной сети кластера. Следовательно, можно отметить, что при первом режиме функционирования кластера можно использовать не больше шести лезвий, при втором девяти и, наконец, при третьем – не больше двенадцати.

Выводы и перспективы последующих исследований

1. Предложен персональный вычислительный кластер *MPP* архитектуры, как модульная многопроцессорная система создана на базе стандартных вычислительных узлов, соединенных высокоскоростной коммуникационной средой.

2. Определены оценки эффективности кластерной системы при организации одностороннего и двустороннего режимов граничного обмена данными. Показано, что двусторонний режим при прочих равных условиях, позволяет в оптимальном варианте функционирование кластерной системы уменьшить количество узлов до четырех и повысить эффективность вычислений до 60 %. Однако, с другой стороны, при этом режиме обмена данными уменьшилось ускорение вычислений за счет увеличения времени граничного обмена данными.

3. Определены оценки эффективности кластерной системы при организации полудуплексного и дуплексного режимов работы системы. Доказано, что в режиме “дуплекс” существенно снизилось время вычислений и, кроме того, значительно выросло ускорение. Однако одновременно нужно подчеркнуть, что несколько снизилась оценка эффективности вычислений (54 % против 60 %).

4. Установлены оценки эффективности кластерной системы при организации многоканальных режимов функционирования вычислительной сети кластера. Такие режимы работы позволяют не только повысить эффективность распараллеливания, но и существенно уменьшить время вычислений, а также значительно ускорить вычисления. Таких результатов удалось достичь за счет уменьшения времени граничного обмена данными между вычислительными узлами кластерной системы.

5. Проведены исследования, направленные на определение загруженности линий связи кластерной системы. Это позволило установить оптимальное число узлов кластерной системы для разных режимов ее работы.

6. Выведены основные аналитические соотношения для определения числовых характеристик эффективности кластерной системы через ее основные параметры.

7. В своих последующих исследованиях авторы намереваются осветить особенности связи нескольких модулей многопроцессорной кластерной системы.

Литература

1. Роуч П. Вычислительная гидромеханика / П. Роуч; пер. с англ. – М.: Мир, 1980. – 616 с.
2. Коздоба Л. А. Вычислительная теплофизика / Л.А. Коздоба. – К.: Наук. думка, 1992. – 224 с.
3. Иващенко В.П. Параллельные вычисления и прикладные задачи металлургической теплофизики / В.П. Иващенко, Г.Г. Швачич, А.А. Шмукин // Системні технології. Регіональний збірник наукових праць. – Вип. 3(56). – Т. 1. – Дніпропетровськ, 2008. – С. 123 – 138.
4. Швачич Г.Г. К вопросу конструирования параллельных вычислений при моделировании задач идентификации параметров окружающей среды / Г.Г. Швачич // Математичне моделювання. – 2006. – № 2 (14). – С. 23 – 34.
5. Швачич Г.Г. О параллельных компьютерных технологиях кластерного типа решения многомерных нестационарных задач / Г.Г. Швачич // Materiály IV mezinárodní vědecko- praktická konference [«Vědecký potenciál světa - 2007»]. – D. 7. – Technické vědy. Matematika. Fyzika. Moderní informační technologie. Výstavba a architektura. – Praha: Publishing House «Education and Science» s.r.o. – P. 35 – 43.
6. Швачич Г.Г. Математическое моделирование скоростных режимов термической обработки длинномерных изделий / Г.Г. Швачич, В.П. Колпак, М.А. Соболенко // Теория и практика металлургии. Общегосударственный научно-технический журнал. – 2007. – № 4 – 5 (59 – 60). – С. 61 – 67.
7. Иващенко В.П. Персональный вычислительный кластер для моделирования многомерных нестационарных задач / В.П. Иващенко, Г.Г. Швачич // Материалы XVI международной научно-технической конференции «Прикладные задачи математики и механики». – Севастополь, 2008. – С. 235 – 238.
8. Сбитнев Ю.І. Дослідження оцінки ефективності багатопроцесорної кластерної системи / Ю.І. Сбінтєв, Г.Г. Швачич, М.О. Ткач // VI International Conference “Strategy of Quality in Industry and Education”, June, 1 – 8 2010, Varna; Bulgaria . – Proceedings. – V. 2. – P. 288 – 296.